CHALMERS

Carbohydrate Synthesis

A synthesis of functionalized glucuronic acids as potential model substrates

Bachelor of Science Thesis

MONA SVANTESSON

Department of Chemistry and Chemical Engineering
Division of Organic Chemistry
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2015

Carbohydrate Synthesis
A synthesis of functionalized glucuronic acids as potential model substrates

Bachelor of Science Thesis

Mona Svantesson

Supervisor: Filip Nylander
Examiner: Gunnar Westman

Department of Chemistry and Chemical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2015

Carbohydrate synthesis – A synthesis of functionalized glucuronic acids as potential model substrates

MONA SVANTESSON

Department of Chemistry and Chemical Engineering
Chalmers Unversity of Technology
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31-772 10 00

Abstract

Lignocellulosic biomass is considered a potential sustainable resource in the production of biofuel and other petroleum derived products. In order to make the process cost-efficient for large scale production technological development is crucial. One important process step that needs improvement is hydrolysis which is used for separation and degradation of the lignocellulosic components. Enzymatic hydrolysis which include degrading enzymes from bacteria and fungi is considered to have potential to improve the hydrolysis process step and therefore more development in this area is needed. In addition lignin-carbohydrate complexes (LCCs) in lignocellulosic biomass are believed to impede the process and therefore more knowledge about their structure is needed.

Model substrates are used in enzymatic studies. The aim of this project is to synthesize three functionalized glucuronic acids by two different reaction routes using various protective groups. These molecules have the potential to be used in enzymatic studies and LCC structure studies.

Molecule **C** was successfully synthesized with good conversion and molecules **A** and **B** needs further work

Acknowledgements

First of all I would like to thank my examiner Gunnar Westman for having me at the division of Organic chemistry and for all support.

I would like to give special thanks to my supervisor Filip Nylander, PhD student at the divison of Organic chemistry who has guided me throughout the project and helped me move forward.

I would also like to thank the remaining members of the division for a good time.

Last but not least I would like to thank my fellow thesis workers at the division of organic chemistry for their enjoyable company.

Table of Contents

1 Introduction

Due to increasing petroleum prices and growing environmental concern, a global progress towards a generation of sustainable sources of fuel and other significant chemical products that are derived from petroleum has been developed in recent years (1). Petroleum products can be derived from carbohydrates from organic sources such as plant material.

Non-food crops, wood and forest residues consists of large carbohydrate macromolecules, i.e. polysaccharides, which include cellulose and hemicellulose as well as lignin. Polysaccharides consist of sugar monomers that can be used in biofuel production and other sugar derived products. Lignin is a complex macromolecule composed of aromatic alcohols and does not contain sugars but has potential in other productions of petroleum derived chemicals.

Biofuels can be derived from food crops, however although these are renewable sources it has limitations. The main disadvantage is that the raw material is also used for food which causes competition between food- and fuel industry (2). This problem has generated the so called second generation biofuels (2) where lignocellulosic biomass has been recognized as the new source material for biofuels and other valuable chemical products(1).

In order to convert lignocellulosics to biofuels and other valuable products it is necessary to first separate the components and break down the polysaccharides into their corresponding sugar monomers. This process is done by hydrolysis, which is also used in cleaving lignin-carbohydrate complexes, LCC. LCCs are covalent bonds between lignin and the polysaccharides of plant material. These are difficult to break and are believed to impede the hydrolysis in lignocellulosic bioconversion processes.

Technical improvement of lignocellulosic bioconversion is necessary in order to make the process more cost-effective. One important improvement is to achieve more efficient hydrolysis. One type of hydrolysis that is considered to have good potential is enzymatic hydrolysis which involve enzymes from fungi and bacteria that naturally produce enzymes that naturally degrade bonds in plant biomass (1).

Model substrates are often used in studies concerning enzymatic hydrolysis in order to investigate in what way the structure of carbohydrates influence the efficiency of enzymes. Structures of LCCs are not entirely understood (3) and therefore more knowledge about them is also necessary to develop the bioconversion processes. Model substrates that mimic the believed naturally occurring LCC bonds in plant material is of importance in structure studies of this complex.

Synthesis of carbohydrates as model substrates is possible by using various reaction routes and protective groups along the way. Two important protective groups used in this project are benzylidene acetal and isopropylidene acetal, which is also called acetonide. Both belong to the type of protective groups called cyclic acetals. Acetonide is not thermodynamically

stable and is usually formed under kinetic control, while benzylidene acetal is thermodynamically stable.

1.1 Aim

Because of the complex nature of polysaccharides and lignin in lignocellulosic material, simple model substrates that structurally mimic the naturally occurring molecules in polysaccharides are needed in enzymatic and LCC studies.

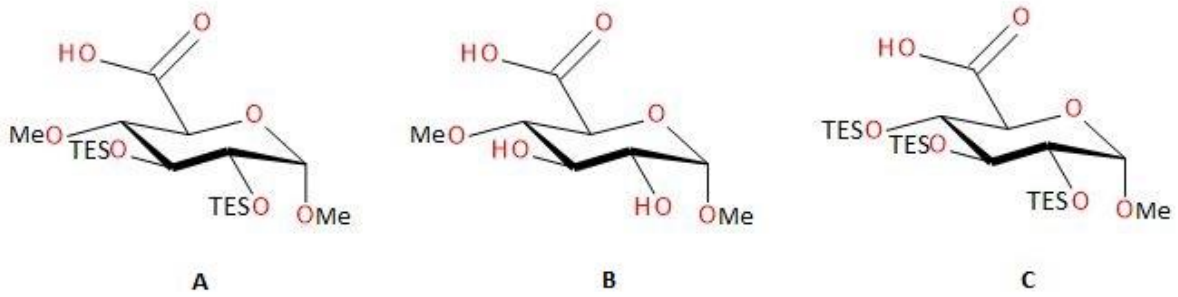

Figure 1. The aim of this project is to reach three target molecules.

The aim of this project is to synthesize carbohydrate molecules with high purity suitable as model substrates. Three target molecules **A**, **B**, and **C** in the form of methylated and functionalized glucuronic acids are to be synthesised in various organic reaction steps involving a number of protective groups. These target molecules are depicted in Figure 1. Selectivity and purity is important and characterization of product molecules from each reaction is made by Nuclear Magnetic Resonance (NMR) spectrometry.

2 Theory and literature review

2.1 Petroleum products from renewable resources

Petroleum is today mostly used in production of fossil fuels, plastics, and various organic chemicals. The majority of petroleum is used as an energy resource for fossil fuels, however, petroleum is not considered sustainable because it is responsible of negative environmental effects such as global warming. In addition petroleum is not renewable and crude oil reserves are declining (2). Increase of petroleum prices and growing environmental concern are the driving forces towards alternative and sustainable resources of fuel and other petroleum products (1). Since petroleum are made up of hydrogen and carbon, similar products can be derived from carbohydrates from organic sources which include plant material. Biofuels derived from plant material and organic residues could help both reduce oil dependence and carbondioxide emissions. The first generation of biofuels involve the conversion of food crops, such as starch, sugarcane juice, animal fats, and vegetable oils, into ethanol and other

biofuels. Although the first generation of biofuels are renewable it has its limitations. The main disadvantage is that the raw material is also used for food which causes competition between food- and fuel industry, thus increasing food prices (2). First generation biofuels can also have a negative environmental impact on various areas such as biodiversity and land use.

This problem has generated the so called second generation biofuels (2) where lignocellulosic biomass has been recognized as a new source material for biofuels and other valuable chemical products (1). Lignocellulosic material is cheap and abundant in the non-food crops from agricultural waste as well as in trees and forest residues. One major advantage is that lignocellulosic biomass does not compete with food crops(2). In order to convert lignocellulosics to biofuels and other value-added products it is necessary to first separate the components and break down the polysaccharides into their corresponding sugar monomers. This process is done by hydrolysis, which is also used in cleaving lignin-carbohydrate complexes, LCCs.

Currently the conversion of lignocellulosic biomass into bio products is not efficient enough to be able to handle billion tons of biomass every year which is needed for the production of biofuels (1). The process is not yet cost-effective enough to compete with the large petroleum based processes and a lot of research and development is still necessary (1).

Significant biofuels and energy products that are being considered in lignocellulosic bioconversion are ethanol, methane and hydrogen. Sources that could be used are wood and agricultural residues (non-food crops) such as wheat straw, rice straw, corn cob, sunflower stalks etc. Ethanol represents the largest market of biofuels (1).

2.2 Components of plant material

Cellulose, which is the main component in the cell wall, is an unsubstituted polysaccharide composed of β-(1,4)-D-glucose monomers (3). Hemicelluloses are heterogeneous polysaccharides that comprise xylans, glucomannans, galactoglucomannans, and xyloglucans (4). Lignin is a large complex phenylpropanoid polymer composed of three main units; p-courmaryl, conferyl, and sinapyl alcohol (3) which can be seen in Figure 2 and occur in all plant material (5). Pectins exist in the primary wall of plant cells and is low in abundance in lignocellulosic biomass (4) and hence not significant compared to the other three building blocks.

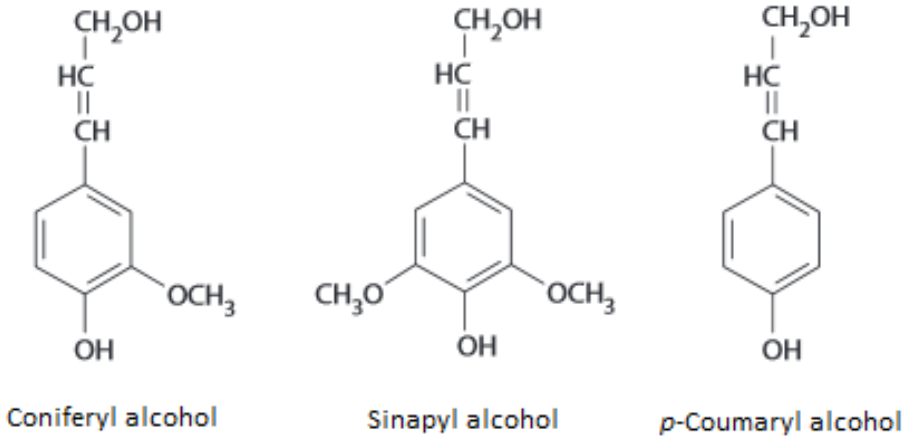

Figure 2. Main monomers of lignin

The main type of hemicelluloses in plant material are xylans.

In hardwood the most abundant xylans are glucuronoxylans which have a backbone of xylose monomers. The xylose backbone can be substituted with acetyl groups and 4-O-methylglucuronic acid side groups. Arabinoglucuronoxylans are main hemicelluloses in non-woody plants such as agricultural crops, and exist as a minor part in softwoods as well. Arabinoglucuronoxylans have a xylose backbone similar to that of glucuronoxylans with glucuronic acid and arabinose substituents. Arabinoxylans are the main hemicellulose in cereal grain and consist of a xylose backbone with arabinose and glucuronic acid substituents (6).

Apart from xylans, galactoglucomannans are the most significant hemicelluloses in plant material. Galactoglucomannan are the major parts of softwoods and has a backbone consisting of glucose and mannose monomers with acetyl groups and galactose substituents. Glucomannans are most common in softwoods and exist in some extent hardwoods and are composed of a backbone similar to that of galactoglucomannans. Xyloglucans are a significant hemicelluloses in hardwoods and occur to some extent in grasses. Their structure is composed of a glucose backbone substituted by xylose in high proportion. In addition arabinose and galactose can be linked to the xylose substituents (6).

