Designing highly delocalized solitons by harnessing the structural parity of π-conjugated polymers

Supporting information

Kalyan Biswas, Jesús Janeiro, Aurelio Gallardo, Marco Lozano, Ana Barragán, Berta Álvarez, Diego Soler-Polo, Oleksandr Stetsovych, Andrés Pinar Solé, Koen Lauwaet, José M. Gallego, Dolores Pérez, Rodolfo Miranda, José I. Urgel, Pavel Jelínek, Diego Peña & David Écija

TABLE OF CONTENTS

1. Experimental details for the synthesis of precursor 1.

General methods for synthesis and characterization in solution

All reactions were carried out under argon using oven-dried glassware. THF, CH_2Cl_2 (DCM) and CH_3CN were purified by a MBraun SPS-800 Solvent Purification System. Commercial reagents were purchased from ABCR GmbH or Sigma-Aldrich.

An anhydrous stock solution of TBAF (1M in CH_3CN) was prepared as described in literature[1]. TLC was performed on Merck silica gel 60 F254 and chromatograms were visualized with UV light (254 and 365 nm). Column chromatography was performed on Merck silica gel 60 (ASTM 230-400 mesh). NMR spectra were recorded on a Varian Mercury-300 or Varian Inova 500 instrument. Low resolution mass spectra (EI) were obtained at 70 eV on a HP-5988 A instrument, while MALDI-TOF were determined on a Bruker Autoflex instrument. 1,4-Dihydro-1,4-epoxynaphthalene (**8**)[2] and 6-(trimethylsilyl)-2,3-dihydro-1*H*-inden-5-yl trifluoromethanesulfonate (**4**)[3] were synthesized following reported procedures. Figure S1 shows the synthetic route followed to obtain compound **1**.

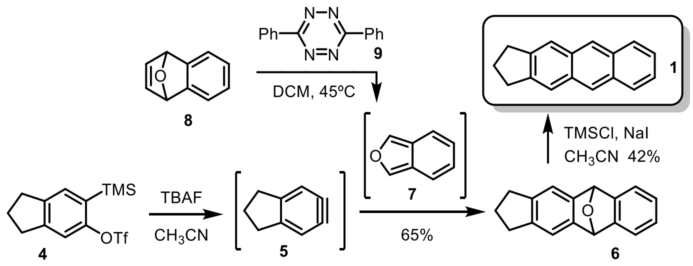
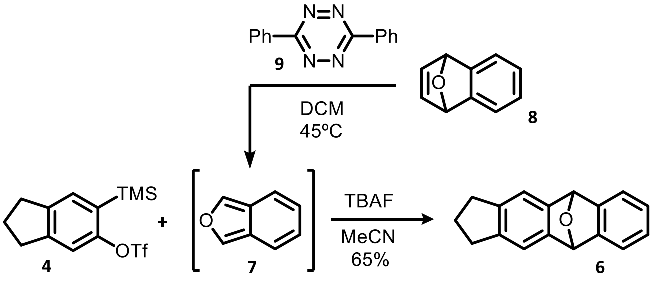


Figure S1 | Synthesis of 2,3-dihydro-1H-cyclopenta[b]anthracene (**1**).

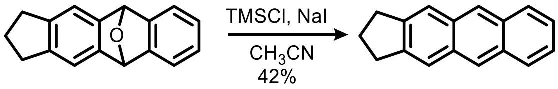
Synthesis of 2,3,5,10-tetrahydro-1H-5, 10-epoxyciclopenta[b]anthracene (6)



A mixture of 1,4-dihydro-1,4-epoxynaphthalene (**8**, 121.4 mg, 0.84 mmol) and 3,6-di(pyridin-2-yl)-1,2,4,5-tetrazine (**9**, 198.9 mg, 0.84 mmol) in dry CH_2Cl_2 (11 mL) were stirred at 45°C for 90 min. Then, triflate **4** (150 mg, 0.44 mmol) was added and a solution of tetrabutylammonium fluoride (1M in CH_3CN, 173.8 mg, 0.66 mmol) was dropwise added. The reaction mixture was stirred at 45°C for 16h. Then, the solvent was removed under reduced pressure and the product was purified by column chromatography (SiO_2, hexane/AcOEt , 4:1) to isolate compound **6** as a white solid (61.4 mg, 59%). ^{1}H-NMR (300 MHz, $CDCl_3$), δ: 7.36 – 7.21 (m, 2H), 7.18 (s, 2H), 7.00 (m, 2H), 6.00 (s, 2H), 2.83 – 2.77 (m, 4H), 2.09 – 2.02 (m, 2H) ppm. ^{13}C-NMR-DEPT (75 MHz, $CDCl_3$), δ: 148.52

(2xC), 146.62 (2xC), 141.77 (2xC), 125.81 (2xCH), 120.12 (2xCH), 116.93 (2xCH), 82.76 (2xCH), 32.56 (2xCH_2), 25.54 (CH_2) ppm. HRMS (APCI-DIP-TOF) for $C_{17}H_{14}O$ ([M+H^+]) Calcd.: 235.1117; Found: 235.1180.

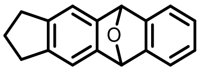
Synthesis of 2,3-dihydro-1H-cyclopenta[b]anthracene (1)



6 **1**

To a stirred solution of compound **6** (37 mg, 0.15 mmol) in dry CH_3CN (1.6 mL), freshly distilled TMSCl (0.12 mL, 0.95 mmol) and NaI (142 mg, 0.95 mmol) were added. The mixture was stirred at rt for 3h. Then, the solvent was evaporated under reduce pressure and the product was purified by column chromatography (SiO_2, hexane/AcOEt, 9:1), affording anthracene **1** as a white solid (14.5 mg, 42%). ^{1}H-NMR (300 MHz, $CDCl_3$), δ: 8.32 (s, 2H), 7.96 (dd, *J* = 6.4, 3.3 Hz, 2H), 7.80 (s, 2H), 7.41 (dd, *J* = 6.6, 3.2 Hz, 2H), 3.09 (t, *J* = 7.4 Hz, 4H), 2.30 – 2.01 (m, 2H) ppm. ^{13}C-NMR-DEPT (75 MHz, $CDCl_3$), δ: 143.55 (2xC), 131.69 (2xC), 131.28 (2xC), 128.02 (2xCH), 125.34 (2xCH), 124.74 (2xCH), 121.63 (2xCH), 32.62 (2xCH_2), 26.27 (CH_2) ppm. HRMS (APCI-DIP-TOF) for $C_{17}H_{14}$ ([M+H]$^+$) Calcd.: 219.1168; Found: 219.1163.

