

# Molecular Level Characterization of the Structure and Interactions in a Peptide-Functionalized Metal-Organic Frameworks

Tanya K. Todorova,<sup>\*†</sup> Xavier Rozanska,<sup>‡</sup> Christel Gervais,<sup>§</sup> Alexandre Legrand,<sup>⊥</sup> Linh N. Ho,<sup>⊥</sup> Pierrick Berruyer,<sup>¶</sup> Anne Lesage,<sup>¶</sup> Lyndon Emsley,<sup>#</sup> David Farrusseng,<sup>⊥</sup> Jérôme Canivet<sup>⊥</sup> and Caroline Mellot-Draznieks<sup>\*†</sup>

<sup>†</sup> Laboratoire de Chimie des Processus Biologiques, UMR 8229 CNRS, UPMC Univ Paris 06, Collège de France, 11 Marcelin Berthelot, 75231 Paris Cedex 05, France

<sup>‡</sup> Materials Design, S.A.R.L. 18 rue de Saisset, 92120 Montrouge, France

<sup>§</sup> Sorbonne Universités - Laboratoire de Chimie de la Matière Condensée de Paris (LCMCP), UPMC Université Paris 06, UMR CNRS 7574, 4 place Jussieu, 75252 Paris cedex 05, France

<sup>⊥</sup> IRCELYON, Université Lyon 1 - CNRS, UMR 5256, 2 Avenue Albert Einstein, 69626 Villeurbanne, France

<sup>¶</sup> Université de Lyon, Institut des Sciences Analytiques, UMR 5280, CNRS, Université Lyon 1, ENS Lyon - 5, rue de la Doua, F-69100 Villeurbanne, France

<sup>#</sup> Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland

## Abstract

We use DFT, newly parameterized Molecular Dynamics simulations and last generation <sup>15</sup>N DNP surface enhanced solid-state NMR spectroscopy to understand graft-host interactions and effects imposed by the MOF host on peptide conformations in a peptide-functionalized MOF. Focusing on two grafts typified by MIL-68-Proline (**–Pro**) and MIL-68-Glycine-Proline (**–Gly-Pro**), we identified the most likely peptide conformations adopted in the functionalized hybrid frameworks. We found that hydrogen bond interactions between the graft and the surface hydroxyl groups of the MOF are key in determining the peptides conformation(s). <sup>15</sup>N DNP SENS methodology shows unprecedented signal enhancements when applied to these peptide-functionalized MOFs. The calculated chemical shifts of selected MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** conformations are in a good agreement with the experimentally obtained <sup>15</sup>N NMR signals. The study shows that the conformations of peptides when grafted in a MOF host are unlikely to be freely distributed, and conformational selection is directed by strong host-guest interactions.

## Introduction

Metal-organic frameworks (MOFs) are attracting ongoing exceptional interest in the scientific community with the demonstration of high value-added applications in photocatalysis,<sup>1</sup> sensing,<sup>2</sup> chiral separation and asymmetric heterogeneous catalysis.<sup>3–5</sup> In particular, MOFs are now recognized as promising platforms to develop features such as enantioselectivity. Their hierarchical porous structures and confined pockets allow protection of catalytic centers and enhanced substrate specificity whilst maintaining efficient diffusion. In recent years researchers have been increasingly inspired by nature for the design of novel MOFs viewed as artificial enzymes.<sup>6–8</sup> Notably tremendous effort has been made to introduce chirality into porous solids<sup>9–12</sup> to allow chiral recognition through conformational adjustment between the substrate and the designed chiral adsorption/catalytic site.<sup>13–19</sup>

One efficient strategy to prepare chiral MOFs is the use of enantiopure ligands<sup>15,20–24</sup> such as amino acids and peptides as constitutive linkers, leading to a well-developed sub-class of amino acid- and peptide-based MOFs.<sup>25–29</sup> Another powerful route relies on post-synthetic modification strategies.<sup>14,30–33</sup> In 2011, Canivet et al. reported the application of solid-phase peptide-coupling to the amino-functionalized In-MIL-68-NH<sub>2</sub> in order to covalently immobilize proline and alanine amino acids within the framework.<sup>34</sup> The successful covalent grafting of larger oligopeptides (up to tetrapeptides) was further reported for a variety of MOFs selected for their non-breathing frameworks and high pore volumes (Zr-UiO-66-NH<sub>2</sub>, In-MIL-68-NH<sub>2</sub> and Al-MIL-101-NH<sub>2</sub>).<sup>35</sup> The enantiomeric purity of these peptide-functionalized porous hybrid solids was demonstrated with the occurrence of a proline-catalyzed asymmetric aldol reaction.

Kaskel and coworkers reported two isorecticular proline-functionalized Zr-based MOF, UiO-67-NH-**Pro** and UiO-68-NH-**Pro**, with similar topology but different window sizes imposed by biphenyl and terphenyl ligands, respectively.<sup>36</sup> They showed an increase in activity in the MOF-catalyzed adol reaction accompanied by an unexpected increase in selectivity when moving from UiO-67 to UiO-68 framework, highlighting the influence of the UiO-type support on catalytic performances. There is thus a need to develop a molecular level understanding of peptide-functionalized MOFs, including the role of the surrounding framework. To the best of our knowledge, no such study on peptide-grafted MOFs has been reported so far. Notably, structural characterization is difficult due to the statistical and dynamical disorder of the grafted species and their low diffraction atomic contrast.

In this work, we investigate the structural features of peptide-functionalized MOFs through a combination of Density Functional Theory (DFT), Molecular Dynamics (MD) and <sup>15</sup>N Dynamic Nuclear Polarization (DNP) surface enhanced NMR spectroscopy (SENS) to unravel host-guest interactions driving the most likely adopted conformation in the experimentally synthesized peptide-functionalized MOFs. Here, we study In-MIL-68-NH<sub>2</sub>, chosen for its reasonable unit cell size for computational modelling,<sup>35</sup> and explore its grafting with proline or glycine-proline, namely MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro**. The glycine residue is conceived here as a small spacer between the MOF wall and the proline. In a first step, dispersion-corrected periodic DFT calculations were coupled with MD simulations to predict the key interactions determining the conformations of the amino acids grafted to the MOF and to explore their dynamics. For that purpose a new parametrized version of the PCFF+ force field, known to yield an excellent description of organic systems,<sup>37–39</sup> was specially developed and validated against *ab initio* MD calculations. The development of this force field was a key step to provide access to nanosecond length MD simulations of the functionalized MOFs. The combination of DFT and MD calculations gave access to the lowest predicted energy conformations of the grafts within the pores. In a second step, the natural abundance <sup>15</sup>N NMR spectra of grafted In-MIL-68-NH-**Pro** and In-MIL-68-NH-**Gly-Pro** were collected by DNP SENS. The low detection limit provided by the DNP NMR technique<sup>40</sup> makes it well suited to study host-guest type interactions in solids,<sup>41</sup> including post-synthetically functionalized MOFs.<sup>42</sup> DFT-calculated <sup>15</sup>N chemical shifts of selected MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** models were compared with <sup>15</sup>N NMR signals obtained from the DNP SENS experiments. We revisited the interpretation of the previously published NMR spectrum of MIL-68-NH-**Pro**<sup>42</sup> and re-assigned the signal emanating from the proline nitrogen. Finally, the difference in terminal proline chemical environment assessed by the

different  $^{15}\text{N}$  NMR chemical shifts is explained by specific interactions with the host framework as enlightened by our combined computational approach.

## Computational and NMR Methods

**Strategy.** The *in silico* study of MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** functionalized MOFs was performed using a multi-step approach. The first step consisted in building a starting pool of initial structural models obtained by standard geometry optimizations ( $T = 0\text{ K}$ ) based on periodic dispersion-corrected DFT calculations. This allowed us to explore various conformations for the anchorage of the **-Pro** and **-Gly-Pro** grafts to the MOF and to obtain an assessment of the relative strengths of interactions at play in the “0K” minimized conformations. The second step consisted of molecular dynamic (MD) simulations run from a selection of DFT-optimized MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** models. This step allowed us to perform an extensive sampling of the conformational space of the grafted amino acid and dipeptide when anchored to the MOF and to establish the role of the MOF framework in their dynamical behavior. In a further step, the lowest energy conformations of MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** were extracted from MD trajectories and further re-optimized at the DFT level and added to our pool of DFT-optimized conformations. In a final step, the pool of DFT-optimized models was used for  $^{15}\text{N}$  NMR chemical shift calculations. The latter were compared to the experimental  $^{15}\text{N}$  NMR DNP spectra of MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** solids, providing assignment of the experimental shifts and revealing the most likely adopted conformations in the experimentally synthesized structure.

**Periodic DFT.** DFT calculations were performed using the Vienna *Ab-initio* Simulation Package (VASP).<sup>43,44</sup> A plane-wave basis set with an energy cutoff of 400 eV was employed, along with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional.<sup>45</sup> The long-range weak dispersion interactions were taken into account using the semi empirical vdW method of Grimme DFT-D3.<sup>46</sup> The electron-ion interactions were described by the projector augmented wave (PAW) method<sup>47</sup> in the implementation of Kresse and Joubert.<sup>48</sup> Lattice parameters and atomic positions were optimized until the forces on all atoms were smaller than  $0.02\text{ eV \AA}^{-1}$ . The primitive cell of MIL-68-NH<sub>2</sub> was taken as a starting point for the construction of all models, using the available crystal data<sup>49,50</sup> and doubling the cell dimension in the *c*-direction, resulting in a relatively large supercell ( $V=5934\text{ \AA}^3$ ) and cell parameters ( $a = b = 21.758\text{ \AA}$ ,  $c = 14.466\text{ \AA}$ ,  $\alpha = \beta = 90^\circ$ ,  $\gamma = 60^\circ$ ). The Brillouin zone was sampled only at the  $\Gamma$ -point, which is sufficient for the supercells calculated in this study. Further details on the construction of the grafted MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** are given in SI.

### Forcefield parametrization.

For the purpose of MD calculations, a set of PCFF+ force field parameters<sup>37-39</sup> were specifically developed to model MIL-68-NH<sub>2</sub>. We benefit from the wide versatility of the PCFF+ force field type to model a wide range of already parameterized organic groups including amino acids and polypeptides with a good accuracy.<sup>37-39</sup> Furthermore, PCFF+ relies on *ab initio* parameterization, making additional *ab initio*-parameterization using this forcefield a coherent choice. The currently available PCFF+ force field was supplemented with the missing non-bonded Lennard-Jones interactions and the missing bonded terms between the inorganic sub-network and the organic benzenedicarboxylate (bdc) linker (i.e.

In-O bonds, O-In-O angles and torsions, In = indium, O = carboxyl oxygen of bdc). In order to generate these parameters, an original and recently reported strategy was applied.<sup>51</sup> The required parameters were fitted using the *ab initio* forces obtained from DFT geometry optimization and *ab-initio* molecular dynamics (AIMD) trajectories of the MOF, thus allowing for high consistency of the structures obtained with the two methods, i.e. AIMD and forcefield-based MD. The extended PCFF+ forcefield reproduced the MOF structure used for the fitting, i.e., the MIL-68-NH<sub>2</sub> DFT-optimized structure, to an accuracy of 0.36 Å r.m.s. atomic displacement excluding hydrogen atoms, making it suitable for further MD simulations of amino acid and peptide-functionalized MOFs. Details about the PCFF+ forcefield, fitting procedure, validation, and resulting parameters are given in the Supporting Information (Tables S1–S5).

**Molecular Dynamics.** Molecular dynamic (MD) simulations on selected models of MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** were performed using extended PCFF+ forcefield specifically designed by us for the purpose of this study (see above). 5 ns NVT simulations (constant number of particles, constant volume, and constant temperature) were carried out at 298 K starting from the DFT optimized structures at 0 K, with an integration time step set to 0.2 fs. Large supercells reconstructed from the VASP primitive cells were used, with cell volumes of 71207 Å<sup>3</sup> and cell parameters of  $a = 43.516$  Å,  $b = 37.705$  Å,  $c = 43.398$  Å and  $\alpha = \beta = \gamma = 90^\circ$  (see Supporting Information). Each computational cell contained 12 grafted amino-acids, thus ensuring a better statistic over the grafts' conformations (Figure S7). The MD simulations were performed with the Large-scale Atomic Massively Parallel Simulator (LAMMPS) software.<sup>52,53</sup> Frames were collected every 5 ps (1000 frames). The Cartesian coordinates root mean square deviation (RMSD) of heavy atoms of the grafted amino-acid (C, N and O) and radial distribution function (RDF) plots of pairs of atoms of interest were analyzed.

**NMR calculations.** The NMR chemical shift calculations were performed with the DFT formalism using the QUANTUM-ESPRESSO software.<sup>54</sup> The PBE generalized gradient approximation<sup>45</sup> was used and the valence electrons were described by norm conserving pseudopotentials in the Kleinman-Bylander form<sup>55</sup> except for In pseudopotential (PAW method in the implementation of Kresse and Joubert) which was generated by Prof. Dal Corso and exported from THEOS web site (<http://theo.srv1.epfl.ch/Main/Pseudopotentials>). The wave functions are expanded on a plane wave basis set with a kinetic energy cut-off of 1088 eV. The integral over the first Brillouin zone are performed using a Monkhorst-Pack (1×1×1) k-point grid for the charge density and chemical shift tensor calculation for all MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** models. The shielding tensor is computed using the GIPAW<sup>56</sup> approach which permits the reproduction of the results of a fully converged all-electron calculation. The isotropic chemical shift  $\delta_{\text{iso}}$  is defined as  $\delta_{\text{iso}} = -[\sigma - \sigma^{\text{ref}}]$ , where  $\sigma$  is the isotropic shielding and  $\sigma^{\text{ref}}$  is the isotropic shielding of the same nucleus in a reference system. In the present case, the comparison between the experimental  $\delta_{\text{iso}}$  and calculated averaged  $\delta_{\text{iso}}$  <sup>15</sup>N chemical shift values for MIL-68-NH<sub>2</sub> ( $\delta_{\text{iso}} = 66$  ppm) allowed to determine  $\sigma^{\text{ref}}$  for this nucleus.

