JOC The Journal of Organic Chemistry

Article

Molecular-Level Understanding of the Major Fragmentation Mechanisms of Cellulose Fast Pyrolysis: An Experimental Approach Based on Isotopically Labeled Model Compounds

Zaikuan J. Yu, Mckay Whetton Easton, Priya Murria, Lan Xu, Duanchen Ding, Yuan Jiang, Jifa Zhang, and Hilkka I. Kenttämaa

J. Org. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.joc.9b00723 • Publication Date (Web): 08 May 2019

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Molecular-Level Understanding of the Major Fragmentation Mechanisms of Cellulose Fast Pyrolysis: An Experimental Approach Based on Isotopically Labeled Model Compounds

Zaikuan J. Yu,‡ Mckay W. Easton,‡ Priya Murria, Lan Xu, Duanchen Ding, Yuan Jiang, Jifa Zhang, and Hilkka I. Kenttämaa*

Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

ABSTRACT: Evaluation of the feasibility of various mechanisms possibly involved in cellulose fast pyrolysis is challenging. Therefore, selectively ^{13}C-labeled cellotriose, ^{18}O-labeled cellobiose, and ^{13}C- and ^{18}O-doubly-labeled cellobiose were synthesized and subjected to fast pyrolysis in an atmospheric pressure chemical ionization source of a linear quadrupole ion trap/orbitrap mass spectrometer. The initial products were immediately quenched, ionized using ammonium cations and subsequently analyzed by the mass spectrometer. The loss or retention of isotope labels upon pyrolysis unambiguously revealed three major competing mechanisms – sequential losses of glycolaldehyde/ethenediol molecules from the reducing end (the reducing-end unraveling mechanism), hydroxymethylene-assisted glycosidic bond cleavage (HAGBC mechanism), and Maccoll elimination. Important discoveries include: 1) reducing-end unraveling is the predominant mechanism occurring at the reducing end; 2) Maccoll elimination cleaves aglyconic bonds and it is the mechanism leading to formation of reducing carbohydrates; 3) HAGBC occurs for glycosides but not at the reducing end of cellodextrins; 4) HAGBC and water loss are the predominant reactions for fast pyrolysis of 1,6-anhydrocellodextrins; and 5) HAGBC can proceed after reducing-end unraveling but unraveling does not occur once the HAGBC reaction pathway is initiated. Moreover, hydrolysis was conclusively ruled out for fast pyrolysis of cellobiose, cellotriose, and 1,6-anhydrocellodextrins up to cellotetraosan.

INTRODUCTION

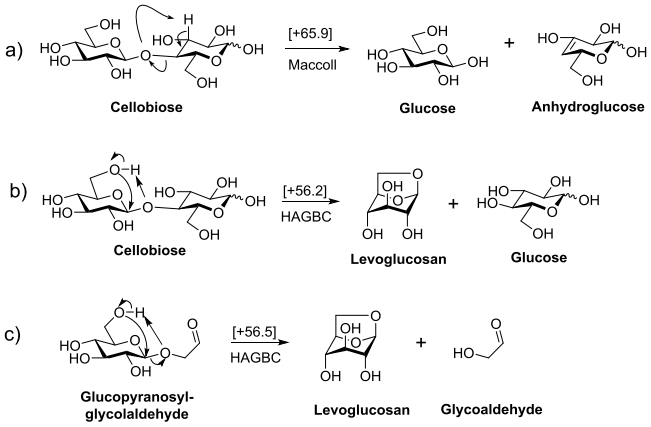
The energy crisis caused by the fast consumption of light fossil fuels has triggered intense research efforts to generate alternative fuels from renewable biomass. Among those efforts, fast pyrolysis of biomass followed by catalytic upgrading to biofuel continues to be one of the most promising approaches.[1–6] Cellulose contributes up to 40∼50% of the weight of biomass depending on the plant species and has therefore attracted great interest.[7,8] However, despite many investigations carried out to delineate the chemical processes occurring during cellulose fast pyrolysis, molecular-level understanding of the mechanisms of these reactions is far from satisfactory. This is partially due to the fact that most studies have utilized gas-chromatography/mass spectrometry (GC/MS) for product analysis.[9–14] This approach does not allow the detection of the initial products but instead downstream secondary and later products due to the long analysis time. Furthermore, this analysis method can only detect volatile compounds.[15]

Depolymerization of cellulose during fast pyrolysis involves cleavages of the β-1,4-glycosidic linkages between adjacent glucose units. Breaking the glycosidic linkages involves the cleavage of either the glycosidic or aglyconic bond, which can occur via different mechanisms, including simple bond cleavages via homolysis[16–19] or heterolysis,[16,20,21] or assisted bond cleavages, as for example via hydrolysis.[16,18,22] Previous studies suggest that formation of an anhydroglucose diradical by homolytic cleavages of two consecutive glycosidic bonds of cellulose may be followed by rearrangement into levoglucosan (LGN) – one of the major fast pyrolysis products of cellulose.[16,23] However, for the cellobiose model system, the enthalpy for the homolytic dissociation of the glycosidic bond has been calculated to be prohibitively high for this to be an important mechanism (101 kcal/mol at the B3LYP/6-31+G(d) level of theory).[23] The enthalpy for the heterolytic cleavage of the glycosidic bond in cellobiose has been reported to be even higher (calculated to be 158 kcal/mol at the B3LYP/6-31+G(d) level of theory).[23] Another possible mechanism for dissociation of the glycosidic bonds is hydrolysis that would ultimately produce glucose, after which glucose may fragment via numerous mechanisms.[24,25] However, no direct experimental evidence has yet been provided to conclusively confirm or reject the occurrence of hydrolysis during fast pyrolysis of cellulose.

Besides aforementioned homolytic and heterolytic bond cleavages and hydrolysis, the glycosidic linkage can also be broken through other mechanisms. Two such mechanisms are the Maccoll elimination that forms a glucose and an anhydrocellodextrin that is not 1,6-anhydrous (Scheme 1a, discussed in more detail later),[18,26] and hydroxymethylene-assisted glycosidic bond cleavage (HAGBC) to form a 1,6-anhydrocellodextrin and a cellodextrin with reduced degree of polymerization (Scheme 1b).[6,16,18,21,27–29] The energetics of HAGBC reactions have been examined previously.[29] It is

Scheme 1. Examples of a) Maccoll elimination cleaving the aglyconic bond in cellobiose that produces glucose

and anhydroglucose, b) hydroxymethylene-assisted glycosidic bond cleavage (HAGBC) for cellobiose that produces levoglucosan and glucose, and c) HAGBC for glucopyranosylglycolaldehyde (a glucoside) that produces levoglucosan and glycolaldehyde.



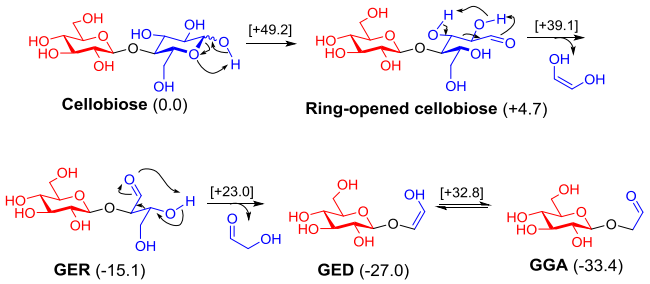
Numbers in brackets are free energies (in kcal/mol) for the transition states: a) the value was calculated at the M06-2X/6-31G+(d,p) level of theory;[30] b) and c) both values were calculated at the M06-2X/6-311++G(d,p) level of theory.[1,31]

worth noting that, theoretically, Maccoll elimination can occur at both glycosidic and aglyconic bonds while HAGBC only cleaves a glycosidic bond. Also, based on literature, HAGBC can occur at any point along the cellulose chain. The resulting 1,6-anhydrocellodextrin can undergo another HAGBC reaction to form other smaller 1,6-anhydrocellodextrins (cellobiosan, cellotriosan, cellotetraosan, etc., or levoglucosan). These 1,6-anhydrocellodextrins are considered to be the major components of "active cellulose" that acts as an intermediate to the production of smaller products.[17,28,24,15,32]

Another reaction mechanism that likely contributes to the fragmentation of cellulose during fast pyrolysis is the "reducing-end unraveling mechanism" that dominates for cellobiose.[1] This mechanism is initiated by ring-opening of the hemiacetal in the reducing end of cellobiose, followed by a concerted reaction leading to the loss of a C2 fragment (glycolaldehyde or ethenediol) to generate glucopyranosylerythrose and glucopyranosylethenediol (the latter can tautomerize to glucopyranosylglycolaldehyde; Scheme 2).[1,6] The resulting intermediate (e.g., glucopyranosylethenediol or glucopyranosylglycolaldehyde) can then undergo HAGBC to generate levoglucosan and glycolaldehyde (Scheme 1c).[1,31] These products have been observed upon fast pyrolysis of cellobiose.[1,33,34] It should be noted here that although cellobiose only contains two glucose units, it yields almost identical fast pyrolysis product distributions as cellohexaose, and these distributions are similar to those measured for cellulose.[34]

Although many reaction mechanisms have been proposed in the literature for cellulose fast pyrolysis, no general consensus has been reached about the most important mechanisms. In order to do this, experimental approaches are needed that can provide conclusive evidence in support or against proposed mechanisms. Herein, a previously reported[1,33] pyrolysis/high-resolution tandem mass spectrometry (Py-MS/MS) experiment, which enables the detection of the initial fast pyrolysis products, was combined with three selectively isotope-labeled cellobioses and cellotrioses to delineate the mechanisms of the major reactions occurring during cellulose fast pyrolysis.

Scheme 2. Reducing-end unraveling mechanism for cellobiose with free energy changes (in parenthesis in kcal/mol) and activation free energies (in brackets) calculated at the M06-2X/6-311++G(d,p) level of theory.[1]



Abbreviations: glucopyranosylerythrose (GER), glucopyranosylethenediol (GED), glucopyranosylglycolaldehyde (GGA),

RESULTS AND DISCUSSION

Model compounds

The nine model compounds that were purchased or synthesized for this study are shown in Figure 1. They are cellotriose (Glc-Glc-Glc-OH; CTS), selectively ^{13}C-labeled cellotriose (Glc[1-^{13}C]-Glc-Glc-OH; ^{13}C-CTS), cellobiose (Glc-Glc-OH; CBS), cellobiosan (CBN), cellotriosan, cellotetraosan, cellobiosylglycolaldehyde (Glc-Glc-glycolaldehyde; CBGA), selectively ^{18}O-labeled cellobiose (Glc-Glc-^{18}OH; ^{18}O-CBS), and ^{13}C- and ^{18}O-doubly-labeled cellobiose (Glc[1-^{13}C]-^{18}O-Glc-OH; ^{13}C-,^{18}O-CBS). A detailed procedure for the synthesis of the above isotope-labeled model compounds as well as cellobiosylglycoaldehyde can be found in the experimental section.

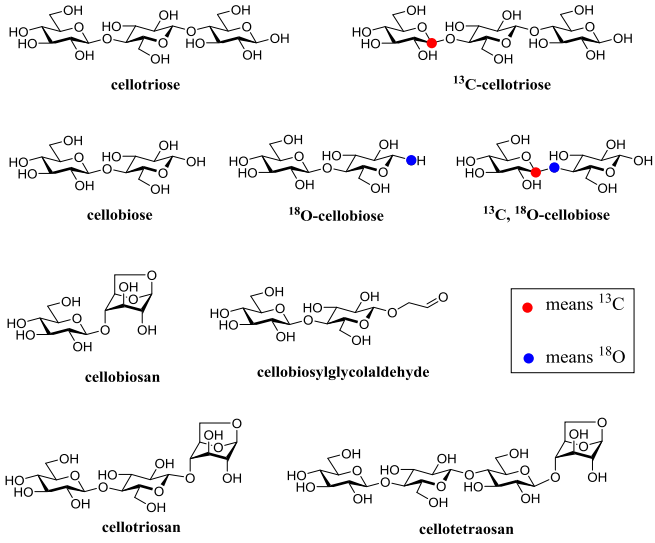


Figure 1. Model compounds used in this study.

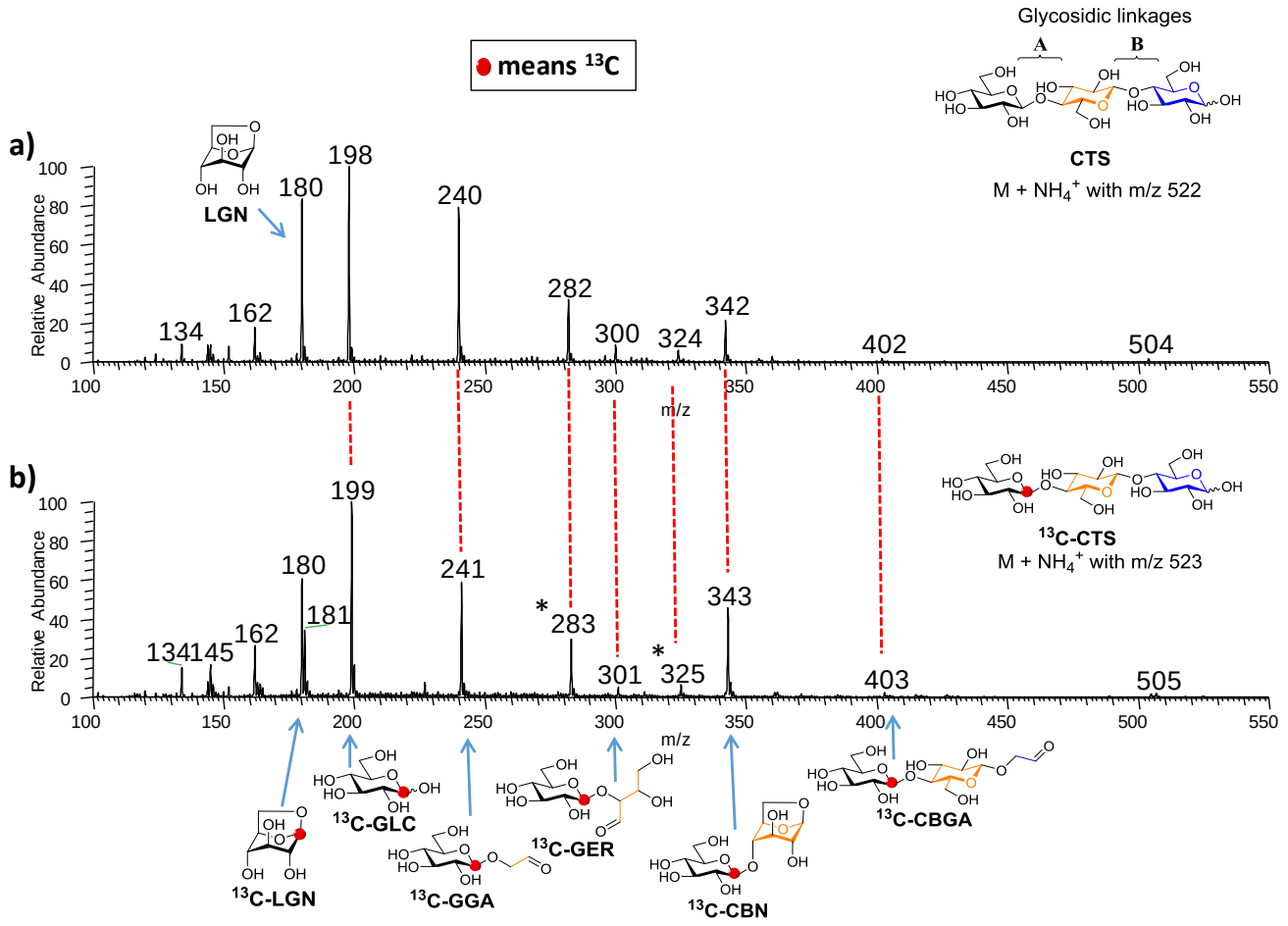


Figure 2. Mass spectra of the NH_4^+ adducts of the initial products formed upon fast pyrolysis of a) unlabeled cellotriose (CTS) and b) ^{13}C-labeled cellotriose. Abbreviations: cellobiosylglycolaldehyde (CBGA), cellobiosan (CBN; anhydrocellobiose), glucopyranosylerythrose (GER), glucopyranosylglycolaldehyde (GGA), glucose (GLC), and levoglucosan (LGN; anhydroglucose). Glycosidic linkages are labeled using A and B. *Indicates dehydration products.