Spectroscopic data



6

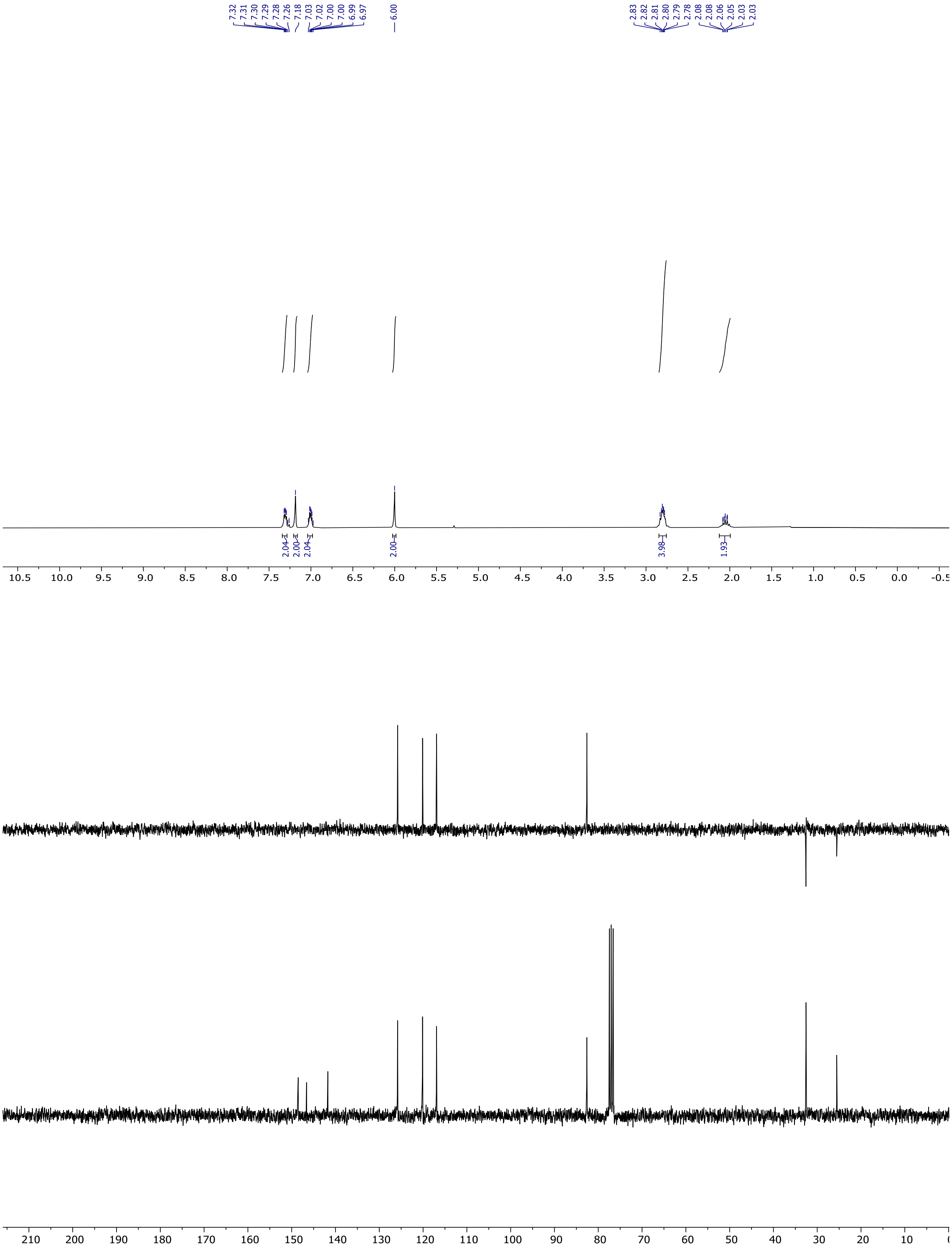
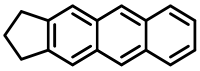


Figure S2 | ^{1}H NMR, ^{13}C NMR and DEPT spectra of compound **5**.



1

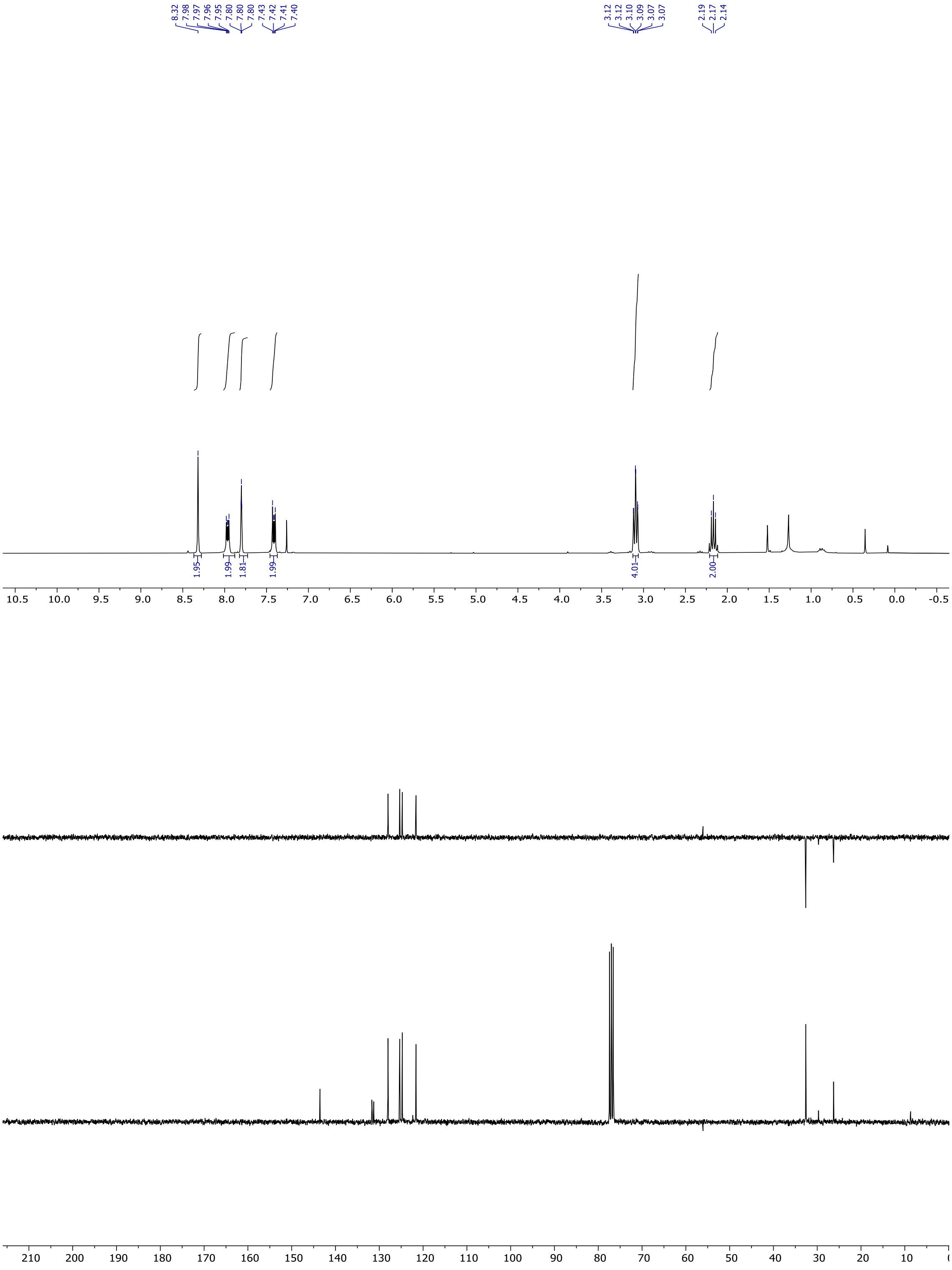


Figure S3 | ^{1}H NMR, ^{13}C NMR and DEPT spectra of compound **1**.

2. Self-assembly of precursor 1 on Au(111).

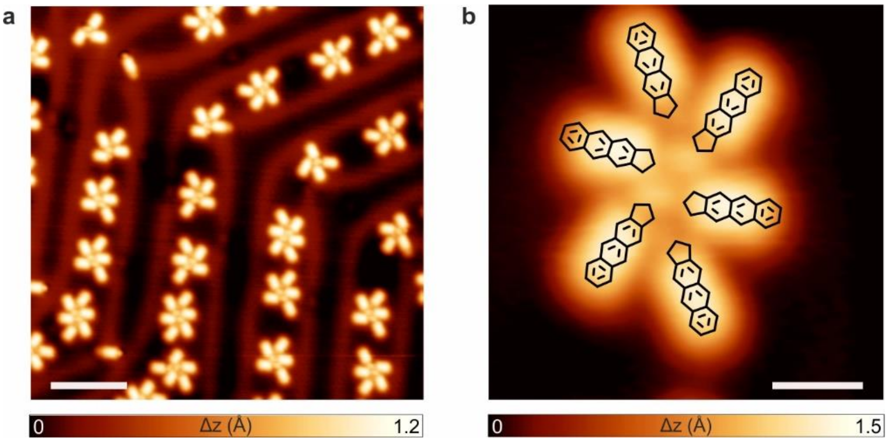


Figure S4 | Self-assembly of precursor 1 on Au(111). a, STM image of the self-assembly of **1** on the Au(111) substrate held at room temperature after deposition of a submonolayer coverage. **b**, High-resolution zoom-in STM image of a supramolecular hexamer with superimposed chemical sketches of precursors. **a**, V_b = -0.5 V, I_t = 50 pA, scale bars = 5 nm; **b,** V_b = -0.5 V, I_t = 50 pA, scale bars = 1 nm.

