CHEMISTRY
A European Journal

Accepted Article

FULL PAPER

Active Site-Mapping of Xylan-Deconstructing Enzymes with Arabinoxylan Oligosaccharides Produced by Automated Glycan Assembly

D. Senf,[a,b] C. Ruprecht,[a] G. H. M. de Kruijff,[a,b,c] S. O. Simonetti,[a,d] F. Schuhmacher,[a,b] P. H. Seeberger,[a,b] and F. Pfrengle[*,a,b]

Abstract: Xylan-degrading enzymes are crucial for the deconstruction of hemicellulosic biomass, making the hydrolysis products available for various industrial applications such as biofuel production. To determine the substrate specificities of these enzymes, we prepared a collection of complex xylan oligosaccharides by automated glycan assembly. Seven differentially protected building blocks provided the basis for the modular assembly of 2-substituted, 3-substituted, and 2-/3-substituted arabino- and glucuronoxylan oligosaccharides. Elongation of the xylan backbone relied on iterative additions of C4-fluorenylmethoxycarbonyl (Fmoc) protected xylose building blocks to a linker-functionalized resin. Arabinofuranose and glucuronic acid residues have been selectively attached to the backbone using fully orthogonal 2-(methyl)naphthyl (Nap) and 2-(azidomethyl)benzoyl (Azmb) protecting groups at the C2- and C3-hydroxyls of the xylose building blocks. The arabinoxylan oligosaccharides are excellent tools to map the active site of glycosyl hydrolases involved in xylan deconstruction. The substrate specificities of several xylanases and arabinofuranosidases were determined by analyzing the digestion products after incubation of the oligosaccharides with glycosyl hydrolases.

Introduction

Xylans are a major component of the plant cell wall and the second most abundant polysaccharide after cellulose.[1] The xylan backbone consists of β-(1→4)-linked D-xylopyranosides with variable substitution patterns across plant species that result in immense structural diversity. Arabinoxylans are an important xylan subclass where the backbone is substituted with L-arabinofuranosyl residues in 2- and/or 3-position.[2] The backbone may further be acetylated, and the 2-position of xylose may carry few additional D-glucuronic acid or 4-O-methylated D-glucuronic acid substituents.

As a major component of biomass, xylans are an attractive resource for various industrial food and non-food applications including the feed, paper, and biofuel industries. Arabinoxylans from cereals and grasses have attracted great attention as a source for advanced biofuels besides cellulose.[3] Such second generation biofuels do not compete with food production, because they use sustainable feedstocks such as waste biomass (e.g. wheat straw) or plants that grow on lands unsuitable for crop cultivation (e.g. switchgrass).[4] The utilization of xylans as pharmaceuticals, ingredients in functional foods, and agricultural applications is also being investigated. Arabinoxylans in the primary cell wall of grasses and cereals[5] make up a dominant component of dietary fibers in our everyday food, improving serum lipid concentrations, lowering blood pressure, and reducing the risk for heart diseases, diabetes, and obesity.[6] Arabinoxylooligosaccharides (AXOs), that are obtained by chemical or enzymatic hydrolysis of arabinoxylans, exhibit beneficial prebiotic effects and antioxidative activities.[7]

Xylan-degrading enzymes are key to hydrolyze arabinoxylan to AXOs or its component monosaccharides. Therefore, a detailed understanding of xylan deconstruction at the molecular level is necessary.[8] Several xylan-degrading enzymes, including xylanases, β-D-xylosidases, arabinofuranosidases, and esterases, act cooperatively to degrade arabinoxylan polysaccharides. The most important enzymatic activity is hydrolysis of the backbone achieved by endo-β-1,4-xylanases that are produced by many organisms such as bacteria and fungi[9]. Many industrial applications such as bread making, animal feed industry, paper manufacturing, and biofuel production depend on the use of xylanases.[8a, 9-10] Xylanases have been classified into glycosyl hydrolase (GH) families 5, 8, 10 and 11 (Carbohydrate-Active Enzyme database)[11] based on the amino acid sequence in the catalytic domain. Most xylanases belong to the families GH10 and GH11.[12] GH10 xylanases tolerate a higher degree of backbone substitution and release smaller hydrolysis products than GH11 xylanases.[13]

The substrate specificities of recently studied xylanases GH5 and GH8 remain to be determined for the most part. The substrate specificities of GH10 and GH11 xylanases were

[a] Deborah Senf, Dr. Colin Ruprecht, Goswinus H. M. de Kruijff, Dr. Sebastian O. Simonetti, Frank Schuhmacher, Prof. Dr. Peter H. Seeberger, Dr. Fabian Pfrengle
Department of Biomolecular Systems, Max-Planck-Institute of Colloids and Interfaces,
Am Mühlenberg 1, 14476 Potsdam, Germany
E-mail: Fabian.Pfrengle@mpikg.mpg.de

[b] Deborah Senf, Goswinus H. M. de Kruijff, Frank Schuhmacher, Prof. Dr. Peter H. Seeberger, Dr. Fabian Pfrengle
Institute of Chemistry and Biochemistry, Freie Universität Berlin, Arnimallee 22, 14195 Berlin, Germany.

[c] Goswinus H. M. de Kruijff
Institute of Organic Chemistry, Johannes Gutenberg Universität Mainz, Duesbergweg 10-14, 55128 Mainz, Germany (current address).

[d] Dr. Sebastian O. Simonetti
Rosario Chemistry Institute (IQUIR, CONICET-UNR), Suipacha 570, (S2002LRK) Rosario, Argentina (current address).
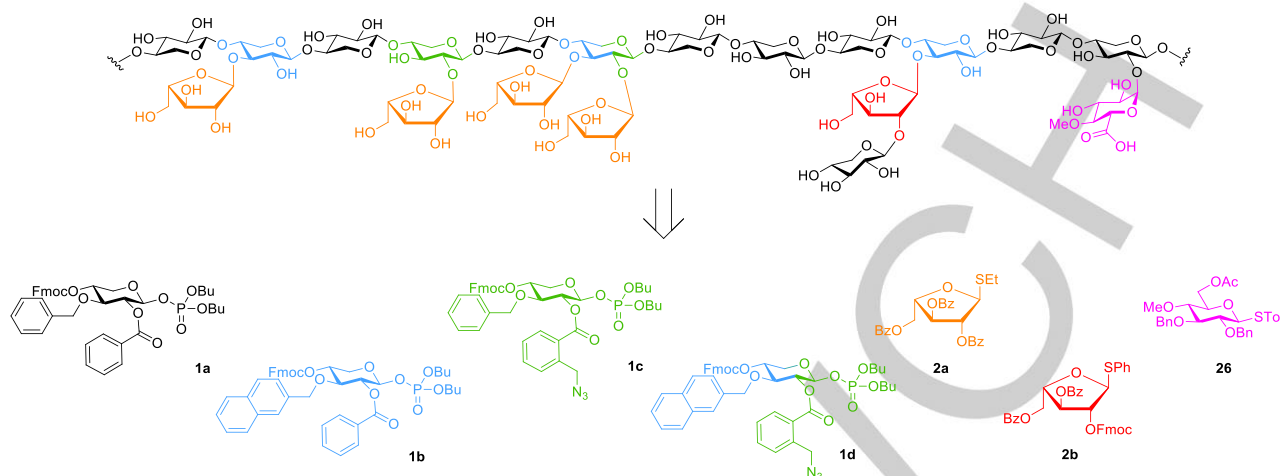
Scheme 1. Chemical structure of an arabinoxylan polysaccharide and the building blocks required for the assembly of representative oligosaccharide fragments.

