Design of efficient sergeant molecules for chiral induction in nano-porous supramolecular assemblies

Andrea Minoia,* Iris Destoop, Elke Ghijsens, Steven De Feyter, Kazukuni Tahara, Yoshito Tobe and Roberto Lazzaroni

ESI-1. MD/Quench iterative scheme with Structure Recombination **1**

ESI-2. Computational Details **2**

ESI-3. Assemblies Characterization **3**
Energetic Analysis *3*
Structural Analysis *4*

ESI-4. On the coupling between supramolecular structures and molecules-molecules and molecules-surface interactions **7**

Reference **9**

ESI-1. MD/Quench iterative scheme with Structure Recombination

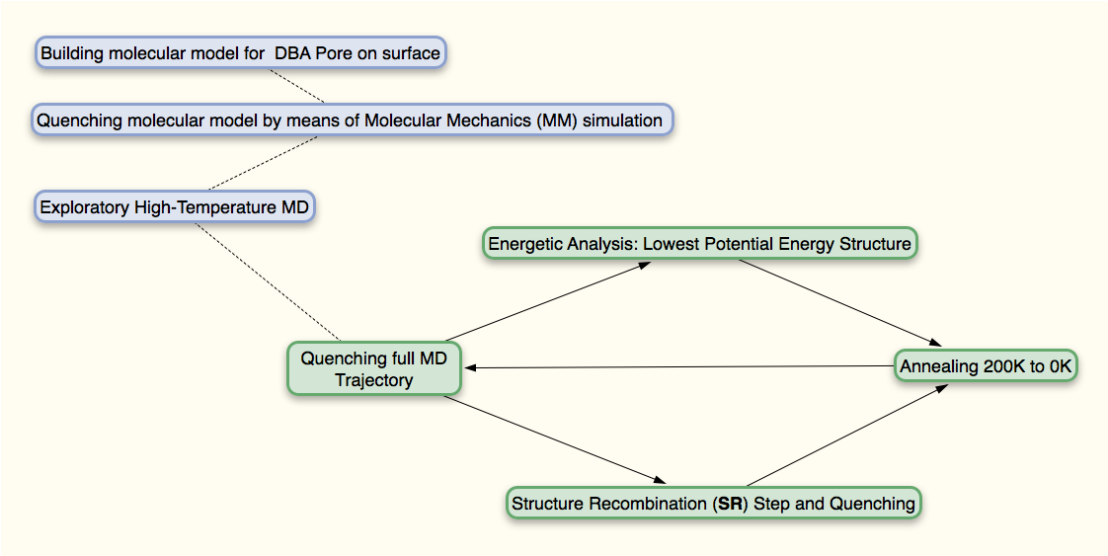

Figure S1. Flowchart for the MD/Quench iterative scheme with SR step.

Figure S1 shows a flowchart to illustrate the main steps of a new MD/Quench iterative scheme, which includes a Structure Recombination (SR) step. The new methodology works as following: the initial structure of the pore is built and minimized to generate a reasonable starting geometry as input for the following step. Next, the minimized pore is treated as a rigid body on the frozen HOPG graphite surface: a 150ps-long MD is performed in the NVT ensemble at 500K in order to sample different orientations of the pore on the surface. The pore must be treated as a rigid body because its molecules are kept assembled by the rather weak van der Waals and electrostatic interactions, which can be easily overcome by the thermal energy available at high temperature and, ultimately, a non-restrained pore would be likely to

