Electronic Supplementary Information

Large-extended 2D supramolecular network of dipoles with parallel arrangement on a Si(111)-B surface

Irma Custovic, [a] Damien Teyssieux, [a] Judicael Jeannoutot, [a] Frank Palmino, [a] Hamed Abbasian,[b] Alain Rochefort,[b] and Frédéric Chérioux *[a]

[a]Univ. Bourgogne Franche-Comté, FEMTO-ST, UFC, CNRS, 15B Avenue des Montboucons, F-25030 Besançon, France
[b]Polytechnique Montréal, Département de génie physique, Montréal (Qué), Canada H3C 3A7

1. Experimental

1.1. Synthesis of LDipCC

All reagents were purchased from Aldrich, except $Pd(PPh_3)_4$ which was purchased from Strem chemical, and used as received. The silica gel used for column chromatography was purchased from Merck. The deuterated NMR solvents were purchased from Euriso-top. The NMR spectra were recorded using a Bruker AC-300 MHz spectrometer.

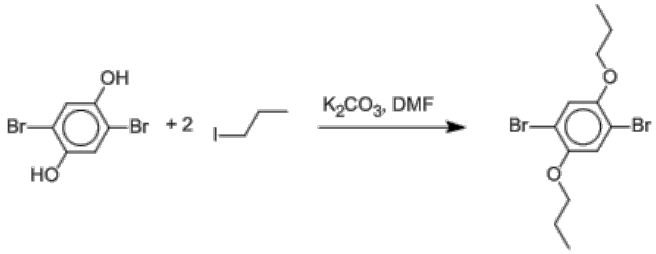

Fig. S1 First step: O-alkylation

This procedure is adapted from S. Mandal *et al.*[1] 1,4-dibromo-2,5-bis(dihydroxy)benzene (2.0 g, 7.5 mmol), 1-iodopropane (6.38 g, 37.5 mmol) and potassium carbonate (3.1 g, 22.4 mmol) were dissolved in 40 mL of DMF. The resulting mixture was heated at 80°C for 10 hours. After cooling to room temperature, 100 mL of Chloroform were added. The organic solution was washed three times with

100 mL of water. The solvent was removed under reduced pressure. The resulting white solid was purified by column chromatography (silica gel, acetone/petroleum ether 1:1). The solid was recrystallized in acetonitrile to give a pure white solid (2 g, 77%). ^{1}H NMR (300 MHz, Chloroform-d) δ = 7.09 (s, 2H), 3.92 (t, J=7.1 Hz, 4H), 1.82 (h, J=7.1 HZ, 4H), 1.06 (t, J=7.1 Hz, 6H).

