

Elucidating the Breathing of the Metal-Organic Framework MIL-53(Sc) with *ab initio* Molecular Dynamics Simulations and *in Situ* X-Ray Powder Diffraction Experiments

Linjiang Chen,^a John P. S. Mowat,^b David Fairen-Jimenez,^c Carole A. Morrison,^d Stephen P. Thompson,^e Paul A. Wright,^b and Tina Düren^{*,a}

^a Institute for Materials and Processes, School of Engineering, The University of Edinburgh, King's Buildings, Edinburgh, EH9 3JL, United Kingdom;

^b EaStCHEM School of Chemistry, University of St. Andrews, Purdie Building, North Haugh, St. Andrews, Fife, KY16 9ST, United Kingdom;

^c Department of Chemical Engineering and Biotechnology, University of Cambridge, Pembroke Street, Cambridge, CB2 3RA, United Kingdom;

^d EaStCHEM Research School, The University of Edinburgh, King's Buildings, Edinburgh, EH9 3JJ, United Kingdom; and

^e Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, United Kingdom.

Computational Setup for DFT Calculations and AIMD Simulations

All *ab initio* molecular dynamics (AIMD) simulations and density functional theory (DFT)-based geometry optimization calculations were performed using the Becke-Lee-Yang-Parr (BLYP)^{1,2} exchange-correlation functional with semiempirical dispersion corrections to the energies and gradients from the DFT-D3³ method, unless stated otherwise. A cutoff radius of 25 Å was used for all dispersion calculations and the three-body contribution to dispersion was explicitly evaluated. A double- ζ valence plus polarization basis set, in conjunction with the relativistic, norm-conserving Goedecker-Teter-Hutter pseudopotentials optimized for the BLYP functional, was used for all elements other than Sc, which was described by the MOLOPT basis set.⁴⁻⁷ The auxiliary plane-wave basis set was defined by an energy cutoff of 350 Ry, accompanied by the relative cutoff of 50 Ry for the Gaussian basis set collocation. During each SCF cycle, the electronic structure was explicitly minimized to a tolerance of 10^{-7} Hartree. Structures were considered to be geometry-optimized when the maximum geometry change, root-mean-square geometry change, maximum force, and root-mean-square force converged to the values of 3.0×10^{-3} Bohr, 1.5×10^{-3} Bohr, 4.5×10^{-4} Hartree/Bohr, and 3.0×10^{-4} Hartree/Bohr, respectively.

The Born-Oppenheimer MD simulations were carried out in the isobaric-isothermal ensemble (NPT; constant number of particles, pressure, and temperature). The equations of motion were integrated using a time step of 0.5 fs. The temperature was controlled by a chain of Nosé-Hoover thermostats⁸ coupled to every degree of freedom (the so-called massive thermostat) with a frequency of 4000 cm^{-1} , which is high enough to sample properly the O-H bond of the hydroxyl groups in the MIL-53 solids.^{9,10} The barostat was set up with a coupling time constant of 300 fs and an external pressure of 1 bar.

Considering the QUICKSTEP module employs Γ -point only calculation, relatively large simulation boxes were used in the AIMD simulations, ensuring the shortest dimension being larger than 14 Å (i.e., a $1 \times 2 \times 2$ unit-cell representation). Alongside this unit-cell representation, a reference cell of constant volume was used to fix the number of grid points used to compute the Coulomb and exchange-correlation energies. This was used to mitigate the effect of varying grid points due to fluctuating volume of the simulation box resulting from the large breathing motion of MIL-53(Sc). It was shown previously,^{11,12} and also seen in our own investigation, that a reference cell was needed to avoid large jumps and/or discontinuities in the potential energy profile when the simulation box was not kept fixed. The use of the reference cell significantly improved the accuracy of the simulation and yielded a good conservation of energy with a drift in energy of less than 3×10^{-7}

⁵ Hartree/atom/step recorded. This reference cell treatment in the context of NPT simulations within the CP2K package was discussed in detail elsewhere.^{11,12} Note the same grid (constant grid density) was used for all the static calculations which were intended for comparing energies. That is, all energy values reported in Tables 2 and 3 in the main text were determined by single-point DFT calculations (i.e., not taken from the NPT-AIMD simulations), where the same exchange-correlation grid (constant density of grid-points) was used.

Grand-Canonical Monte Carlo (GCMC) Simulations

GCMC simulations¹³ were performed to generate CO₂ positions in the MIL-53(Sc)-DMF(*removed*) framework, which were used as the starting configurations for the AIMD simulations. The atomistic representation of the framework was constructed from the experimental crystallographic data¹⁴ with all the atoms kept fixed at their positions during the simulation. The CO₂-framework and CO₂-CO₂ interaction energies were determined according to the Lennard-Jones (LJ) and Coulomb potentials given as:

$$V_{ij} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] + \frac{z_i z_j e^2}{4\pi\epsilon_0 r_{ij}}$$

where ϵ and σ are the LJ potential well depth and diameter, respectively; r_{ij} is the distance between atoms i and j ; z_i and z_j are the fractional charges; e is the elementary charge; and ϵ_0 is the vacuum permittivity.

LJ potential parameters for all framework atoms were taken from the universal force field (UFF).¹⁵ The partial charges for the framework were calculated at the B₃LYP¹⁶/TZVP¹⁷ level of theory and were derived using the Merz-Kollman^{18,19} method with the Gaussian 09 program.²⁰ The model clusters for deriving charges are shown in Figure S1. The CO₂ molecules were modeled according to a fully flexible model, i.e., both the C-O bond and O-C-O angle can vary as described in the literature.²¹ The force-field parameters for the framework and CO₂ are presented in Table S1. The Lorentz-Berthelot combining rules were used to calculate the LJ cross-parameters. Interactions beyond 12.80 Å were neglected. Each simulation consisted of an equilibration period of 2.0×10⁷ iterations followed by a production run of a further 2.0×10⁷ iterations. The snapshot taken at the end of the production run was used as the initial configuration of CO₂ in the AIMD simulation for that particular loading.

Table S1. Force-field parameters for the MIL-53(Sc) framework¹⁴ and CO₂ atoms.²¹

atom type	σ (Å)	ϵ/k_B (K)	nonbonded potential			
H	2.571	22.142	atom type	σ (Å)	ϵ/k_B (K)	q (e)
C	3.431	52.838	O_CO ₂	3.050	80.378	-0.3256
O	3.118	30.193	C_CO ₂	2.800	28.144	0.6512
Sc	2.936	9.561	bonded potential			
			$k(\text{CO})$	1015458.633 K		
			$r_o(\text{CO})$	1.162 Å		
			$k(\text{OCO})$	54351.031 K		
			$\theta_o(\text{OCO})$	180.000 °		

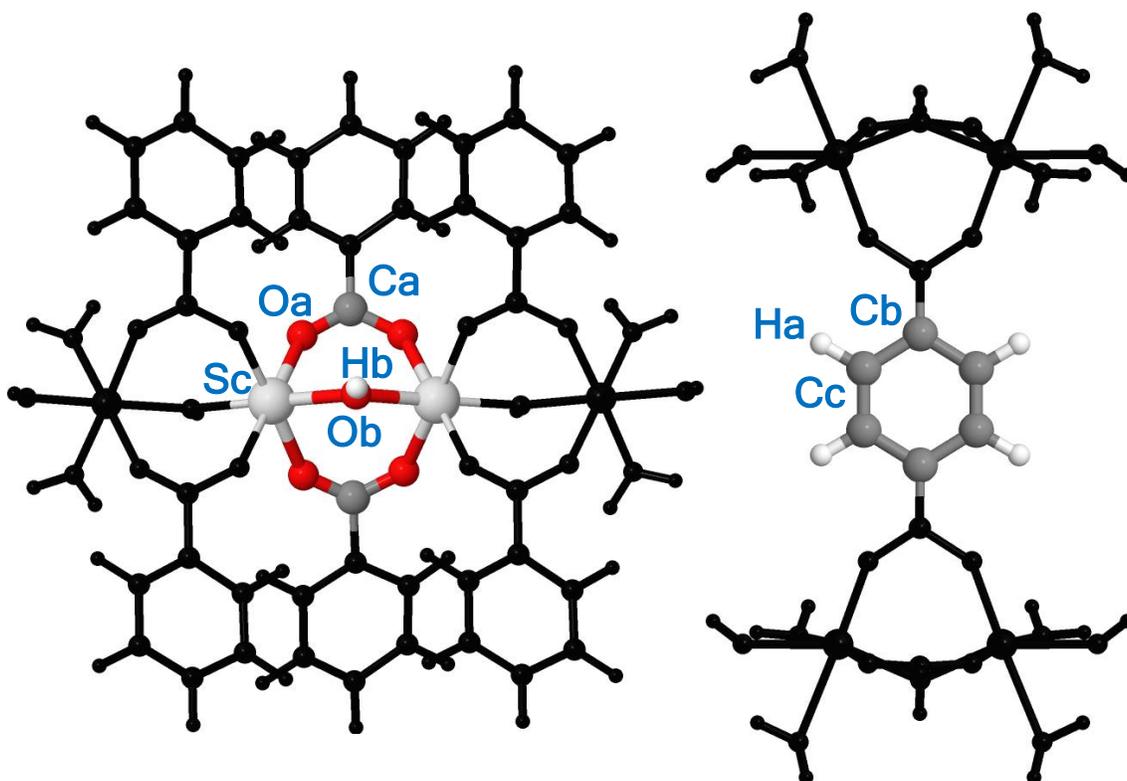


Figure S1. Clusters for calculation of partial atomic charges for the MIL-53(Sc) framework.

Table S2. Partial atomic charges for the MIL-53(Sc) framework.

atom label	Sc	Oa	Ob	Ca	Cb	Cc	Ha	Hb
charge (e)	1.8167	-0.7316	-1.0423	0.9471	-0.0511	-0.1647	0.1298	0.4996

Modeling Procedures for Generating the Various Structural Forms of MIL-53(Sc)

MIL-53(Sc)-*cp* and -*vnp*. The starting configuration of the simulation box consisted of a $1 \times 2 \times 2$ supercell representation of the experimental MIL-53(Sc)-DMF(*removed*) structure,¹⁴ corresponding to 16 $\text{ScO}_4(\text{OH})_2$ units per simulation box. In the AIMD simulations performed at 100 and 293 K, the structure was equilibrated for at least 5 ps, whereas the equilibration was extended to ca. 7 ps for 623 K. Following the equilibration, the simulation was continued for at least 5 ps for production of results at 100 and 293 K. The production run was extended to 13 ps at 623 K. The structures were considered converged to a satisfactory level when all the cell parameters (i.e., a , b , c , α , β , γ , and V) fluctuated around the mean values.

