Advanced model compounds for understanding acid catalyzed lignin depolymerization: identification of renewable aromatics and a lignin-derived solvent

Ciara W. Lahive,‡± Peter J. Deuss,‡± Christopher S. Lancefield,† Zhuohua Sun,‡ David B. Cordes,‡ Claire Young,† Fanny Tran,† Alexandra M. Z. Slawin,† Johannes G. de Vries,‡§ Paul C. J. Kamer,† Nicholas J. Westwood,*‡ Katalin Barta*‡

† School of Chemistry and Biomedical Science Research Complex, University of St Andrews and EaStCHEM, North Haugh, St. Andrews, Fife, KY16 9ST, United Kingdom
‡ Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands
§ Leibniz-Insitut für Katalyse e.V., Albert-Einstein-Straße 29a, 18059 Rostock, Germany

ABSTRACT: The development of fundamentally new approaches for lignin depolymerization is challenged by the complexity of this aromatic biopolymer. While overly simplified model compounds often lack relevance to the chemistry of lignin, the use of lignin streams directly, poses significant analytical challenges to methodology development. Ideally, new methods should be tested on model compounds that are complex enough to mirror the structural diversity in lignin, but still of sufficiently low molecular weight to enable facile analysis. In this contribution we present a new class of advanced (β-O-4)-(β-5) dilinkage models that are highly realistic representations of a lignin fragment. Together with selected β-O-4, β-5 and β-β structures, these compounds provide a detailed understanding of the reactivity of various types of lignin linkages in acid catalysis in conjunction with stabilization of reactive intermediates using ethylene glycol. The use of these new models has allowed for identification of novel reaction pathways and intermediates and led to the characterization of new dimeric products in subsequent lignin depolymerization studies. The excellent correlation between model and lignin experiments highlights the relevance of this new class of model compounds for broader use in catalysis studies. Only by understanding the reactivity of the linkages in lignin at this level of detail can fully optimized lignin depolymerization strategies be developed.

INTRODUCTION

The efficient depolymerization of lignin is one of the major challenges in the full valorization of renewable lignocellulose resources23, and requires fundamentally new catalytic methods.3-4 However, the development of new approaches is particularly challenging due to the complexity of this aromatic polymer.3-5 Methodology development is often done on overly simplified model compounds.6 On the other hand, the work with real lignin streams directly is tedious and leads to extensive analytical challenges including the structural determination of the starting material and the characterization of complex product mixtures.24-38 Therefore, the synthesis of new, more advanced model compounds is highly desired and of general importance in this field.

Lignin contains different aromatic subunits (H, G, S) and various types of linkages (Figure S1).24,34,5 The occurrence of these linkages varies greatly depending on the plant type and pre-treatment methods used. Thus far, most studies have focused on the cleavage of the most abundant β-O-4 linkage using predominantly simple model compounds.23-38,6,8 Much less effort has been devoted to understanding the chemistry of other types of linkages such as the β-β5 and β-β10 (Figure S2).10

It has become increasingly important to develop more sophisticated model compounds911,13 that reflect the complexity of the native lignin structure. To the best of our knowledge, synthetic pathways to model compounds that combine multiple linkage types, contain all lignin-relevant functional groups and at the same time are of limited molecular weight have not yet been developed. In this contribution we provide scalable synthetic paths to access such advanced lignin model compounds and demonstrate their value in understanding the reactivity of the main linkages in real lignin feedstocks under depolymerization conditions.

The new class of advanced model compounds (AB1-4) are a combination of the β-O-4 and the β-5 linkage and contain phenolic and non-phenolic units (Figure 1). Variations on the β-O-4 side include guaiacyl (AB1 and AB3) and syringyl (AB2 and AB4) end groups. The β-5 moiety contains either a non-phenolic (AB1 and AB2) or phenolic end group (AB3 and AB4), whereby the methoxy simulates an internal β-5 linkage, while the phenolic group mimics a terminal β-5 linkage or the result of a cleaved β-O-4 linkage.
The reactivity of these model compounds (AB1-4) was subsequently evaluated in a catalytic method we have previously pioneered, which comprises of acidolysis in conjunction with the stabilization of reactive intermediates under acetal formation conditions. In addition to AB1-4, model compounds representing the β-β lignin linkage (C1-C3) were selected for study. Furthermore, models A6 and B6 were selected for studying the isolated reactivity of the β-O-4 and β-5 linkages, respectively. Using a combination of these models (Figure 1), we were able to gain deeper understanding of the overall reactivity of lignin under these conditions. New reaction pathways and intermediates were established and important products have been identified in actual lignin depolymerization mixtures.

