# Supporting information

# Descriptor-based Rational Design of Two-Dimensional Self-Assembled Nano-architectures Stabilized by Hydrogen Bonds

Pratap Zalake[‡,†], Sukanya Ghosh[§], Shobhana Narasimhan*[,§] and K. George Thomas*[,‡]

[‡]School of Chemistry, Indian Institute of Science Education and Research Thiruvananthapuram
(IISER-TVM), CET Campus, Thiruvananthapuram, India

[§]Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur,
Bangalore 560 064, India

*Corresponding authors: shobhana@jncasr.ac.in*[,§], kgt@iisertvm.ac.in*[,‡]

†Present address: †EPFL STI IMX SUNMIL MXH 130, Station 12, CH-1015 Lausanne, Switzerland

# Contents

| | | |
|---|---|---|
| 1. | Experimental Methods: details of synthesis of phenyleneethynylene derivatives and STM imaging<br>Figure S1: Schematic diagram of synthesis of PE4B and PE3A. | pg *S*3 |
| 2. | Computational Methods: details of DFT calculations.<br>Table S1: Results from DFT for geometries and energetics of the host assemblies<br>Figure S2: Atomistic structures of the host assemblies. | pg *S*6 |
| 3. | Figure S3: Experimental STM images of host assemblies. | pg *S*9 |
| 4. | Figure S4: Experimental STM images and atomistic structures from DFT of PE4A with various guests. | pg *S*10 |
| 5. | Figure S5: Charge redistribution plots from DFT for PE4A with various guest molecules. | pg *S*11 |
| 6. | Figure S6: Experimental STM images and atomistic structures from DFT of PE4B with various guests. | pg *S*12 |
| 7. | Figure S7: Charge redistribution plots from DFT for PE4B with various guest molecules. | pg *S*13 |
| 8. | Figure S8: Experimental STM images and atomistic structures from DFT of PE3A with various guests. | pg *S*14 |
| 9. | Figure S9: Atomistic structures from DFT of host assemblies built up of various ‘test case’ molecules | Pg *S*15 |
| 9. | References | pg *S*16 |

## Synthesis of phenyleneethynylene derivatives

The synthesis of molecule **PE4A** was carried out by following a previously reported procedure.[1] Synthesis and characterization of carboxylic acid derivatives of phenyleneethynylene, **PE4B and PE3A** were carried out by following the reactions shown in Figure S1. The key step for the synthesis of phenyleneethynylene derivatives **PE3A** and **PE4B** is Sonogashira cross coupling reaction, using Pd(II) as catalyst.[2-3] The intermediate compound **3** was prepared by the dimerization of compound **2** by following Glaser reaction using CuCl as a catalyst, which on further reaction with 1 equivalent of phenylacetylene, using Pd(II) as the catalyst, gave the intermediate **4.** Compounds **1** and **4** on reaction with dimethyl 5-ethynylisophthalate (note: 2 equivalents for **1** and 1 equivalent for **4**) yielded products **6** and **5,** respectively. The precursors of carboxylic acid derivatives (**5** and **6**) were purified using recycling HPLC to maintain a high purity of the final products. The purified products **5** and **6** were hydrolyzed using KOH to yield the corresponding acid derivatives **PE4B** and **PE3A**, respectively. All the compounds were characterized using various analytical and spectroscopic techniques. $^{1}$H NMR and $^{13}$C NMR spectra were recorded on Bruker 500 MHz spectrometers. Elemental analysis was carried out using Elementar Vario Micro Cube. All melting points are uncorrected and were determined with a Stuart SMP30 melting point apparatus.

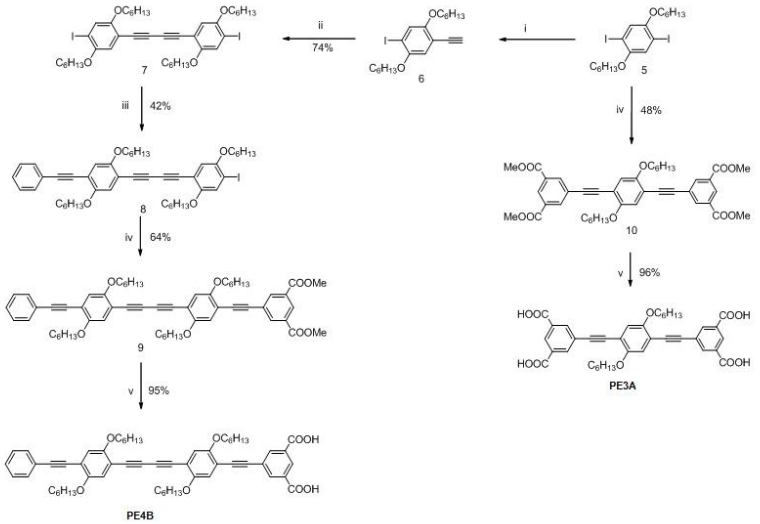

**Figure S1.** i) $Pd(PPh_3)_2Cl_2$, CuI, $PPh_3$, disopropylamine, TMSA; KOH, MeOH. ii) CuCl, diazabicyclo[5.4.0]undec-7-ene; N,N-diisopropylethylamine; iii) $Pd(PPh_3)_2Cl_2$, CuI, $PPh_3$, disopropylamine, phenylacetylene; iv) $Pd(PPh_3)_2Cl_2$, CuI, $PPh_3$, disopropylamine, dimethyl 5-ethynylisophthalate; v) KOH, MeOH.

**Synthesis of 4.** Compound **3** (1.5 g, 1.75 mmol) was kept in a two necked flask and deoxygenated by purging with argon for 15 min, to which $Pd(PPh_3)_2Cl_2$ (24.5 mg, 0.035mmol), CuI (6.6 mg, 0.035 mmol), $PPh_3$ (18 mg, 0.07 mmol), phenyl acetylene (0.178 g, 1.75 mmol), diisopropylamine (10 mL), and THF (30 mL) were added. The mixture was stirred for 12 h at room temperature. After removal of the solvent under reduced pressure, the crude product was purified by column chromatography over silica gel (100-200 mesh) using a mixture (1:10) of $CHCl_3$ and hexane as eluent to yield 0.6 g (42 %) of **4** as a yellow solid. mp 123-124 $^oC$; $^1H$ NMR (500 MHz, $CDCl_3$, TMS): $\delta$ 7.56-7.54 (m, 2H), 7.37-7.35 (d, 3H), 7.31 (s, 1H), 7.00 (d, 2H), 6.89 (s, 1H), 4.01-3.94 (m, 8H), 1.86-1.81 (m, 8H), 1.59-1.52 (m, 7H), 1.38-1.37 (m, 16H), 0.94-0.91 (m, 12H). $^{13}C$ NMR (500 MHz, $CDCl_3$, TMS): $\delta$ 155.59, 155.01, 153.49, 151.79, 131.61, 128..33, 123.99, 123.34, 117.73, 116.95, 116.62, 115.21, 112.48, 112.31, 95.53, 87.09, 85.81, 70. 11, 70.09, 69.80, 69.65, 31.60, 31.53, 31.51, 31.49, 30.92, 29.70, 29.10, 29.08, 25.73, 25.61, 25.57, 22.59, 14.03. Anal. Calcd for $C_{48}H_{61}IO_4$ : C, 69.55; H, 7.42; I, 15.31; O, 7.72 Found: C, 69.62; H, 7.34.

