On-surface synthesis of edge-extended zigzag graphene nanoribbons

Amogh Kinikar° [a], Xiushang Xu° [b,c], Marco Di Giovannantonio [a,f], Oliver Gröning [a], Kristjan Eimre [a,g], Carlo A. Pignedoli [a], Klaus Müllen [c,d], Akimitsu Narita [b,c], Pascal Ruffieux [a], and Roman Fasel [a,e]

[a] Empa, Swiss Federal Laboratories for Materials Science and Technology, nanotech@surfaces Laboratory, 8600 Dübendorf, Switzerland
[b] Okinawa Institute of Science and Technology Graduate University, Organic and Carbon Nanomaterials Unit, 1919-1 Tancha, Onnason, Kunigamigun, Okinawa 904-0495, Japan
[c] Max Planck Institute for Polymer Research, 55128 Mainz, Germany.
[d] Johannes Gutenberg University Mainz, Institute of Physical Chemistry, Duesbergweg 10-14, 55128 Mainz, Germany
[e] University of Bern, Department of Chemistry, Biochemistry and Pharmaceutical Sciences, Freiestrasse 3, 3012 Bern, Switzerland

° These authors contributed equally to this work.

Present address:
[f]Institute of Structure of Matter – CNR (ISM-CNR), via Fosso del Cavaliere 100, 00133 Roma
[g]EPFL, École polytechnique fédérale de Lausanne, NCCR MARVEL, 1015 Lausanne, Switzerland

Contents

Supporting Information

Supporting Information

S1: Nomenclature for edge-extended ZGNRs

3-ZGNR-E(Bisanthene, 7)

Nomenclature:

Alternating fused extension

3-ZGNR-E(Bisanthene, 7)

ZGNR backbone

Periodicity

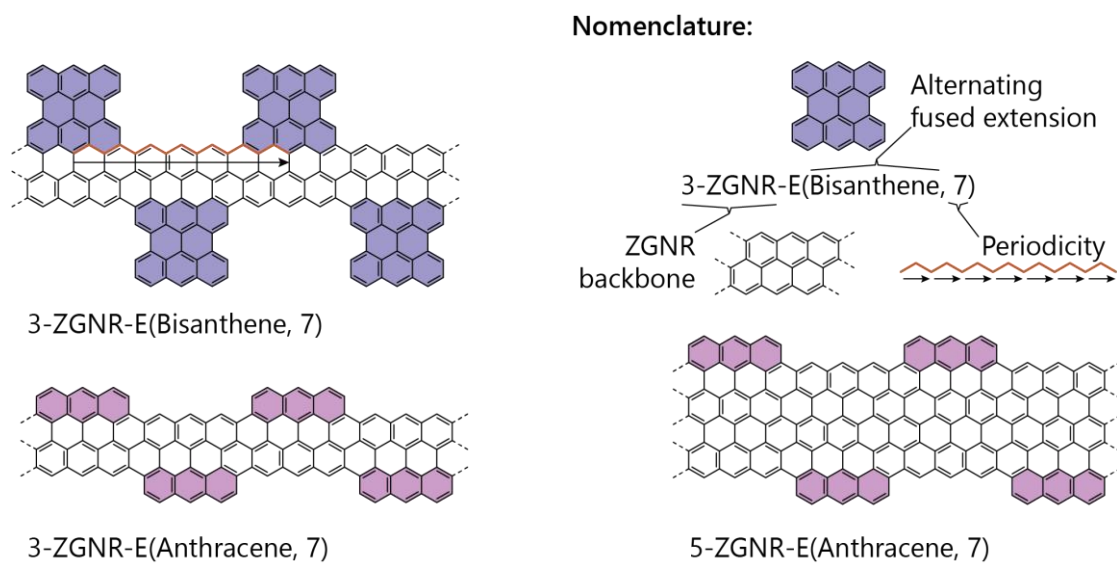3-ZGNR-E(Anthracene, 7)

5-ZGNR-E(Anthracene, 7)

Figure S1: Nomenclature for edge-extended ZGNRs. The ZGNR backbone to which the extensions are fused is identified as the N zigzag rows wide ZGNR (N-ZGNR). The edge extensions that are fused to alternating sides of the N-ZGNR backbone are identified by their chemical name. The periodicity of the fused extensions along one edge is expressed in terms of the axial unit vector of the N-ZGNR. The nomenclature is illustrated with examples.

S2: Synthesis and characterizations of monomer 6

General Methods

All reactions working with air- or moisture-sensitive compounds were carried out under argon atmosphere using standard Schlenk line techniques. Unless otherwise noted, all starting materials and other chemicals were purchased from commercial sources and used without further purification. Thin-layer chromatography (TLC) was done on silica gel coated aluminum sheets with F254 indicator and column chromatography separation was performed with silica gel (particle size 0.063-0.200 mm). Nuclear Magnetic Resonance (NMR) spectra were recorded using Bruker DPX 300 MHz, Avance Neo 400/500 MHz NMR spectrometers. Chemical shifts (δ) were expressed in ppm relative to the residual solvents (CD_2Cl_2, ^1H: 5.32 ppm, ^{13}C: 54.00 ppm; $CDCl_3$, ^1H: 7.26 ppm, ^{13}C: 77.00). Coupling constants (*J*) were recorded in Hertz. Abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. High-resolution mass spectra (HRMS) were recorded on Electrospray ionization (ESI) Thermo Orbitrap Thermo Scientific LTQ-Or-bitrap or on Bruker timsTOF atmospheric pressure chemical ionization (APCI) mass spectrometric (MS) detection.

Synthesis Details

5-bromo-2-(4-bromobenzyl)benzaldehyde (**3**)
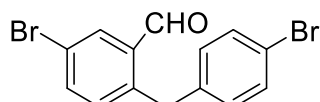
A 50-mL round flask was charged with 5-bromo-2-methylbenzaldehyde (1.20 g, 6.01 mmol), 1-bromo-4-iodoben-zene (4.30 g, 15.0 mmol), silver trifluoroacetate (3.34 g, 15.0 mmol), glycine (360 mg, 4.80 mmol), palladium (II) acetate (136 mg, 0.601 mmol), and CH_3COOH/H_2O (18 mL/2 mL). The reaction mixture was stirred at room temperature for 10 min and then at 100 °C for 36 h. The resulting mixture was cooled to room temperature, diluted with EtOAc (50 mL), and filtered through a silica gel plug. Then, organic phases were combined, washed with brine and saturated aqueous solution of $NaHCO_3$, dried over $MgSO_4$, and evaporated. The residue was purified by silica gel column chromatography (eluent: *n*-hexane: ethyl acetate = 20:1) to give the title compound (1.05 g, 52% yield) as light-yellow liquid. ^1H NMR (400 MHz, $CDCl_3$) δ 10.05 (s, 1H), 7.90 (d, *J* = 2.3 Hz, 1H), 7.58 (dd, *J* = 8.1, 2.2 Hz, 1H), 7.44 – 7.21 (m, 2H), 7.06 (d, *J* = 8.2 Hz, 1H), 6.92 (d, *J* = 8.4 Hz, 2H), 4.27 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 190.82, 141.01, 138.60, 136.80, 135.27, 133.36, 131.78, 130.46, 121.19, 120.46, 37.05. HRMS (ESI, Positive): *m/z* Calcd. For $(C_{14}H_9Br_2)^+$: 334.9086, [M-OH]$^+$, found: 334.9059.
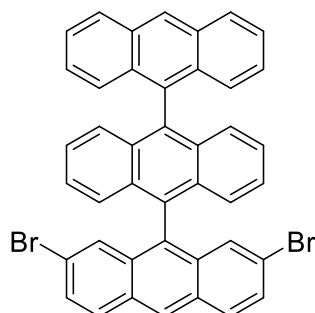
10-bromo-9,9'-bianthracene (**5**)
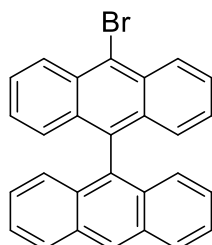
A round-bottom flask was charged with 9,9'-bianthracene (63 mg, 0.18 mmol), *N*-bromosucinimide (NBS) (32 mg, 0.18 mmol), and THF (8.0 mL). After stirring at room temperature for overnight, the reaction was quenched with acetone (10 mL), and evaporated. The residue was purified by silica gel column chromatography (eluent: *n*-hexane: CH_2Cl_2 = 15:1) to give the title compound as yellow solid (31 mg, 40% yield). ^1H NMR (300 MHz, $CDCl_3$) δ 8.81 – 8.65 (m, 3H), 8.18 (d, *J* = 8.6 Hz, 2H), 7.62 – 7.57 (m, 2H), 7.47 (t, *J* = 7.5 Hz, 2H), 7.26 – 6.99 (m, 8H). ^{13}C NMR (75

MHz, CDCl$_3$) δ 133.96, 132.38, 132.33, 131.55, 131.49, 130.48, 128.60, 128.07, 127.58, 127.34, 127.19, 126.65, 126.12, 126.07, 125.40, 123.57. HRMS (ESI, Positive): m/z Calcd. For (C$_{28}$H$_{18}$Br)$^+$: 433.0567, [M+H]$^+$, found:433.0586.

