

Structural analysis of *in vitro* produced microbial α -D-glucans

Sander Sebastiaan van Leeuwen

ISBN	978-90-393-4710-2
Circulation	300
Fonts inside	Garamond, Palatino Linotype
Font cover	Franklin Gothic
Cover image	Rusted cover: edited photo of Bioshock™ case (2K Games™) Stainless glucose inlay: rendered with Blender 2.45
Printer	PrintPartners Ipskamp BV, Enschede

Structural analysis of *in vitro* produced microbial α -D-glucans

Biopolymers synthesised from sucrose by using native and engineered *Lactobacillus reuteri* glucansucrase enzymes

Structuuranalyse van *in vitro* geproduceerde microbiële α -D-glucanen

Biopolymeren gesynthetiseerd uit sucrose met behulp van natieve en gemodificeerde *Lactobacillus reuteri* glucansucrase enzymen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties in het openbaar te verdedigen
op woensdag 12 december 2007 des middags te 2.30 uur

door

Sander Sebastiaan van Leeuwen

geboren op 6 juli 1977 te Rotterdam

Promotor: Prof.dr. J.P. Kamerling

Co-promotor: Dr. G.J. Gerwig

Het onderzoek beschreven in dit proefschrift werd mogelijk gemaakt met financiële steun van Senter Novem (EETK01129).

It is better to keep your mouth closed and let people think you are a fool than to open it and remove all doubt

-Mark Twain-

Light thinks it travels faster than anything but it is wrong. No matter how fast light travels, it finds the darkness has always got there first, and is waiting for it

-Terry Pratchett-

Contents

Chapter 1	General Introduction : Structural analysis and bio-engineering of microbial exopolysaccharides	9
Chapter 2	Development of an NMR structural-reporter-group concept for the primary structural characterisation of α -D-glucans	31
Chapter 3	Structural analysis of the α -D-glucan (EPS180) produced by the <i>Lactobacillus reuteri</i> strain 180 glucansucrase GTF180 enzyme	57
Chapter 4	Structural analysis of the α -D-glucan (EPS35-5) produced by the <i>Lactobacillus reuteri</i> strain 35-5 glucansucrase GTFA enzyme	87
Chapter 5	Structural analysis of the bio-engineered α -D-glucan (mEPS-PNNS) produced by a triple mutant of the glucansucrase GTF180 enzyme from <i>Lactobacillus reuteri</i> strain 180	115
Chapter 6	Structural characterisation of α -D-glucans produced by mutant glucansucrase GTF180 enzymes of <i>Lactobacillus reuteri</i> strain 180	139
Summary		167
Samenvatting		171
Samenvatting in normaal Nederlands		177
Dankwoord		183
Bibliography		187
Curriculum Vitae		189

A common mistake that people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools

-Douglas Adams-

Chapter 1

General Introduction :

Structural analysis and bio-engineering of microbial
exopolysaccharides

Introduction

Micro-organisms have been applied in food processing for many centuries. Even before Marcus Terentius Varro (116-27 BC) described his suspicion of “certain minute creatures which cannot be seen by the eyes, that float in the air and enter the body through the mouth and nose and there cause serious diseases”,¹ micro-organisms were responsible for fermentation processes, leading to improved foods. Beer and wine were probably already known 8,000 years ago, and written records of beer recipes were found in ancient Egypt. Later in history, but still before the existence of micro-organisms was proven, lactic acid bacteria (LAB) were used to ferment milk into yoghurts. Lactic acid bacteria are nowadays commonly used to improve the texture, rheology, and acidity of milk products. Lactate, one of the fermentation products from lactose, is responsible for the acidity, increasing the shelf-life of the fermented milk. The texture and rheology improvements are, among others, caused by exopolysaccharides produced by the lactic acid bacteria. Carbohydrate biopolymers from plants (e.g. cellulose, starch, alginate, carrageenan), animals (e.g. chitin, chitosan), and micro-organisms (e.g. xanthan, acetan, curdlan, dextran), in modified or native form, found their way in food processing.²⁻⁴ These polysaccharides are dissolved, suspended, or dispersed in watery foods, influencing the texture of the product.^{5,6}

The majority of the polysaccharides currently in use are plant polysaccharides, like starches, from cultivated crops. Some chemically modified polysaccharides are being used also,⁶ however, the disadvantage is that their use is restricted. All polysaccharides, even natural ones, applied outside their native crop-product are considered additives and need specific mention on the food packaging. Especially the European Union, Australia, New-Zealand, and Israel are restrictive in allowing food-additives, giving rise to the special indication of E-numbers, which is a subset of the ISN-numbered food-additives recognised worldwide. The application of polysaccharides as thickeners and stabilisers is increasing, beyond the capacity of plant crops, and consequently increasing the price of polysaccharides. This opens a growing market for alternative polysaccharides from microbial sources.

Bacteria can be divided into two classes, Gram-positive and Gram-negative. The division was initially based on a staining method devised in 1884 by the Danish bacteriologist Hans Christian Joachim Gram. The Gram-staining protocol uses a dye that stains some bacteria and leaves others unstained, probably based on interactions with the

peptidoglycan layer which is thicker in Gram-positive bacteria. Nowadays the “Gram-positive” and “Gram-negative” denominations are still used, but the characterisation is based on cell architecture (most notably the differences in the cell-wall) and no longer solely on the original staining method.

Microbial polysaccharides can be divided in capsular polysaccharides (CPSs), lipopolysaccharides (LPSs) which only occur in Gram-negative bacteria, and exopolysaccharides (EPSs). CPSs occur in many micro-organisms and are associated with or covalently bound to the cell envelope. CPSs of pathogenic species (K-antigen) and LPSs (O-antigen) have been extensively researched, due to pharmaceutical interests.⁷ EPSs are excreted into the surroundings of the cell and occur regularly in bacteria and microalgae, but are rarely observed in yeasts and fungi. A selection of EPSs from Gram-negative bacteria has been reviewed by Sutherland on different occasions.⁸⁻¹¹

In this thesis, the structural analysis of α -D-glucans produced by glucansucrase (GTF) enzymes from *Lactobacillus reuteri* strains and bio-engineered EPS structures is described. In this introductory chapter an overview is given of EPS structures, functions, and industrial applications in general, and from LAB in particular. The biosynthesis of exopolysaccharides is described, followed by methods of polysaccharide modification (i.e. chemical, enzymatic, and genetic-engineering). Finally, techniques of EPS structural analysis are described and the challenges of glucan analysis in particular are highlighted.

Exopolysaccharides from lactic acid bacteria

Exopolysaccharides occur in many micro-organisms, particularly bacteria and algae.^{6,12} Some species of yeast and fungi were also found to produce EPSs.⁶ Although the work on Gram-negative species of bacteria is more extensive, due to the immunogenic importance of O-antigen LPSs and K-antigen CPSs, this chapter will focus on LAB, which are Gram-positive. EPSs can be divided into two groups; homopolysaccharides (HoPSs) and heteropolysaccharides (HePSs). Many types of EPS, both HoPSs and HePSs, are excreted by LAB, varying in composition and molecular mass, ranging from 1.0×10^4 to 6.0×10^6 Da.^{2,13} HoPSs consist of one type of monosaccharide residue, most commonly D-glucopyranose or D-fructofuranose.¹⁴ HoPSs can be linear or branched and usually contain one or two types of glycosidic linkage, and are rarely built-up from repeating units.¹⁴⁻¹⁷ However, some repeating unit structures are known for HoPSs (e.g.

pullulan, galactan¹⁸). HePSs are generally built-up of repeating units, consisting of three up to eight monosaccharides.^{2,3} The monosaccharides can be present in both α - and β -anomeric forms. Most importantly, in HePSs more than one type of monosaccharide residue occurs. The first repeating unit of a LAB EPS was elucidated by Doco *et al.* in 1990.¹⁹ Since then many repeating units from LAB HePSs have been analysed and published. Extensive reviews on these structures are available.^{2,3,20}

Structures and functions of microbial EPS

The *in vivo* functions of EPSs are mostly of a protective nature. The slime-layer produced with EPSs seems to provide protection against dehydration by retaining water.²¹ The thickness of the slime also impedes diffusion of potential threats, including toxins (e.g. anti-biotics, sulphur dioxide), but also migration of phagocytes and predators.²¹ Moreover, the EPS structures are often capable of forming a complex with divalent cations, thus preventing diffusion of toxic metal ions completely. The sticky nature of some EPS slimes was found to be responsible for adhering to surfaces. Mutan, a HoPS produced by *Streptococcus mutans* in the oral cavity, was found to function as adherent in dental plaque.²² The EPSs could also be involved in cellular recognition processes; the multivalency of presented epitopes on a long chain of repeating units has a clear advantage over proteins carrying a single N-glycan antenna. The EPS slime is not likely to function as a food-reserve, since most microbes do not have the machinery to catabolise their own EPS product.⁵

The functions of EPSs in nature are determined by their physico-chemical properties. One of the elements of importance is their molecular mass, influencing rheological properties. But also the three-dimensional structure of the EPSs, defined by intra-molecular interactions, charge, and monosaccharide sequence, is of significant influence. For many properties the ability to form inter-molecular interactions between chains, metal-ions (e.g. calcium), peptides, or proteins (e.g. lectin binding), are important in, for example, biofilm formation and quorum sensing.¹¹ All these properties are essentially determined by the monosaccharide sequence of the EPS and understanding of the desired physical properties is based on understanding the structure and interactions of the EPS. Since these properties are ultimately determined by the primary structure of the repeating unit, changes in the repeating unit sequence can have tremendous effect on the properties of an EPS solution.⁶

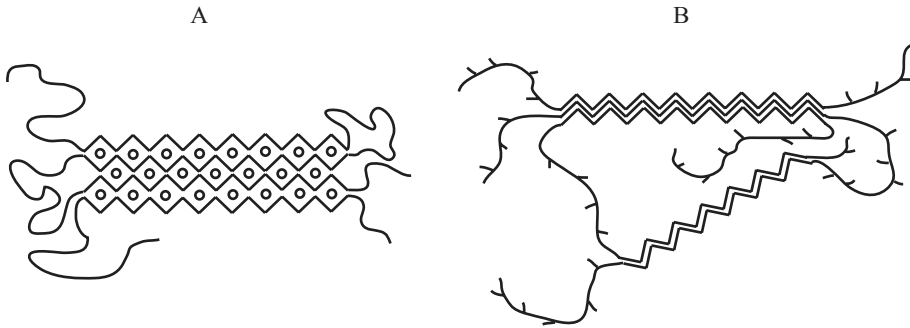


Figure 1. (A) Representation of the egg-box model for association of polyguluronate sequences by interaction between Ca²⁺-ions and alginate strands.²⁹ (B) Representation of gel-forming by aggregation of unsubstituted mannan regions in galactomannans.³¹

Industrial applications of EPS

The physico-chemical properties that lie at the basis of EPS functions in nature are also interesting for certain industrial uses. Most HePSs in use by the industry to date are derived from Gram-negative bacteria. The EPS produced by *Xanthomonas campestris*, for example, is a high-molecular-mass polymer with anionic groups. Xanthan has been well studied and has a low production cost.^{9,23} HePSs produced by LAB have not yet found industrial applications as purified compounds. Difficulties lie in the mass-production of purified EPSs, due to the low biosynthesis rates, which typically amounts to 50-1500 mg/L of culture, as opposed to up to 50 g/L for xanthan.^{5,24-27} Purification of the EPSs from the culture is often a difficult process, which will not benefit the costs of production.

Viscosity - One of the interesting properties of EPS solutions is their viscosity. In most cases, the viscosity is a non-specific property, i.e. function of temperature, concentration, shear, and molecular mass.²⁸ In solution most EPSs act as random coiled chains. At a certain critical concentration these chains get entangled and viscosity increases steeply with increase in EPS concentration. For some EPSs this model does not suffice to explain their viscous properties, in particular pseudoplastic properties (shear thinning). Pseudoplasticity manifests itself in a thinning of the solution upon increase of the shear rate.⁶ Pseudoplasticity is of great importance for some industrial applications. In the oil industry, for example, pseudoplastic solutions are used to improve oil-recovery,²³ but these solutions are also used in dairy products, where pseudoplasticity has been linked to desired texture of the product.

Gel forming - One of the industrially important properties of EPSs is the ability to form strong cross-linked networks, i.e. gels. There are two mechanisms for jellification, (i) interactions of EPSs containing negatively charged groups with cations, and (ii) interactions between neutral EPS chains. An example of interaction between negatively charged EPSs and divalent cations is found in alginate gels. Alginates contain (β 1-4)-linked D-mannuronic acid and (α 1-4)-linked L-guluronic acid residues. Blocks of polyguluronic acid sequences are capable of complexing with Ca^{2+} ions, via the egg-box model (Figure 1A),²⁹ and thus forming connection nodes between strands.³⁰ The mechanism of interaction between neutral polysaccharide strands involves van der Waals interactions and hydrogen bonds. Good examples of gel-forming neutral polysaccharides are the galactomannans. The galactomannans are a family of polysaccharides with different ratios of galactose and mannose residues. Unsubstituted mannose stretches are capable of forming a three-dimensional network of interacting chains (Figure 1B), whereas the galactose stretches or branched regions cannot adopt this conformation. This means that the amount of mannose and length of mannose stretches in the galactomannan determines the ability to form a gel, as well as the strength of that gel.^{31,32}

LAB polysaccharides - Through the use during many centuries of LAB in dairy fermentations all over the world, they have been awarded the GRAS (generally recognised as safe) status.³³ This means that these bacteria are allowed in natural food-production. Addition of bacterial cultures, sometimes in mixed-culture fermentations leads to foods with improved rheological properties.³⁴ Depending on the product intended, so called ropy or mucoid cultures can be applied, leading to the specific desired properties. For yoghurts, cow-milk is fermented with *Streptococcus salivarius* ssp. *thermophilus* and *Lactobacillus delbrueckii* ssp. *bulgaricus*.³⁵ The thermophilic *S. thermophilus* strain is widely applied, since the addition of non-natural stabilisers in yoghurts are prohibited in the EU.² Stability of yoghurt during production processes is an important issue. To confront this, a common industrial technique involves the addition of solids in the form of fat or sugar. The popular demand for low-fat and 'low-carb' foods opens opportunities for non-fat and non-digestible-sugar additions, like HePSs or LAB cultures. The disadvantage of using thermophilic strains, however, lies in the low production rates of HePSs under the fermentation conditions.³⁶

Some HoPSs from LAB have already found use in industrial applications.³⁷ Derivatives of microbial dextrans are being used in gel filtration material (e.g. Sephadex), but also as oral hematinics (e.g. Niferex).³⁸ Levan, a fructan, has interesting thickening

properties for food applications. Recent research into pre- and pro-biotics have led to the use of specific LAB (e.g. *Bifido* bacteria, *Lb. casei* strains) or specific oligosaccharides that are believed to be beneficial for gut flora, to be applied in yoghurt or buttermilk/kefir drinks, which are subsequently advertised as health-foods.³⁹ LAB as well as certain EPS and oligosaccharide products have been implicated in improved resistance to disease,⁴⁰ lowering cholesterol,⁴¹ or preventing cancer.⁴²

Outside the food industry some EPSs are promising for certain applications. In the mining industry certain polysaccharides seem useful in recovery of ultra-fine ore suspended in waste water, and thus improving the yield from the rough ore.⁴³ In medical applications some EPSs seem promising as bio-degradable scaffolding materials for regenerating tissues,⁴⁴ or in complex with gadolinium as a contrast agent in medical scanning techniques.⁴⁵ In the coating industry EPSs seem promising in anti-corrosive applications,⁴⁶ or as an anti-fouling agent.⁴⁷ Recently, chitosan, a charged polysaccharide derived from the chitin shell of crustaceans (e.g. shrimp), was found to have promising properties for electro-conductive applications in fuel-cells.^{48,49} It is possible that certain bacterial EPSs with proper charge distribution may be suitable for similar applications.

Polysaccharide biosynthesis in LAB

Heteropolysaccharides

Synthesis of hetero-exopolysaccharides in LAB can be divided into two parts. The first part is responsible for providing the monosaccharide building-blocks. To this end a sugar carbon source (generally galactose, glucose, and lactose) is transported into the cell from the medium.⁵⁰ The sugars are taken up into the catabolic system of the cell, which is connected to the HePS production via the sugar-1-phosphate pathway. The main carbon source is diversified into a large variety of nucleotide-sugars, required for the HePS synthesis. A large variety of monosaccharide residues and derivatives have been observed in heteropolysaccharides. A detailed review of these processes is available and will not be repeated here.^{19,50} For an overview of monosaccharides and derivatives observed so far, see Table 1.

The second part of the HePS synthesis uses the nucleotide-sugars to build a repeating unit, which is translocated across the membrane and subsequently polymerised. In many cases the study of EPS biosynthetic mechanisms was driven by the implication

Table 1. Monosaccharide residues and derivatives observed in LAB heteropolysaccharides.

Monosaccharides
α,β -D-galactopyranose
α,β -D-galactofuranose
α,β -L-rhamnopyranose
α,β -D-glucopyranose
β -D-ribofuranose
β -D-glucopyranuronic acid
α,β -N-acetyl-D-glucopyranosamine
α,β -N-acetyl-D-galactopyranosamine
L-fucose
D-mannose
D-fructose
3',9'-dideoxy-D-threo-D-altro-nononic acid-(2→6)- α -D-GlcP
<i>m</i> -glycerol-3-phosphate
Derivates
O-acetate
(-3)-phosphate-(1-

of the EPS in virulence of pathogenic Gram-negative bacteria.⁵¹ In these bacteria the basic repeating unit was assembled on the cytoplasmic membrane by sequential glycosyltransferase activity. The first glycosyltransferase in the sequence transfers a glycosyl residue from a glycosyl-1-phosphate onto an undecaprenyl-phosphate carrier.¹⁰ Repeating units are translocated across the membrane, where polymerisation takes place. The polysaccharide is finally released into the medium or attached to the cell wall to form a capsule.

Some research has been directed to understanding similar mechanisms in pathogenic strains of the Gram-positive *Streptococcus* genus.⁵²⁻⁵⁴ These studies showed that in these strains the biosynthesis of the respective repeating units also took place on a lipid carrier at the cytoplasmic side of the membrane. Here, also the repeating unit was translocated across the membrane, polymerised, and either exported or attached to the cell wall.

Making use of homology studies and functional genetic studies of EPS-related genes in *S. thermophilus* and *L. lactis* species showed that the systems here were similar to those in Gram-negative and pathogenic *Streptococci*.^{4,55-61} The first gene described for LAB was that of *S. thermophilus* Sfi6, by Stingle *et al.* in 1996 (Figure 2).⁵⁵ The 14.5-kb gene-cluster comprised of thirteen genes, which are transcribed from a single

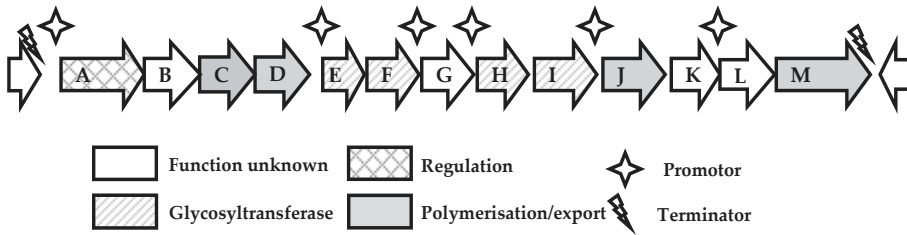


Figure 2. Schematic organisation of the *S. thermophilus* Sfi6 eps gene-cluster.⁵⁵ Functions assigned to the different genes are indicated by arrow filling, promoters and terminators are indicated by icons. Legends to the arrows and icons are included in the figure.

promoter, *epsA*. The EPS genes for other food-grade micro-organisms were studied, including *S. thermophilus* NCFB 2393,⁶² and *Lc. lactis* NIZO B40.⁵⁷ The *eps* gene-clusters showed incredible similarity in all strains studied so far. All gene-clusters have an operon structure. The genes are all oriented in the same direction, leading to a single mRNA strand to be transcribed.^{55,57} Later studies by Stinglee *et al.* showed that in *S. thermophilus* Sfi6 EpsE is responsible for the transfer of the first sugar onto the lipophilic carrier.⁵⁶ The transfer of a Gal unit from Gal-1-P onto the carrier is subsequently followed by the transfer of GalNAc by EpsG, Glc by EpsI, and finally a branching Gal by EpsF (Figure 3). Of the thirteen genes situated on the gene-cluster only four are glycosyltransferases. The other genes are involved in regulation, chain-length determination, polymerisation, and export. Comparison of the *S. thermophilus* Sfi6 gene-cluster with other gene-clusters showed similar organisations.⁶¹ Many of the genes have not yet been completely characterised and no definitive function can be assigned. Especially the mechanisms for polymerisation, chain-length determination, and export are still not well understood.

The consistency in high-molecular masses of EPSs produced indicates the presence of a mechanism that determines the desired chain-length. Most gene-clusters have two or three genes, close to the promoter gene, whose functions are predicted to be involved in chain-length determination and polymerisation of the EPS. This system closely resembles ExoP in *Sinorhizobium meliloti*.⁶³ Most gene-clusters contain proteins with a significant homology with the export proteins Wzx, Wzy, and Wzz, involved in O-antigen polymerisation in *E. coli*.^{61,64} Recent publications of the 3D structure of the Wza translocon for *E. coli* capsular polysaccharides,^{65,66} and its association with the Wzx, Wzy, Wzz complex may provide novel insights into the polymerisation and export mechanisms in *E. coli*, and in homology for LAB.

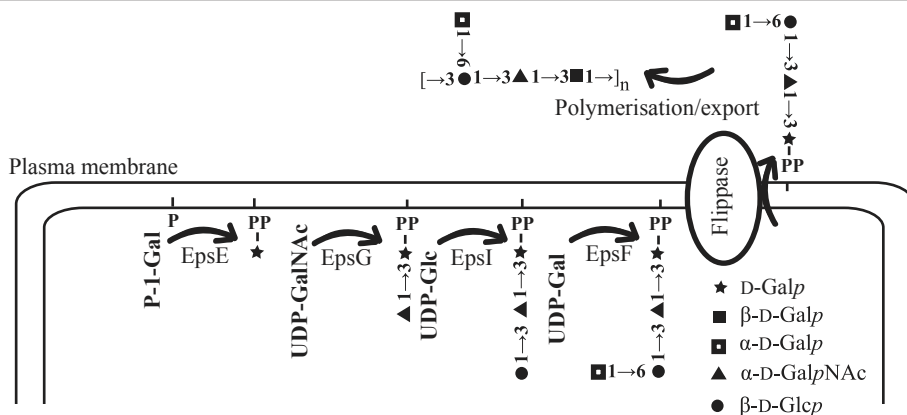


Figure 3. Representation of the proposed biosynthesis of the exopolysaccharide produced by *S. thermophilus* Sf6.⁵⁶

Homopolysaccharides

Some of the LAB HoPSs, like β -glucans produced by *Pediococcus* and *Lactobacillus*,^{67,68} and galactan synthesised by *Lc. lactis* ssp. *cremoris* H414¹⁸ are produced through similar mechanisms as HePSs, described in section 3.1. The majority of LAB HoPSs, however, are β -fructans and α -glucans. These HoPSs are produced by sucrose enzymes that are excreted by the LAB and attached to the cell wall. In the presence of sucrose, these enzymes were also found to be released from the cell wall into the medium. Sucrase enzymes use the energy in the sucrose glycosidic linkage to synthesise their EPS product, in contrast to the ATP-consuming energy-intensive synthesis of HePSs. The Nomenclature Committee of the International Union of Biochemistry and Molecular Biology recognises a classification of sucrase enzymes, based on the reaction catalysed and the product specificity. All sucraes are classified in Enzyme Class (EC) 2, which is the class of transferase enzymes, subclass 4 (glycosyltransferases), sub-subclass 1 (hexosyltransferases), i.e. EC 2.4.1.x.⁶⁹ Alternatively, the <http://www.cazy.org/> (Carbohydrate Active enZyme) website classifies glycoside hydrolase enzymes into 100 families based on amino acid sequences.⁶⁹ In view of their sequence similarities, glucansucraes and fructansucraes have been included in GH70 and GH68, respectively. Families with related mechanisms or structures, but also close evolutionary relation, are grouped into clans.

Fructans - There are two types of fructans produced by LAB, levans and inulins. Levans are synthesised by levansucrase (E.C. 2.4.1.10), which is a cell-wall-associated enzyme catalysing the transfer of D-fructose residues from sucrose. Levans are characterised by a majority of (β 2-6)-glycosidic linkages and have been observed in several *Streptococcus* strains,⁷⁰⁻⁷⁴ *Leuconostoc mesenteroides*,⁷⁵ *Lb. reuteri* 121,^{76,77} and *Lb. sanfranciscensis* LTH2590.⁷⁸ Inulins contain mainly (β 2-1) glycosidic linkages and are synthesised by inulosucrase enzymes found in several species of *Streptococcus*, *Leuconostoc*, and in *Lb. reuteri* 121.^{71,79,80} Other fructan-producing strains have been identified, where the distribution of glycosidic bonds is yet to be determined.⁸¹

Glucans - Glucansucrases (GH70) from LAB occur in a wide variety, producing different types of EPSs. Best known is the dextran class of HoPSs, containing mostly (α 1-6)-linked residues, with (α 1-3), (α 1-2), and sometimes (α 1-4) branches.⁵ Dextran was first identified as a thickener for cane sugar syrups,⁸² the glucan was named dextran for the positive optical rotation of dextran solutions. Many strains of LAB have been found to produce dextran, particularly *Le. mesenteroides* strains.⁵ Mutan was named for the *Streptococcus mutans* OMTZ176 from which it was first identified.⁸³ Mutan consists predominantly of (α 1-3)-linked glucose residues (up to 90%). Côté and Robyt isolated an α -glucan from *Le. mesenteroides* NRRL B-1335,⁸⁴ which turned out to be composed of alternating (α 1-3)- and (α 1-6)-linked glucose residues, and was subsequently named alternan. A recently discovered EPS with mainly (α 1-4)-linked glucose residues from *Lb. reuteri* was named reuteran.⁷⁶ Later studies showed that reuteran contained a significant amount of 6-substituted and 4,6-disubstituted glucopyranose residues, besides 4-substituted glucopyranose residues.⁸⁵ For a schematic overview of these classes of LAB α -D-glucans, see Figure 4.

While most strains produced only one glucansucrase enzyme, some contained genes for more than one glucansucrase, e.g. *Le. mesenteroides* NRRL B1299,^{86,87} *Le. mesenteroides* NRRL B-512F,^{16,75,88-90} *S. mutans* GS-5,^{71,76,91} and *S. salivarius* ATCC 25975.⁹² Glucansucrases from *Streptococci* are generally produced constitutively, whereas *Leuconostoc* usually produce glucansucrases specifically induced by the presence of sucrose.^{69,93} Until recently, best studied glucan producing LAB were *Leuconostoc* and *Streptococcus* species. In 1999, Van Geel-Schutten *et al.* identified reuteran production by *Lb. reuteri* 121,⁷⁶ and since then also dextran, and mutan have been found in *Lactobacilli*.^{94,95} In 2002, Kralj *et al.* identified the glucansucrase enzyme responsible for the production of reuteran.⁸⁵ This was followed by the identification of seven glucansucrase genes in six *Lactobacillus*

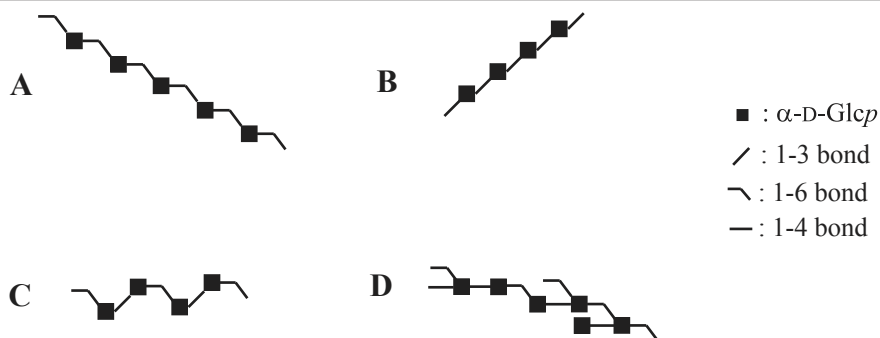


Figure 4. Schematic structures of (A) dextran (α 1-6), (B) mutan (α 1-3), (C) alternan (α 1-3)/(α 1-6), and (D) reuteran (α 1-4)/(α 1-6).

strains, i.e. *Lb. reuteri* 121, *Lb. reuteri* ML1, *Lb. reuteri* 180, *Lb. sakei* Kg15, *Lb. fermentum* Kg3, and *Lb. parabuchneri* 33.^{94,95}

The GH70 family belongs, together with GH77 (4- α -glucanotransferases) and GH13 (starch modifying enzymes) to the GH-H clan, containing a (β / α)₈ barrel catalytic domain, with an Asp, Glu, Asp catalytic triad, situated on β 4, β 5, and β 7 strands, respectively. The GH-H clan has a retaining reaction mechanism. The GH13 family, containing among others α -amylases and cyclodextrin glucosyltransferases (CGTase), is a well studied class of enzymes, with 3D structures available.^{96,97} The α -amylase structure has four conserved regions (I to IV), with seven conserved amino acid residues,⁹⁸ including the catalytic triad. Of these seven conserved residues in GH13, six are also present in GH70, and a different seventh conserved residue is present.⁹⁹ In glucansucrases, like *Lb. reuteri* 121 GTFA, the four conserved regions are circularly permuted, and as a consequence region I is positioned C-terminally from regions II, III, and IV (Figure 5).¹⁰⁰

Making use of homology with the GH13 family, some interesting regions of the GTFA could be identified for further study. Kralj *et al.* identified several non-catalytic amino acid residues adjacent to the catalytic D1133, the putative transition state stabilising residue, that determine the linkage specificity of the glucansucrase GTFA enzyme.¹⁰¹ Through site-directed mutagenesis of these amino acid residues, they were able to transform the reuteransucrase into a dextransucrase.¹⁰¹ Generally, the glucansucrase enzymes have one catalytic domain with a single active site,¹⁰² even though several of them produce highly-branched EPSs, or a linear EPS with two different linkage types.^{103,104} One known exception is the glucansucrase from *Le. mesenteroides* NRRL B-1299, which carries a second C-terminal catalytic domain, responsible for the branching reaction.¹⁰⁴

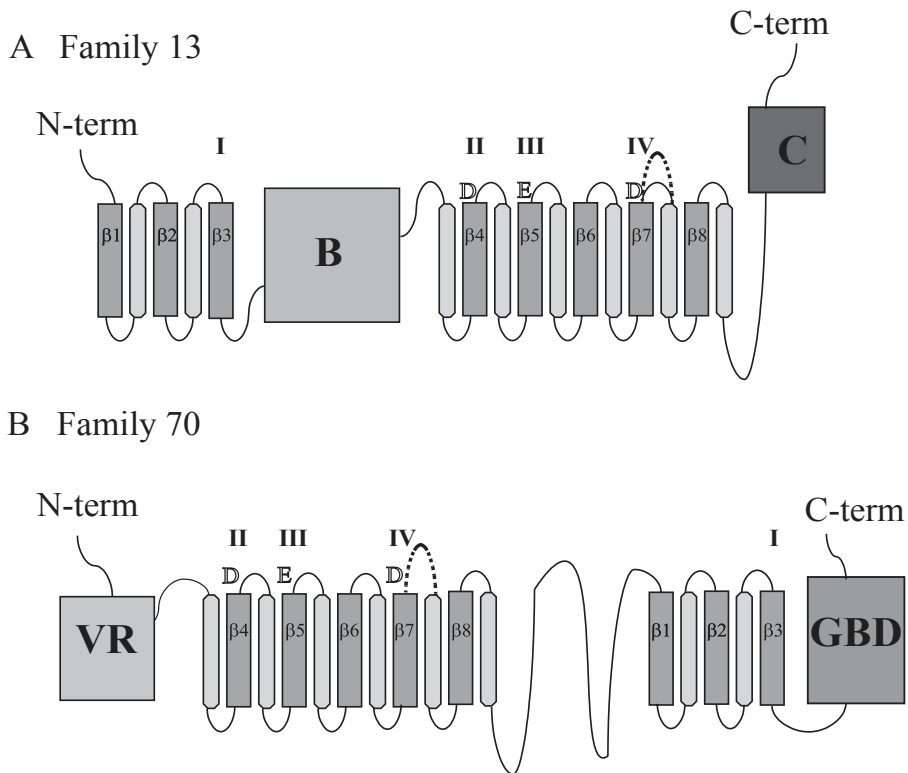


Figure 5. Topology diagrams of members of (A) the GH13 family and (B) the GH70 family. The catalytic domain of family GH13 enzymes have a $(\beta/\alpha)_8$ barrel structure, starting with β -strand 1 and ending with α -helix 8. The B domain is located between β -strand 3 and α -helix 3. Glucan-sucrases have a putative $(\beta/\alpha)_8$ barrel catalytic domain that is circularly permuted.¹⁰⁰ Here the catalytic domain starts with α -helix 3 and ends with β -strand 3. Between α -helix 8 and β -strand 1, a large stretch of unknown function is located. The locations of the four conserved regions (I - IV) are indicated, with the catalytic residues in regions II, III, and IV. B, B domain; C, C domain; GBD, glucan binding domain; VR, variable region.

Polysaccharide modification

Chemical modifications

In order to improve the physico-chemical properties desired for specific applications of the EPSs, several strategies have been employed in altering the EPS structures. The earliest methods of modification were treatments with chemicals, to induce specific alterations. A treatment of branched dextrans with a strong base and iron salts led to the production of dextran-iron complexes (e.g. Niferex), applicable as an iron supplement for anaemic patients.³⁸ Iron-dextran complexes have also found interest as possible contrasting agents in X-ray or MRI scanning diagnostics.⁴⁵ Other modifications

include treatment with periodate, in order to increase the amount of free hydroxyl groups in pestalotan (a (β 1-3) glucan backbone with 60% (β 1-6) branches, some of which short chains), to increase supposed antitumor activity.¹⁰⁵ Periodate oxidation was also applied on scleroglucan (a (β 1-3) glucan backbone with single (β 1-6) side groups), to improve gelling properties.¹⁰⁶ Alteration of dextrans by carboxymethylation, using bromoacetic acid,^{46,107} led to interesting structures that could be immobilised on surfaces, to prevent bio-fouling.¹⁰⁷ TEMPO-mediated oxidation of insoluble chitin, cellulose, and starch was shown to result in more soluble polysaccharides.¹⁰⁸ A chitin modified by this method was found to have interesting cholic-acid binding properties.¹⁰⁹

Enzymatic modifications

Altering the EPS structures with specific enzymes is also an option. However, the application is limited, as enzyme preparations are generally expensive and rarely come in bulk-packaging. To reliably modify a large batch of EPS with specific alterations in mind may be difficult and expensive. The requirement for specific enzymes to be available, with the correct substrate-specificity, has led to limited applications so far. Use of starch-modifying enzymes from plant and bacterial sources, has proven interesting for altering linear starch into amylopectin-like structures, with tailored degrees of branching.¹¹⁰ Treatment of alginates with mannuronan C-5 epimerase, created polysaccharides with high content of guluronic acid,^{111,112} leading to an enhanced ability to form strong gels,¹¹¹ and enhancing solubility at low pH.¹¹² Modifications of LAB HePSs with side-chain trimming enzymes, like β -galactosidase, altered the thickening properties of the EPSs.¹¹³

Bio-engineering

With the advances in elucidating the *eps* gene-clusters in Gram-negative and LAB bacteria, speculations have been made about possible genetic engineering to create tailor-made EPSs.^{2,3,10,11} Some attempts have been made to delete a specific glycosyltransferase from a gene-cluster. One success in Gram-negative bacteria has led to the trimming of the xanthan side-chain of the last residue, via glycosyltransferase knock-out. In LAB an *eps* gene-cluster was successfully transferred from *S. thermophilus* Sfi6 to *L. lactis* MG1363.¹¹⁴ Since *L. lactis* MG1363 did not have the C4-epimerase, required to produce UDP-GalNAc, this residue was not incorporated into the repeating unit, however, Gal was transferred instead. Moreover, it turned out that the *N*-acetyl moiety was involved in the activity of the branching glycosyltransferase on the adjacent residue. As a result

the transfected *L. lactis* strain produced a linear trisaccharide repeating unit, instead of the original, branched tetrasaccharide repeating unit. Remarkably, the *S. thermophilus* Sfi6 gene-cluster enzymes did produce a polysaccharide and successfully exported it. This indicates that the polymerisation and export genes in the *eps* gene-cluster do not have a narrow specificity.

Despite these successful attempts, not many advances have been made in the promised market of tailor-made EPSs. This is not entirely surprising, considering the complexity of the systems involved in HePS production.^{2,3} There are so many genes involved in regulation of the HePS production in LAB, that modifications in the ensemble involved in producing the HePSs might as just as readily disable any EPS production as result in a modified EPS to be produced. The recent 3D structure of the Wza translocon for cell-wall polysaccharides,^{65,66} believed to be similar to the systems used in EPS export, showed only a narrow channel available, i.e. 17 Å in diameter. The Wza translocon is thought to be ubiquitous for polysaccharide export, and it only just accommodates the pentasaccharide repeating unit in the example studied. This would mean that adding to the repeating unit might impede or even prevent export of the bio-engineered EPSs.

For HoPSs, on the other hand, these problems do not apply. In general, HoPSs are produced by a single enzyme, encoded in a single open reading frame (ORF). Studies of glucansucrase enzymes revealed that the *gtf* genes are relatively easy to clone and increase transcription.^{85,94,101} The signal-peptide, targeting the enzyme for excretion and attachment to the membrane can be deleted, as well as the N-terminal variable region (possibly involved in membrane attachment) without loss of function. In some cases the glucansucrase enzyme lost some affinity for the substrate, and subsequently reducing the production rate of EPS. Using the N-terminally truncated version of GTFA, Kralj *et al.* have shown that with a few specific modifications in non-catalytic residues, the reuteransucrase enzyme can change into a dextransucrase enzyme.¹⁰¹ These results are promising for possible bio-engineering targets of other glucansucrase enzymes.

Structural analysis of EPS

In 1990, the first primary structure of a LAB polysaccharide repeating unit (*S. thermophilus*) was determined.¹⁹ Since then many more HePS repeating units have been elucidated.³ Analysis of exopolysaccharide structures involves many chemical techniques. General methods include monosaccharide analysis,¹¹⁵ methylation analysis,¹¹⁵⁻¹¹⁷ partial

acid hydrolysis, and 1D and 2D NMR spectroscopy. Monosaccharide analysis reveals the composition of the repeating unit by identifying and quantifying the monosaccharides present. This analysis can also be used to confirm the homopolysaccharide nature of an EPS or to determine carbohydrate content, to verify purity of the samples. Methylation analysis adds to this the knowledge of substitution patterns. By using ^1H and ^{13}C NMR spectroscopy on the polysaccharide, the anomeric configuration is obtained as well as information on the proton chemical shifts of each residue (from ^1H - ^1H TOCSY NMR), linkage patterns (from ^1H - ^1H ROESY or NOESY NMR), and substitution information (^{13}C - ^1H HSQC and HMBC). Making use of partial acid hydrolysis, a repeating unit can be separated and analysed to verify the repeating unit sequence. In specific cases other techniques may also be useful, like acetolysis, uronic acid degradation, de-*N*-acetylation, periodate oxidation,¹¹⁸ and enzymatic degradation. These methods allow for (semi-)specific cleavage of bonds and reveal structural information of the EPS repeating unit.

In HoPSs, structural analysis is faced with some different challenges. Where HePSs rarely have overlapping anomeric signals in ^1H NMR spectroscopy, in HoPSs overlapping of anomeric signals is guaranteed. The lack of repeating units in HoPSs induces another challenge. So far only initial structural studies have been performed on HoPS structures,¹⁴⁻¹⁷ due to the semi-random distribution of linkage types and branching.

Aim and outline of this thesis

Microbial exopolysaccharides (EPSs) are useful for the food industry as viscosifying, stabilising, emulsifying, and gelling agents. The improvement of texture in fermented dairy products by generally recognised as safe (GRAS) lactic acid bacteria, is caused in part by the *in situ* production of EPSs. Most studied are heteropolysaccharides (HePSs) produced by LAB, and many studies have been performed into their structure-function relationships. Many parts of the biosynthesis of these EPSs have been elucidated in recent years. In case of homopolysaccharides (HoPSs) only initial structural studies have been performed.

The research described in this thesis was part of the EET (Economy, Ecology, Technology) programme “Bioprimer” (EETK01129), researching LAB α -D-glucan EPSs as anti-corrosive additives for use in heavy-duty coatings. The aim of the study was to perform complete structural characterisations of α -D-glucans, produced by

glucansucrase (GTF) enzymes from *Lactobacilli*. Some α -D-glucans have been indicated as promising anti-corrosives, whereas others showed no anti-corrosive activity in a pilot study. To understand the mechanism of anti-corrosive activity, detailed structural knowledge of these glucans is required.

In order to perform the challenging analysis of α -D-glucans, an NMR structural-reporter-group concept was established (Chapter 2), based on known di- and trisaccharide standards. Making use of the structural-reporter-group concept, the primary structure of EPSs produced *in vitro* by the glucansucrase (GTF180) enzyme from *Lb. reuteri* strain 180 (Chapter 3; **EPS180**) and the glucansucrase (GTFA) enzyme from *Lb. reuteri* strain 35-5 (Chapter 4; **EPS35-5**), using sucrose as a substrate, were analysed. Based on all data gathered visual presentations, that include all identified structural features, were formulated. During the analysis of each EPS, new structural-reporter-group signals were added to the concept. Using the enhanced structural-reporter-group concept the EPS product produced from sucrose by a triple mutant of the glucansucrase (GTF180;V1027P:S1137N:A1139S) enzyme from *Lb. reuteri* strain 180 (**mEPS-PNNS**) was analysed (Chapter 5). Again a picture was formulated, that includes all structural elements indentified, and the structural-reporter-group concept was complemented. In Chapter 6, structural characterisations were performed on α -D-glucans produced from sucrose by 12 mutant glucansucrase GTF180 enzymes of *Lb. reuteri* strain 180. Mutations were selected based on sequences observed in other glucansucrase enzymes, i.e. reuteransucrase (NNS), mutansucrase (NNV), alternansucrase (YDA), and dextransucrase (SEV) enzymes, as well as single mutant variants of these sequences. Making use of the structural-reporter-group concept the ^1H NMR spectroscopic data were interpreted, and combined with methylation analysis data to formulate visual representations, depicting all structural elements identified for the 12 mutant EPSs.

References

1. Varro, M. T. *On agriculture* Book I, **27 BC**, Translation: Hooper, W. D.; Ash, H. B. *Loeb Classical Library*, Cambridge, MA, **1934**, 283, Ch. 12, pp 2.
2. De Vuyst, L.; Degeest, B. *FEMS Microbiol. Rev.*, **1999**, *23*, 153-177.
3. De Vuyst, L.; de Vin, F.; Vaningelgem, F.; Degeest, B. *Int. Dairy J.*, **2001**, *11*, 687-707.
4. Van Kranenburg, R.; Boels, I. C.; Kleerebezem, M.; de Vos, W. M. *Curr. Opin. Biotechnol.*, **1999a**, *10*, 498-504.
5. Cerning, J. *FEMS Microbiol. Rev.*, **1990**, *87*, 113-130.
6. Sutherland, I. W. In *Biotechnology of microbial exopolysaccharides. Vol. 9*; Cambridge University Press, Cambridge, **1990**.
7. Whitfield, C.; Roberts, I. S. *Mol. Microbiol.*, **1999**, *31*, 1307-1319.
8. Sutherland, I. W. *Annu. Rev. Microbiol.*, **1985**, *39*, 243-270.
9. Sutherland, I. W. *Biotech. Adv.*, **1994**, *12*, 393-448.
10. Sutherland, I. W. *Int. Dairy J.*, **2001**, *11*, 663-674.
11. Sutherland, I. W. *Trends Microbiol.*, **2001**, *9*, 222-227.
12. Sutherland, I. W. *Adv. Microb. Physiol.*, **1972**, *8*, 143-213.
13. Cerning, J. *Lait*, **1995**, *75*, 463-472.
14. Monsan, P. F.; Bozonnet, S.; Albenne, C.; Joucla, G.; Willemot, R.-M.; Remaud-Simeon, M. *Int. Dairy J.*, **2001**, *11*, 675-685.
15. Monchois, V.; Willemot, R.-M.; Monsan, P. F. *FEMS Microbiol. Rev.*, **1999**, *23*, 131-151.
16. Funane, K.; Ishii, T.; Matsushita, M.; Hori, K.; Mizuno, K.; Takahara, H.; Kitamura, Y.; Kobayashi, M. *Carbohydr. Res.*, **2001**, *334*, 19-25.
17. Argüello-Morales, M. A.; Remaud-Simeon, M.; Pizzut, S.; Sarçabal, P.; Willemot, R.-M.; Monsan, P. F. *FEMS Microbiol. Lett.*, **2000**, *182*, 81-85.
18. Gruter, M.; Leefflang, B. R.; Kuiper, J.; Kamerling, J. P.; Vliegthart, J. F. G. *Carbohydr. Res.*, **1992**, *231*, 273-291.
19. Doco, T.; Wieruszetski, J.-M.; Fournet, B.; Carcano, D.; Ramos, P.; Loones, A. *Carbohydr. Res.*, **1990**, *198*, 313-321.
20. Laws, A. P.; Marshall, V. M. *Int. Dairy J.*, **2001a**, *11*, 709-721.
21. Ceri, H.; McArthur, H. A. I.; Whitfield, C. *Infect. Immun.*, **1986**, *51*, 1-5.
22. Costerton, J. W.; Cheng, K.-J.; Geesey, G. G.; Ladd, T. I.; Nickel, J. C.; Dasgupta, M.; Marrie, T. J. *Annu. Rev. Microbiol.*, **1987**, *41*, 435-464.
23. Sandford, P. A.; Baird, J. In *The Polysaccharides. Vol. 2*; Aspinall, G. O., Ed.; Academic Press: New York, **1983**; pp 411-490.
24. Cerning, J.; Marshall, V. M. E. *Rec. Res. Dev. Microbiol.*, **1999**, *3*, 195-209.
25. De Vuyst, L.; Vanderveken, F.; van de Ven, S.; Degeest, B. *J. Appl. Microbiol.*, **1998**, *84*, 1059-1068.
26. Degeest, B.; de Vuyst, L. *Appl. Env. Microbiol.*, **1999**, *65*, 2863-2870.
27. Degeest, B.; Vaningelgem, F.; de Vuyst, L. *Int. Dairy J.*, **2001**, *11*, 747-757.
28. Morris, E. R.; Cutler, A. N.; Ross-Murphy, S. B.; Rees, D. A. *Carbohydr. Pol.*, **1981**, *1*, 5-21.
29. Grant, G. T.; Morris, E. R.; Rees, D. A.; Smith, P. J. C.; Thom, D. *FEBS Lett.*, **1973**, *32*, 195.
30. Morris, E. R.; Gidley, M. J.; Murray, E. J.; Powell, D. A.; Rees, D. A. *Int. J. Biol. Macromol.*, **1980**, *2*, 327-330.
31. Dea, I. C. M.; Morris, E. R.; Rees, D. A.; Welsh, E. J.; Barnes, H. A.; Price, J. *Carbohydr. Res.*, **1977**, *57*, 249-272.
32. Dea, I. C. M.; Clark, A. H.; McCleary, B. V. *Carbohydr. Res.*, **1986**, *147*, 275-294.
33. Stancioff, D. J.; Renn, D. W. *ACS Symp. Ser.*, **1975**, *15*, 282-295.
34. Bouzar, F.; Cerning, J.; Desmazed, M. *J. Dairy Sci.*, **1997**, *80*, 2310-2317.
35. Berkman, T.; Bozoglu, T. F.; Özilgen, M. *Eng. Micr. Tech.*, **1990**, *12*, 138-140.
36. Degeest, B.; Vaningelgem, F.; de Vuyst, L. *Int. Dairy J.*, **2001**, *11*, 747-757.
37. Roller, S.; Dea, I. C. M. *Crit. Rev. Biotechnol.*, **1992**, *12*, 261-277.
38. Coe, E. M.; Bowern, L. H.; Speer, J. A.; Wang, S.; Sayers, D. E.; Bereman, R. D. *J. inorg. Biochem.*, **1995**, *58*, 269-278.
39. Saier, M. H.; Mansour, N. M. *J. Mol. Microbiol. Biotechnol.*, **2005**, *10*, 22-25.
40. Hosono, A.; Lee, J.; Ametani, A.; Natsume, M.; Hirayama, M.; Adachi, T.; Kaminogawa, S. *Biosci. Biotechnol. Biochem.*, **1997**, *61*, 312-316.
41. Nakajima, H.; Suzuki, Y.; Kaizu, H.; Hirota, T. *J. Food Sci.*, **1992**, *57*, 1327-1329.
42. Kitazawa, H.; Toba, T.; Itoh, T.; Kumano, N.; Adachi, S.; Yamaguchi, T. *Anim. Sci. Technol.*, **1991**, *62*, 277-283.
43. Weissenborn, P. K. *Int. J. Miner. Process.*, **1996**, *47*, 197-211.

44. Suh, J.-K. F.; Matthew, H. W. T. *Biomaterials*, **2000**, *21*, 2589-2598.
45. Rongved, P.; Fritzell, T. H.; Strande, P.; Klaveness, J. *Carbohydr. Res.*, **1996**, *287*, 77-89.
46. Breur, H. J. A. *Fouling and Bioprotection of Metals*, PhD thesis, **2001**.
47. Löfås, S.; Johnsson, B. *Chem. Soc. Chem. Comm.*, **1990**, 1526-1528.
48. Smitha, B.; Sridhar, S.; Khan, A. A. *Eur. Polym. J.*, **2005**, *41*, 1859-1866.
49. Wan, Y.; Peppley, B.; Creber, K. A. M.; Bui, V. T.; Halliop, E. J. *Polym. Sources*, **2006**, *162*, 105-113.
50. Boels, I. C.; van Kranenburg, R.; Hugenholz, J.; Kleerebezem, M.; de Vos, W. M. *Int. Dairy J.*, **2001**, *11*, 723-732.
51. Roberts, I. S. *Annu. Rev. Microbiol.*, **1996**, *50*, 285-315.
52. García, E.; López, R. *FEMS Microbiol. Lett.*, **1997**, *149*, 1-10.
53. García, E.; Arrecubieta, C.; Muñoz, R.; Mollerach, M.; López, R. *Microb. Drug. Resist.*, **1997**, *3*, 72-88.
54. Kolkman, M. A. B.; van der Zeijst, B. A. M.; Nuijten, P. J. M. *Am. Soc. Biochem. Mol. Biol.*, **1997**, *272*, 19502-19508.
55. Stिंगe, F.; Neeser, J.-R.; Mollet, B. *J. Bacteriol.*, **1996**, *178*, 1680-1690.
56. Stिंगe, F.; Newell, J. W.; Neeser, J.-R. *J. Bacteriol.*, **1999**, *181*, 6354-6360.
57. Van Kranenburg, R.; Marugg, J. D.; van Swam, I. I.; Willem, N. J.; de Vos, W. M. *Mol. Microbiol.*, **1997**, *24*, 387-397.
58. Van Kranenburg, R.; van Swam, I. I.; Marugg, J. D.; Kleerebezem, M.; de Vos, W. M. *J. Bacteriol.*, **1999b**, *181*, 338-340.
59. Low, D.; Ahlgren, J. A.; Horne, D.; McMahon, D. J.; Oberg, C. J.; Broadbent, J. R. *Appl. Environ. Microbiol.*, **1998**, *64*, 2147-2151.
60. Bourgoïn, F.; Pluvinet, A.; Gintz, B.; Decaris, B.; Guédon, G. *Gene*, **1999**, *233*, 151-161.
61. Jolly, L.; Stिंगe, F. *Int. Dairy J.*, **2001**, *11*, 733-745.
62. Almirón-Roig, E.; Mulholland, F.; Gasson, M. J.; Griffin, A. M. *Microbiology*, **2000**, *43*, 995-1000.
63. Guidolin, A.; Morona, J. K.; Morona, R.; Hansman, D.; Paton, J. C. *Infect. Immun.*, **1994**, *62*, 5384-5396.
64. Broadbent, J. R.; McMahon, D. J.; Welker, D. L.; Oberg, C. J.; Moineau, S. *Am. Dairy Sci. Assoc.*, **2003**, *86*, 407-423.
65. Dong, C.; Beis, K.; Nesper, J.; Brunkan-LaMontagne, A. L.; Clarke, B. R.; Whitfield, C.; Naismith, J. H. *Nature Lett.*, **2006**, *444*, 226-229.
66. Collins, R. F.; Beis, K.; Dong, C.; Botting, C. H.; McDonnell, C.; Ford, R. C.; Clarke, B. R.; Whitfield, C.; Naismith, J. H. *PNAS*, **2007**, *104*, 2390-2395.
67. Llauberes, R. M.; Richard, B.; Lonvaud, A.; Dubourdieu, D.; Fournet, B. *Carbohydr. Res.*, **1990**, *203*, 103-107.
68. Duenas-Chasco, M. T.; Rodríguez-Carvajal, M. A.; Tejero-Mateo, P.; Espartero, J. L.; Irastorza-Iribas, A.; Gil-Serrano, A. M. *Carbohydr. Res.*, **1998**, *207*, 125-133.
69. Van Hijum, S. A. F. T.; Kralj, S.; Ozimek, L. K.; Dijkhuizen, L.; van Geel-Schutten, G. H. *Microbiol. Mol. Biol. Rev.*, **2006**, *70*, 157-176.
70. Giffard, P. M.; Rathsam, C.; Kwan, E.; Kwan, D. W.; Bunny, K. L.; Koo, S. P.; Jacques, N. A. *J. Gen. Microbiol.*, **1993**, *139*, 913-920.
71. Shiroza, T.; Kuramitsu, H. K. *J. Bacteriol.*, **1988**, *170*, 810-916.
72. Carlsson, J.; Granhén, H.; Johnsson, G.; Wikner, S. *Arch. Oral Biol.*, **1970**, *15*, 1143-1148.
73. Hancock, R. A.; Marschall, K.; Weigel, H. *Carbohydr. Res.*, **1976**, *49*, 351-360.
74. Simms, P. J.; Boyko, W. J.; Edwards, J. R. *Carbohydr. Res.*, **1990**, *208*, 193-198.
75. Robyt, J. F.; Walseth, T. F. *Carbohydr. Res.*, **1979**, *68*, 95-111.
76. Van Geel-Schutten, G. H.; Faber, E. J.; Smit, E.; Bonting, K.; Smith, M. R.; ten Brink, B.; Kamerling, J. P.; Vliegthart, J. F. G.; Dijkhuizen, L. *Appl. Environ. Microbiol.*, **1999**, *65*, 3008-3014.
77. Van Hijum, S. A. F. T.; Bonting, K.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *FEMS Microbiol. Lett.*, **2001**, *205*, 323-328.
78. Korakli, M.; Pavlovic, M.; Gänzle, M. G.; Vogel, R. F. *Appl. Environ. Microbiol.*, **2003**, *69*, 2073-2079.
79. Van Hijum, S. A. F. T.; van Geel-Schutten, G. H.; Rahaoui, H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Appl. Environ. Microbiol.*, **2002**, *68*, 4390-4398.
80. Olivares-Illana, V.; Lopez-Munguia, A.; Olvera, C. *J. Bacteriol.*, **2003**, *185*, 3606-3612.
81. Tiekink, M.; Korakli, M.; Ehrmann, M. A.; Gänzle, M. G.; Vogel, R. F. *Appl. Environ. Microbiol.*, **2003**, *69*, 945-952.
82. Pasteur, L. *Bull. Soc. Chim.*, **1861**, *11*, 30-31.
83. Guggenheim, B. *Helv. Odontol. Acta*, **1970**, *14*, 89-109.
84. Côté, G. L.; Robyt, J. F. *Carbohydr. Res.*, **1982**, *101*, 57-74.
85. Kralj, S.; van Geel-Schutten, G. H.; Rahaoui, H.; Leer, R. J.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Appl. Environ. Microbiol.*, **2002**, *68*, 4283-4291.
86. Monchois, V.; Remaud-Simeon, M.; Monsan, P. F.; Willemot, R.-M. *FEMS Microbiol. Lett.*, **1998**, *159*, 307-315.
87. Monchois, V.; Willemot, R.-M.; Remaud-Simeon, M.; Croux, C.; Monsan, P. F. *Gene*, **1996**, *182*, 23-32.

88. Monchois, V.; Remaud-Simeon, M.; Russell, R. R. B.; Monsan, P. F.; Willemot, R.-M. *Appl. Microbiol. Biotechnol.*, **1997**, *48*, 465-472.
89. Monchois, V.; Reverte, M.; Remaud-Simeon, M.; Monsan, P. F.; Willemot, R.-M. *Appl. Environ. Microbiol.*, **1998**, *64*, 1644-1649.
90. Funane, K.; Mizuno, K.; Takahara, H.; Kobayashi, M. *Biosci. Biotechnol. Biochem.*, **2000**, *64*, 29-38.
91. Fukushima, K.; Ikedat, T.; Kuramitsu, H. K. *Infect. Immun.*, **1992**, *60*, 2815-2822.
92. Simpson, C. L.; Cheetham, N. W. H.; Giffard, P. M.; Jacques, N. A. *Microbiology*, **1995**, *141*, 1451-1460.
93. Kim, D.; Robyt, J. F. *Enzyme Microb. Technol.*, **1994**, *16*, 659-664.
94. Kralj, S.; van Geel-Schutten, G. H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Microbiology*, **2004**, *150*, 2099-2112.
95. Kralj, S.; van Geel-Schutten, G. H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Biocatal. Biotrans.*, **2003**, *21*, 181-187.
96. Jensen, M. H.; Mirza, O.; Albenne, C.; Remaud-Simeon, M.; Monsan, P. F.; Gajhede, M.; Skov, L. K. *Biochemistry*, **2004**, *43*, 3104-3110.
97. Mirza, O.; Skov, L. K.; Remaud-Simeon, M.; Potocki de Montalk, G.; Albenne, C.; Monsan, P. F.; Gajhede, M. *Biochemistry*, **2001**, *40*, 9032-9039.
98. Stevenson, G.; Neal, B.; Liu, D.; Hobbs, M.; Packer, N. H.; Batley, M.; Redmond, J. W.; Lindquist, L.; Reeves, P. J. *Bacteriol.*, **1994**, *176*, 4144-4156.
99. MacGregor, E. A.; Janecek, S.; Svensson, B. *Biochim. Biophys. Acta*, **2001**, *1546*, 1-20.
100. MacGregor, E. A.; Jespersen, H. M.; Svensson, B. *FEBS Lett.*, **1996**, *378*, 263-266.
101. Kralj, S.; van Geel-Schutten, G. H.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Biochemistry*, **2005**, *44*, 9206-9216.
102. Russell, R. B. *Arch. Oral Biol.*, **1990**, *35*, 53-58.
103. Bozonnet, S.; Dols-Laffargue, M.; Fabre, E.; Pizzut, S.; Remaud-Simeon, M.; Monsan, P. F.; Willemot, R.-M. *J. Bacteriol.*, **2002**, *184*, 5753-5761.
104. Fabre, E.; Bozonnet, S.; Arcache, A.; Willemot, R.-M.; Vignon, M. R.; Monsan, P. F.; Remaud-Simeon, M. *J. Bacteriol.*, **2005**, *187*, 296-303.
105. Misaki, A.; Kawaguchi, K.; Miyaji, H.; Nagae, H.; Hokkoku, S.; Kakuta, M.; Sasaki, T. *Carbohydr. Res.*, **1984**, *129*, 209-227.
106. Christensen, B. E.; Aasprong, E.; Stokke, B. T. *Carbohydr. Polym.*, **2001**, *46*, 241-248.
107. McLean, K. M.; Johnson, G.; Chatelier, R. C.; Beumer, G. J.; Steele, J. G.; Griesser, H. J. *Colloids Surf. B Biointerf.*, **2000**, *18*, 221-234.
108. Chang, P.-S.; Robyt, R. F. *J. Carbohydr. Chem.*, **1996**, *15*, 819-830.
109. Yoo, S.-H.; Lee, J.-S.; Park, S. Y.; Kim, Y.-S.; Chang, P.-S.; Lee, H. G. *Int. J. Biol. Macromol.*, **2005**, *35*, 27-31.
110. Garg, S. K.; Alam, Md. S.; Radha Kishan, K. V.; Agrawal, P. *Prot. Expr. Purific.*, **2007**, *51*, 198-208.
111. Skjåk-Bræk, G.; Smidsrød, O.; Larsen, B. *Int. J. Macromol.*, **1986**, *8*, 330-336.
112. Hartmann, M.; Dentini, M.; Draget, K. I.; Skjåk-Bræk, G. *Carbohydr. Polym.*, **2006**, *63*, 257-262.
113. Tuinier, R.; van Casteren, W. H. M.; Looijesteijn, P. J.; Schols, H. A.; Voragen, A. G. J.; Zoon, P. *Biopolymers*, **2001**, *59*, 160-166.
114. Stinglele, F.; Vincent, S. J. F.; Faber, E. J.; Newell, J. W.; Kamerling, J. P.; Neeser, J.-R. *Mol. Microbiol.*, **1999**, *32*, 1287-1295.
115. Kamerling, J. P.; Vliegthart, J. F. G. In *Clinical Biochemistry – Principles, Methods, Applications. Vol. 1, Mass Spectrometry*; Lawson, A. M., Ed.; Walter de Gruyter: Berlin, **1989**; pp 176-263.
116. Ciucanu, I.; Kerek, F. *Carbohydr. Res.* **1984**, *131*, 209-217.
117. Jansson, P.-E.; Kenne, L.; Liedgren, H.; Lindberg, B.; Lönngrén, J. *Chem. Commun. Univ. Stockholm*, **1976**, *8*, 1-74.
118. Hay, G. W.; Lewis, B. A.; Smith, F. *Methods Carbohydr. Chem.*, **1965**, *5*, 357-360.