Reducing-end unraveling mechanism

The ionized (via ammonium attachment) initial products formed upon fast pyrolysis of cellotriose and ^{13}C-labeled cellotriose (Glc[1-^{13}C]-Glc-Glc-OH) (Figure 1) are shown in Figure 2. The ions with the greatest relative abundances derived from unlabeled cellotriose correspond to levoglucosan and its isomers (anhydroglucoses; NH_4^+ adducts with m/z 180), glucose (NH_4^+ adduct with m/z 198), glucopyranosylglycolaldehyde (NH_4^+ adduct with m/z 240), glucopyranosylerythrose (NH_4^+ adduct with m/z 300), cellobiosan (anhydrocellobiose; NH_4^+ adduct with m/z 342), and cellobiosylglycolaldehyde (NH_4^+ adduct with m/z 402). The structures of these ionized products shown in Figure 2 have been[1,33–35] or were confirmed by comparing their CAD mass spectra (Figures 3 and S1) with those of ionized authentic compounds that were synthesized or obtained from commercial sources. For example, the CAD mass spectrum of the ionized pyrolysis product of m/z 343 ($M+NH_4^+$) formed from the ^{13}C-labeled cellotriose (Figure 3b) matches the CAD mass spectrum of the authentic ionized unlabeled cellobiosan (m/z 342) (Figure 3a), indicating that the pyrolysis product is mostly ^{13}C-labeled cellobiosan. Among the unidentified ions are those with m/z values of 144, 162, 282, and 324, which likely correspond to ionized dehydration products of the identified products.

Fast pyrolysis of the ^{13}C-labeled cellotriose produced the same products as unlabeled cellotriose except that all ionized products with m/z value larger than 180 had retained the ^{13}C-label (i.e., their m/z values are one unit greater than those observed for the corresponding unlabeled compound; Figure 2). This label retention mimics the behavior of ^{13}C-labeled cellobiose studied previously,[1] corroborating the reducing-end unraveling mechanism proposed for cellobiose as a primary mechanism also for cellotriose. Therefore, the presence of the middle glucose unit in cellotriose does not change this fragmentation behavior.

HAGBC mechanism

The complete retention of the ^{13}C-label in most reaction products (Figure 2) of ^{13}C-labeled cellotriose demonstrates that the HAGBC mechanism cannot initiate the depolymerization sequence (Scheme 3), and therefore disproves the mechanism shown in Scheme 1b. This result is in agreement with the calculated free energy barriers as that for HAGBC (56.2 kcal/mol; Scheme 1b) is higher than that for reducing-end unraveling (49.2 kcal/mol for the rate limiting step;

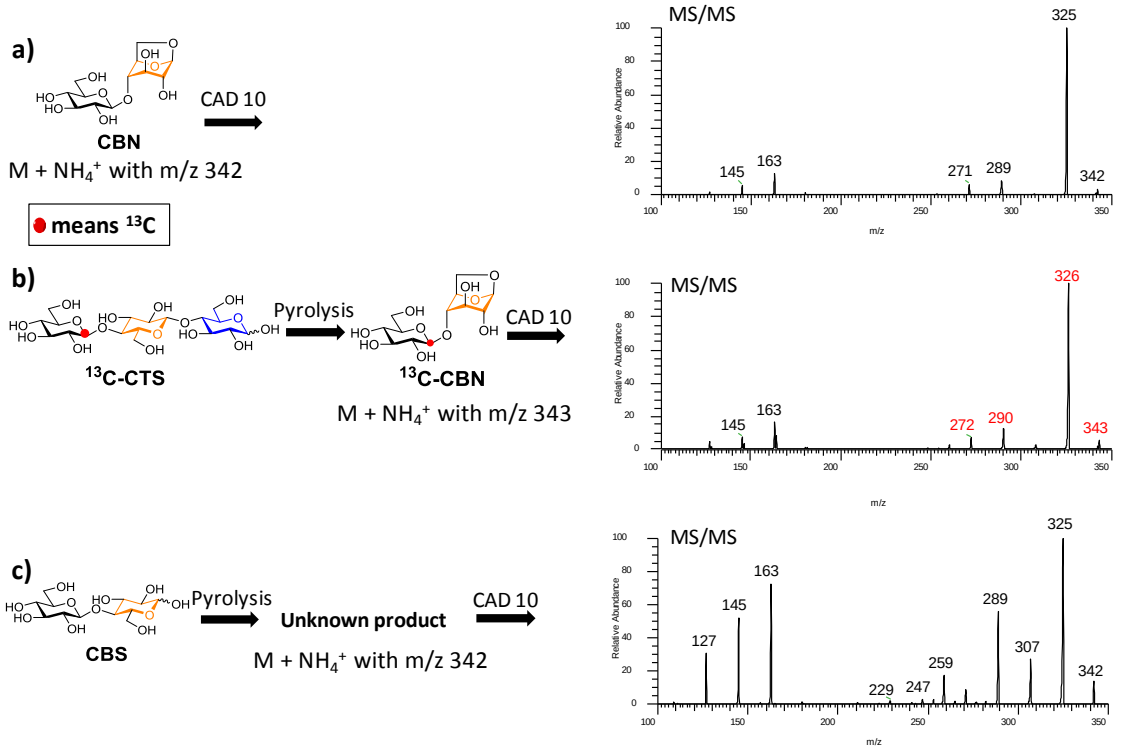


Figure 3. MS/MS CAD mass spectra (measured at a collision energy of 10 (arbitrary units)) of the NH_4^+ adducts of a) authentic cellobiosan (CBN) (m/z 342), b) anhydrocellobiose (^{13}C-CBN) produced upon fast pyrolysis of ^{13}C-labeled cellotriose (^{13}C-CTS) (fragment ions color-coded in red are shifted by one mass unit due to the ^{13}C label), and c) anhydrocellobiose produced upon fast pyrolysis of cellobiose (CBS).

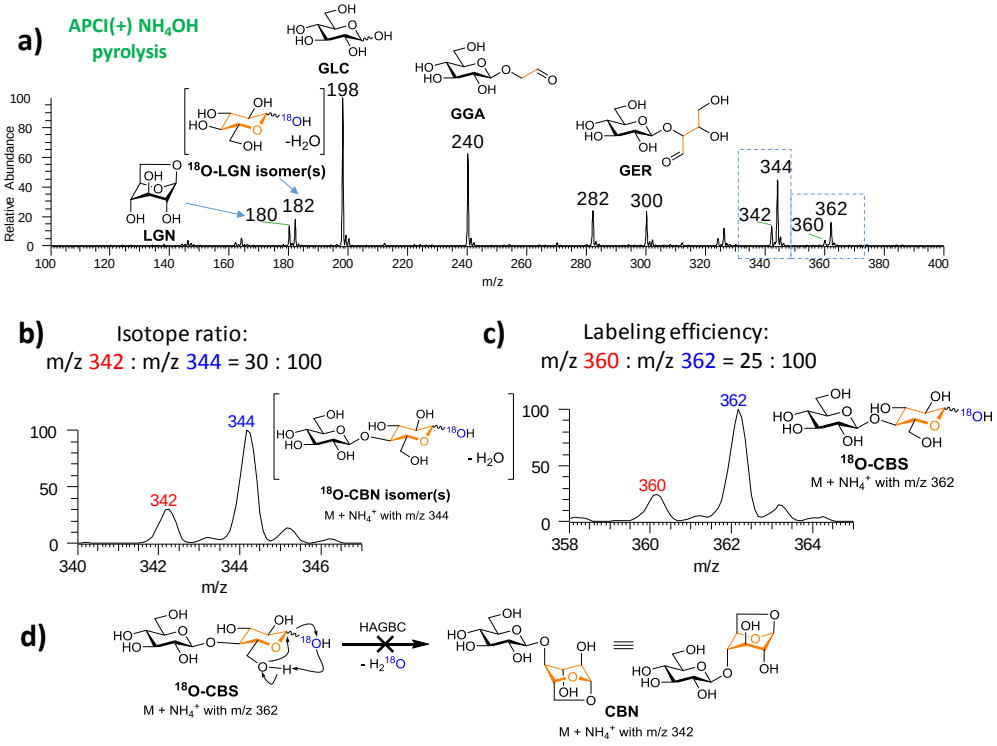
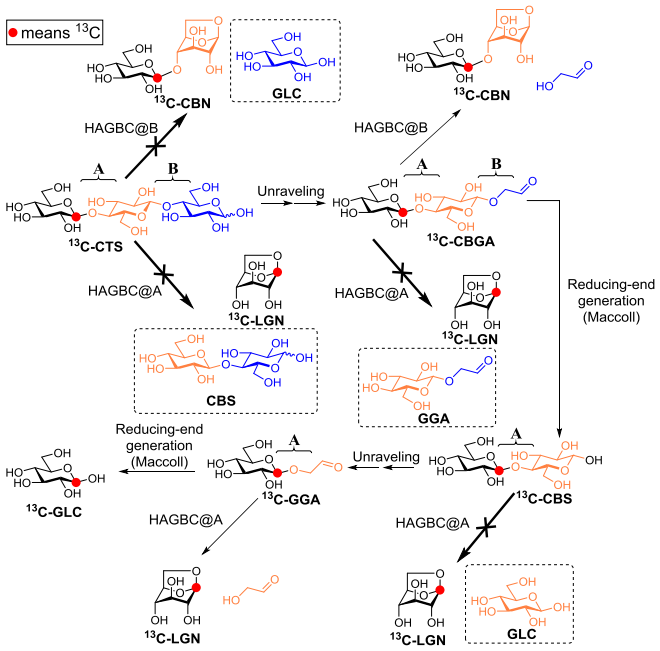


Figure 4. a) Mass spectrum showing NH_4^+ adducts of the initial fast pyrolysis products of cellobiose ^{18}O-labeled at the reducing end (^{18}O-CBS). b) Expanded area showing the ^{18}O-labeled cellobiosan isomers (^{18}O-CBN; NH_4^+ adduct with m/z 344) and unlabeled cellobiosan isomers (CBN; m/z 342; due to incomplete labeling of ^{18}O-cellobiose). c) Expanded area showing the ^{18}O-labeled cellobiose (^{18}O-CBS; NH_4^+ adduct with m/z 362) and unlabeled cellobiose (CBS; m/z 360; due to incomplete labeling of ^{18}O-cellobiose). d) Scheme showing the HAGBC dehydration reaction of ^{18}O-cellobiose (^{18}O-CBS) that gives rise to unlabeled cellobiosan (CBN).

Scheme 3. Three reaction mechanisms proposed for the generation of the major products formed upon fast pyrolysis of ^{13}C-labeled cellotriose. Glycosidic linkages are labeled using A and B. Dashed boxes indicate products that were not observed.



Abbreviations: cellobiosylglycolaldehyde (CBGA), cellobiosan (CBN; anhydrocellobiose), glucopyranosylglycolaldehyde (GGA), glucose (GLC), and levoglucosan (LGN; anhydroglucose. The formation of unlabeled levoglucosan – one of the major products – is not shown in the above mechanism. This additional mechanism is discussed in detail later.

Scheme 2). Furthermore, the label retention results demonstrate that the glycosidic linkage B dissociates prior to A (Scheme 3), in contrast to literature;[29] otherwise an unlabeled cellobiosylglycolaldehyde (NH_4^+ adduct with m/z 240) would have been formed. On the other hand, HAGBC can occur for the products of reducing-end unraveling mechanism (glycosides), which leads to the formation of 1,6-anhydrocellodextrins (Scheme 3 and Figure S2; also discussed in detail later in connection with Figure 7). It is worth noting that both labeled and unlabeled levoglucosan and isomers are produced upon fast pyrolysis of the ^{13}C-labeled cellotriose (NH_4^+ adducts with m/z 181 and 180, respectively); however, the mechanisms shown in Scheme 3 do not account for the formation of unlabeled levoglucosan and isomers. Another mechanism is discussed later to rationalize this observation.

