



Carbohydrate stabilization extends the kinetic limits of chemical polysaccharide depolymerization

Ydna M. Questell-Santiago, Raquel Zambrano-Varela, Masoud Talebi Amiri  and Jeremy S. Luterbacher *

Polysaccharide depolymerization is an essential step for valorizing lignocellulosic biomass. In inexpensive systems such as pure water or dilute acid mixtures, carbohydrate monomer degradation rates exceed hemicellulose—and especially cellulose—depolymerization rates at most easily accessible temperatures, limiting sugar yields. Here, we use a reversible stabilization of xylose and glucose by acetal formation with formaldehyde to alter this kinetic paradigm, preventing sugar dehydration to furans and their subsequent degradation. During a harsh organosolv pretreatment in the presence of formaldehyde, over 90% of xylan in beech wood was recovered as diformylxylose (compared to 16% xylose recovery without formaldehyde). The subsequent depolymerization of cellulose led to carbohydrate yields over 70% and a final concentration of ~5 wt%, whereas the same conditions without formaldehyde gave a yield of 28%. This stabilization strategy pushes back the longstanding kinetic limits of polysaccharide depolymerization and enables the recovery of biomass-derived carbohydrates in high yields and concentrations.

Biomass is the second largest source of renewable carbon after atmospheric CO₂ and, as such, is an important sustainable alternative to fossil resources^{1–3}. The bulk of terrestrial biomass is represented by lignocellulosic biomass, which is mainly composed of 30–50% cellulose, 20–35% hemicellulose and 15–30% lignin^{2,4}. The development of viable biomass conversion platforms or biorefineries will require the selective conversion of biomass's constitutive biopolymers into high-value chemicals and fuels using cost-competitive processes^{3,5–7}. The first step of this conversion often involves the depolymerization of these three biopolymers and the production of single carbohydrates—mostly xylose and glucose—from the hemicellulose and cellulose fraction of the biomass, and phenylpropanoid monomers from the lignin fraction². Biomass-derived carbohydrates are vital platform molecules for the chemical or biological production of renewable fuels and materials, while lignin-derived monomers are useful precursors for renewable aromatic chemicals and drop-in fuels^{4,8,9}.

The production of carbohydrates from lignocellulosic biomass requires the separation of lignin and the cleavage of ether bonds in hemicellulose (mostly xylan) and cellulose chains while minimizing further degradation of the resulting carbohydrates^{10,11}. Most ether cleavage routes are based on acid-catalysed hydrolysis. However, the acidic conditions required for ether bond cleavage can also lead to dehydration to furans and eventual condensation to humins (Fig. 1). This competition between depolymerization and degradation governs most of the carbohydrate production routes being considered in biorefineries. A balance between these reactions is especially difficult to achieve for cellulose because its crystallinity restricts access to the ether bonds¹². Therefore, under most dilute acid (<20 wt% acid) conditions, degradation of glucose outpaces cellulose depolymerization, leading to low achievable yields². Cellulose depolymerization can outpace glucose degradation, but only at temperatures and residence times (10 ms to 1 min at 250 to 400 °C) that are impractical to achieve for solid biomass^{2,13}. These limitations

have led most biorefinery concepts to be developed around sophisticated systems that facilitate the disruption of cellulose crystallinity at low temperature, which in turn significantly increase the rate of cellulose depolymerization versus glucose degradation. These processes use solvents or catalysts such as concentrated mineral acids¹⁴, ionic liquids¹⁵ or cellulase enzymes¹⁶, the recovery or production of which represents a non-negligible fraction of process costs (>10–20%)^{2,16,17}. In the past decade, the costs of ionic liquid production and recovery and of cellulase enzyme production have dropped considerably^{18–20}. However, their cost still offers opportunities for developing alternative depolymerization processes. Recent work has shown that the use of polar aprotic solvents at high concentrations (>60%) with water accelerates the rates of all acid-catalysed reactions, but increases the rate of polysaccharide depolymerization reactions more relative to sugar degradation reactions, thus mitigating but not eliminating this issue^{10,21–23}.

One way to achieve high carbohydrate yields, despite the competing depolymerization and carbohydrate degradation reactions, is to decouple the first reaction (which occurs on the solid) from the second (which occurs in the liquid phase) by using a flow-through reactor (Supplementary Fig. 1). This reactor, which has been used in numerous biomass conversion studies^{24,25}, requires biomass to be immobilized as a packed bed in a heated zone through which the liquid flows. Both water and the aforementioned polar aprotic solvent systems have been used in these types of reactor^{10,25}. Carbohydrate degradation is then controlled by the residence time of the liquid in the packed bed, which, if decreased to limit degradation, increases product dilution and separation costs. The inverse relationship between yield and concentration in such a system usually means that a maximum concentration of sugars can be achieved as the residence time decreases, due to the declining yield of sugars that accompanies the reduced amount of solvent used (see review in ref. ²⁴ and data shown in this work). Although polar aprotic solvent systems allow for higher yields to be achieved with lower acid

