Green Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Green Chem.*, 2022, **24**, 7000

Unravelling stereoisomerism in acid catalysed lignin conversion: an integration of experimental trends and theoretical evaluations†

Zhenlei Zhang, [a] Susanna Monti, [b] Giovanni Barcaro, *[c] Ciaran W. Lahive [a] and Peter J. Deuss *[a]

For the effective valorization of lignin, which is a significant component in agricultural residues, its reactivity needs to be understood in detail. Selective acid-catalysed depolymerisation of the lignin β-O-4 linking motif with stabilization of the formed aldehydes by diols is a promising approach to obtain phenolic monomers in high yields. However, the lignin β-O-4 linking motif exists in both the *erythro* and *threo* isomeric forms, and very little information is available on the influence of stereochemistry on the efficiency of the lignin diol-stabilised acidolysis. This is especially true for the set of intermediates in which the presence of stereochemistry persists. In this study, the stereoisomer ratios of two key intermediates, namely the diol (here ethylene glycol) adducts and C2-vinyl ethers, are monitored carefully in ytterbium(III) trifluoromethanesulfonate [Yb(OTf)$_3$]-catalysed conversion of an *erythro* β-O-4 model compound. The reactions showed the preferential formation and consumption of the ethylene glycol adduct in the *erythro* configuration, and the favored formation of *trans* C2-vinyl ether. Multiscale computational methods (including classical reactive molecular dynamics simulations and quantum chemistry calculations) were applied to elucidate the catalytic origins of the observed stereo-preferences and suggested that a proto-*trans* intermediate complex is stabilised by a hydrogen bond network connecting the carbocation, ethylene glycol, and the anionic [OTf]– species. The synergistic combination of experiments and computational studies disclosed the stereo-preference and the underlying mechanism in triflate-catalysed acidolysis, especially the catalytic role of [OTf]–, which can be helpful for a further improvement of the chemical process.

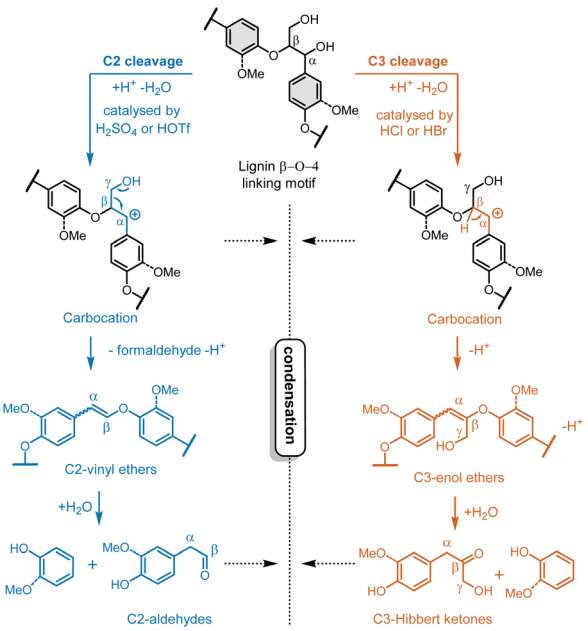
Received 9th May 2022,
Accepted 28th July 2022

DOI: 10.1039/d2gc01740g

rsc.li/greenchem

Introduction

Lignin, which is mainly burned for energy, is increasingly recognised as a potential sustainable source for aromatic chemicals.[1–4] The cleavage of the most abundant β-O-4 linking motif to aromatic monomers can readily be achieved by acidolysis or acid-catalysed degradation with various mineral acids, which is a methodology that has been traditionally used for the structural elucidation of lignin.[5–8] Two β-O-4 cleavage pathways have been recognised (*i.e.*, C3 and C2 cleavage pathways), which differ by the loss of the γ-carbinol group. The pathways are dictated primarily by the type of acid applied (Scheme 1).[9] The C3 cleavage pathway of the β-O-4 linkage, for example,



Scheme 1 The C2 and C3 acidolysis pathways go through the same benzylic carbocation intermediate but yield different subsequent vinyl ether intermediates and final cleavage products.

[a]*Department of Chemical Engineering (ENTEG), University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands. E-mail: p.j.deuss@rug.nl, giovanni.barcaro@cnr.it*
[b]*CNR-ICCOM – Institute of Chemistry of Organometallic Compounds, 56124 Pisa, Italy*
[c]*CNR-IPCF – Institute for Chemical and Physical Processes, 56124 Pisa, Italy*
†Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d2gc01740g

catalysed by HCl or HBr, yields a mixture of C3-ketones known as Hibbert ketones,[10–14] while C2-aldehydes and formaldehyde are generated through the C2 cleavage pathway catalysed, for example, by H_2SO_4 or HOTf. This is a distinction that is poorly understood.[15,16] The obtained cleavage products such as ketones and aldehydes as well as the carbocation intermediate are prone to undergo condensation, thus reducing the final monomer yields.[17,18] The addition of alcohol,[19] or even better a diol (in this study, ethylene glycol (EG)),[20] is shown to be an effective solution to extensively prevent carbon balance losses caused by intermediate and product condensation. Under the acidolysis conditions, the carbocation can be stabilised, and reactive aldehydes can be trapped *in situ* by EG (Scheme 2).[21,22] Recent kinetic studies using $Yb(OTf)_3$ on β-O-4 model compounds allowed us to establish a detailed reaction network incorporating two key reaction intermediates: EG adducts and C2 vinyl ethers. In the identified reaction network, the vinyl ether is found to be generated from both EG adducts and β-O-4 linking motif, although the latter route has higher activation energy.[23] $Yb(OTf)_3$ as well as other metal triflates (*e.g.*, $Zn(OTf)_2$, $Al(OTf)_3$, $Cu(OTf)_2$), was much easier and less hazardous to handle compared to fuming and corrosive superacid HOTf (which delivers much faster reaction rate due to its super acidity). Several advantages[24] of using metal triflates, instead of triflic acid, have also been summarized: (1) better solubility of metal triflates, thus much more precise loading control since solubilising triflic acid in an apolar solvent cannot be readily achieved, especially in small-scale experiments.[25] (2) Better catalytic performance of metal triflates than triflic acid due to a significant decrease of side reactions or product decomposition caused by direct usage of triflic acid.[26] With metal triflates small amounts of HOTf are formed *in situ* to catalyse the reaction,[24,27] therefore, these triflates-catalysed lignin acidolysis actually follow the same reaction paths (*i.e.*, C2 cleavage pathway).[23,28] Cations with different atom radium have different hydrolysis (as well as ethylene glycolysis) constants.[29] This can result in the different formation rate of HOTf,[24] which affects the final *e.g.*, acidolysis reaction rate.[23] The metal triflate-catalysed diol-stabilised acidolysis methodology has been extended to organosolv pine lignin using $Fe(OTf)_3$, where it allowed for up to 35.5 wt% of C2-acetal products.[30] Moreover, related lignin-first mild fractionation strategies ccatalysed by H_2SO_4 (which also catalyses lignin acidolysis in C2 pathway) were developed accordingly, yielding stabilised lignin in which ethylene glycol or other types of diols were incorporated.[22,31–33]

The β-O-4 linking motif has two chiral carbon centers (at the α and β carbons, respectively), thus allowing for four stereoisomers (shown in Fig. 1). These four stereoisomers can be categorised into *erythro* and *threo* configurations (the four stereoisomers and their associated nomenclatures used in literature are summarized in Fig. 1).[34] Their relevant amounts

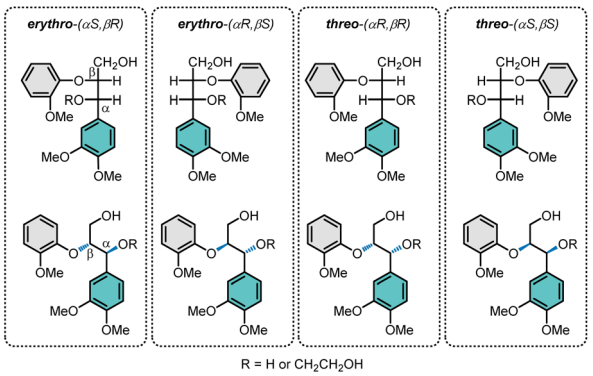
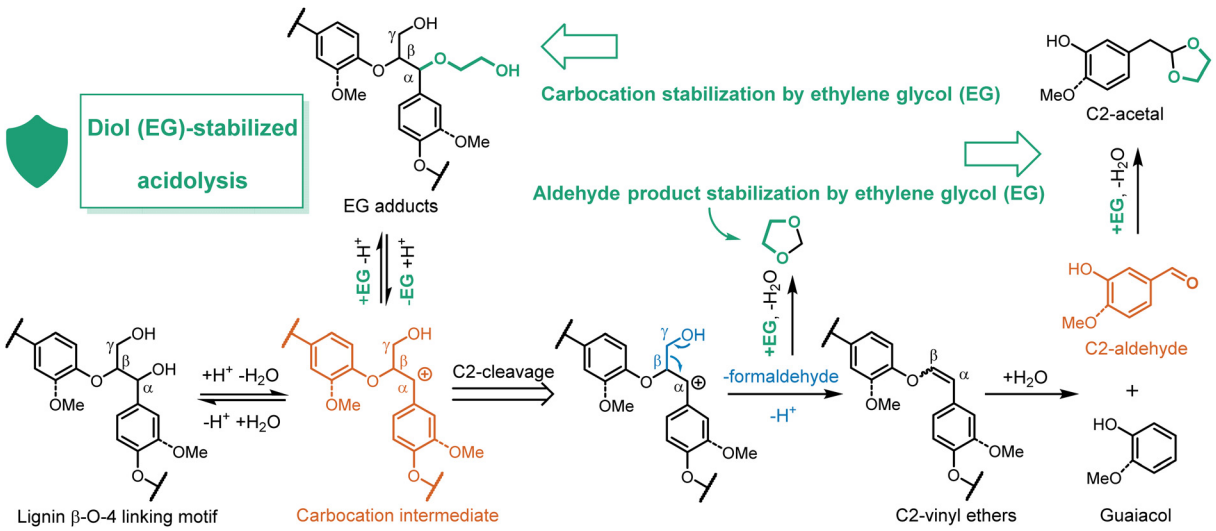


