Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2018

Supporting Information

Building a 22-ring nanographene by combining in-solution and on-surface synthesis

Rafał Zuzak,[†a] Jesus Castro,[†b] Pedro Brandimarte,[c] Mads Engelund,[d] Agustín Cobas,[b] Piotr Piątkowski,[a] Marek Kolmer,[a] Dolores Pérez,[b] Enrique Guitián,[b] Marek Szymoński,[a] Daniel Sánchez-Portal,[c,e] Szymon Godlewski,[*a] Diego Peña[*b]

[a] Centre for Nanometer-Scale Science and Advanced MaterialsNANOSAM, Faculty of Physics, Astronomyand Applied Computer Science, Jagiellonian University, Łojasiewicza 11, 30-348 Kraków (Poland). E-mail: szymon.godlewski@uj.edu.pl.

[b] Centro de Investigación en Química Biolóxica e Materiais Moleculares (CIQUS), Departamento de Química Orgánica, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain. E-mail: diego.pena@usc.es.

[c] Donostia International Physics Center, 20018 Donostia-San Sebastián, Spain.

[d] Espeem S.A.R.L. (espeem.com), 9 rue de Haut-Fournaux, L-4365 Esch-sur-Alzette, Luxembourg.

[e] Centro de Física de Materiales (CSIC-UPV/EHU), 20018 Donostia-San Sebastián, Spain

† These authors contributed equally to this work.

Table of Contents

1. Synthesis of hexapole pentahelicene 2

1.1. General methods

All reactions were carried out under argon using oven-dried glassware. TLC was performed on Merck silica gel 60 F_{254}; chromatograms were visualized with UV light (254 and 360 nm). Flash column chromatography was performed on Merck silica gel 60 (ASTM 230-400 mesh). ^{1}H and ^{13}C NMR spectra were recorded at 300 and 75 MHz or 500 and 125 MHz (Varian Mercury 300 or Bruker DPX-500 instruments), respectively. Low-resolution electron impact mass spectra were determined at 70 eV on a HP-5988A instrument. MALDI-TOF spectra were determined on a Bruker Autoflex instrument. 7,8-Dibromopentahelicene (**5**)[1] was prepared following a published procedure (Figure S1). The synthesis of triflate **4** is similar to a recently published procedure.[2] *n*-BuLi was used in solution in hexane (2.4 M). Commercial reagents and anhydrous solvents were purchased from ABCR GmbH, Aldrich Chemical Co., or Strem Chemicals Inc., and were used without further purification. Solvents (MeCN, THF and CH_2Cl_2) were purified by a MBraun SPS -800 Solvent Purification System. CsF was dried under vacuum at 100ºC, cooled under argon and stored in a glove box.

5 **4**

Figure S1. Structure of pentahelicenes **4** and **5**.

1.2. Synthesis of precursors

*1.2a. Synthesis of 3-bromo-4-methoxypentahelicene (**6**)*

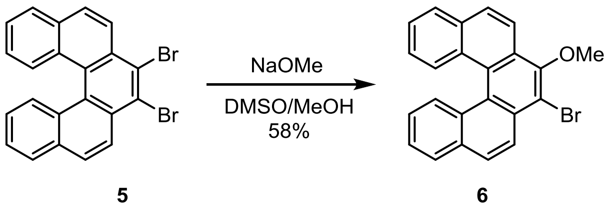

Figure S2. Synthesis of pentahelicene **6**.

To a solution of 7,8-dibromopentahelicene (**5**, 500 mg, 1.15 mmol) in DMSO:MeOH (15:1, 16 mL), a solution of NaOMe (2.40 mL, 1.20 mmol, 0.50 M in MeOH) was added dropwise. The mixture was stirred under argon and reflux overnight. Then, the reaction was quenched by the addition of H_2O and the aqueous phase was extracted with CH_2Cl_2 (2 x 20 mL). The combined organic phases were dried over anhydrous Na_2SO_4 and the solvent was evaporated under reduced pressure. The resulting mixture was purified by column chromatography (SiO_2, CH_2Cl_2/hexane, 3:7) to afford **6** (258 mg, 58%) as an orange solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ: 8.41 (d, *J* = 8.9 Hz, 1H), 8.34 (d, *J* = 8.6 Hz, 2H), 8.24 (d, *J* = 8.9 Hz, 1H), 8.11 – 7.87 (m, 4H), 7.53 (q, *J* = 6.9 Hz, 2H), 7.25 (dd, *J* = 8.9, 6.5 Hz, 2H), 4.11 (s, 3H) ppm. ^{13}C NMR (75 MHz, $CDCl_3$) δ: 152.56 (C), 132.64 (C), 132.09 (C), 130.87 (2xC), 130.49 (C), 129.40 (2xCH), 128.68 (CH), 128.45 (CH), 128.00 (CH), 127.85 (CH), 127.78 (C), 127.72 (C), 126.86 (CH), 126.66 (CH), 125.91 (C), 124.98 (2xCH), 124.94 (CH), 120.08 (CH), 114.81 (C), 61.81(CH_3) ppm. MS (EI) *m/z* (%): 386 (M^+, 76), 307 (44).

*1.2b. Synthesis of 3-bromopentahelicen-4-ol (**7**)*

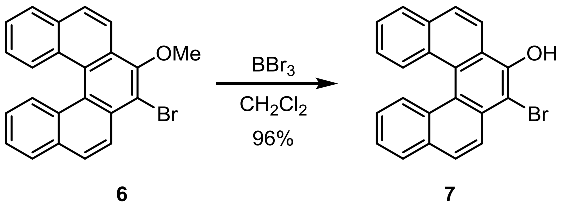

Figure S3. Synthesis of pentahelicene **7**.

A solution of BBr_3 (6.02 mL, 6.02 mmol, 1.0 M in CH_2Cl_2) was dropwise added to a solution of compound **6** (1.17 g, 3.02 mmol) in CH_2Cl_2 (35 mL) at -78 ºC. The mixture was stirred under argon for 1 h at -78 ºC and for 12 h at room temperature. Then, cold H_2O (30 mL) was added, the phases were separated, and the aqueous layer was extracted with CH_2Cl_2 (3 x 30 mL). The combined organic layers were dried over anhydrous Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was filtered through a short pad of silica gel (SiO_2, CH_2Cl_2/hexane, 2:3), affording compound **7** (1.07 g, 96%) as an orange solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ: 8.36 (dd, *J* = 8.7, 3.5 Hz, 2H), 8.29 (d, *J* = 8.6 Hz, 1H), 8.17 (d, *J* = 8.8 Hz, 1H), 7.95 (dd, *J* = 12.0, 8.2 Hz, 4H), 7.59 – 7.43 (m, 2H), 7.27-7.21 (m, 2H), 6.34 (s, 1H) ppm. ^{13}C NMR (75 MHz, $CDCl_3$) δ: 147.72 (C), 132.89 (C), 131.38 (C), 130.75 (C), 130.50 (C), 130.20 (C), 129.59 (CH), 129.10 (CH), 128.93 (CH), 128.03 (2xCH), 127.89 (CH), 127.62 (C), 126.89 (CH), 126.01 (CH), 124.98 (CH), 124.83 (CH), 124.23 (CH), 123.84 (C), 123.37(C), 120.16 (CH), 105.54 (C) ppm. MS (EI) *m/z* (%): 372 (M^+, 44), 293 (49).

1.2c. Synthesis of triflate ***4***

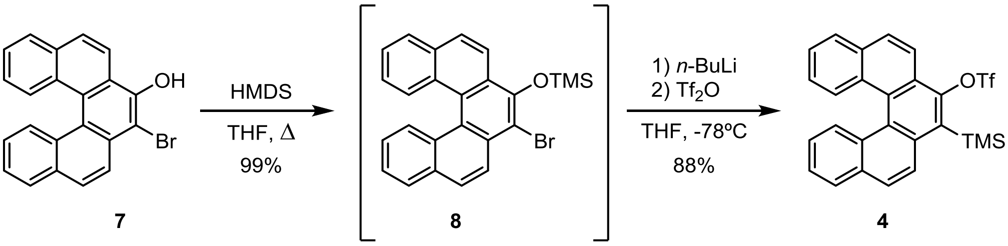

Figure S4. Synthesis of triflate **4**.

