Glucuronic Acid Derivatives in Enzymatic Biomass Degradation: Synthesis and Evaluation of Enzymatic Activity

d'Errico, Clotilde; Madsen, Robert

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Glucuronic Acid Derivatives in Enzymatic Biomass Degradation: Synthesis and Evaluation of Enzymatic Activity

PhD Thesis – May 2016

Clotilde d'Errico

Department of Chemistry
Technical University of Denmark

ACKNOWLEDGEMENTS

This dissertation describes the work produced during my PhD studies. The work was primarily conducted at the DTU Chemistry in the period between November 2012 and January 2016 with an external stay at Novozymes A/S between September 2013 and March 2014.

First and foremost I would like to thank my supervisor Robert Madsen for giving me the opportunity to be part of this project, for the guidance during these years, for the valuable advices and the inspiring discussions.

I am very grateful to Rune Nygaard Monrad for being a great advisor during my experience at Novozymes A/S. Thanks for the patience, for the motivation and for sharing priceless knowledge about both carbohydrate chemistry and biotechnology.

Thanks to Kristian B. R. M. Krogh for the stimulating meetings and for providing the glucuronoyl esterases that I characterized during my project.

The Robert Madsen group, both present and former members, has been special to me, making the daily routine enjoyable and never boring! Especially, I need to thank my officemates Carola, Dominika and Gyrithe for the interesting talks, the sweets in the office, the reciprocal encouragement and for contributing to make the last three years (and a half) unforgettable.

A special thank goes to my colleague Maximilian Böhm for sharing with me his precious knowledge on the glycosylation reactions, and for taking over the glucuronoxylans project making some very remarkable progress.

Thanks to the fellow PhD students at the organic chemistry department for creating a pleasant and lively working environment and for never missing an occasion to share a cake!

My sincere gratitude goes to the technical staff of the chemistry department, without whom working in the lab would be simply impossible. Thanks to the sweet Anne Hector, to Brian Dideriksen, Brian Ekman-Gregersen, Charlie Johansen and Tina Gustafsson for their inestimable help, especially during the chaos of the moving to the new building.

This thesis would be still a formless draft without the invaluable contribution of Martin Jæger Pedersen, Jens Engel-Andreasen, Christine Kinnaert, Lee Fisher and Beatrice Bonora. Thanks for your time and for the thoughtful comments and suggestions.

My greatest thank goes to Beatrice. This experience in Denmark gave me the most amazing gift: your friendship. Thanks for listening, for supporting me unconditionally, for making me feel like home, always. You will always be my favorite vegetarian. *Grazie!*

Thank you Martin, for your love, for being always on my side and for your endless patience. Your presence in my life just made me a better person.

Last but not least, my deepest gratitude goes to my family for their infinite support and love, for their constant presence in my life and for always believing in me. *Vi amo!*

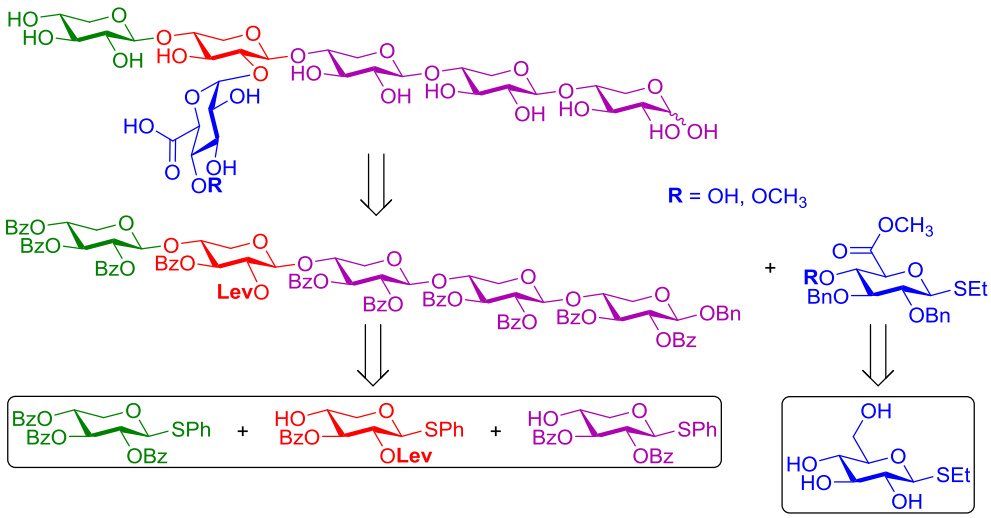
Clotilde d'Errico

May 2016

ABSTRACT

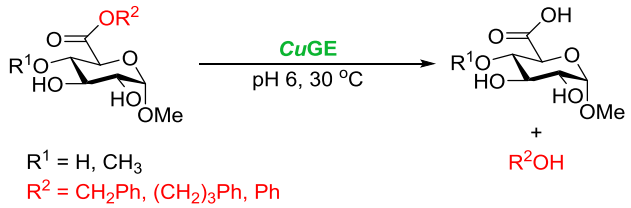
An essential tool for biotechnology companies in enzyme development for biomass delignification is the access to well-defined model substrates. A deeper understanding of the enzymes substrate specificity can be used to address and optimize enzyme mixtures towards natural, complex substrates. Hence, the chemically synthesized substrates often outcompete those isolated from natural sources in terms of reproducibility, homogeneity and purity.

The first part of this work was the synthesis of two glucuronoxylan fragments designed as model substrates for xylanases. The synthesis involved the use of thioxyloside building blocks in an iterative, linear glycosylation strategy.



Two sidechain glucuronate building blocks were synthesized via a divergent strategy from the same thioethylglucose derivative.

In the second part of the project three alkylaromatic and aromatic esters have been prepared as mimics of lignin-carbohydrate complexes found in lignocellulosic biomass, as model substrates for glucuronoyl esterases (GEs). These esters have been used to characterize a novel GE from *Cerrena unicolor* (*Cu*GE), produced by Novozymes, to obtain insights into the substrate specificity of the enzymes.

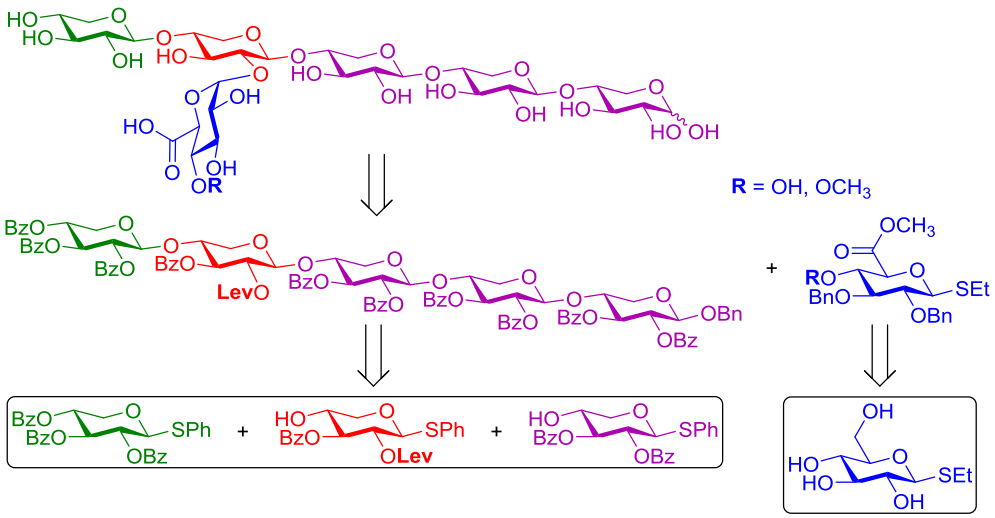


HPLC analysis of the enzymatic reactions led to the determination of kinetic parameters that gave information about both bonding affinity and catalytic efficiency.

RESUMÉ

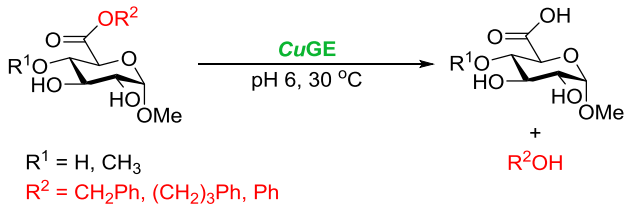
Et vigtigt redskab for bioteknologiske virksomheder, i enzymudvikling til biomasse delignificering er adgang til veldefinerede model substrater. En dybere forståelse af enzymers substratspecificitet kan anvendes til at evaluere og optimere enzymblandinger til naturlige, komplekse substrater. Derfor udkonkurrerer kemisk syntetiserede substrater ofte tilsvarende isoleret fra naturlige kilder i form af reproducerbarhed, homogenitet og renhed.

Den første del af dette arbejde var syntesen af to glucuronoxylanfragmenter udformet som modelsubstrater for xylanaser. Syntesen involverede anvendelsen af thioxylosider som byggeklods i en iterativ lineær glycosylerings strategi.



To sidekæder af glucuronat byggeklodser blev syntetiseret via en divergerende strategi fra samme thioethylglucose derivat.

I anden del af projektet blev tre alkylaromatiske og aromatiske estere forberedt som efterligninger af de lignin-kulhydrat komplekser, der findes i lignocellulose biomasse, som modelsubstrater til glucuronoyl esteraser (GE). Disse estere er blevet anvendt til at karakterisere en hidtil ukendt GE fra *Cerrena unicolor* (*Cu*GE), produceret af Novozymes, for at opnå indsigt i substratspecificiteten af enzymet.



HPLC-analyse af enzymatiske reaktioner førte til bestemmelsen af kinetiskeparametre, der gav oplysninger om både limning affinitet og katalytisk effektivitet.

LIST OF ABBREVIATIONS

1-BBTZ	1-(Benzoyloxy)benzotriazole
Ac	Acetyl
acac	Acetylacetonate
Ar	Aromatic
Ara	Arabinose
All	Allyl
Bn	Benzyl
br	Broad
BSP	1-benzenesulfinyl piperidine
Bu	Butyl
Bz	Benzoyl
CE	Carbohydrate esterase
CSA	Camphorsulfonic acid
Cu	*Cerrena unicolor*
d	Dublet
DAST	(Diethylamino)sulfur trifluoride
DCC	*N*,*N'*-dicyclohexylcarbodiimide
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DMAP	*N*,*N*-dimethyl-4-aminopyridine
DMF	*N*,*N*-dimethylformamide
DMTST	Dimethyl(methylthio)sulfonium triflate
DTBMP	2,6-di-*tert*-butyl-4-methylpyridine
EDC	*N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide
ESI	Electrospray ionization
Et	Ethyl

Fmoc	Fluorenylmethyloxycarbonyl
FT-ICR	Fourier transform ion cyclotron resonance
Gal	Galactose
gCOSY	Gradient-selected correlation spectroscopy
GE	Glucuronoyl esterase
GH	Glycoside hydrolase
Glc	Glucose
GlcA	Glucuronic acid
Glu	Glutamic acid
HexA	Hexenuronic acid
His	Histidine
HMBC	Heteronuclear multiple bond correlation
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation
i	*iso*
IDCP	Iodonium dicollidine perchlorate
LCC	Lignin-carbohydrate complex
Lev	Levulinyl
m	Multiplet; *meta*
MALDI	Matrix-assisted laser desorption/ionization
Man	Mannose
*m*CPBA	*m*-chloroperbenzoic acid
Me	Methyl
MS	Mass spectrometry
Nap	2-Naphthylmethyl
NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide

NMR	Nuclear magnetic resonance
p	*para*
PDC	Pyridinium dichromate
Ph	Phenyl
PMB	*p*-methoxy benzyl
Pr	Propyl
Py	Pyridine
R_f	Retention factor
RRV	Relative reactivity value
s	Singlet
Sc	*Schizophyllum commune*
Ser	Serine
SET	Single electron transfer
t	Triplet
TBAF	Tetrabutylammonium fluoride
TBAI	Tetrabutylammonium iodide
TBS	*tert*-butyldimethylsilyl
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy
Tf	Triflyl; trifluoromethanesulfonyl
TFA	Trifluoroacetic acid
TFAA	Trifluoroacetic anhydride
THF	Tetrahydrofuran
TLC	Thin-Layer Chromatography
TMS	Trimethylsilyl
Tol	Tolyl; *p*-methylphenyl
Tr	Trityl; triphenylmethyl
UV	Ultraviolet
Xyl	Xylose

PUBLICATIONS

- "Enzymatic Degradation of Lignin-Carbohydrate Complexes (LCCs): Model studies Using a Fungal Glucuronoyl Esterase from *Cerrena unicolor*"

- "Improved Biomass Degradation using Fungal Glucuronoyl Esterases – Hydrolysis of Natural Corn Fiber Substrate"

CONTENTS

1 INTRODUCTION

The global interest in renewable resources has exponentially increased in the last decades in conjunction with the depletion of fossil resources and their frightening environmental effects. In this context, biomass and biomass-derived materials have been regarded as promising alternatives since they are the only source of organic carbon suitable for the production of fuels and fine chemicals, with a primary benefit of net zero carbon emission being produced through biological photosynthesis.[1]

The most abundant and renewable biomass source is lignocellulose, which is the non-edible part of plants and therefore not exploitable for the food industry. In fact, since lignocellulosic feedstocks are accumulated as forestry, agricultural and agro-industrial waste, their disposal represents a further environmental problem. As a result, lignocellulosic biomass has a massive potential for sustainable production of bioethanol, paper or fine chemicals that makes it the most interesting renewable natural resource.

Nevertheless the success of lignocellulose applications is hampered by the intrinsic recalcitrance to enzymatic and chemical degradation of the heterogeneous polymers. Therefore extensive research is currently ongoing to address this problem, especially with regard to enzymatic degradation.[2]

Specific enzyme mixtures are prepared in order to degrade the intricate mixture constituting lignocellulose. The factors that affect the hydrolysis can be both enzyme-related and substrate-related, due to the scarce accessibility of enzymes to their designed targets.[3] In order to weaken those factors, a suitable pretreatment

process is usually applied to the raw material. Physical, chemical, physico-chemical and biological processes have been developed in the past decades, aiming at improving the enzyme digestibility and at maximizing the use of isolated lignocellulosic components.[4] However, due to the high cost and resource-consumption of this step,[5] new innovative and sustainable technologies are continuously being investigated.[4]

1.1 Lignocellulose structure

The plant cell wall is composed of three major polymers: cellulose, hemicelluloses and lignin.[2] Cellulose is a linear homogeneous polysaccharide consisting of D–glucose units linked by β–1,4 glycosidic bonds, with a high degree of polymerization. It is present in all the terrestrial plant cell walls, making it the most abundant natural polymer on earth. Hemicelluloses, on the other hand, are a group of heterogeneous polysaccharides including xyloglucans, xylans, mannans and glucomannans, characterized by the same equatorial configuration at C1 and C4.[6] The relative content and construction varies dramatically from species to species, however it is generally acknowledged that the major groups of hemicelluloses are xylans, composed of β–1,4 linked D-xylose units, decorated with arabinoses, glucuronic acids, acetates or xyloses.[7] Their main function is to strengthen the cell wall by creating cross-interactions with both cellulose and lignin.[6]

The last major component in plant cell wall, lignin, is a non-carbohydrate based polymer consisting of three kinds of phenylpropanoid alcohols (*p*-coumaryl, coniferyl and sinapyl alcohols), differing in the degree of substitution at the phenolic ring, combined in an intricate and highly branched pattern. Its main functions are to lead water movements in the plant walls, acting as a barrier to evaporation, defense from pathogens and structural support.[8] Therefore, the lignin content varies a lot from plant to plant, with higher abundance in trees (about 28%

for softwoods and 20% for hardwoods), and within different parts of the plant.[8] While there is no acknowledgement about links between cellulose and lignin, the latter is found to have a variety of covalent bonds to hemicellulose residues, the so-called lignin-carbohydrate complexes (LCCs).[4]

The LCCs present in lignocellulosic material are implied to contribute to its intrinsic recalcitrance towards enzymes, thus rendering the process of delignification more challenging.[4,9] Therefore, it is of fundamental importance to identify and characterize those patterns. The little information about LCCs originally came from model experiments, since lignin extracts were difficult to obtain without degrading LCCs in the process. Recently more efficient preparation processes were developed,[10] allowing the identification[11,12] and quantification[13,14] of these complexes from different plant sources. Among others, LCCs from grasses were identified to be mainly composed of ferulate and *p*-coumarate esters with arabinose residues (Figure 1),[15] while the analyses on soft and hardwoods revealed the presence of benzyl ether, phenyl glycosides and esters of glucuronic acid or 4-*O*-methyl glucuronic acid (γ-esters).[13] No benzyl esters (α-esters) have been found, although they have been observed indirectly by Imamura via 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) oxidation of ester LCCs,[16] in accordance with the uronosyl group rearrangement demonstrated by Li and Helm.[17]

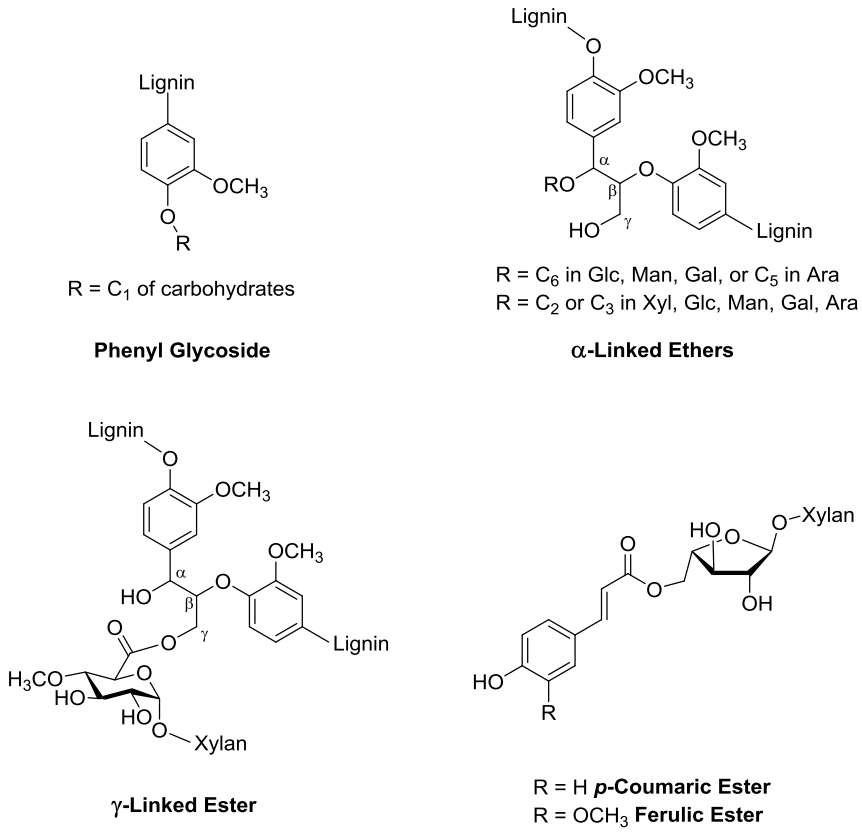


Figure 1 Examples of LCCs found in nature

1.2 Xylans

Xylans are generally known to have a highly complex and heterogeneous structure due to the quantity, distribution and differentiation of branching units. These features, together with the absence of repeating structures, make the classification non-trivial, although it is commonly accepted to divide xylans into homoxylans and heteroxylans, where the latter includes arabinoxylans, glucuronoxylans and (arabino)glucuronoxylans.[18] Beside the saccharidic substituents, most xylans are acetylated to various degree, which mainly takes place on the *O*-3 position, and in fewer cases on the *O*-2 of xylose residues. The relative abundance of the different

xylan classes in hemicelluloses characterizes the structure and properties of the different plant cell walls present in nature.[6,19]

1.2.1 Arabinoxylans

Arabinoxylans are abundantly represented in the cell wall of many cereals (*e.g.* wheat, rye, barley) and they are the main hemicellulose component in flour and wheat bran. Therefore a worldwide interest has been raised in the last decades related to their utilization as dietary fibers, and the many health benefits that they have been found to have.[20] Arabinoxylans present L-arabinofuranosyl residues linked with α-1,2 or α-1,3 bond singularly on a xylose residue, or both on the same unit (Figure 2). In addition, *O*-5 ferulate and *p*-coumarate esters have been observed on some arabinofuranosyl residues, mainly in grass plant cell walls.

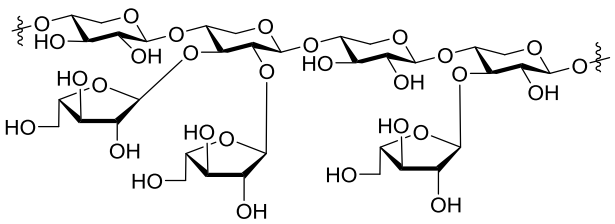


Figure 2 Model representation of arabinoxylan

1.2.2 (Arabino)glucuronoxylan

Arabino-(4-*O*-methylglucurono)xylans are a major component in softwoods, together with arabinoxylans, representing 7-15% of the wood.[21] They present 4-*O*-methyl glucuronic acid (MeGlcA) branches on the *C*-2 position, with a relative abundance of 1:5-6 of xylopyranosyl units,[22] while arabinofuranosyl residues are bonded at the C-3 position of the xylan backbone with an average ratio of 1.3 arabinofuranosyl units per 10 xylopyranosyl units (Figure 3).[10]

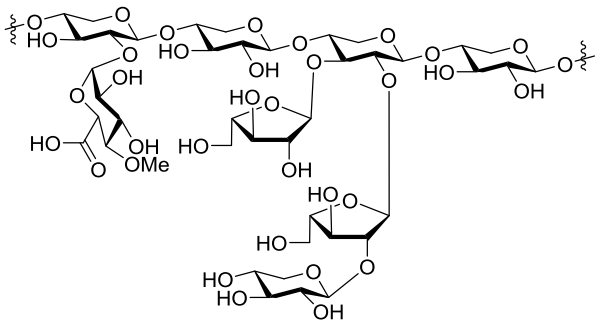


Figure 3 Model of a (arabino)glucuronoxylan

1.2.3 Glucuronoxylans

Glucuronoxylans are predominant in the cell walls of hardwoods. They consist of a xylan backbone substituted at the *O*-2 position with 4-*O*-methyl-α-D-glucuronic acids and/or α-D-glucuronic acids (Figure 4), with an average substitution ratio of 1:10 (one branch every ten xylose residues) and the remaining hydroxyl groups are up to 70% acetylated.[10,21]

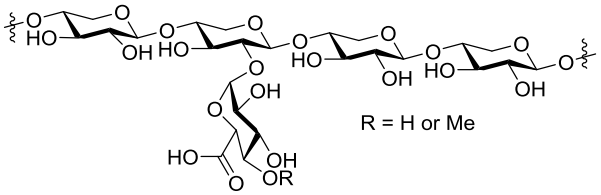


Figure 4 Model representation of glucuronoxylan

An alkaline treatment is used in the Kraft pulping of wood (a preliminary step in the paper production) to convert wood into pulp, that consists essentially of cellulose. This process hydrolyzes the ester groups and partially converts the glucuronic esters into hexenuronic acid (HexA) due to β-elimination, and into 4-*O*-methyl-L-iduronic acid *via* C-5 epimerization. Those conversions are known to interfere with the process in paper production,[23] and therefore it is of high concern to decipher the actual mechanisms involved and the complex oligosaccharides mixture originating from them.

1.3 Enzymatic degradation of lignocellulose

The main challenge in the degradation of lignocellulose has been to efficiently remove the lignin covalently bound to the hemicelluloses.[10] The alkaline pretreatment used in paper production is, at the moment, successful in the hydrolysis of LCCs but this is not applicable or economically feasible for other applications, *e.g.* animal feed or biofuel production.

A possible solution has been envisioned in the screening of suitable enzymatic cocktails, where several differentiated enzymes, defined as hemicellulases, are working synergistically to disrupt the complex architecture of the hemicellulose (Figure 5).

The enzymes involved in xylan degradation can be divided into two groups: main-chain enzymes, which includes β-1,4-D-xylanases and β-1,4-D-xylosidases, and side-chain enzymes, involved in the removal of branches of the xylan backbone, such as α-arabinofuranosidases, acetyl xylan esterases and α-D-glucuronidases.[24,25]

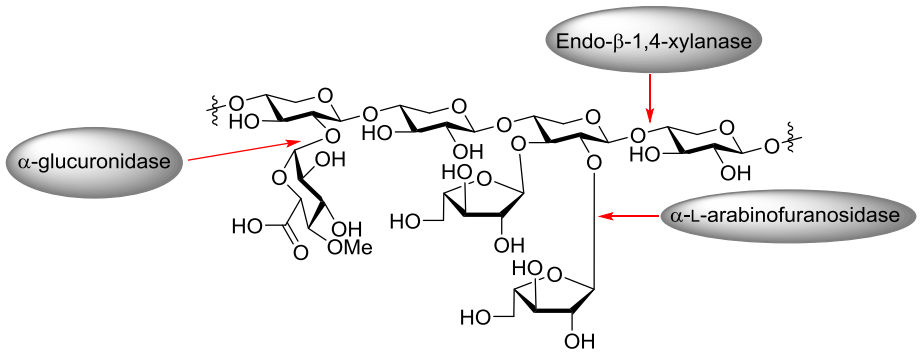


Figure 5 Schematic representation of enzymes involved in the hydrolysis of a substituted xylan

Those enzymes act in a synergistic fashion, defined as heterosynergy, where main-chain enzymes actions are making side-chain enzymes substrates more accessible,

or vice versa. Therefore the combination of their actions is more effective than the sum of their individual interventions.[19]

Moreover, several enzymes used in hemicellulases mixture are also able to hydrolyze LCCs,[26] although a more targeted approach is necessary to optimize the delignification process. This optimization relies on a deeper knowledge about those enzymes required for this process. For that purpose, the design of well-defined and specifically targeted substrates for hemicellulases is found to be an indispensable tool to the goal of extending the understanding and the possible applications.

1.3.1 Xylanases

Xylanases are *O*-glycoside hydrolases involved in the degradation of xylans. Their function is to cleave randomly 1,4-β-D-xylosidic linkages. They are a wide family of diverse enzymes with different substrate specificities, mechanism of action, physicochemical properties and hydrolytic activities.[27] They find application in a wide range of industrial processes from food industry to bioconversion of agricultural wastes to paper industry. Extensive studies have been carried out mainly on two of the xylanase containing glycoside hydrolase (GH) families (GH 10 and 11) which occurred to be active only in presence of three or more consecutive, unsubstituted xylose residues, and therefore they are not useful for highly branched xylans. For example, it is well-known that branches of glucuronic acid are inhibiting the effect of xylanases on the neighboring β-1,4-bond of the xylan backbone of glucuronoxylans.[28,29] Recently a major interest has risen for newly discovered xylanases presenting a binding pocket in their active site which allows a substituent on the xylan backbone to be accommodated very close to the cleavage site. Hence those enzymes are able to deconstruct the xylan backbone of substituted xylans, such as arabinoxylans (GH5)[30] and glucuronoxylans (GH30),[31,32] selectively. GH30 are defined as glucuronoxylanases for their high

specific activity on xylans or xylooligosaccharides containing a 4-OMe-GlcA or GlcA substituent. The enzymes are described to hydrolyze the second glycosidic bond from the branching unit towards the reducing end, releasing a one unit shorter xylan strain.[33]

1.3.2 α-Glucuronidases

Enzymes belonging to the glycoside hydrolase family GH67 (CAZy database, www.cazy.org) were described by the name α-glucuronidases and reported to hydrolyze exclusively the α-1,2-bond between a MeGlcA unit and a xylose residue at the non-reducing end of short oligoxylans.[34] Thus they needed the synergistic effect mentioned above, *i.e.* the co-operation of endoxylanases and β-xylosidases to act on polymeric xylans.[35,36] More recently novel α-glucuronidases were discovered to promote the hydrolysis of the (4-*O*-methyl) glucuronic residues in polymeric xylans,[37] and described to belong to a novel glycoside hydrolase family GH115.[38,39]

The mechanism of action of α-glucuronidases has been thoroughly investigated for both families (GH115 and GH67) and it has been described as an inverting mechanism, meaning that an inversion of configuration is involved in the mechanism of hydrolysis, releasing the glucuronic acid residue as the β-anomer.[40] The same studies by Kolenová et al.[40] gave relevant information about the substrate specificity of this class of enzymes. It was observed that xylooligosaccharides of different lengths were recognized by the active site, and the substrate specificity increased with the number of xylose units suggesting that the enzyme binds not only the glucuronic acid but also the adjacent xylose residues.

1.3.3 Glucuronoyl Esterases

A new class of enzymes belonging to the carbohydrate esterase (CE) family CE15, has recently been discovered and suggested to hydrolyze the ester bond between lignin alcohols and xylan-bound glucuronic acids.[41] They have been named glucuronoyl esterases (GEs). GEs belong to the serine type esterases requiring no metal ion co-factors for catalytic activity.[42,43] Several have been characterized so far[41,43–47] and widely tested on compounds mimicking the naturally occurring LCCs in order to understand their activity.[42,46–48] Preliminary studies showed encouraging results about the specificity of the GEs, which were able to hydrolyze the esters of MeGlcA, whereas substrates of other typical carbohydrate esterases (acetylxylan, feruloyl, pectin methyl esterases) were left untouched.[41] These results were quickly followed by a confirmation of such hydrolytic activity on arylalkyl esters of both 4-*O*-Me glucuronic acid and glucuronic acid. A higher efficiency and faster reactivity on esters of the former compared to the latter suggested a direct role in the enzyme-substrate interaction of the methoxy group which might be indeed recognized by the GE active site.[45,47,48] Within the same scope of research, it has been observed that the enzymes selectively recognized the *gluco*-configuration of the uronic moiety, showing no activity on esters of galacturonic acids.[45] Furthermore, the GE interaction with a glucuronic ester, linked to a short xylan strain with a 1,2-α-bond, was tested to verify the influence of the saccharidic chain on the mechanism of hydrolysis. The results showed no clear difference in reactivity as compared to methyl glucuronates, thus suggesting no recognition of the carbohydrate portion by the enzyme.[48] Nevertheless, the behavior of the enzyme on a more complex, natural substrate could be influenced by the polysaccharic matrix, and therefore a more detailed analysis was conducted on a polymeric substrate made by chemical methyl esterification of alkali-extracted glucuronoxylan. Biely and coworkers[49]

demonstrated that several microbial glucuronoyl esterases were able to deesterify methyl glucuronate residues in a complex structural arrangement, with a similar rate of deesterification for low molecular mass methyl esters of MeGlcA. Most recently, glucuronoyl esterases have been used in hemicellulases mixture on natural lignocellulosic material and a synergic activity has been observed, proving their potential as auxiliary enzymes in the saccharification of lignocellulosic biomass.[50]

1.4 Glucuronate mimics

Before using enzymes on real LCCs preparations, a preliminary analysis is necessary to understand the mechanism of action and the specificity in terms of substrates. For that purpose, mimics of LCCs have been synthesized in the last decades with different degrees of complexity and affinity to the real substrates. As it concerns GE substrates, the synthesis of several esters of glucuronic acid and 4-OMe glucuronic acid has been pursued recently to fulfill the need for ideal substrates for the testing of GE activity both chemically[17,45,51] and enzymatically.[46] Generally, those compounds are glucuronosides featuring an ester moiety, aimed at representing lignin components, that varies from being a methoxy group to aromatic alcohols, to a dimeric lignin-like alcohol.[17] The anomeric group has also been varied, having in the first place a xylose unit or a disaccharide mimicking the xylan chain, then a simple methyl group since it was determined that there was no hampering of the anomeric portion in the mechanism of action of the enzyme.[48] A methoxy group on the *O*-4 position of the sugar component is almost always present due to strong evidences of its abundance.

The first glucuronoyl esterases were tested on methyl glucuronates synthesized by Hirsch (Figure 6)[51,52] starting from the same methyl (benzyl 2,3-di-*O*-benzyl-4-*O*-methyl-β-D-glucopyranosid)uronate which was deprotected and hydrolyzed at the anomeric position to obtain **1**,[52] or modified to the corresponding chloride and

coupled with 1,3,4-tri-*O*-acetyl-α-D-xylopyranose giving the disaccharide which was reacted at the reducing end with *p*-nitrophenol. Deacetylation of the latter gave the final compound **2**, suitable for analysis with UV detection.[51]

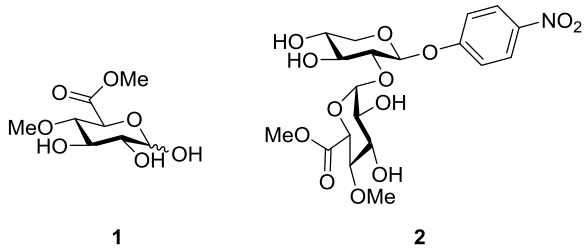


Figure 6 Methyl glucuronates synthesized by Hirsch[51,52]

More substrates were synthesized by esterification with ethereal diazomethane of commercially available D-glucuronic acid, D-galacturonic acid and their *p*-nitrophenyl glycoside relatives to obtain their corresponding methyl esters.[45]

In 2014 three different aryl, alkyl or alkenyl esters of glucuronic acid have been synthesized enzymatically using a lipase B (from *Candida antarctica*) that coupled D-glucuronic acid with cinnamyl alcohol, 3-phenyl-1-propanol and 3-(4-hydroxylphenyl-)-1-propanol in order to have esters as similar as possible to natural LCCs (Figure 7).[46] The main drawback of this technique is that the lipase requires the sugar to be in its open form, excluding the application on glucuronosides, which are supposedly the only way the glucuronates are present in lignocellulosic material.

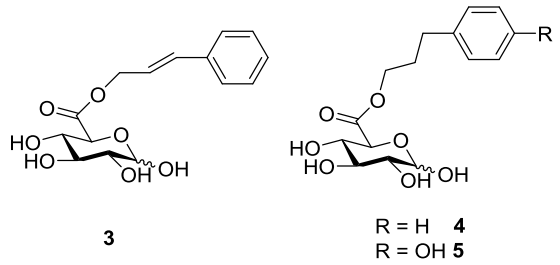


Figure 7 LCCs mimic synthesized via enzymatic route

In order to mimic the lignin components of the LCCs under investigation, aromatic esters of glucuronic acid have been also synthesized.[53] Their synthesis was straightforward and the only challenge was related to the oxidation of the glucoside chosen as the starting material to get to the glucuronic acid and subsequently to the corresponding ester.

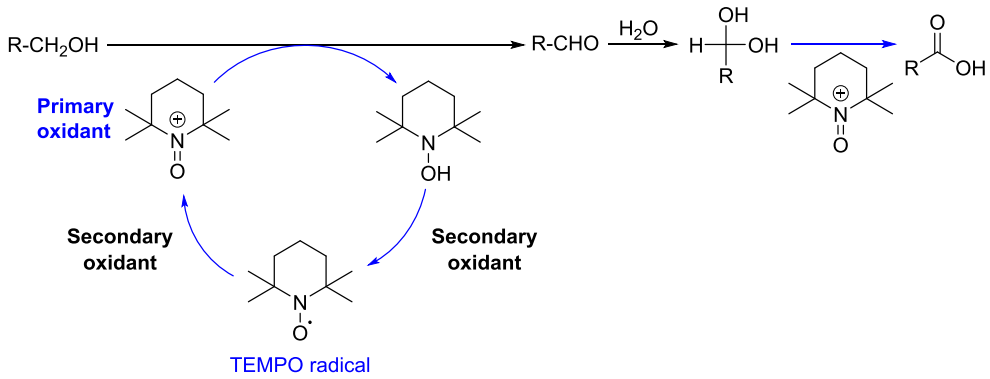
1.5 Oxidation

Organic chemistry includes a large variety of oxidation methods, many of which have been applied to carbohydrates. This work will be focused on those specifically used both on unprotected and protected saccharides to oxidize the primary alcohol to carboxylic acid.

The traditional methods to oxidize a primary alcohol in protected monosaccharides use metals in a high oxidation state, like potassium permanganate or ruthenium (VIII) oxide but they are not compatible with olefins, sulfides or benzyl ethers.[54] Chromium (VI) oxidants have been utilized to obtain uronic acids as well, although the Jones reagent resulted in hydrolysis of acid labile protecting groups and only goes to completion when used in excess (2 to 5 eq).[55] Addition of pyridinium dichromate (PDC) showed improved results when used in larger excess and/or longer reaction times than the conditions reported to obtain aldehydes.[56] PDC could also be used in combination with other oxidants (Swern reagents) to achieve the oxidation in two steps.[57]

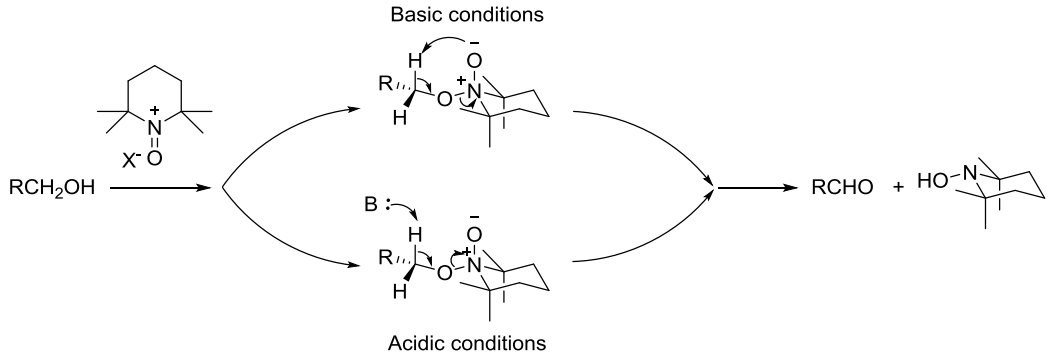
Those examples are typically not compatible with unprotected or partially protected glycosides, and therefore milder reaction conditions are required in order to take advantage of the higher accessibility of the primary alcohol. A mild oxidizing agent widely used for this purpose is 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO), a stable nitroxyl radical used in catalytic amount together with a stoichiometric secondary oxidant.[58]

1.5.1 Oxidation with TEMPO



Scheme 1 Proposed TEMPO mechanism

The proposed mechanism for a TEMPO mediated oxidation (Scheme 1) initiates with the secondary oxidant involved in the transformation of the nitroxyl radical TEMPO into the active oxoammonium salt, which reacts with the alcohol through one of the two different intermediates according to the reaction conditions (Scheme 2). In a basic environment there is a five-membered intermediate, more compact, that accelerates the reaction (and increases selectivity towards primary alcohols) compared to the linear counterpart that occurs in acidic conditions.[59]



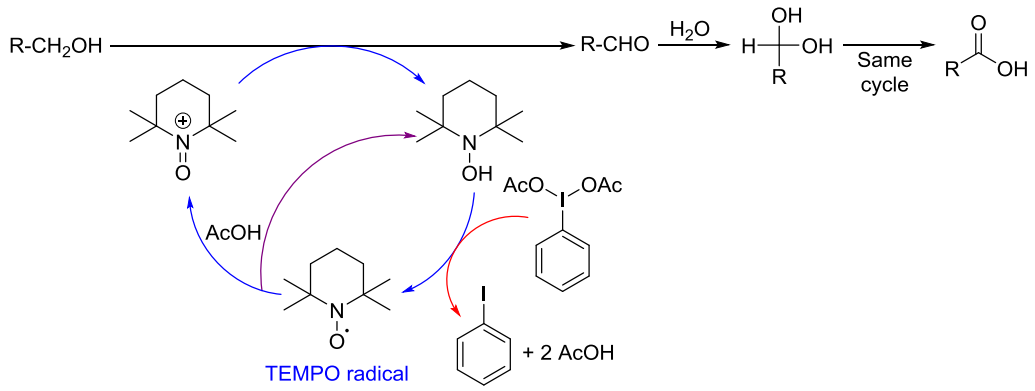
Scheme 2 Suggestion for two different intermediates in TEMPO oxidation, according to the different reaction conditions

This first transformation leads to the formation of the aldehyde intermediate and a hydroxylamine, which is converted again into the TEMPO radical, closing the catalytic cycle. The aldehyde reacts with water to form the hydrated form that is in an equilibrium followed by subsequent oxidation to the carboxylic acid via the regenerated oxoammonium species (the presence of water showed to be necessary to obtain the highest oxidation state) (Scheme 1).[59] Nevertheless it can be affected by the excess of the secondary oxidant, which in such a case becomes the primary oxidant for that oxidation step.

Initially sodium hypochlorite (NaClO) was used as a co-oxidant, in the presence of a bromide source, such as KBr or NaBr, to generate the more powerful oxidant HOBr *in situ*.[60] A wide range of saccharides[59,61–63] have been tested, including complex natural oligosaccharides,[64,65] but the main drawback of this procedure was the high chlorinating properties of NaClO that limited the scope to substrates not sensitive to chlorination. This side effect could be avoided or mitigated if only a catalytic amount of NaClO was used and regenerated with a stoichiometric amount of sodium chlorite ($NaClO_2$), which is responsible for the further oxidation of the aldehyde.[66] In both scenarios the oxidation is selective for primary alcohols, in the

presence of protecting groups of different nature, but in the case of thioglycosides, oxidation at sulfur might occur.[55]

Most recently, several procedures reported the utilization of (diacetoxyiodo)benzene as a stoichiometric co-oxidant,[67–69] in a biphasic solvent mixture of dichloromethane and water. The byproducts are iodobenzene and acetic acid, instead of inorganic salts, making the reaction mixture easier to handle and to work up. De Mico *et al.*[67] suggested a mechanism, shown in Scheme 3, where (diacetoxyiodo)benzene is not oxidizing directly the TEMPO radical, but there is a ligand exchange at the iodine with the alcohol that liberates acetic acid, which is responsible for the disproportionation of TEMPO to the oxoammonium salt and hydroxylamine.[70] Hence, (diacetoxyiodo)benzene reacts with the hydroxylamine to regenerate the corresponding radical and completes the catalytic cycle. The same catalytic cycle is responsible for the further oxidation to the carboxylic acid.[68]



Scheme 3 TEMPO oxidative cycle in the presence of (diacetoxyiodo)benzene as secondary oxidant

This protocol has been widely applied on carbohydrates,[71,72] and it is reported to be compatible with thioglycosides.[69,71]

1.6 Protecting groups

Carbohydrates are complex molecules bearing multiple functional groups which are not always trivial to differentiate among. The anomeric carbon, being a hemiacetal (or acetal in case of a glycoside), has a defined chemistry, easily differentiable from the other groups on the same molecule, which are hydroxyl groups with a very similar reactivity although their biological role could be different according to their relative position.[73] Therefore it is essential to find a way to distinguish among them and for that purpose protecting groups are widely utilized. While it is often straightforward to discriminate between primary and secondary alcohols, the regioselectivity related to the different secondary groups can be very challenging. An additional challenge is represented by the orthogonality among protecting groups which allows for releasing functionalities in a selective fashion but, on the other hand, requires a careful planning with regards to installation/removal reaction conditions.

Even though the carbohydrate protecting groups are generally the same as used in other areas of organic chemistry, it has to be pointed out that those can have additional functions in the general behavior of the entire molecule than just protection, like the direct participation of ester groups on *C*-2 positions of glycosyl donors or their activation or inactivation (arming/disarming) in glycosylation reactions, as it will be further discussed later in this introduction.

In general, the main functionalities utilized for the protection of hydroxyl groups are ethers, esters and acetals.

1.6.1 Ether-type protecting groups

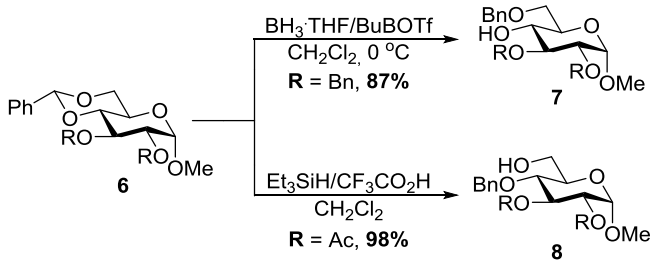
The benzyl ether is probably one of the most applied protecting groups in carbohydrate chemistry due to the stability and the neutral removal conditions.

Thanks to its etheric nature, it can affect the reactivity of the entire saccharide with an activating effect.

The ether bond is generally formed via reaction of the alcohol with benzyl halides in the presence of a base, such as NaH in *N*,*N*-dimethylformamide (DMF),[74] even though milder conditions have been developed in the last decades, like the use of a milder base and a phase-transfer catalyst in tetrahydrofuran (THF).[75,76] On the other hand, the bond is very stable under various conditions but can easily be cleaved by reduction. Hydrogenolysis is usually the preferred method, performed with a Pd catalyst absorbed on charcoal under a hydrogen atmosphere or in the presence of a hydrogen transfer source.[77] In cases where this procedure has not been applicable, various other cleavage procedures have been developed, *e.g.* Na/liquid ammonia (Birch reduction), anhydrous $FeCl_3$ or DDQ.[78]

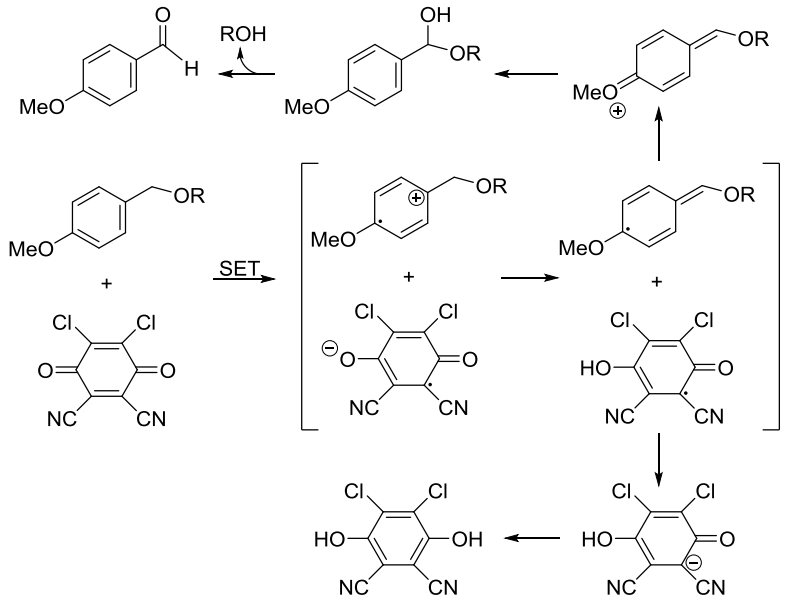
As mentioned above, the main advantage of using protecting groups is to discriminate between the alcohols with similar reactivity and therefore the best use of benzyl groups is the direct installment at the desired position in a regioselective manner. Comprehensive studies have been carried out for this purpose and the most successful strategies include the reductive opening of benzylidene acetals[79,80] and the use of organotin intermediates.[81]

The first method, the reductive opening of benzylidene acetals, refers to the selective cleavage of only one of the two carbon-oxygen bonds involved in the acetal giving a free OH group and a benzyl group. The direction of the opening depends from the reaction conditions together with steric and electronic factors. In the last decades Lewis acids and solvents,[79] together with the substituent on the *O*-3 position, have been screened. In Scheme 4 two different reaction conditions are reported for the synthesis of 6-*O*-benzyl (**7**)[82] and 4-*O*-benzyl (**8**)[80] methyl α-D-glucopyranoside.



Scheme 4 Reductive opening of a benzylidene group on methyl 4,6-*O*-benzylidene-α-D-glucopyranoside, to give the 4-*O*-benzyl[80] and the 6-*O*-benzyl[82] adducts

A way to implement the orthogonality of multiple-protection strategies could be the use of ether protecting groups with different stability, which can be modulated by introduction of substituents on the aromatic ring. The preferred ether of this kind is the *p*-methoxy benzyl ether (PMB) group due to the very convenient installation and removal conditions. While it is installed under similar conditions as the unsubstituted equivalent, the removal generally occurs via oxidation with DDQ, or a Lewis acid ($SnCl_4$). DDQ is commonly used since it is not affecting other protecting groups, including acid sensitive moieties. The oxidation is believed to happen through a single electron transfer (SET) to DDQ to form an oxonium ion which is then neutralized by water (Scheme 5).[83] At the same time DDQ get reduced to 2,3-dichloro-5,6-dicyanohydroquinone which is not soluble in dichloromethane and water, and therefore precipitates and keeps the reaction medium almost neutral through the reaction. This feature is essential in case of deprotection in the presence of acid-sensitive functionalities.



Scheme 5 Mechanism of PMB deprotection with DDQ

A bulkier substituent, that is known to be very useful for the protection of primary hydroxyl groups, is the trityl (Tr, triphenylmethyl) group. Its introduction, with trityl chloride in pyridine, is one of the oldest examples of selective alkylation of a saccharide. This method is still amply adopted for the formation of the bond that can be cleaved in an acidic environment, both in the presence of a Brønsted or a Lewis acid, such as aqueous sulfuric acid, trifluoroacetic acid or BCl_3.[84]

1.6.2 Acetal-type protecting groups

Acetal groups have been extensively used for protection in carbohydrate chemistry for more than a century due to the ease of their formation and the stability in a quite large spectrum of reaction conditions.

Cyclic acetals, such as isopropylidene and benzylidene groups, are the most commonly used for regioselective protection of 1,2- and 1,3-diols of saccharides. They are introduced by direct condensation of the carbonyl equivalent (acetone or benzaldehyde, respectively) or the corresponding dimethoxy acetals under acidic

conditions. The removal is also typically carried out in presence of an acid, either protic (aq. H_2SO_4, trifluoroacetic acid) or a Lewis acid.

As aforementioned, the main advantage is the regioselectivity of these groups, whose reactivity can be easily predicted according to the saccharide of interest. Benzylidene groups have a preference to react with 1,3-diols, which form a more stable six-membered ring. A rather unique conformation is observed, with the phenyl substituent in the equatorial position, and therefore they are mainly used for the protection of the *O*-4 and the *O*-6 position on pyranose moieties. On the other hand, isopropylidene groups would more likely protect vicinal 1,2-diols forming a five-membered ring and the reaction outcome would be strictly dependent on the polyol conformation and the relative thermodynamic stability.

1.6.3 Ester-type protecting groups

The presence of acyl protections in carbohydrate chemistry is ubiquitous, owing to the fair stability under acidic conditions and compatibility with glycosylation chemistry. Furthermore, if the ester is located on the *O*-2 position of a glycosyl donor, it provides the anchimeric assistance on the activation of the latter in the so-called neighboring group effect. As described in the following paragraph (Figure 8b), the ester can direct the stereochemical outcome of the glycosylation coupling.

The principal drawback is related to their utilization in partially protected sugars, and their tendency to migrate between vicinal hydroxyl groups. The general trend is the migration from an axial group to an equatorial position in case of 1,2-*cis* diols, or from secondary to primary OH groups.[73]

The most commonly employed esters for carbohydrates are, by all means, acetyl and benzoyl groups. Their reactivity has been comprehensively explored in the

literature over the last century,[78,85] and therefore it will not be the object of discussion in this report.

Several other esters have widely been used as protecting groups for carbohydrates, among which the levulinyl (Lev) ester functionality has had an increasing interest lately.[86–88] The reasons can be found in the minor aptitude to migration,[89] accompanied by the possibility to be removed with hydrazine monohydrate,[86] and orthogonally to other ester groups (acetates, pivaloates, benzoates). Those properties have resulted in inclusion of the Lev ester in various sets of orthogonal protections for the synthesis of collections of oligosaccharides.[90–92]

1.7 Glycosylation Reaction

Carbohydrates are mainly found in nature as oligo- or polysaccharides, and therefore, there has always been an enormous interest in understanding the mechanism behind the coupling between two single monomers in order to achieve efficient, stereoselective and high-yielding procedures for the assembly of nature-inspired saccharidic structures.

Generally, a glycosylation reaction consists of the generation of a glycosyl donor, preactivating the anomeric position with the installation of a suitable leaving group, and the glycosyl transfer to the glycosyl acceptor, opportunely protected and bearing a free hydroxyl group. This process follows a unimolecular S_N1 mechanism. Nevertheless, the details of the mechanism are hitherto not fully unraveled, as several studies are currently focused on demonstrating the existence of the glycosyl cation, generated by the departure of the leaving group on the donor.[93] Very recently the key ionic intermediate has been isolated in a superacid and the obtained spectroscopic data demonstrated its formation and the conformational analysis.[94] Figure 8 shows the possible pathways, based on

commonly accepted speculations and the state-of-the-art knowledge of the mechanism.

