

Green Chemistry

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



www.rsc.org/greenchem



Green Chemistry

ARTICLE

New degradation compounds from lignocellulosic biomass pretreatment: Routes for formation of potent oligophenolic enzyme inhibitors

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

H. Rasmussen^a, D. Tanner^b, H. R. Sørensen^a and A. S. Meyer^{c*}

In this study 28 new oligophenolic cellulase inhibitors were discovered in wheat straw pretreatment liquors. By consideration of the reaction mechanisms for their formation it is proposed that these oligophenols are formed during hydrothermal biomass pretreatment by pentose self-condensation reactions involving aldol condensations, 1,4 additions to α,β unsaturated carbonyl compounds, 3-keto acid decarboxylations and oxidations. Furthermore, pentose reactions with phenolic lignin components are suggested. The identification of the central role of xylose in the reaction routes for oligophenolic inhibitor formation led to the solution to protect the reactive anomeric center in xylose. It is shown that protection of the anomeric center in *in situ* generated xylose with ethylene glycol monobutyl ether, during pretreatment of wheat straw, reduces the level of oligophenols by 73 %. The results pave the way for implementation of new types of reactions that hinder inhibitor formation in lignocellulosic biomass processing.

Introduction

Lignocellulosic conversion processes have the potential to provide the future society with environmentally friendly energy and platform chemicals¹⁻⁵. A challenge is to make the biorefinery processes economically competitive and a cost driver in biorefinery processes is the enzymes that are utilised in the first lignocellulosic saccharification. Unfortunately, in lignocellulosic biomass conversion, the enzyme catalysed reactions are inhibited by compounds that are generated or liberated during hydrothermal pretreatment of the biomass⁶. Identification of these inhibitory compounds, and notably an understanding of how they are formed, are crucial in order to develop methods to avoid their formation or to remove them. Compound types that are reported to be especially inhibitory towards cellulases, include xylooligosaccharides⁷⁻¹⁰ and several phenolic compounds¹¹⁻¹³. Xylooligosaccharides are biomass structural elements that are liberated from xylan during pretreatment. Although phenolics compounds found in biomass liquors can stem from lignin¹⁴, early reports show that various phenolic compounds may also form as degradation products from glucose, xylose and arabinose¹⁵⁻¹⁷. Recently this formation of phenolics has been confirmed by demonstration of pseudo-lignin generation from model

substrates of cellulose (i.e. Avicel), xylan and xylose¹⁸.

It can be hypothesised that other still unidentified compounds can be potent inhibitors as well, since compounds with multiple chemical functionalities are formed during biomass pretreatment⁶, and give rich possibilities for various chemical reactions to take place and hence form many new potential inhibitor compounds.

The objective of this work was to reveal the chemical identity of the most potent cellulase inhibitors and consider their synthesis routes in order to understand the origin of the compounds and hinder their formation.

Experimental

Chemicals and enzymes

Buffer chemicals, solvents, acids and bases were purchased from Sigma Aldrich. Cellic CTec3 (1920 Biomass Hydrolysis Units/g) obtained from Novozymes A/S (Bagsværd, Denmark) was utilised. It is a commercially available cellulase preparation based on the *Trichoderma reesei* complex. Apart from the cellulolytic enzyme base from *T. reesei* it contains at least the two main cellobiohydrolases EC 3.2.1.91 (Cel6A and Cel7A), five different endo-1,4- β -glucanases EC 3.2.1.4 (Cel7B, Cel5A, Cel12A, Cel61A, and Cel45A), β -glucosidase EC 3.2.1.21, and a β -xylosidase¹⁹, the preparation Cellic CTec3 also contains other proprietary hydrolysis-boosting proteins.

Fractionation of pilot plant liquid from hydrothermal pretreatment (LfHP) of wheat straw

Liquid from pilot scale hydrothermal pretreated (LfHP) wheat straw (183 °C, 12 minutes, 40 % (w/w) dry matter) was

^a DONG Energy, Kraftværksvej 53, DK-7000 Fredericia, Denmark.

^b Dept. of Chemistry, Technical University of Denmark, DK-2800 Lyngby, Denmark.

^c Center for BioProcess Engineering, Dept. of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark.

Electronic Supplementary Information (ESI) available: Tables with LC-MS and LC-MS/MS data and fragments and inhibition mass balance. See DOI: 10.1039/x0xx00000x

extracted with 2-butanone or 2-butanol. 170 g LfHP was extracted with 2 x 160 g solvent. Both extractions were left overnight for complete separation. The emulsion in the middle phase was centrifuged (10 min, 10.000 rcf) for complete separation of water, solvent and insolubles. The water phase and insolubles were freeze dried separately and the organic phase was evaporated at reduced pressure at 50 °C and freeze dried.

Determination of cellulase inhibition from the fractions: Enzyme assay

The freeze dried fractions were dissolved in 0.1 M acetate buffer (pH 5.1) to dry matter 13.6 % (w/w) and avicel was added to 12 % (w/w). When the pretreated wheat straw fiber was used as substrate, it was added to a dry matter content of 12 % (w/w) corresponding to 6 % (w/w) cellulose. (The pretreated wheat straw fiber was the fiber fraction corresponding to the utilised LfHP). Incubation was at 50 °C, 160 rpm and Cellic CTec3 (1920 Biomass Hydrolysis Units/g) was added to 0.6 % (w/w). All assays were made as duplicates with displaced sampling. Upon sampling (200 µL) the enzyme was inactivated at 99 °C for 5 minutes.

Mini pretreatments: Protection of the anomeric center in xylose

Experiments were carried out in a Parr reactor (Parr Instrument Company, Moline, Illinois, USA) with internal cooling loop, on 200 mL scale. 18 g of wheat straw milled to 1 mm size were suspended in 200 mL liquid to give a final dry matter of 8 % (w/w). Agitation speed was 100 rpm.

In the statistical experimental central composite face design the following parameters were tested: Temperatures: 100, 130 and 160 °C. Reaction times: 10, 20, 30 minutes. Solution of glycol ethers with: H₂SO₄ concentrations (% w/w of solution): 0.5, 2.5 and 4.5 %. Water equivalents (to xylose content in the wheat straw): 3.5 (corresponds to the authentic water content in wheat straw i.e. no addition of water), 10.7 (4 mL) and 17.9 (8 mL) equivalents.

In addition to the statistical design, the following parameters were tested in a round bottomed flask equipped with a condenser: Temperatures: 80 and 100 °C. Reaction times: 15, 30 and 60 minutes. H₂SO₄ concentrations (% w/w of solution): 4.5 and 6.5 %. Water equivalents (to xylose content in the wheat straw): 17.9 and 25.1 equivalents.