A solution of compound **7** (1.12 g, 3.0 mmol) and hexamethyldisilazane (HMDS, 0.815 mL, 3.9 mmol) in THF (40 mL) was stirred under argon and refluxed for 2 h. The solvent was evaporated under reduced pressure and the quantitative formation of **8** was checked by ^{1}H NMR. Then, the crude product was dissolved in THF (40 mL), the solution was cooled to -78ºC, *n*-BuLi (1.5 mL, 2.4 M, 3.6 mmol) was added dropwise and the reaction mixture was stirred under argon for 1 h. Then, Tf_2O (0.700 mL, 4.2 mmol) was added dropwise and stirring was kept up for 30 min at -78ºC. Then, saturated aqueous solution of $NaHCO_3$ (30 mL) was added, the phases were separated, and the aqueous layer was extracted with CH_2Cl_2 (3 x 25 mL). The combined organic layers were collected and dried over Na_2SO_4, filtered and concentrated under reduced pressure. The residue was filtered through a short pad of silica gel (SiO_2, CH_2Cl_2/hexane, 1:4), affording compound **4** (1.32 g, 88%) as a yellow solid. Data for **8**: ^{1}H NMR (300 MHz, $CDCl_3$) δ: 8.44 – 8.25 (m, 3H), 8.18 (d, *J* = 8.9 Hz, 1H), 7.96 (dd, *J* = 10.3, 7.9 Hz, 4H), 7.50 (dt, *J* = 10.7, 7.4 Hz, 2H), 7.25 (d, *J* = 6.1 Hz, 2H), 0.44 (s, 9H) ppm. Data for **4**: ^{1}H NMR (300 MHz, $CDCl_3$) δ: 8.33 (dt, *J* = 8.6, 0.6 Hz, 2H), 8.25 (d, *J* = 8.9 Hz, 1H), 8.16 – 7.92 (m, 5H), 7.55 (dddd, *J* = 7.0, 5.8, 4.8, 2.4 Hz, 2H), 7.34 – 7.21 (m, 2H), 0.66 (s, 9H) ppm. ^{13}C NMR (75 MHz, $CDCl_3$) δ: 148.16 (C), 134.99 (C), 132.64 (C), 132.18 (C), 130.87 (C), 130.59 (C), 130.36 (C), 129.83 (CH), 129.74 (C), 129.38 (C), 129.30 (CH), 128.99 (CH), 128.17 (CH), 128.04 (2xCH), 127.81 (CH), 127.44 (CH), 127.06 (CH), 127.00 (C), 125.66 (CH), 125.48 (CH), 125.12 (CH), 2.36 (CH_3) ppm.* MS (EI) *m/z* (%): 498 (M^+, 80), 349 (44).

1.2d. Synthesis of C_2*-symmetric hexapole pentahelicene* ***2***

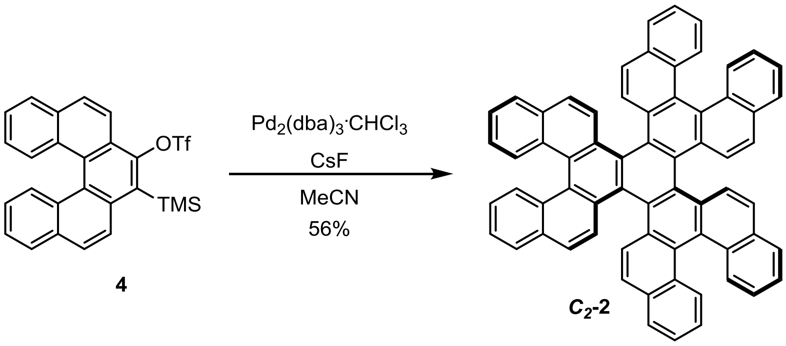

Figure S5. Synthesis of C_2-**2**.

To a solution of **4** (60 mg, 0.120 mmol) in MeCN (3 mL) $Pd_2(dba)_3·CHCl_3$ (12.4 mg, 0.012 mmol) and anhydrous CsF (110 mg, 0.723 mmol) were added and the mixture was stirred under argon atmosphere for 24 h at room temperature. The reaction was quenched by the addition of H_2O (5 mL) and the aqueous phase was extracted with CH_2Cl_2 (3 x 5 mL). The combined organic phases were dried over anhydrous Na_2SO_4 and the solvent was evaporated under reduced pressure. The resulting mixture was purified by column chromatography (SiO_2, CH_2Cl_2/hexane, 1:4) to afford C_2-**2** (18.3 mg, 56%) as an orange solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.82 (d, *J* = 9.0 Hz, 2H), 8.71 (d, *J* = 8.6 Hz, 2H), 8.63 (t, *J* = 8.3 Hz, 4H), 8.16 – 8.06 (m, 6H), 7.94 (d, *J* = 7.9 Hz, 2H), 7.86 (d, *J* = 8.1 Hz, 2H), 7.83 (d, *J* = 3.2 Hz, 2H), 7.82 – 7.78 (m, 4H), 7.66 (t, *J* = 7.4 Hz, 2H), 7.54 (t, *J* = 7.4 Hz, 2H), 7.51 – 7.42 (m, 6H), 7.14 (t, *J* = 7.8 Hz, 2H) ppm. ^{13}C NMR (125 MHz, $CDCl_3$) δ: 132.6 (6xC), 131.89 (4xC), 131.55 (4xC), 131.15 (2xC), 130.59 (4xC), 130.44 (2xC), 130.35 (2xCH), 130.27 (4xC), 129.47 (2xCH), 129.21 (2xCH), 128.43 (2xCH), 128.24 (2xCH), 128.11 (2xCH), 127.75 (2xCH), 127.17 (2xCH), 127.00 (2xCH), 126.91 (2xCH), 126.78 (2xCH), 126.47(2xCH), 126.29 (2xCH), 125.91 (2xCH), 125.53 (2xCH), 124.98 (2xCH), 124.84 (4xC), 124.78 (2xCH), 123.91 (2xCH) ppm. MS (MALDI-TOF) for $C_{66}H_{36}$: 828.167. The spectroscopic data were in agreement with reference 2.

*1.2e. Synthesis of D_3-symmetric hexapole pentahelicene **2***

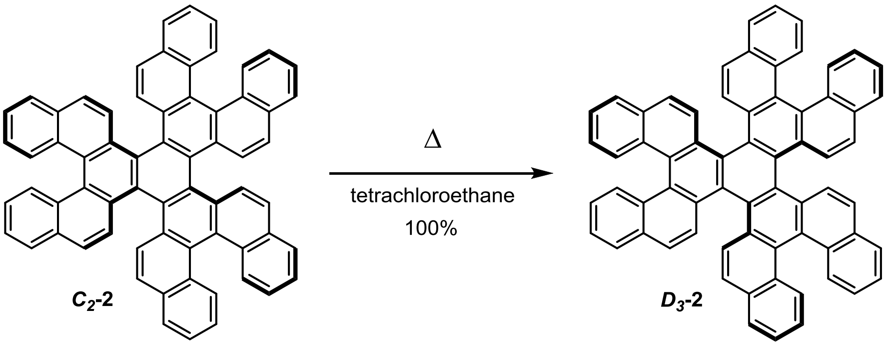

Figure S6. Synthesis of *D_3*-**2**.

A solution of *C_2*-**2** (50 mg, 0.060 mmol) in tetrachloroethane (2 mL) was stirred for 3h at 120ºC. The solvent was removed under reduced pressure to give compound *D_3*-**2** (50 mg, 100%) as a yellow solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.57 (d, *J* = 9.0 Hz, 12H), 7.91 (d, *J* = 8.0 Hz, 6H), 7.79 (d, *J* = 9.0 Hz, 6H), 7.54 (dd, *J* = 8.0, 6.8 Hz, 6H), 7.42 (dd, *J* = 8.5, 7.0 Hz, 6H) ppm. ^{13}C NMR (125 MHz, $CDCl_3$) δ: 132.36 (6xC), 131.68 (6xC), 130.93 (6xC), 129.76 (6xCH), 128.99 (6xC), 128.35 (6xCH), 127.53 (6xCH), 127.00 (6xCH), 126.5 (6xCH), 125.39 (6xCH), 124.94 (6xC) ppm. The spectroscopic data were in agreement with reference 2.