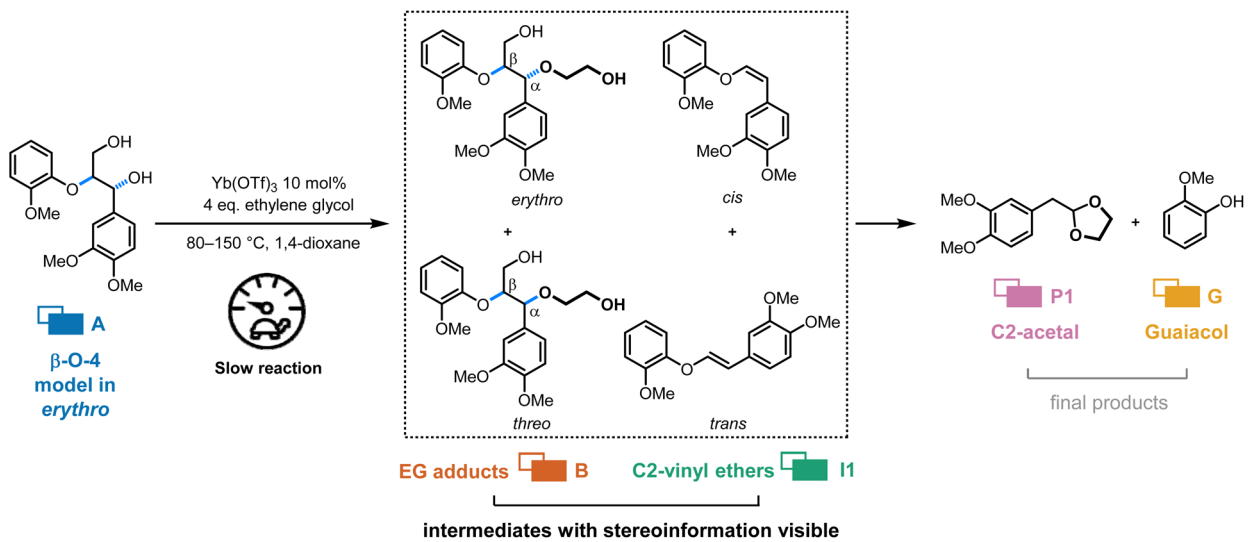
Fig. 1 Fischer projections of *erythro* and *threo* C3 β-O-4 model compounds and their corresponding absolute configuration. Sometimes *erythro* and *threo* isomers are referred as *anti* and *syn* isomers, respectively.



Scheme 2 The C2 cleavage pathway of the β-O-4 linkage motif under the diol-stabilised $Yb(OTf)_3$-catalysed conditions. In this study, ethylene glycol is used.

differ in softwood (nearly equal distributions of β-O-4 linking motif in the *erythro* and *threo* configurations) and hardwood (β-O-4 linking motif in *erythro* configuration prevails).[35–37] It has been shown that *erythro* and *threo* configurations of the β-O-4 linking motif can dramatically influence the spatial geometry conformation of lignin chains, thus leading to differences in reactivity upon depolymerisation.[38–42] For example, under the alkaline pulping conditions, the β-O-4 model in the *erythro* form was more reactive than its *threo* isomer (*i.e.*, 2–8 times faster depending on the aromatic substitution pattern).[43,44] As a result, at alkaline pulping conditions, the *erythro*-enriched hardwood exhibited a higher delignification efficiency than softwood. In addition, the oxidative degradation of lignin showed a stereo-preference: β-O-4 models with *threo* configuration were more reactive under specific reaction conditions, although catalyst types and reaction conditions (*e.g.*, pH) could result in alternating preference or lack of preference.[45–49] Therefore, understanding the relevant differences in lignin reactivity due to the occurrence of different isomeric structures will contribute to more efficient and sustainable process design, *e.g.*, by choosing a reasonable lignin source with desired structures. In addition, it could provide valuable information such as biotechnologically modifications of lignin to reduce lignin recalcitrance and promote biobased chemicals production from lignin.[50]

Regarding diol-stabilised acidolysis, the isomeric ratios for the two key intermediates (*erythro* and *threo* for the EG adduct, *trans* and *cis* for the C2-vinyl ether) can potentially have significant influences on reaction pathways and rates. As far as we know, no data is currently available on the stereo-preference and its origin in the conversion of β-O-4 and intermediate under acidic conditions in the presence of alcohols, while such conditions are also relevant to typical organosolv extraction. This is likely due to the fast reaction rates and the relative instability of intermediates and products, which makes detailed characterization of them difficult. In this study, the acidolysis catalyst ($Yb(OTf)_3$), one of the several triflate salts that are effective for acidolysis in various studies,[23,28,51,52] was used to convert the β-O-4 motif in the *erythro* configuration **A**. This metal triflate can catalyse the diol-stabilised acidolysis of β-O-4 motif slowly without changing reaction pathways (compared with HOTf as catalyst).[23] Therefore, it is possible to obtain quantitative profiles of the separate stereoisomers of intermediates, namely, **B** (the ethylene glycol adducts) in *erythro* and *threo* configurations, and **I1** (the C2-vinyl ethers) in *trans* and *cis* isomeric forms, to elucidate the stereo-preferences during the triflate-catalysed acidolysis of lignin β-O-4 motif (Scheme 3). The mechanism for stereo-preference is elucidated and validated by computational simulations. By using a multiscale/level protocol, which has been successfully applied to investigate the acidolysis mechanisms of lignin model compounds in very similar conditions,[20] we investigate crucial intermediate complexes that account for the stereo-preferred conversion of the EG adducts to the C2-vinyl ethers, as well as the catalytic role of the triflate anion to explain the experimental observations. It is worth to mention that, in contrast to $Yb(OTf)_3$ in 1,4-dioxane used here, for the final application of diol-stabilised lignin acidolysis the aim is to use cheaper triflates *e.g.*, $Fe(OTf)_3$ and even other types of acids *e.g.* H_2SO_4 that can achieve the same reaction pathway with much faster reaction rate.[30,31] The diol-stabilised acidolysis can be performed with H_2SO_4 in dimethyl carbonate as a much greener solvent (Fig. S1, ESI†). However, here the isomers of intermediates are more challenging to monitor due to extremely fast reaction rates and interfering minor side-reactions. Therefore, in this study, $Yb(OTf)_3$ in 1,4-dioxane was used to get insights of the reaction profiles for mechanistic studies (especially the C2 acidolysis pathway).



Scheme 3 Primary identified intermediates and products from the C2 cleavage pathway of the β-O-4 model under the diol-stabilised $Yb(OTf)_3$-catalysed condition.

Results and discussion

C3 β-O-4 model acidolysis catalysed by triflates under various reaction conditions with quantification of intermediates under different configurations and conformations