RESULTS & DISCUSSION

Synthesis of novel (β-O-4)-(β-5) lignin model compounds

To access the novel (β-O-4)-(β-5) models AB1-4, a divergent synthetic methodology was developed that allowed access to both non-phenolic (AB1-2) and phenolic (AB3-4) models (Schemes 1 and 2). Starting from commercially available ferulic acid (1) esterification with MeOH/TMSCl gave methyl ferulate (2) which, when treated with silver(I) oxide, underwent an oxidative dimerization to yield diferylulate (3). This reaction is believed to proceed via a radical mechanism which is under thermodynamic control yielding the racemic trans-diferulate which possesses the same stereochemistry as the β-5 units in lignin. Methylation of the phenol in 3 using CH3I/K2CO3 gave 4 (Table S1) and subsequent oxidative cleavage of the alkene in 4 using the RuCl3/NaIO4 system afforded aldehyde 5. The relative stereochemistry of the β-5 motif in compounds 4 and 5 was confirmed by X-ray crystallography (Section S4.2).

The β-O-4 moiety was installed by aldol reaction between 5 and 6G to afford the di-ester 7G in 82% yield. In this unoptimized aldol protocol, a mixture of both the anti- (erythro) and syn- (threo) stereochemistry at the new stereogenic centres was formed in a 3:1 ratio as determined by quantitative 1H NMR analysis of the crude reaction mixture (Figure S3). Partial separation of the isomers could be achieved by column chromatography (Section S4.1). However, in general, isomeric mixtures at the β-O-4 linkage (A and AB1-4) were prepared and used throughout this work for two main reasons: (i) in real lignin the β-O-4 linkage is known to be present as a mixture of both anti- and syn-isomers and (ii) in acid mediated lignin degradation the reaction proceeds via a common intermediate from both the anti- or syn-isomer.

The diastereomeric mixture of 7G was reduced using NaBH4/MeOH in EtOH to give AB1 in 90% yield without separation of the anti- and syn-isomers. However, anti- and syn-diastereomers of AB1 were obtained on a small scale from the separated isomers of precursor di-ester 7G (Section S4.1). Similarly, an aldol reaction between 5 and 6S provided 7S in 80% yield, which upon reduction gave the desired product AB2 as a diastereomeric mixture in 96% yield.

Figure 1. A summary of model compounds A, B, C1-3 used during our catalytic studies, including novel β-O-4-β-5 dilinkage model compounds (AB1-4) synthesized in this work.
To access the phenolic model compounds AB3 and AB4 a protecting group strategy was employed (Scheme 2). Protection of the phenolic group in 3 with TBSCI/imidazole afforded TBS protected 8 in a quantitative yield with no need for further purification. From 8, following an analogous synthetic route via 9 and 10G or 10S as outlined previously, TBS protected models 11G and 11S were prepared and deprotected (TBAF) to give the phenolic models AB3 and AB4 as mixtures of diastereomers in 80% and 83% yield, respectively over the final two steps. With this set of novel models AB1-4 in hand we decided to study their reactivity in acid mediated lignin depolymerization in the presence of ethylene glycol.