HAGBC has been suggested in the literature to occur at the reducing end of carbohydrates via nucleophilic attack of the primary hydroxyl group at the anomeric carbon, which results in the elimination of the anomeric hydroxyl group in the form of water.[29,36] However, this hypothesis has not been confirmed or rejected based on experimental results. In order to explore this, MS/MS CAD experiments were carried out on the ionized product ($M+NH_4^+$; m/z 342) formed via water elimination upon fast pyrolysis of cellobiose. This ion corresponds to cellobiosan if HAGBC is the only operational mechanism. The CAD mass spectrum measured for this ionized pyrolysis product(s) (Figure 3c) does not match the CAD mass spectrum of the authentic ionized cellobiosan (Figure 3a), indicating that elimination of water from cellobiose mostly produces isomer(s) of cellobiosan that are formed via some other mechanism than HAGBC. This finding indicates that HAGBC does not occur at the reducing end of cellodextrins, in contrast to literature.[29,36]

To further examine this unexpected water-loss product of cellobiose, fast pyrolysis of ^{18}O-labeled cellobiose with the label at the hemiacetal hydroxyl group (Glc-Glc-^{18}OH) was examined. The mass spectrum measured for the ionized fast pyrolysis products is shown in Figure 4a. The ^{18}O-labeling efficiency was 80% as shown in Figure 4c; the synthesis method is detailed in the experimental section. The ionized anhydrocellobiose products (NH_4^+ adducts) have an isotope ratio of 30:100 for ions of m/z 342 and 344 (Figure 4b), similar to that observed for ^{18}O-cellobiose (25:100 for ions of m/z 360 and 362, Figure 4c), supporting the above conclusion that the first water elimination from cellobiose does not proceed through the HAGBC mechanism as this would result in the loss of the anomeric ^{18}O-label (Figure 4d). Therefore, all these results corroborate that HAGBC does not occur at the reducing end of cellodextrins (i.e., cellobiose pyrolysis produces levoglucosan but not cellobiosan; cellotriose pyrolysis produces levoglucosan and cellobiosan but not cellotriosan, etc.). More importantly, reducing-end unraveling has been identified as the predominant mechanism, if not the only mechanism, that occurs at the reducing end.

Maccoll elimination

One of the most abundant products formed upon fast pyrolysis of cellobiose, cellotriose, and other small oligomers is glucose (ion of m/z 198, Figure 2a).[35] This compound cannot be generated via the reducing-end unraveling mechanism as this mechanism does not generate reducing carbohydrates (Scheme 2). Further, the results shown in Figure 2b indicate that glucose cannot be formed via the HAGBC mechanism since HAGBC cannot initiate the depolymerization sequence, in contrast to literature;[18,29,36] otherwise, unlabeled glucose would have been formed for the ^{13}C-labeled cellotriose (Scheme 3). However, instead of unlabeled glucose (NH_4^+ adducts with m/z 198), exclusively ^{13}C-labeled glucose (NH_4^+ adducts with m/z 199) was formed (Figure 2b). Therefore, some other mechanism must be operational. The importance of this mechanism is not limited to glucose formation as reducing-end generation marks the end of one and the beginning of another depolymerization cycle (Scheme 3) and therefore is of significance for extending the fast pyrolysis mechanism delineated for cellobiose to cellotriose, oligomeric cellodextrins and polymeric cellulose.

Since water is produced as a byproduct during fast pyrolysis of cellulose, hydrolysis of the glycosidic bond by external water has been proposed to be the underlying mechanism for the formation of reducing carbohydrates.[16,18,22] Another mechanism proposed in the literature is Maccoll elimination that does not require external water and can cleave both glycosidic and aglyconic bonds.[18,26] However, no experimental evidence has confirmed or rejected these proposed mechanisms. In order to investigate the reducing-end generation mechanism, fast pyrolysis of ^{13}C- and ^{18}O-doubly-labeled cellobiose (Glc[1-^{13}C]-^{18}O-Glc-OH) was studied.

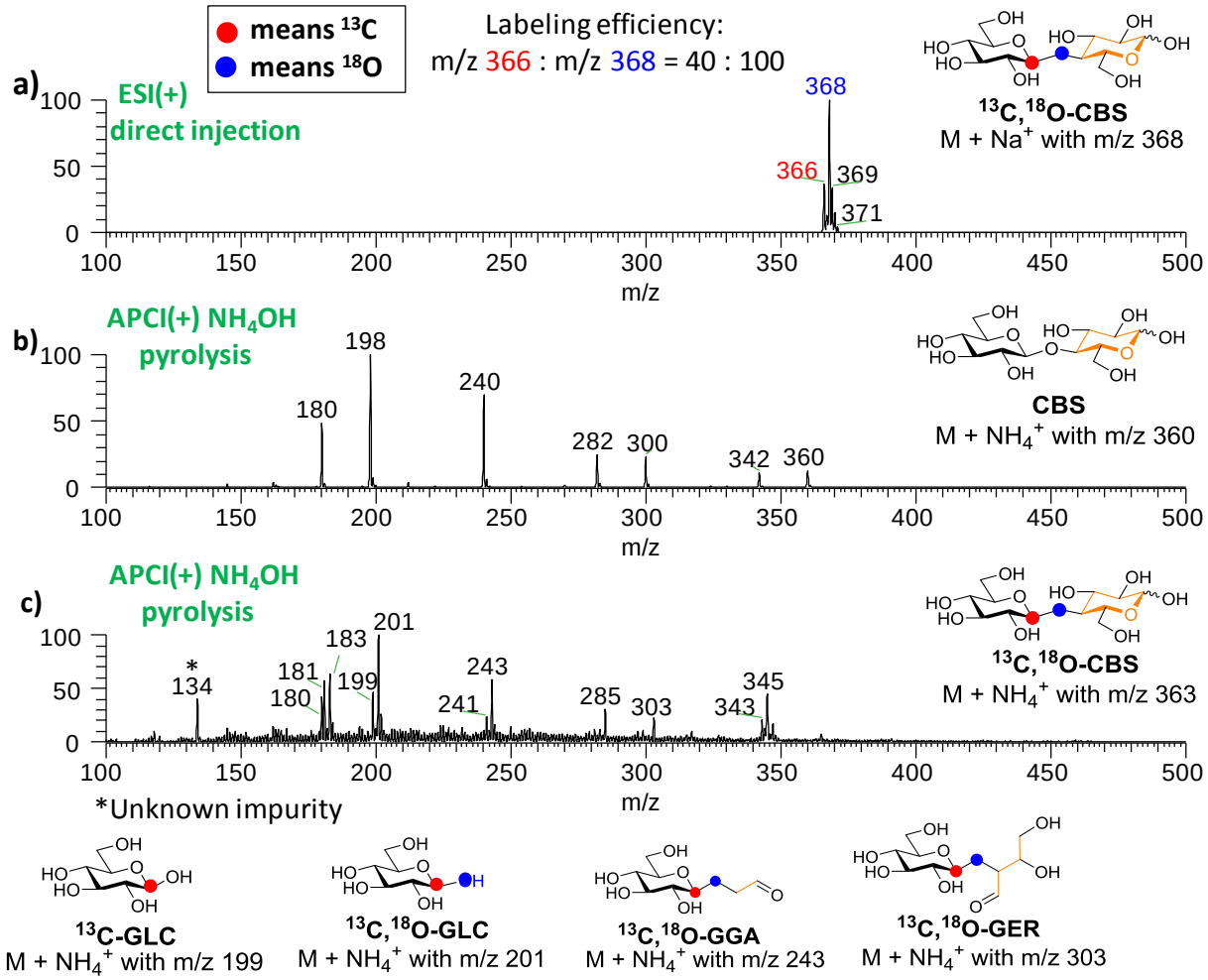


Figure 5. a) (+)ESI mass spectrum of synthetic, ^{13}C- and ^{18}O-doubly-labeled cellobiose ionized by sodium cation attachment (ions with m/z 366 and 368). APCI mass spectra of NH_4^+ adducts of the initial fast pyrolysis products of b) cellobiose and c) ^{13}C- and ^{18}O-doubly-labeled cellobiose. Abbreviations: cellobiose (CBS), glucopyranosylerythrose (GER), glucopyranosylglycolaldehyde (GGA), and glucose (GLC).

This compound contains the ^{18}O-label at the glycosidic oxygen between the two glucose units and the ^{13}C-label on the anomeric carbon of the nonreducing unit (Figure 5), which allows the identification of the cleavage site (glycosidic vs. aglyconic) and the determination of whether external water participates in this reaction.

As shown in Figure 5a, the ^{18}O-labeling efficiency of Glc[1-^{13}C]-^{18}O-Glc-OH was about 70% whereas the ^{13}C-labeling efficiency was 99% (due to the use of commercially available 99% [1-^{13}C]-glucose as the synthesis starting material; detailed synthesis procedure described in the experimental section). The mass spectrum of the NH_4^+ adducts of the fast pyrolysis products of this doubly labeled cellobiose (Figure 5c) shows ^{13}C- and ^{18}O-doubly-labeled and ^{13}C-singly-labeled glucose (NH_4^+ adducts with m/z 201 and 199, respectively) with a similar isotope abundance ratio (m/z 199 : m/z 201 = 40 : 100) as observed for the doubly-labeled cellobiose (Na^+ adducts with m/z 366 and 368 with an abundance ratio of 35 : 100; Figure 5a). Therefore, the formation of the glucose with only the ^{13}C-label (NH_4^+ adduct with m/z 199) was mostly due to the incomplete ^{18}O-labeling in Glc[1-^{13}C]-^{18}O-Glc-OH. The fact that the ^{18}O-labeled glycosidic oxygen was retained in the glucose molecule also containing the ^{13}C-label (NH_4^+ adduct with m/z 201) reveals that the glycosidic bond was not cleaved during the formation of glucose; instead, dissociation of the aglyconic bond must have occurred. This is the first unambiguous identification of the dissociation site between the glucose units during fast pyrolysis of cellobiose. Furthermore, the complete retention of ^{18}O-label rules out the participation of external water in the fast pyrolysis reactions of cellobiose (Figure 6a). This is further confirmed upon examination of the pyrolysis products of oligomeric 1,6-anhydrocellodextrins (e.g., cellobiosan, cellotriosan, and cellotetraosan) that is discussed later. Therefore, Maccoll elimination at the aglyconic bond is identified as the underlying mechanism for reducing-end generation (mechanism shown in Scheme 1a). It is worth noting that the free energy barrier calculated for the aglyconic Maccoll elimination (65.9 kcal/mol; Scheme 1a) is higher than that of HAGBC (56.2 kcal/mol; Scheme 1b). This discrepancy between the experimental results and calculations may be partially explained by gas-phase calculations not accurately representing the actual reaction phase of fast

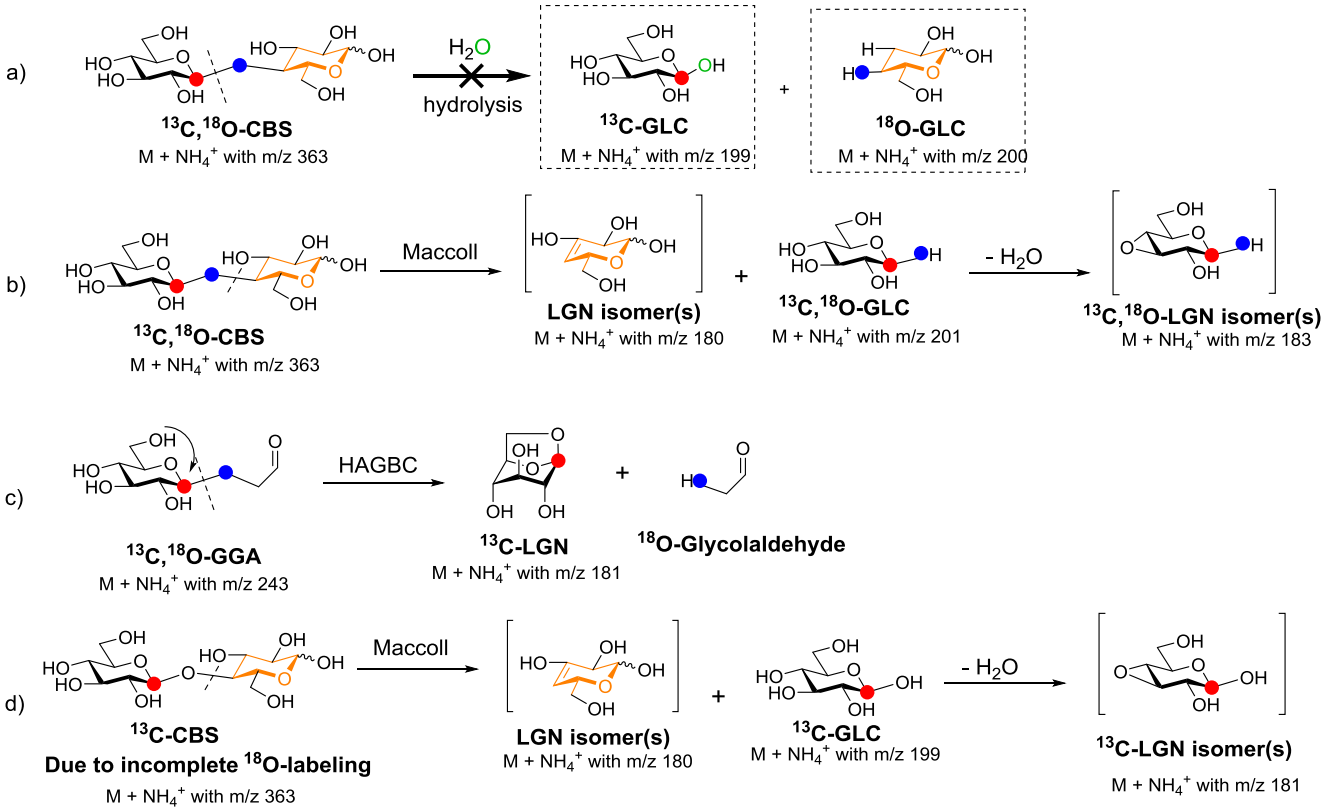


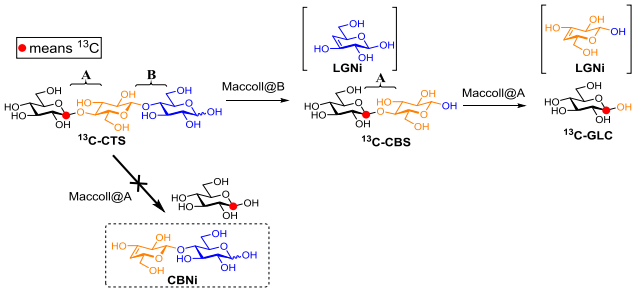
Figure 6. a) Hydrolysis (that does not occur) and b-d) reactions that may give rise to anhydroglucoses (NH_4^+ adducts with m/z 180, 181, and 183). Abbreviations: cellobiose (CBS), glucopyranosylglycolaldehyde (GGA), cellobiosan (CBN), and glucose (GLC), and levoglucosan (LGN). Structures in brackets represent one possible levoglucosan isomer.

pyrolysis. For example, it is possible that the hydroxymethylene group needed for the HAGBC mechanism is involved in hydrogen-bonding with surrounding molecules. The Maccoll mechanism does not involve hydroxyl groups.