determined by analysis of the hydrolysis products after incubation with natural xylans.[14] To identify the individual oligosaccharides in the complex hydrolysis mixtures, sophisticated analytical techniques were needed.[13a, 15] These experiments were aided by crystal structures of ligand-enzyme complexes and mutational analysis of potential binding sites.[16] However, all these methods require particular expertise andequipment and are time consuming. A library of tailored oligosaccharides would greatly simplify the analysis of the hydrolysis products, thus enabling the rapid systematic elucidation of the substrate specificities of the enzymes.[17]

Automated glycan assembly[18] can provide rapid access to collections of defined oligosaccharides. We recently assembled structurally related fragments of different cell wall polysaccharide classes from a few monosaccharide building blocks (BBs).[19] The automated glycan assembly of arabinoxylan oligosaccharides containing α-1,3-linked arabinofuranosyl residues enabled us to characterize the binding specificities of several plant cell wall glycan-directed antibodies.[19a] Herein, we describe the synthesis of structurally diverse arabinoxylan oligosaccharides, including 2-substituted, 3-substituted and 2-/3-substituted glycans. The synthetic oligosaccharides proved valuable for the active-site mapping of GH10 endo-xylanases and arabinofuranosidases.

Results and Discussion

The efficient automated glycan assembly of selected oligosaccharide fragments of arabinoxylan requires differentially protected BBs that are readily synthesized and can be employed in a modular fashion. Four BBs were designed for construction of the xylan backbone (Scheme 1). BB **1a** was used for linear chain elongation. It was equipped with a base-labile Fmoc-protecting group in the C4-position and, for permanent protection of the C2- and C3-positions, a benzoyl ester and a

benzyl ether, respectively. In order to enable substitution of the backbone, temporary protecting groups were selected for the hydroxyls at C2 and C3 that ensure similar reactivity of BBs **1b-d** compared to BB **1a**. Exchange of the benzyl for a (2-methyl)naphtyl (Nap) ether in **1b** allows for installation of an arabinose substituent in the C3-position. Nap ethers are readily cleaved under oxidative conditions.[19a, 20] For temporary protection of the C2-hydroxyl the use of a 2-(azidomethyl)benzoyl (Azmb) ester in BB **1c** is particularly attractive. The Azmb-group provides both the required participating effect for selective β-(1→4)-glycosylation and the potential for chemoselective removal using alkyl phosphines. Although the Azmb group has not been used in solid-phase carbohydrate chemistry previously, it has proven to be a powerful protecting group in the solution-phase synthesis of different complex oligosaccharides.[21] Combination of the Nap and Azmb protecting groups allows for the introduction of doubly substituted xylose units when using BB **1d**. All xylose BBs were equipped with phosphate leaving groups that gave the best results in the synthesis of plant cell wall oligosaccharides previously.[19] Finally, arabinofuranose BBs **2a** and **2b** enable substitution of the backbone either with single arabinose units or elongated oligosaccharide side chains.[19a]

The design of the protecting group pattern in BBs **1a-d** allowed for their synthesis to diverge at a late stage (Scheme 2). Starting from 1,2-*O*-isopropylidene-D-xyloside **3** a simple three step transformation gave the 3-*O*-Bn and 3-*O*-Nap protected derivatives **4a,b** in good yields.[19a, 22] After acidic cleavage of the acetonide with concomitant equilibration to the pyranose form, peracetylated **5a,b** were obtained after treatment with acetic anhydride in pyridine. Subsequently, **5a,b** were converted into the corresponding thioglycosides **6a,b** and the acetyl groups were cleaved by methanolysis. An ytterbium triflate catalyzed selective protection of the C2-position[23] using either Bz₂O or Azmb₂O gave the Bz esters **7a,b** and Azmb esters **7c,d**, respectively. The glycosyl donors **1a-d** were finally obtained by
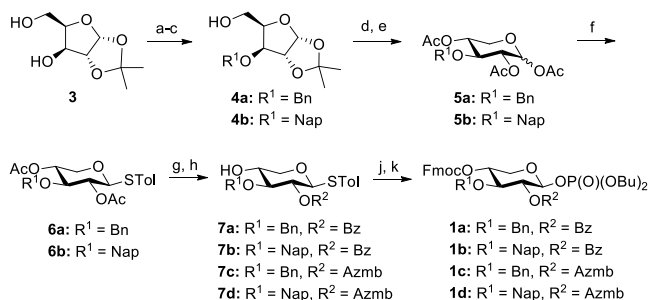
Scheme 2. Synthesis of xylose building blocks **1a-d**. Reagents and conditions: (a) TrCl, DMAP, NEt$_3$, DMF; (b) RBr, NaH, TBAI; (c) TsOH, MeOH/Et$_2$O, H$_2$O (100:10:1) **4a:** 72%, **4b:** 62% (3 steps); (d) H$_2$O/AcOH, reflux; (e) Ac$_2$O, DMAP; (f) HSTol, BF$_3$·OEt$_2$, 0 °C, **6a:** 62%, **6b:** 48% (3 steps); (g) NaOMe, MeOH, CH$_2$Cl$_2$; (h) Bz$_2$O, 5-10 mol% Yb(OTf)$_3$, dioxane, **7a:** 71%, **7b:** 61%, **7c:** 64% , **7d:** 44% (2 steps); (j) FmocCl, pyridine, CH$_2$Cl$_2$, **a:** 79%, **b:** 76%, **c:** 87%, **d:** 68%; (k) HOP(O)(OBu)$_2$, N-iodosuccinimide, triflic acid, **1a:** 87%, **1b:** 95%, **1c:** 60%, **1d:** 63%.

protection of the free hydroxyls at C-4 with FmocCl and introduction of a phosphate ester as the anomeric leaving group. The arabinose BBs **2a** and **2b** were synthesized as reported previously.[24] With all required BBs in hand we produced a collection of 18 xylo- and arabinoxylooligosaccharides using automated glycan assembly (Scheme 3). The selected oligosaccharides represent substructures of natural arabinoxylan polysaccharides including linear oligoxylans (**9-12**), α-1,3-substituted (**13-20**) and α-1,2-substituted oligoarabinoxylans (**22-25**), as well as a structure containing two L-arabinofuranosyl residues attached to the same xylose unit (**21**). We also included two oligosaccharides with a β-1,2-D-xylopyranosyl-α-1,3-L-arabinofuranosyl disaccharide side chain (**16**, **17**) in the target molecules. The biosynthesis of these side chains that are frequently found in grasses was only recently elucidated.[25] Moreover, we synthesized a branched oligosaccharide comprised only of xylopyranoses (**26**), which resembles branched xylans that are present in the mucilage surrounding plant seeds.[26]

The BBs were added to a Merrifield resin equipped with a photocleavable linker[27] (**8**) via alternating cycles of glycosylation and deprotection reactions. Glycosylations with phosphate BBs **1a-d** were originally performed with two times 5 equiv of glycosyl donor and stoichiometric TMSOTf as an activator (module A).[19a] However, we found that two cycles using 1.8 equiv donor each were sufficient to obtain full conversion in the construction of the xylan backbone. Thioglycoside BBs **2a** or **2b** were also reacted using as little as two times 1.8 equiv BB upon activation with stoichiometric NIS and catalytic TfOH (module B). While Fmoc protecting groups are routinely used in automated glycan assembly and readily removed using triethylamine (module C), robust reaction conditions for the removal of the Nap and Azmb protecting groups had yet to be established. After extensive optimization cycles we found that deprotection of the hydroxyl group at 3-OH was accomplished best by treating the resin with a DDQ solution in DCE/MeOH/H$_2$O (64:16:1) at 40 °C in seven

incubation cycles of 20 min (module D).[19a] For deprotection of the Azmb ester several parameters were investigated including reaction temperature, solvent, and the use of different phosphines such as trimethyl, tributyl, and triphenylphosphines. Incubation of the resin with a solution of PBu$_3$ in THF/H$_2$O (5:1) at 45 °C in 6 cycles of 30 min each (module E) gave the best results. Key in this reaction was the use of the right amount of water to hydrolyze potential iminophosphoranes formed during the course of the reaction.[21c]