2,7-dibromo-9,9':10',9''-teranthracene (**6**)

To the solution of 10-bromo-9,9'-bianthracene (**5**) (300 mg, 0.694 mmol) in anhydrous diethyl ether (15 mL) at 0 °C under argon atmosphere, n-butyllithium (1.6 M in hexane, 0.5 mL, 0.8 mmol) was added dropwise. After stirring for 30 min at 0 °C, 5-bromo-2-(4-bromobenzyl)benzaldehyde (244 mg, 0.694 mmol) in dry diethyl ether (10 mL) was added. The resulting mixture was stirred at room temperature overnight, and then quenched with methanol (1.0 mL). After stirring for 10 min, the mixture was extracted with CH$_2$Cl$_2$ (20 mL) for three times. The organic phases were combined, washed with brine, dried over MgSO$_4$, and evaporated. The resulting crude product was dissolved in anhydrous CH$_2$Cl$_2$ (10 mL) and BF$_3$·OEt$_2$ (0.2 mL) was added at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for 2 h, and then quenched with the saturated aqueous solution of NaHCO$_3$ (10 mL). The organic phase was separated, dried over MgSO$_4$, and evaporated. The cyclization product was dissolved in anhydrous CH$_2$Cl$_2$ (10 mL) and 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) (157 mg, 0.694 mmol) was subsequently added to the solution. The resulting mixture was stirred at room temperature for 2 h, and then was poured into water (20 mL), followed by extraction with CH$_2$Cl$_2$ (20 mL) for three times. Then, organic phases were combined, washed with brine, dried over MgSO$_4$, and evaporated. The residue was purified by silica gel column chromatography (eluent: n-hexane: CH$_2$Cl$_2$ = 10:1) to give the title compound (310 mg, 65% yield) as white solid. ^{1}H NMR (400 MHz, CD$_2$Cl$_2$) δ 8.84 (s, 1H), 8.80 (s, 1H), 8.29 (d, J = 8.6 Hz, 2H), 8.16 (d, J = 9.0 Hz, 2H), 7.65 (d, J = 9.0 Hz, 2H), 7.59 – 7.57 (m, 2H), 7.38 – 7.34 (m, 4H), 7.35 (d, J = 3.4 Hz, 2H), 7.29 – 7.26 (m, 6H). ^{13}C NMR (126 MHz, CD$_2$Cl$_2$) δ 134.78, 133.07, 132.87, 132.06, 131.93, 131.71, 131.68, 131.55, 131.47, 130.68, 130.09, 129.50, 128.74, 128.34, 128.19, 127.57, 127.26, 126.71, 126.63, 126.22, 126.15, 125.89, 125.49, 121.41. HRMS (APCI, Positive): m/z Calcd. For (C$_{42}$H$_{25}$Br$_2$)$^+$: 687.0323, [M+H]$^+$, found: 687.0318.

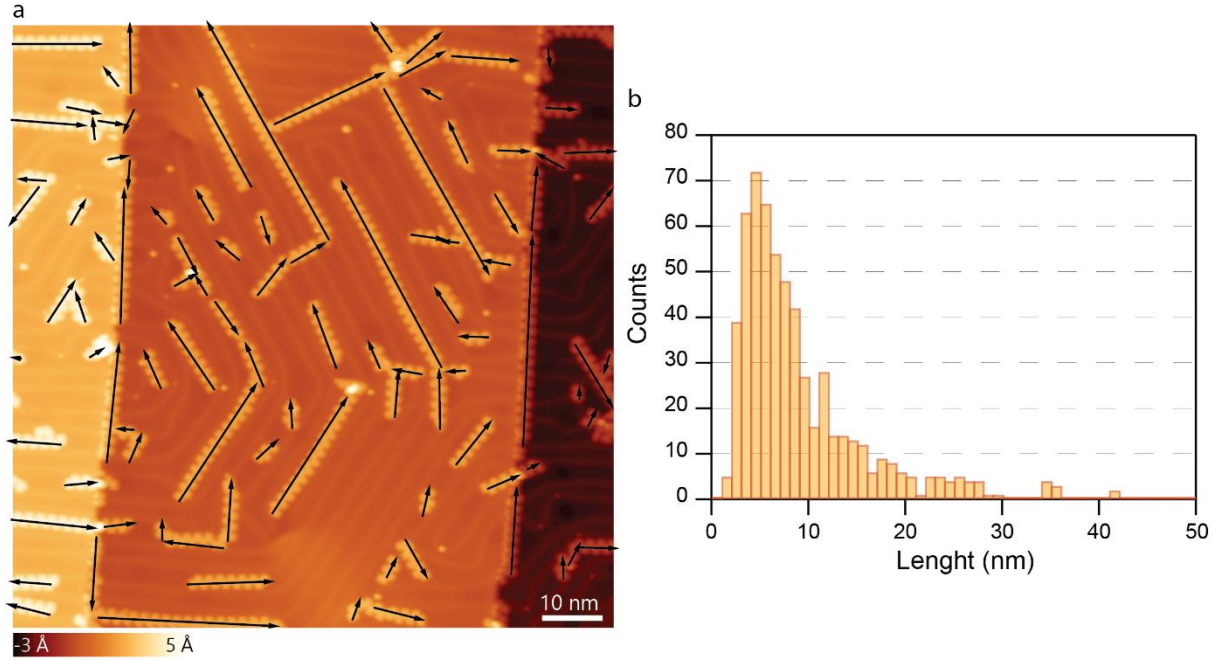## S3: Length distribution of the GNRs

Figure S2: The length of 585 GNRs in six different regions was evaluated. The histogram of GNR lengths is shown in **b**, and a representative image, with the GNRs indicated with black arrows, is shown in **a**. Despite no optimization of GNR length has been attempted, the GNRs are found to grow reasonably well, with an average length of 9 nm. As can be seen in **a**, these results are obtained with samples with a low coverage of GNRs, increasing their coverage presents a straightforward way to increase the average length of the GNRs. (Tunneling parameters for **a**: V = -1 V and I = 50 pA, bin-width for the histogram is 1 nm.)
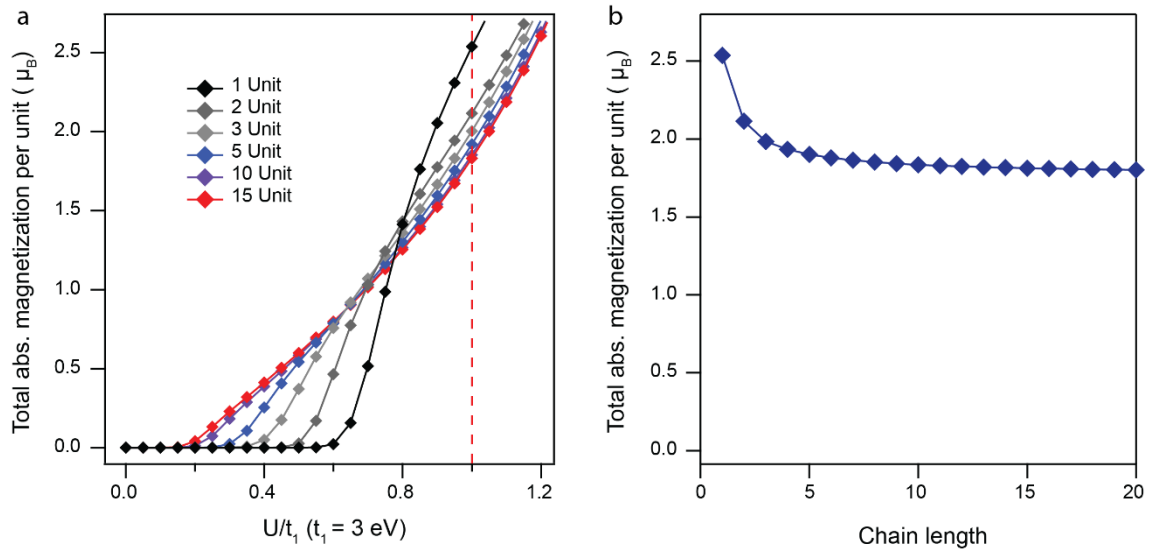
S4: Mean-field Hubbard calculations