disassemble under those conditions. The initial high-energy MD simulation generates 300 structures that are all fully minimized (i.e., all the constraints keeping the pore a rigid body are removed) and their energies are analyzed and compared. From this point on, the search for the most stable structure is done iteratively (green steps in Figure S1). The Structure Recombination (SR) step consists in aligning the lowest-energy structure for the pore alone to the pore geometry that best interacts with the surface. This SR structure, which is uncorrelated to those obtained from the previous MD simulation, is then minimized. This helps to better explore the potential energy landscape of the system. To further refine the search of the lowest-energy structure, both the lowest total potential energy structure found during the previous MD simulation and the SR structure are annealed via a 30ps-long MD simulation, during which the temperature is lowered from 200K to 0K in a sigmoidal way. Here, a low temperature of 200K is used because the pore is no longer treated as a rigid body and 200K is the highest possible temperature for which the pore remains intact for some tens of picoseconds. Each annealing simulation produces 60 structures, which are all minimized and their energy analyzed and compared. A new low-energy structure is found and a new SR structure is created and quenched. These new structures are annealed and the whole procedure is repeated until the improvement of the energy of the system get smaller than the 0.01 kcal mol^{-1}. The lowest-energy structures obtained for the CW and CCW pores are considered to be the most stable aggregates on the surface and are used to investigate the role of chirality on their stability.

ESI-2. Computational Details

This work is based on a combination of force field-based molecular mechanics (MM) and molecular dynamics (MD) simulations. All 'dry' simulations (i.e. without explicit solvent) of the DBA pores on the graphite surface have been performed using the molecular modeling package TINKER 6.1[1] and its MM3(2000)[2] force field. Below, we summarize the main characteristics for the simulations.

- Molecular code: TINKER 6.1;
- Force field: MM3(2000);
- Periodic System: No;
- Surface: Single, frozen square HOPG graphite layer having a surface of about 262 nm^2;
- DBA assembly: Single, isolated pore on surface;
- Solvent phase: No;
- MD time step: 1fs;
- MD ensemble: NVT;
- Temperature: 500K for initial rigid-body calculations, annealing from 200K to 0K;
- Thermostat: Andersen;
- Barostat: None.

ESI-3. Assemblies Characterization

Energetic Analysis

One of the major advantages of force field based simulations over quantum chemical calculations is that the total energy of the system can be easily decomposed in energy contributions. This allows to compute and to compare specific energy contributions in the attempt to understand how chirality affects the energy of the aggregate. In particular, we are interested in the following energies:

1. E1: the total potential energy of the molecular system;
2. E2: the total potential energy of the pore;
3. E3: the pore-surface interaction energy;
4. E4: the interaction of the chiral molecule with the achiral part of the pore;
5. E5: the potential energy of the achiral part of the pore;
6. E6: the potential energy of the chiral molecule in the pore;

The results for the different systems are reported in Table 1 to Table 4.

Table 1. Energetic analysis for -Type CW and +Type CCW **DBA-OC12(2-S)-OC12(2-R)** pores.

Energy Terms	CW pore (kcal mol^{-1})	CCW pore (kcal mol^{-1})	$\Delta E_{(CW-CCW)}$ (kcal mol^{-1})
E1	-348.3	-348.2	-0.1
E2	344.7	344.7	-0.0
E3	-693.0	-693.1	-0.1
E4	-33.6	-33.7	-0.1
E5	296.6	296.6	-0.0
E6	81.7	81.7	0.0

Table 2. Energetic analysis for -Type CW and +Type CCW **cDBA-OC12(2-S)-OC12(3-S)** pores.

Energy Term	CW pore (kcal mol^{-1})	CCW pore (kcal mol^{-1})	$\Delta E_{(CW-CCW)}$ (kcal mol^{-1})
E1	-343.0	-342.3	-0.7
E2	348.7	348.8	-0.1
E3	-691.7	-691.1	-0.6
E4	-33.4	-33.8	0.4
E5	296.5	297.4	-0.9
E6	85.6	85.2	0.4

Table 3. Energetic analysis for -Type CW and +Type CCW **DBA-OC12-OC13** pores.

