Electronic Supplementary Material (ESI) for Journal of Materials Chemistry B.

# SUPPORTING INFORMATION

# Multiscale simulations reveal IRMOF-74-III as a Potent Drug Carrier for Gemcitabine Delivery

Marianna Kotzabasaki,*[,a] Ioannis Galdadas,*[,b] Emmanuel Tylianakis,[a] Emmanuel Klontzas,[a] Zoe Cournia,[b,†] and George E. Froudakis[a,†]

[a]Department of Chemistry, University of Crete, Voutes, 71003, Heraklion, Greece

[b]Biomedical Research Foundation Academy of Athens, 4 Soranou Ephessiou St., 115 27 Athens, Greece

## Contents

## S1. Simulation Details

### *i. DFT Calculations*

Semi-empirical methods (PM7)[1] and *ab initio* calculations within the framework of density functional theory (DFT) were performed in order to study the interaction of GEM with the molecular fragments of IRMOF-74-III and of the functionalized OH-IRMOF-74-III. The molecular model that was used for the IRMOF-74-III was based on its crystal structure (CCDC code 841643), while for the modified OH-IRMOF-74-III the organic linker was modified by replacing two methyl groups with hydroxyl units as is depicted in Fig. S2. The dangling bonds in both molecular fragments were saturated by methyl groups, while lithium atoms were used to cap the terminations connected by metal atoms, similarly to the treatment proposed by Zhang et al.[2] Coordinated water molecules were kept in their crystallographic positions leaving no open metal site.

DFT in the resolution of identity (RI) approximation[3] was applied in our calculations. The PBE exchange–correlation functional[4] along with the def2-TZVP basis set[5] was used (with the corresponding auxiliary basis set for the RI approximation), including the D3 empirical correction term for the dispersion interactions as proposed by Grimme.[6] Tight convergence criteria were enforced on the SCF energy ($10^{-7}$ au) and the norm of the Cartesian gradient ($10^{-4}$ au). On top of this, all binding energies were corrected for the basis set superposition error (BSSE) with the counterpoise (CP) method as proposed by Boys and Bernardi.[7] These corrections have been essential since the BSSE may become critical for non-bonding interactions. In all calculations, we used the Multipole Accelerated Resolution of the Identity (MARI-J), which involves the evaluation of the Coulomb interaction using auxiliary basis sets through the multipole expansion for non-overlapping charge distributions without any significant loss of accuracy.[8,9] PM7 calculations were performed using the MOPAC2012 program package[10] and DFT calculations were performed using the Turbomole v.5.9[11] package.

Initially PM7 method was applied to obtain qualitative molecular geometries of the system studied. The obtained geometries were then used as initial geometries for the DFT calculations. During the DFT geometry optimizations, the molecular fragments of OH-IRMOF-74-III and the unmodified one were kept rigid at their respective optimum geometries, while no constraints were imposed on GEM.

### *MD simulations protocol*

Molecular dynamics (MD) simulations comprise a useful technique that allows the study of the dynamical behavior of GEM molecules inside the pores of IRMOF-74-III and OH-IRMOF-74-III, providing atomic level detailed information of the overall drug-diffusion phenomenon. In our study, the GROMACS v.5.0.7.[12] MD engine was used to perform all MD simulations. Diffusion of drugs through IRMOFs' pores is a dynamic process, thus taking into account the flexibility of both GEM and IRMOFs is important so as to mimic the physiological conditions. Therefore, in our models both GEM and IRMOFs are treated as flexible. To examine the mobility of GEM within the pores of the IRMOF-74-III and OH-IRMOF-74-III, we performed eight independent runs of 200 ns each that correspond to different concentrations of GEM inside the pores (Table S1). The atomistic model of the IRMOF-74-III structure was derived through replication of its crystal structure along the z axis, while the one of the functionalized derivative OH-IRMOF-74-III was obtained through replacement of the methyl groups with hydroxyl units (Table S2, Fig. S2); same as in the DFT calculation. The necessary topology of the IRMOF structures was obtained using the OBGMX[13] tool and it was post-modified so as the parameters for the bonded and non-bonded interactions to be in accordance with the ones described by the Amber99SB-ILDN force field.[14] The force field parameters for GEM were obtained using the Antechamber tool as implemented in the CcpNmr ACPYPE Portal tool [15] and they are based on the general Amber force field (GAFF).[16] The non-bonded interactions between the atoms were modeled using the Coulomb and Lenard-Jones (LJ) potentials:

$$V_{ij} = V_{coul} + V_{LJ} = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

The Lenard Jones (LJ) parameters (σ and ε) for IRMOFs and GEM are given in Tables S3-S4. The Lorentz-Berthelot combining rules were used to calculate cross LJ interaction parameters.

For the Coulombic interactions, the atomic charges of GEM, IRMOF-74-III and OH-IRMOF-74-III were derived by DFT calculations. The geometries that were used for the calculation of the atomic charges are depicted in Fig. S3-S5 and they were obtained after performing geometry optimizations with the DFT/RI-PBE-D3/def2-TZVP method as described above [DFT calculations section]. Partial atomic charges were extracted with the electrostatic potential (ESP)[17,18] fitting method using Turbomole v.5.9 package. In the ESP fitting method, an effective point net charge on each atom is determined by a least square method with constraints so that the sum of the electrostatic potential by effective point charges can reproduce the electrostatic potential which is calculated by the DFT calculation.

Prior to MD simulations, all IRMOF/GEM structures were solvated with explicit water molecules using the TIP3P water model[19] and they were subjected to 10,000 steps of energy minimization using the steepest descent algorithm so as to abolish any steric clash and ensure low energy starting structures. Energy minimization step was followed by a position restrained equilibration in the canonical (NVT) ensemble for 1 ns. Once equilibrated at constant volume, unbiased MD simulations spanning to the μs timescale in total (Table S1) were carried out in the NVT ensemble with the atomic coordinates of the system saved every 2 ps.

Long-range electrostatic interactions were treated using the particle-mesh Ewald scheme[20] with a grid spacing of 1.6 Å and a cut-off 10 Å was applied for the Van der Waals interactions. All bonds were constrained using the LINCS algorithm allowing for a time-step of 2 fs. The temperature during the simulations was kept constant at 310 K by the Nosé-Hoover thermostat[21] using a coupling constant of $\tau_T = 0.5$ ps. Initial velocities were assigned to the atoms using the Maxwell distribution at 310 K.

The analysis of the obtained trajectories was performed using GROMACS tools v.5.0.7 and VMD v.1.9.2[22] was used to visualize the trajectories. Calculation of the Root-mean-square deviation (RMSD) of the IRMOFs in the absence of any GEM molecules (Fig. S6) indicates that both structures are stable and well maintained during the 50 ns simulations, providing credit for the force field parameters used.