Figure S3: **Development of magnetization as a function of on-site Coulomb repulsion and the length of the 3-ZGNR-E(BA, 7) segment. a**, The absolute magnetization is plotted as a function of the on-site repulsion U (in units of the nearest neighbor hopping t_1) for the GNRs of the indicated lengths in terms of precursor units. We find that in the U = t_1 regime (indicated by the dashed red line) GNRs of all lengths are magnetized. **b**, The total absolute magnetization in the specific case of U = t_1 is plotted as a function of chain length, showing that the spin polarization is stable and beyond a size has no dependence on the chain length.
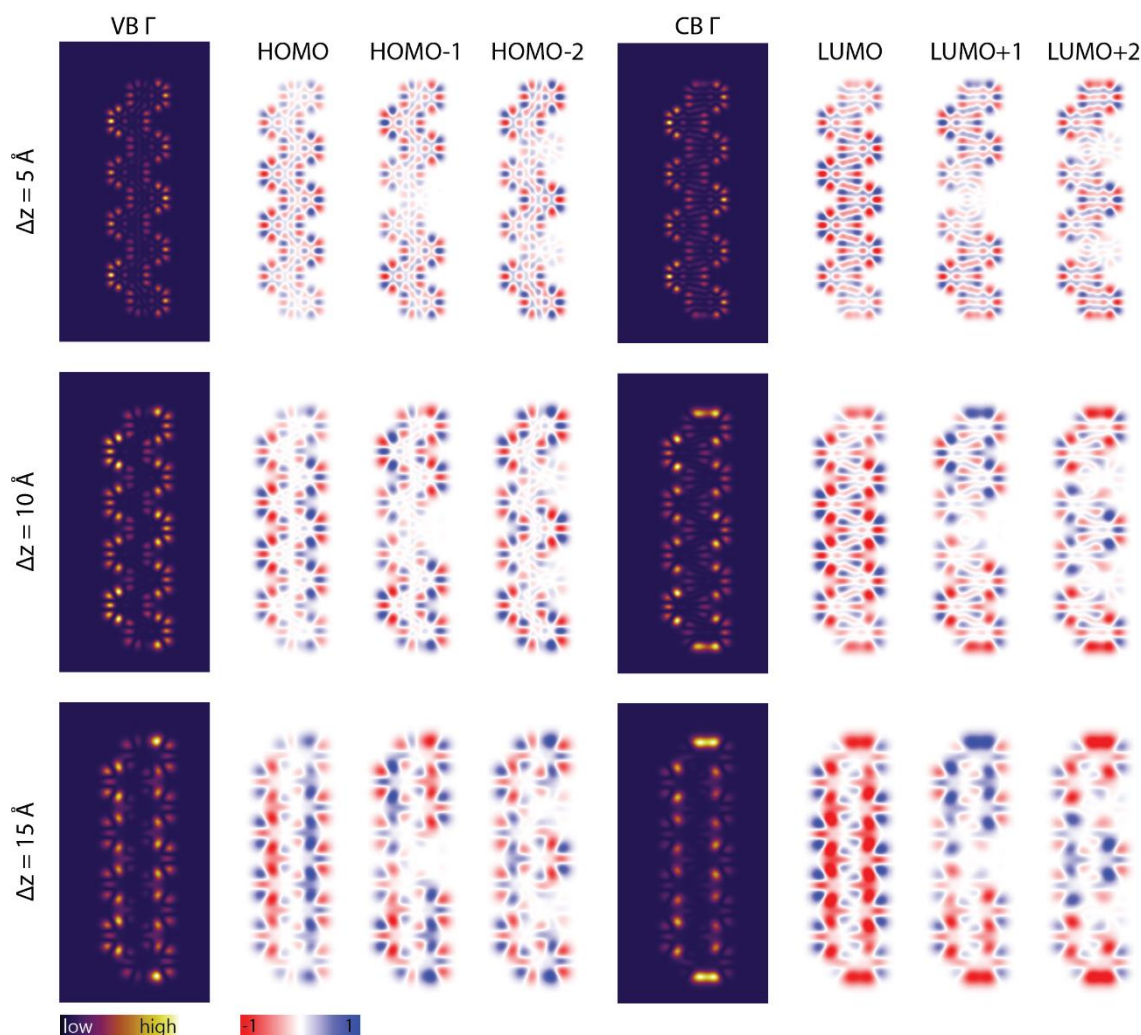
S5: Height-dependence of the LDOS

Figure S4: Tight-binding simulations of the dI/dV maps for a 9-unit segment of 3-ZGNR-E(BA, 7) at three tip-heights (Δz = 5 Å, 10 Å, and 15 Å) at the energies of VB and CB band onsets (i.e. at the Γ point) are shown. To mimic hybridization with the substrate, the dI/dV maps were simulated by including HOMO, HOMO-1 and HOMO-2 for the onset of VB and LUMO, LUMO+1, LUMO+2 for the onset of CB. The corresponding wavefunctions are shown alongside, with red and blue indicating opposite phase.

The simulated dI/dV maps for the CB onset have a clear intensity maximum at the ends of the GNR segment when the tip is far away from the molecule. A similar maximum is absent in the simulations of the VB edge. However, the bright lobes on the ends of the GNR are not due to a specific end-localized state (such as the end-states in topologically non-trivial GNRs). To understand the origin of this apparent state, let us first look at the wavefunction of the HOMO at Δz = 5 Å above the plane of the GNR. It exhibits rapid phase oscillations along the length of the GNR as well as along the width. As the tip is retracted, the rapid phase oscillations at the atomic level lead to more and more pronounced destructive interference of the wavefunction, resulting in weaker and weaker local density of states (LDOS). Consequently, the wavefunction features become less prominent compared to the development of particle-in-a-box states, as seen in the wavefunctions of HOMO, HOMO-1 and HOMO-2 at Δz = 15 Å. However, for the LUMO orbitals, the wavefunctions oscillate rapidly only along the length of the GNR, with the sign of the wavefunction unchanging across the width. The rapid oscillation along the length leads to similar height dependence in the bulk of the GNR as for the HOMO level. At the ends of the GNR, however, the LUMO wavefunctions do not undergo destructive interference, because the outermost wavefunction maximum has no minimum counterpart. This results in much higher intensity of the dI/dV maps and STS spectra at the ends of the GNR for realistic probing distances.