2.3 Second generation ethanol

Most of the raw material for ethanol production is sugar and starch but the most dominating form of sugar in nature are cellulose and hemicellulose (1). The process of converting lignocellulosics into ethanol and other valuable bio products is more difficult than conversion of food crops. Prior to the bioconversion of lignocellulosic material into ethanol the

polysaccharides, cellulose and hemicellulose, must be converted into their monomeric sugar components. The general steps of the process are:

- Pre-treatment of the lignocellulosic biomass to separate cellulose, hemicellulose and lignin.
- Enzymatic hydrolysis of cellulose and hemicellulose into fermentable sugar
- Fermentation of sugars into product
- Recovery and refining of product (7)

During pre-treatment hemicellulose and lignin are separated from the cellulose and then a pre-processing occurs in which hemicellulose and lignin breaks down into smaller fractions which can easily be separated and extracted and further processed. Hemicellulose is converted into xylose which could be converted into ethanol(2) however the conversion of pentoses into ethanol is a difficult process where more research in the area is necessary, and in the meantime xylose is used in the production of other products (8). Lignin is a residue in ethanol production since it does not contain any sugar monomers and could potentially be used to produce high value chemicals or be used as an energy source (1). The cellulose is decomposed into glucose by enzymatic hydrolysis and then processed into ethanol (2).

There are various types of pre-treatment methods which can be divided as mechanical (milling, grinding), physiochemical (e.g. liquid hot water), chemical (alkali, acid, organic solvents, oxidizing agents), and biological (enzyme) processes (7).

Pre-treatment is important to reduce the crystallinity of cellulose and reduce the polymerization of cellulose and the lignin-hemicellulose matrix that embeds the cellulose, thus opening up the structure, creating more surface area for the enzyme to attack during the hydrolysis (2). Pre-treatment also simplifies separation of lignin from the biomass and increases fermentation yields (7).

2.4 Additional bio products from lignocellulosic material

Apart from biofuels, lignocellulosic biomass is also a promising resource in the production of organic chemicals. Ethylene, propylene, benzene, toluene and xylene are five basic organic chemicals that are used to synthesize other chemicals and these could be produced from biomass. Ethylene and propylene can be produced from ethanol which can be produced from the sugars in cellulose and hemicellulose, while the aromatic compounds benzene, toluene and xylene could be derived from lignin. Lignin is also used to produce lignosulfonates which are used in the production of vanillin. Xylose derived from hemicelluloses could be used to produce xylitol, furfural and furfuryl alcohol. Xylitol is used in various applications including its use as a sweetener and furfural is used in the production of furfural phenol plastics, varnishes and pesticides (1).

2.5 Lignin-Carbohydrate Complexes

Covalent bonds connecting lignin with cellulose and hemicellulose in lignocellulosic material are named lignin-carbohydrate complexes, LCCs (8). Nearly all wood have its lignin covalently bond to polysaccharides, mainly hemicelluloses (9). There are different types of

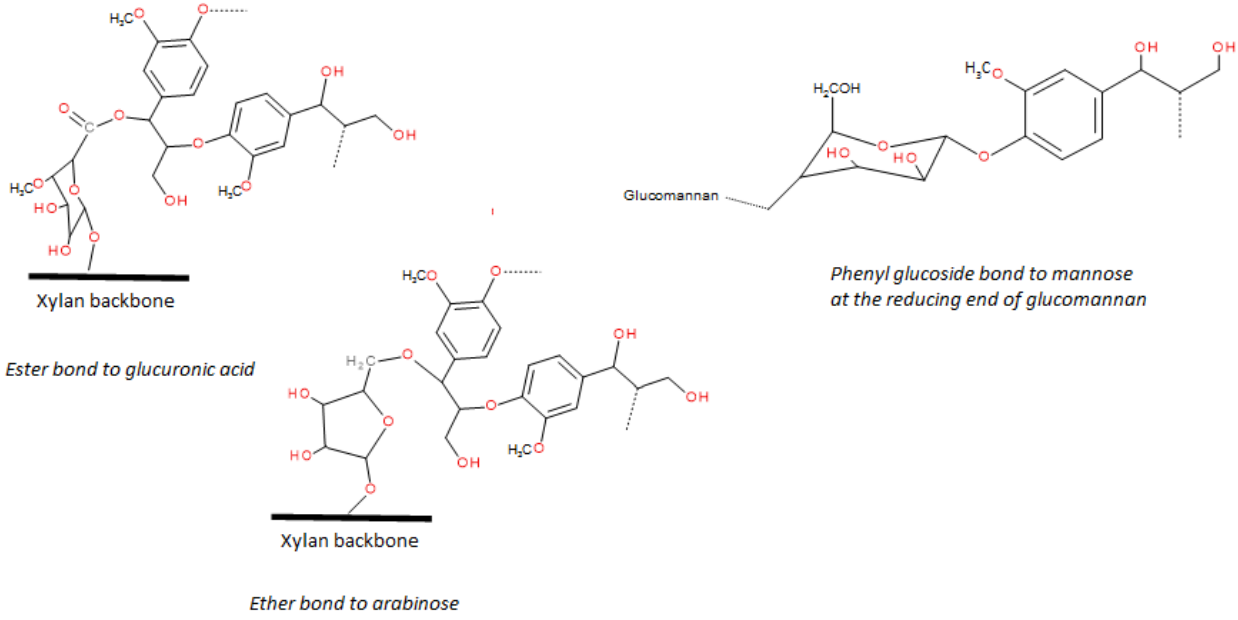

Figure 3. Examples of bonds between lignin and carbohydrates (LCC). The ester and ether bonds are connected to substituents of different types of xylan and the phenyl glucoside bond is linked to glucomannan. Adapted from (8).

LCCs. The exact structure of different variations is not verified, however, ether- ester- and phenyl glucoside bonds are believed to be the main types in wood (3). Examples of these are depicted in Figure 3. Quantification experiments of LCCs have indicated that the main LCC bonds in wood are phenyl glycoside, benzyl ether and ester types, concluding that softwood contains greater quantity of benzyl ether, but lower quantity of phenyl glucoside and ester LCCs than hardwood (9).

The formation and existence of these LCC variations is not definitely proven. The ester and ether bonds are believed to exist with high probability and possible formation mechanisms for them have been identified (8). One type of LCC that has been demonstrated is the ester bond between 4-O-methyl-D-glucuronic acid of glucuronoxylan and the aromatic alcohols of lignin (10). The formation mechanism of phenyl glucoside bond, however, is unclear (8).

The occurrence of LCCs is believed to impede the enzymatic hydrolysis of the lignocellulosic material due to steric hindrance (3). Enzymes that are able to break these bonds between

lignin and hemicellulose are important to make the bioconversion of lignocellulosic material into valuable bio products more efficient (10).

LCCs complicate separation and isolation of lignin and carbohydrates from lignocellulosic material, and additionally they hinder separation of the wood components in biomass conversion processes. Because of these problems it is important to increase our knowledge of the LCC structures in order to improve processes involving biomass (9). .

2.6 Enzymatic hydrolysis

Conversion of the lignocellulosic components into their respective fermentable sugar monomers by enzymatic hydrolysis is a necessary step before initiating the formation of products from lignocellulosic biomass. A range of microorganisms that include bacteria and fungi have been identified to have potential to break down lignocellulosics into their sugar monomers (1).

Cellulases, which are the enzymes that hydrolyse cellulose, are subdivided into three main groups; endoglucanases, exoglucanases (cellobiohydrolases), and β-glucosidases. Endoglucanases cleave internal bonds of the cellulose chain while exoglucanases recognize the chain ends. β-glucosidases recognize low molecular weight polysaccharides of cellulose (e.g cellobiose, cellotriose etc.) (1)

There are numerous enzymes that degrade hemicelluloses into its monomers. Enzymes responsible for example xylan degradation are called xylanases (1).

Bacteria and fungi that produces desirable amounts of endoglucanase, exoglucanase and β-glucosidases could have the potential to effectively degrade lignocellulose. Production of cellulolytic enzymes from various microorganisms has demonstrated that they in general produce none or too little β-glucosidase. β-glucosidase is needed to quickly break down cellobiose which otherwise inhibit the activity of exo- and endoglucanase (1).

Enzymes that has been found to be useful in enzymatic biomass hydrolysis are glucuronoyl esterases (GEs) from wood-decay fungus. GEs are capable of hydrolysizing alkyl and arylalkyl esters of D-glucuronic acid and methyl-D -glucuronic acid (10) and hence have the potential to break the ester LCC between lignin and 4-O-methyl-D-glucuronic acid of hemicelluloses. Studies on synthetic substrates imitating the naturally occurring ester bond in plant material have demonstrated that GEs recognize the glucuronic acid part of the ester bond rather than the lignin alcohol part. Because of the complex and heterogeneous nature of lignocellulosic material the function of GEs on natural substrates have not been demonstrated,,hence the use of model compounds (3).

2.7 Protective groups in carbohydrate synthesis

Modified monosaccharides are often needed. When synthesizing such molecules there are usually two obstacle that need to be overcome: finding a suitable chemical reaction that gives the desired modification, and making sure that modification from this reaction occurs at the desired site of the molecule (11). A common way to ensure a chemical reaction to attack the correct site is to temporarily block other sites that could become involved in the reaction. This is done by using so called protective groups (12).

Hydroxyl groups are abundant in carbohydrates and needs to be protected during various types of reactions such as oxidation or dehydration reactions (12). Ester and ethers are used to protect reactive hydroxyl groups. Esters could trigger chemical reactions, while ethers are less reactive protecting groups. Both reduce the polarity of the molecule and makes it soluble in organic solvents (11).

Ethers are one of the most used type of protective groups (12), however, methyl ethers are stable groups which makes them difficult to remove and are therefore not optimal as a protecting group. Nonetheless it has had great value in structure determination of carbohydrates (11). Silyl ethers were at first used in analytical chemistry to modify carbohydrates into more volatile and less polar compounds. Now they are common in the protecting of hydroxyl groups. Silyl ether groups are stable and endure most common chemical modifications (11), but are nonetheless easily removed with acid- or fluoride ion-catalysed hydrolysis (12).

Cyclic acetals are used in the protection of diols. Isopropylidene acetal, or so called acetonide, and benzylidene acetal are two important cyclic acetals that are mostly used in the protection of 1,2- and 1,3-diols. The classical formation of acetals are done under thermodynamically conditions with either an aldehyde or with ketone such as acetone. Alternative formation exists where reactions are carried out either under termodynamical conditions or under kinetic control where an acetonide is most commonly done under acidic conditions using an dimethoxy acetal such as 2,2-dimethoxypropane or by using 2-methoxypropene. 2-methoxypropene is the most reactive reagent (11) and are commonly used in the formation of acetonide in DMF under kinetic conditions (13).

Bensylidene acetals are normally formed under termodymanic conditions with benzaldehyde and zinc chloride. They can also be formed under kinetic conditions with dimethoxytoluene in DMF with an acid catalyst such as p-toluenesulfonic acid (13) or with bromotoluene (14).

