Mechanistic insights into water-catalysed formation of levoglucosenone from anhydrosugar intermediates by means of high-level theoretical procedures

Wenchao Wan, Li-Juan Yu, Amir Karton*

School of Chemistry and Biochemistry, The University of Western Australia, Perth, WA 6009, Australia

Abstract

Levoglucosenone (LGO) is an important anhydrosugar product from fast pyrolysis of cellulose and biomass. We use the high-level G4(MP2) thermochemical protocol to study the reaction mechanism for the formation of LGO from the 1,4:3,6-dianhydro-α-D-glucopyranose (DGP) pyrolysis intermediate. We find that the DGP-to-LGO conversion proceeds via a multistep reaction mechanism, which involves ring-opening, ring-closing, enol-to-keto tautomerisation, hydration, and dehydration reactions. The rate-determining step (RDS) for the uncatalysed process is the enol-to-keto tautomerisation (\( \Delta G^{\ddagger}_{298} = 68.6 \text{ kcal mol}^{-1} \)). We find that a water molecule can catalyse five of the seven steps in the reaction pathway. In the water-catalysed process the barrier for the enol-to-keto tautomerisation is reduced by as much as 15.1 kcal mol\(^{-1}\), and the hydration step becomes the RDS with an activation energy of \( \Delta G^i_{298} = 58.1 \text{ kcal mol}^{-1} \).

Keywords: Computational thermochemistry, G4(MP2), CCSD(T), levoglucosenone, water catalysis.

*Corresponding author: Amir Karton, E-mail: amir.karton@uwa.edu.au.
1. Introduction

Lignocellulosic biomass – with an annual global production of 220 billion tons per year\[^1\] – is a renewable source for fuels and chemicals and is mainly composed of cellulose, hemicellulose, and lignin\[^2\]. Pyrolysis is a promising technology to convert biomass into char, liquid and gaseous products that could be further upgraded to biofuels and fine chemicals\[^3\]. Pyrolysis liquid products are a mixture of water and hundreds of low-concentration organic compounds, including acids, aldehydes, anhydrosugars, and furans\[^4\]–\[^6\]. In order to produce targeted valuable industrial chemicals, it is important to understand the chemical mechanisms underlying the formation of specific chemicals during pyrolysis.

Levoglucosenone (LGO) is one of the most valuable pyrolysis products due to its unique structure and functional groups (e.g., keto, olefinic, and ether groups)\[^7\],\[^8\]. Accordingly, LGO is the starting material for a wide range of synthetic products including biologically active natural products, sugar mimetics, and agrochemicals\[^9\],\[^10\]. It was found that 5.5 wt% of LGO is produced via fast pyrolysis of cellulose\[^11\], and that the yield of LGO could be increased by using a variety of catalysts including H\(_3\)PO\(_4\)\[^12\]–\[^14\], H\(_2\)SO\(_4\)\[^15\], ionic liquids\[^16\], SO\(_4^2−/ZrO_2\) and SO\(_4^2−/TiO_2\)\[^17\].

The formation of LGO during pyrolysis has been studied both theoretically and experimentally\[^7\],\[^18\]–\[^24\]. However, the molecular mechanisms underlying the formation of LGO during pyrolysis of cellulose and biomass are still largely unknown. A number of precursors for the formation of LGO were proposed and investigated\[^7\],\[^19\],\[^20\],\[^25\]. In an early experimental study it was suggested that levoglucosan (LG) is a precursor of LGO\[^7\]. More recently, Assary and Curtiss\[^19\] proposed a detailed scheme of reaction pathways for the conversion of LG to LGO using density functional theory (DFT) and high-level \textit{ab initio} procedures. They found that a water molecule could catalyse this transformation with an activation barrier of \(\Delta H^{298} = 56.6\text{ kcal mol}^{-1}\) at the G4 level. Qiang Lu et al. studied the
reaction pathways for the formation of LGO from β-D-glucopyranose and cellobiose at the B3LYP/6-31+G(d,p) level of theory\textsuperscript{[20]}. They found that the enol-to-keto tautomerisation is the rate-determining step (RDS) both from β-D-glucopyranose and cellobiose with barriers of $\Delta H^\ddagger_{298} = 69.5 \text{ kcal mol}^{-1}$ and $\Delta H^\ddagger_{298} = 62.7 \text{ kcal mol}^{-1}$, respectively.

