Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2022

Thermal fluctuations and conformational effects on NMR parameters in β-O-4 lignin dimer from QM/MM and machine-learning approaches.

[a]Sonia Milena Aguilera-Segura, [b]Dominik Dragún, [a]Robin Gaumard, [a]Francesco Di Renzo, [c]Irina Malkin Ondík ,[a]Tzonka Mineva*

[a.]ICGM, Univ Montpellier, CNRS, ENSCM, Montpellier, France

[b]FIIT STU in Bratislava, Ilkovičova 2, 842 16 Bratislava, Slovakia

[c]MicroStep-MIS spol. s.r.o. Čavojského 1, 84104 Bratislava, Slovakia

* To whom correspondence should be addressed
Email: tzonka.mineva@enscm.fr

Electronic Supplementary Information

S1. Supporting Figures

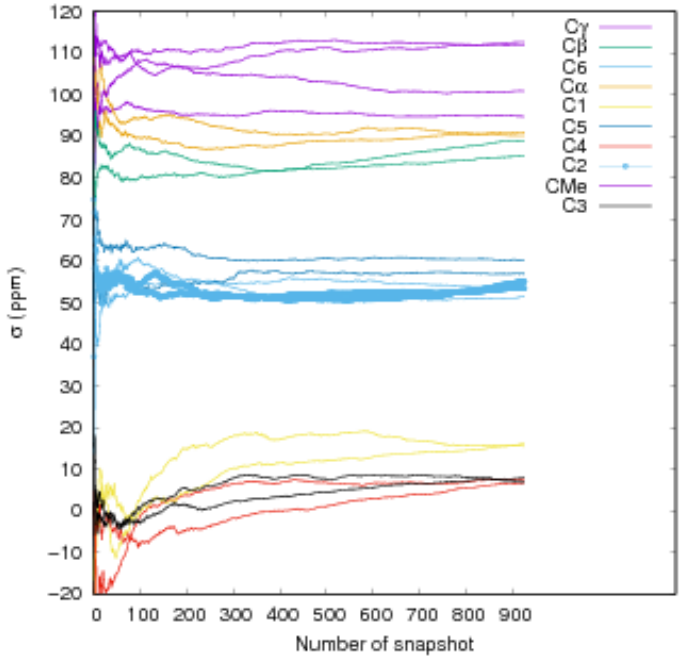

Fig. S1. σ(^{13}C) rolling averaged (step=1) values of β-O-4 guaiacyl dimer model of lignin in water solvent.

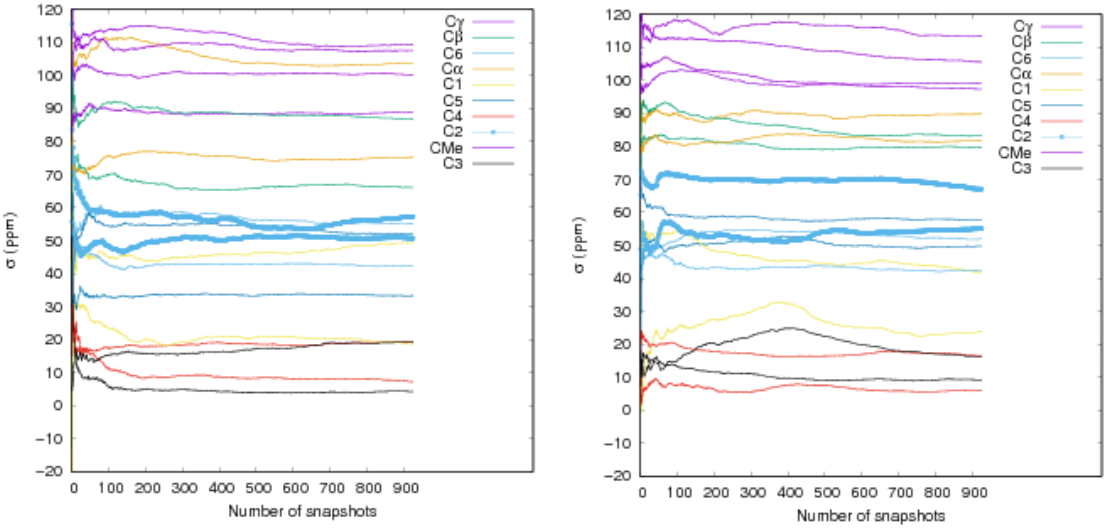

Fig. S2 σ(^{13}C) rolling averaged (step=1) values of β-O-4 guaiacyl dimer model of lignin in pure ethanol (left panel) and 75 wt% ethanol-water (right panel) solvents.