Having established reliable glycosylation and deprotection conditions for all BBs and protecting groups, we prepared a number of xylan oligosaccharides decorated in different positions with either one or two α-1,3-linked arabinofuranoses.[19a] For short oligosaccharides we first synthesized the full xylan backbone using a combination of xylose BBs **1a** and **1b**, and then added the arabinose substitutions after selective deprotection of the Nap group using BB **2a**. For xylan structures with two substitutions, we started with a shorter backbone including one xylose BB (**1a**) after the designated substitution site. After adding the arabinose residue, we continued with the elongation of the backbone and introduction of the second arabinose substituent. We chose a sequential addition of arabinose substituents as the Nap deprotection conditions were optimized for removal of single Nap groups. Fully deprotected oligosaccharides **13-20** were obtained after automated assembly, light-induced cleavage from the resin, and global deprotection consisting of methanolysis and subsequent hydrogenolysis.

The synthesis of oligosaccharides **16** and **17** containing a β-1,2-D-xylopyranosyl-α-1,3-L-arabinofuranosyl disaccharide side chain required capping of the terminal xylose residue after backbone assembly using acetic anhydride and pyridine (module F). Only then, the Nap group was cleaved to enable glycosylation with arabinose BB **2b**, that contains an Fmoc protecting group for elongation of the side chain. Capping of the backbone was thus needed for the subsequent selective glycosylation of the arabinose with xylose BB **1a**.

This set of α-1,3-substituted glycans was supplemented with α-1,2-substituted arabinoxylan oligosaccharides **22-26** by including 2-O-Azmb-substituted BB **1c** in the starting materials. We found that the product yields in the synthesis of 2-substituted compounds decreased with the number of glycosylations performed after introduction of the Azmb-protected BB **1c**. It was advantageous to remove the Azmb-group at an earlier stage of the assembly, and continue with elongation of the backbone after introduction of the substituent (assembly of **24**). Using a combination of the two orthogonal Nap and Azmb protecting groups in a single BB (**1d**) we achieved the synthesis of pentasaccharide **21**, having two substituents linked to same xylose unit. In order to attach these two arabinofuranose residues, first the Nap ether was deprotected and the resulting free hydroxyl was glycosylated with **2a** before deprotection of the Azmb ester allowed for another glycosylation with **2a**.

To showcase the utility of Azmb protected BB **1c** for the automated glycan assembly of other complex xylan oligosaccharides we synthesized 4-OMe-glucuronic acid substituted oligosaccharide **30** (Scheme 4). We chose a post-
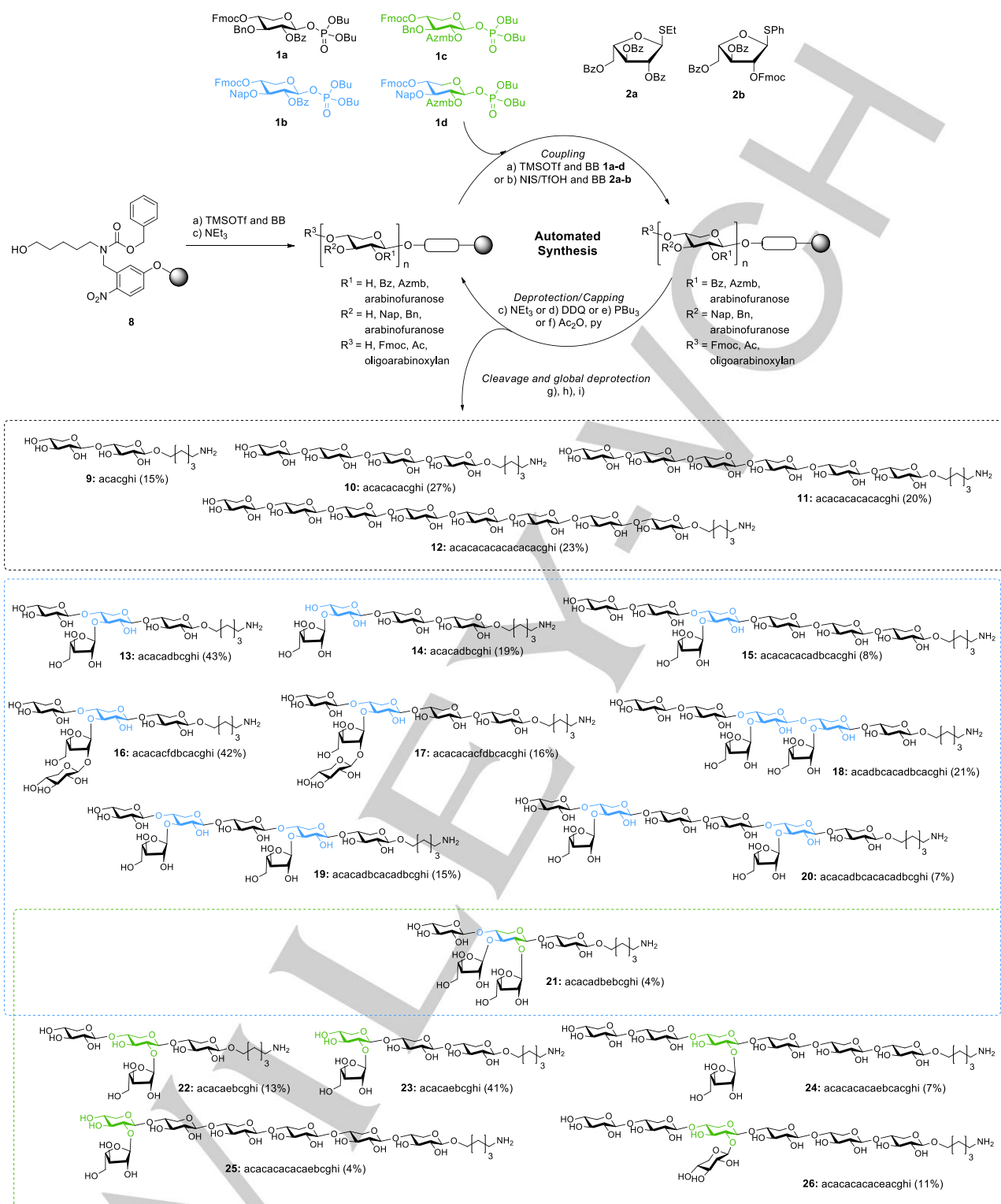
Scheme 3. Automated glycan assembly of α-1,2- and/or α-1,3-substituted arabinoxylan oligosaccharides **9–26**. The letter code below the structures represents the reaction sequence applied in the respective synthesis. Reagents and conditions: a) 2 x 5 equiv, 2 x 3.7 equiv or 2 x 1.8 equiv of BB **1a**, **1b**, **1c**, or **1d**, TMSOTf, DCM, -35 °C (5 min) → -15 °C (30 min) (Module A); b) 2 x 5 equiv or 2 x 1.8 equiv of BB **2a** or **2b**, NIS, TfOH, DCM/dioxane, -40 °C (5 min) → -20 °C (40 min) (Module B); c) 3 cycles of 20% NEt$_3$ in DMF, 25 °C (5 min) (Module C); d) 7 cycles of 0.1 M DDQ in DCE/MeOH/H$_2$O (64:16:1), 40 °C, 20 min (Module D); e) 6 cycles of PBu$_3$, THF/H$_2$O (5:1), 45 °C, 30 min (Module E); f) 3 cycles of Ac$_2$O, pyridine, 25 °C, 30 min (Module F); g) hν (305 nm); h) NaOMe, DCM/MeOH, 12 h; i) H$_2$, Pd/C, EtOAc/MeOH/H$_2$O/HOAc, 12 h (yields are based on resin loading). The letter code below the structures represents the reaction sequence applied in the respective synthesis.
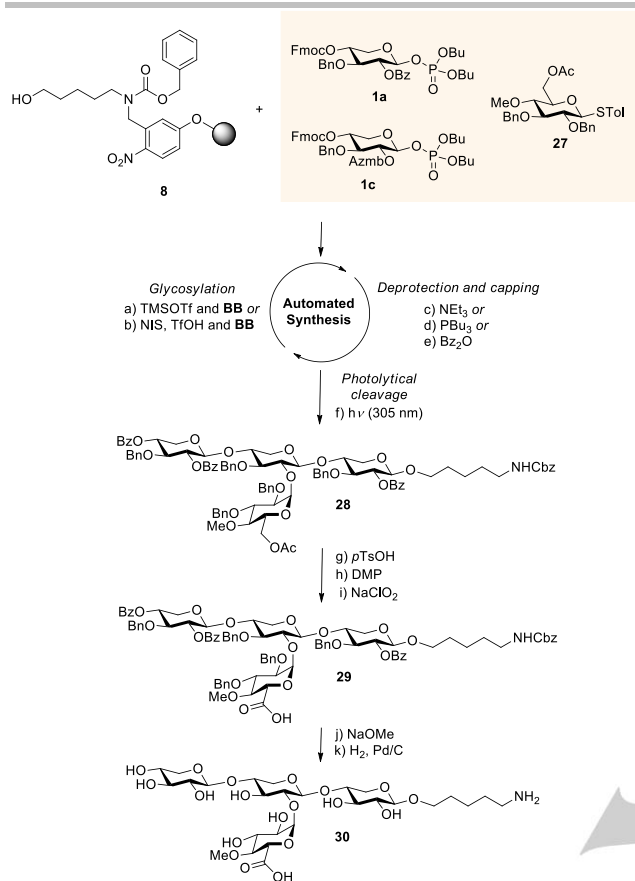
Scheme 4. Automated glycan assembly of glucuronoxylan oligosaccharide **29**. Reagents and conditions: a) 2 x 1.8 equiv of BB **1a** or **1c**, TMSOTf, DCM, -35 °C (5 min) → -15 °C (30 min); b) 2 x 1.8 equiv of BB **26**, NIS, TfOH, DCM/dioxane, -20 °C (5 min) → -5 °C (40 min); c) 3 cycles of 20% NEt₃ in DMF, 25 °C (5 min); d) 6 cylces of PBu₃, THF/H₂O (5:1), 45 °C, 30 min; e) Bz₂O, DMAP, DCE; f) hν (305 nm) **28**: 25%; g) TsOH, DCM/MeOH (9:1), 40 °C, 16 h; h) DMP, DCM, H₂O, 2 d; i) NaOCl₂, NaH₂PO₄, tBuOH/THF/H₂O (2:1:1), 2 h **29**: 55% over 3 steps; h) NaOMe, DCM/MeOH, 12h; i) H₂, Pd/C, EtOAc/MeOH/H₂O/HOAc, 12 h, **30**: 21% over 2 steps.

