

AN INTRODUCTION TO GAS CHROMATOGRAPHY MASS SPECTROSCOPY FOR THE STRUCTURAL ELUCIDATION OF POLYSACCHARIDES FROM SUGAR PROCESSING STREAMS

DU CLOU H^{1,2} AND WALFORD S N²

¹*School of Chemistry, University of KwaZulu-Natal, Westville, 3630, South Africa*

²*Sugar Milling Research Institute, University of KwaZulu-Natal, Durban, 4041, South Africa*
hduclou@smri.org swalford@smri.org

Abstract

The gas chromatograph (GC) coupled with a mass spectrometer (GC-MS) has become a highly versatile tool in determining the structure of compounds such as sugars; including the constituents of oligosaccharides and polysaccharides. For higher saccharides a GC-MS can be used to determine its constituent monosaccharides, the linkage position between these monosaccharides, the general position in the chain (terminal/reducing end, branch point, or intra-chain) as well as the ring size of each sugar (being either a five-membered furanose or six-membered pyranose). The success of such analyses relies on a method of sample preparation which involves a permethylation step, followed by methanolysis and silylation of the individual residues. These residues are separated in the GC, and their structure is determined through the specific, reproducible fragmentation patterns generated from an electron impact process. The resultant ions are separated in the mass spectrometer and the mass spectra are compiled into a database which is used as a reference library. This library is used as an aid in the identification of similarly prepared monosaccharide residues with respect to their type, configuration, and general position within an oligo- or polysaccharide chain. This paper reports the development of a mass spectral library from model compounds in order to qualify the fine structure within complex sugars and uses cane starch (found in sugar factory processing streams) to illustrate the methods used.

Keywords: GC-MS, monosaccharides, polysaccharides, structure, elucidation, library

Introduction

Polysaccharides

Monosaccharides have the general formula $C_n(H_2O)_n$ and exist as either an aldose (aldehyde form) or ketose (ketone form) when in their acyclic (non-cyclic) form. Monosaccharides can also exist as either five-membered (furanose) or six-membered (pyranose) rings as illustrated in Figure 1 (Lindhorst, 2000).

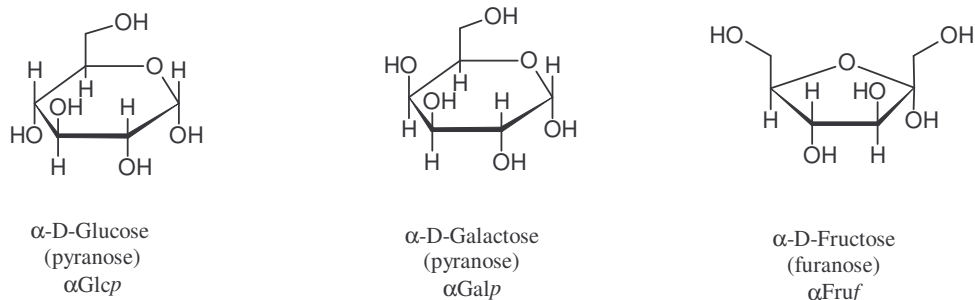


Figure 1. Haworth formulae for some common monosaccharides showing the pyranose (p) and furanose (f) forms.

When sugars are in solution they mutarotate between their alpha (α) and beta (β) forms, as illustrated in Figure 2. The α -conformation occurs if the hydroxy at the first carbon (C-1) is in the axial position, whilst the β -conformation is when this -OH is in the equatorial plane of the carbon ring (Lindhorst, 2000; Robyt, 1998).

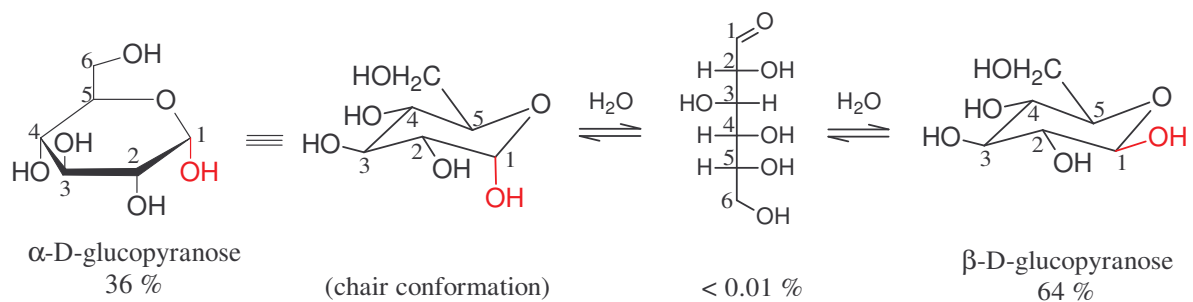


Figure 2. Mutarotation equilibrium between the α - and β -anomers of glucopyranose in aqueous solution. Some hydrogens have not been shown for clarity.

Polysaccharides are high molecular mass biopolymers made up of monomeric sugars. When sugar monomers are bonded together they form disaccharides, trisaccharides, oligosaccharides and polysaccharides. Sucrose is a disaccharide consisting of glucose and fructose, joined at the C-1 of the glucose and the C-2 of the fructose. Figure 3 illustrates a glucopyranose-type polysaccharide and some of the glycosidic linkages that are possible.

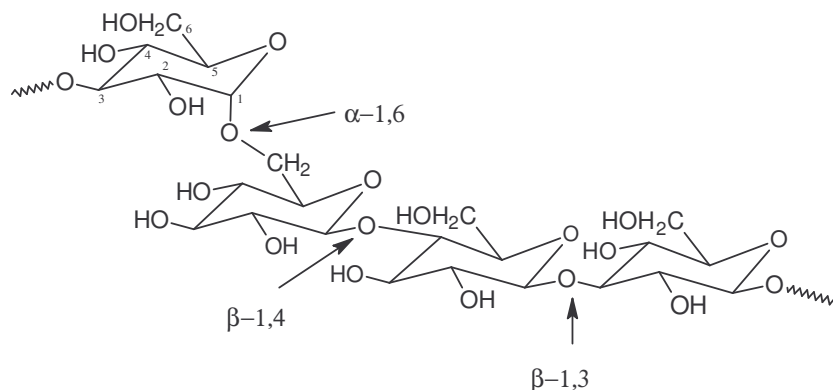


Figure 3. Illustration of some of the different linkages possible in a polysaccharide of pyranose monomers.

Polysaccharides vary according to their monosaccharide composition (i.e. type and ring size), as well as the mode of linkage (α or β), the linkage positions between monomers (e.g. 1,4 or 1,6), and the degree of branching (Lapasin and Pricl, 1995; Patamaporn Sukplang, 2000; Robyt, 1998). Polysaccharides can consist of 10 000 to 100 000 monosaccharides, giving rise to substances with molecular masses that can reach several million daltons (Da). Due to this variety of monomer, coupled with the differential bonding and branching between monomers, a powerful tool is necessary to determine the structure of polysaccharides.

Polysaccharides in the sugar industry

Polysaccharides have been found to be responsible for sugar processing problems; increasing sugar solution viscosities, reducing solution filterability and preventing sugar crystallisation (Morel du Boil, 1991; Ravno and Purchase, 2005). Such polysaccharides include dextran, starch and gums. These polysaccharides can be found throughout the sugar factory; including mixed juice, massecuite, syrup and final molasses.