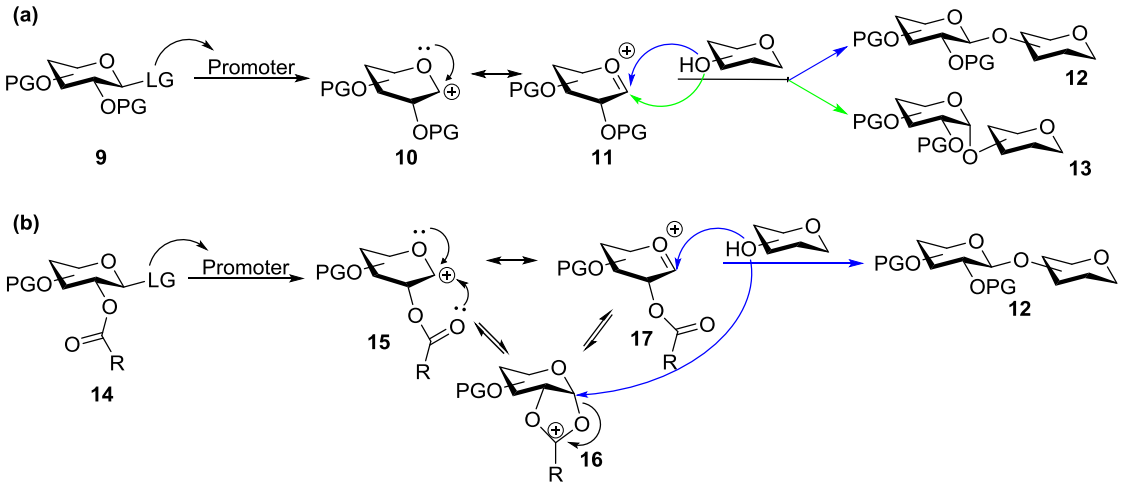


Figure 8 a. Mechanism for glycosylation reaction for a *gluco*- or *galacto*-configured monosaccharide **b.** Mechanism occurring in the presence of an ester group on the *O*-2 position[95]

The removal of the leaving group on the glycosyl donor is usually assisted by a promoter, or catalyst, which is typically a Lewis acid. Thus, the glycosyl cation **10** is formed and stabilized by resonance with *O*-5 generating the oxocarbenium ion **11**, where the sp^2 character of the anomeric carbon allows the nucleophile (acceptor) to attack from both faces of the molecule. Accordingly, the reaction could lead to the formation of two different products: 1,2-*cis* (*i.e.* **12**, α-gluco, β-manno configurations) or 1,2-*trans* (*i.e.* **13**, β-gluco, α-manno configurations) glycosides with a preference, more or less prominent, for the thermodynamically favored α-product due to the anomeric effect.

It must be pointed out that, in most cases, a different pathway is followed in the presence of an ester group on the *O*-2 position (Figure 8b), which would cause the departure of the leaving group by anchimeric assistance according to the neighboring group effect. The subsequent intramolecular stabilization in a bicyclic

intermediate, the acyloxonium ion **16**, would be responsible for guiding the attack towards one side of the molecule and consequently towards the formation of one single adduct, the 1,2-*trans* glycoside **12**.

The initial investigations on glycosylation reactions, by the end of 19th century, already faced the complexity of the process. A first, rational approach was accomplished by Koenigs and Knorr in 1901, whose experiment described the nucleophilic displacement at the anomeric position of a glycosyl chloride or bromide in the presence of Ag_2CO_3 as an acid scavenger.[96]

Those results inspired further experiments and in the next decades, more and disparate conditions were explored.[97] The curiosity of the chemical community towards this transformation never faded and it probably reached the climax in the 80s when a better understanding of the mechanism,[98] driving forces and principles of the glycosylation led to the development of new methods, focused mainly on the design of the novel anomeric leaving groups.[99] Among them, thioglycosides,[100] trichloroacetimidates[99] and fluorides[101] have been conceived in those years to become the most commonly utilized glycosyl donors at present.

1.7.1 Glycosyl donors

Halides

The first glycosylating agents, described by Koenigs and Knorr in 1901, were glycosyl halides.[96] Since then, glycosyl bromides and chlorides have been extensively investigated, whereas, in the last decades, the reactivity of fluorides[102] and then iodides[103] have also been widely explored.

The wide success of glycosyl halides is associated with the versatility of the method that allows to obtain 1,2-*trans* glycosides by exploiting the neighboring group effect, while α-glycosides can be achieved through *in situ* anomerization.

Nevertheless, the control of the stereochemical outcome required a very strict control of the reaction conditions, which is not always convenient.

Glycosyl iodides can be generated from the corresponding bromide in the presence of an iodine source (NaI) and although they are considered very reactive, they showed peculiar characteristics, which made them preferable in some cases over the more stable bromides or chlorides.[99]

On the other hand, for a long time glycosyl fluorides have been considered too stable to be used in glycoside synthesis due to the large bond-dissociation energy of the C—F bond (552 kJ mol^{-1}). Nonetheless a deeper knowledge about their manipulation and the activation mechanisms with weak Lewis acids in the last decades paved the way for their utilization.[101]

Imidates

Proposed for the first time as novel glycosyl donors by Sinaÿ[104] and developed in the 80s by Schmidt,[105] 1-*O*-substituted glycosyl imidates have received a pivotal role in contemporary carbohydrate chemistry, especially in the form of trichloroacetimidates, designed by Schmidt.[106] Their popularity is related to the ease of preparation, high-yielding reactivity and high anomeric stereocontrol. The imidates are commonly prepared from the corresponding hemiacetal by reaction with trichloroacetonitrile and a base, where the latter is determinant for the stereochemical outcome of the preparation. Indeed, NaH or Cs_2CO_3 yield the thermodynamically favored α-glycosyl donor, while K_2CO_3 promotes the formation of the kinetically favored β-product.[107] The glycosylation is usually carried out in the presence of a Lewis acid (trimethylsilyl trifluoromethanesulfonate, TMSOTf, or $BF_3{\cdot}OEt_2$), which is used in catalytic amount, and in this way constitutes a difference to the other current glycosylating

methods. This methodology, broadly implemented in the last decades, has been employed for the synthesis of diverse oligosaccharides both with 1,2-*trans* (using neighboring group participation)[108] and 1,2-*cis* glycosidic bonds.[109]

Thioglycosides

In the plethora of the well-known techniques to create glycosidic bonds, thioglycosides have a leading role since they were first used in 1973 by Ferrier.[110] The ease of preparation and the stability make them suitable candidates for handy and easily-controlled glycosylating procedures. Furthermore, they can be easily converted into other glycosyl donors.[111]

Preparation methods

The traditional and currently the most employed method to prepare 1,2-*trans* thioglycosides is the reaction of the corresponding peracetylated saccharide with the thiol of interest mediated by a Lewis acid (Figure 9c) (typically $BF_3 \cdot OEt_2$, but several others have also been used). Nonetheless, diverse procedures have been explored to achieve both thioalkyl- and thioaryl- glycosides using unprotected reducing sugars as starting material, in a one-pot procedure including acylation and subsequent thioglycosylation (Figure 9a),[112] or acylated glycosyl halides in presence of thiols[113,114] or disulfides (Figure 9b).[115] The latter method was already used in 1919 to prepare a 1-thioglycoside for the first time.[116] It is worth to mention that Hanessian in 1980 obtained the direct conversion of alkyl *O*-glycosides to the corresponding 1-thio-β-D-glycosides by using [alkyl (or aryl) thio]trimethylsilanes.[117]

Interconversion

Thioglycosides are well-known to be very stable under several reaction conditions, working as temporary protections for the anomeric position during protecting group manipulations, or acting as an acceptor and eventually could be converted into different glycosyl donors. A thioglycoside can be used, for example, to achieve the synthesis the corresponding glycosyl halides: a glycosyl bromide can be obtained by reaction with iodine monobromide (Figure 9f),[118] a glycosyl fluoride if treated with *N*-bromosuccinimide/(diethylamino)sulfur trifluoride (NBS/DAST) (Figure 9d)[119] and a glycosyl chloride can be synthesized by reaction with iodine monochloride or Cl_2 (Figure 9e).[120]

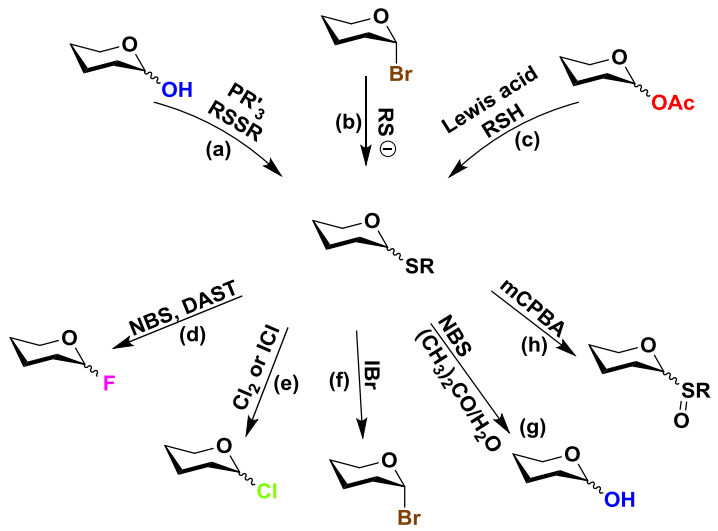


Figure 9 The most common preparation methods for thioglycosides and their conversion into different glycosyl donors

Another common transformation is the hydrolysis of thioglycosides to afford the corresponding hemiacetals (Figure 9g), which are then reacted to give the trichloroacetimidates.[121] The hydrolysis has been performed under several conditions, such as NBS or *N*-iodosuccinimide (NIS) in wet acetone,[122,123] $AgNO_3$

in wet acetone,[124] and tetrabutylammonium periodate/triflic acid (nBu$_4$NIO$_4$/TfOH).[125] A different approach involves the oxidation of the thioether into a sulfoxide, achieved with *m*-chloroperbenzoic acid (*m*CPBA) (Figure 9h)[126] or H_2O_2-acetic anhydride-SiO_2[127] in order to use the oxidized product with triflic anhydride in glycosylation chemistry.[126,128]

Direct use

The anomeric thioether group could, by interaction of the sulfur lone pair with a soft nucleophile, be activated to form a sulfonium intermediate which would be a superior leaving group in a glycosylation reaction. Hence, a wide range of promoters has been investigated in their ability to activate thioglycosides and in all cases, it has been concluded that at least a stoichiometric amount is necessary for the reaction to occur. Ferrier[110] performed for the first time a direct glycosylation of phenyl thioglycosides in the presence of mercury (II) salts and eventually other heavy metals[129,130] were employed as promoters although the yields and the selectivity were not outstanding when the nucleophile was a sugar. The breakthrough in direct thioglycosylation chemistry happened with the introduction of methyl triflate (MeOTf) as a promoting agent, which worked very efficiently due to the high thiophilicity of the reagent.[131] Further research led to the use of dimethyl(methylthio)sulfonium triflate (DMTST),[132] NIS/TfOH or NIS/TMSOTf[133] as the most common promoters, even applicable to very unreactive compounds.[134] Recently, the efficiency of the iodonium system has been confirmed by the development of many variants, such as a system with trityl tetrakis(pentafluorophenyl) borate [TrB(C_6F_5)$_4$]/NIS,[135] iodonium dicollidine perchlorate (IDCP)[136,137] and IPy$_2$BF$_4$/TfOH, which proved to be effective in one-pot sequential glycosylations.[138]

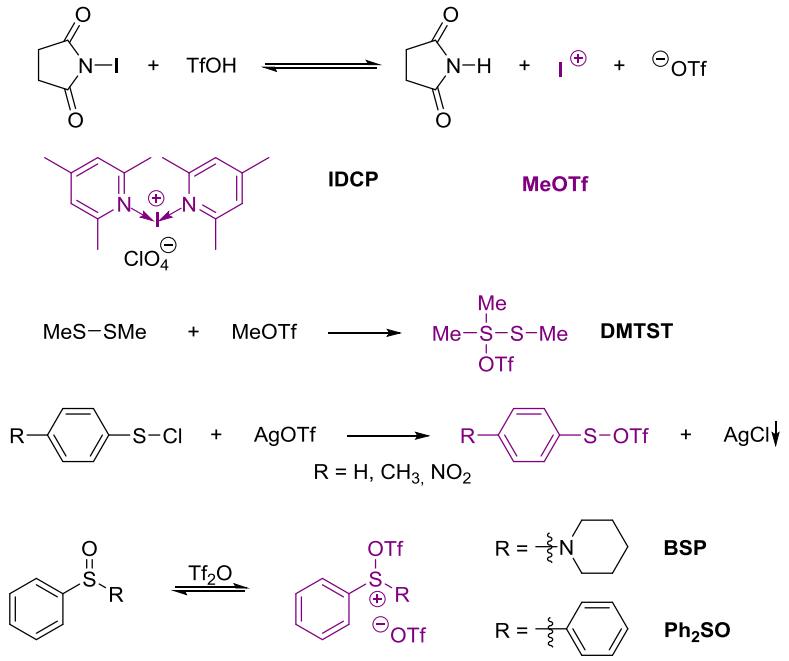


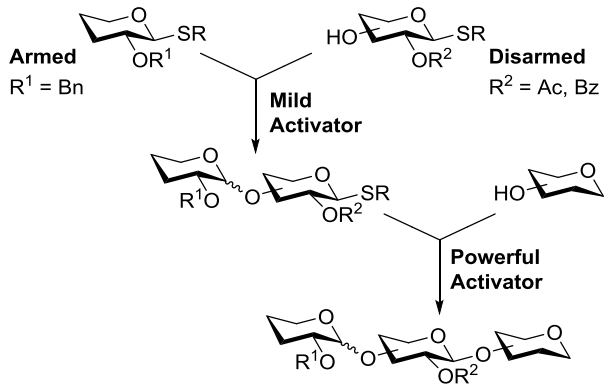
Figure 10 Activating agents for thioglycosides

Similarly, the sulfenyl/sulfonium class of thiophilic promoters was successfully investigated after the popularity of DMTST, with quite remarkable examples such as benzenesulfenyl triflate (PhSOTf),[139] *p*-toluenesulfenyl triflate (*p*-TolSOTf),[140] or the shelf-stable 1-benzenesulfinyl piperidine/triflic anhydride (BSP-Tf_2O)[141] and diphenyl sulfoxide/triflic anhydride (Ph_2SO-Tf_2O),[142] which have been widely employed in oligosaccharide synthesis (Figure 10). The aforementioned versatility of thioglycosyl donors have led to the extensive use of thioglycosides in the various sequential glycosylation strategies developed in the last decades, such as chemoselective, orthogonal and iterative techniques.[143]

1.7.2 Chemoselective glycosylation

With regard to thioglycosides, a chemoselective glycosylation is defined as the condensation between a highly-reactive thioglycosyl donor with a less reactive thioglycosyl acceptor. A pivotal role is certainly played by the protecting groups on

both molecules on the basis of the armed/disarmed concept described by Fraser-Reid.[144,145] His research group coined the term "armed" to describe a benzylated *n*-pentenyl glycoside, more prone to react with a nucleophilic acceptor than the "disarmed" counterpart, fully acylated. In fact, an electron-withdrawing protecting group, *e.g.* an ester, is decreasing the nucleophilicity of the thiofunctionality and destabilizing the oxocarbenium intermediate, leading to a lower reactivity.[146]

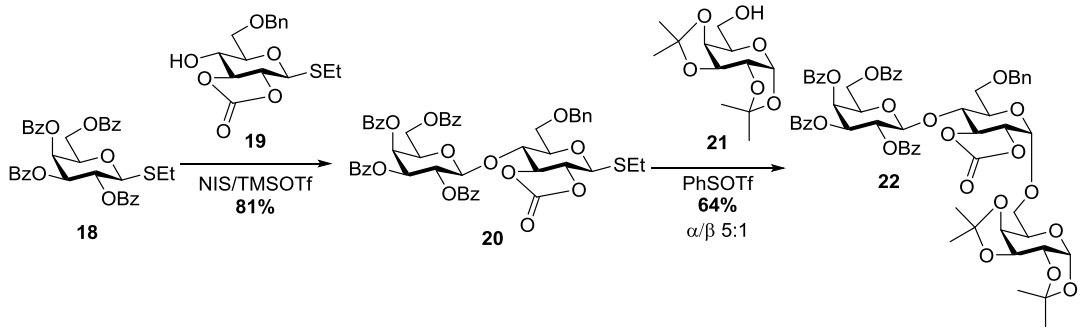


Scheme 6 Armed-Disarmed Strategy

This concept has been extended to different glycosyl donors, including thioglycosides,[137] and it paved the way for a novel glycosylation fashion where the difference in the reactivity between species leads to implement the stereochemical outcome and the efficiency (Scheme 6). For this purpose Wong and coworkers carried out an extensive study in order to classify hundreds of different tolyl thioglycosides on the basis of their relative reactivity values (RRVs)[147] and used this classification in the one-pot synthesis of complex oligosaccharides.[148]

In addition, it has been proven that even the solvent[149] and the substituent on the sulfur atom at the anomeric position[150] could affect the relative reactivity of the glycosides. Therefore this method, although widely used with remarkable results, requires an extremely careful design of the building blocks and of their protection pattern, especially with the perspective of finding the right distribution between

reactivity and stereochemistry in a glycosylation sequence. In order to have a less reactive donor an acyl group would be preferred, but that would lead to the formation, in the end of a sequence, exclusively of a 1,2-*trans* glycosidic bond. To have access to a 1,2-*cis* linkage, Zhu and Boons introduced a 2,3-cyclic carbonate group on a ethyl thioglucoside building block **19**, which was coupled as a disarmed acceptor and consequently able to participate in the next step that yielded the α linked trisaccharide **22** with a ratio of α/β = 5:1 (Scheme 7).[151]

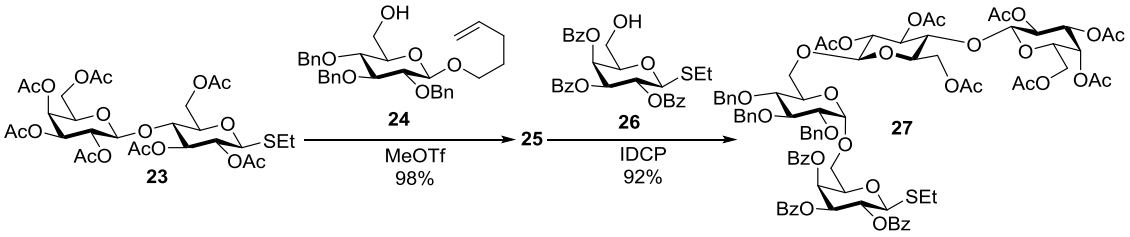


Scheme 7 The use of a disarming, non-participating protecting group in chemoselective glycosylations

1.7.3 Orthogonal glycosylation

Thioglycosides can be successfully employed in orthogonal glycosylations, where the sequential condensation occurs between two different glycosyl donors whose anomeric function can be activated in an orthogonal fashion. The clear advantage as compared to chemoselective glycosylations is the possibility to react compounds independently of their relative reactivities. The pioneer of this technique was Mukaiyama, who reported the use of fluoride donors in the presence of a thioglycosidic acceptor for the synthesis of a complex heptasaccharide,[152] and further broadened the scope with the use of several novel glycosyl donors in combination with thioglycosides.[153–155] A "semi-orthogonality" was exploited by Demchenko and De Meo (Scheme 8)[156] for the selective glycosylation of

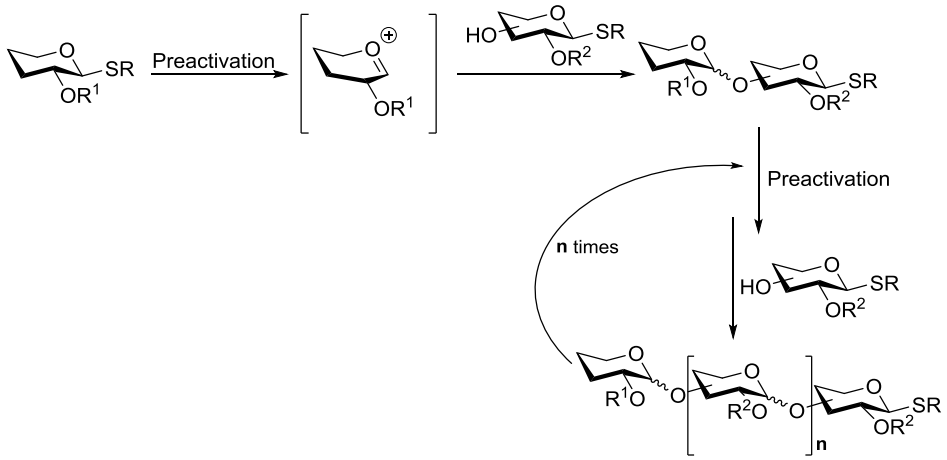
thiodisaccharide **23** with the *n*-pentenyl glycoside **24**, which are usually activated by similar promoting systems. However, they demonstrated that MeOTf could activate both armed and disarmed thioglycosyl donors in the presence of the *O*-glycosides achieving the synthesis of the linear tetrasaccharide **27**, which would not have been possible with a traditional armed/disarmed approach.



Scheme 8 Example of semi-orthogonal glycosylation developed by Demchenko. Adapted from Codée et al.[143]

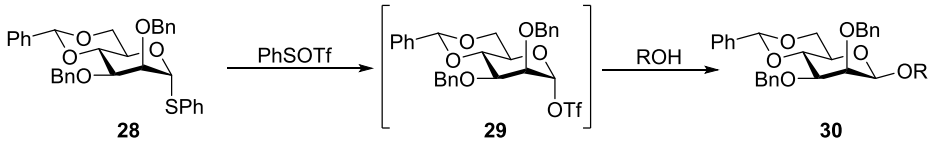
1.7.4 Iterative glycosylation

A further optimization of the glycosylation strategies, moving forward from the chemoselective method, is the iterative glycosylation, defined as a sequential process involving a single type of building block, condensed by using one set of reaction conditions, ideally in the same reaction vessel, *i.e.* in a one-pot fashion.[157]



Scheme 9 Iterative glycosylation strategy

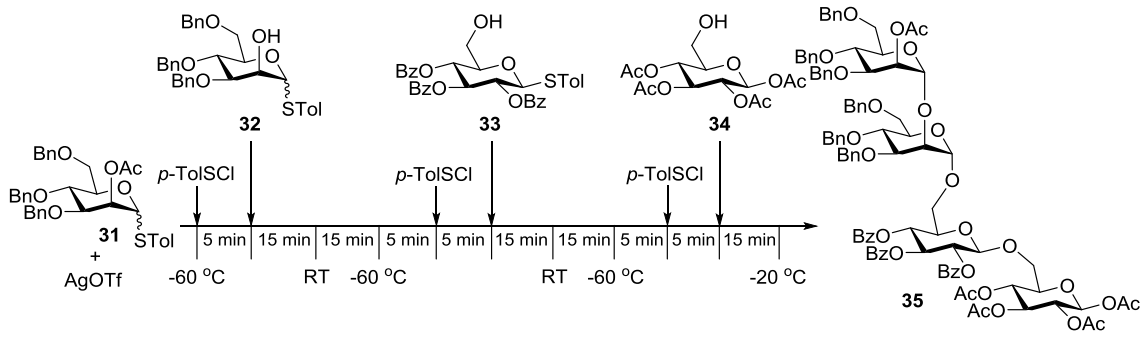
This method would relieve the synthesis from the elaborate design of the building blocks due to the reactivity tuning and, at same time, from time-consuming work up and purification steps. An effective tool for the implementation of this glycosylation is the pre-activation of the glycosyl donor, activated by a promoter in the absence of the acceptor, which would be added subsequently. If the newly formed disaccharide bears the same anomeric function, the glycosylating sequence could then be iterated (Scheme 9).[158] Though it requires that the promoter, used in stoichiometric amount, would be completely consumed and that the generated intermediate would be sufficiently stable to survive until the acceptor is added and, at the same time, sufficiently reactive to undergo glycosylation.[140] Once again, the stability of thioglycosides makes them suitable candidates for this kind of strategy, and an investigation into the different promoter systems led to the synthesis of several challenging oligosaccharides. The first example was reported by Crich, whose group synthesized challenging β-oligomannosides (Scheme 10) via pre-activation of thiomannoside **28**, and subsequent conversion to the α-mannosyl triflate **29** that underwent S_N2-type substitution.[139,159] The activation was achieved by reaction of the thioglycoside with PhSOTf, generated *in situ* by reaction of phenylsulfenyl chloride with silver triflate (AgOTf) in the presence of 2,6-di-*tert*-butyl-4-methylpyridine (DTBMP).[160]



Scheme 10 Pre-activation of thiomannosides operated by Crich

The pre-activation method was also used by the van der Marel group, using Ph_2SO-Tf_2O as the promoter, that was effective also on very disarmed thioglycosides owing to its high thiophilicity.[142] Those promoters were screened by Huang et al[140]

in the development of their iterative one-pot procedure, however, *p*-TolSOTf (generated in situ from *p*-toluenesulfenyl chloride and AgOTf) proved to be superior in their case. They also screened aglycon leaving groups and additives, finally opting for *p*-tolyl thioglycosides in the presence of the dehydrating reagent MS-AW300. This resulted in the development of a novel and efficient one-pot glycosylation approach for the synthesis of diverse oligosaccharides.[161,162] An example is schematized in Scheme 11.



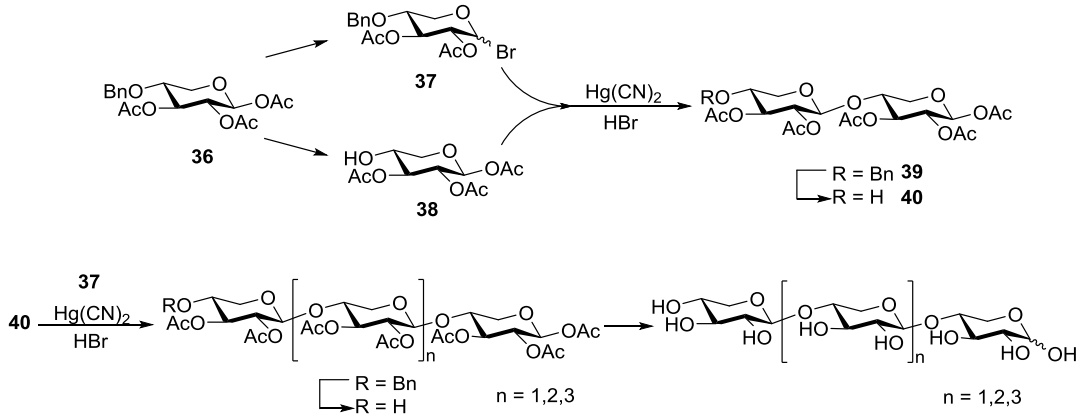
Scheme 11 One-pot synthesis of the tetrasaccharide **35** performed by Huang et al.[140]

A disadvantage of this method is that *p*-TolSCl must be generated *in situ*, due to its limited shelf-life, and therefore a stable, commercially available alternative, sulfenyl chloride, was described by Crich et al.[163] *p*-Nitrobenzenesulfenyl chloride (*p*-NO_2PhSCl) was used in conjunction with AgOTf to effectively activate several thioglycosides at – 78 °C in dichloromethane yielding both 1,2-*cis* and 1,2-*trans* adducts as major products depending on the choice of protecting groups.

1.8 Synthesized xylans

The chemical synthesis of well-defined linear xylans has been described by several research groups in the last decades, by using diverse synthetic approaches and glycosylation methods. The first example is the synthesis of xylobiose, described in 1961,[164] and carried out via Koenigs-Knorr condensation between benzyl 2,3-di-*O*-

benzyl-D-xylopyranoside and a peracetylated xylosyl bromide in the presence of $Hg(CN)_2$. In a similar fashion, twenty years later, Hirsch and Kovac synthesized for the first time a series of oligoxylans of different lengths using a sequential approach (Scheme 12).[165,166] They used 1,2,3-tri-*O*-acetyl-4-*O*-benzyl-β-D-xylopyranose (**36**) as the common building block, which was activated as a donor via reaction with HBr (*i.e.* **37**), or selectively deprotected at position 4 to function as acceptor (*i.e.* **38**). The condensation reaction achieved both α and β products **39** (α/β), with a slight preference for the β adduct (α:β ≈ 1:1.5), which they separated from each other by column chromatography in order to proceed to hydrogenolysis to give **40** and further glycosylation. A final, complete saponification yielded the (1→4)-β-D-pentaxyloside (Scheme 12). Previously, with the same technique, they obtained a series of methyl β-glycosides of xylo-oligosaccharides up to the xylohexaoside.[167]

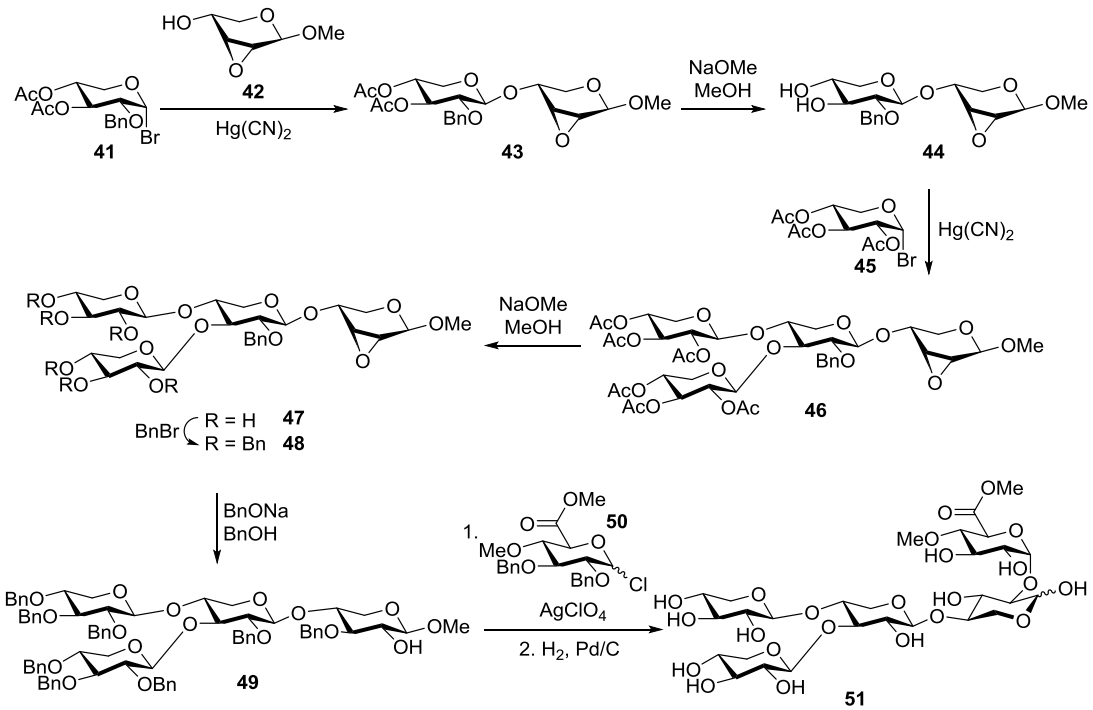


Scheme 12 Hirsch and Kovac strategy for the synthesis of oligohomoxylans

More recently, a blockwise approach was chosen by Takeo et al. to synthesize a series of xylo-oligosaccharides up to a xylodecaose, using thio-xylobiosides, conveniently protected as building blocks, and NIS in combination with silver triflate as promoting agents.[168]

With regard to branched xylans, they are highly represented in lignocellulosic material, as observed after isolation from different natural sources,[10,169] although very few examples of their chemical synthesis are reported in the literature – probably due to the difficulties encountered in obtaining differentiate-protected xylose building blocks.

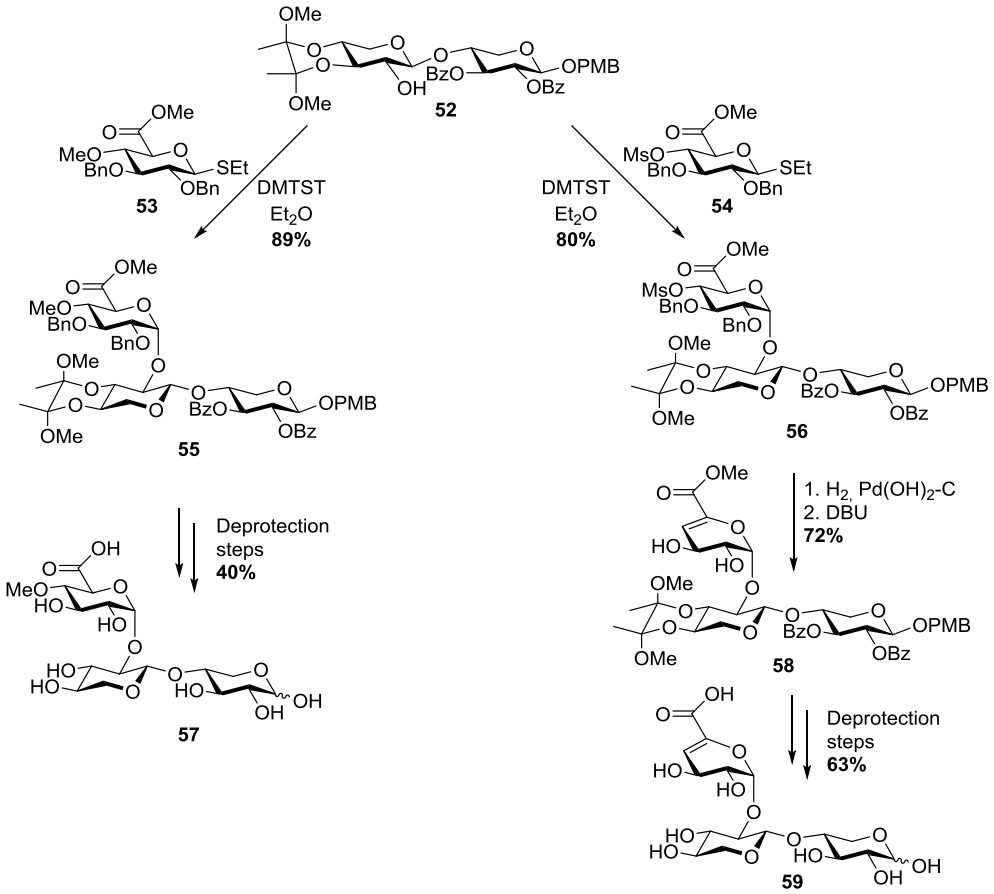
A way to overcome this difficulty was proposed by Hirsch and Kovac, who started a stepwise synthesis of a model oligoxylan from methyl 2,3-anhydro-β-D-ribopyranoside (**42**) (Scheme 13).[170] This epoxide was coupled with xylosyl bromide **41** selectively deprotected at positions *O*-3 and *O*-4. At those positions the disaccharide was condensed with the xylose residues to yield the branched xylotetraose **46**. The epoxide was then stereoselectively opened with benzyl alcohol following the Fürst-Plattner rule, in order to get a *xylo*-configuration in **49**.



Scheme 13 The synthetic strategy adopted by Kovac and coworkers for the synthesis of a branched 4-*O*-methyl-α-D-glucuronic acid-containing xylotetraose.[171]

The free hydroxyl group was coupled with the glycosyl chloride **50** in the presence of silver perchlorate and 2,4,6-collidine resulting in formation of the α adduct as the main product. The complete deprotection produced a model branched xylooligosaccharide **51** constituted by a xylotriose backbone bearing a branching xylose unit linked with a β-(1→3) bond and, for the first time, a α-(1→2) linked glucuronic acid.[171]

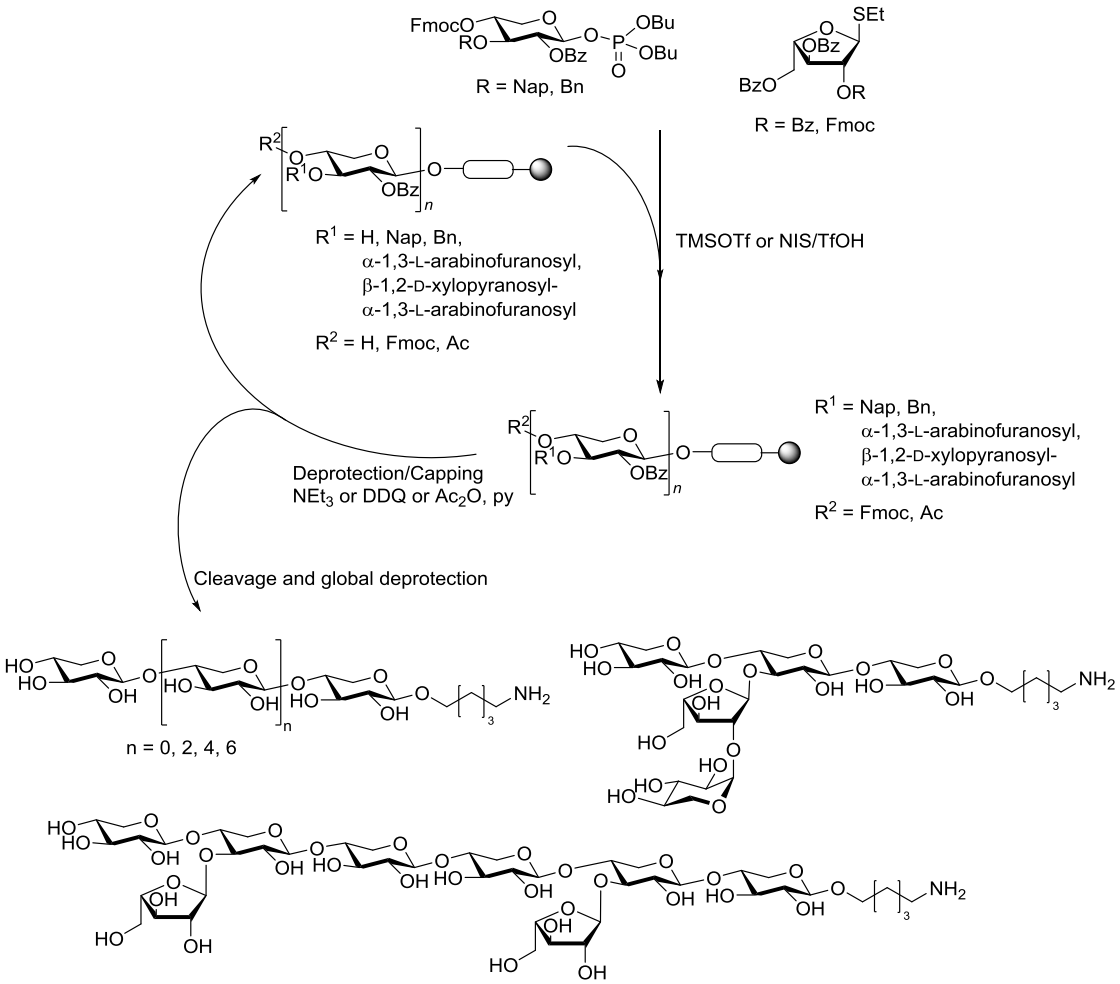
In 2001 Oscarson and Svahnberg[172] showed the synthesis of two uronic acid-containing trisaccharides, **57** and **59**, related to the glucuronoxylan decomposition in wood due to enzymatic cleavage and Kraft pulping (Scheme 14).



Scheme 14 Synthesis of uronic acid-containing trisaccharides **57** and **59** by Oscarson and Svahnberg[172]

The two target molecules were characterized by the common xylobioside backbone **52** acting as an 2'-OH acceptor and coupled with two glucuronic acids differentiated at the 4-position by the presence of a methoxy (in **53**) or a mesyl (in **54**) group. The glycosylations were performed in ethereal solution and in the presence of DMTST as the promoter to achieve only the α-linked products. The trisaccharide **56**, bearing the *O*-mesyl substituent, underwent β-elimination to afford the α,β-unsaturated uronic derivative **58** (HexA derivative), which transformation is relevant since it has been observed in the Kraft pulping process of wood.

A set of arabinoxylan fragments have recently been prepared by Seeberger and coworkers[173] with the assistance of an automated oligosaccharide synthesizer developed in the group (Scheme 15). Two protected xylosides with different protection patterns at position *O*-3 (benzyl as permanent group or (2-naphthyl)methyl substituent for temporary protection to allow for arabinose substitution) were synthesized to serve as building blocks for the linear xylan backbone. Perbenzoylated and 2-Fmoc-L-arabinofuranosides were used for branching. The glycosylation method chosen for these syntheses employed glycosyl dibutylphosphates, activated via TMSOTf or NIS/TfOH and linked at the non-reducing end to a linker-functionalized resin which would provide the oligoxylans as a conjugation tool. The sequential synthesis provided, in short times and overall yields of 7-43%, a collection of eleven arabinoxylan fragments either linear (from a xylobioside to a xylooctaoside) or presenting a naturally occurring pattern of substitutions that included single α-1,3-linked L-arabinofuranosyl and β-1,2-D-xylopyranosyl-α-1,3-L-arabinofuranosyl residues.



Scheme 15 The automated-synthesis method and examples of the synthesized xylans

The glycans were printed on microarray slides and used for probing the binding specificity of anti-xylan monoclonal antibodies.

2 RESULTS AND DISCUSSION

In the following section the results obtained during the course of the studies will be described and discussed. Two different projects were developed, and they both involved the synthesis of model substrates for enzymes related to the biomass degradation processes. In the first part, the synthesis of well-defined glucuronoxylans was designed, aiming at using them for the characterization of both α-glucuronidases and xylanases. In the second part three esters of glucuronic acid, mimicking LCCs, were synthesized and employed as model substrates in kinetic characterization of a novel glucuronoyl esterase.

2.1 Synthesis of Glucuronoxylan Fragments

Two different branched pentaxylans were chosen as targets for this project as mimics of glucuronoxylans fragments (Figure 11). They consisted of a common pentasaccharide formed by β-(1→4)-linked xylose units, and they differed from each other by the substituent on the 4 position of the glucuronic acid branch, which in both cases is linked with a α-(1→2) bond on the penultimate residue from the non-reducing end. In fact, both glucuronic acids and 4-*O*-methyl glucuronic acids have been proved to exist as branches in natural glucuronoxylans.[21] This specific structure was chosen since it represents a valid substrate for glucuronoxylanases, known to hydrolyze the β-(1→4) bond of the xylan backbone exclusively in the presence of a glucuronate branch.[31] The length of the substrates is sufficient to give good enzyme activity and the position of the branch will allows for verifying the substrate specificity of the xylanases. Furthermore the target molecules can be used to investigate the substrate specificity of enzymes belonging to the GH115 family,

described to be able to release a glucuronic acid branch linked to internal xylose residues.[39] Whereas the glucuronidases belonging to the GH67 family cut only terminal residues,[28,29] and therefore the targets could be investigated after the preliminary action of a xylosidase releasing the xylose residue at the non-reducing termini.

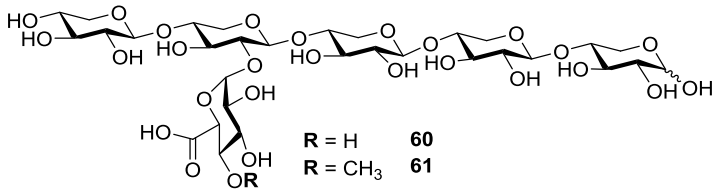


Figure 11 Structure of the two branched pentaxylans chosen as targets in this project

The pentaxylan has been synthesized following an iterative glycosylation process, inspired by the work of Huang and coworkers,[140] using phenyl thioglycosides as building blocks. The advantages of using this procedure are the great stability of these substrates, together with the ease of selective activation, and the possibility of using building blocks with similar reactivity.

2.1.1 Retrosynthetic pathway

The retrosynthesis shown in Figure 12 schematizes the approach that has been chosen to achieve the synthesis of **60** and **61**.

The first scission was foreseen to happen in concomitance with the branch unit and the focus then shifted to the linear pentasaccharide. It was necessary to have it fully protected, to facilitate the glycosylation and the regioselectivity of the final coupling, which gives rise to **62** as a *O*-benzyl perbenzoylated pentaxylan. The only exception is a levulinyl ester regioselectively installed at the *O*-2 position on the fourth residue, in order to have it selectively removed under suitable conditions. The synthesis of **62** was achieved by a linear approach and the first disconnection

would occur from the reducing end to have the glycosyl donor **63** and the acceptor **64**.

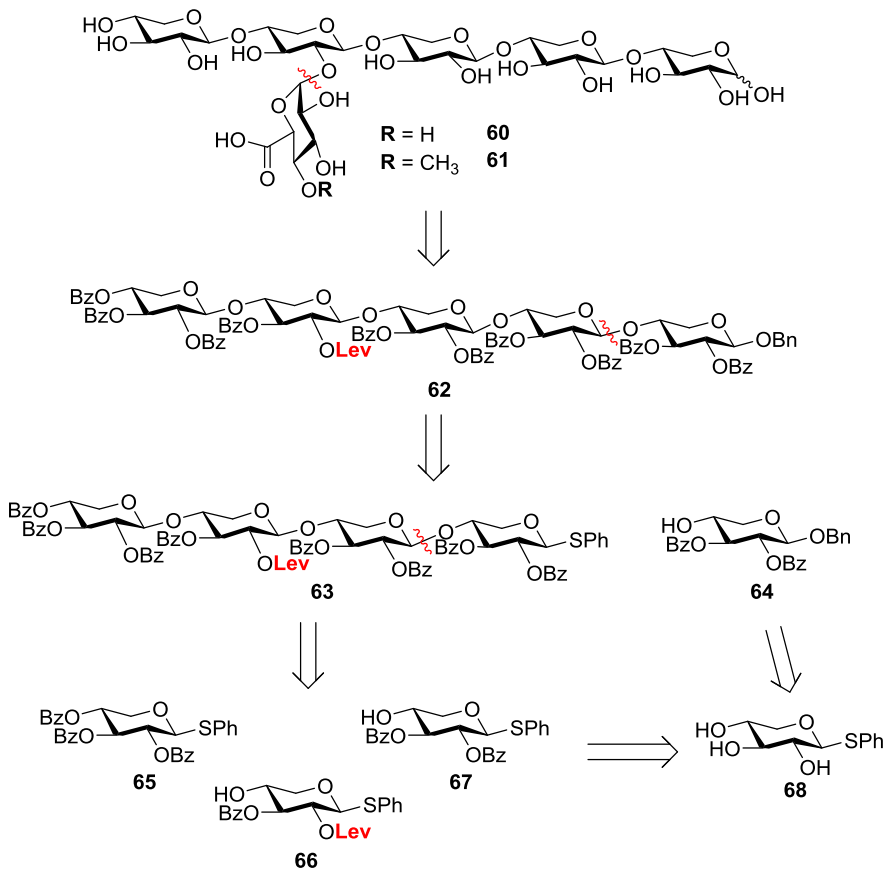


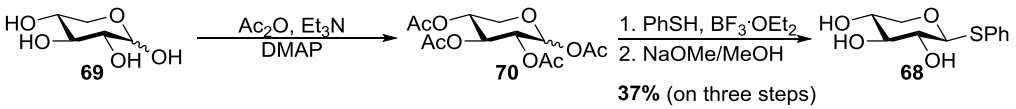
Figure 12 Retrosynthetic pathway followed for the synthesis of target compounds **60** and **61**

In the same fashion, **63** could be split into the corresponding trisaccharide and the acceptor **67**, and so forth until the coupling between the perbenzoylated donor **65** and the alcohol **66** bearing differentiated protection at positions 2 and 3. The monosaccharidic building blocks **64** – **67** could be synthesized altogether starting from the same molecule, phenyl 1-thio-β-D-xylopyranoside (**68**).

2.1.2 Synthesis of the monomeric building blocks

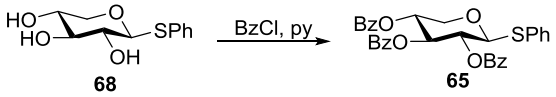
Since the glycosylation process would only contemplate the formation of β-(1→4) xylosidic bonds, the manipulation of building blocks was focused on two features: the protection of position *O*-4 with a selectively removable moiety (a PMB group) and, on the other hand, a stereocontrolling auxiliary group on the 2-position that would assure the right stereochemical outcome during the glycosylation reactions (benzoyl and levulinyl groups).

As aforementioned, the triol **68** represented a common starting material for the preparation of the monomers utilized for the assembly of the selected targets. It was obtained in three steps from commercially available D-xylose by following straightforward and easily scalable procedures. The procedure included a standard peracetylation protocol in the presence of *N*,*N*-dimethyl-4-aminopyridine (DMAP) as catalyst,[174] subsequent thio-glycosylation with thiophenol promoted by $BF_3{\cdot}OEt_2$ and final deprotection with Zemplén conditions (Scheme 16).[175]



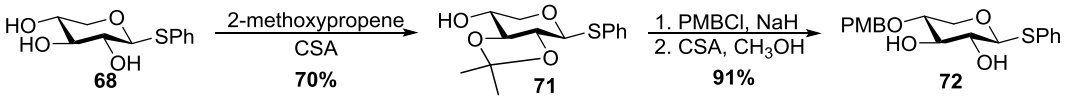
Scheme 16 Synthesis of **68** from commercially available D-xylose according to literature protocols.

A conventional perbenzoylation[176] of **68** afforded the glycosyl donor **65** (Scheme 17), whereas protecting group manipulations were necessary in order to attain the bifunctional building blocks **64**, **66** and **67**.



Scheme 17 Synthesis of phenyl 2,3,4-tri-*O*-benzoyl-1-thio-β-D-xylopyranoside (**65**)

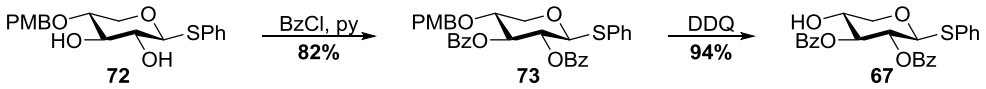
As described above, the PMB ether can be removed in the presence of DDQ, with a SET mechanism which would not compromise the other functionalities.[83] Nevertheless, the regioselective protection of the 4-position is not straightforward due to the similar reactivity towards electrophiles of the other hydroxyl groups. In order to prevent the formation of multiple compounds with consequent decrease of the overall yield and troublesome chromatographic separations, a preliminary protection of the remaining hydroxyl groups was performed (Scheme 18). Initially, the reaction of **68** with 2-methoxypropene in the presence of camphorsulfonic acid (CSA)[177] afforded the 2,3-isopropylidene protected xyloside (**71**) as the major product (70%), even though the 3,4-acetonide and the fully protected adduct, bearing the cyclic acetal and a mixed acetal (1-methoxy-1-methylethyl) in the 4-position, were also isolated.



Scheme 18 Diol **72** was prepared by selective protection and following deprotection of polyol **68**

The protection of the remaining hydroxyl group in a basic environment (NaH and PMBCl), followed by acidic hydrolysis of the acetal moiety, led to the diol **72**.[92]

A standard benzoylation with benzoyl chloride (BzCl) in pyridine yielded **73**, that was isolated, purified and subsequently reacted with DDQ in a mixture of dichloromethane and water (9:1) to give the bifunctional product **67** (Scheme 19).

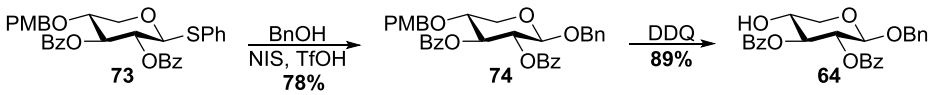


Scheme 19 Synthesis of phenyl 2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (**67**)

The oxidative removal of the ether was generating a byproduct, 2,3-dichloro-5,6-dicyanohydroquinone, that precipitated in the reaction medium causing a tedious

workup. This inconvenience was overcome by quenching the reaction with a buffer solution based on ascorbic acid (0.7%) which was re-oxidizing the hydroquinone and in this way dissolving the emulsions formed.[178]

The fully protected intermediate **73** was employed for the preparation of the glycosyl acceptor **64** which bore a benzyl substituent at the anomeric position. Hence it was used as the reducing end residue of the pentasaccharide in order to avoid any collateral reactivity in the last coupling with the glucuronic acid branch. The glycosylation with BnOH as OH-acceptor was carried out in presence of NIS and TfOH, where the iodonium system promoted the thioglycoside **73** activation, and through neighboring group participation, led to the formation of one single anomer **74**. Despite the presence of a Lewis acid the PMB group survived and could be successfully removed by the oxidative cleavage conditions described above (Scheme 20).



