

Supporting information

Host-guest chemistry under confinement: Peeking at early self-assembly events

Roelof Steeno,^a Andrea Minoia,^b Roberto Lazzaroni,^b Kunal S. Mali,^{a*} and Steven De Feyter^{a*}

^aDivision of Molecular Imaging and Photonics, Department of Chemistry, KU Leuven, Celestijnenlaan 200F, B-3001 Leuven, Belgium, ^bLaboratory for Chemistry of Novel Materials, Materials Research Institute, University of Mons, Place du Parc 20, 7000 Mons, Belgium

A. Materials and methods

A1: Synthesis of ISAOC18

The alkylated isophthalic acid derivative used in this study (5-octadecyloxy-isophthalic acid, ISAOC18) was synthesized according to a previously reported method.¹

A2: Covalent modification of HOPG

Electrochemical measurements were performed using an Autolab PGSTAT101 (Metrohm Autolab BV, The Netherlands). Prior to each experiment, HOPG (working electrode) was freshly cleaved using scotch tape. The electrochemical modification procedure was carried out in a homemade single-compartment three-electrode cell with a working electrode area of 50.3 mm², Pt wire counter and Ag/AgCl/3 M NaCl reference electrodes. 3,5-Bis-*tert*-butylbenzenediazonium (TBD) is unstable and decomposes rapidly: hence it was synthesized from the corresponding aniline precursor immediately prior to electrochemical reduction. This procedure involves addition of 5 mL of a 2 mM 3,5-bis-*tert*-butylaniline (98%, TCI-Tokyo Chemical Industry Co., Ltd.) in 50 mM HCl (Sigma-Aldrich) to 100 μ L aqueous NaNO₂ (0.1 M) for initiating the diazotization reaction. Within 3 to 5 minutes this mixture was gently shaken and pipetted into the EC cell. Cyclic voltammetry (3 cycles, range: -0.5V to -0.6V, scanning rate: 100 mV/s) was used for the electrochemical activation. After modification, 3,5-bis-*tert*-butylphenyl modified samples were rinsed with Milli-Q water (Milli-Q, Millipore, 18.2 M Ω cm, TOC < 3 ppb) to remove any physisorbed material from the surface and dried in a stream of Argon. All compounds were used without further purification.

A3: STM Experiments

All experiments were performed at room temperature (20-22 °C) using a PicoLE (Keysight) or Molecular Imaging STM setup operating in constant current mode at the 1-phenyloctane (98%, Sigma-Aldrich Co. LLC)/HOPG interface. STM tips were mechanically cut from Pt/Ir wire (80%/20%, diameter 0.25mm). Small amounts of coronene (97%, Sigma-Aldrich Co. LLC) /ISAOC18 solutions in 1-phenyloctane were drop casted on the surface of a freshly cleaved (or covalently modified) highly oriented pyrolytic graphite (HOPG, grade ZYB, Advanced Ceramics Inc., Cleveland, OH, USA) substrate. The following concentration ratios were used: 0.5 mM/0.5 mM, 0.375 mM/0.375 mM, 0.3 mM/0.3 mM, 0.5 mM/0.3 mM, 0.3 mM/0.5 mM and 0.05 mM/0.5 mM. Nanoshaving experiments were performed using PicoLITH (version 2.1) software, applying $I_t = 0.2$ nA and $V_s = -1$ mV. Unless otherwise specified, a nanoshaving speed of 0.4 μ m/s was used. Image analysis was performed using Scanning Probe Image Processor (SPIP) software (Image Metrology ApS). Image calibration was performed by obtaining the graphite lattice underneath the monolayer ($I_t = 0.2$ nA and $V_s = -1$ mV). Unit cell parameters were determined from the fast

Fourier transform of the calibrated images averaged over at least 4 different images at different locations.

A4: Molecular Mechanics and Molecular Dynamics (MM/MD)

All simulations were performed using the Biovia molecular modeling code Materials Studio 2018 (MS2018) using the DREIDING force field available in MS2018. This force field was selected for its ability to (i) properly describe the molecule-molecule and molecule-surface interactions acting in this system and (ii) to explicitly take into account H-bonding interactions. Furthermore, we already successfully used DREIDING to study the morphology and stability of ISAOC18 polymorphs at the HOPG graphite/solvent interface.² As per the electrostatic interactions, Gasteiger atomic charges were assigned to all atoms except those in the periodic graphite surface, for which all atomic charges were set to zero to keep the “infinite” surface neutral. For all MD simulations, a time step of 1fs was used and the temperature was controlled with the Nose-Hoover thermostat.

All self-assembled monolayers used to model **P1** and **P2** polymorphs have been first built, relaxed and optimized using a single slab of periodic HOPG graphite and then further refined using two slabs of HOPG graphite, thus providing a more accurate description of the real surface. Note that in all models the surface has been treated as a rigid object frozen in space.

If not explicitly mentioned, MD and MM simulations were conducted in dry, i.e., without explicit solvent. This approximation is acceptable since there is no experimental evidence of the solvent molecules being incorporated in the structure of the two polymorphs. MD simulations in wet conditions, *i.e.*, including explicit solvent molecules, have been used to estimate the affinity of coronene and ISAOC18 molecules with the solvent as well to probe their sorption/desorption dynamics.

B. Supporting data

B1: On the theoretical morphology and relative stability of P1 & P2

Due to the large number of atoms required to model a self-assembled molecular network at the solid/liquid interface, the number of conformations to explore and the time scale involved, it is extremely unlikely that we could ever observe the formation of either polymorph from simply running extremely long MD simulations of coronene and ISAOC18 molecules randomly dispersed at the graphite/solvent interface.

For this reason, we proceed to build the polymorphs in steps, so to obtain a structure loosely based on that visible in the available experimental STM images, to be later refined with an iterative cycle of short MD simulations followed by a geometry optimization of the system.

The starting point is modeling the ISAOC18 hexamer with a coronene in the middle we can see in the STM images for both **P1** and **P2** (Fig. S1): we consider this aggregate to be a building block common to the two polymorphs.

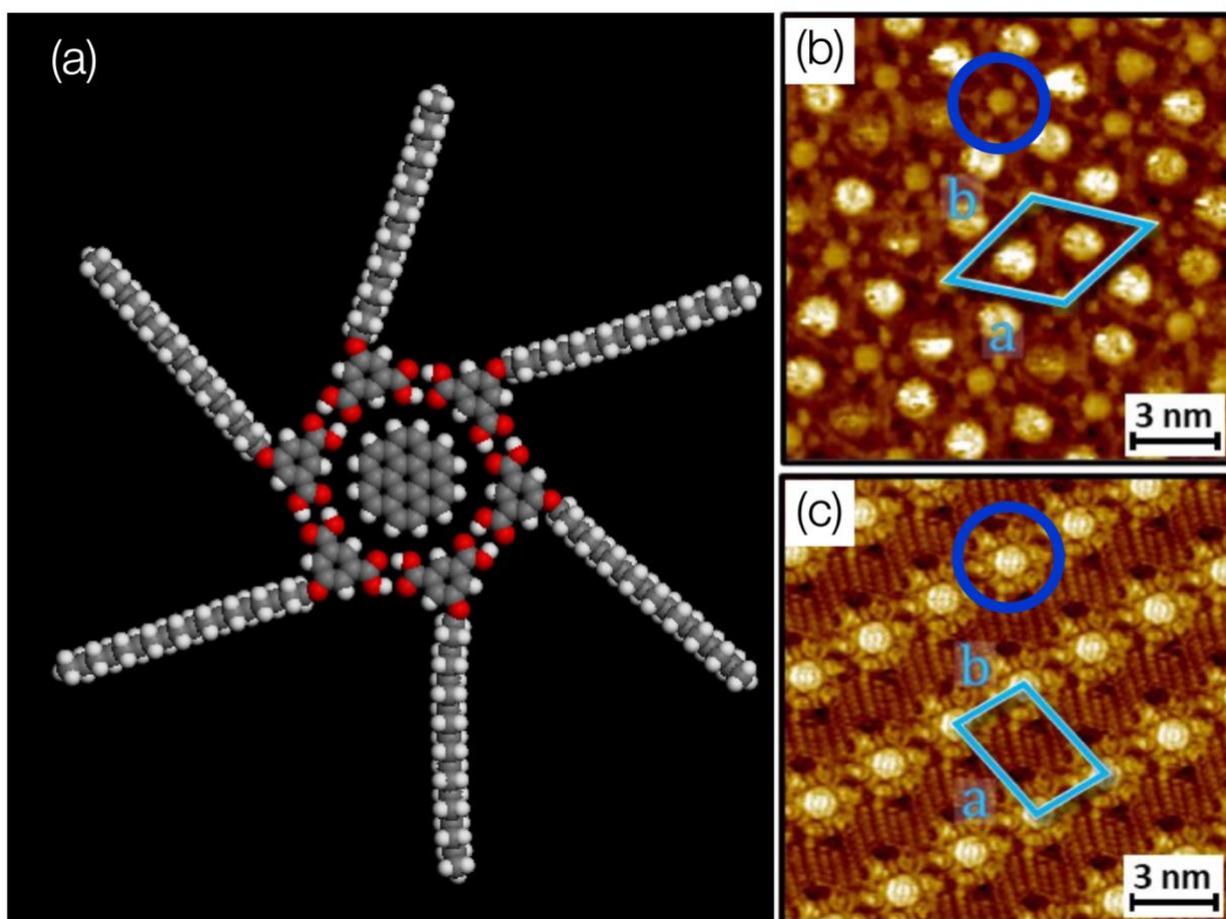


Fig. S1. The building block used to create the polymorphs (a) is clearly visible in both the high-resolution STM images for **P1** (b) and **P2** (c) and is highlighted with blue circles.

Next, we start growing the polymorph by gradually adding and interdigitating more aggregates making sure to relax (and possibly tweak) the structure at each step. Finally, the MD/geometry optimization iterative scheme is used to refine the morphology for the largest system built.

Initial attempts yielded molecular models that matched the symmetry of the unit cell observed in STM images however comparison of the theoretical unit cells with those extracted from the STM images indicated significant differences.

In particular, the significant discrepancies were found in values of vectors a and b for **P1** and vector b for **P2**. This error always occurs in the directions along which two opposite ISAOC18 molecules interdigitate (Fig. S2a.), and we believe that the ISAOC18 conformation must be responsible for it. To build our initial molecular models, we used a somewhat linear ISAOC18 molecule, where the tilt angle between the head and the tail of the molecule is about 146° .