Starch is a common component of sugar solutions. It is a polysaccharide that consists of amylose (20-30%) and amylopectin (80-70%). Amylose is a linear α -1,4 glucan, and amylopectin consists of α -1,4 glucose units with α -1,6 glucose branch points (Gidley, 1985). Figure 4 details the structure of these starch constituents and illustrates the difference in linkage information between the linear and branching glucose components.

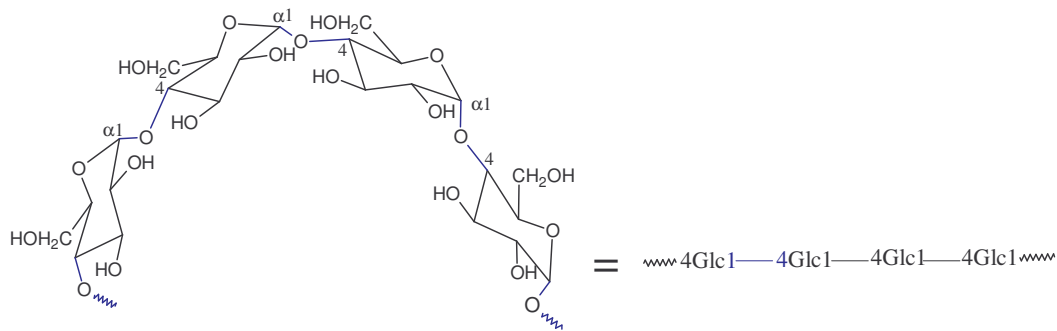
Starch levels vary widely between different cane varieties and growing conditions. A study on starch levels in South African molasses found average levels of 1340 ppm on Bx (0.2% on non-sucrose) (Lionnet and Sahadeo, 1998).

Dextran is a glucose-based bacterial polysaccharide which has always plagued the sugar industry (Bixler *et al.*, 1953; Morel du Boil and Wienese, 2002). These polysaccharides consist of α -1,6-linked D-glycopyranose backbones with a combination of α -1,2-, α -1,3- and α -1,4-linked branches. The branches can consist of a single glucose monomer or α -1,6-linked chains of the D-glucose units (Robyt, 1998). An example of a dextran molecule is depicted in Figure 5.

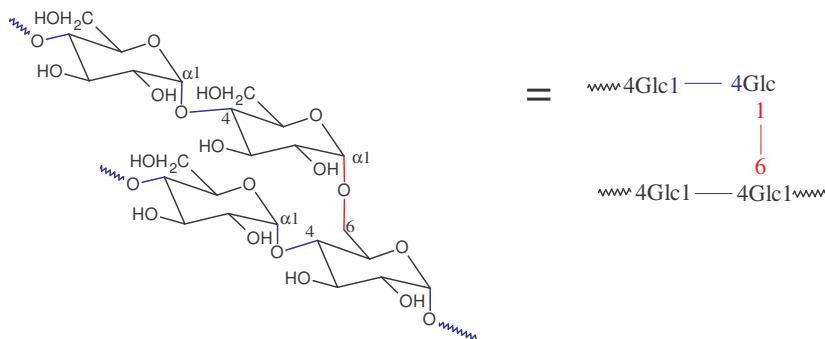
Compositional and structural characterisation of polysaccharides by GC-MS

The characterisation of polysaccharides is considered a complex and challenging task due to their inherent diversity and chain irregularities (Kajiwara and Miyamoto, 2004). Thin-layer chromatography has previously been used to qualify the monosaccharide composition in polysaccharides, but this method has proven to give non-discriminate results (Ruas-Madiedo and de los Reyes-Gavilan, 2005). Nowadays this type of analysis is conventionally performed using gas chromatography (GC) and liquid chromatography (LC) methods. With advancements in chromatographic and spectroscopic techniques such as GC with mass spectroscopy (GC-MS), more information including the monosaccharide composition, anomeric form, ring size, as well as linkage information can be determined (Ruas-Madiedo and de los Reyes-Gavilan, 2005; Walford, 2010). For an in-depth review on the use of GC-MS as a carbohydrate analysis tool in the sugar industry, refer to the work by Walford (2010).

This work presents the permethylation analysis protocol (Ciucanu and Caprita, 2007) used to depolymerise polysaccharides and volatilise the resultant monosaccharides in order to carry out a structural analysis. The steps involved are summarised in Table 1. The fine structure of the totally and partially methylated residues is then identified through their GC retention time indices (RTIs) and mass spectra when compared to authentic standards. Figure 6 outlines the procedural steps of this method.



Representative partial structure of amylose



Representative partial structure of amylopectin

Figure 4. Representative partial structures of amylose (with α -1,4 linkages in blue) and amylopectin (with α -1,6 branches in red), which make up starch (the hydrogens have been eliminated for clarity).

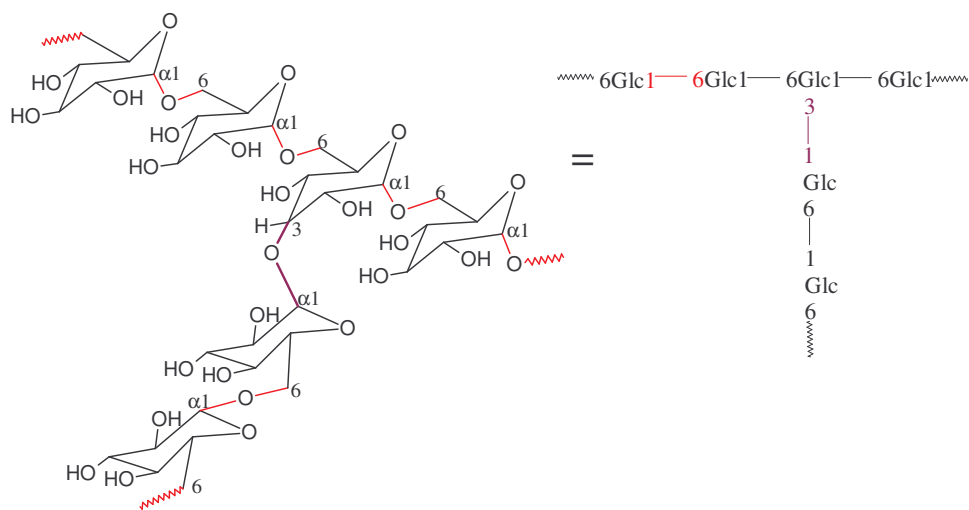


Figure 5. Structure of dextran showing the linear α -1,6 linkages in red with α -1,3 branching shown in purple (Adapted from Lapasin and Prici, 1995).

Table 1. Summary of the steps taken to conduct a permethylation analysis on polysaccharides.

