Electronic Supplementary Material (ESI) for Energy Advances.

Impact of Lignin-carbohydrate complex (LCC) linkages on cellulose pyrolysis chemistry

Arul Mozhi Devan Padmanathan[a], Seth Beck[a], Khursheed B. Ansari[+] and Samir H Mushrif [a]*

**Email: mushrif@ualberta.ca*

[a] Department of Chemical and Materials Engineering, University of Alberta, 9211-116 Street Northwest, Edmonton, Alberta T6G 1H9, Canada.

[+] School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, Singapore

Supplementary Information

Contents

S.1 CPMD-Metadynamics calculations

All CPMD [48] and metadynamics calculations in this study employed CPMD code version 4.3.0 [1] with the plane-wave-pseudopotential implementation of Kohn-Sham density functional theory (DFT) [2]. The Martins-Trouiller (MT) pseudopotential [3]with the revised Perdew-Burke-Ernzerhof (revPBE) functional [4,5] of the generalized gradient approximation (GGA) was used in all CPMD calculations. The plane-wave energy cut-off was set at 70 Ryd. A single k-point (Γ-point) was employed for the integration over the Brillouin zone in reciprocal space. CPMD calculations were performed at a temperature of 298K, and temperature control was achieved using the Nosé-Hoover thermostat [6,7]. The frequency for the ionic thermostat was set to 3000 cm−1 for all systems, corresponding to approximate frequencies of C-H and O-H bond vibrations. To prevent coupling between the ionic and electronic thermostats, the electronic frequency was set to 10000 cm−1 for all systems. The fictitious electron mass was chosen as 300 a.u., the fictitious electronic kinetic energy was within the range of 0.004 a.u. to 0.005 a.u., and a timestep of 4 a.u. (~0.0968 fs) was selected for all calculations. The energies, including the fictitious electronic kinetic energy, were monitored to ensure that the systems did not deviate from the Born-Oppenheimer surface during the molecular dynamics simulations. Two torsion angles were employed as collective variables to enhance the sampling of the CPMD calculations and explore conformations with significant energy barriers using the metadynamics technique [8,9]. Torsion angle 1 was used to rotate the position of the phenyl ring, while torsion angle 2 rotated the cellobiose moiety. This approach ensured that the system sampled all major conformational changes, specifically the variation in the relative position of the quinone methide intermediate to the cellobiose moiety. The height of the potential was maintained at approximately 0.63 kcal/mol. The convergence of the metadynamics simulations was considered complete if at least one of the torsion angles fully sampled the range of 0-360°, and the other angle explored over 90% of the torsional space. The free energy surfaces constructed for LCC-C2, LCC-C3 and LCC-C6 from which the low energy conformers are extracted are shown below in Figure S1. Subsequently, all structures corresponding to the minima were optimized using all-electron DFT calculations.

