

*Supporting information*

## **Modelling the Self-Assembly of Organic Molecules in 2D Molecular Layers with Different Structures**

Joost van der Lit, Jolien L. Marsman, Rik S. Koster, Peter H. Jacobse, Stephan A. den Hartog, Daniel Vanmaekelbergh, Robertus J. M. Klein Gebbink, Laura Filion\*, Ingmar Swart\*

Debye Institute for Nanomaterials Science, Utrecht University, PO Box 80000, 3508 TA Utrecht, The Netherlands

\* to whom correspondence should be addressed: [L.C.Filion@uu.nl](mailto:L.C.Filion@uu.nl); [I.Swart@uu.nl](mailto:I.Swart@uu.nl)

## 1. Synthesis of 1, 2, and 3

### 1.1. Bis-(para-benzoicacid)-acetylene (1):<sup>1</sup>

A catalyst solution was prepared by dissolving Pd(OAc)<sub>2</sub> (23 mg, 0.103 mmol) and tPPTs (342 mg, 0.602 mmol) in degassed H<sub>2</sub>O (10 mL) and stirring for 30 min at room temperature (RT), yielding a dark-brown solution. 4-iodobenzoic acid (1.01 g, 4.073 mmol) and 4-ethynylbenzoic acid (0.589 g, 4.031 mmol) were added to 100 mL H<sub>2</sub>O/MeCN (1:1) yielding a yellow-orange suspension. K<sub>2</sub>CO<sub>3</sub> (2.773 g, 20.09 mmol) was added to the mixture, which was stirred at 50 °C under N<sub>2</sub>. After about 5 min, when a dark orange solution was obtained, a CuI solution (45 mg, 0.24 mmol, in MeCN) and catalyst solution (10mL) were added, turning the mixture yellow-green. After stirring for an hour, the mixture was allowed to cool to RT. HCl (37%) was added to precipitate the product which after filtering yielded 79.9 mg (7.5% yield) of tan solid.

<sup>1</sup>H NMR (400 MHz, d<sub>6</sub>-DMSO): δ=7.96(d,4H, J<sub>1</sub>= 8.61Hz), 7.68 (d, 4H, J<sub>1</sub>= 8.22Hz). <sup>13</sup>C NMR(400 MHz, d<sub>6</sub>-DMSO): δ = 167.06, 132.15, 131.38, 130.03, 126.48, 91.49

### 1.2. Bis-(para-cyanobenzene)acetylene (2):<sup>2</sup>

A mixture of 4-bromobenzonitrile (1.0026 g, 5.505 mmol), 4-ethynylbenzonitrile (0.6662 g, 5.503mmol), Pd(PPh<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub> (79.1 mg, 0.113 mmol) and PPh<sub>3</sub> (27.2 mg, 0.104 mmol) in dry Et<sub>3</sub>N (20 mL) was stirred for 10 min under a N<sub>2</sub>. CuI (38.8 mg, 0.204 mmol) was added and the suspension was heated to 90 °C. The brown-green suspension was stirred for 2 h under reflux. After the mixture had cooled to RT the solvent was removed until a brown solid remained. CHCl<sub>3</sub> (80 mL) was added, and the solution was filtered, washed with a K<sub>2</sub>CO<sub>3</sub> solution, water, and brine. After drying over MgSO<sub>4</sub>, the solvent was evaporated to obtain a brown-yellow solid. The product was purified using column chromatography (Silica gel) (DCM/hexane (6:4)). Further purification was done by recrystallization from acetone yielding 319.1 mg (25.4% yield) of yellow needle crystals.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 7.66 (d, 4H,  $J_1=9.00\text{Hz}$ ), 7.61 (d, 4H,  $J_1 = 8.61$  Hz).  $^{13}\text{C}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  = 132.15, 127.05, 118.21, 112.41, 91.52