3. Statistics of length and structural parity of polymer 3.

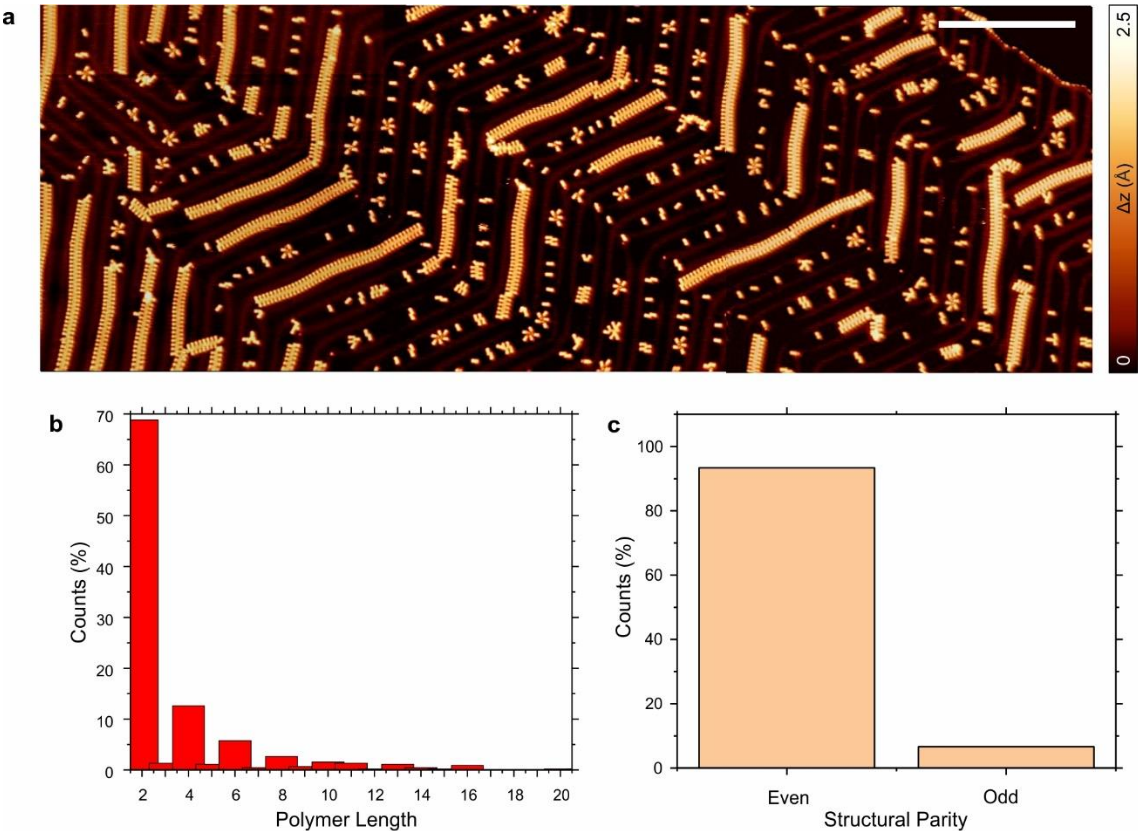


Figure S5 | Statistics of length and structural parity of polymer 3. a, Large area overview STM image after annealing a submonolayer coverage of precursor **1** on Au(111) at 200 °C for 10 minutes. V_b= 0.5 V, I_t = 10 pA, scale bar = 20 nm. **b**, Histogram depicting the distribution of the polymer length. **c**, Histogram of the structural parity (even vs. odd) of the polymer.

4. Extra hydrogenation of the backbone of polymer 3.

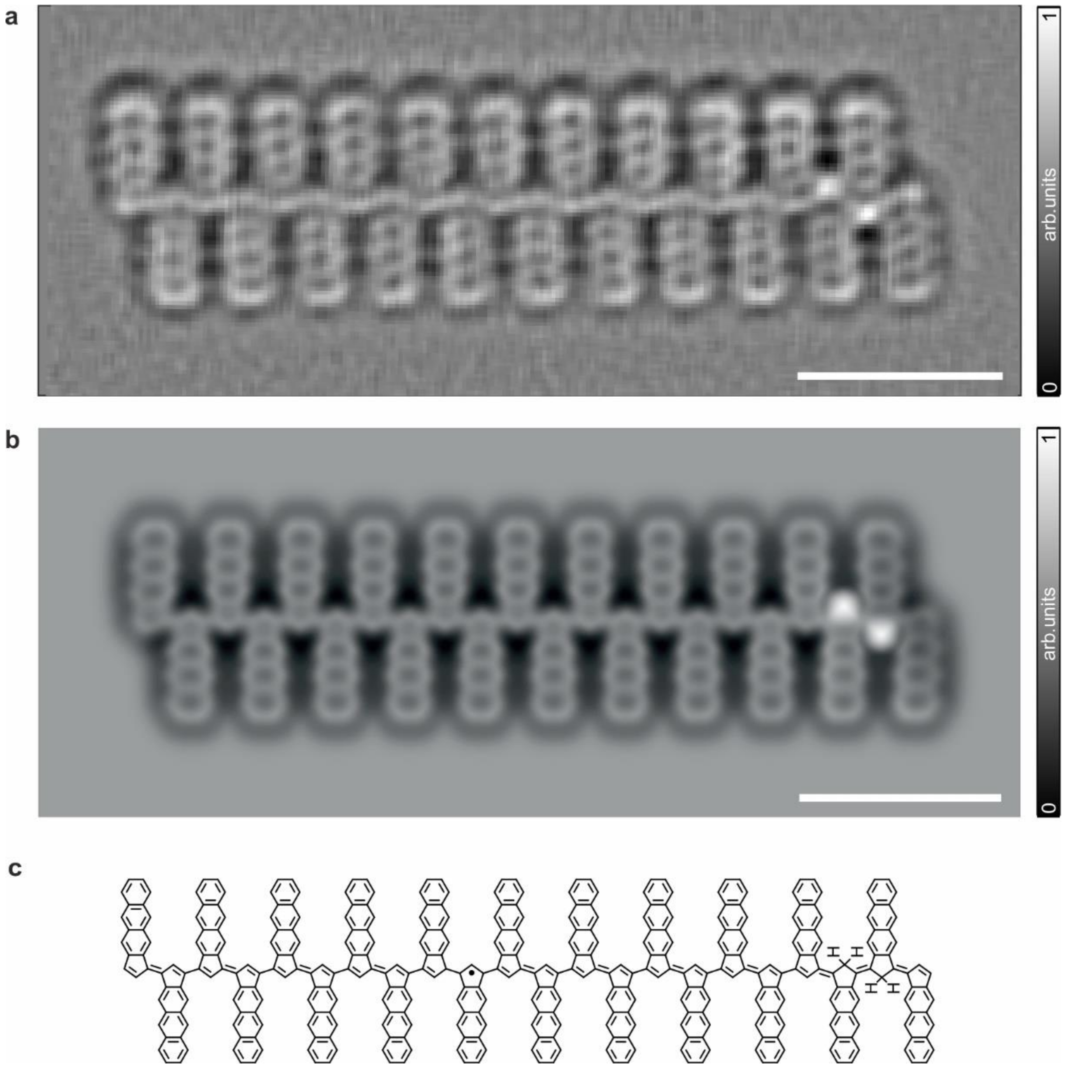


Figure S6 | Extra hydrogenation of the backbone of polymer 3. a, Constant-height frequency-shift, Laplace-filtered, nc-AFM image, acquired with a CO-functionalized tip, of an extra hydrogenated chain of 22 monomer units at two carbon locations (z_{offset} = 120 pm above STM set point: V_b = 5 mV, I_t = 50 pA). **b**, Simulated nc-AFM image of **a**. **a**-**b**, Scale bars = 1 nm. **c**, Corresponding chemical structural model of **b**.

The AFM simulated images were obtained by means of the probe particle model[4] using the molecular structure relaxed in gas phase, given the large size of the polymers. Taking into account the obtained planarity of the polymers and the agreement of the obtained results with the experimental images, we considered this approximation adequate. We characterized the simulated tip with a stiffness parameter of $0.24\, N \cdot m^{-1}$ and charge coefficient of $-0.2\, e$. The tip-sample electrostatic interaction was included in the calculation by means of the Hartree potential calculated with the FHI-AIMS package[5]. The geometry optimization was done at the GGA-PBE level of the theory[6], with energy and force tolerances of $10^{-5} eV$ and $10^{-2} eV/\text{Å}$ respectively, light wavefunctions and the Tkatchenko-Scheffler treartment of the van der Waals interaction[7].

