ASSESSMENT OF COVALENT ORGANIC FRAMEWORK STACKING

Contents

List of Figures

List of Tables

Introduction

Covalent Organic Frameworks are a novel class of materials made from self assembled small organic molecules, imagined and first synthesised by Côté et al. in 2005[1]. First COFs were two-dimensional polymers formed by a dehydration condensation reaction of boronic acid ($RB(OH)_2$) into boroxin rings ($R_3(BO)_3$). Latter, more diverse ligands and structures were elaborated, by adding vertical linkers to yield three- dimensional polymers or the use of triazine or imine condensation instead of boronic ones; imine condensation leading to linear condensation unlike boronic and triazinic ones which leads to tert-substituted cycles.

These materials drew an increasing interest because of their lightness, porosity and optical properties all the while being easy to functionalise, and are therefor excellent candidates for gas separation and storage, electrochemistry, heterogeneous catalysis, and light harvesting; but this very high tunability of the material also makes the number of ligand combination extremely large.

Since experimental chemistry would be unable to assess such amounts of structures, in part due to technical limitations like the high sensitivity to defects of x-ray diffraction, and the low crystallinity of theses materials. In some cases High Resolution Transmission Electron Microscopy was able to give some property like the single layer vector but is unable to asses its stacking. Hence, the advantage of using computational chemistry becomes obvious : it makes it possible to assess the structural property like stacking, in a systematic fashion among a very large number of materials without the need for synthesis. Theses results can then be used to computationally estimate other properties like density, pore volume and even band-gap.

Since COFs' properties are strongly dependant on their stacking structure; wether it be geometric ones for gas adsorption and separation or electronic ones for optoelectronic processes, the first step of the calculation is to compute it as accurately as possible to then converge to an exact structure and obtain more complex properties of the material. The problem is now that a fine layer-to-layer x,y,z offset grid is necessary to achieve reasonable accuracy, and accurate calculations are very expensive machine-time-wise, especially when the number of grid-points is put together with the number of materials to test. To alleviate this cost, the solution chosen here is to make a first estimation of the structure using a method as cheap as possible and still gives a useful starting point for more in-depth optimisation. In this setting, the work detailed below aims at finding the best compromise between calculations too expensive to be

interesting as a pre-screener and the ones not accurate enough to give a useful initial guess for the rest of the treatment. With this idea in mind, different computational technics were tested on a diverse set of COFs in order to evaluate the complexity of calculations needed to achieve reasonable precision. These technics, from simplest to most elaborate are : the Lennard-Jones potential, Lennard-Jones and Coulombic interactions, DFTB+ [2] and GFN2-xTB [3][4]. Finally, the reference optimal structure was obtained obtained using DFT.[5][6]

Furthermore, by scanning all possible x,y,z offsets from one layer to the next it is possible to asses the crystallinity of the material : a steep, smooth potential would yield high crystallinity while irregularities and flatness in the Potential Energy Surface would yield lower crystallinity. This can be of the utmost importance when aiming at applications like light harvesting where irregularity in the crystalline structure can strongly impair the material's efficiency.

1.1 General Considerations

As one can see on the flowchart below (see *fig.1*), the treatment started from a COF selection from the database. The CIF file was then imported in the program, and data were extracted. A single layer was extracted from the structure, and placed at the bottom of the simulation box. The geometry of the polymer is kept the same throughout the process. If the charges were missing, the single layer was passed through an external charge equilibration routine (see *2.3*). Then the list of x and y coordinates to be tested, in this case a 25x25 grid, was generated and distributed among threads. Each of which would then compute the approximate distance between the two sheets at which the the Lennard-Jones potential cancels for the x,y point of interest. A two-stage convergence was used to increase the speed and cheapen the calculations (0.1 unit-cell step for the first one, 0.01 for the second). From that initial distance z_0, 10 points spaced by 0.01 unit-cell were defined, and the z_0 point was stored. For each x,y,z grid-point a copy of the single layer was generated and shifted by the x,y,z offset. The Lennard-Jones interactions between the two layers was then computed. The two structures were then exported under CIF and XYZ format to be fed in LAMMPS (for LJ and Coulombic calculations), xTB and CP2K (for DFT approximation calculations). To obtain the Coulombic energy, the program made use of the fact that LAMMPS is able to compute interactions between two designated sets of atoms, here the two layers, and hence neglect the intra-layer interactions. The structures were then fed into xTB and DFTB+ using simple bash scripts on the FIDIS EPFL cluster. The main program would then collect the energies, find the grid-point with the lowest energy associated and generate the structure associated to that grid-point as CIF file to be finally optimised using CP2K's DFT function.