Acetal cleavage is done by acid hydrolysis. Removal of acetonides are done by aqueous acid, in particular triflouroacetic acid, TFA (11). Benzylidene acetals are removed by acid hydrolysis, hydrogenolysis, or Birch reduction. Removal methods of benzylidene acetals can either result in the original diol, or in a functional group such as a benzyl ether (-OBn) (11).

Isopropylidene acetal and benzylidene acetal are both used to protect the same type of functional groups but differ in site selectivity and deprotection methods, which makes them complementary to each other (14).

3 Experimental

3.1 General methods and materials

The first synthesis route was used in order to reach target molecule **C** which is depicted in Figure 4. An alternative route, as depicted in Figure 5, was begun in order to reach target molecules **A** and **B**.

Commercially available reagents and solvents were used unless stated otherwise and MilliQ-water was used as solvent during reactions. NaOH solution (1.49 µM) used was commercially available NaOH dissolved in distilled water and HCl solution (0.05 M) used in the alternative synthesis route was made by diluting commercially available 1M HCl solution in distilled water. All air and water sensitive reactions were performed under nitrogen atmosphere. An ice bath was used during reactions at 0 °C and oil bath was used during reactions at temperatures above room temperature. Thin layer chromatography (TLC) was performed on silica sheets pre-coated with silica gel (Merck, Silica Gel 60 F_{254}) and TLC results were visualized with a methanolic acid spray reagent (AcOH, H_2SO_4, PAA) prior to heat. ^{1}H and ^{13}C NMR spectra were acquired on a Varian 400 MHz spectrometer. Chemical shifts are reported as d values (ppm) and coupling constants J are given in Hz.

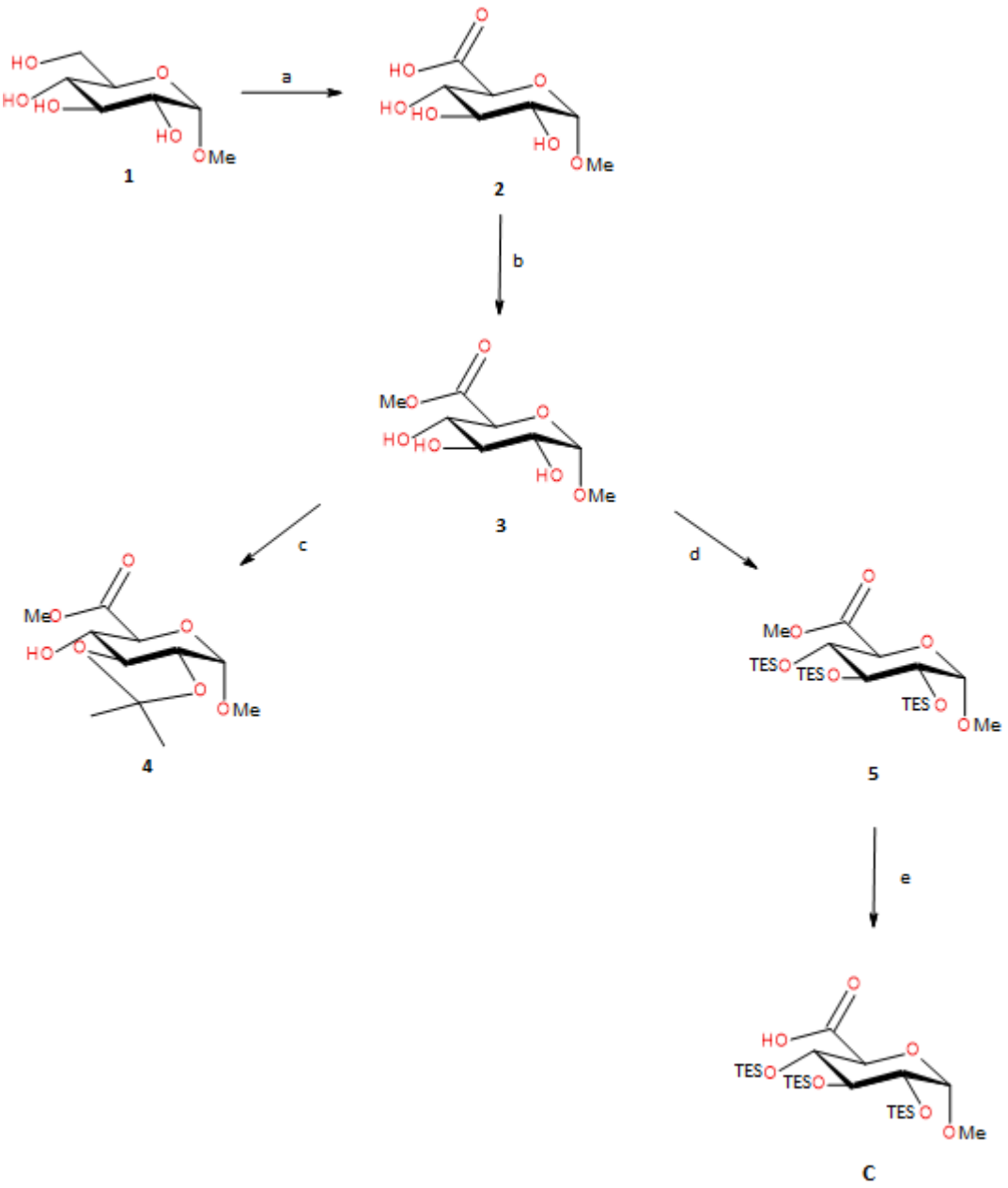

Figure 4. Reagents and conditions: (a) TEMPO, BAIB, H_2O/AcCN, 0°C, 4h; (b) $TMSCHN_2$, MeOH/DCM, rt, 2h; (c) HCl(MeOH), 2-methoxypropene, DMF, rt, 3h; (d) TESCl, Imidazole, DMF, rt, 1h; (e) NaOH, THF, rt/0°C, 58h.

3.1.1 Methyl-α-D-glucopyranosiduronic acid (2)

Methyl-α-D-glucopyranoside **1** (5.017 g, 25.75 mmol) was dissolved in H_2O (13.5 ml) and acetonitrile (41.5 ml) was subsequently added to the solution. While stirred at 0 °C, TEMPO (0.670 g, 4.288 mmol) was added to the solution prior to the addition of BAIB (17.29 g, 53.66 mmol). The reaction was left stirring at 0 °C until TLC (DCM-MeOH 5:1) indicated complete consumption of the reactant (4h). The solution was washed with diethyl ether and evaporated, and then freeze-dried at 0,1 mbar and temperature at -20°C to -5°C with

condenser at -47°C resulting in a yellow, highly viscous product. ^{1}H NMR(400 MHz, D_2O): d = 4.87 (d, 1H, J=3.70, H-1), 4.15 (d, 1H, J=10.02, H-5), 3.70 (dd, 1H, J=8.98, 9.70, H-3), 3.62 (dd, 1H, J=3.70, 9.70, H-2), 3.57 (dd, 1H, J=10.02, 8.98, H-4), 3.44 (s, 3H, $CH_{3\text{-}OMe}$); ^{13}C NMR (400 MHz, D_2O): d = 176.58 ($C{=}O_{COOH}$), 99.49 (C-1), 72.55 (C-5), 71.30 (C-3), 70.71 (C-2), 70.53 (C-4), 55.43 ($CH_{3\text{-}OMe}$).

3.1.2 Methyl (1-*O*-methyl-D-glucopyranose) uronate (3)

Methyl-α-D-glucopyranosiduronic acid **2** (1.00 g, 4.80 mmol) was dissolved in MeOH (23,5ml) and DCM (23,5 ml) was subsequently added. While stirred under nitrogen atmosphere in room temperature, $TMSCHN_2$ (7.50 ml, 15.0 mmol) was slowly added dropwise until the solution turned yellow. The solution was left stirring until TLC-analysis (Aceton-MeOH 1:1) indicated complete consumption of the reactant (~ 2h). The reaction was then quenched with AcOH (0,40 ml) until the solution turned back to colourless and then the solution was diluted with toluene (54 ml). The solution was evaporated and then co-evaporated with hexane (45 ml) yielding the product (0.927 g, 4.17 mmol, 86.85%). ^{1}H NMR(400 MHz, $CDCl_3$): d = 4.84 (d, 1H, J=3.6 Hz, H-1), 4.13 (d, 1H, J=9.6 Hz, H-5), 3.83 (s, 3H, $CH_{3\text{-}COOMe}$), 3.79 (dd, 1H, J=8.98, 9.70, H-3), 3.73 (dd, 1H, J=3.70, 9.70, H-2), 3.61 (dd, 1H, J=3.6, 9.6, H-4), 3.47 (s, 3H, $CH_{3\text{-}OMe}$); ^{13}C NMR (400 MHz, $CDCl_3$): d = 170.24 ($C{=}O_{COOMe}$), 99.73 (C-1), 73.18 (C-5), 71.32 (C-3), 71.15 (C-2), 70.48 (C-4), 55.46 ($CH_{3\text{-}COOMe}$), 52.36 ($CH_{3\text{-}OMe}$).

3.1.3 Methyl (2,3-acetonide-1-*O*-methyl- α-D-glucopyranose) uronate (4)

Methyl (1-*O*-methyl-D-glucopyranose) uronate **3** (0.237 g, 1.066 mmol) was dissolved in dry DMF(0.9 ml). While stirred under nitrogen atmosphere in room temperature methanolic hydrochloric acid (3M, 3.4 µl) was added to the solution. Then 2-methoxypropene (0.18 ml, 1.89 mmol) was slowly added dropwise. TLC (DCM-EtOAc 1:2) indicated incomplete reaction after 3h. The solution was diluted with DCM (6 ml) and washed with $NaHCO_3$ (6 ml) and brine (6 ml). The aqueous layers were extracted twice with DCM (3 ml) and the combined organic phases were dried over $MgSO_4$ and evaporated. NMR-spectra revealed an unsuccessful reaction.

3.1.4 Methyl (1-*O*-methyl-2,3,4-tri(*O*-triethylsilyl)-α-D-glucopyranose) uronate (5)

Methyl (1-*O*-methyl-D-glucopyranose) uronate **3** (0.23 g, 1.035 mmol) was dissolved in dry DMF (5 ml). While stirred under nitrogen atmosphere in room temperature imidazole (0.566 g, 8.314 mmol) was added to the solution prior to the addition of chlorotriethylsilane, TESCl (0.70, 4.17 mmol). The solution was left stirring until TLC (DCM-EtOAc 8:1) indicated the complete consumption of the reactant (1h). The solution was diluted with Et_2O (30 ml) and washed with $NaHCO_3$ (30 ml). Then the aqueous layer, which was a white suspension, was extracted twice with Et_2O (2 x 10 ml). The combined organic phases were dried over $MgSO_4$,

filtered and evaporated. Purification by short-path distillation yielded the product (0.360 g, 0.637 mmol, 61.6 %). ^{1}H NMR (400 MHz, $CDCl_3$): d = 4.71 (d, 1H, J=3.1 Hz, H-1), 4.14-4.06 (m, 1H, H-5), 3.77 (s, 3H, $CH_{3\text{-}COOMe}$), 3.78-3.82 (m, 2H, H-3, H-4), 3.55 (dd, 1H, J=8.3, 3.3, H-2), 3.38 (s, 3H, $CH_{3\text{-}OMe}$), 1.05-0.80 (m, 46H), 0.74-0.45 (m, 32H); ^{13}C NMR (400 MHz, $CDCl_3$): d =170.45 ($C{=}O_{COOMe}$), 99.85 (C-1), 74.62 (C-5), 73.39 (C-3), 73.20 (C-2), 73.14 (C-4), 55.30 ($CH_{3\text{-}COOMe}$), 52.16 ($CH_{3\text{-}OMe}$), 4.96-7.03 ($CH_{3\text{-}O\text{-}Si}$ x 10).