Scheme 20 Synthesis of di-benzoylated acceptor **64**

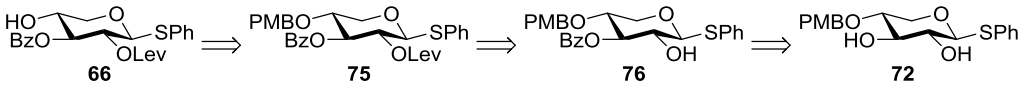
As regards the synthesis of phenyl 3-*O*-benzoyl-2-*O*-levulinyl-1-thio-β-D-xylopyranoside (**66**), it exhibited challenges related to the discrimination between the *O*-2 and *O*-3 positions.

Selective acylation of the diol xyloside 72

The relevance of distinguishing among the several secondary hydroxyl groups on a monosaccharide has been discussed earlier in this work.[73] This feature assumed a pivotal role during the project, since the coupling between the branch and the linear xylan needed to occur on one specific position, *O*-2. Therefore, a specific protecting group was needed on that position, that could be removed selectively in

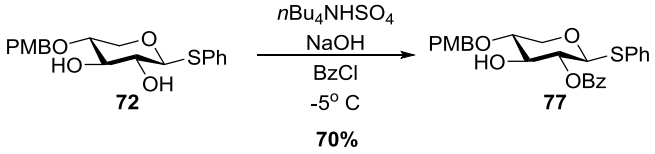
the presence of benzoyl groups and that could assert neighboring group participation in the glycosidic bond formation. The group of choice was the levulinyl ester, since it was reported to be cleaved selectively by reaction with hydrazine, not affecting other ester groups on the same molecule.[179,180]

At this point, the main challenge consisted in the regioselective introduction of the abovementioned protecting group on the diol **72**, preferred as the starting material for the preparation of the desired compound **66** (Scheme 21).



Scheme 21 Pathway initially envisaged for the synthesis of **66**

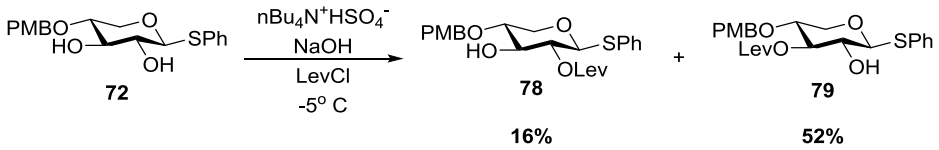
Selective acylation has been thoroughly explored in carbohydrate synthesis,[181–184] and extensively used as a tool in the synthesis of complex oligosaccharides.[92,185,186] Hence, a plethora of different experimental procedures was available for the selective benzoylation on saccharides,[187–190] although only a few examples concerned xylopyranosides,[183,190,191] and fewer thioglycosides.[192] Since the general trend of reactivity for β-xylosides is reported to be O4 > O3 > O2,[184] the initial strategy envisaged a selective monobenzoylation on the 3-position, and a subsequent levulination (Scheme 21). However previous experience in our group proved the benzoylation on compound **72** to be either not selective or preferring the *O*-2 position, as shown in Scheme 22.



Scheme 22 Regioselective benzoylation of the diol **72** under the conditions reported by Garegg and coworkers[192]

The reaction occurred in a two-phase system (CH_2Cl_2/H_2O) so that the base (NaOH) was dissolved in the water phase, whereas the sugar and the acylating agent (BzCl) remained in the organic phase, in the presence of a phase transfer catalyst (nBu$_4$NHSO$_4$) that enabled the interaction among the species.[192]

The high regioselectivity observed suggested to apply the same protocol, using levulinyl chloride instead of benzoyl chloride, on the same substrate but the conversion was poor and the main product was the 3-adduct instead (Scheme 23).

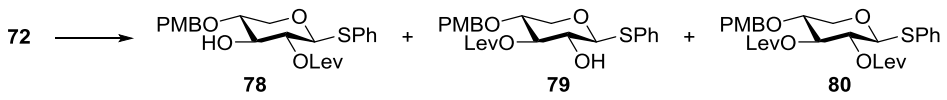


Scheme 23 Attempt at the regioselective levulination of **72**

A possible explanation for those seemingly contradictory results could be found in the well-documented tendency of benzoyl groups to migrate[73,193] hypothesizing that the ester would initially be formed at the more reactive hydroxyl group on *C*-3 and then migrate to *O*-2. On the other hand, the Lev group is tendentially less prone to migrate[89] explaining that the 3-OLev adduct **79** was the main product observed. Furthermore, levulinyl chloride (LevCl) is not commercially available and it was freshly prepared according to literature protocols[194,195] that could indicate lower stability as compared to the shelf-stable benzoyl equivalent and thus compromising the conversion process.

This preliminary experiment paved the way for the optimization of the conditions from both yield and regioselectivity perspectives as reported in Table 1.

Table 1 Screening of reaction conditions for the installation of the Lev group on **72**



	Reaction conditions	78	79	80
1	nBu$_4$NHSO$_4$, NaOH, LevCl	16%	52%	/
2	nBu$_4$NHSO$_4$, NaOH, Lev$_2$O	traces	traces	/
3	LevCl, Py, 0 °C	/	/	/
4	LevOH, DCC, DMAP	10%	27%	9%
5	Me$_2$SnCl$_2$, LevCl, ^{i}Pr$_2$NEt	/	/	/
6	Ag$_2$O, LevCl	traces	traces	/

After the first experience (entry 1) and the speculations about the reactivity of the acylating agent, a further attempt was made by substituting it with levulinic anhydride (Lev$_2$O) which was reported as the reactant of choice for levulinate formation[179] in place of LevCl that was reported to form lactones under basic conditions.[196] Nonetheless, as shown in entry 2, only traces of the single substituted compounds were isolated and starting material was recovered as the main product.

Then, conventional acylation procedures were tried in order to get a general feeling about the Lev group behavior. However, the acylation with freshly prepared LevCl in Py at 0 °C (entry 3) did not yield any product, possibly confirming the hypothesis brought up by Hassner and coworkers[179] about the abovementioned pseudo-ester formation, sensitive to basic conditions. When the esterification was

performed in a neutral environment, with levulinic acid (LevOH), *N*,*N*'-dicyclohexylcarbodiimide (DCC) as coupling agent and DMAP as the promoter (entry 4),[92] a mixture of the two alcohols and the dilevulinate **80** was afforded after 24 hours, although 50% of the starting material was recovered. An approximate ratio of 3:1:1 in favor of the 3-OLev adduct confirmed the tendency that had been noticed beforehand. It is worth noticing that similar conclusions could be drawn by looking at the literature work cited in this regard, where the same procedure was applied on a galactopyranoside[92] and on a galactofuranoside,[197] although in both cases the preference for the 3-position is more pronounced.

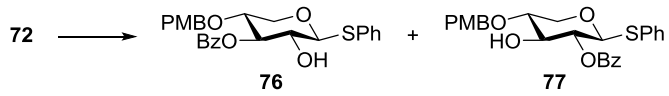
Eventually, metal-catalyzed esterification protocols were explored and applied to the diol of interest (entry 5 and 6), but no remarkable results were obtained. Surprisingly, the utilization of a tin complex in entry 5, which have been previously applied to xylopyranosides in the presence of BzCl,[183,187] did not afford any product of esterification. The TLC analysis of the reaction showed total conversion of the starting material but, after the acidic work-up (3% HCl), the chromatographic purification did not isolate any product, beside starting material.

Since the experiments performed with LevCl, were adapted from benzoylation protocols, a few more procedures were investigated involving selective benzoylation, with methods suggested to direct the reaction towards 3-OH on pyranosides (Table 2).

In 2015 Evtushenko[182] explored the reactivity of transition metals complexes involved in the regioselective benzoylation reaction on 4,6-*O*-benzylidene protected glycopyranosides, and in this context he found out that two complexes, $MoO_2(acac)_2$ and $Cu(CF_3COO)_2$, were selective towards 3-benzoate formation in most cases. Despite the well-documented thiophilicity of both molybdenum and copper, those conditions were applied to **72** and described in entries 2 and 3 of

Table 2. As expected, there was no reactivity in presence of $Cu(CF_3COO)_2$, while $MoO_2(acac)_2$ was more reactive, leading to the 2-benzoate as the major product.

Table 2 Selective benzoylation reaction conditions



	Reaction conditions	76	77
1	nBu_4NHSO_4, NaOH, BzCl[a]	23%	70%
2	$MoO_2(acac)_2$, 2,4,6-collidine, BzCl	8%	21%
3	$Cu(CF_3COO)_2$, 2,4,6-collidine, Bz_2O	/	/
4	1-BBTZ, Et_3N	22%	61%

a. Experiments performed by fellow PhD student Maximiliam Böhm and therefore not reported in this work

Previously,[188] Evtushenko stated that no intermediate complexes were formed between *trans*-vicinal hydroxyl groups, and therefore the complex could be formed between *O*-4 and *O*-3, especially on β-glycosides, favoring the esterification on the 3-position in galactopyranosides. This does not occur on xylopyranosides, where all the vicinal secondary alcohols are indeed *trans* to each other. Consequently, the only plausible intermediate complex is the one with Mo linked to sulfur and *O*-2, induced by the high thiophilicity of Mo, leading to the formation of **77** (Figure 13).

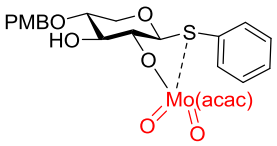


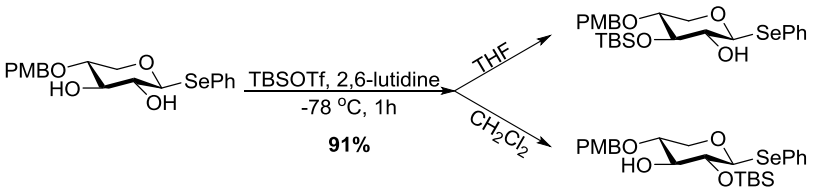
Figure 13 Hypothesis of intermediate complex for the formation of **77**

Eventually the selective benzoylation was attempted by reacting diol **72** with the coupling reagent 1-(benzoyloxy)benzotriazole (1-BBTZ) in the presence of Et_3N as a base (entry 4). The reagent 1-BBTZ was synthesized, according to literature protocol, by reaction of benzoyl chloride with commercially available 1-hydroxybenzotriazole.[189] The reaction outcome confirmed the general trend in regioselectivity with the preference for the adduct **77**, obtained in 61% yield.

Accordingly, this approach to establish a regioselective procedure for the acylation of diol **72** did not lead to the expected results. The differentiation between the two alcohol groups was not always obvious and, by any means, it was not induced in the expected or desired direction. For those reasons, together with a possible instability of the levulinating agents of choice, a new route towards the synthesis of building block **66** was approached.

Solvent-dependent regioselective silylation

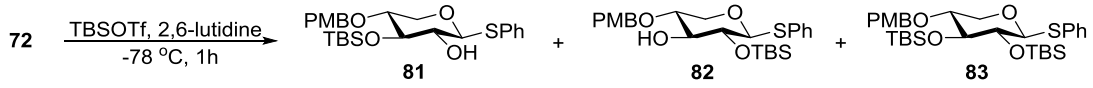
In 2000 Nicolaou and coworkers[198] reported the selective silylation on a selenium-analogue of **72** with *tert*-butyldimethylsilyl triflate (TBSOTf) and 2,6-lutidine and, interestingly, they reported an opposite regioselectivity depending on the solvent used. They obtained primarily a 2-OTBS adduct for the reaction in dichloromethane and silylation at the 3-position when the reaction occurred in THF, with a 91% yield in both cases (Scheme 24).



Scheme 24 Selective, solvent-dependent protection of diol as reported in the literature[198]

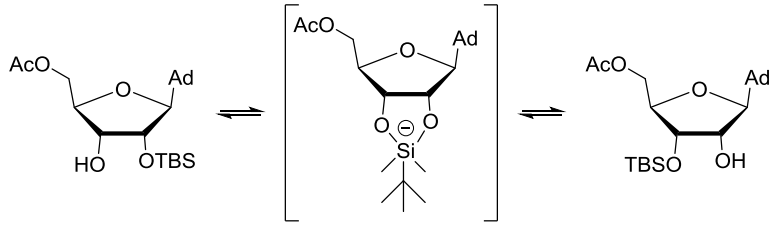
The noticeable regioselectivity with this method and the high similarity of the substrates led to the idea of employing a temporary silyl group protection to bypass the difficulties encountered so far in differentiating among the secondary alcohols of xylopyranosides. Therefore, analogous conditions were applied to **72**, as shown in Table 3, and a remarkable regioselectivity was noticed. In both cases, only one single regioisomer was formed, although in entry 1 the main product was the di-protected product with a yield of 47%.

Table 3 Solvent dependence in regioselective silylation



	Solvent	81	82	83
1	Dichloromethane	/	40%	47%
2	THF	84%	/	/

The authors did not give any possible explanation to this peculiar solvent effect[198] although several studies reported that a kinetic control would induce a reaction on the 2-position, while thermodynamic reaction conditions would lead to the 3-O-silyl adduct. Since the experiments reported in Table 3 were both conducted under kinetic reaction conditions, the discordant results could be explained by implying a silyl group migration.[78,199] The migration of silyl groups is a well-documented phenomenon that has been widely employed in organic synthesis.[200–202] Thermodynamic and kinetic studies on ribonucleosides protected with TBS group, showed that the migration occurs via an intermediate bearing a pentacovalent silicon atom (Scheme 25) and that the process is accelerated by the presence of a base and a protic solvent.[201]



Scheme 25 Proposed mechanism for migration of TBS group. Adapted from Jones and Reese[201]

On this basis it could be assumed that the presence of a polar solvent like THF, although aprotic, and 2,6-lutidine as a base could lead to the migration of the TBS group from the kinetically favored 2-position to the thermodynamically favored 3-position to give **81**. Whereas the typical kinetic conditions (low temperature, short reaction time, apolar solvent) applied for the reaction in dichloromethane led exclusively to the formation of the 3-hydroxyl product **82**.

The promising results observed in Table 3 encouraged the progress towards the synthesis of **66** by using the TBS-protected sugars as an intermediate building block. The following step would be the acylation of the free hydroxyl group on **81** and **82** with a levulinate and a benzoyl moiety, respectively.

It is worth noticing at this point that NMR studies on the two compounds showed how the presence of such a bulky substituent influenced the conformation of the molecule, as widely reported for persilylated monosaccharides.[203] Interestingly, one TBS group on the 3-position generated a distortion of the chair for **81**, which assumed a conformation in between a half-chair and a $^1C_4$ chair conformation. The values of the coupling constant between H-1 and H-2 visibly decreased (J = 3.0 Hz). Similar values could be measured between H-2 – H-3 (J = 4.4 Hz) and H-3 – H-4 (J = 4.4 Hz) indicating a distortion but not a complete inversion of the chair (Figure 14).

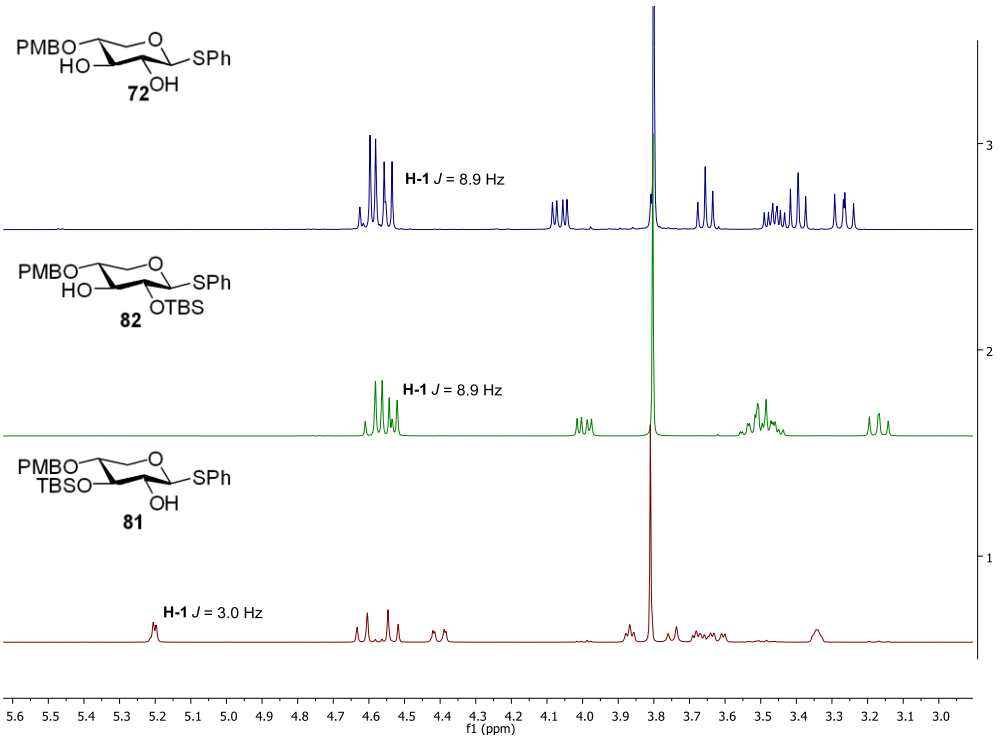
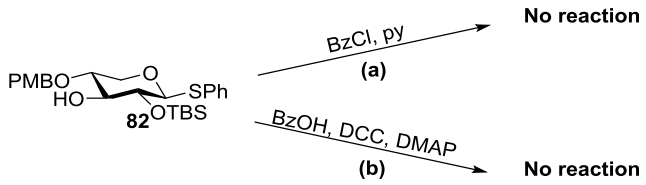


Figure 14 Comparison between ^{1}H NMR spectra of the products of selective silylation and the starting material

On the other hand the conformation of the 2-silylated product **82** was consistent with a $^{4}C_{1}$ chair, according to the *J* values of the anomeric proton (*J* = 8.9 Hz) and the coupling constants among the vicinal protons on the saccharidic moiety, that coincided with axial-axial interactions (Figure 14, and Experimental section). It could be explained by the relative position of the two bulky groups, PMB and TBS, which would have a destabilizing 1,3-diaxial interaction in case of a chair inversion. Therefore no conformational changes occurred since the most stable conformation for **82** is a $^{4}C_{1}$ chair.

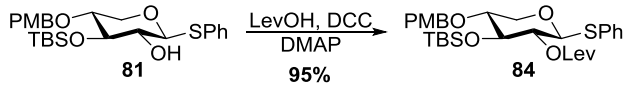
The immediate consequence was the inability to place the benzoyl group on the free alcohol, owing to the steric hindrance of the bulky silylated substituent on the vicinal equatorial position. Neither standard benzoylation conditions (BzCl in

pyridine, Scheme 26a) nor the coupling with benzoic acid in the presence of DCC and DMAP (b) achieved the substitution on the hydroxyl group of **82**.



Scheme 26 Failed attempts of benzoylation of substrate **82**

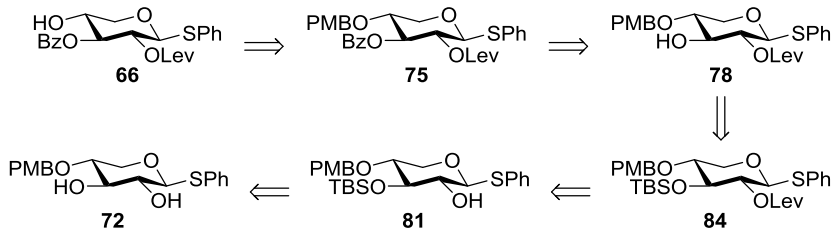
On the contrary, the protection of the alcohol **81** with the levulinyl group went to completion after reaction with LevOH, DCC and DMAP although the free hydroxyl group, in the actual conformation, was (pseudo-)axial and therefore less reactive.[204] Hence longer reaction times or the use of more equivalents of the reactants were necessary to afford the desired fully-protected xylopyranoside **84** in an excellent yield (Scheme 27).



Scheme 27 Levulination of 2-hydroxyl compound **81**

Final steps for the preparation of the levulinated building block 66

The alternative synthetic pathway to afford the building block **66**, which goes through the regioselective installation of a temporary silyl substituent, solved the problems related to the inconvenient poor selectivity of the acylation step, although it added two additional steps to the overall strategy (Scheme 28).



Scheme 28 New retrosynthetic pathway for the synthesis of **66**

In order to reach the target molecule from the fully protected intermediate **84**, three more steps are required. First, the removal of the silyl group was not trivial owing to the presence of both an acid-sensitive ether (PMB group) and a base-sensitive protecting group (Lev group) (Figure 15). Notwithstanding that silyl ethers are well-known to be sensitive to the fluoride ion, due to the higher affinity for fluorine compared to oxygen (Si-F bond strength is 113 kJ/mol greater than the Si-O bond). Therefore they are widely reported to be selectively removed in the presence of a fluoride source.

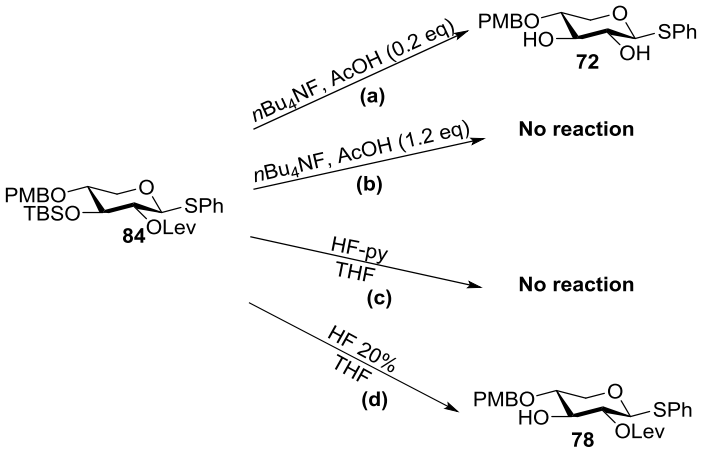
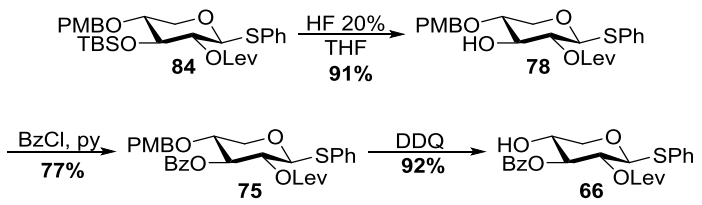


Figure 15 Attempts at TBS-deprotection

Mild conditions are usually applied for the removal of a TBS group, such as tetrabutylammonium fluoride (TBAF) together with acetic acid in catalytic amounts. When those conditions were applied to compound **84**, the main product

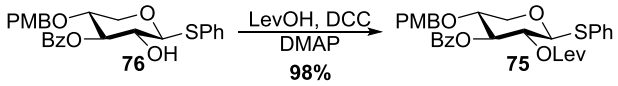
recovered was the diol **72** (Figure 15a), meaning that the conditions were not sufficiently mild to avoid the hydrolysis of the levulinate ester, despite the presence of acetic acid, and probably helped by the presence of residues of water in the TBAF. Hence the deprotection was attempted with a stoichiometric amount of acetic acid, in order to neutralize the reaction medium, but, in those conditions, the reaction outcome after 48 hours was only starting material (Figure 15b). Analogous results were obtained by reaction of the silylated product **84** with the complex HF-pyridine (Figure 15c). Finally, the TBS removal was performed by dissolving the starting material in THF, in low concentration, and then adding a 20% solution of HF in water (Figure 15d).

Eventually, the final 4-hydroxyl building block **66** was prepared by conventional benzoylation and successive oxidative cleavage of the PMB group by DDQ action (Scheme 29).



Scheme 29 Ultimate three steps for the synthesis of **66**

It is worth mentioning that compound **75** was also synthesized from the monobenzoylated product **76** by condensation with LevOH, under the abovementioned reaction conditions, in an almost quantitative yield (Scheme 30).



Scheme 30 Levulination of the intermediate **76**

2.1.3 Assembly of the pentasaccharide

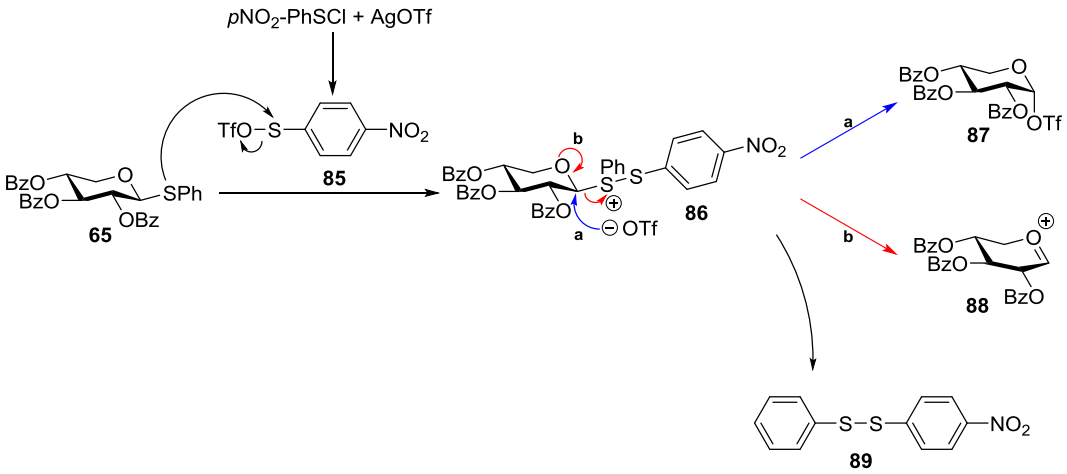
As aforestated, the advantage of adopting a linear, iterative synthetic strategy relies in the optimization of the process by using a single anomeric substituent for all the building blocks and the same glycosylation conditions repeatedly. A prerequisite for such a strategy is the pre-activation of the thioglycosyl donor by reaction with a promoter and the subsequent addition of the acceptor. This method has been explored in recent years and several promoter systems have been proposed.[140,205,206] In the wake of those studies, a novel strategy has been developed in our laboratories. The pre-activation method using thioglycoside/PhSOTf, originally developed by Martichonok & Whitesides,[207] has been modified by using *p*-NO_2PhSCl as the promoter, which is commercially available.[163]

Glycosylation reaction

According to the definition of the pre-activation method, the thioglycosyl donor was activated by the promoter system prior to the addition of the nucleophilic acceptor. The activation occurs by means of *p*-NO_2PhSOTf, which is generated in situ from the corresponding chloride and silver triflate dissolved in dichloromethane and toluene, respectively, at the temperature of -65 °C. In Scheme 31 the plausible mechanism of action for the perbenzoylated glycosyl donor **65** in the presence of the active promoter **85** is represented. After the formation of the disulfidic bond, two possible routes could be followed, which both trigger the departure of the disulfide **89** as an insoluble and colored byproduct.

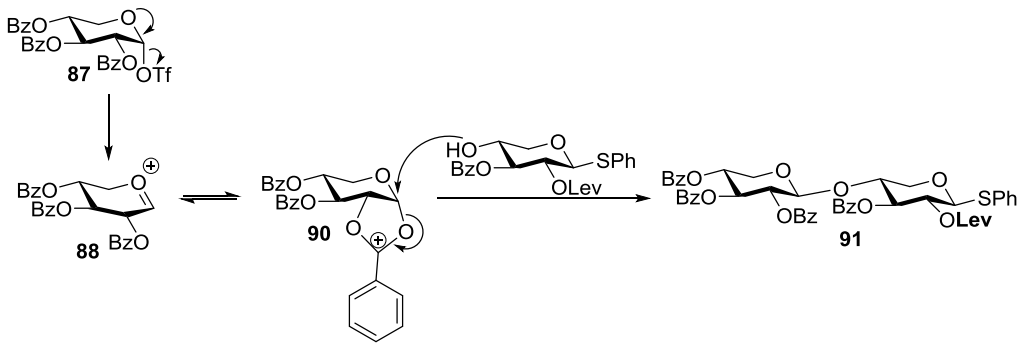
In fact, several recent studies demonstrated the formation of the α-triflate intermediate **87** (Scheme 31, route a),[139,205,208] which has never been isolated, since it degrades rapidly above -50 °C,[205] but it has been observed via low-temperature NMR experiments during triflate-mediated glycosylations.[139] Those studies

revealed that the triflate group influenced the stereochemical outcome by the formation of a transient contact ion pair intermediate enhancing β-selectivity.[209,210] Nevertheless, the debate about the actual kinetically reactive species involved is still open, since divergent theories have been proposed[109] and studies on chemical glycosylation mechanisms are currently going on.[93,94]



Scheme 31 Envisioned activation mechanism of the donor by the promoter system

The alternative would be the direct formation of the oxocarbenium cation **88** (Scheme 31, route b) stabilized by the presence of the participating benzoyl group on 2-position that would generate the acyloxonium ion **90** (Scheme 32).



Scheme 32 Glycosylation with an acyloxonium ion as the reactive specie

Nonetheless **90** is the reactive species in both cases, since the triflate group is released and generates the carbocation with subsequent stabilization by cyclization to form **90**.[209]

The adopted procedure (as schematized in Figure 16) proceeded with the addition of the 4-hydroxyl acceptor **66** dissolved in dichloromethane after the temperature was raised to around -55 °C. For an optimal reaction outcome the temperature was kept in the interval between -55 °C and -50 °C until TLC showed full conversion of the acceptor. The reaction vessel was allowed to warm up to -15 °C and the reaction could be quenched by adding triethylamine to afford exclusively the β-(1→4)-linked disaccharide **91** in 89% yield.

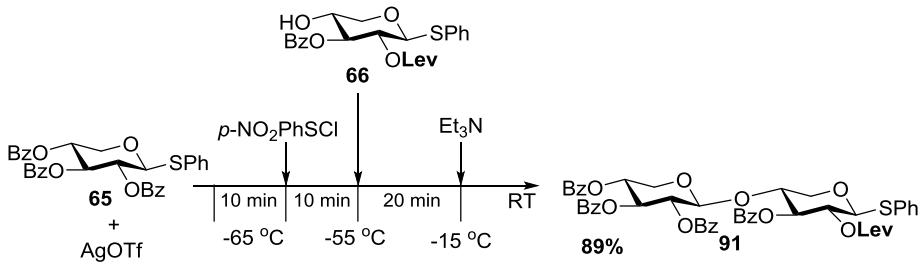
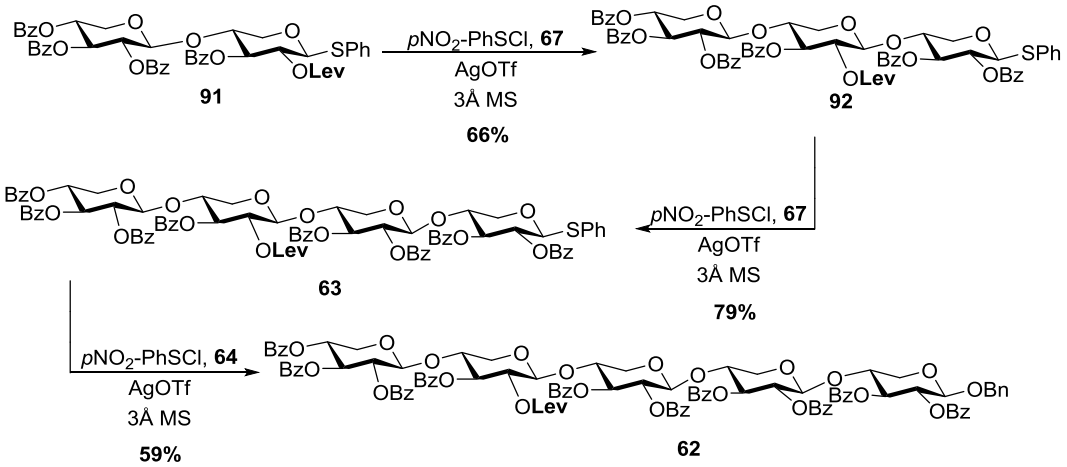


Figure 16 Time-scaled representation of the glycosylation reaction

The use of a stoichiometric quantity of the promoter and a slight excess of the donor, together with a non-interfering byproduct, paved the way for trying the iterative one-pot synthesis procedure, as reported in the literature under analogous conditions.[140] Nevertheless the one-pot approach was not successful for the synthesis of trisaccharide **92**. Several attempts were made, including the addition of a base 2,4,6-tri-*tert*-butylpyrimidine (TTBP) to neutralize the trifluoromethanesulfonic acid produced,[211] but the outcome consisted in a complex and inseparable mixture of saccharides. Hence a sequential strategy was adopted as an alternative and the product of the first glycosylation **91**, was isolated by flash chromatography before it could react with the di-benzoylated acceptor **67**,

according to the same experimental procedure reported in Figure 16, to give the corresponding trisaccharide **92**. Analogous coupling conditions were applied to afford the tetraxylan **63** by reaction of the trisaccharide as donor with the acceptor **67**. As shown in Scheme 33, the intermediate tetrasaccharide **63** was activated at the reducing end by *p*-NO_2PhSOTf and coupled with *O*-glycosidic acceptor **64** to afford the final pentaxylan of interest **62**.



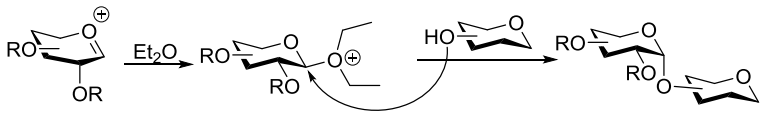
Scheme 33 Sequential reactions performed to achieve pentasaccharide **62**

2.1.4 Synthesis of the glucuronic acid building blocks

The rationale behind the choice of the glucuronic acid moieties is concerned with the characteristics of the glycosidic bond that needs to be formed and the associated reaction conditions. As above-mentioned, the typical linkage encountered in lignocellulosic material is a α-(1→2)-bond and as a result, conditions that generally favor the creation of a α-bond were taken into consideration.

A generally applicable method to obtain α-bonds (1,2-*cis* for *gluco*-configured sugars) has not been developed yet, although a variation of reaction conditions or substrate structures may lead to high α-stereoselectivity.[212]

Among the factors that can influence the stereochemical outcome, the anomeric effect and the solvent are playing a major role. Under conditions where the protecting groups are not interfering with the oxocarbenium ion (neighboring participation), the anomeric effect would favor the formation of the α-glycoside although such effect is less significant due to the irreversible character of glycosylation. On the other hand, some ether-type solvents, *i.e.* Et_2O, THF or dioxane, are found to have a participating effect in some glycosylation processes,[213,214] leading towards the preferential formation of 1,2-*cis* glycosidic bonds. The so-called participating solvent is hypothesized to interact with the oxocarbenium ion to form an equatorial intermediate (Scheme 34), which would be displaced by the hydroxyl group of the acceptor through a S_N2-like mechanism.



Scheme 34 Proposed mechanism for the solvent participation in a 1,2-*cis* glycosylation

Finally, the uronic acid glycosyl donors are generally fairly unreactive due to the presence of an electron-withdrawing carboxylic acid moiety and therefore the use of arming protecting groups such as benzyl was a necessary choice. To reduce the polarity of the compound and limit side reactions the carboxylic acid group was protected as a methyl ester. Eventually inspired by Oscarson's work on the reactivity of glucuronic acid thioglycosides as donors,[57,172,215] the thioethyl group was chosen as the leaving group at the anomeric position. The selected target molecules are displayed in Figure 17. They differ from each other with the substituent at the 4-position which was either a benzyl group in **93**, as with the other secondary alcohols protections, or a methyl ether in **94**, to represent the 4-OMe glucuronate substituents equally abundant in nature on glucuronoxylans.[21]

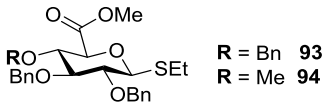
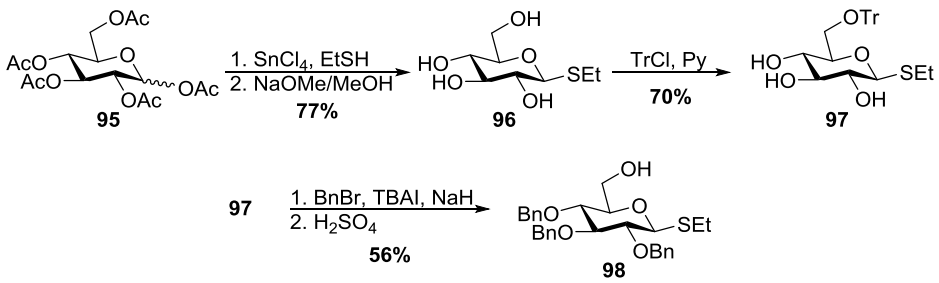


Figure 17 Target glucuronic acid building blocks

The synthesis of methyl (ethyl 2,3,4-tri-*O*-benzyl-1-thio-β-D-glucopyranosid)uronate (**93**) was achieved in a few and straightforward steps starting from commercially available penta-acetylated glucose.

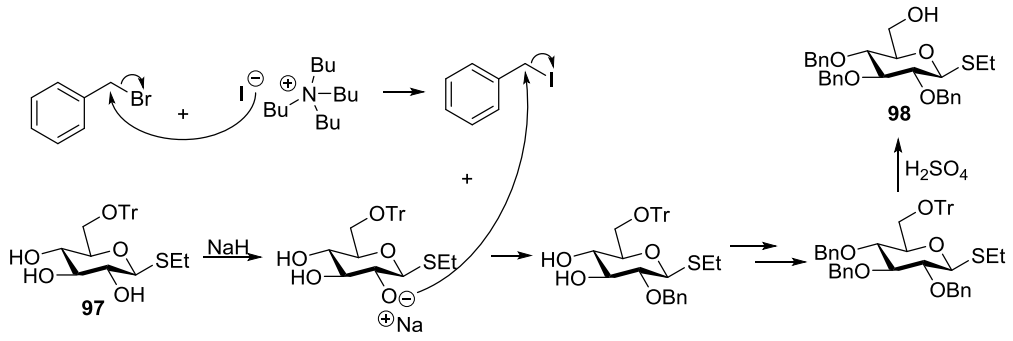
According to a literature protocol,[216] ethyl 1-thio-β-D-glucoside (**96**) was synthesized by condensation of fully acetylated glucose **95** with thioethanol in the presence of $SnCl_4$ as a Lewis acid promoter. As repeatedly mentioned before, an ester group on the 2-position guaranteed the high stereochemical outcome of the glycosylation. The thioglucoside was subjected directly to the standard Zémplen conditions to give the tetraol **96** in an overall yield of 77% (Scheme 35).



Scheme 35 Initial protecting group manipulations in the route for the synthesis of glucuronate glycosyl donor **93**

The following step aimed at the protection of the primary alcohol in order to easily discriminate, afterwards, the hydroxyl group undergoing oxidation. For that purpose, the trityl group was chosen since the bulkiness of the substituent was not compatible with the alkylation of secondary positions. The traditional tritylation conditions yielded the triol **97** in a 70% yield. Benzylation of the remaining hydroxyl groups was achieved via Williamson ether synthesis by reaction with

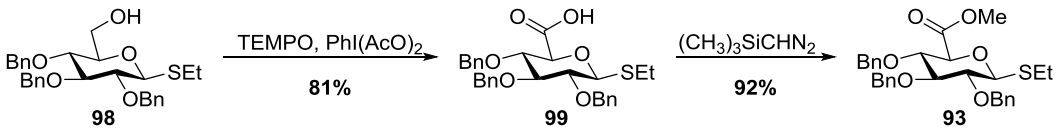
benzyl bromide (BnBr), NaH and tetrabutylammonium iodide (TBAI) that, added in catalytic amount, lowered the reaction time via a halogen exchange process (Scheme 36).



Scheme 36 Benzylation via Williamson ether synthesis in presence of TBAI as a catalyst.

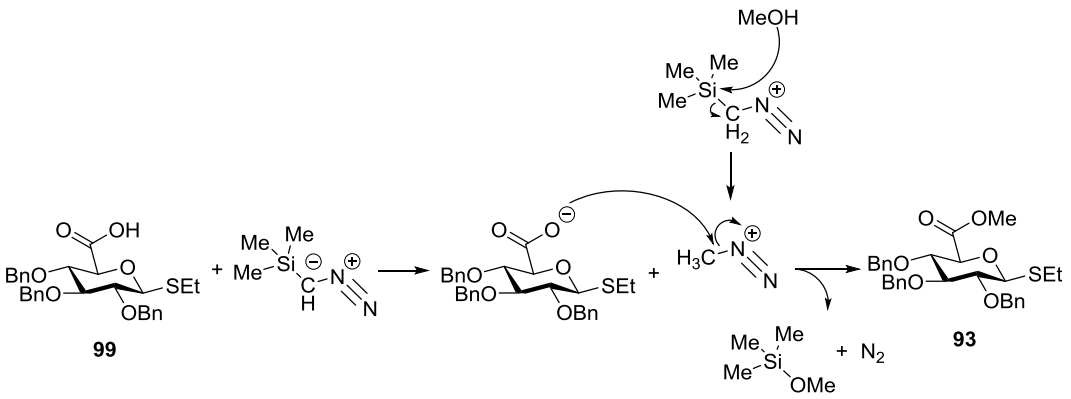
Without further purification, the fully-protected intermediate was dissolved in a mixture of methanol and sulfuric acid to hydrolyze the trityl ether which gave the alcohol **98** in 56% yield over two steps.

At this point, the synthetic route for the synthesis of thioglycosyl donor **93** envisaged the oxidation of the primary alcohol to the carboxylic acid and the subsequent methylation (Scheme 37). As widely reported in the literature,[69,71,72] the oxidation could be performed with a catalytic amount of TEMPO and an excess of (diacetoxyiodo)benzene as the secondary oxidant, despite the presence of a thioether group on the molecule. A short reaction time, with continuous TLC monitoring, and a prompt quenching with aqueous thiosulfate solution prevented the formation of the undesired sulfone or sulfoxide byproducts.



Scheme 37 Oxidation and esterification steps to obtain building block **93**

The carboxylic acid was masked as the methyl ester (Scheme 37) after reaction with (trimethylsilyl)diazomethane (2 M solution in hexane), a shelf-stable, commercially available methylating agent, easier to handle as compared to the non-silylated counterpart.[217] The reaction occurred via the mechanism proposed in Scheme 38, with the *in situ* generation of the very reactive diazomethane species by removal of the trimethylsilyl group in methanol and consequent attack from the carboxylate. The presence of toluene as a co-solvent proved to be essential for the success of the reaction due to the scarce solubility of the reactant in methanol. The driving force for the reaction is the generation of nitrogen gas, which makes the process irreversible.

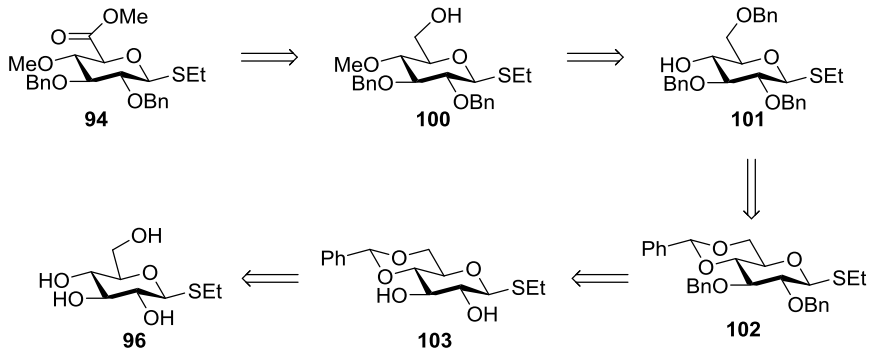


Scheme 38 Mechanism for methylation at the carboxylic acid moiety of **98** with (trimethylsilyl)diazomethane

The above-described methylation gave the methylated glucuronate donor **93** in 92% yield and, from the starting material **95**, a 22% overall yield in 7 steps.

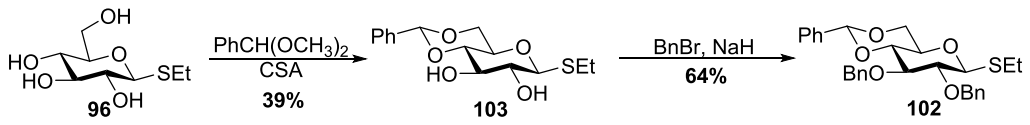
On the other hand, the synthesis of methyl (ethyl 2,3-di-*O*-benzyl-4-*O*-methyl-1-thio-β-D-glucopyranosid)uronate (**94**) required a few more transformations due to the need to methylate regioselectively the hydroxyl group on the 4-position and, afterwards, to release selectively the primary position to undergo oxidation to the carboxylic acid. The retrosynthetic analysis is shown in Scheme 39, where it can be

observed that the strategy envisaged is a two-step oxidation/methylation from the alcohol **100** which would be obtained by methylation of 4-OH and consequent hydrolysis of the primary benzyl ether from the 4-OH intermediate **101**. Compound **101** was obtained by reductive, regioselective opening of the benzylidene group of the fully protected glucoside **102**, which, in turn is derived from the diol **103**. The benzylidene acetal was installed on the tetraol **96**, already utilized for the synthesis of **93**.



Scheme 39 Retrosynthetic strategy for the synthesis of glucuronate glycosyl donor **94**

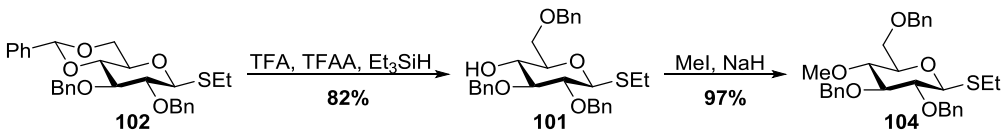
The fully protected intermediate **102** was prepared from the unprotected 1-thio-β-D-glucopyranoside in two steps (Scheme 40). According to standard procedures, the hydroxyl groups in the 4- and 6-position were involved in the formation of the benzylidene acetal by reaction with benzaldehyde dimethylacetal in an acidic environment (CSA). The reaction was stopped after 24 hours but it yielded **103** with 39% yield since starting material was recovered.



Scheme 40 Protection of the hydroxyl functionalities of **96** with ether groups and a cyclic acetal in 2 steps

This transformation was followed by a benzyl protection of the remaining alcohol groups under the above-mentioned conditions.

The reductive ring-opening of the benzylidene moiety on **103** to give the 4-OH glucoside **101** was achieved by reaction with trifluoroacetic acid (TFA), trifluoroacetic anhydride (TFAA) and triethylsilane (Scheme 41).

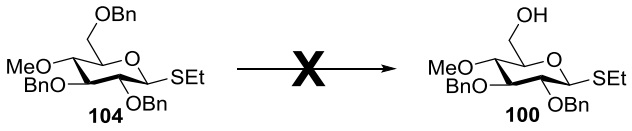


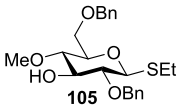
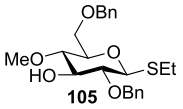
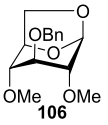
Scheme 41 Regioselective opening of the benzylidene and consequent methylation of the free OH of **101**

The hydroxyl group released by this transformation was methylated via Williamson ether synthesis using NaH as the base and MeI as the methylating agent, yielding the fully protected intermediate **104** in 97% yield. The successive deprotection of the benzyl group on the primary position was challenging owing to the presence of a thioether group at the anomeric position. In fact, standard conditions, such as hydrogenolysis catalyzed by palladium, are known to be ineffective on thioglycosides due to the high thiophilicity of Pd. Hence, different solutions were explored and exposed in Table 4.

A few studies have reported the use of DDQ to hydrolyze the ether bond selectively,[218] and also in the presence of thioethers.[219,220] Those results encouraged two experiments with DDQ to deprotect the benzylated substrate **104**. In the first place, the reaction was performed under analogous conditions to those applied for the PMB deprotection (entry 1) and the main isolated product was, very surprisingly, the 3-OH thioglucoside **105**. A similar reaction outcome was observed after reaction with DDQ under anhydrous conditions and under irradiation with UV light (entry 2).[221]

Table 4 Attempts to remove benzyl group on the 6-position of **104**

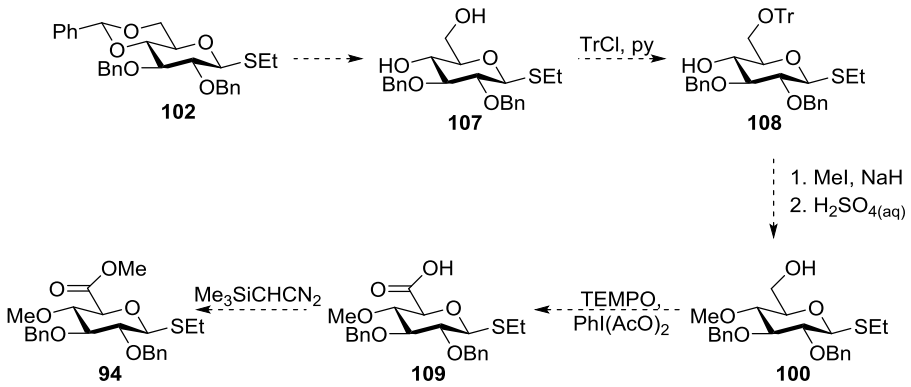


	Reaction Conditions	Main product
1	DDQ in CH_2Cl_2/water	
2	DDQ, hν in anhydrous CH_3CN	
3	$FeCl_3$ in anhydrous CH_2Cl_2	

Fraser-Reid and coworkers developed an alternative method to hydrolyze benzyl ethers using $FeCl_3$ under strictly anhydrous conditions at room temperature. They applied this method to both monosaccharides and complex oligosaccharides, showing its generality and the compatibility with several functionalities, including alkenes and sensitive glycosidic linkages.[222,223] Although such reaction conditions were not applied on thioglycosides, an attempt was made on the fully protected compound **104**, as reported in entry 3. Interestingly, the only isolated product was 1,6-anhydro-2,3-*O*-benzyl-4-*O*-methyl-D-glucopyranoside (**106**). Thus, it can be hypothesized that the deprotection of the primary alcohol occurred first, followed by rapid rearrangement into the bicyclic compound due to nucleophilic attack on the anomeric position, promoted by the Lewis acid $FeCl_3$.

Those unsuccessful results led to the suggestion that a different strategy was necessary to bypass the difficulties encountered in the hydrolysis of benzyl ethers

on a thioglucoside. Therefore a different process was proposed to obtain **94**, based on the synthesis of the same compound accomplished by Oscarson and Svahnberg in 2001.[172] As shown in Scheme 42, the benzylidene acetal in **102** would be completely cleaved, and then the primary alcohol in diol **107** would be selectively protected with the bulky trityl group that, as above-stated, has an excellent regioselectivity towards primary alcohols. The tritylated glucoside **108** would undergo methylation under the same conditions as applied previously, followed by acidic hydrolysis to release the primary hydroxyl group. The remaining steps towards the final target **94**, *i.e.* oxidation with TEMPO and methylation, are expected to proceed as well as for the previous substrate.



Scheme 42 Proposed strategy for the synthesis of **94** based on Oscarson and Svahnberg's work[172]

Unfortunately, the time was not sufficient to complete this second synthesis before the end of this PhD project. Nevertheless, in the meantime, the 4-OMe methyl glucuronate **94** was prepared by a fellow PhD student according to the suggested strategy.

2.2 Glucuronate esters as mimics of LCCs

This section will focus on the synthesis of model substrates for the GE enzymes, which are specific for the hydrolysis of glucuronate esters in the cross-links between glucuronoxylans and lignin in lignocellulose, as described in the introduction. Subsequently several GEs have been tested with the above-mentioned substrates, in order to determine the substrate specificity in relation to both the alcohol part and the 4-*O*-methyl substituent. The design of the targets was inspired by the previous literature reporting studies on the enzymes from the GE family that have been isolated and biochemically characterized so far.[45,46,48,224] Further information were obtained by the crystallization and the solution of the structures by X-ray crystallography for two different enzymes, the *Hypocrea jecorina* Cip_2GE[225] and the *Sporotrichum thermophile* GE2.[226] The latter was even crystallized in complex with a substrate analogue, methyl 4-*O*-methyl-β-D-glucopyranuronate, giving relevant insights into the substrate interaction within the active site.