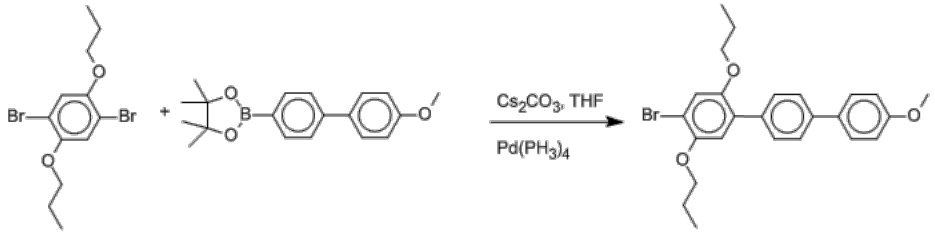

Fig. S2 Second step: Insertion of donor group

This procedure is adapted from N. Miyaura *et al.*[2] 1,4-dibromo-2,5-bis(propyloxy)benzene (448mg, 1.3 mmol) and 4'-methoxybiphenyl pinacol ester (400 mg, 1.3 mmol) were dissolved in 60 mL of THF. Then, an aqueous solution of Cs_2CO_3 (0.5 M, 6 mL) and a $Pd(PPh_3)_4$ catalyst (42 mg, 0.03 mmol) were added. The resulting mixture was heated at 80°C overnight. Then, the solvent was removed under reduced pressure. The crude solid was dissolved in 50 mL of Chloroform and filtered over Celite. Then, the solvent was removed under reduced pressure. The resulting white solid was purified by column chromatography (silica gel, hexane/dichloromethane 1:1) to give a pure white solid (130mg, 22%). ^{1}H NMR (300 MHz, Chloroform-d) δ = 7.6 (m, 6H), 7.18 (s, 1H), 7.00 (d, J=8.0Hz, 2H), 6.85 (s, 1H), 3.99 (t, J=7.1 Hz, 4H), 3.88 (s, 3H), 3.87 (t, J=7.1 Hz, 4H), 1.85 (h, J=7.1 HZ, 2H), 1.71 (h, J=7.1 HZ, 2H), 1.08 (t, J=7.1 Hz, 3H), 0.97 (t, J=7.1 Hz, 3H).

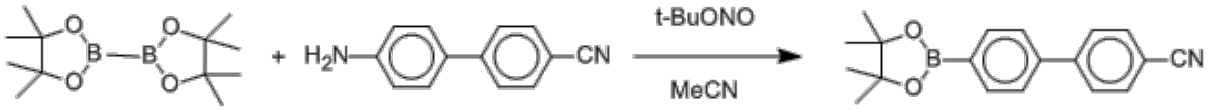

Fig. S3 Synthesis of 4'-cyanobiphenyl pinacol ester

This procedure is adapted from Qui *et al.*[3] Bis(pinacolato)diboron (1.44g, 5.7 mmol) and 4-amino-4'-cyano-biphenyl (1.0, 5.1 mmol), and t-BuONO (789 mg, 7.7 mmol) were dissolved in 25 mL of MeCN. The resulting mixture was heated 2h at 80°C. Then, the solvent was removed under reduced pressure. The resulting mixture was purified by column chromatography (silica gel, hexane/dichloromethane 1:1) to give a pure white solid (500mg, 50%). ^{1}H NMR (300 MHz, Chloroform-d) δ = 7.66 (d, J=8.0 Hz, 2H), 7.63 (d, J=8.0 Hz, 2H), 7.51 (d, J=8.0 Hz, 2H), 7.04 (d, J=8.0 Hz, 2H), 1.30 (s, 12H).