To avoid artefacts, each COF was studied using a naive structure (theoretical bond-angles and bond-lengths) and using this same structure relaxed using CP2K's geometry and cell optimisation.
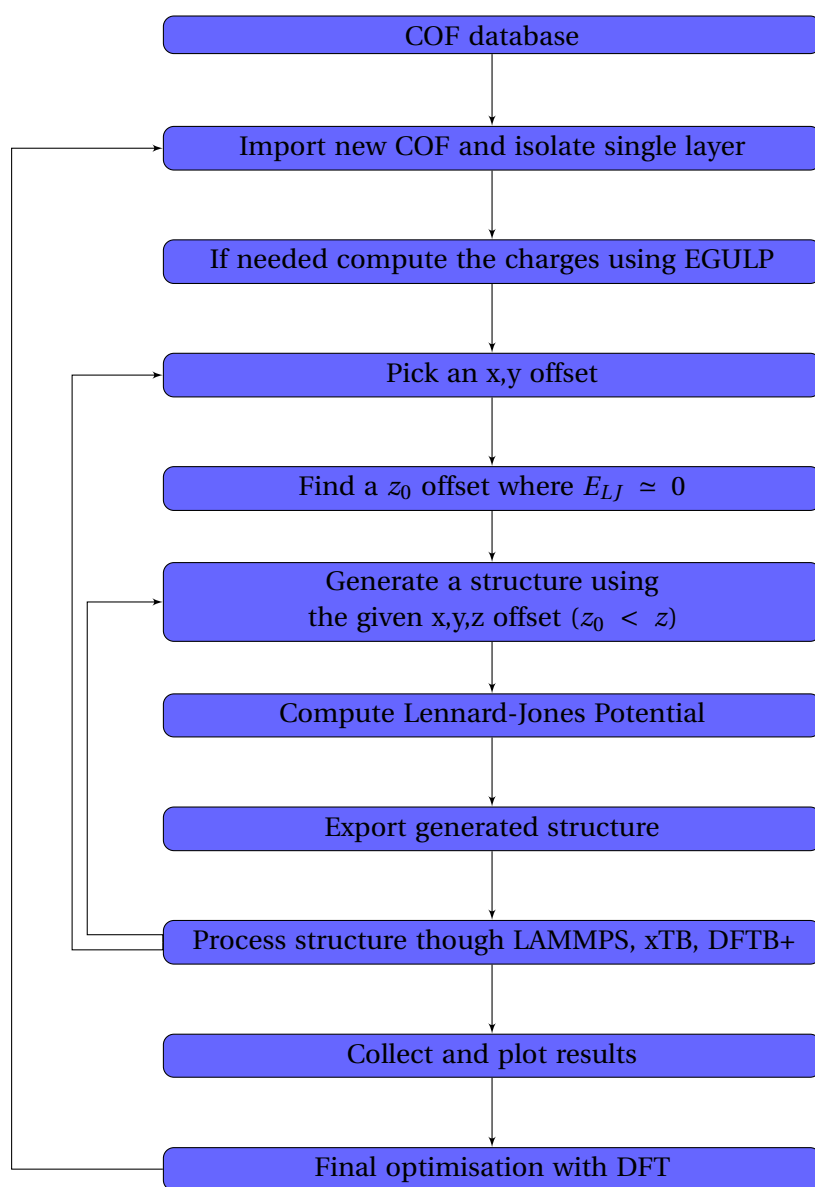
Methodology