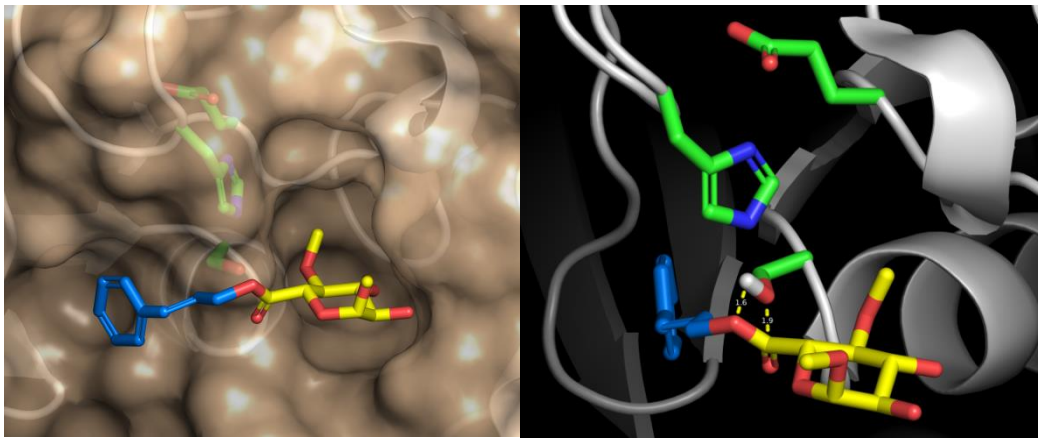
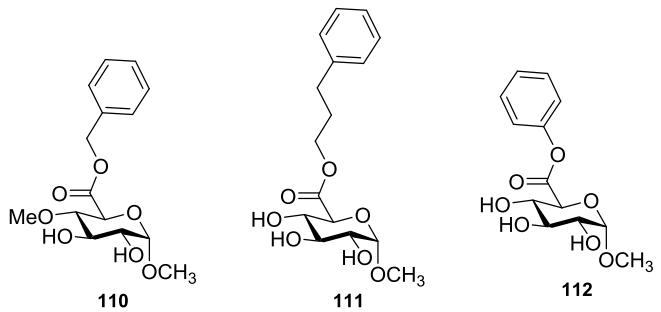


Figure 18 3D reproduction of a model glucuronate in the active site of a GE

A docking study, based on those crystallographic data, has been performed at Novozymes (Figure 18) to investigate the interaction of the substrate with the catalytic site. Furthermore the simulation was used to foresee the position, and potentially the role, of key structural determinants like the 4-OMe group or the aromatic moiety.

On the picture on the left it can be observed how the aromatic part is exposed on the enzyme surface where a hydrophobic interaction can occur with an aromatic residue. On the right, in Figure 18, the catalytic triad (Ser-His-Glu) and the close interaction of the serine residue with the model substrate are highlighted. It is worth noticing that the methoxy group on 4-position points towards the inside of the binding site, deducing that it might generate an additional interaction increasing the binding affinity.

Consequently, three aromatic and alkyl-aromatic esters of methyl α-D-glucuronosides were chosen as target compounds (Figure 19) in order to mimic α-, ɣ- and phenyl ester LCCs (**110** – **112**). While the presence of the first two moieties have been already demonstrated in nature,[13,16] the latter serves to resemble the existence of a phenyl ester as a minor LCC component.

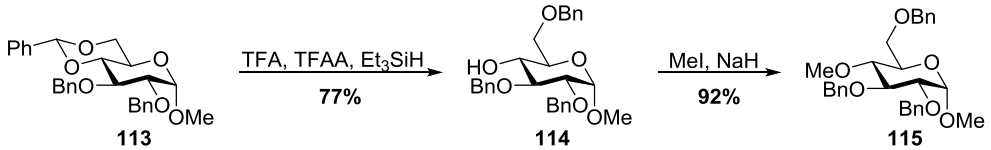


110 **111** **112**

Figure 19 Target glucuronates

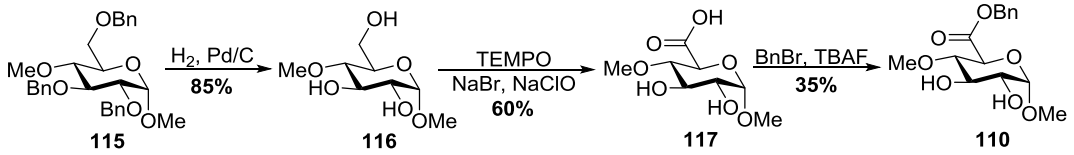
The synthesis of the two thioglucuronate donors **93** and **94** reported in the previous paragraph had a crucial role in the choice of the synthetic route for the

abovementioned targets, since the starting materials would differ only in the anomeric group and the configuration (β-SEt *vs* α-OMe). Otherwise an identical strategy could be used for the protecting group manipulations leading to the free 6-OH to be oxidized and esterified. For the synthesis of benzyl (methyl 4-*O*-methyl-α-D-glucopyranoside) uronate (**110**) a similar approach as to the preparation of thioglucuronic acid **94** was employed. In this case, the presence of a stable and unreactive group at the anomeric position made the entire synthesis more straightforward. The final target was prepared in five steps from the commercially available glucoside **113**.



Scheme 43 Initial synthetic manipulation steps for the preparation of **110**

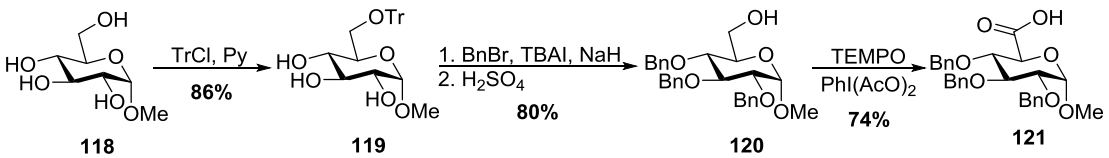
Scheme 43 presents the selective ring-opening of the benzylidene acetal in the fully-protected starting material **113** to give exclusively the 4-OH product **114**, and the subsequent methylation under standard conditions to achieve the corresponding methyl ether **115**.[227] Differently from the strategy previously used, compound **115** could be subjected to standard hydrogenolysis conditions, in the presence of palladium on charcoal, to cleave the three benzyl ethers and release the triol **116** in 85% yield.



Scheme 44 Synthesis of benzyl glucuronate **110**

The conversion of the primary hydroxyl group into the carboxylic acid was accomplished with TEMPO as the primary oxidant. A first attempt involved (diacetoxyiodo)benzene as co-oxidant, under the conditions aforementioned. Surprisingly, no reaction was observed and the starting material was recovered completely after 24 hours. A possible explanation could be the scarce solubility of the triol **116** in organic solvents, such as dichloromethane, which was necessary as a co-solvent, to dissolve the secondary oxidant, which is insoluble in water. Probably the substrate did not have proper access to the oxidant, despite the vigorous stirring applied. Subsequently, the oxidation was tried under the conditions described by Anelli and coworkers,[60] and methyl 4-*O*-methyl-α-D-glucuronic acid **117** was isolated after treatment of **116** with TEMPO, NaClO and NaBr in water, keeping the pH around 10 – 11 by adding a few drops of a NaOH 1.0 M solution.[228] The resulting carboxylic acid **117** was subjected to esterification in the presence of BnBr and TBAF to give **110** according to an analogous literature protocol (Scheme 44).[229]

The strategy regarding the synthesis of phenylpropyl (*i.e.* **111**) and phenyl (*i.e.* **112**) esters of (methyl α-D-glucopyranoside) uronic acid was envisaged to start from the commercially available methyl α-D-glucopyranoside (**118**).

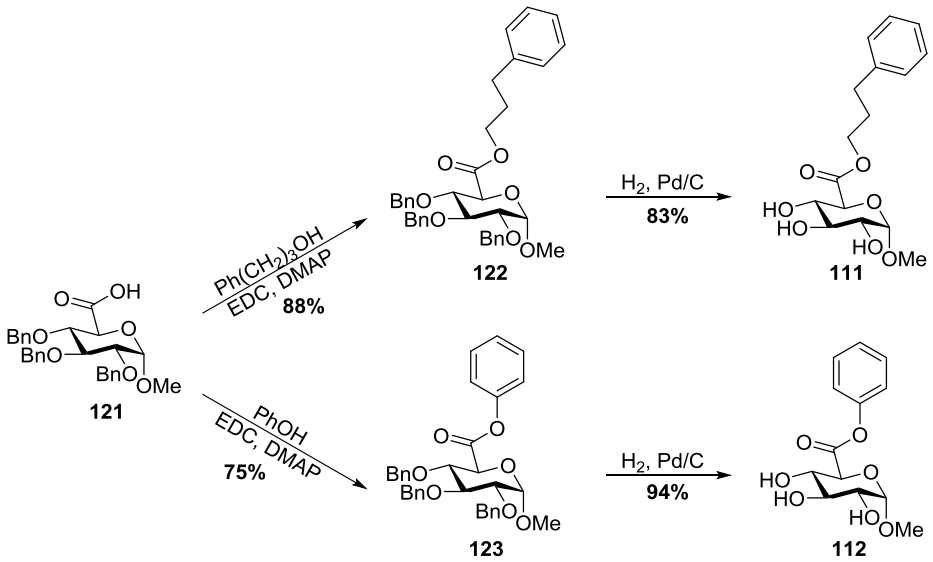


Scheme 45 Synthesis of the glucuronic acid **121**, a key precursor to the target compounds **111** and **112**

The complications experienced in the oxidation of **116** led to the decision that a protection of the secondary hydroxyl groups in tetraol **118** was needed to improve the solubility of the compound under the oxidation conditions. Accordingly, the procedure previously applied to release selectively the primary position for the

synthesis of **93**, was adapted to the current substrate (Scheme 45). The polyol **118** was temporarily and selectively protected at the primary alcohol with a trityl group, benzylated at the remaining positions and de-tritylated directly to yield **120** in a good overall yield by using slightly modified literature protocols.[230] In fact, in this case the subsequent oxidation to the glucuronic acid **121** proceeded smoothly in the presence of TEMPO and (diacetoxyiodo)benzene (Scheme 45).

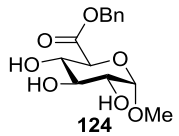
The protected esters 122 and 123 were afforded by condensation of 121 with the corresponding alcohols (3-phenylpropan-1-ol and phenol) under the influence of the coupling agent N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide (EDC) and a catalytic amount of DMAP. Lastly, the deprotection via standard hydrogenolysis gave the desired products 111 and 112, respectively (Scheme 46).



Scheme 46 Oxidation and esterification steps for the synthesis of the glucuronate targets **111** and **112**

2.3 Enzymatic Assays

Compounds **110** – **112**, resembling naturally occurring LCCs, have been employed as screening substrates for the characterization and selection of GEs for industrial delignification of biomass during a seven-month internship in the Enzyme Assay Development Department at Novozymes.



124

Figure 20 Benzyl (methyl α-D-glucopyranoside) uronate **124**

In order to gain more information about the affinity for the 4-OMe substituent on the sugar moiety, compound **124** (Figure 20) was utilized for its analogy in structure with benzyl glucuronate **110**. Glucuronate **124** was previously synthesized by Jonas O. Jørgensen as a part of his bachelor thesis.[47]

2.3.1 Characterization of a novel GE

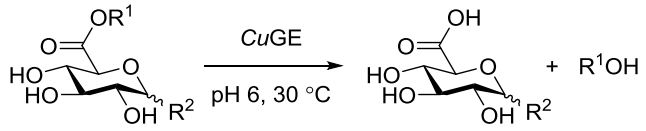
As previously mentioned several GEs have been described and characterized in the last 10 years. A novel GE from *Cerrena unicolor* (*Cu*GE), produced in Novozymes laboratories, was characterized in this study.

The research groups working on GEs, reported that the optimal conditions for obtaining catalytic efficiencies are generally in the pH range 5–7 and 40-60 °C.[46,48,224] Therefore the conditions chosen for these experiments were at pH 6.0 (phosphate buffer) and 30 °C, where two GEs were found to have a k_{cat} of 15-17 s^{-1} on **111**. Such conditions were selected to avoid spontaneous autohydrolysis of the esters that was observed at extreme pH values and high temperatures. Eventually this observation led to seeking an alternative to the regular techniques for the

inactivation of the enzymes, thermic and chemical inactivation, since those methods were affecting the reaction outcome by hydrolysis of the residual starting material. For example, the thermic inactivation of the reaction mixture (5 minutes at 95 °C) raised the autohydrolysis of compound **111** from 0.01 to 0.4 %. Hence it was utilized a 96-wells 10kDa cut-off ultrafiltration plate so that the enzyme was removed mechanically by ultracentrifugation after the reactions were stopped by rapid cooling to 4 °C. A test on the concentration of the eluate was performed to assure that no alterations in concentration were to happen by the process.

A preliminary study on the substrate specifity of *Cu*GE was performed by reacting simple, commercially available esters of glucuronic acid existing as α/β anomeric mixtures with the enzyme at pH 6.0 at 30 °C. The GE activities were measured semi-quantitatively by TLC analysis and are described in Table 5.

Table 5 Activity [a)] of *Cu*GE measured semi-quantitatively by TLC



Entry	R^1	R^2	2 h	18 h	42 h
1	Me	OH	traces	+	+
2	Et	OH	traces	+	++
3	All	OH	+	++	++
4	Bn	OH	++	+++	+++
5	PhPropyl[b)]	αOMe	+++	+++	+++

a) +++: High activity (70-100%); ++: Medium activity (40-70%); +: Low activity (10-40%); trace: ≤ 10% conversion; b) Compound **111**

Overall, a clear preference for the bulkier substrates containing an aryl or alkenyl group in the ester part could be observed, confirming the reported tendency about GEs activity on bulky LCCs mimics.[45]

Consequently, *Cu*GE was subjected to further characterization via Michaelis-Menten kinetics on the synthesized substrates. The GE from *Schizophyllum commune* (*Sc*GE), isolated for the first time from Biely and coworkers in 2006,[41] underwent the same analyses for comparison.

The hydrolysis of the substrates releases aromatic alcohols (phenol, benzyl alcohol or 3-phenylpropanol) allowing to monitor the substrates consumption by UV detection and determining it quantitatively by HPLC.

Table 6 Kinetic parameters for *Sc*GE and *Cu*GE at pH 6.0, 30 °C using synthesized substrates

	*Sc*GE				*Cu*GE			
	K_m [mM]	V_{max} [µmol*min^{-1}*mg^{-1}]	k_{cat} [s^{-1}]	k_{cat}/K_m [mM^{-1}*s^{-1}]	K_m [mM]	V_{max} [µmol*min^{-1}*mg^{-1}]	k_{cat} [s^{-1}]	k_{cat}/K_m [mM^{-1}*s^{-1}]
110	3.7	178	118	32	4.6	161	129	28
124	51	95	64	1.2	80	60	48	0.6
111	66	23	15	0.2	55	21	17	0.3
112[a)]	11	n.d.	n.d.	n.d.	8.9	n.d.	n.d.	n.d.

a) Due to significant autohydrolysis, full kinetic parameters could not be obtained

Table 6 collects the kinetic data obtained for both *Cu*GE and *Sc*GE by fitting Michaelis-Menten kinetics using non-linear regression analysis of V as a function of [S], as showed in Figure 21 for compound **110**.

In general, *Sc*GE and *Cu*GE showed a similar behavior, although the first had a slightly higher catalytic efficiency than the latter. Not surprisingly, the catalytic efficiency (k_{cat}/K_m) for **110**, the substrate bearing the 4-OMe substituent on the glucuronic acid, was 25-50 times higher compared to the values calculated for its analogous **124**. On the other hand a higher K_m value was calculated for **124** compared to **110**, meaning a lower binding affinity of the enzyme to the substrate. Those results confirmed a trend already reported in the literature about the preference of GEs for 4-OMe glucuronate esters.[45] This could be explained by the presence of additional van der Waals interactions between the enzyme and the methoxy group which would result in stronger binding within the active site.[226]

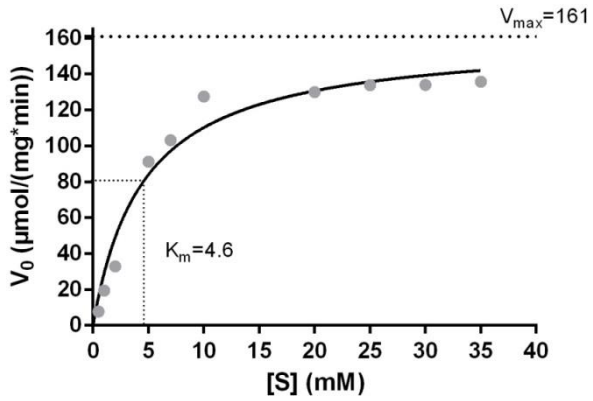


Figure 21 Degradation of **110** with *Cu*GE at pH 6.0 and 30 °C

Binding affinities of both the esterases towards benzyl ester **124** and phenylpropyl ester **111** (*i.e.* mimics of α- and ɣ-esters, respectively) were comparable even though slightly higher catalytic efficiencies were observed for the benzyl ester **124** with both enzymes. The observed preference for bulky esters, as aforementioned, is explained by the fact that the active site of GEs is located on the surface of the enzyme,[225] having easy access to a large substrate on the surface of the enzyme, as

demonstrated from the recent applications on natural[50] or modified polymeric substrates.[49]

Phenyl ester **112** gave a significant autohydrolysis even at pH 6.0, which compromised the calculation of the full kinetic parameters set. Interestingly very high conversions and low binding affinities (K_m) were recorded despite the lack of a 4-*O*-methyl substituent. This result could imply that the existence of phenyl ester LCCs, which has not been reported in the literature so far but could not be ruled out, would not contribute to the recalcitrance of lignocellulose due to the rapidity of the autohydrolysis observed.

In conclusion, it can be asserted that the methoxy group on the 4-position has a key role in the GEs specificity and it is therefore essential for the enzymes to work at their optimal catalytic efficiency. With regard to the alcohol part of the glucuronate esters, considering the observations on a preference for bulky arylalkyl or arylalkenyl[46] groups, an order of GE reactivity on glucuronates was proposed: *benzyl* > *cinnamyl* > *phenylpropyl* > *alkenyl* > *alkyl*.

2.3.2 *Cu*GE and *Sc*GE characterization with a realistic glucuronoyl ester

As a continuation of the studies described in the previous paragraph, the more advanced LCC model compound consisting of a 4-*O*-methyl-glucuronic acid γ-linked to a lignin dimer **125** (Figure 22)[17] was utilized for the kinetic characterization of GEs.

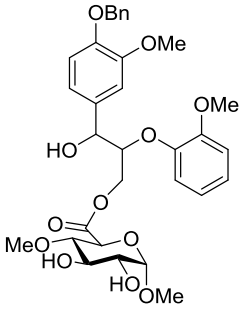


Figure 22 *Threo*-3-[4(benzyloxy)-3-methoxyphenyl]-3-hydroxy-2-(2-methoxyphenoxy)propyl (methyl 4-*O*-methyl-α-D-glucopyranosid)uronate (**125**) synthesized by Li and Helm[17]

The compound was obtained on a generous sample from Prof. Richard F. Helm at Virginia Tech who synthesized the ester in 1995.[17] NMR characterization showed **125** as a mixture of two diastereoisomers since it was prepared from a racemic lignin moiety.

Kinetic characterization of *Cu*GE and *Sc*GE by means of Michaelis-Menten kinetics (Figure 23) was performed by reaction with the ester **125** using the previously reported assay. Although the ester required the addition of 15 V/V% acetonitrile as a co-solvent for the incubation due to limited solubility of both the substrate and the product.

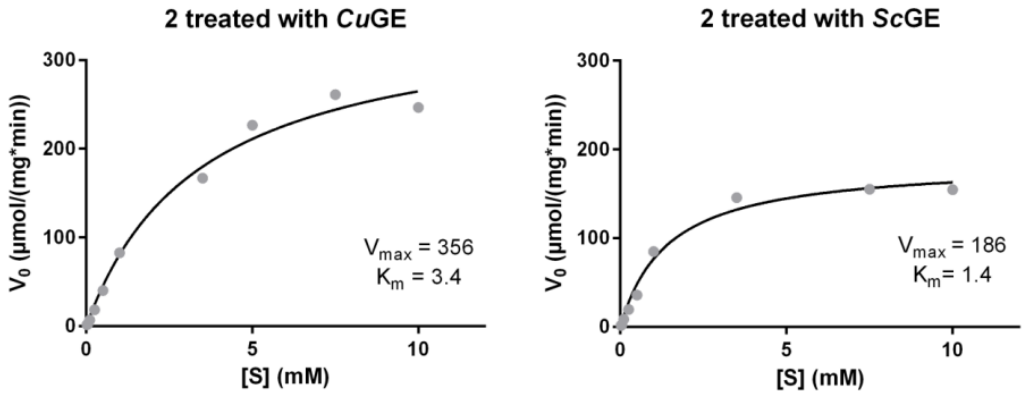


Figure 23 Kinetic curves of **125** with *Cu*GE and *Sc*GE at pH 6.0 and 30 °C

The calculated kinetic data are reported in Table 7 and compared with the most interesting values among the synthesized substrates, **110**.

First and foremost it is worth noticing how binding affinities (K_m) and catalytic efficiencies (k_{cat}/K_m) for both *Sc*GE and *Cu*GE were within the same order of magnitude. However *Cu*GE was found to have a slightly lower binding affinity than *Sc*GE for ester **125**, in accordance with the observations previously made on the relative reactivity of the two esterases.

Table 7 Kinetic parameters for *Cu*GE and *Sc*GE at pH 6.0, 30 °C using **125** and **110**

	*Cu*GE			*Sc*GE		
	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [$mM^{-1}*s^{-1}$]	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [$mM^{-1}*s^{-1}$]
110	4.6	129	28	3.7	118	32
125	3.4	285	83	1.4	125	89

More specifically, higher binding affinities (lower K_m values) are showed for the bulkier glucuronoyl ester **125**, compared to the benzyl ester, for both the enzymes. These results are in accordance for the trend previously reported, confirming the preference of the GEs for bulky arylalkyl alcohols.

3 CONCLUSIONS

During the past three years two different projects have been investigated with the purpose of envisioning the role of glucuronic acid derivatives as substrates for enzymes involved in the biomass degradation. The topic has been explored from two different perspectives: the synthesis of a glucuronoxylan fragment as target for α-glucuronosidases and β-xylanases and the synthesis of aromatic esters of glucuronic esters as targets for glucuronoyl esterases.

First, the synthesis for the (1→4)-β-pentasaccharide **62** was developed. The chosen synthetic strategy was linear and iterative by the use of bifunctional thio-xylosides as building blocks. A protecting-group manipulation strategy was developed for the regioselective protection of the 2-position of the xylose residue on the fourth residue of the pentasaccharide **62** with a Lev group. The glycosylating procedure involved the use of the shelf-stable promoter *p*-NO_2PhSCl and AgOTf. The length of the glycosyl donor, varying from a monosaccharide to a tetrasaccharide, did not affect the outcome and the yields of the different glycosylations. Those results showed that the method was effective and consistent for the type of substrates chosen and led to the desired pentasaccharide with a good 27% overall yield.

In the second part of the project three aromatic esters of glucuronic acid (*i.e.* **110** – **112**) were synthesized from the corresponding methyl glucosides by means of TEMPO oxidation and esterification protocols. They mimic the ester linkage between lignin and hemicellulose fragments in the so-called LCCs. The esters were employed as model substrates for glucuronoyl esterases produced by Novozymes. A novel enzyme of the GE family, *Cu*GE, together with the well-known *Sc*GE, was characterized by kinetic experiments conducted in the Novozymes facilities. The

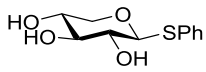
enzymes were treated with the model substrates and the rate of hydrolysis of the esters was measured by HPLC via UV detection of the released aromatic alcohol. The kinetic parameters obtained by means of the Michaelis-Menten equation showed that *Cu*GE has a preference for bulky arylalkyl esters of 4-OMe glucuronic acid, confirming the trends described in literature for *Sc*GE.[45] In order to further support those results a more advanced ester LCC model compound **125** was used as substrate in kinetic experiments and the results compared to those obtained for **110**. The comparison showed values within the same order of magnitude even though a slightly higher binding affinity was observed for **125** with both *Sc*GE and *Cu*GE.

In conclusion, the observed results suggest that GEs could be effective on natural LCCs encouraging further experiments, and therefore their potential utilization in lignocellulosic biomass delignification for forestry, feed and biofuel industries.

4 EXPERIMENTAL

General methods

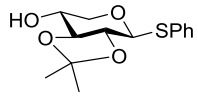
All material, reagents and solvents were purchased from Alfa Aesar, Carbosynth, Sigma-Aldrich or TCI chemicals and used without further purification unless specified otherwise. All solvents were HPLC-grade. The dry solvents were obtained from an Innovative Technology PS-MD-7 Pure-solv solvent purification system. Reactions requiring anhydrous conditions were carried out in flame-dried glassware under inert atmosphere, either using argon or nitrogen. Solvents were removed under vacuum at 30 °C. All reactions were monitored by thin-layer chromatography (TLC), performed on Merck aluminum plates precoated with 0.25 mm silica gel 60 F254. Compounds were visualized under UV irradiation and/or heating after applying a solution of $Ce(SO_4)_2$ (2.5 g) and $(NH_4)_6Mo_7O_{24}$ (6.25 g) in 10% aqueous H_2SO_4 (250 mL). Column chromatography was performed using Geduran silica gel 60 with specified solvents given as volume ratio. 1D (^1H and ^{13}C) and 2D (gCOSY, HSQC, HMBC) NMR spectra were recorded on a Bruker Ascend 400 or a Varian Mercury 300 spectrometer. 2D NMR experiments were performed in order to elucidate the carbohydrate structures. Optical rotations were measured with a Perkin-Elmer Model 241 Polarimeter with a path length of 1 dm. High-resolution mass spectrometry (HRMS) data were recorded on a Bruker SolariX XR 7T ESI/MALDI-FT-ICR MS, with external calibration performed using NaTFA cluster ions. The elemental analyses were performed at the Microanalytic Laboratory Kolbe in Mülheim an der Ruhr (Germany).



Phenyl 1-thio-β-D-xylopyranoside (68)

D-xylose (50.0 g, 0.333 mol) was suspended in dichloromethane (250 ml) together with Et_3N (231 ml, 1.67 mol) and DMAP (8.1 g, 0.067 mol), then acetic anhydride (126 ml, 1.33 mol) was added at 0 °C. The reaction was stirred until TLC indicated full conversion. The reaction mixture was washed with ice-water, 300 ml of 1 M HCl and brine (200 ml). The organic layers were dried over Na_2SO_4, filtered and evaporated under reduced pressure. The crude, without further purification, was dissolved in dichloromethane (300 ml). The stirring mixture was cooled to 0 °C and thiophenol (41 ml, 0.400 mol) and $BF_3 \cdot OEt_2$ (122 ml, 0.999 mol) were added, under inert atmosphere. The solution was stirred at room temperature until disappearance of the starting material on TLC, then diluted with dichloromethane and washed successively with saturated sodium hydrogen carbonate (2x250 ml) and water (2x150 ml), dried over Na_2SO_4, filtrated and concentrated *in vacuo*. The residue was dissolved in methanol (200 ml) and a 0.1 M solution of sodium methoxide in methanol was added. After 15 min the mixture was neutralized with Amberlite IR-120(H+) resin, filtered, and concentrated under reduced pressure. The crude was purified by flash chromatography (ethyl acetate/heptane 7:3, R_f 0.30) to yield **68** (29.5 g, 37%) as white amorphous solid.

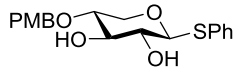
^{1}H NMR (400 MHz, MeOD) δ 7.57 – 7.48 (m, 2H, Ar*H*), 7.37 – 7.23 (m, 3H, Ar*H*), 4.57 (d, *J* = 9.3 Hz, 1H, H-1), 3.95 (dd, *J* = 11.3, 5.2 Hz, 1H, H-5), 3.49 (ddd, *J* = 10.0, 8.8, 5.2 Hz, 1H, H-4), 3.36 (t, *J* = 8.6 Hz, 1H, H-3), 3.24 (dd, *J* = 11.3, 10.1 Hz, 1H, H-5), 3.22 (dd, *J* = 9.3, 8.5 Hz, 1H, H-2). **^{13}C NMR** (101 MHz, MeOD) δ 133.5, 131.7, 128.5, 127.1, 88.7 (C-1), 77.8 (C-3), 72.3 (C-2), 69.5 (C-4), 69.0 (C-5). The data are in accordance with literature.[175]



Phenyl 2,3-*O*-isopropylidine-1-thio-β-D-xylopyranoside (71)

Phenyl thioxyloside **68** (29.5 g, 0.122 mol) was solubilized in DMF (200 ml) with CSA (2.83 g, 0.012 mol) and 2-methoxypropene (37.3 ml, 0.366 mol). The reaction was stirred at 60 °C for 1 hour then cooled to room temperature and quenched with Et_3N (30 ml). The solvent was evaporated and the residue purified over silica gel (heptane/ethyl acetate 7:3, R_f 0.10). 24.1 g (70%) of phenyl 2,3-*O*-isopropylidene thioxyloside **71** was isolated as a colorless oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56 – 7.48 (m, 2H, Ar*H*), 7.36 – 7.27 (m, 3H, Ar*H*), 4.51 (d, *J* = 9.4 Hz, 1H, H-1), 4.11 (dd, *J* = 11.2, 5.2 Hz, 1H, H-5), 3.71 (ddd, *J* = 10.3, 8.8, 5.2 Hz, 1H, H-4), 3.55 (t, *J* = 8.7 Hz, 1H, H-3), 3.34 (dd, *J* = 9.4, 8.6 Hz, 1H, H-2), 3.32 (dd, *J* = 11.2, 10.3 Hz, 1H, H-5), 2.17 (s, 6H, 2xCH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 207.0 ($C(CH_3)_2$), 132.8, 131.6, 129.1, 129.1, 128.3, 127.5, 89.1 (C-1), 77.9 (C-3), 71.9 (C-2), 69.4 (C-4), 69.2 (C-5), 30.9 (2xCH_3). The data are in accordance with literature.[231]

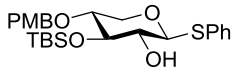


Phenyl 4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (72)

A solution of **71** (17.5 g, 62 mmol), PMBCl (10.9 ml, 80.6 mmol) and NaH (60% oil dispersion, 3.0 g, 74.4 mmol) in DMF (120 ml) was stirred for 16 h at room temperature, then quenched with 10% HCl solution (28 ml). The reaction mixture was diluted with dichloromethane (100 ml) and washed with $NaHCO_3$ (300 ml) and successively brine (200 ml). The organic layers were collected and dried over Na_2SO_4, filtered and concentrated *in vacuo*. The crude compound was dissolved in

CH_2Cl_2/CH_3OH (1:1, 200 ml) and stirred with CSA (14.4 g, 62 mmol) at room temperature overnight. When complete conversion was observed, the reaction was quenched by Et_3N and concentrated. Silica gel purification (heptane/ethyl acetate 6:4, R_f 0.17) afforded **72** (11.1 g, 53%).

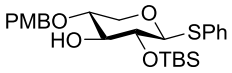
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.59 – 7.46 (m, 2H, Ar*H*), 7.35 – 7.21 (m, 5H, Ar*H*), 6.93 – 6.83 (m, 2H, Ar*H*), 4.61 (d, *J* = 11.2 Hz, 1H, OC*H*$_2$Ph), 4.57 (d, *J* = 11.2 Hz, 1H, OC*H*$_2$Ph), 4.55 (d, *J* = 8.9 Hz, 1H, H-1), 4.06 (dd, *J* = 11.5, 4.8 Hz, 1H, H-5), 3.80 (s, 3H, OC*H*$_3$), 3.66 (t, *J* = 8.6 Hz, 1H, H-3), 3.46 (ddd, *J* = 9.6, 8.6, 4.8 Hz, 1H, H-4), 3.40 (t, *J* = 8.6 Hz, 1H, H-2), 3.27 (dd, *J* = 11.5, 9.7 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.7, 132.8, 132.2, 130.0, 129.7, 129.2, 128.3, 127.6, 114.1, 88.8 (C-1), 76.6 (C-4), 76.5 (C-3), 72.8 (O*C*H$_2$Ph), 72.1 (C-2), 67.1 (C-5), 55.4 (O*C*H$_3$). **HRMS** (MALDI) m/z calcd for $C_{19}H_{22}O_5S$ ($M+Na^+$) 385.1080, found 385.1090.



Phenyl 3-*O*-*tert*-butyldimethylsilyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (81)

Compound **72** (1.0 g, 2.76 mmol) was dissolved in dry THF (30 ml) and the solution was cooled in a dry ice/acetone bath, followed by addition of 2,6-lutidine (0.482 ml, 4.14 mmol) and TBSOTf (0.761 ml, 3.31 mmol). Complete conversion was observed via TLC after 20 min at -78 °C, the reaction mixture was diluted with dichloromethane and washed with water (2x100 ml). The combined organic layers were dried over Na_2SO_4, filtered and concentrated *in vacuo*. The residue was purified over silica gel (heptane/ethyl acetate 9:1, R_f 0.28) to obtain the desired compound **81** as a colorless oil (1.11 g, 84%).

$[\alpha]_D^{20}$ – 130.2° (c 0.086, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.52 – 7.43 (m, 2H, Ar*H*), 7.32 – 7.18 (m, 5H, Ar*H*), 7.04 – 6.70 (m, 2H, Ar*H*), 5.20 (d, *J* = 3.0 Hz, 1H, H-1), 4.62 (d, *J* = 11.5 Hz, 1H, OCH_2Ph), 4.53 (d, *J* = 11.5 Hz, 1H, OCH_2Ph), 4.40 (dd, *J* = 12.4, 1.8 Hz, 1H, H-5), 3.87 (t, *J* = 4.4 Hz, 1H, H-3), 3.81 (s, 3H, OCH_3), 3.75 (d, *J* = 9.4 Hz, 1H, OH), 3.70 – 3.65 (m, 1H, H-2), 3.62 (dd, *J* = 12.4, 3.7 Hz, 1H, H-5), 3.34 (m, 1H, H-4), 0.93 (s, 9H, C(CH_3)$_3$), 0.12 (s, 3H, SiCH_3), 0.06 (s, 3H, SiCH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.6, 131.0, 129.6, 129.5, 129.0, 127.0, 114.1, 89.0 (C-1), 76.1 (C-4), 72.7 (C-2), 71.6 (O*C*H$_2$Ph), 69.4 (C-3), 60.0 (C-5), 55.4 (O*C*H$_3$), 25.9 (C(*C*H$_3$)$_3$), 18.3 (*C*(CH$_3$)$_3$), -4.7 (Si*C*H$_3$), -4.9 (Si*C*H$_3$). **HRMS** (MALDI) m/z calcd for $C_{25}H_{36}O_5SSi$ (M+Na$^+$) 499.1944, found 499.1956.

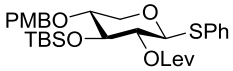


Phenyl 2-*O*-*tert*-butyldimethylsilyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (82)

After dissolving the starting material **72** (200 mg, 0.552 mmol) in dichloromethane (7 ml), the solution was cooled at -78 °C and 2,6-lutidine (0.140 ml, 0.607 mmol) and TBSOTf (0.096 ml, 0.828 mmol) were added. Disappearance of the starting material was observed after 45 minutes and the reaction was washed with water (2x10 ml). The organic phase was dried over Na_2SO_4, filtered and the solvent evaporated. The residue was purified over silica gel (heptane/ethyl acetate 9:1, R_f 0.23) to yield **82** (105 mg, 40%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.51 – 7.44 (m, 2H, Ar*H*), 7.32 – 7.21 (m, 5H, Ar*H*), 6.91 – 6.85 (m, 2H, Ar*H*), 4.60 (d, *J* = 11.3 Hz, 1H, OCH_2Ph), 4.55 (d, *J* = 11.3 Hz, 1H, OCH_2Ph), 4.53 (d, *J* = 8.9 Hz, 1H, H-1), 4.00 (dd, *J* = 11.3, 4.8 Hz, 1H, H-5), 3.80 (s, 3H, OCH_3), 3.54 (dt, *J* = 8.1, 2.2 Hz, 1H, H-3), 3.52 – 3.46 (m,

1H, H-4), 3.49 – 3.42 (m, 1H, H-2), 3.17 (dd, *J* = 11.2, 10.0 Hz, 1H, H-5), 2.52 (d, *J* = 2.2 Hz, 1H, O*H*), 0.94 (s, 9H, C(C*H*$_3$)$_3$), 0.20 (s, 3H, SiC*H*$_3$), 0.16 (s, 3H, SiC*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.6, 134.6, 131.5, 130.2, 129.7, 129.0, 127.4, 114.1, 90.1 (C-1), 78.6 (C-3), 77.2 (C-4), 74.2 (C-2), 72.9 (O*C*H$_2$Ph), 67.4 (C-5), 55.4 (O*C*H$_3$), 26.2 (C(*C*H$_3$)$_3$), 18.6 (*C*(CH$_3$)$_3$), -3.8 (Si*C*H$_3$), -4.1 (Si*C*H$_3$).

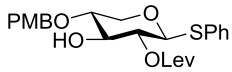


Phenyl 3-*O*-*tert*-butyldimethylsilyl-2-*O*-levulinyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (84)

Compound **81** (1.05 g, 2.21 mmol) was dissolved in dichloromethane (45 ml) followed by addition of DCC (1.38 g, 6.63 mmol), DMAP (0.405 g, 3.32 mmol) and LevOH (0.340 ml, 3.32 mmol). The formation of a white precipitate was observed and full conversion was obtained after 2 h at room temperature. The reaction mixture was filtered through a Celite pad and the filtrate was concentrated to dryness. The residue was purified by column chromatography (heptane/ethyl acetate, 9:1, R_f 0.27) to give **84** as a colorless oil (1.21 g, 95%).

$[\alpha]_D^{20}$ – 33.2° (c 0.28, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.49 – 7.42 (m, 2H, Ar*H*), 7.31 – 7.18 (m, 5H, Ar*H*), 6.89 – 6.83 (m, 2H, Ar*H*), 4.78 (dd, *J* = 6.8, 10.2 Hz, 1H, H-2), 4.85 (d, *J* = 6.8 Hz, 1H, H-1), 4.58 (d, *J* = 11.6 Hz, 1H, OC*H*$_2$Ph), 4.47 (d, *J* = 11.6 Hz, 1H, OC*H*$_2$Ph), 4.13 (dd, *J* = 14.9, 7.0 Hz, 1H, H-5), 3.80 (s, 3H, OC*H*$_3$), 3.79 – 3.76 (m, 1H, H-3), 3.40 – 3.32 (m, 1H, H-4), 3.27 (dd, *J* = 14.9, 5.8 Hz, 1H, H-5), 2.79 – 2.73 (m, 2H, CH_2(β)), 2.67 – 2.57 (m, 2H, CH_2(γ)), 2.17 (s, 3H, C*H*$_3$), 0.89 (s, 9H, C(C*H*$_3$)$_3$), 0.09 (s, 3H, SiC*H*$_3$), 0.06 (s, 3H, SiC*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.4 (Lev, C(O) δ), 171.8 (Lev, C(O) α), 159.5, 135.1, 131.3, 130.2, 129.6, 129.1, 129.0, 127.4, 114.0, 86.9 (C-1), 76.7 (C-4), 72.9 (C-2), 72.4 (O*C*H$_2$Ph), 71.9 (C-3), 64.5 (C-5), 55.4 (O*C*H$_3$), 38.4 (C-β), 29.9 (C-ε),

28.3 (C-γ), 25.8 (C(*C*H_3)$_3$), 18.1 (*C*(CH$_3$)$_3$), -4.4 (Si*C*H_3), -4.7 (Si*C*H_3). **HRMS** (MALDI) m/z calcd for $C_{30}H_{42}O_7SSi$ (M+Na^+) 597.2312, found 597.2325.



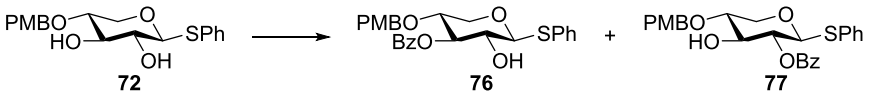
Phenyl 2-*O*-levulinyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (78)

The silylated xyloside **81** (0.100 g, 0.174 mmol) was dissolved in CH_3CN (5 ml) and the mixture was cooled to 0 °C. A 20% HF solution (0.3 ml, 3.48 mmol) was added and the reaction stirred overnight at room temperature. TLC showed conversion of the starting material, so the remaining HF was quenched with methoxytrimethylsilane (0.960 ml, 6.96 mmol) and stirred for 1 hour. The mixture was diluted with dichloromethane (10 ml) and washed with saturated $NaHCO_3$ solution (20 ml) and water (20 ml). The organic layers were combined, dried over Na_2SO_4, filtered and the solvent removed under reduced pressure. The crude material was purified by column chromatography (heptane/ethyl acetate 7:3) yielding the alcohol **78** as a white amorphous solid (0.073 g, 91%).

$[\alpha]_D^{20}$ – 18.8° (c 0.47, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.57 – 7.44 (m, 2H, Ar*H*), 7.36 – 7.22 (m, 5H, Ar*H*), 6.94 – 6.83 (m, 2H, Ar*H*), 4.86 (dd, *J* = 9.6, 9.0 Hz, 1H, H-2), 4.71 (d, *J* = 11.5 Hz, 1H, OC*H*$_2$Ph), 4.63 (d, *J* = 9.8 Hz, 1H, H-1), 4.59 (d, *J* = 11.5 Hz, 1H, OC*H*$_2$Ph), 4.03 (dd, *J* = 11.5, 5.2 Hz, 1H, H-5), 3.81 (s, 3H, OC*H*$_3$), 3.77 (td, *J* = 8.9, 1.6 Hz, 1H, H-3), 3.55 (ddd, *J* = 5.2, 9.0, 10.0 Hz, 1H, H-4), 3.24 (dd, *J* = 11.3, 10.4 Hz, 1H, H-5), 3.17 (d, *J* = 2.1 Hz, 1H, OH), 2.91 (ddd, *J* = 18.4, 7.6, 6.4 Hz, 1H, CH$_2$(β)), 2.81 (dt, *J* = 18.5, 5.9 Hz, 1H, CH$_2$(β)), 2.68 (ddd, *J* = 12.9, 5.9, 3.5 Hz, 1H, CH$_2$(γ)), 2.61 (dt, *J* = 12.7, 4.4 Hz, 1H, CH$_2$(γ)), 2.21 (s, 3H, C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 207.5 (Lev, C(O) δ), 172.1 (Lev, C(O) α), 159.5, 132.9, 132.4, 130.2, 129.6, 129.0, 128.0, 114.0, 86.4 (C-1), 76.5 (C-4), 76.2 (C-3), 73.1 (O*C*H$_2$Ph), 72.7 (C-2), 67.8 (C-5), 55.4 (O*C*H$_3$),

38.4 (C-β), 29.9 (C-ε), 28.3 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{24}H_{28}O_7S$ ($M+Na^+$) 483.1447, found 483.1458.

Selective Benzoylation



Diol **72** (0.200 g, 0.552 mmol) was dissolved in dichloromethane (7 ml) and 1-BBTZ, freshly synthesized from 1-hydroxybenzotriazole and BzCl according to a literature protocol,[189] was added (0.172 g, 0.718 mmol). The reaction mixture was cooled to 0 °C and Et_3N (0.115 ml, 0.828 mmol) was added. The reaction mixture was stirred at 22 °C for 18 hours, diluted with dichloromethane and washed two times with brine. The organic phases were combined, dried with Na_2SO_4, filtered and the solvent evaporated *in vacuo*. The purification via flash chromatography (heptane/ethyl acetate 7:3) gave compound **76** (0.056 g, 22%, R_f 0.23) and compound **77** (0.156 g, 61%, R_f 0.16).

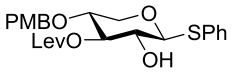
Diol **72** (0.100 g, 0.276 mmol) was dissolved in CH_3CN (2 ml). $MoO_2(acac)_2$ (2 mg, 0.0055 mmol) and 2,4,6-collidine (0.069 ml, 0.552 mmol) were added to the solution at 22 °C, then BzCl (0.045 ml, 0.414 mmol) was added dropwise. The reaction was stirred for 24 hours, and stopped although TLC revealed remaining starting material. The mixture was diluted with dichloromethane, washed with 1 M HCl, $NaHCO_3$ and water. The organic layers were combined and dried over Na_2SO_4, filtered and the solvent was removed under reduced pressure. The crude was purified by column chromatography (heptane/ethyl acetate 7:3) to yield **76** (0.010 g, 8%) and **77** (0.027 g, 21%).

Phenyl 2-*O*-benzoyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (77)

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.11 – 8.05 (m, 2H, Ar*H*), 7.62 – 7.54 (m, 1H, Ar*H*), 7.49 – 7.40 (m, 4H, Ar*H*), 7.30 – 7.22 (m, 5H, Ar*H*), 6.90 – 6.84 (m, 2H, Ar*H*), 5.06 (dd, *J* = 9.4, 8.8 Hz, 1H, H-2), 4.80 (d, *J* = 9.5 Hz, 1H, H-1), 4.64 (d, *J* = 11.4 Hz, 1H, OC*H*$_2$Ph), 4.58 (d, *J* = 11.5 Hz, 1H, OC*H*$_2$Ph), 4.12 (dd, *J* = 11.5, 5.0 Hz, 1H, H-5), 3.86 (t, *J* = 8.7 Hz, 1H, H-3), 3.79 (s, 3H, OC*H*$_3$), 3.60 (ddd, *J* = 9.9, 8.7, 5.1 Hz, 1H, H-4), 3.32 (dd, *J* = 11.4, 10.0 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.1 (Bz), 159.6, 133.5, 132.7, 130.1, 130.0, 129.8, 129.7, 129.1, 128.6, 128.1, 114.2, 114.2, 86.8 (C-1), 77.1 (C-4), 75.9 (C-3), 73.1 (O*C*H$_2$Ph), 73.0 (C-2), 67.6 (C-5), 55.4 (O*C*H$_3$). **Elemental Analysis**: calc. C: 66.94 H: 5.62 S: 6.87; found: C: 66.98 H: 5.59 S: 6.74

Phenyl 3-*O*-benzoyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (76)

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.14 – 8.07 (m, 2H, Ar*H*), 7.60 – 7.52 (m, 1H, Ar*H*), 7.52 – 7.40 (m, 4H, Ar*H*), 7.31 – 7.13 (m, 5H, Ar*H*), 6.81 – 6.74 (m, 2H, Ar*H*), 5.31 (t, *J* = 5.6 Hz, 1H, H-3), 5.08 (d, *J* = 4.9 Hz, 1H, H-1), 4.58 (d, *J* = 11.7 Hz, 1H, OC*H*$_2$Ph), 4.55 (d, *J* = 11.7 Hz, 1H, OC*H*$_2$Ph), 4.37 (dd, *J* = 11.5, 2.2 Hz, 1H, H-5), 3.80 (t, *J* = 5.2 Hz, 1H, H-2), 3.74 (s, 3H, OC*H*$_3$), 3.67 – 3.56 (m, 2H, H-4, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.9 (Bz), 159.7, 134.2, 133.6, 132.2, 130.2, 129.7, 129.5, 129.3, 129.2, 128.6, 127.9, 127.6, 114.1, 89.2 (C-1), 73.1 (C-4), 72.0 (O*C*H$_2$Ph), 71.6 (C-3), 70.4 (C-2), 62.7 (C-5), 55.4 (O*C*H$_3$).

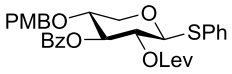


Phenyl 3-*O*-levulinyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (79)

An aqueous solution of NaOH 1.13 M (1.4 ml, 4.137 mmol) was added to a solution of diol **72** (0.200 g, 0.552 mmol) and Bu_4NHSO_4 (0.037 g, 0.110 mmol) in

dichloromethane (10 ml) at –5 °C. After 2 minutes of vigorous stirring, LevCl (0.096 mg, 0.718 mmol) was added and the solution was stirred for further 20 minutes. The reaction mixture was diluted with dichloromethane and washed two times with brine. The organic phase was dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude material was purified by flash chromatography (CH_2Cl_2/ethyl acetate 9:1) to yield **79** (0.133 g, 52%) and **78** (0.041 g, 16%).

79: $[\boldsymbol{\alpha}]_{\mathbf{D}}^{\mathbf{20}}$ – 74.4° (c 0.30, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.64 – 7.42 (m, 2H, Ar*H*), 7.40 – 7.14 (m, 5H, Ar*H*), 6.84 – 6.77 (m, 2H, Ar*H*), 5.09 (t, *J* = 6.4 Hz, 1H, H-3), 4.94 (d, *J* = 6.0 Hz, 1H, H-1), 4.57 (br s, 2H, OC*H*$_2$Ph), 4.22 (dd, *J* = 11.2, 2.5 Hz, 1H, H-5), 3.81 (s, 3H, OC*H*$_3$), 3.65 (q, *J* = 6.3 Hz, 1H, H-2), 3.60 – 3.52 (m, 2H, H-4, H-5), 3.50 (d, *J* = 6.9 Hz, 1H, OH), 2.83 (t, *J* = 6.5 Hz, 2H, CH_2(β)), 2.63 (t, *J* = 6.5 Hz, 2H, CH_2(γ)), 2.20 (s, 3H, C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.9 (Lev, C(O) δ), 172.1 (Lev, C(O) α), 159.6, 133.8, 132.1, 129.6, 129.6, 129.1, 127.8, 114.0, 88.9 (C-1), 73.4 (C-4), 72.9 (C-3), 72.1 (O*C*H$_2$Ph), 70.4 (C-2), 63.8 (C-5), 55.4 (O*C*H$_3$), 38.1 (C-β), 29.9 (C-ε), 28.3 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{25}H_{28}O_7S$ ($M+Na^+$) 483.1447, found 483.1459.

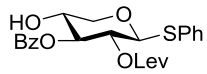


Phenyl 3-*O*-benzoyl-2-*O*-levulinyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (75)

Compound **76** (1.0 g, 2.14 mmol) was dissolved in dichloromethane (30 ml) followed by addition of DCC (0.53 g, 2.57 mmol), DMAP (0.261 g, 2.14 mmol) and LevOH (0.33 ml, 3.22 mmol). A white precipitate formed and complete conversion was observed after 40 minutes at 22 °C. The reaction mixture was filtered through a Celite pad and the filtrate was concentrated to dryness. The

residue was purified by column chromatography (heptane/ethyl acetate, 7:3) to give **75** as a colorless amorphous solid (1.21 g, 98%).

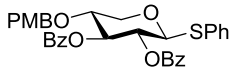
$[\alpha]_D^{20}$ – 13.3° (c 0.27, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.03 – 7.98 (m, 2H, Ar*H*), 7.62 – 7.56 (m, 1H, Ar*H*), 7.52 – 7.42 (m, 4H, Ar*H*), 7.35 – 7.28 (m, 3H, Ar*H*), 7.13 – 7.07 (m, 2H, Ar*H*), 6.74 – 6.68 (m, 2H, Ar*H*), 5.42 (t, *J* = 8.4 Hz, 1H, H-3), 5.04 (t, *J* = 8.6 Hz, 1H, H-2), 4.82 (d, *J* = 8.7 Hz, 1H, H-1), 4.51 (d, *J* = 11.9 Hz, 1H, OC*H*$_2$Ph), 4.47 (d, *J* = 11.9 Hz, 1H, OC*H*$_2$Ph), 4.17 (dd, *J* = 11.8, 4.9 Hz, 1H, H-5), 3.74 (s, 3H, OC*H*$_3$), 3.69 (td, *J* = 9.2, 4.9 Hz, 1H, H-4), 3.44 (dd, *J* = 11.8, 9.2 Hz, 1H, H-5), 2.63 – 2.43 (m, 4H, CH_2(β), CH_2(γ)), 2.04 (s, 3H, C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.8 (Lev, C(O) δ), 171.4 (Lev, C(O) α), 165.6 (Bz), 159.4, 133.3, 132.7, 132.5, 130.0, 129.6, 129.5, 129.0, 128.4, 128.1, 113.8, 86.6 (C-1), 74.2 (C-3), 73.9 (C-4), 72.4 (O*C*H$_2$Ph), 70.3 (C-2), 66.9 (C-5), 55.2 (O*C*H$_3$), 37.8 (C-β), 29.6 (C-ε), 28.0 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{31}H_{32}O_8S$ (M+Na$^+$) 587.1710, found 587.2781.