**NMR spectroscopy.** DNP surface enhanced NMR experiments were performed on a Bruker BioSpin 400 MHz (9.4 T) solid-state DNP NMR spectrometer operating with a Bruker Avance III HD console. The spectrometer is equipped with a triple resonance 3.2 mm low-temperature CPMAS probe set to <sup>1</sup>H-<sup>13</sup>C-<sup>15</sup>N configuration ( $\omega_{\text{1H}}/(2\pi) = 400$  MHz,  $\omega_{\text{13C}}/(2\pi) = 100$  MHz,  $\omega_{\text{15N}}/(2\pi) = 79.4$  MHz). Dynamic

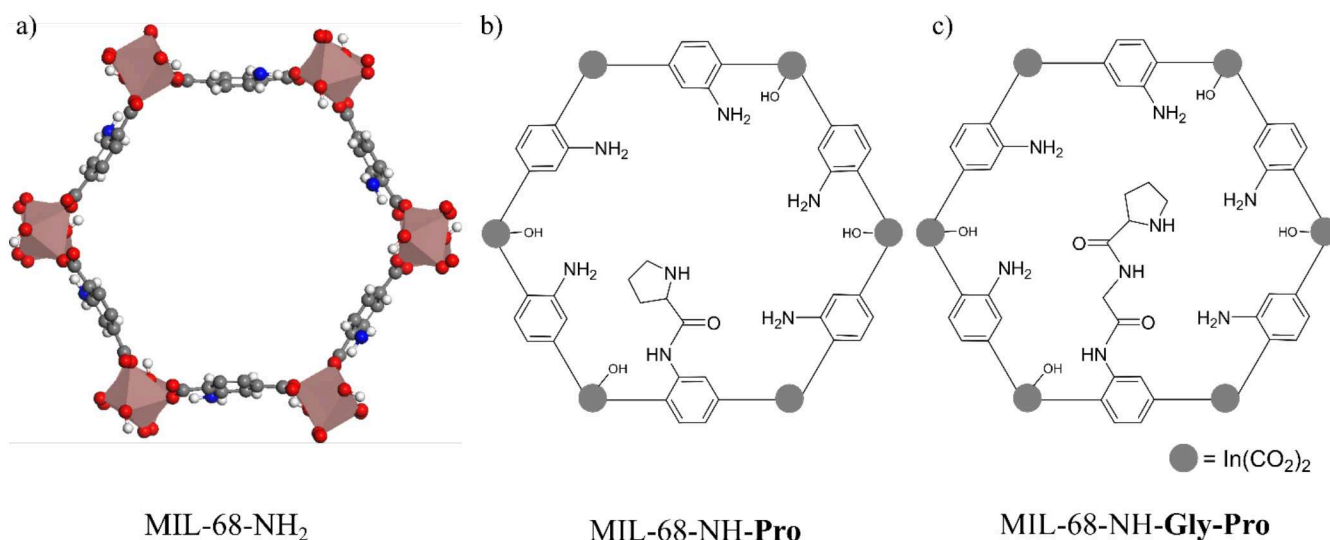
Nuclear Polarization is achieved by irradiating the sample with microwaves at a frequency of 263 GHz generated by a gyrotron and delivered to the sample by a waveguide. The CE matching condition was set to the maximum positive enhancement of AMUPOL polarizing agent<sup>57</sup> using the sweep coil of the spectrometer magnet. Sapphire rotors (endowed with zirconia caps) were used for optimal microwave penetration. Rotors are closed with a polytetrafluoroethylene insert to prevent loss of solvent during the measurements. Spectra acquisition was carried out at ~110 K with a spinning frequency of 8 kHz.

Following the same approach used in Ref. <sup>42</sup>, the samples for DNP SENS were prepared according to the incipient wetness impregnation procedure. MIL-68-NH-**Gly-Pro** was impregnated with a 16 mM solution of TEKPol2<sup>58</sup> in 1,1,2,2 tetrachloroethane (TCE). TEKPol2 is currently the polarizing agent giving the most efficient polarization transfer in organic solutions at 100 K, and has replaced bCTbK<sup>59</sup> used in our previous study. After microwave power optimization, a signal enhancement factor of  $\epsilon_H = 60$  was obtained for the solvent resonance on the impregnated material, as measured by scaling the intensity in one-dimensional spin echo <sup>1</sup>H spectra recorded with microwaves on and off. The <sup>1</sup>H-<sup>15</sup>N DNP SENS CPMAS (cross polarization magic angle spinning) spectrum of MIL-68-NH-**Gly-Pro** at natural abundance has been obtained with a signal to noise ratio of 91 in 11 h of signal averaging. Chemical shifts are referenced to the NH<sup>+</sup> resonance ( $\delta_{iso} = 0$  ppm) of ammonium nitrate.

## Results and Discussion

**Model construction.** Experimental findings indicate that the successful grafting of proline in the 2D hexagonal channels of MIL-68-NH<sub>2</sub> is performed with ~10 % yield providing us a quantitative guideline for the *in silico* grafting.<sup>34,35</sup> MIL-68-NH<sub>2</sub>,<sup>60</sup> which is synthesized with aminated 1,4-benzenedicarboxylate (bdc), is isostructural to MIL-68 (orthorhombic system, *Cmcm* space group,  $a = 21.774$  Å,  $b = 37.677$  Å,  $c = 7.233$  Å).<sup>49,50</sup> MIL-68 was thus used as a starting structural model to build MIL-68-NH<sub>2</sub> (see SI). For that purpose, it was recasted to its primitive cell, thereby reducing the cell volume and facilitating the further grafting of amino acids in a single hexagonal channel. Notably DFT-D3 level calculations showed that optimized cell parameters of MIL-68 ( $a = 21.958$  Å,  $b = 21.958$  Å,  $c = 7.287$  Å) were in excellent agreement with experiment ( $a = 21.758$  Å,  $b = 21.758$  Å,  $c = 7.233$  Å) with less than 1 % error. All further DFT calculations were thus performed with cell parameters fixed at the experimental ones. The unit cell of MIL-68-NH<sub>2</sub> was then doubled along the *c*-direction and grafted with one proline or one glycine-proline dipeptide on one single aminated bdc-NH<sub>2</sub> linker out of the 12 available ones resulting in a 8 % grafting rate. Figure 1 shows a schematic representation of the 6-membered channel in MIL-68-NH<sub>2</sub>, and the peptide-functionalized MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro**.



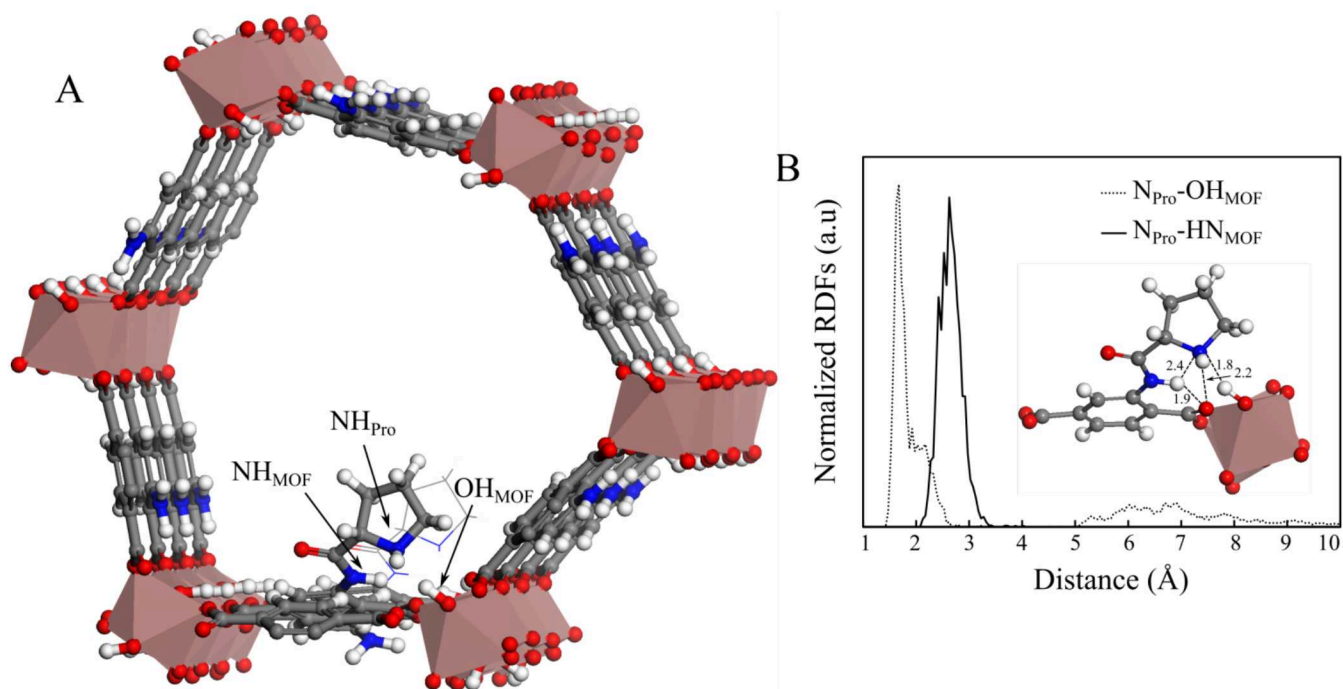


**Figure 1:** a) View along the *c*-axis of MIL-68-NH<sub>2</sub>; In atoms are depicted as brown polyhedra and the bdc-NH<sub>2</sub> linkers in sticks and balls: C is grey, O is red, N is blue and H is white. Schematic representation of b) MIL-68-NH-Pro and c) MIL-68-NH-Gly-Pro viewed along the *c*-axis. Pro = proline, Gly = glycine.

### 1. Key interactions at play in peptide grafted MOFs.

**Conformational search.** In order to identify favorable conformations of –Pro and –Gly-Pro grafts, we computed at the DFT level a large number of periodic models by exploring various conformations of the grafts in terms of i) their position in the channel with respect to terminal –OH groups and ii) the orientation of the grafts. MIL-68-NH-Pro and MIL-68-NH-Gly-Pro models are summarized in Figures S8-S9, respectively, and only representatives of the corresponding most favorable conformations are illustrated in Figures 2 and 3. By construction, the grafted residue or dipeptide establishes a peptide bond with the MOF, the latter playing the role of the N-ter and the residue that of the C-ter, by analogy with the terminology used for describing proteins.

All the computed grafted MIL-68-NH-Pro and MIL-68-NH-Gly-Pro structures exhibit one common feature consisting in a strong hydrogen bond,  $\text{NH}_{\text{MOF}} \cdots \text{O}_{\text{bdc}}$  (1.79–1.90 Å), between the amide H atom of the peptide –NH<sub>MOF</sub>–CO– bond and the carboxyl oxygen atom of the linker which holds the graft. Note that similar, though longer, H-bonds (1.90–1.95 Å) are found in the parent MIL-68-NH<sub>2</sub> and in the functionalized MOFs around non-grafted linkers. The H atom of the NH<sub>MOF</sub>–CO– peptide bond can also form weaker H-bonds with proximal terminal –OH groups (within 2.40–2.90 Å) (see SI for details). Importantly, this local hydrogen-bond network, combined with the planarity of the peptide bond, has two effects: it prohibits rotation of the graft around the C<sub>bdc</sub>–NH<sub>MOF</sub> bond; and it imposes that the carbonyl group of the NH<sub>MOF</sub>–CO– peptide bond points towards the channel rather than the framework. As a result, the first residue is predicted to be anchored to the MOF in a constrained orientation. The key interactions at play in both MIL-68-NH-Pro and MIL-68-NH-Gly-Pro functionalized MOFs are discussed below.



**Figure 2.** MIL-68-NH-Pro: A) Most favorable conformation of the proline graft. B) Normalized radial distribution function (RDF) plots of selected atom pairs in arbitrary units along with selected bond distances (in Å) shown in the inset.

**MIL-68-NH-Pro.** The computed most energetically favorable MIL-68-NH-Pro conformation (Figure 2A) forms a strong hydrogen bond,  $N_{\text{Pro}} \cdots \text{HO}_{\text{MOF}}$  ( $\sim 1.8$  Å), between the N-atom of the proline and a terminal -OH of the hybrid framework. Similar conformations were found in variant models (within  $3.5$  kcal mol $^{-1}$  from DFT calculations, see Figure S8). The absence of this H-bond interaction is systematically associated with less stable models. Importantly, the -Pro graft is predicted to interact with the hybrid framework through additional hydrogen-bonds that emanate from its -NH functional group. They include i) an intramolecular interaction,  $N_{\text{Pro}} \cdots \text{HN}_{\text{MOF}}$ , between  $N_{\text{Pro}}$  and the H-atom of the vicinal  $\text{NH}_{\text{MOF}}\text{-CO}$  peptide bond ( $\sim 2.4$  Å, inset in Figure 2B), ii)  $\text{NH}_{\text{Pro}} \cdots \text{O}_{\text{bdc}}$  contacts between the proline -NH group and the carboxyl oxygen of the linker and iii) possibly  $\text{NH}_{\text{Pro}} \cdots \text{NH}_2\text{-bdc}$  within  $\sim 2.3$  Å with the amino group of a neighboring non-grafted linker (see Fig.S8).

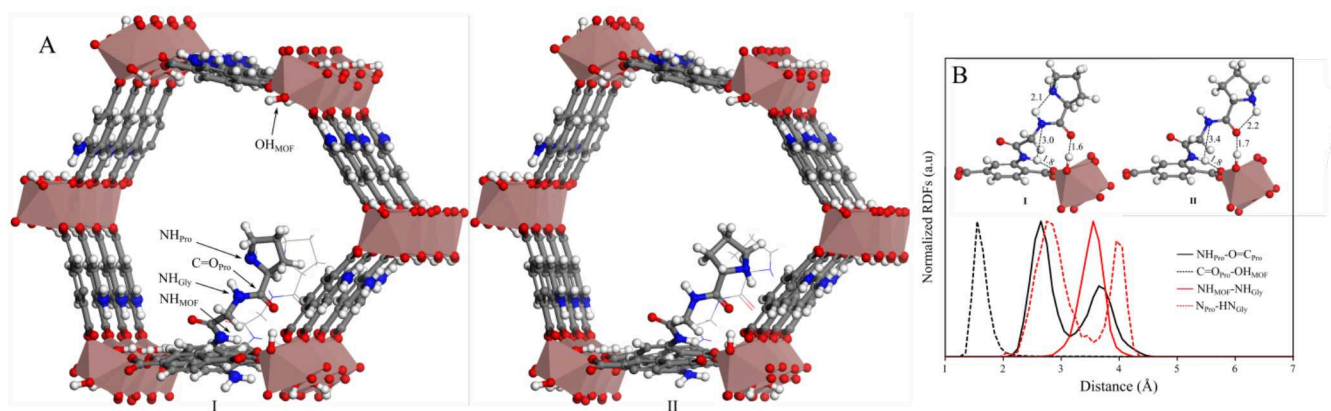
This MIL-68-NH-Pro conformation was further used as the starting point to perform a 5 ns MD calculation at 298 K. The analysis of the MD trajectory showed no significant deviation of the grafted proline conformations from the starting one. The root mean square deviation (RMSD) histogram of the graft's heavy atoms represents the variation of N, C, O atomic coordinates along the MD trajectories, taking the starting DFT-optimized conformation as a reference. RMSDs were calculated from the collected 1000 frames over the 5 ns MD runs and reveal a clear mono-modal distribution of the proline conformations with only a  $1.6$  Å deviation, indicating that only one conformation prevails at room temperature (see Fig.S13(A) in SI). The radial distribution functions (RDFs) of pairs of atoms of interest (Fig.2B) show that the interactions identified above between the proline and the framework are maintained throughout the MD run. They confirm the persistence of this conformation at room temperature as evidenced by the strong H-bond (peak centered at  $\sim 1.8$  Å assigned to  $N_{\text{Pro}} \cdots \text{HO}_{\text{MOF}}$

contacts) supplemented by an intramolecular H-bond interaction with the proximal peptide bond, (peak centered  $\sim 2.5$  Å corresponding to  $N_{\text{Pro}} \cdots \text{HN}_{\text{MOF}}$  contacts). MD runs performed on other favorable conformers yielded very similar trajectories and RDFs.

