Electronic Supporting Information

Assessing the role of chirality in the formation of rosette-like supramolecular assemblies on surfaces

Andrea Minoia*[a], Zongxia Guo[b], Hong Xu[b], S. J. George[c‡],Albertus P. H. J. Schenning[c], Steven De Feyter[b], Roberto Lazzaroni[a]

Modeling protocol: the MD/quench iterative scheme

The MM/MD simulations have been performed with the Forcite module implemented in the Accelrys Materials Studio 5.0 (MS) modeling package.[1] Among the force fields available in MS, we choose Dreiding[2] because it provides an accurate description for H-bonding, which is crucial for describing the interactions between the OPV3T molecules in the rosette-like assemblies.

The simulation cell used is hexagonal and its size is tailored in such a way that the rosettes and their assemblies are isolated aggregates on the graphite surface, which is periodic. Since the Dreiding force field does not use partial atomic charges for the atoms of graphite, non-bonding interactions are described in terms of van der Waals (vdW) interactions. The non-bonding interactions have been cutoff at a distance of 1.5nm and have been treated using the Ewald summation method.
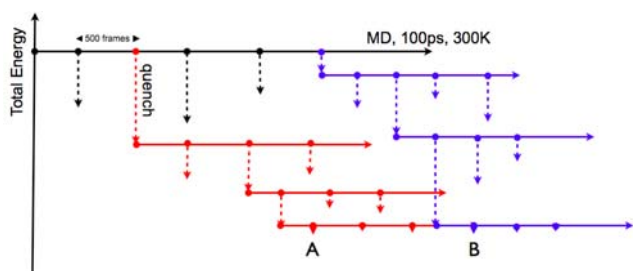

Fig. S1 The scheme illustrates the modeling strategy to evaluate the stabilities of a number of independent structures for each assembly studied, through a series of MD runs and quenching steps. The scheme is iterated until the stability of all the different sampled geometries converge.

Fig. S1 well illustrates the iterative MD/quench modeling strategy used to sample the stability of a number of independent geometries for a given assembly: either an assembly of six molecules, i.e., a rosette, or an assembly of rosettes, i.e, a fraction of a monolayer. For each system considered, an optimized starting geometry is used as starting point for the first 100ps-long MD run. Since a time step of 1fs is used, the resulting MD trajectory contains 100000 different structures, or frames. Every 500 frames, a structure is quenched (its geometry is fully optimised at zero Kelvin) to evaluate its stability based on its total energy. Next, the lowest- and highest-energy structures, chosen among the 200 that have been optimised, are used as starting points for the new MD runs, followed by the quenching procedure; this is represented by the red and blue paths in Fig. S1. From the second iteration on, only the low-energy structure is used as starting point for the next MD runs. The process is iterated until a stable value for the energy of the quenched structure is found. We then consider this structure as a stable structure for the system, and its energy gives the stability of the aggregate. The reason to apply such iterative MD/quench scheme to the low- and high-energy structures from the first MD run is to ensure that the very complex potential energy landscape of the aggregate has been properly sampled.

In table 1, we report the energy decomposition(potential (PE), valence (VE), van der Waals (vdW), H-bonding (HB) and binding (BE) energies) for the stable structures for the LoE and HiE aggregates. LoE and HiE identify the low-energy and hi-energy structures extracted from the first MD run in the MD/quench scheme. Each structure has then been optimized independently. The energy for the LoE and HiE structures for each aggregate converge to very similar value, meaning we have correctly sampled the aggregates' potential energy surfaces. All the MD runs have been performed at room temperature in the NVT ensemble.

Table 1 Energy decomposition (potential (PE), valence (VE), van der Waals (vdW), H-bonding (HB) and binding (BE) energies) of the final structures for the rosettes of chiral OPV3T molecules. All energies are expressed in kcal/mol (1kcal/mol = 4.184 kJ/mol). LoE and HiE referred to the initial structure used in the MD/quench scheme: LoE and HiE indicate the starting point were the low and high energy structures for the considered rosette, respectively.

Structure	PE	VE	vdW	HB	BE
S1 LoE	132.6	227.0	-68.5	-25.9	-4.7
S1 HiE	133.9	225.4	-65.9	-25.4	-4.5
S2F LoE	171.1	244.6	-47.2	-25.5	-5.3
S2F HiE	171.1	239.5	-42.3	-25.3	-5.3
S2 LoE	140.9	229.1	-43.3	-25.7	-3.3
S2 HiE	141.1	230.2	-51.0	-25.3	-3.3
S1F LoE	173.6	242.4	-62.5	-26.3	-4.9
S1F HiE	169.1	245.6	-63.9	-26.1	-5.6

Because we were not interested to follow the time evolution of the systems under study the temperature was controlled by using the direct velocity scale algorithm, as implemented in MS. Once the MD/quench scheme has converged, all the geometrical constraints are removed and the structure optimized once again to evaluate the stability of the system.