Energy Term	CW pore (kcal mol^{-1})	CCW pore (kcal mol^{-1})	$\Delta E_{(CW-CCW)}$ (kcal mol^{-1})
E1	-355.9	-351.5	-4.4
E2	337.5	338.9	-1.4
E3	-693.3	-690.4	-2.9
E4	-34.6	-32.5	-2.1
E5	297.3	296.3	1
E6	74.7	75.1	-0.4

Table 4. Energetic analysis for -Type CW and +Type CCW **cDBA-OC12(2-S)-OC13(2-R)** pores.

Energy Term	CW pore (kcal mol^{-1})	CCW pore (kcal mol^{-1})	$\Delta E_{(CW-CCW)}$ (kcal mol^{-1})
E1	-348.8	-345.9	-2.9
E2	344.3	346.6	-2.3
E3	-693.1	-692.5	-0.6
E4	-35.2	-33.4	-1.8
E5	295.9	296.2	-0.3
E6	83.6	83.9	-0.3

The results from the energetic analysis can be summarize as following: the main reason for the different stabilities of the structures lies in the difference between their internal energies, suggesting some stress occurs in the higher energy structures. To better understand the reasons why certain chiral structures have higher internal strees then their counterparts having opposite chirality, a structural analysis is required.

Structural Analysis

The energetic analysis is useful to determine which pore is more stable and which energy contributions are responsible for the difference in stability. Because most of the difference in the -Type CW/ +Type CCW stability for the different honeycomb structures comes from the internal energy of those structures, a structural analysis is the appropriate tool to pinpoint the causes of the difference in their stability. At first, we macroscopically characterize each pore by replacing the DBA molecule by their centers of mass (Figure S2a): we will refer to this analysis as COM analysis. Next, we use the VMD[5] program to calculate and compare the root mean square deviation (RMSD) between the different COM structures. Finally, the COM structures are characterized by measuring the distances and angles between the centers of mass (Figure S2b). The combination of these two analyses allows us to detect and quantify major deformations and differences between the different pores. We found that a more "fine" structural characterization is too noisy and hence less usefull than the COM analysis due to the numerous local fluctuations.

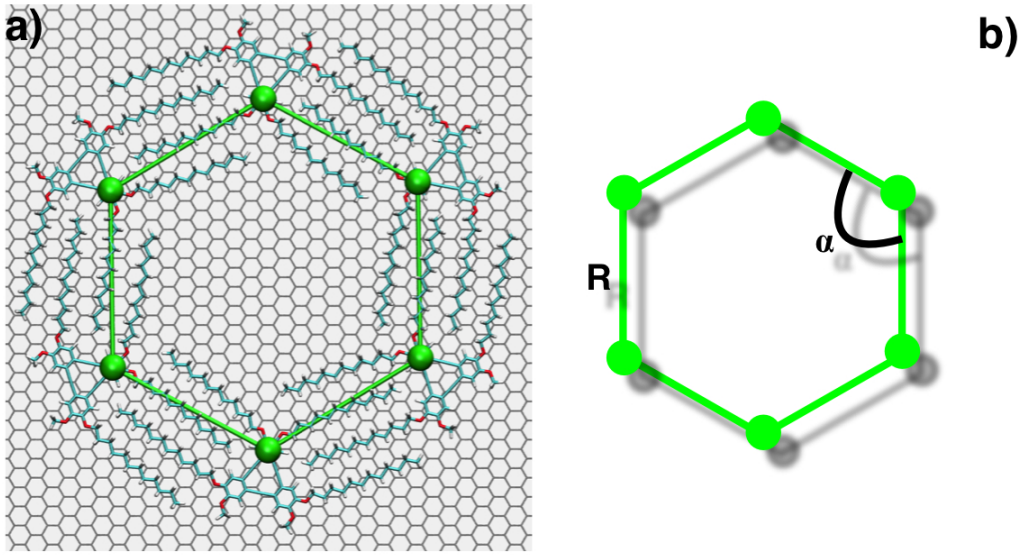

Figure S2. (a) All-atomistic and COM representation for a single pore.