1.3 ^{1}H and ^{13}C NMR spectra

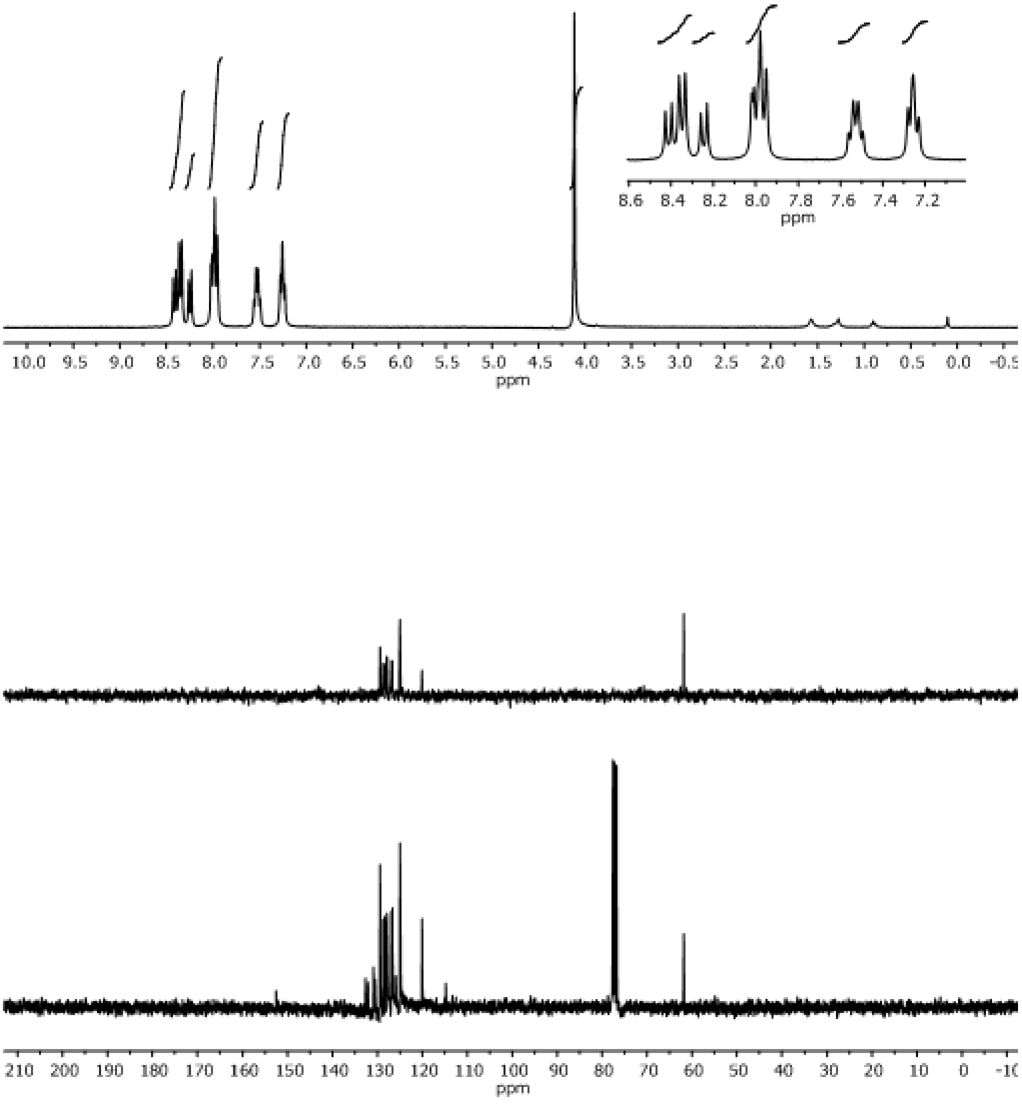

Figure S7. ^{1}H and ^{13}C spectra of compound **6**.

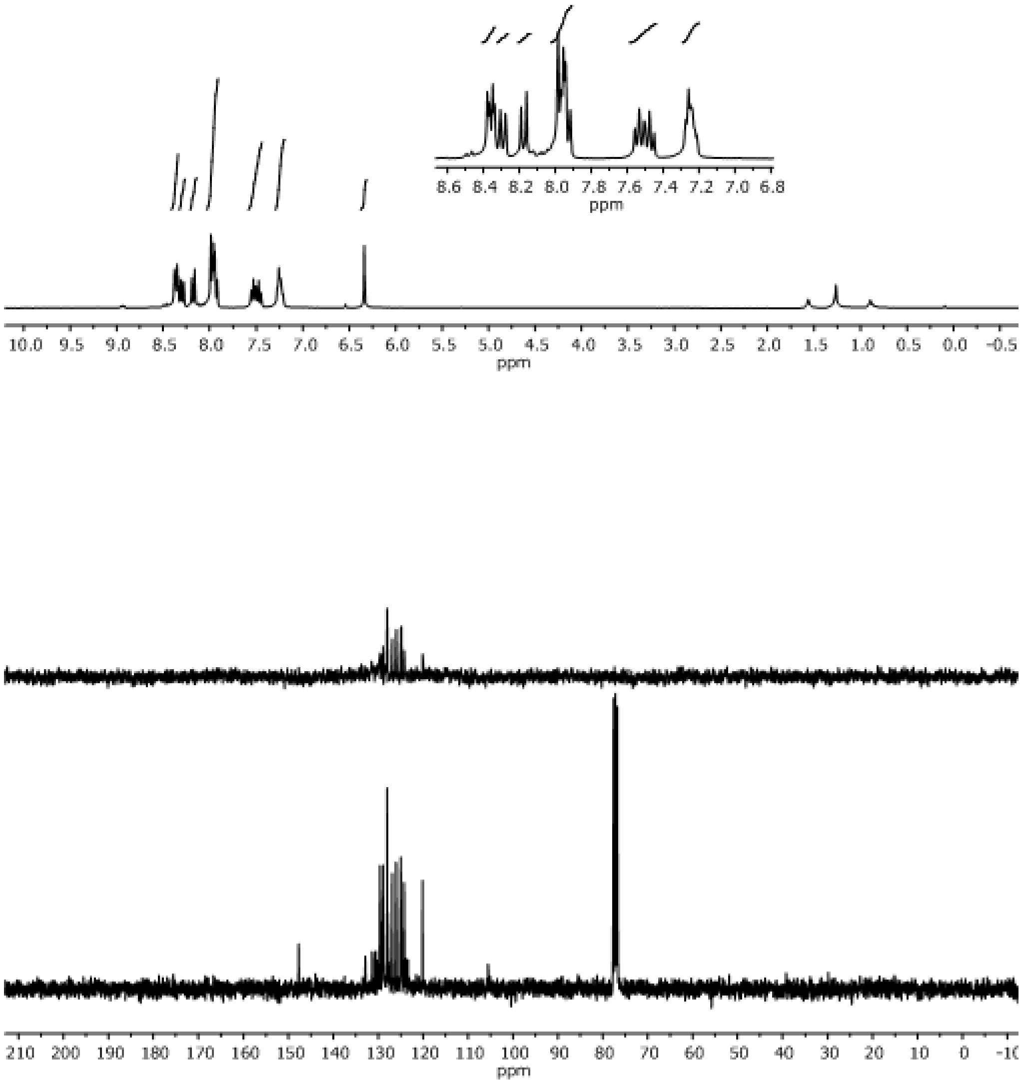

Figure S8. ^{1}H and ^{13}C spectra of compound **7**.