Mini pretreatments: Spiking experiments with xylose and wheat arabinoxylan

Experiments were carried out in a Parr reactor (instrumentation as described above) at 183 °C, 18 minutes with 18 g of wheat straw milled to 1 mm size suspended in 200 mL deionised water. In one experiment xylose (4.64 g, Merck, New Jersey, USA) was added. In another experiment wheat arabinoxylan (4.18 g, medium viscosity ~ 22 cST, Megazyme, Wicklow, Ireland) was added. The added amounts of xylose and arabinoxylan corresponded to the respective amounts of xylose from xylan in 18 g wheat straw and arabinoxylan in 18 g wheat straw. Agitation speed was 100 rpm. The experiments

were compared to a control experiment with no spiking with respect to the amount of oligophenolic compounds measured as total UV count from 3.0-6.0 minutes in the HPLC analysis (evaluation of inhibitor level) described below.

Analytical methods

HPAEC: Glucose analysis. Analysis was performed on a ICS-5000 Dionex ThermoFisher Scientific (ThermoScientific, Sunnyvale, California, USA) analytical system with a PA1 column, 2 x 250 mm (Thermo Scientific). The method was a 2 step isocratic method with an additional final cleaning step with KOH concentrations held constant at 25, 10 and 100 mM for 5, 14 and 7 minutes respectively. The flow rate was 0.25 mL/min.

HPLC: Xylose protection. Analysis was performed on a Dionex (ThermoScientific, Sunnyvale, California, USA) Ultimate 3000 UHPLC with RI detection and a Rezex RHM monosaccharide H+ (8%) column, 7.80 x 300 mm (Phenomenex) with the method according to ²⁰.

HPLC: Evaluation of inhibitor level. Analysis was carried out on a Dionex (ThermoScientific, Sunnyvale, California, USA) Ultimate 3000 UHPLC equipped with a DAD detector and a pentafluorophenyl column 2.1 x 100 mm, particle size 2.6 µm (Phenomenex). The flow was 6.5 mL/min with a gradient from 100 % water (containing 0.025 % TFA) -100 % acetonitrile (containing 0.025 % TFA) over 10 minutes. 100 % acetonitrile (0.025 % TFA) was kept for 2 minutes and changed to 100 % water (0.025 % TFA) over 0.1 minute. The system was equilibrated 3 minutes with 100 % water (0.025 % TFA). The inhibitor level was evaluated as the total UV count at 250 nm at retention time 3.0-4.9 minutes, which corresponds to the retention time for the compounds in the 2-butanone phase in this method.

LC-MS and LC-MS/MS. The freeze dried fractions were dissolved in MilliQ water (water fraction and freeze dried LfHP) and 25 % (vol/vol) acetonitrile (2-butanone and 2-butanol fractions). LC-MS and LC-MS/MS analyses were performed on a Thermo (ThermoFisherScientific, Waltham, MA, USA) Orbitrap Fusion instrument with electrospray ionisation with spray voltage 3500 in positive ionisation and 2500 in negative ionisation. HCD energy for MS/MS was 25 %.

The LC conditions were 1.0 mL/min with a gradient as above (eluent containing 0.1 % formic acid).

The column was a pentafluorophenyl (pfp) discovery HS F5, L x I.D. 15 cm x 4.6 mm, 5 µm particle size (Supelco).

Results and discussion

Determination of inhibition from the fractions: Enzyme assay

Liquid from hydrothermal pilot scale pretreatment (LfHP) of wheat straw was extracted with 2-butanone and 2-butanol. These solvents were chosen because they are very polar, but can be separated from water and have low boiling points and are thus suitable for fractionating of compounds solubilised in

LfHP. Furthermore, they were chosen because a protic and an aprotic solvent could possibly result in further selectively *via* solubility. The fractions were evaporated (only organic phases), freeze dried and subsequently compared at the same dry matter base for their enzyme inhibitor potency.

LfHP inhibited the enzymatic glucose release from avicel with 70 % after 22 hours of reaction compared to the control (Figure 1, top). The organic fractions, notably the 2-butanone fraction, were more inhibitory than the LfHP itself (Figure 1, top). The water fractions showed a decrease in inhibition compared to LfHP, which corresponds to that some of the most inhibitory compounds had been extracted into the organic phases.

The 2-butanone extracts also showed significantly inhibitory action on enzymatic glucose release from authentic pretreated wheat straw (Figure 1, bottom).

No difference in inhibition between freeze dried and non-freeze dried LfHP was observed (data not shown) indicating that potentially inhibitory, volatile compounds, such as furfural and formic acid, had not escaped evaluation. Thus, in agreement with previous data¹³ it was affirmed, that volatile compounds, in the present amounts, where furfural was measured to 0.04 g/l and acetic acid to 5.5 g/L do not contribute to the inhibition from LfHP.

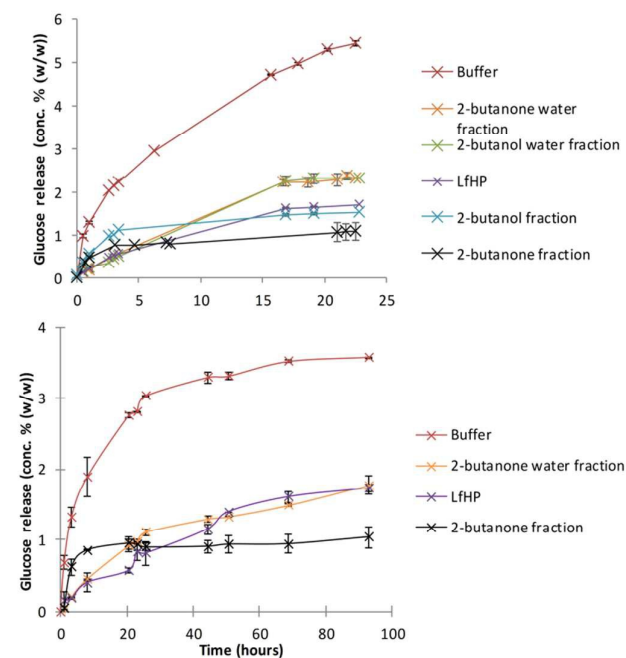


Figure 1. Glucose release during enzymatic hydrolysis of model cellulose substrate (avicel) (top) and authentic pretreated wheat straw (bottom), with compounds added from solvent extractions of liquid from hydrothermal pretreatment (LfHP) of wheat straw. All fractions i.e. compounds from the organic fraction, water fraction or freeze dried LfHP were added to equal concentrations (w/w).