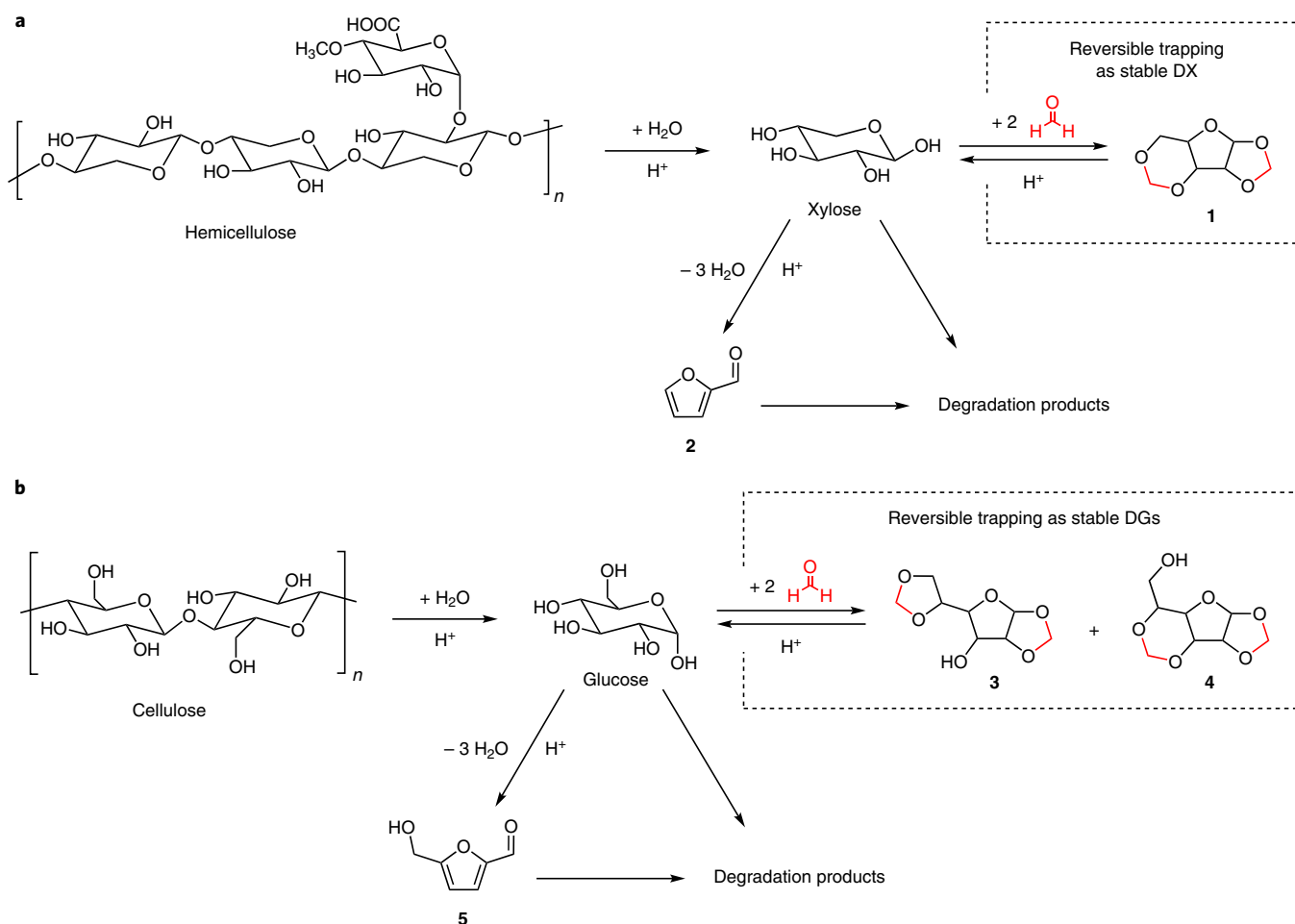


Fig. 1 | Carbohydrate stabilization using formaldehyde (FA). **a**, Preventing xylose degradation by reversibly forming molecule **1** through FA addition during acid-catalysed biomass pretreatment. **b**, Preventing glucose degradation by reversibly forming molecules **3** and **4** through FA addition during acid-catalysed cellulose depolymerization. DX, diformylxylose; DGs, the two main diformylglucose isomers. For simplicity, several intermediate steps are omitted (for example, isomerization of glucose to fructose on the way to **5**) in the dehydration and degradation steps.

concentrations, this tradeoff between dilution and yield still occurs and hampers process economics due to the importance of solvent recovery costs^{10,26–28}. So far, this tradeoff has prevented any commercial implementation of these flow-through systems.

In summary, this kinetic paradigm in biomass polysaccharide depolymerization has produced various strategies that have focused on modifying unfavourable depolymerization kinetics relative to sugar degradation, which have promoted expensive commercial solutions such as those using enzymes, concentrated mineral acids or ionic liquids. In this work, we propose an alternative approach that uses protection group chemistry to ‘trap’ carbohydrates in a stable form that cannot undergo dehydration/degradation. Specifically, we demonstrate that formaldehyde (FA) addition during polysaccharide depolymerization in polar aprotic solvents leads to a reaction with the hydroxyl groups in monosaccharides to form 1,3-dioxolane and 1,3-dioxane rings (Fig. 1). FA can be produced renewably from methanol²⁹ or syngas³⁰. This reaction can be easily reversed in aqueous conditions, leading to quantitative recovery of the FA and high carbohydrate yields compared to non-stabilized controls.

Results and discussion

Biomass pretreatment. We recently discovered that the addition of FA during pretreatment could stabilize lignin, preventing its condensation, and facilitate the conversion of extracted lignin to monomers at close to theoretical yields (47% for wild-type beech and 78%

for transgenic poplar)⁴. In that work we also noted that most of the xylose was recovered as a formylated molecule that we referred to as diformylxylose (DX, **1**). Here, we were interested in studying the ability of FA to act as a stabilizing trap for carbohydrates, including xylose. Using γ -valerolactone (GVL) as a pretreatment solvent, DX could be produced at yields above 90%, minimizing xylose degradation into furfural and other degradation products (Fig. 2c). In comparison, biomass pretreatment without FA (Control) using either 1,4-dioxane or GVL led to significant xylose degradation, with only 16% and 24% xylose recovery, respectively. This contrast in product yields is shown in the two-dimensional heteronuclear single-quantum coherence nuclear magnetic resonance (2D HSQC NMR) spectra of beech wood pretreatment liquors produced without and with the addition of FA. When no FA was added, only the peaks corresponding to lignin methoxy groups and remaining xylose were observed (Fig. 2a). In contrast, when FA was used, we observed peaks corresponding to the 1,3-dioxane and 1,3-dioxolane structures in stabilized lignin and DX (Fig. 2b). All these structures were identified by comparison with spectra of authenticated standards including lignin model compounds⁴, pure xylose and pure DX (Supplementary Fig. 2).

