In silico design and assembly of cage molecules into porous molecular materials.

Marco Bernabei,*[a] Raul Pérez Soto ,[a] Ismael Gómez García [a] and Maciej Haranczyk [a]

[a] IMDEA Materials Institute, C/Eric Kandel 2, 28906 - Getafe, Madrid, Spain

Supporting Information

Content

S1. Porosity characterization of the lowest energy conformers of molecules M1 -M6.

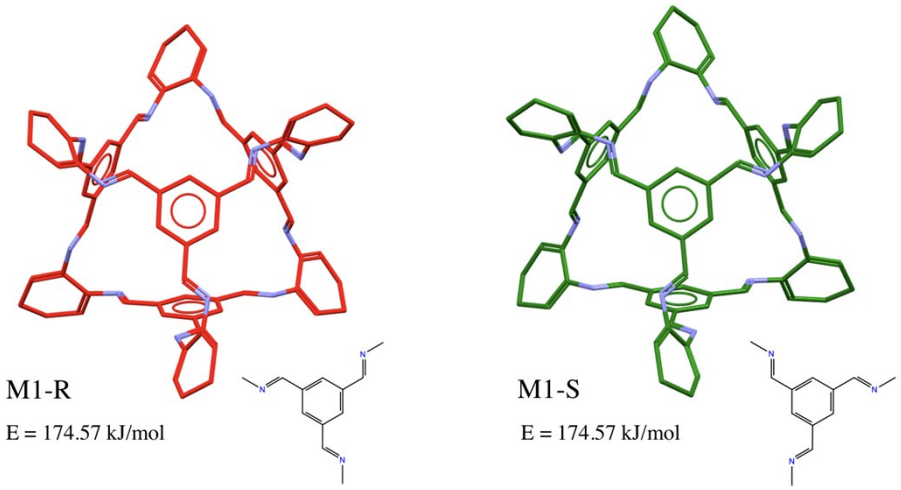

Figure S1. Lowest energy conformers of cage M1. Both enantiomers M1-R and M1-S have the same energy and tetrahedral symmetry. Nitrogen atoms are colored in blue, carbon atoms in red and green for M1-R and M1-S respectively and hydrogen atoms are omitted for clarity

	LCD (Å) DFT-M06-2X	Window size (Å) DFT-M06-2X
CC3-R	5.50	3.77
M1-R	4.94	3.20
M2-R	5.31	3.52
M3-R	5.36	3.60
M4-R	4.97	3.24
M5-R	4.78	3.06
M6-R	5.21	3.48

Table S1. Largest cavity diameters and window size for the lowest energy R-conformers of cage CC3 and molecules M1-M6 after DFT optimizations.

S2. CSP validation: the benchmark covalent cage CC3.

Cage CC3 is probably the most studied cage in the imine cage family with a LCD of 5.50 Å and four windows of equal size of 3.75 Å. The enantiomer CC3-R was synthesized[1] from a mixture of 1,3,5-triformylbenzene and chiral (*R,R*)-1,2-diaminocyclohexane (see Figure 1 in the manuscript for a schematic representation). Similarly, the energetically equivalent enantiomer CC3-S can be made by a mixture of 1,3,5-triformylbenzene and chiral (*S,S*)-1,2-diaminocyclohexane. Applications of porous materials based on cage CC3 in its homochiral polymorph CC3-R(S)-α, are for example: Xenon/Krypton separation[2], enantiomeric and molecular size separation[3], water desalination[4].

A 3D model of CC3 molecule was initially generated with LigPrep starting from a 1D SMILES corresponding to the enantiomer CC3-R(S) and submitted to a conformational search procedure.

The experimentally observed enantiomer CC3-R(S) was correctly predicted as the lowest energy conformer and further optimized in vacuum at the M06-2X/6-311G** level of theory.

The DFT optimized molecular geometry was used as input for the CSP study. Homochiral crystal phases based on CC3-R were searched in the most common 6 space groups for enantiopure crystal structures ($P1$, $P2_1$, $P2_12_12$, $P2_12_12_1$,$C222_1$,$C2$). Similarly, racemic co-crystals CC3-(R,S) were generated in the most common 6 space groups for racemates ($P2_1/c$, $P\bar{1}$,$C2/c$,Cc,$Pna2_1$,$Pbca$). More details on the CSP procedure are provided in the method section in the main text.

The resulting energy landscape, i.e. the plot of relative lattice energy of CC3 as a function of density, is shown in Figure S2.