**Synthesis of 5.** Compound **4** (525 mg, 0.64 mmol) was kept in a two necked flask and deoxygenated by purging with argon for 15 min, to which $Pd(PPh_3)_2Cl_2$ (21 mg, 0.03 mmol), CuI (6 mg, 0.03 mmol), $PPh_3$ (15 mg, 0.06 mmol), dimethyl 5-ethynylisophthalate (140 mg, 0.64 mmol), diisopropylamine (10 mL), and THF (30 mL) were added. The mixture was stirred for 12 h at room temperature. After removal of the solvent under reduced pressure, the crude product was purified by column chromatography over silica gel (100-200 mesh) using a mixture (1:10) of $CHCl_3$ and hexane as eluent to yield 375 mg (64 %) of **5** as a yellow solid. mp 117-118 $^oC$; $^1H$ NMR (500 MHz, $CDCl_3$, TMS): $\delta$ 8.55 (s, 1H), 8.28 (s, 2H), 7.47-4.45 (m, 5H), 6.93-6.91 (d, 4H), 3.96-3.89 (m, 14H), 1.78-1.75 (t, 8H), 1.45-1.43 (m, 8H), 1.30-1.27 (m, 8H), 0.85-0.81 (m, 12H). $^{13}C$ NMR (500 MHz, $CDCl_3$, TMS): $\delta$ 206.97, 165.62, 155.04, 153.04, 136.45, 134.16, 131.62, 130.99, 128.39, 128.33, 127.61, 124.48, 123.34, 117.74, 117.04, 116.95, 113.34, 87.81, 85.81, 69.85, 69.82, 69.66, 52.53, 31.60, 31.53, 30.92, 29.28, 29,18, 29.10, 25.73, 25.61, 22.63, 22.61, 14.04, 14.02. Anal. Calcd for $C_{60}H_{70}O_8$ : C, 78.60; H, 7.68; O, 13.92 Found: C, 78.32; H, 7.39.

**Synthesis of 6.** Compound **1** (700 mg, 1.30 mmol) was kept in a two necked flask and deoxygenated by purging with argon for 15 min, to which $Pd(PPh_3)_2Cl_2$ (91 mg, 0.13 mmol), CuI (12 mg, 0.13 mmol), $PPh_3$ (67 mg, 0.26 mmol),dimethyl 5-ethynylisophthalate (566 mg, 1.60 mmol), diisopropylamine (10 mL), and THF (30 mL) were added. The mixture was stirred for 12 h at room temperature. After removal of the solvent under reduced pressure, the crude product was purified by column chromatography over silica gel (100-200 mesh) using a mixture (3:2) of $CHCl_3$ and hexane as eluent to yield 450 mg (48 %) of **6** as a yellow solid. mp 162-163 °C; $^1H$ NMR (500 MHz, $CDCl_3$, TMS): δ 8.64 (s, 2H), 8.38 (s, 4H), 7.06 (s, 2H), 4.09-4.06 (t, 4H), 3.98 (s, 6H), 1.91-1.88 (t, 4H), 1.60-1.56 (m, 4H), 1.43-1.34 (m, 8H), 0.92-0.89 (m, 6H). $^{13}C$ NMR (500 MHz, $CDCl_3$, TMS): δ 165.62, 153.79, 134.36, 136.44, 130.98, 130.08, 124.52, 117.01, 113.78, 92.92, 87.83, 69.71, 52.52, 31.54, 29.23, 25.76, 22.59, 14.00. Anal. Calcd for $C_{42}H_{46}O_{10}$ : C, 70.97; H, 6.52; O, 22.51 Found: C, 71.12; H, 6.39.

**Synthesis of PE4B and PE3A.** The synthesis of compounds **PE4B** and **PE3A** was achieved through hydrolysis of corresponding ester derivatives (**5** and **6)** by adopting a general procedure given below. The ester derivative (100 mg) was dissolved in THF (10 mL) and to this solution KOH (1.0 g) and methanol (20 mL) were added and refluxed for about 5 h. The reaction mixture was cooled to room temperature, neutralized with dilute HCl and water (50 mL) was added. The precipitated product was either filtered or extracted with chloroform. The solvent was evaporated under reduced pressure to yield the corresponding carboxylic acid as solid product. The spectral details of the individual compounds are provided below.

**PE4B**. 92 mg (95%). mp 216 °C; $^1H$ NMR (500 MHz, THF-$d_8$,TMS): δ 8.50 (s, 1H), 8.20 (s, 2H), 7.39-7.37 (d, 2H), 7.25-7.21 (m, 3H), 7.04-6.96 (m, 4H), 3.95-3.91 (m, 8H), 1.73-1.69 (m, 8H), 1.50-1.42 (m, 8H), 1.29-1.24 (m, 16H), 0.84-0.78 (m, 12H). $^{13}C$ NMR (500 MHz, THF-$d_8$, TMS): δ 165.33, 155.06, 155.03, 153.82, 153.64, 135.75, 131.94, 131.25, 128.22, 128.17, 124.15, 123.60, 117.27, 116.56, 116.47, 114.35, 112.86, 112.25, 85.87, 78.79, 69.25, 69.20, 66.13, 31.58, 31.54, 29.25, 29.10, 29.08, 25.85, 25.73, 22.55, 13.45, 13.41. Anal. Calcd for $C_{58}H_{66}O_8$ : C, 78.17; H, 7.47; O, 14.34 Found: C, 78.53; H, 6.65.

**PE3A** 88 mg (96%). mp above 300 °C; $^1H$ NMR (500 MHz, THF-$d_8$,TMS): δ 8.52 (s, 2H), 8.23 (s, 4H), 7.07 (s, 2H), 4.00-3.98 (t, 4H), 1.78-1.75 (t, 4H), 1.53-1.50- (m, 4H), 1.435-1.30 (m, 8H), 0.92-0.89 (m, 6H). $^{13}C$ NMR (500 MHz, THF-$d_8$, TMS): δ 165.32, 153.95, 135.75, 131.88, 130.08, 124.28, 116.59, 113.75, 92.70, 87.56, 69.20, 31.59, 29.30, 25.86, 22.57, 14.42. Anal. Calcd for $C_{38}H_{38}O_{10}$ : C, 69.71; H, 5.85; O, 24.44 Found: C, 70.01; H, 5.39.