Figure 1.1 – *Overview of the workflow in the setting of this project*

1.2 Lennard-Jones

To asses the complexity of phenomenon involved in the stacking process of COFs, the Lennard-Jones potential was used as a first approach, which is among the simplest classical approximations in computational chemistry to simulate the Van der Waals interactions. The parameters used were the Universal Force Field parameters from Rappé et al. [7]. Since these parameters are element-specific, a mixing rule is needed to combine the parameters. In our case, the Lorentz-Berthelot [8] mixing rules was used, as it is the simplest and most widely used one. It reads as follow:

$$\sigma_{ij} = \sqrt{\sigma_i \sigma_j}$$

and

$$\epsilon_{ij} = \frac{\epsilon_i + \epsilon_j}{2}$$

where indices i denote an element-specific parameter for element i, and indices ij a pair-specific parameter. σ_i and ϵ_i are the distance to $E_{LJ} = 0$ from the nucleus and the depth of the well respectively. The tail correction used here was the truncated-shifted correction :

$$E_{LJ}^{corr}(r) = E_{LJ}(r) + E_{LJ}(r_{cut})$$

Where $E_{LJ}^{corr}(r)$ is the corrected potential, and $r_c ut$ is the cutoff radius; it ensures a smooth potential at cutoff.

1.3 Charge Calculations

For the experimental structures (naive bond angle and bond length), no charges where provided and where hence computed for a single sheet using the charge equilibration from EGULP subroutine [9]. This method is also based on the Universal Force Field's parameters. The core principle behind this method is to use a second-order Taylor expansion, truncated to the second order, of the charge dependant energy ($E(Q)$)[10] :

$$E_A(Q) = E_{A0} + Q_A \left(\frac{\partial E}{\partial Q} \right)_{A0} + \frac{1}{2} Q_{A0}^2 \left(\frac{\partial^2 E}{\partial Q^2} \right)_{A0}$$

This parametrisation makes this problem computationally cheap by avoiding to go through a long and tedious quantum processing of the system and hence fast charge calculations.

Calculations were performed on a single sheet to exclude Coulombic interactions between layers.

Unfortunately, this method was not robust enough and showed some unphysical behaviours, among other yielding strongly asymmetric charge distribution in a symmetric system, or yielding positively charged Fluorine in a fluorinated benzene.

This indicates that the systems under study here needs a more complete treatments, like

DFTB+, in order to obtain the equilibrated charges. An in-depth analysis of the charges as computed by DFTB+ was also performed to asses the strength of the interactions between layers in the repartition of charges. This may enable one to estimate the correlation between charge distribution and stacking and hence the necessity to recompute the charges for every configurations or not.

1.4   Coulombic Potential and Ewald Summation

The Coulombic potential is given by :

$$E_c = k_e \frac{q_i q_j}{r_{ij}}$$

where k_e is the Coulomb's constant, q_i and q_j the charge of the atoms i and j respectively, and r_{ij} the distance between the two atoms. Since this potential converges slowly in periodic systems, it is important to take into account the long range interactions. A natural way to do so would be a very high cutoff but it would increase the computational cost drastically. To tackle this problem, one can take advantage of the periodicity of the system and compute separately the short range interactions by summing them in real space and sum the long range interactions in reciprocal space : this is known as the Ewald summation [11], which is a special case of the Poisson summation. The charge density is split as follow :

$$\rho_i(r) = \rho_i^S(r) + \rho_i^L(r)$$

where $\rho_i^S(r)$, the short range function is as follow

$$\rho_i^S(r) = q_i \delta(r - r_i) - q_i G_\sigma(r - r_i)$$

and $\rho_i^L(r)$, the long range function :

$$\rho_i^L(r) = q_i G_\sigma(r - r_i)$$

with

$$G_\sigma = \frac{1}{(2\pi\sigma^2)^{3/2}} exp \left[-\frac{|r|^2}{2\sigma^2} \right]$$

The substitution above is equivalent to adding and subtracting a Gaussian function around each point charge, described by a Dirac delta function. This will shield long range interactions in the short-range function and is corrected by the long range function to maintain a correct description of the physical system. We then integrate the short range function in real space and the long range in reciprocal space. This technique allows for fast convergence of the Coulombic potential in periodic systems.

For this project, LAMMPS (Large Scale Atomic/Molecular Massively Parallel Simulator)[12] was used for its implementation of the Particle-Particle-Particle-Mesh Ewald Summation

(PPPM) [13] , with a threshold of $10^{-4}\ kcal/mol$

For the CP2K relaxed structures, the charges were provided as computed by DFT in the two-layer structure provided.