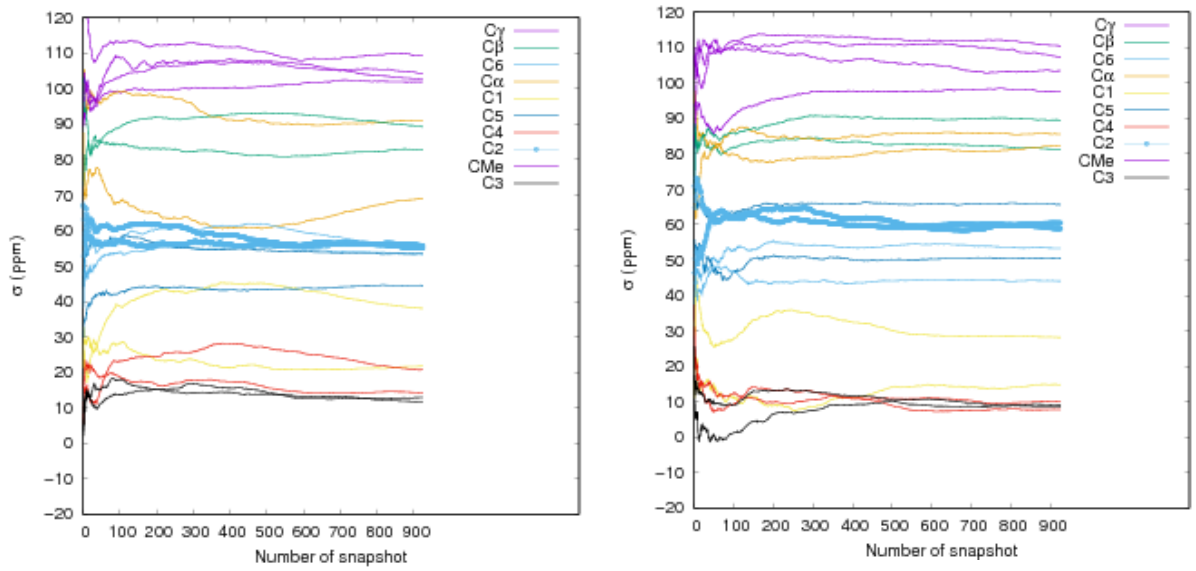

Fig. S3 $\sigma(^{13}C)$ rolling averaged (step=1) values of β-O-4 guaiacyl dimer model of lignin in pure acetonitrile (left panel) and in 75wt% acetonitrile-water (right panel) solvents.