### 1.3. Bis-(para-pyridyl)acetylene (3):<sup>3</sup>

$\text{Br}_2$  (1 mL, 17.26 mmol) was slowly added to a stirred solution of 1,2-bis(4-pyridyl)ethylene (1.01 g, 5.537 mmol) in  $\text{HBr}$  (17 mL, 48%) at 0 °C. The mixture was stirred for 2 h at 120 °C and allowed to cool to RT, before being cooled in ice for 30 min. The resulting suspension was filtered and the residue was washed with cold water and stirred in  $\text{NaOH}$  (35 mL, 2 M) for 1 h before being filtered again. The residue was dried in vacuo, added to a  $t\text{-BuO}^- \text{Na}^+$  solution (60 mL) and left to stir at 80 °C for 4 h. After the mixture was allowed to cool to RT, ethanol (20 mL) and  $\text{H}_2\text{O}$  (20 mL) were added and the resulting orange solution was extracted with  $\text{CDCl}_3$  until the extracts were colorless. After the solvent was removed, a dark orange oil was obtained to which  $\text{CDCl}_3$  and water were added. The resulting solution was again extracted with  $\text{CDCl}_3$  until the extracts were colorless and the solvent was removed to obtain a yellow solid. Purification by silica gel column with a diethyl ether:THF (17:3) eluent yielded 269.4 mg of white solid product (26.9% yield).

$^1\text{H}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  = 8.61 (d, 4H,  $J_{\text{HH}} = 5.09$  Hz), 7.37 (d, 4H,  $J_{\text{HH}} = 5.09$  Hz),  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ):  $\delta$  = 149.83, 130.22, 125.57, 90.62

## 2. MC simulations

### 2.1. Canonical NVT ensemble

We used two-dimensional MC simulations in the canonical NVT ensemble, e.g. with a fixed number of molecules ( $N$ ), fixed volume ( $V$ ), and fixed temperature ( $T$ ), to study the self-assembly of the molecules. Sampling was performed using the Metropolis algorithm.<sup>4</sup> In the initial configuration the molecules are positioned such that they do not overlap and have a random orientation. Each MC step consisted of either a trial displacement in the lateral direction or a trial rotation in the plane. The respective step sizes were chosen such that approximately 30% of the attempted MC steps were accepted. Additionally, the probability of performing a trial displacement was the same as the probability of performing a trial rotation, both equal to 0.5. The molecular density was chosen low enough that the crystal formed in coexistence with a gas. To avoid boundary effects, periodic boundary conditions are applied in the simulation. One MC cycle consisted of  $N$  steps, where  $N$  is the number of molecules in the simulation.

### 2.2. Isotension-isothermal NPT simulations

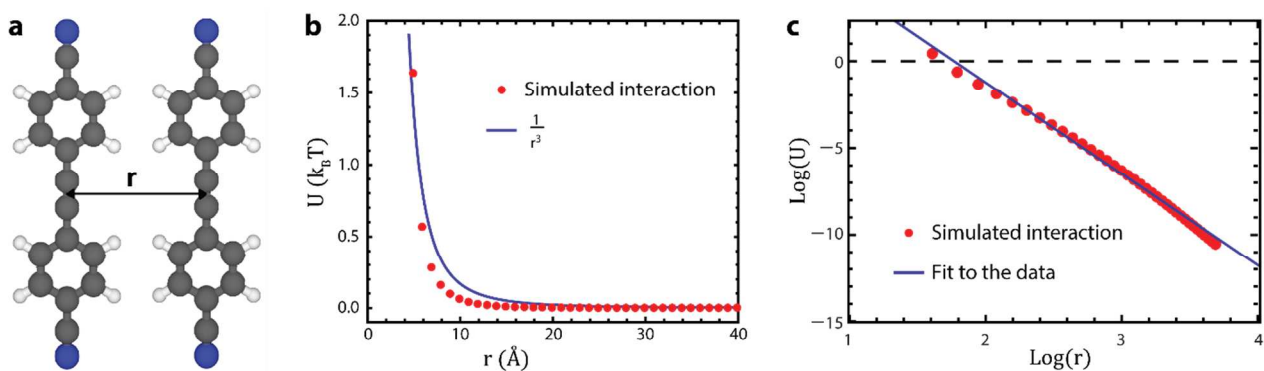
In the NVT simulation the crystals that form are not very large. Furthermore, since the box shape is fixed, the whole box will never be covered with the crystal. We used a different MC method to obtain a larger crystal. In this case, we performed MC simulations in the isobaric – isothermal ensemble, e.g. where the pressure ( $P$ ) was kept constant instead of the volume ( $V$ ), and the box shape was allowed to fluctuate in order to remove stresses caused by the simulation box geometry.<sup>5</sup> In addition to normal MC translation and rotation moves, we also allowed the box volume and shape to fluctuate using trial volume changes. For the initial configuration, the entire box was filled with the crystal found in the NVT simulations. To allow the crystal to fully relax, we fixed the

