A Comparison of Barostats for the Mechanical Characterization of Metal–Organic Frameworks

S.M.J. Rogge,[†] L. Vanduyfhuys,[†] A. Ghysels,[†] M. Waroquier,[†] T. Verstraelen,[†] G. Maurin,[‡] and V. Van Speybroeck[*,†]

Center for Molecular Modeling (CMM), Ghent University, Technologiepark 903, 9052 Zwijnaarde, Belgium, and Institut Charles Gerhardt Montpellier, Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier cedex 05, France

E-mail: Veronique.VanSpeybroeck@UGent.be

Supporting Information

[†]Ghent University
[‡]Université Montpellier

1 Hamiltonian derivation of equations of motion for the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, H)$ ensemble

1.1 Derivation of the Hamiltonian

In this section, a set of equations of motion is proposed, based on a Hamiltonian derivation. To obtain this set of equations, we start from the Lagrangian proposed by Parrinello and Rahman.[1] This Lagrangian describes a system consisting of N masses m_j under influence of a hydrostatic pressure $P\mathbf{1}$ and a potential $\mathcal{V}$, but without temperature control, thus describing the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, H)$ ensemble (see Table 1 in the main text for the used notation):

$$\mathcal{L}_{\mathrm{NPH}}(\boldsymbol{s}^N, \dot{\boldsymbol{s}}^N, \mathbf{h}, \dot{\mathbf{h}}) = \sum_{j=1}^{N} m_j \frac{\dot{\boldsymbol{s}}_j^T \mathbf{h}^T \mathbf{h} \dot{\boldsymbol{s}}_j}{2} - \mathcal{V} - P \det \mathbf{h} + W \frac{\mathrm{Tr}(\dot{\mathbf{h}}^T \dot{\mathbf{h}})}{2}, \tag{S1.1}$$

with W the barostat mass. In this equation, the fractional coordinates $\boldsymbol{s}_j$ are introduced, which are related to the real coordinates $\boldsymbol{r}_j$ of the particles through the unit cell tensor $\mathbf{h} = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$:

$$\boldsymbol{r}_j = \mathbf{h} \boldsymbol{s}_j \tag{S1.2}$$

In the remainder of the derivation, periodic boundary conditions are assumed.

To facilitate the derivation, it is assumed that the time derivative of the unit cell tensor $\mathbf{h}$ is much smaller than the time derivative of the fractional coordinates, so that

$$\dot{\boldsymbol{r}}_j = \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{h} \boldsymbol{s}_j) = \dot{\mathbf{h}} \boldsymbol{s}_j + \mathbf{h} \dot{\boldsymbol{s}}_j \cong \mathbf{h} \dot{\boldsymbol{s}}_j \tag{S1.3}$$

The first two terms of the Lagrangian in Eq. (S1.1) correspond to the usual kinetic and potential energy of the system of N particles, while the third and fourth terms can be regarded as the potential, respectively kinetic energy associated with the unit cell tensor $\mathbf{h}$, with W a fictitious mass associated with the cell tensor motion.

The reciprocal unit cell tensor $\mathbf{G} = [\boldsymbol{g}_a, \boldsymbol{g}_b, \boldsymbol{g}_c]$ with reciprocal basis vectors $\boldsymbol{g}_a, \boldsymbol{g}_b$ and $\boldsymbol{g}_c$, is defined as

$$\mathbf{h}^T \mathbf{G} = \mathbf{1} \Rightarrow \mathbf{G} = [\mathbf{h}^T]^{-1}, \tag{S1.4}$$

so that $\boldsymbol{\mu} \cdot \boldsymbol{g}_\nu = \delta_{\mu,\nu}$.

Using the Lagrangian of Eq. (S1.1), the atomic momenta are derived as

$$\boldsymbol{p}_{s,j} = \nabla_{\dot{\boldsymbol{s}}_j} \mathcal{L}_{\mathrm{NPH}} = m_j \mathbf{h}^T \mathbf{h} \dot{\boldsymbol{s}}_j \cong m_j \mathbf{h}^T \dot{\boldsymbol{r}}_j \tag{S1.5}$$

and the cell momentum tensor $\mathbf{p}_h$ is given by

$$p_{h,\mu\nu} = \frac{\partial \mathcal{L}_{\mathrm{NPH}}}{\partial \dot{h}_{\mu\nu}} = W \dot{h}_{\mu\nu} \Rightarrow \mathbf{p}_h = W \dot{\mathbf{h}} \tag{S1.6}$$

Hence, the Hamiltonian associated with the Lagrangian proposed by Parrinello and Rahman in Eq. (S1.1) takes the following form:

$$\begin{aligned} \mathcal{H}_{\mathrm{NPH}}(\boldsymbol{s}^N, \boldsymbol{p}_s^N, \mathbf{h}, \mathbf{p}_h) &= \sum_{j=1}^{N} \boldsymbol{p}_{s,j} \cdot \dot{\boldsymbol{s}}_j + \mathbf{p}_h : \dot{\mathbf{h}} - \mathcal{L}_{\mathrm{NPH}} && \text{(S1.7)} \\ &= \sum_{j=1}^{N} \frac{\boldsymbol{p}_{s,j}^T \mathbf{G}^T \mathbf{G} \boldsymbol{p}_{s,j}}{2m_j} + \mathcal{V} + \frac{\mathrm{Tr}(\mathbf{p}_h^T \mathbf{p}_h)}{2W} + P \det \mathbf{h} && \text{(S1.8)} \end{aligned}$$

Note that the cell momentum tensor $\mathbf{p}_h$ derived above is not the cell momentum tensor $\mathbf{p}_g$ used in the MTTK and Langevin barostat equations employed in the main text.

1.2 Derivation of the equations of motion

The equations of motion belonging to the Hamiltonian of Eq. (S1.8) are:

$$\begin{aligned} \dot{\boldsymbol{s}}_j &= \nabla_{\boldsymbol{p}_{s,j}} \mathcal{H}_{\mathrm{NPH}} = \mathbf{G}^T \mathbf{G} \frac{\boldsymbol{p}_{s,j}}{m_j} && \text{(S1.9a)} \\ \dot{\boldsymbol{p}}_{s,j} &= -\nabla_{\boldsymbol{s}_j} \mathcal{H}_{\mathrm{NPH}} = -\nabla_{\boldsymbol{s}_j} \mathcal{V} && \text{(S1.9b)} \\ \dot{h}_{\mu\nu} &= \frac{\partial \mathcal{H}_{\mathrm{NPH}}}{\partial p_{h,\mu\nu}} = \frac{p_{h,\mu\nu}}{W} \Rightarrow \dot{\mathbf{h}} = \frac{\mathbf{p}_h}{W} && \text{(S1.9c)} \\ \dot{p}_{h,\mu\nu} &= -\frac{\partial \mathcal{H}_{\mathrm{NPH}}}{\partial h_{\mu\nu}} = -\sum_{j=1}^{N} \frac{\boldsymbol{p}_{s,j}^T}{2m_j} \frac{\partial (\mathbf{G}^T \mathbf{G})}{\partial h_{\mu\nu}} \boldsymbol{p}_{s,j} - \frac{\partial \mathcal{V}}{\partial h_{\mu\nu}} - P G_{\mu\nu} \det \mathbf{h} && \text{(S1.9d)} \end{aligned}$$

where we used that:

$$\frac{\mathrm{d}}{\mathrm{d}A_{\mu\nu}} \det(\mathbf{A}) = \det(\mathbf{A}) \mathrm{Tr}\left(\mathbf{A}^{-1} \boldsymbol{\delta}_{\mu\nu}\right) = \det(\mathbf{A}) [\mathbf{A}^{-1}]_{\nu\mu} \tag{S1.10}$$

in which $\boldsymbol{\delta}_{\mu\nu}$ is the 3×3 matrix for which all elements are zero, except for the element on row μ and column ν.

Eq. (S1.9d) can be further simplified, using the definition of the reciprocal unit tensor $\mathbf{G}$ (S1.4):

$$
\begin{aligned}
-\sum_{j=1}^{N} \frac{\boldsymbol{p}_{s,j}^T}{2m_j} \frac{\partial(\mathbf{G}^T\mathbf{G})}{\partial h_{\mu\nu}} \boldsymbol{p}_{s,j} &= -\sum_{j=1}^{N} \frac{\boldsymbol{p}_{s,j}^T}{m_j} \frac{\partial \mathbf{h}^{-1}}{\partial h_{\mu\nu}} \mathbf{G}\boldsymbol{p}_{s,j} && \text{(S1.11)} \\
&= \sum_{j=1}^{N} \frac{\boldsymbol{p}_{s,j}^T}{m_j} \mathbf{h}^{-1} \frac{\partial \mathbf{h}}{\partial h_{\mu\nu}} \mathbf{h}^{-1} \mathbf{G}\boldsymbol{p}_{s,j} && \text{(S1.12)} \\
&= \sum_{j=1}^{N} \frac{1}{m_j} \left[(\mathbf{G}\boldsymbol{p}_{s,j}) \otimes \left(\mathbf{G}^T\mathbf{G}\boldsymbol{p}_{s,j} \right) \right]_{\mu\nu} && \text{(S1.13)} \\
&= \left[\sum_{j=1}^{N} \frac{\boldsymbol{p}_j \otimes (\mathbf{G}^T\boldsymbol{p}_j)}{m_j} \right]_{\mu\nu} && \text{(S1.14)}
\end{aligned}
$$

Here, the real momenta $\boldsymbol{p}_j = \mathbf{G}\boldsymbol{p}_{s,j}$ were introduced in a similar way as in the derivation of the equations of motion corresponding to a Nosé–Hoover chain.[2]

The potential energy $\mathcal{V}$ only depends on the relative position of the atoms:

$$\mathcal{V} = \mathcal{V}(\{\boldsymbol{r}'_k - \boldsymbol{r}_j\}_{j,k}) \tag{S1.15}$$

In the case of periodic boundary conditions, this dependency on the relative positions can be written out so that only atoms from one unit cell need to be taken into account. Since only those interatomic interactions are counted for which at least one of the atoms is located in the simulation cell, we may assume that $\boldsymbol{r}_j$ lies in this simulation cell, while $\boldsymbol{r}'_k$ is the position of an arbitrary atom. Due to the periodicity, $\boldsymbol{r}'_k$ can always be expressed as $\boldsymbol{r}_k + \boldsymbol{R}_n$, where $\boldsymbol{r}_k$ lies in the simulation cell and

$$\boldsymbol{R}_n = \mathbf{h}\boldsymbol{n}, \qquad \text{with } \boldsymbol{n} \in \mathbb{Z}^3 \tag{S1.16}$$

In this way, the dependence of the potential energy on the relative distances over the borders of the simulation cell transforms into

$$\mathcal{V} = \mathcal{V}(\{\boldsymbol{d}_{jkn}\}_{j,k,\boldsymbol{n}}) \tag{S1.17}$$

with

$$\boldsymbol{d}_{jkn} = \boldsymbol{r}_k - \boldsymbol{r}_j + \boldsymbol{R}_n \tag{S1.18}$$

and both $\boldsymbol{r}_j$ and $\boldsymbol{r}_k$ in the simulation cell. The second term of Eq. (S1.9d) reduces to:

$$-\frac{\partial \mathcal{V}}{\partial h_{\mu\nu}} = -\sum_{\kappa=x,y,z}\sum_{j\leq' k}\sum_{\boldsymbol{n}}\frac{\partial \mathcal{V}}{\partial d_{jkn,\kappa}}\frac{\partial d_{jkn,\kappa}}{\partial h_{\mu\nu}} \tag{S1.19}$$

$$= -\sum_{\kappa=x,y,z}\sum_{j\leq' k}\sum_{\boldsymbol{n}}\frac{\partial \mathcal{V}}{\partial d_{jkn,\mu}}(r_{k,\kappa} - r_{j,\kappa} + R_{n,\kappa})G_{\kappa\nu} \tag{S1.20}$$

making use of the relation $\boldsymbol{s}_j = \mathbf{h}^{-1}\boldsymbol{r}_j = \mathbf{G}^T\boldsymbol{r}_j$. The symbol $\leq'$ indicates that j should be smaller than or equal to k if $\boldsymbol{n} \neq \mathbf{0}$, and both indices should indicate atoms lying in the simulation cell; if $\boldsymbol{n} = \mathbf{0}$, the previous restrictions apply, and moreover, j and k cannot be equal, since they would point to the same atom otherwise.