assembly oxidation strategy because the attempt to use a glucuronic acid BB failed, probably due to low reactivity of the disarmed donor. Iterative additions of **1a**, **1c**, and glucose BB **27** (see SI for synthesis) to linker-functionalized resin **8** provided protected tetrasaccharide **28** in 25% yield after photolyticalcleavage from the solid support. Key for a successful synthesis was C6-protection of the glucose BB with an acetyl group as the respective levulinoyl-protected BB gave poor conversion. Both donors yielded the 1,2-*cis*-glucosylated product with high selectivity.[28] The respective 1,2-*trans* compound could not be identified in the mixture of side products detected by HPLC. The acetyl group was selectively removed using *p*-toluenesulfonic acid[29] and the resulting free hydroxyl was oxidized in a two-step protocol using Dess-Martin periodinane and sodium chlorite.

Subsequent removal of the benzoyl esters and benzyl ethers afforded methylated glucuronoxylan oligosaccharide **30**. Due to the many post-assembly transformations required, we did not pursue the synthesis of larger glucuronoxylan oligosaccharides.

The prepared collection of arabinoxylan oligosaccharides represents suitable substrates for the characterization of xylan-degrading enzymes. We investigated commercially available endo-xylanases and arabinofuranosidases, i.e. a Glycosyl Hydrolase (GH) family 10 xylanase from *Cellvibrio japonicas* (XYNACJ),[30] a GH11 xylanase from rumen microorganism (XYRU6),[31] and two GH43 arabinofuranosidases from *Bacteroides ovatus* (ABFBO17 and ABFBO25)[30, 32] We performed mostly simple end-point measurements, by incubating the enzymes with the respective oligosaccharide for 3 h at 40 °C and then terminating the reaction by heat inactivation. The resulting fragments were analyzed by HPLC coupled to a mass spectrometer and an Evaporative Light Scattering Detector (ELSD). Due to the sensitivity of the ELSD we were able to use as little as 10 nmol of material for each individual reaction.

Exemplary results for the GH10 xylanase XYNACJ are shown in Figure 1a. As expected, linear octasaccharide **12** was readily degraded by the xylanase and di- and trixylosides were obtained. The arabinose substituents in **15** and **24** directed the cleavage to the glycosidic bond, which immediately followed the substituted xylose unit, starting from the reducing end. Depending on whether the arabinose was α-1,2- or α-1,3-linked, the backbone was further hydrolyzed either one or two units before the substituent. The two consecutive arabinose residues in **18** prevented cleavage next to the substituents and instead the linker was cleaved off. However, when the arabinose residues were separated by at least one unsubstituted xylose unit (**19** and **20**) the backbone was readily hydrolyzed between the substituents. Based on the identified fragments, we derived the cutting sites for each of the oligosaccharides and for both types of endo-xylanases investigated (Figure 1b). Summarizing these results, we could deduce general requirements for arabinose substitutions that are tolerated by the GH10 xylanase (Figure 1c). For example we found that the enzyme accommodates arabinose substitution in the +1 position but not in the +2 position relative to the cutting site. Furthermore, our results indicate that decoration of the xylan backbone with an arabinose linked to the C3 of the xylose unit is tolerated in closer proximity to the cutting site (-2 position) than when the arabinose is linked to the C2-position (-3 position). These results are consistent with previously published data on the active site of GH10 xylanases which is conserved between most, but not all, family members.[12] Due to their larger xylan recognition site and the limited size of the preparedoligosaccharides a similar generalization was not possible for the GH11 xylanases.