Reactivity of (β-O-4)-(β-5) model compounds under acetal formation conditions

Acidolysis of lignin has received considerable attention due to the relevance of this method to the biorefinery concept. This approach was originally used to aid structural elucidation and more recently for the production of well-defined aromatic compounds. Using model compounds, two different reaction pathways (C2 and C3 pathways, Scheme 3) have been identified for the cleavage of the β-O-4 linkage and modification of the β-5 linkage. While the C3 pathway provides the Hibbert ketones, the C2 pathway yields C2-aldehydes upon release of formaldehyde, which can then undergo condensation reactions. The balance of these pathways depends on the nature of the mineral acid used. With HBr, the C3 pathway dominates whereas H2SO4 favours the C2 pathway. Similar observations were made regarding the reactivity of the β-5 linkage. Lundquist et al. studied the reactivity of a β-5 model compound with different acids in mixtures of 1,4-dioxane/H2O. While HBr gave mainly the C3-benzofuran product, triflic acid (HOTf) gave predominantly the C2-stilbene product.
Scheme 3. Known pathways for the acid mediated cleavage of the lignin \(\beta\-)O-4 linkage and the modification of the lignin \(\beta\-)5 linkage (\(R = H\) or OMe)

Scheme 4. Products identified in reactions of the (\(\beta\-)O-4)-(\(\beta\-)5) model compounds AB1-4 (See also Sections S6.0 and S9.1)

We have previously described the highly efficient cleavage of \(\beta\-)O-4 lignin model compounds using catalytic amounts of HOTf in conjunction with in situ stabilization of the resulting C2-aldehyde products as their ethylene glycol acetals. \(^{14}\) This concept was also extended to the depolymerization of lignin where re-condensation reactions were markedly suppressed. However, important questions remained unanswered regarding the reactivity of the \(\beta\-\beta\) and \(\beta\-)5 lignin linkages and the products originating from these moieties were not identified. Furthermore, the released formaldehyde was neither detected, nor quantified and its role in recondensation was not clarified. The models AB1-4, were ideally suited to answer these important questions.

General reactivity of (\(\beta\-)O-4)-(\(\beta\-)5) models AB1-4

First, the reactivity of AB1-4 was examined under the reaction conditions we have previously established (HOTf/ethylene glycol). \(^{14}\) Full substrate conversion was seen within 15 minutes resulting in the formation of guaiacol G (from AB1 and AB3) or syringol S (from AB2 and AB4) as determined by HPLC analysis (Scheme 4). These high yields of G and S were very similar to those found for simpler \(\beta\-)O-4 model compounds\(^{14}\) and demonstrated that the chemistry of the \(\beta\-)O-4 linkage was unaffected by the presence of the adjacent \(\beta\-)5 moiety.
Depending on the substrate used (AB1-2 or AB3-4), novel stilbene-acetals P1 or P2 were identified as the other major product (Scheme 4, Section S11.0). These products were likely formed by cleavage of the β-O-4 moiety in AB1-4 to give the C2-aldehyde, which reacted with ethylene glycol (Scheme 3). Subsequent ring opening of the β-5 moiety then occurred also via the C2-pathway. P1 and P2 were isolated and fully characterized with the E-stereochemistry being assigned based on the coupling constants observed between the two alkene protons (16.5 Hz and 16.4 Hz in P1 and P2 respectively, Section S11.0).

In control reactions in the absence of ethylene glycol (Section S9.2), the β-O-4 linkage was cleaved rapidly and the guaiacol G yields were retained. However, a significant difference was seen in the reactivity of the remaining component of AB1, which formed a mixture of oligomeric products (by GPC analysis, Section S7.0). In contrast, GPC analysis of the reaction in the presence of ethylene glycol gave only the desired low MW compounds. HPLC analysis also confirmed these observations (Figures S11 and S12) and similar results were obtained from AB3 (Sections S7.0 and S9.0).

**Product formation profiles and reaction intermediates using (β-O-4)-(β-5) model AB1**

To gain further insight, the acidolysis of AB1 was studied in the presence of ethylene glycol and product formation profiles were recorded (Figure 2a and Sections S8.3). Whilst AB1 was consumed within 15 seconds, guaiacol G and acetal-stilbene P1 were formed at a slower rate, reaching 79% and 56% yields respectively. Two major signals were also observed by UPLC-MS analysis (both with [M+H]+ = 465 g mol⁻¹) prior to the formation of G and P1 (Figures 2a and Section S10.1). These were attributed to the formation of the isomeric alkenes I1, the products of dehydration and deformylation of AB1. Whilst dehydration occurs by loss of the benzyl hydroxyl group in the β-O-4 unit, deformylation could occur in the β-O-4 unit as well as the β-5 unit in AB1. Compounds A⁵ and B⁶ were used to investigate this issue further.