Stabilization of hemicellulose-derived sugars with FA was performed in various polar aprotic solvents, showing comparable product distribution (Fig. 2c). Protic solvents such as ethanol led to poor yields with FA due to the propensity of the FA to react with

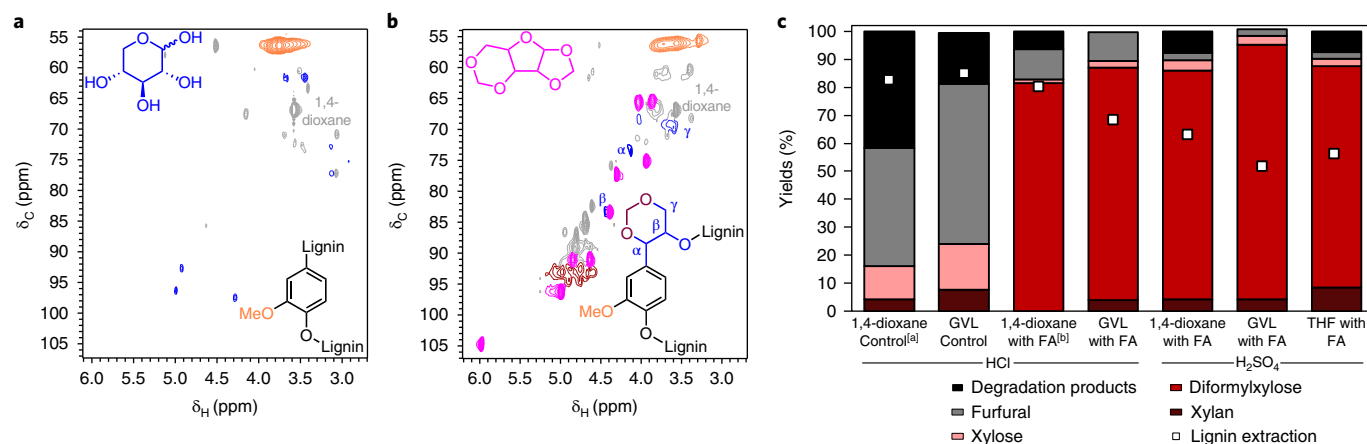


Fig. 2 | Hemicellulose stabilization and lignin extraction during biomass pretreatment. **a**, 2D HSQC NMR spectrum of biomass pretreatment liquor without FA addition (Control). **b**, 2D HSQC NMR spectrum of biomass pretreatment liquor with FA addition. δ_{H} , proton NMR chemical shift (in ppm); δ_{C} , carbon NMR chemical shift (in ppm). **c**, Product distribution resulting from biomass pretreatment in a 90/10 vol/vol organic solvent–water mixture at 100 °C for 2 h with 0.49 M of acid in the absence and presence of FA (~4 wt%). Open squares represent lignin extraction at each pretreatment conditions. [a] Experiment used for the 2D HSQC NMR spectrum showed in **a**. [b] Experiment used for the 2D HSQC NMR spectrum showed in **b**.

the solvent (Supplementary Fig. 3 and Supplementary Section 2.1). Hemicellulose stabilization was similar in GVL, 1,4-dioxane and tetrahydrofuran (THF) solvent systems, although GVL led to the highest DX yield of 91% (Fig. 2). Experiments with pure xylose confirmed that selectivity to DX was slightly better in GVL than 1,4-dioxane (Supplementary Fig. 4). The acidic proton has been shown to have a higher free energy in GVL–water compared to 1,4-dioxane–water systems, due to a decreased solvation free energy, leading to an increase in the acid-catalysed reaction rate, which could explain the selectivity difference^{21,31}. Additionally, when increasing the solid loading to 18 wt% in the GVL–water mixture (that is, decreasing the solvent usage by 60%), 80% of hemicellulose sugars could still be recovered as DX (Supplementary Table 1).

Cellulose depolymerization. After pretreatment, up to 93% of the cellulose fraction was recovered in the remaining solids (Supplementary Table 1). Our previous work has shown that solids extracted under similar conditions could undergo enzymatic hydrolysis⁴. However, for high glucose yields to be obtained, these pretreated solids had to be treated with a dilute acid solution (1 wt% H₂SO₄) before enzymatic treatment to remove grafted FA on the substrate's surface. Here, we instead attempted to depolymerize these pretreated solids using acid hydrolysis with GVL–water solvents, while again using the FA-stabilization strategy. Reactions were performed in a fixed-bed flow-through reactor (Supplementary Fig. 1) packed with rinsed pretreated solids (see composition in Supplementary Table 2) using a progressive temperature increase from 160 °C to 200 °C and 50 mM H₂SO₄ in GVL–water solutions as the solvent.

Similar results to those with xylose were observed with glucose and FA during depolymerization of the cellulose-rich pretreated solids. As glucose was produced, it rapidly formed stable glucose-derived molecules (referred to hereafter as diformylglucose or DG). Specifically, the presence of FA led to the formation of two main DG isomers from glucose by creating 1,3-dioxolane (**3**) and 1,3-dioxane (**4**) structures (Fig. 1b; for characterization see Supplementary Figs. 5 and 6). As with hemicellulose depolymerization, these structures stabilized the glucose that was formed at conditions that previously favoured glucose dehydration and degradation. 2D HSQC spectra of the cellulose hydrolysis liquor exiting the flow-through reactor when FA was used (Fig. 3b) revealed the presence of DG isomers, with peak assignment confirmed by the spectra of purified species (Supplementary Fig. 2). As expected, in the absence of

FA, the 2D HSQC spectra (Fig. 3a) showed chemical shifts that corresponded to pure glucose (Supplementary Fig. 2) instead of DG isomers.