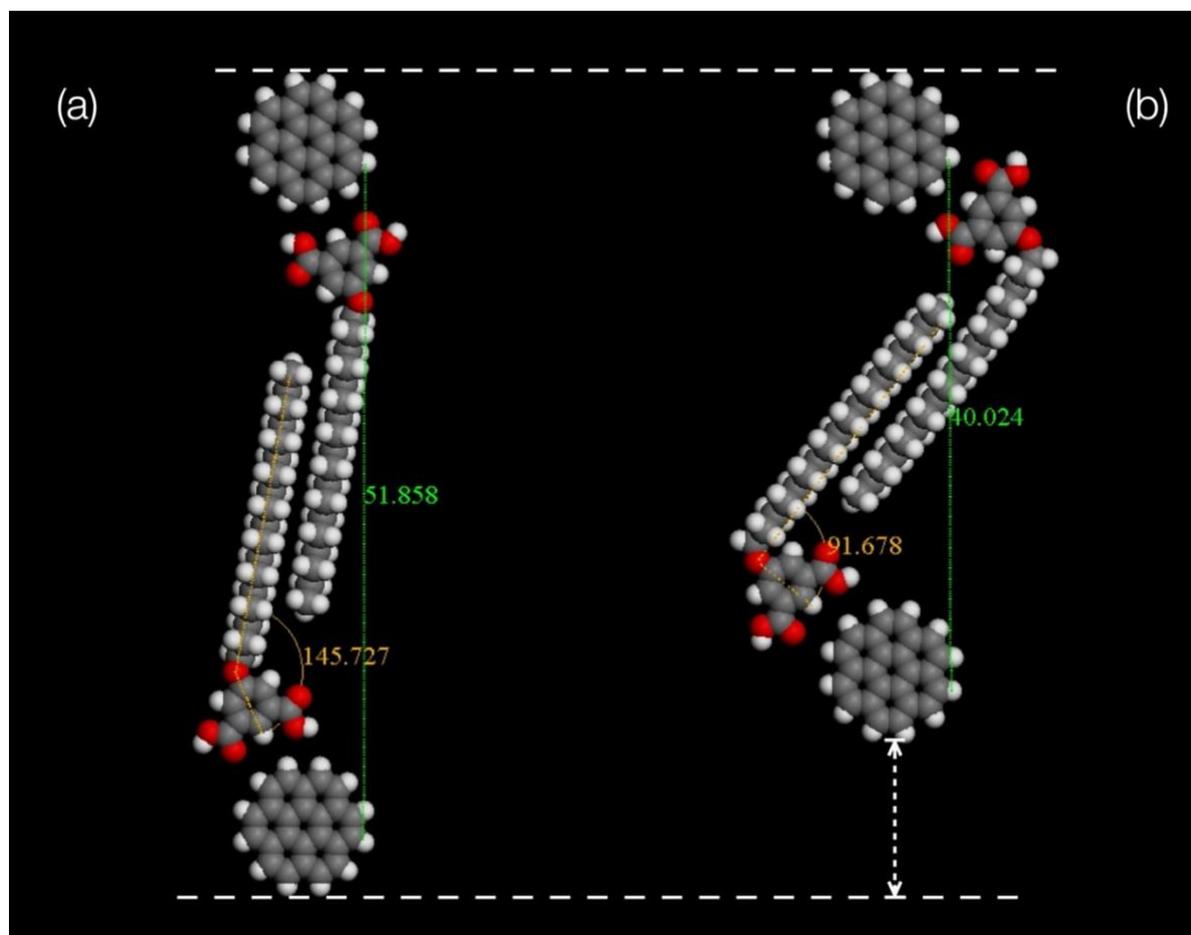


Fig. S2. The effect of interdigitating (a) “linear” vs “bent” (b) ISAOC18 molecule.

By reducing this tilt angle to about 92° (Fig. S2b), we get a strongly bent molecular conformation, but the molecules can still interdigitate and the overall distance between the coronene molecules in the interdigitating aggregates is reduced from 5 nm to about 4 nm, in line with the experimental unit cell parameters.

With this new molecular conformation, we set to build a new set of polymorphs following the same procedure we detailed early on. It is worth noting that the new ISA conformation forms a building block that has a smaller size than before, as shown in Fig. S3.

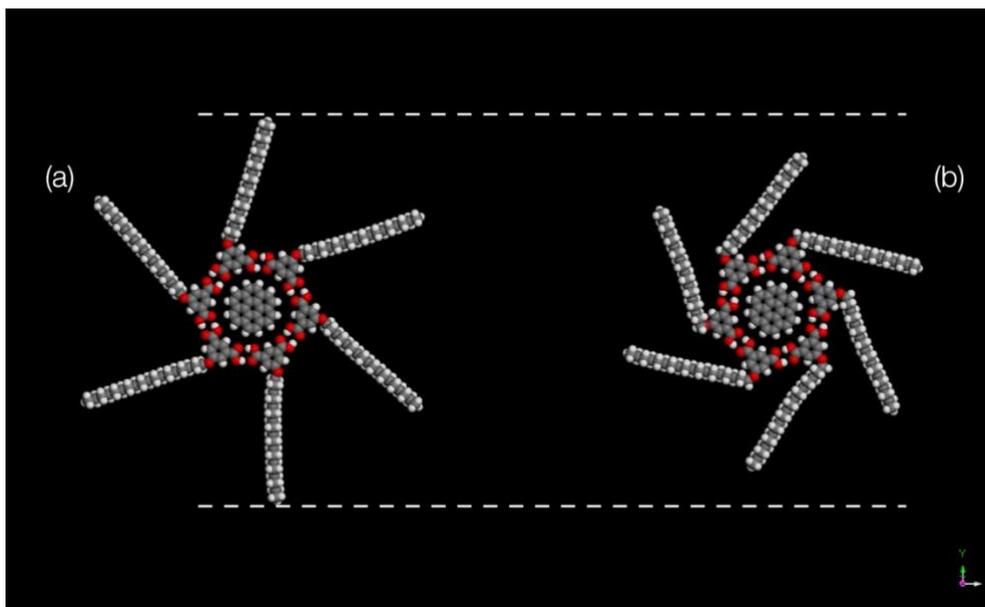


Fig. S3. Comparison of the size and shape between the (a) old and (b) new building block.

With this new tilted ISAOC18 molecule we build a new set of polymorphs and Fig. S4 shows the best candidate models for **P1** and **P2**.

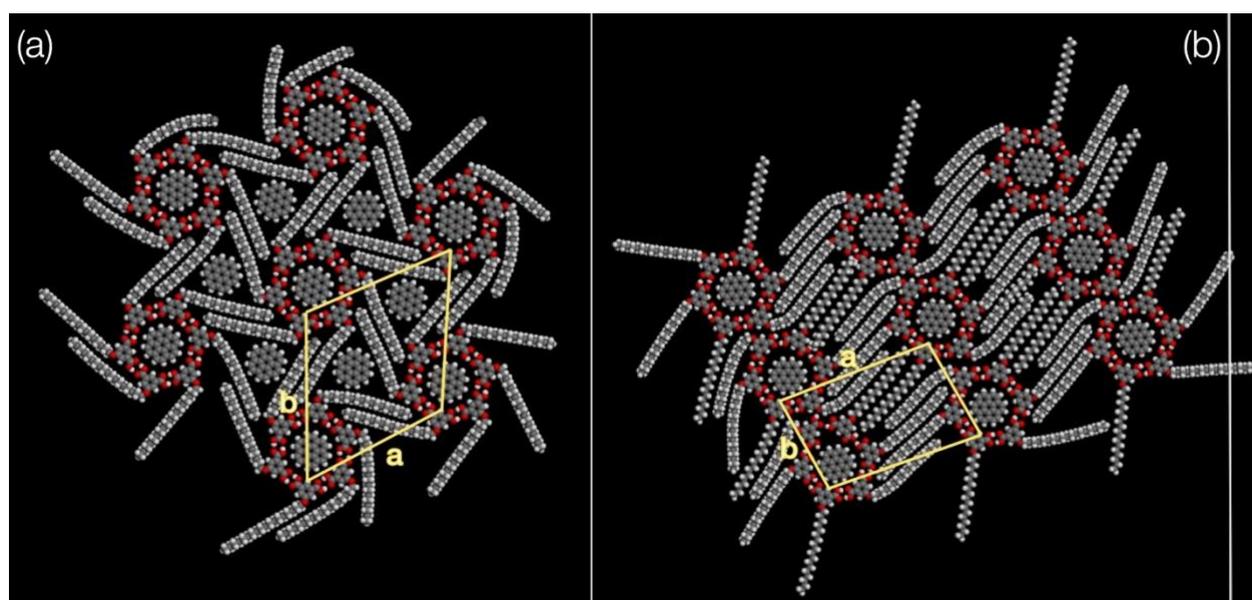


Fig. S4. The new models for (a) **P1** and (b) **P2** polymorphs.

Table S1. Comparison between the theoretical and experimental unit cells for the new polymorphs.

Polymorph	a (nm)	b (nm)	γ ($^\circ$)
P1 (theory)	4.1	4.0	61.0
P1 (exp)	4.05 ± 0.14	4.03 ± 0.16	58.6 ± 3.1
P2 (theory)	4.3	2.7	81.7
P2 (exp)	4.03 ± 0.03	2.65 ± 0.05	78.2 ± 1.5

The molecular model as well as the STM image indicate that **P2** is a dense polymorph and denser than **P1**, with a high degree of ISAOC18 interdigitation (clearly visible in the STM images) where all COR are immobilized in the hexagonal cavity. Both the model and the experimental images also show similar gaps in the interdigitation of the ISAOC18 molecules.

To establish which polymorph is most stable on the surface, we calculated their adsorption energy per unit area. We estimated the surface coverage of the polymorphs with the area of the geometrical figures that best approximate their shape and sizes; a hexagon with an area of 139.5 nm² for **P1** and a parallelogram having an area of 109.5 nm² for **P2**.

Table S2. Comparison of the calculated surface coverage and the adsorption energies per unit area (E_{sta}) for **P1** and **P2**.

Polymorph	Surface Coverage (nm ²)	E_{sta} (kcal mol ⁻¹ nm ⁻²)
P1	103.1	-27.8
P2	88.0	-39.5

B2: On the interpretation of contrast in STM images.

STM images are all about contrast. Molecules should stay still long enough to be visible on the images. In fact, most of the contrast in an STM image comes from immobilized molecules at the surface: the highly mobile ones, like the solvent molecules, are rarely seen even if present on the surface.