As already mentioned above, using $Yb(OTf)_3$ in 1,4-dioxane we could confidently track the evolution of the intermediate species, *viz*. EG adducts **B** and C2-vinyl ethers **I1** (see Scheme 2) in the diol-stabilised acidolysis of the C3 β-O-4 model. We identified two stereoisomers for both **B** and **I1**, respectively, in the HPLC chromatogram (Fig. S2†). The configurations of the two stereoisomers of **B**, were assigned by applying preparative HPLC to separate the isomers, then followed by NMR identification based on coupling constants from the literature ($J_{\alpha,\beta}$ = 7.7 Hz *threo* and $J_{\alpha,\beta}$ = 7.1 Hz *erythro*, ESI, S2.0†).[53–56] For **I1**, a mixture of isomers enriched in the *cis* conformation was synthesised following the procedure reported by Bolm *et al.*[57] using a C2 lignin β-O-4 model and assigned *via* their representative coupling constants (*trans* **I1** $J_{\alpha,\beta}$ 12.5 Hz; *cis* **I1** $J_{\alpha,\beta}$ = 6.8 Hz, details see ESI S2.0†). Based on the stereochemistry data of **B** and **I1**, we calibrated the different isomers of **B** and **I1** separately to monitor these isomeric mixtures during the cleavage reaction (Fig. 2 shows a time course profile at 120 °C). At 120 °C, the C3 β-O-4 model **A** (in the *erythro* configuration) was rapidly consumed within 30 min, while yields of the cleavage products guaiacol **G** and C2-acetal **P1** were continuously increasing to yields of 79% and 54% within 480 min, respectively. The observed yield gap between **G** and **P1** has been discussed in our previous work and, in short, is due to the competing C3 cleavage pathway and condensation reactions, which can be controlled by changing reaction conditions.[23] For example, the addition of higher loading of ethylene glycol can improve the carbon balance by increasingly stabilizing the formed reactive carbocation and aldehydes, albeit at the cost of a lower reaction rate. Water content is also important and high amounts are detrimental to the selectivity of **P1**, likely by promoting the formation C2-aldehyde which tends to undergo the condensation reactions.[21,23] Similarly, the reaction condition differences affect the formation and conversion of intermediates (*vide infra*). Focusing on **B** (in *threo* and *erythro* configuration) and **I1** (in *trans* and *cis* conformation) in their separate quantification, it was found that the **B** in the *threo* configuration had a peak yield of approximately 22% within 20 min, while the *erythro* isomer peaked at approximately 31% (Fig. 2). Regarding **I1**, the *trans* isomer achieved the highest yield of 35% within 90 min, while 24% was the highest yield for the *cis* isomer.

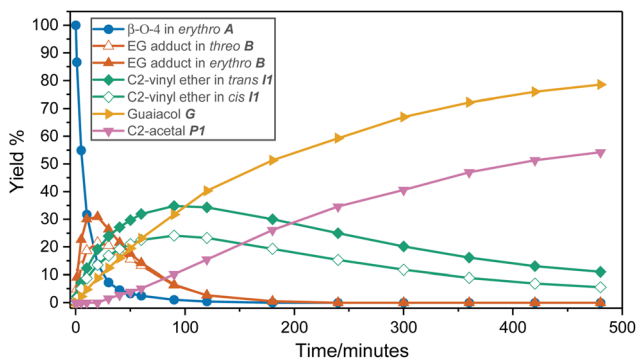


Fig. 2 Evolution of the reaction of the non-phenolic model compound **A** with 10 mol% $Yb(OTf)_3$, 4 eq. ethylene glycol, in 1,4-dioxane at 120 °C. Yields were obtained *via* HPLC analysis using 1,2,4,5-tetramethylbenzene as the internal standard (see Scheme 3 for the corresponding chemical structures).

The *threo*/*erythro* ratio of **B** and the *trans*/*cis* ratio of **I1** were tracked over time during reactions at different temperatures ranging from 80 to 150 °C (Fig. 3 & 4). This revealed some interesting phenomena pointing toward stereospecificity in the lignin β-O-4 acid-catalysed cleavage. At 80 °C, **B** *threo*/*erythro* ratio kept stable between 0.5 to 0.6 (Fig. 3a), and at this temperature, **B** is only formed without conversion. Interestingly when **B** started being converted (*i.e.*, at temperatures from 90 to 150 °C), **B** *threo*/*erythro* ratio steadily increased from approximately 0.5 to 1 or higher (Fig. 3a). The trend in the **B** *threo*/*erythro* ratio change is independent of the reaction temperature as clearly demonstrated by plotting the increase of the **B** *threo*/*erythro* ratio as a function of the **A** conversion percentage (Fig. 3b). This may indicate that the transformation of **B** is kinetically controlled. The plot of the *threo*/*erythro* ratio change *versus* the **B** yield describes the ratio changes during **B**'s formation-dominant and conversion-dominant stages, respectively (Fig. 3c). Before reaching the highest yield of **B** (formation-dominant stage), the ratio increased from approximately 0.5 to 0.68 (*threo*/*erythro*). Afterward, the ratio showed a rapid increase to the range between 0.9 and 1.3. This observation suggests that *erythro* **B** is not only preferably formed but also faster consumed compared to **B** in *threo* configuration.

A stereo-preference was also observed for the formation and conversion of **I1**, respectively. As shown in Fig. 4a, the **I1** in the *trans* conformation was formed preferably, with the *trans*/*cis* ratio starting at 1.3 and increasing as a function of time at all temperatures in the range of 90–150 °C. Since the **I1** can be derived from **B**, and *cis*-**I1** was found to be thermodynamically more stable, *vide infra*, the higher amount of *trans* **I1** indicates *trans* **I1** is the kinetically favored product. The *trans* **I1** was formed in a slightly higher amount than the *cis* **I1** in the data by Yokoyama *et al.*,[15] in the acidolysis of a C3 β-O-4 motif using 0.2 N HCl, HBr, and H_2SO_4 in 82% aqueous (no EG) at 85 °C, although the maximum yield was <1%. In their study, no **B** was obtained (as no diol was used in their system), indicating that a similar *trans*-preference can be present for the reaction of **A** to **I1** as well. In addition, the preferential formation of *trans* **I1** was also reported by Bolm *et al.*[58] in the base-catalysed depolymerisation of the C3 β-O-4 motif, although without a detailed explanation. Interestingly, the *trans*/*cis* ratios were well distributed around 1.3 before the conversion of **A** got close to 100% (Fig. 4b). Afterward, the ratio was remarkably increased, suggesting a much higher conversion rate of the *cis* isomer than the *trans* isomer. This can also

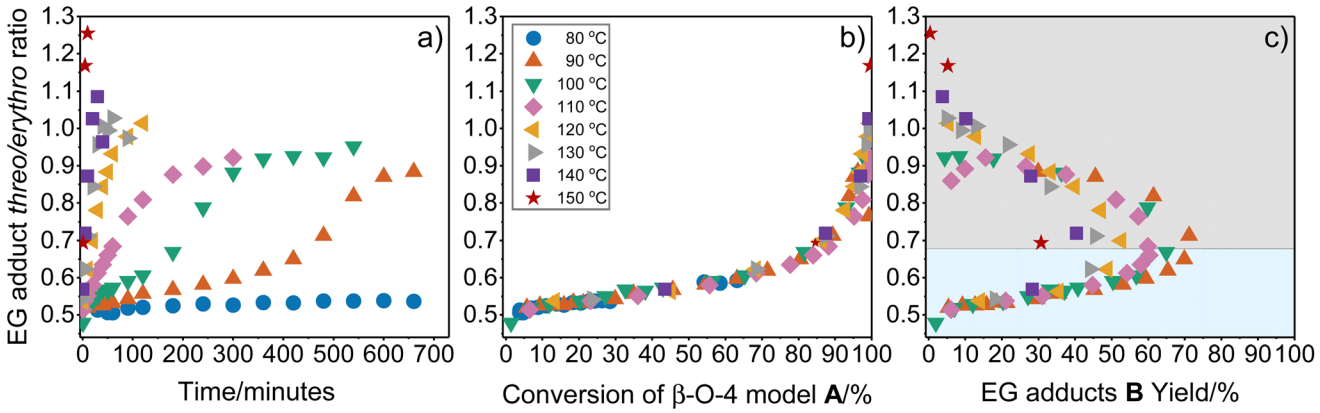


Fig. 3 EG adduct **B** *threo*/*erythro* ratio changes as a function of (a) time, (b) conversion of the β-O-4 model compound **A**, (c) EG adducts **B** yield, showing in light blue for the formation-dominant stage and in gray for the conversion-dominant stage of **B**. Reaction condition: 10 mol% $Yb(OTf)_3$, 4 eq. ethylene glycol, in 1,4-dioxane at 80–150 °C. Yields were obtained *via* HPLC analysis using 1,2,4,5-tetramethylbenzene as the internal standard.

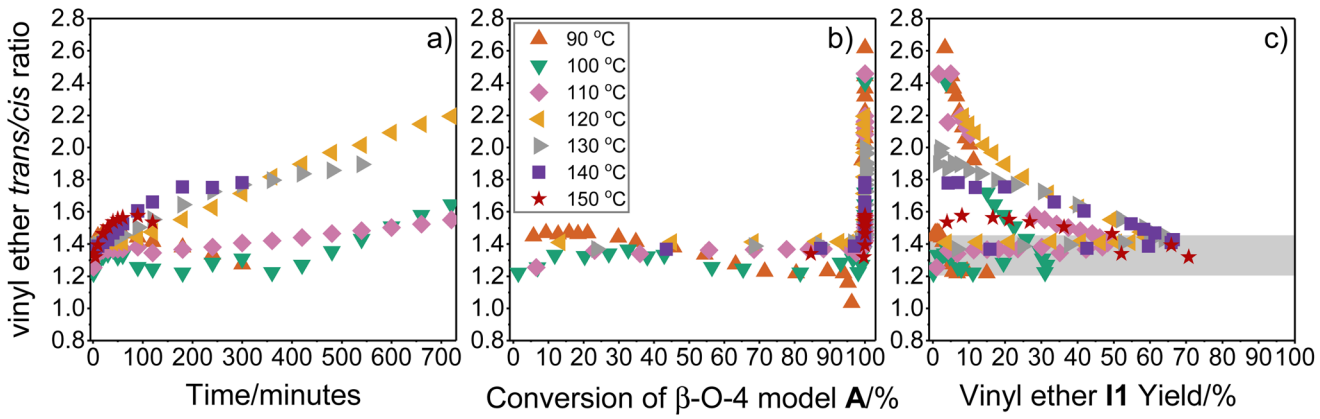


Fig. 4 The C2-vinyl ether **I1** *trans*/*cis* ratio changes as a function of (a) time, (b) conversion of the β-O-4 model compound A, (c) C2-vinyl ethers **I1** yield, showing in gray the formation-dominant stage **I1**. Reaction condition: 10 mol% $Yb(OTf)_3$, 4 eq. ethylene glycol, in 1,4-dioxane at 90–150 °C. Yields were obtained *via* HPLC analysis using 1,2,4,5-tetramethylbenzene as an internal standard.