Eq. (S1.20) can be further simplified using the definition of the virial tensor:[3]

$$\Xi_{\mu\kappa} = \sum_{j\leq' k}\sum_{\boldsymbol{n}}(r_{k,\mu} - r_{j,\mu} + R_{n,\mu})\frac{\partial \mathcal{V}}{\partial d_{jkn,\kappa}} \tag{S1.21}$$

so that

$$-\frac{\partial \mathcal{V}}{\partial h_{\mu\nu}} = -[\boldsymbol{\Xi}^T\mathbf{G}]_{\mu\nu} \tag{S1.22}$$

Substituting Eqs. (S1.20, S1.22) in Eq. (S1.9d) yields the equation of motion for the barostat momentum tensor $\mathbf{p}_h$, with $V = \det\mathbf{h}$:

$$\dot{\mathbf{p}}_h = \left[\frac{1}{V}\sum_{j=1}^{N}\frac{\boldsymbol{p}_j \otimes \boldsymbol{p}_j}{m_j} - \frac{\boldsymbol{\Xi}}{V} - P\mathbf{1}\right]^T V\mathbf{G} \tag{S1.23}$$

where we exploited the symmetry of $\boldsymbol{p}_j \otimes \boldsymbol{p}_j$ and the unit tensor, and the identity $\boldsymbol{p}_j \otimes (\mathbf{G}^T\boldsymbol{p}_j) = (\boldsymbol{p}_j \otimes \boldsymbol{p}_j)\mathbf{G}$.

The first two terms between brackets form the internal pressure tensor $\mathbf{P}_i$, counteracting the external pressure $P\mathbf{1}$. This leads to the expression used in the main text:

$$\mathbf{P}_i = \frac{1}{V}\left[\sum_{j=1}^{N}\frac{\boldsymbol{p}_j \otimes \boldsymbol{p}_j}{m_j} - \boldsymbol{\Xi}\right] \tag{S1.24}$$

and the equations of motion

$$\dot{\boldsymbol{r}}_j = \frac{\boldsymbol{p}_j}{m_j} \tag{S1.25a}$$

$$\dot{\boldsymbol{p}}_j = -\nabla_{\boldsymbol{r}_j}\mathcal{V} = \boldsymbol{F}_j \tag{S1.25b}$$

$$\dot{\mathbf{h}} = \frac{\mathbf{p}_h}{W} \tag{S1.25c}$$

$$\dot{\mathbf{p}}_h = \left(\mathbf{P}_i - P\mathbf{1}\right)^T V\mathbf{G} \tag{S1.25d}$$

where again the assumption expressed in Eq. (S1.3) was used, as well as a similar expression for the atomic momenta. These equations of motion correspond to the ones implemented by Parrinello and Rahman.[1]

1.3 Interpretation of the equations of motion

This last equation of motion is very similar to Newton's second law for the atomic momenta $\boldsymbol{p}_j$, since we can identify $\left(\mathbf{P}_i - P\mathbf{1}\right)^T V\mathbf{G}$ with a force tensor $\mathbf{F}_{\text{press}}$. To see this, we will introduce the net pressure tensor $\mathbf{P}^{\text{net}} = \mathbf{P}_i - P\mathbf{1}$ and assume, for simplicity, a tetragonal cell with unit cell tensor $\mathbf{h} = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}] = \text{diag}(a, b, c)$, so that

$$V\mathbf{G} = \begin{bmatrix} bc & 0 & 0 \\ 0 & ac & 0 \\ 0 & 0 & ab \end{bmatrix} = \begin{bmatrix} |(\boldsymbol{b}\times\boldsymbol{c})\cdot\mathbf{1}_a| & 0 & 0 \\ 0 & |(\boldsymbol{a}\times\boldsymbol{c})\cdot\mathbf{1}_b| & 0 \\ 0 & 0 & |(\boldsymbol{a}\times\boldsymbol{b})\cdot\mathbf{1}_c| \end{bmatrix} \tag{S1.26}$$

Here, $V\mathbf{G}$ contains the projected surface areas. In general, the element $VG_{\mu\nu}$ denotes the area of the surface spanned by the two cell vectors not contained in the μth column of the unit cell tensor $\mathbf{h}$, projected on the surface normal to the ν axis. This results in

$$\dot{\mathbf{p}}_h = \left(\mathbf{P}_i - P\mathbf{1}\right)^T V\mathbf{G} = \begin{bmatrix} bcP^{\text{net}}_{xx} & acP^{\text{net}}_{yx} & abP^{\text{net}}_{zx} \\ bcP^{\text{net}}_{xy} & acP^{\text{net}}_{yy} & abP^{\text{net}}_{zy} \\ bcP^{\text{net}}_{xz} & acP^{\text{net}}_{yz} & abP^{\text{net}}_{zz} \end{bmatrix} \tag{S1.27}$$

Hence, the time derivative of the cell momentum tensor $\mathbf{p}_h$ is given by a tensor whose ele-

ments $\mathbf{F}_{\text{press},\mu\nu}$ denote the pressure parallel to the μ axis, exerted on the surface normal to the ν axis, and multiplied by the signed area of this surface, a result that still holds for less symmetric cell tensors. This latter can be interpreted as a force along the μ axis, on this given surface. The equations of motion (S1.25a–S1.25d) thus consist of two equations of the form $\dot{x} = p/m$ (S1.25a, S1.25c) and two equations of the form $\dot{p} = F$ (S1.25b, S1.25d).

While the here derived equations of motion also yield the correct $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ ensemble when coupled to a thermostat, their form does not allow for an easy implementation. However, they closely resemble the equations of motion as introduced by Martyna, Tuckerman, Tobias and Klein,[4,5] which are discussed in the main text, but cannot be directly derived using this Hamiltonian approach.

2 Characterization of the (meta)stable states

In this section, the stable and metastable states of MIL-53(Al) and the stable state of MOF-5 are analysed. First, it is stated how both the large-pore (LP) and closed-pore (CP) averages for MIL-53(Al) can be determined based on one simulation starting in the LP phase. Furthermore, the average cell parameters of the three structures and the corresponding standard errors are reported. Special attention is given to the cell parameters in the LP phase, where significant discrepancies have been noticed between the Berendsen predictions and those of the two other barostats.

The MD simulations were carried out in the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ ensemble, with $P = 100$ kPa and $T = 300$ K, for a total duration of 800 ps (MIL-53(Al)), respectively 600 ps (MOF-5). The Verlet timestep was 0.5 fs (MIL-53(Al)) or 0.75 fs (MOF-5) and a thermostat relaxation constant τ_T of 0.1 ps was chosen. In the case of MIL-53(Al), the simulation cell contains 152 atoms, corresponding to two conventional unit cells, while for MOF-5 one conventional unit cell of 424 atoms was sufficient.

The same procedure was also carried out at other pressures, from 300 kPa up to 30 MPa for MOF-5, and up to 1 GPa for MIL-53(Al). However, the difference in cell parameters was seen to be smaller than the standard deviation on the simulation results, and these cell parameters are hence not reported. This observation is confirmed by the large bulk modulus, as reported in the main paper.

2.1 Definition of the LP and CP averages

In Figure S1, one of the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations at 1 MPa and 300 K for MIL-53(Al) is shown. Pressure control in that simulation was achieved using the MTTK barostat with a relaxation time $\tau_P = 5$ ps. During this simulation, a LP-to-CP transition is observed after about 320 ps. As mentioned in the methodology section (Eq. 2.8), the time to transition $t_{\mathrm{LP}\to CP}$ is defined as the time needed for the system to first cross the threshold volume V_{tr}, here taken to be 1000 Å^3. This threshold volume, as well as the time to transition, are indicated in gray on Figure S1.

Since we are interested in cell parameters of the stable or metastable state, part of the simulation corresponding to the transition region should be discarded. As mentioned in the computa-

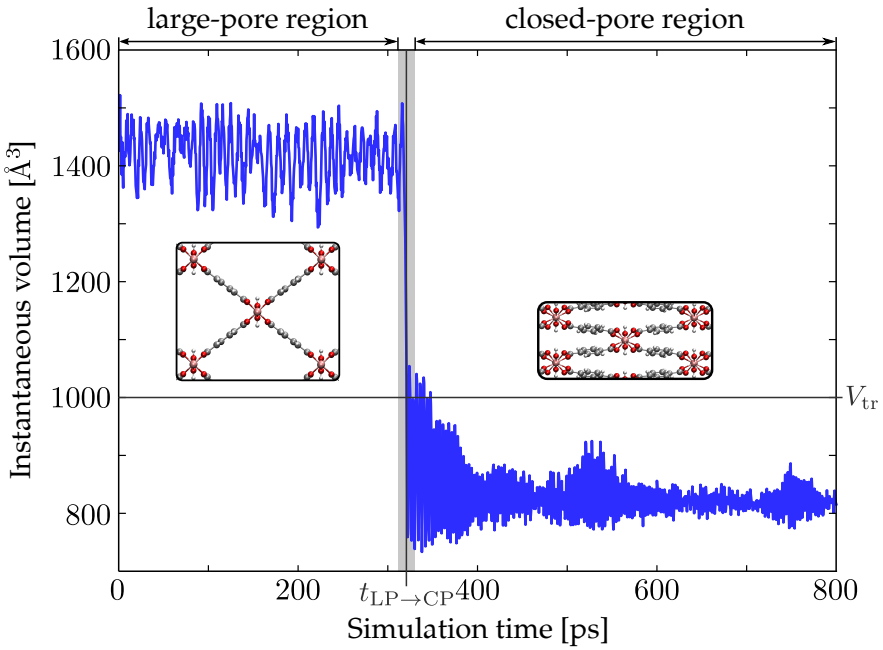


Figure S1: Instantaneous volume as a function of the simulation time for an $(N, P, \boldsymbol{\sigma}_{=}\mathbf{0}, T)$ simulation at 1 MPa and 300 K for MIL-53(Al), using a MTTK barostat with a relaxation time $\tau_P = 5$ ps. Indicated in this figure are the threshold volume V_{tr}, the time to transition $t_{\mathrm{LP}\to\mathrm{CP}}$, and the regions which are considered for the calculation of the LP and CP averages in Section 4.1.

tional details section, a transition interval of $4\tau_P$ is taken into account, centered around $t_{\mathrm{LP}\to\mathrm{CP}}$. This interval, which is omitted in the calculation of the averages, is indicated with the gray box in Figure S1. This ensures that the reported averages of the cell parameters and internal coordinates in Section 4.1 correspond to the large-pore, respectively the closed-pore phase.

2.2 Cell parameters

In Tables S1–S3, the cell parameters are reported for the stable and metastable structures of MIL-53(Al), as well as for the stable structure of MOF-5. The a and c cell lengths are visualised in Figure 3 of the main article, while b is measured perpendicular to those directions. The α cell angle is defined as the angle in between the unit cell vectors $\boldsymbol{b}$ and $\boldsymbol{c}$, while β is the angle in between $\boldsymbol{c}$ and $\boldsymbol{a}$, and γ is enclosed by the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. In these tables, we use the following convention to indicate the standard error: 16.506(4) should be read as 16.506 ± 0.004, while 17516(29) indicates 17516 ± 29.

Table S1: Cell parameters for the large-pore structure of MIL-53(Al), based on the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations depicted in Figure 3 in the main article, calculated at 300 K and 100 kPa. The brackets indicate the standard error on these results.