5. Tip-induced removal of extra hydrogen from odd-membered polymer 3o.

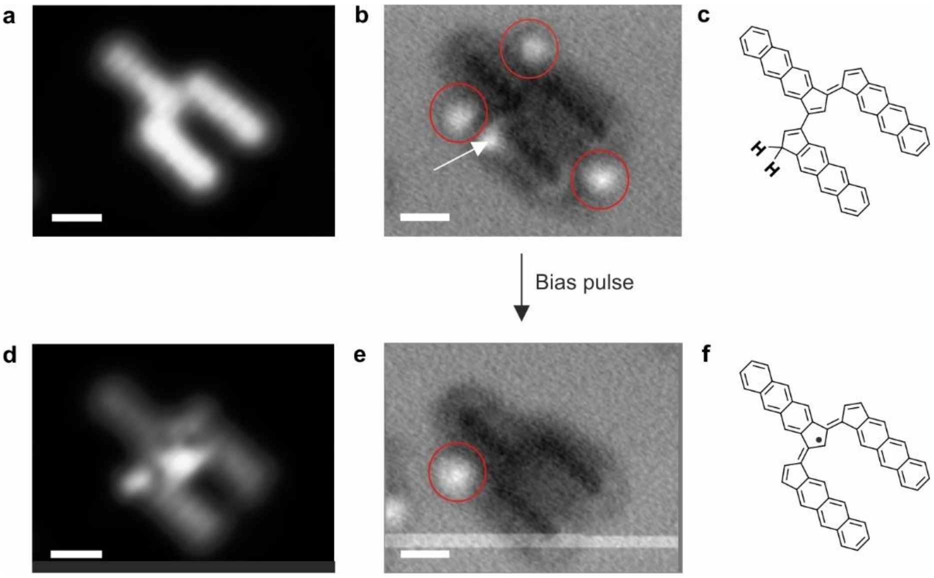


Figure S7 | Tip-induced removal of hydrogen from a doubly hydrogenated position in a three-membered polymer. a,**d**, High-resolution constant-height STM images of a trimer before and after the removal of an extra hydrogen, respectively. V_b = 10 mV, I_t = 50 pA. **b,e**, Simultaneously acquired constant-height frequency-shift nc-AFM images with a CO-functionalized tip. The red circles highlight the presence of CO molecules on the surface. Additionally, the doubly hydrogenated position within the polymer backbone is indicated by a white arrow. Three CO molecules were surrounding the molecule before the C-H cleavage. The extra hydrogen was removed by locating the tip on top of the doubly hydrogenated position, subsequently lifting the tip 200 pm and applying 1.8V until the current dropped and the H was removed. After the C-H cleavage, one CO still remains close to the molecule, and importantly now the trimer displays the soliton. Scale bars = 0.5 nm. **c**,**f**, Chemical structural model of **a**,**b** and **d**,**e**, respectively.

6. Bond length analysis of DFT optimized structures of even- and odd-membered polymers 3.

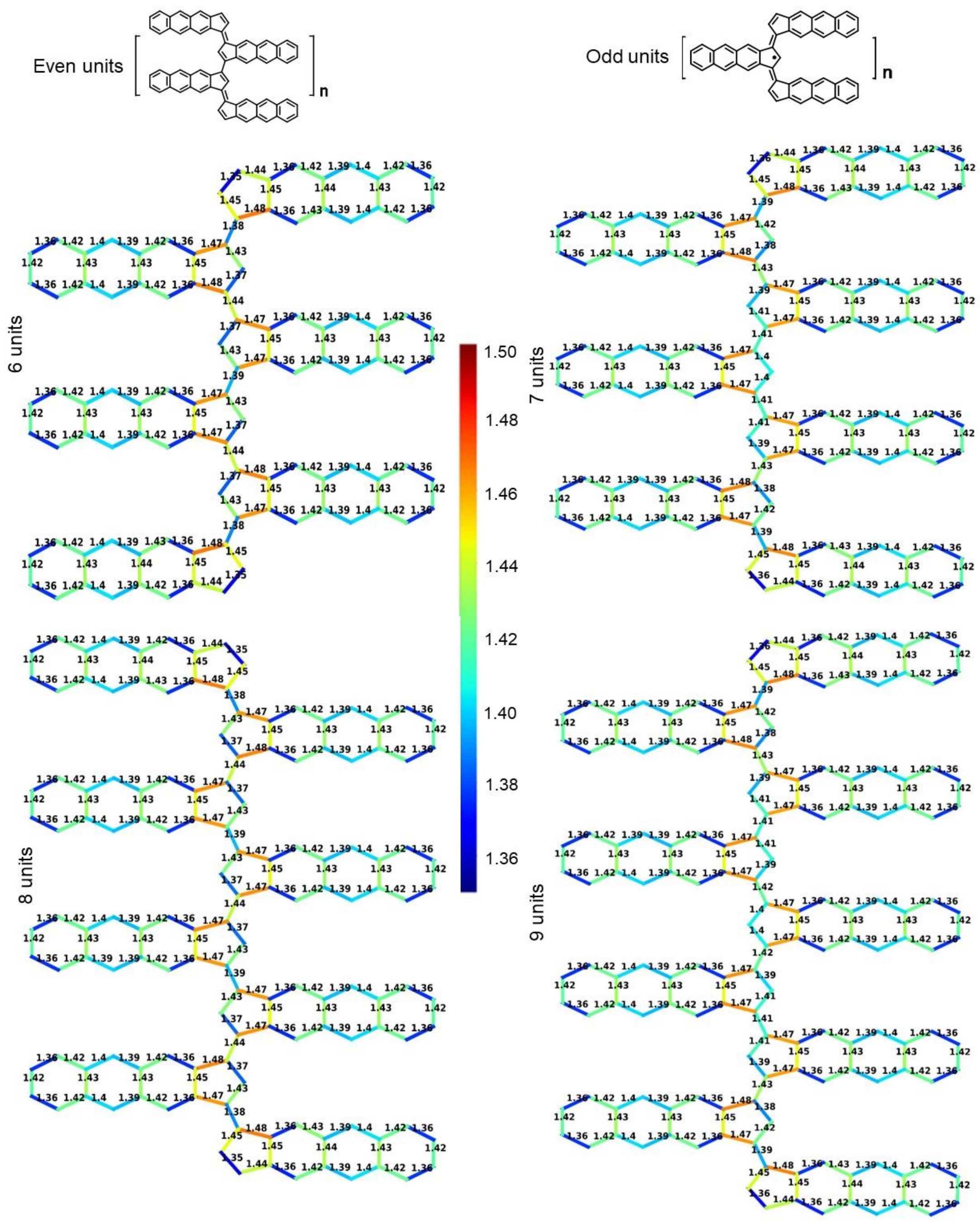


Figure S8 | Bond length analysis of even- and odd-membered polymers. Bond lengths were calculated using the hybrid B3LYP functional and plotted in the selected color range (1.35 to 1.50 Å).

7. Structural details of even- and odd-membered polymers 3 using nc-AFM.

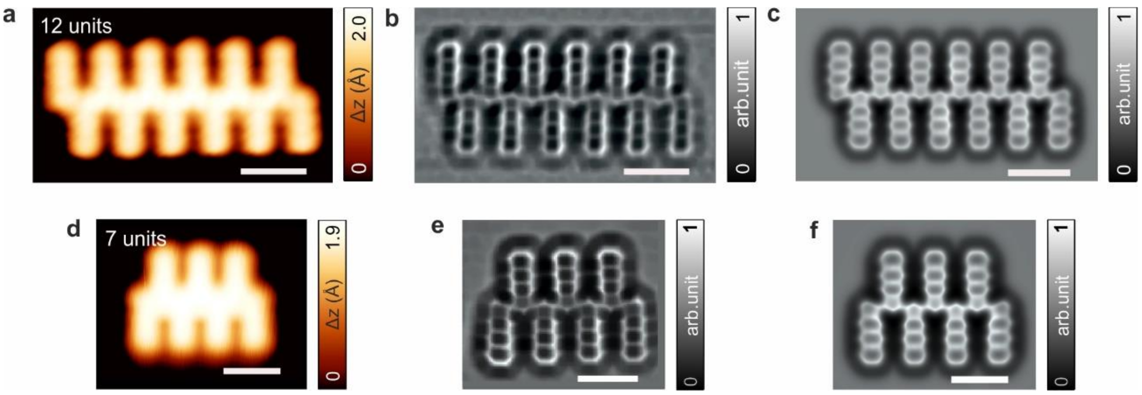


Figure S9 | Structural details of even- and odd-membered polymers 3 using nc-AFM. a,**d**, High-resolution STM images, acquired with a CO-functionalized tip, of isolated 12-membered and 7-membered polymer, respectively. V_b = 5 mV, I_t = 50 pA, scale bars = 1 nm. **b**,**e**, Constant-height frequency-shift nc-AFM images of **a** and **d**, respectively, acquired with a CO-functionalized tip. **b,** V_b = 5 mV, I_t = 50 pA; **e,** V_b = 10 mV, I_t = 50 pA. Scale bars = 1 nm. **c**,**f**, Simulated nc-AFM images with a CO-tip for the corresponding nc-AFM images of **b**,**e** using an optimized molecular structure obtained from total energy DFT calculations. Scale bars = 1 nm.