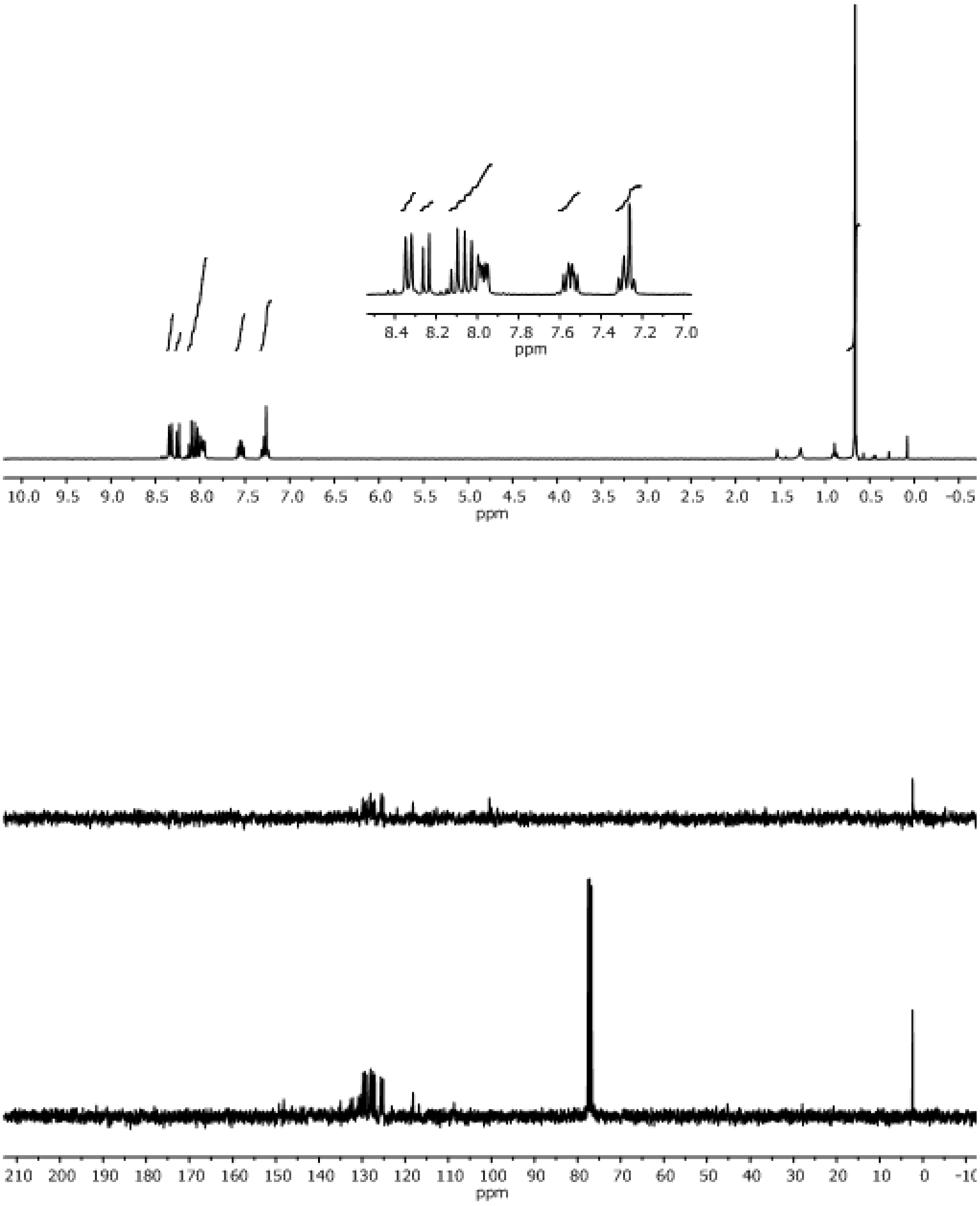

Figure S9. ^{1}H and ^{13}C spectra of compound **4**.

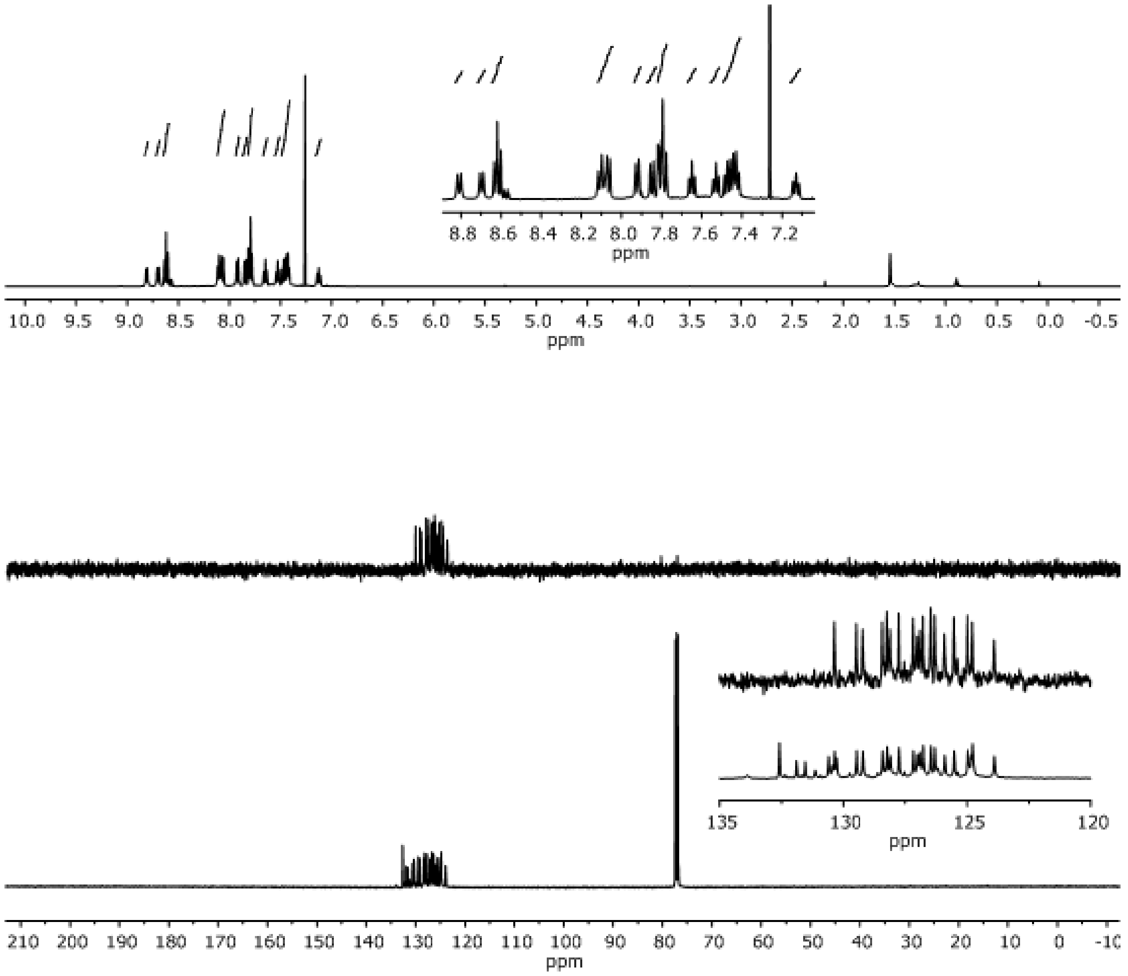

Figure S10. ^{1}H and ^{13}C spectra of compound *C_2*-**2**.