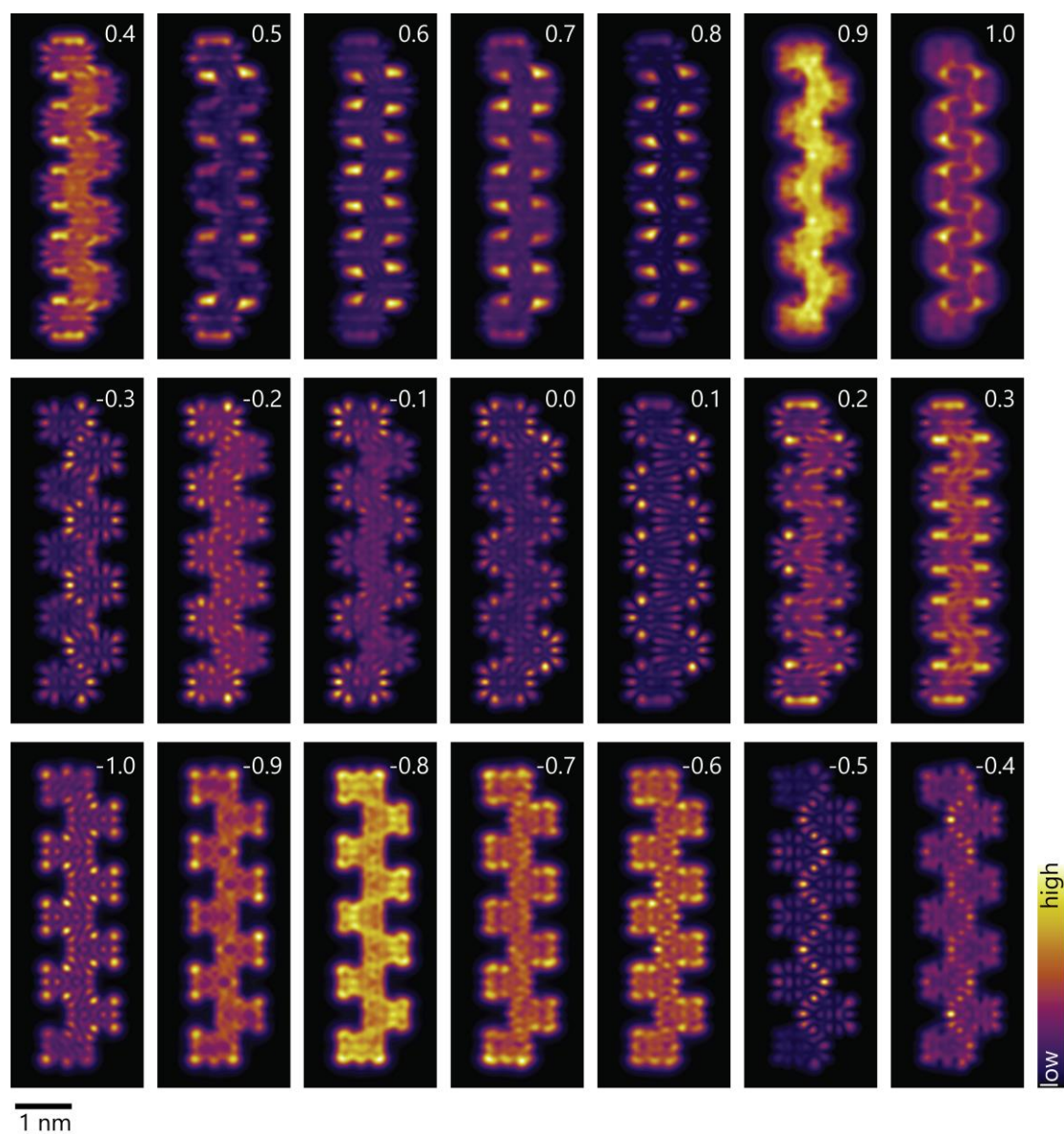## S6: Additional simulated dI/dV maps for a 9-unit segment of 3-ZGNR-E(Bisanthene, 7)

Figure S5: Extended dI/dV simulations of the system in Fig. 4 of the main text are provided for reference. The bias at which the simulation was performed is indicated in eV inside each dI/dV map. dI/dV maps are simulated at a constant-height of 4 Å above the GNR. Each map was normalized individually.
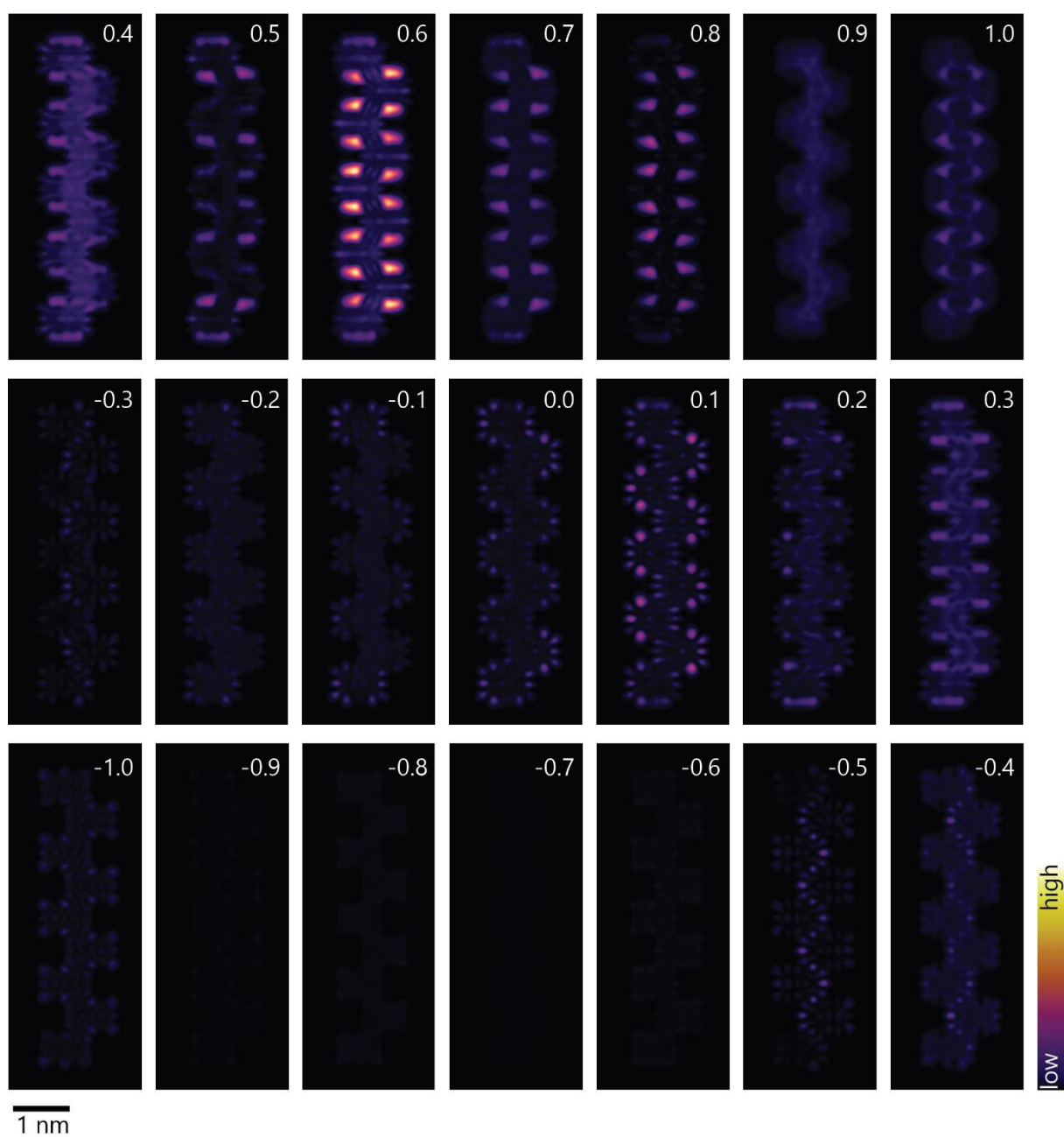
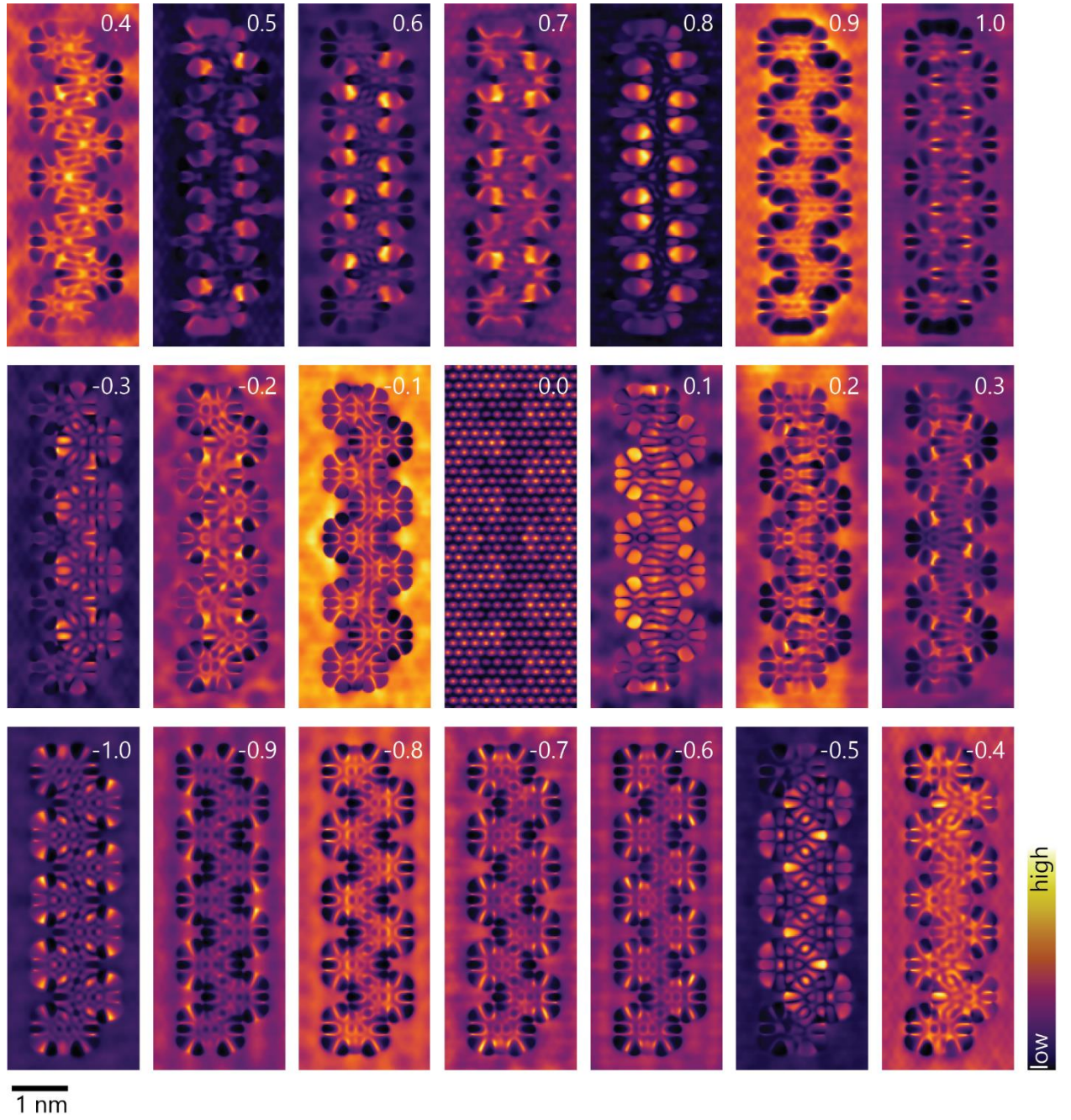**Figure S6.** Same data set as Fig. S5 shown after a global normalization of the intensity.