Step	Method	Description
1	Permethylation of the –OH groups using Hakamori's reagent (methyl iodide with powdered sodium hydroxide (NaOH) in dimethyl sulfoxide (DMSO) (Hakomori, 1965)	Protection of the free –OH groups on the polymer
2	Acidic methanolysis using methanol and acetyl chloride	Hydrolysis of the polymer resulting in free –OH groups on the hydrolysed, partially methylated monosaccharides
3	Silylation of any partially methylated monosaccharides	Addition of silyl groups to the free –OH groups
4	Analysis of the totally and partially methylated residues using GC-MS	

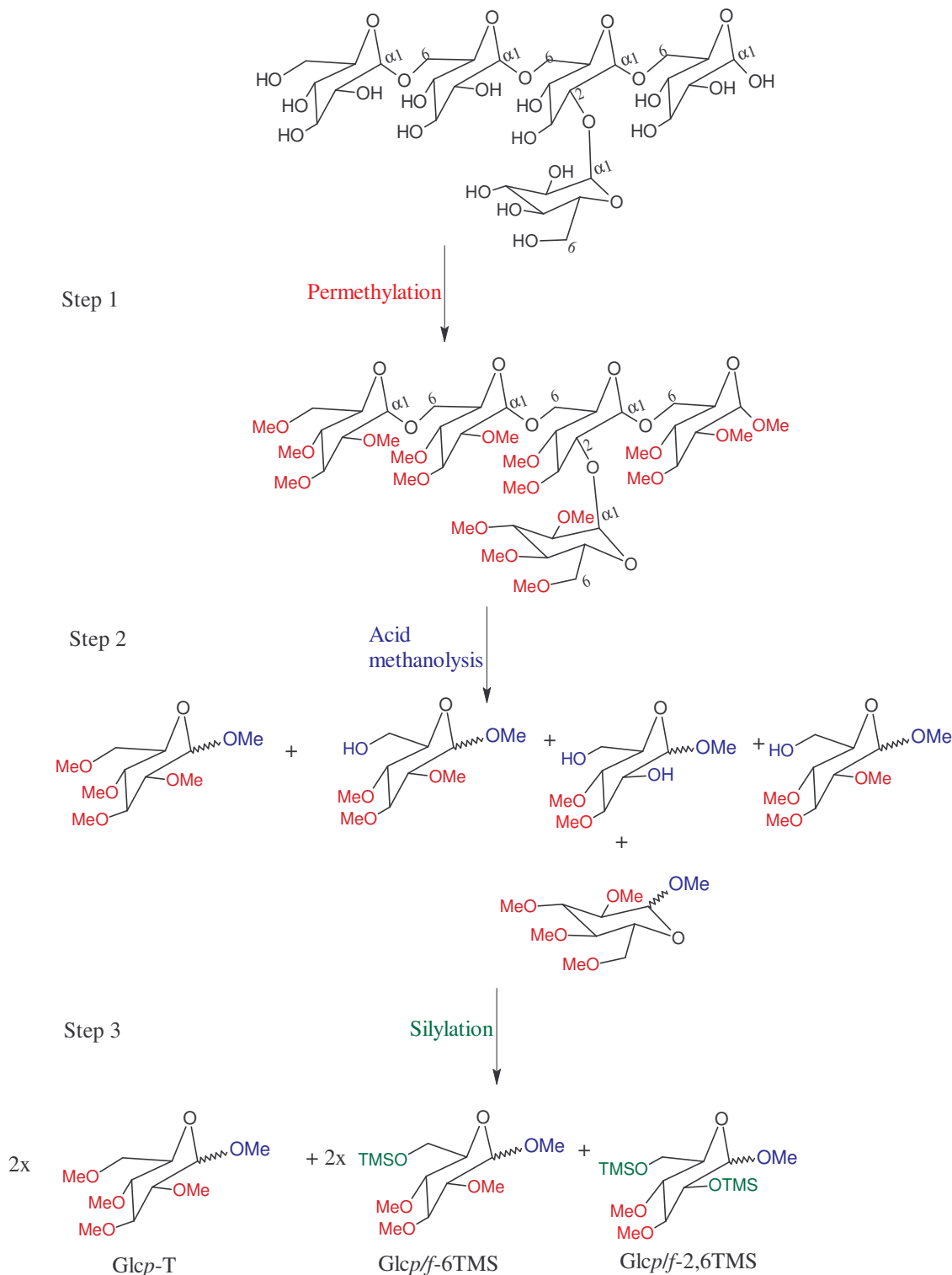


Figure 6. The reactions for the methylation analysis of a dextran-like pentasaccharide. Step 1: a quantitative methylation of the free hydroxyl groups (Red), Step 2: acid methanolysis of the glycosidic bonds. The α - and β -isomers of the resultant methyl glycosides equilibrate (Blue), Step 3: silylation of the free hydroxyl groups (Green). Me=methyl, TMS=trimethylsilyl, T=terminal, p =pyranose, f =furanose (Laine *et al.*, 2002).

Materials and Methods

Permethylation protocol

The cane starch sample and a total of 12 model compounds were permethylated, before undergoing methanolysis followed by silylation, in order to conduct a full structural analysis. For details of the compounds used and a full description of the permethylation protocol, refer to Appendix I.

Development of a mass spectral library

A spectral list for each peak arising from the gas chromatograms of the model compounds was created in the MS data handling software. The mass spectral data for each eluted GC peak was then analysed. The electron ionisation (EI) fragmentation patterns (mass spectra) of these peaks allowed for the identification of the specific monomers of all the model compounds. The RTIs of each eluted monomer on the GC were also saved. The RTIs were calculated as the retention time of each monomer's eluted peak over that of the elution time of the internal standard's peak on the GC.

A mass spectral library was built by importing the spectral lists of each monomer and the internal standard into the library manager. Refer to Appendix II for a summary of the mass spectral data obtained which was used to compile the model compound library.

Sample

Cane starch was previously isolated from NCo376 cane by the Chemical Division of the Sugar Milling Research Institute (SMRI). The sample was subjected to the permethylation protocol and was analysed by GC-MS. The GC elution peaks and MS spectral data obtained were processed using the reference data compiled in the mass spectral library of model compounds. These spectral lists were added as a compound table to the GC-MS method. Processing of the data using this method produced the report of the sample as seen in Appendix III.

Results and Discussion

The GC-MS results of the model compounds dextran, amylopectin and amylose are discussed in detail in this section, and are tabulated in Appendix II. The discussion is then directed to the GC-MS analysis of the cane starch sample and the structural features are compared to those of the model compounds analysed. Finally, the efficacy of the GC-MS model compound library is tested by processing the cane starch sample against it, and comparing the generated results.

Dextran

In Figure 7 the GC chromatogram of the model compounds dextran, amylopectin and amylose are illustrated with the relevant identified residues labelled. This gel permeation grade dextran model compound reveals eight relevant peaks of varying height in the GC chromatogram (Chromatogram 1). The first two peaks (17.086 min and 19.439 min) are identified as the terminal glucose units, methyl 2,3,4,6-tetra-*O*-methyl-glucopyranose (Glc_p-T). The α - and β -anomers of these terminal residues appear with RTIs of 0.41 and 0.46 respectively (refer to Appendix II). The height of these peaks indicates their low abundance compared to the major peaks. The major peaks arise from the liberated partially methylated intra-chain residues in

dextran (Figure 6, Step 2). These peaks (22.012 min and 23.988 min) have RTIs of 0.52 and 0.57 for the α - and β -residues, respectively. Since dextran comprises primarily of α -1,6-linked glucose units, these residues are attributable to the anomers of methyl 2,3,4-tri-*O*-methyl-6-*O*-(trimethylsilyl)-glucopyranoside (Glc_p-6TMS).