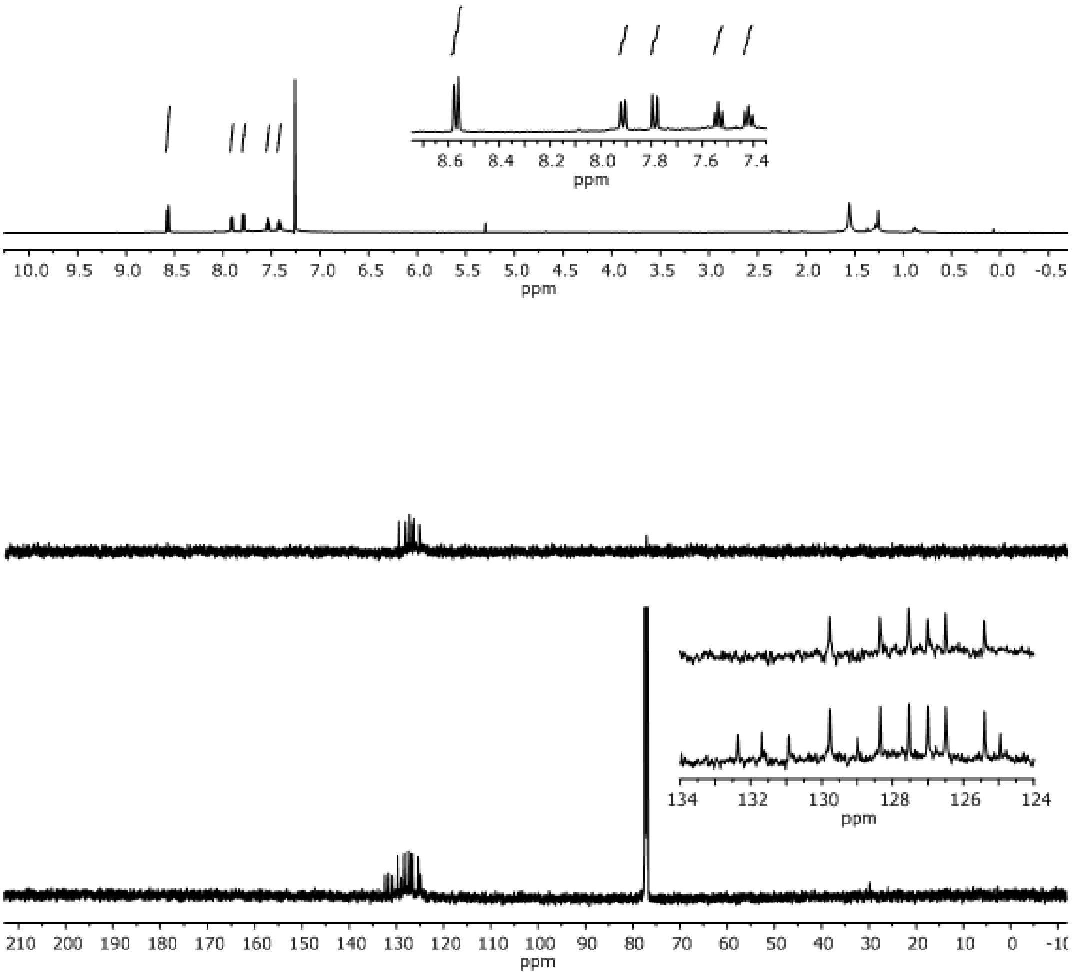

Figure S11. ^{1}H and ^{13}C spectra of compound *D_3*-**2**.

2 Experimental procedure for on-surface studies

2.1 Sample preparation, nc-AFM and STM measurements

All experiments were performed in an ultra-high-vacuum (UHV) multi-chamber system equipped with a low-temperature scanning probe microscope (Omicron GmbH) that could be operated in a scanning tunneling and non-contact atomic force microscopy (STM, nc-AFM) modes. Au(111) samples were prepared in a standard procedure including Ar^+ ion bombardment and subsequent thermal annealing at 720 K. The molecules were evaporated thermally on a sample kept at room temperature from a water-cooled Knudsen cell equipped with 3 crucibles (Kentax GmbH). The sublimation temperature of hexapole pentahelicene **2** was established at 598 K with a flux calibrated using a quartz microbalance (0.3 Hz/min). nc-AFM measurements were performed using a qPlus sensor [3] operated in the constant height frequency modulation mode with the bias voltage V set to 0 V, the Q-factor approximately 18000, and the resonance frequency 23 383 Hz. To obtain submolecular spatial resolution in nc-AFM measurements, CO molecules were picked up by the tungsten nc-AFM tip using a procedure described previously [4]. All STM/STS/nc-AFM measurements were performed at liquid helium (~4.5 K) temperature. For all STM and STS experiments electrochemically etched Pt−Ir tips were applied as probes. All STS data were collected using a lock-in amplifier (Zurich Instruments MFLI) with a central frequency of 680 Hz and amplitude of 25 mV (rms). Thermally induced cyclodehydrogenation of compound **2** was performed in a preparation chamber using a resistive heater. The samples were heated for 10 min at the temperature of 655 K with the temperature monitored by a thermocouple located at a sample stage.

3 Additional STM data

3.1 Additional STM images of molecular millwheels and nanographene ***1***

Figure S12a shows the large-scale STM image of the Au(111) surface with self-assembled millwheels comprising hexapole pentahelicene molecules **2**. Red and yellow arrows indicate millwheels of opposite handedness. Figure S12b shows presence of millwheels and disordered molecules attached to the surface step edge. It is worth to note that in previous studies on helicenes it was indicated that the molecules usually adapt a conformation with several benzene rings located close to the underlying surface and a few remaining pointing slightly upward to form a helical structure [5]. The conformation of the molecules was further inferred from the STM images with the assumption that the most upward located part of the molecule appears as the brightest and the rings interacting with the surface imaged as the faintest in constant current STM topographies [5]. While such reasonable procedure allowed to assign the orientation of helicenes on flat metallic substrates it is however important to note that in our case of hexapole pentahelicenes **2** direct extraction of the molecule configuration on the surface from STM images might be infeasible due to the molecule complexity. Indeed as indicated by Figure S12a,b even at low coverage the molecules tend to localize either in a disordered manner at the surface step edge or within highly ordered millwheels, whereas single molecules could not be recognized.

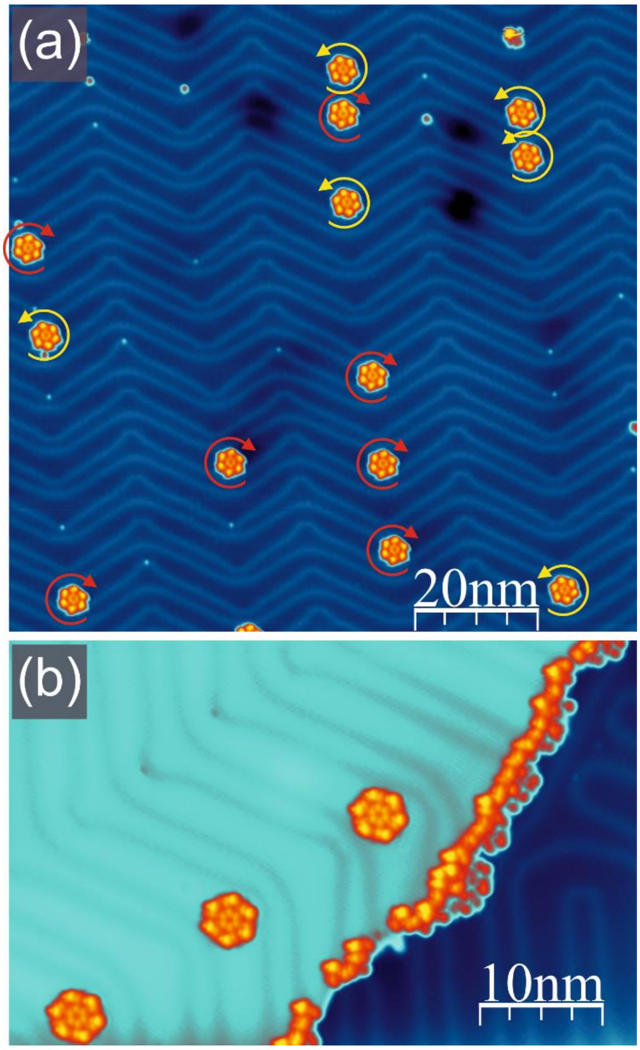

Figure S12. Filled state STM images of millwheels self-assembled from hexapole pentahelicenes ***2*** *on Au(111), (a) millwheels of opposite handedness are marked by red and yellow arrows, (b) millwheels and single molecules attached to the surface step edge; imaging conditions: tunnelling current 50 pA, bias voltage: -1.0 V.*

To recognize the structure of the millwheels formed by the self-assembly of compound **2** we have attempted molecule manipulation to disassemble the structures. Unfortunately, it occurred that the millwheels are relatively strongly fixed and we did not manage to separate them into single molecules. However, in order to shed more light on the internal structure of observed millwheels, we have applied STM imaging with a tip spontaneously decorated by the molecule. The STM image of the Au (111) surface area acquired with the unintentionally functionalized tip is shown in Figure S13a. The surface defects located within the corners of the herringbone pattern are now exhibiting a uniform appearance as a three-lobe structure. We attribute such phenomena to imaging the molecule located at the tip apex by the surface defects. The typically acquired image is shown in the inset of Figure S13a. We have created an artificial image of the millwheel by placing six aforementioned images as shown in Figure S13b (on the right), which correlates reasonably with the recorded STM image of the millwheel (Figure S13b, on the left). Therefore we postulate that each millwheel is created from six hexapole pentahelicenes **2**, as shown schematically in Figure S13c. In such regard the presence of millwheels exhibiting opposite handedness could be understood as the effect of self-assembly into homochiral structures, i.e. we have used the racemic mixture of the hexapole pentahelicenes and the results suggest that the millwheels of opposite handedness are generated from precursors of opposite chirality.