Table 5 summarizes the structural differences between all the different structures by means of their root mean square displacement, RMSD, (normalized to 1).

Table 5. Calculated RMSD values for the different -Type CW and +Type CCW COM structures. In bold the RMSD larger than 0.5. The RMSD values are normalized to 1 and are in arbitrary units. S1=**DBA-OC12**, S2=**DBA-OC12(2-S)-OC12(2-R)**, S3=**cDBA-OC12(2-S)-OC12(3-S)**, S4=**DBA-OC12-OC13** and S5=**cDBA-OC12(2-S)-OC13(2-R)**.

	RMSD	S1		S2		S3		S4		S5	
		CW	CCW	CW	CCW	CW	CCW	CW	CCW	CW	CCW
S1	CW	0.000									
	CCW	0.000	0.000								
S2	CW	0.373	0.376	0.000							
	CCW	0.373	0.376	0.050	0.000						
S3	CW	0.310	0.313	0.143	0.129	0.000					
	CCW	0.367	0.369	0.395	0.393	0.370	0.000				
S4	CW	0.390	0.392	0.372	0.376	0.367	0.088	0.000			
	CCW	**0.992**	**0.994**	**0.824**	**0.821**	**0.835**	**0.700**	**0.668**	0.000		
S5	CW	0.381	0.385	0.124	0.131	0.178	0.331	0.296	**0.768**	0.000	
	CCW	**0.998**	**1.000**	**0.827**	**0.819**	**0.835**	**0.689**	**0.659**	0.141	**0.763**	0.000

The results of the RMSD analysis reported in Table 5 show that CCW assemblies of **DBA-OC12-OC13** (S4 in the table) and **cDBA-OC12(2-S)-OC13(2-R)** (S5 in the table) are the ones whose structures differ the most from all the others, which indicates an important deformations happen for those to assemblies.

For a more accurate description of the COM structures, Figure S5 shows how the average R and α (with their standard deviations as error bars), measured according to Figure S2b, fluctuate in the pores series, with the **cDBA-OC12(2-S)-OC13(2-R)** CCW being the most irregular (highest fluctuations in the averaged R and α structural parameters).

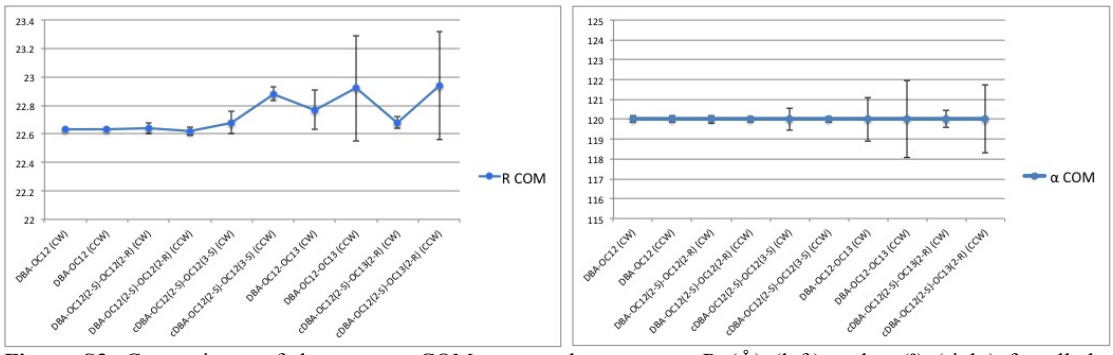

Figure S3. Comparisons of the average COM structural parameters R (Å) (left) and α (°) (right) for all the different pores.

In order to fully characterize the COM structures, the averages of R and α structural parameters shown in Figure S5 are not sufficient to distinguish deformations that are localized from those that are more spreaded over the entire assembly. Figure S6 shows the variations of the parameters R and α within the assembly, when measuring from molecule 1 (the sergeant molecule in the assembly) to molecule 6, following the chirality of the assembly.

