Development and Characterisation of Galactosynthases for Application in Organic Synthesis

Inaugural-Dissertation

Zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Benjamin William Berntsson

aus Canberra, Australien

Düsseldorf, April 2017

aus dem Institut für Bioorganische Chemie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Jörg Pietruszka

2. Prof. Dr. Vlada Urlacher

Tag der mündlichen Prüfung: 12.06.2017

Declaration

Except where acknowledgement has been made, the work published in this thesis is my own. Parts of this work have previously published thus:

Poster Presentations

Benjamin Berntsson, Jörg Pietruszka; Lünen (DE), Feb 2014 at the 5th Annual Retreat of the CLIB-GC. "Learning from Nature: Enzymes in Natural Product Synthesis"

Benjamin Berntsson, Jörg Pietruszka; Wermelskirchen (DE), Mar 2015 at the 6th Annual Retreat of the CLIB-GC. "Learning from Nature: Enzymes in Natural Product Synthesis"

Benjamin Berntsson, Jörg Pietruszka; Herdecke (DE), Feb 2016 at the 7th Annual Retreat of the CLIB-GC. "Developing Non-natural Carbohydrate-active Biocatalysts"

Benjamin Berntsson, Jörg Pietruszka; Mosbach (DE), Mar/Apr 2016 at the 67th Mosbacher Kolloquium. "Development of Non-natural Biocatalysts for the Stereo- and Regioselective Formation of Glycosidic Bonds"

Oral Presentations

Benjamin Berntsson, Jörg Pietruszka; Kerkrade (NL), Oct 2013 at the 3rd Internal CLIB-GC Retreat HHU Düsseldorf. "Learning from Nature: Enzymes in Natural Product Synthesis"

Benjamin Berntsson, Jörg Pietruszka; Mülheim a. d. Ruhr (DE), Oct 2014 at the 4th Internal CLIB-GC Retreat HHU Düsseldorf. "Learning from Nature: Enzymes in Natural Product Synthesis"

Benjamin Berntsson, Jörg Pietruszka; Bonn (DE), Oct 2015 at the 5th Internal CLIB-GC Retreat HHU Düsseldorf. "Development of Glycosynthases from Use in Organic Synthesis"

Benjamin Berntsson, Jörg Pietruszka; Herdecke (DE), Feb 2016 at the 7th Annual Retreat of the CLIB-GC. "Optimising Carbohydrate Synthesis *via* Enzyme Engineering"

Benjamin William Berntsson

Acknowledgements

There are so many people who I must thank and without whom, this would not have been possible. Whether family, friends, colleagues, or miscellaneous, thank you.

Firstly, I'd like to thank my supervisor, Prof. Dr. Jörg Pietruszka, for taking me into his group and giving me the chance to study an incredibly fascinating field of research and sponsor my relocation to Germany. By extension, I would not have been able to stay had it not been for the funding from the Heinrich-Heine-Universität Düsseldorf and the Ministerium für Innovation, Wissenschaft, Forschung und Technologie des Landes Nordrhein-Westfalen *via* the CLIB-GC. I'd also link to extend a heartfelt thank you to the co-ordinators of the program throughout the years, Jessica, Sonja, Jenny, and Martina. I'd also like to thank Prof. Dr. Vlada Urlacher for you helpful discussions and being my second supervisor.

I must also thank my German family, the past and present members of the IBOC, IMET, and fellow CLIB-GC students, for their companionship, camaraderie, and assistance throughout the years. You all helped my adjustment to life in Germany through the sharing of your languages and cultures, and through our many discussions over the years. Thank you to my students, Swetlana, Susan, Wen, and Maggie, and my cooperation partners Lars Freier, Dr. Andreas Küberl, Dr. –Ing. Stephan Noack, and Bianca Klein for their assistance. I would also like to thank the behind-the-scenes team at the IBOC: Moni, Birgit, Vera, Bea, Rainer, and Erik, this group would fall apart without you. I'd especially like to thank Markus, Roza, Sanel, Marc, and Yvonne for their time and assistance in the correction and editing of this tome.

To my supervisor at Henkel, Sascha, thank you for taking me on as an intern in your work group at Henkel. The experience has been overwhelmingly positive for me and I have learned a lot about the industrial side of science. Thank you for your patience during the internship/thesis preparation phase. I could not have achieved this without your understanding and flexibility. Naturally, I must also thank Labor 113 and the rest of LRR for their friendship in the lab and the entertaining work days at Henkel.

To my non-science friends here in Germany, thank you for keeping me sane and level-headed – particularly my billiard buddy, Kai, and my Australian connection, Alex. To my friends back home in Australia, thank you for your support, messages, care packages, and chats over the years. You guys really are the best and there will always be room on my couch for you when you visit.

Finally, my family, especially Dad, Chris, and Aunty Annette. Without your support from the very beginning of my studies some x years ago, I would not be where I am today. There is nobody else in the world to whom I am more indebted. Thank you all.

Abstract

Carbohydrate-containing compounds are one of the three great families of natural products, yet are the least understood, in part due to the complexities of carbohydrate synthesis. Glycosynthases, a non-naturally-occurring family of mutant carbohydrate-active enzymes derived from the glycohydrolases, are just one approach to overcoming the hurdles involved in carbohydrate synthesis. In order to develop a toolbox for the carbohydrate synthesis, this work focussed on the generation of a small library of galactosynthases and the examination of their performance in organic synthesis.

In order to establish a baseline of enzymatic performance, to examine the efficiency of our assays and as a proof-of-principle, we decided to extend upon the work of *Jakeman* and *Withers*[1] and reproduce their LacZ mutants, E537S and E537S/G794D, and characterise them, which to the best of our knowledge has not been previously performed. These enzymes were also used as model in the development of our own, novel galactosynthases derived from Lac4 (*Kluyveromyces lactis*) and LacA (*Aspergillus oryzae*). The required mutations of the nucleophilic residue in these two wild type galactosidases were successfully introduced using PCR-based techniques. Lac4 demonstrated suitability for expression in *E. coli* ER2566, as demonstrated by the expression of a functional wild type galactohydrolase, but the heterogeneous expression of LacA in this system led to the formation of inclusion bodies. Despite being able to characterise the hydrolytic activity of the wild type LacZ enzyme using colourimetric methods, the glycosynthase variants demonstrated such a low activity that they were unable to be characterised despite attempts using colourimetry, ^{1}H NMR kinetics, GC-MS, and HPLC and the kinetic parameters were thus unable to be determined. Characterisation of the wild type Lac4 galactosidase was successful using colourimetric methods, but similarly to the LacZ glycosynthases, it was not possible to characterise the mutants of this galactosidase using established methods. Despite optimisation of the expression of the enzymes in *E. coli* ER2566 and purification using Ni-NTA affinity chromatography, large enough quantities of the enzymes were unable to be expressed and harvested to be able to effectively characterise these galactosynthases.

The synthesis of the α-galactosyl fluoride (**1**), required as both a substrate for galactosynthase reactions and also as a key building block in our syntheses of reference compounds, was achieved in 3 steps from D-galactose with an overall yield of 27% (Figure 0-1). After the observation that this compound undergoes autohydrolysis under standard conditions, kinetic experiments were performed in order to ascertain the half-life of this compound in reaction conditions. It was found that the compound in non-buffered conditions rapidly hydrolyses and ceases to exist after just 18 hours. In buffered conditions, the substrate had a half-life of approximately 4 days depending upon

the pH of the solution. The methodology in the generation of the galactosyl fluoride (**1**) was also applied to a number of other common saccharides.

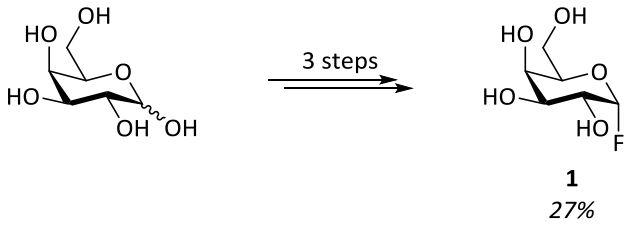


Figure 0-1: The synthesis of α-D-galactopyranosyl fluoride (**1**) was achieved in 3 steps in a 27% yield from the parent carbohydrate.

The other required substrates were a series of *p*-nitrophenyl glycosides for use in colourimetric assays. For this purpose, a library of 4-nitrophenyl glycosides was generated in order to test the hydrolytic substrate range of the wild type enzymes. These compounds were generated from their corresponding D-saccharides, with the exception of L-arabinose, in 3 steps in overall yields ranging from 3-45% (Figure 0-2). These compounds were also important building blocks for the generation of the disaccharide reference compounds.

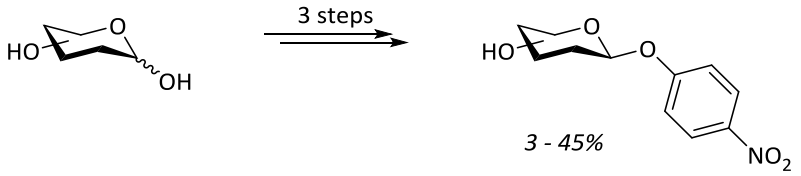


Figure 0-2: The equatorially-configured 4-nitrophenyl glycopyranosides were synthesised in 3 steps with a maximum overall yield of 45% from the parent carbohydrate.

Synthesis of the desired reference compounds was performed and proved to be challenging. Initially, it was attempted to synthesise the β-(1,4)- and β-(1,6)-galactosylglucosides, lactose (**2**) and allolactose (**3**), respectively, *via* a common intermediate, a 4,6-*O*-benzylidene acetal which could be opened selectively generating either a 4-OH or a 6-OH glucoside depending on the method used (Figure 0-3A). Due to the use of a 4-nitrophenyl glucoside, reductive opening proved ineffective as it also reduced the nitro group. Instead, an oxidative method was attempted, although this too proved unsuitable. Instead, it was decided to generate the allolactoside by selective protection and deprotection of the primary alcohol of the glucoside to generate the desired β-(1,6) glycosidic bond, which ultimately proved successful (Figure 0-3B).

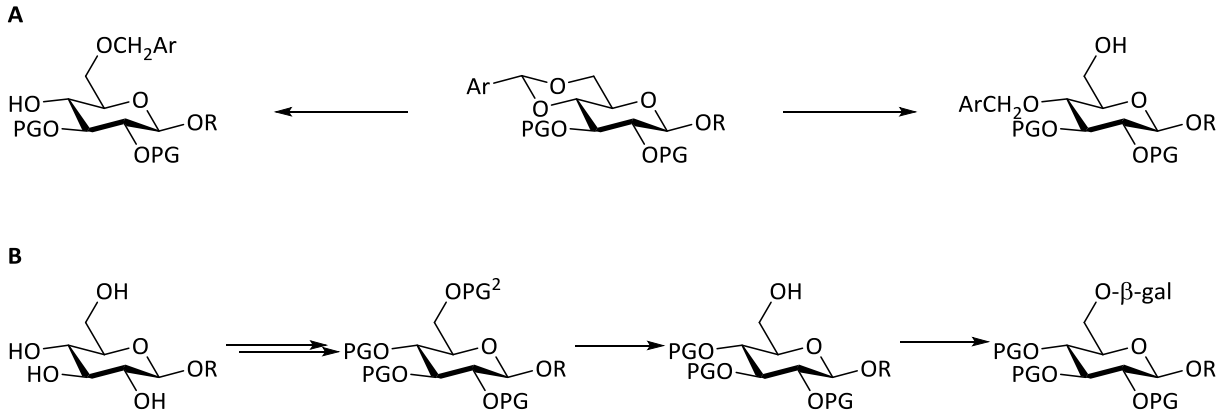


Figure 0-3: A) The selective opening of a 4,6-benzylidenyl glycoside can be selectively opened to return either the 4-hydroxyl or the 6-hydroxyl compound. B) The synthesis of the allolactoside was achieved by selective protection and deprotection of the 6-hydroxy group.

Despite not being able to characterise them, the glycosynthases were investigated as biocatalysts and it was demonstrated that the LacZ mutant, E537S/G794D was successful in generating the desired 4-nitrophenyl β-galactosyl-β-(1,6)-glucopyranoside (allolactose, **5**), although it was only isolable in a 3% yield (Figure 0-4), far from the 70% reported by *Jakeman* and *Withers*.[1] Due to a lack of sufficient analytic methodology, we remain unable to say with certainty whether the novel glycosynthases generated as part of this study are able to perform the glycosylation reaction.

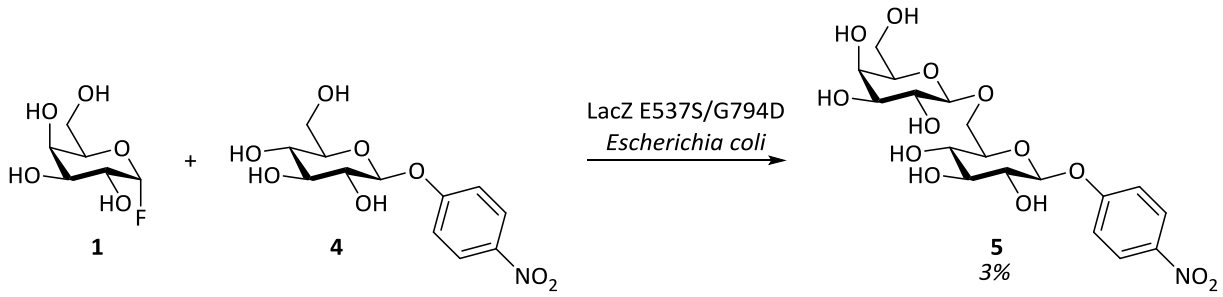


Figure 0-4: The galactosynthase, LacZ E537S/G794D, was able to synthesise the 4-nitrophenyl β-D-allolactoside (**5**), albeit in a 3% yield.

Zusammenfassung

Kohlenhydrathaltige Stoffe sind eine der drei größten Familien der Naturstoffe und doch die am wenigsten verstandenen, teilweise wegen der Komplexität der Synthese von Kohlenhydraten. Glycosynthasen, eine nicht in der Natur vorkommende Familie von kohlenhydrataktiven Enzymen, die von Glycohydrolasen abgeleitet wurden, bieten einen Zugang zu der Synthese von Kohlenhydraten. Der Fokus dieser Arbeit lag in der Synthese einer kleinen Bibliothek von Galaktosynthasen und der Untersuchung von deren Leistung in der organischen Synthese, um einen Plan für die Synthese von Kohlenhydraten zu entwickeln.

Um eine Basislinie der Leistung der entsprechenden Enzyme zu erhalten, die Effizienz der genutzten Essays zu testen und als Bestätigung des Prinzips, wurde entschieden, basierend auf der Arbeit von *Jakeman* und *Withers*[1], deren LacZ-Mutanten E537S und E537S/G794D herzustellen, um diese zu charakterisieren, was nach unserem Wissen bisher nicht geschehen war. Diese Enzyme wurden zudem als Modelle in der Entwicklung unserer eigenen, neuen Galactosynthasen von Lac4 (*Kluyveromyces lactis)* und LacA (*Aspergillus oryzae*) genutzt. Die benötigten Mutationen der nukleophilen Reste konnten erfolgreich durch PCR-basierte Techniken in Lac4 und LacA eingeführt werden. Lac4 erwies sich als passend für die heterogene Expression in *E. coli* ER2566, die heterogene Expression von LacA in diesem System führt jedoch zu der Bildung von Einschlusskörpern. Die hydrolytische Aktivität des Wild-Typs der LacZ- und Lac4-Galaktosidase ließ sich durch colorimetrische Untersuchungen charakterisieren, jedoch zeigten die mutierten Galaktosynthasen so niedrgie Aktivitäten, dass eine Charakteriserung durch Colorimetrie, ^{1}H-NMR-Kinetik, GC-MS und HPLC erfolglos waren. Deswegen war es unmöglich, die kinetischen Parameter zu bestimmen. Eine Charakterisierung der Wild-Typ Lac4-Galactosidase konnte durch colorimetrische Methoden erreicht werden, jedoch war die Charakterisierung der Mutanten, ähnlich der LacZ-Galactosynthasen, durch etablierte Methoden nicht möglich. Trotz Optimierung der Expression der Enzyme in *E. coli* ER2566 und Reinigung durch Ni-NTA Affinitätschromatographie konnte keine ausreichende Menge der Enzyme für die Charakterisierung dieser Galaktosynthasen gewonnen werden.

Die Synthese des α-Galaktosylfluorids (**1**), welches als Substrat für die Reaktion der Galaktosynthasen und auch als Schlüsselbaustein gebraucht wird, konnte in 3-Schritten, ausgehend von D-Galaktose, in einer Gesamtausbeute von 27% erreicht werden (Abbildung 0-1). Es wurde beobachtet, dass dieser Stoff unter normale Bedingungen eine Autohydrolyse durchläuft, weshalb kinetische Untersuchungen zur Bestimmung der Halbwertszeit dieses Stoffes unter Reaktionsbedingungen durchgeführt wurden. Es wurde entdeckt, dass unter ungepufferten Bedingungen schnelle Hydrolyse eintritt und dass α-Galaktosylfluorid (**1**) nach 18 Stunden

vollständig zersetzt ist. In einer gepufferten Lösung jedoch, hat dieser Stoff eine Halbwertszeit von circa 4 Tagen, abhängig vom pH der Lösung. Die Methodik der Synthese des Galaktosylfluorids wurde auch auf andere typische Saccharide angewendet.

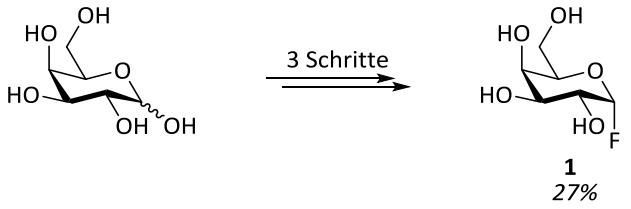


Abbildung 0-1: Die Synthese des α-Galaktofluorids (**1**) wurde über 3 Schritte in einer Ausbeute von 27% erreicht.

Andere, zur Nutzung in colorimetrischen Untersuchungen benötigte Substrate, waren eine Reihe von *p*-Nitrophenylglycosiden. Aus diesem Grund wurde eine Bibliothek von 4-Nitrophenylglycosiden synthetisiert, um die hydrolytische Substratbandbreite des Wild-Typs bestimmen zu können. Diese Stoffe wurden ausgehend von den entsprechenden D-konfigurierten Zuckermolekülen, mit Ausnahme von L-Arabinose, über 3 Schritte in Gesamtausbeuten von 3-45 % aufgebaut (Abbildung 0-2). Sie waren zudem wichtige Bausteine für die Synthese der Disaccharid-Referenzensubstanzen.

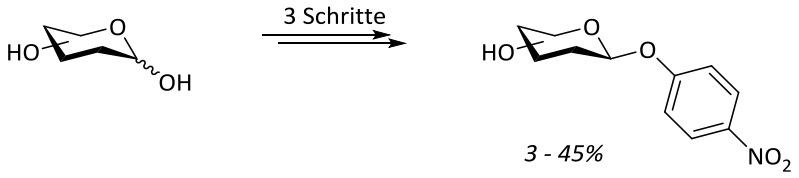


Abbildung 0-2: Die äquatorial konfigurierten 4-Nitrophenylglycopyranoside wurden über 3 Schritte in maximalen Ausbeuten von 45% synthetisiert.

Die Synthese der gewünschte Referenzsubstanzen erwies sich als herausfordernd. Anfänglich wurde versucht, die β-(1,4)- und β-(1,6)-Galactosylglukoside, Lactose (**2**) und Allolactose (**3**), über einen allgemein üblichen Baustein, ein 4,6-*O*-Benzylidenacetal, zu synthetisieren. Letzteres kann selektiv geöffnet werden, um entweder ein 4-OH- oder ein 6-OH-Glucosid zu generieren (Abbildung 0-3A). Da ein 4-Nitrophenylglycosid genutzt wurde, waren reduktive Ringöffnungsmethoden für solche Umsetzungen ungeeignet, da die Nitrogruppe zur Amino-Funktionalität reduziert werden konnte. Stattdessen wurden oxidative Ringöffnungsmethoden versucht, jedoch erwiesen diese sich als ungeeignet. Aus diesem Grund wurde entschieden, das Allolactosid durch selektive Schützung und Entschützung des primären Alkohols zu generieren, um so die gewünschte β-(1,6)-Glykosidbindung erstellen zu können. Letztlich war diese Methode erfolgreich (Abbildung 0-3B).

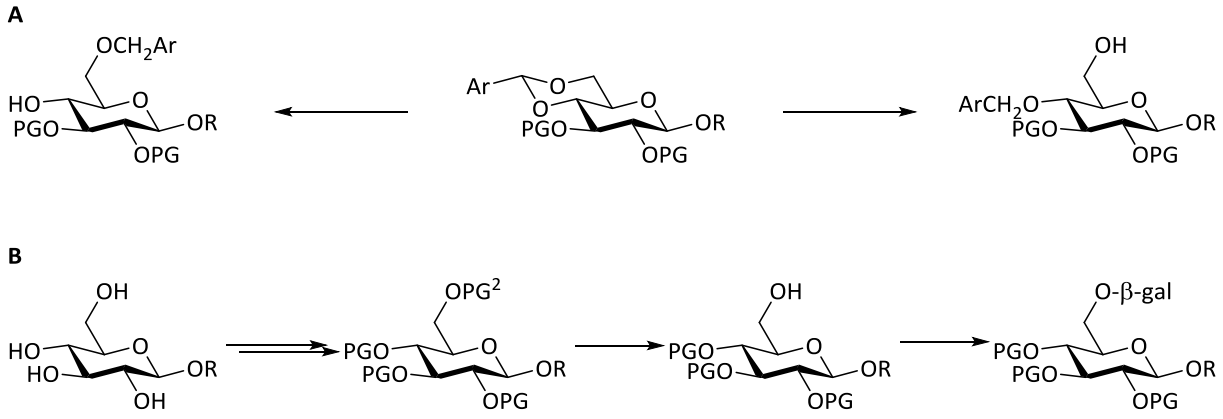


Abbildung 0-3: A) Das 4,6-*O*-Benzylidinylglykosid kann selektiverweise geöffnet werden, um entweder eine 4-Hydroxygruppe oder eine 6-Hydroxygruppe zu erhalten. B) Die Synthese des Allolactosids wurde durch selektive Schützung und Entschützung der 6-Hydroxygruppe erreicht.

Obwohl die Glykosynthasen nicht charakterisierbar waren, wurden sie als Biokatalysatoren geprüft. Es konnte gezeigt werden, dass die E537S/G794D Mutante der LacZ zu der Synthese der 4-Nitrophenyl-β-Galaktosyl-β-(1,6)-Glukopyranosid (Allolactose, **5**) erfolgreich einsetzbar war, die Produkte waren jedoch nur in einer Ausbeute von 3% isolierbar (Abbildung 0-4). Dieser Wert ist in ein starker Gegensatz zu der 70%-igen Ausbeute, die von *Jakeman* und *Withers* berichtet wurde.[1] Wegen Mangel an analytischen Methoden kann immer noch nicht mit Sicherheit gesagt werden, ob die neuen Glykosynthasen, welche in dieser Arbeit hergestellt wurden, die Glykosylierungsreaktion katalysieren können.

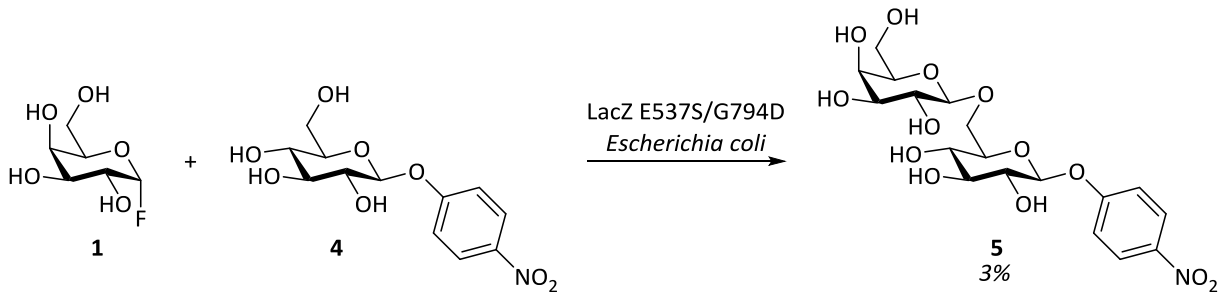


Abbildung 0-4: Die doppelt mutierte Galaktosynthase, LacZ E537S/G794D, konnte das 4-Nitrophenyl-β-D-Allolactosid (**5**) synthetisieren, jedoch nur in einer 3%igen Ausbeute.

General Note

The following figures, tables, schemes, references, and compounds are numbered sequentially. The compounds relevant to this work are also included in the structure summary. Furthermore, all compounds are numbered in accordance with IUPAC recommendations unless otherwise stated.

Abbreviations

Å	Ångstrom (10^{-10} m)
A	alanine
A.U.	absorption unit(s)
Ac	acetyl
Ac_2O	acetic anhydride
AcOH	acetic acid
ADC	antibody-directed catalysis
ADEPT	antibody-directed enzyme-prodrug therapy
adj.	adjusted (to)
ala	alanine
allo	6-*O*-(β-D-galactopyranosyl)-D-glucopyranose
allo	allolactose
amp	ampicillin
anh.	anhydrous
ara	L-arabinopyranose
asp	aspartic acid/aspartate
atm.	atmosphere
BDA	butane-2,3-diacetal
BgaB	*see: LacA*
BglI	β-glucosidase from *Schwanniomyces etchellsii*
BglX	β-glucosidase from *Aspergillus oryzae*
BL21	*Escherichia coli* BL21(DE3)
Bn	benzyl
Boc	*tert*-butoxycarbonyl
brine	saturated aqueous solution of sodium chloride
BSA	bovine serum albumin
^{t}Bu	*tert*-butyl
Bz	benzoyl
BzCl	benzoyl chloride
C	concentration
ca.	circa
cat.	catalyst
CBB	Coomassie Brilliant Blue
CDA	cyclohexane-1,2-diacetal
(Cl)Ac	α-chloroacetate
COSY	correlative spectroscopy
Cp	cyclopentadiene
CSA	(+)-camphorsulfonic acid
$CuSO_4$	copper sulfate
D	aspartic acid/aspartate
d	day(s)
d	doublet
Da	Dalton
DAST	diethylaminosulfur trifluoride
DBU	1,8-diazabicyclo[5.4.0]-7-undecene
DCM	dichloromethane
dH_2O	distilled H2O
DH5α	Escherichia coli DH5α
DMAP	*N,N*-dimethyl-4-aminopyridine
DMS	dimethylsulfone
DMSO	dimethylsulfoxide
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen

DTT	dithiothreitol
E	glutamic acid/glutamate
eda_{EC}	Entner-Douderoff aldolase from *Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EGA	electrogenerated acid
EI	electron ionisation
eq.	equivalent(s)
ESI	electrospray ionisation
Et	ethyl
Et_2O	diethyl ether
Et_3N	triethylamine
Et_3SiH	triethylsilane
ether	diethyl ether
EtOAc	ethyl acetate
EtOH	ethanol
*f*gal	D-galactofuranose
fuc	D-fucopyranose
G	glycine
GABA	γ-aminobutyric acid
gal	D-galactopyranose
galF	β-D-galactopyranosyl fluoride
galNAc	2-acetamido-2-deoxy-D-galactopyranose
GC	gas chromatography
glc	d-glucopyranose
gln	glutamine
glu	glutamic acid/glutamate
gly	glycine
GOI	gene-of-interest
H	histidine
h	hour(s)
HBr	hydrobromic acid
HCl	hydrochloric acid
HF	hydrofluoric acid
HHU	Heinrich Heine University Düsseldorf
HHU	Heinrich-Heine-Universität Düsseldorf
his	histidine
HMBC	heteronuclear multiple bond correlation
HPLC	high-performance liquid chromatography
HR	high-resolution
HSQC	heteronuclear single quantum coherence spectroscopy
IBG-1	Institüt für Bio- und Geowissenschaften: Biotechnologie
IBG-1	Institute of Bio- and Geosciences: Biotechnology
IBOC	Institut für Bioorganische Chemie der Heinrich-Heine-Universität Düsseldorf
IBOC	Institute of Bioorganic Chemistry of the Heinrich Heine University Düsseldorf
IDCP	iodonium dicollidine perchlorate
imid.	imidazole
IR	infrared
IUPAC	International Union of Pure and Applied Chemistry
K_2CO_3	potassium carbonate
K_2HPO_4	dipotassium hydrogen phosphate
kan	kanamycin
kDa	kiloDalton
KH_2PO_4	potassium dihydrogen phosphate
λ	wavelength (nm)
L	leucine
L	litre
lac	4-*O*-(β-D-galactopyranosyl)-D-glucopyranose
lac	lactose

Lac4	β-galactosidase from *Kluyveromyces lactis*
LacA	β-galactosidase from *Aspergillus oryzae*
LacZ	β-galactosidase from *Escherichia coli*
LB	lysogeny broth
LC	liquid chromatography
leu	leucine
lysozyme	human lysozyme
µL	microlitre
µM	micromolar
µmol	micromole(s)
m	mass
m	multiplet
M	molar
m.p.	melting point
m/z	mass-charge ratio
MALDI	matrix-assisted laser desorption/ionisation
man	D-mannopyranose
Me	methyl
MeCN	acetonitrile
MeOH	methanol
mg	milligram
$MgSO_4$	magnesium sulfate
MHz	megahertz
min	minute(s)
mL	millilitre
mM	millimolar
mmol	millimole(s)
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MTBE	methyl *tert*-butyl ether
MW	molar mass
MW	molecular weight
MWCO	molecular weight cut-off
n	mole(s)
N_3	azide
Na_2CO_3	sodium carbonate
Na_2HPO_4	disodium hydrogen phosphate
NaF	sodium fluoride
NaH_2PO_4	sodium dihydrogen phosphate
$NaHCO_3$	sodium hydrogen carbonate
NaOAc	sodium acetate
NaOMe	sodium methoxide
nHept	*n*-heptane
$\cdot nH_2O$	n-hydrate
NHTroc	2,2,2-trichloroethylcarbamate
Ni-NTA	nitrilotriacetic acid-immobilized nickel
NIS	*N*-iodosuccinimide
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
NPhth	phtalimido
NTA	nitrilotriacetic acid
OD	optical density
OD_{600}	optical density at λ= 600 nm
oePCR	overlap extension PCR
OH	hydroxyl
$\cdot OEt_2$	diethyl etherate
oNP	2-nitrophenyl
oNP	*ortho*-nitrophenyl
PCR	polymerase chain reaction
PE	petroleum ether (60-80 °C fraction)
Pfu	DNA polymerase from *Pyrococcus furiosus*
PG	protecting group
*p*gal	D-galactopyranose
Ph	phenyl
$PhBCl_2$	dichlorophenylborane
PhOAc	phenoxyacetic

PhSe	phenyl selenide
PMB	*para*-methoxybenzyl
PMP	*para*-methoxyphenyl
pNP	*para*-nitrophenyl
POI	protein-of-interest
poly.	polymerase
ppm	parts-per-million
^{i}Pr	*iso*-propyl
PTC	phase-transfer catalyst
py	pyridine
Q	glutamine
q	quartet
qcSDM	QuikChange® SDM
quant.	quantitative
R^2	coefficient of determination
R_f	retention factor
rpm	revolutions-per-minute
rthSDM	round-the-horn SDM
S	serine
s	second (s)
s	singlet
S.A.	specific activity
sacc.	saccharide
sat.	saturated
SDM	side-directed mutagenesis
SDS	sodium dodecyl sulfate
ser	serine
Sg	sugar
sH_2O	sterile, distilled H2O
T	temperature
t	time
t	triplet
T4	T4 DNA Ligase
TAE	Tris-Acetate buffer
Taq	DNA polymerase from *Thermus aquaticus*
TB	terrific broth
TBDPS	*tert*-butyldiphenylsilyl
TCA	trichloroacetimidate
Temp.	temperature
tet	tetracyclin
THF	tetrahydrofuran
THP	tetrahydropyran
TIC	total ion current
TIPS	triisopropylsilyl
TLC	thin-layer chromatography
TMP	2,4,6-trimethylpyridine
TMS	trimethylsilyl
TOF	time-of-flight
tol	toluene
Tr	trityl
Tris	Tris(hydroxymethyl)amino-methane
UV	ultraviolet
V	volume
Vis	visible
WT	wild type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
xyl	D-xylopyranose

Table of Contents

Chapter 1 - Introduction and Aims

1.1 The Significance of Carbohydrate-Containing Molecules

Carbohydrates are referred to as the 'third great class' of biological molecules due to their sheer abundance and importance in all natural systems.[2] There are a number of common, naturally-occurring carbohydrates found in nature (Figure 1-1) and it is these and derivatives and combinations thereof which are responsible for the many activities of carbohydrates. Compared to nucleic acids and proteins, carbohydrates have been historically overlooked in terms of biological activities despite having demonstrated roles in the structural integrity of cells, toxins and antibodies, half-life control of proteins, cell-matrix interactions, cell adhesion, cell-cell recognition, tumour metastasis, signal transduction, pathological processes, fertilisation, neurobiology, vaccines, xenotransplantation, and modulation of protein functions *via* post-translational glycosylation.[3-7] Recently, interest in the field of glycobiology, and more importantly, knowledge of glycobiological processes, has increased and carbohydrate-containing molecules are being more widely recognised as having industrial significance.

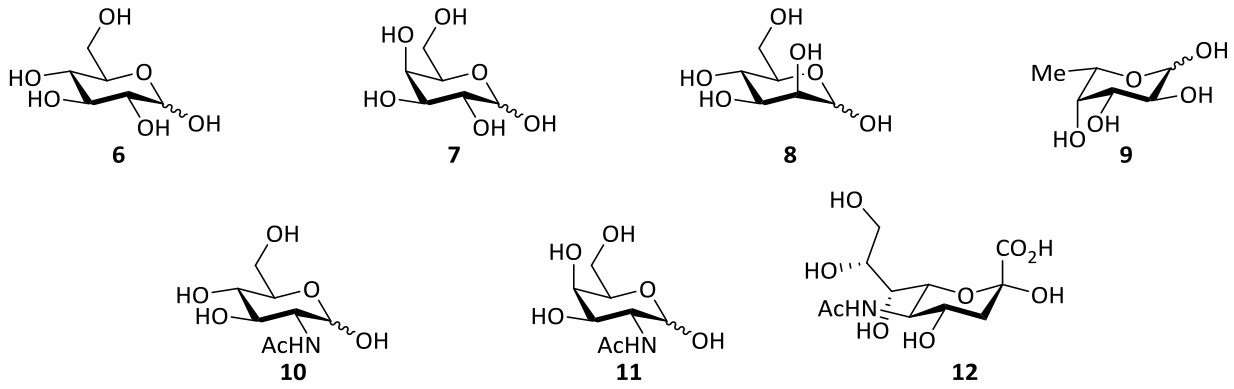
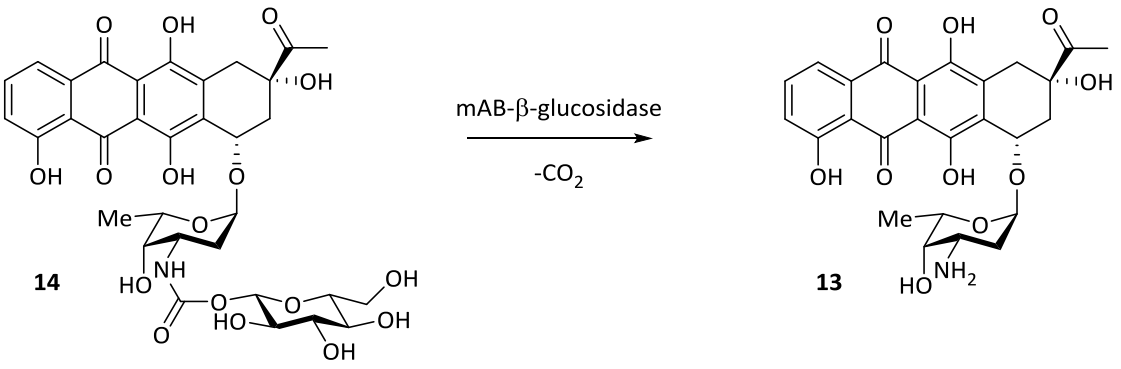


Figure 1-1: A selection of commonly-occurring carbohydrates in natural systems. D-glucopyranoside (**6**), D-galactopyranoside (**7**), D-mannopyranoside (**8**), L-fucopyranoside (**9**), 2-acetamino-2-deoxy-D-glucopyranoside (**10**), 2-acetamino-2-deoxy-D-galactopyranoside (**11**), *N*-acetyl-α-D-neuraminic acid (**12**).

Indeed, there are many studies which have shown the importance of not only the presence of a saccharide moiety, but also the exact identity and configuration thereof. In the anthracycline class of anti-tumour antibiotics, it has been shown that only the α-configured glycoside is active.[8] The potency of these compounds is so great that they typically show toxic side-effects, which renders them clinically undesirable.[3] However, it has been shown that a C4-epimer of the daunosamine (**13**) residue shows reduced cardiac toxicity without significant loss in activity.[8] Alteration of the identity of the attached carbohydrate has also demonstrated desirable anti-tumour activities and lower

toxicities.[9,10] In the case of aureolic acids, a group of antimicrobial and anti-tumour agents,[3] it has been shown that the aglycone moiety alone is inactive and the pendant polysaccharides are critical for the activity of these compounds, assisting in the binding of the aglycone to DNA.[11-13] In fact, the identity of attached carbohydrate moieties has been shown to be crucial for the activity of a range of bioactive molecules such as, avermectins,[14] enediyne antibiotics,[15-18] macrolide antibiotics,[19-25] polyene macrolides,[26] glycopeptides,[27] angucyclines,[28] and cholestane glycosides.[29]

In the search for new oncological treatments, efficient methods for glycosylation are also of great use and importance. The branched tetrasaccharide, sialyl Lewis^x, is a carbohydrate binding ligand important for the recruitment of neutrophils and monocytes from the blood to damaged tissues[30] and has abnormally-high presence in certain tumours.[31] This makes it an attractive target for a range of anti-inflammatory drugs or anti-cancer therapeutics. The over-production of certain receptors in cancers is well-known and L1210 leukaemia cells for example, are known to overproduce fucose receptors,[32] as such drug-makers can design prodrug treatments able to be directed to the site of infection using a method known as antibody-directed enzyme-prodrug therapy (ADEPT)[33] or antibody-directed catalysis (ADC).[34] In a study by *Leenders et al.*,[35] it was demonstrated that the use of a β-glucuronyl carbamate daunomycin compound (**14**) for ADEPT experiments showed up to 2000-times lower toxicity than the parent drug (Scheme 1-1).



Scheme 1-1: ADEPT can direct drug targets, such as daunomycin (**13**), to carbohydrate-specific receptors in order to deliver drugs to the site of infection before being enzymatically activated.

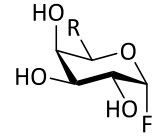
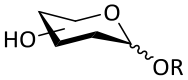
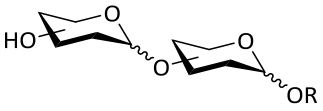
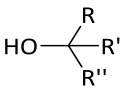
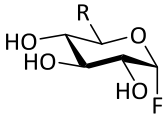
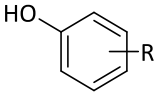
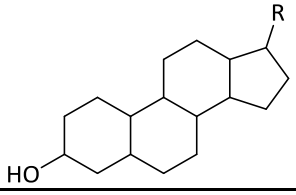
Despite the evidence that carbohydrates play an important role in nature, particularly in bioactive compounds, there is a lack in methodology for the efficient synthesis of carbohydrates. Due to the complex and similar reactivities of this group of polyols, the synthesis of these compounds still faces a number of challenges, the most prominent of which is the extensive and often wasteful protection/deprotection strategies. It is this gap in methodology that is addressed in this body of work through the development of novel carbohydrate-active enzymes capable of stereo-, regio-, and enantioselective glycosylation of unprotected carbohydrates.

1.2 Aims

The goals of the project are thus follows:

1. The work performed initially by *Jakeman* and *Withers*[1] on the β-galactosynthases, E537S and E537S/G794D, derived from *Escherichia coli* appeared promising and was therefore used as an established model glycosynthase upon which we could compare and contrast results obtained from any novel biocatalysts developed. The use of this established enzyme would also allow for the determination of baseline activity and allow for the development and assessment of suitable activity assays.
2. *Jakeman* and *Withers* tested the mutant glycosynthases using only one glycosyl donor, α-galactosyl fluoride (**1**), and a small selection of carbohydrate glycosyl acceptors. Therefore, it was desired to further explore the donor substrate range of these enzymes. Also, we wanted to expand upon the range of glycosyl acceptors for this enzyme, including the exploration of non-carbohydrate acceptors (Table 1-1).
3. To generate new, complementary glycosynthases which demonstrate a similar β-(1,6) selectivity to the *LacZ* derivatives, but which may have a different substrate range. Additionally, the use of a semi-promiscuous enzyme may allow for selectivity or reactivity tuning, similar to the change in reactivity observed by *Lundemo et al.*[36] who observed the absence of hydrolytic activity, yet retention of transglycosylation activity of a wild type glycosidase enzyme at high pH.

Table 1-1: Examples of potential glycosyl donors and acceptors for investigation for use with the glycosynthases generated within the context of this work.

Fluoride Donors		Glycosyl Acceptors	
		Monosaccharides	
$R = CH_2OH$	(D-Galactopyranosyl, **1**)	Disaccharides	
$R = CH_3$	(D-Fucopyranosyl, **15**)		
R = H	(L-Arabinopyranosyl, **16**)		
		Alkyl alcohols	
		Phenols	
$R = CH_2OH$	(D-Glucopyranosyl, **17**)	Steroids	
R = H	(D-Xylopyranosyl, **18**)		

Chapter 2 - Background

2.1 Basic Structure of Carbohydrates

The name 'carbohydrate' is derived from the stoichiometric formula for monosaccharides, $(CH_2O)_n$, or 'hydrated carbon'. This formula is also quite oversimplified, as we know now that many modifications on the traditional saccharide form exist, including *S*-, *N*-, *C*-, and *O*-glycosides, aminosaccharides such as *N*-acetyl glucosamine, and deoxy saccharides which are required by all cells in the form of the deoxyribonucleic acids. There also exist two major tautomeric forms of carbohydrates, aldoses and ketoses, the most simple of which are the 3-carbon saccharides, glyceraldehyde and dihydroxyacetone, and are interconvertible *via* an unstable enediol intermediate (Figure 2-1A). Due to the high number of chiral centres in carbohydrates, there exist two enantiomers for each, D- and L-carbohydrates (Figure 2-1B), with the D-enantiomer being the most predominant in nature.

Figure 2-1: A) the trialdose, glyceraldehyde (left), and the triketose, dihydroxyacetone (right). B) The two enantiomers of glucopyranose, D- and L- (left and right, respectively)

The aforementioned trioses are most commonly found as phosphorylated intermediates within the glycolysis pathways of many organisms, though these are not the only naturally-occurring carbohydrates. As previously mentioned, carbohydrates are one of the most prevalent naturally-occurring molecules, and tetroses, pentoses, hexoses, and heptoses are found throughout biological systems.

As the number of carbons in the carbohydrate increases, so too does the reactivity of the molecule. This is most notable in pentoses and hexoses where the molecules prefer to take on ring structures as opposed to a linear structure.[2,37] Arguably, the most important carbohydrate is D-glucose and as an aldohexose, it has four stereogenic centres. However, when it forms a ring, a fifth stereogenic centre is built from the prochiral aldehyde, this centre is commonly referred to as the anomeric centre and the conformation of the hemiacetal therefore provides a further stereodescriptor for all cyclic carbohydrates (Figure 2-2). The ratio of tautomeric forms, α- or β-, is influenced by electronic and steric factors and the natural ratio of the forms of various carbohydrates is given in Table 2-1.

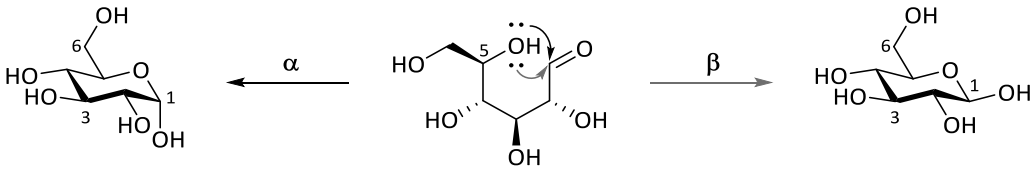


Figure 2-2: The electron pairs of 5-OH in glucose form either the α- or β-pyranose depending upon which face of the aldehyde is attacked.

Table 2-1: Relative amounts of the tautomeric forms of various monosaccharides in water at 40 °C.[38]

Carbohydrate	α-pyranose [%]	β-pyranose [%]	α-furanose [%]	β-furanose [%]
Ribose	20	56	6	18
Lyxose	71	29	<1	<1
Altrose	27	40	20	13
Glucose	36	64	<1	<1
Mannose	67	33	<1	<1
Fructose	3	57	9	31

The preference for the ring form, lactol, over the straight chain is generally considered to be more stable in the 5- and 6-membered lactols, furanoses and pyranoses respectively, as the bonds are free to adopt angles close to the natural 109.5° bonding angle of a tetrahedral, sp^3-hybridised atom. The deceptive drawing of these molecules as flat would suggest that the internal angles in the ring are 108° and 120° respectively. However, in 3D space, these molecules tend to adopt an envelope- and chair-form respectively, as their most stable conformations. Further structural conformers exist for 6-membered rings (Figure 2-3).

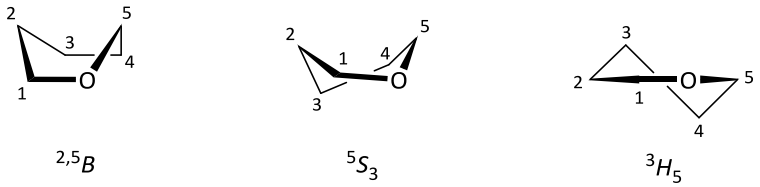
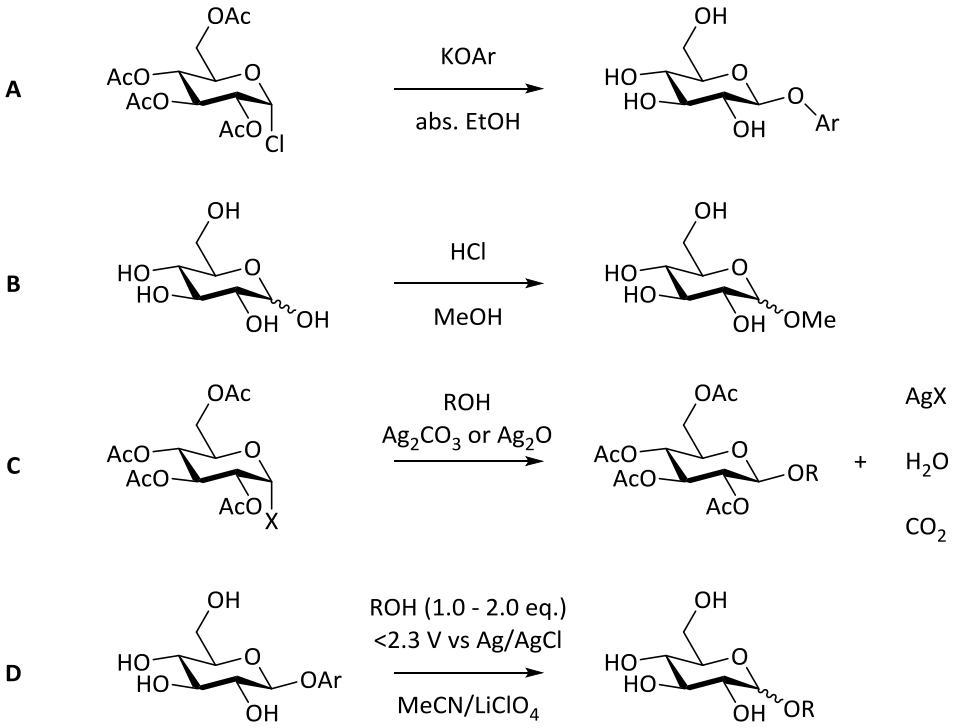


Figure 2-3: In addition to the chair conformation, 6-membered rings can exist in other conformations. Left to right: boat, twist (also known as skew or skew-boat), and half-chair forms.[39]

2.2 Methods for Glycosylation

Although the French chemist, *Paul Schützenberger*, published the first synthesis of glycosides, the analysis of the products was inconclusive.[40] It wasn't until 1879 when *Arthur Michael* reported the first glycosylation reaction of a protected glycosyl halide with a potassium phenolate salt to yield a phenyl glycoside, that the glycosylation reaction was established as a method in organic chemistry (Scheme 2-1A).[40] This was then followed up in 1893 by *Emil Fischer* who took a different approach in that he used the unprotected carbohydrate in acidic conditions to return a methyl glycoside (Scheme 2-1B).[41] To demonstrate the applicability of his method, he applied it again the following

year in the synthesis of phenyl glycosides.[42] The biggest drawback of the *Fischer* glycosylation, was that both the pyranose and furanose forms were returned as were both the α- and β-anomers of each and their corresponding ratios were determined by the stability of each isomer. In 1901, *Koenigs* and *Knorr*,[43] and independently also *Fischer* and *Armstrong*,[44] coupled methanol and ethanol to per-acetylated glycosyl halides using Ag_2CO_3 as a weak base to scavenge the halide leaving group. The *Koenigs-Knorr* reaction, as it's now known, has a high selectivity for generating the anomerically-inverted product as would be expected for a S_N2 nucleophilic substitution reaction (Scheme 2-1C). Recently, *Noyori* and *Kurimoto* have also reported the electrochemical synthesis of alkyl glycosides from unprotected phenyl glycosides (Scheme 2-1D).[45]



Scheme 2-1: Methods for glycosylation. A) The *Michael* glycosylation. B) The *Fischer* glycosylation. C) The *Koenigs-Knorr* glycosylation. D) Electrochemical glycosylation.

Since the initial glycosylations at the turn of the 20th century, many new methods for glycosylation have been developed. Importantly, these methods typically allow for stereoselective glycosylation returning either the axially- or equatorially-configured glycosides. This is achieved by a number of techniques, which rely on the alteration of reactivity. The most common influences are: the identity of the anomeric leaving group,[46,47] the protecting groups on the carbohydrate,[47-50] the choice of activator,[51,52] the choice of solvent,[53-58] and the use of anchimeric groups in the 2-position, i.e. acetyl groups.[59] Furthermore, by combining these various elements, the reactivity of each glycoside can be tuned allowing for selective glycosylation and thus rapid formation of oligosaccharides.[60-62]

The combination of these factors can allow for the synthesis of large oligosaccharides and these methods typically fall into one of two categories: orthogonal glycosylation, and reactivity tuning. Although lying outside the scope of this project, impressive demonstrations of these two approaches to oligosaccharide synthesis are found in the synthesis of everinomectin by *Nicolaou et al.*[63-65] and the one-pot synthesis of the envelope glycoprotein gp120 of HIV by *Grice et al.*[66] respectively.

Modern chemical (as opposed to biocatalysed) glycosylation methods are classed based on the identity of the glycosyl donor, a number of which are known. As demonstrated by *Michael*[40] and later *Koenigs* and *Knorr*[43] and *Fischer* and *Armstrong*,[44] the most effective method for glycosylation involves the use of a leaving group to facilitate the reaction without proceeding *via* the open chain form which can lead to more complex product mixtures, as demonstrated by the early versions of the *Fischer* procedure. Anomeric leaving groups have therefore been a topic of special interest to carbohydrate chemists over the years, with the more common glycosyl donors including acetates, glycosyl halides, phenyl selenoglycosides, thioglycosides, glycal-derived epoxides, and acetimidates.

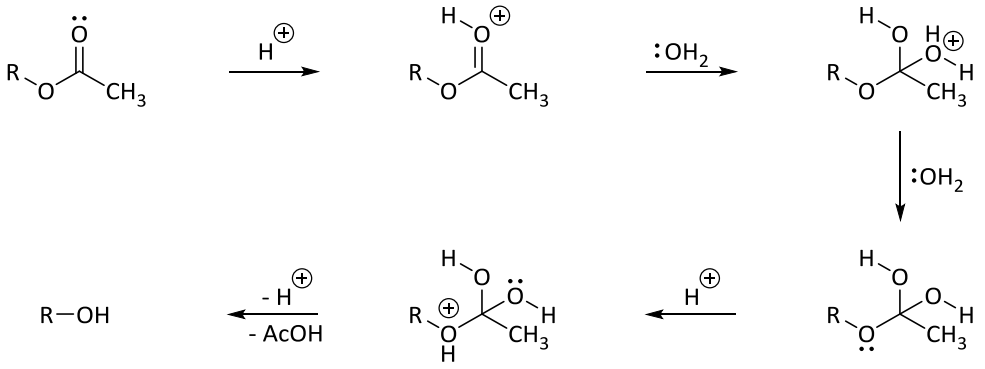
2.3 Glycosyl Donors

Acetates

Anomeric acetates are common intermediates used for the preparation of other glycosyl donors. Synthesis of these compounds is typically a standard procedure using acetic anhydride.[46] The variation of reaction conditions can favour the formation of one anomer over the other. For example, using a hard Lewis acid, such as $ZnCl_2$, results in thermodynamic control of the reaction leading to high selectivity for the axial product.[51] Conversely, the use of Ac_2O and NaOAc results in the near-selective formation of the equatorial anomer.[52] When the per-*O*-acetate is used as an intermediate, the anomeric purity is of lesser importance due to the anchimeric assistance from the 2-*O*-acetate during subsequent glycosylation steps. Therefore, reaction in a solution of Ac_2O/pyridine usually suffices for per-*O*-acetylation.[67] When starting from the unprotected carbohydrate, the aforementioned reaction conditions only generate the pyranoside. In order to obtain the furanose derivative, *Furneaux et al.*[68] have described a method using boric acid which forms a complex to trap the furanose form. This soluble intermediate is then reacted with Ac_2O and a catalytic amount of H_2SO_4 to return the furanose as the major product.

The anomeric acetate can then serve as a leaving group for glycosylation using a strong, hard Lewis acid.[46] For the preparation of glycosyl halides, the per-*O*-acetylated glycoside is reacted with the hydrogen halide (HX) wherein the latter acts as both the acid catalyst and the halide source.[69] For the generation of thioglycosides, boron trifluoride diethyl etherate ($BF_3{\cdot}OEt_2$) is used as the Lewis acid.[70] However, one major consideration in using such reagents is the need for acid stability in the donor

and acceptor, particularly with regards to side products being formed by acid-catalysed $A_{Ac}2$ hydrolysis (Scheme 2-2).



Scheme 2-2: Acid-catalysed $A_{Ac}2$ hydrolysis of alkyl acetates

Glycosyl Halides

Glycosyl halides have been used in glycosylation since the pioneering reactions by *Michael*, *Koenigs* and *Knorr*, and *Fischer*,[40-44,71] all of whom used either the glycosyl chloride or bromide in their studies. The glycosyl chlorides and bromides are easily generated by reaction of the per-*O*-acetylated glycoside with the corresponding hydrogen halide. The drawback of these compounds is their high reactivity and as such, care is required in their handling and storage. The iodides are known to react with alcohols in the absence of a promotor and must be generated *in situ* as they are typically too reactive to be isolated.[72] Glycosyl fluorides are the most stable and were for a long time considered too inert to be used as glycosyl donors.[46]

Glycosyl chlorides and bromides are typically reacted with halophilic metal salts, such as those from silver(I) or mercury(II). The drawback of this method is the hazard and inconvenience of working with heavy metal salts – this is particularly true for mercury(I) salts. Furthermore, the metal activators are required in at least equimolar amounts, but a large excess is typically used.[46]

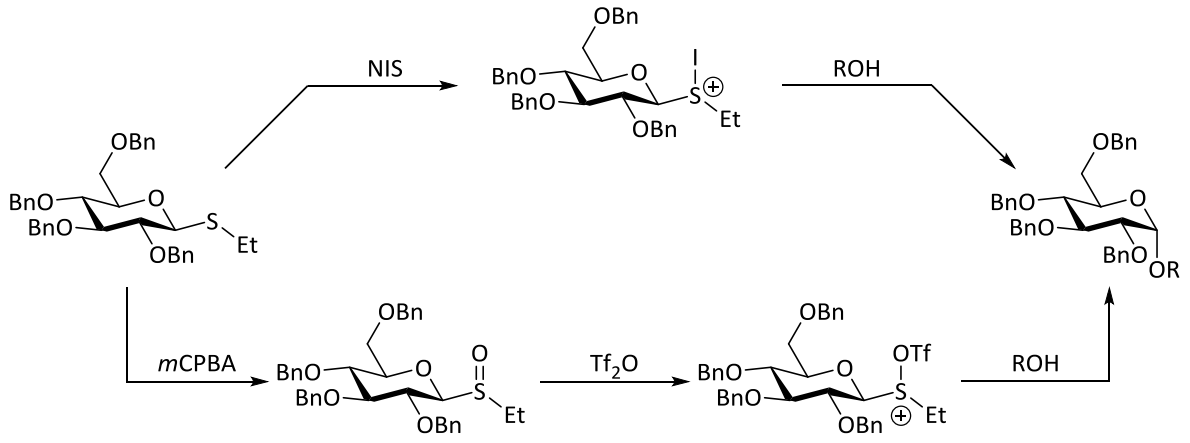
In contrast to the other halide donors, glycosyl fluorides did not react using the same activators as for the chlorides and bromides. In 1981, *Mukaiyama* was able to perform glycosylation of a per-*O*-acetylated glycosyl fluoride using $SnCl_2$ and $AgClO_4$ as a glycosylation promoter combination.[73] Glycosyl fluorides are traditionally obtained by reaction of the per-*O*-acetylated glycoside in a HF/pyridine solution. These reaction conditions are quite harsh and the use of highly corrosive HF requires special reaction vessels to be used, making this reaction less desirable. *Rosenbrook et al.*[74] reported the use of diethylaminosulfur trifluoride (DAST) as an alternative to these harsh conditions. DAST was previously known to convert a primary or secondary hydroxyl group to the fluoride with stereoinversion; as such, this reaction relies on the use of an anomerically-deprotected saccharide. Alternatively, equatorially-configured fluorides have been

selectively synthesised by *Gordon* and *Danishefsky* from glycals *via* an epoxide intermediate using tetrabutylammonium fluoride (TBAF).[75]

Glycosyl fluorides, like the other glycosyl halides, are reacted with an alcohol in the presence of a promoter, such as Cp_2HfCl_2-$AgClO_4$,[76] $BF_3{\cdot}OEt_2$,[77] $TiCl_4$,[78] and various Brønsted acids.[79] Trimethylsilyl ethers of the desired alcohol have also been used as the glycosyl acceptor in research performed by *Hashimoto et al.*[80] wherein they describe the use of a catalytic silicon-based Lewis acid to facilitate glycosylation.

Thioglycosides

One of the most successful and popular glycosyl donors in oligosaccharide synthesis, thioglycosides have earned their place as one of the most important intermediates in carbohydrate chemistry. The reason for their popularity is the versatility of the aryl- or alkylthio group as an anomeric protecting group as it is stable under a broad range of conditions.[81] Since the introduction of thioglycosides by *Fischer* and *Delbrück* in 1909,[82] a number of modifications and methods for the introduction of this group have been investigated. Anomeric thio groups can be installed from a broad range of starting materials using a number of methods. For per-*O*-acetylated glycosyl halides, the use of a thiol and a strong base, such as sodium hydroxide,[82,83] or the use of an alkali thiolate salt[84] are proven methods of generating the desired thioglycoside. Thioglycosyl donors must be activated prior to glycosylation, either directly or indirectly (Scheme 2-3). Direct activation is a one-step method which requires the addition of an *S*-specific electrophile, such as iodonium ions from *N*-iodosuccinimide (NIS) or iodonium dicollidine perchlorate,[85] which results in a sulfonium salt. Alternatively, indirect activation of thioglycosides first requires conversion of the 1-thio moiety to a more appropriate leaving group, typically a halide.[59] Treatment of the thioglycoside with iodine bromide, iodine chloride,[86] DAST,[87] or hypervalent iodoarene difluoride,[88] results in the corresponding glycosyl halide. This can then be activated in an analogous fashion to glycosyl halides. Alternatively, the thioglycosides can be activated *via* conversion to their corresponding sulfoxide using a variety of oxidising agents, such as peroxides.[89,90] Activation of the resulting sulfoxides is popularly performed using triflic anhydride,[91-93] although other activation methods are known.[90,94-98]



Scheme 2-3: Glycosylation of thioglycosides using either the direct activation (top) or the indirect activation *via* the sulfoxide intermediate (bottom).

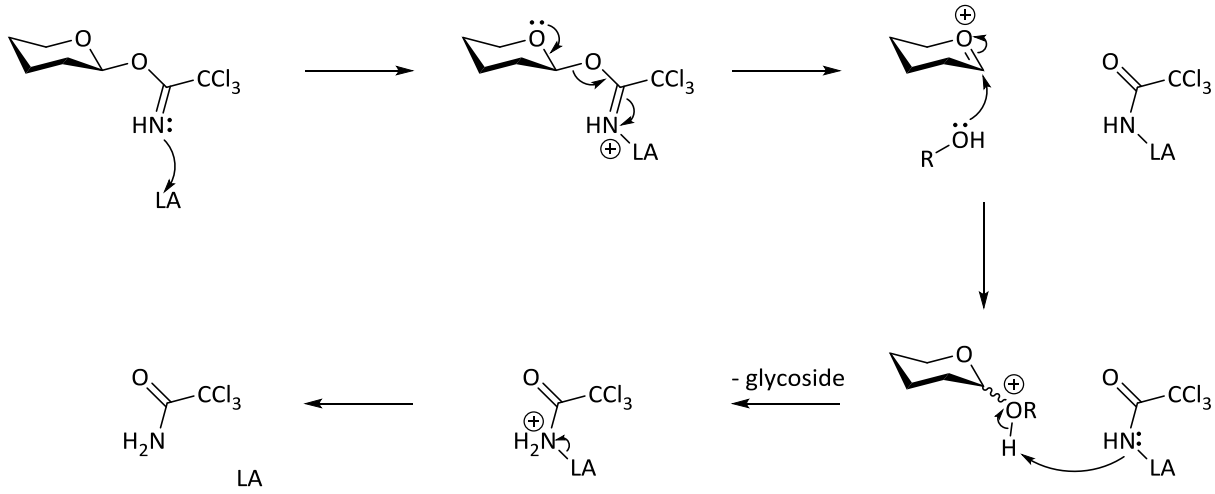
Selenoglycosides

Despite appearing in carbohydrate chemistry a century ago,[99] selenoglycosides were not employed as glycosyl donors until 1991, as pioneered by *Mehta* and *Pinto*.[100] Mostly taking the form of phenyl selenoglycosides, these donors gained popularity as it was observed that they could be activated using techniques previously used for thioglycosides. Additionally, it was found that they could be activated selectively in the presence of other glycosyl donors including thioglycosides,[100] and remained stable under conditions used for the activation of glycosyl halides and trichloroacetimidates.[101] An advantage of using the selenoglycosides is their stability to both acidic[66] and basic[102] conditions, making the formation of selenoglycosides a viable early step or precursor for large, multi-step reactions.

Trichloroacetimidates

Imidates were first reported as glycosyl donors in 1977,[103] but it wasn't until the trichloroacetimidate (TCA) group was introduced a few years later that their use became more prevalent.[104] The attractiveness of the TCA group lies in its high reactivity under mild activating conditions in catalytic amounts (Scheme 2-4)[53,54,105-110] and has been compared to glycosyl phosphates, nature's glycosyl donors.[111] In comparison to glycosyl bromides, as utilised in classical *Koenigs-Knorr* chemistry, the TCAs are more stable. In contrast to thio- or selenoglycosides, the trichloroacetimidates are less stable and require installation prior to coupling. Furthermore, the addition of the TCA group can only take place to the hemi-acetal form of the saccharide and as such, any anomeric protecting groups must first be removed before formation of the TCA. This can prove problematic in extended syntheses as it necessarily requires additional steps, thereby decreasing overall yields. Trichloroacetimidate is generated by reaction of the hemi-acetal with trichloroacetonitrile in the presence of a base. The identity of the base in this step is crucial, as it is known to influence the configuration of the resulting TCA. A weak base, such as K_2CO_3, and shorter reaction times result in

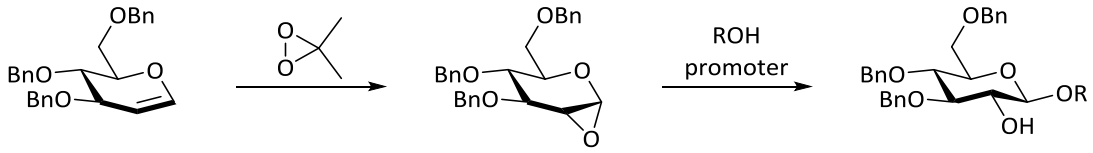
near-exclusive formation of the equatorially-configured product.[112] In contrast, the use of a strong base, such as NaH or DBU, results in formation of the thermodynamic, axially-configured product.[104,113-115]



Scheme 2-4: Trichloroacetimidates are activated by catalytic amounts of a Lewis acid (LA). Only the anomeric centre is shown for clarity.

Glycals

Glycals (1,5-anhydro-2-deoxy-1-hexenitols) are a powerful building block in the carbohydrate chemist's arsenal as these compounds are unable to act as a glycosyl donor until activated due to the presence of a double bond between *C*-1 and *C*-2 (Scheme 2-5). Once activated, these compounds always form a 1,2-*trans* glycoside. Studies by *Halcomb* and *Danishefsky*[116] found that conversion of glucosyl and galactosyl glycals to the 1,2-oxirane (epoxide) using 2,2-dimethyldioxirane (DMDO) showed high selectivity for the α-epoxide, which yielded the β-glycoside nearly exclusively after reaction with MeOH (α/β 1:20). This same study showed that the stereoselectivity of epoxidation is primarily influenced by the configuration at the 3-position, wherein the 2,3-*trans* epoxide is preferentially formed. *Danishefsky et al.* also reported that the use of acceptors larger than methanol resulted in longer reaction times and required promotion.[117,118] As a result, promotors for the glycosylation of 1,2-epoxidised glycals have been utilised, i.e. $ZnCl_2$[119,120] or K_2CO_3/18-crown-6,[121] the latter being typically used when using phenols as acceptors. 1,2-*cis* glycosides can be formed from the epoxide by using a stannyl ether/$AgBF_4$ system in a double-displacement mechanism.[122]



Scheme 2-5: Glycals are readily transformed to the 1,2-epoxide using DMDO, which can subsequently be reacted with an appropriate glycosyl acceptor to generate the 1,2-*trans*-configured glycoside.

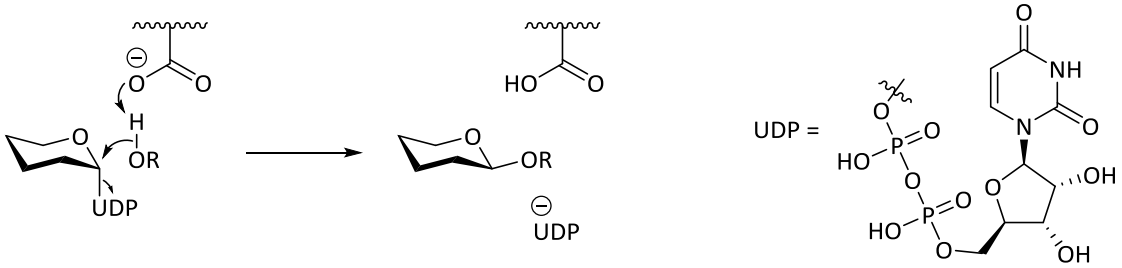
2.4 Carbohydrate-Active Enzymes as Biocatalysts

The glycosidic bond is one of the most important bonds in naturally-occurring compounds, accounting for up to two-thirds of carbon in the biosphere[123,124] and is estimated to have a half-life of around 5 million years at room temperature and neutral pH.[125,126] This level of stability makes the glycosidic bond more stable than the amide bond in proteins, $t_{½}$ = 460 years, and the phosphodiester bond of the nucleic acids DNA and RNA, $t_{½}$ = 140,000 years and 4 years, respectively.[127,128] The enzymes which are able to catalyse glycosidic bond hydrolysis are arguably among the most proficient of enzyme classes, able to cleave bonds at a rate of up to 1000 s^{-1} – a 10^{17}-fold enhancement.[125,126,129] Carbohydrate-active enzymes (CAZys) are a broad group of enzymes can be divided into four major classes: glycoside hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), and carbohydrate esterases (CEs). These are responsible for hydrolytic glycosidic bond cleavage, glycosidic bond formation, non-hydrolytic uronic acid cleavage, and the hydrolysis of sugar-ester bonds, respectively. Therefore, the use of CAZys in carbohydrate chemistry has a distinct advantage over traditional chemical methods, as these biocatalysts don't typically require the extensive protecting group chemistry required of purely chemical syntheses. Furthermore, as these enzymes are from living organisms requiring highly-specific glycosidic patterns, the enzymes also display a high degree of specificity, usually preferring one regio- and stereoisomer over all others.

2.4.1 Glycosyltransferases

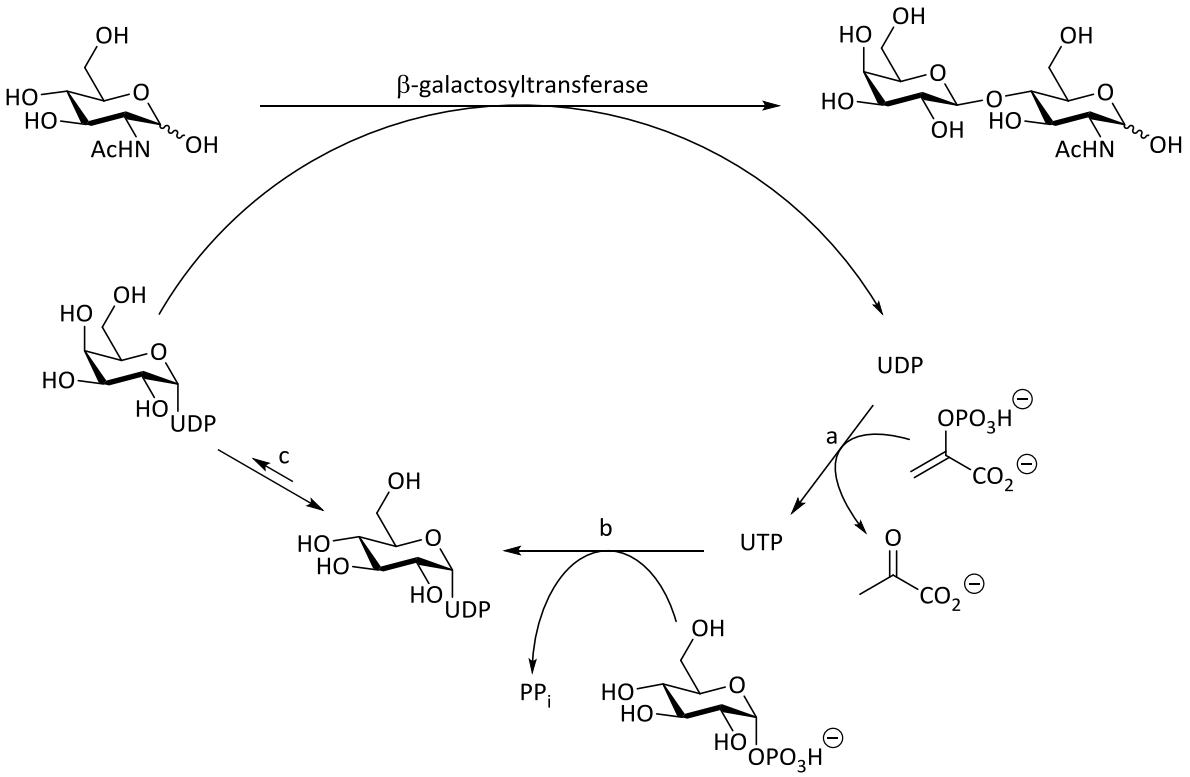
Glycosyltransferases [EC 2.4.-.-] are anabolic enzymes responsible for glycan synthesis in nature and are generally highly selective for both the donor and acceptor and can be divided into two broad classes, Leloir and non-Leloir glycosyltransferases.[130,131] Due to the myriad of biological functions of carbohydrates and their complex and specific coupling patterns, it is not surprising that each organism expresses a large number of glycosyltransferases, each with their own specificities with respect to donor, acceptor, and linkage.[132] They function by transferring an activated glycosyl donor – typically glycosyl nucleotides or glycosyl phospholipids – to acceptor substrates, such as oligosaccharides, antibiotics, lipids, and proteins (Scheme 2-6).[7,130] This results in a diverse range of glycans, which possess many varying roles within organisms including pathogenesis, signalling, cell-cell interactions and signalling, energy storage, cell wall structure, and important post-translational

glycosylation of proteins.[124,133,134] Regardless of role, the specificity of ordering within the structures is of critical importance and the deficiency or malfunctioning of these enzymes is the underlying cause of a large number of congenital disorders.[135-137] So prevalent are these enzymes that they make up approximately 2% of an organism's genome, the exceptions being obligate symbionts and parasites which contain few to no detectable GT-encoding genes. [126,133,137] Despite their clear utility in nature, these enzymes are not so easy to use as biocatalysts in chemical synthesis, as they are typically limited by their high specificity and the high cost of the glycosyl donor substrates, which although available commercially, can cost upwards of 1000 € g^{-1}. This unfortunate reality undermines their use in chemical synthesis and many researchers prefer to work with more promiscuous CAZys, such as glycosidases.



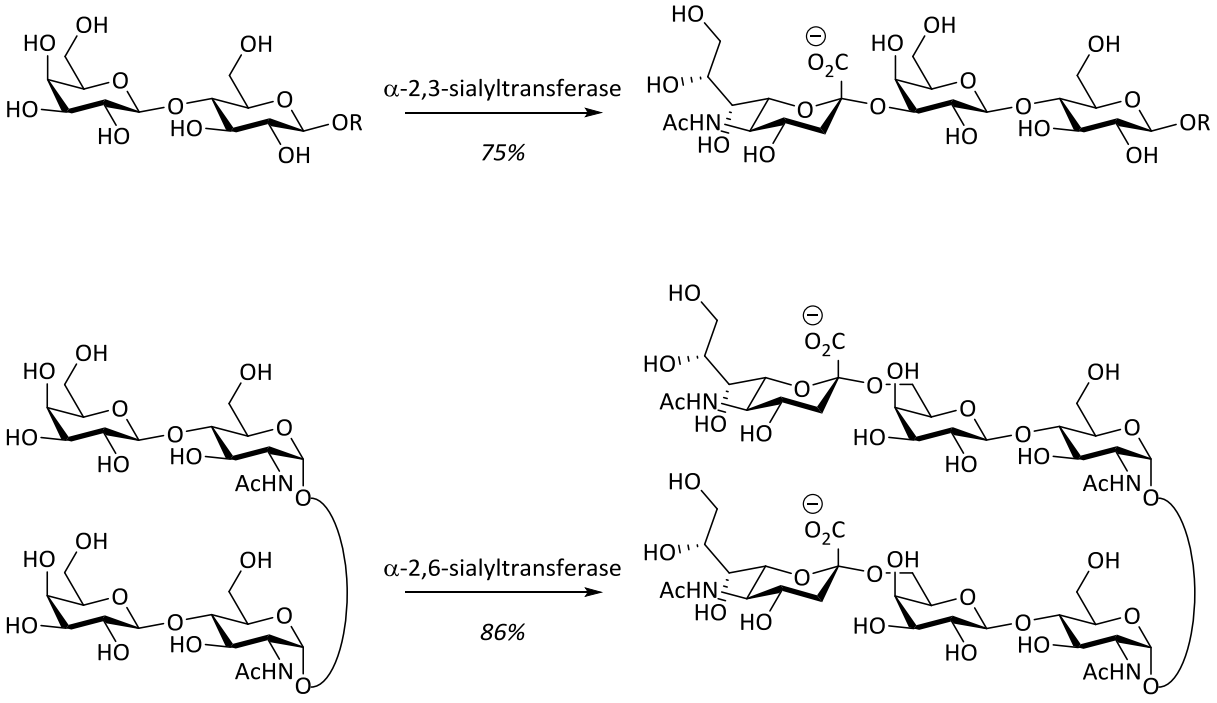
Scheme 2-6: Simple schematic of the mechanism of an inverting glycosyltransferase.

N-acetyllactosamine is a common motif in glycoproteins and plays an important role in cell adhesion, immune response, and tumour metastasis and progression.[138] In 1982, *Wong et al.*[139] reported the first preparative-scale glycosyltransferase-catalysed synthesis of this compound using a combination of four enzymes – galactosyltransferase, pyruvate kinase, UDP-glucose pyrophosphorylase, and UDP-galactose 4'-epimerase (Scheme 2-7). This approach overcame the problem of sourcing large quantities of expensive substrate thereby making glycosyltransferase-catalysed synthesis of oligosaccharides a more attractive tool to carbohydrate chemists.



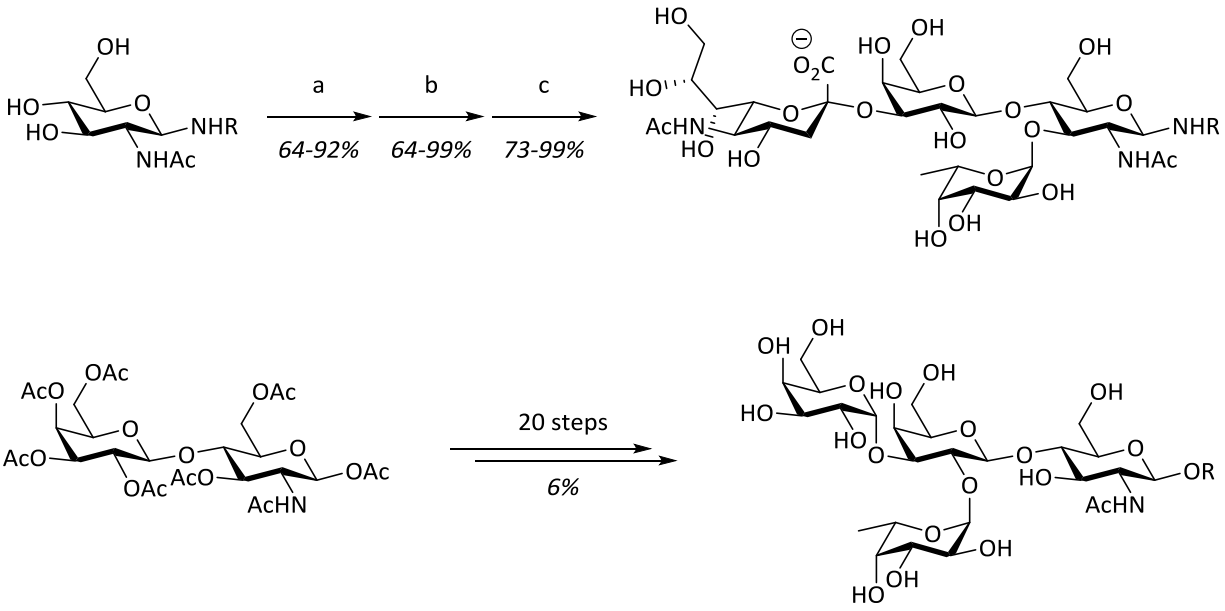
Scheme 2-7: Synthesis of *N*-lactosamine by *Wong et al.* using a phosphate recycling system.[139] a) pyruvate kinase; b) UDP-glucose pyrophosphorylase; c) UDP-galactose 4'-epimerase. Additional elements of the recycling system removed for clarity.

Another synthetic problem for which glycosyltransferases are suited is the installation of sialic acid moieties. These are commonplace in glycoconjugates, particularly gangliosides, and are biologically-important in the binding of bacteria, lectins, toxins, and viruses.[140] As the presence of the carboxyl group reduces the reactivity of the anomeric hydroxyl group, chemical sialylation can be challenging, as the 2,3-dehydro product is a common side product under chemical conditions.[141] *Danishefsky et al.*[142] used a sialyltransferase to overcome this problem in their total synthesis of the GM3 cell-surface ganglioside. They utilised an α-2,3-sialyltransferase to sialylate the lactosyl ceramide in yields of 75% (Scheme 2-8). That same year, *Unversagt* reported the sialylation of a decasaccharide-asparagine conjugate using an α-2,6-sialyltransferase in 86% yield.[143]



Scheme 2-8: Sialylations using sialyltransferases. Upper: *Danishefsky et al.*'s total synthesis of the cell-surface ganglioside GM3 used an α-2,3-sialyltransferase to install terminal sialic acid moiety.[142] Lower: Impressive synthesis of a sialylated undecasaccharide by *Unversagt*[143] using an α-2,6-sialyltransferase with the linking pentasaccharide represented by the loop.

Fucosylation is another ubiquitous biological modification in the generation of glycoconjugates[144,145] and is one of the last reactions involved in biological blood group antigen synthesis.[140] As blood group antigens are important in a number of biological roles, such as in the inflammation response,[146] it is natural that fucosyltransferases are of interest to researchers. The first preparative-scale, enzymatic synthesis of sialyl Lewis^x was reported by *Ichikawa et al.*,[146] wherein they performed the last step, the fucosylation of a sialylated lactosamine, using an α-1,3-fucosyltransferase. This work was extended upon by *Baisch* and *Öhrlein*[147] who synthesised a complex glycopeptide cluster to increase the binding of sialyl Lewis^x to E-selectin. This synthesis was purely enzymatic and took place in three steps from a *N*-glucosamine backbone through successive transformations using a β-1,4-galactosyltransferase, an α-2,3-sialyltransferase, and finally an α-1,3-fucosyltransferase in an overall yield of up to 60% (9 examples, 33-60% over 3 steps; Scheme 2-9). This demonstrates a clear benefit of using glycosyltransferases as biocatalysts, even on simple tetrasaccharides, particularly when comparing this to the results of *Hara et al.* who recently synthesised a similar tetrasaccharide, the B group blood antigen, chemically in 20 steps and obtained an overall yield of just 6%.[148]



Scheme 2-9: The use of glycosyltransferases in the synthesis of complex oligosaccharides demonstrates clear advantages over traditional chemical synthesis as demonstrated in the comparison of the syntheses of tetrasaccharides by *Baisch* and *Öhrlein*[147] (upper) and *Hara et al.*[148] (lower).

2.4.2 Glycohydrolases

Glycoside hydrolases, also known as glycosidases [EC 3.2.1.-], can be broadly defined as having one of two possible activities; those which invert, and those which retain the anomeric bond configuration after hydrolysis. Furthermore, glycosidases can also be defined as either *endo*- or *exo*-depending on where in an oligosaccharide they catalyse hydrolysis (Figure 2-4).

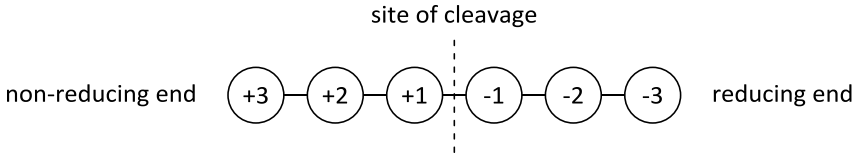


Figure 2-4: Basic nomenclature of glycohydrolases. *Exo*-glycohydrolases cut a single, terminal saccharide unit, typically at the reducing end. *Endo*-glycohydrolases cut in the middle of poly/oligosaccharide chains. Positive sites are those non-reducing carbohydrate units on the non-reducing side of hydrolysis, whereas negative sites are those carbohydrates cleaved from the reducing end of an oligosaccharide.

For both of these general cases, the enzymes have a pair of carboxylic acid residues within the active site which facilitate hydrolysis. Although, there are glycosidases known, which utilise alternative mechanisms, such as those which utilise neighbouring-group participation mechanisms,[149-161] are ascorbate-[162-164] or NAD-dependent,[165-169] as well as the highly-specialised sialidases.[170-174] With the exception of the myrosinase class of glycosidases,[162] it is widely accepted that the transition state of glycosidase-catalysed hydrolysis proceeds *via* an oxocarbenium ion-like transition state (Figure

2-5),[125,129,149,169,173-176] although, due to the instability of glycosyl cations, which have lifetime estimates of between 10^{-10} and 10^{-12} s, this is yet to be directly observed.[123]

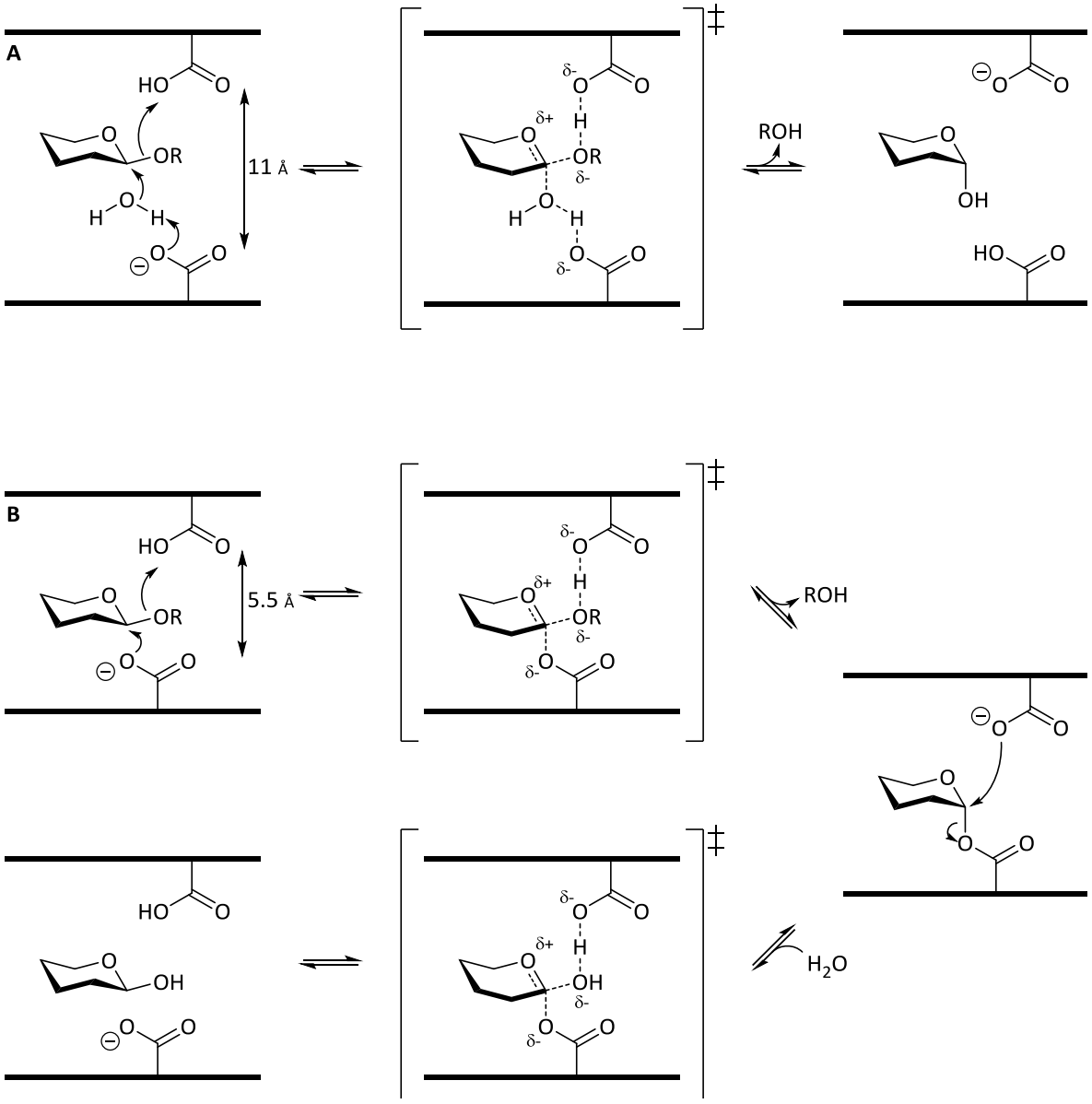
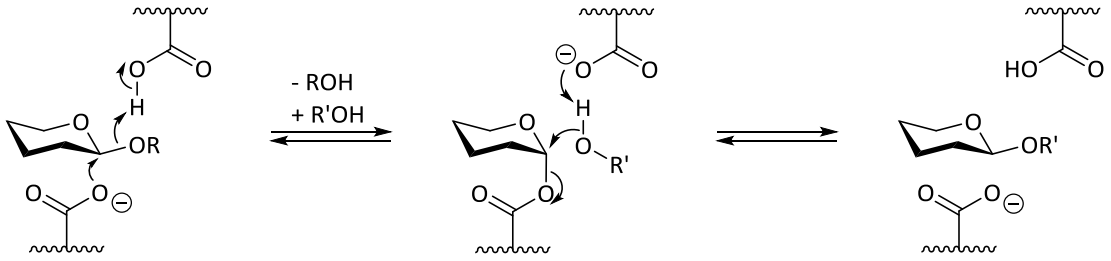


Figure 2-5: With the exception of myrosinases, it is widely accepted that the glycosidase mechanism proceeds *via* an oxocarbenium mechanism as demonstrated in the simple mechanisms of an inverting β-glycosidase (A) and a retaining β-glycosidase (B).[129]

In inverting hydrolases, the two carboxylic acid residues are approximately 11 Å apart which allows the substrate and a water molecule to bind between them before reacting *via* a single-displacement, S_N2-like mechanism. The two carboxylic residues act as an acid/base tandem, whereby one residue abstracts a proton from the water molecule, which then attacks at the anomeric centre, whilst the other assists the leaving of the generated alcohol by donating a proton to the generated oxyanion (Figure 2-5A). In retaining hydrolases on the other hand, the two catalytic residues are located 5.5 Å

apart, allowing only the substrate to bind, and act *via* a double-displacement mechanism, as first postulated by *Koshland*.[177] In this case, one carboxylic acid residue acts as a nucleophile, attacking the anomeric centre, with the other acting as a general acid, donating a proton to the outgoing alcohol. The result is a reactive sugar-enzyme intermediate with inverted anomeric stereochemistry. The second step involves the newly-generated conjugate base stabilising the incoming water molecule *via* hydrogen bonding, allowing the water molecule to react with the sugar-enzyme intermediate to once more invert the stereochemistry, leading to overall retention of stereochemistry of the anomeric centre (Figure 2-5B).

Glycosidases can be used for glycosylation in one of two ways; either by reverse hydrolysis or by transglycosylation (Scheme 2-10).[178] For reverse hydrolysis-type reactions, a prerequisite of these reactions is the reduction of water concentration compared to a normal aqueous system.[179] Enzymes which are stable under these conditions are of particular interest, as are those which are active in organic solvents.[180-185] Studies using those few glycohydrolases which have been studied at low water concentrations have revealed that although hydrolysis is typically increased in higher water concentrations, the equilibrium can be shifted towards synthesis by decreasing the water content, or more specifically, the water activity of the aqueous medium. Furthermore, it has been demonstrated in these few studies that different enzymes have different requirements for their respective water activities as highlighted by the 0.67 water activity required by almond β-glucosidase[186] and 0.29 required by the β-glycosidase from *Pyrococcus furiosus*.[187]



Scheme 2-10: Simple schematic of the mechanism of a glycohydrolase which can either hydrolyse glycosides (R' = H) or perform transglycosylation reaction (R' ≠ H).

Currently, the reasons for such large differences in water activity dependence is unknown, but it is hypothesised that these necessary water molecules play a functional role and necessarily their absence leads to cessation of enzymatic activity.[179] *Teze* and co-workers[188] investigated this by comparison of GH1 glycosidases and identified a number of conserved water molecules. They identified that 90% of those structures contained at least seven water molecules. They concluded that these conserved water molecules, which were typically clustered, formed water channels and aided the transportation of water to the active site. This insight has interesting implications for the use of these enzymes for synthesis and is a reasonable explanation as to why those enzymes, which are active in solvents with lower water concentrations, are better at forming glycosidic bonds. With this

in mind, the engineering of hydrolases to be more active at lower water activities would be an interesting area of investigation for the creation of biocatalysts from hydrolase enzymes.

An alternative to this is the immobilisation of hydrolases, such as performed by *Lundemo et al.*[189] In their study, they compared the activity of lyophilised glycosidase against those previously immobilised and found that the immobilised enzyme had 17- to 70-fold higher activity than the lyophilised enzyme at lower water activity despite the two having similar activities at full hydration. This was rationalised as being due to the immobilisation process facilitating the retention of crucial water molecules and thus helping to retain its activity at lower water concentrations. High-alcohol solutions have also been demonstrated as useful in both reverse hydrolysis and transglycosylation systems for the synthesis of alkyl glycosides.[179] Typically-used alcohols for this are hexanol and octanol, large chain alcohols which are immiscible with water. The aqueous phase of the two-phase system is saturated with the alcohol in addition to the hydrophilic saccharide and enzyme. Higher alcohols are disadvantaged by their lower solubility in water and therefore reaction times are typically longer and yields lower.

Directed evolution has also been used to improve selectivity for transglycosylation over hydrolysis. *Feng et al.*[190] used a β-glycosidase from *Thermus thermophilus*, which had a transglycosylation yield of 8%. Following mutation, they were able to increase the yield to 75%. An analysis into the structural differences between wild type enzymes and their transglycosylase mutants by *Teze et al.*[191] showed no major structural changes. However, they noted that the majority of beneficial mutations were located in the -1 site. Using this semi-rational approach, *Tellier* and co-workers[192] were able to create an α-transglycosylase from the α-galactosidase of *Geobacillus stearotherophilus*. Investigations by *Hansson* and *Adlercreutz*[187] found that a single mutation in the *P. furiosus* β-glucosidase resulted in a two-fold increase in the transglycosylation/hydrolysis ratio for the hexanol catalyst. Single mutation of the β-glucosidase from *T. neapolitana* accounted for a 7-fold increase in this ratio, increasing the yield to 58%.[193] In a follow-up study, *Lundemo et al.*[36] were able to eliminate hydrolytic activity in the mutant by increasing the pH of the reaction to 10. Importantly, this had no significant effect on the rate of transglycosylation. There do exist natural transglycosylases, which despite having high homology to glycohydrolases, preferentially catalyse transglycosylation.[194] Characteristic of these natural transglycosylases are xyloglucan *endo*-transglycosylases, sucrose-type enzymes of GH 13 and GH 70, cyclodextrin glucanotransferases, α-amylases, and sialidases. Although, as transglycosylases are nearly indistinguishable from their hydrolysing counterparts, it remains difficult to pinpoint what features improve transglycosylation for further mutation studies.

Finally, glycosidases show perfect stereoselectivity, in that only the α- or β-product is formed, as a result of the active site architecture which prevents the formation of the undesired anomer. This is a crucial advantage over non-enzymatic techniques which often show great, but not perfect, selectivity;

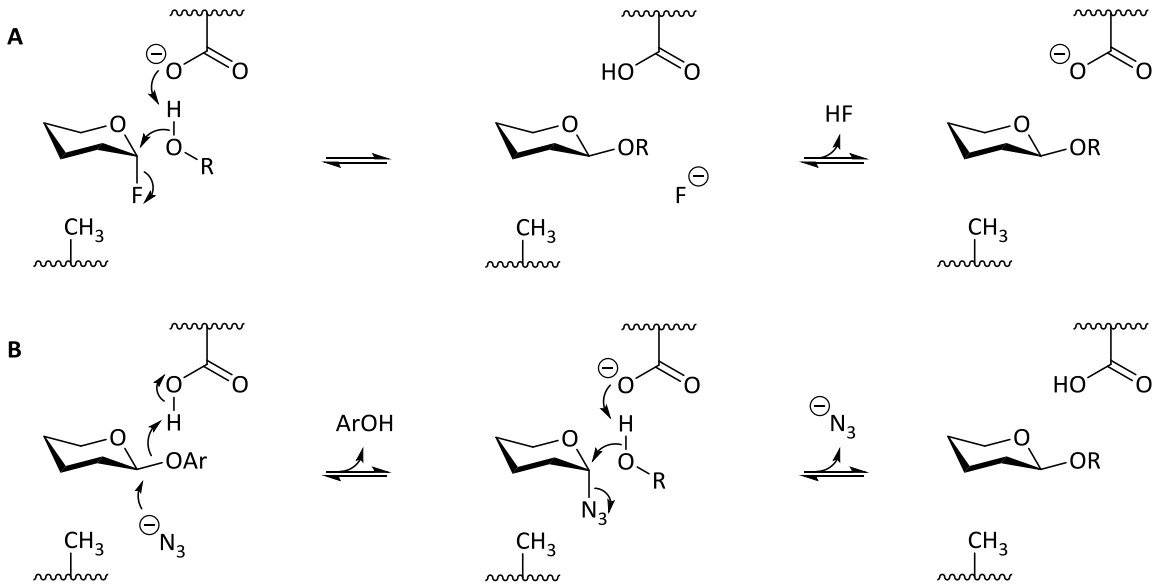
achieved with appropriate reagents and reaction conditions, leaving groups, and neighbouring groups. The other advantage that enzymes have over chemical glycosylation methods, is that they avoid the use of protecting group chemistry, which is used to ensure regioselectivity. The regioselectivity of enzymes is dependent upon the class of glycosidase used. *Exo*-glycosidases cleave/add only a single saccharide at the non-reducing end of the carbohydrate and this can give rise to a mixture of regioisomers. The use of *endo*-glycosidases for this typically results in higher regioselectivity as a result of the cleft in which the substrate needs to fit. However, this method is only suitable for the addition of large polysaccharides and not small monomeric units.[195]

2.4.3 Glycosynthases

Glycosynthases are a recent invention and do not occur naturally. These enzymes are not a new enzyme class, but rather a specific family of active-site mutants of glycohydrolases. The mutation of the nucleophilic residue in the active site results in a structurally identical enzyme that lacks hydrolytic activity.[196] Despite this, they are able to catalyse the formation of glycosidic bonds when activated substrates are used.[197] Invented in 1998, *Withers* and co-workers[198] observed that catalytic nucleophile mutants, despite being catalytically inactive, maintained the architecture of the wild type enzyme. By exposure of an activated glycoside which mimics the anomeric configuration of the sugar-enzyme intermediate, they postulated that this would allow for the synthesis of polysaccharides without concern for product hydrolysis.[198] Using the β-glucosidase, *Abg*, from *Agrobacterium* sp., the E538A mutant was generated and when α-glucopyranosyl fluoride was used as a donor substrate, a variety of polysaccharides of various orders were generated using a number of different carbohydrate-acceptors resulting in yields of up to 92%.[198] Critical to these reactions is the use of an activated donor, as it mimics the enzyme-sugar intermediate found in the native enzymes. The use of glycosyl fluorides as donors is a well-established method,[198-203] but other donors such as azides[204,205] and formates[206,207] are known. Larger leaving groups on the donors are typically avoided as they create extra steric bulk in the active site, which would hinder the incoming acceptor molecule. An exception to this is the use of aryl sialosides as donors in a neuraminidase-based glycosynthase.[208]

Classically, the glycosynthase mechanism proceeds in a S_N1-type reaction wherein the fluoride is displaced by the incoming acceptor, activated by the acid/base residue (Scheme 2-11A).[198] An alternative method, typically used to activate donors *in situ* is the so-called chemical rescue method. This reaction mechanism follows a double-displacement mechanism in which the donor, with an appropriate anomeric leaving group, is reacted with a non-enzymatic nucleophile, such as an azide or a formate.[207] This generates the "true" donor species, which then follows a classical glycosynthase mechanism (Scheme 2-11B). As there is no hydrolytic back-reaction, *Park et al.*[209] were able to demonstrate that the use of a glycosynthase also increases the regioselectivity as only one

regioisomer was formed in comparison to the 3 regioisomers generated by the wild type glycosidase. This difference arises due to the nature of the reaction, which is kinetically controlled in the synthases, compared to thermodynamic in glycosidases and transglycosidases.[210]



Scheme 2-11: Glycosynthases have two main mechanisms. A depicts the 'classical' synthase approach using an activated substrate which mimics the glycosyl-enzyme intermediate. B shows the alternative reaction mechanism, the so-called chemical rescue method, wherein a small nucleophile encourages departure of the leaving group, generating the reactive intermediate *in situ*.