Next, we aimed at identifying the minimal length of the linear xylan oligosaccharides required for hydrolysis by the xylanases. Our end-point measurements with the series of linear oligoxylans (**9-12**) suggested that a trisaccharide might
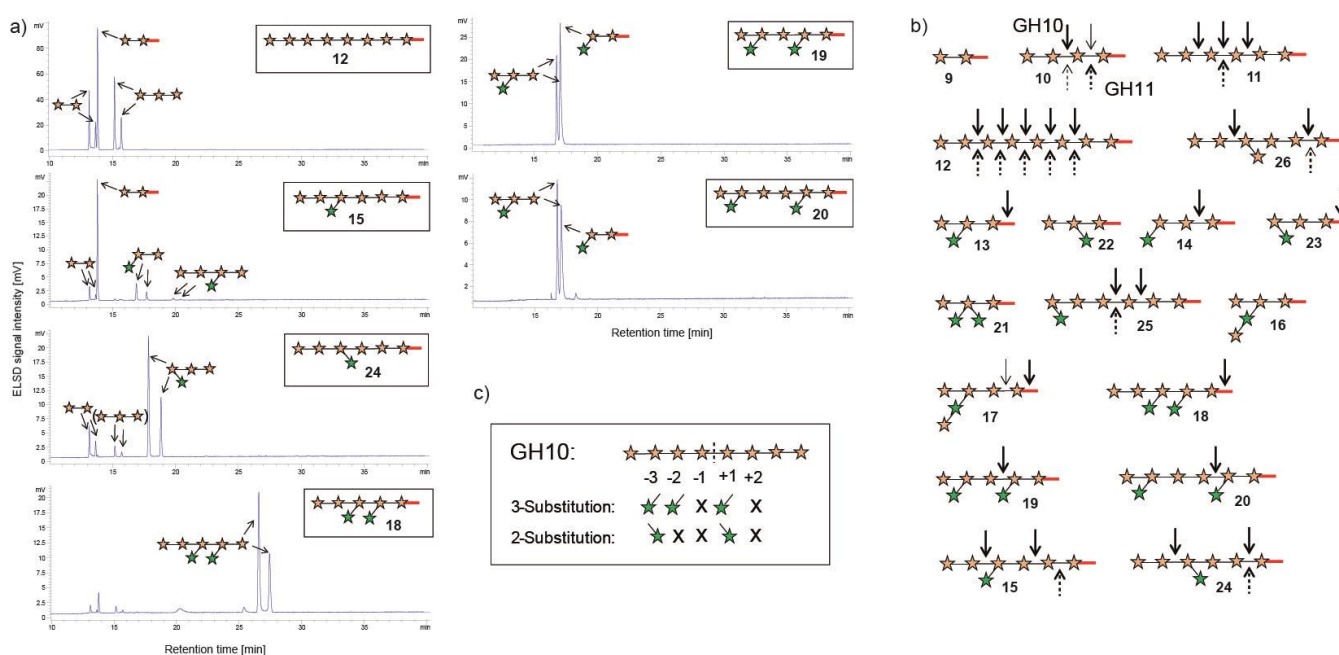
Figure 1. Investigation of GH10 and GH11 endo-xylanases using synthetic xylan oligosaccharides. (a) HPLC analyses of reactions of a GH10 xylanase with different substrates (indicated by boxes). Peaks are annotated with xylan fragments with aminopentyl linker or with free reducing end (with or without red bar,), Note that the α- and the β-form of the fragments with free reducing end elute separately. (b) Cutting sites of GH10 and GH11 xylanases in all prepared arabinoxylan oligosaccharides. Thin arrows indicate minor activity at the respective site. (c) General requirements for arabinose substitutions relative to the cutting site of the GH10 xylanase derived from the results obtained in (b). The "X" denotes xylose residue positions that must not be substituted. The reducing end of the structures is located on the right.

be the smallest linear structure hydrolyzed (Figure 1b and Supplementary Figure 1). Indeed, using de-arabinosylated compound **22**, we found partial hydrolysis of the xylotrioside by the GH10 but not the GH11 enzyme (Supplementary Figure 2).

Intriguingly, in a few cases we also identified small amounts of oligosaccharide fragments that cannot result from the simple hydrolytic activity of GH10 xylanases, such as the linear xylotrioside found after reaction with compounds **24** and **26** (Figure 1a and Supplementary Figure 1). We hypothesize that endo-transglycosylation reactions generated longer oligosaccharides, which were then digested to this xylotrioside. Endo-transglycosylation has previously been described as a side reaction of β-retaining GH10 enzymes,[16a] in which, instead of water, an oligoxyloside serves as the acceptor during the enzymatic reaction.

Another important class of enzymes in the deconstruction of arabinoxylans are arabinofuranosidases that can act on mono-[33] or double-substituted[34] xylose residues. To investigate the substrate specificities of these enzymes, we performed experiments with two different GH 43 arabinofuranosidases (see Supplementary Figure 3 in the SI). Our results indicated a strong preference of ABFBO25 for cleavage of α-1,2- compared to α-1,3-linked arabinoses (Figure 2a,b). To analyze this observation more in detail, we carried out time-course experiments using compounds **13**, **14**, **22**, and **23** (Figure 2c). Our results corroborated the strong preference of ABFBO25 for α-1,2- compared to α-1,3-linked arabinoses, and revealed a slight

preference of the enzyme for non-terminal arabinoses. Compound **21** with two arabinose substitutions at the same xylose residue was exclusively not digested by ABFBO25, but representedthe only substrate for ABFBO17 (Figure 2d). In agreement with previous studies,[8b, 30, 32] ABFBO17 selectively removed the α-1,3-linked arabinose, leaving the α-1,2-substituted arabinose compound as the product (Figure 2d). We determined the identity of the product simply by correlating its retention time with compound **22** that we used as a standard. Doubly arabinosylated xylose residues are frequently found in arabinoxylans of grasses and correspond to a large portion of the arabinose content of grass cell walls.[3b, 8b] To tap this monosaccharide pool, microorganisms, such as *Bacteroides ovatus,* have developed a two-step mechanism, involving digestion by two different types of enzymes.[8b, 32] The exclusive selectivity of ABFBO17 and ABFBO25, which belong to the same GH family, for either mono- or disubstituted arabinoxylosides indicate distinct properties in the catalytic domain of these enzymes.

Conclusions

A set of 19 xylan oligosaccharides with arabinofuranose and glucuronic acid substituents attached to the C2- and/or C3-position of the backbone xylose units was prepared by automated glycan assembly. Three fully orthogonal temporary
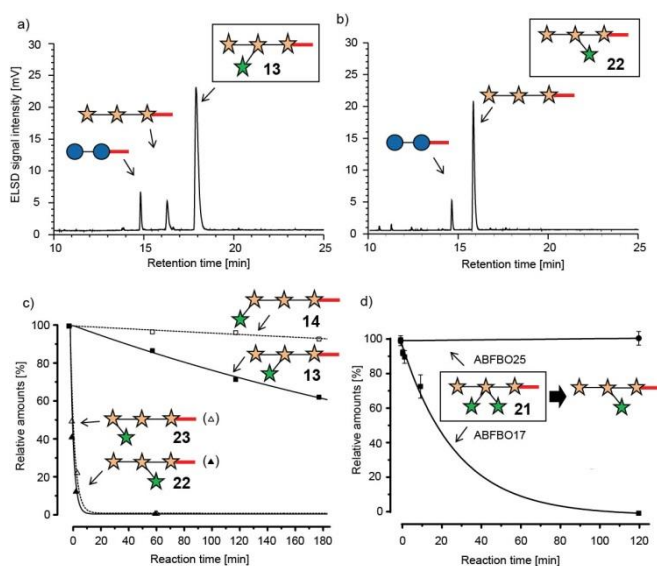
Figure 2. Arabinofuranosidase activities on different arabinoxylan oligosaccharide substrates. (a,b) HPLC analysis of a reaction of a GH43 arabinofuranosidase from Bacteroides ovatus (ABFBO25) with xylotriose substituted with arabinose at the C3- (a) or the C2- (b) position after 60 min incubation time. The structure of the substrates is indicated by boxes. The glucose disaccharide was used as internal standard for quantification. (c) Time-course experiment using four different tetrasaccharides as substrates for ABFBO25. Shown is one representative experiment for each compound. (d) Time-course experiment for ABFBO17, another GH43 arabinofuranosidase from Bacteroides ovatus, and ABFBO25 using a C2-, C3-disubstituted arabinoxylan oligosaccharide as substrate (indicated by a box). Note that only ABFBO17 specifically removes C3-substituted arabinose from C2-, C3-disubstituted xyloses. These results were obtained at an enzyme concentration of 0.1 U/ml.