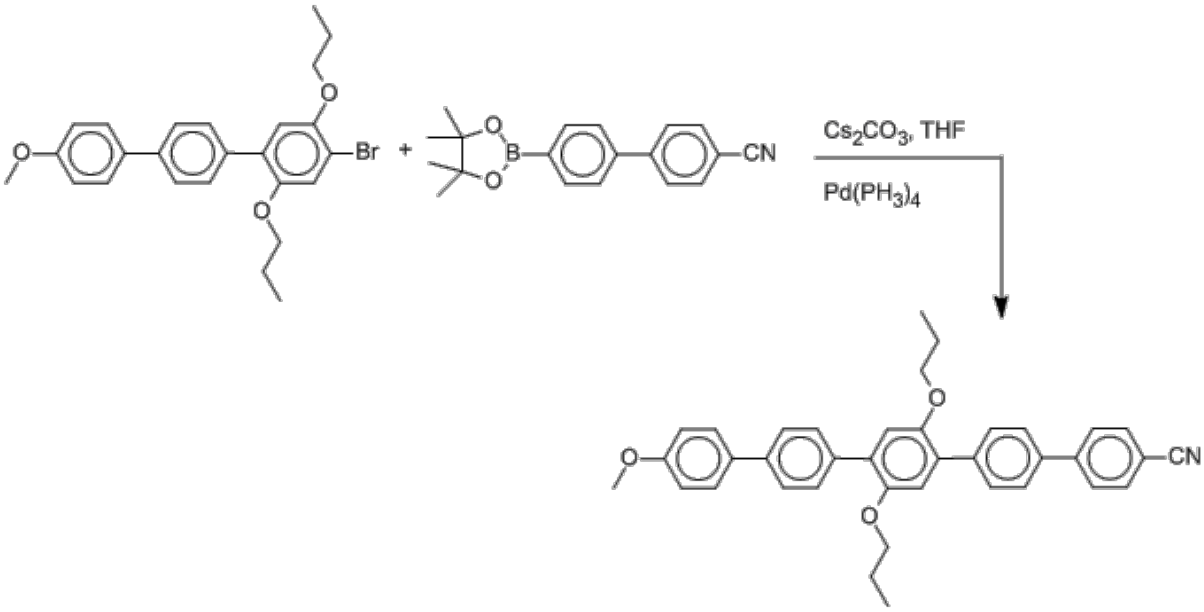

Fig. S4 Insertion of electron withdrawing group leading to LdipCC molecules

This procedure is adapted from N. Miyaura *et al.*[2] 1-bromo-4'-(4"-methoxybiphenyl)-2,5-bis(propyloxy)benzene (120mg, 0.26 mmol) and 4'-cyanobiphenyl pinacol ester (121mg, 0.4 mmol) were dissolved in 60 mL of THF. Then, an aqueous solution of Cs_2CO_3 (0.5 M, 6 mL) and a $Pd(PPh_3)_4$ catalyst (42 mg, 0.03 mmol) were added. The resulting mixture was heated at 80°C overnight. Then, the solvent was removed under reduced pressure. The crude solid was dissolved in 40 mL of Chloroform and filtered over Celite. Then, the solvent was removed under reduced pressure. The resulting white solid was purified by column chromatography (silica gel, hexane/dichloromethane 1:1) to give a white solid. Eventually, the crude solid is recrystallized in 20mL of toluene-ethanol 1/1 mixture to give a pure with solid (90mg, 60%). ^{1}H NMR (300 MHz, Chloroform-d) δ = 7.68 (m, 14H), 7.03 (m, 4H), 3.93 (t, J=7.1 Hz, 4H), 3.88 (s, 3H), 1.74 (h, J=7.1 HZ, 4H), 0.98 (t, J=7.1 Hz, 6H). ^{13}C NMR (75 MHz, Chloroform-d) δ = 159.1, 150.4, 150.3, 145.4, 139.4, 137.5, 133.4, 132.6, 130.8, 130.2, 129.9, 129.6, 128.1, 127.8, 127.6, 126.7, 126.2, 116.1, 116.0, 114.2, 71.2, 55.4, 22.7, 10.8.

1.2. STM experiments

After carefully outgassing (until pressure reaches $5{\cdot}10^{-10}$ mbar) the highly B-doped Si(111) wafer (0.001 Ohm.cm), the sample was cleaned by a series of rapid in situ heating to 1473 K under a pressure lower than $5{\cdot}10^{-10}$ mbar and subsequent annealing (1 h at 1123 K). The Si(111)-B √3·√3 reconstruction surface was obtained with a maximum boron atom concentration. In these conditions, we obtained a passivated Si surface because the boron atoms induce a charge transfer that depopulates the Si dangling bonds.
STM experiments were performed in a UHV chamber with a base pressure lower than $2{\cdot}10^{-10}$ mbar using a variable temperature Omicron scanning tunnelling microscope (VT-STM). STM images were acquired in a constant-current mode at room temperature or at 110 K.
LDipCC molecules were deposited from a quartz crucible at 478 K. The Si(111)-B surface was kept at room temperature during the sublimation. Each image was treated with the WSXM software.[4] The structural models are created with Blender 3D.[5] The size of Si adatoms are arbitrary magnified for clarity.