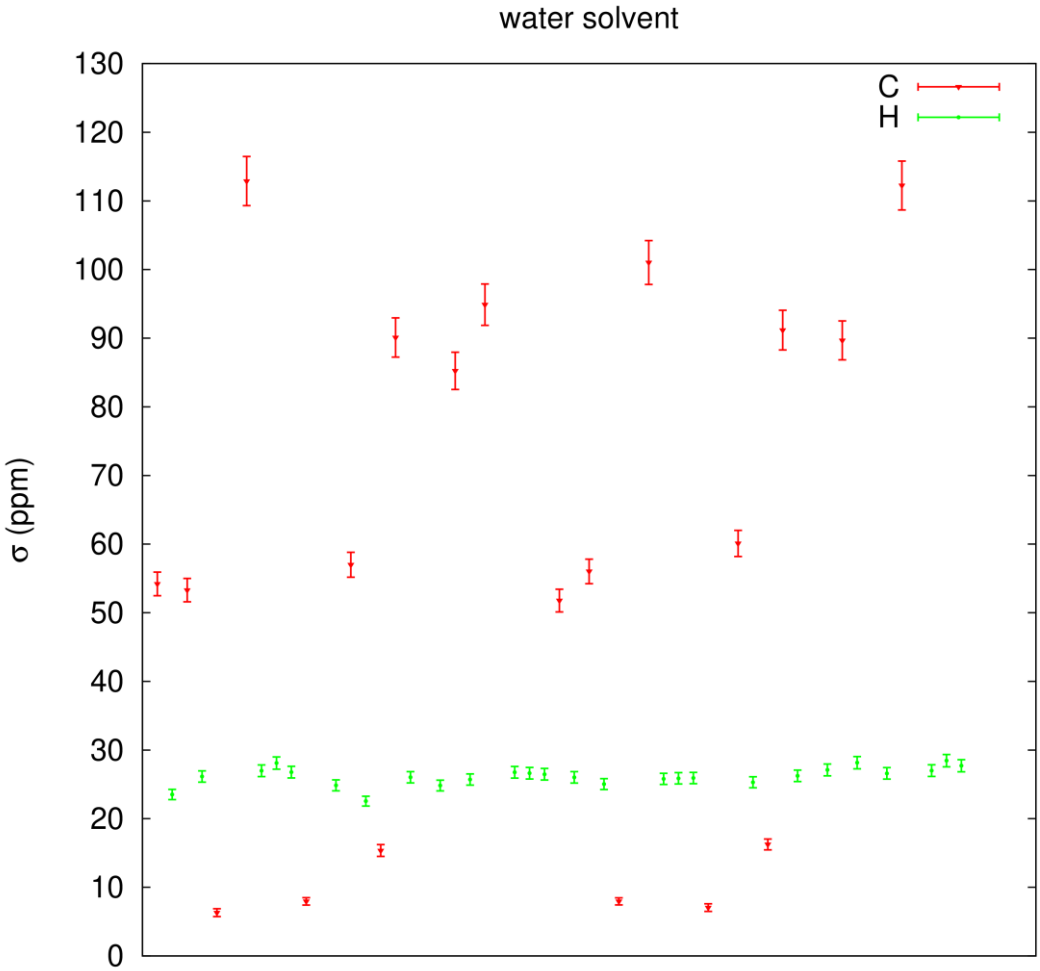

Figure S4. Averaged σ_{iso} (^{13}C) (red symbols) and σ_{iso} (^{1}H) (green symbols) with the standard deviation errors of β-O-4 guaiacyl dimer model of lignin in water solvent.