To determine the self-diffusion coefficient of GEM molecules we used the Einstein relation of the Mean Square Displacement (MSD),

$$\langle MSD \rangle = \lim_{t \to \infty} \left\langle \left\| \vec{r}_i(t) - \vec{r}_i(t_0) \right\|^2 \right\rangle = 6Dt$$

where $\vec{r}_i(t) - \vec{r}_i(t_0)$ is the (vector) distance traveled by the center of mass of a GEM molecule and t is the time. The MSD is calculated using the GROMACS tool gmx_msd. For the calculation, an index file containing GEM atom numbers is used and the MSD is averaged over these atoms. The self-diffusion coefficient is extracted by fitting a straight line to the average MSD of GEM molecules at the linear region of MSD graph. In order to avoid the ballistic region at the beginning of the MSD calculation, as well as large statistical errors due to poor sampling, we calculated the diffusion coefficient through linear regression only between 10 – 50 ns of the production run where there is linear correlation between the MSD and time (Fig. 2, main text).

To investigate the positioning of GEM molecules inside the IRMOFs' pores we calculated the radial distribution function (RDF) of specific atoms of GEM with respect to atoms of the functional groups of IRMOFs (Fig. S9-S12). RDF, $g_{AB}(r)$, is a useful metric that reflects the probability of finding an atom at a spherical shell of certain thickness at a distance *r* from a reference atom and is defined in the following way:

$$g_{AB}(r) = \frac{\langle \rho_B(r) \rangle}{\langle \rho_B \rangle_{total}} = \frac{1}{\langle \rho_B \rangle_{total}} \frac{1}{N_A} \sum_{i \in A}^{N_A} \sum_{j \in B}^{N_B} \frac{\delta(r_{ij} - r)}{4\pi r^2}$$

where $\langle \rho_B(r) \rangle$ is the particle density of type B at a distance r around particles A, and $\langle \rho_B \rangle_{total}$ the particle density of type B averaged overall spheres around particles A with radius $r_{max}$ (in our study $r_{max} = 1.0$ nm). RDF analyses were carried out using the full 200 ns trajectory of each system.

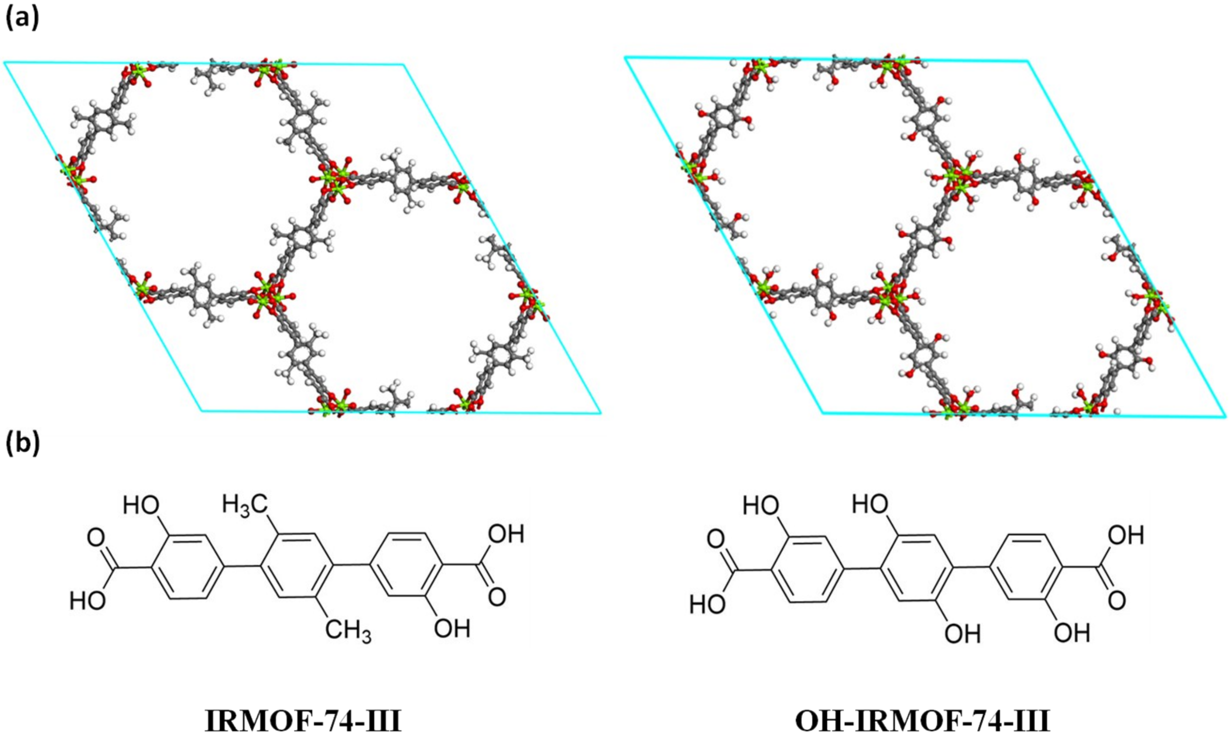
| System | Composition | Simulation Time (ns) |
|---|---|---|
| 1 | IRMOF-74-III | 50 |
| 2 | OH-IRMOF-74-III | 50 |
| 3 | IRMOF-74-III (32 wt% GEM) | 200 |
| 4 | IRMOF-74-III (63 wt% GEM) | 200 |
| 5 | IRMOF-74-III (95 wt% GEM) | 200 |
| 6 | OH-IRMOF-74-III ( 32 wt% GEM) | 200 |
| 7 | OH-IRMOF-74-III ( 63 wt% GEM) | 200 |
| 8 | OH-IRMOF-74-III ( 95 wt% GEM) | 200 |

**Table S1.** List of simulated systems. In systems 1 and 2 (control systems) no GEM molecule has been incorporated. In systems 3&6, 4&7 and 5&8 (diffusion monitor systems) we have encapsulated 20, 40 and 60 GEM molecules inside the MD cell respectively. The mass fraction (wt%) for each system has been estimated as the mass of GEM molecules over the mass of each IRMOF.

| | **IRMOF-74-III** | **OH-IRMOF-74-III** |
|---|---|---|
| **Crystal cell size ($nm^3$)** | 4.58×4.58×0.65 | - |
| **MD cell size ($nm^3$)** | 4.58×4.58×2.59 | 4.58×4.58×2.59 |
| **Crystal cell angle ($^o$)** | $(\alpha, \beta, \gamma) = (90, 90, 120)$ | $(\alpha, \beta, \gamma) = (90, 90, 120)$ |

| | | |
|---|---|---|
| **CCDC code** | 841643 | - |
| **$V_V$ (nm$^3$)** | 21.163 | 22.034 |
| **$V_T$ (nm$^3$)** | 47.123 | 47.123 |
| **Porosity ($V_V$/$V_T$)** | 0.45 | 0.47 |

**Table S2.** Crystal and MD information of the IRMOFs studied. $V_V$: void volume, $V_T$: bulk volume of IRMOF. Void space volume and bulk volume was calculated using Poreblazer v.3.0.2[23] using a 1.4 Å solvent probe.