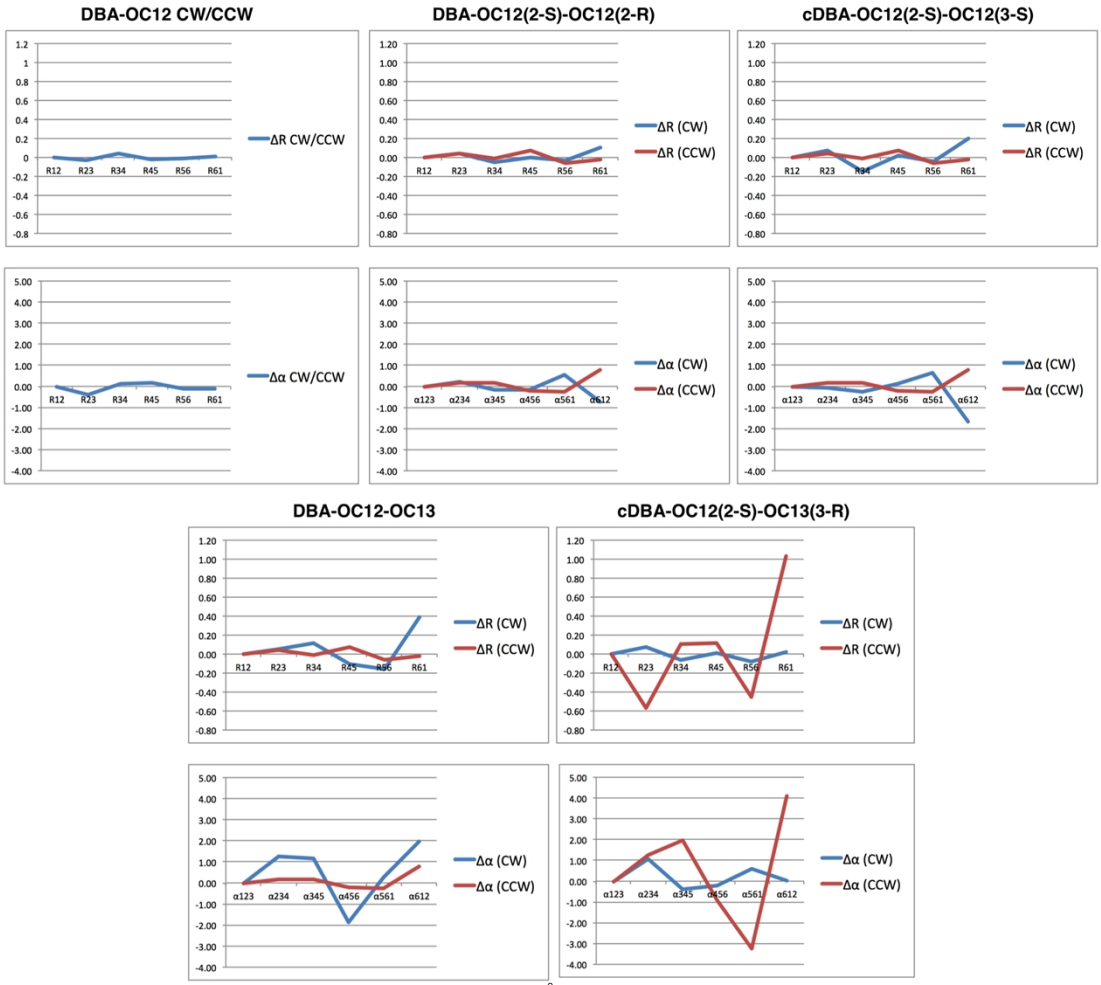

Figure S4. Trends of the COM structural parameters R (Å) and α (°) for the different COM pores.