Phenyl 3-*O*-benzoyl-2-*O*-levulinyl-1-thio-β-D-xylopyranoside (66)

Compound **75** (3.33 g, 5.89 mmol) was dissolved in CH_2Cl_2/H_2O (9:1, 30 ml) and DDQ (2.01 g, 8.84 mmol) was added. The mixture was vigorously stirred for 4 hours at 22 °C, until TLC indicated full conversion and the remaining DDQ was quenched with a buffer solution of ascorbic acid (0.7%, 1.86 g, 8.84 mmol), citric acid (1.5%) and sodium hydroxide (1%). The mixture was diluted with dichloromethane and washed with saturated aqueous $NaHCO_3$ and water. The organic phase was dried (Na_2SO_4), filtered and the solvent removed under vacuum. Purification with column chromotography yielded **66** (2.41 g, 92%) as a white amorphous solid.

$[\alpha]_D^{20}$ + 25.6° (c 0.70, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.23 – 7.89 (m, 2H, Ar*H*), 7.70 – 7.28 (m, 8H, Ar*H*), 5.20 – 5.13 (m, 2H, H-2, H-3), 4.96 – 4.90 (m, 1H, H-1), 4.38 (dd, *J* = 11.9, 4.4 Hz, 1H, H-5), 3.95 – 3.87 (m, 1H, H-4), 3.54 (dd, *J* = 11.9, 7.9 Hz, 1H, H-5), 2.72 – 2.65 (m, 2H, CH_2(β)), 2.65 – 2.50 (m, 2H, CH_2(γ)), 2.09 (s, 3H, C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.9 (Lev, C(O) δ), 171.2 (Lev, C(O) α), 166.9 (Bz), 133.7, 132.9, 132.5, 130.2, 129.1, 128.9, 128.6, 128.1, 86.4 (C-1), 75.7 (C-3), 69.7 (C-2), 68.2 (C-4), 67.3 (C-5), 37.9 (C-β), 29.6 (C-ε), 28.0 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{23}H_{24}O_7S$ ($M+Na^+$) 467.1134, found 467.1146.

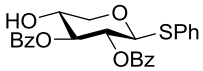


Phenyl 2,3-di-*O*-benzoyl-4-*O*-*p*-methoxybenzyl-1-thio-β-D-xylopyranoside (73)

The diol **72** (0.680 g, 1.88 mmol) was dissolved in pyridine (5 ml) and BzCl (0.436 ml, 3.75 mmol) was added. The reaction mixture was stirred at 22 °C for 2 h, then it was diluted with dichloromethane and washed with 1 M HCl (2x20 ml) and water (2x20 ml). The organic phase was dried over Na_2SO_4, filtered and the solvent removed under vacuum. The residue was purified by column chromatography to afford **73** (0.877, 82%).

$[\alpha]_D^{20}$ + 55.0° (c 1.00, $CHCl_3$). **^{1}H NMR (**400 MHz, $CDCl_3$) δ 7.99 – 7.95 (m, 4H), 7.56 – 7.51 (m, 2H), 7.50 – 7.46 (m, 2H), 7.42 –7.34 (m, 4H), 7.31 – 7.27 (m, 3H), 7.15 – 7.12 (m, 2H), 6.74 – 6.71 (m, 2H), 5.60 (t, *J* = 8.0 Hz, 1H, H-3), 5.34 (t, *J* = 8.0 Hz, 1H, H-2), 5.04 (d, *J* = 8.1 Hz, 1H, H-1), 4.55 (d, *J* = 11.8 Hz, 1H, OC*H*$_2$Ph), 4.52 (d, *J* = 11.8 Hz, 1H, OC*H*$_2$Ph), 4.28 (dd, *J* = 11.9, 4.6 Hz, 1H, H-5), 3.79 – 3.74 (m, 1H, H-4), 3.75 (s, 3H, OC*H*$_3$), 3.57 (dd, *J* = 11.9, 8.5 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.7 (Bz), 165.4 (Bz), 159.5, 133.4, 133.2, 132.5,

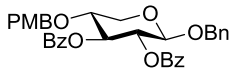
130.1, 130.0, 129.7, 129.7, 129.5, 129.4, 129.1, 128.5, 128.5, 128.1, 113.9, 86.9 (C-1), 73.9 (C-3), 73.6 (C-4), 72.5 (O*C*H_2Ph), 70.6 (C-2), 66.4 (H-5), 55.3 (O*C*H_3). **HRMS** (MALDI) m/z calcd for $C_{33}H_{30}O_7S$ (M+Na^+) 593.1604, found 593.1617.



Phenyl 2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (67)

Compound **73** (0.620 g, 1.09 mmol) was dissolved in CH_2Cl_2/H_2O 9:1 (3 ml) and DDQ (0.370 g, 1.63 mmol) was added. The mixture was vigorously stirred for 4 h at room temperature, the remaining DDQ was quenched using a buffer solution composed of ascorbic acid (0.7%, 0.287 g, 1.63 mmol), citric acid (1.5%) and sodium hydroxide (1%). It was afterwards diluted with dichloromethane, washed with saturated aqueous $NaHCO_3$ and water. The organic phase was dried with Na_2SO_4, filtered, the solvent evaporated under reduced pressure and purified by column chromatography (heptane/ethyl acetate 7:3, R_f 0.18) yielded **67** (0.458 g, 94%) as white amorphous solid.

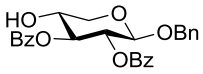
$[\alpha]_D^{20}$ + 63.7° (c 1.00, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.07 – 7.97 (m, 4H, Ar*H*), 7.58 – 7.47 (m, 4H, Ar*H*), 7.45 – 7.38 (m, 4H, Ar*H*), 7.34 – 7.29 (m, 3H, Ar*H*), 5.43 (t, *J* = 7.4 Hz, 1H, H-2), 5.33 (t, *J* = 7.4 Hz, 1H, H-3), 5.09 (d, *J* = 7.3 Hz, 1H, H-1), 4.45 (dd, *J* = 12.0, 4.4 Hz, 1H, H-5), 4.00 (td, *J* = 7.6, 4.4 Hz, 1H, H-4), 3.61 (dd, *J* = 12.0, 7.6 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.1 (Bz), 165.2 (Bz), 133.8, 133.6, 133.0, 132.8, 130.2, 130.0, 129.3, 129.2, 128.9, 128.7, 128.6, 128.3, 86.8 (C-1), 76.0 (C-3), 70.2 (C-2), 68.4 (C-4), 67.6 (C-5). **HRMS** (MALDI) m/z calcd for $C_{33}H_{30}O_7S$ (M+Na^+) 593.1604, found 593.1617.



Benzyl 2,3-di-*O*-benzoyl-4-*O*-*p*-methoxybenzyl-β-D-xylopyranoside (74)

Starting material **73** (200 mg, 0.350 mmol) and the acceptor, benzyl alcohol (0.044 ml, 0.420 mmol), were mixed in the reaction flask and dried overnight on a vacuum line. The reactants were dissolved in dichloromethane (6 ml) and cooled to -40 °C, NIS (94 mg, 0.420 mmol) and triflic acid (9 µl, 0.105 mmol) were added to the stirring mixture. Full conversion of the starting material was observed via TLC analysis after 20 minutes and the reaction was neutralized with Et_3N (0.145 ml, 1.05 mmol). The resulting mixture was stirred with 6 ml of 1 M $Na_2S_2O_3$ until the yellow color disappeared. The organic phase was diluted with dichloromethane and washed with brine, dried over Na_2SO_4, filtered, concentrated and purified with flash chromatography (heptane/ethyl acetate 7:3) yielding **74** (0.155 g, 78%) as a colorless oil.

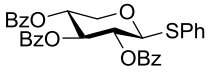
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 – 7.95 (m, 4H, Ar*H*), 7.61 – 7.52 (m, 2H, Ar*H*), 7.46 – 7.36 (m, 5H, Ar*H*), 7.34 – 7.13 (m, 6H, Ar*H*), 6.83 – 6.75 (m, 2H, Ar*H*), 5.60 (t, *J* = 8.1 Hz, 1H, H-3), 5.39 (dd, *J* = 8.3, 6.5 Hz, 1H, H-2), 4.92 (d, *J* = 12.3 Hz, 1H, OC*H*$_2$Ph, OBn), 4.78 (d, *J* = 6.5 Hz, 1H, H-1), 4.68 (d, *J* = 12.3 Hz, 1H, OC*H*$_2$Ph, OBn), 4.61 (d, *J* = 11.8 Hz, 1H, OC*H*$_2$Ph), 4.57 (d, *J* = 11.8 Hz, 1H, OC*H*$_2$Ph), 4.19 (dd, *J* = 11.9, 4.6 Hz, 1H, H-5), 3.88 – 3.81 (m, 1H, H-4), 3.80 (s, 3H, OC*H*$_3$), 3.55 (dd, *J* = 11.9, 8.5 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.7 (Bz), 165.4 (Bz), 159.4, 141.1, 137.1, 133.2, 130.0, 129.9, 129.8, 129.7, 129.6, 129.6, 129.5, 128.6, 128.4, 128.4, 128.4, 127.9, 127.8, 127.7, 127.1, 113.9, 99.4 (C-1), 74.0 (C-4), 72.6 (C-3), 72.4 (*C*H$_2$Ph), 71.1 (C-2), 70.4 (*C*H$_2$Ph, OBn), 63.1 (C-5), 55.3 (O*C*H$_3$).



Benzyl 2,3-di-*O*-benzoyl-β-D-xylopyranoside (64)

Compound **74** (0.155 g, 0.272 mmol) was dissolved in CH_2Cl_2/H_2O 9:1 (3 ml) and DDQ (0.093 g, 0.408 mmol) was added. The reaction mixture was vigorously stirred for 4 h at room temperature and the remaining DDQ was quenched using a buffer solution composed of ascorbic acid (0.7%, 0.072 g, 0.408 mmol), citric acid (1.5%) and sodium hydroxide (1%). The mixture was diluted with dichloromethane and washed with saturated aqueous $NaHCO_3$ and water, the organic phase was dried with Na_2SO_4, filtered and concentrated to afford **64** (0.113 g, 93%) as a white amorphous solid after purification by column chromatography.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.99 – 7.94 (m, 4H, Ar*H*), 7.59 – 7.19 (m, 11H, Ar*H*), 5.43 (dd, *J* = 7.8, 6.0 Hz, 1H, H-2), 5.26 (t, *J* = 7.6 Hz, 1H, H-3), 4.89 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.78 (d, *J* = 6.0 Hz, 1H, H-1), 4.65 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.26 (dd, *J* = 12.0, 4.5 Hz, 1H, H-5), 4.01 (td, *J* = 7.6, 4.5 Hz, 1H, H-4), 3.54 (dd, *J* = 12.0, 7.8 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.2 (Bz), 165.3 (Bz), 137.0, 133.7, 133.5, 130.1, 130.0, 129.4, 129.0, 128.6, 128.5, 128.5, 128.0, 99.1 (C-1), 75.4 (C-3), 70.5 (C-2), 70.4 (O*C*H$_2$Ph), 68.7 (C-4), 64.6 (C-5).



Phenyl 2,3,4-tri-*O*-benzoyl-1-thio-β-D-xylopyranoside (65)

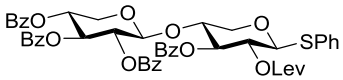
The building block **68** (5.12 g, 21.13 mmol) was dissolved in pyridine (45 ml) and BzCl (7.4 ml, 63.40 mmol). The reaction mixture was stirred at room temperature for 1 h and the excess of BzCl was quenched by adding 10 ml of methanol and the mixture was stirred for additional 10 minutes. The disappearance of a white precipitate was observed. The reaction mixture was diluted with dichloromethane

and washed with 1 M HCl (2x100 ml) and water (2x100 ml). The organic phase was dried over Na_2SO_4, filtered and the solvent evaporated *in vacuo*. The residue was purified by column chromatography to afford **65** (5.1 g, 44%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06 – 8.03 (m, 2H, Ar*H*), 8.01 – 7.98 (m, 4H, Ar*H*), 7.56 – 7.51 (m, 5H, Ar*H*), 7.42 – 7.32 (m, 9H, Ar*H*), 5.78 (t, *J* = 6.6 Hz, 1H, H-3), 5.46 (t, *J* = 6.3 Hz, 1H, H-2), 5.32-5.27 (m, 2H, H-1, H-4), 4.71 (dd, *J* = 12.3, 4.0 Hz, 1H, H-5), 3.83 (dd, *J* = 12.3, 6.5 Hz, 1H, H-5). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.6 (Bz), 165.3 (Bz), 165.3 (Bz), 133.6, 133.5, 133.5, 133.2, 132.8, 130.2, 130.1, 130.1, 129.3, 129.2, 129.0, 128.6, 128.6, 128.5, 128.3, 86.5 (C-1), 70.6 (C-3), 70.1 (C-2), 68.8 (C-4), 63.7 (C-5). The data are in accordance with literature.[232]

General procedure of glycosylation

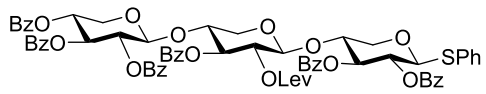
Silver triflate (0.722 mmol), dissolved in toluene (2 ml), was added to a solution of the donor (0.361 mmol) and stirred with freshly activated molecular sieves 4 Å (1 g) in dichloromethane (3 ml). The mixture was stirred for 10 minutes and cooled down to -65 °C. *p*NO_2PhSCl (0.361 mmol) was dissolved in anhydrous dichloromethane (0.5 ml) and added dropwise to the mixture. The acceptor (0.324 mmol) was dissolved in dichloromethane (0.5 ml) and added to the mixture after complete consumption of donor was observed by TLC. The temperature was raised to -55 °C and kept between -55 and -50 °C until TLC showed full conversion of the acceptor. The mixture was warmed up to -15 °C within 10 minutes and the reaction was neutralized with Et_3N (1.08 mmol). The mixture was diluted with dichloromethane and filtered through a Celite pad. The filtrate was concentrated under reduced pressure and purified by flash chromatography.



Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinyl-1-thio-β-D-xylopyranoside (91)

Compound **91** was synthesized from donor **65** and acceptor **66** following the general procedure of glycosylation, and isolated in 89% yield after silica gel purification (toluene/heptane/ethyl acetate 2:4:3).

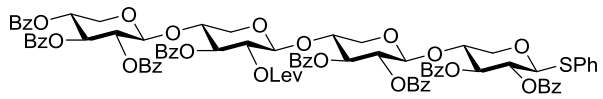
$[\boldsymbol{\alpha}]_{\mathbf{D}}^{\mathbf{20}}$ – 43.6° (c 0.38, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09 – 7.88 (m, 8H, Ar*H*), 7.63 – 7.28 (m, 17H, Ar*H*), 5.64 (t, *J* = 6.6 Hz, 1H, H-3'), 5.50 (t, *J* = 8.2 Hz, 1H, H-3), 5.22 (dd, *J* = 6.6, 4.9 Hz, 1H, H-2'), 5.09 (t, *J* = 8.5 Hz, 1H, H-2), 5.10 – 5.06 (m, 1H, H-4'), 4.91 (d, *J* = 4.8 Hz, 1H, H-1'), 4.83 (d, *J* = 8.6 Hz, 1H, H-1), 4.19 (dd, *J* = 12.0, 4.9 Hz, 1H, H-5), 4.09 – 3.95 (m, 2H, H-4, H-5'), 3.45 (dd, *J* = 12.0, 9.1 Hz, 1H, H-5), 3.43 (dd, *J* = 12.5, 6.1 Hz, 1H, H-5'), 2.72 – 2.41 (m, 4H, CH_2(β), CH_2(γ)), 2.08 (s, 3H, C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.9 (Lev, C(O) δ), 171.3 (Lev, C(O) α), 165.5 (Bz), 165.4 (Bz), 165.3 (Bz), 165.0 (Bz), 133.5, 133.5, 133.5, 132.8, 132.4, 130.0, 130.0, 129.9, 129.9, 129.5, 129.3, 129.1, 129.0, 128.6, 128.6, 128.5, 128.5, 128.2, 99.7 (C-1'), 86.4 (C-1), 75.2 (C-4), 73.4 (C-3), 70.3 (C-2), 70.1 (C-2'), 69.5 (C-3'), 68.5 (C-4'), 66.0 (C-5), 60.8 (C-5'), 37.8 (C-β), 29.7 (C-ε), 28.0 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{49}H_{44}O_{14}S$ ($M+Na^+$) 911.2343, found 911.2362.



Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (92)

Compound **92** was synthesized from donor **91** and acceptor **67** following the general procedure of glycosylation, and isolated in 66% yield after purification (toluene/heptane/ethyl acetate 2:1:1).

$[\alpha]_D^{20}$ – 43.3° (c 0.24, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 – 7.84 (m, 12H, Ar*H*), 7.62 – 7.21 (m, 23H, Ar*H*), 5.60 (m, 2H, H-3, H-3’’), 5.36 (t, *J* = 8.3 Hz, 1H, H-3’), 5.32 (t, *J* = 7.7 Hz, 1H, H-2), 5.13 (dd, *J* = 6.6, 5.0 Hz, 1H, H-2’’), 5.06 (d, *J* = 7.9 Hz, 1H, H-1), 5.07 – 4.99 (m, 1H, H-4’’), 4.91 (dd, *J* = 8.4, 6.8 Hz, 1H, H-2’), 4.68 (d, *J* = 4.8 Hz, 1H, H-1’’), 4.60 (d, *J* = 6.7 Hz, 1H, H-1’), 4.37 (dd, *J* = 12.1, 4.6 Hz, 1H, H-5), 4.04 – 3.94 (m, 2H, H-4, H-5’’), 3.76 (td, *J* = 8.5, 5.0 Hz, 1H, H-4’), 3.64 (dd, *J* = 12.1, 8.3 Hz, 1H, H-5), 3.52 (dd, *J* = 12.2, 4.9 Hz, 1H, H-5’), 3.35 (dd, *J* = 12.4, 6.2 Hz, 1H, H-5’’), 3.10 (dd, *J* = 12.2, 9.0 Hz, 1H, H-5’), 2.68 – 2.34 (m, 4H, CH_2(β), CH_2(γ)), 2.07 (s, 3H, CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.0 (Lev, C(O) δ), 171.3 (Lev, C(O) α), 165.6 (Bz), 165.5 (Bz), 165.4 (Bz), 165.3 (Bz), 165.3 (Bz), 165.0 (Bz), 133.5, 133.5, 133.4, 133.3, 133.0, 132.6, 130.1, 130.0, 130.0, 129.9, 129.8, 129.6, 129.5, 129.5, 129.3, 129.2, 129.1, 129.0, 128.6, 128.5, 128.5, 128.5, 128.4, 128.1, 100.8 (C-1’), 99.5 (C-1’’), 86.7 (C-1), 75.5 (C-4), 75.0 (C-4’), 72.8 (C-3), 72.5 (C-3’), 71.6 (C-2’), 70.4 (C-2), 70.1 (C-2’’), 69.6 (C-3’’), 68.6 (C-4’’), 65.6 (C-5), 62.4 (C-5’), 60.9 (C-5’’), 37.8 (C-β), 29.8 (C-ε), 27.9 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{68}H_{60}O_{20}S$ ($M+Na^+$) 1251.3290, found 1251.3310.

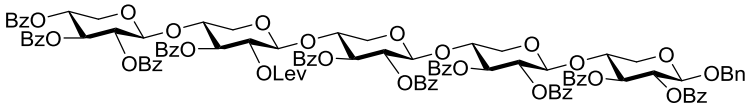


Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (63)

Compound **63** was synthesized from donor **92** and acceptor **67** following the general procedure of glycosylation, and isolated in 79% yield after purification (toluene/heptane/ethyl acetate 3:3:2).

$[\boldsymbol{\alpha}]_{\mathbf{D}}^{\mathbf{20}}$ – 42.1° (c 0.38, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.23 – 7.78 (m, 15H, Ar*H*), 7.72 – 6.98 (m, 30H, Ar*H*), 5.63 (t, *J* = 7.9 Hz, 1H, H-3), 5.57 (t, *J* = 6.7 Hz, 1H, H-3’’’), 5.46 (t, *J* = 7.7 Hz, 1H, H-3’), 5.35 – 5.25 (m, 2H, H-2, H-3’’), 5.15 (dd, *J* = 7.9, 6.0 Hz, 1H, H-2’), 5.11 (dd, *J* = 9.1, 4.2 Hz, 1H, H-2’’’), 5.03 (td, *J* = 6.4, 4.1 Hz, 1H, H-4’’’), 4.99 (d, *J* = 8.0 Hz, 1H, H-1), 4.83 (dd, *J* = 8.7, 6.8 Hz, 1H, H-2’’), 4.78 (d, *J* = 6.1 Hz, 1H, H-1’), 4.66 (d, *J* = 4.9 Hz, 1H, H-1’’’), 4.35 (d, *J* = 6.8 Hz, 1H, H-1’’), 4.18 (dd, *J* = 12.1, 4.7 Hz, 1H, H-5), 4.07 – 3.99 (m, 1H, H-4), 3.97 (dd, *J* = 12.4, 3.9 Hz, 1H, H-5’’’), 3.79 – 3.63 (m, 3H, H-4’’, H-4’, H-5’), 3.49 (dd, *J* = 8.5, 5.2 Hz, 1H, H-5), 3.46 (t, *J* = 6.1 Hz, 1H, H-5’’), 3.33 (dd, *J* = 12.4, 6.3 Hz, 1H, H-5’’’), 3.28 (dd, *J* = 11.8, 8.0 Hz, 1H, H-5’), 3.02 (dd, *J* = 12.2, 9.1 Hz, 1H, H-5’’), 2.63 – 2.45 (m, 2H, CH_2(β)), 2.44 – 2.28 (m, 2H, CH_2(γ)), 2.06 (s, 3H, CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.8 (Lev, C(O) δ), 171.2 (Lev, C(O) α), 165.6 (Bz), 165.5 (Bz), 165.5 (Bz), 165.4 (Bz), 165.3 (Bz), 165.3 (Bz), 165.2 (Bz), 165.0 (Bz), 133.5, 133.5, 133.4, 133.4, 133.3, 133.2, 132.9, 132.6, 130.1, 130.0, 130.0, 129.9, 129.9, 129.8, 129.8, 129.7, 129.6, 129.5, 129.5, 129.4, 129.3, 129.2, 129.1, 129.0, 128.6, 128.6, 128.5, 128.5, 128.4, 128.4, 128.1, 100.8 (C-1’), 100.6 (C-1’’), 99.5 (C-1’’’), 86.7 (C-1), 75.4 (C-4), 75.3 (C-4’), 75.1 (C-

4’’), 73.0 (C-3), 72.5 (C-3’’), 71.9 (C-3’), 71.4 (C-2’), 71.3 (C-2’’), 70.5 (C-2), 70.1 (C-2’’’), 69.6 (C-3’’’), 68.6 (C-4’’’), 65.7 (C-5), 62.4 (C-5’’), 62.3 (C-5’), 60.9 (C-5’’’), 37.8 (C-β), 29.8 (C-ε), 27.9 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{87}H_{76}O_{26}S$ ($M+Na^+$) 1591.4237, found 1591.4254.



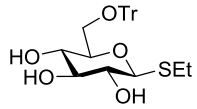
Benzyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside (62)

Compound **62** was synthesized from donor **63** and acceptor **64** following the general procedure of glycosylation, and isolated in 59% yield after purification (toluene/heptane/ethyl acetate 3:3:2).

$[\boldsymbol{\alpha}]_{\mathbf{D}}^{\mathbf{20}}$ – 43.0° (c 0.10, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.00 – 7.82 (m, 20H, Ar*H*), 7.60 – 7.47 (m, 6H, Ar*H*), 7.46 – 7.12 (m, 29H, Ar*H*), 5.56 (t, *J* = 6.7 Hz, 1H, H-3’’’’), 5.54 – 5.50 (dd, *J* = 7.5, 8.1 Hz, 1H, H-3), 5.47 (t, *J* = 8.1 Hz, 1H, H-3’), 5.39 (t, *J* = 8.0 Hz, 1H, H-3’’), 5.29 (dd, *J* = 8.4, 6.6 Hz, 1H, 2), 5.28 (t, *J* = 8.3 Hz, 1H, H-3’’’), 5.14 (dd, *J* = 8.4, 6.5 Hz, 1H, H-2’), 5.09 (dd, *J* = 6.9, 5.1 Hz, 1H, H-2’’’’), 5.04 (dd, *J* = 8.2, 6.3 Hz, 1H, H-2’’), 5.02 (m, 1H, H-4’’’’), 4.81 (dd, *J* = 8.5, 6.7 Hz, 1H, H-2’’’), 4.79 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.69 (d, *J* = 6.3 Hz, 1H, H-1’), 4.65 (d, *J* = 5.0 Hz, 1H, H-1’’’’), 4.63 (d, *J* = 6.7 Hz, 1H, H-1), 4.55 (d, *J* = 12.3 Hz, 1H, OC*H*$_2$Ph), 4.53 (d, *J* = 6.1 Hz, 1H, H-1’’), 4.32 (d, *J* = 6.7 Hz, 1H, H-1’’’), 4.00 – 3.92 (m, 3H, H-4, H-5, H-5’’’’), 3.80 – 3.63 (m, 3H, H-4’, H-4’’, H-4’’’), 3.59 (dd, *J* = 12.2, 4.8 Hz, 1H, H-5’’), 3.45 (dd, *J* = 12.2, 4.6 Hz, 1H, H-5’’’), 3.44 – 3.38 (m, 1H, H-5’), 3.36 (dd, *J* = 12.2, 8.9 Hz, 1H, H-5), 3.32 (dd, *J* =

12.3, 6.2 Hz, 1H, H-5’’’’), 3.19 (dd, *J* = 12.1, 8.4 Hz, 1H, H-5’’), 3.08 (dd, *J* = 12.2, 8.6 Hz, 1H, H-5’), 2.99 (dd, *J* = 12.2, 9.0 Hz, 1H, H-5’’’), 2.60 – 2.45 (m, 2H, CH_2(β)), 2.42 – 2.28 (m, 2H, CH_2(γ)), 2.05 (s, 3H, CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.7 (Lev, C(O) δ), 171.1 (Lev, C(O) α), 165.5 (Bz), 165.4 (Bz), 165.3 (Bz), 165.3 (Bz), 165.3 (Bz), 165.2 (2xBz), 165.0 (Bz), 164.9 (Bz), 164.9 (Bz), 136.8, 133.4, 133.4, 133.3, 133.3, 133.2, 133.1, 133.0, 133.0, 129.9, 129.9, 129.9, 129.8, 129.7, 129.7, 129.6, 129.6, 129.5, 129.5, 129.4, 129.3, 129.3, 129.2, 129.0, 128.9, 128.5, 128.5, 128.4, 128.4, 128.4, 128.3, 128.3, 128.2, 128.2, 127.8, 127.8, 101.0 (C-1’), 100.4 (C-1’’’), 100.4 (C-1’’), 99.4 (C-1’’’’), 99.3 (C-1), 76.0 (C-4), 75.2 (C-4’’), 75.1 (C-4’), 74.9 (C-4’’’), 72.4 (C-3’’’), 72.1 (C-3), 72.0 (C-3’), 71.9 (C-3’’), 71.4 (C-2’’’), 71.3 (C-2’), 71.1 (C-2’’), 71.0 (C-2), 70.3 (O*C*H_2Ph), 70.0 (C-2’’’’), 69.5 (C-3’’’’), 68.5 (C-4’’’’), 62.6 (C-5), 62.3 (C-5’’’), 62.2 (C-5’, C-5’’), 60.7 (C-5’’’’), 37.7(C-β), 29.6 (C-ε), 27.7 (C-γ). **HRMS** (MALDI) m/z calcd for $C_{107}H_{94}O_{33}$ ($M+Na^+$) 1930.5603, found 1930.5619.

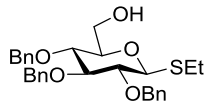
Branching units



Ethyl 6-*O*-trityl-1-thio-β-D-glucopyranoside (97)

Ethyl 1-thio-β-D-glucopyranoside (22.0 g, 0.0891 mol) was dissolved in 120 ml of pyridine and trityl chloride (32.8 g, 0.118 mol) was added. The mixture was stirred at 90 °C for 3 hours then cooled to room temperature and diluted with ethyl acetate (300 ml). The mixture was washed with 1 M HCl (2x300 ml) and the organic layer isolated, dried over Na_2SO_4, filtered and concentrated. The crude material was purified by flash chromatography (heptane/acetone 6:4, R_f 0.23) to yield **97** (31.8 g, 70%) as a white amorphous solid.

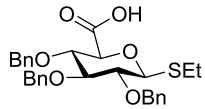
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 7.22 (m, 15H, Ar*H*), 4.34 (d, *J* = 9.6 Hz, H-1), 3.66 – 3.48 (m, 2H, H-3, H-4), 3.41 – 3.29 (m, 4H, H-2, H-5, H-6, H-6), 2.91 (s, 1H, OH), 2.88 (s, 1H, OH), 2.84 – 2.59 (m, 2H, S*CH*$_2$CH$_3$), 2.54 (d, *J* = 1.3 Hz, 1H, OH), 1.27 (t, *J* = 7.4 Hz, 3H, SCH$_2$*CH*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.6, 128.6, 128.0, 127.3, 85.9 (C-1), 77.8, 77.8, 72.3, 72.3, 64.7 (C-6), 24.4 (S*CH*$_2$CH$_3$), 22.8 (SCH$_2$*CH*$_3$). The data are in accordance with literature.[233]



Ethyl 2,3,4-tri-*O*-benzyl-1-thio-β-D-glucopyranoside (98)

Triol **97** (22.6 g, 484 mmol) was dissolved in anhydrous DMF (200 ml) and NaH (9.68 g, 242 mmol, 60% oil dispersion) was added. The reaction mixture was cooled to 0 °C and stirred for 10 minutes. TBAI (1.25 g, 3.39 mmol) and BnBr (28.7 ml, 242 mmol) were added slowly to the mixture. The reaction was stirred for 16 hours at room temperature then the excess of NaH and BnBr was quenched with methanol. The mixture was diluted with Et_2O (650 ml) and washed with water (2x400 ml). The organic phases were collected and dried over Na_2SO_4, filtered and concentrated under reduced pressure. The crude product was dissolved directly in methanol (450 ml) and concentrated H_2SO_4 (4.5 ml) was added. The reaction mixture was stirred 1 hour then Na_2CO_3 (38.1 g) was added to neutralize the reaction. After 2 hours the salts were filtered off and the filtrate was diluted with dichloromethane (500 ml) and washed with water (2x450 ml). The organic layers were combined and dried over Na_2SO_4. Filtration and evaporation of the solvent under vacuum gave the crude material which was purified by column chromatography (heptane/acetone 6:4, R_f 0.38) to yield **98** (13.3 g, 56%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.21 (m, 15H, Ar*H*), 4.93 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.92 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.90 – 4.83 (m, 2H, OC*H*$_2$Ph), 4.75 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.66 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.51 (d, *J* = 9.8 Hz,1H, H-1), 3.87 (dd, *J* = 12.0, 2.6 Hz, 1H, H-6), 3.76 – 3.65 (m, 2H, H-3, H-6), 3.58 (t, *J* = 9.4 Hz, 1H, H-4), 3.41 (t, *J* = 9.5 Hz, 1H, H-2), 3.40 – 3.34 (m, 1H, H-5), 2.84 – 2.66 (m, 2H, SC*H*$_2$CH$_3$), 1.33 (t, *J* = 7.4 Hz, 3H, SCH$_2$*CH*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.4, 137.9, 128.5, 128.5, 128.4, 128.3, 128.1, 128.0, 127.9, 127.8, 127.7, 86.5 (C-3), 85.3 (C-1), 81.8 (C-2), 79.3 (C-5), 77.7 (C-4), 75.8 (O*C*H$_2$Ph), 75.6 (O*C*H$_2$Ph), 75.2 (O*C*H$_2$Ph), 62.2 (C-6), 25.2 (S*CH*$_2$CH$_3$), 15.2 (SCH$_2$*CH*$_3$). The data are in accordance with literature.[192]

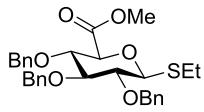


Ethyl 2,3,4-tri-*O*-benzyl-1-thio-β-D-glucopyranosiduronic acid (99)

The alcohol **98** (0.500 g, 1.01 mmol) was dissolved in a solvent mixture CH_2Cl_2/H_2O (2:1, 5 ml) and stirred vigorously. TEMPO (0.032 g, 0.202 mmol) and $PhI(OAc)_2$ (0.815 g, 2.53 mmol) were added to the mixture. Full conversion of the starting material was observed after 50 minutes with TLC and the remaining oxidant was quenched using $Na_2S_2O_3$ (20 ml, 10% solution). The water phase was extracted with ethyl acetate (2x20 ml), then organic layers were dried over $NaSO_4$, filtered and the solvent evaporated under vacuum. The carboxylic acid **99** (0.416 g, 81%) was isolated by flash chromatography (heptane/ethyl acetate/AcOH 7:3:0.5, R_f 0.40).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40 – 7.20 (m, 15H, Ar*H*), 4.90 (d, *J* = 10.3 Hz, 1H, OC*H*$_2$Ph), 4.86 (d, *J* = 11.1 Hz, 1H, OC*H*$_2$Ph), 4.81 (d, *J* = 11.1 Hz, 1H, OC*H*$_2$Ph), 4.78 (d, *J* = 10.7 Hz, 1H, OC*H*$_2$Ph), 4.72 (d, *J* = 10.3 Hz, 1H, OC*H*$_2$Ph),

4.66 (d, *J* = 10.7 Hz, 1H, OC*H*$_2$Ph), 4.59 (d, *J* = 9.7 Hz, 1H, H-1), 3.98 (d, *J* = 9.1 Hz, 1H, H-5), 3.83 (t, *J* = 8.9 Hz, 1H, H-4), 3.72 (t, *J* = 8.4 Hz, 1H, H-3), 3.49 (dd, *J* = 9.4, 8.3 Hz, 1H, H-2), 2.84 – 2.67 (m, 2H, S*CH*$_2$CH$_3$), 1.32 (t, *J* = 7.4 Hz, 3H, SCH$_2$*CH*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.4 (*C*OOH), 138.0, 137.7, 137.3, 128.5, 128.5, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 85.5 (C-1), 85.0 (C-3), 81.0 (C-2), 78.6 (C-4), 77.2 (C-5), 75.6 (O*C*H$_2$Ph), 75.4 (O*C*H$_2$Ph), 75.0 (O*C*H$_2$Ph), 25.4 (S*CH*$_2$CH$_3$), 15.0 (SCH$_2$*CH*$_3$). The data are in accordance with literature.[69]

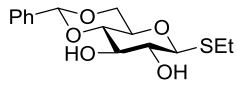


Methyl (ethyl 2,3,4-tri-*O*-benzyl-1-thio-β-D-glucopyranosid) uronate (93)

Glucuronic acid **99** (0.200 g, 0.393 mmol) was dissolved in a solvent mixture CH_3OH/toluene (1:1, 6 ml) and a solution of Me_3SiCHN_2 (0.77 ml, 2M in hexane) was slowly added to the mixture. After 3 h TLC showed full conversion of the starting material and the excess reagent was quenched with acetic acid (2 ml). The reaction mixture was concentrated and the crude material was purified by column chromatography (heptane/ethyl acetate 8:2, R_f 0.36) to yield the methyl ester **93** as a white amorphous solid (0.189 g, 92%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.44 – 6.98 (m, 15H, Ar*H*), 4.92 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.91 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.85 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.78 (d, *J* = 10.8 Hz, 1H, OC*H*$_2$Ph), 4.74 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.61 (d, *J* = 10.8 Hz, 1H, OC*H*$_2$Ph), 4.50 (d, *J* = 9.7 Hz, 1H, H-1), 3.89 (d, *J* = 9.7 Hz, 1H, H-5), 3.84 (t, *J* = 9.2 Hz, 1H, H-4), 3.72 (s, 3H, OC*H*$_3$), 3.71 (m, 1H, H-3), 3.40 (t, *J* = 9.2 Hz, 1H, H-2), 2.76 – 2.59 (m, 2H, S*CH*$_2$CH$_3$), 1.24 (t, *J* = 7.3 Hz, 3H, SCH$_2$*CH*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 168.7 (*C*(O)-OCH$_3$), 138.3, 137.8,

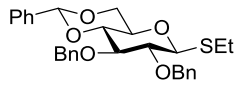
128.5, 128.4, 128.3, 128.0, 127.9, 127.9, 129.8, 127.8, 85.9 (C-1), 85.8 (C-3), 81.2 (C-2), 79.3 (C-4), 78.1 (C-5), 75.9 (O*C*H_2Ph), 75.6 (O*C*H_2Ph), 75.1 (O*C*H_2Ph), 52.5 (O*C*H_3), 25.2 (S*C*$H_2$$CH_3$), 15.0 (S$CH_2$*C*$H_3$). The data are in accordance with literature.[69]



Ethyl 4,6-*O*-benzylidene-1-thio-β-D-glucopyranoside (103)

Ethyl 1-thio-β-D-glucopyranoside (8.26 g, 36.8 mmol) was dissolved in anhydrous CH_3CN (250 ml) together with $PhCH(OMe)_2$ (16.6 ml, 110 mmol). CSA (4.3 g, 18.4 mmol) was added and the reaction was stirred at 22 °C for 24 h. The mixture was neutralized with Et_3N (3.6 ml, 25.8 mmol), filtered and the solvent evaporated under reduced pressure. The crude was purified by flash chromatography to yield **103** (4.5 g, 39%) as a white amorphous solid.

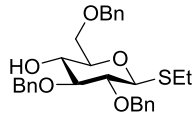
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.54 – 7.44 (m, 2H, Ar*H*), 7.42 – 7.33 (m, 3H, Ar*H*), 5.53 (s, 1H, C*H*Ph), 4.45 (d, *J* = 9.8 Hz, 1H, H-1), 4.34 (dd, *J* = 10.5, 4.8 Hz, 1H, H-6), 3.81 (t, *J* = 8.8 Hz, 1H, H-3), 3.76 (t, *J* = 10.1 Hz, 1H, H-6), 3.56 (t, *J* = 9.2 Hz, 1H, H-4), 3.49 (t, *J* = 9.8 Hz, 1H, H-2), 3.52 – 3.45 (m, 1H, H-5), 2.75 (qd, *J* = 7.4, 2.0 Hz, 2H, SC*H*$_2$$CH_3$), 1.32 (t, *J* = 7.4 Hz, 3H, SCH_2C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 136.9, 129.3, 128.4, 126.3, 101.9 (*C*HPh), 86.6 (C-1), 80.4 (C-4), 74.6 (C-3), 73.2 (C-2), 70.6 (C-5), 68.6 (C-6), 24.8 (S*C*$H_2$$CH_3$), 15.3 ($SCH_2$*C*$H_3$). The data are in accordance with literature.[234]



Ethyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-1-thio-β-D-glucopyranoside (102)

Diol **103** (4.2 g, 13.4 mmol) was dissolved in DMF (120 ml), then NaH (1.2 g, 51.6 mmol, 60% oil dispersion) was added and the mixture was stirred at 0 °C. BnBr (4.00 ml, 33.6 mmol) was added and the temperature was raised to room temperature. The reaction mixture was stirred for 20 h, and the remaining NaH and BnBr were quenched with methanol (35 ml) and diluted with Et_2O (200 ml). The organic phase was washed with water (2x250 ml), dried over Na_2SO_4, filtered, concentrated and purified by flash chromatography (heptane/ethyl acetate 9:1, R_f 0.24) to give **102** (4.25 g, 64%) as a colorless oil.

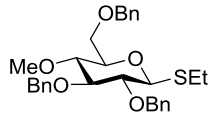
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.52 – 7.46 (m, 2H, Ar*H*), 7.42 – 7.27 (m, 13H, Ar*H*), 5.59 (s, 1H, C*H*Ph), 4.96 (d, *J* = 11.3 Hz, 1H, OCH_2Ph), 4.89 (d, *J* = 10.2 Hz, 1H, OCH_2Ph), 4.82 (d, *J* = 10.2 Hz, 1H, OCH_2Ph), 4.81 (d, *J* = 11.3 Hz, 1H, OCH_2Ph), 4.57 (d, *J* = 9.8 Hz, 1H, H-1), 4.36 (dd, *J* = 10.5, 5.0 Hz, 1H, H-6), 3.82 (dd, *J* = 8.2, 9.3 Hz, 1H, H-3), 3.77 (t, *J* = 10.5 Hz, 1H, H-6), 3.72 (t, *J* = 9.3 Hz, 1H, H-4), 3.47 (dd, *J* = 9.7, 8.2 Hz, 1H, H-2), 3.49 – 3.40 (m, 1H, H-5), 2.85 – 2.68 (m, 2H, SC$H_2$$CH_3$), 1.33 (t, *J* = 7.4 Hz, 3H, SCH_2CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.4, 138.0, 137.3, 129.0, 128.4, 128.3, 128.3, 128.1, 127.9, 127.7, 126.0, 101.1 (*C*HPh), 85.9 (C-1), 82.8 (C-3), 81.6 (C-4), 81.3 (C-2), 76.0 (O*C*H_2Ph), 75.2 (O*C*H_2Ph), 70.2 (C-5), 68.7 (C-6), 25.2 (S*C*H_2CH_3), 15.1 (SCH_2*C*H_3). The data are in accordance with literature.[235]



Ethyl 2,3,6-tri-*O*-benzyl-1-thio-β-D-glucopyranoside (101)

Compound **102** (1.64 g, 3.33 mmol) was dissolved in dichloromethane (7 ml) and the solution was cooled to 0 °C. TFAA (1.4 ml, 9.99 mmol) and Et_3SiH (2.7 ml, 16.6 mmol) were added and the mixture was stirred for 5 minutes, and TFA (1.3 ml, 16.6 mmol) was added dropwise. The temperature was slowly raised to room temperature. TLC showed full conversion of the starting material after 3 h. The reaction mixture was diluted with ethyl acetate (30 ml), washed with saturated $NaHCO_3$ (25 ml) and brine (30 ml). The organic layers were combined, dried over Na_2SO_4, filtered, concentrated and purified by flash chromatography to obtain **101** (1.35 g, 82%) as a white powder.

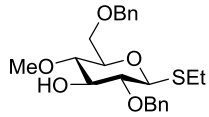
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.27 (m, 15H, Ar*H*), 4.93 (d, *J* = 11.4 Hz, 1H, OC*H*$_2$Ph), 4.92 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.80 (d, *J* = 11.4 Hz, 1H, OC*H*$_2$Ph), 4.75 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.60 (d, *J* = 11.9 Hz, 1H, OC*H*$_2$Ph), 4.56 (d, *J* = 11.9 Hz, 1H, OC*H*$_2$Ph), 4.49 (d, *J* = 9.6 Hz, 1H, H-1), 3.80 – 3.68 (m, 2H, H-6), 3.64 (t, *J* = 9.2 Hz, 1H, H-4), 3.52 (t, *J* = 8.7 Hz, 1H, H-3), 3.49 – 3.44 (m, 1H, H-5), 3.42 (dd, *J* = 8.7, 9.6 Hz, 1H, H-2), 2.84 – 2.67 (m, 2H, SC*H*$_2$CH$_3$), 1.33 (t, *J* = 7.4 Hz, 3H, SCH$_2$C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.6, 137.9, 137.8, 128.6, 128.4, 128.4, 128.4, 127.9, 127.9, 127.8, 127.7, 86.0 (C-3), 85.2 (C-1), 81.3 (C-2), 77.9 (C-5), 75.5 (O*C*H$_2$Ph), 75.4 (O*C*H$_2$Ph), 73.7 (O*C*H$_2$Ph), 72.1 (C-4), 70.6 (C-6), 25.1 (S*C*H$_2$CH$_3$), 15.2 (SCH$_2$*C*H$_3$). The data are in accordance with literature.[192]



Ethyl 2,3,6-tri-*O*-benzyl-4-*O*-methyl-1-thio-β-D-glucopyranoside (104)

NaH (0.230 g, 5.78 mmol, 60% oil dispersion) was suspended in DMF (15 ml) and stirred at 0 °C. A solution of **101** (1.30 g, 2.63 mmol) in DMF (5 ml) was added dropwise and the suspension was stirred for 1 h at 0 °C. Then MeI (0.245 ml, 3.94 mmol) was added. The mixture was stirred for 1 h and the remaining reagents were quenched by addition of ice-cold water. The resulting mixture was extracted several times with ethyl acetate. The combined organic layers were dried over Na_2SO_4, filtered and the solvent evaporated under reduced pressure. The crude product was purified by column chromatography (heptane/ethyl acetate 85:15) to give the final compound **104** (1.30 g, 95%).

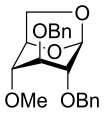
$[\alpha]_D^{20}$ + 11.2° (c 0.41, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.42 – 7.27 (m, 15H, Ar*H*), 4.91 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.90 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.84 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.74 (d, *J* = 10.2 Hz, 1H, OC*H*$_2$Ph), 4.65 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.58 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.45 (d, *J* = 9.7 Hz, 1H, H-1), 3.77 (dd, *J* = 10.9, 1.7 Hz, 1H, H-6), 3.70 (dd, *J* = 10.9, 4.7 Hz, 1H, H-6), 3.59 (t, *J* = 8.8 Hz, 1H, H-3), 3.51 (s, 3H, OC*H*$_3$), 3.41 (t, *J* = 9.4 Hz, 1H, H-2), 3.42 – 3.36 (m, 1H, H-5), 3.32 (dd, *J* = 8.8, 9.7 Hz, 1H, H-4), 2.88 – 2.67 (m, 2H, SC*H*$_2$CH$_3$), 1.34 (t, *J* = 7.4 Hz, 3H, SCH$_2$C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.6, 138.3, 138.0, 128.4, 128.4, 128.3, 127.9, 127.8, 127.7, 127.7, 127.6, 86.6 (C-3), 85.0 (C-1), 81.6 (C-2), 80.0 (C-4), 79.2 (C-5), 75.7 (O*C*H$_2$Ph), 75.5 (O*C*H$_2$Ph), 73.5 (O*C*H$_2$Ph), 69.3 (C-6), 60.7 (O*C*H$_3$), 25.0 (S*C*H$_2$CH$_3$), 15.2 (SCH$_2$*C*H$_3$). **HRMS** (MALDI) m/z calcd for $C_{30}H_{36}O_5S$ ($M+Na^+$) 531.2175, found 531.2187.



Ethyl 2,6-di-*O*-benzyl-4-*O*-methyl-1-thio-β-D-glucopyranoside (105)

Compound **104** (0.200 g, 0.393 mmol) and DDQ (0.134 g, 0.590 mmol) were dissolved in dry CH_3CN (6 ml) and the reaction flask irradiated with a UV lamp (364 nm) for 48 h. The reaction mixture was washed with a buffer solution of 0.7% ascorbic acid (0.104 g, 0.590 mmol), 1.5% citric acid (0.225 g) and 1% NaOH (0.150 g) in water (15 ml). The organic layer was diluted with dichloromethane (15 ml) and washed with brine (25 ml). The organic phase was dried over Na_2SO_4, filtered and the solvent evaporated under vacuum. The compound **105** (0.084 g, 51%) was isolated as the major product from flash chromatography (heptane/ethyl acetate 7:3, R_f 0.35).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.45 – 7.27 (m, 10H, Ar*H*), 4.97 (d, *J* = 11.1 Hz, 1H, OC*H*$_2$Ph), 4.67 (d, *J* = 11.1 Hz, 1H, OC*H*$_2$Ph), 4.64 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.57 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.42 (d, *J* = 9.7 Hz, 1H, H-1), 3.76 (dd, *J* = 10.9, 1.9 Hz, 1H, H-6), 3.68 (dd, *J* = 10.8, 5.0 Hz, 1H, H-6), 3.64 (t, *J* = 8.8 Hz, 1H, H-3), 3.50 (s, 3H, OC*H*$_3$), 3.37 (ddd, *J* = 9.7, 5.0, 1.9 Hz, 1H, H-5), 3.28 (dd, *J* = 8.8, 9.6 Hz, 1H, H-2), 3.18 (dd, *J* = 9.0, 9.7 Hz, 1H, H-4), 2.88 – 2.60 (m, 2H SC*H*$_2$CH$_3$), 1.34 (t, *J* = 7.4 Hz, 3H, SCH$_2$C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.2, 138.2, 138.0, 128.6, 128.3, 128.3, 128.1, 127.7, 127.6, 84.6 (C-1), 81.3 (C-2), 79.3 (C-4), 78.8 (C-5), 78.4 (C-3), 75.1 (O*C*H$_2$Ph), 73.5 (O*C*H$_2$Ph), 69.3 (C-6), 60.5 (O*C*H$_3$), 25.1 (S*C*H$_2$CH$_3$), 15.1 (SCH$_2$*C*H$_3$). The structure was elucidated by means of 2D NMR experiments, specifically an HMBC experiment was necessary to identify which of the Bn groups was hydrolyzed.

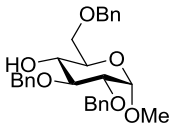


1,6-anhydro-2,3-*O*-benzyl-4-*O*-methyl-D-glucopyranose (106)

Compound **104** (0.200 g, 0.393 mmol) was dissolved in dry dichloromethane (8 ml) and the mixture was cooled to 0 °C. $FeCl_3$ (0.127 g, 0.786 mmol) was added and the temperature was slowly raised to room temperature over 2 h. The reaction was quenched with H_2O (0.5 ml), diluted with dichloromethane and the mixture was washed two times with brine (2x20 ml). The organic layers were combined, dried over Na_2SO_4, filtered and the solvent evaporated under reduced pressure. The crude product was purified by silica gel chromatography (heptane/ethyl acetate 7:3, R_f 0.30) to yield **106** as the main product.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33 – 7.17 (m, 10H, Ar*H*), 5.37 (br s, 1H, H-1), 4.57 (dd, *J* = 5.9, 1.1 Hz, 1H, H-5), 4.52 (d, *J* = 12.5 Hz, 1H, OC*H*$_2$Ph), 4.46 (d, *J* = 12.5 Hz, 1H, OC*H*$_2$Ph), 4.46 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.40 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 3.92 (dd, *J* = 7.1, 0.9 Hz, 1H, H-6), 3.67 (dd, *J* = 7.0, 6.0 Hz, 1H, H-6), 3.53 – 3.46 (m, 1H, H-3), 3.34 (s, 3H, OC*H*$_3$), 3.30 – 3-24 (m, 1H, H-2), 3.11 (br s, 1H, H-4). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 137.9, 137.8, 128.5, 128.1, 127.9, 127.9, 127.7, 100.6 (C-1), 79.7 (C-4), 75.8 (C-3), 75.5 (C-2), 73.4 (C-5), 72.0 (O*C*H$_2$Ph), 71.8 (O*C*H$_2$Ph), 65.2 (C-6), 57.2 (OC*H*$_3$).

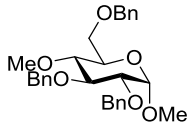
Glucuronate esters



Methyl 2,3,6-tri-*O*-benzyl-α-D-glucopyranoside (114)

Compound **113** (5 g, 10.8 mmol) was dissolved in dry CH_2Cl_2 (20 ml) and cooled to 0 °C before TFAA (4.58 ml, 32.4 mmol) and Et_3SiH (8.65 ml, 54 mmol) were added. After 5 min at 0 °C TFA (4.13 ml, 54.0 mmol) was added dropwise. The reaction mixture was stirred for 4 hours at room temperature and diluted with ethyl acetate, washed with saturated $NaHCO_3$, brine then dried over $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The crude product was purified by flash chromatography (heptane/ethyl acetate 7:3). Compound **114** was obtained as a colorless oil (3.86 g, 77%).