IRMOF-74-III

OH-IRMOF-74-III

**Fig. S1. (a)** Initial configurations of the IRMOFs studied in this work, **(b)** Organic linkers that coordinate to the Mg clusters of the studied IRMOFs.

| **Atom type** | **σ (nm)** | **ε (kj/mol)** |
|---|---|---|
| C or CA | 0.339967 | 0.359824 |
| CT | 0.339967 | 0.457730 |

| | | |
|---|---|---|
| HA | 0.259964 | 0.062760 |
| HC | 0.264953 | 0.065688 |
| HW | 0.000000 | 0.000000 |
| HO | 0.000000 | 0.000000 |
| OW | 0.315061 | 0.636386 |
| OH | 0.306647 | 0.880314 |
| O2 | 0.295992 | 0.878640 |
| Mg | 0.141225 | 3.743420 |

**Table S3.** Lenard-Jones (LJ) parameters in the Amber99SB-ILDN force field corresponding to the atom types that were used to model the studied IRMOFs. CA, C and HA were used to model the aromatic rings and carboxyl groups atoms, CT and HT were used to model the methyl groups of IRMOF-74-III, OH and HO were used to model the hydroxyl groups of OH-IRMOF-74-III, O2 was used to model the coordinated oxygen atoms, and OW and HW were used to model the water molecules.

| **Atom type** | **σ (nm)** | **ε (kj/mol)** |
|---|---|---|
| C3 | 0.339967 | 0.457730 |
| C | 0.339967 | 0.359824 |
| HA | 0.259964 | 0.062760 |
| HN | 0.106908 | 0.065688 |
| H2 | 0.229317 | 0.065688 |
| H1 | 0.247135 | 0.065688 |
| HO | 0.000000 | 0.000000 |
| OS | 0.300001 | 0.711280 |

| | | |
|---|---|---|
| O | 0.295992 | 0.878640 |
| OH | 0.306647 | 0.880314 |
| N | 0.325000 | 0.711280 |
| NC | 0.325000 | 0.711280 |
| F | 0.311815 | 0.255224 |

**Table S4.** Lenard-Jones (LJ) parameters in the GAFF force field corresponding to the atom types that are used to model GEM. C3 was used to model sp$^3$ carbon atoms, C was used to model sp$^2$ carbon atoms, HA was used to model hydrogen atoms connected to an aromatic carbon, HN was used to model hydrogen atoms of the amino group, H2 was used to model hydrogen atoms that are connected to a sp$^2$ carbon, H1 was used to model hydrogen atoms that are connected to a sp$^3$ carbon, HO was used to model hydroxyl hydrogen atoms, OS was used to model the furanose oxygen, O was used to model the carbonyl oxygen, OH was used to model the hydroxyl oxygen atoms, N was used to model amino group nitrogen, NC was used to model aromatic nitrogen atoms and F was used to model fluorine atoms.

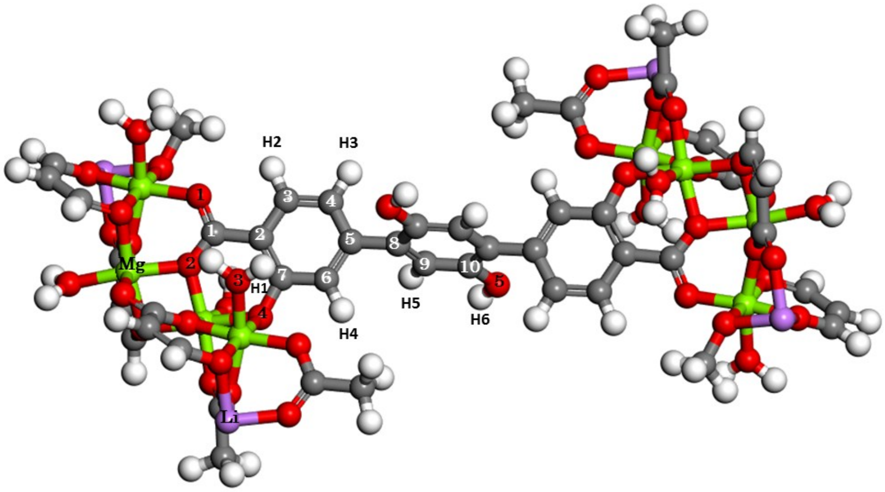


**Fig. S2.** OH-IRMOF-74-III structure used for derivation of atomic charges. Dangling bonds were saturated by methyl groups and hydrogen atoms. Carbon, oxygen, hydrogen, magnesium and lithium atoms are represented in gray, red, white, lime and purple spheres respectively.

| **Atom** | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 |
|---|---|---|---|---|---|---|---|---|---|
| **Charge (e)** | 0.792 | -0.483 | 0.010 | -0.497 | 0.457 | -0.516 | 0.585 | -0.067 | -0.436 |
| **Atom** | C10 | H1 | H2 | H3 | H4 | H5 | H6 | O1 | O2 |
| **Charge (e)** | 0.321 | 0.449 | 0.111 | 0.162 | 0.198 | 0.146 | 0.413 | -0.860 | -0.853 |
| **Atom** | O3 | O4 | Mg | | | | | | |
| **Charge (e)** | -0.898 | -0.789 | 1.897 | | | | | | |

**Table S5.** Partial atomic charges for OH-IRMOF-74-III as derived from ESP-fit.

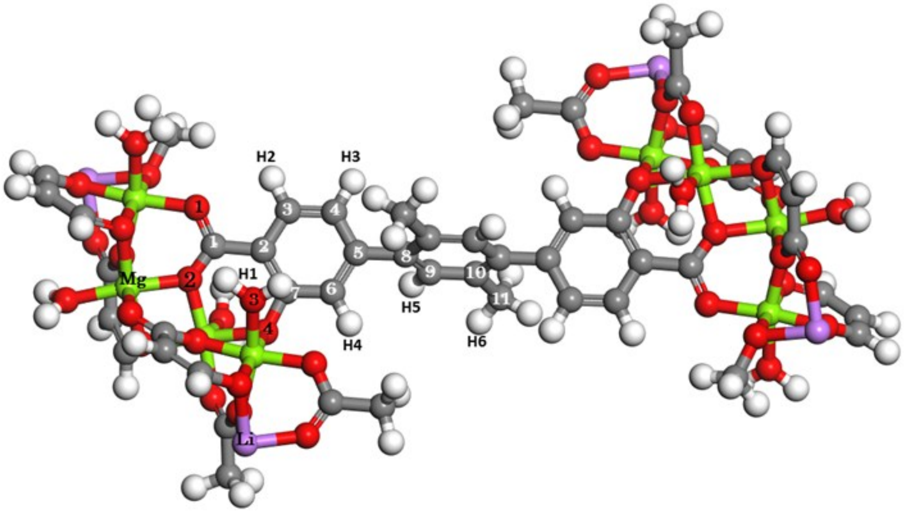


**Fig. S3.** IRMOF-74-III structure used for derivation of atomic charges. Dangling bonds were saturated by methyl groups and hydrogen atoms. Carbon, oxygen, hydrogen, magnesium and lithium atoms are represented in gray, red, white, lime and purple spheres respectively.