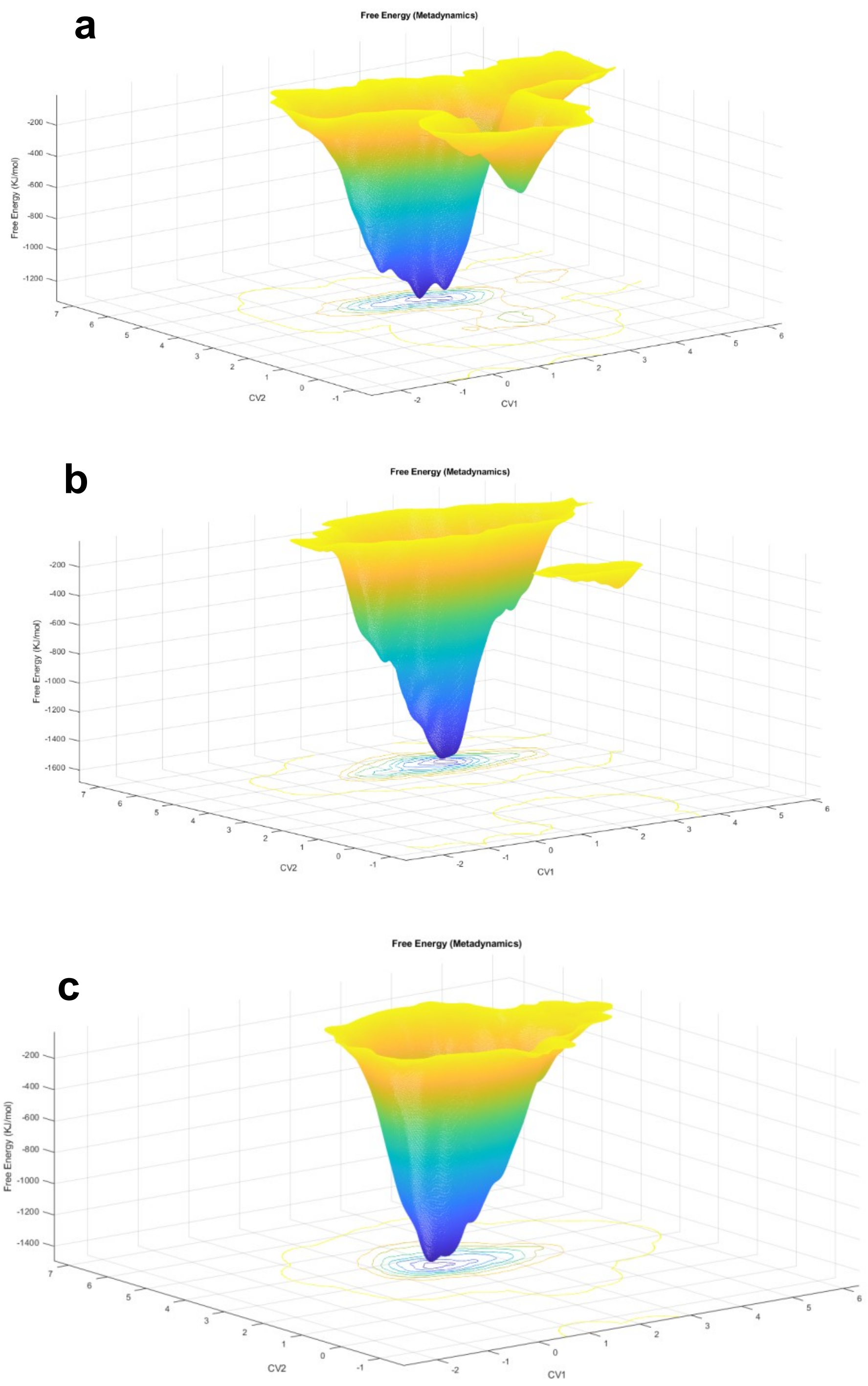

Figure S1: Free energy surfaces constructed by rotating the LCC molecule along CV1 (Torsional angle along the phenyl ring in lignin) and CV2 (Torsional angle along the LCC bond). a) LCC-C2, b) LCC-C3, c) LCC-C6

S.2 Thin-film Preparation

Thin-film techniques are used to maintain the sample at isothermal and kinetically limited regime.