Compound distribution in the fractions

The peaks with retention time from 1 to 8 minutes were distributed differently in the phases after the solvent

extractions (Figure 2, S1 Figure 1). Peaks beyond 8 minutes were also observed in blank control samples (data not shown) and thus interpreted as extractables. The compounds with *rt* from 1 to 3.7 minutes were left in the water phases (Figure 2A and S1 Figure 1A) and the compounds with *rt* from 5.5 to 8 minutes were extracted into the organic phases (Figure 2C and S1 Figure 1D) most effectively into the 2-butanone phase (Figure 2C). The 2-butanone phase also contained compounds with retention time 1-3.7 minutes similar to those in the water phases (S1 Figure 1).

The compound distribution in the fractions corresponded well with the enzyme inhibition assay (Figure 1), that showed less inhibition from the 2-butanone fraction than the 2-butanone fraction. Thus, the difference between strong inhibition and relieved inhibition is likely due to the compounds with *rt* 5.5-8 and 1-3.7 minutes respectively. There was an overlap of compounds with *rt* 3.7-5.5 in the water and organic solvent fractions, indicating that many of the compounds in this region of the chromatogram could be similar in the different fractions.

Identification of compounds with LC-MS/MS

2-butanone fraction. More than 100 different compounds were eluting in the region 5.5-8 minutes in the 2-butanone fraction (Figure 2C). Thus it was found that LC-MS/MS was more suitable for analytical evaluation than NMR, and the approach was to structurally elucidate as many compounds as possible, to gain information of general structures and compound types. 30 compounds had sufficient intensity and quality of fragmentation. The compounds eluting in the region 5.5-8 minutes in the 2-butanone fraction (Figure 2C) had a high degree of double bond equivalents (DBE) compared to their masses (S1 Table 1) i.e. they were quite condensed structures. Common mass losses that gave rise to fragments (S1 Table 1, S1 Table 3) were 18, 44 and 15, which corresponded to H₂O from elimination of hydroxyl, CO₂ from carboxylic acids and CH₃• radical from cleavage of phenol methoxy groups²¹ respectively (S1 Table 3). Lactones with hydroxyl substituents were strongly suggested by i) multiple sequential neutral mass losses involving [M-H⁺-44], [M-H⁺-18] and in some cases also [M-H⁺-28] CO loss, as well as by ii) the structures being so condensed, that they cannot contain carboxylic acids and hydroxyl groups to explain the multiple fragmentation pattern. Compounds with fused ring systems were strongly suggested by minor fragmentation, again consistent with a condensed structure. Additional fragmentation (S1 Table 1, S1 Table 3) arose from straightforward heterobond cleavage and cleavage of bonds between non-fused ring systems.

The compounds in the 2-butanone fraction were thus mainly oligophenolic compounds, including lactone and flavonoid like structures, flavonoids in addition to monophenolic compounds (Figure 3). Out of the 30 proposed structures, 26 are new (search in Chemical Abstract Service *via* SciFinder, similarity search down to 70 % based on molecular formula).

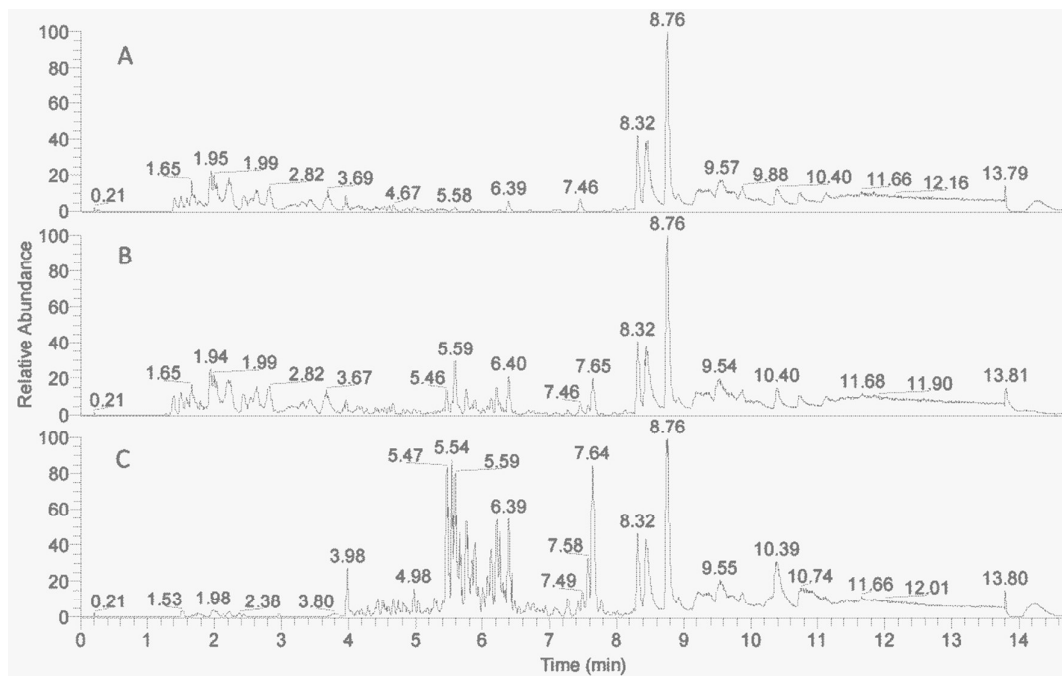


Figure 2. Base peak chromatograms (negative ionisation) of fractions from solvent extractions. A: Water fraction from 2-butanone extraction. B: Non extracted LfHP. C: 2-butanone fraction. The fractions exhibit increasing inhibition from top to bottom.

Three have been reported earlier (cumaroyl substituted pentose (**8**) and feruloyl substituted pentose (**12**) and the flavonoid **15**). A regioisomer (**9**) of a previously described flavonoid structure was also found, but it has not been reported previously from wheat straw. Apart from the structures shown in Figure 3, 10 already known monophenolic compounds and several of their regioisomers were also found (not shown). In relation to cellulase inhibition, the data strongly indicate that the higher inhibition of the 2-butanone fraction is contributed by the new oligophenolic compounds that we propose are a group of "super inhibitory" compounds in LfHP. Many of the proposed structures in Figure 3 are Michael acceptors and they are thus readily susceptible for nucleophilic attack from nucleophiles in amino acids in proteins. This may explain why they are potent enzyme inhibitors. Compound **8**, **12**, **17** and **22** are examples of Michael acceptors with the α,β unsaturated carbonyl moiety determined as described above and in this context the structure of the compound apart from the α,β unsaturated carbonyl moiety can be speculated to be less important. Due to the presence of the many different compounds, several enzyme activities in the catalytic cascade of enzymatic cellulose conversion are expected to be affected. The super inhibitory action of oligophenols agree with earlier studies⁸, reporting that oligophenols, such as tannic acid and oligomeric proanthocyanidin isolated from a grape seed extract, are more potent cellulase inhibitors than monophenolic compounds