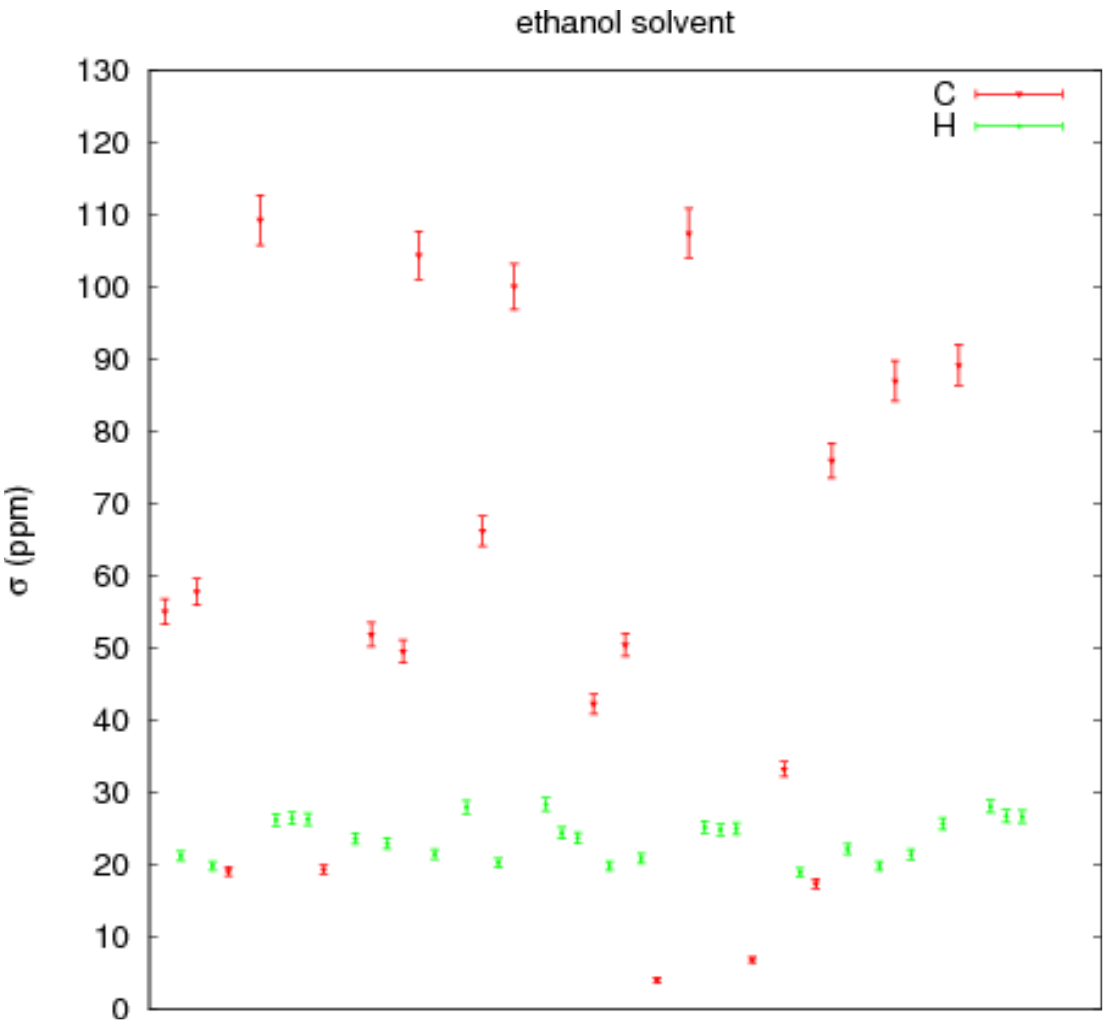

Figure S5. Averaged σ_{iso} (^{13}C) (red symbols) and σ_{iso} (^{1}H) (green symbols) with the standard deviation errors of β-O-4 guaiacyl dimer model of lignin in ethanol solvent.

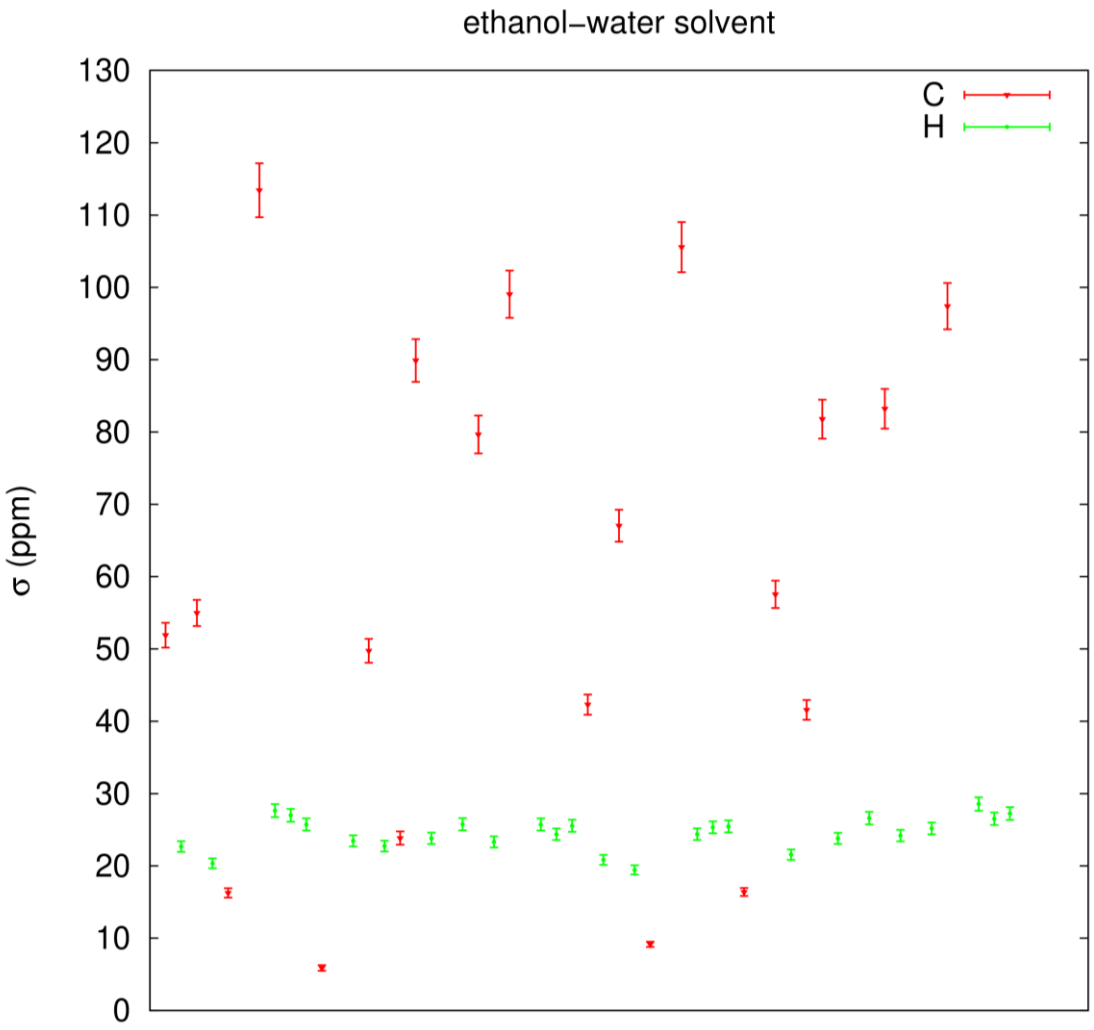

Figure S6. Averaged σ_{iso} (^{13}C) (red symbols) and σ_{iso} (^{1}H) (green symbols) with the standard deviation errors of β-O-4 guaiacyl dimer model of lignin in ethanol-water solvent.