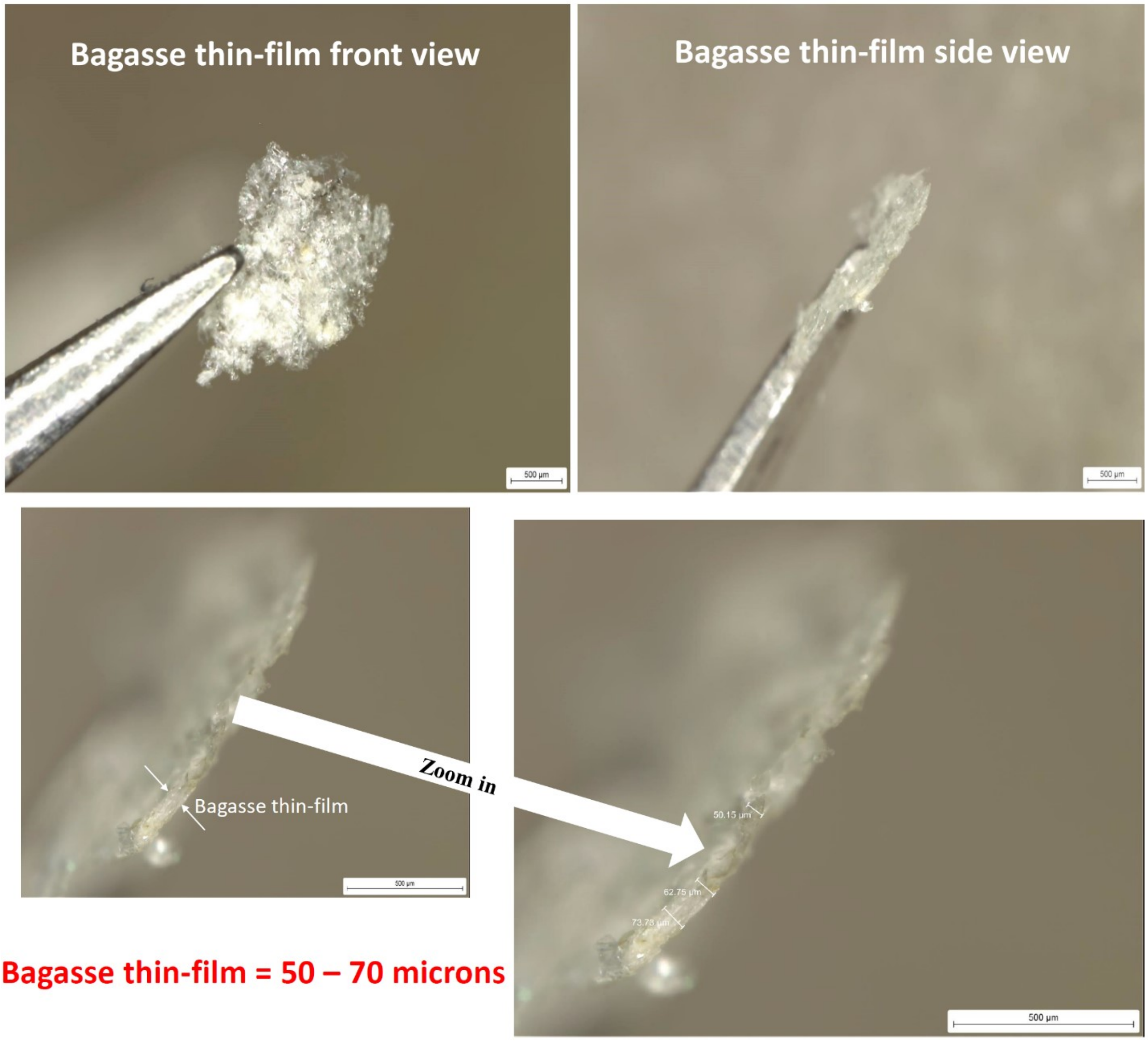

Figure S2: Digital microscopy images of Bagasse thin-films – front, side and zoomed in views.

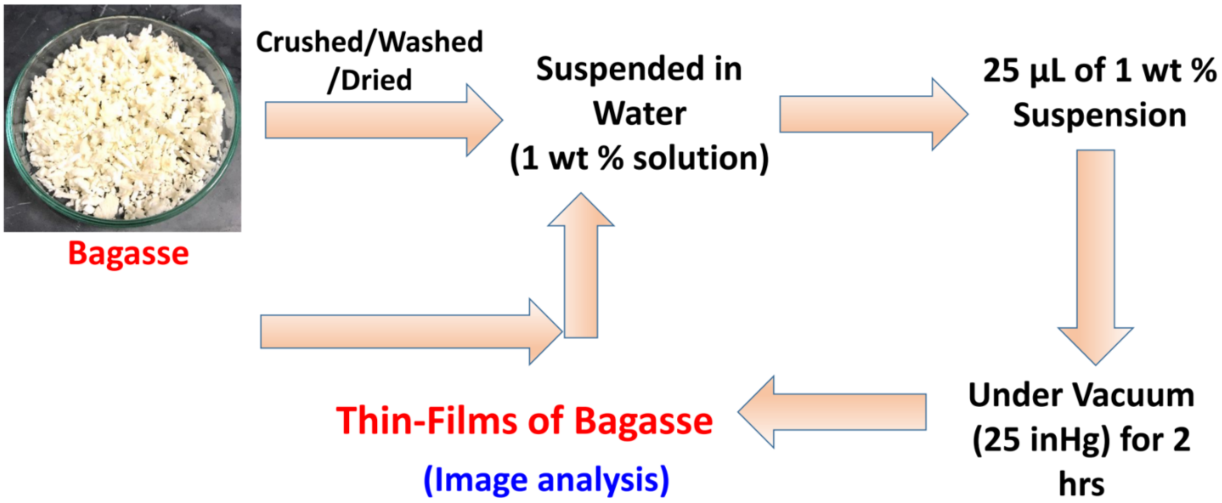

Figure S3: Schematic of the steps involved in thin-film preparation from crushed bagasse.