Figure S7. Simulated constant-current dI/dV maps. Constant-current dI/dV maps inherently mix integrated LDOS and the LDOS at the applied-bias. At constant-current the tip-sample separation increases in regions with high LDOS between the Fermi level and applied bias. Consequently, the measured LDOS signal is smaller leading to an apparent contrast-inversion as is also seen in the experimental images in Fig. 4e. The bias at which the map is obtained is indicated in the top-right of each image, which are individually normalized. The constant-current corresponds to an iso-value of $1 \times 10^{-7}/a_0^3$, where a_0 is the Bohr radius.
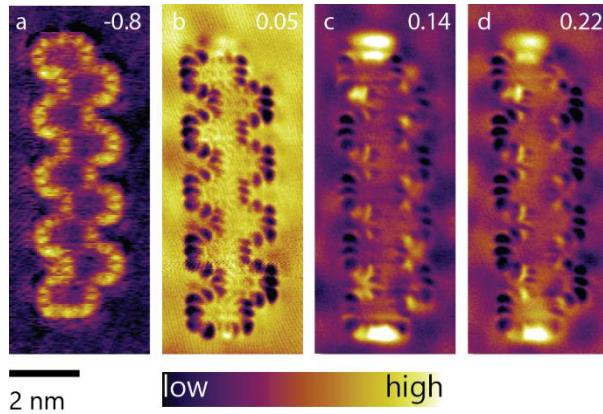
S7: Additional experimental dI/dV maps

Figure S8: dI/dV maps taken at constant-current. a-d, dI/dV maps acquired in constant-current mode, with I = 200 pA and the bias voltages (in V) indicated in the image. The GNR states hybridize with the electrons of the gold substrate. As a consequence, individual orbitals of the GNR are smeared out in energy. The map acquired at 0.05 V (**b**) has no bright lobes at the end of the GNR. It also exhibits the intensity inversion that afflicts such measurements when the molecular orbitals are hybridized (New J. Phys. 19 113033, New J. Phys. 19 113033). The bright lobes appear when the bias is increased to 0.14 V (**c**) and persist at even higher bias voltages (**d**). These observations indicate that, due to the hybridization with the substrate, the VB extends up to at least 0.05 V.

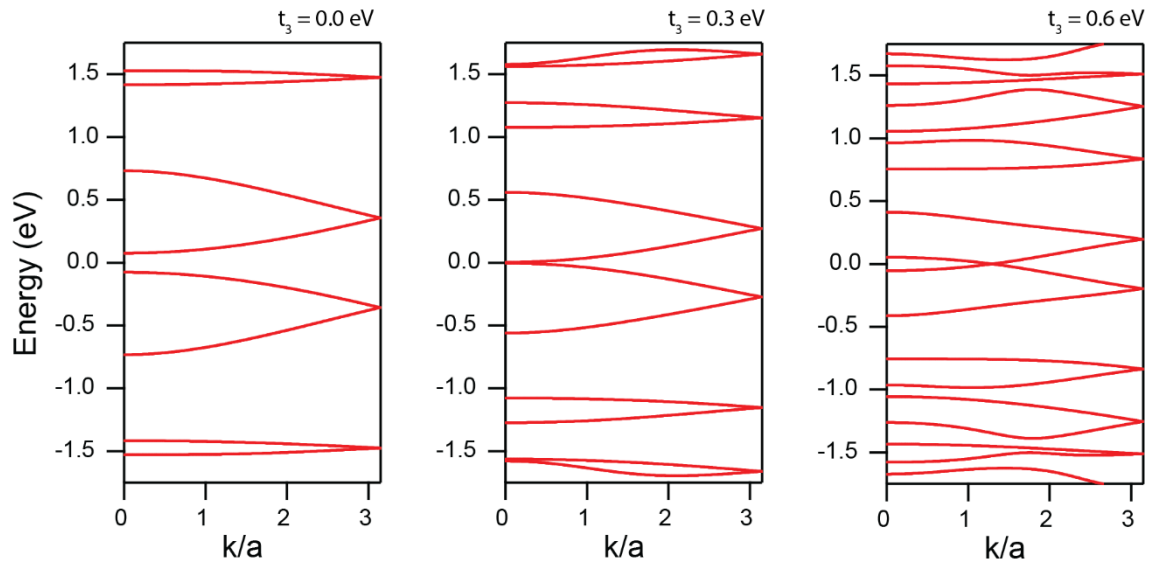## S8: Topologically trivial nature of 3-ZGNR-E(Bisanthene, 7)

Figure S9: 3-ZGNR-E(BA, 7) is topologically trivial. The presence of near Fermi-level peaks in the STS (see Fig. 4b in the main text), localized at the ends of the GNRs, warrants a careful elucidation of the topological class of the GNRs. To this end, we perform tight-binding (TB) simulations with t_1 = 3 eV, t_2 = 0 eV and t_3 = 0.3 eV. At the most basic level, the TB simulations of a finite GNR segment do not show mid-gap end states, showing that the 3-ZGNR-E(BA, 7) is topologically trivial. Additionally, the Zak phases of the four frontier bands VB −1, VB, CB and CB+1 are found to be identical. To change the topological invariant of the GNR, we would have to switch bands with different Zak phases. In the present case, it would necessarily involve the bands more distant from the Fermi level than the four frontier bands. Therefore, the topologically trivial phase observed here is very robust. To illustrate this further, we attempt to tune the GNR into a topologically non-trivial phase by changing the 3rd nearest neighbor hopping (t_3) from 0 eV to 0.6 eV . At t_3 = 0 eV the system has a bandgap. A topological transition would be indicated by the closing of the bandgap and its reopening as a tuning parameter is adiabatically varied. However, we do observe the

bandgap closing at t_3 = 0.3 eV, however the band gap does not re-open as t_3 is increased further. Therefore, the topological class of the GNR does not change with t_3. As the system is established to be topologically trivial at t_3 = 0.3 eV, it therefore must remain topologically trivial across the studied values of t_3.

S9: Charging of the 3-ZGNR-E(Bisanthene, 7) on Au(111)

Projected density of states¶

Here, we provide a didactic discussion on the projected density of states (PDOS).