The total product yields exiting the reactor were difficult to quantify directly due to the presence of oligomers. Therefore, we had the entire mixture undergo aqueous acid hydrolysis at 120 °C, which is the preferred method to hydrolyse oligomers, and reversed the formation of DG isomers to yield glucose. The total soluble carbohydrate yield and the resulting total carbohydrate concentrations were measured for different solvent residence times (that is, different quantities of solvents used) by varying the bed height in the reactor (represented in Fig. 3c as the final solvent-to-solid ratio). Without FA, we observed that carbohydrate yields decreased as soluble carbohydrates underwent longer residence times in larger packed-bed volumes. This decrease in yield eventually overtook the increase in concentration associated with a greater amount of biomass, such that we reached a maximum achievable concentration value at a solvent-to-solid ratio of 15:1 (in line with previous reports)^{10,24}. When FA was used, the resulting stabilization of monosaccharides minimized the formation of furans and humins and led to significant yield increases at high solid loadings (from 28 to 70% carbohydrate yield at a solvent-to-solid ratio of 11:1). As a result, under the conditions we explored, we no longer hit a maximum carbohydrate concentration, reaching up to 45 mg per g carbohydrates (compared to 19 mg per g achieved without FA). Increasing concentrations in solvent-based processes is one of the most economically critical aspects of such technologies, due to the high cost of solvent recovery²⁶. Interestingly, the effect of stabilization could be clearly observed by comparing the coloration of the stabilized and unstabilized solutions after reaction. The unstabilized solution was systematically significantly darker than the solution with FA, which was indicative of increased humins formation (Supplementary Fig. 7).

The difference in yield and concentration was slightly less pronounced when pure dried cellulose was used (Avicel PH 101, Supplementary Fig. 8). This smaller difference could be attributed to the significantly lower surface area of Avicel¹² compared to non-dried pretreated solids due to pore collapse during drying. However, in a biorefinery context, pretreated solids are a much more relevant substrate than Avicel, which has undergone extensive drying and compacting.

These results demonstrate a shift in the typical kinetic tradeoff seen for cellulose hydrolysis and glucose degradation when stabilization is used. The achieved yield of 70% carbohydrates with a

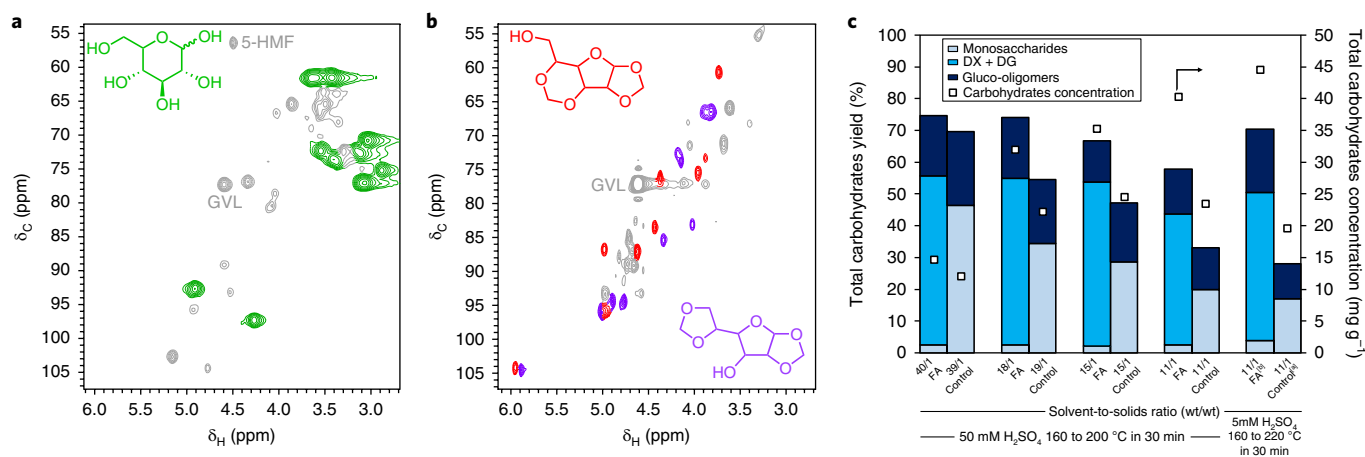


Fig. 3 | Soluble carbohydrates produced from cellulose-rich pretreated solids using a flow-through reactor. a, 2D HSQC NMR spectrum of the liquor exiting the flow-through reactor without FA addition. **b**, 2D HSQC NMR spectrum of the liquor exiting the flow-through reactor with FA addition. **c**, Total yield and concentration of soluble carbohydrates in the absence (Control) and presence of FA (7 wt%) as a function of total amount of solvent used. [a] Experiment used to produce the liquor used for the 2D HSQC NMR spectrum showed in **a**. [b] Experiment used to produce the liquor used for the 2D HSQC NMR spectrum showed in **b**. GVL, γ -valerolactone; 5-HMF, 5-hydroxymethylfurfural. Black arrow indicates the relevant axis for the data points.

concentration of 45 mg g^{-1} far exceeds what has been attained under aqueous conditions. Lee et al.²⁴ published an extensive review on the development and optimization of flow-through or ‘percolation’ reactors featuring the hydrolysis of cellulose in aqueous conditions and reported that, for a yield of 70% glucose, the expected carbohydrate concentration was, depending on the acid concentration, between 1.1 and 1.4 g per 100 ml ($\sim 11\text{--}14 \text{ mg g}^{-1}$) in a similar reactor—or three to four times lower than our result (Supplementary Fig. 9). Importantly, we also found that FA could be quantitatively recovered after deprotection in aqueous conditions ($>98\%$ recovery, which is within experimental error, Supplementary Section 2.3).

Despite FAs near-quantitative regeneration, its large-scale use will probably involve several challenges. Notably, several FA recovery steps would need to be implemented in a hypothetical process. Preliminary economic calculations (for details see Supplementary Section 2.4) based on previously published techno-economic analyses (TEA)²⁶ show that FA use could be worthwhile despite additional recovery steps in the context of diminishing solvent use in a GVL-based process. These additional steps also lead to costs within the range of those predicted for enzyme-based processes¹⁹. FA has been previously used in large-scale cellulose processing to increase the stretchability of viscose fibres^{32,33}, further demonstrating that its large-scale use is feasible. However, these past processes had to deal with the formation of insoluble FA side products in their spinning bath. Such possible side reactions, as well as regulatory issues linked to FA use and FA toxicity for subsequent biological sugar processing, are all aspects that remain unexplored and that will have to be addressed if FA is to be used in a biorefinery.