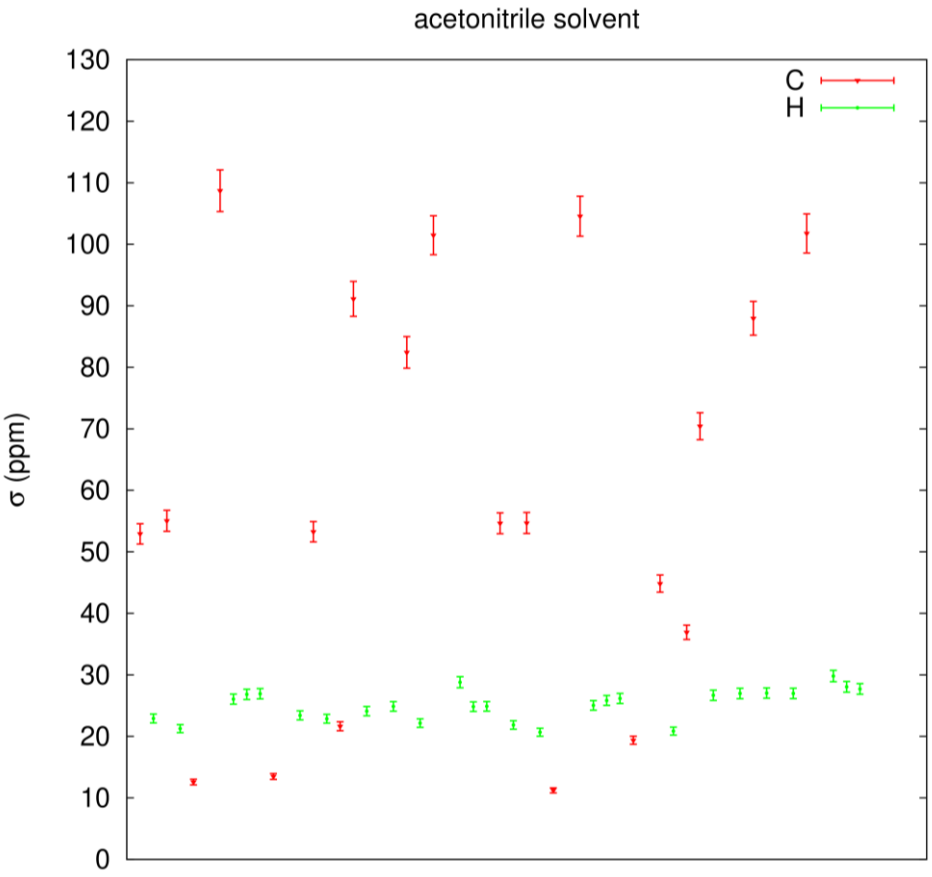

Figure S7. Averaged σ_{iso} (^{13}C) (red symbols) and σ_{iso} (^{1}H) (green symbols) with the standard deviation errors of β-O-4 guaiacyl dimer model of lignin in acetonitrile solvent.