S.3 Minimal change in kinetics between LCCs with different linkage sites (C2, C3, C6)

Table S1: Inter- and Intra-moiety hydrogen bond lengths (Å) made by the cellobiose moiety when binding at C2, C3 and C6 positions with the lignin dimer.

Hydrogen bonds in cellulose dimer	**Bong length**		
	Reactant		
	LCC site -C2	**LCC site- C3**	**LCC site - C6**
C4-C3	2.42		2.42
C5'-C4'	2.29		2.29
C3-C2	2.36	2.53	2.36
O5-C3'	1.93	2.03	1.93
C3'-C4'	2.37	2.38	2.37
O5'-C6'	2.35	2.39	2.36
C2-C1		2.4	

In addition to the difference in activation barriers between pure cellobiose and cross-linked cellobiose in LCC, there are minimum differences in kinetics among various LCC molecules (LCC-C2, LCC-C3 and LCC-C6). C2 and C3 carbons in cellobiose are attached to a hydroxyl (-OH) group while C6 is attached to a hydroxymethyl (-CH3OH) group. It is interesting to note that the LCC linkage made at different carbon atoms (C2 and C6) exhibit similar activation barriers for ring opening mechanism (c.f. Figure 2). Moreover, for all three mechanisms, LCC-C3 has a marginally higher activation barrier. Activation barriers in cellobiose have been shown to be influenced by the hydrogen bonds made in the reaction environment [10,11]. Hydroxyl groups from neighboring cellulose sheets can catalyze and promote glycosidic C-O bond cleavage. Therefore, any conformational advantages in making such hydrogen bonds for LCCs linked at the C3 position to the sugar is investigated by identifying all intra- and inter-moiety hydrogen bonds made by cross-linked cellobiose in LCC. Their bond distances have been calculated and reported in Table 2. It can be noticed that for LCC bonding at the C2 and C6, 6 identical hydrogen bonds are identified with almost the same bond distances. However, when the lignin dimer and cellobiose are bound at the C3 position, there are fewer number of hydrogen bonds with longer (weaker) bond distances. This small difference in the hydrogen bonding made by cross-linked cellobiose between the three bonding sites can potentially explain the difference in their activation energies (c.f. Figure 2). The availability of only weaker hydrogen bonds in the reaction site for cellobiose cross-linked at C3, could lead to less favored TS for the three mechanisms without access to hydrogen bonds to stabilize it. However, the difference in barriers between LCCs binding at different positions is < 5 kcal/mol, indicating that the site of linking between cellobiose and lignin does not have a highly significant role in cellulose decomposition. On the other hand, the deviation in activation energy between the same mechanism in cellobiose and in LCC are huge, indicating that LCC linkage plays a prominent role in cellulose decomposition kinetics.

S.4 Molecule and system visualizations

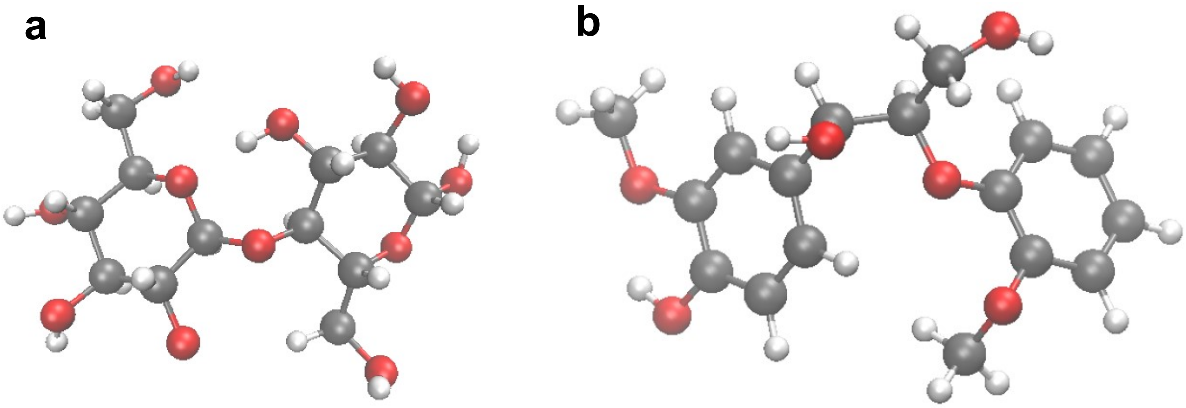

Figure S4: DFT calculated structures of model molecules (a) Cellobiose (cellulose dimer) (b) Quinone methide intermediate (lignin dimer).