pressure at zero. Hence, these simulations are performed in the isotension-isothermal ensemble at zero pressure.

### 2.3. Distance dependence of the Coulomb interactions.

Electrostatic forces can be long range in nature. This implies that the energy of a molecule can, in principle, depend on the structure of the material at arbitrarily large distance. In practice, this means that it is not fully safe to cutoff the potential energy at a finite distance and specialized numeric techniques are necessary.<sup>7</sup> In two dimensions, an interaction which is proportional to  $r^{-\alpha}$ , with  $\alpha$  an integer, is considered to be a short range interaction when  $\alpha > 2$ , while for  $\alpha \leq 2$  it is considered a long range interaction.<sup>6,7</sup> The molecules in our model are charge neutral so there is no net monopole-monopole interaction between the molecules. Hence, the interaction between two molecules has overall only dipole ( $r^{-3}$ ), quadrupole ( $r^{-5}$ ) and higher order interactions. These interactions are in principle all short range interactions. To verify this, we calculated the energy of two molecules **2** with the same orientation positioned side by side, see Figure S1a. The as calculated Coulomb interactions for different distances are given in Figure S1b. For comparison, a fit to an  $r^{-3}$  function is also plotted. The interaction between the molecules decays faster than  $r^{-3}$ . Figure S1c shows a double logarithmic plot of the interaction together with a fit. The slope is found to be -5.27. Hence, the interaction between the two molecules decays faster than  $r^{-5}$ .

In our model, we therefore neglected electrostatic forces if the center-to-center distance between a pair of molecules was larger than 30Å. As can be seen in Figure S1b, the potential energy between these molecules approaches zero at this distance.



**Figure S1.** Distance dependence of the Coulomb interactions. **(a)** Definition of the lateral distance between two molecules of **2**. **(b)** Strength of the coulomb interaction between the molecules as a function of distance (red). For comparison a plot of  $r^{-3}$  is also given (blue). **(c)** Double logarithmic plot of the energy-distance relationship (red), together with a fit (blue). The slope of the fit (-5.27) indicates that the interaction decays faster than  $r^{-5}$ .

### 3. Robustness of the model

We examined how robust the results of the model are with respect to changes in the distance between the charges ( $\delta$ ) and the distance between the positive charges and the lone pairs ( $\Delta_L$ ). In addition, we examined what the influence of the presence of a conductive surface is. Finally, we studied the self-assembly of **2** to establish if the model can also be applied to self-assembly on other surfaces.

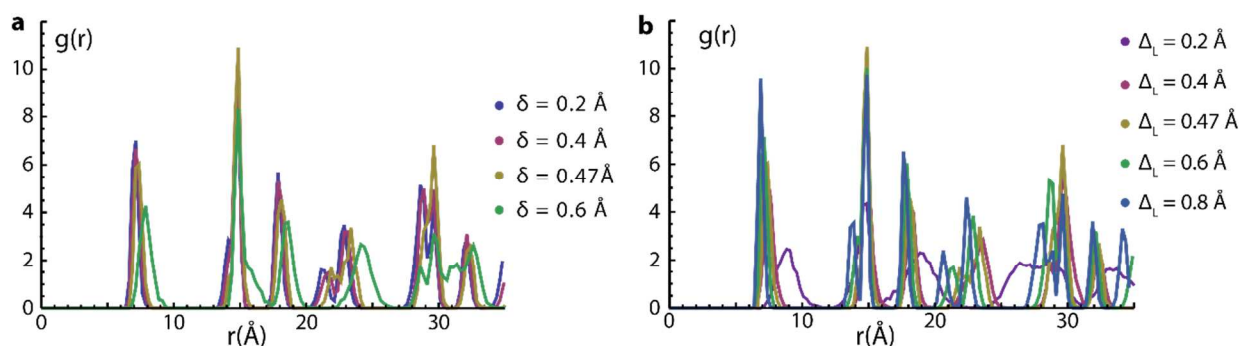
#### 3.1. Influence of $\delta$ and of $\Delta_L$