MIL-53(Sc)-*int* and -*np* with a CO_2 loading of 2.2 mmol g^{-1} (8 CO_2 per simulation box). $1 \times 2 \times 2$ unit cells of the MIL-53(Sc)-DMF(*removed*) structure were used as the starting configuration for the framework in the AIMD simulations. The initial positions of adsorbate CO_2 molecules inside the AIMD simulation box were obtained from GCMC simulations of CO_2 adsorption in the same framework performed at 196 K. In the simulation box aiming at generating the *int* topology, a total of 16 CO_2 molecules were placed initially into the pores by GCMC simulation, followed by manually removing 8 CO_2 from the box, leading to one-half of the channels being empty. For the *int* structure, the AIMD simulation was run for 10 ps as equilibration and 10 ps as production (5 + 5 ps for the *np* structure).

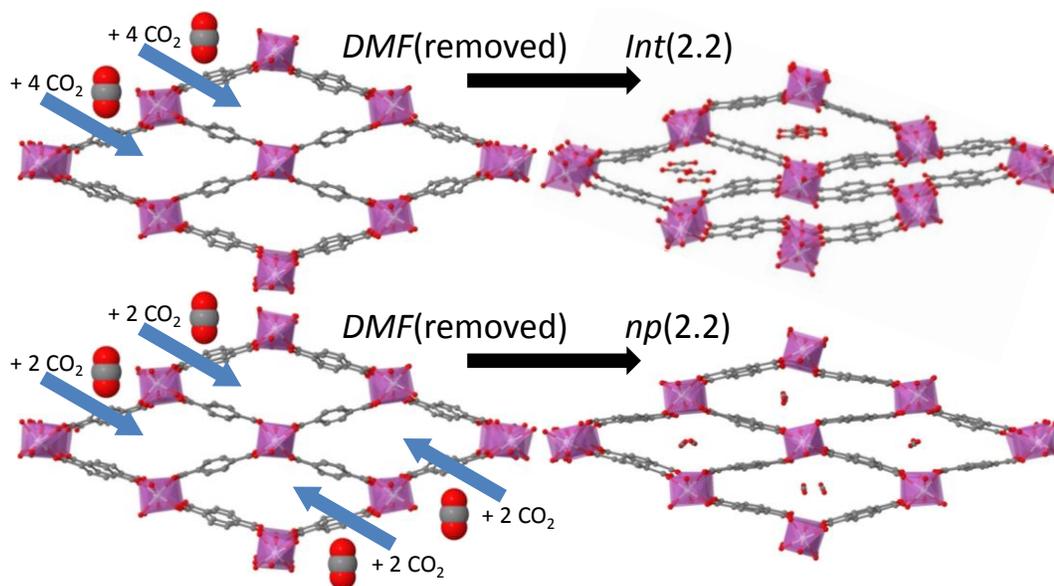


Figure S2. Set-up of simulations to investigate the structural forms of MIL-53(Sc) for a CO_2 loading of 2.2 mmol g^{-1} . The simulations correspond to the different pore-opening mechanisms for the loading as specified in the main text. Hydrogen atoms are omitted for clarity.

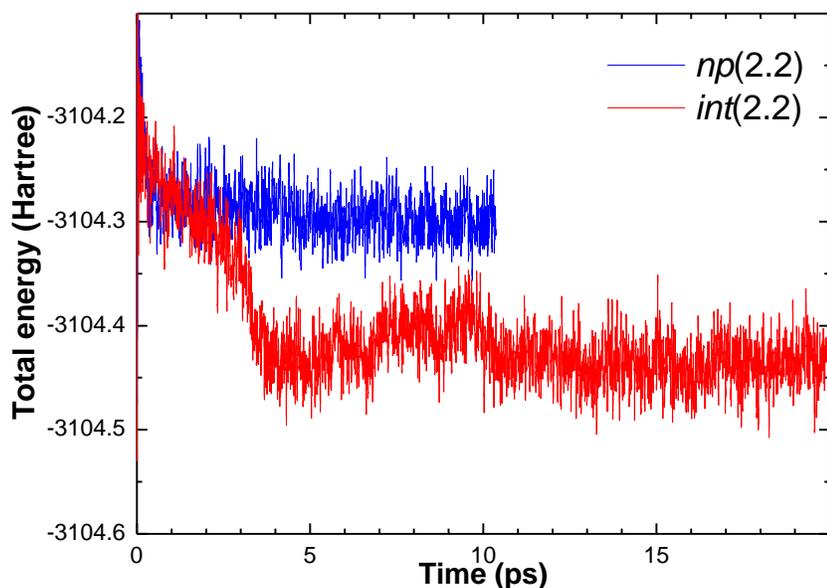


Figure S3. Time evolution of the total (= kinetic + potential) energy during the course of the AIMD simulations for the MIL-53(Sc)-*int* (red) and -*np* (blue) structures at a CO₂ loading of 2.2 mmol g⁻¹. It is clear that the *int* structure lies lower in energy than the *np* structure, indicating the former is energetically favored.

MIL-53(Sc)-*int* and -*np* with a CO₂ loading of 3.3 mmol g⁻¹ (12 CO₂ per simulation box). 4 additional CO₂ molecules were added into the open channels of the above-obtained *int* structure (with 2.2 mmol g⁻¹ CO₂) manually based on chemical intuitions (Figure S4). Since the *int* topology is preserved, i.e., half of the channels are open while the other half are closed, it is denoted as *int*(3.3) both here and in the main text. In contrast, the *np*(3.3) structure represents the situation when the additional adsorption occurs in the initially closed channels of the *int* structure (Figure S4). Two types of partially open channels are formed. One is similar to the open channels of the *int*(2.2) structure while the other is similar to the *np*(2.2) structure. This particular configuration was chosen to allow for determination of the pore-opening energy for the *int*(2.2) → *np*(3.3) structural transformation. For both the *int*(3.3) and *np*(3.3) structures, the simulations were run for 7 ps and the last configurations were used for calculations.

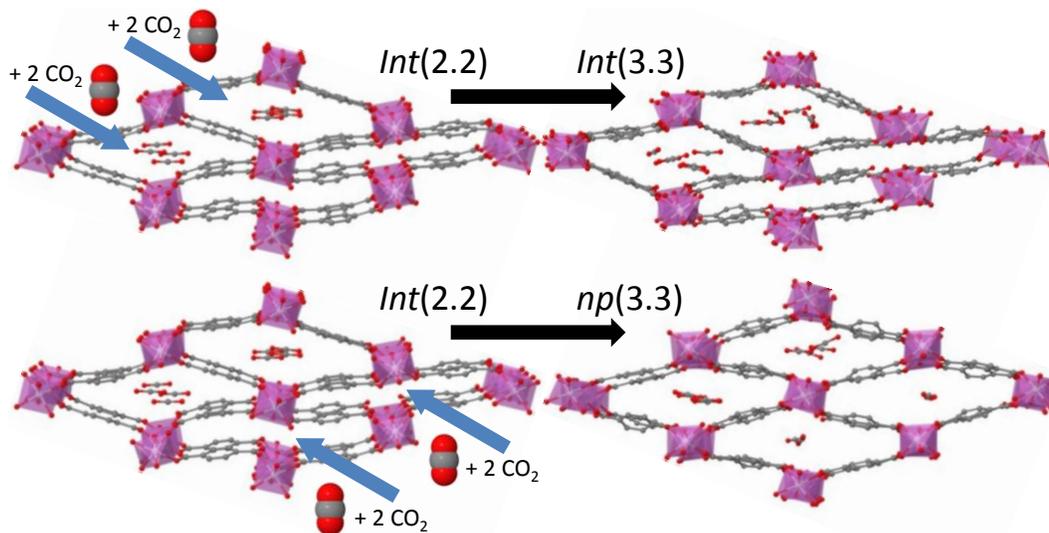


Figure S4. Set-up of simulations to investigate the structural transformations of MIL-53(Sc) from a CO₂ loading of 2.2 mmol g⁻¹ to with a CO₂ loading of 3.3 mmol g⁻¹. Hydrogen atoms are omitted for clarity.

MIL-53(Sc)-*lp*. Since the maximum number of CO₂ molecules that could be fitted in the MIL-53(Sc)-*DMF(removed)* structure using GCMC simulation was only 7 CO₂ per unit cell, higher loadings were realized by progressively adding CO₂ molecules into the framework expanded by AIMD simulations at the current loading, virtually analogous to the adsorption experiment. Note that these supplementary structures (i.e., MIL-53(Sc) with a loading of 8, 10, 12, or 14 CO₂ per unit cell of the *DMF(removed)* structure) were not necessarily equilibrated before the new CO₂ molecules were inserted into the framework manually based on chemical intuitions. The structure with a loading of 16 CO₂ per unit cell of the *DMF(removed)* structure, denoted MIL-53(Sc)-*lp*, was equilibrated for 5 ps, followed by the production run of another 5 ps.

Note that each of the AIMD structures reported in the main text was generated by first averaging the corresponding AIMD trajectory of the production run. The thus-obtained time-averaged structure was then energy minimized at 0 K where the atomic coordinates were fully optimized with respect to the time-averaged cell parameters.

Calculation of Framework Potential Energy

The potential energy for all the empty framework configurations (i.e., the CO₂ molecules were removed from the *int*, *np*, and *lp* structures) was calculated at the BLYP-D₃ level of theory. Note that these calculations were performed on the energy-minimized, time-averaged structures for the *cp*, *vnp*, *int*, and *lp* forms, whereas the energy-minimized, experimental structure was used for determining the framework potential for MIL-53(Sc)-DMF(*removed*).

Calculation of Sorption-Induced Pore-Opening Energy

$$\Delta E = E(\text{MIL} \cdots n_2 \text{CO}_2) - E(\text{MIL} \cdots n_1 \text{CO}_2) - (n_2 - n_1)E(\text{CO}_2) \quad (\text{S1})$$

where $E(\text{MIL} \cdots n_2 \text{CO}_2)$ and $E(\text{MIL} \cdots n_1 \text{CO}_2)$ are the total energies of the MIL-53(Sc) framework with a loading of n_2 and n_1 CO₂ ($n_2 > n_1$), respectively; and $E(\text{CO}_2)$ is the energy of an isolated, geometry-optimized CO₂ molecule in a supercell with the same dimensions as the framework loaded with n_2 CO₂.