| **Atom** | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 |
|---|---|---|---|---|---|---|---|---|---|
| **Charge (e)** | 0.830 | -0.432 | -0.064 | -0.417 | 0.479 | -0.519 | 0.588 | -0.265 | -0.334 |
| **Atom** | C10 | C11 | H1 | H2 | H3 | H4 | H5 | H6 | O1 |
| **Charge (e)** | 0.410 | -0.295 | 0.436 | 0.129 | 0.171 | 0.227 | 0.155 | 0.146 | -0.964 |
| **Atom** | O2 | O3 | O4 | Mg | | | | | |
| **Charge (e)** | -1.057 | -0.913 | -0.967 | 1.930 | | | | | |

**Table S6.** Partial atomic charges for IRMOF-74-III as derived from ESP-fit.

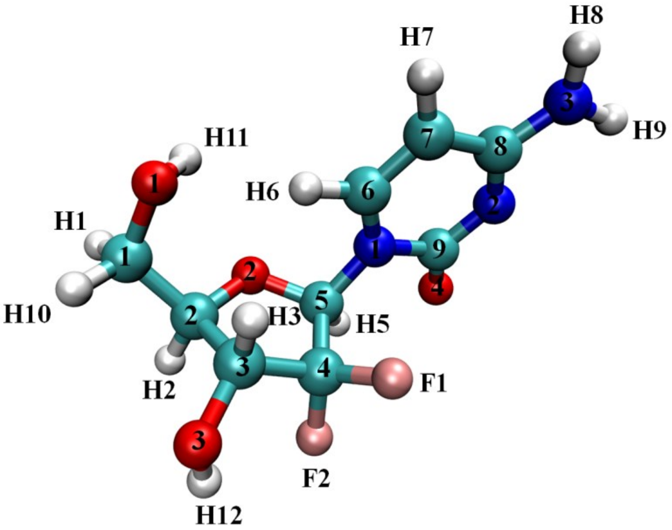


**Fig. S4.** GEM structure used for derivation of atomic charges. Carbon, oxygen, hydrogen, nitrogen and fluorine atoms are represented in cyan, red, white, blue and pink spheres respectively.

| **Atom** | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 |
|---|---|---|---|---|---|---|---|---|---|
| **Charge (e)** | 0.064 | 0.072 | 0.205 | 0.215 | 0.166 | 0.157 | -0.628 | 0.858 | 0.874 |
| **Atom** | H1 | H2 | H3 | H5 | H6 | H7 | H8 | H9 | H10 |
| **Charge (e)** | 0.045 | 0.030 | 0.101 | 0.108 | 0.083 | 0.217 | 0.383 | 0.398 | 0.089 |
| **Atom** | H11 | H12 | O1 | O2 | O3 | O4 | F1 | F2 | N1 |
| **Charge (e)** | 0.375 | 0.372 | -0.511 | -0.331 | -0.543 | -0.547 | -0.157 | 0.162 | -0.295 |
| **Atom** | N2 | N3 | | | | | | | |
| **Charge (e)** | -0.797 | -0.846 | | | | | | | |

**Table S7.** Partial atomic charges for GEM as derived from ESP-fit.

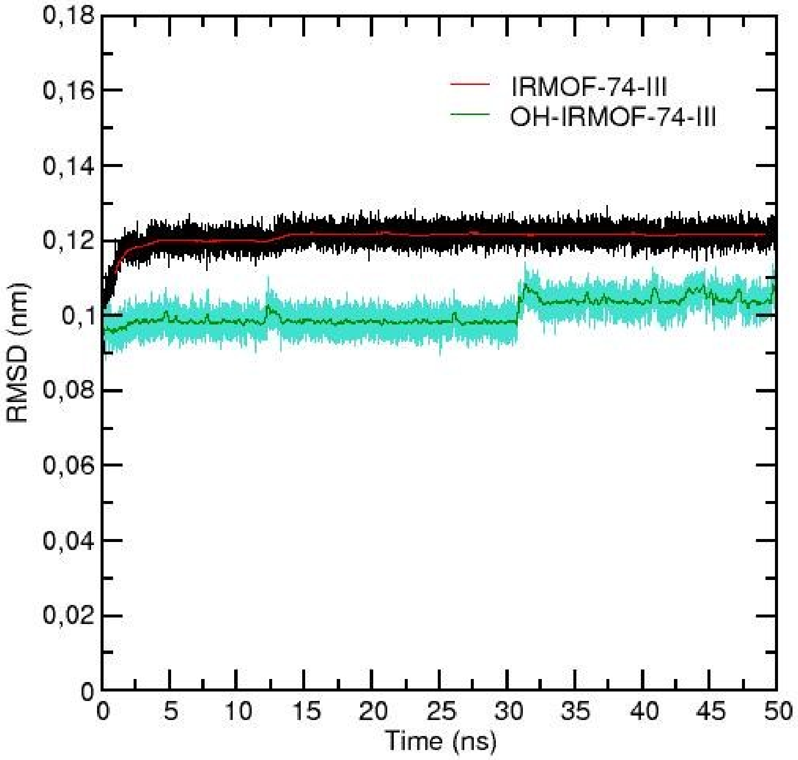


**Fig. S5.** Root-mean-square deviation (RMSD) of the heavy atoms of the studied IRMOFs from their initial structure during 50 ns simulations ("control systems 1 & 2", Table S1).

### *ii. GCMC simulations*

GCMC simulations were conducted at 310 K (37°) and under a wide range of fugacities i.e. $10^{-15}$-$10^{5}$ Pa. Interactions between guest molecules and between framework and guest molecules were treated using Lennard-Jones and Coulomb potentials for non-bonded and electrostatic interactions, respectively. Lorenz-Berthelot mixing rules were used for off-diagonal interactions. For long-range VdW interactions the potential was cut at a distance beyond $r_{VdW}$=12.8 Å, whereas electrostatic interactions were treated using Ewald summation. MOF unit cell as mentioned earlier (page S5, Table S2) had cell lengths a=b=45.8 Å, c= 6.5 Å, and angles α=β=90°, γ=120°. 1x1x4 box was used for both systems, allowing the cell lengths to be larger than two times the potential cut-off distance. Framework atoms were held constant at their crystallographic positions. Dreiding forcefield parameters ε and σ were used for framework atoms, whereas the partial charges were determined using the ESP charge fitting method (Table S5 & S6).