**Study of the relative reactivity of β-O-4 and β-5 units in AB1**

In a reaction with 10 mol% HOTf β-O-4 model A yielded 87% G and 54% acetal P3 (Scheme 5a). Next, the reaction was monitored for 2 hours (Figure 2b, and Sections S8.1 and S10.2). This revealed that A was rapidly consumed and two main products were formed ([M+H]+ = 287 g mol⁻¹ by UPLC-MS). This reactivity pattern was analogous to that observed for AB1 and the detected mass of the products confirmed the formation of the isomeric enol ethers I3, formed by acid catalysed dehydration/deformylation of the β-O-4 moiety en route to the C2-aldehyde. I3 was further converted to G in 80% yield and P3 in 61% yield.
Scheme 5. Reactions with HOTf and ethylene glycol with a) β-O-4 model compound A b) β-5 model compound B

When no ethylene glycol was added, G was still obtained in good yield (69%), but the C2-aldehyde was not observed due to its conversion to a complex mixture of products, as seen for AB1 under these conditions (Section S9.0). During these reactions, ketal P4, the ethylene glycol ketal of the Hibbert ketone, was also identified (UPLC-MS, Section S10.2). Its formation provided evidence for the functioning of the C3 cleavage pathway in these reactions. Since this pathway also leads to the formation of guaiacol G, this explains the discrepancies between the yields of G and P3 from A (and analogously the differences between the yields of G and P1 formed from AB1 above). Dehydrated intermediate I4 (Figure 2b), the most likely precursor of P4, was previously observed when water was used as solvent but could not be detected under our reaction conditions.24c,24b

Next, the reactivity of the β-5 model B was investigated. Upon reaction of B with 10 mol% HOTf and 4 eq. ethylene glycol, E-stilbene P5 was obtained in 76% yield (Scheme 5b). However, the consumption of B was slow compared to A and AB1, and full conversion of B was only achieved after 30 minutes in contrast to 15 seconds for A and AB1 (Figure 2c, Section S8.2). The rates of formation of P5 corresponded to the rates of B consumption and no other reaction intermediates were identified. This is consistent with either the concerted deformylation/ring opening of B or the formation of short-lived intermediates en route to P5 (β-5 C2 pathway shown in Scheme 3). Dehydrated benzofuran P6 (Figure 2c), was identified as minor side product (UPLC-MS, Section S10.3). P6 originates from the C3-pathway previously identified on acid catalysed modification of the β-5 linkage (Scheme 3).28
Proposed reaction pathways in acidolysis of AB1

Returning to the reactivity of AB1 under acidolysis and acetal forming conditions, a series of reaction pathways were constructed (Scheme 6) and rate analysis provided the curve fits shown in the corresponding figures (on rate modelling see Section S8.0). The AB1 acidolysis products ([M+H]⁺ = 465 g mol⁻¹) were assigned to the E- and Z-isomers of enol ether I1, products of the reverse Prins reaction of AB1 in which the β-5 linkage remains unmodified. This is consistent with the very fast formation of I3 from A. The subsequent cleavage of I1 to form G and an elusive intermediate I1a (calculated rate of consumption I1 = 0.35 min⁻¹ vs I3 = 0.22 min⁻¹) is the subsequent step followed by the modification of the β-5 linkage via C2 pathway to give the final acetal stilbene product P1 (rate of formation = 0.14 min⁻¹ for both P1 and P5). The C3 pathway for the β-5 modification also occurs as a minor side reaction providing traces of P8 similar to the traces of P6 formed from B. The second existing route by which G is formed from AB1 is the C3 pathway analogous to that identified using the β-O-4 model compound A. This route leads to P7 ([M+H]⁺ = 403 g mol⁻¹), the corresponding Hibbert ketal analogue (Section S10.1). For the β-O-4 cleavage, the C2 pathway is dominant over the C3 pathway under these reaction conditions (a 3:1 ratio based on the modelled rates and the P1 to G yield discrepancy).

The ring opening of the β-5 linkage occurs nearly exclusively via the C2 pathway.