The AFM simulated images were obtained by means of the probe particle model[4] using the molecular structure relaxed in gas phase, given the large size of the polymers. Taking into account that the obtained planarity of the polymers and the agreement of the obtained results with the experimental images, we considered this approximation adequate. We characterized the simulated tip with a stiffness parameter of $0.24\ N \cdot m^{-1}$ and charge coefficient of $-0.2\ e$. The tip-sample electrostatic interaction was included in the calculation by means of the Hartree potential calculated with the FHI-AIMS package[5]. The geometry optimization was done at the GGA-PBE level of the theory[6], with energy and force tolerances of $10^{-5} eV$ and $10^{-2} eV/\text{Å}$, respectively, as well as light wavefunctions and the Tkatchenko-Scheffler treatment of the van der Waals interaction[7].

8. Additional calculations for the reaction mechanism.

In this section, we provide additional simulations using QM/MM as well as DFT CI-NEB methods to support the discussion in the main text. Figure S10 summarizes possible Au adatom-assisted dehydrogenation processes discussed in the main text using QM/MM-based free energy MD calculations. The results presented in Figure S10 demonstrate the higher probability of the dehydrogenation reaction to occur following the mechanism proposed in Figure 3a-d over all the studied alternatives. We also observe that the further dehydrogenation of the intermediate S_4 will be an endothermic process.

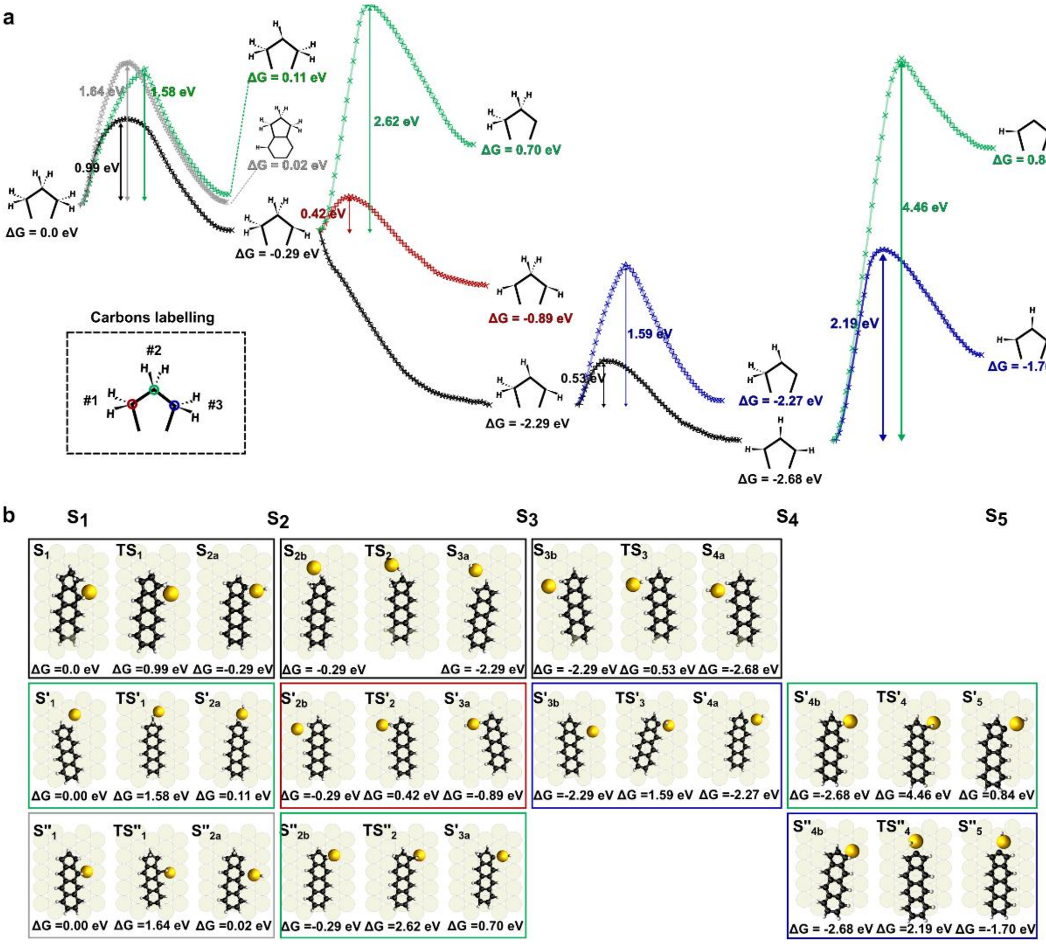


Figure S10 | Free energy barriers calculated in the frame of the QM/MM formalism and structures of the initial, intermediate, and final stages of alternative considered reaction pathways for certain stages of the C-H cleavage processes shown in Fig.2. a, Concatenation of the most favorable energy barrier (black line) as shown in Figure 3 of the main manuscript, compared with all the explored alternative reaction pathways. The alternative processes are depicted in red, green, and blue, corresponding to the C-H cleavage of the carbons #1, #2, and #3, or in grey for the C-H cleavage of carbon on a benzene ring. **b**, Initial, transitional, and final structures of all the free energy barriers shown in panel a.

Given that the dehydrogenation processes shown in Figure 3a-d and Figure S10 will involve the continuous creation and passivation of radicalic states, we had studied the alignment of the frontier orbitals HOMO/LUMO of the intermediates S_{1-4} with respect to the Au(111) surface, in order to identify the possible presence of charge transfer from intermediates to the surface.

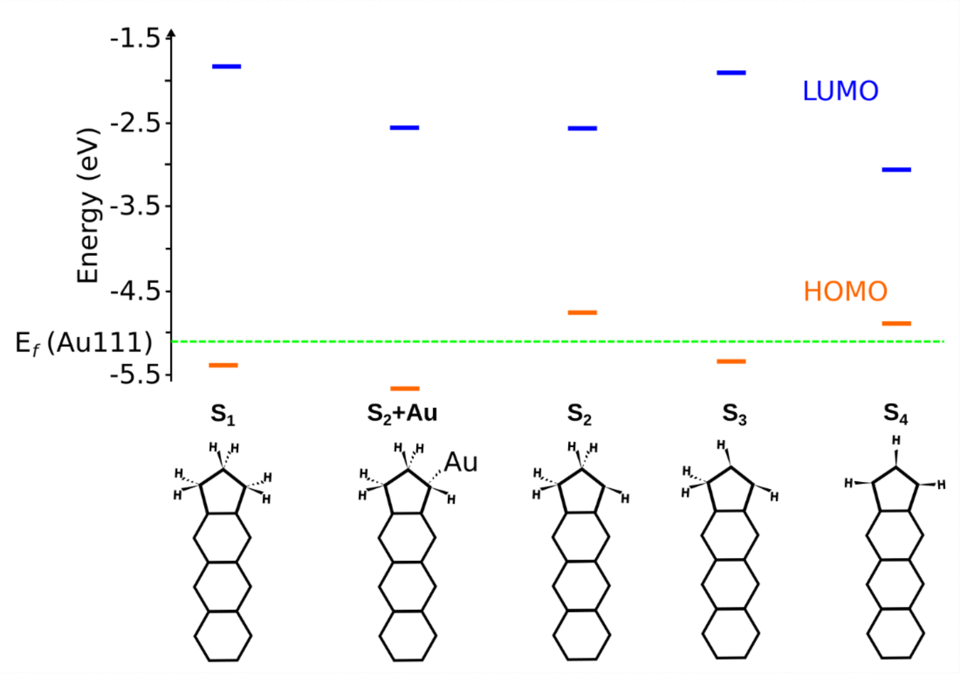


Figure S11 | Study of the molecule-substrate hybridization. Frontier orbitals HOMO/LUMO of the precursor S_1 intermediates S_{2-4}, and the intermediate S_2, passivated by a gold adatom, calculated with the spin-resolved, GGA-based, hybrid potential PBE0, implemented in the FHI-AIMS package, in gas phase plotted alongside the fermi energy of the Au(111) surface, placed at -5.1eV with respect to the vacuum level.