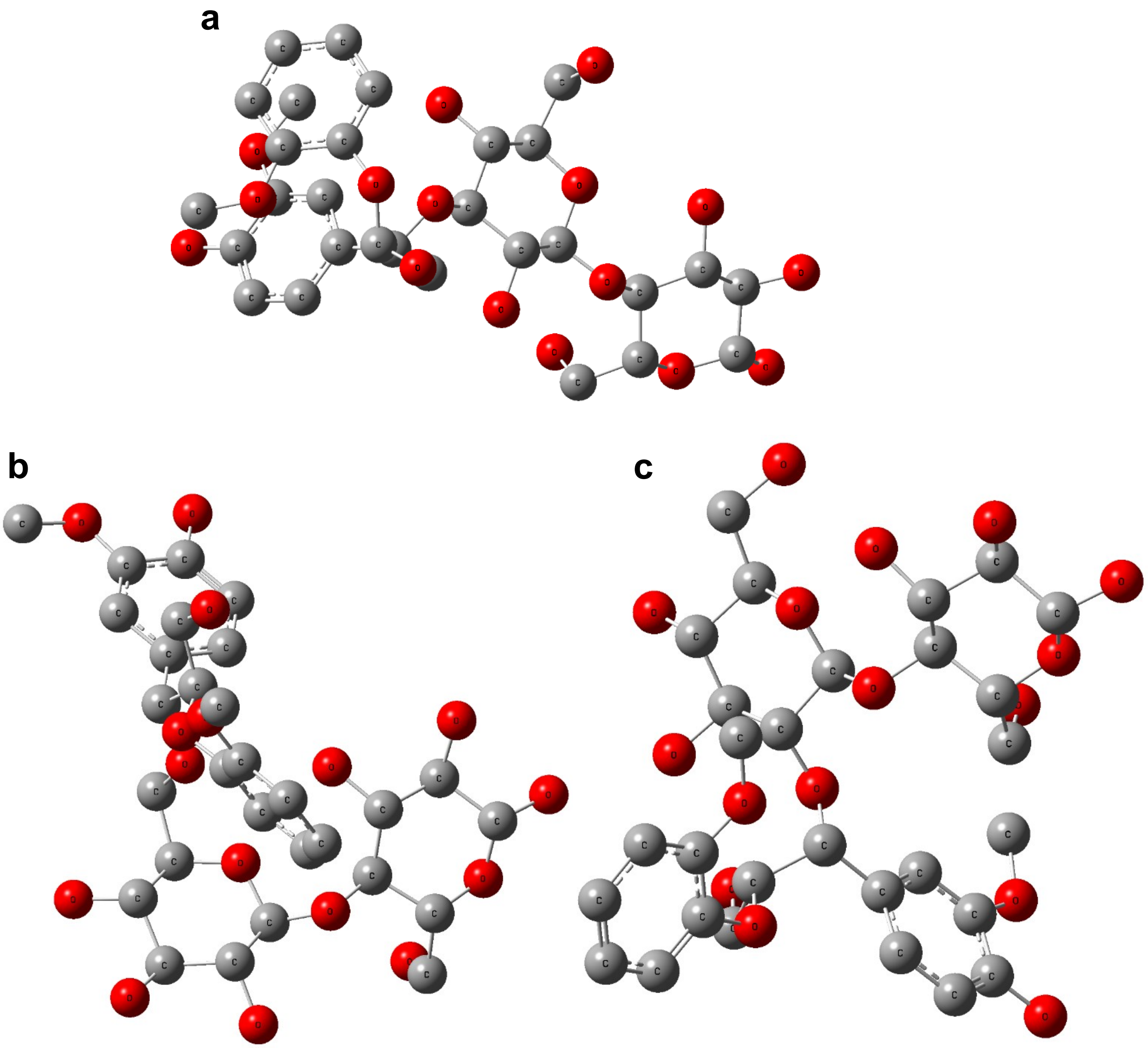

Figure S5: DFT calculated structures of Lignin-carbohydrate complex molecules (a) LCC-C3 (b) LCC-C6 (c) LCC-C2