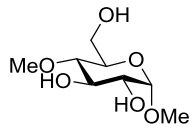
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 – 7.17 (m, 15H, Ar*H*), 4.91 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.80 (d, *J* = 11.0 Hz, 1H, OC*H*$_2$Ph), 4.76 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.71 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.58 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.56 (d, *J* = 11.0 Hz, 1H, OC*H*$_2$Ph), 4.49 (d, *J* = 3.6 Hz, 1H, H-1), 3.93 (t, *J* = 9.3 Hz, 1H, H-3), 3.68 (dd, *J* = 11.5, 2.4 Hz, 1H, H-6), 3.62 (dd, *J* = 11.5, 3.9 Hz, 1H, H-6), 3.60 – 3.53 (m, 1H, H-5), 3.44 (t, *J* = 9.3 Hz, 1H, H-4), 3.42 (dd, *J* = 9.6, 3.6 Hz, 1H, H-2), 3.28 (s, 3H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.8, 138.2, 138.1, 128.5, 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.7, 98.2 (C-1), 82.0, 80.0, 75.8 (O*C*H$_2$Ph), 75.1 (O*C*H$_2$Ph), 73.5 (O*C*H$_2$Ph), 70.7, 61.8, 55.2 (O*C*H$_3$). The data are in accordance with literature.[227]



Methyl 2,3,6-tri-*O*-benzyl-4-*O*-methyl-α-D-glucopyranoside (115)

A solution of **114** (1 g, 2.15 mmol) in DMF (6 ml) was added dropwise to a suspension of NaH (0.240 g, 4.73 mmol) in dry DMF (10 ml) at 0 °C. The suspension was stirred for 1 h at 0 °C and then MeI was added. The mixture was stirred for 2 h and the remaining reagents were quenched by addition of ice-cold water. The resulting mixture was extracted several times with ethyl acetate. The combined organic layers were dried over $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The crude product was purified through by column chromatography (heptane/ethyl acetate 7:3) to give the final compound **115** (0.946 g, 92%) as a colorless oil.

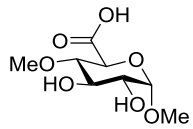
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.54 – 7.16 (m, 15H, Ar*H*), 4.95 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.81 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.80 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.65 (d, *J* = 12.2 Hz, 1H, OC*H*$_2$Ph), 4.64 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.62 (d, *J* = 3.6 Hz, 1H, H-1), 4.52 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 3.87 (t, *J* = 9.3 Hz, 1H, H-3), 3.71 (dd, *J* = 10.8, 4.2 Hz, 1H, H-6), 3.68 – 3.57 (m, 2H, H-5, H-6), 3.52 (dd, *J* = 9.7, 3.6 Hz, 1H, H-2), 3.47 (s, 3H, OC*H*$_3$), 3.38 (s, 3H, OC*H*$_3$), 3.34 (t, *J* = 9.3 Hz, 1H, H-4). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 144.0, 138.7, 138.1, 128.7, 128.5, 128.4, 128.1, 128.1, 128.0, 127.8, 127.6, 126.9, 98.2 (C-1), 82.0, 80.0, 75.8 (O*C*H$_2$Ph), 75.1 (O*C*H$_2$Ph), 73.5 (O*C*H$_2$Ph), 70.7, 61.9, 55.2 (O*C*H$_3$), 52.1 (O*C*H$_3$). The data are in accordance with literature.[228]



Methyl 4-*O*-methyl-α-D-glucopyranoside (116)

A solution of **115** (0.625 g, 1.31 mmol) in methanol (20 ml) was stirred and Pd 10% on activated charcoal (50 mg) was added. The suspension was degassed and backfilled with H_2 three times then stirred under a hydrogen atmosphere (1 atm). After 18 h at room temperature the reaction mixture was filtered through a Celite pad with methanol. The filtrate was concentrated under reduced pressure and purified by flash chromatography (CH_2Cl_2/CH_3OH 8:2) to yield the triol **116** (0.270 g, 99%)

^{1}H NMR (400 MHz, $CDCl_3$) δ 4.75 (d, *J* = 3.9 Hz, 1H, H-1), 3.85 (dd, *J* = 11.9, 2.8 Hz, 1H, H-6), 3.78 (t, *J* = 9.2 Hz, 1H, H-3), 3.75 (dd, *J* = 11.9, 4.2 Hz, 1H, H-6), 3.58 (s, 3H, OC*H*$_3$), 3.57 – 3.53 (m, 1H, H-5), 3.52 – 3.49 (dd, *J* = 9.5, 3.9 Hz, 1H, H-2), 3.41 (s, 3H, OC*H*$_3$), 3.23 – 3.12 (dd, *J* = 9.2, 9.7 Hz, 1H, H-4). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 99.1 (C-1), 79.2, 74.7, 72.6, 70.8, 61.9, 60.7 (O*C*H$_3$), 55.9 (O*C*H$_3$). The data are in accordance with literature.[228]

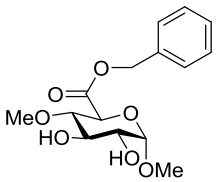


Methyl 4-*O*-methyl-α-D-glucopyranosidic acid (117)

The compound **116** (0.750 g, 3.60 mmol) was dissolved in water (10 ml) together with TEMPO (0.112 g, 0.720 mmol) and NaBr (0.180 g, 1.80 mmol) and stirred at 0 °C. NaClO solution (3.8 ml, 15.8 mmol) was added dropwise to the mixture. The pH was kept around 10-11 by addition of 3 drops of NaOH (2 M solution). After 3 h the reaction was neutralized with HCl and concentrated, redissolved with

methanol and filtered. The resulting solution was concentrated and purified by column chromatography (CH_2Cl_2/CH_3OH/AcOH 8:1.5:0.5) to give **117** as a white solid (0.726 g, 90%).

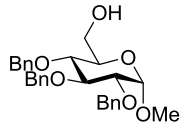
^{1}H NMR (400 MHz, CD_3OD) δ 4.79 (d, *J* = 3.6 Hz, 1H, H-1), 3.93 (d, *J* = 9.9 Hz, 1H, H-5), 3.70 (t, *J* = 9.3 Hz, 1H, H-3), 3.53 (s, 3H, OC*H*$_3$), 3.58 – 3.50 (m, 1H, H-2), 3.43 (s, 3H, OC*H*$_3$), 3.26 (t, *J* = 9.4 Hz, 1H, H-4). **^{13}C NMR** (101 MHz, CD_3OD) δ 175.6 (*C*OOH), 100.0 (C-1), 82.3, 73.2, 71.7, 59.4, 54.7 (O*C*H$_3$), 20.2. The data are in accordance with literature.[228]



Benzyl (methyl 4-*O*-methyl-α-D-glucopyranoside) uronate (110)

Glucuronic acid **117** (250 mg, 1.13 mmol) was dissolved in dry DMF (4 ml) and cooled to 0 °C. TBAF (1 M solution in THF, 1.24 ml, 1.24 mmol) and BnBr (0.147 ml, 1.24 mmol) were added to the solution and the reaction mixture let to stir at room temperature for 20 hours. The solvent was evaporated and the residue purified by silica gel column chromatography (heptane/ethyl acetate 7:3) to afford **110** as a colorless oil (67.5 mg, 20%).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43 – 7.30 (m, 5H, Ar*H*), 5.26 (br s, 2H, OC*H*$_2$Ph), 4.82 (d, *J* = 3.8 Hz, 1H, H-1), 4.11 (d, *J* = 9.8 Hz, 1H, H-5), 3.77 (t, *J* = 9.2 Hz, 1H, H-3), 3.59 (m, 1H, H-2), 3.45 (s, 3H, OC*H*$_3$(anom.)), 3.37 (s, 3H, OC*H*$_3$), 3.37 (t, *J* = 9.4 Hz, 1H, H-4), 2.58 (s, 1H, OH), 2.08 (s, 1H, OH). **^{13}C NMR** (100 MHz, $CDCl_3$): δ 169.4 (*C*OOBn), 128.6, 128.5, 99.4 (C-1), 80.9, 74.5 (*O*CH$_2$Ph), 72.1, 70.3, 67.4, 60.4 (O*C*H$_3$), 55.9 (O*C*H$_3$). **HRMS** (ESIMS) m/z calcd for $C_{15}H_{20}NaO_7$ ($M+Na^+$) 335.2946, found 335.1101.

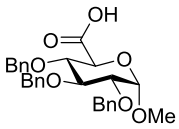


Methyl 2,3,4-tri-*O*-benzyl-α-D-glucopyranoside (120)

A solution of methyl α-D-glucopyranoside (**118**) (15.0 g, 77.3 mmol) and trityl chloride (23.6 g, 92.8 mmol) in dry pyridine (200 ml) was stirred at 90 °C until disappearance of starting material. After 4 h, the solvent was evaporated under reduced pressure, the reaction mixture diluted with dichloromethane (250 ml) and then washed with water (2x300 ml). The combined organic layers were dried over $MgSO_4$, filtered and evaporated under reduced pressure. The tritylated compound was crystallized from toluene, filtered and washed with heptane (28.9 g, 86%). The compound **119** (5.02 g, 11.4 mmol) was then dissolved in dry DMF with NaH (2.31 g, 57.0 mmol). After stirring for 20 minutes BnBr and TBAI are added at 0 °C. Reaction mixture was stirred at room temperature for 18 h and diluted with ethyl acetate. The resulting mixture was washed with brine, dried with $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The fully protected crude product was dissolved in methanol (300 ml) with 1% of H_2SO_4 and stirred at room temperature for 1 h until the full conversion of the starting material was observed by TLC. Na_2CO_3 (7.30 g) was added to the reaction mixture which was stirred until pH 7 was reached. After 1 h the mixture was filtered, concentered, then diluted with CH_2Cl_2, washed two times with brine. The organic layers were dried over $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The crude material was purified by flash chromatography (heptane/ethyl acetate, 7:3) to give **120** as a white solid (4.20 g, 80%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 – 7.25 (m, 15H, Ar*H*), 5.01 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.91 (d, *J* = 11.0 Hz, 1H, OC*H*$_2$Ph), 4.86 (d, *J* = 10.9 Hz, 1H,

OC*H*$_2$Ph), 4.82 (d, *J* = 12.0 Hz, 1H, OC*H*$_2$Ph), 4.68 (d, *J* = 12.0 Hz, 1H, OC*H*$_2$Ph), 4.66 (d, *J* = 11.0 Hz, 1H, OC*H*$_2$Ph), 4.59 (d, *J* = 3.6 Hz, 1H, H-1), 4.03 (t, *J* = 9.3 Hz, 1H, H-3), 3.79 (dd, *J* = 11.5, 2.4 Hz, 1H, H-6), 3.71 (dd, *J* = 11.7, 3.8 Hz, 1H, H-6), 3.70 – 3.61 (m, 1H, H-5), 3.55 (t, *J* = 9.3 Hz, 1H, H-4), 3.52 (dd, *J* = 9.5, 3.6 Hz, 1H, H-2), 3.38 (s, 3H, OC*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.8, 138.2, 138.1, 138.0, 128.5, 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 128.0, 127.9, 127.7, 98.2 (C-1), 82.0, 80.0, 75.8 (O*C*H$_2$Ph), 75.1 (O*C*H$_2$Ph), 73.5 (O*C*H$_2$Ph), 70.7, 61.8, 55.2 (O*C*H$_3$). The data are in accordance with literature.[236]

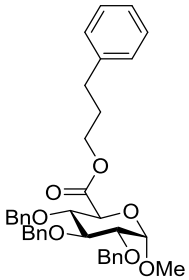


Methyl 2,3,4-tri-*O*-benzyl-α-D-glucopyranosidic acid (121)

A solution of **120** (4.20 g, 9.04 mmol) in CH_2Cl_2 (40 ml) was vigorously stirred when water (20 ml) was added. $PhI(OAc)_2$ (7.18 g, 22.6 mmol) and TEMPO (0.282 g, 1.81 mmol) were added. After 2 hours the remaining oxidants were quenched with a solution of $Na_2S_2O_3$ (10%). The mixture was diluted with ethyl acetate and washed two times with brine. The organic phase was dried over Na_2SO_4, filtered, concentrated and purified by column chromatography (heptane/ethyl acetate/AcOH 6:4:0.3). The product **121** was obtained as a colorless oil (3.20 g, 74%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34 – 7.10 (m, 15H, Ar*H*), 4.88 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.74 (d, *J* = 11.0 Hz, 1H, OC*H*$_2$Ph), 4.72 (d, *J* = 12.1 Hz, 2H, OC*H*$_2$Ph), 4.56 (d, *J* = 12.0 Hz, 1H, OC*H*$_2$Ph), 4.55 (d, *J* = 12.0 Hz, 1H, OC*H*$_2$Ph), 4.54 (d, *J* = 3.6 Hz, 1H, H-1), 4.14 (d, *J* = 10.1 Hz, 1H, H-5), 3.93 (t, *J* = 9.3 Hz, 1H, H-3), 3.63 (dd, *J* = 9.8, 9.2 Hz, 1H, H-4), 3.49 (dd, *J* = 9.6, 3.5 Hz, 1H, H-2), 3.32 (s, 3H, OC*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 173.6 (*C*OOH), 138.4, 137.8, 137.5, 128.6, 128.5, 128.4, 128.2, 128.1, 128.0, 128.0, 127.8, 98.6 (C-1), 81.4,

79.2, 79.1, 75.9 (O*C*H_2Ph), 75.3 (O*C*H_2Ph), 73.7 (O*C*H_2Ph), 69.6, 55.8 (O*C*H_3). The data are in accordance with literature.[237]

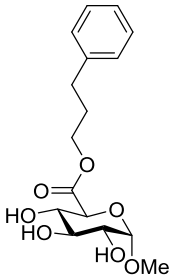


Phenylpropyl (methyl 2,3,4-tri-*O*-benzyl-α-D-glucopyranoside) uronate (122)

Compound **121** (3.00 g, 6.27 mmol) was dissolved in dry CH_2Cl_2 (50 ml) and the solution cooled to 0 °C. 3-Phenylpropan-1-ol (4.26 ml, 31.3 mmol) and DMAP (0.077 g, 0.627 mmol) were added, followed by EDC (1.44 g, 7.52 mmol). After 4 hours at room temperature the mixture was diluted with CH_2Cl_2 and washed twice with brine. The organic layer was dried over $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The residue was purified by silica gel column chromatography (heptane/ethyl acetate 7:3) to give **122** as a colorless oil (3.29 g, 88%).

^{1}H NMR (400 MHz, $CDCl_3$): δ 7.59 – 7.09 (m, 20H, Ar*H*), 5.02 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.89 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.88 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.86 (d, *J* = 12.0 Hz, 1H, OC*H*$_2$Ph), 4.71 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.67 (d, *J* = 3.5 Hz, 1H, H-1), 4.65 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.25 (d, *J* = 10.0 Hz, 1H, H-5), 4.18 (tdd, *J* = 10.9, 6.7, 4.2 Hz, 2H, H-α), 4.06 (t, *J* = 9.3 Hz, 1H, H-3), 3.81 (dd, *J* = 9.9, 9.1 Hz, 1H, H-4), 3.64 (dd, *J* = 9.6, 3.5 Hz, 1H, H-2), 3.47 (s, 3H, OC*H*$_3$), 2.73 – 2.63 (m, 2H, H-γ), 1.98 (tt, *J* = 13.4, 6.7 Hz, 2H, H-β). **^{13}C NMR** (100 MHz, $CDCl_3$): δ 169.8 (*C*OOR), 140.9, 138.6, 138.0, 138.0, 128.6, 128.5, 128.5, 128.4, 128.3, 128.1, 128.0, 127.8, 127.7, 126.1, 98.8 (C-1), 81.5,

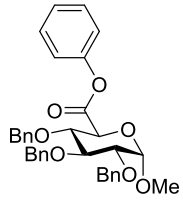
79.7, 79.4, 75.9 (O*C*H_2Ph), 75.2 (O*C*H_2Ph), 73.7 (O*C*H_2Ph), 70.3, 65.0, 55.7 (O*C*H_3), 32.0, 30.0.



Phenylpropyl (methyl α-D-glucopyranoside) uronate (111)

Compound **122** (3.29 g, 5.52 mmol) was dissolved in dry THF (35 ml) and an excess of Pd 10% on activated charcoal was added (220 mg). The suspension was degassed and backfilled with H_2 three times then stirred under a hydrogen atmosphere (1 atm) for 24 hours at room temperature. The resulting mixture was filtered through a Celite pad and rinsed with two volumes of THF (15 ml). The filtrate was concentrated *in vacuo* and purified on a silica gel column (CH_2Cl_2/CH_3OH 95:5) giving **111** as a colorless oil (1.50 g, 83%).

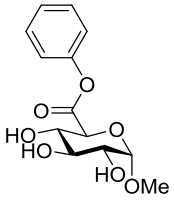
^{1}H NMR (400 MHz, $CDCl_3$): δ 7.22 – 7.02 (m, 5H, Ar*H*), 4.73 (d, *J* = 3.6 Hz, 1H, H-1), 4.12 (t, *J* = 6.6 Hz, 2H, H-α), 4.03 (d, *J* = 9.7 Hz, 1H, H-5), 3.73 (t, *J* = 9.2 Hz, 1H, H-3), 3.65 (t, *J* = 9.4 Hz, 1H, H-4), 3.54 (dd, *J* = 9.5, 3.7 Hz, 1H, H-2), 3.36 (s, 3H, OC*H*$_3$), 2.60 (dd, *J* = 9.1, 6.2 Hz, 2H, H-γ), 1.92 (dq, *J* = 9.2, 6.7 Hz, 2H, H-β). **^{13}C NMR** (100 MHz, $CDCl_3$): δ 170.2 (*C*OOR), 141.1, 128.5, 128.4, 126.1, 100.0 (C-1), 73.5, 71.6, 71.5, 70.8, 65.1, 55.7 (O*C*H_3), 31.9, 30.0. **HRMS** (ESIMS) m/z calcd for $C_{16}H_{22}NaO_7$ (M+Na^+) 349.3205, found 349.1264.



Phenyl (methyl 2,3,4-tri-*O*-benzyl-α-D-glucopyranoside) uronate (123)

Compound **121** (3.00 g, 6.27 mmol) was dissolved in dry CH_2Cl_2 (50 ml) and the solution cooled to 0 °C. Phenol (2.95 g, 31.3 mmol) and DMAP (0.077 g, 0.627 mmol) were added, followed by EDC (1.44 g, 7.52 mmol). The reaction was stirred at room temperature for 3 hours after which the starting material had been completely consumed. The mixture was diluted with CH_2Cl_2 and washed twice with brine. The organic layer was dried over $MgSO_4$, filtered and the solvent evaporated under reduced pressure. The residue was purified by silica gel column chromatography (toluene/acetone 9:1) to give **123** as a colorless oil (2.61 g, 75%).

^{1}H NMR (300 MHz, $CDCl_3$): δ 7.39 – 6.91 (m, 20H, Ar*H*), 4.93 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.82 (d, *J* = 10.7 Hz, 1H, OC*H*$_2$Ph), 4.78 (d, *J* = 10.9 Hz, 1H, OC*H*$_2$Ph), 4.76 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.61 (d, *J* = 10.7 Hz, 1H, OC*H*$_2$Ph), 4.60 (d, *J* = 12.1 Hz, 1H, OC*H*$_2$Ph), 4.60 (d, *J* = 3.3 Hz, 1H, H-1), 4.34 (d, *J* = 9.9 Hz, 1H, H-5), 4.00 (t, *J* = 9.3 Hz, 1H, H-3), 3.81 (dd, *J* = 9.1, 9.9 Hz, 1H, H-4), 3.56 (dd, *J* = 9.6, 3.5 Hz, 1H, H-2), 3.40 (s, 3H, OC*H*$_3$). **^{13}C NMR** (75 MHz, $CDCl_3$): δ 167.2 (*C*OOPh), 149.3, 137.5, 136.9, 136.8, 128.5, 127.5, 127.4, 127.4, 127.2, 127.0, 126.9, 125.2, 120.2, 97.9 (C-1), 80.4, 78.6, 78.3, 74.9 (O*C*H$_2$Ph), 74.3 (O*C*H$_2$Ph), 72.7 (O*C*H$_2$Ph), 69.4, 54.8 (O*C*H$_3$).



Phenyl (methyl α-D-glucopyranoside) uronate (112)

Compound **123** (2.61 g, 4.70 mmol) was dissolved in dry THF (30 ml) and an excess of Pd 10% on activated charcoal was added (0.200 g). The suspension was degassed and backfilled with H_2 three times then stirred under hydrogen atmosphere (1 atm) for 24 hours at room temperature. The resulting mixture was filtered through a Celite pad and rinsed with two volumes of THF (15 ml). The filtrate was concentrated under reduced pressure to give **112** as a white solid (1.26 g, 94%).

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.45 – 6.98 (m, 5H, Ar*H*), 4.71 (d, *J* = 3.7 Hz, 1H, H-1), 4.28 (d, *J* = 9.8 Hz, 1H, H-5), 3.82 (t, *J* = 9.2 Hz, 1H, H-3), 3.68 (dd, *J* = 9.0, 9.8 Hz, 1H, H-4), 3.62 (dd, *J* = 9.5, 3.7 Hz, 1H, H-2), 3.41 (s, 3H, OCH_3). **^{13}C NMR** (100 MHz, $CDCl_3$): δ 168.7 (*C*OOPh), 150.3, 129.5, 126.3, 121.4, 100.0 (C-1), 73.5, 71.7, 71.5, 71.0, 55.9 (O*C*H_3). **HRMS** (ESIMS) m/z calcd for $C_{13}H_{16}NaO_7$ (M+Na^+) 307.2428, found 307.0774.

Enzymatic methods

Semi-quantitative detection of glucuronoyl esterase activity was conducted by TLC analysis of aliquots from incubation mixtures (35 °C) containing the tested enzymes (0.025 mg/ml) and the substrates (8 mM) in 50 mM sodium phosphate buffer, pH 6. Reactions were run for 42 hours and aliquots were withdrawn for TLC analysis after 2, 18 and 42 hours. Aliquots were chromatographed on aluminium TLC plates coated with silica gel 60 (Merck) in CH_2Cl_2/CH_3OH/H_2O

(80:25:4) and the conversion of glucuronoyl esters into the corresponding alcohols and glucuronic acids was visualized by development with 1 M sulfuric acid and heating.

Kinetic parameters were determined by enzymatic hydrolysis at 30 °C in 96-well MultiScreen 10 kDa cut-off ultrafiltration plates (Millipore) for 10 minutes in 50 mM sodium phosphate buffer, pH 6. Substrate concentrations varied from 0.025 to 150 mM, while enzyme concentrations were in the range of 0.001 to 0.1 mg/ml (0.02 to 2 µM) depending on the compound studied. After incubation, reactions were stopped by rapid cooling to 4 °C followed by mechanical removal of the enzyme from the solution by ultracentrifugation in a pre-cooled centrifuge for 20 minutes at 4 °C. The degree of substrate hydrolysis was determined on the basis of integrated areas of the UV-absorbing alcohols produced within different time intervals as quantified by HPLC (ICS-5000 Dionex system, Thermo Fisher Scientific) using a Luna C18 3 µm column (100 Å, 150x4.6 mm, Phenomenex) and UV detection at 210 nm. Elution was carried out with a mixture of acetonitrile/0.01% formic acid solution at pH 3.6 (isocratic, 35:65 V/V) at a flow rate of 0.7 ml/min. The obtained data were fitted to the Michaelis-Menten equation to estimate the values for K_m, V_{max} and k_{cat}.

5 BIBLIOGRAPHY

(1) Zhou, C.-H.; Xia, X.; Lin, C.-X.; Tong, D.-S.; Beltramini, J. *Chem. Soc. Rev.* **2011**, *40*, 5588–5617.

(2) Albersheim, P.; Darvill, A.; Roberts, K.; Sederoff, R.; Staehelin, A. *Plant cell walls*; Albersheim, P., Darvill, A., Roberts, K., Sederoff, R., Staehelin, A., Eds.; Taylor & Francis Group: New York Abingdon, 2011.

(3) Alvira, P.; Tomás-Pejó, E.; Ballesteros, M.; Negro, M. J. *Bioresour. Technol.* **2010**, *101*, 4851–4861.

(4) Sun, S.; Sun, S.; Cao, X.; Sun, R. *Bioresour. Technol.* **2016**, *199*, 49–58.

(5) Aita, G. A.; Salvi, D. A.; Walker, M. S. *Bioresour. Technol.* **2011**, *102*, 4444–4448.

(6) Scheller, H. V.; Ulvskov, P. *Annu. Rev. Plant Biol.* **2010**, *61*, 263–289.

(7) Sun, X.-F.; Sun, R.; Fowler, P.; Baird, M. S. *J. Agric. Food Chem.* **2005**, *53*, 860–870.

(8) Dimmel, D. In *Lignin and Lignans - Advances in Chemistry*; Hetiner, C., Dimmel, D., Schmidt, J., Eds.; CRC Press, Taylor & Francis Group: Boca Raton, FL, 2010; pp 1–10.

(9) Du, X.; Pérez-Boada, M.; Fernández, C.; Rencoret, J.; del Río, J. C.; Jiménez-Barbero, J.; Li, J.; Gutiérrez, A.; Martínez, A. T. *Planta* **2014**, *239*, 1079–1090.

(10) Balakshin, M.; Capanema, E. A.; Berlin, A. *Isolation and analysis of lignin-carbohydrate complexes preparations with traditional and advanced*

methods: A review, 1st ed.; Elsevier B.V., 2014; Vol. 42.

(11) Balakshin, M. Y.; Capanema, E. A.; Chang, H. M. *Holzforschung* **2007**, *61*, 1–7.

(12) Ralph, J.; Hatfield, R. D.; Quideau, S.; Helm, R. F.; Grabber, J. H.; Jung, H. G.; West, L. D. *J. Am. Chem. Soc.* **1994**, *116*, 9448–9456.

(13) Balakshin, M.; Capanema, E. A.; Gracz, H.; Chang, H. M.; Jameel, H. *Planta* **2011**, *233*, 1097–1110.

(14) Yuan, T. Q.; Sun, S. N.; Xu, F.; Sun, R. C. *J. Agric. Food Chem.* **2011**, *59*, 10604–10614.

(15) Bunzel, M. *Phytochem. Rev.* **2010**, *9*, 47–64.

(16) Imamura, T.; Watanabe, T.; Kuwahara, M.; Koshijima, T. *Phytochemistry* **1994**, *37*, 1165–1173.

(17) Li, K.; Helm, R. F. *J. Agric. Food Chem.* **1995**, *43*, 2098–2103.

(18) Ebringerová, A.; Thomas, H. *Adv. Polym. Sci.* **2005**, *186*, 1–67.

(19) Monclaro, A. V.; Ferreira Filho, E. X. In *Fungal Enzymes*; Polizeli, M. de L. T. M., Rai, M., Eds.; CRC Press, Taylor & Francis Group: Boca Raton, FL, 2014; pp 280–293.

(20) Mendis, M.; Simsek, S. *Food Hydrocoll.* **2014**, *42*, 239–243.

(21) Whistler, R. L.; Chen, C.-C. In *Wood structure and composition*; Lewin, M., Goldstein, L. S., Eds.; Marcel Dekker Inc.: New York Basel Hong Kong, 1991; pp 287–319.

(22) Johnson, S.; Samuelson, O. *Anal. Chim. Acta* **1966**, *36*, 1–11.

(23) Chakar, F. S.; Allison, L.; Ragauskas, A. J.; McDonough, T. J. *Tappi J.*

2000, *83*, 1–11.

(24) de Siqueira, F. G.; Ferreira Filho, E. X. *Mini. Rev. Org. Chem.* **2010**, *7*, 54–60.

(25) Jordan, D. B.; Bowman, M. J.; Braker, J. D.; Dien, B. S.; Hector, R. E.; Lee, C. C.; Mertens, J. A.; Wagschal, K. *Biochem. J.* **2012**, *442*, 241–252.

(26) Jeffries, T. W. *Biodegradation* **1990**, *1*, 163–176.

(27) Collins, T.; Gerday, C.; Feller, G. *FEMS Microbiol. Rev.* **2005**, *29*, 3–23.

(28) de Wet, B. J. M.; van Zyl, W. H.; Prior, B. A. *Enzyme Microb. Technol.* **2006**, *38*, 649–656.

(29) Nagy, T.; Emami, K.; Fontes, C. M. G. A.; Ferreira, L. M. A.; Humphry, D. R.; Gilbert, H. J. *J. Bacteriol.* **2002**, *184*, 4925–4929.

(30) Correia, M. A. S.; Mazumder, K.; Brás, J. L. A.; Firbank, S. J.; Zhu, Y.; Lewis, R. J.; York, W. S.; Fontes, C. M. G. A.; Gilbert, H. J. *J. Biol. Chem.* **2011**, *286*, 22510–22520.

(31) Nishitani, K.; Nevins, D. J. *J. Biol. Chem.* **1991**, *266*, 6539–6543.

(32) Padilha, I. Q. M.; Valenzuela, S. V.; Grisi, T. C. S. L.; Diaz, P.; de Araújo, D. A. M.; Pastor, F. I. J. *Int. Microbiol.* **2014**, *17*, 175–184.

(33) Biely, P.; Puchart, V.; Stringer, M. A.; Krogh, K. B. R. M. *FEBS J.* **2014**, *281*, 3894–3903.

(34) Puls, J.; Schmidt, O.; Granzow, C. *Enzyme Microb. Technol.* **1987**, *9*, 83–88.

(35) Choi, I.-D.; Kim, H.-Y.; Choi, Y.-J. *Biosci. Biotechnol. Biochem.* **2000**, *64*, 2530–2537.

(36) Uchida, H.; Nanri, T.; Kawabata, Y.; Kusakabe, I.; Murakami, K. *Biosci.*

Biotechnol. Biochem. **1992**, *56*, 1608–1615.

(37) Tenkanen, M.; Siika-Aho, M. *J. Biotechnol.* **2000**, *78*, 149–161.

(38) Chong, S.; Battaglia, E.; Coutinho, P. M.; Henrissat, B.; Tenkanen, M.; de Vries, R. P. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 1323–1332.

(39) Ryabova, O.; Vršanská, M.; Kaneko, S.; van Zyl, W. H.; Biely, P. *FEBS Lett.* **2009**, *583*, 1457–1462.

(40) Kolenová, K.; Ryabova, O.; Vršanská, M.; Biely, P. *FEBS Lett.* **2010**, *584*, 4063–4068.

(41) Špániková, S.; Biely, P. *FEBS Lett.* **2006**, *580*, 4597–4601.

(42) Li, X. L.; Špániková, S.; de Vries, R. P.; Biely, P. *FEBS Lett.* **2007**, *581*, 4029–4035.

(43) Topakas, E.; Moukouli, M.; Dimarogona, M.; Vafiadi, C.; Christakopoulos, P. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 1765–1772.

(44) Vafiadi, C.; Topakas, E.; Biely, P.; Christakopoulos, P. *FEMS Microbiol. Lett.* **2009**, *296*, 178–184.

(45) Duranová, M.; Hirsch, J.; Kolenová, K.; Biely, P. *Biosci. Biotechnol. Biochem.* **2009**, *73*, 2483–2487.

(46) Katsimpouras, C.; Bénarouche, A.; Navarro, D.; Karpusas, M.; Dimarogona, M.; Berrin, J. G.; Christakopoulos, P.; Topakas, E. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 5507–5516.

(47) d'Errico, C.; Jørgensen, J. O.; Krogh, K. B. R. M.; Spodsberg, N.; Madsen, R.; Monrad, R. N. *Biotechnol. Bioeng.* **2015**, *112*, 914–922.

(48) Špániková, S.; Poláková, M.; Joniak, D.; Hirsch, J.; Biely, P. *Arch.*

Microbiol. **2007**, *188*, 185–189.

(49) Biely, P.; Malovíková, A.; Uhliariková, I.; Li, X.; Wong, D. W. S. *FEBS Lett.* **2015**, *589*, 2334–2339.

(50) d'Errico, C.; Börjesson, J.; Ding, H.; Krogh, K. B. R. M.; Spodsberg, N.; Madsen, R.; Monrad, R. N. *J. Biotechnol.* **2016**, *219*, 117–123.

(51) Hirsch, J.; Koós, M.; Kovác, P. *Carbohydr. Res.* **1998**, *310*, 145–149.

(52) Hirsch, J.; Langer, V.; Koos, M. *Molecules* **2005**, *10*, 251–258.

(53) Poláková, M.; Joniak, D.; Ďuriš, M. *Collect. Czech. Chem. Commun.* **2000**, *65*, 1609–1618.

(54) Madsen, R. In *Glycoscience: Chemistry and Chemical Biology*; Fraser-Reid, B. O., Tatsuta, K., Thiem, J., Eds.; Springer-Verlag: Berlin Heidelberg New York, 2008; pp 179–225.

(55) Allanson, N. M.; Liu, D.; Chi, F.; Jain, R. K.; Chen, A.; Ghosh, M.; Hong, L.; Sofia, M. J. *Tetrahedron Lett.* **1998**, *39*, 1889–1892.

(56) Vogel, C.; Gries, P. *J. Carbohydr. Chem.* **1994**, *13*, 37–46.

(57) Garegg, P. J.; Olsson, L.; Oscarson, S. *J. Org. Chem.* **1995**, *60*, 2200–2204.

(58) Montanari, F.; Quici, S.; Henry-Riyad, H.; Tidwell, T. T.; Studer, A.; Vogler, T. *e-EROS Encyclopedia of Reagents for Organic Synthesis*; 2007.

(59) de Nooy, A. E. J.; Besemer, A. C.; van Bekkum, H. *Tetrahedron* **1995**, *51*, 8023–8032.

(60) Anelli, P. L.; Biffi, C.; Montanari, F.; Quici, S. *J. Org. Chem.* **1987**, *52*, 2559–2562.

(61) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1993**, *34*, 1181–1184.

(62) Garegg, P. J.; Oscarson, S.; Tedebark, U. *J. Carbohydr. Chem.* **1998**, *17*, 587–594.

(63) Haller, M.; Boons, G.-J. *J. Chem. Soc. Perkin Trans. 1* **2001**, 814–822.

(64) de Nooy, A. E. J.; Besemer, A. C.; van Bekkum, H. *Carbohydr. Res.* **1995**, *269*, 89–98.

(65) Chauvin, A.-L.; Nepogodiev, S. A.; Field, R. A. *J. Org. Chem.* **2005**, *70*, 960–966.

(66) Zhao, M.; Li, J.; Mano, E.; Song, Z.; Tschaen, D. M.; Grabowski, E. J. J.; Reider, P. J. *J. Org. Chem.* **1999**, *64*, 2564–2566.

(67) De Mico, A.; Margarita, R.; Parlanti, L.; Vescovi, A.; Piancatelli, G. *J. Org. Chem.* **1997**, *62*, 6974–6977.

(68) Epp, J. B.; Widlanski, T. S. *J. Org. Chem.* **1999**, *64*, 293–295.

(69) van den Bos, L. J.; Codee, J. D. C.; van der Toorn, J. C.; Boltje, T. J.; van Boom, J. H.; Overkleeft, H. S.; van der Marel, G. A. *Org. Lett.* **2004**, *6*, 2165–2168.

(70) Golubev, V. A.; Sen, V. D.; Kulyk, I. V.; Aleksandrov, A. L. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **1975**, *24*, 2119–2126.

(71) Bera, S.; Linhardt, R. J. *J. Org. Chem.* **2011**, *76*, 3181–3193.

(72) Christina, A. E.; Muns, J. A.; Olivier, J. Q. A.; Visser, L.; Hagen, B.; van den Bos, L. J.; Overkleeft, H. S.; Codée, J. D. C.; van der Marel, G. A. *Eur. J. Org. Chem.* **2012**, 5729–5737.

(73) Haines, A. H. *Adv. Carbohydr. Chem. Biochem.* **1976**, *33*, 11–109.

(74) Czernecki, S.; Georgoulis, C.; Provelenghiou, C. *Tetrahedron Lett.* **1976**, *17*,

3535–3536.

(75) Van Hijfte, L.; Little, R. D. *J. Org. Chem.* **1985**, *50*, 3940–3942.

(76) Hatakeyama, S.; Mori, H.; Kitano, K.; Yamada, H.; Nishizawa, M. *Tetrahedron Lett.* **1994**, *35*, 4367–4370.

(77) Weissman, S. A.; Zewge, D. *Tetrahedron* **2005**, *61*, 7833–7863.

(78) Wuts, P. G. M. In *Greene's Protective Groups in Organic Synthesis*; Wuts, P. G. M., Ed.; John Wiley & Sons, Inc., 2014; pp 17–471.

(79) Ek, M.; Garegg, P. J.; Hultberg, H.; Oscarson, S. *J. Carbohydr. Chem.* **1983**, *2*, 305–311.

(80) Garegg, P. J.; Hultberg, H.; Wallin, S. *Carbohydr. Res.* **1982**, *108*, 97–101.

(81) Grindley, T. B. *Adv. Carbohydr. Chem. Biochem.* **1998**, *53*, 17–142.

(82) DeNinno, M.; Etienne, J.; Duplantier, K. *Tetrahedron Lett.* **1995**, *36*, 669–672.

(83) Oikawa, Y.; Yoshioka, T.; Yonemitsu, O. *Tetrahedron Lett.* **1982**, *23*, 885–888.

(84) Wuts, P. G. M.; Greene, T. W. *Greene's Protective Groups in Organic Synthesis*. 2007, pp 16–366.

(85) Gómez, A. M. In *Glycoscience: Chemistry and Chemical Biology*; Fraser-Reid, B. O., Tatsuta, K., Thiem, J., Eds.; Springer-Verlag: Berlin Heidelberg New York, 2008; pp 105–177.

(86) van Boom, J. H.; Burgers, P. M. J. *Tetrahedron Lett.* **1976**, *52*, 4875–4878.

(87) Chery, F.; Cronin, L.; O'Brien, J. L.; Murphy, P. V. *Tetrahedron* **2004**, *60*, 6597–6608.

(88) Adamo, R.; Kováč, P. *Eur. J. Org. Chem.* **2006**, 2803–2809.

(89) Rej, R. N.; Glushka, J. N.; Chew, W.; Perlin, A. S. *Carbohydr. Res.* **1989**, *189*, 135–148.

(90) Wong, C. H.; Ye, X. S.; Zhang, Z. *J. Am. Chem. Soc.* **1998**, *120*, 7137–7138.

(91) Zhu, T.; Boons, G.-J. *J. Am. Chem. Soc.* **2000**, *122*, 10222–10223.

(92) Yang, B.; Yoshida, K.; Yin, Z.; Dai, H.; Kavunja, H.; El-Dakdouki, M. H.; Sungsuwan, S.; Dulaney, S. B.; Huang, X. *Angew. Chem. Int. Ed.* **2012**, *51*, 10185–10189.

(93) Bohé, L.; Crich, D. *Carbohydr. Res.* **2015**, *403*, 48–59.

(94) Martin, A.; Arda, A.; Désiré, J.; Martin-Mingot, A.; Probst, N.; Sinaÿ, P.; Jiménez-Barbero, J.; Thibaudeau, S.; Blériot, Y. *Nat. Chem.* **2016**, *8*, 186–191.

(95) Demchenko, A. V. In *Handbook of Chemical Glycosylation: Advances in Stereoselectivity and Therapeutic Relevance*; Demchenko, A. V., Ed.; Wiley-VCH: Verlag, 2008; pp 1–27.

(96) Koenigs, W.; Knorr, E. *Chem. Ber.* **1901**, *34*, 957–981.

(97) Igarashi, K. *Adv. Carbohydr. Chem. Biochem.* **1977**, *34*, 243–283.

(98) Lemieux, R. U.; Hendriks, K. B.; Stick, R. V; James, K. *J. Am. Chem. Soc.* **1975**, *97*, 4056–4062.

(99) Zhu, X.; Schmidt, R. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 1900–1934.

(100) Oscarson, S. In *Glycoscience: Chemistry and Chemical Biology*; Fraser-Reid, B. O., Tatsuta, K., Thiem, J., Eds.; Springer: Berlin, 2008; pp 661–697.

(101) Mukaiyama, T. *Angew. Chem. Int. Ed.* **2004**, *43*, 5590–5614.

(102) Mukaiyama, T.; Murai, Y.; Shoda, S. *Chem. Lett.* **1981**, *10*, 431–432.

(103) Gervay, J.; Hadd, M. J. *J. Org. Chem.* **1997**, *62*, 6961–6967.

(104) Pougny, J. R.; Jacquinet, J. C.; Nassr, M.; Duchet, D.; Milat, M. L.; Sinay, P. *J. Am. Chem. Soc.* **1977**, *99*, 6762–6763.

(105) Schmidt, R. R.; Michel, J. *Angew. Chem. Int. Ed.* **1980**, *19*, 731–732.

(106) Schmidt, R. R.; Zhu, X. In *Glycoscience: Chemistry and Chemical Biology*; Fraser-Reid, B. O., Tatsuta, K., Thiem, J., Eds.; Springer: Berlin, 2008; pp 451–524.

(107) Schmidt, R. R. *Angew. Chem. Int. Ed.* **1986**, *25*, 212–235.

(108) Schmidt, R. R.; Castro-Palomino, J. C.; Retz, O. *Pure Appl. Chem.* **1999**, *71*, 729–744.

(109) Abdel-Rahman, A. A. H.; Jonke, S.; El Ashry, E. S. H.; Schmidt, R. R. *Angew. Chem. Int. Ed.* **2002**, *41*, 2972–2974.

(110) Ferrier, R. J.; Hay, R. W.; Vethaviyasar, N. *Carbohydr. Res.* **1973**, *27*, 55–61.

(111) Zhong, W.; Boons, G.-J. In *Handbook of Chemical glycosylation*; Demchenko, A. V., Ed.; Wiley-VCH: Verlag, 2008; pp 261–303.

(112) Agnihotri, G.; Tiwari, P.; Misra, A. K. *Carbohydr. Res.* **2005**, *340*, 1393–1396.

(113) Shah, R. H.; Bahl, O. P. *Carbohydr. Res.* **1978**, *65*, 153–158.

(114) Tsvetkov, Y. E.; Byramova, N. F.; Backnowsky, L. V. *Carbohydr. Res.* **1983**, *115*, 254–258.

(115) Mukherjee, C.; Tiwari, P.; Misra, A. K. *Tetrahedron Lett.* **2006**, *47*, 441–445.

(116) Schneider, W.; Stiehler, O. *Chem. Ber.* **1918**, *51*, 220–234.

(117) Hanessian, S.; Guindon, Y. *Carbohydr. Res.* **1980**, *86*, C3–C6.

(118) Kartha, K. P. R.; Field, R. A. *Tetrahedron Lett.* **1997**, *38*, 8233–8236.

(119) Nicolaou, K. C.; Dolle, R. E.; Papahatjis, D. P.; Randall, J. L. *J. Am. Chem. Soc.* **1984**, *106*, 4189–4192.

(120) Wolfrom, M. L.; Groebke, W. *J. Org. Chem.* **1963**, *28*, 2986–2988.

(121) Schmidt, R. R.; Kinzy, W. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 21–123.

(122) Oshitari, T.; Shibasaki, M.; Yoshizawa, T.; Tomita, M.; Takao, K.; Kobayashi, S. *Tetrahedron* **1997**, *53*, 10993–11006.

(123) Motawia, M. S.; Marcussen, J.; Møller, B. L. *J. Carbohydr. Chem.* **1995**, *14*, 1279–1294.

(124) Gómez, A. M.; Company, M. D.; Agocs, A.; Uriel, C.; Valverde, S.; López, J. C. *Carbohydr. Res.* **2005**, *340*, 1872–1875.

(125) Uchiro, H.; Wakiyama, Y.; Mukaiyama, T. *Chem. Lett.* **1998**, *27*, 567–568.

(126) Kahne, D.; Walker, S.; Cheng, Y.; Van Engen, D. *J. Am. Chem. Soc.* **1989**, *111*, 6881–6882.

(127) Kakarla, R.; Dulina, R. G.; Hatzenbuhler, N. T.; Hui, Y. W.; Sofia, M. J. *J. Org. Chem.* **1996**, *61*, 8347–8349.

(128) Raghavan, S.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 1580–1581.

(129) Garegg, P. J.; Henrichson, C.; Norberg, T. *Carbohydr. Res.* **1983**, *116*, 162–

165.

(130) Woodward, R. B.; Logusch, E.; Nambiar, K. P.; Sakan, K.; Ward, D. E.; Au-Yeung, B.-W.; Balaram, P.; Browne, L. J.; Card, P. J.; Chen, C. H.; Chenevert, R. B.; Fliri, A.; Frobel, K.; Gais, H.-J.; Garratt, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajan Babu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsuta, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, A. T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N.-C. *J. Am. Chem. Soc.* **1981**, *103*, 3215–3217.

(131) Lönn, H. *Carbohydr. Res.* **1985**, *139*, 105–113.

(132) Fügedi, P.; Garegg, P. J. *Carbohydr. Res.* **1986**, *149*, C9–C12.

(133) Veeneman, G. H.; van Leeuwen, S. H.; van Boom, J. H. *Tetrahedron Lett.* **1990**, *31*, 1331–1334.

(134) Konradsson, P.; Udodong, U. E.; Fraser-Reid, B. O. *Tetrahedron Lett.* **1990**, *31*, 4313–4316.

(135) Takeuchi, K.; Tamura, T.; Mukaiyama, T. *Chem. Lett.* **2000**, *29*, 124–125.

(136) Lemieux, R. U.; Morgan, A. R. *Can. J. Chem.* **1965**, *43*, 2190–2197.

(137) Veeneman, G. H.; van Boom, J. H. *Tetrahedron Lett.* **1990**, *31*, 275–278.

(138) Huang, K.-T.; Winssinger, N. *Eur. J. Org. Chem.* **2007**, 1887–1890.

(139) Crich, D.; Sun, S. *Tetrahedron* **1998**, *54*, 8321–8348.

(140) Huang, X.; Huang, L.; Wang, H.; Ye, X.-S. *Angew. Chem. Int. Ed.* **2004**, *43*, 5221–5224.

(141) Crich, D.; Smith, M. *J. Am. Chem. Soc.* **2001**, *123*, 9015–9020.

(142) Codée, J. D. C.; Litjens, R. E. J. N.; Den Heeten, R.; Overkleeft, H. S.; van Boom, J. H.; van der Marel, G. A. *Org. Lett.* **2003**, *5*, 1519–1522.

(143) Codée, J. D. C.; Litjens, R. E. J. N.; van den Bos, L. J.; Overkleeft, H. S.; van der Marel, G. A. *Chem. Soc. Rev.* **2005**, *34*, 769–782.

(144) Fraser-Reid, B.; Udodong, U. E.; Wu, Z.; Ottosson, H.; Merritt, R. J.; Rao, C. S.; Roberts, C.; Madsen, R. *Synlett* **1992**, 927–942.

(145) Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraser-Reid, B. *J. Am. Chem. Soc.* **1988**, *110*, 5583–5584.

(146) Fraser-Reid, B.; Wu, Z.; Udodong, U. E.; Ottosson, H. *J. Org. Chem.* **1990**, *55*, 6068–6070.

(147) Zhang, Z.; Ollmann, I. R.; Ye, X. S.; Wischnat, R.; Baasov, T.; Wong, C. H. *J. Am. Chem. Soc.* **1999**, *121*, 734–753.

(148) Mong, K. K. T.; Wong, C. H. *Angew. Chem. Int. Ed.* **2002**, *41*, 4087–4090.

(149) Lahmann, M.; Oscarson, S. *Org. Lett.* **2000**, *2*, 3881–3882.

(150) Sliedregt, L. A. J. M.; Zegelaar-Jaarsveld, K.; van der Marel, G. A.; van Boom, J. H. *Synlett* **1993**, 335–337.

(151) Zhu, T.; Boons, G.-J. *Org. Lett.* **2001**, *3*, 4201–4203.

(152) Hashihayata, T.; Ikegai, K.; Takeuchi, K.; Jona, H.; Mukaiyama, T. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 1829–1848.

(153) Hashihayata, T.; Mandai, H.; Mukaiyama, T. *Chem. Lett.* **2003**, *32*, 442–443.

(154) Mukaiyama, T.; Wakiyama, Y.; Miyazaki, K.; Takeuchi, K. *Chem. Lett.*

1999, *28*, 933–934.

(155) Takeuchi, K.; Tamura, T.; Mukaiyama, T. *Chem. Lett.* **2000**, *29*, 122–123.

(156) Demchenko, A. V.; De Meo, C. *Tetrahedron Lett.* **2002**, *43*, 8819–8822.

(157) Yu, B.; Yang, Z.; Cao, H. *Curr. Org. Chem.* **2005**, *9*, 179–194.

(158) Wang, Z.; Wasonga, G.; Swarts, B. M.; Huang, X. In *Carbohydrate Chemistry: Proven Synthetic methods*; Kovac, P., Ed.; CRC Press, Taylor & Francis Group: Boca Raton, FL, 2012; pp 43–51.

(159) Crich, D. *J. Carbohydr. Chem.* **2002**, *21*, 663–686.

(160) Crich, D.; Sun, S. *J. Am. Chem. Soc.* **1997**, *119*, 11217–11223.

(161) Wang, Z.; Zhou, L.; El-Boubbou, K.; Ye, X.; Huang, X. *J. Org. Chem.* **2007**, *72*, 6409–6420.

(162) Miermont, A.; Zeng, Y.; Jing, Y.; Ye, X.; Huang, X. *J. Org. Chem.* **2007**, *72*, 8958–8961.

(163) Crich, D.; Cai, F.; Yang, F. *Carbohydr. Res.* **2008**, *343*, 1858–1862.

(164) Myhre, D. V.; Smith, F. *J. Org. Chem.* **1961**, *26*, 4609–4612.

(165) Hirsch, J.; Kováč, P. *Carbohydr. Res.* **1981**, *90*, C5–C7.

(166) Hirsch, J.; Kováč, P.; Petráková, E. *Carbohydr. Res.* **1982**, *106*, 203–216.

(167) Kováč, P.; Hirsch, J. *Carbohydr. Res.* **1982**, *100*, 177–193.

(168) Takeo, K.; Ohguchi, Y.; Hasegawa, R.; Kitamura, S. *Carbohydr. Res.* **1995**, *278*, 301–313.

(169) Puchart, V.; Biely, P. *J. Biotechnol.* **2008**, *137*, 34–43.

(170) Kováč, P.; Hirsch, J.; Kováčik, V. *Carbohydr. Res.* **1979**, *75*, 109–116.

(171) Kováč, P.; Hirsch, J.; Kovacik, V.; Kocis, P. *Carbohydr. Res.* **1980**, *85*, 41–49.

(172) Oscarson, S.; Svahnberg, P. *J. Chem. Soc. Perkin Trans. 1* **2001**, 873–879.

(173) Schmidt, D.; Schuhmacher, F.; Geissner, A.; Seeberger, P. H.; Pfrengle, F. *Chem. Eur. J.* **2015**, *21*, 5709–5713.

(174) Steglich, W.; Höfle, G. *Angew. Chem. Int. Ed.* **1969**, *8*, 981–981.

(175) Toyooka, N.; Nakazawa, A.; Himiyama, T.; Nemoto, H. *Heterocycles* **2003**, *59*, 75–79.

(176) Kim, J. M.; Roy, R. *Carbohydr. Res.* **1997**, *298*, 173–179.

(177) Rio, S.; Beau, J.-M.; Jacquinet, J. C. *Carbohydr. Res.* **1991**, *219*, 71–90.

(178) Vasan, M.; Rauvolfova, J.; Wolfert, M. A.; Leoff, C.; Kannenberg, E. L.; Quinn, C. P.; Carlson, R. W.; Boons, G.-J. *ChemBioChem* **2008**, *9*, 1716–1720.

(179) Hassner, A.; Strand, G.; Rubinstein, M.; Patchornik, A. *J. Am. Chem. Soc.* **1975**, *97*, 1614–1615.

(180) Jeker, N.; Tamm, C. *Helv. Chim. Acta* **1988**, *71*, 1895–1903.

(181) Jiang, L.; Chan, T.-H. *J. Org. Chem.* **1998**, *63*, 6035–6038.

(182) Evtushenko, E. V. *J. Carbohydr. Chem.* **2015**, *34*, 41–54.