3.1.5 1-*O*-Methyl-2,3,4-tri(*O*-triethylsilyl)-α-D-glucuronic acid (C)

Methyl (1-*O*-methyl-2,3,4-tri(*O*-triethylsilyl)-α-D-glucopyranose) uronate **5** (0.23 g, 0.407 mmol) was dissolved in THF (5 ml) and subsequently a NaOH solution (1.49 µM, 1.0 ml, 1.49 mmol) was added causing a slightly yellow two phase suspension. The suspension was left stirring at rt under 24h. TLC was unclear and the suspension continued while vigoursly stirred at 40°C under reflux. After 24h work up was initiated and the suspension was diluted with diethyl ether (30 ml) and washed with $NaHCO_3$. The yellow aqueous layers were extracted twice with diethyl ether (2x10 ml) and the combined transparent organic layers were dried over $MgSO_4$ and filtered. The filtrate was subsequently evaporated yielding the product (0.27 g, 0.49 mmol, 120.5%) NMR result showed partial consumption of the starting material. The reaction was resumed using the partially converted starting material (0.23, 0.407 mmol) in THF (5 ml). NaOH solution (1.6 µM, 1.0 ml, 1.625 mmol) was added while stirred causing a slightly yellow two phase suspension. The reaction was left stirring at 50°C under reflux during 24h. The suspension was then diluted with diethyl ether (30 ml) and washed with $NaHCO_3$ (30 ml) causing two transparent layers. The aqueous layers were extracted twice with diethyl ether (2x10 ml) and the combined organic layers were dried over $MgSO_4$ and filtered. The filtrate was subsequently evaporated yielding the crude product as a slightly yellow oil (0.257 g). ^{1}H NMR (400 MHz, $CDCl_3$): d = 4.71 (m, 1H, H-1), 4.08 (m, 1H, H-5), 3.78-3.82 (m, 2H, H-3, H-4), 3.55 (m, 1H, H-2), 3.38 (s, 3H, $CH_{3\text{-}OMe}$), 1.05-0.80 (m, 46H), 0.74-0.45 (m, 32H); ^{13}C NMR (400 MHz, $CDCl_3$): d =170.08 ($C{=}O_{COOH}$), 99.82 (C-1), 74.62 (C-5), 73.39 (C-3), 73.20 (C-2), 73.14 (C-4), 55.25 ($CH_{3\text{-}OMe}$), 4.96-7.03 ($CH_{3\text{-}O\text{-}Si}$ x 16).

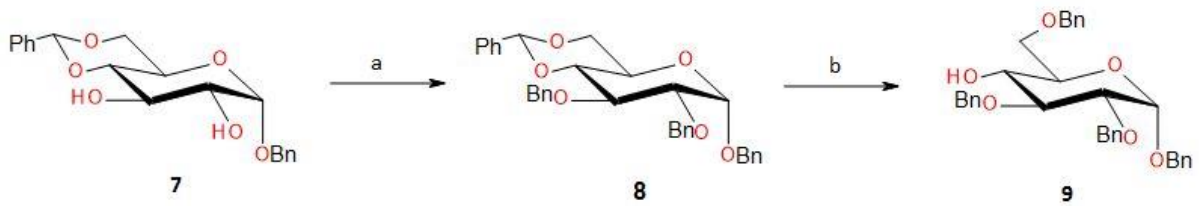

Figure 5. Alternative synthesis route. Reagents and conditions: (a) NaH, BnBr, DMF, rt, 3h; (b) Triethylsilane, Triflouroacetic acid, DCM, rt

3.1.6 Benzyl 2,3-di-O-benzyl-4,6-O-benzylidene-α-D-glucopyranoside (8)

Benzyl 4,6-O-benzylidene-a-D-glucopyranoside **7** (0.206 g, 0.575 mmol) was dissolved in dry DMF(2.3 ml). While stirred under nitrogen atmosphere at 0°C 0,084 g NaH (60% suspension in oil, 2,06 mmol) was slowly added dropwise which caused a gas development (white foam) and made the solution yellow. Subsequently BnBr (0.15 ml, 1.26 mmol) was added to the solution. The reaction was then stirred in room temperature until TLC-analysis indicated complete consumption of reactant (~3h). 2-propanol (0.2 ml, 2.62 mmol) was then added and the solution was diluted with ethyl acetate. The solution was then extracted 3 times with brine (3 x 8ml) and the aqueous layer, which was a white suspension, was extracted with diethyl ether(4,5 ml). The organic phase was dried over $MgSO_4$ and evaporated. ^{1}H NMR (400 MHz, $CDCl_3$): d = 7.52 – 7.23 (m, 25H), 4.93 (d, J = 11.2 Hz, 1H), 4.86 (d, J=12.04 Hz,1H), 4.83 (d, J = 4.0 Hz, 1H), 4.76 (d, J = 7.0 Hz, 1H), 4.73 (d, J = 7.3 Hz, 1H), 4.59 (d, J = 12.2 Hz, 2H), 4.20 (dd, J = 10.2, 4.9 Hz, 1H), 4.11 (t, J = 9.3 Hz, 1H), 3.91 (td, J = 10.0, 4.9 Hz, 1H), 3.69 (t, J = 10.3 Hz, 1H), 3.61 (t, J = 9.4 Hz, 1H), 3.56 (dd, J = 9.3, 3.8 Hz, 1H); ^{13}C NMR (400 MHz, $CDCl_3$): d = 138-126 (Ph), 101.21 (C-1), 96.55 (CHPh), 82.2 (C-2), 79.25 (C-4), 78.68 (C-3), 75.33, 73.45, 69.33 (CH_2Ph), 69.02 (C-6), 62.62 (C-5).

3.1.7 Methyl 2,3,6-tri-*O*-benzyl-α-D-glucopyranoside (9)

Benzyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-α-D-glucopyranoside **8** was dissolved in anhydrous DCM (2 ml). While stirred under nitrogen atmosphere at 0^{o}C triethylsilane (28 µl, 0.0179 mmol) and triflouroacetic acid (14 µl, 0.178 mmol) was added to the solution. The solution was left stirring in room temperature for 3h subsequent to a second addition of triethylsilane (24 µl, 0.149 mmol) and triflouroacetic acid (11.4 µl, 0.149 mmol) at 0^{o}C. The solution was once again stirred in room temperature for 2h subsequent to a third addition of triethylsilane (24 µl, 0.149 mmol) and triflouroacetic acid (11.4 µl, 0.149 mmol) at 0^{o}C. The solution was left stirring in room temperature until TLC (EtOAc-Toluene 1:18) indicated complete reaction (3h). The solution was then diluted with ethyl acetat (4 ml), neutralized with $NaHCO_3$ (5 ml) and washed with HCl (0.05 M, 5 ml). The solution was then neutralized with $NaHCO_3$ (5 ml), washed with brine (5 ml). The organic layer was dried over Na_2SO_4 and then evaporated yielding the crude product **9** (0.159 g). An attempt to purify the product was made by flash column chromatography (EtOAc-toluene 1:13) however the chosen eluant would not dissolve the product. . ^{1}H NMR (400 MHz, $CDCl_3$): d = 7.52 – 7.23 (m, 24H, Ph), 5.06 (d, 2H, CH_2Ph), 4.85 (d, 1H, H-1), 4.73-4.5 (m, 6H, CH_2Ph x 3), 3.87 (dd, 1H, H-3), 3.8-3.55 (m, 5H, H-5, H-6a, H-4, H-6b, H-2); ^{13}C NMR (400 MHz, $CDCl_3$); d = 139-126 (Ph), 95.55 (C-1), 81.33 (C-3), 79.79 (C-2), 75.34, 72.30, 72.11, 70.97 (CH_2Ph), 70.52 (C-5), 69.25 (C-4), 70.14 (C-6).

4 Results and Discussion

In this section the results in terms of analysis of NMR spectra, yield and purity of the various compound products in the two synthesis routes will be discussed.

As depicted in the reaction scheme in Figure 4 Methyl (1-*O*-methyl-D-glucopyranose) uronate **3** was used as a common starting material for of target molecules **A**, **B** and **C**. **3** was prepared in two steps from the methyl-α-D-glucopyranoside **1**. An oxidation of alcohol group at C-5 on the pyranose ring of **1** with 2,2,6,6-tetramethylpiperidin-1-yloxyl (TEMPO) together with bis(acetoxy)iodo benzene (BAIB) as co-oxidant resulted in a carboxylic acid group which provided the methylated glucuronic acid **2** with satisfactory purity according to NMR spectra (Figure A1, Figure A2 in appendix). Methylation of **2** followed in order to reduce the reactivity of the hydroxyl group at the carboxylic acid moiety which by experience is believed to interfere with the formation of isopropylidene acetal in a following step. The carboxylate was methylated with TMS-diazomethane. These two steps resulted in the common starting point of Methyl (1-*O*-methyl-D-glucopyranose) uronate **3** in 86,9% yield. NMR spectra (Figure A3, Figure A4 in appendix) indicated a good result apart from some solvent remainder which was eliminated through evaporation.

Reaction steps that included both 2-methoxypropen and 2,2-dimethoxypropane was made in order to protect **3** with an isopropylidene acetal. This protective group would have protected the 1,2-diol at C-2 and C-3 prior to ester cleaving reaction steps that could lead to target molecules **A** and **B**. The hydroxyl groups at C-2 and C-3 would otherwise interfere in future methylation reaction of hydroxyl group at C-4. According to NMR result no acetal had been formed.

The formation of cyclic acetals can either be carried out under equilibrating conditions where the product that is most thermodynamically stable is formed or under kinetic control where the product formed depends on competing rates of product formation. If the reaction takes a kinetic or thermodynamic route depends on various factors such as solvent, type of acid catalyst, reaction time, temperature and type of reagent used (15). The reason for the defaulted reaction could be that the isopropylidene acetal that was supposed to be synthesized was not thermodynamically stable and that it would have needed kinetic control to avoid the reaction to return to the more stable starting material. Another factor that could have been the reason is that air or moisture interfered with the reaction and caused immediate cleavage of the acetonides formed.

Silylation with imidazole and chlorotriethylsilane of **3** provided silyl ether groups on C-2, C-3 and C-4 resulting in a methylated target molecule **C** (compound **5**). NMR spectra indicated an excess of ether silyl groups which indicated remaining unbound reagent in the sample and consequently the sample had to be purified. Due to high boiling point of triethylsilanol purification was done by using short-path distillation which provided the compound **5** in 61.6

% yield. In a subsequent step the ester moiety of **5** was cleaved by basic hydrolysis involving NaOH at 40-50°C to deliver crude target molecule **C**.

^{1}H NMR spectrum of target molecule C in Figure 6 depicts that the desired product is present in the sample. When comparing it with the NMR spectrum of the starting material **5,** which can be seen in Figure A5 in appendix, it has similar peaks apart from the singlet peak from the methyl group at the ester moiety at 3.77 ppm which, as expected, is absent. This indicates that the desired product **C** has been formed, however, the spectrum is indistinct presumably due to impurities and remaining starting material. Therefore the sample should be purified to dispose of impurities and to be able to obtain a clearer spectrum. The ^{13}C NMR spectrum is also rather indistinct and could improve by purification. When measuring conversion it can be assumed, relying on NMR integral measurements, that about 90% of the starting material has converted to product which is considered to be a good result.