2. Additional STM image

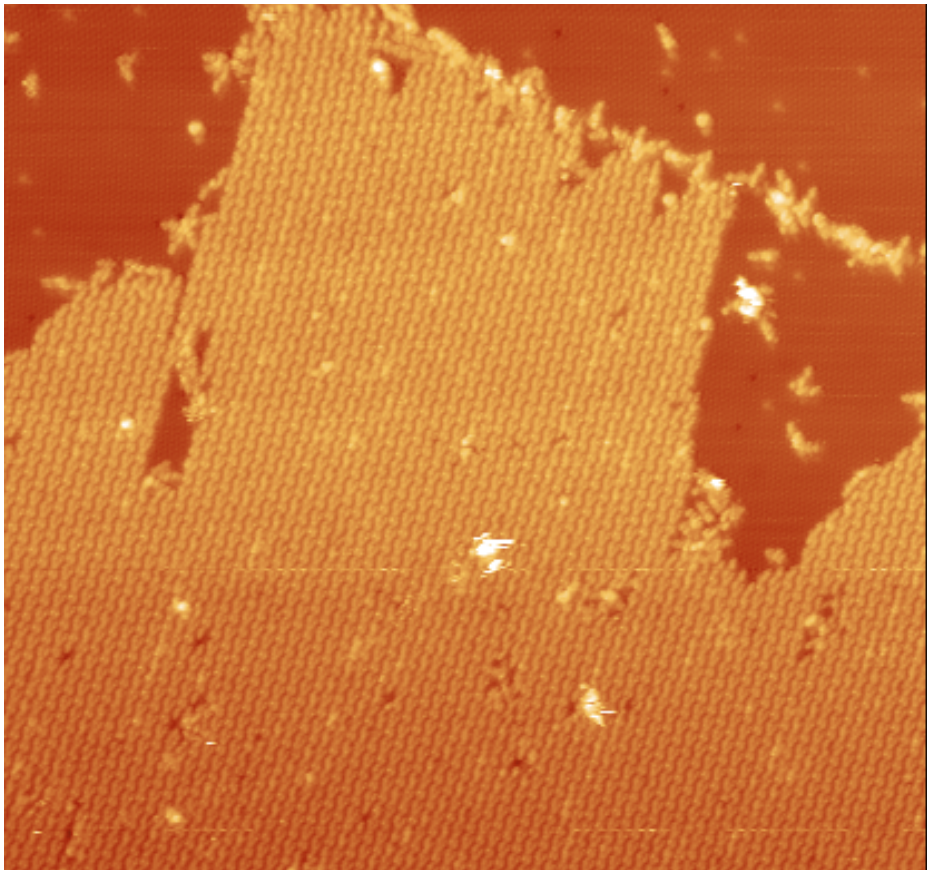

Fig. S5 Large-scale STM image of a 0.9 monolayer of LDipCC on Si(111)-B surface with submolecular resolution on each molecule (98x106 nm^2, V_S = +1.9 V, I_t = 10 pA, T = 110 K).

