Electronic Supplementary Material (ESI) for Chemical Science.


# Supplementary material

# From cellulose to kerogen: molecular simulation of a geological process

Lea Atmani, Christophe Bichara, Roland J.-M. Pellenq, Henri Van Damme, Adri C. T. van Duin, Zamaan Raza, Lionel A. Truflandier, Amaël Obliger, Paul G. Kralert, Franz J. Ulm, and Jean-Marc Leyssale*

E-mail: jean-marc.leyssale@u-bordeaux.fr.

# Complementary results on kerogen growth

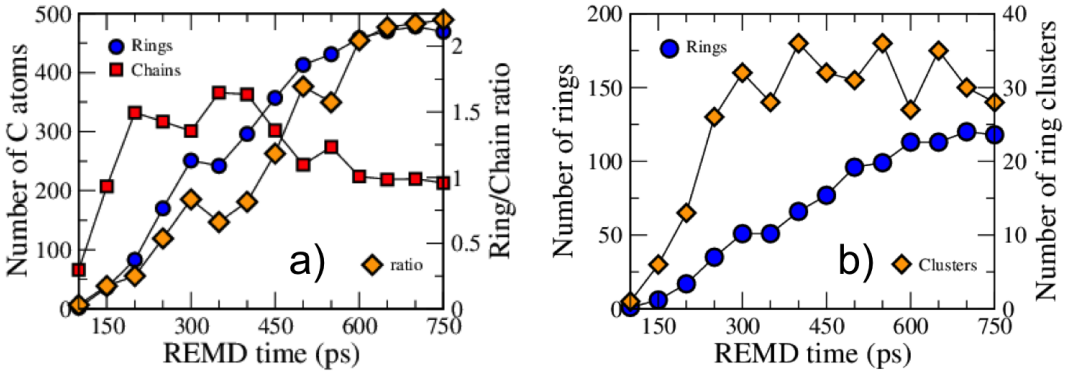


Figure S1: Evolution with REMD time of (a) the numbers of atoms in rings and chains and (b) the number of rings and ring clusters. All data are from the 423 K replica in the first REMD simulation.

Fig. S1 highlights the growth of the macromolecular (or kerogen) carbon-rich phase during the first REMD simulation. We observe (Fig. S1(a)) the increase in the total number of carbon atoms (first in chains then in rings) in this phase up to around 400 ps, followed by the conversion from chain C atoms to ring C atoms. The total number of ring clusters increases and saturates around 300 ps.

# Complementary results on kerogen maturation

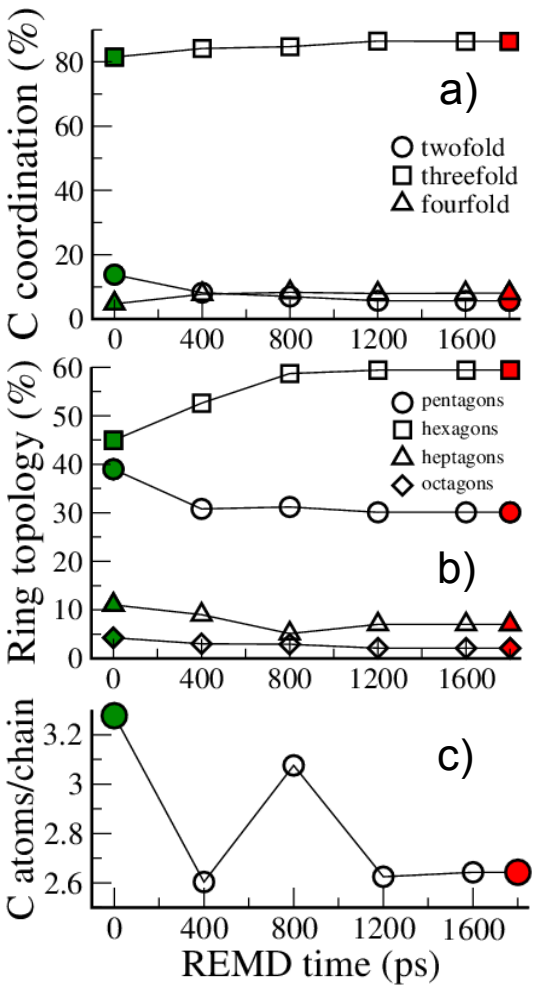


Figure S2: Evolution with REMD time of (a) the carbon atoms coordination, (b) the ring topology and (c) the size of chain clusters (in number of carbon atoms). All data are from the 423 K replica in the second REMD simulation.