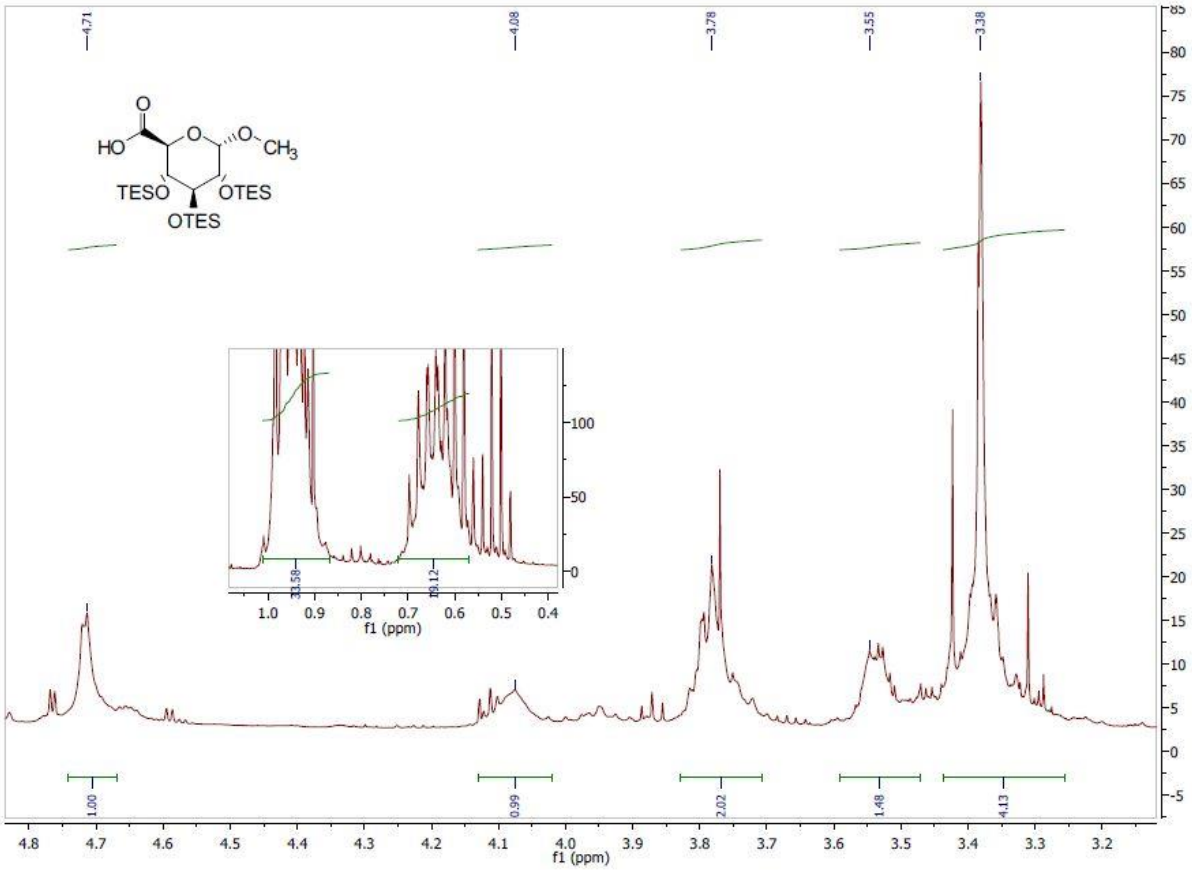

*Figure 6. 1H NMR of compound **C** recorded in $CDCl_3$.*

The unsuccessful protection step of compound **3** that should produce the acetylated compound **4** generated the necessity of an alternative synthesis route towards the remaining two target molecules **A** and **B** that is depicted in Figure 5. In this route the 1-*O*-benzyl-α-D-glucopyranoside **7** protected with a benzylidene acetal is used as starting material which

subsequently has its two OH groups at C-2 and C-3 benzylated with sodium hydride, NaH and benzyl bromide, BnBr provided **8** in 78% yield. The benzylidene acetal of **8** was then deprotected which yielded crude compound **9**. Compound **9** contained some impurities and remaining starting material from compound **8.** Results are depicted in NMR spectra in Figure A10 and Figure A11 in appendix. When measuring conversion it can be assumed, relying on NMR integral measurements, that about 60% of the starting material has converted to product. The reaction step yielding compound **9** would have needed more time to increase the conversion and purification with flash column chromatography is necessary. An attempt of purification was made with EtOAc-Toluene as eluant, however, the chosen eluant would not dissolve the sample. Purification with flash column chromatography could be done in the future by choosing another eluant such as petroleum ether-DCM in which ability to dissolve **9** has been established by TLC.

5 Conclusion and future work

In summary, two synthesis routes has been used in order to synthesize three target molecules that can be used as model substrates. A methylated glucopyranoside in the original synthesis route was used as starting material which reached target molecule **C** with good conversion.

The protective group strategy that included an isopropylidene acetal on the path towards the two other target molecules **A** and **B** was unsuccessful. If this protective group is desired an air-free technique should be tested to eliminate air and moisture contamination and in addition further work might be needed in order to achieve the kinetic control necessary for this reaction step to succeed. Another alternative synthesis route has been initiated that includes benzylidene acetal and benzyl ethers as protective groups. With the appropriate reaction and protective group strategy, this synthesis route could reach target molecules **A** and **B**.

References

1. Kumar R, Singh S, Singh O V. Bioconversion of lignocellulosic biomass: Biochemical and molecular perspectives. J Ind Microbiol Biotechnol. 2008;35(5):377–91.

2. Naik SN, Goud V V., Rout PK, Dalai AK. Production of first and second generation biofuels: A comprehensive review. Renew Sustain Energy Rev. 2010;14(2):578–97.

3. Katsimpouras C, Bénarouche A, Navarro D, Karpusas M, Dimarogona M, Berrin JG, et al. Enzymatic synthesis of model substrates recognized by glucuronoyl esterases from Podospora anserina and Myceliophthora thermophila. Appl Microbiol Biotechnol. 2014;98(12):5507–16.

4. Gibson LJ. The hierarchical structure and mechanics of plant materials. J R Soc Interface. 2012;9(76):2749–66.

5. Sun R-C. Cereal Straw as a Resource for Sustainable Biomaterials and Biofuels - Chemistry, Extractives, Lignins, Hemicelluloses and Cellulose [Internet]. Elsevier. Elsevier; 2010 [cited 2015 Apr 29]. Available from: http://app.knovel.com/web/toc.v/cid:kpCSRSBBC1/viewerType:toc/root_slug:cereal-straw-as

6. Gírio FM, Fonseca C, Carvalheiro F, Duarte LC, Marques S, Bogel-Łukasik R. Hemicelluloses for fuel ethanol: A review. Bioresour Technol [Internet]. Elsevier Ltd; 2010;101(13):4775–800. Available from: http://dx.doi.org/10.1016/j.biortech.2010.01.088

7. Aita GA, Salvi DA, Walker MS. Enzyme hydrolysis and ethanol fermentation of dilute ammonia pretreated energy cane. Bioresour Technol [Internet]. Elsevier Ltd; 2011;102(6):4444–8. Available from: http://dx.doi.org/10.1016/j.biortech.2010.12.095

8. Gellerstedt G. Wood Chemistry. Göteborg: Vasastadens bokbinderi; 2008.

9. Balakshin M, Capanema E, Gracz H, Chang H, Jameel H. Quantification of lignin-carbohydrate linkages with high-resolution NMR spectroscopy. Planta. 2011;233(6):1097–110.

10. Li X-L, Špániková S, de Vries RP, Biely P. Identification of genes encoding microbial glucuronoyl esterases. FEBS Lett. 2007;581(21):4029–35.

11. Stick R V. Carbohydrates: The Sweet Molecules of Life. San Diego, California: Academic Press; 2001.

12. Greene TW, Wuts PGM. Greene´s Protective Groups in Organic Synthesis. 4th ed. Hoboken, N.J: Wiley-Interscience; 2007.

13. Kahn SH, O'Neill RA, editors. Modern Methods in Carbohydrate Synthesis. Amsterdam: Hardwood Academic Publishers; 1996. 226-228 p.

14. Osborn H, editor. Carbohydrates Best synthetic methods [Internet]. Academic Press; 2003. 29-33 p. Available from: https://books.google.se/books?id=S4sSyb6NsNcC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