S.5 Molecule Coordinates

Cellulose

C	14.95	10.4	14.25
C	19.06	13.32	15.52
C	19.47	14.69	14.99
C	21	14.81	15.02
C	16.38	9.86	14.09
C	16.5	8.33	14.23
C	17.97	7.87	14.07
C	18.89	8.73	14.94
C	18.65	10.21	14.6
C	22.25	11.5	14.13
C	21.36	12.52	14.85
C	19.84	12.24	14.75
H	19.26	13.23	16.59
H	14.54	10.15	15.22
H	19.08	14.84	13.98
H	14.29	9.96	13.51
H	21.37	14.82	16.05
H	16.65	10.07	13.08
H	16.11	8.02	15.19
H	18.26	7.96	13.02
H	18.66	8.57	15.99
H	18.83	10.33	13.52
H	21.88	10.49	14.29
H	22.27	11.68	13.06
H	21.65	12.55	15.89
H	19.59	12.3	13.69
H	15.38	6.91	13.5
H	20.67	16.64	14.85
H	18.48	6.03	13.79
H	15.29	12.23	14.87
H	20.43	8.4	13.75
H	17.52	12.19	15.35
H	23.54	11.17	15.53
H	18.11	15.51	16.12
O	15.75	7.74	13.19
O	21.6	13.8	14.3
O	21.26	16	14.43
O	18.07	6.53	14.5
O	14.94	11.81	14.09
O	20.26	8.39	14.69
O	17.33	10.52	14.94
O	19.55	10.98	15.33

O	17.68	13.13	15.26
O	23.56	11.57	14.67
O	19.02	15.71	15.86

Quinone Methide intermediate (Lignin dimer)

C	16.19	15.02	16.99
C	16.22	17.14	16.07
C	16.5	17.32	14.7
C	17.36	18.39	14.32
C	17.87	19.28	15.29
C	17.55	19.11	16.65
C	16.73	18.04	17.04
C	16.14	16.55	12.4
C	15.14	17.57	11.83
C	15.92	15.17	11.73
C	16.74	14	12.27
C	17.72	13.38	11.46
C	18.44	12.27	11.94
C	19.54	11.89	9.85
C	18.19	11.8	13.24
C	17.22	12.4	14.05
C	16.49	13.5	13.56
H	16.74	15.3	17.89
H	16.92	14.66	16.26
H	15.54	14.19	17.24
H	17.64	18.56	13.3
H	18.5	20.1	14.99
H	17.95	19.8	17.38
H	16.49	17.89	18.08
H	17.17	16.81	12.17
H	14.17	17.46	12.31
H	14.97	17.35	10.77
H	14.96	19.4	12.44
H	16.19	15.28	10.68
H	14.46	14.02	11.19
H	17.88	13.76	10.46
H	18.6	11.79	9.3
H	20.25	11.2	9.39
H	19.93	12.9	9.7
H	19.51	10.52	12.96
H	17.01	12.01	15.04
H	15.72	13.95	14.17
O	15.43	16.1	16.48
O	15.93	16.42	13.82

O	15.58	18.9	11.9
O	14.56	14.8	11.74
O	19.39	11.57	11.22
O	18.92	10.74	13.67

Lignin-carbohydrate complex (LCC) molecule

C	16.25	1.32	2.45
C	18.27	2.59	2.9
C	19.05	3.27	3.86
C	20.22	3.94	3.47
C	20.61	3.97	2.12
C	19.84	3.3	1.15
C	18.67	2.61	1.54
C	18.65	4.4	5.93
C	19.54	4.23	7.17
C	17.18	4.79	6.24
C	16.22	4.55	5.09
C	16	5.56	4.12
C	14.98	5.4	3.15
C	15.61	7.37	1.96
C	16.66	4.15	8.98
C	13.52	6.26	6.67
C	13.14	7.65	6.14
C	12.29	7.54	4.86
C	14.31	3.3	4.21
C	15.38	3.42	5.11
C	15.27	3.55	9.21
C	15.28	2.01	9.44
C	13.89	1.48	9.26
C	13.33	1.92	7.9
C	13.28	3.46	7.92
C	10.1	4.74	5.62
C	11.43	5.47	5.46
C	12.3	5.32	6.73
C	14.14	4.28	3.22
H	15.65	2.09	1.98
H	15.55	0.71	3.01
H	16.72	0.69	1.7
H	20.86	4.42	4.21
H	21.52	4.48	1.84
H	20.14	3.31	0.12
H	18.12	2.1	0.78
H	19.04	5.23	5.34
H	19.3	3.33	7.72