GEM molecules were treated as rigid. Their coordinates were taken using ESP charge fitting method on the minimized geometry at the DFT level of theory using PBE as a functional and the TZVP basis set as described earlier in this text [*Section i. DFT calculations*]. For the Lennard-Jones parameters ε and σ it was modeled using the TraPPE force field.[24] Partial charges were calculated using the ESP method on the optimized geometry (Table S7). Potential parameters are given in Table S8 according to the notation of Fig. S5.

| **Atom** | **σ (Å)** | **ε/T (K)** |
|---|---|---|
| C1 | 3.950 | 46.000 |
| C2 | 4.680 | 10.000 |
| C3 | 4.680 | 10.000 |
| C4 | 4.730 | 25.700 |
| C5 | 4.680 | 10.000 |
| C6 | 3.600 | 30.700 |
| C7 | 3.600 | 30.700 |
| C8 | 3.600 | 30.700 |
| C9 | 3.450 | 30.700 |
| F1 | 0.000 | 0.000 |
| F2 | 0.000 | 0.000 |
| H1 | 0.000 | 0.000 |
| H2 | 0.000 | 0.000 |
| H3 | 0.000 | 0.000 |
| H5 | 0.000 | 0.000 |
| H6 | 2.360 | 25.450 |
| H7 | 2.360 | 25.450 |
| H8 | 0.000 | 0.000 |
| H9 | 0.000 | 0.000 |
| H10 | 0.000 | 0.000 |
| H11 | 0.000 | 0.000 |
| H12 | 0.000 | 0.000 |
| N1 | 2.600 | 141.000 |
| N2 | 3.200 | 57.000 |

|  |  |  |
|---|---|---|
| N3 | 3.340 | 111.000 |
| O1 | 3.020 | 93.000 |
| O2 | 2.200 | 190.000 |
| O3 | 3.020 | 93.000 |
| O4 | 2.950 | 90.000 |

**Table S8.** Potential parameters ε and σ for GEM molecule used for the GCMC simulations of this work. Atom numbers correspond to the ones shown in Fig. S5.

## S2. Supplementary results

| System | Total number of H-bonds |
|---|---|
| IRMOF-74-III (32 wt% GEM) | 6.8 ± 0.5 |
| OH-IRMOF-74-III (32 wt% GEM) | 5.1 ± 0.3 |
| IRMOF-74-III (63 wt% GEM) | 25.6 ± 1.1 |
| OH-IRMOF-74-III (63 wt% GEM) | 22.8 ± 0.9 |
| IRMOF-74-III (95 wt% GEM) | 62 .0 ± 2.1 |
| OH-IRMOF-74-III (95 wt% GEM) | 51.6 ± 1.8 |

**Table S9.** Average number of the total hydrogen bonds formed between GEM molecules in GEM/water/IRMOF-74-III, and GEM/water/OH-IRMOF-74-III systems.

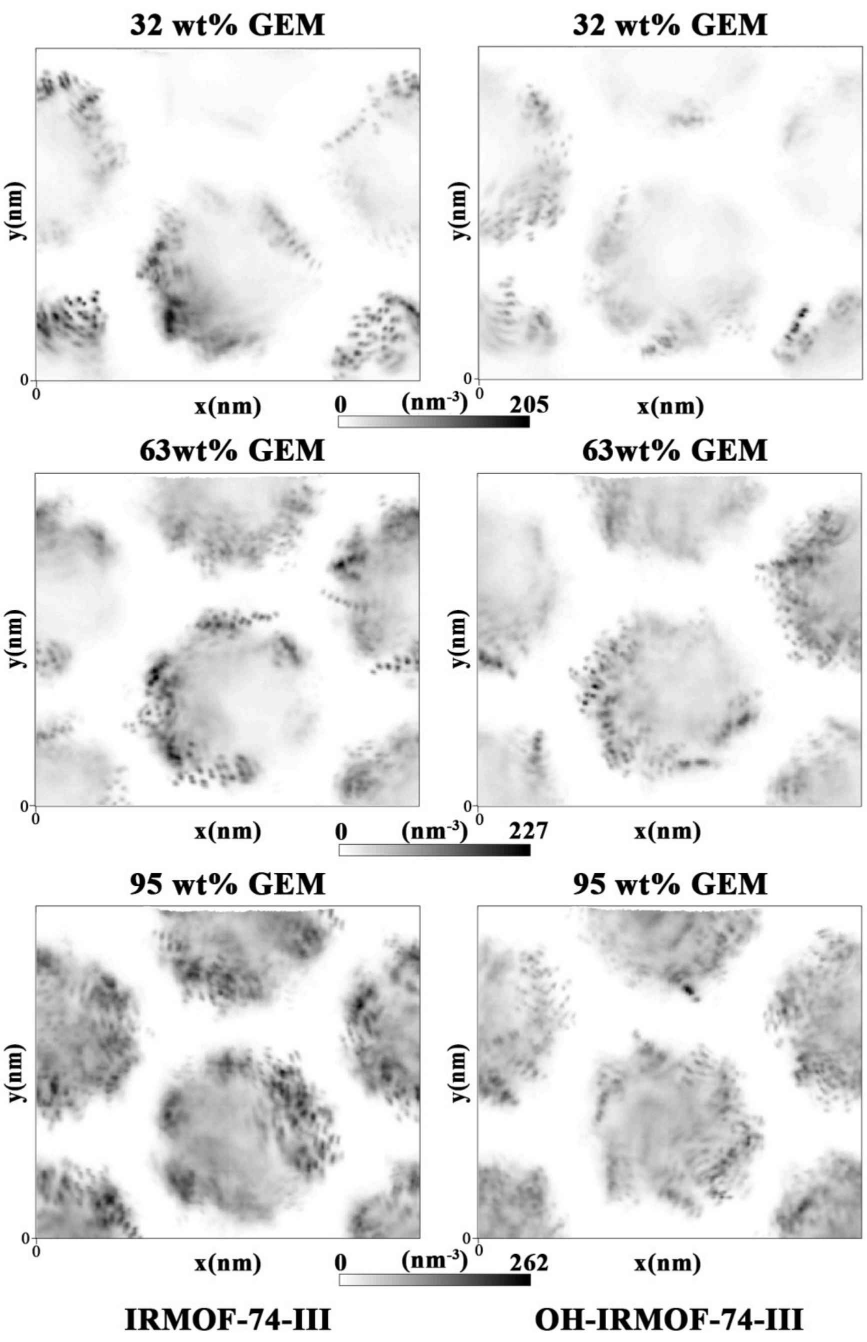
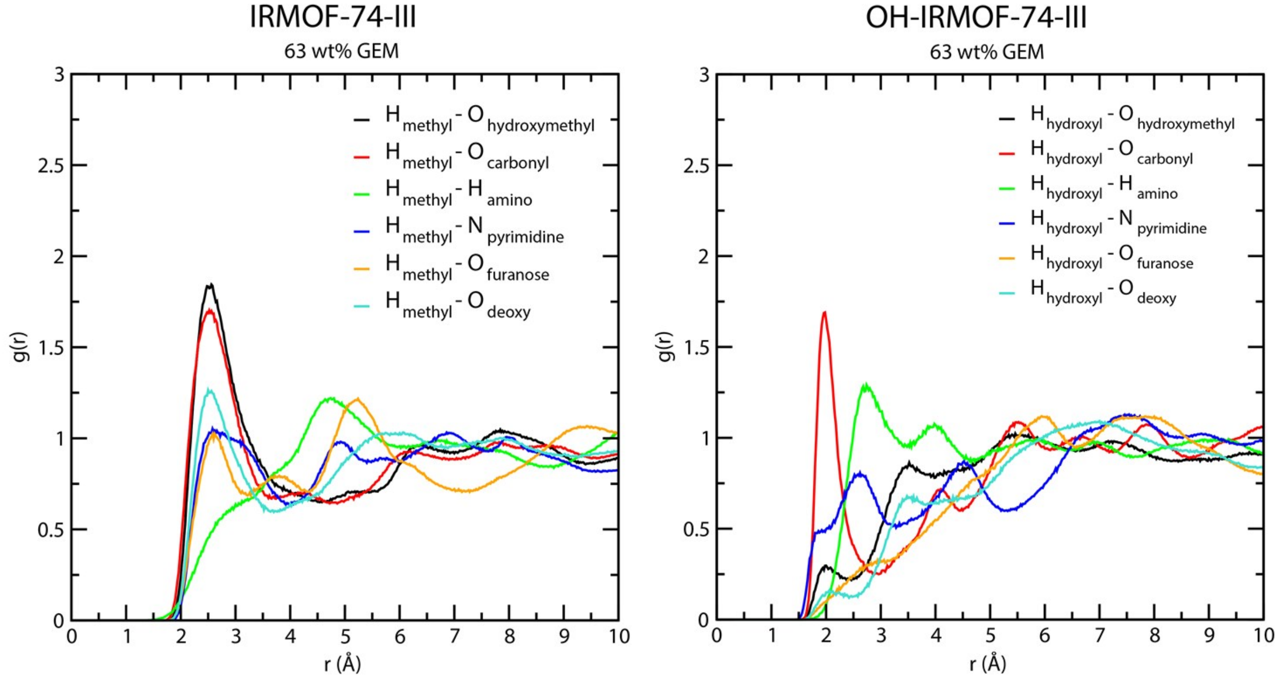