be reflected by plotting the ratio change *versus* the **I1** yield (Fig. 4c). At the accumulating stage of **I1**, the *trans*/*cis* ratio increased only slightly from 1.2 to 1.4, but a dramatic increase up to 2.6 was observed at its conversion stage. This can readily be explained by the expected higher reactivity of the *cis*-isomer as it has a more accessible C=C bond for ethylene glycol or water addition.[17,21] In order to see whether the above stereo-preference extends to reactions catalysed by other metal triflate salts, $Zn(OTf)_2$ was used as the acidolysis catalyst. As shown in Fig. S3,† very similar stereo-preferences for both **B** and **I1** were observed. In addition, phenolic β-O-4 linking motif, representing an end-group in lignin, showed clear stereo-preferences under the diol-stabilised acidolysis conditions (Fig. S4†).

The reaction rate can be manipulated by increasing the catalyst and lowering the EG loading. However, lowering the EG loading also has a negative effect on product selectivity due to the increased formation of condensation products.[23] Furthermore, water is unavoidably involved in the reaction, either introduced by the biomass itself or during the dehydration step in the reaction (see Scheme 2). A high water content in acidolysis reactions negatively affects the yield of the final products by promoting side cleavage reactions of **I1** to the reactive C2 aldehyde, as well as reverting the acetal **P1** back to the same aldehyde under aqueous acidic conditions.[23] The **B** *threo*/*erythro* ratio was investigated under different reaction conditions (*i.e.*, different EG loading, catalyst loading, and water content). In terms of the effect of EG loading, mainly differences in changing rate of *threo*/*erythro* were observed, besides a higher peak ratio of *threo*/*erythro* was observed at lower EG loading (Fig. 5a). This could be due to a change in the thermodynamic stability of the isomers induced by the solvent environment or changes due to the kinetics of the equilibrium between **A** and **B** in relation to the conversion of **B** to **I1** and side products. The increase of catalyst loading from 2.5 to 20 mol% had nearly no effect on the peak *threo*/*erythro* ratio, though it accelerated the ratio changing rate following

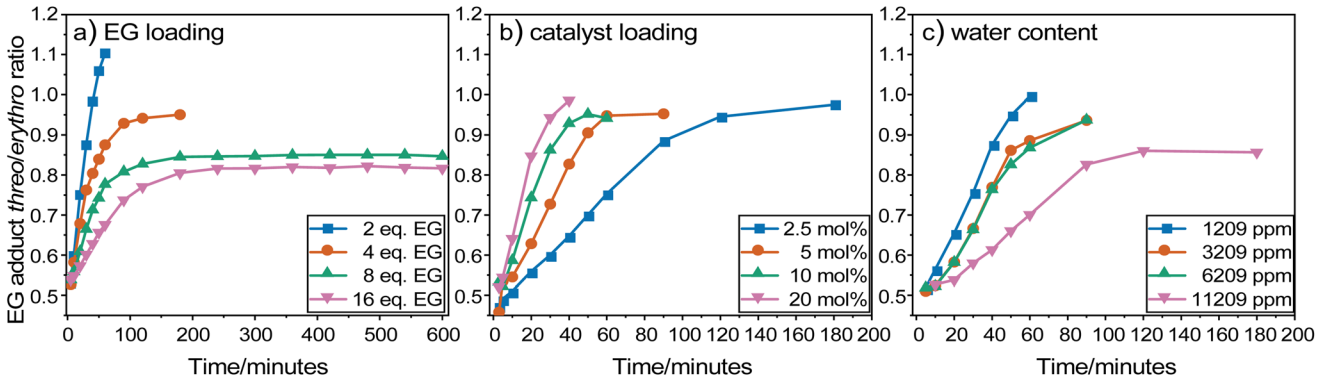


Fig. 5 The change of *threo*/*erythro* ratio of EG adducts **B** along time under different (a) EG loading, (b) catalyst loading, (c) water content. Reaction condition with different EG loading: 10 mol% $Yb(OTf)_3$, in dry 1,4-dioxane, at 130 °C. Reaction condition with different catalyst loading: 4 eq. EG, in dry 1,4-dioxane, at 130 °C. Reaction condition with different water content: 10 mol% $Yb(OTf)_3$, 4 eq. EG, in dry, at 130 °C. Yields were obtained *via* HPLC analysis using 1,2,4,5-tetramethylbenzene as the internal standard.

the increased reaction rate (Fig. 5b). It was found that an increase in water decreased the acidolysis rate by slowing down the carbocation formation rate.[23] Thus, the formation and conversion rates of **B** were decreased resulting in a slower change rate of the *threo*/*erythro* ratio at higher water content (Fig. 5c). It is apparent that the *threo*/*erythro* ratio starts from approximately 0.5 then increases over time despite the change of EG, catalyst, and water loading, showing that *erythro* **B** remains the preferred product under all these conditions. The **I1** *trans*/*cis* ratio fluctuated more under different conditions (Fig. 6). Interestingly, the **I1** *trans*/*cis* ratio started >1 under 2, 4, and 8 eq. of EG, while the initial ratio was <1 at 16 eq. of EG (Fig. 6a). This suggests a slight alternation of the preference of the formation of **I1** based on the EG content. The difference in catalyst loading resulted in no alternation and still gave *trans*/*cis* ratio >1 (Fig. 6b). Similar to the effect of increasing EG content, the higher water content in reaction made the preferential formation of *trans* **I1** less obvious. For example, the initial *trans*/*cis* ratio was approximately 1 at a very high water content of 11 209 ppm (Fig. 6c). This can be an effect of changes to the solvent environment affecting the relative thermodynamic stability of the isomers. Alternatively, reaction kinetics are affected and the *trans*–*cis* isomerization reaction of **I1** is relatively promoted at high water and ethylene glycol levels compared to other reaction steps. The selectivity of the reaction is known to be significantly affected due to alternative reaction pathways as previously shown.[23] Indeed, under high water or EG loading, the reaction rate was lowered: the high water content was detrimental to the **P1** selectivity by promoting **I1** cleavage to C2-aldehyde, while the high EG content (enhancing reactive intermediates stabilization) was beneficial for selective depolymerisation.[23]

In summary, stereo-preferences for the formation and conversion of intermediates **B** and **I1** were obtained upon the triflate-catalysed diol (EG)-stabilised acidolysis. In terms of **B**, its *erythro* configuration was preferably formed and converted almost regardless of the reaction conditions. The formation of *trans* **I1** was favored, although only when the EG and water

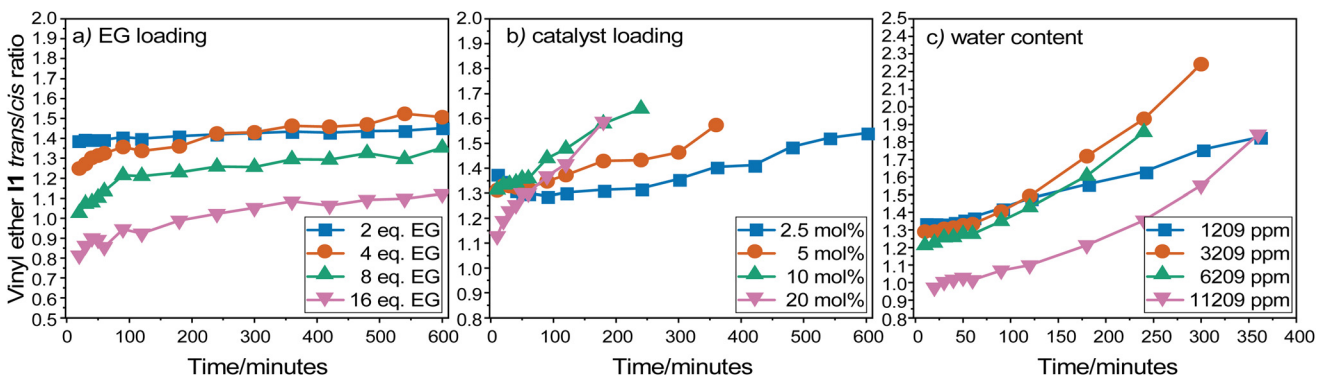


Fig. 6 The change of *trans*/*cis* ratio of vinyl ethers **I1** along time under different (a) EG loading, (b) catalyst loading, (c) water content. Yields obtained *via* HPLC analysis using 1,2,4,5-tetramethylbenzene as the internal standard. Reaction condition with different EG loading: 10 mol% Yb $(OTf)_3$, in dry 1,4-dioxane, at 130 °C. Reaction condition with different catalyst loading: 4 eq. EG, in dry 1,4-dioxane, at 130 °C. Reaction condition with different water content: 10 mol% $Yb(OTf)_3$, 4 eq. EG, in dry, at 130 °C.

content were not too high; its *cis* counterpart was preferably consumed. However, the catalytic origin of the observed stereo-preferences, and whether these preferences are controlled kinetically or thermodynamically, remains unanswered.