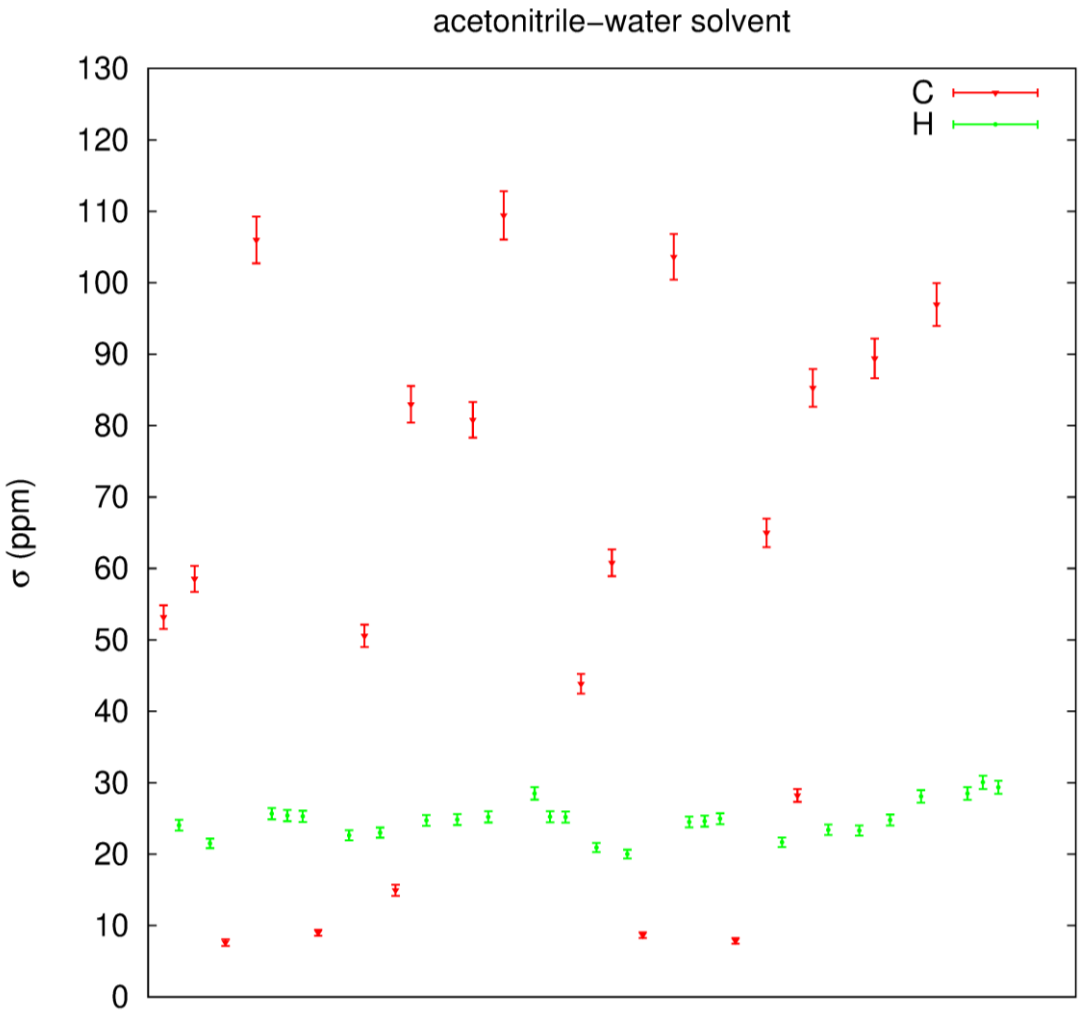

Figure S8. Averaged σ_{iso} (^{13}C) (red symbols) and σ_{iso} (^{1}H) (green symbols) with the standard deviation errors of β-O-4 guaiacyl dimer model of lignin in acetonitrile-water solvent.