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
a [Å]	16.506(4)	16.542(3)	16.505(2)	17.096(165)	17.077(63)	17.060(48)	16.990(108)	17.060(48)	17.042(23)
b [Å]	6.624(3)	6.617(1)	6.625(1)	6.589(6)	6.589(3)	6.590(2)	6.597(10)	6.590(3)	6.591(1)
c [Å]	13.478(5)	13.442(3)	13.479(2)	12.610(265)	12.665(104)	12.697(76)	12.791(168)	12.686(71)	12.721(23)
α [°]	90.001(6)	90.001(3)	90.001(2)	90.011(41)	90.003(10)	89.996(7)	90.071(110)	89.989(32)	89.994(32)
β [°]	90.001(3)	90.000(1)	90.000(1)	90.001(9)	90.000(2)	89.999(4)	89.983(14)	89.998(9)	90.000(5)
γ [°]	90.03(6)	90.00(2)	89.99(2)	89.91(45)	90.03(14)	90.02(8)	89.98(44)	90.08(25)	89.97(18)
V [Å^3]	1473.6(9)	1471.2(4)	1473.9(3)	1416.6(190)	1422.2(76)	1425.0(54)	1430.3(121)	1423.2(45)	1426.2(26)

Table S2: Cell parameters for the closed-pore structure of MIL-53(Al), based on the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations depicted in Figure 3 in the main article, calculated at 300 K and 100 kPa. The brackets indicate the standard error on these results.

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
a [Å]	19.4906(19)	19.4958(32)	19.4977(11)	19.4824(22)	19.4787(31)	19.4710(104)	19.4820(9)	19.4829(13)	19.4842(14)
b [Å]	6.5159(3)	6.5172(3)	6.5183(2)	6.5196(3)	6.5189(10)	6.5182(20)	6.5195(2)	6.5194(3)	6.5193(4)
c [Å]	6.506(4)	6.471(2)	6.442(2)	6.538(3)	6.568(15)	6.626(62)	6.548(3)	6.541(2)	6.537(5)
α [°]	89.998(11)	89.999(2)	89.998(2)	89.998(3)	90.000(5)	90.005(10)	89.997(14)	89.997(33)	90.009(18)
β [°]	90.001(4)	89.999(2)	89.998(2)	89.998(3)	90.000(5)	90.005(10)	89.997(14)	89.997(33)	90.009(18)
γ [°]	98.312(16)	96.368(18)	96.416(9)	96.220(20)	96.188(16)	96.036(213)	96.195(11)	96.210(16)	96.213(4)
V [Å^3]	821.3(5)	817.2(3)	813.6(2)	824.1(4)	827.8(18)	834.9(72)	825.4(3)	824.6(3)	824.2(6)

Table S3: Cell parameters for the stable MOF-5 structure, based on the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations depicted in Figure 3 in the main article, calculated at 300 K and 100 kPa. The brackets indicate the standard error on these results.

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
a [Å]	25.972(21)	25.935(11)	25.899(16)	26.004(12)	26.001(7)	26.001(7)	25.986(6)	25.987(4)	25.989(6)
b [Å]	25.965(18)	25.935(7)	25.892(13)	26.008(8)	25.995(9)	26.001(8)	25.985(5)	25.989(4)	25.989(8)
c [Å]	25.974(17)	25.930(16)	25.893(17)	26.001(13)	25.998(13)	25.994(14)	25.988(4)	25.985(5)	25.987(4)
α [°]	90.002(9)	90.001(3)	90.000(2)	90.002(12)	90.001(18)	90.005(15)	90.007(44)	89.989(33)	90.006(20)
β [°]	90.004(14)	90.001(5)	90.000(2)	90.006(19)	89.996(16)	90.003(20)	90.008(47)	90.006(43)	90.000(38)
γ [°]	90.000(8)	90.002(4)	89.998(2)	90.002(9)	89.995(14)	89.997(17)	90.002(34)	90.007(17)	90.018(46)
V [Å^3]	17516(29)	17442(17)	17364(27)	17571(16)	17558(17)	17563(15)	17533(7)	17535(5)	17539(8)

2.3 Distribution of the MIL-53(Al) LP averages

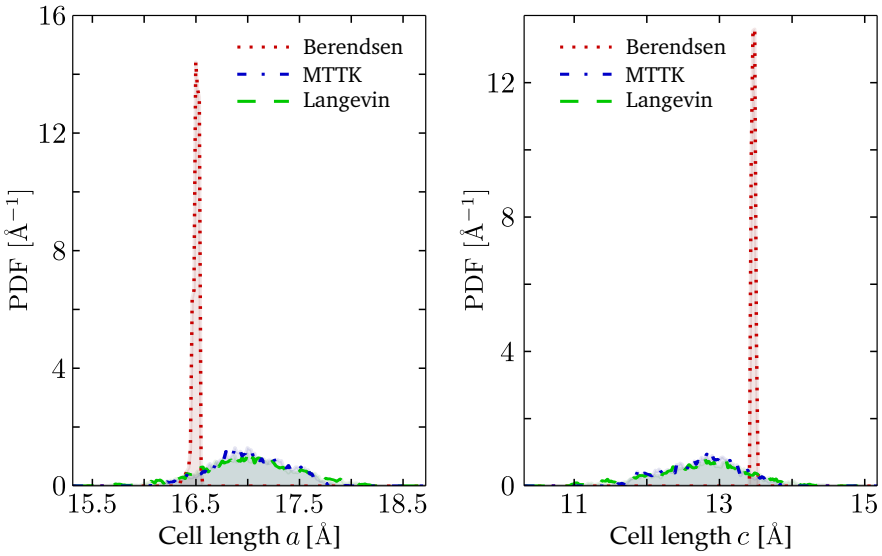


Figure S2: Probability density functions for the cell lengths a (left) and c (right) in the LP phase during one of the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ MD simulations of MIL-53(Al) at 300 K and 0.1 MPa used to generate Figure 3 of the main text. A barostat relaxation time of 10 ps has been used.

Figure 3 in the main manuscript displays the cell parameters for the LP and CP equilibrium states of MIL-53(Al) and MOF-5 as derived from $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations. While the three barostats predict similar unit cell vectors for the MIL-53(Al) CP phase and the stable state of MOF-5, the discrepancy noticed in the LP phase of MIL-53(Al) in the reproduction of the soft directions a and c requires some elucidation. In Figure S2, we plot the PDFs of these two cell lengths a and c as resulting from an $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulation at 300 K and 0.1 MPa for the three barostats and a barostat relaxation time of 10 ps. It was shown in Figure 3 of the main manuscript that the barostat relaxation time does not influence the average cell parameters.

Two apparent conclusions emerge from Figure S2: (i) the in the PDFs corresponding with the Berendsen barostat is much narrower, and (ii) the shift of the positions of the peaks with respect to the mean values of the cell vectors a and c predicted by the MTTK and Langevin barostats.

The sharp peak observed in both PDF plots suggests an extremely small area of possible (a, c) values encountered during the MD simulation when using the Berendsen barostat. The small red strip in the correlation diagram of Figure S3, corresponding to the Berendsen barostat, highly contrasts with the broad blue scatter region belonging to the MTTK barostat. This is

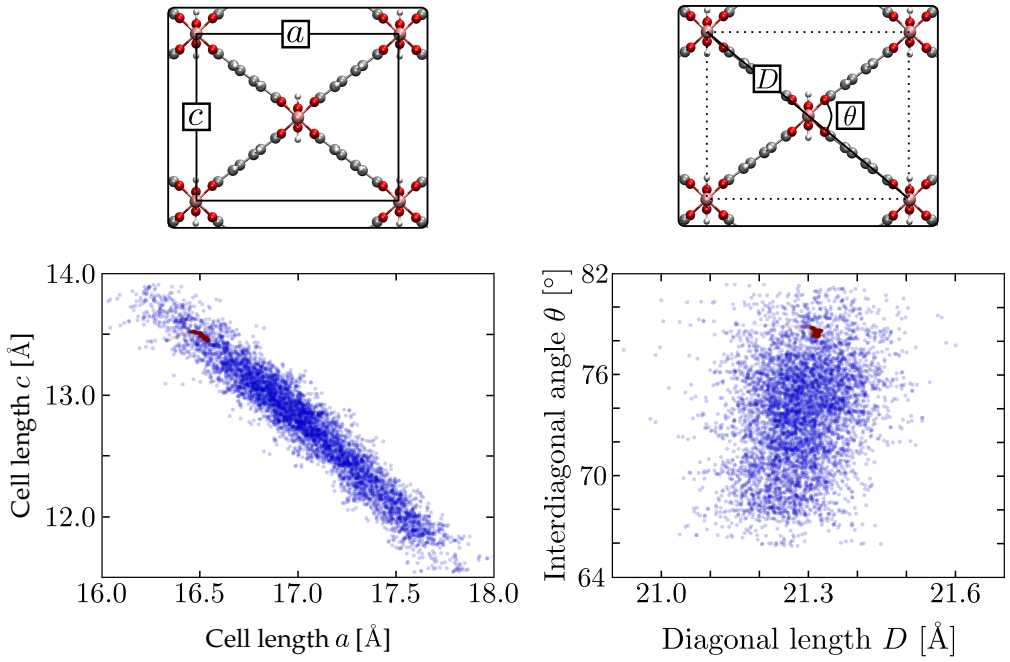


Figure S3: Scatter plot for the plane spanned by the cell lengths a and c (left) and the plane spanned by the diagonal length D and the interdiagonal angle θ (right) in the LP phase during one of the $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ MD simulations of MIL-53(Al) at 300 K and 0.1 MPa used to generate Figure 3 of the main text. For both the Berendsen (red) and MTTK (blue) barostat, a barostat relaxation time of 10 ps has been used.

a direct consequence of the reduced pressure fluctuations of the Berendsen barostat, since the large pressure fluctuations are responsible for the large fluctuations in cell lengths. We note the striking linear correlation observed between the a and c cell lengths with a correlation coefficient of -0.97. Hence, a reduction in the unit cell length a will lead to an increase in the cell length c, and vice versa, indicating a clear coupling of these soft modes. The corresponding free energy profile in the (a, c) plane corresponding to the LP phase is hence characterised by an elongated, narrow minimum, shown in the left pane of Figure S3.

Instead of the (a, c) correlation diagram, we could also construct a scatter diagram displaying the diagonal length D and interdiagonal angle θ, as shown in the first row of Figure S3. Although the two scatter plots of Figure S3 contain the same information, the two observables D and θ are less correlated, leading to a better defined global minimum in the free energy (or maximum in the probability distribution). Also here, the Berendsen results are confined to a small strip succeeding in determining the correct mean diagonal length D, but failing in reproducing the

interdiagonal angle θ. From the right pane of Figure S3, it is clear that the Berendsen barostat only allows a very small range of diagonal lengths around 21.3 Å, which corresponds to the minimum obtained with the MTTK barostat. To achieve the global minimum in the (D, θ) plane, θ should be decreased, while the allowed change in D is limited for the Berendsen barostat. Hence, a should be increased and c should be decreased simultaneously to decrease θ, while limiting the change in D. Any deviation from this estimate corresponds to an increase in free energy. This free energy barrier can be overcome by sufficiently large pressure fluctuations. However, these fluctuations are exactly what is suppressed when using the Berendsen scheme. As a result, the Berendsen scheme will be noticeably slower to converge in a narrow, elongated free energy minimum as present here.

3 Free energy profile

To obtain the pressure and free energy profiles of Figure 6 and Figure 7, $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations were carried out on a volume grid with a spacing amounting to 5 Å^3 (MIL-53(Al)) or 10 Å^3 (MOF-5), with the same computational details as in the previous section. The average internal pressure was fitted to polynomial as a function of volume, and the resulting data were extracted from this fit. For MIL-53(Al), this polynomial was of eleventh order, while a fifth-order polynomial sufficed for MOF-5. The coefficients of these polynomials are reported in the next paragraph. Subsequently, the properties derived from these simulations are reported. Finally, we also investigate the fluctuations in cell shape during $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations for MIL-53(Al), focussing on the cell lengths a and c. We also analyze the internal pressure as a function of these cell lengths, and explain how the linear behavior of the internal pressure can explain the observed coinciding pressure profiles of Figures 6 and 7 of the main paper. Finally, we investigate the free enthalpy profile as an alternative to the pressure and the free energy profiles provided in the main text, determining the different stable and metastable states at a given pressure. Moreover, we prove via a committor analysis that the local maximum of the free energy profile corresponds to a transition state.