Analysis of MD Statistics

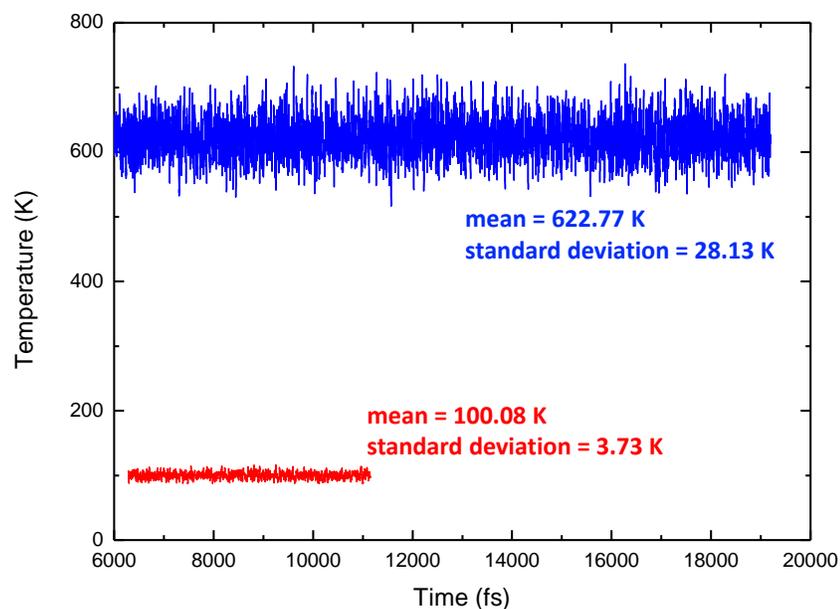


Figure S5. Time evolution of the temperature over the course of AIMD simulations performed at 100 (red) and 623 (blue) K. The average temperatures and standard deviations indicate good MD statistics.

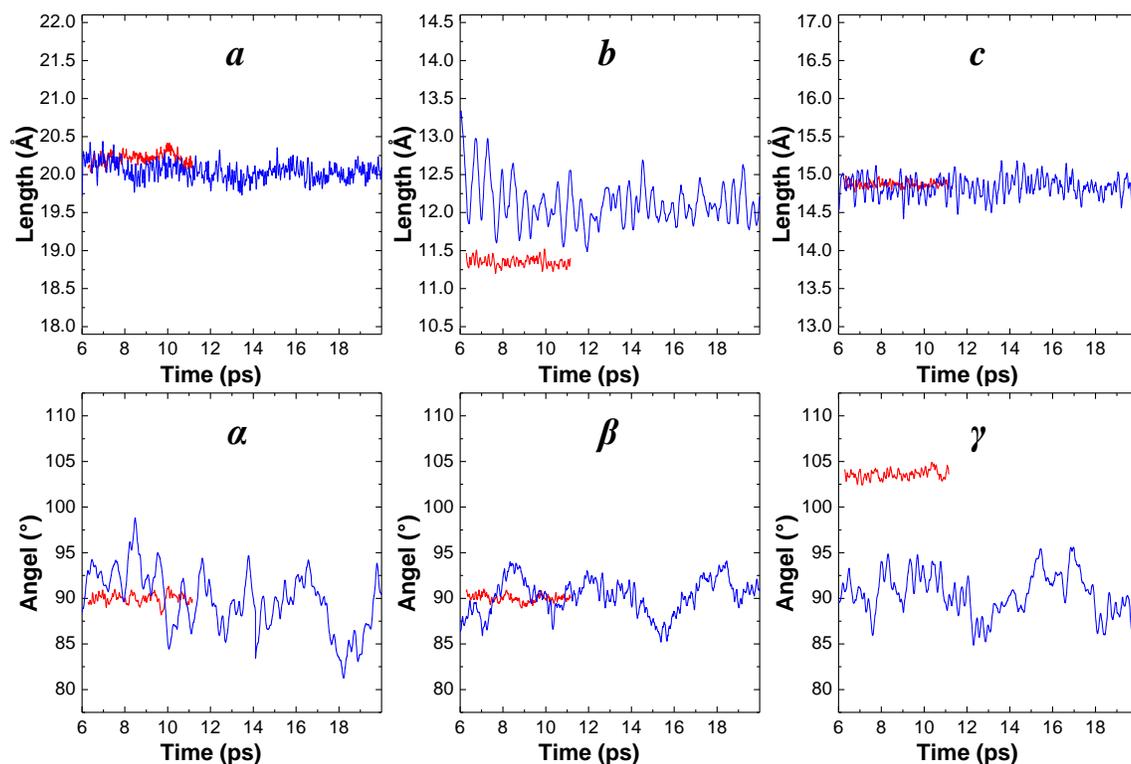


Figure S6. Time evolution of the cell parameters during the course of AIMD simulations performed at 100 (red) and 623 (blue) K. Both structures were considered equilibrated as the parameters fluctuated around mean values. It is also clear that the higher the temperature, the larger the amplitudes of fluctuations, in accordance with basic thermodynamic principles.

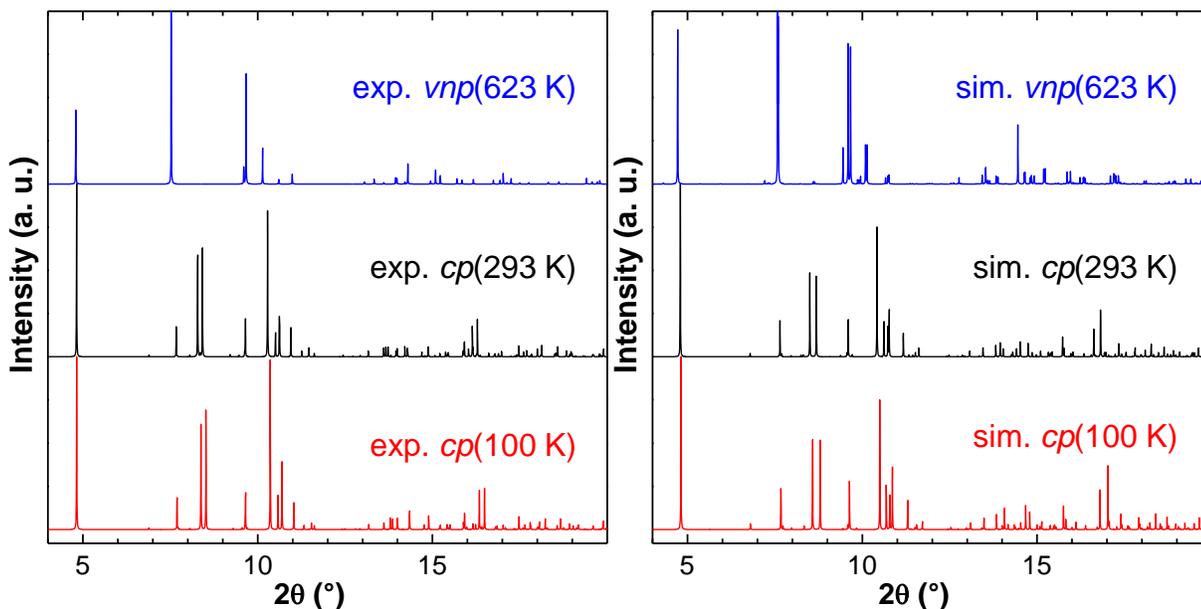


Figure S7. Comparison of the X-ray powder diffraction patterns of dehydrated MIL-53(Sc) at 100, 293, and 623 K between the AIMD-simulated and experimental structures.

Importance of Dispersion Interactions

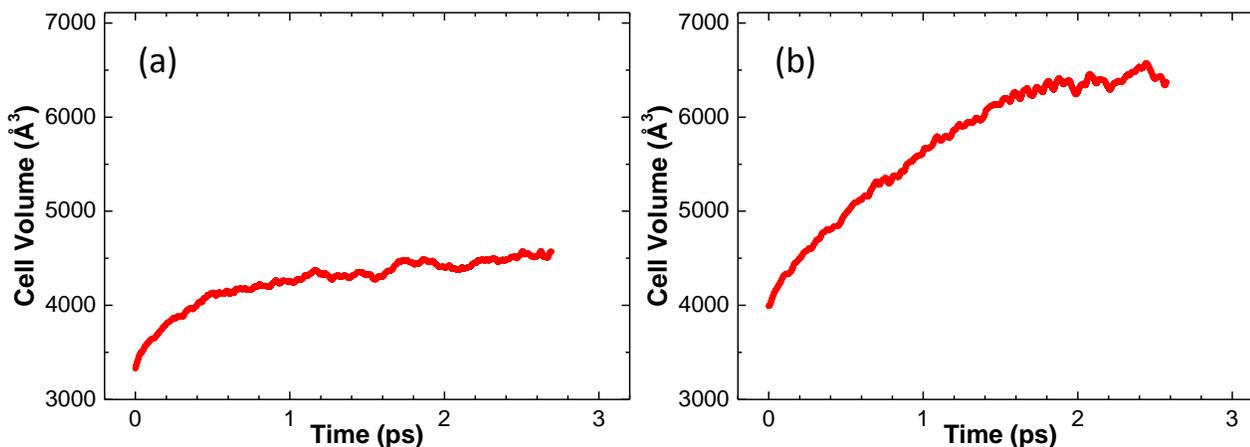
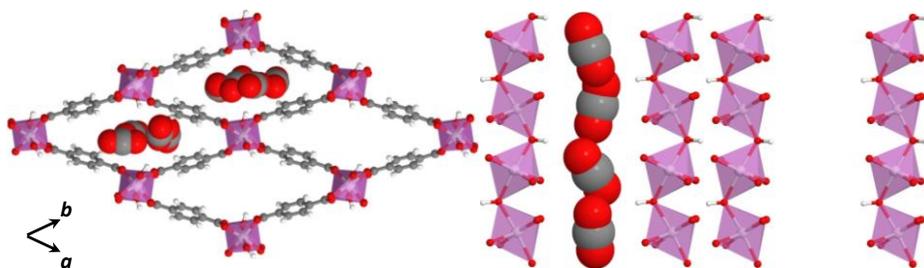


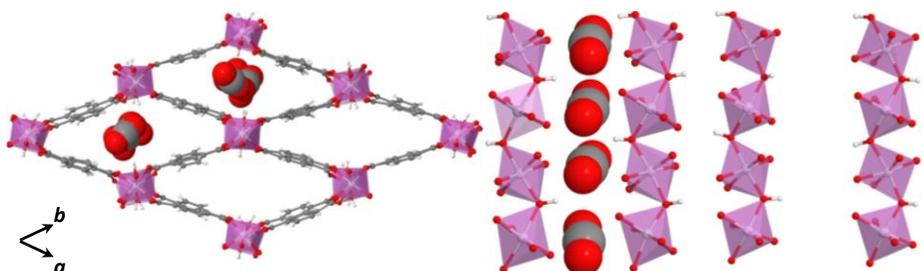
Figure S8. Time evolution of the cell volume (per 16 $\text{ScO}_4(\text{OH})_2$ units) during the AIMD simulations performed employing the BLYP functional, i.e., dispersion corrections were not applied. (a) The simulation was performed at 100 K with starting configuration being the MIL-53(Sc)-*cp* structure. (b) The simulation was performed at 196 K with starting configuration being the MIL-53(Sc)-*int* structure (with 2.2 mmol g^{-1} CO_2 loading). It is clear that explicit inclusion of dispersion corrections in the DFT calculations was required to predict correctly the *cp* and *int* structures.