(183) Helm, R. F.; Ralph, J.; Anderson, L. *J. Org. Chem.* **1991**, *56*, 7015–7021.

(184) Thorsheim, K.; Siegbahn, A.; Johnsson, R. E.; Stålbrand, H.; Manner, S.; Widmalm, G.; Ellervik, U. *Carbohydr. Res.* **2015**, *418*, 65–88.

(185) Gu, G.; An, L.; Fang, M.; Guo, Z. *Carbohydr. Res.* **2014**, *383*, 21–26.

(186) Hu, Y.-P.; Lin, S.-Y.; Huang, C.-Y.; Zulueta, M. M. L.; Liu, J.-Y.; Chang, W.; Hung, S.-C. *Nat. Chem.* **2011**, *3*, 557–563.

(187) Demizu, Y.; Kubo, Y.; Miyoshi, H.; Maki, T.; Matsumura, Y.; Moriyama, N.; Onomura, O. *Org. Lett.* **2008**, *10*, 5075–5077.

(188) Evtushenko, E. V. *J. Carbohydr. Chem.* **2010**, *29*, 369–378.

(189) Kim, S.; Chang, H.; Kim, W. J. *J. Org. Chem.* **1985**, *50*, 1751–1752.

(190) Sivakumaran, T.; Jones, J. K. N. *Can. J. Chem.* **1967**, *45*, 2493–2500.

(191) Kondo, Y. *Carbohydr. Res.* **1982**, *107*, 303–311.

(192) Garegg, P. J.; Kvarnström, I.; Niklasson, A.; Niklasson, G.; Svensson, S. C. T. *J. Carbohydr. Chem.* **1993**, *12*, 933–953.

(193) Danishefsky, S. J.; DeNinno, M. P.; Chen, S. *J. Am. Chem. Soc.* **1988**, *110*, 3929–3940.

(194) Wissner, A.; Grudzinskas, C. V. *J. Org. Chem.* **1978**, *43*, 3972–3974.

(195) Kordes, M.; Winsel, H.; de Meijere, A. *Eur. J. Org. Chem.* **2000**, 3235–3245.

(196) Newman, M. S.; Gill, N.; Darré, B. *J. Org. Chem.* **1966**, *31*, 2713–2714.

(197) Kashiwagi, G. A.; Mendoza, V. M.; de Lederkremer, R. M.; Gallo-Rodriguez, C. *Org. Biomol. Chem.* **2012**, *10*, 6322–6332.

(198) Nicolaou, K. C.; Mitchell, H. J.; Fylaktakidou, K. C.; Rodríguez, R. M.; Suzuki, H. *Chem. Eur. J.* **2000**, *6*, 3116–3148.

(199) Mulzer, J.; Schollhorn, B. *Angew. Chem. Int. Ed.* **1990**, *29*, 431–432.

(200) Ogilvie, K. K.; Entwistle, D. W. *Carbohydr. Res.* **1981**, *89*, 203–210.

(201) Jones, S. S.; Reese, C. B. *J. Chem. Soc. Perkin Trans. 1* **1979**, 2762–2764.

(202) Mori, Y.; Futamura, Y.; Horisaki, K. *Angew. Chem. Int. Ed.* **2008**, *47*, 1091–1093.

(203) Pedersen, C. M.; Nordstrøm, L. U.; Bols, M. *J. Am. Chem. Soc.* **2007**, *129*, 9222–9235.

(204) Boons, G.-J.; Zhu, T. *Synlett* **1997**, 809–811.

(205) Yamago, S.; Yamada, T.; Maruyama, T.; Yoshida, J. I. *Angew. Chem. Int. Ed.* **2004**, *43*, 2145–2148.

(206) Nguyen, H. M.; Poole, J. L.; Gin, D. Y. *Angew. Chem. Int. Ed.* **2001**, *40*, 414–417.

(207) Martichonok, V.; Whitesides, G. M. *J. Org. Chem.* **1996**, *61*, 1702–1706.

(208) Crich, D.; Cai, W. *J. Org. Chem.* **1999**, *64*, 4926–4930.

(209) Crich, D.; Chandrasekera, N. S. *Angew. Chem. Int. Ed.* **2004**, *43*, 5386–5389.

(210) Frihed, T. G.; Bols, M.; Pedersen, C. M. *Chem. Rev.* **2015**, *115*, 4963–5013.

(211) Liao, G.; Burgula, S.; Zhou, Z.; Guo, Z. *Eur. J. Org. Chem.* **2015**, 2942–2951.

(212) Nigudkar, S. S.; Demchenko, A. V. *Chem. Sci.* **2015**, *6*, 2687–2704.

(213) Eby, R.; Schuerch, C. *Carbohydr. Res.* **1974**, *34*, 79–90.

(214) Wulff, G.; Röhle, G. *Angew. Chem. Int. Ed.* **1974**, *13*, 157–170.

(215) Krog-Jensen, C.; Oscarson, S. *Carbohydr. Res.* **1998**, *308*, 287–296.

(216) Houdier, S.; Vottero, P. J. A. *Carbohydr. Res.* **1993**, *248*, 377–384.

(217) Hashimoto, N.; Aoyama, T.; Shioiri, T. *Chem. Pharm. Bull.* **1981**, *29*, 1475–1478.

(218) Ikemoto, N.; Schreiber, S. L. *J. Am. Chem. Soc.* **1992**, *114*, 2524–2536.

(219) Shih, T.; Fang, Y. *Synth. Commun.* **2007**, *37*, 3337–3349.

(220) Crich, D.; Vinogradova, O. *J. Org. Chem.* **2007**, *72*, 3581–3584.

(221) Rahim, M. A.; Matsumura, S.; Toshima, K. *Tetrahedron Lett.* **2005**, *46*, 7307–7309.

(222) Debenham, J. S.; Rodebaugh, R.; Fraser-Reid, B. *J. Org. Chem.* **1997**, *62*, 4591–4600.

(223) Rodebaugh, R.; Debenham, J. S.; Fraser-Reid, B. *Tetrahedron Lett.* **1996**, *37*, 5477–5478.

(224) Sunner, H.; Charavgi, M.; Olsson, L.; Topakas, E.; Christakopoulos, P. *Molecules* **2015**, *20*, 17807–17817.

(225) Pokkuluri, P. R.; Duke, N. E. C.; Wood, S. J.; Cotta, M. A.; Li, X. L.; Biely, P.; Schiffer, M. *Proteins* **2011**, *79*, 2588–2592.

(226) Charavgi, M. D.; Dimarogona, M.; Topakas, E.; Christakopoulos, P.; Chrysina, E. D. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 63–73.

(227) Yoneda, Y.; Kawada, T.; Rosenau, T.; Kosma, P. *Carbohydr. Res.* **2005**, *340*, 2428–2435.

(228) Li, K.; Helm, R. F. *Carbohydr. Res.* **1995**, *273*, 249–253.

(229) Bowkett, E. R.; Harding, J. R.; Maggs, J. L.; Park, B. K.; Perrie, J. A.; Stachulski, A. V. *Tetrahedron* **2007**, *63*, 7596–7605.

(230) Guan, Z.; Zhang, L. H.; Sinaÿ, P.; Zhang, Y. *J. Org. Chem.* **2012**, *77*, 8888–

8895.

(231) Phanumartwiwath, A.; Hornsby, T. W.; Jamalis, J.; Bailey, C. D.; Willis, C. L. *Org. Lett.* **2013**, *15*, 5734–5737.

(232) Crich, D.; Dai, Z. *Tetrahedron* **1999**, *55*, 1569–1580.

(233) Ray, A.; Roy, N. *Carbohydr. Res.* **1990**, *196*, 95–100.

(234) Verduyn, R.; Douwes, M.; van der Klein, P. A. M.; Mösinger, E. M.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron* **1993**, *49*, 7301–7316.

(235) Kumar, A.; Geng, Y.; Schmidt, R. R. *Eur. J. Org. Chem.* **2012**, 6846–6851.

(236) Matwiejuk, M.; Thiem, J. *Eur. J. Org. Chem.* **2011**, 5860–5878.

(237) Cheng, K.; Liu, J.; Liu, X.; Li, H.; Sun, H.; Xie, J. *Carbohydr. Res.* **2009**, *344*, 841–850.

APPENDIX – PUBLICATIONS

- "Enzymatic Degradation of Lignin-Carbohydrate Complexes (LCCs): Model studies Using a Fungal Glucuronoyl Esterase from *Cerrena unicolor*" C. d'Errico, J. O. Jørgensen, K. B. R. M. Krogh, N. Spodsberg, R. Madsen, R. N. Monrad, *Biotechnol. Bioeng.* **2015**, *112*, 914–922.

- "Improved Biomass Degradation using Fungal Glucuronoyl Esterases – Hydrolysis of Natural Corn Fiber Substrate" C. d'Errico, J. Börjesson, H. Ding, K. B. R. M. Krogh, N. Spodsberg, R. Madsen, R. N. Monrad, *J. Biotechnol.* **2016**, *219*, 117–123.

ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Enzymatic Degradation of Lignin-Carbohydrate Complexes (LCCs): Model Studies Using a Fungal Glucuronoyl Esterase from *Cerrena Unicolor*

Clotilde d'Errico,[1] Jonas O. Jørgensen,[1] Kristian B. R. M. Krogh,[2] Nikolaj Spodsberg,[2] Robert Madsen,[1] Rune Nygaard Monrad[2]

[1]Department of Chemistry, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

[2]Novozymes A/S, Krogshøjvej, 36, 2880 Bagsværd, Denmark; telephone: (+45) 44460000; e-mail: rnmo@novozymes.com

ABSTRACT: Lignin-carbohydrate complexes (LCCs) are believed to influence the recalcitrance of lignocellulosic plant material preventing optimal utilization of biomass in e.g. forestry, feed and biofuel applications. The recently emerged carbohydrate esterase (CE) 15 family of glucuronoyl esterases (GEs) has been proposed to degrade ester LCC bonds between glucuronic acids in xylans and lignin alcohols thereby potentially improving delignification of lignocellulosic biomass when applied in conjunction with other cellulases, hemicellulases and oxidoreductases. Herein, we report the synthesis of four new GE model substrates comprising α- and -arylalkyl esters representative of the lignin part of naturally occurring ester LCCs as well as the cloning and purification of a novel GE from *Cerrena unicolor* (*Cu*GE). Together with a known GE from *Schizophyllum commune* (*Sc*GE), *Cu*GE was biochemically characterized by means of Michaelis–Menten kinetics with respect to substrate specificity using the synthesized compounds. For both enzymes, a strong preference for 4-*O*-methyl glucuronoyl esters rather than unsubstituted glucuronoyl esters was observed. Moreover, we found that α-arylalkyl esters of methyl α-D-glucuronic acid are more easily cleaved by GEs than their corresponding -arylalkyl esters. Furthermore, our results suggest a preference of *Cu*GE for glucuronoyl esters of bulky alcohols supporting the suggested biological action of GEs on LCCs. The synthesis of relevant GE model substrates presented here may provide a valuable tool for the screening, selection and development of industrially relevant GEs for delignification of biomass.

Biotechnol. Bioeng. 2015;112: 914–922.

KEYWORDS: enzymatic delignification; lignin-carbohydrate complexes; glucuronoyl esterase; *Cerrena unicolor*; *Schizophyllum commune*; substrate specificity

Correspondence to: R. N. Monrad

Contract grant sponsor: Danish Council for Strategic Research

Received 10 October 2014; Revision received 14 November 2014; Accepted 17 November 2014

Accepted manuscript online 25 November 2014;

Article first published online 7 January 2015 in Wiley Online Library (http://onlinelibrary.wiley.com/doi/10.1002/bit.25508/abstract).

DOI 10.1002/bit.25508

Introduction

In lignocellulosic plant material, lignin is known to be intimately associated with hemicellulose as covalently linked macromolecular structures known as lignin-carbohydrate complexes (LCCs). The main LCCs present in wood are believed to be esters, benzyl ethers and phenyl glycosides (Balakshin et al., 2011; Watanabe, 1995), whereas in grasses ester linkages between arabinosyl residues in xylan and *p*-coumaric and ferulic acids are abundant (Bunzel M, 2010). The formation of LCCs is believed to take place constantly during lignin biosynthesis in growing plants (Watanabe, 1995). During lignification, random oxidative coupling of phenoxy radicals of monolignols (coniferyl, coumaryl and sinapyl alcohols) generates unstable quinone methide intermediates which are prone to nucleophilic attack of water, alcohols and carboxylates. In the case of sugar alcohols and uronic acids, nucleophilic attack on quinone methide intermediates leads to benzyl ether and benzyl ester LCCs, respectively, thereby cross-linking lignin and hemicellulose. Based on the quinone methide pathway it is commonly accepted that benzyl ether and benzyl ester LCCs represent initial LCC structures formed during lignin biosynthesis (Watanabe, 1995), however, although benzyl ether LCCs have indeed been observed directly (Balakshin and Capanema, 2003), only indirect evidence for the benzyl ester (α-ester) LCCs (Fig. 1, structure A) have so far been reported (2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) oxidation of ester LCCs) (Imamura et al., 1994). Instead, -ester LCCs (Fig. 1, structure B) were recently identified directly by advanced 2D NMR spectroscopy in soft- and hardwoods as major LCC components (Balakshin and Capanema, 2007; Balakshin et al., 2011; Yuan et al., 2011). The apparent occurrence of -ester LCCs instead of the commonly believed benzyl esters may be explained by migration of the uronosyl group from the α to the position once formed during lignification (or during sample preparation) as indicated by NMR analysis of complex α- and -ester LCC model compounds (Li and Helm, 1995b).

In Kraft pulping of wood (e.g. paper making) ester LCCs are easily cleaved under the alkaline conditions currently used, however, in other applications where alkaline (pre)-treatment of biomass is not possible or economically feasible (e.g. animal feed or

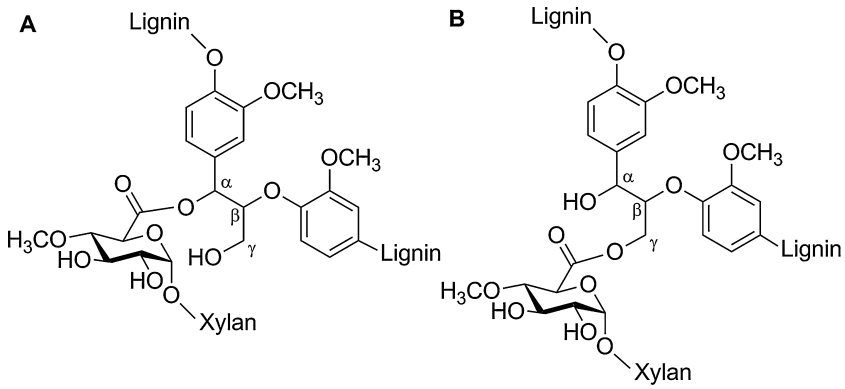


Figure 1. Representative structures of α- and -linked ester LCCs (**A** and **B**, respectively) connecting lignin alcohols and glucuronic acid residues in xylans.

biofuel production) development of efficient sustainable (enzyme) technologies to separate lignin and hemicellulose is of key importance. Glucuronoyl esterases (GEs) are a recently discovered class of carbohydrate esterases (CEs) which have been proposed to be able to degrade ester LCCs between glucuronic acids in xylans and lignin alcohols. GEs were first described in 2006 by (Špániková and Biely, 2006) and belong to the CE15 family in the continuously updated CAZy database (www.cazy.org). Currently, seven GEs have been purified and biochemically characterized (Duranová et al., 2009; Katsimpouras et al., 2014; Li et al., 2007; Špániková and Biely, 2006; Topakas et al., 2010; Vafiadi et al., 2009).

Herein, we report the synthesis of four GE substrates mimicking the natural ester LCCs in lignocellulosic plant material and the cloning and biochemical characterization of a novel fungal GE from *Cerrena unicolor*.

Materials and Methods

Chemicals and Enzymes

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Commercial glucuronate esters (methyl D-glucuronate, allyl D-glucuronate, benzyl D-glucuronate) were purchased from Carbo-Synth (Compton, UK) as anomeric mixtures. CALB (immobilized *Candida antarctica* lipase B, NZ435) is a product of Novozymes A/S. The solvents were freshly dried by Puresolv equipment.

Molecular Cloning of CuGE

The DNA sequence (accession number: GenBank: KM875459) encoding the *Cu*GE (residues 1–474) was amplified from the *Cerrena unicolor* MS01356 cDNA library by PCR using the primers CE-F: 5'-TAAGAATTCCAAAATGTTCAAGCCATCTTTCGT-3' and CE-R: 5'-TATGCGGCCGCCTCAATCAGGTCAAAGTGGGAGT-3' (see Supplementary material). The amplification reaction was composed of 1 μl of *Cerrena unicolor* MS01356 cDNA, 12.5 μl of 2 × Reddymix PCR Buffer, 1 μl of 5 μM primer CE-F, 1 μl of 5 μM primer CE-R, and 9.5 μl of H_2O. The amplification reaction was incubated in a PTC-200 DNA Engine Thermal Cycler programmed for one cycle at 94 °C for 2 min; and 35 cycles each at 94 °C for 15 s and 60 °C for 1.5 min. A 1.4 kb PCR reaction product was isolated by 1% agarose gel electrophoresis using TAE buffer and staining with SYBR Safe DNA gel stain. The DNA band was visualized with the aid of an Eagle Eye Imaging System and a Darkreader Transilluminator. The 1.4 kb DNA band was excised from the gel and purified using a GFX PCR DNA and Gel Band Purification Kit according to the manufacturer's instructions. The 1.4 kb fragment was cleaved with *Eco* RI and *Not* I and purified using a GFX PCR DNA and Gel Band Purification Kit according to the manufacturer's instructions. The cleaved 1.4 kb fragment was then directionally cloned by ligation into *Eco* RI-*Not* I cleaved pXYG1051 (Patent WO2005080559) using T4 ligase (Promega) according to the manufacturer's instructions. The ligation mixture was transformed into *E. coli* TOP10F competent cells (Invitrogen) according to the manufacturer's instructions. The transformation mixture was plated onto LB plates supplemented with 100 μg of ampicillin per ml. Plasmid minipreps were prepared from several transformants and sequenced. One plasmid with the correct *Cerrena unicolor* CE15 gGE coding sequence was chosen.

Expression of Recombinant Protein

The *Aspergillus oryzae* strain BECh2 (Patent WO200039322) was transformed with pXYG1051-cuCE15 using standard techniques (Christensen et al., 1988). To identify transformants producing the recombinant GE, the transformants and BECh2 were cultured in 10 ml of YP + 2% glucose medium at 30 °C and 200 RPM. Samples were taken after 4 days growth and resolved with SDS PAGE to identify recombinant GE production. A novel band of about 50 kDa was observed in cultures of transformants that was not observed in cultures of the untransformed BECh2. Several transformants that appeared to express the recombinant GE at high levels were further cultured in 100 ml of YP + 2% glucose medium in 500 ml shake flasks at 30 °C and 200 RPM. Samples were taken after 2, 3, and

4 days growth and expression levels compared by resolving the samples with SDS PAGE. A single transformant that expressed the recombinant GE at relatively high levels was selected and isolated twice by dilution streaking conidia on selective medium containing 0.01% Triton X-100 to limit colony size and fermented in YP + 2% glucose medium in shake flasks as described above to provide material for purification. The shake flask cultures were harvested after 4 days growth and fungal mycelia were removed by filtering the cultivation broth through Miracloth (Calbiochem).

Purification of CuGE

Sterile filtered cultivation broth was concentrated and buffer exchanged to buffer A (25 mM acetate, pH 4.5) using a Sartorius crossflow system equipped with a polyethersulfone 50 kDA cut-off Sartocon Slice membrane (Sartorius). The concentrate was applied onto a cation exchange SP Sepharose Fast Flow column XK 26/20 (GE Healthcare). The column volume (CV) was 20 ml. The column was equilibrated in buffer A. Unbound protein was washed off with 5 CVs of buffer A. The column was eluted with a linear gradient of buffer B (1.0 M NaCl in buffer A) over 5 CVs. Fractions were analyzed by SDS PAGE. The enzyme was recovered in the eluate. The eluate was buffer exchanged to buffer C (25 mM TRIS, pH 8.5) and applied onto an anion exchange Q Sepharose Fast Flow column XK 26/20 (GE Healthcare). The CV was 50 ml. The column was equilibrated in buffer C. Unbound protein was washed off with 5 CVs of buffer C. The column was eluted with a linear gradient of buffer D (1.0 M NaCl in buffer C) over 5 CVs. Fractions were analyzed by SDS PAGE, and the enzyme was recovered in the eluate fraction (Fig. 2).

Cloning, Expression and Purification of ScGE

*Sc*GE (Špániková and Biely, 2006) (swissprot:D8QLP9) was recombinantly expressed at Novozymes in the host *Aspergillus oryzae* and subsequently purified to homogeneity using standard techniques.

Deglycosylation Using Endoglycosidase H

The enzyme was diluted to 1 mg/ml in buffer (50 mM MES, pH 6.0). Endoglycosidase H (5 U/ml, Roche Diagnostics) was added on a volume basis to 50 mU/ml. The reaction mixture was incubated for 1 h at room temperature and overnight at 4 °C.

Intact Molecular Weight Analyses

Intact molecular weight analyses were performed using a Bruker microTOF focus electrospray mass spectrometer. The samples were diluted to 1 mg/ml in MQ water. The diluted samples were online washed on a Waters MassPREP On-Line Desalting column (2.1 × 10 mm) and introduced to the electrospray source with a flow of 200 μl/h by an Agilent LC system. Data analysis was performed with DataAnalysis version 3.4 (Bruker Daltonik). The molecular weight of the samples was calculated by deconvolution of the raw data in the range 30 to 70 kDa.

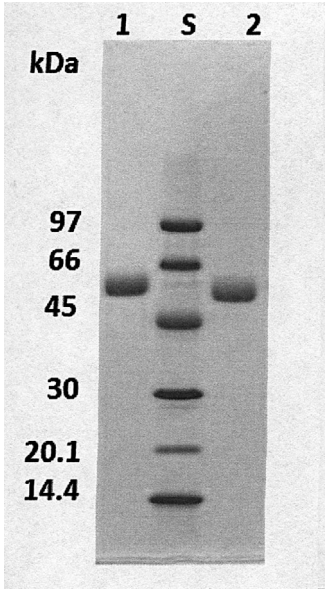


Figure 2. SDS PAGE analysis of recombinant *Cu*GE before (lane 1) and after treatment with endoglycosidase H (lane 2). Molecular weight in kDa of the standards (lane S) are indicated on the left side.

Differential Scanning Calorimetry

Thermostability of *Cu*GE was determined by Differential Scanning Calorimetry (DSC) using a MicroCal VP-Capillary Differential Scanning Calorimeter. The thermal denaturation temperature, T_d (°C), was defined as the top of the denaturation peak (major endothermic peak) in the thermogram (Cp vs. T) obtained after heating the enzyme solution (approximately 0.5 mg/ml) in buffer (50 mM acetate, pH 5.0) at a constant programmed heating rate of 90 K/hr. Sample- and reference-solutions (approximately 0.2 ml) were loaded into the calorimeter (reference: buffer without enzyme) from storage conditions at 10 °C and thermally pre-equilibrated for 20 min at 20 °C prior to DSC scan from 20 °C to 120 °C. The denaturation temperature was determined at an accuracy of approximately +/− 1°C.

Synthesis of Glucuronate Esters

^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$ at 25 °C on a Bruker Ascend 400 spectrometer operating at 400 MHz and 100 MHz, respectively, and on a Varian Mercury 300 spectrometer operating at 300 MHz for ^{1}H NMR and at 75 MHz for the ^{13}C NMR.

*Benzyl (methyl 4-O-methyl-α-D-glucopyranoside) uronate (**1**)*

4-*O*-Methyl-glucuronic acid **8** (Li and Helm, 1995a) (250 mg, 1.13 mmol) was dissolved in dry *N*,*N*-dimethylformamide (DMF) (4 ml) and cooled to 0 °C. TBAF (1 M solution in tetrahydrofuran

(THF), 1.24 ml, 1.24 mmol) and BnBr (0.147 ml, 1.24 mmol) were added to the solution and the reaction mixture let to stir at room temperature for 20 h. The solvent was evaporated and the residue purified by silica gel column chromatography (heptane/ethyl acetate 7:3) to afford **1** as a colorless oil (67.5 mg, 20%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.43–7.30 (m, 5H, Ar*H*), 5.26 (s, 2H, OCH_2Ph), 4.82 (d, $J = 3.8$ Hz, 1H, H-1), 4.11 (d, $J = 9.8$ Hz, 1H, H-5), 3.77 (t, $J = 9.2$ Hz, 1H, H-3), 3.59 (m, 1H, H-2), 3.45 (s, 3H, OCH_3(anom)), 3.37 (s, 3H, OCH_3), 3.37 (t, $J = 9.4$ Hz, 1H, H-4), 2.58 (s, 1H, OH), 2.08 (s, 1H, OH). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.4, 128.6, 128.5, 99.4, 80.9, 74.5, 72.1, 70.3, 67.4, 60.4, 55.9. ESIMS m/z: $[M + Na]^+$ calcd for $C_{15}H_{20}NaO_7$: 335.2946; found: 335.1101.

*Benzyl (methyl α-D-glucopyranoside) uronate (**2**)*

Glucuronic acid **10** (1.0 g, 4.80 mmol) was dissolved in dry DMF (15 ml) and TBAF (1 M THF solution, 5.48 ml) was added at 0 °C. BnBr (0.650 ml, 5.25 mmol) was added over 1 minute. After 20 h at room temperature the solvent was co-evaporated with toluene (4 × 25 ml). The residue was purified by silica gel column chromatography (acetone/ethyl acetate 1:5) to give a colorless oil (300 mg, 21%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.32–7.20 (m, 5H, Ar*H*), 5.13 (d, $J = 12.4$, 1H, OCH_2Ph), 5.08 (d, $J = 12.4$ Hz, 1H, OCH_2Ph), 4.76 (d, $J = 3.5$ Hz, 1H, H-1), 4.07 (d, $J = 9.2$ Hz, 1H, H-5), 3.72 (t, $J = 9.4$ Hz, 1H, H-3), 3.64 (t, $J = 9.3$ Hz, 1H, H-4), 3.52 (dd, $J = 9.1$, 3.6 Hz, 1H, H-2), 3.36 (s, 3H, OCH_3). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.1, 135.2, 128.6, 128.3, 128.2, 100.0, 73.3, 71.6, 71.3, 71.1, 67.3, 55.7. ESIMS m/z: $[M + Na]^+$ calcd for $C_{14}H_{18}NaO_7$: 321.2687; found: 321.0942.

*Phenylpropyl (methyl 2,3,4-tri-O-benzyl-α-D-glucopyranoside) uronate (**13**)*

Compound **12** (3.00 g, 6.27 mmol) was dissolved in dry CH_2Cl_2 (50 ml) and the solution cooled to 0 °C. 3-Phenylpropan-1-ol (4.26 ml, 31.3 mmol) and DMAP (76.6 mg, 0.627 mmol) were added, followed by *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (1.44 g, 7.52 mmol). After 4 h at room temperature the mixture was diluted with CH_2Cl_2 and washed twice with brine. The organic layer was dried over $MgSO_4$ and the solvent evaporated at reduced pressure. The residue was purified by silica gel column chromatography (heptane/ethyl acetate 7:3) to give **13** as a colorless oil (3.29 g, 88%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.59–7.09 (m, 20H, Ar*H*), 5.02 (d, $J = 10.9$ Hz, 1H, OCH_2Ph), 4.89 (d, $J = 10.9$ Hz, 1H, OCH_2Ph), 4.88 (d, $J = 10.9$ Hz, 1H, OCH_2Ph), 4.86 (d, $J = 12.0$ Hz, 1H, OCH_2Ph), 4.71 (d, $J = 12.1$ Hz, 1H, OCH_2Ph), 4.67 (d, $J = 3.5$ Hz, 1H, H-1), 4.65 (d, $J = 10.9$ Hz, 1H, OCH_2Ph), 4.25 (d, $J = 10.0$ Hz, 1H, H-5), 4.18 (tdd, $J = 10.9$, 6.7, 4.2 Hz, 2H, H-α), 4.06 (t, $J = 9.3$ Hz, 1H, H-3), 3.81 (dd, $J = 9.9$, 9.1 Hz, 1H, H-4), 3.64 (dd, $J = 9.6$, 3.5 Hz, 1H, H-2), 3.47 (s, 3H, OCH_3), 2.73–2.63 (m, 2H, H-γ), 1.98 (tt, $J = 13.4$, 6.7 Hz, 2H, H-β). ^{13}C NMR (100 MHz, $CDCl_3$): δ 169.8, 140.9, 138.6, 138.0, 138.0, 128.6, 128.5, 128.5, 128.4, 128.3, 128.1, 128.0, 127.8, 127.7, 126.1, 98.8, 81.5, 79.7, 79.4, 75.9, 75.2, 73.7, 70.3, 65.0, 55.7, 32.0, 30.0.

*Phenylpropyl (methyl α-D-glucopyranoside) uronate (**3**)*

Compound **13** (3.29 g, 5.52 mmol) was dissolved in dry THF (35 ml) and an excess of Pd/C was added (220 mg). The suspension was degassed and backfilled with H_2 three times then stirred under hydrogen atmosphere (1 atm) for 24 h at room temperature. The resulting mixture was filtered through a Celite pad and rinsed with two volumes of THF (15 ml). The filtrate was concentrated *in vacuo* and purified on a silica gel column (CH_2Cl_2/CH_3OH 95:5) giving 3 as a colorless oil (1.50 g, 83%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.22–7.02 (m, 5H, Ar*H*), 4.73 (d, $J = 3.6$ Hz, 1H, H-1), 4.12 (t, $J = 6.6$ Hz, 2H, H-α), 4.03 (d, $J = 9.7$ Hz, 1H, H-5), 3.73 (t, $J = 9.2$ Hz, 1H, H-3), 3.65 (t, $J = 9.4$ Hz, 1H, H-4), 3.54 (dd, $J = 9.5$, 3.7 Hz, 1H, H-2), 3.36 (s, 3H, OCH_3), 2.60 (dd, $J = 9.1$, 6.2 Hz, 2H, H-γ), 1.92 (dq, $J = 9.2$, 6.7 Hz, 2H, H-β). ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.2, 141.1, 128.5, 128.4, 126.1, 100.0, 73.5, 71.6, 71.5, 70.8, 65.1, 55.7, 31.9, 30.0. ESIMS m/z: $[M + Na]^+$ calcd for $C_{16}H_{22}NaO_7$: 349.3205; found: 349.1264.

*Phenyl (methyl 2,3,4-tri-O-benzyl-α-D-glucopyranoside) uronate (**14**)*

Compound **12** (3.00 g, 6.27 mmol) was dissolved in dry CH_2Cl_2 (50 ml) and the solution cooled to 0 °C. Phenol (2.95 g, 31.3 mmol) and DMAP (76.6 mg, 0.627 mmol) were added, followed by *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (1.44 g, 7.52 mmol). The reaction was stirred at room temperature for 3 h after which the starting material had been consumed completely. The mixture was diluted with CH_2Cl_2 and washed twice with brine. The organic layer was dried over $MgSO_4$ and the solvent evaporated at reduced pressure. The residue was purified by silica gel column chromatography (toluene/acetone 9:1) to give **14** as a colorless oil (2.61 g, 75%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.39–6.91 (m, 20H, Ar*H*), 4.93 (d, $J = 10.9$ Hz, 1H, OCH_2Ph), 4.82 (d, $J = 10.7$ Hz, 1H, OCH_2Ph), 4.78 (d, $J = 10.9$ Hz, 1H, OCH_2Ph), 4.76 (d, $J = 12.1$ Hz, 1H, OCH_2Ph), 4.61 (d, $J = 10.7$ Hz, 1H, OCH_2Ph), 4.60 (d, $J = 12.1$ Hz, 1H, OCH_2Ph), 4.60 (d, $J = 3.3$ Hz, 1H, H-1), 4.34 (d, $J = 9.9$ Hz, 1H, H-5), 4.00 (t, $J = 9.3$ Hz, 1H, H-3), 3.81 (dd, $J = 9.1$, 9.9 Hz, 1H, H-4), 3.56 (dd, $J = 9.6$, 3.5 Hz, 1H, H-2), 3.40 (s, 3H, OCH_3). ^{13}C NMR (75 MHz, $CDCl_3$): δ 167.2, 149.3, 137.5, 136.9, 136.8, 128.5, 127.5, 127.4, 127.4, 127.2, 127.0, 126.9, 125.2, 120.2, 97.9, 80.4, 78.6, 78.3, 74.9, 74.3, 72.7, 69.4, 54.8.

*Phenyl (methyl α-D-glucopyranoside) uronate (**4**)*

Compound **14** (2.61 g, 4.70 mmol) was dissolved in dry THF (30 ml) and an excess of Pd/C was added (200 mg). The suspension was degassed and backfilled with H_2 three times then stirred under a hydrogen atmosphere (1 atm) for 24 h at room temperature. The resulting mixture was filtered through a Celite pad and rinsed with two volumes of THF (15 ml). The filtrate was concentrated under reduced pressure to give **4** as a white solid (1.26 g, 94%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.45–6.98 (m, 5H, Ar*H*), 4.71 (d, $J = 3.7$ Hz, 1H, H-1), 4.28 (d, $J = 9.8$ Hz, 1H, H-5), 3.82 (t, $J = 9.2$ Hz, 1H, H-3), 3.68 (dd, $J = 9.0$, 9.8 Hz, 1H, H-4), 3.62 (dd, $J = 9.5$, 3.7 Hz, 1H, H-2), 3.41 (s, 3H, OCH_3). ^{13}C NMR (100 MHz, $CDCl_3$): δ 168.7, 150.3,

129.5, 126.3, 121.4, 100.0, 73.5, 71.7, 71.5, 71.0, 55.9. ESIMS m/z: $[M + Na]^+$ calcd for $C_{13}H_{16}NaO_7$: 307.2428; found: 307.0774.

Ethyl D-glucuronate

To a 100 ml conical flask containing 3.5 g of dried 4 Å molecular sieves were added D-glucuronic acid (490 mg, 2.5 mmol), 99.9% ethanol (292 μL, 5.0 mmol), immobilized CALB (NZ435, 200 mg) and *tert*-butanol (10 ml). Under shaking the mixture was heated to 60 °C for 192 h after which thin-layer chromatography (TLC) indicated formation of the desired product. The reaction mixture was filtered and the filtrate concentrated *in vacuo*. The residue was redissolved in 1:1 $CHCl_3/CH_3OH$ (50 ml), concentrated with Celite and purified by silica gel column chromatography ($CHCl_3/CH_3OH/H_2O$ 80:25:2) to produce ethyl D-glucuronate (80 mg, 14%) contaminated with small amounts of D-glucofuranurono-6,3-lactone (glucuronolactone) (Bock and Pedersen, 1983; Wang et al., 2010) as an inseparable impurity (~1:6 ratio as compared to ethyl D-glucuronate). ^{13}C NMR (100 MHz, D_2O) δ 172.6, 171.7, 97.3 (β), 93.5 (α), 76.3, 75.8, 74.8, 73.4, 72.6, 72.4, 72.1, 71.9, 64.0, 64.0, 14.4. ESIMS m/z: $[M + Na]^+$ calcd for $C_8H_{14}NaO_7$: 245.1768; found: 245.0629. D-glucofuranurono-6,3-lactone (Bock and Pedersen, 1983) (predominantly β): ^{13}C NMR (100 MHz, D_2O) δ 178.6, 103.9 (β), 84.9, 78.7, 77.9, 70.2.

Enzymatic Methods

Semi-quantitative detection of glucuronoyl esterase activity was conducted by TLC analysis of aliquots from incubation mixtures (35 °C) containing the tested enzymes (0.025 mg/ml) and the substrates (8 mM) in 50 mM sodium phosphate buffer, pH 6. Reactions were run for 42 h and aliquots were withdrawn for TLC analysis after 2, 18, and 42 h. Aliquots were chromatographed on aluminum TLC plates coated with silica gel 60 (Merck) in $CH_2Cl_2/CH_3OH/H_2O$ (80:25:4) and the conversion of glucuronoyl esters into the corresponding alcohols and glucuronic acids was visualized by development with 1 M sulfuric acid and heating.

Kinetic parameters were determined by enzymatic hydrolysis at 30 °C in 96-well MultiScreen 10 kDa cut-off ultrafiltration plates (Millipore) for 10 min in 50 mM sodium phosphate buffer, pH 6. Substrate concentrations varied from 0.025 to 150 mM, while enzyme concentrations were in the range of 0.001 to 0.1 mg/ml (0.02–2 μM) depending on the compound studied. After incubation, reactions were stopped by rapid cooling to 4 °C followed by mechanical removal of the enzyme from the solution by ultracentrifugation (4000 g) through a 10 kDa membrane in a pre-cooled centrifuge for 20 min at 4 °C. The degree of substrate hydrolysis was determined on the basis of integrated areas of the UV-absorbing alcohols produced within different time intervals as quantified by HPLC (ICS-5000 Dionex system, Thermo Fisher Scientific) using a Luna C18 3 μm column (100 Å, 150 × 4.6 mm, Phenomenex) and UV detection at 210 nm. Elution was carried out with a mixture of acetonitrile/0.01% formic acid solution at pH 3.6 (isocratic, 35:65 V/V) at a flow rate of 0.7 ml/min. The obtained data were fitted to the Michaelis–Menten equation to estimate the values for K_m, V_{max} and k_{cat}.

Results

Substrate Synthesis

In order to mimic the proposed α- and -ester LCCs, methyl glycoside esters **1**–**3** were prepared by chemical synthesis (Fig. 3). In addition to α- and -esters, one could envision the existence of phenyl esters as minor LCC components, and the phenyl ester **4** was also chosen as a synthetic target. Hereby four LCC model compounds were prepared for characterization of GEs with respect to substrate specificity both in relation to the alcohol part and the 4-*O*-methyl substituent.

Benzyl (methyl 4-*O*-methyl-α-D-glucopyranoside) uronate (**1**) was synthesized in five steps from commercially available methyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-α-D-glucopyranoside (**5**) (Scheme 1). First, selective ring-opening of the benzylidene acetal was performed with triethylsilane and trifluoroacetic acid which was followed by methylation under standard conditions to afford methyl ether **7** (Yoneda et al., 2005). Conventional hydrogenolysis of the benzyl ether in the presence of palladium on charcoal gave a triol which was selectively oxidized at the 6-position with (2,2,6,6-tetramethylpiperidin-1-yl) oxy (TEMPO), sodium bromide and sodium hypochlorite in water at pH 10–11 (Li and Helm, 1995a). The resulting 4-*O*-methyl-glucuronic acid **8** was esterified with benzyl bromide (BnBr) and tetrabutylammonium fluoride (TBAF) to give **1** following an analogous literature protocol (Bowkett et al., 2007).

The remaining phenyl, benzyl and phenylpropyl esters of (methyl α-D-glucopyranoside) uronate, i.e. compounds **2**–**4**, were prepared from commercially available methyl α-D-glucopyranoside in a few and straightforward steps (Scheme 2). The benzyl ester **2** was obtained in two steps by subjecting the starting glucoside (**9**) to oxidation with TEMPO (Li and Helm, 1995a) followed by benzylation with benzyl bromide (Bowkett et al., 2007). The phenyl and phenylpropyl esters **3** and **4** required protection of the secondary alcohols as benzyl ethers in the starting glucoside. Thus, by using slightly modified literature protocols (see Supplementary material) methyl α-D-glucopyranoside (**9**) was converted into methyl 2,3,4-tri-*O*-benzyl-α-D-glucopyranuronic acid (**12**) (Guan et al., 2012). Subsequent reaction with the corresponding alcohols (3-phenylpropan-1-ol and phenol), *N*,

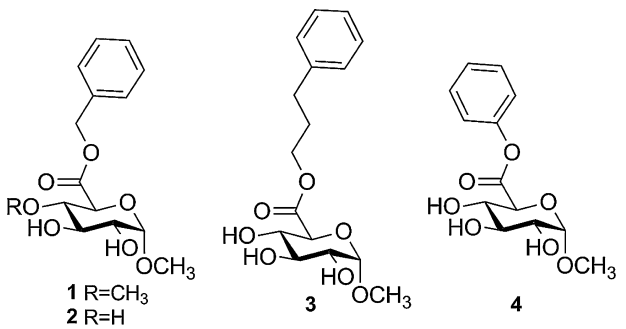
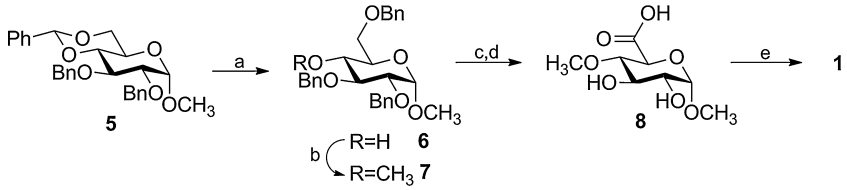


Figure 3. Synthesized ester LCC model compounds.



Scheme 1. Reagents and conditions: a) $(CH_3CH_2)_3SiH$, CF_3CO_2H, CH_2Cl_2, 0°C to rt, 77% b) NaH, CH_3I, THF, 0°C, 92% c) H_2, Pd/C, CH_3OH, rt, 85% d) TEMPO, NaBr, NaClO, H_2O, rt, 60% e) BnBr, TBAF, DMF, rt, 35%.

N-dimethyl-4-aminopyridine (DMAP) and the coupling agent *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide (EDAC) gave protected esters **13** and **14** which were deprotected by hydrogenolysis to afford **3** and **4**, respectively.

Ethyl glucuronate was prepared enzymatically in a single step in 20% yield from ethanol and glucuronic acid using immobilized *Candida antarctica* lipase B (CALB, NZ435) in *tert*-butanol according to published procedures (Katsimpouras et al., 2014; Moreau et al., 2004).

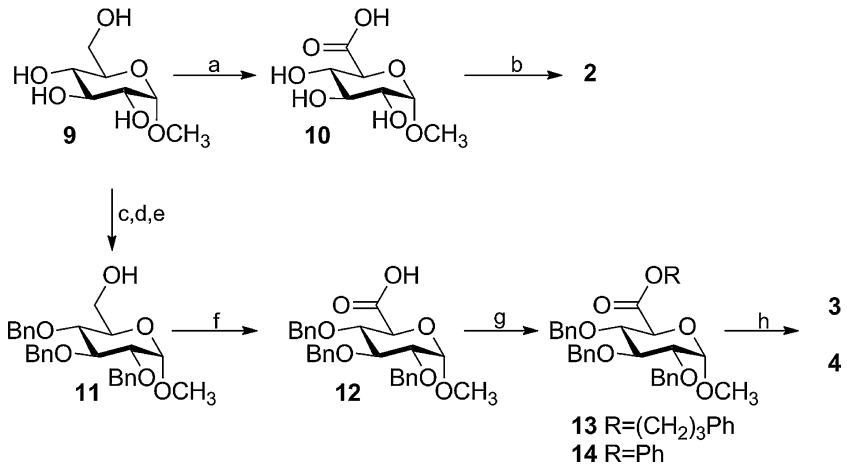
Identification and Purification of CuGE

The *Cerrena unicolor* strain was isolated from fungal spores collected in Kamchatka, Russia in 1997 and its DNA was extracted and sequenced (see Supplementary material). The gene for the full length protein encoding 474 amino acids was cloned and recombinantly expressed in *Aspergillus oryzae* and subsequently purified by standard techniques. Similarly, the known *Sc*GE (Špániková and Biely, 2006) (swissprot:D8QLP9) was recombinantly expressed in *A. oryzae* and purified to homogeneity using conventional procedures.

Physico-chemical Properties of CuGE

The *Cu*GE was purified to homogeneity as visualized by SDS PAGE and the molecular mass was found to be 58 kDa (Fig. 2). After treatment with Endoglycosidase H, the molecular weight on SDS PAGE decreased to 55 kDa clearly indicating that the enzyme is *N*-glycosylated which is also supported by the presence of a predicted *N*-glycosylation site in the protein sequence. However, the Mw is still 7 kDa higher than the predicted molecular weight of 48 kDa of the mature protein. This is in accordance with *Cu*GE being modular with a catalytic core and a family 1 carbohydrate binding module (CBM) linked together with a serine and threonine rich linker. Such linker regions are known to be prone to *O*-glycosylation, and a clear glycosylation pattern with 162 Da spacing was indeed observed around 51 kDa (Fig. S1). *Cu*GE was found to have good stability at pH 5.0 with a thermal denaturation temperature of 70 °C ensuring that the enzyme is fully stable during the kinetic analysis.

The esterase from *Cerrena unicolor* was found to share highest homology values to proteins classified as esterases within CE15. Compared to previously identified glucuronoyl esterases from *Hypocrea jecorina* Cip2_GE (Li et al., 2007) (swissprot:G0RV93) and



Scheme 2. Reagents and conditions: a) TEMPO, NaBr, NaClO, H_2O, rt, 81% b) BnBr, TBAF, DMF, 0°C, 21% c) Ph_3CCl, pyridine, 90°C, 86% d) BnBr, $(C_4H_9)_4NI$, NaH, DMF, rt e) H_2SO_4, CH_3OH, rt, 80% (over 2 steps) f) TEMPO, $PhI(OAc)_2$, CH_2Cl_2, H_2O, rt, 74% g) (**13**) 3-phenylpropan-1-ol, DMAP, EDAC, CH_2Cl_2, 88%. (**14**) PhOH, DMAP, EDAC, CH_2Cl_2, rt, 75% h) (**3**) H_2, Pd/C, THF, 83%. (**4**) H_2, Pd/C, THF, 94%.

from *Schizophyllum commune* (Špániková and Biely, 2006) (swissprot:D8QLP9) *Cu*GE shared 55% and 62% sequence identity, respectively.

Characterization of GEs

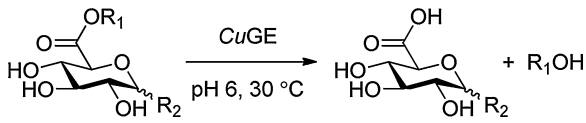
Initial characterization of the substrate specificity of *Cu*GE was conducted at pH 6.0 at 30°C using simple esters of glucuronic acid existing as α/β anomeric mixtures. pH 6.0 was chosen in order to avoid spontaneous autohydrolysis of the substrates observed at extreme pH values. GE activities were judged semi-quantitatively by TLC analysis and are depicted in Table I. Overall, a clear preference for the more bulky substrates containing an aryl or alkenyl group in the alcohol part was observed. The same trend was observed for a number of other proprietary fungal GEs (data not shown). This seems to support the proposed activity of GEs on bulky lignin carbohydrate ester linkages. Further characterization by means of Michaelis–Menten kinetics on the synthesized substrates was performed with *Cu*GE as well as with the well-known GE from *Schizophyllum commune* (*Sc*GE) for comparison. Kinetic parameters were determined quantitatively by HPLC (UV detection) by monitoring formation of phenol, benzyl alcohol or 3-phenylpropanol (Table II). The obtained kinetic data were fitted to Michaelis–Menten kinetics using non-linear regression analysis of V as a function of [S] (Fig. 4). For both enzymes, a 25–50 times higher catalytic efficiency (k_{cat}/K_m) was observed for substrates carrying a 4-*O*-methyl substituent in the glucuronic acid (**1** versus **2**, Table II); a trend which has also been observed previously (Duranová et al., 2009). Significant autohydrolysis of **4** even at pH 6.0 prevented us from obtaining full kinetic parameters for this compound, but surprisingly low binding affinities (K_m) and very high conversions were found despite the lack of a 4-*O*-methyl substituent. Comparison of the substrate specificity for esters carrying a benzyl versus a phenylpropyl alcohol (i.e. mimics of α- and -esters, respectively) revealed comparable binding affinities of *Sc*GE and *Cu*GE towards the two substrates **2** and **3**, however, higher catalytic efficiencies were observed for the benzyl ester **2** with both enzymes. In general, *Sc*GE showed slightly higher catalytic efficiency than *Cu*GE, but the two enzymes seem to behave quite similarly.

Discussion

In general, detailed knowledge of the macromolecular architecture of LCCs is hampered by the complex nature of such structures and the inability to isolate well-defined fragments. Due to this there is still debate on the exact identity and abundance of ester LCCs in different plant tissues, however, both α- and -glucuronoyl esters of lignin alcohols do indeed seem to exist in a variety of lignocellulosic plants (Balakshin and Capanema, 2007; Balakshin et al., 2011; Imamura et al., 1994; Watanabe, 1995; Yuan et al., 2011). As a result of limited substrate availability most biochemical characterization of GEs to date has been conducted using methyl esters of (4-*O*-methyl)-glucuronic acid, however, we believe that α- and -esters of (4-*O*-methyl)-glucuronic acids resembling also the lignin part of naturally occurring LCCs would constitute better screening substrates in the search for an industrially relevant GE. We are convinced that the concise synthesis of application relevant GE substrates such as **1** reported here may provide a valuable tool for screening, characterization and selection of GEs for industrial delignification of biomass.

Currently, seven CE15 GEs of fungal origin have been purified and biochemically characterized, whereas only the *Hypocrea jecorina* Cip2_GE (Pokkuluri et al., 2011) and the *Sporotrichum thermophile* GE2 (Charavgi et al., 2013) have been crystallized and have had their structures solved by X-ray crystallography. An inactive variant of the latter was even crystallized in a complex with the substrate analogue methyl 4-*O*-methyl-D-glucopyranuronate providing valuable insights into substrate binding within the active site (Charavgi et al., 2013). The GE from *Cerrena unicolor* as well as the majority of the GEs reported in the literature are bimodular consisting of a catalytic domain, a linker region and an *N*-terminal family 1 CBM, whereas the originally discovered GE from the wood-rotting fungus *Schizophyllum commune* is comprised only of a catalytic domain (Li et al., 2007). The CE15 GEs belong to serine type esterases requiring no metal ion co-factors for catalytic activity (Li et al., 2007; Topakas et al., 2010). As reported by several groups, optimal catalytic efficiencies of the currently described GEs are generally achieved in the range pH 5–7 and 40–60°C. *Cu*GE will find use in the higher range of this temperature interval with the determined denaturation temperature of 70 °C. *Cu*GE and *Sc*GE were found to have a k_{cat} of 15–17 s^{-1} on **3** at pH 6.0 (phosphate

Table I. Activity[a] of *Cu*GE measured semi-quantitatively by TLC.



Entry	R_1	R_2	2 h	18 h	42 h
1	Methyl	OH	trace	+	+
2	Ethyl	OH	trace	+	++
3	Allyl	OH	+	++	++
4	Benzyl	OH	++	+++	+++
5	$Ph(CH_2)_3$[b]	αOCH$_3$	+++	+++	+++

[a]+++: High activity (70–100%); ++: Medium activity (40–70%); +: Low activity (10–40%); trace: ≤ 10% conversion.

Table II. Kinetic parameters for *Sc*GE and *Cu*GE at pH 6.0, 30 °C using synthesized substrates.

	*Sc*GE			*Cu*GE		
	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [$mM^{-1} s^{-1}$]	K_m [mM]	k_{cat} [s^{-1}]	k_{cat}/K_m [$mM^{-1} s^{-1}$]
1	3.7 (1.2)	118 (9.4)	32	4.6 (1.0)	129 (7.6)	28
2	51 (8.0)	64 (5.0)	1.2	80 (24)	48 (7.5)	0.6
3	66 (22)	15 (2.7)	0.2	55 (14)	17 (1.8)	0.3
4[a]	11 (5.8)	n.d.	n.d.	8.9 (3.2)	n.d.	n.d.