In a recent computational investigation, Satorri\textsuperscript{[25]} investigated the reaction mechanism for the conversion of LG to LGO, and suggested that 1,4:3,6-dianhydro-α-D-glucopyranose (DGP) – an important anhydrosugar product (3.3 wt\%\textsuperscript{[11]}) from cellulose pyrolysis – might be a possible precursor for the formation of LGO. However, the reaction mechanism for the DGP-to-LGO conversion was not investigated. Figure 1 summarises the two pathways for the conversion of cellulose to LGO.

**Fig. 1.** A schematic illustration of the two possible pathways for the conversion of cellulose to LGO.

Nearly four decades ago Shafizadeh et al.\textsuperscript{[26]} showed that pyrolysis of DGP under acidic conditions generated significant amounts of LGO. They proposed a general reaction scheme for the formation of LGO from DGP. This scheme is illustrated in Figure 2. This scheme involves diol, triol, and tetraol intermediates, however the reaction mechanisms leading to
their formation were not discussed. In the present work we investigate the reaction mechanism for the uncatalysed formation of LGO from DGP in more detail using the high-level ab-initio G4(MP2) theory \cite{27}. We also consider the water-catalysed reaction and show that a water catalyst can reduce the Gibbs free-activation energy at 298 K ($\Delta G^\ddagger_{298}$) for the entire process by as much as 10.5 kcal mol$^{-1}$. In addition, we show that the water catalyst leads to a change in the RDS for the overall process.

![Diagram](image)

**Fig. 2.** Reaction scheme for the conversion of LGO to DGP proposed by Shafizadeh et al. \cite{26}. For the sake of consistency the labels of the intermediates are taken from Figure 3.

### 2. Computational Details

High-level ab initio \cite{28} and density functional theory \cite{29} calculations were carried out using the Gaussian 09 program suite \cite{30}. All geometry optimisations were performed at the B3LYP/6-31G(2df,p) level of theory as prescribed in the G4(MP2) protocol \cite{27}. Zero point vibrational energies, enthalpic, and entropic corrections have been obtained from such calculations. The equilibrium structures were verified to have all real harmonic frequencies and the transition structures to have only one imaginary frequency. The connectivities of the transition structures were confirmed by performing intrinsic reaction coordinate (IRC) calculations \cite{31,32}.
Gas-phase Gibbs free energies at 298 K ($\Delta G_{298}$) were obtained using the G4(MP2) variant of the Gaussian-4 (G4) composite thermochemical protocol [27, 33]. The G4(MP2) protocol is an efficient composite procedure for approximating the CCSD(T) (coupled cluster energy with singles, doubles, and quasiperturbative triple excitations) energy in conjunction with a large triple-ζ-quality basis set [34, 35]. This protocol is widely used for the calculation of thermochemical and kinetic properties (for a recent review see reference [34]). The G4(MP2) procedure has been found to produce gas-phase thermochemical properties (such as reaction energies, bond dissociation energies, and enthalpies of formation) with a mean absolute deviation (MAD) of 1.04 kcal mol$^{-1}$ from the 454 experimental energies of G3/05 test set [36]. It has been recently found that G4(MP2) shows a similarly good performance for barrier heights [37–40].

3. Results and Discussion

3.1 Uncatalysed mechanism for the formation of LGO from DGP

Shafizadeh et al. proposed a general reaction mechanism for the conversion of DGP to LGO (Figure 2) [26]. This scheme involves three polyol intermediates, however, the steps leading to the formation of these intermediates were not specified or discussed in detail. We begin by proposing a more detailed seven-step reaction pathway for the formation of LGO from DGP. Our proposed reaction mechanism is illustrated schematically in Figure 3.
Fig. 3. Detailed reaction mechanism for the conversion of DGP to LGO proposed in the present work.