To study the robustness of the model, we studied the self-assembly of bis-(para-cyanobenzene)-acetylene (molecule **2** in the main text) with various values of  $\delta$  and  $\Delta_L$ . The results are given in Figure S2. We varied  $\delta$  from 0.2 to 0.8 Å. For values of  $\delta$  smaller than approximately 0.6 Å, the molecules form the close-packed layer described in the main text. In this regime, an increase in  $\delta$  results in a small

increase in the distance between the molecules. This increase can be rationalized from a decrease in the inter-molecular interactions (due to a larger distance between the negative charges of one atom and the positive charge of another atom). Above  $\delta = 0.6 \text{ \AA}$ , the molecules self-assemble into a different crystal structure.

Increasing  $\Delta_L$  has the same effect as increasing  $\delta$ : the distance between the molecules increases but the symmetry of the crystal is unaffected. For very small values of  $\Delta_L$ , there are significant differences in the crystal structure, as evidenced by the different radial distribution function (purple curve in Figure S2b). This can be rationalized from the smaller effect of the negative point charge used to model the lone pair at smaller  $\Delta_L$ .

Considering that  $\delta$  corresponds to distances in a quadrupole, we conclude that the model is robust with respect to small variations in  $\delta$  and  $\Delta_L$ .

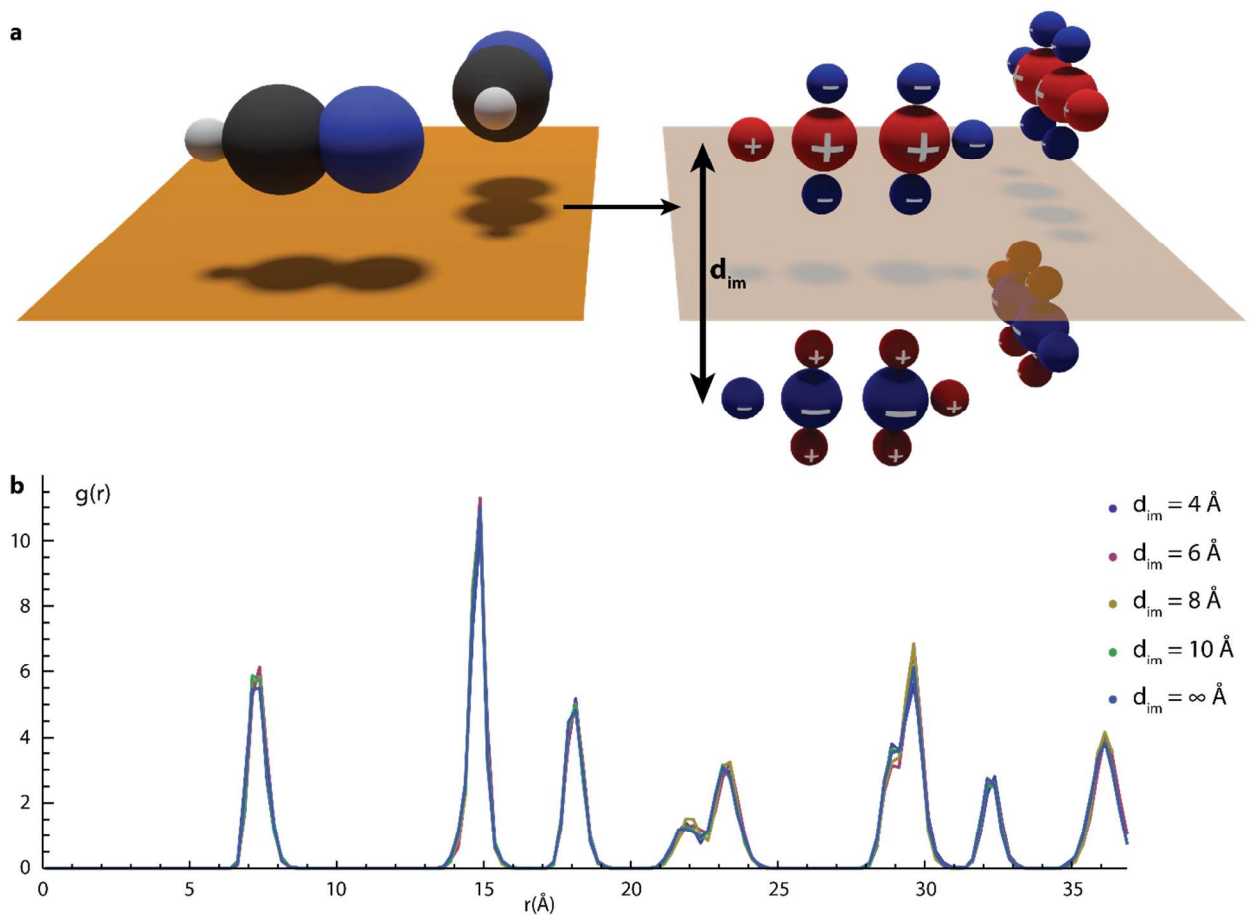