protecting groups (Fmoc, Nap and Azmb) allowed for the modular assembly of tailored molecular probes for biological applications. These include the characterization of xylan-directed antibodies[19a] as well as glycosyltransferases involved in xylan biosynthesis.[25]

We investigated the substrate specificities of xylan-degrading enzymes by incubating the arabinoxylan oligosaccharides with glycosyl hydrolases followed by HPLC-MS analysis of the digestion products. Simple end-point measurements revealed the substrate specificities of one GH10 xylanase and two arabinofuranosidases. More detailed information on the preferences of the enzymes for specific oligosaccharides was available through time-course experiments. The HPLC-based assay using synthetic arabinoxylan oligosaccharides rather than natural polysaccharide substrates shortens analysis time to a single day and provides a powerful tool for screening xylan-degrading enzymes.

Future syntheses of even larger xylan oligosaccharides will enable the active site-mapping of xylanases belonging to the other GH families 5, 8, and 11. These xylanase families are less studied than the GH10 xylanases and insights into substrate binding properties are urgently needed.

the need for more sophisticated analytical techniques and thus provides a powerful tool for screening xylan-degrading enzymes.

Future efforts are directed towards the synthesis of even larger xylan oligosaccharides that will enable the active site-mapping of xylanases belonging to other GH families such as 5, 8, and 11. These xylanase families are less studied than the GH10 xylanases and new information about their substrate binding properties is urgently needed.

Experimental Section

Synthesizer Modules and Conditions. Linker-functionalized resin 8 (12.5 - 33.8 µmol of hydroxyl groups) was placed in the reaction vessel of the automated oligosaccharide synthesizer and swollen for at least 30 min in DCM. Before every synthesis the resin was washed with DMF, THF and DCM. Subsequently the glycosylation (Module A and D) and deprotection (Module B and C) steps were performed. Mixing of the components was accomplished by bubbling Argon through the reaction mixture.

Module A: Glycosylation with Glycosyl Phosphates. The resin (12.5 - 33.8 µmol of hydroxyl groups) was swollen in DCM (2 mL) and the temperature of the reaction vessel was adjusted to -30 °C. Prior to the glycosylation reaction the resin was washed with TMSOTf in DCM and then DCM only. For the glycosylation reaction the DCM was drained and a solution of phosphate BB (1.8 - 5 equiv in 1 mL DCM) was delivered to the reaction vessel at -35 °C. The reaction was started by the addition of TMSOTf in DCM (1.8 – 5 equiv in 1 mL DCM). The glycosylation was performed for 5 min at -35 °C and then for 30 minutes at -15 °C. Subsequently the solution was drained and the resin washed three times with DCM. The whole procedure was repeated once to increase conversion of the acceptor sites. Afterwards the resin was washed three times with DCM at 25 °C.

Module B: Glycosylation with Thioglycosides. The resin (12.5 - 33.8 µmol of hydroxyl groups) was swollen in DCM (2 mL) and the temperature of the reaction vessel was adjusted to -30 °C. Prior to the glycosylation reaction the resin was washed with TMSOTf in DCM and then DCM only. For the glycosylation reaction the DCM was drained and a solution of thioglycoside BB (1.8 - 5 equiv in 1 mL DCM) was delivered to the reaction vessel. After the set temperature was reached, the reaction was started by the addition of NIS (4.44 equiv or 8.88 equiv) and TfOH (0.44 equiv or 0.88 equiv) in DCM/dioxane (2:1). The glycosylation was performed for 5 min at -40 °C and then for 40 min at -20 °C. Subsequently the solution was drained and the resin was washed with DCM. The whole procedure was repeated once to ensure full conversion of all acceptor sites. Afterwards the resin was washed three times with DCM at 25 °C

Module C: Fmoc Deprotection. The resin was washed with DMF, swollen in 2 mL DMF and the temperature of the reaction vessel was adjusted to 25 °C. Prior to the deprotection reaction the DMF was drained and the resin was washed with DMF three times. For Fmoc deprotection 2 mL of a solution of 20% Et3N in DMF was delivered to the reaction vessel. After 5 min the solution was drained and the whole procedure was repeated another two times. After Fmoc deprotection was complete the resin was washed with DMF, THF and DCM.

Module D: Nap Deprotection. The resin was washed with DCM three times and the temperature of the reaction vessel was adjusted to 40 °C. For Nap deprotection the DCM was drained and 1.5 mL of of a 0.1 M DDQ solution in DCE/MeOH/H2O (64:16:1) was delivered to the reaction

vessel. After 20 min the reaction solution was drained and the whole procedure was repeated another six times. After Nap deprotection was complete the resin was washed with DMF, THF and DCM.

Module E: Azmb Deprotection. The resin was washed with THF, swollen in 2 mL THF and the temperature of the reaction vessel was adjusted to 25 °C. Prior to the deprotection step the THF was drained and the resin was washed with THF three times. For Azmb deprotection, 2 mL of a solution of PBu3 (7.5 or 15 equiv.) in THF containing 5% H_2O was delivered to the reaction vessel at 45 °C. After 30 min the solution was drained and the whole procedure was repeated another five times. After Azmb deprotection was complete the resin was washed with THF and DCM.

Module F: Acetylation. The temperature of the reaction vessel was adjusted to 25 °C and the resin washed with pyridine (2 mL) three times. For capping, 1 mL pyridine and 1 mL acetic anhydride were delivered to the reaction vessel. After 30 min the reaction solution was drained and the whole procedure was repeated another two times. After capping was complete the resin was washed with pyridine and DCM.

Module G: Benzoylation. The temperature was adjusted to 25 °C and the resin washed with pyridine (2 mL) three times. For benzoylation temperature was set to 40 °C and 2 mL pyridine and 1 mL of a solution containing 0.5 M benzoic anhydride and 0.25 M DMAP in DCE were delivered to the reaction vessel. After 30 min the reaction solution was drained and the whole procedure was repeated another two times. After capping was complete the resin was washed with DCM.

Cleavage from the solid support.[27] After assembly of the oligosaccharides cleavage from the solid support was accomplished by modification of a previously published protocol[27] using the Vapourtec E-Series UV-150 photoreactor Flow Chemistry System. The medium pressure metal halide lamp is filtered using the commercially available red filter. The resin, suspended in DCM, was loaded into a plastic syringe. The suspension was then pumped using a syringe pump (PHD2000, Harvard Aparatus) at 1 mL/min through a 10 mL reactor, constructed of 1/8 inch o.d. FEP tubing. The total volume within the photoreactor was 9 mL. The temperature of the photoreactor was maintained at 20 °C and the lamp power was 80%. The exiting flow was deposited in a 10 mL syringe containing a filter, with a collection flask beneath the syringe.