**MIL-68-NH-Gly-Pro.** We find that the **Gly** spacer profoundly affects the interactions at play for the next proline residue and particularly the nature of its accessible conformations. Overall, both DFT and MD calculations tend to show that the **-Gly** spacer, by moving the proline “away” from the pore walls, allows a larger conformational freedom of the proline in MIL-68-NH-**Gly-Pro** than observed in MIL-68-NH-**Pro**. At the DFT level, two families of favored conformations, noted as I and II, were identified as stable minima at 0 K (Figure 3A). We find that the site chosen for the grafting step in the channel causes their relative energy to span a range of 3-6 kcal mol<sup>-1</sup>. The two families differ in that the **-Pro** residue is rotated in relation to the **-Gly** residue due to a “flip” around the  $C_{\text{GLY}}-C_{\text{PRO}}$  bond between the glycine and the proline. Depending on its orientation, the N-ter proline may thus be oriented towards either the channels voids or the MOF’s wall. In conformer I, the nitrogen of the proline establishes an intramolecular H-bond (2.1 Å) with the H-atom of the vicinal  $\text{NH}_{\text{GLY}}-\text{CO}$  peptide bond,  $N_{\text{Pro}} \cdots \text{HN}_{\text{Gly}}$ . In conformer II, the dipeptide exhibits an extended conformation upon the flip of the proline cycle, so that the intramolecular interaction (2.2 Å) takes place with the carbonyl oxygen of the peptide bond,  $\text{NH}_{\text{Pro}} \cdots \text{O}=\text{C}_{\text{Pro}}$ . Moreover, both conformers are characterized by strong hydrogen bond interactions between the carbonyl group of the **-Pro** residue and one terminal hydroxyl group,  $\text{C}=\text{O}_{\text{Pro}} \cdots \text{HO}_{\text{MOF}}$ , at  $\sim 1.7$  Å. This H-bond plays a key role in the stabilization of the whole dipeptide as it anchors the spacer to the MOF’s framework, while leaving the proline rotational freedom. In the absence of such an H-bond between the spacer and the MOF, the conformations of the **-Gly-Pro** graft are found to be significantly less stable by  $\sim 15$  kcal mol<sup>-1</sup> (Fig. S9, models C and F). Finally, the graft could form additional hydrogen bonds with the  $-\text{NH}_2$ -bdc or the oxygens of the carboxyl groups from the linker (see Fig.S9).

The RDFs collected for MIL-68-NH-**Gly-Pro** over MD runs exhibit more complex features than those collected for MIL-68-NH-**Pro**. This is consistent with having a longer dipeptide graft and the occurrence of two favored conformers for the graft at room temperature. Similarly to the MIL-68-NH-**Pro** case study, the key interactions in MIL-68-NH-**Gly-Pro** are maintained throughout the dynamics run. The RDF peak centered around 1.7 Å (Fig.3B, black dotted line) captures the common feature of both conformers, I and II, i.e. the anchorage of the peptide bond through a strong H-bond,  $\text{C}=\text{O}_{\text{Pro}} \cdots \text{HO}_{\text{MOF}}$ , between its carbonyl group and a terminal hydroxyl group of the MOF, with the orientation of the amide group towards the center of the pore rather than towards the pore walls. Intramolecular interactions at play in both conformers are also well recapitulated in the  $N_{\text{Pro}} \cdots \text{HN}_{\text{Gly}}$  RDF peak, typical of conformer I (Fig.3B, red dotted line) and in the  $\text{NH}_{\text{Pro}} \cdots \text{O}=\text{C}_{\text{Pro}}$  RDF peak, typical of conformer II (Fig.3B, black solid line).





**Figure 3. MIL-68-NH-Gly-Pro:** A) Representative low energy conformations of the glycine-proline graft. B) Normalized RDFs plots of selected atom pairs in arbitrary units along with selected bond distances (in Å) of conformers I and II, shown in the inset.

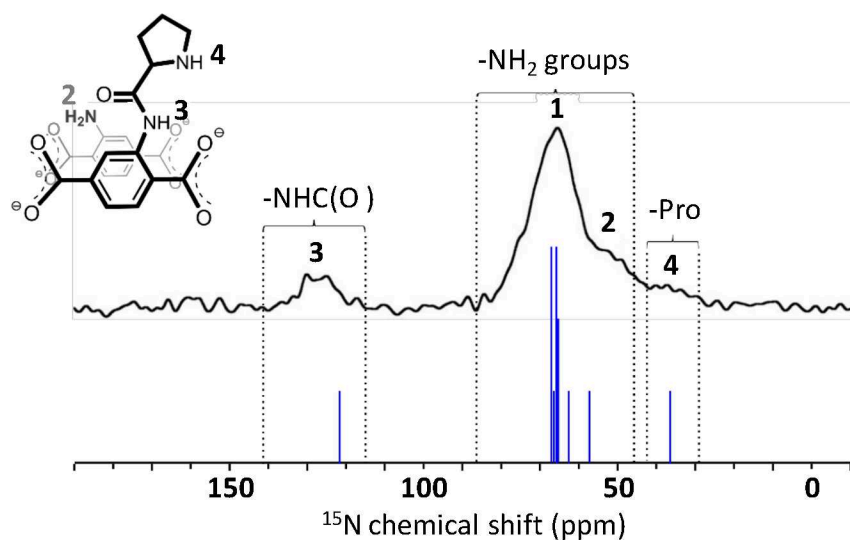
Hence, the force field results are in line with the DFT analysis based of the most favorable conformers, in support of the chemical significance of our extended PCFF+ all-atom force field method. DFT and MD calculations indicate that the conformations of both fragments are not freely distributed, but they are strongly directed by H-bonding between the amino-acids and the framework hydroxyl group. For the -Pro graft, this leads to only one type of stable conformation, and for the -Gly-Pro graft, two types of stable conformations with similar energies are identified.

## 2. <sup>15</sup>N chemical shifts in peptide functionalized MOFs

Natural abundance <sup>15</sup>N solid-state NMR experiments on hybrid materials still remain considerably challenging due to the low gyromagnetic ratio ( $\nu_0 = 40.58$  MHz at 9.4 T) and low natural abundance (N.A. = 0.37%) of <sup>15</sup>N. Therefore higher sensitivity and shorter acquisition time make DNP-enhanced solid-state NMR a very powerful tool to unravel the local structure of widely used N-containing active groups such as amino acids<sup>61</sup> and catalyst ligands like bipyridine or N-heterocyclic carbenes grafted on hybrid solids.<sup>62,63</sup>

Having identified the key interactions predicted to be at play in the peptide-functionalized MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro**, and in particular the stabilization of the grafted amino acids through interactions with the MOF framework, we compare the <sup>15</sup>N NMR signatures generated from combined DFT/MD computations with those recorded experimentally from <sup>1</sup>H-<sup>15</sup>N DNP SENS CPMAS spectra. This allowed us to (i) validate the computational models and (ii) confirm the experimentally assigned chemical shifts of nitrogen species in amino acids-functionalized MOFs.

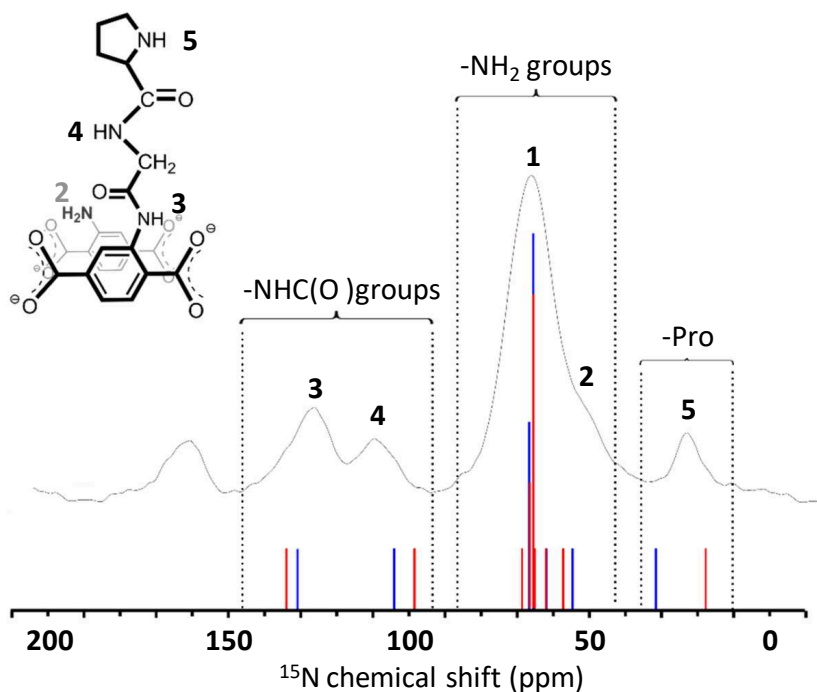
We recall here that the MIL-68-NH-**Pro** unit cell used to model all NMR spectra contains twelve bdc linkers as part of two successive hexagonal windows and includes one functionalized bdc. The simulated <sup>15</sup>N NMR pattern thus gives access to thirteen signals in the case of -**Pro** and fourteen in the case of -**Gly-Pro** (indicated as bars in Fig.4) accounting for all nitrogen atoms of the MOF. The computed <sup>15</sup>N chemical shifts for the most favorable MIL-68-NH-**Pro** structure (Fig.2A) is plotted along with the experimental spectrum<sup>42</sup> in Figure 4. Full set of computed spectra are shown in Fig. S10.



**Figure 4:** Assignment of the experimental  $^{15}\text{N}$  DNP SENS CPMAS spectrum of MIL-68-NH-Pro from Ref. <sup>42</sup> according to the simulated  $^{15}\text{N}$  spectrum shown as bars, for the most favorable conformation of the grafted proline as shown in Figure 2A. The height of the bars is proportional to the number of corresponding N atoms.

As expected, one signal corresponding to the amido N atom appears at high chemical shift value (121 ppm) in the simulated pattern, in a good agreement with the experimental peak 3 recorded at  $\delta_{\text{iso}} \sim 127$  ppm. Similarly, the  $-\text{NH}_2$  groups of the non-grafted bdc linkers give rise to intense peaks at 62-67 ppm which correspond to the recorded most intense broad peak 1 at  $\delta_{\text{iso}} \sim 66$  ppm. Interestingly the assignment of the two remaining peaks from DFT calculations differs from the early assignment proposed by Rossini *et al.* from their pioneering experimental study.<sup>42</sup> Indeed, the  $^{15}\text{N}$  NMR shift calculations here suggest that the shoulder of peak 1 recorded at  $\delta_{\text{iso}} \sim 55$  ppm (computed at  $\delta_{\text{iso}} = 57.7$  ppm) noted as peak 2 in the Figure 4, is not the signal of the amino N atom of the proline but rather a signal of a  $-\text{NH}_2$  group of a bdc linker. This linker is in the vicinity of the graft, in which the  $-\text{NH}_2$  group is displaced towards the 3-membered channel and forms an H-bond with a carboxyl oxygen belonging to another linker of this channel (see Fig S12). Finally our calculations find that the signal from the N atom of the grafted proline is at  $\delta_{\text{iso}} = 37.3$  ppm. This signal corresponds in the recorded  $^{15}\text{N}$  NMR spectrum of MIL-68-NH-Pro to a very weak signal in the 35 ppm region, now referred to as peak 4, which was considered as noise in the earlier interpretation.<sup>42</sup> The position of this  $^{15}\text{N}$  signal demonstrates the strong influence of H-bonding between the proline moiety and the hydroxyls of MIL-68 host as calculations show a shift of this peak 4 to  $\delta_{\text{iso}} = 19$  ppm when an unfavorable conformation is adopted whereby the proline residue does not form any hydrogen bonds with the MOF framework (Fig.S8, model F). This new assignment was further confirmed in the case of a peptide-functionalized MOF with the larger and more flexible glycine-proline dipeptide. The  $^1\text{H}$ - $^{15}\text{N}$  DNP SENS CPMAS spectrum of MIL-68-NH-Gly-Pro reported here was recorded following a novel protocol based on the use of the recently introduced biradical TEKPol2, which allows to further increase the DNP signal enhancement of the experiment from  $\epsilon_{\text{H}}=16$  as previously reported for MIL-68-NH-Pro in ref.<sup>42</sup> to  $\epsilon_{\text{H}} = 60$  in the present work. Figure 5 shows the experimental spectrum along with the computed  $^{15}\text{N}$  chemical shifts (bars) for the most favorable conformations I and II identified through our combined DFT/MD approach (see Fig. 3A). It should be noted that the signal at 161 ppm, which has no equivalent

in computed NMR spectra, is assigned to pyridine/pyridinium species<sup>64</sup> from the coupling agent used in the MIL-68-NH-Gly-Pro synthesis procedure (see ESI).<sup>35</sup>



**Figure 5:** Experimental  $^{15}\text{N}$  DNP NMR spectrum of MIL-68-NH-Gly-Pro along with the simulated  $^{15}\text{N}$  spectra corresponding to conformers I (blue bars) and II (red bars) presented in Figure 3A. The height of the bars is proportional to the number of corresponding N atoms.

The peaks **1** at  $\delta_{\text{iso}} \sim 66$  ppm and **3** at  $\delta_{\text{iso}} \sim 127$  ppm corresponds to an average of those calculated for the two conformations of MIL-68-NH-Gly-Pro and can be assigned respectively to free  $-\text{NH}_2$  groups and amido N atom of the MOF grafted to the glycine residue. Like in the case of proline, the average of computed chemical shifts  $\delta_{\text{iso}} = 56.5$  ppm reveals that the peak **2** recorded at  $\delta_{\text{iso}} \sim 55$  ppm corresponds to one non-grafted  $-\text{NH}_2$  group in the proximity of the graft. As expected the additional peak **4** calculated at  $\delta_{\text{iso}} = 105$  ppm for the conformation I and  $\delta_{\text{iso}} = 99$  ppm for the conformation II is in a good agreement with the recorded one at 107 ppm and can be assigned to amido N atom of the glycine residue. Unlike with the MIL-68-NH-Pro spectrum, the good spectral sensitivity here allowed for an unambiguous detection of the high field chemical shift signal at  $\delta_{\text{iso}} \sim 25$  ppm in the MIL-68-NH-Gly-Pro  $^{15}\text{N}$  spectrum. Based on the computed spectra, this signal can be assigned to the terminal proline N atom of the -Gly-Pro graft, as an average of two peaks calculated for conformations I and II, at  $\delta_{\text{iso}} = 31$  ppm and 18 ppm respectively.

Noteworthy, the 10 ppm  $^{15}\text{N}$  NMR shift for the signal of the terminal proline N atom in MIL-68-NH-Pro (35 ppm) and MIL-68-NH-Gly-Pro (25 ppm) is a fingerprint of a different chemical environment of the terminal amino acid. The DFT/MD calculations allow unambiguously attributing this difference to strong H-bonding interactions of the grafted peptide with the MIL-68 host framework: in MIL-68-NH-Pro proline interacts with the surface  $-\text{OH}$  groups, while in MIL-68-NH-Gly-Pro such interactions



are prevented by the glycine, which fulfills its role of a spacer by putting the proline moiety away from the MOF surface.

## Conclusion

Molecular level characterization of the structure and interactions at play in the peptide-functionalized MOFs MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** reveal that the conformations of both fragments are not freely distributed, but they are strongly directed by H-bonding interactions between the amino acids and the framework. As a result, we have predicted only one family of stable conformations for MIL-68-NH-**Pro** and two families of stable conformations with similar computed energies for MIL-68-NH-**Gly-Pro**. This highlights the key role of the surface –OH groups of the MOFs in anchoring the graft through strong hydrogen bonding interactions. Comparing calculated  $^{15}\text{N}$  chemical shifts to  $^{15}\text{N}$  obtained with DNP SENS, allows us to assign the spectra of MIL-68-**Pro** and MIL-68-**Gly-Pro**. Moreover, we re-assigned the signal emanating from the proline nitrogen as compared to the previously published spectrum of MIL-68-NH-**Pro**.<sup>42</sup> Highlighting differences in the dynamic behavior of small peptide grafts in a MOF host, we anticipate that the structure of the MOF host will be as essential as the amino acids sequence in the peptide in determining the accessible conformations of the grafts.