Let the molecule be defined by: $H_M = \begin{bmatrix} M_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_p \end{bmatrix}$ with $|M_i\rangle$ as the eigenvectors with M_i eigenvalues. Each gas-phase molecular orbital would be one eigenvector.

We can define the Hamiltonian and eigenvectors for the substrate as: $H_S = \begin{bmatrix} S_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S_q \end{bmatrix}$ and $|S_i\rangle$.

The combined Hamiltonian of the molecule and the slab in a correspondingly expanded Hilbert basis would be: $H_C = \begin{bmatrix} H_M & T \\ T^\dagger & H_S \end{bmatrix}$ with $|C_i\rangle$ as the eigenvectors, where T is a $p \times q$ matrix that introduces the coupling between the substrate and the molecule, with $T^\dagger$ as its transpose conjugate.

When there is no coupling between the molecule and the slab: $T = 0_{p \times q}$. In this scenario, H_C is a diagonal matrix, and its eigenvectors are simply those of the molecule and the substrate (expressed in the expanded Hilbert space):

$$|C^0{}_i\rangle = \begin{cases} |M_i\rangle, & i \leq p \\ |S_{i-p}\rangle, & p < i \leq p + q \end{cases}$$

Moreover, these eigenvectors are also orthonormal to each other. The dot product of the gas-phase molecular orbitals (expressed in the expanded Hilbert space) onto the eigenvectors of the molecule and the substrate is:

$$\sum_i^{p+q} |C^0{}_i\rangle \langle C^0{}_i | M_j \rangle$$

However, as $|C^0{}_i\rangle = |M_i\rangle$ for $i \leq p$:

$$\therefore \langle C^0{}_i | M_j \rangle = \delta_{ij}$$

The above result is trivial. We have shown that molecular orbitals map to themselves when they do not hybridize with the surface. However, if the molecule and the slab can hybridize, T is not a null matrix and H_C is no longer diagonal. Therefore, the eigenvectors of the coupled system, $|C^{hy}{}_i\rangle$ are in general:

$$|C^{hy}{}_i\rangle = \sum_k n_k^i |C^0{}_k\rangle$$

with $\sum_k |n_k^i|^2 = 1$ for normalization. Therefore, the dot product of the gas-phase molecular orbitals with all states of the hybridized molecule and slab system is:

$$\sum_i^{p+q} |C^{hy}{}_i\rangle \langle C^{hy}{}_i | M_j \rangle = \sum_i^{p+q} |C^{hy}{}_i\rangle \sum_k n_k^i \langle C^0{}_k | M_j \rangle$$

$$\Rightarrow \langle C^{hy}{}_i | M_j \rangle = \sum_k n_k^i \langle C^0{}_k | M_j \rangle$$

$$\therefore \left\langle C^{hy}{}_i | M_j \right\rangle = n_j^i$$

Thus upon hybridization, the jth gas-phase molecular orbital can have non-zero projections on all states of the adsorbed molecule-on-slab system.

As the $|C^{hy}{}_i\rangle$ form an orthonormal basis set: $\sum_i |n_j^i|^2 = 1$.

Therefore $|n_j^i|^2$ is a probability mass function. Since the slab is large, we can treat the discrete energy eigenvalues labeled with the variable i as a continuous function of energy. Thus, $|n_j^i|^2$ can then be expressed as a probability density function: $\int_{-\infty}^{\infty} n_j(E). dE = 1$.

This probability density function $n_j(E)$ is the projected density of states (PDOS) of the jth molecular orbital projected onto the hybridized system. PDOS is a useful parameter to reflect the hybridization of orbitals. For a non-hybridized system:

$$n_j(E) = \delta(E - M_j)$$

Any hybridization would broaden the delta function.

Moreover, consider if instead of hybridization between the substrate and the molecules, there was simply a rigid shift in the energy of the molecular orbitals when the molecule is adsorbed on the system:

$$H_C = \begin{bmatrix} H_M + u.\, \mathbb{I} & 0 \\ 0 & H_S \end{bmatrix}$$

In such a scenario, $\langle C^0{}_i | M_j \rangle = \delta_{ij}$ would still be valid. However:

$$\langle C^0{}_i | H_C | C^0{}_i \rangle = \begin{cases} M_i + u, & i \le p \\ S_i , & p < i \le p + q \end{cases}$$

The PDOS in this situation would be:

$$n_j(E) = \delta(E - (M_j + u))$$

Thus, the peak of the PDOS can also reflect the shift in the energy of the orbitals once they have adsorbed on the surface.

The PDOS for the gas phase orbitals of the 9-unit 3-ZGNR-E(BA, 7)a segment and the adsorbed same segment adsorbed on an Au(111)-slab are shown in Fig. S10. Most orbitals reveal a significantly broadened PDOS revealing a clear hybridization with the surface. Hole doping induced by interaction with the surface is evidenced by systematic shifts of the prominent peak of each PDOS to higher energy than the corresponding gas-phase orbital. Consequently, the HOMO, HOMO-1, and HOMO-2 are substantially depopulated, thus rendering the GNR closed shell.
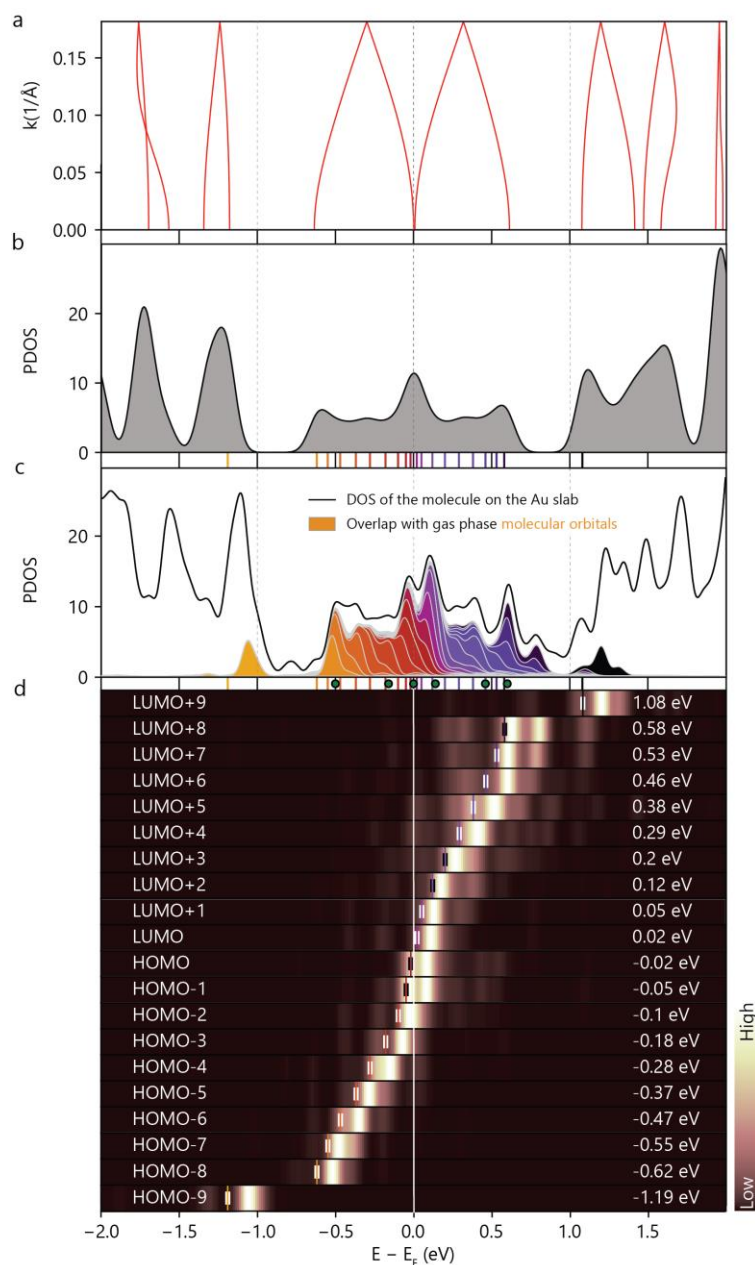
Figure S10: Hybridization of the GNR orbitals with the Au slab. **a**, Gas phase DFT band structure of the 3-ZGNR-E(Bisanthene, 7). **b**, The corresponding gas-phase density of states (DOS). **c**, Total DOS (black line) of a 3-ZGNR-E(BA, 7) segment consisting of 9 precursor units adsorbed on Au(111). PDOS of 10 occupied (HOMO to HOMO-9) and 10 unoccupied (LUMO to LUMO+9) orbitals are shown in colored overlaid peaks. The energies of all the calculated orbitals are indicated by colored lines on the x-axis of panels **b** and **c**. For reference, the energies of the simulated dI/dV maps shown in Fig. 4 of the main text are indicated by green dots below the x-axis. **d**, Color maps representing the PDOS as a function of energy for the different orbitals. Colored short vertical lines indicate the gas-phase energy of the corresponding orbital; these are also written on the right-hand-side.
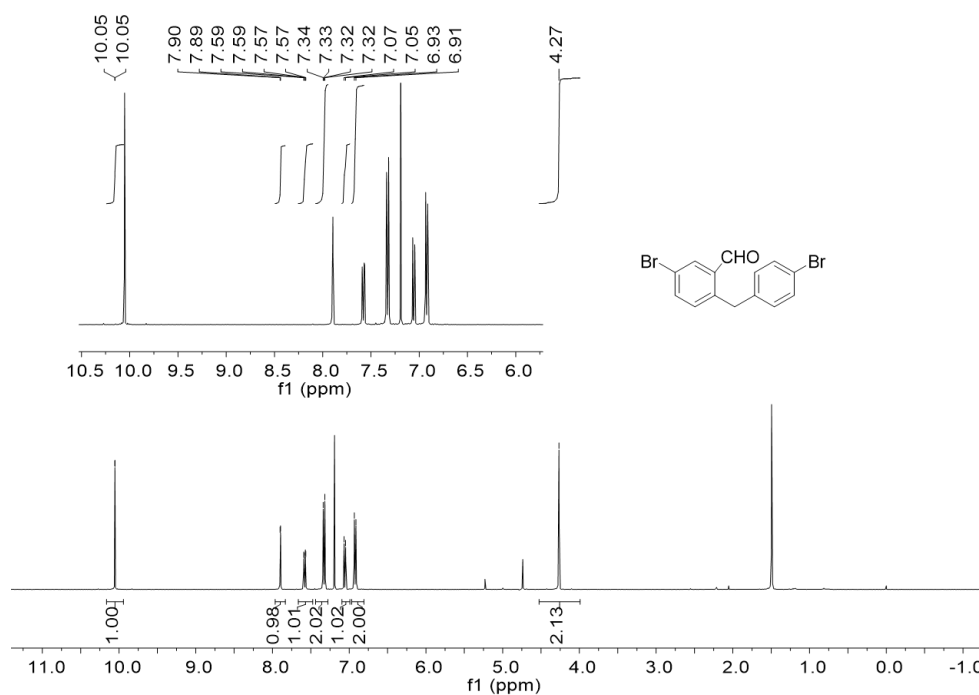
S10: NMR Spectra