15. Kociency, Phillip J. Protecting Groups. 3rd ed. 2005.

Appendix

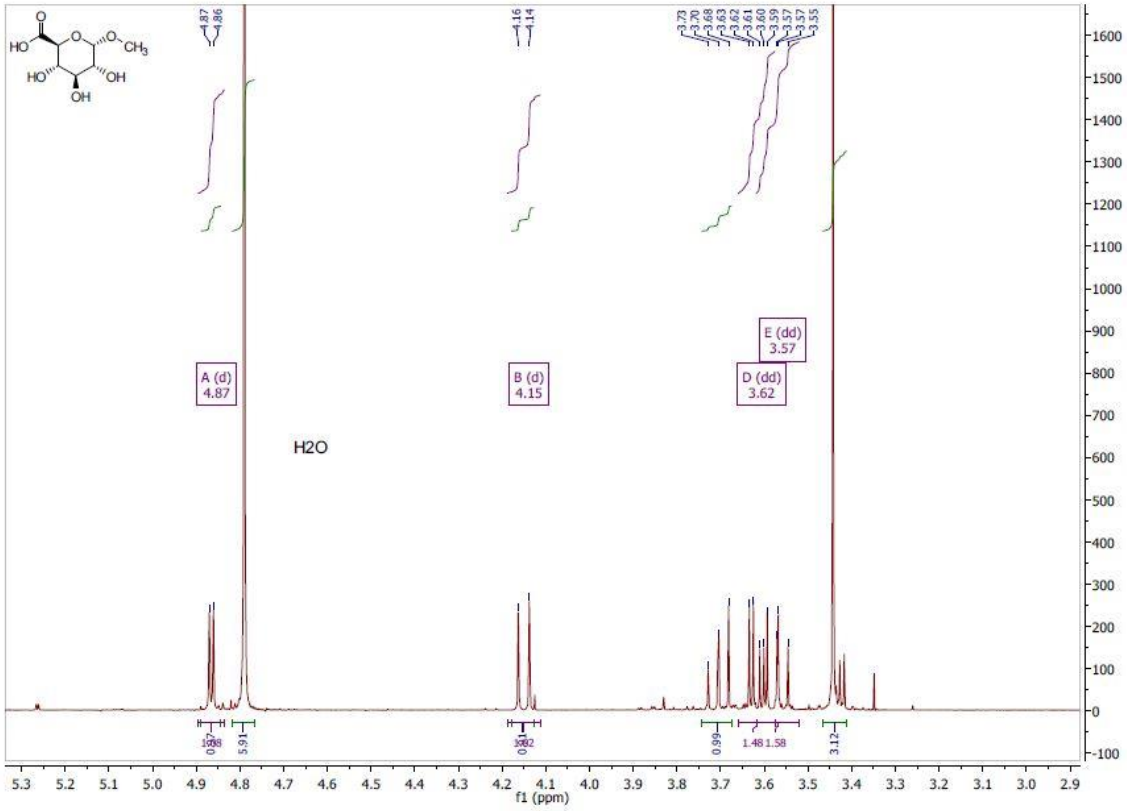

*Figure A1. ^{1}H NMR of compound **2** recorded in D_2O.*

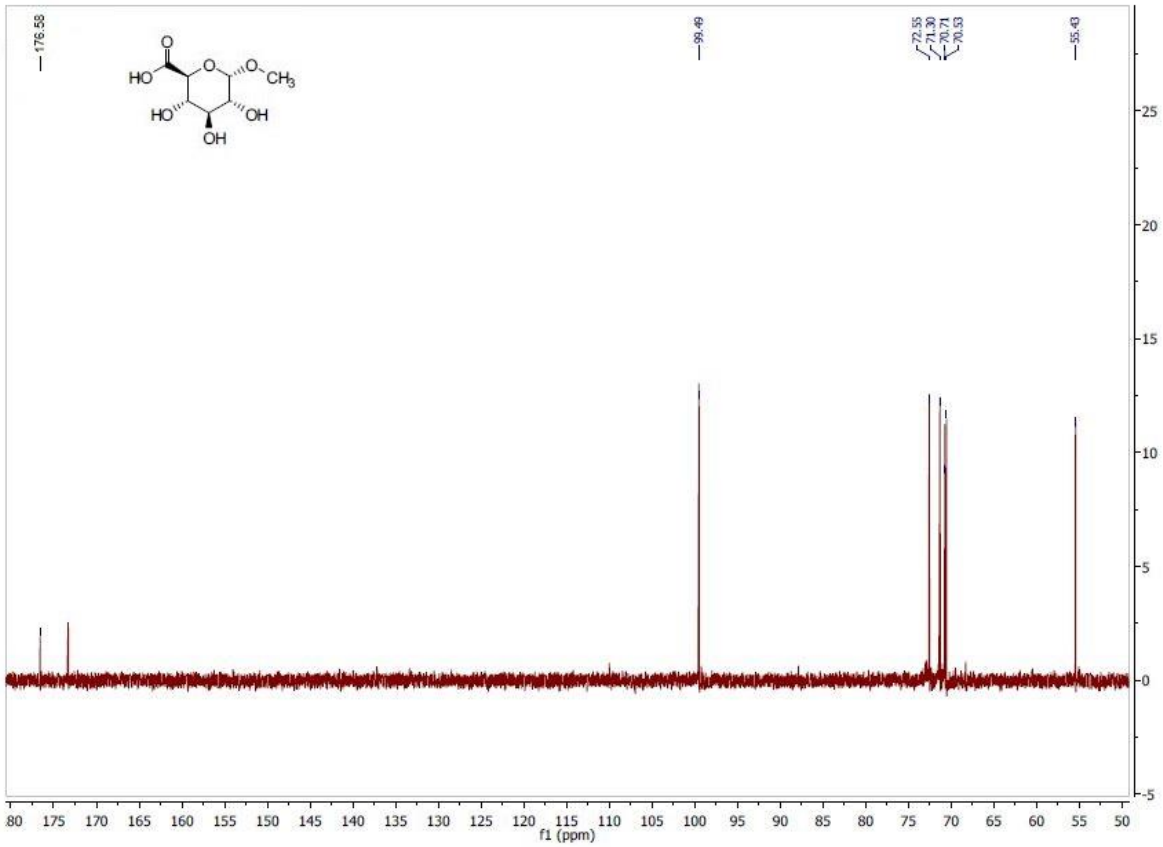

*Figure A2. ^{13}C NMR of compound **2** recorded in D_2O.*

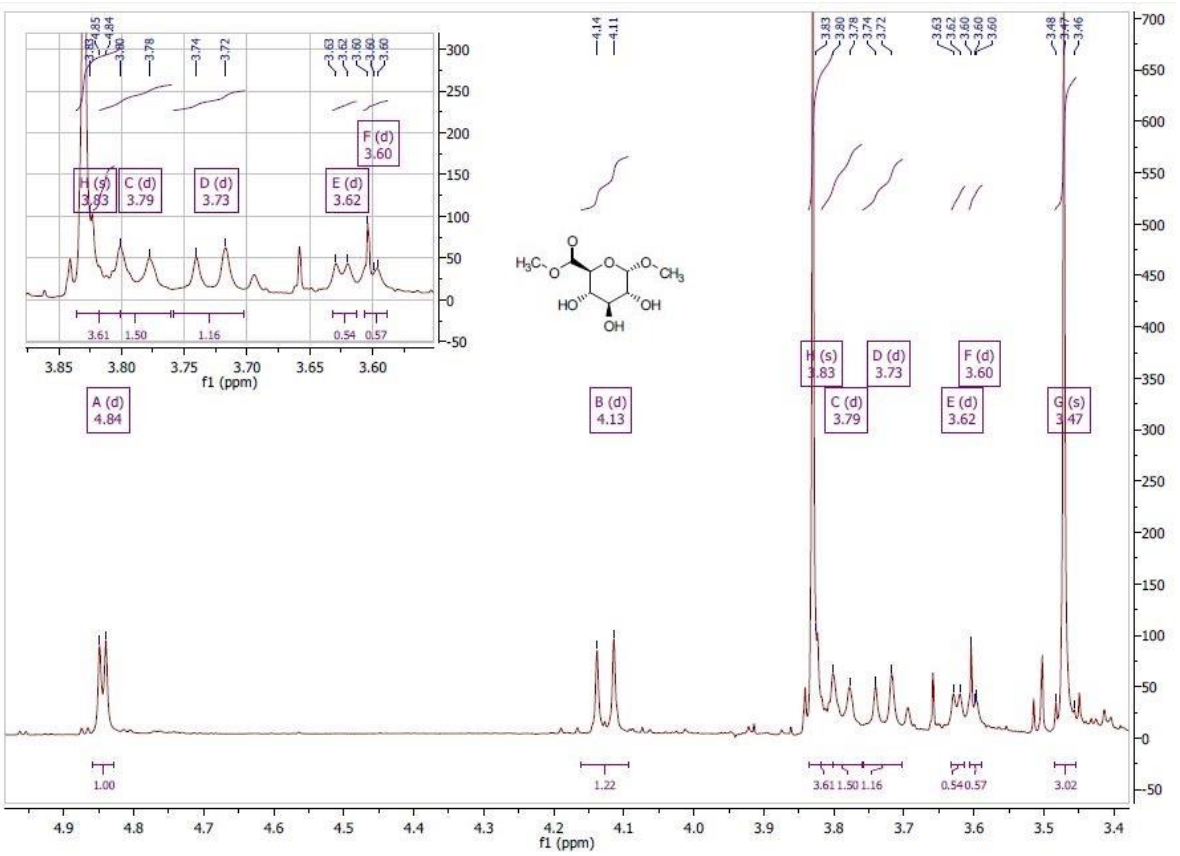

*Figure A3. ^{1}H NMR of compound **3** recorded in $CDCl_3$.*

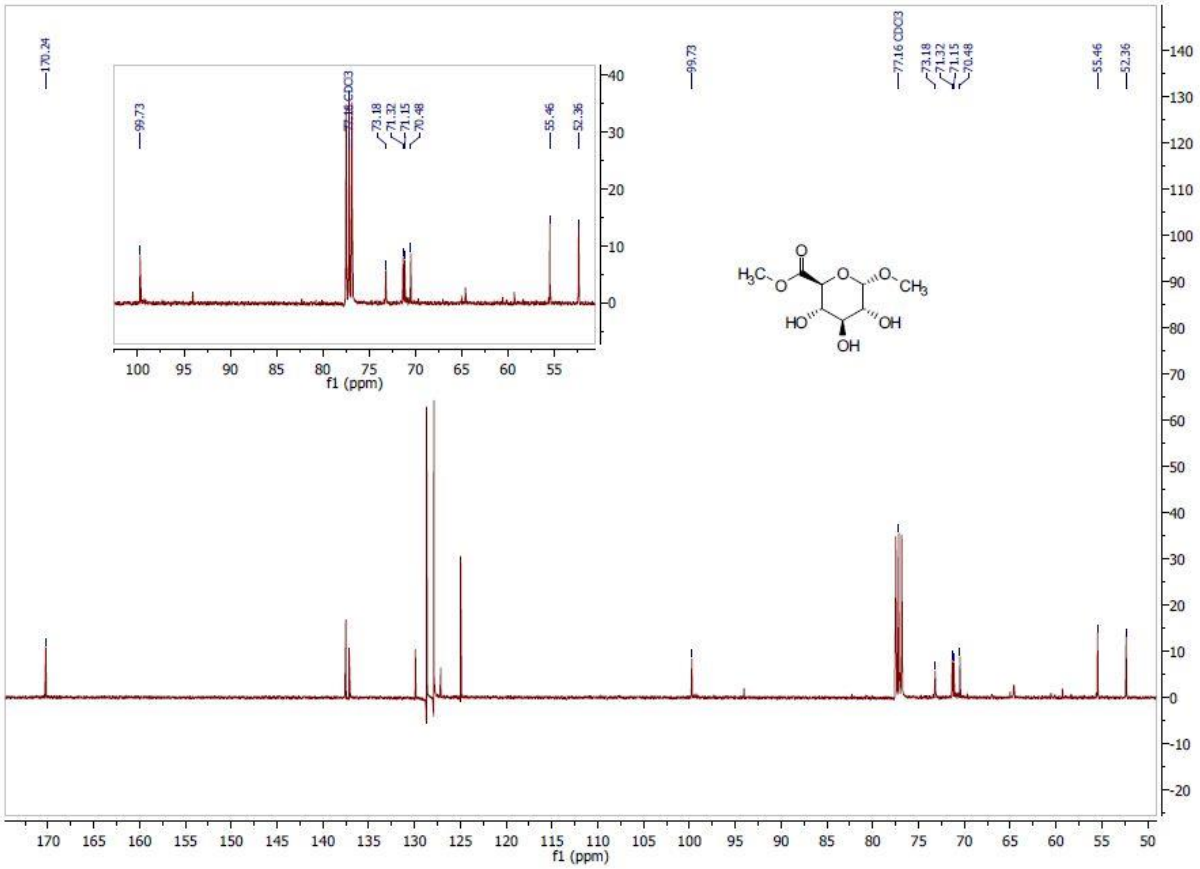

*Figure A4. ^{13}C NMR of compound **3** recorded in $CDCl_3$*