The energy landscape of CC3 has been already investigated by Day's group employing a Monte Carlo simulated annealing for initial structure generation and a subsequent final minimization procedure was carried out with the DMACRYS software for rigid molecules, which make use of anisotropic atom-atom model potentials.[5]

Our results are in quantitative agreement with the results of Day's group even using an isotropic Coulomb potential for the electrostatic interactions. Only 12 structures are found in a 50 kJ/mol range from the global minimum, and five of them, color-filled circles in Figure S2, feature a tetrahedral porous network connecting the cavities of the cages.

The global minimum, black filled point in Figure S2, corresponds to the racemic co-crystal cubic phase CC3-(R,S)α. The latter is not available in the Cambridge structure database since the structure and unit cell parameters were refined from PXRD experiments[6] therefore, a direct geometric comparison with the predicted phases is not possible.

The homochiral phase CC3-Rα, yellow point in Figure S2, is correctly predicted as the global minimum of the structures generated in the six enantiopure space groups. The latter, initially predicted in the space group $P2_12_12_1$ shows the full experimental symmetry $F4_132$ and is a good geometrical match to the experimental structure PUDXES in Table 1.

Due to the large size of the imine cages our approach is limited to generating structures with Z'=1 molecules in the asymmetric unit, therefore the polymorph PUDXES02 of CC3, also known as CC3-Rβ phase (see Table 1) was not predicted since it was experimentally observed in the space group $P3$ with Z'=3.

Although our CSP approach does not account for solvent inclusion, the structures of solvated polymorphs can still be predicted as local minima in the energy landscape even if they incur structural rearrangement upon solvent removal.
The solvated framework of PUDXES02, i.e. CC3-Rβ-solvated with CSD-Refcode NODVIN, was experimentally observed in space group *R*3 with Z'=1 cage in the asymmetric unit.
The structure corresponding to the CC3-Rβ-solvated framework (not shown in Figure S2) showed indeed the correct full *R*3 symmetry and was predicted to be a high energy structure found approximately 50 kJ/mol above the CC3-Rα phase.
Structure overlay with experimental structures and comparison of the cell parameters for the predicted phases CC3-Rα and CC3-Rβ-solvated are presented in Figure S3 and Table S2, respectively.

Finally, to validate the accuracy of our energy model in predicting the landscape of porous organic cages we made use of DFT-D3 solid state periodic calculations to fully relax the lowest energy structures of the landscape. Results for cage CC3 are listed in Table S3. Although the relative energy of the phase CC3-Rα increase of about 50 % after DFT-D3 relaxation, the global energy ranking is preserved and porosity, characterized by the PLD, is maintained.

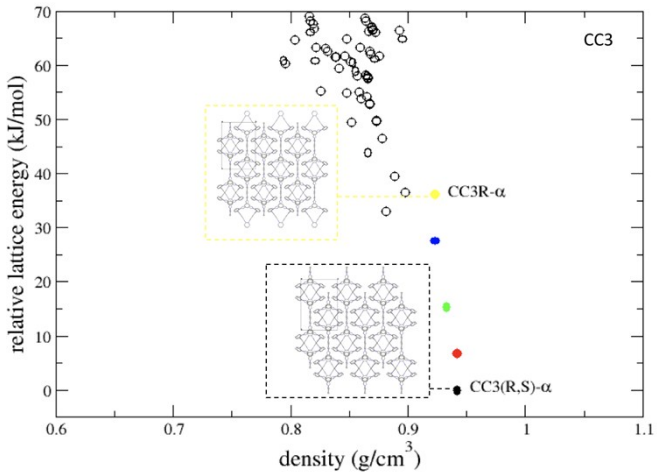

Figure S2. Crystal energy landscape of CC3. Energies are w.r.t the global minimum. Both homochiral and racemic polymorphs are correctly predicted as the lowest energy structures in the enantiopure and racemic space groups, respectively.

		Cell axes length (Å)			Cell angles			Density (g/cm^3)
Material	Method	a	b	c	α	β	γ	
CC3-Rα (*F*4$_1$32)	Exp.	24.80	24.80	24.80	90.00	90.00	90.00	0.97
	CSP	25.24	25.24	25.24	90.00	90.00	90.00	0.92
CC3-Rβ-solv (*R*3)	Exp.	25.63	25.63	11.18	90.00	90.00	120.00	0.87
	CSP	26.36	26.35	11.39	90.03	90.01	120.00	0.81

Table S2. Cell parameters from our CSP study compared to the experimental values for CC3-Rα (CSD-Refcode PUDXES) and CC3-Rβ-solvated (CSD-Refcode NODVIN). Both CC3-Rα and CC3-Rβ-solvated were initially predicted in space groups $P2_12_12_1$ and *P*1 respectively. Full symmetry was revealed after analysis with PLATON.