Evolution of MIL-53(Sc)-*int* during the Course of the AIMD Simulation

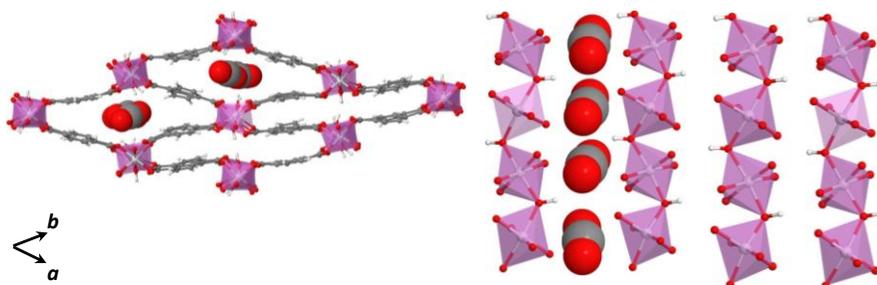
$t = 0$; $a = 21.412 \text{ \AA}$, $b = 21.412 \text{ \AA}$, $c = 14.603 \text{ \AA}$, $\alpha = 90.000^\circ$, $\beta = 90.000^\circ$, $\gamma = 50.776^\circ$, $V = 5186.34 \text{ \AA}^3$



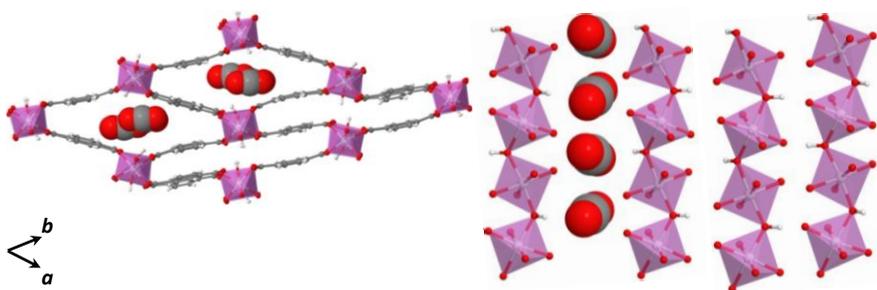
$t = 1 \text{ ps}$; $a = 21.691 \text{ \AA}$, $b = 21.788 \text{ \AA}$, $c = 14.756 \text{ \AA}$, $\alpha = 90.117^\circ$, $\beta = 89.946^\circ$, $\gamma = 50.141^\circ$, $V = 5353.36 \text{ \AA}^3$



$t = 5 \text{ ps}$; $a = 20.799 \text{ \AA}$, $b = 21.390 \text{ \AA}$, $c = 14.923 \text{ \AA}$, $\alpha = 87.883^\circ$, $\beta = 87.122^\circ$, $\gamma = 37.493^\circ$, $V = 4035.834 \text{ \AA}^3$



$t = 10 \text{ ps}$; $a = 20.589 \text{ \AA}$, $b = 21.384 \text{ \AA}$, $c = 14.897 \text{ \AA}$, $\alpha = 90.930^\circ$, $\beta = 87.176^\circ$, $\gamma = 39.010^\circ$, $V = 4107.99 \text{ \AA}^3$



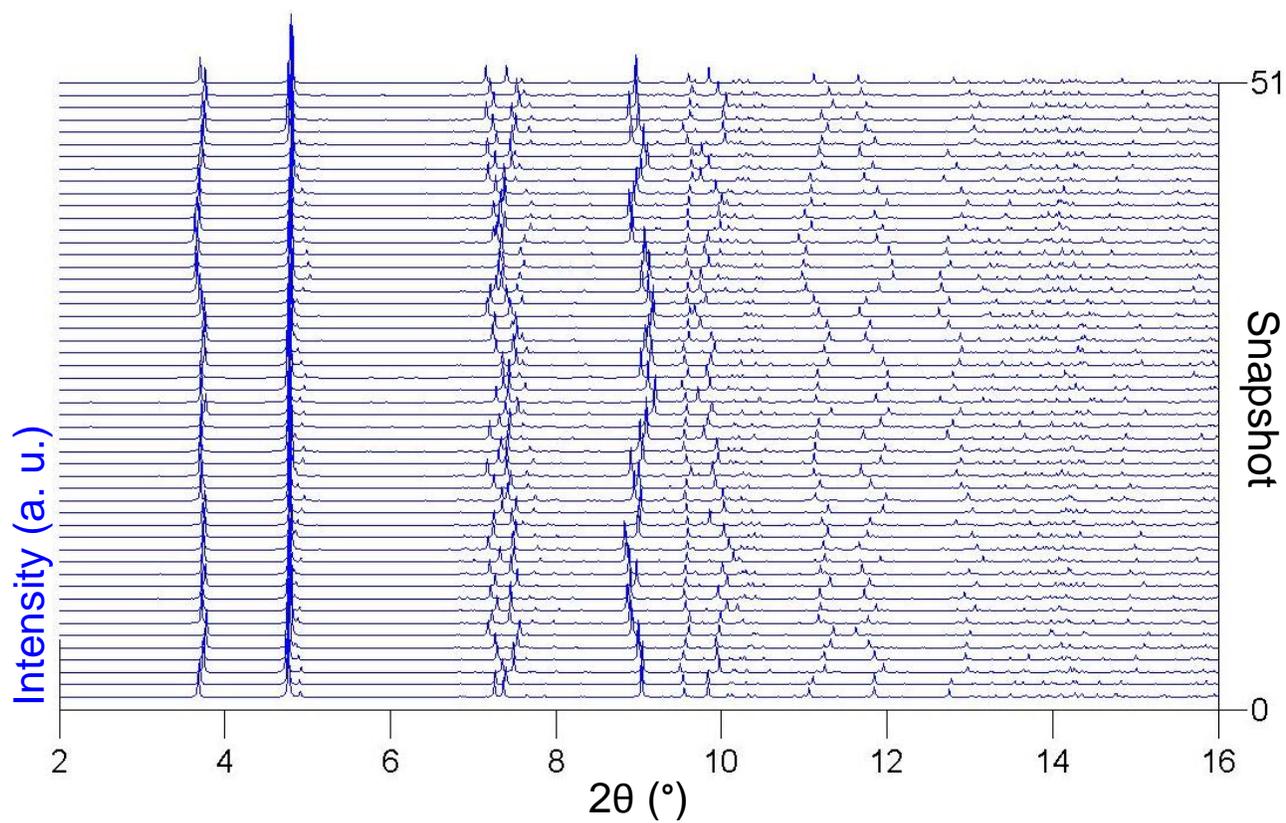


Figure S9. A series of simulated patterns based on the snapshots (with an interval of 0.2 ps) taken during the AIMD simulation of the MIL-53(Sc)-*int* structure. These patterns were averaged to obtain the averaged pattern shown in Figure 5b in the main text.

AIMD-Simulated MIL-53(Sc)-*np* Structure for a CO₂ Loading of 2.2 mmol g⁻¹

In contrast to the MIL-53(Sc)-*int* structure (Figure S10a) for the same loading, the CO₂ molecules are uniformly distributed among all the channels in the *np* form (Figure S10b), leading to the partial opening of all pores. This structure was previously observed for some other MIL-53 solids.^{9,22}

It should be pointed out that the MIL-53(Sc)-*int* and *np* structures correspond to the different hypotheses (specified in the main text) on the pore-opening mechanism for the same CO₂ uptake. Comparing the X-ray powder diffraction (XRPD) patterns calculated on the two simulated structures with the experiment (Figure S10c), it is clear that the solid adopted the *int* form and not the *np* form at the first plateau of the CO₂ adsorption isotherm (2 – 3 mmol g⁻¹). In light of this finding, both structures were further scrutinized on the molecular level in order to compare the different conformations of the CO₂ inside the two frameworks and their influences on the resulting adsorption mechanisms at work. The simulation results summarized in Figure S11 and Table S3 indicate that the two structural forms exhibited remarkably different arrangements of the CO₂ molecules located in the open channels. The different natures of the CO₂-CO₂ interactions consequently influenced the nature of the CO₂-framework interactions. In particular, the simulations suggested that the electron donor-acceptor complexes formed between the C(CO₂)•••O(μ₂-OH) in the *int* structure only can be realized in the presence of the strong CO₂-CO₂ interactions (Table S3).

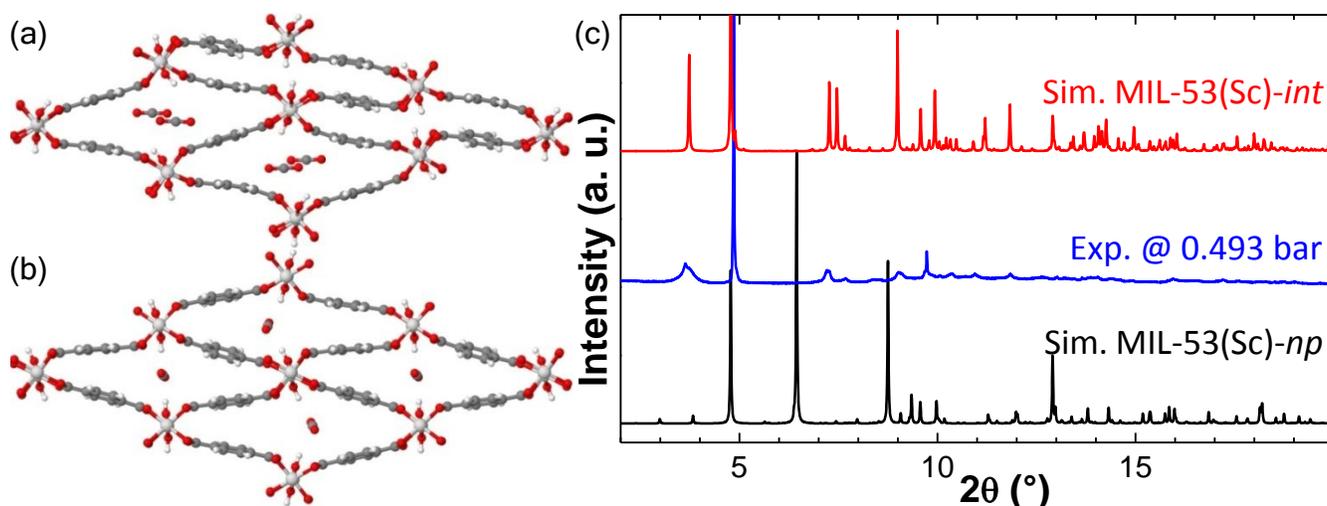


Figure S10. Two possible structural forms for MIL-53(Sc) at a CO₂ loading of 2.2 mmol g⁻¹ generated by AIMD simulations: MIL-53(Sc)-*int*, (a); and MIL-53(Sc)-*np*, (b). Comparison of the calculated XRPD patterns based on the *int* and *np* structures with the experimental one corresponding to the CO₂ uptake of ca. 2 mmol g⁻¹ is given in (c).