Glycosynthases have been used in a number of syntheses towards complex polysaccharides. In 2012, *Sakurama et al.*[211] reported their use of a bacterial α-L-fucosynthase to generate Lewis antigens A and X. *Danby* and *Withers*[210] note in their most recent review on glycosynthases that the combination of the aforementioned L-fucosidase with a previously reported 1,2-α-fucosidase[212] would be beneficial to access the full range of Lewis antigens. Similarly, *Kwan et al.*[213] were able to produce the type 2 blood group A antigen in just 3 steps using a combination of glycosynthases and glycosyl transferases. This synthetic route demonstrates clearly the advantages of biocatalysts in complex carbohydrate synthesis when one compares this to the synthesis by *Hara et al.* who generated the B antigen in 20 steps with an overall yield of just 6%.[148] A more impressive demonstration of the power of enzymatic synthesis of oligosaccharides is the recent synthesis of lactosamine oligomers by *Henze et al.*[201] who synthesised a decasaccharide in a one-pot reaction using a glycosyltransferase and a glycosynthase, reporting yields of up to 31%.

Initially, glycosynthases were limited to the formation of oligosaccharides. This disadvantaged them in comparison to the glycosidases and glycosyltransferases, which had been demonstrated to couple aglycone acceptors.[179,214] *Fort et al.*[215] successfully utilised an endocellulosynthase from *Humicola insolens* (Cel7B) which accepted 6'-bromo, -amino, and -*S*-glycosyl cellobiosyl fluorides generating functionalised glycans. This enzyme was later found to accept 6'-azides allowing for further

functionalisation of the cellulose scaffold.[216] *Li* and *Kim*[217] were also able to generate a range of di- and trisaccharides using a series of aryl glycosides as acceptors with yields of up to 98% on multi-milligram scales to generate biologically-important molecules. More recently, focus has shifted towards aglycone acceptors such as flavonoids,[218] glycosphingolipids,[219] and steroids,[220] all of which are important biomolecules.

Since their invention in 1998, a number of glycosynthases have been generated and used to catalyse the formation of new glycosidic bonds. Both *endo*- and *exo*-glycosidases have been converted to their respective synthases and in addition, thioglycoligases and thioglycosynthases have also been developed which catalyse the formation of *S*-glycosides.[221] With this in mind, we chose a number of potential glycosynthases to pursue in the context of this work. These were identified based upon the reactivities of their corresponding wild type glycohydrolases, and the reaction products generated from hydrolase-catalysed transglycosylation.

Table 2-2: Potential glycosynthases and their reported activities

Enzyme	Organism	Type	Activity	Ref.
LacZ	*Escherichia coli*	β-galactosidase	(1,6)	[1]
Lac4	*Kluyveromyces lactis*	β-galactosidase	(1,4), (1,6)	[222,223]
Bgl1	*Pichia etchellsii*	β-glucosidase	(1,4), (1,6)	[224]
BglJ	*Aspergillus oryzae*	β-glucosidase	(1,6)	[225]
LacA	*Aspergillus oryzae*	β-galactosidase	(1,4), (1,6)	[222]
LacC	*Aspergillus oryzae*	β-galactosidase	(1,4), (1,6)	[222]

2.5 Structure, Form, and Function of Galactosidases

2.5.1 LacZ: a β-galactosidase from *Escherichia coli*

The β-galactosidase from *Escherichia coli*, LacZ, is one of the best characterised and well-studied enzymes to date as it is not only one of the first enzymes characterised,[226] but also is one of the proteins encoded as part of the *lac* (lactose) operon, one of the first genetic regulatory mechanisms described.[227] LacZ is a galactosyl hydrolase, and is responsible for the hydrolysis of lactose (**2**) into its fundamental components, glucose (**6**) and galactose (**7**) (Figure 2-6). LacZ displays optimal activity for catalytic purposes in slightly basic conditions (ca. pH 7.5) and 37 °C.[228] LacZ is frequently used in molecular biology as a reporter enzyme, relying upon α-complementation to indicate transformed colonies – typically referred to as a "blue-white" screen. This method is a qualitative indicator, with transformed colonies appearing blue due to the formation of indigo, a vivid, blue compound, and untransformed colonies appearing white. Early studies into the mechanism of LacZ identified its transglycosylation activity,[229] which converts the β-(1,4)-coupled lactose (**2**) to β-(1,6) allolactose

(**3**), subsequently identified as the true inducer of the *lac* operon and true substrate of LacZ (Figure 2-6).[230]

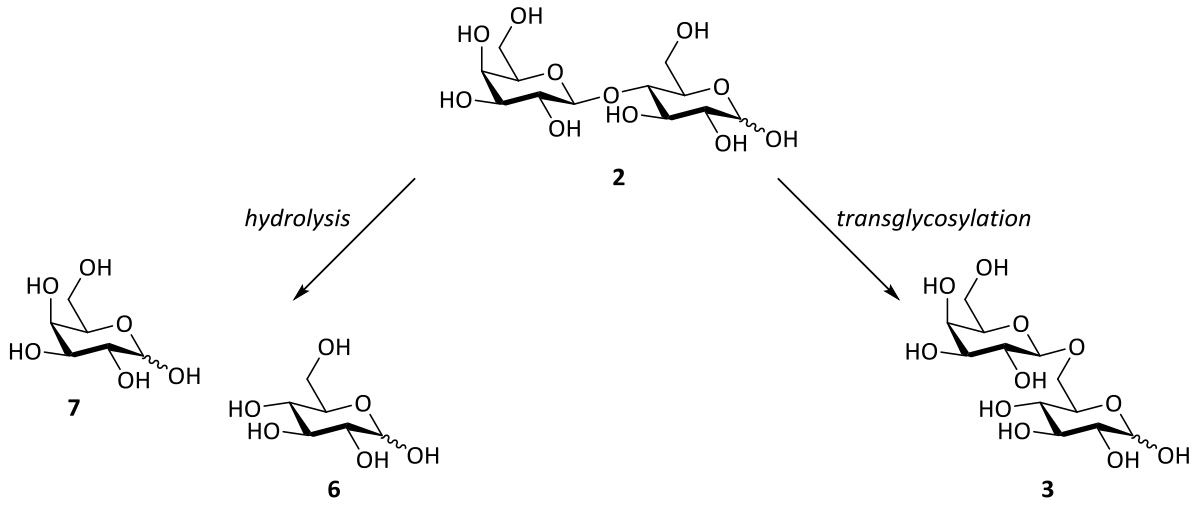


Figure 2-6: Two main reactions of LacZ on lactose (**2**): hydrolysis to form glucose (**6**) and galactose (**7**), and the transglycosylation reaction to form allolactose (**3**).

Structure of LacZ

Tertiary Structure

Studies on the β-galactosidase from *E. coli* have demonstrated that the major species of the enzyme is a homotetramer (Figure 2-7, left).[231,232] Each monomer consists of 1023 amino acid residues and the catalytically active tetramer contains 4 active sites. Interestingly, despite requiring a tetramer for activity, the active sites show no cooperativity or allostery.[232] The monomer contains 5 distinct domains made up of the majority of the 1023 proteomic residues (Figure 2-7, right).[233] Domain 1 (D1) is a jelly-roll β-barrel (residues 51-217) and leads into the fibronectin type-III-like barrel of domain 2 (D2) (residues 220-333), reminiscent of those found in immunoglobulin. The third domain (D3) is a distorted TIM barrel which is centrally located within the monomer (residues 334-627) and contains the majority of the catalytic machinery, located at the *C*-terminal end of the barrel. Unlike normal TIM barrels, the third domain lacks a 5th helix and contains a distorted β-strand. Domain 4 (D4) is a topologically-identical fibronectin-type barrel (residues 628-736) and domain 5 (D5), a unique 18-stranded β-sandwich (residues 737-1023).[232,233] The 50 residues of the *N*-terminus compose the α-peptide and it is this region which is involved in α-complementation.

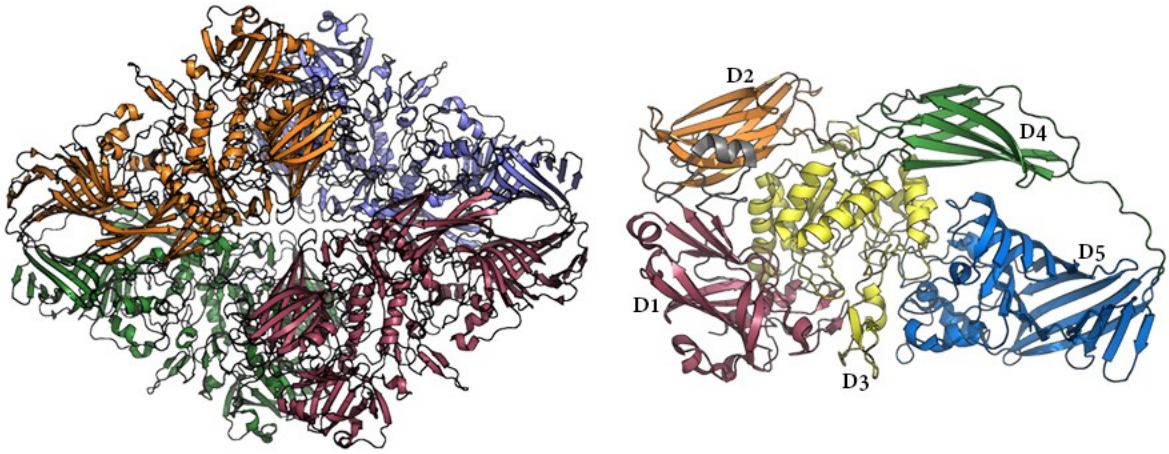


Figure 2-7: Quaternary and tertiary structures of LacZ. Left: Four monomers come together to form the active tetrameric structure. Right: The LacZ monomer is made up of five domains and the α-peptide (grey). PDB: 1JYX.

Quaternary Structure

The quaternary structure of LacZ is a homotetramer, or rather a dimer of dimers, and it is widely reported that only this tetrameric form is active,[231,234] although it has been reported to have low activity in a dimeric form.[235] Within the tetramer, there are 3 major intermonomeric interactions responsible for the formation and stability of the tetramer. Two of these are regions associated with the so-called "long interface" and the third belongs to the "activating interface" (Figure 2-8).[233] The one of the regions associated with the long interface involves three intramonomeric interactions between Arg561 and the backbone carbonyls of Leu524, Thr530, and Arg531; one intermonomeric interaction between Arg561 and Ser525, and a water-bridge directly between Arg561 of the neighbouring chain (Figure 2-9, bottom left).[232] The other long interface interaction involves intermonomeric hydrogen-bonding interactions between D4 and D5.[232] The activating interface occurs along the vertical axis of symmetry and is responsible for the activity of the LacZ tetramer.[234] A four-helix bundle generated by helices α13 and α14 of D3 (residues 429-448 and 465-479 respectively) is the largest contributor to the activating interface (Figure 2-9, top right). In addition, interactions involving D2 and the α-peptide complete the interactions required for formation of the tetramer and subsequent activity.[233] In particular, the completion of the active site by the D2 loop (residues 272-288) of the neighbouring monomeric unit is responsible for the activity of the tetramer (Figure 2-9, bottom right).[232] This also explains the lack of activity of dimeric and monomeric forms of LacZ.

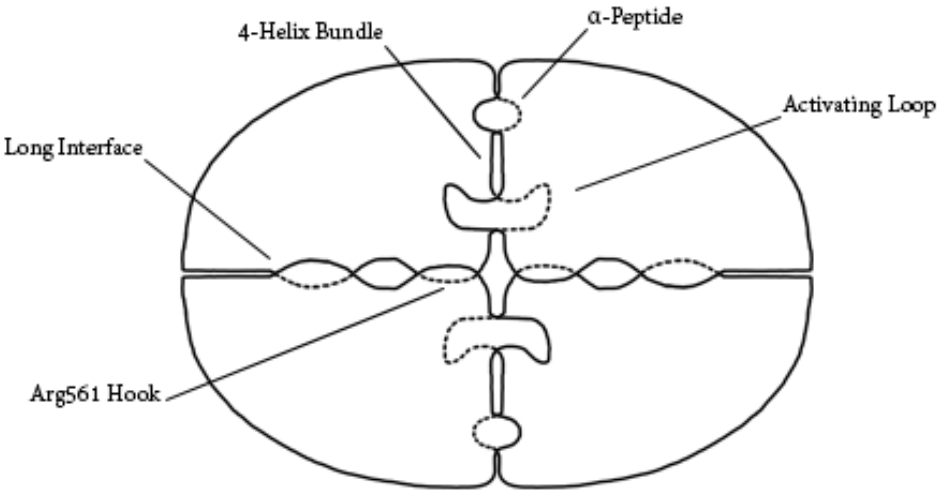


Figure 2-8: Schematic of the interactions between the monomeric units of LacZ which contribute to the oligomerisation.

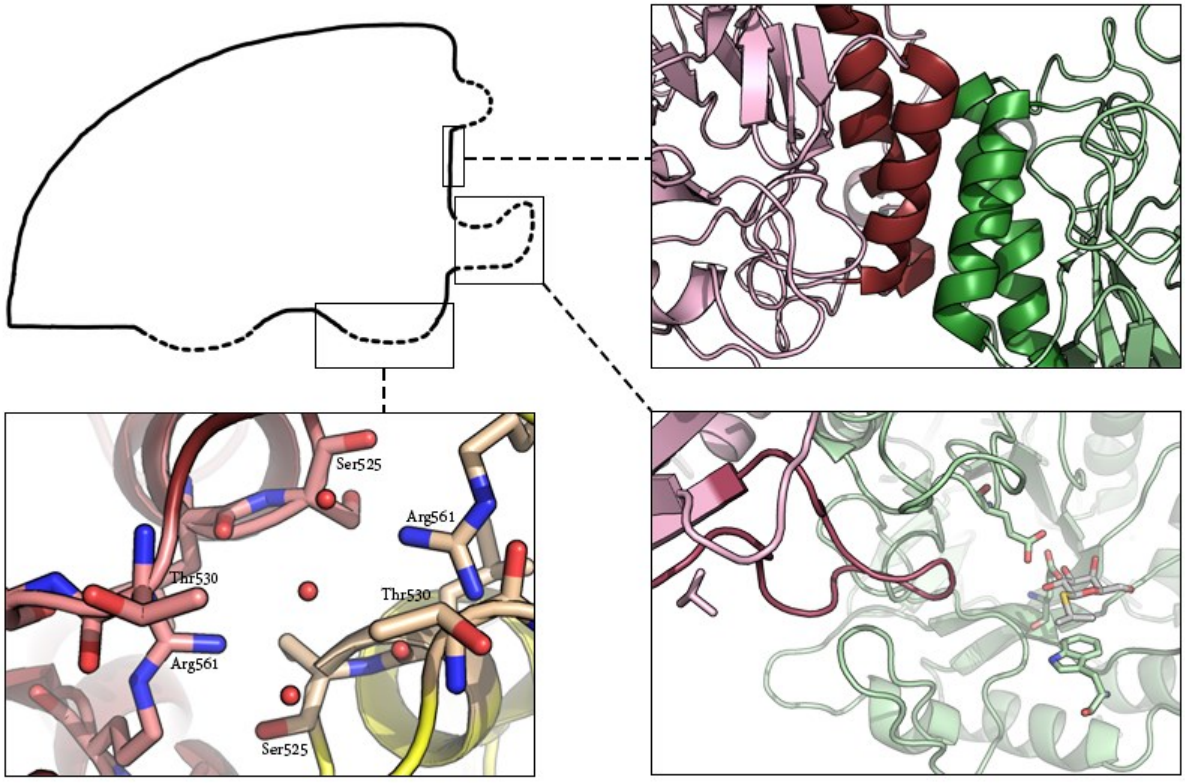


Figure 2-9: Expansions of the three of the major interactions of the monomeric unit. Clockwise from top left: Schematic drawing of the individual monomeric unit; the 4-helix bundle interaction which stabilises the active tetrameric form between monomer **a** (red) and monomer **b** (green). Positioning of the loop of domain 2 of the neighbouring monomeric unit (red) completes the active centre. The Arg561 "hook", which binds two monomers together along the long face to complete a dimer between monomer **a** (pink) and monomer **b** (pale yellow). PDB: 1JYX

α-Complementation

In early studies of the LacZ enzyme, α-complementation was observed during *N*-terminus deletion experiments. In these deletion mutants, Δ23-31 (M112)[236] and Δ11-41 (M15),[237] tetrameric LacZ was seen to dissociate into inactive dimers.[238,239] Though, when these deletion mutants were supplemented with a cyanogen bromide-derived peptide of residues 3-92 ($CNBr_2$), a functional tetramer would be re-formed.[240] Other studies have found that the α-donor 3-41 is also accepted by the α-acceptors M15 and M112, enabling the reformation of active tetramer.[241] It should also be noted that while the α-complemented enzyme displays comparable activity to that of the wild type, it is more susceptible to thermal and chaotropic denaturation.[232,241,242]

Since its discovery, α-complementation has now become a popular selection method, wherein an α-peptide fragment is co-expressed with a protein-of-interest, on a plasmid, in an *E. coli* strain possessing one of the aforementioned LacZ deletion mutants. This allows for easy identification of plasmid-containing *E. coli* cells using a blue-white screen.[234] Importantly, whilst this α-peptide is critical for the formation of the active tetramer, the identity of residues 1-26 are unimportant.[243] Mutations of residues 27 onwards are more detrimental to the formation and stability of the tetramer. Residues 13-20 in adjacent monomers contact each other directly.[234] Of particular importance in this region is Glu17 as mutations of this residue increase the rate of denaturation of the tetramer to the dimer.[242] The dimers themselves are also shown to be unstable and will rapidly dissociate further to the monomeric forms unless sufficient Na^+, Mg^{2+}, and thiols are present.[242,244,245] Residues 22-31, while positioned close to the activating interface, do not contact the neighbouring dimer, but rather interact with D1 and contribute to the stabilisation of the D3-D3 four-helix bundle (Figure 2-10).[232,233] As mentioned before, the identity of residues 1-26 appear unimportant for dimerisation, residues 27-31 do appear important as they pass through a tunnel created by D1, D2, and D3, which assist in anchoring the α-peptide. This increases the overall stability of the tetramer.[232] As such, alteration of the identity of these residues lowers the thermodynamic favourability of the α-peptide to occupy this tunnel space, and in turn the favourability of tetramer formation.

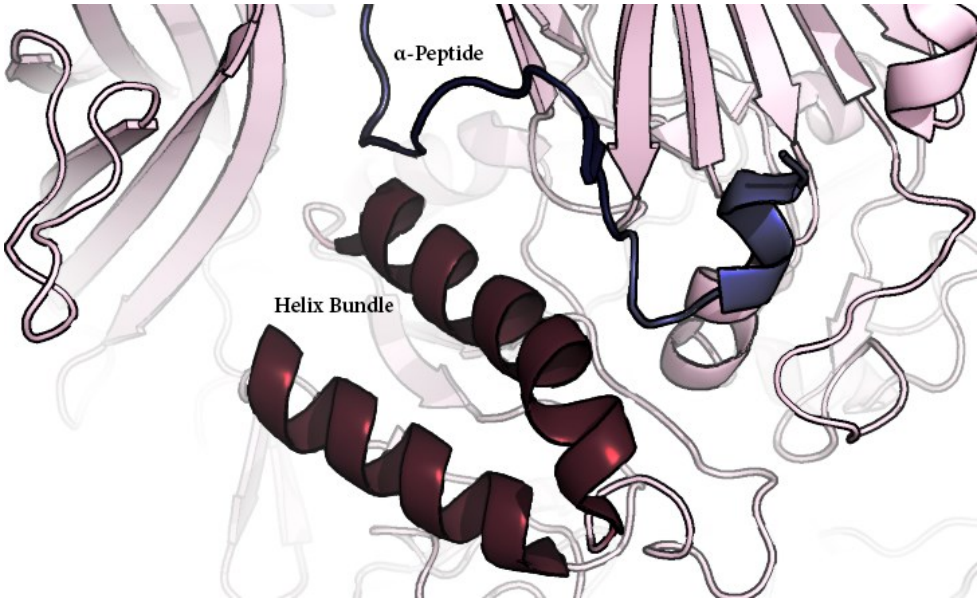


Figure 2-10: The α-peptide (dark blue) is positioned directly over the 2 helices (dark red) which contribute to the 4-helix bundle responsible for the stability of the tetrameric form. PDB: 1JYX.

Catalytic Centre

The active site of LacZ is located at the *C*-terminal end of the D3 TIM-barrel (Figure 2-11).[233] As LacZ is of the family GH 2, it follows a simple retaining mechanism comprising a nucleophilic residue and a catalytic acid/base residue, Glu537 and Glu461 respectively.[246,247] In contrast to early models of the catalytic pocket and mechanism, the two residues are not positioned 180° apart from one another, but rather 90° and in close proximity. The adjacency of these catalytically-important carboxylic residues is crucial to the pK_a cycling required for catalysis in glycosidases.[248] As a result, Glu461 must be protonated in the ground state due to the carboxylate anion of Glu537. This positioning within the molecule increases the acidity of the Glu461 proton during the first step by lowering the effective pH of the active pocket to less than 4.[249] This enables Glu461 to more easily donate it's proton to the leaving group and quench the increasing negative charge generated by formation of the sugar-enzyme complex.[249,250] Indeed, computational studies show that the pK_a of Glu461 in the free enzyme is high and significantly decreased in the substrate-enzyme complex, favouring transfer of water to the galactosyl moiety.[249]

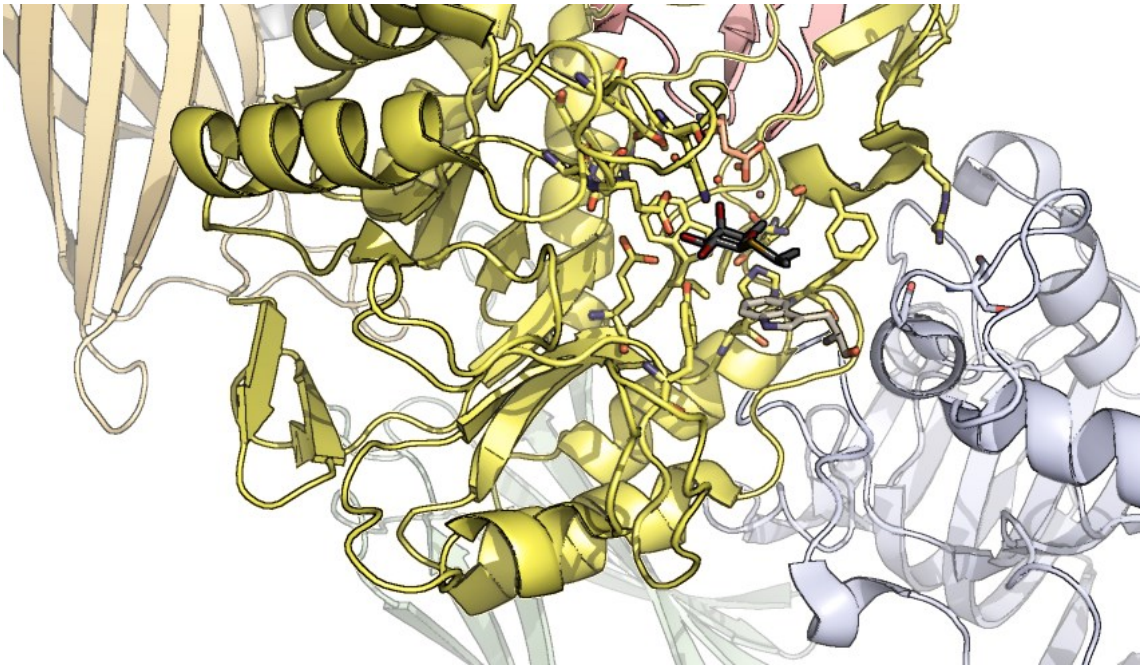


Figure 2-11: Domain 3 (yellow) highlighted in comparison to the rest of the monomeric unit with the active site located at the *C*-terminal end of the β-barrel is easily identified by the IPTG (dark grey). PDB: 1JYX.

Surrounding the two catalytic residues are a number of residues, ions, and small molecules important for pocket structure,[251-256] stabilisation of transition states,[257-260] substrate positioning,[260-263] and optimisation of catalytic chemistry.[247,249,264-268] One of the most curious of features is the so-called "active-site loop", which consists of residues 794-803 from D5. The loop, which can have either open or closed conformations, is known to be important for the activity of the enzyme, particularly through its interactions with Arg599.[255] It appears to play a role in modulating substrate binding affinities by encouraging substrate binding when open, and transition state stabilisation when closed. Additionally, it also appears to play a role in transglycosylation and the formation of allolactose as it has been suggested that it plays an important role in glucose binding (Figure 2-12).[234,254,269] When substrate is unbound, the loop takes on the flexible, open conformation and closes when the substrate is bound to the deep binding site.[256] Mutations of Gly794, the loop hinge, have led to more active galactosidic activity, particularly when mutated to an acidic residue.[251] The conformation dependence of the loop to the presence and location of substrate can be linked to the 6-hydroxyl group and its neighbours. When no substrate is bound or substrate is bound in shallow mode, Phe601 interacts hydrophobically with the galactosyl C6 and Met542, edge-wise and facially respectively, and Arg599 interacts with the aromatic system edge-wise *via* a cation-π bond.[254] This guanidinium group also interacts with Glu797 and forms hydrogen-bonds with the carbonyls of Ser796 and Gly794. As the carbonyl group of Phe601 is coordinated to Na^+, it undergoes a conformational change upon deep-binding of the substrate. The disruption of the cation-π bond between Phe601 and Arg599 results in Arg599 moving away and distorting, in turn severing the

interactions of the guanidinium group. Ser796 then moves towards the active pocket, where the α and β carbons can interact hydrophobically with Phe601, closing the loop.[255]

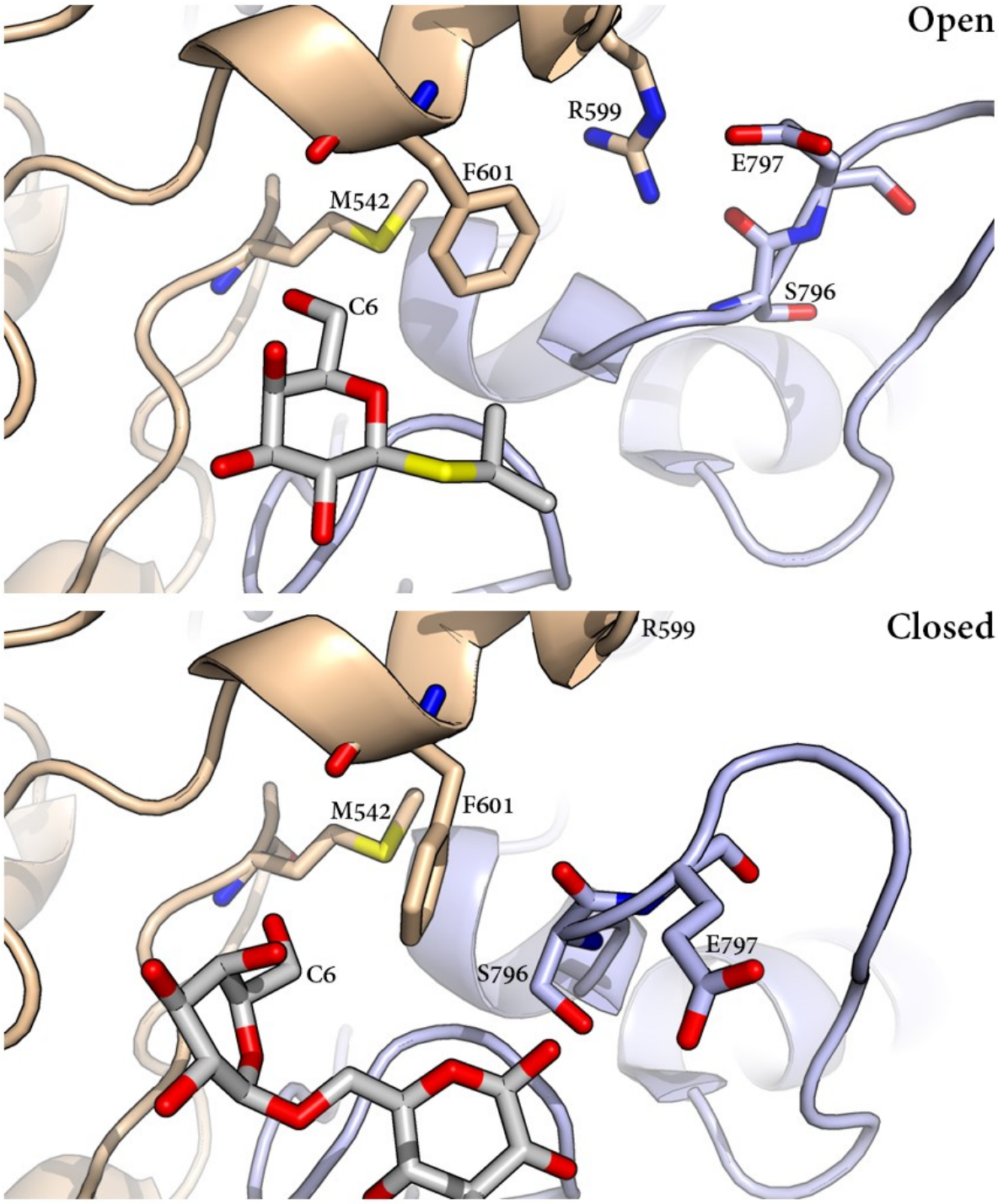


Figure 2-12: Comparison of the active site loop region between the open conformation and the closed conformation. PDB: 1JYX, 4DUW

Mechanistic Studies

The active site requires more than just the two catalytic glutamic acid residues to function correctly as the enzyme is also specific to galactopyranose and thus, substrate positioning within the active site is critical to the correct function of the enzyme. It has been demonstrated *via* crystallographic studies that LacZ possesses two, overlapping substrate binding modes: shallow (Figure 2-13) and deep (Figure 2-14). As the natural substrate of LacZ is allolactose, it could be suggested that the two binding sites correspond to glucose and galactose, respectively. However, LacZ is specific only to *exo*-galactopyranosyl moieties and more promiscuous in its acceptance of aglycones.[226,228,270-272]

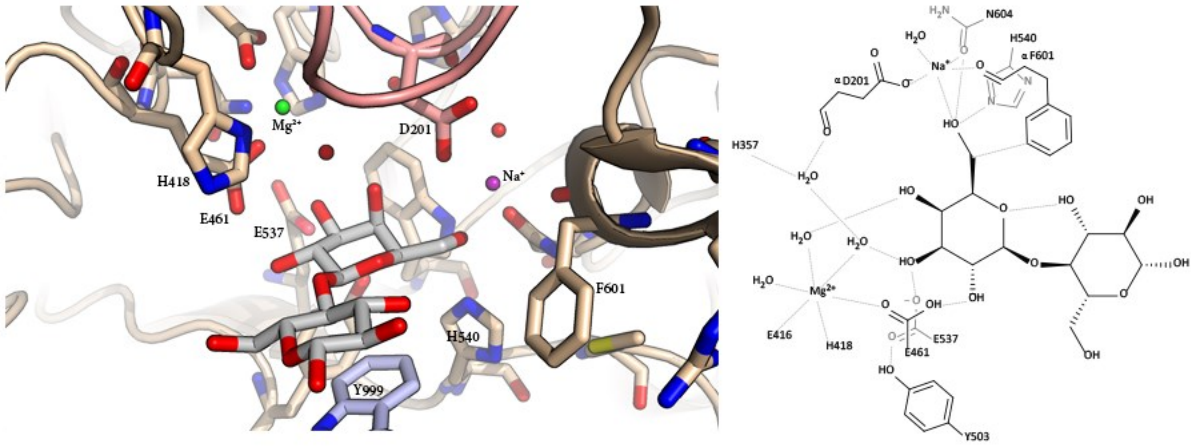


Figure 2-13: The binding of the substrate, lactose, in the "shallow" site of LacZ. PDB: 1JYN

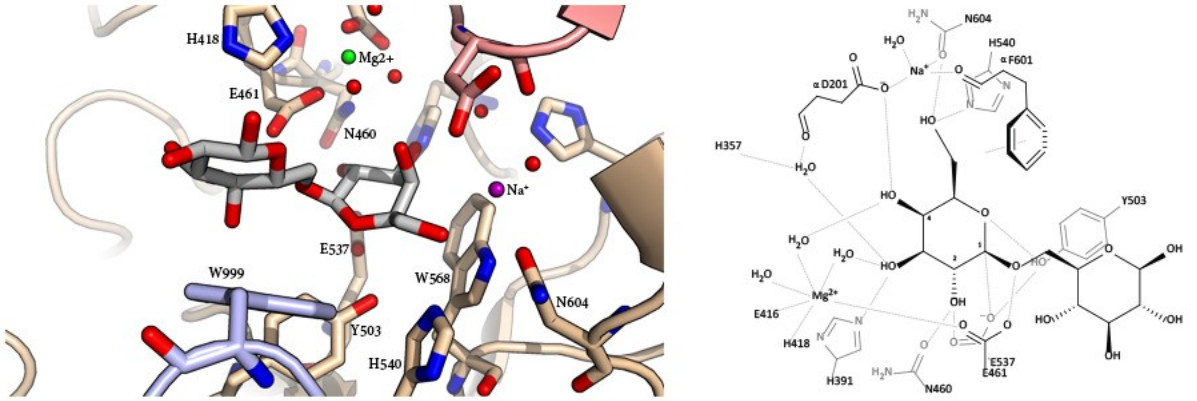


Figure 2-14: The binding of the transglycosylation product, allolactose, in the "deep" site of LacZ. (PDB: 4DUW)

Binding Modes

Upon entering the shallow site, the galactosyl unit stacks on Trp999, a highly conserved residue in related β-galactosidases, with the oxygen of the galactosyl ring positioned above the indole and the glycosyl carbons within van der Waals range.[234,261] Following this initial interaction, a web of specific hydrogen-bonding interactions between the galactosyl moiety and the active site are formed. It

should be noted that the vast majority of interactions in this shallow mode are between the galactosyl unit and the enzyme, very few interactions involve the aglycone.[250] The 6-hydroxyl group forms the strongest interactions with the active pocket, bonding with His540, Asn604, and Na^+. These interactions then presumably act as a hinge, allowing further hydrogen bonds to develop between 4-OH and Asp201 and a water molecule within the coordination sphere of Mg^{2+}. The 2- and 3-hydroxyl groups also interact directly with Glu537 and Glu461, but also with Tyr503, Asn460 and Mg^{2+} indirectly.

Anchored by the interactions between the 4- and 6-hydroxyl groups, the galactosyl moiety is moved into the deep mode binding site, undergoing a 90° change in orientation (Figure 2-15), possibly *via* an inverted boat conformation.[250] This transition between binding sites is concurrent with the closing of the active-site loop, which undergoes a shift of about 11 Å.[255] This creates a new series of hydrogen bonds around the 2- and 3-hydroxyl groups as well as moving the anomeric centre into an optimal position for the catalysis to occur. In this deep binding mode, the 3-hydroxyl has displaced a water molecule and now interacts with a water molecule within the coordination sphere of magnesium. It now also interacts directly with His391 and a water molecule coordinated to His357 and the carboxylic oxygen of Asp201. The 2-hydroxyl position now interacts with Asn460 and, importantly, Glu537. This presumed hydrogen bond between Glu537 and 2-OH has been calculated to contribute >10 kcal mol^{-1} to the stability of the transition state.[175]

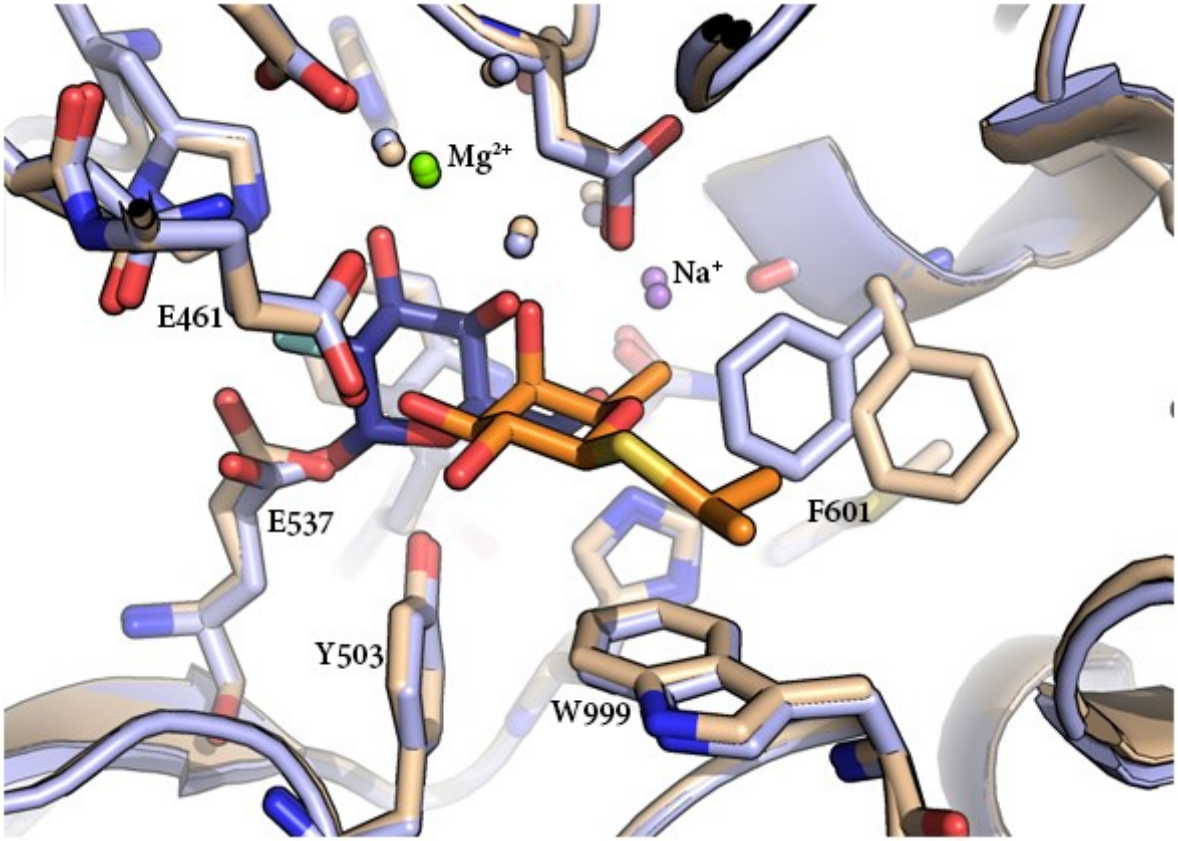


Figure 2-15: The change in position of the substrate between the shallow and deep binding modes created by an overlay of the "shallow" bound substrate, IPTG (orange), and the "deep" bound 2-deoxy-2-fluorogalactose (blue) in the active site. PDB: 1JYX, 1JZ2

Currently, it is not possible to view transition states and docking of the lactose/allolactose moieties in the native enzyme as the substrate undergoes hydrolysis. Furthermore, crystallisation studies using lactose and allolactose as substrates of catalytically inactive mutants only show these molecules binding in the shallow mode.[250,269] Therefore, substrate mimics, such as IPTG or 2-fluoro-2-deoxy-β-D-galactosyl substrates, or transition state mimics, like galactonolactone or galactotetrazole, are soaked into the crystals.[250,253-256,259,265] Despite the crystal structures showing the galactosyl substrates to maintain a $^4C_1$ conformation in both the shallow and deep binding modes, it has been suggested that the pyranose ring may be flattened during the oxocarbenium transition state. This hypothesis is supported by the high binding affinities of the transition state analogues galactonolactone and galactotetrazole, which demonstrate 10^3-10^8 better binding affinities compared to galactose.[250] Transferred nuclear Overhauser effect (TR-NOE) ^{1}H NMR studies using the inhibitor *C*-lactose, the *C*-glycoside derivative of lactose, revealed that in the deep binding mode, lactose most likely takes on a high-energy, gauche-gauche conformer.[258] This hypothesis was recently confirmed by crystallisation of allolactose in a Gly794 mutant of the enzyme.[269]

The Role of Metal Ions in LacZ

The first crystal structures of LacZ indicated a number of metal-binding sites within the enzyme which are both structurally and catalytically important.[233] Structural metal ions, such as the Mg^{2+} cation that interacts with the α-peptide and the 4-helix bundle, are not explicitly required for formation of the tetramer and, by extension, activity. However, it has been shown that a Mg^{2+} cation and a Na^{+} cation are of catalytic importance (Figure 2-16).[273] The two metal ions are located around 8 Å apart and are linked by His418, Val103, Asn102, and Asp201. Moreover, the affinity of the protein for either ion appears to be linked to its affinity for the other.[265]

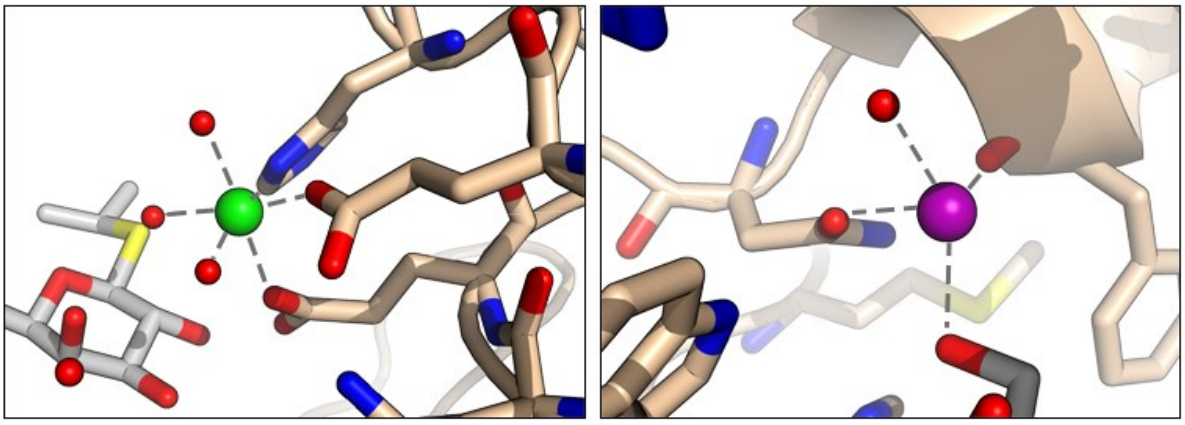


Figure 2-16: Coordination spheres of the metal ions in the active site of LacZ. Left: Mg^{2+}. Right: Na^{+}. PDB: 1JYX

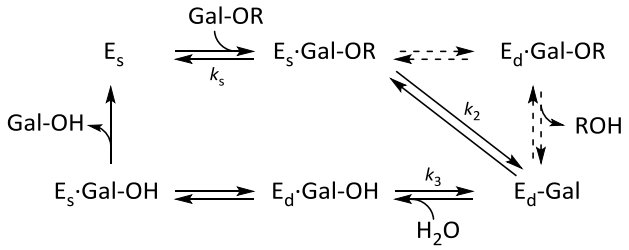
LacZ requires a single Mg^{2+} cation per subunit for activity; however, up to 4 others can bind to other parts of the monomer.[264] The critical Mg^{2+} cation is octahedrally-coordinated in the rear of the binding pocket by three water residues, Glu416, His418, and the catalytic residue, Glu461.[233,252] Initially proposed to directly participate in catalysis, acting as a catalytic electrophile,[274] this was revised to infer properties relating to the maintenance of structural integrity.[275] Recently, site-directed mutagenesis studies of Glu416[264] and His418[265] have shown that Mg^{2+} is important in modulating active site chemistry by modulating the Brønsted acid/base properties of Glu461.[264] Early studies found that LacZ could accommodate a number of divalent cations,[276,277] although this was later found to be due to metal impurities and it was demonstrated that only Mn^{2+} could be substituted with near-to-native activity.[273] The enzyme also displayed activity with Ca^{2+}, albeit at a significantly lower level. Calculations have revealed that Mg^{2+} lowers the activation energy required for glycosylation of the enzyme by close to 15 kcal mol^{-1}.[278]

The role of Na^{+} in LacZ is believed to be two-fold. Firstly, as Na^{+} coordinates in both shallow and deep binding modes with the 6-hydroxyl group of galactopyranosides, its role in substrate positioning is clear.[233,250,263] Secondly, Na^{+} is believed to play a role in stabilisation of the protein and facilitate the conformational switch between open and closed forms.[250] In the active pocket, Na^{+} is

coordinated square pyramidally by two water molecules, Asp201, the backbone carbonyl of Phe601, and Asn604 with one of the water molecules being replaced by the 6-hydroxyl group upon binding of galactose.[233,264] Substitution of Na^+ with monovalent K^+ shows no significant difference in enzyme activity and both metals provide only modest reductions in activation energy.[264,278]

Mechanism of Hydrolysis

The mechanism of hydrolysis goes through a two-step mechanism in which the formation of the enzyme-substrate complex is rate-determining (Scheme 2-12).[249] The reaction goes through two distinct steps, "galactosylation" [of the enzyme] and "degalactosylation", represented by the rate constants k_2 and k_3, respectively. Galactosylation comprises all steps between the binding of the substrate to the covalent intermediate (E_s·Gal-OR to E_d-Gal), whereas the degalactosylation refers to the transfer of the galactosyl unit to the acceptor (E_d-Gal to E_d·Gal-OH). Substrate binding and release are represented by the constant, k_s.



Scheme 2-12: Reaction mechanism of β-galactosidase-catalysed hydrolysis.[234] Binding of the substrate (Gal-OR) to the enzyme (E_s) initially occurs in the shallow site (E_s·Gal-OH) before the substrate is moved to the deep site (E_d·Gal-OH). Galactosylation of the enzyme then occurs to form the enzyme-sugar intermediate (E_d-Gal). Hydrolysis of the intermediate by water yields the product complex (E_d·Gal-OH), which is then moved to the shallow site (E_s·Gal-OH) and released to return the free enzyme. In cases of (trans)glycosylation, the galactosyl acceptor is another alcohol rather than water.

During substrate binding, the active-site loop is in an open position, granting the substrate easy access to the catalytic pocket. Substrate binding occurs in the shallow site by interactions between the aromatic system of Trp999 and the partially charged hydrogens on the underside of the galactosyl chair.[261] At this point, a network of hydrogen bonds are formed between the galactosyl moiety and the enzyme, in addition, hydrophobic interactions between C6 and Phe601, and an intramolecular hydrogen bonding interaction between O5 of galactose and the 3-hydroxyl of the glucoside.[250] Once complexed to the enzyme, the galactoside is moved approximately 3 Å to the deep site and stacked upon Trp568, relying on the same interactions between the aromatic system and the partially-positive ring hydrogens.[234] The movement around an axis created by 4-OH and 6-OH displaces a water molecule and forms a new, closer network of hydrogen bonds between the enzyme and the remainder of the carbohydrate, allowing for effective stabilisation of the transition state, and the active-site loop is concurrently closed.

The energy required for galactosylation (k_2) includes the energy for substrate transition from the shallow site (E_s·Gal-OH) to the deep site (E_d·Gal-OH) as well as that of the formation of the covalent intermediate (E_d-Gal). The movement from shallow to deep mode is an unfavourable equilibrium as *Juers et al.*[250] noted an absence of electron density in crystallographic analyses of E537Q crystals, which had been soaked with the common hydrolytic substrates lactose, allolactose, and *o*-nitrophenyl β-D-galactopyranoside.[250] They also observed a lack of deep binding when native enzyme was incubated with the substrate mimic, IPTG, adding weight to this hypothesis. Computational studies suggested that galactose exists in a "pre-transitional", $^4H_3$ conformation within the deep pocket and is converted to a ^4E conformation during the transition state (Figure 2-17).[278] However, the proposed $^4H_3$ conformation does not agree with NMR studies of *C*-lactose,[258] nor with observations of galactose in crystallisation studies.[250] This suggests that distortion associated with transition state formation occurs as Glu461 donates its proton to the glycosidic bond. As the result of proton donation, the glycosidic oxygen develops a positive charge, polarising the C1-O1 bond, shifting charge to C1. C1 is now partially positive and as such, O5, the heterocyclic oxygen donates its electrons, forming a partial double bond between the two. At this point, C1 and O5 are exhibiting partial sp^2 hybridisation, and as a result, the ring shifts towards a more planar, ^4E conformation. This also gives credence to the generation of an oxocarbenium-like cationic species which is widely accepted as an intermediate in the mechanism.[123]

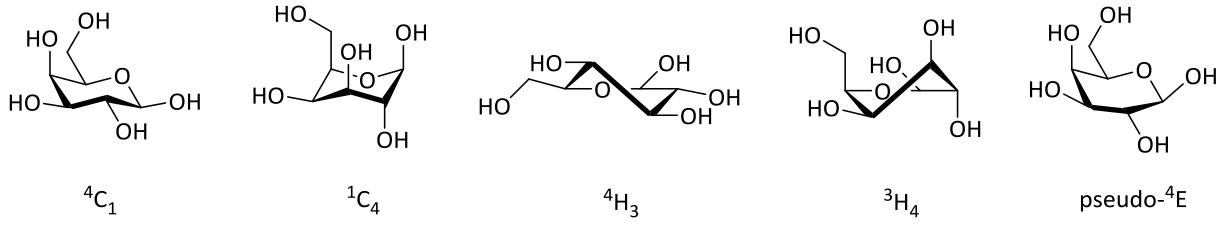
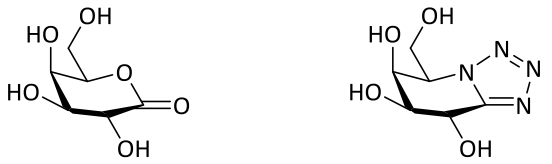


Figure 2-17: Different ring conformations of galactopyranose.

The reaction itself most likely proceeds *via* an S_N2-like mechanism due to the planarity of the ring transition state, a requirement of S_N2 mechanisms. Furthermore, should the reaction proceed *via* an S_N1 mechanism, the electron donation from O5 to C1 would require the electrons to be antiperiplanar to the breaking C1-O1 bond, which would require significant distortion of the ring.[279] Two molecules, which exhibit ^4E planarity and bind well in the deep mode, are the transition state analogues, galactonolactone (**19**) and galactotetrazole (**20**), and have properties which support the planarity of the transition state and furthermore, an S_N2-like mechanism (Figure 2-18).[250] Evidence of the quality and nearness of these analogues to the true transition state comes from a number of site-directed mutagenesis experiments on the transition state-stabilising residues, His357,[262] His391,[257] and Asn 460,[259] which demonstrated that the binding of these analogues correlated with the k_{cat}/K_M values of the mutant enzymes.



19 **20**

Figure 2-18: Transition state analogues of galactosidase hydrolysis, galactonolactone (**19**) and galactotetrazole (**20**) which exhibit ^4E conformations thought to be reminiscent of the oxocarbenium intermediate.

The enzyme-galactosyl covalent intermediate occupies a $^4C_1$ chair conformation as evidenced by crystallographic studies where the enzyme was reacted with 2,4-dinitrophenyl 2-fluoro-2-deoxy-β-D-galactopyranoside.[250] The leaving group properties of 2,4-dinitrophenyl make it ideal for acceleration of k_2, and 2-fluoro-2-deoxy greatly decreases the rate of k_3, allowing for trapping of the covalent intermediate.[280] This substitution allows for the maintenance of any stabilising effects *via* hydrogen bonding, whilst simultaneously exploiting the charge build-up at C1. Due to the highly electronegative nature of fluorine, the oxocarbenium-like transition state is destabilised, making 2-fluoro-2-deoxy substrates more likely to remain in the covalently-stabilised intermediate form.

Before galactosylation of the acceptor can occur (Figure 2-19), a negative charge needs to first be restored to Glu537. It is believed that this is assisted by partial proton donation from Tyr503.[268] The now generated trigonal oxocarbenium cation is stabilised by the now negatively-charged Glu537, Tyr503, and O5. Furthermore, now that Glu537 is again charged, Glu461 is now a stronger base, more easily deprotonating the acceptor, better facilitating nucleophilic attack at the anomeric centre.[249,250] Due to the trigonal nature of the oxocarbenium species, the galactosyl ring is naturally distorted back to something akin to a ^4E configuration. Whether the incoming water molecule is positioned whilst the enzyme is covalently bound to the galactosyl unit, or whether it is positioned during the formation of the second oxocarbenium-like transition state in degalactosylation is unknown. Regardless of when it is positioned, it is activated by deprotonated Glu461 *via* general base catalysis. Secondary deuterium isotope effect studies for this step indicated that the rate of k_3 is related to the identity of the acceptor molecule.[234] Precise alignment of the incoming acceptor with the galactosyl moiety would also be expected for degalactosylation; an experimentally-supported expectation.[281] The puckering of the ring to the ^4E conformation rotates the galactosyl residue slightly, altering the interaction of the 6-hydroxyl group with Na^+. This adjustment appears to concurrently induce a conformational change to re-open the active-site loop.[250]

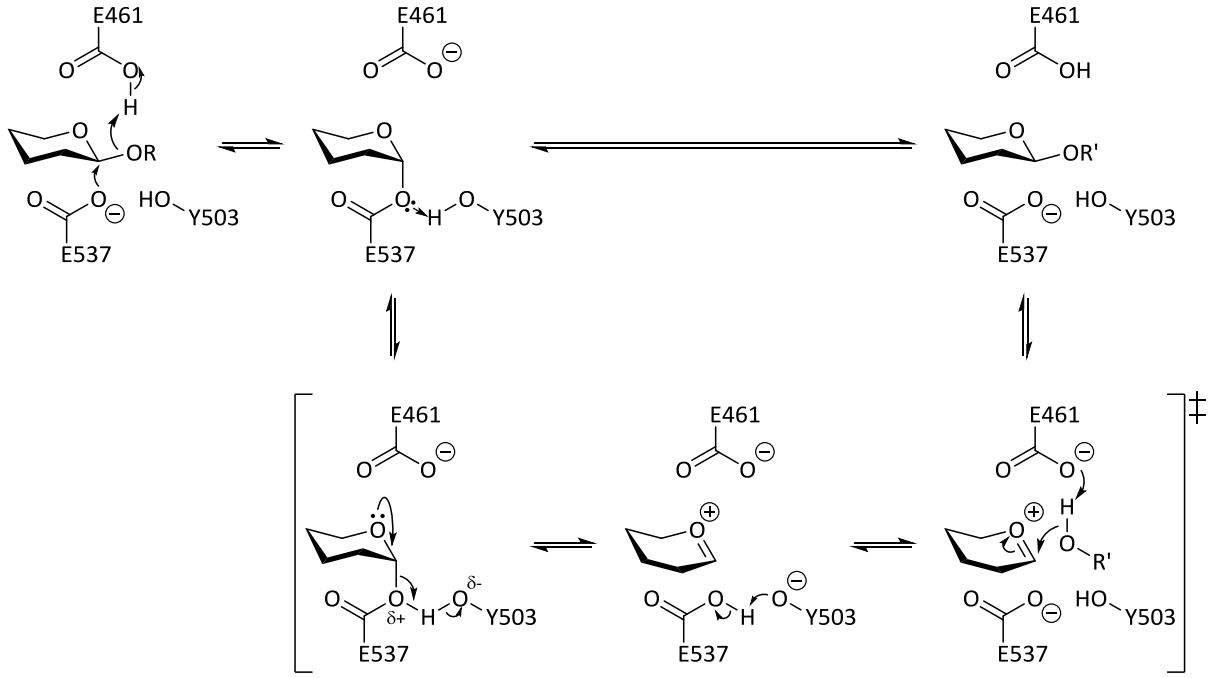


Figure 2-19: Expanded mechanism highlighting the role of Tyr503 in the mechanism of degalactosylation of the enzyme/galactosylation of the acceptor. After the formation of the enzyme-sugar intermediate, Y503 partially donates its proton to facilitate the formation of the oxocarbenium intermediate. After reformation of the negative charge on E537, E461 is better able to activate the galactosyl acceptor due to the presence of two anions in the active site.[249]

In the case of hydrolysis, the aglycone product, glucose, diffuses out of the active pocket. In the case of transglycosylation, the reaction whereby allolactose is generated, the glucoside is required to stay longer in the active pocket and kinetic studies have shown that glucose release occurs slowly.[269,270] It is known that interactions with Asn102, His 418, Lys517, Ser796, Glu797, and Trp999 bind the glucose long enough to keep it in the pocket (Figure 2-20).[250,269] The selectivity of the 6-hydroxyl as the galactosyl acceptor as opposed to a re-galactosylation of the 4-hydroxyl is due to repositioning of the glucoside unit following cleavage to a more energetically-favourable site within the shallow site, also referred to as the "glucose site" in terms of transgalactosylation. Furthermore, the primary 6-hydroxyl group is more flexible than the secondary hydroxyls of the 2-, 3-, or 4- positions, especially considering that the 6-hydroxyl group can reach further into the active site, increasing the likelihood of a (1,6)-glycosyl bond forming.[250]

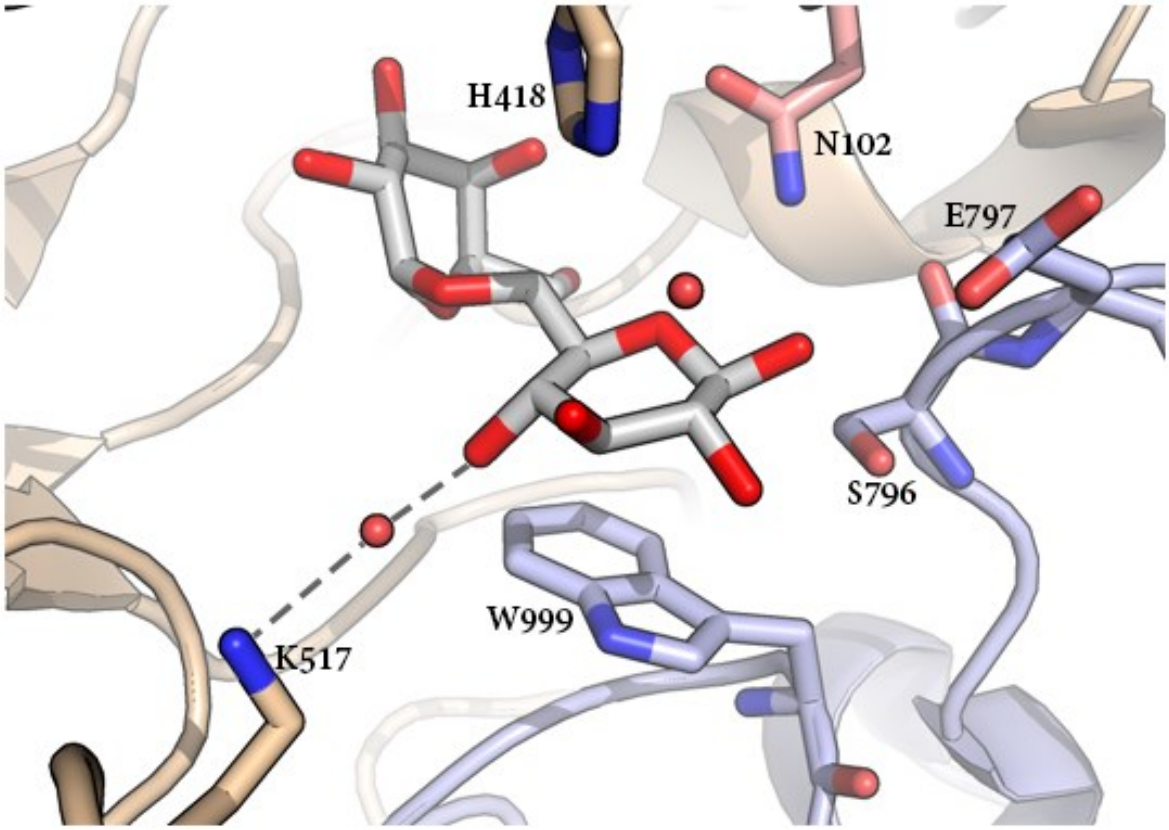


Figure 2-20: The "glucose" site of LacZ involving a number of residues important in the shallow binding of substrates. Noticeably, K517 forms a water bridge (grey dashed line) to the 4-OH lending itself to the specificity for equatorial 4-OH of glucose. PDB: 1JZ8

Trp999 is a key residue in the encouragement of allolactose formation, as the π···H-C interactions with hydrogens in the 3- and 5-positions are required for the retention of glucopyranose in the active site, and substitutions of this residue display a lower rate of allolactose production.[261] In the wild type enzyme, it was shown that the initial rate of hydrolysis-to-transgalactosylation is 47:53,[270] and substitution experiments for Trp999 show a greatly decreased affinity for glucose in the shallow site.[261] Only W999F mutations showed any significant allolactose production, albeit greatly reduced in comparison to the WT enzyme. Other studies have demonstrated the specificity of the "glucose site" using C3 and C4 anomers, demonstrating specificity for equatorial hydroxyls in these positions.[282] Crystal structures of deep-bound allolactose indicate that water bridges between glucose and His418 and Lys517 are most likely responsible for this observation.[269] These same structural analyses indicated specificity for β-glucopyranose due to specific interactions between the anomeric hydroxyl and three residues: Asn102, Ser796, and Glu797 – the last two being members of the active-site loop, giving credence to the loop's role in glucose binding and by extension allolactose production. Finally, *Wheatley et al.* identified a so-called "allolactose synthesis motif" consisting of the residues 517, 599, the loop 795-805, and 808.[269] This motif was conserved in 53 β-galactosidases

and the majority of the organisms in which this motif was conserved also possessed repressors homologous to the *lac* repressor of *E. coli*.

2.5.2 Lac4: a β-galactosidase from *Kluyveromyces lactis*

The yeast β-galactosidase from *Kluyveromyces lactis*, Lac4, is another biotechnologically-important enzyme and is naturally found in dairy environments and hydrolyses lactose. Despite this biotechnological relevance, the enzyme and its properties have not been studied to the same extent as the β-galactosidase from *E. coli*. As *K. lactis* is a natural inhabitant of dairy products, it is unsurprising that the enzymatic activity is naturally optimised for such conditions. By virtue of this, Lac4 is of vital importance within the dairy industry for the de-lactosification of milk products.[283,284] Specifically, the native enzyme displays optimal activity at neutral pH and 35 °C.[285] Reduced-lactose dairy foods are industrially important, as it is estimated that over 70% of the world's population is lactose-sensitive or -intolerant, with the highest proportion being of East Asian, central and southern African, and South American descent.[284] Treatment of lactose-containing products by β-galactosidases can either be performed during production of dairy products, or can be administered orally as tablet preparations by those who are lactose-sensitive. Furthermore, crystallisation of lactose in dairy products, such as condensed milk and ice cream, can be problematic for the overall product quality. As a result, hydrolysis of lactose into its base components, galactose and glucose, not only reduces crystallisation due to higher product solubilities, but also increases sweetness of the products and improves digestibility.[284,286]

Structure of Lac4

Like LacZ from *E. coli*, Lac4 is also a galactohydrolase from GH 2 and therefore shares a number of comparable properties with other GH 2 enzymes. These two enzymes are homologous and show 30% sequence identity and 47% similarity with one another.[287] Despite this high similarity, there are key differences in Lac4 which affect substrate binding and specificity and protein stability. Lac4 consists of 1025 residues with a molecular mass of 119 kDa,[283,288] though being of fungal origin, the true mass is closer to 200 kDa due to post-translational glycosylation of the enzyme.[289] Similar to LacZ, the tertiary structure of Lac4 is folded into 5 domains (denoted D1-5) (Figure 2-21). Lac4 D1 (residues 32-204) is a jelly-roll fold and is classified as a GH2 carbohydrate binding domain.[283] D2 and D4 (residues 205-332 and 643-720, respectively) are immunoglobulin-like β-sandwiches. D3 (residues 333-642), as with LacZ, is a TIM barrel, which comprises the majority of the active site. However, unlike LacZ, Lac4 displays a more typical morphology. D5 (residues 741-1025) is a large, 18-strand β-sandwich.[283]

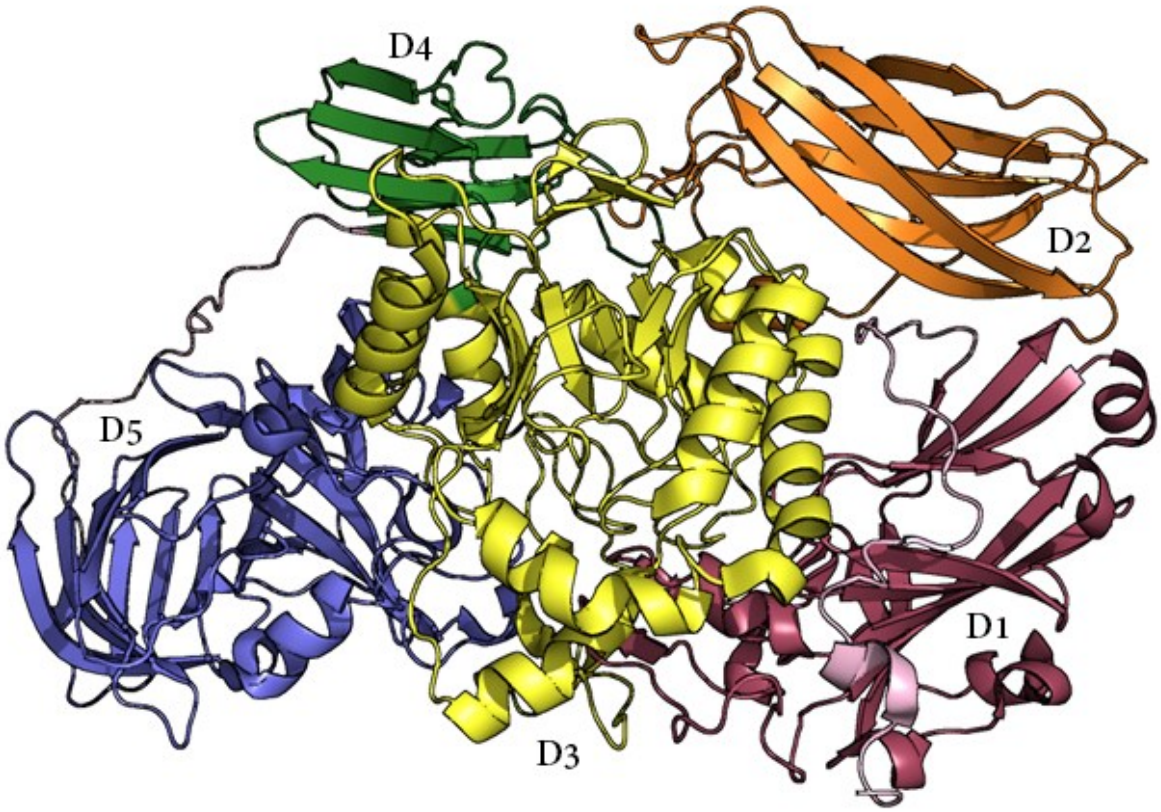


Figure 2-21: The five domains (D1-5) of the Lac4 monomer. PDB: 3OB8

The quaternary structure of Lac4 is also a dimer of dimers (Figure 2-22), although unlike LacZ in which only the tetrameric form is active, Lac4 is active in both the dimeric and tetrameric forms, and furthermore, these two forms display different activities.[283] Ultracentrifugal analyses suggest that the dimeric form is the major form of the multimeric enzyme and that there is a likely equilibrium between the dimeric and tetrameric forms.[283] Three main intermonomeric interfaces exist in the Lac4 tetramer, dubbed surfaces 1-3 (S1-3). S1, between monomers A and C, is the largest interface and is responsible for the stability of the dimer. The majority, approximately 75%, of contacts in S1 are hydrophobic and involves residues from D1, D3, D5, and unassigned *N*-terminus residues – similar to those of the α-peptide in LacZ. α-complementation has not been observed for Lac4 despite this region being shown to be critical for activity.[283] The S2 and S3 interactions are responsible for the formation and stability of the Lac4 tetramer and while still being predominantly hydrophobic, there are a higher proportion of non-hydrophobic contacts in these two regions.[283] S2, slightly smaller than S1, is present in monomers A and B and is believed to stabilise the formation of the tetramer. The interface here involves contact from all five domains and the unassigned *N*-terminal region.

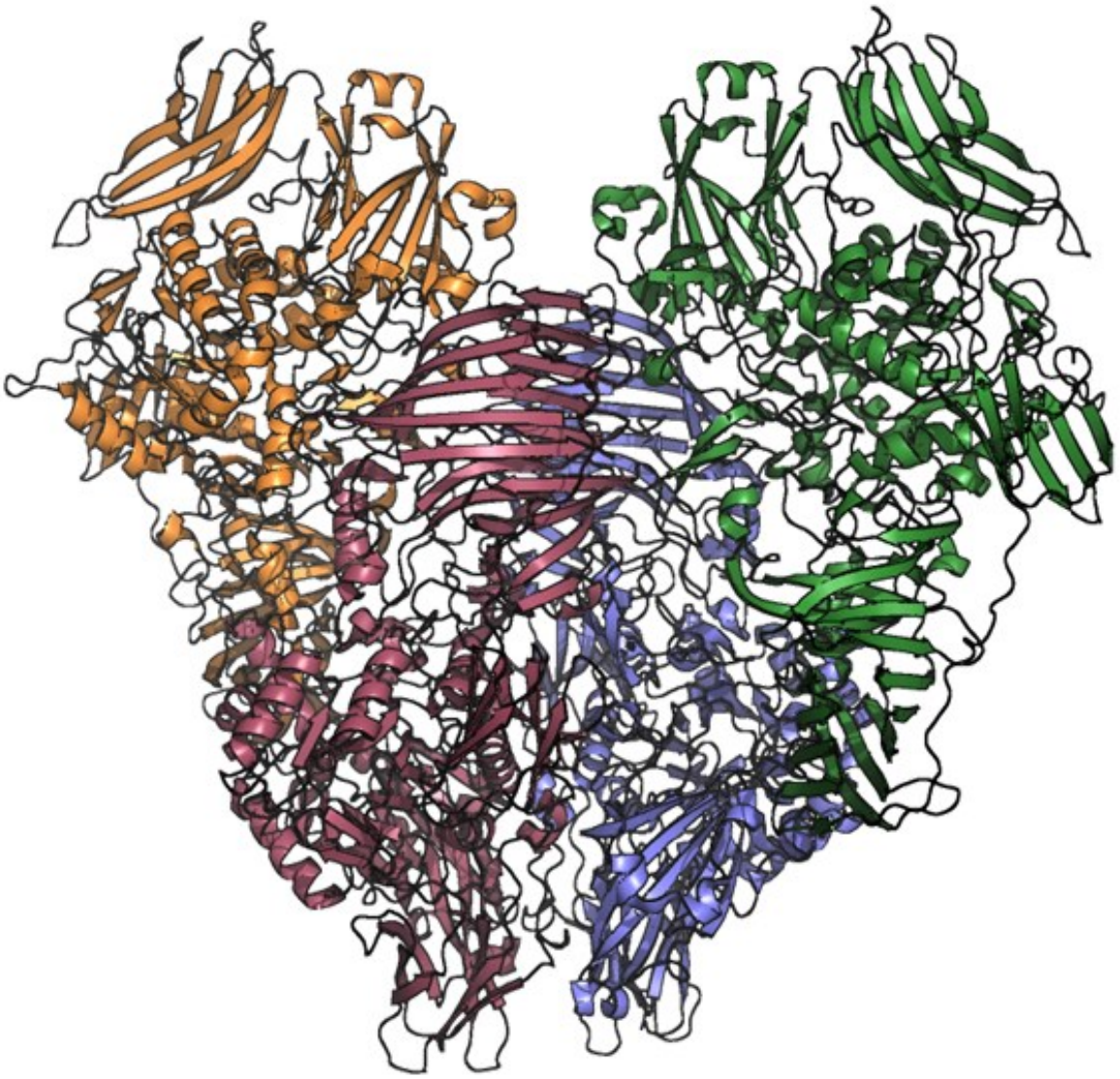


Figure 2-22: The quaternary structure of Lac4 displaying a dimer of dimers. PDB: 3OB8

Post-Translational Modification in *K. lactis*

Post-translational modification of enzymes expressed in yeast systems is well-documented and is the subject of many studies.[290] In *K. lactis*, there are a number of genes known and cells that lack these enzymes exhibit defects in protein glycosylation, altering the morphology of the cell.[291,292] It is also known that changes in glycosylation patterns in proteins can affect the stability and activity of expressed enzymes, which can present problems in the biotechnological applications thereof.[290] Like other fungi and yeasts, *N*- and *O*-glycosylation of enzymes expressed in *K. lactis* typically produce glycans with a high mannose content.[290] Although the majority of enzymes produced by higher-order eukaryotes require some form of post-translational glycosylation for structural integrity and activity, this is not always necessary as some eukaryotic enzymes, particularly those originating

from yeasts, are known to be functional when expressed in prokaryotic systems, which do not contain the machinery required for post-translational glycosylation of enzymes.[293]

Catalytic Centre

Lac4, being a GH2 galactosidase, has a simple retaining mechanism of hydrolysis and thus has two catalytic carboxylic acid-containing residues which act as a nucleophile and a catalytic acid/base, Glu551 and Glu482 respectively. These were identified *via* sequence analysis with other GH2 galactosidases, LacZ from *E. coli*[294] and the β-galactosidase from *Arthrobacter* sp.[283] The catalytic machinery of Lac4, like in LacZ, is located at the *C*-terminal of the β-strands of the D3 TIM barrel and comprises also residues from D1 (Asn88, Val89, and Asp187) and D5 (Ala1000 and Cys1001). Crystallographic studies revealed three metal-binding centres in and near the active site; one for a divalent Mg or Mn cation, and two for monovalent Na cations (Figure 2-23).[283]

The Mg^{2+} site is located in a position which would suggest that it plays a role in activation of the catalytic acid/base and is coordinated by His416, Glu414, and Glu482. The crystal structure does not reveal any coordinated water, but based on the coordination of these three residues around the cation, it can be assumed that the remaining 3 coordination sites of Mg^{2+} are occupied by water molecules in an octahedral fashion. The first Na^{+} site is coordinated by Asp187, the backbone carbonyl of Phe620, Asp623, and the 6-hydroxyl of the substrate. Again as with the Mg^{2+} site and due to the similarities with the LacZ active site, it can also be presumed that Na^{+} is coordinated trigonal bipyramidally with the fifth coordination site being occupied by a water molecule. The second Na^{+} site is not found in LacZ and shows similarity with another GH2 galactosidase from *Arthrobacter* sp. and suggests that it plays a structural role in maintaining the shape of the catalytic pocket.[283] This structural Na^{+} cation is in a pseudo-square pyramidal geometry and coordinated with the backbone carbonyl and the sidechain of Thr87, the backbone carbonyl of Asn88, Asp185, and the backbone carbonyl of neighbouring Gln186. These same studies revealed an additional Mn^{2+} and two Na^{+} binding sites, which are suspected to play structural roles within the enzyme by shaping and stabilising the dimer-dimer interfaces, S2 and S3.

The entrance to the catalytic pocket is quite narrow and is approximately 20 Å deep, though upon dimerisation, this pocket is buried even deeper as the cavities of the adjacent monomeric units face each other within the interface. This face-to-face arrangement of the monomers combined with the D3 loop (residues 420-443) results in a narrow opening of 10 Å to the pocket.[283]

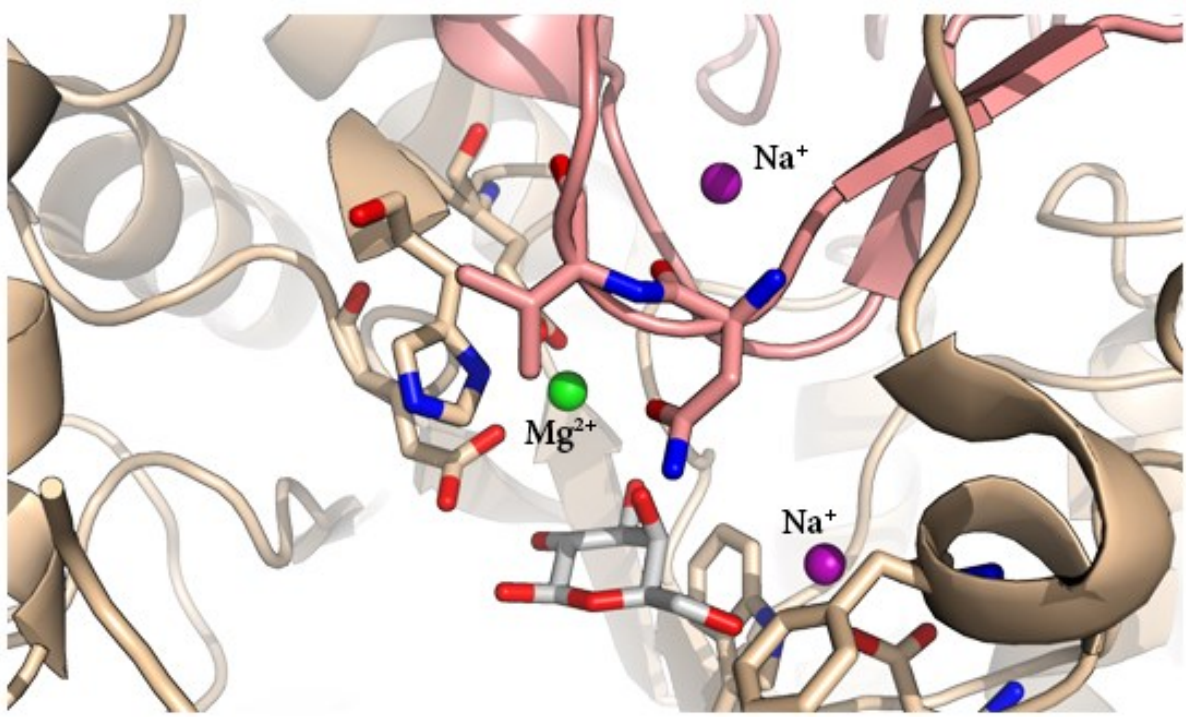


Figure 2-23: The three catalytically-important metal centres. Mg^{2+} and the lower right Na^{+} cations position the substrate and mediate the active site chemistry as in LacZ. The upper Na^{+} cation plays a structural role. PDB: 3OB8

Active Site Specificity

As the natural environment for the yeast *K. lactis* is dairy products, it is unsurprising that Lac4 has evolved to preferentially bind and hydrolyse the *exo*-β-glycosyl bond of lactose and separates the galactosyl moiety from the glucoside. Like a number of galactohydrolases, Lac4 also exhibits transgalactosidic activity and will form β-(1,6) bonds, although β-(1,4) and β-(1,3) linkages are known.[222,223] In addition, it has shown that this enzyme prefers di- and tri-saccharide formation with higher order oligosaccharides representing a minority of the transgalactosylation products.[222]

As the number of crystal structures for this enzyme is limited, it is hard to elucidate conclusively the method of substrate binding and interactions within the active site. However, due to the homology of Lac4 with LacZ, a number of properties can be inferred by overlaying and comparing the crystal structures. The crystal structure reveals that the galactose moiety is positioned above Trp582 during hydrolysis and is presumably held in place by C-H···π interactions with the partially-charged hydrogens of the galactosyl ring. Also, it appears that a number of the important stabilising interactions found in LacZ are present in the Lac4 crystal structure. Despite the absence of water molecules in the active site of the substrate-bound model, the high conservation observed in the galactosyl binding portion of the pocket – including the conservation of metal-binding motifs – would suggest a similar network of hydrogen bonds, including water bridges, would be present within the catalytic centre (Figure 2-24). Manual docking studies of lactose within the active pocket

suggest that there are a number of vital hydrogen bonding interactions between the glucoside and the D3 loop, specifically with Glu431, Lys436, and Tyr440, and these would contribute to lactose specificity.[283]

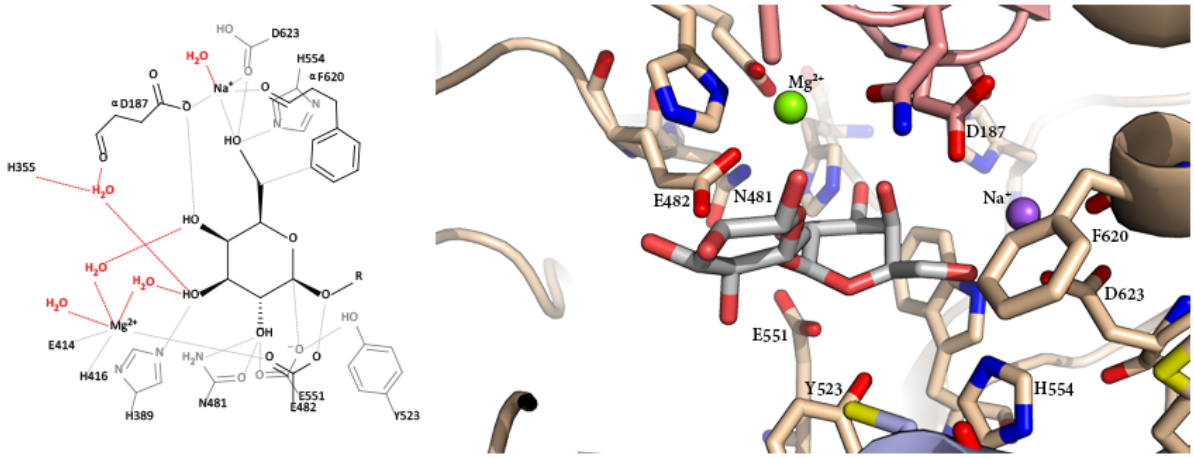


Figure 2-24: Schematic of the bonding network of galactopyranoside in the active site of Lac4. Allolactose is 'docked' in the active site by overlaying and aligning the crystal structure of Lac4 (PDB: 4OB8) and a LacZ structure (PDB: 4DUW).

2.5.3 LacA: a β-galactosidase from *Aspergillus oryzae*

Unlike LacZ and Lac4, the β-galactosidase of the fungus, *Aspergillus oryzae*, does not belong to GH2, but rather GH35.[295] Whereas GH2 typically consists of hydrolases from microorganisms, GH35 enzymes are found more commonly in higher-order, eukaryotic organisms – particularly plants.[284] LacA, the β-galactosidase from *A. oryzae*, like Lac4, is biotechnologically important in the dairy industry and is used for the treatment of milk.[295] Unlike Lac4, which prefers neutral pH and moderate temperatures, LacA prefers an acidic environment and has maximal activity in the pH range of 4.5-4.8, and temperatures of 45-55 °C depending on substrate,[222] which, among other reasons, would allow the enzyme to survive some pasteurisation processes.[296]

Structure of LacA

LacA is 1005 amino acids-long with a molecular mass of 110kDa.[295] The enzyme exists as an active monomer, in contrast to the aforementioned LacZ and Lac4 enzymes, which both require the oligomerisation for activity. The tertiary structure of LacA displays 6 distinct domains (D1-6) with domains 2-6 forming a crescent around D1 (Figure 2-25).[295] Domain 1 (residues 40-397) is a TIM barrel and contains the catalytic residues and active site. However, the TIM barrel of D1, is atypical in that the fifth helix is absent – a feature also observed in D3 of LacZ.[297] In similar crystal structures, domains 2 and 3 (residues 398-475 and 476-577 respectively) are considered as two parts of a whole domain.[297] Combined, D2/3 consists of 16 antiparallel β-sheets with a C-terminus α-helix. The reason for the splitting of the domain is presumably due to the resemblance of D3 to an I-type

immunoglobulin fold.[297] Domain 4 (residues 572-664) is dominated by a Greek key β-sandwich consisting of 8 antiparallel β-strands, which lead into a short, 18-residue strand, dubbed D5a. Domain 5, as alluded to, is made up of two parts, 5a (residues 665-683) and 5b (residues 859-1005) and these two parts come together to form a jelly roll-like structure formed by a 5-strand and a 3-strand β-sheet. The discontiguous nature of D5 results in one of the β-strands of the 5-strand sheet being formed by D5a. Domain 6 (residues 684-858) sits at the apex of the crescent and is a right-handed, 8-strand jelly roll β-sandwich. Again, this sandwich consists of a 3-strand and a 5-strand sheet and contains a large loop between β-strands 6 and 7, which juts over D1 and assists in formation of the active site.[295,297] The enzyme also contains two disulfide bridges, C205-206 and C266-C315, which contribute to the overall stability and structure of LacA.

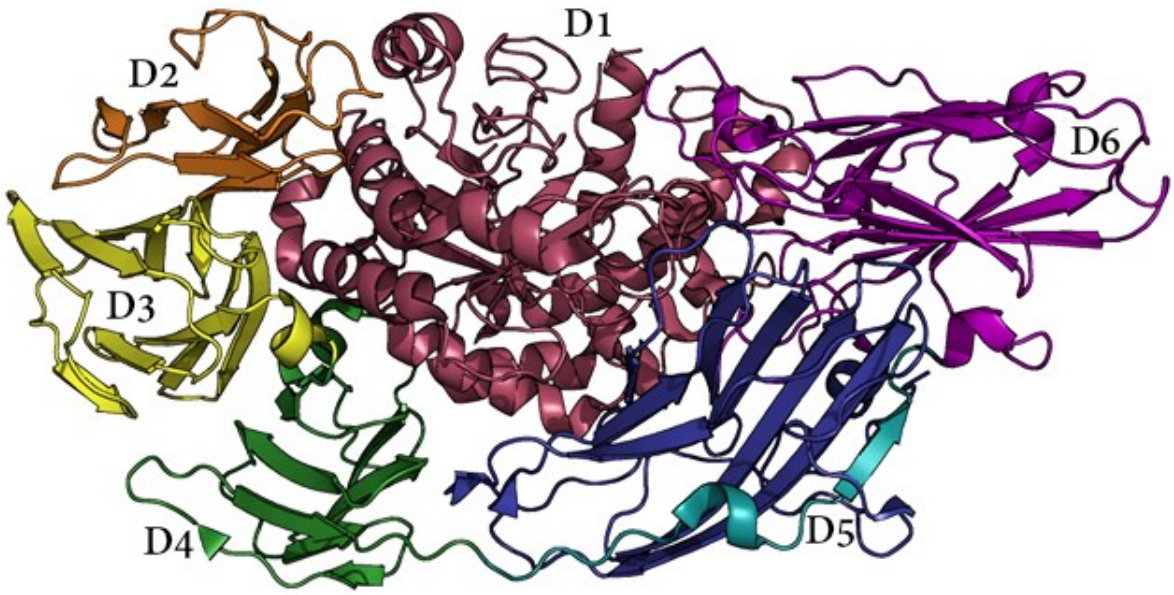


Figure 2-25: The tertiary structure of LacA shows the catalytic domain, D1 (dark red), flanked in a crescent form by the remaining domains. PDB: 4IUG

N-Glycosylation

Post-translational modification is well-known for many eukaryotic organisms and is important for folding, structure, and stability.[290] Filamentous fungi, such as *A. oryzae*, typically express enzymes, which have undergone post-translational modification, typically being small, high-mannose *N*-glycans and *O*-mannosylation.[298,299] The crystallised LacA from *A. oryzae* was heterologously expressed in the yeast, *Kluyveromyces lactis*, produced by *Maiksimainen et al.* would therefore not exhibit an identical glycosylation pattern of the native, homologously-expressed enzyme.[295]

The sequence of *lacA* indicates 11 potential sites for *N*-glycosylation, one of which is excluded by crystal structure as it is located within the protein. The heterologously-expressed LacA exhibited six

glycans, three of which exhibit typically-fungal, high-mannose structures (Figure 2-26). These glycan moieties are presumably related to the stability of the protein, as a number of these chains interact with the protein *via* hydrogen bonding to residues. These chains reach across the surface of the enzyme and adhere to the interfaces between the various domains. This could also protect the enzyme from proteolysis as aromatic and hydrophobic residues preferred by pepsin are hidden underneath two of these large, high-mannose chains.[295]

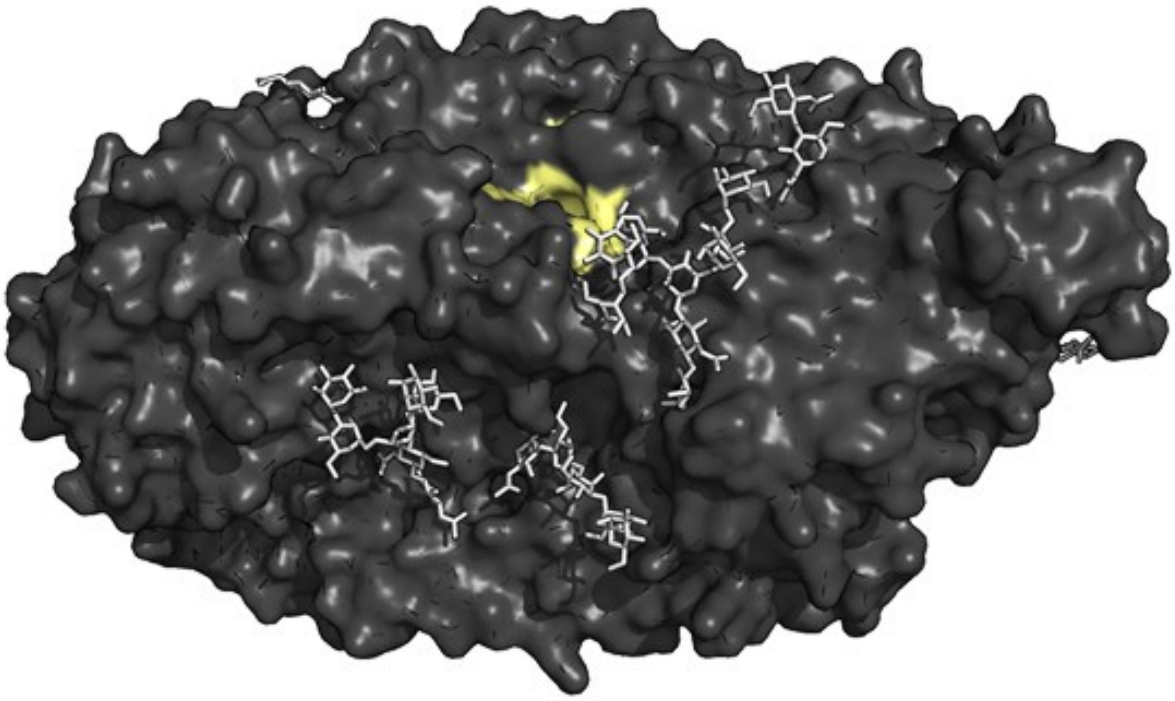


Figure 2-26: The native LacA undergoes extensive post-translational *N*-glycosylation (white). The active site is highlighted in yellow. PDB: 4IUG

Catalytic Centre

Mechanistically, LacA follows the same retentive activity as both LacZ and Lac4. Thus, the residues Glu200 and Glu298 have been identified as the acid/base catalyst and the nucleophile respectively.[295] Despite its biotechnological importance, the structural features of LacA have not been as extensively investigated as other β-galactosidases. However, an alignment of the active site of LacA with that of Lac4 reveals a number of putative interactions between the enzyme and the substrate (Figure 2-27). Furthermore, although no water molecules were directly observed to be bound in the active site along with galactose. The similarity of pocket structure between the two indicates the presence of water-bridges and is expected to be of importance in the stabilisation of reactive intermediates. The galactose molecule stacks in the active site above the aromatic system of Tyr342 *via* the same C-H···π interactions observed in other galactosidases. C-6 is oriented by hydrophobic interactions between itself and Phe304 whilst the 6-hydroxyl interacts with Glu142 and Tyr364. Glu142 is also within

hydrogen bond distance of the 4-hydroxyl, which combined with the 4-OH···Asn140 interaction, presumably contributes to the specificity of LacA for galactopyranose. The 3-hydroxyl group hydrogen bonds with Tyr96 and forms a weak, hydrostatic interaction with the backbone carbonyl of Ile139. Asn199, Glu200, and Glu298 bind the 2-hydroxyl group and it can be inferred that these interactions are important for the stabilisation of the oxocarbenium-like transition state as in other β-galactosidases. Glu200 and Glu298 interact with the anomeric hydroxyl group and C1 respectively, as expected. Furthermore, these residues appear to be stabilised by Asp258 and, like in LacZ and Lac4, a tyrosine residue, Tyr260, is observed under the oxygen of the pyranose ring and in proximity to the nucleophilic residue which suggests that it may be involved in transition state stabilisation and/or as a proton shuttle.

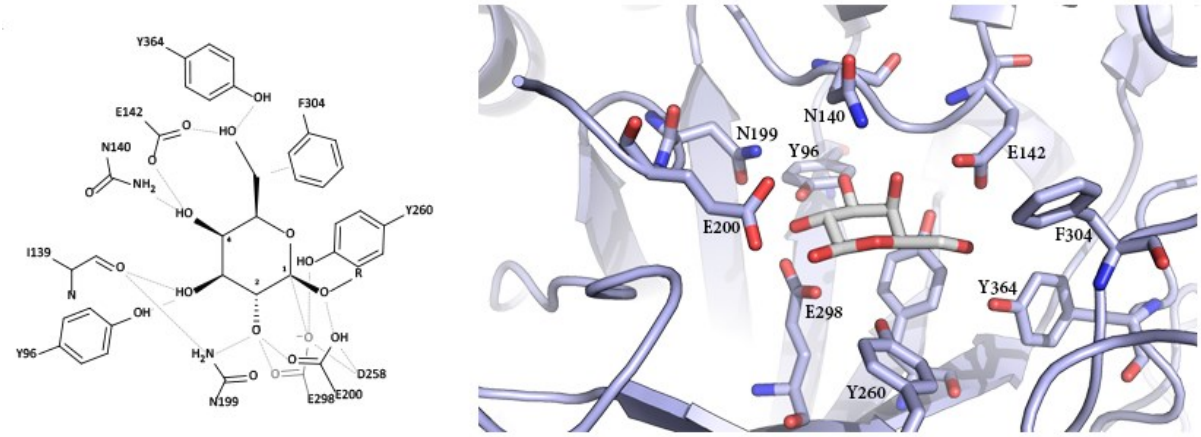


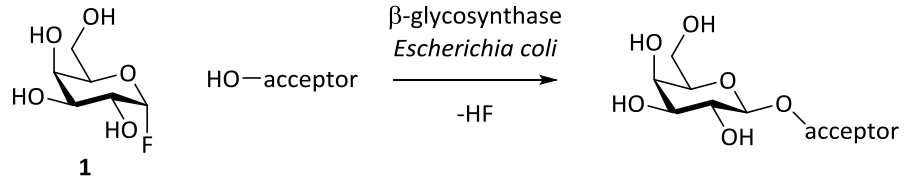
Figure 2-27: Binding of the substrate in the active site of LacA. PDB: 4IUG

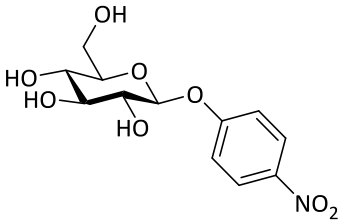
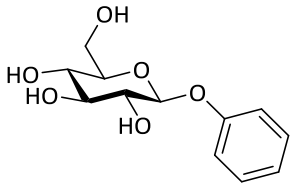
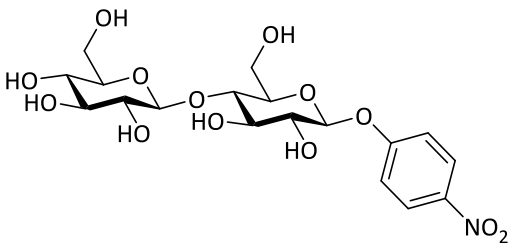
Chapter 3 - Results and Discussion

3.1 LacZ from *Escherichia coli* as a Model Glycosynthase

The LacZ enzyme from *E. coli* was chosen as a model glycosynthase for two main reasons. Firstly, this enzyme is one of the best-studied enzymes to date and as such, there is a wealth of information available on its activity, structure, and mechanism. Secondly, this enzyme has also been previously developed and used as a galactosynthase by *Jakeman et al.* in 2002.[1] In the aforementioned study, two mutants of LacZ were used: E537S, a simple substitution of the nucleophilic residue; and the double mutant, E537S/G794D, which was chosen based on an increased hydrolytic activity which had been previously reported for this enzyme.[251] Using α-D-galactopyranosyl fluoride (**1**, galF) as a glycosyl donor, they were able to produce three, small oligosaccharides using both synthase variants. In each case, a glucoside was used as the acceptor molecule with varying yields of 40 – 85% (Table 3-1).[1] While these yields vary wildly from moderate to good, the enzymes were selective for the β-(1,6) product and required no protecting group chemistry on the acceptor molecule which is a requirement in classical synthesis of oligosaccharides.

Table 3-1: General scheme for the reaction of the LacZ galactosynthase and the results of the initial glycosylations performed by *Jakeman et al.* using LacZ mutants demonstrating the catalytic ability of these glycosynthases.[1]



Acceptor	% Yield$_{E537S}$	% Yield$_{E537S/G794D}$
	40	70
	61	85
	51	80

Although two LacZ-derived galactosynthases have previously been generated and their potential demonstrated, there exists still the question of substrate range. Studies into the hydrolytic activity of the native enzyme would suggest that indeed, LacZ is capable of hydrolysing both D-fucopyranosides and L-arabinopyranosides.[226,277,300] Furthermore, based on hydrolytic studies of the wild type enzyme, the β-galactosidase from *E. coli* is able to hydrolyse more than its natural substrates, lactose (**2**) and allolactose (**3**), 4-*O*-(β-D-galactopyranosyl)- and 6-*O*-(β-D-galactopyranosyl)-D-glucopyranose respectively.[226,270,277] Indeed, LacZ has been shown to hydrolyse a number of other substrates[226,272,277] and of particular use are those which induce a detectable colour change which, most famously, led to blue-white screening – one of the most ubiquitously used techniques in molecular biology.[301] This, combined with its transglycosylation activity with non-glycosidic glycosyl acceptors,[302] makes it an excellent candidate for further study as a glycosynthase.

3.1.1 Isolation of *lacZ* from *E. coli*

After identification of LacZ as a candidate glycosynthase, isolation of the *lacZ* gene from *E. coli* was performed. Initially, it was considered to isolate this gene from the *E. coli* strain, DH5α, using PCR techniques, but as DH5α contains the truncated *lacZ*ΔM15 variant of the gene,[303] gene isolation was instead attempted using the BL21(DE3) strain.[304] Ultimately, Q5 DNA polymerase demonstrated non-specific amplification at an annealing temperature of 50 °C from whole-cell BL21(DE3) templates as evidenced by streaking of the PCR product by agarose gel electrophoresis. Accordingly, the annealing temperature of the PCR was increased to 55 and 60 °C and gene amplification from whole-cell templates was performed using Q5 polymerase, both of which yielded clear bands by agarose gel electrophoresis, which corresponded to the ca. 3 kbp length of *lacZ*.

Following DNA extraction, the sample was prepared for insertion into the vector *via* digestion with the restriction enzymes, XbaI and XhoI. Subsequent ligation into a similarly digested, linearised pET28a(+) plasmid gave a mixture of gene-containing plasmids and re-ligated vector. The resulting DNA was used to transform competent *E. coli* MACH1 cells, which were then plated onto kanamycin-selective agar and colonies were selected. These colonies were used as templates for a subsequent round of PCR using *Taq* polymerase to check the insertion of gene-containing plasmids (Figure 3-1, left). The positive colonies, 1, 4, and 6, were grown, thereby amplifying the plasmid, and the DNA extracted and sequenced, which confirmed the successful isolation of *lacZ* from *E. coli* BL21(DE3) in the form pET28a::*lacZ* (Figure 3-1, right).

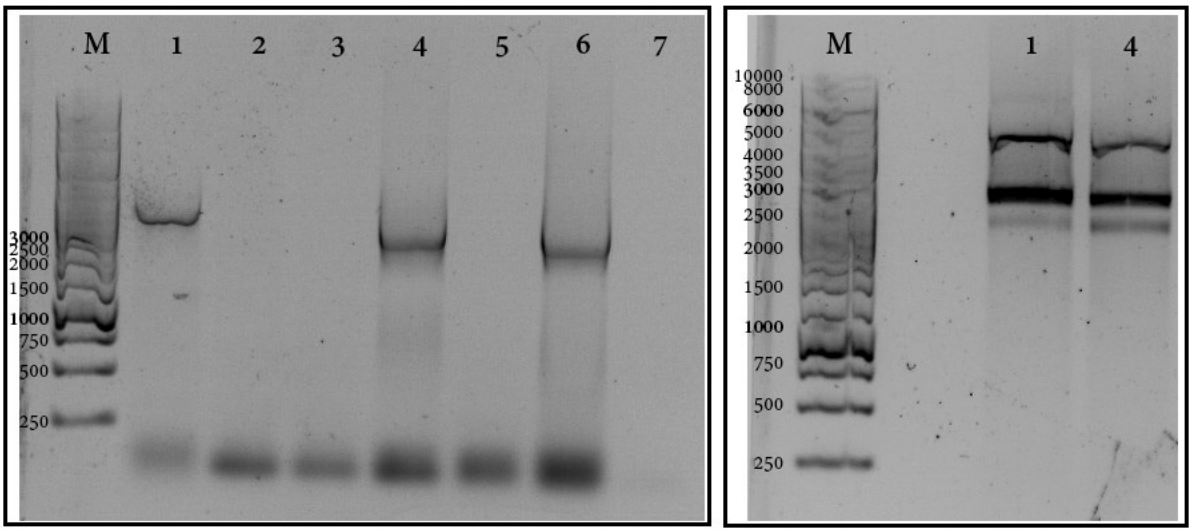


Figure 3-1: Left: Colony PCR of 7 colonies of pET28a::*lacZ*-transformed *E. coli* MACH1 using *RedTaq*®. Right: Digestion analysis of the isolated DNA of colonies 1 and 4 using *XbaI* and *EcoRI*. NOTE: All images of agarose gel electrophoresis have been edited for clarity using Adobe® Photoshop® CS6 Extended and the originals are available on request.