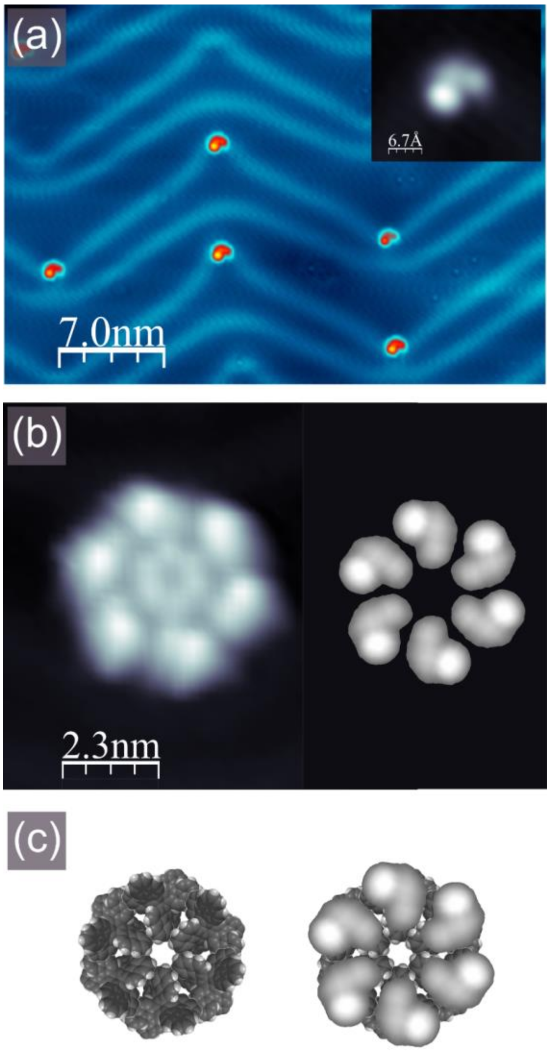

*Figure S13. (a) filled state STM image of Au(111) with a decorated tip, the contrast is assigned to the presence of a hexapole pentahelicene molecule **2** on a tip apex, inset shows magnification of the surface defect imaged by the functionalized tip; imaging conditions: tunnelling current 30 pA, bias voltage: +0.5 V; (b) filled state STM image of the millwheel (left) and the artificial image generated from six images shown in the inset of panel (a); c) proposed structural model of the millwheel (left) and the model with superimposed artificial image generated from six images as in (b); imaging conditions: tunnelling current 50 pA, bias voltage: -1.0 V.*

Figure S14 shows the nanographenes **1** generated by on-surface cyclodehydrogenation at higher molecule coverage compared to Figure 2 from the main text. Right-hand upper corner of Figure S14b shows the area of Au(111) surface, from which the molecules have been laterally removed by scanning with slightly elevated tunneling current indicating on a small molecule-substrate interaction potential corrugation.

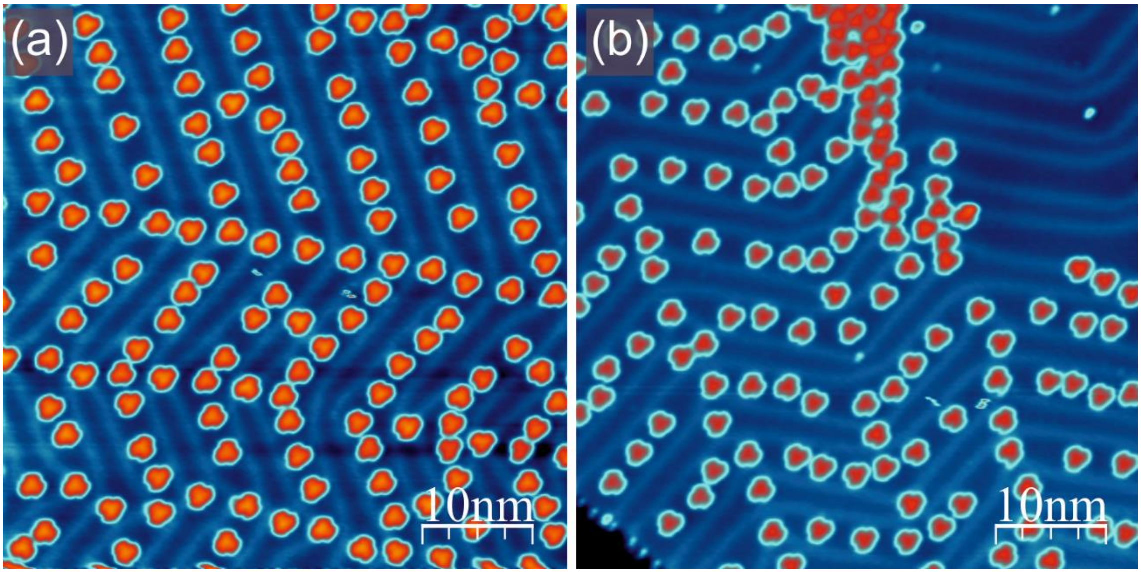

Figure S14. Filled-state STM images of nanographenes ***1*** *generated by on-surface cyclodehydrogenation of hexapole pentahelicenes* ***2*** *on Au(111), (a) nanographenes* ***1*** *arranging themselves along surface reconstruction pattern, (b) part of the surface area, from which the molecules have been removed by scanning at elevated up to 500 pA tunnelling current; imaging conditions: tunnelling current 50 pA, bias voltage: -1.0 V.*

4. Calculation details

4.1. Nanographene 1

First-principles electronic structure calculations were performed using density functional theory as implemented in SIESTA package [6, 7]. We used the van der Waals density functional by Dion *et al* [8] with the modified exchange by Klimeš, Bowler and Michaelides [9]. A double-ζ plus polarization (DZP) basis set was adopted to expand the valence-electron wave functions, where the orbital radii were defined using a 10 meV energy shift [7]. Core electrons were described using norm-conserving Troullier-Martins pseudopotentials [10]. The real space grid for integrations was defined using a 350 Ry energy cutoff [7]. We considered an electronic temperature of 300 K. The self-consistency cycle was stopped when variations on the elements of the density matrix were lower than 10^{-6} as well as 10^{-4} eV for the Hamiltonian matrix elements.

The freestanding nanographene **1** (Figure S15a) was computed within a simulation cell of dimensions 40x35x25 $Å^3$ and geometry optimizations were performed using the conjugate gradient (CG) method until all forces were lower than 5 meV/Å. For the calculations involving the metal substrate, the Au(111) surface was represented by a 3-layer thick slab within a 28.85x29.98x30 $Å^3$ simulation cell, being the third dimension perpendicular to the Au surface. We adopted for the top Au atomic layer an extended DZP basis optimized for the description of the Au(111) surface [11]. Moreover, a Hydrogen passivation was employed at the slab bottom in order to prevent spurious effects due to interaction between surface states belonging to the top and bottom surfaces of the slab [12] (see Figure S15b). A 4x4 *k*-point mesh was used to sample the bidimensional Brillouin zone. The full system represented in Figure S15b comprises 570 atoms where, except by the two bottom Au layers which were kept fixed, all other atoms were allowed to relax with the CG method until forces were lower than 10 meV/Å. In the experiment the nanographenes were found to be located at different positions, however a significant amount of the molecules was found to have one of the arms along the <1-10> direction. Therefore, we chose this configuration as our initial structure before geometry optimization. Spin polarized simulations of the gas-phase nanographene indicated no magnetization consistent with the relatively large electronic gap. Therefore, spin polarization was disregarded for the calculations of the supported molecule.