Fig. S2 describes the structural evolution, or maturation, of the kerogen during the second REMD simulation (*i. e.*, after removal of the fluid). It shows that the hybridization of carbon atoms does not evolve much but that the fraction of 6-member rings (hexagons) significantly increases. The size of aliphatic clusters (chains) is slightly reduced.

## Pore size distributions of the kerogen models

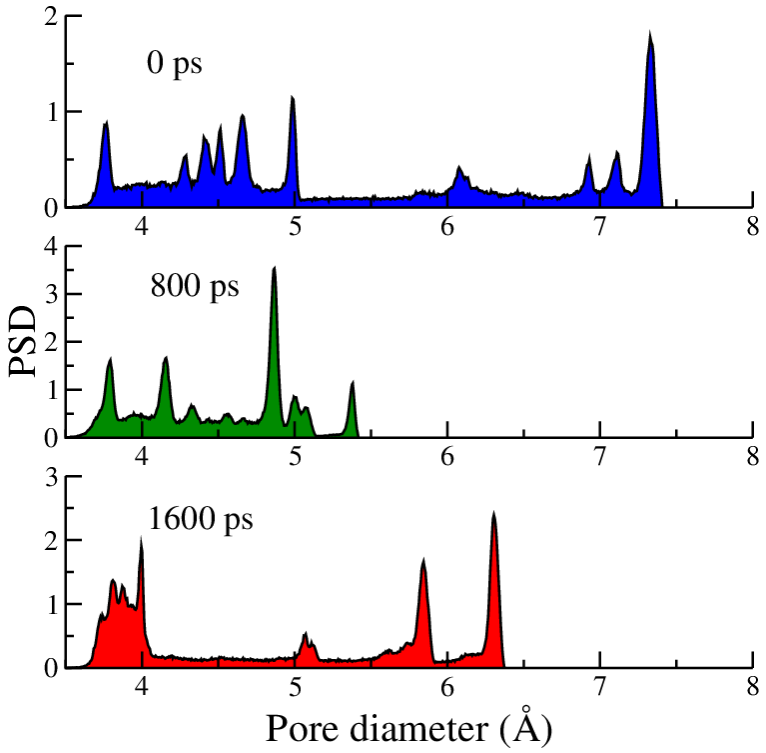


Figure S3: Pore size distribution at different REMD times. Data computed using the Gelb-Gubbins technique[1] limited to the pore space accessible to $CO_2$. All data are from the 423 K replica in the second REMD simulation (0 and 1800 ps correspond to kerogen models $A$ and $B$, respectively).

Figure S3 shows the evolution of the pore size distribution (PSD) during the second REMD simulation. In agreement with the evolution of porosity (Fig. 6(b)), a reduction in pore size is first observed, followed by a slight increase towards the end of the simulation. In any cases, PSDs remain in the 3.5-7.5 Å range, characteristic of ultramicropores.