Water fraction from 2-butanone extraction. The compounds eluting in the region 1-4 minutes in the water fraction (Figure

2A) were identified as xylooligosaccharides (S1 Table 2). They were identified from their accurate mass and fragmentation pattern, which were compared to xylooligosaccharide standards. Pentoseoligosaccharides up to DP6 with 0-3 acetyl substituents were identified with MS2, but in the MS1 spectra pentoseoligosaccharides up to DP 12 were also seen - still with only 1-3 acetyl substituents (data not shown). The presence of acetylated xylooligosaccharides in LfHP is in accordance with previously reported data⁷. Feruloyl substituted pentoseoligosaccharides were also found in the water phase as well as the 2-butanone phase *vide infra*.

Other compounds distributed across the 2-butanone and water phase fractions. A closer assessment of the compounds with retention time from 3.7-5.5 minutes, where the 2-butanone fraction and its water fraction had an overlap in the base peak chromatograms (Figure 2), indicated that the majority was feruloyl and/or acetyl substituted pentose oligomers (S1 Table 4). The feruloyl substituted compounds were thus not completely separated in the extraction. From the accurate masses there was a clear tendency, that the 2-butanone fraction contained more of the shorter feruloyl substituted pentoseoligosaccharides, while the water fraction contained more of the longer feruloyl substituted pentoseoligosaccharides (S1 Table 4). In a corresponding manner mono coumaroylated pentose oligomers were found in the overlap region with retention time from 3.7-5.5 (S1 Table 4).

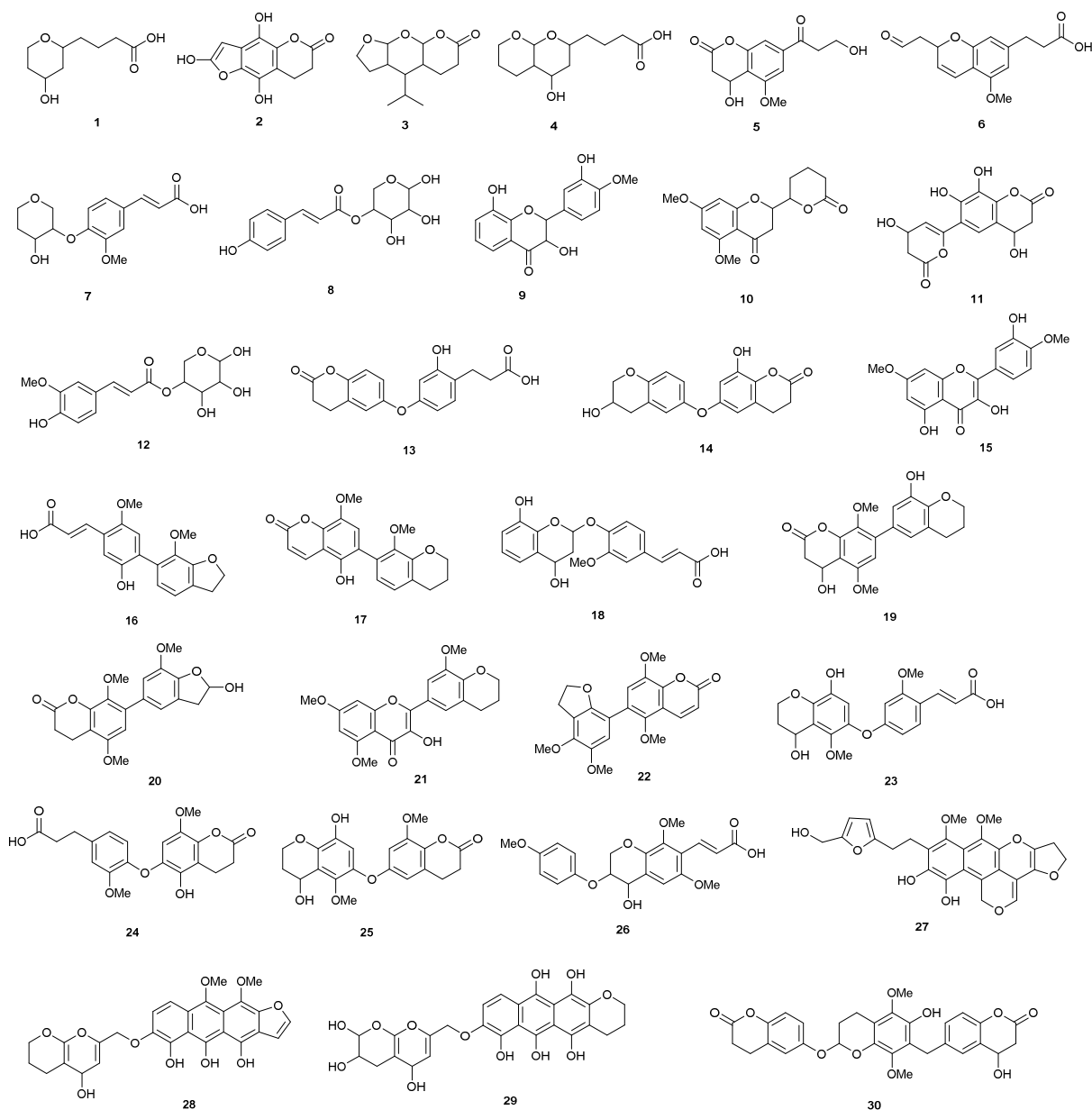


Figure 3. Proposed structures from LC-MS/MS. Note that ring substitution positions are given for clarity, but regioisomerism was not determined.

The finding of mono feruloylated pentoseoligomers and feruloylated and acetylated pentoseoligomers (of DP 1-9) agreed with data reported earlier²² that thermochemically (0.22 w/w H₂SO₄) pretreated corn fiber contains feruloylated xylooligomer analogues with large side chains, ethyl glycosides as well as acetyl substituents

The MS analysis in positive ionisation mode also revealed the presence of compounds containing nitrogen in both the water fraction and the 2-butanone fraction. The vast majority of these compounds were equal in accurate mass, retention time and had approximately equal intensity and were thus likely not decisive for enzyme inhibition. An initial analysis did show one

compound with accurate mass corresponding to acetyl azepinone. This compound was reported previously²³ to be a Maillard reaction product formed in acidic glucose/glycine rich solutions.