More broadly, the determination of the effect between the support and the functional group is of high interest for applications such as catalysis. Indeed the MOF host should not be considered as an inert carrier, but as an active support, whose surface functional groups (here –OH groups) can interact with anchored active species, as has been widely acknowledged for heterogeneous catalysis on inorganic supports such as silica, alumina and zeolites.<sup>65–67</sup> The existence of such interactions in the peptide-grafted MOF is clearly supported by the DNP-enhanced NMR results, and the combined DFT/MD approach points to their exact nature. The Pro $\cdots$ HO-MOF interactions highlighted here further demonstrate the active role of the MIL-68 framework support for catalytic applications, as previously observed by Canivet and coworkers for an asymmetric aldol reaction using peptide-grafted MIL-68.<sup>35</sup>

We anticipate that this work will help to direct the choice of platforms for the design of MOF-based hybrids such as heterogeneous catalysts. Commonly used MOF supports, chosen for their high accessible volume and stability such as MIL-101 or UiO-type structures, often present OH groups on their inner surface. Here we have shown that such hydroxyls can be involved in strong hydrogen bonding interactions with the grafted active sites at their vicinity, even though such interactions are often neglected or underestimated.

## ACKNOWLEDGEMENTS

We acknowledge financial support from the French National Research Agency (ANR project: HOPFAME ANR-13-BS07-0002-01), the Fondation de l'Orangerie and the ERC Advanced Grant No. 320860. We thank Prof. Paul Tordo and Dr Olivier Ouari for providing TEKPol2. The calculations have been performed using the HPC resources from GENCI (CINES/TGCC/IDRIS) through Grant (2016-097343 and -097535).



## References:

- (1) Zhang, T.; Lin, W. *Chem. Soc. Rev.* **2014**, *43* (16), 5982–5993.
- (2) Hu, Z.; Deibert, B. J.; Li, J. *Chem. Soc. Rev.* **2014**, *43* (16), 5815–5840.
- (3) Gascon, J.; Corma, A.; Kapteijn, F.; Llabrés i Xamena, F. X. *ACS Catal.* **2014**, *4* (2), 361–378.
- (4) Farrusseng, D.; Aguado, S.; Pinel, C. *Angew. Chem. Int. Ed.* **2009**, *48* (41), 7502–7513.
- (5) Liu, J.; Chen, L.; Cui, H.; Zhang, J.; Zhang, L.; Su, C.-Y. *Chem. Soc. Rev.* **2014**, *43* (16), 6011–6061.
- (6) Lillerud, K. P.; Olsbye, U.; Tilset, M. *Top. Catal.* **2010**, *53* (13–14), 859–868.
- (7) Gu, Z.-Y.; Park, J.; Raiff, A.; Wei, Z.; Zhou, H.-C. *ChemCatChem* **2014**, *6* (1), 67–75.
- (8) Zhang, M.; Gu, Z.-Y.; Bosch, M.; Perry, Z.; Zhou, H.-C. *Coord. Chem. Rev.* **2015**, *293–294*, 327–356.
- (9) Wu, C.-D.; Hu, A.; Zhang, L.; Lin, W. *J. Am. Chem. Soc.* **2005**, *127* (25), 8940–8941.
- (10) Bradshaw, D.; Claridge, J. B.; Cussen, E. J.; Prior, T. J.; Rosseinsky, M. J. *Acc. Chem. Res.* **2005**, *38* (4), 273–282.
- (11) Morris, R. E.; Bu, X. *Nat. Chem.* **2010**, *2* (5), 353–361.
- (12) Chen, L.; Reiss, P. S.; Chong, S. Y.; Holden, D.; Jelfs, K. E.; Hasell, T.; Little, M. A.; Kewley, A.; Briggs, M. E.; Stephenson, A.; Thomas, K. M.; Armstrong, J. A.; Bell, J.; Busto, J.; Noel, R.; Liu, J.; Strachan, D. M.; Thallapally, P. K.; Cooper, A. I. *Nat. Mater.* **2014**, *13* (10), 954–960.
- (13) Liu, Y.; Xuan, W.; Cui, Y. *Adv. Mater.* **2010**, *22* (37), 4112–4135.
- (14) Yoon, M.; Srirambalaji, R.; Kim, K. *Chem. Rev.* **2012**, *112* (2), 1196–1231.
- (15) Ma, L.; Abney, C.; Lin, W. *Chem. Soc. Rev.* **2009**, *38* (5), 1248–1256.
- (16) Ma, L.; Falkowski, J. M.; Abney, C.; Lin, W. *Nat. Chem.* **2010**, *2* (10), 838–846.
- (17) Song, F.; Wang, C.; Falkowski, J. M.; Ma, L.; Lin, W. *J. Am. Chem. Soc.* **2010**, *132* (43), 15390–15398.
- (18) Dang, D.; Wu, P.; He, C.; Xie, Z.; Duan, C. *J. Am. Chem. Soc.* **2010**, *132* (41), 14321–14323.
- (19) Peng, Y.; Gong, T.; Zhang, K.; Lin, X.; Liu, Y.; Jiang, J.; Cui, Y. *Nat. Commun.* **2014**, *5*, 4406.
- (20) Liu, T.; Liu, Y.; Xuan, W.; Cui, Y. *Angew. Chem. Int. Ed.* **2010**, *49* (24), 4121–4124.
- (21) Lin, Z.-J.; Lü, J.; Hong, M.; Cao, R. *Chem. Soc. Rev.* **2014**, *43* (16), 5867–5895.
- (22) Bailey, A. J.; Lee, C.; Feller, R. K.; Orton, J. B.; Mellot-Draznieks, C.; Slater, B.; Harrison, W. T. A.; Simoncic, P.; Navrotsky, A.; Grossel, M. C.; Cheetham, A. K. *Angew. Chem. Int. Ed.* **2008**, *47* (45), 8634–8637.
- (23) Appelhans, L. N.; Kosa, M.; Radha, A. V.; Simoncic, P.; Navrotsky, A.; Parrinello, M.; Cheetham, A. K. *J. Am. Chem. Soc.* **2009**, *131* (42), 15375–15386.
- (24) Zhang, S.-Y.; Wojtas, L.; Zaworotko, M. J. *J. Am. Chem. Soc.* **2015**, *137* (37), 12045–12049.
- (25) Vaidyanathan, R.; Bradshaw, D.; Rebilly, J.-N.; Barrio, J. P.; Gould, J. A.; Berry, N. G.; Rosseinsky, M. J. *Angew. Chem. Int. Ed.* **2006**, *45* (39), 6495–6499.
- (26) Rabone, J.; Yue, Y.-F.; Chong, S. Y.; Stylianou, K. C.; Bacsa, J.; Bradshaw, D.; Darling, G. R.; Berry, N. G.; Khimyak, Y. Z.; Ganin, A. Y.; Wiper, P.; Claridge, J. B.; Rosseinsky, M. J. *Science* **2010**, *329* (5995), 1053–1057.
- (27) Manton, A.; Massüger, L.; Rabu, P.; Palivan, C.; McCusker, L. B.; Taubert, A. *J. Am. Chem. Soc.* **2008**, *130* (8), 2517–2526.
- (28) Katsoulidis, A. P.; Park, K. S.; Antypov, D.; Martí-Gastaldo, C.; Miller, G. J.; Warren, J. E.; Robertson, C. M.; Blanc, F.; Darling, G. R.; Berry, N. G.; Purton, J. A.; Adams, D. J.; Rosseinsky, M. J. *Angew. Chem. Int. Ed.* **2014**, *53* (1), 193–198.
- (29) Wang, M.; Xie, M.-H.; Wu, C.-D.; Wang, Y.-G. *Chem. Commun.* **2009**, No. 17, 2396–2398.
- (30) Banerjee, M.; Das, S.; Yoon, M.; Choi, H. J.; Hyun, M. H.; Park, S. M.; Seo, G.; Kim, K. *J. Am. Chem. Soc.* **2009**, *131* (22), 7524–7525.
- (31) Tanabe, K. K.; Cohen, S. M. *Chem. Soc. Rev.* **2011**, *40* (2), 498–519.
- (32) Cohen, S. M. *Chem. Rev.* **2012**, *112* (2), 970–1000.
- (33) Zhao, J.; Li, H.; Han, Y.; Li, R.; Ding, X.; Feng, X.; Wang, B. *J. Mater. Chem. A* **2015**, *3* (23), 12145–12148.

- (34) Canivet, J.; Aguado, S.; Bergeret, G.; Farrusseng, D. *Chem. Commun.* **2011**, 47 (42), 11650–11652.
- (35) Bonnefoy, J.; Legrand, A.; Quadrelli, E. A.; Canivet, J.; Farrusseng, D. *J. Am. Chem. Soc.* **2015**, 137 (29), 9409–9416.
- (36) Kutzscher, C.; Nickerl, G.; Senkovska, I.; Bon, V.; Kaskel, S. *Chem. Mater.* **2016**, 28 (8), 2573–2580.
- (37) Yiannourakou, M.; Ungerer, P.; Leblanc, B.; Rozanska, X.; Saxe, P.; Vidal-Gilbert, S.; Gouth, F.; Montel, F. *Oil Gas Sci. Technol. – Rev. D'IFP Energ. Nouv.* **2013**, 68 (6), 977–994.
- (38) Ungerer, P.; Rigby, D.; Leblanc, B.; Yiannourakou, M. *Mol. Simul.* **2014**, 40 (1-3), 115–122.
- (39) Rozanska, X.; Ungerer, P.; Leblanc, B.; Saxe, P.; Wimmer, E. *Oil Gas Sci. Technol. – Rev. D'IFP Energ. Nouv.* **2015**, 70 (3), 405–417.
- (40) Rossini, A. J.; Zagdoun, A.; Lelli, M.; Lesage, A.; Copéret, C.; Emsley, L. *Acc. Chem. Res.* **2013**, 46 (9), 1942–1951.
- (41) Hoffmann, H. C.; Debowski, M.; Müller, P.; Paasch, S.; Senkovska, I.; Kaskel, S.; Brunner, E. *Materials* **2012**, 5 (12), 2537–2572.
- (42) Rossini, A. J.; Zagdoun, A.; Lelli, M.; Canivet, J.; Aguado, S.; Ouari, O.; Tordo, P.; Rosay, M.; Maas, W. E.; Copéret, C.; Farrusseng, D.; Emsley, L.; Lesage, A. *Angew. Chem. Int. Ed.* **2012**, 51 (1), 123–127.
- (43) Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996**, 6 (1), 15–50.
- (44) Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, 54 (16), 11169–11186.
- (45) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, 77 (18), 3865–3868.
- (46) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, 132 (15), 154104.
- (47) Blöchl, P. E. *Phys. Rev. B* **1994**, 50 (24), 17953–17979.
- (48) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, 59 (3), 1758–1775.
- (49) Barthelet, K.; Marrot, J.; Férey, G.; Riou, D. *Chem. Commun.* **2004**, No. 5, 520–521.
- (50) Volkringer, C.; Meddouri, M.; Loiseau, T.; Guillou, N.; Marrot, J.; Férey, G.; Haouas, M.; Taulelle, F.; Audebrand, N.; Latroche, M. *Inorg. Chem.* **2008**, 47 (24), 11892–11901.
- (51) France-Lanord, A.; Soukiasian, P.; Glatli, C.; Wimmer, E. *ArXiv151005416 Cond-Mat* **2015**.
- (52) Plimpton, S. J. *Comput. Phys.* **1995**, 117 (1), 1–19.
- (53) <http://lammps.sandia.gov>.
- (54) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Davide Ceresoli; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Corso, A. D.; Gironcoli, S. de; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Anton Kokalj; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Stefano Paolini; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. *J. Phys. Condens. Matter* **2009**, 21 (39), 395502.
- (55) Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett.* **1982**, 48 (20), 1425–1428.
- (56) Pickard, C. J.; Mauri, F. *Phys. Rev. B* **2001**, 63 (24), 245101.
- (57) Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P. *Angew. Chem. Int. Ed.* **2013**, 52 (41), 10858–10861.
- (58) Kubicki, D. J.; Casano, G.; Schwarzwälder, M.; Abel, S.; Sauvée, C.; Ganesan, K.; Yulikov, M.; Rossini, A. J.; Jeschke, G.; Copéret, C.; Lesage, A.; Tordo, P.; Ouari, O.; Emsley, L. *Chem. Sci.* **2015**, 7 (1), 550–558.
- (59) Zagdoun, A.; Casano, G.; Ouari, O.; Lapadula, G.; Rossini, A. J.; Lelli, M.; Baffert, M.; Gajan, D.; Veyre, L.; Maas, W. E.; Rosay, M.; Weber, R. T.; Thieuleux, C.; Coperet, C.; Lesage, A.; Tordo, P.; Emsley, L. *J. Am. Chem. Soc.* **2012**, 134 (4), 2284–2291.
- (60) Pera-Titus, M.; Savonnet, M.; Farrusseng, D. *J. Phys. Chem. C* **2010**, 114 (41), 17665–17674.
- (61) Werner, M.; Heil, A.; Rothermel, N.; Breitzke, H.; Groszewicz, P. B.; Thankamony, A. S.; Gutmann, T.; Buntkowsky, G. *Solid State Nucl. Magn. Reson.* **2015**, 72, 73–78.
- (62) Kobayashi, T.; Perras, F. A.; Slowing, I. I.; Sadow, A. D.; Pruski, M. *ACS Catal.* **2015**, 5 (12), 7055–7062.
- (63) Grüning, W. R.; Rossini, A. J.; Zagdoun, A.; Gajan, D.; Lesage, A.; Emsley, L.; Copéret, C. *Phys. Chem. Chem. Phys.* **2013**, 15 (32), 13270–13274.
- (64) Duthaler, R. O.; Roberts, J. D. *J. Am. Chem. Soc.* **1978**, 100 (16), 4969–4973.
- (65) Schauerermann, S.; Freund, H.-J. *Acc. Chem. Res.* **2015**, 48 (10), 2775–2782.
- (66) Schoenbaum, C. A.; Schwartz, D. K.; Medlin, J. W. *Acc. Chem. Res.* **2014**, 47 (4), 1438–1445.

- (67) Nørskov, J. K.; Bligaard, T.; Hvolbæk, B.; Abild-Pedersen, F.; Chorkendorff, I.; Christensen, C. H. *Chem. Soc. Rev.* **2008**, 37 (10), 2163–2171.