Elucidation of the reaction mechanism leading to the stereo-preferred formation and conversion of EG adducts B and C2-vinyl ethers I1 *via* ReaxFF-molecular dynamics and DFT-NEB calculations

These observed results were investigated further by computational methodologies with an aim to find their mechanistic origins considering molecular models at the atomic scale. Obtaining of the representative structures of different isomers is of vital importance to perform accurate simulations. Possible molecular conformations were characterized, and dynamics/reactivity of the different species during the multiple stages of the reactions were disclosed. Through the conformational search of the intermediates **B** and **I1**, 170 conformers for **B** in *threo* and *erythro* configurations and 100 structures for **I1** in *cis* and *trans* conformations were identified. These were grouped and selected for DFT optimisation to extract the stereoisomers with the lowest energies, which included four stereoisomers of **B** and two stereoisomers of **I1** (Fig. 7 and Fig. S6–S8†). The minimum energy structure of **B** in *erythro* configuration is approximately 2 kcal mol^{-1} more stable than the one in the *threo* configuration, which means that the **B** in the *erythro* configuration is thermodynamically more stable (Table S1†). Considering that the *threo*/*erythro* ratio of **B** starts at 0.5 under different reaction conditions, this is in line with the preferential formation of *erythro* **B**. The structure optimisation results suggest a thermodynamically controlled step for the formation of **B** from the carbocation originated by the dehydration of **A** (Scheme 2). Regarding **I1**, the *cis* conformation was thermodynamically more stable than the *trans* (also approximately 2 kcal mol^{-1} difference). Experimentally, *trans* **I1** had a higher yield than the *cis* **I1** showing that this step is not under thermodynamic control. Additionally, **B** in the *erythro* configuration was consumed faster. The reaction mechanism leading to the experimental stereo-preference observation (*i.e.*, from *erythro* **B** to *trans* **I1**) was further explored through RMD and DFT-NEB calculations. The release of Brønsted acids from metal salts with weakly coordinating anions readily occurs.[29] Recent studies reveal that the active species of metal triflates for various reactions is triflic acid.[24,27] Traces of triflic acid are readily generated *in situ* under the reaction conditions using metal triflate salts. Therefore, triflic acid was used in the simulation as the catalyst. The triflate anion is acknowledged as a weak base, but recent reviews demonstrate that the triflate anion can be directly involved in different reactions.[59] The formed triflate types counter ions can facilitate proton shuttling, analogous to hydrogensulfate counter ions.[17,21,28,60–62] Examining the sulfur–oxygen pair distribution functions displayed in Fig. 8a indicates that triflic acid molecules are mainly located near O3, O4, and O5 and closer to O3 of the β-O-4 motif. The inspection of the dynamics of the O3 and O4 side chains, which is reflected in the distance distribution plots shown in Fig. 8b, reveals that O3 essentially preserved its positions, whereas O4 showed the tendency to adopt many different conformations in response to the local environment. Indeed, this longer side chain can freely fluctuate in solution. The spatial distribution functions (SDFs) shown in Fig. 9 confirmed the trend of the S–O distance distribution, suggesting again that the most probable acid binding sites are those close to O3 far from the aromatic rings.

The formation mechanism of C2-acetal **P1** from C2-vinyl ethers **I1** has been exhaustively investigated in an earlier

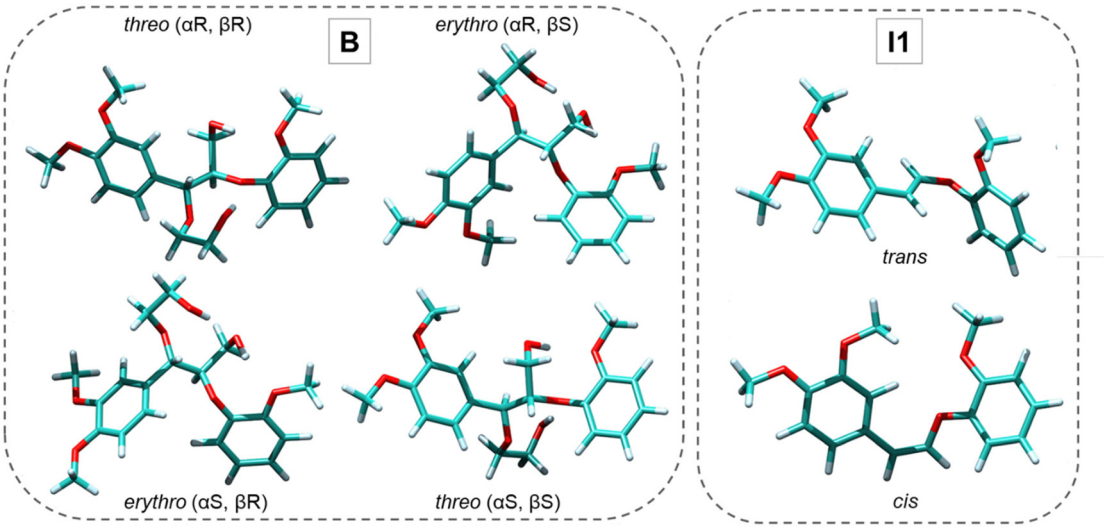


Fig. 7 Ball and stick representation of the minimum energy structures of the **B** and **I1** intermediates singled out *via* computational CS searches. Carbon atoms in cyan, oxygen in red, and hydrogen in white.

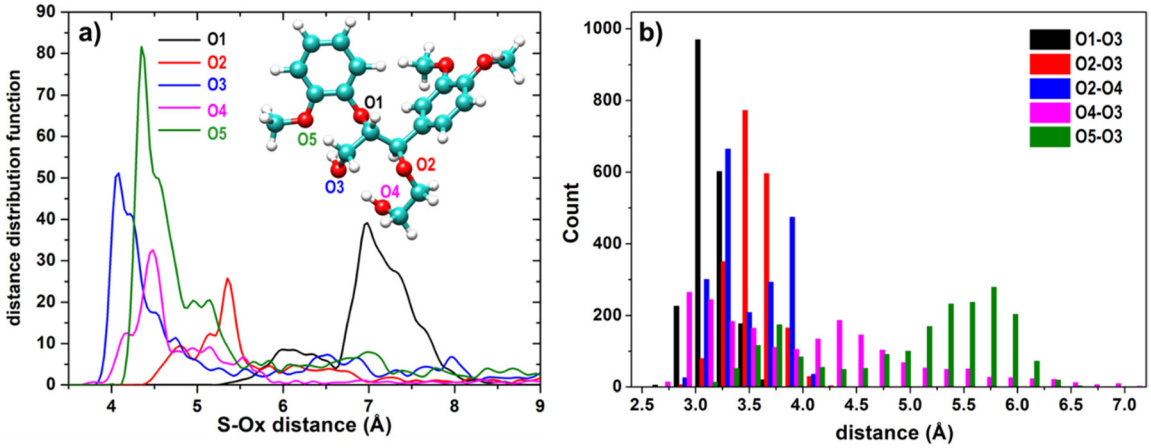


Fig. 8 (a) Distance distribution functions of the sulfur atom of triflic acid from the oxygen atoms of the *erythro* EG adduct **B** (absolute configuration in α*R*,β*S*). (b) Distributions of the interatomic distances of the *erythro* EG adduct **B** (in α*R*,β*S* configuration) oxygen atoms.