3.1 Coefficients of the polynomial fit

To reduce the noise, the average internal pressure $\langle P_i \rangle$ as a function of the volume V for each set of barostat and barostat relaxation time is fitted to a polynomial, following:

$$\langle P_i(V) \rangle = \sum_{i=0}^{n} c_i \left(\frac{V}{\overline{V}} \right)^i \tag{S3.1}$$

Here, $\overline{V}$ indicates the average of the volume range over which the fit is carried out, c_i are the expansion coefficients, and n is the order of the polynomial. While the inclusion of the scaling factor $\overline{V}$ is optional, it results in dimensionless coefficients c_i. As indicated before, $n = 11$ for MIL-53(Al), while $n = 5$ for MOF-5. In Table S4 (MIL-53(Al)) and Table S5 (MOF-5), these coefficients and the average volume are reported.

Table S4: Expansion coefficients and average volume of Eq. (S3.1) for the eleventh-order fitted polynomial $\langle P_i(V) \rangle$ of MIL-53(Al).

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
c_0 [–]	−1.59008	−1.25024	0.91957	−2.22197	−3.58651	−0.57022	−1.90592	−2.30559	−1.50672
c_1 [–]	19.12706	14.80045	−10.93514	26.18793	42.25686	6.93607	22.49824	27.11946	17.68089
c_2 [–]	−103.12052	−78.43614	59.38026	−138.56336	−223.99605	−37.27945	−119.12735	−143.26768	−92.90548
c_3 [–]	329.52481	246.14169	−193.68855	434.97511	705.67489	117.46774	373.98608	449.19608	289.00235
c_4 [–]	−694.37948	−508.93296	420.55134	−900.88182	−1468.89647	−241.84621	−774.18818	−929.48110	−591.96741
c_5 [–]	1014.06102	728.74214	−636.99586	1293.32802	2122.14156	342.19782	1110.35507	1333.50885	838.92170
c_6 [–]	−1048.00038	−737.93957	685.86139	−1313.87089	−2172.06746	−339.84218	−1126.36733	−1354.11535	−839.72810
c_7 [–]	766.85596	528.73549	−524.46125	944.80555	1575.45281	236.92467	808.43677	973.53906	593.82656
c_8 [–]	−389.51075	−262.80094	278.95406	−471.42169	−793.76823	−113.57925	−402.42557	−485.75119	−290.77173
c_9 [–]	130.83415	86.32168	−98.25441	155.46777	264.62466	35.61672	132.33401	160.22067	93.88720
c_{10} [–]	−26.16137	−16.86700	20.62257	−30.50189	−52.54437	−6.56210	−25.87467	−31.44532	−17.98788
c_{11} [–]	2.35958	1.48539	−1.95398	2.69724	4.70830	0.53639	2.27886	2.78212	1.54861
$\overline{V}$ [Å^3]	1135.3	1134.4	1133.2	1130.0	1130.0	1130.0	1130.0	1130.0	1130.0

Table S5: Expansion coefficients and average volume of Eq. (S3.1) for the fifth-order fitted polynomial $\langle P_i(V) \rangle$ of MOF-5.

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
c_0 [–]	−3.26938	5.11326	0.35147	5.74757	6.24227	6.13908	7.72948	7.51857	7.06447
c_1 [–]	16.15398	−25.85593	−1.99040	−28.87692	−31.33841	−30.88933	−38.84610	−37.73010	−35.57075
c_2 [–]	−31.94106	52.24869	4.41575	57.98740	62.88434	62.12116	78.03961	75.68469	71.58951
c_3 [–]	31.59294	−52.74324	−4.81967	−58.17694	−63.04588	−62.41886	−78.33749	−75.85985	−71.98903
c_4 [–]	−15.63147	26.5981	2.59660	29.16181	31.58131	31.33641	39.29343	37.99345	36.17049
c_5 [–]	3.09498	−5.36089	−0.55376	−5.84292	−6.32363	−6.28847	−7.87893	−7.60676	−7.26468
$\overline{V}$ [Å^3]	17807.7	17798.9	17802.7	17800.1	17804.6	17796.7	17809.1	17780.4	17816.1

3.2 Derived properties

For both MIL-53(Al) and MOF-5, we generated one pressure profile for each barostat and barostat relaxation time. One such profile consists of a set of MD trajectories, one per volume grid point. Afterwards, the bootstrap method was applied to generate 99 extra trajectories per MD trajectory.[6,7] These bootstrapped trajectories are generated by random sampling of the configurations from the original trajectory with replacement. In this way, bootstrapped pressure profiles can be generated, and an estimate can be made of the average of and the sampling error on the simulated results, such as the equilibrium volumes and transition pressures. The sampling errors provided in Tables S6–S7 are calculated in this way, and give a reliable estimate of the uncertainty on the equilibrium volumes and the transition pressures. For the bulk moduli and differences in free energy, one should also take into account that there is not only an uncertainty in the calculated property, but also in the volume at which this property is calculated, so that the tabulated uncertainties for these properties are only a lower bound. Moreover, note that this sampling error only consists of the error which could be introduced by a nonconverged simulation. Other errors, such as those introduced by the force field or nonergodicity of the algorithms, are contained in the statistical error, which is not calculated here.

Table S6: Equilibrium volumes, transition pressures, bulk moduli and energy differences between the stable and metastable state of MIL-53(Al), based on the $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations depicted in Figure 6 of the main article. The reported errors are the standard deviations calculated based on 100 bootstrapped trajectories.

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
V_{CP} [Å^3]	821.91(18)	821.39(22)	821.26(26)	819.69(26)	819.48(26)	820.25(25)	820.70(26)	820.61(9)	820.70(10)
V_{tr} [Å^3]	1269.5(12)	1268.9(12)	1272.3(13)	1273.3(16)	1272.0(14)	1271.6(17)	1268.1(16)	1268.4(5)	1270.1(4)
V_{LP} [Å^3]	1453.4(4)	1453.5(4)	1454.1(4)	1448.0(7)	1448.4(6)	1447.4(7)	1450.0(6)	1449.0(2)	1447.7(2)
$-P_{\mathrm{CP}\rightarrow\mathrm{LP}}$ [MPa]	178.0(7)	178.4(8)	177.0(7)	183.0(9)	183.0(9)	183.1(9)	179.1(9)	180.3(3)	179.6(3)
$P_{\mathrm{LP}\rightarrow\mathrm{CP}}$ [MPa]	33.6(5)	36.8(6)	38.3(6)	28.4(8)	28.2(6)	27.4(6)	30.2(6)	30.0(2)	29.8(2)
K_{CP} [GPa]	3.333(20)	3.417(18)	3.488(22)	3.658(29)	3.641(27)	3.664(28)	3.566(25)	3.623(8)	3.625(9)
K_{LP} [GPa]	2.58(4)	2.52(4)	2.44(4)	1.63(4)	1.67(4)	1.59(4)	1.73(4)	1.70(1)	1.64(1)
$\Delta E_{\mathrm{tr}\rightarrow\mathrm{CP}}$ [kJ/mol]	27.55(3)	27.72(8)	27.68(8)	28.79(9)	28.74(11)	28.62(12)	27.85(11)	28.05(3)	28.02(4)
$\Delta E_{\mathrm{LP}\rightarrow\mathrm{CP}}$ [kJ/mol]	24.54(7)	24.75(7)	24.75(7)	25.57(8)	25.52(10)	25.52(10)	24.79(10)	25.03(3)	24.97(3)

Table S7: Equilibrium volume and bulk modulus as calculated for MOF-5 based on the $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations depicted in Figure 7 in the main article. The reported errors are the standard deviations calculated based on 100 bootstrapped trajectories.

	Berendsen			MTTK			Langevin		
	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps	1 ps	5 ps	10 ps
V_0 [Å^3]	17527.3(4)	17541.6(4)	17535.5(4)	17538.5(4)	17542.3(5)	17540.8(4)	17544.2(4)	17544.1(4)	17545.9(4)
K_0 [GPa]	7.033(10)	7.265(9)	7.263(10)	6.612(11)	6.738(12)	6.735(12)	6.863(11)	6.629(10)	7.076(11)

3.3 Cell shape fluctuations

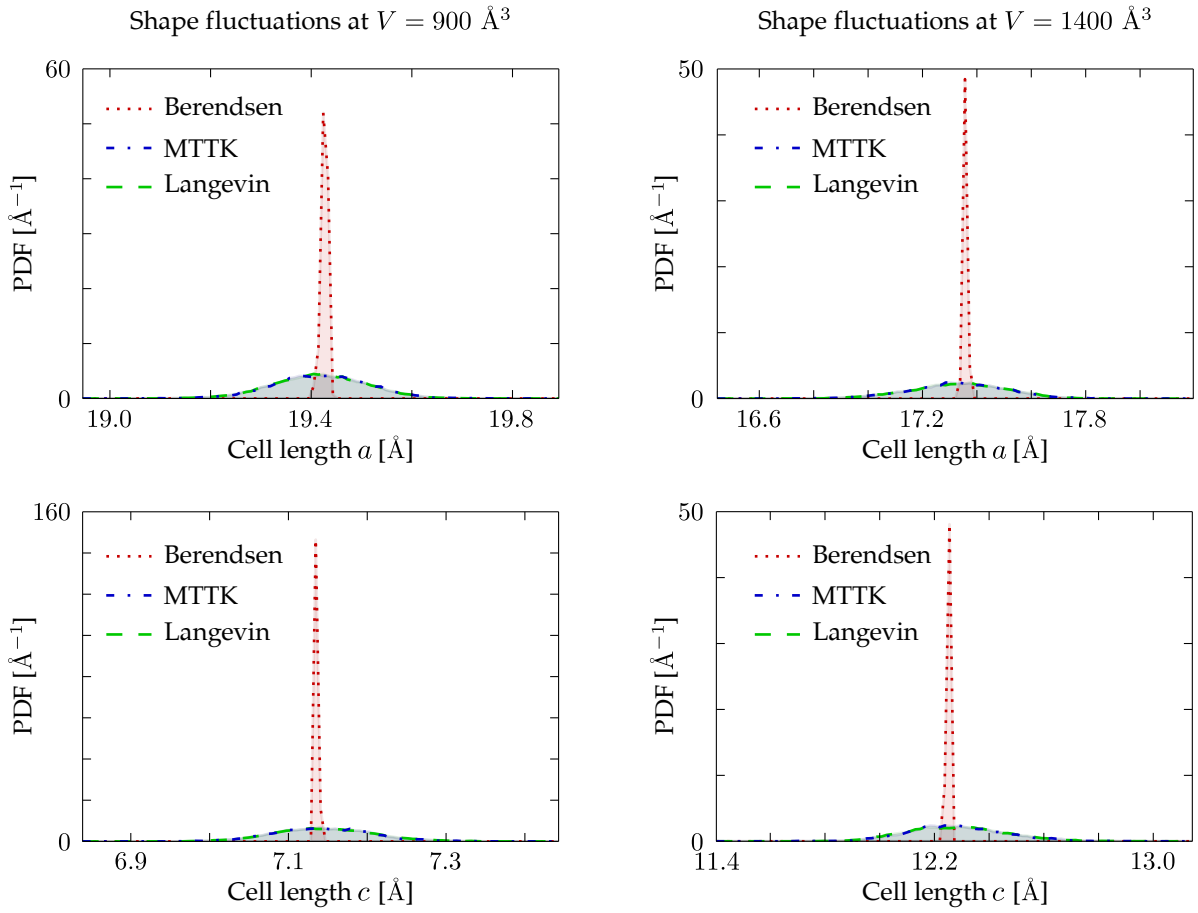


Figure S4: Probability density functions for the cell lengths a (top row) and c (bottom row) during an $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ MD simulation of MIL-53(Al) at 300 K. The two left panes are generated for a unit cell volume of 900 Å^3 (close to the CP phase), while the two right panes correspond to $V = 1400$ Å^3 (close to the LP phase). For all distributions, a barostat relaxation time of 10 ps has been used.