Diformylxylose conversion. Diformylxylose is a versatile molecule that is volatile and soluble in hexane—properties that could make it an interesting bio-derived fuel or lubricant additive. DX can also be converted back to xylose in aqueous environments (71% xylose yield; Supplementary Fig. 10). Because of its modified structure, we hypothesized that diformylxylose might be a more favourable substrate for furfural production. Furfural is typically produced by the Brønsted acid-catalysed dehydration of pentose, such as xylose, but this reaction can be further facilitated (especially at lower temperatures) by the combination of Brønsted and Lewis acids^{2,34,35}. The rate is accelerated with Lewis acids because they catalyse the isomerization of xylose via a hydrogen transfer (1,2-hydride shift), which easily leads to the furanose configuration necessary for producing

furfural (Fig. 4a)^{36,37}. This pathway has a lower activation energy than one-step xylose dehydration using only a Brønsted acid³⁵. We verified this by comparing the xylose production in a two-phase system at 160°C in the presence $0.15 \text{ M H}_2\text{SO}_4$ with and without a Lewis acid present (Fig. 4b). We observed significantly less xylose conversion and very low furfural production when no Lewis acid was present, whereas furfural yields around 70% could be achieved with a Lewis acid.

When DX was used as a starting product, similar furfural yields (65%) were achieved without Lewis acid addition (Fig. 4c). Interestingly, no difference in the reaction kinetics was observed when a Lewis acid was added at these conditions (Fig. 4c), which suggested that DX proceeded to furfural through a new mechanism (proposed in Fig. 4a, blue arrows). In the aqueous environment in which DX was partially present within the two-phase reaction system, DX hydrolysed, releasing one of its protective groups as FA and forming **6** (for characterization see Supplementary Fig. 11). Although **6** was not quantified by high-performance liquid chromatography (HPLC), it was identified as the main intermediate during operando ^{13}C NMR studies of the aqueous phase with labelled $1\text{-}^{13}\text{C}$ -diformylxylose (Supplementary Section 2.5 and Supplementary Figs. 12 and 13). Product analyses of these labelled reactions also showed that the furfural produced by the dehydration of DX had its labelled carbon in the same position as that produced from xylose (aldehyde carbon). These observations led us to propose a new mechanism for the production of furfural where **6** is protonated and undergoes ring-opening to form a carbocation in C1 (Fig. 4a), as is the case for unmodified xylose. Generally, the resulting hydroxyl group in C1 of unmodified xylose (or other carbohydrates) will rapidly deprotonate to form an aldehyde (aldose) that stabilizes the positive dipole induced by the ring-opening³⁸. The transition state (TS) was identified based on that of the analogous reaction of protonation of O5 in unmodified xylose³⁹. In the case of **6**, the aldehyde formation is blocked, and the ring-opening creates a carbocation in C1, forming **7**. This carbocation probably rearranges to a more stable configuration via a similar 1,2-hydride shift to that observed for Lewis acid-catalysed reactions (proton shown in bold in Fig. 4a, blue pathway). This intermediate is potentially more stable due its ability to resonate with the hydroxyl group in C2. We believe that this carbocation rearrangement occurs before any FA elimination reaction. Afterwards, these resonance structures could deprotonate, creating an aldehyde in C2 which could rapidly