1.5   Geometry, Frequency, Non-bonding - Tight Binding method (GFN2-xTB)

GFN2-xTB is a method approximating the Density Functional Theory by using atom-specific parametrisation. It is mostly designed for vibrational analysis, geometry optimisation and non covalent interactions (the latest being the one of interest here). The theory it relies on reads as follow, starting from the total energy :

$$E = E_{el} + E_{rep} + E_{disp} + E_{XB}$$

with E_{rep} the repulsion energy, E_{disp} the dispersion energy and E_{XB} the halogen bonding and E_{el} the electronic potential is based on a corrected third-order-truncated Taylor expansion of the Hamiltonian :

$$E_{el} = \sum_i^{occ} n_i \Psi_i |H_0| \Psi_i + \frac{1}{2} \sum_{A,B} \sum_{l(A)} \sum_{l'(B)} p_l^A p_{l'}^B \gamma_{AB,ll'} + \frac{1}{3} \sum_A \Gamma_A q_A^3 - T_{el} S_{el}$$

The first order therm is simply the energy associated with each electron in its Molecular Orbital, the second order and third order term are the SCC contribution taking only in account diagonal terms and the additional term $T_{el}S_{el}$ is the electronic thermal-entropic terms. Ψ_i is the valence Molecular Orbital with occupation n_i, H_0 the zero-order Hamiltonian, q_A the Mulliken charges of atom A, Γ_A the derivative of the Hubbard parameter η_A, l and l' are the electronic shells of atoms A and B respectively and p_l^A the charge distribution on atom A's shells with angular momentum l and $\gamma_{AB,ll'}$ is a Coulombic space-dependant damping factor. The xTB was run using the D3 dispersion correction and the SPME (Smooth Particle Mesh Ewald) summation, the Hessian matrix was estimated using the Broyden mixing scheme, for at least 8 steps before alternating with the Direct Inversion in the Iterative Subspace [14][15][16].

1.6   DFTB+

The Density Functional-Based Tight Binding (DFTB) is a more heavily parametrised method. Like xTB, it relies on an expansion around the density functional energy but uses atomic parametrisation for the electronic Hamiltonian diagonal terms (element specific) as well as well as for the off diagonal ones (pair specific). For this process, were used two sets of parameters : the 3ob-3-1 (parametrisation for Organic and Biological System) set [17][18][19][20] and the Matsci-0-3 [21] (parametrisation for Material Sciences) set for COFs containing Boron (COF-1 and COF-5). The two sets of parameters were compared for a COF were both sets could

be used. The static grid-point calculations were performed with a self-consistent Hamiltonian, with a force and self-consistency threshold of $10^{-5} a.u$, the Broyden mixer was used, and the k-points were set to x,y,z={1,1,1}, the dispersion calculations used the D3 method, using the Becke-Johnson damping factor and including the Hubbard derivatives and a correction was applied on the damping factor in the special case of Hydrogen to mimic Hydrogen-bonding.

The cell and geometry optimisation were performed using the same parameters except with a lower threshold for the forces and SCC : 10^{-6}.

1.7   Density Functional Theory

Density Functional theory is a technic relying on the physical equivalence of a set of N one-electron wavefunction $\Psi(r)$ and the sum of its position dependant module $n(r)$, the density function. A functional is then applied on the density function to extract observables as the energy. It was here used as the last stage of our process in oder to obtain the reference COF stacking starting from each technic's guess in order to asses its accuracy. The functional used was the PBE (Perdew−Burke−Ernzerhof) one, a nonempirical GGA (Generalized Gradient Approximation) functional. In addition, the DFT-D3 correction term for dispersion forces (London Forces) was used, and set to included the 9^{nth} order term, since theses play a key role in the non-bonding interactions which will determine the optimum stacking. For all elements, including Hydrogen, a Gaussian Double-Zeta function was used for the basis set [22] to mimic accurately polarisation.

1. Côté, A. P., Benin, A. I., Ockwig, N. W., O'Keeffe, M., Matzger, A. J. & Yaghi, O. M. Porous, Crystalline, Covalent Organic Frameworks. *Science* **310,** 1166–1170. ISSN: 0036-8075, 1095-9203 (Nov. 2005).