Figure S11 shows the frontier HOMO/LUMO orbital alignment of the precursor S_1 and intermediates $S_{2-4}$with respect to the Fermi level of the Au(111) surface. We observe that HOMO orbitals of intermediates S_2 and S_4 are located above the Fermi level, which indicates possible charge transfer from the intermediates to the surfaces. Nevertheless, according to our QM/MM simulations, during the dehydrogenation processes, gold adatom passivates the radical character of the resulting intermediate by forming organometallic complexes, stabilizing the final intermediates energetically. Importantly, the HOMO orbital of the gold complex ($S_2 + Au$) is located below the Fermi level, which indicates negligible charge transfer, as shown in Figure S11.

To explore alternative reaction mechanisms involving organometallic intermediates, we calculated activation free energies barriers of the dehydrogenation $S_2 + Au$ using QM/MM.

To confirm that our conclusions about the dehydrogenation of the intermediate S_2 were not affected by the expectable charge transference, we repeated the S_2 to S_3 and S'_2 to S'_3 transitions (shown in Figure S10) in the presence of an extra gold and hydrogen atoms from the previous dehydrogenation, using the QM/MM-based MD methodology. Figure 12 shows $S_2 + Au$ to $S_3 + Au$ and $S'_2 + Au$ to $S'_3 + Au$ transitions mediated by an extra Au adatom. We found in both cases a good qualitative agreement with the results shown in Figure S10, with relatively low actiation barriers and thermodynamycally stable final products.

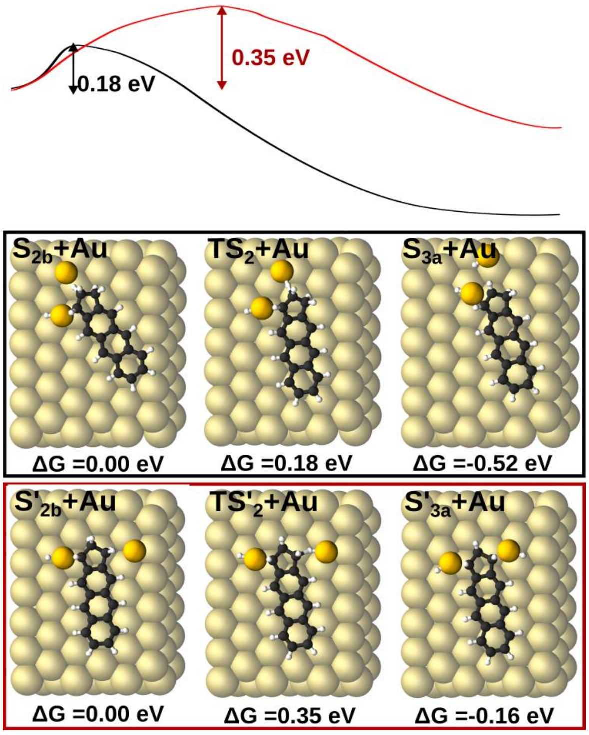


Figure S12 | Two alternative reaction mechanisms on complex $S_2 + Au$ intermediates. The activation free energy corresponding to the most probable barriers for the second dehydrogenation shown in Fig. S10 in the presence of the residual adatom and hydrogen from the first dehydrogenation using QM/MM method.

Next, we discuss the transition to S_5. First, we carried out CI-NEB DFT calculations of the S_4 to S_5 and S'_4 to S'_5 transitions, including one layer of Au(111) substrate (Figure S13), confirming the endothermic character of the reaction. Note that the application of 1 layer of Au(111) slab can be justified by comparison of adsorption energies using a slab made of 1 and 3 layers. Figure S16 compares the total energies of the four stages of the dimerization reaction shown in Fig. 3 (initial, final, and the two intermediate stages) between one-layer and three-layer Au(111) substrate. We found the maximum deviation from the original results to be 80 meV, supporting the validity of the one-layer slab model.

Second, we repeated the free energy calculation of the S_4 to S_5 transition, using the QM/MM method, in the presence of an extra hydrogen in the adatom, so the cleaved H atoms would be able to evaporate as H_2. Figure S14 displays the free energy barrier, which confirms the endothermic character of the S_4 to S_5 transition.

In conclusion, we found that the S_4 to S_5 transition to be endothermic independently of computational approaches.

Finally, we present an alternative reaction mechanism for the dimerization of the intermediates S_4 shown in the Figure 3e of the main text. Figure S15 shows the reaction mechanism, including coordinated cleavage of the two hydrogens from the intermediates S_4, as well as the initial and final energies for two extra alternative processes. The results indicate that the dimerization of the intermediates S_4 will most likely happen as a two-step dehydrogenation of the lateral carbons of the 5-member carbon ring, followed by the formation of the covalent bond between them, as shown in the Figure 3e of the main text.

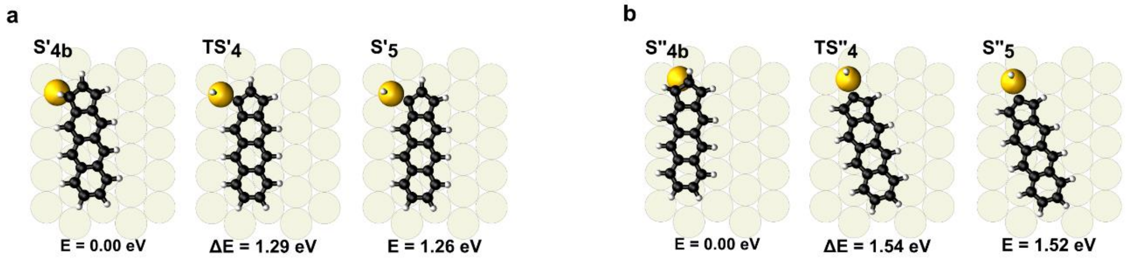


Figure S13 | Energy barriers for the last C-H cleavage process shown in Fig. S10 recalculated with CI-NEB DFT calculations, including a substrate for comparison. The panel's a) and b) show the process and energy barriers of dehydrogenating the carbons #1 and #2, respectively from the intermediates S'_{4b} and S''_{4b}. Including a substrate allows the charge transference between the molecules and the substrate, lowering the energy barrier of the process. Nevertheless, the quantitative difference between the energies shown in this figure and the last column of Fig. S10 might also be given by geometrical differences in the used models, such as the different number of Au(111) layers used of fundamental differences in the full DFT-based NEB and the QM/MM based-MD methods.

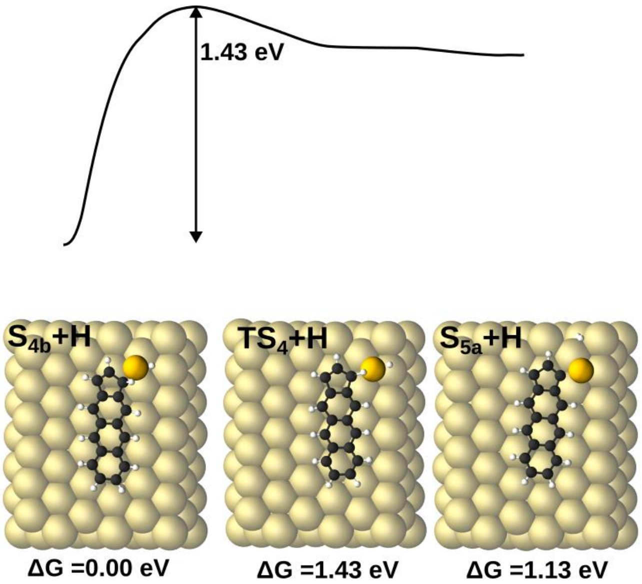


Figure S14 | Alternative reaction mechanism for the S_4 to S_5 transition. a, Activation free energy barrier modelling of the dehydrogenation process of the intermediate S_4 shown in Fig. S10, in the presence of an extra hydrogen to form molecular hydrogen, using QM/MM method.