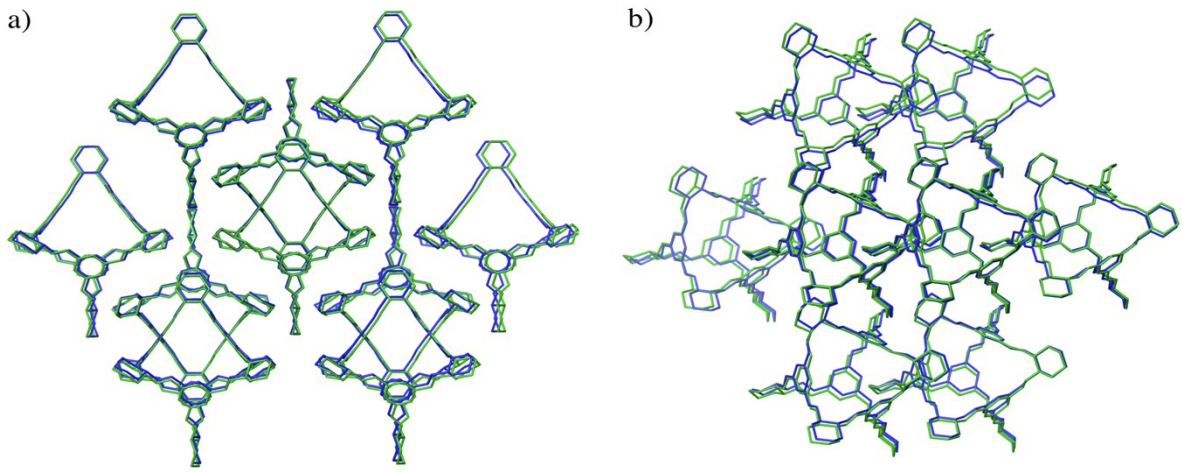

Figure S3 Structural Overlay of cluster of 15 molecules for experimental (blue) and predicted (green) CC3-Rα (Panel a) and CC3-Rβ-solvated (Panel b). Solvent molecules in panel b are omitted for clarity.

Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
CC3-(R,S)α	$P2_1/c$	0.00	3.74	0.94	0.00	4.05	0.99
CC3-(R,S)	C_2/c	6.71	3.74	0.94	7.86	3.93	0.99
CC3-(R,S)	$Pna2_1$	15.40	3.74	0.93	22.60	4.00	0.98
CC3-(R,S)	C_2/c	27.51	3.74	0.92	30.93	3.86	0.99
CC3-Rα	$F4_132$	36.21	3.74	0.92	53.38	3.89	0.95

Table S3. Comparison between CSP and DFT-D3 results for cage CC3. ΔE are given w.r.t the energy of the global minimum found in CSP

S3. Predicted α-phases of molecules M1-M6.