2. Aradi, B., Hourahine, B. & Frauenheim, T. DFTB+, a Sparse Matrix-Based Implementation of the DFTB Method. *The Journal of Physical Chemistry A* **111,** 5678–5684. ISSN: 1089-5639 (July 2007).

3. Grimme, S., Bannwarth, C. & Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86). *Journal of Chemical Theory and Computation* **13,** 1989–2009. ISSN: 1549-9618 (May 2017).

4. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. en. *J. Chem. Theory Comput.* 20 (2019).

5. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. en. *Physical Review* **136,** B864–B871. ISSN: 0031-899X (Nov. 1964).

6. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. en. *Physical Review* **140,** A1133–A1138. ISSN: 0031-899X (Nov. 1965).

7. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. en. *Journal of the American Chemical Society* **114,** 10024–10035. ISSN: 0002-7863 (Dec. 1992).

8. Lorentz, H. A. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. en. *Annalen der Physik* **248,** 127–136. ISSN: 1521-3889 (1881).

9. Kadantsev, E. S., Boyd, P. G., Daff, T. D. & Woo, T. K. Fast and Accurate Electrostatics in Metal Organic Frameworks with a Robust Charge Equilibration Parameterization for High-Throughput Virtual Screening of Gas Adsorption. *The Journal of Physical Chemistry Letters* **4,** 3056–3061. ISSN: 1948-7185 (Sept. 2013).

10. Rappe, A. K. & Goddard, W. A. Charge equilibration for molecular dynamics simulations. en. *The Journal of Physical Chemistry* **95,** 3358–3363. ISSN: 0022-3654, 1541-5740 (Apr. 1991).

11. Lee, H. & Cai, W. Ewald Summation for Coulomb Interactions in a Periodic Supercell. en, 12.

12. Plimpton, S. Fast Parallel Algorithms for Short–Range Molecular Dynamics. en, 42 (1993-95).

13. Hockney, R. W. & Eastwood, j. a. James. *Computer simulation using particles* English. ISBN: 978-0-07-029108-9. <https://trove.nla.gov.au/version/11367413> (visited on 06/11/2019) (New York : McGraw-Hill International Book Co, 1981).

14. Pulay, P. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters* **73,** 393–398. ISSN: 0009-2614 (July 1980).

15. Pulay, P. Improved SCF convergence acceleration. en. *Journal of Computational Chemistry* **3,** 556–560. ISSN: 1096-987X (1982).

16. Shepard, R. & Minkoff, M. Some comments on the DIIS method. *Molecular Physics* **105,** 2839–2848. ISSN: 0026-8976 (Oct. 2007).

17. Kubillus, M., Kubař, T., Gaus, M., Řezáč, J. & Elstner, M. Parameterization of the DFTB3 Method for Br, Ca, Cl, F, I, K, and Na in Organic and Biological Systems. en. *Journal of Chemical Theory and Computation* **11,** 332–342. ISSN: 1549-9618, 1549-9626 (Jan. 2015).

18. Lu, X., Gaus, M., Elstner, M. & Cui, Q. Parametrization of DFTB3/3OB for Magnesium and Zinc for Chemical and Biological Applications. en. *The Journal of Physical Chemistry B* **119,** 1062–1082. ISSN: 1520-6106, 1520-5207 (Jan. 2015).

19. Gaus, M., Lu, X., Elstner, M. & Cui, Q. Parameterization of DFTB3/3OB for Sulfur and Phosphorus for Chemical and Biological Applications. en. *Journal of Chemical Theory and Computation* **10,** 1518–1537. ISSN: 1549-9618, 1549-9626 (Apr. 2014).

20. Gaus, M., Goez, A. & Elstner, M. Parametrization and Benchmark of DFTB3 for Organic Molecules. en. *Journal of Chemical Theory and Computation* **9,** 338–354. ISSN: 1549-9618, 1549-9626 (Jan. 2013).

21. Frenzel, J, Oliveira, A. F., Jardillier, N, Heine, T & Seifert, G. Semirelativistic, selfconsistent charge SlaterKoster tables for densityfunctional based tightbinding (DFTB) for materials science simulations. en, 2 (Sept. 2004).

22. VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. en. *The Journal of Chemical Physics* **127,** 114105. ISSN: 0021-9606, 1089-7690 (Sept. 2007).