STM images of **P2** domains are very detailed and even the alkyl tails of ISAOC18 molecules can be easily distinguished, since they are highly immobilized to being interdigitated. The individual ISA head and the coronene are visible too, for the same reason.

But STM images of **P1** (see Fig. 1b main text) have a striking peculiarity. In **P1** there are two different types of coronene, COR1 and COR2, that differ by their positions in the ISAOC18 network: COR1 coronene molecules sit at the center of the ISAOC18 hexamer while the COR2 ones are trapped between interdigitating hexamers that form triangular cavities (see Fig. 1c main text). What is puzzling is the fact that COR2 coronene molecules look slightly larger than the COR1 ones and, even more so, they are all consistently the brightest ones.

A large and fuzzy spot is often associated with a poor molecular immobilization but, if we look at the model, both types of coronene fit tightly into the ISAOC18 network and MD simulations show they have about the same mobility. Plus, this would not help much explain the different brightness observed in the images. To try explaining both the size and brightness of the spots seen in the STM images, we decided to entertain the possibility of a second coronene molecule stacked atop a COR2 coronene or to have the adsorption on the surface of a coronene-coronene dimer formed in

solution, rather than single coronene. This top molecule could have greater mobility than the one below, while interacting with it *via* π - π stacking, which could explain the increased contrast.

Either way, the question now becomes what makes COR1 and COR2 coronene molecules different and why only one of them is a dimer. To answer this question, we built models stacking a coronene molecules on top of a COR1 or a COR2 site in the **P1** polymorph to form a D1 or D2 dimers, respectively (Fig. S5a). Since the ISAOC18 alkyl tails have a zigzag “up-down” orientation on the surface (Fig. S5b), they are slightly taller than the coronene molecules in the monolayer, providing a small ridge that could help trapping the top molecules of both D1 and D2 dimers.

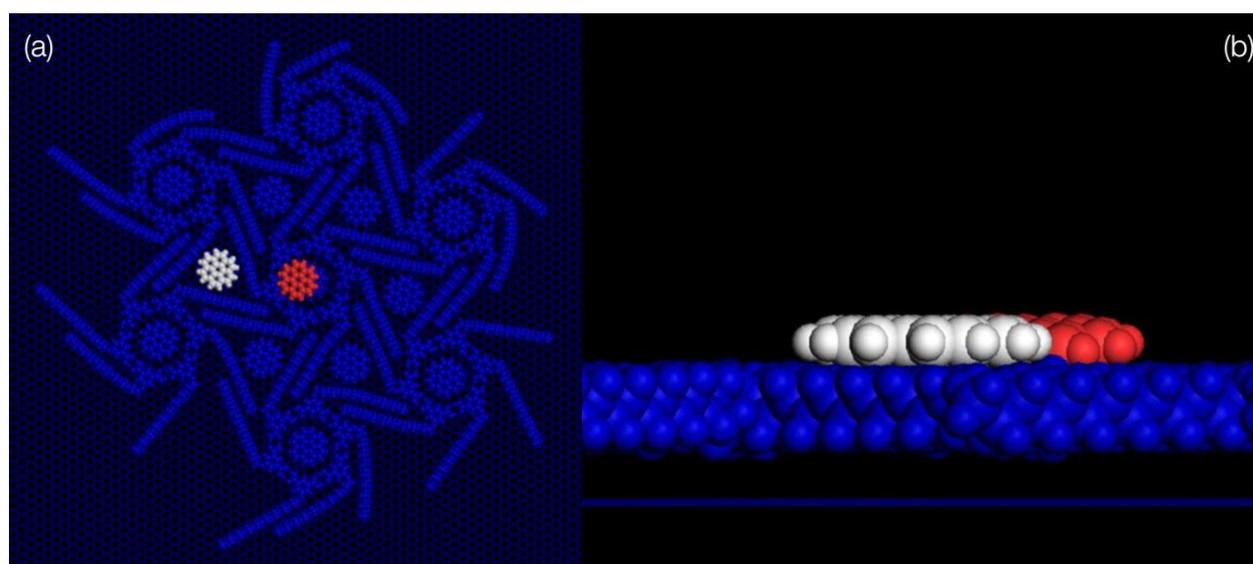


Fig. S5. (a) Top view showing the top molecule of the D1 (in red) and D2 (in white) dimers. (b) Side view showing how the top coronene molecule in the dimers sits slightly below the top part of the ISAOC18 tails (in blue).

While the strength of the intermolecular interactions within those dimers are comparable, dry MD simulations at room temperature show that the top molecule in D1 is more mobile than that in D2. The reason is that the top coronene in D1 has a larger area to explore before “hitting” the ridge provided by the ISAOC18 alky tails than that in D2, as the ISAOC18 heads are flat and do not help confining it.

But more interestingly, wet MD simulations at room temperature show that in just few hundreds of picoseconds the top coronene of the D1 dimer was able to jump over the ridge only to be immediately trapped for the rest of the 1ns-long MD simulation in a new, more stable, D2 dimer (Fig. S6). Also, during this simulation, the original D2 dimer remained intact.

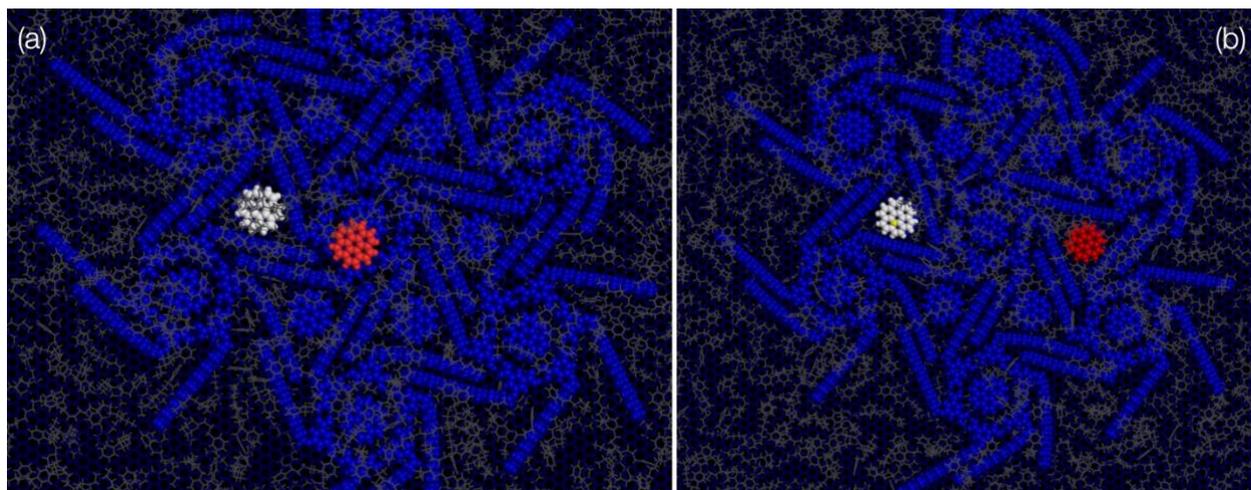


Fig. S6. (a) Beginning and (b) end of a 1-ns long wet MD simulation at room temperature. The solvent is color-coded in gray.

This suggests that only D2 dimers are stable, while the D1 dimers are unstable and quickly disappear, causing the difference in contrast observed in the STM images. And since in the **P2** polymorph we would only have the more unstable D1 dimers, this could also explain why coronene molecules are not brighter but have the same contrast than the rest of the monolayer.

B3: On the thermal stability of P1 and P2 on pristine HOPG

To validate the higher stability of **P2**, an annealing experiment was performed. Addition of an equimolar solution (0.5 mM) of coronene/ISAOC18 results in a monolayer where both polymorphs are present on the surface (Fig. S7a). After confirming the presence of both polymorphs with STM, the sample was annealed for 15 minutes at 80 °C and left to cool down for an additional 15 minutes. After adding a small droplet of 1-PO, STM revealed the surface to be solely covered by many small domains of **P2** (Fig. S7b). This result further corroborates the higher thermodynamic stability of **P2** compared to **P1** as inferred from the calculations.

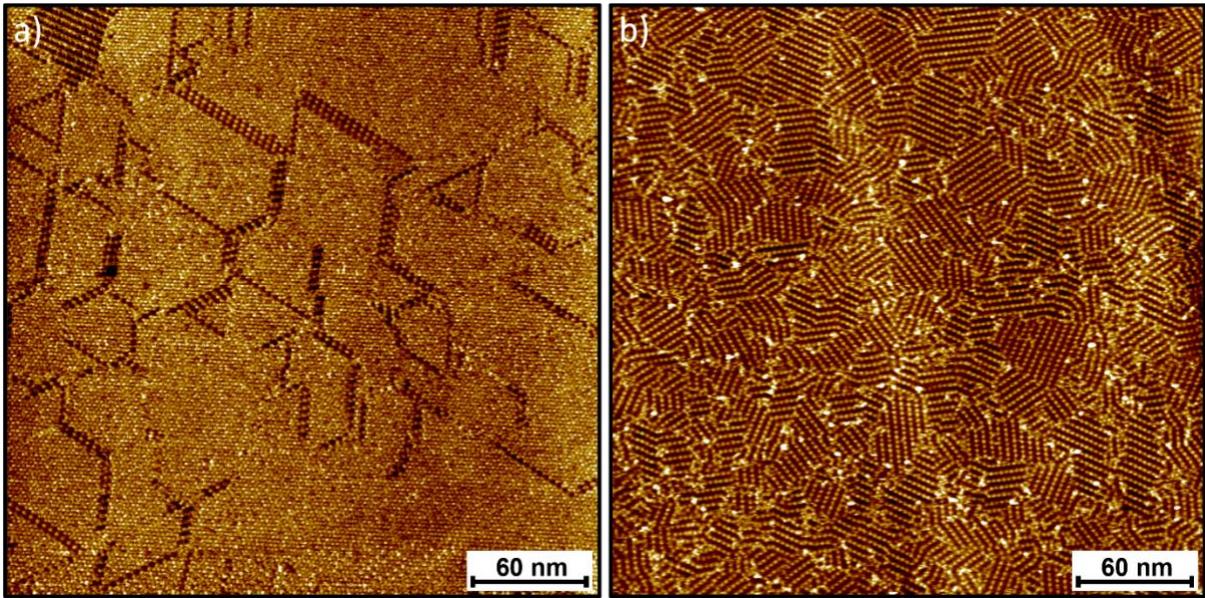


Fig. S7. STM images ($300 \times 300 \text{ nm}^2$) (a) before and (b) after annealing 15 minutes at $80 \text{ }^\circ\text{C}$.