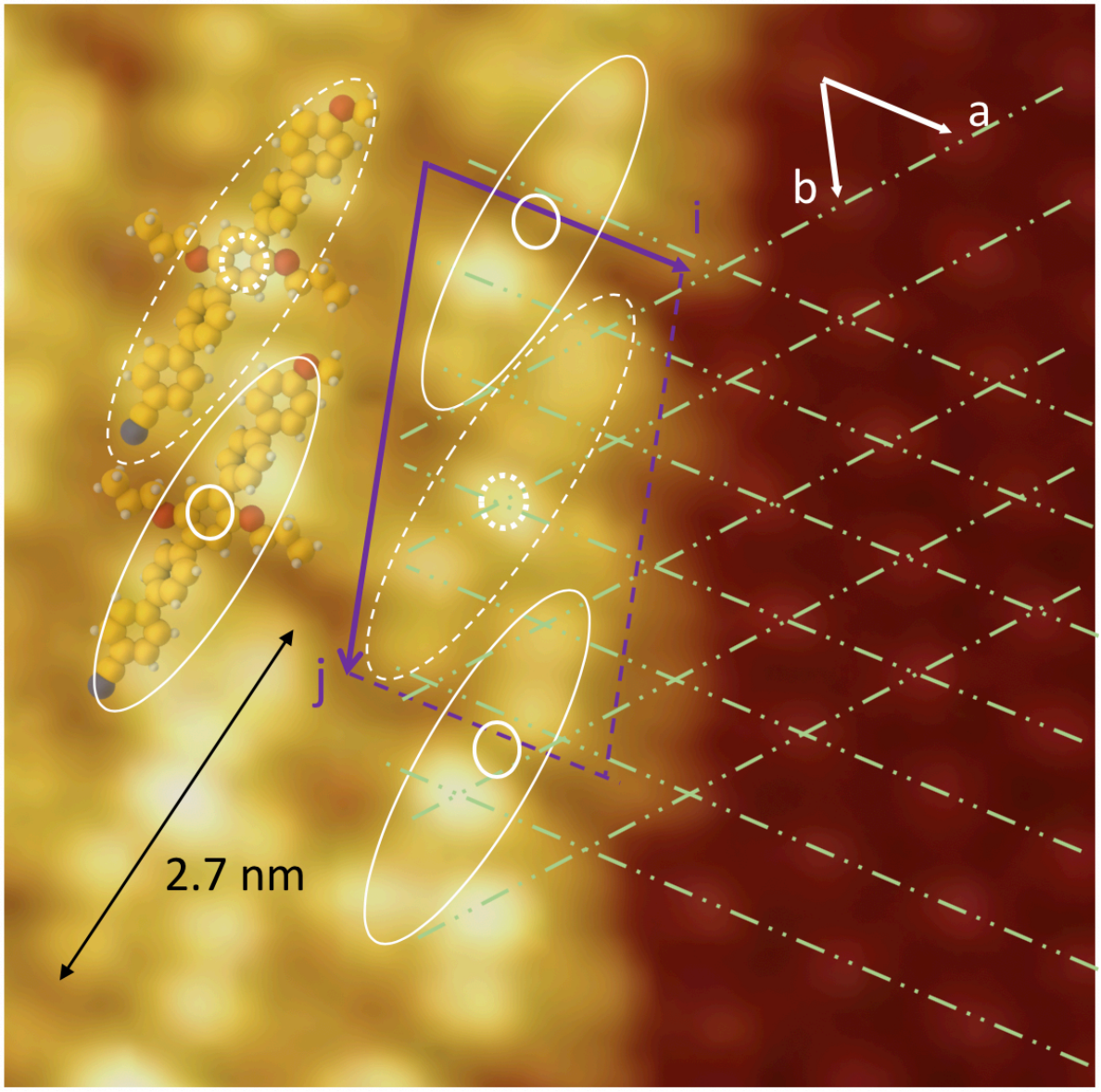

Fig. S6 STM image of a step island of LDipCC on Si(111)-B surface with submolecular resolution on each molecule (4.5x4.5 nm^2, V_S = +2.3 V, I_t = 10 pA, T = 110 K) with superimposed molecules. The central phenyl ring of one molecule is adsorbed over a silicon adatom (surrounded by the dashed white circle), while the second one is adsorbed between two silicon adatoms (surrounded by the white circle), as shown by the dashed lines corresponding to rows of silicon adatoms. a and b are the vectors describing the lattice of the Si(111)-B surface. i and j are the vectors describing the unit cell containing two LDipCC molecules and used for calculations.