Fig. S15c shows the calculated molecular levels of a freestanding nanographene **1**, aligned with respect to the vacuum level. The DFT estimation for the energy gap between the highest occupied and lowest unoccupied molecular orbitals ("HOMO" and "LUMO", respectively) is ~1.5 eV, which corresponds to a value 40% lower than the experimentally measured. This is expected due to the well known underestimation of the HOMO'-LUMO' gap of standard (local and semilocal) DFT functionals. Note that we include a prime as a notation to distinguish the assignment of the levels from the theory. As described in the main text, the HOMO'-1 and LUMO'+1 states are doubly-degenerate, which is consistent with the three-fold symmetry of the planar molecule. The calculated band structure of the full system comprising the nanographene on Au(111) slab is presented in Fig. S15d (left) for a path in the supercell's Brillouin zone from Γ to X (X here is the zone-boundary along the supercell's y-direction and the energies are referred to the vacuum level). Note that, since the simulation cell is several times larger than the surface unit cell, the bands associated with the gold substrate appear folded several times with respect to its more standard appearance. The goal here, however, is to verify whether the nanographene energy levels can be distinguished among those bands. This can be done by projecting on the Carbon atomic orbitals, in particular on the p_z as shown in Fig. S15d (center). Localized levels can be easily distinguished below and above the Fermi energy, with an energy difference equivalent to the HOMO'-LUMO' gap found for the gas-phase nanographene calculation. The interaction with the substrate is sufficiently low that the splitting between HOMO' and HOMO'-1 and between the LUMO' and LUMO'+1 does not seem to be substantially modified with respect to the free molecule, nor the degeneracies lifted. This justifies the use of the free molecule to simulate the STS images in the main text. It is also interesting to note that there is a non-negligible hybridization with the Au states, particularly clear for the unoccupied states, that will contribute to broaden the molecular orbitals as indicated by the experimental observation.

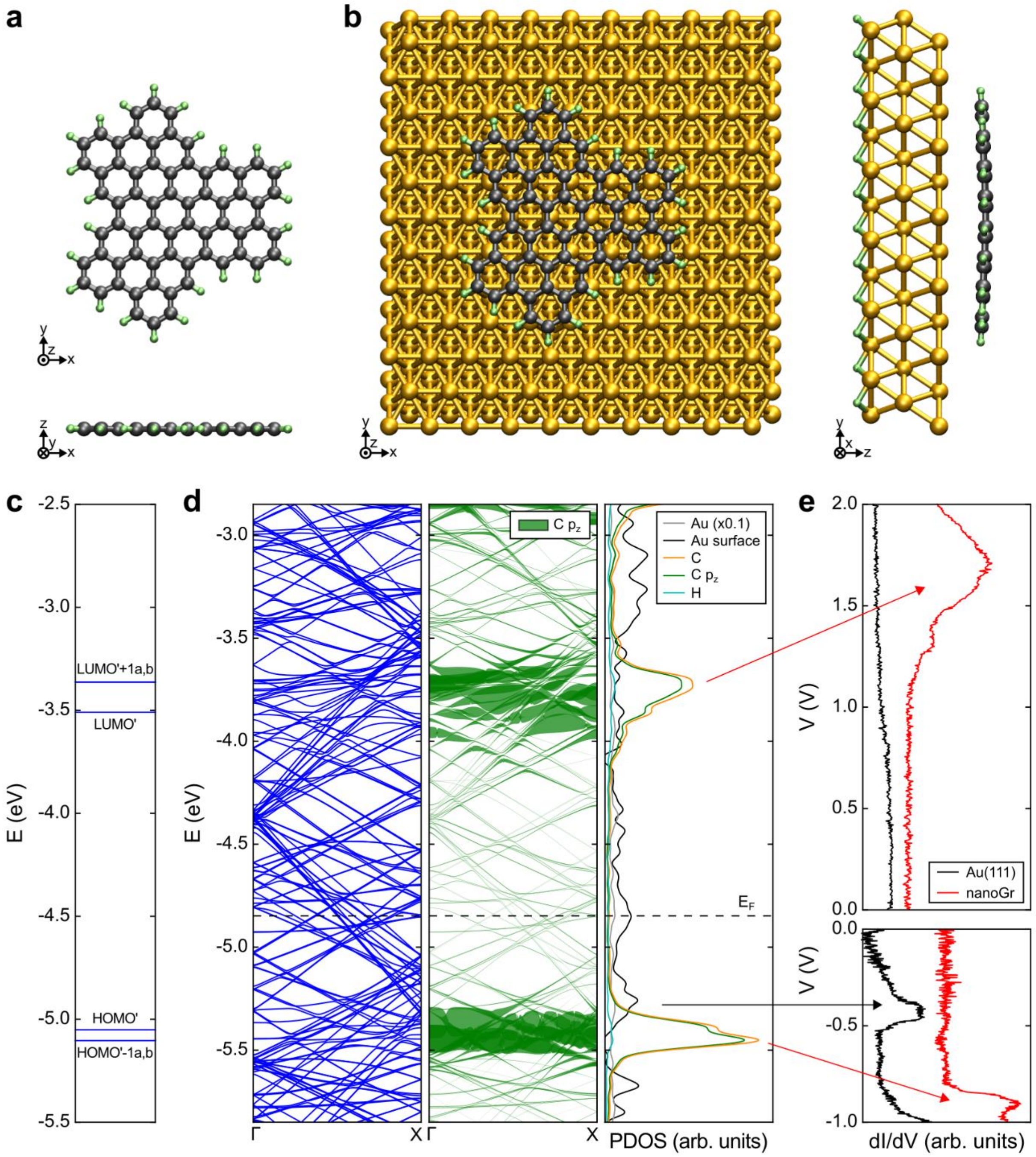

Figure S15. (a) Top and side views of a relaxed freestanding nanographene ***1****. (b) Top and side views of a nanographene* ***1*** *on Au(111) slab after geometry optimization. (c) Electronic structure of the isolated nanographene molecule, where the highest occupied and lowest unoccupied molecular orbitals ("HOMO" and "LUMO") are indicated. (d) From left to right: band structure of the full system shown in b; band structure projected on Carbon p_z orbitals (the width of the green shaded area indicates the amount of C p_z character of each state); projected density of states (PDOS) on all Au atoms (gray), surface Au atoms far from the nanographene (black), Carbon atoms (orange), Carbon p_z orbitals (green) and Hydrogen atoms (cyan), with a Lorentzian broadening of 35 meV. (e) Scanning tunneling spectroscopy (STS) recorded on top of a nanographene (red) and on the Au(111) surface (black). The energies in the electronic structure calculations are given with respect to the vacuum energy.*

The densities of states (DOS) projected on the different atomic species or on parts of the system are shown in the right side of Figure S4.1d. The DOS projected on the Carbon atoms exhibit two clear resonances. The first at ~-5.45 eV (600 meV below the Fermi level) is assigned to the HOMO', HOMO'-1a and HOMO'-1b levels, while the second at ~-3.75 eV (~1.1 eV above the Fermi level) corresponds to the LUMO', LUMO'+1a and LUMO'+1b levels. Interestingly this alignment of the molecular levels with respect to the substrate is similar to that observed experimentally, in which the

HOMO’ is substantially closer to the Fermi level than the LUMO’. The level alignment is a complex issue that depends critically, among other factors, on the size of the molecular gap (and how it is modified upon adsorption due to polarization effects) and the ability of the computational method to properly describe the modifications of the surface dipole upon molecular adsorption. Here we used an extended basis set for the surface Au atoms that guarantees a proper description of the surface properties. Among them, in spite of the very thin slab utilized, we have a reasonable description of the surface state. This is possible thanks to the Hydrogen passivation of the bottom surface of the slab. The Au(111) surface state appears as a broad peak at ~-5.3 eV (*i.e.*, ~450 meV below the Fermi level) in the DOS projected on the gold surface. To obtain this signal and isolate the surface character, we took the difference between the PDOS projected on the Au atoms in the surface layer and far from the nanographene and the PDOS projected on the Au atoms in middle slab layer. The position of the surface state in the calculation at this energy, in excellent agreement with experiment, was further confirmed by a calculation for the clean Au(111) surface using a slab identical to that used for the nanographene adsorption.