The seven steps in our proposed reaction mechanism are described as follows. (i) In the first step a water molecule is added to the pyranose ring. Specifically, the hydrogen atom is added to the pyranose oxygen and a hydroxyl group is added to the C₁ carbon atom. This ring-opening reaction leads to the formation of the first intermediate (INT₁, Figure 3). (ii) A hydrogen-transfer reaction, in which the hydrogen from the hydroxyl group at the C₁ position is transferred to the cyclic furanose oxygen. In this step the C₁–O bond of the furanose ring breaks, resulting in the formation of INT₂ (Figure 3). (iii) A hydrogen-transfer from C₂ to the pyranose oxygen leading to the breaking of the pyranose ring and the formation of an acyclic structure (INT₃). (iv) An enol-to-keto tautomerisation leading to the formation of INT₄ (Figure 3). (v) A dehydration reaction, in which the hydrogen attached to C₃ is eliminated with the hydroxyl group attached to C₄, leading to the formation of the C₃=C₄ bond in INT₅. (vi) A hydrogen-transfer cyclisation reaction, in which the hydrogen of the hydroxyl group at C₅ shifts to the carbonyl oxygen in the C₁ position, forming a pyranose ring. (vii) A dehydration-cyclisation reaction leading to the formation of the final LGO product (Figure 3).
We note that Shafizadeh et al. also proposed an additional reaction mechanism (Scheme 1 in reference [26]). We considered this reaction mechanism and found that the first hydrogen-transfer step has a higher Gibbs free-activation energy than the RDS for the mechanism illustrated in Figure 3. For the uncatalysed reaction, the barrier for the first step of this alternative mechanism is $\Delta G^\ddagger_{298} = 73.6 \text{ kcal mol}^{-1}$, whilst the barrier for the RDS for the mechanism in Figure 3 is 68.6 kcal mol$^{-1}$ (vide infra). Similarly, for the water-catalysed reaction, the barrier for the first step of the alternative mechanism is $\Delta G^\ddagger_{298} = 71.1 \text{ kcal mol}^{-1}$, whilst the barrier for the RDS for the mechanism in Figure 3 is 58.1 kcal mol$^{-1}$ (Section 3.2). Therefore, we will not consider this alternative mechanism here (see Supporting Information for further details).

The G4(MP2) Gibbs free-energy profile ($\Delta G_{298}$) for the uncatalysed reaction pathway is depicted in Figure 4 (black line), whilst Figure 5 shows the optimised TSs located along the uncatalysed reaction profile. The lowest energy conformation of DGP is chosen as the zero-energy reference point of the reaction profile. The first hydration step has an activation energy of $\Delta G^\ddagger_{298}(\text{TS1}) = 58.1 \text{ kcal mol}^{-1}$. The relatively high activation energy for this step may be attributed to the fact that the bonds that are being broken and formed in TS1 are far from their equilibrium distances. In particular, the bond length of the breaking C$_1$–O bond of the pyranose ring is 2.603 Å, whilst the bond lengths of the forming O–H and C$_1$–OH bonds are 1.472 and 2.288 Å, respectively. Interestingly, the product of this hydration reaction (INT1) is essentially isoenergetic with the DGP reactant.
Fig. 4. Reaction profile ($\Delta G_{298}$, kcal mol$^{-1}$) for the uncatalysed (black line), water-catalysed (blue line), and self-catalysed (red line) conversion of DGP to LGO.

Fig. 5. B3LYP/6-31G(2df,p) optimised transition structures TS1–TS7 in the reaction profile for the uncatalysed conversion of DGP to LGO (Figure 4). The bonds being broken and formed in the TSs are represented by dashed lines. Atomic color scheme: H, white; C, gray; O, red. The carbon atoms involved in the TSs are labeled in the same way as in Figure 3.