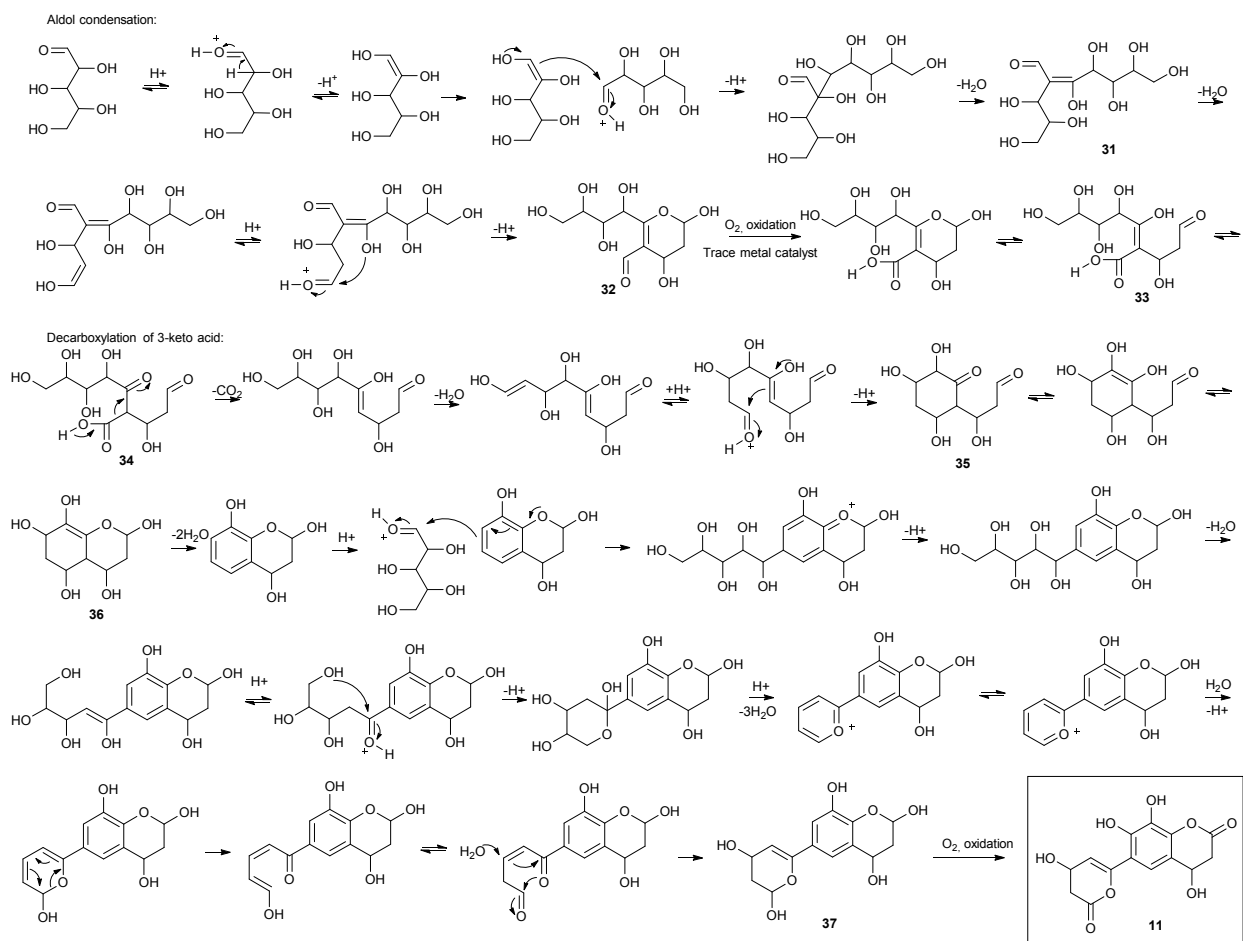


Figure 4. Proposed reaction mechanisms involving only pentose self-condensation reactions.

Origin of the identified compounds - reaction mechanisms

Consideration of the possible reaction mechanisms involved in the formation of the identified oligophenols indicates that the oligophenols arise from a wide array of reactions and reactants. Some of the compounds appear to arise from pentose self-condensation reactions alone (Figure 4, Figure 5A), while other compounds are most likely reaction products from pentose reactions with lignin degradation compounds (Figure 5B and Figure 5C) or result from additional routes involving only lignin degradation compounds (Figure 5D). These considerations were confirmed with spiking experiments with xylose and wheat arabinoxylan in pretreatment of wheat straw (Section 2) which increased the amount of oligophenolic compounds with 70 % and 21 % respectively.

Reactions involving only pentoses. The reaction mechanism for formation of **11**, as an example (Figure 4), is proposed to be initiated with an acid catalysed aldol condensation between two pentoses resulting in compound **31**. Xylose was abundant in the LfHP (data not shown) and is therefore likely to be involved in the reaction (arabinose was present in less

amounts, but may also be involved). In the suggested reaction, the aldol condensation is followed by further dehydration and cyclisation, resulting in the aldehyde substituted oxane **32**. The aldehyde is oxidised to the carboxylic acid by molecular oxygen, present during the reactions, presumably catalysed by trace metals present. The oxane hemiacetal is in equilibrium with its open chain form **33**, where the enol function is proposed to tautomerise to its keto form, resulting in the 3-keto acid **34**, that readily undergoes decarboxylation under acidic conditions. A dehydration gives a terminal carbonyl group, that is subject to nucleophilic attack from the enol, that acts as a carbon nucleophile, resulting in the ketone substituted cyclic compound **35**. The ketone is proposed to tautomerise to its enol form, resulting in the possibility of a new ring closure to the bicyclic hemiacetal **36**. Compound **36** is suggested to undergo further dehydration and protonation and subsequently undertake nucleophilic attack on a third pentose molecule. A cyclisation of the pentose moiety leads to a tricyclic compound that further dehydrates and undergoes an electrocyclic ring opening.