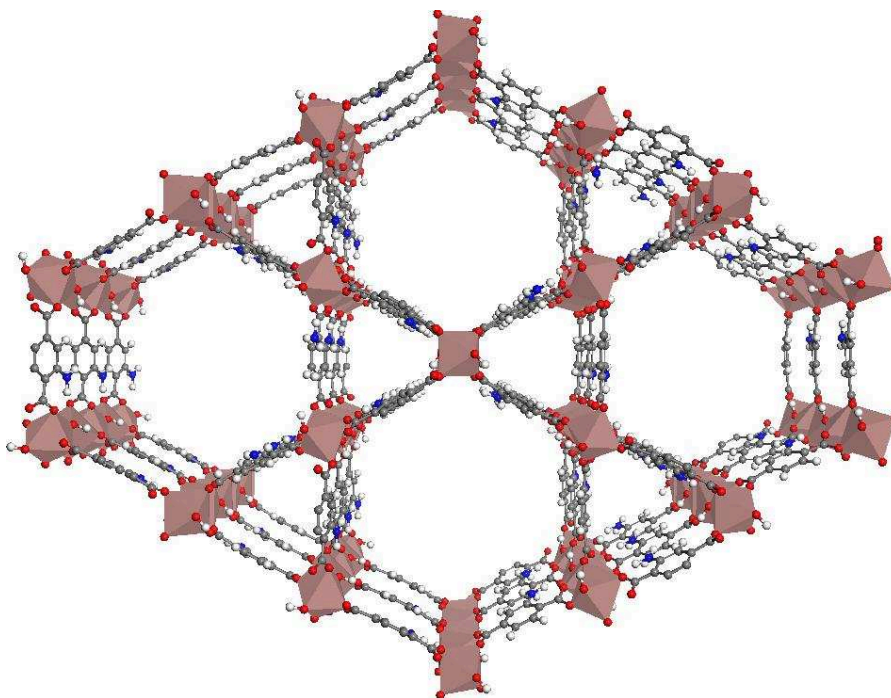
## Supporting Information

- SI-1: In-MIL-68-NH<sub>2</sub>
- SI-2: In-MIL-68-NH<sub>2</sub> forcefield parameterization
- SI-3: PCFF+ fitted parameters for In-MIL-68-NH<sub>2</sub>
- SI-4: Typical supercell used in the LAMMPS simulations
- SI-5: MIL-68-NH<sub>2</sub>-**Pro** models
- SI-6: MIL-68-NH<sub>2</sub>-**Gly-Pro** models
- SI-7: <sup>15</sup>N spectra of MIL-68-NH-**Pro** models
- SI-8: <sup>15</sup>N spectra of MIL-68-NH-**Gly-Pro** models
- SI-9: Grafted amino acid in the proximity of the 3-membered MOF channel
- SI-10: RMSDs distributions for MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro**
- SI-11: Experimental data for <sup>15</sup>N DNP SENS CPMAS acquisition



### SI-1: In-MIL-68-NH<sub>2</sub>

The In-MIL-68-NH<sub>2</sub> used for generating all MIL-68-NH-**Pro** and MIL-68-NH-**Gly-Pro** models is depicted in Figure S1. It contains mono-aminated benzene dicarboxylates (NH<sub>2</sub>-bdc) as linkers and is isostructural to MIL-68<sup>1,2</sup> (*Cmcm* space group,  $a = 21.774 \text{ \AA}$ ,  $b = 37.677 \text{ \AA}$ ,  $c = 7.233 \text{ \AA}$ ) as shown by Farrusseng and co-workers.<sup>3</sup> It is built up from infinite straight chains of corner-sharing metal-centered InO<sub>4</sub>(OH)<sub>2</sub> octahedra connected to each other through NH<sub>2</sub>-bdc, generating 1D channels and two hydroxyl groups oriented in *trans* positions.<sup>1,2</sup> The three-dimensional network of MIL-68-NH<sub>2</sub> features two types of channels, i.e., hexagonal and triangular 1D-channels with diameters of 16 and 6 Å, respectively. 1/3 of the hydroxyl groups point towards the voids of the triangular channels, while 2/3 is oriented towards the hexagonal ones. For the purpose of DFT calculations, In-MIL-68-NH<sub>2</sub> was recast to its primitive cell ( $a = 21.758 \text{ \AA}$ ,  $b = 21.758 \text{ \AA}$ ,  $c = 7.233 \text{ \AA}$ ), thereby reducing the volume of the crystallographic unit cell to a half of the conventional setting, and facilitating the further grafting of amino acids in a single hexagonal channel. For the construction of the MIL-68-NH<sub>2</sub> structural model, we have started from the experimental single-crystal structure of MIL-68<sup>2</sup> in which we substitute one single H atom on each phenyl ring by an NH<sub>2</sub>-group. In order to determine the most likely positions of -NH<sub>2</sub> functional groups, four models were generated where the amino group replaces in turn each of the four hydrogen atoms of the bdc linker. After geometry optimization, all four models differ in less than 0.02 eV (< 0.5 kcal mol<sup>-1</sup>), suggesting that all four are equally possible. Our calculations recurrently show that the six bdc-NH<sub>2</sub> linkers in a (1 × 1) unit cell of MIL-68-NH<sub>2</sub> are involved in four H-bond interactions between one H atom of the NH<sub>2</sub>-group and a proximal terminal hydroxyl groups (NH<sub>2</sub> ⋯ OH<sub>MOF</sub>): two H-bonds with OH-groups in the 6-membered channel and two H-bonds with OH-groups in the 3-membered channel. In addition, within each NH<sub>2</sub>-bdc linker, each -NH<sub>2</sub> group systematically forms H-bonds with the proximal carboxyl oxygen atoms from the bdc linker (NH<sub>2</sub> ⋯ O<sub>bdc</sub>). Interatomic distances are in the range of 2.70–2.90 Å and 2.40–2.48 Å for the former H-bonds in the 6- and 3-membered channels, respectively, and 1.90 – 1.95 Å for the interactions with the O<sub>bdc</sub>.



**Figure S1:** View of the In-MIL-68-NH<sub>2</sub> crystal structure shown along the *c*-axis. In atoms are depicted as brown polyhedral and bdc-NH<sub>2</sub> bridging linkers as balls and sticks (C: grey; H: white, N: blue, O: red).

## SI-2: In-MIL-68-NH<sub>2</sub> forcefield parameterization

PCFF+ is an extension of the Polymer Consistent Force Field (PCFF)<sup>4-9</sup> with adjustments of the non-bonding parameters to improve the agreement between computed and experimental available properties of fluids, namely density and heat of vaporization. PCFF is a second generation force field that is based on CFF91.<sup>4-6</sup> The parameterization of these forcefields was derived from *ab initio* calculations for all parameters. All systems that are described by PCFF and CFF91 are also inherited by PCFF+. However, transition metal and carboxyl oxygens as encountered in In-MIL-68-NH<sub>2</sub> are currently not described in PCFF+. In order to generate these missing parameters, a strategy was developed based on i) *ab initio* molecular dynamic (AIMD) trajectories of the solid of interest using VASP<sup>10,11</sup> followed by ii) a parameter fitting procedure to minimize the differences between the forcefield forces to those along the VASP trajectories, using the forcefield optimizer module as available in MedeA. The starting structure for AIMD simulations was the primitive cell of In-MIL-68-NH<sub>2</sub> as obtained from an initial geometry optimization using VASP as described in the

computational details section of the main text, and then reconstructed in its orthorhombic structure ( $a = 43.516 \text{ \AA}$ ,  $b = 37.705 \text{ \AA}$ ,  $c = 43.398 \text{ \AA}$ , and  $\alpha = \beta = \gamma = 90^\circ$ ) to the molecular mechanic forces estimations. Then, a series of AIMD simulations with different initial temperatures, varying from 300 to 800 K with a temperature step of 50 K, were performed in the canonical ensemble (NVT) during 30 fs with a time step of 1 fs. These different initial temperatures used for the AIMD simulations allowed sampling of various parts of the potential energy surface. Soft potentials were used for oxygen, carbon, and nitrogen, along with an energy cutoff of 283 eV. Calculations were done using the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional.<sup>12</sup> The long-range weak dispersion interactions were taken into account using the semiempirical van der Waals method of Grimme.<sup>13</sup> The electron-ion interactions were described by the projector augmented wave (PAW) method<sup>14</sup> in the implementation of Kresse and Joubert.<sup>15</sup> The Brillouin zone was sampled at the  $\Gamma$ -point.

All 330 frames obtained from the ab initio molecular dynamics simulations constituted the training and validation sets on which the forcefield parameters fitting procedure relied. To perform the parameter fitting procedure, we used the forcefield optimizer module available in the MedeA<sup>®</sup> simulation environment.<sup>16–19</sup> A more detailed description of the fitting procedures (Figure S2), capabilities and methods related to this module can be found in the study of France-Lanord *et al.*<sup>17</sup> The same module was successfully applied in other studies.<sup>18–19</sup> For each frame obtained from the AIMD simulations, total potential energy, forces along x, y, and z for each atom, and stress along x, y, and z are recorded. However, only atomic forces were used in the current forcefield parameters optimization. Typically, 50 frames were used in the *training* set pool, while all other frames were considered in the *validation* set pool. The forcefield parameters fitting starts by the introduction of the missing atom types, the labels of which are shown in Figure S3. The long-range interaction parameters for Lenard-Jones and coulombic interactions were not reoptimized after initially set as described next. These parameters are reported in Table S1. Lenard-Jones parameters were chosen from equivalent existing atom types in PCFF+ to conserve the overall transferability of the new atom type parameters. For omil and opmil oxygen atoms of the MIL-68

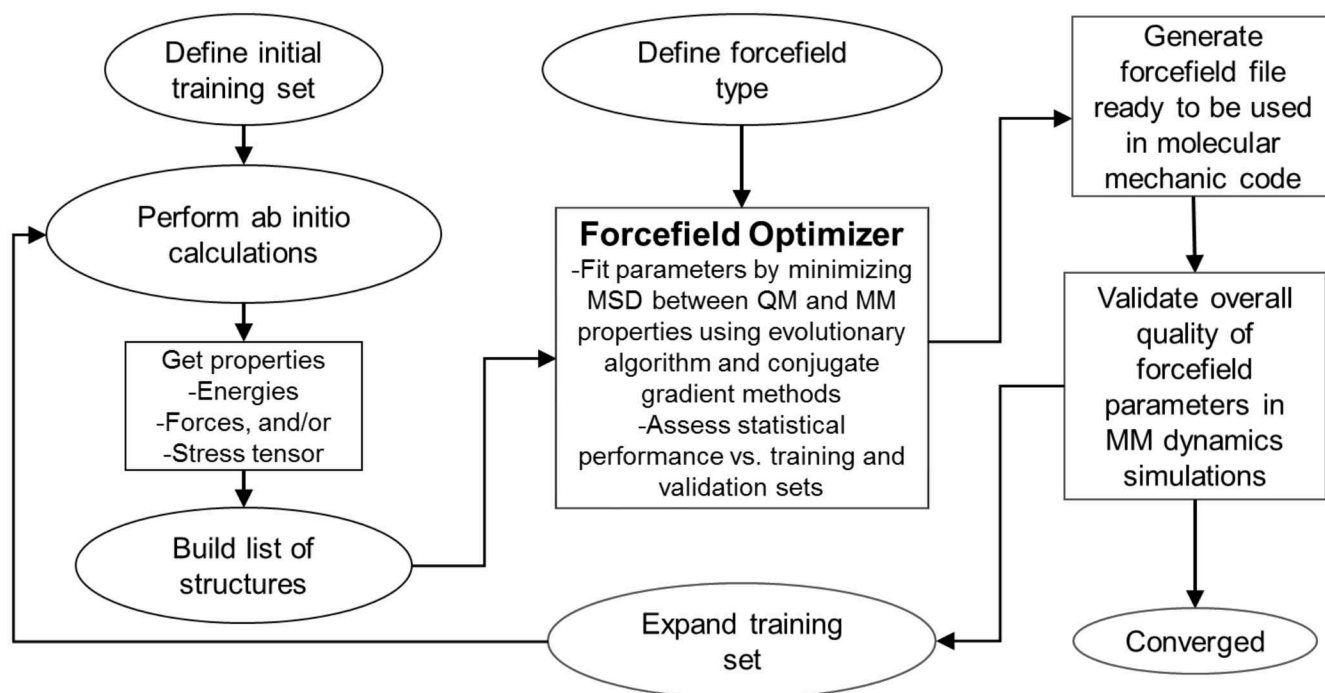
framework, Lenard-Jones parameters of oxygen in zeolite were chosen: indium oxide oxygen atom is closer to this atom type than an oxygen atom in an organic molecule. For the carbon atom cmil, the Lenard-Jones parameters of generic carbon atom were used. For the indium atom, we used the same parameters as for Rb<sup>+</sup>. The coulombic charges were determined from electrostatic potential fitted charges (ESP)<sup>20</sup> using Turbomole quantum chemistry package.<sup>21</sup> A large cluster was built from the In-MIL-68-NH<sub>2</sub> periodic structure and terminated by hydrogen atoms (Figure S4). A single point energy calculation at the PBE/SVP level<sup>12,22</sup> was done to obtain the ESP charges. Charges for the omil, opmil, hpmil, cmil and inmil atom types were averaged over the cluster, however not taking into account the atom charges close the cluster edges. Validation of these long range parameters was performed by simulation of the CO<sub>2</sub> adsorption isotherm using the classical Monte-Carlo software Gibbs<sup>23</sup> as implemented in MedeA.<sup>16,24–26</sup> Rigid molecule model is used for CO<sub>2</sub>: we first optimized the geometry of the CO<sub>2</sub> molecules with its PCFF+ parameters using the LAMMPS molecular mechanic software. In the subsequent Monte Carlo adsorption isotherm simulations performed with the Gibbs code, the solid is assumed to be rigid, and the Lennard-Jones and electrostatic potential energy grids characterizing the solid are generated prior the Monte-Carlo sampling. The grids were generated for an In-MIL-68-NH<sub>2</sub> supercell with dimensions 43.52 × 37.71 × 43.40 Å<sup>3</sup>. To avoid important CO<sub>2</sub> adsorption in the MOF as observed in the study of Yang *et al.*,<sup>27</sup> side channels that are not experimentally accessible were blocked using rare gas atoms. The Monte-Carlo simulations were done in the Grand Canonical ensemble<sup>28</sup> at 300 K with CO<sub>2</sub> fugacity of 1, 3, 5, 10, 13, 15, 20, 23, 25, 30, 33, 35, 40, 43, and 45 bar. For each simulation at a given CO<sub>2</sub> fugacity, an equilibration run of one million steps took place, followed by five million steps production run. The starting configuration for a given CO<sub>2</sub> fugacity was taken from the final configuration of the previous CO<sub>2</sub> fugacity simulation. A configurational bias was used in the Monte-Carlo sampling.<sup>29</sup> The Lenard-Jones pair energy cutoff was set to 18.8 Å, and the electrostatic potential energy was evaluated via Ewald summation, with the default Ewald summation parameters.<sup>30</sup> The CO<sub>2</sub> adsorption isotherm is reported in Figure S4 and compares favorably to the experimental and computed CO<sub>2</sub> adsorption isotherms reported in ref.<sup>27,31</sup> that were obtained for Al-MIL-68-NH<sub>2</sub>. The computed isosteric heat of



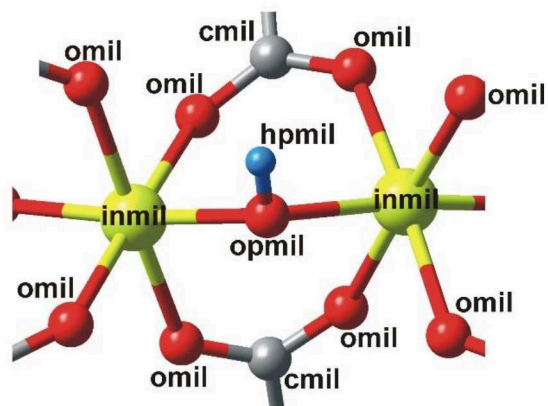
CO<sub>2</sub> adsorption is 5.2 kcal mol<sup>-1</sup> at 273 K when extrapolated at zero coverage, while the reported experimental value is 5.8 kcal mol<sup>-1</sup>.<sup>31</sup>

### SI-3: PCFF+ fitted parameters for In-MIL-68-NH<sub>2</sub>

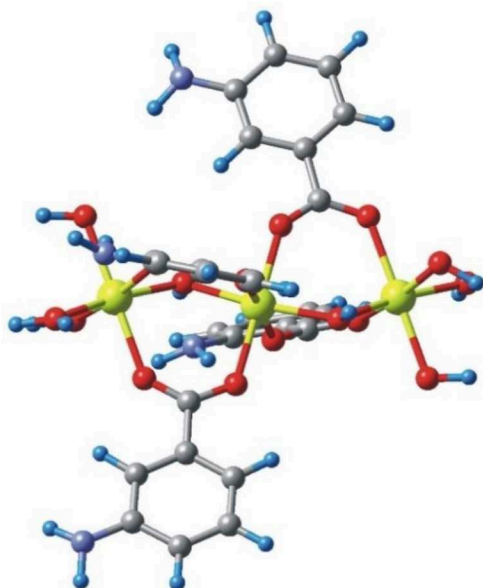
The bond and angle potential energy terms were fitted using the integrated forcefield optimizer (Figure S2).<sup>17-19</sup> Due to the limited number of parameters to optimize, each one was individually fitted iteratively through mean-square minimization of the forces computed with VASP (from the training set) and LAMMPS. First, we optimized the equilibrium bond distance or angle, followed by optimization of the force constant. The omil-inmil-omil angle parameter was an exception since it shows two minima at 90 and 180°. Therefore, we fixed the ratio of the quadratic force constants to get minima at these values and optimize them together through systematical sampling to best reproduce the VASP forces. For the torsion potential energy terms, we used equivalent terms available in the PCFF+ parameterization. A set of four dihedral angles and two Wilson out-of-plane terms were used to fine-tune the density of In-MIL-68-NH<sub>2</sub> because of the difficulty to control them through coulombic and Lenard-Jones terms. We systematically scaled this small number of dihedral and out-of-plane force constants parameters, performed series of isothermal-isobaric (NPT) simulation at 300 K, and computed the density for each. The torsion force constants that were rescaled were inmil-omil-cmil-cp, inmil-omil-cmil-omil, cmil-cp-cp-cp, and cp-cmil-omil-omil. We added two Wilson out-of-plane terms for cmil-cp-cp-cp and cp-cmil-omil-omil and initially fixed them to the same value than the PCFF+ cp-cp-cp-h out-of-plane parameter, *i.e.*  $K_{\psi}=15$  kcal mol<sup>-1</sup> rad<sup>-2</sup>. The results of this systematic search are reported in Figure S6. The best match is obtained with the scaling parameter is 2.10. All In-MIL-68-NH<sub>2</sub> parameters are summarized in Tables S1 - S5. Finally, geometry optimization of In-MIL-68-NH<sub>2</sub> performed with LAMMPS using the set of newly derived PCFF+ parameters is compared to the DFT optimized structure obtained with VASP. The root mean square deviation of the atomic coordinates of all atoms except the hydrogen ones amounts to 0.37 Å, indicating the quality of our parametrized version of PCFF+ forcefield, making it applicable in molecular dynamics simulations of peptide-functionalized In-MIL-68-NH<sub>2</sub> systems.