*It was a glimpse into some kind of wonderful world where electricity and mathematics
and engineering and nice diagrams all came together*

-Edward M. Purcell-

Chapter 2

Development of an NMR structural-reporter-group concept for the primary structural characterisation of α -D-glucans

Sander S. van Leeuwen, Bas R. Leeflang, Gerrit J. Gerwig, and
Johannis P. Kamerling

Abstract – An NMR study of proton and carbon chemical shift patterns of known, linear α -D-glucopyranose di- and trisaccharide structures was carried out. Chemical shift patterns for (α 1-2)-, (α 1-3)-, (α 1-4)-, and (α 1-6)-linked D-glucose residues were analysed and compared to literature data. A structural-reporter-group concept was formulated to function as a tool in the structural analysis of α -D-glucans. The structural-reporter-group signals identified allow elucidation of glucan structures, when the available material is not sufficient for ^{13}C correlation spectroscopy.

Introduction

In the structural analysis of polysaccharides, partial acid hydrolysis and subsequent structure determination of the generated fragments by NMR spectroscopy is a much employed technique. The specific substitution pattern of a carbohydrate residue can often be elucidated from the exact chemical shifts of well-resolved proton resonances, observed outside the proton bulk region of δ 4.00-3.50. These protons have been termed structural-reporter groups. Heteropolysaccharides with a repeating unit generally have well-separated anomeric ^1H and ^{13}C signals for each residue of the repeating unit. The latter signals usually form the start for a complete analysis of the ^1H and ^{13}C NMR spectra, making use of a variety of 2D techniques, such as ^1H - ^1H TOCSY, ^1H - ^1H NOESY, ^1H - ^1H ROESY, ^{13}C - ^1H HSQC, and ^{13}C - ^1H HMBC. Partial acid hydrolysis fragments of the repeating unit can be easily assigned with the same NMR spectroscopic techniques. In case of homopolysaccharides such as glucans with different types of glycosidic linkages, but no repeating unit, the analysis of both the polysaccharide and the oligosaccharide fragments is more complicated. Here, to distil structural information from the various applied NMR techniques, it would be essential to know how ^1H and ^{13}C chemical shifts are influenced by different substitution patterns.

Besides the well-known dextran- and mutansucrase enzymes, a number of novel glucansucrase enzymes and glucan products have been characterised in recent years.¹ This includes enzymes synthesising unique α -D-glucan products, with (α 1-4) or (α 1-2) glycosidic bonds. The arrangement of glycosidic linkages strongly contributes to specific glucan properties such as solubility and rheology. The identification of the exact order of different glycosidic bonds in various glucan products remains a major challenge in carbohydrate research. Recently, it has been shown that the α -D-glucan produced from sucrose by glucansucrase GTFA from *Lactobacillus reuteri* strain 35-5 (**EPS35-5**) contains (α 1-6) and (α 1-4) linkages,² whereas the α -D-glucan produced from sucrose by glucansucrase GTF180 from *Lb. reuteri* strain 180 (**EPS180**) contains (α 1-6) and (α 1-3) linkages.³ Mutations in the GTFA glucansucrases have led to different ratios of the linkage types.⁴ Here, we report the basics of an NMR structural-reporter-group concept for the analysis of such branched α -D-glucans, making use of an NMR model study of α -D-glucan di- and trisaccharides.

Results

Kojibiose, α -D-Glcp-(1 \rightarrow 2)-D-Glcp (**B** \rightarrow **A**; **1**)

The 1D ^1H NMR spectrum of kojibiose (Figure 1) showed four anomeric ^1H signals at δ 5.427 (**A α** H-1, $^3J_{1,2}$ 3.4 Hz), 5.376 (**B β** H-1, $^3J_{1,2}$ 3.4 Hz), 5.084 (**B α** H-1, $^3J_{1,2}$ 2.9 Hz), and 4.785 (**A β** H-1, $^3J_{1,2}$ 8.3 Hz); the anomeric ratio amounts to 51:49 (note that the intensity of the **A β** signal is influenced by the pre-saturation of the HOD signal). The splitting of the **B** H-1 doublet into two doublets (**B α** H-1 and **B β** H-1) corresponds with the α - and β -anomeric forms of residue **A** (**A α** and **A β**). The great difference between the δ values of the two **B** H-1 signals (**B α** \ll **B β**) reflects directly the nearby orientation of the β -anomeric OH group, compared with the α -anomeric OH group, of residue **A** in the (α 1-2) linkage. Starting from the anomeric signals in the 2D TOCSY spectrum (Figure 1/180 ms), the chemical shifts of all non-anomeric protons could be determined (Table 1). In the 2D ROESY spectrum (Figure 1) strong inter-residual couplings were observed between **B α** H-1 and **A α** H-2, between **B β** H-1 and **A β** H-2, and between **B α** H-1 and **A α** H-1. The peaks at δ 3.932 (**B α** H-5) and δ 4.017 (**B β** H-5) in the 1D ^1H NMR spectrum have the distinctive shape of H-5 signals.

Elucidation of the ^{13}C - ^1H HSQC spectrum (Figure 1) yielded the ^{13}C chemical shifts (Table 2). The substitution of **A** O-2 by residue **B** is reflected by the downfield shifts of **A α** C-2 ($\delta_{\text{C-2}}$ 77.1; α -D-Glcp, $\delta_{\text{C-2}}$ 73.1) and **A β** C-2 ($\delta_{\text{C-2}}$ 79.9; β -D-Glcp, $\delta_{\text{C-2}}$ 75.7).^{5,6}

Nigerose, α -D-Glcp-(1 \rightarrow 3)-D-Glcp (**B** \rightarrow **A**; **2**)

The 1D ^1H NMR spectrum of nigerose (Figure 2) showed four anomeric ^1H signals at δ 5.371 (**B α** H-1, $^3J_{1,2}$ 3.4 Hz), 5.355 (**B β** H-1, $^3J_{1,2}$ 3.4 Hz), 5.233 (**A α** H-1, $^3J_{1,2}$ 3.4 Hz), and 4.663 (**A β** H-1, $^3J_{1,2}$ 7.7 Hz); the anomeric ratio amounts to 39:61. The chemical shifts of all non-anomeric ^1H resonances of residues **A** and **B**, as established by TOCSY measurements (Figure 2/180 ms), are presented in Table 1. The strong inter-residual ROESY cross-peaks (Figure 2) of **B** H-1 with **A α** H-3 and of **B** H-1 with **A β** H-3 are in accordance with an (α 1-3) linkage. The **B** H-5 signal at δ 4.014 shows the characteristic splitting pattern due to coupling with **B** H-6a, H-6b, and H-4.

The ^1H chemical shifts determined were correlated to ^{13}C chemical shifts in the ^{13}C - ^1H HSQC spectrum (Figure 2; Table 2). The substitution of **A** O-3 by residue **B** is reflected by the downfield shifts of **A α** C-3 ($\delta_{\text{C-3}}$ 80.6; α -D-Glcp, $\delta_{\text{C-3}}$ 74.4) and **A β** C-3 ($\delta_{\text{C-3}}$ 83.2; β -D-Glcp, δ 77.3).^{5,6}

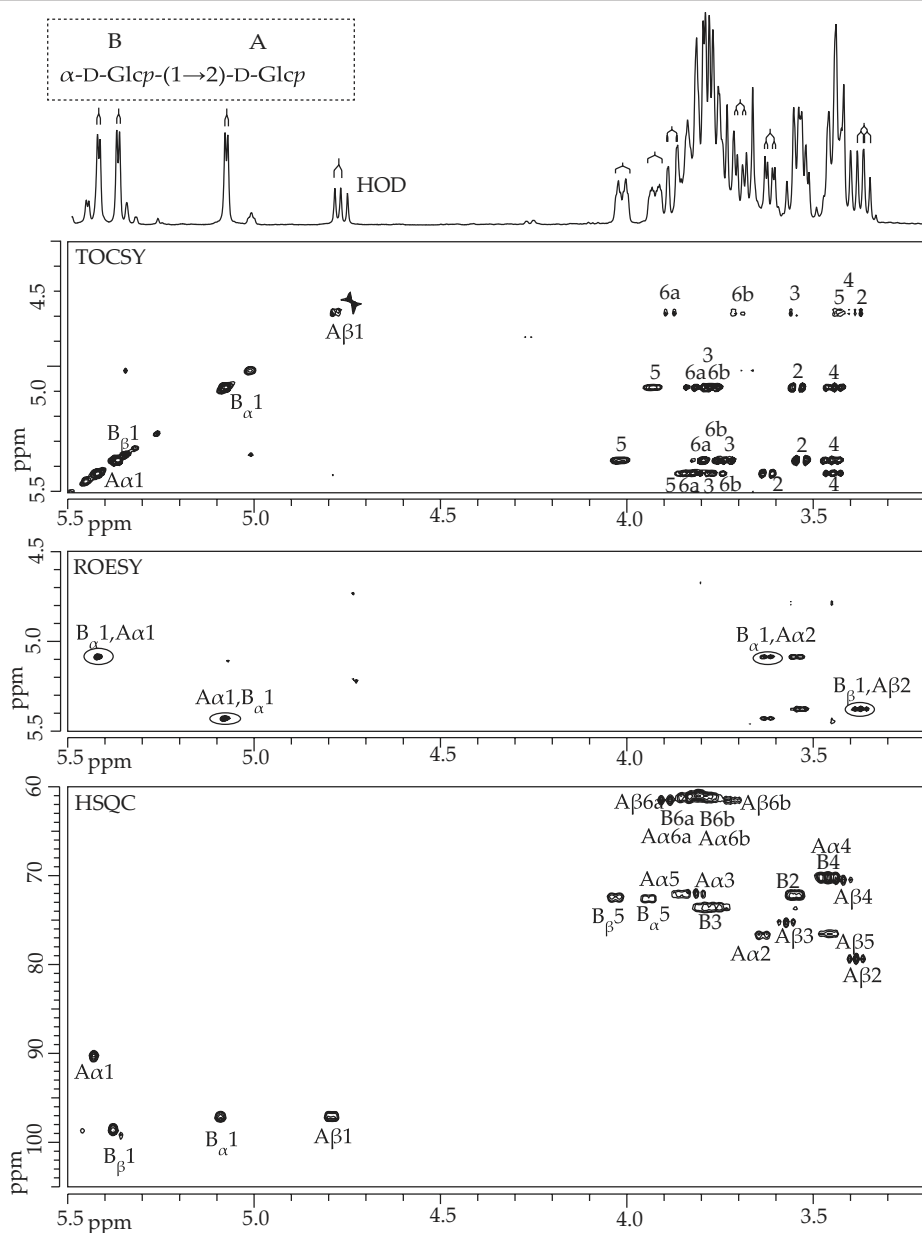


Figure 1. Kojibiose (**1**): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . Anomeric protons in the TOCSY spectrum ($\text{A}\alpha 1$, etc.) have been indicated on the diagonal; numbers in the horizontal tracks belong to the cross-peaks of the scalar-coupling network of the residues indicated. In the ROESY spectrum inter-residual couplings ($\text{B}\alpha 1, \text{A}\alpha 2$ means a cross-peak between $\text{B}\alpha$ H-1 and $\text{A}\alpha$ H-2, etc.) have been indicated with circles. In the ^{13}C - ^1H HSQC spectrum $\text{A}\alpha 1$ denotes the cross-peak between H-1 and C-1 of residue $\text{A}\alpha$, etc.

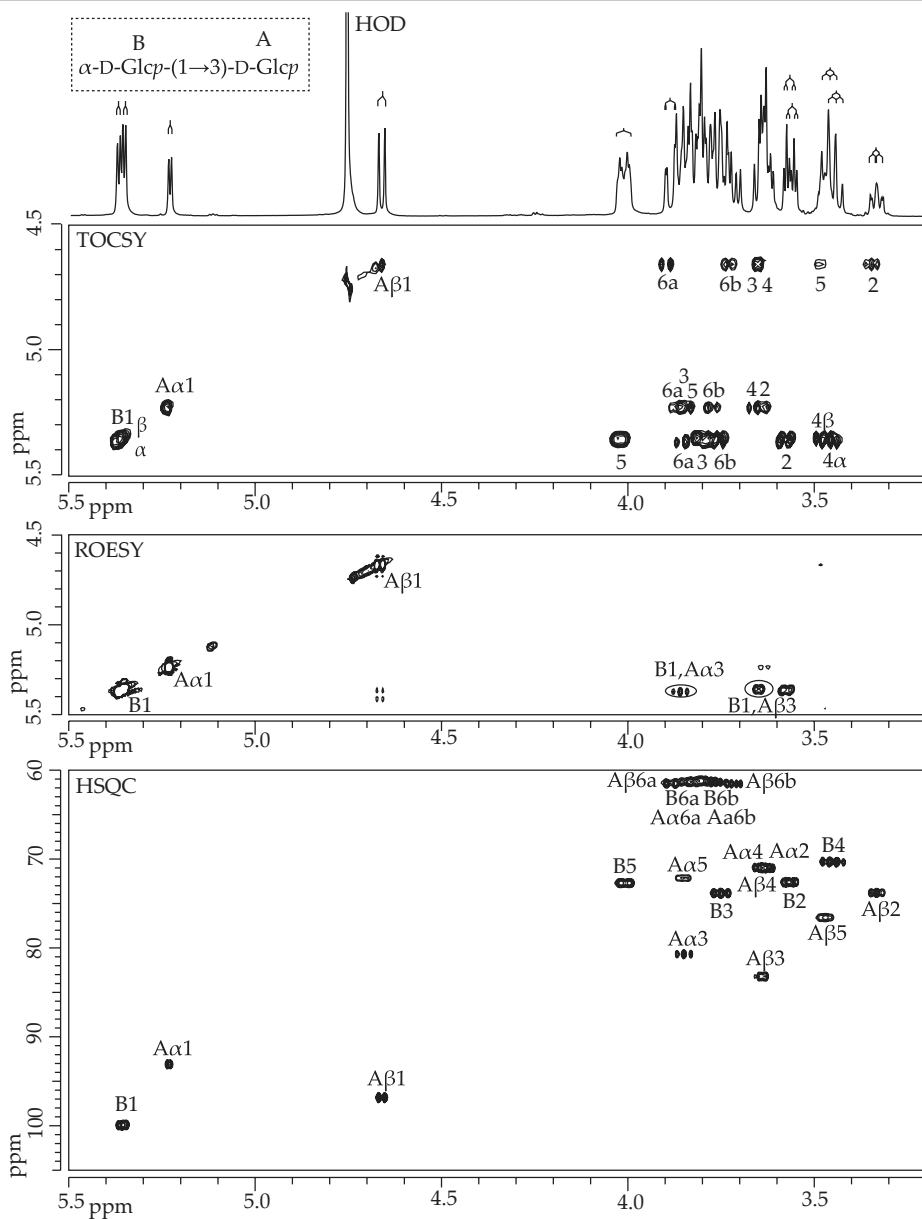


Figure 2. Nigerose (2): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

Maltose, α -D-Glc β -(1 \rightarrow 4)-D-Glc β (**B** \rightarrow **A**; **3**)

The 1D ^1H NMR spectrum of maltose (Figure 3) showed three anomeric ^1H signals at δ 5.404 (**B** H-1, $^3J_{1,2}$ 3.7 Hz), 5.225 (**A α** H-1, $^3J_{1,2}$ 3.7 Hz), and 4.649 (**A β** H-1, $^3J_{1,2}$ 8.0 Hz); the anomeric ratio amounts to 31:69. In contrast to the anomeric protons of the α -D-Glc β -(1 \rightarrow 2)-, α -D-Glc β -(1 \rightarrow 3)-, and α -D-Glc β -(1 \rightarrow 6)- residues, the anomeric proton of the α -D-Glc β -(1 \rightarrow 4)- residue gives rise to only one signal, indicating that the influence of the anomeric configuration of the reducing Glc residue is negligible. A full assignment of the non-anomeric proton signals (Table 1) followed from TOCSY experiments (Figure 3/180 ms). In the ROESY spectrum (Figure 3) a strong inter-residual ^1H - ^1H correlation was observed between **B** H-1 and **A α,β** H-4, reflecting the (1-4) linkage.

The ^{13}C chemical shifts of all carbon atoms were determined by correlating ^{13}C - ^1H HSQC spectroscopic data (Figure 3) with ^1H chemical shifts determined by TOCSY spectroscopy (Table 2). The substitution on O-4 of residue **A** is reflected by the downfield shifts of **A α** C-4 ($\delta_{\text{C-4}}$ 77.7; α -D-Glc β , $\delta_{\text{C-4}}$ 71.2) and **A β** C-4 ($\delta_{\text{C-4}}$ 77.7; β -D-Glc β , $\delta_{\text{C-4}}$ 71.2).^{5,6}

Isomaltose, α -D-Glc β -(1 \rightarrow 6)-D-Glc β (**B** \rightarrow **A**; **4**)

The 1D ^1H NMR spectrum of isomaltose (Figure 4) showed four anomeric signals at δ 5.239 (**A α** H-1, $^3J_{1,2}$ 3.5 Hz), 4.955 (**B β** H-1, $^3J_{1,2}$ 3.5 Hz), 4.948 (**B α** H-1, $^3J_{1,2}$ 3.5 Hz), and 4.672 (**A β** H-1, $^3J_{1,2}$ 7.6 Hz); the anomeric ratio amounts to 35:65. The signals of the anomeric protons in the TOCSY spectrum (Figure 4/180 ms) were used as starting point to identify all non-anomeric proton signals (Table 1). The **A α** H-5 and **A β** H-5 signals are shifted downfield to δ 4.01 (α -D-Glc β , $\delta_{\text{H-5}}$ 3.84) and δ 3.643 (β -D-Glc β , $\delta_{\text{H-5}}$ 3.45), respectively, caused by the attached residue **B** at O-6. The substitution at O-6 also causes strong downfield shifts of **A α** H-6b to δ 4.00 (α -D-Glc β , $\delta_{\text{H-6b}}$ 3.74) and **A β** H-6b to δ 3.962 (β -D-Glc β , $\delta_{\text{H-6b}}$ 3.71), and upfield shifts of **A α** H-6a to δ 3.70 (α -D-Glc β , $\delta_{\text{H-6a}}$ 3.83) and **A β** H-6a to δ 3.78 (β -D-Glc β , $\delta_{\text{H-6a}}$ 3.89). The ROESY spectrum (Figure 4) showed strong **B** H-1,**A α** H-6a and **B** H-1,**A β** H-6a cross-peaks, and weak **B** H-1,**A α** H-6b and **B** H-1,**A β** H-6b connectivities, as expected for the (1-6) linkage.

By combining the ^1H chemical shifts derived from the TOCSY measurements and the ^{13}C - ^1H correlation data from the 2D HSQC spectrum (Figure 4), the ^{13}C chemical shifts of all carbon atoms were identified (Table 2). The substitution on O-6 of residue **A** is reflected by the downfield shifts of **A α** C-6 ($\delta_{\text{C-6}}$ 66.7; α -D-Glc β , $\delta_{\text{C-6}}$ 62.2) and **A β** C-6 ($\delta_{\text{C-6}}$ 66.7; β -D-Glc β , $\delta_{\text{C-6}}$ 62.3).^{5,6}

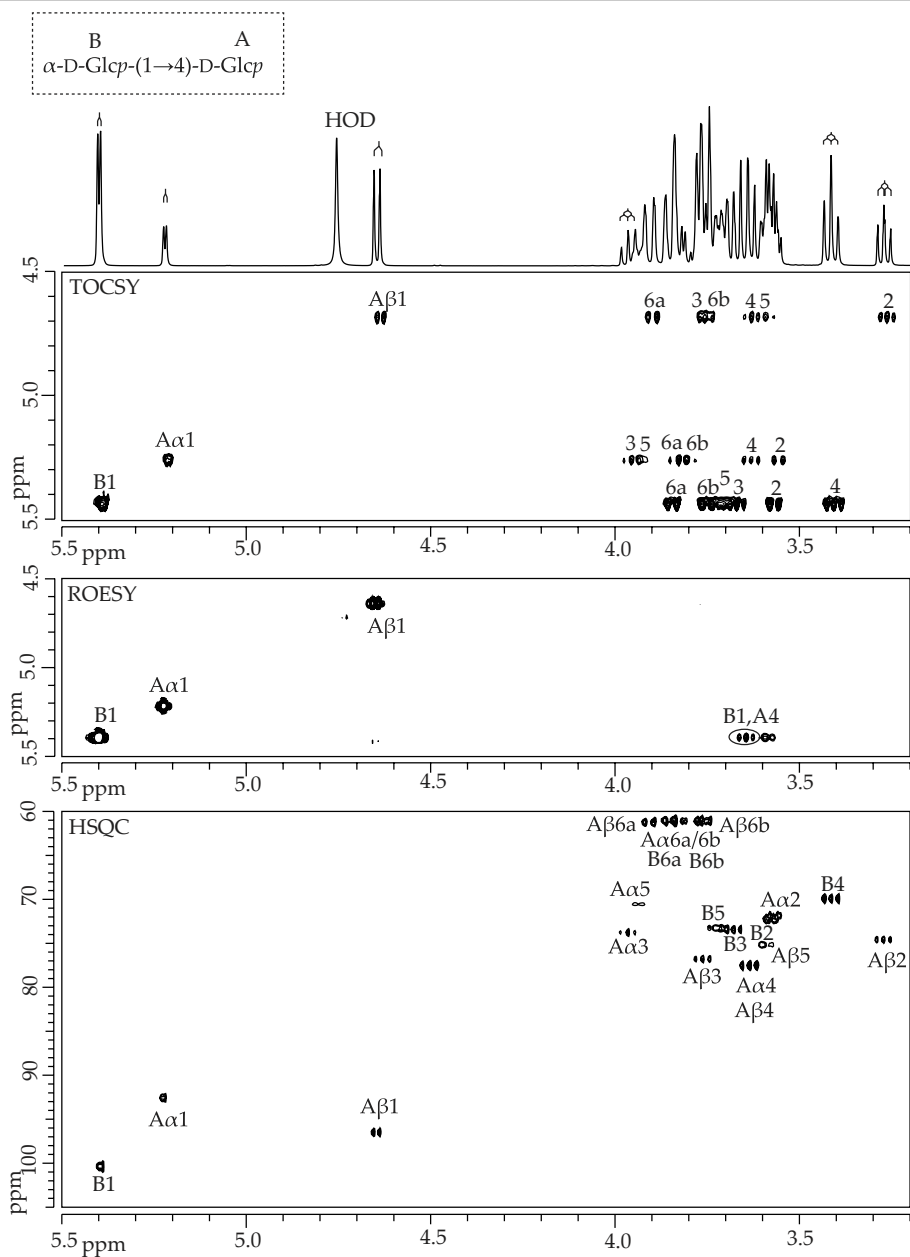


Figure 3. Maltose (3): 500-MHz 1D 1H NMR spectrum, 2D 1H - 1H TOCSY spectrum (mixing time 180 ms), 2D 1H - 1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - 1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

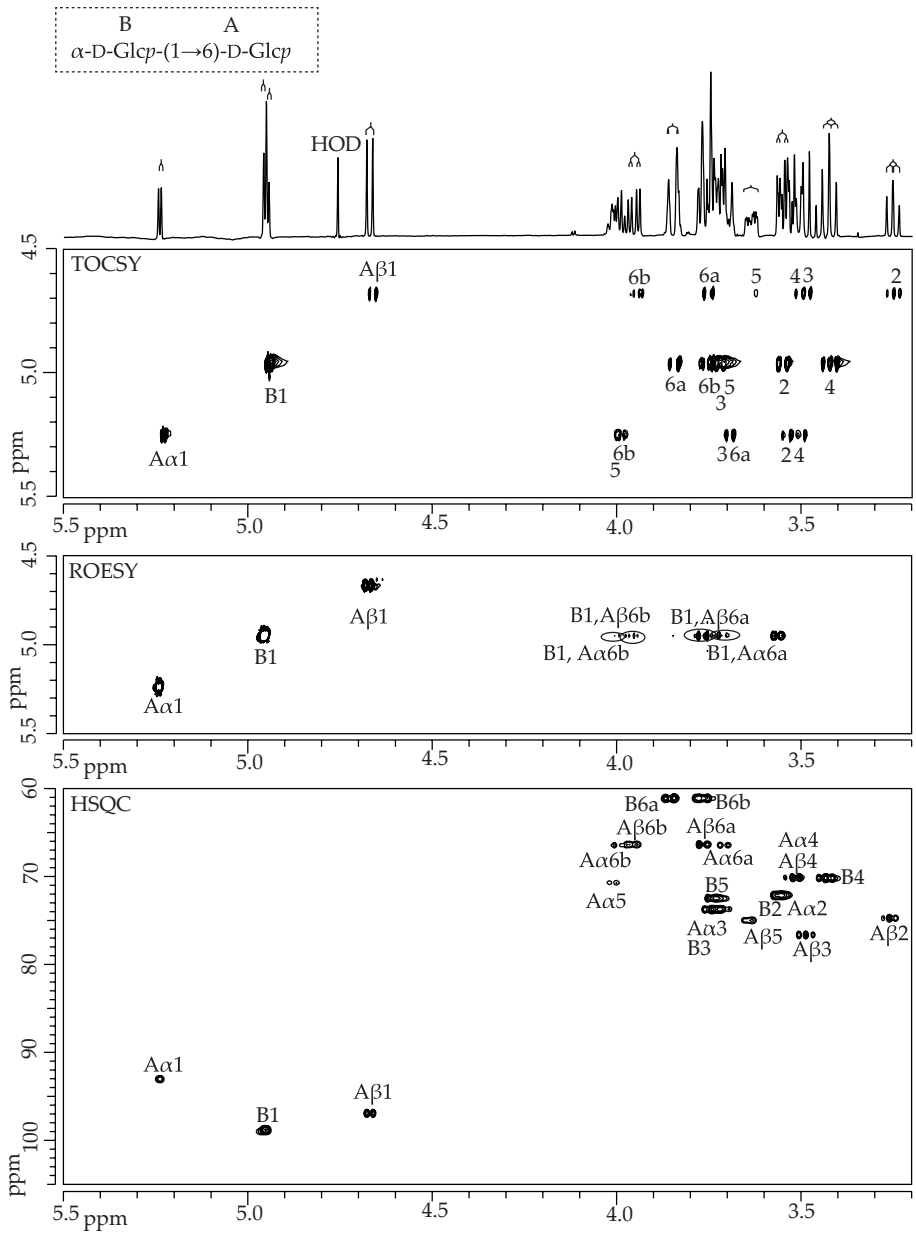


Figure 4. Isomaltose (4): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

Table 1. ¹H chemical shifts of D-glucopyranose residues of free D-glucose and disaccharides **1** to **4**, referenced to internal acetone (δ 2.225). ³J_{1,2} couplings are included between brackets, where measured.

Residue ^a	D-glucose	1	2	3	4
Aα					
H-1	5.224 (3.3)	5.427 (3.4)	5.233 (3.4)	5.225 (3.7)	5.239 (3.5)
H-2	3.524 (9.3)	3.632 (9.7)	3.63	3.56	3.55
H-3	3.71	3.80	3.85	3.972 (9.6)	3.71
H-4	3.402 (9.4)	3.45	3.65	3.64	3.53
H-5	3.84	3.85	3.84	3.94	4.01
H-6a	3.83	3.84	3.86	3.848 (1.4;-12.1)	3.70
H-6b	3.74	3.76	3.75	3.796 (4.4)	4.00
Aβ					
H-1	4.637 (7.9)	4.785 (8.3)	4.663 (7.7)	4.649 (8.0)	4.672 (7.6)
H-2	3.236 (9.3)	3.373 (9.2)	3.332 (8.2)	3.272 (9.1)	3.253 (8.3)
H-3	3.480 (9.4)	3.55	3.64	3.76	3.48
H-4	3.393 (9.4)	3.43	3.64	3.62	3.52
H-5	3.449	3.45	3.47	3.60	3.643
H-6a	3.890 (1.9;-12.1)	3.866 (1.2;-12.5)	3.887 (1.8;-11.7)	3.902 (1.9;-12.1)	3.78
H-6b	3.71	3.687 (4.8)	3.723 (4.8)	3.75	3.962 (4.5;-11.7)
Bα					
H-1	-	5.084 (2.9)	5.371 (3.4)	5.404 (3.7)	4.948 (3.5)
H-2	-	3.55	3.571 (9.6)	3.57	3.554 (9.8)
H-3	-	3.78	3.75	3.68	3.75
H-4	-	3.45	3.443 (9.6)	3.414 (9.6)	3.421 (9.4)
H-5	-	3.932	4.014	3.71	3.75
H-6a	-	3.84	3.84	3.848 (1.4;-12.1)	3.849 (1.4;-11.7)
H-6b	-	3.77	3.78	3.76	3.78
Bβ					
H-1	-	5.376 (3.4)	5.355 (3.4)	5.404 (3.7)	4.955 (3.5)
H-2	-	3.53	3.557 (9.6)	3.57	3.554 (9.8)
H-3	-	3.74	3.75	3.68	3.75
H-4	-	3.46	3.455 (9.6)	3.414 (9.6)	3.421 (9.4)
H-5	-	4.017	4.014	3.71	3.75
H-6a	-	3.81	3.84	3.848 (1.4;-12.1)	3.849 (1.4;-11.7)
H-6b	-	3.78	3.78	3.76	3.78

^aA α and A β stand for the anomeric configuration of residue **A**; B α and B β stand for residue **B**, connected to the α - and β -anomer of residue **A**, respectively.

Table 2. ^{13}C chemical shifts of D-glucopyranose residues of disaccharides **1** to **4**, referenced to the methyl-carbon of internal acetone (δ 31.08).

Residue ^a		1	2	3	4
Aα	C-1	90.2	93.2	92.6	93.0
	C-2	77.1	70.8	72.4	72.4
	C-3	72.2	80.6	74.0	73.9
	C-4	70.7	70.8	77.7	70.3
	C-5	72.2	72.0	70.8	70.9
	C-6	61.0	61.2	61.4	66.7
Aβ	C-1	97.4	96.8	96.6	97.1
	C-2	79.9	73.6	75.0	75.0
	C-3	75.4	83.2	77.0	76.9
	C-4	70.9	70.8	77.7	70.3
	C-5	76.6	76.5	75.3	75.2
	C-6	61.5	61.5	61.5	66.7
Bα	C-1	97.6	99.9	100.4	98.9
	C-2	72.3	72.5	72.6	72.4
	C-3	73.8	73.7	73.6	73.9
	C-4	70.5	70.3	70.1	70.5
	C-5	72.9	72.5	73.4	72.5
	C-6	61.0	61.2	61.4	61.4
Bβ	C-1	98.6	99.9	100.4	98.9
	C-2	72.3	72.5	72.6	72.4
	C-3	73.8	73.7	73.6	73.9
	C-4	70.5	70.3	70.1	70.5
	C-5	72.9	72.5	73.4	72.5
	C-6	61.0	61.2	61.4	61.4

^a**A α** and **A β** stand for the anomeric configuration of residue **A**; **B α** and **B β** stand for residue **B**, connected to the α - and β -anomer of residue **A**, respectively.

Nigerotriose, α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 3)-D-Glcp (**C** \rightarrow **B** \rightarrow **A**; **5**)

The 1D ^1H NMR spectrum of nigerotriose (Figure 5) showed five anomeric ^1H signals at δ 5.386 (**B α** H-1, $^3J_{1,2}$ 3.8 Hz), 5.371 (**B β** H-1, $^3J_{1,2}$ 3.8 Hz), 5.359 (**C** H-1, $^3J_{1,2}$ 3.8 Hz), 5.233 (**A α** H-1, $^3J_{1,2}$ 3.8 Hz), and 4.663 (**A β** H-1, $^3J_{1,2}$ 7.6 Hz); the anomeric ratio amounts to 32:68. Starting from the anomeric signals in the TOCSY spectrum (Figure 5/180 ms), the signals of all non-anomeric protons could be assigned (Table 3). The ^1H chemical shifts of **A α** and **A β** closely resembled those found for the reducing unit in **2** (nigerose). The splitting of the anomeric signal of **B** is characteristic for the influence of the α/β -configuration of the reducing residue **A**. The strong inter-residual **B** H-1,**A α** H-3, **B** H-1,**A β** H-3, and **C** H-1,**B** H-3 cross-peaks in the ROESY spectrum (Figure 5) reflect the two (1-3) linkages.

The ^1H chemical shifts determined were correlated to ^{13}C chemical shifts in the ^{13}C - ^1H HSQC spectrum (Figure 5; Table 4). The substitution of **A** O-3 by residue **B**

is reflected by the downfield shifts of **A α** C-3 (δ_{C-3} 80.6; α -D-Glc p , δ_{C-3} 74.4) and **A β** C-3 (δ_{C-3} 83.6; β -D-Glc p , δ_{C-3} 77.3), and the substitution of **B** O-3 by residue **C** by the downfield shift of **B** C-3 (δ_{C-3} 80.6; α -D-Glc p -OMe, δ_{C-3} 74.7).⁶

Maltotriose, α -D-Glc p -(1 \rightarrow 4)- α -D-Glc p -(1 \rightarrow 4)-D-Glc p (**C** \rightarrow **B** \rightarrow **A**; **6**)

The 1D ¹H NMR spectrum of maltotriose (Figure 6) revealed three anomeric ¹H signals at δ 5.399 (**B** H-1 and **C** H-1, $^3J_{1,2}$ 3.4 Hz), 5.224 (**A α** H-1, $^3J_{1,2}$ 3.4 Hz), and 4.650 (**A β** H-1, $^3J_{1,2}$ 7.8 Hz); the anomeric ratio amounts to 37:63. The assignments of all proton signals, as deduced from TOCSY measurements (Figure 6/180 ms), are presented in Table 3. Strong inter-residual connectivities were observed in the ROESY spectrum (Figure 6) between **C** H-1 and **B** H-4, between **B** H-1 and **A α** H-4, and between **B** H-1 and **A β** H-4, in accordance with two (1-4) linkages.

Correlation of ¹H and ¹³C signals in the HSQC spectrum (Figure 6) allowed the assignment of all ¹³C resonances (Table 4). The substitution of **A** O-4 by residue **B** gives rise to large downfield shifts of **A** C-4 (δ_{C-4} 77.8; α -D-Glc p , δ_{C-4} 71.2) and **A β** C-4 (δ_{C-4} 77.8; β -D-Glc p , δ_{C-4} 71.2), and that of **B** O-4 by residue **C** to a large downfield shift of **B** C-4 (δ_{C-4} 77.8; α -D-Glc p -OMe, δ_{C-4} 71.2).⁶

Isomaltotriose, α -D-Glc p -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)-D-Glc p (**C** \rightarrow **B** \rightarrow **A**; **7**)

The 1D ¹H NMR spectrum of isomaltotriose (Figure 7) showed four anomeric signals at δ 5.241 (**A α** H-1, $^3J_{1,2}$ 3.5 Hz), 4.969 (**B β** H-1, $^3J_{1,2}$ 3.5 Hz), 4.961 (**B α** H-1 and **C** H-1, $^3J_{1,2}$ 3.5 Hz), and 4.672 (**A β** H-1, $^3J_{1,2}$ 8.0 Hz); the anomeric ratio amounts to 35:65. Starting from the anomeric signals in the TOCSY spectrum (Figure 7/180 ms) the chemical shifts of all non-anomeric protons were determined (Table 3). A strong ROESY cross-peak (Figure 7) reflected the **C** H-1, **B** H-6a and **B** H-1, **A** H-6a inter-residual connectivities, and therefore the (1-6) linkages. Weak inter-residual ROESY connectivities were observed between **C** H-1 and **B** H-6b and/or between **B** H-1 and **A** H-6b.

The ¹H chemical shifts were combined with ¹³C-¹H correlation data from the HSQC spectrum (Figure 7) to assign all ¹³C chemical shifts (Table 4). The substitution of **A** O-6 by residue **B** is reflected by the downfield shift of **A α** C-6 (δ_{C-6} 66.8; α -D-Glc p , δ_{C-6} 62.2) and **A β** C-6 (δ_{C-6} 66.8; β -D-Glc p , δ_{C-6} 62.3), and the substitution of **B** O-6 by residue **C** by the downfield shift of **B** C-6 (δ_{C-6} 66.5; α -D-Glc p -OMe, δ_{C-6} 62.2).⁶

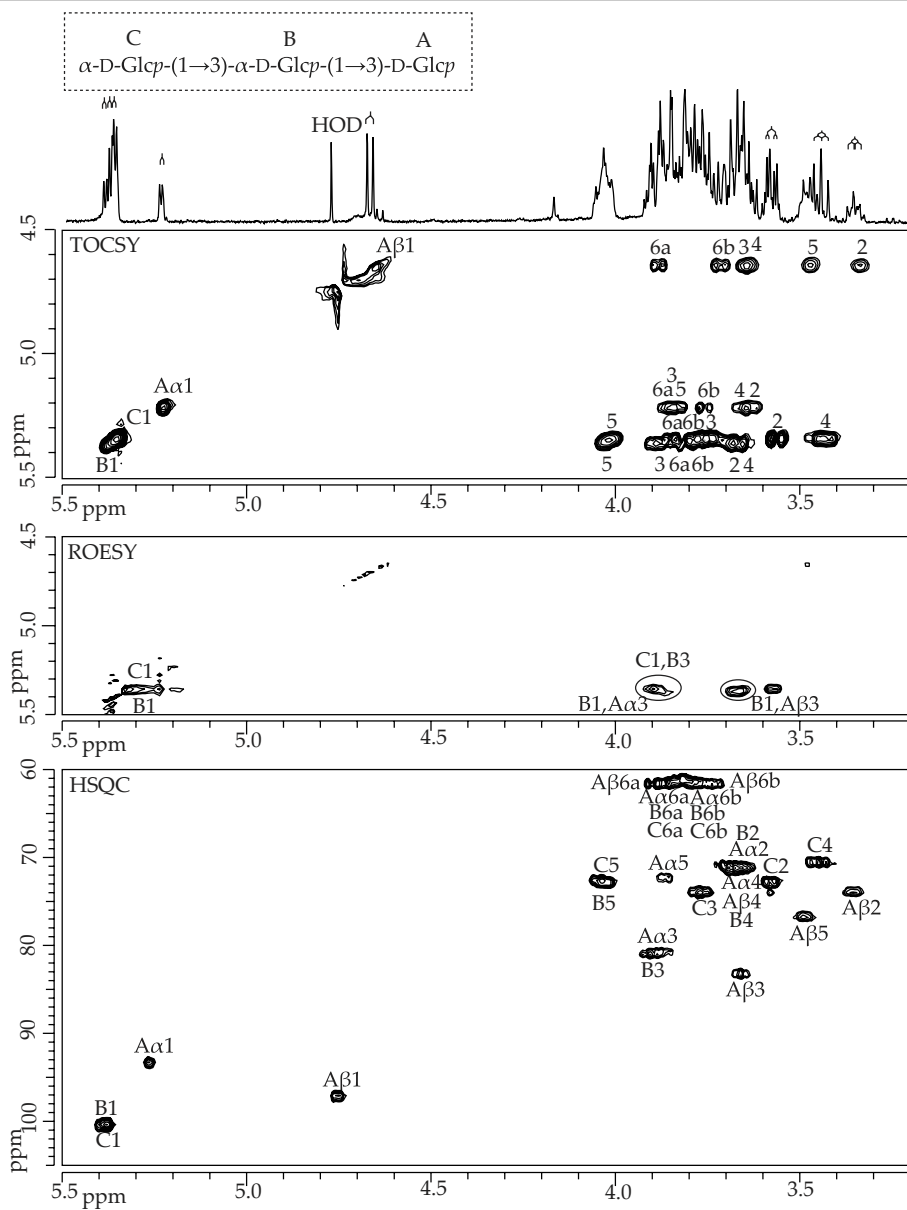


Figure 5. Nigerotriose (5): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

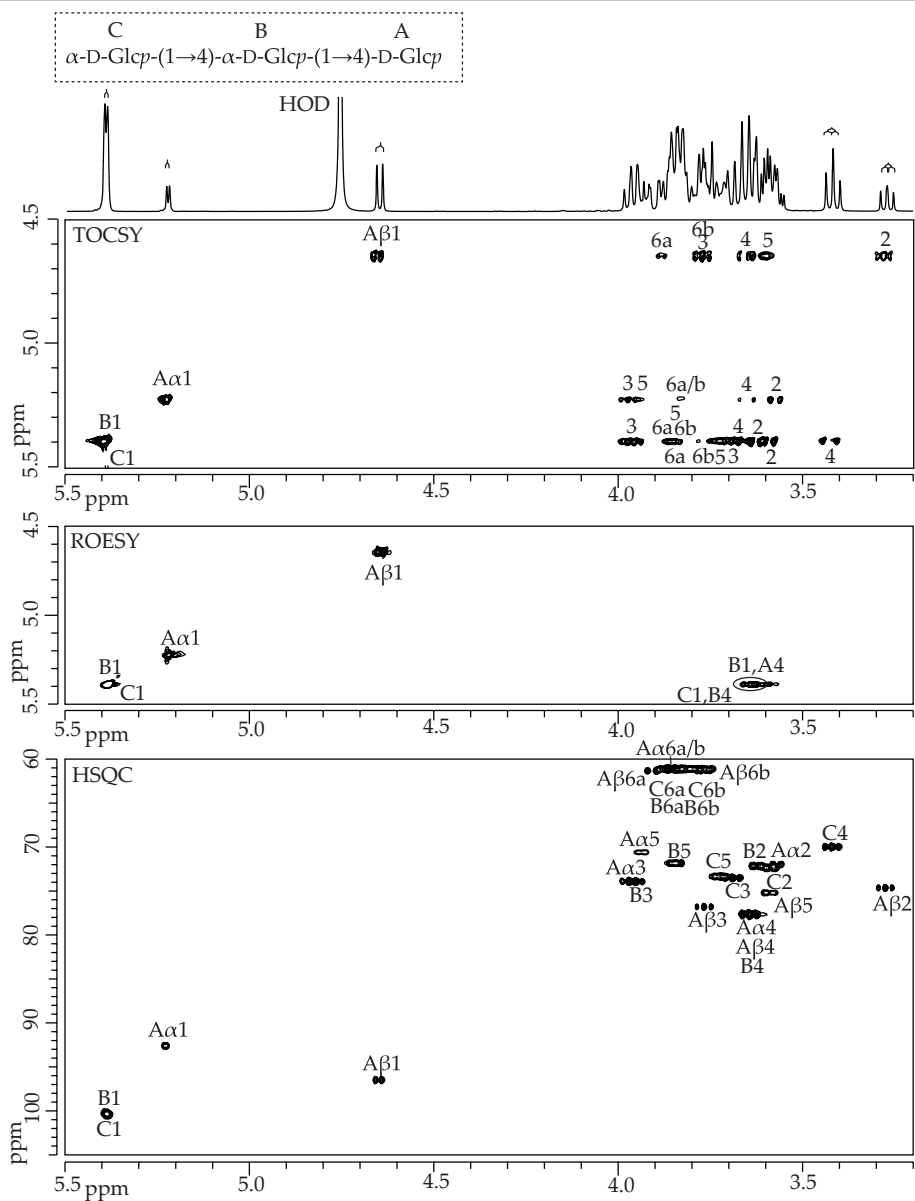


Figure 6. Maltotriose (6): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

Panose, α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp (C** \rightarrow **B** \rightarrow **A**; **8**)**

The 1D ^1H NMR spectrum of panose (Figure 8) showed four anomeric signals at δ 5.399 (**B** H-1, $^3J_{1,2}$ 3.4 Hz), 5.225 (**A α** H-1, $^3J_{1,2}$ 3.9 Hz), 4.955 (**C** H-1, $^3J_{1,2}$ 3.4 Hz), and 4.650 (**A β** H-1, $^3J_{1,2}$ 8.3 Hz); the anomeric ratio amounts to 38:62. The assignments of all protons, as deduced from TOCSY experiments (Figure 8/180 ms), are presented in Table 3. Inspection of the ROESY spectrum (Figure 8) clearly indicated the expected cross-peaks for the (1-6) (**C** H-1,**B** H-6a) and (1-4) (**B** H-1,**A** H-4) linkages.

Correlation of ^1H and ^{13}C signals in the HSQC spectrum (Figure 8) allowed the assignment of all ^{13}C chemical shifts (Table 4). The substitution of **A** O-4 by residue **B** is demonstrated by the downfield shifts of **A α** C-4 ($\delta_{\text{C-4}}$ 78.2; α -D-Glcp, $\delta_{\text{C-4}}$ 71.2) and **A β** C-4 ($\delta_{\text{C-4}}$ 78.2; β -D-Glcp, $\delta_{\text{C-4}}$ 71.2), and the substitution of **B** O-6 by residue **C** by the downfield shift of **B** C-6 ($\delta_{\text{C-6}}$ 66.8; α -D-Glcp-OMe, $\delta_{\text{C-6}}$ 62.2).⁶

Discussion

Substitution of the reducing D-Glcp unit

Comparison of the δ values of the H-1 atoms of the reducing D-Glcp units **A** in the oligosaccharides **1-8** shows that the introduction of an α -D-Glcp substituent at **A** O-2, O-3, O-4, or O-6 influences the exact position of the **A** H-1 resonance (Tables 1 and 3). A 2-substituted reducing Glcp residue [**A**; -(1 \rightarrow 2)-D-Glcp; **1**] is reflected by **A α** and **A β** H-1 signals at δ 5.427 and 4.785, respectively. For a 3-substituted reducing Glcp residue [**A**; -(1 \rightarrow 3)-D-Glcp; **2** and **5**] these values are δ 5.233 and 4.663, respectively; for a 4-substituted reducing Glcp residue [**A**; -(1 \rightarrow 4)-D-Glcp; **3**, **6**, and **8**] δ 5.225 and 4.650, respectively; and for a 6-substituted reducing Glcp residue [**A**; -(1 \rightarrow 6)-D-Glcp; **4** and **7**] δ 5.240 and 4.672, respectively. It may be clear that such differences are useful indicators of the type of substitutions in linear α -D-glucan oligosaccharide analysis.

Besides the chemical shift positions of the H-1 atoms of the reducing Glcp residues, also the positions of reducing β -D-Glcp H-2's, occurring outside the bulk signal (δ 4.00-3.50) at the high-field side (δ 3.25-3.37), are similarly useful indicators of the substitution pattern of the reducing residue (Tables 1 and 3).

The importance of the position of the D-Glcp H-3 signal

The H-3 signal has a unique position for 2-, 3-, and 4-substituted D-Glcp residues (Tables 1 and 3). In **1** the -(1 \rightarrow 2)- α,β -D-Glcp residue **A** has an H-3 signal that is shifted downfield in reference to free α,β -D-Glcp (α -anomer, $\delta_{\text{H-3}}$ 3.71; β -anomer, $\delta_{\text{H-3}}$ 3.48): to

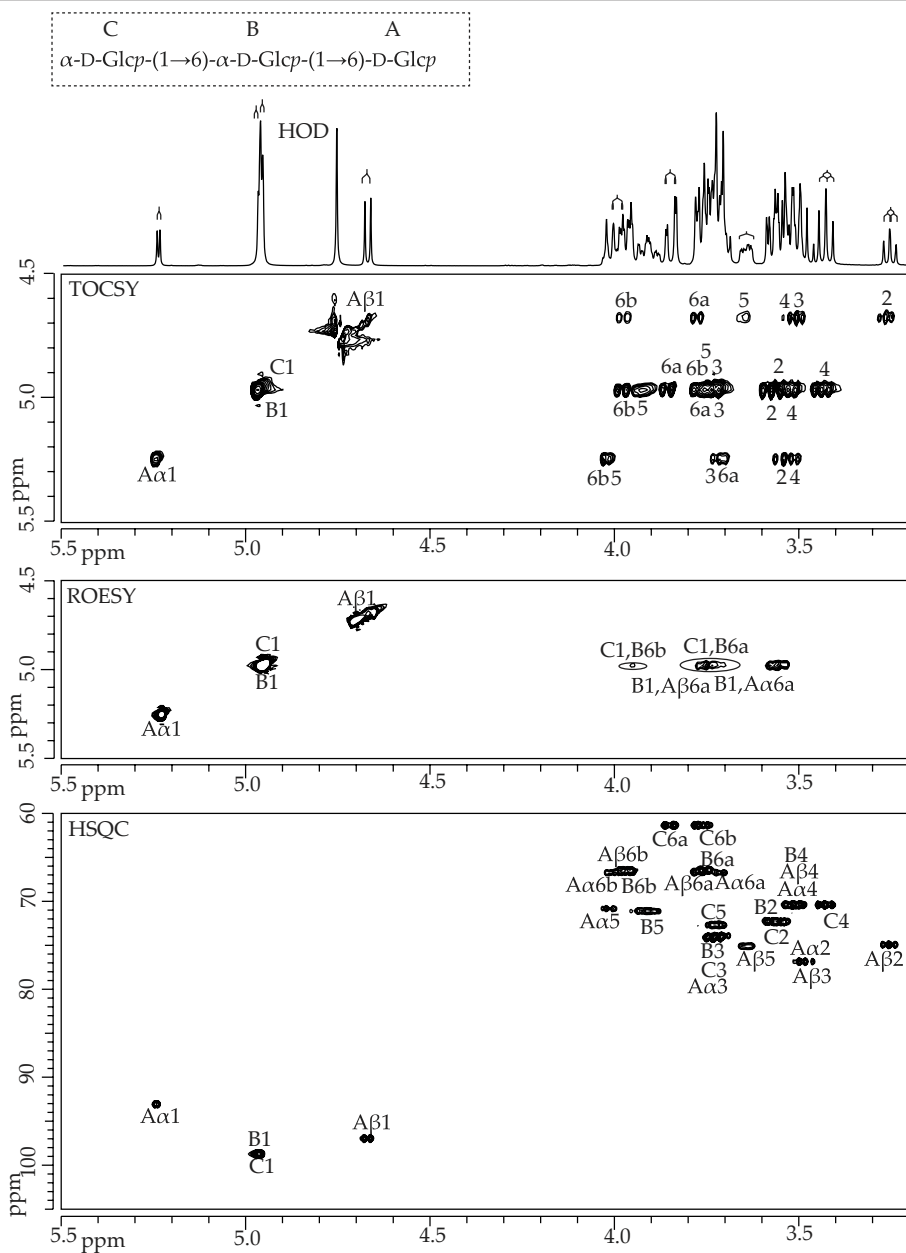


Figure 7. Isomaltotriose (7): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

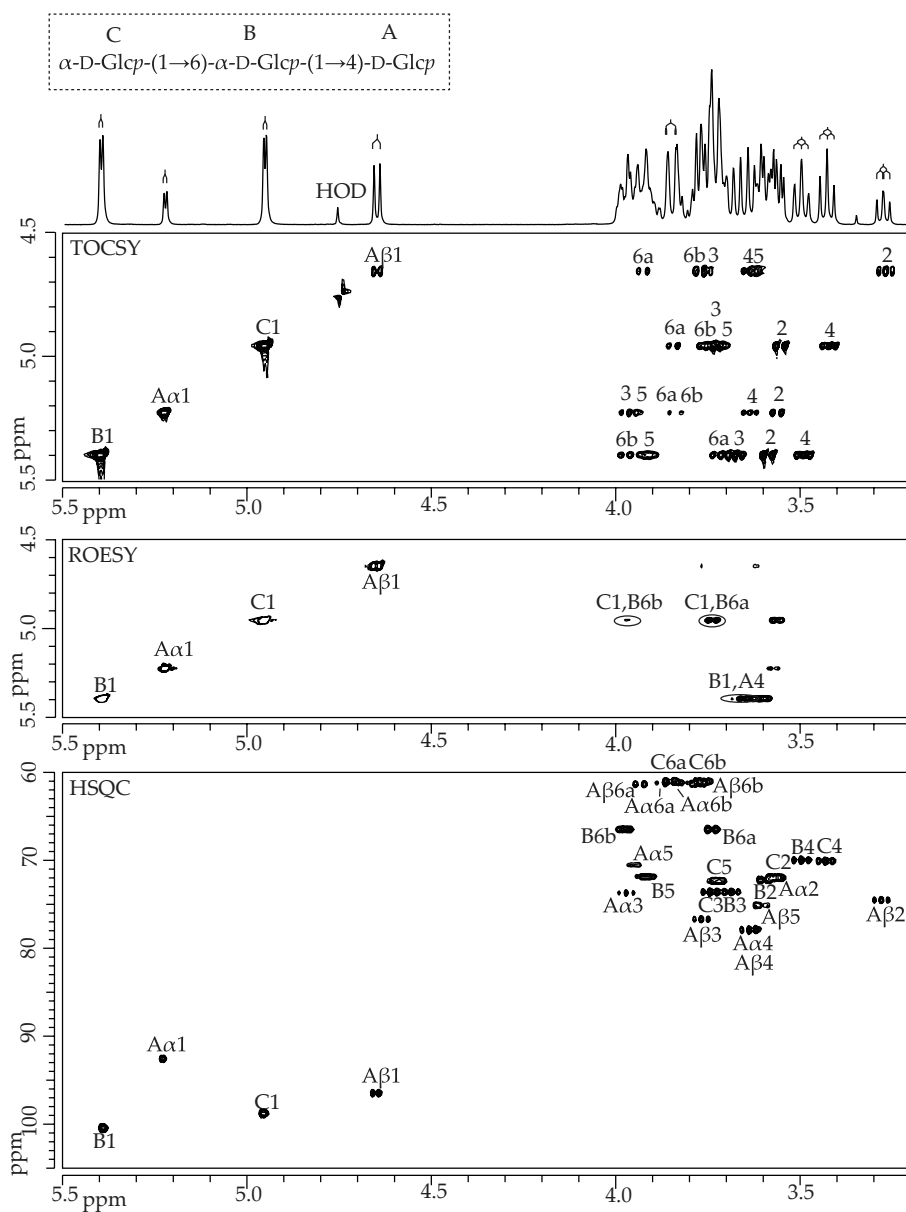


Figure 8. Panose (8): 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 1.

Table 3. ¹H chemical shifts of D-glucopyranose residues of trisaccharides 5 to 8, referenced to internal acetone (δ 2.225). ³J_{1,2} couplings are included between brackets, where measured.

Residue ^a	5	6	7	8
Aα				
H-1	5.233 (3.8)	5.224 (3.4)	5.241 (3.5)	5.225 (3.9)
H-2	3.64	3.57	3.54	3.58
H-3	3.85	3.97	3.72	3.96
H-4	3.66	3.63	3.50	3.63
H-5	3.84	3.94	4.01	3.93
H-6a	3.86	n.d.	3.70	3.86
H-6b	3.77	3.81	4.019 (4.7-12.3)	3.80
Aβ				
H-1	4.663 (7.6)	4.650 (7.8)	4.672 (8.0)	4.650 (8.3)
H-2	3.343 (9.3)	3.278 (9.8)	3.253 (9.2)	3.276 (9.3)
H-3	3.65	3.77	3.48	3.77
H-4	3.64	3.65	3.50	3.66
H-5	3.47	3.63	3.647	3.63
H-6a	3.89	3.91	3.78	3.91
H-6b	3.73	3.75	3.967 (4.7-12.3)	3.78
B				
H-1	5.386/5.371 (3.8)	5.399 (3.4)	4.961/4.969 (3.5)	5.399 (3.4)
H-2	3.68	3.63	3.57	3.60
H-3	3.90	3.96	3.72	3.69
H-4	3.66	3.65	3.49	3.507 (9.3)
H-5	4.03	3.85	3.93	3.91
H-6a	3.85	3.85	3.78	3.72
H-6b	3.74	3.82	3.983 (4.7-12.3)	3.97
C				
H-1	5.359 (3.8)	5.399 (3.4)	4.961 (3.5)	4.955 (3.4)
H-2	3.562 (9.7)	3.59	3.55	3.56
H-3	3.76	3.67	3.72	3.74
H-4	3.431 (9.2)	3.418 (9.8)	3.422 (9.6)	3.432 (9.3)
H-5	4.01	3.72	3.72	3.73
H-6a	3.87	3.84	3.848 (1.9-12.3)	3.844 (2.0-12.3)
H-6b	3.80	3.74	3.78	3.76

^aAα and Aβ stand for the anomeric configuration of residue A

Table 4. ^{13}C chemical shifts of D-glucopyranose residues of trisaccharides **5** to **8**, referenced to the methyl-carbon of internal acetone (δ 31.08).

Residue ^a		5	6	7	8
Aα	C-1	93.2	92.6	93.2	92.8
	C-2	70.9	72.0	72.3	72.3
	C-3	80.6	74.0	73.8	73.9
	C-4	70.9	77.8	70.4	78.2
	C-5	72.3	70.4	70.9	70.7
	C-6	61.4	61.4	66.8	61.3
Aβ	C-1	96.9	96.5	97.0	96.7
	C-2	73.9	74.7	75.0	74.9
	C-3	83.6	76.9	77.1	76.9
	C-4	71.0	77.8	70.4	78.2
	C-5	76.7	75.3	75.1	75.3
	C-6	61.4	61.4	66.8	61.5
B	C-1	100.2	100.4	98.8	100.7
	C-2	70.9	72.2	72.3	72.5
	C-3	80.6	74.0	74.1	73.8
	C-4	70.9	77.8	70.4	70.1
	C-5	72.7	71.8	71.1	72.1
	C-6	61.4	61.4	66.5	66.8
C	C-1	100.2	100.4	98.8	99.0
	C-2	72.9	72.5	72.3	72.4
	C-3	73.9	73.6	74.1	73.8
	C-4	70.6	70.1	70.5	70.4
	C-5	72.9	73.1	72.7	72.5
	C-6	61.4	61.4	61.4	61.3

^a**A α** and **A β** stand for the anomeric configuration of residue **A**

δ 3.80 for **A α** ($\Delta\delta + 0.09$ ppm) and to δ 3.55 for **A β** ($\Delta\delta + 0.07$ ppm). In **2** and **5** the H-3 signal of the $-(1\rightarrow3)\text{-}\alpha,\beta\text{-D-Glcp}$ residue **A** is shifted downfield to δ 3.85 for **A α** ($\Delta\delta + 0.14$ ppm) and to δ 3.64-3.65 for **A β** ($\Delta\delta + 0.16\text{-}0.17$ ppm). The chemical shift of H-3 in the $-(1\rightarrow3)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow3)\text{-}$ unit **B** in **5** shows a similar downfield shift (δ 3.90, $\Delta\delta + 0.15$ ppm) in reference to H-3 of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow3)\text{-}$ unit **B** in **2** (δ 3.75). In **3**, **6**, and **8** the $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}$ residue **A** has an H-3 signal that is shifted downfield to δ 3.96-3.97 for **A α** ($\Delta\delta + 0.25\text{-}0.26$ ppm) and δ 3.76-3.77 for **A β** ($\Delta\delta + 0.28\text{-}0.29$ ppm). Here, the particularly strong downfield shift of **A α** H-3 and **A β** H-3 is the result of an inter-residual hydrogen bond between **B** O-2 and **A** H(O-3).⁷ The H-3 resonance of the $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit **B** in **6** is shifted downfield, in reference to H-3 of $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit **B** in **3** (δ 3.68), to δ 3.96 ($\Delta\delta + 0.28$ ppm).

The importance of the position of the α -D-Glcp H-4 signal

In free α -D-glucopyranose, the H-4 signal resonates at δ 3.40, which is situated outside the bulk region. In non-reducing terminal α -D-Glcp units [α -D-Glcp-(1 \rightarrow x)-], this ^1H signal shows, depending on the glycosidic linkage, only a slight upfield shift (for **1-8**, δ 3.41-3.45). In substituted α -D-Glcp units the H-4 signals may show specific shifts downfield into the bulk region (Tables 1 and 3). In case of a -(1 \rightarrow 2)- α -D-Glcp unit, H-4 resonates at δ 3.45 (**1**), which overlaps with the region of an α -D-Glcp-(1 \rightarrow x)- unit. However, the H-4 signal of -(1 \rightarrow 6)- α -D-Glcp(-) resonates at δ 3.49-3.53 (**4, 7, 8**) and that of -(1 \rightarrow 3)- α -D-Glcp(-) and -(1 \rightarrow 4)- α -D-Glcp(-) at δ 3.63-3.66 (**2, 3, 5, 6, 8**). This means, that as long as there are no (α 1-2) linkages present in an α -D-glucan oligo- or polysaccharide, a peak in the range of δ 3.41-3.45 corresponds with an α -D-Glcp-(1 \rightarrow x) unit. In oligosaccharides, derived from a partial acid hydrolysis of polysaccharides, the surface area under this peak can be used to determine whether the fragment is a branched oligosaccharide or not. In polysaccharides, this peak may be a useful indicator of the degree of branching.