Figure S11. ^{1}H NMR spectrum of compound **3** in CDCl$_3$ (400 MHz, 298 K).
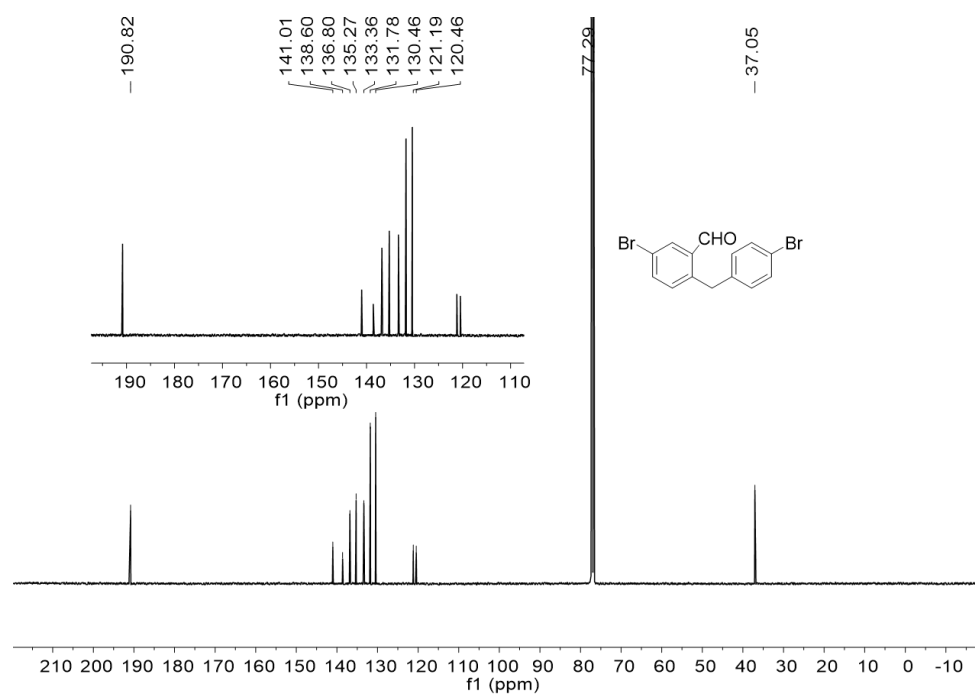

Figure S12. ^{13}C NMR spectrum of compound **3** in CDCl$_3$ (126 MHz, 298 K).
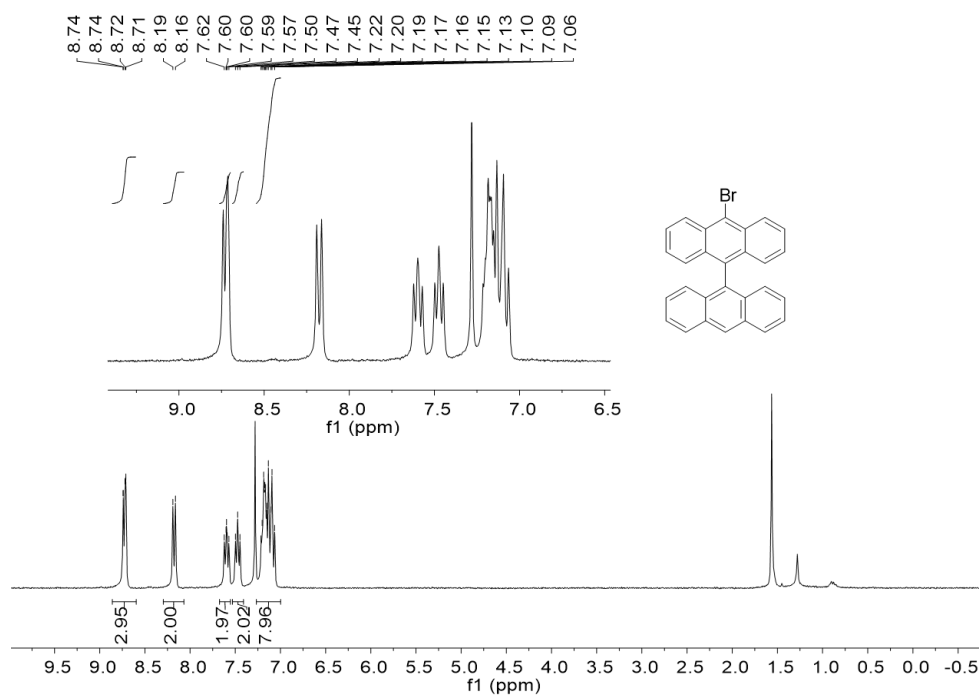
Figure S13. ^{1}H NMR spectrum of compound **5** in CDCl$_3$ (300 MHz, 298 K).