**STM imaging:** The STM imaging was carried out in the constant-height mode under ambient conditions at the liquid-solid interface immediately after drop casting the solution on HOPG surface using a multimode scanning tunneling microscope.

## Computational Methods

### Density functional theory calculations

Density functional theory (DFT) calculations were performed using the SIESTA package.[4-5] We used a double zeta polarized basis set with confinement energy of 0.01 Ry. The mesh cutoff size of the real-space grid was taken to be 250 Ry. Troullier-Martin norm-conserving pseudopotentials were used.[6] Exchange-correlation interactions were treated within the PBE form of the generalized gradient approximation.[7] Van der Waals interactions were incorporated by using the DFT-D2 treatment.[8] A vacuum with a spacing of 16 Å was introduced along non-repeating directions. For the large unit cells used in this study, the Brillouin zone was sampled at the zone center Γ only. Geometry optimization was performed using the conjugate gradient algorithm. In the calculations, highly oriented pyrolitic graphite (HOPG) was replaced by a single atomic layer of graphene. All atomic coordinates were relaxed until the forces on all atoms were smaller than 0.03 eV/Å. Basis set superposition errors were eliminated by applying the counterpoise correction procedure.[9-10] The simulated STM images were obtained using the Tersoff-Hamann approach in the constant height mode.[11]

In all cases, the unit cells used corresponded to the primitive unit cell of the corresponding HEX or LIN pattern. These primitive unit cells contain three host molecules in the HEX arrangements of PE4A and PE3A, six host molecules in the HEX arrangement of PE4B, and two host molecules in the LIN arrangements of all three host molecules (see Fig. S2, where the black rhombi and rectangles mark the boundaries of the unit cells for the HEX and LIN, respectively). In all cases, the unit cell contains only one guest molecule (when present).

The initial coordinates of the molecules in the gas phase were obtained using the Gaussview software package. The unit cell parameters were varied stepwise, with the constraint that they be commensurate with the underlying lattice constant of graphene. The coordinates of all molecules were allowed to relax, following a three-step procedure: (i) the coordinates of the isolated molecules were then relaxed using DFT. Next, for each set of cell parameters: (ii) a monolayer of molecules was assembled in either the HEX or LIN pattern in the gas phase (i.e., no graphene substrate) and the coordinates were further relaxed, with the constraint that they remain in the x-y plane, (iii) this monolayer was then placed on graphene and the coordinates was further relaxed, with no geometric constraints. The optimized cell parameters thus obtained, along with the corresponding values of the adsorption energy $E_{\text{ads}}$ are listed in Table S1 below.

**Table S1.** Results from DFT for geometries and energetics of the HEX and LIN arrangements of the host molecules when adsorbed on graphene. $N_{\text{host}}$ is the number of host molecules contained in the 2D primitive unit cell, whose optimal size is given by the cell parameter $a$ (in HEX) or $a$,$b$ (in LIN). We see that the optimal lattice parameters as obtained from DFT are very close to those determined from the STM experiments. For ease of reference, the corresponding geometries are given in Fig. S2 below. $E_{\text{ads}}$ is the adsorption energy of the host assemblies on graphene.

| Host molecule → | | PE4A | PE4B | PE3A |
|---|---|---|---|---|
| HEX | $N_{\text{host}}$ (molecules/cell) | 3 | 6 | 3 |
| | $a$ [DFT](nm) | 4.23 | 5.23 | 3.24 |
| | $a$ [expt](nm) | 4.2 ± 0.1 | 5.2 ± 0.1* | 3.2 ± 0.1* |
| | $E_{\text{ads}}$ (meV/Å$^2$) | -13.1 | -14.2 | -16.8 |
| LIN | $N_{\text{host}}$ (molecules/cell) | 2 | 2 | 2 |
| | $a$ [DFT](nm) | 1.49 | 1.49 | 1.49 |
| | $a$ [expt](nm) | NA | 1.5 ± 0.1 | 1.5 ± 0.1 |
| | $b$ [DFT] (nm) | 6.38 | 5.17 | 3.88 |
| | $b$ [expt](nm) | NA | 5.2 ± 0.1 | 3.9 ± 0.1 |
| | $E_{\text{ads}}$ (meV/Å$^2$) | -12.0 | -15.2 | -17.1 |

* In presence of coronene

### Determination of form of Host Descriptor $\eta$

The form of $\eta$ was determined by demanding that it involve a combination of powers of the three physical parameters $L$, $N_{\text{COOH}}$ and $N_{\text{alkoxy}}$; this combination should correlate linearly with the difference in $\Delta G$ between the HEX and LIN patterns.[12] Here, $L$ is the length of the central backbone of the host molecule, measured between the centers of the terminal phenyl rings, $N_{\text{COOH}}$ is the number of terminal carboxyl groups, and $N_{\text{alkoxy}}$ is the number of alkoxy side chains. The power of $N_{\text{COOH}}$ was set = 1 (based on the physical argument that the main stabilization of the assembly comes from dimeric hydrogen bonds between COOH groups). Thus the trial form chosen was $\eta = N_{\text{COOH}} \times (L)^{\alpha} \times [1 + (N_{\text{alkoxy}})^{\beta}]$; this form ensures that $\eta$ does not become zero or infinite in the absence of alkoxy side chains. The powers $\alpha$ and $\beta$ were required to be integers or simple fractions and then determined so as to minimize the error obtained by a linear fit; this led to $\alpha = 1$ and $\beta = 1/8 = 0.125$.