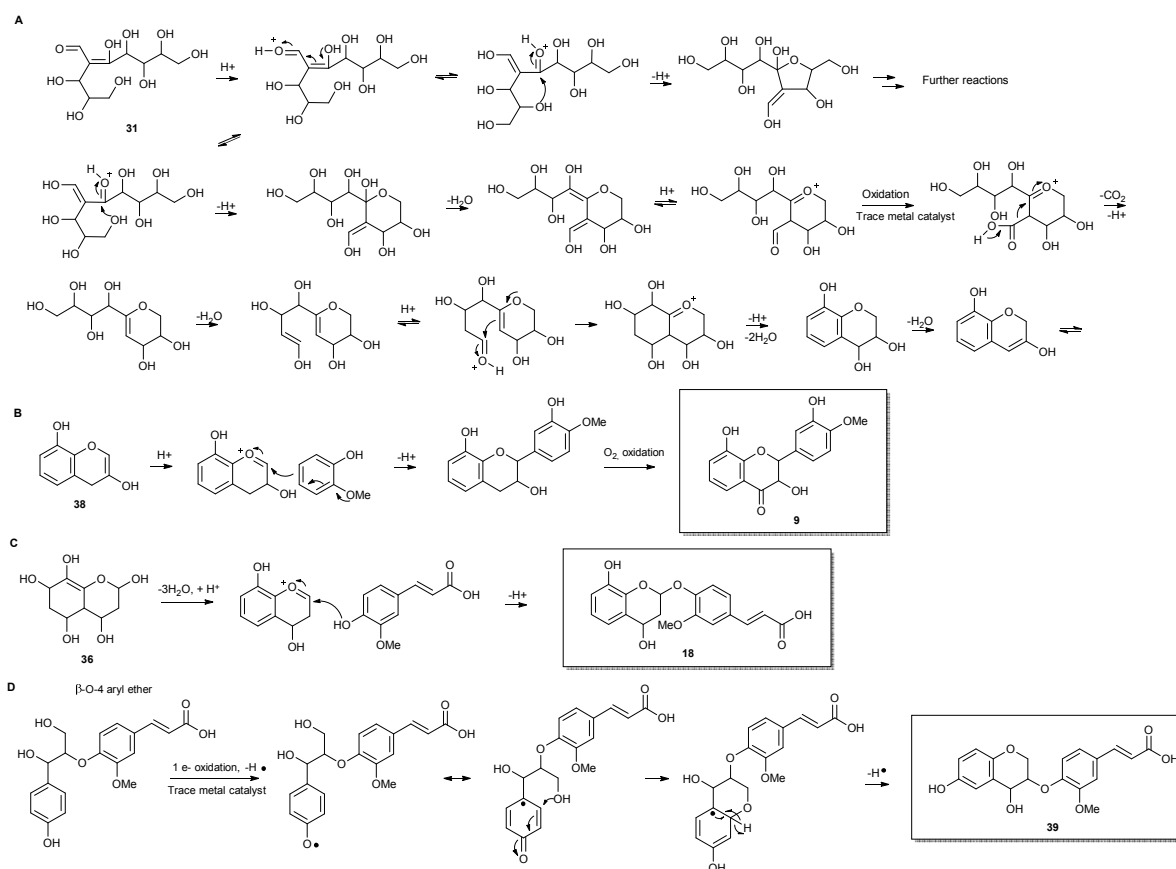


Figure 5. Examples of reactions leading to other of the discovered compounds. A) Variations in pentose self-condensation reactions. B) Reaction with methoxylated phenol resulting in methoxylated compounds and C) Reaction with ferulic acid resulting in the common (methoxyphenyl)prop-2-enoic acid motif. D) The common β -O-4 aryl ether in lignin - here the aryl is ferulic acid.

Nucleophilic attack from water facilitates ring closure to compound **37**, which undergoes oxidation to the final product **11** due to atmospheric oxygen present during pretreatment. Previously reported data²⁴, show that phenols readily undergo oxidation under similar reaction conditions.

Many variations of the reaction mechanism described for **11**, as an example, are possible. For example, the first cyclisation of compound **31**, will lead to a six membered ring with a different substituent pattern, if the nucleophilic attack takes place as a 1,4 addition to an α,β unsaturated carbonyl compound (Figure 5A). In this case the molecule will lack oxygen in the 2-position. Furthermore, a five membered ring i.e. an oxolane can be formed by nucleophile attack from the appropriate hydroxyl group (Figure 5A). Variations of the mechanisms in the proposed reaction schemes and reaction with other nucleophiles as for example lignin degradation compounds (Figure 5B and Figure 5C), explain the many different reaction products formed.

Reactions involving pentoses and lignin degradation compounds. In addition to pentose self-condensation

reactions, reactions involving other compounds in the reaction mixture such as methoxylated phenols, ferulic and coumaric acids arising from lignin degradation may take place (Figure 5B and Figure 5C). Methoxylated phenols explain the common methoxylations as exemplified with **9** (Figure 5B). It is worth noting that flavonoids can actually be formed as a reaction product during pretreatment.

Another common structural repeater was (methoxyphenyl)prop-2-enoic acid, which can be explained by reaction of ferulic acid with pentoses (Figure 5C) as exemplified with **18**. The ferulic acid moiety can in turn give rise to the corresponding bicyclic lactone by intramolecular esterification when a hydroxyl substituent is in *ortho* position. The lactone in the bicyclic moiety can be further dehydrated and/or react with other components in the reaction mixture. Furthermore, non methoxylated (phenyl)prop-2-enoic acid moieties are also proposed to be present and can, apart from reactions involving coumaric acid from lignin degradation, be explained by acid catalysed ring opening of lactones origination from pentose reactions alone (Figure 4).

Another possible isomer of compound 18 was found. Compound 39 can be explained by lignin degradation, where a β -O-aryl ether is proposed to undergo a one electron oxidation catalysed by trace metals as depicted in Figure 5D. A radical mechanism is involved in the ring closure resulting in the bicyclic compound with a ferulic acid substituent.

Protection of pentoses at the anomeric position

The above findings make it very clear that pentoses are heavily involved in formation of the identified inhibitors. Many of the new oligophenol compounds are possibly also precursors to pseudo lignin. This proposition is supported by the findings from ²⁵ where pretreated holocellulose, isolated from bamboo, was found to generate pseudolignin which was more rich in alicyclic and hydroxyl substituted structures, than reference lignin. In addition, the pseudolignin went from being rich in aliphatic structures, to being rich in aromatic structures, with increasing pretreatment time. Furthermore ¹⁸ have shown that pseudo lignin is formed from carbohydrates and that especially xylan and xylose are prone to undergo degradation, even at low severities, and precipitate as spheres of pseudo lignin on cellulose surfaces.

The degradation of pentoses is a double negative, because pentoses are now also becoming desired products in biorefineries. The pentoses are thus both lost, and at the same time unwanted products as inhibitors and/or pseudolignin are formed.