The bundle of anhydroglucose (levoglucosan isomers) peaks (NH_4^+ adducts with m/z 180, 181, and 183 in Figure 5c) observed for the labeled cellobioses can be explained as follows: 1) NH_4^+ adduct with m/z 180 must have been generated from the reducing-end glucose unit since it does not contain either ^{13}C- or ^{18}O label. This can be explained by aglyconic Maccoll elimination in the intact, singly and doubly labeled cellobioses (Figures 6b and 6d). 2) The NH_4^+ adduct with m/z 183, containing both the ^{13}C- and ^{18}O-labels, was formed by an aglyconic Maccoll elimination in the doubly-labeled cellobiose to generate doubly-labeled glucose (NH_4^+ adduct with m/z 201), followed by dehydration with retention of the anomeric hydroxyl group (Figure 6b). 3) The NH_4^+ adduct with m/z 181, containing only the ^{13}C-label, can be produced upon HAGBC reactions of ^{13}C-labeled intermediates (for example as shown in Figure 6c) formed upon reducing-end unraveling (e.g., glucopyranosyl-erythrose and glucopyranosylglycolaldehyde; NH_4^+ adducts with m/z 303 and 243, respectively, in Figure 5c). This NH_4^+ adduct with m/z 181 can also be formed as shown in Figure 6d.

It was mentioned above that both labeled and unlabeled levoglucosan (and isomers; NH_4^+ adducts with m/z 181 and 180, respectively) were formed upon pyrolysis of ^{13}C-labeled cellotriose (Figure 2b). However, the mechanisms presented in Scheme 3 only show the formation of labeled levoglucosan (or isomers) and therefore is not complete. The formation of the unlabeled levoglucosan (or isomers) can be explained by Maccoll elimination occurring on the intact cellobiose. Initiation of the reaction by Maccoll elimination at the aglyconic bond B generates an unlabeled levoglucosan isomer from the reducing-end glucose unit; the resulting cellobiose can then undergo another Maccoll elimination to form the ^{13}C-labeled glucose and another unlabeled levoglucosan isomer (Scheme 4). It is worth noting that Maccoll occurs at the aglyconic bond B prior to bond A; otherwise, an unlabeled cellobiosan *isomer* (NH_4^+ adduct

Scheme 4. Proposed mechanisms for the generation of unlabeled levoglucosan isomers upon fast pyrolysis of ^{13}C-labeled cellotriose. Glycosidic linkages are labeled using A and B. Dashed box indicates a product that was not observed. Structures in brackets represent one possible levoglucosan isomer.



Abbreviations: cellotriose (CTS), cellobiose (CBS), glucose (GLC), cellobiosan isomers (CBNi) and levoglucosan isomers (LGNi).

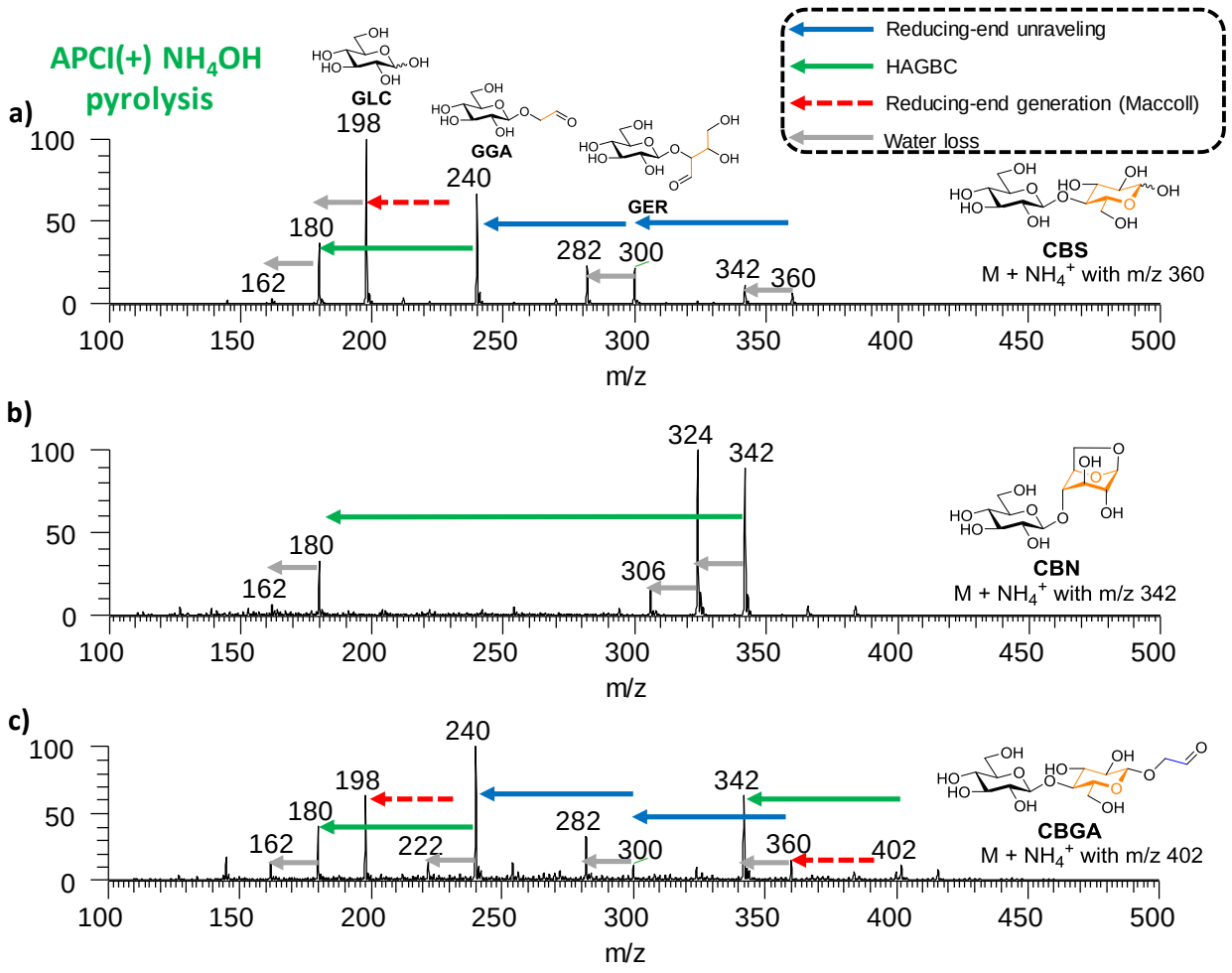


Figure 7. Mass spectra of NH_4^+ adducts of the fast pyrolysis products of a) cellobiose, b) cellobiosan, and c) cellobiosylglycolaldehyde. Abbreviations: cellobiose (CBS), cellobiosan (CBN), cellobiosylglycolaldehyde (CBGA) glucopyranosylerythrose (GER), glucopyranosylglycolaldehyde (GGA), and glucose (GLC). The pyrolysis mass spectra of cellobiose and cellobiosan were remeasured here and agree with previously published data.[1,34]

with m/z 342) would have been formed. Instead, only cellobiosan, *not* an isomer, was formed with complete retention of the ^{13}C-label (NH_4^+ adduct with m/z 343), confirmed by MS/MS CAD experiments (Figures 3a and 3b). Therefore, similar to HAGBC discussed above, the bond (glycosidic for HAGBC and aglyconic for Maccoll) closest to the reducing-end glucose unit dissociates prior to the other one. Together, these results indicate that fast pyrolysis of oligomeric cellodextrins follows highly ordered mechanisms.

Fast pyrolysis of various reaction intermediates

Having established HAGBC, reducing-end unraveling, and Maccoll elimination as important mechanisms in the depolymerization reactions, it was of interest to examine the pyrolysis reactions of specific fast pyrolysis intermediates in order to obtain more information on the entire reaction network. First, fast pyrolysis of 1,6-anhydrocellodextrins was inspected and compared with pyrolysis of their reducing counterparts – cellodextrins. For example, fast pyrolysis of cellobiosan (a 1,6-anhydrocellodextrin; Figure 7b), as opposed to cellobiose (the reducing counterpart; Figure 7a), did not produce reducing carbohydrates (glucose in this case) via Maccoll elimination nor the intermediates commonly observed for reducing-end unraveling mechanism, such as glucopyranosylerythrose (NH_4^+ adduct with m/z 300) and glucopyranosylglycolaldehyde (NH_4^+ adduct with m/z 240). Instead, water loss (NH_4^+ adduct with m/z 324) and formation of levoglucosan (NH_4^+ adduct with m/z 180) via HAGBC mechanism were observed (the levoglucosan structure was confirmed by MS/MS CAD experiments, Figure S3). Similarly, it has been demonstrated earlier that pyrolysis of cellotriosan produces cellobiosan.[34] Further, it was found here that pyrolysis of cellotetraosan produces cellotriosan (MS/MS CAD data for authentic cellotriosan and cellotriosan produced from fast pyrolysis of cellotetraosan are shown in Figure S4). Hence, none of these larger 1,6-anhydrocellodextrins (up to cellotetraosan) produced reducing carbohydrates upon fast pyrolysis (mass spectrum of the ionized fast pyrolysis products of cellotetraosan is shown in Figure S5). This finding demonstrates that HAGBC is

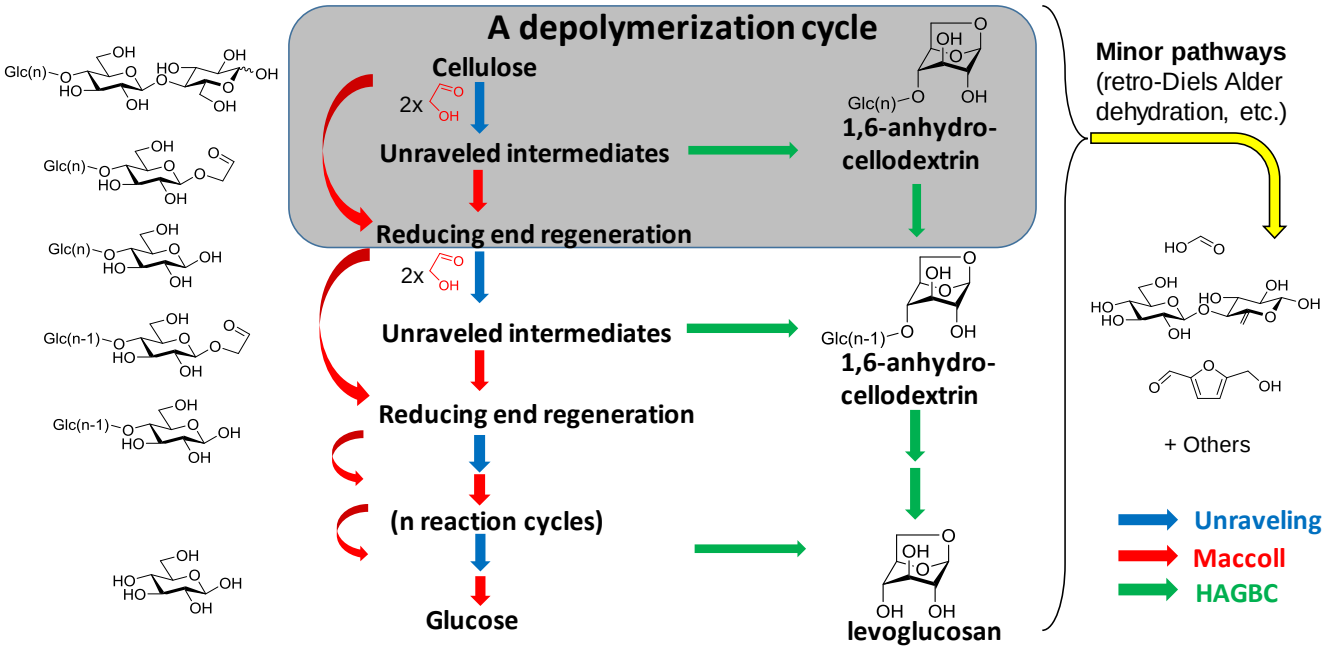


Figure 8. The network of proposed major reactions for cellulose fast pyrolysis

favored over Maccoll during the pyrolysis of 1,6-anhydrocellodextrins, and that no unraveling occurs in the absence of a reducing end. Furthermore, the lack of formation of reducing carbohydrates (e.g., glucose and cellobiose; NH_4^+ adducts with m/z198 and 360, respectively) upon fast pyrolysis of cellobiosan (Figure 7b) confirms that no hydrolysis involving external water took place. More conclusively, even when an aqueous solution of cellobiosan was loaded into a hollow quartz tube for fast pyrolysis, no hydrolysis products (i.e., glucose and cellobiose) were observed (Figure S6). These results clearly exclude the participation of external water in the fast pyrolysis reactions of cellobiosan, and the same is expected to be true for larger 1,6-anhydrocellodextrins.

Second, fast pyrolysis products of cellobiosylglycolaldehyde – an intermediate formed upon reducing-end unraveling – was compared with the fast pyrolysis products of cellobiose and cellobiosan. In contrast to pyrolysis of cellobiosan (Figure 7b), pyrolysis of cellobiosylglycolaldehyde resulted in a similar product distribution (Figure 7c) as observed for cellobiose (Figure 7a). A detailed inspection of the differences and similarities in these fast pyrolysis product distributions lead to three important discoveries: 1) fast pyrolysis of cellobiosylglycolaldehyde produces cellobiosan (Figure S2), while fast pyrolysis of cellobiose does not (Figure 3a and 3c), further supporting the conclusion made above that HAGBC only occurs for glycosides and does not occur at the reducing end of cellodextrins; 2) fast pyrolysis of cellobiosylglycolaldehyde produces both unraveling intermediates, such as glucopyranosylerythrose (NH_4^+ adduct with m/z 300) and glucopyranosylglycolaldehyde (NH_4^+ adduct with m/z 240), as well as cellobiosan (NH_4^+ adduct with m/z 342) and levoglucosan (NH_4^+ adduct with m/z 180), while fast pyrolysis of cellobiosan only produces cellobiosan and levoglucosan, suggesting that HAGBC can proceed subsequent to reducing-end unraveling, but unraveling does not occur once the HAGBC reaction sequence is initiated; and 3) fast pyrolysis of cellobiosylglycolaldehyde produces reducing carbohydrates (i.e., cellobiose and glucose; NH_4^+ adducts with m/z 360 and 198, respectively), while fast pyrolysis of cellobiosan and other, larger 1,6-anhydrocellodextrins does not, indicating that intermediates generated upon reducing-end unraveling can undergo reducing-end generation via aglyconic Maccoll elimination, while 1,6-anhydrocellodextrins cannot (instead, they predominantly undergo HAGBC and water loss).