3. Computational procedures

The Density Functional Theory (DFT) calculations were carried out at T = 0K with NWChem[6] and Siesta packages, and the results were visualized with the Chimera software[7]. With NWChem, we used the generalized gradient approximation using the PBE0 functional[8] and include energy correction for the van der Waals interactions[9]. We used 6-311G basis sets to describe all H, C, O and N atoms. First, we have determined the more stable geometry of gas phase LDipCC molecule by performing series of full optimization cycles where the starting geometry was modified by rotating different segments of the molecule. The obtained optimized geometry is shown in **Figure S7**, with a calculated dipole moment of 7.0 Debye orientated along the molecular axis toward the methoxy fragment. The calculated dipole moments are used to perform a qualitative analysis of the LDipCC network, a more accurate approach to describe charge distribution need to be used for a quantitative description.

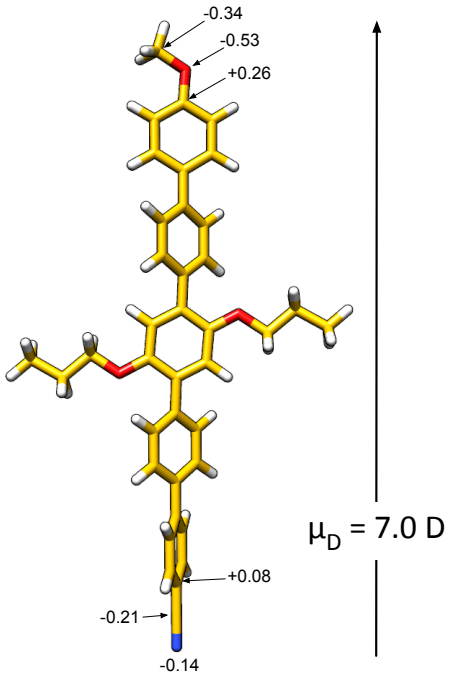

Fig. S7. Most stable geometry of LDipCC molecule obtained from DFT calculations. Mulliken population (a negative value indicates electron accumulation) of different atoms among the terminal groups in LDipCC indicates that both -CN and -OCH_3 groups are electron rich regions. A geometry where one propyloxyl side-chain is rotated out-of-plane was found 10 meV more stable.

This optimized geometry was then used to build the self-assembly of LDipCC on the Si(111)-B surface models shown in Figure 2. The position of the different LDipCC molecules with respect to surface Si adatoms was guided by the analysis of experimental STM images near the assembly boundaries where the surface adatoms can be clearly identified. From those static models, we considered small island of four LDipCC molecules to determine the nature of HOMO/LUMO for parallel and antiparallel molecular arrangement. Our aim was to evaluate the degree of wavefunction overlap among the different assembly without the presence of a substrate. **Figure S8** shows series of four nearly degenerate orbitals below (V < 0) and above (V > 0) Fermi level (E_F = 0 eV). The wavefunction of these small islands reproduce the original nature of individual molecule, and the overlap between the units is weak supporting a relatively small orbital mixing.