**Fig. S6.** 2D number-density maps of GEM molecules inside IRMOF-74-III and OH-IRMOF-74-III with respect to the different loading profiles. Dark gray and black areas correspond to areas of high concentration of GEM. White space between the dark areas indicates the hexagonal IRMOF cells.

| **Linker atom – GEM atom** | **IRMOF-74-III** | | **OH-IRMOF-74-III** | |
|---|---|---|---|---|
| | first shell | first minimum (Å) | first shell | First minimum (Å) |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{hydroxymethyl})}$ | 1.7 | 3.8 | 0.1 | 2.3 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{carbonyl})}$ | 1.2 | 3.0 | 0.9 | 2.2 |
| $H_{(\text{methyl or hydroxyl})} - H_{(\text{amino})}$ | 0.3 | 2.8 | 1.2 | 3.1 |
| $H_{(\text{methyl or hydroxyl})} - N_{(\text{pyrimidine})}$ | 0.3 | 2.6 | 1.6 | 2.9 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{furanose})}$ | 0.5 | 2.8 | 0.6 | 3.5 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{deoxy})}$ | 1.6 | 3.8 | 0.3 | 2.6 |

**Table S10.** First average solvation shell for the various functional groups in the 32 wt% GEM simulation. The first solvation shell was calculated by integrating the corresponding pair radial distribution function up to the first minimum (Fig. 4). H, hydrogen atom; C, carbon atom; O, oxygen atom.


IRMOF-74-III

63 wt% GEM

OH-IRMOF-74-III

63 wt% GEM

**Fig. S7.** Radial distribution function of representative atoms of the different functional groups of 63 wt% GEM with respect to (a) hydrogen atom of the methyl group of IRMOF-74-III organic linker, and with respect to (b) hydrogen atom of the hydroxyl group of OH-IRMOF-74-III organic linker.

| **Linker atom – GEM atom** | **IRMOF-74-III** | | **OH-IRMOF-74-III** | |
|---|---|---|---|---|
| | first shell | first minimum (Å) | first shell | first minimum (Å) |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{hydroxymethyl})}$ | 2.3 | 4.1 | 0.2 | 2.4 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{carbonyl})}$ | 1.9 | 3.7 | 1.0 | 3.0 |
| $H_{(\text{methyl or hydroxyl})} - H_{(\text{amino})}$ | 0.7 | 3.5 | 1.3 | 3.4 |
| $H_{(\text{methyl or hydroxyl})} - N_{(\text{pyrimidine})}$ | 1.6 | 4.0 | 1.0 | 3.3 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{furanose})}$ | 0.9 | 3.2 | 0.3 | 3.2 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{deoxy})}$ | 1.3 | 3.5 | 0.1 | 2.3 |

**Table S11.** First average solvation shell for the various functional groups in the 63 wt% GEM simulation. The first solvation shell was calculated by integrating the corresponding pair radial distribution function up to the first minimum (Fig. S7). H, hydrogen atom; C, carbon atom; O, oxygen atom.

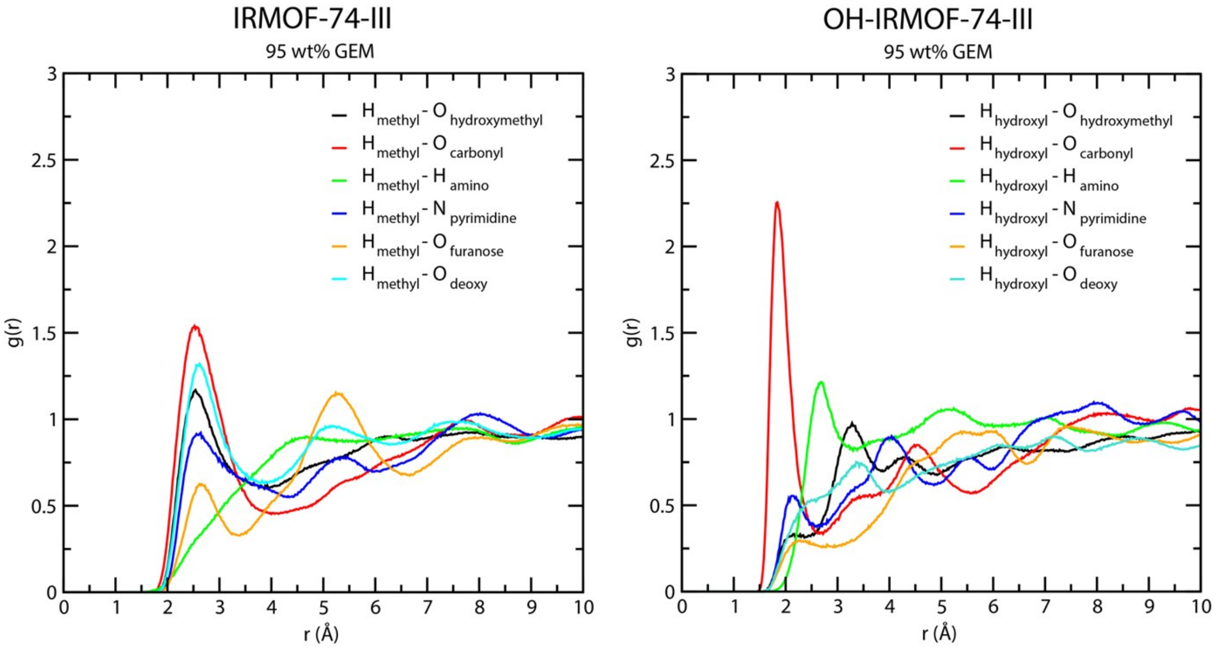


**Fig. S8.** Radial distribution function of representative atoms of the different functional groups of 95 wt% GEM with respect to (a) hydrogen atom of the methyl group of IRMOF-74-III organic linker, and with respect to (b) hydrogen atom of the hydroxyl group of OH-IRMOF-74-III organic linker.