Figure S16 presents the induced charge after adsorption, *i.e.*, $\rho_{\mathrm{ind}}(\mathbf{r}) = \rho_{\mathrm{Tot}}(\mathbf{r}) - \rho_{\mathrm{Au}}(\mathbf{r}) - \rho_{\mathrm{nnGr}}(\mathbf{r})$, where $\rho_{\mathrm{Tot}}(\mathbf{r})$ is the charge density of the full system (nanographene plus Au slab), $\rho_{\mathrm{Au}}(\mathbf{r})$ is the charge density of the isolated slab and $\rho_{\mathrm{nnGr}}(\mathbf{r})$ for the isolated nanographene. The distribution of the induced charge determines the modification of the surface dipole layer and, thus, it is crucial to understand the energy level alignment. To help on visualization, the same calculated induced charges are presented in Fig. S16 separated on the region around the nanographene (top) and near the surface (bottom). As expected due to the so-called Pauli pushback or cushion effect of the surface charges [13], electrons tend to accumulate near the surface, while at height of the molecular plane we mostly find a depletion of electrons. Interestingly, however, the electron accumulation takes place along the molecular contour and particularly in the “cove” regions between the molecule arms. These regions act as strong electron attractors as seen by the three big red blobs in the Au surface. In the surface region directly underneath the molecular plane we find mostly a depletion of electron charge. We also performed a Bader population analysis [14, 15] to complete our understanding of the charge redistribution. Although the band structure and PDOS in Figure S15 (with the HOMO’ and LUMO’ levels far from the Fermi level) clearly indicate that there cannot be a substantial charge transfer between the nanographene and the surface, the population analysis gave a depopulation of ~0.25 electrons of the nanographene. We understand this intriguing result as an additional signature of the charge rearrangement observed in Figure S16, where the whole area covered by the molecular plane becomes electron deficient (with the compensating charge located around the molecule, mostly in the coves). A reminiscent of this situation was observed in recent calculations of armchair graphene ribbons of different widths on Au(111), and it was related to the p-like alignment of the ribbons [16]. Our nanographenes also present a p-like alignment with the HOMO’ level substantially closer to the gold Fermi level, so it is tempting to postulate a similar connection in the present case.

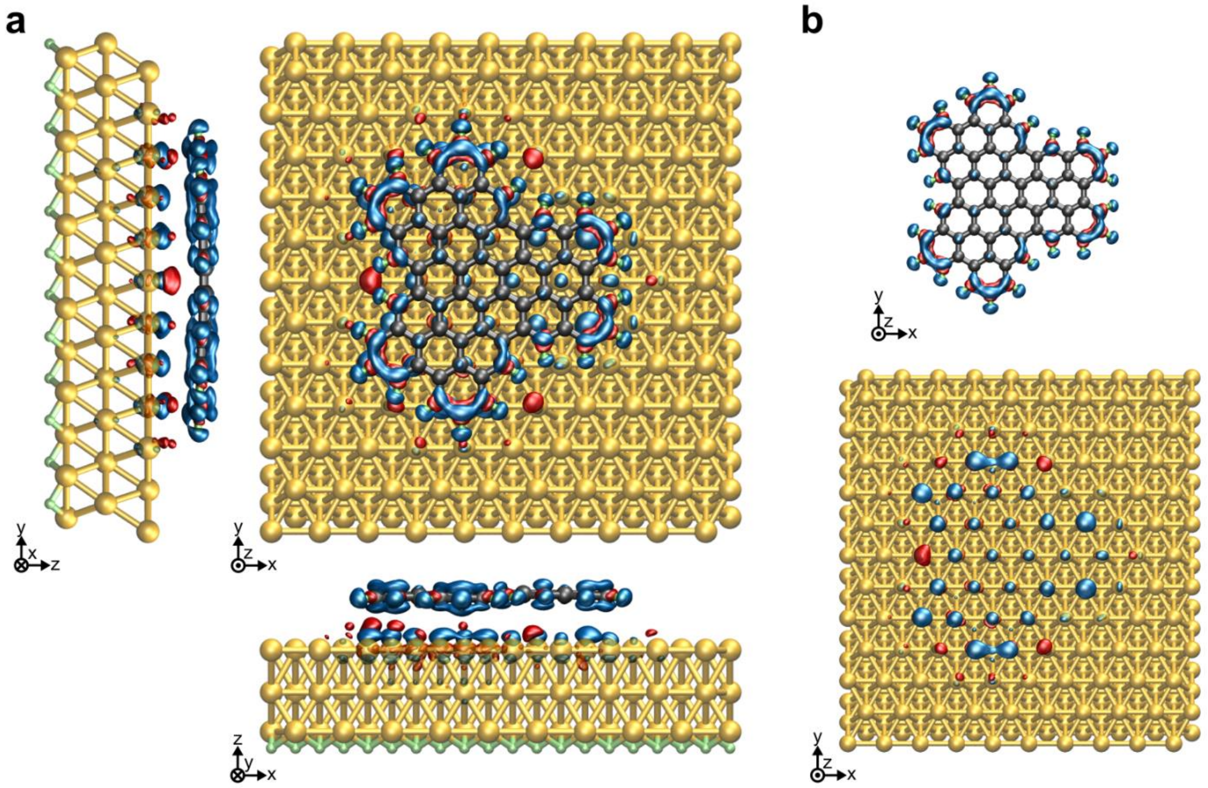

*Figure S16 (**a**) Induced charge upon adsorption of a nanographene **1** on Au(111) surface (top and side views), where red (blue) indicates electron accumulation (deficiency). (**b**) Same as **a**, but decomposed into the regions around the nanographene (top) and close to the Au surface (bottom). All isosurfaces correspond to $5x10^{-4}$ electron/$Bohr^3$.*

4.2. Simulations of dI/dV images

To compute the STM images we followed the surface integration technique of Paz and Soler[17]. In addition, we used the Tersoff–Hamann approximation [18] assuming a proportionality factor reaching 1 $nA \cdot Å^{-3}$ for the ratio between the tunneling current and the local density of states. Finally, in order to mimic the effect of spatial uncertainty in the measurement which reduces the resolution of the STM images, we have convoluted our currents with a Gaussian kernel $K(r,r_0) = (pi \cdot \sigma^2)^{3/2} \cdot \exp(|r-r_0|^2/(2s^2))$ with $\sigma = 0.5$ Å.

5. References

[1] S. Goretta, C. Tasciotti, S. Mathieu, M. Smet, W. Maes, Y. M. Chabre, W. Dahaen, R. Giasson, J. Raimundo, C. R. Henry, C. Barth, M. Gingras, *Org. Lett.*, 2009, **17**, 3846

[2] T. Hosokawa, Y. Takahashi, T. Matsushima, S. Watanabe, S. Kikkawa, I. Azumaya, A. Tsurasaki, A. K. Kamikawa, *J. Am. Chem. Soc.* 2017, **139**, 18512

[3] F. J. Giessibl, *Appl. Phys. Lett.* 2000, **76**, 1470.

[4] L. Gross, F. Mohn, N. Moll, P. Liljeroth, G. Meyer, *Science* 2009, **325**, 1110.

[5] K.-H. Ernst, *Acc. Chem. Res.* 2015, **49**, 1182.

[6] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García, J. M. Soler, *Phys. Stat. Solidi (b)* 1999, **215**, 809.

[7] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, *J. of Phys.: Cond. Matt.* 2002, **14**, 2745.

[8] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, B. I. Lundqvist, *Phys. Rev. Lett.* 2004, **92**, 246401.

[9] J. Klimeš, D. R. Bowler, A. Michaelides, *J. Phys.: Condens. Matter* 2010, **22**, 022201.

[10] N. Troullier, J. L. Martins, *Phys. Rev. B* 1991, **43**, 1993.

[11] S. García-Gil, A. García, N. Lorente, P. Ordejón, *Phys. Rev. B* 2009, **79**, 075441.

[12] N. Gonzalez-Lakunza, I. Fernández-Torrente, K. J. Franke, N. Lorente, A. Arnau, J. I. Pascual, *Phys. Rev. Lett.* 2008, **100**, 156805.

[13] G. Witte, S. Lukas, P. S. Bagus, C. Wöll, *Appl. Phys. Lett.* 2005, **87**, 263502.

[14] R. F. W. Bader, "Atoms in molecules: A quantum theory," Oxford University Press, New York, 1990.

[15] G. Henkelman, A. Arnaldsson, H. Jónsson, *Comput. Mater. Sci.* 2006, **36**, 254.

[16] N. Merino-Díez, A. García-Lekue, E- Carbonell-Sanromà, J. Li, M. Corso, L. Colazzo, F. Sedona, D. Sánchez-Portal, J. I. Pascual, D. G. de Oteyza, *ACS Nano* 2017, **11**, 11661.

[17] O. Paz, J. M. Soler, *Phys. Status Solidi (b)* 2006, **243**, 1080.

[18] J. Tersoff, D. R. Hamann, *Phys. Rev. Lett.* 1983, **50**, 1998.