The importance of the position of the D-Glcp H-5 signal

Analysis of the oligosaccharides **1-8** showed interesting features when focusing on the H-5 signals (Tables 1 and 3). Taking free α -D-glucopyranose as a starting point ($\delta_{\text{H-5}}$ 3.84), the H-5 signal of the -(1 \rightarrow 2)- α -D-Glcp unit **A α** in **1** is found at a similar position (δ 3.85), but the H-5 signals of the α -D-Glcp-(1 \rightarrow 2)- units **B α** and **B β** are shifted downfield to δ 3.93 ($\Delta\delta + 0.09$ ppm) and 4.02 ($\Delta\delta + 0.18$ ppm), respectively. Taking free β -D-glucopyranose as a starting point ($\delta_{\text{H-5}}$ 3.45), the H-5 signal of the -(1 \rightarrow 2)- β -D-Glcp unit **A β** in **1** is found at a similar position (δ 3.45). The relatively high downfield position of the **B** H-5 signals is caused by an inter-residual hydrogen bond between **B** O-5 and **A** H(O-3) in both anomeric forms.⁷

Also in **2** the H-5 signals of the reducing -(1 \rightarrow 3)- α,β -D-Glcp unit **A** are found close to the positions of free α,β -D-Glcp H-5. However, H-5 of the α -D-Glcp-(1 \rightarrow 3)-unit **B** is observed at δ 4.01 ($\Delta\delta + 0.17$ ppm). The relatively high downfield position of **B** H-5 is due to the inter-ring hydrogen bond between **B** O-5 and **A** H(O-2).⁷ Such a downfield shift is also found in **5**, showing α -D-Glcp-(1 \rightarrow 3)- unit **C** and -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 3)- unit **B** H-5 at δ 4.01 [**C** O-5...**B** H(O-2)] and 4.03 [**B** O-5...**A** H(O-2)], respectively, whereas the H-5 signals of the -(1 \rightarrow 3)- α,β -D-Glcp **A** unit are similar to those of free α,β -D-Glcp H-5.

In contrast to the H-5 positions of the reducing units in **1** and **2**, those in **3** and **6** show downfield shifts with respect to free α,β -D-Glcp, i.e. $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}$ unit **A α** H-5 appears at δ 3.94 ($\Delta\delta + 0.10$ ppm), and $-(1\rightarrow4)\text{-}\beta\text{-D-Glcp}$ unit **A β** H-5 at δ 3.60-3.63 ($\Delta\delta + 0.15\text{-}0.18$ ppm). The H-5 signals of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit **B** in **3** and **C** in **6** are detected at the upfield positions δ 3.71-3.72 ($\Delta\delta - 0.13\text{-}0.12$ ppm); the H-5 resonance of the $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit **B** in **6** shifts significantly downfield (δ 3.85, $\Delta\delta + 0.14$ ppm) in reference to the H-5 of $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit **B** in **3** (δ 3.71).

In a similar way as shown for **3** and **6**, the H-5 signals of the $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}$ unit **A α** and the $-(1\rightarrow6)\text{-}\beta\text{-D-Glcp}$ unit **A β** in **4** and **7** show downfield shifts: **A α** H-5 at δ 4.01 ($\Delta\delta + 0.17$ ppm) and **A β** H-5 at δ 3.64-3.65 ($\Delta\delta + 0.19\text{-}0.20$ ppm). The H-5 signal of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit **B** in **4** and **C** in **7** are detected at the upfield positions δ 3.75-3.72 ($\Delta\delta - 0.09\text{-}0.12$ ppm); the H-5 resonance of the $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit **B** in **7** shifts significantly downfield (δ 3.93, $\Delta\delta + 0.18$ ppm) in reference to H-5 of $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit **B** in **4** (δ 3.75).

The deduced information with respect to H-5 fits completely with the situation in **8** (panose). Here, **A α** and **A β** H-5 have chemical shifts of δ 3.93 and 3.63, respectively, in accordance with a reducing $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}$ **A** unit (compare with **4** and **6**). The **B** H-5 signal appears at δ 3.91, reflecting a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4/6)\text{-}$ unit (in this case $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit **B**) and the **C** H-5 resonance is observed at δ 3.73, reflecting an $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4/6)\text{-}$ unit (in this case $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit **C**).

The linearity of ^1H chemical shifts

Comparison of the H-1 – H-6a,b chemical shift patterns (i.e. TOCSY patterns) of the non-reducing terminal $\alpha\text{-D-Glcp}\text{-}(1\rightarrow x)\text{-}$ residues in the analysed di- and trisaccharides reveals overlapping features for the same x. The chemical shift pattern of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow3)\text{-}$ residue **C** in trisaccharide **5** corresponds with that of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow3)\text{-}$ residue **B** in disaccharide **2**. The chemical shift pattern of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ residue **C** in trisaccharide **6** corresponds with that of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ residue **B** in disaccharide **3**. The chemical shift pattern of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ residue **C** in trisaccharides **7** and **8** corresponds with that of the $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ residue **B** in disaccharide **4**.

Compared to free $\alpha\text{-D-Glcp}$, in disaccharide **2** (nigerose), the **A α** H-2, H-3, and H-4 signals are shifted downfield to δ 3.63 ($\delta_{\text{H-2}}$ 3.52; $\Delta\delta + 0.11$ ppm), 3.85 ($\delta_{\text{H-3}}$ 3.71; $\Delta\delta + 0.14$ ppm), and 3.65 ($\delta_{\text{H-4}}$ 3.40; $\Delta\delta + 0.25$ ppm), respectively. In trisaccharide **5** (nigerotriose) the **B** H-2, H-3, and H-4 resonances are shifted downfield in reference to

α -D-Glcp-(1 \rightarrow 3)- unit **B** in **2**, to δ 3.68 ($\delta_{\text{H-2}}$ 3.57; $\Delta\delta$ + 0.11 ppm), 3.90 ($\delta_{\text{H-3}}$ 3.75; $\Delta\delta$ + 0.15 ppm), and 3.66 ($\delta_{\text{H-4}}$ 3.44; $\Delta\delta$ + 0.22 ppm), respectively.

Compared to free α -D-Glcp, in disaccharide **3** (maltose), the **A α** H-3, H-4, and H-5 signals are shifted downfield to δ 3.97 ($\delta_{\text{H-3}}$ 3.71; $\Delta\delta$ + 0.26 ppm), 3.64 ($\delta_{\text{H-4}}$ 3.40; $\Delta\delta$ + 0.24 ppm), and 3.94 ($\delta_{\text{H-5}}$ 3.84; $\Delta\delta$ + 0.10 ppm), respectively. In trisaccharide **6** (maltotriose) the **B** H-3, H-4, and H-5 resonances are shifted downfield, in reference to α -D-Glcp-(1 \rightarrow 4)- unit **B** in **3**, to δ 3.96 ($\delta_{\text{H-3}}$ 3.68; $\Delta\delta$ + 0.28 ppm), 3.65 ($\delta_{\text{H-4}}$ 3.41; $\Delta\delta$ + 0.24 ppm), and 3.85 ($\delta_{\text{H-5}}$ 3.71; $\Delta\delta$ + 0.14 ppm), respectively.

Compared to free α -D-Glcp, in disaccharide **4** (isomaltose), the **A α** H-4, H-5, H-6a, and H-6b signals are shifted to δ 3.53 ($\delta_{\text{H-4}}$ 3.40; $\Delta\delta$ + 0.13 ppm), 4.01 ($\delta_{\text{H-5}}$ 3.84; $\Delta\delta$ + 0.17 ppm), 3.70 ($\delta_{\text{H-6a}}$ 3.83; $\Delta\delta$ - 0.13 ppm), and 4.00 ($\delta_{\text{H-6b}}$ 3.74; $\Delta\delta$ + 0.26 ppm), respectively. In trisaccharide **7** (isomaltotriose) the **B** H-4, H-5, H-6a, and H-6b resonances are shifted, in reference to α -D-Glcp-(1 \rightarrow 6)- unit **B** in **4**, to δ 3.49 ($\delta_{\text{H-4}}$ 3.42; $\Delta\delta$ + 0.07 ppm), 3.93 ($\delta_{\text{H-5}}$ 3.75; $\Delta\delta$ + 0.18 ppm), 3.78 ($\delta_{\text{H-6a}}$ 3.85; $\Delta\delta$ - 0.07 ppm), and 3.98 ($\delta_{\text{H-6b}}$ 3.78; $\Delta\delta$ + 0.20 ppm), respectively. In trisaccharide **8** (panose) the **B** H-4, H-5, H-6a, and H-6b signals are shifted, in reference to α -D-Glcp-(1 \rightarrow 4)- unit **B** in **3**, to δ 3.51 ($\delta_{\text{H-4}}$ 3.41; $\Delta\delta$ + 0.10 ppm), 3.91 ($\delta_{\text{H-5}}$ 3.71; $\Delta\delta$ + 0.20 ppm), 3.72 ($\delta_{\text{H-6a}}$ 3.85; $\Delta\delta$ - 0.13 ppm), and 3.97 ($\delta_{\text{H-6b}}$ 3.76; $\Delta\delta$ + 0.21 ppm), respectively.

Comparison of the chemical shifts of residue **B** in trisaccharides **5-8** with those in disaccharides **1-4**, shows that the size and direction (up- or downfield) of the shifts upon substitution are similar to the size and direction of the shifts of residue **A α** of **1-8** in reference to free α -D-Glcp.

Conclusions

The generated ^1H and ^{13}C NMR data, recorded at 300K, fit earlier reported data.^{5,6,8-11} It should be noted that for the development of the CASPER computer program the ^1H NMR data were recorded at 340K,¹²⁻¹⁴ yielding some differences, especially where inter-residual hydrogen bonds are involved. In principle, the CASPER computer program was developed to predict repeating-unit structures and oligosaccharides when all chemical shifts are available, in combination with substitution pattern information from methylation analysis. For homopolysaccharides, lacking a repeating unit, the CASPER computer program cannot be used for structure prediction. However, the CASPER computer program was found to be a useful tool for verification of the assigned ^1H resonances.

For the analysis of α -D-glucan oligo- and polysaccharides, in this study a series of structural reporters have been defined that can be used in 1D ^1H NMR spectroscopy. The substitution of the reducing residue $[-(1\rightarrow x)\text{-}\alpha\text{-D-Glc}p]$ can be derived directly from the anomeric ^1H chemical shifts. The amount of branching can be derived from the surface area under the peak around δ 3.40, provided there are no (α 1-2) linkages. Although most of the chemical shifts fall inside the bulk-region of δ 4.00-3.50, the 2D TOCSY H-1 – H-6a,b patterns of each residue can easily be used for the identification of substitution patterns. Even when using 2D TOCSY experiments with short mixing times (7 ms – 60 ms), the specific resonances of H-2, H-3, and H-4 will often suffice to elucidate an oligosaccharide structure. This means that when only minute amounts of oligosaccharide or polysaccharide sample are available, or low solubility impedes 2D NMR spectroscopy, structures can still be solved.

Even though ^{13}C - ^1H HSQC spectra do not seem to be strictly necessary for identification of α -D-glucans, the data from these experiments are useful for verification of assignments made by 2D TOCSY. The ^{13}C chemical shift is more predictable than the ^1H chemical shift, but can only be used when sufficient amounts of material are available, or with ^{13}C -enriched samples.

The developed ^1H structural-reporter-group concept has been used successfully in the primary structural analysis of α -D-glucans secreted by *Lb. reuteri* strains, as will be reported elsewhere.

Materials and Methods

Materials

Nigerose, maltose, maltotriose, isomaltose, isomaltotriose, and panose were purchased from Sigma ($\geq 95\%$, HPAEC-PAD). Nigerotriose was purchased from Dextra Laboratories Ltd., Reading, UK. Kojibiose was a gift of Dr. B.H. Koeppen (South Africa). D_2O (99.9 atom%) was acquired from Cambridge isotope laboratories, Inc., Andover, MA.

NMR Spectroscopy

^1H NMR spectra, including ^1H - ^1H and ^{13}C - ^1H correlation spectra were recorded at a probe temperature of 300K on a Bruker DRX500 spectrometer (Bijvoet Center, Department of NMR Spectroscopy). Samples were exchanged once with 99.9 atom% D_2O , lyophilised, and dissolved in 650 μL D_2O . ^1H chemical shifts are expressed in ppm

by reference to internal acetone (δ 2.225), and ^{13}C chemical shifts are expressed in ppm by reference to the methyl-carbon of internal acetone (δ 31.08). One-dimensional 500-MHz ^1H NMR spectra were recorded with a spectral width of 5000 Hz in 16k complex data sets and zero filled to 32k. A WEFT pulse sequence was applied to suppress the HOD signal.¹⁵ When necessary, a fifth order polynomial baseline correction was applied. Two-dimensional TOCSY spectra were recorded using MLEV17 mixing sequences with spin-lock times of 10, 30, 60, 120, and 180 ms.¹⁶ The spin-lock field strength corresponded to a 90° pulse width of about 28 μs at 13 dB. The spectral width in the 2D TOCSY experiments was 4006 Hz at 500 MHz in each dimension. 400-1024 spectra of 2k data points with 8-32 scans per t_1 increment were recorded. 2D rotating-frame nuclear Overhauser enhancement spectra (ROESY) were recorded with 300 ms mixing time.¹⁷ The spectral width was 4006 Hz at 500 MHz in each dimension. Suppression of the HOD signal was performed by a 1 s pre-saturation during the relaxation delay. Between 400 and 1024 data sets of 2k data points were recorded with 8-16 scans per t_1 increment. 2D ^{13}C - ^1H HSQC spectroscopy was carried out at a ^1H frequency of 500.0821 MHz and a ^{13}C frequency of 125.7552 MHz.¹⁸ Spectra were recorded with a spectral width of 4006 Hz for t_2 and 10,000 Hz for t_1 . The HOD signal was pre-saturated for 1 s, and ^{12}C -bound protons were suppressed using a TANGO pulse sequence. During acquisition of the ^1H FID, a ^{13}C decoupling pulse was applied. 128-256 experiments of 2k data points were recorded with 128 scans per t_1 increment. 2D NMR spectroscopic data were analysed by applying a sinus multiplication window and zero filling to spectra of 4k by 1k dimensions. In case of ^{13}C - ^1H HSQC data, the spectra were zero filled to 4k by 512 data points. A Fourier transform was applied, and where necessary, a fifth to fifteenth order polynomial baseline function was applied. All NMR data were processed using in-house developed software (J.A. van Kuik, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University).

References

1. Van Hijum, S. A. F. T.; Kralj, S.; Ozimek, L. K.; Dijkhuizen, L.; van Geel-Schutten, G. H. *Microbiol. Mol. Biol. Rev.*, **2006**, *70*, 157-176.
2. Kralj, S.; van Geel-Schutten, G. H.; Rahaoui, H.; Leer, R. J.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Appl. Env. Microbiol.*, **2002**, *68*, 4283-4291.
3. Kralj, S.; van Geel-Schutten, G. H.; Dondorff, M. M. G.; Kirsanovs, S.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Microbiology*, **2004**, *150*, 3681-3690.
4. Kralj, S.; van Geel-Schutten, G. H.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Biochemistry*, **2005**, *44*, 9206-9216.
5. Bock, K.; Pedersen, C. *Adv. Carbohydr. Chem. Biochem.*, **1983**, *41*, 27-66.

6. Bock, K.; Thøgersen, H. *Ann. Rep. NMR Spectr.*, **1982**, *13*, 2-57.
7. Dowd, M. K.; Zeng, J.; French, A. D.; Reilly, P. J. *Carbohydr. Res.*, **1992**, *230*, 223-244.
8. Backman, I.; Erbing, B.; Jansson, P.-E.; Kenne, L. *J. Chem. Soc. Perkin Trans. 1*, **1988**, 889-898.
9. Jansson, P.-E.; Kenne, L.; Schweda, F. *J. Chem. Soc. Perkin Trans. 1*, **1988**, 2792-2736.
10. Adeyeye, A.; Jansson, P.-E.; Kenne, L.; Widmalm, G. *J. Chem. Soc. Perkin Trans. 2*, **1991**, 963-973.
11. Jansson, P.-E.; Kenne, L.; Kolare, I. *Carbohydr. Res.*, **1994**, *257*, 163-176.
12. Jansson, P.-E.; Kenne, L.; Widmalm, G. *Carbohydr. Res.*, **1987**, *168*, 67-77.
13. Jansson, P.-E.; Kenne, L.; Widmalm, G. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 508-516.
14. Stenutz, R.; Jansson, P.-E.; Widmalm, G. *Carbohydr. Res.*, **1998**, *306*, 11-17.
15. Hård, K.; van Zadelhoff, G.; Moonen, P.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.*, **1992**, *209*, 895-915.
16. Bax, A.; Davies, D. G. *J. Magn. Reson.*, **1985**, *65*, 355-360.
17. Bax, A.; Davies, D. G. *J. Magn. Reson.* **1981**, *63*, 207-213.
18. Wilder, G.; Wüthrich, K. *J. Magn. Reson.*, **1991**, *102*, 239-214.

*It is inevitable that many ideas of the young mind will later have to give way to the
hard realities of life*

-Felix Bloch-

Chapter 3

Structural analysis of the α -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme

Sander S. van Leeuwen,¹ Slavko Kralj,^{2,3} Ineke H. van Geel-Schutten,^{3,4}
Gerrit J. Gerwig,¹ Lubbert Dijkhuizen,^{2,3} and Johannes P. Kamerling¹

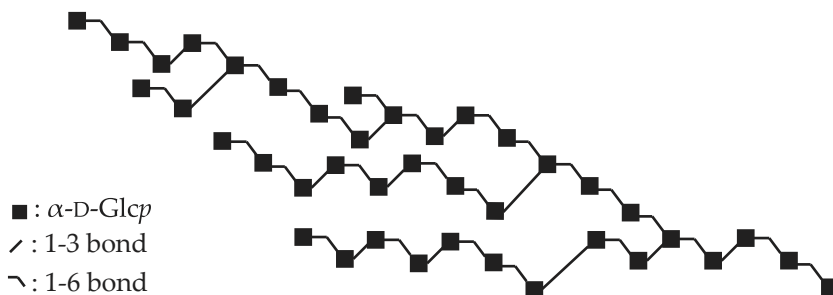
¹*Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, 3584 CH Utrecht, The Netherlands*

²*Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands*

³*Centre for Carbohydrate Bioprocessing, TNO - University of Groningen, 9750 AA Haren, The Netherlands*

⁴*Current address: Materials Innovation Centre B.V., 2628 BL Delft, The Netherlands*

Abstract - The neutral exopolysaccharide **EPS180** produced from sucrose by the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* 180 was found to be a (1→3, 1→6)- α -D-glucan, with no repeating units present. Based on linkage analysis, periodate oxidation, and 1D/2D ^1H and ^{13}C NMR spectroscopy of the intact **EPS180**, as well as MS and NMR analysis of oligosaccharides obtained by partial acid hydrolysis of **EPS180**, a visual representation, that includes all identified structural features, was formulated as follows.



Introduction

Lactic acid bacteria (LAB), like *Lactobacilli*, produce exopolysaccharides (EPSs) that are excreted into their surroundings. The complex functions of the EPSs *in vivo* are not fully understood but seem to be mostly of a protective nature. Exopolysaccharides have been shown to be involved in adhesion,¹ certain cellular recognition processes,² and protection against dehydration, phagocytosis or toxins.² It is unlikely that these EPSs are being produced as a food reserve, as most LAB do not have the enzymes to catabolise these EPSs.³ Because of their physical properties, that lie at the basis of their protective nature, the food and dairy industry is interested in these exopolysaccharides as thickeners, stabilisers, and gelling agents.⁴

So far, most structurally characterised EPSs from LAB are heteropolysaccharides with repeating-unit structures.⁵ Structures of LAB homopolysaccharides produced by specific glucansucrase enzymes have not been extensively investigated. Only initial structural studies have been performed on dextrans (polysaccharides containing mostly (α 1-6) linkages) with (α 1-2), (α 1-3), and (α 1-4) branches,^{6,7} or mostly linear homopolysaccharides.⁸

Recently, a family of glucansucrases was discovered in *Lb. reuteri*, which converts sucrose into large, heavily branched α -glucans. One of these glucansucrases (GTF180) produces EPS180, an α -glucan with (α 1-3) and (α 1-6) glycosidic linkages.⁹ The GTF180 enzyme shows large similarity with other glucansucrase enzymes, but has a relatively large N-terminal variable region. Truncation of the enzyme, by deletion of the variable region, had no effect on the linkage distribution of the α -glucan produced.⁹ The unique polysaccharide structure produced was suggested to be pre-biotic.¹⁰ Here, we describe the structural analysis of this homopolysaccharide, leading to the formulation of a visual representation that includes the various structural elements established.

Results

Composition of EPS180

Monosaccharide analysis of **EPS180** revealed the presence of glucose only, and a carbohydrate content of 100% (w/w). Methylation analysis of **EPS180** showed the presence of terminal, 3-substituted, 6-substituted, and 3,6-disubstituted glucopyranose in a molar percentage of 12, 24, 52, and 12%. 1D ^1H NMR spectroscopy (Figure 1A) indicated an α -anomeric configuration for all glucose residues. Of these residues, 69% is involved in (α 1-6) linkages ($\delta_{\text{H-1}} \sim 4.96$) and 31% in (α 1-3) linkages ($\delta_{\text{H-1}} 5.33$), which is in agreement with the linkage distribution found by methylation analysis (see also Ref. 9). In the 1D ^1H NMR spectrum recorded at 330K (Figure 1B), the (α 1-6) anomeric signal is split into two overlapping peaks, while the (α 1-3) anomeric signal remains a single peak. This may suggest that the $(-)\alpha\text{-D-Glc}\beta\text{-}(1\rightarrow 3)^*$ residues in **EPS180** are present in a uniform structural element. The splitting of the (α 1-6) anomeric ^1H signal in the spectrum recorded at 330K indicates at least two specific structural elements for $(-)\alpha\text{-D-Glc}\beta\text{-}(1\rightarrow 6)\text{-}$ residues.

Partial acid hydrolysis

After having evaluated different acid hydrolysis conditions, making use of MALDI-TOF-MS and ^1H NMR analyses, a large batch of **EPS180** (500 mg) was subjected to partial acid hydrolysis, using 0.5 M TFA (3 h, 90 °C). The latter conditions generated a pool of oligosaccharides with different degrees of polymerisation, as evidenced by ^1H NMR spectroscopy (Figure 1C) and HPAEC on CarboPac PA-100 (Figure 2A). The 1D ^1H NMR spectrum of the pool (Figure 1C) showed a strong reduction of the anomeric signal corresponding to (α 1-3)-linked Glc residues, while the relative amount of (α 1-6)-linked Glc residues was only slightly reduced, indicating that the (α 1-3) linkages are more susceptible to acid hydrolysis with TFA than the (α 1-6) linkages. The glycan pool was pre-fractionated on Bio-Gel P-4, and three subpools were isolated, denoted **I** - **III**. The fragment size distribution within each subpool was determined by MALDI-TOF-MS (data not shown). Subpool **I** contained fragments in a size range up to seven Glc units, subpool **II** fragments from five up to nine glucose units, and subpool **III** from seven Glc units and larger. As subpool **III** consisted mainly of oligosaccharides too

*For Glc residues at semi-defined places in the structure $(-)\alpha\text{-D-Glc}\beta\text{-}(1\rightarrow x)\text{-}$ or $-(1\rightarrow x)\text{-}\alpha\text{-D-Glc}\beta\text{-}(-)$ is used. When the structural context of the residue is precisely known, this is indicated as follows: $-(1\rightarrow x)\text{-}\alpha\text{-D-Glc}\beta\text{-}(1\rightarrow y)\text{-}$ describing an x-substituted residue with an (1-y) linkage. In case of a non-reducing terminal residue $\alpha\text{-D-Glc}\beta\text{-}(1\rightarrow x)\text{-}$ is used, a reducing terminal residue is indicated with $-(1\rightarrow x)\text{-D-Glc}\beta\text{-}$.

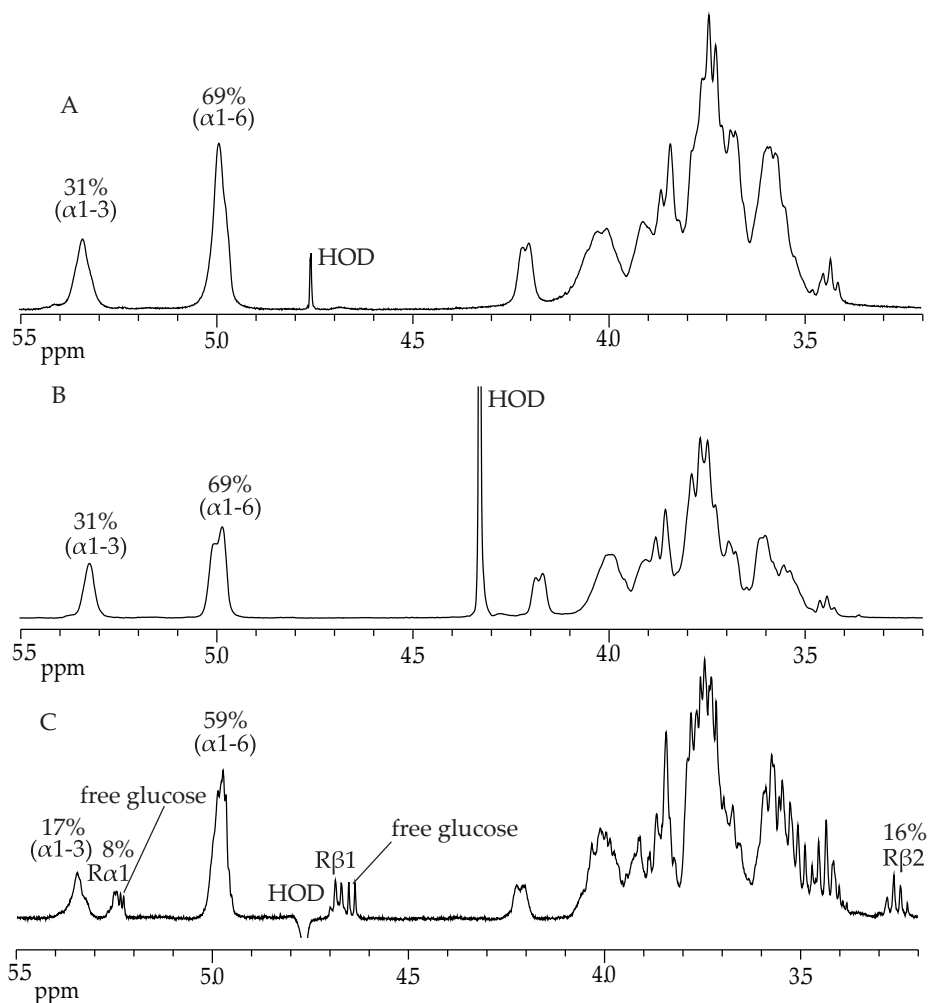


Figure 1. 500-MHz 1D ^1H NMR spectra of (A) intact EPS180 recorded at 300K in D_2O , (B) intact EPS180 recorded at 330K in D_2O , and (C) EPS180 partial acid hydrolysate recorded at 300K in D_2O .

large for full structural analysis, it was not further studied. Subpools **I** and **II** were subfractionated by HPAEC on CarboPac PA-1, using a 0 - 300 mM NaOAc gradient in 100 mM NaOH (Figure 2B), and fractions **1** to **6** were isolated from subpool **I**, and **7** and **8** from subpool **II**.

Fraction 1 - ^1H NMR analysis of fraction **1** revealed the presence of only one compound, that corresponds with isomaltotriose, α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1) (Chapter 2).¹¹

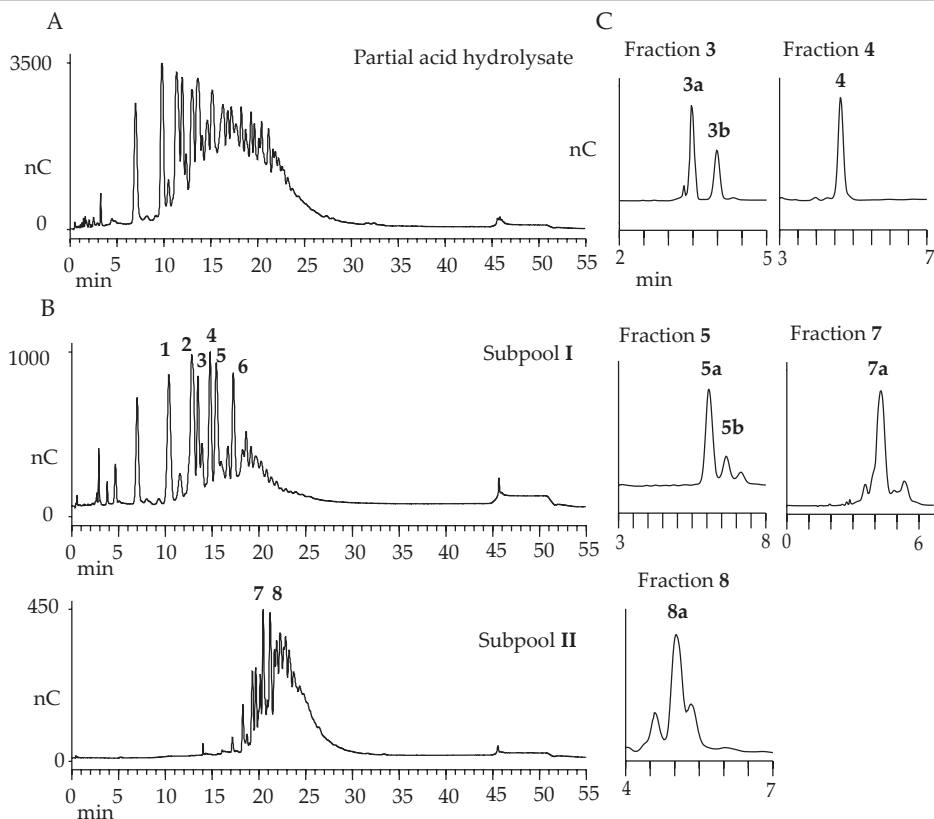


Figure 2. (A) HPAEC profile of **EPS180** partial acid hydrolysate on CarboPac PA-100; (B) HPAEC profiles of Bio-Gel P-4 subpool **I** and subpool **II** on CarboPac PA-1, using a linear gradient; and (C) HPAEC profiles of HPAEC fractions **3**, **4**, **5**, **7**, and **8** on CarboPac PA-1, using isocratic conditions. For experimental details, see Materials and Methods.

Fraction 2 - MALDI-TOF-MS analysis of fraction **2** showed an $[M+Na]^+$ pseudomolecular ion at m/z 689, corresponding with Hex_4 . The 1H NMR spectrum showed a multiple anomeric signal, corresponding to three protons, at $\delta \sim 4.96$, reflecting the presence of α -D-Glcp-(1 \rightarrow 6)- (residue **D**) and $-\alpha$ -D-Glcp-(1 \rightarrow 6)- (residues **C**) units (Chapter 2).¹¹ Furthermore, the H-1 α and H-1 β resonances of the reducing unit **R** (δ 5.240 and 4.669) are in agreement with a $-(1\rightarrow6)$ -D-Glcp unit (Chapter 2).¹¹ Therefore, fraction **2** contains isomaltotetraose, α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

Fraction 3 - The MALDI-TOF mass spectrum of fraction **3** revealed $[M+Na]^+$ pseudomolecular ions at m/z 689 and 851, corresponding with Hex_4 and Hex_5 , respectively. Fraction **3** was further separated by HPAEC on CarboPac PA-1, isocratically

eluted with 100 mM NaOAc in 100 mM NaOH (Figure 2C), yielding two fractions denoted **3a** and **3b**.

The MALDI-TOF mass spectrum of fraction **3a** showed one peak at m/z 851, corresponding to the $[M+Na]^+$ pseudomolecular ion of Hex₅. The ¹H NMR spectrum (Figure 3A) revealed one multiple anomeric signal at δ ~4.96 (4 H) and two anomeric signals corresponding to the reducing residue **R** at δ 5.241 and 4.671. On guidance of the earlier developed structural-reporter-group concept (Chapter 2),¹¹ the signal at δ ~4.96 indicates the presence of α -D-Glcp-(1 \rightarrow 6)- (residue **D**) and $-\alpha$ -D-Glcp-(1 \rightarrow 6)- (residues **C**) units. The chemical shift positions of H-1 α and H-1 β of the reducing residue **R** are in agreement with the occurrence of a -(1 \rightarrow 6)-D-Glcp unit. Considering all analytical data, fraction **3a** is identified as isomaltopentaose, α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

MALDI-TOF-MS analysis of fraction **3b** gave rise to one $[M+Na]^+$ pseudomolecular ion at m/z 689, corresponding with Hex₄. The 1D ¹H NMR spectrum (Figure 4) showed five anomeric signals at δ 5.345/5.336 (**B** H-1, ³J_{1,2} 3.8 Hz), 5.249 (**R α** H-1, ³J_{1,2} 3.8 Hz), 4.965 (**D** H-1, ³J_{1,2} 3.8 Hz), 4.958 (**A** H-1, ³J_{1,2} 3.8 Hz), and 4.680 (**R β** H-1, ³J_{1,2} 7.6 Hz). The splitting of the **B** H-1 signal into two doublets is due to the influence of the α/β configuration of the reducing residue **R**. Note that the intensity of the **R β** H-1 signal is influenced by the pre-saturation of the HOD signal. Therefore, in quantifications the surface area of **R β** H-2 (δ ~3.27) is used. Starting from the anomeric signals in the 2D ¹H-¹H TOCSY spectrum (Figure 4/180 ms), all chemical shifts of the non-anomeric protons could be determined (Table 1). The **R α** and **R β** H-1 values fit best with those of a -(1 \rightarrow 6)-D-Glcp unit (H1 α , δ 5.239-5.241, H1 β , δ 4.671-4.672) (Chapter 2),¹¹ which is further supported by the chemical shifts of **R α** and **R β** H-5 at δ 4.00 and 3.63, respectively (library data: δ 4.01 and 3.64-3.65);¹¹ compare also the H-6a and H-6b values with library data.¹¹ Although the anomeric signals of **A** and **D** strongly overlap, the difference in H-2, H-3, and H-4 between the two residues can clearly be observed in the built-up series of mixing times (data not shown). The chemical shifts of the set of **A** H-2, H-3, and H-4 at δ 3.64, 3.85, and 3.67, respectively, correspond with those of a -(1 \rightarrow 3)- α -D-Glcp(-) unit. The presence of **D** H-4 at δ 3.42 (dd, 1 H) reflects the occurrence of one terminal α -D-Glcp-(1 \rightarrow x)- unit, most probably α -D-Glcp-(1 \rightarrow 6)- instead of α -D-Glcp-(1 \rightarrow 3)- (Chapter 2),¹¹ indicating a linear structure; note that the H-4 signals of residues **A** and **B** do not resonate at this value. Taking into account that **R** is -(1 \rightarrow 6)-D-Glcp (see above), the significant downfield shift of **B** H-5 (δ 4.20) is indicative for an internal $-\alpha$ -D-Glcp-**B**-(1 \rightarrow 3)- α -D-Glcp- element (Chapter 2).¹¹ It

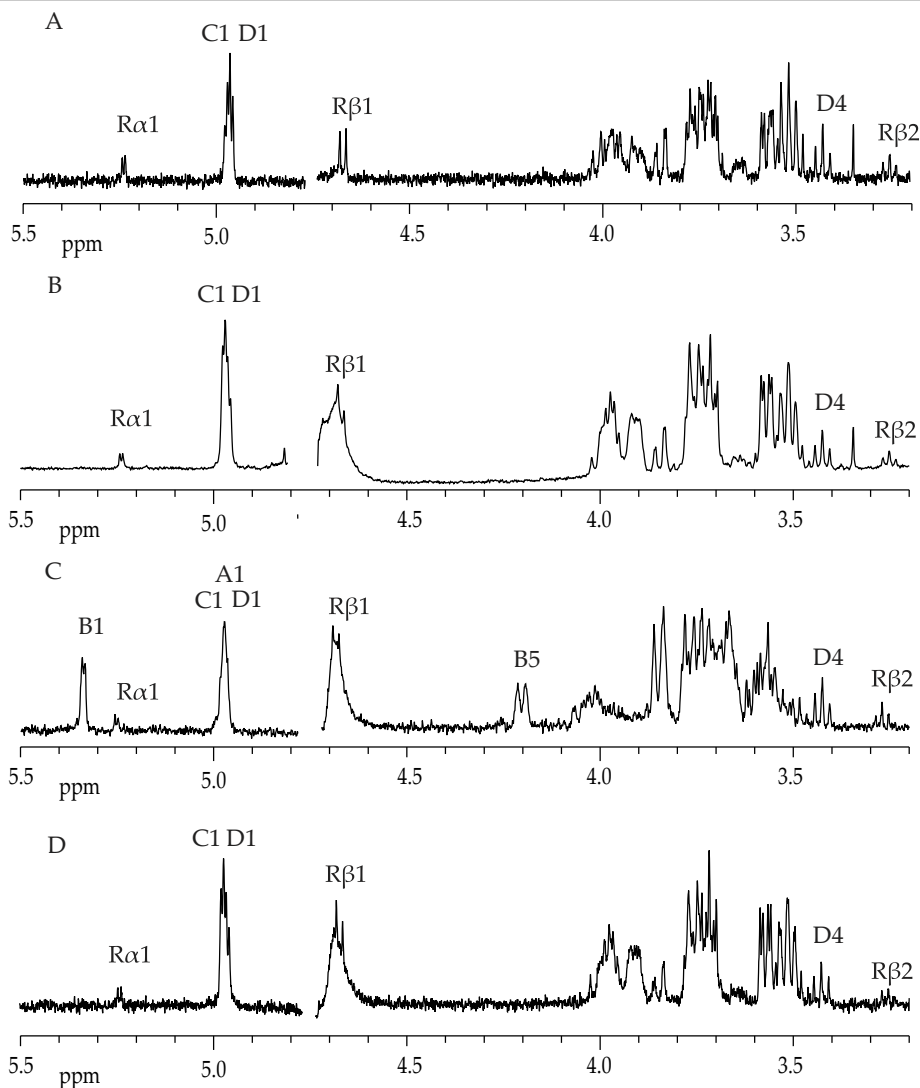


Figure 3. 500-MHz 1D ^1H NMR spectra of (A) fraction **3a**, (B) fraction **5a**, (C) fraction **5b**, and (D) fraction **6**, recorded at 300K in D_2O .

should be noted that the downfield shift of **B** H-5 is caused by the occurrence of a hydrogen bond between **B** O-5 and H(O)-2 of the neighbouring Glc residue.¹² When compared with the chemical shift of H-5 of the internal Glc residue in nigerotriose (δ 4.01),¹¹ a further downfield shift is seen for H-5, which is in favour of a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp-B-(1}\rightarrow3)\text{-}$ instead of a $-(1\rightarrow3)\text{-}\alpha\text{-D-Glcp-B-(1}\rightarrow3)\text{-}$ element. As residue **D** occurs in terminal position, the **B** H-5 signal indicates a **B1** \rightarrow **3A** sequence, which means that the total sequence has to be **D1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6R**, i.e. $\alpha\text{-D-Glcp-(1}\rightarrow6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow6)\text{-D-Glcp}$ (Scheme 1).

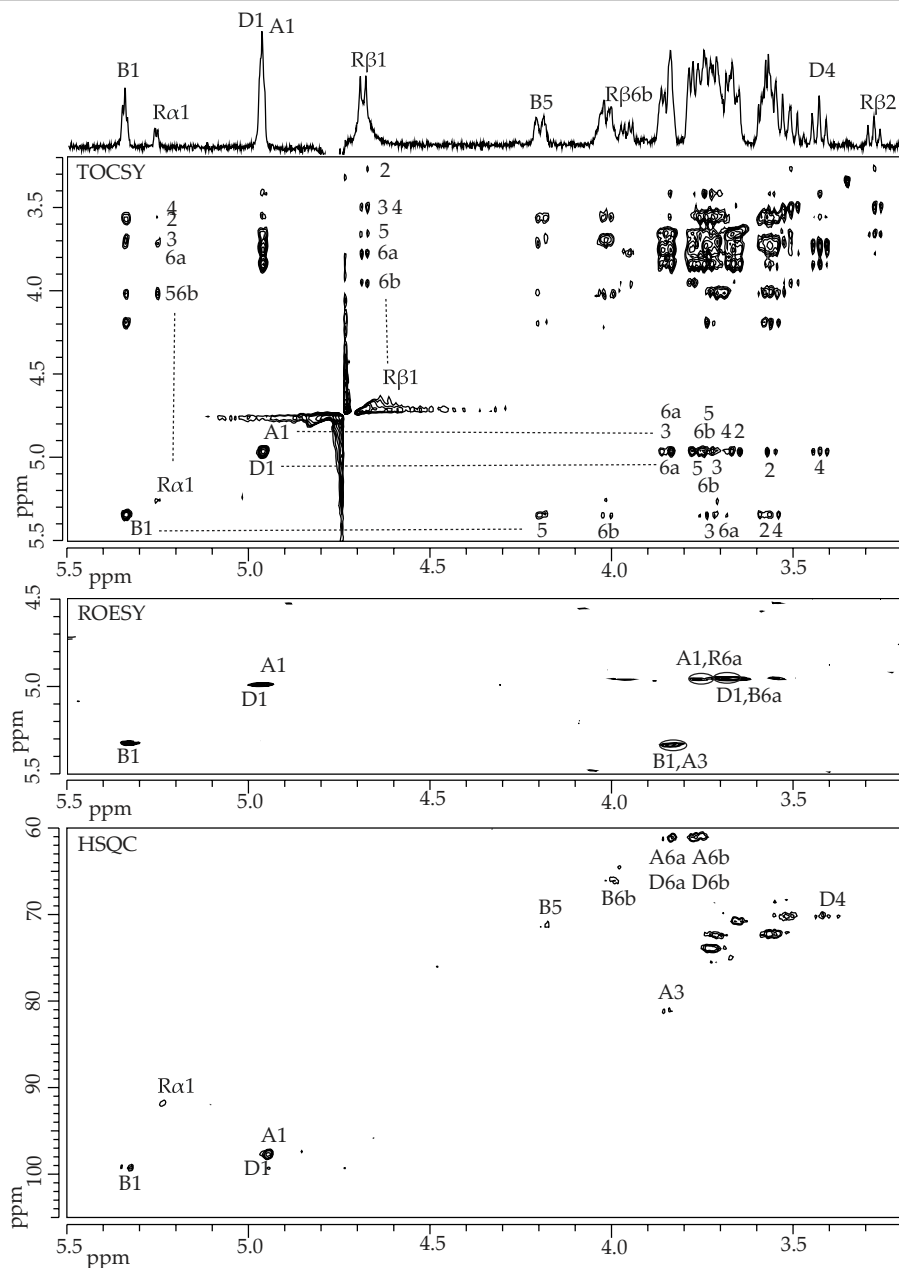


Figure 4. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum of fraction 3b, recorded at 300K in D_2O . Anomeric protons in the TOCSY spectrum (R α 1, etc.) have been indicated on the diagonal; numbers in the horizontal and vertical tracks belong to the cross-peaks of the scalar-coupling network of the residues indicated. In the ROESY spectrum inter-residual couplings (A1,R6a means a cross-peak between A H-1 and R H-6a, etc.) have been indicated with circles. In the ^{13}C - ^1H HSQC spectrum A1 denotes the cross-peak between H-1 and C-1 of residue A, etc.

The established structure of the tetrasaccharide was verified by 2D ^{13}C - ^1H HSQC and ROESY measurements. Interpretation of the HSQC spectrum (Figure 4) yielded $\delta_{\text{C-6}}$ values of 66.5 ppm for both residue **R** and **B**, and 61.2 ppm for residues **A** and **D**, indicating 6-substituted **R** and **B** units (Chapter 2).¹¹ The $\delta_{\text{C-3}}$ value of 81.2 ppm of residue **A** indicated a 3-substituted **A** unit (Chapter 2).¹¹ In the ROESY spectrum (Figure 4) inter-residual cross-peaks were observed between **A** H-1 and **R** H-6a, between **B** H-1 and **A** H-3, and between **D** H-1 and **B** H-6a.

Fraction 4 - MALDI-TOF-MS analysis of fraction **4** showed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 851, corresponding with Hex₅. HPAEC analysis on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH) showed only one peak (Figure 2C). The 1D ^1H NMR spectrum (Figure 5) revealed six anomeric signals at δ 5.334 (**B** H-1, $^3J_{1,2}$ 3.6 Hz), 5.241 (**R α** H-1, $^3J_{1,2}$ 3.6 Hz), 4.982 (**A** H-1, $^3J_{1,2}$ 3.6 Hz), 4.982 (**C** H-1, $^3J_{1,2}$ 3.6 Hz), 4.963 (**D** H-1, $^3J_{1,2}$ 3.6 Hz), and 4.670 (**R β** H-1, $^3J_{1,2}$ 8.0 Hz).

Starting from the anomeric signals in the 2D ^1H - ^1H TOCSY spectrum (Figure 5/180 ms), the chemical shifts of all non-anomeric signals could be determined (Table 1). The H-1 α and H-1 β chemical shifts of the reducing residue **R** correspond to the values established for a -(1 \rightarrow 6)-D-Glc p unit.¹¹ The 6-substitution of residue **R** is also reflected in the chemical shift of H-5 at δ 4.00 and 3.63 for the α - and β -configuration, respectively.¹¹ Note also the positions of H-6a and H-6b, corresponding to the occurrence of a -(1 \rightarrow 6)-D-Glc p unit. Although the anomeric signals of **A** and **C** overlap, the H-2, H-3, and H-4 signals could be assigned by evaluating the built-up series in TOCSY experiments with incremental mixing-times (data not shown). The chemical shifts of the set of **A** H-2, H-3, and H-4 are δ 3.65, 3.84, and 3.67, respectively, corresponding with the values of a -(1 \rightarrow 3)- α -D-Glc p (-) unit.¹¹ The anomeric proton of residue **A** has a chemical shift in the range established for (-) α -D-Glc p -(1 \rightarrow 6)- units in the structural-reporter-group concept (library data: δ 4.94 - 4.96¹¹), indicating the occurrence of a -(1 \rightarrow 3)- α -D-Glc p -(1 \rightarrow 6)- unit. The chemical shifts of the set of **C** H-2, H-3, and H-4 at δ 3.58, 3.76, and 3.51, combined with the H-1 signal at δ 4.982 (α 1-6), indicate the occurrence of a -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- unit (Chapter 2).¹¹ This is further confirmed by the **C** H-5 chemical shift at δ 3.93, which corresponds to the $\delta_{\text{H-5}}$ of a 6-substituted residue.

The chemical shift of residue **B** H-5 at δ 4.20 was shown to indicate the occurrence of a -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 3)- unit (see compound **3b**). Residue **D** has a

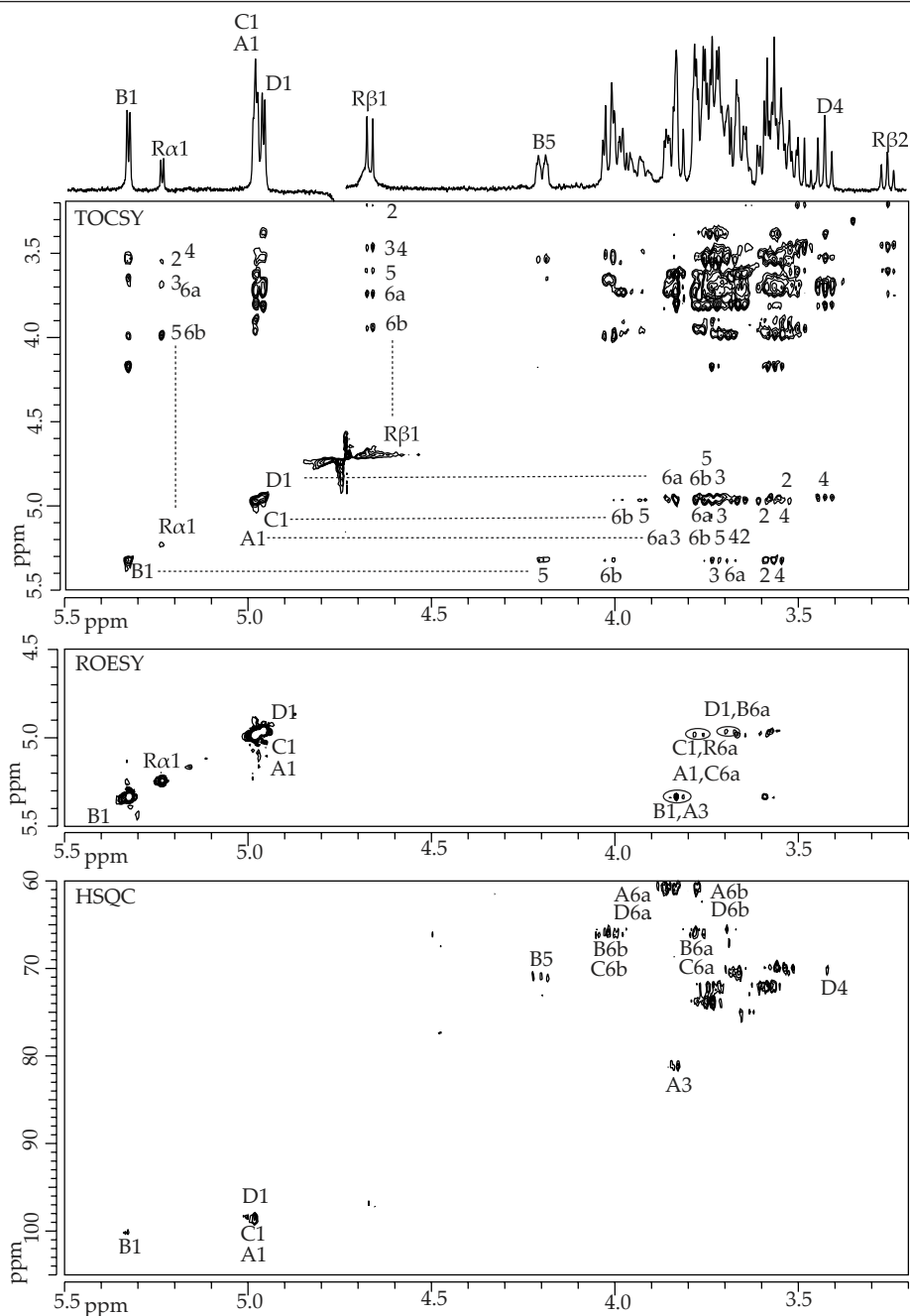


Figure 5. 500-MHz 1D ¹H NMR spectrum, 2D ¹H-¹H TOCSY spectrum (mixing time 180 ms), 2D ¹H-¹H ROESY spectrum (mixing time 300 ms), and 2D ¹³C-¹H HSQC spectrum of fraction 4, recorded at 300K in D₂O. For an explanation of the coding systems, see Figure 4.

well-separated anomeric peak at δ 4.963, reflecting an $(-)\alpha$ -D-Glcp-(1 \rightarrow 6)- unit. Guided by the structural-reporter-group concept established,¹¹ **D** H-4 at δ 3.42 indicates a terminal unit. The other chemical shift values of residue **D** also correspond to those expected for a terminal α -D-Glcp-(1 \rightarrow 6)- unit.¹¹

Since residue **A** is a $-(1\rightarrow3)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow6)\text{-}$ unit and residue **B** was shown to be a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}$ unit, it can be concluded that compound **4** has a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow6)\text{-}$ element, i.e. **X1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6Y**. The position of residues **R** and **D** in the structure as reducing and non-reducing termini, respectively, are clear, however, the position of residue **C** cannot be determined from the chemical shift pattern. In order to assign the position of residue **C** (**D1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6C1** \rightarrow **6R** or **D1** \rightarrow **6C1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6R**) the inter-residual cross-peaks in the ROESY spectrum (Figure 5) were evaluated. Taking into account the conclusions so far, the shared H-1 track of **A** and **C** showed inter-residual cross-peaks between H-1 and H-6a at δ 3.76-3.77, corresponding to **R** and **C** H-6a. Since residue **C** cannot substitute residue **C**, it can only be residue **A** H-1 interacting with **C** H-6a, and **C** H-1 with **R** H-6a. This finding is confirmed by the presence of the **D** H-1, **B** H-6a (δ 3.71) cross-peak (absence of **D** H-1, **C** H-6a cross-peak), leading unequivocally to **D1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6C1** \rightarrow **6R**, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

The substitution pattern of the established structure of compound **4** was verified by 2D ¹³C-¹H HSQC spectroscopy. In the HSQC spectrum (Figure 5) the δ_{C_6} value of 66.8 ppm reflects the 6-substitution of residues **B**, **C**, and **R**, in contrast to the δ_{C_6} value of 61.3 ppm for residues **A** and **D**.¹¹ The 3-substitution of residue **A** is evident from the δ_{C_3} value of 81.3 ppm.

Fraction 5 - Analysis by MALDI-TOF-MS revealed two $[M+Na]^+$ pseudomolecular ions with m/z 1013 and 1175, corresponding with Hex₆ and Hex₇, respectively. Fraction **5** was further separated on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH), yielding a major fraction **5a** and a minor fraction **5b** (Figure 2C).

The MALDI-TOF mass spectrum of fraction **5a** showed an $[M+Na]^+$ pseudomolecular ion at m/z 1013, in accordance with Hex₆. The ¹H NMR spectrum of fraction **5a** (Figure 3B) showed a multiple anomeric signal at δ ~4.96 (5 H), indicating five $(-)\alpha$ -D-Glcp-(1 \rightarrow 6)- units.¹¹ The H-1 α and H-1 β values of the reducing residue **R** at δ 5.241 and 4.671, respectively, indicate the occurrence of a $-(1\rightarrow6)\text{-D-Glcp}$ residue. On guidance of the structural-reporter-group concept,¹¹ the signal at δ 3.42 (dd, 1 H) was identified as H-4 of a α -D-Glcp-(1 \rightarrow x)- residue, in this case an α -D-Glcp-(1 \rightarrow 6)- unit.

The combined data identified compound **5a** as isomaltohexaose, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

The MALDI-TOF-MS analysis of fraction **5b** revealed an $[M+Na]^+$ pseudomolecular ion at m/z 1175, corresponding with Hex₆. The sample contained insufficient material for full 2D NMR spectroscopic analysis. The 1D ¹H NMR spectrum (Figure 3C) showed anomeric peaks at δ 5.334 (residues **B**, 2 H; (-) α -D-Glcp-(1 \rightarrow 3)-), 5.248 (residue **R α**), \sim 4.96 (residues **A**, **C**, and **D**, 4 H; (-) α -D-Glcp-(1 \rightarrow 6)-), and 4.678 (residue **R β**). The H-1 α and H-1 β values of residue **R** correspond best with the occurrence of a -(1 \rightarrow 6)-D-Glcp residue, and resemble the values found for **3b**. Furthermore, **R β** H-2 at δ 3.273 corresponds with the value of a -(1 \rightarrow 6)-D-Glcp unit and not with that of a -(1 \rightarrow 3)-D-Glcp unit.¹¹ The signal at δ 3.42 (dd), shown to correspond to H-4 of a terminal unit, has a surface area of 1 H, indicating a linear structure for fraction **5b**. Furthermore, the peak at δ 4.20, outside the bulk-region, has a surface area corresponding to 2 H. This peak, with the characteristic shape of an H-5 signal, was established to indicate the presence of -(1 \rightarrow 6)- α -D-Glcp-**B**-(1 \rightarrow 3)- units (see **3b** and **4**). This indicates that both residues **B** are 6-substituted, i.e. **X1** \rightarrow **6B1** \rightarrow **3Y**. In view of the latter conclusion, the single non-reducing terminal unit **D** must have an α -D-Glcp-(1 \rightarrow 6)- character. The slight downfield shifts of **R α** H-1 and **R β** H-1 compared to all **R** H-1 signals of isomalto-oligosaccharides, suggest, like in **3b**, a **B1** \rightarrow **3A1** \rightarrow **6R** element. In case of the occurrence of a **B1** \rightarrow **3A1** \rightarrow **6R** element a splitting of **B** H-1, due to the influence of the α/β configuration of residue **R**, is expected. However, the low signal-to-noise ratio of the NMR spectrum as well as the overlapping of two **B** H-1 signals, may interfere with the observation of the expected splitting pattern.

These data lead to two possible structures for **5b**: **D1** \rightarrow **6C1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6R**, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp, or **D1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6C1** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6R**, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1). The available data do not make a distinction possible.

Fraction 6 - The MALDI-TOF-MS analysis of fraction **6** revealed an $[M+Na]^+$ pseudomolecular ion at m/z 1175, corresponding with Hex₆. HPAEC-PAD analysis on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH) showed one peak. Inspection of the 1D ¹H NMR spectrum (Figure 3D) showed a multiple anomeric

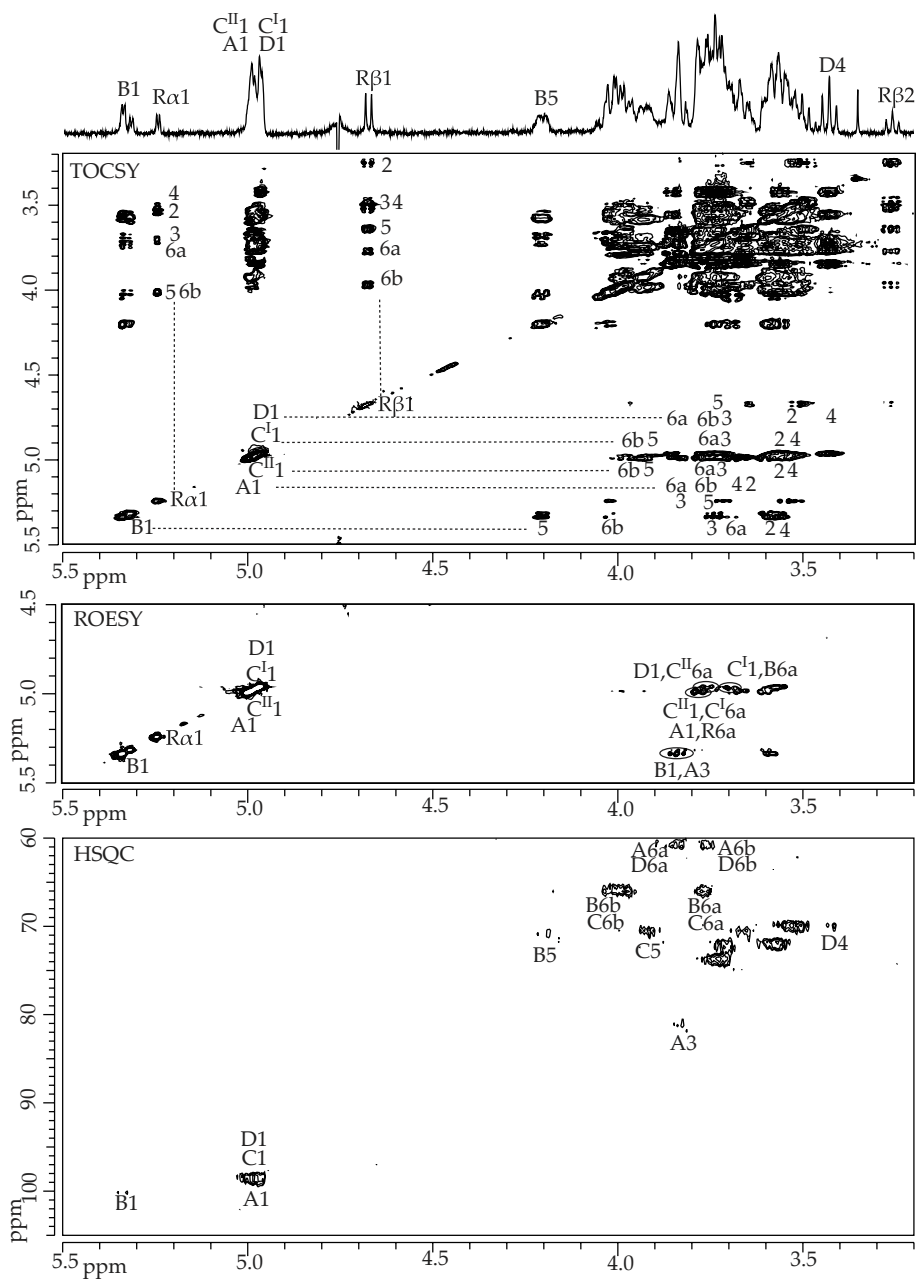


Figure 6. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms), and 2D ^{13}C - ^1H HSQC spectrum of fraction 7a, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 4.

signal at $\delta \sim 4.96$ (6 H; (-) α -D-Glc p -(1 \rightarrow 6)-), reducing residue anomeric signals at δ 5.240 (**R α** H-1) and 4.668 (**R β** H-1), reflecting the presence of a -(1 \rightarrow 6)-D-Glc p unit (Chapter 2),¹¹ and a signal at δ 3.42 (dd, 1 H), established as a structural reporter for terminal α -D-Glc p -(1 \rightarrow x)- units.¹¹ These data identify **6** as isomaltoheptaose, i.e. α -D-Glc p -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)-D-Glc p (Scheme 1).