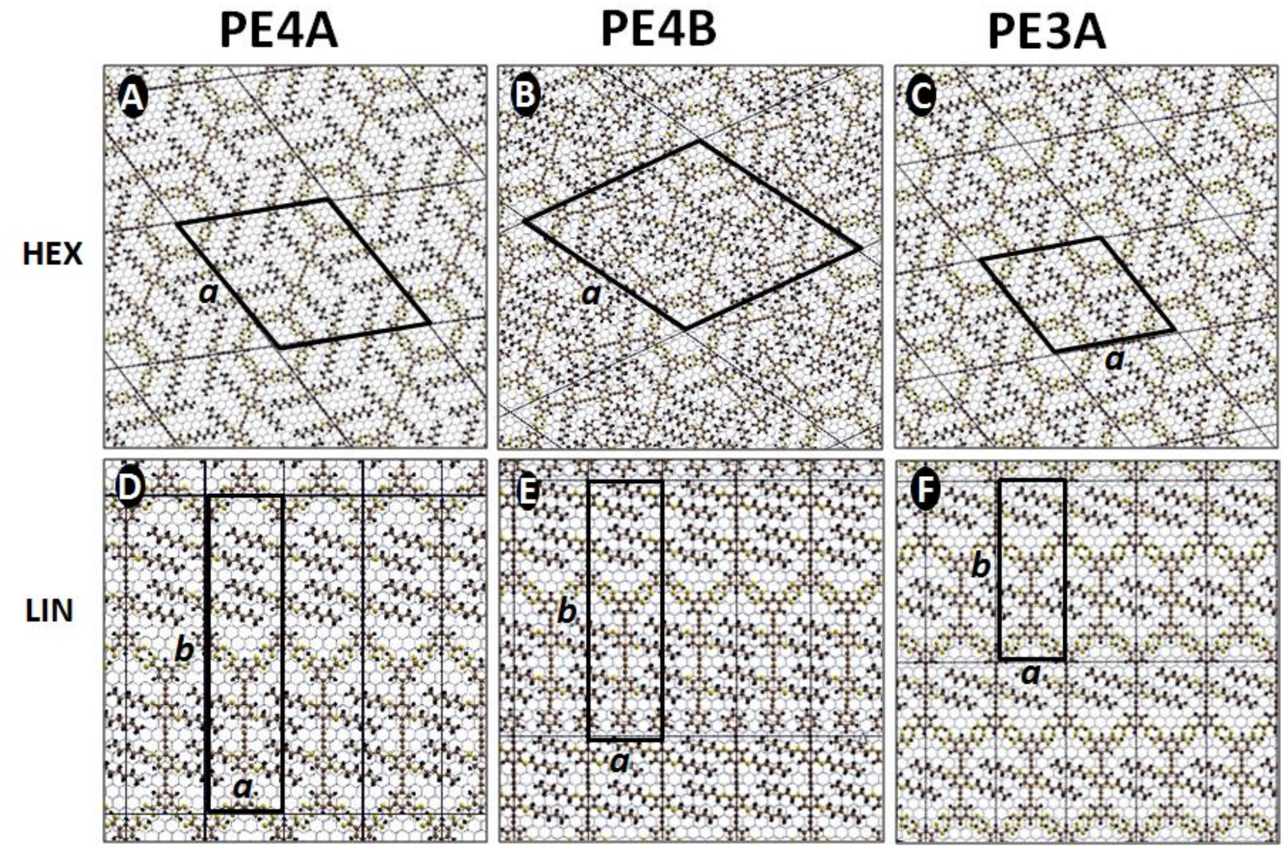

**Figure S2**: Atomistic structures of host assemblies on graphene. All coordinates have been relaxed using DFT calculations. Black rhombi and rectangles mark the boundaries of the unit cells for the HEX and LIN patterns, respectively. Atomic colors: C of host: brown, H: black, O: yellow, C of graphene: gray, C of guest: green.

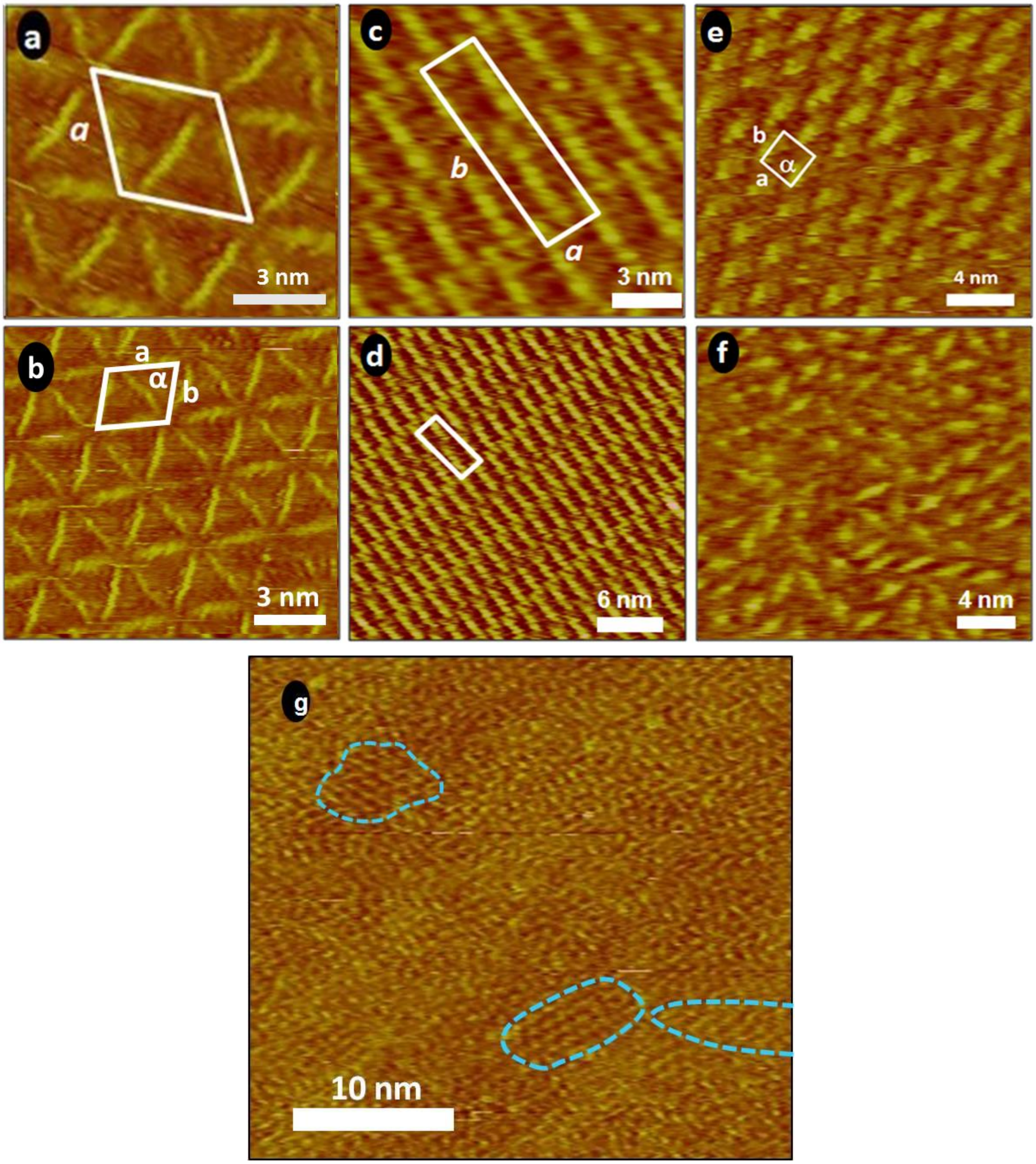

**Figure S3**. Experimental STM images of host assemblies. PE4A in HEX pattern: (a) zoomed in view (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –1500 mV; $I_t$ = 60 pA) and (b) larger area STM image (scan size: 30 × 30 nm$^2$; $V_{\text{bias}}$ = –850 mV; $I_t$ = 50 pA). PE4B in LIN pattern: (c) zoomed in view (scan size: 15 × 15 nm$^2$; $V_{\text{bias}}$ = –1450 mV; $I_t$ = 100 pA) and (d) larger area STM image (scan size: 40 × 40 nm$^2$; $V_{\text{bias}}$ = –1450 mV; $I_t$ = 100 pA). PE3A: (e) local LIN domains (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –1925 mV; $I_t$ =230 pA), (f) random 2D glass arrangement (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –1750 mV; $I_t$ =230 pA) of PE3A and (g) larger area, showing 2D glass but with some domains of ordered LIN (outlined by blue dashed lines).