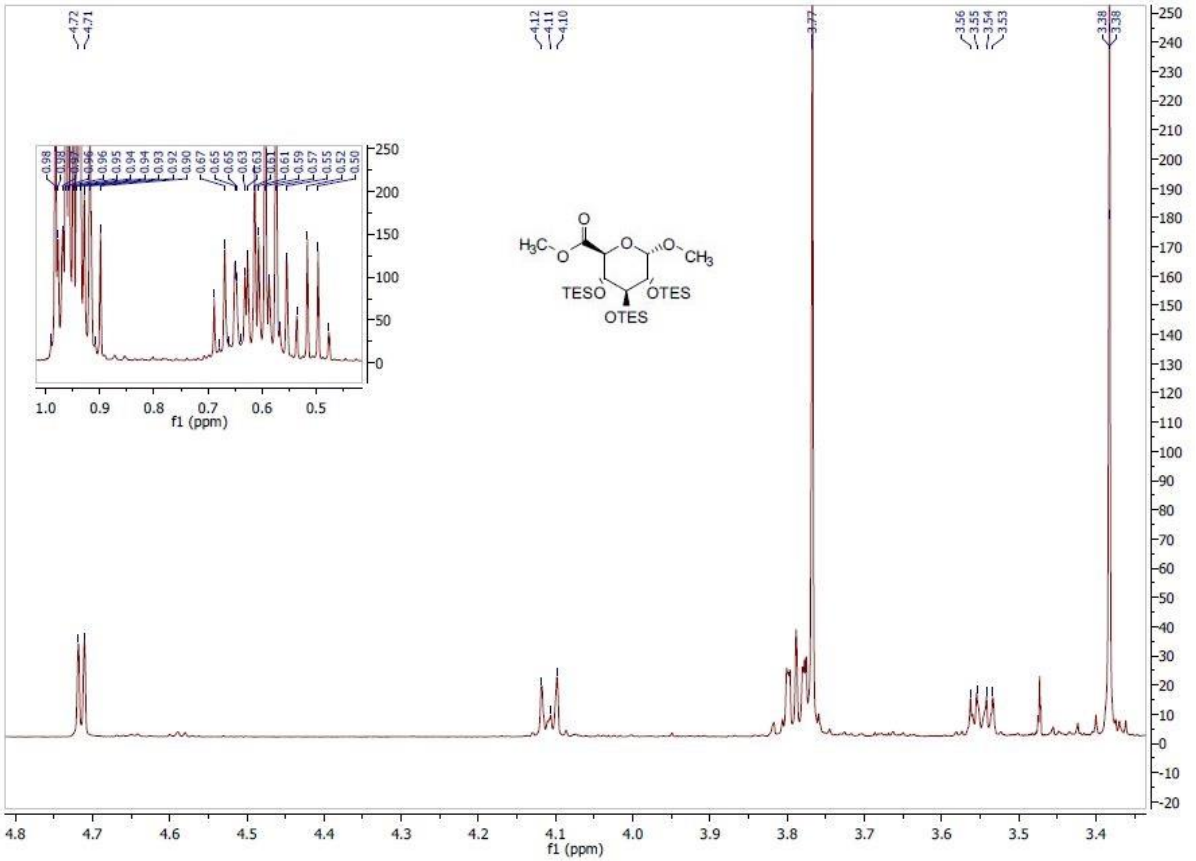

*Figure A5. ^{1}H NMR of compound **5** recorded in $CDCl_3$.*

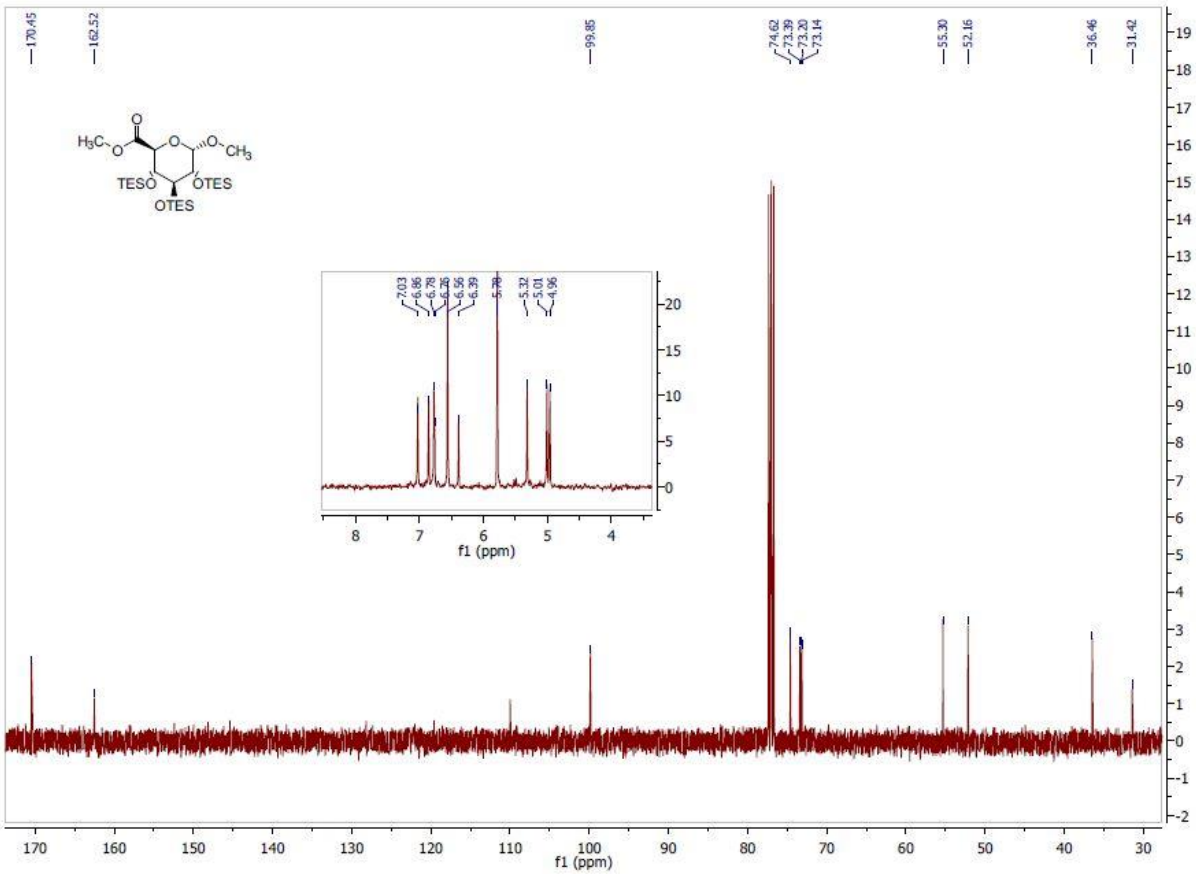

*Figure A6. ^{13}C NMR of compound **5** recorded in $CDCl_3$.*

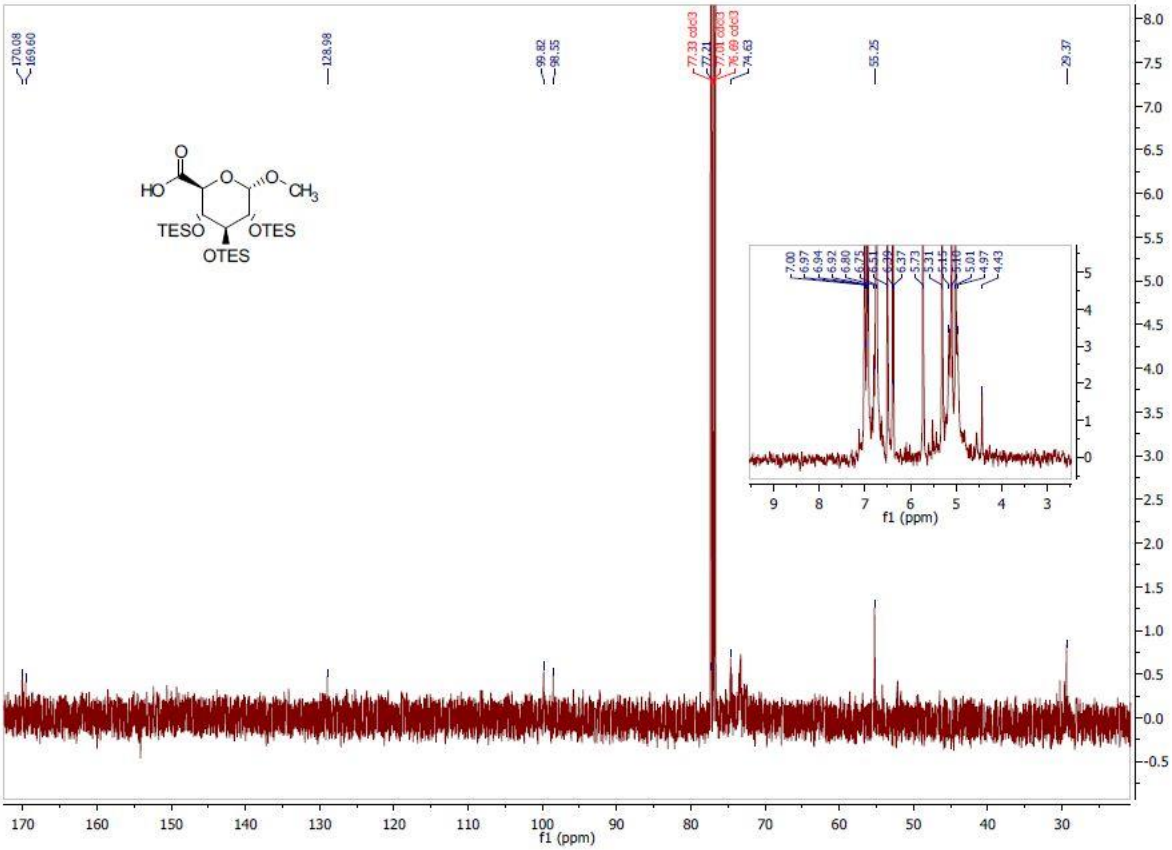

*Figure A7. ^{13}C NMR of compound **C** recorded in $CDCl_3$ recorded in $CDCl_3$.*

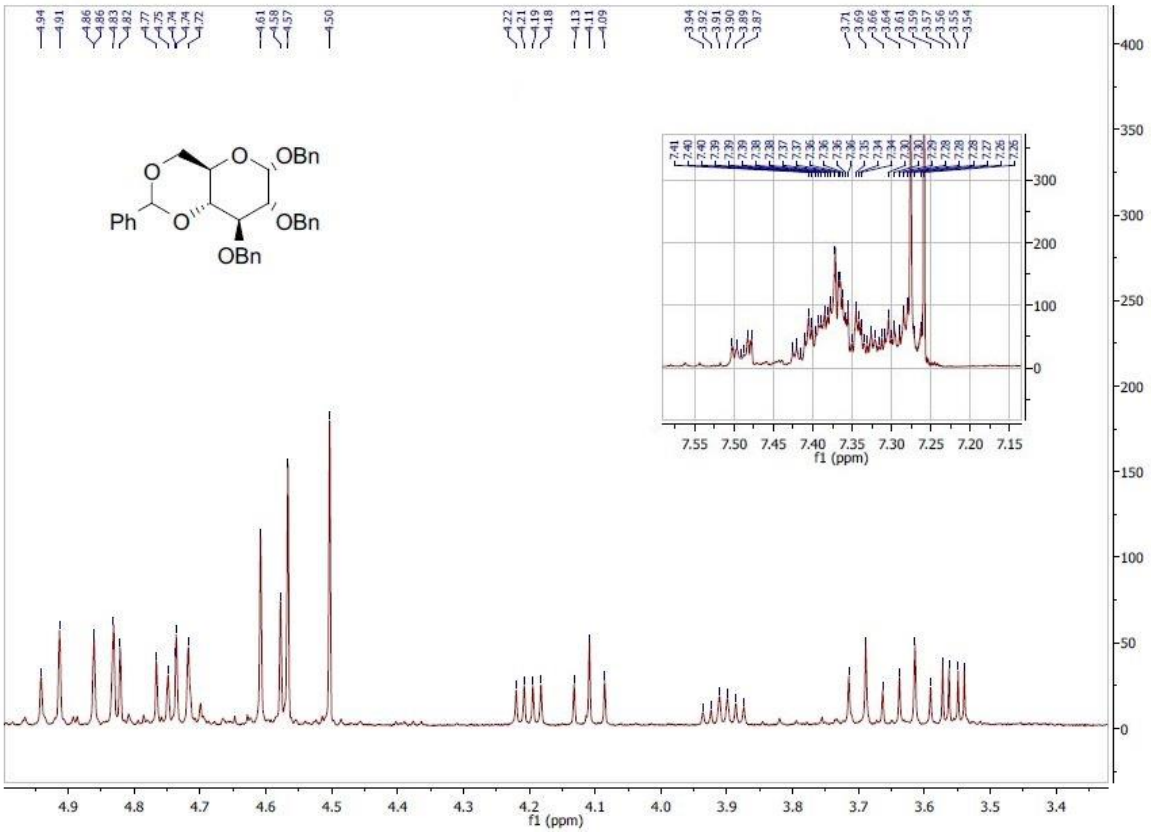

Figure A8. ^{1}H NMR of compound 8 recorded in $CDCl_3$.