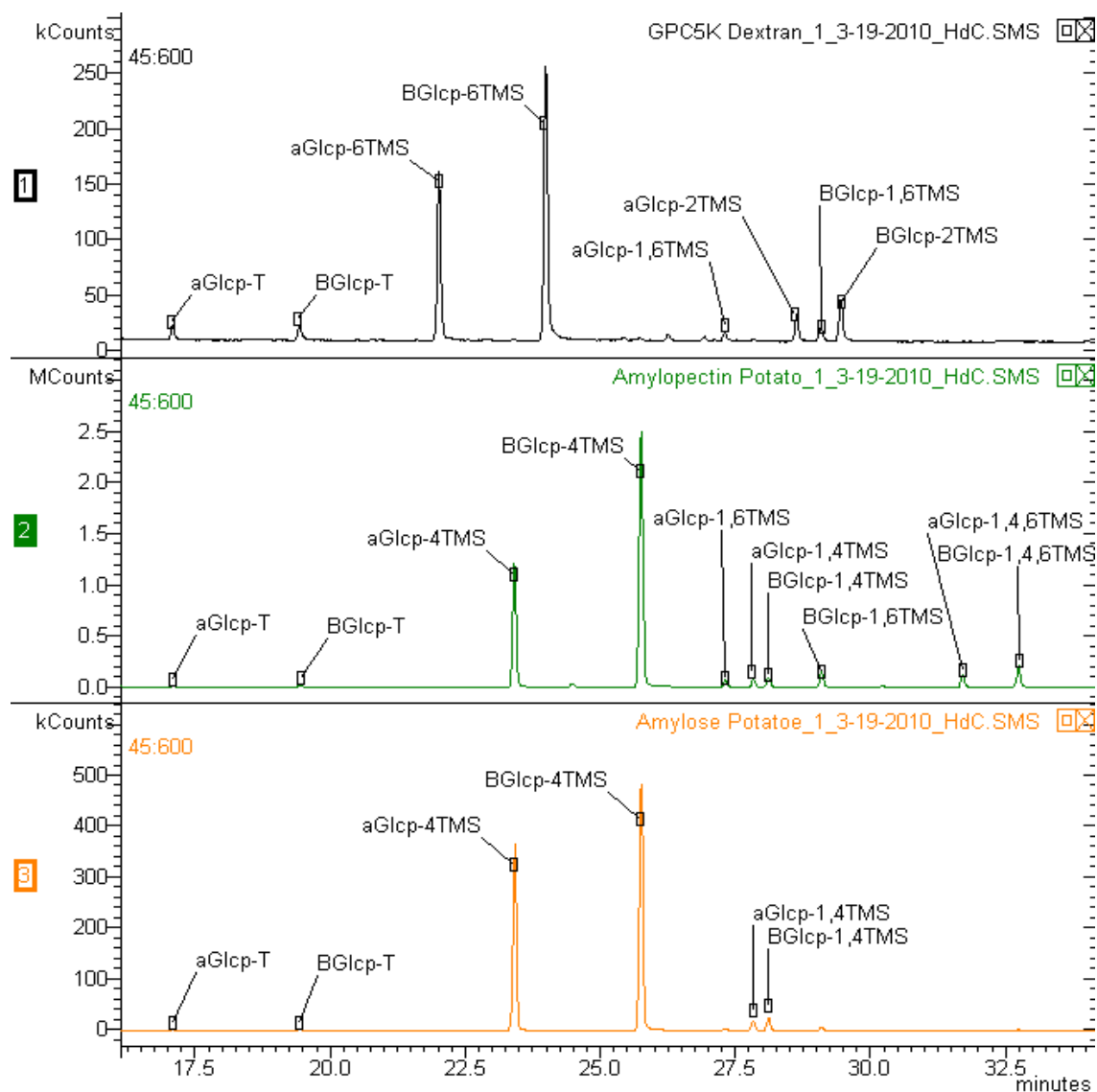


Figure 7. The GC chromatograms for the compositional analysis of model compounds gel permeation grade dextran (1), potato amylopectin (2) and potato amylose (3) are illustrated with the relevant identified residues shown. The mass spectra for these peaks were stored in the reference library.

Residues which arise from the intra-chain cleavage of **branch points** in the polysaccharide give rise to combinations of methyl 3,4,6-tri-*O*-methyl-2-*O*-(trimethylsilyl)-glucopyranoside (Glc_p-2TMS), methyl 2,4,6-tri-*O*-methyl-3-*O*-(trimethylsilyl)-glucopyranoside (Glc_p-3TMS)

or methyl 2,3,6-tri-*O*-methyl-4-*O*-(trimethylsilyl)-glucopyranoside (Glc p -4TMS) residues, depending on the branching present. The elution peaks with RTIs of 0.68 and 0.70 (28.638 min and 29.450 min) were identified by their mass spectra as α - and β residues of Glc p -2TMS, respectively. The additional set of eluted residues, with RTIs of 0.65 and 0.69 (27.314 min and 29.089 min) respectively, were first believed to be due to α -1,3- or α -1,4-linkages. However, on analysis of the mass spectra of these residues, two trimethylsilyl groups were revealed instead of only one - which arise due to the presence of moisture under these reaction conditions (Mikkola and Oivanen, 2009). Consequently, both the C-1 and C-6 positions of the intra-chain monosaccharides are hydroxylated and silylated (Figure 6, Steps 2 and 3), giving rise to the α - and β -anomers of 2,3,4-tri-*O*-methyl-1,6-bis-*O*-(trimethylsilyl)-glucopyranoside (Glc p -1,6TMS). The presence of moisture (due to normal waters of crystallisation and/or insufficiently dried starting materials) and the investigation into the mechanistic pathways of hydrolysis, however, will be the topic of further investigation.

Amylopectin

The amylopectin model compound reveals ten relevant peaks of varying height in the GC chromatogram (Figure 7, Chromatogram 2). As with dextran, the first two peaks are identified as the terminal α - and β -Glc p -T residues. Unlike dextran, the major peaks in amylopectin arise from the intra-chain α -1,4-linked glucose residues liberated during hydrolysis (Figure 6, Step 2). These peaks (23.413 min and 25.754 min) have RTIs of 0.52 and 0.57, and represent the α - and β -anomers of methyl 2,3,6-tri-*O*-methyl-4-*O*-(trimethylsilyl)-glucopyranoside (Glc p -4TMS), respectively.

Although amylopectin contains α -1,6-branch points, these particular residues (i.e. Glc p -6TMS) are not detected in the GC chromatogram as expected. However, the presence of these linkages is indicated by four of the remaining six residues. The next four consecutive peaks (and lack of peaks for Glc p -6TMS) in the amylopectin GC chromatogram arise again due to the interaction of moisture during the hydrolysis of the polysaccharide. Consequently, both anomers of Glc p -1,6-TMS branch residues and 2,3,4-tri-*O*-methyl-1,4-bis-*O*-(trimethylsilyl)-glucopyranoside (Glc p -1,4TMS) intra-chain residues arise. The former residues are equivalent to those which arise from dextran, whilst the latter residues (27.827 min and 28.112 min) with RTIs of 0.66 and 0.67, respectively, were not identified in dextran.