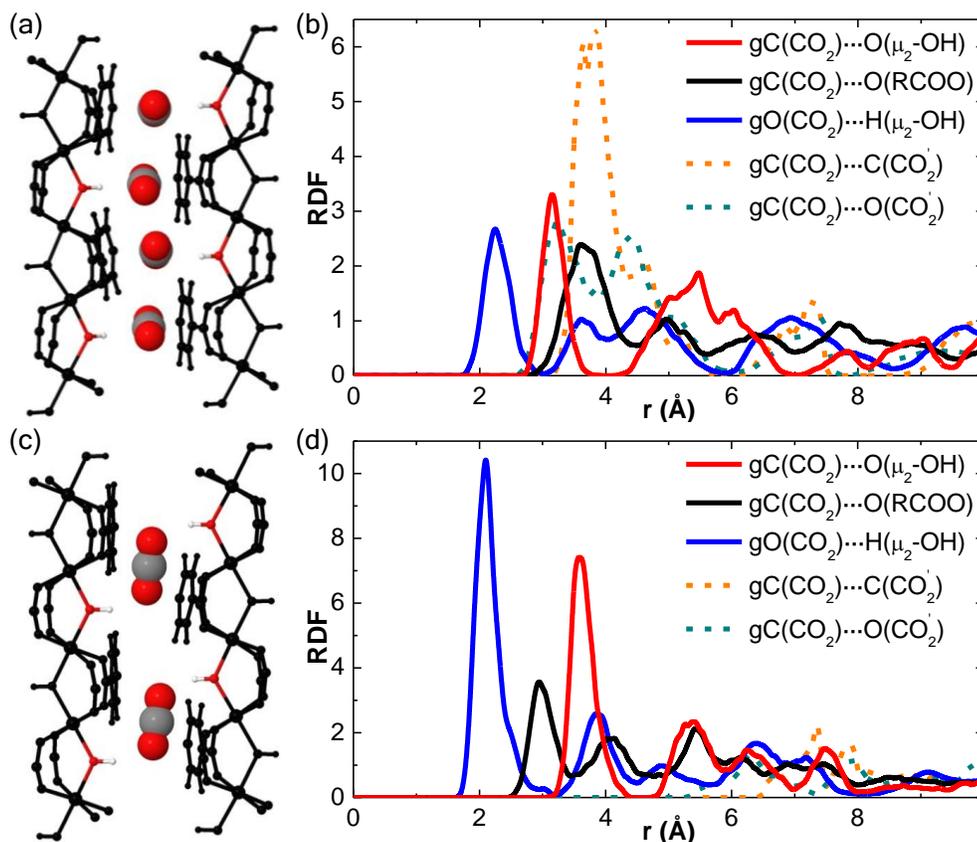


Figure S11. View of the partially open channels in the *int* (a) and *np* (c) structures obtained from the AIMD simulations. The hydroxyl groups and CO₂ molecules in the central channel are color coded: white, hydrogen; black, carbon; and red, oxygen. The corresponding radial distribution functions were computed from the AIMD-NPT simulations: (b) and (d) for the *int* and *np* structures, respectively.

Table S3. Characteristic distance (in Å) for the selected atom pairs observed in the MIL-53(Sc)-*int* and -*np* structures.

	MIL-53(Sc)- <i>int</i>	MIL-53(Sc)- <i>np</i>
C(CO ₂)...O(μ ₂ -OH)	3.15	3.60
C(CO ₂)...O(RCOO)	3.55	3.00
O(CO ₂)...H(μ ₂ -OH)	2.25	2.10
C(CO ₂)...C(CO ₂ ')	3.65	7.35
C(CO ₂)...O(CO ₂ ')	3.30	6.30

MIL-53-*lp*

Indexing the high-pressure phase using the WinPlotr program in the fullprof suite and refinement using the Le Bail method yielded an orthorhombic unit cell with parameters of $a = 7.310(1)$ Å, $b = 17.029(2)$ Å, and $c = 13.504(1)$ Å. Comparison of the indexed cell parameters with the literature indicated good agreement with the orthorhombic *Imma* cell observed for the *large-pore* MIL-53(Cr)-ht and MIL-53(Al)-ht.^{9,23} The structure MIL-53-open with the *Imma* space group was used as the initial model for building MIL-53(Sc)-*lp* and the cell parameters of the new model were modified using the experimental Le Bail refined ones. Thereafter, the structures were subject to energy minimization based on molecular mechanics calculations as implemented in Materials Studio,²⁴ modifying all the atomic positions. The bonded and the short-range (van der Waals) nonbonded interactions between the atoms were modeled using the LJ potential with the UFF, whereas bond stretching was described by a harmonic term, angle bending by a three-term Fourier cosine expansion and torsions and inversions by cosine-Fourier expansion terms. A cutoff distance of 12 Å was used for the LJ interactions. The long-range, electrostatic, interactions, arising from the presence of partial atomic charges, were modeled using a Coulombic term. The Ewald summation method was used to compute the electrostatic interactions. Partial atomic charges were derived from the charge equilibration method (Q_{Eq}).²⁵ To identify the positions of the adsorbed CO₂, we used GCMC simulations on the MIL-53(Sc)-*lp* model identified above. Four different models were obtained:

(i) Model *A* was obtained after GCMC equilibration at high pressure to allow the saturation of the pores with CO₂ (i.e., 14 molecules per unit cell).

(ii) Models *B* and *C* were obtained after the insertion of a number of CO₂ molecules based on the calculated uptake, 12 and 14 molecules per unit cell respectively, and using the *Imma* symmetry. The models were then energy minimized using the approach used to obtain the MIL-53(Sc)-*lp* model.

(iii) Model *D* is similar to model *C*, but the optimization of the CO₂ molecules localization using the energy minimization procedure was performed with symmetry reduced to *P* 1.

Comparison of the patterns with experimental data showed good agreement with the model optimized in *P* 1 (model *D*) with the calculated number of CO₂ molecules (12 per unit cell). To obtain the final model, the calculated positions for the adsorbed CO₂ molecules were included in the *Imma* framework model (with the occupancy of each site reduced to 0.125 to offset the site

multiplicity). The framework was refined with distance constraints applied to the Sc-O: 2.095(5) Å; O-C: 1.275(5) Å; and C-C: 1.45(2) Å and 1.39(2) Å for the ring-carboxylate and the aromatic bonds, respectively. Also constrained were the through space O-O: 2.95(5) Å distances to maintain the octahedral geometry around the scandium and C-C across the phenyl ring 2.78(5) Å to maintain the rigid nature of the terephthalate linker. The refinement of MIL-53(Sc)-*lp* against the experimental data converged with $R_{wp} = 0.0805$, $R_p = 0.0543$ in an orthorhombic unit cell with *Imma* symmetry and unit cell $a = 7.31067(17)$ Å, $b = 17.0297(4)$ Å, $c = 13.5068(5)$ Å with a total of 13 constraints applied to the framework as described above, chemical formula $\text{Sc}(\text{OH})(\text{BDC}) \cdot 3.5\text{CO}_2$.

$\lambda = 0.825072$ Å, L-S cycle 2064

Obsd. and Diff. Profiles

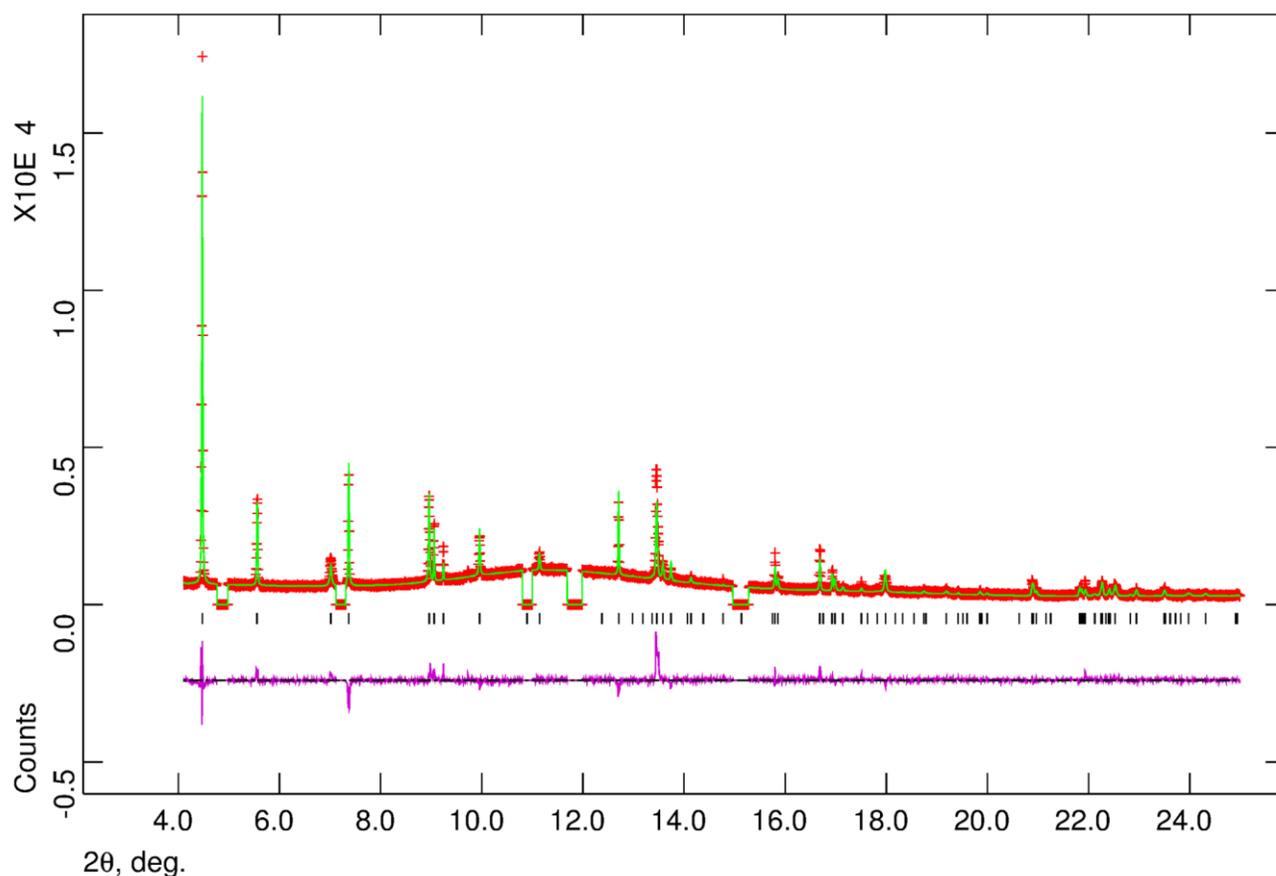


Figure S12. Rietveld profile fit to synchrotron powder X-ray diffraction data for MIL-53(Sc)-*lp*. Excluded regions are residual MIL-53(Sc)-*int* phase present in all experimental patterns of MIL-53(Sc)-*lp* which could not be fitted as a two-phase refinement (due to the complexity of the structure and the observed broadening effects). Experimental data, red markers; fitted profile green; difference plot purple.