# Chemical description of the largest ring clusters

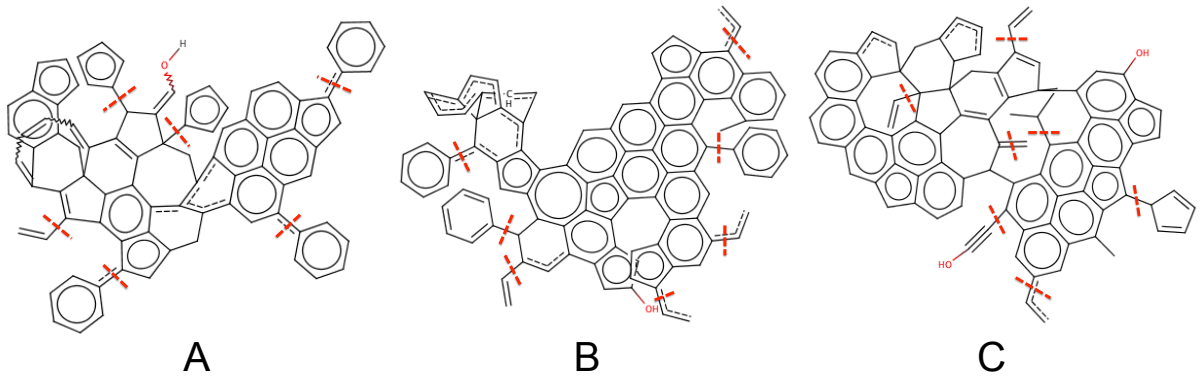


Figure S4: Chemical structure of the three largest ring clusters in model B. Cluster limits are displayed with red dashed lines. The structures were drawn using Marvin (version 17.23.0, 2017, ChemAxon (http://www.chemaxon.com) from the observation of the actual clusters in the model and are as close as possible from the exact structures.

# Gas phase potential energy surface

Table S1: 0K energies (with respect to isolated atoms) of the main constituents of the fluid phase computed with ReaxFF and PW-DFT (all results are expressed in kcal/mol). Errors with respect to DFT data $\left(\frac{E_{DFT}-E_{Reax}}{E_{DFT}} \times 100\right)$ are given in parentheses.

| Formula | Usual name | ReaxFF | PW-DFT | (Err(%)) |
|---|---|---|---|---|
| $H_2O$ | water | -248.9 | -241.4 | (-3.1) |
| $H_2$ | hydrogen | -108.7 | -104.7 | (-3.8) |
| CO | carbon monoxide | -261.8 | -275.8 | (5.1) |
| $CO_2$ | carbon dioxide | -374.1 | -431.1 | (13.2) |
| $CH_4$ | methane | -404.6 | -421.4 | (4.0) |
| $CH_4O$ | methanol | -530.0 | -528.3 | (-0.3) |
| $CH_4O_2$ | methanediol | -648.9 | -650.1 | (0.18) |
| $C_2H_2O_2$ | acetylenediol | -673.7 | -636.6 | (-5.8) |
| $C_2H_2O$ | ethynol | -538.3 | -529.4 | (-1.7) |
| $H_3O^+$ | hydronium | -304.8[a] | -98.3 | (-210) |
| $C_2H_3O^+$ | ketenium | -604.2[a] | -414.4 | (-45.8) |

[a] ReaxFF energy correpsonds to the equivalent neutral molecule and cannot be compared to DFT results.

Plane wave DFT calculations and structure relaxations were performed with the PWscf code as part of the Quantum-Espresso package[2] using the PBE functional.[3] A plane wave cutoff of 60 Ry was employed for the electronic wavefunctions and 480 Ry for the augmentation charge density, using the ultra-soft pseudopotentials[4] recommended by Castelli et al.[5] The Brillouin zone was sampled at the $\Gamma$ point only. For isolated molecule calculations a cubic supercell of size 30 Å was considered. The electronic self-consistent field calculations were converged with an energy tolerance of $10^{-8}$ Ry. Structure optimizations were perfomed without symmetry contraints using a tight convergence threshold of $0.5 \times 10^{-4}$ Ry/Bohr on the root mean square of the atomic forces. Final total energies were computed using the same box size including the Makov-Payne (MP) correction[6] to attenuate further periodic image interactions. The conventional static jellium background was used to compensate the charge of ionic molecules.