B4: Growth of P1 and P2 inside nanocorrals

At relatively high COR/ISAOC18 ratio (> 0.1), filling of the corral with P1 and P2 domains could take several minutes to several tens of minutes and could therefore be followed by STM (Fig. S8). After nanoshaving, no voltage pulses were applied to the STM tip to avoid any influence on the assembly. Using the draw polygon shape in the SPIP software, the domains could be highlighted (P1 = blue, P2 = green) whilst calculating the surface coverage (nm^2). Videos following the growth for different corrals can be found [Insert reference for the videos].

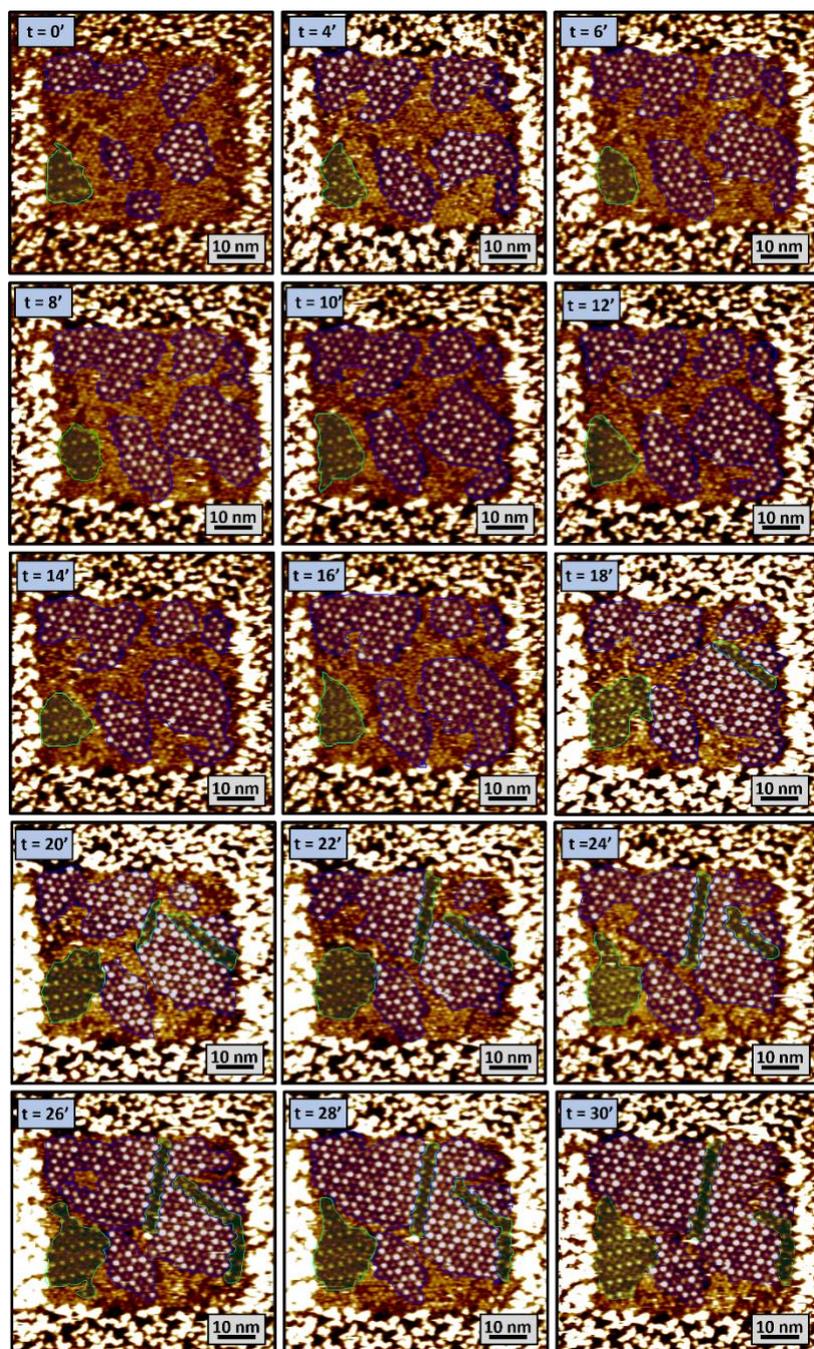


Fig. S8. Sequential STM images displaying the growth of P1 and P2 domains in corrals upon nanoshaving, highlighted in respectively blue and green, for a 0.5 mM/0.5 mM COR/ISAOC18 solution.

Averaging over 12 corrals, the increase in percentage of **P1**, **P2** and **P1 + P2** could be followed over time as well as the area. In Fig. S9a, it can be seen, starting from about 15 % occupation of the corral with **P1** and **P2** domains, the coverage increases gradually over time to reach about 60 %. Following the area occupied by **P1** and **P2** domains over time and applying a linear fit, shown in Fig. S9b, a higher growth rate for **P1** could be found ($25.4 \pm 1.5 \text{ nm}^2/\text{min}$ vs. $18.2 \pm 1.0 \text{ nm}^2/\text{min}$).

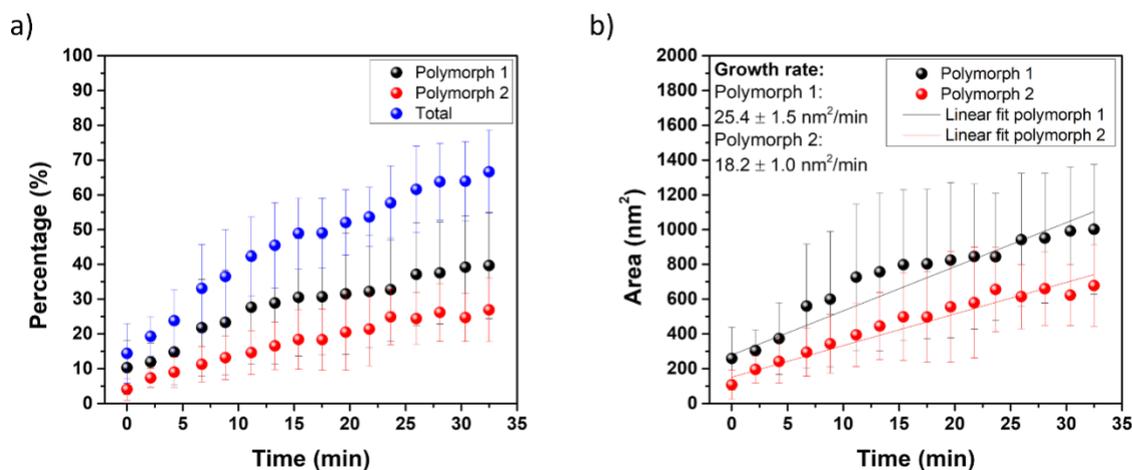


Fig. S9. Growth of **P1** and **P2** over time in corrals for 0.5 mM/0.5 mM COR/ISAOC18. (a) Increase in corral coverage (%) **P1** (black), **P2** (red) and **P1 + P2** (blue). (b) Increase in area (nm²) for **P1** (black) and **P2** (red).

B5: Self-assembly of COR in the absence of ISAOC18

To verify whether ISAOC18 is necessary to observe coronene molecules immobilized in the early self-assembly, 0.5 mM coronene in 1-phenyloctane was drop casted on a modified HOPG substrate. Upon nanoshaving, no COR self-assembly was observed. Instead, columnar self-assembly could be imaged. We ascribe this to contamination that gets adsorbed into the corrals (Fig. S10), which is often seen for nanoshaving experiments with low solute concentrations or if the solute does not adsorb strongly and might originate from the grafting procedure or is present in the solvent itself. Note that, in the present case, although COR adsorbs strongly relative to ISAOC18, it is probably mobile and is “pushed around” by the STM tip in the absence of any templating molecule. This is possibly related to the poor lateral interactions, which are mostly expected to be van der Waals based, holding the self-assembled monolayer of COR together.

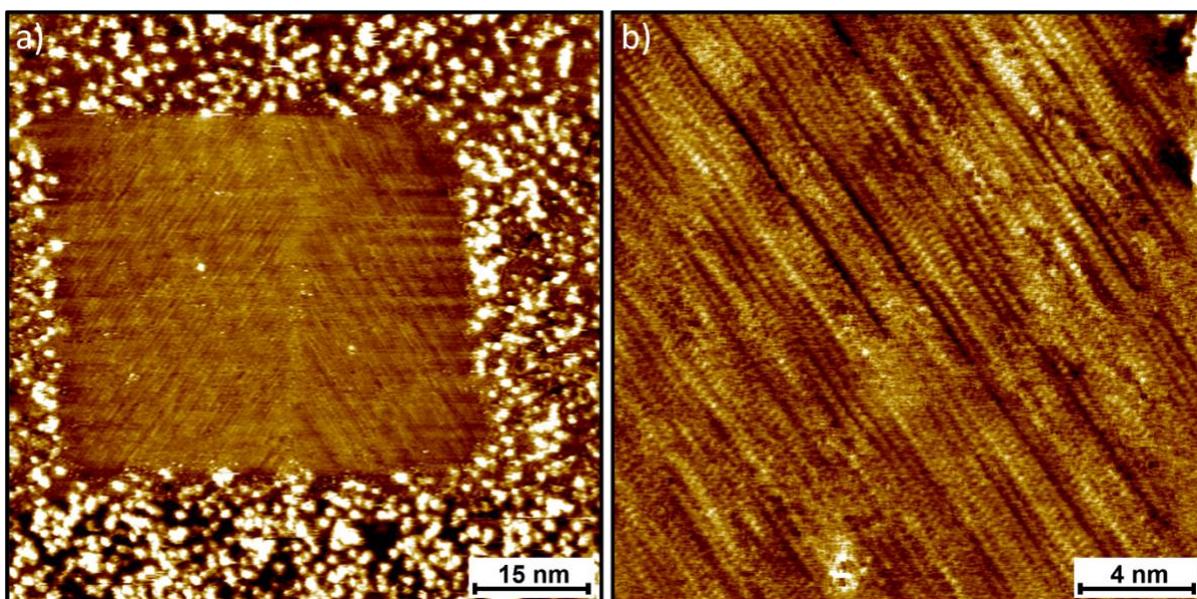


Fig. S10. STM images displaying linear row-like assembly after nanoshaving in a 0.5 mM coronene solution in 1-phenyloctane, which we ascribe to the adsorption of contamination.

A lower grafting density (0.5 mM TBD) was used to potentially avoid such contamination and possibly immobilize the coronene assembly in between. While the contamination was not present, potentially blocking the surface for coronene assembly, no indication for the semi-ordered coronene assembly was observed (Fig. S11), which indicates that the observed early self-assembly in corrals (see Fig. 2c-e main text) cannot consist of coronene assembled on its own.