H	19.33	5.03	7.87
H	21.36	3.5	7.27
H	17.18	5.87	6.43
H	16.62	6.43	4.15
H	15.61	8.05	2.82
H	15.31	7.94	1.08
H	16.62	7.01	1.79
H	14.26	5.82	6.02
H	17.26	4.13	9.89
H	12.59	8.17	6.91
H	17.17	3.59	8.2
H	12.83	7.02	4.06
H	13.29	4.91	1.7
H	13.64	2.46	4.28
H	15.54	2.67	5.87
H	14.79	4.04	10.05
H	16.04	1.64	8.75
H	13.23	1.74	10.09
H	13.97	1.6	7.09
H	12.71	3.8	8.79
H	9.55	5.14	6.47
H	10.31	3.69	5.82
H	11.94	5.03	4.61
H	11.68	5.57	7.6
H	15.6	0.85	10.99
H	11.38	8.74	3.7
H	15.36	0.14	9.33
H	11.53	1.46	8.51
H	14.54	5.58	8.09
H	8.6	4.19	4.56
H	13.88	9.01	5.1
O	17.17	1.9	3.36
O	18.67	3.19	5.19
O	20.9	4.2	6.79
O	16.68	4.24	7.47
O	14.7	6.3	2.15
O	15.69	1.78	10.79
O	11.14	6.85	5.2
O	11.98	8.81	4.43
O	13.15	4.17	2.29
O	14.38	0.12	9.27
O	12.05	1.36	7.71
O	14.6	3.87	8
O	12.68	3.94	6.76
O	14.05	6.41	7.97

O	9.3	4.83	4.45
O	14.27	8.39	5.71

S.6 Sample Simulation files

CPMD Conformation Search (.inp file)

```
&INFO
Sample File
Conformation Search
&END

&CPMD
MOLECULAR DYNAMICS CP
RESTART ALL
TRAJECTORY SAMPLE XYZ
 1
NOSE IONS MASSIVE
 298.0d0   3000.0d0
NOSE ELECTRONS
 0.005  10000.0d0
MAXSTEP
200000
TIMESTEP
 4.0
EMASS
 300
FILEPATH
/scratch/"FILEPATH"
MEMORY BIG
&END

&SYSTEM
 SYMMETRY
  1
 ANGSTROM
 CELL
  15 1.0 1.0  0.0  0.0  0.0
 DUAL
  4.0
 CUTOFF
  70.0
 CHARGE
  0
&END

&DFT
```

```
 FUNCTIONAL REVPBE
 GC-CUTOFF
 1.0d-08
&END

&ATOMS
*C_MT_revPBE.psp KLEINMAN-BYLANDER
 LMAX=P LOC=P
  16
2.15111425      -3.20865339     -2.88012490
2.69188825      -2.53244939     -1.77306690
...
…
…
…
…
…
…

*O_MT_revPBE.psp KLEINMAN-BYLANDER
 LMAX=P LOC=P
  10
3.74176425      -3.21418439     -1.16875290
2.87807125      3.18503861      -2.85819090
...
…
…
…
…
…
…

*H_MT_revPBE.psp KLEINMAN-BYLANDER
 LMAX=S LOC=S
  24
3.27956225      3.46697461      -3.69719990
...
…
…
…
…
…
…

METADYNAMICS COLLECTIVE VARIABLES
DEFINE VARIABLES
2
TORSION 9 5 6 26 SCF 1.2 KCV 1.0 MCV 150
TORSION 6 26 16 13 SCF 1.0 KCV 1.0 MCV 150
```

```
END DEFINE
META_RESTART RFILE
LAGRANGE TEMPERATURE
298.0
LAGRANGE TEMPCONTROL
298.0 200.0
HILLS = 0.1 0.001
METASTEPNUM
2000
MOVEMENT CHECK
0.02
MINSTEPNUM INTERMETA
100
MAXSTEPNUM INTERMETA
200
CHECK DELAY
10
METASTORE NO TRAJECTORY
20 1 100000
MONITOR
1
END METADYNAMICS

&END
```