Fraction 7 - The MALDI-TOF mass spectrum of fraction **7** showed [M+Na]⁺ pseudomolecular ions at m/z 851, 1013, and 1175, corresponding with Hex₅, Hex₆, and Hex₇, respectively. The Hex₆ pseudomolecular ion was more intense than those of Hex₅ and Hex₇. Fraction **7** was further separated on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH), rendering one major fraction **7a** (Figure 2C).

MALDI-TOF-MS analysis of fraction **7a** revealed an [M+Na]⁺ pseudomolecular ion at m/z 1013, corresponding with Hex₆. The ¹H NMR spectrum (Figure 6) showed seven anomeric signals at δ 5.343/5.319 (**B** H-1, ³ $J_{1,2}$ 3.7 Hz), 5.249 (**R α** H-1, ³ $J_{1,2}$ 3.7 Hz), 4.993 (**A** H-1, **C^{II}** H-1, ³ $J_{1,2}$ 3.7 Hz), 4.972 (**C^I** H-1, **D** H-1, ³ $J_{1,2}$ 3.7 Hz), and 4.679 (**R β** H-1, ³ $J_{1,2}$ 7.8 Hz). The splitting of the **B** H-1 signal is due to the influence of the α/β configuration of the reducing residue **R** (compare with **3b**). The H-1 α and H-1 β chemical shift values of **R** correspond to a -(1 \rightarrow 6)- α -D-Glc p unit, in an (-) α -D-Glc p -(1 \rightarrow 3)- α -D-Glc p -(1 \rightarrow 6)-D-Glc p element (compare with **3b** and **5b**). The peak at δ 3.43, identified as a structural-reporter-group signal for non-reducing terminal residues,¹¹ has a surface area corresponding to 1 H, suggesting a linear structure for **7a**.

Starting from the anomeric signals in the 2D ¹H-¹H TOCSY spectrum (Figure 6), all chemical shifts of the non-anomeric protons were determined (Table 1). The H-1 signals of **A** and **C^{II}**, and of **C^I** and **D** overlapped, making it difficult to distinguish between the two residues in each TOCSY track. However, observing the built-up of the scalar coupling network in experiments with increasing mixing times (data not shown), the chemical shifts of each residue could be assigned. On the **A** H-1 track the set of H-2, H-3, and H-4 is observed at δ 3.64, 3.85, and 3.66, respectively, corresponding with a -(1 \rightarrow 3)- α -D-Glc p (-) unit, in this case a -(1 \rightarrow 3)- α -D-Glc p -(1 \rightarrow 6)- unit.¹¹ Residue **C^{II}**, in the same track, showed the proton pattern of a -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- unit, with H-5, H-6a, and H-6b at δ 3.94, 3.77, and 3.98, respectively. The **D** H-1 track revealed the chemical shift pattern of a terminal residue, with the most noticeable structural-reporter-group signal H-4 at δ 3.43.¹¹ Residue **C^I**, in the same track, showed a proton pattern similar to **C^{II}**.

The chemical shift of **B** H-5 at δ 4.21, was established as a structural-reporter-group signal for a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow3)\text{-}$ unit (see **3b**). This is further supported by the chemical shift positions of H-6a (δ 3.70) and H-6b (δ 4.01).¹¹ The finding that **7a** is a linear structure (see above), having an internal **X1** \rightarrow **6B1** \rightarrow **3Y** element, leads to the conclusion that **7a** is built up of two isomalto-oligosaccharides, interconnected by an (α 1-3) linkage. In view of the **B** H-1 splitting, as well as the specific chemical shift values of **R** H-1 α and H-1 β , indicating a **B1** \rightarrow **3A1** \rightarrow **6R** element (see **3b** and **5b**), a **D1** \rightarrow **6C^{II}** \rightarrow **6C^I** \rightarrow **6B1** \rightarrow **3A1** \rightarrow **6R** sequence is indicated, i.e. $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow3)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-D-Glcp}$ (Scheme 1).

The established structure was verified by 2D ¹³C-¹H HSQC and ROESY measurements. Interpretation of the HSQC spectrum (Figure 6) yielded $\delta_{\text{C-6}}$ values of 66.8 ppm for residues **B**, **C^I**, **C^{II}**, and **R**, and 61.2 ppm for residues **A** and **D**, indicating 6-substituted **B**, **C^I**, **C^{II}**, and **R** units (Chapter 2).¹¹ The $\delta_{\text{C-3}}$ value of 81.3 ppm for residue **A** reflects the 3-substitution of this residue.¹¹ In the ROESY spectrum (Figure 6) inter-residual cross-peaks were observed. On the **B** H-1 track a cross-peak was found with **A** H-3 at δ 3.85. On the H-1 track of **D** and **C^I** cross-peaks were observed with $\delta_{\text{H-6a}}$ 3.70 and 3.77, corresponding to **B** H-6a and **C^{II}** or **R** H-6a, respectively. On the H-1 track of **A** and **C^{II}** cross-peaks were found with $\delta_{\text{H-6a}}$ 3.76, corresponding to **C^I** or **R** H-6a. These cross-peaks fit with the structure proposed above.

Fraction 8 - The MALDI-TOF mass spectrum of fraction **8** revealed $[\text{M}+\text{Na}]^+$ pseudomolecular ions at m/z 1013, 1175, and 1337, corresponding with Hex₆, Hex₇, and Hex₈, respectively, with Hex₇ as the major component. Fraction **8** was further separated on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH); major fraction **8a** was isolated (Figure 2C).

MALDI-TOF-MS analysis of fraction **8a** revealed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 1175, corresponding with Hex₇. The 1D ¹H NMR spectrum (Figure 7) showed eight anomeric signals at δ 5.343/5.320 (**B** H-1, ³J_{1,2} 4.0 Hz), 5.246 (**R α** H-1, ³J_{1,2} 4.0 Hz), 4.990 (**A** H-1 and **C^{III}** H-1, ³J_{1,2} 4.0 Hz), 4.968 (**C^I** H-1, **C^{II}** H-1, and **D** H-1, ³J_{1,2} 4.0 Hz), and 4.676 (**R β** H-1, ³J_{1,2} 7.5 Hz). The H-1 α and H-1 β chemical shifts of reducing residue **R** correspond to a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}$ unit, matching the values observed in fraction **3b**, **5b**, and **7a**, and suggests a **B1** \rightarrow **3A1** \rightarrow **6R** element. The splitting of the **B** H-1 signal is due to the influence of the α/β configuration of the reducing residue **R**, similar to the splitting observed in **3b** and **7a**. The peak at δ 3.43, identified as a structural-reporter-group signal for non-reducing terminal residues,¹¹ has a surface area corresponding to 1 H, suggesting a linear structure for **8a**.

Starting from the anomeric signals in the 2D ^1H - ^1H TOCSY spectrum (Figure 7), all chemical shifts of the non-anomeric protons could be determined (Table 1). Residues **C**^I, **C**^{II}, and **C**^{III} showed the same chemical shift pattern, with H-5, H-6a, and H-6b at δ 3.93, 3.76, and 3.98, respectively. These values were shown to indicate a $-(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-}$ residue. Residue **B** shows the chemical shift pattern of a $-(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)\text{-}$ unit, with H-5 at δ 4.20 as the most obvious structural-reporter-group signal (see **3b**, **4**, etc.). Residue **B** H-6a and H-6b at δ 3.70 and 4.01, respectively, confirm this observation. Residue **D** shows the chemical shift pattern of a non-reducing terminal residue, with H-4 at δ 3.43 being the most notable structural-reporter-group signal. Taking together all data, this results in a total sequence for compound **8a** of **D1** \rightarrow **6C**^{III}**1** \rightarrow **6C**^{II}**1** \rightarrow **6C**^I**1** \rightarrow **6B****1** \rightarrow **3A****1** \rightarrow **6R**, i.e. $\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 6)\text{-D-Glcp}$ (Scheme 1).

The established structure was verified by 2D ^{13}C - ^1H HSQC and ROESY measurements. In the HSQC spectrum (Figure 7) the δ_{C_6} values of residues **B**, **C**^{III,II,I} and **R** at 66.8 ppm reflect the 6-substitution of these units. The δ_{C_6} values of residues **A** and **D** were determined at 61.2 ppm. The δ_{C_3} value of 81.2 ppm of residue **A** indicated a 3-substituted residue **A** (Chapter 2).¹¹ In the ROESY spectrum (Figure 7), inter-residual cross-peaks were observed. Residue **B** H-1 showed one inter-residual cross-peak with **A** H-3. On the **D**, **C**^{III} H-1 track cross-peaks were observed with $\delta_{\text{H-6a}}$ 3.70 and 3.76-3.77. On the **A**, **C**^{III} H-1 track one inter-residual cross-peak was observed at $\delta_{\text{H-6a}}$ 3.77. These data fit with the proposed sequence for **8a**.

Smith degradation

In order to investigate the degree of polymerisation of (α 1-3) glycosidic bonds, **EPS180** was subjected to a Smith degradation, comprising of a periodate oxidation, followed by reduction with NaBH_4 , and mild acid hydrolysis with formic acid. The formed products were analysed by GLC-EI-MS (after trimethylsilylation) and HPAEC-PAD. In view of the linkage analysis (see section 2.1), $\alpha\text{-D-Glcp}\text{-}(1\rightarrow 1)\text{-Gro}$ and $[\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)]_n\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 1)\text{-Gro}$, but also, due to overhydrolysis, Gro, D-Glc, and $[\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)]_n\text{-D-Glc}$ can be expected.

GLC-EI-MS analysis (Figure 8A) of the trimethylsilylated residue showed three major products, namely, Gro, D-Glcp, and $\alpha\text{-D-Glcp}\text{-}(1\rightarrow 1)\text{-Gro}$. Typical fragments for the latter compound comprise m/z 451 (aA_1), 361 (aA_2), 271 (aA_3), 217, and 204 for the Glc part, and m/z 219 (bA_1) and 337 (abJ_1) for the Gro part.^{14,15,16,17}

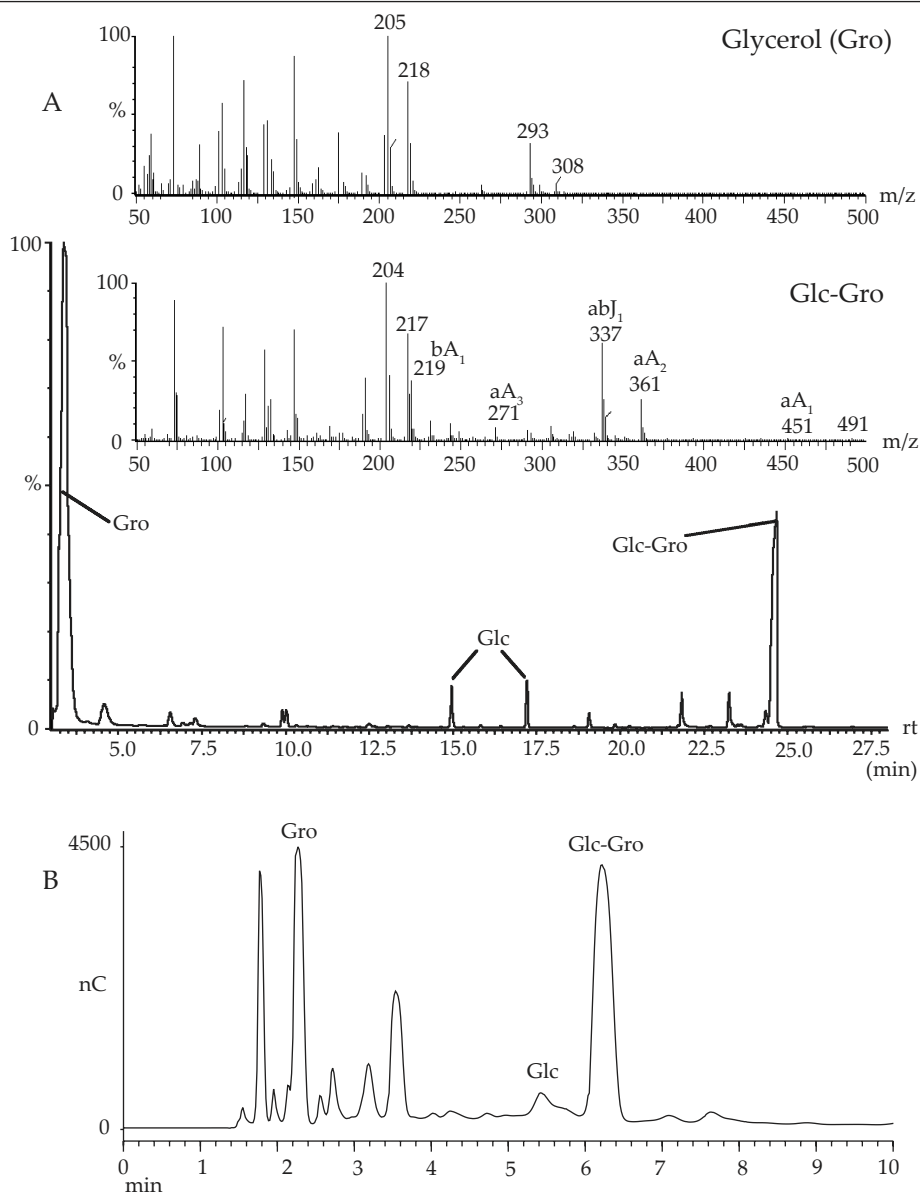


Figure 8. (A) GLC-EI-MS analysis and (B) HPAEC-PAD analysis of fragments obtained by Smith degradation of EPS180.

Table 1. ^1H chemical shifts of D-glucopyranose residues of oligosaccharide fragments obtained by partial acid hydrolysis of **EPS180** and of intact **EPS180**. Residue labels correspond to those used in Scheme 1.

Residue	3b	4	7a	8a	EPS180
R α -1	5.249	5.241	5.249	5.246	-
R α -2	3.55	3.55	3.56	3.56	-
R α -3	3.74	3.74	3.75	3.76	-
R α -4	3.50	3.50	3.51	3.52	-
R α -5	4.00	4.00	4.01	3.99	-
R α -6a	3.74	3.76	3.77	3.77	-
R α -6b	4.01	4.01	4.00	4.01	-
R β -1	4.680	4.670	4.679	4.676	-
R β -2	3.26	3.27	3.26	3.27	-
R β -3	3.48	3.48	3.48	3.49	-
R β -4	3.52	3.52	3.51	3.52	-
R β -5	3.63	3.63	3.64	3.63	-
R β -6a	3.78	3.75	3.75	3.76	-
R β -6b	3.98	3.97	3.98	3.97	-
A-1	4.958	4.982	4.993	4.990	4.99
A-2	3.64	3.65	3.64	3.64	3.65
A-3	3.85	3.84	3.85	3.85	3.87
A-4	3.67	3.67	3.66	3.67	3.66
A-5	3.72	3.73	3.74	3.72	3.70
A-6a	3.85	3.85	3.85	3.84	3.86
A-6b	3.76	3.75	3.75	3.76	3.75
B-1	5.345/5.336	5.334	5.343/5.319	5.343/5.320	5.32
B-2	3.58	3.58	3.59	3.58	3.59
B-3	3.75	3.75	3.74	3.75	3.75
B-4	3.50	3.50	3.50	3.51	3.52
B-5	4.20	4.20	4.21	4.20	4.17
B-6a	3.70	3.71	3.70	3.70	3.72
B-6b	4.01	4.00	4.01	4.01	4.00
C-1	-	4.982	4.972/4.993	4.990/4.968	4.96
C-2	-	3.58	3.58	3.59	3.57
C-3	-	3.76	3.77	3.76	3.76
C-4	-	3.51	3.50	3.51	3.51
C-5	-	3.93	3.94	3.93	3.89
C-6a	-	3.77	3.77	3.77	3.76
C-6b	-	3.99	3.98	3.98	3.95
D-1	4.965	4.963	4.972	4.968	4.96
D-2	4.56	3.55	3.55	3.54	3.55
D-3	3.75	3.75	3.75	3.76	3.76
D-4	3.42	3.42	3.43	3.43	3.41
D-5	3.77	3.77	3.78	3.78	3.78
D-6a	3.85	3.85	3.85	3.85	3.86
D-6b	3.76	3.76	3.75	3.76	3.77
E-1	-	-	-	-	4.99
E-2	-	-	-	-	3.65
E-3	-	-	-	-	3.86
E-4	-	-	-	-	3.77
E-5	-	-	-	-	3.90
E-6a	-	-	-	-	3.75
E-6b	-	-	-	-	3.95

HPAEC analysis on CarboPac PA-100 (Figure 8B) of the residue revealed two major peaks that could be related to Gro (R_t 2.3 min) and α -D-Glcp-(1 \rightarrow 1)-Gro (R_t 6.2 min), which was identified by its elution position, being slightly later than D-Glcp (R_t 5.5 min). Since α -D-Glcp-(1 \rightarrow 3)-D-Glcp under the same conditions has an R_t value $>$ 10 min, the presence of [α -D-Glcp-(1 \rightarrow 3)] $_n$ α -D-Glcp-(1 \rightarrow 1)-Gro with $n=1$ or higher could be excluded.

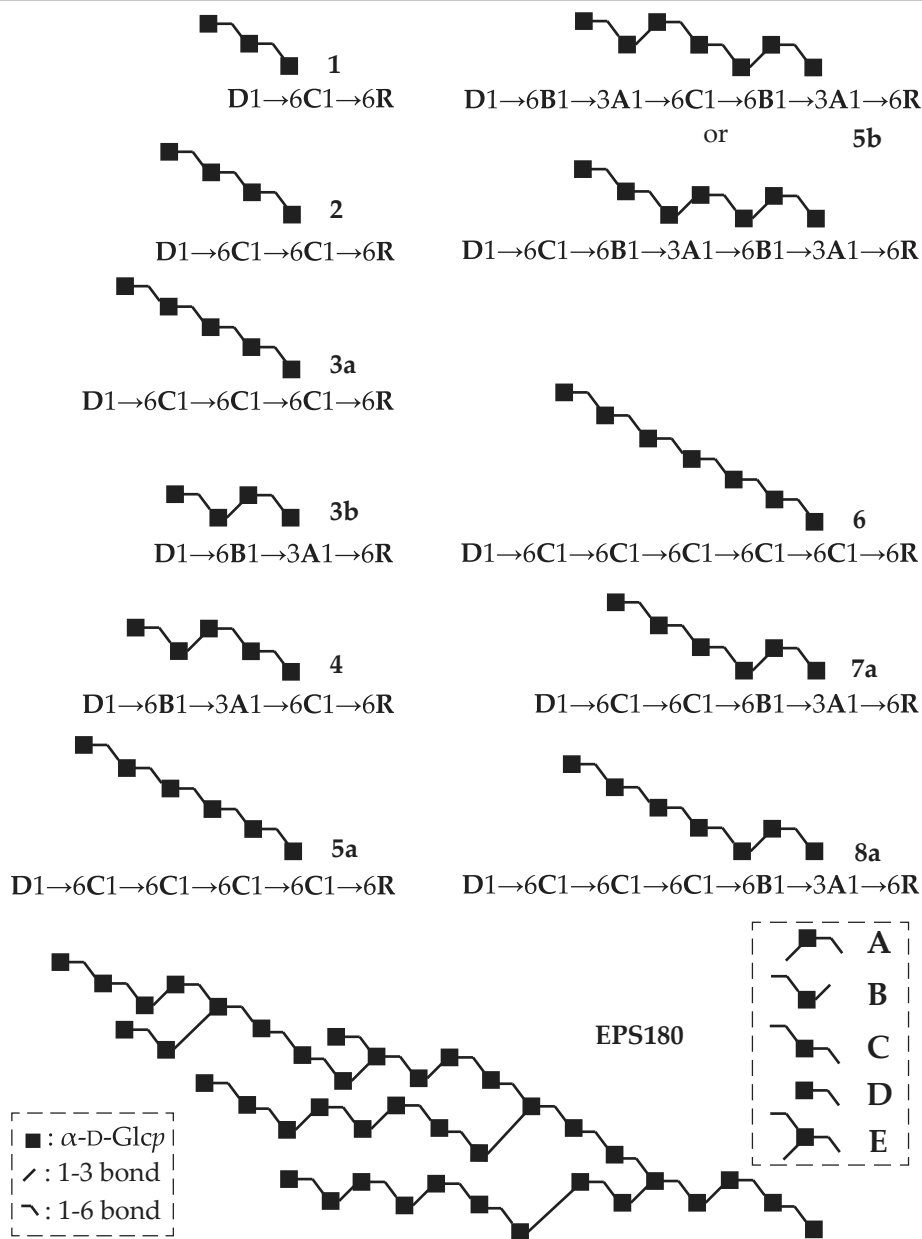
The absence of larger structures than glucosyl-glycerol indicates that the **EPS180** structure does not contain two or more consecutive (α 1-3) linkages, but is built up from isomalto-oligosaccharides, interconnected by single (α 1-3) linkages.

2D NMR spectroscopy of **EPS180**

For the unravelling of the 1D ^1H NMR spectrum of **EPS180** (Figure 1B), 2D TOCSY experiments with a short mixing time (60 ms) and with a long mixing time (120 ms), as well as data from the ^{13}C - ^1H HSQC measurements of **EPS180** (Figure 9, Table 1) were interpreted using the data gathered from the oligosaccharides obtained by partial acid hydrolysis of **EPS180** (section 2.2).

In the 60 ms TOCSY spectrum (330K, Figure 9), the H-1 track of residue **B** showed only H-2 and H-3 at δ 3.59 and 3.75, respectively. In the 120 ms TOCSY spectrum (330K, Figure 9) the complete scalar coupling network is detected with H-5 at δ 4.17, corresponding with the value found for -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units (see residue **B** in **3b**, **4**, **5b**, **7a**, and **8a**). As the surface area of the H-5 peak at δ 4.17 in the 1D ^1H NMR spectrum recorded at 330K (Figure 1B) is influenced by the suppression of the HOD peak, this signal cannot be used to quantify the occurrence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units in **EPS180**. Fortunately, the 1D ^1H NMR spectrum recorded at 300K (Figure 1A) can be used for this goal; here, the H-5 signal at δ 4.20 has a surface area matching that of the anomeric peak at δ 5.32, corresponding with (-) α -D-Glcp-(1 \rightarrow 3)- units. This finding leads to the conclusion that all (-) α -D-Glcp-(1 \rightarrow 3)- residues are 6-substituted, which is further supported by the absence of an H-4 signal at δ \sim 3.43 (structural-reporter-group signal for terminal residues) and the presence of an H-4 signal at δ 3.52 on the **B** H-1 track (-(1 \rightarrow 6)- α -D-Glcp(-)).¹¹ Further support for the conclusion that all (α 1-3)-linked residues occur in a -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- element (residue **B**), stems from the absence of cross-peaks in the δ 3.78 - 3.90 region, typical for H-3 of a 3-substituted residue (nigerotriose),¹¹ as well as H-6a of a non-6-substituted residue.¹¹

The **A/E** H-1 track in the 60 ms TOCSY spectrum showed H-2 and H-3 at δ 3.65 and 3.86-3.87, respectively, corresponding to a -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- unit (see



Scheme 1. Structures of oligosaccharide fragments, obtained by partial acid hydrolysis of **EPS180** and a visual representation that includes all structural elements identified for **EPS180**. Residue labels correspond with those used in the figures and are presented on the right.

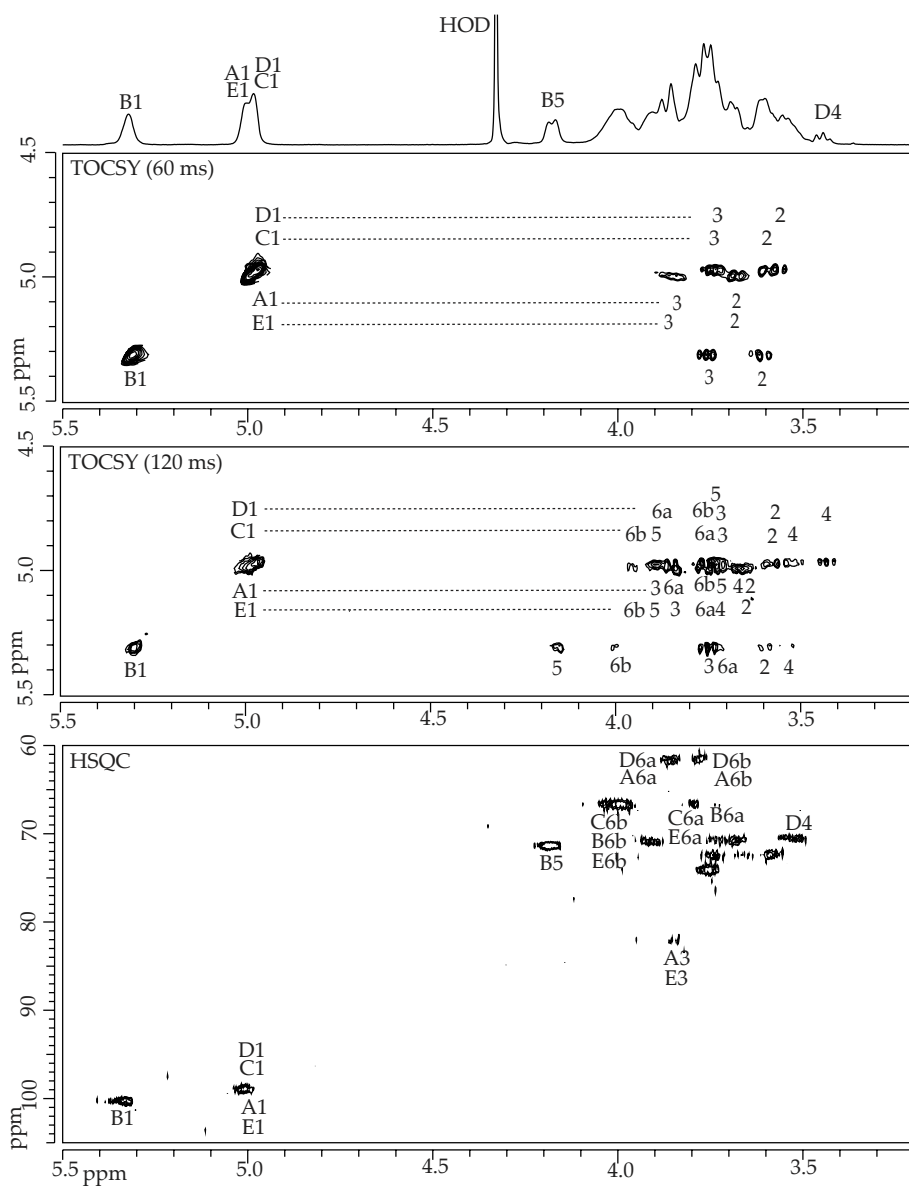


Figure 9. 500-MHz 1D ¹H NMR spectrum, 2D ¹H-¹H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms), and 2D ¹³C-¹H HSQC spectrum of EPS180 recorded at 330K in D₂O. Anomeric protons in the TOCSY spectrum (A1, etc.) have been indicated on the diagonal; numbers in the horizontal tracks belong to the cross-peaks of the scalar-coupling network of the residues indicated. In the ¹³C-¹H HSQC spectrum A1 denotes the cross-peak between H-1 and C-1 of residue A, etc.

residue **A** in **3b**, **4**, **7a**, and **8a**). However, the **A/E** H-1 track in the 120 ms TOCSY spectrum showed two different H-5 signals at δ 3.70 and 3.90, respectively. The H-5 resonance at δ 3.70 corresponds with the value expected for an $(-)\alpha$ -D-Glcp-(1 \rightarrow 6)-residue that is not 6-substituted (see residue **A** in **3b**, **4**, **7a**, and **8a**),¹¹ in this case a $(-)(1\rightarrow 3)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 6)\text{-}$ unit (residue **A**). The H-5 signal at δ 3.90 corresponds with a $(-)(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 6)\text{-}$ element (see residue **C** in **4**, **7a**, and **8a**, for example). Combined with the H-2 and H-3 δ values, that reflect a 3-substitution, this means that a branching $(-)(1\rightarrow 3,6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 6)\text{-}$ unit is indicated (residue **E**).

The H-1 track labelled **C/D** in the 60 ms TOCSY spectrum showed two overlapping peaks for H-2 at δ 3.57 and 3.55, corresponding with residue types **C** and **D**, respectively (compare with **3b**, **4**, **7a**, and **8a**). The **C/D** H-1 track revealed one H-3 signal at δ 3.76. These δ values indicate that there is no 3-substitution on these residues (library data: $\delta_{\text{H-2}}$ 3.63-3.68 and $\delta_{\text{H-3}}$ 3.84-3.90 in nigerose and nigerotriose).¹¹ In the 120 ms TOCSY spectrum the **C/D** H-1 track showed two distinctly different H-4 resonances at δ 3.41 and 3.51. The H-4 signal at δ 3.41 reflects a terminal residue (**D**) and that at δ 3.51 a non-terminal, 6-substituted residue.¹¹ These chemical shifts allow the distinction between $(-)(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 6)\text{-}$ (residue **C**) and $\alpha\text{-D-Glcp-(1}\rightarrow 6)\text{-}$ (residue **D**). In the 120 ms TOCSY spectrum, the **C/D** H-1 track showed also peaks for H-5 and H-6b that are shifted downfield, in favour of the 6-substitution of residue **C**.

The substitution patterns established from the TOCSY spectra were verified using $^{13}\text{C-}^1\text{H}$ HSQC spectroscopy (330K, Figure 9). Residues **B**, **C**, and **E** gave rise to a $\delta_{\text{C-6}}$ value of 65.9 ppm, reflecting the 6-substitution of these residues.¹¹ The $\delta_{\text{C-6}}$ value of 61.2 ppm confirmed the absence of a 6-substitution in residues **A** and **D**.¹¹ Finally, the $\delta_{\text{C-3}}$ value of 81.3 ppm of residues **A** and **E** indicated a 3-substitution.¹¹

Conclusions

In order to formulate a picture of the native **EPS180** α -D-glucan, ^1H NMR data of **EPS180** were combined with methylation analysis data (section 2.1) and periodate oxidation data (section 2.3) of **EPS180**, and ^1H NMR data of the **EPS180** fragments described in section 2.2. The visual representation is depicted in Scheme 1. It can be concluded that the $(1\rightarrow 3, 1\rightarrow 6)\text{-}\alpha\text{-D-glucan}$ of *Lb. reuteri* strain 180 contains only $\alpha\text{-D-Glcp-(1}\rightarrow 6)\text{-}$ units in terminal position. All $(-)\alpha\text{-D-Glcp-(1}\rightarrow 3)\text{-}$ units were shown to be 6-substituted, and the polysaccharide is built-up from different lengths of isomalto-

oligosaccharides, interconnected by single (α 1-3) bridges. The methylation analysis data define the statistical distribution of the five residue types that are the building blocks of **EPS180**. The partial acid hydrolysis fragments all fit into the picture, and reflect the structural possibilities within the determined boundaries.

Materials and Methods

Materials

EPS180 was a gift from TNO Quality of Life, Zeist, The Netherlands. D₂O (99.9 atom%) was purchased from Cambridge isotope laboratories, Inc., Andover, MA.

Monosaccharide analysis

A polysaccharide sample was subjected to methanolysis (1.0 M methanolic HCl, 24 h, 85 °C), followed by re-*N*-acetylation and trimethylsilylation (1:1:5 hexamethyldisilazane-trimethylchlorosilane-pyridine; 30 min, room temperature). The mixture of trimethylsilylated methyl glycosides was analysed by GLC on an EC-1 column (30 m x 0.32 mm, Alltech Associates Inc., Illinois, USA) using a Chrompack CP 9002 gas chromatograph (temperature gradient, 140-240 °C at 4 °C/min). The identification of the monosaccharide derivatives was confirmed by GLC-EI-MS.¹⁵

Linkage analysis

A polysaccharide sample was permethylated using CH₃I and solid NaOH in DMSO, as described earlier.¹⁸ After hydrolysis with 2 M TFA (2 h, 120 °C), the partially methylated monosaccharides were reduced with NaBD₄ (2 h, room temperature). Conventional work-up, involving neutralisation with HOAc and removal of boric acid by co-evaporation with MeOH, followed by acetylation with 1:1 acetic anhydride:pyridine (3 h, 120 °C), yielded a mixture of partially methylated alditol acetates, which was analysed by GLC-EI-MS.^{15,19}

Partial acid hydrolysis

In pilot experiments, 5-mg aliquots of **EPS180** were hydrolysed for 30, 60, 90, and 120 min (0.5 M TFA, 90 °C). After concentration under reduced pressure, the residues were analysed by MALDI-TOF-MS and 1D ¹H NMR spectroscopy. Based on the initial results, **EPS180** (500 mg) was hydrolysed in 25 mL 0.5 M TFA for 3 h at 90 °C. The solution was concentrated under reduced pressure at a rotary evaporator, and the

residue was separated on a Bio-Gel P-4 column (400 x 12 mm, BioRad), eluted with 25 mM NH_4HCO_3 ; 0.9-mL fractions were collected at a flow rate of 13.5 mL/h. Fractions were tested for the presence of carbohydrate by a TLC spot-test with orcinol/ H_2SO_4 staining. Carbohydrate-containing fractions were analysed by MALDI-TOF-MS.

Smith degradation

EPS180 (50 mg) was incubated with 10 mL 50 mM sodium periodate in 0.1 M NaOAc (pH 4.3) for 96 h at 4 °C in the dark. Then, the excess of periodate was destroyed by addition of 0.5 mL ethylene glycol. The oxidised polysaccharide solution was dialysed against tap water (24 h, room temperature), treated with excess NaBH_4 (18 h, room temperature), and subsequently neutralised with 4 M HOAc.²⁰ After co-evaporation of boric acid with MeOH, the residue was hydrolysed with 90% HCOOH (30 min, 90 °C). Finally, the solution was concentrated under a stream of N_2 , and the products were analysed by GLC-EI-MS and HPAEC-PAD.

HPAEC-PAD

High-pH anion-exchange chromatography was performed on a Dionex DX500 workstation, equipped with an ED40 pulsed amperometric detection (PAD) system. A triple-pulse amperometric waveform (E_1 0.1 V, E_2 0.7 V, E_3 -0.1 V) was used for detection with the gold electrode.²¹ Analytical separations were performed on a CarboPac PA-100 column (250 x 4 mm, Dionex), using a linear gradient of 0 - 300 mM NaOAc in 100 mM NaOH (1 mL/min). Samples were fractionated on a CarboPac PA-1 column (250 x 9 mm, Dionex), using a linear gradient of 0 - 300 mM NaOAc in 100 mM NaOH (4 mL/min) or isocratic conditions of 100 mM NaOAc in 100 mM NaOH (4 mL/min). Collected fractions were immediately neutralised with 4 M HOAc, desalted on CarboGraph SPE columns (150 mg graphitised carbon, Alltech) using 1:3 acetonitrile: H_2O as eluent, and lyophilised.

Mass spectrometry

GLC-EI-MS was performed on a Fisons Instruments GC 8060/ MD 800 system (Interscience BV; Breda, The Netherlands) equipped with an AT-1 column (30 m x 0.25 mm, Alltech), using a temperature gradient of 140-240 °C at 4 °C/min.¹⁵

Matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) was carried out on a Voyager-DE Pro (Applied Biosystems; Nieuwerkerk aan de IJssel, The Netherlands) instrument in the reflector mode at an

accelerating voltage of 24 kV, using an extraction delay of 90 ns, in a resolution of 5000-9000 FWHM. Samples (1 μ L) were mixed in a 1:1 ratio with a mixture of 7.5 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 1:1 acetonitrile:H₂O, and spectra were recorded in the positive-ion mode.

NMR spectroscopy

¹H NMR spectra, including ¹H-¹H and ¹³C-¹H correlation spectra, were recorded on a Bruker DRX500 spectrometer (Bijvoet Center, Department of NMR spectroscopy, Utrecht University). Oligosaccharide ¹H NMR spectra were recorded at a probe temperature of 300K, and intact **EPS180** ¹H NMR spectra were recorded at both 300K and 330K. Samples were exchanged once with 99.9 atom% D₂O, lyophilised, and dissolved in 650 μ L D₂O. ¹H chemical shifts (δ) are expressed in ppm by reference to internal acetone (δ 2.225), and ¹³C chemical shifts (δ) are expressed in ppm by reference to the methyl-carbon of internal acetone (δ 31.08). 1D ¹H NMR spectra were recorded with a spectral width of 5000 Hz in 16k complex data sets and zero filled to 32k. A WFT pulse sequence was applied to suppress the HOD signal.²² When necessary, a fifth order polynomial baseline correction was applied. 2D TOCSY spectra were recorded using MLEV17 mixing sequences with spin-lock times of 10, 30, 60, 120, and 180 ms. The spin-lock field strength corresponded to a 90° pulse width of about 28 μ s at 13 dB. The spectral width in 2D TOCSY experiments was 4006 Hz at 500 MHz in each dimension. 400-1024 spectra of 2k data points with 8-32 scans per t₁ increment were recorded. 2D rotating-frame nuclear Overhauser enhancement spectra (ROESY) were recorded with 300 ms mixing time. The spectral width was 4006 Hz at 500 MHz in each dimension. Suppression of the HOD signal was performed by 1 s pre-saturation during the relaxation delay. Between 400 and 1024 data sets of 2k data points were recorded with 8-16 scans per t₁ increment. 2D ¹³C-¹H HSQC spectroscopy was carried out at a ¹H frequency of 500.0821 MHz and a ¹³C frequency of 125.7552 MHz. Spectra were recorded with a spectral width of 4006 Hz for t₂ and 10 kHz for t₁. The HOD signal was pre-saturated for 1 s, and ¹²C-bound protons were suppressed using a TANGO pulse sequence. During acquisition of the ¹H FID, a ¹³C decoupling pulse was applied. 128-256 experiments of 2k data points were recorded with 128 scans per t₁ increment. 2D NMR spectroscopic data were analysed by applying a sinus multiplication window and zero filling to spectra of 4k by 1k dimensions. In the case of ¹³C-¹H HSQC data, the spectra were zero filled to 4k by 512 data points. A Fourier transform was applied, and where necessary, a fifth to fifteenth order polynomial baseline function was applied. All

NMR data were processed using in-house developed software (J.A. van Kuik, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University).

References

1. Costerton, J. W.; Cheng, K.-J.; Geesey, G. G.; Ladd, T. I.; Nickel, J. C.; Dasgupta, M.; Marric, T. J. *Annu. Rev. Microbiol.*, **1987**, *41*, 435-464.
2. Ceri, H.; McArthur, H. A. I.; Whitfield, C. *Infect. Immun.*, **1986**, *51*, 1-5.
3. Cerning, J. *FEMS Microbiol. Rev.*, **1990**, *87*, 113-130.
4. Sandford, P. A.; Baird, J. In *The Polysaccharides. Vol. 2*; Aspinall, G. O., Ed.; Academic Press: New York, **1983**; pp 411-490.
5. De Vuyst, L.; Degeest, B. *FEMS Microbiol. Rev.*, **1999**, *23*, 153-177.
6. Monchois, V.; Willemot, R.-M.; Monsan, P. *FEMS Microbiol. Rev.*, **1999**, *23*, 131-151.
7. Funane, K.; Ishii, T.; Matsushita, M.; Hori, K.; Mizuno, K.; Takahara, H.; Kitamura, Y.; Kobayashi, M. *Carbohydr. Res.*, **2001**, *334*, 19-25.
8. Argüello-Morales, M. A.; Remaud-Simeon, M.; Pizzut, S.; Sarçabal, P.; Willemot, R.-M.; Monsan, P. *FEMS Microbiol. Lett.*, **2000**, *182*, 81-85.
9. Kralj, S.; van Geel-Schutten, G. H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Microbiology*, **2004**, *150*, 2099-2112.
10. Gibson, G. R.; Roberfroid, M. B. *J. Nutr.*, **1995**, *125*, 1401-1412.
11. Van Leeuwen, S. S.; Leeftang, B. R.; Gerwig, G. J.; Kamerling, J. P. (Chapter 2) (to be published)
12. Dowd, M. K.; Zeng, J.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *230*, 223-244.
13. Stenutz, R.; Jansson, P.-E.; Widmalm, G. *Carbohydr. Res.*, **1998**, *306*, 11-17.
14. Kovacic, V.; Bauer, S.; Rosik, J.; Kovac, P. *Carbohydr. Res.*, **1968**, *8*, 282-290.
15. Kamerling, J. P.; Vliegthart, J. F. G. In *Clinical Biochemistry – Principles, Methods, Applications. Vol. 1, Mass Spectrometry*; Lawson, A. M., Ed.; Walter de Gruyter: Berlin, **1989**; pp 176-263.
16. Van der Kaaden, A.; van Doorn-van Wakeren, J. I. M.; Kamerling, J. P.; Vliegthart, J. F. G.; Tiesjema, R. H. *Eur. J. Biochem.*, **1984**, *141*, 513-519.
17. Veerkamp, J. H.; van Schaik, F. W. *Biochim. Biophys. Acta*, **1974**, *348*, 370-387.
18. Ciucanu, I.; Kerek, F. *Carbohydr. Res.* **1984**, *131*, 209-217.
19. Jansson, P.-E.; Kenne, L.; Liedgren, H.; Lindberg, B.; Lönngrén, J. *Chem. Commun. Univ. Stockholm*, **1976**, *8*, 1-74.
20. Hay, G. W.; Lewis, B. A.; Smith, F. *Methods Carbohydr. Chem.*, **1965**, *5*, 357-360.
21. Lee, Y. C. *Anal. Biochem.*, **1990**, *189*, 151-162.
22. Hård, K.; van Zadelhoff, G.; Moonen, P.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.*, **1992**, *209*, 895-915.

One sometimes finds what one is not looking for

-Alexander Fleming-

Chapter 4

Structural analysis of the α -D-glucan (EPS35-5) produced by the *Lactobacillus reuteri* strain 35-5 glucansucrase GTFA enzyme

Sander S. van Leeuwen,¹ Slavko Kralj,^{2,3} Ineke H. van Geel-Schutten,^{3,4}
Gerrit J. Gerwig,¹ Lubbert Dijkhuizen,^{2,3} and Johannes P. Kamerling¹

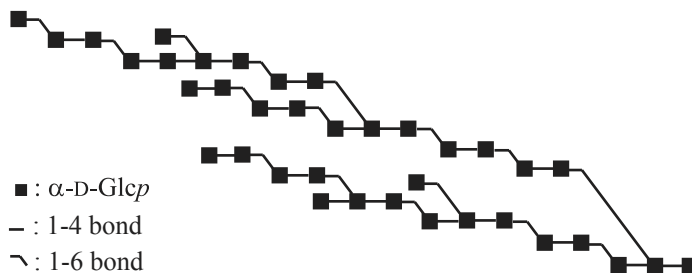
¹*Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, 3584 CH Utrecht, The Netherlands*

²*Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands*

³*Centre for Carbohydrate Bioprocessing, TNO - University of Groningen, 9750 AA Haren, The Netherlands*

⁴*Current address: Materials Innovation Centre B.V., 2628 BL Delft, The Netherlands*

Abstract - The neutral exopolysaccharide **EPS35-5** produced from sucrose by the glucansucrase GTFA enzyme from *Lactobacillus reuteri* 35-5 was found to be a (1→4, 1→6)- α -D-glucan, with no repeating units present. Based on linkage analysis and 1D/2D ^1H and ^{13}C NMR spectroscopy of intact **EPS35-5**, as well as MS and NMR analysis of oligosaccharides obtained by partial acid hydrolysis and enzymatic hydrolysis, using pullulanase M1 (*Klebsiella planticola*), of **EPS35-5**, a visual representation, that includes all identified structural elements, was formulated as follows.



Introduction

The use of lactic acid bacteria (LAB) in food production is commonplace, due to their status as generally recognised as safe (GRAS). One of the reasons for the use of LAB in food (especially dairy) is the generation of large exopolysaccharides (EPSs) in the food medium. The physico-chemical properties of the EPSs lie at the basis of their clearly defined functions in food as stabilisers, gelling agents, and thickeners.¹ Whereas *in vivo* the function of these EPSs is not completely clear, they seem to be involved in adhesion,² certain cellular processes,³ and protection against dehydration, phagocytosis or toxins.³

Until recently, EPSs from lactic acid bacteria that were analysed, were generally heteropolysaccharides with repeating-unit structures.⁴ Structures of LAB homopolysaccharides produced by specific glucansucrase enzymes have not been extensively analysed. Only initial structural studies have been performed on dextrans (polysaccharides containing mostly $(\alpha 1-6)$ linkages) with $(\alpha 1-2)$, $(\alpha 1-3)$, and $(\alpha 1-4)$ branches,^{5,6} or linear homopolysaccharides.⁷ Recently, a structural-reporter-group concept for α -glucans was established,⁸ and applied to the analysis of an α -glucan produced by the glucansucrase GTF180 enzyme, resulting in a detailed structural model.⁹

Previous studies of GTFA, a glucansucrase enzyme with large similarity to GTF180, revealed that it produces **EPS35-5**, an α -glucan with $(\alpha 1-4)$ and $(\alpha 1-6)$ glycosidic linkages.¹⁰ Here, we describe the structural analysis of this homopolysaccharide, using the structural-reporter groups established previously (Chapter 2),⁸ and enhanced by application on **EPS180** (Chapter 3).⁹ The various data were combined to formulate a picture, that includes all structural elements identified in **EPS35-5**.

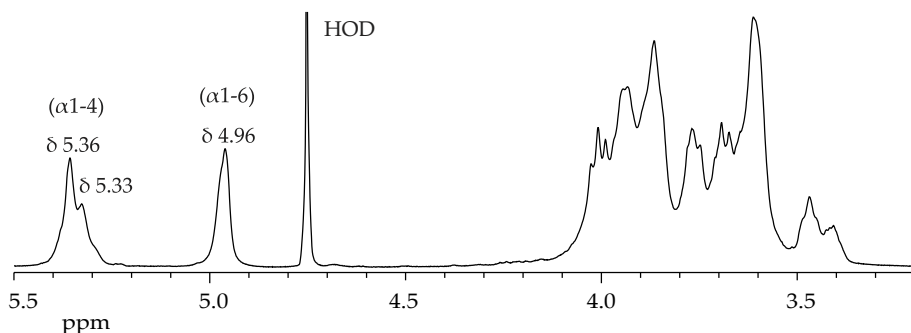


Figure 1. 500-MHz 1D ^1H NMR spectrum of intact **EPS35-5** recorded at 300K in D_2O .

Results

Composition of EPS35-5

Monosaccharide analysis of **EPS35-5** revealed the presence of glucose only, and a carbohydrate content of 100% (w/w). Methylation analysis of **EPS35-5** showed the presence of terminal, 4-substituted, 6-substituted, and 4,6-disubstituted glucopyranose in a molar percentage of 14, 46, 26, and 14%. 1D ^1H NMR spectroscopy (Figure 1) showed an α -anomeric configuration for all glucose residues. Of these residues, 58% is involved in (α 1-4) linkages ($\delta_{\text{H-1}} \sim 5.36$ and ~ 5.33) and 42% in (α 1-6) linkages ($\delta_{\text{H-1}} \sim 4.96$), which is in agreement with the linkage ratios found by methylation analysis (see also Ref. 10). The (α 1-4) anomeric ^1H signal is split into two, overlapping broad peaks, indicating that ($-\alpha$ -D-Glc p -(1 \rightarrow 4)-* residues in **EPS35-5** exist in at least two significantly different structural elements. The (α 1-6) anomeric ^1H signal is represented by a single peak, which may reflect a more uniform structural element for the ($-\alpha$ -D-Glc p -(1 \rightarrow 6)- residues.

Partial acid hydrolysis

After the evaluation of various partial acid hydrolysis conditions, making use of MALDI-TOF-MS and 1D ^1H NMR spectroscopy, a large batch of **EPS35-5** (500 mg) was subjected to partial acid hydrolysis with 0.5 M TFA (2 h, 90 $^\circ\text{C}$). The latter conditions rendered a pool of oligosaccharides in a broad distribution of degrees of polymerisation, as indicated by MALDI-TOF-MS (data not shown), 1D ^1H NMR

*For Glc residues at semi-defined places in the structure ($-\alpha$ -D-Glc p -(1 \rightarrow x)- or -(1 \rightarrow x)- α -D-Glc p (-)) is used. When the structural context of the residue is precisely known, this is indicated as follows: -(1 \rightarrow x)- α -D-Glc p -(1 \rightarrow y)- describing an x-substituted residue with an (1-y) linkage. In case of a non-reducing terminal residue α -D-Glc p -(1 \rightarrow x)- is used, a reducing terminal residue is indicated with -(1 \rightarrow x)-D-Glc p .

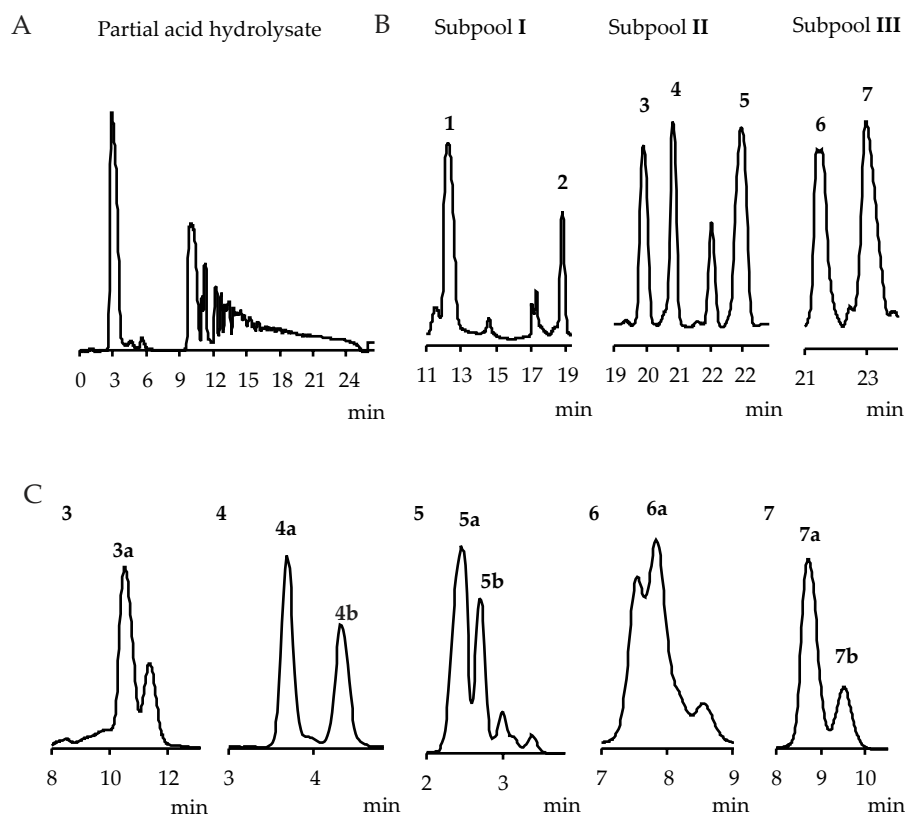


Figure 2. (A) HPAEC-PAD profile of EPS35-5 partial acid hydrolysate on CarboPac PA-100; (B) HPAEC-PAD profiles of Bio-Gel P-4 subpools I, II, and III on CarboPac PA-1, using a linear gradient; and (C) HPAEC-PAD profiles of HPAEC fractions 3-7 on CarboPac PA-1, using isocratic conditions. For experimental details, see Materials and Methods.

spectroscopy (data not shown), and HPAEC analysis on CarboPac PA-100 (Figure 2A). The glycan pool was pre-fractionated on Bio-Gel P-4, yielding four subpools, denoted I - IV. The fragment size distribution in each subpool was determined by MALDI-TOF-MS (data not shown). Subpool I consisted of fragments smaller than three Glc units. Subpool II contained fragments of three up to five Glc residues. Subpool III consisted of fragments of five up to seven Glc units. Finally, subpool IV consisted of fragments of seven residues and larger. The fragments in subpool IV were mainly too large for full structural analysis, and were not studied further. Subpools I - III were subfractionated by HPAEC on CarboPac PA-1, using a 0 - 300 mM NaOAc gradient in 100 mM NaOH (Figure 2B), and fractions 1 and 2 were isolated from subpool I, fractions 3 - 5 from subpool II, and fractions 6 and 7 from subpool III.

Fraction 1 - Profiling of fraction **1** on CarboPac PA-100 showed one peak. The 1D ^1H NMR spectrum of fraction **1** corresponded with that of isomaltose, $\alpha\text{-D-Glcp-(1}\rightarrow\text{6)-D-Glcp}$ (Scheme 1).⁸

Fraction 2 - Analysis on CarboPac PA-100 revealed that fraction **2** consisted of one compound. The 1D ^1H NMR spectrum corresponded with that of maltose, $\alpha\text{-D-Glcp-(1}\rightarrow\text{4)-D-Glcp}$ (Scheme 1).⁸

Fraction 3 - Analysis of fraction **3** on CarboPac PA-100 showed two component peaks. Fraction **3** was further separated on CarboPac PA-1, isocratically eluted with 100 mM NaOAc in 100 mM NaOH (Figure 2C), yielding one major fraction **3a**. The MALDI-TOF mass spectrum of fraction **3a** revealed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 527, corresponding with Hex_3 . The 1D ^1H NMR spectrum of **3a** (Figure 3) showed five anomeric signals at δ 5.412 (**D** H-1, $^3J_{1,2}$ 3.7 Hz), 5.240 (**R α** H-1, $^3J_{1,2}$ 3.8 Hz), 4.958 (**B β** H-1, $^3J_{1,2}$ 3.8 Hz), 4.950 (**B α** H-1, $^3J_{1,2}$ 3.8 Hz), and 4.669 (**R β** H-1, $^3J_{1,2}$ 8.1 Hz). Starting from the anomeric signals in the 2D ^1H - ^1H TOCSY spectrum (Figure 3/180 ms), all non-anomeric signals could be assigned (Table 1). The **R α** and **R β** H-1 values correspond with those found for the $\text{-(1}\rightarrow\text{6)-D-Glcp}$ unit of isomaltose, which was established as a structural-reporter-group signal for this type of unit.⁸ This was further confirmed by the **R** H-5 (δ 4.00 for **R α** and δ 3.64 for **R β**) and H-6b (δ 4.01 for **R α** and δ 3.93 for **R β**) resonances.⁸ Taking into account the ROESY inter-residual cross-peak between **D** H-1 and **B** H-4 (Figure 3), the set of H-3 and H-4 signals of residue **B** at δ 4.01 and 3.66, respectively, indicated a $\text{-(1}\rightarrow\text{4)-}\alpha\text{-D-Glcp-(1}\rightarrow\text{6)-}$ unit for residue **B**.⁸ The most notable signal on the **D** H-1 track of the TOCSY spectrum is the H-4 signal at δ 3.417, which was identified as a marker for a terminal $\alpha\text{-D-Glcp-(1}\rightarrow\text{x)-}$ unit;⁸ the H-1 value of **D** indicates the occurrence of an $\alpha\text{-D-Glcp-(1}\rightarrow\text{4)-}$ unit. These data lead to an isopanose structure **D1** \rightarrow **4B1** \rightarrow **6R**, i.e. $\alpha\text{-D-Glcp-(1}\rightarrow\text{4)-}\alpha\text{-D-Glcp-(1}\rightarrow\text{6)-D-Glcp}$ (Scheme 1).

Fraction 4 - The MALDI-TOF mass spectrum of fraction **4** showed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 527, corresponding with Hex_3 . Fraction **4** was further separated on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH) (Figure 2C), yielding major fraction **4a** and minor fraction **4b** (non-carbohydrate contaminant).

The 1D ^1H NMR spectrum of **4a** (Figure 4) showed four anomeric ^1H signals at δ 5.354/5.344 (**D** H-1, $^3J_{1,2}$ 3.9 Hz), 5.222 (**R α** H-1, $^3J_{1,2}$ 3.6 Hz), 4.985/4.976 (**E** H-1, $^3J_{1,2}$

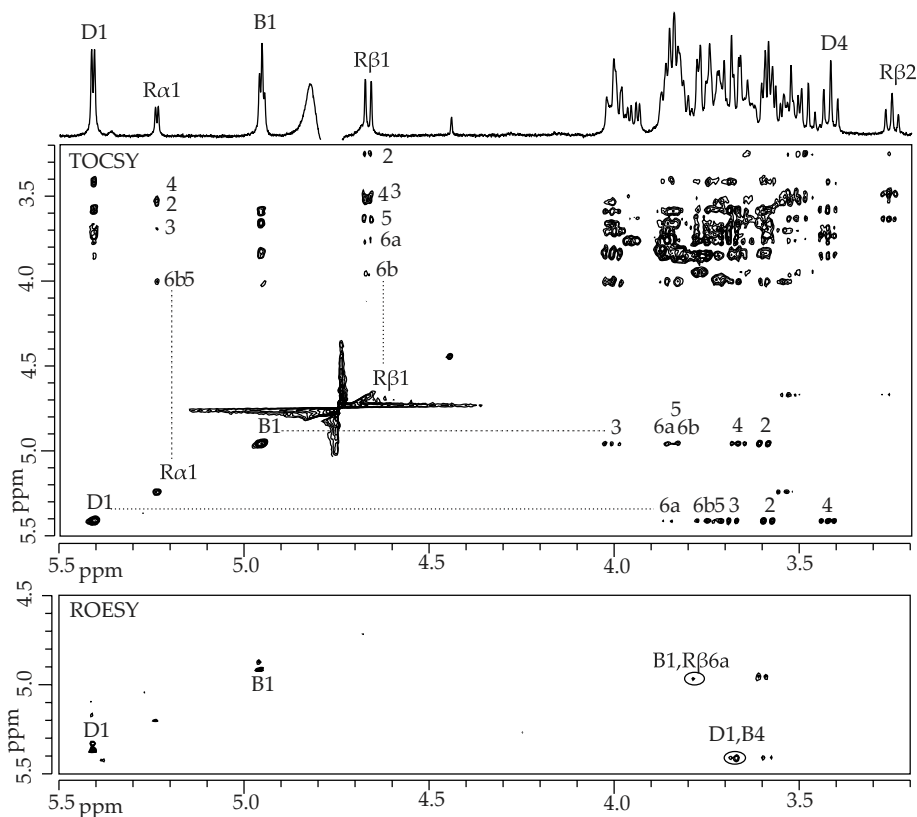


Figure 3. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), and 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms) of fraction **3a**, recorded at 300K in D_2O . Anomeric protons in the TOCSY spectrum (Ra1 , etc.) have been indicated on the diagonal; numbers in the horizontal and vertical tracks belong to the cross-peaks of the scalar-coupling network of the residues indicated. In the ROESY spectrum inter-residual couplings (D1, B4 means a cross-peak between **D** H-1 and **B** H-4, etc.) have been indicated with circles.

3.8 Hz), and 4.660 ($\text{R}\beta$ H-1, $^3J_{1,2}$ 8.1 Hz). Analysis of the 2D ^1H - ^1H TOCSY spectrum (Figure 4/180 ms) yielded all chemical shifts of the non-anomeric protons for residues **D**, **E**, and **R** (Table 1). In the absence of (α 1-2) linkages, both the **D** H-4 (δ 3.41) and **E** H-4 (δ 3.43) values are indicative of a terminal α -D-Glcp-(1 \rightarrow x)- residue. Taking into account the methylation analysis data of **EPS35-5** (section 2.1), **D** H-4 fits best with the occurrence of an α -D-Glcp-(1 \rightarrow 4)- unit and **E** H-4 with that of an α -D-Glcp-(1 \rightarrow 6)- unit.⁸ The chemical shifts of residue **R** H-3 (Ra δ 3.99; $\text{R}\beta$ 3.78) are indicative of a 4-substitution, confirming the occurrence of a **D1** \rightarrow **4R** structural element,⁸ which is further supported by the **D** H-1,**R** H-4 cross-peak in the 2D ROESY spectrum (Figure 4). In the case of a 4-substitution, the H-5 signal is to be expected at $\delta \sim 3.94$ for Ra and

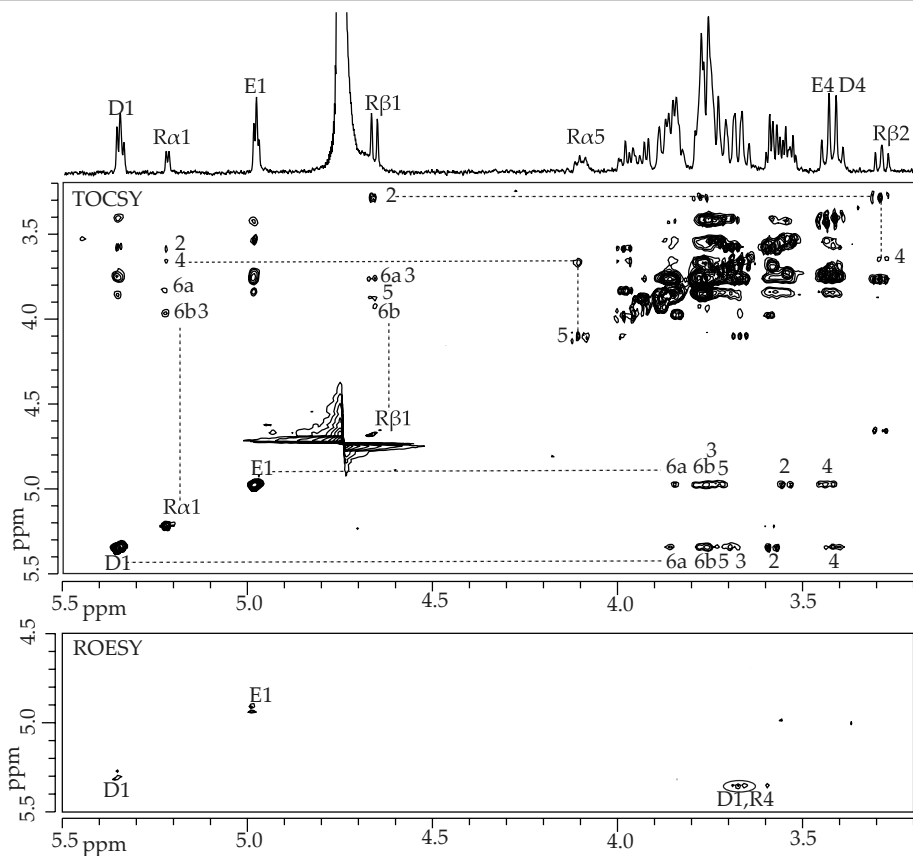


Figure 4. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), and 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms) of fraction **4a**, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 3.