While for DBA-OC12 CW/CCW pores the fluctuations of the structural parameters are spread throughout the pore, for the **DBA-OC12-OC13** and **cDBA-OC12(2-S)-OC13(2-R)** pores (both CW and CCW) the largest deformations are mostly localized in the neighborhood of the chiral molecule, with the largest fluctuations being measured in the **cDBA-OC12(2-S)-OC13(2-R)** CCW pore. The main reason for the large, highly localized deformations those two systems is due to the fact that with the alternation of -OC12 and -OC13 alkoxy group in the sergeant molecular structure, -Type and +Type interdigitations are not equivalent: while in the -Type CW structure the -OC13 akoxy groups (in green in Figure S5) lie along the inner and outer perimeter of the pore, in the +Type CCW structure they are pointing directly at the conjugated, rigid core of the adjacent DBA molecules, leading to the local deformation of the assembly.

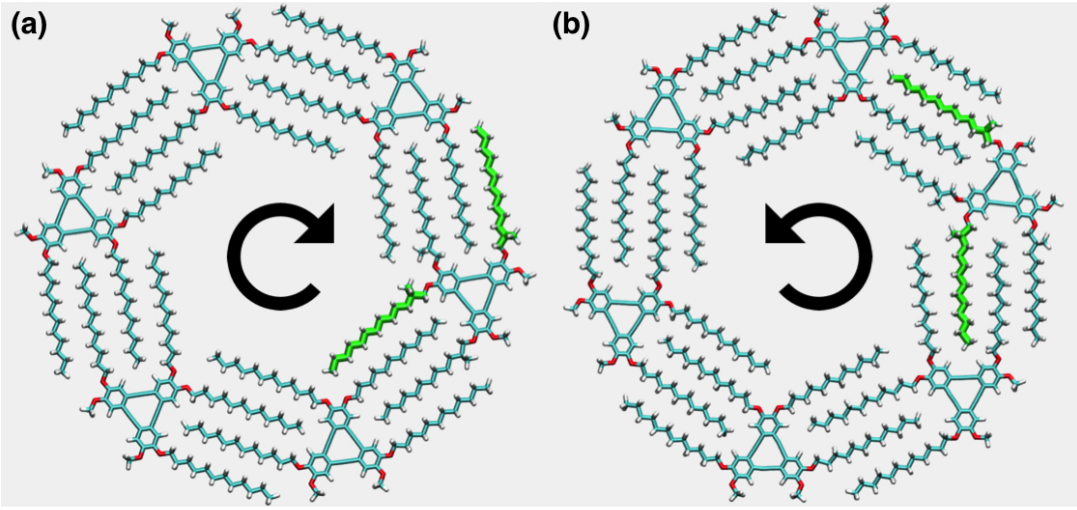

Figure S5. Comparison between the position of the –OC13 alkoxy groups in the CW and CCW assemblies involving the **cDBA-OC12(S)-OC13(R)** sergeant molecule.

ESI-4. On the coupling between supramolecular structures and molecules-molecules and molecules-surface interactions

To show how the interaction between the pore and the surface, $E_{pore\text{-}HOPG}$, and the interactions between the molecules in the pore, E_{pore}, are mutually influenced, we plot in Figure S6a the time evolution of the **DBA-OC12(2-S)-OC12(2-R)** CW pore orientation on the surface (i.e., the alignment of the two dashed lines in Figure S6b), of its $E_{pore\text{-}HOPG}$ interactions (with and without rigid-body constraints on the pore) and of its ${}^{Q}E_{pore}$ during the initial high-temperature MD simulation.