3.1.2 Generation of *lacZ* Mutants

Following the work of *Jakeman* and *Withers*,[1] it was decided to replicate their two *lacZ* mutants, E537S and E537S/G794D in order to establish our assays with literature-proven galactosynthases. Additionally, with these two enzymes in hand, we could extend upon the work by *Jakeman* and *Withers* by further investigating their range of glycosyl donors and acceptors. The mutants were generated using a site-directed mutagenesis technique based on *QuikChange*®.[305,306] Following the successful generation of *lacZ* E537S, this mutant was used as a template for the generation of the double mutant, *lacZ* E537S/G974D. With exception of the mutagenic primers, this mutagenesis was performed according to the optimal conditions identified when creating the single mutant, *lacZ* E537S. Sequencing of this step confirmed the presence of the double mutant. The sequencing results also confirmed something unexpected; the ribosomal binding site (RBS) on the plasmid was absent in the pET28a::*lacZ* variants. As a result, it was identified that the *XbaI* site within the multiple cloning site (MCS) of pET28a(+) vector cuts upstream of the RBS and therefore during subcloning, this important segment is removed from the final gene construct (Figure 3-2). Re-analysis of the MCS of the pET series vectors identified the *NdeI* restriction site, 5'-CA$^{\vee}$TATG-3' being downstream from the RBS. Furthermore, as this restriction site contains an ATG motif, there would be no excessive elongation of the gene after cutting; only two base pairs. However, the 5'-sequence of pET28a::*lacZ*, 5'-AGAATGACC…, meant that this enzyme was not compatible for cutting.

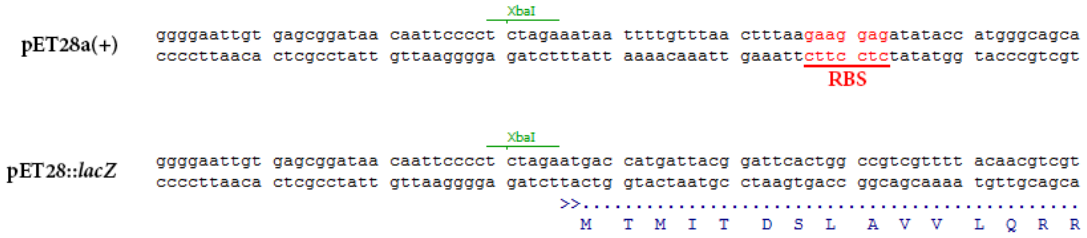


Figure 3-2: Comparison of the sequences of the empty pET28a(+) vector and the initial pET28a::*lacZ* construct.

To remedy this, a restriction enzyme which had both compatible ends with the *NdeI* site and the 5'-end of *lacZ* was required. Based on compatible sticky ends methodology, a suitable, readily-available enzyme, *AseI*, was identified. As this sequence did not match the 5'-end of the pET28a::*lacZ* constructs, SDM was required to insert the 5'-ATT nucleotides in the -3 position. After unsuccessful attempts to achieve this using qcSDM and re-extraction of the gene from *E. coli* BL21(DE3) using primers with appropriate codons, the correct mutation was eventually introduced using round-the-horn (RtH) SDM. After successful determination of the optimal conditions using *pET28a::lacZ* WT as a model, these conditions were applied successfully to all three variants (Figure 3-3).

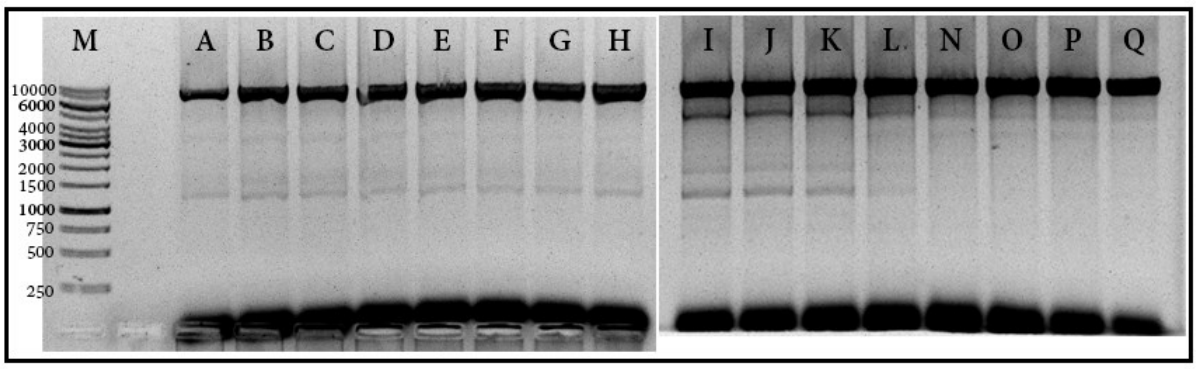


Figure 3-3: Screen of DNA polymerases for the RtH SDM of the pET28a::*lacZ* variants. M - Marker (GeneRuler™ 1 kb DNA Ladder); A-H – Phusion DNA polymerase; I-L/N-Q – KAPA HiFi DNA polymerase. Temperatures [°C]: 49.5 (A, I); 49.8 (B, J); 51.7 (C, K); 54.7 (D, L); 59.4 (E, N); 62.7 (F, O); 64.4 (G, P); 64.5 (H, Q).

After using the PCR products to transform competent *E. coli* MACH1 cells, triplicate colony picking, then DNA amplification and isolation, a restriction analysis was performed. The restriction analysis of the amplified DNA was performed using three restriction enzymes, *AseI*, *XhoI*, and *BspHI*, as the fragments produced by these would result in much more defined bands as they would be more separate from one another. The expected fragment sizes of 3075, 2079, 1285, 708, and 679 were visible in all but one sample (Figure 3-4). These mutants, named pET28a*::*lacZ*, were then digested on a larger scale for the subcloning of the *lacZ* genes into pET22b(+) vectors. Using the same restriction enzymes as for the analysis, the 9 samples were again digested and the 3 kbp band in the eight correctly-fragmented samples isolated. The isolated genes were then ligated into a

NdeI/*XhoI*-linearised pET22b(+) vector using T4 DNA ligase. Three colonies were picked from these plates and a digestion analysis using *EcoRV* showed that in all but two cases, the expected fragments, 6036 and 2404, were present (Figure 3-5). This was then confirmed by sequencing of the plasmid. With the three pET22b::*lacZ* variants (Table 3-2) now in hand, an adequate expression system was desired.

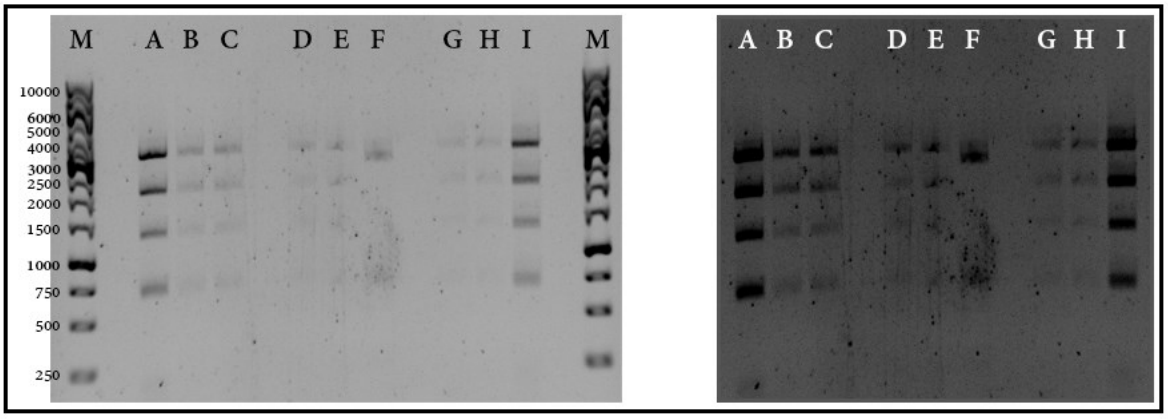


Figure 3-4: Restriction analysis of three colonies of each mutation of *lacZ* after RtH SDM using the three enzymes, *BspHI*, *AseI*, and *XhoI*. The identical, darker image (right) is shown to highlight difficult-to-see fragmentations. M – Marker (GeneRuler™ 1 kb DNA Ladder); A-C – *lacZ* WT; D-F – *lacZ* E537S; G-I – *lacZ* E537S/G794D.

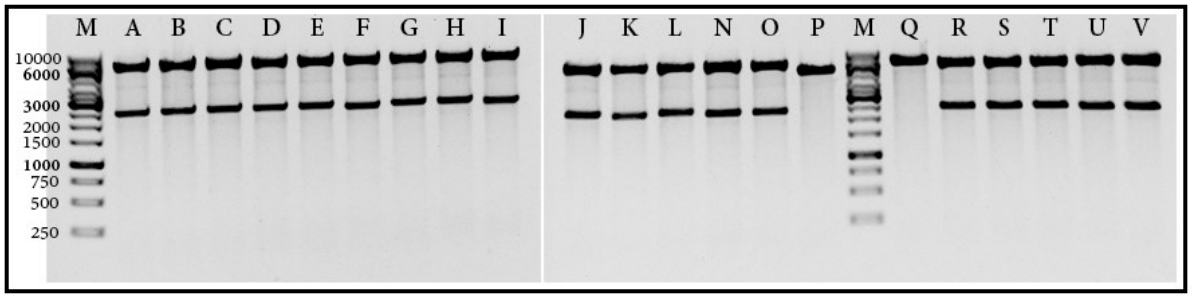


Figure 3-5: Restriction analysis of the pET28a*::*lacZ* variant colonies using *EcoRV*. M – Marker (GeneRuler™ 1 kb DNA Ladder); A-C – *lacZ* WT (template 1); D-F – *lacZ* WT (template 2); G-I – *lacZ* E537S (template 1); J-L – *lacZ* E537S (template 2); N-P – *lacZ* E537S/G794D (template 1); Q-S – *lacZ* E537S/G794D (template 2); T-V – *lacZ* E537S/G794D (template 3).

Table 3-2: Summary of *lacZ* constructs.

Gene	Plasmid	Method	Construct
lacZ	pET28a(+)	Isolation from *E. coli* BL21(DE3)	pET28a::*lacZ* WT
	pET28a(+)	RtH SDM from pET28a::*lacZ* WT	pET28a*::*lacZ* WT
	pET22b(+)	Subcloning from pET28a*::*lacZ* WT	pET22::*lacZ* WT
lacZ E537S	pET28a(+)	qcSDM from pET28a::*lacZ* WT	pET28a::*lacZ* E537S
	pET28a(+)	RtH SDM from pET28a::*lacZ* E537S	pET28a*::*lacZ* E537S
	pET22b(+)	Subcloning from pET28a*::*lacZ* E537S	pET22::*lacZ* E537S
lacZ E537S/G794D	pET28a(+)	qcSDM from pET28a::*lacZ* E537S	pET28a::*lacZ* E537S/G794D
	pET28a(+)	RtH SDM from pET28a::*lacZ* E537S/G794D	pET28a*::*lacZ* E537S/G794D
	pET22b(+)	Subcloning from pET28a*::*lacZ* E537S/G794D	pET22::*lacZ* E537SG794D

3.1.3 Host Selection for the Expression of *lacZ*

As the *lacZ* gene was isolated from *E. coli*, it necessarily requires the bare minimum in terms of expression machinery. With no post-translational modification required and *E. coli* as a well-established expression system, it was decided to express *lacZ* in an *E. coli* system. It is known that the glycosynthase enzymes, although catalytically-inactive, retain the structure of the native enzyme.[250,307] Combined with the well-established methods for measuring the hydrolytic activity of enzymes, *lacZ* in particular, it was decided to use the wild type enzyme for expression tests and optimisation thereof. Initially, homologous expression in BL21(DE3) was planned. However, use of this strain resulted in the inability to determine hydrolytically inactive enzymes without committing to purification of the enzyme from the crude mixture. Therefore, the expression of *lacZ* gene in another *E. coli* strain, XL1-Blue, was attempted. Analysis of the genotype of this strain showed that it possesses the non-functional *lacZ*ΔM15 gene. For the purposes of a blue-white screen, this would mean that only when the wild type *lacZ* of the plasmid is expressed, would a colony appear blue. However, as the plasmid is of the pET family, it requires a T7 RNA polymerase to correctly transcribe genes located within its multiple cloning site. As XL1-Blue is not a strain that contains the T7 phage RNA polymerase, it was attempted to transform our cells with a second plasmid which contains the T7 gene. To this end, two T7-containing vectors were obtained, pML5::T7 and pBBR1MCS3::T7.[308] These vectors were transformed into competent XL1-Blue cells, before being made competent a second time and subsequently transformed a second time with pET28a::*lacZ*. Unfortunately, after blue-white screening assays, this dual-vector approach proved to be

unsuccessful in the production of *lacZ* in XL1-Blue, presumably due to problems in the expression of T7 polymerase. As a result, a new strain was sought and eventually ER2566 was identified. This strain was particularly interesting as it had a genotype which included *lacZ*::T7. This means that the expression of T7 was under the control of the *lac* operon. As such, it could be expected that induction of the system by IPTG would only result in the production of a β-galactosidase when there is a β-galactosidase-containing plasmid in the cell; a hypothesis which was confirmed by exposure of a sample of crude lysate to a solution of X-gal (Figure 3-6).

Figure 3-6: X-Gal assay of the crude lysate of expression cultures of *E. coli* ER2566. Left – LacZ WT; right – pET22b(+) empty vector.

3.1.4 Optimisation of *lacZ* Expression in *E. coli* ER2566

Many strains of *E. coli* are frequently used for expression of β-galactosidases and the conditions for expression of LacZ and variants thereof in a selection of host strains has been previously performed. However, excluded from this list is the target strain ER2566 and optimisation of expression conditions was therefore required. The majority of protein expressions within the *Pietruszka* group are performed using variations upon standard protocol which has been successfully used to express aldolases,[309] P450 monooxygenases,[310] alcohol dehydrogenases,[311] and enoate reductases.[312] Logically therefore, this method was referenced as a baseline (Table 3-3, "Control"). In conjunction with *Magdalena Sommer* (IBOC) and *Lars Freier* (IBG-1), optimisation of LacZ expression was performed by variation of four independent variables: pH of the growth medium, temperature at which the cells are grown, the concentration of IPTG used for induction of expression, and the time at which protein expression is induced in the growing cells.

As TB was used for the growth medium, pH could be controlled by varying the pH of the buffering solution. *Escherichia coli* is a known inhabitant of the human intestine and it is known that the pH of the intestinal tract can vary between 5.7 in the caecum to 7.4 in the terminal ileum.[313] Furthermore, it is also known that the optimal pH for LacZ is between 7.5 and 8.0.[228] Therefore, it was decided to

vary the pH of the expression medium between pH 6.5 and 8.0. Similarly, despite *E. coli* and LacZ preferring 37 °C for growth and activity respectively, it is known that lower expression temperatures can increase protein quality which translates to higher specific activities, and decrease aggregation of proteins which can lead to cell death.[314] To this, a temperature range from 18 °C-37 °C was trialled. The concentration of IPTG used for induction is typically between 0.1 and 3.0 mM final concentration. The exact concentration of IPTG which is used for induction has previously been demonstrated to affect the amount of protein produced during over-expression.[315] Furthermore, the over-expression of proteins has been believed to vary depending upon the point in the growth phase the cells are in at the time of induction. Low cell densities can stress the producing organism and the resulting slower growth can yield lower enzyme titres. Conversely, high cell densities mean that the cells reach the stationary growth phase too soon after induction, thereby reducing the time the cells have to express the enzyme. Thus, optimisation of the time at which the cells were induced was explored by analysis of the cell density of the culture, as measured by OD_{600}, at the time of induction. As such, these four variables were analysed in a 96-deep-well plate-based screen and the data analysed on the absolute hydrolytic activity of the well (Figure 3-7). The expressions were monitored using both a positive control containing a genomic copy of *lacZ*, *E. coli* BL21(DE3), and a negative control which contains the *lacZ*::T7 mutation, *E. coli* ER2566. These were both transformed with the protein-producing vector, pET22b::eda_{EC},[316] so as to mimic the cellular stress induced by over-expression of a protein.

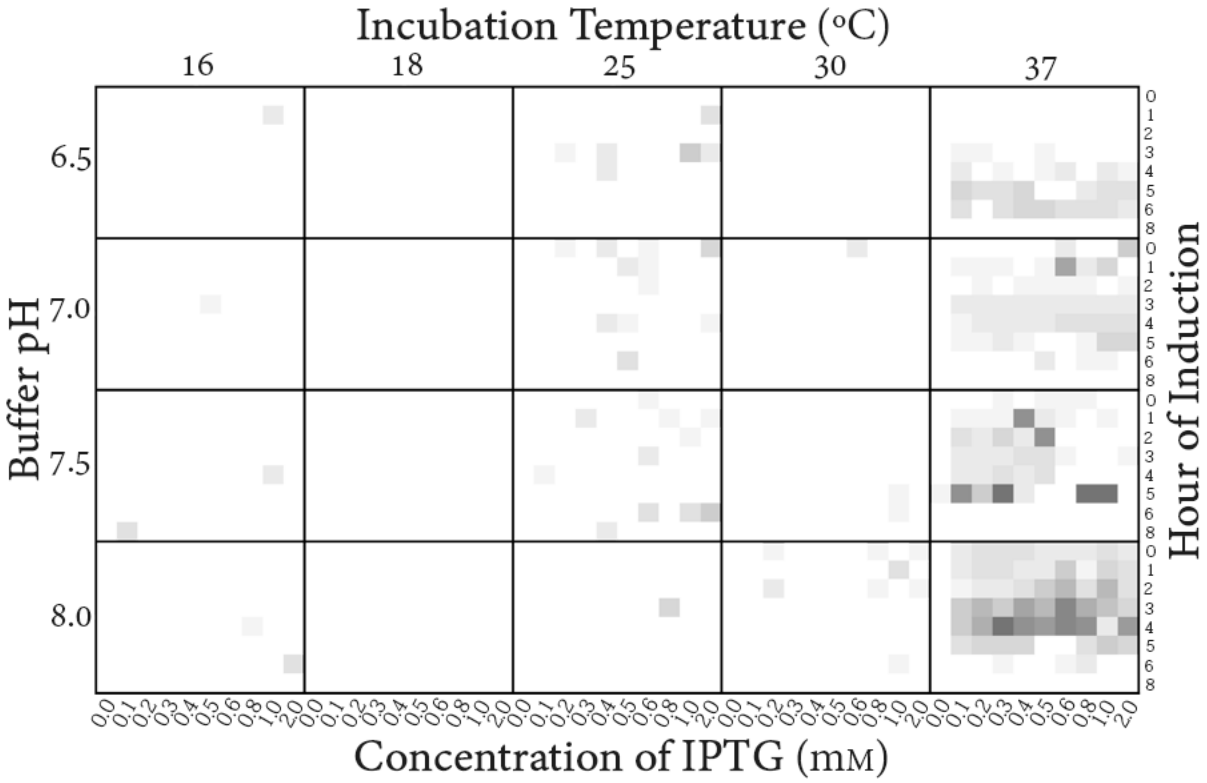


Figure 3-7: Qualitative view of the activity of LacZ expressions using a white-grey colour scale from +1.5 standard deviations to +7 standard deviations above the mean.

Trends in the observed data were extrapolated and analysed using the bioinformatics software, KriKit,[317] which revealed that the time of induction, as measured by OD_{600}, and the temperature of expression were the two factors which contributed most significantly to the amount of protein produced, whereas the pH of the medium contributed little and the concentration of IPTG had no significant effect (Figure 3-8). The optimal OD_{600} was found to be between 1.3 and 1.7, which corresponds to the end of the lag phase/beginning of the exponential phase of growth. Furthermore, the optimal temperature for expression was found to be 37 °C. This could be due to a number of reasons, but most likely was the increase in protein production at this temperature due to a higher density of cells being produced and by extension, more protein. In order to determine whether this is due to a higher level of expression per cell, or whether this is due to a higher density of cells, which produce the same amount of protein per cell, the specific activity of the lysate would need to be examined. The minimal effect of the pH of the growth medium can be explained by the natural habitat of *E. coli* in that it lives in the intestinal tract of mammals and the pH varies between different regions within the tract. Surprisingly, the concentration of IPTG had no significant effect on the amount of protein expressed.

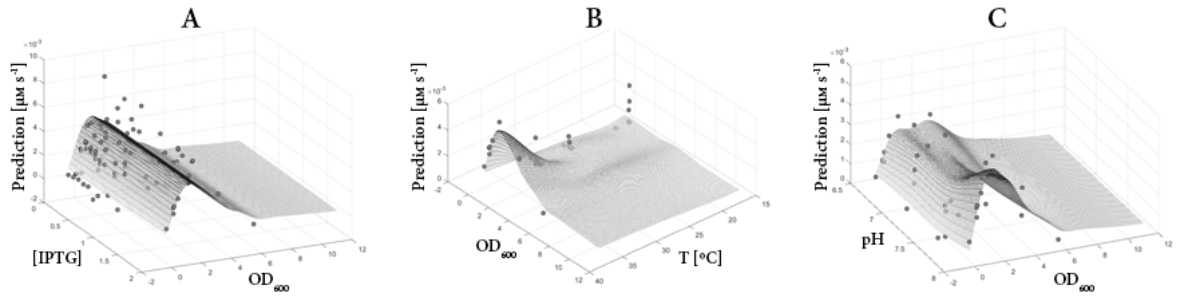


Figure 3-8: KriKit analysis of the optimisation of LacZ production in *E. coli* ER2566. The vertical axis represents the volumetric activity of the samples given. The values expressed above are shown without confidence intervals and the values given are taken when the remaining two variables are held at their maximum values. A - Concentration of IPTG against the OD_{600} at the time of induction. B – The OD_{600} at time of induction against the incubation temperature. C – The pH of the growth medium against the OD_{600} at the time of induction.

Based on the results of the 96-well plate screen, seven conditions were chosen for up-scaling to 100 mL expression cultures and these were compared to the control expression, the original protocol (Table 3-3). The analysis for this series was two-fold. Primary analysis would reveal the absolute amount of activity of the expression cultures by analysing the volumetric activity of the crude lysate and secondary analysis would be based on the specific activity of the aforementioned lysate (Figure 3-9). Those cultures which did not pass the primary analysis, would not be eligible for further consideration in the secondary analysis as the goal was to maximise enzyme yield and the secondary analysis would help to negate any false positives due to higher cell densities. All cultures were induced at an identical time-point and grown at 37 °C, with the exception of the control expression. The pH of the medium was either 7.5 or 8.0. The IPTG concentration used for induction was 0.1, 0.6, 1.0, or 1.5 mM. All cultures were harvested after 24 hours as a straight-forward, overnight procedure was sought.

Table 3-3: Conditions used for the optimisation of expression of LacZ in *E. coli* ER2566.

Culture	Temperature [°C]	Induction [h (OD_{600})]	Medium [pH]	IPTG [mM]
Control	25	8.0 (N/A)	7.5	0.1
1	37	4.75 (2.4)	7.5	0.1
2	37	4.75 (2.2)	7.5	0.6
3	37	4.75 (2.4)	7.5	1.5
4	37	4.75 (2.4)	8.0	0.1
5	37	4.75 (2.2)	8.0	0.6
6	37	4.75 (2.2)	8.0	1.0
7	37	4.75 (2.8)	8.0	1.5

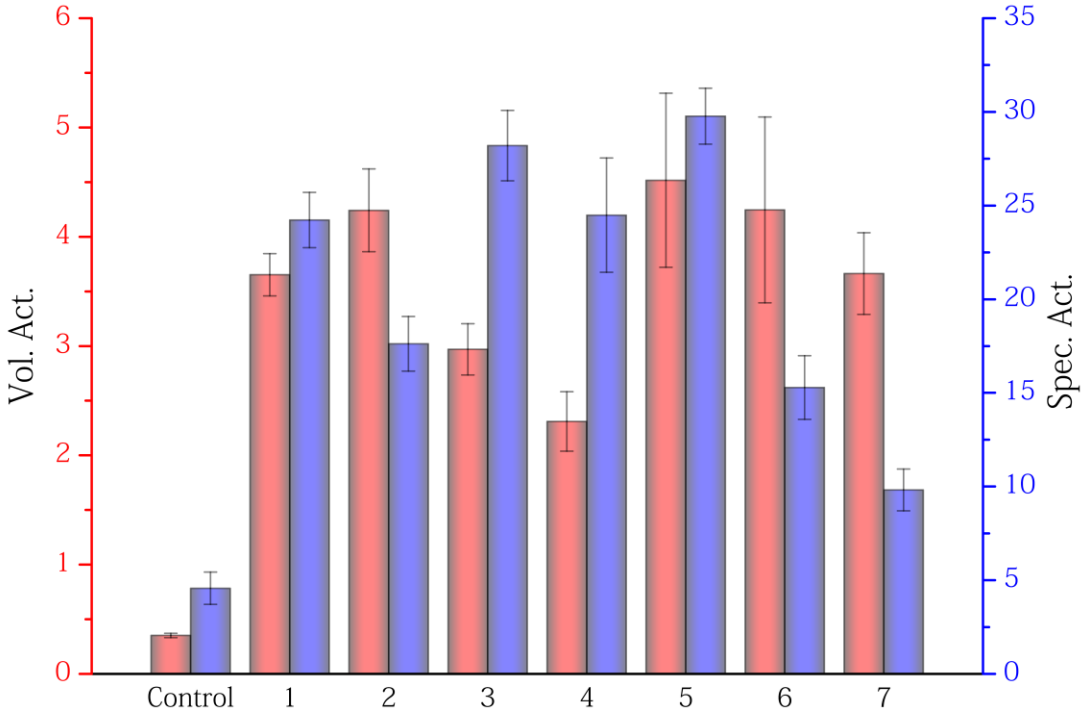


Figure 3-9: Volumetric activity (red) and specific activity (blue) of the expression of LacZ in *E. coli* ER2566 in a 100 mL scale.

All cultures produced at least 5-fold more protein than the control based on volumetric activity. However, conditions 3 and 4 demonstrated the poorest protein yields and were thus excluded from further consideration. Conditions 1, 2, 5, 6, and 7, all produced statistically identical amounts of protein. After the application of the secondary analysis, it became clear that growth condition 5 was the best for our system. In terms of volumetric activity, it produced 14-fold more protein than previous conditions and 6-fold more protein-per-cell by comparison of specific activity. Resultingly, these conditions were adopted for all subsequent expressions of LacZ in *Escherichia coli* ER2566.

3.1.5 Purification of LacZ

With optimised expression conditions for LacZ and related mutants in the *E. coli* strain ER2566, a purification strategy was required. Due to the insertion of the gene of interest into a pET vector, a terminal His_6-tag was deemed a logical choice for a purification tag, as it allows for purification of the enzyme of interest (EOI) from a crude lysate with high selectivity.[318] The method of cloning used to insert the *lacZ* gene into the pET22 vector resulted in the creation of a C-terminal His_6-tag and that was targeted for purification. Furthermore, for ease of measurement, the wild type enzyme was used as opposed to the mutants, as hydrolytic activity is easily measured using colourimetric and fluorimetric assays.

After expression, opening of the cells *via* French press proved to be quick and effective, although also non-reproducible due to wear in the needle valve. Therefore, two rounds of sonication was taken as the preferred mode of cell opening as the inevitable wear of the sonotrodes was both slower in comparison to the needle valve of the French press, and resulted in more consistent cell opening. The isolation of LacZ from the crude lysate was then performed using immobilised metal affinity chromatography (IMAC), specifically nickel (II) nitrilotriacetic acid (Ni-NTA). During the first purification round using this system, all fractions across all steps were collected and analysis of the activity of the individual fractions showed that LacZ was eluted at an imidazole concentration of 250 mM. This indicated that these enzymes are interacting strongly with the column, which in turn demonstrated that the use of affinity chromatography for the purification of LacZ is viable. The imidazole-containing buffer was then exchanged for activity assay buffer using PD10 size-exclusion column to allow for characterisation of the *C*-terminal His_6-tagged variant of wild type LacZ. Despite having a good purification method, the enzyme yields were low, at only 5 mg L^{-1}. Colourimetric and SDS-PAGE analysis of the entire process was performed and showed that there was residual activity in a number of different fractions indicating a number of points in the purification method which could be improved upon (Figure 3-10). Therefore, optimisation of the purification method was attempted using a peristaltic pump and all further purifications were performed using this method.

Fraction	Protein Concentration [mg mL^{-1}]	Activity [µM min^{-1} mg^{-1}]
Resuspended Cells	0.25 ± 0.02	1.39 ± 0.01
Disrupted Cells	10.96 ± 0.56	3.50 ± 0.04
Lysate Supernatant	10.81 ± 0.77	3.45 ± 0.02
Pellet	1.45 ± 0.04	2.13 ± 0.01
Flow-Through	6.84 ± 0.43	1.89 ± 0.01
Column Wash	0.84 ± 0.02	0.11 ± 0.06
Elution (250 mM)	0.37 ± 0.04	51.74 ± 0.19
Elution (1 M)	0.15 ± 0.01	N.D.
Size-Exclusion	0.30 ± 0.03	133.18 ± 0.61

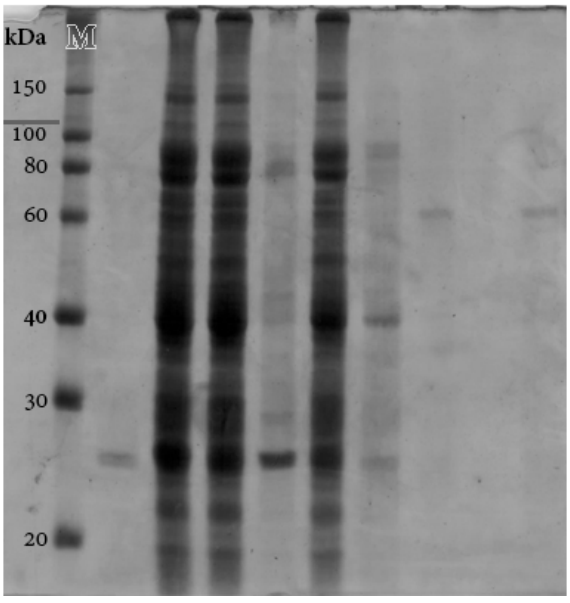


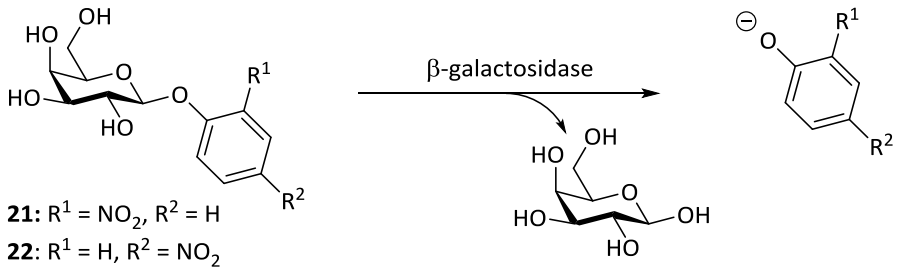
Figure 3-10: Purification table of LacZ variants expressed in *E. coli* ER2566 using the wild type as a model. The grey line at ca. 110 kDa indicates the expected mass of the His_6-tagged LacZ protein.

Looking at the SDS-PAGE, it is clearly seen that no significant amount of LacZ is secreted from the cells into the supernatant as seen by the lane, 'resuspended cells'. Furthermore, we also do not see any enzyme present in the pellet fraction, which indicates that the enzyme is soluble and does not form inclusion bodies. The hydrolytic assay results also support this observation. As the amount of LacZ produced by the *E. coli* cells is very little, as evidenced by an apparent lack of enzyme in the fraction eluted with 250 mM of imidazole and in the size-exclusion fraction. However, we know that LacZ is in these fractions as evidenced by the hydrolytic activity of these fractions. We used this information to investigate points of enzyme loss within our purification system.

As the pellet showed activity, after separation from the supernatant, the pellet was resuspended in lysis buffer and re-centrifuged. This supernatant fraction was then purified using the standard method and yielded an extra 4.4 mg L^{-1} of LacZ. The other notable source of protein loss was the flow-through – the lysate supernatant mixture, which passed through the column during the loading step but did not bind to the column. To this, the speed of loading was reduced from 4.1 mL min^{-1} to 2.3 mL min^{-1} and the loading time was extended to 45 min. As a result of this, and by performing two additional rounds of loading and purification of the lysate, an extra 1.8 mg L^{-1} of protein was obtained. Combined, this increased protein yield to around 11 mg L^{-1}. This yield could still be improved and there are still many parameters, which could contribute to this, the most obvious of which would be by altering the position of and/or the identity of the purification tag.

3.1.6 Quantification of the Activity of LacZ

The kinetic parameters of LacZ wild type were examined using a hydrolytic, colourimetric assay (Scheme 3-1) using the two common substrates, *o*-nitrophenyl and *p*-nitrophenyl β-D-galactopyranoside, gal-oNP (**21**) and gal-pNP (**22**) respectively, and in different buffers (Figure 3-11 and Figure 3-12, Table 3-4).



Scheme 3-1: Enzymatic hydrolysis of nitrophenyl galactosides produces the nitrophenolate anion which absorbs strongly at λ = 405 nm in basic conditions.

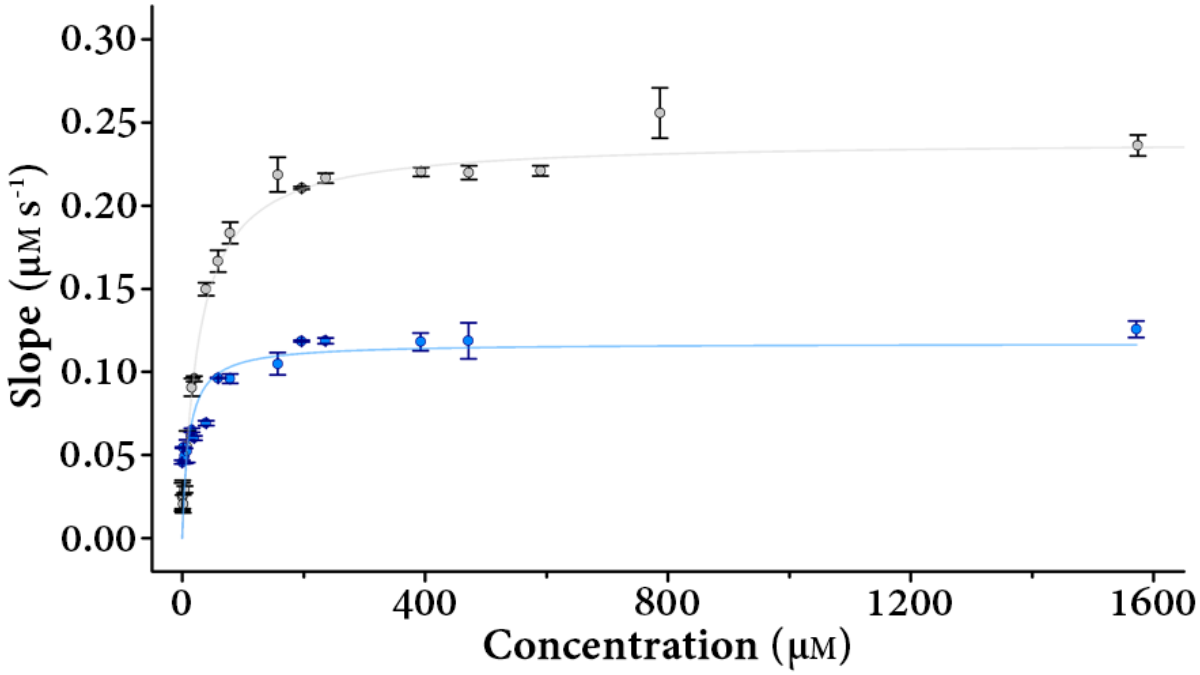


Figure 3-11: Michaelis-Menten curves for the hydrolysis of gal-oNP (**21**) by LacZ WT in 200 mM phosphate buffer (pH 7.5) (grey) and TES buffer (blue) at 25 °C.

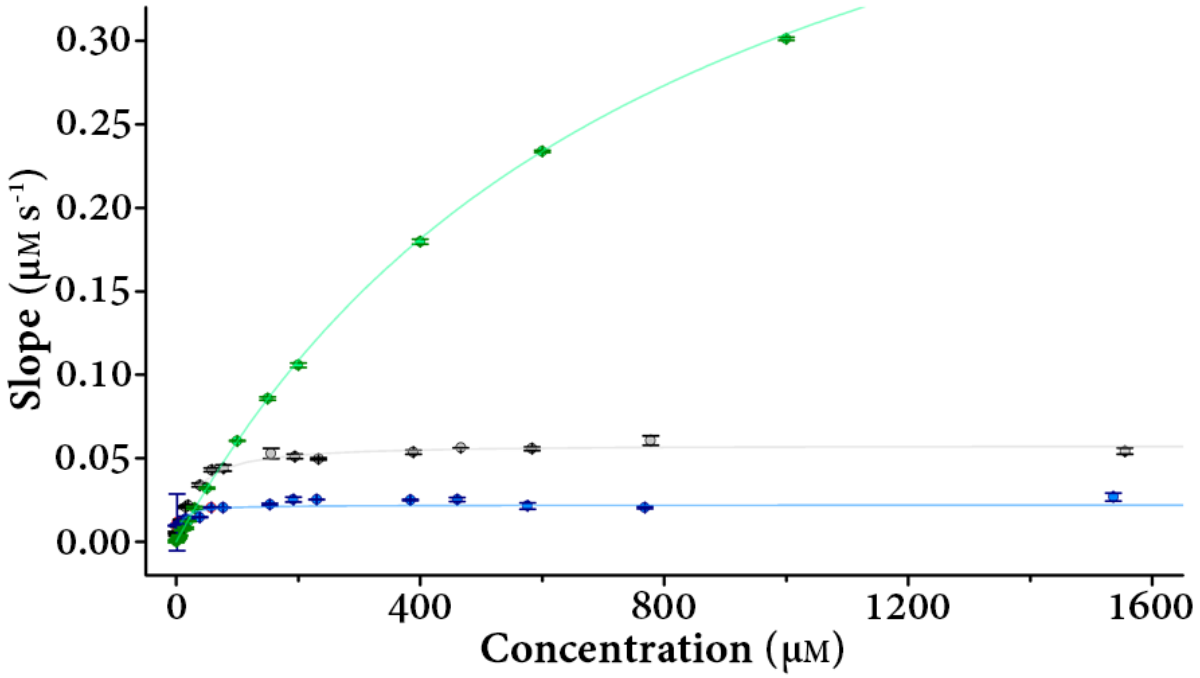


Figure 3-12: Michaelis-Menten curves for the hydrolysis of gal-pNP (**22**) by LacZ WT in 200 mM phosphate buffer (pH 7.5) (grey), TES buffer (blue), and ammonium buffer (green) at 25 °C.

Table 3-4: Kinetic parameters of LacZ in various conditions. [a] Outliers excluded from fitting analysis. phos – 200 mM mixed phosphate, pH 7.5, 1 mM $MgCl_2$; TES – 30 mM TES, 145 mM NaCl, pH 7.0; NH_4 – 50 mM NH_4HCO_3, pH 7.5 (AcOH), 1 mM $MgCl_2$.

Substrate	Buffer	K_M [mM]	k_{cat} [s^{-1}]	k_{cat}/K_M
21	phos	0.026 ± 0.002	126.8 ± 2.2	4803 ± 84
21	TES	0.015 ± 0.006	170.8 ± 13.8	11370 ± 919
22	phos	0.027 ± 0.002	30.6 ± 0.5	1136 ± 20
22[a]	TES	0.015 ± 0.002	35.2 ± 0.8	2388 ± 53
22	NH_4	0.810 ± 0.030	44.0 ± 0.6	54.4 ± 0.7

The TES buffer was tested as it is the common buffer used by the *Huber* group[253-255,264,265] who have done extensive point-mutagenesis studies on the activity of the LacZ enzyme and therefore, made for an excellent point of reference to determine the accuracy of our measurements. The values obtained for the hydrolysis of the *o*- and *p*-nitrophenyl galactosides are comparable to those obtained by *Juers et al.*,[265] K_M = 0.12 mM, k_{cat} = 620 s^{-1} and K_M = 0.04 mM, k_{cat} = 90 s^{-1} respectively. However, this group uses His_6-tag-free enzymes, preferring instead to isolate the enzyme *via* size-exclusion chromatography. Despite delivering the best activity for the hydrolysis of the galactoside, unfortunately the *N*-tris(hydroxymethyl)methyl group of the buffering agent could act as a glycosidic acceptor and as such is unsuitable for use with the galactosynthases. An ammonium-based buffer was also trialled as outlined by *Withers* and co-workers,[198] which would allow for evaporative removal of salts from the product mixtures. Sadly though, the use of this buffer dramatically decreased the affinity of the substrate and reduced the turnover, so again, this was deemed unsuitable for use with the LacZ-derived synthases. Inspired again by *Withers* and co-workers,[1] a sodium/potassium mixed phosphate buffer was instead used as it gave results comparable to the TES buffer, but contained no potentially-reactive hydroxyl groups. During these experiments, a substrate inhibition effect was also observed: galactose acts as a competitive inhibitor of LacZ by interacting with the pocket and preventing hydrolysable substrate from entering (Table 3-5). This effect is presumably due to the presence of Trp999 in the shallow site which binds saccharides and aromatics with high specificity.[261]

Table 3-5: Inhibition of galactose on LacZ. phos – 200 mM mixed phosphate, pH 7.5, 1 mM $MgCl_2$; TES – 30 mM TES, 145 mM NaCl, pH 7.0.

Substrate	Buffer	k_i [mM]
gal-oNP	TES	15.3 ± 5.6
gal-oNP	phos	90.7 ± 13.8
gal-pNP	phos	94.5 ± 14.7

Having identified conditions in which the wild type showed good activity, the hydrolytic specific activity of the mutants was also checked. Firstly, by exposure of a sample of the crude lysate to X-gal, which turned blue when exposed to the wild type, but remained colourless in both of the mutant

galactosidases (Figure 3-13). On the chance that the hydrolytic activity was just much slower when compared to the wild type, hydrolysis was measured photometrically using the aforementioned gal-pNP assay in phosphate buffer and as expected, no hydrolytic activity was measurable.

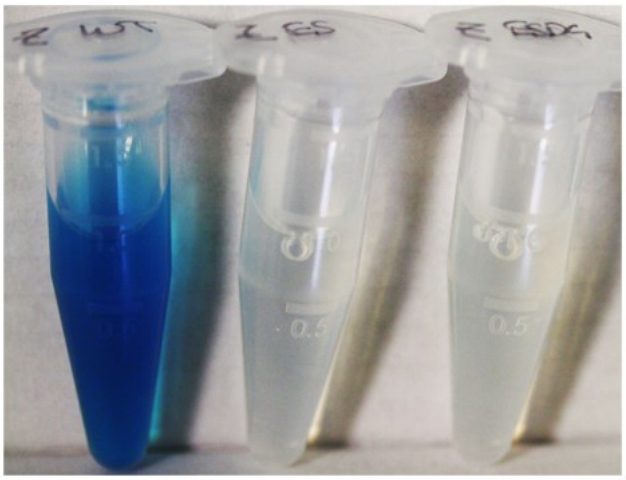


Figure 3-13: X-Gal analysis of the crude lysates of the LacZ variants expressed in *E. coli* ER2566. Left to right: LacZ WT, LacZ E537S, LacZ E537S/G794D.

The hydrolytic substrate range of LacZ was also examined using the *p*-nitrophenyl β-glycosides: 2-acetamido-2-deoxy-D-galactopyranoside (galNAc, **23**), D-galactopyranoside (*p*gal, **22**), D-galactofuranoside (*f*gal, **24**), D-fucopyranoside (fuc, **25**), L-arabinopyranoside (ara, **26**), D-glucopyranoside (glc, **4**), D-xylopyranoside (xyl, **27**), and D-mannopyranoside (man, **28**). These were all tested at a concentration of 1 mM for both the lysate supernatant and the purified enzyme (Table 3-6). No hydrolytic activity was observed for substrates **4**, **23**, **27**, or **28**.

Table 3-6: Hydrolytic activity of LacZ on a range of *p*-nitrophenyl β-glycosides. Those entries with N.D. indicate that no activity was detectable.

pNP β-glycoside	Lysate [μm min^{-1} mg^{-1}]	Purified LacZ [μm min^{-1} mg^{-1}]
*p*gal **22**	1.27 ± 0.06	98.41 ± 1.25
*f*gal **24**	N.D.	2.93 ± 0.25
fuc **25**	N.D.	3.52 ± 0.34
ara **26**	N.D.	24.22 ± 0.21

Galactosamine-pNP (**23**) was difficult to measure due to its low solubility in the buffered solution, but despite its low solubility there was no observable colour change indicating no hydrolytic activity. This is unsurprising though, as the 2-OH group typically forms interactions with a number of residues within the active site and the presence of an acetate group would likely hinder access to the active pocket. It has previously been shown that the enzyme +1 site is highly selective for galactose and therefore, glucopyranose and xylopyranose – those with an equatorial 4-OH position – were not ideal substrates for LacZ.[250,269] Fucose, although a six-carbon sugar, differs from galactose by the

absence of the 6-OH. The 6-OH group has been found to interact with a number of residues and the sodium ion in the active pocket and is also critical in the positioning of the substrate from the shallow to deep binding mode. Perhaps then unsurprisingly, as D-fucopyranose can only form hydrophobic interactions at this position, it is a poor substrate for LacZ. L-arabinopyranose fared better in comparison to fucose and the rate of hydrolysis was only about 4-fold lower than galactose. This is most likely as the 5-methyl group in fucose is not present in arabinose and this non-polar moiety is not placed in a highly polar region making interactions between the enzyme and substrate more favourable. Interestingly, the enzyme shows high specificity for the pyranose form of galactose, with the furanose being less preferred. The furanose form of galactose is a ring formed by the cyclisation of the 4-OH with the aldehyde whereas the pyranose is the result of a 5-OH ring closure. Thus, the furanose has a long, branched side chain which would probe too deeply into the active site and most likely form disfavourable interactions with His540 and Phe601 as it entered the site. However, this is a speculative assessment based on the shapes of the pocket and of the substrate. Confirmation *via* docking studies would be necessary to confirm this hypothesis. The low hydrolytic activity would suggest that there is a high energy barrier required for the substrate to correctly dock into the active site, which in itself suggests some distortion of the substrate and/or the active pocket would be requisite.

3.2 Lac4 from Kluyveromyces lactis

The β-galactosidase from the yeast *Kluyveromyces lactis* (Lac4), an enzyme that already sees extensive use in industry as a dietary supplement, has been shown to be a strong producer of galactooligosaccharides.[223] Furthermore, as it belongs to the same glycohydrolase family as the extensively-studied LacZ from *E. coli*, much can be inferred about this enzyme in terms of structure and function, despite only two crystal structures being available (PDB: 3OBA, 3OB8). Recombinant expression of Lac4 in *E. coli* has also been previously achieved by *Kim et al.*[293] and heterologous expression of this enzyme in our established *E. coli* ER2566 system was expected to be unproblematic.

To date, synthesis using Lac4 as a biocatalyst has been limited to transglycosylation.[222,223,319,320] However, its efficacy as a biocatalyst is limited as glucose is a non-competitive inhibitor of the enzyme (K_i = 758 mM) and galactose is a competitive inhibitor (K_i = 45 mM).[321] It has also been demonstrated that the hydrolytic activity of Lac4 is increased when produced in the presence of β-lactoglobulin or bovine serum albumin.[322] Its product ratio is dominated by the formation of β-(1,6) bonds with minimal amounts of β-(1,4) and -(1,3) products being formed.[222] In addition, it also appears that Lac4 prefers to add only a single galactosyl unit to the carbohydrate galactosyl acceptor and disaccharide formation is preferred.[319] Although, as this enzyme has been explored in the

context of exploitation of the transglycosylation properties, trisaccharides are also known in the form of β-D-galactopyranosyl-(1→6)-D-lactose. This preference for addition to a lactose unit occurs more frequently in commercial preparations of the enzyme as opposed to permeabilised cells.[223] Investigations into the effect of temperature and pH upon the transglycosylation properties of Lac4 have also revealed that disaccharide formation was maximised at a temperature of 50 °C and at a pH of 7.5, but in all cases, the mixture was never more than 15% disaccharide.[320]

3.2.1 Optimisation of the *lac4* Gene

It's well known that although the majority of organisms share the same codon 'language', the frequency of usage of these codons can vary, usually due to environmental factors, with some species, such as *Thermus thermophilus*, avoiding some codons entirely.[323] As such, when expressing a gene heterologously, this discrepancy in the codon usage needs to be considered. As the gene *lac4* is originally from the ascomycetous yeast, *Kluyveromyces lactis*, and it was envisioned that this gene would be expressed in the bacterium, *Escherichia coli*, the codon usages between these two organisms needed to be examined (Figure 3-14).

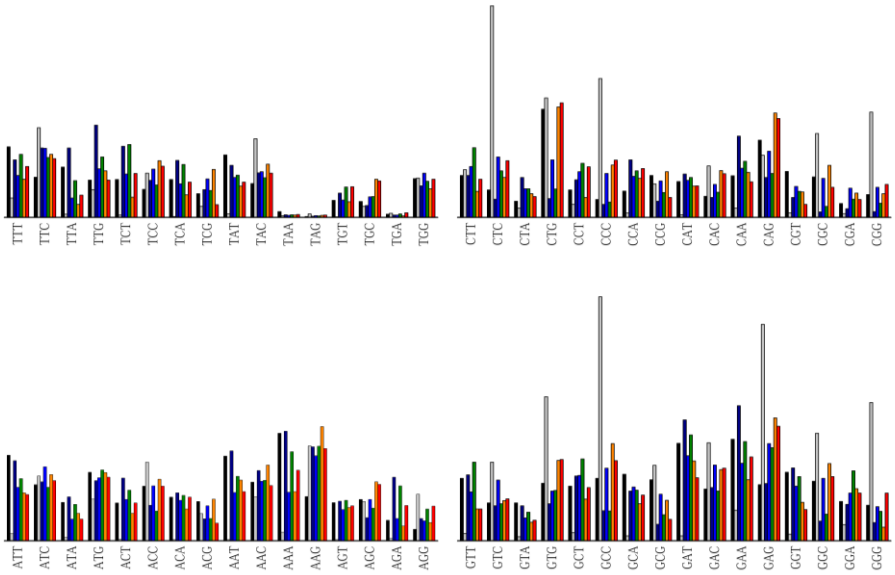


Figure 3-14: Codon usage of various organisms. *E. coli* (black), *T. thermophilus* (grey), *K. lactis* (dark blue), *A. oryzae* (blue), *A. thaliana* (green), *D. melanogaster* (orange), *H. sapiens* (red).

Optimisation of a gene is typically performed by choosing the codon which is most frequently used by the organism and saturating the gene with these high-frequency codons. Polyalanine- and polyglutamine-coding regions are present in a number of eukaryotic genes[324-327] and it is also known that highly repetitive regions can result in ribosomal slippage during translation thereby causing mutations.[328] Organisms with high GC content such as *T. thermophilus* have evolved to ensure the genome remains stable in extreme environmental conditions and these may not be suitable in

mesophilic organisms such as *E. coli*. Taking these factors into consideration, the gene was optimised so as to be harmonious with the homologously-expressed Lac4. This would mimic the rate at which the mRNA is transcribed in *K. lactis* which may be important for correct folding of the monomeric units. Furthermore, as the *lac4* gene is quite large, 3075 bp, opting for a frequency-optimised gene could starve the cell of certain tRNAs during gene overexpression, which may have a negative effect on cell health. The harmonious optimisation of the gene therefore was performed *in silico*, which was designed so that the frequency of the codon used in *K. lactis* was used at a similar frequency in *E. coli* (Figure 3-15). Once gene optimisation had been performed, the gene was then scanned for undesired restriction sites using *Clone Manager* and these restriction sites were manually removed following the principles used for gene optimisation.

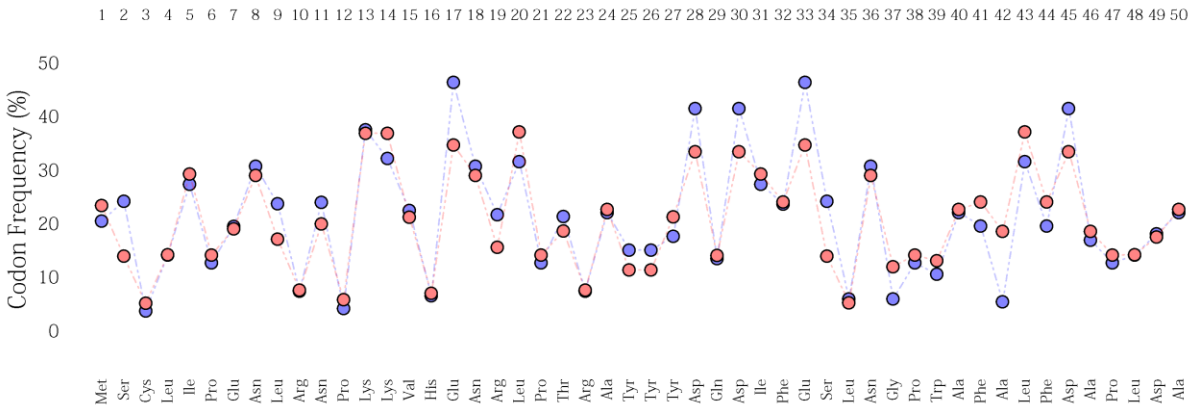


Figure 3-15: Harmonious codon optimisation of *lac4* (blue) for expression in *E. coli* (red) shown for the first 50 amino acids only.

3.2.2 Subcloning of *lac4* into pET22b(+)

After gene optimisation, the synthetic gene in the pUC57 vector was ordered (GenScript, USA) with the 5'-*NdeI* and 3'-*XhoI* restriction sites included. Initially, it was decided to clone the gene into the pET21a(+) vector as this would result in a *C*-terminally His_6-tagged enzyme. The vector, along with the pUC57::*lac4* construct, was then digested using *NdeI* and *XhoI* (Figure 3-16, left). When analysing the agarose gel, it was observed that although the pUC construct displayed correct fragmentation, the pET21a(+) empty vector obtained did not display the expected 5365 bp band as expected, but rather one which was around 6000 bp. This raised three questions: Was the digestion incomplete or unsuccessful? Or was one restriction enzyme accidentally excluded? Was this indeed the pET21a(+) empty vector? As the digestion of the pUC57 construct and pET21 were digested at the same time, it was taken that the digestion was complete, and furthermore, as the restriction enzyme mix was created as a master mix, which was added to the DNA, the both enzymes were definitely included in the digestion as pUC57 displayed correct fragmentation. Therefore, the question of whether the vector was truly empty appeared the most logical. The digested pET21

vector was attempted in a ligation with *lac4* anyway. However, restriction analysis of pUC57::*lac4* and pET21::*lac4* was performed again using *NdeI* and *XhoI* and in all cases, the expected pET21::*lac4* construct did not fragment correctly (Figure 3-16, right). It was therefore determined that the vector was not empty and instead cloning into pET22b(+), which would still return a *C*-terminally His_6-tagged protein, was attempted.

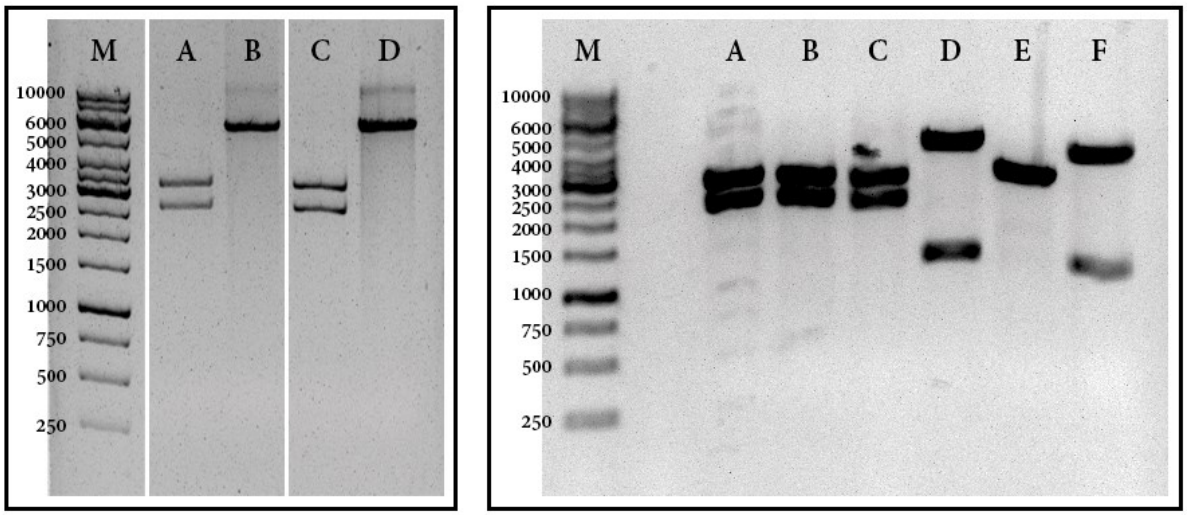


Figure 3-16: Restriction analyses of *lac4* variants using *NdeI* and *XhoI*. M – Marker (GeneRuler™ 1 kb DNA Ladder). Left: the digestion of pUC57::*lac4* (A and C) and pET21a(+) (B and D) in the presence of the phosphatase, FastAP (A and B) and the without (C and D). Right: digestion of the pUC57::*lac4* vectors (A-C, individual colonies) against the intended pET21a::*lac4* ligates (D-F, individual colonies).

The first restriction analysis of pET22b(+) with *NdeI* and *XhoI* showed the expected band size of 5365 bp. With that, pET22b and pUC57::*lac4* were digested using *NdeI* and *XhoI* and the desired fragments isolated. The desired fragments were recombined using T4 DNA ligase and the ligated DNA used to transform competent *E. coli* MACH1 cells and plated onto ampicillin-selective LB agar. Eight colonies were picked from each transformation plate and the plasmids amplified and isolated. These DNA isolates were then analysed using restriction enzymes to identify the various possible recombinations of the DNA fragments confidently (Table 3-7). As the colonies all grew on ampicillin-selective agar, it was assumed that all colonies must contain the ampicillin resistance gene found in either the pUC57 fragment or the pET22 vector. Restriction analysis of the ligation products identified 6 of the 18 clones had the expected fragmentation pattern and these were then confirmed by sequencing (Figure 3-17).

Table 3-7: Theoretical restriction patterns of pET22b::*lac4* ligation and side products by *PaeI*.

Fragment A	Fragment B	Restriction Pattern (*PaeI*)
pET22b (vector)	*lac4*	6483, 1959
pET22b (vector)	pET22b (MCS)	5493
pUC57 (vector)	*lac4*	3945
pET22b (vector)	pUC57 (vector)	8080
pUC57 (vector)	pET22b (MCS)	no cut

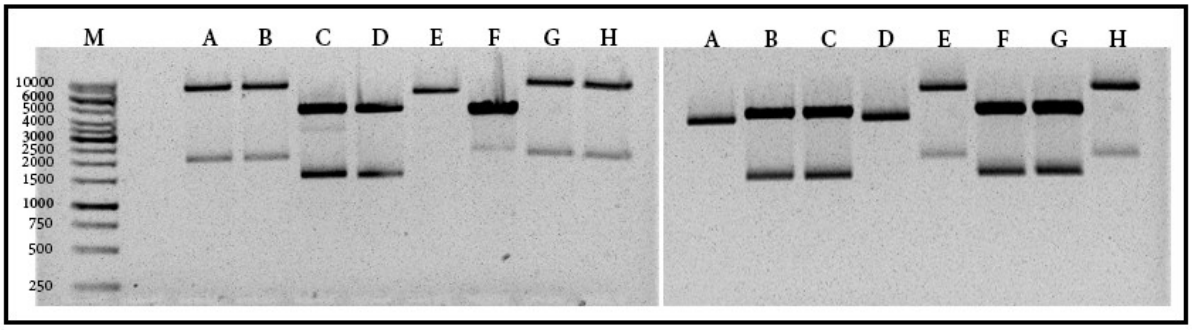


Figure 3-17: Restriction analysis of the ligation products using *NdeI* and *XhoI*. M – Marker (GeneRuler™ 1 kb DNA Ladder). A – Colonies picked from those samples created using the 1:8 mixture; B – colonies picked created by the 1:16 ratio of plasmid to gene.

3.2.3 Mutant Selection and Mutation of *lac4*

To the best of our knowledge, glycosynthases had not been generated from the β-galactosidase from *K. lactis*. Furthermore, the majority of literature-known glycosynthase mutants for retaining glycosidases were Gly, Ala, and Ser mutants, although a few are known with Asn, Gln, Val, Thr, Ile, and Cys mutations.[221] As the nucleophilic residue in Lac4 is Glu, five amino acid substitutions were interesting to investigate: Gly, Ala, Leu, Ser, and Gln (Figure 3-18).

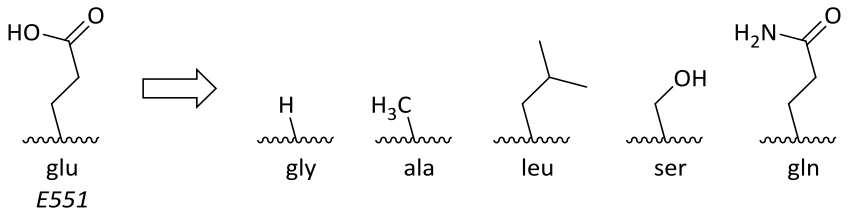


Figure 3-18: Desired mutations for E551 in Lac4.

Firstly, glycine was chosen as it is the smallest of all amino acids. This would presumably free up space within the active pocket to accommodate the leaving group of the activated glycosyl donor. As glycine has no side chain, it is highly flexible and could reduce the stability of the active pocket *via* removal of the rigidity borne of having a side chain. Secondly, alanine was chosen as it is the smallest amino acid with a side chain. Therefore, the presence of the methyl group was expected to maintain

active pocket stability, yet still be small enough so as not to hinder any leaving group associated with the activated glycosyl substrate. Leucine was selected as it is sterically similar to the glutamate residue being substituted. Compared to glutamic acid, leucine is lacking a methylene group and although a more fitting substitution would be homoleucine, we were limited to naturally-occurring amino acids and leucine was determined to be a good approximation. It was also expected that leucine would alter the chemistry of the active site as the side chain possesses no polar functionality; combined with the bulk of the residue, this could result in hindrance of the substrate *via* steric interactions with the α-configured leaving group. However, this negative interaction could encourage the expulsion of the leaving group from the sugar as it is taken into the active pocket thereby encouraging a quicker reaction with an acceptor molecule due to the decreased stability of the oxocarbenium intermediate. Serine contains a polar side chain, although it is quite short. The interaction in the native enzyme between E551 and Y523 is believed to stabilise the oxocarbenium intermediate, specifically the charge, which builds on the O5 ring oxygen. Therefore, the presence of the hydroxyl group of the serine residue could help to stabilise Y523 by sharing the charge build up through proton donation. In addition, as serine is both small and polar, it could more easily accommodate the highly polar galactosyl fluoride when compared to a less polar residue such as valine. Finally, glutamine was chosen as a substitute for the glutamate residue as it is a poor nucleophile, yet retains similar electronic and steric properties of its nucleophilic cousin. As the chemistry of the active site of any enzyme is highly controlled and defined by the residues within, it was thought that the glutamine residue could fill the stabilising role of the glutamate in the active site; specifically, the stabilising hydrogen bonding interactions which E551 typically has with both Y523 and the 2-OH group of the substrate.

The mutants were prepared from the pET22a::*lac4* plasmid using PCR-based site-directed mutagenesis methods. After optimisation of conditions for qcSDM, the *lac4* mutants were prepared and the PCR product used to transform competent *E. coli* MACH1 cells. After colony picking, amplification of the plasmid, and subsequent isolation thereof, the mutants were visualised by agarose gel electrophoresis and restriction analysis (Figure 3-19) and confirmed by sequencing.

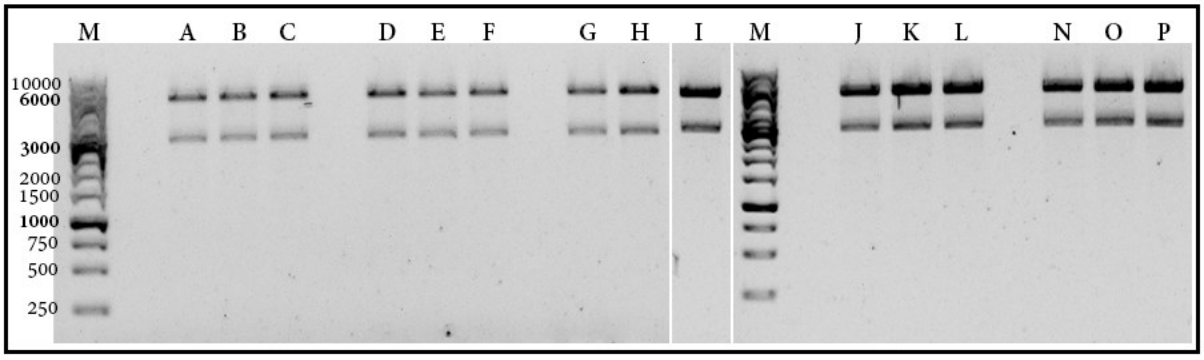


Figure 3-19: Restriction analysis of triplicate clones of each of the *lac4* mutants generated by qcSDM. M – Marker (GeneRuler™ 1 kb DNA Ladder); E551G (A-C); E551A (D-F); E551L (G-I); E551S (J-L); E551Q (N-P).

3.2.4 Recombinant Expression and Purification of Lac4

As the heterologous expression of *lac4* in *E. coli* has previously been demonstrated,[293] it was determined that expression in the developed *E. coli* ER2566 system would be applicable for *lac4*. Furthermore, as with LacZ, Lac4 is a β-galactosidase and a *lacZ* deletion strain such as ER2566 would allow for easy identification even from crude extracts. The optimised expression system developed for *lacZ* was used for the expression of *lac4* unmodified. This was confirmed by exposure of a diluted sample of crude lysate to X-gal (Figure 3-20, left). With confirmation of hydrolytic activity, large-scale purification was performed. Initially, following lysis *via* sonication, the purification of Lac4 from the lysate supernatant was performed using IMAC on the Äkta chromatography robot where the Lac4 was eluted at 250 mM imidazole, as for LacZ. Comparison of the lysate supernatant to the protein fraction eluted using 250 mM imidazole revealed two bands at approximately 120 kDa (Figure 3-20, right).

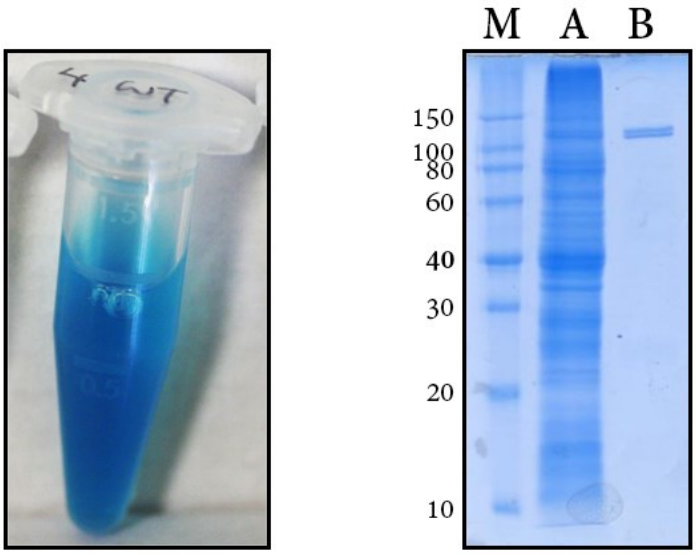


Figure 3-20: Left - X-Gal analysis of the crude lysate of Lac4 WT expressed in *E. coli* ER2566. Right - SDS-PAGE analysis of the cell-free crude extract of Lac4 heterologously expressed in *E. coli* ER2566 (A) compared to the enzyme fraction eluted from IMAC at 250 mM imidazole (B). M – Marker (Roti®-Mark 10-150).

The reason for two bands was not investigated, but presumably relates to two isoforms of Lac4. The efficiency of the purification was determined by checking the activity in each fraction and comparing that with the protein observable by SDS-PAGE (Figure 3-21). The purification analysis was performed using a peristaltic pump for convenience; although, this eventually became the method-of-choice for protein purification.

Fraction	Protein Concentration [mg mL^{-1}]	Activity [µM min^{-1} mg^{-1}]
Resuspended Cells	0.28 ± 0.19	27.05 ± 0.09
Disrupted Cells	43.82 ± 7.19	5.10 ± 0.09
Lysate Supernatant	13.61 ± 0.71	19.67 ± 0.19
Pellet	2.38 ± 0.36	5.79 ± 0.04
Flow-Through	6.36 ± 0.28	0.06 ± 0.01
Column Wash	1.48 ± 0.09	0.17 ± 0.09
Elution (250 mM)	0.30 ± 0.01	36.95 ± 0.69
Elution (1 M)	0.14 ± 0.004	N.D.
Size-Exclusion	0.33 ± 0.01	113.37 ± 0.58

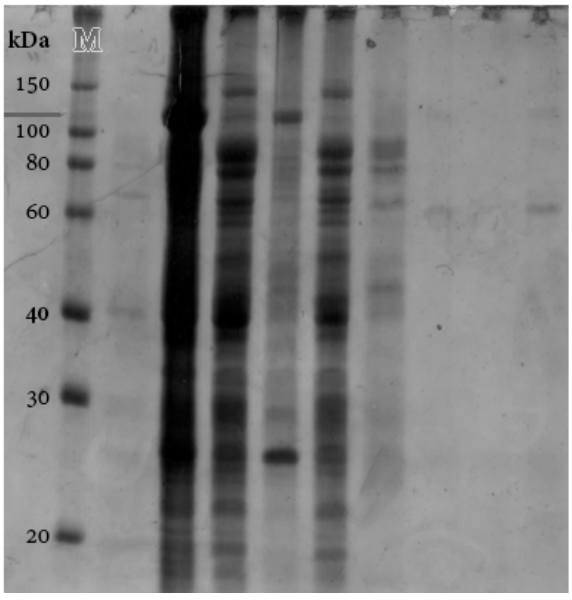


Figure 3-21: Purification table constructed for the IMAC purification of pET22::*lac4* variants using the wild type as a model. The grey line at ca. 110 kDa indicates the expected mass of the His_6-tagged LacZ protein.

The initial purification technique resulted in a low amount of enzyme, approximately 5.9 mg L^{-1}, which seemed to indicate one of two things; either the enzyme was not being properly overexpressed in the *E. coli* ER2566 system, or that a large amount of enzyme was being lost in the purification. Analysis of the preliminary purification scheme indicated there was still active enzyme present in the pellet. Therefore, the pellet was resuspended in buffer and centrifuged a second time and the supernatant collected. The pellet supernatant was then collected and processed as normal which yielded an extra 0.25 mg of purified protein; however, this protein only had a specific activity of 0.04 U mg^{-1}. The other weakness of the original purification method is in the column loading. It was apparent from the flow-through and wash fractions that not all protein was binding to the column as there was still protein present in the supernatant after loading the column for 30 minutes on ice, but more importantly, Lac4 was being eluted in the washing fraction – the fraction eluted by low imidazole concentrations. This could indicate one of two things: a) the protein is not binding strongly to the column, or b) the His_6-tag is hindered in binding to the column and the protein did not have an opportunity to bind to the column. The former is most likely an extension of the latter, and could be investigated by optimising the location and/or identity of the tag, but this was not explored. Rather, this problem was addressed by performing a second and third round of loading and purification, increasing the loading time to 45 minutes, and decreasing the pump speed from 4.1 mL min^{-1} to 2.3 mL min^{-1}. This seemingly small change dramatically increased the amount of protein eluted in the first elution step; a 2.5-fold increase from 0.3 to 0.8 mg mL^{-1}. Furthermore, the protein eluted from this round appeared to be more pure, as evidenced by the 10-fold increase in specific activity when compared to the pre-optimised conditions.

After analysing both of these apparent weaknesses in the purification methods, the yield was increased from 6 mg L^{-1} to approximately 46 mg L^{-1}; an 8-fold increase. The results of this optimisation are summarised in Table 3-8. It should be noted though that this more than doubles the time required for the purification of protein. Optimal storage conditions of the purified protein should be considered and investigated in order to properly benefit from this yield increase. In addition, attempts to scale-up the protein purification of Lac4 resulted also in lower yields of protein and the recovered protein also displayed a lower activity. A brief investigation into this phenomenon resulted in the realisation that imidazole is an irreversible inhibitor of $Lac4^{293}$ and the use of His_6-tags for the purification of Lac4 is therefore less-than-optimal and other methods of purification of this enzyme should be investigated in subsequent studies.

Table 3-8: Summary of the effect of optimisation of the purification of pET22::*lac4* using Ni-NTA.

	Before Optimisation	After Optimisation
Pellet Mass [g L^{-1}]	10 – 12	10 – 12
Mass of Pellet Used [g]	2.38	2.54
Mass Enzyme (Lysate) [g]	188	193
Mass Enzyme (Purified) [g]	4.9 – 5.8	39 – 46

3.2.5 Hydrolytic Activity of Lac4 and Kinetic Parameters

In a similar fashion to that of LacZ, the activity and kinetic parameters of Lac4 and associated mutants were determined *via* colourimetric analysis using gal-pNP as a hydrolytic substrate. The range of conditions were not checked as thoroughly as for LacZ. The WT and mutant enzyme activities checked using the phosphate buffer (Table 3-9).

Table 3-9: Summary of the hydrolytic activity of the Lac4 variants on 1 mM gal-pNP in phosphate buffer. Buffer: mixed phosphate (200 mM), pH 7.5, $MgCl_2$ (1 mM).

Enzyme	Specific Activity [U mg^{-1}]
WT	226.7 ± 12.2
E551G	0.2 ± 0.0
E551L	0.6 ± 0.3
E551S	0.5 ± 0.0
E551Q	0.0 ± 0.0

When comparing the activity of Lac4 WT to LacZ WT, a number of observations were made (Table 3-10). Of note, Lac4 has a much lower substrate specificity in comparison to LacZ indicated by the 50-fold higher K_M of Lac4 and the 10-fold lower catalytic efficiency as determined by k_{cat}/K_M.

Table 3-10: Experimentally-determined kinetic parameters of LacZ-His_6 and Lac4-His_6.

	Lac4 WT	LacZ WT
K_M [mM]	1.354 ± 0.087	0.027 ± 0.002
k_{cat} [s^{-1}]	209.1 ± 4.1	30.8 ± 0.5
k_{cat}/K_M [s^{-1} mM^{-1}]	154 ± 3	1137 ± 20
$k_{i\ (gal\text{-}pNP)}$ [mM]	N.D.	95.0 ± 14.7

One of the major differences in the active site architecture between the two enzymes is the presence of an aromatic residue in LacZ (W999), which is exchanged for a cysteine in Lac4 (C1001). In a study by *Huber et al.* in 2003,[261] the tryptophan residue was demonstrated to be critical for the activity of LacZ. Mutations of this residue for the non-polar, aromatic phenylalanine resulted in a 2-fold decrease in specificity and a slight decrease in overall catalytic ability. Glycine mutants resulted in a 2.5-fold decrease in specificity a 4.5-fold decrease in catalytic ability. The polar, aromatic tyrosine resulted in a dramatic, 20-fold decrease in specificity and 24-fold decrease in catalytic ability, while a

leucine mutant was 27.5- and nearly 100-fold worse than the tryptophan-containing wild type. Crystallographic studies on LacZ,[250] revealed an interaction between the aromatic electrons of tryptophan and the C-H bonds of the underside of the carbohydrate. The orientation of the substrate in the 'shallow' site of LacZ also suggested that there was a polar interaction between the 5-O of the substrate and the indolic nitrogen. The relatively small variation between the Trp and Phe residues in the *Huber* study would suggest that the aromatic interaction dominates in the positioning of the substrate. Due to the similarity between the two enzymes, the decreased activity of Lac4 in comparison would suggest the lack of an aromatic system in the mouth of the Lac4 active site retards the binding of the substrate in the pocket prior to hydrolysis. Trp999 has also been shown to bind to the aromatic *o*-nitrophenyl and *p*-nitrophenyl aglycones, affecting the rate of hydrolysis.[234] As such, a C1001W mutation in Lac4 could be beneficial in increasing the hydrolytic activity of Lac4 by increasing substrate specificity. On the other hand, as galactose is bound in a highly similar fashion in Lac4 as in the 'deep' site of LacZ, the lack of an aromatic system in this position may increase the variety of glycosidic acceptor molecules in the galactosynthase. This hypothesis could be tested in the wild type by analysing the relative hydrolytic activity of the two enzymes for different galactosides.

The other interesting difference between the two enzymes is the inhibitory effect of gal-pNP on LacZ which appears absent in Lac4. The fact that this effect is also observed in the hydrolysis of gal-oNP would indicate that rather than gal-pNP or pNP having the inhibitory effect, the hydrolysed galactose is responsible for this activity. Again, *Huber et al.*[300] demonstrated that galactose is a good inhibitor of galactosidase (K_i = 20 mM) and presumably acts as a competitive inhibitor by blocking the active site by interacting with Trp999. Galactose has also been shown to be a competitive inhibitor of Lac4,[321] but the K_i (45 mM) is more than double that of LacZ. Galactose inhibition of LacZ was observed in TES buffer for the oNP substrate (K_i = 15.3 ± 5.6 mM), although the inhibitory activity is reduced 6-fold in the phosphate buffer (K_i = 90.7 ± 13.7 mM). Regardless, from a synthetic perspective, this lower inhibitory value of galactose is useful, as it would allow a higher concentration of activated substrate to be present in the reaction mixture without having a dramatic effect on the activity. As previously stated, Trp999 does interact strongly with the aglycone, but as the reaction mixture is slightly basic (pH 7.5), the charged phenolate is expected to interact more with the anions in solution than entering the negatively-charged active site. As such, the inhibitory effect of galactose is expected to dominate.

3.3 Glycosidases from *Aspergillus oryzae*

Industrially, filamentous fungi like *Aspergillus oryzae*, *A. niger*, and *Trichoderma reesei*, are a highly interesting source for hydrolytic enzymes due to their ability to secrete high titres of protein into growth media.[298] The glycosidases of *A. oryzae* are known for their tolerance to higher temperatures and lower pHs which can be valuable attributes for synthetic applications.[225,286,329-332] With respect to

glycosynthases, these properties are quite interesting as our glycosyl donor is a glycosyl fluoride, which upon reaction with a glycosyl acceptor generates HF, thus decreasing the pH of the system in the absence of a buffer. Furthermore, as thermostable enzymes are typically more tolerant to organic solvents,[333] this may have the added effect of allowing a broader substrate range.

The β-galactosidase, LacA, has been in use for over 40 years as a supplement for lactose-intolerant people and it was therefore considered an interesting enzyme to compare to LacZ of *E. coli* and Lac4 of *K. lactis*. Furthermore, β-glucosidases from *A. oryzae* had previously been shown to be hydrolytically active for the isoflavones, diadzin (**29**) and genistin (**30**).[334] These two compounds are interesting due to their cardio-protective and anti-cancer effects and an efficient chemoenzymatic synthesis thereof would be of interest.

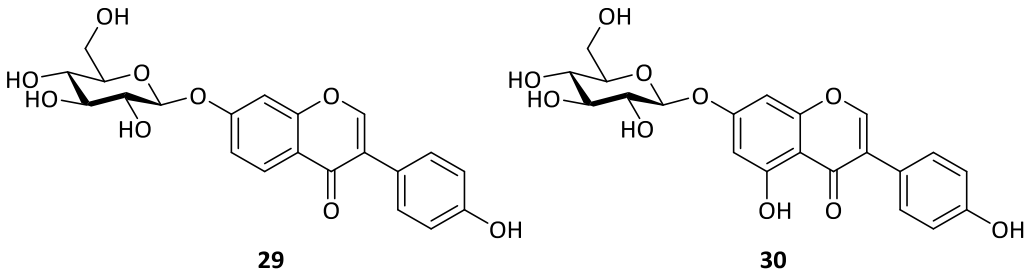


Figure 3-22: The structures of the isoflavones diadzin (**29**) and genistin (**30**).

3.3.1 Isolation of Genomic DNA from *Aspergillus oryzae*

Having been identified as proteins of interest, isolation of the genes, *bglJ* and *lacA,* from *A. oryzae* was attempted in a similar fashion to that of *E. coli*. To this end, a culture of *A. oryzae* RIB40 was obtained from the *Deutsche Sammlung von Mikroorganismen und Zellkulturen* (DSMZ; DSM 63303) and this was grown on potato dextrose agar for 5 days at 30 °C. The first attempt at extracting DNA was a direct gene extraction of *bglJ* using colony PCR. Two samples were therefore prepared, in which a similar mass of mycelia was suspended in H_2O, one of which was then incubated at 99 °C for 10 minutes to degrade the cell wall. These two samples were then used as the template for colony PCR and an annealing temperature screen was performed across the range of 55-70 °C. Unfortunately, this PCR experiment was unsuccessful and raised two questions: were the cells not lysed? Or was it a case of the primers not binding to the DNA properly?

To test this, a DNA extraction was performed following a literature-described method.[335] This method returned a small amount of DNA (0.5 μg) as determined spectrophotometrically. Analysis of the extraction product by agarose gel electrophoresis revealed that a small amount of DNA was obtained as evidenced by the upper band. The lower band indicated that there was also an amount of RNA extracted. The extraction of *bglJ* by PCR was then re-trialled in an identical fashion as

previously but using the extracted DNA as a template, but this gave the same results. Two questions again came to mind: is the DNA fragment too long to amplify? Or are the primers not binding correctly? As the *bglJ* gene contains introns, it was already planned that these would be excised using overlap extension PCR (oePCR). So to answer the question of fragment length, each oe-fragment was attempted to be isolated separately using the same PCR method as above.

Again, there was no product returned. Digestion of the DNA was then atempted using RNase A and the restriction enzyme *XbaI* and incubating for a further 3.5 hours. This was then used as a template for PCR-based gene isolation using the same method; although one sample was 2% DMSO (v/v). Despite this, agarose gel electrophoresis indicated that there was no gene isolated (Figure 3-23).

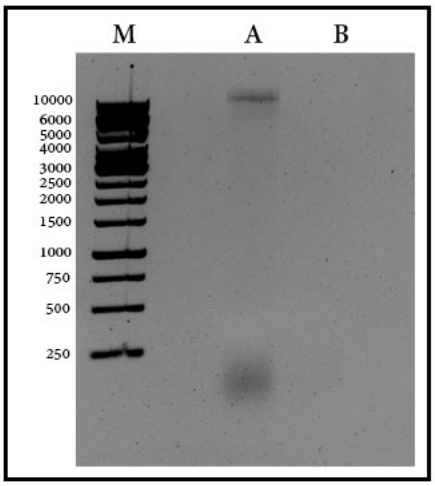


Figure 3-23: Comparison of isolated *A. oryzae* DNA (A) and the same DNA after digestion with RNase A (B). M – Marker (GeneRuler™ 1 kb DNA Ladder).

In light of these results, it was decided to attempt to reisolate the DNA using the same method and comparison to a different method.[336] These two samples were then split into four. The first sample was the crude genomic extraction. A second sample was digested using RNase A. Sample 3 was digested using the restriction enzyme *XbaI* directly following digestion by RNase A. Finally, sample 4 was digested using *DpnI* following exposure to RNase A. All samples were then analysed photometrically to determine the DNA concentration in each and it was found that there was no DNA extracted using the new protocol.

Regardless, as the gene *bglJ* proved difficult to isolate using PCR, isolation of another target gene, *lacA*, was attempted. Conditions for the PCR-based isolation of *lacA* was then screened using a temperature gradient, with and without DMSO, and three DNA polymerases. The isolation of DNA using this method of *Prabha et al.*[335] appeared then to be the biggest problem. Thus, isolation was once again tried using the method of *Bir et al.*[336] and this time, 4.8 µg of DNA was isolated. A sample of this was then treated using RNase A analogously to previous extractions, and then was used as the

template for further screening of *lacA* isolation methods. Despite extensive testing for optimal PCR conditions, we were unable to return a product. Because of this, it was decided to optimise and obtain a synthetically-created gene.

3.3.2 Optimisation and Subcloning of the *lacA* Gene

As *bglJ* and *lacA* were frustratingly difficult to isolate *via* PCR, it was determined to proceed only with a synthesised version of the β-galactosidase gene, *lacA*. It was determined that in order to directly compare results, and therefore *lacA* should also be expressed in *E. coli* ER2566. Similar to that of the *lac4* gene, it was decided to optimise the *lacA* gene using a codon-harmonised, rather than most-used-codon approach. As the natural gene has 3 introns, as identified by sequence alignment,[337-339] these regions were removed prior to optimisation. Following optimisation, the gene was analysed for undesired restriction sites and repetitious regions and these were manually altered following the same harmonious pattern where possible. The gene was then submitted for synthesis and arrived in the pUC57 vector having been cloned into it using *EcoRV*.

As with the synthetically-ordered *lac4* gene, the synthetic *lacA* gene was also designed with a 5' *NdeI* restriction site and a 3' *XhoI* site in order to clone it into the pET22b(+) vector. Restriction of the pUC57::*lacA* construct was performed by *NdeI* and *XhoI* and the restriction pattern was as expected. The gene was then cloned into the pET22b(+) vector.

3.3.3 Mutation of *lacA* and Activity

Now that the pET22::*lacA* plasmid had been obtained, site-directed mutagenesis was required to generate the library of galactosynthases. In an identical fashion to *lac4*, five mutations were chosen: gly, ala, leu, ser, and gln. Due to the previous successes with qcSDM, it was decided to apply this method to the *lacA* constructs. This approach was ultimately successful and the mutants were analysed by agarose gel electrophoresis, and a restriction analysis using *NdeI* and *XhoI* (Figure 3-24). The mutant genes were then confirmed by sequencing.

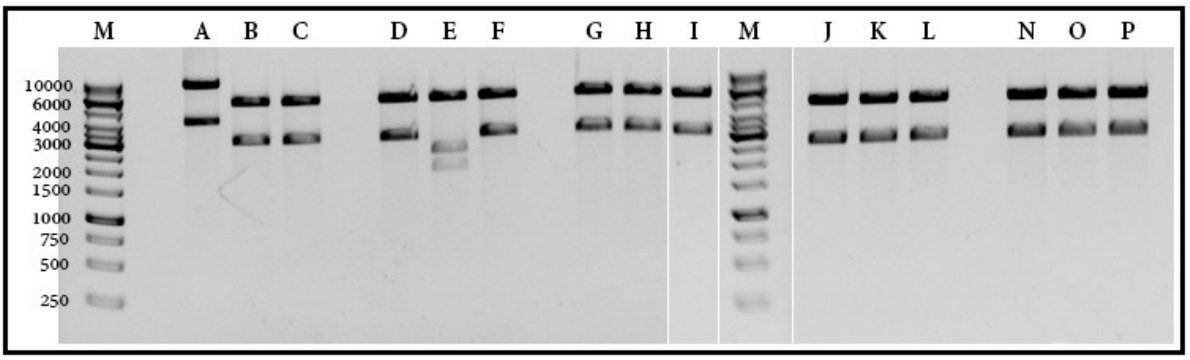


Figure 3-24: Restriction analysis of the triplicate clones produced from the qcSDM of each *lacA* mutant. M - Marker (GeneRuler™ 1 kb DNA Ladder); E298G (A-C); E298A (D-F); E298L (G-I); E298S (J-L); E298Q (N-P).

The hydrolytic activity of the wild type galactosidase was tested using a dilute sample of the crude lysate and exposing it to X-gal (Figure 3-25, left). Previous studies into native LacA have shown that it has an optimal pH of 4.5 and the crude lysate was thus diluted into a citrate-phosphate buffer adjusted to this pH. Unfortunately, this experiment showed no observable β-galactosidase activity. The presence of enzyme was checked *via* SDS-PAGE, which showed that although the enzyme was being overexpressed, it was located within the pellet fraction and absent from the supernatant fraction (Figure 3-25, right). This indicated the presence of inclusion bodies. This was a disappointing result and it was decided to shift focus to the two enzymes, which displayed hydrolytic activity, LacZ and Lac4.

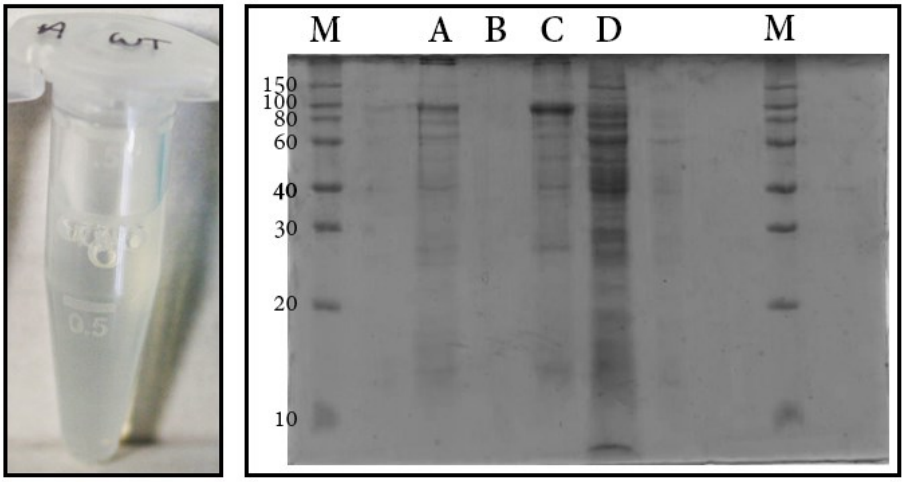


Figure 3-25: Left - X-Gal hydrolysis analysis of the crude lysate of LacA heterologously expressed in *E. coli* ER2566. Right - SDS-PAGE analysis of the expression of LacA in *E. coli* ER2566. M - Marker (Roti®-Mark 10-150); A – Culture; B – Culture Supernatant; C – Disrupted Pellet; D – Cell-free crude lysate.

3.4 Quantification of Synthase Activity

Glycosynthases need two components for a reaction: a glycosyl acceptor and a glycosyl donor (Figure 3-26). The donors come in many forms, but in this work, we used exclusively a galactopyranosyl fluoride donor (galF, **1**). The acceptors are a much broader class of molecules, but all require nucleophilic capability. These are therefore limited to heteroatoms, typically oxygen, sulfur, and nitrogen. This work focussed on the use of alcohols as glycosyl acceptors.

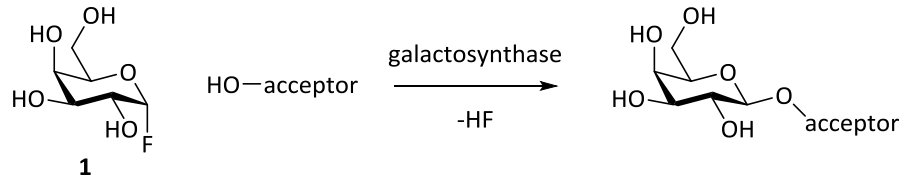


Figure 3-26: General reaction scheme for a galactosynthase using galF (**1**) as the glycosyl donor.

3.4.1 Photometric Quantification

The ease and success of the spectrophotometric determination of hydrolytic activity of the β-galactosidases made this method interesting for investigating the activity of the reverse-hydrolysis reaction. As our hydrolytic assay used *p*-nitrophenyl β-D-galactopyranoside (gal-pNP), it seemed therefore logical to attempt to measure the synthase activity by reacting *p*-nitrophenol (pNP) with galF in a 96-well plate and measuring a decrease in absorbance at λ = 405 nm in pH 7.5 buffer.

This methodology was attempted using the Lac4 library of galactosynthases and were tested at pH 7.5 at 25 °C and 40 °C, the optimal temperature of the wild type enzyme for hydrolysis of lactose.[293] Background reactions between pNP and galF were not observed at either of these temperatures. One sample series with the removal of galF also gave a baseline for the background absorbance of pNP in the reaction buffer. The hydrolytic activity of the wild type was checked concurrently as both a control for the spectrophotometer and as a control for the viability of the enzymes, as they had been harvested a week prior. As the week-old WT enzyme showed hydrolytic activity, and it is known that the architecture of the active site is unaffected by point mutation, it was taken as evidence that all enzyme variants were still viable. Each enzyme was added undiluted in triplicates to reaction wells prepared in triplicate containing pNP and galF in buffer with final concentrations of 0.33 μM and 25 mM respectively. Unfortunately, in all cases, this method showed no decrease in absorbance. As we were searching for a universal method, and not enzyme-specific, and this method did not appear to be appropriate for Lac4, the effectiveness of this method was re-assessed.

As a direct measurement of the substrates was not viable, yet colourimetric analysis provided a convenient method for measuring reaction rates, an indirect method was sought. Furthermore,

96-well plate-based assays provided a platform for collecting a large number of data points simultaneously, allowing for rapid assessment of a variety of conditions. Furthermore, by allowing a large number of reactions on a miniaturised scale, the amount of materials required for each measurement was vastly reduced. In collaboration with a colleague, *Marc Hayes*, a discontinuous, colourimetric fluoride-measuring assay was devised.[340] This method exploited the lability of silyl ethers to fluoride and the assay was designed in three parts (Figure 3-27). Firstly, the reaction is carried out using galF and glc-pNP in various concentrations. After a certain amount of time, a sample of the reaction is transferred to a second plate containing TIPS-protected pNP in MeCN. This is allowed to react for a further set period, before a fraction of the MeCN solution is transferred into a third plate where the absorbance is measured.

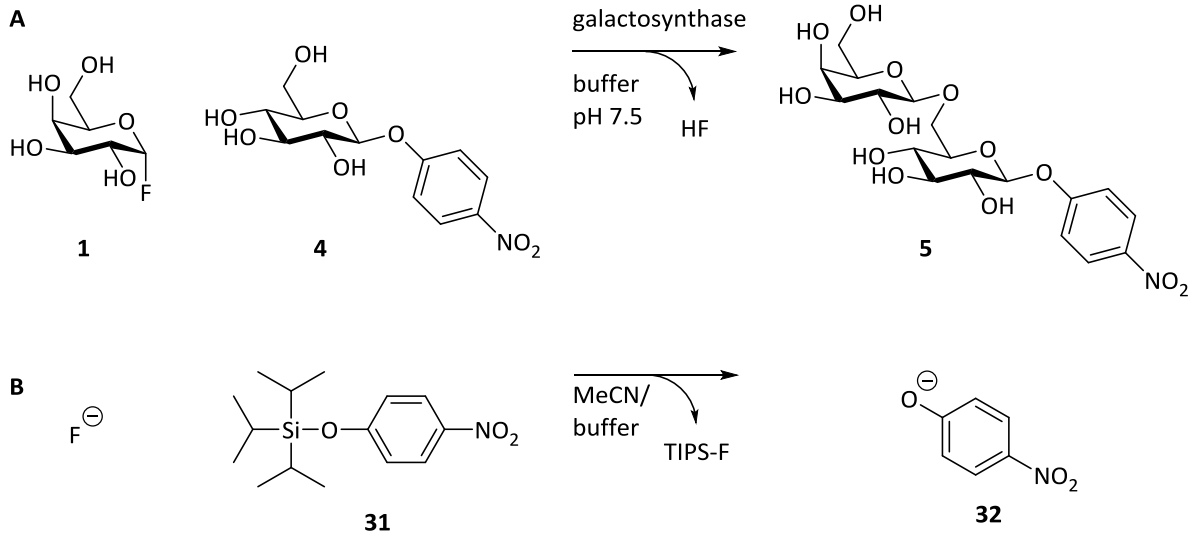


Figure 3-27: Two-part reaction schematic for the measurement of the activity of galactosynthases *via* fluoride measurement analysis. A - The enzyme-catalysed synthase reaction between galF (**1**) and glc-pNP (**4**) generates a disaccharide and a molecule of HF. B - The fluoride ion reacts with TIPS-pNP (**31**) to release the *p*-nitrophenolate anion (**32**) which has an absorption maximum at λ = 405 nm under basic conditions.

The synthesis of TIPS-pNP (**31**) is described in Section 5.4.8 and analysis *via* NMR showed only minor impurities. Firstly, the optimal reaction time for the deprotection of TIPS-pNP was determined by adding, at various time points, an amount of NaF to wells containing 1 mM solution of TIPS-pNP in MeCN so that the final concentration of NaF was 5 mM (Figure 3-28). The result of this assay showed that a reaction time of 10 minutes was sufficient for deprotection of pNP as evidenced by the flattened curve between 6 and 11 minutes (slope = 0.0010 ± 0.0001, R^2 = 0.984)

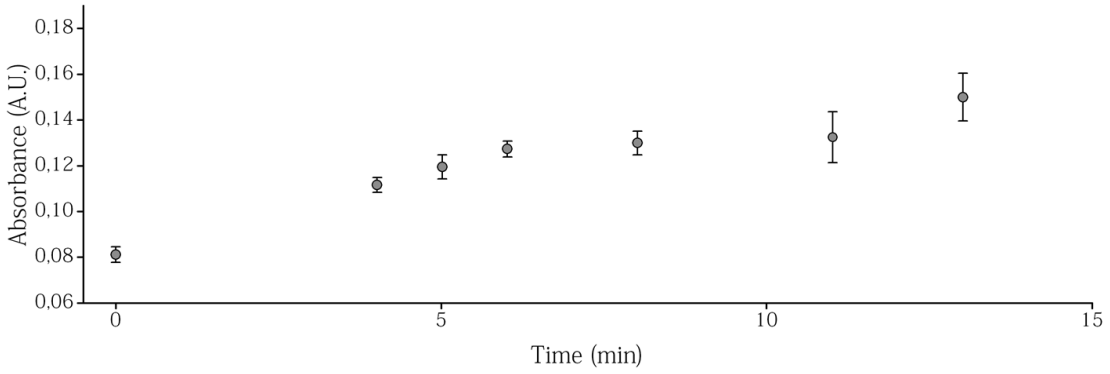


Figure 3-28: Determination of the optimal reaction time for the deprotection of 1 mM solution of TIPS-pNP (**31**) in MeCN by NaF. Absorbance was measured at λ = 405 nm. Reaction conditions: TIPS-pNP (1 mM in MeCN, 180 µL), $NaF_{(aq)}$ (10 mM, 20 µL), room temperature

With the reaction time for the deprotection step set, the calibration curve could be constructed. It was created by adding concentrations of NaF (0-10 mM) in reaction buffer to a 1 mM solution of TIPS-pNP (**31**) in MeCN (Figure 3-29). The calibration was constructed using the data points between 0.1 and 2.0 mM of stock solution and shows excellent correlation (R^2 = 0.992). The large standard deviation of the higher concentrations is concerning, as is the clear deviation in the graph. The absorbance observed is towards the lower limits of the spectrophotometer, and it was attempted to increase the sensitivity of the reading.

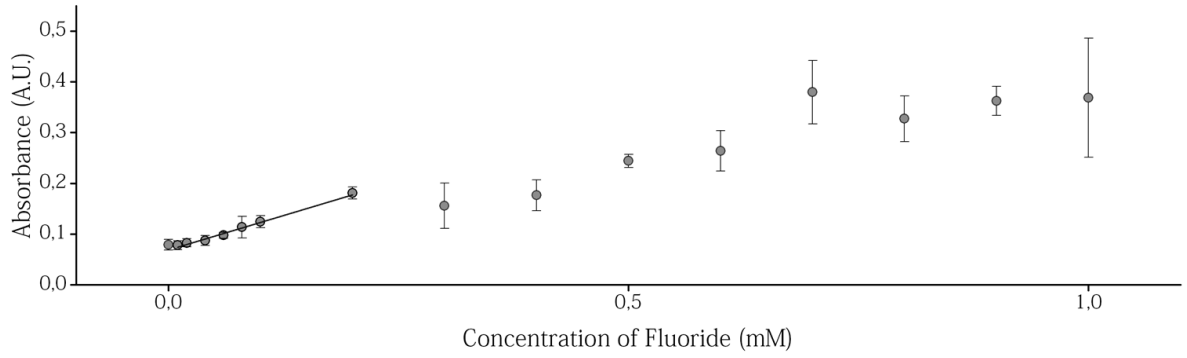


Figure 3-29: Calibration curve of the NaF-mediated degradation of various concentrations of TIPS-pNP (**31**) in MeCN. Concentrations shown reflect the final concentration of the well, not the concentration of the stock solution used. Absorbance was measured at λ = 405 nm at pH 7.5 (200 mM phosphate buffer). Reaction conditions: TIPS-pNP (1 mM in MeCN, 180 µL), $NaF_{(aq)}$ (various concentrations, 20 µL), 10 min, room temperature R^2 = 0.979 (0.0 mM-1.0 mM); R^2 = 0.997 (10 µM-200 µM).