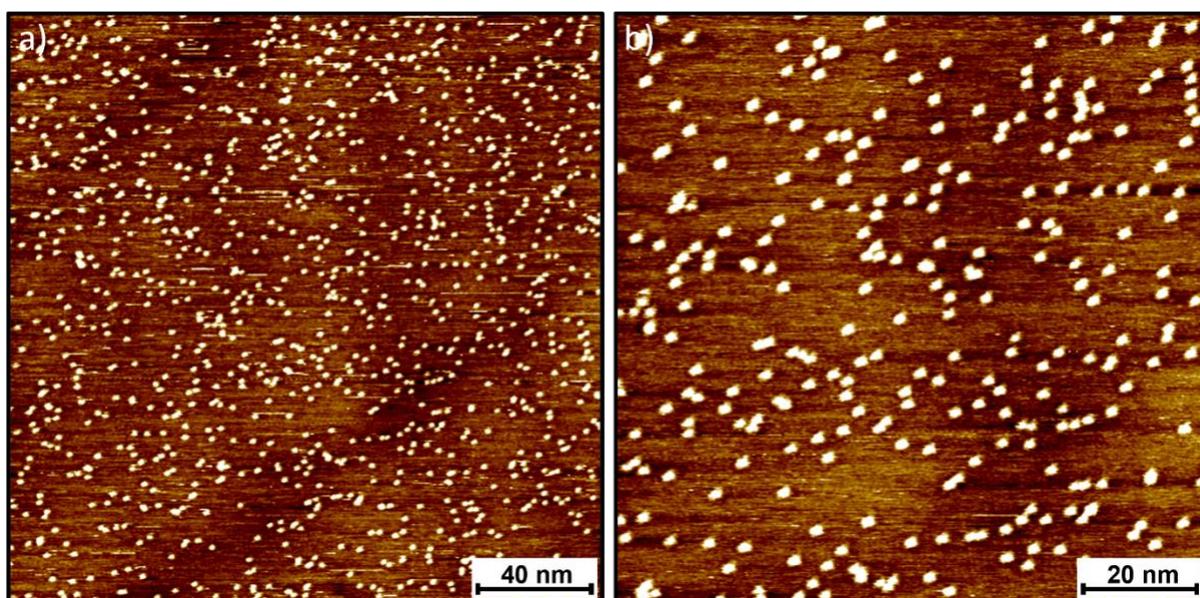


Fig. S11. STM images highlighting a medium density grafted HOPG (0.5 mM TBD) with no visual assembly in the open spaces, measured in 0.5 mM coronene in 1-phenyloctane.

The fact that the embryonic clusters are observed only under nanoconfinement is plausibly related to the reduced lateral diffusion in the corrals, originating from its sidewalls acting as barriers. This slows down the nucleation and growth of P1 and P2 (see Figure 2c, d, main text) drastically such that it takes several tens of minutes for the corral to fill up. Also, the slower adsorption/diffusion of ISA-OC18/COR host-guest nuclei is also related to the fast adsorption of coronene in the cavities. The combined effect is the trapping of coronene in between meandering isophthalic acid molecules, which allows their observation with STM. On pristine surface however, no such lateral barriers exist decreasing the possibility of observation of the embryonic stages of the assembly process.

We also suspect that the amount of time it takes between deposition and STM observation could also be a factor. In the case of a typical experiment involving deposition of the solution on pristine graphite and subsequent imaging with STM, it takes typically several minutes (15-20 minutes under ideal conditions) to get an image, whereas in the nanoshaving experiments, the time interval between corral creation and subsequent imaging, is relatively shorter (2-3 minutes). Thus, a combination of slower lateral diffusion in corrals and the ability of imaging quickly, compared to that on pristine graphite leads to possibility of imaging the early stages of self-assembly under nanoconfinement conditions.

B6: Molecular adsorption energies of coronene and ISAOC18 on HOPG

To obtain more insight into the experimental observations the modeling strategy was aimed at getting to know how the different molecules behave and to identify the crucial interactions that take place in the systems under study.

Let us first consider how strongly the different molecules adsorb on pristine HOPG graphite. Such adsorption energy, AE , is calculated as the difference between the total potential energy of the system where a single molecule is adsorbed on the surface, E_{sys} , and the sum of the potential energies for the isolated molecule, E_{mol} , and isolated surface E_{surf} :

$$AE = E_{sys} - (E_{mol} + E_{surf}) \quad (\text{Eq. 1})$$

We should note here that being the surface treated as a rigid body, its potential energy E_{surf} is zero.

We also calculate the adsorption energy per unit area (E_{ads}). For this we estimate the molecular footprint as the area of the surface that is not accessible to a solvent modelled as a spherical probe having a radius of 1.4Å. These Solvent Accessible Surface Area (SASA) estimations were performed with VMD and the results are reported in Table S3.

Table S3. Comparison between molecular footprint, adsorption energy, AE , and density of adsorption energy, E_{ads} , for the three molecular species involved in this work.

Molecule	Footprint (nm ²)	AE (kcal mol ⁻¹)	E_{ads} (kcal mol ⁻¹ nm ⁻²)
ISAOC18	2.46	-77.0	-31.3
CORONENE	1.45	-50.3	-34.7
PHENYLOCTANE	1.30	-33.3	-25.6

Of the three molecular species considered here, 1-phenyloctane has, as expected, the least tendency to adsorb on the surface. Being the largest of the three molecules, it is not surprising that ISAOC18 has the strongest adsorption energy per molecule, while coronene is the molecule with the largest adsorption energy per nm² of surface coverage of all.

Next, we want to know which molecule between coronene and ISAOC18 has the best affinity with the solvent: this will give us an indication on the propensity of a molecule to move from the surface back into the solvent phase and vice versa. Two models have been built (Fig. S12), one for ISAOC18 and one for coronene where a single molecule was soaked into the liquid phase. This phase was modelled with 164 molecules of 1-phenyloctane, equilibrated at room pressure and temperature (resulting in a theoretical density of 0.8 g/cm³, in good agreement with the experimental value of 0.85g/cm³).

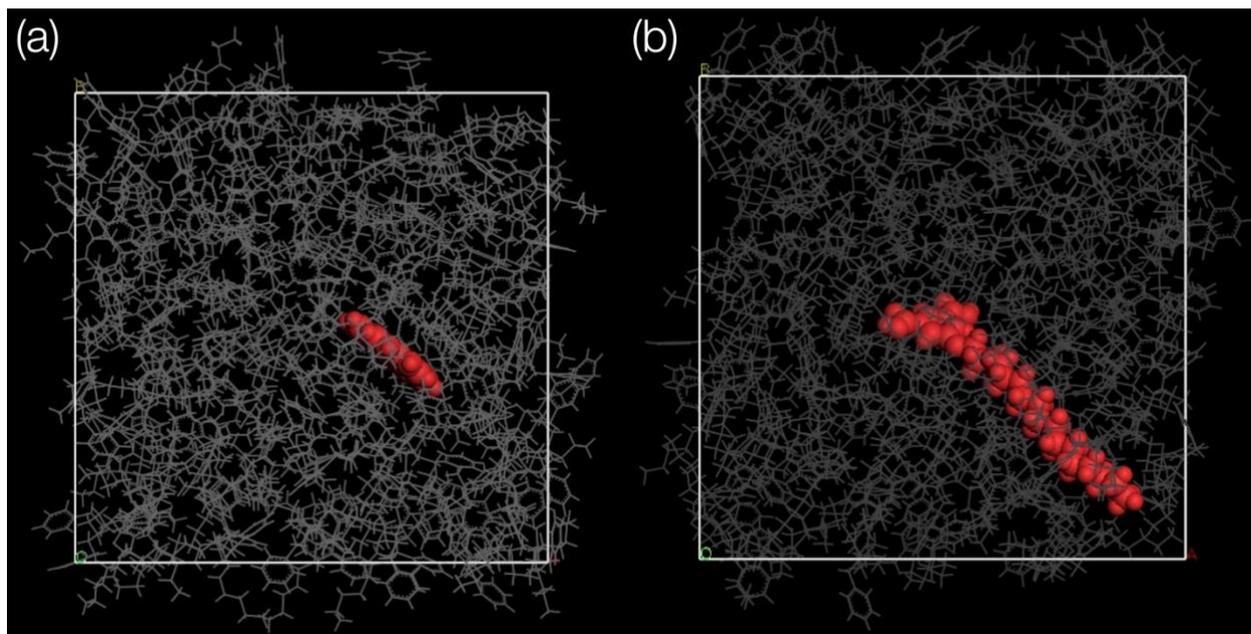


Fig. S12. Solvated coronene (a) and ISAOC18 (b). For the sake of clarity, 1-phenyloctane molecules are color-coded in gray while ISAOC18 and coronene molecules are shown in red.

The molecular energy of solvation, E_{solv} , was calculated with Eq.2:

$$E_{solv} = E_{sys} - (E_{mol} + E_{solvent}) \quad (\text{Eq. 2})$$

where E_{sys} is total potential energy of the system, and E_{mol} and $E_{solvent}$ are the potential energies for the non-interacting molecule and solvent phase, respectively. For all the energy terms used in Eq. 2 we consider their average values collected during a 1ns-long MD simulation at 300K and 1 atm.

These results show that ISAOC18 has a much better affinity with the solvent than coronene, being E_{solv} for ISAOC18 and coronene equal to $-163.0 \text{ kcal mol}^{-1}$ and $-103.7 \text{ kcal mol}^{-1}$, respectively. Also considering the much higher flexibility of the ISAOC18 molecules compared to coronene, we can suggest that, despite having the higher adsorption energy per molecule, ISAOC18 molecules can easily (partially and even totally) desorb from the surface, making them a much more dynamic system than coronene.

To test this hypothesis, we consider a system composed of 28 coronene and 24 ISAOC18 molecules adsorbed on periodic HOGP graphite (treated as a rigid body) and “wetted” with 261 solvent molecules. After the initial minimization to relax the system, we perform a 500ps-long MD simulation at 600K, examining coronene/ISAOC18 desorption (such high temperature is used to speed up the dynamics and keep the computational effort to a manageable level).

Despite the high temperature used and the presence of explicit solvent, coronene molecules never fully (nor partially) desorb, while ISAOC18 molecules can be seen leaving the surface and entering

the liquid phase, shown in Fig. S13 (same figure as in Fig 3 in the main text, provided here as an enlarged version for the sake of clarity).