In the left pane of Figure 2 of the main paper, it was shown that the pressure fluctuations were suppressed when using the Berendsen barostat in $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations. This also leads to a reduction of the true volume fluctuations in the simulations, of the fluctuations in the distribution function for the internal coordinates in the LP phase, shown in Figure 4 and Figure 5 of the main paper, and of the fluctuations in the distribution function for the cell parameters in the LP phase, as shown in Figure S2 of the Supporting Information. For the $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations discussed in this section, the volume is kept constant. However, the reduced pressure fluctuations are still expected to affect the distribution of the five cell shape parameters. Here,

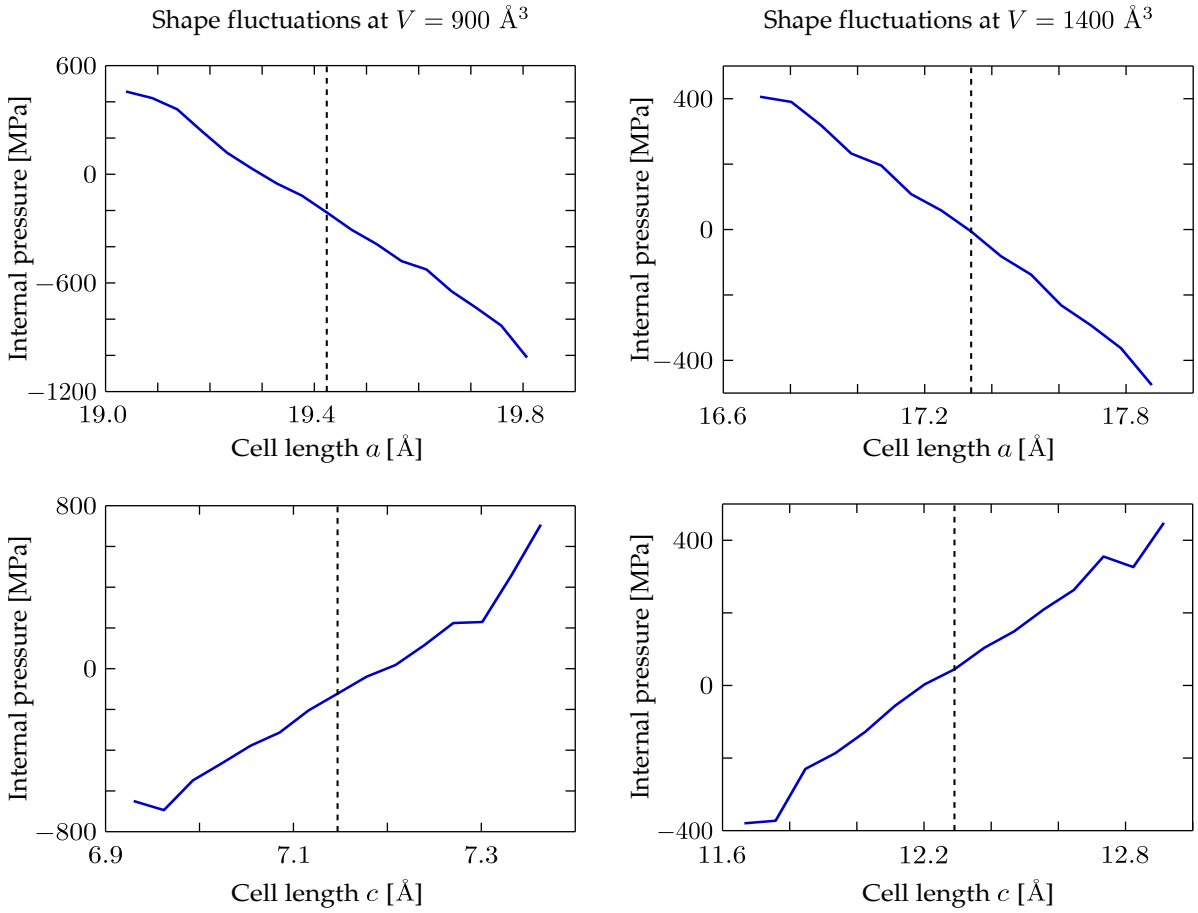


Figure S5: Internal pressure as a function of the cell lengths a (top row) and c (bottom row) for the $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations with the MTTK barostat corresponding to Figure S4. The results are obtained by discretising the cell lengths during the simulation, and averaging the internal pressure for each of the discretised intervals.

we will again focus on the two cell parameters a and c of MIL-53(Al), which can still fluctuate independently since only the volume is kept fixed. As shown in Figure S4, the fluctuations of these cell lengths in the $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ ensemble are largely underestimated when using a Berendsen barostat, while the MTTK and Langevin barostat yield coinciding distributions.

Based on this observation, it might seem quite surprising that the pressure profiles obtained with the three barostats coincide, since the Berendsen barostat does not sample the complete distribution of the cell parameters. However, as shown in Figure S5, the internal pressure is in good approximation a linear function of the cell lengths a and c. Hence, any symmetric distribution of the cell length PDFs will yield the same result for the average internal pressure, provided that the mean of the distribution is correct. This is indeed the case for Berendsen, as

can be seen in Figure S4, which explains the coinciding internal pressure profiles of Figure 6 and Figure 7 of the main paper. However, as stated in the main paper, any nonlinear function of these cell lengths, for instance including the variation in the cell lengths, will be able to distinguish between the Berendsen barostat on the one hand, and the MTTK and Langevin barostats on the other hand.

3.4 Free enthalpy profile of MIL-53(Al)

In Figure 6 of the main text, both the pressure and the free energy profile are shown as a function of the unit cell volume for MIL-53(Al). A completely similar discussion as in Section 4.2 for MIL-53(Al) can be carried out using the free enthalpy profile as a function of the unit cell volume. The free enthalpy G as a function of the unit cell volume V, defined as

$$G(P, T; V) = F(T; V) + PV, \tag{S3.2}$$

is the thermodynamic potential whose minima as a function of V indicate the stable and metastable states under isothermal-isobaric conditions at a temperature T and pressure P.

In Figure S6, different free enthalpy profiles are provided, corresponding to different values of the external pressure P and at 300 K. The global minimum of each of these profiles corresponds to the stable state at the given (P, T) conditions for MIL-53(Al). For some intermediate pressures, another local minimum appears, which corresponds to a metastable state at the given (P, T) conditions. For these profiles, also a local maximum arises between the local and the global minima, indicating a transition state, as will be explicitly shown with a committor analysis below.

From the free enthalpy profiles in Figure S6, we easily locate the stable state at each pressure, and in addition also the transition pressures are readily derived. We observe that for some pressures these free enthalpy profiles show multiple minima, while for others only one minimum is observed. In the case that two local minima arise, one corresponds with the stable state (red circle, global minimum) and the other with the metastable state (black circle, local minimum). Both the CP and the LP state can be stable or metastable, depending on the applied pressure. The LP-to-CP transition pressure corresponds to that pressure for which the LP metastable phase disappears, around 29 MPa. Likewise, the CP metastable phase disappears for pressure below

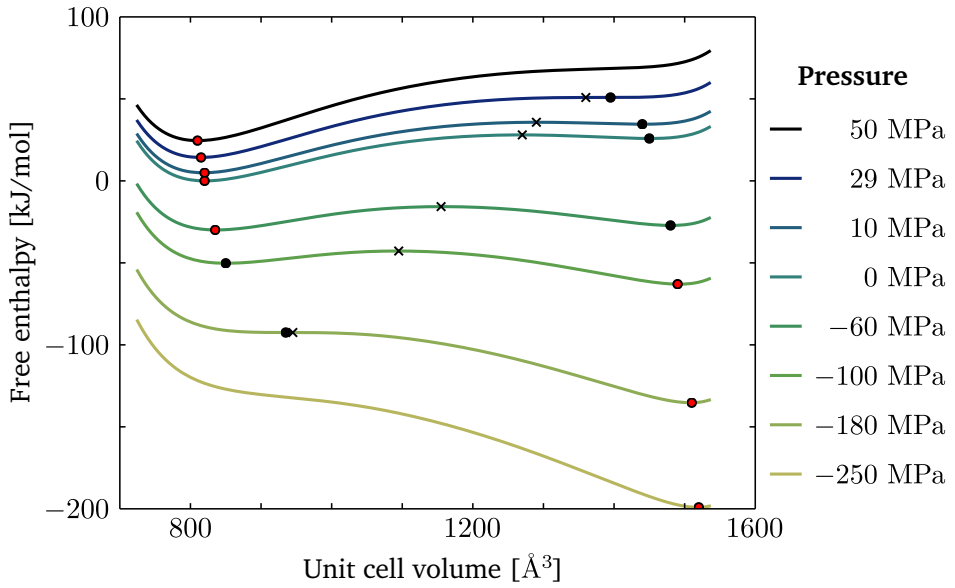


Figure S6: Free enthalpy G as a function of the constrained unit cell volume V for MIL-53(Al) at different pressures, resulting from $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations at $T = 300$ K. This plot was generated using the Langevin barostat with a relaxation time of 1 ps. Stable states (global minima) are indicated with filled red circles, metastable states (local minima) are indicated with black filled circles, and transition states (local maxima) are indicated with black crosses.

circa −180 MPa, indicating the CP-to-LP transition pressure.

For instance, at 300 K and starting from a pressure of −250 MPa, we see that the single minimum near the LP phase represents the global minimum. When increasing the pressure to −180 MPa, a second minimum arises near the CP state. However, this is only a metastable state at this pressure, which is separated from the global minimum by means of a local maximum: the transition state (black cross). When increasing the presure to about −60 MPa, the CP state becomes the global minimum and thus the stable state, while the LP state is now a local minimum. For higher pressures the LP minimum completely disappears, and only a stable CP minimum remains.

From this discussion, we can also determine the transition pressure, assuming that a transition is a collective motion, so that any nonzero barrier cannot be overcome.[8,9] Indeed, suppose we start in the LP phase at 0 MPa, and increase the pressure. As long as the pressure is sufficiently low, this state will be metastable, and will be retained if no steering force is present. Only when the pressure is increased above 29 MPa, the LP minimum disappears, and the structure will transform towards the CP state. Hence, this pressure corresponds to the LP-to-CP transition

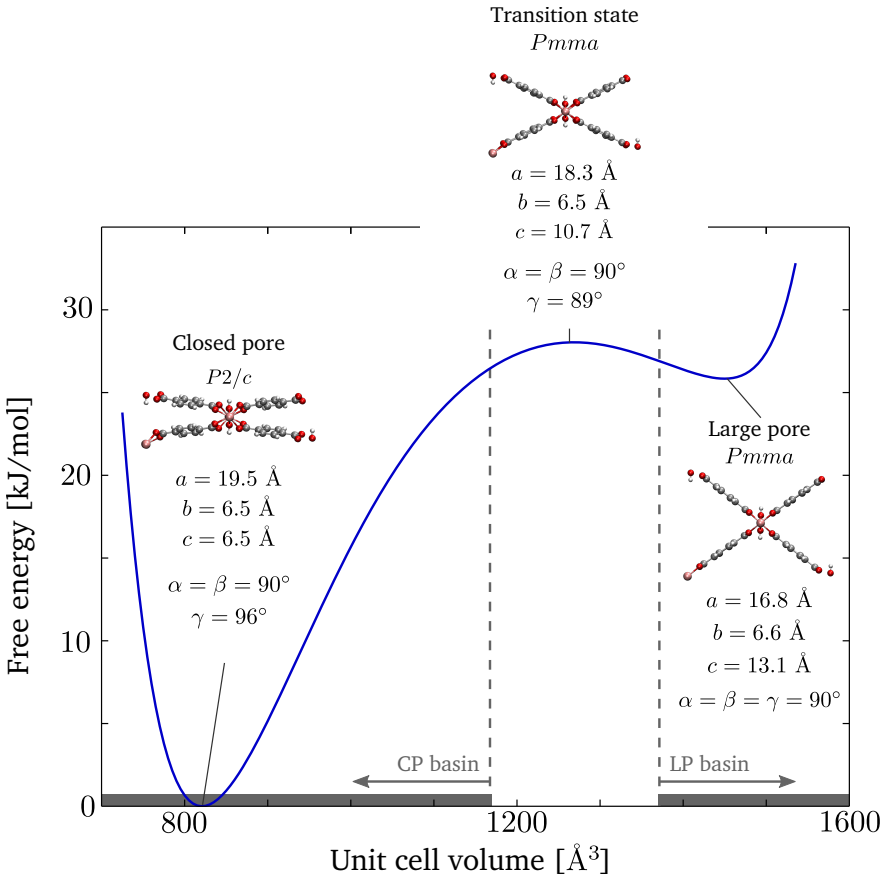


Figure S7: Overview of the stable CP state, the metastable LP state, and the transition state of MIL-53(Al): cell parameters and symmetry groups. All information is obtained from $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations with the Langevin barostat and a barostat relaxation time of 1 ps. The reported structures are averaged over the whole trajectory.

pressure. Likewise, when starting in the CP phase at about 0 MPa, and decreasing the pressure, the structure will only transform to the LP phase when the CP minimum disappears, around -180 MPa. Hence, $P_{\text{CP}\rightarrow\text{LP}} = -180$ MPa. These observations all agree with the discussion in Section 4.2 of the main text.