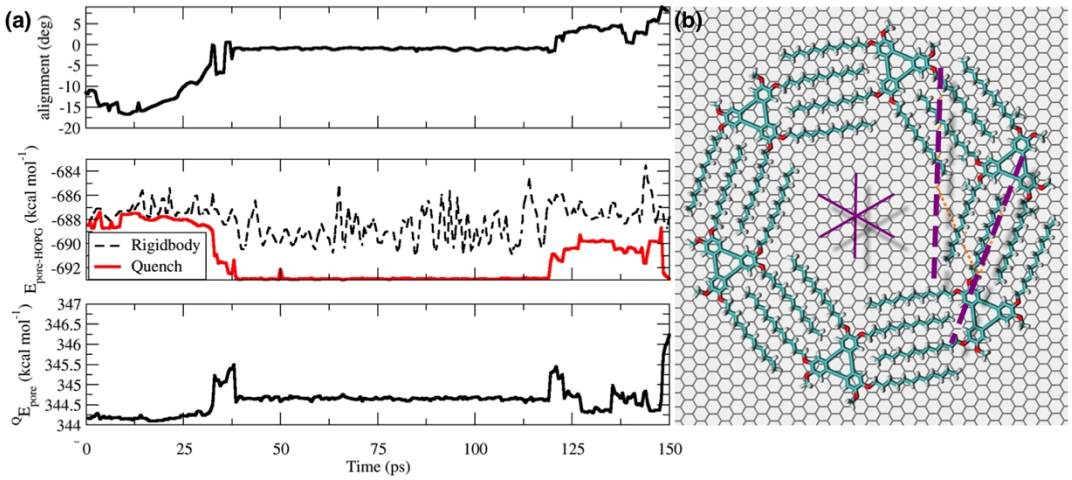

Figure S6. (a) Time-evolutions of the pore orientation on the surface, and the $E_{pore\text{-}HOPG}$, ${}^{Q}E_{pore\text{-}HOPG}$ and ${}^{Q}E_{pore}$ energy terms. (b) The angle between the dashed lines defines the pore orientation on the surface. The solid lines indicate the three main axes of symmetry of HOPG graphite.

During the MD runs, the pore (treated as rigid body) diffuses and rotates on the surface while its interactions with the graphite, $E_{pore\text{-}HOPG}$ (in red in Figure S6a), fluctuate around an average value of -688.1 ± 1.3 kcal mol^{-1}. After removing the rigid body constrains, the system is quenched and the quenched $E_{pore\text{-}HOPG}$, ${}^{Q}E_{pore\text{-}HOPG}$, drops to an average value of -691.3 ± 2.0 kcal mol^{-1}. The ${}^{Q}E_{pore\text{-}HOPG}$ time evolution shows that when the pore is oriented so that the alkoxy group of reference aligns to

the graphite main axis of reference (dashed lines in Figure S6b) the pore-HOPG interaction is maximized while its strength is reduced when the pore is oriented otherwise. This is due to the fact that most of the alkoxy groups in the pore are properly aligned with one of the three main axes of HOPG graphite (solid purple lines in Figure S6b). The energy of the pore (without rigid body constrains), $^{Q}E_{pore}$, also changes due to the fact the molecules react to their different orientations on the surface adopted during the MD simulation by changing their local conformations to maximize their interactions both with the surface and with the other molecules in the pore. Interestingly, the lowest-energy structures for the pore are not those that maximize its interaction with the surface; this suggests that, in order for the pore to better interact with the surface, some stress is generated in the pore. Figure S7 relates the time evolutions of $^{Q}E_{pore}$, $^{Q}E_{pore\text{-}HOPG}$ and the total potential energies, E_{tot}, with the changes in the structure of the pore, which is quantified by calculating the RMSD of each pore with respect to that of the pore in the lowest total potential energy structure (color coded in Figure S7). Figure S7 also shows that there is not a direct link between the RMSD and the energy difference calculated between two structures: pores having small RMSD may be close in energy or not, depending on how well/bad they interact with the surface. In fact, the evolution of the total potential energy (Figure S7b) is very similar to that of the interaction energy between the pore and the surface: this suggests that the pore-surface energy term dominates over the internal energy of the pore, hence the molecules in the pore tend modify their conformations according to the local structure of the surface underneath, so to maximize their interactions with the surface, regardless of the fact that the internal energy of the pore may increase. The interplay we uncovered between the molecule-surface and the molecule-molecule interactions in determining the pore most stable structure is a general characteristic found in any supramolecular system that is interacting with a corrugated surface and cannot be decoupled.