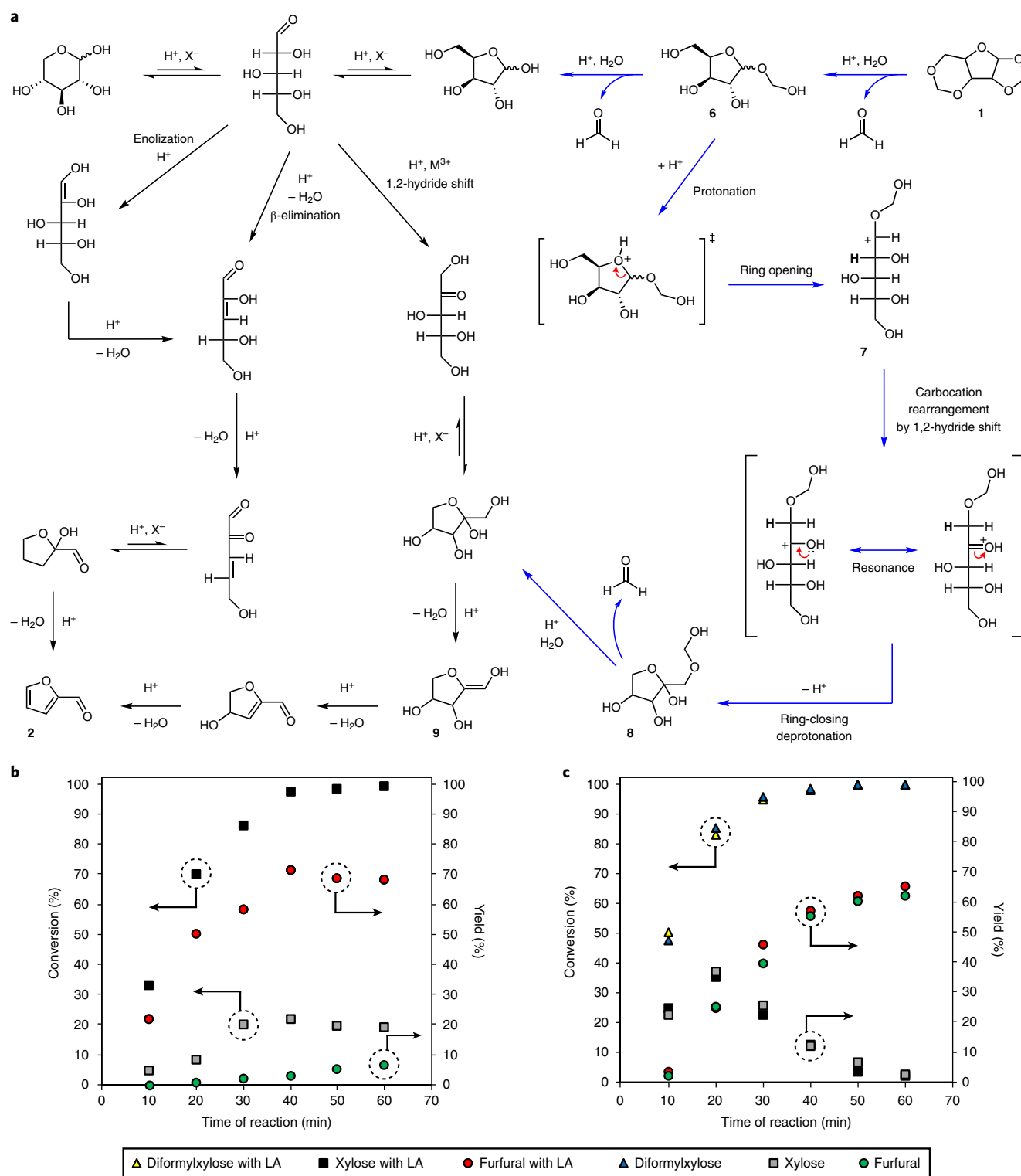


Fig. 4 | Pentose dehydration to furfural in biphasic water-alkylphenol reaction systems. **a**, Previously reported and proposed pentose dehydration mechanisms. Black arrows depict the conventional routes of xylose dehydration via the 1,2-hydride shift, enolization and β -elimination reactions⁴⁰. Blue arrows highlight a proposed mechanism from diformylxylose via the tandem hydrolysis-dehydration of diformylxylose to furfural. X^- , halides ions; M^{3+} , metal cations. **b,c**, Brønsted acid-catalysed production of furfural at 160 °C in a biphasic aqueous alkylphenol reaction system from xylose (with and without Lewis acid (LA) addition) (**b**) and diformylxylose (with and without Lewis acid addition) (**c**). Black arrows indicate the relevant axis for the set(s) of associated data points.

undergo ring-closing by a nucleophilic attack of O5 to C2, forming **8**. O5 also could attack before deprotonation, also forming **8**. Whether **8** hydrolysed first, releasing FA and forming xylulose

before subsequently dehydrating (drawn route), or if dehydration occurs first, followed by hydrolysis, is unknown. Nevertheless, both routes lead to **9**, which readily dehydrates to furfural.

The rate data suggest that this pathway's rate-determining step has a comparable activation energy to that of the tandem isomerization–dehydration of xylose–xylulose via the 1,2-hydrate shift, which is catalysed by the presence of Lewis acid (Fig. 4b,c), and could allow the production of furfural at lower temperatures without additional catalysts. Preliminary results using a mixture of DG and glucose did not show significant differences in 5-hydroxymethylfurfural (5) yields compared to pure glucose with only a Brønsted acid as the catalyst (Supplementary Fig. 14).

Conclusions

We have shown that the use of protection group chemistry during biomass depolymerization could stabilize carbohydrates and enable their production at conditions that were previously unfavourable due to dehydration and degradation reactions. Specifically, formaldehyde, a cheap bulk chemical that can easily be produced renewably, was used to stabilize xylose during biomass pretreatment and glucose during cellulose depolymerization in a flow-through reactor by formation of acetal groups. Because these reactions can easily be reversed under aqueous conditions, we systematically and significantly increased sugar yields compared to unstabilized controls. Using a flow-through reactor, we were able to produce a concentrated carbohydrate solution (~5 wt% before solvent separation) at a yield of 70%, which was almost three times the yield of the unstabilized control and three to four times more concentrated than the highly optimized aqueous system.

Currently available biomass depolymerization strategies for biomass largely rely on finding (often expensive) catalysts or ionic environments that favourably modify the kinetics of depolymerization versus degradation. The strategy presented here relies instead on 'trapping' the desired product while it is being produced in previously kinetically unfavourable conditions. Furthermore, it has the advantage of being fully compatible with lignin stabilization, which could facilitate the integrated chemical depolymerization of the three major biopolymers in biomass. This approach could lead to new processes for depolymerizing and valorizing biomass-derived carbohydrates or their protected equivalents.

Methods

Beech wood pretreatment. In a 60 ml glass reactor, 1.0–2.5 g of air-dried beech wood particles (<0.45 mm) were mixed with 9 ml of solvent (GVL, 1,4-dioxane, THF or EtOH), 420 μ l of 37 wt% HCl solution and 1 ml of 37 wt% FA solution. When sulfuric acid was used, 280 μ l of 95–97 wt% sulfuric acid and 340 μ l of water were added instead of the HCl solution. For control experiments (without FA), 690 μ l of water was added instead of 1 ml of the 37 wt% FA solution. The reactor was heated in a silicone oil bath set at 100 °C and mixed by a stir bar at 600 r.p.m. After reaction, the slurry was filtered and washed with 10 ml of the same solvent solution used for the reaction and then with 50 ml of water. The resulting solution was analysed by HPLC.

Solid residue depolymerization. Pretreated solids (1–5 g) mixed with 8 g of crushed silicon dioxide fused granules separated by quartz wool plugs (to avoid significant bed compacting and plugging) were packed into a flow-through reactor set-up (Supplementary Fig. 1). The reactor's heated zone was placed between two aluminium blocks within an insulated furnace. The solvent mixture was flowed through the system using a piston pump. The pressure was maintained constant at 15 bar by flowing nitrogen in the headspace of the liquid collector through a back-pressure regulator. At the beginning of the reaction, the packed reactor was heated to 150 °C in flowing nitrogen with a 20 min ramp. The temperature was allowed to stabilize from 150 to 160 °C for 3 min. Immediately afterwards, the solvent was flowed through the reactor at 2 ml min⁻¹ for 5 min, then changed to 1.5 ml min⁻¹, while a 30 min linear temperature ramp was applied from 160 to 200 °C or 160 to 220 °C, depending on the experiment. The collected liquid was analysed by HPLC and gas chromatography (GC) to quantify sugar monomers and DGs. Acid hydrolysis of the collected liquid was then performed and the resulting solution was analysed by HPLC to quantify the sugar monomers. The difference in monosaccharide yields between the hydrolysed and unhydrolysed samples was assumed to be sugar oligomers.