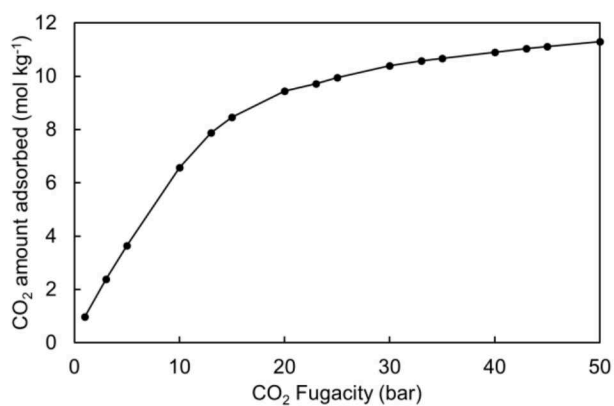
**Figure S2.** Schematic overall forcefield parameters fitting protocol.



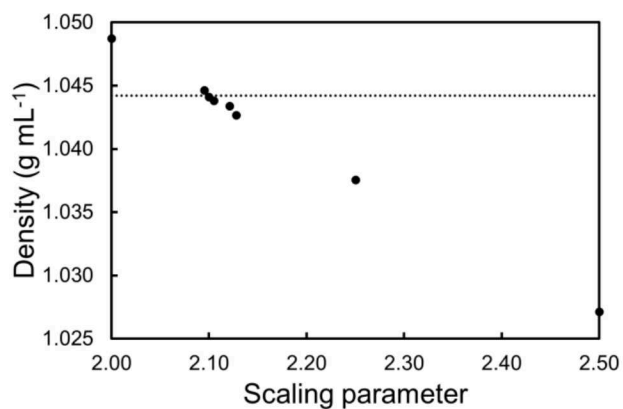
**Figure S3.** Atom types newly introduced in PCFF+ for In-MIL-68-NH<sub>2</sub>.



**Figure S4.**  $\text{In}_3\text{C}_{28}\text{N}_4\text{O}_{16}\text{H}_{33}$  cluster used in the ESP charge calculation.



**Figure S5.**  $\text{CO}_2$  adsorption isotherm in In-MIL-68- $\text{NH}_2$  at 300 K.



**Figure S6.** Effect of selected dihedral angle force constant rescaling on In-MIL-68- $\text{NH}_2$  density in NPT simulation at 300 K. The dotted line is the experimental value.

**Table S1.** Non-bonded parameters for In-MIL-68-NH<sub>2</sub> atoms obtained in this work.<sup>a</sup>

Atom type	$q$	$r^0$	$\varepsilon$
cmil	0.60	3.9147	0.0668
hpmil	0.40	1.087	0.008
inmil	1.40	1.9927	4.8433
omil	-0.60	2.3	0.086
opmil	-0.70	2.3	0.086

<sup>a</sup>:  $r^0$  is in Å,  $\varepsilon$  in kcal mol<sup>-1</sup>. The coulombic energy is  $E_C = \frac{q_i q_j}{\varepsilon_0 r_{ij}}$  and the Lenard-Jones potential

$$\text{energy } E_{LJ} = \varepsilon_{ij} \left( 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^9 - 3 \left( \frac{r_{ij}^0}{r_{ij}} \right)^6 \right), \text{ where } r_{ij} = \sqrt[6]{\frac{r_i^6 + r_j^6}{2}} \text{ and } \varepsilon_{ij} = \frac{2r_i^3 r_j^3 \sqrt{\varepsilon_i \varepsilon_j}}{r_i^6 + r_j^6}.$$

**Table S2.** Harmonic bond parameters for In-MIL-68-NH<sub>2</sub> atoms obtained in this work.<sup>a</sup>

Pair type	$R_0$	$K_2$
cmil-omil	1.29	420.0
cmil-cp	1.48	386.4
omil-inmil	2.30	294.0
opmil-inmil	2.20	279.3
opmil-hpmil	1.00	900.0

<sup>a</sup>:  $R_0$  is in Å, and  $K_2$  in kcal mol<sup>-1</sup> Å<sup>-2</sup>. The harmonic bond potential energy is  $E_{\text{bond}} = K_2(R - R_0)^2$ .

**Table S3.** Quartic angle parameters for In-MIL-68-NH<sub>2</sub> atoms obtained in this work.<sup>a</sup>

Angle type	$\theta_0$	$K_2$	$K_4$
cp-cp-cmil	120.0	199.5	0
cp-cmil-omil	120.0	136.5	0
cmil-omil-cmil	126.0	60.0	0
cmil-omil-inmil	126.0	136.5	0
omil-inmil-omil	135.0	-2000	1500
omil-cmil-omil	126.0	262.5	0
omil-inmil-opmil	85.0	262.5	0
inmil-opmil-inmil	116.0	283.5	0
inmil-opmil-hpmil	116.0	40.0	0
inmil-omil-inmil	117.02	60.0	0
opmil-inmil-opmil	170.0	336.0	0

<sup>a</sup>:  $\theta_0$  is in degree,  $K_2$  in kcal mol<sup>-1</sup> rad<sup>-2</sup>, and  $K_4$  in kcal mol<sup>-1</sup> rad<sup>-4</sup>. The quartic angle potential energy is  $E_{\text{angle}} = K_2(\theta - \theta_0)^2 + K_4(\theta - \theta_0)^4$ .

**Table S4.** Torsion parameters for In-MIL-68-NH<sub>2</sub> atoms obtained in this work.<sup>a</sup>

Torsion type	$V_1$	$\varphi_1^0$	$V_2$	$\varphi_2^0$	$V_3$	$\varphi_3^0$
cmil-cp-cp-n_3	0.073468	0	0.16643	0	0.1	0
cmil-cp-cp-n_2	0.073468	0	0.16643	0	0.1	0
cmil-cp-cp-cp	8.3667	0	1.1932	0	0	0
cmil-cp-cp-h	0.14893	2.701	0.13858	0.026333	0.01	0.21406
cmil-cp-cp-na	0.073468	0	0.16643	0	0.1	0
cp-cp-cmil-omil	0	0	1.1923	0	0	0
inmil-omil-cmil-cp	21.0	0	0	0	0	0
inmil-omil-cmil-omil	13.13	0	5.25	0	1.31	0
*-inmil-opmil-*	0.001	0	0	0	0	0
*-inmil-omil-*	0.001	0	0	0	0	0

<sup>a</sup>:  $V_1$ ,  $V_2$ , and  $V_3$  are in kcal mol<sup>-1</sup>,  $\varphi_1^0$ ,  $\varphi_2^0$ , and  $\varphi_3^0$  are in rad. The torsion potential energy is  $E_{\text{torsion}} = \sum_{i=1}^3 V_i (1 + \cos(n\varphi - \varphi_i^0))$ . n\_2 is the nitrogen atom type in urethane, carbamate like functional groups, e.g. N in -NH-C(=O)- groups. n\_3 is the nitrogen atom type in primary or secondary amide functional groups, e.g. N in -NR(R')-C(=O)-R" groups. na is the nitrogen atom type in sp<sup>3</sup> amine. cp is the carbon atom type in aromatic ring. h is the hydrogen atom type.

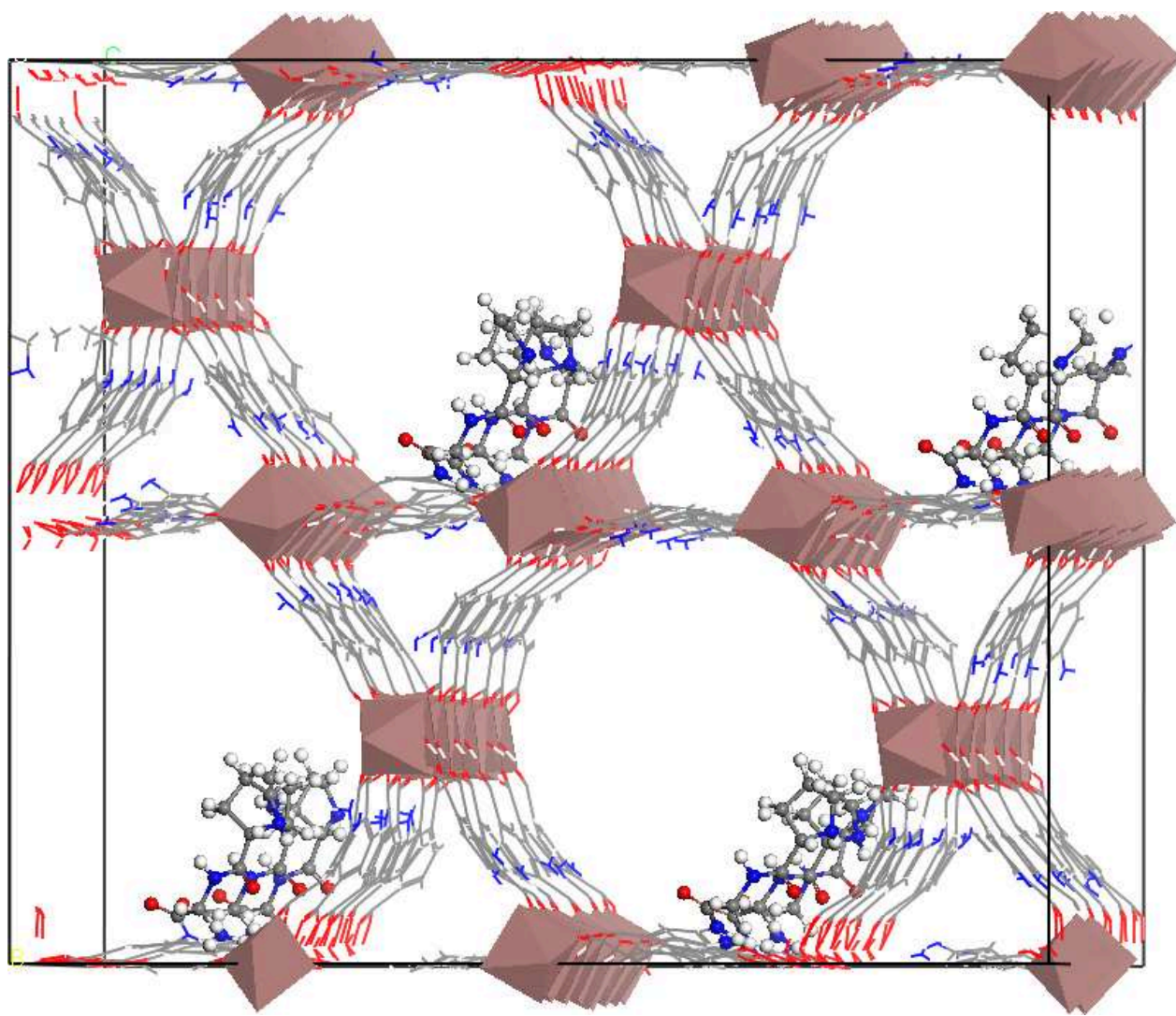
**Table S5.** Wilson out of plane parameters for In-MIL-68-NH<sub>2</sub> atoms obtained in this work.<sup>a</sup>

Out-of-plane type	$K_{\psi}$
cmil-cp-cp-cp	30.5
cp-cmil-omil-omil	30.5

<sup>a</sup>:  $K_{\psi}$  is in kcal mol<sup>-1</sup> rad<sup>-2</sup>. The Wilson out-of-plane potential energy is  $E_{\text{oop}} = K_{\psi}(\psi - \psi_0)^2$