**Figure S2. The radial distribution of BPCA with varying  $\delta$  and lone pair distance.** **a**, The radial distribution of **2** for several values of delta. **b**, The radial distribution of **2** with several distances for the lone pair.  $\delta = 0.47 \text{ \AA}$ .

### 3.2. Modelling the surface via image charges

It is well known that the electrostatic interaction between a point charge and a conductive surface can be modelled using the method of image charges.<sup>8</sup> To examine the influence of image charges on the self-assembly, we expanded our model as indicated in Figure S3a. The image charges are located at a distance  $d_{im}$  below the molecule. When two atoms interact in the expanded model, not only the interaction between the charges are taken into account but also the interaction between the charges of one atom and the image charges of the other atom. Interactions between image charges are not included. As a starting point for the distance between the image charges, we set it equal to twice the typical adsorption height of 3.3 – 3.4 Å on Au(111). Hence,  $d_{im}$  should be approximately 6.6 Å. We varied  $d_{im}$  between 4 Å and  $\infty$ . The latter corresponds to neglecting image charges (as in the original model). Figure S3b shows the radial distribution function for different  $d_{im}$  of a close packed layer of **2**. The radial distribution functions overlap for all values of  $d_{im}$ , demonstrating that image charges have a negligible influence on the self-assembly and can therefore be neglected.