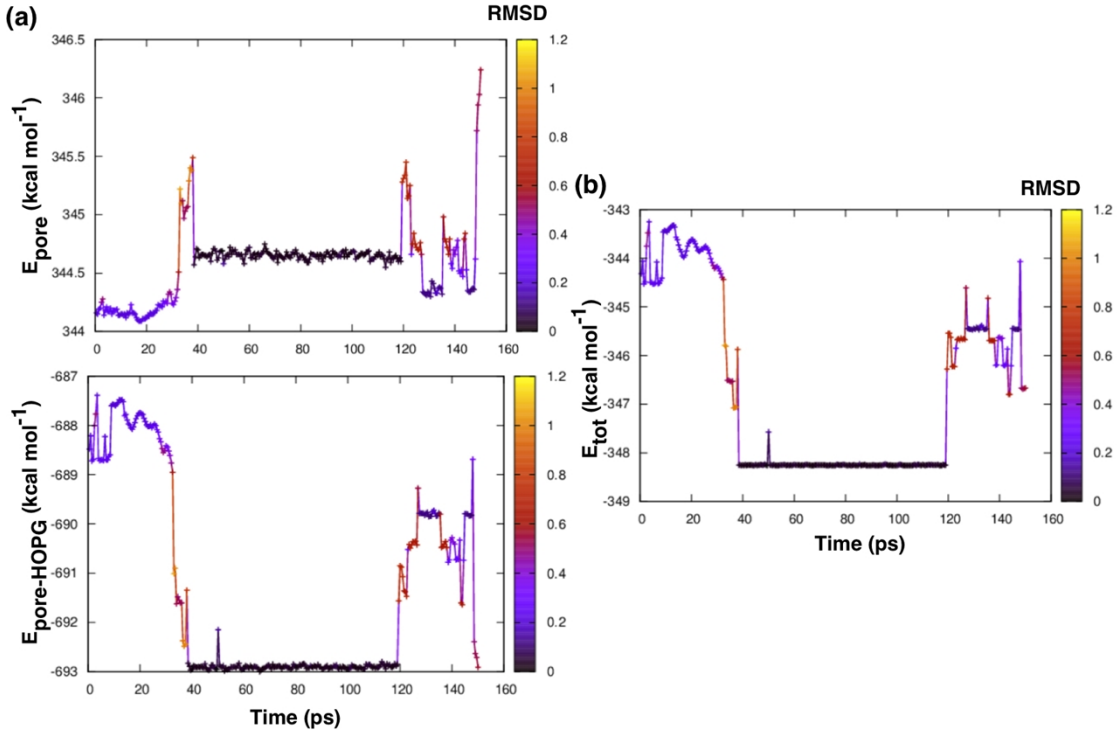

Figure S7. (a) Time evolution of $^{Q}E_{pore}$ and $^{Q}E_{pore\text{-}HOPG}$ for the **DBA-OC12(2-S)-OC12(2-R)** CW pore. (b) Time-evolution of the total potential energy of the system. All lines are color-coded according to the value of the RMSD

calculated for each quenched structure with respect to the structure of the pore in the lowest total potential energy system.

Reference

[1] http://dasher.wustl.edu/tinker/

[2] MM3 J.-H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, 1989, **111**, 8566; J.-H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, 1989, **111**, 8576; N. L. Allinger, F. Li and L. Yan, *J. Comput. Chem.*, 1990, **11**, 848; N. L. Allinger, F. Li, L. Yan and J. C. Tai, *J. Comput. Chem.*, 1990, **11**, 868; N. L. Allinger, Y. H. Yuh and J.-H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551.