The transition structures for the succeeding five steps (TS2–TS6) share similar structural features – they all involve a strained four-membered ring formed between a migrating hydrogen atom and three heavy C and O atoms (Figure 5). The Gibbs free-activation energies for these steps spread over a wide range from 38.8 (TS6) to 68.6 (TS4).
kcal mol\(^{-1}\). The last dehydration step of the reaction mechanism is associated with a relatively low activation Gibbs free energy of \(\Delta G^{\dagger}_{298}(TS7) = 45.5\) kcal mol\(^{-1}\) (or \(\Delta G^{\dagger}_{298}(TS7) = 53.0\) kcal mol\(^{-1}\), relative to the stable INT6 intermediate). The TSs with the highest Gibbs free-activation energies are TS3–TS5, therefore we will now concentrate on these steps.

The TS for the ring-opening, hydrogen-transfer step (TS3, Figure 5) involves a hydrogen transfer from C\(_2\) to the furanose oxygen, breaking of the C\(_3\)–O bond, and formation of the C\(_2\)=C\(_3\) double bond. This TS has an activation energy of \(\Delta G^{\dagger}_{298}(TS3) = 61.4\) kcal mol\(^{-1}\). The lengths of the breaking C\(_3\)–O and C\(_2\)–H bonds are 1.576 Å and 1.554 Å, respectively, whilst the lengths of the forming C\(_2\)=C\(_3\) and O–H bonds are 1.500 and 1.175 Å, respectively.

The enol-to-keto tautomerisation (TS4, Figure 5) is the RDS for the entire process with an activation Gibbs free energy of 68.1 kcal mol\(^{-1}\) relative to the DGP reactant. We note that the reaction barrier relative to the enol tautomer (INT3) is 57.5 kcal mol\(^{-1}\). This reaction barrier is almost identical to the G4(MP2) reaction barrier for the enol-to-keto tautomerisation in vinyl alcohol (56.9 kcal mol\(^{-1}\))\(^{[38]}\). It has been previously shown that this reaction can be catalysed by a range of catalysts, including water, organic acids, and inorganic acids \(^{[41, 42]}\), which alleviate the high strain energy involved in the TS. We also note that the G4(MP2) reaction barrier for the enol-to-keto tautomerisation in vinyl alcohol is in very good agreement with the barrier height calculated at the higher W1-F12 level (56.7 kcal mol\(^{-1}\)) \(^{[41, 43]}\). This close agreement increases our confidence in the G4(MP2) reaction barrier heights.

The transition structure for the first dehydration step (TS5, Figure 5) involves the elimination of the hydroxyl group at the C\(_3\) position with the hydrogen at the C\(_4\) position (Figure 3). Similarly to the TSs for the 1,3-hydrogen transfer steps (TS2, TS3, TS4, and TS6), TS5 involves a strained four-membered ring as shown in Figure 5. In TS5, the hydroxyl group at the C\(_4\) position has a strong hydrogen bond with the carbonyl oxygen at the
C₁ position. This is evident from a relatively short length of the OH•••OC₁ hydrogen bond (namely, 1.686 Å). The Gibbs free-activation energy for this step is 62.1 kcal mol⁻¹. For comparison, this barrier is significantly lower than the uncatalysed barrier reported by Assary et al.[19] for the initial dehydration step in the conversion of LG to LGO (ΔG‡₉₀⁺ = 68.7 kcal mol⁻¹, MP2/6-311++G(3df,3pd)). Another point of reference is the dehydration reaction in ethanol, the G4(MP2) barrier for this reaction is ΔG‡₂⁹₈ = 66.7 kcal mol⁻¹. The lower Gibbs free-activation energy for TS₅ may be partially attributed to the abovementioned hydrogen-bonding interaction.