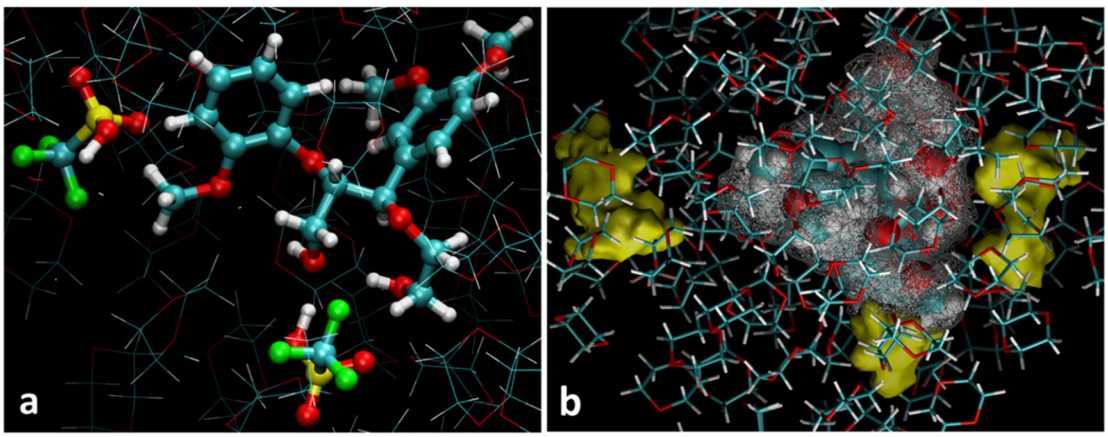


Fig. 9 (a) Snapshots selected for the NEB simulations. (b) Spatial distribution functions representing the most probable positions of triflic acid molecules (yellow contours) around the *erythro* EG adduct **B**. The last 2000 configurations of *erythro* EG adduct **B** have been superimposed on the average structure and displayed as vdW spheres inside the density contour plot of the *erythro* EG adduct **B** (grey wireframe) to give an idea of the various orientations/configurations of the molecule. The closest solvent molecules are also shown (sticks). Carbon atoms cyan, oxygen red, hydrogen white, sulfur yellow, and fluorine green.

study;[21] thus here, the mechanism leading to the favored transformation from *erythro* **B** to *trans* **I1** is clarified by checking the abundance of stereoisomers. In the traditional lignin acidolysis reaction, distinct from HCl or HBr catalysed lignin acidolysis (C3 acidolysis pathway), the triflic acid- (metal triflates-) and sulfuric acid-catalysed lignin acidolysis undergoes the C2 cleavage pathway. To illustrate this, the acidolysis was firstly run for 90 min catalysed by HCl, after which $Yb(OTf)_3$ was added (Fig. S5†). Clearly, we could observe that the C2 pathway was activated after the addition of $Yb(OTf)_3$, while there is nearly no acidolysis reaction in the first 90 min with HCl as strong Brønsted acid. Therefore, the anion must play a crucial role in the reaction, so we are particularly interested in understanding its role in governing the reaction mechanisms that result in stereo preferences. Starting from the *erythro* **B** shown in Fig. 9a, we simulated the reaction mechanism through NEB calculations. These could provide a qualitative (even semi-quantitative) description of the energy pathway of the reaction mechanism, reproducing the enhanced stability of *cis* **I1** relative to *trans* **I1** (of about 1.6 kcal mol^{-1}). The reaction energy profile and mechanism (Fig. 10 and Scheme 4) is driven by proton transfers between triflic acid and the ethylene glycol side-chain (O2) of **B**, then the proton on the γ-OH (O3) is coordinated to the formed triflate anion releasing formaldehyde and **I1**. The reaction from *erythro* **B** to *trans* **I1** can be formally divided into two separate steps. First, protonation of **B** on the O2 atom by triflic acid leads to the removal of EG *via* TS1, which interacts *via* a network of hydrogen bonds with the resulting β-O-4 that has an α-carbocation (CC^+), and the anionic triflate species (Fig. 10b and Scheme 4). This stabilised

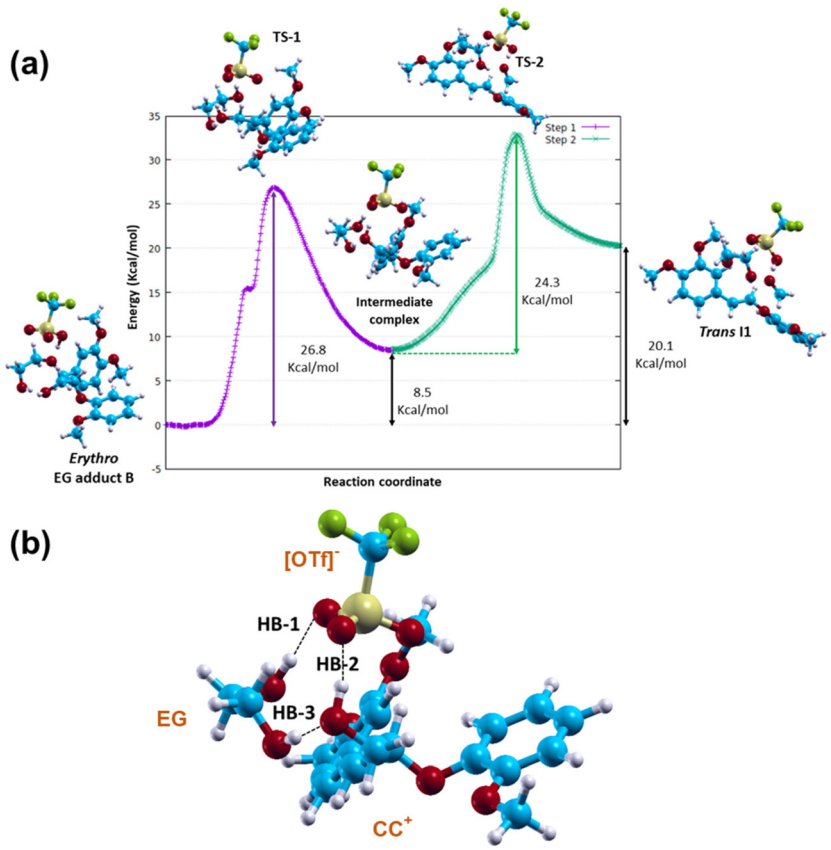
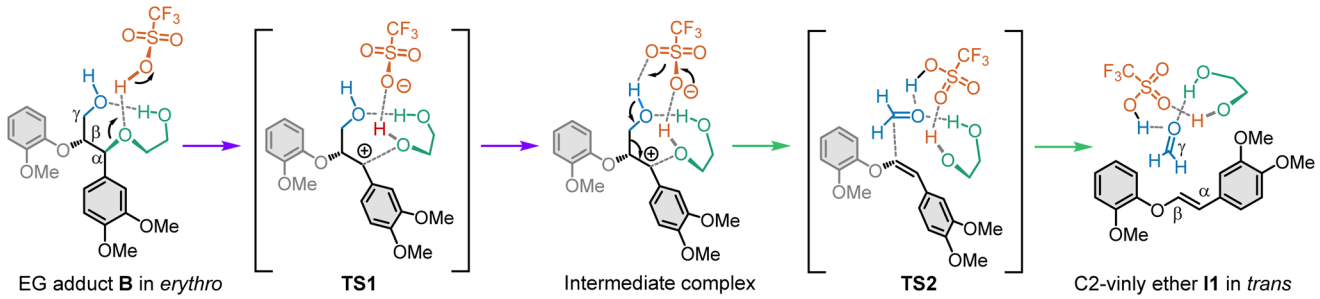


Fig. 10 (a) Reaction energy profile of the *erythro* EG adduct **B** with the acid catalyst HOTf leading to the formation of the *trans* C2-vinyl ether **I1** (and an equivalent of EG and formaldehyde), as simulated at the DFT/PBE-D2 level of theory. The path has been divided into two separate steps. Initial, intermediate, and final states are shown together with the structure of the TS (transition states) of the two steps. (b) Structure and configuration of the intermediate complex (CC^+, EG adduct, and [OTf]−) highlighting the formation of the three hydrogen bonds responsible for the complex stabilization as ionic pair. Carbon atoms cyan, oxygen red, hydrogen white, sulfur yellow, and fluorine green. A schematic representation can be found in Scheme 4.



Scheme 4 Schematic representation of the structural changes according to the energy profile displayed in Fig. 10 involving the reaction mechanism of *erythro* EG adduct **B** with the acid catalyst HOTf leading to the formation of *trans* C2-vinyl ether **I1**.

structure, named "intermediate complex", links the two steps of the reaction.

The analysis of this intermediate complex *via* estimation of Mulliken charges reveals that the CC^+, and EG carry a total charge of approximately +0.7*e* balanced by an equal negative charge on the anionic triflate. Deprotonation from the B^+ carbocation to the triflate restores the overall neutrality of the fragments, then the *trans* **I1** and a formaldehyde molecule were formed *via* TS2. The barrier of the whole process, which is the sum of the energy difference between the starting **B** and the intermediate complex (8.5 kcal mol^{-1}) and the energy barrier of the second step, is approximately 32.8 kcal mol^{-1}, in line with the previously published experimental kinetics.[23] Such a mechanism reasonably explains the preferred for-

mation of the *trans* **I1**, regardless of its higher energy relative to the *cis* **I1** isomer. A detailed analysis of the mechanism reveals that the dihedral angle at the formation site of the C=C double bond adopts values between 166° and 179°. This confirms the preferential transformation of the *erythro* **B** to the *trans* **I1** *via* the intermediate complex adopting a proto-*trans* vinyl ether conformation. As can be deduced from the structures shown in Table S1† and Fig. 7, the (α*S*,β*R*) and (α*R*,β*S*) configurations of **B** (both are in the *erythro* configuration) show an analogous spatial arrangement in the active region of the acidolysis (β-O-4 cleavage) process; thus, a similar reactive scenario for **B** in the (α*R*,β*S*) configuration is highly expected. On the other hand, if we consider the structure of **B** in the *threo* configurations, *i.e.*, (α*R*,β*R*) and (α*S*,β*S*), which, analogously, are very similar in the region of the β-O-4 cleavage (as shown in Fig. 7), a competitive path for the formation of **I1** was not identified. For **B** in the *threo* configuration, in fact, the carbocationic/anionic pair stabilization was not achieved due to the missing network of hydrogen bonds, which, in turn, allows for the stabilization of the intermediate complex of the *erythro* configurations. The path toward the *trans* **I1** is thus kinetically favored by specific stabilizing interactions, which explains the observed higher concentration of *trans* **I1** than *cis* **I1** during the formation phase. These calculations also reveal a possible explanation of the influence of the hydrogensulfate and triflate species on the reaction (both can facilitate the C2 pathway over the C3 one). The proton shuttling that occurs over two of the oxygens of these counterions that are spatially separated suggesting that a single counterion can donate and receive protons from different positions of the β-O-4 motif, which can enhance the reverse Prins reaction required for the C2 catalytic pathway. On the other hand, we cannot exclude the EG and water in the reaction function as the final acceptor of the proton transfer. How these are further involved the reaction is the subject of future studies.