To further characterise the three states of interest, the stable, metastable and transition state, we will look at the free enthalpy profile as a function of the unit cell volume at 0 MPa (corresponding to the free energy), shown in Figure S7. Due to the interpolation scheme, the volumes V_{CP}, V_{tr} and V_{LP} reported in Table S6 differ slightly from those reported in Figure S7, which correspond to the values for which $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations were carried out. For each of these three simulations, the average structure throughout the simulation is determined. This leads to

the structures and cell parameters as shown in Figure S7. The symmetry group is determined using the PLATON software.[10]

To verify that the local maximum indeed corresponds to a transition state, a committor analysis is performed.[11,12] For this, 100 snapshots are taken from the $(N, V, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulation with volume closest to V_{tr}. For each of these snapshots, an $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulation is carried out at $P = 0$ MPa and $T = 300$ K, where the velocities are randomly distributed according to the Maxwell-Boltzmann distribution. The basins of the closed pore and large pore are chosen to be symmetric with respect to V_{tr}. The volume is regarded as the collective variable to determine if the structure belongs to the CP or LP basin:

$$\Omega_{\text{CP}} = \left\{ V | V < 1170 \text{ Å}^3 \right\} \qquad \Omega_{\text{LP}} = \left\{ V | V > 1370 \text{ Å}^3 \right\}. \tag{S3.3}$$

For 57 of the 100 simulations, the LP is first encountered before transforming to the CP. This corresponds to a committor value $p_{\text{LP}} = 0.57$, confirming that the found maximum is a transition state.

4 Dynamic determination of the transition pressure

In this section, the transition from the large-pore to the closed-pore structure in MIL-53(Al) is further investigated. The simulation results reported here are obtained from 100 $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations with $T = 300$ K and various pressures P ranging from 0.1 MPa to 1 GPa, using both the MTTK and Langevin barostat and different barostat relaxation times τ_P. We will propose two models, a Poisson and a Gaussian model, to explain the observed times to transitions. In the Poisson process, it is assumed that the LP-to-CP transition is stochastic and occurs very rarely. This corresponds to the situation where the applied pressure is low, so that the large energy barrier between the CP and LP states cannot be crossed based on this mean applied pressure. In the Gaussian process, however, the applied pressure steers the material over the barrier, so that the time to transition is a peaked distribution, with low variance due to the different initial configurations.

In the first two paragraphs, we will determine how we can best fit the data to a Poisson or Gaussian process, and how estimators for the parameters describing these distributions can be derived based on the principle of maximum entropy.[13,14] Afterwards, a criterion is established to determine whether the data can best be explained by a Gaussian or Poisson process. Finally, we report the simulated probability density functions for the MTTK barostat with $\tau_P = 1$ ps. By comparing these distributions with the best exponential fit to the data, we can visually confirm the validity of this criterion.

4.1 Derivation of the parameters of the Poisson process

A first possible distribution which can be used to model the time to transition is given by the exponential distribution $\mathcal{E}$. This distribution is completely determined by one parameter, the decay constant t_0. The variable X is exponentially distributed with parameter t_0 if

$$\text{prob}(X|t_0) = \mathcal{E}(X; t_0) = \frac{1}{t_0} \exp\left(-\frac{X}{t_0}\right) \tag{S4.1}$$

and $X \geq 0$. Here, we used the notation $\text{prob}(X|t_0)$ to denote the conditional probability of finding a time to transition equal to X given the exponential parameter t_0. This distribution

hence completely determines the probability density function for the time to transition, given this last parameter.

However, in this work, we do not a priori know this parameter t_0, and will need to estimate this parameter from the given sampling data $\boldsymbol{X} = \{X_1, ..., X_N\}$ with sample size N. In accordance with the principle of maximum entropy, we will estimate this parameter as the value of t_0 which maximises the conditional probability $\text{prob}(t_0|\boldsymbol{X}, I)$, where I contains all available background information—for instance earlier experiments. This probability is proportional to $\text{prob}(\boldsymbol{X}|t_0, I)$ through Bayes' theorem:

$$\text{prob}(t_0|\boldsymbol{X}, I) = \frac{\text{prob}(\boldsymbol{X}|t_0, I)\ \text{prob}(t_0|I)}{\text{prob}(\boldsymbol{X}|I)} \propto \text{prob}(\boldsymbol{X}|t_0, I)\ \text{prob}(t_0|I) \tag{S4.2}$$

While the proportionality constant $\text{prob}(\boldsymbol{X}|I)$, which does not depend on t_0, is important for normalization, we will omit it here, since we are only interested in finding the value of t_0 corresponding to the maximum probability. The factor $\text{prob}(t_0|I)$, the so-called prior, expresses how we expect the parameter t_0 to be distributed given the background information I. For instance, it seems straightforward to assume that the time to transition should be nonnegative, and hence also $t_0 > 0$. However, to avoid any bias, we will assume that this prior is a flat distribution. As such, this factor can be adsorbed as a constant of proportionality, yielding

$$\text{prob}(t_0|\boldsymbol{X}, I) \propto \text{prob}(\boldsymbol{X}|t_0, I) \tag{S4.3}$$

If each of the data points contained in $\boldsymbol{X}$ are independent, this latter probability can be expressed as a product of the contributing partial probabilities, which we assumed to be exponentially distributed:

$$\text{prob}(t_0|\boldsymbol{X}, I) \propto \prod_{k=1}^{N} \text{prob}(X_k|t_0, I) = \prod_{k=1}^{N} \mathcal{E}(X_k; t_0) \tag{S4.4}$$

After some algebraic manipulations, we arrive at a workable expression for the distribution

of t_0:

$$\begin{aligned} \text{prob}(t_0|\boldsymbol{X}, I) &\propto \prod_{k=1}^{N} \frac{1}{t_0} \exp\left(-\frac{X_k}{t_0}\right) && \text{(S4.5)} \\ &\propto \frac{1}{t_0^N} \exp\left(-\frac{\sum_{k=1}^{N} X_k}{t_0}\right) && \text{(S4.6)} \\ &\propto \frac{1}{t_0^N} \exp\left(-\frac{N\overline{\boldsymbol{X}}}{t_0}\right) && \text{(S4.7)} \end{aligned}$$

where $\overline{\boldsymbol{X}}$ denotes averaging over the sample size:

$$\overline{\boldsymbol{X}} = \frac{1}{N} \sum_{k=1}^{N} X_k \tag{S4.8}$$

To find the maximum of this probability distribution, it is easier to work with the logarithm of this expression:

$$L = \ln\{\text{prob}(t_0|\boldsymbol{X}, I)\} = c - N \ln t_0 - \frac{N\overline{\boldsymbol{X}}}{t_0} \tag{S4.9}$$

where c expresses that this probability is only known up to a constant of proportionality. The maximum of this distribution is then found by differentiating this expression with respect to t_0, the parameter we are trying to optimise:

$$\left.\frac{\mathrm{d}L}{\mathrm{d}t_0}\right|_{t_0^*} = -\frac{N}{t_0^*} + \frac{N\overline{\boldsymbol{X}}}{t_0^{*2}} = 0 \Rightarrow t_0^* = \overline{\boldsymbol{X}} \tag{S4.10}$$

Hence, the best estimator for t_0 is given by the sample mean $\overline{\boldsymbol{X}}$, which is independent of the sample size. This value indeed corresponds to the maximum in the distribution, since

$$\left.\frac{\mathrm{d}^2L}{\mathrm{d}t_0^2}\right|_{t_0^*} = \frac{N}{t_0^{*2}} - \frac{2N\overline{\boldsymbol{X}}}{t_0^{*3}} = -\frac{N}{\overline{\boldsymbol{X}}^2} < 0 \tag{S4.11}$$

Moreover, the standard deviation on this estimator can be derived from this second order derivative:[15]

$$\sigma_{t_0^*}^2 = \left[-\left.\frac{\mathrm{d}^2L}{\mathrm{d}t_0^2}\right|_{t_0^*}\right]^{-1} = \frac{\overline{\boldsymbol{X}}^2}{N} \tag{S4.12}$$

Hence, we see that the 1σ confidence interval for the best t_0^*, the estimator of t_0, is given by

$$t_0^* = \overline{\boldsymbol{X}} \pm \frac{\overline{\boldsymbol{X}}}{\sqrt{N}} \tag{S4.13}$$

and hence becomes narrower the higher the sample size N, as could be expected.

4.2 Derivation of the parameters of the Gaussian process

The second model for the time to transition is given by the Gaussian or normal distribution $\mathcal{N}$, characterised by a mean μ and standard deviation σ:

$$\text{prob}(X|\mu,\sigma) = \mathcal{N}(X;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right) \tag{S4.14}$$

As in the previous section, we are interested in the best estimate for the parameters μ and σ given the data $\boldsymbol{X}$. Assuming that these two parameters are independent, we can tackle this problem sequentially, starting with μ. Hence, we are interested in the marginal distribution, by integrating out σ:

$$\text{prob}(\mu|\boldsymbol{X},I) = \int_0^\infty \text{prob}(\mu,\sigma|\boldsymbol{X},I)\mathrm{d}\sigma \propto \int_0^\infty \text{prob}(\boldsymbol{X}|\mu,\sigma,I)\mathrm{d}\sigma \tag{S4.15}$$

For this expression, we again applied Bayes' theorem and assumed a flat prior $\text{prob}(\mu,\sigma|I)$. Assuming that the different samples are independent, we arrive at the counterpart of Eq. (S4.4):

$$\text{prob}(\mu|\boldsymbol{X},I) \propto \int_0^\infty \prod_{k=1}^N \text{prob}(X_k|\mu,\sigma,I)\mathrm{d}\sigma = \int_0^\infty \prod_{k=1}^N \mathcal{N}(X_k;\mu,\sigma)\mathrm{d}\sigma \tag{S4.16}$$

Inserting the analytical expression for the normal distribution then yields:

$$\begin{aligned}
\text{prob}(\mu|\boldsymbol{X},I) &\propto \int_0^\infty \prod_{k=1}^N \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(X_k-\mu)^2}{2\sigma^2}\right)\mathrm{d}\sigma & \text{(S4.17)}\\
&\propto \frac{1}{(2\pi)^{N/2}}\int_0^\infty \frac{1}{\sigma^N}\exp\left(-\frac{\sum_{k=1}^N (X_k-\mu)^2}{2\sigma^2}\right)\mathrm{d}\sigma & \text{(S4.18)}
\end{aligned}$$