Diformylxylose synthesis. Approximately 15 g of pure xylose, 31.5 ml of 37 wt% HCl and 75 ml of 37 wt% FA were mixed in 675 ml of 1,4-dioxane and reacted at

80 °C for 30 min. The resulting solution was neutralized with sodium bicarbonate and dried under reduced pressure with a rotary evaporator at 60 °C. The residue was extracted five times with 300 ml of n-hexane. All organic phases were combined and dried under reduced pressure with a rotary evaporator at 60 °C. The residue was distilled at 120 °C and 0.1 mbar to obtain a yellowish solid. This solid was recrystallized in ethanol three times and analysed by NMR to confirm its structure by comparison with previously reported data¹ (Supplementary Fig. 2). This white solid crystal was then used as the diformylxylose standard.

Diformylglucose synthesis. Pure glucose (15 g) was reacted with 31.5 ml of 37 wt% HCl and 75 ml of 37 wt% FA in 675 ml of 1,4-dioxane at 80 °C for 30 min. Before separation, the solution was neutralized with sodium bicarbonate and dried under reduced pressure with a rotary evaporator at 60 °C. The residue was extracted five times with 500 ml of ethyl acetate. All organic phases were combined and dried under reduced pressure with a rotary evaporator set a 40 °C. The resultant residue was distilled at 125 °C and ~0.06 mbar to obtain a yellowish paste. The paste was diluted with MQ-water and filtered with a 0.2 μ m polytetrafluoroethylene (PTFE) filter. To obtain pure diformylglucose isomers, the solution was separated by HPLC and targeted peaks were collected using an automated fraction collector. Isomer 1 was completely isolated from the mixture by this technique, but isomer 2 always contained a certain amount of isomer 1, due to their similar retention times. Collected fractions were neutralized with sodium bicarbonate and dried under reduced pressure in a rotary evaporator at 40 °C. The solid residue was diluted with DMSO-*d*₆, filtered with a 0.2 μ m PTFE filter and used for gas chromatography–mass spectrometry and NMR analyses.

Pentose dehydration to furfural. Pentose dehydration was carried out in a 10 ml thick-walled glass reactor, into which 0.12 mmol of substrate (xylose or diformylxylose), 1.5 g of 0.15 M sulfuric acid in MQ-water and 3 g of 2-SBP were loaded. To start the reaction, the reactor was placed in a silicone oil bath at 160 °C and stirred at 600 r.p.m. For reactions with AlCl₃, 67 mol substrate per mol AlCl₃ was added. Reactors were cooled after the specified reaction times by placing them in a water bath at room temperature. For better phase separation, the reactor was centrifuged for 3 min at 4,000 r.p.m (3,005 g-force) by placing it inside a 50 ml centrifuge tube. The aqueous phase was collected with a glass pipette and analysed by HPLC. The organic phase was analysed by GC.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 29 January 2018; Accepted: 2 August 2018;

Published online: 17 September 2018

References

- Luterbacher, J. S. et al. Solvent-enabled nonenzymatic sugar production from biomass for chemical and biological upgrading. *ChemSusChem* **8**, 1317–1322 (2015).
- Luterbacher, J. S., Alonso, D. M. & Dumesic, J. A. Targeted chemical upgrading of lignocellulosic biomass to platform molecules. *Green. Chem.* **16**, 4816–4838 (2014).
- Questell-Santiago, Y. M. & Luterbacher, J. S. in *High Pressure Technologies in Biomass Conversion* (ed. Lukasik, R. M.) 9–36 (Royal Society of Chemistry, Crolydon, 2017).
- Shuai, L. et al. Formaldehyde stabilization facilitates lignin monomer production during biomass depolymerization. *Science* **354**, 329–333 (2016).
- Wyman, C. E., Cai, C. M. & Kumar, R. in *Encyclopedia of Sustainability Science and Technology* (ed. Meyers, R. A.) 1–27 (Springer, New York, 2017).
- Kumar, A. K. & Sharma, S. Recent updates on different methods of pretreatment of lignocellulosic feedstocks: a review. *Bioresour. Bioprocess.* **4**, 7 (2017).
- Alonso, D. M. et al. Increasing the revenue from lignocellulosic biomass: maximizing feedstock utilization. *Sci. Adv.* **3**, e1603301 (2017).
- Lan, W., Amiri, M. T., Hunston, C. M. & Luterbacher, J. S. Protection group effects during α,γ -diol lignin stabilization promote high-selectivity monomer production. *Angew. Chem. Int. Ed.* **57**, 1356–1360 (2018).
- Nyoo Putro, J., Edi Soetaredjo, F., Lin, S.-Y., Ju, Y.-H. & Ismadji, S. Pretreatment and conversion of lignocellulose biomass into valuable chemicals. *RSC Adv.* **6**, 46834–46852 (2016).
- Luterbacher, J. S. et al. Nonenzymatic sugar production from biomass using biomass-derived γ -valerolactone. *Science* **343**, 277–280 (2014).
- Haghighi Mood, S. et al. Lignocellulosic biomass to bioethanol, a comprehensive review with a focus on pretreatment. *Renew. Sustain. Energy Rev.* **27**, 77–93 (2013).
- Neuman, R. P. & Walker, L. P. Solute exclusion from cellulose in packed columns: experimental investigation and pore volume measurements. *Biotechnol. Bioeng.* **40**, 218–225 (1992).