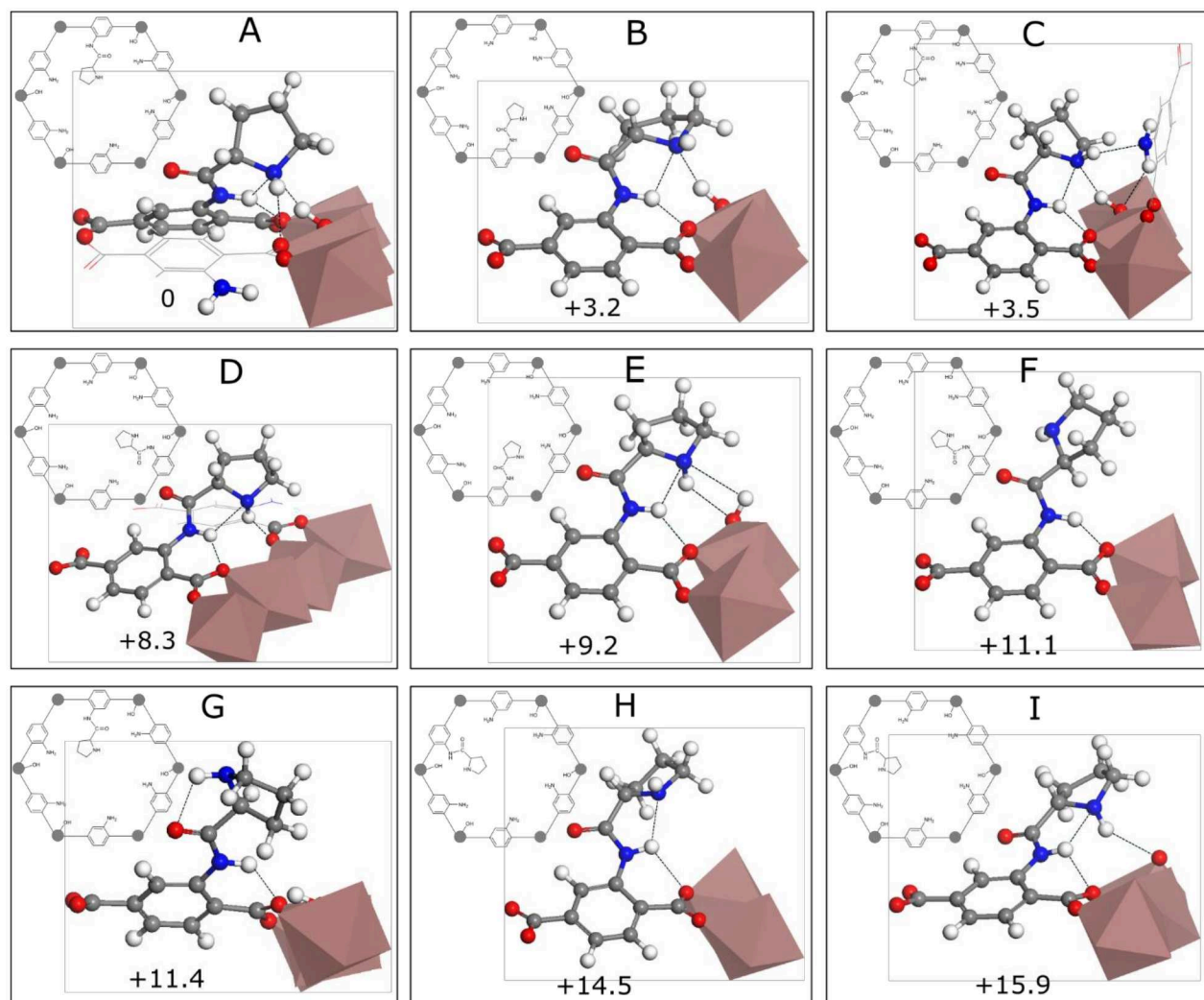


**SI-4: Typical supercell used in the LAMMPS simulations**



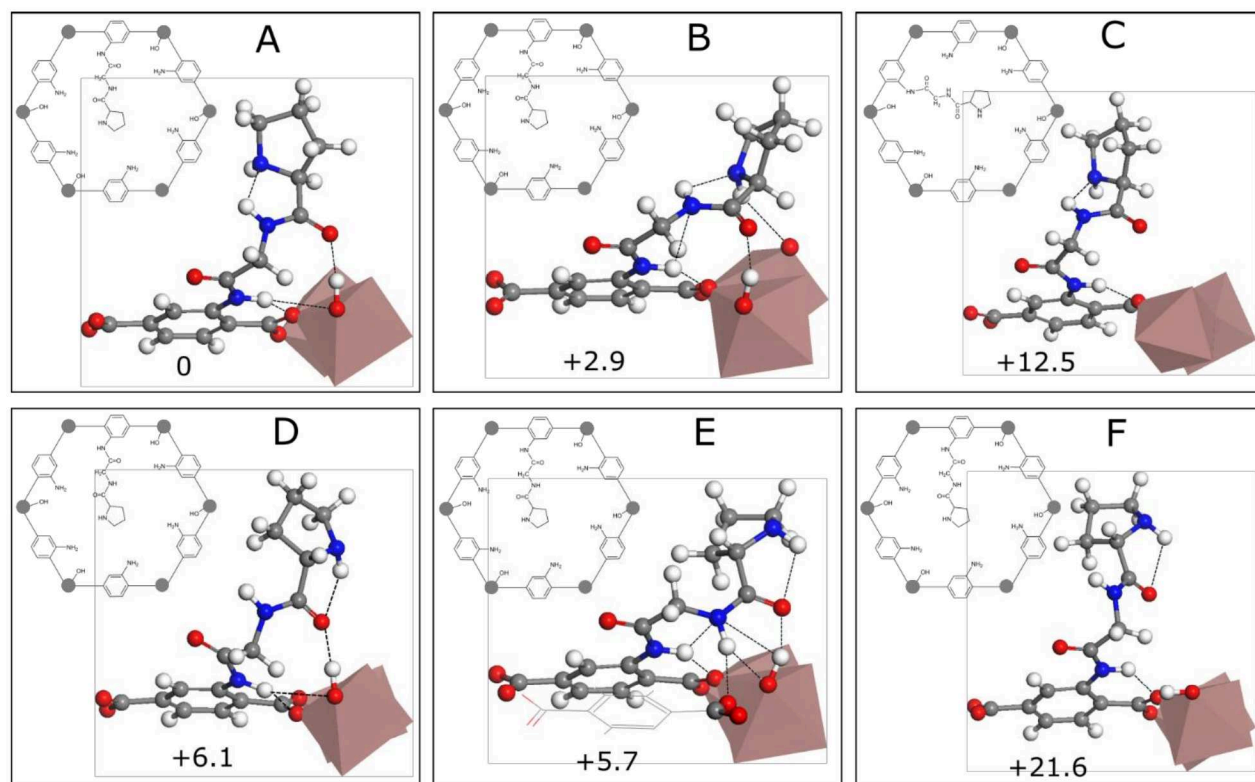
**Figure S7:** Supercell containing 12 grafted amino acids used in the MD simulations.

## SI-5: MIL-68-NH-Pro models



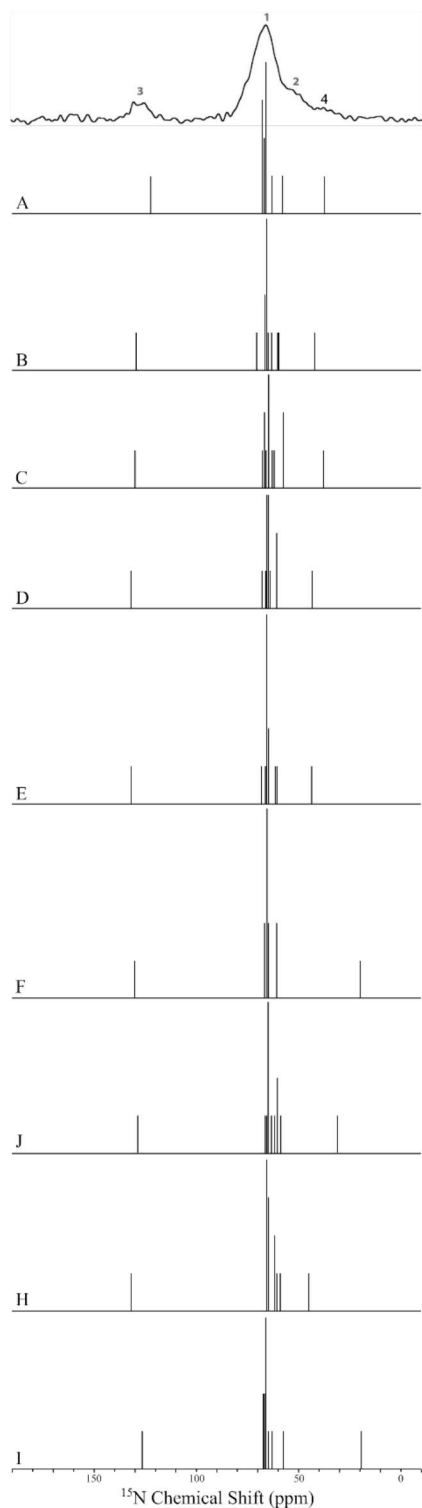
**Figure S8:** MIL-68-NH-Pro models along with the schematic representation of the **-Pro** grafting sites in the 6-membered channel viewed along the *c*-axis. In atoms are depicted as brown polyhedra, N is blue, O is red, C is gray, and H is white. Relative energies of the models are given in kcal mol<sup>-1</sup> relative to the most stable model **A**.

## SI-6: MIL-68-NH-Gly-Pro models



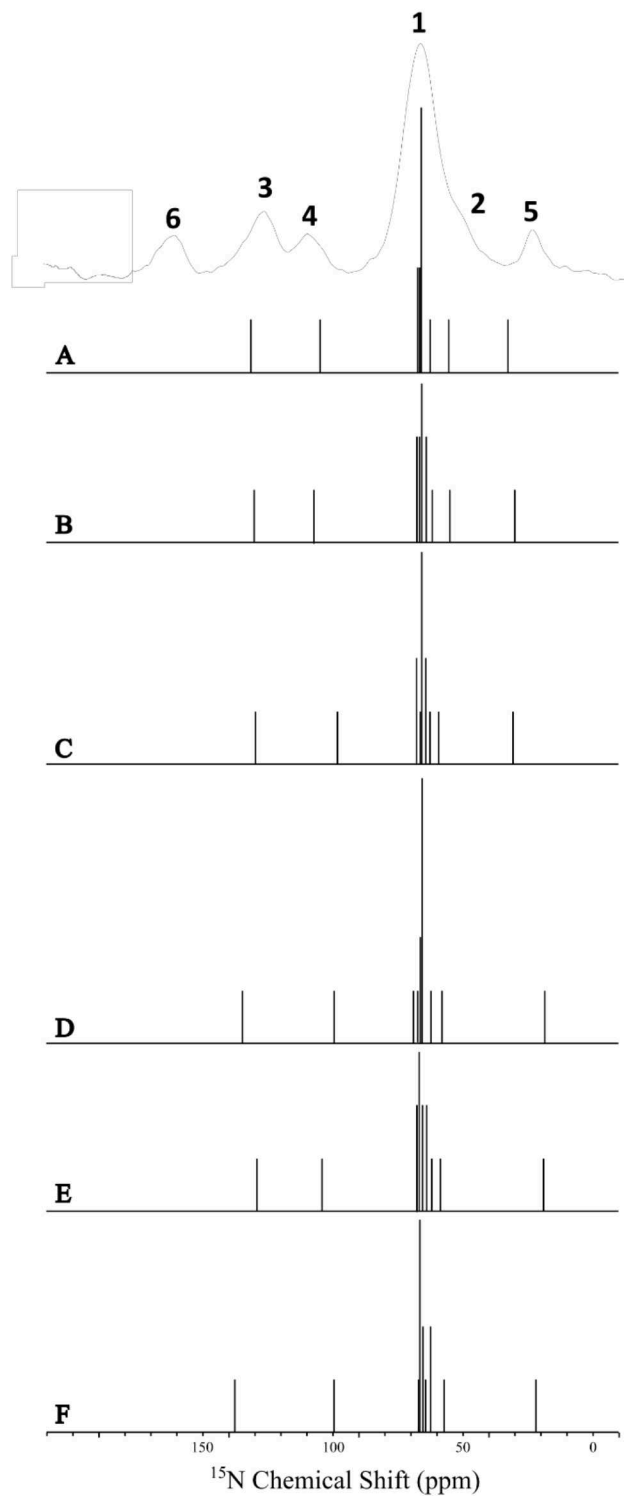
**Figure S9:** MIL-68-NH-Gly-Pro models along with the schematic representation of the -Gly-Pro grafting sites in the 6-membered channel viewed along the *c*-axis. See Figure S8 for color coding. Relative energies of the models are given in kcal mol<sup>-1</sup> relative to the most stable model **A**. Models A to C belong to conformation family **I**, while models D to F belong to conformation family **II**.

### SI-7: $^{15}\text{N}$ spectra of MIL-68-NH-Pro models



**Figure S10:** Experimental  $^{15}\text{N}$  DNP SENS NMR spectrum of MIL-68-NH-Pro from Ref.<sup>[32]</sup> along with the simulated  $^{15}\text{N}$  spectra shown as bars, for the models shown in **Figure S8**.

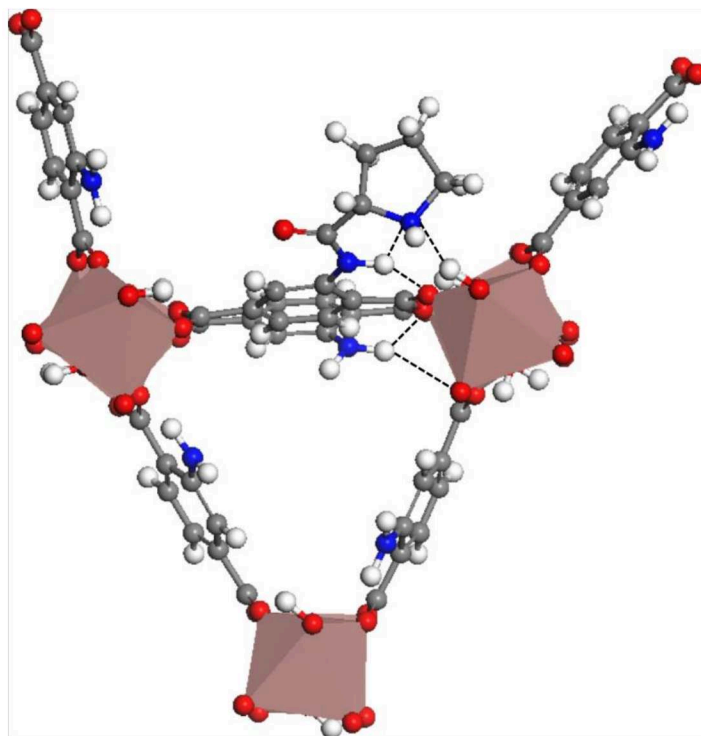
**SI-8:  $^{15}\text{N}$  spectra of MIL-68-NH-Gly-Pro models**



**Figure S11:** Experimental  $^{15}\text{N}$  DNP SENS NMR spectrum of MIL-68-NH-Gly-Pro along with the simulated  $^{15}\text{N}$  spectra shown as bars, for the models shown in **Figure S9**.

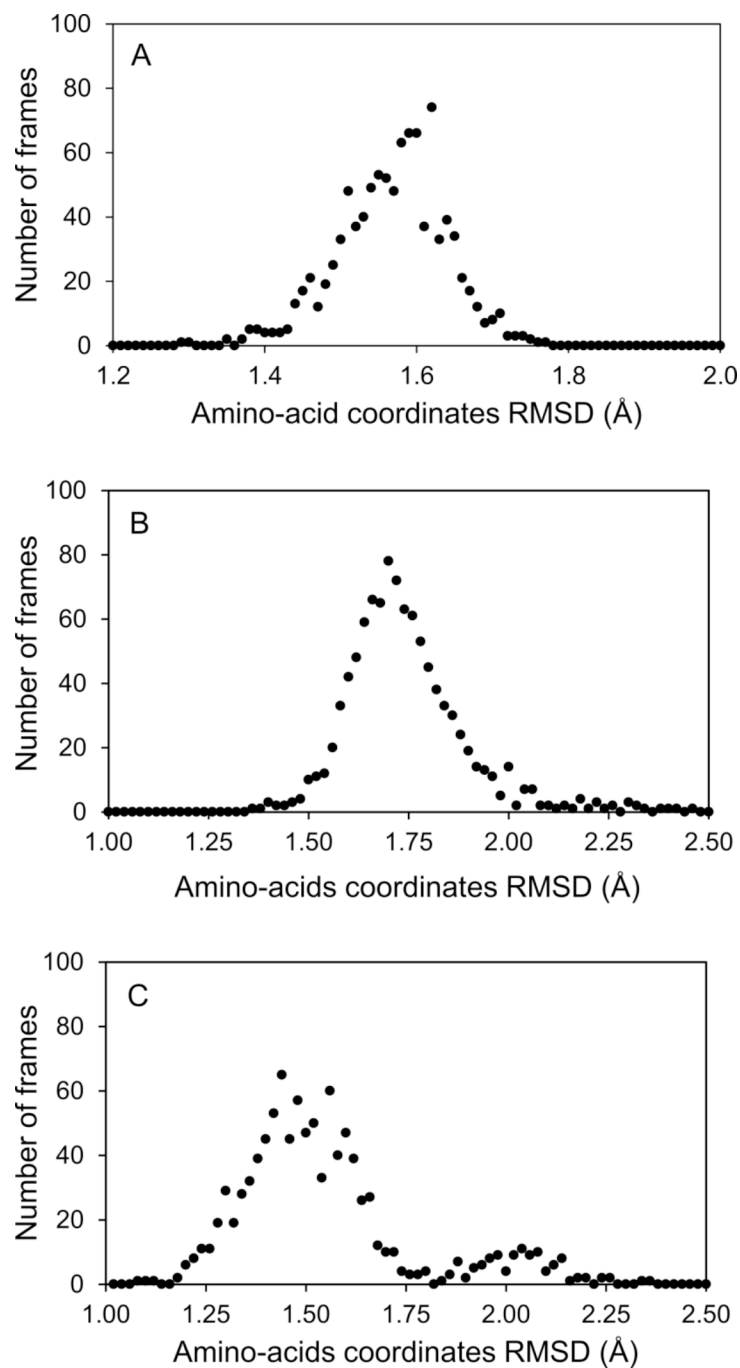


**SI-9: Grafted amino acid in the proximity of the 3-membered MOF channel**



**Figure S12:** MIL-68-NH-Pro model (Figure 2A) illustrating the  $\text{-NH}_2$  group in the vicinity of the graft being displaced towards the 3-membered channel, and as a result the H-bond formed with a carboxyl oxygen belonging to another bdc linker of this 3-membered channel.