| **Linker atom – GEM atom** | **IRMOF-74-III** | | **OH-IRMOF-74-III** | |
|---|---|---|---|---|
| | first shell | first minimum (Å) | first shell | first minimum (Å) |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{hydroxymethyl})}$ | 1.5 | 3.9 | 0.2 | 2.5 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{carbonyl})}$ | 1.8 | 5.0 | 1.2 | 2.7 |
| $H_{(\text{methyl or hydroxyl})} - H_{(\text{amino})}$ | 2.0 | 5.3 | 1.1 | 3.4 |
| $H_{(\text{methyl or hydroxyl})} - N_{(\text{pyrimidine})}$ | 1.6 | 5.5 | 0.4 | 2.8 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{furanose})}$ | 0.6 | 3.5 | 0.2 | 2.8 |
| $H_{(\text{methyl or hydroxyl})} - O_{(\text{deoxy})}$ | 1.3 | 3.5 | 0.8 | 3.9 |

**Table S12.** First average solvation shell for the various functional groups in the 95 wt% GEM simulation. The first solvation shell was calculated by integrating the corresponding pair radial distribution function up to the first minimum (Fig. S8). H, hydrogen atom; C, carbon atom; O, oxygen atom.

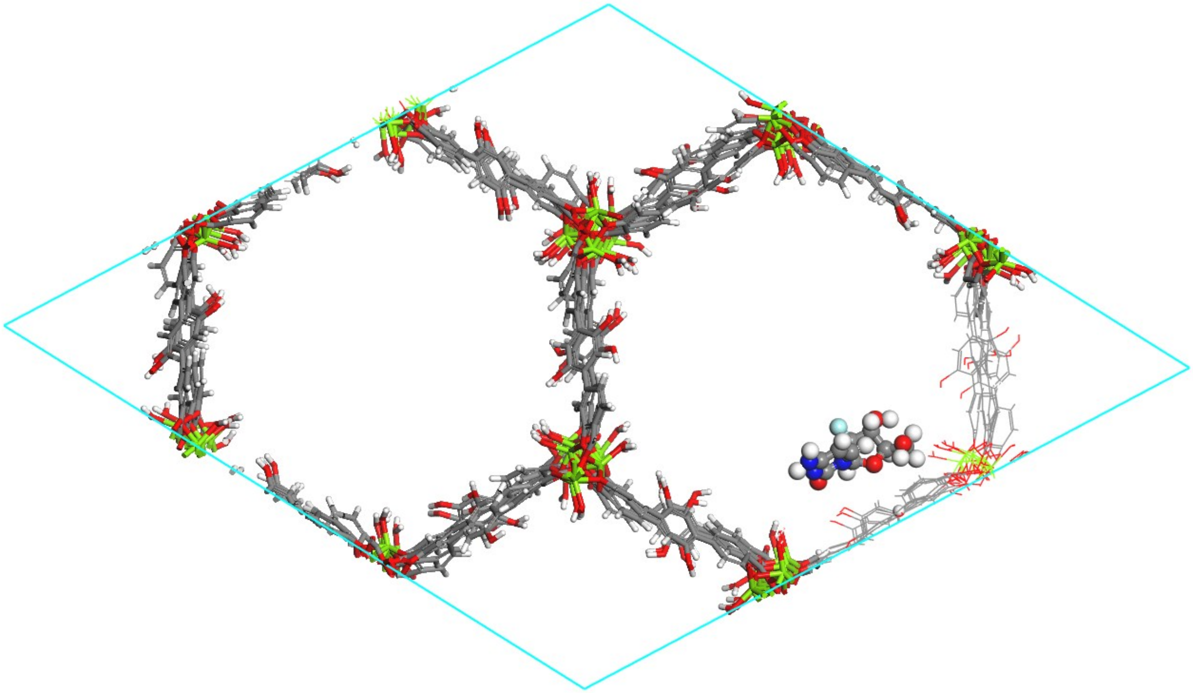


**Fig. S9.** MD simulation snapshot (t=192 ns) of one GEM molecule of 63 wt% in OH-IRMOF-74-III in the presence of water. Water molecules and the rest of GEM molecules are not illustrated here for clarity reasons. As can be seen GEM interacts with the hydroxyl group of OH-IRMOF-74-III organic linker via the groups of its 4-amino-pyrimidine moiety.

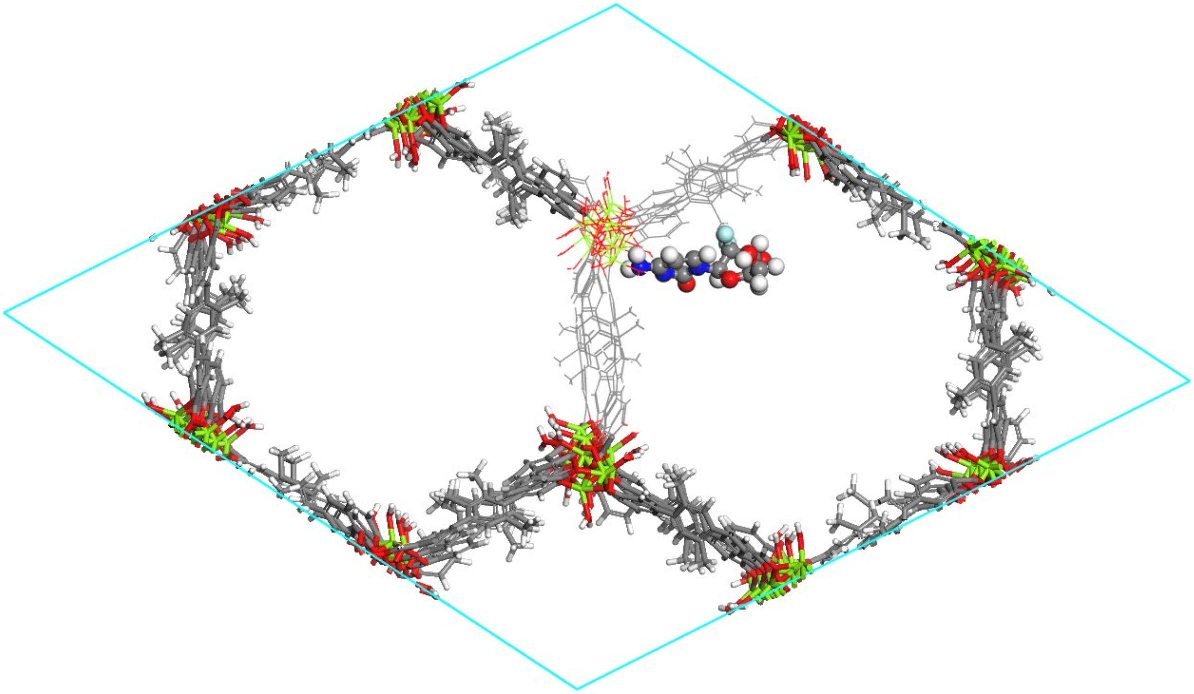


**Fig. S10.** MD simulation snapshot (t=129 ns) of one GEM molecule of 63 wt% in IRMOF-74-III in the presence of water. Water molecules and the rest of GEM molecules are not illustrated here for clarity reasons. As can be seen, the amino group of GEM mainly interacts with the corners of IRMOF-74-III leaving the rest of the molecule free to rotate in the bulk.