On this basis we speculated, that protection of the anomeric center could reduce the level of pentose degradation, even with no protection of the other hydroxyl groups. Anomeric protection could be possible in an industrial process of biomass pretreatment with a Fischer-type glycosylation with glycol ethers as protection groups. Glycol ethers have the desired high boiling points for biomass pretreatment and are water miscible, which is also important, because some water has to be present in order to hydrolyse the pentose oligosaccharides in the biomass to monomers, but water may also obstruct the protection reaction. An acid catalyst must be present in order for the protection reaction to take place, but the presence of acid will only additionally facilitate structural breakage of the biomass, which is the purpose of the biomass pretreatment. Ethylene glycol monobutyl ether (Figure 6) was tested in a pretreatment of wheat straw in the presence of water. It showed good results with respect to protection of the anomeric center and in reduction in the level of degradation compounds.

The optimal parameters for the pretreatment with anomeric protection results from an interplay of several factors in the pretreatment reaction mixture. 1) Disruption of xylan in the biomass to xylooligosaccharides. 2) Hydrolysis of xylooligosaccharides to xylose. 3) Reaction of xylose with the glycol ether to both the α and β anomer of the xyloside. 4) Possibly also reaction at the reducing ends of xylooligosaccharides with the glycol ether to form both anomers, followed by hydrolysis/cleavage of a protected xylose from the particular xylooligomer. As such the

parameters were optimised toward the maximum amount of protected xylose (sum of both anomers evaluated from HPLC, Supplementary Information, Figure 2), rather than the actual yield, based on the strategy that optimal protection would also result in minimum oligophenolic inhibitor formation.

A statistical experimental design was conducted in order to identify the optimal parameters for the protection reaction. Protection of the anomeric center (100 °C, 30 minutes, 18 equivalents of water (to xylose), 4.5 % H₂SO₄ (w/w of solution)) reduced the level of degradation compounds by 73 % compared to the original pretreatment method investigated in this paper (183 °C, 18 minutes, 100 % water) (Figure 6). The control experiment to the ethylene glycol monobutyl ether pretreatment, with no glycol ether present, only reduced the level of degradation compounds by 54 %. These are very promising results with respect to protection of xylose to hinder xylose degradation to the group of degradation compounds/inhibitors identified in this paper. However, the deprotection of ethylene glycol monobutyl ether protected xylose, by addition of water, as well as the separation of ethylene glycol monobutyl ether from the aqueous reaction mixture need further optimisation. Only after a full technology development of the separation can the xylose recovery and options for ethylene glycol monobutyl ether recycling be fully quantitatively evaluated, but quantitative estimates are given in the next section.

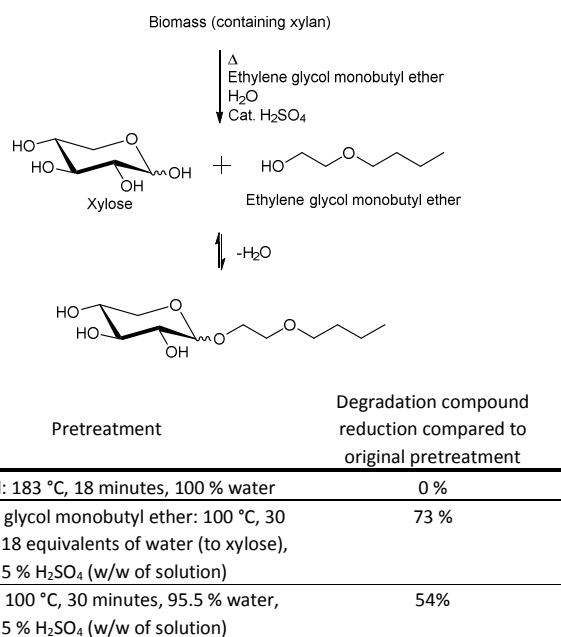


Figure 6. Protection of the anomeric center in *in situ* generated xylose from biomass with ethylene glycol monobutyl ether and reduction in the level of degradation compounds/inhibitors during a pretreatment of wheat straw.

Pretreatment outcome	Hydrothermal pretreatment	Pretreatment with anomeric protection
Cellulose digestibility for the enzymes	Low	High
Inhibitor level	High	Low
Xylose degradation	High	Low

Figure 7. Concept of pretreatment with anomeric protection compared to the current pilot plant hydrothermal pretreatment (please refer to Figure 6 for pretreatment parameters).

Quantitative considerations and perspectives

The current pilot plant biomass pretreatment (183 °C, 18 minutes, 100 % water) has been optimised towards minimum pentose degradation and minimum enzyme inhibitor formation, whilst at the same time aiming at enhancing the cellulose degradability. As such the oligophenolic enzyme inhibitor compounds in the 2-butanone fraction from extraction of pilot plant LfHP (S1 Figure 3) constituted 11 ± 2 % (w/w) of the total weight of compounds in the 2-butanone and water fractions (the mass balance of the 2-butanone extraction was 88 ± 4 %). However, the trade-off with optimisation towards minimum pentose degradation and minimum enzyme inhibitor formation is that cellulose digestibility for the enzymes is not optimal, due the mild pretreatment conditions. For the pretreatment with anomeric protection, calculations based on the reduced level of inhibitors of 73 % determined above and an estimated increase in cellulose digestibility implies an increase in glucose yield of approximately 15 % after the enzymatic hydrolysis. The pretreatment with anomeric protection can be increased further towards cellulose digestibility without the negative effect of pentose degradation to oligophenolic inhibitor compounds due to the protected pentose. Another, at least as important consideration is the loss of xylose, which for example is utilised in mixed fermentation of xylose and glucose to ethanol in biorefineries. A quantitative scenario as follows may be realistic when considering that in the current pilot plant pretreatment process approximately 25 % of the xylose is lost due to degradation to inhibitor compounds during pretreatment. Calculation based on the reduced level of inhibitor compounds in pretreatment with anomeric protection, implies a xylose recovery of up to 93 %. Thus approximately 20 % more xylose can be deployed in a mixed glucose-xylose fermentation to ethanol which adds up to a total increase in ethanol from both xylose and glucose fermentation to 30 %. Even if only half of these estimated levels are attained in practice, the gains would be significant.

These calculations illustrate the concept of pretreatment with anomeric protection, which is illustrated in Figure 7. The commercial and economic prospects of the new pretreatment with anomeric protection are currently under evaluation.

Conclusions

The formation of inhibitors during pretreatment of lignocellulosic feedstocks is a persistent problem, and notably the compounds that retard enzymatic cellulose conversion represent an obstacle for achieving optimal enzymatic productivity and high glucose yields.

A solvent extraction approach was employed for liquid separation of compounds of different polarity from pilot scale hydrothermal pretreated wheat straw. The compounds extracted into 2-butanone were found to exert particularly profound inhibition of cellulolytic glucose release on both pure cellulose and pretreated wheat straw fibers, but also the compounds left in the aqueous phase exerted profound inhibition. The inhibitors were identified with LC-MS/MS as oligophenols and xylooligosaccharides respectively. Although the xylooligosaccharide were more abundant, the oligophenols were significantly more potent cellulase inhibitors when compared on a dry matter basis.