A general reaction network for fast pyrolysis of cellulose

Based on above experimental results, a general reaction network for fast pyrolysis of cellulose is proposed in Figure 8. Cellulose pyrolysis is likely initiated by either Maccoll elimination to cleave off an anhydroglucose (red arrows in Figure 8), or by reducing-end unraveling to produce unraveled intermediates (blue arrows in Figure 8). These intermediates can either undergo HAGBC to generate 1,6-anhydrocellodextrin (green arrows in Figure 8) or Maccoll elimination to regenerate a reducing end. However, once the HAGBC mechanism is "selected", the 1,6-anhydrocellodextrins will continue to undergo reactions via the HAGBC mechanism and eventually produce levoglucosan. On the other hand, reducing end is generated when Maccoll occurs for an intermediate formed upon the reducing-end unraveling, which will start a new depolymerization cycle via reducing-end unraveling and reducing-end generation or is diverted to the HAGBC pathway. This network is in a good agreement with the fact that the abundance of glucose formed upon fast pyrolysis deceases as the number of monomer units increases[34] in the pyrolyzed compound since the more monomer units the pyrolyzed carbohydrate contains, the greater is the probability that the HAGBC mechanism gets "selected" and therefore outcompetes the reducing-end unraveling mechanism. The same rationale explains the

high yield of levoglucosan observed for cellulose fast pyrolysis: since a greater number of monomers leads to a greater probability for the HAGBC pathway being "selected", more 1,6-anhydrocellodextrins are generated that eventually produce levoglucosan as the end product. Therefore, these results allow the prediction of product distributions based on the degree of polymerization of the pyrolyzed compounds. It should be noted that additional reaction mechanisms might be operational for long-chain polymers, and that an increased crystallinity may also play a role in the depolymerization reactions during fast pyrolysis. However, oligomeric cellodextrins should be governed by the above highly ordered mechanisms.

CONCLUSIONS

Many reaction mechanisms have been proposed in literature in an effort to better understand the depolymerization reactions of cellulose upon fast pyrolysis. The experimental approach described here enabled the identification of the three most important reaction mechanisms: reducing-end unraveling, hydroxymethylene-assisted bond cleavage (HAGBC), and aglyconic bond cleavage via Maccoll elimination. The process is likely initiated either by reducing-end unraveling or Maccoll elimination that occurs next to the reducing end glucose unit. Maccoll elimination generates a new reducing end, which starts a new depolymerization cycle via reducing-end unraveling. On the other hand, HAGBC only occurs after reducing-end unraveling. HAGBC and dehydration are the dominant reactions for degradation of 1,6-anhydrocellodextrins. The observation of highly organized reaction pathways suggests that radical reactions are minor or nonexistent. Several of the discoveries reported here are in contrary to or are subjects of debate in current literature. These discoveries include the following: 1) hydrolysis does not take place; 2) HAGBC reaction does not occur at the reducing end of cellodextrins; and 3) glucose and smaller cellodextrins are formed by aglyconic bond cleavages via the Maccoll mechanism. Altogether, these results provide important insights into fast pyrolysis mechanisms of cellulose, which can guide computational studies aimed at improved understanding of the thermodynamics of cellulose fast pyrolysis at the molecular level.

EXPERIMENTAL SECTION

Reagents

Aqueous ammonium hydroxide as well as unlabeled cellotriose, cellobiose, cellobiosan, and levoglucosan were purchased from Sigma Aldrich (St. Louise, MO, USA). Cellotetraosan and cellotriosan were purchased from Synthose (Concord, Ontario, Canada). All chemicals were used as received.

Pyrolysis tandem mass spectrometry (Py-MS/MS) setup

The Py-MS/MS instrument has been detailed previously.[33] Briefly, a pyroprobe (CDS analytical) made from a flat platinum ribbon (a platinum coil along with a quartz tube were used when loading liquid samples) was positioned inside the IonMax box of a Thermo LQIT-Orbitrap high-resolution mass spectrometer. The IonMax box was filled with dry nitrogen. Commercial and synthesized samples were loaded onto the pyrobrobe that was then heated up to 600 °C in 1 ms and held at 600 °C for one second. The evaporating pyrolysis products (M) were instantly quenched in the 100 °C nitrogen atmosphere and ionized via atmospheric pressure chemical ionization (APCI) by using ammonium hydroxide as a dopant. The resulting ammonium adducts ($M+NH_4^+$) were immediately transferred into the LQIT/orbitrap mass spectrometer for analysis. The overall pyrolysis and analysis time was less than 125 ms.[1,33] Mass spectra were collected just before, during, and immediately after pyrolysis. Background ions were removed from the pyrolysis mass spectra by subtracting the mass spectra measured before and after pyrolysis from the mass spectra measured during pyrolysis. All reported mass spectra are the average of at least 10 mass spectra. The MS/MS CAD experiments were conducted in the linear quadrupole ion trap by isolating an ion of interest with an isolation width of 2 m/z values and kinetically excited for 30 ms. Nominal collision energy of 15 (arbitrary units) was used for CAD of ionized cellobiosan and isomers ($[M + NH_4]^+$ with m/z 342) and for ionized cellotriosan and isomers ($[M + NH_4]^+$ with m/z 504) while 20 was used for ionized cellobiosylglycolaldehyde. The reported MS/MS CAD mass spectra are an average of at least 10 mass spectra.

As no ion signal was obtained after CAD of NH_4^+ adducts of levoglucosan and isomers, these compounds were also ionized by deprotonation in negative ion mode with ammonium hydroxide dopant, which resulted in $[M - H]^-$ anions with m/z 161. These ions were subjected to CAD at a nominal collision energy of 18 (arbitrary units) to obtain MS/MS CAD mass spectra.

Synthesis of ^{13}C-labeled cellotriose (T10) (Scheme S1)

Synthesis of phenyl 2,3,4,6–tetra–O–acetyl–1-thio-β–D–glucopyranoside (T2). This procedure was adapted from literature.[37] Boron trifluoride diethyl etherate (9 ml, 75 mmol) was added into a solution of **T1** (10 g, 25 mmol) and thiophenol (3.3 mL ml, 30.6 mmol) in dichloromethane (DCM) (60 mL) at 0 °C. The reaction mixture was stirred at room temperature for 20 h and diluted with DCM (60 mL). The resulting solution was washed with a saturated solution of $NaHCO_3$ in water, dried (Na_2SO_4), filtered, and the filtrate was concentrated *in vacuo*. The resulting residue was purified by recrystallization in ethanol to afford **T2** (6.8 g, 51 %) as white needle-shaped crystals. ^{1}H NMR (800 MHz, Chloroform-*d*) δ 7.51 – 7.45 (m, 2H), 7.34 – 7.28 (m, 3H), 5.21 (t, *J* = 9.4 Hz, 1H), 5.03 (dd, *J* = 10.1, 9.5 Hz, 1H), 4.97 (dd, *J* = 10.0, 9.3 Hz, 1H), 4.70 (d, *J* = 10.1 Hz, 1H), 4.21 (dd, *J* = 12.2, 5.2 Hz, 1H), 4.17 (dd, *J* = 12.2, 2.4 Hz, 1H), 3.72 (ddd, *J* = 10.1, 5.3, 2.5 Hz, 1H), 2.08 (s, 3H), 2.07 (s, 3H), 2.01 (s, 3H), 1.98 (s, 3H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-*d*) δ 170.7, 170.3, 169.5, 169.4, 133.2, 131.8, 129.1, 128.5, 85.8, 75.9, 74.1, 70.1, 68.3, 62.3, 20.9, 20.8, 20.7, 20.7. The NMR spectral data are in accordance with literature.[38] ESI HRMS: m/z calcd for $C_{20}H_{24}NaO_9S$ $[M + Na]^+$ 463.1033, found 463.1031.

Synthesis of phenyl 4,6-benzylidene-1-thio-β–D–glucopyranoside (T3). This procedure was adapted from literature.[37] Thioglucoside **T2** (6.8 g, 15.5 mmol) was dissolved in anhydrous methanol (100 mL). Catalytic amount of sodium methoxide was added into the reaction and stirred overnight at room temperature. The reaction was then neutralized with Amberlyst 15 which was removed

thereafter by filtration. The product residue was collected *in vacuo* and dissolved again in dimethylformamide (DMF) (20 mL), into which benzaldehyde dimethylacetal (mL, 23.1 mmol) and camphorsulfonic acid (430 mg, 1.85 mmol) were added and the reaction was allowed to proceed at 50 °C overnight. The reaction was quenched with Et_3N, diluted with DCM (100 mL) and washed with a saturated water solution of $NaHCO_3$ followed by saturated brine, and then dried over Na_2SO_4, filtered, and concentrated *in vacuo*. Due to limited solubility, the resulting residue was used in the next step without purification.

Synthesis of phenyl 2,3-di-O-acetyl-4,6-O-benzylidene-1-thio-β–D–glucopyranoside (T4). The crude diol **T3** from the previous synthesis was dissolved in anhydrous pyridine (10 mL) into which acetic anhydride (10 mL) was added. The reaction was stirred overnight at room temperature and quenched with methanol. Solvent was removed at reduced pressure and the residue was washed with 4M aqueous HCl, saturated aqueous $NaHCO_3$, and brine, and dried over anhydrous Na_2SO_4. The residue was purified by silica gel column chromatography (hexane:ethyl acetate = 4:1) to afford **T4** (5.1 g, 74% three steps) as a white foam. ^{1}H NMR (800 MHz, Chloroform-*d*) δ 7.51 – 7.46 (m, 2H), 7.44 – 7.40 (m, 2H), 7.35 (ddd, *J* = 8.1, 6.3, 3.7 Hz, 6H), 5.50 (s, 1H), 5.34 (t, *J* = 9.3 Hz, 1H), 5.01 (dd, *J* = 10.1, 8.9 Hz, 1H), 4.81 (d, *J* = 10.0 Hz, 1H), 4.39 (dd, *J* = 10.6, 5.0 Hz, 1H), 3.79 (t, *J* = 10.3 Hz, 1H), 3.67 (t, *J* = 9.6 Hz, 1H), 3.58 (td, *J* = 9.7, 5.0 Hz, 1H), 2.10 (s, 3H), 2.03 (s, 3H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-*d*) δ 170.1, 169.5, 136.7, 133.0, 131.7, 129.2, 129.1, 128.4, 128.3, 126.2, 101.5, 86.6, 78.1, 77.2, 77.1, 76.9, 72.9, 70.8, 70.7, 68.5, 20.8, 20.8. The NMR spectral data are in accordance with literature.[39] ESI HRMS: m/z calcd for $C_{23}H_{24}NaO_7S$ [M + Na]$^+$ 467.1135, found 467.1131.

Synthesis of benzyl 2,3-di-O-acetyl-4,6-O-benzylidene-β–D–glucopyranosyl-(1→4)-2,3-di-O-acetyl-6-O-benzyl-β-D-glucopyranoside (T6). A solution of glycosyl donor **T4** (1.5 g, 3.4 mmol) and acceptor **T5** (synthesized according to literature[40]) (1.8 g, 4.8 mmol) in anhydrous DCM (10 mL) was stirred with activated molecular sieves (4 Å) under argon for 30 min. N-iodosuccinimide (0.765 g, 3.4 mmol) was then added and stirring was continued for 30 min more. After addition of a catalytic amount of silver triflate (AgOTf, 0.009 g, 0.034 mmol), the color of the reaction mixture turned yellow brown. The reaction was quenched by addition of triethyl amine (Et_3N) and stirred for an additional 20 min. This reaction mixture was diluted with DCM and washed with sodium thiosulfate to remove iodine. The resulting residue was collected *in vacuo* and purified by silica gel column chromatography (hexane:ethyl acetate = 3:1) to afford **T6** (1.8 g, 70%) as colorless syrup. ^{1}H NMR (800 MHz, Chloroform-*d*) δ 7.47 – 7.41 (m, 6H), 7.39 – 7.34 (m, 7H), 7.32 – 7.30 (m, 2H), 5.47 (s, 1H), 5.19 (t, *J* = 9.5 Hz, 1H), 5.14 (t, *J* = 9.4 Hz, 1H), 5.03 (dd, *J* = 9.7, 8.0 Hz, 1H), 4.91 (d, *J* = 12.3 Hz, 1H), 4.88 (dd, *J* = 9.3, 8.0 Hz, 1H), 4.80 (d, *J* = 12.0 Hz, 1H), 4.64 (d, *J* = 12.3 Hz, 1H), 4.60 – 4.50 (m, 2H), 4.35 (dd, *J* = 10.4, 5.0 Hz, 1H), 4.13 (q, *J* = 7.2 Hz, 1H), 4.00 (t, *J* = 9.5 Hz, 1H), 3.78 (qd, *J* = 11.1, 2.8 Hz, 2H), 3.70 (t, *J* = 10.2 Hz, 1H), 3.61 (t, *J* = 9.6 Hz, 1H), 3.45 (ddd, *J* = 9.9, 3.4, 1.9 Hz, 1H), 3.32 (td, *J* = 9.7, 5.0 Hz, 1H), 2.03 (s, 3H), 2.02 (s, 3H), 2.02 (s, 3H), 1.99 (s, 3H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-*d*) δ 170.0, 169.6, 169.5, 169.0, 137.6, 136.9, 136.7, 129.1, 128.6, 128.3, 128.2, 128.2, 128.1, 128.0, 127.8, 127.6, 126.1, 105.6, 101.4, 101.3, 100.7, 99.3, 78.0, 75.5, 74.6, 73.6, 73.1, 72.5, 71.9, 71.5, 70.5, 68.4, 67.1, 65.9, 60.3, 20.9, 20.7, 20.6, 20.6. ESI HRMS: m/z calcd for $C_{41}H_{46}NaO_{15}$ [M + Na]$^+$ 801.2729, found 801.2733.