Approximations used in the model
Some approximations have been introduced in the model, in order to reduce its complexity. Despite the experiment is conducted at the solvent/substrate interface, we have not taken explicitly into account the solvent since we do not aim to study the dynamics of formation of the rosettes and their self-assembling, but rather to obtain their energetic stabilities. Furthermore, no solvent/molecule co-adsorption takes place at the surface.
In order to obtain a better estimation of the stability of the rosettes, we neglect the energetic noise coming from the conformational degrees of freedom in the three dodecyloxy end groups of each OPV molecule. Their main effect is to improve the adhesion of the molecules to the surface via the so-called CH-π interactions (about 1.0 kcal/mol per CH-π interaction)[3] when assembling at the solvent/graphite interface. Since no solvent is included in our models, the molecules, once adsorbed, have no tendency to desorb from the surface. Furthermore, when modeling the isolated rosette, the presence of the three dodecyloxy end groups does not play a role in determining the internal structure of the rosettes, nor they can affect the stability of the hydrogen-bond pattern found at the core of the rosette. For these reasons, we have replaced each dodecyloxy group by an ethoxy group. The dodecyloxy end goups will be reintroduced when modeling the assembly of rosettes since in this case the steric hindrance produced by the adsorbed dodecyloxy end groups can strongly influence the stability of the different assemblies and therefore it cannot be neglected in the model.
The graphite surface has been modeled by a single periodic graphene slab, which has been treated as a rigid body fixed in space. This approximation is based on the assumption that physisorption of molecules does not significantly affect the geometry of the surface.

How the molecular binding energy is calculated
The binding energy, i.e., the energy gained by each OPV3T molecule when it is part of the rosette with respect to that of an isolated molecule adsorbed on the graphite surface, is calculated using eq. 1.

$$E_b = \frac{(E_{ass} - 6 \times E_{OPV})}{6} \quad \text{(eq. 1)}$$

where E_b is the binding energy, E_{ass} the total potential energy for the assembly and E_{OPV} the potential energy for an isolated OPV3T, respectively. For the achiral OPV, the energy of the isolated molecule adsorbed on surface, E_{OPV}, is 20.6kcal/mol and the binding energy is about -4 kcal/mol for both CW and CCW rosettes (see table 1 in the paper).
Considering the chiral OPV, since the chirality of the molecule breaks the symmetry with respect to the way it adsorbs on the surface, two reference systems must be used in order to evaluate those binding energies. For the S1 and S2 rosettes, we consider the energy of the isolated molecule when it adsorbs having the methyl groups pointing away from the surface (E_{OPV}=26.8kcal/mol), whereas for S1F and S2F rosettes, we consider the OPV molecule adsorbed having the methyl groups pointing towards the surface (E_{OPV}=33.8kcal/mol). The binding energy obtained (see table 2 in the paper) is more favorable for S1F and S2F rosettes, meaning that isolated molecules having the chiral groups pointing at the surface are better stabilized when assembled in the rosettes, but, still, their aggregation gives structures that are not energetically favorable with respect the to S1 and S2 aggregates. Regarding the stable S1 and S2 rosettes, a difference in the binding energy is found in favour of the clockwise (S1) rosette.

On the end-to-end distance between OPVs lateral groups in the rosette aggregates.
The CW and CCW molecular arrangements, i.e., S1/A1 and S2/A2 rosettes, not only have different H-bonding patterns linking the molecules to form the cores of the rosettes, but also differ for the orientation of the OPVs lateral groups. Fig. S2 shows both CW and CCW rosettes, and a black line is used to highlight the orientation of the OPVs lateral groups. While in the CW rosettes those groups are pointing to different directions, in the CCW arrangement they are pointing directly to each other.
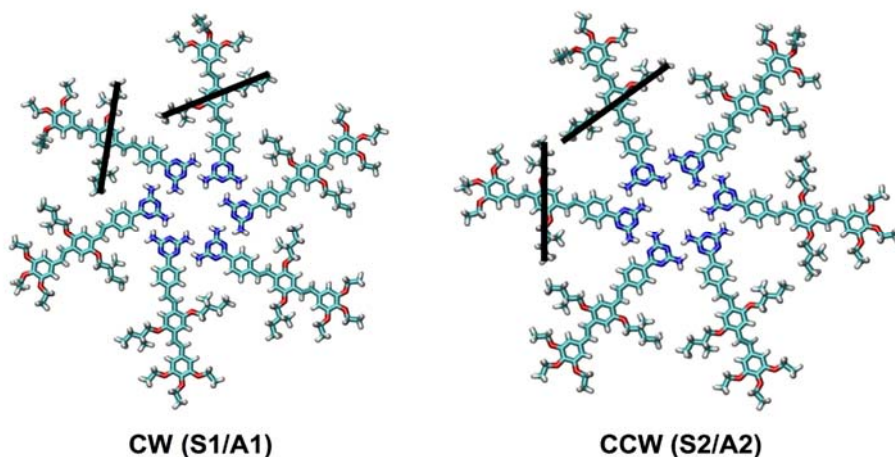

CW (S1/A1)   **CCW (S2/A2)**

Fig. S2 CW (S1/A1) and CCW (S2/A2) rosettes:the black lines highlight the orientation of the OPVs lateral groups. The end-to-end distance is calculated as the shortest distances between the lateral groups of nearest OPV molecules.