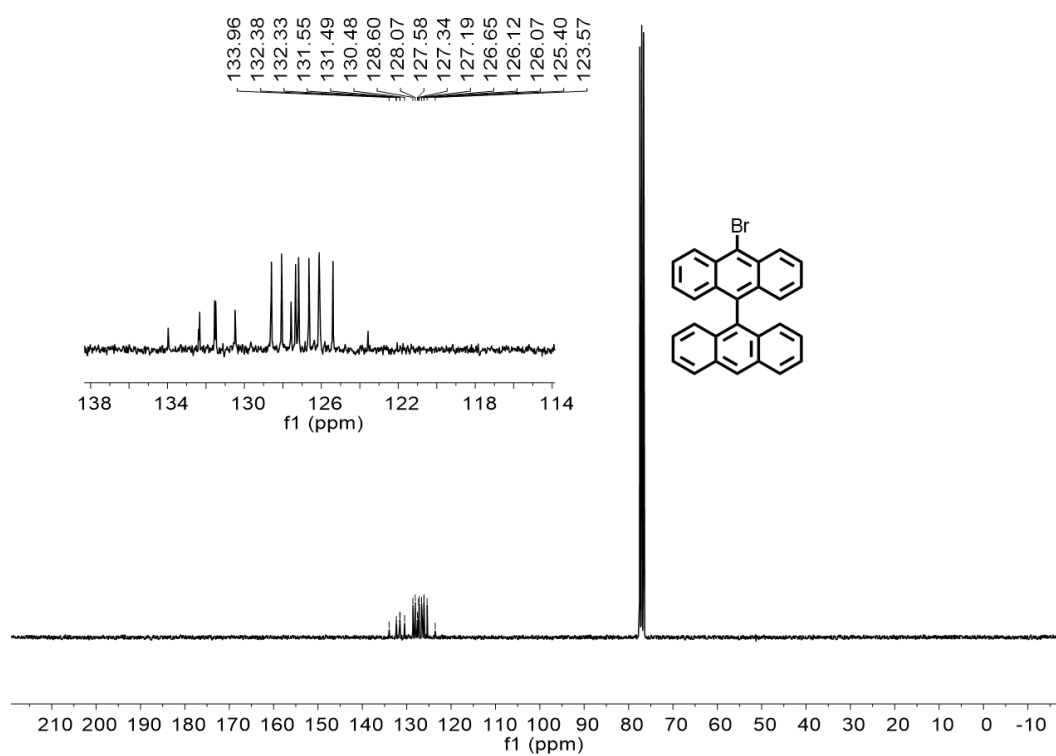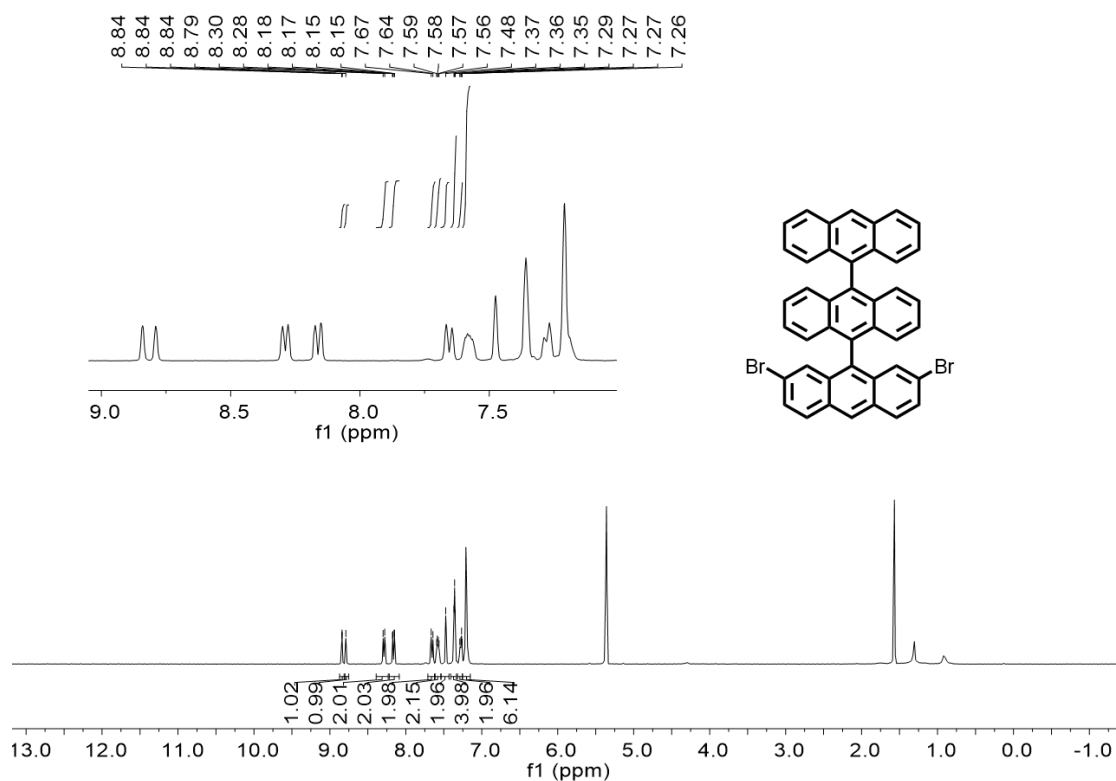**Figure S14**. ^{13}C NMR spectrum of compound **15** in CDCl$_3$ (75 MHz, 298 K).

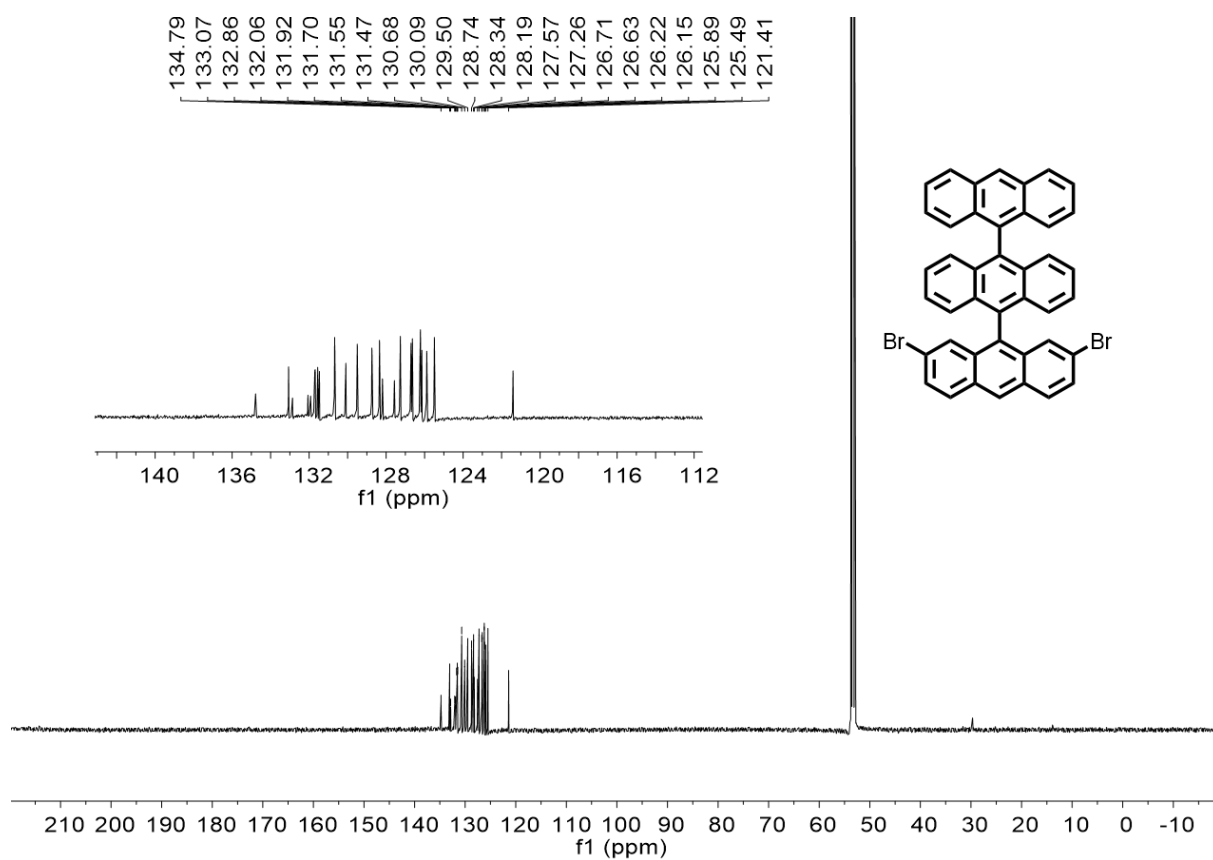**Figure S15**. ^{1}H NMR spectrum of compound **10** in CD$_2$Cl$_2$ (400 MHz, 298 K).

Figure S16. ^{13}C NMR spectrum of compound **10** in CD$_2$Cl$_2$ (126 MHz, 298 K).