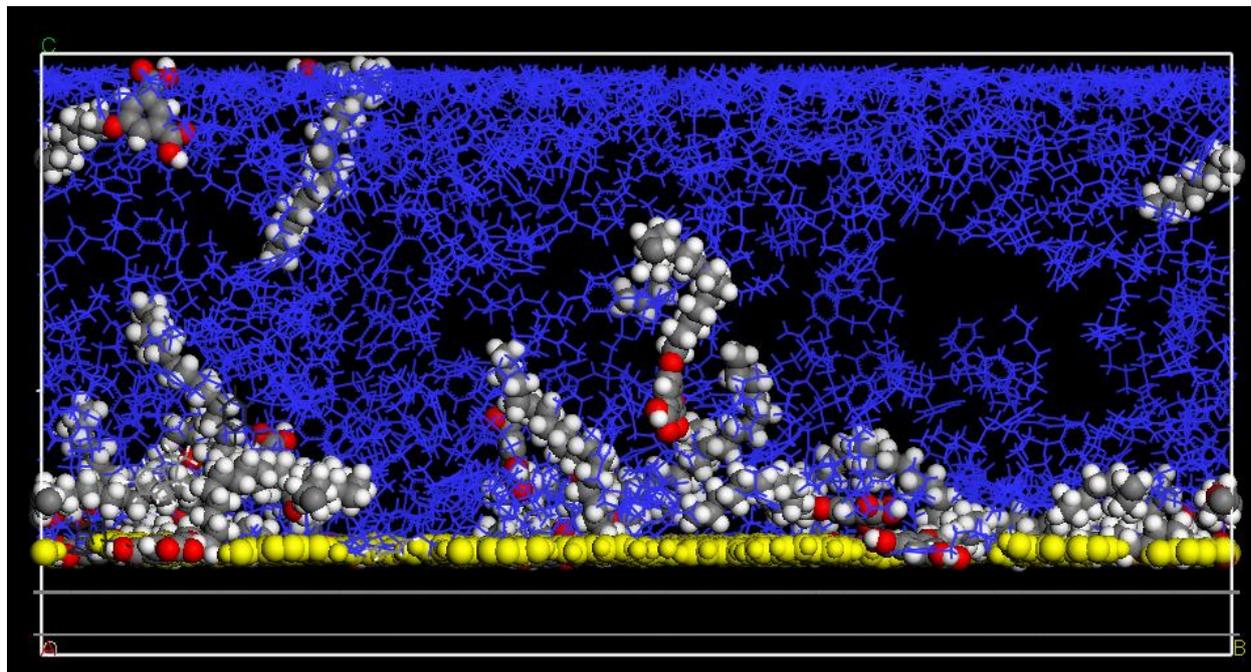


Fig. S13. At least 3 ISAOC18 molecules are seen to totally desorb from the surface, with many more being partially into the liquid phase. Coronene molecules are shown in yellow while solvent molecules are in blue.

It must be noted that ISAOC18 molecules have also been observed to desorb during MD simulations conducted at room temperature, thus supporting the idea they form a dynamic system.

This supports our energetic analysis, indicating that the solvent molecules at the surface, having low AE, can easily be replaced by either coronene or ISAOC18 molecules. And that despite ISAOC18 having the strongest AE per molecule, its flexibility and higher affinity with the solvent phase make it a much more dynamic molecule than coronene, being able to desorb and re-adsorb on the surface. Finally, it seems clear that coronene, once adsorbed on surface, sticks to it and it cannot be easily replaced by either solvent or an ISAOC18 molecules.

B7: On P1 & P2 polymorphs formation

High-resolution STM images can provide a direct observation of a monolayer morphology, and measurements performed in nanocorrals can let us peek into the early stages of formation of those polymorphs (Fig. 2, main text). In general, to study the mechanisms that lead to the formation of a monolayer one should design a set of experiments to observe, for example, how changing the solution concentration, relative abundance of the different molecules and experimental conditions alter the monolayer morphology and/or its relative abundance.

In this work it has been observed that an excess of coronene in solution, as well as an increase(?) in the degree of confinement at the surface via slow creation of nanocorrals, promotes the formation of **P1** with respect to **P2**. We then turned to modelling to shed light on the mechanism.

We have already shown that, in the presence of solvent, ISAOC18 molecules are quite dynamic, being able to partially (and even fully) desorb from the surface, while we have never observed in our simulations the desorption of coronene molecules (Fig. S13). Since coronene is likely to be the most stable molecule on the surface, it seems reasonable to hypothesize that a slow-growing nanocorral will favor on the exposed HOPG surface a local excess of coronene, hence promoting the same polymorph one would get from using an excess of coronene molecules in solution.

To investigate that situation, we modeled a mixture of coronene and ISAOC18 molecules in a 1:1 ratio (28 molecules of each species) on a comparatively large surface, which represents a low adsorption density. After the initial minimization (Fig. S14a) followed by a 10-ns long MD simulation at room temperature, we observe an evolution towards phase separation of the two components. As shown in Fig. 14b, ISAOC18 molecules cluster together under the strong vdW interactions between their tails, forming quite ordered aggregates with interdigitated alkyl tails and extended H-bonding networks between ISA heads. In a way, they look like precursor of the classic ISAOC18 monolayer usually observed at the graphite/liquid interface.²⁻⁷ Coronene molecules, on the other hand, are not often included in the ISAOC18 clusters and are mostly located in a different region of the surface.

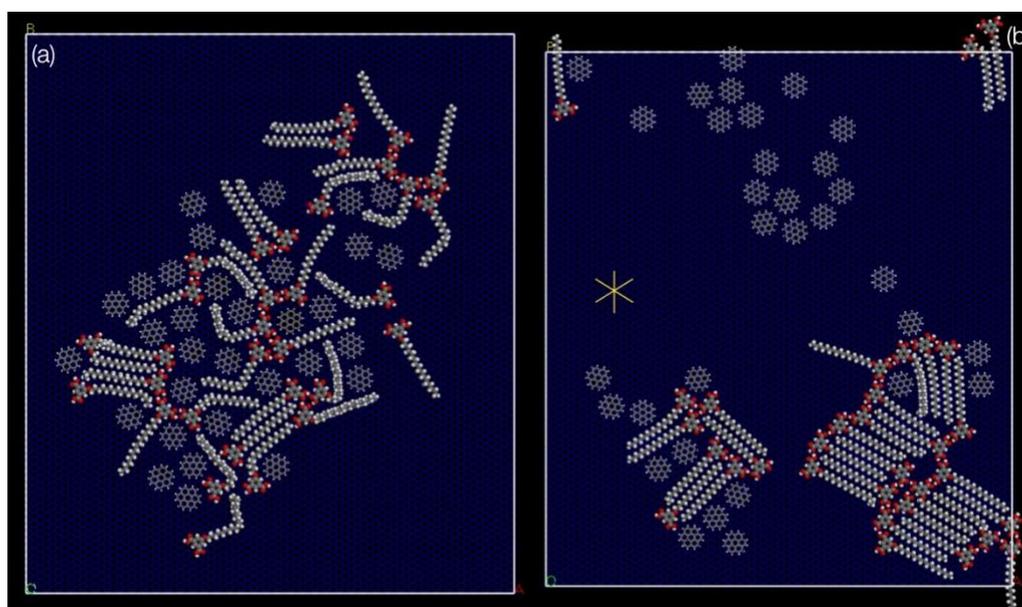


Fig. S14. From an initial mix of molecules on the surface (a), separate clusters of coronene and ISAOC18 tend to emerge (b) when a low density, 1:1 molecular ratio is used. The main axes of graphite are shown by the yellow lines in panel b.

Thanks to their long alkyl tail, the ISA molecules tend to align to the main axes of graphite (shown in yellow in Fig. S14b), forming domains that are rotated with respect to each other by 60°. Coronene molecules, on the other hand, show a high mobility on the surface, but no preferential orientations on graphite and form close packed aggregates. In those aggregates, the distance between the centers of adjacent coronene molecules is about 1.2 nm.

Fig. S15 shows the effect of increasing the molecular density at the surface while keeping the same 1:1 ratio. Since this high molecular density at the surface prevents phase separation, most of the ISAOC18 molecules are forced to “reptate” around the coronene molecules. Because separate domains cannot be formed, different proto-structures begin to emerge. We can recognize aggregates that could evolve into the polymorph building blocks (the ISAOC18 hexamer with a coronene in the centre) and others that could promote the formation of **P1** and **P2** polymorphs.