A statistical analysis of the end-to-end distance between the OPVs lateral groups in the most stable structures, i.e., the shortest distance between the lateral groups of nearest OPV molecules, has been performed and the results are reported in Table 3.

Table 3 Average end-to-end distances between OPVs lateral groups in the different rosettes

Rosette (orientation)	Distance/nm
S1/A1 (CW)	0.82 ± 0.05 / 0.85 ± 0.06
S2/A2 (CCW)	0.45 ± 0.05 / 0.45 ± 0.01

For the CW rosettes, the average end-to-end distance is 0.82 nm, almost twice the value found for the CCW rosette, i.e., 0.45nm. The small standard deviation obtained for both the molecular arrangements shows how regular is the end-to-end distance in the rosettes.

Building the monolayers
The first step in building the aggregation of rosettes to form a monolayer is to assemble S1 and S2 rosettes according to the organisation extracted from the STM images for the CW layer. After having relaxed these preliminary assemblies, the dodecyloxy end groups have been reintroduced in the OPV molecular structure (fig. 2 top). In contrast to the case of the isolated rosettes, here the steric hindrance produced by the adsorbed dodecyloxy end groups can strongly influence the stability of the different assemblies and therefore it cannot be neglected in the model. We reintroduce only two dodecyloxy end groups per molecule, since it is known experimentally that the third dodecyloxy group is not adsorbed and remains in the liquid phase, probably due to the steric hindrance at the surface caused by the other end groups. The usual MD/quench iterative scheme is then applied to obtain the most stable structures for the monolayers of CW and CCW rosettes (fig. 2 bottom).
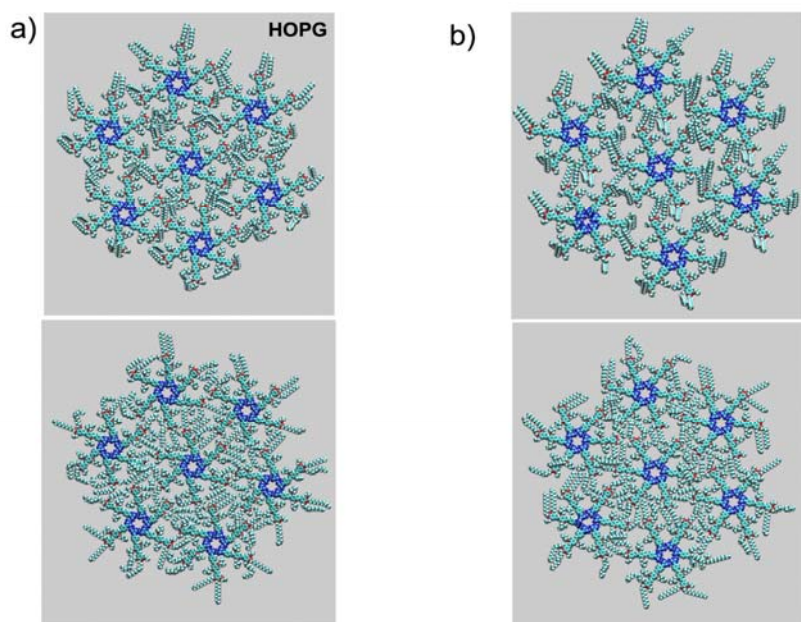

Fig. S2 Model monolayers composed of seven CW S1 (a) and CCW S2 (b) rosettes. The top images represent the starting geometries, while those at the bottom show two low energy structures resulting from the MD/quench iterative scheme. For the sake of clarity the graphite surface is represented as a continuous light gray background.

References
[1] Product description: http://accelrys.com/products/materials-studio/
[2] S.L. Mayo, D. B. Olafson and W.A. Goddard III, *J. Phys. Chem.*, 1990, **94**, 8897
[3] M. Nishio, M. Hirota, Y. Umezawa, *The CH-π interaction: Evidence, Nature and Consequences*, Wiley-VCH:New York, 1998; K. Shibasaki, A. Fujii, N. Mikami, S.J. Tsuzuki, *J. Phys. Chem. A.*, 2006, **110**, 4397.