Conclusions

In summary, the stereo-preferences in the formation and consumption of two important intermediates were identified during the triflate-catalysed diol (EG)-stabilised acidolysis of the lignin β-O-4 motif by their accurate quantification in the reaction under different temperatures (80–150 °C), catalyst concentration (2.5–20 mol%), ethylene glycol content (2–16 eq.), and water content (1209–11 209 ppm). The stereo-preferred formation of the EG adduct **B** in the *erythro* configuration from the *erythro* β-O-4 model compound **A** is revealed. From **B** in the *erythro* configuration, the formation of *trans* **I1** is favored. Despite rate differences, the stereo-preferred formation and conversion of *erythro* **B** are independent of temperatures, catalyst loading, water, and ethylene glycol content. The stereo-preferred formation of *trans* **I1** is more sensitive to extreme conditions (*i.e.*, very high water and ethylene glycol content in the reaction), under which *cis* **I1** becomes slightly preferred to be formed.

The computational simulations elucidated that the formation of **B** in the *erythro* configuration is thermodynamically favored. The reaction from **B** in the *erythro* configuration to the *trans* **I1** is shown to be a kinetically controlled step driven by proton transfers. The computational simulations reveal that the triflate anion plays an important catalytic role by facilitating proton transfers. These consist of two separate steps (*i.e.*, protonation and deprotonation) with an activation energy of 32.8 kcal mol^{-1}. These two steps are linked *via* an "intermediate complex" stabilised by a network of hydrogen bonds between the anionic [OTf]– species, carbocation, and the release of ethylene glycol. The dihedral angle at the formation site of **I1** ends up between 166° and 179°, revealing a proto-*trans* vinyl ether accounting for the preference for the *trans* **I1**. In contrast, **B** in the *threo* configuration could not lead to an analogous formation of **I1** due to the lack of an intermediate complex that is stabilised to a similar extent. To the best of our knowledge, this is the first study that dives into the stereo-preference and its catalytic origins encountered in lignin β-O-4 acidolysis, which is relevant not only for depolymerisation but also for typical lignin fractionation processes under acidic conditions. This study may also provide a basis for understanding the preference of triflate and sulfate-type catalysts for the C2 cleavage pathway over the C3 cleavage pathway, which is essential to steer product selectivity.

Experimental

Materials and experimental procedures

Ytterbium(III) trifluoromethanesulfonate $Yb(OTf)_3$ was obtained from Sigma-Aldrich. Other chemicals supplied by Acros, Sigma-Aldrich, Fisher Chemicals, or Fluorochem were used as received. The 1,4-dioxane utilized in this work was acquired anhydrous under argon and used following standard Schlenk techniques. The model compounds **A**, **B**, **I1**, C2-acetal **P1** (Scheme 3) were synthesised according to the published procedures (ESI, S2.0† for details). The typical procedure of running the time course reaction is provided in S3.0, ESI.† Analytical procedures for HPLC, preparative HPLC, UPLC-MS, NMR, GC-MS, *etc*. are provided in the ESI (see S1.0† for details).

Computational methods

Three computational approaches at different levels of theory and scale, namely conformational search (CS) (classical method based on a force field (FF) followed by DFT optimisation – single molecule), reactive molecular dynamics simulations (RMD) (classical method based on a reactive FF – supramolecular system), and nudged elastic band (NEB) calculations (DFT – reduced supramolecular model) were used sequentially to identify the most probable conformations of the intermediates, appropriate arrangements of the molecules during the reaction process in solution, and possible reaction mechanisms with the lowest energy barrier.

The CSs of the **B** and **I1** intermediates were carried out with the Balloon software,[63,64] selecting the MMFF94 force field, conformation-dependent electronegativity equalization charges, and a multi-objective genetic algorithm. The identified structures (approximately 170 geometries for **B** and 100 structures for **I1**) were optimised at the DFT level using the M06-2X functional and the 6-31G(d) basis set, employing the Gaussian 09 (Revision A.02) package,[43] as done in a previous investigation for similar types of lignin derivatives.[44] The optimised structures were grouped into families based on the energy difference to the minimum energy conformer determined in each case, considering an energy cutoff of 6 and 1 kcal mol^{-1} for **B** and **I1** compounds, respectively. Then, all the torsional angles were analyzed to classify the structures (see Fig. S6–S8 and Table S1† for a detailed analysis of the torsional angles).

The RMD simulations were carried out through the Amsterdam Density Functional (ADF)/ReaxFF[45] software with an explicit representation of the 1,4-dioxane solution consisting of a pre-equilibrated/randomized solvent box that produced results in line with the reference[65] (see, for example, Fig. S9† where the radial distribution function of 1,4-dioxane oxygen is displayed). Solute and solvent (which contained a few molecules of triflic acid) were described through a FF used in earlier investigations.[20] The lowest energy structure of **B** in the *erythro* form, picked out from the CS, was inserted in the simulation box ($34 \times 22 \times 26$ $Å^3$), which contained 70 molecules of 1,4-dioxane and three molecules of triflic acid. After a standard preliminary equilibration procedure in the NVT ensemble, the simulations were then carried out, first, in the NPT ensemble (200 ps) at constant pressure (1 bar), using the Berendsen's barostat[47] with isotropic molecule-based scaling and a time constant of 1 ps, and then in the NVE ensemble (800 ps). The temperature (T = 300 K) was maintained by coupling the system to a thermal bath with the Andersen algorithm and a time constant of 1 ps. The integration step was set to 0.25 fs, and configurations were collected every 0.1 ps. The last 100 ps of the simulations were used for the system characterization and the extraction of promising structures. The analysis was focused on interatomic distances and solute-solvent interactions, which were rendered through atom-atom distance distribution functions (DDFs) and spatial distribution functions (SDFs).

The NEB simulations were performed with the Quantum Espresso (QM) software[48] using periodic boundary conditions, plane-waves basis sets, ultrasoft pseudopotentials, and a description of exchange/correlation with a gradient-corrected functional (PBE)[49] corrected for vdW interactions.[50] Cutoffs on the one-particle wave function and electronic density were set to 40 and 400 Ry, respectively. The first Brillouin cell in the reciprocal space was sampled at the Gamma point only. All the calculations were spin-unrestricted with Gaussian smearing of the one-particle energy levels of 0.002 Ry. Twelve images along the minimum energy path (MEP) were chosen for each reaction.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

Z. Zhang acknowledges the China Scholarship Council for funding (grant no. 201704910922). The work performed by C. W. L. has partly been conducted within the framework of the Dutch TKI-BBEI project "CALIBRA", reference TEBE117014.

References

1 Z. Sun, B. Fridrich, A. de Santi, S. Elangovan and K. Barta, *Chem. Rev.*, 2018, **118**, 614–678.
2 C. Zhang and F. Wang, *Acc. Chem. Res.*, 2020, **53**, 470–484.
3 W. Schutyser, T. Renders, S. Van den Bosch, S. F. Koelewijn, G. T. Beckham and B. F. Sels, *Chem. Soc. Rev.*, 2018, **47**, 852–908.
4 Z. Zhang, J. Song and B. Han, *Chem. Rev.*, 2017, **117**, 6834–6880.
5 E. Adler, *Wood Sci. Technol.*, 1977, **11**, 169–218.
6 K. Lundquist, *Acta Chem. Scand.*, 1967, **21**, 1750–1754.
7 K. Lundquist, *Acta Chem. Scand.*, 1970, **24**, 889–907.
8 K. Lundquist and L. Ericsson, *Acta Chem. Scand.*, 1970, **24**, 3681–3686.
9 O. Karlsson, K. Lundquist, S. Meuller and K. Westlid, *Acta Chem. Scand., Ser. B*, 1988, **42**, 48–51.
10 M. Kulka and H. Hibbert, *J. Am. Chem. Soc.*, 1943, **65**, 1180–1185.
11 T. Yokoyama and Y. Matsumoto, *Holzforschung*, 2008, **62**, 164–168.
12 H. Ito, T. Imai, K. Lundquist, T. Yokoyama and Y. Matsumoto, *J. Wood Chem. Technol.*, 2011, **31**, 172–182.
13 T. Imai, T. Yokoyama and Y. Matsumoto, *J. Wood Chem. Technol.*, 2012, **32**, 165–174.
14 K. Lundquist, R. Lundgren, J. Danielsen, A. Haaland and S. Svensson, *Acta Chem. Scand.*, 1972, **26**, 2005–2023.
15 T. Imai, T. Yokoyama and Y. Matsumoto, *J. Wood Sci.*, 2011, **57**, 219–225.
16 T. Yokoyama, *J. Wood Chem. Technol.*, 2014, **35**, 27–42.
17 M. R. Sturgeon, S. Kim, K. Lawrence, R. S. Paton, S. C. Chmely, M. Nimlos, T. D. Foust and G. T. Beckham, *ACS Sustainable Chem. Eng.*, 2013, **2**, 472–485.
18 Y. M. Questell-Santiago, M. V. Galkin, K. Barta and J. S. Luterbacher, *Nat. Rev. Chem.*, 2020, **4**, 311–330.
19 A. Kaiho, M. Kogo, R. Sakai, K. Saito and T. Watanabe, *Green Chem.*, 2015, **17**, 2780–2783.
20 P. J. Deuss, M. Scott, F. Tran, N. J. Westwood, J. G. de Vries and K. Barta, *J. Am. Chem. Soc.*, 2015, **137**, 7456–7467.
21 A. De Santi, S. Monti, G. Barcaro, Z. Zhang, K. Barta and P. J. Deuss, *ACS Sustainable Chem. Eng.*, 2021, **9**, 2388–2399.