# DFT evaluation of (small-scale) REMD configurations

In this section we evaluate the effects of DFT relaxation on a few configurations taken from a REMD simulation of the decomposition of a 420 atoms cellulose crystal. The 370 ps long REMD simulation was performed with exactly the same MD simulation parameters (interatomic potential, time step, thermostat parameters). Temperature was evenly spaced (i.e. by steps of 100 K) among 33 replicas, ranging from 300 to 3500 K. The number of replicas and their spacing being less sensitive because of broader energy distributions in smaller sized systems. Seven configurations were extracted from the 300 K replica trajectory and relaxed with DFT. The plane wave DFT calculations were performed using the same computational parameters as those detailed for the molecular case (see above) with the exception of the supercell approach and the MP correction. The lattice parameters and ionic positions were allowed to relax to a target pressure of 0.5 kbar, where the cell was constrained to be orthorhombic. In this case, geometry relaxations were converged using a coarser threshold of at least 0.04 Ry (12.5 kcal/mol), $10^{-3}$ Ry/Bohr and 0.1 kBar for the total enthalpy, atomic forces and pressure, respectively.

Table S2: Evolution of chemical bonding upon DFT relaxation of REMD (Reax) configurations. I, B and F respectively stand for the initial (before relaxation) number of bonds, the number of broken bonds and the number of formed bonds after relaxation.

| Time (ps) | C-C | | | C-O | | | C-H | | | O-H | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I | B | F | I | B | F | I | B | F | I | B | F |
| 29 | 76 | 1 | 0 | 77 | 1 | 1 | 102 | 1 | 2 | 97 | 9 | 11 |
| 57 | 78 | 0 | 3 | 66 | 3 | 0 | 78 | 0 | 8 | 118 | 12 | 8 |
| 86 | 84 | 0 | 3 | 56 | 1 | 2 | 75 | 0 | 5 | 124 | 10 | 8 |
| 114 | 94 | 3 | 3 | 50 | 2 | 0 | 63 | 0 | 6 | 130 | 15 | 14 |
| 143 | 100 | 0 | 0 | 44 | 0 | 0 | 59 | 2 | 2 | 132 | 5 | 8 |
| 171 | 96 | 1 | 1 | 41 | 0 | 1 | 60 | 0 | 5 | 133 | 12 | 12 |
| 200 | 105 | 0 | 2 | 40 | 0 | 1 | 50 | 0 | 5 | 143 | 16 | 13 |

Fig. S5 shows snapshots of these configurations before and after DFT relaxation. The overall structure of the different configurations is relatively well conserved upon relaxation. However, as can be seen in table S2, chemical bonding can be slightly modified. This does not