$\delta \sim 3.63$ for **R β** , however, in **4a** the H-5 signal is found at δ 4.11 for **R α** and δ 3.89 for **R β** , indicating the additional effect of a 6-substitution.⁸ The occurrence of the **E1** \rightarrow **6R** structural element is further supported by the H-6b chemical shifts of **R**. These data lead to a **D1** \rightarrow **4[E1** \rightarrow **6]R** structure for compound **4a** (Scheme 1). The splitting of the anomeric signal of the α -D-Glcp-(1 \rightarrow 6)- unit is also observed in isomaltose.⁸ The anomeric splitting observed across the (α 1-4) bond in residue **D** has not been observed previously.

It should be noted that the 4,6-disubstitution of **R** yields a different set of **R α** H-1 and **R β** H-1 chemical shift values. Also the **R α** H-5 resonance at δ 4.11 can be considered as a new structural-reporter-group signal. These findings lead to a further fine-tuning of the structural-reporter-group concept, similar to the new structural reporters established for **EPS180**.⁹

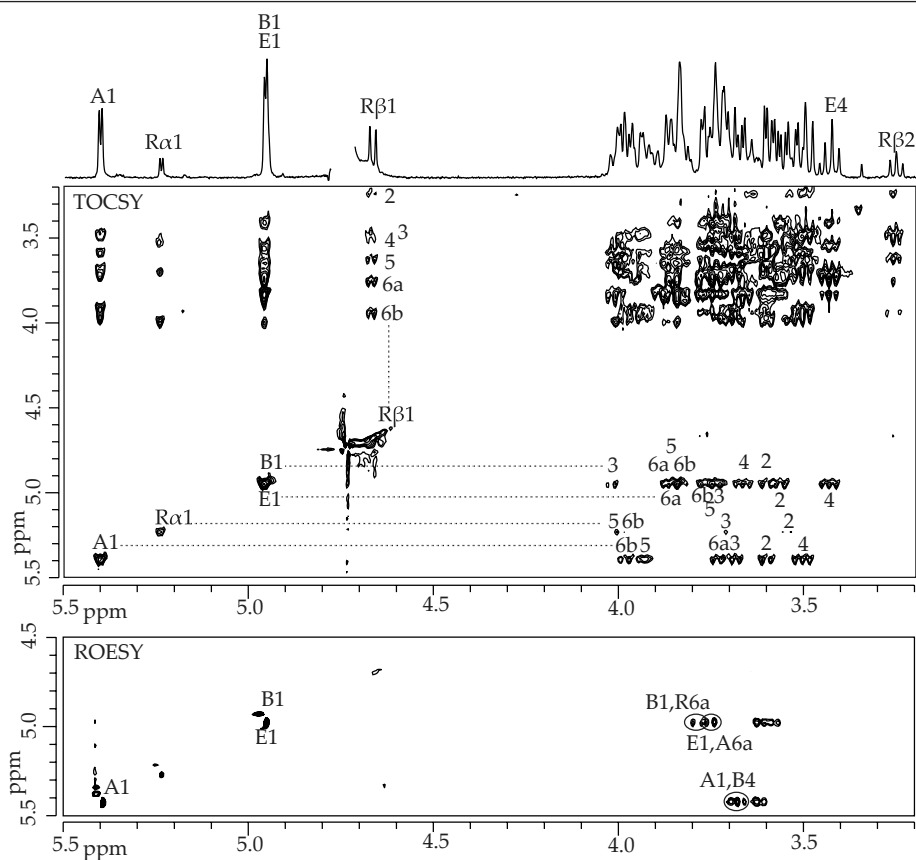


Figure 5. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), and 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms) of fraction **5a**, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 3.

Fraction 5 - MALDI-TOF-MS analysis of fraction **5** revealed $[\text{M}+\text{Na}]^+$ pseudomolecular ions at m/z 527 and 689, corresponding with Hex_3 and Hex_4 , respectively. Fraction **5** was further separated on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH) (Figure 2C), rendering a major fraction **5a** and a minor fraction **5b**.

MALDI-TOF-MS analysis of **5a** revealed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 689, corresponding with Hex_4 . The 1D ^1H NMR spectrum of **5a** (Figure 5) showed anomeric signals at δ 5.404 (**A** H-1, $^3J_{1,2}$ 3.8 Hz), 5.241 (**R α** H-1, $^3J_{1,2}$ 3.6 Hz), 4.957 (**B** H-1 and **E** H-1, $^3J_{1,2}$ 3.4 Hz), and 4.669 (**R β** H-1, $^3J_{1,2}$ 8.1 Hz). Starting from the anomeric signals in the 2D ^1H - ^1H TOCSY spectrum (Figure 5/180 ms), all chemical shifts of the non-anomeric protons could be determined (Table 1). The chemical shift positions of H-1 α and H-1 β of the reducing residue **R** at δ 5.241 and 4.669, respectively, are in agreement with the occurrence of a $-(1\rightarrow6)\text{-D-Glc}$ unit.^{8,9} This is further reflected by

the **R α** and **R β** H-5 resonances at δ 4.00 and 3.65, respectively (library data: δ 3.99-4.01 and 3.63-3.65).^{8,9} The **A** H-1 value at 5.404 ppm is indicative of an (-) α -D-Glcp-(1 \rightarrow 4)-unit.⁸ On the **A** H-1 track the set of H-5, H-6a, and H-6b is observed at δ 3.90, 3.72, and 3.97, respectively, corresponding with a -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- unit (compare with panose),⁸ leading to the occurrence of a **X1 \rightarrow 6A1 \rightarrow 4Y** element.

The overlapping anomeric signals of residue **B** and **E**, indicative of (-) α -D-Glcp-(1 \rightarrow 6)- units, form a single track in the TOCSY spectrum. However, the difference between the two sets of H-2, H-3, and H-4 can clearly be observed in the built-up series of mixing times (data not shown). The chemical shifts of the set of **B** H-2, H-3, and H-4 at δ 3.60, 3.99, and 3.65, respectively, are indicative of the occurrence of a -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)- unit (see compound **3a**),⁸ leading to an **X1 \rightarrow 6A1 \rightarrow 4B1 \rightarrow 6Z** structure. The presence of **E** H-4 at δ 3.427 (dd, 1 H) reflects the occurrence of one terminal α -D-Glcp-(1 \rightarrow 6)- unit.⁸ Considering that **E** is a terminal residue and **R** the reducing residue, leads unequivocally to the conclusion that compound **5a** has the structure **E1 \rightarrow 6A1 \rightarrow 4B1 \rightarrow 6R**, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

The structure established for tetrasaccharide **5a** was verified by 2D ¹H-¹H ROESY measurements (Figure 5), showing inter-residual cross-peaks between **A** H-1 and **B** H-4, **B** H-1 and **R** H-6a, and **E** H-1 and **A** H-6a. It should be noted that the ¹H NMR data of **E1 \rightarrow 6A1 \rightarrow 4** fit perfectly those observed in the **E1 \rightarrow 6A1 \rightarrow 4** element of panose.⁸

The MALDI-TOF mass spectrum of **5b** revealed an [M+Na]⁺ pseudomolecular ion at m/z 527, corresponding with Hex₃. The 1D ¹H NMR spectrum matches that of panose,⁸ i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp (Scheme 1).

Fraction 6 - MALDI-TOF-MS analysis of fraction **6** gave rise to a single [M+Na]⁺ pseudomolecular ion at m/z 851, corresponding with Hex₅. Fraction **6** was further separated on CarboPac PA-1 under isocratic conditions (Figure 2C), yielding one major fraction **6a** (Hex₃).

The 1D ¹H NMR spectrum of **6a** (Figure 6) revealed six anomeric signals at δ 5.361 (**C** H-1, ³*J*_{1,2} 3.8 Hz), 5.341 (**D** H-1, ³*J*_{1,2} 3.6 Hz), 5.241 (**R α** H-1, ³*J*_{1,2} 3.8 Hz), 4.986 (**E** H-1, ³*J*_{1,2} 3.6 Hz), 4.960/4.951 (**B** H-1, ³*J*_{1,2} 3.6 Hz), and 4.670 (**R β** H-1, ³*J*_{1,2} 7.9 Hz). The chemical shifts of the non-anomeric protons, as included in Table 1, were obtained from 2D TOCSY spectra with increasing mixing times (Figure 6/180 ms). The chemical shift values of **R α** and **R β** H-1 at δ 5.241 and 4.670, respectively, correspond

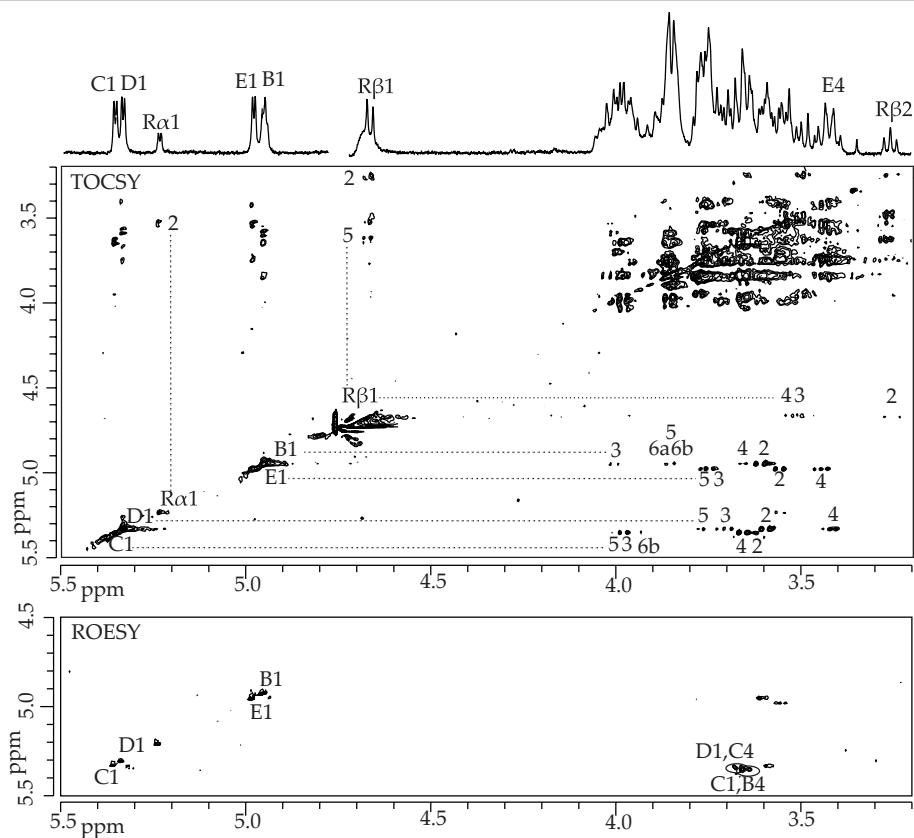


Figure 6. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), and 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms) of fraction **6a**, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 3.

with the values established for the occurrence of a $-(1\rightarrow6)\text{-D-Glcp}$ unit.⁸ The δ values of **R α** H-5 at 4.00 ppm and **R β** H-5 at 3.64 ppm confirm the 6-substitution of residue **R**.⁸ The H-1 signal of residue **B**, split into **B α** and **B β** at δ 4.951 and 4.960, respectively, corresponds with an $(\text{-})\alpha\text{-D-Glcp}(1\rightarrow6)\text{-}$ unit, influenced by the reducing residue anomeric configuration, suggesting a **B1** \rightarrow **6R** element. The set of **B** H-2, H-3, and H-4 at δ 3.61, 4.01, and 3.65, respectively, revealed that **B** is a $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}(1\rightarrow6)\text{-}$ unit (see also compounds **3a** and **5a**). In a similar way as discussed for compound **4a**, the **D** H-4 (δ 3.418) and **E** H-4 (δ 3.435) chemical shifts indicate the presence of a terminal $\alpha\text{-D-Glcp-D}(1\rightarrow4)\text{-}$ and a terminal $\alpha\text{-D-Glcp-E}(1\rightarrow6)\text{-}$ unit, respectively. In fact, the **D** and **E** H-1 signals at δ 5.341 and 4.986, respectively, match those of the **D** and **E** residues in **4a**, occurring in a **D1** \rightarrow **4[E1** \rightarrow **6]X** element. Taking into account that **EPS35-5** only contains $(\alpha 1\text{-}4)$ and $(\alpha 1\text{-}6)$ linkages, the **C** H-1 value (δ 5.361) has

to be correlated with an (α 1-4)-linked residue. Moreover, the set of **C** H-2, H-3, and H-4 values indicates a $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit. The H-5 signal, however, is found at δ 4.05 (library maltotriose residue **B**: $\delta_{\text{H-5}}$ 3.85),⁸ being in favour of an additional 6-substitution (see compound **4a**), and leading to the conclusion that residue **C** is a $-(1\rightarrow4,6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ unit. Taken together, these data lead to a **D1** \rightarrow **4**[**E1** \rightarrow **6**]**C1** \rightarrow **4****B1** \rightarrow **6****R** sequence for compound **6a**, i.e. $\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}[\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}]\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-D-Glcp}$ (Scheme 1). It is interesting to note that branching has a specific influence on the position of the H-1 signal of the branched residue: δ 5.399 for $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ (maltotriose)⁸ and δ 5.361 for $-(1\rightarrow4,6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ (compound **6a**). Remarkably, the downfield shift of H-5 when going from $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}$ to $-(1\rightarrow4,6)\text{-}\alpha\text{-D-Glcp}\text{-}$ in case of residue **C** ($\Delta\delta$ + 0.20 ppm) is similar to that observed for residue **R α** in **4a** ($\Delta\delta$ + 0.17 ppm).

The established structure of the pentasaccharide was verified by 2D ¹H-¹H ROESY measurements (Figure 6), showing inter-residual cross-peaks between **C** H-1 and **B** H-4, and between **D** H-1 and **C** H-1. Inter-residual cross-peaks for **B** H-1 and **E** H-1 were not observed, due to low signal intensity.

Fraction 7 - The MALDI-TOF mass spectrum of fraction **7** revealed $[\text{M}+\text{Na}]^+$ pseudomolecular ions at m/z 851 and 1013, corresponding with Hex₅ and Hex₆, respectively. Fraction **7** was further separated on CarboPac PA-1 (eluent: 100 mM NaOAc in 100 mM NaOH) (Figure 2C), yielding major fraction **7a** (Hex₆) and minor fraction **7b** (Hex₅).

The 1D ¹H NMR spectrum of **7a** (Figure 7) showed anomeric signals at δ 5.387 (**A**^I H-1, $^3J_{1,2}$ 3.9 Hz), 5.380 (**A**^{II} H-1, $^3J_{1,2}$ 3.9 Hz), 5.239 (**R α** H-1, $^3J_{1,2}$ 3.6 Hz), 4.955 (**B**^I H-1 and **E** H-1, $^3J_{1,2}$ 3.7 Hz), and 4.669 (**R β** H-1, $^3J_{1,2}$ 7.8 Hz). Table 1 includes the chemical shifts of the various residual protons as determined by 2D ¹H-¹H TOCSY analysis (Figure 7/180 ms). In TOCSY experiments with different mixing times (data not shown) the built-up of the scalar coupling network could be followed. The set of H-1 α and H-1 β chemical shifts of the reducing residue **R** corresponds with the value of a $-(1\rightarrow6)\text{-D-Glcp}$ unit.⁸ The **E** H-4 signal at δ 3.432 (dd, 1 H) corresponds with a terminal $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit (compare with compounds **3a**, **4a**, **5a**, and **6a**), thereby indicating a linear structure. Taking into account the methylation analysis data of **EPS35-5**, the **A** H-1 signals at δ 5.387 and 5.380 reflect the occurrence of two (α -D-Glcp-(1 \rightarrow 4)- units, denoted **A**^I and **A**^{II}.⁸ The chemical shift patterns of **A**^I and **A**^{II} are the same and fit best with the values found for residue **A** in compound **5a**, indicating

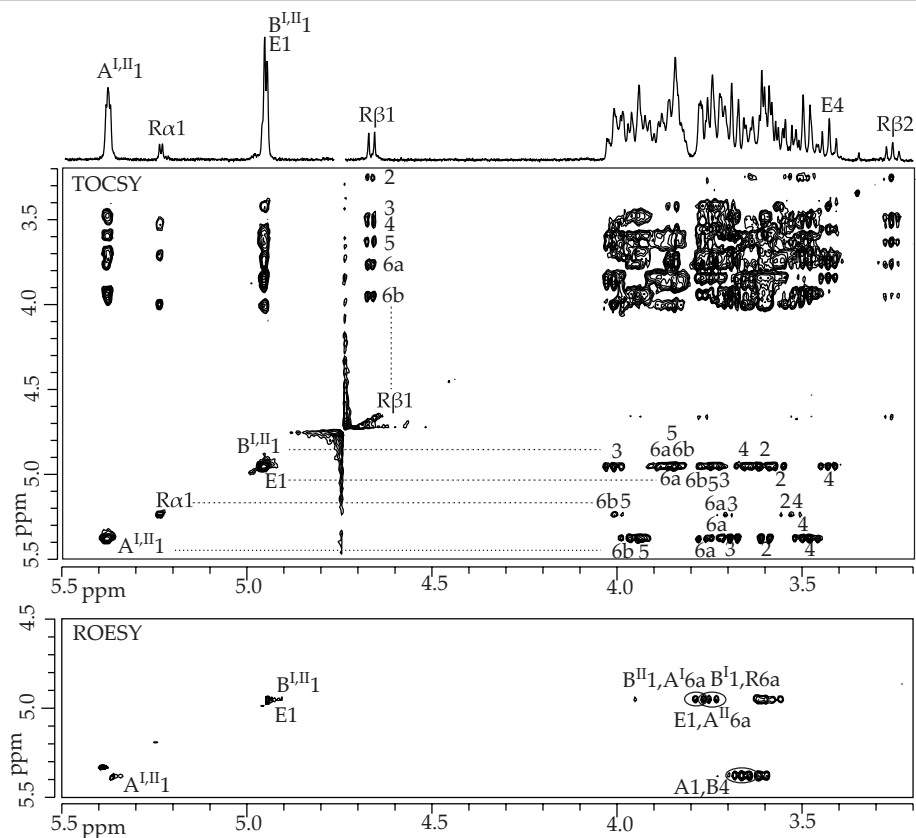


Figure 7. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), and 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms) of fraction **7a**, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 3.

two $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ units in compound **7a** (**A** H-5 at δ 3.92). The surface area of the **B** and **E** H-1 signals at δ 4.955 matches 3 H, indicating the presence of three $(\text{-})\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ units. In the TOCSY experiments with incremental mixing times (data not shown), two sets of chemical shifts could be observed. The set **B** H-2, H-3, and H-4 at δ 3.60, 4.01, and 3.66, respectively, matches the values of a $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit (see compounds **3a**, **5a**, and **6a**). Residue **E** showed the typical proton chemical shift pattern of a terminal $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}$ unit (see above). Combination of the various data indicates that the hexasaccharide contains twice the **X** $1\rightarrow$ **4B** $1\rightarrow$ **6Y** sequence and twice the **X'** $1\rightarrow$ **6A** $1\rightarrow$ **4Y'** sequence. With residue **R** in an **X''** $1\rightarrow$ **6R** sequence and **E** as a terminal unit, these data lead to a **E** $1\rightarrow$ **6A** $^{II}1\rightarrow$ **4B** $^{II}1\rightarrow$ **6A** $^I1\rightarrow$ **4B** $^I1\rightarrow$ **6R** sequence for **7a**, i.e. $\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow6)\text{-D-Glcp}$ (Scheme 1).

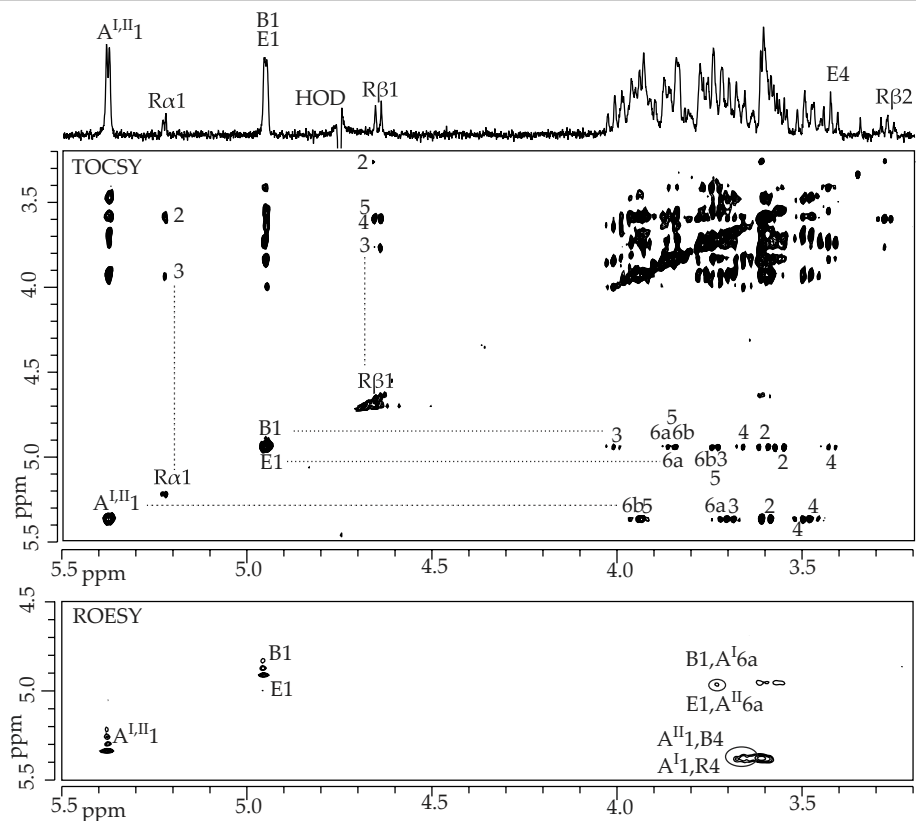


Figure 8. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectrum (mixing time 180 ms), and 2D ^1H - ^1H ROESY spectrum (mixing time 300 ms) of fraction **7b**, recorded at 300K in D_2O . For an explanation of the coding systems, see Figure 3.

The structure of fraction **7a** was verified by 2D ^1H - ^1H ROESY measurements (Figure 7), showing inter-residual cross-peaks between **A** H-1 and **B** H-4, between **B** H-1 and **A** H-6a and **R** H-6a, and between **E** H-1 and **A** H-6a.

1D ^1H NMR analysis of fraction **7b** (Figure 8) revealed anomeric signals at δ 5.381 (**A**^{I,II} H-1, $^3J_{1,2}$ 3.8 Hz), 5.225 (**R** α H-1, $^3J_{1,2}$ 3.6 Hz), 4.954 (**B** H-1 and **E** H-1, $^3J_{1,2}$ 3.8 Hz), and 4.649 (**R** β H-1, $^3J_{1,2}$ 8.0 Hz). The chemical shift values as determined by 2D ^1H - ^1H TOCSY spectroscopy (Figure 8/180 ms) are shown in Table 1. The set of anomeric signals of **R** α and **R** β at δ 5.225 and 4.649, respectively, corresponds with the occurrence of a $-(1\rightarrow4)\text{-D-Glcp}$ unit.⁸ Similar to compound **7a**, the two **A** residues of **7b** that overlap on the **A**^{I,II} TOCSY track develop exactly the same built-up chemical shift patterns in the TOCSY experiments with different mixing times (data not shown), demonstrating the presence of two $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow4)\text{-}$ units. Note also the similarity with residue **A** in compound **5a**. The H-1 signal at δ 4.954 (2 H) indicates

the occurrence of two (-) α -D-Glcp-(1 \rightarrow 6)- units, i.e. residues **B** and **E**. In TOCSY experiments with incremental mixing times (data not shown) two distinctly different proton chemical shift patterns develop. The **B** H-3 resonance is observed at δ 4.00, indicating a -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)- unit. On guidance of the structural-reporter-group concept, the **E** H-4 signal at δ 3.429 identified residue **E** as a terminal α -D-Glcp-(1 \rightarrow 6)- unit. With twice the **X1** \rightarrow **6A1** \rightarrow **4Y** sequence and a 4-substituted **R** unit, it can be concluded that compound **7b** has a **E1** \rightarrow **6A^{II}1** \rightarrow **4B1** \rightarrow **6A^I1** \rightarrow **4R** sequence, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp (Scheme 1).

The sequence established for pentasaccharide **7b** was verified by 2D ^1H - ^1H ROESY measurements (Figure 8), revealing inter-residual cross-peaks between **A** H-1 and **B** H-4 and **R β** H-4, between **B** H-1 and **A^I** H-6a, and between **E** H-1 and **A^{II}** H-6a.

2D NMR spectroscopy of **EPS35-5**

In order to unravel the 1D ^1H NMR spectrum of **EPS35-5** (Figure 1), 2D TOCSY experiments with increasing mixing times of 10, 30, 60, 120, and 180 ms, as well as data from the ^{13}C - ^1H HSQC measurements of **EPS35-5** were interpreted (Table 1) using the data collected from the oligosaccharides obtained by partial acid hydrolysis of **EPS35-5** (section 2.2).

In the 60 ms TOCSY spectrum (Figure 9), the **A/D** H-1-track at δ 5.36 (indicating (α 1-4)-linked D-Glcp)⁸ showed one chemical shift value for H-2 at δ 3.60 and one value for H-3 at δ 3.70, indicating that these residues are not 4-substituted (library data from maltotriose: $\delta_{\text{H-2}}$ 3.63 and $\delta_{\text{H-3}}$ 3.96).⁸ In the 120 ms TOCSY spectrum (Figure 9) the **A/D** H-1 track revealed two different H-4 signals at δ 3.48 (residue **A**), indicating -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- units (see residue **A** in **5a**, **7a**, and **7b**), and 3.42 (residue **D**), indicating terminal α -D-Glcp-(1 \rightarrow 4)- units (see residue **D** in **3a**, **4a**, and **6a**).

The **C/F** H-1 track in the 60 ms TOCSY spectrum at δ 5.33 (indicating (α 1-4)-linked D-Glcp residues) showed two different H-2 signals for residues **C** and **F** at δ 3.65 (compare with residue **C** in **6a**) and 3.60, respectively, reflecting different substitution patterns (Scheme 1). The H-3 signal, however, is observed at δ 3.98 for both residues, indicating the occurrence of -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 4)- units. The **C** H-5 resonance at δ 4.05 fits with the occurrence of a -(1 \rightarrow 4,6)- α -D-Glcp-(1 \rightarrow 4)- unit (see residue **C** in **6a**).

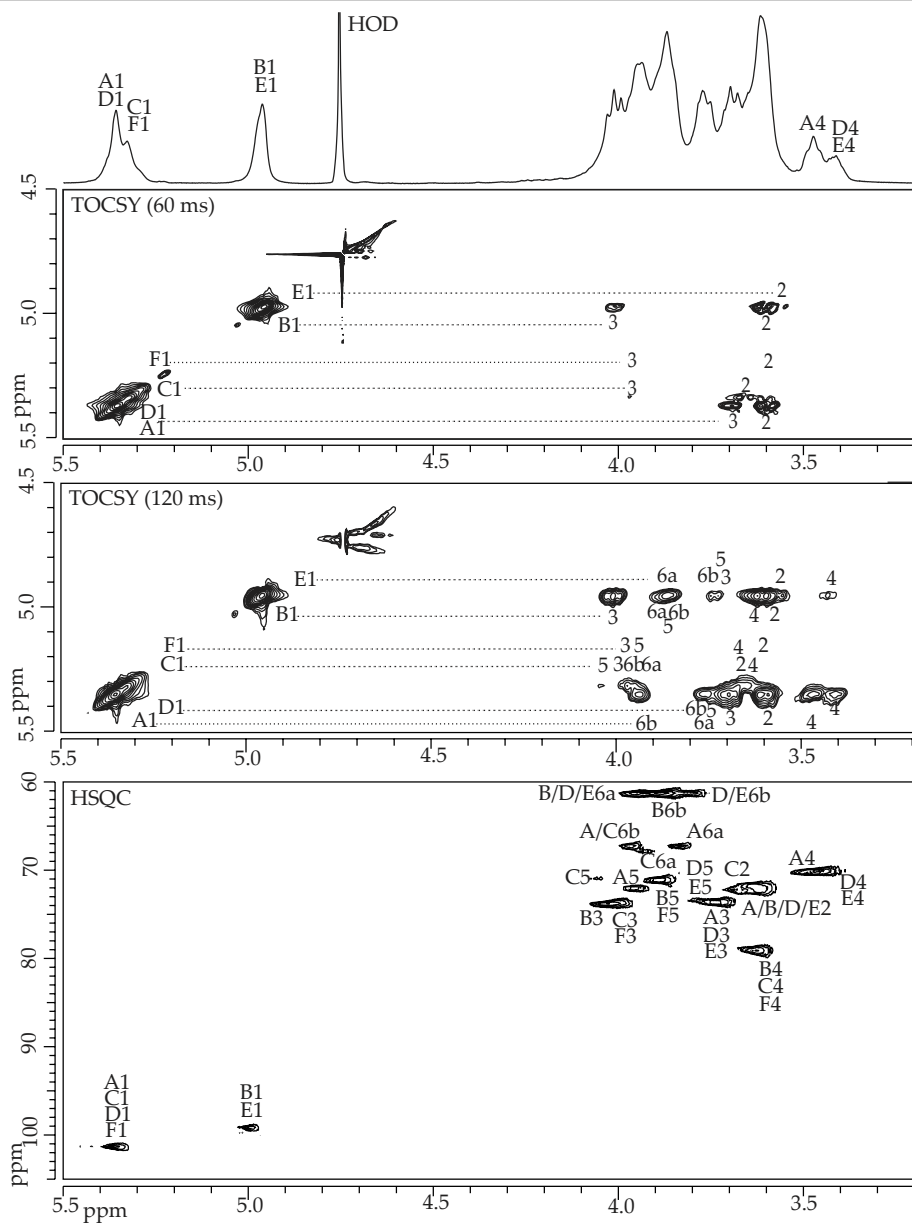
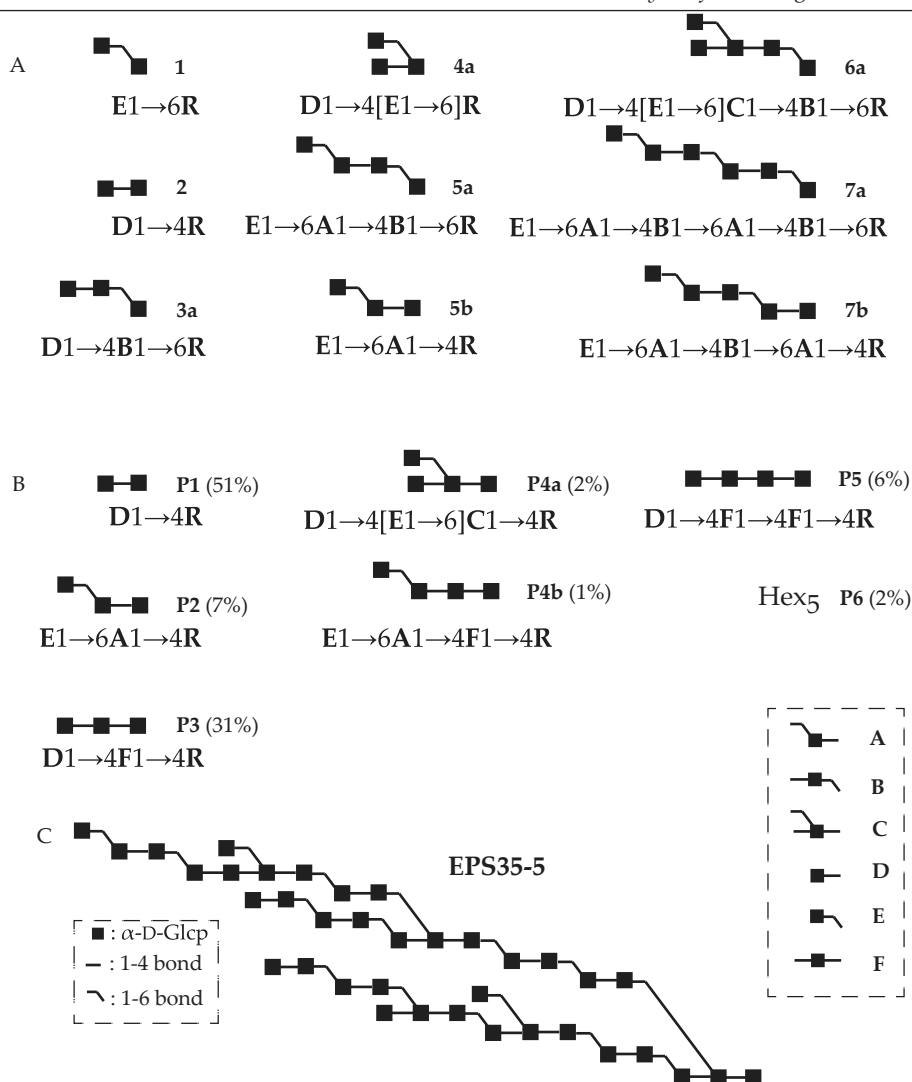


Figure 9. 500-MHz 1D ^1H NMR spectrum, 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms), and 2D ^{13}C - ^1H HSQC spectrum of EPS35-5 recorded at 300K in D_2O . Anomeric protons in the TOCSY spectrum (A1, etc.) have been indicated on the diagonal; numbers in the horizontal tracks belong to the cross-peaks of the scalar-coupling network of the residues indicated. In the ^{13}C - ^1H HSQC spectrum A1 denotes the cross-peak between H-1 and C-1 of residue A, etc.



Scheme 1. Structures of oligosaccharide fragments, obtained by (A) partial acid hydrolysis and (B) pullulanase M1 hydrolysis of EPS35-5, and (C) visual representation that includes all structural features identified for EPS35-5 including all structural features established. Quantities determined for the oligosaccharides obtained by pullulanase M1 hydrolysis are included between brackets. Residue labels correspond with those used in the figures and are presented on the right.

On the **B/E** H-1 track in the 60 ms TOCSY spectrum at δ 4.96 (indicating (α 1-6)-linked D-Glcp residues) two different H-2 signals are observed, a strong signal at δ 3.61 (residue **B**) and a weak signal at δ 3.56 (residue **E**). The set of **B** H-2 and H-3 at δ 3.61 and 4.01 corresponds with the occurrence of a -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)- unit (compare with residue **B** in **3a**, **5a**, **6a**, **7a**, and **7b**). In the 120 ms TOCSY spectrum residue **E**

Table 1. ^1H chemical shifts of oligosaccharide fragments of **EPS35-5** obtained by partial acid hydrolysis and of intact **EPS35-5**, measured at 300K in D_2O . Residue labels correspond to those used in Scheme 1.

Residue	3a	4a	5a	6a	7a	7b	EPS35-5
R α -1	5.240	5.222	5.241	5.241	5.239	5.225	-
R α -2	3.55	3.59	3.54	3.55	3.55	3.58	-
R α -3	3.70	3.99	3.67	3.71	3.70	4.01	-
R α -4	3.52	3.67	3.50	3.50	3.50	n.d.	-
R α -5	4.00	4.113	4.00	4.00	4.00	n.d.	-
R α -6a	3.73	3.84	3.73	3.72	3.73	n.d.	-
R α -6b	4.01	4.01	4.03	4.01	4.01	n.d.	-
R β -1	4.669	4.660	4.669	4.670	4.669	4.649	-
R β -2	3.253	3.285	3.253	3.257	3.255	3.271	-
R β -3	3.48	3.78	3.46	3.48	3.49	3.77	-
R β -4	3.51	3.67	3.51	3.51	3.53	3.66	-
R β -5	3.64	3.89	3.65	3.64	3.63	3.63	-
R β -6a	3.75	3.78	3.80	3.76	3.77	n.d.	-
R β -6b	3.927	3.95	3.97	3.96	3.96	n.d.	-
A-1	-	-	5.404	-	5.387/80	5.381	5.36
A-2	-	-	3.60	-	3.60	3.60	3.60
A-3	-	-	3.68	-	3.70	3.70	3.70
A-4	-	-	3.496	-	3.492	3.497/92	3.48
A-5	-	-	3.90	-	3.92	3.92	n.d.
A-6a	-	-	3.72	-	3.75	3.74	3.74
A-6b	-	-	3.97	-	3.97	3.96	3.94
B-1	4.958/50	-	4.957	4.960/51	4.955	4.954	4.96
B-2	3.60	-	3.60	3.61	3.60	3.61	3.61
B-3	4.01	-	3.99	4.01	4.01	4.00	4.01
B-4	3.66	-	3.65	3.65	3.66	3.65	3.65
B-5	3.85	-	3.85	3.85	3.85	3.85	3.85
B-6a	3.86	-	3.85	3.86	3.86	3.87	3.89
B-6b	3.80	-	3.82	3.83	3.82	3.84	3.85
C-1	-	-	-	5.361	-	-	5.33
C-2	-	-	-	3.65	-	-	3.65
C-3	-	-	-	3.98	-	-	3.98
C-4	-	-	-	3.67	-	-	3.65
C-5	-	-	-	4.05	-	-	4.05
C-6a	-	-	-	n.d.	-	-	3.91
C-6b	-	-	-	3.97	-	-	3.97
D-1	5.412	5.354/44	-	5.341	-	-	5.36
D-2	3.59	3.58	-	3.59	-	-	3.59
D-3	3.70	3.71	-	3.70	-	-	3.70
D-4	3.417	3.410	-	3.418	-	-	3.42
D-5	3.75	3.76	-	3.75	-	-	3.75
D-6a	3.86	3.86	-	3.86	-	-	n.d.
D-6b	3.78	3.77	-	3.77	-	-	3.77

(Table 1 continued)

E-1	-	4.985/76	4.957	4.986	4.955	4.954	4.96
E-2	-	3.54	3.56	3.55	3.56	3.57	3.56
E-3	-	3.75	3.72	3.76	3.75	3.75	3.75
E-4	-	3.431	3.427	3.435	3.432	3.429	3.43
E-5	-	3.74	3.73	3.75	3.75	3.74	3.75
E-6a	-	3.85	3.85	n.d.	3.85	3.85	3.85
E-6b	-	3.76	3.74	3.77	3.77	3.75	3.75
F-1	-	-	-	-	-	-	5.33
F-2	-	-	-	-	-	-	3.60
F-3	-	-	-	-	-	-	3.98
F-4	-	-	-	-	-	-	3.65
F-5	-	-	-	-	-	-	3.97
F-6a	-	-	-	-	-	-	n.d.
F-6b	-	-	-	-	-	-	n.d.

H-4 is observed at δ 3.43, in agreement with a terminal α -D-Glc p -(1 \rightarrow 6)- unit. Compare also the other chemical shift values with those found for residue **E** in compounds **4a**, **5a**, **6a**, **7a**, and **7b** (Table 1). On the same track the set of **B** H-4, H-5, H-6a, and H-6b supports the occurrence of a -(1 \rightarrow 4)- α -D-Glc p -(1 \rightarrow 6)- unit. The absence of an H-4 cross-peak at δ 3.50 indicates that a -(1 \rightarrow 6)- α -D-Glc p -(1 \rightarrow 6)- element does not occur in **EPS35-5**.^{8,9} The fact that the H-6a and H-6b values on the (α 1-6) anomeric track exclude 6-substitution, further confirms the conclusion that only single (α 1-6) bridges occur in **EPS35-5**.

The substitution patterns established from the TOCSY spectra were verified using ¹³C-¹H HSQC spectroscopy (Figure 9), yielding δ_{C-4} values of 78.9 ppm for residues **B**, **C**, and **F**, reflecting the 4-substitution of these residues.¹¹ Residues **A** and **C** gave rise to δ_{C-6} values of 67.1 ppm, reflecting the 6-substitution of these residues.^{8,11} For the residues **B**, **D**, and **E** the δ_{C-6} value of 61.3 ppm (**F** was not detected, due to minor occurrence) confirmed the absence of a 6-substitution.

Enzymatic hydrolysis

The occurrence of both (α 1-4) and (α 1-6) linkages in **EPS35-5** may render the polysaccharide susceptible to hydrolysis by *Klebsiella planticola* pullulanase M1, which specifically cleaves the (α 1-6) linkage in (-) α -D-Glc p -(1 \rightarrow 4)- α -D-Glc p -(1 \rightarrow 6)-sequences.¹² The substrate specificity of the commercially available enzyme was checked using a series of model compounds (data not shown).

A sample of **EPS35-5** (20 mg) was incubated with pullulanase M1, and the initial progress of the hydrolysis was followed by analysing aliquots at 30, 60, 90, and 120 min by MALDI-TOF-MS and 1D ¹H NMR spectroscopy. As clear indications were obtained

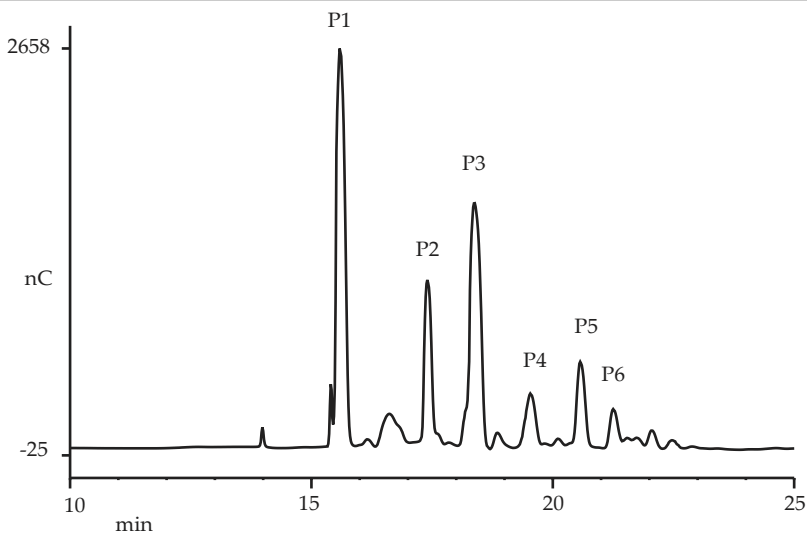


Figure 10. HPAEC-PAD profile of pullulanase M1 hydrolysate on CarboPac PA-100, using a linear gradient.

for the formation of oligosaccharide fragments (MALDI-TOF-MS), accompanied by a decrease in the anomeric signal of $(-)\alpha\text{-D-Glcp-(1}\rightarrow\text{6)-}$ units ($\delta_{\text{H-1}}$ 4.96; 1D $^1\text{H NMR}$ spectroscopy), the hydrolysis was continued for 96 h. The resulting mixture was analysed by MALDI-TOF-MS, revealing $[\text{M}+\text{Na}]^+$ pseudomolecular ions at m/z 527, 689, and 851, corresponding with Hex_3 , Hex_4 , and Hex_5 , respectively. An aliquot of the mixture was analysed on CarboPac PA-100 (Figure 10), using 0 - 300 mM NaOAc in 100 mM NaOH as eluent; for identification and quantification purposes a mixture of maltose, maltotriose, panose, maltotetraose, maltopentaose, and maltohexaose standards was subjected to the same pullulanase hydrolysis and desalting procedure as the **EPS35-5** sample, and also analysed on CarboPac PA-100, using 0 - 300 mM NaOAc in 100 mM NaOH as eluent. The bulk of the pullulanase M1 hydrolysate was separated on CarboPac PA-1, using the same elution conditions, yielding fractions **P1** to **P6**.

Fraction P1 - HPAEC and 1D $^1\text{H NMR}$ analysis showed fraction **P1** to contain maltose,⁸ $\alpha\text{-D-Glcp-(1}\rightarrow\text{4)-D-Glcp}$. Using the response factor of maltose in the standard mixture, the relative amount of **P1** in the hydrolysate was determined to be 51 mol % (Scheme 1).

Fraction P2 - MALDI-TOF-MS analysis of fraction **P2** revealed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 527, corresponding with Hex_3 . The retention time of **P2** on CarboPac PA-100 matched that of panose, whereas the 1D $^1\text{H NMR}$ spectrum

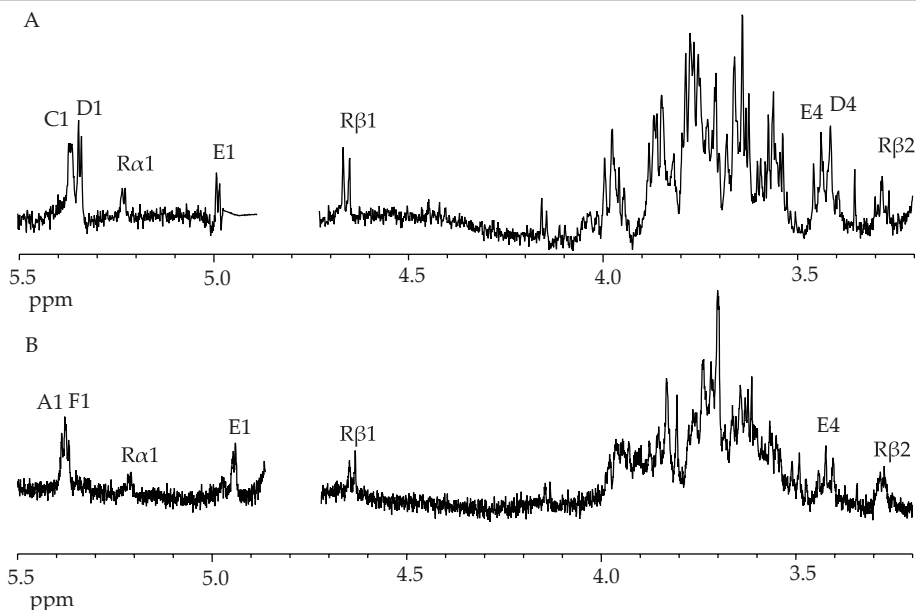


Figure 11. 500-MHz ^1H NMR spectra of (A) compound **P4a** and (B) compound **P4b** obtained by pullulanase M1 hydrolysis of **EPS35-5**, recorded at 300K in D_2O .

of **P2** corresponded with panose,⁸ α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp. Using the response factor of panose in the standard mixture, the relative amount of **P2** was determined to be 7 mol % (Scheme 1).

Fraction P3 - The MALDI-TOF mass spectrum of fraction **P3** showed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 527, corresponding with Hex₃. The retention time of **P3** on CarboPac PA-100 matched that of maltotriose, and the 1D ^1H NMR spectrum of **P3** also corresponded with that of maltotriose, α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp.⁸ The amount of **P3** was quantified, using the response factor of maltotriose, yielding 31 mol % (Scheme 1).

Fraction P4 - MALDI-TOF-MS analysis of fraction **P4** revealed a single $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 689, corresponding with Hex₄. Using the response factor of maltotetraose, the relative amount of **P4** was estimated to be 3 mol %. The 1D ^1H NMR spectrum of **P4** indicated the presence of a mixture of two components in a 2:1 ratio. Fraction **P4** was further separated on CarboPac PA-1, using 100 mM NaOAc in 100 mM NaOH, yielding a major fraction **P4a** (Hex₄) and a minor fraction **P4b** (Hex₄).

The 1D ^1H NMR spectrum of **P4a** (Figure 11A) showed 5 anomeric signals at δ 5.363 (**C** H-1, $^3J_{1,2}$ 3.2 Hz), 5.340 (**D** H-1, $^3J_{1,2}$ 3.2 Hz), 5.226 (**R α** H-1, $^3J_{1,2}$ 3.2 Hz),

4.985 (**E** H-1, $^3J_{1,2}$ 3.2 Hz), and 4.655 (**Rβ** H-1, $^3J_{1,2}$ 7.9 Hz). The set of anomeric signals of **Rα** and **Rβ** corresponds with the occurrence of a reducing $-(1\rightarrow4)\text{-D-Glc}p$ unit.⁸ It is clear that one $(\alpha1\text{-}6)$ -linked and two $(\alpha1\text{-}4)$ -linked D-Glc p residues are present in Hex₄. The H-1 signals of residues **C**, **D**, and **E** fit with the values observed in compound **6a** for the **D1** \rightarrow **4**[**E1** \rightarrow **6**]**C** branch-point. Note the two overlapping H-4 signals at δ 3.41 and 3.43, corresponding with the occurrence of both $\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ and $\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ units. These data lead to the sequence **D1** \rightarrow **4**[**E1** \rightarrow **6**]**C1** \rightarrow **4R** for **P4a**, i.e. $\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}[\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}]\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-D-Glc}p$ (Scheme 1).

The 1D ^1H NMR spectrum of **P4b** (Figure 11B) revealed two $(-)\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ units (residues **A** and **F**, $\delta_{\text{H-1}}$ 5.382, 2 H) and one $(-)\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ unit (residue **E** $\delta_{\text{H-1}}$ 4.944, 1 H). Furthermore, only one H-4 signal of a terminal residue at δ 3.43 (residue **E**) is observed, indicating a linear structure. The reducing residue **R** H-1 α and H-1 β resonances at δ 5.225 and 4.650, respectively, correspond with the occurrence of a reducing $-(1\rightarrow4)\text{-D-Glc}p$ unit. These data lead to an **E1** \rightarrow **6A1** \rightarrow **4F1** \rightarrow **4R** sequence for the linear tetrasaccharide **P4b**, i.e. $\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-D-Glc}p$ (Scheme 1).

Fraction P5 - The MALDI-TOF-MS spectrum of fraction **P5** showed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 689, corresponding with Hex₄. The retention time of **P5** on CarboPac PA-100, as well as its 1D ^1H NMR spectrum, matched that of maltotetraose, $\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-D-Glc}p$ (Scheme 1). Using the response factor of the maltotetraose standard, the relative amount of **P5** was determined to be 6 mol %.

Fraction P6 - Fraction **P6** was analysed by MALDI-TOF-MS, revealing an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 851, corresponding with Hex₅. The retention time of **P6** on CarboPac PA-100 did not match that of maltopentaose, which eluted later than **P6**, indicating a structure with at least one $(\alpha1\text{-}6)$ linkage. Using the response factor of maltopentaose the relative abundance of **P6** was estimated to be 2 mol %. Fraction **P6** did not contain sufficient material for further analysis.

Conclusion

A visual representation, that includes all identified structural features of the native **EPS35-5** α -D-glucan was formulated by combining ^1H NMR data of **EPS35-5** with methylation analysis data (section 2.1), as well as with structural information of oligosaccharides obtained by partial acid hydrolysis (section 2.2) and pullulanase M1 hydrolysis (section 2.4) of **EPS35-5**. The visual representation of **EPS35-5** is depicted in Scheme 1. It can be concluded that the (1 \rightarrow 4, 1 \rightarrow 6)- α -D-glucan of *Lb. reuteri* strain 35-5 is built-up from maltose, maltotriose, and maltotetraose elements with single (α 1-6) bridges. The majority of the terminal residues are α -D-Glc p -(1 \rightarrow 4)- units, however, a small amount of α -D-Glc p -(1 \rightarrow 6)- units also occurs. The significant occurrence of compound **P1** leads to a structure with large amounts of alternating (α 1-4) and (α 1-6) linkages. The methylation analysis data dictate the boundaries for the statistical distribution of the six types of building blocks in **EPS35-5**. All oligosaccharides obtained by partial acid hydrolysis and pullulanase M1 hydrolysis of **EPS35-5** fit into the picture, reflecting the validity of the representation.

Materials and Methods

Materials

EPS35-5 was a gift from TNO Quality of Life, Zeist, The Netherlands. D₂O (99.9 atom%) was purchased from Cambridge isotope laboratories, Inc., Andover, MA. A suspension of pullulanase M1 (*Klebsiella planticola*) in 3.2 M (NH₄)₂SO₄ was purchased from Megazyme International Ireland Ltd., Bray, Ireland.

Monosaccharide analysis

A polysaccharide sample was subjected to methanolysis (1.0 M methanolic HCl, 24 h, 85 °C), followed by re-*N*-acetylation and trimethylsilylation (1:1:5 hexamethyldisilazane-trimethylchlorosilane-pyridine; 30 min, room temperature). The mixture of trimethylsilylated methyl glycosides was analysed by GLC on an EC-1 column (30 m x 0.32 mm, Alltech Associates Inc., Illinois, USA) using a Chrompack CP 9002 gas chromatograph (temperature gradient, 140-240 °C at 4 °C/min). The identification of the monosaccharide derivatives was confirmed by GLC-EI-MS.¹³

Linkage analysis

A polysaccharide sample was permethylated using CH_3I and solid NaOH in DMSO , as described earlier.¹⁴ After hydrolysis with 2 M TFA (2 h, 120 °C), the partially methylated monosaccharides were reduced with NaBD_4 (2 h, room temperature). Conventional work-up, involving neutralisation with HOAc and removal of boric acid by co-evaporation with MeOH, followed by acetylation with 1:1 acetic anhydride:pyridine (3 h, 120 °C), yielded a mixture of partially methylated alditol acetates, which was analysed by GLC-EI-MS.^{13,15}

Partial acid hydrolysis

In pilot experiments, 5-mg aliquots of **EPS35-5** were hydrolysed for 30, 60, 90, and 120 min (0.5 M TFA, 90 °C). After concentration under reduced pressure, the residues were analysed by MALDI-TOF-MS and 1D ^1H NMR spectroscopy. Based on the initial results, **EPS35-5** (500 mg) was hydrolysed in 25 mL 0.5 M TFA for 2 h at 90 °C. The solution was concentrated under reduced pressure at a rotary evaporator, and the residue was separated on a Bio-Gel P-4 column (400 x 12 mm, BioRad), eluted with 25 mM NH_4HCO_3 ; 0.9-mL fractions were collected at a flow rate of 13.5 mL/h. Fractions were tested for the presence of carbohydrate by a TLC spot-test with orcinol/ H_2SO_4 staining. Carbohydrate-containing fractions were analysed by MALDI-TOF-MS.

Enzymatic hydrolysis

An equimolar mixture of maltose, maltotriose, panose, maltotetraose, maltopentaose, and maltohexaose (50 mM, 1 mL) was dissolved in assay buffer (50 mM NaOAc, pH 5.0 with 4 M HOAc), and subsequently 5 μL pullulanase M1 suspension was added. A sample of 20 mg **EPS35-5** was dissolved in 2 mL assay buffer (see above), and 30 μL pullulanase M1 suspension was added. Hydrolysis of samples was carried out for 96 h at 37 °C. From each sample aliquots were taken at $t = 30, 60, 90,$ and 120 min, and subsequently boiled for 2 min to terminate enzymatic hydrolysis. Each sample was desalted on CarboGraph SPE columns (150 mg graphitised carbon, Alltech) using 1:3 acetonitrile: H_2O as eluent, lyophilised, and subsequently analysed by HPAEC-PAD, MALDI-TOF-MS, and 1D ^1H NMR spectroscopy. The bulk of the EPS35-5 sample was finally desalted, lyophilised and re-dissolved in 300 μL H_2O . Pullulanase hydrolysis fragments were separated by HPAEC and analysed by MALDI-TOF-MS and 1D ^1H NMR spectroscopy.

HPAEC-PAD

High-pH anion-exchange chromatography was performed on a Dionex DX500 workstation, equipped with an ED40 pulsed amperometric detection (PAD) system. A triple-pulse amperometric waveform (E_1 0.1 V, E_2 0.7 V, E_3 -0.1 V) was used for detection with the gold electrode.¹⁶ Analytical separations were performed on a CarboPac PA-100 column (250 x 4 mm, Dionex), using a linear gradient of 0 - 300 mM NaOAc in 100 mM NaOH (1 mL/min). Samples were fractionated on a CarboPac PA-1 column (250 x 9 mm, Dionex), using a linear gradient of 0 - 300 mM NaOAc in 100 mM NaOH (4 mL/min) or isocratic conditions of 100 mM NaOAc in 100 mM NaOH (4 mL/min). Collected fractions were immediately neutralised with 4 M HOAc, desalted on CarboGraph SPE columns (150 mg graphitised carbon, Alltech) using 1:3 acetonitrile:H₂O as eluent, and lyophilised.

Mass spectrometry

GLC-EI-MS was performed on a Fisons Instruments GC 8060/ MD 800 system (Interscience BV; Breda, The Netherlands) equipped with an AT-1 column (30 m x 0.25 mm, Alltech), using a temperature gradient of 140-240 °C at 4 °C/min.¹³

Matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) was carried out on a Voyager-DE Pro (Applied Biosystems; Nieuwerkerk aan de IJssel, The Netherlands) instrument in the reflector mode at an accelerating voltage of 24 kV, using an extraction delay of 90 ns, in a resolution of 5000 - 9000 FWHM. Samples (1 μ L) were mixed in a 1:1 ratio with a mixture of 7.5 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 1:1 acetonitrile:H₂O, and spectra were recorded in the positive-ion mode.

NMR spectroscopy

¹H NMR spectra, including ¹H-¹H and ¹³C-¹H correlation spectra, were recorded on a Bruker DRX500 spectrometer (Bijvoet Center, Department of NMR spectroscopy, Utrecht University) at a probe temperature of 300 K. Samples were exchanged once with 99.9 atom% D₂O, lyophilised, and dissolved in 650 μ L D₂O. ¹H chemical shifts (δ) are expressed in ppm by reference to internal acetone (δ 2.225), and ¹³C chemical shifts (δ) are expressed in ppm by reference to the methyl-carbon of internal acetone (δ 31.08). 1D ¹H NMR spectra were recorded with a spectral width of 5000 Hz in 16k complex data sets and zero filled to 32k. A WEFT pulse sequence was applied to suppress the HOD signal.¹⁷ When necessary, a fifth order polynomial baseline correction was applied. 2D

TOCSY spectra were recorded using MLEV17 mixing sequences with spin-lock times of 10, 60, 120, and 180 ms. The spin-lock field strength corresponded to a 90° pulse width of about 28 μ s at 13 dB. The spectral width in 2D TOCSY experiments was 4006 Hz at 500 MHz in each dimension. 400-1024 spectra of 2k data points with 8-32 scans per t_1 increment were recorded. 2D rotating-frame nuclear Overhauser enhancement spectra (ROESY) were recorded with 300 ms mixing time. The spectral width was 4006 Hz at 500 MHz in each dimension. Suppression of the HOD signal was performed by 1 s pre-saturation during the relaxation delay. Between 400 and 1024 data sets of 2k data points were recorded with 8-16 scans per t_1 increment. 2D ^{13}C - ^1H HSQC spectroscopy was carried out at a ^1H frequency of 500.0821 MHz and a ^{13}C frequency of 125.7552 MHz. Spectra were recorded with a spectral width of 4006 Hz for t_2 and 10 kHz for t_1 . The HOD signal was pre-saturated for 1 s, and ^{12}C -bound protons were suppressed using a TANGO pulse sequence. During acquisition of the ^1H FID, a ^{13}C decoupling pulse was applied. 128-256 experiments of 2k data points were recorded with 128 scans per t_1 increment. 2D NMR spectroscopic data were analysed by applying a sinus multiplication window and zero filling to spectra of 4k by 1k dimensions. In the case of ^{13}C - ^1H HSQC data, the spectra were zero filled to 4k by 512 data points. A Fourier transform was applied, and where necessary, a fifth to fifteenth order polynomial baseline function was applied. All NMR data were processed using in-house developed software (J.A. van Kuik, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University).

References

1. Sandford, P. A.; Baird, J. In *The Polysaccharides. Vol. 2.*; Aspinall, G. O., Ed.; Academic Press: New York, **1983**; pp 411-490.
2. Costerton, J. W.; Cheng, K.-J.; Geesey, G. G.; Ladd, T. I.; Nickel, J. C.; Dasgupta, M.; Marrie, T. J. *Annu. Rev. Microbiol.*, **1987**, *41*, 435-464.
3. Ceri, H.; McArthur, H. A. I.; Whitfield, C. *Infect. Immun.*, **1986**, *51*, 1-5.
4. De Vuyst, L.; Degeest, B. *FEMS Microbiol. Rev.*, **1999**, *23*, 153-177
5. Monchois, V.; Willemot, R.-M.; Monsan, P. *FEMS Microbiol. Rev.*, **1999**, *23*, 131-151.
6. Funane, K.; Ishii, T.; Matsushita, M.; Hori, K.; Mizuno, K.; Takahara, H.; Kitamura, Y.; Kobayashi, M. *Carbohydr. Res.*, **2001**, *334*, 19-25.
7. Argüello-Morales, M. A.; Remaud-Simeon, M.; Pizzut, S.; Sarçabal, P.; Willemon, R.-M.; Monsan, P. *FEMS Microbiol. Lett.*, **2000**, *182*, 81-85.
8. Van Leeuwen, S. S.; Leeftang, B. R.; Gerwig, G. J.; Kamerling, J. P. (Chapter 2) (to be published)
9. Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P. (Chapter 3) (to be published)
10. Kralj, S.; van Geel-Schutten, G. H.; Rahaoui, H.; Leer, R. J.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Appl. Env. Microbiol.*, **2002**, *68*, 4283-5291.
11. Bock, K.; Thøgersen, H. *Ann. Rep. NMR Spectr.* **1982**, *13*, 2-57.
12. Domań-Pytka, M.; Bardowski, J. *Crit. Rev. Microbiol.*, **2004**, *30*, 107-121.