22 Y. Liu, N. Deak, Z. Wang, H. Yu, L. Hameleers, E. Jurak, P. J. Deuss and K. Barta, *Nat. Commun.*, 2021, **12**, 5424.
23 Z. L. Zhang, C. W. Lahive, J. G. M. Winkelman, K. Barta and P. J. Deuss, *Green Chem.*, 2022, **24**, 3193–3207.
24 E. T. Sletten, Y. J. Tu, H. B. Schlegel and H. M. Nguyen, *ACS Catal.*, 2019, **9**, 2110–2123.
25 T. T. Dang, F. Boeck and L. Hintermann, *J. Org. Chem.*, 2011, **76**, 9353–9361.
26 D. C. Rosenfeld, S. Shekhar, A. Takemiya, M. Utsunomiya and J. F. Hartwig, *Org. Lett.*, 2006, **8**, 4179–4182.
27 J. D. Steen, S. Stepanovic, M. Parvizian, J. W. de Boer, R. Hage, J. Chen, M. Swart, M. Gruden and W. R. Browne, *Inorg. Chem.*, 2019, **58**, 14924–14930.
28 P. J. Deuss, C. W. Lahive, C. S. Lancefield, N. J. Westwood, P. C. Kamer, K. Barta and J. G. de Vries, *ChemSusChem*, 2016, **9**, 2974–2981.
29 C. F. Baes and R. S. Mesmer, *The Hydrolysis of Cations*, 489 Seiten, Preis: £ 18.60. Berichte der Bunsengesellschaft für physikalische Chemie, John Wiley & Sons, New York, London, Sydney, Toronto, 1976, 81, pp. 245–246, DOI: 10.1002/bbpc.19770810252.
30 P. J. Deuss, C. S. Lancefield, A. Narani, J. G. de Vries, N. J. Westwood and K. Barta, *Green Chem.*, 2017, **19**, 2774–2782.
31 A. De Santi, M. V. Galkin, C. W. Lahive, P. J. Deuss and K. Barta, *ChemSusChem*, 2020, **13**, 4468–4477.
32 Z. Wang, Z. Zhang, H. Wang and P. J. Deuss, *Chem. Catal.*, 2022, **2**, 1407–1427.
33 C. Dong, X. Meng, C. S. Yeung, H.-Y. Tse, A. J. Ragauskas and S.-Y. Leu, *Green Chem.*, 2019, **21**, 2788–2800.
34 C. W. Lahive, P. C. J. Kamer, C. S. Lancefield and P. J. Deuss, *ChemSusChem*, 2020, **13**, 4238–4265.
35 T. Akiyama, H. Goto, D. S. Nawawi, W. Syafii, Y. Matsumoto and G. Meshitsuka, *Holzforschung*, 2005, **59**, 276–281.
36 M. Hauteville, K. Lundquist and S. von Unge, *Acta Chem. Scand., Ser. B*, 1986, **40**, 31–35.
37 T. Akiyama, K. Magara, G. Meshitsuka, K. Lundquist and Y. Matsumoto, *J. Wood Chem. Technol.*, 2014, **35**, 8–16.
38 V. Langer, K. Lundquist and J. Parkås, *BioResources*, 2007, **2**, 590–597.
39 K. Lundquist, V. Langer and J. Parkås, *BioResources*, 2009, **4**, 529–536.
40 K. Lundquist, V. Langer, S. Li and R. Stomberg, *Lignin Stereochemistry and its Biosynthetic Implications*, 12th International Symposium on Wood and Pulping Chemistry, Madison, Wisconsin, USA, 2003, vol. 1, p. 239–244.
41 S. Besombes, D. Robert, J.-P. Utille, F. R. Taravel and K. Mazeau, *J. Agric. Food Chem.*, 2003, **51**, 34–42.
42 R. Stomberg and K. Lundquist, *Nord. Pulp Pap. Res. J.*, 1994, **9**, 37–43.
43 S. Shimizu, T. Yokoyama, T. Akiyama and Y. Matsumoto, *J. Agric. Food Chem.*, 2012, **60**, 6471–6476.
44 S. Shimizu, P. Posoknistakul, T. Yokoyama and Y. Matsumoto, *BioResources*, 2013, **8**, 4312–4322.
45 C. Bohlin, K. Lundquist and L. J. Jönsson, *Enzyme Microb. Technol.*, 2008, **43**, 199–204.
46 C. Bohlin, P.-O. Andersson, K. Lundquist and L. J. Jönsson, *J. Mol. Catal. B: Enzym.*, 2007, **45**, 21–26.
47 P. Posoknistakul, T. Akiyama, T. Yokoyama and Y. Matsumoto, *J. Wood Chem. Technol.*, 2016, **36**, 288–303.
48 P. Posoknistakul, T. Akiyama, T. Yokoyama and Y. Matsumoto, *J. Wood Chem. Technol.*, 2016, **37**, 87–98.
49 P. Posoknistakul, S. Akiho, T. Akiyama, T. Yokoyama and Y. Matsumoto, *J. Wood Sci.*, 2018, **64**, 451–457.
50 J. Ralph, C. Lapierre and W. Boerjan, *Curr. Opin. Biotechnol.*, 2019, **56**, 240–249.
51 X. Huang, O. M. Morales Gonzalez, J. Zhu, T. I. Korányi, M. D. Boot and E. J. M. Hensen, *Green Chem.*, 2017, **19**, 175–187.
52 R. Jastrzebski, S. Constant, C. S. Lancefield, N. J. Westwood, B. M. Weckhuysen and P. C. Bruijnincx, *ChemSusChem*, 2016, **9**, 2074–2079.
53 A. C. H. Braga, S. Zacchino, H. Badano, M. G. Sierra and E. A. Rúveda, *Phytochemistry*, 1984, **23**, 2025–2028.
54 S.-S. Lee, N.-I. Baek, Y.-S. Baek, D.-K. Chung, M.-C. Song and M.-H. Bang, *Molecules*, 2015, **20**, 5616–5624.
55 T. Miyase, A. Ueno, N. Takzawa, H. Kobayashi and H. Oguchi, *Chem. Pharm. Bull.*, 1987, **35**, 3713–3719.
56 T. Kikuchi, S. Matsuda, S. Kadota and T. Tai, *Chem. Pharm. Bull.*, 1985, **33**, 1444–1451.
57 S. Dabral, J. Mottweiler, T. Rinesch and C. Bolm, *Green Chem.*, 2015, **17**, 4908–4912.
58 S. Dabral, J. Engel, J. Mottweiler, S. S. M. Spoehrle, C. W. Lahive and C. Bolm, *Green Chem.*, 2018, **20**, 170–182.
59 B. Dhakal, L. Bohe and D. Crich, *J. Org. Chem.*, 2017, **82**, 9263–9269.
60 Y. Jing, X. Mu, Z. Han, C. Liu and D. Zhang, *Mol. Catal.*, 2019, **463**, 140–149.
61 G. Kovács, G. Ujaque and A. Lledós, *J. Am. Chem. Soc.*, 2008, **130**, 853–864.
62 B. Yuan, R. He, W. Shen and M. Li, *Eur. J. Org. Chem.*, 2017, **2017**, 3947–3956.
63 M. J. Vainio and M. S. Johnson, *J. Chem. Inf. Model.*, 2007, **47**, 2462–2474.
64 J. S. Puranen, M. J. Vainio and M. S. Johnson, *J. Comput. Chem.*, 2010, **31**, 1722–1732.
65 R. Barata-Morgado, M. L. Sánchez, I. Fdez. Galván, J. C. Corchado, M. E. Martín, A. Muñoz-Losa and M. A. Aguilar, *Theor. Chem. Acc.*, 2013, **132**, 1390.