Determination and quantification of the formaldehyde released from the (β-O-4)-(β-5) models

During the acidolysis of models AB1, A and B, the C2 reaction pathways for the β-5 and β-O-4 linkages both involve the formal loss of a carbinol group. Although previous studies agree that this is achieved through the release of formaldehyde²⁴⁻²⁵, there has been little direct evidence to support this or attempts to quantify the amount of formaldehyde released, likely due to experimental difficulties. Our unique reaction conditions, however, allow for identification and quantification of the released formaldehyde trapped as its ethylene glycol acetal, 1,3-dioxolane Z (Scheme 7).

Scheme 7. 1,3-dioxolane Z formation from the reactions of A, B, or AB1 with HOTf and ethylene glycol.
Reactions of AB1, A and B were repeated in d7-1,4-dioxane. In all cases the corresponding 1,3-dioxolane Z was clearly identified (signals at δ 4.77 and δ 3.76 in 1H-NMR spectra) and the amounts of Z as well as acetal products P1 and P3 were quantified using an internal standard (Figure 3, for details see Section S12.1). Also, for the β-5 model B, the amount of Z (81% yield) was consistent with the corresponding C2 product, P5 (76% yield by HPLC from a separate reaction, Figure 3b and Section S12.2). Finally, for AB1 an 85% yield of Z based on the release of two equivalents of formaldehyde was found (Figure 3c and Section S12.3). The amount of P1 was slightly lower than expected based on the yield of Z (62% P1 vs 85% Z), but is consistent with the HPLC yields discussed above (Scheme 4) combined with the observation that the C3 pathway for the cleavage of the β-O-4 linkage still leads to a product in which the β-5 unit has been modified according to the C2 pathway leading to additional Z (Scheme 6). The observed quantities of Z, together with the identified products of the complementary C2 pathways, are strong indications that most of the released formaldehyde is trapped as its corresponding acetal. Formaldehyde has been previously implicated in condensation reactions\(^{46,59}\), thus the use of ethylene glycol in our catalytic system contributes to eliminating the adverse effects of formaldehyde. This, together with the trapping of other reactive intermediates (aldehydes) explains the success of this methodology when applied to lignin.\(^{46}\)

**Examination of the reactivity of β-β model compounds**

The effect of our standard acidolysis conditions on the β-β motif was studied using the model C1a (sesamin, Scheme 8) as C1a has the same relative configuration as the β-β linkage in lignin.\(^{5,9}\) Acidolysis of C1a led to a remarkably clean reaction (Section S13.1) with the main products being epimers C1b (asarinin/episesamin) and C1c (epiasararin/diasesamin, Scheme 8).\(^{9,31}\) The ratio of C1a : C1b : C1c was 1 : 1 : 0.1 (1H NMR, Figure S19) with a >95% mass balance (GC-FID) being observed. Reaction of C1a in the absence of ethylene glycol provided the same product mixture indicating little influence of the diol on this reaction (Figure S20). The same product distribution was also observed when C2a (yangambin) was reacted under these conditions (Figure S21). Epimerization reactions for similar compounds have been previously reported using different Lewis acids.\(^{39,40,41}\) Phenolic versions of these compounds (e.g. pinoresinol and syringaresinol C3, Figure 1) and their epimers were previously obtained during lignin acidolysis\(^{4,23,26} \) and were again identified in this work (vide infra). These results indicate no effect of ethylene glycol on the products formed via acidolysis of the β-β motif in lignin.

**Identification of dimeric products in lignin derived product mixtures**

This work culminated in our analysis of lignin derived product mixtures to assess if the reactions observed in the model compounds translated to the natural material itself. A typical organosolv lignin consists predominantly of the most abundant β-O-4 linkage and the less abundant (about 10%) β-5 and β-β linkages (other minor linkages were not considered).\(^{5}\) Therefore, it is very likely that the β-5 linkages will be flanked by β-O-4 linkages, a situation that inspired the design of AB1-4. The same will hold true for the β-β linkages. Exposure of lignin to our catalytic acidolysis conditions would therefore be expected to give phenolic acetals P9-11 as the major products via the C2-pathways as they result from the cleavage of neighbouring β-O-4 linkages (Scheme 9)\(^{44}\) as well as small amounts of Hibbert ketals via the C3 pathway. A β-5 dimer flanked by two β-O-4 linkages should result in stilbene compounds such as P2 via the C2 β-O-4 cleavage pathway plus smaller amounts of ketal structures such as P7 (Scheme 6) through the C3 β-O-4 cleavage pathway. A β-β dimer flanked by two β-O-4 linkages should give epimerized diphenolic β-β fragments like C3 (Scheme 9).\(^{23,24}\)
Scheme 9: Schematic showing of specific linkages as they would appear in lignin and expected cleavage products. A hypothetic lignin structure is shown containing β-O-4, β-5 and β-β linkages.