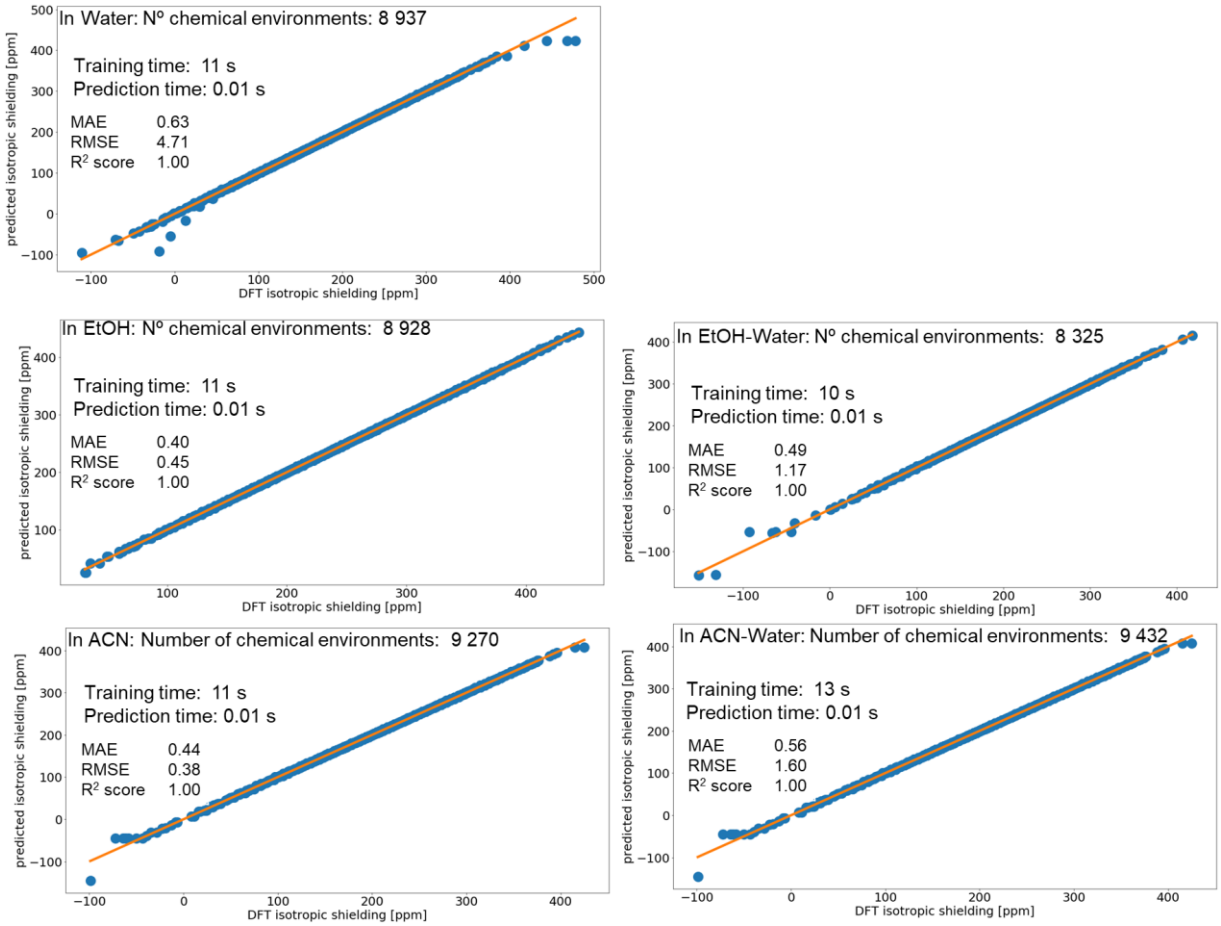

Figure S9. Correlation between the σ_{iso} (^{17}O) values predicted with ML – GBR model and those computed with DFT for the G-b-G lignin dimer in water (top); ethanol (middle left), aqueous 75wt% ethanol (middle right) - acetonitrile (down left), and 75wt% acetonitrile –water (down right). Number of chemical environments, training and prediction time, mean averaged error (MAE), and root mean square errors (RMSE) are reported for each solvent.

The outliers in Figure S9 are as follows:

In water:

Outlier 1 element: O coordinates: 1.315819 1.116589 5.51375
Real value = 478.16 ppm； predicted value = 423.12 ppm
Frame: *lig-Water-frame_93*
Outlier 2 element: O coordinates: 1.356145 1.377420 5.489409
Real value: 443.66 ppm； predicted value: 423.47 ppm
Frame: *lig-Water-frame_99*
Outlier 3 element: O coordinates: 1.367365 1.436707 5.530867
Real value = 467.97 ppm; predicted value = 423.18 ppm
Frame: *lig-Water-frame_100*
Outlier 4 element: O coordinates: 4.268215 0.492121 -1.225273
Real value = 12.84 ppm; predicted value = -16.82 ppm
Frame: *lig-Water-frame_720*

Outlier 5 element: O coordinates: 1.369903 3.840386 5.132885
Real value = -4.77 ppm; predicted value = -54.97 ppm
Frame: *lig-Water-frame_928*
Outlier 6 element: O coordinates: 1.421624 3.825311 5.161767
Real value= -18.54 ppm; predicted value = -95.92 ppm
Frame: lig-Water-frame_932
Outlier 7 element: O coordinates: 1.557944 4.066079 4.528399
Real value = -110.97 ppm； predicted value: -91.84 ppm
Frame: *lig-Water-frame_957*

In ethanol – no outliers found

In ethanol-water:

Outlier 1: element: O coordinates: 5.740399 -3.811082 5.029433
Real value = -131.22 ppm； predicted value = -156.02 ppm
Frame: *lig-etoh-frame_262*
Outlier 2: element: O 5.679339 -3.851260 5.043047
Real value = -92.86 ppm; predicted value = -52.90 ppm
Frame: lig-etoh-frame_261
Outlier 3: element: O coordinates: 5.731097 -4.055655 4.998866
Real value = -66.57 ppm; predicted value = -56.20 ppm

In acetonitrile – no outliers found

In acetonitrile-water:

Outlier 1 element: O coordinates: -3.841373 3.569311 -4.411906
Real value = -72.29 ppm; predicted value = -44.57ppm
Frame: *lig-ACN-Water-frame_63*
Outlier 2 element: O coordinates: 6.915452 0.071544 -0.501983
Real value = 425.06 ppm； predicted value = 408.10 ppm
Frame: *lig-ACN-Water-frame_64*
Outlier 3 element: O coordinates: -4.141404 3.887675 -4.337876
Real value = -98.84 ppm; predicted value = -145.60 ppm
Frame: *lig-ACN-Water-frame_98*
Outlier 4 element: O coordinates: -4.179555 3.942128 -4.343942
Real value = -64.22 ppm; predicted value = -44.86
Frame: *lig-ACN-Water-frame_107*
Outlier 5 element: O; coordinates: -2.551616 4.481367 -4.553537
Real value = -61.47 ppm; predicted value = -44.74
Frame: *lig-ACN-Water-frame_405*