**SI-10: RMSDs distributions for MIL-68-NH-Pro and MIL-68-NH-Gly-Pro**



**Figure S13:** Amino-acid heavy atoms coordinate RMSDs distributions over 1000 frames collected from a 5 ns trajectory with the bin step set to 0.02 Å for the most stable MIL-68-NH-Pro model (A) and MIL-68-NH-Gly-Pro conformations I and II shown in (B) and (C), respectively.

## SI-11: Experimental data for $^{15}\text{N}$ DNP SENS NMR acquisition

Experimental parameters for  $^{15}\text{N}$  CPMAS spectrum of MIL-68-NH-**Gly-Pro** (Figure 5):

	MIL-68-NH- <b>Gly-Pro</b>
MAS frequency	8 kHz
Number of scans	8192
Complex points	888
Acquisition time	14.9 ms
Recycle delay	5 s
CP contact time	1 ms
Linear Ramp	70 % to 100 %
CP power level	
$^1\text{H}^{\text{a}}$	68 kHz
$^{15}\text{N}$	40 kHz
Heteronuclear decoupling scheme	SPINAL-64 (100 kHz)
Exponential line broadening	200 Hz

<sup>a</sup> The  $^1\text{H}$  CP power is reported for 100% of the linear ramp.

Experimental parameters for  $^{15}\text{N}$  CPMAS spectra of MIL-68-NH<sub>2</sub> and MIL-68-NH-**Gly-Pro** (Figure S11):

	MIL-68-NH- <b>Gly-Pro</b> (second TEKPol2 batch)	MIL-68-NH <sub>2</sub>
MAS frequency	8 kHz	8 kHz
Number of scans	1024	256
Complex points	888	888
Acquisition time	14.9 ms	14.9 ms
Recycle delay	5 s	5 s
CP contact time	2 ms	2 ms
Linear Ramp	70 % to 100 %	70 % to 100 %
CP power level		
$^1\text{H}^{\text{a}}$	68 kHz	68 kHz
$^{15}\text{N}$	40 kHz	40 kHz
Heteronuclear decoupling scheme	SPINAL-64 (100 kHz)	SPINAL-64 (100 kHz)
Exponential line broadening	200 Hz	200 Hz

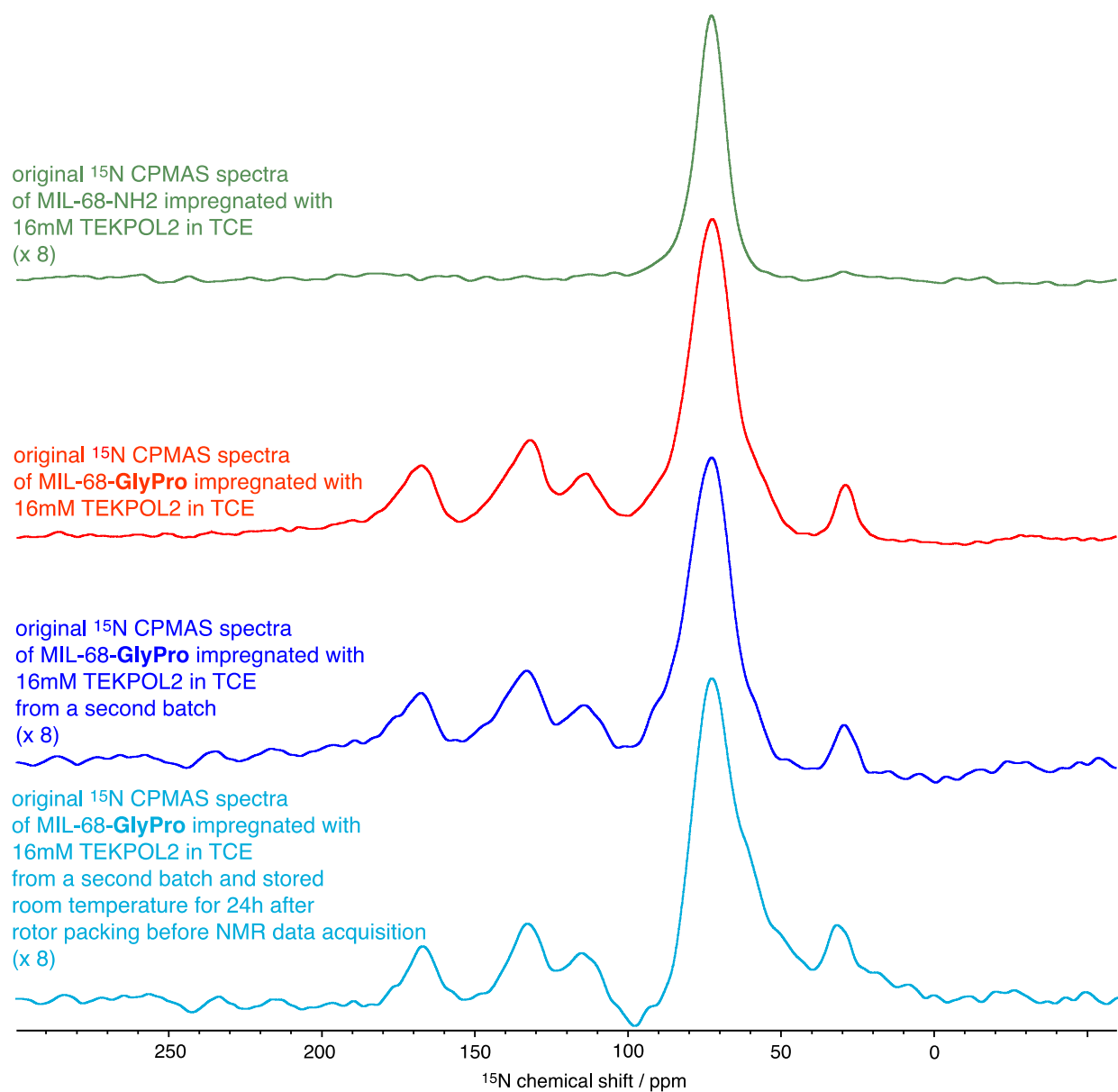
<sup>a</sup> The  $^1\text{H}$  CP power is reported for 100% of the linear ramp.

### Solvent/reactant detail:

**1,1,2,2 tetrachloroethane**, reactant grade, was purchased from Sigma-Aldrich and used to prepare TEKPOL2 solution without further purification.

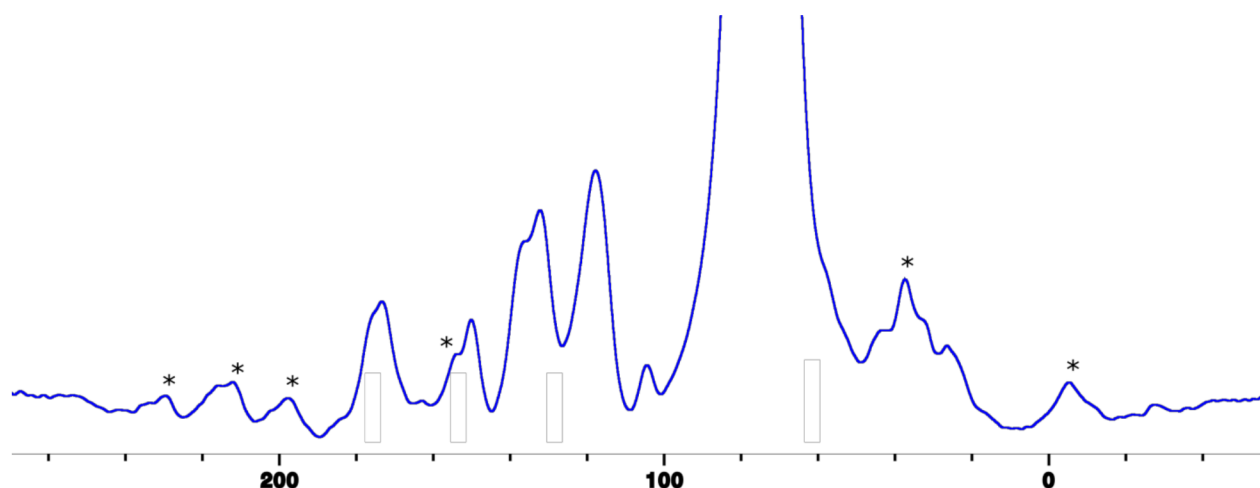
**TEKPol2** was provided by the group of Paul Tordo and Olivier Ouari (Aix-Marseille University) and prepared according to the synthesis previously reported.<sup>[33]</sup>

### <sup>15</sup>N CPMAS spectra for MIL-68-NH<sub>2</sub> and MIL-68-NH-Gly-Pro



**Figure S14:** Comparison of <sup>15</sup>N CPMAS spectra recorded for the parent material, MIL-68-NH<sub>2</sub> (green) and different MIL-68-NH-Gly-Pro samples. The peak corresponding to the highest chemical shift (161 ppm) is not present in the spectrum of MIL-68-NH<sub>2</sub> but is observed in the spectra recorded on MIL-68-NH-Gly-Pro as shown here.

The  $^{15}\text{N}$  NMR spectra simulated in this work (see also Figure S11) have no features above 140 ppm, indicating that the signal at 161 ppm is the signature of an external species. To ensure that none of the  $^{15}\text{N}$  signals observed in this spectrum can be assigned to a product potentially coming from the degradation of the polarizing agent TEKPol2 (which is a nitroxide biradical), two additional  $^{15}\text{N}$  CPMAS spectra were recorded with MIL-68-NH-**Gly-Pro** impregnated with a fresh solution of 16 mM TEKPol2 in TCE prepared from a different batch of the polarizing agent: one just after rotor packing and another one after 24 hours of storage at room temperature (the spectra are shown in Figure S14). Both spectra have a signal at 161 ppm. In contrast, a rotor of the parent material MIL-68-NH<sub>2</sub> impregnated with 16 mM TEKPol2 in TCE does not show the peak (Figure S14).



**Figure S15:**  $^{13}\text{C}$  CPMAS spectrum recorded for MIL-68-NH-**Gly-Pro**.

## References

- [1] K. Barthelet, J. Marrot, G. Férey, D. Riou, *Chem. Commun.* **2004**, 520–521.
- [2] C. Volkringer, M. Meddouri, T. Loiseau, N. Guillou, J. Marrot, G. Férey, M. Haouas, F. Taulelle, N. Audebrand, M. Latroche, *Inorg. Chem.* **2008**, *47*, 11892–11901.
- [3] M. Pera-Titus, M. Savonnet, D. Farrusseng, *J. Phys. Chem. C* **2010**, *114*, 17665–17674.
- [4] H. Sun, S. J. Mumby, J. R. Maple, A. T. Hagler, *J. Am. Chem. Soc.* **1994**, *116*, 2978–2987.
- [5] H. Sun, *J. Comput. Chem.* **1994**, *15*, 752–768.
- [6] H. Sun, *Macromolecules* **1995**, *28*, 701–712.
- [7] D. Rigby, H. Sun, B. E. Eichinger, *Polym. Int.* **1997**, *44*, 311–330.
- [8] J. Ennari, J. Hamara, F. Sundholm, *Polymer* **1997**, *38*, 3733–3744.



- [9] M. Entrialgo-Castaño, A. Lendlein, D. Hofmann, *Adv. Eng. Mater.* **2006**, *8*, 434–439.
- [10] G. Kresse, J. Furthmüller, *Comput. Mater. Sci.* **1996**, *6*, 15–50.
- [11] G. Kresse, J. Furthmüller, *Phys. Rev. B* **1996**, *54*, 11169–11186.
- [12] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [13] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- [14] P. E. Blöchl, *Phys. Rev. B* **1994**, *50*, 17953–17979.
- [15] G. Kresse, D. Joubert, *Phys. Rev. B* **1999**, *59*, 1758–1775.
- [16] *MedeA: Materials Exploration and Design Analysis*, Materials Design, Inc., **n.d.**
- [17] A. France-Lanord, P. Soukiassian, C. Glattli, E. Wimmer, *ArXiv151005416 Cond-Mat* **2015**.
- [18] R. Asahi, C. M. Freeman, P. Saxe, E. Wimmer, *Model. Simul. Mater. Sci. Eng.* **2014**, *22*, 75009.
- [19] M. Christensen, W. Wolf, C. Freeman, E. Wimmer, R. B. Adamson, L. Hallstadius, P. E. Cantonwine, E. V. Mader, *J. Nucl. Mater.* **2015**, *460*, 82–96.
- [20] B. H. Besler, K. M. Merz, P. A. Kollman, *J. Comput. Chem.* **1990**, *11*, 431–439.
- [21] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- [22] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- [23] *MedeA-Gibbs: License IFPEN-CNRS-Université Paris-Sud (2003)*, **n.d.**
- [24] M. Yiannourakou, P. Ungerer, B. Leblanc, X. Rozanska, P. Saxe, S. Vidal-Gilbert, F. Gouth, F. Montel, *Oil Gas Sci. Technol. – Rev. D'IFP Energ. Nouv.* **2013**, *68*, 977–994.
- [25] P. Ungerer, D. Rigby, B. Leblanc, M. Yiannourakou, *Mol. Simul.* **2014**, *40*, 115–122.
- [26] X. Rozanska, P. Ungerer, B. Leblanc, M. Yiannourakou, *Oil Gas Sci. Technol. – Rev. D'IFP Energ. Nouv.* **2013**, *68*, 299–307.
- [27] Q. Yang, S. Vaesen, M. Vishnuvarthan, F. Ragon, C. Serre, A. Vimont, M. Daturi, G. D. Weireld, G. Maurin, *J. Mater. Chem.* **2012**, *22*, 10210–10220.
- [28] A. Z. Panagiotopoulos, *Mol. Phys.* **1987**, *61*, 813–826.
- [29] B. Smit, S. Karaborni, J. I. Siepmann, *J. Chem. Phys.* **1995**, *102*, 2126–2140.
- [30] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- [31] L. Wu, M. Xue, S.-L. Qiu, G. Chaplais, A. Simon-Masseron, J. Patarin, *Microporous Mesoporous Mater.* **2012**, *157*, 75–81.
- [32] A. J. Rossini, A. Zagdoun, M. Lelli, J. Canivet, S. Aguado, O. Ouari, P. Tordo, M. Rosay, W. E. Maas, C. Copéret, et al., *Angew. Chem. Int. Ed.* **2012**, *51*, 123–127.
- [33] D. J. Kubicki, G. Casano, M. Schwarzwälder, S. Abel, C. Sauvée, K. Ganesan, M. Yulikov, A. J. Rossini, G. Jeschke, C. Copéret, et al., *Chem. Sci.* **2016**, *7*, 550–558.