The final two residues (31.700 min and 32.732 min) also arise from inherent moisture interactions. During hydrolysis the C-1, C-4 and C-6 positions of branch point residues are all hydroxylated and subsequently silylated, giving rise to both the α - and β -anomers of 2,3-bis-*O*-methyl-1,4,6-tri-*O*-(trimethylsilyl)-glucopyranoside (Glc p -1,4,6TMS). In this instance this is considered a unique feature compared to this dextran model compound, as amylopectin has α -1,4-linkages with some α -1,6-branch points attached. These peaks are only possible in different dextran samples, if that particular dextran has α -1,4-branch points instead of α -1,2- or α -1,3-branches.

Amylose

The amylose model compound reveals six relevant peaks of varying intensity in the GC chromatogram (Figure 7, Chromatogram 3). As with dextran and amylopectin, the first two peaks are identified as the terminal α - and β -Glc p -T residues. The remaining four peaks are equivalent to the α - and β -Glc p -4TMS and α - and β -Glc p -1,4TMS residues established in amylopectin, and arise for the same reasons. Only these identified residues are expected in amylose, since this α -1,4-linked polysaccharide is without any branches.

Cane starch

The cane starch sample was treated in the same manner as the model compounds. On inspection of the GC chromatogram (Figure 8), the cane starch sample reveals ten relevant peaks. It appears that the features which the cane starch has in common with all the model compounds are the anomers of *Glc*p-T and *Glc*p-1,6TMS. Furthermore, the cane starch sample has the *Glc*p-4TMS and *Glc*p-1,4TMS features specific to amylose, in addition to the amylopectin specific residue of *Glc*p-1,4,6TMS. While it is evident that the only residues not present in the cane starch compound are the *Glc*p-2TMS anomers of dextran, this fact cannot preclude the presence of dextran in the cane starch sample, as a dextran with α -1,4-branches cannot be distinguished from amylopectin and amylose residues.

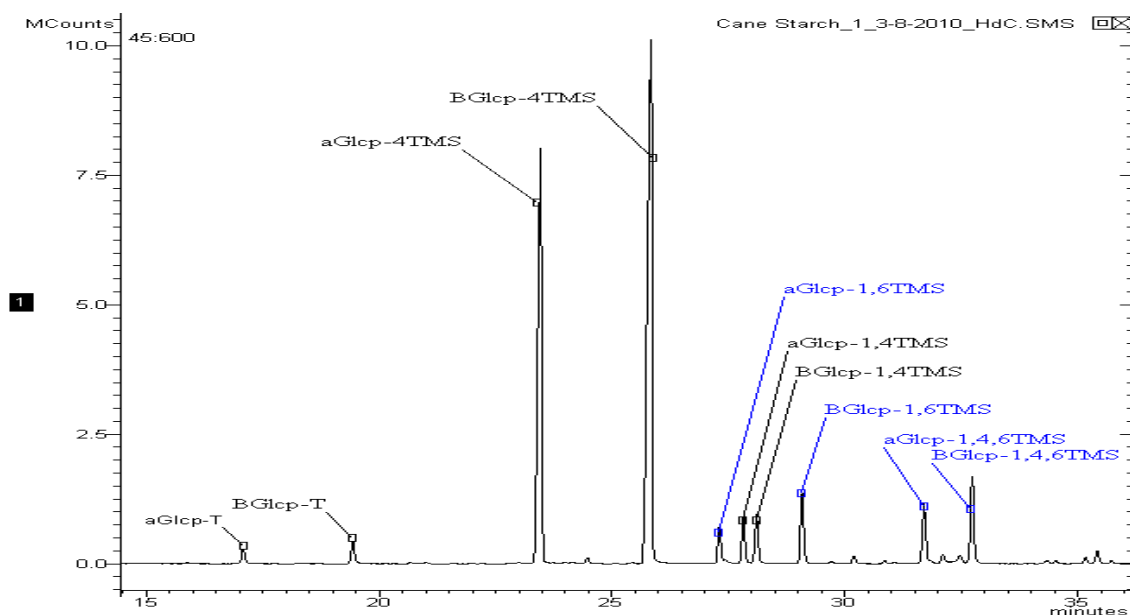


Figure 8. The structural analysis of cane starch reveals all the components found in amylopectin (all peaks) and amylose (black peaks only), highlighting the special features which arise from the permethylation treatment of the amylopectin component (blue) in this sample.

The efficacy of the GC-MS model compound library was then tested by processing the cane starch data against it. The identified residues listed in Table 2 reveal that all the peaks detected by observation were matched by the results generated from the GC-MS model compound library. The processed data confirms (what was originally deduced from the RTIs and mass spectra) that all ten residues of the amylopectin model compound are indeed present in the cane starch sample. However, on manual inspection alone the amylose-specific residues cannot be distinguished from amylopectin-specific residues in the sample. Once the data is processed against the GC-MS library, however, the areas of the identified peaks can be used to uncover more information regarding the sample's structure. For instance, after the identification of the *Glc*p-1,4TMS and *Glc*p-1,6TMS residues in the cane starch sample, and after comparison with the same in the amylopectin model compound (refer to Appendix III), the change in ratio of the *Glc*p-1,4TMS:*Glc*p-1,6TMS residues was determined. It was established that the cane starch sample consists of 20% more *Glc*p-1,4TMS to *Glc*p-1,6TMS residues, and confirms a composition of 80% amylopectin and 20% amylose.

Table 2. Summary of the processed GC chromatogram peak data for the cane starch sample treated by the permethylation method, and the GC-MS model library residues identified.

Retention time index/peak area % of anomers of the linkage type	Anomer*	Residue identified by inspection	Residue identified by the GC-MS model compound library
0.40/36	α	Glc p -T	Glc p -T
0.46/64	β	Glc p -T	Glc p -T
0.52/30	α	Glc p -6	Glc p -6
0.56/15	α	Glc p -4	Glc p -4
0.57/70	β	Glc p -6	Glc p -6
0.61/85	β	Glc p -4	Glc p -4
0.65/43	α	Glc p -1,6	Glc p -1,6
0.66/39	α	Glc p -1,4	Glc p -1,4
0.67/61	β	Glc p -1,4	Glc p -1,4
0.69/57	β	Glc p -1,6	Glc p -1,6
0.75/40	α	Glc p -1,4,6	Glc p -1,4,6
0.78/60	β	Glc p -1,4,6	Glc p -1,4,6

*Anomeric information was derived by work performed by Bleton *et al.* (1996).

Conclusions

It is shown in this study that the permethylation protocol enables depolymerisation of polysaccharides, allowing for the analysis of the hydrosylates by GC-MS. Together these methods allow the individual monosaccharides to be identified, not only by type and ring size, but also by their position and participation (terminal, intra-chain, or intra-chain branches) in the polysaccharide chain. It is also shown, whilst the presence of inherent (or other) moisture affects the type of residue produced during hydrolysis, that these specific residues produced are identifiable and useful in polysaccharide-specific determinations. It is demonstrated that, through the understanding of the permethylation protocol and the GC-MS method, that the RTIs and mass spectra of all the relevant model compound-specific features can be identified and stored in a reference library. This GC-MS model compound reference library developed at the SMRI proved effective in identifying all the residues in the cane starch sample, and further proved useful to aid in distinguishing differences between the amylopectin standard and cane starch samples by establishing the 80% amylopectin and 20% amylose composition of the latter. The future use of such a reference library would be to rapidly and accurately assist the researcher in identifying or confirming the structural components in other isolated oligo- or polysaccharide samples.