Gaussian TS Search (.com file)

```
%nprocshared=16

%mem=64GB

%chk=ts.chk

# opt=(calcfc,ts,noeigen) freq=noraman rm062x/6-311+g(d) temperature=500


Title Card Required


0 1

 C(PDBName=C,ResName=0,ResNum=0_C)                 -2.71946900    0.83900900   -0.57632600

 C(PDBName=C,ResName=0,ResNum=0_C)                 -4.18082800    0.86415500   -1.01235500

 C(PDBName=C,ResName=0,ResNum=0_C)                 -2.48615500   -0.00919500    0.67097900

 ...
…
…
…
…
```

```
…
…
 O(PDBName=O,ResName=0,ResNum=0_C)                   -4.37911900    1.78372400   -2.05522500
 O(PDBName=O,ResName=0,ResNum=0_C)                   -5.62270300   -3.84564400    1.13034200
 O(PDBName=O,ResName=0,ResNum=0_C)                   -4.00494100   -5.37221200   -0.28986100
 ...
…
…
…
…
…
…
 H(PDBName=H,ResName=0,ResNum=0_C)                    2.44885500    1.92649200    3.69237600
 H(PDBName=H,ResName=0,ResNum=0_C)                    2.08825100    2.58482400    0.73324000
 H(PDBName=H,ResName=0,ResNum=0_C)                   -0.19463800    2.82466200    4.22837400
 H(PDBName=H,ResName=0,ResNum=0_C)                    1.38525900    4.56801700   -3.05128700
...
…
…
…
…
…
…
```

References

(1) Bridgwater, A. V; Meier, D.; Radlein, D. An Overview of Fast Pyrolysis of Biomass. *Org. Geochem.* **1999**, *30* (12), 1479–1493. https://doi.org/https://doi.org/10.1016/S0146-6380(99)00120-5.

(2) Zhang, Y.-H. P.; Ding, S.-Y.; Mielenz, J. R.; Cui, J.-B.; Elander, R. T.; Laser, M.; Himmel, M. E.; McMillan, J. R.; Lynd, L. R. Fractionating Recalcitrant Lignocellulose at Modest Reaction Conditions. *Biotechnol. Bioeng.* **2007**, *97* (2), 214–223. https://doi.org/10.1002/bit.21386.

(3) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys.Rev.Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(4) Zhang, Y.; Yang, W. Comment on ``Generalized Gradient Approximation Made Simple''. *Phys. Rev. Lett.* **1998**, *80* (4), 890. https://doi.org/10.1103/PhysRevLett.80.890.

(5) Guo, Z.; Wang, J.; Qin, F.; Ge, L.; Li, Z.; Shen, W. Facile Adjusting of a Right-Handed Helical Structure of Cellulose-Based Carbon Material for Chiral Separation. *ACS Sustain. Chem. Eng.* **2020**, *8* (8), 3401–3411. https://doi.org/10.1021/acssuschemeng.9b07540.

(6) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *\jcp* **1984**, *81* (1), 511–519. https://doi.org/10.1063/1.447334.

(7) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695–1697. https://doi.org/10.1103/PhysRevA.31.1695.

(8) Jia, W.; Wang, H.; Chen, M.; Lu, D.; Lin, L.; Car, R.; E, W.; Zhang, L. Pushing the Limit of Molecular Dynamics with Ab Initio Accuracy to 100 Million Atoms with Machine Learning. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*; SC '20; IEEE Press, 2020.

(9) Iannuzzi, M.; Laio, A.; Parrinello, M. Efficient Exploration of Reactive Potential Energy Surfaces Using Car-Parrinello Molecular Dynamics. *Phys. Rev. Lett.* **2003**, *90*, 238302. https://doi.org/10.1103/PhysRevLett.90.238302.

(10) Padmanathan, A. M. D.; Mushrif, S. H. Pyrolytic Activation of Cellulose: Energetics and Condensed Phase Effects. *React. Chem. Eng.* **2022**, *7* (5), 1136–1149. https://doi.org/10.1039/D1RE00492A.

(11) Maliekkal, V.; Maduskar, S.; Saxon, D. J.; Nasiri, M.; Reineke, T. M.; Neurock, M.; Dauenhauer, P. Activation of Cellulose via Cooperative Hydroxyl-Catalyzed Transglycosylation of Glycosidic Bonds. *ACS Catal.* **2019**, *9* (3), 1943–1955. https://doi.org/10.1021/acscatal.8b04289.