As the absorption at λ = 405 nm of the *p*-nitrophenolate anion increases with pH, it was decided to add a greater volume (100 µL) of the MeCN reaction to a small amount of an aqueous base. Due to the low absorbance, it was also considered that adding a higher volume of $NaF_{(aq)}$ to a more concentrated solution of TIPS-pNP would allow for a faster reaction rate. While this theory does not seem relevant to the calibration curve, it would allow for a higher titre of fluoride from the reactions

to be transferred to the TIPS-pNP solution, thereby allowing a higher amount of TIPS-pNP to react and this would hopefully result in measurements that are more sensitive. As such, these two modifications were implemented in a subsequent calibration using a 100 mM aqueous solution of Na_2CO_3 as the base (Figure 3-30).

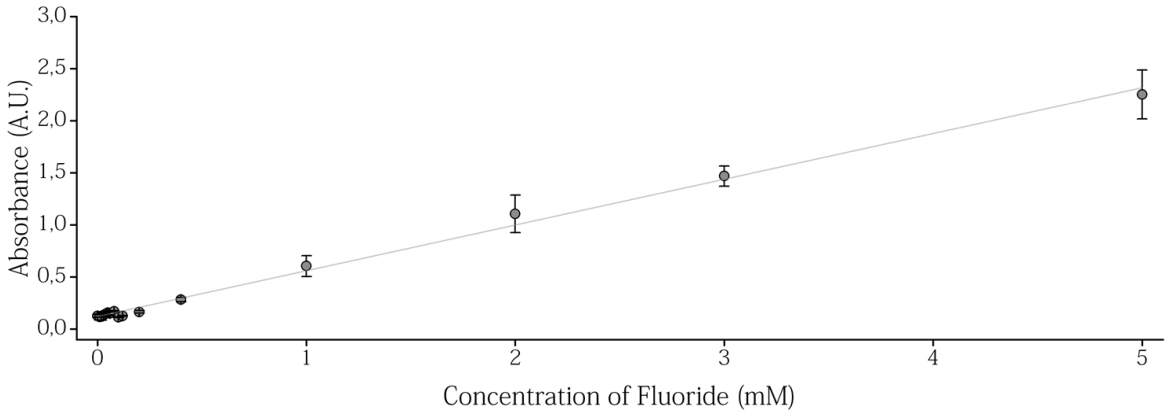


Figure 3-30: Calibration curve of the deprotection of TIPS-pNP (**31**) using base-enhancement to improve sensitivity. Concentrations shown reflect the final concentration of the well, not the concentration of the stock solution used. Absorbance was measured at λ = 405 nm in 100 mM solution of Na_2CO_3. Reaction conditions: TIPS-pNP (1.20 mM in MeCN, 160 µL), $NaF_{(aq)}$ (various concentrations, 40 µL), 10 min, room temperature, *then* 100 µL of MeCN solution transferred to wells containing $Na_2CO_{3\,(aq)}$ (100 mM, 50 µL). R^2 = 0.995 (0.0 mM-5.0 mM).

As hoped, this proof-of-concept experiment did indeed improve the sensitivity of the measurement, resulting in high correlation of absorbance at λ = 405 nm to the concentration of fluoride present in solution (R^2 = 0.998). These modifications were then applied in the measurement of the kinetic parameters of the galactosynthases for the substrates galF and glc-pNP. This method was therefore tested using the literature-known, but as-yet-uncharacterised LacZ E537S mutant in a concentration of 5 µM of enzyme. The concentrations of the glycosyl donor, galF, were varied between 0 and 20 mM, as were the concentrations of the glycosyl acceptor, glc-pNP, and these were mixed in different ratios to a final volume of 200 µL in a 96-well plate. They were allowed to react in the presence of the galactosynthase for 15 minutes, before 100 µL of the reaction mixture was transferred to a new 96-well plate where each well contained 100 µL of TIPS-pNP (2.0 mM in MeCN) and these were allowed to react for 15 minutes. 100 µL of this solution was then transferred to a new 96-well plate where each well contained 100 µL of $Na_2CO_{3\,(aq)}$ (100 mM) which was then measured using a plate-reading spectrophotometer. These readings were then compared to concurrently prepared sample for the enzyme solution was replaced with buffer in order to negate the background reaction of the two reagents in their various ratios. Disappointingly, this method did not appear to be sensitive enough for measuring the activity of the enzymes at this concentration. Logically therefore, subsequent experiments would need to be performed at higher enzyme concentrations. This approach is limited as the overall enzyme yield of our expression system is low and the amount of

enzyme available for use in these assays is restricted as a result. As such, this method, while showing potential, is not applicable to our system and more sensitive methods were explored.

3.4.2 Fluoride-Selective Electrode

Developed concurrently to the TIPS-pNP-based colourimetric assay and inspired by previous characterisations of glycosynthases using fluoride-selective electrodes,[341-343] an electrochemical approach to fluoride quantification was also assessed for viability. Using a fluoride-selective electrode coupled to a pH meter, it was hoped to allow for fine measurements of fluoride concentration in solution. This method would allow for the determination of efficacy and reaction rates on a scale similar to that of the enzyme-catalysed reactions.

The first step in assessing the validity of the process is the construction of a calibration curve. To this, NaF dissolved in reaction buffer was used as the fluoride source. The curve was constructed by measuring fluoride concentrations between the limit of detection, 5 μM, and 500 mM (Figure 3-31). As the relationship between the millivoltaic response and the concentration is logarithmic, the base-10 logarithm of the concentration was measured against the response. After this modification to the values of concentration, a linear relationship was observed across this range (R^2 = 0.995).

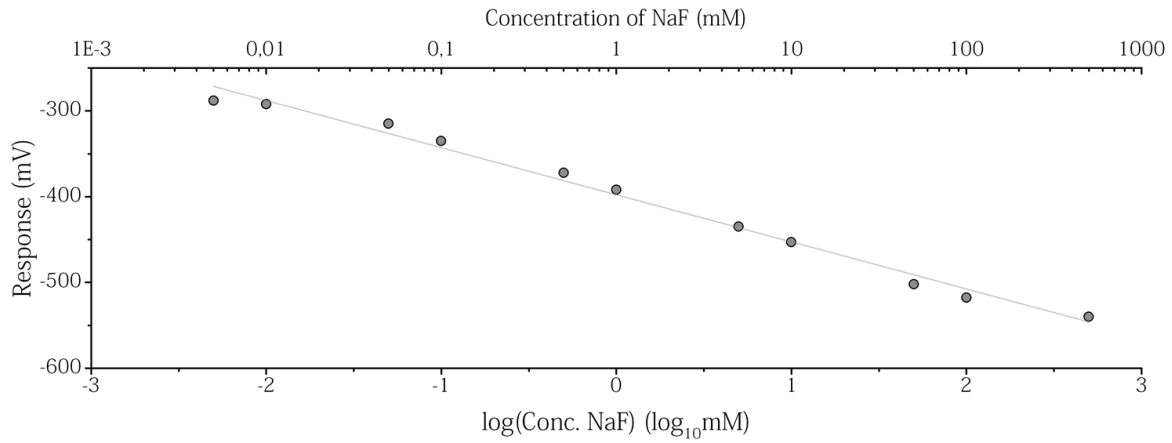


Figure 3-31: Calibration curve for the fluoride-selective electrode across a range of 5 μM to 500 mM in reaction buffer using NaF as the fluoride source.

With the fluoride concentrations now calibrated to the voltaic output of the pH meter, the background degradation rate of the galF substrate needed to be investigated. As steady-state kinetic measurements need to be measured at the start of a reaction, it was determined that the stability of galF would be measured over a 10 minute period at 30 second intervals. To maximise the likelihood of measuring the degradation of galF, three background measurements were taken at 84.0, 89.5, and 131.1 mM concentrations of galF in reaction buffer (Figure 3-32). In all cases, there was no observable degradation of galF, indicating that the stability of the compound in the buffered aqueous

system was stable enough over the measurement period. Any variations in the concentration of fluoride over the reaction time period could therefore be attributed to the enzyme.

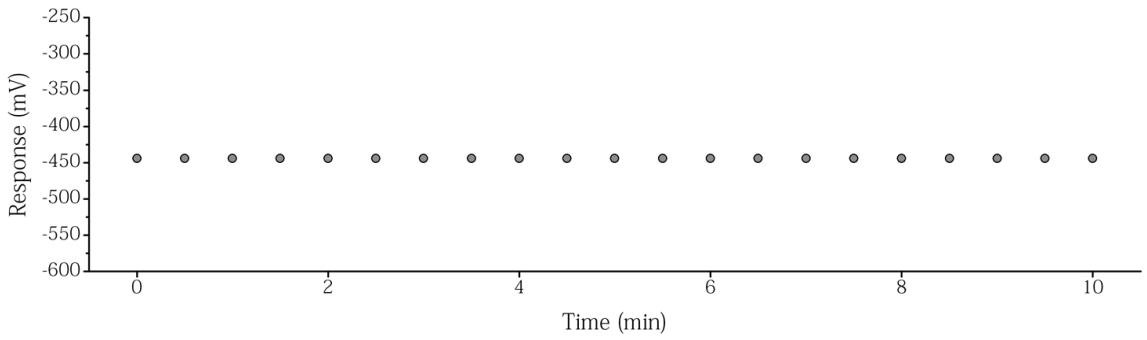


Figure 3-32: Degradation curve of galF (131.1 mM in reaction buffer) as measured by the fluoride-selective electrode.

As with the 96-well plate assay, LacZ E537S enzyme was taken as a first test. First, the enzyme was analysed to see if the enzyme reacted with just galF. There are two possible outcomes for this reaction. One, the enzyme increases the hydrolytic rate of galF (**1**) in aqueous solutions. Two, the enzyme can catalyse the formation of homopolysaccharides, although this is dispreferred due to the architecture of the active site.[250,269] For this, the enzyme stock solution (1.0 mg mL^{-1}, 0.1 mL) was mixed with a series of stock solutions of galF (**1**) (0-10 mM, 1.9 mL) and the voltage of the reaction was measured as for the background degradation, 10 min at 30 s intervals, and the measurements performed in triplicate. Disappointingly, every measurement returned no observable change in the concentration of fluoride. This suggested that the enzyme may be too slow a catalyst for any changes to be observed. The concentrations of galF were well below the inhibition range of galactose and it is known that the shallow site of LacZ is designed for binding carbohydrates with high specificity.[269] Regardless, perhaps glc (**6**) or glc-pNP (**4**) would be a better glycosyl acceptor due to the extra stability afforded it in the shallow site through interactions with Asn102, Ser796, and Glu797. The experiment was then re-designed in an effort to investigate this. With the final concentration of the enzyme in each reaction tube identical to the previous experiment (0.45 μM), the concentration of galF (**1**) was consistent across all samples as well (1.0 mM). The variable here was the concentration of either glc (**6**) or glc-pNP (**4**) (1.0, 2.0, and 4.0 mM) and the experiment repeated. Sadly, these experiments also yielded no observable concentration in the concentration of fluoride. Because of these experiments, it was determined that the measurement of fluoride in galactosynthase reactions *via* an ion-selective electrode was not sensitive enough for our purposes. As a result, more sensitive methods were investigated.

3.4.3 NMR Studies

Having used NMR to characterise synthesised molecules, reference spectra were available for the starting materials, the degradation product, and the two expected disaccharide products. Determination of enzyme kinetics *via* ^{1}H NMR spectroscopy was attempted. However, during the synthesis of galF (**1**), it was observed that deprotection of the per-*O*-acetylated parent molecule (**33**) resulted in product degradation, which raised the question: is galF stable enough for the galactosynthase reactions? To answer this question, a series of ^{1}H NMR kinetic experiments were undertaken. It was decided to include an internal standard (IS) to allow for quantification of degradation rates. The samples were measured at regular intervals over 2 days at 25 °C; as per conditions outlined by the *Withers* group.[1] The desired criteria for an internal standard was that the signal of the standard must not overlap with any of the peaks in the sugar range of δ (ppm) 3.2-5.9. The first standard chosen was maleic acid which contains two sp^2-hybridised protons which resonate at δ (ppm) 6.41. This signal lay outside our range and was therefore investigated as an internal standard (Figure 3-33).

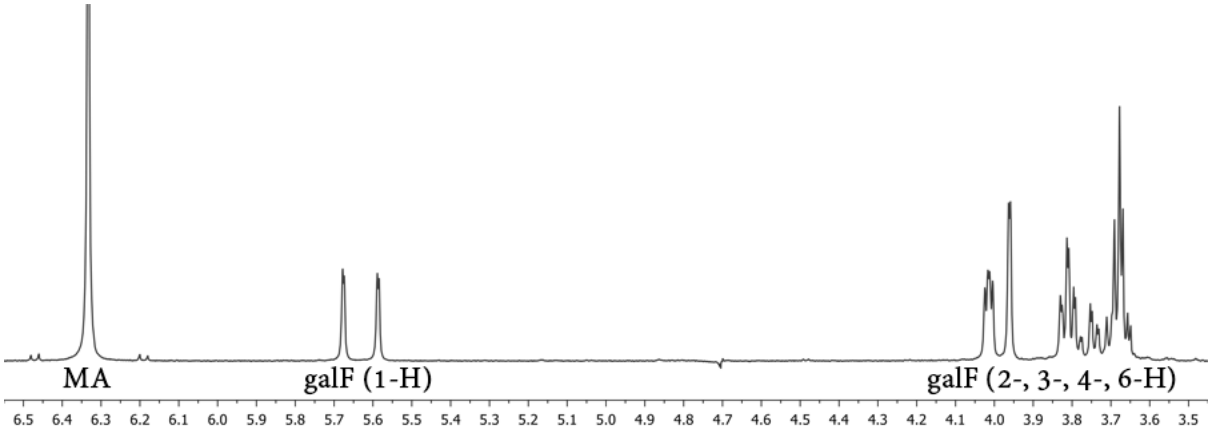


Figure 3-33: ^{1}H NMR of maleic acid (MA) and galF (**1**) in D_2O using a solvent suppressed method.

As an internal standard, the maleic acid worked perfectly as the peak was well-defined and distinct from those of galF (**1**). What raised concerns, was the sigmoidal nature of the curve which suggested a mixed reaction type. Furthermore, the degradation of galF proceeded within the span of 18 hours which conflicted with the claim by the *Withers* group[1] that the reaction was allowed to proceed over 48 hours. This suggested there were more factors in play and to this we hypothesised that the internal standard was interfering somehow with the rate of degradation; possibly by facilitating the degradation as a result of its acidity. Therefore, investigation of an alternative internal standard was undertaken. Dimethylsulfone (DMS, **34**) was chosen as it has no acidic properties, is readily soluble in water, and has a resonance outside the sugar range, δ (ppm) 3.15. Furthermore, as this signal is the result of geminal methyl groups, the 6-proton resonance meant that less DMS could be used to obtain a strong peak in the ^{1}H NMR spectrum (Figure 3-34).

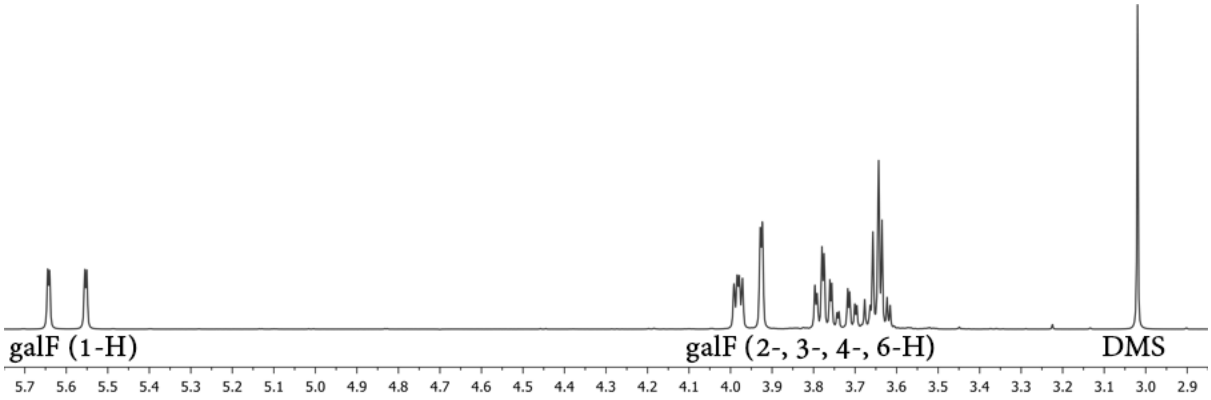


Figure 3-34: ^{1}H NMR of dimethylsulfone (DMS) and galF (#) in D_2O using a solvent suppressed method.

The use of DMS as an internal standard did decrease the rate of degradation when compared to the results obtained using the maleic acid IS, but the complete degradation of galF still occurred well within the 48 hour window as reported for *Jakeman* and *Withers*.[1] Furthermore, the degradation curve was still sigmoidal. This was attributed to one major factor; pH. As the degradation reaction is S_N1-type reaction in which the fluoride is displaced by a molecule of water generating the natural carbohydrate and HF, it would be expected that the degradation follows an exponential decay curve. In addition, the degradation rate appeared dependent upon the identity of the IS used. With this in mind, it was decided to measure the degradation rate of galF (**1**) in deuterated phosphate buffer at pH 7.5 in the presence of DMS (**34**) (Figure 3-35).

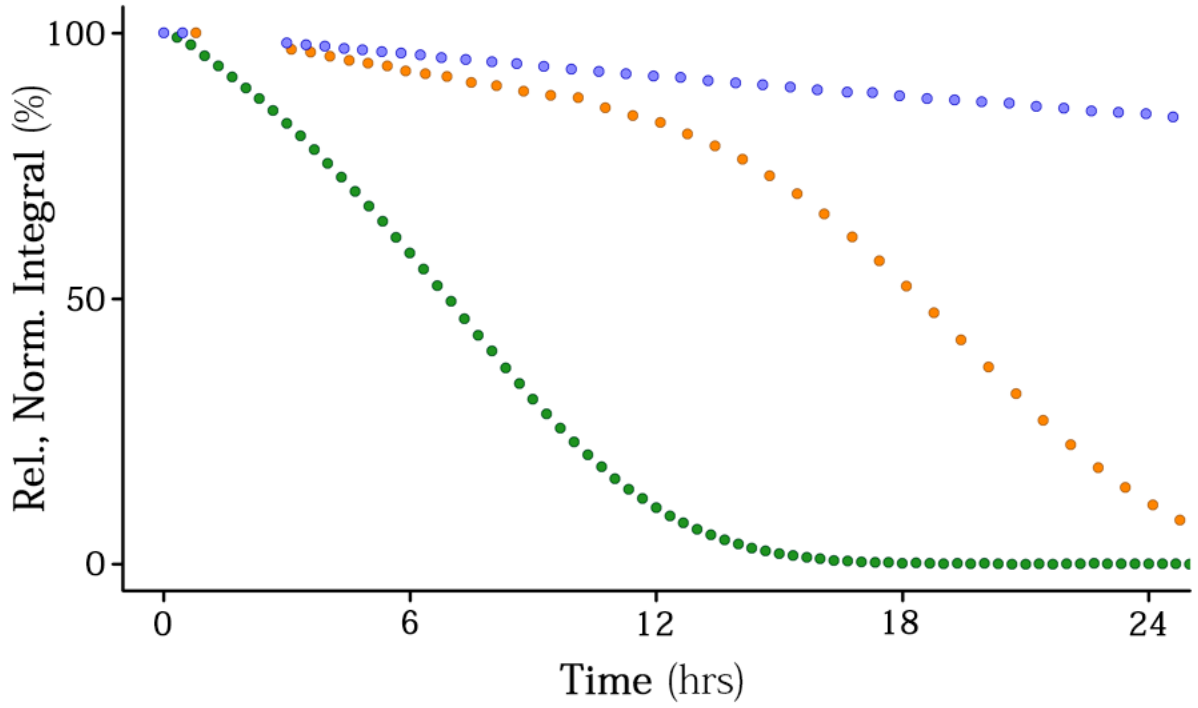


Figure 3-35: Comparison of the relative degradation rates of galF (**1**) over 24 hrs in the presence of the internal standards in both unbuffered – MA (green) and DMS (orange) – and buffered – DMS in a 200 mm phosphate buffer at pH 7.5 (blue) – systems.

This showed that the reaction rate of degradation follows the expected first-order, meaning that the degradation of galF is concentration-dependent. Additionally, the half-life of galF in our pH 7.5-buffered system was 101.4 ± 0.4 hours (R^2 = 0.999); demonstrating the suitability of the substrate for synthesis. Concerning, was the fact that the enzymes themselves would need to be quite active in order to return a significant amount of product before the competing autohydrolysis reaction could consume the substrate. In order to properly demonstrate that the degradation of galF is truly pH-dependent, an experiment was performed wherein galF was dissolved in phosphate buffers in the near-neutral pH range of 6.0-8.0 with DMS again being used as the IS (Figure 3-36). Gratifyingly, the hypothesis was shown to be correct as it was observed that galF was most stable at pH 7.0 and deviations from neutral pH accelerated the autohydrolysis of the compound. These degradation experiments also allowed us to determine the lower limit of detection: 2.5 mM.

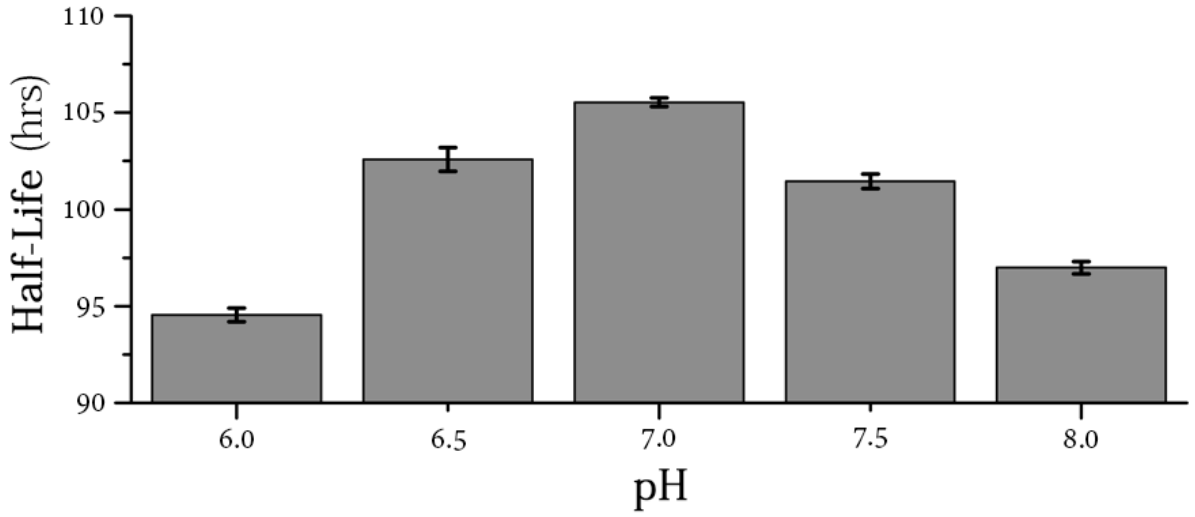


Figure 3-36: The half-life of galF (**1**) in 200 mM phosphate buffer at different pH values.

With the stability of galF (**1**) had been determined, we used this information to construct a calibration curve. The curve was then trialled for use in the enzyme-catalysed reactions. Using LacZ E537S (25 μM), the reaction was performed and the increase in catalysis was measured as determined by deviation of the degradation rate from the known background rate over 5 hours (Figure 3-37). It was observed that there is a significant deviation in the rate of galF conversion in the synthase-catalysed reaction which equated to this enzyme having a specific activity of 0.58 ± 0.05 U mg^{-1}. This result was exciting as, to the best of our knowledge, this was the first characterisation of the activity of this galactosynthase.

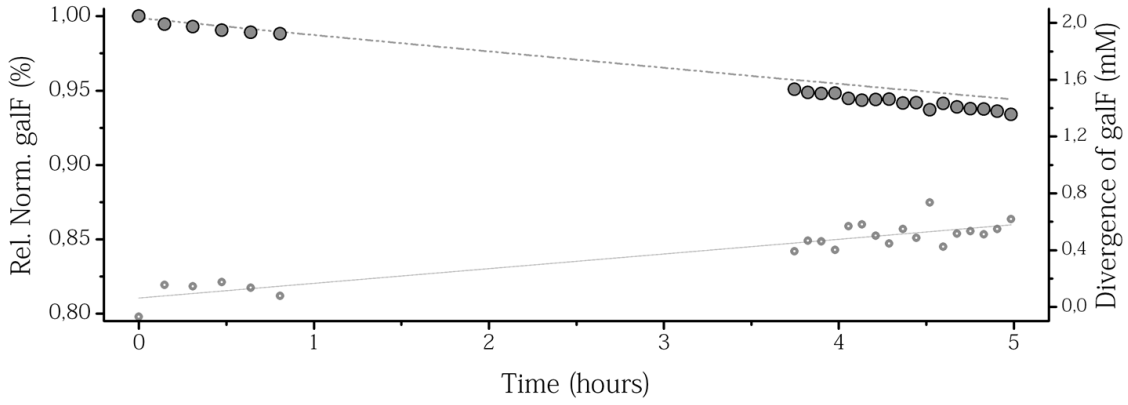


Figure 3-37: Double Y-axis graph of the relative amount of galF in the reaction as normalised to DMS. The left Y-axis compares the background degradation rate of galF (dotted line) against the degradation of galF in the enzyme-catalysed reaction (filled circles). The right Y-axis quantifies this deviation (hollow circles) in mM of galF converted by the enzyme against the amount of galF, which would naturally be degraded without the enzyme present.

When the experiment was repeated using a higher enzyme concentration (80.9 μM), a lower overall activity was observed (0.07 ± 0.01 U mg^{-1}). Naturally, this led us to question of reproducibility using this NMR-based method. The reaction was therefore repeated again using the lysate supernatant of the enzyme in case the enzyme lost activity during purification. This experiment was performed in duplicate from the same batch of cells, so as to ensure the ratio of the enzyme is statistically identical in each reaction. The results of these experiments were in disagreement with one another (0.00 ± 0.00 U mg^{-1} and 3.35 ± 0.20 U mg^{-1}, respectively) and this method of determination of the activity of the glycosynthases was reevaluated.

^{19}F NMR was also considered as an alternative to ^{1}H NMR, as the spectra were cleaner due to the presence of less different fluorine atoms. The biggest problem encountered when attempting to apply ^{19}F NMR in an aqueous solution, was finding a suitable internal standard which is inert under these conditions and was soluble in water. As it is known that the conjugate base of a strong acid is a weak base, it was thought that using the salt of a fluoride-containing superacid, such as sodium trifluoroacetate (Na-TFA), could solve the problem. It was hoped that this would not alter the pH so significantly as to have noticeable effects on the degradation of galF. This compound was then trialled as an internal standard against an identical sample containing no IS (Figure 3-38).

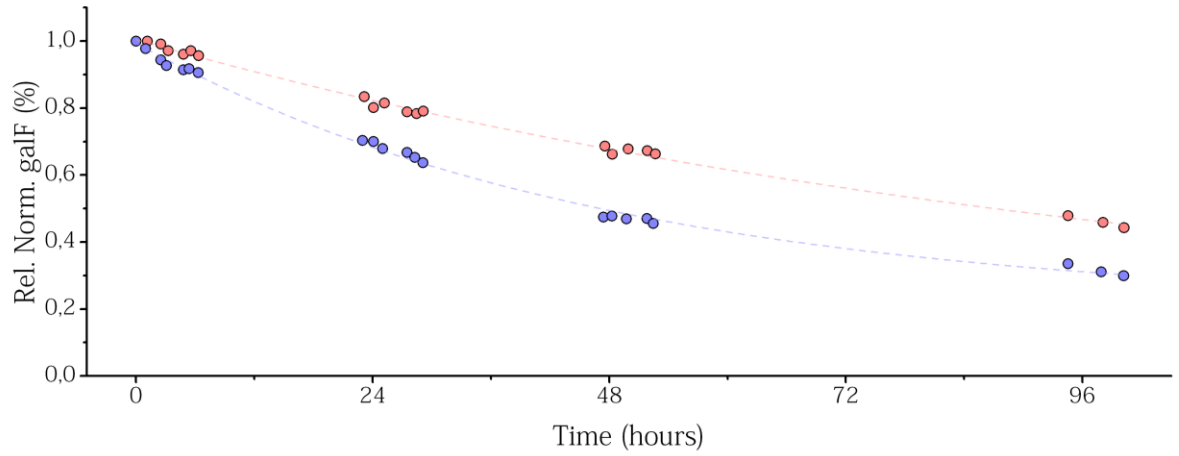


Figure 3-38: ^{19}F NMR degradation kinetics of galF in the absence of an internal standard (red) and in the presence of Na-TFA as an internal standard (blue). Reaction conditions: galF (20.0 mg) in deuterated reaction buffer (pH 7.5, 500 µL). In the case of the IS-positive sample, 1.0 mg of Na-TFA was added to the reaction tube prior to measurement.

Although the use of Na-TFA as an internal standard in ^{19}F NMR kinetics did appear possible, the rate of degradation of the galF under these conditions was significantly influenced with the $t_{½}$ was decreased by more than a day ($t_{½}$ = 86.6 ± 3.6 hrs against 57.6 ± 2.4 hrs) in the presence of the trifluoroacetate anion. As was earlier demonstrated, the degradation rate of galF is influenced by the pH and it was presumed that the presence of this increased the pH enough to alter the reactivity of the glycosyl fluoride. It is worth mentioning here though that the internal standard was added at the time of the reaction start and the pH of the buffer was unadjusted after the addition of Na-TFA.

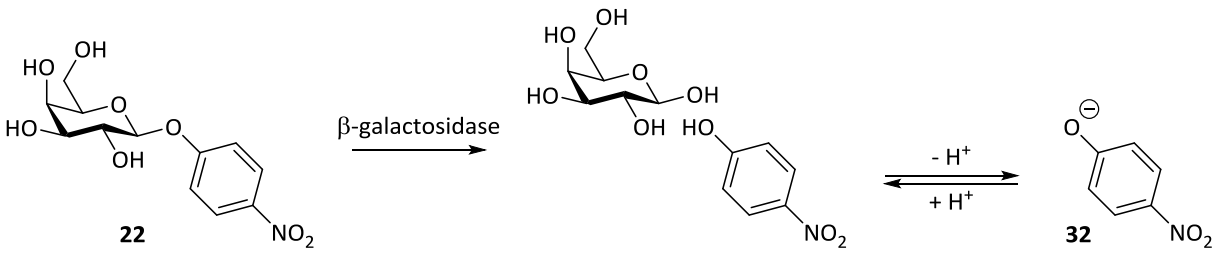
3.5 Carbohydrate Synthesis

There is a large range of carbohydrates available from commercial sources, but due to the typically restrictive costs involved in the mass-production of non-naturally occurring carbohydrates, the costs involved in purchasing these starting materials can be prohibitive. Therefore, within the scope of this thesis, methods were devised in order to obtain the starting materials required for the exploration of the substrate range and scope of the developed glycosynthases. However, once obtained, these substrates were to be used in glycosynthase reactions and the reaction products subsequently analysed and purified.

3.5.1 Synthesis of the *p*-Nitrophenyl Glycosides

In order to measure the hydrolytic activity of the envisaged library of glycosidases, the first test of their successful expression is typically a hydrolytic screening to examine the activity of the wild type enzyme. Through this, we aim to gather information of possible glycosyl donor substrates available to the enzyme as it is known that point mutation of the nucleophilic residue does not affect the overall architecture of the active site.[198,307] This was envisioned to be achieved through the use of the

p-nitrophenyl group as it is deprotonated in basic conditions due to the electron-withdrawing nitro group and the resulting phenolate (**32**) is a vivid yellow colour (Scheme 3-2). This means that the concentration of phenolate in a sample can be easily detected using colourimetric methods. Furthermore, it was envisaged to use these *p*-nitrophenyl glycosides as glycosyl acceptor molecules as the UV-active phenyl ring is easily detectable, even in complex mixtures, allowing for easy detection *via* UV spectrometry.



Scheme 3-2: Hydrolysis of the *p*-nitrophenyl glycoside by a glycosidase yields the sugar and *p*-nitrophenol. In basic conditions, the phenol is deprotonated to the *p*-nitrophenolate (**32**) anion which absorbs strongly at λ = 405 nm giving it a vivid yellow appearance.

A library of *p*-nitrophenyl glycosides was generated synthetically starting from either the commercially available sugar or per-acetylated glycoside as appropriate. As the glycosidases to be examined were all β-glycosidases, the library consisted solely of equatorially-configured *p*-nitrophenyl glycosides, which are β-configured, with the exception of the α-arabinopyranoside. In order to achieve this, a synthetic method was required in order to force the reaction to overcome the anomeric effect, which prefers the axial anomer. It was hoped that this could be achieved *via* use of the per-acetylated carbohydrate starting materials by taking advantage of the neighbouring group directing effect of the acetate group (Figure 3-39).

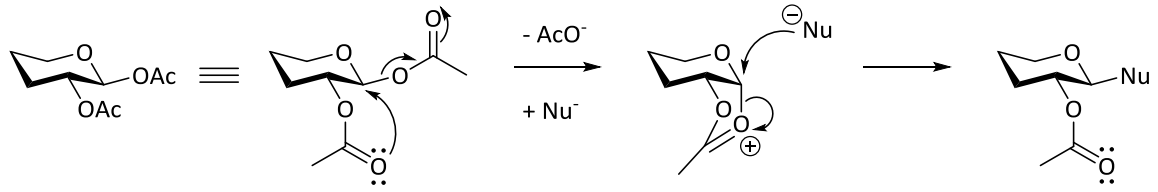
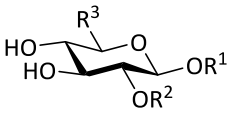
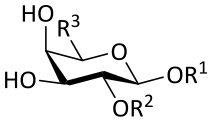


Figure 3-39: A simplified version of the neighbouring group effect wherein the 2-acetate blocks the axial site from nucleophilic attack by the incoming nucleophile, forcing the reaction to prefer the formation of the β-anomer. The remaining hydroxyl groups of the carbohydrate are omitted for clarity.

As previously stated, a library of β-glycosides was required and a selection of carbohydrates was therefore desired for assay purposes as hydrolytic substrates, reference compounds, or glycosidic acceptors (Figure 3-40).



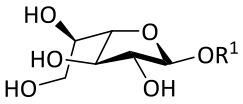
4: R^1=pNP, R^2=H, R^3=CH_2OH



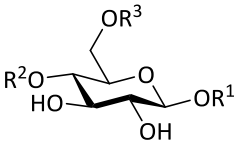
22: R^1=pNP, R^2=H, R^3=CH_2OH
25: R^1=pNP, R^2=H, R^3=CH_3

27: R^1=pNP, R^2=H, R^3=H
35: R^1=pNP, R^2=NHAc, R^3=CH_2OH

26: R^1=pNP, R^2=H, R^3=H
23: R^1=pNP, R^2=NHAc, R^3=CH_2OH



24: R^1=pNP

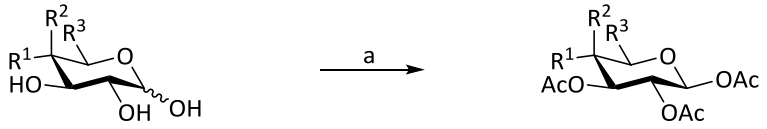


36: R^1=pNP, R^2=β-gal, R^3=H
5: R^1=pNP, R^2=H, R^3=β-gal

Figure 3-40: Desired *p*-nitrophenyl glycosides desired for testing of hydrolytic activity of enzymes, reference compounds for reaction mixture analysis, and as glycosidic acceptors.

For those carbohydrates which were both commercially available and financially viable, these were ordered as the unprotected sugar (xyl, gal, ara, lac). Per-*O*-acetylated-β-D-glucopyranoside was readily available from commercial suppliers and the synthesis of compound **4** was commenced from this advanced starting material. Initially, per-*O*-acetylation of the unprotected sugars was performed overnight using Ac_2O in pyridine, which worked well for small-scale reactions, up to 5 g. However, for larger scale reactions, the reaction workup became too unwieldy due to the large volumes of solvent, and particularly $CuSO_{4\ (aq)}$ required for the removal of pyridine from the organic phase, as per the protocol outlined by *Hoffmann et al.*[344] A method of per-*O*-acetylation for scales over 5 g was sought, which required no metals. In 2013, *Percec* and co-workers[345] required per-*O*-acetylated galactose for the generation of complex glycan structures and utilised a method which involved the portion-wise addition of galactose to a refluxing solution of NaOAc in Ac_2O (Scheme 3-3). As this method does not use pyridine as a solvent and required no Cu^{2+} to complex the pyridine to remove it to the aqueous phase, it was trialled as an alternative method for large scale per-*O*-acetylations. This method proved highly effective and resulted in complete conversion of the galactose within 15 minutes. Once obtaining the crude product, it was observed that a complex mix of products was present and purification *via* column chromatography was unable to separate the anomeric mixture. As the commercially-obtained, anomerically-pure per-acetylated glucose was crystalline, it was attempted to selectively crystallise the β-anomer out of solution. Following the method of *Wolfrom* and *Groebke,*[346] after removing the extraction solvent under reduced pressure, the mixture was triturated using 95% ethanol, which resulted in the β-anomer precipitating out of solution, while the α-anomer remained soluble, as evidenced by the ^{1}H NMR spectrum of the crystalline material. This

method was modified for xylose, as both anomers remained soluble in ethanol. Instead, the β-product was triturated instead using methanol.



7: R^1=H, R^2=OH, R^3=CH_2OH
37: R^1=H, R^2=OH, R^3=CH_3
38: R^1=H, R^2=OH, R^3=H
6: R^1=OH, R^2=H, R^3=CH_2OH
39: R^1=OH, R^2=H, R^3=H
2: R^1=O-β-gal, R^2=H, R^3=CH_2OH

41: R^1=H, R^2=OAc, R^3=CH_2OAc (*48%*)
42: R^1=H, R^2=OAc, R^3=CH_3 (*45%*)
43: R^1=H, R^2=OAc, R^3=H (*quant., α/β mix)*
44: R^1=OAc, R^2=H, R^3=CH_2OAc (*59%*)
45: R^1=OAc, R^2=H, R^3=H (*52%*)
46: R^1=O-β-(Ac_4)gal, R^2=H, R^3=CH_2OAc (*72%*)

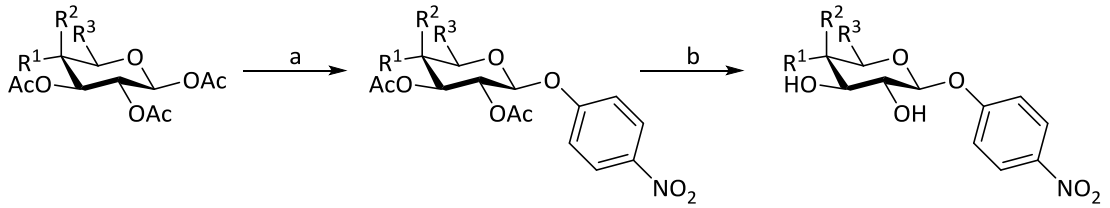
Scheme 3-3: The per-*O*-acetylated, β-configured glycosides were generated from the corresponding unprotected saccharides. a) NaOAc (1.1 eq.), Ac_2O, refl., 15 min.

Having solved the problem of anomeric separation, the next step was to convert the per-acetates to their corresponding *p*-nitrophenyl glycosides. In all cases except arabinose, this is the β-anomer. Using $BF_3{\cdot}OEt_2$ as a Lewis acid to form glycosidic linkages with phenols is not a new strategy, having first been pioneered by *Bretschneider* and *Beran* in 1949[347] using per-acetylated β-D-glucopyranose. This method was then investigated further in 1990 by *Yamaguchi* and co-workers[348] who used tetra-*O*-acetyl-α-glucopyranosyl fluoride as the starting reagent and coupled this to a number of different phenols using $BF_3{\cdot}OEt_2$. In their experiments, it was revealed that the use of $BF_3{\cdot}OEt_2$ as a Lewis acid in the glycosylation reaction, the β-anomer was preferred when the reaction was performed in the presence of 1,1,3,3-tetramethylguanidine, an organic base. When performed in the absence of this base, the α-anomer dominated. Our desired 4-nitrophenol is a poor nucleophile as a result of the strongly electron-withdrawing nitro functionality in the *para* position to the hydroxyl group, thus reducing the electron density and subsequently nucleophilicity of the hydroxyl moiety. As a result, the reaction using 4-nitrophenol as a glycoside acceptor is retarded when compared to other phenolic nucleophiles, as observed by *Lee et al.*[349] Analysis of the reaction after 3 days *via* TLC showed two distinct spots when run in an EtOAc/PE system and appeared as though all starting material had been reacted. The reaction was then worked up and the products separated using column chromatography. However, analysis of the isolated compounds by ^{1}H NMR showed impurities which were unable to be removed by our chromatographic method. An elution system screen was run using a number of common solvents and to our surprise, when Et_2O was the major component in the mobile phase, a third spot was visible. Column chromatography was then implemented and the third spot was revealed to be unreacted starting material. It should be noted, that after 3 days of reaction time, the per-acetylated starting material was not fully converted to the

phenyl glycoside. Extension of this reaction time did not increase the conversion amount nor the yields obtained.

It was observed that the desired β-phenyl compound was crystalline and inspired by our previous success of anomeric separation *via* trituration, we also applied this method to this compound. Fortunately, selective recrystallisation of the β-product was possible from hot ethanol. However, considering that the starting material would co-crystallise in ethanol, the crude mixture was first triturated in Et_2O to remove the per-acetylated glycoside prior to recrystallisation. It was on occasion observed that the β-glycoside did not triturate from Et_2O as previously. In these cases, analysis of the ethereal solution revealed high concentrations of free *p*-nitrophenol still in solution. These mixtures were rewashed with saturated hydrogen carbonate solution to remove the excess phenol and subsequently triturated as had been previously done. This method of trituration and recrystallisation was only applicable to the galactopyranoside and the glucopyranoside. The pentoses, disaccharides, and D-fucopyranoside required separation *via* chromatographic methods.

With the successful purification of the β-glycoside, the tetra-*O*-acetylated compounds were deacetylated under *Zemplén* conditions and were obtained in quantitative yields (Scheme 3-4).



41: R^1=H, R^2=OAc, R^3=CH_2OAc
42: R^1=H, R^2=OAc, R^3=CH_3
43: R^1=H, R^2=OAc, R^3=H
44: R^1=OAc, R^2=H, R^3=CH_2OAc
45: R^1=OAc, R^2=H, R^3=H
46: R^1=O-β-(Ac_4)gal, R^2=H, R^3=CH_2OAc

47: R^1=H, R^2=OAc, R^3=CH_2OAc (*88%*)
48: R^1=H, R^2=OAc, R^3=CH_3 (*19%*)
49: R^1=H, R^2=OAc, R^3=H (*3%*)
50: R^1=OAc, R^2=H, R^3=CH_2OAc (*77%*)
51: R^1=OAc, R^2=H, R^3=H (*29%*)
52: R^1=O-β-(Ac_4)gal, R^2=H, R^3=CH_2OAc (*27%*)

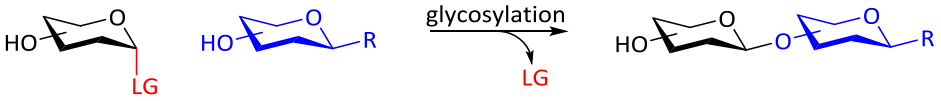
22: R^1=H, R^2=OH, R^3=CH_2OH (*quant.*)
25: R^1=H, R^2=OH, R^3=CH_3 (*quant.*)
26: R^1=H, R^2=OH, R^3=H (*quant.*)
4: R^1=OH, R^2=H, R^3=CH_2OH (*quant.*)
27: R^1=OH, R^2=H, R^3=H (*quant.*)
36: R^1=O-β-gal, R^2=H, R^3=CH_2OH

Scheme 3-4: Reaction scheme for the generation of 4-nitrophenyl β-glycosides* from their per-acetylated precursors. Reaction conditions: a) pNP (1.5 eq.), $BF_3{\cdot}OEt_2$ (2.5 eq.), Et_3N (0.5 eq.), anh. CH_2Cl_2, room temperature, 3 d (1-67%); b) NaOMe/anh. MeOH (0.5 M), anh. MeOH/CH_2Cl_2 (2:1), r.t, 16 h (*quant.*). *in the case of arabinopyranose, the final conformation is α-L-.

3.5.2 Synthesis of Glycosyl Fluorides as Glycosyl Donors

Glycosylation requires two components: a glycosyl acceptor, which is usually in the form of a nucleophilic alcohol, amine, or thiol; and a glycosyl donor, which takes the form of an anomerically-protected glycoside where the protecting group is a suitable leaving group (Scheme 3-5). For the purposes of our research, we focussed on the use of glycosyl fluorides as our donor

substrates. Due to the small size of the fluorine atom, this was ideal as a leaving group for use with the glycosynthase biocatalysts because it is able to fit into the pocket easily and the steric interactions between the axial fluorine and the two hydrogen atoms under the ring would be minimal. Furthermore, as the fluoride anion is stable in aqueous solutions, it is a useful leaving group.



Scheme 3-5: An overall schematic of the glycosylation reaction wherein the anomeric functional group (red) also acts as a leaving group, allowing coupling of the glycosyl donor (black) to the acceptor (blue).

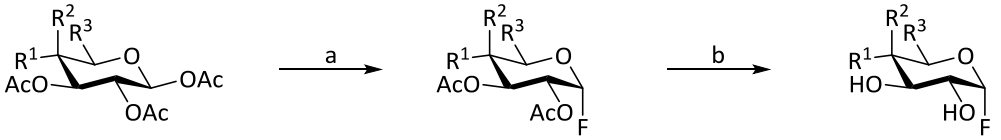
In a similar fashion to the *p*-nitrophenyl glycosides, it was decided to use the per-*O*-acetylated glycosides as the starting material for the fluoridation reaction. As the optimal conditions for the per-acetylation of the sugars was determined when optimising the reaction conditions for the generation of the β-phenyl glycosides, we already had the ideal conditions at hand. Thus, the next step was the anomeric fluoridation of the sugars. The conversion of per-*O*-acetylated glycosides to their anomerically-fluoridated counterparts using hydrofluoric acid is a technique which dates back nearly 100 years to experiments first performed by *Brauns* using freshly-distilled HF.[350] Recently, *Zacharia* and *Hayashi*[69] used 70% HF in pyridine at room temperature to achieve the fluoridation of per-acetylated galactopyranoside. They found this reaction to go to completion with an α/β ratio of 96:4. As the axially-configured α-anomer was our target due to the use of galactosynthases expected to invert the configuration at the anomeric centre, this result was ideal for our purposes. A range of per-*O*-acetylated glycosides were fluoridated at this position and the anomeric product ratio is given in Table 3-11. Following fluoridation, the products were separated using column chromatography.

Table 3-11: Anomeric ratio of the fluoridation reaction. *for L-arabinopyranoside, the β-anomer is the axially- and the α is equatorially-configured.

Compound		R^1	R^2	R^3	α:β
33	D-gal	H	OAc	CH_2OAc	10 : 90
53	D-fuc	H	OAc	CH_3	24 : 76
54	L-ara	H	OAc	H	71 : 29*
55	D-glc	OAc	H	CH_2OAc	16 : 84
56	D-xyl	OAc	H	H	17 : 83
57	D-lac	*O*-β-$(Ac)_4$gal	H	CH_2OAc	6 : 94

For the synthesis of reference compounds, the deacetylated sugar fluorides were not desirable as the presence of free hydroxyl groups would result in undesired polymer formation. Furthermore, as the

per-*O*-acetylated glycosyl fluorides were stable, these compounds were stored at room temperature until deprotection was required. It was expected that deprotection using the previously implemented *Zemplén* conditions using NaOMe in MeOH would proceed smoothly. However, after leaving the reaction for the deprotection of 2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl fluoride (**33**) overnight, it was observed that no fluoridated product could be obtained as observed by ^{1}H NMR. The reaction conditions were modified after a procedure outlined by *Steinmann* and co-workers[351] wherein they performed the reaction at 0 °C and for only 2 hours. The solvent, dichloromethane, was then removed under reduced pressure at 0 °C and the remaining solvent, MeOH, was then flash frozen using $N_{2\,(l)}$ and removed *via* lyophilisation. The resulting white powder was then stored at -20 °C and the product appeared stable under these conditions for up to 3 months as observed by ^{1}H NMR (Scheme 3-6).



47: R^1=H, R^2=OAc, R^3=CH_2OAc
48: R^1=H, R^2=OAc, R^3=CH_3
49: R^1=H, R^2=OAc, R^3=H
50: R^1=OAc, R^2=H, R^3=CH_2OAc
51: R^1=OAc, R^2=H, R^3=H
52: R^1=O-β-(Ac_4)gal, R^2=H, R^3=CH_2OAc

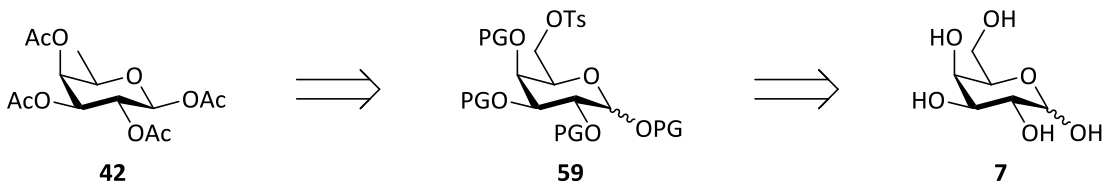
33a: R^1=H, R^2=OAc, R^3=CH_2OAc (*56%*)
53a: R^1=H, R^2=OAc, R^3=CH_3 (*26%*)
54b: R^1=H, R^2=OAc, R^3=H (*25%*)
55a: R^1=OAc, R^2=H, R^3=CH_2OAc (*91%*)
56a: R^1=OAc, R^2=H, R^3=H (*43%*)
57a: R^1=O-β-(Ac_4)gal, R^2=H, R^3=CH_2OAc (*36%*)

1: R^1=H, R^2=OH, R^3=CH_2OH (*quant.*)
15: R^1=H, R^2=OH, R^3=CH_3
16: R^1=H, R^2=OH, R^3=H
17: R^1=OH, R^2=H, R^3=CH_2OH
18: R^1=OH, R^2=H, R^3=H
58: R^1=O-β-gal, R^2=H, R^3=CH_2OH

Scheme 3-6: Synthesis of the axially-configured glycosyl fluorides. Reaction conditions: a) 70% HF/py (1 mL $mmol^{-1}$), N_2, 0 °C → room temperature, 4 h (25-42%); b) NaOMe/MeOH (0.5 M), anh. CH_2Cl_2/MeOH (1:2), 0 °C, 2 h (*quant.*).

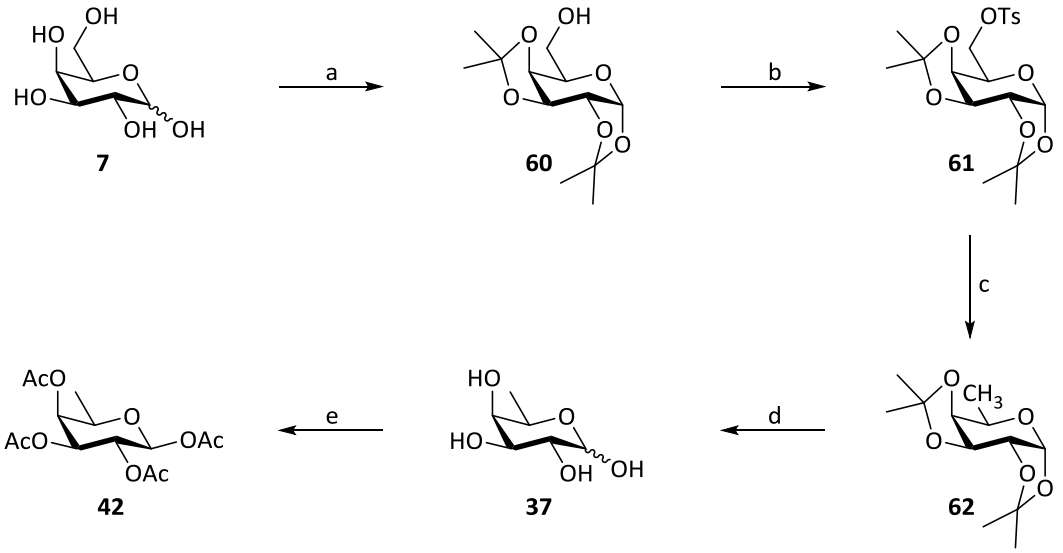
3.5.3 Synthesis of 1,2,3,4-Tetra-*O*-acetyl-β-D-fucopyranose

At over 100 € g^{-1} (*Sigma Aldrich*, Nov 2016), the testing of fucose as a substrate for expressed glycohydrolases and glycosynthases was expensive considering that the reactions for the generation of fluoride or *p*-nitrophenyl derivatives were relatively low yielding. Instead, and with the assistance of *Vera Ophoven* (IBOC), we decided to access this group of compounds by synthesis from 6-hydroxyfucose, more commonly known as galactose. Starting from the readily available and inexpensive D-galactose (**7**), the plan was to generate the per-acetylated D-fucopyranoside (**42**) *via* the 6-*O*-tosylated galactose intermediate (**59**) (Scheme 3-7).



Scheme 3-7: Proposed synthetic route towards per-acetylated D-fucopyranoside (**42**) from D-galactopyranoside (**7**) *via* a 6-tosylated intermediate (**59**).

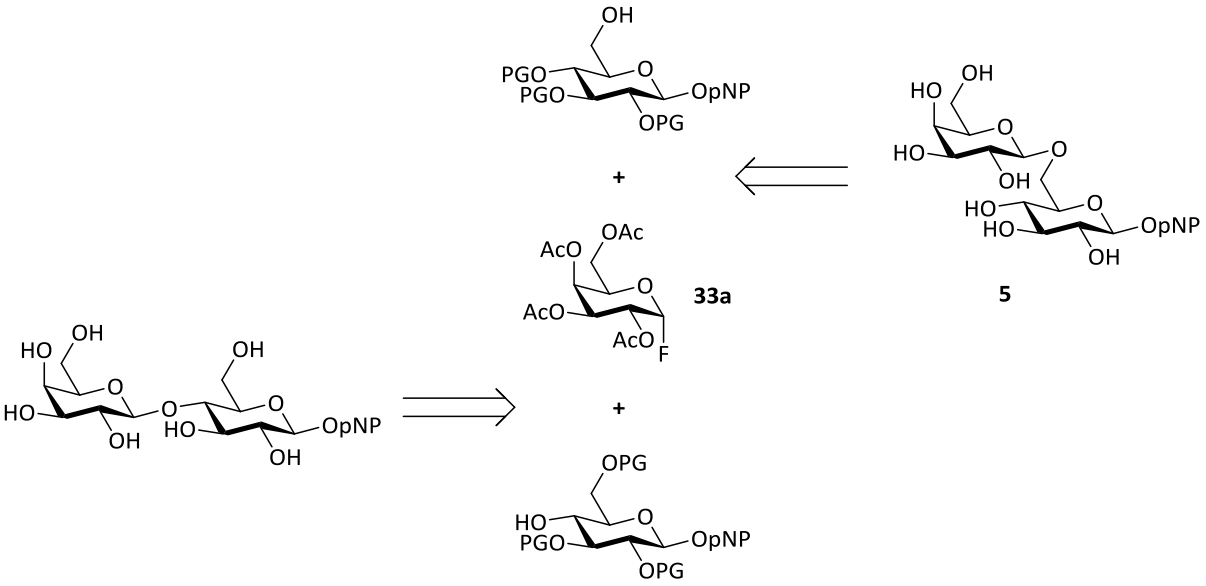
It was planned that by selectively tosylating the 6-hydroxyl group of galactose and subsequent reduction of this moiety would generate a 5-methyltetrahydropyran scaffold, rather than the 5-hydroxymethyltetrahydropyran of the galactose. To achieve this, selective protection of the secondary hydroxyl groups of galactose would need to be implemented before tosylation can be performed. To this end, it was decided to proceed using the acetonide acetal which prefers 1,2-diols over 1,3- and 1,4-systems.[352-354] It was expected that this selectivity would lead to the formation of the 1,2:3,4-di-*O*-isopropylidene compound (**60**) leaving the primary alcohol free. Using conditions outlined by *Sennari* and co-workers,[354] galactose (**7**) was successfully diisopropylidenated to return the expected α-anomer in quantitative yields as evidenced by the 3 peaks with an integral of 12 protons in the range 1.3-1.6 ppm and the $^{3}J_{1,2}$ coupling constant of 1.6 Hz (Scheme 3-8). With the secondary alcohols protected, C-6 could now be dehydroxylated *via* a 6-*O*-tosyl intermediate. The tosylation of the primary 6-hydroxyl group was carried out in a straight-forward manner using a catalytic amount of DMAP to increase the electrophilicity of the tosyl group and the 6-tosyl galactoside (**61**) was returned in good yields. Subsequently, compound **61** was reduced to **62** using a 1M solution of lithium triethylborohydride. The use of $LiBHEt_3$, or 'super hydride', as outlined by *Weigel* and *Liu*,[355] increased the yield of deoxygenation from 67% over two steps using NaI then $NaBH_4$, to 88% in a single transformation. Unfortunately, *Weigel* and *Liu* did not report specific reaction conditions for the conversion using super hydride and thus reaction conditions were adapted from papers by *Denmark et al.*[356] and *Taber et al.*[357] Despite claims by both of the authors that the reaction should proceed rapidly at room temperature within a matter of hours, it was found that overnight reaction at room temperature was insufficient for the reaction to proceed. A number of experiments into the reactivity of $LiBHEt_3$ have shown that the reaction takes place at a slower reaction in carbohydrate compounds[358,359] compared to other, less polar alcohols such as those used by *Denmark* and *Taber*. It has also been shown that the reaction takes place at a slower rate in THF compared to other solvents, such as 1,4-dioxane.[360] It was decided to reflux the reaction for 1 h subsequent to the overnight, room temperature reaction, which returned compound **62** in an excellent yield of 94%. Following the reduction of C-6, the isopropylidene protecting groups were removed to return D-fucose (**37**) in quantitative yields. The carbohydrate **37** was then per-acetylated using Ac_2O in pyridine to return the desired per-*O*-acetylated fucose compound, **42**, in an overall yield of 28% over five steps.



Scheme 3-8: Synthetic approach to the generation of per-*O*-acetylated β-D-fucopyranoside (**42**). Reaction conditions: a) $ZnCl_2$ (1.5 eq.), H_2SO_4 (3.3 ‰), acetone, room temperature, 16h, *quant.*; b) TsCl (1.5 eq.), DMAP (0.1 eq.), py/CH_2Cl_2, room temperature, 5 h, 66%; c) $LiBHEt_3$ (6.0 eq.), THF, refl., 16 h, 94%; d) $AcOH_{(aq)}$ (80%), 80 °C, 6 h, *quant.*; e) Ac_2O (6.0 eq.), py, r.t, 16 h, 45%.

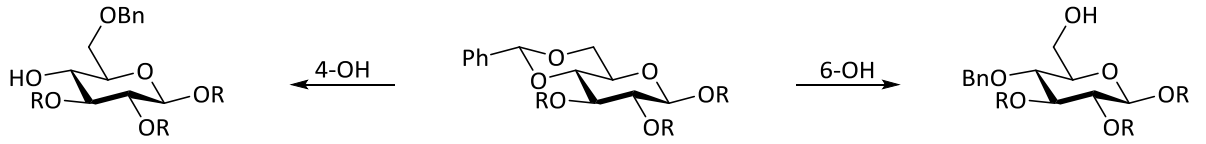
3.5.4 Synthesis of *p*-Nitrophenyl β-D-Allolactoside

The glycosynthases chosen were created from galactohydrolases which hydrolyse lactose (**2**) and can perform the transglycosylation reaction to generate allolactose (**3**). Therefore, reference compounds were required in order to analyse galactosynthase reactions properly and to assist in the development of methods for reaction product analysis. The proposed synthesis of **3** was therefore envisaged using the building blocks already available to us, namely the tetra-*O*-acetyl-α-D-galactopyranosyl fluoride (**33a**) and a 4-nitrophenyl tri-*O*-protected-β-D-glucopyranoside (Scheme 3-9). The glycosylation was expected to proceed *via* the existing $BF_3{\cdot}OEt_2$/amine method. A protocol published by *Wang* and co-workers[361] inspired the creation of a universal glucoside to create the series of possible β-disaccharides formed between a galactosyl donor and the glucoside acceptor. The use of this method would allow the synthesis of new references in the event that our galactosynthases were able to use different glycosyl donor molecules, such as D-galactofuranose, D-fucopyranose, or L-arabinopyranose. In their paper, *Wang* and co-workers produced fully-protected glucopyranosides using orthogonal protecting groups allowing for selective deprotection of hydroxyl groups. Thus, it was attempted to generate a 4-nitrophenyl β-D-glucopyranoside which could be selectively opened to give the free 4-OH or the 6-OH leaving the other groups protected.



Scheme 3-9: Our desired galactosyl glucopyranoside references can be generated from differently tri-*O*-protected glucopyranosides and per-*O*-acetylated galF (**33a**).

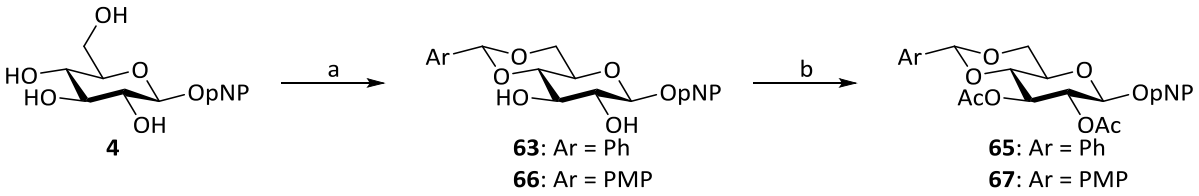
In order to achieve this, our method hinged on the use of a 4,6-*O*-benzylidenyl acetal and selective acetal opening. Selective protection of 1,3-diols using arylidene acetals is not new in carbohydrate chemistry with extensive studies done into 1,3-dioxolane rings and selective cleavage thereof beginning in 1951 with *Doukas* and *Fontaine*.[362] Since then, a number of studies have generated a variety of methods for regioselective ring opening of benzylidenyl-protected glycosides to generate either 6-aryl-4-hydroxyl-[363-366] or 4-aryl-6-hydroxyl-[367,368] 2,3-di-*O*-protected glycosides (Scheme 3-10).[369,370] These reactions take place under reductive conditions and employ a combination of reducing agents, Brønsted acids, and Lewis acids to affect selective opening. These conditions though have been shown to be affected by the choice of solvent,[368] the choice of acid,[369,371] and by the identity of the protecting group on the 3-hydroxyl.[365,368]



Scheme 3-10: Regioselective ring opening of a benzylidenyl acetal can free either the 4- or the 6-hydroxyl group depending on the reaction conditons used.

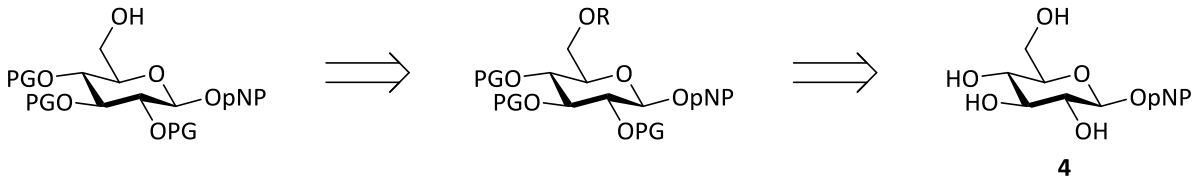
Starting with compound **4** (Scheme 3-11), the reaction with benzaldehyde dimethyl acetal in the presence of CSA to return the desired 4,6-benzylidenyl acetal **63** in an 87% yield. This product was then subsequently diacetylated to protect the 2- and 3-OH groups to return a fully-protected glucoside, **64**. At this point, it was decided to attempt the regioselective dioxolane opening to return the free 6-OH compound. This reaction was performed using Et_3SiH activated by $PhBCl_2$ as outlined

by *Rising et al.*[372] However, it was observed that under these conditions, the nitro functionality of the aglycone was also reduced to the amine. It was decided to use the *p*-methoxybenzylidene as this should allow for oxidative opening as a result of the electron-donating methoxy group. This should activate the benzylic position by making it more electrophilic rendering it more susceptible to nucleophilic attack. The *p*-methoxybenzaldehyde dimethyl acetal (**65**) was synthesised and isolated by distillation. This was then subsequently used to protect the 4- and 6-OH groups and returned the acetal **66** in an 86% yield. Again, the 2- and 3-OHs were acetylated to return the fully-protected glucoside, **67**. Fortunately, compound **67** precipitated out of solution during the course of the reaction and was collected in a 78% yield. However, this blessing turned sour when it was discovered that this compound was insoluble in most solvents with the exception of MeCN. For the oxidative acetal opening, a number of different oxidative conditions were trialled such as: metal salts with *tert*-butyl hydrogen peroxide,[373] sodium bromate/sodium dithionite,[374] CAN,[375] and DDQ.[376] Disappointingly, these reaction conditions resulted in either no reaction, or the generation of a large number of unwanted by-products, which when observed by ^{1}H NMR, appeared to be mostly over-oxidised as attested to by the large number of aldehyde signals. In light of this, our attempts to synthesise a universal glucoside were abandoned.



Scheme 3-11: The synthesis of 4-nitrophenyl 4,6-arylidenyl-2,3-di-*O*-acetyl-β-D-glucopyranoside. Reaction conditions: a) $ArCH_2(OMe)_2$ (4.0 eq.), CSA (20 mol%), anh. MeCN, 50 °C, 4 h (69-88%); b) Ac_2O (3.0 eq.), py, room temperature, 16 h (55-78%).

As our attempts to synthesise a soluble and selectively-deprotectable glucopyranosyl acceptor were unsuccessful, we turned our attention to a direct synthesis of a glucoside possessing a free and stable 6-hydroxyl moiety (Scheme 3-12).



Scheme 3-12: Retrosynthetic analysis of a 4-nitrophenyl 2,3,4-tri-*O*-protected-β-D-glucopyranoside from glc-pNP (**4**) *via* a selectively removable protecting group, (R), on the primary hydroxyl group.

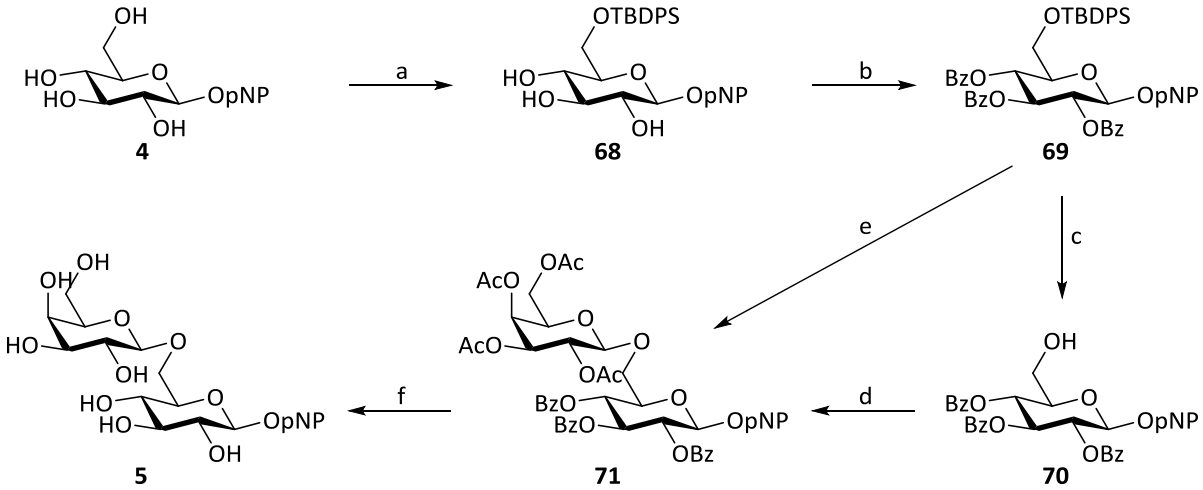
Starting from 4-nitrophenyl β-D-glucopyranoside (**4**), we endeavoured to first selectively-protect the primary 6-hydroxyl group using a large protecting group, such as the trityl (Tr) and the

tert-butyldiphenylsilyl (TBDPS) groups. These two groups have strong preference for the primary hydroxyl group over the secondary groups, a result of their steric bulk.[377] For the protection of the secondary hydroxyl groups, a benzoate ester (Bz) was chosen as, unlike acetate esters, the Bz group has a lower tendency for migration.[378] In addition to this, the Bz protecting group can be cleaved under similar conditions as Ac groups, which are present on the glycosyl fluoride donors. This translates to a one-step deprotection for the entire disaccharide, which should result in higher yields as there is one less workup and purification step.

With this synthetic approach, primary hydroxyl of compound **4** was first protected using TBDPS. It was decided to use this silyl ether, as the silyl ethers are well known for being selectively cleavable using fluorides under mild conditions and stable under the conditions required for the subsequent benzoylation.[379,380] Using a procedure developed by *Fylaktakidou* and co-workers,[381] compound **4** was selectively silylated under inert conditions using TBPDSCl, Et_3N, and DMAP in anhydrous DMF. This reaction was quite slow, but after 24 h and column chromatography, a good yield of product **68** was obtained (60%, Scheme 3-13). Following isolation, product **68** was benzoylated using BzCl and DMAP and refluxed in dichloromethane overnight. The target compound **69** was then isolated *via* column chromatography. Repetition of this reaction on a larger scale indicated a total conversion to the desired product by TLC. Isolation using the same chromatographic conditions and purity analysis by ^{1}H NMR indicated a large impurity in the fraction containing our product. A number of conditions were trialled for separation using TLC which eventually led to a second spot appearing underneath the product spot. This spot, although not isolated, was presumed to be a di-*O*-benzoylated by-product. A second purification was performed using a strongly non-polar eluent mixture of PE, toluene, and dichloromethane. The use of this mixture as the first purification step always led to mixed fractions, which was presumably the result of intermolecular π-π stacking interactions between the many aromatic rings. A first column using EtOAc/PE was required for a crude purification before compound **69** could be properly isolated using column chromatography.

The isolated tri-*O*-benzoyl-6-silyl glucoside required now selective deprotection of the TBDPS group. Although the use of TBAF as a fluoride source for cleavage of silyl ethers is a well-established method, TBAF itself is quite corrosive to the mucosa, respiratory tract,[382] and the environment in general.[383] Furthermore, the silyl fluoride by-product is wasted.[384] The use of recyclable, silica-supported triflic acid catalyst as outlined by *Yan et al.*[385] was trialled as an alternative method for selective desilylation. The synthesis of the catalyst required only anhydrous conditions, and once completely dry, the free-flowing powder was stable at room temperature. In addition, work-up of the reaction required only filtration to remove the SiO_2·TfOH which was shown in their research to be reusable and did not encourage migration of the Bz group. The use of this reagent returned the 6-hydroxy tri-*O*-benzoyl glucoside **70** in good yields and subsequent galactosylation of this position using fluoride **33a** in the presence of BF_3·OEt_2 returned the disaccharide **71**, albeit in a low yield of

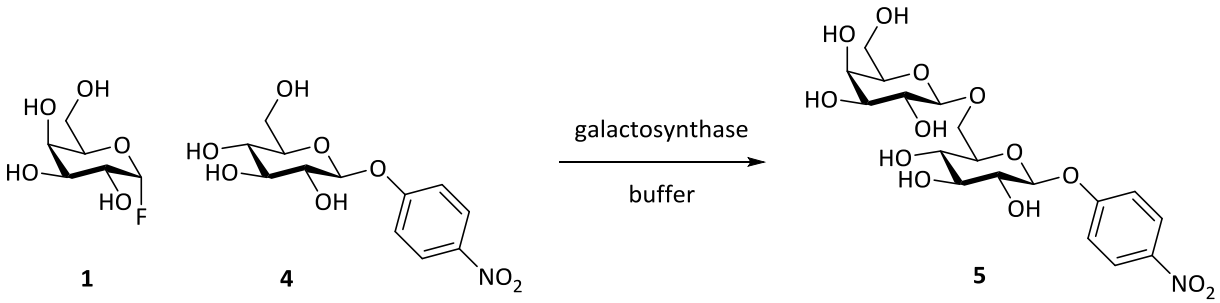
25%; A two-step yield of just 15%. This led to a reconsideration of the methodology behind the glycosylation. Keeping in mind that silyl ethers can be removed using fluoride sources, it was considered that direct reaction of the silylated glucoside **69** with the galactosyl fluoride **33a** in the presence of $BF_3{\cdot}OEt_2$ may also result in glycosylation as fluoride is released when fluoride **33a** is reacted. Additionally, it has previously been shown that $BF_3{\cdot}OEt_2$ can also be used for desilylation.[386] Given that there is also literature precedent for the combination deprotection/glycosylation,[77,387] this reaction was then trialled using the fluoride **33a** and the glucoside **69** in a ratio of 1.1:1.0. This reaction was successful and although the yield was only 46%, it still provided a better alternative to the previous two-step method in terms of time and yield. Deacetylation and debenzoylation of the resulting disaccharide, **71**, was performed in one step using *Zemplén* conditions and generated the desired allolactoside **5** in quantitative yields.



Scheme 3-13: The synthesis of 4-nitrophenyl β-D-allolactopyranose (**5**). Reaction conditions: a) TBPDSCl (1.1 eq.), Et_3N (1.4 eq.), DMAP (0.1 mol%), N_2, anh. DMF, room temperature, 24 h (60%); b) BzCl (6.0 eq.), TMP (4.0 eq.), DMAP (10 mol%), anh. CH_2Cl_2, reflux, 16 h (60%); c) $TfOH{\cdot}SiO_2$ (10 mol%), MeCN, 50 °C, 30 min (61%); d) $(Ac)_4$-β-galF (**33a**, 2.0 eq.), $BF_3{\cdot}OEt_2$ (2.5 eq.), Et_3N (0.5 eq.), anh. CH_2Cl_2, room temperature, 16 h (25%); e) $(Ac)_4$-β-galF (**33a**, 1.1 eq.), $BF_3{\cdot}OEt_2$ (2.2 eq.), Et_3N (1.1 eq.), anh. CH_2Cl_2, room temperature, 30 min (46%); f) NaOMe (0.5 M), anh. MeOH/ CH_2Cl_2 (2:1), room temperature, 16 h (*quant.*).

3.5.5 Application of Glycosynthases

Using the work of *Jakeman* and *Withers*[1] as a basis, it was first attempted to replicate their results. The LacZ E537S/G794D enzyme was expressed and purified according to common procedures and was added to a reaction tube containing galF (**1**) and glc-pNP (**4**) in phosphate buffer (Scheme 3-14).



Scheme 3-14: Reaction of the galactosyl donor, (**1**), and the galactosyl acceptor, (**4**), mediated by a galactosynthase to generate the desired β-(1,6)-disaccharide (**5**). Conditions: 200 mM phosphate buffer (pH 7.5), room temperature, 48 hrs.

Following the reaction and removal of solvent, the products were derivatised *via* per-acetylation to allow for easy quantification in the absence of the buffering salts. The product mixture was then visualised using TLC, which revealed a spot that did not correspond to either starting material, pNP, or galactose. After purification using preparative TLC (pTLC), this product spot was isolated in a minimal, 2% yield. The product was identified as 4-nitrophenyl hepta-*O*-acetylallolactose (**74**) *via* NMR and was consistent with those values reported by *Jakeman* and *Withers*.[1] This allowed us to be confident that our enzyme is capable of producing the β-(1,6) disaccharide. However, with a yield of the per-*O*-acetylated compound (**74**) of 0.6 mg (2%), this reaction clearly was not as efficient as the 9.4 mg (12.4 mg per-*O*-acetylated) allolactoside claimed by *Jakeman* and *Withers*.[1] This result was surprising, as the conditions for the reaction were identical to those reported. The reaction was repeated using the LacZ E537S enzyme, again under identical conditions to those reported, but after derivatisation, none of the allolactoside, **74**, was isolable.

In order to avoid derivatisation of the product mixtures, it was attempted to exploit the partial solubility of the sugar in organic solvents to separate them from the salts and enzymes remaining in the lyophilisate. As a proof-of-theory, a mixture of galactose (**7**), glc-pNP (**4**), and lactose (**2**) was tested in an attempt to extract them using a small selection of readily-available polar solvents: acetone, methyl *tert*-butyl ether (MTBE), diethyl ether (Et_2O), acetonitrile (MeCN), methanol (MeOH), and ethanol (EtOH) (Figure 3-41). The crystalline carbohydrates were combined and then extracted using the aforementioned solvents for one hour at reflux before the extraction was stopped and the solvent removed under reduced pressure. The extract was then compared to the standards *via* ^{1}H NMR (Table 3-12).

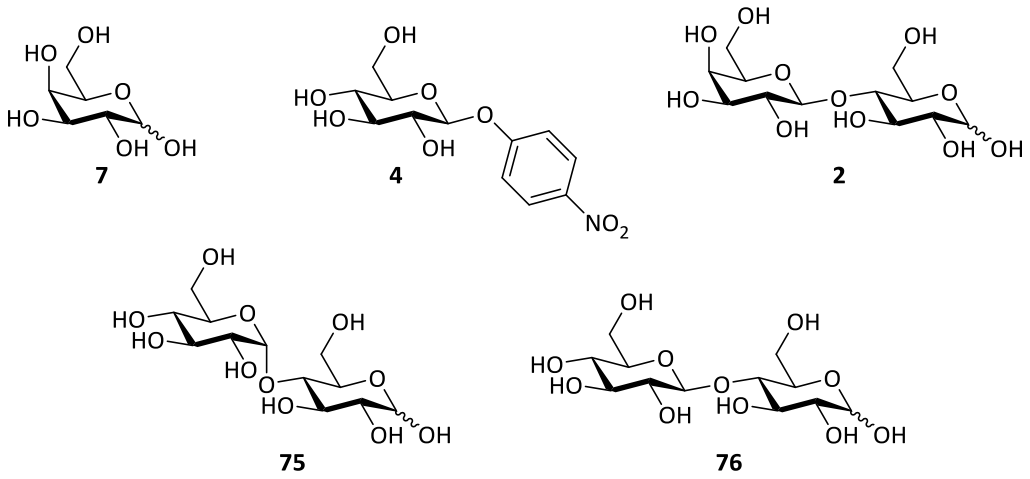


Figure 3-41: Compounds tested for extraction using the Soxhlet apparatus.

Table 3-12: Results of the continuous extraction of carbohydrates from various solvents after being refluxed for one hour. The masses given correspond to the amount of carbohydrate added to the extraction cup. Those masses given in italics are those which were soluble in the given solvent as determined by ^{1}H NMR. The insoluble carbohydrate was excluded from the yield calculations and as such, the recovery yields are given as percentage of soluble carbohydrate.

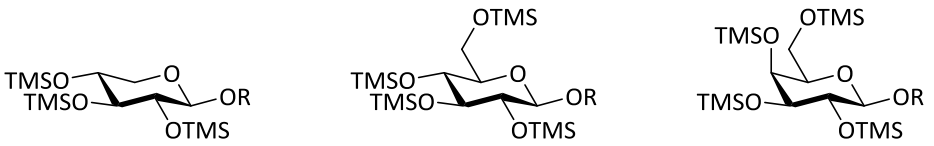
Solvent	bp [°C]	gal (**7**) [mg]	glc-pNP (**4**) [mg]	lac (**2**) [mg]	Recovered Mass [mg] (*yield*)
Acetone	56	9.5	*7.7*	10.8	8.1 (*quant.*)
Et_2O	35	9.9	*9.9*	11.8	neg.
MTBE	55	10.0	*8.7*	9.6	7.6 (*87%*)
MeCN	82	*9.7*	*13.6*	13.3	17.9 (*77%*)
MeOH	65	*9.1*	*10.3*	*11.0*	25.9 (*85%*)
EtOH	78	*9.7*	*10.1*	*10.2*	25.9 (*86%*)

Interestingly, the 4-nitrophenyl glucoside **4** was selectively isolable when using acetone and the ethers. This suggested that this method may be applicable to selective extraction of the phenyl compound from complex mixtures. Furthermore, MeCN was unable to extract the disaccharide, and it was considered that sequential extraction could allow for easy separation of the various compounds. Attempts to up-scale these extractions to 100 mg of each carbohydrate whilst simultaneously expanding the selection to include maltose (**75**) and cellobiose (**76**) using Et_2O, MeCN, and EtOH sequentially were performed. Disappointingly, only mixtures of the saccharides were extracted and chromatographic separation of products would be required regardless. Therefore, we turned our focus into developing and investigating methods for monitoring the reaction, preferably quantitatively.

3.5.6 Characterisation of Glycosynthase Product Mixtures

Gas Chromatography

The first approach was to use GC-MS to monitor the product mixture. Due to the polar nature of carbohydrates, it was necessary to first derivatise them before analysis *via* GC.[388-391] It was also hoped that by using a coupled mass spectrometer we could use the fragmentation patterns generated by the known standards for analysis of unknown compounds by fragmentation analysis. Monosaccharide standards were per-*O*-(trimethyl)silylated according to a known procedure[388] and analysed for method development (Figure 3-42).



xyl: R = TMS
77: R = 4-NO_2Ph

glc: R = TMS
78: R = 4-NO_2Ph

gal: R = TMS
79: R = 4-NO_2Ph

Figure 3-42: Trimethylsilylated standards used for GC-MS analysis. As the per-*O*-silylation of unprotected carbohydrates can result in multiple products (i.e. open chain, furanose, pyranose, and axially- or equatorially-configured anomeric centres), only the β-pyranose form is shown although a mixture of these products will be obtained, albeit in a predictable ratio as discussed by *Angyal*.[38]

The first GC method used was adapted from literature[389] and the individual saccharide samples were measured (Figure 3-43). This method showed good separation of the compounds and a number of observations were made. Firstly, the monosaccharides galactose (gal), glucose (glc), and xylose (xyl) showed multiple peaks which were attributed to various conformers. Secondly, the *p*-nitrophenyl β-derivatives of the aforementioned saccharides had quite high boiling points, all eluting around 270 °C. Finally, although the peaks could be clearly observed, no molecular peak could be observed in the mass spectra associated with them. This result meant that fragmentation analysis for the determination of the identity of any unknown compounds detected was unsuitable for these compounds and conditions.

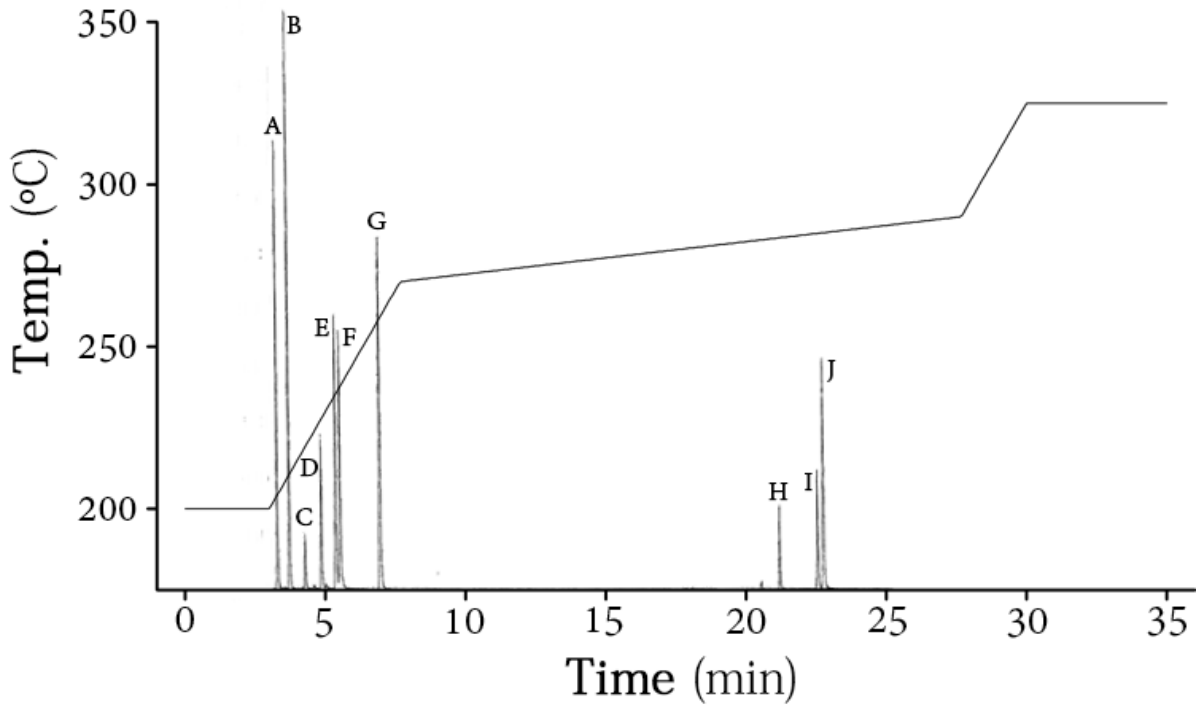


Figure 3-43: Overlaid chromatogram and temperature gradient of GC method A. Peak information (Compound, R_t (min), Temp. (°C): A – (**xyl**), 3.32, 204.8; B – (**xyl**), 3.75, 211.2; C – (**gal**), 4.31, 219.6; D – (**gal**), 4.89, 228.4; E – (**glc**), 5.41, 236.2; F – (**gal**), 5.57, 238.5; G – (**glc**), 6.92, 258.8; H – (**77**), 21.23, 283.6; I – (**78**), 22.58, 284.9; J – (**79**), 22.79, 285.1.

In light of this, a new program was developed with a lower starting temperature and gentler temperature gradient (Figure 3-44). With regard to better separation, this was achieved using this method. However, the elution of the pNP derivatives was not significantly lowered and their still occurred in the range of 250 to 260 °C. This raised some concern over the probable elution temperature of pNP disaccharides which were expected to be higher.

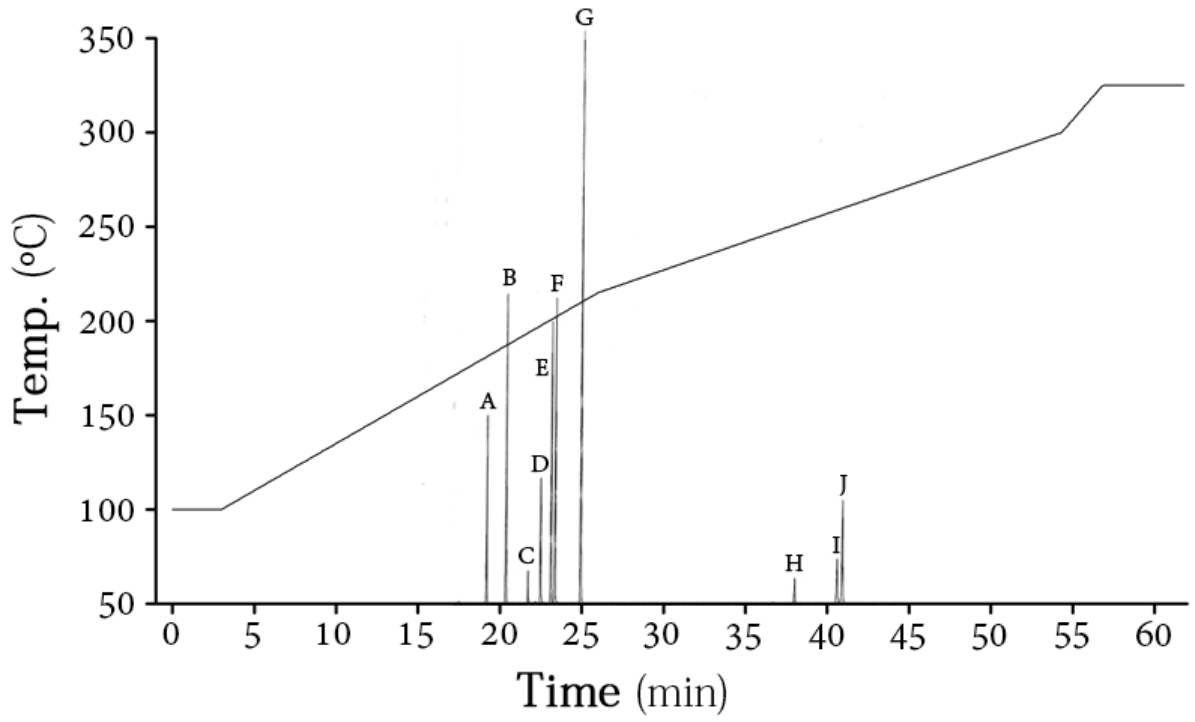


Figure 3-44: Overlaid chromatogram and temperature gradient of GC method B. Peak information (Compound, R_t (min), Temp. (°C): A – (**xyl**), 19.17, 180.9; B – (**xyl**), 20.35, 186.7; C – (**gal**), 21.70, 193.5; D – (**gal**), 22.46,197.3; E – (**glc**), 23.10, 200.5; F – (**gal**), 23.33, 201.7; G – (**glc**), 24.90, 209.5; H – (**77**), 37.97, 250.9; I – (**78**), 40.56, 258.7; J – (**79**), 40.84, 259.5.

The applicability of this method to the analysis of glycosynthase-catalysed reaction mixtures was determined to be limited, and therefore of no further use. The TMS derivatisation of the reaction samples would first require lyophilisation, as the TMS group is highly sensitive to water. This meant that taking regular samples and obtaining information in real-time was not possible. Coupled with the high temperatures required to elute the pNP monosaccharides and the lack of fragmentation analysis, the use of this method was reconsidered.

HPLC-DAD

HPLC separation of small molecules is an advanced field of analysis and typically involves only small quantities of product. Carbohydrates are notoriously difficult to separate using standard columns as the similarity in their polarities generally lead to inseparable mixtures. To address this, a number of specialised columns with chiral stationary phases are predominantly used for their separation *via* HPLC, namely amine and amide columns. Furthermore, the standard detector on most HPLC devices, a diode array detector (DAD), which measures the UV-Vis absorbance of analytes in a sample, is unsuitable for the analysis of carbohydrates as they lack UV-active properties under standard conditions. Additionally, other standard detectors, such as mass spectrometers, require post-column addition of solvents to encourage and stabilise ion formation of carbohydrates.[392] This makes the detection of carbohydrates difficult using standard techniques.

As our model glycosyl acceptor, glc-pNP, contained an aromatic system in the aglycone, it was thought that this moiety would be visible when analysing using a DAD coupled to the HPLC column. Due to the high polarity of the glycosyl compounds, derivatisation of the products was also desired for use with normal phase (np) HPLC columns, particularly as we had access to preparative np columns with a chiral stationary phase. To this, the glycosides were derivatised *via* per-*O*-acetylation and submitted for analysis (Figure 3-45).

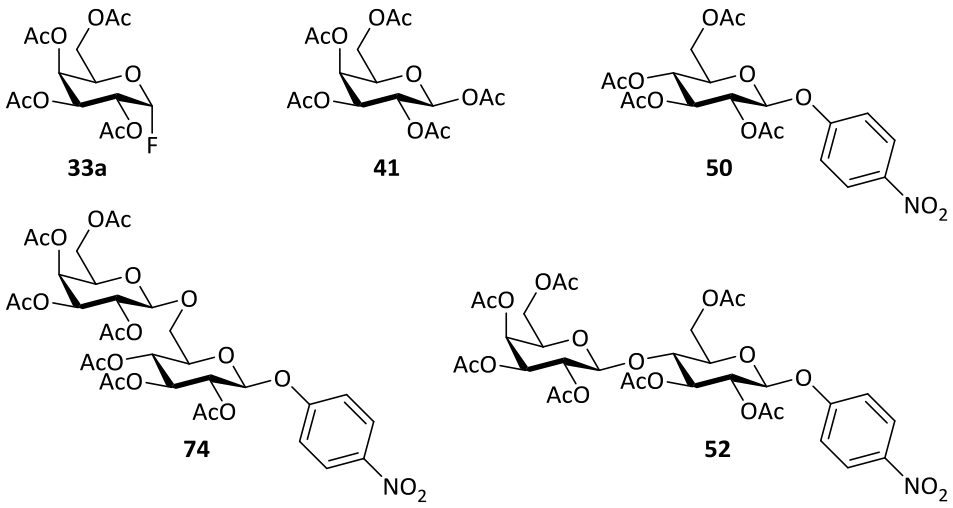


Figure 3-45: Standard per-acetylated samples used for the development of a HPLC-UV-Vis-based analysis of glycosynthase reaction mixtures.

Initially, 5 samples were submitted: $(Ac)_4$galF (**33a**), $(Ac)_5$gal (**41**), $(Ac)_4$glc-pNP (**50**), $(Ac)_7$allo-pNP (**74**), and a crude sample of $(Ac)_7$lac-pNP (**52**). These were then run individually and thought to be separable with the disaccharides eluting after nearly two hours. Closer analysis of the samples, demonstrated that the disaccharides absorbed weakly across all wavelengths when measured by HPLC-DAD. A further disappointment was when $(Ac)_7$allo-pNP (**74**) was analysed as it overlapped with the lactose derivative **52.** In light of these results, this method was deemed inappropriate for the analysis of reaction mixtures and was therefore abandoned.

HPLC-MS/MS

Mass spectrometric analysis of small molecules typically requires only minimal amounts of sample and coupling this with HPLC seemed ideal for following the reactions of galactosynthases. In collaboration with the *Noack* working group of the *IBG-1* at the *Forschungszentrum Jülich*, a method of HPLC-MS/MS, specifically an ESI/TOF tandem system, was planned for the analysis of the mixtures of these reactions. ESI-MS, or electrospray ionisation mass spectrometry, is a sensitive and reliable method of analysis for studying femtomolar amounts of analytes.[393] This method is applicable to both large and small molecules with various polarities. In contrast to other mass spectrometers which measure the charge/mass (*m*/*z*) ratio of molecules, TOF (time-of-flight)

instruments measure the time required for a molecule or fragment to arrive at a plane parallel to the plane of origin (Figure 3-46). The advantage of this combination is therefore two-fold. The ESI method generates fragments, allowing us to study both the fragmentation of our analytes whilst utilising the higher sensitivity of TOF instruments. Furthermore, this method did not require the derivatisation of samples. Additionally, as this method required only micrograms of sample, this meant that samples of the reactions should be able to be taken without major disruption of the reaction, which could occur by depletion of solvent and disruption of the mixing process.

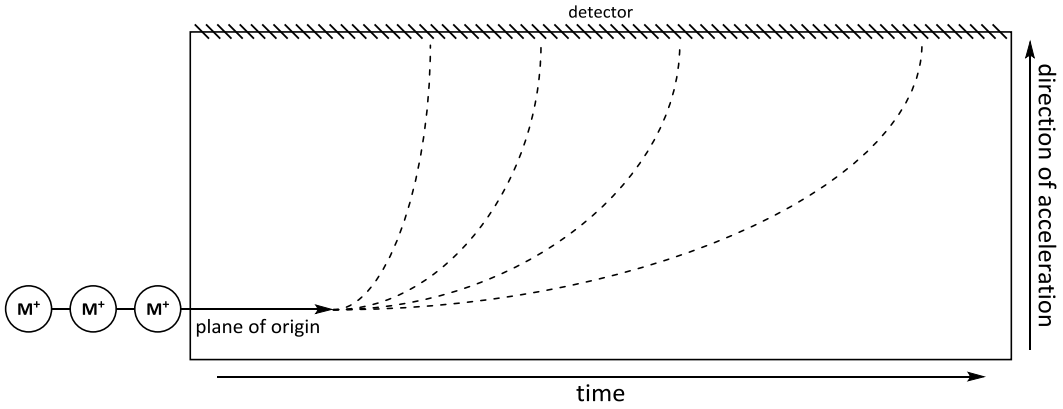


Figure 3-46: Simplified schematic of time-of-flight (TOF) mass spectrometry. The dotted lines represent the trajectory of analytes of different mass-charge ratios. The smallest compound impact the detector first (most left) and the largest impact later (most right).

In order to develop a method, a number of non-derivatised samples were provided (Figure 3-47). These samples consisted of the two starting materials, the autohydrolysis product of compound **1**, and the two expected 4-nitrophenyl disaccharides. In addition, the proposed internal standard DMS (**34**) was also included. As the software allows for mass selection, the compound only needed to be eluted in the measured range in order for quantities to be measured.

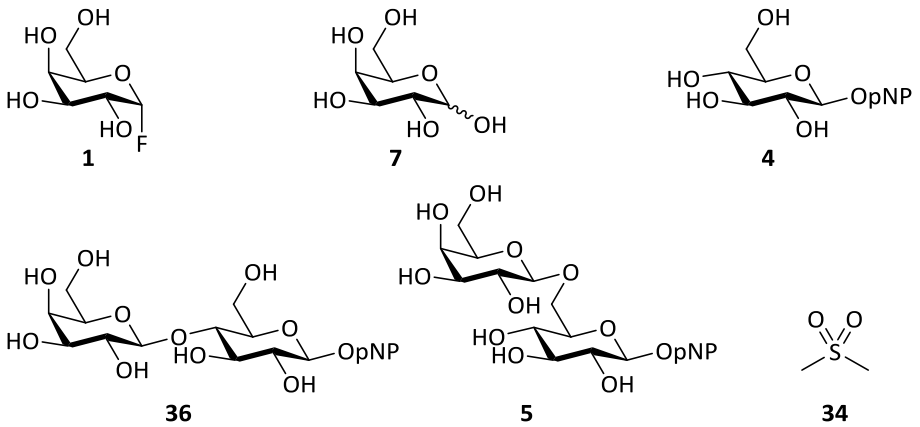


Figure 3-47: Standards provided for the development of the HPLC-MS/MS method of reaction mixture analysis.

These samples were then measured using a HILIC (hydrophobic interaction liquid chromatography) column, a technique used for the separation of polar and hydrophilic compounds.[395,396] Although compound identity in chromatography is typically determined by the retention time, as our analysis involved mass spectrometry, we were able to isolate peaks by their *m/z* from the total ion current (TIC). Consequently, only compounds with the same mass needed to be separable using this method rather than all compounds in the sample. This proved successful as 4-nitrophenyl β-D-lactoside (**36**) and 4-nitrophenyl β-D-allolactoside (**5**), which both have the same *m/z* (465 $[M+H]^+$) were clearly separable using this method and eluted at 8.75 min and 17.44 min, respectively (Figure 3-48). Additionally, as we had the fragmentation patterns from our standard samples, we could identify with certainty the analyte.

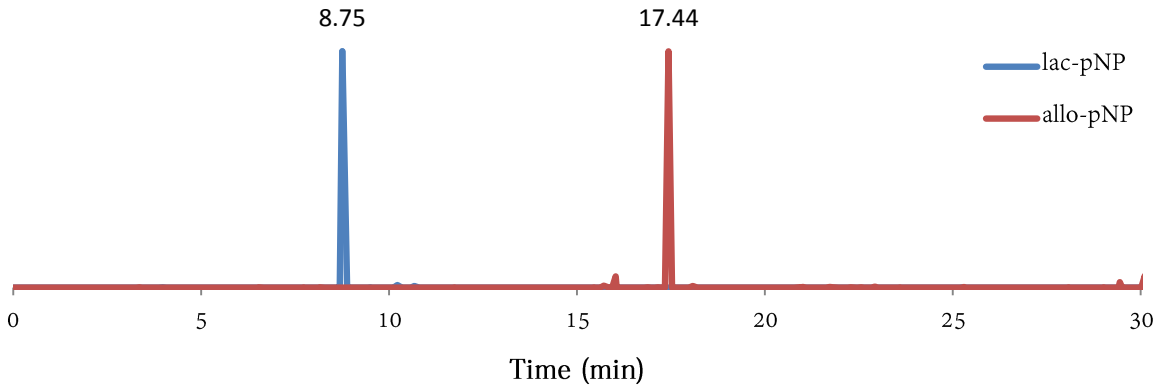


Figure 3-48: Extracted ion chromatogram of *m/z* = 465 from the HPLC-MS/MS TIC of the separation of 4-nitrophenyl β-D-lactopyranoside (**36**, blue) and 4-nitrophenyl β-D-allolactopyranoside (**5**, red).