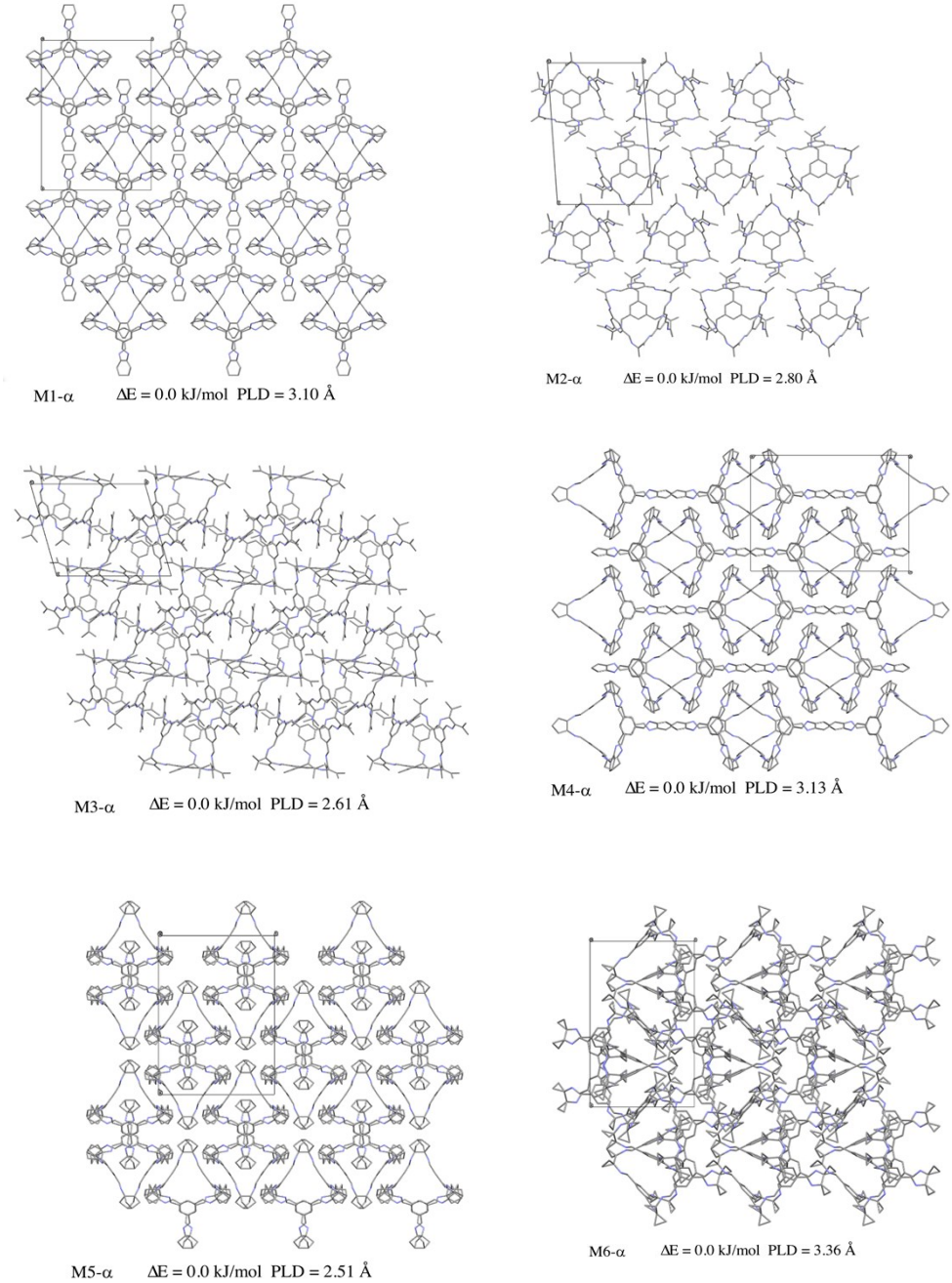

Figure S4 Lowest energy predicted structures for M1-M6. The information in the captions includes the relative lattice energy (ΔE) and the PLD.

S4. Comparison between CSP and DFT-D3 results for the first five predicted structures of molecules M1-M6.

Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
M1-α	*F*d-3	0.00	3.10	0.96	0.00	3.12	1.00
M1-2	$P\bar{1}$	6.37	3.10	0.93	-8.80	2.73	1.00
M1-3	$P\bar{1}$	7.22	3.08	0.94	-25.66	2.88	1.07
M1-4	$P\bar{1}$	9.35	3.08	0.93	-4.42	2.53	1.00
M1-5	$P\bar{1}$	10.29	3.23	0.93	-2.27	2.80	1.00
Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
M2-α	$P\bar{1}$	0.00	2.80	0.97	0.00	2.09	1.04
M2-2	$P2_1/c$	2.66	2.88	0.95	0.71	2.48	1.03
M2-3	$P\bar{1}$	3.30	2.47	0.97	-6.34	2.06	1.05
M2-4	$P\bar{1}$	3.52	2.60	0.97	-6.44	2.04	1.05
M2-5	$P2_1/c$	3.74	2.82	0.96	-8.80	2.55	1.03
Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
M3-α	$P\bar{1}$	0.00	2.61	0.84	0.00	1.70	0.98
M3-2	$P2_1/c$	1.65	2.44	0.83	37.09	1.64	0.92
M3-3	$P2_1/c$	6.66	3.07	0.80	46.64	2.44	0.89
M3-4	$P\bar{1}$	7.30	2.09	0.84	31.56	1.80	0.95
M3-5	C_2/c	10.90	3.01	0.81	45.01	2.17	0.90
Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
M4-α	*F*d-3	0.00	3.14	1.08	0.00	3.30	1.09
M4-2	$P\bar{1}$	45.54	3.14	0.97	49.40	3.13	1.01
M4-3	$P\bar{1}$	50.22	3.13	0.97	17.63	3.27	1.09
M4-4	$P2_1/c$	51.50	3.12	0.97	51.51	3.20	1.04
M4-5	$P2_1/c$	52.03	3.13	0.98	13.78	3.84	1.09
Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
M5-α	*F*d-3	0.00	2.51	1.04	0.00	1.45	1.22
M5-2	C_2/c	21.48	2.59	1.03	21.62	1.81	1.18
M5-3	$P2_1/c$	38.90	2.95	1.00	38.22	2.84	1.07
M5-4	$P\bar{1}$	39.40	2.94	0.99	38.96	2.84	1.06
M5-5	$P\bar{1}$	40.38	2.94	1.00	39.06	2.82	1.06
Structure	space group	CSP ΔE(kJ/mol)	CSP PLD(Å)	CSP density(g/cm^3)	DFT-D3 ΔE(kJ/mol)	DFT-D3 PLD(Å)	DFT-D3 density(g/cm^3)
M6-α	$P2_1/c$	0.00	3.37	0.98	0.00	3.13	1.04
M6-2	$P\bar{1}$	2.62	2.22	0.95	7.45	1.93	1.00
M6-3	$P\bar{1}$	5.83	2.28	0.93	23.24	2.00	0.98
M6-4	$P2_1/c$	7.52	3.30	0.94	22.77	3.03	0.98
M6-5	*R*3	13.04	2.19	0.95	-8.80	1.96	1.07