Adsorbing the constant of proportionality, this expression reduces to

$$\begin{aligned}
\mathrm{prob}(\mu|\boldsymbol{X}, I) &\propto \int_0^\infty w^{N-2} \exp\left(-\frac{w^2}{2}\sum_{k=1}^{N}(X_k-\mu)^2\right) \mathrm{d}w && \text{(S4.19)}\\
&\propto \int_0^\infty w^{N-2} \exp\left(-w^2\frac{N}{2}(\overline{\boldsymbol{X}^2}-2\mu\overline{\boldsymbol{X}}+\mu^2)\right) \mathrm{d}w && \text{(S4.20)}\\
&\propto \left[\frac{2}{N(\overline{\boldsymbol{X}^2}-2\mu\overline{\boldsymbol{X}}+\mu^2)}\right]^{(N-1)/2} \int_0^\infty z^{N-2}\exp(-z^2)\mathrm{d}z && \text{(S4.21)}\\
&\propto \left[\frac{1}{\overline{\boldsymbol{X}^2}-2\mu\overline{\boldsymbol{X}}+\mu^2}\right]^{(N-1)/2} && \text{(S4.22)}
\end{aligned}$$

In this derivation, we adsorbed all constants of proportionality, in particular the last integral which does not depend on the mean, since these constants only result in a global scaling of the probability distribution. Furthermore, we used the transformations $w = \sigma^{-1}$ and

$$z = w\sqrt{\frac{N}{2}(\overline{\boldsymbol{X}^2}-2\mu\overline{\boldsymbol{X}}+\mu^2)} \tag{S4.23}$$

Again considering the logarithm of this expression, we arrive at

$$L = \ln\{\mathrm{prob}(\mu|\boldsymbol{X}, I)\} = c - \frac{N-1}{2}\ln\left[\overline{\boldsymbol{X}^2}-2\mu\overline{\boldsymbol{X}}+\mu^2\right] \tag{S4.24}$$

The derivative of this expression yields:

$$\left.\frac{\mathrm{d}L}{\mathrm{d}\mu}\right|_{\mu^*} = -(N-1)\frac{\mu^*-\overline{\boldsymbol{X}}}{\overline{\boldsymbol{X}^2}-2\mu^*\overline{\boldsymbol{X}}+\mu^{*2}} = 0 \Rightarrow \mu^* = \overline{\boldsymbol{X}} \tag{S4.25}$$

Hence, the best estimator for the mean of the normal distribution, μ, is again given by the sample mean $\overline{\boldsymbol{X}}$, as long as the denominator is nonzero. This restriction is always fulfilled if $\overline{\boldsymbol{X}} \neq 0$, since

$$\overline{\boldsymbol{X}^2}-2\mu^*\overline{\boldsymbol{X}}+\mu^{*2} \geq \overline{\boldsymbol{X}}^2-2\mu^*\overline{\boldsymbol{X}}+\mu^{*2} = (\overline{\boldsymbol{X}}-\mu^*)^2 \geq 0 \tag{S4.26}$$

If however $\overline{\boldsymbol{X}} = 0$, Eq. (S4.24) reduces to $L = c - (N-1)\ln(\overline{\boldsymbol{X}^2}+\mu^2)/2$, with a maximum for $\mu^* = 0 = \overline{\boldsymbol{X}}$.

To estimate the error on this estimator, we need the second order derivative:

$$\left.\frac{\mathrm{d}^2 L}{\mathrm{d}\mu^2}\right|_{\mu^*} = -(N-1)\frac{\overline{\boldsymbol{X}^2} - 2\mu^*\overline{\boldsymbol{X}} + \mu^{*2} - 2(\mu^* - \overline{\boldsymbol{X}})^2}{\left(\overline{\boldsymbol{X}^2} - 2\mu^*\overline{\boldsymbol{X}} + \mu^{*2}\right)^2} \tag{S4.27}$$

$$= -(N-1)\frac{\overline{\boldsymbol{X}^2} - \overline{\boldsymbol{X}}^2}{\left(\overline{\boldsymbol{X}^2} - \overline{\boldsymbol{X}}^2\right)^2} \tag{S4.28}$$

$$= -\frac{N-1}{s^2_{\boldsymbol{X}}} \tag{S4.29}$$

where we introduced the sample variance

$$s^2_{\boldsymbol{X}} = \overline{\boldsymbol{X}^2} - \overline{\boldsymbol{X}}^2 = \frac{1}{N}\sum_{k=1}^{N} X_k^2 - \left[\frac{1}{N}\sum_{k=1}^{N} X_k\right]^2 \tag{S4.30}$$

This yields:

$$\sigma^2_{\mu^*} = \left[-\left.\frac{\mathrm{d}^2 L}{\mathrm{d}\mu^2}\right|_{\mu^*}\right]^{-1} = \frac{s^2_{\boldsymbol{X}}}{N-1} \tag{S4.31}$$

Hence, the 1σ interval for estimator of μ is given by

$$\mu^* = \overline{\boldsymbol{X}} \pm \sqrt{\frac{s^2_{\boldsymbol{X}}}{N-1}} \tag{S4.32}$$

and can again be narrowed down by increasing the sample size N.

In an analogous fashion an estimator for σ can be considered. With the same manipulations as used to arrive at Eq. (S4.18), we find

$$\mathrm{prob}(\sigma|\boldsymbol{X}, I) = \int_{-\infty}^{\infty} \mathrm{prob}(\mu, \sigma|\boldsymbol{X}, I)\mathrm{d}\mu \tag{S4.33}$$

$$\propto \int_{-\infty}^{\infty} \mathrm{prob}(\boldsymbol{X}|\mu, \sigma, I)\mathrm{d}\mu \tag{S4.34}$$

$$\propto \frac{1}{\sigma^N}\int_{-\infty}^{\infty} \exp\left(-\frac{\sum_{k=1}^{N}(X_k - \mu)^2}{2\sigma^2}\right)\mathrm{d}\mu \tag{S4.35}$$

$$\propto \frac{1}{\sigma^N}\int_{-\infty}^{\infty} \exp\left(-\frac{N(\overline{\boldsymbol{X}^2} - 2\mu\overline{\boldsymbol{X}} + \mu^2)}{2\sigma^2}\right)\mathrm{d}\mu \tag{S4.36}$$

$$\propto \frac{1}{\sigma^N}\exp\left(-\frac{N}{2}\frac{s^2_{\boldsymbol{X}}}{\sigma^2}\right)\int_{-\infty}^{\infty} \exp\left(-\frac{N(\overline{\boldsymbol{X}} - \mu)^2}{2\sigma^2}\right)\mathrm{d}\mu \tag{S4.37}$$

Carrying out the remaining integration, we find

$$\text{prob}(\sigma|\boldsymbol{X}, I) \propto \frac{1}{\sigma^{N-1}} \exp\left(-\frac{N}{2}\frac{s_{\boldsymbol{X}}^2}{\sigma^2}\right) \tag{S4.38}$$

In this derivation, we used the identity $\overline{\boldsymbol{X}^2} = \overline{\boldsymbol{X}}^2 + s_{\boldsymbol{X}}^2$ to introduce the sample variance. The logarithm of Eq. (S4.38) is then given by

$$L = \ln\{\text{prob}(\sigma|\boldsymbol{X}, I)\} = c - (N-1)\ln\sigma - \frac{N}{2}\frac{s_{\boldsymbol{X}}^2}{\sigma^2} \tag{S4.39}$$

with derivative

$$\left.\frac{\mathrm{d}L}{\mathrm{d}\sigma}\right|_{\sigma^*} = -\frac{N-1}{\sigma^*} + N\frac{s_{\boldsymbol{X}}^2}{\sigma^{*3}} = 0 \Rightarrow \sigma^* = \sqrt{\frac{N}{N-1}s_{\boldsymbol{X}}^2} \tag{S4.40}$$

For the error on this estimate, we consider the second order derivative:

$$\left.\frac{\mathrm{d}^2L}{\mathrm{d}\sigma^2}\right|_{\sigma^*} = \frac{N-1}{\sigma^{*2}} - 3N\frac{s_{\boldsymbol{X}}^2}{\sigma^{*4}} = -2\frac{(N-1)^2}{N}\frac{1}{s_{\boldsymbol{X}}^2} \tag{S4.41}$$

Hence:

$$\sigma_{\sigma^*}^2 = \left[-\left.\frac{\mathrm{d}^2L}{\mathrm{d}\sigma^2}\right|_{\sigma^*}\right]^{-1} = \frac{N}{2(N-1)^2}s_{\boldsymbol{X}}^2 \tag{S4.42}$$

and the 1σ confidence interval is given by

$$\sigma^* = \sqrt{\frac{N}{N-1}s_{\boldsymbol{X}}^2} \pm \sqrt{\frac{N}{2(N-1)^2}s_{\boldsymbol{X}}^2} \tag{S4.43}$$

4.3 Choosing the appropriate model

In the previous paragraphs, we derived estimators for each of the possible models describing the time to transition. However, we did not yet mention how we can determine which of these models best explains the simulated data. To do so, let G denote the hypothesis that the data can be explained by a Gaussian, and P the hypothesis that the data supports the Poisson process. We will investigate the ratio R of the two conditional probabilities

$$R = \frac{\text{prob}(G, \mu^*, \sigma^*|\boldsymbol{X}, I)}{\text{prob}(P, t_0^*|\boldsymbol{X}, I)} \tag{S4.44}$$

Using Bayes' theorem, this ratio can be further expressed as:[15]

$$R = \frac{\text{prob}(G|I)}{\text{prob}(P|I)} \times \frac{\text{prob}(\boldsymbol{X}|G, \mu^*, \sigma^*, I)}{\text{prob}(\boldsymbol{X}|P, t_0^*, I)} \times \frac{\sigma_{\sigma^*}}{\Delta\sigma^*} \frac{\sigma_{\mu^*}}{\Delta\mu^*} \frac{\Delta t_0^*}{\sigma_{t_0^*}} \tag{S4.45}$$

This ratio is thus the product of three factors. The first factor expresses the preference for one hypothesis over the other, given only the background information I. Since we want to focus on the most probable model given the data, without interference of the a priori known distribution, we will set this factor equal to one. The second factor, expressing that the data can be explained by a Gaussian, respectively a Poisson process with the best estimate for its parameters, was already encountered in the previous paragraphs, and can readily be inserted. Finally, the last factor, called the Ockham parameter, corrects for the fact that the Gaussian model has two parameters, whereas the Poisson process is described by only one parameter. Since we expect that increasing the number of parameters will improve the goodness of fit, this extra parameter should be penalised to avoid a model with a vast amount of parameters. The factors σ_{x^*} express the uncertainty on the estimator x^*, and were calculated before. The factors $\Delta x^* = x^*_{\text{max}} - x^*_{\text{min}}$ contain a priori information on the range in which we expect to find the best estimator x^* for x. To simplify this Ockham parameter, we will first assume that $\Delta\sigma^* \cong \sigma_{\sigma^*}$, hence that we have a good idea about the width of the Gaussian distribution. While this assumption can be relieved, it is not expected that this factor becomes very small, and will hence not largely influence the final result. Furthermore, since t_0^* and μ^* are both estimated by the mean of the sample data, we can expect that our a priori information about both is the same: $\Delta t_0^* = \Delta\mu^*$. Inserting Eq. (S4.12) for $\sigma_{t_0^*}$ and Eq. (S4.31) for σ_{μ^*} then yields:

$$R = \frac{\prod_{k=1}^{N} \text{prob}(X_k, G, \mu^*, \sigma^*, I)}{\prod_{k=1}^{N} \text{prob}(X_k, P, t_0^*, I)} \sqrt{\frac{s^2_{\boldsymbol{X}}}{N-1}} \frac{\sqrt{N}}{\overline{\boldsymbol{X}}} \tag{S4.46}$$

Since we know the Gaussian and Poisson distribution with the optimal estimators as calcu-