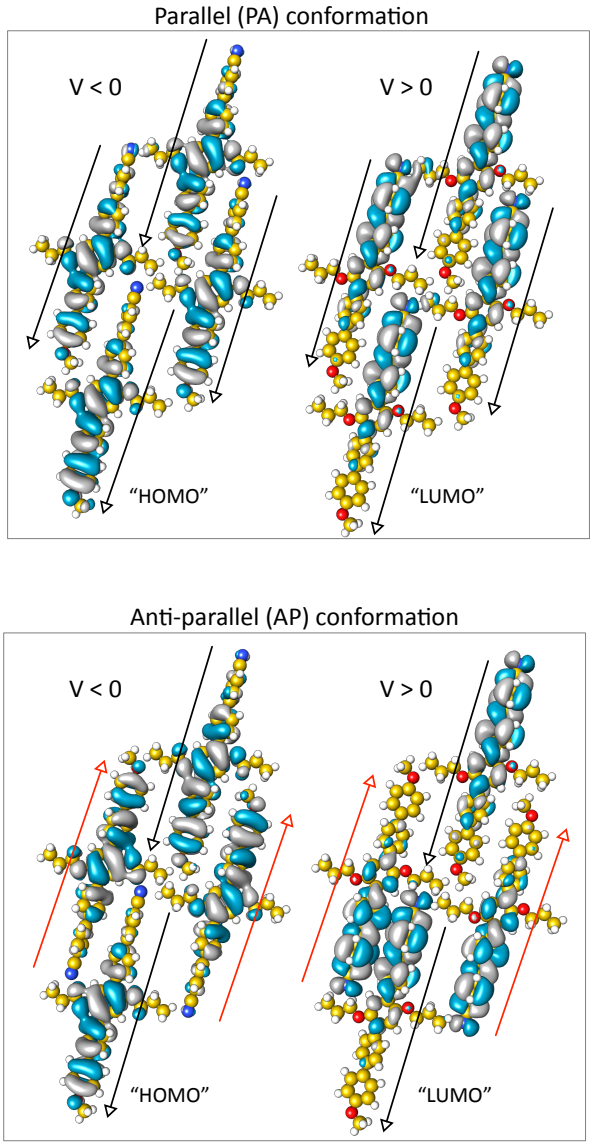

Fig. S8 Wavefunction contours of HOMO (V < 0) and LUMO (V > 0) for LDipCC molecule organized into a parallel (upper panel) and anti-parallel (lower panel) arrangement. The calculated dipole moment for the PA and AP arrangements are 28.9 D and 1.7 D, respectively. The black and red arrows indicate the direction of the dipole moment of individual LDipCC molecule. (isosurface = 0.01 $Å^3$)

For calculations with Siesta, we used periodic boundary conditions with vdW-DF2 (or LMKLL) functional that includes van der Waals corrections. The supercells models shown at Fig. S9 were built from Fig. 2. The LDipCC molecules in the supercell were fully optimized while the reconstructed Si(111)-B(√3×√3) R30° surface slab was fixed to the bulk geometry.[10,11] The computations were performed with norm-conserving Trouillier−Martins pseudopotentials and double-ζ polarized atomic basis sets. The mesh cut-off used to form the real space grid in the DFT calculations was 300 Ry, and the *k*-point grid was limited to the Γ point. We considered a vacuum region of 35 Å to minimize the interactions between periodic images in the direction normal to the slab. The structural relaxation and geometry optimization were carried out using conjugate-gradient method until the forces and the variation of total energy were less than 0.04 eV/Å and 0.001 eV, respectively. To discriminate between molecule-molecule and molecule-surface interactions, we calculated the adsorption energy of PA and AP arrangement such as:

$$\boldsymbol{E}_{\mathrm{ads}} = \boldsymbol{E}_{\mathrm{LDipCC/SiB}} - \boldsymbol{E}_{\mathrm{SiB}} - \boldsymbol{E}_{\mathrm{LDipCC}}$$

where, E_{LDipCC} is the total DFT energy of the free standing LDipCC network with the same geometry as in the LDipCC/Si(111)-B complex, E_{SiB} is the total energy of the substrate and $E_{\mathrm{LDipCC/SiB}}$ is for the optimized complex. The comparison of E_{LDipCC} for PA and AP arrangement is used to estimate the contribution from molecule-molecule interactions to their relative stability. The calculated E_{ads} values are -5.32 and -4.90 eV/molecule for the PA and AP arrangement, respectively. As given the large size of the LDipCC molecule, the magnitude of these values appears reasonable.