Table S4. Selected bond distances and angles of the *cp* (100 and 293 K) and *vnp* (623 K) forms of MIL-53(Sc) from AIMD simulations and experiments.²⁶

MIL-53(Sc)- <i>i</i>	Sc-O(μ_2 -OH) (\AA) ^a		Sc-O(RCOO) (\AA) ^a		\angle Sc-(μ_2 -O)-Sc ($^\circ$) ^b	
	sim.	exp.	sim.	exp.	sim.	exp.
<i>cp</i> (100 K)	2.105	2.090	2.134	2.093	124.4	122.8
<i>cp</i> (293 K)	2.108	2.109	2.142	2.105	124.2	120.8
<i>vnp</i> (623 K)	2.099	2.068	2.125	2.075	123.5	123.2

^a The differences between simulated and experimental bond distances for Sc-O(μ_2 -OH) and Sc-O(RCOO) are smaller than 1.5% and 2.4%, respectively, for all three temperatures. ^b The differences between simulated and experimental Sc-(μ_2 -O)-Sc angles are smaller than 2.8% for all three temperatures. These very small discrepancies between simulation and experiment demonstrate the very good performance of our AIMD simulation in describing the chemistry involved in the MIL-53(Sc) solid responding to temperature variation. The very good reproduction of the geometry of the chain of μ_2 -OH corner-sharing $\text{ScO}_4(\text{OH})_2$ octahedra demonstrates that the BLYP-D₃ scheme, together with the computational setup used here, can treat the metal-oxygen interactions with a high level of accuracy.

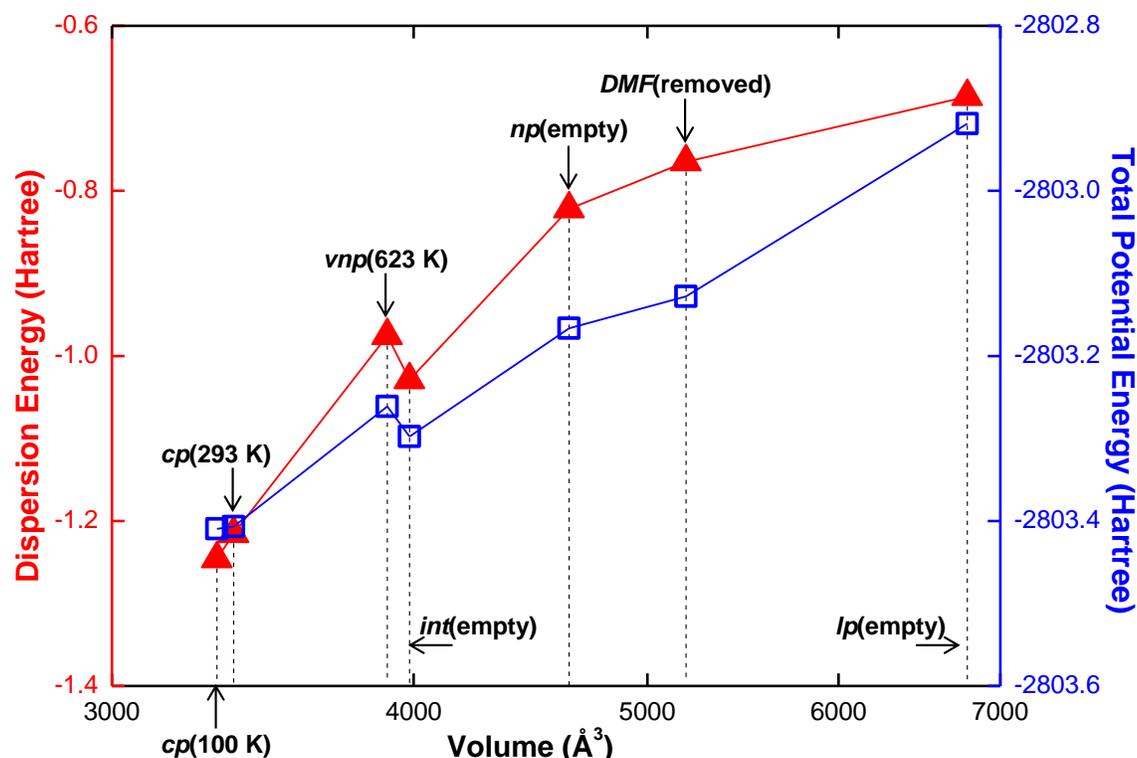


Figure S13. Total potential energies, together with the corresponding dispersion contributions according to the DFT-D₃ correction, are plotted as a function of the framework volume for all empty MIL-53(Sc) structures generated by the AIMD simulations. Both the energies and the volumes are per framework consisting of 8 $\text{ScO}_4(\text{OH})_2$ units, and the energies were calculated using the same exchange-correlation grid (constant density of grid-points). It is clearly shown that as the MIL-53(Sc) framework expands the structure becomes increasingly energetically unfavorable, which is mainly due to the decrease of dispersion in the increasingly less dense framework. Note that the volume is in logarithmic scale and the two vertical axes have the same magnitude (i.e., 0.8 Hartree), while the lines were added to guide the eye.

Further Validation of the AIMD Simulations in the NPT Ensemble

The AIMD simulation-based modeling scheme demonstrated in this contribution can be considered a valid tool for studying the breathing phenomena of the MIL-53(Sc) solid, on the basis of its accurate reproductions of the experimental *cp* and *vnp* phases induced by temperature variation, and its successful predictions of the *int* and *lp* phases due to CO₂ adsorption, confirmed by the XRPD experiments in this work. The methodology is thus successful with regard to the modeling objectives of this work, which are to assist structure determinations of the different phases of MIL-53(Sc) and to elucidate the structural transitions on the molecular level. In addition to validating the modeling scheme applied here by means of comparing simulation to experiment, we further present detailed analyses of the thermodynamic data obtained from our AIMD simulations in the isobaric-isothermal (NPT) ensemble to justify our choice of Gaussian basis set and charge-density cutoff.

It has been pointed out in the literature that a large basis set, in conjunction with a high cutoff, may be needed to achieve accurate pressure/stress computations for simulations in ensembles with fluctuating volume.^{11,12} Using NPT-Monte Carlo simulations, McGrath et al. investigated the effect of charge-density cutoff on the energy-volume (U - V) curve for liquid water at ambient conditions.¹¹ It was found that a progressively increased cutoff (280 – 1200 Ry), together with a TZV2P basis set, systematically decreased the magnitude of the jumps in the U - V curves. The study highlights the importance of choosing a basis set and a charge-density cutoff to ensure the accuracy in computation of internal pressure for isobaric simulations.

Herein, the large system sizes – ranging from 304 to 496 atoms per simulation box – entail the need for using a relatively small basis set and charge-density cutoff, in order to allow for tractable simulations with long enough time frames to develop system dynamics. Nevertheless, the U - V curve obtained from our NPT-AIMD simulation (Figure S14) does not exhibit any significant jumps/discontinuities for a volume change as large as 36% (larger than the 24% volume change studied by McGrath et al.),¹¹ although we used a DZVP basis set and a cutoff of 350 Ry. This may be surprising at first sight if compared to the findings made by McGrath et al.¹¹ We, however, tend to attribute this good behavior of our NPT simulations to the fact that our system is a solid-state framework rather than a liquid system as studied by McGrath et al. Indeed, for framework systems a DZVP basis set has been shown to allow simulations based upon it to describe accurately a variety of properties,^{27,28} and is a good compromise between quality of basis set and speed of computation. Moreover, it is generally known that the choice of simulation parameters is often

system-specific. For example, Siepmann and co-workers were able to study the ionic-liquid systems using NPT-AIMD simulations with a DZVP basis set and a cutoff of 300 Ry.^{29,30}

As a further step to justify the validity of the DZVP basis set and 350-Ry cutoff for our specific system, we assessed the effect of basis-set size and cutoff value on our simulation cell. We performed a sensitivity test where three NPT-AIMD simulations were carried out in parallel, all starting with the same system configuration. Tests 1, 2, and 3 were run using (DZVP & 350 Ry), (TZV2P & 350 Ry), and (DZVP & 700 Ry), respectively, while all other simulation parameters were the same. For all three simulations, the cell parameters were averaged over 1 ps and are summarized in Table S5. Note that an intended test with (TZV2P & 700 Ry) setup was proven intractable given our computing resources. It can be clearly seen from Table S5 that the three simulations with the different combinations of basis set and cutoff yielded mostly identical results, thus demonstrating that the (DZVP & 350 Ry) setup, which gave the fastest computation, is a valid and sensible choice for NPT-AIMD simulation of our system. It, however, should be noted that the basis set/cutoff-dependence of our system in long term cannot be inferred from this comparison, since these simulations only ran for 1 ps. Whereas a more rigorous investigation is worth conducting, a comparative, theoretical study is beyond the scope of this work.

Finally, as noted by Schmidt et al.,¹² the conserved quantity is an important quality-indicator for MD simulations in ensembles other than NVE (where the total energy is conserved). All our NPT-AIMD simulations achieved a very good conservation of energy with almost negligible drift, which can be seen from Figure S15, for example. Furthermore, Figure S16 and Table S6 show the good pressure behaviors of our NPT simulations. The cumulative average pressure converged to the externally applied pressure reasonably well with only small fluctuations in an order of 10's of bars. The amplitude of the fluctuations of the instantaneous pressure is considerably larger, which, however, is known to be mainly a result of the finite system size.¹²

To conclude, we have validated our choice of the (DZVP & 350 Ry) combination for the NPT-AIMD simulations of MIL-53(Sc) in the context of the several technical issues about isobaric simulations raised in the literature. We have shown our computational setup – particularly concerning basis set and charge-density cutoff – yielded a very good thermodynamic behavior of the NPT-AIMD simulations performed on a framework system. The applicability of the NPT-AIMD methodology to studying the MIL-53(Sc) solid thus can be affirmed.