Figure 4: 2D HSQC NMR spectrum of walnut methanosolv lignin showing areas used for the quantification of visible linkages and determination of S : G : H ratios.
Table 1: Lignin characteristics determined by GPC and 2D-HSQC analysis.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Lignin</th>
<th>$M_n$(Da), $M_w$(Da), D</th>
<th>S, G, H (%)</th>
<th>Linkages (per 100 C₅ units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pine methanosolv</td>
<td>1075, 2088, 1.9</td>
<td>0, 100, trace</td>
<td>β-O-4</td>
</tr>
<tr>
<td>2</td>
<td>Beech ethanosolv</td>
<td>928, 2016, 2.2</td>
<td>68, 32, 0</td>
<td>φ-O-4-OR</td>
</tr>
<tr>
<td>3</td>
<td>Walnut methanosolv</td>
<td>808, 1518, 2.2</td>
<td>65, 29, 6</td>
<td>β-5</td>
</tr>
</tbody>
</table>

a determined by GPC (THF) against polystyrene standards (Section S15.1). b Determined by 2D-HSQC using signal intensities of the corresponding aromatic signals corrected for the amount of protons (Section S15.2). c Determined by 2D-HSQC by comparing the signal intensities of the aromatic signals to the intensities of the benzylic protons of the linkages corrected for the amount of protons (Section S15.2). d Total number of β-O-4 linkages e amount of α-methoxylated/ethoxylated units.

Table 2: Product distribution P9-P11 (Shown in Scheme 9) obtained from lignin acidolysis reaction using HOTf and in the presence of ethylene glycol.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Lignin</th>
<th>P9 b</th>
<th>P10 b</th>
<th>P11 b</th>
<th>Total P9-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pine methanosolv</td>
<td>-</td>
<td>4.4 Wt%</td>
<td>0.1 Wt%</td>
<td>4.5 Wt%</td>
</tr>
<tr>
<td>2</td>
<td>Beech ethanosolv</td>
<td>4.8 Wt%</td>
<td>2.6 Wt%</td>
<td>-</td>
<td>7.4 Wt%</td>
</tr>
<tr>
<td>3</td>
<td>Walnut methanosolv</td>
<td>7.0 Wt%</td>
<td>3.9 Wt%</td>
<td>0.4 Wt%</td>
<td>11.3 Wt%</td>
</tr>
</tbody>
</table>

a Conditions: 50 mg lignin, 60 mg ethylene glycol, 7.5 Wt% HOTf, 1 mL 1,4-dioxane, 30 min., 140 °C, in sealed pressure vessel, n-octadecane as GC internal standard. Low MW fraction was obtained by extraction of dried reaction solid with 9:1 toluene : DCM.

b Determined by GC-FID referring to the starting lignin.

In order to confirm this, catalytic depolymerization reactions were carried out using pine, beech and walnut shell organosolv lignins. These lignins were obtained by standard organosolv processing and characterized using 2D HSQC NMR (Figure 4) and GPC for which the most relevant data are summarized in Table 1 (Isolation and characterization details in S14.0 and S15.0).