Synthesis of benzyl 1-ol-2,3-di-O-acetyl-6-O-benzyl-β–D–glucopyranosyl-(1→4)-2,3-di-O-acetyl-6-O-benzyl-β-D-glucopyranoside (T7). Triethylsilane (0.835 g, 7.2 mmol) and trifluoromethanesulfonic acid (TfOH, 1.08 g, 7.2 mmol) were sequentially added into a cooled (-78 °C) solution of **T6** (1.8 g, 2.4 mmol) in anhydrous DCM. The reaction mixture was stirred at -78 °C for 1 h and then quenched with methanol and triethyl amine. The resulting mixture was washed with saturated aqueous $NaHCO_3$ and water, and dried over Na_2SO_4. The filtrate was concentrated *in vacuo* and the resulting residue was purified by silica gel column chromatography (hexane:ethyl acetate:dichloromethane = 3:1:1) to afford compound **T7** (1.68 g, 90%) as colorless syrup. ^{1}H NMR (800 MHz, Chloroform-*d*) δ 7.46 – 7.22 (m, 15H), 5.10 (t, *J* = 9.4 Hz, 1H), 5.03 (dd, *J* = 9.8, 7.9 Hz, 1H), 4.94 – 4.88 (m, 2H), 4.77 (dd, *J* = 11.1, 6.0 Hz, 2H), 4.63 (d, *J* = 12.4 Hz, 1H), 4.53 (t, *J* = 11.9 Hz, 3H), 4.50 (d, *J* = 8.0 Hz, 1H), 4.47 (d, *J* = 8.0 Hz, 1H), 3.96 (t, *J* = 9.5 Hz, 1H), 3.80 – 3.74 (m, 2H), 3.72 (td, *J* = 10.3, 9.7, 3.8 Hz, 2H), 3.43 (dt, *J* = 10.0, 2.5 Hz, 1H), 3.32 (dt, *J* = 9.2, 4.4 Hz, 1H), 3.10 (d, *J* = 3.7 Hz, 1H), 2.06 (s, 3H), 2.01 (s, 3H), 1.96 (s, 3H), 1.96 (s, 3H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-*d*) δ 171.2, 170.2, 169.7, 169.1, 137.7, 137.4, 137.0, 128.6, 128.6, 128.4, 128.1, 128.1, 128.0, 127.9, 127.7, 100.1, 75.8, 75.0, 74.8, 73.7, 73.6, 72.8, 71.8, 71.6, 70.7, 70.2, 69.8, 67.3, 20.9, 20.7, 20.7, 20.7. ESI HRMS: m/z calcd for $C_{41}H_{48}NaO_{15}$ [M + Na]$^+$ 803.2885, found 803.2876.

Synthesis of benzyl 2,3,4,6-tetra-O-acetyl-β-D-[1-^{13}C]glucopyranosyl-(1→4)-1-2,3-di-O-acetyl-6-O-benzyl-β–D–glucopyranosyl-(1→4)-2,3-di-O-acetyl-6-O-benzyl-β-D-glucopyranoside (T8). A mixture of glycosyl acceptor **T7** (1.68 g, 2.16 mmol) and donor peracetyl-glucopyranosyl-trichloroacetimidate (1.169 g, 2.376 mmol) was stirred in 20 mL anhydrous diethyl ether/dichloromethane (1:1) with molecular sieves (4Å) at room temperature under argon for 30 min. The mixture was then cooled to –20 °C and trimethylsilyl trifluoromethanesulfonate (TMSOTf, 12 µL, 0.068 mmol) was added into it. The reaction was slowly warmed to room temperature over 1 h. TLC analysis showed complete conversion of starting material to a major product. The reaction was quenched by the addition of 0.1 mL triethylamine and filtered. The filtrate was concentrated in vacuum and purified by silica gel chromatography (hexane: ethyl acetate = 4:1) to yield protected trisaccharide T8 (1.56 g, 65%) as colorless syrup. 1H NMR (800 MHz, Chloroform-d) δ 7.45 – 7.30 (m, 15H), 5.08 (t, J = 9.5 Hz, 1H), 5.03 – 4.98 (m, 1H), 4.93 (q, J = 9.7, 9.3 Hz, 2H), 4.88 (d, J = 12.3 Hz, 1H), 4.80 (q, J = 8.3, 7.6 Hz, 1H), 4.75 – 4.72 (m, 1H), 4.65 (d, J = 11.7 Hz, 1H), 4.60 (d, J = 12.4 Hz, 1H), 4.54 – 4.44 (m, 3H), 4.39 – 4.34 (m, 1H), 4.36 – 4.29 (m, 1H), 4.28 (dt, J = 16.6, 8.4 Hz, 1H), 4.10 (q, J = 7.2 Hz, 1H), 3.99 – 3.89 (m, 2H), 3.90 – 3.81 (m, 1H), 3.74 (d, J = 2.5 Hz, 1H), 3.72 – 3.66 (m, 1H), 3.65 (dt, J = 10.8, 3.3 Hz, 1H), 3.41 (dt, J = 9.9, 2.6 Hz, 1H), 3.30 (d, J = 9.4 Hz, 1H), 3.26 – 3.21 (m, 1H), 3.10 – 3.05 (m, 1H), 2.06 (s, 3H), 2.01 (s, 3H), 2.00 (s, 3H), 1.98 (s, 4H), 1.97 (s, 3H), 1.96 (s, 3H), 1.96 (s, 3H), 1.93 (s, 3H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-d) δ 170.5, 170.3, 170.2, 170.1, 169.6, 169.3, 169.0, 168.7, 137.6, 137.1, 137.0, 128.8, 128.7, 128.6, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 127.9, 127.6,

127.6, 127.6, 127.4, 99.9, 99.8, 74.7, 74.6, 74.4, 74.0, 73.6, 73.3, 73.0, 72.9, 72.8, 72.6, 71.8, 71.6, 71.5, 71.4, 71.3, 70.7, 67.8, 67.2, 61.5, 60.4, 20.7, 20.65, 20.60, 20.57, 20.54, 20.52, 20.45. ESI HRMS: m/z calcd for $C_{54}{}^{13}CH_{66}NaO_{24}$ $[M + Na]^+$ 1134.3870, found 1134.3888.

Synthesis of benzyl β-D-[1-^{13}C]glucopyranosyl-(1→4)-6-O-benzyl-β–D–glucopyranosyl-(1→4)-6-O-benzyl-β-D-glucopyranoside (T9). Global deacetylation of trisaccharide **T8** was achieved under Zempler conditions.[41] Briefly, $NaOCH_3$ was added into a solution of **T8** (0.4 g, 0.36 mmol) in 10 mL anhydrous methanol to adjust pH to 9. The reaction was allowed to continue overnight. The reaction was then quenched by Amberlyst 15 acid resin, and the filtrate was concentrated *in vacuo*. This deacetylated intermediate **T9** was characterized by NMR before the removal of benzyl acetal protection to avoid anomerization of reducing end. ^{1}H NMR (800 MHz, Methanol-d_4) δ 7.45 – 7.24 (m, 15H), 4.91 (d, *J* = 11.8 Hz, 1H), 4.69 (d, *J* = 11.7 Hz, 1H), 4.65 – 4.58 (m, 3H), 4.58 (s, 1H), 4.45 – 4.36 (m, 2H), 4.28 – 4.18 (m, 1H), 3.95 – 3.85 (m, 5H), 3.70 – 3.53 (m, 5H), 3.51 (t, *J* = 9.0 Hz, 1H), 3.38 (dd, *J* = 9.2, 7.8 Hz, 1H), 3.35 (d, *J* = 2.2 Hz, 1H), 3.34 – 3.31 (m, 3H), 3.29 (ddd, *J* = 9.1, 5.4, 2.5 Hz, 1H), 3.25 (td, *J* = 8.6, 6.1 Hz, 1H). ^{13}C{^{1}H} NMR (201 MHz, Methanol-d_4) δ 138.6, 137.9, 129.0, 128.5, 128.2, 128.1, 128.0, 104.4, 104.1, 101.4, 79.0, 77.3, 76.6, 76.1, 75.9, 74.1, 74.0, 73.2, 72.5, 72.1, 71.6, 71.2, 67.7, 62.5. ESI HRMS: m/z calcd for $C_{38}{}^{13}CH_{50}NaO_{16}$ $[M + Na]^+$ 798.3025, found 798.3011.

The above intermediate was dissolved in 10 mL methanol:water (1:1) and a catalytic amount of palladium on activated charcoal (2 mg) was added into the solution. The reaction was carried out at 50 psi H_2 atmosphere for 24 hours. The catalyst was removed by filtration and the filtrate was collected *in vacuo* to afford ^{13}C-labeled cellotriose **T10** as white powder (0.163 g, 90 % yield in two steps). ESI HRMS: m/z calcd for $C_{17}{}^{13}CH_{32}NaO_{16}$ $[M + Na]^+$ 528.1616, found 528.1628. The ^{1}H and ^{13}C{^{1}H} NMR (800 MHz, and 201 MHz, respectively, Methanol-d_4), complicated spectra due to mutarotation, please see in SI.

Synthesis of cellobiosylglycolaldehyde (CBGA) (Scheme S2)

Synthesis of 2,3,4,6-tetra-O-acetyl-glucopyranosyl-(1→4)-2,3,6-tri-O-acetyl–D–glucopyranose (CB2). Hydrazine acetate (271 mg, 2.9 mmol) was added into a solution of peracetylated cellobiose **CB1** (2 g, 2.9 mmol) in dry DMF. After stirring for 2 hours at 50 °C, the solvent was removed under vacuum. The residue was diluted with ethyl acetate and washed with saturated LiCl to remove residual DMF. The product was chromatographed (hexane:ethyl acetate = 2:1) to give hemiacetal **CB2** (1.569 g 84%), as colorless syrup. Due to hemicetal anomerization resulting in a mixture of isomers, NMR was not measured at this point.

Synthesis of 2,3,4,6-tetra-O-acetyl-glucopyranosyl-(1→4)-2,3,6-tri-O-acetyl–D–glucopyranosyl trichloroacetimidate (CB3). Hemiacetal **CB2** (200 mg, 0.256 mmol) was dissolved in a mixture of 2 mL trichloroacetonitrile and 10 mL DCM, followed by addition of catalytic amount of 1,8-diazabicyclo(5.4.0)undec-7-ene (1.52 mg, 0.01 mmol). The reaction was allowed to proceed at room temperature for an hour, after which the solvent was removed and the product was chromatographed (hexane:ethyl acetate = 1:1) to give **CB3** (173 mg, 87%) as white foam. Due to compound existing as a mixture of anomers, NMR was not measured.

Synthesis of allyl 2,3,4,6-tetra-O-acetyl-glucopyranosyl-(1→4)-2,3,6-tri-O-acetyl-β–D–glucopyranoside (CB4). The glycosyl donor **CB3** (173 mg, 0.223 mmol) and acceptor allyl alcohol were dissolved in 10 mL anhydrous DCM and stirred with molecular sieves (4Å) for 30 min at room temperature. Then a catalytic amount of trimethylsilyl trifluoromethanesulfonate (TMSOTf, 12 μL, 0.068 mmol) was added at -20 °C and the reaction was allowed to continue for 45 min. The reaction was quenched by addition of triethyl amine. Molecular sieves were filtered and the filtrate concentrated *in vacuo*. The product was purified by silica gel chromatography (hexane:ethyl acetate = 3:2) to afford **CB4** (110 mg, 73%) as colorless syrup. ^{1}H NMR (800 MHz, Chloroform-*d*) δ 5.81 (ddt, *J* = 16.4, 10.8, 5.6 Hz, 1H), 5.23 (d, *J* = 17.3 Hz, 1H), 5.19 – 5.14 (m, 3H), 5.12 (t, *J* = 9.4 Hz, 1H), 5.04 (t, *J* = 9.7 Hz, 1H), 4.90 (q, *J* = 8.3 Hz, 2H), 4.50 (s, 1H), 4.49 (s, 1H), 4.35 (dd, *J* = 12.5, 4.3 Hz, 1H), 4.28 (dd, *J* = 13.4, 4.8 Hz, 1H), 4.06 (ddd, *J* = 19.4, 12.5, 5.6 Hz, 2H), 4.02 (d, *J* = 12.4 Hz, 1H), 3.76 (t, *J* = 9.5 Hz, 1H), 3.67 – 3.62 (m, 1H), 3.57 (dd, *J* = 10.2, 4.9 Hz, 1H), 2.11 (s, 3H), 2.06 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H), 2.00 (s, 3H), 1.99 (s, 3H), 1.96 (s, 3H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-*d*) δ 170.4, 170.3, 170.2, 169.8, 169.5, 169.3, 169.0, 133.3, 117.6, 100.7, 99.3, 76.5, 72.9, 72.6, 72.5, 71.9, 71.6, 71.5, 70.0, 67.8, 61.9, 61.5, 20.8, 20.7, 20.6, 20.5. The NMR spectral data are in accordance with literature.[42] ESI HRMS: m/z calcd for $C_{29}H_{40}NaO_{18}$ $[M + Na]^+$ 699.2107, found 699.2119.

Synthesis of allyl glucopyranosyl-(1→4)-β–D–glucopyranoside (CB5). Peracetylated glucoside **CB4** (110 mg, 0.163 mmol) was dissolved in 10 mL anhydrous methanol and NaOMe was added to adjust pH to 9. The reaction was carried out at room temperature overnight and quenched with Amberlyst 15. After filtration, the product was collected in vacuo to afford **CB5** in quantitative yield and used in the next step without further purification. ^{1}H NMR (800 MHz, Methanol-d_4) δ 6.00 (ddq, *J* = 16.7, 11.1, 6.0 Hz, 1H), 5.42 – 5.33 (m, 1H), 5.20 (dt, *J* = 10.5, 3.3 Hz, 1H), 4.45 (dd, *J* = 8.0, 5.6 Hz, 1H), 4.39 (ddd, *J* = 18.2, 6.8, 4.5 Hz, 2H), 4.19 (dt, *J* = 12.6, 6.0 Hz, 1H), 3.98 – 3.84 (m, 3H), 3.70 (dt, *J* = 12.0, 5.9 Hz, 1H), 3.65 – 3.53 (m, 2H), 3.45 – 3.38 (m, 2H), 3.37 – 3.30 (m, 3H), 3.27 (td, *J* = 8.5, 4.6 Hz, 1H). ^{13}C{^{1}H} NMR (201 MHz, Methanol-d_4) δ 134.2, 116.4, 103.2, 101.8, 79.4, 76.6, 76.4, 75.0, 74.9, 73.5, 73.4, 69.9, 69.8, 61.0, 60.5, 48.0, 47.9, 47.8, 47.7, 47.6, 47.5, 47.4. The NMR spectral data are in accordance with literature.[42] ESI HRMS: m/z calcd for $C_{15}H_{26}NaO_{11}$ $[M + Na]^+$ 405.1367, found 405.1355.

Synthesis of glycolaldehyde glucopyranosyl-(1→4)-β–D–glucopyranoside (CBGA). A solution of allyl cellobioside CB5 in 10 mL methanol was cooled to -78 °C and was then saturated with O3. The reaction was stirred for 1 hour and purged with O2, and the remaining ozonide was quenched with excess methyl sulfide, forming DMSO as a byproduct. The reaction was concentrated in vacuo and the residue was re-dissolved in 5 mL methanol. The product was precipitated by a treatment with ice-cold diethyl ether, and collected by centrifuge. This procedure was repeated three times to remove DMSO, which afforded CBGA as white

powder (54 mg, 87% yield in two steps). ESI HRMS: m/z calcd for C14H24NaO12 [M + Na]+ 407.1160, found 407.1133. The ^{1}H and ^{13}C{^{1}H} NMR (800 MHz, and 201 MHz, respectively, Methanol-d_4), complicated spectra due to mutarotation, please see in SI.