Global Deprotection. The protected oligosaccharide was dissolved in DCM (1 mL) and MeOH (2 mL) and NaOMe (0.5 M in MeOH, 0.5 mL) were added. The reaction mixture was stirred overnight and subsequently neutralized by addition of prewashed Amberlite IR-120 resin. The resin was filtered off and the solvents were removed in vacuo. The crude product was purified by preparative HPLC, dissolved in a mixture of EtOAc/MeOH/AcOH/H_2O (4:2:2:1, 3 mL) and the resulting solution was added to a round-bottom flask containing Pd/C (10% Pd, 10-20 mg). The suspension was saturated with H2 for 30 min and stirred under an H_2-atmosphere overnight. After filtration of the reaction mixture through a syringe filter the solvents were evaporated to provide the fully deprotected oligosaccharide.

Analysis of Glycosyl Hydrolase Substrate specificities. The endo-β-1,4-xylananse and α-L-arabinofuranosidase enzymes were purchased from Megazyme (Bray, Ireland) and used in the following buffers that were suggested by the manufacturer: Cellvibrio japonicas GH10 xylanase (XYNACJ), in 100 mM sodium acetate buffer, pH 5; rumen microorganism GH11 xylanase M6 (XYRU6) in 100 mM sodium phosphate buffer, pH 6; *Bacteroides ovatus* GH43 arabinofuranosidases B17 (ABFBO17) and B25 (ABFBO25) in sodium phosphate buffer, pH

6.5. The enzymes were used at a concentration of 1 U/mL for the end-point measurements and at 0.1 U/mL for the time-course experiments. The oligosaccharides were used at a concentration of 1 mM. All reactions were carried out at 40 °C and terminated by incubation at 80 °C for 5 min. The reactions were analyzed on an Agilent 1200 Series HPLC equipped with an Agilent 6130 quadruple mass spectrometer (MS) and an Agilent 1200 Evaporative Light Scattering Detector (ELSD). The oligosaccharides were separated on a Hypercarb column (150 x 4.6 mm, Thermo Scientific) using a water (including 0.1% formic acid)-acetonitrile (ACN) gradient at a flow-rate of 0.7 mL/min starting at 2.5% ACN for 5 min, ramping up to 15% ACN at 8 min, followed by a slow increase of ACN to 30% at 40 min, a steep ramp to 100% ACN at 43.5 min, a decline back to 2.5% ACN from 46 min to 47 min, and equilibration until 55 min at 2.5% ACN. The peaks in the ELSD traces were assigned based on their retention time and the corresponding masses in the MS.

[1]   a) E. A. Rennie, H. V. Scheller, *Current Opinion in Biotechnology* **2014**, *26*, 100-107; b) H. V. Scheller, P. Ulvskov, *Annu. Rev. Plant Biol.* **2010**, *61*, 263-289.
[2]   A. Ebringerová, T. Heinze, *Macromol. Rapid Commun.* **2000**, *21*, 542-556.
[3]   a) B. C. Saha, *Journal of Industrial Microbiology and Biotechnology* **2003**, *30*, 279-291; b) M. Pauly, K. Keegstra, *Plant J.* **2008**, *54*, 559-568
[4]   J. Fargione, J. Hill, D. Tilman, S. Polasky, P. Hawthorne, *Science* **2008**, *319*, 1235-1238.
[5]   J. Vogel, *Curr. Opin. Plant Biol.* **2008**, *11*, 301-307.
[6]   a) A. M. Neyrinck, V. F. Van Hee, N. Piront, F. De Backer, O. Toussaint, P. D. Cani, N. M. Delzenne, *Nutrition and Diabetes* **2012**, *2*, e28; b) B. A. Williams, D. Zhang, A. T. Lisle, D. Mikkelsen, C. S. McSweeney, S. Kang, W. L. Bryden, M. J. Gidley, *Nutrition* **2016**, *32*, 491-497.
[7]   a) J. Liu, S. Willför, C. Xu, *Bioactive Carbohydrates and Dietary Fibre* **2015**, *5*, 31-61; b) M. Mendis, E. Leclerc, S. Simsek, *Carbohydr. Polym.* **2016**, *139*, 159-166.
[8]   a) D. Dodd, I. K. O. Cann, *Glob Change Biol Bioenergy* **2009**, *1*, 2-17; b) S. Lagaert, A. Pollet, C. M. Courtin, G. Volckaert, *Biotechnology Advances* **2014**, *32*, 316-332.
[9]   T. Collins, C. Gerday, G. Feller, *FEMS Microbiology Reviews* **2005**, *29*, 3-23.
[10]   a) Q. Beg, M. Kapoor, L. Mahajan, G. Hoondal, *Appl. Mircorbiol. Biotechnol.* **2001**, *56*, 326-338; b) L. Viikari, A. Kantelinen, J. Sundquist, M. Linko, *FEMS Microbiology Reviews* **1994**, *13*, 335-350.
[11]   V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **2014**, *42*, D490-D495.
[12]   A. Pollet, J. A. Delcour, C. M. Courtin, *Crit. Rev. Biotechnol.* **2010**, *30*, 176-191.
[13]   a) S. L. Maslen, F. Goubet, A. Adam, P. Dupree, E. Stephens, *Carbohydr. Res.* **2007**, *342*, 724-735; b) M. Vardakou, P. Katapodis, M. Samiotaki, D. Kekos, G. Panayotou, P. Christakopoulos, *Int. J. Biol. Macromol.* **2003**, *33*, 129-134; c) P. Biely, M. Vršanská, M. Tenkanen, D. Kluepfel, *J. Biotechnol.* **1997**, *57*, 151-166.
[14]   a) J. Beaugrand, G. Chambat, V. W. K. Wong, F. Goubet, C. Rémond, G. Paës, S. Benamrouche, P. Debeire, M. O'Donohue, B. Chabbert, *Carbohydr. Res.* **2004**, *339*, 2529-2540; b) F. J. M. Kormelink, H. Gruppen, R. J. Viëtor, A. G. J. Voragen, *Carbohydr. Res.* **1993**, *249*, 355-367.

[15]   F. J. M. Kormelink, R. A. Hoffmann, H. Gruppen, A. G. J. Voragen, J. P. Kamerling, J. F. G. Vliegenthart, *Carbohydr. Res.* **1993**, *249*, 369-382.

[16]   a) S. J. Charnock, T. D. Spurway, H. Xie, M.-H. Beylot, R. Virden, R. A. J. Warren, G. P. Hazlewood, H. J. Gilbert, *J. Biol. Chem.* **1998**, *273*, 32187-32199; b) Z. Fujimoto, S. Kaneko, A. Kuno, H. Kobayashi, I. Kusakabe, H. Mizuno, *J. Biol. Chem.* **2004**, *279*, 9606-9614; c) G. Pell, L. Szabo, S. J. Charnock, H. Xie, T. M. Gloster, G. J. Davies, H. J. Gilbert, *J. Biol. Chem.* **2004**, *279*, 11777-11788; d) G. Pell, E. J. Taylor, T. M. Gloster, J. P. Turkenburg, C. M. G. A. Fontes, L. M. A. Ferreira, T. Nagy, S. J. Clark, G. J. Davies, H. J. Gilbert, *J. Biol. Chem.* **2004**, *279*, 9597-9605; e) B. Svensson, M. Søgaard, *J. Biotechnol.* **1993**, *29*, 1-37.