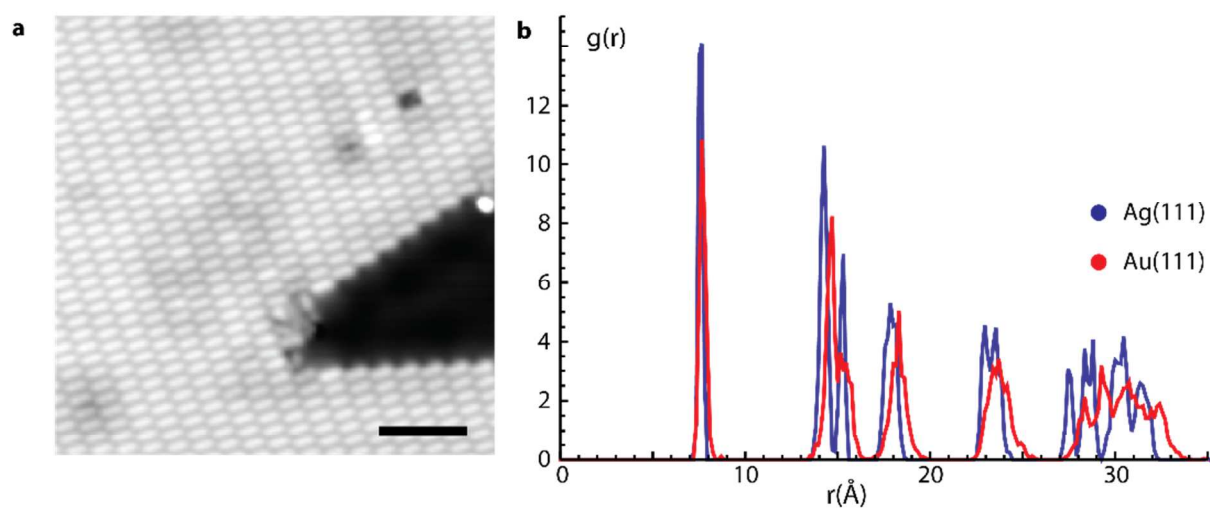




**Figure S3. Effect of including image charges in the model.** **a**, a model of two HCN molecules on a surface (left) and their analog using our model (right). Images charges are included below the real molecule as opposite point charges. **b**, the radial distribution function of the close packed structure of **2** calculated at different distances  $d_{im}$ . At infinite distance the model is similar to the one presented in the main text.

### 3.3. Adsorption of **2** on Ag(111)

To investigate if the model can also be used for molecules on other surfaces, we studied the self-assembly of **2** on Ag(111). In Figure S4a a STM image of a self-assembled structure of **2** on Ag(111) is shown. As in the case of Au(111), the molecules form a close-packed structure. A comparison of the radial distribution function,  $g(r)$ , of the self-assembled layer on Ag(111) and Au(111) shows only minor differences (Figure S4b). These might be due to changes in the charge distribution in the molecule caused by the stronger molecule-substrate interaction on Ag(111).



**Figure S4. Self-assembly of **2** on a Ag(111) surface.** **a**, STM image of self-assembled layer ( $V = 0.1$  V,  $I = 20$  pA). Scale bar: 5 nm. **b**,  $g(r)$  on **2** on Ag(111) and Au(111).

## Supplementary references

- (1) Dibowski, H.; Schmidtchen, F. P. Synthesis and Investigation of Phosphine Ligands Containing Cationic Guanidino Functions in Aqueous Heck Reactions. *Tetrahedron* **1995**, *51*, 2325–2330.
- (2) Berger, O.; Kaniti, A.; van Ba, C. T.; Vial, H.; Ward, S. A.; Biagini, G. A.; Bray, P. G.; O'Neill, P. M. Synthesis and Antimalarial Activities of a Diverse Set of Triazole-Containing Furamidine Analogues. *ChemMedChem* **2011**, *6*, 2094–2108.
- (3) Coe, B. J.; Harries, J. L.; Harris, J. A.; Brunschwig, B. S.; Coles, S. J.; Light, M. E.; Hursthouse, M. B. Syntheses, Spectroscopic and Molecular Quadratic Nonlinear Optical Properties of Dipolar ruthenium(II) Complexes of the Ligand 1,2-Phenylenebis(dimethylarsine). *Dalton Trans.* **2004**, 2935–2942.
- (4) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087.
- (5) Daan Frenkel, B. S. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press, 2001.
- (6) Dauxois, T.; Ruffo, S.; Arimondo, E.; Wilkens, M. *Dynamics and Thermodynamics of Systems with Long-Range Interactions*; Dauxois, T.; Ruffo, S.; Arimondo, E.; Wilkens, M., Eds.; Lecture Notes in Physics; Springer Berlin Heidelberg: Berlin, Heidelberg, 2002; Vol. 602.
- (7) Campa, A.; Dauxois, T.; Fanelli, D.; Ruffo, S. *Physics of Long-Range Interacting Systems*; Oxford University Press, 2014.
- (8) Perdew, J. P.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. Atoms, Molecules, Solids, and Surfaces: Applications of the Generalized Gradient Approximation for Exchange and Correlation. *Phys. Rev. B* **1992**, *46*, 6671–6687.