REFERENCES

- Bixler GH, Hines GE, McGhee RM and Shurter RA (1953). Dextran. *Industrial and Engineering Chemistry* 45(4): 692-705.
- Bleton J, Mejanelle P, Sansoulet J, Goursaud S and Tchaplal A (1996). Characterization of neutral sugars and uronic acids after methanolysis and trimethylsilylation for recognition of plant gums. *Journal of Chromatography A* 720(1-2): 27-49.
- Ciucanu I and Caprita R (2007). Per-O-methylation of neutral carbohydrates directly from aqueous samples for gas chromatography and mass spectrometry analysis. *Analytica Chimica Acta* 585: 81-85.
- Gidley MJ (1985). Quantification of the structural features of starch polysaccharides by n.m.r. spectroscopy. *Carbohydrate Research* 139: 85-93.
- Hakomori S-I (1965). A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. *Journal of Biochemistry* 55: 205-208.
- Kajiwara K and Miyamoto T (2004). Progress in Structural Characterization of Functional Polysaccharides. pp 1-11. In: S Dumitriu (Ed). *Polysaccharides: Structural Diversity and Functional Versatility* CRC Press, New York, USA.
- Laine C, Tamminen T, Vikkula A and Vuorinen T (2002). Methylation analysis as a tool for structural analysis of wood polysaccharides. *Holzforschung* 56(6): 607-614.
- Lapasin R and Prich S (1995). *Rheology of Industrial Polysaccharides: Theory and Applications*. First edition. Chapman & Hall, Glasgow. pp 1-46, 134-150, 163-214, 507-522.
- Lindhorst TK (2000). *Essentials of Carbohydrate Chemistry and Biochemistry*. Wiley-VCH, Weinheim. pp 3-31, 195-208.
- Lionnet GRE and Sahadeo P (1998). An analytical survey of final molasses from 15 cane producing countries. Sugar Milling Research Institute, Durban, South Africa.
- Mikkola S and Oivanen M (2009). Hydrolytic decomposition of glycosides in aqueous acids. *ARKIVOC*. Part iii: 39-53.
- Morel du Boil PG (1991). The role of oligosaccharides in crystal elongation. *Proc S Afr Sug Technol Ass* 65: 171-178.
- Morel du Boil PG and Wienese S (2002). Enzymatic reduction of dextran in process - laboratory evaluation of dextranases. *Proc S Afr Sug Technol Ass* 76: 435-443.
- Patamaporn Sukplang BS (2000). Production and characterization of a novel extracellular polysaccharide produced by *Paenibacillus valaei*, sp. Nov. Dissertation, University of North Texas, Denton, Texas, USA.
- Ravno AB and Purchase BS (2005). Dealing with dextran in the South African Sugar Industry. *Proc S Afr Sug Technol Ass* 79: 28-47.
- Robyt JF (1998). *Essentials of Carbohydrate Chemistry*. Springer-Verlag, Iowa. pp 22-39, 48, 157-202, 345-359
- Ruas-Madiedo P and de los Reyes-Gavilan CG (2005). Invited Review: Methods for the screening, isolation, and characterization of exopolysaccharides produced by lactic acid bacteria. *J Dairy Sci* 88(3): 843-856.
- Walford SN (2010). GC-MS as a tool for carbohydrate analysis in a research environment. *Proc Int Soc Sug Cane Technol* 27: 1-15.

Appendix I

This section details the permethylation protocol which is used as a standard procedure at the Sugar Milling Research Institute to determine the structural features of polysaccharide samples.

Method

All reagents used were supplied by Merck (except for the dimethyl sulfoxide, which is from The British Drug House, Ltd.) and were of analytical grade.

Table I.i. Model compounds which were permethylated and analysed by GC-MS.

Model compound	Type of saccharide	Type of linkages	Supplier
Glucose (Glc)	Monosaccharide	Terminal	Unknown
Cellobiose (Glc-Glc)	Disaccharide	β 1,4	Sigma
Amylose from potato	Polysaccharide containing Glc	α 1,4	Sigma
Amylopectin from potato	Polysaccharide containing Glc	α 1,4 with α 1,6 branches	Sigma
Maltotriose (Glc-Glc-Glc)	Trisaccharide	α 1,4	Sigma
Dextran from <i>Leuconostoc mesenteroides</i>	Polysaccharide containing Glc	α 1,6 (with α 1,4/ α 1,3/ α 1,2 branches)	Sigma
Panose (Glc-Glc-Glc)	Trisaccharide	α 1,4 and α 1,6	Sigma
Lichenan from <i>Cetraria islandica</i>	Polysaccharide containing Glc	β 1,3 and β 1,4	Sigma
Galactose (Gal)	Monosaccharide	N/A	Sigma
Melibiose (Glc-Gal)	Disaccharide	α 1,6	BDH
Lactose (Glc-Gal)	Disaccharide	β 1,4	Saarchem

Each sample and model compound (0.5 mg) was weighed into separate 4 cm³ screw cap vials. To each of these, 0.5 mL dry dimethyl sulfoxide (dried over 3Å molecular sieves), was added together with 100 μ L iodomethane (3 molar equivalents per millimole replaceable hydrogen). To this solution 60 mg of finely ground sodium hydroxide was added (in excess). The pellets of sodium hydroxide were ground in a hot and dry (heated to 120°C in an oven) mortar and pestle.

The capped vials were stirred on a vortex stirrer to achieve a suspension and were then left in a thermo-regulated sonicator bath at 50°C for 30 minutes. The derivatised samples were then left to cool. The per-*O*-methylated standards were then extracted using 1 cm³ water together with 1 cm³ dichloromethane. Once added the mixture was vortexed and left for 10 minutes to separate into two layers. The top aqueous layer was removed and discarded. The process was repeated by addition of 1 cm³ water. Excess water in the extracted permethylated standards was removed by addition of anhydrous sodium sulphate. The standards were transferred to 1 cm³ reactor vials (conical inserts with a screw cap). To each standard 5 μ L of sorbitol solution (7.5 mg in 25 cm³ methanol) was added as the internal standard. The standards were dried under a flow of dry nitrogen.

Methanolysis and silylation

The following steps (methanolysis and silylation) were required for all model compounds other than monosaccharides.

Methanolysis: A methanolic hydrochloride reagent was prepared by adding 140 μL acetylchloride to 4 cm^3 dried methanol (dried over 3 \AA molecular sieves). An aliquot of this solution (0.5 cm^3) was added to the dried sample. The vials were sealed with a silicone plug and then tightly capped with the screw-on lid. The vials were placed in a heating module, the base at each compartment being lined with a drop of cooking oil, and then they were left at 80°C for 20 hours. After the samples had cooled they were dried under a flow of dry nitrogen.