Number of chemical environments: 274 340

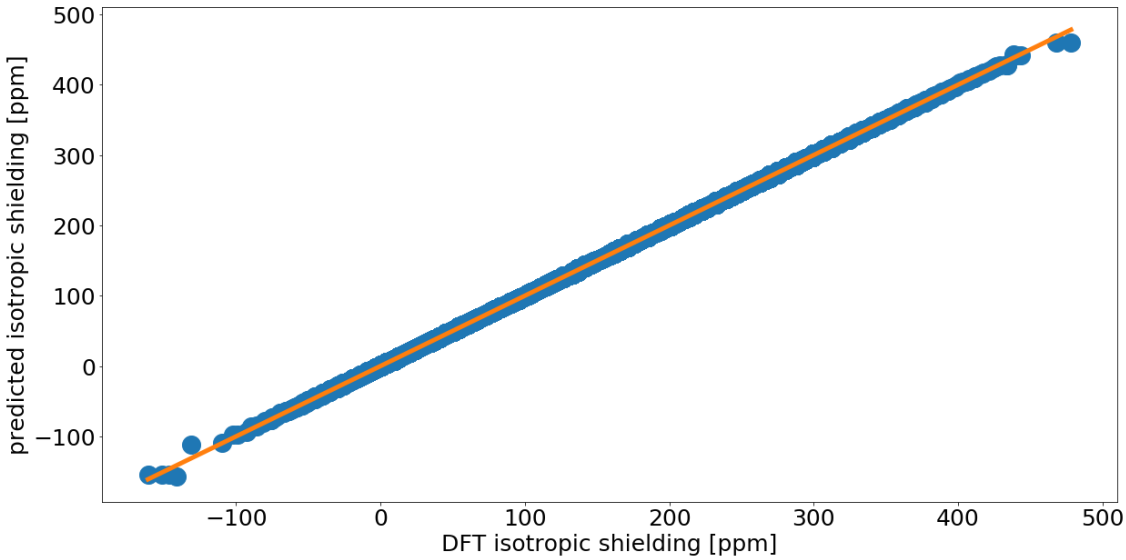

Figure S10. Correlation between the isotropic shielding values of ^{17}O, ^{13}C and ^{1}H predicted with ML-GBR model and those computed with DFT in all the solvents. The training and prediction time is, respectively, 555 and 0.55 s. MAE=0.35 ppm; RMSE=0.30 ppm and the correlation parameter R^2=1.

Outlier 1: element: O coordinates: 5.740399 -3.811082 5.029433
Real value = -131.22 ppm； predicted value = -156.36 ppm
Frame: *lig-etoh-frame_262*

Outlier 2: element: O coordinates: 1.315819 1.116589 5.513751
Real value = 478.16 ppm; predicted value = 460.03 ppm
Frame: *lig-Water-frame_93*

Outlier 3: element: C coordinates: -0.009662 1.964282 3.886950
Real value = -141.31 ppm； predicted value = -111.45 ppm
Frame: lig-Water-frame_876

S2. Supporting Tables

Table S1. Dynamic (time-averaged) ^{13}C chemical shifts in ppm in water, ethanol (EtOH), acetonitrile (ACN), ethanol-water (EtOH-water) and acetonitrile-water (ACN-water) solvents in ring A and ring B in G-b-G dimer model of lignin. The reference isotropic shielding is the DFT computed for TMS with the same level of theory ^{13}C σ_{iso} =162.72 ppm

Carbon atom	water	EtOH	ACN	EtOH-water	ACN-water
2A	109.42	104.92	107.69	107.72	104.22
2B	106.72	112.22	108.02	95.72	101.92
5A	105.72	110.82	109.42	113.02	112.12
5B	102.62	129.32	117.92	105.12	97.72
6A	108.52	107.72	109.82	110.82	109.52
6B	110.92	120.42	108.12	120.42	118.92
1A	147.32	141.12	141.12	138.92	147.82
1B	146.52	125.82	125.82	121.12	134.52
3A	156.42	143.52	150.12	146.42	155.12
3B	154.72	158.62	151.52	153.52	154.02
4A	154.72	143.12	149.22	156.82	153.72
4B	155.72	145.32	143.32	146.32	154.82
AMe	50.52	53.52	54.02	49.32	56.72
BMe	61.72	55.32	58.22	57.12	59.12