Chapter 4 - Summary and Outlook

4.1 Summary

The importance of carbohydrate-containing molecules for biological systems cannot be understated. It has been repeatedly shown that these organisms are critical to sustaining life and have a broad range of important functions. As glycosides are found in many natural food sources, particularly in plants, and their presence or absence can contribute to the overall taste and flavour of the final product – for example naringin, a compound responsible for the bitterness in grapefruits, which is tasteless after cleavage of glycosides[397] – these compounds are of great importance in the food processing industry. Additionally, carbohydrate chemistry is of increasing importance within the pharmaceutical and cosmetic industries as carbohydrate-containing compounds have been linked with many disease,[31,398] infection,[399,400] and ageing processes.[401,402] Consequently, researchers are constantly looking for methods of efficient glycosylation of small and large molecules alike. Since the advent of carbohydrate chemistry in the late 1870s by *Franchimont*[52] and *Michael*[40], many methods have been developed for the chemical synthesis of glycosides. These methods have the advantage of being applicable to a broad range of systems, but are limited on large scales due to the extensive protection and deprotection strategies required for stereo-, regio-, and chemoselective glycosylation. With the discovery and application of enzymes as tools in carbohydrate synthesis in the latter half of last century, it was found that enzymes possess high substrate specificity, which allows chemists to avoid such protection/deprotection strategies. This high specificity, while the biggest strength of biocatalysts, is also their largest weakness as it can become difficult to create non-natural and unusual glycosides. It has been found on numerous occasions that a blending of the two synthetic methods, chemoenzymatic syntheses, can be highly beneficial for the construction of glycosides on various scales.[143,403-406]

Three β-galactosidases were identified as potential galactosynthases: LacZ from *Escherichia coli*, Lac4 from *Kluyveromyces lactis*, and LacA from *Aspergillus oryzae* and constructs and mutants thereof created (Table 4-1). These all have been demonstrated to have β-(1,6) specificity. In the case of LacZ, the *lacZ* gene was successfully isolated from the BL21(DE3) strain of *E. coli* using PCR techniques and then cloned into the pET22b(+) vector and two mutants, E537S and E537S/G794D, were generatedusing site-directed mutagenesis, both of which are literature known.[1] Lac4, to our knowledge, has not yet been developed and used as a galactosynthase and the gene was optimised and synthesised and five mutants of the nucleophilic residue, Glu551, were generated: Gly, Ala, Leu, Ser, and Gln. Isolation of viable genetic material from the fungus, *A. oryzae*, proved challenging and ultimately synthesis of the gene, *lacA*, was required. Analogously to Lac4, the *lacA* gene was optimised, synthesised, and subcloned into pET22b(+). The same five nucleophile mutants as Lac4

were also created for this enzyme at residue Glu298. When expressed in *E. coli*, this enzyme proved to be inactive in solution and further attempts to work with it were abandoned for the purposes of this work.

Table 4-1: Summary of constructs created during the progress of this work.

Gene	Construct
lacZ WT	pET28a::*lacZ* WT
	pET28a*::*lacZ* WT
	pET22::*lacZ* WT
lacZ E537S	pET28a::*lacZ* E537S
	pET28a*::*lacZ* E537S
	pET22::*lacZ* E537S
lacZ E537S/G794D	pET28a::*lacZ* E537S/G794D
	pET28a*::*lacZ* E537S/G794D
	pET22::*lacZ* E573S/G794D
lac4 WT	pUC57::*lac4* WT
	pET22::*lac4* WT
lac4 E551G	pET22::*lac4* E551G
lac4 E551A	pET22::*lac4* E551A
lac4 E551L	pET22::*lac4* E551L
lac4 E551S	pET22::*lac4* E551S
lac4 E551Q	pET22::*lac4* E551Q
lacA WT	pUC57::*lacA* WT
	pET22::*lacA* WT
lacA E551G	pET22::*lacA* E551G
lacA E551A	pET22::*lacA* E551A
lacA E551L	pET22::*lacA* E551L
lacA E551S	pET22::*lacA* E551S
lacA E551Q	pET22::*lacA* E551Q

Initial expressions of these enzymes in *E. coli* ER2566 resulted in poor titres and an investigation into optimising the expression conditions was undertaken. Four parameters were therefore varied in attempts to improve the titre of LacZ and it was also hoped that this could be applied to the expression of Lac4. By measuring the hydrolytic activity using established colourimetric assays, it was observed that the concentration of IPTG used to induce the sample had a negligible effect on the amount of enzyme produced. It was observed that *E. coli* had a preference for basic growth conditions as evidenced by high hydrolytic activity from those cultures grown at pH 7.5 and 8.0. The largest influences came from the time of induction and the growth temperature (Figure 3-8).

The induction time, linked directly to the cell mass, showed that induction was required to take place at the beginning of the exponential growth phase. Coupled with an increase in the incubation temperature to 37 °C, we were able to improve the production level by approximately 6-fold as evidenced by the increased specific activity of the clarified supernatant of lysed cells. In addition, it

was also attempted to improve the titre by means of a purification table. This table revealed that the enzyme was not binding very strongly to the column, which is most likely a result of the *C*-terminal His_6-tag. This tag could be obstructed from interacting with the immobilised Ni^{2+} due to the bulk of the enzyme. An investigation into the ideal position of the tag should be pursued. Furthermore, it was revealed that imidazole is a strong inhibitor of Lac4, and therefore, an investigation into the identity of the purification tag should also be further explored. This could, by extension, also improve the activity of the already-slow mutant enzymes. However, after optimising the purification technique, we were able to increase the yield of purified LacZ 2-fold, from 5 to 11 mg L^{-1}, and of purified Lac4 8-fold, from 6 to 46 mg L^{-1}.

Attempts were made to characterise the LacZ and Lac4 enzymes using a number of methods. Characterisation of hydrolysis was readily applicable and easy to follow due to the many well-established procedures available in literature. However, the characterisation of the reverse reaction, the formation of *O*-glycosyl species, proved quite elusive. All attempts to characterise the synthetic capabilities of these enzymes photometrically were unsuccessful, as were those methods developed for the detection of the generated fluoride ion. Both of these methods indicated an extremely slow turnover rate and these methods were not sensitive enough to detect such small changes. ^{1}H NMR investigations into the stability of galF (**1**) were performed and it was found that this compound is the most stable at neutral pH. This provided the base for NMR-based approaches to activity quantification, which showed promising signs of being applicable to the analysis of these reactions, particularly when coupled with an internal standard. Unfortunately, the low reproducibility is a problem that desperately needs to be addressed before continuing with this method of carbohydrate reaction analysis. Furthermore, a large sample volume (500 µL), the use of deuterated solvent, and the slow rate of data collection, when compared to plate readers for example, make this method more applicable for determination of specific activity, rather than Michaelis-Menten kinetic analysis, which requires many samples to be prepared. Perhaps, with further development in maximising enzyme titre, by changing host organism for example, or by redesigning the enzyme by changing the position and identity of the purification tag, the titre can be increased further.

The synthesis of the reference compounds and reagents proved to be for the most part uncomplicated, although a number of transformations were low yielding. We were able to increase the time-efficiency of the reactions by avoiding column chromatography and utilising instead crystallisation techniques for purification of a number of the glycosides. During the synthesis of the fluoridated glycosyl donors, it was observed that these compounds were instable and had a tendency to undergo autohydrolysis under normal deprotection reaction and storage conditions. By modifying the deprotection to react at a lower temperature and for a substantially shorter time, we were able to successfully isolate the desired glycosyl donor, galF (**1**). Our hypothesis that the original degradation was accelerated in unbuffered aqueous systems, including room temperature storage, by the *in situ*

generation of HF was given credibility by the results of the ^{1}H NMR stability studies. The stability of the other generated glycosyl fluorides was not investigated in this manner, but it is expected that they too would be most stable in dry conditions or neutral, aqueous buffer systems.

In our attempts to generate a selectively-deprotectable glucoside, we were unsuccessful in applying this strategy to our 4-nitrophenyl system. The literature-known methods of selective, reductive acetal opening of 4,6-arylidenyl glycosides also resulted in the reduction of the nitro functionality. Oxidative methods were also unsuccessful in selectively opening our acetal as hoped. While this method was not applicable to our nitrophenyl compounds, it may be useful in the generation of other benzylidenyl substrates such as umbelliferonyl glucosides, which can be used in fluorimetric assays. Regardless, we were able to synthesise the *p*-nitrophenyl allolactoside, **5**, using a more directed approach wherein the 6-hydroxyl group was selectively protected and deprotected to allow for specific β-(1,6) coupling with a galactosyl fluoride.

Investigations into the generated galactosynthases is ongoing, but we were unable to replicate the results presented by *Jakeman* and *Withers*.[1] Though the galactosynthases generated from the galactohydrolase from *Escherichia coli*, LacZ, were able to catalyse the formation of a β-(1,6) glycosidic linkage between galF (**1**) and glc-pNP (**4**), the yield of this experiment was much lower than that reported in literature. These results indicate that further research into the optimisation of reaction conditions and analytic methods is required. Disappointingly, we were unable to fully investigate the reactions of our novel galactosynthases generated from the yeast β-galactosidase, Lac4 from *Kluyveromyces lactis*.

Our brief foray into continuous extraction as a method of selective isolation of various glycosides from complex mixtures appeared promising, but ultimately proved unreliable on larger scales. Consequently, attention was focussed upon the development and application of techniques to analyse the product mixtures without prior separation. To this end, we trialled GC-EI-MS and HPLC-DAD as a method of analysis for per-*O*-silylated and per-*O*-acetylated product mixtures, respectively. Unfortunately, neither of these methods appeared promising. Collaboration with the *Noack* working group with regards to the implementation of HPLC-MS/MS for the separation and identification of carbohydrate mixtures shows promise.

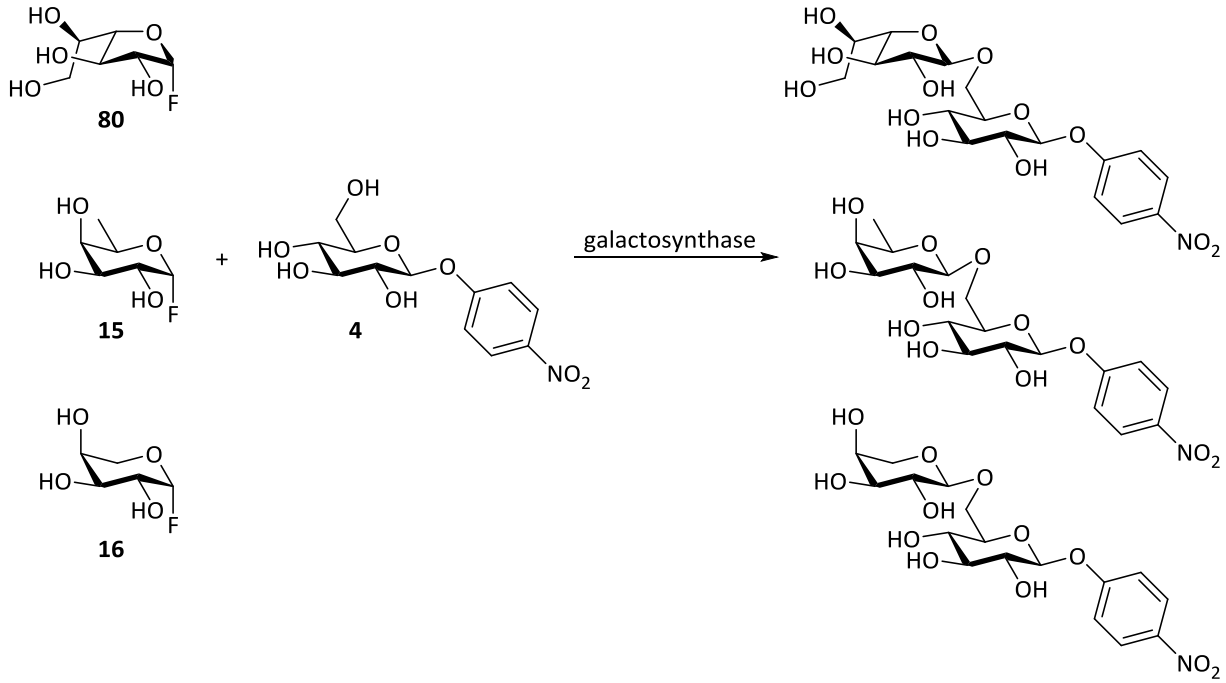
4.2 Outlook

While we were able to generate a number of new glycoside mutants, in the case of LacA, we were unable to express an active version of the enzyme, which suggests that post-translations modification of these enzymes is necessary for activity. Investigation into an effective and proven eukaryotic system, such as *Saccharomyces* sp., *Kluyveromyces* sp., *Hansenula* sp.,[290] or *Aspergillus* sp.,[407] should

be therefore be explored. These systems all have post-translational glycosylation systems, which should assist in the solubility of the LacA enzyme.

Isolation of the LacZ and Lac4 enzyme variants showed low titres. This problem could be addressed and rectified by use of an alternative expression system, such as *Burkholderia* sp.[408] or *Bacillus* sp.[409] It would be useful to investigate an organism capable of secreting any enzymes of interest to allow for easy isolation. Purification of our enzymes could also be improved by assessing and altering both the position and identity of the purification tag used, particularly as we were unable to reproduce the results reported by *Jakeman* and *Withers*.[1] This could be achieved by moving the existing *C*-terminus His_6-tags to the *N*-terminus or by removing the tag completely and instead using a size-exclusion chromatographic or precipitation method for enzyme purification. Alternative tags could also be used, such as *Strep*-[410] or NE tags.[411] Alternatively, the EOI could be coupled to maltose binding protein,[412] which could later be removed through use of a cleavable linker.[413] The use of such purification systems would also overcome the identified issue of imidazole inhibition.

One aspect which was touched upon, but not further investigated was the use of alternative glycosyl fluorides as donors. As we were able to observe hydrolysis of the 4-nitrophenyl fucopyranoside, arabinopyranoside, and the galactofuranosyl pyranoside, it suggests that these glycosyl scaffolds may be appropriate for glycosynthase reactions using our LacZ or Lac4 variants (Scheme 4-1). It would also be desirable to investigate the range of available glycosyl acceptor molecules, but further development of appropriate and reliable assays is first required.



Scheme 4-1: Alternative glycosyl fluoride donors could be used in the glycosynthase reaction as hinted at by results of the hydrolytic assays using 4-nitrophenyl glycosides and the wild type enzyme.

One of the biggest problems frequently encountered was that of analytical techniques for the quantitative determination of reaction kinetics and reaction product mixtures. We demonstrated that while colourimetry was appropriate for hydrolytic activity determination, it was not sensitive enough for the determination of synthase activity. GC-based analysis of product mixtures also proved unsuitable, as did HPLC-DAD and measurement using a fluoride-ion selective electrode. The use of ^{1}H NMR did show promise as a method of reaction monitoring and the determination of kinetic parameters, but large reaction volumes and multiple samples were required rendering this system appropriate for specific activity determination, rather than determination of Michaelis-Menten kinetic parameters. Additionally, in collaboration with the *Noack* group (IBG-1), we made a promising start on the use of HPLC-tandem mass spectrometry as a method for the quantification and separation of complex carbohydrate reaction mixtures. Alternative methods to investigate could include the use of capillary electrophoresis, a well-established method of carbohydrate and biomolecule analysis.[414-417]

Chapter 5 - Experimental

5.1 Biological Protocols

5.1.1 General Experimental Conditions

Various recipes, protocols, and procedures herein are taken from or adapted from a collection of general protocols and procedures used within the IBOC by present and past co-workers.

Equipment

Instrument	Supplier	Model
Centrifuge	Eppendorf AG	5424 R
Centrifuge	Hettich Holding GmbH	Rotanta 460R
Centrifuge	Thermo Fisher Scientific, Inc.	Sorvall™ RC6+
Centrifuge Rotor	Piramoon Technologies, Inc.	FIBERLite® F10-6x500y
Centrifuge Rotor	Thermo Fisher Scientific, Inc.	Sorvall™ SLC-4000
Chromatography Robot	GE Healthcare Life Sciences	ÄKTApurifier™
Column	Macherey-Nagel GmbH & Co. KG	Protino® Ni-NTA 5 mL
Column	QIAGEN GmbH	Ni-NTA Superflow 5 mL Cartridge
Cuvette	ratiolab	ratiolab® Q-VETTES semi-micro
Electrophoresis Chamber	Bio-Rad Laboratories, Inc.	Mini PROTEAN® Tetra Vertical Electrophoresis Cell
Electrophoresis Chamber	PEQLAB Biotechnologie GmbH	PerfectBlue™ Gel System Mini M
Electrophoresis Power Supply	Bio-Rad Laboratories, Inc.	PowerPac™ HC
Electrophoresis Power Supply	VWR International	Power Source 300V
French Press	Thermo Fisher Scientific, Inc.	French Press cell disruptor FA-078A-E
French Press	Thermo Fisher Scientific, Inc.	French® Pressure Cell FA-032
Incubator	Heidolph Instruments GmbH & Co. KG	Inkubator 1000
Incubator	LTF Labortechnik GmbH	Incubation Hood RCS
Incubator	New Brunswick Scientific Co, Inc.	Innova® 42R Incubator Shaker
Orbital Shaker	Edmund Bühler GmbH	TiMix
Orbital Shaker	Edmund Bühler GmbH	VKS-75 Control
Orbital Shaker	Eppendorf AG	Thermomixer compact
Orbital Shaker	Heidolph Instruments GmbH & Co. KG	Unimax 1010
PCR Thermal Cycler	Biometra GmbH	TProfessional Basic
PCR Thermal Cycler	VWR International	Doppio
Peristaltic Pump	Pharmacia Fine Chemicals	Pump P-1
pH Meter	Knick Elektronische Messgeräte GmbH & Co. KG	766 Calimatic

pH Meter	Knick Elektronische Messgeräte GmbH & Co. KG	Multi-Calimatic®
Pipette	Eppendorf AG	Research® 0.1 - 2.5 µL
Pipette	Eppendorf AG	Research® 1 - 10 µL
Pipette	Eppendorf AG	Research® 10 - 100 µL
Pipette	Eppendorf AG	Research® plus, single-channel, variable, 100 - 1000 µL
Pipette	Eppendorf AG	Research® plus, 8-channel, variable, 100 - 1000 µL
Pipette	Eppendorf AG	Research® plus, 12-channel, variable, 100 - 1000 µL
Pipette, electronic	Eppendorf AG	Research® pro, 8-channel, variable, 50 - 1200 µL
Rotator	Bibby Scientific Ltd.	Stuart® SB2
Rotator Tube Holder	Bibby Scientific Ltd.	SB3/1 Micro Tube Holder
Rotator Tube Holder	Bibby Scientific Ltd.	SB3/4 Culture Tube Holder
Sonicator	BANDELIN electronic GmbH & Co. KG	Sonopuls
Sonotrode	BANDELIN electronic GmbH & Co. KG	Sonopuls KE76
Spectrophotometer	Tecan Trading AG	Infinite® M1000 Pro
Spectrophotometer	Thermo Fisher Scientific, Inc./PEQLAB Biotechnologie GmbH	NanoDrop 2000c Spectrophotometer
Ultrafiltration Tube	Sartorius AG	Vivaspin 20

Software

Software	Version
Adobe® Photoshop® CS6 Extended[418]	v 13.0 x64
Clone Manager[419]	v 9.4, Professional Edition
GENtle[420]	v 1.9.4
KriKit[317,421]	v 1.0.3
Origin[422]	v 9.1.0
PyMOL[423]	v 1.3
UCSF Chimera[424,425]	v 1.8

Web Tools

Tool	Function
BLAST®[337]	Sequence alignment tool
ClustalW2[426]	Multiple sequence alignment tool
Oligo Calc[427,428]	Oligonucleotide properties calculator
Peptide Properties Calculator[429]	Polypeptide properties calculator
Phylogeny.fr[338,430]	Phylogenic analysis tool
PrimerX[431]	Primer design tool for qcSDM
Sequence Manipulation Suite[432,433]	Collection of DNA and protein analysis tools

Databases

Database	Description
BRENDA[434]	Comprehensive enzyme information system
CAZy[435]	Carbohydrate-active enzyme database
Kazusa[436]	Codon Usage Database
NCBI[437]	National Center for Biotechnology Information
PDB[438,439]	RSCB Protein Data Bank

Kits

Kit	Supplier	Product Description
DNA Purification	Analytik Jena AG	innuPREP DOUBLEpure Kit
Gel Extraction	QIAGEN GmbH	QIAquick® Gel Extraction Kit
Plasmid Isolation	Analytik Jena AG	innuPREP Plasmid Mini Kit

5.1.2 Growth Media

Media	Recipe	
2YT	16 g	tryptone
	10 g	yeast extract
	5 g	NaCl
		to 1000 mL with dH_2O
Lysogeny Broth (LB)	10 g	tryptone
	5 g	yeast extract
	10 g	NaCl
		to 1000 mL with dH_2O
Terrific Broth (TB), nutrients	12 g	tryptone
	24 g	yeast extract
	5 g	glycerol
		to 900 mL with dH_2O
Terrific Broth (TB), buffer (pH 6.5)	0.95 g	KH_2PO_4
	0.53 g	K_2HPO_4
		to 100 mL with dH_2O
Terrific Broth (TB), buffer (pH 7.0)	0.58 g	KH_2PO_4
	1.00 g	K_2HPO_4
		to 100 mL with dH_2O
Terrific Broth (TB), buffer (pH 7.5)	0.26 g	KH_2PO_4
	1.41 g	K_2HPO_4
		to 100 mL with dH_2O
Terrific Broth (TB), buffer (pH 8.0)	0.09 g	KH_2PO_4
	1.62 g	K_2HPO_4
		to 100 mL with dH_2O

After preparation, all media was autoclaved to sterilise. In the case of TB media, the two parts, nutrients and buffer, were prepared, autoclaved, and stored separately and were combined prior to use. Any antibiotics were added to the media after sterilisation when the liquid was warm to the touch.

For the preparation of LB agar plates, agarose (2% w/v) was added to the above mixture prior to autoclaving. Unless otherwise stated, the TB buffer was adjusted to pH 8.0 was used for expression.

5.1.3 Buffers and Standard Solutions

Unless otherwise stated, all buffers, solutions, and dyes were prepared in dH_2O. Buffers of mixed phosphate were prepared as a mixture of KH_2PO_4 and Na_2HPO_4.

Buffer	Recipe	
McIlvaine Buffer	100 mM	citric acid
		pH 4.5 (Na_2HPO_4)
Reaction Buffer	200 mM	mixed phosphate buffer
		pH 7.5
	1 mM	$MgCl_2$

The antibiotics and inducing agents used were prepared as 1000x stock solutions and were in all cases filter sterilised using a 0.2 µm cellulose acetate sterile syringe filter (VWR).

Stock Solution	Concentration	
Ampicillin (x1000)	100 mg mL^{-1}	(dH_2O)
IPTG (x1000)	1 M	(dH_2O)
Kanamycin (x1000)	10 mg mL^{-1}	(dH_2O)
Tetracycline (x1000)	10 mg mL^{-1}	(EtOH)

5.1.4 Organisms

Escherichia coli

E. coli strains used in this work were supplied by Thermo Fischer Scientific, Inc., New England Biolabs, Inc., or Agilent Technologies, Inc.

Strain	Genotype
BL21(DE3)[304]	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*
DH5α[303]	F⁻ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^- m_K^+$), λ⁻
ER2566[440]	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::mini*Tn*10--*Tet^S*)2 [dcm] R(zgb-::*Tn*10--*Tet^S*) endA1 Δ(mcrC-mrr)114::IS10*
MACH1[441]	F⁻ Φ80(*lacZ*)ΔM15 Δ*lac*X74 *hsd*R($r_K^- m_K^+$) Δ*rec*A1398 *end*A1 *ton*A
XL1-Blue[442]	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q}Z*ΔM15* Tn*10* (Tet^r)]

Aspergillus oryzae

Aspergillus oryzae 63303 was obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) as a freeze-dried culture.

5.1.5 Plasmids, Vectors, and Primers

Empty pET-series vectors were obtained from Novagen. pMI5:T7 and pBBR1MCS3::T7 were kindly provided by Dr. Anita Loeschke of the Institut für Molekulare Enzymtechnologie (IMET) and have been previously published.[308] pUC57 vectors were supplied as carrier vectors for synthesised genes from GenScript. pET22b(+)::eda_{EC} was generously provided by Dr. Thomas Classen of the IBOC and is previously published.[316] Primers were supplied as desalted lyophilisates from Sigma Aldrich.

The following list is a summary of all gene-containing plasmid constructs used in this work, the full details of which can be found in Section A-1.

Empty Vector	Genotype
pET21a(+)	amp^R, P_{T7}, *lacI*
pET22b(+)	amp^R, P_{T7}, *lacI*
pET28a(+)	kan^R, P_{T7}, *lacI*

Plasmid	Description
pET22b::*lacZ*	*lacZ* wild type with *C*-terminal His-tag
pET22b::*lacZ* E537S	*lacZ* E537S mutant (*via* qcSDM) with *C*-terminal His-tag
pET22b::*lacZ* E537S/G794D	*lacZ* E537S/G794D double mutant (*via* qcSDM) with *C*-terminal His-tag
pET28a::*lacZ*	*lacZ* wild type with *C*-terminal His-tag
pET28a::*lacZ* E537S	*lacZ* E537S mutant (*via* qcSDM) with *C*-terminal His-tag
pET28a::*lacZ* E537S/G794D	*lacZ* E537S/G794D mutant (*via* qcSDM) with *C*-terminal His-tag
pET28a*::*lacZ*	pET28a::*lacZ* with tag(-4)att mutant (*via* rthSDM)
pET28a*::*lacZ* E537S	pET28a::*lacZ* E537S with tag(-4)att mutant (*via* rthSDM)
pET28a*::*lacZ* E537S/G794D	pET28a::*lacZ* E537S/G794D with tag(-4)att mutant (*via* rthSDM)
pUC57::*lac4*	*lac4* wild type (codon harmonised)
pET22a::*lac4*	*lac4* wild type (codon harmonised) with *C*-terminal His-tag
pET22a::*lac4* E551G	*lac4* E551G mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lac4* E551A	*lac4* E551A mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lac4* E551L	*lac4* E551L mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lac4* E551S	*lac4* E551S mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lac4* E551Q	*lac4* E551Q mutant (*via* qcSDM) with *C*-terminal His-tag
pUC57::*lacA*	*lacA* wild type (codon harmonised)
pET22a::*lacA*	*lacA* wild type (codon harmonised) with *C*-terminal His-tag
pET22a::*lacA* E298G	*lacA* E298G mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lacA* E298A	*lacA* E298A mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lacA* E298L	*lacA* E298L mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lacA* E298S	*lacA* E298S mutant (*via* qcSDM) with *C*-terminal His-tag
pET22a::*lacA* E298Q	*lacA* E298Q mutant (*via* qcSDM) with *C*-terminal His-tag

Plasmid	Description
pML5::T7[308]	T7 polymerase
pBBR1MCS3::T7[308]	T7 polymerase
pET22b::eda$_{EC}$[316]	Entner-Douderoff aldolase from *E. coli*

All primers are presented in the direction 5' → 3'. For those primers used for site-directed mutagenesis, the mutations are **marked**.

Primer	Sequence
LacZ_XbaI_fwd	CGTCTAGAATGACCATGATTACGGATTCAC
LacZ_XhoI_rev	CGCTCGAGTTTTTGACACCAGACCAACTG
Lac4_E551A_fwd	CTGATCCTGTGT**GCT**TACGGGCATGCCATG
Lac4_E551A_rev	CATGGCATGCCCGTA**AGC**ACACAGGATCAG
Lac4_E551G_fwd	CTGATCCTGTGT**GGC**TACGGGCATGCCATG
Lac4_E551G_rev	CATGGCATGCCCGTA**GCC**ACACAGGATCAG
Lac4_E551L_fwd	CTGATCCTGTGT**TTA**TACGGGCATGCCATG
Lac4_E551L_rev	CATGGCATGCCCGTA**TAA**ACACAGGATCAG
Lac4_E551Q_fwd	CTGATCCTGTGT**CAG**TACGGGCATGCCATG
Lac4_E551Q_rev	CATGGCATGCCCGTA**CTG**ACACAGGATCAG
Lac4_E551S_fwd	CTGATCCTGTGT**AGC**TACGGGCATGCCATG
Lac4_E551S_rev	CATGGCATGCCCGTA**GCT**ACACAGGATCAG
LacA_E298A_fwd	CGCCGTACTCTTTATTA**GCA**TTTCAAGCTGGCG
LacA_E298A_rev	CGCCAGCTTGAAA**TGC**TAATAAAGAGTACGGCG
LacA_E298G_fwd	CGCCGTACTCTTTATTA**GGT**TTTCAAGCTGGCG
LacA_E298G_rev	CGCCAGCTTGAAA**ACC**TAATAAAGAGTACGGCG
LacA_E298L_fwd	CGCCGTACTCTTTATTA**CTG**TTTCAAGCTGGCG
LacA_E298L_rev	CGCCAGCTTGAAA**CAG**TAATAAAGAGTACGGCG
LacA_E298Q_fwd	CGCCGTACTCTTTATTA**CAG**TTTCAAGCTGGCG
LacA_E298Q_rev	CGCCAGCTTGAAA**CTG**TAATAAAGAGTACGGCG
LacA_E298S_fwd	CGCCGTACTCTTTATTA**AGC**TTTCAAGCTGGCG
LacA_E298S_rev	CGCCAGCTTGAAA**GCT**TAATAAAGAGTACGGCG
LacZ_BspHI_insert_fwd	CAATTCCCCTCTA**TCA**TGACCATGATTAC
LacZ_BspHI_insert_rev	GTAATCATGGTCA**TGA**TAGAGGGGAATTG
LacZ_E461S_fwd	CGCTGGGGAAT**AGT**TCAGGCCACG
LacZ_E461S_rev	CGTGGCCTGA**ACT**ATTCCCCAGCG
LacZ_E537S_fwd	CGCTGATCCTTTGC**AGT**TACGCCCACGCGATG
LacZ_E537S_rev	CATCGCGTGGGCGTA**ACT**GCAAAGGATCAGCG
LacZ_G794D_fwd	GGATAACGACATT**GAC**GTAAGTGAAGCGAC
LacZ_G794D_rev	GTCGCTTCACTTAC**GTC**AATGTCGTTATCC
LacZ_AseI_rth_fwd	[phos]**ATT**AATGACCATGATTACGGATTCACTGGCC
LacZ_AseI_rth_rev	[phos]GAGGGGAATTGTTATCCGCTCACAATTCC
Lac4_seqfwd_646	GGATCAGTGGTGGCTAAG
Lac4_seqrev_1252	CCATTATCCGAACCATCC

Lac4_seqrev_2442	CTGCTACCCTCAAATTTC
LacA_seqfwd_597	CGGATAACTACATTGCAAAC
LacA_seqrev_1259	GCTAGTCCGAGCTACTTAAC
LacA_seqrev_2416	GCCATCCATAGCGTAATC
LacZ_seqfwd_1254	CGCTGTTCGCATTATCC
LacZ_seqfwd_593	GCCGTCTGAATTTGACC
LacZ_seqrev_2403	CAATCCACATCTGTGAAAG
pET_RP*	CTAGTTATTGCTCAGCGG
T7*	TAATACGACTCACTATAGGG

*Primers supplied by GATC

5.1.6 Commercial Enzymes

DNA Polymerases

Polymerase	Supplier	Product Description
Herc II	Agilent Technologies, Inc.	Herculase II Fusion DNA Polymerase (1 U μL^{-1})
KAPA HiFi	Kapa Biosystems	KAPA HiFi HotStart DNA Polymerase (1 U μL^{-1})
Pfu	IMET	DNA polymerase from *Pyrococcus furiosus*
Phire II	Thermo Fisher Scientific, Inc.	Phire Hot Start II DNA Polymerase (1 U μL^{-1})
Phusion	New England Biolabs, Inc.	Phusion® High-Fidelity DNA Polymerase (2 U μL^{-1})
Q5	New England Biolabs, Inc.	Q5® Hot Start High-Fidelity DNA Polymerase (2 U μL^{-1})
Taq	Thermo Fisher Scientific, Inc.	Taq DNA Polymerase, native (5 U μL^{-1})
Vent	New England Biolabs, Inc.	Vent® DNA Polymerase (2 U μL^{-1})

Restriction Enzymes

Polymerase	Supplier	Product Description
AseI	Thermo Fisher Scientific, Inc.	*VspI* (*AseI*) (10 U μL^{-1})
BspHI	Thermo Fisher Scientific, Inc.	*PagI* (*BspHI*) (10 U μL^{-1})
DpnI, FD	Thermo Fisher Scientific, Inc.	FastDigest *DpnI* (1 U μL^{-1})
EcoRI	Thermo Fisher Scientific, Inc.	*EcoRI* (10 U μL^{-1})
EcoRV	Thermo Fisher Scientific, Inc.	*EcoRV* (10 U μL^{-1})
HindIII	Thermo Fisher Scientific, Inc.	*HindIII* (10 U μL^{-1})
NcoI	Thermo Fisher Scientific, Inc.	*NcoI* (10 U μL^{-1})
NdeI	Thermo Fisher Scientific, Inc.	*NdeI* (10 U μL^{-1})
PaeI	Thermo Fisher Scientific, Inc.	*PaeI* (*SphI*) (10 U μL^{-1})
RNase A	Thermo Fisher Scientific, Inc.	RNase A, DNase and protease-free (10 mg mL^{-1})
XbaI	Thermo Fisher Scientific, Inc.	*XbaI* (10 U μL^{-1})
XhoI	Thermo Fisher Scientific, Inc.	*XhoI* (10 U μL^{-1})

Other Enzymes

Polymerase	Supplier	Product Description
Chitinase	Zymo Research Corp.	Zymolyase – Yeast Lytic Enzyme
FastAP	Thermo Fisher Scientific, Inc.	FastAP Thermosensitive Alkaline Phosphatase ($1\ U\ \mu L^{-1}$)
T4 Ligase	Thermo Fisher Scientific, Inc.	T4 DNA Ligase ($5\ U\ \mu L^{-1}$)

5.1.7 Culturing Cells

Preculture

The culture was prepared in the desired medium with antibiotic as appropriate. A volume of 5 mL for gene amplification in *E. coli* MACH1 and small-scale protein expression, or 10 mL for large-scale protein expression cultures was used. The culture was then grown overnight at 37 °C on a culture rotator.

Chemical Competency

Buffer	Recipe	
TFB1	30 mM	KOAc
		pH 5.8 (AcOH)
	100 mM	$RbCl_2$
	10 mM	$CaCl_2$
	50 mM	$MnCl_2$
	15%	glycerol (w/v)
		in MilliQ water, filter sterilised
TFB2	10 mM	MOPS
		pH 6.5 (NaOH)
	75 mM	$CaCl_2$
	10 mM	$RbCl_2$
	15%	glycerol (w/v)
		in MilliQ water, filter sterilised

A preculture (5 mL) was transferred into fresh LB (50 mL) with antibiotic as appropriate and the culture incubated on an orbital shaker at 37 °C and 120 rpm. The OD_{600} of the cultures was monitored and once the optical density reached 0.6, the culture was split between 2x 50 mL Falcon™ tubes and centrifuged for 10 min at 3000 × *g* and 4 °C. The supernatant was then discarded and the cells were gently resuspended in ice-cold TFB1 (10 mL, 0.2 volumes) and incubated on ice for 40 min. The samples were then recentrifuged for a further 10 min at 3000 × *g* and 4 °C. The supernatant was again discarded and the pellet was resuspended in ice-cold TFB2 (2 mL, 0.04 volumes) before being incubated on ice for 30 min. The cells were then aliquoted (50 µL) into sterile, labelled Eppendorf tubes and flash frozen in $N_{2\ (l)}$. The chemically competent cells were then stored at -80 °C until required.

Chemically competent cells were tested by transforming two aliquots; one with sH_2O, and the other with an antibiotic resistance-containing vector, and were plated onto LB agar with antibiotic and incubated for 16 h at 37 °C.

Transformation of Cells

An aliquot of chemically competent *E. coli* cells was placed on ice and a plasmid sample (2 µL) was gently swirled through the cell suspension which was then incubated on ice for 30 min. The cells were then heat-shocked at 42 °C for 35 s (BL21(DE3), XL1-Blue) or 90 s (ER2566, MACH1) before being returned to ice. Fresh LB (1 mL) was then added and the culture incubated at 37 °C for 1 h on a culture rotator. Following this the cells were plated according to standard procedures.

Plating Cells

Following transformation, 100 µL of the cell culture was transferred to a LB agar plate (with desired antibiotic or screening substance). The cells were then distributed across the plate by addition of sterile glass beads and gentle shaking of the Petri dish. After the cells had been distributed across the plate, the beads were removed and the plate incubated for 16 h at 37 °C, or for 3 d at room temperature in the absence of light.

Long-Term Cell Storage

For long-term storage of cell cultures, a 1:1 mixture of the cell culture with sterile glycerol was created and up to 1 mL is stored in a 1.8 mL Nunc® CryoTube® at -80°C indefinitely.

5.1.8 Genetic Isolation

Isolation of Genomic DNA from E. coli BL21(DE3) Whole Cells

A pipette tip was scraped across a colony of *E. coli* BL21(DE3) on LB agar and the resulting cells suspended in 500 µL of sterile H_2O to create the template DNA mixture for PCR isolation.

Component	Volume
sH_2O	35 µL
Q5 buffer (5x conc.)	10 µL
dNTPs [10 mM]	1 µL
*lacZ_XbaI*_fwd [10 µM]	1 µL
*lacZ_XhoI*_rev [10 µM]	1 µL
Cell suspension	1 µL
Q5 polymerase	1 µL

The PCR program was as follows at two different elongation (step 3) temperatures, with steps 2-4 repeated 30x:

Step	T [°C]	*t*
1	95	12.5 min
2	95	1 min
3	55 / 60	45 s
4	72	3 min
5	72	5 min
6	10	HOLD

Following the PCR, all reaction product were visualised using agarose gel electrophoresis.

DNA Purification

DNA samples were purified by removal of small nucleotides and primers (≤60 bp) using a DNA purification kit and following the instructions provided.

Plasmid Amplification and Extraction

Plasmids were amplified *in vivo* using *E. coli* MACH1 due to its higher-than-average replication rate.[441] Chemically competent *E. coli* MACH1 cells were transformed using the standard transformation protocol and this transformed culture was used to inoculate a 5 mL preculture in LB with the appropriate antibiotic in a 15 mL Falcon™ tube and incubated for 16 h at 37 °C. The cells were then centrifuged at 12000 rpm for 1 min. The plasmid was then isolated using a plasmid extraction kit using the provided protocols.

Fragment Extraction from Agarose Gel

The agarose gel after electrophoresis was visualised and the target DNA-containing slice of gel was carefully removed and weighed. The extraction of genetic material from agarose gel was performed using a gel extraction kit using the provided protocols.

Codon Optimisation

Codon optimisation was performed using *Microsoft Excel*. The codon frequency data was manually added into the document from the online codon usage database, *Kazusa*. The original coding sequence was imported from the online database of the *NCBI*. Once the gene was optimised, the output was viewed and manipulated with *Clone Manager* and the proposed gene was analysed for restriction sites and highly-repetitive sequences. These undesired elements were, where possible, modified to another codon with a similar usage to prevent unwanted digestion activities.

DNA Sequencing

DNA samples were sent for sequencing to either SEQLAB Sequence Laboratories Göttingen GmbH or to GATC Biotech AG. Sequencing samples were prepared according to guidelines provided by the respective sequencing companies.

5.1.9 Site-Directed Mutagenesis

Following the PCR steps, the products were visualised using agarose gel electrophoresis. The PCR product was then digested using *DpnI* to degrade the template and subsequently used to transform competent *E. coli* MACH1 cells and plated onto appropriate antibiotic-selective agar.

QuikChange® Site-Directed Mutagenesis (qcSDM)

Primer Design

qcSDM requires the design of two complementary primers 25-45 bp in length containing a centrally-located mutation. The mutation is flanked by regions of 10-20 bp so the primer has a T_m of ≥78 °C and the GC content is between 40 and 60%. The design was assisted by the use of the web tool, *PrimerX*, and by the primer design function in *Clone Manager*.

Generation of pET28a::lacZ library

The pET28a::*lacZ* plasmid was mutated in two-steps; first to the single mutant E537S, then subsequently the double mutant, E537S/G794D, with pET28a::*lacZ* E537S being used as the template for the double mutant.

Component	Volume
H_2O	39.5 µL
5x Phusion HF Buffer	5 µL
DMSO	1 µL
dNTPs [10 mM]	1 µL
pET28a::*lacZ* / pET28a::*lacZ* E537S [50 ng]	1 µL
*lacZ*_E537S_fwd / *lacZ*_G794D_fwd [10 µM]	1 µL
*lacZ*_E537S_rev / *lacZ*_G794D_rev [10 µM]	1 µL
Phusion	0.5 µL

The following program was used for mutagenesis, where steps 2-4 were repeated 17x:

Step	T [°C]	t
1	98	30 s
2	95	10 s
3	60	30 s
4	72	8.5 min
5	72	5 min
6	10	HOLD

Generation of pET22b::lac4 library

The plasmid, pET22b::*lac4* was used as a template for the generation of all *lac4* mutants. The mutation of the nucleophilic residue E551 was the only mutant considered. Five alternative residues were chosen, G, A, L, S, and Q.

Component	Volume
H_2O	35 µL
5x KAPA HiFi GC Buffer	10 µL
*lac4*_E551X_fwd [10 µM]	1 µL
*lacA*_E551X_rev [10 µM]	1 µL
pET22b::*lac4* [50 ng]	0.5 µL
dNTPs [10 mM]	1.5 µL
KAPA HiFi	1 µL

The mixture was then submitted to PCR where steps 2-4 were repeated 17x.

Step	T [°C]	*t*
1	98	4 min
2	98	30 s
3	60	30 s
4	72	12 min
5	72	10 min
6	10	HOLD

Generation of pET22b::lacA library

The plasmid, pET22b::*lacA* was used as a template for the generation of all *lacA* mutants. The mutation of the nucleophilic residue E298 was the only mutant considered. Five alternative residues were chosen, G, A, L, S, and Q.

Component	Volume
H_2O	36.5 µL
5x HF Buffer	10 µL
*lacA*_E298X_fwd [10 µM]	0.5 µL
*lacA*_E298X_rev [10 µM]	0.5 µL
pET22b::*lacA* [50 ng]	0.5 µL
dNTPs [10 mM]	1 µL
Phire II	1 µL

The mixture was then submitted to PCR where steps 2-4 were repeated 17x.

Step	T [°C]	*t*
1	98	4 min
2	98	30 s
3	60	30 s
4	72	12 min
5	72	10 min
6	10	HOLD

Round-the-Horn Site-Directed Mutagenesis (rthSDM)

Primer Design

Two primers were designed with the desired mutation located at the 5'-end only one of the primers. The other primer is fully complementary to the plasmid. The primers were designed with a $T_m = 60\pm1$ °C and a length of 30-35 bp. Furthermore, the primers must be phosphorylated at the 5'-terminal to allow for subsequent ligation with T4 DNA ligase.

pET28b::lacZ → pET28b::lacZ*

rthSDM was used to install an *AseI* restriction site 4bp upstream (-4) of the *lacZ* gene to allow for excision and religation of the gene into a new vector. An identical protocol was followed for all three *lacZ* variants.

Component	Volume
H_2O	35 µL
5x KAPA HiFi Fidelity buffer	10 µL
pET28a::*lacZ* constructs [50 ng]	0.5 µL
*lacZ_AseI*_rth_fwd [10 µM]	1 µL
*lacZ_AseI*_rth_rev [10 µM]	1 µL
dNTPs [10 mM]	1.5 µL
KAPA HiFi	1 µL

The mixtures was prepared in PCR tubes where the following program was used, where steps 2-4 were repeated 17x.

Step	T [°C]	*t*
1	98	2 min
2	98	20 s
3	65	30 s
4	72	5 min
5	72	5 min
6	10	HOLD

5.1.10 Digestion

To check the correct insertion of a GOI into a vector, a sample of the DNA was digested by one or more restriction enzymes. These enzymes were selected based on the expected size of the resulting DNA fragments; the size of which is checked *via* gel electrophoresis. The enzymes also needed to be checked for compatibility (DoubleDigest tool; Thermo Scientific). Should the enzymes not be compatible with one another, subsequent digestion steps are required. The sample mixtures are then prepared according to guidelines supplied by the enzyme supplier. Subsequent digestion, the products were visualised using agarose gel electrophoresis. In the case that fragments required isolation, they were isolated from the gel using the Gel Extraction kit.

Template Digestion of a PCR Amplicon

Following SDM, the template DNA was digested by addition of FD *DpnI* (1 µL), which was added directly to the PCR product mixture and subsequently incubated at 37 °C for 1 h. Following incubation, the mixture was then treated at 80 °C for 10 min to inactivate the FD *DpnI*.

Double Digest of pET21a(+) for Subcloning

The empty vector, pET21a(+), was opened for subcloning. FastAP, a phosphatase, was added to sample B to check whether re-ligation of the cut genes presented a problem in terms of subcloning.

Component	Volume A	Volume B
dH_2O	3 µL	2 µL
10x 'O' Buffer	1 µL	1 µL
pET28a(+)::*lacZ*	4 µL	4 µL
XbaI	1 µL	1 µL
EcoRI	1 µL	1 µL
FastAP	--	1 µL

Digestion of pET22b(+)

The empty vector was digested for the insertion of the synthetic *lac4* and *lacA* genes using *NdeI* and *XhoI*.

Component	Volume
sH_2O	8 µL
'O' Buffer	2 µL
pET22b(+)	8 µL
NdeI	1 µL
XhoI	1 µL

Digestion of pET28a(+)

pET28a(+) was digested using *XbaI* and *XhoI* for insertion of *lacZ* into the vector using the following mixture:

Component	Volume
pET28a(+)	4 µL
10x Tango Buffer	2 µL
XbaI	1 µL
XhoI	1 µL
dH_2O	2 µL

Digestion of Isolated lacZ WT

The DNA sample isolated from *E. coli* BL21 (DE3) by whole-cell isolation was digested using *XbaI* and *XhoI* using the following mixture:

Component	Volume
lacZ	10 µL
10x Tango Buffer	6 µL
XbaI	1 µL
XhoI	1 µL
dH_2O	12 µL

Double Digest of pET28a::lacZ: confirmation of transformation

Following the transformation of pET28a::*lacZ* into MACH1, the vector was amplified and purified using the standard protocol and digested.

Component	Volume
dH_2O	2 µL
10x Tango Buffer	2 µL
pET28a(+)::*lacZ*	3 µL
XbaI	2 µL
EcoRI	1 µL

Double Digest of pET28a::lacZ Library: confirmation of gene

Prior to the introduction of a restriction site at the head of the gene for subcloning, the pET28a::*lacZ* library was confirmed *via* digestion with restriction enzymes.

Component	Volume
dH_2O	5 µL
10x Tango Buffer	1 µL
pET28a(+)::*lacZ* library	2 µL
XbaI	1 µL
XhoI	1 µL

Digestion of pET28a::lacZ*

Following the rthSDM insertion of the -4 *AseI* restriction site into the pET28a::*lacZ* mutants, the pET28*::*lacZ* constructs were all digested using *AseI*, *XhoI*, and *BspHI* in an identical fashion and the following protocol was identical for the three variants.

Component	Volume
sH_2O	7 µL
'O' Buffer	5 µL
pUC57::*lac4*	35 µL
AseI	1 µL
XhoI	1 µL
BspHI	1 µL

Digest of pET22b::lacZ Subclones

After subcloning the *lacZ* library into pET22b(+), three colonies were picked from each plate and the plasmid amplified and extracted. The resulting DNA was then digested to confirm insertion of the pET22b::*lacZ* plasmid.

Component	Volume
dH_2O	1.2 µL
10x Tango Buffer	0.5 µL
pET22b::*lacZ*	3 µL
EcoRV	0.3 µL

Digestion of pUC57 Constructs Using FastAP

The synthetic gene constructs, pUC57::*lacA* and pUC57::*lac4*, were digested for subcloning. The digestions were performed identically for each gene, one digestion using an alkaline phosphatase, FastAP, and one without.

Component	Volume A	Volume B
sH_2O	15 µL	14 µL
'O' Buffer	3 µL	3 µL
FastAP	--	1 µL
NdeI	1 µL	1 µL
XhoI	1 µL	1 µL
pET22b(+)	10 µL	10 µL

Digestion of pUC57::lac4

The synthetic gene construct, pUC57::*lac4* was digested for insertion into pET22b(+) using *NdeI* and *XhoI*.

Component	Volume
sH_2O	39 µL
'O' Buffer	5 µL
pUC57::*lac4*	4 µL
NdeI	1 µL
XhoI	1 µL

Digestion of pUC57::lacA

The synthetic gene construct, pUC57::*lacA* was digested for insertion into pET22b(+) using *NdeI* and *XhoI* in an identical fashion to that of the *lac4* gene.

Double Digest of pET21a(+) Constructs: confirmation of vector identity

To check the success of the ligation and subsequent transformation of pET21a::*lacA* and pET21a::*lac4*, the plasmids were amplified and purified using a standard protocol and subsequently digested.

Component	pET21a::*lacA*	pET21a::*lac4*
dH_2O	6.5 µL	6.5 µL
10x 'O' Buffer	1 µL	1 µL
template DNA	0.5 µL	0.5 µL
NdeI	1 µL	1 µL
XhoI	1 µL	1 µL

Double Digest of pET22b::lac4: confirmation of transformation

To confirm successful transformation of pET22b::*lac4* into *E. coli* MACH1 after subcloning, eight clones from each plate were picked and the vector amplified for plasmid extraction. Two separate mixtures were made depending upon whether or not the DNA concentration was less-than-or-equal-to 50 ng μL^{-1} or whether it was greater than 50 ng μL^{-1}.

Component	≤50 ng	>50 ng
sH_2O	5 µL	7 µL
'B' Buffer	1 µL	1 µL
template	4 µL	1 µL
PaeI	1 µL	1 µL

Double Digest of pET22b::lac4 E551 Mutants

After qcSDM of pET22b::*lac4*, three colonies were picked from each plate and the plasmid amplified and extracted. The resulting DNA was then digested to confirm insertion of the pET22b::*lac4* E551X plasmid.

Component	Volume
dH_2O	3.4 µL
10x Tango Buffer	2 µL
pET22b::*lac4* E551X	4 µL
NdeI	0.3 µL
XhoI	0.3 µL

Double Digest of pET22b::lacA: confirmation of transformation

To confirm successful transformation of pET22b::*lac4* into *E. coli* MACH1 after subcloning, eight clones from each plate were picked and the vector amplified for plasmid extraction. Two separate mixtures were made depending upon whether or not the DNA concentration was less-than-or-equal-to 50 ng μL^{-1} or whether it was greater than 50 ng μL^{-1}.

Component	≤50 ng	>50 ng
sH_2O	4 µL	6 µL
'R' Buffer	1 µL	1 µL
template	3 µL	1 µL
HindIII	1 µL	1 µL
XhoI	1 µL	1 µL

Double Digest of pET22b::lacA E298 Mutants

After qcSDM of pET22b::*lacA*, three colonies were picked from each plate and the plasmid amplified and extracted. The resulting DNA was then digested to confirm insertion of the pET22b::*lacA* E298X plasmid.

Component	Volume
dH_2O	3.4 µL
10x Tango Buffer	2 µL
pET22b::*lacA* E298X	4 µL
NdeI	0.3 µL
XhoI	0.3 µL

5.1.11 Ligation

Following ligation, the product mixture was used to transform chemically competent *E. coli* MACH1 cells and then plated onto antibiotic-selective LB agar.

Ligation of rthSDM Products

After purification of DNA generated from round-the-horn SDM, the linear DNA was circularised using the following protocol:

Component	Volume
H_2O	7 µL
T4 DNA Ligase Buffer	2 µL
linearised DNA	10 µL
T4 DNA Ligase	1 µL

Ligation of lacZ into pET28a(+)

After purification of the digested DNA from the agarose gel, the *lacZ* gene and the cut plasmid were ligated using the following procedure:

Component	Volume
H_2O	1 µL
T4 DNA Ligase Buffer	1 µL
pET28a(+)	2 µL
lacZ	5 µL
T4 DNA Ligase	1 µL

Ligation of lacZ into pET22b(+)

After purification of digested *lacZ*, *lacZ* E537S, and *lacZ* E537S/G794D and pET22b(+) from an agarose electrophoresis gel, the two fragments were ligated. They were ligated against a control and mixed with an excess of the gene fragment (molar ratio given) to minimise self-religation of the cut vector.

Component	control	1:5
sH_2O	7 µL	3 µL
T4 DNA Ligase Buffer	1 µL	1 µL
pET22b(+)	1 µL	1 µL
lacZ	--	4 µL
T4 DNA Ligase	1 µL	1 µL

Ligation of lacA and lac4 into pET21b(+)

After purification of the digested *lacA*, *lac4*, and pET21b(+) from an agarose electrophoresis gel, the two fragments were ligated in a 1:8 molar ratio of vector:gene to minimise self-ligation of the vector.

Component	Volume
sH_2O	3 µL
T4 DNA Ligase buffer	1 µL
pET21b(+)	1 µL
purified gene	4 µL
T4 DNA Ligase	1 µL

Ligation of lac4 into pET22b(+)

After purification of the digested *lac4* and pET22b(+) from an agarose electrophoresis gel, the two fragments were ligated. They were ligated against a control (addition of sH_2O instead of gene) and mixed with an excess of the gene (molar ratios given) to minimise self-religation of the cut vector.

Component	control	1:8	1:16
sH_2O	7 µL	5 µL	3 µL
T4 DNA Ligase buffer	1 µL	1 µL	1 µL
pET22b(+)	1 µL	1 µL	1 µL
lac4	--	2 µL	4 µL
T4 DNA Ligase	1 µL	1 µL	1 µL

Ligation of lacA into pET22b(+)

After purification of the digested *lacA* and pET22b(+) from an agarose electrophoresis gel, the two fragments were ligated in an analogous fashion as for *lac4*.

5.1.12 Protein Expression

All expression cultures were inoculated from a preculture which had been grown overnight. The shaking cultures were performed in sterile, baffled, 500 mL Erlenmeyer flasks, or sterile, 3 L Fernbach flasks. Unless stated otherwise, the cultures were incubated at 37 °C and shaken at 120 rpm. They were induced at various time points using IPTG and incubated for a total of 24 hours before being harvested *via* centrifugation at 4 °C and 10000 rpm for 30 min. The supernatant was decanted off and the pellet collected. Pellets were stored at -20 °C until required.

96-Well Deep Well Plate Expression

This method was developed for the screening of growth conditions for the expression of LacZ WT in *E. coli* ER2566. The temperature, pH of KP_i, time of induction, and concentration of IPTG used for induction was varied according to the reaction parameters described in Section 3.1.4 .

To a sterile 96-well deep well plate (2 mL capacity) was added LB expression media containing ampicillin (0.1% v/v, 990 µL). Each well was the inoculated with a diluted sample of the preculture (OD_{600} = 0.01, 10 µL) and the plate transferred to an orbital shaker. The cells were then grown over a 24 h period at 1000 rpm and induced with IPTG at various time intervals.

Small-scale Shaking Culture for Growth Optimisation

The best-performing growth conditions from the 96-well deep well plate assay were assessed for their improvement over the initial conditions by scaling up to a 100 mL expression.

Control

The control replicates the pre-optimisation growth conditions. TB (pH 7.5, 100 mL), ampicillin stock (100 µL), and preculture (1 mL) were combined and incubated at 25 °C for 8 h at which point it was induced with IPTG (10 µL).

Optimisation

TB (100 mL), ampicillin stock (100 µL), and preculture (1 mL) were incubated until the $OD_{600} \approx 1.2$ ($t \approx 4.5$ h), at which point it was induced with IPTG.

Large-scale Expression

Original 24-Hour Method

TB (pH 7.5, 1 L), antibiotic stock (1 mL), and preculture (10 mL) were incubated at 25 °C for 8 h before induction with IPTG stock (100 µL).

Improved 24-Hour Method

TB (1 L), antibiotic stock (1 mL), and preculture (10 mL) were combined in a sterile, baffled, 3L Fernbach flask. The culture was incubated for 4.5 h before induction with IPTG stock (600 µL).

5.1.13 Lysis

Buffer	Recipe	
Lysis Buffer	50 mM	mixed phosphate buffer
	300 mM	NaCl
		pH 7.5
Lysozyme Buffer	50 mM	K-phosphate buffer
		pH 8.0
	25 mM	NaCl

Mechanical Cell Lysis

Cell Lysis via French-Press

Cells were resuspended in Lysis Buffer (20% w/v) and cooled in an ice-bath. They were opened by passing the cells three times through a French-press at 800 psi. The resulting cell slurry was then centrifuged at 10000 rpm for 30 min at 4 °C.

Cell Lysis via Sonication

Cells were resuspended in Lysis Buffer (20% w/v) and cooled in an ice-bath. They were opened by sonication for 2x 10 min at 50% power. After which time, the cells were centrifuged at 10000 rpm for 30 min at 4 °C.

Enzymatic Cell Lysis

Cells were weighed into a 50 mL Falcon™ tube and resuspended in Lysozyme Buffer (20% w/v) and placed in an ice-bath. A solution of lysozyme in Lysozyme Buffer (10 mg mL^{-1}) was then pipetted into the Falcon™ tube so the final concentration of lysozyme was 1 mg mL^{-1}. This was then left to react for between 30 min and 1 h. After which time, the cells were centrifuged at 10000 rpm for 30 min at 4 °C.

5.1.14 Protein Purification

Cell cultures used for expression can either be used fresh or stored at -20 °C until required.

Ni-NTA Affinity Chromatography

The columns used for NiNTA chromatography packed with Ni-NTA Agarose from QIAGEN GmbH and are stored wet in 20% $EtOH_{(aq)}$ at 4 °C.

Column Regeneration

Buffer	Recipe	
Stripping Buffer	50 mM	Na-phosphate buffer
		pH 8.0
	300 mM	NaCl
	100 mM	EDTA
		pH 8.0 (NaOH)

The columns were first stripped of nickel and trace metal residues using Stripping Buffer (8 CV, 5.0 mL min^{-1}). The column was then regenerated using a 100 mM solution of $NiCl_{2\ (aq)}$ (8 CV, 2.7 mL min^{-1}) and subsequently washed of any unbound nickel with dH_2O (4 CV, 5.0 mL min^{-1}).

Column Loading

Buffer	Recipe	
Equilibration Buffer	50 mM	mixed phosphate buffer
		pH 8.0
	1 M	NaCl
	10 mM	imid.

The column was connected to a peristaltic pump and the crude lysate was stored on ice. Onto a column which had been equilibrated using Equilibration Buffer or Äkta Buffer A (6 CV, 4.1 mL min^{-1}) was loaded the crude lysate. For loading, the column was placed directly above the Falcon™ tube containing the crude lysate so as to create a circuit; ensuring no loss of enzyme during the loading process. The loading was performed at 0 °C for 45 min at 1.8 mL min^{-1}.

Enzyme Purification (Manual)

Buffer	Recipe	
Wash Buffer	50 mM	mixed phosphate buffer
		pH 8.0
	1 M	NaCl
	20 mM	imid.
Elution Buffer	50 mM	mixed phosphate buffer
		pH 8.0
	300 mM	NaCl
	250 mM	imid.
Cleanse Buffer	50 mM	mixed phosphate buffer
		pH 8.0
	300 mM	NaCl
	1 M	imid.

After loading, the column was washed with Wash Buffer (6 CV, 4.1 mL min^{-1}) to elute any unbound and non-specifically bound proteins. The POI was eluted from the column using Elution Buffer (2.5 CV, 4.1 mL min^{-1}) and collected and the buffer changed using a PD10 column. The NiNTA column was then washed with Cleansing Buffer (3 CV, 5.0 mL min^{-1}), then with dH_2O (6 CV, 5.0 mL min^{-1}) and 20% $EtOH_{(aq)}$ (6 CV, 5.0 mL min^{-1}).

Enzyme Purification (Åkta)

Buffer	Recipe	
Äkta Buffer A	50 mM	Na-phosphate buffer
	300 mM	NaCl
	10 mM	imid.
		pH 8.0
		degassed and *filtered*
Äkta Buffer B	50 mM	Na-phosphate buffer
	50 mM	NaCl
	2 M	imid.
		pH 8.0
		degassed and *filtered*

After loading, the column was attached to the Äkta and the column was exposed to the following conditions:

Step	CV	Flow Rate	Äkta Buffers A:B	[imid.]
1	5	5 mL min^{-1}	100:0	10 mM
2	6	5 mL min^{-1}	98.5:1.5	40 mM
3	6	5 mL min^{-1}	87.5:12.5	259 mM
4	2	5 mL min^{-1}	0:100	2 M

The absorption of the solution at λ=280 nm was measured and indicated the elution of protein-containing fractions. Fractions were collected from step 3, and these were found to contain the POI. The POI-containing fractions were then pooled and concentrated and the buffer exchanged using a PD10 column. The NiNTA column was then washed with dH_2O (6 CV, 5.0 mL min^{-1}) and 20% $EtOH_{(aq)}$ (6 CV, 5.0 mL min^{-1}).

PD-10 Desalting Columns

Following equilibration of the PD-10 columns (Disposable PD-10 Desalting Column, GE Life Sciences) with 8 CV of Reaction Buffer, no more than 2 mL of the POI-containing solution. The POI was then eluted from the column using 3.5 mL of Reaction Buffer and collected in a 15 mL Falcon™ tube and stored at 4 °C until required.

Concentrating Protein

Concentration of the POI was obtained by ultracentrifugation (Vivaspin 20 MWCO 50000, GE Healthcare Life Sciences). Centrifugation at 8000 rpm for 20 minutes at 10 °C typically resulted in concentration of the POI-containing solution to ca. 1 mL. If the sample was not concentrated to the desired volume after this step, it was re-centrifuged using the same settings until the desired volume was achieved.

5.2 Biological Analytical Protocols

5.2.1 DNA Concentration

A_{260} Assay for DNA Concentration Determination

The concentration of DNA in an isolated sample was determined by absorption at 260 nm using a NanoDrop spectrophotometer.

5.2.2 Cell Density

OD_{600}

The cell density in a culture sample was determined by absorption at 600 nm using a NanoDrop spectrophotometer.

96-Well Plate-based OD_{600} Measurement

This assay was designed to allow for high-throughput determination of the OD_{600} of a cell culture.

A culture of *E. coli* ER2566 with the pET22b(+) plasmid was grown at 37 °C on a culture rotator overnight. The culture was then removed, and the OD_{600} measured using a NanoDrop spectrophotometer to generate the "true" OD_{600}. From this culture of known optical density was created a series of dilutions using H_2O. This allowed for the generation of a calibration curve for measurement of OD in a 96-well plate with the Tecan spectrophotometer at 600 nm.

5.2.3 Protein Concentration

Bradford Assay for Protein Concentration Determination

The concentration of protein in a solution was measured by the ratio of absorbance of the enzyme-dye complex at 594 nm to the absorbance of the free dye at 466 nm, shown to increase sensitivity of the assay.[443] A calibration curve was constructed using BSA at various concentrations and a linear relationship between concentration and absorbance ratio 594/466 for enzyme concentrations of up to 300 μg mL^{-1} was observed ($R^2 = 0.9926$).

A_{280} Assay for Protein Concentration Determination

The concentration of a sample of purified enzyme was determined using the absorbance at 280 nm, which corresponds to the absorbance of aromatic residues in a protein. The concentration of the solution is then determined using Beer's law:

$$C = \frac{A_{280}}{\varepsilon \times l}$$

Where l is the pathlength of the spectrophotometer in cm.

The extinction coefficient (ε) and the MW of the POI (Da) were estimated using the online tool, *Peptide Property Calculator*. These values were then entered into the NanoDrop's automated protein concentration program and the concentration of the protein was subsequently measured.

5.2.4 Colony PCR

Colony PCR of pET28a::lacZ

Colony PCR was performed to check if *lacZ* was correctly inserted into the vector, pET28a(+) and that the vector-gene construct was taken up by the cell. For this, 7 colonies were selected and suspended in 10 µL sH2O and the PCR samples were prepared as follows:

Component	Volume
sH_2O	4 µL
RedTaq	5 µL
*lacZ_Xba1*_fwd [10 µM]	0.3 µL
*lacZ_Xho1*_rev [10 µM]	0.3 µL
Colony Suspension	0.4 µL

The following PCR program was used with steps 2-4 repeated 30x:

Step	T [°C]	*T*
1	95	12.5 min
2	95	30 s
3	60	30 s
4	72	3 min
5	72	4.5 min
6	10	HOLD

The PCR products were then visualised by agarose gel electrophoresis.

5.2.5 Agarose Gel Electrophoresis

All agarose gels used GeneRuler™ 100 bp DNA Ladder (Thermo Fisher Scientific), GeneRuler™ 1 kb DNA Ladder (Thermo Fisher Scientific), or 1 kb DNA Ladder (New England Biolabs, Inc.) as markers.

Sample Preparation

Samples for gel electrophoresis are easily prepared by mixing 5 parts of target DNA solution with 1 part DNA Loading Dye (6x conc., Thermo Fisher Scientific, Inc.). The sample, once thoroughly mixed, is then ready for electrophoresis.

Gel Preparation

Buffer	Recipe	
Tris-Acetate Buffer	40 mM	Tris
	19 mM	AcOH
	1 mM	EDTA
		pH 8.0

Agarose gels for electrophoresis should be freshly prepared before use. A 0.8% agarose gel was prepared in tris-acetate buffer using the following method:

Component	Concentration
Agarose	0.8% w/v
GelRed™	0.1‰ v/v

Once homogeneously mixed (may require warming), the mixture was poured carefully into a gel mould sealed at each end with masking tape, and a well-comb inserted at one end. It is important during gel pouring not to allow air bubbles to form, as this will disturb the separation of the nucleic acids. The gel is then left at room temperature until it is properly set. Once the gel is set, the masking tape is removed before carefully lowering the gel into the electrophoresis chamber.

Running a Gel and Visualisation

A 0.8% w/v agarose gel was prepared and placed into the electrophoresis chamber oriented so that the wells were at the cathode end of the chamber, as the negatively charged DNA will migrate to the positive anode. The electrophoresis buffer is then gently poured into the chamber until the gel is completely submerged. It is at this point the DNA sample, as well as any markers, are carefully added to each well with a Hamilton syringe, taking care not to overfill the wells. The gels were subsequently run at 180 V for 25 min, at which point the current was switched off, and the gels visualised using under UV light.

5.2.6 SDS-PAGE

All SDS-PAGEs used Roti®-Mark 10-150 (Carl Roth) and/or PageRuler™ Prestained Protein Ladder (10 to 180kDa, Thermo Fisher Scientific) as markers.

SDS-PAGE Preparation

Buffer	Recipe	
Collection Gel Buffer	0.5 M	Tris/HCl
		pH 6.8
SDS Collection Gel (4.5% mixture)	27.4 mL	SDS (10% w/v)
	66.7 mL	Collection Gel Buffer
		to 250 mL with dH_2O
Separation Gel Buffer	1.5 M	Tris/HCl
		pH 8.8
SDS Separation Gel (12% mixture)	4.1 mL	SDS (10% w/v)
	102.2 mL	Separation Gel Buffer
		to 250 mL with dH_2O

The SDS-PAGE gel is made in two parts. The first part, the lower part of the gel, is the separation gel. The separation properties of the gel can be modified by altering the polyacrylamide content, whereby

thinner gels allow for better separation of higher MW proteins, and inversely, higher acrylamide contents separate low MW proteins better. The separation layer is made as follows (makes 4 gels):

Component	10%	12%
acrylamide/bisacrylamide (30%)	6.68 mL	7.93 mL
SDS Separation Gel (12% mix.)	13.21 mL	11.96 mL
TEMED	50 µL	50 µL
APS [10% w/v]	100 µL	100 µL

It is important to work quickly whilst preparing these gels as the polymerisation reaction occurs readily upon addition of the APS solution. The separation gel is carefully pipetted into the SDS-PAGE mould, ensuring a 1.5 cm at the top for the collection gel (ca. 4 mL for each gel). After addition of the separation gel to the mould, it is covered in ^{i}PrOH and allowed to set.

Once set, the collection gel is prepared. The collection gel has a low acrylamide content and a high SDS content, as this allows for the well contents to collect at the top of the separation gel, allowing for samples of different volumes to be directly comparable to one another and reduces sample spreading. The collection gel is made according to the following method (makes 4 gels):

Component	Volume
acrylamide/bisacrylamide (30%)	1.2 mL
SDS Collection Gel (4.5% mix.)	6.84 mL
TEMED	50 µL
APS [10% w/v]	100 µL

After removal of ^{i}PrOH from the top of the now-set separation gel, the collection gel is added to the top of the SDS-PAGE mould, followed by quick, yet careful insertion of the well comb. Once the comb has been inserted, the gel is again left to set. Once the gel is set, it can be used freshly, or stored in a sealed, moist container at 4 °C. The gel remains viable for up to 2 weeks under these conditions.

SDS-PAGE Sample Preparation

Stock Solution	Recipe	
Loading Dye (5x)	100 mM	EDTA
	43%	glycerol (v/v)
	0.05%	Bromophenol blue (v/v)

Samples for SDS are prepared in a volume of 75 µL, with 20% (v/v) of the mixture being the SDS loading dye. The remaining sample is a mix of the sample diluted in dH_2O; as follows:

Component	Volume
protein sample	x µL
dH_2O	60-x µL
Loading Dye (5x)	15 µL

The mixture is then thermally denatured at 99 °C for 20 min and it is at this point ready for SDS-PAGE analysis.

Running an SDS-PAGE

Buffer	Recipe	
SDS Running Buffer	25 mM	Tris
	192 mM	glycine
	0.1%	SDS (w/v)

The electrophoresis chamber is prepared firstly by firmly attaching the SDS-PAGE to the cell. SDS Running Buffer is then added to the chamber so it is just overflowing the cell. The area of the chamber surrounding the cell is filled to the line marked on the exterior of the chamber. The sample comb is then carefully removed from the gel, ensuring there is no tearing of the wells. A small aliquot of the sample(s) (ca. 10 µL) was then added to the well using a Hamilton syringe and run against a protein marker.

Once all samples have been loaded in the wells, the SDS-PAGE was run at 150 V for 18 min, then directly again at 200 V for 45 min. The loading dye contains bromophenol blue, which is eluted before proteins under SDS-PAGE running. This tracking dye allowed for visual determination of when the current to the gel should be switched off – when the dye reaches the bottom of the gel.

After running, the gel was removed from the chamber and the mould removed. The collection gel was removed from the separation gel, and the separation gel was carefully rinsed with dH_2O. After rinsing with water, the gel was ready for staining.

Staining with Coomassie Brilliant Blue G-250

Standard Solution	Recipe	
Fixing Solution	30%	EtOH (v/v)
	10%	AcOH (v/v)
Colloidal Coomassie Solution	5%	$Al_2(SO_4)_3 \cdot 16H_2O$ (w/v)
	0.02%	Coomassie brilliant blue G250 (w/v)
	10%	EtOH (v/v)
	2%	H_3PO_4 (v/v)

After running the SDS-PAGE, the gel was first submersed in Fixing Solution and incubated at room temperature for 10 min on an orbital shaker at low speed and subsequently rinsed with dH_2O. The gel was then submersed in Colloidal Coomassie Solution and incubated overnight at room temperature on an orbital shaker at low speed. The Coomassie solution was the carefully poured off and the gel was once more rinsed with dH_2O. A digital image of the gel was then created.

Staining with Silver Stain

Standard Solution	Recipe	
Fixing Solution	30%	EtOH (v/v)
	10%	AcOH (v/v)
Incubating Solution	30%	EtOH (v/v)
	0.5%	NaOAc (w/v)
	0.2%	$Na_2S_2O_3$ (w/v)
Silver Nitrate Solution	0.1%	$AgNO_3$ (w/v)
Developing Solution	2.5%	Na_2CO_3 (w/v)
Stop Solution	2.3 M	citric acid

After running the SDS-PAGE, the gel was submerged in Fixing Solution and incubated for 10 min at room temperature on an orbital shaker at low speed. After fixing, the gel was rinsed with dH_2O before being incubated in Incubating Solution for 10 min at room temperature on an orbital shaker at low speed. The gel was subsequently rinsed with dH_2O before being incubated 3x in dH_2O for 10 min each time at room temperature with gentle shaking. After removal of the water, the gel was submerged in Silver Nitrate Solution containing 0.2% CH_2O (v/v) which was added prior to use. This was incubated for 10 min at room temperature with gentle shaking before the $AgNO_3$ solution was removed and the gel rinsed with water. The gel was then developed for 10 min at room temperature with gentle shaking on an orbital shaker using Developing Solution containing 0.2% CH_2O (v/v) which was added prior to use. After development, Stop Solution (1V) was added to the developing tray. After removal of the liquid, the gel was digitally imaged.

5.3 Enzyme Activity Assays

5.3.1 X-gal Assay (Blue-White Screening)

Standard Solution	Concentration	
X-gal Mix	2%	X-gal (DMSO)

Plate-based Assay

A sterile LB agar plate with the desired antibiotic selection is taken and, whilst ensuring sterility, IPTG stock (20 µL) and X-gal Mix (100 µL) were added to the plates. The IPTG/X-gal mixture was then evenly spread across the surface of the agarose and the plate then incubated upside-down at 37 °C for 1 h. The plate was then inoculated normally and the cells allowed to grow overnight at 37 °C.

E. coli transformed with pET22b::eda_{EC} was used as a negative control, which would mimic the stress the cell undergoes during protein overexpression. *E. coli* transformed with either pET28a(+) or pET22b(+) containing the respective WT gene was used as the positive control.

Eppendorf® tube-based Assay

This qualitative test is scalable and works with both purified enzyme solutions and crude lysate mixtures. The results are always compared against a blank sample whereby the enzyme sample is substituted for an appropriate buffer in the following ratio:

Component	Volume
dH_2O	890 µL
Enzyme sample	100 µL
X-gal Mix	10 µL

The samples are then incubated for 2 h and room temperature

E. coli transformed with pET22b::edaEC was used as a negative control which would mimic the stress the cell undergoes during protein overexpression. *E. coli* transformed with either pET28a(+) or pET22b(+) containing the respective WT gene was used as the positive control.

5.3.2 96-Well Plate-based pNP Hydrolysis Assay

The hydrolysis assay was performed using a plate-reading spectrophotometer and was used to qualitatively determine the hydrolytic activity of an enzyme by measuring the increase in absorption at 405 nm, which corresponds to the absorption maxima of the *para*-nitrophenolate anion. Each well contained the following mixture: POI-containing solution (40 µL), gal-pNP (20 µL, 10 mM), and Reaction Buffer (140 µL). These were measured in triplicate and referenced to calibration standards and controls.

Ten measurements at 12 s intervals were then taken at 405 nm at 25±1 °C. This yielded an absorption gradient whereby a positive gradient indicated hydrolytic activity. The results were converted from A.U. to µM using a calibration curve which was constructed using pNP in Reaction Buffer at various concentrations which showed a linear correlation between absorbance at 405 nm and the concentration of pNP^- for concentrations up to 500 µM (R^2 = 0.9996). This allowed for determination of activity in the form µM min^{-1}.

When the concentration of the POI-containing solution is also known, the specific activity (nmol min^{-1} mg^{-1} or U mg^{-1}) can also be determined using the following equation:

$$S.A. = \frac{m \times V_{well}}{[enz.] \times V_{enz.}}$$

Where m is the gradient (µM min^{-1}), V_{well} is the total volume of the well (mL), [*enz.*] is the concentration of the POI-containing solution (mg mL^{-1}), and $V_{enz.}$ is the volume of POI-containing solution in the well.

The approximate error of the S.A. was calculated using the following equation:

$$S.A._{err} = \left(\frac{(m + m_{err}) \times V_{well}}{[enz.] \times V_{enz.}}\right) - S.A.$$

The slope and associated error were determined by standard linear regression using *Microsoft Excel 2010*.

This method was also used to determine the Michaelis-Menten kinetic parameters of an enzyme by varying concentrations of the substrate in each well triplicate and plotting the slope against substrate concentration. The following formula was used for the standard determination of K_M and V_{max}:

$$y = \frac{V_{max} \times x}{K_M + x}$$

Where y is the slope in μM sec^{-1} and x is the concentration of the substrate in μM.

The following formula was used for the determination of K_M and V_{max} with a modifier for substrate inhibition:[444]

$$y = \frac{V_{max}}{1 + \frac{K_M}{x} + \frac{x}{k_i}}$$

Where y is the slope in μM sec^{-1} and x is the concentration of the substrate in μM.

The values for slope and concentration were entered into *Origin*[422] and the parameters K_M, K_i, and V_{max} were determined by the program.

5.3.3 96-Well Plate-based pNP Synthase Assay

The activity of the synthase enzymes was determined *via* use of a non-continuous assay. By determining the amount of F^- released through autohydrolysis of the substrate, galF, and comparing that to the amount of F^- released when the enzyme is present over a set time period, the difference is then taken to be the enzyme activity. The concentration of F^- is determined colourimetrically by measurement at 405 nm, the wavelength corresponding to the *para*-nitrophenolate anion, which is protected by the fluoride-labile TIPS group. As the enzyme-catalysed reaction is disubstratic, the K_M is dependent upon the concentrations of both. Therefore, this assay is performed as a mixture of two substrate gradients – one gradient for the acceptor, galF, and the other for the donor, glc-pNP – and the output is a 3-dimensional plane from which the K_M for both substrates is determined.

Each sample was measured in triplicate and contained the two substrates in various concentrations. These were compared to the background rate of reaction. To these wells was added either enzyme (reaction) or buffer (background) with the total volume for each well being 200 μL. The plate was

incubated at 25 °C for 15 min, after which time, 100 µL was transferred to another 96-well plate. To each well was added TIPSpNP in MeCN (2 mM, 100 µL) and this was allowed to react for a further 15 min. 100 µL of the MeCN/buffer mixture was then carefully transferred to a new 96-well plate and to each well was subsequently added $Na_2CO_{3\ (aq)}$ (100 mM, 100 µL) and the plate was measured at 405 nm using a spectrophotometer. The values were entered into *Origin*[422] and the 2D Michaelis-Menten kinetics were determined.

5.3.4 ^{1}H NMR Reaction Kinetics

To an NMR tube was added galF and glc-pNP in various concentrations. To this was added DMS (1.0 mg) as an internal standard. The components were then dissolved in deuterated Reaction Buffer and enzyme dissolved in non-deuterated Reaction Buffer so that the total volume was 500 µL. The samples were then measured at 25 °C using a solvent-suppressing ^{1}H method (64 scans, D1=4 s). The integral of the anomeric peak of galF ($\delta_{ppm} = 5.73$) was compared to the integral of DMS ($\delta_{ppm} = 3.16$) and this ratio was used to determine the concentration of galF in solution for any given time point. Kinetic measurements were recorded over 3 h.

5.4 Chemistry Protocols

5.4.1 General Experimental Conditions

All chemicals and general equipment used in this work were purchased from Sigma-Aldrich Co. LLC., Carl Roth GmbH & Co. KG, Merck KGaA, Carbosynth Ltd., or VWR International.

According to standard procedures, all solvents were either: dried and distilled either immediately prior to use,[445] obtained after distillation by a solvent drying system from MBraun (Trocknungssystem MB SPS-800), or stored appropriately.

Nuclear Magnetic Resonance (NMR) Spectroscopy

^{1}H NMR spectra were recorded at 298 K on a Bruker Avance/DRX 600 operating at 600 MHz, or a Bruker ARX 300 operating at 300 MHz. Chemical shifts are reported in ppm (δ) and the experiments run in deuterated solvents and are referenced to the central, residual proton resonance (Table 5-1). ^{1}H resonance multiplicities are described using the following abbreviations; s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and combinations thereof for systems of higher coupling. ^{1}H signals are described by chemical shift δ (ppm) (multiplicity, |*J* (Hz)|, integration, assignment).

^{13}C NMR spectra were recorded at 298 K on a Bruker Avance/DRX 600 operating at 151 MHz. The proton-uncoupled signals are described by chemical shift δ (ppm) (multiplicity, |*J* (Hz)|, assignment), relative to the central peak of the solvent-dependent reference. Multiplicity is only included for ^{13}C-^{19}F coupled signals.

^{19}F NMR spectra were recorded at 298 K on a Bruker ARX 300 operating at 282 MHz. The signals are described by chemical shift δ (ppm) (assignment).

Signal assignment was assisted with the following 2D-NMR experiments: ^{1}H-^{1}H COSY, ^{1}H-^{1}H NOESY, ^{1}H-^{13}C HSQC, ^{1}H-^{13}C HMBC. NMR kinetics experiments were referenced to the internal standards rather than the solvent signal.

Table 5-1: NMR reference data with values given in δ (ppm).[446]

Solvent	^{1}H	^{13}C	^{19}F
$CDCl_3$	7.26	77.16	-
CD_3OD	3.31	49.00	-
d_6-DMSO	2.50	39.52	-
d_3-MeCN	1.94	1.32	-
		118.26	
D_2O	4.79	-	-
Internal Standards			
DMS	3.16	-	-
$NaC(O)CF_3$	-	-	-75.55

High-Performance Liquid Chromatography (HPLC)

HPLC-MS/MS (tandem mass spectrometry)

Samples were measured using an Agilent 1260 Affinity liquid chromatography system using a SeQuant® ZIC®-pHILIC column (5 μM, polymeric, 150 x 2.1 mm) and acetonitrile/NH_4Ac (10 mM, pH 9.2) as the eluent. Tandem mass spectrometry data was recorded on an AB Sciex TripleTOF™ 6600 system using positive ion detection. Mass spectra are recorded in mass/charge ratios (*m*/*z*) and relative abundance (% base peak intensity).

HPLC-UV-Vis

Samples were measured using a Dionex™ UltiMate® 3000 liquid chromatography system using a Chiralpak® IC column (250 x 4.6 mm) and ^{n}Hept/^{i}PrOH (7:3) as the eluent. UV-Vis spectra were recorded using a Dionex™ UltiMate™ 3000 Diode Array Detector at 288 nm.

Infrared (IR) Spectroscopy

IR spectra were recorded on a PerkinElmer SpectrumOne IR spectrometer and peaks given in wavenumbers, $\tilde{\nu}$ (cm^{-1}). Samples were measured neat and peak strength is described using the following abbreviations: s (strong), m (medium), w (weak), br (broad).

Optical Rotation

Optical rotation was measured on a PerkinElmer Model 341 Polarimeter at 298 K. Concentrations are expressed as g dL^{-1} in the indicated solvent.

Melting Point

Melting points (m.p. (°C)) are uncorrected and were recorded on a Stuart® SMP3 melting point apparatus.

Gas Chromatography (GC)

GCEI-MS samples were measured using a Hewlett-Packard HP 6890 Series gas chromatography system using an Optima-5-MS column (30 m x 0.25 mm, 5% diphenyl, 95% dimethylpolysiloxane, 0.25 μm) with He as the carrier gas (1 mL min^{-1}) using positive ion detection. Mass spectra were recorded on a HP 5973 Mass Selective Detector using electron impact (EI) ionisation at 70 eV. Mass spectra are recorded in mass/charge ratio (*m/z*) and relative abundance (% base peak intensity). Two methods were used, denoted either R_t (A) or (B). Method A is based upon a literature-published protocol[389] – initial temperature: 200 °C, initial time and solvent delay: 3 min. Oven ramp at 15 °C min^{-1} to 270 °C; then oven ramp at 1 °C min^{-1} to 290 °C; then oven ramp at 15 °C min^{-1} to 325 °C for 5 min. Method B – initial temperature: 100 °C, initial time and solvent delay: 3 min. Oven ramp at 5 °C min^{-1} to 215 °C; oven ramp at 3 °C min^{-1} to 300 °C; oven ramp at 10 °C min^{-1} to 325 °C for 5 min.

Other Chromatography

Thin-layer chromatography (TLC) was run on either normal phase or reversed-phase (RP) TLC plates and run in the eluting system described for each plate. Unless otherwise stated, it is assumed that TLC analysis was performed on normal phase plates. Normal phase TLC was run on POLYGRAM® SIL G/UV_{254} pre-coated polyester sheets (40 x 80 mm; 0.2 mm silica gel with fluorescent indicator) and RP-TLC run on ALUGRAM® RP-18 W/UV_{524} pre-coated aluminium sheets (40 x 80 mm; 0.15 mm silica gel C_{18} with fluorescent indicator). Plates were viewed under UV light (254 nm) and then developed in ceric phosphomolybdic acid (12.5 g phosphomolybdic acid, 5.0 g $Ce(SO_4)_2 \cdot 4\ H_2O$ in 6% H_2SO_4/H_2O (v/v, 500 mL)) or vanillin (12 g in 1% H_2SO_4/EtOH (v/v, 200 mL)).

Preparative thin-layer chromatography (pTLC) was run on pre-coated TLC plates silica gel 40 F-254 (20 x 20 cm, glass backed) in the eluting system described for each plate and visualised under UV light (254 nm).

Column chromatography were run on Silica gel 60 (Merck KGaA; 0.040-0.063 mm, 230-400 mesh) using the eluent system described for each separation.

Software

MestReNova[447]	v 8.0.1-10878
ChemDraw[448]	v 15.0.0.106
Origin[422]	v 9.1.0 (64-bit) Sr2

5.4.2 General Procedures

General Procedure A for the Per-*O*-Acetylation of Carbohydrates (GPA)

This procedure is based on a previously published protocol.[344]

To a stirring solution of carbohydrate (1.00 eq.) in pyridine (1.00 mL/mmol carbohydrate) was added Ac_2O (1.50 eq./carbohydrate OH-group). This was allowed to stir for 16 h at room temperature. The mixture was then diluted with EtOAc and washed sequentially with H_2O [1x reaction volume (RV)], sat. $NaHCO_{3\ (aq)}$ (3x RV), H_2O (1x RV), sat. $CuSO_{4\ (aq)}$ (3x RV), H_2O (2x RV), and brine (1x RV). The organic phase was then dried with $MgSO_4$, filtered, the filtrate collected, and the solvent was removed under reduced pressure. The products were then purified either by recrystallisation or column chromatography (see individual entries).

General Procedure B for the Per-*O*-Acetylation of Carbohydrates (GPB)

This procedure is based on a previously published protocol.[345]

To a stirring, refluxing solution of NaOAc (1.10 eq.) in Ac_2O (1.00 mL/mmol carbohydrate) was added carbohydrate (1.00 eq., min. 100 mmol) portionwise taking care not to allow the reaction to boil over between addition of portions. Concluding complete addition of carbohydrate to reaction, it was allowed to react for 15 min before being poured into a vigourously-stirring flask of ice-water (7RV) and was left until no ice remained. The product was then extracted into CH_2Cl_2 (2RV) and the organic phase was sequentially washed with ice-cold water (6x 2RV), sat. $NaHCO_{3\ (aq)}$ (4x 2RV), and brine (1x 2RV). The organic phase was then dried with $MgSO_4$, filtered, the filtrate collected, and the solvent removed under reduced pressure. The products were then purified either by recrystallisation or column chromatography (see individual entries).

General Procedure C for Anomeric 4-Nitrophenylation of Per-*O*-Acetylated Glycosides (GPC)

This procedure is based upon a previously published protocol.[349]

To a dry flask containing per-*O*-acetylated glycoside (1.00 eq.) and pNP (1.20 eq.) dissolved in anh. CH_2Cl_2 (2.00 mL/mmol glycoside) was added Et_3N (0.50 eq.). $BF_3{\cdot}OEt_2$ (2.50 eq.) was then added using a syringe pump over 1 h, following which, the reaction was left to stir at room temperature for

3 d. The reaction was then quenched by the addition of sat. $NaHCO_{3\,(aq)}$ (0.5RV) and the CH_2Cl_2 removed under reduced pressure. The products were then extracted into EtOAc (2RV) and the organic phase was washed sequentially with sat. $Na_2CO_{3\,(aq)}$ (4x RV) and brine (1x RV). The washing step was then repeated until the aqueous phase remained pale yellow upon addition of base. The organic phase was then dried with $MgSO_4$, filtered, the filtrate collected, and the solvent removed under reduced pressure to yield a yellow oil. The product was then triturated from hot Et_2O, the solid collected and washed with *n*-pentane and H_2O. The products were then purified either by recrystallisation or column chromatography (see individual entries).

General Procedure D for Deacetylation of Protected Glycosides (GPD)

This procedure is based upon a previously published protocol.[351,449]

Acetylated glycoside (1.00 eq.) was dissolved in an anh. mixture of MeOH and CH_2Cl_2 (2:1, 10 mL/mmol glycoside). To this stirring mixture was added a suspension of NaOMe in anh. MeOH (0.50 M, 0.67 mL/mmol glycoside) and the reaction left for 4 h, monitoring by TLC. After complete consumption of the acetylated starting material, the reaction was quenched by addition of a cation-exchange resin, Dowex™ Monosphere™ 650C, and the mixture was left stirring a further 15 min until no longer cloudy. The mixture was then filtered, the filtrate collected, and the solvent removed under reduced pressure. No further purification was performed.

General Procedure E for Anomeric Fluorination of Glycosides (GPE)

This procedure is based on a previously published protocol.[69]

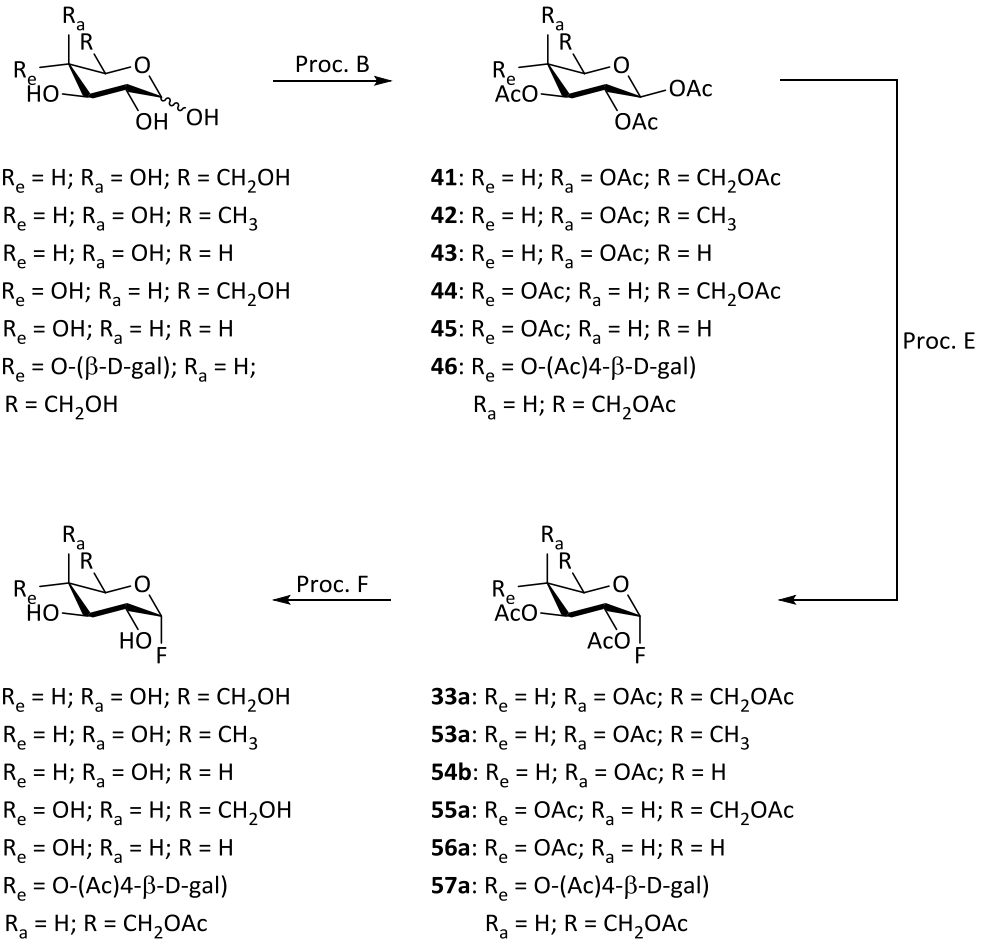
Per-acetylated glycoside (1.00 eq.) was dissolved in 70% HF/py (1.00 mL/mmol glycoside) at 0 °C and the reaction was stirred at 0 °C for 30 min before being allowed to warm to room temperature and monitored by TLC. After complete consumption of the starting material, the reaction was diluted with EtOAc (2RV) and H_2O (2RV), and the reaction quenched by careful addition of $Na_2CO_{3\,(S)}$ until the pH is between 7 and 8. Once neutralised, the mixture was then diluted with a further 3RV of EtOAc, and the organic phase was isolated and sequentially washed with H_2O (2x RV), sat. $CuSO_{4\,(aq)}$ (5x 2RV, or until aqueous phase doesn't change colour upon mixing), H_2O (2x RV), and brine (1x RV). The organic phase was then dried with $MgSO_4$, filtered, the filtrate collected, and the solvent removed under reduced pressure. Products were purified using column chromatography (see individual entries).

General Procedure F for Deacetylation of Protected Glycosyl Fluorides (GPF)

This procedure is a modified version of the above procedure.[351]

Acetylated glycosyl fluoride (1.00 eq.) was dissolved in an anh. mixture of MeOH and CH_2Cl_2 (2:1, 10.0 mL/mmol glycosyl fluoride) at 0 °C. To this stirring mixture was added a suspension of NaOMe in anh. MeOH (0.50 M, 0.67 mL/mmol glycosyl fluoride) and the reaction left stirring at 0 °C for 2 h. After complete consumption of the starting material, the reaction was quenched by addition of a cation-exchange resin, Dowex™ Monosphere™ 650C, and the mixture left stirring a further 15 min until no longer cloudy. The mixture was then filtered, the filtrate collected, and the CH_2Cl_2 removed under reduced pressure at 0 °C. Following complete removal of CH_2Cl_2 from the mixture, the methanolic residue was then flash frozen in $N_{2\,(l)}$ and the methanol removed *via* lyophilisation. No further purification was performed.

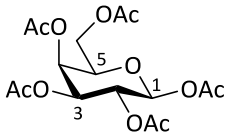
5.4.3 Synthesis of Glycosyl Fluorides



1,2,3,4,6-Penta-*O*-acetyl-β-D-galactopyranoside (41)

Carbohydrate **7** (20.0 g, 111.0 mmol) was reacted according to GPB and the desired product crystallised from the crude mixture using hot EtOH (200 mL). The crystals were then collected *via* filtration and washed with PE to return product **41** as colourless, amorphous crystals in a yield of 48% (20.7 g, 53.1 mmol).

Data was in agreement with that previously published.[345]



R_f = 0.38 (ethyl acetate/petroleum ether 6:4)

m.p. (°C) = 143.0 – 143.7 °C (lit.[345] 142 – 143 °C)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.70 (d, $^3J_{1,2}$ = 8.3 Hz, 1 H, 1-H), 5.43 (dd, $^3J_{4,3}$ = 3.4 Hz, $^3J_{4,5}$ = 1.0 Hz, 1 H, 4-H), 5.33 (dd, $^3J_{2,3}$ = 10.4 Hz, $^3J_{2,1}$ = 8.3 Hz, 1 H, 2-H), 5.08 (dd, $^3J_{3,2}$ = 10.4 Hz, $^3J_{3,4}$ = 3.4 Hz, 1 H, 3-H), 4.16 (dd, $^2J_{6a,6b}$ = 11.3 Hz, $^3J_{6a,5}$ = 6.8 Hz, 1 H, 6-H_a), 4.12 (dd, $^2J_{6b,6a}$ = 11.3 Hz, $^3J_{6b,5}$ = 6.4 Hz, 1 H, 6-H_b), 4.05 (ddd, $^3J_{5,6a}$ = 6.7 Hz, $^3J_{5,6b}$ = 6.6 Hz, $^3J_{5,4}$ = 1.2 Hz, 1 H, 5-H), 2.16 (s, 3 H, 4-OC(O)C**H$_3$**), 2.12 (s, 3 H, 1-OC(O)C**H$_3$**), 2.04 (s, 3 H, 2-OC(O)C**H$_3$**), 2.04 (s, 3 H, 6-OC(O)C**H$_3$**), 1.99 (s, 3 H, 3-OC(O)C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.5 (6-O**C**(O)CH_3), 170.3 (4-O**C**(O)CH_3), 179.1 (3-O**C**(O)CH_3), 169.5 (2-O**C**(O)CH_3), 169.1 (1-O**C**(O)CH_3), 92.3 (C-1), 71.9 (C-5), 71.0 (C-3), 68.0 (C-2), 66.9 (C-4), 61.2 (C-6), 21.0 (1-OC(O)**C**H_3), 20.8 (6-OC(O)**C**H_3, 2-OC(O)**C**H_3), 20.8 (4-OC(O)**C**H_3), 20.7 (3-OC(O)**C**H_3).

HRMS (ESI, +ve): 413.10494 [M+Na] ($C_{16}H_{22}O_{11}Na$, calc. 413.10598)

IR (neat): 1756, 1739, 1376, 1213, 1079, 1053 cm^{-1}

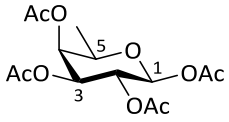
$[\alpha]_D^{25}$(c, solv.): +23.4 (c 1.04, $CHCl_3$), lit.[345] $[\alpha]_D^{25}$: +23.0 (c 1.0, $CHCl_3$)

Elem. Analysis (%) calc. for $C_{16}H_{22}O_{11}$: C 49.23, H 5.68. Found: C 49.44 ± 0.07, H 5.74 ± 0.01.

1,2,3,4-Tetra-*O*-acetyl-β-D-fucopyranoside (42)

Carbohydrate **37** (10.0 g, 60.6 mmol) was reacted according to GPA and the desired product purified *via* column chromatography (ethyl acetate/petroleum ether 1:1) to give **42** in a yield of 45% (9.0 g, 27.1 mmol).

Data was in agreement with that previously published.[453]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.66 (d, $^3J_{1,2}$ = 8.3 Hz, 1 H, 1-H), 5.29 (dd, $^3J_{2,3}$ = 10.4 Hz, $^3J_{2,1}$ = 8.3 Hz, 1 H, 2-H), 5.25 (dd, $^3J_{4,3}$ = 3.6 Hz, $^3J_{4,5}$ = 1.1 Hz, 1 H, 4-H), 5.05 (dd, $^3J_{3,2}$ = 10.4 Hz, $^3J_{3,4}$ = 3.5 Hz, 1 H, 3-H), 3.94 (dq, $^3J_{5,6}$ = 6.4 Hz, $^3J_{5,4}$ = 1.0 Hz, 1 H, 5-H), 2.17 (s, 3 H, 4-OC(O)C**H$_3$**), 2.09 (s, 3 H, 1-OC(O)C**H$_3$**), 2.02 (s, 3 H, 2-OC(O)C**H$_3$**), 1.97 (s, 3 H, 3-OC(O)C**H$_3$**), 1.20 (d, $^3J_{5,6}$ = 6.5 Hz, 3 H, 6-C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

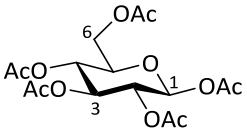
δ (ppm): 170.6 (4-O**C**(O)CH_3), 170.1 (3-O**C**(O)CH_3), 169.6 (2-O**C**(O)CH_3), 169.3 (1-O**C**(O)CH_3), 92.3 (C-1), 71.4 (C-3), 70.4 (C-5), 70.0 (C-4), 68.0 (C-2), 20.9 (1-OC(O)**C**H_3), 20.8 (2-OC(O)**C**H_3), 20.7 (4-OC(O)**C**H_3), 20.7 (3-OC(O)**C**H_3), 16.0 (C-6).

HRMS (ESI): 355.09995 [M+Na] ($C_{14}H_{20}O_9Na$, calc. 355.09995)

1,2,3,4,6-Penta-*O*-acetyl-β-D-glucopyranoside (44)

Carbohydrate **6** (20.0 g, 111.0 mmol) was reacted according to GPB and the desired product crystallised from the crude mixture using hot Et_2O (150 mL). The crystals were then collected *via* filtration and washed with ice-cold Et_2O to return product **44** as colourless, amorphous crystals in a yield of 59%. (25.4 g, 65.1 mmol).