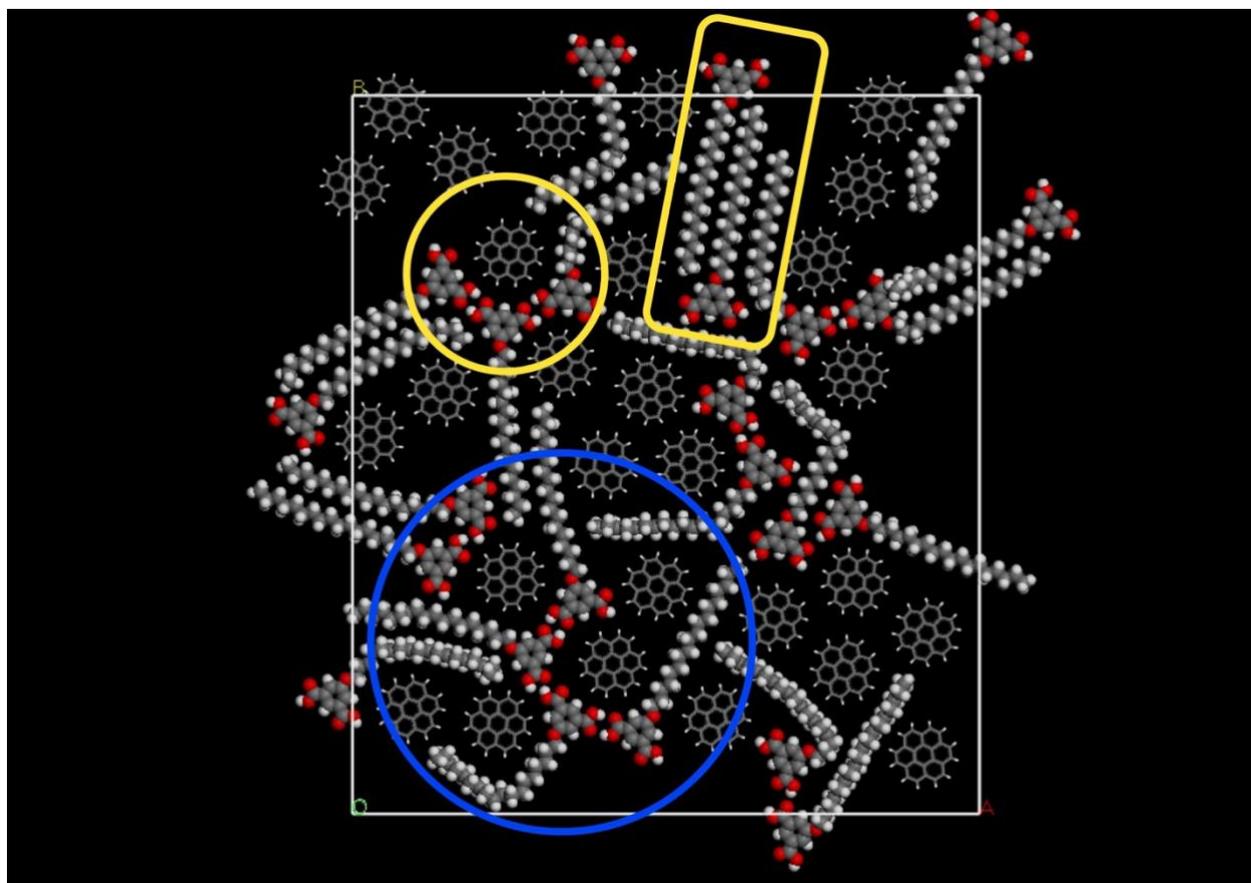


Fig. S15. By modelling ISAOC18 and coronene molecules at high adsorption density in a 1:1 ratio we can identify precursor structures to the polymorphs building blocks, circled in yellow, and areas that could evolve either towards **P1**, circled in blue, or towards **P2** (in the yellow rectangle).

It is also worth noting that the coronene-coronene separation is increased to about 1.5-1.7 nm when separated by one ISAOC18 molecule (either by the head or the tail of the molecule). This is

in agreement with the STM measurements on early-stage assembly inside nanocorrals, shown in Fig. 2e of the main text.

By increasing the amount of coronene at the surface with respect to the number of ISAOC18 (Fig. S16a), it is unlikely for ISAOC18 molecules to be able to form extended domains of interdigitating molecules. Such conditions should favor the formation of the **P1** polymorph. In contrast, by increasing the amount of ISAOC18 (Fig. S16b) we favor the formation of close-packed ISAOC18 aggregates, possibly promoting the formation of the **P2** polymorph.

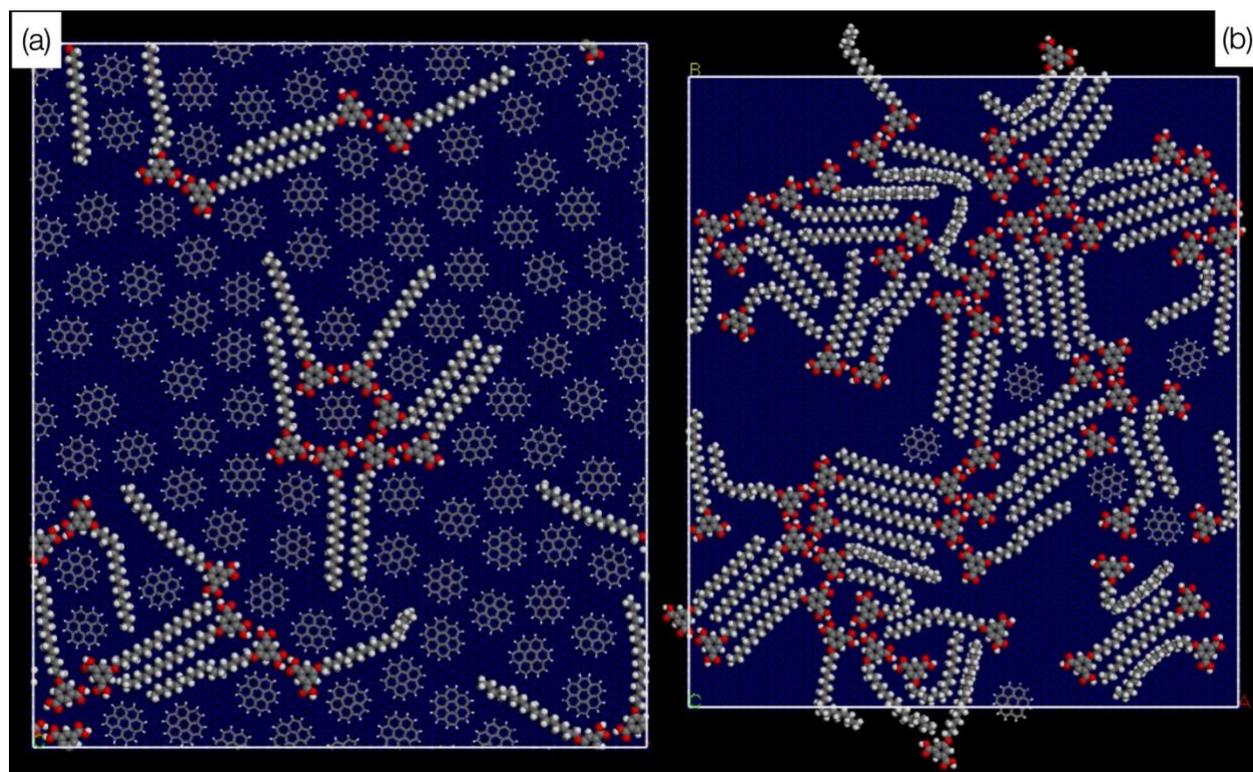


Fig. S16. The effect on molecular aggregation at the surface of introducing an excess of (a) coronene or (b) ISAOC18 molecules in the system.

By considering the stronger adsorption of coronene molecules on the surface, their tendency to stick longer to it than ISAOC18, and the results from with different coronene/ISAOC18 ratio, we can propose the following mechanism for the nucleation and growth of the observed polymorph. If the ISAOC18 molecules cannot diffuse sufficiently around coronene molecules (either because they are trapped in coronene aggregates or because of nanoconfinement), large aggregates of closely-packed ISAOC18 molecules cannot form. In this scenario, it seems more likely that the H-bonds between the ISA heads become dominant (with respect to side-by-side alignment of the alkyl chains), thus favoring the formation of ISAOC18 hexamers encompassing a coronene molecule. Eventually, these aggregates will interact with each other and interdigitate to give rise to the **P1** polymorph. In contrast, if the local concentration of coronene is sufficiently low, the

vdW interactions between the ISAOC18 tails will be dominant, thus driving the formation of most favorable close-packed structures. In this scenario, the most energetic favorable way for ISAOC18 to incorporate the coronene molecules in the close-packed aggregate is to encompass them by building the cyclic hexamer around them, forming the intermolecular H-bond network between the ISA heads while keeping a high degree of ISAOC18 tail interdigitation. This would result in the more stable, and dense, **P2** polymorph.

B8: On the effect of the COR/ISAOC18 ratio

Depending on the COR/ISAOC18 ratio in solution, in the initial state right after nanoshaving, the corral was either completely filled with 2-component polymorphs (Fig. S17a, ratio of 0.1), had an intermediate coverage (Fig. S17b, ratio of 0.6) or contained only small domains (Fig. S17c-d, ratio of 1-1.67). For growth of the 2-component domains over time see Fig. S8/ supporting videos.

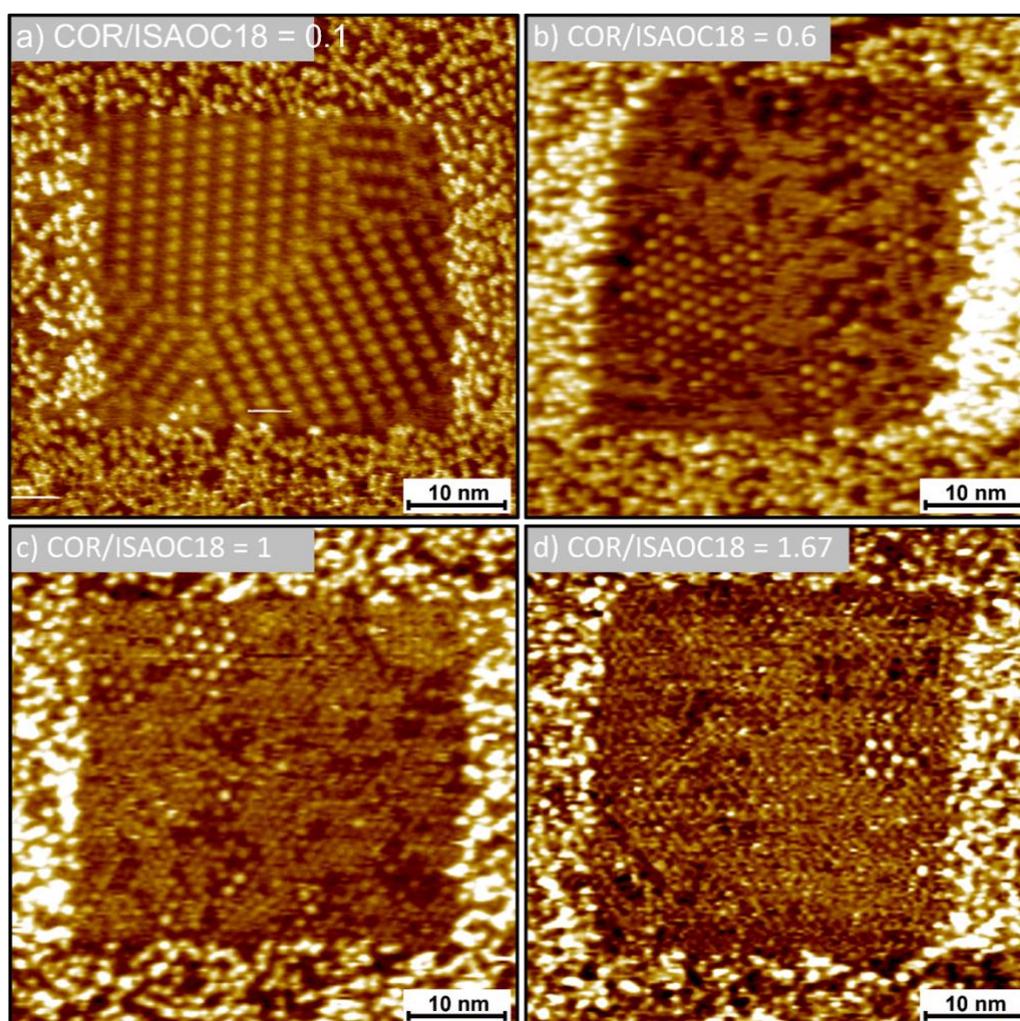


Fig. S17. STM images immediately after nanoshaving for a COR/ISAOC18 ratio of (a) 0.1, (b) 0.6, (c) 1 and (d) 1.67, highlighting the decrease in initial corral filling with increasing COR/ISAOC18 ratio.