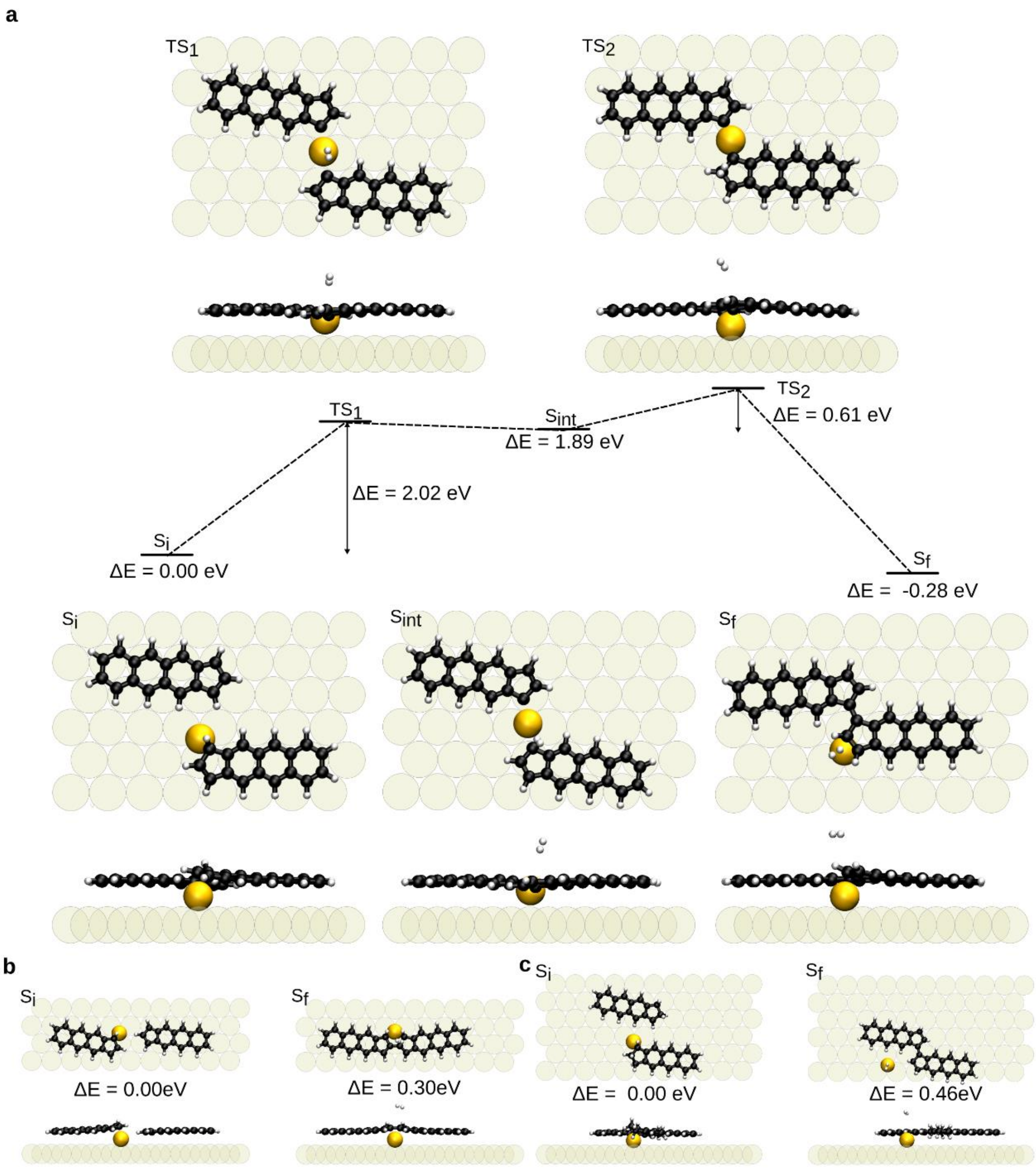


Figure S15 | Proposed alternatives to the dimerization process. a, Polymerization process, similar to that shown in Fig. 3e, where the C-H cleavage of the carbons occurs simultaneously. This alternative will present a potential energy barrier of 2.02 eV for the complete process of creation of the organo-metallic bond between the molecules and the two C-H cleavages that are necessary for its formation, instead of the two proposed barriers of 1.28eV and 1.17eV proposed in Fig. 3e for such purpose. **b**, Energy balance of the polymerization occurring through the carbons at the apex of the five-member carbon ring that would lead to the creation of linear dimers instead of the experimentally observed chains. This process was found to be thermodynamically unfavorable. **c,** Energy balance of the polymerization occurring from intermediates $S2_a$ after the first C-H cleavage. This process was found to be thermodynamically unfavorable.

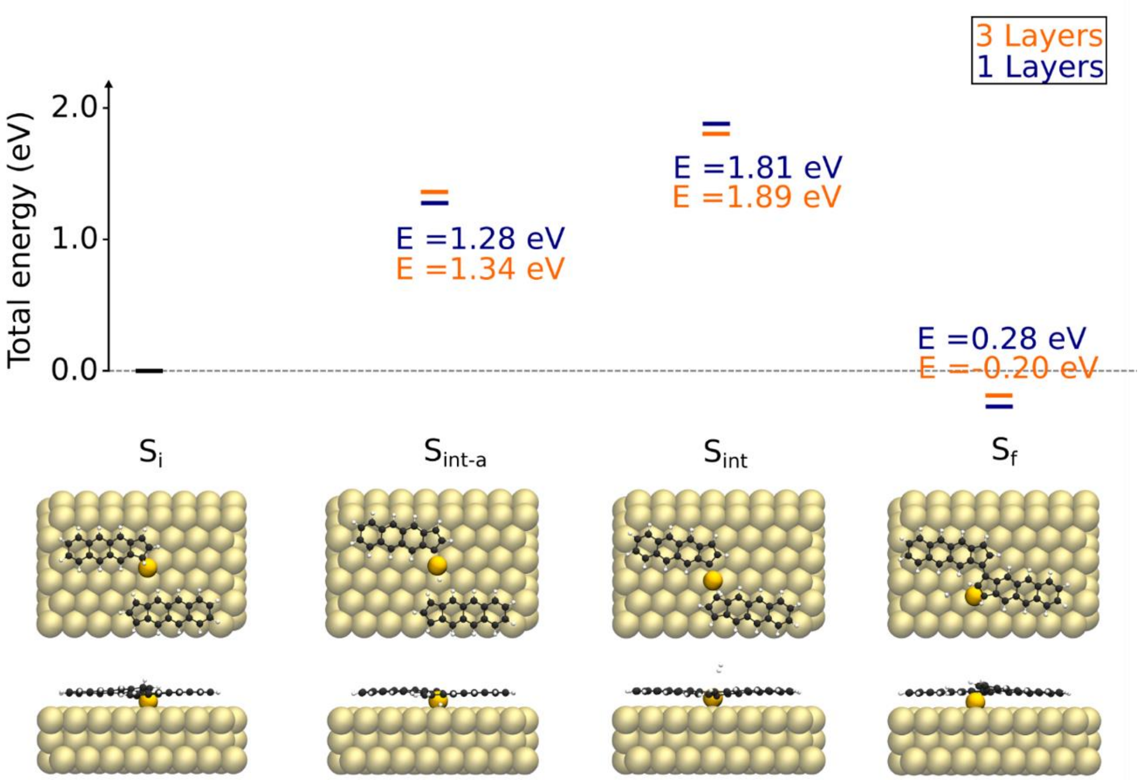


Figure S16 | Comparison between dimerization reaction stages taking place in a one-layered and three-layered Au(111) substrate. The total energy of the initial, two intermediate, and final stages of the dimerization process shown in Fig. 3e recalculated on a three-layered Au(111) substrate plotted against the one-layered results shown in Fig. 3e and the top and side geometries of the mentioned systems optimized on the three-layered substrate.

9. Correlation between the topographic contrast of the backbone and the structural parity of polymers 3.

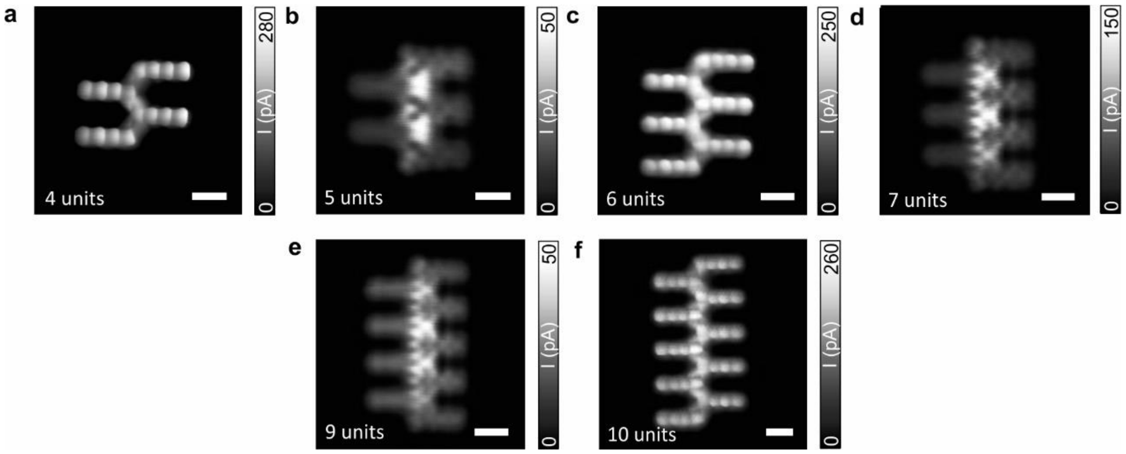


Figure S17 | Correlation between topographic contrast of the polymer backbone and structural parity. a-f, High-resolution constant-height STM images with a CO-functionalized tip of even- and odd-membered polymers V_b = 5 mV; I_t = 50 pA; z_{offset} = 70 pm, 200 pm, 70 pm, 150 pm, 180 pm, and 80 pm, respectively. Scale bars = 0.5 nm.