Data was in agreement with that previously published.[450]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.72 (d, $^{3}J_{1,2}$ = 8.3 Hz, 1 H, 1-H), 5.25 (dd, $^{3}J_{3,2}$ = 9.4 Hz, $^{3}J_{34}$ = 9.4 Hz, 1 H, 3-H), 5.14 (dd, $^{3}J_{2,3}$ = 9.6 Hz, $^{3}J_{2,1}$ = 8.3 Hz, 1 H, 2-H), 5.13 (dd, $^{3}J_{4,5}$ = 10.1 Hz, $^{3}J_{4,3}$ = 9.4 Hz, 1 H, 4-H), 4.29 (dd, $^{2}J_{6a,6b}$ = 12.5 Hz, $^{3}J_{6a,5}$ = 4.5 Hz, 1 H, 6-H_a), 4.11 (dd, $^{2}J_{6b,6a}$ = 12.5 Hz, $^{3}J_{6b,5}$ = 2.2 Hz, 1 H, 6-H_b), 3.84 (ddd, $^{3}J_{5,4}$ = 10.1 Hz, $^{3}J_{5,6a}$ = 4.6 Hz, $^{3}J_{5,6b}$ = 2.2 Hz, 1 H, 5-H), 2.10 (s, 3 H, -OC(O)C**H$_3$**), 2.07 (s, 3 H, -OC(O)C**H$_3$**), 2.02 (s, 3 H, -OC(O)C**H$_3$**), 2.02 (s, 3 H, -OC(O)C**H$_3$**), 2.00 (s, 3 H, -OC(O)C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

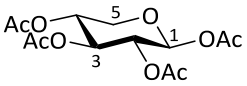
δ (ppm): 170.7 (6-O**C**(O)CH_3), 170.2 (3-O**C**(O)CH_3), 169.5 (4-O**C**(O)CH_3), 169.4 (2-O**C**(O)CH_3), 169.1 (1-O**C**(O)CH_3), 91.8 (C-1), 72.9 (C-3), 72.9 (C-5), 70.4 (C-2), 67.9 (C-4), 61.6 (C-6), 21.0 (1-OC(O)**C**H_3), 20.8 (6-OC(O)**C**H_3), 20.7 (2-OC(O)**C**H_3, 4-OC(O)**C**H_3), 20.7 (3-OC(O)**C**H_3).

$[\alpha]_D^{25}$(c, solv.): +4.1 (c 1.03, $CHCl_3$), lit.[450] **$[\alpha]_D^{25}$**: +3.9 (c 1.00, $CHCl_3$)

1,2,3,4-Tetra-*O*-acetyl-β-D-xylopyranoside (45)

Carbohydrate **39** (20.0 g, 133.2 mmol) was reacted according to GPB and the desired product crystallised from the crude mixture using hot EtOH (200 mL). The crystals were then collected *via* filtration and washed with PE to return product **45** as colourless, amorphous crystals in a yield of 52% (22.1 g, 69.5 mmol).

Data was in agreement with that previously published.[451,452]



R_f = 0.47 (ethyl acetate/petroleum ether 1:1)

m.p. (°C) = 126.1 – 127.4 °C (lit.[452] 126 – 127 °C)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.72 (d, $^3J_{1,2}$ = 6.9 Hz, 1 H, 1-H), 5.21 (dd, $^3J_{3,2}$ = 8.3 Hz, $^3J_{3,4}$= 8.3 Hz, 1 H, 3-H), 5.04 (dd, $^3J_{2,3}$ = 8,4 Hz, $^3J_{2,1}$ = 6.9 Hz, 1 H, 2-H), 4.98 (ddd, $^3J_{4,5b}$ = 8.3 Hz, $^3J_{4,3}$ = 8.3 Hz, $^3J_{4,5a}$ = 5.0 Hz, 1 H, 4-H), 4.15 (dd, $^2J_{5a,5b}$ = 12.0 Hz, $^3J_{5a,4}$ = 5.0 Hz, 1 H, 5-H_a), 3.53 (dd, $^2J_{5b,5a}$ = 12.1 Hz, $^3J_{5b,4}$ = 8.4 Hz, 1 H, 5-H_b), 2.11 (s, 3 H, -OC(O)C**H**$_3$), 2.06 (s, 3 H, -OC(O)C**H**$_3$), 2.05 (s, 3 H, -OC(O)C**H**$_3$), 2.05 (s, 3 H, -OC(O)C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 169.9 (3-O**C**(O)CH_3), 169.9 (4-O**C**(O)CH_3), 169.4 (2-O**C**(O)CH_3), 169.1 (1-O**C**(O)CH_3), 92.2 (C-1), 71.1 (C-3), 69.6 (C-2), 68.5 (C-4), 62.9 (C-5), 20.9 (1-OC(O)**C**H_3), 20.8 (4-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

HRMS (ESI, +ve): 341.08387 [M+Na] ($C_{13}H_{18}O_9Na$, calc. 341.08430); 357.05791 [M+K] ($C_{13}H_{18}O_9K$, calc. 357.05824)

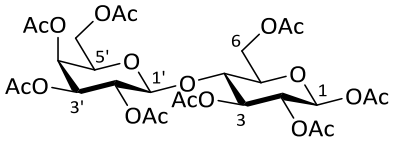
$[\alpha]_D^{25}$(c, solv.): -32.2 (c 0.53, $CHCl_3$), -26.3 (c 1.03, $CHCl_3$); lit.[452] $[\alpha]_D^{20}$: -38.5 (c 2.3, CH_2Cl_2)

Elem. Analysis (%) calc. for $C_{13}H_{18}O_9$: C 49.06, H 5.70. Found: C 49.18 ± 0.02, H 5.74 ± 0.01.

1,2,3,6-Tetra-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl)-β-D-galactopyranosyl-β-D-glucopyranoside (46)

Disaccharide **2** (12.7 g, 35.2 mmol) was reacted according to GPB and the desired product crystallised from the crude mixture using hot Et_2O (200 mL). The crystals were then collected *via* filtration and washed successively with ice-cold EtOH (2x 100 mL) and H_2O (2x 100 mL) to return product **46** as colourless, amorphous crystals which were dried in a desiccator in a yield of 72% (17.1 g, 25.2 mmol).

Data was in agreement with that previously published.[345]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.67 (d, $^3J_{1,2}$ = 8.2 Hz, 1 H, 1-H), 5.35 (d, $^3J_{4',3'}$ = 3.6 Hz, 1 H, 4'-H), 5.24 (dd, $^3J_{3,4}$ = 9.2 Hz, $^3J_{3,2}$ = 9.2 Hz, 1 H, 3-H), 5.11 (dd, $^3J_{2',3'}$ = 10.4 Hz, $^3J_{2',1'}$ = 7.9 Hz, 1 H, 2'-H), 5.04 (dd, $^3J_{2,3}$ = 8.9 Hz, $^3J_{2,1}$ = 8.9 Hz, 1 H, 2-H), 4.94 (dd, $^3J_{3',2'}$ = 10.4 Hz, $^3J_{3',4'}$ = 3.4 Hz, 1 H, 3'-H), 4.47 (d, $^3J_{1',2'}$ = 7.9 Hz, 1 H, 1'-H), 4.45 (dd, $^2J_{6a,6b}$ = 12.0 Hz, $^3J_{6a,5}$ = 2.0 Hz, 1 H, 6-H_a), 4.17-4.09 (m, 2 H, 6-H_b, 6'-H_a), 4.07 (dd, $^2J_{6'b,6'a,}$ = 11.2 Hz, $^3J_{6'b,5'}$ = 7.3 Hz, 1 H, 6'-H_b), 3.90-3.85 (m, 1 H, 5'-H), 3.84 (dd, $^3J_{4,5}$ = 9.9 Hz, $^3J_{4,3}$ = 9.1 Hz, 1 H, 4-H), 3.75 (ddd, $^3J_{5,4}$ = 9.9 Hz, $^3J_{5,6b}$ = 4.9 Hz, $^3J_{5,6a}$ = 2.0 Hz, 5-H), 2.15 (s, 3 H, 4'-OC(O)C**H$_3$**), 2.12 (s, 3 H, 6-OC(O)C**H$_3$**), 2.09 (s, 3 H, 1-OC(O)C**H$_3$**), 2.06 (s, 3 H, 6'-OC(O)C**H$_3$**), 2.05 (s, 3 H, 3-OC(O)C**H$_3$**), 2.04 (s 3 H, 2'-OC(O)C**H$_3$**), 2.03 (s, 3 H, 2-OC(O)C**H$_3$**), 1.96 (s, 3 H, 3'-OC(O)C**H$_3$**).

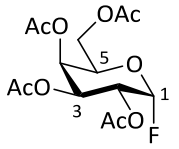
^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.5 (6'-O**C**(O)CH_3), 170.4 (6-O**C**(O)CH_3), 170.3 (3'-O**C**(O)CH_3), 170.2 (4'-O**C**(O)CH_3), 169.8 (2-O**C**(O)CH_3), 169.7 (3-O**C**(O)CH_3), 169.1 (2'-O**C**(O)CH_3), 169.0 (1-O**C**(O)CH_3), 101.1 (C-1'), 91.7 (C-1), 75.8 (C-4), 73.6 (C-5), 72.8 (C-3), 71.1 (C-2), 70.9 (C-5'), 70.6 (C-3'), 69.1 (C-2'), 66.7 (C-4'), 61.9 (C-6), 61.0 (C-6'), 21.0 (-OC(O)**C**H_3), 21.0 (-OC(O)**C**H_3), 20.9 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

2,3,4,6-Tetra-*O*-acetyl-α-D-galactopyranosyl fluoride (33a)

Glycoside **41** (7.83 g, 20.1 mmol) was reacted as outlined in GPE. The crude mixture was purified *via* column chromatography (ethyl acetate/petroleum ether 4:6) to give **33a** in a yield of 56% (3.96 g, 11.3 mmol).

Data was in agreement with that previously published.[351]



R_f = 0.53 (ethyl acetate/petroleum ether 1:1)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.80 (dd, $^2J_{1\text{-}H,1\text{-}F}$ = 53.2 Hz, $^3J_{1,2}$ = 2.7 Hz, 1 H, 1-H), 5.53 (dd, $^3J_{4,3}$ = 3.4 Hz, $^3J_{4,5}$ = 1.4 Hz, 1 H, 4-H), 5.36 (dd, $^3J_{3,2}$ = 10.8 Hz, $^3J_{3,4}$ = 3.2 Hz, 1 H, 3-H), 5.19 (ddd, $^3J_{2,1\text{-}F}$ = 23.8 Hz, $^3J_{2,3}$ = 10.8 Hz, $^3J_{2,1}$ = 2.8 Hz, 1 H, 2-H), 4.41 (ddd, $^3J_{5,6b}$ = 6.6 Hz, $^3J_{5,6a}$ = 6.5 Hz, $^3J_{5,4}$ = 1.3 Hz, 1 H, 5-H), 4.16 (dd, $^2J_{6a,6b}$ = 11.4 Hz, $^3J_{6a,5}$ = 6.3 Hz, 1 H, 6-H_a), 4.12 (dd, $^2J_{6b,6a}$ = 11.4 Hz, $^3J_{6b,5}$ = 6.8 Hz, 1 H, 6-H_b), 2.15 (s, 3 H, 4-OC(O)C**H$_3$**), 2.12 (s, 3 H, 2-OC(O)C**H$_3$**), 2.06 (s, 3 H, 6-OC(O)C**H$_3$**), 2.01 (s, 3 H, 3-OC(O)C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.5 (-O**C**(O)CH_3), 170.4 (-O**C**(O)CH_3), 170.1 (-O**C**(O)CH_3), 170.0 (-O**C**(O)CH_3), 104.5 (d, 228.5, C-1), 69.0 (d, 3.5, C-5), 67.6 (d, 23.8, C-2), 67.5 (C-4), 67.1 (C-3), 61.4 (C-6), 20.8 (2x -OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

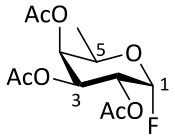
^{19}F NMR (282 MHz, $CDCl_3$): **δ** -150.8 (F-1).

$[\alpha]_D^{25}$ (c, solv.): +91.3 (c 1.04, $CHCl_3$), lit.[351] $[\alpha]_D^{20}$ +100 (c 1, $CHCl_3$)

2,3,4-Tri-*O*-acetyl-α-D-fucopyranosyl fluoride (53a)

Glycoside **42** (3.59 g, 10.8 mmol) was reacted as outlined in GPE. The crude mixture was purified *via* column chromatography (ethyl acetate/petroleum ether 3:7) to give **53a** in a yield of 26% (0.83 g, 2.80 mmol).

Data was in agreement with that previously published.[351]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.75 (dd, $^{2}J_{1\text{-}H,1\text{-}F}$ = 53.7 Hz, $^{3}J_{1,2}$ = 2.8 Hz, 1 H, 1-H), 5.40-5.32 (m, 2 H, 3-H, 4-H), 5.17 (ddd, $^{3}J_{2\text{-}H,1\text{-}F}$ = 23.8 Hz, $^{3}J_{2,3}$ = 11.6 Hz, $^{3}J_{2,1}$ = 2.7 Hz, 1 H, 2-H), 4.34 (q, $^{3}J_{5,6}$ = 6.5 Hz, 1 H, 5-H), 2.17 (s, 3 H, 4-OC(O)C**H**$_3$), 2.11 (s, 3 H, 2-OC(O)C**H**$_3$), 2.00 (s, 3 H, 3-OC(O)C**H**$_3$), 1.19 (d, $^{3}J_{6,5}$ = 6.5 Hz, 3 H, 6-C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

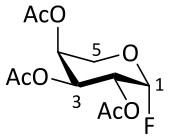
δ (ppm): 170.5 (-O**C**(O)CH_3), 170.4 (-O**C**(O)CH_3), 170.1 (-O**C**(O)CH_3), 104.7 (d, $^{1}J_{C\text{-}1,F\text{-}1}$ = 226.5 Hz, C-1), 70.5 (C-4), 67.6 (d, $^{2}J_{C\text{-}2,F\text{-}1}$ = 24.3 Hz, C-2), 67.6 (C-3), 67.4 (d, $^{3}J_{C\text{-}5,F\text{-}1}$ = 3.5 Hz, C-5), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3), 15.9 (C-6).

^{19}F NMR (282 MHz, $CDCl_3$): **δ** -150.0 (F-1).

$[\alpha]_D^{25}$ (c, solv.): +113.3 (c 1.03, $CHCl_3$), lit.[351] $[\alpha]_D^{20}$ +100 (c 1, $CHCl_3$)

2,3,4-Tri-*O*-acetyl-β-L-arabinopyranosyl fluoride (54a)

Carbohydrate **38** (20.0 g, 134.8 mmol) was reacted according to GPB and was isolated an inseparable crude mixture of isomers. The crude form was obtained in a quantitative yield (46.9 g, 147.5 mmol). A sample of crude glycoside **43** (3.33 g, 10.5 mmol) was directly reacted as outlined in GPE. The crude mixture was purified *via* column chromatography (diethyl ether/*n*-pentane 3:7) and subsequently recrystallised from hot diethyl ether/*n*-pentane (1:2). The crystals were washed with *n*-pentane to give **54a** in a yield of 25% (0.71 g, 2.60 mmol).



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.78 (dd, $^{2}J_{1\text{-}H,1\text{-}F}$ = 53.6 Hz, $^{3}J_{1,2}$ = 2.7 Hz, 1 H, 1-H), 5.41 (ddd, $^{3}J_{4,3}$ = 3.4 Hz, $^{3}J_{4,5a}$ = 1.6 Hz, $^{3}J_{4,5b}$ = 1.6 Hz, 1 H, 4-H), 5.36 (dd, $^{3}J_{3,2}$ = 10.8 Hz, $^{3}J_{3,4}$ = 3.4 Hz, 1 H, 3-H), 5.22 (ddd, $^{3}J_{2\text{-}H,1\text{-}F}$ = 23.7 Hz, $^{3}J_{2,3}$ = 10.8 Hz, $^{3}J_{2,1}$ = 2.7 Hz, 1 H, 2-H), 4.13 (dd, $^{2}J_{5a,5b}$ = 13.3 Hz, $^{3}J_{5a,4}$ = 1.2 Hz, 1 H, 5-H_a), 3.88 (dd, $^{2}J_{5b,5a}$ = 13.3 Hz, $^{3}J_{5b,4}$ = 1.9 Hz, 1 H, 5-H_b), 2.16 (s, 3 H, 4-OC(O)C**H**$_3$), 2.12 (s, 3 H, 2-OC(O)C**H**$_3$), 2.03 (s, 3 H, 3-OC(O)C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.4 (2-O**C**(O)CH_3), 170.3 (4-O**C**(O)CH_3), 170.1 (3-O**C**(O)CH_3), 105.0 (d, $^{1}J_{C\text{-}1,F\text{-}1}$ = 227.8 Hz, C-1), 68.4 (C-4), 67.8 (d, $^{2}J_{C\text{-}2,1\text{-}F}$ = 24.2 Hz, C-2), 66.7 (C-3), 62.7 (d, $^{3}J_{C\text{-}5,1\text{-}F}$ = 3.9 Hz, C-5), 21.0 (4-OC(O)**C**H_3), 20.8 (2 -OC(O)**C**H_3), 20.8 (3 -OC(O)**C**H_3).

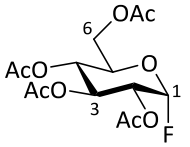
^{19}F NMR (282 MHz, $CDCl_3$): **δ** -153.3 (F-1).

$[\alpha]_D^{25}$ (c, solv.): +138.7 (c 0.67, $CHCl_3$); +147.9 (c 1.01, $CHCl_3$)

2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl fluoride (55a)

Glycoside **44** (3.90 g, 10.0 mmol) was reacted as outlined in GPE. The crude mixture was purified *via* column chromatography (ethyl acetate/petroleum ether 7:3) to give **55a** in a yield of 91% (3.20 g, 9.1 mmol).

Data was in agreement with that previously published.[351,454,455]



R_f = 0.51 (ethyl acetate/petroleum ether 1:1)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.75 (dd, $^2J_{1\text{-}H,1\text{-}F}$ = 52.8 Hz, $^3J_{1,2}$ = 2.7 Hz, 1 H, 1-H), 5.50 (dd, $^3J_{3,4}$ = 9.9 Hz, $^3J_{3,2}$ = 9.9 Hz, 1 H, 3-H), 5.16 (dd, $^3J_{4,5}$ = 9.9 Hz, $^3J_{4,3}$ = 9.9 Hz, 1 H, 4-H), 4.96 (ddd, $^3J_{2\text{-}H,1\text{-}F}$ = 24.2 Hz, $^3J_{2,3}$ = 10.2 Hz, $^3J_{2,1}$ = 2.8 Hz, 1 H, 2-H), 4.29 (dd, $^2J_{6a,6b}$ = 12.5 Hz, $^3J_{6a,5}$ = 4.1 Hz, 1 H, 6-H_a), 4.19 (ddd, $^3J_{5,4}$ = 10.3 Hz, $^3J_{5,6a}$ = 4.1 Hz, $^3J_{5,6b}$ = 2.2 Hz, 1 H, 5-H), 4.15 (dd, $^2J_{6b,6a}$ = 12.5 Hz, $^3J_{6b,5}$ = 2.3 Hz, 1 H, 6-H_b), 2.11 (s, 3 H, 2-OC(O)C**H$_3$**), 2.11 (s, 3 H, 6-OC(O)C**H$_3$**), 2.05 (s, 3 H, 4-OC(O)C**H$_3$**), 2.03 (s, 3 H, 3-OC(O)C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

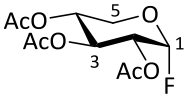
δ (ppm): 170.7 (6-O**C**(O)CH_3), 170.1 (-O**C**(O)CH_3), 170.1 (-O**C**(O)CH_3), 169.6 (4-O**C**(O)CH_3), 103.9 (d, $^1J_{C\text{-}1,F\text{-}1}$ = 229.6 Hz, C-1), 70.4 (d, $^2J_{C\text{-}2,F\text{-}1}$ = 24.6 Hz, C-2), 69.9 (d, $^3J_{C\text{-}5,F\text{-}1}$ = 4.3 Hz, C-5), 69.5 (C-3), 67.5 (C-4), 61.3 (C-6), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

^{19}F NMR (282 MHz, $CDCl_3$): **δ** -149.8 (F-1)

2,3,4-Tri-*O*-acetyl-α-D-xylopyranosyl fluoride (56a)

Glycoside **45** (3.18 g, 10.0 mmol) was reacted as outlined in GPE. The crude mixture was purified *via* column chromatography (ethyl acetate/petroleum ether 4:6) to give **56a** in a yield of 43% (1.21 g, 4.3 mmol).

Data was in agreement with that previously published.[454]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.69 (dd, $^2J_{1\text{-}H,1\text{-}F}$ = 53.0 Hz, $^3J_{1,2}$ = 2.7 Hz, 1 H, 1-H), 5.49 (dd, $^3J_{3,2}$ = 9.9 Hz, $^3J_{3,4}$ = 9.9 Hz, 1 H, 3-H), 5.04 (ddd, $^3J_{4,5b}$ = 10.7 Hz, $^3J_{4,3}$ = 9.8 Hz, $^3J_{4,5a}$ = 5.9 Hz, 1 H, 4-H), 4.89 (ddd, $^3J_{2,1\text{-}F}$ = 24.1 Hz, $^3J_{2,3}$ = 10.1 Hz, $^3J_{2,1}$ = 2.8 Hz, 1 H, 2-H), 3.98 (dd, $^2J_{5a,5b}$ = 11.2 *Hz*, $^3J_{5a,4}$ = 6.0 Hz, 1 H, 5-H_a), 3.78 (dd, $^2J_{5b,5a}$ = 11.1 Hz, $^3J_{5b,4}$ = 11.1 Hz, 1 H, 5-H_b), 2.10 (s, 3 H, -OC(O)C**H$_3$**), 2.05 (s, 3 H, -OC(O)C**H$_3$**), 2.04 (s, 3 H, -OC(O)C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

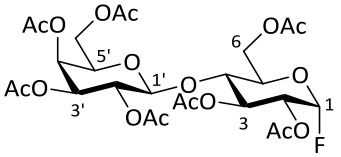
δ (ppm): 170.2 (-O**C**(O)CH_3), 170.0 (-O**C**(O)CH_3), 170.0 (-O**C**(O)CH_3), 104.1 (d, $^1J_{C\text{-}1,F\text{-}1}$ = 229.2 Hz, C-1), 70.6 (d, $^2J_{C\text{-}2,F\text{-}1}$ = 24.8 Hz, C-2), 69.1 (C-3), 68.4 (C-4), 60.4 (d, $^3J_{C\text{-}5,F\text{-}1}$ = 4.7 Hz, C-5), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

^{19}F NMR (282 MHz, $CDCl_3$): **δ** -151.5 (F-1)

2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl)-β-D-galactopyranosyl-α-D-glucopyranosyl fluoride (57a)

Glycoside **46** (3.39 g, 5.00 mmol) was reacted as outlined in GPE. The crude mixture was purified *via* column chromatography (ethyl acetate/petroleum ether 3:7) to give **57a** in a yield of 36% (1.13 g, 1.80 mmol).

Data was in agreement with that previously published.[456]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.67 (dd, $^2J_{1\text{-}H,1\text{-}F}$ = 53.0 Hz, $^3J_{1,2}$ = 2.8 Hz, 1 H, 1-H), 5.49 (dd, $^3J_{3,2}$= 9.8 Hz, $^3J_{3,4}$ = 9.8 Hz, 1 H, 3-H), 5.35 (d, $^3J_{4',3'}$ = 3.4 Hz, 1 H, 4'-H), 5.11 (dd, $^3J_{2',3'}$ = 10.4 Hz, $^3J_{2',1'}$ = 7.9 Hz, 1 H, 2'-H), 4.96 (dd, $^3J_{3',2'}$ = 10.4 Hz, $^3J_{3',4'}$ = 3.4 Hz, 1 H, 3'-H), 4.88 (ddd, $^3J_{2\text{-}H,1\text{-}F}$ = 24.2 Hz, $^3J_{2,3}$ = 10.2 Hz, $^3J_{2,1}$ = 2.8 Hz, 1 H, 2-H), 4.52 (dd, $^2J_{6a,6b}$ = 12.3 Hz, $^3J_{6a,5}$ = 1.9 Hz, 1 H, 6-H_a), 4.50 (d, $^3J_{1',2'}$ = 7.9 Hz, 1 H, 1'-H), 4.18-4.13 (m, 2 H, 6-H_b, 6'-H_a), 4.11 (ddd, $^3J_{5,6b}$ = 10.4 Hz, $^3J_{5,4}$ = 4.5 Hz, $^3J_{5,6a}$ = 1.8 Hz, 1 H, 5-H), 4.07 (dd, $^2J_{6'b,6'a}$ = 11.2 Hz, $^3J_{6'b,5'}$ = 7.2 Hz, 1 H, 6'-H_b), 3.90-3.82 (m, 2 H, 4-H, 5'-H), 2.15 (s, 3 H, 4'-OC(O)C**H**$_3$), 2.14 (s, 3 H, 6-OC(O)C**H**$_3$), 2.10 (s, 3 H, 2-OC(O)C**H**$_3$), 2.06 (s, 6 H, 3-OC(O)C**H**$_3$, 6'-OC(O)C**H**$_3$), 2.04 (s, 3 H, 2'-OC(O)C**H**$_3$), 1.96 (s, 3 H, 3'-OC(O)C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.5 (-O**C**(O)CH_3), 170.4 (-O**C**(O)CH_3), 170.3 (-O**C**(O)CH_3), 170.3 (-O**C**(O)CH_3), 170.2 (-O**C**(O)CH_3), 169.5 (3-O**C**(O)CH_3), 169.1 (2'-O**C**(O)CH_3), 103.8 (d, $^1J_{C\text{-}1,F\text{-}1}$ = 229.6 Hz, C-1), 100.9

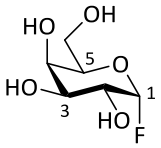
(C-1‘), 75.2 (C-4), 71.1 (C-3‘), 70.9 (C-5‘), 70.8 (d, $^{3}J_{C\text{-}5,F\text{-}1}$ = 4.2 Hz, C-5), 70.6 (d, $^{2}J_{C\text{-}2,F\text{-}1}$ = 24.4 Hz, C-2), 69.2 (C-2‘), 69.1 (C-3), 66.8 (C-4‘), 61.4 (C-6), 61.0 (C-6‘), 21.0 (-OC(O)**C**H_3), 20.9 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

^{19}F NMR (282 MHz, $CDCl_3$): **δ** -149.1 (F-1).

α-D-Galactopyranosyl fluoride (1)

Fluoride **33a** (1.22 g, 3.50 mmol) was reacted as outlined in GPF. The crude mixture was not purified further and stored at -20 °C. Product **1** was obtained in quantitative yield (0.67 g, 3.70 mmol).

Data was in agreement with that previously published.[351]



R_f = 0.70 (RP-TLC, methanol/ethyl acetate/water 7:2:1)

^{1}H NMR (600 MHz, D_2O):

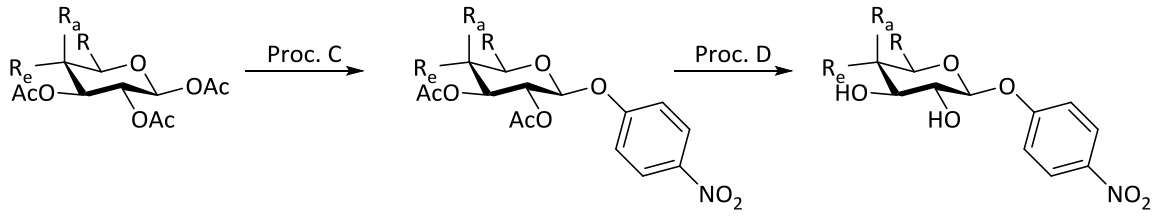
δ (ppm): 5.73 (dd, $^{2}J_{1\text{-}H,1\text{-}F}$ = 53.7 Hz, $^{3}J_{1,2}$ = 2.8 Hz, 1 H, 1-H), 4.11 (dd, $^{3}J_{5,6a}$ = 7.7 Hz, $^{3}J_{5,6b}$ = 4.5 Hz, 1 H, 5-H), 4.06 (d, $^{3}J_{4,3}$ = 3.1 Hz, 1 H, 4-H), 3.92 (dd, $^{3}J_{3,2}$ = 10.3 Hz, $^{3}J_{3,4}$ = 3.2 Hz, 1 H, 3-H), 3.86 (ddd, $^{3}J_{2,1\text{-}F}$ = 26.0 Hz, $^{3}J_{2,3}$ = 10.3 Hz, $^{3}J_{2,1}$ = 2.9 Hz, 1 H, 2-H), 3.81-3.74 (m, 2 H, 6-H_a, 6-H_b).

^{13}C NMR (151 MHz, D_2O):

δ (ppm): 107.5 (d, 222.8, C-1), 73.4 (d, 2.7, C-5), 68.9 (C-4), 68.8 (C-3), 67.9 (d, 24.5, C-2), 61.0 (C-6).

^{19}F NMR (282 MHz, $CDCl_3$): **δ** -152.3 (F-1).

5.4.4 Synthesis of *para*-Nitrophenyl Glycosides



41: R_e = H; R_a = OAc; R = CH_2OAc
42: R_e = H; R_a = OAc; R = CH_3
43: R_e = H; R_a = OAc; R = H
44: R_e = OAc; R_a = H; R = CH_2OAc
45: R_e = OAc; R_a = H; R = H
46: R_e = OAc; R_a = H; R = CH_2O-((Ac)$_4$-b-D-gal)

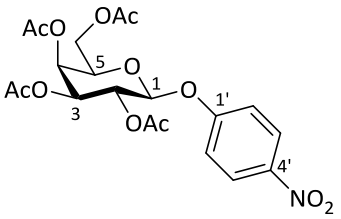
47: R_e = H; R_a = OAc; R = CH_2OAc
48: R_e = H; R_a = OAc; R = CH_3
49: R_e = H; R_a = OAc; R = H
50: R_e = OAc; R_a = H; R = CH_2OAc
51: R_e = OAc; R_a = H; R = H
52: R_e = OAc; R_a = H; R = CH_2O-((Ac)$_4$-b-D-gal)

22: R_e = H; R_a = OH; R = CH_2OH
25: R_e = H; R_a = OH; R = CH_3
26: R_e = H; R_a = OH; R = H
4: R_e = OH; R_a = H; R = CH_2OH
27: R_e = OH; R_a = H; R = H
36: R_e = OAc; R_a = H; R = CH_2O-((Ac)$_4$-b-D-gal)

4-Nitrophenyl 2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside (47)

Glycoside **41** (19.5 g, 50.0 mmol) was reacted as outlined in GPC. The desired product was then crystallised from hot EtOH as colourless filaments. The crystals were then collected by filtration and washed with ice-cold dH_2O and dried in a desiccator to obtain **47** in a yield of 88% (20.7 g, 44.1 mmol)

Data was in agreement with that previously published.[457,458,461]



m.p. (°C) = 143.6 – 144.5 °C (lit.[457] 141 – 143 °C)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.21 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.09 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 5.52 (dd, $^3J_{2,3}$ = 10.4 Hz, $^3J_{2,1}$ = 7.9 Hz, 1 H, 2-H), 5.48 (dd, $^3J_{4,3}$ = 3.5 Hz, $^3J_{4,5}$ = 1.0 Hz, 1 H, 4-H), 5.17 (d, $^3J_{1,2}$ = 7.9 Hz, 1 H, 1-H), 5.14 (dd, $^3J_{3,2}$ = 10.5 Hz, $^3J_{3,4}$ = 3.4 Hz, 1 H, 3-H), 4.22 (dd, $^2J_{6a,6b}$ = 11.3 Hz, $^3J_{6a,5}$ = 7.0 Hz, 1 H, 6-H_a), 4.16 (dd, $^2J_{6b,6a}$ = 11.3 Hz, $^3J_{6b,5}$ = 5.9 Hz, 1 H, 6-H_b), 4.13 (ddd, $^3J_{5,6a}$ = 7.1 Hz, $^3J_{5,6b}$ = 6.0 Hz, $^3J_{5,4}$ = 1.1 Hz, 1 H, 5-H), 2.19 (s, 3 H, -OC(O)C**H**$_3$), 2.07 (s, 3 H, -OC(O)C**H**$_3$), 2.07 (s, 3 H, -OC(O)C**H**$_3$), 2.02 (s, 3 H, -OC(O)C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.4 (-O**C**(O)CH_3), 170.2 (-O**C**(O)CH_3), 170.2 (-O**C**(O)CH_3), 169.4 (-O**C**(O)CH_3), 161.3 (C-1‘), 143.4 (C-4‘), 125.9 (C-3‘), 116.7 (C-2‘), 98.8 (C-1), 71.6 (C-5), 70.7 (C-3), 68.4 (C-2), 66.8 (C-4), 61.5 (C-6), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

HRMS (ESI): 492.11084 [M+Na] ($C_{20}H_{23}NO_{12}Na$, calc. 492.11179)

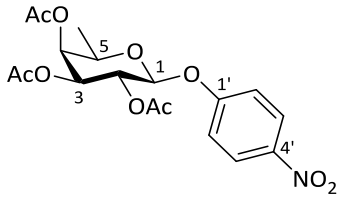
IR (neat): 1734, 1591, 1516, 1345, 1220, 1044 cm^{-1}

$[\alpha]_D^{25}$ (c, solv.): -9.3 (c 2.03, $CHCl_3$); -9.8 (c 1.01, $CHCl_3$), lit.[458] $[\alpha]_D^{20}$ -9.7 ($CHCl_3$)

Elem. Analysis (%) calc. for $C_{20}H_{23}NO_{12}$: C 51.18, H 4.94, N 2.98. Found: C 50.39 ± 0.21, H 4.90 ± 0.03, N 3.07 ± 0.02.

4-Nitrophenyl 2,3,4-tri-*O*-acetyl-β-D-fucopyranoside (48)

Glycoside **42** (1.69 g, 5.1 mmol) was reacted as outlined in GPC. The desired product was isolated *via* column chromatography (Et_2O/*n*-pentane 4:6) to obtain **48** in a yield of 19% (0.39 g, 0.90 mmol).



^{1}H NMR (600 MHz, $CDCl_3$):

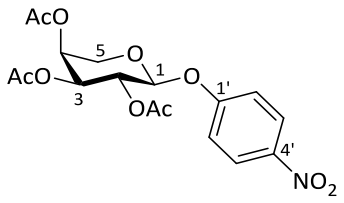
δ (ppm): 8.22 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3‘-H), 7.16 d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2‘-H), 5.85 (d, $^3J_{1,2}$ = 3.6 Hz, 1 H, 1-H), 5.56 (dd, $^3J_{3,2}$ = 10.9 Hz, $^3J_{3,4}$ = 3.4 Hz, 1 H, 3-H), 5.37 (dd, $^3J_{4,3}$ = 3.4 Hz, $^3J_{4,5}$ = 1.4 Hz, 1 H, 4-H), 5.31 (dd, $^3J_{2,3}$ = 11.0 Hz, $^3J_{2,1}$ = 3.6 Hz, 1 H, 2-H), 4.18 (dq, $^3J_{5,6}$ = 6.4 Hz, $^3J_{5,4}$ = 1.4 Hz, 1 H, 5-H), 2.20 (s, 3 H, 4-OC(O)C**H**$_3$), 2.07 (s, 3 H, 3-OC(O)C**H**$_3$), 2.04 (s, 3 H, 2-OC(O)C**H**$_3$), 1.13 (d, $^3J_{6,5}$ = 6.5 Hz, 3 H, 6-C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.6 (4-O**C**(O)CH_3), 170.5 (3-O**C**(O)CH_3), 170.2 (2-O**C**(O)CH_3), 161.3 (C-1‘), 143.0 (C-4‘), 126.0 (C-3‘), 116.5 (C-2‘), 94.9 (C-1), 70.7 (C-4), 67.8 (C-3), 67.6 (C-2), 66.3 (C-5), 20.9 (3-OC(O)**C**H_3), 20.8 (2-OC(O)**C**H_3), 20.8 (4-OC(O)**C**H_3), 16.0 (C-6).

4-Nitrophenyl 2,3,4-tri-*O*-acetyl-α-L-arabinopyranoside (49)

A sample of crude glycoside **43** (1.59 g, 5.00 mmol) was reacted as outlined in GPC. Product **49** was isolated *via* column chromatography (diethyl ether/*n*-pentane 4:6) in a yield of 3% (51.0 mg, 0.10 mmol).



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.22 (d, $^{3}J_{3',2'}$ = 9.2 Hz, 2 H, 3‘-H), 7.16 (d, $^{3}J_{2',3'}$ = 9.2 Hz, 2 H, 2‘-H), 5.88 (d, $^{3}J_{1,2}$ = 3.5 Hz, 1 H, 1-H), 5.57 (dd, $^{3}J_{3,2}$ = 10.8 Hz, $^{3}J_{3,4}$ = 3.5 Hz, 1 H, 3-H), 5.42 (ddd, $^{3}J_{4,3}$ = 3.4 Hz, $^{3}J_{4,5b}$ = 1.9 Hz, $^{3}J_{4,5a}$ = 1.4 Hz, 1 H, 4-H), 5.36 (dd, $^{3}J_{2,3}$ = 10.8 Hz, $^{3}J_{2,1}$ = 3.5 Hz, 1 H, 2-H), 3.97 (dd, $^{2}J_{5a,5b}$ = 13.3 Hz, $^{3}J_{5a,4}$ = 1.4 Hz, 1 H, 5-H_a), 3.79 (dd, $^{2}J_{5b,5a}$ = 13.3 Hz, $^{3}J_{5b,4}$ = 2.0 Hz, 1 H, 5-H_b), 2.18 (s, 3 H, 4-OC(O)C**H$_3$**), 2.08 (s, 3 H, 2-OC(O)C**H$_3$**), 2.06 (s, 3 H, 3-OC(O)C**H$_3$**).

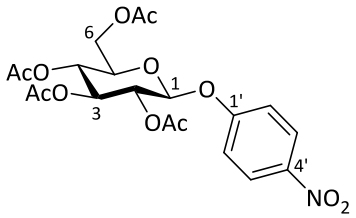
^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.5 (-O**C**(O)CH_3), 170.3 (-O**C**(O)CH_3), 170.2 (-O**C**(O)CH_3), 161.2 (C-1‘), 143.1 (C-4‘), 126.1 (C-3‘), 116.6 (C-2‘), 95.2 (C-1), 68.7 (C-4), 67.8 (C-2), 66.9 (C-3), 61.9 (C-5), 21.1 (-OC(O)**C**H_3), 20.9 (-OC(O)**C**H_3).

4-Nitrophenyl 2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (50)

Glycoside **44** (19.5 g, 50.0 mmol) was reacted according to GPC. The desired product was then crystallised from hot EtOH as colourless filaments. The crystals were then collected by filtration and washed with ice-cold dH_2O and dried in a desiccator to give **50** in a yield of 77% (18.0 g, 38.3 mmol).

Data was in agreement with that previously published.[457,458]



m.p. (°C) = 177.1 – 177.8 °C (lit.[458] 177 – 178.5 °C)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.21 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.07 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 5.36-5.28 (m, 2 H, 2-H, 3-H), 5.22 (d, $^3J_{1,2}$ = 7.3 Hz, 1 H, 1-H), 5.18 (dd, $^3J_{4,3}$ = 9.5 Hz, $^3J_{4,5}$ = 9.5 Hz, 1 H, 4-H), 4.28 (dd, $^2J_{6a,6b}$ = 12.4 Hz, $^3J_{6a,5}$ = 5.5 Hz, 1 H, 6-H_a), 4.18 (dd, $^2J_{6b,6a}$ = 12.4 Hz, $^3J_{6b,5}$ = 2.4 Hz, 1 H, 6-H_b), 3.93 (ddd, $^3J_{5,4}$ = 10.1 Hz, $^3J_{5,6a}$ = 5.5 Hz, $^3J_{5,6b}$ = 2.4 Hz, 1 H, 5-H), 2.07 (s, 3 H, 6-OC(O)C**H$_3$**), 2.07 (s, 3 H, 2-OC(O)C**H$_3$**), 2.06 (s, 3 H, 4-OC(O)C**H$_3$**), 2.05 (s, 3 H, 3-OC(O)C**H$_3$**).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.6 (6-O**C**(O)CH_3), 170.3 (-O**C**(O)CH_3), 169.5 (4-O**C**(O)CH_3), 169.3 (-O**C**(O)CH_3), 161.3 (C-1'), 143.4 (C-4'), 125.9 (C-3'), 116.8 (C-2'), 98.2 (C-1), 72.6 (C-5), 72.6 (C-3), 71.1 (C-2), 68.2 (C-4), 62.0 (C-6), 20.8 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3).

HRMS (ESI): 492.11068 [M+Na] ($C_{20}H_{23}NO_{12}Na$, calc. 492.11125)

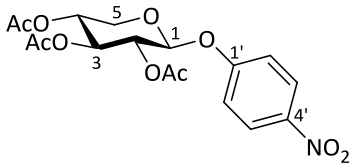
$[\alpha]_D^{25}$ (c, solv.): -43.9 (c 1.00, $CHCl_3$), lit.[458] **$[\alpha]_D^{22}$**-38.6 ($CHCl_3$)

Elem. Analysis (%) calc. for $C_{20}H_{23}NO_{12}$: C 51.18, H 4.94, N 2.98. Found: C 51.39 ± 0.03, H 4.93 ± 0.00, N 3.21 ± 0.01.

4-Nitrophenyl 2,3,4-tri-*O*-acetyl-α-D-xylopyranoside (51)

Glycoside **45** (15.9 g, 50.0 mmol) was reacted according to GPC. The desired product was then crystallised from hot EtOH as colourless filaments. The crystals were then collected by filtration and washed with ice-cold dH_2O and dried in a desiccator to obtain **51** in a yield of 29% (5.80 g, 14.6 mmol)

Data was in agreement with that previously published.[459,460]



m.p. (°C) = 136.5 – 137.4 °C (lit.[459] 138 – 140 °C)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.21 (d, $^3J_{3',2'}$ = 9.1 Hz, 2 H, 3'-H), 7.08 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 5.33 (d, $^3J_{1,2}$ = 5.3 Hz, 1 H, 1-H), 5.24 (dd, $^3J_{3,2}$ = 7.2 Hz, $^3J_{3,4}$ = 7.2 Hz, 1 H, 3-H), 5.19 (dd, $^3J_{2,3}$ = 7.4 Hz, $^3J_{2,1}$ = 5.3 Hz, 1 H, 2-H), 5.00 (ddd, $^3J_{4,3}$ = 6.8 Hz, $^3J_{4,5b}$ = 6.8 Hz, $^3J_{4,5a}$ = 4.3 Hz, 1 H, 4-H), 4.23 (dd, $^2J_{5a,5b}$ = 12.3 Hz, $^3J_{5a,4}$ = 4.3 Hz, 1 H, 5-H_a), 3.62 (dd, $^2J_{5b,5a}$ = 12.3 Hz, $^3J_{5b,4}$ = 6.7 Hz, 1 H, 5-H_b), 2.11 (s, 3 H, -OC(O)C**H**$_3$), 2.10 (s, 3 H, -OC(O)C**H**$_3$), 2.10 (s, 3 H, -OC(O)C**H**$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

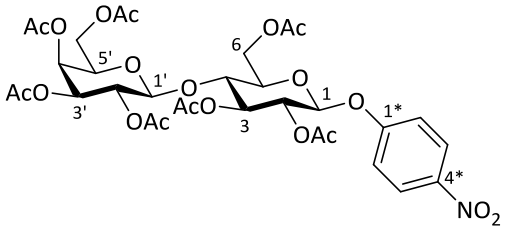
δ (ppm): 169.9 (-O**C**(O)CH_3), 169.8 (-O**C**(O)CH_3), 169.4 (-O**C**(O)CH_3), 161.1 (C-1'), 143.2 (C-4'), 126.0 (C-3'), 116.7 (C-2'), 97.6 (C-1), 69.9 (C-3), 69.5 (C-2), 68.2 (C-4), 61.9 (C-5), 20.9 (-OC(O)**C**H_3), 20.9 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3).

$[\alpha]_D^{25}$ (c, solv.): -69.3 (c 1.03, $CHCl_3$), lit.[459] $[\alpha]_D^{20}$ -74.2 (c 1,00 $CHCl_3$)

4-Nitrophenyl 2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl)-β-D-galactopyranosyl-β-D-glucopyranoside (52)

Glycoside **46** (10.2 g, 15.0 mmol) was reacted as outlined in GPC. The desired product was isolated *via* column chromatography (ethyl acetate/petroleum ether 3:7) to obtain **52** in a yield of 27% (3.03 g, 4.00 mmol).

Data was in agreement with that previously published.[462]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.18 (d, $^{3}J_{3*,2*}$ = 9.2 Hz, 2 H, 3*-H), 7.04 (d, $^{3}J_{2*,3*}$ = 9.2 Hz, 2 H, 2*-H), 5.35 (dd, $^{3}J_{4',3'}$ = 3.5 Hz, $^{3}J_{4',5'}$ = 1.1 Hz, 1 H, 4'-H), 5.29 (ddd, $^{3}J_{H,H}$ = 8.8 Hz, $^{3}J_{H,H}$ = 6.7 Hz, $^{3}J_{H,H}$ = 2.4 Hz, 1 H, 3-H), 5.21-5.16 (m, 2 H, 1-H, 2-H), 5.11 (dd, $^{3}J_{2',3'}$ = 10.5 Hz, $^{3}J_{,2',1'}$ = 7.9 Hz, 1 H, 2'-H), 4.96 (dd, $^{3}J_{3',2'}$ = 10.4 Hz, $^{3}J_{3',4'}$ = 3.5 Hz, 1 H, 3'-H), 4.51 (d, $^{3}J_{1',2'}$ = 7.9 Hz, 1 H, 1'-H), 4.50 (dd, $^{2}J_{6a,6b}$ = 11.1 Hz, $^{3}J_{6a,5}$ = 2.1 Hz, 1 H, 6-H_a), 4.17-4.10 (m, 2 H, 6-H_b, 6'-H_a), 4.07 (dd, $^{2}J_{6'b,6'a}$ = 11.2 Hz, $^{3}J_{6'b,5'}$ = 7.4 Hz, 1 H, 6'-H_b), 3.92-3.87 (m, 2 H, 4-H, 5'-H), 3.84 (ddd, $^{3}J_{5,6b}$ = 9.9 Hz, $^{3}J_{5,4}$ = 5.8 Hz, $^{3}J_{5,6a}$ = 2.2 Hz, 1 H, 5-H), 2.14 (s, 3 H, 4'-OC(O)C**H$_3$**), 2.07 (s, 3 H, 3-OC(O)C**H$_3$**), 2.06 (s, 3 H, 6-OC(O)C**H$_3$**), 2.05 (s, 3 H, 6'-OC(O)C**H$_3$**), 2.04 (s, 6 H, 2-OC(O)C**H$_3$**, 2'-OC(O)C**H$_3$**), 1.95 (s, 3 H, 3'-OC(O)C**H$_3$**).

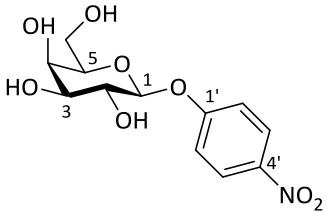
^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.4 (6-O**C**(O)CH_3), 170.2 (6'-O**C**(O)CH_3), 170.2 (4'-O**C**(O)CH_3), 170.1 (3'-O**C**(O)CH_3), 169.8 (3-O**C**(O)CH_3), 169.6 (2-O**C**(O)CH_3), 169.2 (2'-O**C**(O)CH_3), 161.3 (C-1*), 143.3 (C-4*), 125.9 (C-3*), 116.7 (C-2*), 101.3 (C-1'), 97.8 (C-1), 76.1 (C-4), 73.2 (C-5), 72.7 (C-3), 71.3 (C-2), 71.0 (C-3'), 70.9 (C-5'), 69.2 (C-2'), 66.7 (C-4'), 62.0 (C-6), 60.9 (C-6'), 20.9 (-OC(O)**C**H_3), 20.8 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3), 20.7 (-OC(O)**C**H_3), 20.6 (-OC(O)**C**H_3).

4-Nitrophenyl β-D-galactopyranoside (22)

Glycoside **47** (4.69 g, 10.0 mmol) was reacted as outlined in GPD. The crude mixture was not purified any further. Product **22** was obtained as colourless, amorphous solid in a quantitative yield (3.08 g, 10.2 mmol).

Data was in agreement with that previously published.[459]



m.p. (°C) = 167.5 – 168.3 °C

^{1}H NMR (600 MHz, d_6-DMSO):

δ (ppm): 8.20 (d, $^3J_{3',2'}$ = 8.9 Hz, 2 H, 3'-H), 7.22 (d, $^3J_{2',3'}$ = 8.9 Hz, 2 H, 2'-H), 5.30 (d, $^3J_{2\text{-}OH,2\text{-}H}$ = 5.2 Hz, 1 H, 2-OH), 5.03 (d, $^3J_{1,2}$ = 7.6 Hz, 1 H, 1-H), 4.94 (d, $^3J_{3\text{-}OH,3\text{-}H}$ = 5.7 Hz, 1 H, 3-OH), 4.69 (dd, $^3J_{6\text{-}OH,\,6\text{-}Ha}$ = 5.5 Hz, $^3J_{6\text{-}OH,\,6\text{-}Hb}$ = 5.5 Hz, 1 H, 6-OH), 4.58 (d, $^3J_{4\text{-}OH,4\text{-}H}$ = 4.6 Hz, 1 H, 4-OH), 3.73 (ddd, $^3J_{4\text{-}H,4\text{-}OH}$ = 4.6 Hz, $^3J_{4,3}$ = 3.3 Hz, $^3J_{4,5}$ = 1.0 Hz, 1 H, 4-H), 3.67 (ddd, $^3J_{5,6a}$ = 6.2 Hz, $^3J_{5,6b}$ = 6.2 Hz, $^3J_{5,4}$ = 1.0 Hz, 1 H, 5-H), 3.63 (ddd, $^3J_{2,3}$ = 9.5 Hz, $^3J_{2.1}$ = 7.7 Hz, $^3J_{2\text{-}H,2\text{-}OH}$ = 5.2 Hz, 1 H, 2-H), 3.56 (ddd, $^2J_{6a,6b}$ = 11.3 Hz, $^3J_{6a,5}$ = 5.7 Hz, $^3J_{6\text{-}Ha,6\text{-}OH}$ = 5.7 Hz, 1 H, 6-H_a), 3.50 (ddd, $^2J_{6b,6a}$ = 11.8 Hz, $^3J_{6b,5}$ = 5.8 Hz, $^3J_{6\text{-}Hb,6\text{-}OH}$ = 5.8 Hz, 1 H, 6-H_b), 3.45 (ddd, $^3J_{3,2}$ = 9.3 Hz, $^3J_{3\text{-}H,3\text{-}OH}$ = 5.8 Hz, $^3J_{3,4}$ = 3.3 Hz, 1 H, 3-H).

^{13}C NMR (151 MHz, d_6-DMSO):

δ (ppm): 162.5 (s, C-1'), 141.6 (s, C-4'), 125.8 (d, 15.6, C-3'), 116.6 (d, 10.0, C-2'), 100.5 (d, 24.6, C-1), 75.8 (d, 17.3, C-5), 73.2 (s, C-3), 70.1 (d, 14.4, C-2), 68.1 (d, 23.5, C-4), 60.3 (s, C-6).

^{1}H NMR (600 MHz, D_2O):

δ (ppm): 8.27 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.26 (d, $^3J_{2',3'}$ = 8.9 Hz, 2 H, 2'-H), 5.22 (d, $^3J_{1,2}$ = 7.7 Hz, 1 H, 1-H), 4.04 (d, $^3J_{4,3}$ = 3.3 Hz, 1 H, 4-H), 3.95 (dd, $^3J_{5,6a}$ = 6.2 Hz, $^3J_{5,6b}$ = 6.2 Hz, 1 H, 5-H), 3.88 (dd, $^3J_{2,3}$ = 10.0 Hz, $^3J_{2,1}$ = 7.7 Hz, 1 H, 2-H), 3.84-3.77 (m, 3 H, 3-H, 6-H_a, 6-H_b).

^{13}C NMR (151 MHz, D_2O):

δ (ppm): 161.8 (C-1'), 142.5 (C-4'), 142.5 (C-3', C-5'), 116.4 (C-2', C-6'), 100.0 (C-1), 75.6 (C-5), 72.4 (C-3), 70.3 (C-2), 68.4 (C-4), 60.7 (C-6).

HRMS (ESI): 324.06868 [M+Na] ($C_{12}H_{15}NO_8Na$, calc. 324.06899)

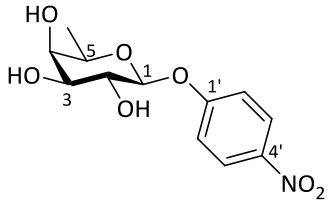
$[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\mathbf{25}}$ (c, solv.): -89.2 (c 1.01, H_2O); -93.4 (c 0.50, H_2O); -84.5 (c 1.00, MeOH), lit.[459] $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\mathbf{20}}$ -80.6 (c 0.5, H_2O)

Elem. Analysis (%) calc. for $C_{12}H_{15}NO_8$: C 47.84, H 5.02, N 4.65. Found: C 46.84 ± 0.16, H 5.13 ± 0.02, N 4.67 ± 0.05.

4-Nitrophenyl β-D-fucopyranoside (25)

Glycoside **48** (116.0 mg, 0.28 mmol) was reacted as outlined in GPD. The crude mixture was not purified any further. Product **25** was obtained as colourless, amorphous solid in a quantitative yield (88.0 mg, 0.31 mmol).

Data was in agreement with that previously published.[466]



^{1}H NMR (600 MHz, D_2O):

δ (ppm): 8.27 (d, $^3J_{3',2'}$ = 8.9 Hz, 2 H, 3'-H), 7.26 (d, $^3J_{2',3'}$ = 8.9 Hz, 2 H, 2'-H), 5.21 (d, $^3J_{1,2}$ = 7.2 Hz, 1 H, 1-H), 4.06 (q, $^3J_{5,6}$ = 6.6 Hz, 1 H, 5-H), 3.85 (d, $^3J_{4,3}$ = 3.4 Hz, 1 H, 4-H), 3.84-3.78 (m, 2 H, 2-H, 3-H), 1.30 (d, $^3J_{6,5}$ = 6.5 Hz, 3 H, 6-C**H**$_3$).

^{13}C NMR (151 MHz, D_2O):

δ (ppm): 161.8 (C-1'), 142.4 (C-4'), 126.1 (C-3'), 116.3 (C-2'), 99.7 (C-1), 72.5 (C-3), 71.5 (C-5), 71.1 (C-4), 70.1 (C-2), 15.3 (C-6).

^{1}H NMR (600 MHz, d_6-DMSO):

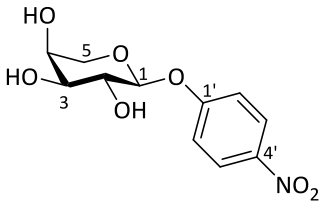
δ (ppm): 8.21 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.21 (d, $^3J_{2',3'}$ = 9.3 Hz, 2 H, 2'-H), 5.24 (d, $^3J_{2\text{-}OH,2\text{-}H}$ = 5.2 Hz, 1 H, 2-OH), 5.03 (d, $^3J_{1,2}$ = 7.7 Hz, 1 H, 1-H), 4.87 (d, $^3J_{3\text{-}OH,3\text{-}H}$ = 5.9 Hz, 1 H, 3-OH), 4.60 (d, $^3J_{4\text{-}OH,4\text{-}H}$ = 4.8 Hz, 1 H, 4-OH), 3.84 (q, $^3J_{5,6}$ = 6.4 Hz, 1 H, 5-H), 3.58 (ddd, $^3J_{2,3}$ = 9.6 Hz, $^3J_{2,1}$ = 7.6 Hz, $^3J_{2\text{-}H,2\text{-}OH}$ = 5.2 Hz, 1 H, 2-H), 3.49 (ddd, $^3J_{4\text{-}H,4\text{-}OH}$ = 4.7 Hz, $^3J_{4,3}$ = 3.4 Hz, $^3J_{4,5}$ = 1.1 Hz, 1 H, 4-H), 3.44 (ddd, $^3J_{3,2}$ = 9.4 Hz, $^3J_{3\text{-}H,3\text{-}OH}$ = 5.9 Hz, $^3J_{3,4}$ = 3.3 Hz, 1 H, 3-H), 1.14 (d, $^3J_{6,5}$ = 6.4 Hz, 3 H, 6-C**H**$_3$).

^{13}C NMR (151 MHz, d_6-DMSO):

δ (ppm): 162.5 (C-1‘), 141.5 (C-4‘), 125.7 (C-3‘), 116.4 (C-2‘), 100.2 (C-1), 73.2 (C-3), 70.8 (C-4), 70.6 (C-5), 69.7 (C-2), 16.5 (C-6).

4-Nitrophenyl α-L-arabinopyranoside (26)

Glycoside **49** (51.0 mg, 0.13 mmol) was reacted as outlined in GPD. The crude mixture was not purified any further. Product **26** was obtained as colourless, amorphous solid in a quantitative yield (40.0 g, 0.15 mmol).



^{1}H NMR (600 MHz, D_2O):

δ (ppm): 8.28 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3‘-H), 7.26 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2‘-H), 5.18 (d, $^3J_{1,2}$ = 7.5 Hz, 1 H, 1-H), 4.05 (ddd, $^3J_{4,3}$ = 3.6 Hz, $^3J_{4,5a}$ = 2.6 Hz, $^3J_{4,5b}$ = 1.3 Hz, 1 H, 4-H), 4.02 (dd, $^2J_{5a,5b}$ = 13.0 Hz, $^3J_{5a,4}$ = 2.2 Hz, 1 H, 5-H_a), 3.93-3.87 (m, 2 H, 2-H, 5-H_b), 3.83 (dd, $^3J_{3,2}$ = 9.7 Hz, $^3J_{3,4}$ = 3.4 Hz, 1 H, 3-H).

^{13}C NMR (151 MHz, D_2O):

δ (ppm): 161.7 (C-1‘), 142.5 (C-4‘), 126.1 (C-3‘), 116.4 (C-2‘), 100.1 (C-1), 71.9 (C-3), 70.2 (C-2), 68.0 (C-4), 66.6 (C-5).

^{1}H NMR (600 MHz, d_6-DMSO):

δ (ppm): 8.21 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3‘-H), 7.21 (d, $^3J_{2',3'}$ = 9.3 Hz, 2 H, 2‘-H), 5.32 (d, $^3J_{2\text{-}OH,2\text{-}H}$ = 5.1 Hz, 1 H, 2-OH), 5.05 (d, $^3J_{1,2}$ = 6.7 Hz, 1 H, 1-H), 4.90 (d, $^3J_{4\text{-}OH,4\text{-}H}$ = 5.6 Hz, 1 H, 4-OH), 4.70 (d, $^3J_{3\text{-}OH,3\text{-}H}$ = 4.1 Hz, 1 H, 3-OH), 3.78-3.70 (m, 2 H, 3-H, 5-H_a), 3.69-3.62 (m, 2 H, 2-H, 5-H_b), 3.49 (ddd, $^3J_{4,3}$ = 8.8 Hz, $^3J_{4,5a}$ = 5.6 Hz, $^3J_{4,5b}$ = 3.3 Hz, 1 H, 4-H).

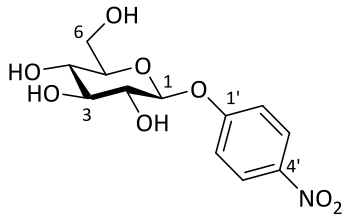
^{13}C NMR (151 MHz, d_6-DMSO):

δ (ppm): 162.3 (C-1‘), 141.6 (C-4‘), 125.7 (C-3‘), 116.5 (C-2‘), 100.2 (C-1), 72.2 (C-4), 70.1 (C-2), 67.3 (C-3), 65.8 (C-5).

4-Nitrophenyl β-D-glucopyranoside (4)

Glycoside **50** (2.35 g, 5.00 mmol) was reacted as outlined in GPD. The crude mixture was then recrystallised from hot EtOH to return product **4** as a colourless, amorphous solid in a quantitative yield.

Data was in agreement with that previously published.[459,463]



m.p. (°C) = 166.8 – 167.8 °C (lit.[459] 166 – 167 °C)

^{1}H NMR (600 MHz, D_2O):

δ (ppm): 8.28 (d, $^3J_{3',2'}$ = 9.1 Hz, 2 H, 3'-H), 7.26 (d, $^3J_{2',3'}$ = 9.1 Hz, 2 H, 2'-H), 5.28 (d, $^3J_{1,2}$ = 6.8 Hz, 1 H, 1-H), 3.96 (dd, $^2J_{6a,6b}$ = 12.5 Hz, $^3J_{6a,5}$ = 2.2 Hz, 1 H, 6-H_a), 3.77 (dd, $^2J_{6b,6a}$ = 12.5 Hz, $^3J_{6b,5}$ = 5.7 Hz, 1 H, 6-H_b), 3.71 (ddd, $^3J_{5,4}$ = 10.0 Hz, $^3J_{5,6b}$ = 5.9 Hz, $^3J_{5,6a}$ = 2.5 Hz, 1 H, 5-H), 3.67-3.62 (m, 2 H, 2-H, 3-H), 3.57-3.50 (m, 1 H, 4-H).

^{13}C NMR (151 MHz, D_2O):

δ (ppm): 161.7 (C-1'), 142.6 (C-4'), 126.1 (C-3'), 116.4 (C-2'), 99.4 (C-1), 76.2 (C-5), 75.4 (C-3), 72.7 (C-2), 69.3 (C-4), 60.4 (C-6).

^{1}H NMR (600 MHz, d_6-DMSO):

δ (ppm): 8.21 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.22 (d, $^3J_{2',3'}$ = 9.3 Hz, 2 H, 2'-H), 5.43 (d, $^3J_{2\text{-}OH,2\text{-}H}$ = 4.7 Hz, 1 H, 2-OH), 5.14 (d, $^3J_{3\text{-}OH,3\text{-}H}$ = 4.6 Hz, 1 H, 3-OH), 5.08 (d, $^3J_{1,2}$ = 7.9 Hz, 1 H, 1-H), 5.07 (d, $^3J_{4\text{-}OH,4\text{-}H}$ = 5.4 Hz, 1 H, 4-OH), 4.57 (dd, $^3J_{6\text{-}OH,\ 6\text{-}Ha}$ = 5.8 Hz, $^3J_{6\text{-}OH,\ 6\text{-}Hb}$ = 5.8 Hz, 1 H, 6-OH), 3.69 (ddd, $^2J_{6a,6b}$ = 11.8 Hz, $^3J_{6\text{-}Ha,6\text{-}OH}$ = 5.4 Hz, $^3J_{6a,5}$ = 2.0 Hz, 1 H, 6-H_a), 3.47 (ddd, $^2J_{6b,6a}$ = 11.8 Hz, $^3J_{6\text{-}Hb,6\text{-}OH}$ = 5.9 Hz, $^3J_{6b,5}$ = 5.9 Hz, 1 H, 6-H_b), 3.41 (ddd, $^3J_{5,4}$ = 9.8 Hz, $^3J_{5,6b}$ = 5.6 Hz, $^3J_{5,6a}$= 2.0 Hz,

1 H, 5-H), 3.32-3.25 (m, 2 H, 2-H, 3-H), 3.19 (ddd, $^{3}J_{4,5}$ = 9.6 Hz, $^{3}J_{4,3}$ = 8.3 Hz, $^{3}J_{4\text{-}H,4\text{-}OH}$ = 5.2 Hz, 1 H, 4-H).

^{13}C NMR (151 MHz, d_6-DMSO):

δ (ppm): 162.4 (C-1‘), 141.6 (C-4‘), 125.7 (C-3‘), 116.5 (C-2‘), 99.8 (C-1), 77.2 (C-5), 76.4 (C-3), 73.1 (C-2), 69.5 (C-4), 60.5 (C-6).

HRMS (ESI): 324.06891 [M+Na] ($C_{12}H_{15}NO_8Na$, calc. 324.06899)

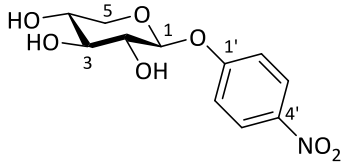
$[\alpha]_D^{25}$ (c, solv.): -96.2 (c 1.01, H_2O); -91.9 (c 1.00, MeOH), lit.[459] $[\alpha]_D^{20}$ -102.3 (c 1,00 H_2O)

Elem. Analysis (%) calc. for $C_{12}H_{15}NO_8$: C 47.84, H 5.02, N 4.65. Found: C 45.36 ± 0.02, H 5.38 ± 0.00, N 4.57 ± 0.03.

4-Nitrophenyl β-D-xylopyranoside (27)

Glycoside **51** (1.20 g, 3.00 mmol) was reacted as outlined in GPD. The crude mixture was not purified any further. Product **27** was obtained as colourless, amorphous solid in a quantitative yield (1.04 g, 3.8 mmol).

Data was in agreement with that previously published.[459,464,465]



m.p. (°C) = 159.9 – 160.1 °C (lit.[464] 158 °C)

^{1}H NMR (600 MHz, D_2O):

δ (ppm): 8.27 (d, $^{3}J_{3',2'}$ = 9.3 Hz, 2 H, 3‘-H), 7.25 (d, $^{3}J_{2',3'}$ = 9.3 Hz, 2 H, 2‘-H), 5.23 (d, $^{3}J_{1,2}$ = 7.3 Hz, 1 H, 1-H), 4.06 (dd, $^{2}J_{5a,5b}$ = 11.6 Hz, $^{3}J_{5a,4}$ = 5.4 Hz, 1 H, 5-H_a), 3.74 (ddd, $^{3}J_{4,5b}$ = 10.3 Hz, $^{3}J_{4,3}$ =

8.7 Hz, $^{3}J_{4,5a}$ = 5.4 Hz, 1 H, 4-H), 3.63 (dd, $^{3}J_{2,3}$ = 9.3 Hz, $^{3}J_{2,1}$ = 7.4 Hz, 1 H, 2-H), 3.59 (dd, $^{3}J_{3,2}$ = 9.1 Hz, $^{3}J_{3,4}$ = 9.1 Hz, 1 H, 3-H), 3.55 (dd, $^{2}J_{5b,5a}$ = 11.0 Hz, $^{3}J_{5b,4}$ = 11.0 Hz, 1 H, 5-H_b).

^{13}C NMR (151 MHz, D_2O):

δ (ppm): 161.5 (C-1‘), 142.6 (C-4‘), 126.1 (C-3‘), 116.4 (C-2‘), 99.9 (C-1), 75.3 (C-3), 72.6 (C-2), 68.9 (C-4), 65.3 (C-5).

^{1}H NMR (600 MHz, d_6-DMSO):

δ (ppm): 8.21 (d, $^{3}J_{3',2'}$ = 9.2 Hz, 2 H, 3‘-H), 7.21 (d, $^{3}J_{2',3'}$ = 9.2 Hz, 2 H, 2‘-H), 5.46 (d, $^{3}J_{2\text{-}OH,2\text{-}H}$ = 4.9 Hz, 1 H, 2-OH), 5.18 (d, $^{3}J_{3\text{-}OH,3\text{-}H}$ = 4.6 Hz, 1 H, 3-OH), 5.11 (d, $^{3}J_{4\text{-}OH,4\text{-}H}$ = 4.5 Hz, 1 H, 4-OH), 5.08 (d, $^{3}J_{1,2}$ = 7.1 Hz, 1 H, 1-H), 3.77 (dd, $^{2}J_{5a,5b}$ = 9.4 Hz, $^{3}J_{5a,4}$ = 3.4 Hz, 1 H, 5-H_a), 3.41-3.34 (m, 2 H, 4-H, 5-H_b), 3.31-3.23 (m, 2 H, 2-H, 3-H).

^{13}C NMR (151 MHz, d_6-DMSO):

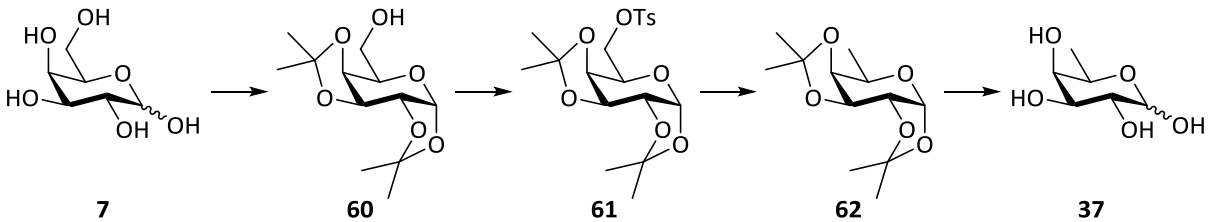
δ (ppm): 162.7 (C-1‘), 142.2 (C-4‘), 126.2 (C-3‘), 117.0 (C-2‘), 100.8 (C-1), 76.7 (C-4), 73.3 (C-2), 69.7 (C-3), 66.3 (C-5).

HRMS (ESI): 294.05835 [M+Na] ($C_{11}H_{13}NO_7Na$, calc. 294.05824)

$[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\mathbf{25}}$ (c, solv.): -57.2 (c 1.02, MeOH), lit.[459] $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\mathbf{20}}$-56.8 (c 0.50, H_2O)

Elem. Analysis (%) calc. for $C_{11}H_{13}NO_7$: C 48.71, H 4.83, N 5.16. Found: C 48.72 ± 0.18, H 4.88 ± 0.02, N 5.29 ± 0.05.

5.4.5 Synthesis of D-fucopyranose



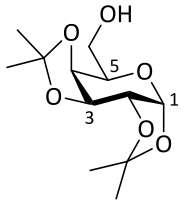
7 **60** **61** **62** **37**

1,2:3,4-Di-*O*-isopropylidine-α-D-galactopyranose (60)

This procedure is adapted from a previously published protocol.[354]

In a round-bottomed flask containing $ZnCl_2$ (90.0 g, 67.0 mmol, 0.16 eq.) in acetone (900 mL) was added a catalytic amount of H_2SO_4 (3.00 mL) and D-galactose (7; 75.0 g, 416.3 mmol). The mixture was allowed to stir at room temperature overnight. Following completion as determined by TLC, the reaction was cooled to 0 °C and quenched using sat. $NaHCO_{3\,(aq)}$ (900 mL) and subsequently allowed to warm to room temperature and stirred a further 2 hrs. The reaction mixture was then filtered over celite and concentrated under reduced pressure. The product was then extracted using Et_2O (3x 300 mL) until no yellow colour was observable in the organic layer. The organic phase was then dried with $MgSO_4$, filtered, and the solvent removed under reduced pressure to obtain **60** in a quantitative yield (110 g, 423 mmol).

Data was in agreement with that previously published.[354,467]



R_f = 0.38 (ethyl acetate/petroleum ether 1:1)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.57 (d, $^3J_{1,2}$ = 5.0 Hz, 1 H, 1-H), 4.62 (dd, $^3J_{3,4}$ = 7.9 Hz, $^3J_{2,3}$ = 2.4 Hz, 1 H, 3-H), 4.34 (dd, $^3J_{2,1}$ = 5.0 Hz, $^3J_{2,3}$ = 2.4 Hz, 1 H, 2-H), 4.27 (dd, $^3J_{4,3}$ = 8.1 Hz, $^3J_{4,5}$ = 1.5 Hz, 1 H, 4-H), 3.91-3.83 (m, 2 H, 5-H, 6-H_a), 3.79-3.70 (m, 1 H, 6-H_b), 1.53 (s, 3 H, 1,2-C(CH_3)$_a$), 1.46 (s, 3 H, 3,4-C(CH_3)$_a$), 1.34 (s, 6 H, 1,2-C(CH_3)$_b$, 3,4-C(CH_3)$_b$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 109.6 (3,4-**C**(CH_3)$_2$), 108.8 (1,2-**C**(CH_3)$_2$), 96.5 (C-1), 71.8 (C-4), 70.9 (C-3), 70.7 (C-2), 68.2 (C-5), 62.6 (C-6), 26.2 (1,2-C(**C**H_3)$_a$), 26.1 (3,4-C(**C**H_3)$_a$), 25.1 (1,2-C(**C**H_3)$_b$), 24.5 (3,4-C(**C**H_3)$_b$).

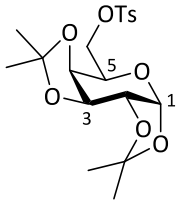
HRMS (ESI): 283.11518 [M+Na] ($C_{12}H_{20}O_6Na$, calc. 283.11521)

1,2:3,4-Di-*O*-isopropylidene-6-*O*-tosyl-α-D-galactopyranose (61)

This procedure is adapted from a previously published protocol.[468,469]

Glycoside **60** (108.0 g, 414.9 mmol) and DMAP (5.12 g, 41.5 mmol, 0.10 eq.) were dissolved in anh. py (800 mL). To this was added dropwise a solution of TsCl (118.6 g, 622.4 mmol, 1.50 eq.) dissolved in anh. CH_2Cl_2 (80 mL). After complete addition of the TsCl solution, the mixture was allowed to stir at room temperature overnight. Following complete conversion as monitored by TLC, the reaction was quenched with water and the reaction concentrated under reduced pressure with toluene. Once sufficiently concentrated, the residue was redissolved in CH_2Cl_2 (100 mL) and washed with water (3x 50 mL), sat. $HCO_{3\,(aq)}$ (3x 50 mL), and brine (1x 100 mL). The organic phase was then dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure to obtain compound **61** in a yield of 66% (114 g, 275 mmol).

Data was in agreement with that previously published.[468]



R_f = 0.36 (*n*-pentane/ethyl acetate 5:1)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 7.81 (d, $^3J_{3',2'}$ = 8.1 Hz, 2 H, 3'-H), 7.33 (d, $^3J_{2',3'}$ = 8.0 Hz, 2 H, 2'-H), 5.45 (d, $^3J_{1,2}$ = 4.9 Hz, 1 H, 1-H), 4.59 (dd, $^3J_{3,4}$ = 8.1 Hz, $^3J_{3,2}$ = 2.5 Hz, 1 H, 3-H), 4.29 (dd, $^3J_{2,1}$ = 5.0 Hz, $^3J_{2,3}$ = 2.5 Hz, 1 H, 2-H), 4.23-4.17 (m, 2 H, 4-H, 6-H_a), 4.09 (dd, $^2J_{6b,6a}$ = 10.2 Hz, $^3J_{6b,5}$ = 6.8 Hz, 1 H, 6-H_b), 4.04 (ddd, $^3J_{5,6b}$ = 6.6 Hz, $^3J_{5,6a}$ = 6.1 Hz, $^3J_{5,4}$ = 1.7 Hz, 1 H, 5-H), 2.44 (s, 3 H, 4'-CH_3), 1.50 (s, 3 H, 1,2-C(CH_3)$_a$), 1.34 (s, 3 H, 3,4-C(CH_3)$_a$), 1.31 (s, 3 H, 1,2-C(CH_3)$_b$), 1.28 (s, 3 H, 3,4-C(CH_3)$_b$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 144.9 (C-4'), 133.0 (C-1'), 129.9 (C-2'), 128.3 (C-3'), 109.7 (3,4-**C**(CH_3)$_2$). 109.1 (1,2-**C**(CH_3)$_2$), 96.3 (C-1), 70.7 (C-3), 70.6 (C-4), 70.5 (C-2), 68.3 (C-6), 66.0 (C-5), 26.1 (1,2-C(**C**H_3)$_a$), 26.0 (3,4-C(**C**H_3)$_a$), 25.1 (1,2-C(**C**H_3)$_b$), 24.5 (3,4-C(**C**H_3)$_b$), 21.8 (4'-CH_3).

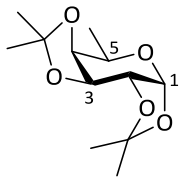
HRMS (ESI): 437.12396 [M+Na] ($C_{19}H_{26}O_8SNa$, calc. 437.12406)

1,2:3,4-Di-*O*-isopropylidene-α-D-fucopyranose (62)

This procedure is adapted from a previously published protocol.[355-357]

Glycoside **62** (2.00 g, 5.00 mmol) was dissolved in THF (5 mL) in a round-bottomed flask. To this was added a 1 M aqueous solution of Super-Hydride (Li[Et_3BH], 20.0 mL, 10.0 eq.) and the reaction was allowed to stir at room temperature overnight before being heated to reflux for a further 3 hrs. Following completion, the reaction mixture was poured into water (100 mL) and extracted with EtOAc (3x 50 mL). The organic phases were then combined and dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. The resulting mixture was then purified using column chromatograpy (ethyl acetate/petroleum ether 2:8) to obtain product **62** in a yield of 94% (1.08 g, 4.7 mmol).

Data was in agreement with that previously published.[470]



$\mathbf{R_f}$ = 0.62 (ethyl acetate/petroleum ether 4:6)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 5.48 (d, $^3J_{1,2}$ = 5.3 Hz, 1 H, 1-H), 4.55 (dd, $^3J_{3,4}$ = 7.9 Hz, $^3J_{3,2}$ = 2.4 Hz, 1 H, 3-H), 4.25 (dd, $^3J_{2,1}$ = 5.2 Hz, $^3J_{2,3}$ = 2.4 Hz, 1 H, 2-H), 4.04 (dd, $^3J_{4,3}$ = 8.0 Hz, $^3J_{4,5}$ = 2.0 Hz, 1 H, 4-H), 3.87 (dq, $^3J_{5,6}$ = 6.6 Hz, $^3J_{5,4}$ = 1.9 Hz, 1 H, 5-H), 1.48 (s, 3 H, 1,2-C(CH_3)$_a$), 1.43 (s, 3 H, 3,4-C(CH_3)$_a$), 1.31 (s, 3 H, 1,2-C(CH_3)$_b$), 1.29 (s, 3 H, 3,4-C(CH_3)$_b$), 1.21 (d, $^3J_{6,5}$ = 6.6 Hz, 3 H, 6-C**H$_3$**).

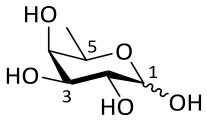
^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 109.0 (3,4-**C**(CH_3)$_2$). 108.3 (1,2-**C**(CH_3)$_2$), 96.6 (C-1), 73.6 (C-4), 71.0 (C-3), 70.4 (C-2), 63.5 (C-5), 26.1 (1,2-C(**C**H_3)$_a$), 26.1 (3,4-C(**C**H_3)$_a$), 25.0 (1,2-C(**C**H_3)$_b$), 24.5 (3,4-C(**C**H_3)$_b$), 16.0 (C-6).

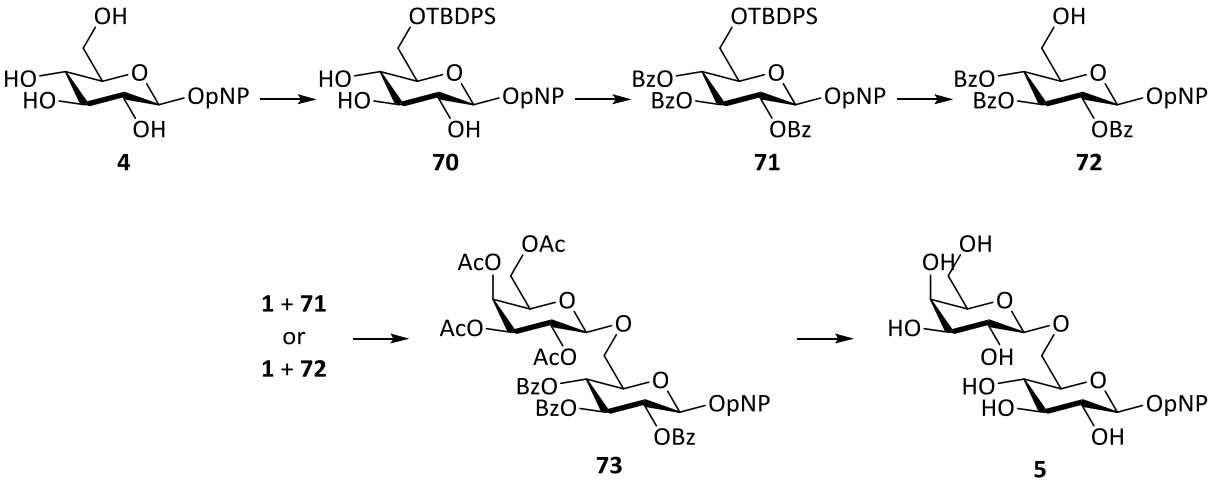
D-Fucose (37)

This procedure is adapted from a previously published protocol.[471]

To a round-bottomed flask containing glycoside **62** (17.0 g, 73.8 mmol) was added AcOH (80%, 400 mL) and the reaction was allowed to stir for 3 days at 80 °C. The mixture was subsequently concentrated under reduced pressure to a crude foam, which subsequently formed a gum under atmospheric pressure. Compound **37** was obtained in a quantitative yield (11.6 g, 77.0 mmol) and used in its crude form in further experiments.



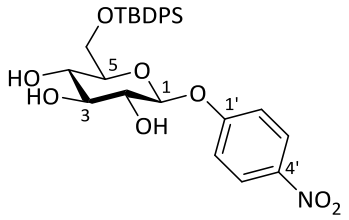
5.4.6 Chemical Synthesis of *para*-Nitrophenyl Allolactose



4-Nitrophenyl 6-*O*-(*tert*-butyldiphenyl)silyl-β-D-glucopyranoside (70)

This procedure is adapted from a previously published protocol.[381]

To a dry round-bottomed flask with a nitrogen atmosphere was added glc-pNP (**4**; 0.60 g, 2.00 mmol) and anh. DMF (40 mL). The stirring solution was then cooled to 0 °C. Et_3N (0.39 mL, 2.80 mmol, 1.40 eq.) and DMAP (19.5 mg, 0.20 mmol, 8.00 mol%) were then added. A solution of TBDPSCl (0.55 mL, 2.10 mmol, 1.10 eq.) was then added *via* syringe pump over 1 h. Subsequently, the reaction was heated to 50 °C and reacted for 3 d. The reaction was then diluted with EtOAc (100 mL) and washed successively with dH_2O (4x 100 mL) and brine (2x 100 mL). The organic layer was then dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. The crude material was purified using column chromatography (EtOAc) to obtain **70** in a yield of 60% (0.64 g, 1.20 mmol).



R_f = 0.23 (ethyl acetate/petroleum ether 8:2)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.13 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.67-7.61 (m, 4 H, arom.-C**H**), 7.45-7.37 (m, 2 H, arom.-C**H**), 7.36-7.27 (m, 4 H, arom.-C**H**), 7.09 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 5.01 (d, $^3J_{1,2}$ = 7.2 Hz, 1 H, 1-H), 3.98 (dd, $^2J_{6a,6b}$ = 11.0 Hz, $^3J_{6a,5}$ = 4.1 Hz, 1 H, 6-H_a), 3.90 (dd, $^2J_{6b,6a}$ = 11.0 Hz, $^3J_{6b,5}$ = 5.5 Hz, 1 H, 6-H_b), 3.75-3.66 (m, 3 H, 2-H, 3-H, 4-H), 3.61 (ddd, $^3J_{5,4}$ = 9.5 Hz, $^3J_{5,6b}$ = 5.5 Hz, $^3J_{5,6a}$ = 4.1 Hz, 1 H, 5-H), 1.06 (s, 9 H, -SiC(C**H$_3$**)$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

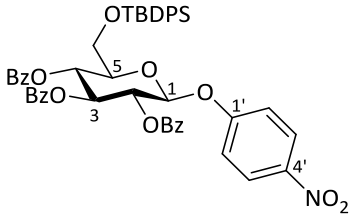
δ (ppm): 161.8 (C-1'), 143.0 (C-4'), 135.7 (arom.-**C**H), 135.7 (arom.-**C**H), 132.9 (arom-**C**$_{ipso}$), 132.6 (arom-**C**$_{ipso}$), 130.1 (arom.-**C**H), 127.9 (arom.-**C**H), 127.9 (arom.-**C**H), 125.9 (C-3'), 116.7 (C-2'), 100.0 (C-1), 76.4 (C-3), 76.0 (C-5), 73.4 (C-2), 71.3 (C-4), 64.2 (C-6), 26.9 (-SiC(**C**H$_3$)$_3$), 19.4 (-Si**C**(CH$_3$)$_3$).

4-Nitrophenyl 2,3,4-tri-*O*-benzoyl-6-*O*-(*tert*-butyldiphenyl)silyl-β-D-glucopyranoside (71)

This procedure is adapted from a previously published protocol.[472]

To a stirring solution of glycoside **70** (0.54 g, 1.00 mmol) dissolved in anh. CH_2Cl_2 (10 mL) was added TMP (0.53 mL, 4.00 mmol, 4.00 eq.). BzCl (0.70 mL, 6.00 mmol, 6.00 eq.) was then added dropwise, followed by DMAP (12.2 mg, 10.0 mol%) and the reaction was left stirring at room temperature overnight. The reaction was quenched using ice-cold dH_2O (5 mL) and the product extracted into CH_2Cl_2 (50 mL). The organic extract was then washed sequentially with dH_2O (40 mL), sat. $NaHCO_{3\,(aq)}$ (2x 40 mL), and brine (20 mL) before being dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. The crude mixture was purified using column chromatography (ethyl acetate/petroleum ether 2:8) to give **71** in a yield of 60% (0.51 g, 0.60 mmol).

Data was in agreement with that previously published.[385]



R_f = 0.25 (diethyl ether/petroleum ether 4:6), 0.39 (petroleum ether/toluene/CH_2Cl_2 2:3:5)

^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.10 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.96 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 8.3 Hz, $^4J_{o\text{-}Bz,p\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**$_o$), 7.87 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 7.5 Hz, $^4J_{o\text{-}Bz,p\text{-}Bz}$ = 1.1 Hz, 2 H, arom.-C**H**$_o$), 7.85 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 7.3 Hz, $^4J_{o\text{-}Bz,p\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**$_o$), 7.61 (dd, $^3J_{o\text{-}Ph,m\text{-}Ph}$ = 8.0 Hz, $^4J_{o\text{-}Ph,p\text{-}Ph}$ = 1.5 Hz, 2 H, arom.-C**H**$_o$), 7.58-7.50 (m, 4 H, arom.-C**H**$_o$, arom.-C**H**$_p$), 7.47 (m, 1 H, arom.-C**H**$_p$), 7.41-7.33 (m, 5 H, arom.-C**H**$_m$, arom.-C**H**$_p$), 7.33-7.28 (m, 3 H, arom.-C**H**$_m$, arom.-C**H**$_p$), 7.29-7.15 (m, 4 H, arom.-C**H**$_m$), 7.14 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 5.96 (dd, $^3J_{3,4}$ = 9.5 Hz, $^3J_{3,2}$ = 9.5 Hz, 1 H, 3-H), 5.80 (dd, $^3J_{2,3}$ = 9.6 Hz, $^3J_{2,1}$ = 7.7 Hz, 1 H, 2-H), 5.66 (dd, $^3J_{4,5}$ = 9.7 Hz, $^3J_{4,3}$ = 9.7 Hz, 1 H, 4-H), 5.50 (d, $^3J_{1,2}$ = 7.6 Hz, 1 H, 1-H), 4.11 (ddd, $^3J_{5,4}$ = 9.9 Hz, $^3J_{5,6a}$ = 5.6 Hz, $^3J_{5,6b}$ = 2.6 Hz, 1 H, 5-H), 3.92 (dd, $^2J_{6a,6b}$ = 11.7 Hz, $^3J_{6a,5}$ = 5.6 Hz, 1 H, 6-H_a), 3.89 (dd, $^2J_{6b,6a}$ = 11.7 Hz, $^3J_{6b,5}$ = 2.6 Hz, 1 H, 6-H_b), 1.05 (s, 9 H, 6-OSiC(C**H**$_3$)$_3$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 165.9 (3-O**C**(O)Ph), 165.2 (-O**C**(O)Ph), 165.1 (-O**C**(O)Ph), 161.6 (C-1'), 143.3 (C-4'), 135.7 (arom.-**C**H), 135.5 (arom.-**C**H), 133.6 (arom.-**C**H), 133.6 (arom.-**C**H), 133.5 (arom.-**C**H), 133.0 (arom.-**C**H), 132.6 (arom.-**C**H), 129.9 (arom.-**C**H), 129.9 (arom.-**C**H), 129.9 (arom.-**C**H), 129.2 (arom.-**C**$_{ipso}$), 129.0 (arom.-**C**$_{ipso}$), 129.0 (arom.-**C**$_{ipso}$), 128.8 (arom.-**C**$_{ipso}$), 128.6 (arom.-**C**$_{ipso}$), 128.6 (arom.-**C**H), 128.5 (arom.-**C**H), 128.4 (arom.-**C**$_{ipso}$), 127.8 (arom.-**C**H), 127.8 (arom.-**C**H), 125.9 (C-3'), 117.1 (C-2'), 98.9 (C-1), 76.3 (C-5), 73.0 (C-3), 71.8 (C-2), 69.0 (C-4), 62.8 (C-6), 26.8 (6-OSiC(**C**H$_3$)$_3$), 19.3 (6-OSi**C**(CH$_3$)$_3$).

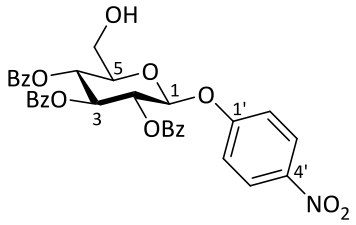
4-Nitrophenyl 2,3,4-tri-*O*-benzoyl-β-D-glucopyranoside (72)

This procedure is adapted from a previously published protocol.[385]

Glycoside **71** (426.0 mg, 0.50 mmol, 1.00 eq.) was dissolved in MeCN (2.50 mL). To this was added triflic acid-on-silica (TfOH·SiO_2, 2.00 mmol/g; 32.0 mg, 10.0 mol%) and the reaction stirred at 50 °C

for 30 min. The reaction was then filtered and the solid washed with MeCN. The filtrate was collected and the solvent removed under reduced pressure. The crude mixture was purified using column chromatography (ethyl acetate/petroleum ether 3:7) to obtain **72** in a yield of 61% (187.0 mg, 0.30 mmol).

Data was in agreement with that previously published.[385]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.18 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.96 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 8.3 Hz, $^4J_{o\text{-}Bz,\ p\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**$_o$), 7.94 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 8.3 Hz, $^4J_{o\text{-}Bz,\ p\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**$_o$), 7.86 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 8.4 Hz, $^4J_{o\text{-}Bz,\ p\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**$_o$), 7.55 (tt, $^3J_{p\text{-}Bz,\ m\text{-}Bz}$ = 7.5 Hz, $^4J_{\ p\text{-}Bz,\ o\text{-}Bz}$ = 1.3 Hz, 1 H, arom.-C**H**$_p$), 7.52 (tt, $^3J_{p\text{-}Bz,\ m\text{-}Bz}$ = 7.4 Hz, $^4J_{\ p\text{-}Bz,\ o\text{-}Bz}$ = 1.3 Hz, 1 H, arom.-C**H**$_p$), 7.45 (tt, $^3J_{p\text{-}Bz,\ m\text{-}Bz}$ = 7.5 Hz, $^4J_{\ p\text{-}Bz,\ o\text{-}Bz}$ = 1.3 Hz, 1 H, arom.-C**H**$_p$), 7.41 (dd, $^3J_{m\text{-}Bz,\ o\text{-}Bz}$ = 8.3 Hz, $^3J_{m\text{-}Bz,\ p\text{-}Bz}$ = 7.3 Hz, 2 H, arom.-C**H**$_m$), 7.37 (dd, $^3J_{m\text{-}Bz,\ o\text{-}Bz}$ = 8.3 Hz, $^3J_{m\text{-}Bz,\ p\text{-}Bz}$ = 7.4 Hz, 2 H, arom.-C**H**$_m$), 7.31 (dd, $^3J_{m\text{-}Bz,\ o\text{-}Bz}$ = 8.3 Hz, $^3J_{m\text{-}Bz,\ p\text{-}Bz}$ = 7.4 Hz, 2 H, arom.-C**H**$_m$), 7.10 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 6.06 (dd, $^3J_{3,4}$ = 9.6 Hz, $^3J_{3,2}$ = 9.6 Hz, 1 H, 3-H), 5.82 (dd, $^3J_{2,3}$ = 9.7 Hz, $^3J_{2,1}$ = 7.8 Hz, 1 H, 2-H), 5.60 (dd, $^3J_{4,5}$ = 9.7 Hz, $^3J_{4,3}$ = 9.7 Hz, 1 H, 4-H), 5.55 (d, $^3J_{1,2}$ = 7.8 Hz, 1 H, 1-H), 4.02 (ddd, $^3J_{5,4}$ = 9.9 Hz, $^3J_{5,6b}$ = 4.6 Hz, $^3J_{5,6a}$ = 2.3 Hz, 1 H, 5-H), 3.92 (dd, $^2J_{6a,6b}$ = 13.0 Hz, $^3J_{6a,5}$ = 2.3 Hz, 1 H, 6-H$_a$), 3.81 (dd, $^2J_{6b,6a}$ = 13.0 Hz, $^3J_{6b,5}$ = 4.6 Hz, 1 H, 6-H$_b$).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 166.2 (4-O**C**(O)Ph), 165.9 (3-O**C**(O)Ph), 165.1 (2-O**C**(O)Ph), 161.4 (C-1'), 143.4 (C-4'), 134.0 (arom.-**C**H), 133.7 (arom.-**C**H), 133.6 (arom.-**C**H), 130.1 (arom.-**C**H), 129.9 (arom.-**C**H), 129.9 (arom.-**C**H), 128.9 (arom.-**C**$_{ipso}$), 128.7 (arom.-**C**H), 128.7 (arom.-**C**$_{ipso}$), 128.6 (arom.-**C**H), 128.5 (arom.-**C**H), 128.4 (arom.-**C**$_{ipso}$), 126.0 (C-3'), 116.8 (C-2'), 98.7 (C-1), 75.5 (C-5), 72.5 (C-3), 71.6 (C-2), 69.2 (C-4), 61.3 (C-6).

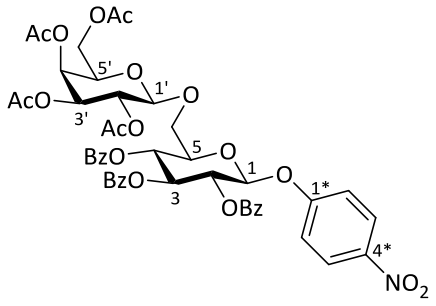
4-Nitrophenyl 2,3,4-tri-*O*-benzoyl-6-*O*-(2,3,4,6-tetra-*O*-acetyl)-β-D-galactopyranosyl-β-D-glucopyranoside (73)

Method A (adapted from a known literature procedure)[349]

To a stirring mixture of glycoside **72** (175.1 mg, 0.50 mmol, 1.00 eq.) and fluoride **33a** (153.4 mg, 0.25 mmol, 2.00 eq.) in anh. CH_2Cl_2 (2. mL) was added Et_3N (17.4 µL, 0.13 mmol, 0.50 eq.). Subsequently, $BF_3{\cdot}OEt_2$ (77.1 µL, 0.63 mmol, 2.50 eq.) was added dropwise. The reaction was then stirred at room temperature overnight and monitored by TLC. Once no more of the glucoside (72) was observable, the reaction was diluted with CH_2Cl_2 (20 mL) and washed sequentially with sat. $NaHCO_{3\,(aq)}$ (50 mL) and brine (20 mL). The organic phase was then dried over $MgSO_4$, filtered, and the solvent remove under reduced pressure. The crude mixture was purified using column chromatography (ethyl acetate/petroleum ether 2:8) to obtain **73** in a yield of 25% (60.0 mg, 0.06 mmol).

Method B (adapted from a known literature procedure)[386]

To a stirring mixture of glycoside **71** (0.60 g, 0.70 mmol, 1.00 eq.), fluoride **33a** (0.26 g, 0.80 mmol, 1.10 eq.), and Et_3N (0.11 mL, 0.80 mmol, 1.10 eq.) in anh. CH_2Cl_2 (5 mL) was slowly added $BF_3{\cdot}OEt_2$ (0.19 mL, 1.50 mmol, 2.20 eq.). The reaction was monitored by TLC and was diluted with CH_2Cl_2 (30 mL) after full conversion of **71** was observed. This was then washed successively with sat. $NaHCO_{3\,(aq)}$ (50 mL) and brine (50 mL). The organic phase was then dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. The crude mixture was purified using column chromatography (ethyl acetate/petroleum ether 2:8) to obtain **73** in a yield of 46% (0.30 g, 0.30 mmol).



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.25 (d, $^3J_{3^*,2^*}$ = 9.2 Hz, 2 H, 3*-H), 7.98-7-79 (m, 4 H, arom.-C**H**$_o$, arom.-C**H**$_o$), 7.82 (dd, $^3J_{o\text{-}Bz,m\text{-}Bz}$ = 8.3 Hz, $^4J_{o\text{-}Bz,p\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**$_o$), 7.55 (tt, $^3J_{p\text{-}Bz,m\text{-}Bz}$ = 7.3 Hz, $^4J_{p\text{-}Bz,o\text{-}Bz}$ = 1.3 Hz,

1 H, arom.-C**H**$_p$), 7.52 (tt, $^3J_{p\text{-}Bz,m\text{-}Bz}$ = 7.3 Hz, $^4J_{p\text{-}Bz,o\text{-}Bz}$ = 1.3 Hz, 1 H, arom.-C**H**$_p$), 7.45 (tt, $^3J_{p\text{-}Bz,m\text{-}Bz}$ = 7.4 Hz, $^4J_{p\text{-}Bz,o\text{-}Bz}$ = 1.3 Hz, 1 H, arom.-C**H**$_p$), 7.41 (ddt, $^3J_{m\text{-}Bz,o\text{-}Bz}$ = 9.0 Hz, $^3J_{m\text{-}Bz,p\text{-}Bz}$ = 7.7 Hz, $^4J_{m\text{-}Bz,m\text{-}Bz}$ = 1.5 Hz, 2 H, arom.-C**H**$_m$), 7.37 (ddt, $^3J_{m\text{-}Bz,o\text{-}Bz}$ = 8.8 Hz, $^3J_{m\text{-}Bz,p\text{-}Bz}$ = 7.6 Hz, $^4J_{m\text{-}Bz,m\text{-}Bz}$ = 1.4 Hz, 2 H, arom.-C**H**), 7.30 (ddt, $^3J_{m\text{-}Bz,o\text{-}Bz}$ = 8.8 Hz, $^3J_{m\text{-}Bz,p\text{-}Bz}$ = 7.5 Hz, $^4J_{m\text{-}Bz,m\text{-}Bz}$ = 1.3 Hz, 2 H, arom.-C**H**$_m$), 7.13 (d, $^3J_{2*,3*}$ = 9.3 Hz, 2 H, 2*-H), 5.96 (dd, $^3J_{3,2}$ = 9.6 Hz, $^3J_{3,4}$ = 9.4 Hz, 1 H, 3-H), 5.78 (dd, $^3J_{2,3}$ = 9.8 Hz, $^3J_{2,1}$ = 7.9 Hz, 1 H, 2-H), 5.44 (dd, $^3J_{4,5}$ = 10.1 Hz, $^3J_{4,3}$ = 9.4 Hz, 1 H, 4-H), 5.43 (d, $^3J_{1,2}$ = 7.8 Hz, 1 H, 1-H), 5.38 (dd, $^3J_{4',3'}$ = 3.5 Hz, $^3J_{4',5'}$ = 1.1 Hz, 1 H, 4'-H), 5.26 (dd, $^3J_{2',3'}$ = 10.5 Hz, $^3J_{2',1'}$ = 8.0 Hz, 1 H, 2'-H), 4.98 (dd, $^3J_{3',2'}$ = 10.5 Hz, $^3J_{3',4'}$ = 3.4 Hz, 1 H, 3'-H), 4.52 (d, $^3J_{1',2'}$ = 8.0 Hz, 1 H, 1'-H), 4.30 (ddd, $^3J_{5,4}$ = 10.3 Hz, $^3J_{5,6b}$ = 8.5 Hz, $^3J_{5,6a}$ = 2.0 Hz, 1 H, 5-H), 4.09 (dd, $^2J_{6'a,6'b}$ = 11.4 Hz, $^3J_{6'a,5'}$ = 6.9 Hz, 1 H, 6'-H$_a$), 4.05 (dd, $^2J_{6'b,6'a}$ = 11.4 Hz, $^3J_{6'b,5'}$ = 6.3 Hz, 1 H, 6'-H$_b$), 4.04 (dd, $^2J_{6a,6b}$ = 11.3 Hz, $^3J_{6a,5}$ = 2.2 Hz, 1 H, 6-H$_a$), 3.88 (ddd, $^3J_{5',6'a}$ = 7.0 Hz, $^3J_{5',6'b}$ = 6.2 Hz, $^3J_{5',4'}$ = 1.2 Hz, 1 H, 5'-H), 3.85 (dd, $^2J_{6b,6a}$ = 11.2 Hz, $^3J_{6b,5}$ = 8.5 Hz, 1 H, 6-H$_b$), 2.18 (s, 3 H, 4'-OC(O)C**H**$_3$), 1.99 (s, 3 H, 3'-OC(O)C**H**$_3$), 1.91 (s, 3 H, 6'-OC(O)C**H**$_3$), 1.85 (s, 3 H, 2'-OC(O)C**H**$_3$).