13. Kamerling, J. P.; Vliegthart, J. F. G. In *Clinical Biochemistry – Principles, Methods, Applications. Vol. 1, Mass Spectrometry*; Lawson A. M., Ed.; Walter de Gruyter: Berlin, **1989**; pp 176-263.
14. Ciucanu, I.; Kerek, F. *Carbohydr. Res.* **1984**, *131*, 209-217.
15. Jansson, P.-E.; Kenne, L.; Liedgren, H.; Lindberg, B.; Lönnngren, J. *Chem. Commun. Univ. Stockholm*, **1976**, *8*, 1-74.
16. Lee, Y. C. *Anal. Biochem.* **1990**, *189*, 151-162.
17. Hård, K.; van Zadelhoff, G.; Moonen, P.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.*, **1992**, *209*, 895-915.

*In questions of science, the authority of a thousand is not worth the humble reasoning
of a single individual*

-Galileo Galilei-

Chapter 5

Structural analysis of the bio-engineered α -D-glucan (mEPS-PNNS) produced by a triple mutant of the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180

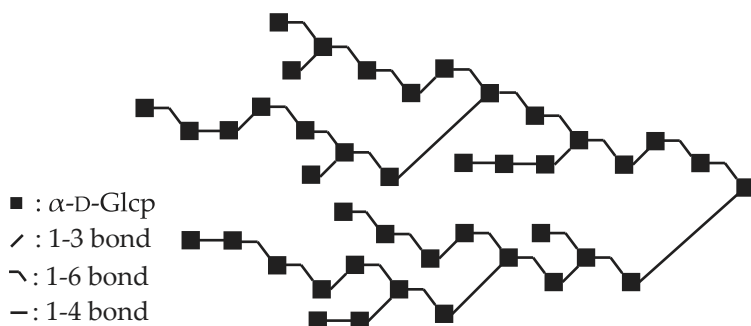
Sander S. van Leeuwen,¹ Slavko Kralj,^{2,3} Gerrit J. Gerwig,¹ Lubbert Dijkhuizen,^{2,3} and Johannes P. Kamerling¹

¹Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, 3584 CH Utrecht, The Netherlands

²Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands

³Centre for Carbohydrate Bioprocessing, TNO - University of Groningen, 9750 AA Haren, The Netherlands

Abstract—Site-directed mutagenesis of the glucansucrase GTF180 gene from *Lactobacillus reuteri* strain 180 was used to transform the active site region. The α -D-glucan (**mEPS-PNNS**) produced by the triple mutant V1027P:S1137N:A1139S differed in structure from that of the wild type α -D-glucan (EPS180). Besides (α 1-3) and (α 1-6) linkages, as present in **EPS180**, **mEPS-PNNS** also contained (α 1-4) linkages. Linkage analysis, periodate oxidation, and 1D/2D ^1H NMR spectroscopy of the intact **mEPS-PNNS**, as well as MS and NMR analysis of oligosaccharides obtained by partial acid hydrolysis of **mEPS-PNNS**, were combined to formulate the following visual representation of **mEPS-PNNS**, which includes all identified structural features.



Introduction

Lactic acid bacteria (LAB), like *Lactobacilli*, excrete exopolysaccharides (EPSs) into their surroundings. Exopolysaccharides have been found as adhesives,¹ participants in certain cellular recognition processes,² and as slime forming agents for protection against dehydration, phagocytosis or toxins.² The physical properties that are important for their *in vivo* functions, also make these polysaccharides suitable for the food and dairy industry. Bacterial exopolysaccharides have been used as thickeners, stabilisers, and gelling agents.³ To improve these properties, attention has been paid to the engineering of polysaccharide structures, via chemical⁴ or enzymatic derivatisations,⁵⁻⁷ and by genetic modification of source micro-organisms.^{8,9}

Recently, a family of glucansucrases was discovered in *Lb. reuteri*, which converts sucrose into large, heavily branched α -D-glucans. Structural analysis of the homopolysaccharides produced by the glucansucrases GTFA and GTF180 revealed highly complex structures, and visual representations of the structural elements have been proposed.^{10,11} Previous studies on the site-directed mutagenesis near the catalytic D1133 (putative transition state stabilising residue) of GTFA have shown that small mutations in the glucansucrase can give rise to large changes in the EPS structure and its properties.¹² Similar modifications in the homologous region of the GTF180 gene have been performed (unpublished data). The EPS of the triple-mutant V1027P:S1137N:A1139S (**mEPS-PNNS**) showed the most extensive changes in linkage distribution compared to the wild type polysaccharide (**EPS180**). Here, we report a detailed structural analysis of **mEPS-PNNS**, identifying the structural elements and their quantities, and finally formulating a visual representation containing these structural elements.

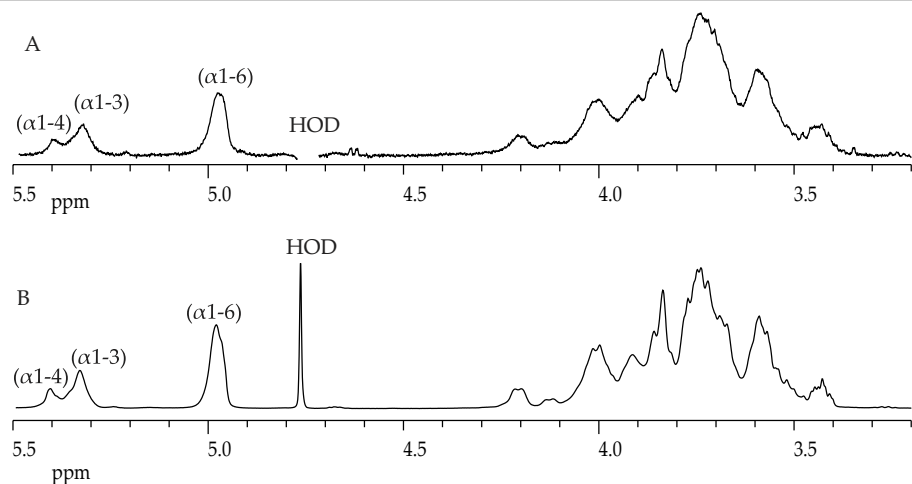


Figure 1. 500-MHz 1D ^1H NMR spectra of (A) **mEPS-PNNS** and (B) subpool **IV**, recorded at 300K in D_2O .

Results

Composition of **mEPS-PNNS**

Methylation analysis of **mEPS-PNNS** showed the presence of terminal, 3-substituted, 4-substituted, 6-substituted, and 3,6-disubstituted glucopyranose in a molar percentage of 18, 10, 12, 42, and 18%. 1D ^1H NMR spectroscopy of intact **mEPS-PNNS** (Figure 1A) indicated an α -anomeric configuration for all glucose residues. Integration of the anomeric signals revealed 60% (α 1-6)-linked D-Glcp ($\delta_{\text{H-1}} \sim 4.96$), 28% (α 1-3)-linked D-Glcp ($\delta_{\text{H-1}} \sim 5.33$), and 12% (α 1-4)-linked D-Glcp ($\delta_{\text{H-1}} \sim 5.39$) residues, which is in agreement with the linkage distribution found by methylation analysis. The low solubility of **mEPS-PNNS** influenced the quality of the 1D ^1H NMR spectrum. The broader peaks, compared to the 1D ^1H NMR spectrum of wild type **EPS180**,¹² may be the result of a more structural diversity.

Partial acid hydrolysis

A sample of **mEPS-PNNS** (800 mg) was incubated 10 times with 0.5 M TFA (30 min, 90 °C), with intermediate centrifugation and collection of supernatant. Analysis of the 10 supernatant samples by 1D ^1H NMR spectroscopy showed that the proton patterns all resembled that of intact **mEPS-PNNS** (Figure 1A), suggesting that all the structural elements present in **mEPS-PNNS** are also present in the hydrolysates. In each case, the linkage distribution, as determined from the 1D ^1H NMR spectra of the 10 batches, amounts to 59% (α 1-6) linkages, 28% (α 1-3) linkages, and 13% (α 1-4)

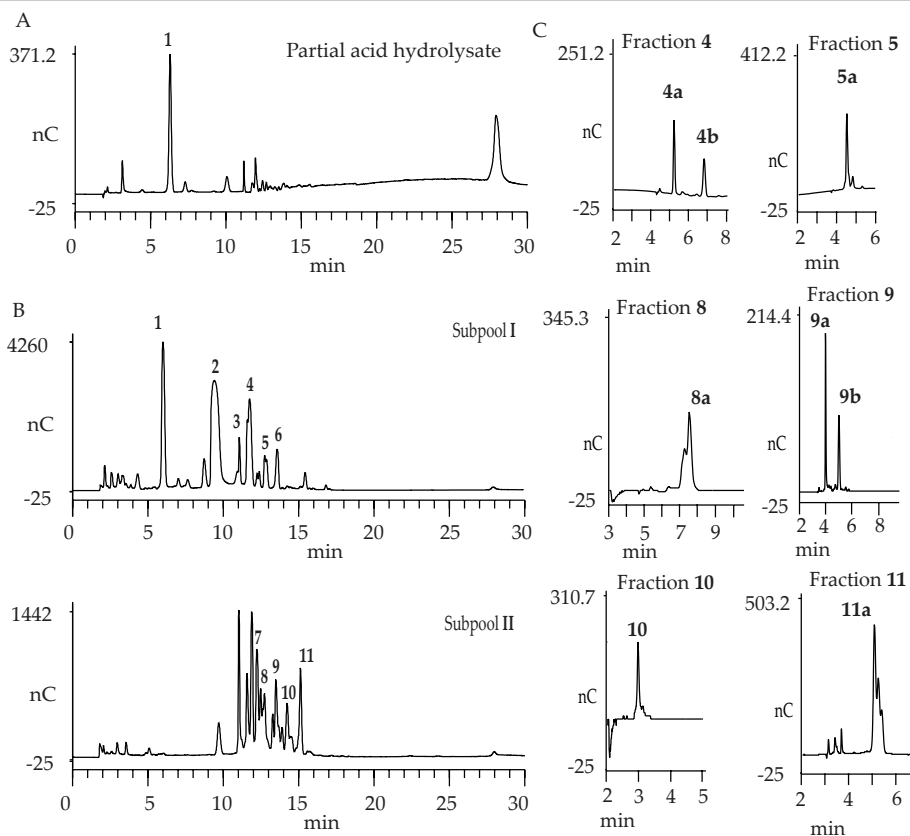
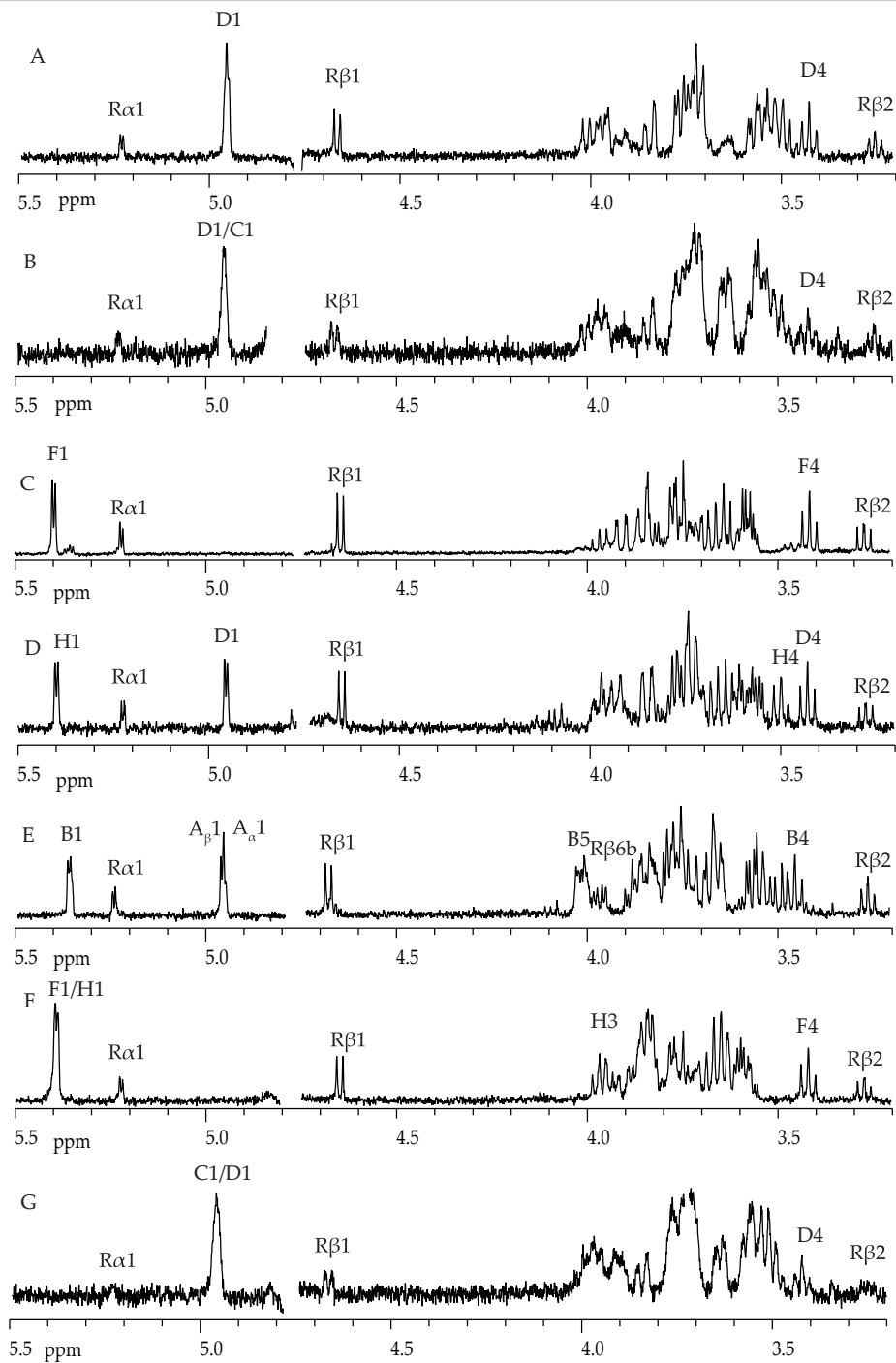


Figure 2. (A) HPAEC profile of *mEPS-PNNS* partial acid hydrolysate on CarboPac PA-100; (B) HPAEC profiles of Bio-Gel P-2 subpools I and II on CarboPac PA-100 using a linear gradient; and (C) HPAEC profiles of HPAEC fractions 4, 5, 8, 9, and 11 on CarboPac PA-100, using isocratic conditions. For experimental details, see Materials and Methods.

linkages. This indicates that, in contrast with wild type **EPS180**,¹⁰ all three linkage types are equally susceptible to acid hydrolysis under the conditions selected. For further investigations the 10 batches were pooled.

HPAEC profiling of the pool on CarboPac PA-100 (eluent: 0 - 300 mM NaOAc in 100 mM NaOH) showed free glucose (fraction 1) as the major component (Figure 2A). To obtain suitable fractions for further analysis, a pre-fractionation was performed on Bio-Gel P-2, yielding 16 overlapping fractions with a broad distribution of fragment sizes (data not shown). On the guidance of MALDI-TOF-MS analysis, the fractions were combined in four subpools, as follows: Fractions containing fragments with $DP \leq 3$ were combined in subpool I (Figure 2B), fractions containing mostly fragments with $DP4 - DP7$ in subpool II (Figure 2B), fractions containing mostly fragments with $DP8$



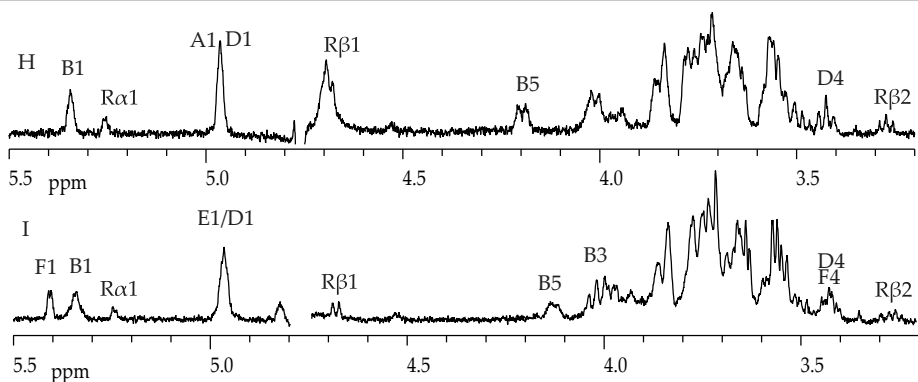


Figure 3. 500-MHz 1D ^1H NMR spectra of (A) fraction 2, (B) fraction 3, (C) fraction 4a, (D) fraction 4b, (E) fraction 5a, (F) fraction 6, (G) fraction 7, (H) fraction 8a, and (I) fraction 11a, recorded at 300K in D_2O . Structural-reporter-group signals that could be distinguished are indicated in the spectra. Labels correspond with those used in Table 1 and Scheme 1.

- DP11 in subpool **III**, and fractions containing fragments with DP > 11 in subpool **IV**. Subpools **I** and **II** were further separated on CarboPac PA-100, yielding fractions 1 to 6 and 7 to 11, respectively. Subpools **III** and **IV**, containing fragments too large for a complete ^1H NMR analysis were not separated.

Fraction 2 - The retention time of fraction 2 on CarboPac PA-100, as well as the 1D ^1H NMR spectrum (Figure 3A) were in agreement with the presence of isomaltose, i.e. α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).¹³

Fraction 3 - The MALDI-TOF mass spectrum of fraction 3 revealed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 527, corresponding with Hex₃. The 1D ^1H NMR spectrum (Figure 3B) matched that of isomaltotriose, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).¹³

Fraction 4 - MALDI-TOF-MS analysis of fraction 4 showed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 527, corresponding with Hex₃. However, the separation profile on CarboPac PA-100 (Figure 2B) showed two overlapping peaks. Therefore, fraction 4 was further separated on CarboPac PA-100, using 100 mM NaOAc in 100 mM NaOH as eluent (Figure 2C), and fractions 4a and 4b were isolated. The 1D ^1H NMR spectrum of compound 4a (Figure 3C) corresponded with that of maltose, i.e. α -D-Glcp-(1 \rightarrow 4)-D-Glcp.¹³ The 1D ^1H NMR spectrum of compound 4b (Figure 3D), corresponding with Hex₃ (MALDI-TOF-MS), indicated the presence of panose, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp (Scheme 1).¹³

Fraction 5 - MALDI-TOF-MS analysis of fraction **5** revealed $[M+Na]^+$ pseudomolecular ions at m/z 527 and 689, corresponding with Hex₃ and Hex₄, respectively. Fraction **5** was further separated on CarboPac PA-100 (Figure 2C), using isocratic conditions as described in 2.2.3, rendering one major fraction **5a** (Hex₃, MALDI-TOF-MS).

The 1D ¹H NMR spectrum of fraction **5a** (Figure 3E) showed five anomeric signals at δ 5.366 (**B** H-1, $^3J_{1,2}$ 3.4 Hz), 5.246 (**R α** H-1, $^3J_{1,2}$ 3.4 Hz), 4.962 (**A β** H-1, $^3J_{1,2}$ 3.4 Hz), 4.955 (**A α** H-1, $^3J_{1,2}$ 3.4 Hz), and 4.681 (**R β** H-1, $^3J_{1,2}$ 7.8 Hz). The set of **R α** and **R β** H-1 values corresponds with the occurrence of a -(1 \rightarrow 6)-D-Glcp unit in a -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp sequence (library data: -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp, δ 5.246-5.249 and 4.676-4.680; -(1 \rightarrow 4/6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp, δ 5.240-5.241 and 4.667-4.672).^{10,11,13} The presence of **R α** H-6b at δ 3.966 is in agreement with a 6-substituted residue **R**. The splitting of **A** H-1 is due to the influence of the α/β configuration of the reducing residue **R**. The **B** H-1 signal at δ 5.366 corresponds with that of an (-) α -D-Glcp-(1 \rightarrow 3)-* unit.¹³ The **B** H-4 signal at δ 3.451 was established as a structural-reporter-group signal of a terminal α -D-Glcp-(1 \rightarrow x)- unit,¹³ and has a surface area corresponding to 1 H, indicating a linear structure. These data lead to a **B1 \rightarrow 3A1 \rightarrow 6R** sequence for compound **5a**, i.e. α -D-Glcp-(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

Fraction 6 - MALDI-TOF-MS and ¹H NMR analysis (Figure 3F) of fraction **6** demonstrated the presence of maltotriose, i.e. α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp (Scheme 1).¹³

Fraction 7 - The MALDI-TOF mass spectrum of fraction **7** revealed an $[M+Na]^+$ pseudomolecular ion at m/z 689, corresponding with Hex₄. The 1D ¹H NMR spectrum (Figure 3G) matched that of isomaltotetraose, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp.¹⁰

Fraction 8 - MALDI-TOF-MS analysis of fraction **8** showed $[M+Na]^+$ pseudomolecular ions at m/z 527 and 689, corresponding with Hex₃ and Hex₄, respectively. Fraction **8** was further separated on CarboPac PA-100, using 100 mM NaOAc in 100 mM NaOH (Figure 2C), yielding major fraction **8a** (Hex₄, MALDI-TOF-MS).

* For Glc residues at semi-defined places in the structure (-) α -D-Glcp-(1 \rightarrow x)- or -(1 \rightarrow x)- α -D-Glcp(-) is used. When the structural context of the residue is precisely known, this is indicated as follows: -(1 \rightarrow x)- α -D-Glcp-(1 \rightarrow y)- describing an x-substituted residue with an (1-y) linkage. In case of a non-reducing terminal residue α -D-Glcp-(1 \rightarrow x)- is used, a reducing terminal residue is indicated with -(1 \rightarrow x)-D-Glcp.

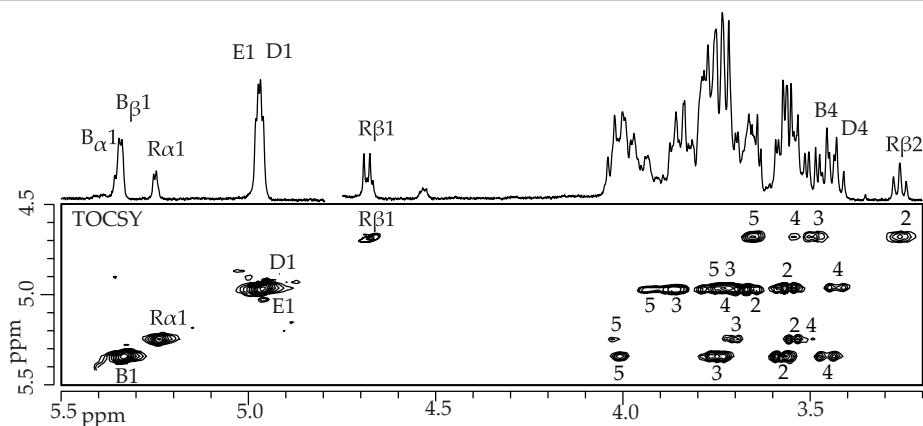


Figure 4. 500-MHz 1D ^1H NMR spectrum and the relevant part of the 2D ^1H - ^1H TOCSY spectrum (60 ms) of fraction **9a**, recorded at 300K in D_2O . Anomeric protons in the TOCSY spectrum ($\text{R}\alpha 1$, etc.) have been indicated on the diagonal; numbers in the horizontal tracks belong to the cross-peaks of the scalar-coupling network of the residues indicated.

The 1D ^1H NMR spectrum (Figure 3H) of fraction **8a** showed five anomeric signals at δ 5.345/5.336 (**B** H-1, $^3J_{1,2}$ 3.6 Hz), 5.249 (**R α** H-1, $^3J_{1,2}$ 3.6 Hz), 4.965 (**D** H-1, $^3J_{1,2}$ 3.6 Hz), 4.958 (**A** H-1, $^3J_{1,2}$ 3.6 Hz), and 4.680 (**R β** H-1, $^3J_{1,2}$ 7.8 Hz). Assignments of non-anomeric proton chemical shifts (Table 1) were obtained from 2D ^1H - ^1H TOCSY measurements (data not shown). In fact, the various NMR data match exactly with those obtained from **EPS180** compound **3b**, identified previously:¹⁰ **D1**→**6B1**→**3A1**→**6R**, i.e. α -D-Glcp-(1→6)- α -D-Glcp-(1→3)- α -D-Glcp-(1→6)-D-Glcp (Scheme 1).

Fraction 9 - Analysis of fraction **9** by MALDI-TOF-MS revealed $[\text{M}+\text{Na}]^+$ pseudomolecular ions at m/z 527 and 689, corresponding with Hex_3 and Hex_4 , respectively. Fraction **9** was further separated on CarboPac PA-100, using isocratic conditions described in 2.2.3 (Figure 2C), yielding a major fraction **9a** (Hex_4 ; MALDI-TOF-MS) and a minor fraction **9b** (Hex_3 ; MALDI-TOF-MS).

The 1D ^1H NMR spectrum of Hex_4 **9a** (Figure 4) showed six anomeric signals at δ 5.353 (**B α** H-1, $^3J_{1,2}$ 3.8 Hz), 5.343 (**B β** H-1, $^3J_{1,2}$ 3.8 Hz), 5.249 (**R α** H-1, $^3J_{1,2}$ 3.8 Hz), 4.975 (**E** H-1, $^3J_{1,2}$ 3.8 Hz), 4.965 (**D** H-1, $^3J_{1,2}$ 3.8 Hz), and 4.683 (**R β** H-1, $^3J_{1,2}$ 7.6 Hz). 2D ^1H - ^1H TOCSY measurements (Figure 4/60 ms) delivered most of the δ values of the non-anomeric protons (Table 1). The chemical shift positions of **R α** and **R β** H-1 at δ 5.249 and 4.683, respectively, correspond with the occurrence of a -(1→6)-D-Glcp unit in a -(1→3)- α -D-Glcp-(1→6)-D-Glcp sequence.^{10,11,13} This is further corroborated by the **R β** H-2 signal at δ 3.258, indicating that residue **R** is not 3- or 4-substituted.¹³ The

split **B** H-1 signal at δ 5.353/5.343 fits best with the occurrence of an $(-)\alpha$ -D-Glcp-(1 \rightarrow 3)- unit.^{10,13} The set of **B** H-4 and H-5 at δ 3.45 and 4.00, respectively, identified residue **B** as a terminal α -D-Glcp-(1 \rightarrow 3)- unit (see nigerose residue **B** and nigerotriose residue **C** in Ref. 13). Residue **D** H-1 at δ 4.965 is indicative of an $(-)\alpha$ -D-Glcp-(1 \rightarrow 6)- unit, and together with **D** H-4 at δ 3.43 a terminal α -D-Glcp-(1 \rightarrow 6)- unit is demonstrated.¹³ The occurrence of two terminal units indicates a branched structure for **9a**. Finally, the **E** H-1 signal at δ 4.975 revealed the occurrence of an $(-)\alpha$ -D-Glcp-(1 \rightarrow 6)- unit. Since residue **R** is not 3-substituted, the remaining internal residue **E** has to be a 3,6-disubstituted unit, yielding the sequence of **9a**: **B**1 \rightarrow 3[**D**1 \rightarrow 6]**E**1 \rightarrow 6**R**, i.e. α -D-Glcp-(1 \rightarrow 3)-[α -D-Glcp-(1 \rightarrow 6)-] α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

The 1D ^1H NMR spectrum of trisaccharide **9b** matched that of maltotriose, i.e. α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 4)-D-Glcp (Scheme 1).¹³

Fraction 10 - MALDI-TOF-MS analysis of fraction **10** revealed an $[\text{M}+\text{Na}]^+$ pseudomolecular ion at m/z 1013, corresponding with Hex₆. The 1D ^1H NMR spectrum (Figure 5) showed anomeric signals at δ 5.36 (**B α** H-1), 5.34 (**B β** H-1), 5.246 (**R α** H-1), 4.96-4.98 (**C^I** H-1, **D** H-1, **E** H-1), and 4.679 (**R β** H-1, $^3J_{1,2}$ 7.8 Hz). The chemical shift values of the non-anomeric protons were obtained from 2D ^1H - ^1H TOCSY measurements (Figure 5/60 ms). The **R α** and **R β** H-1 δ -values at 5.246 and 4.679 ppm, respectively, correspond with the presence of a -(1 \rightarrow 6)-D-Glcp unit, in a -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)-D-Glcp sequence.^{10,11,13} The **R β** H-2 signal at δ 3.259, indicating that residue **R** is not 3- or 4-substituted (library data: **R β** in nigerose $\delta_{\text{H-2}}$ 3.332, and **R β** in maltose $\delta_{\text{H-2}}$ 3.272), supports this conclusion.¹³ The set of **B** H-1, H-2, H-3, and H-4 values corresponds with that of **B** in **9a**, revealing the occurrence of a terminal α -D-Glcp-(1 \rightarrow 3)- unit. The presence of the set of **D** H-1 and H-4 at δ \sim 4.97 and 3.42, respectively, indicates the presence of a terminal α -D-Glcp-(1 \rightarrow 6)- unit. The presence of two terminal units implicates a branched structure for compound **10** (Hex₆). Since the H-1 signal at δ 4.96-4.98, typical for $(-)\alpha$ -D-Glcp-(1 \rightarrow 6)- residues, corresponds with 4 H, and residue **R** is not 3-substituted, the structure has to contain an internal 3,6-disubstituted residue **E** (compare with **9a**). Then, the remaining two residues can only be internal α -D-Glcp-(1 \rightarrow 6)- units, denoted residue **C^I** and **C^{II}**. Since the **B** H-1 signal is split, due to the influence of the reducing residue α/β configuration, the most probable sequence for compound **10** is **D**1 \rightarrow 6**C^{II}**1 \rightarrow 6**C^I**1 \rightarrow 6[**B**1 \rightarrow 3]**E**1 \rightarrow 6**R**, i.e. α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)[α -D-Glcp-(1 \rightarrow 3)]- α -D-Glcp-(1 \rightarrow 6)-D-Glcp (Scheme 1).

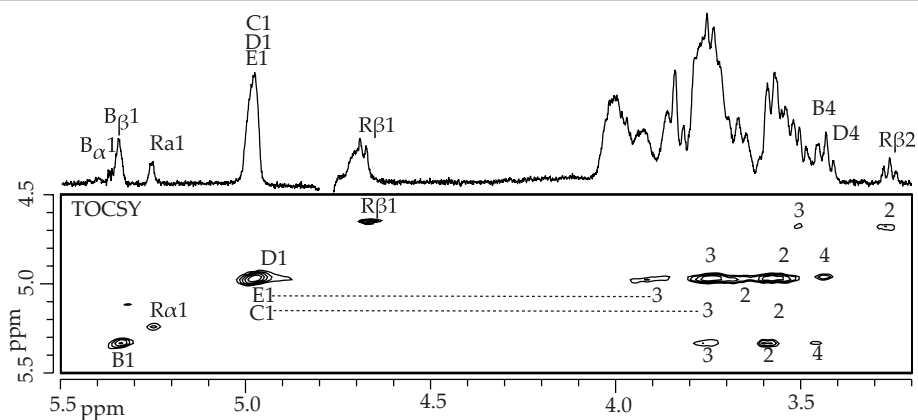
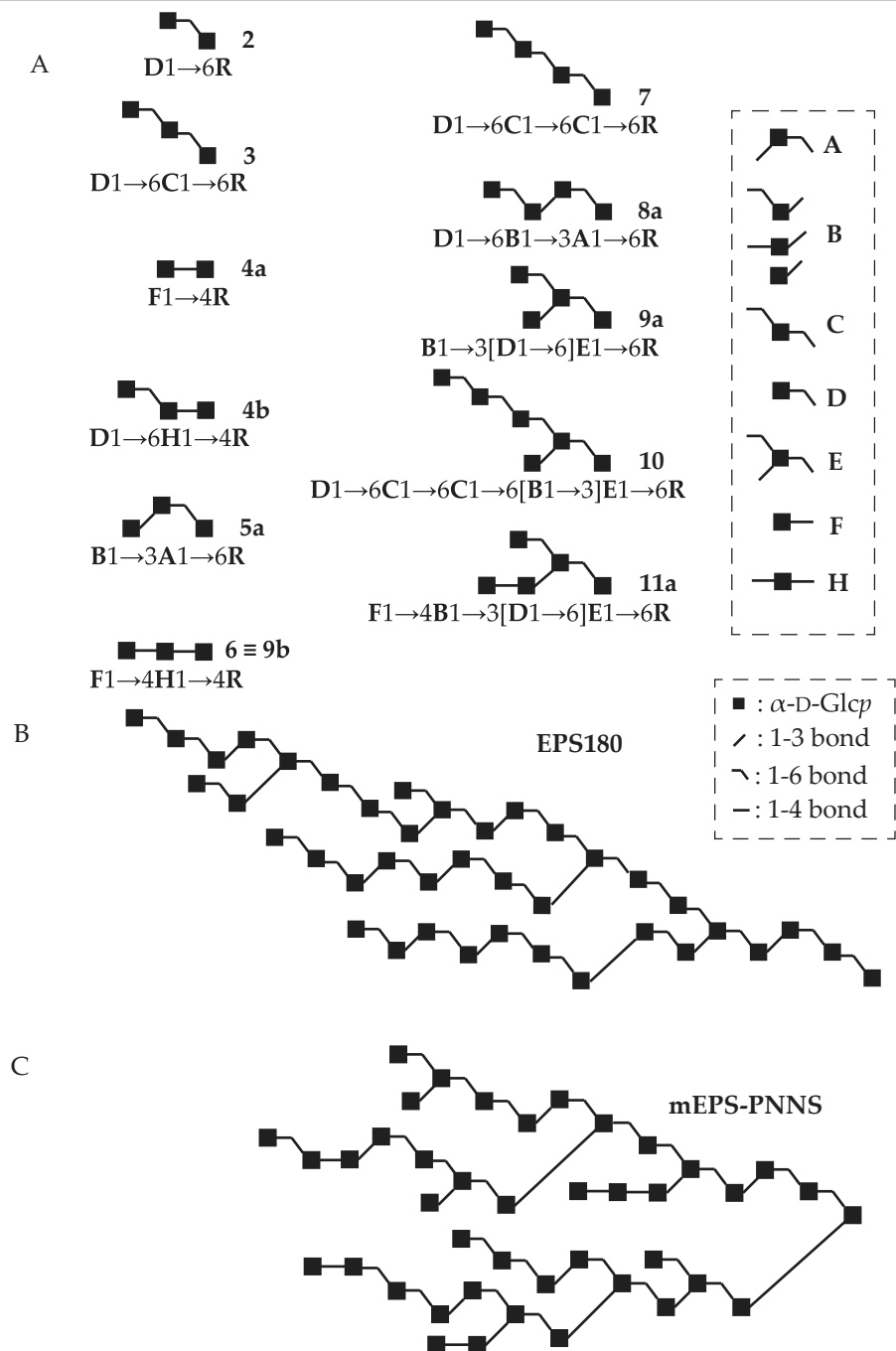


Figure 5. 500-Hz 1D ^1H NMR spectrum and the relevant part of the 2D ^1H - ^1H TOCSY spectrum (60 ms) of fraction **10**, recorded at 300K in D_2O . For an explanation of the coding system, see Figure 4.

Fraction 11 - MALDI-TOF-MS analysis of fraction **11** showed $[\text{M}+\text{Na}]^+$ pseudomolecular ions at m/z 851 and 1013, corresponding with Hex_5 and Hex_6 , respectively. Fraction **11** was further separated on CarboPac PA-100, using isocratic conditions described in 2.2.3 (Figure 2C), yielding one major fraction **11a** (Hex_5 , MALDI-TOF-MS).

The 1D ^1H NMR spectrum of fraction **11a** (Figure 3I) showed anomeric signals at δ 5.410 (**F** H-1, $^3J_{1,2}$ 3.6 Hz; (α 1-4) linkage), 5.35-5.36 (**B** H-1; (α 1-3) linkage), 5.247 (**R α** H-1), 4.97-4.99 (**D** H-1, **E** H-1; (α 1-6) linkages), and 4.681 (**R β** H-1, $^3J_{1,2}$ 7.6 Hz). The H-1 α and H-1 β chemical shifts of the reducing residue **R** at δ 5.247 and 4.681, respectively, correspond with the occurrence of a $-(1\rightarrow6)\text{-D-Glcp}$ unit in a $-(1\rightarrow3)\text{-}\alpha\text{-D-Glcp}$ -(1 \rightarrow 6)-D-Glcp sequence.^{10,11,13} The occurrence of the signal at $\delta_{\text{H-4}}$ 3.42 (dd, 2 H), being indicative for an $\alpha\text{-D-Glcp}$ -(1 \rightarrow x)- unit,¹³ reflects a branched structure. As an H-5 signal at δ 4.20 is missing, the occurrence of a $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp}$ -(1 \rightarrow 3)- unit can be excluded.¹⁰ However, a signal at δ 4.12 is observed with a characteristic peak-shape of an H-5 signal, significantly downfield from H-5 of an $\alpha\text{-D-Glcp}$ -(1 \rightarrow 3)- unit (see also residue **B** in **9a**),¹³ which can be explained by the effect of a 4-substitution on an $\alpha\text{-D-Glcp}$ -(1 \rightarrow 3)- unit ($\Delta\delta_{\text{H-5}}$ +0.10-0.14 ppm, Ref. 13), leading to the presence of a $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp}$ -**B**-(1 \rightarrow 3)- unit. This observation adds a new structural-reporter-group signal to the concept.^{10,11,13} These data lead to a structural element **F1** \rightarrow **4B1** \rightarrow **3E1** \rightarrow **6R**. The final residue **D** has to be positioned to cause a branched structure, which can only occur at residue **E**, resulting in an **F1** \rightarrow **4B1** \rightarrow **3[D1** \rightarrow **6]E1** \rightarrow **6R** sequence for compound **11a**, i.e. $\alpha\text{-D-Glcp}$ -(1 \rightarrow 4)- $\alpha\text{-D-Glcp}$ -(1 \rightarrow 3)-[$\alpha\text{-D-Glcp}$ -(1 \rightarrow 6)]- $\alpha\text{-D-Glcp}$ -(1 \rightarrow 6)-D-Glcp (Scheme 1).



Scheme 1. (A) Structures of oligosaccharides obtained by partial acid hydrolysis, visual representations that include all identified structural elements (B) of wild type EPS180, and (C) of mEPS-PNNS.

Smith degradation

In order to investigate the degree of polymerisation of (α 1-3) glycosidic bonds, **mEPS-PNNS** was subjected to a Smith degradation, comprising a periodate oxidation, followed by reduction with NaBH_4 , and mild acid hydrolysis with formic acid. The formed products were analysed by GLC-EI-MS and HPAEC-PAD. In view of the linkage analysis (see section 2.1), α -D-Glc p -(1 \rightarrow 1)-Gro, and [α -D-Glc p -(1 \rightarrow 3)] $_n$ - α -D-Glc p -(1 \rightarrow 1)-Gro, but also, due to excessive hydrolysis, Gro, D-Glc, and [α -D-Glc p -(1 \rightarrow 3)] $_n$ -D-Glc can be expected.

GLC-EI-MS analysis (Figure 6A) of the trimethylsilylated residue showed three major products: Gro, D-Glc p , and α -D-Glc p -(1 \rightarrow 1)-Gro. Typical fragments for α -D-Glc p -(1 \rightarrow 1)-Gro are m/z 451 (aA_1), 361 (aA_2), 271 (aA_3), 217, and 204 for the Glc moiety, and m/z 219 (bA_1) and 337 (abJ_1) for the Gro moiety.^{14,15,16,17}

HPAEC analysis on CarboPac PA-100 (Figure 6B) of the residue revealed two major peaks that could be related to Gro (R_t 2.2 min) and α -D-Glc p -(1 \rightarrow 1)-Gro (R_t 6.2 min), which was identified by its elution position, being slightly later than D-Glc p (R_t 5.6 min). Since α -D-Glc p -(1 \rightarrow 3)-D-Glc p under the same conditions has an R_t value $>$ 10 min, the presence of [α -D-Glc p -(1 \rightarrow 3)] $_n$ - α -D-Glc p -(1 \rightarrow 1)-Gro with $n=1$ or higher could be excluded.

The absence of larger structures than glucosyl-glycerol indicates that the **mEPS-PNNS** structure does not contain two or more consecutive (α 1-3) linkages, similar to the wild type **EPS180**.¹⁰

2D NMR analysis of **mEPS-PNNS**, using subpool IV

Intact **mEPS-PNNS** does not dissolve in suitable amounts to allow 2D NMR analysis. As the ^1H NMR spectrum of subpool **IV**, isolated from a partial hydrolysate of **mEPS-PNNS** and containing fragments of DP $>$ 11 is similar to that of intact **mEPS-PNNS** (Figure 1A and 1B), this subpool was subjected to 2D NMR analysis. Especially, the larger oligosaccharides in subpool **IV** are suitable to represent all the structural elements that are present in the full length polysaccharide and do not have interfering signals from reducing residues. 2D ^1H - ^1H TOCSY experiments with increasing mixing-times (10, 30, 60, 120, and 150 ms) were interpreted (Figure 7; 60 and 120 ms) to unravel the structural elements present in **mEPS-PNNS**. In several cases, use is made of the ^1H NMR data collected from the oligosaccharides obtained by partial acid hydrolysis of **mEPS-PNNS** (section 2.2), as well as from data reported in Refs. 10, 11, and 13.

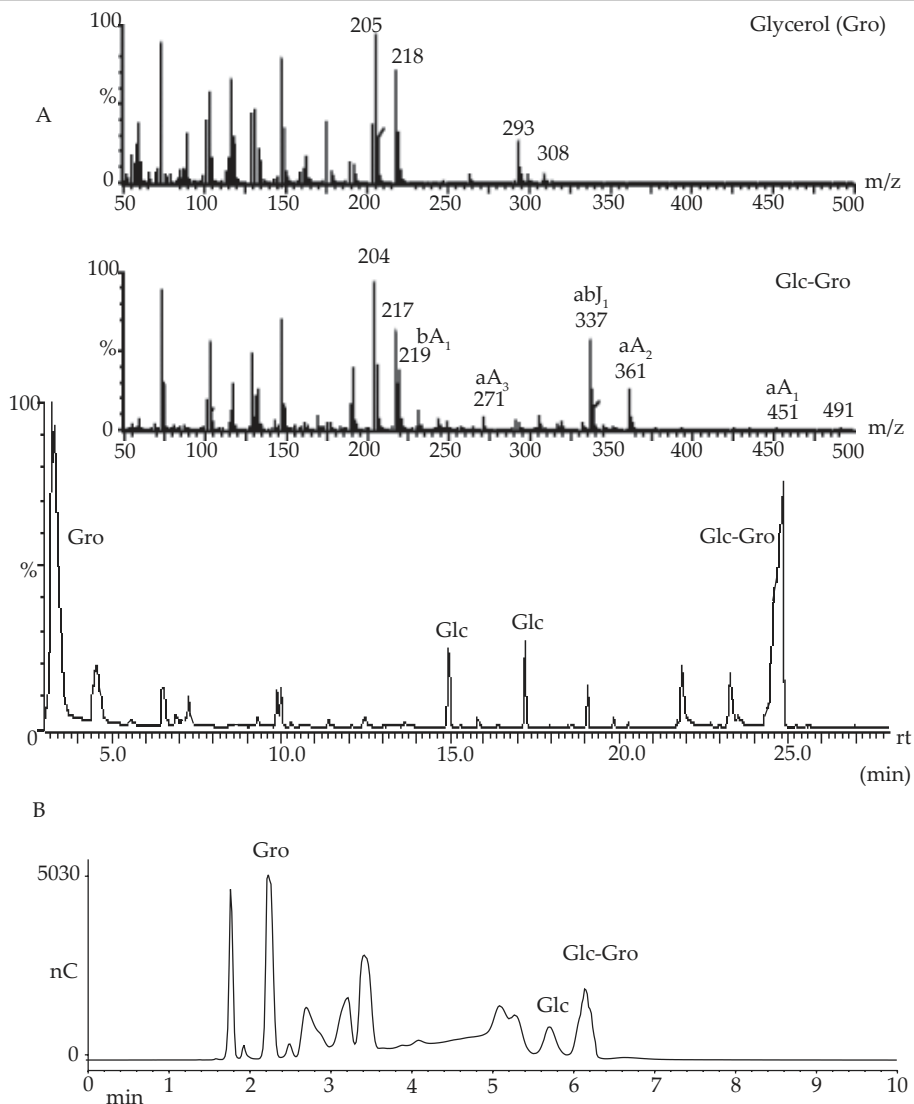


Figure 6. (A) GLC-EI-MS analysis and (B) HPAEC-PAD analysis of fragments obtained by Smith degradation of mEPS-PNNS.

So far (Refs. 11, 13, and this study), depending on the microenvironment of the unit, the δ value of the H-1 signal of a $(-)\alpha\text{-D-Glcp-(1}\rightarrow\text{4)-}$ unit varies between 5.41 and 5.33 ppm. In a similar way, the δ value of the H-1 signal of a $(-)\alpha\text{-D-Glcp-(1}\rightarrow\text{3)-}$ unit was found to lie between 5.39 and 5.32 ppm (Refs. 10, 13, and this study). This means that the anomeric signal of $(\alpha\text{1-4})$ -linked residues could overlap with the anomeric signal of $(\alpha\text{1-3})$ -linked residues. In the 1D ^1H NMR spectrum of subpool **IV** (Figure 1B), the H-1 signal between δ 5.41 and 5.39 has a surface area matching the

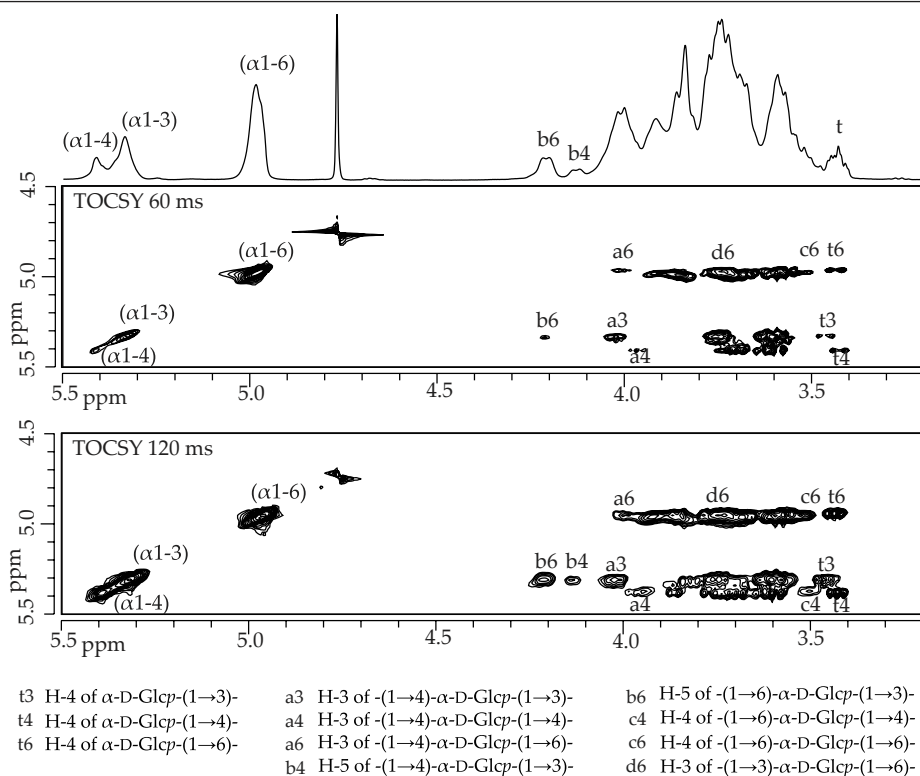


Figure 7. 500-MHz 1D ^1H NMR spectrum and the relevant parts of the 2D ^1H - ^1H TOCSY spectra (upper 60 ms, lower 120 ms) of subpool **IV**, recorded at 300K in D_2O . Anomeric signals and structural-reporter-group signals have been indicated with labels; the legend to the labels used is included.

amount of $(\alpha 1\text{-}4)$ -linked residues, as determined by methylation analysis (section 2.1), and can therefore be considered as the $(\alpha 1\text{-}4)$ -anomeric signal. This means that the peak between δ 5.39 and 5.32 corresponds with $(\alpha 1\text{-}3)$ -linked residues. The peak at δ 4.96 is a typical structural reporter for $(-)\alpha\text{-D-Glcp-(1}\rightarrow 6)$ - units (δ -range between 4.99 and 4.95 ppm, Refs. 10, 11, 13, this study).

In the 60 ms TOCSY spectrum of subpool **IV** (Figure 7), the H-1 track of $(-)\alpha\text{-D-Glcp-(1}\rightarrow 3)$ - residues (δ 5.39-5.32) showed a characteristic H-3 signal at δ 4.01, reflecting the occurrence of $(-)(1\rightarrow 4)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 3)$ - units (compare with $(-)(1\rightarrow 4)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 4)$ -, $\delta_{\text{H-3}}$ 3.96; $(-)(1\rightarrow 4)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 6)$ -, $\delta_{\text{H-3}}$ 4.01).^{11,13} This is further confirmed by the H-5 resonance at δ 4.12 (see compound **11a**) in the 120 ms TOCSY spectrum (Figure 7). On the same track in the 120 ms TOCSY spectrum another H-5 resonance at δ 4.20 is observed, corresponding with $(-)(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow 3)$ - units.¹⁰

Table 1. ^1H chemical shifts of oligosaccharide fragments of **mEPS-PNNS** obtained by partial acid hydrolysis. Residue labels correspond with those used in Scheme 1.

Residue	8a	9a	10
R α -1	5.249	5.249	5.246
R α -2	3.54	3.54	n.d.
R α -3	3.71	3.71	n.d.
R α -4	3.51	3.52	n.d.
R α -5	4.00	4.01	n.d.
R β -1	4.680	4.683	4.679
R β -2	3.258	3.258	3.259
R β -3	3.49	3.49	n.d.
R β -4	3.52	3.51	n.d.
R β -5	3.65	3.65	n.d.
A-1	4.958	-	-
A-2	3.65	-	-
A-3	3.84	-	-
A-4	3.67	-	-
B-1	5.345/336	5.353/343	5.36/34
B-2	3.58	3.57	3.58
B-3	3.75	3.75	3.75
B-4	3.51	3.45	3.45
B-5	4.20	4.00	n.d.
C-1	-	-	~4.97
C-2	-	-	3.57
C-3	-	-	3.74
D-1	4.965	4.965	~4.97
D-2	3.55	3.55	3.57
D-3	3.75	3.75	3.74
D-4	3.42	3.43	3.42
D-5	n.d.	3.76	n.d.
E-1	-	4.975	~4.97
E-2	-	3.65	3.65
E-3	-	3.86	3.88
E-4	-	3.76	n.d.
E-5	-	3.92	n.d.

The presence of terminal α -D-Glcp-(1 \rightarrow 3)- units is indicated by the presence of the signal at δ 3.45 (H-4), an established structural reporter for terminal residues.¹³ The absence of $\delta_{\text{H-3}} \sim 3.90$ on this track, indicates that -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 3)- units do not occur in **mEPS-PNNS**.¹³

On the H-1 track of the (-) α -D-Glcp-(1 \rightarrow 4)- residues in the 60 ms TOCSY spectrum (Figure 7) a weak H-3 resonance was detected at δ 3.96, indicating a minor occurrence of -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 4)- units (see residue **B** in maltotriose,¹³ and residue **H** in **6** and **9b**). A stronger H-3 signal at δ 3.70 indicates a major occurrence

of $-(1\rightarrow6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ and/or terminal $\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ units. The structural-reporter-group signal for terminal residues at $\delta_{\text{H-4}}$ 3.42 is also observed on this track, confirming the presence of $\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ units. The occurrence of minor amounts of $-(1\rightarrow6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ is reflected by the weak H-4 resonance at δ 3.50, observed in the 120 ms TOCSY spectrum (Figure 7) (see also residue **B** in panose,¹³ and residue **D** in **4b**). The absence of $\delta_{\text{H-3}} \sim 3.90$ (see residue **B** in nigerotriose)¹³ suggests that $-(1\rightarrow3)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ units do not occur in **mEPS-PNNS**.

On the $(-)\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ H-1 track in the 30 ms TOCSY spectrum (data not shown) $\delta_{\text{H-3}}$ -values were observed at 3.75, 3.85, and 4.00 ppm, respectively. The first value corresponds with the occurrence of $\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ and/or $-(1\rightarrow6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ units,¹³ the second indicates the occurrence of $-(1\rightarrow3)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ and/or $-(1\rightarrow3,6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ units,¹⁰ and the third is indicative for the presence of $-(1\rightarrow4)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ units (compare residue **B** in compounds **3a**, **5a**, **6a**, and **7a** in Ref. 11). The occurrence of terminal $\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ is further supported by the presence of an H-4 resonance at δ 3.43 on the $(\alpha 1\text{-}6)\text{-}$ anomeric track. The $-(1\rightarrow6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ moiety was shown to occur by the presence of $\delta_{\text{H-4}}$ 3.50 on the same track.^{10,13}

Discussion and conclusions

Discussion

Building blocks and their quantities were determined by combining data obtained from methylation analysis (section 2.1), Smith degradation analysis (section 2.3), and ¹H NMR spectroscopy of intact **mEPS-PNNS**, as well as from MS and ¹H NMR analysis of oligosaccharides obtained by partial acid hydrolysis of **mEPS-PNNS** (section 2.2).

As stated already (section 2.1), subpool **IV** contained 28% $(\alpha 1\text{-}3)\text{-}$ linked D-Glc_p. From integration of the H-5 signals of $-(1\rightarrow6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow3)\text{-}$ (δ 4.20) and $-(1\rightarrow4)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow3)\text{-}$ (δ 4.12) in the 1D ¹H NMR spectrum, the occurrence of these building blocks was quantified at 17 and 7%, respectively (Table 2). This leaves 4% for the remaining $\alpha\text{-D-Glc}p\text{-}(1\rightarrow3)\text{-}$ unit (Table 2).

The original structural limitation of single $(\alpha 1\text{-}3)$ bridges in **EPS180**¹⁰ exists also in **mEPS-PNNS**. Since $(-)\alpha\text{-D-Glc}p\text{-}(1\rightarrow4)\text{-}$ units were not 3-substituted, all 3-substituted Glc_p residues (10%, section 2.1) are in fact $-(1\rightarrow3)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$. Furthermore, methylation analysis indicated 18% 3,6-disubstituted Glc_p, which means 18% of all residues are $-(1\rightarrow3,6)\text{-}\alpha\text{-D-Glc}p\text{-}(1\rightarrow6)\text{-}$ (Table 2).

Table 2. Percentages of different building blocks present in **mEPS-PNNS** and wild type **EPS180**.¹⁰

Building block	mEPS-PNNS	EPS180
α -D-Glcp-(1→3)-	4	0
-(1→4)- α -D-Glcp-(1→3)-	7	0
-(1→6)- α -D-Glcp-(1→3)-	17	31
α -D-Glcp-(1→4)-	6	0
-(1→4)- α -D-Glcp-(1→4)-	3	0
-(1→6)- α -D-Glcp-(1→4)-	3	0
α -D-Glcp-(1→6)-	8	12
-(1→3)- α -D-Glcp-(1→6)-	10	19
-(1→4)- α -D-Glcp-(1→6)-	2	0
-(1→6)- α -D-Glcp-(1→6)-	22	26
-(1→3,6)- α -D-Glcp-(1→6)-	18	12

In accordance with 18% branching, also 18% terminal residues do occur (section 2.1). As mentioned above, 4% of these residues were α -D-Glcp-(1→3)- units. Since the H-4 resonance at δ 3.42-3.43, indicative for terminal residues,^{10,11,13} on the (α 1-6)-anomeric track was stronger than on the (α 1-4)-anomeric track, the occurrence of α -D-Glcp-(1→6)- and α -D-Glcp-(1→4)- will be around 8 and 6%, respectively.

With a total of 12% (α 1-4)-linked residues (section 2.1) and the estimated 6% for α -D-Glcp-(1→4)- units, an amount of 6% is left for -(1→4)- α -D-Glcp-(1→4)- and -(1→6)- α -D-Glcp-(1→4)-. Methylation analysis showed 12% 4-substituted Glcp. Taking into account that already 7% -(1→4)- α -D-Glcp-(1→3)- units were assigned (see above), 5% is left for both -(1→4)- α -D-Glcp-(1→4)- units and -(1→4)- α -D-Glcp-(1→6)- units. Based on TOCSY cross-peak intensities, the amounts -(1→4)- α -D-Glcp-(1→4)- (H-3), -(1→6)- α -D-Glcp-(1→4)- (H-4), and -(1→4)- α -D-Glcp-(1→6)- (H-3) units will be about 3, 3, and 2%, respectively (Table 2).

With the building blocks that could be determined from the available ¹H NMR and methylation analysis data, only one possible building block is left undetermined, i.e. -(1→6)- α -D-Glcp-(1→6)- units, which can then be determined indirectly (22%) from the methylation analysis.

Conclusions

Using the building block quantities and the sequences found for the compounds in section 2.2, a visual representation of **mEPS-PNNS**, that includes all identified structural features, was formulated as presented in Scheme 1. For comparison the earlier generated representation of wild type **EPS180**¹⁰ is included in Scheme 1, whereas the percentages of building blocks are given in Table 2.

The resulting picture shows a complex structure for **mEPS-PNNS** compared to wild type **EPS180**. The presence of (α 1-4) di- and trisaccharide elements shows that the mutations in the GTF180 enzyme have significantly changed its structural selectivity. However, the original structural limitation of single (α 1-3) bridges still exists. The presence of (α 1-6)-linked elements up to five residues in a row (compound **10**) indicates that the isomalto-oligosaccharide basis of the structure also still exists. The maximum length of (α 1-6)-linked stretches was not determined in wild type **EPS180**, and also cannot be determined in the case of **mEPS-PNNS**. The different linkage distribution in the **mEPS-PNNS** structure puts a limit on the maximum length of the (α 1-6)-linked elements, however, it must be shorter than that in wild type **EPS180**. The **mEPS-PNNS** polysaccharide is less soluble than the wild type **EPS180**. The lower solubility is probably the result of a more rigid structure, since (α 1-6) linkages are flexible, whereas (α 1-3) and (α 1-4) are more rigid.¹⁸ Increased rigidity of the structure may also cause the higher resistance to acid hydrolysis that was observed.

Materials and Methods

Materials

mEPS-PNNS was produced by incubation of the GTF180 (V1027P:S1137N:A1139S) mutant enzyme from *Lb. reuteri* strain 180 with 146 mM sucrose, containing 1% Tween 80 and 0.02% sodium azide, for 7 days. The **mEPS-PNNS** produced was isolated by precipitation with ethanol, as described previously.¹⁹

Linkage analysis

A polysaccharide sample was permethylated using CH_3I and solid NaOH in DMSO, as described earlier.²⁰ After hydrolysis with 2 M TFA (2 h, 120 °C), the partially methylated monosaccharides were reduced with NaBD_4 (2 h, room temperature). Conventional work-up, involving neutralisation with HOAc and removal of boric acid by co-evaporation with MeOH, followed by acetylation with 1:1 acetic anhydride:pyridine (3 h, 120 °C), yielded a mixture of partially methylated alditol acetates, which was analysed by GLC-EI-MS.^{15,21}

Partial acid hydrolysis

A sample of **mEPS-PNNS** (800 mg) was treated with 0.5 M TFA (2 mL) for 30 min at 90 °C. After centrifugation (1500 *g*, 5 min) the supernatant was collected, and

the pellet was treated again with 0.5 M TFA under the same conditions. This procedure was repeated 10 times. Each supernatant was investigated by 1D ^1H NMR spectroscopy. Subsequently, the supernatant samples were pooled, profiled on CarboPac PA-100 and separated on Bio-Gel P-2 (400 x 15 mm, BioRad), eluted with 25 mM NH_4HCO_3 ; 1.2-mL fractions were collected at a flow rate of 11.5 mL/h. Fractions were tested for the presence of carbohydrates by a TLC spot-test with orcinol/ H_2SO_4 staining. Carbohydrate-containing fractions were analysed by MALDI-TOF-MS.