Synthesis of ^{13}C- and ^{18}O-doubly labeled cellobiose (G9) (Scheme S3)

Synthesis of benzyl 2,3-di-O-benzyl-4,6-ol-β–D–galactopyranoside (G2). The starting material **G1** (1.1 g, 2 mmol) was dissolved in 20 mL of DCM:TFA:water (10:1:0.1, v:v:v) and was stirred for 15 min. TLC indicated full conversion of the starting material to product. The reaction was quenched by saturated sodium bicarbonate, and extracted twice with DCM. The residue was collected in vacuo and purified by silica gel column chromatography (ethyl acetate:hexane = 4:1) to give **G2** (770 mg, 86%) as syrup. ^{1}H NMR (800 MHz, Methanol-d_4) δ 7.37 (d, J = 7.6 Hz, 4H), 7.30 – 7.23 (m, 11H), 4.94 (d, J = 11.8 Hz, 1H), 4.90 – 4.82 (m, 1H), 4.74 (d, J = 11.8 Hz, 1H), 4.71 (d, J = 11.0 Hz, 1H), 4.66 (d, J = 11.9 Hz, 1H), 4.61 (d, J = 11.8 Hz, 1H), 4.47 (d, J = 7.8 Hz, 1H), 4.05 (s, 1H), 3.82 (dd, J = 11.3, 6.8 Hz, 1H), 3.77 (dd, J = 11.3, 5.3 Hz, 1H), 3.69 (dd, J = 9.6, 7.8 Hz, 1H), 3.51 – 3.43 (m, 2H). ^{13}C{^{1}H} NMR (201 MHz, Methanol-d_4) δ 138.7, 138.4, 137.7, 127.9, 127.9, 127.8, 127.7, 127.7, 127.6, 127.3, 127.2, 127.1, 102.6, 80.9, 78.9, 75.0, 74.6, 71.3, 70.6, 65.9, 61.1. ESI HRMS: m/z calcd for $C_{27}H_{30}NaO_6$ [M + Na]$^+$ 473.1935, found 473.1910.

Synthesis of benzyl 2,3-di-O-benzyl-4,6-ol-β–D–galacto-4-ulopyranoside (G3). This procedure was adapted from literature.[43] A solution of diol **G2** (770 mg, 1.7 mmol), dibutyltin oxide (426 mg, 1.7 mmol), and 3 Å molecular sieves were refluxed for 3 h in anhydrous toluene. The solvent was removed under vacuum and the crude stannylene derivative was dissolved in dry chloroform. 1,3-Dibromo-5,5-dimethylhydantoin (242 mg, 0.855 mmol) was added into this mixture and the reaction was stirred for 30 min at room temperature. The reaction mixture was then diluted with chloroform and washed with sodium thiosulfate. The organic phase was dried and concentrated *in vacuo*. The residue was chromatographed (toluene:acetone = 3:1) to give **G3** (475 mg, 62%) as colorless syrup. ^{1}H NMR (800 MHz, Chloroform-d) δ 7.47 – 7.23 (m, 15H), 5.05 – 4.92 (m, 3H), 4.82 – 4.74 (m, 2H), 4.73 (d, J = 11.9 Hz, 1H), 4.67 (d, J = 11.6 Hz, 1H), 4.21 (d, J = 9.4 Hz, 1H), 4.08 (t, J = 5.0 Hz, 1H), 3.93 (qd, J = 12.2, 5.1 Hz, 2H), 3.86 (dd, J = 9.4, 6.2 Hz, 1H). ^{13}C{^{1}H} NMR (201 MHz, Chloroform-d) δ 204.5, 137.8, 137.4, 136.6, 128.7, 128.5, 128.4, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 102.0, 82.9, 82.5, 78.4, 77.3, 77.1, 76.9, 74.7, 73.9, 71.3, 61.7. ESI HRMS: m/z calcd for $C_{27}H_{28}NaO_6$ [M + Na]$^+$ 471.1778, found 471.1750.

Synthesis of benzyl 2,3-di-O-benzyl-6-ol-β–D–galacto-4[^{18}O]-ulopyranoside (G4). The unlabeled galactouloside **G3** (475 mg, 1.06 mmol) was dissolved in 5 mL 1,4-dioxane and 1 mL $H_2{}^{18}O$ and Amberlyst 15 were added. The reaction was stirred for 24 hours and the equilibrium between labeled and unlabeled species was monitored by LTQ-MS. Upon equilibration, the reaction mixture was filtered and the filtrate was concentrated in vacuo and used in the next step without further purification. ESI HRMS: m/z calcd for $C_{27}H_{28}NaO_5{}^{18}O$ [M + Na]$^+$ 473.1821, found 473.1799.

Synthesis of benzyl 2,3-di-O-benzyl-4[^{18}O],6-ol-β–D–glucopyranoside (G5). This procedure was adapted from literature.[43] A solution of labeled galactouloside **G4** in 10 mL dichloroethane was cooled to 0 °C and triacetoxylborohydride (633 mg, 3.18 mmol) was added. The reaction was carried out for an hour and allowed to warm to room temperature. The reaction mixture was concentrated in vacuo and chromatographed (hexane:ethyl acetate = 1:1) to give **G5** (355 mg, 74% yield in two steps) to yield a colorless syrup. ^{1}H NMR (800 MHz, Methanol-d_4) δ 7.39 – 7.20 (m, 15H), 4.95 (dd, J = 12.0, 3.5 Hz, 1H), 4.91 – 4.83 (m, 2H), 4.81 (d, J = 3.3 Hz, 1H), 4.66 (ddd, J = 23.1, 11.4, 3.5 Hz, 2H), 4.54 (dd, J = 8.1, 3.4 Hz, 1H), 3.92 (dd, J = 11.9, 2.5 Hz, 1H), 3.71 (dt, J = 9.7, 4.9 Hz, 1H), 3.49 (dd, J = 11.1, 7.5 Hz, 1H), 3.45 (td, J = 9.0, 3.4 Hz, 1H), 3.35 (t, J = 8.4 Hz, 1H), 3.32 – 3.28 (m, 1H). ^{13}C{^{1}H} NMR (201 MHz, Methanol-d_4) δ 140.2, 139.9, 139.0, 129.4, 129.4, 129.3, 129.3, 129.2, 129.2, 129.1, 129.0, 128.9, 128.9, 128.8, 128.5, 128.4, 103.6, 85.9, 83.1, 77.8, 76.4, 75.7, 72.0, 71.7, 71.7, 62.7. The NMR spectral data are in accordance with literature.[44] ESI HRMS: m/z calcd for $C_{27}H_{30}NaO_5{}^{18}O$ [M + Na]$^+$ 475.1977, found 475.1989.

Synthesis of benzyl 2,3-di-O-benzyl-4[^{18}O]-ol-6-benzoyl-β–D–glucopyranoside (G6). A solution of labeled galactoside **G5** (175 mg, 0.387 mmol) in 5 mL DCM was cooled to -40 °C. Benzoyl chloride (59.6 mg, 0.426 mmol) was added dropwise under vigorous stirring and the reaction was continued for 30 min. The reaction was quenched by addition of 1mL methanol, and the solvent was removed under vacuum. The residue was chromatographed (hexane:ethyl acetate = 3:1) to give **G6** (100 mg, 47%) as colorless syrup. ^{1}H NMR (800 MHz, Methanol-d_4) δ 8.08 – 8.03 (m, 2H), 7.56 (q, J = 6.0, 4.5 Hz, 1H), 7.44 (q, J = 7.8 Hz, 2H), 7.35 – 7.14 (m, 14H), 4.89 – 4.78 (m, 5H), 4.69 (ddd, J = 9.4, 6.5, 2.9 Hz, 1H), 4.62 (ddd, J = 26.2, 10.9, 4.9 Hz, 2H), 4.54 (t, J = 5.6 Hz, 1H), 4.51 (dt, J = 9.8, 4.9 Hz, 1H), 3.65 – 3.58 (m, 2H), 3.49 (t, J = 8.5 Hz, 1H), 3.39 (q, J = 8.0, 7.4 Hz, 1H). ^{13}C{^{1}H} NMR (201 MHz, Methanol-d_4) δ 166.5, 138.8, 138.5, 137.3, 132.9, 129.9, 129.3, 128.2, 128.0, 127.8, 127.8, 127.7, 127.5, 127.4, 127.2, 127.1, 102.1, 84.4, 81.7, 75.1, 74.4, 73.9, 70.6, 70.5, 63.8. ESI HRMS: m/z calcd for $C_{34}H_{34}NaO_6{}^{18}O$ [M + Na]$^+$ 579.2239, found 579.2211.

Synthesis of benzyl 2,3,4,6-tetra-O-acetyl-β–D–[1-^{13}C]glucopyranosyl-(1→4[^{18}O])-2,3-di-O-benzyl-6-benzoyl-β–D–glucopyranoside (G7). A solution of glycosyl acceptor **G6** (100 mg, 0.18 mmol) and donor peracetylated glucopyranosyl trichloroacetimidate (107 mg, 0.216 mmol) in 10 mL anhydrous DCM was stirred with activated molecular sieves (4Å) for 30 min at room temperature. Catalytic amount of trimethylsilyl trifluoromethanesulfonate (TMSOTf, 12 µL, 0.068 mmol) was added into the mixture and the reaction was continued for 30 min. TLC indicated complete conversion of the starting material to the product. The reaction was quenched by addition of triethyl amine. The solvent was removed under vacuum and the residue was chromatographed to give **G7** (130 mg, 81%) as colorless syrup. ^{1}H NMR (800 MHz, Chloroform-d) δ 8.09 (d, J = 7.8 Hz, 2H), 7.60 (t, J = 7.5 Hz, 1H), 7.48 (t, J = 7.7 Hz, 2H), 7.42 – 7.18 (m, 15H), 5.08 (p, J = 9.3 Hz, 2H), 5.00 (q, J = 7.9 Hz, 1H), 4.96 (d, J = 11.4 Hz, 1H), 4.89 (dt, J = 11.7, 8.3 Hz, 3H), 4.75 (d, J = 11.6 Hz, 1H), 4.69 – 4.61 (m, 3H), 4.53 (d, J = 7.7 Hz, 1H), 4.40 (dt, J = 11.5, 6.2 Hz, 1H), 4.10 (dt, J = 12.3,

5.6 Hz, 1H), 3.90 (td, *J* = 9.0, 4.4 Hz, 1H), 3.84 (d, *J* = 12.6 Hz, 1H), 3.71 – 3.63 (m, 2H), 3.56 – 3.51 (m, 1H), 3.41 (s, 1H), 2.06 (s, 3H), 1.98 (s, 3H), 1.96 (s, 3H), 1.95 (s, 3H). $^{13}C\{^1H\}$ NMR (201 MHz, Chloroform-*d*) δ 170.5, 170.2, 169.4, 169.3, 166.1, 138.9, 138.2, 137.0, 133.4, 129.7, 128.6, 128.5, 128.3, 128.3, 128.1, 128.1, 127.0, 101.9, 101.4, 100.9, 100.7, 100.5, 98.1, 92.4, 91.7, 89.1, 82.5, 81.9, 80.2, 76.2, 74.9, 74.8, 73.1, 73.1, 72.7, 71.9, 71.8, 70.9, 67.9, 63.2, 61.5, 61.4, 20.8, 20.7, 20.66, 20.61, 20.58. ESI HRMS: m/z calcd for $C_{47}{}^{13}CH_{52}NaO_{15}{}^{18}O$ $[M + Na]^+$ 910.3224, found 910.3212.

Synthesis of benzyl β–D–[1-^{13}C]glucopyranosyl-(1→4[^{18}O])-2,3-di-O-benzyl-β–D–glucopyranoside (G8). The procedure for global deacetylation was detailed in the description of the synthesis of **CB5**. The reaction product was purified by reverse column chromatography (water:acetonitrile = 1:1) to give **G8** (64 mg, 71% from **G7**)) as colorless syrup. ^1H NMR (800 MHz, Methanol-d_4) δ 4.95 (d, *J* = 11.7 Hz, 1H), 7.42 – 7.23 (m, 15H), 4.98 (d, *J* = 10.2 Hz, 1H), 4.89 (d, *J* = 11.0 Hz, 1H), 4.75 – 4.65 (m, 3H), 4.57 (d, *J* = 7.7 Hz, 1H), 3.95 – 3.91 (m, 3H), 3.75 (d, *J* = 12.0 Hz, 1H), 3.63 (t, *J* = 9.4 Hz, 1H), 3.52 (dd, *J* = 12.2, 6.5 Hz, 1H), 3.41 (d, *J* = 9.8 Hz, 1H), 3.36 (q, *J* = 9.0 Hz, 2H), 3.22 (p, *J* = 9.2, 8.2 Hz, 4H). $^{13}C\{^1H\}$ NMR (201 MHz, Methanol-d_4) δ 138.4, 138.2, 137.5, 128.2, 127.9, 127.8, 127.8, 127.7, 127.42, 127.40, 127.2, 102.3, 96.8, 92.5, 83.1, 81.6, 77.2, 76.5, 75.7, 75.5, 75.3, 74.6, 74.4, 74.3, 70.66, 70.64, 61.7, 59.9. ESI HRMS: m/z calcd for $C_{32}{}^{13}CH_{40}NaO_{10}{}^{18}O$ $[M + Na]^+$ 638.2539, found 638.2526.

Synthesis of β–D–[1-^{13}C]glucopyranosyl-(1→4[^{18}O])–D–glucopyranose (G9). The procedure for global debenzylation was detailed in the description of the synthesis of **T9**. Quantitative yield of **G9** was obtained from **G8**. ESI HRMS: m/z calcd for $C_{11}{}^{13}CH_{22}NaO_{10}{}^{18}O$ $[M + Na]^+$ 368.1130, found 368.1156.

Synthesis of ^{18}O-labeled cellobiose (Scheme S4)

Cellobiose was dissolved in ^{18}O-enriched water (97%) and Amberlyst acid resin was added as the catalyst. The reaction was stirred overnight and terminated by removing the immobilized acid. Solid product was collected in vacuo after removal of the solvent water. The labeling efficiency was determined to be 80%.

ASSOCIATED CONTENT

Supporting Information. Synthesis schemes for cellobiosylglycoaldehyde and the isotopically labeled model compounds, NMR spectra, and complemental mass spectra. This material is available free of charge via the Internet at http://pubs.acs.org."

AUTHOR INFORMATION

Corresponding Author

*E-mail: hilkk@purdue.edu.

Author Contributions

‡These authors contributed equally.

Notes

The authors declare no conflict of interest

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support by the Center for Direct Catalytic Conversion of Biomass to Biofuels (C3Bio), an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award Number DE-SC0000997.