The last two steps of the reaction are highly exergonic. Namely, INT₆ lies 7.5 kcal mol⁻¹ below the reactant and the final LGO product lies 11.6 kcal mol⁻¹ below the reactant. Both steps are associated with relatively low reaction barrier heights, namely of 38.8 (TS₆) and 45.5 (TS₇) kcal mol⁻¹ (or 53.0 kcal mol⁻¹ relative to the stable intermediate INT₆).

Two important features of the overall reaction profile for the uncatalysed conversion of DGP to LGO are:

- The enol-to-keto tautomerisation step is clearly the RDS. In particular, TS₄ is higher by amounts ranging from 6.5 (TS₅) to 29.8 (TS₆) kcal mol⁻¹ than the other TSs involved in the process.
- The overall process is highly exergonic; the final product lies lower in energy than the reactant by 11.6 kcal mol⁻¹. This provides a thermodynamic driving force for the entire process.[44]

It is well known that the presence of water molecules can significantly enhance the reaction rates of hydrogen transfer reactions in the gas phase relative to the rates of the uncatalysed reactions[38, 41, 42, 44–47]. Thus, TS₂–TS₆ – which involve a strained four-membered ring formed between a migrating hydrogen and three heavy C and O atoms
(Figure 5) – could potentially be catalysed by a water catalyst. In the next section we will explore the water-catalysed conversion of DGP to LGO.

3.2 Water catalysed mechanism for the formation of LGO from DGP

Water is not only a major constituent of raw biomass material \cite{48, 49}, but also a significant pyrolytic product along pyrolysis processes such as the hydroperoxide decomposition step \cite{50}. It has been found that water could catalyse certain important reactions in pyrolysis of glucose such as ring-opening \cite{51}, tautomerisation \cite{52} and dehydration reactions \cite{52}. In this section we investigate the reaction profile for the conversion of DGP to LGO catalysed by a water molecule.

We find that an explicit water molecule can act as a hydrogen bridge in the transition structures TS2–TS6, thereby lowering the activation energies for these steps by amounts ranging from 4.1 (TS3) to 15.1 (TS4) kcal mol\(^{-1}\). As a consequence of this catalytic activity, the enol-to-keto tautomerisation (TS4) is no longer the RDS. The RDS for the catalysed process becomes the first dehydration step (TS1), which cannot be catalysed by a water molecule.

The water-catalysed reaction profile is shown in Figure 4 (blue line) and the optimised TSs are depicted in Figure 6. In all of these TSs the water molecule assists the hydrogen transfer by converting the four-membered ring TS into a less strained six-membered ring TS (Figures 5 and 6).
In the second step a water molecule acts as a hydrogen shuttle that transports a hydrogen from the hydroxyl group to the furanose oxygen (WTS2, Figure 6). The activation energy for the catalysed step is reduced by 7.4 kcal mol$^{-1}$ relative to the uncatalysed reaction. We also located a self-catalysed TS for this step (STS2, Figure 6) which is associated with a lower activation energy $\Delta G^\ddagger_{298}(STS2) = 35.5$ kcal mol$^{-1}$ (red line, Figure 4). In the self-catalysed TS the hydroxyl group attached to C$_5$ acts as the hydrogen shuttle.

The activation energy for the enol-to-keto tautomerisation (WTS4, Figure 6) is reduced by as much as 15.1 kcal mol$^{-1}$ upon an inclusion of the water catalyst. This step, which is the RDS step of the uncatalysed process, is now associated with an activation energy of $\Delta G^\ddagger_{298}(WTS4) = 53.5$ kcal mol$^{-1}$. Thus, in the catalysed mechanism the reaction barrier for this step is lower by 4.6 kcal mol$^{-1}$ than the reaction barrier for the hydration step ($\Delta G^\ddagger_{298}(TS1) = 58.1$ kcal mol$^{-1}$). Upon inclusion of the water catalyst, the activation energy
for TS5 is reduced by 8.0 kcal mol\(^{-1}\), whilst the reaction barrier for TS6 is reduced by 10.2 kcal mol\(^{-1}\).