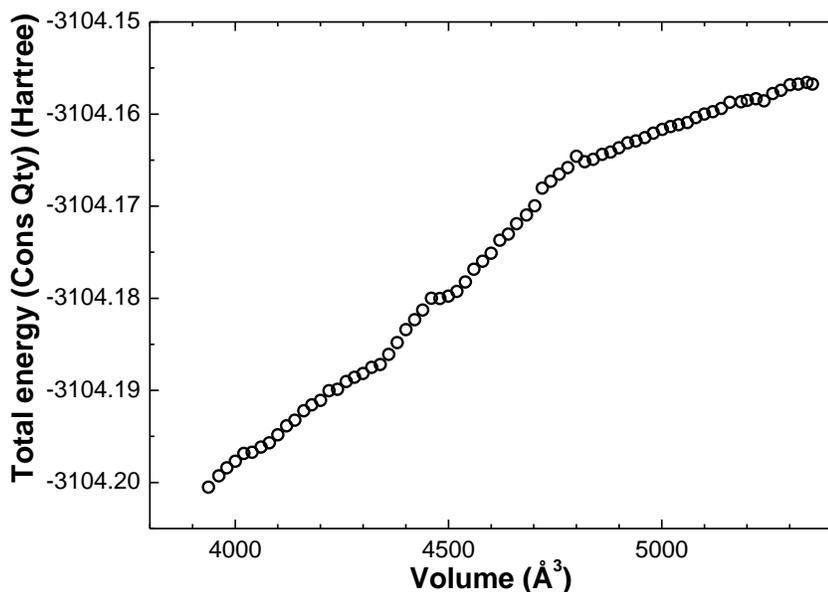


Figure S14. Total energy (conserved quantity) of a single snapshot from the NPT-AIMD simulation (for MIL-53(Sc)-*int*) as a function of the volume of the simulation box for a charge-density cutoff of 350 Ry, used in conjunction with a double- ζ valence plus polarization Gaussian basis set (DZVP). For clarity, the spacing between two adjacent data-points is approximately 20 \AA^3 (wherever possible), and the data used for plotting are given in Table S7. Clearly, despite the use of a DZVP basis set and a relatively low cutoff, the energy-volume curve obtained from our NPT-AIMD simulation does not exhibit any significant jumps/discontinuities for a volume change as large as 36%.

Table S5. Comparison of simulated cell parameters from the NPT-AIMD simulations with the different combinations of basis set and charge-density cutoff. ^a

test i	sim. variables ^b	a (\AA)	b (\AA)	c (\AA)	α ($^\circ$)	β ($^\circ$)	γ ($^\circ$)	V (\AA^3)	cost ^c (s)
1	DZVP & 350 Ry	20.589	21.304	14.821	92.656	88.438	38.377	4010.390	40.0
2 ^d	TZV2P & 350 Ry	20.540	21.241	14.879	92.336	88.248	38.826	4046.460	90.0
		(-0.24)	(-0.29)	(+0.39)	(-0.35)	(-0.21)	(+1.17)	(+0.90)	(+125.0)
3 ^d	DZVP & 700 Ry	20.713	21.296	14.906	92.621	88.671	38.384	4059.960	85.0
		(+0.60)	(-0.04)	(+0.57)	(-0.04)	(+0.26)	(+0.02)	(+1.23)	(+112.5)

^a All three tests started with the same system configuration, i.e., a snapshot taken from the NPT-AIMD simulation (using DZVP & 350 Ry) of MIL-53(Sc)-*int* at 196 K. All three tests were run for 1 ps and all cell parameters reported here were averaged over that period (1 ps). Note, a too low simulation temperature (e.g., 100K) would not allow meaningful comparison of the simulated cells based on the different basis sets and cutoffs, as the structure itself does not change significantly at very low temperatures. On the other hand, a too high simulation temperature (e.g., 623 K) would need much longer simulation time until the time-averaged cell yields a correct representation of the structure. ^b All other simulation parameters were kept the same. ^c Calculation time for one MD step using 512 cores on HECTOR (<http://www.hector.ac.uk>). ^d Results are compared to the corresponding ones obtained from test 1. The difference between test 2(3) and test 1 is given in parentheses and is defined as $(\text{test}_i - \text{test}_1)/\text{test}_1 \times 100\%$, where $i = 2, 3$. It is clear that neither using a larger basis set nor increasing the cutoff resulted in significant changes in the cell parameters, whereas the computational cost was more than doubled in both cases. Thus, the (DZVP & 350 Ry) combination is the obvious choice based on this comparison.

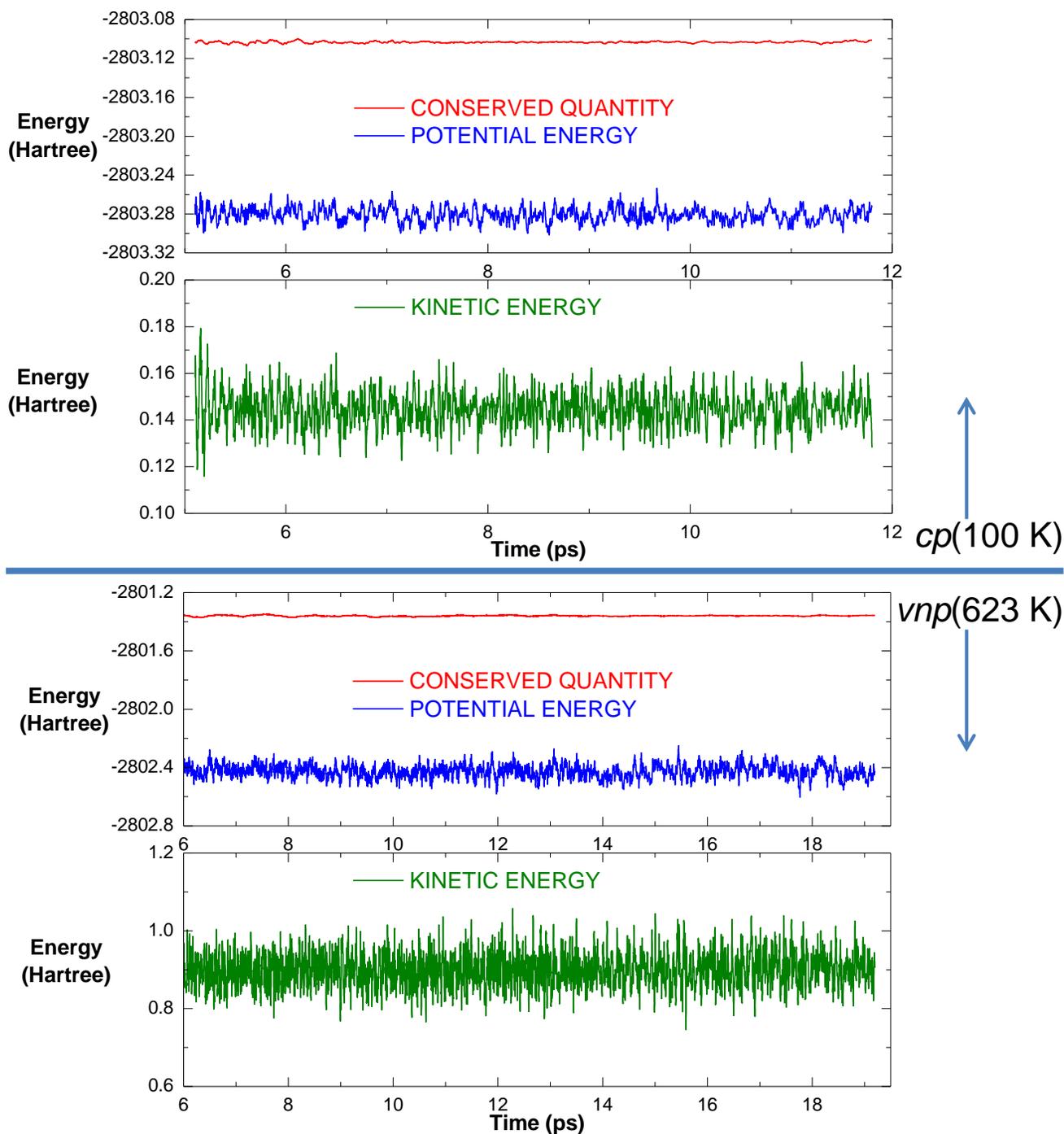


Figure S15. Time evolution of the total (conserved quantity), potential, and kinetic energies over the course of the production runs of the NPT-AIMD simulations at 100 and 623 K. In both cases, very good MD statistics can be observed: i.e., both simulations show very good conservation of the total energy with almost negligible drift in energy.