REFERENCES

(1) Degenstein, J. C.; Murria, P.; Easton, M.; Sheng, H.; Hurt, M.; Dow, A. R.; Gao, J.; Nash, J. J.; Agrawal, R.; Delgass, W. N.; et al. Fast Pyrolysis of 13C-Labeled Cellobioses: Gaining Insights Into the Mechanisms of Fast Pyrolysis of Carbohydrates. *J. Org. Chem.* **2015**, *80* (3), 1909–1914. https://doi.org/10.1021/jo5025255.

(2) Parsell, T.; Yohe, S.; Degenstein, J.; Jarrell, T.; Klein, I.; Gencer, E.; Hewetson, B.; Hurt, M.; Kim, J. I.; Choudhari, H.; et al. A Synergistic Biorefinery Based on Catalytic Conversion of Lignin Prior to Cellulose Starting from Lignocellulosic Biomass. *Green Chem.* **2015**, *17* (3), 1492–1499. https://doi.org/10.1039/C4GC01911C.

(3) Venkatakrishnan, V. K.; Degenstein, J. C.; Smeltz, A. D.; Delgass, W. N.; Agrawal, R.; Ribeiro, F. H. High-Pressure Fast-Pyrolysis, Fast-Hydropyrolysis and Catalytic Hydrodeoxygenation of Cellulose: Production of Liquid Fuel from Biomass. *Green Chem.* **2014**, *16* (2), 792–802. https://doi.org/10.1039/C3GC41558A.

(4) Agrawal, R.; Singh, N. R. Synergistic Routes to Liquid Fuel for a Petroleum-Deprived Future. *AIChE J.* **2009**, *55* (7), 1898–1905. https://doi.org/10.1002/aic.11785.

(5) French, R.; Czernik, S. Catalytic Pyrolysis of Biomass for Biofuels Production. *Fuel Process. Technol.* **2010**, *91* (1), 25–32. https://doi.org/10.1016/j.fuproc.2009.08.011.

(6) Evans, R. J.; Milne, T. A. Molecular Characterization of the Pyrolysis of Biomass. *Energy Fuels* **1987**, *1* (2), 123–137. https://doi.org/10.1021/ef00002a001.

(7) Zheng, Y.; Pan, Z.; Zhang, R. Overview of Biomass Pretreatment for Cellulosic Ethanol Production. *Int. J. Agric. Biol. Eng.* **2009**, *2* (3), 51–68. https://doi.org/10.3965/ijabe.v2i3.168.

(8) Zhang, J.; Choi, Y. S.; Yoo, C. G.; Kim, T. H.; Brown, R. C.; Shanks, B. H. Cellulose–Hemicellulose and Cellulose–Lignin Interactions during Fast Pyrolysis. *ACS Sustain. Chem. Eng.* **2015**, *3* (2), 293–301. https://doi.org/10.1021/sc500664h.

(9) Ralph, J.; Hatfield, R. D. Pyrolysis-GC-MS Characterization of Forage Materials. *J. Agric. Food Chem.* **1991**, *39* (8), 1426–1437. https://doi.org/10.1021/jf00008a014.

(10) Wang, S.; Guo, X.; Liang, T.; Zhou, Y.; Luo, Z. Mechanism Research on Cellulose Pyrolysis by Py-GC/MS and Subsequent Density Functional Theory Studies. *Bioresour. Technol.* **2012**, *104*, 722–728. https://doi.org/10.1016/j.biortech.2011.10.078.

(11) Patwardhan, P. R.; Dalluge, D. L.; Shanks, B. H.; Brown, R. C. Distinguishing Primary and Secondary Reactions of Cellulose Pyrolysis. *Bioresour. Technol.* **2011**, *102* (8), 5265–5269. https://doi.org/10.1016/j.biortech.2011.02.018.

(12) Lu, Q.; Dong, C.; Zhang, X.; Tian, H.; Yang, Y.; Zhu, X. Selective Fast Pyrolysis of Biomass Impregnated with ZnCl2 to Produce Furfural: Analytical Py-GC/MS Study. *J. Anal. Appl. Pyrolysis* **2011**, *90* (2), 204–212. https://doi.org/10.1016/j.jaap.2010.12.007.

(13) Patwardhan, P. R.; Satrio, J. A.; Brown, R. C.; Shanks, B. H. Influence of Inorganic Salts on the Primary Pyrolysis Products of Cellulose. *Bioresour. Technol.* **2010**, *101* (12), 4646–4655. https://doi.org/10.1016/j.biortech.2010.01.112.

(14) Branca, C.; Giudicianni, P.; Di Blasi, C. GC/MS Characterization of Liquids Generated from Low-Temperature Pyrolysis of Wood. *Ind. Eng. Chem. Res.* **2003**, *42* (14), 3190–3202. https://doi.org/10.1021/ie030066d.

(15) Wang, S.; Guo, X.; Liang, T.; Zhou, Y.; Luo, Z. Mechanism Research on Cellulose Pyrolysis by Py-GC/MS and Subsequent Density Functional Theory Studies. *Bioresour. Technol.* **2012**, *104*, 722–728. https://doi.org/10.1016/j.biortech.2011.10.078.

(16) Zhang, X.; Yang, W.; Dong, C. Levoglucosan Formation Mechanisms during Cellulose Pyrolysis. *J. Anal. Appl. Pyrolysis* **2013**, *104*, 19–27. https://doi.org/10.1016/j.jaap.2013.09.015.

(17) Shen, D. K.; Gu, S. The Mechanism for Thermal Decomposition of Cellulose and Its Main Products. *Bioresour. Technol.* **2009**, *100* (24), 6496–6504. https://doi.org/10.1016/j.biortech.2009.06.095.

(18) Zhang, Y.; Liu, C.; Chen, X. Unveiling the Initial Pyrolytic Mechanisms of Cellulose by DFT Study. *J. Anal. Appl. Pyrolysis* **2015**, *113*, 621–629. https://doi.org/10.1016/j.jaap.2015.04.010.

(19) Zhang, X.; Yang, W.; Blasiak, W. Kinetics of Levoglucosan and Formaldehyde Formation during Cellulose Pyrolysis Process. *Fuel* **2012**, *96*, 383–391. https://doi.org/10.1016/j.fuel.2012.01.006.

(20) Sullivan, A. L.; Ball, R. Thermal Decomposition and Combustion Chemistry of Cellulosic Biomass. *Atmos. Environ.* **2012**, *47*, 133–141. https://doi.org/10.1016/j.atmosenv.2011.11.022.

(21) Agarwal, V.; Dauenhauer, P. J.; Huber, G. W.; Auerbach, S. M. Ab Initio Dynamics of Cellulose Pyrolysis: Nascent Decomposition Pathways at 327 and 600 °C. *J. Am. Chem. Soc.* **2012**, *134* (36), 14958–14972. https://doi.org/10.1021/ja305135u.

(22) Liang, X.; Montoya, A.; Haynes, B. S. Local Site Selectivity and Conformational Structures in the Glycosidic Bond Scission of Cellobiose. *J. Phys. Chem. B* **2011**, *115* (36), 10682–10691. https://doi.org/10.1021/jp204199h.

(23) Zhang, X.; Li, J.; Yang, W.; Blasiak, W. Formation Mechanism of Levoglucosan and Formaldehyde during Cellulose Pyrolysis. *Energy Fuels* **2011**, *25* (8), 3739–3746. https://doi.org/10.1021/ef2005139.

(24) Vinu, R.; Broadbelt, L. J. A Mechanistic Model of Fast Pyrolysis of Glucose-Based Carbohydrates to Predict Bio-Oil Composition. *Energy Environ. Sci.* **2012**, *5* (12), 9808–9826. https://doi.org/10.1039/C2EE22784C.

(25) Seshadri, V.; Westmoreland, P. R. Concerted Reactions and Mechanism of Glucose Pyrolysis and Implications for Cellulose Kinetics. *J. Phys. Chem. A* **2012**, *116* (49), 11997–12013. https://doi.org/10.1021/jp3085099.

(26) Assary, R. S.; Curtiss, L. A. Thermochemistry and Reaction Barriers for the Formation of Levoglucosenone from Cellobiose. *ChemCatChem* **2012**, *4* (2), 200–205. https://doi.org/10.1002/cctc.201100280.

(27) Hosoya, T.; Nakao, Y.; Sato, H.; Kawamoto, H.; Sakaki, S. Thermal Degradation of Methyl β-d-Glucoside. A Theoretical Study of Plausible Reaction Mechanisms. *J. Org. Chem.* **2009**, *74* (17), 6891–6894. https://doi.org/10.1021/jo900457k.

(28) Lin, Y.-C.; Cho, J.; Tompsett, G. A.; Westmoreland, P. R.; Huber, G. W. Kinetics and Mechanism of Cellulose Pyrolysis. *J. Phys. Chem. C* **2009**, *113* (46), 20097–20107. https://doi.org/10.1021/jp906702p.

(29) Mayes, H. B.; Broadbelt, L. J. Unraveling the Reactions That Unravel Cellulose. *J. Phys. Chem. A* **2012**, *116* (26), 7098–7106. https://doi.org/10.1021/jp300405x.

(30) Easton, M. W.; Nash, J. J.; Kenttämaa, H. I. Dehydration Pathways for Glucose and Cellobiose During Fast Pyrolysis. *J. Phys. Chem. A* **2018**, *122* (41), 8071–8085. https://doi.org/10.1021/acs.jpca.8b02312.

(31) Easton, M. W. Density Functional Theory Calculations Complement Mass Spectrometry Experiments in the Investigation of Biomass Fast Pyrolysis and Ion-Molecule Reaction Mechanisms, Purdue University: West Lafayette, IN, 2016.

(32) Richards, G. N. Glycolaldehyde from Pyrolysis of Cellulose. *J. Anal. Appl. Pyrolysis* **1987**, *10* (3), 251–255. https://doi.org/10.1016/0165-2370(87)80006-2.

(33) Hurt, M. R.; Degenstein, J. C.; Gawecki, P.; Borton II, D. J.; Vinueza, N. R.; Yang, L.; Agrawal, R.; Delgass, W. N.; Ribeiro, F. H.; Kenttämaa, H. I. On-Line Mass Spectrometric Methods for the Determination of the Primary Products of Fast Pyrolysis of Carbohydrates and for Their Gas-Phase Manipulation. *Anal. Chem.* **2013**, *85* (22), 10927–10934. https://doi.org/10.1021/ac402380h.

(34) Degenstein, J. C.; Hurt, M.; Murria, P.; Easton, M.; Choudhari, H.; Yang, L.; Riedeman, J.; Carlsen, M. S.; Nash, J. J.; Agrawal, R.; et al. Mass Spectrometric Studies of Fast Pyrolysis of Cellulose. *Eur. J. Mass Spectrom.* **2015**, *21* (3), 321–326. https://doi.org/10.1255/ejms.1335.

(35) Fang, T. T.; Bendiak, B. The Stereochemical Dependence of Unimolecular Dissociation of Monosaccharide-Glycolaldehyde Anions in the Gas Phase: A Basis for Assignment of the Stereochemistry and Anomeric Configuration of Monosaccharides in Oligosaccharides by Mass Spectrometry via a Key Discriminatory Product Ion of Disaccharide Fragmentation, m/z 221. *J. Am. Chem. Soc.* **2007**, *129* (31), 9721–9736. https://doi.org/10.1021/ja0717313.

(36) Zhou, X.; Nolte, M. W.; Mayes, H. B.; Shanks, B. H.; Broadbelt, L. J. Experimental and Mechanistic Modeling of Fast Pyrolysis of Neat Glucose-Based Carbohydrates. 1. Experiments and Development of a Detailed Mechanistic Model. *Ind. Eng. Chem. Res.* **2014**, *53* (34), 13274–13289. https://doi.org/10.1021/ie502259w.

(37) Wang, Z.; Chinoy, Z. S.; Ambre, S. G.; Peng, W.; McBride, R.; Vries, R. P. de; Glushka, J.; Paulson, J. C.; Boons, G.-J. A General Strategy for the Chemoenzymatic Synthesis of Asymmetrically Branched N-Glycans. *Science* **2013**, *341* (6144), 379–383. https://doi.org/10.1126/science.1236231.

(38) Xu, C.; Liu, H.; Li, X. Thioglycosylation of 1,2-Cis-Glycosyl Acetates: A Long-Standing Overlooked Issue in Preparative Carbohydrate Chemistry. *Carbohydr. Res.* **2011**, *346* (9), 1149–1153. https://doi.org/10.1016/j.carres.2011.03.033.

(39) Tatina, M.; Yousuf, S. K.; Mukherjee, D. 2,4,6-Trichloro-1,3,5-Triazine (TCT) Mediated One-Pot Sequential Functionalisation of Glycosides for the Generation of Orthogonally Protected Monosaccharide Building Blocks. *Org. Biomol. Chem.* **2012**, *10* (28), 5357–5360. https://doi.org/10.1039/C2OB25452B.

(40) Autar, R.; Liskamp, R. M. J.; Pieters, R. J. A Facile Synthesis of the GalNAcβ1→4Gal Target Sequence of Respiratory Pathogens. *Carbohydr. Res.* **2005**, *340* (15), 2436–2442. https://doi.org/10.1016/j.carres.2005.08.001.

(41) Reina, J. J.; Rojo, J. Synthesis of 2-Azidoethyl α-d-Mannopyranoside Orthogonally Protected and Selective Deprotections. *Tetrahedron Lett.* **2006**, *47* (15), 2475–2478. https://doi.org/10.1016/j.tetlet.2006.02.063.

(42) Sakakibara, K.; Nakatsubo, F.; French, A. D.; Rosenau, T. Chiroptical Properties of an Alternatingly Functionalized Cellotriose Bearing Two Porphyrin Groups. *Chem. Commun.* **2012**, *48* (62), 7672–7674. https://doi.org/10.1039/C2CC30805C.

(43) Söderman, P.; Widmalm, G. Stereospecific Deuteration in the Synthesis of Methyl α-(4-2H)-Cellobioside. *J. Org. Chem.* **1999**, *64* (11), 4199–4200. https://doi.org/10.1021/jo982256h.

(44) Fügedi, P. Synthesis of 4-O-(α-L-Rhamnopyranosyl-D-Glucopyranuronic Acid. *J. Carbohydr. Chem.* **1987**, *6* (3), 377–398. https://doi.org/10.1080/07328308708057927.

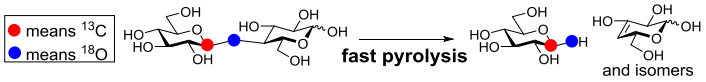


Table of Content