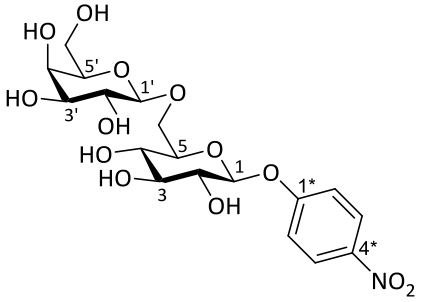
^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 170.4 (6'-O**C**(O)CH$_3$), 170.3 (4'-O**C**(O)CH$_3$), 170.2 (3'-O**C**(O)CH$_3$), 169.4 (2'-O**C**(O)CH$_3$), 165.8 (3-O**C**(O)Ph), 165.5 (4-O**C**(O)Ph), 165.1 (2-O**C**(O)Ph), 161.7 (C-1*), 143.4 (C-4*), 134.0 (arom.-**C**H), 133.7 (arom.-**C**H), 133.6 (arom.-**C**H), 130.0 (arom.-**C**H), 129.9 (arom.-**C**H), 129.9 (arom.-**C**H), 128.9 (arom.-**C**$_{ipso}$), 128.8 (arom.-**C**H), 128.7 (arom.-**C**$_{ipso}$), 128.6 (arom.-**C**H), 128.5 (arom.-**C**$_{ipso}$, arom.-**C**H), 126.2 (C-3*), 117.0 (C-2*), 101.5 (C-1'), 99.0 (C-1), 74.5 (C-5), 72.6 (C-3), 71.5 (C-2), 71.1 (C-5'), 70.9 (C-3'), 69.6 (C-4), 68.7 (C-6), 68.7 (C-2'), 67.2 (C-4'), 61.5 (C-6'), 20.8 (-OC(O)**C**H$_3$), 20.7 (-OC(O)**C**H$_3$), 20.7 (-OC(O)**C**H$_3$), 20.7 (-OC(O)**C**H$_3$).

4-Nitrophenyl β-D-galactopyranosyl-β-D-glucopyranoside (5)

Disaccharide **73** (0.23 g, 0.20 mmol, 1.00 eq.) was reacted as outlined in GPD. The crude mixture was not purified further. Product **5** was obtained in a quantitative yield (130.0 mg).

Data was in agreement with that previously published.[473]



$\mathbf{R_f}$ = 0.64 (ethyl acetate/petroleum ether 4:6, aluminium oxide)

^{1}H NMR (600 MHz, $CDCl_3$):

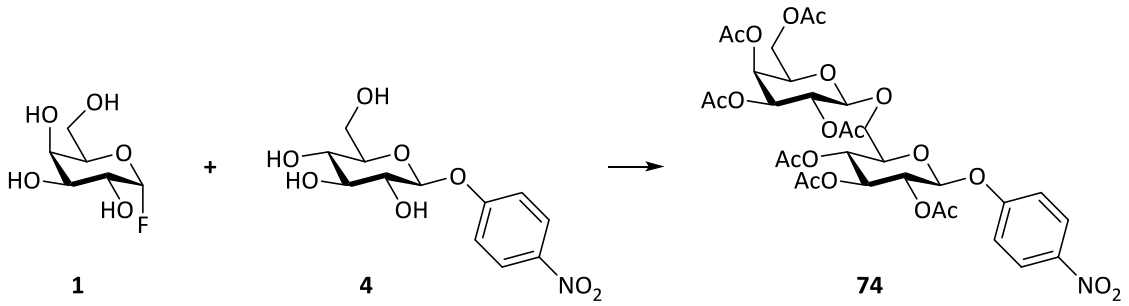
δ (ppm): 8.30 (d, $^3J_{3*,2*}$ = 9.3 Hz, 2 H, 3*-H), 7.31 (d, $^3J_{2*,3*}$ = 9.3 Hz, 2 H, 2*-H), 5.31 (d, $^3J_{1,2}$ = 7.7 Hz, 1 H, 1-H), 4.42 (d, $^3J_{1',2'}$ = 7.7 Hz, 1 H, 1'-H), 4.24 (dd, $^2J_{6a,6b}$ = 14.4 Hz, $^3J_{6a,5}$ = 4.8 Hz, 1 H, 6-H_a), 3.95-3-87 (m, 3 H, 5-H, 6-H_b, 4'-H), 3.76 (dd, $^2J_{6'a,6'b}$ = 11.7 Hz, $^3J_{6'a,5'}$ = 7.8 Hz, 1 H, 6'-H_a), 3.72 (dd, $^2J_{6'b,6'a}$ = 11.7 Hz, $^3J_{6'b,5'}$ = 4.4 Hz, 1 H, 6'H_b), 3.69-3.58 (m, 5 H, 2-H, 3-H, 4-H, 3'H, 5'-H), 3.55 (dd, $^3J_{2',1'}$ = 9.9 Hz, $^3J_{2',3'}$ = 7.7 Hz, 1 H, 2'-H).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 161.6 (C-1*), 142.6 (C-4*), 126.1 (C-3*), 116.6 (C-2*), 103.3 (C-1'), 99.4 (C-1), 75.4 (C-5), 75.3 (C-3), 75.1 (C-5'), 72.7 (C-2), 72.6 (C-3'), 70.7 (C-2'), 69.1 (C-4), 68.6 (C-4'), 68.3 (C-6), 60.9 (C-6').

5.4.7 Enzymatic Synthesis of *para*-Nitrophenyl Allolactose

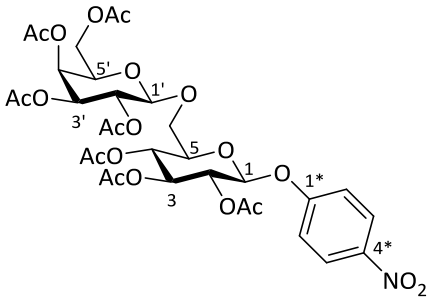
This procedure is adapted from a previously published protocol.[1]



Glycoside **4** (8.7 mg, 29 µmol, 1.0 eq.) and fluoride **1** (11 mg, 58 µmol, 2.0 eq.) were dissolved in buffer (200 mM mixed phosphate, pH 7.5, 1 mM $MgCl_2$; 1 mL) containing galactosidase. The reaction was then shaken at 300 rpm on an orbital shaker for 2 days at 25 °C. The reaction products were then flash frozen using $N_{2\ (l)}$ and the solvent was removed *via* lyophilisation. The carbohydrate/salt mixture was then derivatised according to GPA and the crude product mixture was then purified using pTLC (ethyl acetate/petroleum ether 6:4).

Product was only observed for this reaction using the galactosidase variant *LacZ E537S/G794D* in an isolable yield of 2% (0.3 mg).

Data obtained was in agreement with the literature.[1]



R_f = 0.62 (ethyl acetate/petroleum ether 4:6)

^{1}H NMR (600 MHz, $CDCl_3$):

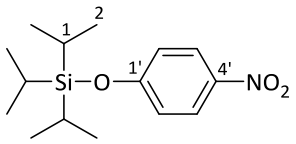
δ (ppm): 8.27 (d, $^3J_{3*,2*}$ = 9.2 Hz, 2 H, 3*-H), 7.10 (d, $^3J_{2*,3*}$ = 9.3 Hz, 2 H, 2*-H), 5.41 (dd, $^3J_{4',3'}$ = 3.5 Hz, $^3J_{4',5'}$ = 1.0 Hz, 1 H, 4'-H), 5.33-5.29 (m, 1 H, 2-H), 5.27 (dd, $^3J_{4,5}$ = 9.7 Hz, $^3J_{4,3}$ = 7.5 Hz, 1 H, 4-H), 5.23 (dd, $^3J_{2',3'}$ = 10.5 Hz, $^3J_{2',1'}$ = 8.0 Hz, 1 H, 2'-H), 5.14 (d, $^3J_{1,2}$ = 7.6 Hz, 1 H, 1-H), 5.00-4.95 (m, 2 H, 3-H, 3'-H), 4.49 (d, $^3J_{1',2'}$ = 8.0 Hz, 1 H, 1'-H), 4.20 (dd, $^2J_{6a,6b}$ = 11.4 Hz, $^3J_{6a,5}$ = 6.7 Hz, 1 H, 6-H_a), 4.11 (dd, $^2J_{6b,6a}$ = 11.5 Hz, $^3J_{6b,5}$ = 6.3 Hz, 1 H, 6-H_b), 3.98 (ddd, $^3J_{5',6'a}$ = 10.2 Hz, $^3J_{5',6'b}$ = 8.1 Hz, $^3J_{5',4'}$ = 2.1 Hz, 1 H, 5'-H), 3.92-3.88 (m, 2 H, 5-H, 6'-H_a), 3.67 (dd, $^2J_{6'b,6'a}$ = 11.0 Hz, $^3J_{6'b,5'}$ = 8.1 Hz, 1 H, 6'-H_b), 2.20 (s, 3 H, -OC(O)C**H**$_3$), 2.08 (s, 3 H, -OC(O)C**H**$_3$), 2.06 (s, 3 H, -OC(O)C**H**$_3$), 2.06 (s, 3 H, -OC(O)C**H**$_3$), 2.04 (s, 3 H, -OC(O)C**H**$_3$), 1.99 (s, 3 H, -OC(O)C**H**$_3$), 1.83 (s, 3 H, -OC(O)C**H**$_3$).

5.4.8 Synthesis of TIPS-pNP

Triisopropyl (4-nitrophenoxy)silane (31)

To a dry flask containing *p*-nitrophenol (0.70 g, 5.00 mmol, 1.00 eq.), Et_3N (1.20 mL, 5.60 mmol, 1.10 eq.), and imidazole (1.00 g, 14.7 mmol, 2.90 eq.), was added anh. CH_2Cl_2 (5 mL). The stirring mixture was cooled to 0 °C and to this was added TIPSCl (0.84 mL, 6.00 mmol, 1.20 eq.) over 5 min using a syringe pump. The reaction was then left to gradually warm to room temperature and reacted for a total of 16 hrs before being quenched with dH_2O (5 mL). The reaction was stirring vigorously for a further 5 min before being diluted with CH_2Cl_2 (20 mL). The phases were then separated and the organic phase was washed successively with dH_2O (20 mL) and brine (20 mL). The organic phase was then dried over $MgSO_4$ and the solvent removed under reduced pressure. The crude mixture was purified using flash column chromatography (diethyl ether/*n*-pentane 2:8) to give **31** in a yield of 95% (1.42 g, 4.8 mmol).

Data was in agreement with that previously published.[474,475]



^{1}H NMR (600 MHz, $CDCl_3$):

δ (ppm): 8.15 (d, $^{3}J_{3',2'}$ = 9.1 Hz, 2 H, 3'-H), 6.93 (d, $^{3}J_{2',3'}$ = 9.1 Hz, 2 H, 2'-H), 1.30 (dq, $^{3}J_{1,2}$ = 7.5 Hz, $^{3}J_{1,2}$ = 7.2 Hz, 3 H, 1-H), 1.11 (d, $^{3}J_{2,1}$ = 7.5 Hz, 18 H, 2-H).

^{13}C NMR (151 MHz, $CDCl_3$):

δ (ppm): 162.2 (C-1'), 141.9 (C-4'), 126.0 (C-3'), 120.1 (C-2'), 18.0 (C-2), 12.8 (C-1).

5.4.9 Synthesis of Other Compounds

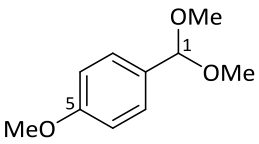
4-Methoxybenzaldehyde dimethylacetal (67)

This method was adapted from a known literature prodcedure.[476]

To a stirring suspension of Dowex 650C (2.50 g) in MeOH (250 mL) was added *p*-methoxybenzaldehyde (62.0 mL, 510 mmol, 1.00 eq.) and trimethylorthoformate (66.0 mL, 603.3 mmol, 1.20 eq.). The mixture was then refluxed under anhydrous conditions for 4 hours. The solids were removed by filtration and the remaining solvent removed under reduced pressure. The title compound was then isolated *via* fractional distillation at 10 mbar and 80 °C. The fractions are given below as a ratio of aldehyde to acetal as determined by ^{1}H NMR.

F1 – 4.27 g (1:13); F2 – 58.3 g (1:19); F3 – 21.0 g (1:44); F4 – 0.07 g (1:75)

Data was in agreement with that previously published.[477]



^{1}H NMR (600 MHz, $CDCl_3$):

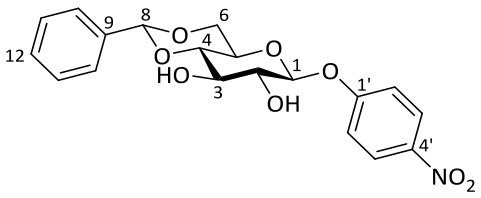
δ (ppm): 7.37 (d, $^3J_{4,3}$ = 8.6 Hz, 2 H, 4-H), 6.90 (d, $^3J_{3,4}$ = 8.7 Hz, 2 H, 3-H), 5.35 (s, 1 H, 1-H), 3.81 (s, 3 H, 5-OC**H**$_3$), 3.31 (s, 6 H, 1-(OC**H**$_3$)$_2$).

4-Nitrophenyl 4,6-benzylidenyl-β-D-galactopyranoside (65)

This method was adapted from a known literature procedure.[478]

Glycoside **4** (220.0 mg, 0.73 mmol, 1.00 eq.), benzaldehyde dimethyl acetal (0.438 mL, 2.90 mmol, 4.00 eq.) and CSA (33.9 mg, 0.15 mmol, 20 mol%) were dissolved in anh. MeCN (1.00 mL). The reaction was then stirred at 30 °C for 4 hrs before being quenched with Et_3N (0.10 mL). The reaction was then filtered and the solid washed with MeCN. The filtrate was collected and the solvent was removed under reduced pressure. The resulting crude product was then purified using column chromatography (ethyl acetate/petroleum ether 6:4) to give **65** in a yield of 87% (246.0 mg, 0.6 mmol).

Data was in agreement with that previously published.[479]



m.p. (°C) = 187.3 – 187.6 °C (lit.[479] 180 – 182 °C)

^{1}H NMR (600 MHz, d_3-MeCN):

δ (ppm): 8.20 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3'-H), 7.51-7.47 (m, 2 H, 11-H), 7.43-7.37 (m, 3 H, 10-H, 12-H), 7.21 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2'-H), 5.59 (s, 1 H, 8-H), 5.18 (d, $^3J_{1,2}$ = 7.7 Hz, 1 H, 1-H), 4.30 (dd, $^2J_{6a,6b}$ = 10.2 Hz, $^3J_{6a,5}$ = 4.9 Hz, 1 H, 6-H$_a$), 3.92 (d, $^3J_{2\text{-}OH,2\text{-}H}$ = 4.4 Hz, 1 H, 2-OH), 3.77-3.70 (m, 3 H, 3-H, 3-OH, 6-H$_b$), 3.67 (ddd, $^3J_{H,H}$ = 9.9 Hz, $^3J_{H,H}$ = 9.7 Hz, $^3J_{5,6a}$ = 4.9 Hz, 1 H, 5-H), 3.61-3.51 (m, 2 H, 2-H, 4-H).

^{13}C NMR (151 MHz, d_3-MeCN):

δ (ppm): 162.9 (C-1‘), 143.8 (C-4‘), 138.7 (C-9), 129.9 (C-11), 129.1 (C-10), 127.2 (C-12), 126.7 (C-3‘), 117.4 (C-2‘), 102.2 (C-8), 101.3 (C-1), 81.0 (C-4), 75.1 (C-2), 73.9 (C-3), 69.0 (C-6), 67.4 (C-5).

HRMS (ESI): 412.10012 [M+Na] ($C_{19}H_{19}NO_8Na$, calc. 412.10029)

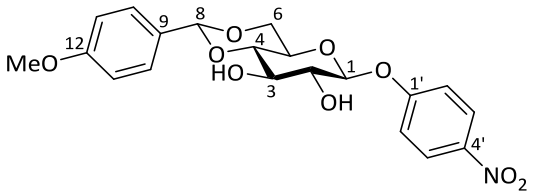
4-Nitrophenyl 4,6-(4-methoxy)benzilidenyl-β-D-galactopyranoside (68)

This method was adapted from a known literature procedure.[478]

Glycoside **4** (1.09 g, 3.62 mmol, 1.00 eq.), *p*-methoxybenzaldehyde dimethyl acetal* (2.55 g, 13.9 mmol, 3.80 eq.) and CSA (0.18 g, 0.79 mmol, 20 mol%) were dissolved in anh. MeCN (35 mL). The reaction was then stirred at 50 °C for 4 hrs before being quenched with Et_3N (2.00 mL). The reaction was then filtered and the solid washed with MeCN. The filtrate was collected and the solvent was removed under reduced pressure. The product was then recrystallised from hot EtOH to give **68** in a yield of 88% (1.33 g, 3.20 mmol).

*dimethyl acetal used had a 1:44 ratio of aldehyde to acetal.

Data was in agreement with that previously published.[480]



m.p. (°C) = 196.2 – 196.7 °C (lit.[480] 190 – 192 °C)

^{1}H NMR (600 MHz, d^3-MeCN):

δ (ppm): 8.20 (d, $^3J_{3',2'}$ = 9.2 Hz, 2 H, 3‘-H), 7.41 (d, $^3J_{10,11}$ = 8.7 Hz, 2 H, 10-H), 7.21 (d, $^3J_{2',3'}$ = 9.2 Hz, 2 H, 2‘-H), 6.93 (d, $^3J_{11,10}$ = 8.7 Hz, 2 H, 11-H), 5.54 (s, 1 H, 8-H), 5.18 (d, $^3J_{1,2}$ = 7.6 Hz, 1 H, 1-H), 4.28 (dd, $^2J_{6a,6b}$ = 10.1 Hz, $^3J_{6a,5}$ = 4.8 Hz, 1 H, 6-H_a), 3.88 (d, $^3J_{2\text{-}OH,2\text{-}H}$ = 4.4 Hz, 1 H, 2-OH), 3.79 (s, 3 H, 12-OC**H**$_3$), 3.75-3.68 (m, 3 H, 3-H, 3-OH, 6-H_b), 3.65 (ddd, $^3J_{H,H}$ = 9.8 Hz, $^3J_{H,H}$ = 9.7 Hz, $^3J_{5,6b}$ = 4.9 Hz, 1 H, 5-H), 3.59-3.50 (m, 2 H, 2-H, 4-H).

^{13}C NMR (151 MHz, d^3-MeCN):

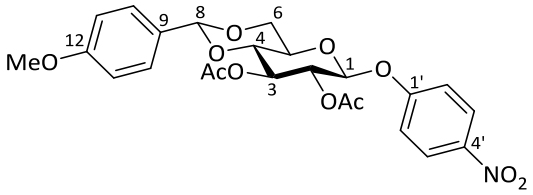
δ (ppm): 163.0 (C-1‘), 161.1 (C-12), 143.8 (C-4‘), 131.1 (C-9), 128.5 (C-10), 126.7 (C-3‘), 117.4 (C-2‘), 114.4 (C-11), 102.2 (C-8), 101.3 (C-1), 81.0 (C-4), 75.1 (C-2), 73.9 (C-3), 68.9 (C-6), 67.5 (C-5), 55.9 (12-O**C**H$_3$).

HRMS (ESI): 420.12864 [M+H] ($C_{20}H_{21}NO_9Na$, calc. 120.39345)

4-Nitrophenyl 2,3-diacetyl-4,6-(4-methoxy)benzylidenyl-β-D-galactopyranoside (69)

Following GPA, glycoside **68** (0.42 g, 1.00 mmol, 1.00 eq.) was reacted. Over the course of the reaction, a white precipitate formed which was collected and washed with Et_2O to reveal compound **69** in a yield of 78% (0.40 g, 0.8 mmol).

Data was in agreement with that previously published.[480]



m.p. (°C) = 247.1 – 248.6 °C (lit.[480] 252 – 253 °C)

^{1}H NMR (600 MHz, *d*^{6}-DMSO):

δ (ppm): 8.23 (d, $^3J_{3',2'}$ = 8.7 Hz, 2 H, 3‘-H), 7.31 (d, $^3J_{10,11}$ = 8.2 Hz, 2 H, 10-H), 7.25 (d, $^3J_{2',3'}$ = 8.8 Hz, 2 H, 2‘-H), 6.93 (d, $^3J_{11,10}$ = 8.2 Hz, 2 H, 11-H), 5.83 (d, $^3J_{1,2}$ = 7.9 Hz, 1 H, 1-H), 5.60 (s, 1 H, 8-H), 5.43 (dd, $^3J_{3,4}$ = 9.5 Hz, $^3J_{3,2}$= 9.5 Hz, 1 H, 3-H), 5.14 (dd, $^3J_{2,3}$ = 8.6 Hz, $^3J_{2,1}$ = 8.6 Hz, 1 H, 2-H), 4.29 (dd, $^2J_{6a,6b}$ = 10.1 Hz, $^3J_{6a,5}$ = 4.9 Hz, 1 H, 6-H$_a$), 4.02 (ddd, $^3J_{5,6b}$ = 9.7 Hz, $^3J_{5,4}$ = 9.6 Hz, $^3J_{5,6a}$ = 4.8 Hz, 1 H, 5-H), 3.95 (dd, $^3J_{4,5}$ = 9.6 Hz, $^3J_{4,3}$ = 9.6 Hz, 1 H, 4-H), 3.82-3.71 (m, 4 H, 6-H$_b$, 12-O**C**H$_3$), 2.02 (s, 6 H, 2-OC(O)**C**H$_3$, 3-OC(O)**C**H$_3$).

^{13}C NMR (151 MHz, *d*^{6}-DMSO):

δ (ppm): 169.6 (3-O**C**(O)CH$_3$), 169.2 (2-O**C**(O)CH$_3$), 161.0 (C-1‘), 159.7 (C-12), 142.5 (C-4‘), 129.5 (C-9), 127.5 (C-10), 125.9 (C-3‘), 116.7 (C-2‘), 113.5 (C-11), 100.5 (C-8), 97.1 (C-1), 77.0 (C-4), 71.3 (C-2), 71.1 (C-3), 67.3 (C-6), 65.8 (C-5), 55.1 (*p*-O**C**H$_3$), 20.5 (-OC(O)**C**H$_3$), 20.3 (-OC(O)**C**H$_3$).

HRMS (ESI): 504.14984 [M+H] ($C_{24}H_{26}NO_{11}$, calc. 504.15004), 542.10578 [M+K] ($C_{24}H_{25}NO_{11}K$, calc. 526.10592)

References

[1] Jakeman, D. L., Withers, S. G. *Can. J. Chem.* **2002**, *80 (8)*, 866-870; 'On expanding the repertoire of glycosynthases: Mutant β-galactosidases forming β-(1,6)-linkages'.

[2] Mathews, C. K., van Holde, K. E., Ahern, K. G. *Biochemistry*; 3rd ed.; Pearson Education, Inc.: Upper Saddle River, **2000**; 'Biochemistry'

[3] Weymouth-Wilson, A. C. *Nat. Prod. Rep.* **1997**, *14 (2)*, 99-110; 'The role of carbohydrates in biologically active natural products'.

[4] Zhang, Y., Wang, F. *Drug Discoveries & Therapeutics* **2015**, *9 (2)*, 79-87; 'Carbohydrate drugs: current status and development prospect'.

[5] Varki, A. *Glycobiology* **1993**, *3 (2)*, 97-130; 'Biological Roles of Oligosaccharides - All of the Theories Are Correct'.

[6] Bertozzi, C. R., Kiessling, L. L. *Science* **2001**, *291 (5512)*, 2357-2364; 'Chemical glycobiology'.

[7] Xu, L., Qi, T., Xu, L., Lu, L., Xiao, M. *J. Carbohydr. Chem.* **2016**, *35 (1)*, 1-23; 'Recent progress in the enzymatic glycosylation of phenolic compounds'.

[8] Arcamone, F., Penco, S., Vigevani, A., Redaelli, S., Franchi, G., Di Marco, A., Casazza, A. M., Dasdia, T., Formelli, F. *J. Med. Chem.* **1975**, *18 (7)*, 703-707; 'Synthesis and antitumor properties of new glycosides of daunomycinone and adriamycinone'.

[9] Lothstein, L., Sweatman, T. W., Priebe, W. *Bioorg. Med. Chem. Lett.* **1995**, *5 (16)*, 1807-1812; 'Hydroxylation at C-3′ of doxorubicin alters the selected phenotype of cellular drug resistance'.

[10] Baer, H. H., Siemsen, L. *Can. J. Chem.* **1988**, *66 (1)*, 187-190; 'Synthesis and biological activity of (S)-2′-fluorodaunorubicin'.

[11] Kaziro, Y., Kamiyama, M. *J. Biochem.* **1967**, *62 (4)*, 424-429; 'Mechanism of Action of Chromomycin A_3 II. Inhibition of RNA Polymerase Reaction'.

[12] Hayasaka, T., Inoue, Y. *Biochemistry* **1969**, *8 (6)*, 2342-2347; 'Chromomycin A_3 studies in aqueous solutions. Spectrophotometric evidence for aggregation and interaction with herring sperm deoxyribonucleic acid'.

[13] Wakisaka, G., Uchino, H., Nakamura, T., Sotobayashi, H., Shirakawa, S., Adachi, A., Sakurai, M. *Nature* **1963**, *198 (4878)*, 385-386; 'Selective Inhibition of Biosynthesis of Ribonucleic Acid in Mammalian Cells by Chromomycin A_3'.

[14] Mrozik, H., Linn, B. O., Eskola, P., Lusi, A., Matzuk, A., Preiser, F. A., Ostlind, D. A., Schaeffer, J. M., Fisher, M. H. *J. Med. Chem.* **1989**, *32 (2)*, 375-381; 'Syntheses and biological activities of 13-substituted avermectin aglycons'.

[15] Nicolaou, K. C., Tsay, S. C., Suzuki, T., Joyce, G. F. *J. Am. Chem. Soc.* **1992**, *114 (19)*, 7555-7557; 'DNA-carbohydrate interactions. Specific binding of the calicheamicin γ_1^I oligosaccharide with duplex DNA'.

[16] Ikemoto, N., Kumar, R. A., Ling, T. T., Ellestad, G. A., Danishefsky, S. J., Patel, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92 (23)*, 10506-10510; 'Calicheamicin-DNA complexes: warhead alignment and saccharide recognition of the minor groove'.

[17] Christner, D. F., Frank, B. L., Kozarich, J. W., Stubbe, J., Golik, J., Doyle, T. W., Rosenberg, I. E., Krishnan, B. *J. Am. Chem. Soc.* **1992**, *114 (23)*, 8763-8767; 'Unmasking the chemistry of

DNA cleavage by the esperamicins: modulation of 4'-hydrogen abstraction and bistranded damage by the fucose-anthranilate moiety'.

[18] Long, B. H., Golik, J., Forenza, S., Ward, B., Rehfuss, R., Dabrowiak, J. C., Catino, J. J., Musial, S. T., Brookshire, K. W., Doyle, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86 (1)*, 2-6; 'Esperamicins, a class of potent antitumor antibiotics: mechanism of action'.

[19] Flynn, E. H., Sigal, M. V., Wiley, P. F., Gerzon, K. *J. Am. Chem. Soc.* **1954**, *76 (12)*, 3121-3131; 'Erythromycin. I. Properties and Degradation Studies'.

[20] Flynn, E. H., Murphy, H. W., McMahon, R. E. *J. Am. Chem. Soc.* **1955**, *77 (11)*, 3104-3106; 'Erythromycin. II. Des-*N*-methylerythromycin and *N*-Methyl-C14-erythromycin'.

[21] Jones, P. H., Rowley, E. K. *J. Org. Chem.* **1968**, *33 (2)*, 665-670; 'Chemical modifications of erythromycin antibiotics. I. 3'-De(dimethylamino)erythromycin A and B'.

[22] Williams, J. D., Sefton, A. M. *J. Antimicrob. Chemother.* **1993**, *31 (suppl_C)*, 11-26; 'Comparison of macrolide antibiotics'.

[23] Ōmura, S., Tsuziki, K., Sunazuka, T., Toyota, H., Takahashi, I., Itoh, Z. *J. Antibiot.* **1985**, *38 (11)*, 1631-1632; 'Gastrointestinal Motor-Stimulating Activity of Macrolide Antibiotics and the Structure-Activity Relationship'.

[24] Ōmura, S., Tsuzuki, K., Sunazuka, T., Marui, S., Toyoda, H., Inatomi, N., Itoh, Z. *J. Med. Chem.* **1987**, *30 (11)*, 1941-1943; 'Macrolides with gastrointestinal motor stimulating activity'.

[25] Lartey, P. A., Nellans, H. N., Faghih, R., Petersen, A., Edwards, C. M., Freiberg, L., Quigley, S., Marsh, K., Klein, L. L., Plattner Jacob, J. *J. Med. Chem.* **1995**, *38 (10)*, 1793-1798; 'Synthesis of 4''-Deoxy Motilides: Identification of a Potent and Orally Active Prokinetic Drug Candidate'.

[26] Hervé, M., Debouzy, J. C., Borowski, E., Cybulska, B., Gary-Bobo, C. M. *Biochim. Biophys. Acta, Biomembr.* **1989**, *980 (3)*, 261-272; 'The role of the carboxyl and amino groups of polyene macrolides in their interactions with sterols and their selective toxicity. A ^{31}P-NMR study'.

[27] Boger, D. L., Honda, T., Menezes, R. F., Colletti, S. L. *J. Am. Chem. Soc.* **1994**, *116 (13)*, 5631-5646; 'Total Synthesis of Bleomycin A_2 and Related Agents. 3. Synthesis and Comparative Evaluation of Deglycobleomycin A_2, Epideglycobleomycin A_2, Deglycobleomycin A_1, and Desacetamido-, Descarboxamido-, Desmethyl-, and Desimidazolyldeglycobleomycin A_2'.

[28] Rohr, J., Thiericke, R. *Nat. Prod. Rep.* **1992**, *9 (2)*, 103-137; 'Angucycline group antibiotics'.

[29] Shi, B., Wu, H., Yu, B., Wu, J. *Angew. Chem.* **2004**, *116 (33)*, 4424-4427; '23-Oxa-Analogues of OSW-1: Efficient Synthesis and Extremely Potent Antitumor Activity'.

[30] Springer, T. A. *Nature* **1990**, *346 (6283)*, 425-434; 'Adhesion receptors of the immune system'.

[31] Hanisch, F.-G., Hanski, C., Hasegawa, A. *Cancer Res.* **1992**, *52 (11)*, 3138-3144; 'Sialyl Lewis^x Antigen as Defined by Monoclonal Antibody AM-3 Is a Marker of Dysplasia in the Colonic Adenoma-Carcinoma Sequence'.

[32] Duncan, R., Kopecková, P., Strohalm, J., Hume, I. C., Lloyd, J. B., Kopecek, J. *Br. J. Cancer* **1988**, *57 (2)*, 147-156; 'Anticancer agents coupled to *N*-(2-hydroxypropyl)methacrylamide copolymers. II. Evaluation of daunomycin conjugates in vivo against L1210 leukaemia'.

[33] Senter, P. D., Saulnier, M. G., Schreiber, G. J., Hirschberg, D. L., Brown, J. P., Hellström, I., Hellström, K. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85 (13)*, 4842-4846; 'Anti-tumor effects of antibody-alkaline phosphatase conjugates in combination with etoposide phosphate'.

[34] Bagshawe, K. D. *Br. J. Cancer* **1987**, *56 (5)*, 531-532; 'Antibody directed enzymes revive anti-cancer prodrugs concept'.

[35] Leenders, R. G. G., Scheeren, H. W., Houba, P. H. J., Boven, E., Haisma, H. J. *Bioorg. Med. Chem. Lett.* **1995**, *5 (24)*, 2975-2980; 'Synthesis and evaluation of novel daunomycin-phosphate-sulfate -β-glucuronide and -β-glucoside prodrugs for application in ADEPT'.

[36] Lundemo, P., Karlsson, E. N., Adlercreutz, P. *Appl. Microbiol. Biotechnol.* **2017**, *101 (3)*, 1121-1131; 'Eliminating hydrolytic activity without affecting the transglycosylation of a GH1 β-glucosidase'.

[37] Clayden, J., Greeves, N., Warren, S., Wothers, P. *Organic Chemistry*; Oxford University Press, Inc.: New York City, **2001**; 'Organic Chemistry'

[38] Angyal, S. J. *Angew. Chem. Int. Ed.* **1969**, *8 (3)*, 157-166; 'The Composition and Conformation of Sugars in Solution'.

[39] Karlson, P., Dixon, H. B. F., Jeannin, Y., Liébeqc, C., Lindberg, B., Loening, K. L., Moss, G. P., Velick, S. F. *Eur. J. Biochem.* **1980**, *111 (2)*, 295-298; 'Conformational Nomenclature for Five and Six-Membered Ring Forms of Monosaccharides and Their Derivatives'.

[40] Michael, A. *Am. Chem. J* **1879**, *1*, 305-312; 'On the synthesis of helicin and phenolglucoside'.

[41] Fischer, E. *Ber. Dtsch. Chem. Ges.* **1893**, *26 (3)*, 2400-2412; 'Ueber die Glucoside der Alkohole'.

[42] Fischer, E., Beensch, L. *Ber. Dtsch. Chem. Ges.* **1894**, *27 (2)*, 2478-2486; 'Ueber einige synthetische Glucoside'.

[43] Koenigs, W., Knorr, E. *Ber. Dtsch. Chem. Ges.* **1901**, *34 (1)*, 957-981; 'Ueber einige Derivate des Traubenzuckers und der Galactose'.

[44] Fischer, E., Armstrong, E. F. *Ber. Dtsch. Chem. Ges.* **1901**, *34 (2)*, 2885-2900; 'Ueber die isomeren Acetohalogen-Derivate des Traubenzuckers und die Synthese der Glucoside'.

[45] Noyori, R., Kurimoto, I. *J. Org. Chem.* **1986**, *51 (22)*, 4320-4322; 'Electrochemical glycosylation method'.

[46] Davis, B. G., Chambers, D., Cumpstey, I., France, R., Gamblin, D. In *Carbohydrates*; Academic Press: Oxford, **2003**, p 69-120; 'Synthesis and Activation of Carbohydrate Donors: Acetates, Halides, Phenyl selenides and Glycals'.

[47] Paulsen, H. *Angew. Chem. Int. Ed. Engl.* **1982**, *21 (3)*, 155-173; 'Advances in Selective Chemical Syntheses of Complex Oligosaccharides'.

[48] Mootoo, D. R., Konradsson, P., Udodong, U., Fraser-Reid, B. *J. Am. Chem. Soc.* **1988**, *110 (16)*, 5583-5584; 'Armed and disarmed n-pentenyl glycosides in saccharide couplings leading to oligosaccharides'.

[49] Fraser-Reid, B., Konradsson, P., Mootoo, D. R., Udodong, U. *J. Chem. Soc., Chem. Commun.* **1988**, *0 (12)*, 823-825; 'Direct elaboration of pent-4-enyl glycosides into disaccharides'.

[50] Pietruszka, J. In *Carbohydrates*; Academic Press: Oxford, **2003**, p 195-218; 'Modern Glycosidation Methods: Tuning of Reactivity'.

[51] Limousin, C., Cléophax, J., Petit, A., Loupy, A., Lukacs, G. *J. Carbohydr. Chem.* **1997**, *16 (3)*, 327-342; 'Solvent-Free Synthesis of Decyl D-Glycopyranosides Under Focused Microwave Irradiation'.

[52] Franchimont, A. P. N. *Ber. Dtsch. Chem. Ges.* **1879**, *12 (2)*, 1938-1942; 'Ueber Kohlehydrate'.

[53] Schmidt, R. R., Michel, J. *J. Carbohydr. Chem.* **1985**, *4 (2)*, 141-169; '*O*-(α-D-Glucopyranosyl)trichloroacetimidate as a Glucosyl Donor'.

[54] Wegmann, B., Schmidt, R. R. *J. Carbohydr. Chem.* **1987**, *6 (3)*, 357-375; 'The Application of the Trichloroacetimidate Method to the Synthesis of α-D-Gluco- and α-D-Galactopyranosides'.

[55] Rathore, H., Hashimoto, T., Igarashi, K., Nukaya, H., Fullerton, D. S. *Tetrahedron* **1985**, *41 (23)*, 5427-5438; 'Cardiac glycosides: 5. Stereoselective syntheses of digitoxigenin α-D, β-D, α-L, and β-L-glucosides'.

[56] Rathore, H., From, A. H. L., Ahmed, K., Fullerton, D. S. *J. Med. Chem.* **1986**, *29 (10)*, 1945-1952; 'Cardiac glycosides. 7. Sugar stereochemistry and cardiac glycoside activity'.

[57] Vankar, Y. D., Vankar, P. S., Behrendt, M., Schmidt, R. R. *Tetrahedron* **1991**, *47 (48)*, 9985-9992; 'Synthesis of β-*O*-glycosides using enol ether and imidate derived leaving groups. Emphasis on the use of nitriles as a solvent'.

[58] Schmidt, R. R., Behrendt, M., Toepfer, A. *Synlett* **1990**, *1990 (11)*, 694-696; 'Nitriles as Solvents in Glycosylation Reactions: Highly Selective β-Glycoside Synthesis'.

[59] Kartha, K. P. R., Field, R. A. In *Carbohydrates*; Academic Press: Oxford, **2003**, p 121-145; 'Synthesis and Activation of Carbohydrate Donors: Thioglycosides and Sulfoxides'.

[60] Lönn, H. *Carbohydr. Res.* **1985**, *139*, 105-113; 'Synthesis of a tri- and a hepta-saccharide which contain α-L-fucopyranosyl groups and are part of the complex type of carbohydrate moiety of glycoproteins'.

[61] Lönn, H. *Carbohydr. Res.* **1985**, *139*, 115-121; 'Synthesis of a tetra- and a nona-saccharide which contain α-L-fucopyranosyl groups and are part of the complex type of carbohydrate moiety of glycoproteins'.

[62] Shimizu, M., Togo, H., Yokoyama, M. *Synthesis* **1998**, *1998 (06)*, 799-822; 'Chemistry of Glycosyl Fluorides'.

[63] Nicolaou, K. C., Mitchell, H. J., Suzuki, H., Rodríguez, R. M., Baudoin, O., Fylaktakidou, K. C. *Angew. Chem. Int. Ed.* **1999**, *38 (22)*, 3334-3339; 'Total Synthesis of Everninomicin 13,384-1—Part 1: Synthesis of the $A_1B(A)C$ Fragment'.

[64] Nicolaou, K. C., Rodríguez, R. M., Fylaktakidou, K. C., Suzuki, H., Mitchell, H. J. *Angew. Chem. Int. Ed.* **1999**, *38 (22)*, 3340-3345; 'Total Synthesis of Everninomicin 13,384-1—Part 2: Synthesis of the $FGHA_2$ Fragment'.

[65] Nicolaou, K. C., Mitchell, H. J., Rodríguez, R. M., Fylaktakidou, K. C., Suzuki, H. *Angew. Chem. Int. Ed.* **1999**, *38 (22)*, 3345-3350; 'Total Synthesis of Everninomicin 13,384-1—Part 3: Synthesis of the DE Fragment and Completion of the Total Synthesis'.

[66] Grice, P., Ley, S. V., Pietruszka, J., Osborn, H. M. I., Priepke, H. W. M., Warriner, S. L. *Chem. Eur. J.* **1997**, *3 (3)*, 431-440; 'A New Strategy for Oligosaccharide Assembly Exploiting Cyclohexane-1,2-diacetal Methodology: An Efficient Synthesis of a High Mannose Type Nonasaccharide'.

[67] Behrend, R., Roth, P. *Justus Liebigs Ann. Chem.* **1904**, *331 (3)*, 359-382; 'Ueber die Birotation der Glucose'.

[68] Furneaux, R. H., Rendle, P. M., Sims, I. M. *J. Chem. Soc., Perkin Trans. 1* **2000**, *(13)*, 2011-2014; 'The influence of boric acid on the acetylation of aldoses: 'one-pot' syntheses of penta-O-acetyl-β-D-glucofuranose and its crystalline propanoyl analogue'.

[69] Zacharia, J. T., Hayashi, M. *Carbohydr. Res.* **2012**, *348*, 91-94; 'Facile synthesis of acacetin-7-*O*-β-D-galactopyranoside'.

[70] Balcerzak, A. K., Ferreira, S. S., Trant, J. F., Ben, R. N. *Bioorg. Med. Chem. Lett.* **2012**, *22 (4)*, 1719-1721; 'Structurally diverse disaccharide analogs of antifreeze glycoproteins and their ability to inhibit ice recrystallization'.

[71] Fischer, E. *Ber. Dtsch. Chem. Ges.* **1895**, *28 (1)*, 1145-1167; 'Ueber die Verbindungen der Zucker mit den Alkoholen und Ketonen'.

[72] Uchiyama, T., Hindsgaul, O. *Synlett* **1996**, *1996 (06)*, 499-501; 'Per-*O*-Trimethylsilyl-α-L-Fucopyranosyl Iodide: A Novel Glycosylating Agent for Terminal α-L-Fucosylation'.

[73] Mukaiyama, T., Murai, Y., Shoda, S.-i. *Chem. Lett.* **1981**, *10 (3)*, 431-432; 'An efficient method for glycosylation of hydroxy compounds using glycopyranosyl fluoride'.

[74] Rosenbrook, W., Riley, D. A., Lartey, P. A. *Tetrahedron Lett.* **1985**, *26 (1)*, 3-4; 'A new method for the synthesis of glycosyl fluorides'.

[75] Gordon, D. M., Danishefsky, S. J. *Carbohydr. Res.* **1990**, *206 (2)*, 361-366; 'Displacement reactions of a 1,2-anhydro-α-D-hexopyranose: installation of useful functionality at the anomeric carbon'.

[76] Suzuki, K., Maeta, H., Matsumoto, T. *Tetrahedron Lett.* **1989**, *30 (36)*, 4853-4856; 'An improved procedure for metallocene-promoted glycosidation. Enhanced reactivity by employing 1:2-ratio of Cp_2HfCl_2-$AgClO_4$'.

[77] Nicolaou, K. C., Chucholowski, A., Dolle, R. E., Randall, J. L. *J. Chem. Soc., Chem. Commun.* **1984***(17)*, 1155-1156; 'Reactions of glycosyl fluorides. Synthesis of *O*-, *S*-, and *N*-glycosides'.

[78] Kreuzer, M., Thiem, J. *Carbohydr. Res.* **1986**, *149 (2)*, 347-361; 'Aufbau von Oligosacchariden mit Glycosylfluoriden unter Lewissäure-Katalyse'.

[79] Hideki, J., Hiroki, M., Teruaki, M. *Chem. Lett.* **2001**, *30 (5)*, 426-427; 'A Catalytic and Stereoselective Glycosylation with Glucopyranosyl Fluoride by Using Various Protic Acids'.

[80] Hashimoto, S., Hayashi, M., Noyori, R. *Tetrahedron Lett.* **1984**, *25 (13)*, 1379-1382; 'Glycosylation using glucopyranosyl fluorides and silicon-based catalysts. Solvent dependency of the stereoselection'.

[81] Toshima, K., Tatsuta, K. *Chem. Rev.* **1993**, *93 (4)*, 1503-1531; 'Recent progress in *O*-glycosylation methods and its application to natural products synthesis'.

[82] Fischer, E., Delbrück, K. *Ber. Dtsch. Chem. Ges.* **1909**, *42 (2)*, 1476-1482; 'Über Thiophenol-glucoside'.

[83] Schneider, W., Sepp, J., Stiehler, O. *Ber. Dtsch. Chem. Ges.* **1918**, *51 (1)*, 220-234; 'Synthese zweier isomerer Reihen von Alkyl-thioglucosiden'.

[84] Purves, C. B. *J. Am. Chem. Soc.* **1929**, *51 (12)*, 3619-3627; 'Relations between rotatory power and structure in the sugar group. XXI. β-Thiophenol glycosides of glucose, xylose, lactose and cellobiose'.

[85] Veeneman, G. H., van Leeuwen, S. H., van Boom, J. H. *Tetrahedron Lett.* **1990**, *31 (9)*, 1331-1334; 'Iodonium ion promoted reactions at the anomeric centre. II An efficient thioglycoside mediated approach toward the formation of 1,2-*trans* linked glycosides and glycosidic esters'.

[86] Kartha, K. P. R., Field, R. A. *Tetrahedron Lett.* **1997**, *38 (47)*, 8233-8236; 'Glycosylation chemistry promoted by iodine monobromide: Efficient synthesis of glycosyl bromides from thioglycosides, and *O*-glycosides from 'disarmed' thioglycosides and glycosyl bromides'.

[87] Nicolaou, K. C., Dolle, R. E., Papahatjis, D. P., Randall, J. L. *J. Am. Chem. Soc.* **1984**, *106 (15)*, 4189-4192; 'Practical synthesis of oligosaccharides. Partial synthesis of avermectin B_{1a}'.

[88] Caddick, S., Gazzard, L., Motherwell, W. B., Wilkinson, J. A. *Tetrahedron* **1996**, *52 (1)*, 149-156; 'Preparation of 1-fluoroglycosides from 1-arylthio and 1-arylselenoglycosides using 4-methyl(difluoroiodo)benzene'.

[89] Kakarla, R., Dulina, R. G., Hatzenbuhler, N. T., Hui, Y. W., Sofia, M. J. *J. Org. Chem.* **1996**, *61 (23)*, 8347-8349; 'Simple and Efficient Method for the Oxidation of Sulfides to Sulfoxides: Application to the Preparation of Glycosyl Sulfoxides'.

[90] Crich, D., Sun, S. *Tetrahedron* **1998**, *54 (29)*, 8321-8348; 'Direct chemical synthesis of β-mannopyranosides and other glycosides via glycosyl triflates'.

[91] Kahne, D., Walker, S., Cheng, Y., Van Engen, D. *J. Am. Chem. Soc.* **1989**, *111 (17)*, 6881-6882; 'Glycosylation of unreactive substrates'.

[92] Yang, D., Kim, S. H., Kahne, D. *J. Am. Chem. Soc.* **1991**, *113 (12)*, 4715-4716; 'Construction of glycosidic N-O linkages in oligosaccharides'.

[93] Kim, S. H., Augeri, D., Yang, D., Kahne, D. *J. Am. Chem. Soc.* **1994**, *116 (5)*, 1766-1775; 'Concise Synthesis of the Calicheamicin Oligosaccharide Using the Sulfoxide Glycosylation Method'.

[94] Raghavan, S., Kahne, D. *J. Am. Chem. Soc.* **1993**, *115 (4)*, 1580-1581; 'A one step synthesis of the ciclamycin trisaccharide'.

[95] Alonso, I., Khiar, N., Martín-Lomas, M. *Tetrahedron Lett.* **1996**, *37 (9)*, 1477-1480; 'A new promoter system for the sulfoxide glycosylation reaction'.

[96] Chanteloup, L., Beau, J.-M. *Tetrahedron Lett.* **1992**, *33 (37)*, 5347-5350; 'Phenylsulfenyl-D-ribofuranosides as efficient ribosyl donors: Application to the synthesis of [1′,-^{13}C]-(deoxy)nucleosides'.

[97] Crich, D., Sun, S. *J. Am. Chem. Soc.* **1998**, *120 (2)*, 435-436; 'Direct Formation of β-Mannopyranosides and Other Hindered Glycosides from Thioglycosides'.

[98] Kartha, K. P. R., Aloui, M., Field, R. A. *Tetrahedron Lett.* **1996**, *37 (48)*, 8807-8810; 'Iodine: A versatile reagent in carbohydrate chemistry III. Efficient activation of glycosyl halides in combination with DDQ'.

[99] Schneider, W., Wrede, F. *Ber. Dtsch. Chem. Ges.* **1917**, *50 (1)*, 793-804; 'Synthese eines schwefelhaltigen und eines selenhaltigen Disaccharides'.

[100] Mehta, S., Mario Pinto, B. *Tetrahedron Lett.* **1991**, *32 (35)*, 4435-4438; 'Phenylselenoglycosides as novel, versatile glycosyl donors. Selective activation over thioglycosides'.

[101] Mehta, S., Pinto, B. M. *J. Org. Chem.* **1993**, *58 (12)*, 3269-3276; 'Novel glycosidation methodology. The use of phenyl selenoglycosides as glycosyl donors and acceptors in oligosaccharide synthesis'.

[102] Aloui, M., Chambers, D. J., Cumpstey, I., Fairbanks, A. J., Redgrave, A. J., Seward, C. M. P. *Chem. Eur. J.* **2002**, *8 (11)*, 2608-2621; 'Stereoselective 1,2-*cis* Glycosylation of 2-*O*-Allyl Protected Thioglycosides'.

[103] Pougny, J. R., Jacquinet, J. C., Nassr, M., Duchet, D., Milat, M. L., Sinay, P. *J. Am. Chem. Soc.* **1977**, *99 (20)*, 6762-6763; 'A novel synthesis of 1,2-*cis*-disaccharides'.

[104] Schmidt, R. R., Michel, J. *Angew. Chem. Int. Ed. Engl.* **1980**, *19 (9)*, 731-732; 'Facile Synthesis of α- and β-*O*-Glycosyl Imidates; Preparation of Glycosides and Disaccharides'.

[105] Schmidt, R. R., Grundler, G. *Angew. Chem. Int. Ed. Engl.* **1982**, *21 (10)*, 781-782; 'α-Linked Disaccharides from *O*-(β-D-Glycopyranosyl) Trichloroacetimidates using Trimethylsilyl Trifluoromethanesulfonate as Catalyst'.

[106] Dobarro-Rodriguez, A., Trumtel, M., Wessel, H. P. *J. Carbohydr. Chem.* **1992**, *11 (3)*, 255-263; 'Triflic Anhydride: An Alternative Promoter in Glycosidations'.

[107] Douglas, S. P., Whitfield, D. M., Krepinsky, J. J. *J. Carbohydr. Chem.* **1993**, *12 (1)*, 131-136; 'Silver Trifluoromethanesulfonate(Triflate) Activation of Trichloroacetimidates in Glycosylation Reactions'.

[108] Wang, Z.-G., Douglas, S. P., Krepinsky, J. J. *Tetrahedron Lett.* **1996**, *37 (39)*, 6985-6988; 'Polymer-supported syntheses of oligosaccharides: Using dibutylboron triflate to promote glycosylations with glycosyl trichloroacetimidates'.

[109] Urban, F. J., Moore, B. S., Breitenbach, R. *Tetrahedron Lett.* **1990**, *31 (31)*, 4421-4424; 'Synthesis of tigogenyl β-*O*-cellobioside heptaacetate and glycoside tetraacetate via Schmidt's trichloroacetimidate method; some new observatons'.

[110] Nicolaou, K. C., Daines, R. A., Ogawa, Y., Chakraborty, T. K. *J. Am. Chem. Soc.* **1988**, *110 (14)*, 4696-4705; 'Total synthesis of amphotericin B. 3. The final stages'.

[111] Fairbanks, A. J., Seward, C. M. P. In *Carbohydrates*; Academic Press: Oxford, **2003**, p 147-194; 'Synthesis and Activation of Carbohydrate Donors: Acetimidates, n-Pentenyl and Vinyl Glycosides'.

[112] Schmidt, R. R., Michel, J., Roos, M. *Liebigs Ann. Chem.* **1984**, *1984 (7)*, 1343-1357; 'Glycosylimidate, 12 Direkte Synthese von *O*-α- und *O*-β-Glycosyl-imidaten'.

[113] Schmidt, R. R., Michel, J. *Tetrahedron Lett.* **1984**, *25 (8)*, 821-824; 'Direct *O*-glycosyl trichloroacetimidate formation, nucleophilicity of the anomeric oxygen atom'.

[114] Schmidt, R. R., Stumpp, M. *Liebigs Ann. Chem.* **1983**, *1983 (7)*, 1249-1256; 'Glycosylimidate, 8. Synthese von 1-Thioglycosiden'.

[115] Numata, M., Sugimoto, M., Koike, K., Ogawa, T. *Carbohydr. Res.* **1987**, *163 (2)*, 209-225; 'Total synthesis of sialosylcerebroside, GM4'.

[116] Halcomb, R. L., Danishefsky, S. J. *J. Am. Chem. Soc.* **1989**, *111 (17)*, 6661-6666; 'On the direct epoxidation of glycals: application of a reiterative strategy for the synthesis of β-linked oligosaccharides'.

[117] Gervay, J., Danishefsky, S. *J. Org. Chem.* **1991**, *56 (18)*, 5448-5451; 'A stereospecific route to 2-deoxy-β-glycosides'.

[118] Gallant, M., Link, J. T., Danishefsky, S. J. *J. Org. Chem.* **1993**, *58 (2)*, 343-349; 'A stereoselective synthesis of indole-β-*N*-glycosides: an application to the synthesis of rebeccamycin'.

[119] Timmers, C. M., van der Marel, G. A., van Boom, J. H. *Chem. Eur. J.* **1995**, *1 (3)*, 161-164; 'Synthesis of a Methyl Heptaglucoside with Phytoalexin Elicitor Activity Based on Oxidative Coupling of Glucals'.

[120] Danishefsky, S. J., Bilodeau, M. T. *Angew. Chem. Int. Ed. Engl.* **1996**, *35 (13-14)*, 1380-1419; 'Glycals in Organic Synthesis: The Evolution of Comprehensive Strategies for the Assembly of Oligosaccharides and Glycoconjugates of Biological Consequence'.

[121] Dushin, R. G., Danishefsky, S. J. *J. Am. Chem. Soc.* **1992**, *114 (9)*, 3471-3475; 'Stereospecific synthesis of aryl β-glucosides: an application to the synthesis of a prototype corresponding to the aryloxy carbohydrate domain of vancomycin'.

[122] Liu, K. K. C., Danishefsky, S. J. *J. Org. Chem.* **1994**, *59 (7)*, 1895-1897; 'A Direct Route from 1α,2α-Anhydroglucose Derivatives to α-Glucosides'.

[123] Sinnott, M. L. *Chem. Rev.* **1990**, *90 (7)*, 1171-1202; 'Catalytic Mechanisms of Enzymatic Glycosyl Transfer'.

[124] Campbell, J. A., Davies, G. J., Bulone, V., Henrissat, B. *Biochem. J.* **1997**, *326*, 929-939; 'A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities'.

[125] Zechel, D. L., Withers, S. G. *Acc. Chem. Res.* **2000**, *33 (1)*, 11-18; 'Glycosidase mechanisms: Anatomy of a finely tuned catalyst'.

[126] Lairson, L. L., Henrissat, B., Davies, G. J., Withers, S. G. *Annu. Rev. Biochem.* **2008**, *77*, 521-555; 'Glycosyltransferases: Structures, functions, and mechanisms'.

[127] Wolfenden, R., Lu, X. D., Young, G. *J. Am. Chem. Soc.* **1998**, *120 (27)*, 6814-6815; 'Spontaneous hydrolysis of glycosides'.

[128] Thompson, J. E., Kutateladze, T. G., Schuster, M. C., Venegas, F. D., Messmore, J. M., Raines, R. T. *Bioorg. Chem.* **1995**, *23 (4)*, 471-481; 'Limits to Catalysis by Ribonuclease A'.

[129] Rye, C. S., Withers, S. G. *Curr. Opin. Chem. Biol.* **2000**, *4 (5)*, 573-580; 'Glycosidase mechanisms'.

[130] Lairson, L. L., Watts, A. G., Wakarchuk, W. W., Withers, S. G. *Nat. Chem. Biol.* **2006**, *2 (12)*, 724-728; 'Using substrate engineering to harness enzymatic promiscuity and expand biological catalysis'.

[131] Rich, J. R., Szpacenko, A., Palcic, M. M., Bundle, D. R. *Angew. Chem. Int. Ed.* **2004**, *43 (5)*, 613-615; 'Glycosyltransferase-Catalyzed Synthesis of Thiooligosaccharides'.

[132] Ünligil, U. M., Rini, J. M. *Curr. Opin. Struct. Biol.* **2000**, *10 (5)*, 510-517; 'Glycosyltransferase structure and mechanism'.

[133] Coutinho, P. M., Deleury, E., Davies, G. J., Henrissat, B. *J. Mol. Biol.* **2003**, *328 (2)*, 307-317; 'An evolving hierarchical family classification for glycosyltransferases'.

[134] Lee, S. S., Hong, S. Y., Errey, J. C., Izumi, A., Davies, G. J., Davis, B. G. *Nat. Chem. Biol.* **2011**, *7 (9)*, 631-638; 'Mechanistic evidence for a front-side, S_Ni-type reaction in a retaining glycosyltransferase'.

[135] Jaeken, J., Matthijs, G. *Annu. Rev. Genomics Hum. Genet.* **2007**, *8 (1)*, 261-278; 'Congenital Disorders of Glycosylation: A Rapidly Expanding Disease Family'.

[136] Leroy, J. G. *Pediatr. Res.* **2006**, *60 (6)*, 643-656; 'Congenital Disorders of *N*-Glycosylation Including Diseases Associated With *O*- as Well as *N*-Glycosylation Defects'.

[137] Freeze, H. H. *Glycobiology* **2001**, *11 (12)*, 129R-143R; 'Update and perspectives on congenital disorders of glycosylation'.

[138] Rech, C., Rosencrantz, R. R., Krenek, K., Pelantová, H., Bojarová, P., Römer, C. E., Hanisch, F.-G., Kren, V., Elling, L. *Adv. Synth. Catal.* **2011**, *353 (13)*, 2492-2500; 'Combinatorial One-Pot Synthesis of Poly-*N*-acetyllactosamine Oligosaccharides with Leloir-Glycosyltransferases'.

[139] Wong, C. H., Haynie, S. L., Whitesides, G. M. *J. Org. Chem.* **1982**, *47 (27)*, 5416-5418; 'Enzyme-catalyzed synthesis of N-acetyllactosamine with in situ regeneration of uridine 5'-diphosphate glucose and uridine 5'-diphosphate galactose'.

[140] Crout, D. H. G., Vic, G. *Curr. Opin. Chem. Biol.* **1998**, *2 (1)*, 98-111; 'Glycosidases and glycosyl transferases in glycoside and oligosaccharide synthesis'.

[141] Okamoto, K., Goto, T. *Tetrahedron* **1990**, *46 (17)*, 5835-5857; 'Glycosidation of sialic acid'.

[142] Liu, K. K. C., Danishefsky, S. J. *Chem. Eur. J.* **1996**, *2 (11)*, 1359-1362; 'A Paradigm Case for the Merging of Glycal and Enzymatic Assembly Methods in Glycoconjugate Synthesis: A Highly Efficient Chemo-Enzymatic Synthesis of GM3'.

[143] Unverzagt, C. *Angew. Chem. Int. Ed. Engl.* **1996**, *35 (20)*, 2350-2353; 'Chemoenzymatic Synthesis of a Sialylated Undecasaccharide–Asparagine Conjugate'.

[144] Palcic, M. M. *Methods Enzymol.* **1994**, *230*, 300-316; 'Glycosyltransferases in glycobiology'.

[145] Engels, L., Elling, L. *Glycobiology* **2014**, *24 (2)*, 170-178; 'WbgL: a novel bacterial α-1,2-fucosyltransferase for the synthesis of 2′-fucosyllactose'.

[146] Ichikawa, Y., Lin, Y. C., Dumas, D. P., Shen, G. J., Garcia-Junceda, E., Williams, M. A., Bayer, R., Ketcham, C., Walker, L. E., Paulson, J. C., Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114 (24)*, 9283-9298; 'Chemical-enzymic synthesis and conformational analysis of sialyl Lewis X and derivatives'.

[147] Baisch, G., Öhrlein, R. *Angew. Chem. Int. Ed. Engl.* **1996**, *35 (16)*, 1812-1815; 'Chemoenzymatic Synthesis of Sialyl Lewis^{x} Glycopeptides'.

[148] Hara, A., Imamura, A., Ando, H., Ishida, H., Kiso, M. *Molecules* **2014**, *19 (1)*, 414; 'A New Chemical Approach to Human ABO Histo-Blood Group Type 2 Antigens'.

[149] Terwisscha van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrissat, B., Dijkstra, B. W. *Biochemistry* **1995**, *34 (48)*, 15619-15623; 'Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and x-ray structure of a complex with allosamidin evidence for substrate assisted catalysis'.

[150] Synstad, B., Gåseidnes, S., van Aalten, D. M. F., Vriend, G., Nielsen, J. E., Eijsink, V. G. H. *Eur. J. Biochem.* **2004**, *271 (2)*, 253-262; 'Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase'.

[151] van Aalten, D. M. F., Komander, D., Synstad, B., Gåseidnes, S., Peter, M. G., Eijsink, V. G. H. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98 (16)*, 8979-8984; 'Structural insights into the catalytic mechanism of a family 18 *exo*-chitinase'.

[152] Williams, S. J., Mark, B. L., Vocadlo, D. J., James, M. N. G., Withers, S. G. *J. Biol. Chem.* **2002**, *277 (42)*, 40055-40065; 'Aspartate 313 in the *Streptomyces plicatus* Hexosaminidase Plays a Critical Role in Substrate-assisted Catalysis by Orienting the 2-Acetamido Group and Stabilizing the Transition State'.

[153] Hash, J. H., Rothlauf, M. V. *J. Biol. Chem.* **1967**, *242 (23)*, 5586-5590; 'The *N*,*O*-Diacetylmuramidase of *Chalaropsis* Species : I. Purification and Crystallization'.

[154] Vollmer, W. *FEMS Microbiol. Rev.* **2008**, *32 (2)*, 287-306; 'Structural variation in the glycan strands of bacterial peptidoglycan'.

[155] Martinez-Fleites, C., Korczynska, J. E., Davies, G. J., Cope, M. J., Turkenburg, J. P., Taylor, E. J. *Carbohydr. Res.* **2009**, *344 (13)*, 1753-1757; 'The crystal structure of a family GH25 lysozyme from *Bacillus anthracis* implies a neighboring-group catalytic mechanism with retention of anomeric configuration'.

[156] An, S., Chen, L., Wei, J.-F., Yang, X., Ma, D., Xu, X., Xu, X., He, S., Lu, J., Lai, R. *PLoS One* **2012**, *7 (2)*, e31920; 'Purification and Characterization of Two New Allergens from the Venom of *Vespa magnifica*'.

[157] Biner, O., Trachsel, C., Moser, A., Kopp, L., Langenegger, N., Kämpfer, U., von Ballmoos, C., Nentwig, W., Schürch, S., Schaller, J., Kuhn-Nentwig, L. *PLoS One* **2015**, *10 (12)*, e0143963; 'Isolation, *N*-glycosylations and Function of a Hyaluronidase-Like Enzyme from the Venom of the Spider *Cupiennius salei*'.

[158] Umekawa, M., Huang, W., Li, B., Fujita, K., Ashida, H., Wang, L.-X., Yamamoto, K. *J. Biol. Chem.* **2008**, *283 (8)*, 4469-4479; 'Mutants of Mucor hiemalis Endo-β-*N*-acetylglucosaminidase Show Enhanced Transglycosylation and Glycosynthase-like Activities'.

[159] Fujita, K., Sato, R., Toma, K., Kitahara, K., Suganuma, T., Yamamoto, K., Takegawa, K. *J. Biochem.* **2007**, *142 (3)*, 301-306; 'Identification of the Catalytic Acid–Base Residue of *Arthrobacter* Endo-β-*N*-Acetylglucosaminidase by Chemical Rescue of an Inactive Mutant'.

[160] Li, B., Song, H., Hauser, S., Wang, L.-X. *Org. Lett.* **2006**, *8 (14)*, 3081-3084; 'A Highly Efficient Chemoenzymatic Approach toward Glycoprotein Synthesis'.

[161] Hamilton, S. R., Davidson, R. C., Sethuraman, N., Nett, J. H., Jiang, Y., Rios, S., Bobrowicz, P., Stadheim, T. A., Li, H., Choi, B.-K., Hopkins, D., Wischnewski, H., Roser, J., Mitchell, T., Strawbridge, R. R., Hoopes, J., Wildt, S., Gerngross, T. U. *Science* **2006**, *313 (5792)*, 1441-1443; 'Humanization of Yeast to Produce Complex Terminally Sialylated Glycoproteins'.

[162] Burmeister, W. P., Cottaz, S., Rollin, P., Vasella, A., Henrissat, B. *J. Biol. Chem.* **2000**, *275 (50)*, 39385-39393; 'High Resolution X-ray Crystallography Shows That Ascorbate Is a Cofactor for Myrosinase and Substitutes for the Function of the Catalytic Base'.

[163] Bones, A. M., Rossiter, J. T. *Physiol. Plant.* **1996**, *97 (1)*, 194-208; 'The myrosinase-glucosinolate system, its organisation and biochemistry'.

[164] Björkman, R., Lönnerdal, B. *Biochim. Biophys. Acta, Enzymol.* **1973**, *327 (1)*, 121-131; 'Studies on myrosinases III. Enzymatic properties of myrosinases from *Sinapis alba* and *Brassica napus* seeds'.

[165] Burstein, C., Kepes, A. *Biochim. Biophys. Acta, Gen. Subj.* **1971**, *230 (1)*, 52-63; 'The α-galactosidase from *Escherichia coli* K12'.

[166] Lodge, J. A., Maier, T., Liebl, W., Hoffmann, V., Sträter, N. *J. Biol. Chem.* **2003**, *278 (21)*, 19151-19158; 'Crystal Structure of Thermotoga maritima α-Glucosidase AglA Defines a New Clan of NAD+-dependent Glycosidases'.

[167] Raasch, C., Streit, W., Schanzer, J., Bibel, M., Gosslar, U., Liebl, W. *Extremophiles* **2000**, *4 (4)*, 189-200; '*Thermotoga maritima* AglA, an extremely thermostable NAD^+-, Mn^{2+}-, and thiol-dependent α-glucosidase'.

[168] Rajan, S. S., Yang, X., Collart, F., Yip, V. L. Y., Withers, S. G., Varrot, A., Thompson, J., Davies, G. J., Anderson, W. F. *Structure* **2004**, *12 (9)*, 1619-1629; 'Novel Catalytic Mechanism of Glycoside Hydrolysis Based on the Structure of an NAD^+/Mn^{2+}-Dependent Phospho-α-Glucosidase from *Bacillus subtilis*'.

[169] Yip, V. L. Y., Varrot, A., Davies, G. J., Rajan, S. S., Yang, X., Thompson, J., Anderson, W. F., Withers, S. G. *J. Am. Chem. Soc.* **2004**, *126 (27)*, 8354-8355; 'An Unusual Mechanism of Glycoside Hydrolysis Involving Redox and Elimination Steps by a Family 4 β-Glycosidase from *Thermotoga maritima*'.

[170] Kim, S., Oh, D.-B., Kang, H. A., Kwon, O. *Appl. Microbiol. Biotechnol.* **2011**, *91 (1)*, 1-15; 'Features and applications of bacterial sialidases'.

[171] Buschiazzo, A., Alzari, P. M. *Curr. Opin. Chem. Biol.* **2008**, *12 (5)*, 565-572; 'Structural insights into sialic acid enzymology'.

[172] Watson, J. N., Dookhun, V., Borgford, T. J., Bennet, A. J. *Biochemistry* **2003**, *42 (43)*, 12682-12690; 'Mutagenesis of the Conserved Active-Site Tyrosine Changes a Retaining Sialidase into an Inverting Sialidase'.

[173] Amaya, M. F., Watts, A. G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A., Paris, G., Frasch, A. C., Withers, S. G., Alzari, P. M. *Structure* **2004**, *12 (5)*, 775-784; 'Structural Insights into the Catalytic Mechanism of *Trypanosoma cruzi trans*-Sialidase'.

[174] Watts, A. G., Damager, I., Amaya, M. L., Buschiazzo, A., Alzari, P., Frasch, A. C., Withers, S. G. *J. Am. Chem. Soc.* **2003**, *125 (25)*, 7532-7533; '*Trypanosoma cruzi* Trans-sialidase Operates through a Covalent Sialyl–Enzyme Intermediate: Tyrosine Is the Catalytic Nucleophile'.

[175] Withers, S. G. *Carbohydr. Polym.* **2001**, *44 (4)*, 325-337; 'Mechanisms of glycosyl transferases and hydrolases'.

[176] Davies, G., Henrissat, B. *Structure* **1995**, *3 (9)*, 853-859; 'Structures and Mechanisms of Glycosyl Hydrolases'.

[177] Koshland, D. E., Jr. *Biol. Rev. Camb. Philos. Soc.* **1953**, *28 (4)*, 416-436; 'Stereochemistry and the Mechanism of Enzymatic Reactions'.

[178] Nilsson, K. G. I. *Carbohydr. Res.* **1987**, *167*, 95-103; 'A simple strategy for changing the regioselectivity of glycosidase-catalysed formation of disaccharides'.

[179] Adlercreutz, P. *Appl. Microbiol. Biotechnol.* **2017**, *101 (2)*, 513-519; 'Comparison of lipases and glycoside hydrolases as catalysts in synthesis reactions'.

[180] Gupta, M. N. *Eur. J. Biochem.* **1992**, *203 (1-2)*, 25-32; 'Enzyme function in organic solvents'.

[181] Klibanov, A. M. *Nature* **2001**, *409 (6817)*, 241-246; 'Improving enzymes by using them in organic solvents'.

[182] Vic, G., Thomas, D. *Tetrahedron Lett.* **1992**, *33 (32)*, 4567-4570; 'Enzyme-catalysed synthesis of alkyl β-D-Glucosides in organic media'.

[183] Vic, G., Crout, D. H. G. *Tetrahedron: Asymmetry* **1994**, *5 (12)*, 2513-2516; 'Synthesis of glucosidic derivatives with a spacer arm by reverse hydrolysis using almond β-D-glucosidase'.

[184] Vic, G., Biton, J., Le Beller, D., Michel, J.-M., Thomas, D. *Biotechnol. Bioeng.* **1995**, *46 (2)*, 109-116; 'Enzymatic glucosylation of hydrophobic alcohols in organic medium by the reverse hydrolysis reaction using almond-β-D-glucosidase'.

[185] Vic, G., Crout, D. H. G. *Carbohydr. Res.* **1995**, *279*, 315-319; 'Synthesis of allyl and benzyl β-D-glucopyranosides, and allyl β-D-galactopyranoside from D-glucose or D-galactose and the corresponding alcohol using almond β-D-glucosidase'.

[186] Ljunger, G., Adlercreutz, P., Mattiasson, B. *Enzyme Microb. Technol.* **1994**, *16 (9)*, 751-755; 'Enzymatic synthesis of octyl-β-glucoside in octanol at controlled water activity'.

[187] Hansson, T., Adlercreutz, P. *Biotechnol. Bioeng.* **2001**, *75 (6)*, 656-665; 'Enhanced transglucosylation/hydrolysis ratio of mutants of *Pyrococcus furiosus* β-glucosidase: Effects of donor concentration, water content, and temperature on activity and selectivity in hexanol'.

[188] Teze, D., Hendrickx, J., Dion, M., Tellier, C., Woods, V. L., Tran, V., Sanejouand, Y.-H. *Biochemistry* **2013**, *52 (34)*, 5900-5910; 'Conserved Water Molecules in Family 1 Glycosidases: A DXMS and Molecular Dynamics Study'.

[189] Lundemo, P., Karlsson, E. N., Adlercreutz, P. *J. Mol. Catal. B: Enzym.* **2014**, *108*, 1-6; 'Preparation of two glycoside hydrolases for use in micro-aqueous media'.

[190] Feng, H.-Y., Drone, J., Hoffmann, L., Tran, V., Tellier, C., Rabiller, C., Dion, M. *J. Biol. Chem.* **2005**, *280 (44)*, 37088-37097; 'Converting a β-Glycosidase into a β-Transglycosidase by Directed Evolution'.

[191] Teze, D., Hendrickx, J., Czjzek, M., Ropartz, D., Sanejouand, Y.-H., Tran, V., Tellier, C., Dion, M. *Protein Eng. Des. Sel.* **2014**, *27 (1)*, 13-19; 'Semi-rational approach for converting a GH1 β-glycosidase into a β-transglycosidase'.

[192] Teze, D., Daligault, F., Ferrières, V., Sanejouand, Y.-H., Tellier, C. *Glycobiology* **2015**, *25 (4)*, 420-427; 'Semi-rational approach for converting a GH36 α-glycosidase into an α-transglycosidase'.

[193] Lundemo, P., Adlercreutz, P., Karlsson, E. N. *Appl. Environ. Microbiol.* **2013**, *79 (11)*, 3400-3405; 'Improved Transferase/Hydrolase Ratio through Rational Design of a Family 1 β-Glucosidase from *Thermotoga neapolitana*'.

[194] Bissaro, B., Monsan, P., Fauré, R., O'Donohue, Michael J. *Biochem. J.* **2015**, *467 (1)*, 17-35; 'Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases'.

[195] Shoda, S.-i., Uyama, H., Kadokawa, J.-i., Kimura, S., Kobayashi, S. *Chem. Rev.* **2016**, *116 (4)*, 2307-2413; 'Enzymes as Green Catalysts for Precision Macromolecular Synthesis'.

[196] Withers, S. G., Rupitz, K., Trimbur, D., Warren, R. A. J. *Biochemistry* **1992**, *31 (41)*, 9979-9985; 'Mechanistic Consequences of Mutation of the Active-Site Nucleophile Glu 358 in *Agrobacterium* β-Glucosidase'.

[197] Bojarová, P., Kren, V. *Chimia* **2011**, *65 (1-2)*, 65-70; 'Glycosidases in Carbohydrate Synthesis: When Organic Chemistry Falls Short'.

[198] Mackenzie, L. F., Wang, Q. P., Warren, R. A. J., Withers, S. G. *J. Am. Chem. Soc.* **1998**, *120 (22)*, 5583-5584; 'Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis'.

[199] Jahn, M., Chen, H. M., Müllegger, J., Marles, J., Warren, R. A. J., Withers, S. G. *Chem. Commun.* **2004***(3)*, 274-275; 'Thioglycosynthases: double mutant glycosidases that serve as scaffolds for thioglycoside synthesis'.

[200] Goddard-Borger, E. D., Tysoe, C., Withers, S. G. *Carbohydr. Res.* **2016**, *435*, 97-99; 'Glycosynthase mediated synthesis of psychosine'.

[201] Henze, M., Schmidtke, S., Hoffmann, N., Steffens, H., Pietruszka, J., Elling, L. *ChemCatChem* **2015**, *7 (19)*, 3131-3139; 'Combination of Glycosyltransferases and a Glycosynthase in Sequential and One-Pot Reactions for the Synthesis of Type 1 and Type 2 *N*-Acetyllactosamine Oligomers'.

[202] Malet, C., Planas, A. *FEBS Lett.* **1998**, *440 (1–2)*, 208-212; 'From β-glucanase to β-glucansynthase: glycosyl transfer to α-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile'.

[203] Okuyama, M., Mori, H., Watanabe, K., Kimura, A., Chiba, S. *Biosci. Biotechnol. Biochem.* **2002**, *66 (4)*, 928-933; 'α-glucosidase mutant catalyzes "α-glycosynthase"-type reaction'.

[204] Cobucci-Ponzano, B., Conte, F., Bedini, E., Corsaro, M. M., Parrilli, M., Sulzenbacher, G., Lipski, A., Dal Piaz, F., Lepore, L., Rossi, M., Moracci, M. *Chem. Biol.* **2009**, *16 (10)*, 1097-1108; 'β-Glycosyl Azides as Substrates for β-Glycosynthases: Preparation of Efficient α-L-Fucosynthases'.

[205] Cobucci-Ponzano, B., Zorzetti, C., Strazzulli, A., Carillo, S., Bedini, E., Corsaro, M. M., Comfort, D. A., Kelly, R. M., Rossi, M., Moracci, M. *Glycobiology* **2011**, *21 (4)*, 448-456; 'A novel α-D-galactosynthase from *Thermotoga maritima* converts β-D-galactopyranosyl azide to α-galacto-oligosaccharides'.

[206] Moracci, M., Trincone, A., Perugino, G., Ciaramella, M., Rossi, M. *Biochemistry* **1998**, *37 (49)*, 17262-17270; 'Restoration of the Activity of Active-Site Mutants of the Hyperthermophilic β-Glycosidase from *Sulfolobus solfataricus*: Dependence of the Mechanism on the Action of External Nucleophiles'.

[207] Pozzo, T., Plaza, M., Romero-García, J., Faijes, M., Karlsson, E. N., Planas, A. *J. Mol. Catal. B: Enzym.* **2014**, *107*, 132-139; 'Glycosynthases from *Thermotoga neapolitana* β-glucosidase 1A: A comparison of α-glucosyl fluoride and *in situ*-generated α-glycosyl formate donors'.

[208] Cheng, L. L., Shidmoossavee, F. S., Bennet, A. J. *Biochemistry* **2014**, *53 (24)*, 3982-3989; 'Neuraminidase Substrate Promiscuity Permits a Mutant *Micromonospora viridifaciens* Enzyme To Synthesize Artificial Carbohydrates'.

[209] Park, I., Lee, H., Cha, J. *Biotechnol. Lett.* **2014**, *36 (4)*, 789-796; 'Glycoconjugates synthesized *via* transglycosylation by a thermostable α-glucosidase from *Thermoplasma acidophilum* and its glycosynthase mutant'.

[210] Danby, P. M., Withers, S. G. *ACS Chem. Biol.* **2016**, *11 (7)*, 1784-1794; 'Advances in Enzymatic Glycoside Synthesis'.

[211] Sakurama, H., Fushinobu, S., Hidaka, M., Yoshida, E., Honda, Y., Ashida, H., Kitaoka, M., Kumagai, H., Yamamoto, K., Katayama, T. *J. Biol. Chem.* **2012**, *287 (20)*, 16709-16719; '1,3-1,4-α-L-Fucosynthase That Specifically Introduces Lewis a/x Antigens into Type-1/2 Chains'.

[212] Wada, J., Honda, Y., Nagae, M., Kato, R., Wakatsuki, S., Katayama, T., Taniguchi, H., Kumagai, H., Kitaoka, M., Yamamoto, K. *FEBS Lett.* **2008**, *582 (27)*, 3739-3743; '1,2-α-L-Fucosynthase: A glycosynthase derived from an inverting α-glycosidase with an unusual reaction mechanism'.

[213] Kwan, D. H., Ernst, S., Kötzler, M. P., Withers, S. G. *Glycobiology* **2015**, *25 (8)*, 806-811; 'Chemoenzymatic Synthesis of a Type 2 Blood Group A Tetrasaccharide and Development of High-throughput Assays Enables a Platform for Screening Blood Group Antigen-cleaving Enzymes'.

[214] Zhou, M., Hou, Y., Hamza, A., Zhan, C.-G., Bugni, T. S., Thorson, J. S. *Org. Lett.* **2012**, *14 (21)*, 5424-5427; 'Probing the Regiospecificity of Enzyme-Catalyzed Steroid Glycosylation'.

[215] Fort, S., Boyer, V., Greffe, L., Davies, G. J., Moroz, O., Christiansen, L., Schülein, M., Cottaz, S., Driguez, H. *J. Am. Chem. Soc.* **2000**, *122 (23)*, 5429-5437; 'Highly Efficient Synthesis of β-(1→4)-Oligo- and -Polysaccharides Using a Mutant Cellulase'.

[216] Codera, V., Edgar, K. J., Faijes, M., Planas, A. *Biomacromolecules* **2016**, *17 (4)*, 1272-1279; 'Functionalized Celluloses with Regular Substitution Pattern by Glycosynthase-Catalyzed Polymerization'.

[217] Li, C., Kim, Y.-W. *ChemBioChem* **2014**, *15 (4)*, 522-526; 'Characterization of a Galactosynthase Derived from *Bacillus circulans* β-Galactosidase: Facile Synthesis of D-Lacto- and D-Galacto-*N*-bioside'.

[218] Yang, M., Davies, G. J., Davis, B. G. *Angew. Chem. Intl. Ed.* **2007**, *46 (21)*, 3885-3888; 'A glycosynthase catalyst for the synthesis of flavonoid glycosides'.

[219] Rich, J. R., Cunningham, A. M., Gilbert, M., Withers, S. G. *Chem. Commun.* **2011**, *47*, 10806-10808; 'Glycosphingolipid synthesis employing a combination of recombinant glycosyltransferases and an endoglycoceramidase glycosynthase'.

[220] Wilkinson, S. M., Watson, M. A., Willis, A. C., McLeod, M. D. *J. Org. Chem.* **2011**, *76 (7)*, 1992-2000; 'Experimental and Kinetic Studies of the *Escherichia coli* Glucuronylsynthase: An Engineered Enzyme for the Synthesis of Glucuronide Conjugates'.

[221] Cobucci-Ponzano, B., Strazzulli, A., Rossi, M., Moracci, M. *Adv. Synth. Catal.* **2011**, *353 (13)*, 2284-2300; 'Glycosynthases in Biocatalysis'.

[222] Urrutia, P., Rodriguez-Colinas, B., Fernandez-Arrojo, L., Ballesteros, A. O., Wilson, L., Illanes, A., Plou, F. J. *J. Agric. Food Chem.* **2013**, *61 (5)*, 1081-1087; 'Detailed Analysis of Galactooligosaccharides Synthesis with β-Galactosidase from *Aspergillus oryzae*'.

[223] Rodriguez-Colinas, B., de Abreu, M. A., Fernandez-Arrojo, L., de Beer, R., Poveda, A., Jimenez-Barbero, J., Haltrich, D., Olmo, A. O. B., Fernandez-Lobato, M., Plou, F. J. *J. Agric. Food Chem.* **2011**, *59 (19)*, 10477-10484; 'Production of Galacto-oligosaccharides by the β-Galactosidase from *Kluyveromyces lactis*: Comparative Analysis of Permeabilized Cells versus Soluble Enzyme'.

[224] Rather, M. Y., Mishra, S., Aravinda, S. *J. Biotechnol.* **2013**, *165 (1)*, 63-68; 'Exploring the synthetic potential of cell bound β-glycosidase of *Pichia etchellsii*'.

[225] Watanabe, A., Suzuki, M., Ujiie, S., Gomi, K. *J. Biosci. Bioeng.* **2016**, *121 (3)*, 259-264; 'Purification and enzymatic characterization of a novel β-1,6-glucosidase from *Aspergillus oryzae*'.

[226] Cohn, M. *Bacteriol. Rev.* **1957**, *21 (3)*, 140-168; 'Contributions of studies on the β-galactosidase of *Escherichia coli* to our understanding of enzyme synthesis'.

[227] Jacob, F., Monod, J. *J. Mol. Biol.* **1961**, *3 (3)*, 318-356; 'Genetic regulatory mechanisms in the synthesis of proteins'.

[228] Wutor, V. C., Togo, C. A., Pletschke, B. I. *Chemosphere* **2007**, *68 (4)*, 622-627; 'The effect of physico-chemical parameters and chemical compounds on the activity of β-D-galactosidase (B-GAL), a marker enzyme for indicator microorganisms in water'.

[229] Burstein, C., Cohn, M., Kepes, A., Monod, J. *Biochim. Biophys. Acta, Nucl. Acids Prot. Synth.* **1965**, *95 (4)*, 634-639; 'Rôle du lactose et de ses produits métaboliques dans l'induction de l'opéron lactose chez *Escherichia coli*'.

[230] Müller-Hill, B., Rickenberg, H. V., Wallenfels, K. *J. Mol. Biol.* **1964**, *10 (2)*, 303-318; 'Specificity of the induction of the enzymes of the lac operon in *Escherichia coli*'.

[231] Marchesi, S. L., Steers, E., Jr., Shifrin, S. *Biochim. Biophys. Acta* **1969**, *181 (1)*, 20-34; 'Purification and Characterization of Multiple Forms of β-Galactosidase of *Escherichia coli*'.

[232] Juers, D. H., Jacobson, R. H., Wigley, D., Zhang, X. J., Huber, R. E., Tronrud, D. E., Matthews, B. W. *Protein Sci.* **2000**, *9 (9)*, 1685-1699; 'High resolution refinement of β-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for α-complementation'.

[233] Jacobson, R. H., Zhang, X.-J., DuBose, R. F., Matthews, B. W. *Nature* **1994**, *369 (6483)*, 761-766; 'Three-dimensional structure of β-galactosidase from *E. coli*'.

[234] Juers, D. H., Matthews, B. W., Huber, R. E. *Protein Sci.* **2012**, *21 (12)*, 1792-1807; 'LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance'.

[235] Pilipenko, O. S., Atyaksheva, L. F., Poltorak, O. M., Chukhrai, E. S. *Russ. J. Phys. Chem. A* **2007**, *81 (6)*, 990-994; 'Dissociation and catalytic activity of oligomer forms of β-galactosidases'.

[236] Welply, J. K., Fowler, A. V., Beckwith, J. R., Zabin, I. *J. Bacteriol.* **1980**, *142 (2)*, 732-734; 'Positions of early nonsense and deletion mutations in *lacZ*'.

[237] Langley, K. E., Villarejo, M. R., Fowler, A. V., Zamenhof, P. J., Zabin, I. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72 (4)*, 1254-1257; 'Molecular basis of β-galactosidase α-complementation'.

[238] Ullmann, A., Jacob, F., Monod, J. *J. Mol. Biol.* **1967**, *24 (2)*, 339-343; 'Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of *Escherichia coli*'.

[239] Ullmann, A. *BioEssays* **1992**, *14 (3)*, 201-205; 'Complementation in β-Galactosidase - from Protein-Structure to Genetic-Engineering'.

[240] Langley, K. E., Fowler, A. V., Zabin, I. *J. Biol. Chem.* **1975**, *250 (7)*, 2587-2592; 'Amino acid sequence of β-galactosidase. IV. Sequence of an α-complementing cyanogen bromide peptide, residues 3 to 92'.

[241] Welply, J. K., Fowler, A. V., Zabin, I. *J. Biol. Chem.* **1981**, *256 (13)*, 6804-6810; 'β-Galactosidase α-complementation. Overlapping sequences'.

[242] Welply, J. K., Fowler, A. V., Zabin, I. *J. Biol. Chem.* **1981**, *256 (13)*, 6811-6816; 'β-Galactosidase α-complementation. Effect of single amino acid substitutions'.

[243] Fowler, A. V., Zabin, I. *J. Biol. Chem.* **1983**, *258 (23)*, 14354-14358; 'Purification, structure, and properties of hybrid β-galactosidase proteins'.

[244] Gallagher, C. N., Huber, R. E. *Biochemistry* **1997**, *36 (6)*, 1281-1286; 'Monomer−Dimer Equilibrium of Uncomplemented M15 β-Galactosidase from *Escherichia coli*'.

[245] Gallagher, C. N., Huber, R. E. *J. Protein Chem.* **1998**, *17 (2)*, 131-141; 'Studies of the M15 β-Galactosidase Complementation Process'.

[246] Gebler, J. C., Aebersold, R., Withers, S. G. *J. Biol. Chem.* **1992**, *267 (16)*, 11126-11130; 'Glu-537, not Glu-461, is the nucleophile in the active site of (*lacZ*) β-galactosidase from *Escherichia coli*'.

[247] Richard, J. P., Huber, R. E., Lin, S., Heo, C., Amyes, T. L. *Biochemistry* **1996**, *35 (38)*, 12377-12386; 'Structure−Reactivity Relationships for β-Galactosidase (*Escherichia coli, lac Z*). 3. Evidence that Glu-461 Participates in Brønsted Acid−Base Catalysis of β-D-Galactopyranosyl Group Transfer'.

[248] McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., Withers, S. G. *Biochemistry* **1996**, *35 (31)*, 9958-9966; 'The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: A ^{13}C-NMR study of *Bacillus circuluns* xylanase'.

[249] Richard, J. P. *Biochemistry* **1998**, *37 (13)*, 4305-4309; 'The enhancement of enzymatic rate accelerations by Brønsted acid-base catalysis'.

[250] Juers, D. H., Heightman, T. D., Vasella, A., McCarter, J. D., Mackenzie, L., Withers, S. G., Matthews, B. W. *Biochemistry* **2001**, *40 (49)*, 14781-14794; 'A Structural View of the Action of *Escherichia coli* (*lacZ*) β-Galactosidase'.

[251] Martinez-Bilbao, M., Holdsworth, R. E., Edwards, L. A., Huber, R. E. *J. Biol. Chem.* **1991**, *266 (8)*, 4979-4986; 'A Highly Reactive β-Galactosidase (*Escherichia coli*) Resulting from a Substitution of an Aspartic-Acid for Gly-794'.

[252] Sutendra, G., Wong, S., Fraser, M. E., Huber, R. E. *Biochem. Biophys. Res. Commun.* **2007**, *352 (2)*, 566-570; 'β-Galactosidase (*Escherichia coli*) has a second catalytically important Mg^{2+} site'.

[253] Jancewicz, L. J., Wheatley, R. W., Sutendra, G., Lee, M., Fraser, M. E., Huber, R. E. *Arch. Biochem. Biophys.* **2012**, *517 (2)*, 111-122; 'Ser-796 of β-galactosidase (*Escherichia coli*) plays a key role in maintaining a balance between the opened and closed conformations of the catalytically important active site loop'.

[254] Dugdale, M. L., Dymianiw, D. L., Minhas, B. K., D'Angelo, I., Huber, R. E. *Biochem. Cell Biol.* **2010**, *88 (5)*, 861-869; 'Role of Met-542 as a guide for the conformational changes of Phe-601 that occur during the reaction of β-galactosidase (*Escherichia coli*)'.

[255] Dugdale, M. L., Vance, M. L., Wheatley, R. W., Driedger, M. R., Nibber, A., Tran, A., Huber, R. E. *Biochem. Cell Biol.* **2010**, *88 (6)*, 969-979; 'Importance of Arg-599 of β-galactosidase (*Escherichia coli*) as an anchor for the open conformations of Phe-601 and the active-site loop'.

[256] Juers, D. H., Hakda, S., Matthews, B. W., Huber, R. E. *Biochemistry* **2003**, *42 (46)*, 13505-13511; 'Structural basis for the altered activity of Gly794 variants of *Escherichia coli* β-galactosidase'.

[257] Huber, R. E., Hlede, I. Y., Roth, N. J., McKenzie, K. C., Ghumman, K. K. *Biochem. Cell Biol.* **2001**, *79 (2)*, 183-193; 'His-391 of β-galactosidase (*Escherichia coli*) promotes catalyes by strong interactions with the transition state'.

[258] Espinosa, J. F., Montero, E., Vian, A., García, J. L., Dietrich, H., Schmidt, R. R., Martín-Lomas, M., Imberty, A., Cañada, F. J., Jiménez-Barbero, J. *J. Am. Chem. Soc.* **1998**, *120 (6)*, 1309-1318; '*Escherichia coli* β-galactosidase recognizes a high-energy conformation of *C*-lactose, a nonhydrolizable substrate analogue. NMR and modeling studies of the molecular complex'.

[259] Wheatley, R. W., Kappelhoff, J. C., Hahn, J. N., Dugdale, M. L., Dutkoski, M. J., Tamman, S. D., Fraser, M. E., Huber, R. E. *Arch. Biochem. Biophys.* **2012**, *521 (1-2)*, 51-61; 'Substitution for Asn460 Cripples β-galactosidase (*Escherichia coli*) by increasing substrate affinity and decreasing transition state stability'.

[260] Roth, N. J., Huber, R. E. *J. Biol. Chem.* **1996**, *271 (24)*, 14296-14301; 'The β-galactosidase (*Escherichia coli*) reaction is partly facilitated by interactions of His-540 with the C6 hydroxyl of galactose'.

[261] Huber, R. E., Hakda, S., Cheng, C., Cupples, C. G., Edwards, R. A. *Biochemistry* **2003**, *42 (6)*, 1796-1803; 'Trp-999 of β-galactosidase (*Escherichia coli*) is a key residue for binding, catalysis, and synthesis of allolactose, the natural *lac* operon inducer'.

[262] Roth, N. J., Rob, B., Huber, R. E. *Biochemistry* **1998**, *37 (28)*, 10099-10107; 'His-357 of β-galactosidase (*Escherichia coli*) interacts with the C3 hydroxyl in the transition state and helps to mediate catalysis'.

[263] Xu, J., McRae, M. A. A., Harron, S., Rob, B., Huber, R. E. *Biochem. Cell Biol.* **2004**, *82 (2)*, 275-284; 'A study of the relationships of interactions between Asp-201, Na^+ or K^+, and galactosyl C6 hydroxyl and their effects on binding and reactivity of β-galactosidase'.

[264] Lo, S., Dugdale, M. L., Jeerh, N., Ku, T., Roth, N. J., Huber, R. E. *Protein J.* **2010**, *29 (1)*, 26-31; 'Studies of Glu-416 Variants of β-Galactosidase (*E. coli*) Show that the Active Site Mg^{2+} is Not Important for Structure and Indicate that the Main Role of Mg^{2+} is to Mediate Optimization of Active Site Chemistry'.

[265] Juers, D. H., Rob, B., Dugdale, M. L., Rahimzadeh, N., Giang, C., Lee, M., Matthews, B. W., Huber, R. E. *Protein Sci.* **2009**, *18 (6)*, 1281-1292; 'Direct and indirect roles of His-418 in metal binding and in the activity of β-galactosidase (*E. coli*)'.

[266] Ring, M., Bader, D. E., Huber, R. E. *Biochem. Biophys. Res. Commun.* **1988**, *152 (3)*, 1050-1055; 'Site-Directed Mutagenesis of β-Galactosidase (*Escherichia coli*) Reveals That Tyr-503 Is Essential for Activity'.

[267] Ring, M., Huber, R. E. *Arch. Biochem. Biophys.* **1990**, *283 (2)*, 342-350; 'Multiple Replacements Establish the Importance of Tyrosine-503 in β-Galactosidase (*Escherichia coli*)'.

[268] Penner, R. M., Roth, N. J., Rob, B., Lay, H., Huber, R. E. *Biochem. Cell Biol.* **1999**, *77 (3)*, 229-236; 'Tyr-503 of β-galactosidase (*Escherichia coli*) plays an important role in degalactosylation'.

[269] Wheatley, R. W., Lo, S., Jancewicz, L. J., Dugdale, M. L., Huber, R. E. *J. Biol. Chem.* **2013**, *288 (18)*, 12993-13005; 'Structural Explanation for Allolactose (*lac* Operon Inducer) Synthesis by *lacZ* β-Galactosidase and the Evolutionary Relationship between Allolactose Synthesis and the *lac* Repressor'.

[270] Huber, R. E., Kurz, G., Wallenfels, K. *Biochemistry* **1976**, *15 (9)*, 1994-2001; 'A quantitation of the factors which affect the hydrolase and transgalactosylase activities of β-galactosidase (*E. coli*) on lactose'.

[271] Craig, D. B., Haslam, A. M., Coombs, J. M. L., Nichols, E. R. *Biochem. Cell Biol.* **2010**, *88 (3)*, 451-458; 'Kinetic studies of unmodified individual *Escherichia coli* β-galactosidase molecules in free solution'.

[272] Nichols, E. R., Gavina, J. M. A., McLeod, R. G., Craig, D. B. *Protein J.* **2007**, *26 (2)*, 95-105; 'Single molecule assays of β-galactosidase from two wild-type strains of *E. coli*: Effects of protease inhibitors on microheterogeneity and different relative activities with differing substrates'.

[273] Huber, R. E., Parfett, C., Woulfe-Flanagan, H., Thompson, D. J. *Biochemistry* **1979**, *18 (19)*, 4090-4095; 'Interaction of divalent cations with β-galactosidase (*Escherichia coli*)'.

[274] Selwood, T., Sinnott, M. L. *Biochem. J.* **1990**, *268 (2)*, 317-323; 'A solvent-isotope-effect study of proton transfer during catalysis by *Escherichia coli* (*lacZ*) β-galactosidase'.

[275] Roth, N. J., Huber, R. E. *Biochem. Biophys. Res. Commun.* **1996**, *219 (1)*, 111-115; 'Glu-416 of β-galactosidase (*Escherichia coli*) is a Mg^{2+} ligand and beta-galactosidases with substitutions for Glu-416 are inactivated, rather than activated, by Mg^{2+}'.

[276] Rickenberg, H. V. *Biochim. Biophys. Acta* **1959**, *35*, 122-129; 'The effect of metal ions and proteins on the stability of the β-galactosidase of *Escherichia coli*'.

[277] Wallenfels, K., Weil, R. In *The Enzymes*; Paul, D. B., Ed.; Academic Press: **1972**; Vol. 7, p 617-663; '20 β-Galactosidase'.

[278] Brás, N. F., Fernandes, P. A., Ramos, M. J. *J. Chem. Theory Comput.* **2010**, *6 (2)*, 421-433; 'QM/MM Studies on the β-Galactosidase Catalytic Mechanism: Hydrolysis and Transglycosylation Reactions'.

[279] Deslongchamps, P. *Pure Appl. Chem.* **1993**, *65 (6)*, 1161-1178; 'Intramolecular Strategies and Stereoelectronic Effects - Glycosides Hydrolysis Revisited'.

[280] Withers, S. G., Rupitz, K., Street, I. P. *J. Biol. Chem.* **1988**, *263 (17)*, 7929-7932; '2-Deoxy-2-fluoro-D-glycosyl fluorides. A new class of specific mechanism-based glycosidase inhibitors'.

[281] Kappelhoff, J. C., Liu, S. Y. J., Dugdale, M. L., Dymianiw, D. L., Linton, L. R., Huber, R. E. *Protein J.* **2009**, *28 (2)*, 96-103; 'Practical Considerations When Using Temperature to Obtain Rate Constants and Activation Thermodynamics of Enzymes with Two Catalytic Steps: Native and N460T-β-Galactosidase (*E. coli*) as Examples'.

[282] Huber, R. E., Gaunt, M. T., Hurlburt, K. L. *Arch. Biochem. Biophys.* **1984**, *234 (1)*, 151-160; 'Binding and Reactivity at the Glucose Site of Galactosyl-β-Galactosidase (*Escherichia coli*)'.

[283] Pereira-Rodríguez, Á., Fernández-Leiro, R., González-Siso, M. I., Cerdán, M. E., Becerra, M., Sanz-Aparicio, J. *J. Struct. Biol.* **2012**, *177 (2)*, 392-401; 'Structural basis of specificity in tetrameric *Kluyveromyces lactis* β-galactosidase'.

[284] Husain, Q. *Crit. Rev. Biotechnol.* **2010**, *30 (1)*, 41-62; 'β-Galactosidases and their potential applications: a review'.

[285] Rodríguez, Á. P., Leiro, R. F., Trillo, M. C., Cerdán, M. E., Siso, M. I. G., Becerra, M. *Microb. Cell Fact.* **2006**, *5*; 'Secretion and properties of a hybrid *Kluyveromyces lactis-Aspergillus niger* β-galactosidase'.

[286] Todorova-Balvay, D., Stoilova, I., Gargova, S., Vijayalakshmi, M. A. *J. Mol. Recognit.* **2006**, *19 (4)*, 299-304; 'An efficient two step purification and molecular characterization of β-galactosidases from *Aspergillus oryzae*'.

[287] Rice, P., Longden, I., Bleasby, A. *Trends Genet.* **2000**, *16 (6)*, 276-277; 'EMBOSS: The European Molecular Biology Open Software Suite'.

[288] Pereira-Rodríguez, Á., Fernández-Leiro, R., Siso, M. I. G., Cerdán, M. E., Becerra, M., Sanz-Aparicio, J. *Acta Crystallogr F* **2010**, *66*, 297-300; 'Crystallization and preliminary X-ray crystallographic analysis of β-galactosidase from *Kluyveromyces lactis*'.

[289] Athès, V., Lange, R., Combes, D. *Eur. J. Biochem.* **1998**, *255 (1)*, 206-212; 'Influence of polyols on the structural properties of *Kluyveromyces lactis* β-galactosidase under high hydrostatic pressure'.

[290] Morka, K., Pietruszka, J., Meyer zu Berstenhorst, S. *J. Biotechnol.* **2014**, *191*, 176-186; 'Comparative expression of lipase CAL-A in the yeasts *Saccharomyces cerevisiae*, *Kluyveromyces lactis* and *Hansenula polymorpha* to investigate a possible host influence'.

[291] Farina, F., Uccelletti, D., Goffrini, P., Butow, R. A., Abeijon, C., Palleschi, C. *Biochem. Biophys. Res. Commun.* **2004**, *318 (4)*, 1031-1038; 'Alterations of *O*-glycosylation, cell wall, and mitochondrial metabolism in *Kluyveromyces lactis* cells defective in KlPmr1p, the Golgi Ca^{2+}-ATPase'.

[292] Uccelletti, D., Farina, F., Morlupi, A., Palleschi, C. *Res. Microbiol.* **1999**, *150 (1)*, 5-12; 'Mutants of Kluyveromyces lactis with altered protein glycosylation are affected in cell wall morphogenesis'.

[293] Kim, C. S., Ji, E.-S., Oh, D.-K. *Biotechnol. Lett.* **2003**, *25 (20)*, 1769-1774; 'Expression and characterization of *Kluyveromyces lactis* β-galactosidase in *Escherichia coli*'.