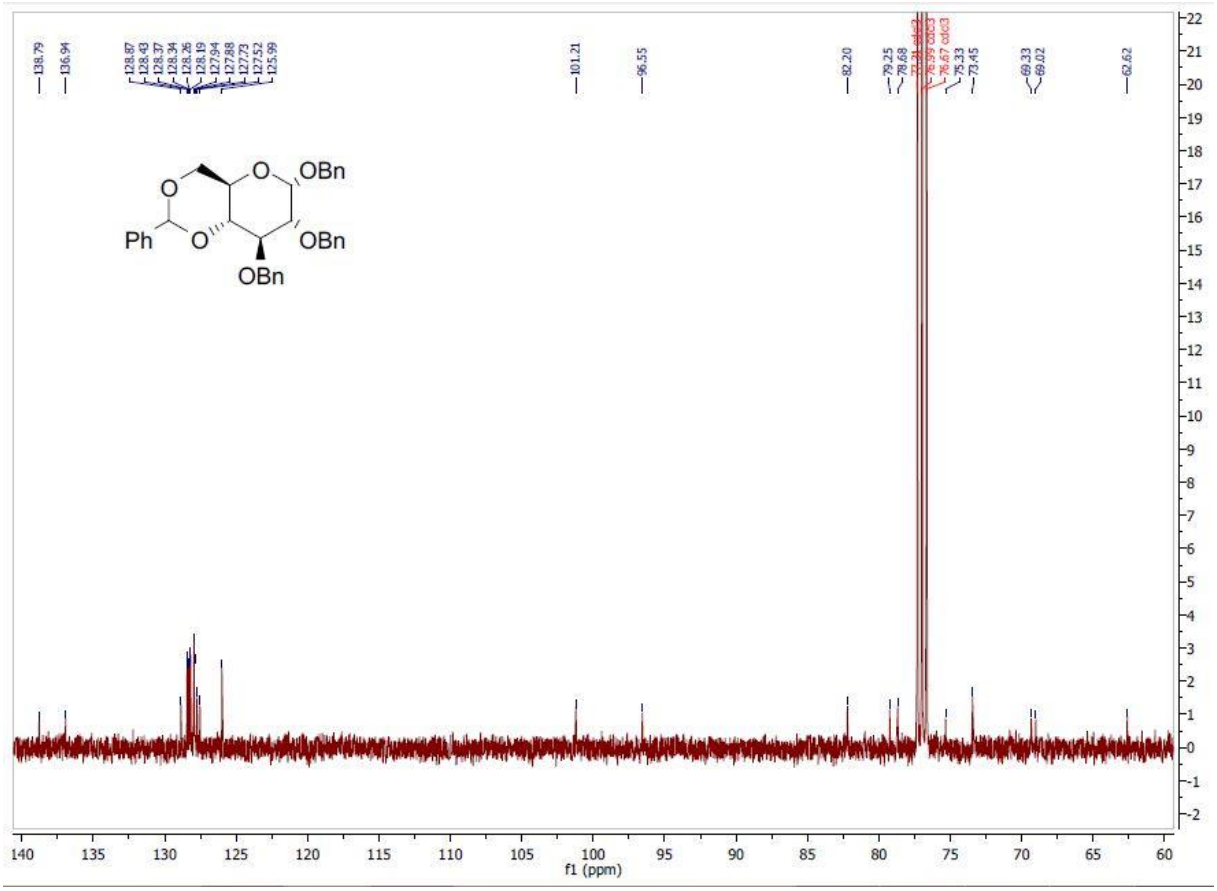

Figure A9. ^{13}C NMR of compound 8 recorded in $CDCl_3$.

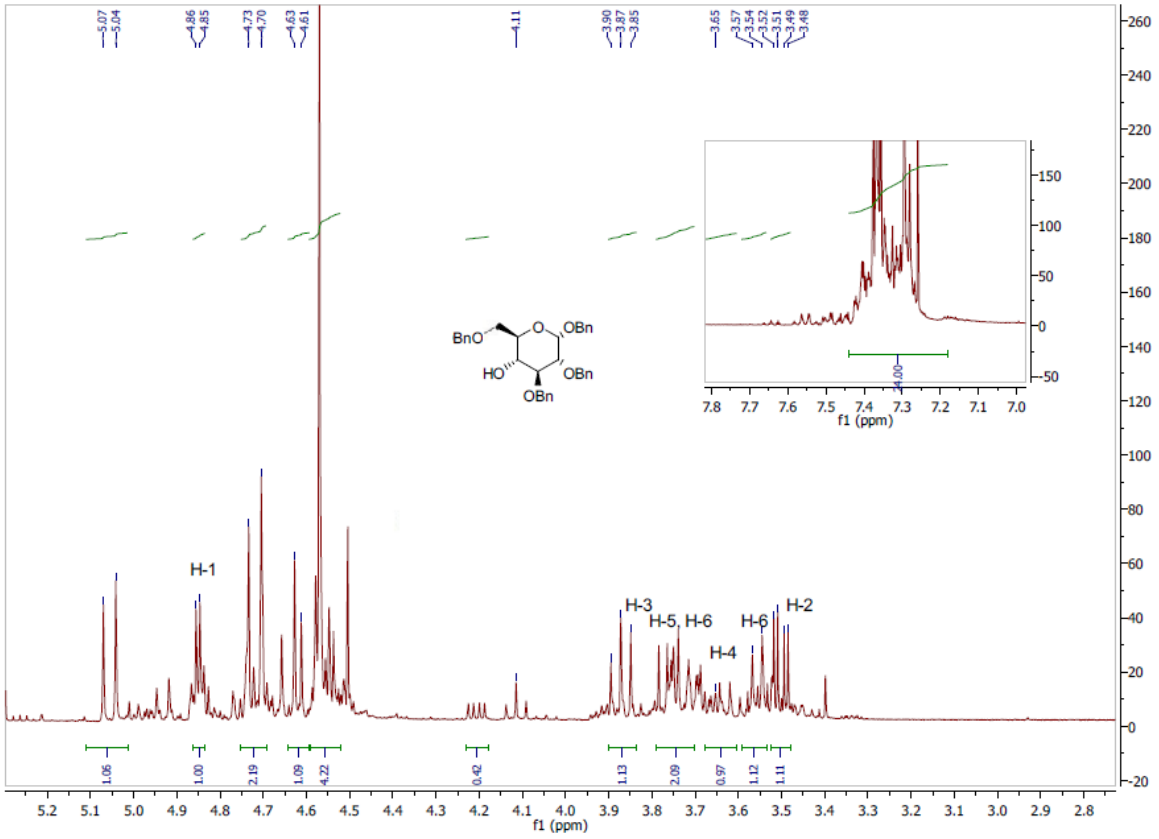

Figure A10. ^{1}H NMR of compound 9 recorded in $CDCl_3$.

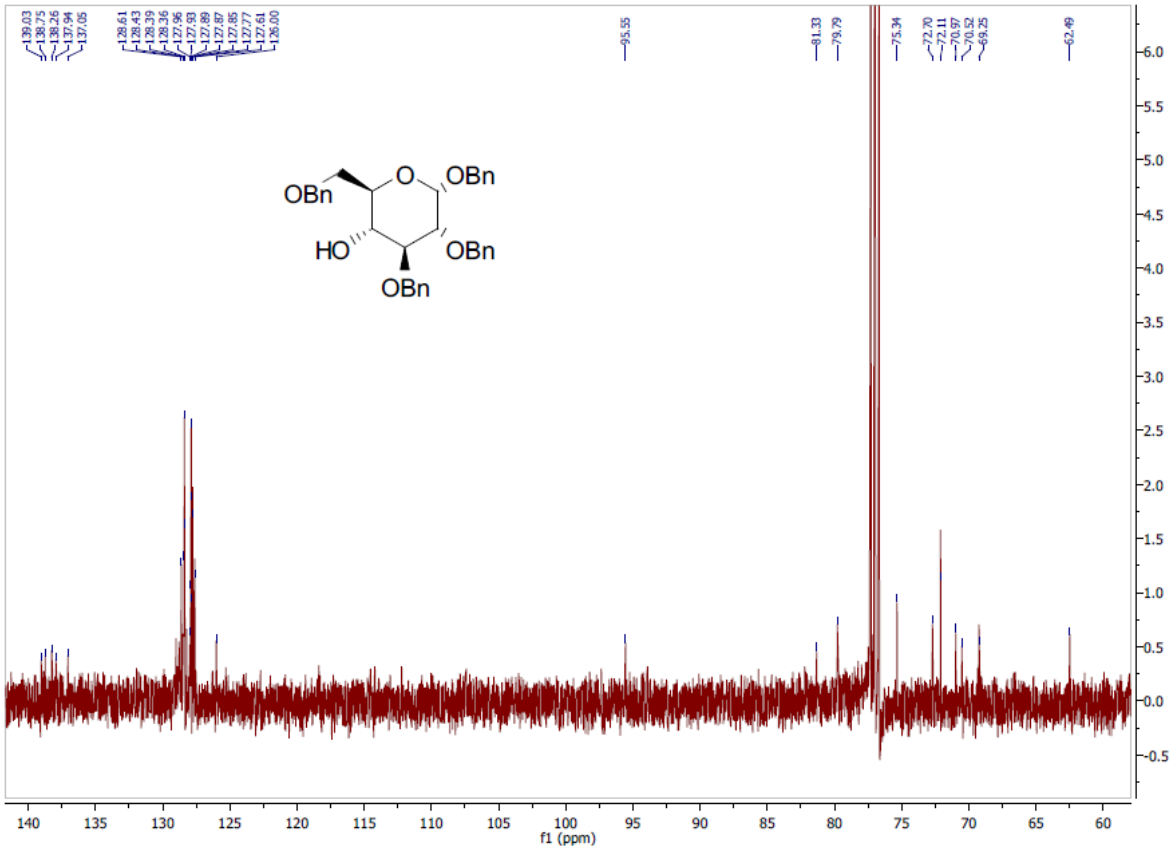

Figure A11. ^{13}C NMR of compound 9 recorded in $CDCl_3$.