Numbers in parentheses are the estimates of the standard errors.
[a]Due to significant autohydrolysis, full kinetic parameters could not be obtained.

buffer) and 30 °C. On an almost identical substrate (Substrate V in (Katsimpouras et al.,2014), the two GEs under investigation (from *Podospora anserina* and *Sporotrichum thermophile*) had a k_{cat} of 0.8–2.8 s^{-1}, respectively, at 50 °C and pH 6.0 (phosphate buffer) (the two substrates differ by **3** existing as a single α-anomer as opposed to an α/β mixture). With respect to GE substrate specificity, the observed difference in reactivity targeting glucuronoyl esters with and without a 4-*O*-methyl substituent (Table II) may be associated with additional van der Waals interactions between the enzyme and the 4-*O*-methyl substituent resulting in stronger substrate binding within the active site as reported by (Charavgi et al. 2013). An order of magnitude lower binding affinity (higher K_m values) was indeed obtained with **2** not carrying a 4-*O*-methyl substituent as compared to its 4-*O*-methyl counterpart (**1**). Although not directly comparable, some of the data reported by Biely and co-workers (Li et al., 2007; Špániková et al., 2007) indicate a slightly higher preference of *Sc*GE for methyl esters of glucuronic acid as opposed to bulkier arylalkyl esters. In contrast, we observed a clear preference of *Cu*GE and a number of other fungal GEs for esters of bulky alcohols. These results fit well with the observation that the active site of CE15 GEs is exposed to the surface of the enzyme (Pokkuluri et al., 2011) potentially providing access to large substrates such as lignin ester LCCs. Based on the quantitative kinetic data presented in Table II it is obvious that α-esters are cleaved more easily than -esters for both *Sc*GE and *Cu*GE. Among different -esters, (Katsimpouras et al., 2014) reported a preference of GEs from the fungi *Podospora anserina* and *Sporotrichum thermophile* for -esters carrying an α-β conjugated double bond representative of the cinnamyl monolignols which are found in lignin. Although not described in literature, the existence of lignin phenol ester LCCs via glucuronic acid cannot be ruled out, but based on the rapid autohydrolysis of **4** (the instability of phenol esters is known in literature (Stefanidis and Jencks, 1993), such LCC structures (if occurring in nature) would not be expected to play a major role in the recalcitrance of lignocellulose.

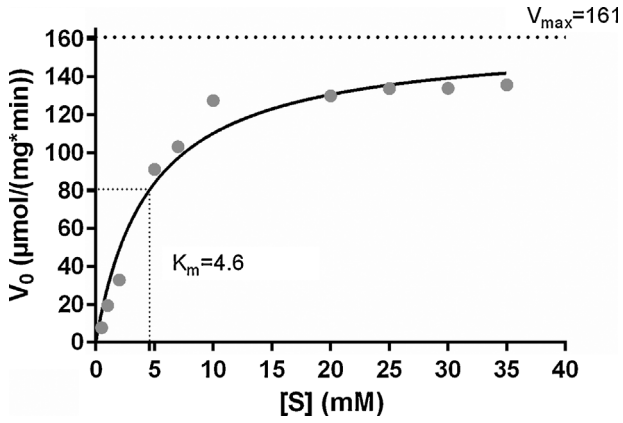


Figure 4. Degradation of 1 with *Cu*GE at pH 6.0 and 30°C.

To summarize the current knowledge on GE specificity, this enzyme class recognizes esters of glucuronic acid as their substrates, whereas other esters including esters of galacturonic acid are not recognized (Duranová et al., 2009). Based on the available literature (Duranová et al., 2009; Špániková et al., 2007) and the data obtained here, we can firmly conclude that the 4-*O*-methyl substituent in the glucuronic acid residue is the key structural determinant for the specificity of GEs and is thereby essential in order for these enzymes to work at optimal catalytic efficiency. The anomeric substitution of glucuronoyl esters seems to be of lesser importance (Špániková et al., 2007), and although most GE work has been conducted on α-anomeric glucuronoyl esters, there has even been a report showing activity on a β-anomer of a glucuronic acid ester (Duranová et al., 2009). Within the alcohol part of glucuronoyl esters, bulky arylalkyl or arylalkenyl substituents seem to be favored, and the following order of GE reactivity on glucuronoyl esters can be proposed: benzyl > cinnamyl > phenylpropyl > alkenyl > alkyl.

In conclusion, we have reported a novel fungal glucuronoyl esterase from *Cerrena unicolor* which shows a preference for bulky arylalkyl esters of 4-*O*-methyl-glucuronic acid supporting the hypothesis that GEs are involved in plant cell wall delignification by degradation of ester LCCs and may find biotechnological use in forestry, feed and biofuel industries. Although proposed already in 2006 (Špániková and Biely, 2006) and supported by our data, the suggested action of GEs on LCCs needs yet to be demonstrated in real biotechnological applications. We are currently investigating such effects of GEs for enzymatic biomass degradation using natural substrates.

This work was funded by the Danish Council for Strategic Research (SET4Future project). The authors declare that they have no conflicts of interest.

REFERENCES

Balakshin M, Capanema E, Gracz H, Chang H-M, Jameel H. 2011. Quantification of lignin–carbohydrate linkages with high-resolution NMR spectroscopy. Planta 233:1097–1110.

Balakshin MY, Capanema EA, Chang H-M. 2007. MWL fraction with a high concentration of lignin-carbohydrate linkages: Isolation and 2D NMR spectroscopic analysis. Holzforschung 61:1–7.

Balakshin MY, Capanema EA, Chen C-L, Gracz HS. 2003. Elucidation of the structures of residual and dissolved pine kraft lignins using an HMQC NMR technique. J Agric Food Chem 51:6116–6127.

Bock K, Pedersen C. 1983. Carbon-13 nuclear magnetic resonance spectroscopy of monosaccharides. Adv Carbohydr Chem Biochem 41:27–66.

Bowkett ER, Harding JR, Maggs JL, Park BK, Perrie JA, Stachulski AV. 2007. Efficient synthesis of 1β-*O*-acyl glucuronides via selective acylation of allyl or benzyl D-glucuronate. Tetrahedron 63:7596–7605.

Bunzel M. 2010. Chemistry and occurrence of hydroxycinnamate oligomers. Phytochem Rev 9:47–64.

Charavgi M-D, Dimarogona M, Topakas E, Christakopoulos P, Chrysina ED. 2013. The structure of a novel glucuronoyl esterase from *Myceliophthora thermophila* gives new insights into its role as a potential biocatalyst. Acta Crystallogr D Biol Crystallogr 69:63–73.

Christensen T, Woeldike H, Boel E, Mortensen SB, Hjortshoej K, Thim L, Hansen MT. 1988. High level expression of recombinant genes in *Aspergillus oryzae*. Nat Biotechnol 6:1419–1422.

Duranová M, Hirsch J, Kolenová K, Biely P. 2009. Fungal glucuronoyl esterases and substrate uronic acid recognition. Biosci Biotechnol Biochem 73:2483–2487.

Guan Z, Zhang L-H, Sinaÿ P, Zhang Y. 2012. Study on metal-induced reactions of α-diazocarbonyl glucosides. J Org Chem 77:8888–8895.

Imamura T, Watanabe T, Kuwahara M, Koshijima T. 1994. Ester linkages between lignin and glucuronic acid in lignin-carbohydrate complexes from *Fagus crenata*. Phytochemistry 37:1165–1173.

Katsimpouras C, Bénarouche A, Navarro D, Karpusas M, Dimarogona M, Berrin J-G, Christakopoulos P, Topakas E. 2014. Enzymatic synthesis of model substrates recognized by glucuronoyl esterases from *Podospora anserina* and *Myceliophthora thermophila*. Appl Microbiol Biotechnol 98:5507–5516.

Li K, Helm RF. 1995a. A practical synthesis of methyl 4-*O*-methyl-α-D-glucopyranosiduronic acid. Carbohydr Res 273:249–253.

Li K, Helm RF. 1995b. Synthesis and rearrangement reactions of ester-linked lignin-carbohydrate model compounds. J Agric Food Chem 43:2098–2103.

Li X-L, Špániková S, de Vries RP, Biely P. 2007. Identification of genes encoding microbial glucuronoyl esterases. FEBS Lett 581:4029–4035.

Moreau B, Lognay GC, Blecker C, Brohée J-C, Chéry F, Rollin P, Paquot M, Marlier M. 2004. Synthesis of novel D-glucuronic acid fatty esters using *Candida antarctica* lipase in *tert*-butanol. Biotechnol Lett 26:419–424.

Pokkuluri PR, Duke NEC, Wood SJ, Cotta MA, Li X-L, Biely P, Schiffer M. 2011. Structure of the catalytic domain of glucuronoyl esterase Cip2 from *Hypocrea jecorina*. Proteins 79:2588–2592.

Špániková S, Biely P. 2006. Glucuronoyl esterase—Novel carbohydrate esterase produced by *Schizophyllum commune*. FEBS Lett 580:4597–4601.

Špániková S, Poláková M, Joniak D, Hirsch J, Biely P. 2007. Synthetic esters recognized by glucuronoyl esterase from *Schizophyllum commune*. Arch Microbiol 188:185–189.

Stefanidis D, Jencks WP. 1993. General base catalysis of ester hydrolysis. J Am Chem Soc 115:6045–6050.

Topakas E, Moukouli M, Dimarogona M, Vafiadi C, Christakopoulos P. 2010. Functional expression of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*: identification of the nucleophilic serine. Appl Microbiol Biotechnol 87:1765–1772.

Vafiadi C, Topakas E, Biely P, Christakopoulos P. 2009. Purification, characterization and mass spectrometric sequencing of a thermophilic glucuronoyl esterase from *Sporotrichum thermophile*. FEMS Microbiol Lett 296:178–184.

Wang R, Neoh TL, Kobayashi T, Miyake Y, Hosoda A, Taniguchi H, Adachi S. 2010. Degradation kinetics of glucuronic acid in subcritical water. Biosci Biotechnol Biochem 74:601–605.

Watanabe T. 1995. Important properties of lignin-carbohydrates complexes (LCCs) in environmentally safe paper making. Trends Glycosci Glycotechnol 7:57–68.

Yoneda Y, Kawada T, Rosenau T, Kosma P. 2005. Synthesis of methyl 4'-*O*-methyl-$^{13}C_{12}$-β-D-cellobioside from $^{13}C_6$-D-glucose. Part 1: Reaction optimization and synthesis. Carbohydr Res 340:2428–2435.

Yuan T-Q, Sun S-N, Xu F, Sun R-C. 2011. Characterization of lignin structures and lignin–carbohydrate complex (LCC) linkages by quantitative ^{13}C and 2D HSQC NMR spectroscopy. J Agric Food Chem 59:10604–10614.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Journal of Biotechnology 219 (2016) 117–123

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Improved biomass degradation using fungal glucuronoyl—esterases—hydrolysis of natural corn fiber substrate

Clotilde d'Errico[a], Johan Börjesson[b], Hanshu Ding[b], Kristian B.R.M. Krogh[b], Nikolaj Spodsberg[b], Robert Madsen[a], Rune Nygaard Monrad[b,*]

[a] Department of Chemistry, Technical University of Denmark, Kgs. Lyngby 2800, Denmark
[b] Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Denmark

ARTICLE INFO

Article history:
Received 16 September 2015
Received in revised form
26 November 2015
Accepted 15 December 2015
Available online 19 December 2015

Keywords:
Enzymatic delignification
Lignin-carbohydrate complexes
Glucuronoyl esterase
Corn fiber
Lignocellulose
Biofuel

ABSTRACT

Lignin-carbohydrate complexes (LCCs) are in part responsible for the recalcitrance of lignocellulosics in relation to industrial utilization of biomass for biofuels. Glucuronoyl esterases (GEs) belonging to the carbohydrate esterase family 15 have been proposed to be able to degrade ester LCCs between glucuronic acids in xylans and lignin alcohols. By means of synthesized complex LCC model substrates we provide kinetic data suggesting a preference of fungal GEs for esters of bulky arylalkyl alcohols such as ester LCCs. Furthermore, using natural corn fiber substrate we report the first examples of improved degradation of lignocellulosic biomass by the use of GEs. Improved C5 sugar, glucose and glucuronic acid release was observed when heat pretreated corn fiber was incubated in the presence of GEs from *Cerrena unicolor* and *Trichoderma reesei* on top of different commercial cellulase/hemicellulase preparations. These results emphasize the potential of GEs for delignification of biomass thereby improving the overall yield of fermentable sugars for biofuel production.

1. Introduction

The interest for utilization of lignocellulosic biomass as primary energy source in society has increased enormously in the last decades due to its high abundance and low cost and significant concerns about depletion of fossil fuel resources (Sánchez and Cardona, 2008; Xu and Huang, 2014). Industrial conversion of lignocellulosic biomass into ethanol consists of three main steps: pretreatment of the biomass, enzymatic saccharification to generate fermentable sugars and microbial fermentation to produce ethanol (Xu and Huang, 2014). The inherent recalcitrance of lignocellulosic biomass requires severe pretreatment necessitating consumption of energy and chemicals and the concerted action of several enzymes at a high dosage as compared to production of starch-based ethanol (Humbird et al., 2011; Kwiatkowski et al., 2006). The recalcitrance of lignocellulosics may in part be ascribed to LCCs which are stable, covalent linkages between lignin and polysaccharides (mainly between lignin and hemicelluloses) (Aita et al., 2011; Balakshin et al., 2014; Du et al., 2014). A number of LCC linkages are implied to complicate the separation of lignin from cellulose and hemicellulose such as esters, benzyl ethers and phenyl glycosides (Balakshin et al., 2011, 2014; Watanabe, 1995). Among these the ester bonds between lignin alcohols and 4-*O*-methyl-α-D-glucuronic acid residues in xylans (Fig. 1A) are susceptible to enzymatic degradation, and GEs belonging to the CE 15 family have been proposed to degrade such ester LCCs (Špániková and Biely, 2006) thereby potentially improving the degradability of lignocellulosic plant material by enhancing access of cellulolytic and hemicellulolytic enzymes to cellulose and hemicellulose, respectively. It has indeed been shown that GEs are able to degrade simple ester LCC mimics comprising glucuronoyl esters of alkyl and arylalkyl alcohols, (Katsimpouras et al., 2014; Li et al., 2007; Špániková et al., 2007) and we have previously semi-quantitatively observed a preference of GEs for esters of bulky arylalkyl alcohols (d'Errico et al., 2015) supporting the hypothesis that GEs are capable of degrading large molecules such as ester LCCs. However, the high complexity and heterogeneity of lignocellulose has prevented thorough testing of GEs on natural substrates and thus

Abbreviations: CBM, carbohydrate binding module; CE, carbohydrate esterase; CV, column volume; *Cu*GE, *Cerrena unicolor* glucuronoyl esterase; DDQ, 2,3-dichloro-5,6-dicyano-1,4-benzoquinone; DM, dry matter; GE, glucuronoyl esterase; HL-NREL-PCS, high-liquor NREL pretreated corn stover; LCC, lignin-carbohydrate complex; NREL-PCS, NREL pretreated corn stover; *Sc*GE, *Schizophyllum commune* glucuronoyl esterase; TLC, thin-layer chromatography; *Tr*GE, *Trichoderma reesei* glucuronoyl esterase; TrβX, *Trichoderma reesei* β-xylosidase.

* Corresponding author.
E-mail address: rnmo@novozymes.com (R.N. Monrad).

http://dx.doi.org/10.1016/j.jbiotec.2015.12.024
0168-1656/

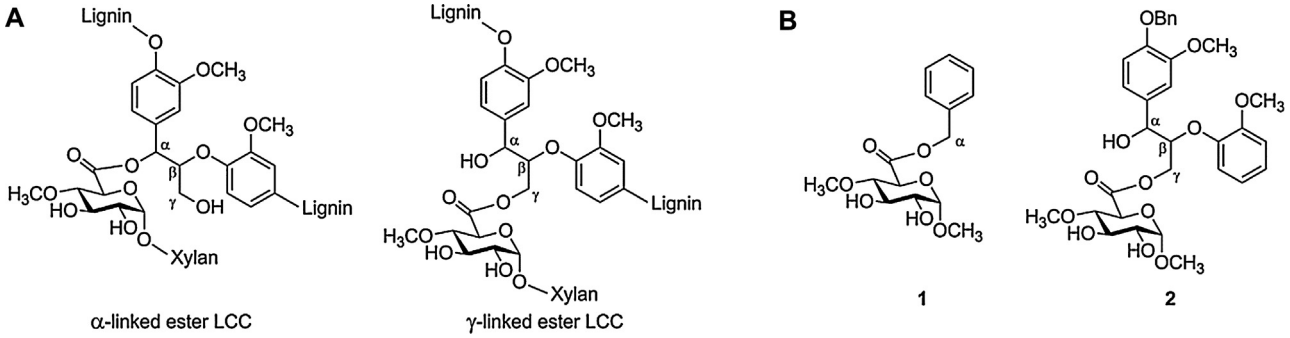


Fig. 1. (A) Representative structures of α- and γ-linked ester LCCs connecting lignin alcohols and glucuronic acid residues in xylans; (B) Synthesized advanced ester LCC model compounds.

not much is known about the biological action of GEs. Tsai et al. (2012) found that constitutive expression of *Phanerochaete carnosa* GE in *Arabidopsis thaliana* led to altered cell wall composition and improved xylose recovery in transgenic plants and recently Biely et al. (2015) showed the first example of GE action on a polymeric substrate made by chemical methyl esterification of alkali extracted beechwood glucuronoxylan. Herein we report the first examples of improved degradation of lignocellulosic biomass from natural sources by the use of GEs (*Cerrena unicolor* (*Cu*GE) and *Trichoderma reesei* (*Tr*GE, syn. *Hypocrea jecorina* Cip2 GE)) providing direct evidence for the action of GEs on plant cell wall ester LCCs. Furthermore, using advanced, synthetic ester LCC model substrates and GEs from *C. unicolor* and *Schizophyllum commune* (*Sc*GE), we present kinetic data supporting the previously suggested preference of GEs for esters of bulky lignin alcohols.

2. Materials and methods

Benzyl D-glucuronate was purchased from CarboSynth (Compton, UK). All other chemicals were purchased from Sigma–Aldrich (St. Louis, MO). Ultraflo® L (*Humicola insolens* β-glucanase preparation) and Cellic® CTec (*T. reesei* cellulase preparation) are commercial products of Novozymes A/S. GH3 β-xylosidase from *T. reesei* (*Trβ*X) was obtained recombinantly by expression in *Aspergillus oryzae* using standard techniquies as described in Rasmussen et al. (2006).

2.1. Synthesis of advanced glucuronate esters

The synthesis of benzyl (methyl 4-*O*-methyl-α-D-glucopyranoside) uronate (**1**) was recently reported in the literature (d'Errico et al., 2015), while the advanced ester LCC model compound **2** was synthesized in 1995 (Li and Helm, 1995) and received as a generous gift from Professor Richard F. Helm. The structure of **2** was confirmed by NMR spectroscopy (Fig. S1). ^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$ at 25 °C on a Bruker Ascend 400 spectrometer operating at 400 and 100 MHz, respectively.

2.2. Pretreatment of corn fiber

Raw corn fiber containing 85% DM was dried at room temperature and milled to a particle size of less than 1 mm. The material was added milliQ water to approximately 12% DM and subjected to autoclaving at 140 °C for 150 min (no washing step). The resulting heat pretreated corn fiber was used directly in incubation experiments with GEs as described below.

2.3. Cloning, expression and purification of GEs

The known GEs *Cu*GE (d'Errico et al., 2015) (GenBank accession no. KM875459), *Sc*GE (Špániková and Biely, 2006) (EMBL: EFI91386) and *Tr*GE (syn. *H. jecorina* Cip2 GE, Li et al., 2007) (EMBL: AY281368) were cloned into the expression vector pDAU109 (Schnorr and Christensen, WO2005042735) and recombinantly expressed in the host *A. oryzae* MT3568, an amdS (acetamidase) disrupted gene derivative of *A. oryzae* JaL355 (Lehmbeck and Wahlbom, WO2005070962), and subsequently purified to homogeneity using standard techniques.

2.4. Purification and characterization of TrGE

Sterile filtered *Tr*GE cultivation broth was adjusted to pH 7.0 and applied onto a hydrophobic charged induction chromatography MEP Hypercel (Pall Corporation) column XK 26/20 (GE Healthcare). The column volume (CV) was 20 mL. The column was equilibrated in buffer A (25 mM HEPES, pH 7.0). Unbound protein was washed off with 7CVs of buffer A. The column was eluted with a linear gradient of buffer B (50 mM AcOH, pH 4.0) over 1CV followed by 5CVs of buffer B. Fractions were collected and analyzed by SDS PAGE. The fractions were pooled and adjusted to pH 7.5 with 3 M TRIS and applied onto an anion exchange Q Sepharose Fast Flow column XK 16/20 (GE Healthcare) with a CV of 20 mL. The column was equilibrated in buffer C (25 mM HEPES, pH 7.5). Unbound protein was washed off with 5CVs of buffer C, and the column was eluted with a linear gradient of buffer D (0.5 M NaCl in buffer C) over 5CVs. Fractions were analyzed by SDS PAGE, and the enzyme was recovered in the run through. The run through was concentrated and buffer exchanged with 25 mM HEPES/150 mM NaCl, pH 6.5 using a Sartorius crossflow system equipped with a polyethersulfone 5 kDA cut-off Sartocon Slice membrane (Sartorius) (Fig. 2). Deglycosylation of the purified protein using Endoglycosidase H (Roche Diagnostics), characterization by means of intact molecular weight analyses (microTOF electrospray mass spectrometry) and determination of the thermal denaturation temperature (T_d) of the protein (Differential Scanning Calorimetry) was carried out according to previously reported procedures (d'Errico et al., 2015).

2.5. Semi-quantitative GE activity measurements

Semi-quantitative detection of GE activity was conducted by TLC analysis of aliquots from incubation mixtures (35 °C) containing *Cu*GE or *Tr*GE (0.025 mg/mL) or Ultraflo® or Cellic® CTec (0.125 mg/mL) and the substrate benzyl D-glucuronate (8 mM) in 50 mM sodium phosphate buffer, pH 6.0. Aliquots were withdrawn for TLC analysis after 2 h and chromatographed on aluminum TLC plates coated with silica gel 60 (Merck) in $CH_2Cl_2/CH_3OH/H_2O$ (80:25:4). The conversion of benzyl D-glucuronate into benzyl alco-

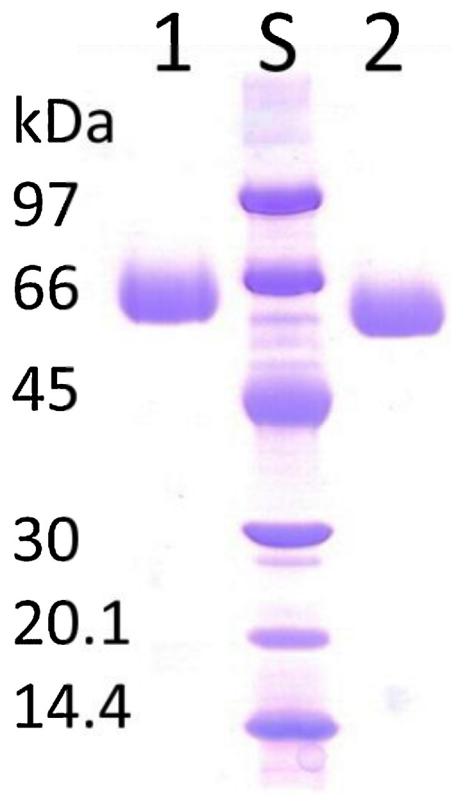


Fig. 2. SDS PAGE analysis of recombinant *Tr*GE before (lane 1) and after treatment with Endoglycosidase H (lane 2). Molecular weight in kDa of the standards (lane S) are indicated on the left side.

hol and glucuronic acid was visualized by development with 1 M sulfuric acid and heating.

2.6. Determination of kinetic parameters

Kinetic parameters were determined by enzymatic hydrolysis at pH 6 at 30 °C according to a previously reported procedure (d'Errico et al., 2015). Substrate and enzyme were incubated in 96-well multiscreen 10 kDa cut-off ultracentrifugation plates (Millipore) for 10 min in 50 mM sodium phosphate buffer, pH 6.0. Kinetic parameters for **1** were recently determined and reported by d'Errico et al. (2015). For compound **2** substrate concentrations varied from 0.025 to 10 mM, and the enzyme concentration was 0.0025 mg/mL (0.05 μM). 15 V/V% acetonitrile was used as a co-solvent due to the limited solubility of **2** and its hydrolysis product. Briefly, stock solutions of **2** were prepared by dissolving different amounts of **2** in 30:70 V/V% acetonitrile/water. Aliquots of the homogeneous stock solutions were then mixed with buffer and enzyme solution to give a final concentration of 15 V/V% acetonitrile during the hydrolysis step. After incubation, reactions were stopped by rapid cooling to 4 °C followed by mechanical removal of the enzyme from the solution by ultracentrifugation (4000 × *g*) through a 10 kDa membrane in a precooled centrifuge for 20 min at 4 °C. After ultracentrifugation samples were diluted 10 times with water to obtain relevant concentrations for HPLC analysis. The degree of substrate hydrolysis was determined on the basis of integrated areas of the UV-absorbing alcohols produced within different time intervals as quantified by HPLC (ICS-5000 Dionex system, Thermo Fisher Scientific) using a Luna C18 3 μm column (100 Å, 150 × 4.6 mm, Phenomenex) and UV detection at 210 nm. Elution was carried out with a mixture of acetonitrile/0.01% formic acid solution at pH 3.6 (isocratic, 44:56 V/V) at a flow rate of 0.7 mL/min. The obtained data were fitted to the Michaelis–Menten equation to estimate the values for K_m and k_{cat}.

2.7. Enzymatic hydrolysis of pretreated corn fiber

The enzymatic hydrolysis of pretreated corn fiber (50 mg) was conducted in Eppendorf tubes containing 50 mM succinic acid buffer pH 5.0 in a total volume of 2 mL (2.5% DM in assay). The incubation mixtures contained a base enzyme load of Ultraflo® L (5 g/kg DM) and *Tr*βX (1 g/kg DM) or Cellic® CTec (5 g/kg DM) and *Tr*βX (1 g/kg DM). *Cu*GE and *Tr*GE were supplemented to the incubation mixtures at a level of 1 g/kg DM. The Eppendorf tubes were incubated at 50 °C for 24 h in a Thermomixer (Eppendorf AG) under continuous mixing at 1300 rpm. The reactions were terminated thermally by heating at 100 °C for 10 min in a preheated Thermomixer. Release of C5 sugars (xylose and arabinose), glucose and glucuronic acid was quantified by HPLC (Dionex BioLC system, Thermo Fisher Scientific) using a CarboPac PA1 analytical column (4 × 250 mm, Dionex) combined with a CarboPac PA1 guard column (4 × 50 mm, Dionex). Monosaccharides were separated isocratically with 10 mM potassium hydroxide (xylose, arabinose and glucose) or 101 mM sodium hydroxide and 160 mM sodium acetate (glucuronic acid) at a flow rate of 1 mL/min and detected by a pulsed electrochemical detector in the pulsed amperiometric detection mode. The potential of the electrode was programmed for +0.1 volt (t = 0–0.4 s) to −2.0 volt (t = 0.41–0.42 s) to 0.6 volt (t = 0.43 s) and finally −0.1 volt (t = 0.44–0.50 s), while integrating the resulting signal from t = 0.2–0.4 s. Pure monosaccharides were used as standards.

3. Results and discussion

3.1. Physico-chemical properties of TrGE

Purified *Tr*GE was found to have a molecular mass of 60 kDa as visualized by SDS PAGE (Fig. 2). After treatment with Endoglycosidase H, the molecular weight on SDS PAGE decreased to 55 kDa clearly indicating that the enzyme is *N*-glycosylated, which is also supported by the presence of a predicted *N*-glycosylation site in the protein sequence. However, the Mw is still 9 kDa higher than the predicted molecular weight of 46 kDa of the mature protein. As also reported previously for *Cu*GE (d'Errico et al., 2015), this is in accordance with *Tr*GE being modular with a catalytic core and a family 1 CBM linked together with a serine and threonine rich linker. Such linker regions are known to be prone to *O*-glycosylation, and a clear glycosylation pattern with 162 Da spacing was indeed observed around 56 kDa in the Endoglycosidase H treated enzyme (Fig. S2). *Tr*GE was found to be blocked for N-terminal sequencing as experienced for many extracellular enzymes involved in plant cell wall degradation. The characteristics of *Tr*GE (mature protein and *N*- and *O*-glycosylation pattern) were in agreement with characterization data of the same enzyme previously expressed in *T. reesei*, and were thus found to be the same irrespective of the chosen expression host (*A. oryzae* (this study) versus *T. reesei* (Li et al., 2007)). With a thermal denaturation temperature of 71 °C *Tr*GE was found to have good stability at pH 5.0 ensuring that the enzyme is fully stable during the hydrolysis of pretreated corn fiber. The thermal stability of *Tr*GE is similar to that of *Cu*GE (70 °C, d'Errico et al., 2015). *Tr*GE was found to have the highest sequence homologies to proteins classified as esterases within CE15. Compared to previously identified GEs, *Tr*GE shared highest sequence identity with *Pa*GE1 from the coprophile *Podospora anserina* (Katsimpouras et al., 2014; 59.6% sequence identity), whereas it shared 54.6 and 52.4% sequence identity with *Cu*GE and *Sc*GE, respectively.

Table I
Kinetic parameters for *Cu*GE and *Sc*GE at pH 6.0, 30 °C using synthesized substrates.

	*Cu*GE			*Sc*GE			Reference
Substrate	K_m[mM]	k_{cat} [s^{-1}]	k_{cat}/K_m[$mM^{-1} \times s^{-1}$]	K_m[mM]	k_{cat} [s^{-1}]	k_{cat}/K_m[$mM^{-1} \times s^{-1}$]	
1	4.6 (1.0)	129 (7.6)	28	3.7 (1.2)	118 (9.4)	32	d'Errico et al. (2015)
2[a]	3.4 (0.7)	285 (22)	83	1.4 (0.3)	125 (6.9)	89	This study

Numbers in parentheses are standard deviations.
[a] 15 V/V% acetonitrile was used as a co-solvent for the incubation due to limited solubility of the substrate and product.

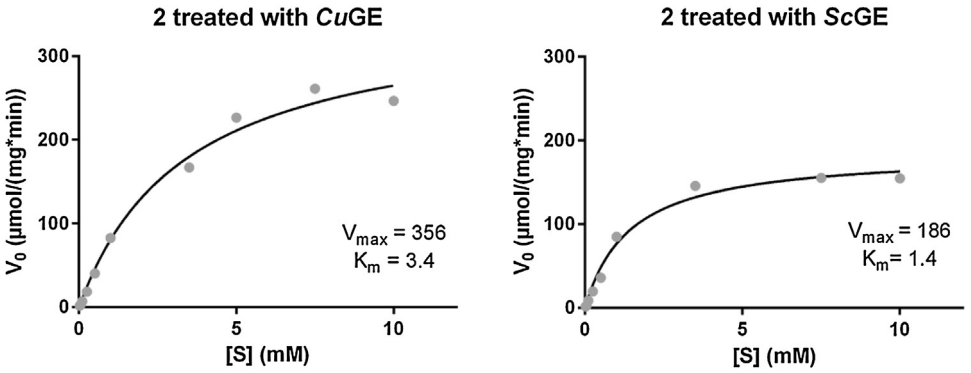


Fig. 3. Degradation of **2** with *Cu*GE and *Sc*GE at pH 6.0 and 30 °C. Experimental data points are shown with circles; best fits to Michaelis–Menten kinetics are shown with full lines.

3.2. Characterization of GEs on model substrates

For the kinetic characterization of GEs, two model compounds were selected. The first compound was the simple α-linked benzyl 4-*O*-methyl-glucuronic acid ester **1** (d'Errico et al., 2015) (Fig. 1B), easily accessible by synthesis in five steps, whereas the second compound was the advanced ester LCC model compound **2**, (Li and Helm, 1995) consisting of a 4-*O*-methyl-glucuronic acid γ-linked to a lignin dimer. The latter model compound comprises a very detailed model of the lignin part of ester LCCs thereby representing the structural complexity of natural glucuronoyl ester LCCs. Both α- and γ-linked glucuronoyl ester LCCs (Fig. 1A) are believed to exist in various lignocellulosic plants, (Balakshin et al., 2011, 2014; Imamura et al., 1994; Li and Helm, 1995; Watanabe 1995), however, while γ-linked glucuronoyl esters have been observed directly in a number of hard- and softwoods by advanced 2D NMR spectroscopy, (Balakshin et al., 2007, 2011; Du et al., 2014; Yuan et al., 2011) evidence for the originally proposed α-linked glucuronoyl esters has only been shown indirectly by DDQ oxidation of model compounds (Imamura et al., 1994).

Kinetic characterization of *Cu*GE and *Sc*GE by means of Michaelis–Menten kinetics on the synthetic α- and γ-linked model substrates **1** and **2** was performed with a previously reported assay (d'Errico et al., 2015) by monitoring formation of UV-active alcohols by HPLC after incubation of GE and model substrate at pH 6.0 at 30 °C (Table I and Fig. 3). pH 6.0 was selected in order to minimize autohydrolysis of the substrate observed at extreme pH values. *Sc*GE and *Tr*GE have been reported to have pH optima at pH 7.0 (Špániková and Biely, 2006) and pH 5.5 (Li et al., 2007), respectively, whereas no pH optimum has been reported for *Cu*GE. With both *Cu*GE and *Sc*GE, binding affinities (K_m) and catalytic efficiencies (k_{cat}/K_m) for **1** and **2** were within the same order of magnitude emphasizing that the simple ester substrate **1** is a good mimic of the more realistic glucuronoyl ester LCC model **2** and may thus find use as a screening substrate in the search for industrially relevant GEs. *Sc*GE was found to have slightly higher binding affinity (lower K_m value) for compound **2** than *Cu*GE, however, more or less similar catalytic efficiencies of *Cu*GE and *Sc*GE were obtained which is in line with our previous observations (d'Errico et al., 2015). Furthermore, from a specificity point of view our results suggest a slight preference of both *Cu*GE and *Sc*GE for glucuronoyl esters of more bulky lignin alcohols i.e., displaying higher binding affinities (lower K_m values) and higher catalytic efficiencies for the more bulky substrate **2** as compared to the less sterically hindered substrate **1** which supports the previously observed preference of GEs for glucuronoyl esters of bulky arylalkyl alcohols (d'Errico et al., 2015). Although this is also in agreement with previous observations by Pokkuluri et al. (2011) reporting the active site of GEs to be situated at the surface of the enzyme and thus enabling access to large substrates, more GEs need to be tested in order to establish firm conclusions on this.

3.3. Activity of GEs on natural corn fiber substrate

*Cu*GE and *Tr*GE were further employed in the enzymatic saccharification of a natural corn fiber substrate in small scale. Together, corn stover (stalks, leaves, husk, cobs) and corn fiber make up the most abundant agricultural residue in the US constituting the lignocellulosic portion of the corn plant left as a by-product from the processing of corn (targeting starch, oil and protein in the corn kernels) (Sánchez and Cardona, 2008; Wyman, 1993, 2001). Corn fiber accounts for around 10% of the corn kernels (Wyman, 1993) and it is the fraction of the corn kernels which is left after removal of starch by wet milling. It consists of the seed coat and the residual endosperm after the starch has been removed for further processing (Sánchez and Cardona, 2008). Corn fiber is mainly composed of glycans (glucan 37.2%, xylan 17.6%, arabinan 11.2%, mannan 3.6%), but significant amounts of lignin (7.8%) as well as protein and smaller amounts of fat are also found (11.0 and 2.5%, respectively) (Wyman, 2001).

Corn fiber was homogenized by milling followed by heat pretreatment (autoclaving, 140 °C) (no washing) to partially disrupt the plant cell wall components and finally subjected to enzymatic hydrolysis using two cellulase/hemicellulase formulations and two different glucuronoyl esterases, *Cu*GE and *Tr*GE (Table II).

Table II
Sugar release from enzymatic hydrolysis of pretreated corn fiber at pH 5.0, 50 °C using GEs on top of different cellulase/hemicellulase preparations.

	Ultraflo® L + *Trβ*X +			Cellic® CTec + *Trβ*X +		
	no GE	*Cu*GE	*Tr*GE	no GE	*Cu*GE	*Tr*GE
Xyl/Ara (g/kg DM)	100.9 (4.7)	110.9 (4.6)	107.5 (13.6)	42.9 (3.2)	51.5 (6.0)	53.5 (9.8)
Glc (g/kg DM)	165.8 (2.1)	172.7 (7.6)	168.1 (14.4)	196.5 (6.0)	205.1 (5.9)	202.6 (14.8)
GlcA[a] (g/kg DM)	4.0 (0.2)	4.5 (0.1)	4.3 (0.4)	n.d.	n.d.	n.d.

Numbers in parentheses are standard deviations; n.d.: not determined.
[a] Total glucuronic acids (glucuronic acid and 4-*O*-methyl glucuronic acid) measured as glucuronic acid equivalents.

The enzymatic hydrolysis was conducted at 50 °C, pH 5.0 containing 50 mg dry weight corn fiber (2.5% DM in assay) in 2 mL incubation vessels. For the incubations of corn fiber pH 5.0 was selected to obtain a slightly acidic environment as found in industrial saccharification processes. Both cellulolytic and hemicellulolytic enzymes would potentially benefit from GE mediated LCC cleavage potentially leading to increased access to cellulose and hemicellulose, respectively, and *Cu*GE and *Tr*GE were thus tested in combination with a commercial β-glucanase/hemicellulase product Ultraflo® L as well as the commercial cellulase product Cellic® CTec. A total of 6.0 g cellulase/hemicellulase (Ultraflo®/*Trβ*X or Cellic® CTec/*Trβ*X) per kg DM was used as a base enzyme load and GEs were supplemented at 1.0 g per kg DM (16.7% GE load compared to the base cellulase/hemicellulase load) for 24 h (cf. Materials and methods). Hydrolyses were terminated thermally (heat inactivation at 100 °C for 10 min) followed by filtration and quantification of C5 sugars (xylose and arabinose), glucose and glucuronic acid by HPLC (Table II). In general, both total sugar and glucose release were found to increase by 5–10% on top of the base cellulase/hemicellulase preparations, whereas release of C5 sugars increased by 10–20% on top of Ultraflo®/*Trβ*X and Cellic® CTec/*Trβ*X when *Cu*GE was employed for the hydrolysis. More or less similar increments of sugar release were observed with *Tr*GE, however, generally with slightly higher standard deviations.

As expected the highest release of C5 sugars was observed with the commercial β-glucanase/hemicellulase preparation Ultraflo®, whereas the highest release of glucose was observed using the commercial cellulase Cellic® CTec.

Similar overall trends of *Cu*GE were observed with HL-NREL-PCS as the substrate, however, large standard deviations prevented firm conclusions to be made using this substrate (Data not shown). (In order to prepare HL-NREL-PCS, NREL-PCS (Schell et al., 2003) received from the U.S. Department of Energy, National Renewable Energy Laboratory (NREL) was separated into liquor and high-solid fractions. The liquor fraction was then recombined with a smaller portion of the high-solid fraction thereby obtaining an increased liquor level as compared to the original NREL-PCS. In assay 5% DM was used). The composition of corn stover is similar to corn fiber, however, the total amount of hemicellulose is generally lower (Corn stover: glucan 40.9%, xylan 21.5%, arabinan 1.8%, galactan 1.0%, lignin 11.0%, protein 8–8.9%, fat 1.3%, ash 7.2% (Wyman, 2001)).

More direct evidence of the action of GEs on natural ester LCCs was observed by improved release of glucuronic acids after treatment of corn fiber with GEs on top of Ultraflo® (Table II). The release of glucuronic acids is probably effected by initial GE-mediated ester LCC cleavage leaving a non-substituted glucuronic or 4-*O*-methyl glucuronic acid substituent on xylan which is then removed by inherent *H. insolens* alpha-glucuronidases in the commercial Ultraflo® hemicellulase blend used. The presence of alpha-glucuronidase activity in Ultraflo® was confirmed by proteomics by mass spectrometry. Release of glucuronic acids (glucuronic acid and 4-*O*-methyl-glucuronic acid; measured as glucuronic acid equivalents) was only quantified using Ultraflo®/*Trβ*X as the base enzyme preparation and was found to increase approximately 10% when *Cu*GE was used.

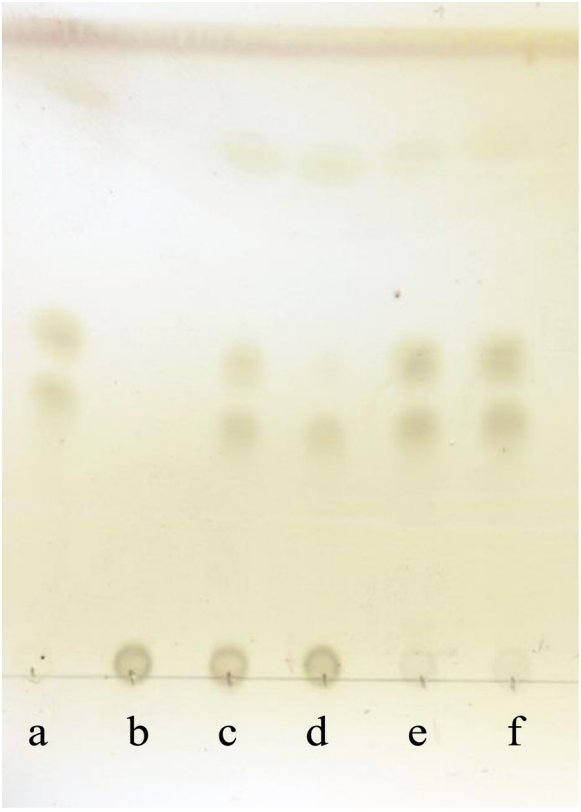


Fig. 4. Semi-quantitative detection of GE activity in Ultraflo® and Cellic® CTec by TLC analysis. (a) Benzyl D-glucuronate (isomeric mixture, 8 mM), (b) D-glucuronic acid (8 mM), (c–f) Benzyl D-glucuronate (8 mM) treated with (c) *Cu*GE or (d) *Tr*GE (both at 0.025 mg/mL) or (e) Ultraflo® or (f) Cellic® CTec (both at 0.125 mg/mL) for 2 h, 35 °C, pH 6.

The commercial cellulase/hemicellulase products Ultraflo® and Cellic® CTec were found only to possess very little GE activity by themselves. In a simple TLC assay, GE activities of Ultraflo® and Cellic® CTec were assessed by measuring formation of glucuronic acid upon incubation of the cellulase/hemicellulase products with benzyl D-glucuronate. Even when dosed five times higher than *Cu*GE and *Tr*GE, only trace GE activity of Ultraflo® and Cellic® CTec was observed under conditions where *Cu*GE and *Tr*GE showed >50% conversion of benzyl D-glucuronate (Fig. 4). (The presence of CE15 GEs in Ultraflo® and Cellic® CTec was also confirmed by proteomics). The lack of noticeable GE activity in the commercial cellulase/hemicellulase products concludes that the observed ester LCC cleavage is effected exclusively by the added GE monocomponents rather than a synergistic action of multiple GEs.

3.4. Industrial applicability of GEs

In relation to industrial bioethanol production, pretreatment of lignocellulosic biomass is necessary in order to disrupt the

recalcitrant lignin–carbohydrate matrix thereby improving the overall efficiency of the subsequent enzymatic saccharification step. The main goals of biomass pretreatment are to remove and/or break down lignin and increase the enzyme accessibility of cellulose and hemicellulose, while minimizing formation of enzyme inhibitors (Alvira et al., 2010; Xu and Huang, 2014). Inclusion of GEs in the enzymatic digestion of lignocellulosic biomass seems particularly beneficial in combination with less severe physical or physico-chemical pretreatment methods having less extreme pH values such as thermal or hydrothermal pretreatments (liquid hot water, steam explosion or heat pretreatments). These methods only result in partial cleavage of (ester) LCCs as opposed to the more extreme chemical pretreatment methods such as the alkaline lime, sodium hydroxide or ammonia pretreatments which facilitate saponification of ester LCCs (as well as acetate and lignin esters) or the acidic dilute acid and organosolv pretreatment methods using sulfuric acid or other acids which result in ester LCC cleavage, but on the other hand may also lead to increased lignin and lignin-carbohydrate condensation at elevated temperatures (i.e., ether bond formation) (Xu and Huang, 2014). By degrading ester LCCs, GEs seem capable of improving the separation of lignin and carbohydrates thus increasing the enzyme accessibility of cellulose and hemicellulose. Furthermore, no sugar- or lignin-derived enzyme inhibitors are generated by GE treatment thus making GEs interesting in connection with future sustainable physical or physico-chemical pretreatment methods.

In the present study a relatively high dosage of 16.7% GE as compared to the base cellulase/hemicellulase load was necessary to demonstrate significant effects on C5 sugar, glucose and glucuronic acid release. Based on this it seems likely that ester LCC bonds are not easily accessible to enzymatic attack due to the intimate association of lignin, hemicellulose and cellulose as both covalent and non-covalent structures. As a result, an efficient breakdown of lignocellulosic material can only be achieved by the concerted action of multiple enzymes, and although currently dosed too high to meet the overall economic requirements for industrial scale saccharification, (more efficient) GEs may play an important role in connection with less severe, sustainable pretreatment methods thereby lowering the needs for energy and chemicals in the processing of lignocellulosics.

4. Conclusion

By means of model compound studies using GEs from *C. unicolor, S. commune* and *T. reesei* the present study provides insights into the substrate specificity of fungal GEs supporting the previously observed preference of GEs for glucuronoyl esters of bulky arylalkyl alcohols such as ester LCCs. Furthermore, we report the first examples of activity of GEs on natural lignocellulosic biomass confirming the initially proposed activity of GEs on ester LCCs and emphasizing the potential of this class of carbohydrate esterases as auxiliary enzymes in the saccharification of lignocellulosic biomass in particular in connection with the transition toward more sustainable pretreatment methods.

Conflict of interest

The authors JB, HD, KBRMK, NS and RNM are employees of Novozymes A/S.

Supporting information

Additional supporting information (additional figures) may be found in the online version of this article at the publisher's web-site.

Acknowledgements

The authors thank Prof. Richard F. Helm (Department of Biochemistry, Virginia Tech, Virginia) for a generous sample of the glucuronoyl esterase substrate 2. This work was funded by the Danish Council for Strategic Research (SET4Future project).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbiotec.2015.12.024

References

Aita, G.A., Salvi, D.A., Walker, M.S., 2011. Enzyme hydrolysis and ethanol fermentation of dilute ammonia pretreated energy cane. Bioresour. Technol. 102, 4444–4448.

Alvira, P., Tomás-Pejó, E., Ballesteros, M., Negro, M.J., 2010. Pretreatment technologies for an efficient bioethanol production process based on enzyme hydrolysis: a review. Bioresour. Technol. 101, 4851–4861.

Balakshin, M.Y., Capanema, E.A., Chang, H.-m., 2007. MWL fraction with a high concentration of lignin–carbohydrate linkages: isolation and 2D NMR spectroscopic analysis. Holzforschung 61, 1–7.

Balakshin, M.Y., Capanema, E., Berlin, A., 2014. Isolation and analysis of lignin–carbohydrate complexes preparations with traditional and advanced methods: a review. Stud. Nat. Prod. Chem. 42, 83–115.

Balakshin, M., Capanema, E., Gracz, H., Chang, H.-m., Jameel, H., 2011. Quantification of lignin–carbohydrate linkages with high-resolution NMR spectroscopy. Planta 233, 1097–1110.

Biely, P., Malovíková, A., Uhliariková, I., Li, X.-L., Wong, D.W.S., 2015. Glucuronoyl esterases are active on the polymeric substrate methyl esterified glucuronoxylan. FEBS Lett. 589, 2334–2339.

d'Errico, C., Jørgensen, J.O., Krogh, K.B.R.M., Spodsberg, N., Madsen, R., Monrad, R.N., 2015. Enzymatic degradation of lignin–carbohydrate complexes (LCCs): model studies using a fungal glucuronoyl esterase from *Cerrena unicolor*. Biotechnol. Bioeng. 112, 914–922.

Du, X., Pérez-Boada, M., Fernández, C., Rencoret, J., del Rió, J.C., Jiménez-Barbero, J., Li, J., Gutiérrez, A., Martínez, A.T., 2014. Analysis of lignin–carbohydrate and lignin–lignin linkages after hydrolase treatment of xylan–lignin, glucomannan–lignin and glucan–lignin complexes from spruce wood. Planta 239, 1079–1090.

Imamura, T., Watanabe, T., Kuwahara, M., Koshijima, T., 1994. Ester linkages between lignin and glucuronic acid in lignin–carbohydrate complexes from *Fagus crenata*. Phytochemistry 37, 1165–1173.

Humbird, D., Davis, R., Tao, C., Kinchin, C., Hsu, D., Aden, A., Schoen, P., Lukas, J., Olthof, B., Worley, M., Sexton, D., Dudgeon, D., 2011. Process Design and Economics for Biochemical Conversion of Lignocellulosic Biomass to Ethanol. National Renewable Energy Laboratory, Golden, Colorado.

Katsimpouras, C., Bénarouche, A., Navarro, D., Karpusas, M., Dimarogona, M., Berrin, J.-G., Christakopoulos, P., Topakas, E., 2014. Enzymatic synthesis of model substrates recognized by glucuronoyl esterases from *Podospora anserina* and *Myceliophthora thermophila*. Appl. Microbiol. Biotechnol. 98, 5507–5516.

Kwiatkowski, J.R., McAloon, A.J., Taylor, F., Johnston, D.B., 2006. Modeling the process and costs of fuel ethanol production by the corn dry-grind process. Ind. Crops Prod. 23, 288–296.

Li, K., Helm, R.F., 1995. Synthesis and rearrangement reactions of ester-linked lignin–carbohydrate model compounds. J. Agric. Food Chem. 43, 2098–2103.

Li, X.-L., Špániková, S., de Vries, R.P., Biely, P., 2007. Identification of genes encoding microbial glucuronoyl esterases. FEBS Lett. 581, 4029–4035.

Pokkuluri, P.R., Duke, N.E.C., Wood, S.J., Cotta, M.A., Li, X.-L., Biely, P., Schiffer, M., 2011. Structure of the catalytic domain of glucuronoyl esterase Cip2 from *Hypocrea jecorina*. Proteins 79, 2588–2592.

Rasmussen, L.E., Sørensen, H.R., Vind, J., Viksø-Nielsen, A., 2006. Mode of action and properties of the β-xylosidases from *Talaromyces emersonii* and *Trichoderma reesei*. Biotechnol. Bioeng. 94, 869–876.

Sánchez, Ó.J., Cardona, C.A., 2008. Trends in biotechnological production of fuel ethanol from different feedstocks. Bioresour. Technol. 99, 5270–5295.

Schell, D.J., Farmer, J., Newman, M., McMillan, J.D., 2003. Dilute-sulfuric acid pretreatment of corn stover in pilot-scale reactor. In: Davison, B.H., Lee, J.W., Finkelstein, M., McMillan, J.D. (Eds.), Biotechnology for Fuels and Chemicals. Humana Press, pp. 69–85.

Špániková, S., Biely, P., 2006. Glucuronoyl esterase–novel carbohydrate esterase produced by *Schizophyllum commune*. FEBS Lett. 580, 4597–4601.

Špániková, S., Poláková, M., Joniak, D., Hirsch, J., Biely, P., 2007. Synthetic esters recognized by glucuronoyl esterase from *Schizophyllum commune*. Arch. Microbiol. 188, 185–189.

Tsai, A.Y.-L., Canam, T., Gorzsás, A., Mellerowicz, E.J., Campbell, M.M., Master, E.R., 2012. Constitutive expression of a fungal glucuronoyl esterase in *Arabidopsis* reveals altered cell wall composition and structure. Plant Biotechnol. J. 10, 1077–1087.

Watanabe, T., 1995. Important properties of lignin–carbohydrates complexes (LCCs) in environmentally safe paper making. Trends Glycosci. Glycotechnol. 7, 57–68.
Wyman, C.E., 1993. Cellulosic biomass conversion technology and its application to ethanol production from corn. Fuel Reformulation, 67–74.
Wyman, C.E., 2001. Applications of corn stover and fiber. In: Johnson, L.A., White, P. (Eds.), Corn Chemistry and Technology. American Association of Cereal Chemists, St. Paul, Minnesota, Chapter 20.
Xu, Z., Huang, F., 2014. Pretreatment methods for bioethanol production. Appl. Biochem. Biotechnol. 174, 43–62.
Yuan, T.-Q., Sun, S.-N., Xu, F., Sun, R.-C., 2011. Characterization of lignin structures and lignin–carbohydrate complex (LCC) linkages by quantitative ^{13}C and 2D HSQC NMR spectroscopy. J. Agric. Food Chem. 59, 10604–10614.