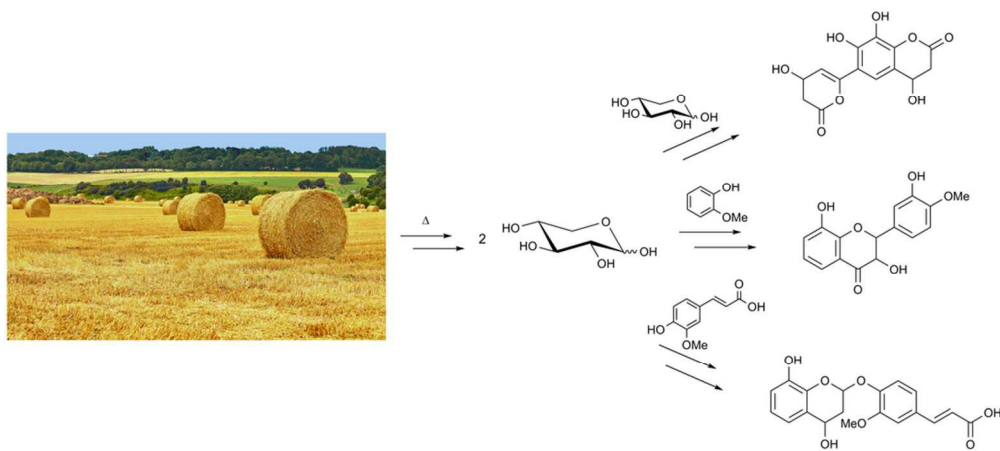
In total, 28 new inhibitory oligophenol compounds were identified from hydrothermally pretreated wheat straw liquor and their reaction mechanisms for formation were considered. The most important conclusion from this work was that reactions arising from xylose (or pentoses in general) are heavily involved in formation of the new compounds. The solution to avoid degradation of pentoses to form inhibitors is to protect the reactive anomeric center. The amount of degradation compounds, identified in this study, was reduced with 73 % by pretreatment with ethylene glycol monobutyl ether, compared to the original pretreatment. The results obtained not only provided new information of the compounds present in pretreated wheat straw biomass liquor, but pave the way for implementation of new types of reactions that prevent inhibitor formation in biomass processing.

Acknowledgements

This work was supported by the Danish National Advanced Technology Foundation *via* the Technology Platform "Biomass for the 21st century – B21st".

Notes and references

- 1 M. J. Climent, A. Corma and S. Iborra. *Green Chem.*, 2014, **16**, 516–547.
- 2 J. Han, S. M. Sen, D. M. Alonso, J. A. Dumesic and C. T. Maravelias. *Green Chem.*, 2014, **16**, 653–661.
- 3 J. S. Luterbacher, J. M. Rand, D. M. Alonso, J. Han, J. T. Youngquist, C. T. Maravelias, B. F. Pfleger and J. A. Dumesic. *Science*, 2014, **343**, 277–280.
- 4 T. D. Matson, K. Barta, A. V. Iretskii and P. C. Ford. *J. Am. Chem. Soc.*, 2011, **133**, 14090–14097.
- 5 R. Rinaldi. *Angew. Chem. Int. Ed.*, 2014, **53**, 8559–8560.
- 6 H. Rasmussen, H.R. Sørensen and A.S. Meyer. *Carbohydr. Res.*, 2014, **385**, 45–57.
- 7 R. Kont, M. Kurasin, H. Teugjas and P. Väljemäe. *Biotechnol. Biofuels.*, 2013, **6**, 135.
- 8 J. Zhang and L. Viikari. *Bioresour. Technol.*, 2012, **117**, 286–291.
- 9 M. J. Baumann, K. Borch and P. Westh. *Biotechnol. Biofuels*, 2011, **4**, 45.
- 10 Q. Qing, B. Yang, and C. E. Wyman, C. E. *Bioresour. Technol.*, 2010, **101**, 9624–9630.
- 11 A. Tejirian and F. Xu. *Enzyme Microb. Technol.*, 2011, **48**, 239–247.
- 12 S. I. Mhlongo, R. d. Haan, M. Viljoen-Blooma and W. H. van Zyla. *Enzyme Microb. Technol.*, 2015, **81**, 16–22.
- 13 Y. Kim, E. Ximenes, N. S. Mosier and M. R. Ladisch. *Enzyme Microb. Technol.*, 2011, **48**, 408–415.
- 14 V. D. Mitchell, C. M. Taylor and S. Bauer. *Bioenerg. Res.*, 2014, **7**, 654–669.
- 15 I. Forsskåhl, T. Popoff and O. Theander. *Carbohydr. Res.*, 1976, **48**, 13–21.
- 16 T. Popoff and O. Theander. *Carbohydr. Res.*, 1972, **22**, 135–149.
- 17 T. Popoff and O. Theander. *Acta Chem. Scand., Ser. B*, 1976, **30**, 397–402.
- 18 R. Kumar, F. Hu, P. Sannigrahi, S. Jung, A. J. Ragauskas and C.E. Wyman. *Biotechnol. Bioeng.*, 2013, **110**, 737–753.
- 19 L. Rosgaard, S. Pedersen, J. Langston, D. Akerhielm, J.R. Cherry and A.S. Meyer. *Biotechnol. Prog.*, 2007, **23**, 1270–1276.
- 20 A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter and D. Templeton. Determination of Sugars, Byproducts, and Degradation Products in Liquid Fraction Process Samples. 2006, NREL - Biomass Program.
- 21 F. Sánchez-Rabaneda, O. Jáuregui; I. Casals, C. Andrés-Lacueva, M. Izquierdo-Pulido and R. M. Lamuela-Raventós. *J. Mass Spectrom.*, 2003, **38**, 35–42.
- 22 M. M. Appeldoorn, P. de Waard, M. A. Kabel, H. Gruppen and H. A. Schols. *Carbohydr. Res.*, 2013, **381**, 33–42.
- 23 H. Wang, H. Qian and W. Yao. *Food Chem.*, 2011, **128**, 573–584.
- 24 H. R. Devlin and I. J. Harris. *Ind. Eng. Chem. Fundam.*, 1984, **23**, 387–392.
- 25 X. Ma, X. Yang, X. Zheng, L.Chen, L. Huang, S. Cao and H. Akinosho. *Cellulose*, 2015, **22**, 1687–1696



92x41mm (300 x 300 DPI)