[17]   a) M. P. Bartetzko, F. Schuhmacher, P. H. Seeberger, F. Pfrengle, *J. Org. Chem.,* in press, DOI: 10.1021/acs.joc.6b02745; b) P. Dallabernardina, F. Schuhmacher, P. H. Seeberger, F. Pfrengle, *Chem. Eur. J.*, in press, DOI: 10.1002/chem.201605479.

[18]   a) P. H. Seeberger, *Acc. Chem. Res.* **2015**, *48*, 1450-1463; b) O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523-1527; c) L. Kröck, D. Esposito, B. Castagner, C.-C. Wang, P. Bindschadler, P. H. Seeberger, *Chem. Science* **2012**, *3*, 1617-1622.

[19]   a) D. Schmidt, F. Schuhmacher, A. Geissner, P. H. Seeberger, F. Pfrengle, *Chem. Eur. J.* **2015**, *21*, 5709-5713; b) M. P. Bartetzko, F. Schuhmacher, H. S. Hahm, P. H. Seeberger, F. Pfrengle, *Org. Lett.* **2015**, *17*, 4344-4347; c) P. Dallabernardina, F. Schuhmacher, P. H. Seeberger, F. Pfrengle, *Org. Biomol. Chem.* **2016**, *14*, 309-313.

[20]   a) J. A. Wright, J. Yu, J. B. Spencer, *Tetrahedron Lett.* **2001**, *42*, 4033-4036; b) J. Xia, S. A. Abbas, R. D. Locke, C. F. Piskorz, J. L. Alderfer, K. L. Matta, *Tetrahedron Lett.* **2000**, *41*, 169-173; c) J. Xia, J. L. Alderfer, C. F. Piskorz, K. L. Matta, *Chem. Eur. J.* **2001**, *7*, 356-367.

[21]   a) A. Arasappan, P. L. Fuchs, *J. Am. Chem. Soc.* **1995**, *117*, 177-183; b) K. R. Love, R. B. Andrade, P. H. Seeberger, *J. Org. Chem.* **2001**, *66*, 8165-8176; c) T. Wada, A. Ohkubo, A. Mochizuki, M. Sekine, *Tetrahedron Lett.* **2001**, *42*, 1069-1072.

[22]   L. A. Paquette, L. Barriault, D. Pissarnitski, J. N. Johnston, *J. Am. Chem. Soc.* **2000**, *122*, 619-631.

[23]   T.-Y. Huang, M. M. L. Zulueta, S.-C. Hung, *Org. Lett.* **2011**, *13*, 1506-1509.

[24]   a) P. I. Abronina, N. M. Podvalnyy, S. L. Sedinkin, K. G. Fedina, A. I. Zinin, A. O. Chizhov, V. I. Torgov, L. O. Kononov, *Synthesis* **2012**, *44*, 1219-1225; b) J. Kandasamy, M. Hurevich, P. H. Seeberger, *Chem. Comm.* **2013**, *49*, 4453-4455.

[25]   a) N. Anders, M. D. Wilkinson, A. Lovegrove, J. Freeman, T. Tryfona, T. K. Pellny, T. Weimar, J. C. Mortimer, K. Stott, J. M. Baker, M. Defoin-Platel, P. R. Shewry, P. Dupree, R. A. C. Mitchell, *Proc. Natl. Acad. Sci.* **2012**, *109*, 989-993; b) D. Chiniquy, V. Sharma, A. Schultink, E. E. Baidoo, C. Rautengarten, K. Cheng, A. Carroll, P. Ulvskov, J. Harholt, J. D. Keasling, M. Pauly, H. V. Scheller, P. C. Ronald, *Proc. Natl. Acad. Sci.* **2012**, *109*, 17117-17122.

[26]   C. Voiniciuc, M. Günl, M. H.-W. Schmidt, B. Usadel, *Plant. Physiol.* **2015**, *169*, 2481-2495.

[27]   S. Eller, M. Collot, J. Yin, H. S. Hahm, P. H. Seeberger, *Angew. Chem. Int. Ed.* **2013**, *52*, 5858-5861.

[28]   H. S. Hahm, M. Hurevich, P. H. Seeberger, *Nature Communications* **2016**, *7*, 12482.

[29]   A. G. González, I. Brouard, F. León, J. I. Padrón, J. Bermejo, *Tetrahedron Lett.* **2001**, *42*, 3187-3188.

[30]   B. V. McCleary, V. A. McKie, A. Draga, E. Rooney, D. Mangan, J. Larkin, *Carbohydr. Res.* **2015**, *407*, 79-96.

[31]   O. Al-Widyan, M. H. Khataibeh, K. Abu-Alruz, *J. Appl. Sci.* **2008**, *8*, 672-676.

[32]   A. Rogowski, J. A. Briggs, J. C. Mortimer, T. Tryfona, N. Terrapon, E. C. Lowe, A. Basle, C. Morland, A. M. Day, H. Zheng, T. E. Rogers, P. Thompson, A. R. Hawkins, M. P. Yadav, B. Henrissat, E. C. Martens, P. Dupree, H. J. Gilbert, D. N. Bolam, *Nat. Commun.* **2015**, *6*.

[33]   a) F. J. M. Kormelink, M. J. F. Searle-Van Leeuwen, T. M. Wood, A. G. J. Voragen, *Appl. Mircobiol. Biotechnol.* **1991**, *35*, 231-232; b) T. M. Bourgois, V. Van Craeyveld, S. Van Campenhout, C. M. Courtin, J. A. Delcour, J. Robben, G. Volckaert, *Appl. Mircobiol. Biotechnol.* **2007**, *75*, 1309-1317.

[34]   a) V. K. M. J. Laere, G. Beldman, J. A. G. Voragen, *Appl. Mircobiol. Biotechnol.* **1997**, *47*, 231-235; b) L. A. M. van den Broek, R. M. Lloyd, G. Beldman, J. C. Verdoes, B. V. McCleary, A. G. J. Voragen, *Appl. Mircobiol. Biotechnol.* **2005**, *67*, 641-647; c) H. R. Sørensen, C. T. Jørgensen, C. H. Hansen, C. I. Jørgensen, S. Pedersen, A. S. Meyer, *Appl. Mircobiol. Biotechnol.* **2006**, *73*, 850-861.

FULL PAPER

FULL PAPER
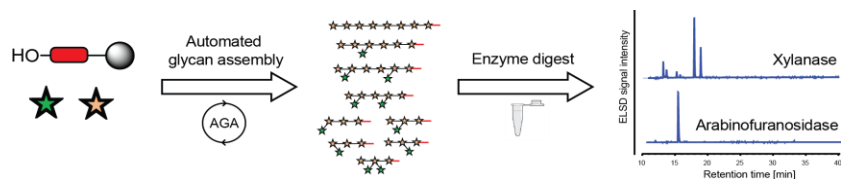
Automated glycan assembly enables characterization of glycosyl hydrolases:
We report the automated glycan assembly of a collection of arabinoxylan oligosaccharides with diverse substitution patterns. The synthetic arabinoxylan oligosaccharides were used for analyzing the substrate specificities of xylan-degrading enzymes which are critical reagents for biotechnological applications such as the production of second generation biofuels.

D. Senf, C. Ruprecht, G. H. M. de Kruijff, S. O. Simonetti, F. Schuhmacher, P. H. Seeberger, F. Pfrengle*

Page No. – Page No.

Active Site-Mapping of Xylan-Deconstructing Enzymes with Arabinoxylan Oligosaccharides Produced by Automated Glycan Assembly