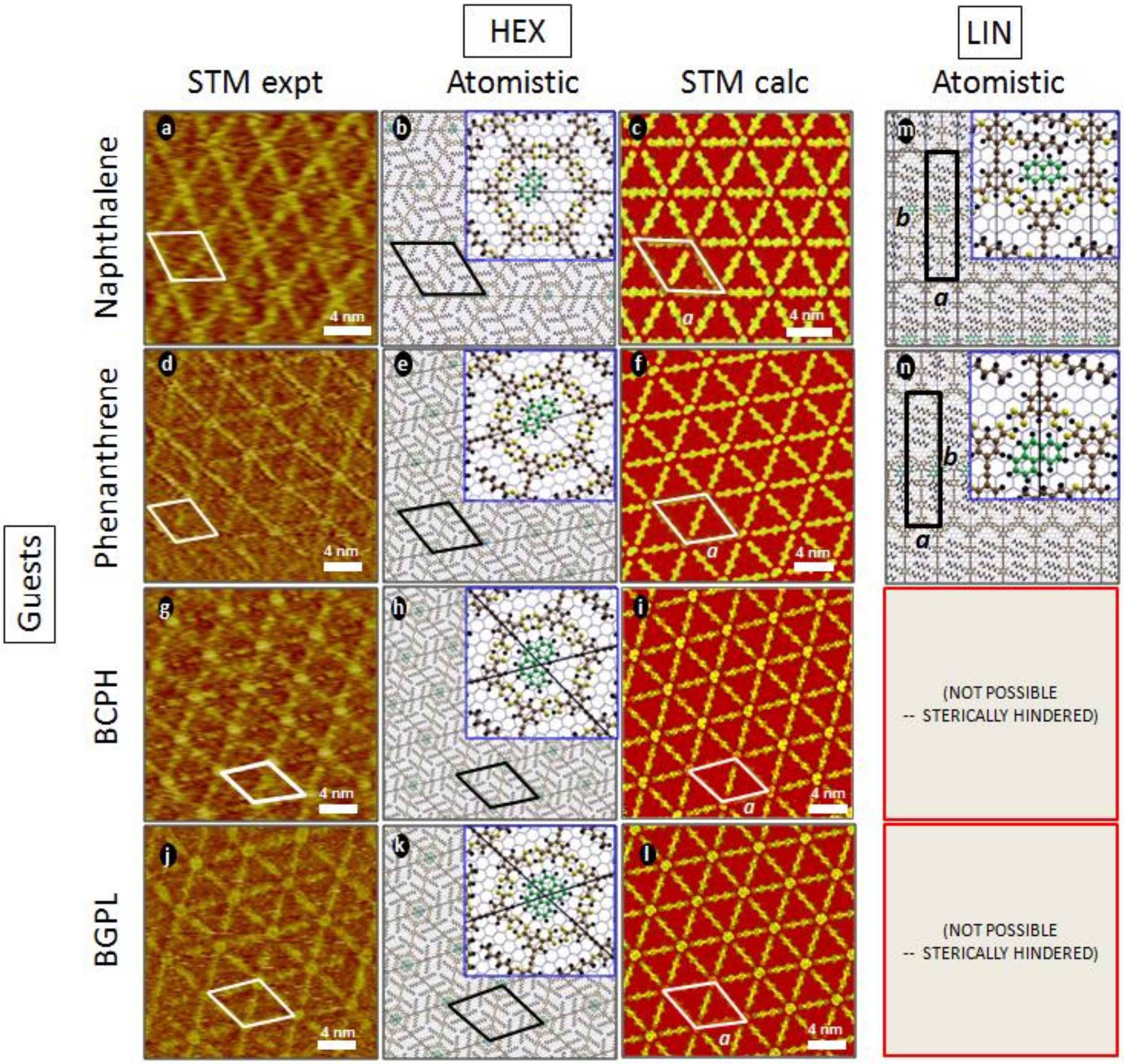

**Figure S4.** STM images and atomistic structures of PE4A with guest molecules on HOPG or graphene in HEX (a)-(l) and LIN (m-n) pattern. Experimental STM images of PE4A with (a) naphthalene (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –600 mV; It = 120 pA), (d) phenanthrene (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –1150 mV; $I$t = 153 pA) , (g) ) BCPH (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –539 mV; $I_t$ = 60 pA and (j) BGPL (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –1400 mV; $I_t$ = 100 pA) as guest in HEX pattern. The corresponding atomistic structures of PE4A with (b) naphthalene, (e) phenanthrene, (h) BCPH and (k) BGPL as guest. The insets are zoomed in views. Simulated STM images of PE4A with (c) naphthalene ($V_{\text{bias}}$ = –2720 mV), (f) phenanthrene ($V_{\text{bias}}$ = –2720 mV), (i) BCPH ($V_{\text{bias}}$ = –2720 mV) and (l) BGPL ($V_{\text{bias}}$ = –2720 mV). Atomistic structures of PE4A with (m) naphthalene and (n) phenanthrene in LIN pattern. The black/white lines indicate the unit cell; with a and b are the unit cell parameter. Atomic colors: C of host: brown, H: black, O: yellow, C of graphene: gray, C of guest: green.