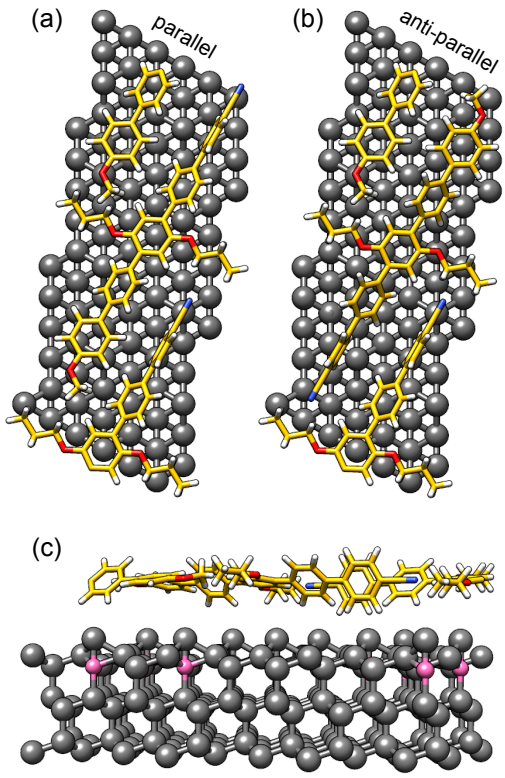

Fig S9. Unit cell of the (a) parallel (PA) and (b) anti-parallel (AP) arrangement on a Si(111)-B surface used in DFT calculations. A side view of the PA arrangement is given in (c) to show the Si-B substrate, where the B atoms (in pink) are sitting under every Si surface adatoms

The STM simulations were performed with the DyFlex software[12-13] developed in our group. The present calculations used the Tersoff-Hamann[14] theory in conjunction with the ASED-MO theoretical framework[15]. The systems are represented by large-scale molecular models that contain 15 LDipCC molecules (1155 atoms) and a Si(111)-B surface (1700 atoms). Such large and finite models allow us to control individual atoms and open the possibility to explore the influence of single/multiple/random structural or chemical defects within a nearly periodic arrangement. The models were built from the relaxed geometries obtained from the DFT calculations. The STM images were visualized with the help of Python 3.7 and Mayavi for 3D visualization[16].

[1] S. Mandal, S. Mandal, S. K. Ghosh, P. Sar, A. Ghosh, R. Saha, B. Saha*, RSC Adv.* 2016, **6**, 69005.
[2] N. Miyaura, A. Suzuki, *Chem. Rev.* 1995, **95**, 2457.
[3] D. Qui, L. Jin, Z. Zheng, H. Meng, F. Mo, X. Wang, Y. Zhang, J. Wang, *J. Org. Chem.* 2013, **78**, 1923.
[4] I. Horcas, R. Fernandez, J. M. Gomez-Rodriguez, J. Colchero, J. Gomez-Herrero, A. M. Baro, *Rev. Sci. Instrum.* 2007, **78**, 013705.
[5] https://blender.org
[6] M. Valiev, E.J. Bylaska, N. Govind, K. Kowalski, T.P. Straatsma, H.J.J. van Dam, D. Wang, J. Nieplocha, E. Apra, T.L. Windus, W.A. de Jong, *Comput. Phys. Commun*. 2010, **181**, 1477.
[7] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, *J. Comput. Chem.* 2004, **25**, 1605.
[8] C.Adamo, V.Barone, *J. Chem. Phys.* 1999, **110**, 6158.
[9] S. Grimme *J. Comp. Chem*. 2006, 271787.
[10] F. Cleri, *Phys. Rev. B* 2009, **80**, 235406.
[11] G. Copie, Y. Makoudi, C. Krzeminski, F. Chérioux, F. Palmino, S. Lamare, B. Grandidier, F. Cleri, *J. Phys. Chem. C* 2014, **118**, 12817.
[12] M.-A. Dubois, X. Bouju, A. Rochefort, *J. Appl. Phys*. 2018, **124**, 044301.
[13] N. Boulanger-Lewandowski, A. Rochefort, *Phys. Rev. B* 2011, **83**, 115430.
[14] J. Tersoff, D.R. Hamann, *Phys. Rev. B* 1985, **31**, 805.
[15] A. B. Anderson, R. W. Grimes, S. Y. Hong, *J. Phys. Chem*. 1987, **91**, 4245.
[16] P. Ramachandran, G. Varoquaux, *IEEE Computing in Science & Engineering* 2011, **13**, 40.