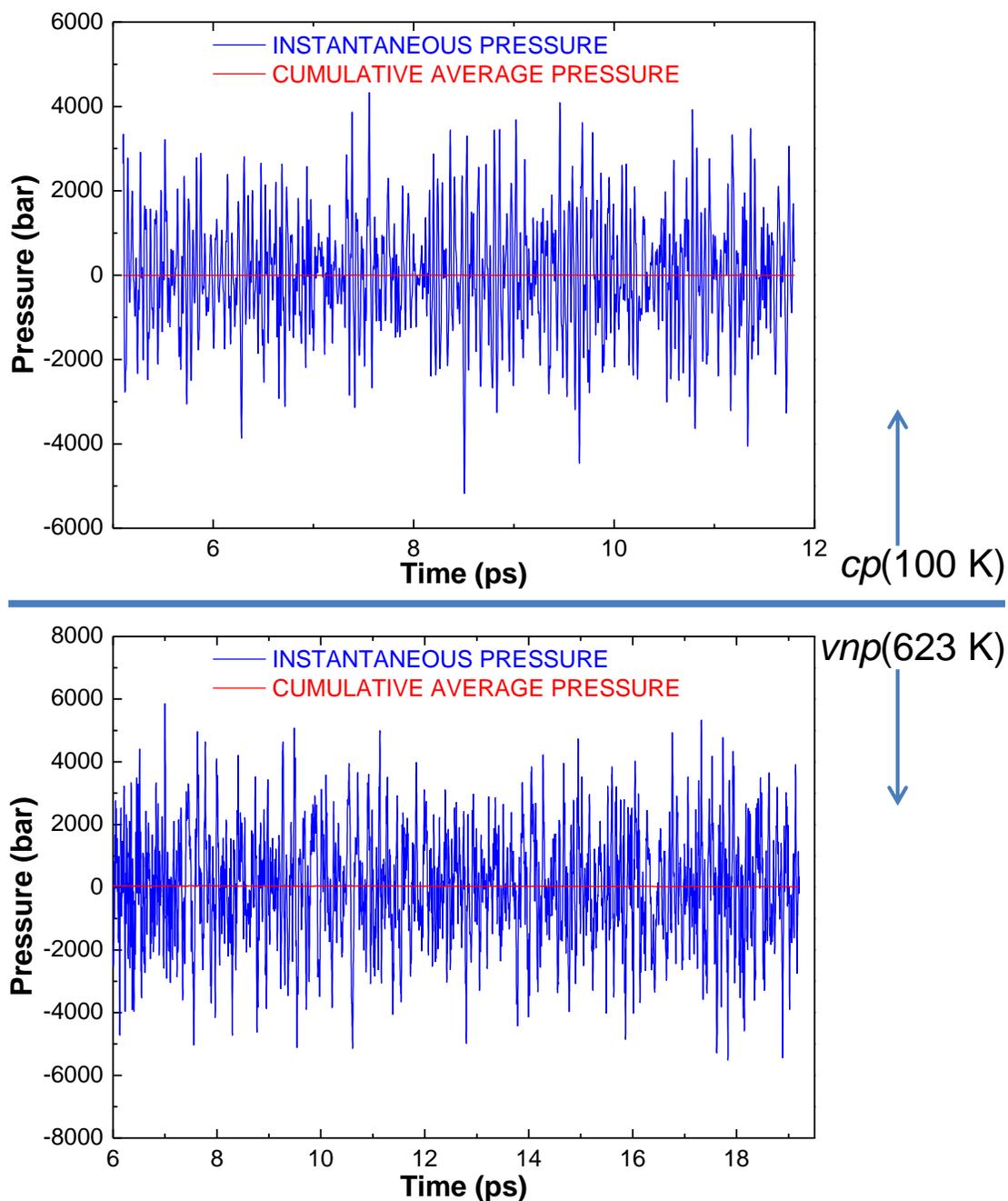


Figure S16. Time evolution of the instantaneous and cumulative average pressures over the course of the production runs of the NPT-AIMD simulations at 100 and 623 K. In both simulations, the externally applied pressure was 1 bar. The mean value of all instantaneous pressures for simulation at 100 and 623 K is 17 and -7 bar, respectively. Combining the results shown in Figure S15, these data indicate the computational setup adopted in this work is adequate enough to yield accurate results in NPT-AIMD simulations.

Table S6. Simulation parameters and corresponding average pressures from NPT-AIMD simulations.^a

system	basis set & cutoff ^b	\bar{p} (bar) ^c
MIL-53(Sc) at 1 bar & 100 K ^d	DZVP & 350 Ry	17 (1330)
MIL-53(Sc) at 1 bar & 623 K ^d	DZVP & 350 Ry	-7 (3715)
liquid water at 1 bar & 330 K ^e	TZV2P & 700 Ry ^f	-30 (3025)
		14 (2752)
		111 (2915)
	TZV2P & 1200 Ry ^f	46 (3494)
		-57 (3563)
		-175 (3439)
		16 (2913)
		22 (3514)

^a The pressure convergences in the NPT-AIMD simulations in this work are compared to those reported in the literature. These data demonstrate the general understanding of MD simulations in the NPT ensemble: i.e., converging the internal pressure is strongly dependent on the system size and on the simulation duration rather than a critical issue of the basis set/cutoff. Note, to our best knowledge, among the very few literature studies on applying AIMD to flexible MOFs,^{27,31} pressure data were not reported when the NPT ensemble was used.³¹ ^b DZVP(TZV2P) stands for a double(triple)- ζ valence basis set augmented with one set(two sets) of polarization functions. ^c \bar{p} is the average pressure, and the corresponding root-mean-square deviation (RMSD) value is given in parentheses. Internal pressure can be either positive or negative, as it represents the tendency of the simulation cell to expand or contract.¹² ^d This work. ^e Data were taken from ref 12. ^f Using the same basis set and cutoff, other simulation parameters were varied: i.e., DFT functional and the reference-cell treatment.

Table S7. Data used in Figure S14.

$V (\text{\AA}^3)$	<i>Cons. Qty.</i> (Hartree)	$V (\text{\AA}^3)$	<i>Cons. Qty.</i> (Hartree)	$V (\text{\AA}^3)$	<i>Cons. Qty.</i> (Hartree)
3937.726	-3104.20051328	4421.134	-3104.18233123	4900.180	-3104.16366831
3962.004	-3104.19929964	4440.049	-3104.18129217	4920.507	-3104.16314656
3981.011	-3104.19842529	4460.561	-3104.18002032	4940.092	-3104.16292429
4000.767	-3104.19769967	4480.167	-3104.18004518	4960.841	-3104.16258042
4020.876	-3104.19686256	4500.000	-3104.17978140	4980.930	-3104.16208419
4040.096	-3104.19672539	4520.012	-3104.17926282	5001.882	-3104.16166752
4060.879	-3104.19617234	4540.204	-3104.17823685	5021.083	-3104.16137024
4080.217	-3104.19569950	4560.141	-3104.17687212	5040.212	-3104.16116726
4100.064	-3104.19483689	4580.344	-3104.17597964	5060.258	-3104.16093324
4120.522	-3104.19384974	4600.038	-3104.17512334	5080.208	-3104.16039883
4140.250	-3104.19325476	4620.701	-3104.17371298	5100.502	-3104.16001616
4161.244	-3104.19221114	4640.726	-3104.17304411	5120.419	-3104.15976515
4180.076	-3104.19157492	4660.091	-3104.17192606	5140.005	-3104.15938655
4200.462	-3104.19110124	4683.419	-3104.17097368	5160.237	-3104.15872756
4220.134	-3104.19005485	4702.762	-3104.16996910	5186.312	-3104.15867854
4240.024	-3104.18990041	4720.273	-3104.16806047	5200.677	-3104.15853466
4260.503	-3104.18904191	4740.029	-3104.16730680	5221.195	-3104.15833755
4280.291	-3104.18858527	4760.075	-3104.16655077	5240.352	-3104.15858751
4300.037	-3104.18817020	4780.182	-3104.16581728	5260.574	-3104.15778225
4320.279	-3104.18750609	4800.275	-3104.16460027	5280.103	-3104.15743570
4340.070	-3104.18721375	4820.030	-3104.16519587	5300.781	-3104.15684033
4360.066	-3104.18611507	4840.024	-3104.16491612	5320.418	-3104.15676244
4380.029	-3104.18483589	4860.573	-3104.16438482	5340.930	-3104.15657647
4400.663	-3104.18341361	4880.552	-3104.16414053	5353.365	-3104.15675647

REFERENCES

- (1) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (2) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (3) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- (4) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703.
- (5) Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641.
- (6) Krack, M. *Theor. Chem. Acc.* **2005**, *114*, 145.
- (7) VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2007**, *127*, 114105.
- (8) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635.
- (9) Serre, C.; Millange, F.; Thouvenot, C.; Noguès, M.; Marsolier, G.; Louër, D.; Férey, G. *J. Am. Chem. Soc.* **2002**, *124*, 13519.
- (10) Vimont, A.; Travert, A.; Bazin, P.; Lavalley, J.-C.; Daturi, M.; Serre, C.; Férey, G.; Bourrelly, S.; Llewellyn, P. L. *Chem. Commun.* **2007**, 3291.
- (11) McGrath, M. J.; Siepmann, J. I.; Kuo, I. F. W.; Mundy, C. J.; VandeVondele, J.; Hutter, J.; Mohamed, F.; Krack, M. *ChemPhysChem* **2005**, *6*, 1894.
- (12) Schmidt, J.; VandeVondele, J.; Kuo, I. F. W.; Sebastiani, D.; Siepmann, J. I.; Hutter, J.; Mundy, C. J. *J. Phys. Chem. B* **2009**, *113*, 11959.
- (13) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: from algorithms to applications*; 2nd ed.; Academic Press: San Diego, CA 92101-4495, USA, 2002.
- (14) Mowat, J. P. S.; Miller, S. R.; Slawin, A. M. Z.; Seymour, V. R.; Ashbrook, S. E.; Wright, P. A. *Microporous Mesoporous Mater.* **2011**, *142*, 322.
- (15) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.
- (16) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (17) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.
- (18) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
- (19) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
- (20) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford CT, 2009.
- (21) Cygan, R. T.; Romanov, V. N.; Myshakin, E. M. *J. Phys. Chem. C* **2012**, *116*, 13079.
- (22) Whitfield, T. R.; Wang, X.; Liu, L.; Jacobson, A. J. *Solid State Sci.* **2005**, *7*, 1096.
- (23) Loiseau, T.; Serre, C.; Huguenard, C.; Fink, G.; Taulelle, F.; Henry, M.; Bataille, T.; Férey, G. *Chem. – Eur. J.* **2004**, *10*, 1373.
- (24) *Materials Studio 5.0*, Accelrys Software Inc.: San Diego, CA 92121, USA.
- (25) Rappe, A. K.; Goddard, W. A. *The Journal of Physical Chemistry* **1991**, *95*, 3358.

- (26) Mowat, J. P. S.; Seymour, V. R.; Griffin, J. M.; Thompson, S. P.; Slawin, A. M. Z.; Fairen-Jimenez, D.; Düren, T.; Ashbrook, S. E.; Wright, P. A. *Dalton Trans.* **2012**, *41*, 3937.
- (27) Kanoo, P.; Reddy, S. K.; Kumari, G.; Haldar, R.; Narayana, C.; Balasubramanian, S.; Maji, T. K. *Chem. Commun.* **2012**, *48*, 8487.
- (28) Li, W.; Probert, M. R.; Kosa, M.; Bennett, T. D.; Thirumurugan, A.; Burwood, R. P.; Parinello, M.; Howard, J. A. K.; Cheetham, A. K. *J. Am. Chem. Soc.* **2012**, *134*, 11940.
- (29) Mallik, B. S.; Siepmann, J. I. *J. Phys. Chem. B* **2010**, *114*, 12577.
- (30) Mallik, B. S.; Kuo, I. F. W.; Fried, L. E.; Siepmann, J. I. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4884.
- (31) Hoffmann, H. C.; Assfour, B.; Epperlein, F.; Klein, N.; Paasch, S.; Senkovska, I.; Kaskel, S.; Seifert, G.; Brunner, E. *J. Am. Chem. Soc.* **2011**, *133*, 8681.