[294] Poch, O., L'Hôte, H., Dallery, V., Debeaux, F., Fleer, R., Sodoyer, R. *Gene* **1992**, *118 (1)*, 55-63; 'Sequence of the *Kluyveromyces lactis* β-galactosidase: comparison with prokaryotic enzymes and secondary structure analysis'.

[295] Maksimainen, M. M., Lampio, A., Mertanen, M., Turunen, O., Rouvinen, J. *Int. J. Biol. Macromol.* **2013**, *60*, 109-115; 'The crystal structure of acidic β-galactosidase from *Aspergillus oryzae*'.

[296] Holsinger, V. H., Rajkowski, K. T., Stabel, J. R. *Rev. Sci. Tech. Off. Int. Epizoot.* **1997**, *16 (2)*, 441-451; 'Milk Pasteurisation and Safety: a brief history and update'.

[297] Rojas, A. L., Nagem, R. A. P., Neustroev, K. N., Arand, M., Adamska, M., Eneyskaya, E. V., Kulminskaya, A. A., Garratt, R. C., Golubev, A. M., Polikarpov, I. *J. Mol. Biol.* **2004**, *343 (5)*, 1281-1292; 'Crystal Structures of β-Galactosidase from *Penicillium* sp. and its Complex with Galactose'.

[298] Deshpande, N., Wilkins, M. R., Packer, N., Nevalainen, H. *Glycobiology* **2008**, *18 (8)*, 626-637; 'Protein glycosylation pathways in filamentous fungi'.

[299] Kainz, E., Gallmetzer, A., Hatzl, C., Nett, J. H., Li, H., Schinko, T., Pachlinger, R., Berger, H., Reyes-Dominguez, Y., Bernreiter, A., Gerngross, T., Wildt, S., Strauss, J. *Appl. Environ. Microbiol.* **2008**, *74 (4)*, 1076-1086; 'N-Glycan Modification in *Aspergillus* Species'.

[300] Huber, R. E., Gaunt, M. T. *Arch. Biochem. Biophys.* **1983**, *220 (1)*, 263-271; 'Importance of Hydroxyls at Positions 3, 4, and 6 for Binding to the Galactose Site of β-Galactosidase (*Escherichia coli*)'.

[301] Horwitz, J. P., Chua, J., Curby, R. J., Tomson, A. J., Da Rooge, M. A., Fisher, B. E., Mauricio, J., Klundt, I. *J. Med. Chem.* **1964**, *7 (4)*, 574-575; 'Substrates for Cytochemical Demonstration of Enzyme Activity. I. Some Substituted 3-Indolyl-β-D-glycopyranosides'.

[302] Boos, W. In *Methods Enzymol.*; Academic Press: **1982**; Vol. Volume 89, p 59-64; '[10] Synthesis of (2R)-glycerol-*O*-β-D-galactopyranoside by β-galactosidase'.

[303] Grant, S. G., Jessee, J., Bloom, F. R., Hanahan, D. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87 (12)*, 4645-4649; 'Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants'.

[304] New England Biolabs, I.; New England Biolabs, Inc.: Ipswich, **2014**; 'C2527 Datasheet'.

[305] Stratagene **2007**; 'QuikChange® Site-Directed Mutagenesis Kit'.

[306] Wang, W. Y., Malcolm, B. A. *BioTechniques* **1999**, *26 (4)*, 680-682; 'Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange (TM) site-directed mutagenesis'.

[307] Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J., Withers, S. G. *J. Am. Chem. Soc.* **1994**, *116 (25)*, 11594-11595; 'Changing Enzymic Reaction Mechanisms by Mutagenesis: Conversion of a Retaining Glucosidase to an Inverting Enzyme'.

[308] Loeschke, A. Doctoral, Heinrich-Heine-Universität, **2012**; 'Funktionelle Koexpression geclusterter Gene in verschiedenen Bakterien mithilfe des TREX-Systems'.

[309] Dick, M., Hartmann, R., Weiergraber, O. H., Bisterfeld, C., Classen, T., Schwarten, M., Neudecker, P., Willbold, D., Pietruszka, J. *Chemical Science* **2016**, *7 (7)*, 4492-4502; 'Mechanism-based inhibition of an aldolase at high concentrations of its natural substrate acetaldehyde: structural insights and protective strategies'.

[310] Holec, C., Neufeld, K., Pietruszka, J. *Adv. Synth. Catal.* **2016**, *358 (11)*, 1810-1819; 'P450 BM3 Monooxygenase as an Efficient NAD(P)H-Oxidase for Regeneration of Nicotinamide Cofactors in ADH-Catalysed Preparative Scale Biotransformations'.

[311] Holec, C., Sandkuhl, D., Rother, D., Kroutil, W., Pietruszka, J. *ChemCatChem* **2015**, *7 (19)*, 3125-3130; 'Chemoenzymatic Synthesis towards the Active Agent Travoprost'.

[312] Rüthlein, E., Classen, T., Dobnikar, L., Schölzel, M., Pietruszka, J. *Adv. Synth. Catal.* **2015**, *357 (8)*, 1775-1786; 'Finding the Selectivity Switch – A Rational Approach towards Stereocomplementary Variants of the Ene Reductase YqjM'.

[313] Fallingborg, J. *Dan. Med. Bull.* **1999**, *46 (3)*, 183-196; 'Intraluminal pH of the human gastrointestinal tract'.

[314] Rosano, G. L., Ceccarelli, E. A. *Frontiers in Microbiology* **2014**, *5*, 172; 'Recombinant protein expression in *Escherichia coli*: advances and challenges'.

[315] Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., van Oudenaarden, A. *Nat. Genet.* **2002**, *31 (1)*, 69-73; 'Regulation of noise in the expression of a single gene'.

[316] Classen, T. Doctoral, Heinerich-Heine-Universität Düsseldorf, **2013**; 'Über Enzyme als facettenreiche Katalysatoren in der Organischen Chemie'.

[317] Freier, L., von Lieres, E., **2016**, 'KriKit' (v 1.0.3)

[318] Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R., Stuber, D. *Nat Biotech* **1988**, *6 (11)*, 1321-1325; 'Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent'.

[319] Kim, C. S., Ji, E.-S., Oh, D.-K. *Biochem. Biophys. Res. Commun.* **2004**, *316 (3)*, 738-743; 'A new kinetic model of recombinant β-galactosidase from *Kluyveromyces lactis* for both hydrolysis and transgalactosylation reactions'.

[320] Martínez-Villaluenga, C., Cardelle-Cobas, A., Corzo, N., Olano, A., Villamiel, M. *Food Chem.* **2008**, *107 (1)*, 258-264; 'Optimization of conditions for galactooligosaccharide synthesis during lactose hydrolysis by β-galactosidase from *Kluyveromyces lactis* (Lactozym 3000 L HP G)'.

[321] Mateo, C., Monti, R., Pessela, B. C. C., Fuentes, M., Torres, R., Manuel Guisán, J., Fernández-Lafuente, R. *Biotechnol. Prog.* **2004**, *20 (4)*, 1259-1262; 'Immobilization of Lactase from Kluyveromyces lactis Greatly Reduces the Inhibition Promoted by Glucose. Full Hydrolysis of Lactose in Milk'.

[322] Del Moral-Ramírez, E., Domínguez-Ramírez, L., Cruz-Guerrero, A. E., Rodríguez-Serrano, G. M., García-Garibay, M., Gómez-Ruiz, L., Jiménez-Guzmán, J. *J. Agric. Food Chem.* **2008**, *56 (14)*, 5859-5863; 'Role of lysine ε-amino groups of β-lactoglobulin on its activating effect of *Kluyveromyces lactis* β-galactosidase'.

[323] Plotkin, J. B., Kudla, G. *Nat. Rev. Genet.* **2011**, *12 (1)*, 32-42; 'Synonymous but not the same: the causes and consequences of codon bias'.

[324] Akarsu, A. N., Stoilov, I., Yilmaz, E., Sayli, B. S., Sarfarazi, M. *Hum. Mol. Genet.* **1996**, *5 (7)*, 945-952; 'Genomic structure of *HOXD13* gene: a nine polyalanine duplication causes synpolydactyly in two unrelated families'.

[325] Amiel, J., Laudier, B., Attie-Bitach, T., Trang, H., de Pontual, L., Gener, B., Trochet, D., Etchevers, H., Ray, P., Simonneau, M., Vekemans, M., Munnich, A., Gaultier, C., Lyonnet, S. *Nat. Genet.* **2003**, *33 (4)*, 459-461; 'Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome'.

[326] Warrick, J. M., Chan, H. Y. E., Gray-Board, G. L., Chai, Y., Paulson, H. L., Bonini, N. M. *Nat. Genet.* **1999**, *23 (4)*, 425-428; 'Suppression of polyglutamine-mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70'.

[327] Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D. W., Amos, C., Dobyns, W. B., Subramony, S. H., Zoghbi, H. Y., Lee, C. C. *Nat. Genet.* **1997**, *15 (1)*, 62-69; 'Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the α_{1A}-voltage-dependent calcium channel'.

[328] Toulouse, A., Au-Yeung, F., Gaspar, C., Roussel, J., Dion, P., Rouleau, G. A. *Hum. Mol. Genet.* **2005**, *14 (18)*, 2649-2660; 'Ribosomal frameshifting on MJD-1 transcripts with long CAG tracts'.

[329] Akasaki, M., Suzuki, M., Funakoshi, I., Yamashina, I. *J. Biochem.* **1976**, *80 (6)*, 1195-1200; 'Characterization of β-Galactosidase from a Special Strain of *Aspergillus oryzae*'.

[330] Tanaka, Y., Kagamiishi, A., Kiuchi, A., Horiuchi, T. *J. Biochem.* **1975**, *77 (1)*, 241-247; 'Purification and Properties of β-Galactosidase from *Aspergillus oryzae*'.

[331] Kudo, K., Watanabe, A., Ujiie, S., Shintani, T., Gomi, K. *J. Biosci. Bioeng.* **2015**, *120 (6)*, 614-623; 'Purification and enzymatic characterization of secretory glycoside hydrolase family 3 (GH3) aryl β-glucosidases screened from *Aspergillus oryzae* genome'.

[332] Langston, J., Sheehy, N., Xu, F. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764 (5)*, 972-978; 'Substrate specificity of *Aspergillus oryzae* family 3 β-glucosidase'.

[333] Cowan, D. A. *Comp. Biochem. Physiol. A Physiol.* **1997**, *118A (3)*, 429-438; 'Thermophilic proteins: Stability and function in aqueous and organic solvents'.

[334] Ito, J., Sahara, H., Kaya, M., Hata, Y., Shibasaki, S., Kawata, K., Ishida, S., Ogino, C., Fukuda, H., Kondo, A. *J. Mol. Catal. B: Enzym.* **2008**, *55 (1–2)*, 69-75; 'Characterization of yeast cell surface displayed *Aspergillus oryzae* β-glucosidase 1 high hydrolytic activity for soybean isoflavone'.

[335] Prabha, T. R., Revathi, K., Vinod, M. S., Shanthakumar, S. P., Bernard, P. *Curr. Sci.* **2013**, *104 (3)*, 345-347; 'A simple method for total genomic DNA extraction from water moulds'.

[336] Bir, N., Paliwal, A., Muralidhar, K., Reddy, P., Sarma, P. U. *Prep. Biochem.* **1995**, *25 (4)*, 171-181; 'A Rapid Method for the Isolation of Genomic DNA from *Aspergillus fumigatus*'.

[337] National Center for Biotechnology Information, **2016**, 'BLAST®: Basic Local Alignment Search Tool' (*http://blast.ncbi.nlm.nih.gov/Blast.cgi*)

[338] Dereeper, A., Audic, S., Claverie, J. M., Blanc, G. *BMC Evol. Biol.* **2010**, *10*; 'BLAST-EXPLORER helps you building datasets for phylogenetic analysis'.

[339] Ito, Y., Sasaki, T., Kitamoto, K., Kumagai, C., Takahashi, K., Gomi, K., Tamura, G. *J. Gen. Appl. Microbiol.* **2002**, *48 (3)*, 135-142; 'Cloning, nucleotide sequencing, and expression of the β-galactosidase-encoding gene (lacA) from *Aspergillus oryzae*'.

[340] Hayes, M., Personal Communication.

[341] Van Lieshout, J., Faijes, M., Nieto, J., Van Der Oost, J., Planas, A. *Archaea* **2004**, *1 (4)*; 'Hydrolase and glycosynthase activity of endo-1,3-β-glucanase from the thermophile *Pyrococcus furiosus*'.

[342] Hancock, S. M., Rich, J. R., Caines, M. E. C., Strynadka, N. C. J., Withers, S. G. *Nat. Chem. Biol.* **2009**, *5 (7)*, 508-514; 'Designer enzymes for glycosphingolipid synthesis by directed evolution'.

[343] Honda, Y., Kitaoka, M. *J. Biol. Chem.* **2006**, *281 (3)*, 1426-1431; 'The first glycosynthase derived from an inverting glycoside hydrolase'.

[344] Hoffmann, N., Pietruszka, J., Söffing, C. *Adv. Synth. Catal.* **2012**, *354 (6)*, 959-963; 'From Sophorose Lipids to Natural Product Synthesis'.

[345] Percec, V., Leowanawat, P., Sun, H.-J., Kulikov, O., Nusbaum, C. D., Tran, T. M., Bertin, A., Wilson, D. A., Peterca, M., Zhang, S. D., Kamat, N. P., Vargo, K., Moock, D., Johnston, E. D., Hammer, D. A., Pochan, D. J., Chen, Y. C., Chabre, Y. M., Shiao, T. C., Bergeron-Brlek, M., André, S., Roy, R., Gabius, H.-J., Heiney, P. A. *J. Am. Chem. Soc.* **2013**, *135 (24)*, 9055-9077; 'Modular Synthesis of Amphiphilic Janus Glycodendrimers and Their Self-Assembly into Glycodendrimersomes and Other Complex Architectures with Bioactivity to Biomedically Relevant Lectins'.

[346] Wolfrom, M. L., Groebke, W. *J. Org. Chem.* **1963**, *28 (11)*, 2986-2988; 'Tetra-*O*-acylglycosyl Chlorides from 1-Thioglycosides and Their Conversion to Penta-*O*-acyl Esters'.

[347] Bretschneider, H., Beran, K. *Monatsh. Chem. Verw. Teile Anderer Wiss.* **1949**, *80 (2)*, 262-270; 'Überführung ein- und zweiwertiger Phenole in acetylierte β-D-Glucoside mit β-Pentaacetyl-D-glucose und Borfluorid'.

[348] Yamaguchi, M., Horiguchi, A., Fukuda, A., Minami, T. *J. Chem. Soc., Perkin Trans. 1* **1990***(4)*, 1079-1082; 'Novel synthesis of aryl 2,3,4,6-tetra-*O*-acetyl-D-glucopyranosides'.

[349] Lee, Y. S., Rho, E. S., Min, Y. K., Kim, B. T., Kim, K. H. *J. Carbohydr. Chem.* **2001**, *20 (6)*, 503-506; 'Practical β-stereoselective *O*-glycosylation of phenols with penta-*O*-acetyl-β-D-glucopyranose'.

[350] Brauns, D. H. *J. Am. Chem. Soc.* **1923**, *45 (3)*, 833-835; 'Fluoro-acetyl derivatives of sugars. I.'.

[351] Steinmann, A., Thimm, J., Matwiejuk, M., Thiem, J. *Macromolecules* **2010**, *43 (8)*, 3606-3612; 'Formation of Homooligosaccharides Using Base-Promoted Glycosylation of Unprotected Glycosyl Fluorides'.

[352] Williams, D. R., Sit, S.-Y. *J. Am. Chem. Soc.* **1984**, *106 (10)*, 2949-2954; 'Total synthesis of (+)-phyllanthocin'.

[353] Lavallée, P., Ruel, R., Grenier, L., Bissonnette, M. *Tetrahedron Lett.* **1986**, *27 (6)*, 679-682; 'Convenient access to two enantiomeric oxirane synthons bearing a quaternary gem-dimethyl carbon center: Synthesis of 3*S*-(+) and 3*R*-(−)-2,2-dimethyl-3,4-oxo-1-butanol from *R*-(−)-pantolactone'.

[354] Sennari, G., Hirose, T., Iwatsuki, M., Ōmura, S., Sunazuka, T. *Chem. Commun.* **2014**, *50 (63)*, 8715-8718; 'A concise total synthesis of puberulic acid, a potent antimalarial agent'.

[355] Weigel, T. M., Liu, H. *Tetrahedron Lett.* **1988**, *29 (34)*, 4221-4224; 'Synthesis of Stereospecifically Labeled Carbohydrates II: Preparation of (3*S*)- and (3*R*)-[3-$^{2}H_{1}$]Abequose'.

[356] Denmark, S. E., Regens, C. S., Kobayashi, T. *J. Am. Chem. Soc.* **2007**, *129 (10)*, 2774-2776; 'Total Synthesis of Papulacandin D'.

[357] Taber, D. F., Reddy, P. G., Ameson, K. O. *J. Org. Chem.* **2008**, *73 (9)*, 3467-3474; 'A potential route to neuroprostanes and isoprostanes: Preparation of the four enantiomerically pure diastereomers of 13-F_{4t}-NeuroP'.

[358] Brown, H. C., Krishnamurthy, S. *J. Am. Chem. Soc.* **1973**, *95 (5)*, 1669-1671; 'Lithium triethylborohydride. Exceptionally powerful nucleophile in displacement reactions with organic halides'.

[359] Krishnamurthy, S., Brown, H. C. *J. Org. Chem.* **1976**, *41 (18)*, 3064-3066; 'Facile reduction of alkyl tosylates with lithium triethylborohydride. An advantageous procedure for the deoxygenation of cyclic and acyclic alcohols'.

[360] Srivastava, V. K., Lerner, L. M. *Carbohydr. Res.* **1978**, *64*, 263-265; 'A facile synthesis of 6-deoxy-D-glucose. Reduction of a peimary suifonate with lithium triethylborohydride'.

[361] Wang, C.-C., Kulkarni, S. S., Lee, J.-C., Luo, S.-Y., Hung, S.-C. *Nat. Protoc.* **2008**, *3 (1)*, 97-113; 'Regioselective one-pot protection of glucose'.

[362] Doukas, H. M., Fontaine, T. D. *J. Am. Chem. Soc.* **1951**, *73 (12)*, 5917-5918; 'Cleavage of sapogenin terminal rings with lithium aluminum hydride'.

[363] Leggetter, B. E., Brown, R. K. *Can. J. Chem.* **1964**, *42 (5)*, 990-1004; 'The influence of substituents on the ease and direction of ring opening in the $LiAlH_4$-$AlCl_3$ reductive cleavage of substituted 1,3-dioxolanes'.

[364] Lipták, A., Jodál, I., Nánási, P. *Carbohydr. Res.* **1975**, *44 (1)*, 1-11; 'Stereoselective ring-cleavage of 3-*O*-benzyl- and 2,3-di-*O*-benzyl-4,6-*O*-benzylidenehexopyranoside derivatives with the $LiAlH_4$·$AlCl_3$, reagent'.

[365] J. Garegg, P., Hultberg, H., Wallin, S. *Carbohydr. Res.* **1982**, *108 (1)*, 97-101; 'A novel, reductive ring-opening of carbohydrate benzylidene acetals'.

[366] Debenham, S. D., Toone, E. J. *Tetrahedron: Asymmetry* **2000**, *11 (2)*, 385-387; 'Regioselective reduction of 4,6-O-benzylidenes using triethylsilane and BF_3·Et_2O'.

[367] Garegg, P. J., Hultberg, H. *Carbohydr. Res.* **1981**, *93 (1)*, C10-C11; 'A novel, reductive ring-opening of carbohydrate benzylidene acetals, with unusual regioselectivity'.

[368] Oikawa, M., Liu, W.-C., Nakai, Y., Koshida, S., Fukase, K., Kusumoto, S. *Synlett* **1996**, *1996 (12)*, 1179-1180; 'Regioselective Reductive Opening of 4,6-*O*-Benzylidene Acetals of Glucose or Glucosamine Derivatives by BH_3·Me_2NH-BF_3·OEt_2'.

[369] Sakagami, M., Hamana, H. *Tetrahedron Lett.* **2000**, *41 (29)*, 5547-5551; 'A selective ring opening reaction of 4,6-*O*-benzylidene acetals in carbohydrates using trialkylsilane derivatives'.

[370] Shie, C.-R., Tzeng, Z.-H., Kulkarni, S. S., Uang, B.-J., Hsu, C.-Y., Hung, S.-C. *Angew. Chem. Int. Ed.* **2005**, *44 (11)*, 1665-1668; 'Cu(OTf)$_2$ as an Efficient and Dual-Purpose Catalyst in the Regioselective Reductive Ring Opening of Benzylidene Acetals'.

[371] Johnsson, R., Ohlin, M., Ellervik, U. *J. Org. Chem.* **2010**, *75 (23)*, 8003-8011; 'Reductive Openings of Benzylidene Acetals Revisited: A Mechanistic Scheme for Regio- and Stereoselectivity'.

[372] Rising, T. W. D. F., Heidecke, C. D., Moir, J. W. B., Ling, Z. L., Fairbanks, A. J. *Chem. Eur. J.* **2008**, *14 (21)*, 6444-6464; 'Endohexosaminidase-catalysed glycosylation with oxazoline donors: Fine tuning of catalytic efficiency and reversibility'.

[373] Sato, K., Igarashi, T., Yanagisawa, Y., Kawauchi, N., Hashimoto, H., Yoshimura, J. *Chem. Lett.* **1988**, *17 (10)*, 1699-1702; 'Oxidative Cleavage of 4,6-*O*-Benzylidene Ring with *t*-Butyl Hydroperoxide and Copper(II) Chloride - Preparation of Methyl 4-*O*- and 6-*O*-Benzoylhexopyranoside Derivatives'.

[374] Adinolfi, M., Barone, G., Guariniello, L., Iadonisi, A. *Tetrahedron Lett.* **1999**, *40 (48)*, 8439-8441; 'Facile cleavage of carbohydrate benzyl ethers and benzylidene acetals using the $NaBrO_3/Na_2S_2O_4$ reagent under two-phase conditions'.

[375] Kim, S. S., Jung, H. C. *Synthesis* **2003***(14)*, 2135-2137; 'An Efficient Aerobic Oxidation of Alcohols to Aldehydes and Ketones with TEMPO/Ceric Ammonium Nitrate as Catalysts'.

[376] Zhang, Z. Y., Magnusson, G. *J. Org. Chem.* **1996**, *61 (7)*, 2394-2400; 'DDQ-mediated oxidation of 4,6-*O*-methoxybenzylidene-protected saccharides in the presence of various nucleophiles: Formation of 4-OH, 6-Cl, and 6-Br derivatives'.

[377] Wuts, P. G. M., Greene, T. W. In *Greene's Protective Groups in Organic Synthesis*; John Wiley & Sons, Inc.: **2006**, p 16-366; 'Protection for the Hydroxyl Group, Including 1,2- and 1,3-Diols'.

[378] Haines, A. H. In *Adv. Carbohydr. Chem. Biochem.*; Tipson, R. S., Derek, H., Eds.; Academic Press: **1976**; Vol. 33, p 11-109; 'Relative Reactivities of Hydroxyl Groups in Carbohydrates'.

[379] Nelson, T. D., Crouch, R. D. *Synthesis* **1996**, *1996 (09)*, 1031-1069; 'Selective Deprotection of Silyl Ethers'.

[380] Crouch, R. D. *Tetrahedron* **2004**, *60 (28)*, 5833-5871; 'Selective monodeprotection of bis-silyl ethers'.

[381] Fylaktakidou, K. C., Duarte, C. D., Koumbis, A. E., Nicolau, C., Lehn, J.-M. *ChemMedChem* **2011**, *6 (1)*, 153-168; 'Polyphosphates and Pyrophosphates of Hexopyranoses as Allosteric Effectors of Human Hemoglobin: Synthesis, Molecular Recognition, and Effect on Oxygen Release'.

[382] Bothwell, J. M., Angeles, V. V., Carolan, J. P., Olson, M. E., Mohan, R. S. *Tetrahedron Lett.* **2010**, *51 (7)*, 1056-1058; 'A mild and chemoselective method for the deprotection of tert-butyldimethylsilyl (TBDMS) ethers using iron(III) tosylate as a catalyst'.

[383] Kumar, G. D. K., Baskaran, S. *J. Org. Chem.* **2005**, *70 (11)*, 4520-4523; 'A Facile, Catalytic, and Environmentally Benign Method for Selective Deprotection of tert-Butyldimethylsilyl Ether Mediated by Phosphomolybdic Acid Supported on Silica Gel'.

[384] Wang, B., Sun, H.-X., Chen, B., Sun, Z.-H. *Green Chem.* **2009**, *11 (8)*, 1112-1114; 'Practical, environment-benign and atom economic KOAc-catalysed deprotection of aryl TIPS ethers under mild fluoride-free conditions'.

[385] Yan, S. Q., Ding, N., Zhang, W., Wang, P., Li, Y. X., Li, M. *Carbohydr. Res.* **2012**, *354*, 6-20; 'An efficient and recyclable catalyst for the cleavage of *tert*-butyldiphenylsilyl ethers'.

[386] Ziegler, T., Eckhardt, E., Pantkowski, G. *J. Carbohydr. Chem.* **1994**, *13 (1)*, 81-109; 'Regioselective Glycodesilylation of Silylated Glycosides as a Useful Tool for the Preparation of Oligosaccharides'.

[387] Kunz, H., Sager, W. *Helv. Chim. Acta* **1985**, *68 (1)*, 283-287; 'Stereoselective glycosylation of Alcohols and Silyl Ethers Using Glycosyl Fluorides and Boron Trifluoride Etherate'.

[388] Troyano, E., Olano, A., Fernández-Díaz, M., Sanz, J., Martínez-Castro, I. *Chromatographia* **1991**, *32 (7-8)*, 379-382; 'Gas Chromatographic Analysis of Free Monosaccharides in Milk'.

[389] Soria, A. C., Sanz, M. L., Villamiel, M. *Food Chem.* **2009**, *114 (2)*, 758-762; 'Determination of minor carbohydrates in carrot (*Daucus carota* L.) by GC-MS'.

[390] Medeiros, P. M., Simoneit, B. R. T. *J. Chromatogr. A* **2007**, *1141 (2)*, 271-278; 'Analysis of sugars in environmental samples by gas chromatography-mass spectrometry'.

[391] Isidorov, V. A., Szczepaniak, L. *J. Chromatogr. A* **2009**, *1216 (51)*, 8998-9007; 'Gas chromatographic retention indices of biologically and environmentally important organic compounds on capillary columns with low-polar stationary phases'.

[392] Rogatsky, E., Tomuta, V., Stein, D. T. *Anal. Chim. Acta* **2007**, *591 (2)*, 155-160; 'LC/MS quantitative study of glucose by iodine attachment'.

[393] Ho, C. S., Lam, C. W. K., Chan, M. H. M., Cheung, R. C. K., Law, L. K., Lit, L. C. W., Ng, K. F., Suen, M. W. M., Tai, H. L. *Clin. Biochem. Rev.* **2003**, *24 (1)*, 3-12; 'Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications'.

[394] Guilhaus, M. *J. Mass Spectrom.* **1995**, *30 (11)*, 1519-1532; 'Special feature: Tutorial. Principles and instrumentation in time-of-flight mass spectrometry. Physical and instrumental concepts'.

[395] Buszewski, B., Noga, S. *Anal. Bioanal. Chem.* **2012**, *402 (1)*, 231-247; 'Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique'.

[396] Teleki, A., Sánchez-Kopper, A., Takors, R. *Anal. Biochem.* **2015**, *475*, 4-13; 'Alkaline conditions in hydrophilic interaction liquid chromatography for intracellular metabolite quantification using tandem mass spectrometry'.

[397] Ribeiro, M. H. *Appl. Microbiol. Biotechnol.* **2011**, *90 (6)*, 1883-1895; 'Naringinases: occurrence, characteristics, and applications'.

[398] García-Borrón, J. C., Solano, F. *Pigm. Cell Res.* **2002**, *15 (3)*, 162-173; 'Molecular Anatomy of Tyrosinase and its Related Proteins: Beyond the Histidine-Bound Metal Catalytic Center'.

[399] Feizi, T., Larkin, M. *Glycobiology* **1990**, *1 (1)*, 17-23; 'AIDS and glycosylation'.

[400] Moran, A. P., Gupta, A., Joshi, L. *Gut* **2011**; 'Sweet-talk: role of host glycosylation in bacterial pathogenesis of the gastrointestinal tract'.

[401] Chang, R. C.-C., So, K.-F. *Cell. Mol. Neurobiol.* **2008**, *28 (5)*, 643-652; 'Use of Anti-aging Herbal Medicine, *Lycium barbarum*, Against Aging-associated Diseases. What Do We Know So Far?'.

[402] Alarcón de la Lastra, C., Villegas, I. *Mol. Nutr. Food Res.* **2005**, *49 (5)*, 405-430; 'Resveratrol as an anti-inflammatory and anti-aging agent: Mechanisms and clinical implications'.

[403] Vic, G., Hastings, J. J., Crout, D. H. G. *Tetrahedron: Asymmetry* **1996**, *7 (7)*, 1973-1984; 'Glycosidase-catalysed synthesis of glycosides by an improved procedure for reverse hydrolysis: application to the chemoenzymatic synthesis of galactopyranosyl-(1→4)-*O*-α-galactopyranoside derivatives'.

[404] Yasukochi, T., Fukase, K., Kusumoto, S. *Tetrahedron Lett.* **1999**, *40 (36)*, 6591-6593; '3-Nitro-2-pyridyl glycoside as donor for chemical glycosylation and its application to chemoenzymatic synthesis of oligosaccharide'.

[405] Zanini, D., Roy, R. *Bioconjugate Chem.* **1997**, *8 (2)*, 187-192; 'Chemoenzymatic Synthesis and Lectin Binding Properties of Dendritic *N*-Acetyllactosamine'.

[406] Qian, X., Sujino, K., Otter, A., Palcic, M. M., Hindsgaul, O. *J. Am. Chem. Soc.* **1999**, *121 (51)*, 12063-12072; 'Chemoenzymatic Synthesis of α-(1→3)-Gal(NAc)-Terminating Glycosides of Complex Tertiary Sugar Alcohols'.

[407] Kim, S., Matsuo, I., Ajisaka, K., Nakajima, H., Kitamoto, K. *Biosci. Biotechnol. Biochem.* **2002**, *66 (10)*, 2168-2175; 'Cloning and Characterization of the nagA Gene that Encodes β-*N*-Acetylglucosaminidase from *Aspergillus nidulans* and Its Expression in *Aspergillus oryzae*'.

[408] Knapp, A., Voget, S., Gao, R., Zaburannyi, N., Krysciak, D., Breuer, M., Hauer, B., Streit, W. R., Müller, R., Daniel, R., Jaeger, K.-E. *Appl. Microbiol. Biotechnol.* **2016**, *100 (3)*, 1265-1273; 'Mutations improving production and secretion of extracellular lipase by *Burkholderia glumae* PG1'.

[409] Vogt, C. M., Schraner, E. M., Aguilar, C., Eichwald, C. *Microb. Cell Fact.* **2016**, *15*, 137; 'Heterologous expression of antigenic peptides in *Bacillus subtilis* biofilms'.

[410] Schmidt, T. G. M., Koepke, J., Frank, R., Skerra, A. *J. Mol. Biol.* **1996**, *255 (5)*, 753-766; 'Molecular Interaction Between the *Strep*-tag Affinity Peptide and its Cognate Target, Streptavidin'.

[411] Ho, P. W.-L., Tse, Z. H.-M., Liu, H.-F., Lu, S., Ho, J. W.-M., Kung, M. H.-W., Ramsden, D. B., Ho, S.-L. *PLoS One* **2013**, *8 (9)*, e74065; 'Assessment of Cellular Estrogenic Activity Based on Estrogen Receptor-Mediated Reduction of Soluble-Form Catechol-*O*-Methyltransferase (COMT) Expression in an ELISA-Based System'.

[412] Bedouelle, H., Duplay, P. *Eur. J. Biochem.* **1988**, *171 (3)*, 541-549; 'Production in Escherichia coli and one-step purification of bifunctional hybrid proteins which bind maltose'.

[413] Leriche, G., Chisholm, L., Wagner, A. *Bioorg. Med. Chem.* **2012**, *20 (2)*, 571-582; 'Cleavable linkers in chemical biology'.

[414] Khandurina, J., Blum, D. L., Stege, J. T., Guttman, A. *Electrophoresis* **2004**, *25 (14)*, 2326-2331; 'Automated carbohydrate profiling by capillary electrophoresis: A bioindustrial approach'.

[415] Moini, M. In *Capillary Electrophoresis of Biomolecules: Methods and Protocols*; Volpi, N., Maccari, F., Eds.; Humana Press: Totowa, NJ, **2013**, p 79-119; 'High-Throughput Capillary Electrophoresis–Mass Spectrometry: From Analysis of Amino Acids to Analysis of Protein Complexes'.

[416] Oefner, P. J., Chiesa, C. *Glycobiology* **1994**, *4 (4)*, 397-412; 'Capillary electrophoresis of carbohydrates'.

[417] Li, X., Jackson, P., Rubtsov, D. V., Faria-Blanc, N., Mortimer, J. C., Turner, S. R., Krogh, K. B., Johansen, K. S., Dupree, P. *Biotechnol. Biofuels* **2013**, *6*, 94-94; 'Development and application of a high throughput carbohydrate profiling technique for analyzing plant cell wall polysaccharides and carbohydrate active enzymes'.

[418] Adobe, **2012**, 'Adobe® Photoshop® CS6 Extended' (v 13.0 x64)

[419] Scientific & Educational Software, **2015**, 'Clone Manager' (v 9.4)

[420] Manske, M., **2003**, 'GENtle' (v 1.9.4)

[421] Freier, L., Hemmerich, J., Schöler, K., Wiechert, W., Oldiges, M., von Lieres, E. *Eng. Life Sci.* **2016**, 538-549; 'Framework for Kriging-based iterative experimental analysis and design: Optimization of secretory protein production in *Corynebacterium glutamicum*'.

[422] OriginLab Corporation, **2013**, 'Origin®' (v 9.1.0 (64-bit) Sr2)

[423] Schrödinger, **2010**, 'PyMOL™' (v 1.3)

[424] Ferrin, T., Huang, C., Couch, G., Petterson, E., Goddard, T., Meng, E., Yang, Z., Hertig, S., Weber, D., Mischel, D., Robles, V. M., Maréchal, J.-D., Greenblatt, D., Konerding, D., Pintilie, G., Zhang, W., Lang, T., Schreiber, S., Chen, J. E., Morris, B., Margraf, T., **2013**, 'UCSF Chimera' (v 1.8)

[425] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., Ferrin, T. E. *J. Comput. Chem.* **2004**, *25 (13)*, 1605-1612; 'UCSF Chimera - A visualization system for exploratory research and analysis'.

[426] European Bioinformatics Institute, **2016**, 'ClustalW2' (*http://www.ebi.ac.uk/Tools/msa/clustalw2/*)

[427] Kibbe, W. A., Buehler, E., Cao, Q., **2015**, 'Oligo Calc: oligonucleotide properties calculator' (*http://biotools.nubic.northwestern.edu/OligoCalc.html*)

[428] Kibbe, W. A. *Nucleic Acids Res.* **2007**, *35*, W43-W46; 'OligoCalc: an online oligonucleotide properties calculator'.

[429] Chazan, A., **1998**, 'Peptide Property Calculator' (*http://biotools.nubic.northwestern.edu/proteincalc.html*)

[430] Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., Gascuel, O. *Nucleic Acids Res.* **2008**, *36*, W465-W469; 'Phylogeny.fr: robust phylogenetic analysis for the non-specialist'.

[431] Lapid, C., Gao, Y., **2003**, 'PrimerX' (*http://www.bioinformatics.org/primerx/*)

[432] Stothard, P., **2004**, 'The Sequence Manipulation Suite' (*http://www.bioinformatics.org/sms2/reference.html*)

[433] Stothard, P. *BioTechniques* **2000**, *28 (6)*, 1102-1104; 'The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences'.

[434] Schomburg, D., **2016**, 'BRENDA: the comprehensive enzyme information system' (*http://brenda-enzymes.info/*)

[435] Henrissat, B., **1998**, 'Carbohydrate-active enzymes database (CAZy)' (*http://www.cazy.org/*)

[436] Nakamura, Y., **2007**, 'Codon Usage Database' (*http://www.kazusa.or.jp/codon/*)

[437] National Center for Biotechnology Information, **2016**, 'NCBI (National Center for Biotechnology Information)' (*http://www.ncbi.nlm.nih.gov/*)

[438] Research Collaboratory for Structural Bioinformatics, **2016**, 'RSCB Protein Data Bank' (*http://www.rcsb.org/pdb/*)

[439] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. *Nucleic Acids Res.* **2000**, *28 (1)*, 235-242; 'The Protein Data Bank'.

[440] New England Biolabs, I.; 3.0 ed.; New England Biolabs, Inc.: Ipswich, **2014**; 'IMPACT™ Kit Instruction Manual'.

[441] Invitrogen Corporation; Thermo Fisher Scientific, Inc.: Carlsbad, **2003**; 'Mach1™ T1 Phage-Resistant Chemically Competent *E. coli*'.

[442] Stratagene; Agilent Technologies: Santa Clara, **2004**; 'XL1-Blue Competent Cells'.

[443] Zor, T., Seliger, Z. *Anal. Biochem.* **1996**, *236 (2)*, 302-308; 'Linearization of the bradford protein assay increases its sensitivity: Theoretical and experimental studies'.

[444] Murray, D. R. P. *Biochem. J.* **1930**, *24 (6)*, 1890-1896; 'The inhibition of esterases by excess substrate.'.

[445] Armarego, W. L. F., Perrin, D. D. *Purification of Laboratory Chemicals*; 4 ed.; Butterworth-Heinemann: Bath, **1996**; 'Purification of Laboratory Chemicals'

[446] Fulmer, G. R., Miller, A. J. M., Sherden, N. H., Gottlieb, H. E., Nudelman, A., Stoltz, B. M., Bercaw, J. E., Goldberg, K. I. *Organometallics* **2010**, *29 (9)*, 2176-2179; 'NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist'.

[447] Mestrelab Research, **2012**, 'MestReNova' (v 8.0.1-10878)

[448] PerkinElmer, **2015**, 'ChemDraw Professional' (v 15.0.0.106)

[449] Osborn, H. M. I. *Carbohydrates*; Academic Press: Reading, **2003**; 'Carbohydrates'

[450] Michihata, N., Kaneko, Y., Kasai, Y., Tanigawa, K., Hirokane, T., Higasa, S., Yamada, H. *J. Org. Chem.* **2013**, *78 (9)*, 4319-4328; 'High-Yield Total Synthesis of (−)-Strictinin through Intramolecular Coupling of Gallates'.

[451] Wei, X., Ma, Y., Wu, Q., Zhang, J., Cai, Z., Lu, M. *Molecules* **2015**, *20 (12)*, 19789; 'An Improved Helferich Method for the α/β-Stereoselective Synthesis of 4-Methylumbelliferyl Glycosides for the Detection of Microorganisms'.

[452] Šardzík, R., Noble, G. T., Weissenborn, M. J., Martin, A., Webb, S. J., Flitsch, S. L. *Beilstein J. Org. Chem.* **2010**, *6*, 699-703; 'Preparation of aminoethyl glycosides for glycoconjugation'.

[453] Hou, S. J., Kovác, P. *Eur. J. Org. Chem.* **2008***(11)*, 1947-1952; 'A Convenient Synthesis of Furanose-Free D-Fucose Per-*O*-Acetates and a Precursor for Anthrose'.

[454] Takanashi, S., Mori, K. *Liebigs Ann-Recl* **1997***(5)*, 825-838; 'Synthesis of lurlenic acid and lurlenol, the sex pheromones of the green flagellate *Chlamydomonas*'.

[455] Pelletier, G., Zwicker, A., Allen, C. L., Schepartz, A., Miller, S. J. *J. Am. Chem. Soc.* **2016**, *138 (9)*, 3175-3182; 'Aqueous Glycosylation of Unprotected Sucrose Employing Glycosyl Fluorides in the Presence of Calcium Ion and Trimethylamine'.

[456] Jünnemann, J., Thiem, J., Pedersen, C. *Carbohydr. Res.* **1993**, *249 (1)*, 91-94; 'Facile synthesis of acetylated glycosyl fluorides derived from di- and tri-saccharides'.

[457] Burke, H. M., Gunnlaugsson, T., Scanlan, E. M. *Org. Biomol. Chem.* **2016**, *14 (38)*, 9133-9145; 'Glycosylated lanthanide cyclen complexes as luminescent probes for monitoring glycosidase enzyme activity'.

[458] Temeriusz, A., Gubica, T., Rogowska, P., Paradowska, K., Cyranski, M. K. *Carbohydr. Res.* **2005**, *340 (6)*, 1175-1184; 'Crystal structure and solid state ^{13}C NMR analysis of nitrophenyl 2,3,4,6-tetra-*O*-acetyl-β-D-gluco- and D-galactopyranosides'.

[459] Gubica, T., Temeriusz, A., Pawłowski, P., Stroka, J. *J. Phys. Org. Chem.* **2010**, *23 (9)*, 853-858; 'Molecular structure of nitrophenyl *O*-glycosides in relation to their redox potentials'.

[460] Gubica, T., Stepien, D. K., Ostrowski, A., Pisklak, D. M., Temeriusz, A., Głowacka, E., Paradowska, K., Cyranski, M. K. *J. Mol. Struct.* **2012**, *1007*, 227-234; 'Crystal and molecular structure of nitrophenyl 2,3,4-tri-*O*-acetyl-β-D-xylopyranosides'.

[461] Kumar, V., Talisman, I. J., Malhotra, S. V. *Eur. J. Org. Chem.* **2010**, *2010 (18)*, 3377-3381; 'Application of Halide Molten Salts as Novel Reaction Media for O-Glycosidic Bond Formation'.

[462] Patil, P. R., Ravindranathan Kartha, K. P. *J. Carbohydr. Chem.* **2008**, *27 (7)*, 411-419; 'Solvent-Free Mechanochemical Synthesis of Aryl Glycosides'.

[463] Zhou, K., Luan, H.-w., Hu, Y., Ge, G.-b., Liu, X.-b., Ma, X.-c., Hou, J., Wang, X.-l., Yang, L. *J. Mol. Catal. B: Enzym.* **2012**, *80*, 48-57; 'Isolation and characterization of a novel α-glucosidase with transglycosylation activity from *Arthrobacter* sp. DL001'.

[464] Shafizadeh, F., McGinnis, G. D., Susott, R. A., Tatton, H. W. *J. Org. Chem.* **1971**, *36 (19)*, 2813-2818; 'Thermal reactions of α-D-xylopyranose and β-D-xylopyranosides'.

[465] López, R., Fernández-Mayoralas, A. *J. Org. Chem.* **1994**, *59 (4)*, 737-745; 'Enzymic β-Galactosidation of Modified Monosaccharides: Study of the Enzyme Selectivity for the Acceptor and Its Application to the Synthesis of Disaccharides'.

[466] Matta, K. L., Rana, S. S., Piskorz, C. F., Barlow, J. J. *Carbohydr. Res.* **1983**, *112 (2)*, 213-220; 'Synthesis of *p*-nitrophenyl 2-*O*-α- and -β-L-fucopyranosyl-β-D-fucopyranoside'.

[467] Zhang, Q., Kang, X., Long, L., Zhu, L., Chai, Y. *Synthesis* **2015**, *47 (01)*, 55-64; 'Mild and Selective Deprotection of tert-Butyl(dimethyl)silyl Ethers with Catalytic Amounts of Sodium Tetrachloroaurate(III) Dihydrate'.

[468] Andersen, S. M., Heuckendorff, M., Jensen, H. H. *Org. Lett.* **2015**, *17 (4)*, 944-947; '3-(Dimethylamino)-1-propylamine: A Cheap and Versatile Reagent for Removal of Byproducts in Carbohydrate Chemistry'.

[469] Joosten, J. A. F., Evers, B., van Summeren, R. P., Kamerling, J. P., Vliegenthart, J. F. G. *Eur. J. Org. Chem.* **2003***(18)*, 3569-3586; 'Synthesis of β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→2)-α-D-Man*p*(1→*O*)$(CH_2)_7CH_3$ Mimics to Explore the Substrate Specificity of Sialyltransferases and *trans*-Sialidases'.

[470] Kobayashi, S., Kuroda, H., Ohtsuka, Y., Kashihara, T., Masuyama, A., Watanabe, K. *Tetrahedron* **2013**, *69 (10)*, 2251-2259; 'Evaluation of cyclopentyl methyl ether (CPME) as a solvent for radical reactions'.

[471] Guchhait, G., Misra, A. K. *J. Carbohydr. Chem.* **2010**, *29 (1)*, 1-9; 'Synthesis of the Trisaccharide Repeat of the *O*-Antigen of *Rhanella aquatilis* 1-95'.

[472] Nashed, E. M., Glaudemans, C. P. J. *J. Org. Chem.* **1987**, *52 (23)*, 5255-5260; 'Selective silylation of β-D-galactosides. A new approach to the synthesis of (1→6)-β-D-galactopyranooligosaccharides'.

[473] Bojarová, P., Petrásková, L., Ferrandi, E. E., Monti, D., Pelantová, H., Kuzma, M., Simerská, P., Kren, V. *Adv. Synth. Catal.* **2007**, *349 (8-9)*, 1514-1520; 'Glycosyl Azides – An Alternative Way to Disaccharides'.

[474] Ellington, J. C., Arnett, E. M. *J. Am. Chem. Soc.* **1988**, *110 (23)*, 7778-7785; 'Kinetics and thermodynamics of phenolate silylation and alkylation'.

[475] Khalafi-Nezhad, A., Fareghi Alamdari, R., Zekri, N. *Tetrahedron* **2000**, *56 (38)*, 7503-7506; 'Efficient and Selective Protection of Alcohols and Phenols with Triisopropylsilyl Chloride/Imidazole Using Microwave Irradiation'.

[476] Schöne, N. Doctoral, Universität Stuttgart, **2007**; 'Enantiomerenreine Allylboronsäureester - Synthese und Anwendung'.

[477] Myles, L., Gathergood, N., Connon, S. J. *Chem. Commun.* **2013**, *49 (46)*, 5316-5318; 'The catalytic versatility of low toxicity dialkyltriazolium salts: *in situ* modification facilitates diametrically opposed catalysis modes in one pot'.

[478] Brown, L. J., Bouvet, D. R., Champion, S., Gibson, A. M., Hu, Y. L., Jackson, A., Khan, I., Ma, N. C., Millot, N., Wadsworth, H., Brown, R. C. D. *Angew. Chem. Intl. Ed.* **2007**, *46 (6)*, 941-944; 'A Solid-Phase Route to F^{18}-Labeled Tracers, Exemplified by the Synthesis of [F^{18}]2-Fluoro-2-deoxy-D-glucose'.

[479] Murphy, P. V., O'Brien, J. L., Smith III, A. B. *Carbohydr. Res.* **2001**, *334 (4)*, 327-335; 'Stereospecific synthesis of β-D-allopyranosides by dihydroxylation of β-D-erythro-2,3-dideoxyhex-2-enopyranosides'.

[480] Khan, S. H., Abbas, S. A., Matta, K. L. *Carbohydr. Res.* **1989**, *193*, 125-139; 'Synthesis of Some Oligosaccharides Containing the *O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-(1→2)-*O*-α-D-mannopyranosyl Unit. Potential Substrates for UDP-GlcNAc-α-D-mannopyranosyl-(1→6)-*N*-Acetyl-β-D-glucosaminyl-Transferase (GnT-V)'.

[481] Sheu, C., Kang, P., Khan, S., Foote, C. S. *J. Am. Chem. Soc.* **2002**, *124 (15)*, 3905-3913; 'Low-temperature photosensitized oxidation of a guanosine derivative and formation of an imidazole ring-opened product'.

[482] Reese, C. B., Stewart, J. C. M., van Boom, J. H., de Leeuw, H. P. M., Nagel, J., de Rooy, J. F. M. *J. Chem. Soc., Perkin Trans. 1* **1975***(10)*, 934-942; 'The synthesis of oligoribonucleotides. Part XI. Preparation of ribonucleoside 2'-acetal 3'-esters by selective deacylation'.

Appendix A - Molecular Biology

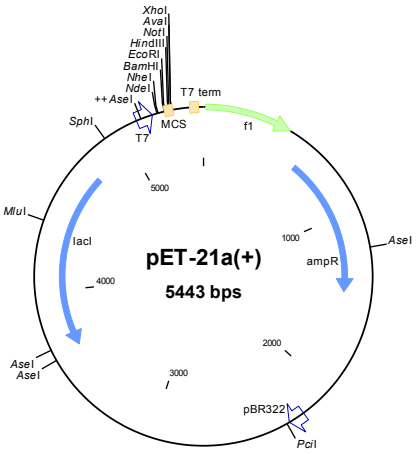
Original images of SDS-PAGE and agarose electrophoresis gels are available on request.

A-1 Vector List

pET21a(+)

This vector was obtained from Novagen.

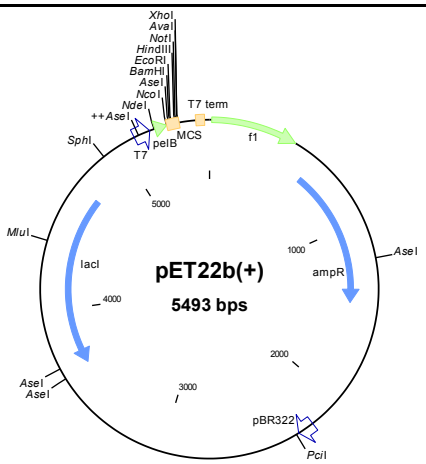
f1 ORI	12-467
amp^R	599-1456
pBR322 ORI	2217
lacI	4730-3651
T7 promoter	5133-5117
T7 transcription start	5134
T7-tag	5237-5205
MCS	5241-5286
His-tag	5304-5287
T7 terminator	5372-5418



pET22b(+)

This vector was obtained from Novagen.

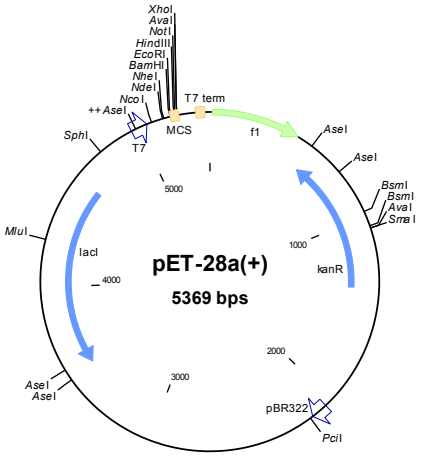
f1 ORI	12-467
amp^R	599-1456
pBR322 ORI	2217
lacI	4700-3651
T7 promoter	5133-5117
T7 transcription start	5134
pelB leader sequence	5205-5270
MCS	5269-5336
His-tag	5354-5337
T7 terminator	5422-5468



pET28a(+)

This vector was obtained from Novagen.

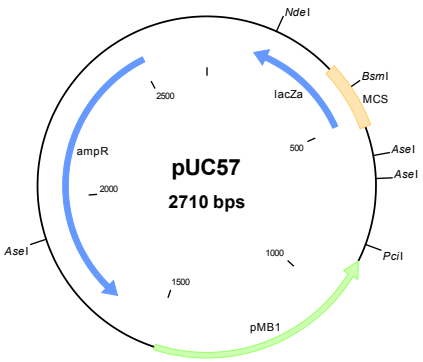
f1 ORI	12-467
kan^R	1375-563
pBR322 ORI	2084
lacI	4597-3518
T7 promoter	5000-4984
T7 transcription start	5001
His-tag	5083-5100
T7-tag	5131-5163
MCS	5167-5212
His-tag	5213-5230
T7 terminator	5298-5344



pUC57

This vector was supplied as a part of a synthetic construct with ordered genes from GenScript.

MCS	352-524
lacZα	493-146
pMB1	1490-876
amp^R	2510-1650



A-2 Plasmid Cards and Sequences

The following sequences are marked for *non-coding nucleotides*, **mutation sites**, and **modifications**.

A-2.1 LacZ: β-galactosidase from *Escherichia coli* BL21 (DE3)

lacZ-His(x6)

atgaccatgattacggattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttgcagcacatccccct
ttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctgaatggcgaatggcgctttgcctggtttccgg
caccagaagcggtgccggaaagctggctggagtgcgatcttcctgaggccgatactgtcgtcgtcccctcaaactggcagatgcacggttacgatg
cgcccatctacaccaacgtgacctatcccattacggtcaatccgccgtttgttcccacggagaatccgacgggttgttactcgctcacatttaatgttg
atgaaagctggctacaggaaggccagacgcgaattatttttgatggcgttaactcggcgtttcatctgtggtgcaacgggcgctgggtcggttacggc
caggacagtcgtttgccgtctgaatttgacctgagcgcatttttacgcgccggagaaaaccgcctcgcggtgatggtgctgcgctggagtgacggca
gttatctggaagatcaggatatgtggcggatgagcggcatttttccgtgacgtctcgttgctgcataaaccgactacacaaatcagcgatttccatgttg
ccactcgctttaatgatgatttcagccgcgctgtactggaggctgaagttcagatgtgcggcgagttgcgtgactacctacgggtaacagtttctttatg
gcagggtgaaacgcaggtcgccagcggcaccgcgcctttcggcggtgaaattatcgatgagcgtggtggttatgccgatcgcgtcacactacgtct
gaacgtcgaaaacccgaaactgtggagcgccgaaatcccgaatctctatcgtgcggtggttgaactgcacaccgccgacggcacgctgattgaag
cagaagcctgcgatgtcggtttccgcgaggtgcggattgaaaatggtctgctgctgctgaacggcaagccgttgctgattcgaggcgttaaccgtca
cgagcatcatcctctgcatggtcaggtcatggatgagcagacgatggtgcaggatatcctgctgatgaagcagaacaactttaacgccgtgcgctgt
tcgcattatccgaaccatccgctgtggtacacgctgtgcgaccgctacggcctgtatgtggtggatgaagccaatattgaaacccacggcatggtgc
caatgaatcgtctgaccgatgatccgcgctggctaccggcgatgagcgaacgcgtaacgcgaatggtgcagcgcgatcgtaatcacccgagtgtg
atcatctggtcgctggggaatgaatcaggccacggcgctaatcacgacgcgctgtatcgctggatcaaatctgtcgatccttcccgcccggtgcagt
atgaaggcggcggagccgacaccacggccaccgatattatttgcccgatgtacgcgcgcgtggatgaagaccagcccttcccggctgtgccgaaa
tggtccatcaaaaaatggctttcgctacctggagagacgcgcccgctgatcctttgc**gaa**tacgcccacgcgatgggtaacagtcttggcggtttcg
ctaaatactggcaggcgtttcgtcagtatccccgtttacagggcggcttcgtctgggactgggtggatcagtcgctgattaaatatgatgaaaacggc
aacccgtggtcggcttacggcggtgattttggcgatacgccgaacgatcgccagttctgtatgaacggtctggtctttgccgaccgcacgccgcatc
cagcgctgacggaagcaaaacaccagcagcagtttttccagttccgtttatccgggcaaaccatcgaagtgaccagcgaatacctgttccgtcatag
cgataacgagctcctgcactggatggtggcgctggatggtaagccgctggcaagcggtgaagtgcctctggatgtcgctccacaaggtaaacagtt
gattgaactgcctgaactaccgcagccggagagcgccgggcaactctggctcacagtacgcgtagtgcaaccgaacgcgaccgcatggtcagaa
gccggacacatcagcgcctggcagcagtggcgtctggctgaaaacctcagcgtgacactccccgccgcgtcccacgccatcccgcatctgaccac
cagcgaaatggatttttgcatcgagctgggtaataagcgttggcaatttaaccgccagtcaggctttctttcacagatgtggattggcgataaaaaaca
actgctgacgccgctgcgcgatcagttcacccgtgcaccgctggataacgacatt**ggc**gtaagtgaagcgacccgcattgaccctaacgcctgggt
cgaacgctggaaggcggcgggccattaccaggccgaagcagcgttgttgcagtgcacggcagatacacttgctgatgcggtgctgattacgaccg
ctcacgcgtggcagcatcaggggaaaaccttatttatcagccggaaaacctaccggattgatggtagtggtcaaatggcgattaccgttgatgttgaa
gtggcgagcgatacaccgcatccggcgcggattggcctgaactgccagctggcgcaggtagcagagcgggtaaactggctcggattagggccgc
aagaaaactatcccgaccgccttactgccgcctgttttgaccgctgggatctgccattgtcagacatgtataccccgtacgtcttcccgagcgaaaac
ggtctgcgctgcgggacgcgcgaattgaattatggcccacaccagtggcgcggcgacttccagttcaacatcagccgctacagtcaacagcaactg

atggaaaccagccatcgccatctgctgcacgcggaagaaggcacatggctgaatatcgacggtttccatatggggattggtggcgacgactcctgg
agcccgtcagtatcggcggaattccagctgagcgccggtcgctaccattaccagttggtctggtgtcaaaaactcgagcaccaccaccaccaccact
ga

LacZ-His(x6)

MTMITDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLRSLNGEWRFA
WFPAPEAVPESWLECDLPEADTVVVPSNWQMHGYDAPIYTNVTYPITVNPPFVPTENPTGCYS
LTFNVDESWLQEGQTRIIFDGVNSAFHLWCNGRWVGYGQDSRLPSEFDLSAFLRAGENRLAVM
VLRWSDGSYLEDQDMWRMSGIFRDVSLLHKPTTQISDFHVATRFNDDFSRAVLEAEVQMCGE
LRDYLRVTVSLWQGETQVASGTAPFGGEIIDERGGYADRVTLRLNVENPKLWSAEIPNLYRAV
VELHTADGTLIEAEACDVGFREVRIENGLLLLNGKPLLIRGVNRHEHHPLHGQVMDEQTMVQ
DILLMKQNNFNAVRCSHYPNHPLWYTLCDRYGLYVVDEANIETHGMVPMNRLTDDPRWLPA
MSERVTRMVQRDRNHPSVIIWSLGNESGHGANHDALYRWIKSVDPSRPVQYEGGGADTTATD
IICPMYARVDEDQPFPAVPKWSIKKWLSLPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYP
RLQGGFVWDWVDQSLIKYDENGNPWSAYGGDFGDTPNDRQFCMNGLVFADRTPHPALTEA
KHQQQFFQFRLSGQTIEVTSEYLFRHSDNELLHWMVALDGKPLASGEVPLDVAPQGKQLIELPE
LPQPESAGQLWLTVRVVQPNATAWSEAGHISAWQQWRLAENLSVTLPAASHAIPHLTTSEMD
FCIELGNKRWQFNRQSGFLSQMWIGDKKQLLTPLRDQFTRAPLDNDIGVSEATRIDPNAWVER
WKAAGHYQAEAALLQCTADTLADAVLITTAHAWQHQGKTLFISRKTYRIDGSGQMAITVDVE
VASDTPHPARIGLNCQLAQVAERVNWLGLGPQENYPDRLTAACFDRWDLPLSDMYTPYVFPS
ENGLRCGTRELNYGPHQWRGDFQFNISRYSQQQLMETSHRHLLHAEEGTWLNIDGFHMGIGG
DDSWSPSVSAEFQLSAGRYHYQLVWCQKLEHHHHHH

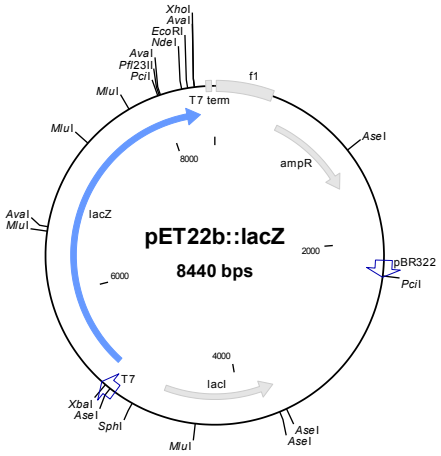
Mutations

The following *lacZ* mutants were generated using the codons shown with the mutations being performed on the vector shown.

Mutant	Codon (5'→3')	Original Vector
E537S	agt	pET28a::*lacZ* WT
E537S/G794D	gac	pET28a::*lacZ* E537S

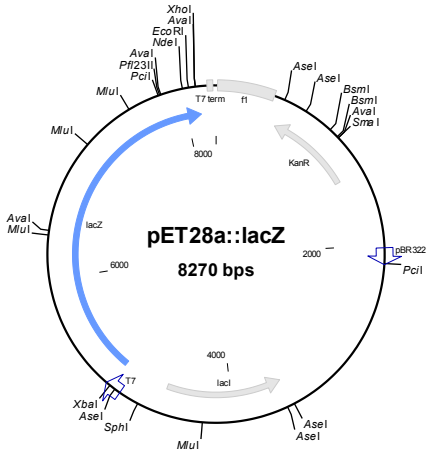
pET22b::*lacZ*

f1 ORI	12-467
amp^R	599-1456
pBR322 ORI	2217
lacI	4700-3651
T7 promoter	5133-5117
T7 transcription start	5134
lacZ-His(x6)	5206-8304
T7 terminator	8369-8415



pET28a::*lacZ*

f1 ORI	12-467
kan^R	1375-563
pBR322 ORI	2084
lacI	4597-3518
T7 promoter	5000-4984
T7 transcription start	5001
lacZ-His(x6)	5036-8134
T7 terminator	8245-8199



A-2.2 Lac4: β-galactosidase from *Kluyveromyces lactis*

lac4 (optimised gene)

*cat*atgagctgccttattccggagaatttacggaaccccaaaaaagttcacgaaaatcgtctgccgacccgggcatactactatgatcaagatattttt
gaaagcctaaatgggccgtgggcatttgctctgtttgatgctccgcttgacgcacccgatgcaaaaaatttagactgggaaactgctaaaaaatggtc
gactattagcgtaccgtctcattgggaacttcaagaagactggaaatacggtaaaccgatttacactaacgtacaatacccgatcccgatcgacatcc
cgaatccgcccaccgtaaatccgaccggtgtttatgcacgtacctttgaattagattccaaatccattgagtcctttgagcaccgtctgcgttttgaggg
tgtagacaattgttacgagctttatgttaatggtcagtatgtaggttttaataaagggtctagaaacggggcagaatttgatatccagaaatacgttagc
gaggggggaaaacttagtagtagtaaaagtttttaaatggtctgattctacctatatcgaggaccaggatcagtggtggctaagcggtatttaccgtgac
gttagcttacttaaactgccgaaaaaagctcatattgaagacgttcgggtaaccacaacctttgtagacagccaatatcaagatgctgagcttagcgt
aaaagttgatgtacaaggtagcagctatgatcacatcaattttacactttacgaaccggaagatgggagcaaagtttacgatgcttcgagcctgctga
acgaggagaatgggaacactacctttagcaccaaagaatttatttctttttctactaaaaaaaacgaagaaacagcatttaaaatcaacgtaaaagctc
cggaacattggactgctgaaaatccgaccctgtacaaataccaactggatttaattgggagcgatgggagtgtaattcagagcattaaacaccatgtt
ggttttcgtcaggtagagctgaaagacggtaacattaccgttaatgggaaagacattctatttcgtggtgtaaaccgtcatgatcaccatccgcggttt
ggtcgtgcagtaccgttagattttgttgttcgggacctgattcttatgaaaaaatttaacatcaatgcagttagaaactcccattatccgaaccatccga
aagtatatgacctatttgataaactagggttttgggtaattgacgaggctgatcttgaaacccacggtgttcaggagccgtttaatagacatactaacct
ggaggcagaatatccggataccaaaaataaactatacgatgttaatgctcattacttaagcgataatccggagtacgaggtagcttacttagaccgtg
catctcagcttgtacttcgtgatgtaaatcatccgtccattattatctggtctctgggtaacgaagcatgttatgggcgtaaccacaaagctatgtacaa
attaattaaacagctggatccgactcgtcttgtacattatgagggtgacctgaacgcactgagtgctgatatctttagttttatgtacccgacatttgaaa
ttatggaacggtggcggaaaaaccacaccgatgaaaatggtaaatttgaaaaaccgctgatcctgtgt**gag**tacgggcatgccatgggtaacggtc
cggggagcctgaaagaatatcaggagctgttttacaaagagaaattttaccagggtgggtttatctgggaatgggctaatcacggtattgaatttgaa
gatgttagtaccgctgatggtaaactgcataaagcctatgcctatggtggtgactttaaagaagaggttcatgacggggtatttatcatggatggtctgt
gtaacagtgagcataatccgacccccgggcttgtagagtataaaaaagttattgaacccgttcatattaaaattgctcacgggagcgtaacaatcaca
aataaacacgactttattactacagaccacttactgtttatcgacaaagacactgggaaaacaatcgacgttccgagcttaaaaccggaagaaagcg
ttaccattccgagcgatacaacctatgttgttgctgtactgaaagatgatgcaggtgttcttaaagctggtcatgaaattgcttgggggcaggcagaac
ttccgctgaaagtacccgattttgttacagagacagctgaaaaagcagctaaaatcaacgacggtaaaagatatgtaagcgttgaatctagtgggct
gcattttatcctggacaaactgctgggtaaaattgaatcgcttaaagtaaaaggtaaagaaatttcttcgaaatttgagggtagcagcatcacctttgg
cgtccgccgactaataatgatgaaccgcgggactttaaaaactggaaaaaatacaatattgatttaatgaaacagaacatccacggggtaagtgtag
aaaaaggtagcaatggtagccttgcagtagtaactgttaacagcagaatatctccggttgtattttactatgggtttgagaccgttcaaaaatacactat
ctttgcaaacaaaataaacctgaacaccagcatgaaacttaccggggaatatcaaccgccggattttccgcgtgttgggtacgaattttggcttgggg
atagttatgaaagctttgaatggttaggtagggggcccggggaaagctatcccgataaaaaagaaagccagcgttttggtctttacgattctaaagat
gtagaggaatttgtatatgactatccgcaggaaaatgggaatcatacagatactcactttctgaacatcaaatttgaaggtgctgggaaactttccatct
ttcagaaagagaaaccgtttaactttaaaattagcgacgaatacggggttgatgaagcagctcacgcatgtgacgttaaacgttacgggcgtcactat
cttcggctggaccatgctatccacggtgttggttcggaagcttgcgggccggcagttctagaccaataccgtctgaaagcacaggattttaactttga
gtttgatctagcatttgaa*ctcgag*

lac4-His(x6)

atgagctgccttattccggagaatttacggaaccccaaaaaagttcacgaaaatcgtctgccgacccgggcatactactatgatcaagatatttttgaa
agcctaaatgggccgtgggcatttgctctgtttgatgctccgcttgacgcacccgatgcaaaaaatttagactgggaaactgctaaaaaatggtcgac
tattagcgtaccgtctcattgggaacttcaagaagactggaaatacggtaaaccgatttacactaacgtacaatacccgatcccgatcgacatcccga
atccgcccaccgtaaatccgaccggtgtttatgcacgtacctttgaattagattccaaatccattgagtcctttgagcaccgtctgcgttttgagggtgt
agacaattgttacgagctttatgttaatggtcagtatgtaggttttaataaagggtctagaaacggggcagaatttgatatccagaaatacgttagcga
gggggaaaacttagtagtagtaaaagtttttaaatggtctgattctacctatatcgaggaccaggatcagtggtggctaagcggtatttaccgtgacgt
tagcttacttaaactgccgaaaaaagctcatattgaagacgttcgggtaaccacaacctttgtagacagccaatatcaagatgctgagcttagcgtaa
aagttgatgtacaaggtagcagctatgatcacatcaattttacactttacgaaccggaagatgggagcaaagtttacgatgcttcgagcctgctgaac
gaggagaatgggaacactacctttagcaccaaagaatttatttctttttctactaaaaaaaacgaagaaacagcatttaaaatcaacgtaaaagctcc
ggaacattggactgctgaaaatccgaccctgtacaaataccaactggatttaattgggagcgatgggagtgtaattcagagcattaaacaccatgttg
gttttcgtcaggtagagctgaaagacggtaacattaccgttaatgggaaagacattctatttcgtggtgtaaaccgtcatgatcaccatccgcggtttg
gtcgtgcagtaccgttagattttgttgttcgggacctgattcttatgaaaaaatttaacatcaatgcagttagaaactcccattatccgaaccatccgaa
agtatatgacctatttgataaactagggttttgggtaattgacgaggctgatcttgaaacccacggtgttcaggagccgtttaatagacatactaacctg
gaggcagaatatccggataccaaaaataaactatacgatgttaatgctcattacttaagcgataatccggagtacgaggtagcttacttagaccgtgc
atctcagcttgtacttcgtgatgtaaatcatccgtccattattatctggtctctgggtaacgaagcatgttatgggcgtaaccacaaagctatgtacaaat
taattaaacagctggatccgactcgtcttgtacattatgagggtgacctgaacgcactgagtgctgatatctttagttttatgtacccgacatttgaaatt
atggaacggtggcggaaaaaccacaccgatgaaaatggtaaatttgaaaaaccgctgatcctgtgtgagtacgggcatgccatgggtaacggtccg
gggagcctgaaagaatatcaggagctgttttacaaagagaaattttaccagggtgggtttatctgggaatgggctaatcacggtattgaatttgaagat
gttagtaccgctgatggtaaactgcataaagcctatgcctatggtggtgactttaaagaagaggttcatgacggggtatttatcatggatggtctgtgta
acagtgagcataatccgacccccgggcttgtagagtataaaaaagttattgaacccgttcatattaaaattgctcacgggagcgtaacaatcacaaat
aaacacgactttattactacagaccacttactgtttatcgacaaagacactgggaaaacaatcgacgttccgagcttaaaaccggaagaaagcgtta
ccattccgagcgatacaacctatgttgttgctgtactgaaagatgatgcaggtgttcttaaagctggtcatgaaattgcttgggggcaggcagaacttc
cgctgaaagtacccgattttgttacagagacagctgaaaaagcagctaaaatcaacgacggtaaaagatatgtaagcgttgaatctagtgggctgca
ttttatcctggacaaactgctgggtaaaattgaatcgcttaaagtaaaaggtaaagaaatttcttcgaaatttgagggtagcagcatcacctttggcgt
ccgccgactaataatgatgaaccgcgggactttaaaaactggaaaaaatacaatattgatttaatgaaacagaacatccacggggtaagtgtagaaa
aaggtagcaatggtagccttgcagtagtaactgttaacagcagaatatctccggttgtattttactatgggtttgagaccgttcaaaaatacactatcttt
gcaaacaaaataaacctgaacaccagcatgaaacttaccggggaatatcaaccgccggattttccgcgtgttgggtacgaattttggcttggggata
gttatgaaagctttgaatggttaggtaggggccccggggaaagctatcccgataaaaaagaaagccagcgttttggtctttacgattctaaagatgta
gaggaatttgtatatgactatccgcaggaaaatgggaatcatacagatactcactttctgaacatcaaatttgaaggtgctgggaaactttccatctttc
agaaagagaaaccgtttaactttaaaattagcgacgaatacggggttgatgaagcagctcacgcatgtgacgttaaacgttacgggcgtcactatctt
cggctggaccatgctatccacggtgttggttcggaagcttgcgggccggcagttctagaccaataccgtctgaaagcacaggattttaactttgagttt
gatctagcatttgaactcgagcaccaccaccaccaccactga

Lac4-His(x6)

MSCLIPENLRNPKKVHENRLPTRAYYYDQDIFESLNGPWAFALFDAPLDAPDAKNLDWETAKK
WSTISVPSHWELQEDWKYGKPIYTNVQYPIPIDIPNPPTVNPTGVYARTFELDSKSIESFEHRLRF
EGVDNCYELYVNGQYVGFNKGSRNGAEFDIQKYVSEGENLVVVKVFKWSDSTYIEDQDQWW
LSGIYRDVSLLKLPKKAHIEDVRVTTTFVDSQYQDAELSVKVDVQGSSYDHINFTLYEPEDGSKV
YDASSLLNEENGNTTFSTKEFISFSTKKNEETAFKINVKAPEHWTAENPTLYKYQLDLIGSDGSVI
QSIKHHVGFRQVELKDGNITVNGKDILFRGVNRHDHHPRFGRAVPLDFVVRDLILMKKFNINA
VRNSHYPNHPKVYDLFDKLGFWVIDEADLETHGVQEPFNRHTNLEAEYPDTKNKLYDVNAH
YLSDNPEYEVAYLDRASQLVLRDVNHPSIIIWSLGNEACYGRNHKAMYKLIKQLDPTRLVHYEG
DLNALSADIFSFMYPTFEIMERWRKNHTDENGKFEKPLILC**E**YGHAMGNGPGSLKEYQELFYKE
KFYQGGFIWEWANHGIEFEDVSTADGKLHKAYAYGGDFKEEVHDGVFIMDGLCNSEHNPTPG
LVEYKKVIEPVHIKIAHGSVTITNKHDFITTDHLLFIDKDTGKTIDVPSLKPEESVTIPSDTTYVVA
VLKDDAGVLKAGHEIAWGQAELPLKVPDFVTETAEKAAKINDGKRYVSVESSGLHFILDKLLG
KIESLKVKGKEISSKFEGSSITFWRPPTNNDEPRDFKNWKKYNIDLMKQNIHGVSVEKGSNGSLA
VVTVNSRISPVVFYYGFETVQKYTIFANKINLNTSMKLTGEYQPPDFPRVGYEFWLGDSYESFE
WLGRGPGESYPDKKESQRFGLYDSKDVEEFVYDYPQENGNHTDTHFLNIKFEGAGKLSIFQKEK
PFNFKISDEYGVDEAAHACDVKRYGRHYLRLDHAIHGVGSEACGPAVLDQYRLKAQDFNFEF
DLAFELE**HHHHHH**

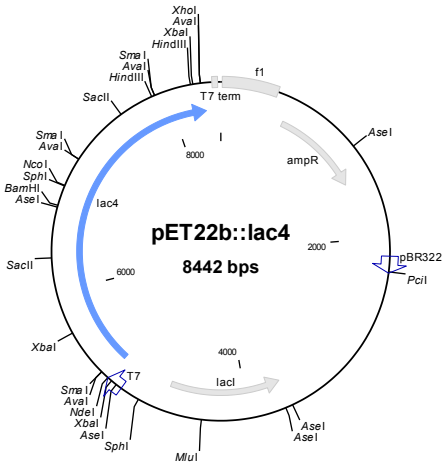
Mutations

The following *lac4* mutants were generated using the codons shown with the mutations being performed on the vector shown.

Mutant	Codon (5'→3')	Original Vector
E551G	ggc	pET22a::*lac4* WT
E551A	gct	pET22a::*lac4* WT
E551L	tta	pET22a::*lac4* WT
E551S	agc	pET22a::*lac4* WT
E551Q	cag	pET22a::*lac4* WT

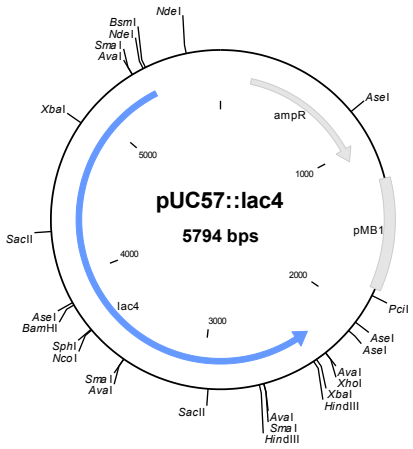
pET22b::*lac4*

f1 ORI	12-467
amp^R	599-1456
pBR322 ORI	2217
lacI	4700-3651
T7 promoter	5133-5117
T7 transcription start	5134
lac4-His(x6)	5205-8306
T7 terminator	8371-8417



pUC57::*lac4*

amp^R	599-1456
pMB1 ORI	1221-1835
lac4	5360-2286



A-2.3 LacA: β-galactosidase from *Aspergillus oryzae*

lacA (optimised gene)

*cat*atgaaattattaagcgtggcagcagtggcattattagctgctcaggctgctggcgcaagcattaaacatcggttaaacggctttactattttagag
catccggatccggctaagcgggatttattacaagatatcgtgacgtgggatgataagagcttatttattaacggagaaagaatcatgctatttagcgga
gaggtgcatccgtttcggttaccggtaccgagtttatggttagatatttttcataaaattcgggcattaggctttaactgtgtaagcttttacatcgattgg
gcattattagaaggaaaaccgggcgattaccgggctgagggcattttcgcattagagccgtttttcgatgctgcaaaagaggctggcatctacttaatt
gcacgcccgggcagttacattaacgcagaagtttctggcggcggctttccgggatggttacagagagttaacggcacgttacgcagtagcgatgaa
ccgtttttaaaggcaacggataactacattgcaaacgcagcagcagcagtggctaaagcacaaattactaacggagggccggtaatcttataccagc
cggagaacgagtacagcggcggctgttgtggcgttaagtacccggatgctgattacatgcagtacgtgatggatcaggcacggaaagcagatatcg
tggtaccgtttattagcaacgatgcatctccgagcgggcataacgcaccgggatccggaactagcgcagtggatatctacggccatgatagctaccc
gttaggcttcgattgtgctaacccgagcgtatggccggaaggcaagttaccggataactttcgcactttacatttagaacagagcccgtcgacgccgt
actctttatta**gaa**tttcaagctggcgcattcgatccgtggggcggaccgggcttcgagaagtgttacgcattagtgaaccatgagtttagtcgggtgt
tttacagaaacgatttatcctttggagtgagcaccttcaacctatacatgacgtttggcggaacgaactggggcaacttaggacatccgggcggatac
acgagctacgattacggaagtccgataacggagactcggaacgtgactcgggaaaaatacagcgatataaaattattagcaaacttcgttaaggcta
gtccgagctacttaaccgcaacgccgcggaacttaacgacgggcgtgtacacggatacgagcgatttagcagttaccccgctaatcggcgattccc
cgggctctttttttgtggttcggcatactgattacagcagccaagaatctaccagttacaagttaaaattaccgacctccgcaggcaacttaacgatcc
cgcagttagaaggcacgctctccttaaacggacgggattctaagatccatgtggtggattacaacgtgagcggaactaacatcatttacagtacggc
agaggtttttacctggaaaaaattcgatggcaacaaagttttagtgctatacggcggaccgaaagagcatcatgagttagcaatcgcaagcaaatcta
acgtgaccattattgagggcagtgatagcggaatcgtttctactagaaaaggcagcagcgtgattatcggctgggatgttagcagcacgcggcgcat
tgtgcaagttggcgatttacgggtgtttttattagatagaaacagcgcatacaactactgggttccggagttaccgacggagggcacgagcccgggg
tttagcacgagtaaaactaccgcaagcagcatcatcgtgaaagcaggctacttattacggggcgcacatttagatggcgcagatttacatttaacggc
agattttaacgcaaccaccccgatcgaggtgattggcgcaccgacgggcgcaaaaaacttatttgtgaacggcgagaaagcaagccatacggttga
taaaaacggcatttggagctccgaagttaaatacgctgcaccggaaattaaattaccgggcttaaaagatttagattggaaatacttagatactttacc
ggagatcaaaagcagctacgatgatagtgcatgggtgagtgctgatttaccgaaaacgaaaaacacgcatcggccgttagatacgccgacgagtct
ctacagcagcgattacggctttcatacgggctacttaatttacagaggccattttgtggcaaacggcaaagagagcgagttcttcatccgcacgcaag
gcggcagcgcttttggatccagcgtatggttaaacgaaacttacttaggcagctggacgggcgcagattacgctatggatggcaacagcacctaca
aactcagccagttagaaagtggcaaaaactacgttattacggtggtgatcgataacttaggcttagatgaaaactggactgttggcgaagagaccat
gaaaaacccgcggggcatcttaagctacaaattaagcggacaagatgcaagcgctattacctggaaattaacgggcaacttaggaggagaggatt
accaggataaagtgcggggaccgttaaacgagggcggattatacgctgaacgccagggctttcatcagccgcagccgccgagcgagagctggga
aagtggctccccgttagagggcttaagtaaaccgggcattggattttacacggcacagtttgatttagatttaccgaaaggctgggatgtgccgttata
ctttaacttcggcaacaacacccaggctgcacgggcacagttatacgttaacggctaccagtacggcaaatttacgggaaacgtggggccgcagac
cagctttccggtgccggagggcattttaaactaccgcggaaccaactacgtggctttatccttatgggctttagaaagtgatggcgcaaaattaggca
gctttgagttaagctacaccaccccggtgttaaccggatacgggaacgtggaatctccggaacagccgaaatacgaacagcggaaaggagcttac
ctcgag

lacA-His(x6)

atgaaattattaagcgtggcagcagtggcattattagctgctcaggctgctggcgcaagcattaaacatcggttaaacggctttactattttagagcat
ccggatccggctaagcgggatttattacaagatatcgtgacgtgggatgataagagcttatttattaacggagaaagaatcatgctatttagcggaga
ggtgcatccgtttcggttaccggtaccgagtttatggttagatatttttcataaaattcgggcattaggctttaactgtgtaagcttttacatcgattgggc
attattagaaggaaaaccgggcgattaccgggctgagggcattttcgcattagagccgtttttcgatgctgcaaaagaggctggcatctacttaattgc
acgcccgggcagttacattaacgcagaagtttctggcggcggctttccgggatggttacagagagttaacggcacgttacgcagtagcgatgaacc
gtttttaaaggcaacggataactacattgcaaacgcagcagcagcagtggctaaagcacaaattactaacggagggccggtaatcttataccagcc
ggagaacgagtacagcggcggctgttgtggcgttaagtacccggatgctgattacatgcagtacgtgatggatcaggcacggaaagcagatatcgt
ggtaccgtttattagcaacgatgcatctccgagcgggcataacgcaccgggatccggaactagcgcagtggatatctacggccatgatagctaccc
gttaggcttcgattgtgctaacccgagcgtatggccggaaggcaagttaccggataactttcgcactttacatttagaacagagcccgtcgacgccgt
actctttattaa**gaa**tttcaagctggcgcattcgatccgtggggcggaccgggcttcgagaagtgttacgcattagtgaaccatgagtttagtcgggtgt
tttacagaaacgatttatcctttggagtgagcaccttcaacctatacatgacgtttggcggaacgaactggggcaacttaggacatccgggcggatac
acgagctacgattacggaagtccgataacggagactcggaacgtgactcgggaaaaatacagcgatataaaattattagcaaacttcgttaaggcta
gtccgagctacttaaccgcaacgccgcggaacttaacgacgggcgtgtacacggatacgagcgatttagcagttaccccgctaatcggcgattccc
cgggctctttttttgtggttcggcatactgattacagcagccaagaatctaccagttacaagttaaaattaccgacctccgcaggcaacttaacgatcc
cgcagttagaaggcacgctctccttaaacggacgggattctaagatccatgtggtggattacaacgtgagcggaactaacatcatttacagtacggc
agaggtttttacctggaaaaaattcgatggcaacaaagttttagtgctatacggcggaccgaaagagcatcatgagttagcaatcgcaagcaaatcta
acgtgaccattattgagggcagtgatagcggaatcgtttctactagaaaaggcagcagcgtgattatcggctgggatgttagcagcacgcggcgcat
tgtgcaagttggcgatttacgggtgtttttattagatagaaacagcgcatacaactactgggttccggagttaccgacggagggcacgagcccgggg
tttagcacgagtaaaactaccgcaagcagcatcatcgtgaaagcaggctacttattacggggcgcacatttagatggcgcagatttacatttaacggc
agattttaacgcaaccaccccgatcgaggtgattggcgcaccgacgggcgcaaaaaacttatttgtgaacggcgagaaagcaagccatacggttga
taaaaacggcatttggagctccgaagttaaatacgctgcaccggaaattaaattaccgggcttaaaagatttagattggaaatacttagatactttacc
ggagatcaaaagcagctacgatgatagtgcatgggtgagtgctgatttaccgaaaacgaaaaacacgcatcggccgttagatacgccgacgagtct
ctacagcagcgattacggctttcatacgggctacttaatttacagaggccattttgtggcaaacggcaaagagagcgagttcttcatccgcacgcaag
gcggcagcgcttttggatccagcgtatggttaaacgaaacttacttaggcagctggacgggcgcagattacgctatggatggcaacagcacctaca
aactcagccagttagaaagtggcaaaaactacgttattacggtggtgatcgataacttaggcttagatgaaaactggactgttggcgaagagaccat
gaaaaacccgcggggcatcttaagctacaaattaagcggacaagatgcaagcgctattacctggaaattaacgggcaacttaggaggagaggatt
accaggataaagtgcggggaccgttaaacgagggcggattatacgctgaacgccagggctttcatcagccgcagccgccgagcgagagctggga
aagtggctccccgttagagggcttaagtaaaccgggcattggattttacacggcacagtttgatttagatttaccgaaaggctgggatgtgccgttata
ctttaacttcggcaacaacacccaggctgcacgggcacagttatacgttaacggctaccagtacggcaaatttacgggaaacgtggggccgcagac
cagctttccggtgccggagggcattttaaactaccgcggaaccaactacgtggctttatccttatgggctttagaaagtgatggcgcaaaattaggca
gctttgagttaagctacaccaccccggtgttaaccggatacgggaacgtggaatctccggaacagccgaaatacgaacagcggaaaggagcttac
ctcgag**caccaccaccaccaccac**tga

LacA-His(x6)

MKLLSVAAVALLAAQAAGASIKHRLNGFTILEHPDPAKRDLLQDIVTWDDKSLFINGERIMLFS
GEVHPFRLPVPSLWLDIFHKIRALGFNCVSFYIDWALLEGKPGDYRAEGIFALEPFFDAAKEAGI
YLIARPGSYINAEVSGGGFPGWLQRVNGTLRSSDEPFLKATDNYIANAAAAVAKAQITNGGPVI
LYQPENEYSGGCCGVKYPDADYMQYVMDQARKADIVVPFISNDASPSGHNAPGSGTSAVDIY
GHDSYPLGFDCANPSVWPEGKLPDNFRTLHLEQSPSTPYSLL**E**FQAGAFDPWGGPGFEKCYALV
NHEFSRVFYRNDLSFGVSTFNLYMTFGGTNWGNLGHPGGYTSYDYGSPITETRNVTREKYSDIK
LLANFVKASPSYLTATPRNLTTGVYTDTSDLAVTPLIGDSPGSFFVVRHTDYSSQESTSYKLKLPT
SAGNLTIPQLEGTLSLNGRDSKIHVVDYNVSGTNIIYSTAEVFTWKKFDGNKVLVLYGGPKEHH
ELAIASKSNVTIIEGSDSGIVSTRKGSSVIIGWDVSSTRRIVQVGDLRVFLLDRNSAYNYWVPELPT
EGTSPGFSTSKTTASSIIVKAGYLLRGAHLDGADLHLTADFNATTPIEVIGAPTGAKNLFVNGEK
ASHTVDKNGIWSSEVKYAAPEIKLPGLKDLDWKYLDTLPEIKSSYDDSAWVSADLPKTKNTHRP
LDTPTSLYSSDYGFHTGYLIYRGHFVANGKESEFFIRTQGGSAFGSSVWLNETYLGSWTGADYA
MDGNSTYKLSQLESGKNYVITVVIDNLGLDENWTVGEETMKNPRGILSYKLSGQDASAITWKL
TGNLGGEDYQDKVRGPLNEGGLYAERQGFHQPQPPSESWESGSPLEGLSKPGIGFYTAQFDLDL
PKGWDVPLYFNFGNNTQAARAQLYVNGYQYGKFTGNVGPQTSFPVPEGILNYRGTNYVALSL
WALESDGAKLGSFELSYTTPVLTGYGNVESPEQPKYEQRKGAYLE**HHHHHH**

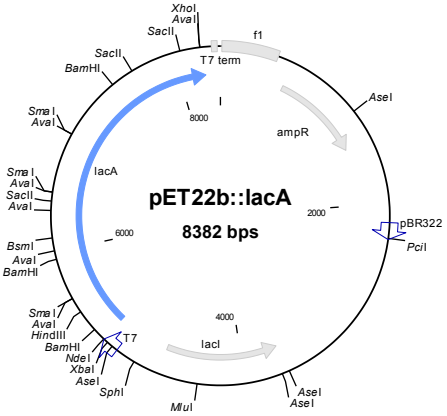
Mutations

The following *lac4* mutants were generated using the codons shown with the mutations being performed on the vector shown.

Mutant	Codon (5'→3')	Original Vector
E298G	ggt	pET22a::*lacA* WT
E298A	gca	pET22a::*lacA* WT
E298L	ctg	pET22a::*lacA* WT
E298S	agc	pET22a::*lacA* WT
E298Q	cag	pET22a::*lacA* WT

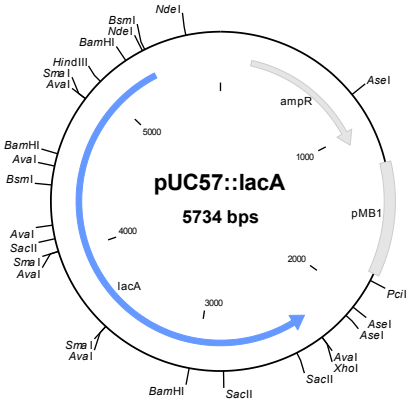
pET22b::*lacA*

f1 ORI	12-467
amp^R	599-1456
pBR322 ORI	2217
lacI	4700-3651
T7 promoter	5133-5117
T7 transcription start	5134
lacA-His(x6)	5205-8246
T7 terminator	8311-8357



pUC57::*lacA*

amp^R	599-1456
pMB1 ORI	1221-1835
lacA	5300-2286

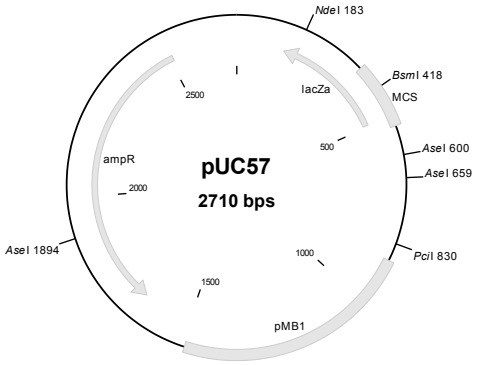


A-3 Obtained Constructs

pML5::T7

This construct was kindly provided by Dr. Anita Loeschke of the IMET and is previously published.[308]

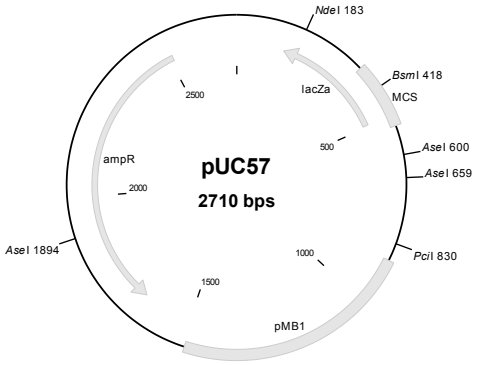
MCS	352-524
lacZα	493-146
pMB1	1490-876
amp^R	2510-1650



pBBR1MCS3::T7

This construct was kindly provided by Dr. Anita Loeschke of the IMET and is previously published.[308]

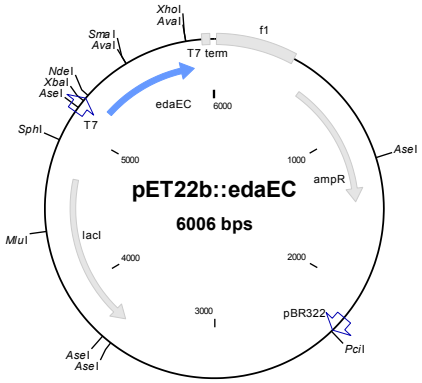
MCS	352-524
lacZα	493-146
pMB1	1490-876
amp^R	2510-1650



pET22b(+)::edaEC

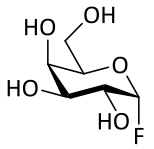
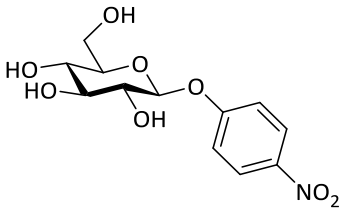
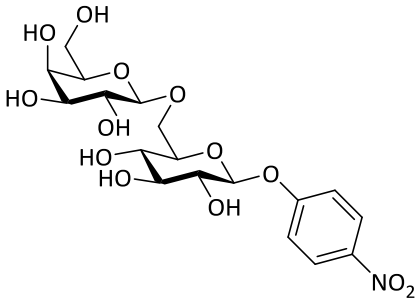
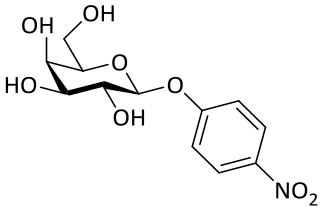
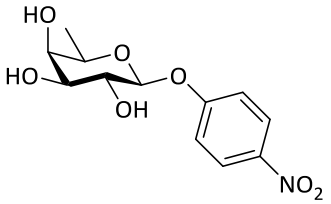
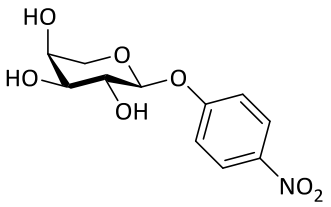
This construct was kindly provided by Dr. Thomas Classen of the IBOC and is previously published.[316]

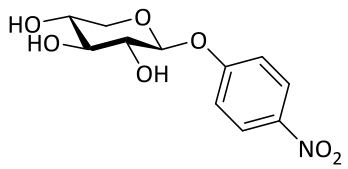
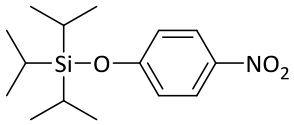
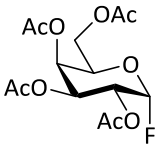
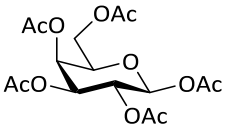
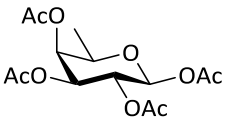
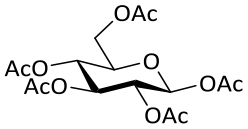
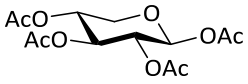
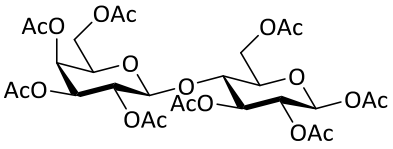
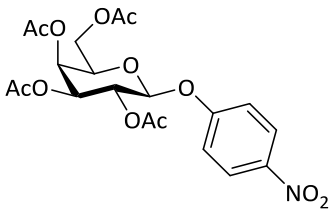
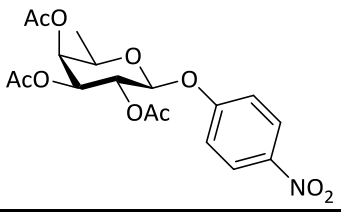
F1 ORI	12-467
amp^{R}	599-1456
pBR322 ORI	2217
lacI	4700-3651
T7 promoter	5133-5117
T7 transcription start	5134
eda^{EC}	5205-5870
T7 terminator	5935-5981

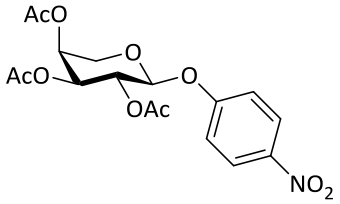
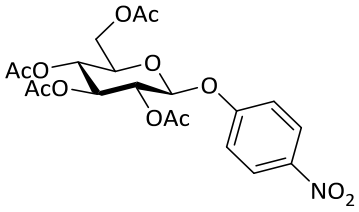
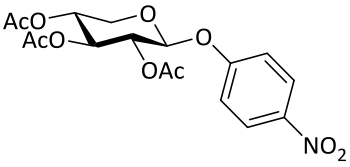
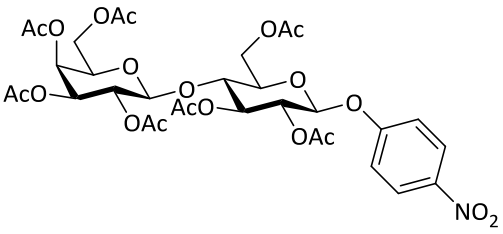
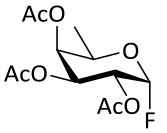
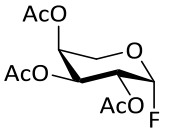
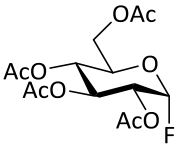
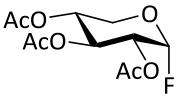
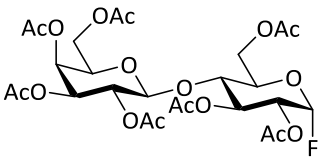


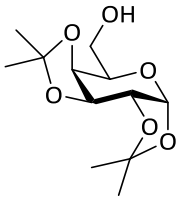
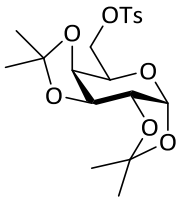
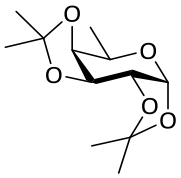
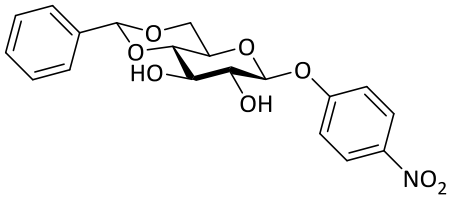
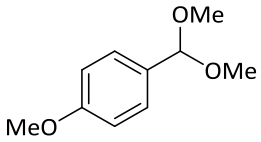
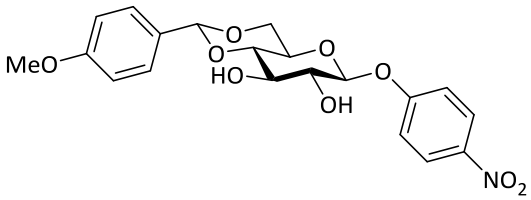
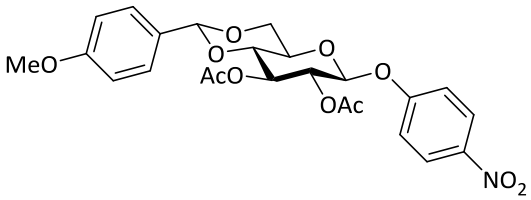
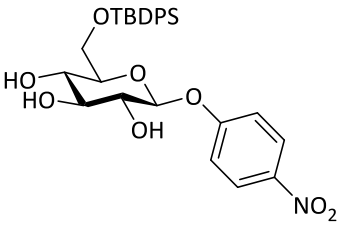
Appendix B - Compound List and Spectral Data

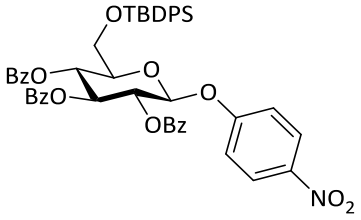
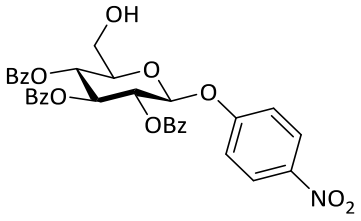
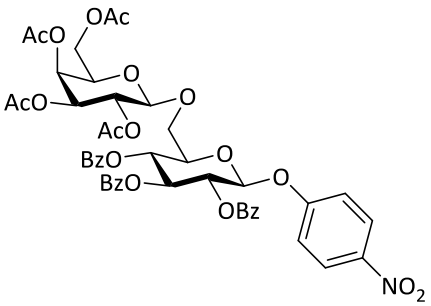
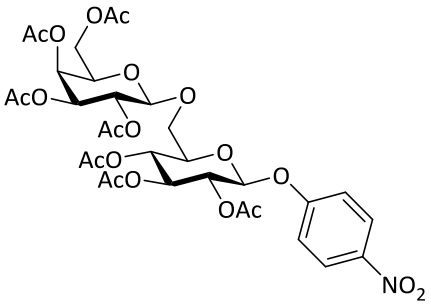
Spectral data is available on request.

Number	Structure	Lab Book Entry
1		BB2-30
4		BB1-81
5		BB2-90
22		BB1-82
25		BB2-29
26		BB2-28

27		BB1-83
31		BB2-72
33		BB2-49
41		BB1-92
42		VO216_BB
44		BB1-87
45		BB1-86
46		BB2-07
47		BB2-37
48		BB2-18

49		BB2-22
50		BB2-36
51		BB1-76
52		BB2-10
53		BB2-20
54		BB2-19
55		BB1-02
56		BB1-07
57		BB2-24

60		VO202_BB
61		VO207_BB
62		VO208_1_BB
65		BB1-29
67		BB1-38
68		BB1-39
69		BB1-41
70		BB2-79

71		BB2-80
72		BB2-82
73		BB2-83
74		BB2-35_glu