Table S4 Comparison between CSP and DFT-D3 results for the first five structures w.r.t. the predicted global minimum. ΔE are given w.r.t the energy of the the α-phases at CSP and DFT-D3 level.

S5. H-bond patterns in M1-γ and M2-γ.

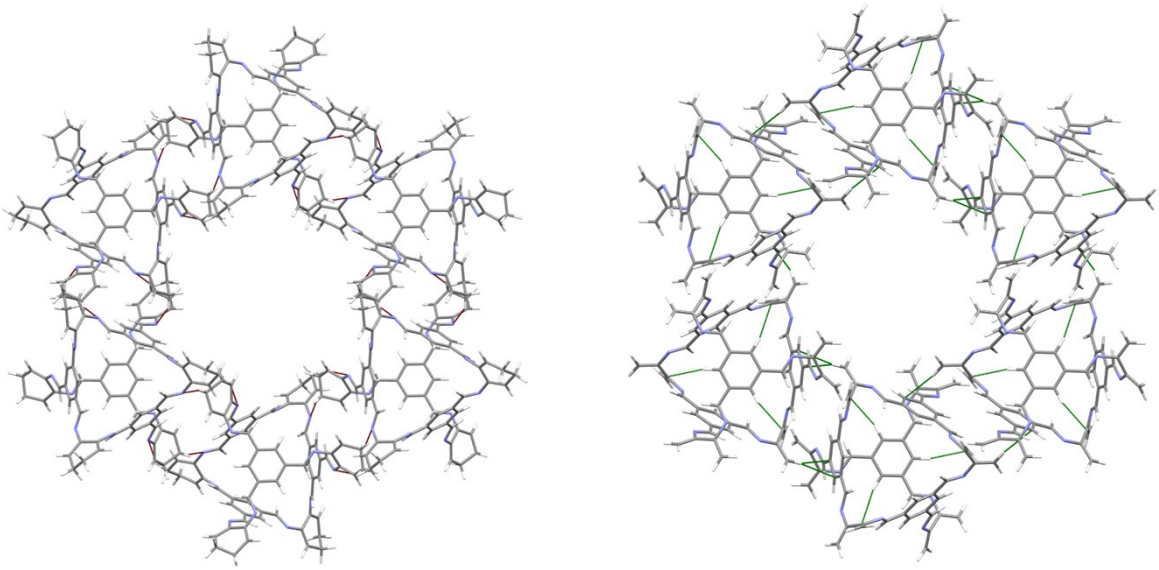

Figure S5 Left panel: Small cluster of the structure M1-γ. Intermolecular C(sp^3)H---N bonds are displayed as red lines. Right panel: Small cluster of the structure M2-γ. Intermolecular C(sp^2)H--- π bonds are displayed as green lines.

S6. MD simulations to assess the thermal stability of M1-γ and M2-γ.