Silylation: Pyridine (0.5 cm^3), hexamethyldisilazane (0.45 cm^3) and trifluoroacetic acid (50 μL) were added to each vial, the vials were capped with Teflon-lined lids and placed in the sonicator bath at 80°C for 20 minutes. The cooled samples were then manually centrifuged and the supernatant transferred to 1 cm^3 gas chromatography vials, which were crimped closed.

Analysis

A Varian CP-3800 gas chromatograph fitted with a Varian 4000 GC-MS/MS mass detector was used for analysis. The details of the GC-MS conditions used are outlined in Table I.ii.

Table I.ii. GC-MS conditions used to analyse the derivatised model compounds.

Column	Factor four capillary column, VF-5ms (30 m x 0.25 mm x 0.25 μm)			
Injector temperature	260°C			
Injection mode	standard split/splitless			
Injection volume	1 μL			
Carrier gas	Helium			
Gas flow rate	1 cm^3/min			
Split ratio	Time (min)	Split	Split ratio	
	Initial	On	10	
	0.01	Off	Off	
	0.75	On	50	
Oven programme	Temp (°C)	Rate (°C/min)	Hold (min)	Total (min)
	100	-	2.00	2.00
	180	2.0	0.00	42.00
	300	40.0	2.00	47.00
Ion source	Electron ionisation			
Ionisation	5.00 min to 47.00 min			
Detected mass range	45 m/z to 600 m/z			

Appendix II

The information from the gas chromatograms and mass spectra of the eluted components of the model compounds is summarised in this section. This information was used to compile a reference library on the GC-MS software which can be used when determining the fine structures in unknown carbohydrate samples.

Model Compound GC-MS data

Model carbohydrate compounds were permethylated according to the protocol described in Appendix I, and the information summarised in Table II.i was used to compile the GC-MS reference library.

Table II.i. Summary of the GC retention times and mass spectral data obtained from permethylated model compounds used to compile a GC-MS reference library.

Model compound	Retention time (min)	Anomer	Linkage	Me positions	TMS positions	Ring size	Retention time index/peak area % of anomers of the linkage type
Glucose (Glc)	16.079	α	Glc-T	1,2,3,4,6	none	pyranose	0.41/45
	18.427	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/55
Amylose	17.090	α	Glc-T	1,2,3,4,6	none	pyranose	0.41/43
	19.447	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/57
	23.419	α	Glc-4	1,2,3,6	4	pyranose	0.56/41
	25.758	β	Glc-4	1,2,3,6	4	pyranose	0.61/59
	27.831	α	Glc-1,4	2,3,6	1,4	pyranose	0.66/46
	28.113	β	Glc-1,4	2,3,6	1,4	pyranose	0.67/54
Amylopectin	17.093	α	Glc-T	1,2,3,4,6	none	pyranose	0.40/36
	19.446	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/64
	23.413	α	Glc-4	1,2,3,6	4	pyranose	0.56/15
	25.754	β	Glc-4	1,2,3,6	4	pyranose	0.61/85
	27.303	α	Glc-1,6	2,3,4	1,6	pyranose	0.65/31
	27.827	α	Glc-1,4	2,3,6	1,4	pyranose	0.66/45
	28.112	β	Glc-1,4	2,3,6	1,4	pyranose	0.67/55
	29.092	β	Glc-1,6	2,3,4	1,6	pyranose	0.69/69
	31.700	α	Glc-1,4,6	2,3	1,4,6	pyranose	0.75/12
32.732	β	Glc-1,4,6	2,3	1,4,6	pyranose	0.78/88	
Cellobiose	16.074	α	Glc-T	1,2,3,4,6	none	pyranose	0.41/23
	18.435	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/77
	22.255	α	Glc-4	1,2,3,6	4	pyranose	0.56/40
	24.616	β	Glc-4	1,2,3,6	4	pyranose	0.61/60
Maltotriose	16.088	α	Glc-T	1,2,3,4,6	none	pyranose	0.41/40
	18.432	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/60
	22.300	α	Glc-4	1,2,3,6	4	pyranose	0.56/87
	24.698	β	Glc-4	1,2,3,6	4	pyranose	0.61/13
	24.638	β	Glc-4	1,2,3,6	4	pyranose	0.61/60

Continued...

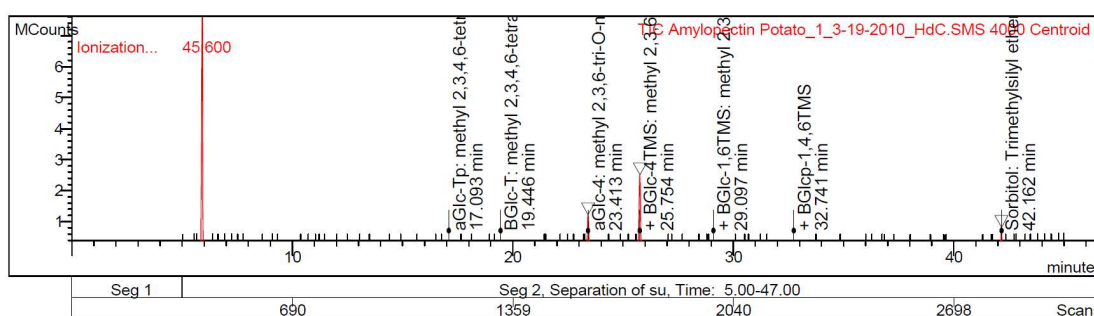
Model compound	Retention time (min)	Anomer	Linkage	Me positions	TMS positions	Ring size	Retention time index/peak area % of anomers of the linkage type
Dextran	17.086	α	Glc-T	1,2,3,4,6	none	pyranose	0.41/47
	19.439	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/53
	22.012	α	Glc-6	1,2,3,4	6	pyranose	0.52/38
	23.988	β	Glc-6	1,2,3,4	6	pyranose	0.57/62
	27.314	α	Glc-1,6	2,3,4	1,6	pyranose	0.65/35
	28.638	α	Glc-2	1,3,4,6	2	pyranose	0.68/39
	29.089	β	Glc-1,6	2,3,4	1,6	pyranose	0.69/65
	29.450	β	Glc-2	1,3,4,6	2	pyranose	0.70/61
Panose	16.066	α	Glc-T	1,2,3,4,6	none	pyranose	0.41/65
	18.386	β	Glc-T	1,2,3,4,6	none	pyranose	0.46/35
	20.865	α	Glc-6	1,2,3,4	6	pyranose	0.52/38
	22.248	α	Glc-4	1,2,3,6	4	pyranose	0.56/48
	22.853	β	Glc-6	1,2,3,4	6	pyranose	0.57/62
Lichenan	24.603	β	Glc-4	1,2,3,6	4	pyranose	0.61/52
	22.230	α	Glc-4	1,2,3,6	4	pyranose	0.56/30
	22.951	α	Glc-3	1,2,4,6	3	pyranose	0.57/26
	24.266	β	Glc-3	1,2,4,6	3	pyranose	0.60/75
Galactose (Gal)	24.585	β	Glc-4	1,2,3,6	4	pyranose	0.61/70
	18.159	α	Gal-T	1,2,3,5,6	none	furanose	0.44/82
	18.599	α	Gal-T	1,2,3,5,6	none	pyranose	0.45/28
	19.329	β	Gal-T	1,2,3,5,6	none	pyranose	0.47/72
Melibiose	19.468	β	Gal-T	1,2,3,5,6	none	furanose	0.48/18
	18.128	α	Gal-T	1,2,3,5,6	none	furanose	0.44/100
	18.618	α	Gal-T	1,2,3,4,6	none	pyranose	0.45/67
	19.359	β	Gal-T	1,2,3,4,6	none	pyranose	0.47/33
	20.890	α	Glc-6	1,2,3,4	6	pyranose	0.52/32
Lactose	22.892	β	Glc-6	1,2,3,4	6	pyranose	0.57/68
	18.082	α	Gal-T	1,2,3,5,6	none	furanose	0.44/100
	18.587	α	Gal-T	1,2,3,4,6	none	pyranose	0.45/22
	19.303	β	Gal-T	1,2,3,4,6	none	pyranose	0.47/78
	22.221	α	Glc-4	1,2,3,6	4	pyranose	0.56/55
	24.571	β	Glc-4	1,2,3,6	4	pyranose	0.61/45