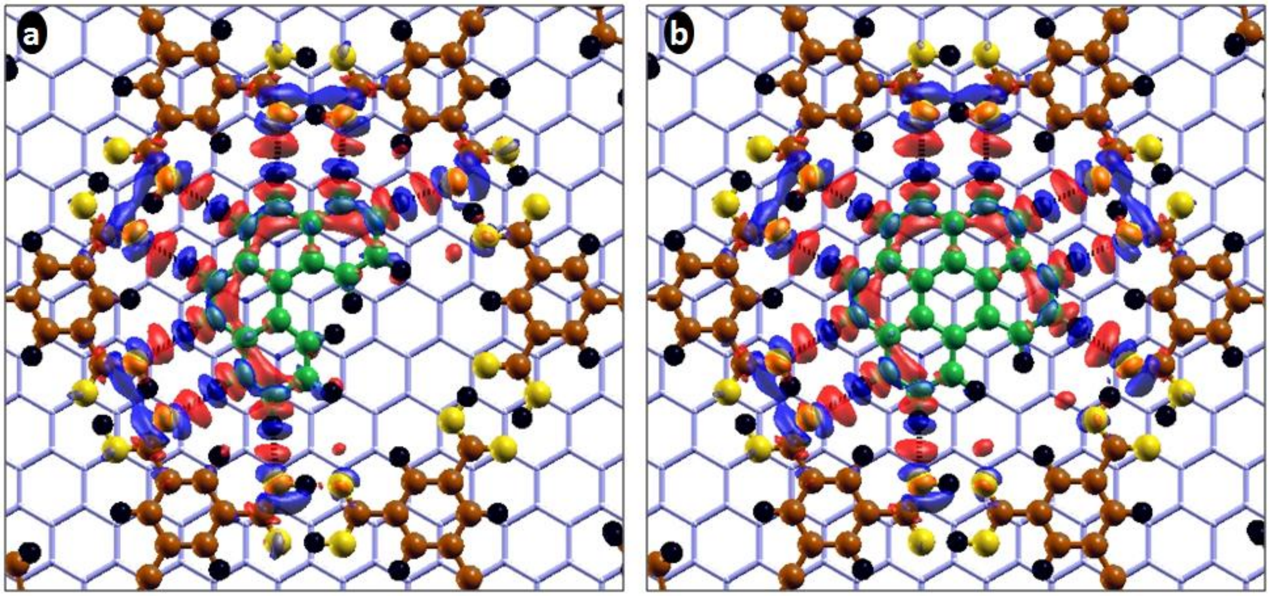

**Figure S5.** Theoretically computed charge redistribution plots between PE4A and different guest molecules. PE4A (a) with benzo-c-phenanthrene (BCPH) and (b) benzo-ghi-perylene (BGPL) on graphene. Red and blue lobes indicate electron accumulation and depletion respectively, plotted at an isosurface value of 0.0002 e/bohr$^3$. Note the clear signatures of eight and ten host-guest hydrogen bonds in (a) and (b), respectively. Color scheme: brown: C of PE4A; yellow: O, black: H; green: C of guest. Gray lines indicate bonds between C atoms of graphene.

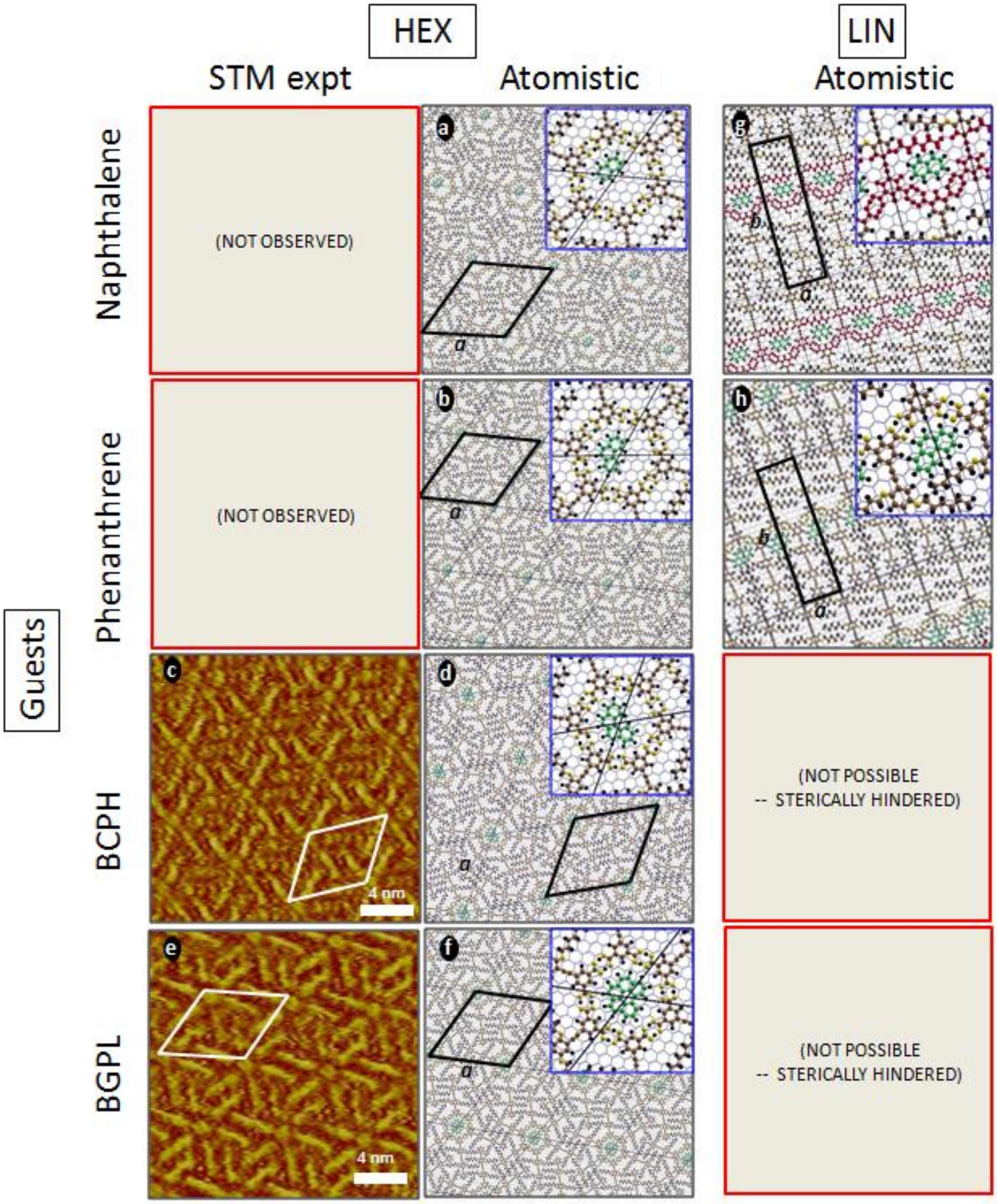

**Figure S6.** Atomistic structures from DFT and experimental STM images of PE4B with various guests in (a)-(f) HEX and (g)-(h) LIN patterns. Atomistic structures of PE4B with (a) naphthalene (b) phenanthrene, (d) BCPH and (f) BGPL (PE4B in HEX); (g) naphthalene and (h) phenanthrene (PE4B in LIN). Experimental STM images of PE4B with (c) BCPH (scan size: 20 × 20 $\text{nm}^2$; $V_{\text{bias}}$ = −1018 mV; $I_t$ = 120 pA) and (e) BGPL (scan size: 20 × 20 $\text{nm}^2$; $V_{\text{bias}}$ = −1500 mV; $I_t$ = 150 pA). The black/white lines indicate the unit cell; with *a* and *b* are the unit cell parameter. Atomic colors: C of host: brown, H: black, O: yellow, C of graphene: gray, C of guest: green.