The results above indicate that the involvement of an explicit water molecule in the hydrogen shifts significantly affects the dynamics of the overall reaction. In the uncatalysed mechanism, the RDS is the enol-to-keto tautomerisation ($\Delta G^\ddagger_{298}(\text{TS4}) = 68.6$ kcal mol\(^{-1}\)). In the catalysed mechanism the ring-opening hydration step becomes the RDS ($\Delta G^\ddagger_{298}(\text{TS1}) = 58.1$ kcal mol\(^{-1}\)), and the ring-opening hydrogen-transfer step is associated with a slightly lower activation energy of $\Delta G^\ddagger_{298}(\text{WTS3}) = 57.3$ kcal mol\(^{-1}\). Overall, the involvement of the water catalyst reduces the Gibbs free-activation energy for the DGP-to-LGO transformation by 10.5 kcal mol\(^{-1}\).

4. Conclusions

We use the high-level ab initio G4(MP2) thermochemical procedures to investigate the uncatalysed and water-catalysed reaction mechanism for the transformation of DGP to LGO. We draw the following conclusions:

1. The reaction mechanism has seven elementary steps, including ring-opening, ring-closing, enol-to-keto tautomerisation, dehydration, and hydration reactions.
2. In the uncatalysed mechanism, the enol-to-keto tautomerisation is the RDS with an activation Gibbs free energy of $\Delta G^\ddagger_{298} = 68.6$ kcal mol\(^{-1}\) relative to the DGP reactant.
3. In the catalysed mechanism, the enol-to-keto tautomerisation has a reaction barrier of $\Delta G^\ddagger_{298} = 53.5$ kcal mol\(^{-1}\) and is no longer the RDS. In this scenario the ring-opening hydration step becomes the RDS with a barrier of $\Delta G^\ddagger_{298} = 58.1$ kcal mol\(^{-1}\), this barrier is closely followed by the barrier for the ring-opening hydrogen-transfer step ($\Delta G^\ddagger_{298} = 57.3$ kcal mol\(^{-1}\)$).
4. The transformation from DGP to LGO is thermodynamically favored as the Gibbs free energy of LGO lies 11.6 kcal mol$^{-1}$ below DGP.

Our results provide important insights into the molecular mechanisms of biomass pyrolysis. Most notably, that this conversion is water-catalysed and that the ring-opening hydration step is the RDS. We hope that our high-level theoretical results will inspire additional experimental measurements.

**Supporting Information**

Gibbs free energies for the first step in the competing reaction mechanism proposed by Shafizadeh *et al.* [26] (Table S1 and Figure S1); G4(MP2) reaction profile on the enthalpic surface for the uncatalysed and water-catalysed conversion of DGP to LGO, including zero-point vibrational energy (ZPVE), heat content function ($H_{298}-H_0$), and entropic ($\Delta S$) corrections from G4(MP2) theory (Table S2); absolute energies required to calculate all the data in Table S2 (Table S3); and B3LYP/6-31G(2df,p) optimised geometries for all the local minima and transition structures considered in this work (Table S4).

**Acknowledgements**

This study is dedicated to the late Professor Allan H. White (1938–2016) who made numerous contributions to the fields of inorganic chemistry and crystallography [53]. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government. We also acknowledge system administration support provided by the Faculty of Science at UWA to the Linux cluster of the Karton group. We gratefully acknowledge the provision of a Scholarship for International Research Fees (SIRF) and a University International Stipend
scholarship (to W.W.), the provision of an Australian Postgraduate Award (to L-J.Y.), and an Australian Research Council (ARC) Discovery Early Career Researcher Award (to A.K., project number: DE140100311).

References


Graphical TOC

\[ \Delta G_{298} = 68.6 \text{ kcal mol}^{-1} \]
Uncatalysed RDS

\[ \Delta G_{298} = 58.1 \text{ kcal mol}^{-1} \]
Water Catalysed RDS