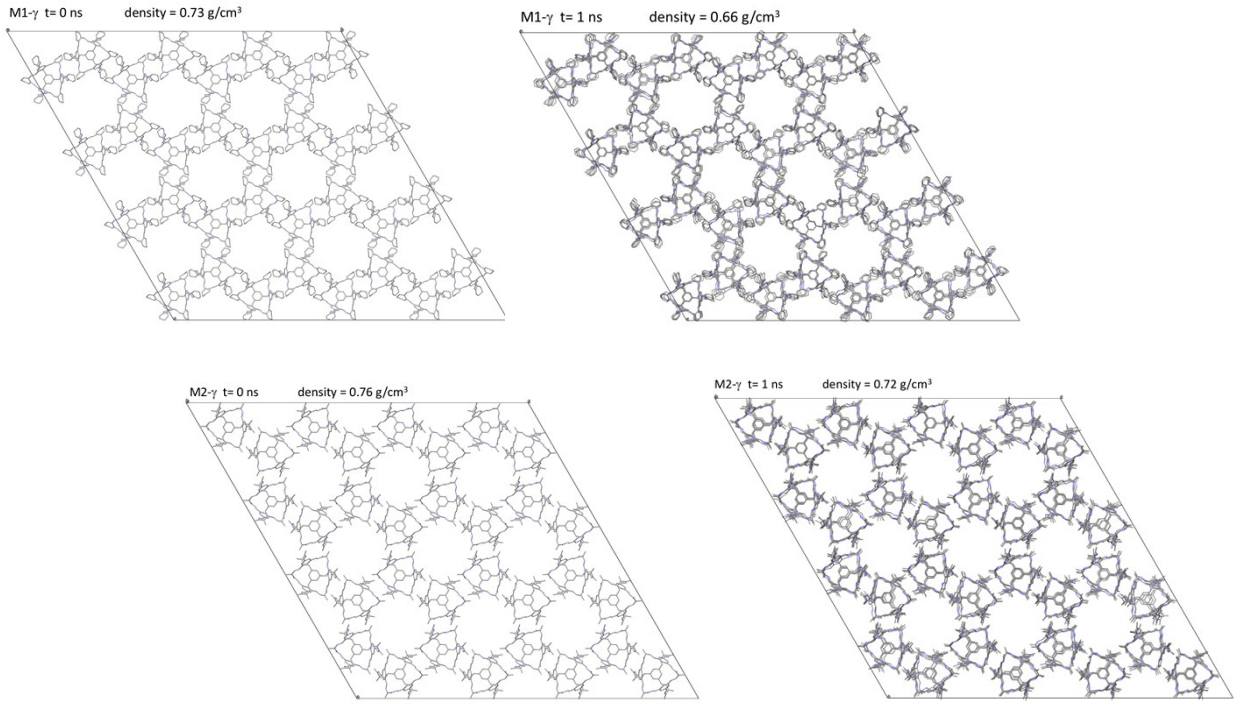

Figure S6 **Top panels**: Projection of a supercell of structure M1-γ at the beginning (left) and at the end (right) of a 1ns MD simulation. **Bottom panels:** Projection of a supercell of structure M2-γ at the beginning (left) and at the end (right) of a 1ns MD simulation

S7. References.

1 T. Tozawa, J. T. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespear *et al. Nature Materials* 2009, **8**, 973.
2 L. Chen, P. S. Reiss, S. Y. Chong, D. Holden, K. E. Jelfs, T. Hasell *et al., Nature* Materials, 2014, **13**, 954.
3 A. Kewley, A. Stephenson, L. Chen, M. E. Briggs, T. Hasell and A. I. Cooper, *Chemistry of Materials* 2015, **27**, 3207.;T. Mitra, K. E. Jelfs, M. Schmidtmann, A. Ahmed, S. Y. Chong, D. J. Adamas and A. I. Cooper, *Nature Chemistry,* 2013, **5**, 276.
4 X. Kong and J. Jiang, *Phys. Chem. Chem. Phys.,* 2017, **19**, 18178.;X. Kong and J. Jiang, *J. Phys. Chem. C,* 2018, **122**, 1732.
5 E. O. Pyzer-Knapp, H. P. G. Thompson, F. Schiffmann, K. E. Jelfs, S. Y. Chong, M. A. Little, A. I. Cooper and G. M. Day. *Chem. Sci.* 2014, **5**, 2235 and references therein.
6 J. A. T. Jones, T. Hasell, X. Wu, J. Bacsa, K. E. Jelfs, M. Schmidtmann *et al., Nature,* 2011, **474**, 367.