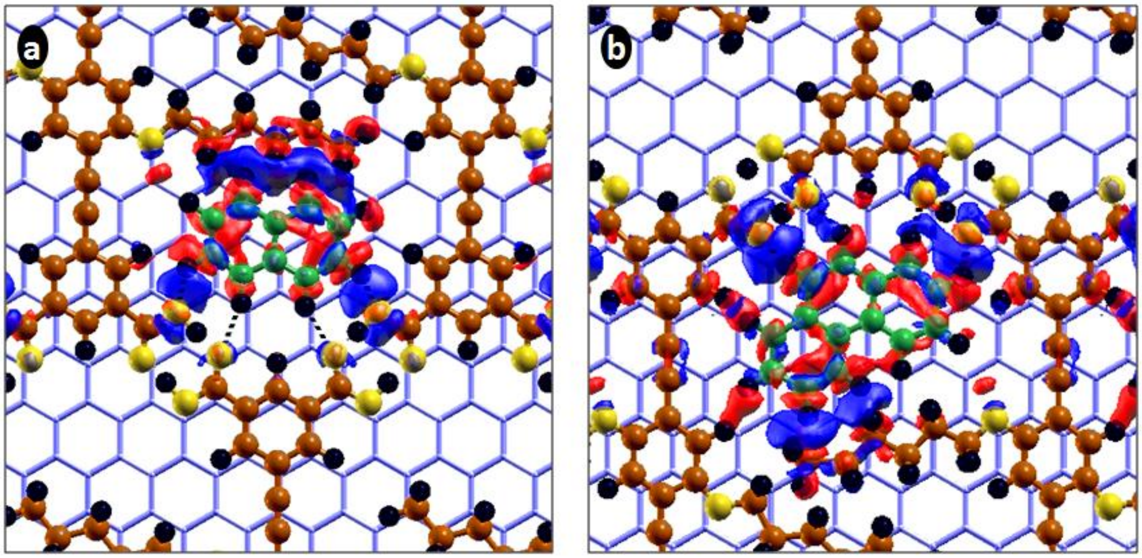

**Figure S7.** Theoretically computed charge redistribution plots between PE4B and different guest molecules. PE4B with (a) naphthalene and (b) phenanthrene on graphene in LIN pattern. Red and blue lobes indicate electron accumulation and depletion respectively, plotted at an isosurface value of 0.00045 e/bohr$^3$. Note the total number of host-guest bonds are hard to define precisely in both (a) and (b), due to lower symmetry of LIN pattern. Color scheme: brown: C of PE4B; yellow: O, black: H; green: C of guests. Gray lines indicate bonds between C atoms of graphene.