An initial step to study the effect of confinement on the polymorphic preference was carried out by investigating, in corrals, the same COR/ISAOC18 ratios (0.1, 0.6, 1, 1.67) that were studied on pristine HOPG. To effectively compare the relative coverages of **P1** and **P2** inside corrals for the different COR/ISAOC18 ratios, the analysis was only performed once the 2-component domains inside reached a coverage of more than 50 %, which took a few minutes for the intermediate (0.6) and about 20 minutes for high (1-1.67) COR/ISAOC18 ratios. For the confinement experiments, the average was taken over 40 corrals of $50 \times 50 \text{ nm}^2$ at a nanoshaving speed of $0.4 \text{ }\mu\text{m/s}$. Calculation of the polymorphic surface coverages for each COR/ISAOC18 ratio on pristine HOPG was performed by probing 5 different macroscopic areas, by physically moving the STM sample plate, and taking at least three $300 \times 300 \text{ nm}^2$ images. SPIP draw polygon shape was used to calculate the area occupied by the domains. No strong effect of confinement was observed.

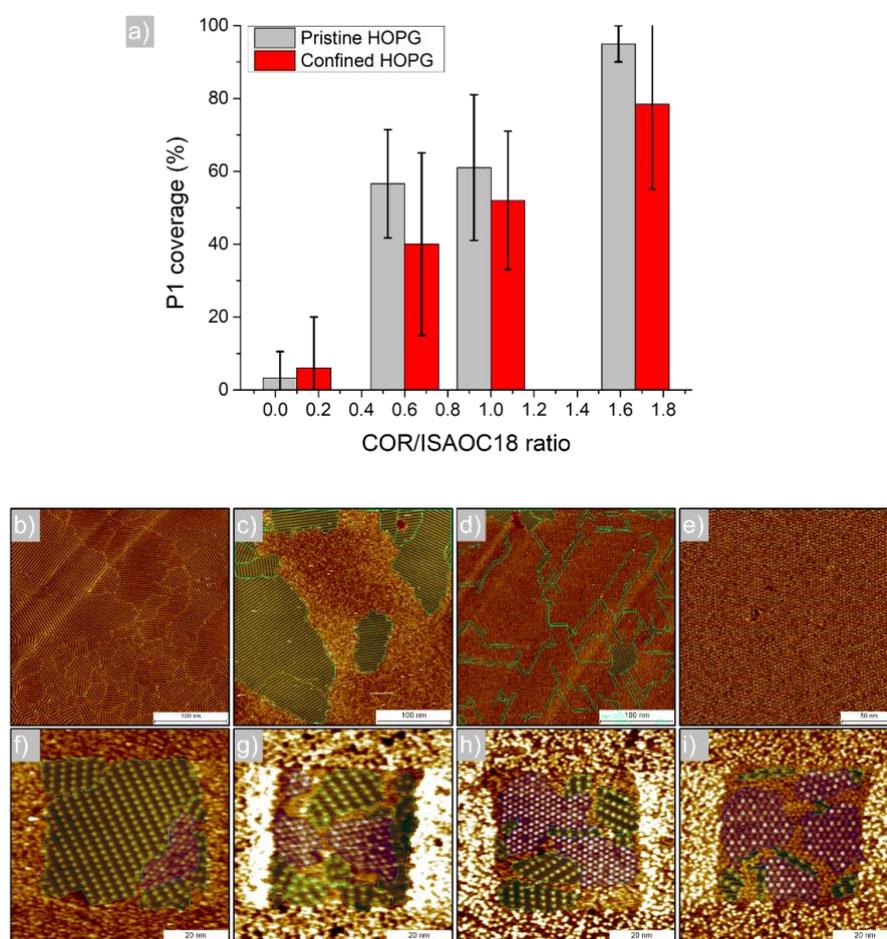


Fig. S18. (a) Influence of COR/ISAOC18 ratio on the **P1** coverage (%) for pristine HOPG (grey, Fig. 4a in the main text) and in corrals (red). Representative STM images for a COR/ISAOC18 ratio (b, f) 0.1, (c, g) 0.6, (d, h) 1 and (e, i) 1.67. P2 domains are highlighted in green for pristine HOPG (b-e). **P1** and P2 domains are highlighted in respectively green and blue in corrals (f-i), respectively.

Finally, the effect of the total concentration of COR + ISAOC18 in solution was investigated by decreasing the total concentration of the equimolar COR/ISAOC18 solution ($C_{\text{tot}} = 1 \text{ mM}$) by 25 % and 40 %. Lower total concentrations did not produce self-assembly on pristine nor confined HOPG.

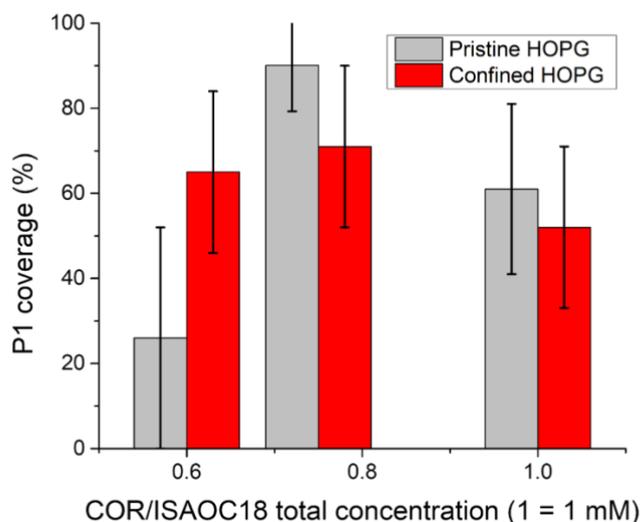


Fig. S19. Influence of COR/ISAOC18 total concentration.

B9: On the effect of the size of the nanocorrals

Fig. S20 provided below clearly shows the influence of the size of the nanocorrals on the outcome of the self-assembly process of the COR/ISA-OC18 system. It is evident that for smaller nanocorrals the surface coverage of the two polymorphs is low. Only 40-50% of the corral area is covered by P1 and P2 immediately after the completion of the nanoshaving process. On the other hand, a full surface coverage comprising both P1 and P2 was observed in the first STM image obtained for the $150 \times 150 \text{ nm}^2$ corrals.

The difference in surface coverage after nanoshaving could be related to the thickness of the nanochannel formed in the initial stages of nanoshaving. While each second of the nanoshaving process exposes the same area in both cases, the thickness of the nanochannel is significantly reduced for the larger nanocorrals since it needs to travel longer distances before passing onto the next scan line. Therefore, the growth of the nanochannel in the slow nanoshaving direction is more gradual for larger corrals compared to smaller. Since nuclei need to overcome a critical size to grow, it is not unreasonable that a more gradual increase in the nanochannel thickness results in a more gradual crossing of this size, resulting in a localization of the nuclei in the initial exposed

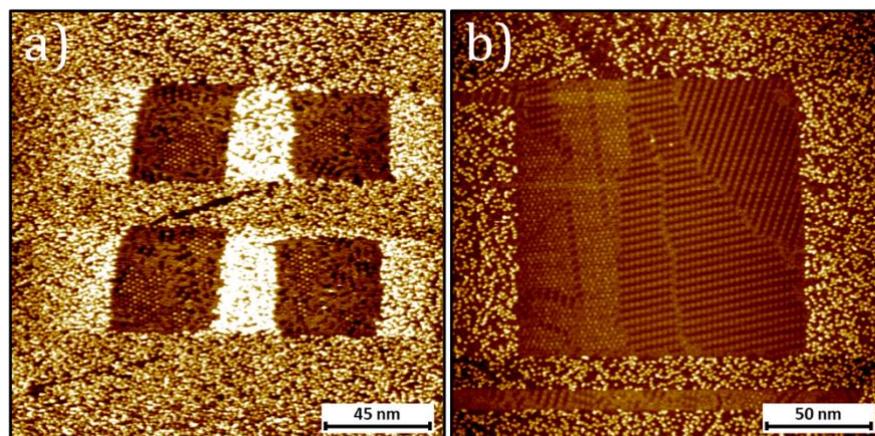


Fig. S20. Influence of corral size on the outcome of the COR/ISA-OC18 assembly. STM images represent the first scan after the nanoshaving was completed. (a) Corral size $\sim (50 \text{ nm} \times 50 \text{ nm})$. (b) Corral size $\sim (150 \text{ nm} \times 150 \text{ nm})$. ($C_{\text{ISA}} = 0.5 \text{ mM}$, $C_{\text{CoR}} = 0.3 \text{ mM}$).

graphite, followed by subsequent growth of the nuclei upon further nanoshaving leading to a full surface coverage. For smaller corrals, characterized by a relatively faster nanoshaving along the slow scan direction, the critical nucleus size can already be surpassed before nucleation takes place and could explain the nucleation of P1 and P2 at random locations in the corral from the embryonic mixture of coronene and ISAOC18 occupying the space.

C. References

- 1 S. Valiyaveetil, C. Gans, M. Klapper, R. Gereke and K. Müllen, *Polym. Bull.*, 1995, **34**, 13–19.
- 2 A. M. Bragança, A. Minoia, R. Steeno, J. Seibel, B. E. Hirsch, L. Verstraete, O. Ivasenko, K. Müllen, K. S. Mali, R. Lazzaroni and S. De Feyter, *J. Am. Chem. Soc.*, 2021, **143**, 11080–11087.
- 3 F. Tao and S. L. Bernasek, *J. Am. Chem. Soc.*, 2005, **127**, 12750–12751.
- 4 F. Tao and S. L. Bernasek, *Surf. Sci.*, 2007, **601**, 2284–2290.
- 5 P. N. Dickerson, A. M. Hibberd, N. Oncel and S. L. Bernasek, *Langmuir*, 2010, **26**, 18155–18161.
- 6 S. De Feyter, A. Gesquière, M. Klapper, K. Müllen and F. C. De Schryver, *Nano Lett.*, 2003, **3**, 1485–1488.
- 7 K. W. Park, J. Adisojoso, J. Plas, J. Hong, K. Müllen and S. De Feyter, *Langmuir*, 2014, **30**, 15206–15211.