Next, 50 mg samples of these lignins were subjected to the catalytic acidolysis conditions. The crude reaction mixtures were processed by extraction to obtain low MW and high MW fractions (Sections S16.0 and S17.0). The low MW fractions were analysed by GC-FID and GC-MS and the expected main product acetics (P9-11, Scheme 9) were quantified using an internal standard (Table 2). The P9 vs P10 ratios corresponded well to the amount of S and G units in the lignin starting material. Moreover, the total acetal yields for the respective lignins were dependent on the number of β-O-4 linkages in the original lignin (compare Tables 1 and 2). In the case of ethanosolv beech lignin, the β-O-4 moiety showed increased ethanol incorporation as a result of the organosolv procedure.21 This is a likely explanation of the slightly higher than expected acetal yields based on the overall β-O-4 content determined by NMR. All acetal yields corresponded well to the isolated yields we have previously reported (Section S17.0 for analysis details).24 In these reactions small amounts of products were also seen that correlate to cleavage of the β-O-4 moiety via the C3-pathway, including P12 (Figure 5a).

The product mixtures from pine lignin were investigated first. Gratifyingly, acetal stilbene P2 could be identified by GC-MS analysis and its presence verified by spiking with an authentic sample of P2 (Figure 5a and Table S8). The yield of P2 was determined as 2 wt%, in agreement with the relatively high percentage of β-5 linkages (10 per 100 aromatic units) in this lignin. Since pine lignin contains only G units, none of the corresponding S containing acetal stilbenes were observed. Compound P13 (analogous to P8) was also detected (Figure 5a). No β-β dimer fragments were identified in this reaction mixture given the limited amount of such linkages present in this lignin (<1 β-β linkages per 100 aromatic units, Table 1).

The beech organosolv and the walnut shell methanosolv lignins were richer in β-β linkages (4 and 8 β-β linkages per 100 aromatic units respectively) thus β-β-containing fragments derived from these lignins were successfully identified. The presence of syringaresinol C3a was verified by spiking with an authentic sample for both lignins (Figure 5b and Figure S29). C3a and epimer C3b were found as a 1:1 mixture and identified based on their identical MW and fragmentation patterns. The observation of C3, a β-β dimer of two S units, is consistent with the relatively high amount of S units in these lignins. In addition, it is known that S units are more likely to undergo β-β dimer formation during lignin biosynthesis.34 The combined yields of these epimers from beech and walnut lignin was 2.6 wt% and 5.5 wt% respectively, which is in line with the amount of the respective linkages in these lignins (GC-MS analysis see Tables S9 and S10). Additionally, in the samples obtained from the walnut methanosolv lignin, trace quantities of P2 and P13 were observed.
Figure 5: GC-MS traces of product mixtures obtained from the depolymerization of a) methanosolv pine lignin and the same sample spiked with an authentic sample of compound P2 and b) beech wood ethanosolv lignin and the same sample spiked with an authentic sample of compound C3a. Reaction conditions: 50 mg lignin, 60 mg ethylene glycol, 7.5 Wt% HOTf, 1 mL 1,4-dioxane, 30 min., 140 °C, in sealed pressure vessel, n-octadecane as GC internal standard (For more detailed analysis of the GC-MS trace see Section S17.2).

Figure 6: 1H-NMR spectrum of the crude reaction mixture obtained from the depolymerisation of 50 mg walnut methanosolv lignin demonstrating the formation of 1,3-dioxolane Z. Reaction conditions: 7.5 wt% HOTf, 60 mg ethylene glycol at 140 °C for 30 minutes, quenched by the addition of 5 µL Et3N.

The above results clearly demonstrate that the chemistry established using the (β-O-4)-(β-5) model compounds AB1-4, as well as the β-β model compounds C1 and C2 using acetal formation conditions can be directly extrapolated to the depolymerization of lignin under the same conditions. The unambiguous identification of structurally diverse dimeric compounds such as P2 or C3 in complex lignin derived product mixtures would prove extremely challenging solely based on GC-MS or UPLC-MS analysis. With lignin-relevant model compounds such as AB1-4, however, the formation of these compounds can be rationalized. Analysis of the product mixtures also confirmed the dominance of the C2 reaction pathways, which should coincide with formaldehyde release from the β-O-4 and β-5 motifs. A separate set of experiments was conducted to confirm this using beech ethanosolv and walnut methanosolv lignin in d8-1,4-dioxane. The 1H-NMR analysis of these reactions revealed the formation of 1,3-dioxolane Z (Figure 6 and Section S18.0). The yields of Z from beech and walnut lignin were 1 wt% and 4.2 wt% respectively (quantified using an internal standard). This corresponds to the amounts of acetals P9-P11 detected. It is remarkable, that the reactivity trends established using our new models AB1-4 were also in good agreement in terms of formaldehyde release with the results obtained with actual lignin samples. A further advantage of trapping the released formaldehyde is that it leads to a more complete overall carbon mass balance of the lignin depol-
ymerization reaction. The high yield of 1,3-dioxolane \(Z\) bodes well for the large-scale production of this compound from lignin, in addition to the valuable aromatics, since 1,3-dioxolane \(Z\) already finds use as a solvent.