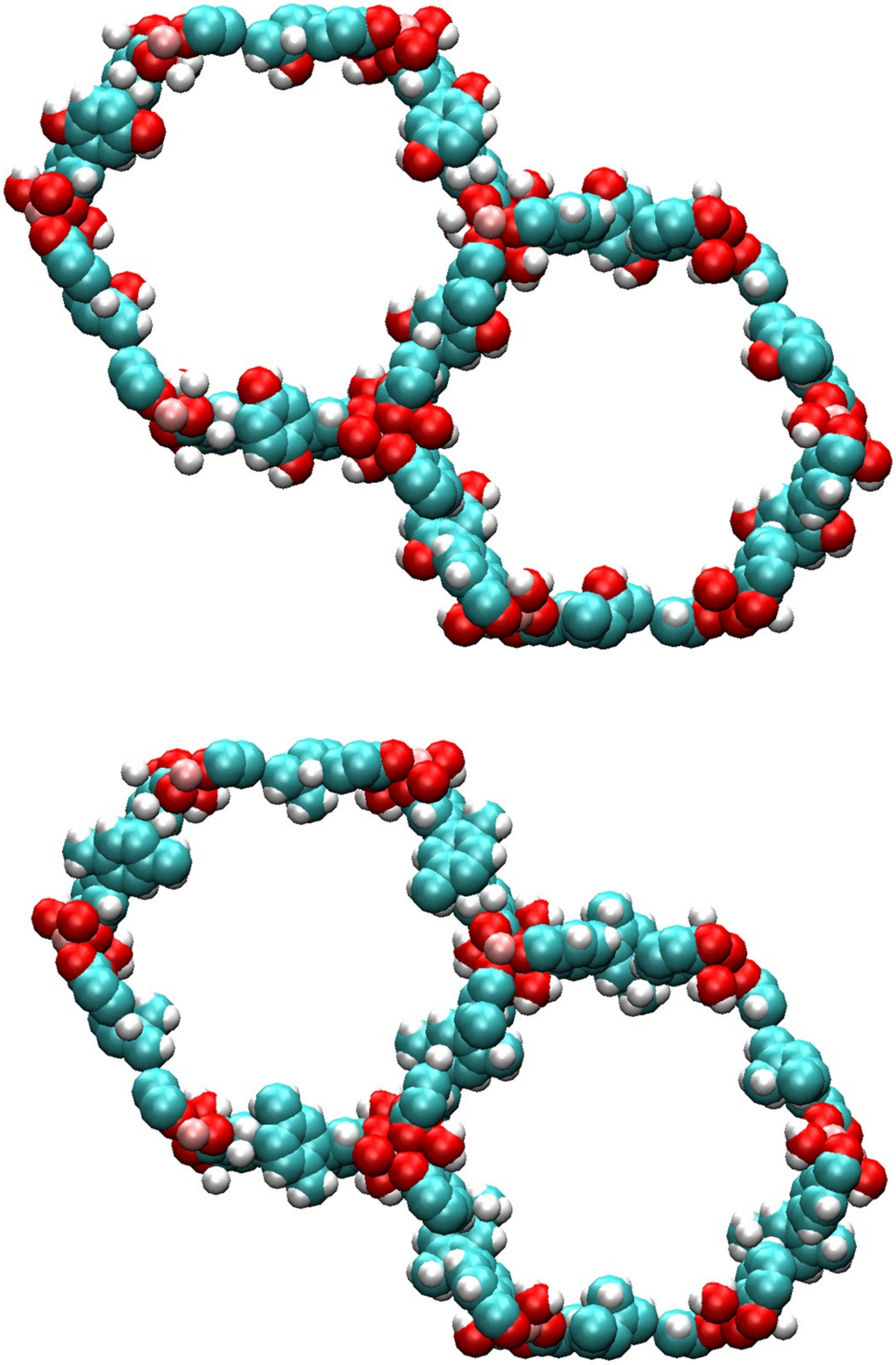


**Fig. S11.** MD simulation cells of OH-IRMOF-74-III (top) and IRMOF-74-III (down) as represented in VDW spheres. The bulkier –CH3 group compared to the –OH group leads to a decrease in the free volume of IRMOF-74-III.

## References

---

[1] J. J. Stewart, *J Mol Model.*, 2013, **19**(1), 1-32.

[2] H. Zhang, Y. Lv, T. Tan and D. van der Spoel, *J. Phys. Chem. B*, 2016, **120**(3), 477-484.

[3] F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Letters.*, 1998, **294** (1-3), 143-152.

[4] J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865−3868.

[5] F. Weigend, and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297−3305.

[6] S. Grimme, J. Antony, S. Ehrlick, S. Krieg, *J. Chem. Phys.*, 2010, **132,** 154104.

[7] S. Boys, F. Bernardi, *Mol. Phys.*, 1970, **19**, 53–566.

[8] O. Vahtras, J. Almlöf, M. W. Feyereisen*, Chem Phys Lett*, 1993, **213**(5:6), 514-518.

[9] M. Sierka, A. Hogekamp, R. Ahlrichs, *R. J. Chem. Phys.*, 2003, **118**, 9136-9148

[10] J. J. P. Stewart, MOPAC2012, Stewart Computational Chemistry, Colorado Springs, CO, USA, 2012.

[11] R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. M. Kölmel, *Chem. Phys. Lett.*, 1989, **162** (3), 165–169.

[12] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX*, 2015, **1–2**, 19-25.

[13] G. Garberoglio, *J. Comput. Chem.*, 2012, **33**, 2204-2208.

[14] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror and D.E. Shaw, *Proteins*, 2010, **78(**8), 1950-8.

[15] A. W. Sousa da Silva and W. F Vranken, *BMC Res. Notes*, 2012, **5**, 367.

[16] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* 2004, **25**, 1157−1174.

[17] U. C. Singh and P. A. Kollman, *Comput. Chem.*, 1984, **5**, 129.

[18] B. H Besler, K. M. Merz, P. A. Kollman, *J. Comput. Chem.*, 1990, **11**, 431.

[19] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein *J. Chem. Phys.*, 1983, ***79***, 926.

[20] U. Essmann, L. Perera, M. L Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577.

[21] S. Nose, *Mol. Phys.*, 1984, **52**(2), 255−268.

[22] W. Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, 1996, **14**, 33-38.

[23] L. Sarkisov, *Mol. Simul.*, 2011, **37**, 1248-1257.

[24] *"The Transferable Potentials for Phase Equilibria Family of Force Fields".* The Siepmann Group. University of Minnesota. Retrieved August 27, 2014.