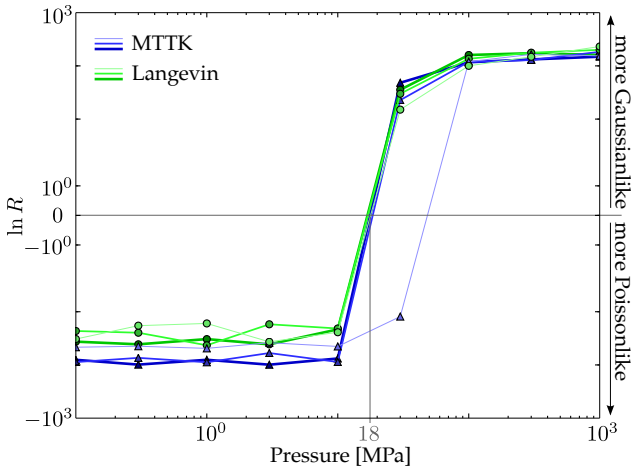


Figure S8: Logarithm of the ratio as expressed in Eq. (S4.51) with $N = 100$ as a function of the pressure for different barostat and barostat relaxation times. The more negative the value, the more the LP-to-CP transition can be seen as a Poisson process, while positive values indicate a Gaussian process. The darker the color, the higher the barostat relaxation time, chosen from 1, 5 and 10 ps.

lated in the previous section, this can be simplified to:

$$R = \sqrt{\frac{N}{N-1}}\frac{s_{\boldsymbol{X}}}{\overline{\boldsymbol{X}}}\frac{\prod_{k=1}^{N}\frac{1}{\sqrt{2\pi}\sigma^*}\exp\left(-\frac{(X_k-\mu^*)^2}{2\sigma^{*2}}\right)}{\prod_{k=1}^{N}\frac{1}{t_0^*}\exp\left(-\frac{X_k}{t_0^*}\right)} \tag{S4.47}$$

$$= \frac{1}{(2\pi)^{N/2}}\sqrt{\frac{N}{N-1}}\frac{s_{\boldsymbol{X}}}{\overline{\boldsymbol{X}}}\left(\frac{t_0^*}{\sigma^*}\right)^N\frac{\exp\left(-\frac{1}{2\sigma^{*2}}\sum_{k=1}^{N}(X_k-\mu^*)^2\right)}{\exp\left(-\frac{1}{t_0^*}\sum_{k=1}^{N}X_k\right)} \tag{S4.48}$$

$$= \frac{1}{(2\pi)^{N/2}}\sqrt{\frac{N}{N-1}}\frac{s_{\boldsymbol{X}}}{\overline{\boldsymbol{X}}}\left(\frac{t_0^*}{\sigma^*}\right)^N\frac{\exp\left(-\frac{N(\overline{\boldsymbol{X}^2}-2\mu^*\overline{\boldsymbol{X}}+\mu^{*2})}{2\sigma^{*2}}\right)}{\exp\left(-\frac{N\overline{\boldsymbol{X}}}{t_0^*}\right)} \tag{S4.49}$$

$$= \frac{1}{(2\pi)^{N/2}}\left(\frac{N-1}{N}\right)^{(N-1)/2}\left(\frac{\overline{\boldsymbol{X}}^2}{s_{\boldsymbol{X}}^2}\right)^{(N-1)/2}\exp\left(\frac{N+1}{2}\right) \tag{S4.50}$$

This ratio can vary hugely. Hence, it is more appropriate to express the logarithm of this ratio:

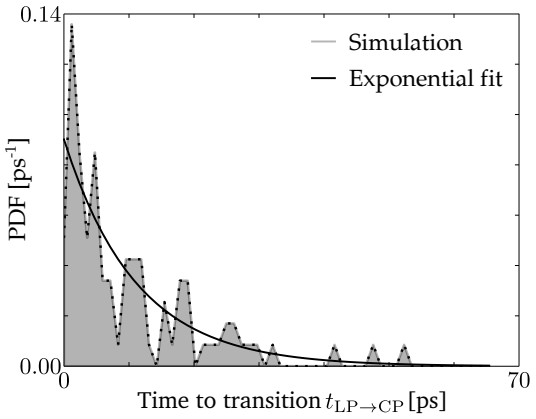
$$\ln R = -\frac{N}{2}\ln(2\pi) + \frac{N-1}{2}\left[\ln(N-1) - \ln(N) + 2\ln\overline{\boldsymbol{X}} - \ln s_{\boldsymbol{X}}^2\right] + \frac{N+1}{2} \tag{S4.51}$$

In Figure S8, this logarithm is shown as a function of the applied pressure P for the different barostat and barostat relaxation times of Figure 8 of the main paper. We observe a clear switch

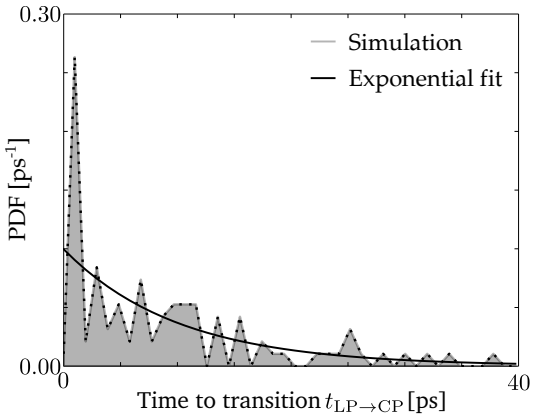
between the Poissonlike process up to 10 MPa, and the Gaussianlike process as of 30 MPa—only the MTTK simulation with a barostat time constant of 1 ps deviates from this result. Note that, to obtain the ratio R, we should exponentiate this result, indicating that adapting our assumptions regarding some of the factors in Eq. (S4.45)—the Ockham factor and the a priori ratio—will not largely influence the result. If we define the transition pressure as the pressure needed to switch the LP-to-CP transition from a rare event (Poissonlike) to a normal event (Gaussianlike), it is observed from Figure S8 that this pressure is found in the range $[10\,\text{MPa}, 30\,\text{MPa}]$. Both processes are equally as probable to explain the given data for a pressure of about 18 MPa, in accordance with previous results in this paper and elsewhere.[16] As a last remark, note that while the Langevin barostat for small pressures indicates a Poisson process, this tendency is less pronounced than when using the MTTK barostat. This could be due to the stochastic nature of the Langevin barostat, in which one of the terms in the equation of motion for the cell momentum tensor $\mathbf{p}_g$ is normally distributed.

4.4 Overview of the distributions

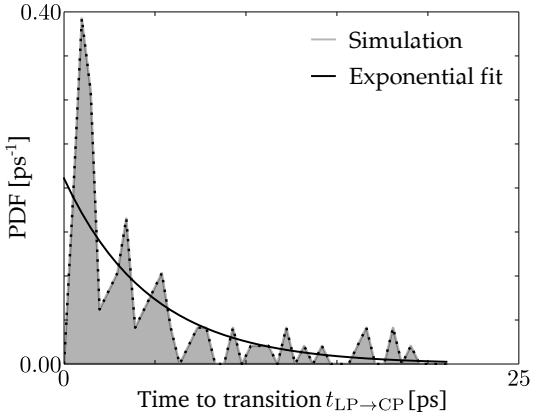
In Figure S9, the probability density functions for the time to transition simulated with the MTTK barostat with $\tau_P = 1$ ps and at various pressures are reported. The distribution for pressures of 1 MPa and 1 GPa are also shown in Figure 8 of the main paper. Each of the figures is complemented by the analytical exponential distribution, for which the parameter t_0^* is taken as the best estimator for this parameter as defined above. These distributions visually confirm the conclusion expressed in Figure S8: up to (and, for this combination of barostat and relaxation time, including) 30 MPa, the given data can be well fitted to an exponential function with the given estimator t_0^*, while this is no longer the case for higher pressures.



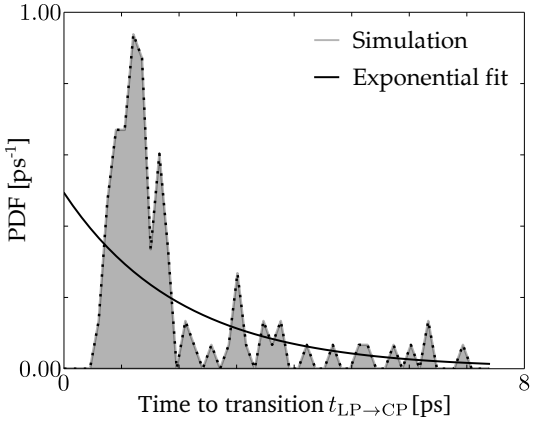
(a) 1 MPa



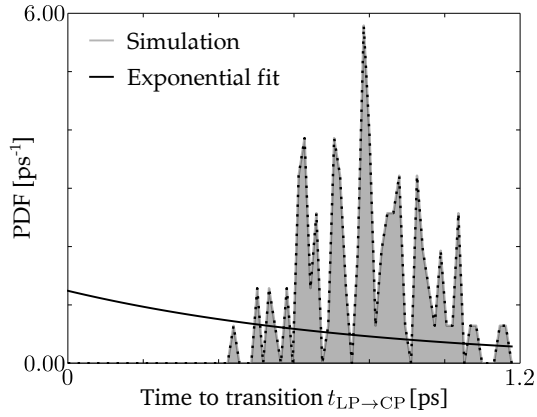
(b) 3 MPa



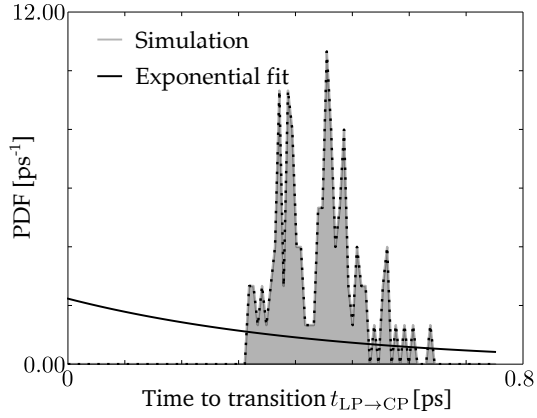
(c) 10 MPa



(d) 30 MPa



(e) 100 MPa



(f) 300 MPa

Figure S9: Probability density function (PDF) of the time to transition from 100 $(N, P, \boldsymbol{\sigma}_a = \mathbf{0}, T)$ simulations at $T = 300$ K for MIL-53(Al) and different pressures applied with the MTTK barostat with a relaxation time of 1 ps.

References

(1) Parrinello, M.; Rahman, A. *Phys. Rev. Lett.* **1980**, *45*, 1196–1199.

(2) Martyna, G. J.; Klein, M. L.; Tuckerman, M. E. *J. Chem. Phys.* **1992**, *97*, 2635–2643.

(3) McEvoy, P. *Classical Theory*; MicroAnalytix: San Diego, 2002; pp 232–234.

(4) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(5) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117–1157.

(6) Efron, B. *Ann. Statist.* **1979**, *7*, 1–26.

(7) Chernick, M. R. *Bootstrap Methods: A Guide for Practitioners and Researchers*, 2nd ed.; Wiley: Hoboken, 2008; pp 1–77.

(8) Ghysels, A.; Vanduyfhuys, L.; Vandichel, M.; Waroquier, M.; Van Speybroeck, V.; Smit, B. *J. Phys. Chem. C* **2013**, *117*, 11540–11554.

(9) Vanduyfhuys, L.; Ghysels, A.; Rogge, S. M. J.; Demuynck, R.; Van Speybroeck, V. *Mol. Simul.* **2015**, *41*, 1311–1328.

(10) Spek, A. L. *J. Appl. Cryst.* **2003**, *36*, 7–13.

(11) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.

(12) E, W.; Vanden-Eijnden, E. *J. Stat. Phys.* **2006**, *123*, 503–523.

(13) Jaynes, E. T. *Phys. Rev.* **1957**, *106*, 620–630.

(14) Jaynes, E. T. *Phys. Rev.* **1957**, *108*, 171–190.

(15) Sivia, D. S.; Skilling, J. *Data Analysis: A Bayesian Tutorial*; Oxford University Press: Oxford, 2006; pp 3–126.

(16) Yot, P. G.; Boudene, Z.; Macia, J.; Granier, D.; Vanduyfhuys, L.; Verstraelen, T.; Van Speybroeck, V.; Devic, T.; Serre, C.; Férey, G.; Stock, N.; Maurin, G. *Chem. Commun.* **2014**, *50*, 9462–9464.