Smith degradation

A sample of **mEPS-PNNS** (10 mg) was incubated with 2 mL 50 mM sodium periodate in 0.1 M NaOAc (pH 4.3) for 96 h at 4 °C in the dark. Then, the excess of periodate was destroyed by addition of 0.2 mL ethylene glycol. The oxidised polysaccharide solution was dialysed against tap water (24 h, room temperature), treated with excess NaBH_4 (18 h, room temperature), and subsequently neutralised with 4 M HOAc.²² After co-evaporation of boric acid with MeOH, the residue was hydrolysed with 90% HCOOH (30 min, 90 °C). Finally, the solution was concentrated under a stream of N_2 , and the products were analysed by GLC-EI-MS and HPAEC-PAD.

HPAEC-PAD

High-pH anion-exchange chromatography was performed on a Dionex DX500 workstation, equipped with an ED40 pulsed amperometric detection (PAD) system. A triple-pulse amperometric waveform (E_1 0.1 V, E_2 0.7 V, E_3 -0.1 V) was used for detection with the gold electrode.²³ Analytical separations were performed on a CarboPac PA-100 column (250 x 4 mm, Dionex), using a linear gradient of 0 - 300 mM NaOAc in 100 mM NaOH (1 mL/min). Samples were fractionated on a CarboPac PA-100 column (250 x 9 mm, Dionex), using a linear gradient of 0 - 300 mM NaOAc in 100 mM NaOH (4 mL/min) or isocratic conditions of 100 mM NaOAc in 100 mM NaOH (4 mL/min). Collected fractions were immediately neutralised with 4 M HOAc, desalted on CarboGraph SPE columns (150 mg graphitised carbon, Alltech) using 1:3 acetonitrile: H_2O as eluent, and lyophilised.

Mass spectrometry

GLC-EI-MS was performed on a Fisons Instruments GC 8060/ MD 800 system (Interscience BV; Breda, The Netherlands) equipped with an AT-1 column (30 m x 0.25 mm, Alltech), using a temperature gradient of 140-240 °C at 4 °C/min.¹⁵

Matrix-assisted laser desorption ionisation time-of-flight mass spectrometry (MALDI-TOF-MS) was carried out on a Voyager-DE Pro (Applied Biosystems; Nieuwerkerk aan de IJssel, The Netherlands) instrument in the reflector mode at an accelerating voltage of 24 kV, using an extraction delay of 90 ns, in a resolution of 5000-9000 FWHM. Samples (1 μ L) were mixed in a 1:1 ratio with a mixture of 7.5 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 1:1 acetonitrile:H₂O, and spectra were recorded in the positive-ion mode.

NMR spectroscopy

1D/2D ¹H NMR spectra were recorded on a Bruker DRX500 spectrometer (Bijvoet Center, Department of NMR spectroscopy, Utrecht University) at a probe temperature of 300K. Samples were exchanged once with 99.9 atom% D₂O, lyophilised, and dissolved in 650 μ L D₂O. ¹H chemical shifts (δ) are expressed in ppm by reference to internal acetone (δ 2.225). 1D ¹H NMR spectra were recorded with a spectral width of 5000 Hz in 16k complex data sets and zero filled to 32k. A WFT pulse sequence was applied to suppress the HOD signal.²⁴ When necessary, a fifth order polynomial baseline correction was applied. 2D TOCSY spectra were recorded using MLEV17 mixing sequences with spin-lock times of 10, 30, 60, 120, and 150 ms. The spin-lock field strength corresponded with a 90° pulse width of about 28 μ s at 13 dB. The spectral width in 2D TOCSY experiments was 4006 Hz at 500 MHz in each dimension. 400-1024 spectra of 2k data points with 8-32 scans per t₁ increment were recorded. 2D NMR spectroscopic data were analysed by applying a sinus multiplication window and zero filling to spectra of 4k by 1k dimensions. A Fourier transform was applied, and where necessary, a fifth to fifteenth order polynomial baseline function was applied. All NMR data were processed using in-house developed software (J.A. van Kuik, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University).

References

1. Costerton, J. W.; Cheng, K.-J.; Geesey, G. G.; Ladd, T. I.; Nickel, J. C.; Dasgupta, M.; Marrie, T. J. *Annu. Rev. Microbiol.*, **1987**, *41*, 435-464.
2. Ceri, H.; McArthur, H. A. I.; Whitfield, C. *Infect. Immun.*, **1986**, *51*, 1-5.
3. Sandford, P. A.; Baird, J. In *The Polysaccharides. Vol. 2*; Aspinall, G. O., Ed.; Academic Press: New York, **1983**; pp 411-490.
4. Sutherland, I. W. *Pure & Appl. Chem.* **1997**, *69*, 1911-1917.
5. Staaf, M.; Yang, Z.; Huttunen, E.; Widmalm, G. *Carbohydr. Res.*, **2000**, *236*, 113-119.
6. Skjåk-Bræk, G.; Smidsrod, O.; Larsen, B. *Int. J. Biol. Macromol.*, **1986**, *8*, 330-336.
7. McCleary, B. V.; Dea, I. C. M.; Windust, J.; Cooke, D. *Carbohydr. Polym.*, **1984**, *4*, 253-270.
8. Breedveld, M.; Bonting, K.; Dijkhuizen, L. *FEMS Microbiol. Lett.*, **1998**, *169*, 241-249.

9. Stingle, F.; Neeser, J.-R.; Mollet, B. *J. Bacteriol.*, **1996**, *178*, 1680-1690.
10. Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P., (Chapter 3) (To be published).
11. Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P., (Chapter 4) (To be published).
12. Kralj, S.; van Geel-Schutten, G. H.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Biochemistry*, **2005**, *44*, 9206-9216.
13. Van Leeuwen, S. S.; Leeftang, B. R.; Gerwig, G. J.; Kamerling, J. P., (Chapter 2) (To be published).
14. Kovacic, V.; Bauer, S.; Rosik, J.; Kovac, P. *Carbohydr. Res.*, **1968**, *8*, 282-290.
15. Kamerling, J. P.; Vliegthart, J. F. G. In *Clinical Biochemistry – Principles, Methods, Applications. Vol. 1, Mass Spectrometry*; Lawson, A. M., Ed.; Walter de Gruyter: Berlin, **1989**; pp 176-263.
16. Van der Kaaden, A.; van Doorn-van Wakeren, J. I. M.; Kamerling, J. P.; Vliegthart, J. F. G.; Tiesjema, R. H. *Eur. J. Biochem.*, **1984**, *141*, 513-519.
17. Veerkamp, J. H.; van Schaik, F. W. *Biobim. Biophys. Acta*, **1974**, *348*, 370-387.
18. Dowd, M. K.; Zeng, J.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *230*, 223-244.
19. Van Geel-Schutten, G. H.; Faber, E. J.; Smit, E.; Bonting, K.; Smith, M. R.; Ten Brink, B.; Kamerling, J. P.; Vliegthart, J. F. G.; Dijkhuizen, L. *Appl. Environ. Microbiol.*, **1999**, *276*, 44557-44562.
20. Ciucanu, I.; Kerek, F. *Carbohydr. Res.*, **1984**, *131*, 209-217.
21. Jansson, P.-E.; Kenne, L.; Liedgren, H.; Lindberg, B.; Lönnngren, J. *Chem. Commun. Univ. Stockholm*, **1976**, *8*, 1-74.
22. Hay, G. W.; Lewis, B. A.; Smith, F. *Methods Carbohydr. Chem.*, **1965**, *5*, 357-360.
23. Lee, Y. C. *Anal. Biochem.*, **1990**, *189*, 151-162.
24. Hård, K.; van Zadelhoff, G.; Moonen, P.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1992**, *209*, 895-915.

In every real man a child is hidden that wants to play

-Friedrich Nietzsche-

Chapter 6

Structural characterisation of α -D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180

Sander S. van Leeuwen*,¹ Slavko Kralj*,^{2,3} Wieger Eeuwema,^{2,3} Gerrit J. Gerwig,¹ Lubbert Dijkhuizen,^{2,3} and Johannes P. Kamerling¹

¹*Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, 3584 CH Utrecht, The Netherlands*

²*Department of Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands*

³*Centre for Carbohydrate Bioprocessing, TNO - University of Groningen, 9750 AA Haren, The Netherlands*

*The authors S. S. van Leeuwen and S. Kralj have contributed equally to the work described in this paper.

Abstract – Mutagenesis of specific amino acid residues of the glucansucrase (GTF180) enzyme from *Lactobacillus reuteri* strain 180 yielded 12 mutant enzymes that produced modified exopolysaccharides (mEPSs) from sucrose. Ethanol-precipitated and purified mEPSs were subjected to linkage analysis, Smith degradation analysis, and 1D/2D ^1H NMR spectroscopy. Comparison of the results with structural data of the previously described wild type **EPS180** and triple mutant **mEPS-PNNS** revealed a broad variation of structural elements between mEPS molecules. The amount of (α 1-3) linkages varied from 14 - 40%, the amount of (α 1-4) linkages (not present in the wild type) from 0 - 12%, and the amount of (α 1-6) linkages from 52 - 86%. The average molecular weight (M_w) ranged from 9.4 MDa to 32.3 MDa and the degree of branching varied from 8 - 20%. Using a previously established ^1H NMR structural-reporter-group concept, visual representations, that include all identified structural features, were formulated for all mEPS molecules. Variations in the mEPS structures strongly affected the physical properties of the mEPSs.

Introduction

Exopolysaccharides (EPSs) produced by lactic acid bacteria (LAB) are regularly used in the food industry,¹⁻⁴ as their physical properties influence viscosity and mouth-feel of dairy products. Many of the EPSs produced by LAB strains have been subjected to structural analysis, leading to postulated structure-function relationships.⁵⁻⁷ The next challenge in the food technology process is the design of engineered EPSs with improved physical properties.⁸⁻¹¹ Since the majority of the studied EPSs from LAB are repeating-unit-composed heteropolysaccharides,¹² the modifications investigated are mostly in the ensemble of glycosyltransferase enzymes, using nucleotide activated sugars as substrate, that produce these repeating units.⁸⁻¹⁰

Structures of non-repeating-unit-composed homopolysaccharide products of glucansucrase enzymes have not been extensively analysed. Glucansucrase enzymes of lactic acid bacteria use sucrose to synthesise a diversity of α -D-glucans, i.e. dextrans with (α 1-6) linkages, or with a majority of (α 1-6) linkages and (α 1-2), (α 1-3), and/or (α 1-4) branches (mainly found in *Leuconostoc*),^{13,14} mutan with a majority of (α 1-3) linkages (found in *Streptococcus*),¹⁵ alternan with alternating (α 1-3) and (α 1-6) linkages (only reported in *Le. mesenteroides*),¹⁶ and reuteran a highly branched structure with mainly (α 1-4) linkages (found in *Lactobacillus reuteri*).¹⁷

Recently, two glucansucrases enzymes have been described in *L.b. reuteri*, converting sucrose into large, heavily branched α -D-glucans.^{18,19} The structures of two of these α -D-glucans, i.e. **EPS180** and **EPS35-5**, have been extensively analysed (Chapters 3 and 4).^{20,21} Modifications in amino acid residues (N1134:N1135:S1136) adjacent to the catalytic D1133 (putative transition state stabilising residue) of the reuteransucrase (GTFA) enzyme, responsible for the synthesis of **EPS35-5**, were shown to introduce changes into the glycosidic linkage patterns of the EPS products of the mutant GTFA enzymes. Subsequent introduction of modifications of P1026V and I1029V, close to the catalytic D1024 (nucleophile) increased the (α 1-6) specificity of GTFA.¹¹ Glucansucrase enzymes modified in this tripeptide sequence (usually a conserved SEV sequence) immediately following the putative transition state stabilising residue have also been reported for the dextransucrase enzyme of *Le. mesenteroides* NRRL B-512F and the alternansucrase enzyme of *Le. mesenteroides* NRRL B-1335.²²

Here, based on linkage analysis, Smith degradation analysis, and 1D/2D ¹H NMR spectroscopy making use of a recently established ¹H NMR structural-reporter-group

concept,²³ we formulate a series of visual representations, that include the identified structural features of the mutant EPSs (mEPSs), produced by genetically engineered glucansucrase GTF180 enzymes using sucrose as a substrate (wild type is responsible for the synthesis of **EPS180**). Modifications were made in the GTFA-equivalent residues (S1137:N1138:A1139), and in a fourth residue (Q1140), which has not been targeted before. Mutations were selected based on sequences observed in other glucansucrase enzymes, i.e. reuteransucrase (NNS), mutansucrase (NNV), alternansucrase (YDA), and dextransucrase (SEV) enzymes, as well as single mutant variants of these sequences. For comparison the structural data of a recently analysed mEPS produced by a triple mutant of the glucansucrase GTF180 enzyme (**mEPS-PNNS**)²⁴ are included.

Results

Mutants

The glucansucrase GTF180 enzyme has a unique S1137:N1138:A1139 (SNA) sequence following the putative transition state stabilising residue, followed by a fully conserved Q1140.²⁵ This region differs from the consensus sequence SEVQ, conserved in many glucansucrase enzymes from *Streptococcus*, *Leuconostoc*, and *Lactobacillus* species.^{25,26} To study the effect on the **EPS180** structure this GTF180 region was mutated, substituting the unique tripeptide SNA for (1) the conserved SEV sequence, found in many other glucansucrase enzymes;²⁵ (2) the conserved NNS sequence, found in the reuteransucrase enzymes GTFEA from *Lb. reuteri* 121,²⁷ and GTFO from *Lb. reuteri* ATCC 55730;¹⁹ (3) NNA, SNS, and SNL sequences, as single mutations; (4) the NNV sequence, present in the mutansucrase GTFJ enzyme of *S. salivarius* ATCC 25975;²⁸ (5) the unique YDA motif, found in the alternansucrase enzyme of *Le. mesenteroides* NRRL B-1355,¹⁶ as well as the single mutant variant sequences YNA and SDA.

The completely conserved glutamine residue Q1140 has not been targeted before in any glucansucrase enzyme. This residue, directly following the tripeptide sequence was therefore also mutated: Q1140A (small), Q1140E (similar size, different side chain), and Q1140H (basic side chain).

A total of 12 mutant genes were constructed, expressed in *E. coli* BL21 Star (DE3), and mutant enzymes were used to produce mEPSs by incubating with sucrose (146 mM; 37 °C, 7 days). All mutant enzymes were able to synthesise sufficient polysaccharide material for a detailed structural analysis.

Analysis of mutant GTF180 exopolysaccharides (mEPSs)

Using a combination of methylation analysis and 1D ¹H NMR spectroscopy (structural-reporter-group concept),²³ structural elements have been identified for the 12 mutant EPS structures. Surveys of the methylation analysis data are presented in Table 1, whereas ¹H NMR data are summarised in Tables 2 and 3. For comparison the earlier data of wild type **EPS180**²⁰ and triple mutant **mEPS-PNNS**²⁴ are included in these Tables. The various 1D ¹H NMR spectra reflected only the presence of α -D-Glcp residues in all mEPSs. The H-1 signal of (-) α -D-Glcp-(1 \rightarrow 4)-[†] units can be

[†]For Glc residues at semi-defined places in the structure (-) α -D-Glcp-(1 \rightarrow x)- or -(1 \rightarrow x)- α -D-Glcp(-) is used. When the structural context of the residue is precisely known, this is indicated as follows: -(1 \rightarrow x)- α -D-Glcp-(1 \rightarrow y)- describing an x-substituted residue with an (1-y) linkage. In case of a non-reducing terminal residue α -D-Glcp-(1 \rightarrow x)- is used, a reducing terminal residue is indicated with -(1 \rightarrow x)-D-Glcp.

Table 1. Methylation analysis data of EPS180 and mEPSs. Linkage distribution data are shown in molar percentages based on GLC intensities.

EPS	Glc β (1-	-3)Glc β (1-	-4)Glc β (1-	-6)Glc β (1-	-3,6)Glc β (1-
Wild type ²⁰	12	24	-	52	12
mEPS-PNNS ²⁴	18	10	12	42	18
mEPS-SNS	12	24	-	51	13
mEPS-SNL	14	23	-	47	16
mEPS-SDA	10	24	-	56	10
mEPS-SNAE	12	16	2	52	18
mEPS-SNAA	11	6	-	69	14
mEPS-SNAH	8	8	-	76	8
mEPS-NNV	11	30	2	46	11
mEPS-SEV	13	24	-	48	15
mEPS-NNA	12	26	3	47	12
mEPS-YNA	18	21	4	39	18
mEPS-NNS	15	27	4	39	15
mEPS-YDA	19	23	7	31	20

observed between δ 5.41-5.33,^{21,23,24} while (-) α -D-Glc β -(1 \rightarrow 3)- H-1 can vary between δ 5.39-5.32,^{23,23,24} causing possible overlap between anomeric signals. In this paper the surface area under the peak between δ 5.41 and 5.39 (when present) always corresponds with the amount of 4-substituted glucopyranose residues as determined by methylation analysis. This means that in these mEPSs there is a clear separation between the signals from (α 1-3)-linked and (α 1-4)-linked glucose residues. The peak at δ \sim 4.96 is a typical structural reporter for (-) α -D-Glc β -(1 \rightarrow 6)- units (δ -range between 4.99 and 4.95 ppm).^{21,21,23,24}

2D ¹H-¹H TOCSY experiments with increasing mixing times (10, 30, 60, 120, and 150 ms; 60 and 120 ms are shown in figures as examples) were interpreted, where available, to unravel the structural elements present in the mEPSs. Starting from the anomeric signals in the TOCSY spectrum with 10 ms mixing time, cross-peaks are observed with H-2 signals. In each incremental step subsequent ¹H signals arise on the anomeric track and can be assigned to H-3, H-4, H-5, H-6a, and H-6b. Previous analysis of di- and trisaccharide standards revealed clear chemical-shift patterns for specific -(1 \rightarrow x)- α -D-Glc β -(1 \rightarrow y)- residues.²³ Making use of structural-reporter-group signals and knowledge of these chemical-shift patterns, the mEPS structures were analysed.

S1137N (mEPS-NNA) - Methylation analysis of **mEPS-NNA** revealed the presence of terminal, 3-substituted, 4-substituted, 6-substituted, and 3,6-disubstituted glucopyranose residues in a molar percentage of 12, 26, 3, 47, and 12% (Table 1). Integration of the anomeric signals in the 1D ¹H NMR spectrum (Figure 1) showed 35% (α 1-3)-linked glucose (δ_{H-1} \sim 5.35), 4% (α 1-4)-linked glucose (δ_{H-1} \sim 5.40), and 61% (α 1-6)-linked glucose (δ_{H-1} \sim 4.96) residues (Table 2), which is in accordance with the methylation analysis data.

Table 2. Surface area ratios of peaks in the 500-MHz 1D ^1H NMR spectra of **EPS180** and mEPSs recorded at 300K in D_2O . (columns 2-4) ratios of H-1 signals of (α 1-3)-, (α 1-4)-, and (α 1-6)-linked Glcp residues; (columns 5 and 6) ratios of H-5 signals of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- and -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- units. Ratios are shown in percentages relative to the total amount of glucose residues.

EPS	(α 1-3) $\delta_{\text{H-1}}$ 5.35	(α 1-4) $\delta_{\text{H-1}}$ 5.40	(α 1-6) $\delta_{\text{H-1}}$ 4.96	6G ³ $\delta_{\text{H-5}}$ 4.20	4G ³ $\delta_{\text{H-5}}$ 4.12
Wild type ²⁰	31	0	69	31	0
mEPS-PNNS ²⁴	28	12	60	23	7
mEPS-SNS	34	0	66	34	0
mEPS-SNL	38	0	62	31	0
mEPS-SDA	35	< 1	65	31	< 1
mEPS-SNAE	29	3	68	24	3
mEPS-SNAA	16	0	84	11	0
mEPS-SNAH	14	0	86	12	0
mEPS-NNV	40	2	58	35	2
mEPS-SEV	34	1	65	29	1
mEPS-NNA	35	4	61	28	4
mEPS-YNA	36	6	58	28	6
mEPS-NNS	37	4	59	30	6
mEPS-YDA	40	8	52	21	8

Note: 6G³ represents -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units; 4G³ represents -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- units.

The structural-reporter-group concept identified the peaks at $\delta_{\text{H-5}}$ 4.20 and 4.12 as markers for -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- and -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- units, respectively.^{21,24} Integration of these peaks in the 1D ^1H NMR spectrum revealed that these moieties were present at 28 and 4%, respectively, in **mEPS-NNA**.

On the (α 1-3)-anomeric track in the 60 ms TOCSY spectrum (Figure 1) of **mEPS-NNA** $\delta_{\text{H-3}}$ 4.01 was observed, confirming the occurrence of -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- units (compare a3 in **mEPS-PNNS**).²⁴ On the same track in the TOCSY spectrum the structural reporter $\delta_{\text{H-4}}$ 3.45 was detected, indicating the occurrence of α -D-Glcp-(1 \rightarrow 3)- units (compare with residue **B** in compounds **5a**, **9a**, and **10** in Ref. 24). On the (α 1-3)-anomeric track in the 120 ms TOCSY spectrum (Figure 1) the H-5 signal at δ 4.20 confirmed the presence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units.²⁰

On the (α 1-4) anomeric track in the 120 ms TOCSY spectrum a cross-peak at $\delta_{\text{H-4}}$ 3.42 is present, indicating the occurrence of α -D-Glcp-(1 \rightarrow 4)- units.²³ The specific values $\delta_{\text{H-4}}$ 3.50 and $\delta_{\text{H-6b}}$ 4.00 for -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- units, and $\delta_{\text{H-3}}$ 3.96 for -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 4)- units are not observed in the TOCSY spectra, making the occurrence of such elements highly unlikely. If present, they occur in negligible amounts.

On the (α 1-6)-anomeric track in the 60 ms TOCSY spectrum a cross-peak at $\delta_{\text{H-4}}$ 3.42 occurred, indicating the presence of α -D-Glcp-(1 \rightarrow 6)- units. In the 60 ms TOCSY spectrum the $\delta_{\text{H-3}}$ signal at 4.01 ppm, specific for -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)- units is not observed. On the same track in the 120 ms TOCSY spectrum the unique

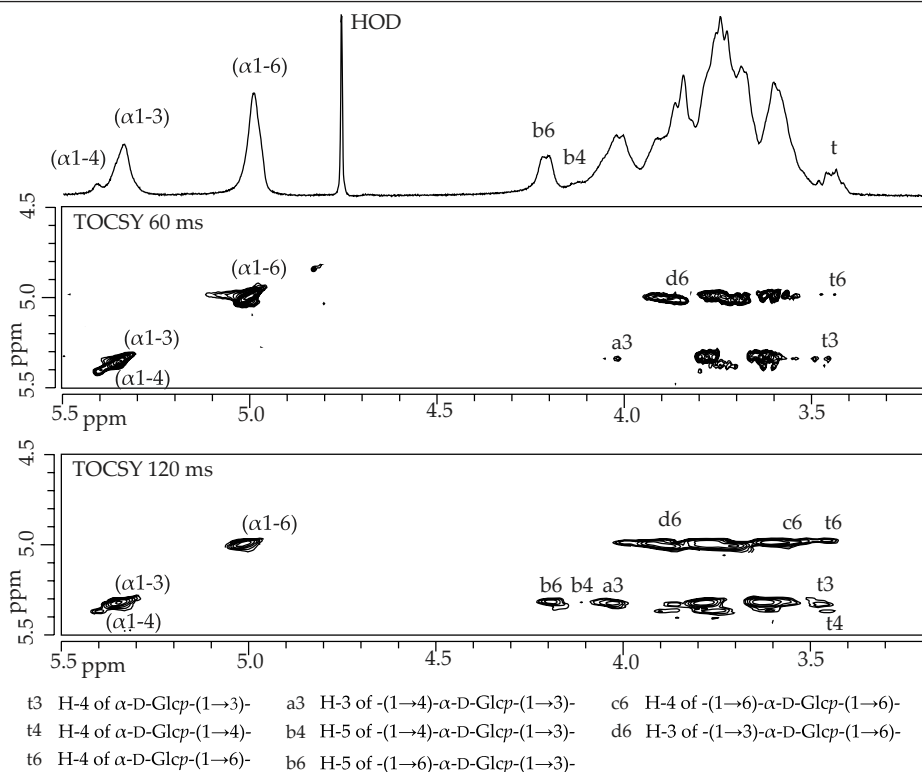


Figure 1. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-NNA**, recorded at 300K in D_2O . Anomeric signals in the TOCSY spectra have been indicated on the diagonal. In the horizontal tracks structural-reporter-group signals have been indicated with labels; legends of the labels are included in the figure.

$\delta_{\text{H-3}} \sim 3.87$ signal, indicating the presence of -(1→3)- α -D-Glcp-(1→6)- and/or -(1→3,6)- α -D-Glcp-(1→6)- units, is detected. The absence of these $\delta_{\text{H-3}}$ resonances on the (α 1-3)- and (α 1-4)-anomeric tracks indicates that -(1→3)- α -D-Glcp-(1→3)- and -(1→3)- α -D-Glcp-(1→4)- units do probably not occur.^{20,24} On the (α 1-6)-anomeric track in the 120 ms spectrum $\delta_{\text{H-4}} 3.51$, indicative of -(1→6)- α -D-Glcp-(1→6)- units, is observed, confirming the presence of this structural element.

Q1140H (mEPS-SNAH) - Methylation analysis of **mEPS-SNAH** indicated the occurrence of terminal, 3-substituted, 6-substituted, and 3,6-disubstituted glucopyranose residues in a molar percentage of 8, 8, 76, and 8% (Table 1). In the 1D ^1H NMR spectrum (Figure 2) 14% (α 1-3)-linked glucose ($\delta_{\text{H-1}} \sim 5.35$) and 86% (α 1-6)-linked glucose ($\delta_{\text{H-1}} \sim 4.96$) residues were observed (Table 2), which fits with the methylation analysis data. The surface area of $\delta_{\text{H-5}} 4.20$ (structural reporter for -(1→6)- α -D-Glcp-(1→3)- units)²⁰ corresponded with 12% of all residues in **mEPS-SNAH**.

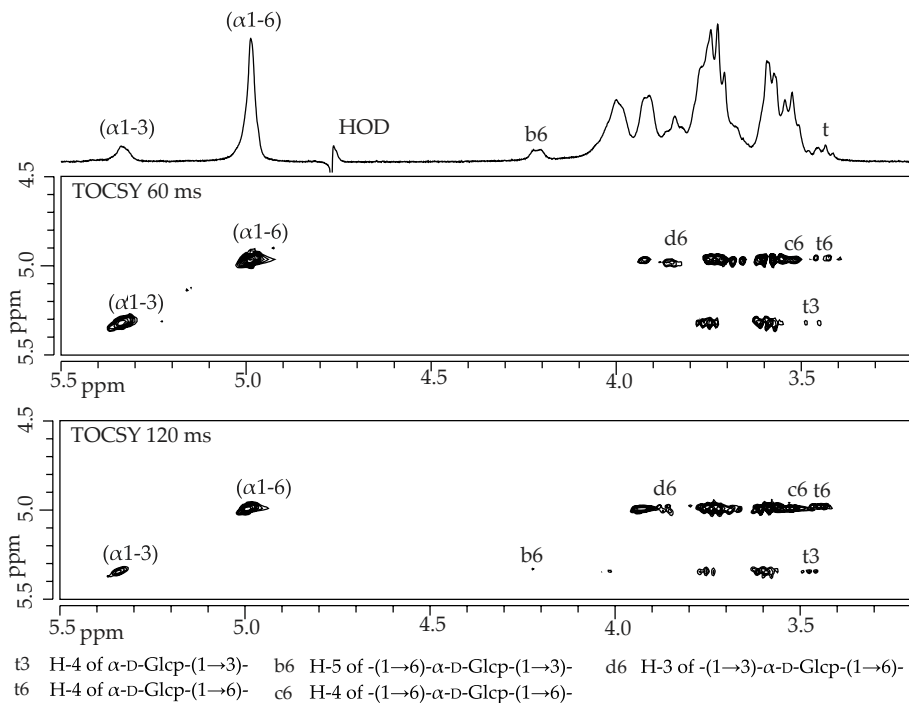


Figure 2. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SNAH**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

On the (α 1-3)-anomeric track in the 60 ms TOCSY spectrum (Figure 2) $\delta_{\text{H-4}}$ 3.45 was observed, indicating the occurrence of terminal α -D-Glcp-(1 \rightarrow 3)- units. In the 120 ms TOCSY spectrum the H-5 signal at δ 4.20 is detected, confirming the occurrence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units.²⁰ The resonance at $\delta_{\text{H-3}}$ 3.90, specific for a -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 3)- unit was not seen on this track, which means that this element probably does not occur.

On the (α 1-6)-anomeric track in the 60 ms TOCSY spectrum H-4 signals were observed at δ 3.50 and 3.42, indicating the occurrence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- and α -D-Glcp-(1 \rightarrow 6)- units, respectively. The occurrence of the former was further supported by the observation of $\delta_{\text{H-5}}$ 3.90 on the same track in the 120 ms TOCSY spectrum.^{20,23} The H-3 resonance at δ \sim 3.87 revealed the presence of -(1 \rightarrow 3,6)- α -D-Glcp-(1 \rightarrow 6)- and/or -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- units.²⁰

A1139L (mEPS-SNL) - Methylation analysis of **mEPS-SNL** showed the presence of terminal, 3-substituted, 6-substituted, and 3,6-disubstituted glucopyranose residues in a

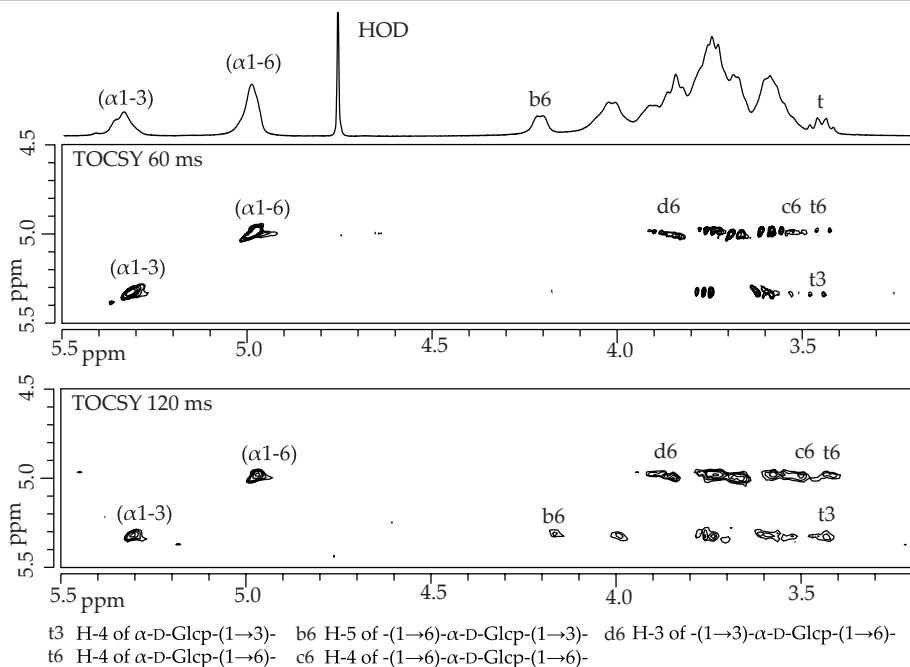


Figure 3. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SNL**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

molar percentage of 14, 23, 47, and 16% (Table 1). Integration of the anomeric peaks in the 1D ^1H NMR spectrum (Figure 3) showed 38% (α 1-3)-linked glucose ($\delta_{\text{H-1}} \sim 5.35$) and 62% (α 1-6)-linked glucose ($\delta_{\text{H-1}} \sim 4.96$) residues, matching the linkage analysis data (Table 2). A small signal seemed to be present at $\delta_{\text{H-1}}$ 5.40, suggesting the occurrence of (α 1-4)-linked glucose residues. However, in the case that this small peak really reflects such residues, the amount is $< 0.5\%$ of the residues in **mEPS-SNL** (see also the absence in the methylation analysis). Therefore, this observation was neglected for the formulation of the visual representation (see section 3.1.1).

On the H-1 track of the (-) α -D-Glcp-(1 \rightarrow 3)- units in the 60 ms TOCSY spectrum (Figure 3) two H-4 signals were observed at δ 3.51 and 3.45, indicating the occurrence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- and α -D-Glcp-(1 \rightarrow 3)- units, respectively.^{20,23} The occurrence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units is further supported by the detection of $\delta_{\text{H-5}}$ 4.20 in the 120 ms TOCSY spectrum on the same track.²⁰ Integration of the $\delta_{\text{H-5}}$ signal at 4.20 ppm in the 1D ^1H NMR spectrum revealed that 31% of all residues are -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units.

On the (α 1-6)-anomeric track in the 60 ms TOCSY spectrum the $\delta_{\text{H-4}}$ signals at 3.50 and 3.42 ppm reflected the occurrence of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- and α -D-

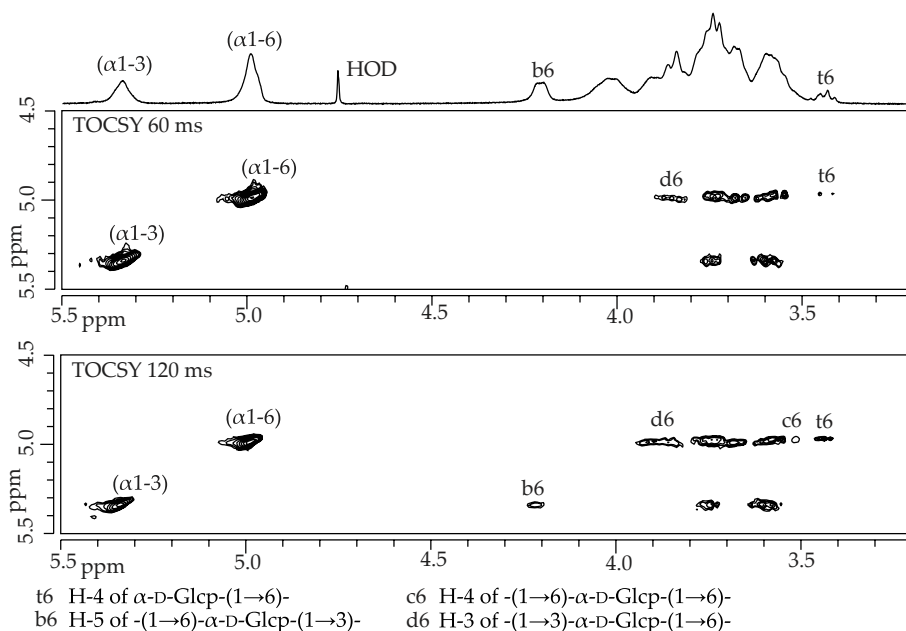


Figure 4. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SNS**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

Glcp-(1 \rightarrow 6)- units. The H-3 signal at $\delta \sim 3.87$ corresponds with the presence of -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- and/or -(1 \rightarrow 3,6)- α -D-Glcp-(1 \rightarrow 6)- units.²⁰

Other mEPSs - In a similar fashion as described in sections 2.2.1. - 2.2.3. the other mEPSs were analysed, and the relevant data are presented in Tables 1, 2, and 3. The 1D/2D ^1H NMR spectra are depicted in Figures 4 - 12.

All ^1H NMR data of **mEPS-SNS** (Figure 4) match those of wild type **EPS180**.²⁰ In the 1D ^1H NMR spectrum of **mEPS-SDA** (Figure 5) a small signal at $\delta_{\text{H-1}}$ 5.40 was observed, corresponding with the occurrence of $\sim 0.5\%$ of (α 1-4)-linked glucose residues. On the (α 1-4)-anomeric track in the 60 ms TOCSY spectrum weak H-2 and H-3 signals at δ 3.58 and 3.70, respectively, were observed. These values fit with the occurrence of α -D-Glcp-(1 \rightarrow 4)- or -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- units.^{21,24} Since no further data are available, no distinction between these two types could be made. On the (α 1-4)-anomeric track in the 120 ms TOCSY spectrum of **mEPS-SNAE** (Figure 6) a weak $\delta_{\text{H-5}}$ 4.12 was observed, indicating the occurrence of -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- units. In the 1D ^1H NMR spectrum this signal is not clear, however, the surface area determined for that region is sufficient to correspond with all (α 1-4)-linked residues.

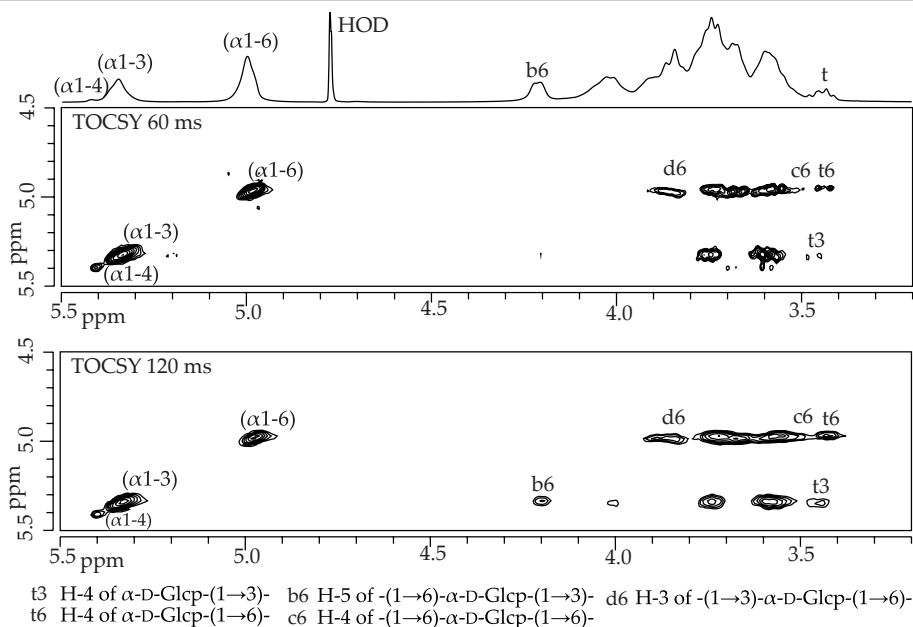


Figure 5. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SDA**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

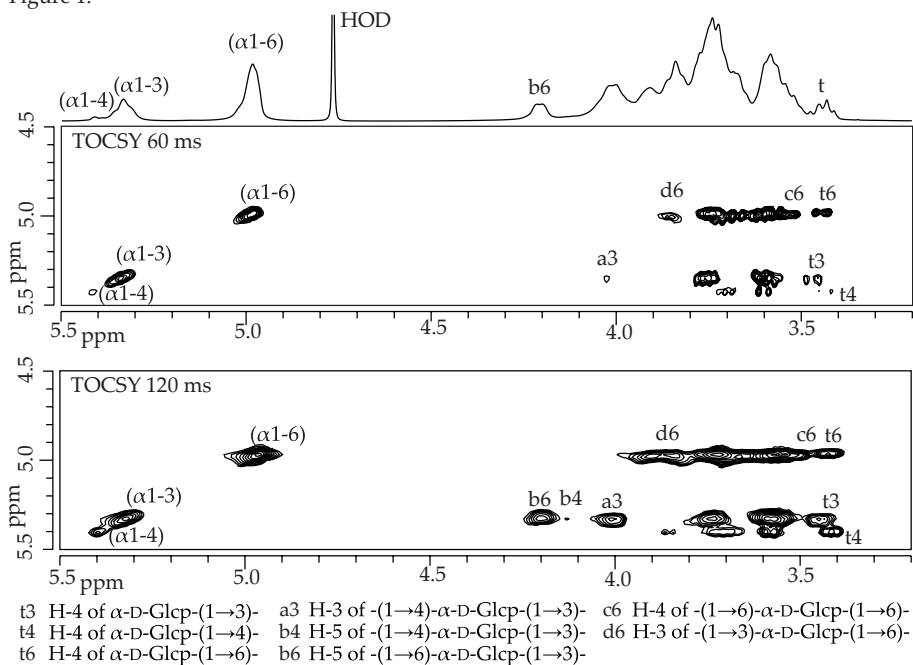


Figure 6. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SNAE**, recorded at 300K in D_2O . For details of the coding system, see Figure 1

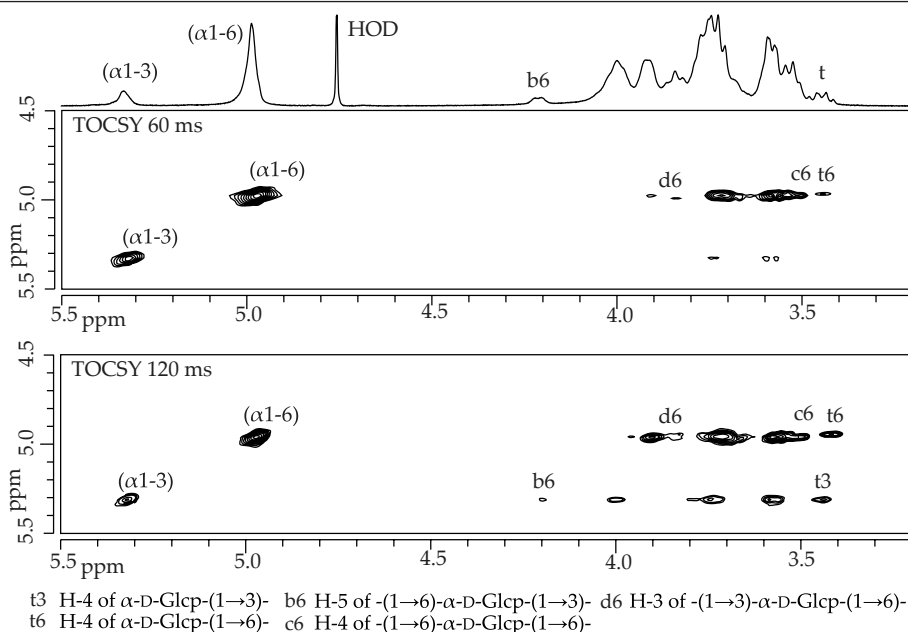


Figure 7. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SNAA**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

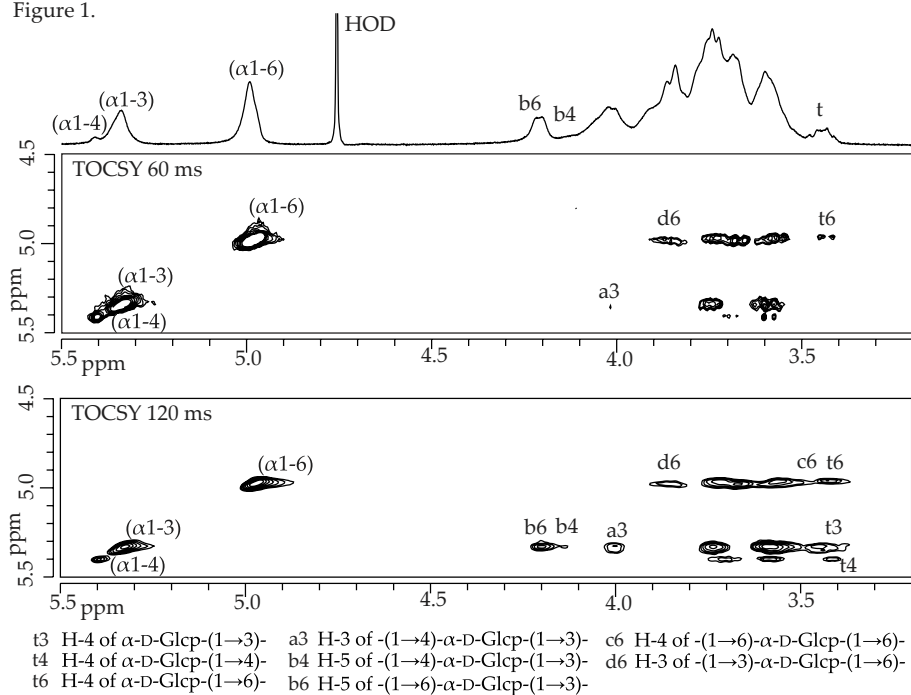


Figure 8. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-NNV**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

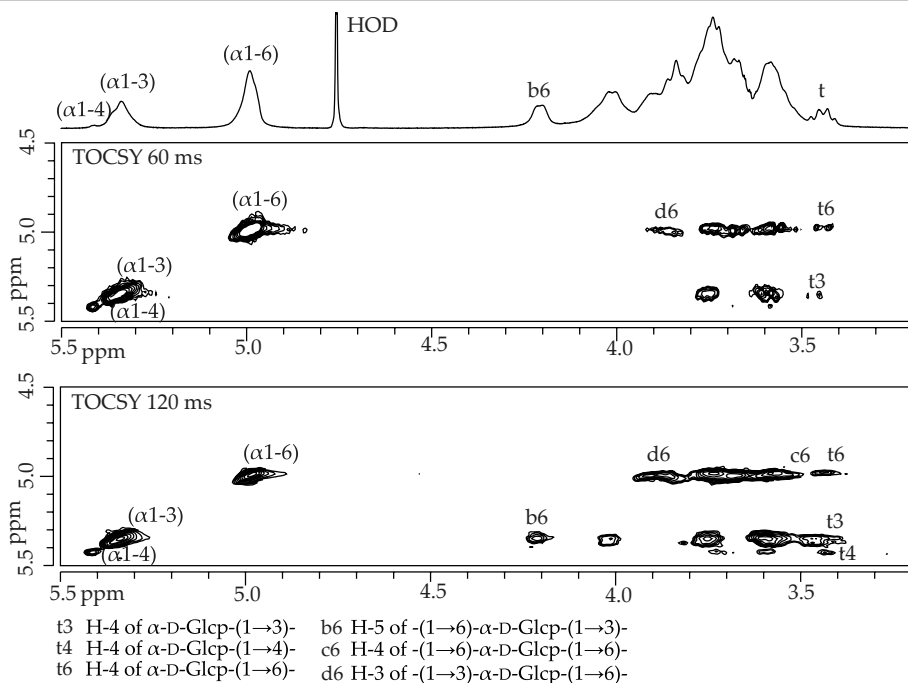


Figure 9. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectra (top, mixing time 60 ms; bottom, 120 ms) of **mEPS-SEV**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

For **mEPS-SNAA** (Figure 7), **mEPS-NNV** (Figure 8), and **mEPS-SEV** (Figure 9) no special remarks need to be made. The 1D ^1H NMR spectra of saturated solutions of **mEPS-YNA** (Figure 10), **mEPS-NNS** (Figure 11), and **mEPS-YDA** (Figure 12) showed a low signal-to-noise ratio. For **mEPS-YNA** and **mEPS-NNS** TOCSY spectra could be recorded with mixing times up to 60 ms. For **mEPS-YDA** only a 1D ^1H NMR spectrum could be recorded.

Smith degradation

Previous studies of wild type **EPS180** and **mEPS-PNNS** revealed that only single (α 1-3) bridges occurred in these polysaccharides.^{20,24} Since **mEPS-YDA** represented one of the highest amounts of (α 1-3)-linked glucose residues (40%), it is one of the most likely mEPSs to deviate from this structural limitation. Therefore, **mEPS-YDA** was subjected to Smith degradation to investigate the degree of polymerisation of (α 1-3) glycosidic bonds. Possible products include: α -D-Glcp-(1 \rightarrow 1)-Gro, [α -D-Glcp-(1 \rightarrow 3)]- α -D-Glcp-(1 \rightarrow 1)-Gro, and, due to overhydrolysis, Gro, D-Glc and [α -D-Glcp-(1 \rightarrow 3)]- α -D-Glc.

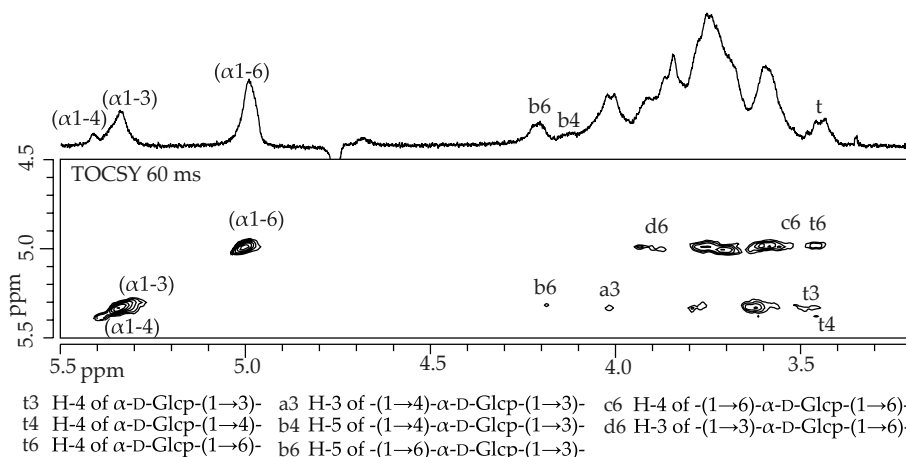


Figure 10. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectrum (mixing time 60 ms) of **mEPS-YNA**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

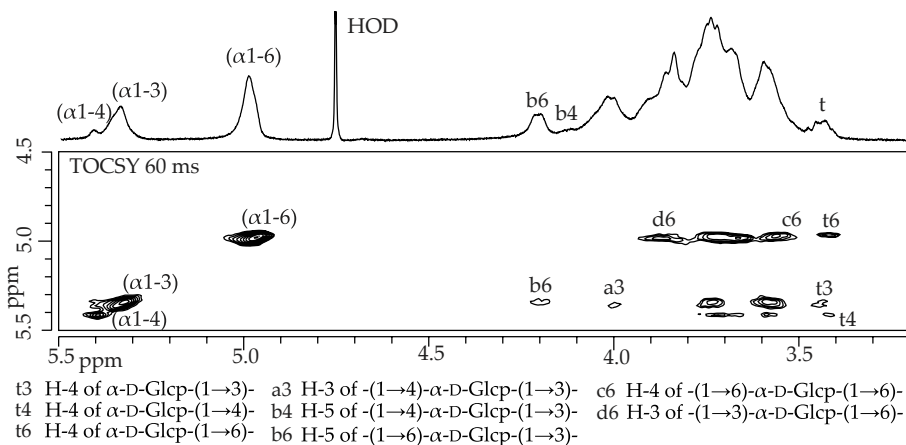


Figure 11. 500-MHz 1D ^1H NMR spectrum and 2D ^1H - ^1H TOCSY spectrum (mixing time 60 ms) of **mEPS-NNS**, recorded at 300K in D_2O . For details of the coding system, see Figure 1.

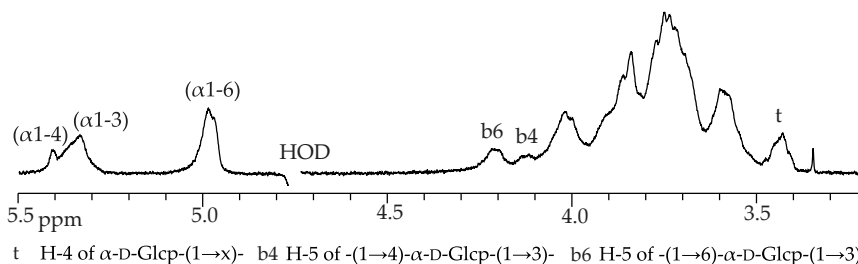


Figure 12. 500-MHz 1D ^1H NMR spectrum of **mEPS-YDA** recorded at 300K in D_2O . For details of the coding system, see Figure 1.

Table 4. Average molecular weights of **EPS180** and mEPSs shown in MDa and in percentages relative to wild type **EPS180**, as determined by HPSEC-MALLS.

EPS	M_w (MDa)	Relative M_w (%)
Wild type	32.3	100
mEPS-PNNS	9.4	29
mEPS-SNS	30.5	94
mEPS-SNL	19.3	60
mEPS-SNAE	16.2	50
mEPS-SNAA	23.9	74
mEPS-SEV	21.2	66
mEPS-NNA	24.2	75
mEPS-YNA	12.5	39
mEPS-NNS	20.0	62
mEPS-YDA	10.6	33

HPAEC analysis on CarboPac PA-100 (data not shown) of the residue revealed two major peaks that were identified as Gro (R_t 2.3 min) and α -D-Glc β -(1 \rightarrow 1)-Gro (R_t 6.2 min).^{20,24} Since α -D-Glc β -(1 \rightarrow 3)-D-Glc β under the same conditions has an R_t value > 10 min, the presence of [α -D-Glc β -(1 \rightarrow 3)]_n α -D-Glc β -(1 \rightarrow 1)-Gro with n=1 or higher could be excluded.^{20,24}

The absence of larger structures than glucosyl-glycerol indicates that the **mEPS-YDA** structure does not contain two or more consecutive (α 1-3) linkages. Since two mEPSs with significant changes in structure (i.e. **mEPS-YDA** and **mEPS-PNNS**²⁴) show this feature, similar to the wild type **EPS180**,²⁰ it is probable that this holds true for all mEPS structures.

Average molecular weights

Using HPSEC-MALLS, with standard mixtures of linear dextrans and pullulans for calibration, the size of a selection of the mEPSs was determined. The wild type **EPS180** was taken as a reference, set to an average molecular weight of 100%, and the average molecular weights of the mEPSs were expressed in percentages of that of the reference (Table 4).

The wild type **EPS180** and **mEPS-SNS**, having similar methylation and ¹H NMR analysis data, have also comparable average molecular weights. The average molecular weights of **mEPS-SNAA** and **mEPS-NNA** are decreased to 74 and 75%, respectively. The readily soluble **mEPS-SNL** and **mEPS-SEV** and the poorly soluble **mEPS-NNS** have similar average molecular weights at 60, 66, and 62%, respectively, of that of **EPS180**. The highly soluble **mEPS-SNAE** polysaccharide has an average molecular weight of 50% of that of **EPS180**. Finally, the mEPSs with the most reduced average molecular weight are the very poorly soluble **mEPS-YNA**, **mEPS-YDA**, and **mEPS-PNNS**, with 39, 33, and 29%, respectively, of that of **EPS180**.

Although the data for **mEPS-YNA**, **mEPS-YDA**, and **mEPS-PNNS** suggest a correlation between solubility and average molecular weight, the relatively low molecular weight (50%) for the highly soluble **mEPS-SNAE** and the relatively high molecular weight (62%) for the poorly soluble **mEPS-NNS** indicate that no clear conclusions can be made yet.

Discussion and Conclusions

Discussion

As described recently,²³ based on the $^1\text{H}/^{13}\text{C}$ NMR data of a series of model α -D-gluco-oligosaccharides, an ^1H NMR structural-reporter-group concept could be developed that was used in the characterisation of fragments derived from bacterial α -D-glucans produced by native (GTF180,²⁰ GTFA²¹) and mutant (triple mutant of GTF180²⁴) glucansucrase enzymes from *Lb. reuteri* strains. The developed and further expanded concept turned out to be also applicable on the polysaccharide level, and visual representations, that include all identified structural features, were formulated for **EPS180**,²⁰ **EPS35-5**,²¹ and **mEPS-PNNS**,²⁴ whereby the methylation analysis of each of the EPSs dictate the structural boundaries of the representations. Making use of the established building blocks as generated by the structural-reporter-group ^1H NMR library, and taking into account the representations of **EPS180** and **mEPS-PNNS**, Scheme 1 surveys similar representations for the 12 α -D-glucan mEPSs produced by the different mutant glucansucrase GTF180 enzymes, which are based on the quantification of the different building blocks present, as shown in Table 5.

Starting from the integrations of the anomeric signals in the 1D ^1H NMR spectra, the amounts of (α 1-3)-linked, (α 1-4)-linked, and (α 1-6)-linked residues are found (Table 2). Also the amounts of $-(1\rightarrow6)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}$ and $-(1\rightarrow4)\text{-}\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}$ units can be deduced from the 1D ^1H NMR spectra via the integration of their H-5 signals (Table 2). Since two consecutive (α 1-3) linkages do not occur in any of the mEPSs, as evidenced from TOCSY measurements and Smith degradation analysis, there are only three types of (α 1-3)-linked glucose residues possible, i.e. the two internal units mentioned above and terminal $\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}$ units. Therefore, the amount of $\alpha\text{-D-Glcp-(1}\rightarrow3)\text{-}$ can be determined from the difference in surface area between the (α 1-3) H-1 signal at $\delta \sim 5.35$ and the H-5 signals at δ 4.20 and 4.12.

To determine the amounts of $\alpha\text{-D-Glcp-(1}\rightarrow4)\text{-}$ and $\alpha\text{-D-Glcp-(1}\rightarrow6)\text{-}$ units in the mEPSs, use was made of the methylation analysis data and the $\delta_{\text{H-4}}$ signals at 3.41-

Table 5. Percentages of the different building blocks determined for the wild type **EPS180** and the mEPSs.

EPS	G ³	4G ³	6G ³	G ⁴	4G ⁴	6G ⁴	G ⁶	3G ⁶	4G ⁶	6G ⁶	3,6G ⁶
Wt ²⁰	0	0	36	0	0	0	12	24	0	16	12
PNNS ²⁴	4	7	17	6	3	3	8	10	2	22	18
SNS	0	1	31	1	0	0	12	24	0	19	13
SNL	7	0	31	0	0	0	8	23	0	16	15
SDA	3	1	31	1	0	0	6	25	0	23	10
SNAE	6	2	24	2	0	0	10	14	0	24	18
SNAA	6	0	12	0	0	0	6	6	0	58	12
SNAH	3	0	13	0	0	0	5	8	0	63	8
NNV	4	2	33	2	0	0	5	28	0	15	11
SEV	5	1	32	1	0	0	8	24	0	15	14
NNA	3	3	32	3	0	0	6	26	0	15	12
YNA	6	4	29	4	0	0	8	21	0	10	18
NNS	4	4	28	4	0	0	7	21	0	17	15
YDA	11	8	21	8	0	0	1	22	0	9	20

Note: G³ represents α -D-Glcp-(1 \rightarrow 3)- units, 4G³ represents -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- units, etc.

3.45 ppm on the (α 1-3)-, (α 1-4)-, and (α 1-6)-anomeric tracks in the TOCSY spectra. The total amount of terminal residues follows directly from the methylation analysis (Table 1). Evaluation of the δ_{H-4} signals showed that all (α 1-4)-linked glucose residues are in fact terminal residues; and they occur only in α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- sequences. Having determined the amounts of terminal α -D-Glcp-(1 \rightarrow 3)- and α -D-Glcp-(1 \rightarrow 4)- units, the amounts of terminal α -D-Glcp-(1 \rightarrow 6)- units can be calculated via subtraction from the total amount of terminal residues.

As revealed by methylation analysis, all mEPSs contain only 3,6-branch points. The amount of the 3,6-branching (and terminal) residues follows directly from the integration of the GLC peaks. The 2D NMR spectra showed that none of the (-) α -D-Glcp-(1 \rightarrow 4)- units was 3-substituted, which means that all 3,6-disubstituted residues are in fact -(1 \rightarrow 3,6)- α -D-Glcp-(1 \rightarrow 6)- units. In a similar fashion the amount of 3-substituted residues is obtained from the methylation analysis data, and for the reasons mentioned above all 3-substituted residues are in fact -(1 \rightarrow 3)- α -D-Glcp-(1 \rightarrow 6)- units.

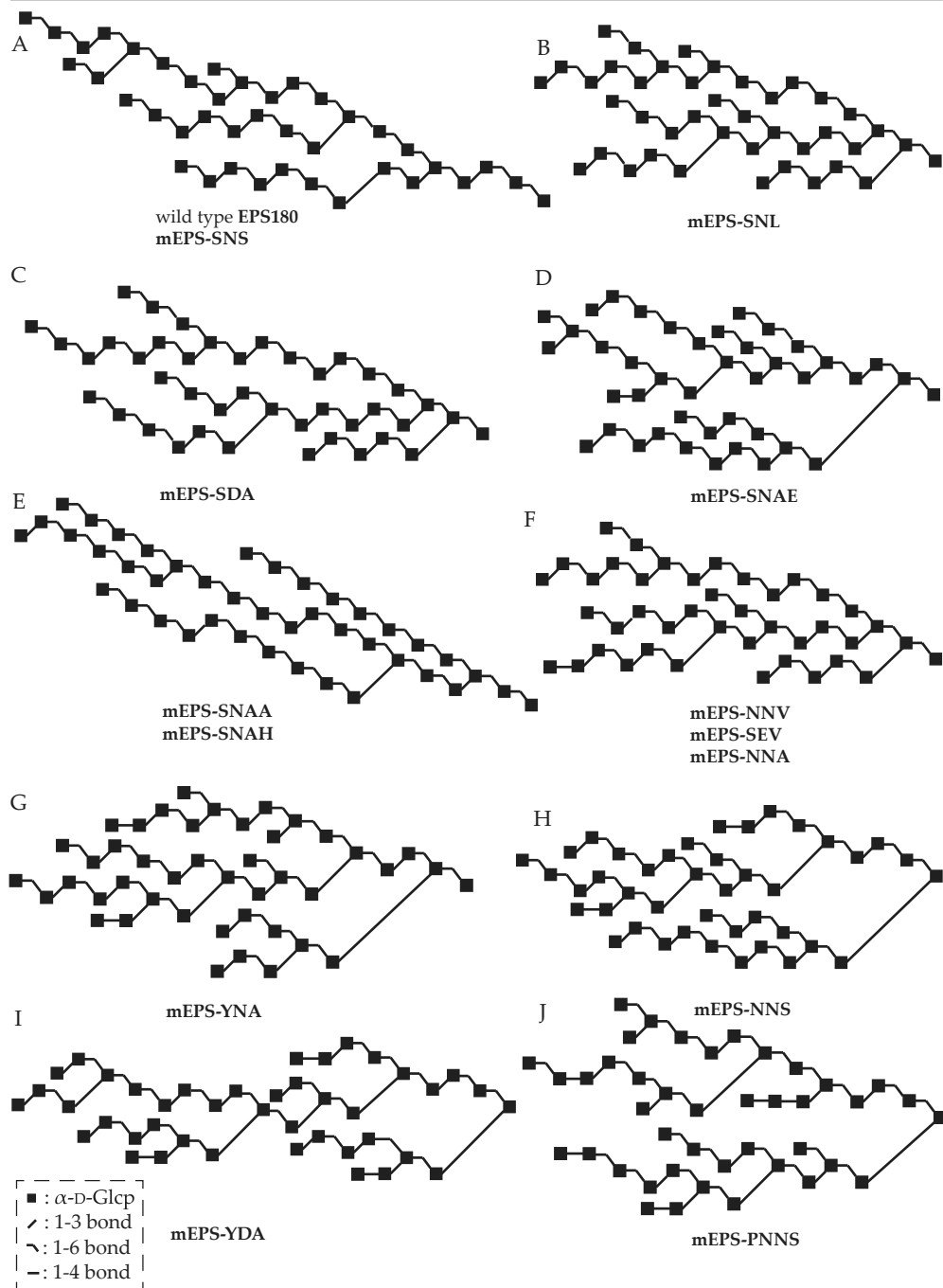
These data leave only one type of building block for the mEPSs, except for **mEPS-YDA** and **mEPS-SDA**, undetermined, i.e. the -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 6)- units. The latter can be deduced from the relative percentages of all the other building blocks, and verified by the difference between the total amount of 6-substituted glucose residues (methylation analysis) and the amount of -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 3)- units (δ_{H-5} 4.20 surface area in the 1D ¹H NMR spectrum).