13. Peterson, A. A. et al. Thermochemical biofuel production in hydrothermal media: a review of sub- and supercritical water technologies. *Energy Environ. Sci.* **1**, 32–65 (2008).
14. Bergius, F. Conversion of wood to carbohydrates. *Ind. Eng. Chem.* **29**, 247–253 (1937).
15. Binder, J. B. & Raines, R. T. Fermentable sugars by chemical hydrolysis of biomass. *Proc. Natl Acad. Sci. USA* **107**, 4516–4521 (2010).
16. Tao, L. et al. NREL 2012 Achievement of Ethanol Cost Targets: Biochemical Ethanol Fermentation via Dilute-Acid Pretreatment and Enzymatic Hydrolysis of Corn Stover (NREL, 2014); <https://doi.org/10.2172/1129271>
17. Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B. A. & Blanch, H. W. The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnol. Bioeng.* **109**, 1083–1087 (2012).
18. Gschwend, F. J. V., Brandt-Talbot, A., Chambon, C. L. & Hallett, J. P. in *Ionic Liquids: Current State and Future Directions* Vol. 1250 (eds Shiflett, M. B. & Scurto, A. M.) 209–223 (American Chemical Society, Washington DC, 2017).
19. Moriarty, K. L., Milbrandt, A. R., Warner, E., Lewis, J. E. & Schwab, A. A. 2016 Bioenergy Industry Status Report (NREL, Golden, 2018).
20. Liu, G., Zhang, J. & Bao, J. Cost evaluation of cellulase enzyme for industrial-scale cellulosic ethanol production based on rigorous Aspen Plus modeling. *Bioprocess. Biosyst. Eng.* **39**, 133–140 (2016).
21. Mellmer, M. A. et al. Solvent effects in acid-catalyzed biomass conversion reactions. *Angew. Chem. Int. Ed.* **53**, 11872–11875 (2014).
22. Mellmer, M. A., Alonso, D. M., Luterbacher, J. S., Gallo, J. M. R. & Dumesic, J. A. Effects of γ -valerolactone in hydrolysis of lignocellulosic biomass to monosaccharides. *Green Chem.* **16**, 4659–4662 (2014).
23. Ghosh, A., Bai, X. & Brown, R. C. Solubilized carbohydrate production by acid-catalyzed depolymerization of cellulose in polar aprotic solvents. *Chem. Select* **3**, 4777–4785 (2018).
24. Lee, Y. Y., Iyer, P. & Torget, R. W. in *Recent Progress in Bioconversion of Lignocellulosics* (ed. Tsao, G. T.) 93–115 (Springer, Berlin, 1999).
25. Yang, B., Tao, L. & Wyman, C. E. Strengths, challenges, and opportunities for hydrothermal pretreatment in lignocellulosic biorefineries. *Biofuels Bioprod. Bioref.* **12**, 125–138 (2018).
26. Han, J., Luterbacher, J. S., Alonso, D. M., Dumesic, J. A. & Maravelias, C. T. A lignocellulosic ethanol strategy via nonenzymatic sugar production: process synthesis and analysis. *Bioresour. Technol.* **182**, 258–266 (2015).
27. Eerhart, A. J. J. E. et al. Fuels and plastics from lignocellulosic biomass via the furan pathway; a technical analysis. *RSC Adv.* **4**, 3536–3549 (2013).
28. Eerhart, A. J. J. E., Patel, M. K. & Faaij, A. P. C. Fuels and plastics from lignocellulosic biomass via the furan pathway: an economic analysis. *Biofuels Bioprod. Bioref.* **9**, 307–325 (2015).
29. Qian, M., Liauw, M. A. & Emig, G. Formaldehyde synthesis from methanol over silver catalysts. *Appl. Catal. Gen.* **238**, 211–222 (2003).
30. Bahmanpour, A. M., Hoadley, A. & Tanksale, A. Formaldehyde production via hydrogenation of carbon monoxide in the aqueous phase. *Green Chem.* **17**, 3500–3507 (2015).
31. Shuai, L. & Luterbacher, J. Organic solvent effects in biomass conversion reactions. *ChemSusChem* **9**, 133–155 (2016).
32. Mössinger, D., Scheytt, H., Uihlein, K., Wunderlich, D. & Zimmerer, B. High-tenacity viscose multifilament yarn with low yarn linear density. US Patent US20150322595A1 (2015).
33. Ronald, R., Martin, R. T. & Richardson, W. C. Filaments of regenerated cellulose. US Patent US3388117A (1968).
34. Pagán-Torres, Y. J., Wang, T., Gallo, J. M. R., Shanks, B. H. & Dumesic, J. A. Production of 5-hydroxymethylfurfural from glucose using a combination of Lewis and Brønsted acid catalysts in water in a biphasic reactor with an alkylphenol solvent. *ACS Catal.* **2**, 930–934 (2012).
35. Choudhary, V., Sandler, S. I. & Vlachos, D. G. Conversion of xylose to furfural using lewis and brønsted acid catalysts in aqueous media. *ACS Catal.* **2**, 2022–2028 (2012).
36. Román-Leshkov, Y., Moliner, M., Labinger, J. A. & Davis, M. E. Mechanism of glucose isomerization using a solid Lewis acid catalyst in water. *Angew. Chem. Int. Ed.* **49**, 8954–8957 (2010).
37. Choudhary, V., Pinar, A. B., Sandler, S. I., Vlachos, D. G. & Lobo, R. F. Xylose isomerization to xylulose and its dehydration to furfural in aqueous media. *ACS Catal.* **1**, 1724–1728 (2011).
38. Robyt, J. F. *Essentials of Carbohydrate Chemistry* (Springer, New York, 1998).
39. Nimlos, M. R., Qian, X., Davis, M., Himmel, M. E. & Johnson, D. K. Energetics of xylose decomposition as determined using quantum mechanics modeling. *J. Phys. Chem. A* **110**, 11824–11838 (2006).
40. Danon, B., Marcotullio, G. & Jong, W. Mechanistic and kinetic aspects of pentose dehydration towards furfural in aqueous media employing homogeneous catalysis. *Green Chem.* **16**, 39–54 (2013).

Acknowledgements

This work was supported by the Swiss National Science Foundation through grant PYAPP2_154281, by the Swiss Competence Center for Energy Research: Biomass for a Swiss Energy Future through the Swiss Commission for Technology and Innovation grant KTI.2014.0116, and by EPFL. The authors thank M. Studer from Bern University of Applied Sciences (Switzerland) for providing beech wood.

Author contributions

Y.M.Q.S. and J.S.L. conceived of the idea, designed the study and wrote the manuscript. Y.M.Q.S. performed all experiments unless noted otherwise, and analysed the data. R.Z.V. performed xylose and diformylxylose dehydration reactions, and assisted in purifying diformylxylose and diformylglucose. M.T.A. produced some of the pretreated solids and performed the preliminary techno-economic calculations. All authors edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41557-018-0134-4>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.S.L.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.