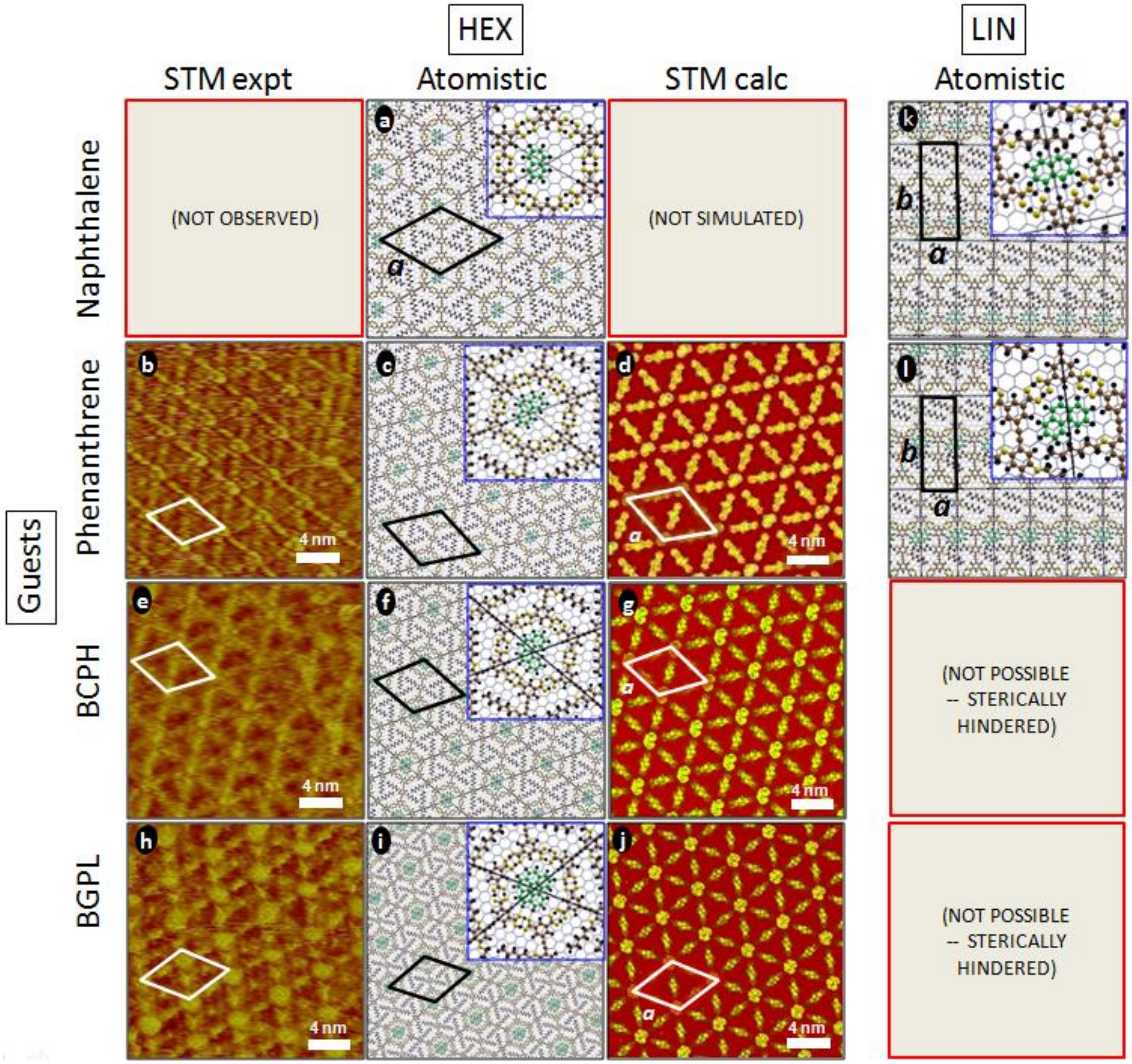

**Figure S8.** STM images and atomistic structures of PE3A with various guests in HEX (a)-(j) and LIN (k)-(l) patterns. Experimental STM images of PE3A with (b) phenanthrene (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –553 mV; $I_t$ = 229 pA), (e) BCPH (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –600 mV; $I_t$ = 450 pA) and (h) BGPL (scan size: 20 × 20 nm$^2$; $V_{\text{bias}}$ = –1214 mV; $I_t$ = 122 pA) as guests in HEX pattern. Atomistic structures from DFT of PE3A with (a) naphthalene, (c) phenanthrene, (f) BCPH and (i) BGPL in HEX pattern; (k) naphthalene and (l) phenanthrene in LIN pattern. Simulated STM images of PE3A with (d) phenanthrene, (g) BCPH and (j) BGPL as guests. Simulated STM images are obtained at $V_{\text{bias}}$ = –2720 mV. The black/white lines indicate the unit cell; with *a* and *b* are the unit cell parameters. Atomic colors: C of host: brown, H: black, O: yellow, C of graphene: gray, C of guest: green.

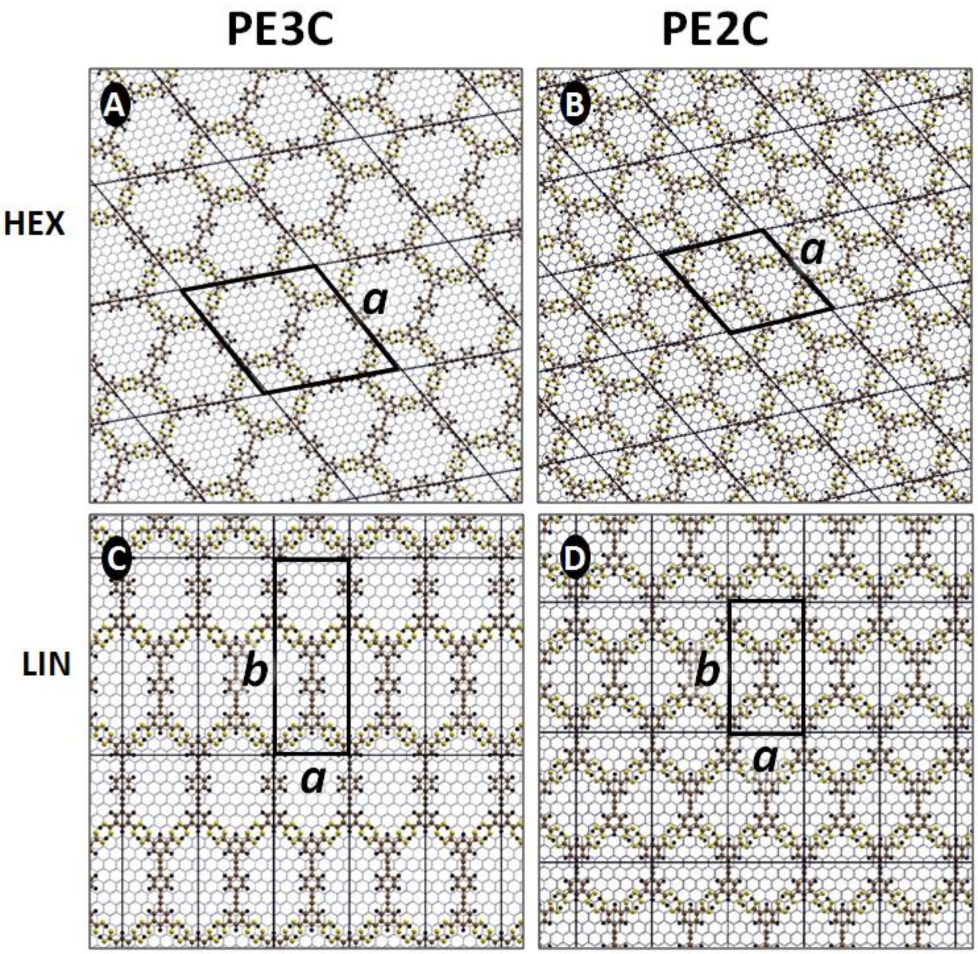

**Figure S9**: Atomistic structures of host assemblies of the test molecules on graphene. All coordinates have been relaxed using DFT calculations. Black rhombi and rectangles mark the boundaries of the unit cells for the HEX and LIN patterns, respectively. Atomic colors: C of host: brown, H: black, O: yellow, C of graphene: gray, C of guest: green.

# References

1. Zalake, P.; Thomas, K. G. Role of Hydrogen Bonding on the Self-Organization of Phenyleneethynylenes on Surfaces. *Langmuir* **2013,** *29* (7), 2242-2249.
2. Samori, P.; Francke, V.; Enkelmann, V.; Müllen, K.; Rabe, J. P. Synthesis and Solid State Structures of Functionalized Phenyleneethynylene Trimers in 2D and 3D. *Chem. Mater.* **2003,** *15* (5), 1032-1039.
3. Ramesh, A. R.; Thomas, K. G. Directional Hydrogen Bonding Controlled 2d Self-Organization of Phenyleneethynylenes: From Linear Assembly to Rectangular Network. *Chem. Commun.* **2010,** *46* (20), 3457-3459.
4. Jose, M. S.; Emilio, A.; Julian, D. G.; Alberto, G.; Javier, J.; Pablo, O.; Daniel, S.-P. The Siesta Method for Ab Initio Order- N Materials Simulation. *J. Phys. Condens. Matter* **2002,** *14* (11), 2745.
5. Sánchez-Portal, D.; Ordejón, P.; Artacho, E.; Soler, J. M. Density-Functional Method for Very Large Systems with Lcao Basis Sets. *Int. J. Quantum Chem.* **1997,** *65* (5), 453-461.
6. Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991,** *43* (3), 1993-2006.
7. Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996,** *77* (18), 3865-3868.
8. Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006,** *27* (15), 1787-1799.
9. Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970,** *19* (4), 553-566.
10. Brauer, B.; Kesharwani, M. K.; Martin, J. M. L. Some Observations on Counterpoise Corrections for Explicitly Correlated Calculations on Noncovalent Interactions. *J. Chem. Theory Comput.* **2014,** *10* (9), 3791-3799.
11. Tersoff, J.; Hamann, D. R. Theory and Application for the Scanning Tunneling Microscope. *Phys. Rev. Lett.* **1983,** *50* (25), 1998-2001.
12. Ghiringhelli, L. M.; Vybiral, J.; Levchenko, S. V.; Draxl, C.; Scheffler, M. Big Data of Materials Science: Critical Role of the Descriptor. *Phys. Rev. Lett.* **2015,** *114* (10), 105503-105508.