CONCLUSION

We have described the synthesis of a new class of \((\beta-O-4)-(\beta-5)\) lignin models \(\text{Ab1-4}\) that are realistic representations of an abundant lignin fragment (particularly in softwoods). These models allowed for in-depth catalysis studies and enabled a detailed understanding of the controlled catalytic depolymerization of lignin itself. This was possible as \(\text{Ab1-4}\) are sufficiently complex to mimic lignin reactivity but still enable product analysis. We also gained detailed insight into the acid catalyzed cleavage of \(\text{Ab1-4}\) as well as other \(\beta-O-4, \beta-5\) and \(\beta-\beta\) model compounds. It was demonstrated that the mild depolymerization strategies presented herein were highly efficient in the cleavage of C-O bonds, while the main C-C linkages in the \(\beta-5\) and \(\beta-\beta\) were left intact, the only C-C bond scission being the release of formaldehyde. Therefore, in order to obtain high yields of aromatic monomers, lignins with high \(\beta-O-4\) content are desired. The structure and quantity of dimeric products on the other hand relates to the type and number of C-C bonds present in the starting lignin structure. Major reaction pathways (C2 and C3, Scheme 3 and Scheme 6) and important intermediates were identified. In addition, novel dimeric products, such as the E-acetel stilbenes \(P_1\) and \(P_2\) were isolated. This has, for the first time, allowed the identification of these products in depolymerization mixtures generated from pine and walnut lignins.

Recently, Sels and coworkers found the use of ethylene glycol beneficial in reductive lignin depolymerization.\(^{38}\) Our previous studies also addressed the advantages of using ethylene glycol under acidolysis conditions.\(^{44}\) Herein, we further specified the benefits of using ethylene glycol, in our reactions. Firstly, ethylene glycol stabilizes the various C2-aldehydes formed on cleavage of the \(\beta-O-4\) linkages. Further, ethylene glycol plays a role in “trapping” the formaldehyde released both from the \(\beta-O-4\) as well as the \(\beta-\beta\) linkage. Importantly, we were able to quantify the amount of released formaldehyde in model and lignin reactions via the corresponding 1,3-dioxolane \(Z\) formed.

Overall, a close correlation between the reactivity of \(\text{Ab1-4}\) and lignin was found. Thus, our novel \((\beta-O-4)-(\beta-5)\) lignin models should find general use in future catalytic lignin depolymerization studies and will enable further improvements in our understanding of the reactivity of lignin. This is an essential component of establishing financially viable biorefineries.

ASSOCIATED CONTENT

Supporting Information. Synthetic procedures and analytical data for described model compounds; crystallographic data for compounds 4 and 5; procedures and analytical data for reactions with model compounds as well as product isolation and characterization; lignin isolation procedures and characterization; lignin depolymerization procedures; analytical data for product mixtures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Authors
* k.barta@rug.nl
* njw2@st-andrews.ac.uk

Author Contributions
‡ These authors contributed equally.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by the EP/J01819/1, EP/K00445X/1 grants (NJW and PCJK), an EPSRC Doctoral Prize Fellowship (CSL), and the European Union (Marie Curie ITN ‘SubBiCat’ PITN-GA-2013-607044, CWL, NJW, PCJK, PJD, KB, JdeV). We acknowledge the EPSRC UK National Mass Spectrometry Facility at Swansea University for mass spectrometry analysis.

REFERENCES

changing the base to Cs, CO₂, the ring-opened product Si was produced in 98% yield (Table S1).


Graphic entry for the Table of Contents (TOC)