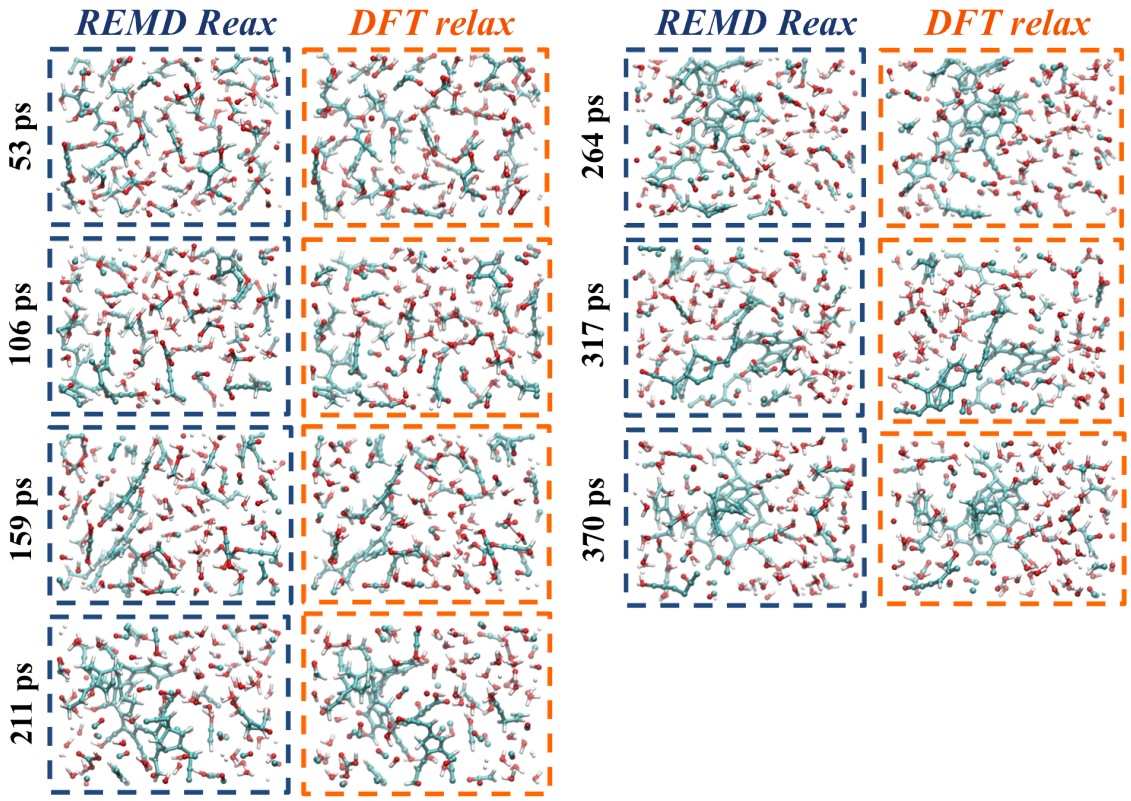


Figure S5: Snapshots of seven configurations taken from the 300 K replica of a small-scale (420 atoms) REMD simulation are compared before and after DFT relaxation (color code: C (cyan), O (red) and H (white)).

concern much the carbon backbone as well as C-O bonds, which remain almost unaffected. However, we observe many creations of C-H bonds essentially through the transfer of H atoms from hydronium and ketenium ions to initially sp hybridized C atoms. This induces breaking many O-H bonds. Proton transfer between hydronium and water is also responsible for numerous destructions and creations of O-H bonds.

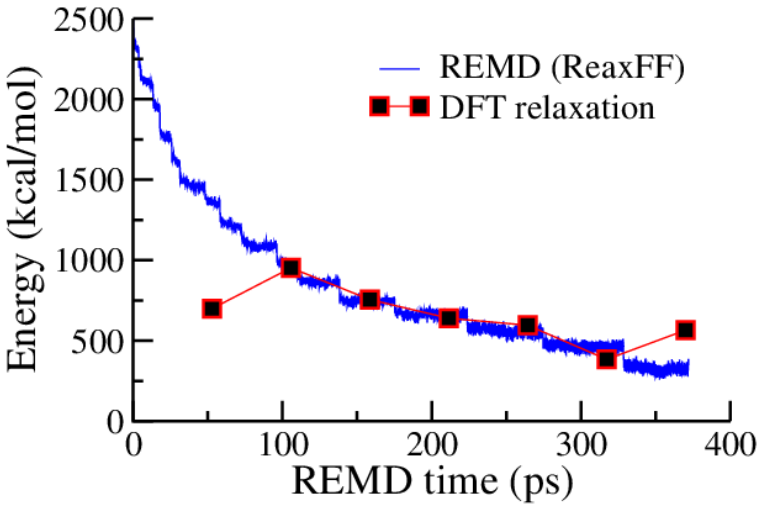


Figure S6: DFT enthalpies obtained after relaxation, superimposed to the evolution of the Reax potential energy with REMD time from the 300 K replica. Energy origins have been adjusted to improve the superimposition.

Finally, Fig. S6 compares the DFT enthalpies of these seven configurations to the evolution of the Reax potential energy during the REMD simulation. As can be seen, the evolutions of these two energies compare reasonably well, aside from the initial configuration, significantly stabilized by the DFT relaxation because of significant changes in bonding. These results confirm, in any case a global progressive decrease in energy during the decomposition process.

# References

(1) Gelb, L. D.; Gubbins, K. E. Pore Size Distributions in Porous Glasses: A Computer Simulation Study. *Langmuir* **1999**, *15*, 305–308.

(2) Giannozzi, P. et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matt.* **2009**, *21*, 395502.

(3) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(4) Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **1990**, *41*, 7892.

(5) Lejaeghere, K. et al. Reproducibility in density functional theory calculations of solids. *Science* **2016**, *351*, aad3000.

(6) Makov, G.; Payne, M. C. Periodic boundary conditions in ab initio calculations. *Phys. Rev. B* **1995**, *51*, 4014–4022.