Where the previously described **mEPS-PNNS** contained 11 different building blocks,²⁴ the 12 mEPSs described here contain between 5 and 8 different building blocks. In case of **mEPS-SDA** it was uncertain whether the (-) α -D-Glcp-(1 \rightarrow 4)- units were situated terminally, or as internal -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- units as well. Since all other similar mEPSs that contain (α 1-4)-linked glucose residues showed only terminal α -D-Glcp-(1 \rightarrow 4)- residues it seems likely that this is also the case for **mEPS-SDA**, therefore the representation shown in Scheme 1 assumes only terminal α -D-Glcp-(1 \rightarrow 4)- units to occur. For **mEPS-YDA** it was most difficult determining the existing building blocks, since 2D NMR spectroscopy was not possible. For the latter polysaccharide it is possible that all (α 1-4)-linked residues are terminal, however, it could also resemble the **mEPS-PNNS** structure,²⁴ where some -(1 \rightarrow 6)- α -D-Glcp-(1 \rightarrow 4)- units occur, but also -(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 6)- units are found. Since all (α 1-4)-linked glucoses in **mEPS-YDA** are situated in (-) α -D-Glcp-(1 \rightarrow 4)- α -D-Glcp-(1 \rightarrow 3)- sequences (Table 2), it seems likely that **mEPS-YDA** resembles the other mEPSs, leading to a structure with almost no α -D-Glcp-(1 \rightarrow 6)- units (Scheme 1).

Conclusions

Analysis of the EPS products (mEPSs) of mutant GTF180 enzymes revealed that simple modifications in a specific region can have large effects on the structure of the glucan product. The amount of branching could be varied from 8% in **mEPS-SNAH** to 20% in **mEPS-YDA**. The amount of (α 1-6)-linked glucose residues varied from 52% in **mEPS-YDA** to 86% in **mEPS-SNAH**, while the amount of (α 1-3)-linked glucose residues varied from 14% in **mEPS-SNAH** to 39% in **mEPS-YDA**. The amount of (α 1-4)-linked glucose residues (not present in wild type **EPS180**,²⁰ and 12% in **mEPS-PNNS**²⁴) occurred up to 8% in **mEPS-YDA**. Some mEPSs were difficult to dissolve (**mEPS-YDA**, **mEPS-YNA**, and **mEPS-NNS**) resulting in difficult recording of NMR spectra. All the mEPSs showed the presence of a mix of terminal α -D-Glcp-(1 \rightarrow 3)-, α -D-Glcp-(1 \rightarrow 6)-, and/or α -D-Glcp-(1 \rightarrow 4)- units, while the wild type **EPS180** contained only terminal α -D-Glcp-(1 \rightarrow 6)- residues. Despite all the structural changes in the mutant EPSs, one structural element of the original **EPS180** remained: all mEPSs contain single (α 1-3) bridges.

The presented results show that a wide range of products can be made, through simple modifications of a single glucanase enzyme.



Scheme 1. Visual representations that include all identified structural features of (A) wild type EPS180²⁰ and mEPS-SNS, (B) mEPS-SNL, (C) mEPS-SDA, (D) mEPS-SNAE, (E) mEPS-SNAA and mEPS-SNAH, (F) mEPS-NNA, mEPS-NNV, and mEPS-SEV, (G) mEPS-YNA, (H) mEPS-NNS, (I) mEPS-YDA, and (J) mEPS-PNNS²⁴.

Materials and Methods

Bacterial strains, plasmids, media, and growth conditions

Escherichia coli TOP 10 (Invitrogen, Carlsbad, CA) was used as host for cloning purposes, and plasmid pET15b (Novagen, Madison, WI) for expression of the different (mutant) *gtf* genes in *E. coli* BL21 Star (DE3) (Invitrogen). *E. coli* strains were grown aerobically at 37 °C in LB medium. *E. coli* strains containing recombinant plasmids were cultivated in LB medium with 100 µg/mL ampicillin. Agar plates were made by adding 1.5% agar to the LB medium.

Cloning and mutagenesis

To facilitate future manipulation the N-terminally truncated GTF180 gene,¹⁸ obtained using PCR with primers 180expCF 5'-GATGCATGAGCTCCCATGGGCATTAACGGCCAACAATATATATATGACCC-3', containing *SacI* (Italics) and *NcoI* (bold) restriction sites and 180expR 5'-ATATCGATGGGCCCCGGATCCTATTAGTGATGGTGATGGTGATGTTTITGGCCGTTTAAATCACCAGGTTTAAATGG-3', containing *ApaI* (italics) and *BamHI* (bold) restriction sites and a C-terminal His-tag (underlined),¹⁸ was digested with *SacI/ApaI*, and cloned in the corresponding sites of pBluescript II SK+, yielding plasmid pBGTF180-ΔN. Furthermore, two unique restrictions (*SacI*, 2802bp; *XhoI*, 3606bp) were introduced using the QuickChange™ site-directed mutagenesis kit (Stratagene; LaJolla, CA) and the following primers 5'-GGAATAAGGATAGTGAAAATGTCGACTACGGTGGTTTIGC-3' 5'-ATAAAGATTCAGTTCCTCGAGTTTACTATGGAGACC-3', and the complementary reverse primers, containing an introduced *SacI* and *XhoI* site, respectively (underlined, silent mutation by change of base shown in bold face), yielding plasmid pBGTF180-ΔNSX. This plasmid was digested with *NcoI* and *BamHI* and ligated in the corresponding sites of pET15b yielding p15GTF180-ΔNSX. pBGTF180-ΔNSX was used for site-directed mutagenesis, sequencing and rapid exchange (using *NcoI* and *XhoI* restriction sites) with p15GTF180-ΔNSX.

The various single and double mutants were generated using The QuickChange™ site-directed mutagenesis kit (Stratagene, LaJolla, CA) and appropriate primer pairs to introduce mutations in pBGTF180SX. After successful mutagenesis (confirmed by restriction analysis and/or nucleotide sequencing), pBGTF180SX (containing mutation) was digested with *NcoI* and *XhoI* and ligated in the corresponding sites of p15GTF180-ΔNSX.

Purification of GTF180 mutant proteins

GTF180 mutant proteins were produced and Ni-NTA purified as described previously.²⁹

In vitro glucan production by mutant GTF180 enzymes

Polymers were produced by incubation of mutant enzyme preparations with 146 mM sucrose for 7 days, using the conditions described previously,²⁷ and addition of 1% Tween 80 and 0.02% sodium azide. Glucans produced were isolated by precipitation with ethanol as described previously.²⁷

Linkage analysis

Polysaccharide samples were permethylated using CH_3I and solid NaOH in DMSO, as described earlier.³⁰ After hydrolysis with 2 M TFA (2 h, 120 °C), the partially methylated monosaccharides were reduced with NaBD_4 (2 h, room temperature). Conventional work-up, involving neutralisation with HOAc and removal of boric acid by co-evaporation with MeOH, followed by acetylation with 1:1 acetic anhydride: pyridine (3 h, 120 °C), yielded mixtures of partially methylated alditol acetates, which were analysed by GLC-EI-MS.^{31,32}

Smith degradation

A sample of **mEPS-YDA** (10 mg) was incubated with 2 mL 50 mM sodium periodate in 0.1 M NaOAc (pH 4.3) for 96 h at 4 °C in the dark. Then, the excess of periodate was destroyed by addition of 0.1 mL ethylene glycol. The oxidised polysaccharide solution was dialysed against tap water (24 h, room temperature), treated with excess NaBH_4 (18 h, room temperature), and subsequently neutralised with 4 M HOAc.³³ After co-evaporation of boric acid with MeOH, the residue was hydrolysed with 90% HCOOH (30 min, 90 °C). Finally, the solution was concentrated under a stream of N_2 , and the products were analysed by HPAEC-PAD.

HPAEC-PAD

High-pH anion-exchange chromatography was performed on a Dionex DX500 workstation, equipped with an ED40 pulsed amperometric detection (PAD) system. A triple-pulse amperometric waveform (E_1 0.1 V, E_2 0.7 V, E_3 -0.1 V) was used for detection with the gold electrode.³⁴ Analytical separations were performed on a CarboPac PA-100 column (250 x 4 mm; Dionex), using a linear gradient of 0 - 200 mM NaOAc in 100 mM NaOH (1 mL/min).

Mass spectrometry

GLC-EI-MS was performed on a Fisons Instruments GC 8060/ MD 800 system (Interscience BV; Breda, The Netherlands) equipped with an AT-1 column (30 m x 0.25 mm, Alltech Associates Inc.; Illinois, USA), using a temperature gradient (140-240 °C at 4 °C/min).³¹

NMR spectroscopy

¹H NMR spectra, were recorded on a Bruker DRX500 spectrometer (Bijvoet Center, Department of NMR spectroscopy) at a probe temperature of 300K. Samples were exchanged once with 99.9 atom% D₂O (Cambridge isotope laboratories, Inc.; Andover, MA), lyophilised, and dissolved in 650 μL D₂O. ¹H chemical shifts (δ) are expressed in ppm by reference to internal acetone (δ 2.225). 1D ¹H NMR spectra were recorded with a spectral width of 5000 Hz in 16k complex data sets and zero filled to 32k. A WEFT pulse sequence was applied to suppress the HOD signal.³⁵ When necessary, a fifth order polynomial baseline correction was applied. 2D ¹H-¹H TOCSY spectra were recorded using MLEV17 mixing sequences with spin-lock times of 10, 30, 60, 120, and 150 ms. The spin-lock field strength corresponded to a 90° pulse width of about 28 μs at 13 dB. The spectral width in 2D TOCSY experiments was 4006 Hz at 500 MHz in each dimension. 400-1024 spectra of 2k data points with 8-32 scans per t₁ increment were recorded. 2D NMR spectroscopic data were analysed by applying a sinus multiplication window and zero filling to spectra of 4k by 1k dimensions. A Fourier transform was applied, and where necessary, a fifth to fifteenth order polynomial baseline function was applied. All NMR data were processed using in-house developed software (J.A. van Kuik, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University).

HPSEC-MALLS

The molecular weights of the glucans were determined by high-performance size-exclusion chromatography (HPSEC) coupled on-line with a multi angle laser light scattering (MALLS) and differential refractive index detection (RI, Schambeck SDF). The HPSEC system consisted of an isocratic pump, an injection valve, a guard column and a set of two SEC columns in series (Shodex SB806MHQ and TSK gel 6000PW). A Dawn-DSP-F (Wyatt Technology; St. Barbara, CA) laser photometer HeNe (λ = 632.8 nm), equipped with a K5 flow cell, and thermostatted by a Peltier heating system, was used as MALLS detector. Samples were filtered through a 0.45 μm filter (MILLEX) and the injection volume was 220 μL. Na₂SO₄ (0.1 M) was used as eluent at a flow rate of 0.8

mL/min. Pullulan and dextran samples with Mw ranging from 4.10^4 to 2.10^6 Da were used as standards. Determinations were performed in duplicate.

References

- Jolly, L.; Vincent, S. J. F.; Duboc, P.; Neeser, J.-R. *Antonie van Leeuwenhoek*, **2002**, *82*, 367-374.
- Sutherland, I. W. In *Biotechnology of microbial exopolysaccharides*, Cambridge University Press, Cambridge, **1990**; pp 117-125.
- Cerning, J. *FEMS Microbiol. Rev.*, **1990**, *87*, 113-130.
- Shepherd, R.; Rockey, J.; Sutherland, I. W.; Roller, S. J. *Biotechnol.*, **1995**, *40*, 207-217.
- Van Marle, M. E.; Zoon, P. *Neth. Milk Dairy J.*, **1995**, *49*, 47-65.
- Morris, R.; Gidley, M. J.; Murray, E. J.; Powell, D. A.; Rees, D. A. *Int. Biol. Macromol.*, **1980**, *2*, 327-330.
- Dea, I. C. M.; Morris, E. R.; Rees, D. A.; Welsch, E. J.; Price, J. *Carbohydr. Res.*, **1997**, *57*, 249-272.
- Van Kranenburg, R.; Boels, I. C.; Kleerebezem, M.; de Vos, W. M. *Current Opin. Biotechnol.*, **1999**, *10*, 498-504.
- Kleerebezem, M.; van Kranenburg, R.; Tuinier, R.; Boels, I. C.; Zoon, P.; Looijensteijn, E.; Hugenholtz, J.; de Vos, W. M. *Antonie van Leeuwenhoek*, **1999**, *76*, 357-365.
- Edwards, K. J.; Jay, A. J.; Colquhoun, I. J.; Morris, V. J.; Dasso, M. J.; Griffin, A. M. *Microbiology*, **1999**, *145*, 1499-1506.
- Kralj, S.; van Geel-Schutten, G. H.; Faber, E. J.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Biochemistry*, **2005**, *44*, 9206-9216.
- De Vuyst, L.; Degeest, B. *FEMS Microbiol. Rev.*, **1999**, *23*, 153-177.
- Monchois, V.; Willemot, R.-M.; Monsan, P. *FEMS Microbiol. Rev.*, **1999**, *23*, 131-151.
- Funane, K.; Ishii, T.; Matsushita, M.; Hori, K.; Mizuno, K.; Takahara, H.; Kitamura, Y.; Kobayashi, M. *Carbohydr. Res.*, **2001**, *334*, 19-25.
- Guggenheim, B. *Helv. Odontol. Acta*, **1970**, *14*, 89-109.
- Argüello-Morales, M. A.; Remaud-Simeon, M.; Pizzut, S.; Sarçabal, P.; Willemot, R.-M.; Monsan, P. *FEMS Microbiol. Lett.*, **2000**, *182*, 81-85.
- Van Hijum, S. A. F. T.; Kralj, S.; Ozimek, L. K.; Dijkhuizen, L.; van Geel-Schutten, G. H. *Microbiol. Mol. Biol. Rev.*, **2006**, *70*, 157-176.
- Kralj, S.; van Geel-Schutten, G. H.; Dondorff, M. M.; Kirsanovs, S.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Microbiology*, **2004**, *150*, 3681-3690.
- Kralj, S.; Stripling, E.; Sanders, P.; van Geel-Schutten, G. H.; Dijkhuizen, L. *Appl. Environ. Microbiol.*, **2005**, *71*, 3942-3950.
- Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P. (Chapter 3) (to be published)
- Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P. (Chapter 4) (to be published)
- Moulis, C.; Joucla, G.; Harrison, D.; Fabre, E.; Potocki-Veronese, G.; Monsan, P.; Remaud-Simeon, M. *J. Biol. Chem.*, **2006**, *281*, 31254-31267.
- Van Leeuwen, S. S.; Leeftang, B. R.; Gerwig, G. J.; Kamerling, J. P. (Chapter 2) (to be published)
- Van Leeuwen, S. S.; Kralj, S.; Eeuwema, W.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P. (Chapter 5) (to be published)
- Kralj, S.; van Geel-Schutten, G. H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Biotrans.*, **2003**, *21*, 181-187.
- Monchois, V.; Willemot, R.-M.; Monsan, P. *FEMS Microbiol. Rev.*, **1999**, *23*, 131-151.
- Van Geel-Schutten, G. H.; Faber, E. J.; Smit, E.; Bonting, K.; Smith, M. R.; Ten Brink, B.; Kamerling, J. P.; Vliegthart, J. F. G.; Dijkhuizen, L. *Appl. Environ. Microbiol.* **1999**, *276*, 44557-44562.
- Simpson, C. L.; Cheetham, N. W.; Giffard, P. M.; Jacques, N. A. *Microbiology*, **1995**, *141*, 1451-1460.
- Kralj, S.; van Geel-Schutten, G. H.; van der Maarel, M. J. E. C.; Dijkhuizen, L. *Microbiology*, **2004**, *150*, 2099-2112.
- Ciucanu, I.; Kerek, F. *Carbohydr. Res.* **1984**, *131*, 209-217.
- Kamerling, J. P.; Vliegthart, J. F. G. In *Clinical Biochemistry – Principles, Methods, Applications. Vol. 1, Mass Spectrometry*; Lawson, A. M., Ed.; Walter de Gruyter: Berlin, **1989**; pp 176-263.
- Jansson, P.-E.; Kenne, L.; Liedgren, H.; Lindberg, B.; Lönnngren, J. *Chem. Commun. Univ. Stockholm*, **1976**, *8*, 1-74.
- Hay, G. W.; Lewis, B. A.; Smith, F. *Methods Carbohydr. Chem.*, **1965**, *5*, 357-360.
- Lee, Y. C. *Anal. Biochem.*, **1990**, *189*, 151-162.
- Hård, K.; van Zadelhoff, G.; Moonen, P.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.*, **1992**, *209*, 895-915.

Appendices

*In the beginning the Universe was created. This has made a lot of people very angry
and been widely regarded as a bad move*

-Douglas Adams-

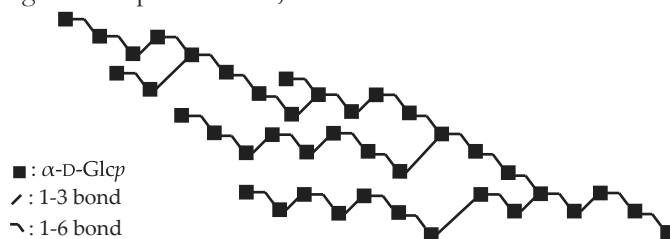
Summary

Microbial exopolysaccharides play an important role in the food industry, particularly in fermented dairy formulations. Lactic acid bacteria have been widely used to improve shelf life and texture of dairy products for thousands of years, thus earning the status as generally recognised as safe (GRAS). In recent years polysaccharides have been implicated for other applications as well. These applications range from health-agents (e.g. preventing cancer, stimulating the immune system, or biodegradable surgical materials) to heavy-industrial applications (e.g. ultra-fine ore flocculants, anti-corrosive agent, anti-fouling agent, or electro-conductive component of fuel-cells).

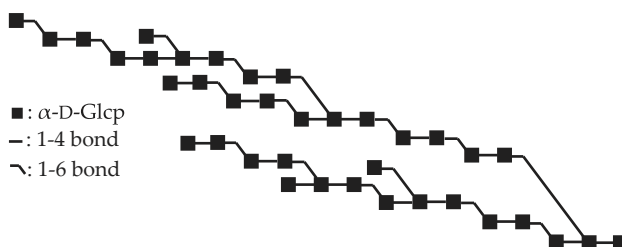
As part of an EET-programme (Economy, Ecology, Technology) named “Bioprimer” (EETK01129), α -D-glucans produced from sucrose by *Lactobacillus reuteri* glucansucrase enzymes were identified as possible anti-corrosive pigments for heavy-duty coatings. These exopolysaccharides were also implicated as possible pre-biotic additives and can be produced in high quantity. The aim of the study described in this thesis was to perform detailed structural analysis of different α -D-glucans produced by native and engineered glucansucrase enzymes.

In Chapter 2 a set of di- and trisaccharide standards was subjected to 1D/2D ^1H and ^{13}C NMR spectroscopy, to analyse the chemical shift patterns of differently substituted glucose residues. From these analyses a structural-reporter-group concept was established, to help analyse the α -D-glucans and oligosaccharides obtained by partial acid hydrolysis and enzymatic hydrolysis of the glucans, as described in Chapters 3-6.

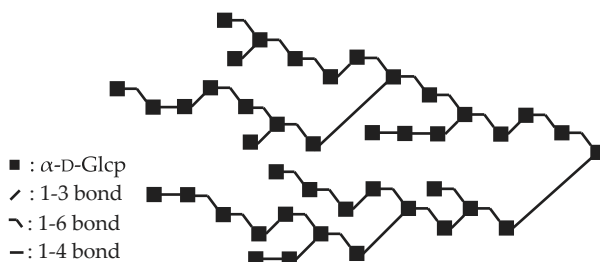
Chapter 3 describes the structural analysis of the exopolysaccharide (**EPS180**) produced from sucrose by the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180. ^1H NMR data of the intact EPS, as well as from fragments obtained by partial acid hydrolysis, were interpreted using the structural-reporter-group concept established in Chapter 2. The ^1H NMR data were combined with data from 2D ^{13}C - ^1H HSQC, methylation analysis, monosaccharide analysis, and Smith degradation, resulting in the following visual representation, that includes all identified structural features:



In Chapter 4 the structural analysis of **EPS35-5**, produced by the glucansucrase GTFA enzyme from *Lactobacillus reuteri* strain 35-5 using sucrose as substrate, is described. A visual representation, including all identified structural elements (below) of the native **EPS35-5** α -D-glucan was generated by combining ^1H NMR data of **EPS35-5** with methylation analysis data, as well as with structural information of oligosaccharides obtained by partial acid hydrolysis and pullulanase M1 (*Klebsiella planticola*) hydrolysis of **EPS35-5**.

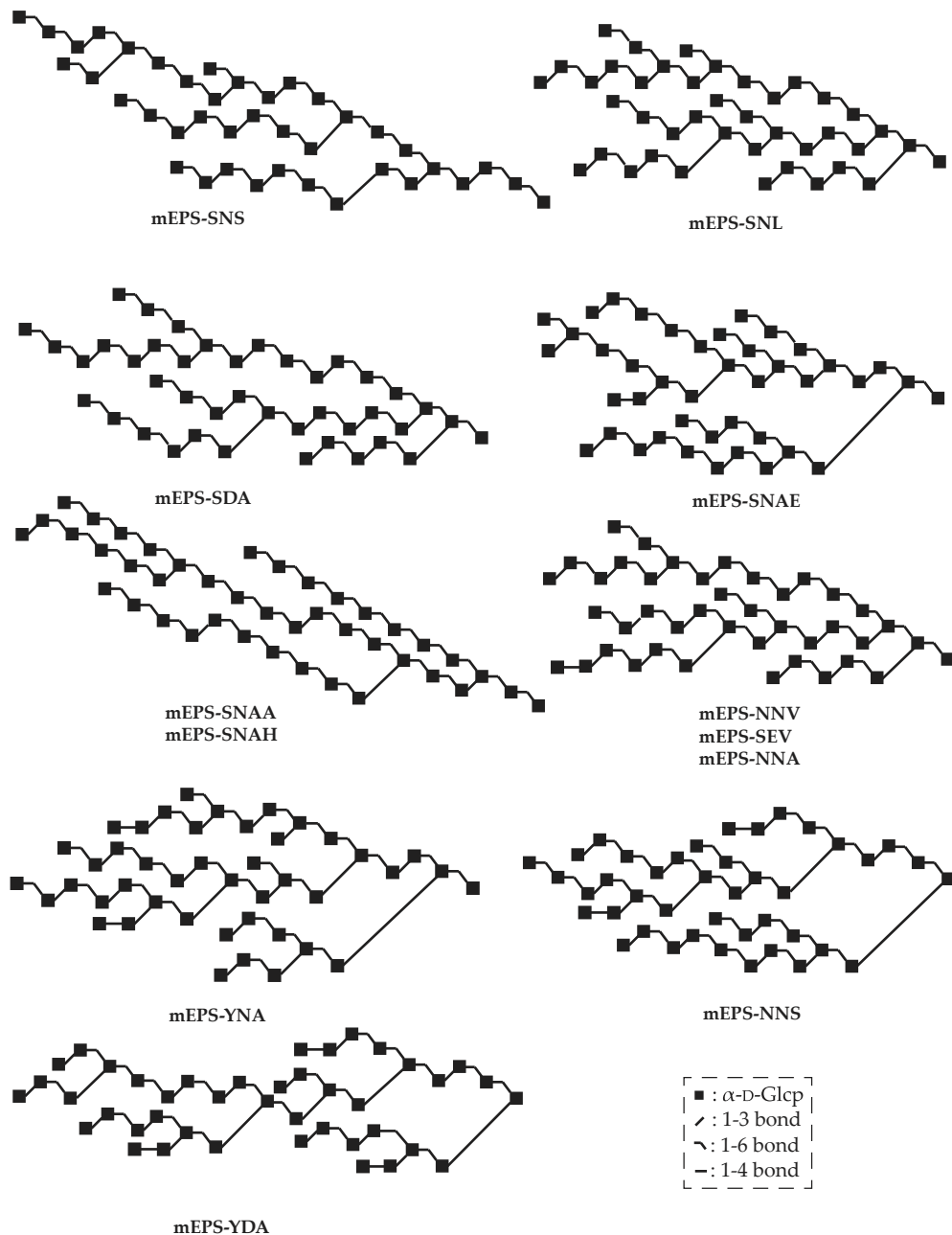


The α -D-glucan (**mEPS-PNNS**) produced from sucrose by the triple mutant V1027P:S1137N:A1139S of the glucansucrase GTF180 enzyme was analysed as described in Chapter 5. Beside (α 1-3) and (α 1-6) linkages, as present in wild type **EPS180**, **mEPS-PNNS** also contained (α 1-4) linkages. Based on linkage analysis, Smith degradation, and 1D/2D ^1H NMR spectroscopy of the intact **mEPS-PNNS**, as well as MS and NMR analysis of oligosaccharides obtained by partial acid hydrolysis of **mEPS-PNNS**, a visual representation, that includes all identified structural elements, was formulated as follows:



Finally, in Chapter 6, the tripeptide S1137:N1138:A1139 (SNA) sequence following the catalytic D1136 (putative transition state stabilising residue) of the glucansucrase GTF180 enzyme was modified into unique sequences found in mutansucrase, alternansucrase, and reuteransucrase enzymes, as well as the common SEV sequence and single mutant variations on these sequences. A fourth residue (Q1140), which has not been targeted before, was also modified. The mutant glucansucrase enzymes were expressed and used to produce mEPSs from sucrose. The structural-reporter-group concept was applied to ^1H NMR analyses of the α -D-glucans. Linkage analysis, Smith

degradation, and 1D/2D ^1H NMR spectroscopic analyses of the mEPSs resulted in visual representations to be formulated (below), showing a wide diversity in structural elements. Analysis of mEPSs by size-exclusion chromatography coupled with MALLS showed a variety of average molecular weights.



*Having a child is surely the most beautifully irrational act that two people in love
can commit*

-Bill Cosby-

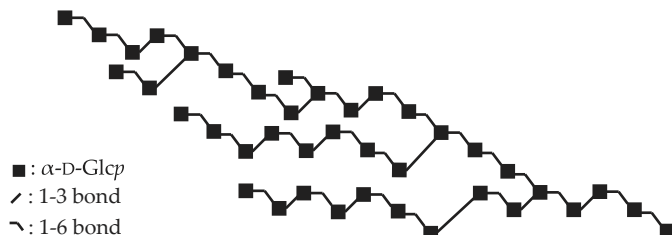
Samenvatting

Microbiële exopolysachariden spelen een belangrijke rol in de voedingsmiddelenindustrie, in het bijzonder in gefermenteerde zuivelformuleringen. Melkzuurbacteriën vinden al duizenden jaren brede toepassing in het verbeteren van de houdbaarheid en de textuur van zuivelproducten, waarmee ze de status als Algemeen Beschouwd als Veilig (“generally recognised as safe”, GRAS) hebben verdiend. In recente jaren zijn polysachariden tevens aangeduid voor andere toepassingen. Deze toepassingen variëren van gezondheidsbevorderaars (bijv. het voorkomen van kanker, stimuleren van het immuun-systeem, bio-afbreekbare chirurgische materialen) tot toepassingen in de zware industrie (bijv. flocculatie van ultra-fijne ertsdeeltjes, anti-corrosieve pigmenten, anti-fouling agentia, of electro-conductieve componenten in katalytische cellen).

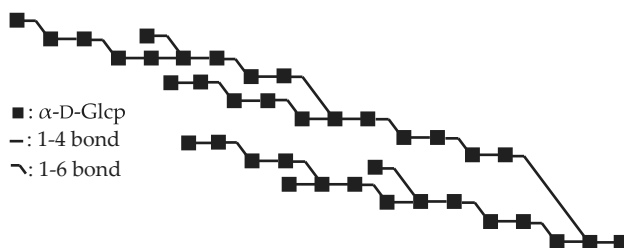
Als onderdeel van een EET-programma (Economie, Ecologie, Technologie) “Bioprimer” genaamd (EETK01129), zijn α -D-glucanen geproduceerd uit sucrose door *Lactobacillus reuteri* glucansucrase enzymen, geïdentificeerd als mogelijke anti-corrosieve pigmenten voor “heavy-duty” coatings. Deze exopolysachariden zijn ook aangemerkt als mogelijke pre-biotische additieven en kunnen in grote hoeveelheden worden geproduceerd. De doelstelling van de studie beschreven in dit proefschrift was het uitvoeren van gedetailleerde structuuranalyses van verschillende α -D-glucanen geproduceerd door natieve en gemodificeerde glucansucrase enzymen.

In Hoofdstuk 2 is een verzameling di- en trisacharide standaarden onderworpen aan 1D/2D ^1H en ^{13}C NMR spectroscopie, om zo de chemical shift patronen van verschillend gesubstitueerde glucose residuen te analyseren. Uit deze analyses werd een structural-reporter-group concept geformuleerd, om te helpen bij de analyse van de α -D-glucanen en oligosachariden verkregen door partiële zure en enzymatische hydrolyse van de glucanen, zoals beschreven in de Hoofdstukken 3-6.

In Hoofdstuk 3 staat de structuuranalyse beschreven van het exopolysacharide (**EPS180**) geproduceerd uit sucrose door het glucansucrase GTF180 enzym uit de *Lactobacillus reuteri* 180 stam. ^1H NMR data van zowel het intacte EPS, alsook van fragmenten verkregen door partiële zure hydrolyse, zijn geïnterpreteerd met behulp van het structural-reporter-group concept zoals vastgesteld in Hoofdstuk 2. De ^1H NMR data zijn gecombineerd met data uit 2D ^{13}C - ^1H HSQC, methyleringsanalyse, monosacharide analyse en Smith degradatie, met als resultaat de volgende visuele representatie, waarin alle geïdentificeerde structurele facetten zijn opgenomen:

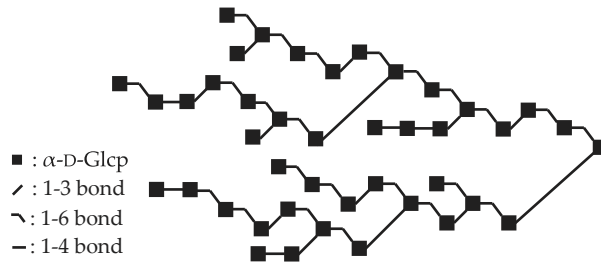


In Hoofdstuk 4 staat de structuuranalyse van **EPS35-5**, geproduceerd door het glucansucrase GTFA enzym uit de *Lactobacillus reuteri* 35-5 stam met sucrose als substraat, beschreven. Een visuele representatie (zie beneden), die alle geïdentificeerde structurelelementen bevat, van het natieve **EPS35-5** α -D-glucaan is gegenereerd door het combineren van ^1H NMR data van **EPS35-5** met methyleringsanalyse data, alsmede de structuurinformatie uit oligosachariden verkregen door partiële zure hydrolyse en pullulanase M1 (*Klebsiella planticola*) hydrolyse van **EPS35-5**.

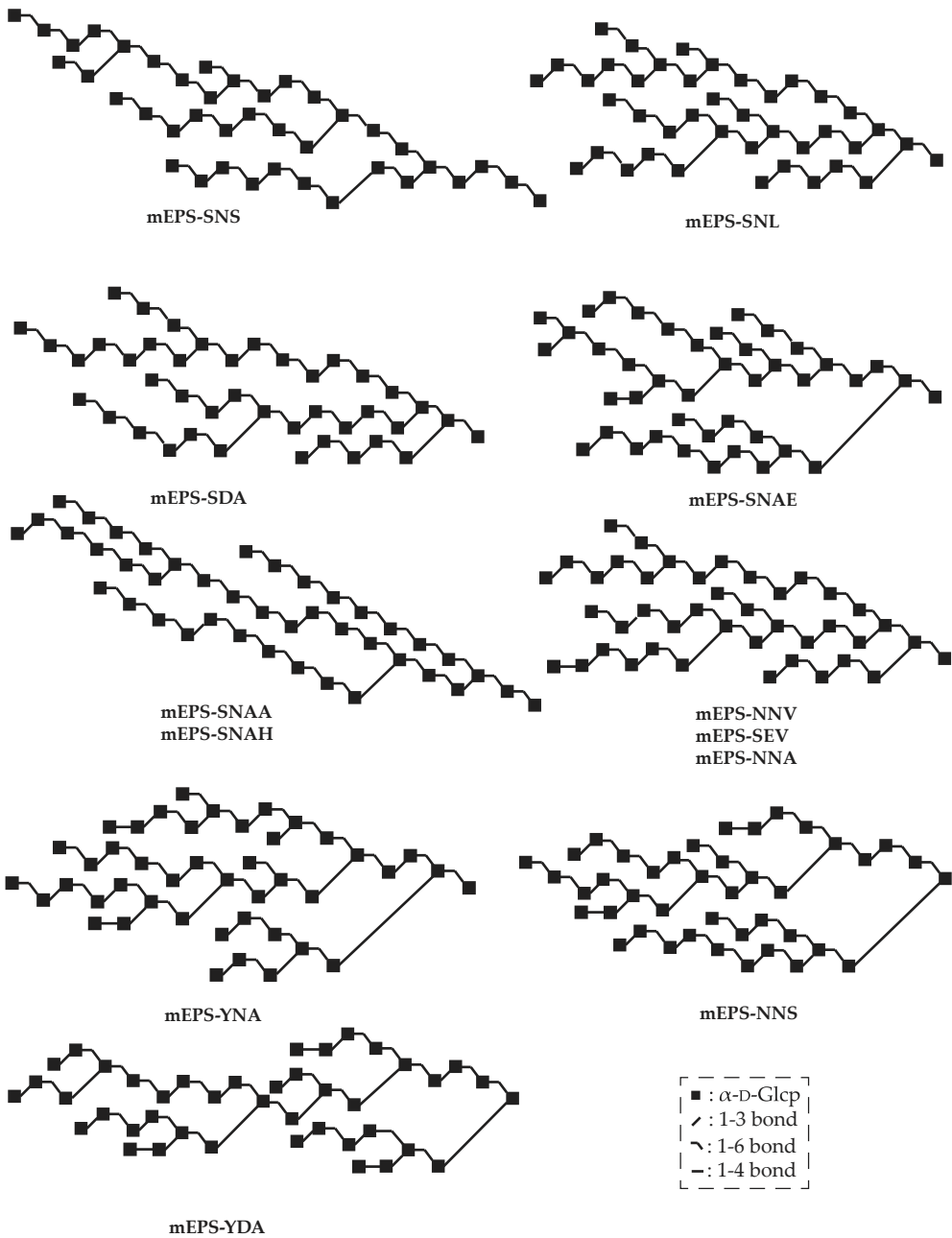


Het α -D-glucaan (**mEPS-PNNS**) geproduceerd uit sucrose door de tripelmutant V1027P:S1137N:A1139S van het glucansucrase GTF180 enzym is geanalyseerd zoals beschreven in Hoofdstuk 5. Naast (α 1-3) en (α 1-6) bindingen, zoals aanwezig in het wild type **EPS180**, bevatte **mEPS-PNNS** tevens (α 1-4) bindingen. Gebaseerd op methyleringsanalyse, Smith degradatie, en 1D/2D ^1H NMR spectroscopie van het intacte **mEPS-PNNS**, alsook MS en NMR analyse van oligosachariden verkregen

door partiële zure hydrolyse van **mEPS-PNNS**, is een visuele representatie waarin alle geïdentificeerde structurelementen zijn opgenomen als volgt geformuleerd:



Tot slot, in Hoofdstuk 6, is de tripeptide S1137:N1138:A1139 (SNA) sequentie volgend op de catalytische D1136 (vermeend overgangstoestand stabiliserend residu) van het glucansucrase GTF180 enzym gemodificeerd naar de unieke sequenties zoals gevonden in mutansucrase, alternansucrase en reuteransucrase enzymen, alsmede de vaak voorkomende SEV sequentie en enkel-mutant variaties op deze sequenties. Een vierde aminozuur (Q1140) die nog niet eerder is geselecteerd, is eveneens gemodificeerd. De mutant glucansucrase enzymen zijn tot expressie gebracht en gebruikt om mEPS te produceren uit sucrose. Het structural-reporter-group concept is toegepast op ^1H NMR analyses van de α -D-glucanen. Methyleringsanalyse, Smith degradatie, en 1D/2D ^1H NMR spectroscopische analyses van de mEPSsen resulteerden in de formulering van visuele representaties (zie beneden), die een brede variatie in structurelementen tonen. Analyse van de mEPSsen met gelpermeatiechromatografie, gekoppeld aan MALLS, toonde een variatie in de gemiddelde molecuul gewichten.



If you think in terms of a year, plant a seed; if in terms of ten years, plant trees; if in terms of 100 years, teach the people

-Confucius-

Deze samenvatting is speciaal bedoeld voor mensen die na de eerste paar zinnen moesten opgeven. Dat is geen schande, want de mensen die het wel kunnen begrijpen, hebben daar eerst jaren voor gestudeerd.

Samenvatting in normaal Nederlands

Wat is suiker?

Deze vraag lijkt heel simpel te beantwoorden. Suiker is van dat spul dat je in de koffie doet, of met kaneel op een beschuitje; lekker en zoet. Hoewel dat spul inderdaad suiker is, is suiker meer dan alleen die kristalletjes. Die kristalletjes zijn een bepaalde soort suiker die sucrose genoemd wordt. De meeste mensen kennen ook meer soorten suiker dan alleen sucrose. Lactose bijvoorbeeld, vooral bekend uit melk en melkproducten; fructose, ook wel vruchtsuiker of fruitsuiker genoemd; en ook glucose, een van de belangrijke bestanddelen van de meeste frisdranken. Glucose en fructose zijn monosacchariden, ofwel enkele suikers, wat aangeeft dat het enkele bouwblokjes zijn. Sucrose en lactose zijn disacchariden; suikers opgebouwd uit twee bouwblokjes. Sucrose is opgebouwd uit een glucose en een fructose; lactose is opgebouwd uit glucose en galactose, een ander bouwblokje, dat hier niet van belang is. Om de zaken simpel te houden hebben we het daar dan ook verder niet over. Voor het werk in dit proefschrift is vooral glucose van belang, en ook sucrose wordt hier en daar genoemd.

Bouwen met suiker

Suiker bouwblokjes kunnen aan elkaar worden vastgemaakt op verschillende koppelplaatsen. Op zo'n koppelplaats kan het ene bouwblokje aan het andere gemaakt worden. Glucose heeft 5 van die plaatsen, die 1, 2, 3, 4, en 6, genummerd zijn. Nummer 5 is overgeslagen, maar dat heeft een chemische reden die voor deze samenvatting niet relevant is. Het is als een haven die 6 meerpalen heeft, maar aan meerpaal 5 zit altijd de boot van de kustwacht, waardoor die niet beschikbaar is.

De voorgenoemde koppelplaatsen kan je vergelijken met klittenband. Koppelplaats 1 heeft de haakjes, en 2, 3, 4, en 6 hebben dat wollige. Je kan dus nooit een glucose met

klittenband 2 aan de klittenband 3 van een andere plakken. Je kan wel een glucose met 1 plakken aan de 2 van een andere glucose, dat heet een (1-2)-binding. Je begrijpt dat er dan heel wat mogelijkheden zijn om structuren te bouwen met deze glucose blokjes.

Enzymen

Enzymen zijn kleine fabriekjes die in elke cel van elk levend wezen voorkomen. Er zijn er heel veel verschillende soorten van. Sommige enzymen maken suikerstructuren. De klittenband van suiker plakt namelijk niet vanzelf aan elkaar, zoals echte klittenband, maar moet ertoe gedwongen worden. De enzymen die met suikers werken kunnen dit. De meeste enzymen bouwen slechts met één bouwblokje en kunnen maar één soort binding maken.

Melkzuurbacteriën en polysachariden

In het onderzoek beschreven in dit proefschrift spelen melkzuurbacteriën ook een rol. We gebruiken deze bacteriën om bijvoorbeeld yoghurt te maken. De melkzuurbacteriën maken een goedje dat zorgt dat de yoghurt niet zo snel bederft en dat tevens de zurige smaak van yoghurt veroorzaakt. Echter, het verschil tussen yoghurt en de melk waar het van gemaakt wordt, is niet alleen dat yoghurt zuur is; yoghurt is ook dik. Dit wordt onder andere veroorzaakt door polysachariden, zo noem je grote ketens van suiker bouwblokjes.

Er zijn twee groepen polysachariden die een melkzuurbacterie kan maken; heteropolysachariden en homopolysachariden. Heteropolysachariden zijn opgebouwd uit meer dan één soort bouwblokje; homopolysachariden bestaan maar uit één soort bouwblokje. Onder het kopje “Enzymen” had ik al genoemd dat de meeste enzymen maar een soort bouwblokje kunnen gebruiken en die maar in een soort binding kunnen vastmaken. Bijvoorbeeld een glucose met een (1-3)-binding vast aan een galactose. Dat betekent dat een heteropolysacharide in een melkzuurbacterie door meerdere enzymen gebouwd moet worden. De enzymen zetten ieder op hun beurt een bouwblokje in vastgestelde volgorde vast en maken zo een repeterende eenheid. De repeterende eenheden worden door een enzym naar buiten getransporteerd en daar door nog een ander enzym aan elkaar geplakt tot een lange reeks herhalingen van hetzelfde element.

Homopolysachariden worden doorgaans op een andere manier gemaakt. Hiervoor maakt de bacterie een enzym, dat aan de buitenkant van de bacterie wordt vastgezet. Als de bacterie dan in een omgeving zwemt waar sucrose aanwezig is, maakt dit enzym van die sucrose een polysacharide. Dit doet het enzym door de glucose en fructose van

elkaar te splitsen. Sommige enzymen gebruiken dan de fructose en bouwen daar een lange keten mee, stap voor stap. Anderen gebruiken de glucose en bouwen een lange glucose keten. Wat bijzonder is aan deze enzymen is dat ze vaak meer dan een soort binding kunnen maken, bijvoorbeeld (1-3) en (1-6), en maken vaak ook vertakkingen.

Glucanen

Nu zijn we aangekomen bij het onderwerp van mijn onderzoek: glucanen. Dit zijn polysachariden gebouwd van alleen glucose. Het zijn dus homopolysachariden. Tot nu toe zijn de precieze structuren van heteropolysachariden al vaak opgehelderd. Met onderzoek kon men de repeterende eenheid achterhalen en meten hoe groot de polysacharide was. Dan kan een polysacharide beschreven worden door te vertellen welke repeterende eenheid gevonden is, en hoe veel herhalingen er in de heteropolysacharide aanwezig zijn. De homopolysachariden hebben geen repeterende structuur en kunnen dus niet op zo'n manier beschreven worden.

Structuuranalyse

Omdat je een polysacharide niet even onder de microscoop kan leggen om te bekijken hoe het in elkaar zit, moet je andere trucjes en indirecte meetmethodes toepassen. De precieze methode van de structuuranalyse gaat te ver om hier helemaal uit te leggen. De belangrijkste techniek die we hebben toegepast, heet NMR spectroscopie. Bij NMR spectroscopie plaats je wat je bestuderen wil in een sterk magneetveld. Nu zit er naast elke koppelplaats van een suiker bouwblokje een klein tolletje met een stip midden op de bovenkant. Normaal ligt zo'n tolletje gewoon niets bijzonders te doen. In dit magneetveld gedragen de tolletjes zich wat anders, ze gaan allemaal rechtop staan tollen. Kijkend vanaf opzij zie je dan alleen een tol, kijkend vanaf boven zou je een cirkel met een stip in het midden zien. Vervolgens passen we een trukje met een tweede magneetveld toe waardoor de tolletjes een beetje vlakker gaan draaien, dan kan je het stipje op de bovenkant wel zien als je van de zijkant kijkt, en je ziet het een cirkelbeweging trekken vanaf de bovenkant. En omdat je het stipje kan zien bewegen, kan je ook bepalen hoe snel de tol rondjes draait. Nu draait iedere tol naast zo'n koppelplaats met een andere snelheid. Dus als je glucose in de NMR meet, krijg je voor de 6 koppelplaatsen een getal dat de draaisnelheid uitdrukt, waarbij we ook kunnen uitpuzzelen bij welke koppelplaats welk getal hoort.

Zo hebben we een patroon met getallen voor de koppelplaatsen van glucose. Maar als je twee glucose bouwblokjes aan elkaar plakt, dan veranderen een paar van

die draaisnelheden. Het patroon van de draaisnelheden is afhankelijk van de binding die een glucose heeft, echter ook op welke koppelplaatsen andere glucoses zitten, zoals 4 en 6. Een glucose die een (1-3)-binding maakt, heeft een ander patroon dan eentje die een (1-4)-binding maakt. Het is nog ingewikkelder, want een glucose die een (1-3)-binding maakt en zelf op koppelplaats 4 bezet is, heeft een ander patroon dan een glucose die een (1-3)-binding maakt en op koppelplaats 6 bezet is. Eerst hebben we deze patronen onderzocht (Hoofdstuk 2). Met behulp van de patronen konden de structuren van twee homopolysacchariden worden opgehelderd (Hoofdstuk 3, **EPS180**; Hoofdstuk 4, **EPS35-5**).

Het enzym verantwoordelijk voor het bouwen van **EPS180** heeft een holte waar sucrose wordt gesplitst in glucose en fructose, en waar vervolgens met glucose de polysaccharide gebouwd wordt. Die holte is vrij krap en de manier waarop de bouwblokjes erin passen is beperkt. Dit maakt dat de bouwblokjes alleen dusdanig passen dat er alleen (1-3)- en (1-6)-bindingen kunnen worden gemaakt. Projectpartners in Groningen hebben 13 aangepaste versies van dit enzym gemaakt, met net een andere vorm holte. Dit heeft invloed op de polysaccharide die gemaakt kan worden. De structuren van deze polysacchariden zijn met dezelfde methoden opgehelderd (Hoofdstukken 5 en 6).

Hoe zit dat nu met die roest?

De meesten van jullie hebben begrepen dat we dit onderzoek gedaan hebben, omdat de polysaccharide iets tegen roesten deed. In de gepresenteerde hoofdstukken komt dit helemaal niet terug. Dat klopt, het polysaccharide uit Hoofdstuk 3 (**EPS180**) werd interessant gevonden, omdat het roest kon voorkomen in een simpele proef. Het polysaccharide uit Hoofdstuk 4 (**EPS35-5**), dat in veel opzichten op **EPS180** lijkt, had dit effect helemaal niet. Het is dan interessant om te kijken wat het verschil in structuur is, zodat je wellicht kan achterhalen waardoor het anti-corrosieve effect veroorzaakt wordt.

Nu we de structuren van deze twee polysacchariden hebben bepaald, is er nog geen definitief antwoord. Wel zien we verschillende structuren en daarmee kunnen we hypothesen stellen. De (1-6)-bindingen hebben een knik erin, waardoor ze meer flexibel zijn dan (1-4)- en (1-3)-bindingen. Gezien **EPS180** vooral (1-6)-bindingen heeft, kan het zijn dat deze flexibiliteit belangrijk is voor de activiteit. Doordat ze door die flexibiliteit makkelijker kunnen draaien voor een gunstige aanhechting aan het metaaloppervlak, bijvoorbeeld. Dit blijven hypothesen, want we hebben nog niet genoeg informatie om deze vragen te beantwoorden. **EPS180** en **EPS35-5**, die in eerste instantie veel met

elkaar gemeen leken te hebben, blijken onvergelykbare structuren hebben. Om de vraag van de activiteit beter te beantwoorden, moeten we dus kijken naar verschillen in anti-corrosieve activiteit tussen polysachariden die wat meer op **EPS180** lijken.

Het is dus heel interessant om de anti-corrosieve werking van de gemodificeerde polysachariden uit Hoofdstukken 5 en 6 ook te bepalen. Helaas had het project daartoe geen tijd meer. Inmiddels zijn een paar chemische behandelde varianten op **EPS180** gemaakt die prima lijken te werken in een coating. Daarmee is het project verder gegaan om verf te mengen. Terwijl ik dit schrijf, zijn er proefvlakken blootgesteld aan de elementen, om te kijken of de nieuwe coatingsystemen met toegevoegde polysacharide inderdaad goed werken. Als dat zo is, en ze doen het net zo goed als de traditionele coating met het vervuilende zink erin, dan is het project op zich geslaagd, ook al weten we nog niet hoe het precies werkt. Het onderzoek in dit proefschrift is ook geslaagd, omdat we polysacharide structuren hebben opgehelderd die voorheen nog niet in zoveel detail bestudeerd zijn.

*Build a man a fire, and he'll be warm for a day. Set a man on fire, and he'll be warm
for the rest of his life*

-Terry Pratchett-

Dankwoord

And so we arrive at the traditional end of the thesis. We talked about science in professional and layman terms. Now it is time for a sort of ‘Thanksgiving’. Doing PhD research and writing a PhD thesis is something that often seems to be a one-man show. Reality is that you cannot do it alone. Therefore I should thank many people. Most acknowledgement chapters are filled with lists of names, taking the risk of forgetting to mention someone. I shall not do this. I would like to thank everyone that was around. The people from BOC and CPC with whom I shared coffee, cake, borrels, and labdays. The people of the Bijvoet AiO platform, with whom I organised the AiO evenings and shared general support tasks around Bijvoet events. The project partners with whom I had regular meetings. Thank you for shaping the environment in which I was capable of performing my research and completing my PhD thesis.

Daar zou ik het natuurlijk bij kunnen laten, iedereen bedankje gehad en klaar ermee. Maar daarmee zou ik een aantal mensen tekort doen. Het is nu eenmaal zo, zonder iemand te willen beledigen, dat sommige mensen meer impact hebben in die vier en een half jaar dan anderen. Mensen die meer betrokken zijn geweest bij het werk, of waarmee je een sterkere band ontwikkeld hebt. Deze mensen wil ik toch nog even apart benoemen.

Te beginnen natuurlijk mijn promotor Hans Kamerling, die mij gered heeft van een leven lang pensioenen uitrekenen. Hoewel dat best even leuk was, heeft de wetenschap toch meer te bieden; meer uitdaging, meer voldoening, meer plezier, om het maar in chronologische volgorde te plaatsen. Maar ook informatieve gesprekken over de interne werking van de Universiteit, wetenschapsfinanciering, en meer leerzame stof, tijdens menig lange auto- of treinreis naar een meeting of congres.

Als tweede Gerrit Gerwig, mijn co-promotor. Met uitspraken als, ‘Wat doen wij daar?!’, ‘Wat is het hier een zoi!’, ‘Kijk nou, allemaal bagger!’, ‘Is er al verse koffie?!’, en ‘Ja, dat weet ik allemaal niet!’ (merk op dat overal een uitroepteken achter staat). En natuurlijk altijd geïnteresseerd in hoe de danswedstrijden gingen. Dit alles droeg bij aan een zeer vermakelijke tijd tussen de experimenten door.

Dan kunnen we Mayken (Klets!) natuurlijk niet vergeten, bijna de hele tijd heb je tegenover me gezeten. Leuke websites, grappige plaatjes en maffe filmpjes werden snel gedeeld. De roddels en geruchten in en rond het lab hoefde ik nooit op te zoeken, op een of andere manier wist jij altijd alles en was graag bereid mij op de hoogte te houden.

Ook adviezen over promotie regelen, proefschrift lay-outen en drukker uitzoeken waren zeer nuttig. Het was voor mij erg handig dat jij een kleine 4 weken eerder dan ik aan de beurt was.

Menig eenzaam uurtje pieken vangen werd verlicht door Koen, die tegenover de deur van het Dionex-hok zijn zuurkasten (jawel, meervoud!) had. Maar niet alleen bij de Dionex heb ik van die gezelligheid kunnen genieten, ook bij borrels en het squashen. We gaan elkaar weer wat vaker treffen in het Noorden.

Dat brengt me meteen bij Lizette en Adriana die ook aan de wekelijkse squash meededen. Daarnaast was Whistler ook erg gezellig. Een mooi, zonnig weekendje voordat het congres begon, om de omgeving te voet, op de mountainbike, en per kano te verkennen.

De laatste anderhalf jaar werd het secretariaat bemand door Walter. Handig als je wat te regelen had, of 's-ochtends gewoon even koffie drinken en nadenken hoe we middels een Ministerie voor Onredelijke zaken het land weer op de rails konden krijgen.

In het bijzonder wil ik ook de projectpartners in Groningen bedanken. De samenwerking met Lubbert Dijkhuizen en Slavko Kralj was uitermate prettig. We hebben hele mooie wetenschap kunnen doen en gelukkig gaan we daar in de toekomst nog even mee door. Ik zou graag denken dat de band ook meer dan puur professioneel is geworden.

Er is meer in het leven dan werken alleen. Je hebt tussendoor ook afleiding en rust nodig om wetenschap te kunnen bedrijven. Even de repeterende experimenten, successen en mislukkingen, achter je laten en het eens heel ergens anders over hebben. Het is ook mooi om goede vrienden te hebben om die rust mee te kunnen delen. Bas en Jolanda, altijd goed voor gezellig een spelletje, of kletsen over de rare dingen die je meemaakt, ziet op televisie of leest in een krant. En natuurlijk met Lara onze voorganger op het avontuurlijke pad van de kindertjes. We hebben gelukkig al veel kunnen afkijken, waardoor de afgelopen weken misschien iets minder zwaar waren dan ze hadden kunnen zijn.

Mijn schoonouders en mijn nieuwste broer, vragend naar het onderzoek, wat je doet, hoe het gaat. Ik hoop dat de "Samenvatting in Normaal Nederlands" nu eindelijk eens de uitleg gaf waar ik eigenlijk nooit toe kwam. Jullie hebben er allemaal toch regelmatig naar gevraagd.

And in Australia, so far away, my brother and practically-sister-in-law. I realise that you came already for our wedding and cannot possibly be here for the defence of

my thesis. It is sad to have to miss you on these occasions. But I hope that we can find time (and money!) to make the trip to Australia soon. I think it is our turn to visit you now, it's been 6 years. Despite the distance you have endeavoured to stay in touch and understand what I was doing in the lab. Jerry also always interested in what my future plans were (secretly hoping it would migrate me to Oz?), and Abbey cooing over the baby and other social achievements.

Mijn lieve ouders wil ik natuurlijk niet vergeten. Een tweepolige opvoeding. De moeder met de zweep, geen genoeg nemen met middelmatigheid, niet te snel opgeven. En de vader van de zachte hand, voor de balans. Maar samen altijd achter mij staan, geloven in wat ik kan, vaak nog voordat ik dat zelf door had. Mijn wetenschappelijke carrière was zonder dit geloof al op de middelbare school in de kiem gesmoord. Maar niet alleen moreel ben ik gesteund, ook financieel hebben jullie bijgedragen aan het boekje waar dit in te lezen staat.

Op de dag zelf word ik natuurlijk bijgestaan door twee Paranimfen. Robert, broertje, motormuis. Je wilde graag een manuscript en kon niet wachten op meer informatie over de taken die je als Paranimf te wachten stonden. Je hebt me ook immens geholpen door het leveren van de software waarmee de lay-out en omslag design een koud (of eigenlijk louw) kunstje waren. Dat het boekje er zo mooi uitziet is ook deels aan jou en je grafische adviezen te danken! En Eldert als goede vriend. Ik ben blij dat jij ook bij de verdediging naast me zal staan. Als de agenda's niet te druk waren wisten we toch meestal elkaar wel met regelmaat te treffen. Even samen eten, een spelletje, praten over zinnige of onzinnige onderwerpen. Om een of andere reden meestal aan de voordeur bij vertrek, we hebben daar dan inmiddels ook een stoel staan.

En tot slot dan mijn lieve Jozefien. Laatst had ik het er al over, hoe snel dingen soms gaan. Precies 6 jaar geleden was ik net afgestudeerd. Unhappy single, geen richting in carrière en dus ook maar even fysiek verdwaald aan de andere kant van onze planeet. Ik had toen nooit kunnen denken dat ik 6 jaar later zoveel bereikt zou hebben. Ik werd van vriendje man, van man vader, van huurder huizenbezitter en van doctorandus straks dus doctor. En zonder jouw steun en liefde had ik niet één van deze mijlpalen kunnen bereiken. Jij maakt dat ik kan zijn wie ik wilde worden, iets waarvan ik 6 jaar geleden slechts kon dromen.

Bibliography

Structural characterisation of α -D-glucans produced by mutant glucansucrase GTF180 enzymes of *Lactobacillus reuteri* strain 180.

Van Leeuwen, S. S.; Kralj, S.; Eeuwema, W.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P., manuscript in preparation (Chapter 6, this thesis).

Structural analysis of the bio-engineered α -D-glucan (mEPS-PNNS) produced by a triple mutant of the glucansucrase GTF180 enzyme from *Lactobacillus reuteri* strain 180.

Van Leeuwen, S. S.; Kralj, S.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P., manuscript in preparation (Chapter 5, this thesis).

Structural analysis of the α -D-glucan (EPS35-5) produced by the *Lactobacillus reuteri* strain 35-5 glucansucrase GTFA enzyme.

Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P., manuscript in preparation (Chapter 4, this thesis).

Structural analysis of the α -D-glucan (EPS180) produced by the *Lactobacillus reuteri* strain 180 glucansucrase GTF180 enzyme.

Van Leeuwen, S. S.; Kralj, S.; van Geel-Schutten, G. H.; Gerwig, G. J.; Dijkhuizen, L.; Kamerling, J. P., manuscript in preparation (Chapter 3, this thesis).

Development of an NMR structural-reporter-group concept for the primary structural characterisation of α -D-glucans.

Van Leeuwen, S. S.; Leeftang, B. R.; Gerwig, G. J.; Kamerling, J. P., manuscript in preparation (Chapter 2, this thesis).

Molecular cloning and characterization of the alkaline ceramidase from *Pseudomonas aeruginosa* PA01.

Nieuwenhuizen, W. F.; **van Leeuwen, S.**; Jack, R. W.; Egmond, M. R.; Götz, F., *Protein Expr. Purif.*, **2003**, *30*, 94-104.

Synthesis of a novel fluorescent ceramide analogue and its use in the characterization of recombinant ceramidase from *Pseudomonas aeruginosa* PA01.

Nieuwenhuizen, W. F.; **van Leeuwen, S.**; Götz, F.; Egmond, M. R., *Chem. Phys. Lip.*, **2002**, *114*, 181-191.

Curriculum vitae

Sander Sebastiaan van Leeuwen werd op 6 juli 1977 geboren te Rotterdam. In 1996 behaalde hij zijn VWO diploma aan het Newman College in Breda en startte datzelfde jaar zijn studie Scheikunde aan de Universiteit Utrecht. Een bijvak werd gevolgd binnen de sectie Bio-Organische Chemie van het Bijvoet Centrum voor Biomoleculair Onderzoek onder begeleiding van Drs. P. Erbel. In september 2001 behaalde hij zijn doctoraalexamen op basis van onderzoek verricht bij de sectie Enzymology & Protein Engineering onder begeleiding van Dr. W.F. Nieuwenhuizen en Prof.dr. M.R. Egmond. Dit onderzoek betrof de expressie, zuivering en karakterisering van het recombinant ceramidase uit *Pseudomonas earuginosa* PA01. Na een korte zwerftocht door Australië en een 8 maanden durende carrière in de pensioenverzekeringen keerde hij terug tot de wetenschap. Onder begeleiding van Prof.dr. J.P. Kamerling en Dr. G.J. Gerwig werd van april 2003 tot april 2007 promotieonderzoek gedaan bij de sectie Bio-Organische Chemie van het Bijvoet Centrum voor Biomoleculair Onderzoek, zoals beschreven in dit proefschrift. Sinds juli 2007 is hij werkzaam als PostDoc onderzoeker bij de afdeling Microbiologie van het Groningen Biomolecular Sciences and Biotechnology Institute aan de Universiteit Groningen.