The linkage information was deduced from knowledge of the model compound, whilst the ring size information was derived from the fragmentation patterns. The anomeric form of the monosaccharides was derived from work by done by Bleton *et al.* (1996).

Appendix III

This section shows a typical processed report using the compound table generated from the model compounds library search of the GC chromatogram of the cane starch sample. This report shows the cane starch sample.

Sample ID:	Amylopectin Potato	Operator:	HdC
Instrument ID:	Varian CP-3800	Last Calibration:	None
Acquisition Date:	3/19/2010 6:01 PM	Data File:	..._1_3-19-2010_hdc.sms
Calculation Date:	5/11/2010 3:54 PM	Method:	...compound table 5.mth
Inj. Sample Notes:	Amylopectin from potatoe permethylated, methanolysis, silylation with sorbitol internal std		

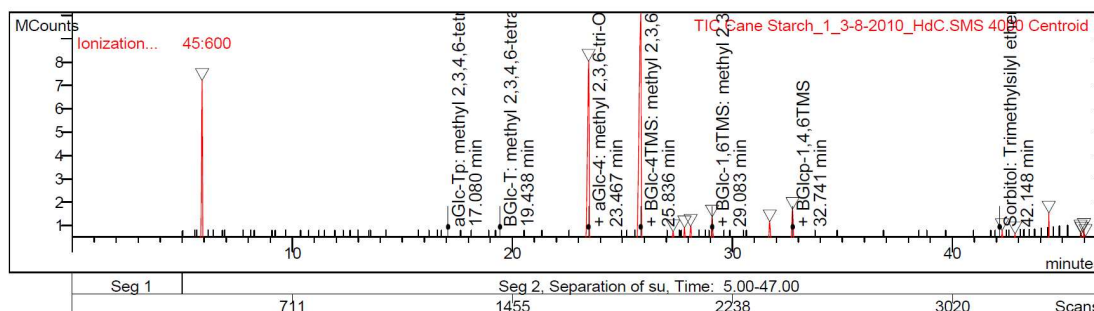


Target Compounds

#	RT	Peak Name	Res Type	Quan Ions	Area	Amount/RF
1	42.162	Sorbitol: Trimethylsilyl	Id.	73.1	580343	1 Counts
2	19.446	BGlc-T: methyl 2,3,4,6-t	Id.	88.2	34663	34663 Counts
3	23.413	aGlc-4: methyl 2,3,6-tri	Id.	88.1	1041620	1041620 Counts
4	25.754	BGlc-4TMS: methyl 2,3,6-	Id.	88.0	1854405	1854405 Counts
5	27.313	aGlc-1,6TMS	Id.	88.1	47760	47760 Counts
6	27.832	aGlc-1,4TMS: methyl 2,3,	Id.	88.1	44863	44863 Counts
7	28.115	BGlc-1,4TMS: methyl 2,3,	Id.	73.1	56226	56226 Counts
8	29.097	BGlc-1,6TMS: methyl 2,3,	Id.	73.1	99626	99626 Counts
9	31.707	aGlc-1,4,6TMS	Id.	146.3	99450	99450 Counts
10	32.741	BGlc-1,4,6TMS	Id.	146.3	142184	142184 Counts
11	22.011	aGlc-6p: methyl-2,3,4-tr	Miss.	88.1	0	0
12	23.988	BGlc-6p: methyl 2,3,4-t	Miss.	88.1	0	0
13	28.639	aGlc-2p: methyl 3,4,6 -t	Miss.	146.3	0	0
14	29.455	BGlc-2p: methyl 3,4,6-tr	Miss.	146.3	0	0
15	17.093	aGlc-Tp: methyl 2,3,4,6-	Id.	88.1	19414	19414 Counts

Figure III.i. Processed amylopectin data on the GC-MS software. Processing was based on the compound table compiled from the model compounds library search.

Sample ID: Cane Starch Operator: HdC
 Instrument ID: Varian CP-3800 Last Calibration: None
 Acquisition Date: 3/8/2010 11:39 PM Data File: ...h_1_3-8-2010_hdc.sms
 Calculation Date: 5/11/2010 3:51 PM Method: ...compound table 5.mth
 Inj. Sample Notes: Cane starch permethylated methanolysis and silylated with sorbitol.



Target Compounds

#	RT	Peak Name	Res Type	Quan Ions	Area	Amount/RF
1	42.148	Sorbitol: Trimethylsilyl	Id.	73.1	187481	1 Counts
2	19.438	BGlc-T: methyl 2,3,4,6-t	Id.	88.2	636336	636336 Counts
3	23.467	aGlc-4: methyl 2,3,6-tri	Id.	88.1	9178381	9178381 Counts
4	25.836	BGlc-4TMS: methyl 2,3,6-	Id.	88.0	12540035	12540035 Counts
5	27.309	aGlc-1,6TMS	Id.	88.1	467718	467718 Counts
6	27.824	aGlc-1,4TMS: methyl 2,3,	Id.	88.1	502781	502781 Counts
7	28.099	BGlc-1,4TMS: methyl 2,3,	Id.	73.1	628689	628689 Counts
8	29.083	BGlc-1,6TMS: methyl 2,3,	Id.	73.1	894377	894377 Counts
9	31.699	aGlc-1,4,6TMS	Id.	146.3	1017848	1017848 Counts
10	32.741	BGlc-1,4,6TMS	Id.	146.3	1212843	1212843 Counts
11	22.008	aGlc-6p: methyl-2,3,4-tr	Id.	88.1	7254	7254 Counts
12	23.988	BGlc-6p: methyl 2,3,4-t	Miss.	88.1	0	0
13	28.639	aGlc-2p: methyl 3,4,6 -t	Miss.	146.3	0	0
14	29.455	BGlc-2p: methyl 3,4,6-tr	Miss.	146.3	0	0
15	17.080	aGlc-Tp: methyl 2,3,4,6-	Id.	88.1	391276	391276 Counts

Figure III.ii. Processed cane starch data on the GC-MS software. Processing was based on the compound table compiled from the model compounds library search.