10. D*I*/d*V* spectra obtained using different STM tips for various polymers.

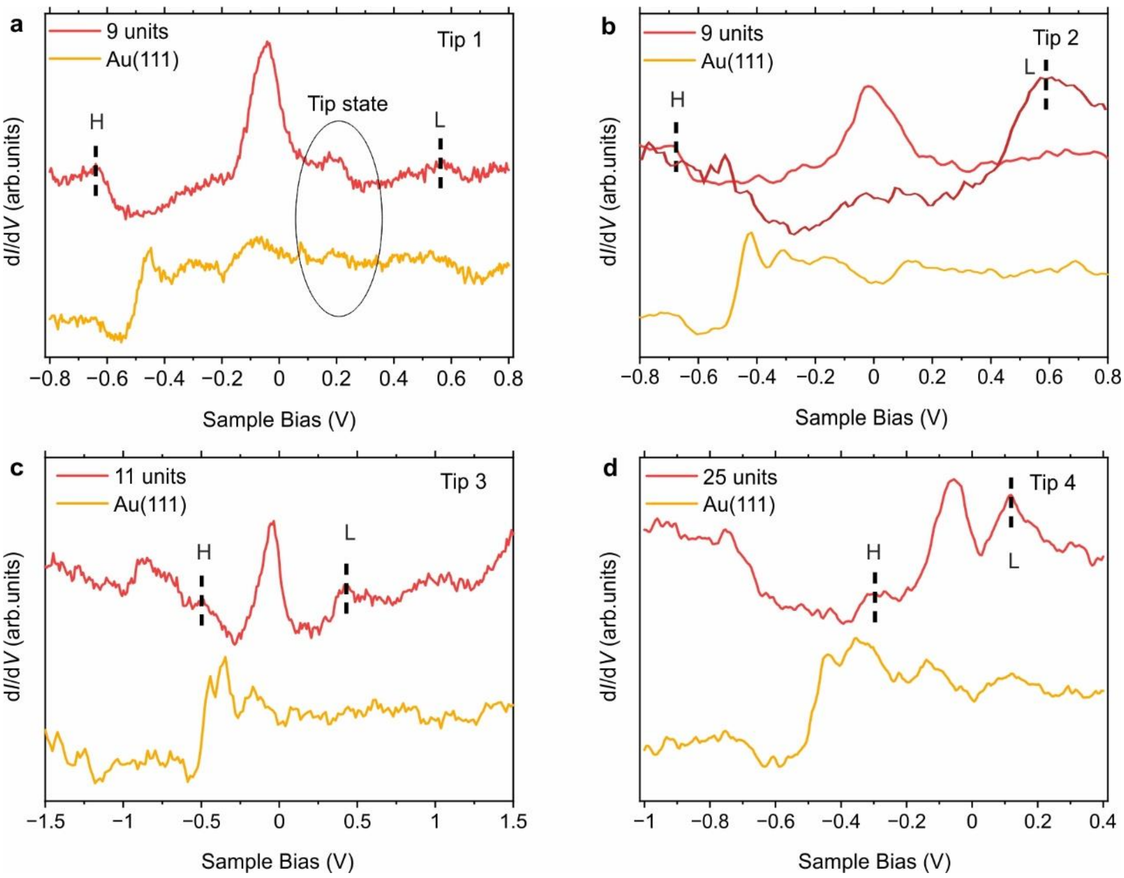


Figure S18 | D*I*/d*V* spectra obtained using different STM tips for various odd-membered polymers. The d*I*/d*V* spectra for 3, 5, 7, and 9-membered polymers in Figure 4j were acquired using Tip 1, revealing a kink at around 0.2 V in the spectra attributed to the tip state. However, this kink is not observed in the spectra that were obtained with Tip 2 for a 9-membered polymer (panel **b**) and with Tip 3 or Tip 4 for the 11- and 25-membered polymers, panels **c** and **d**, respectively.

11. Electronic structure of a 10-membered and a 9-membered polymer 3.

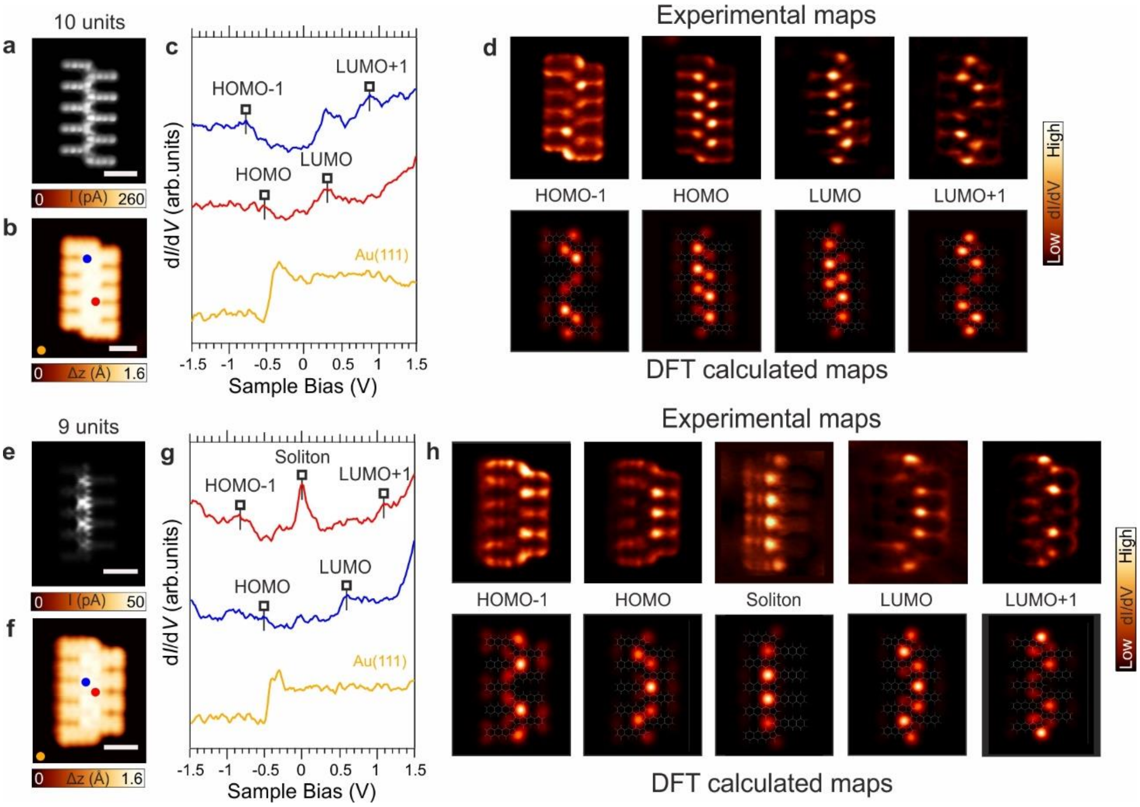


Figure S19 | Electronic structure of a 10-membered and a 9-membered oligo-naphthoindenylidene polymer. a,**e**, Constant-height high-resolution STM images of an even-membered and an odd-membered polymer acquired with a CO-functionalized tip, respectively. V_b = 5 mV, scale bars = 0.5 nm. **b**,**f**, High-resolution constant-current STM images of a 10-membered and a 9-membered polymer acquired with a CO-functionalized tip, respectively. V_b = −2.0 V, I_t = 100 pA, scale bars = 0.5 nm. **c**,**g**, d*I*/d*V* spectra of 10-membered and 9-membered polymer acquired at the positions indicated by the colored circles in **b**,**f**, respectively. Reference spectra taken on the bare Au(111) surface is depicted in orange and the acquisition positions marked with an orange circle. Open feedback parameters for d*I*/d*V* spectra of **b**,**e**: V_b = −2.0 V, I_t = 100 pA, V_{rms} = 20 mV. **d**,**h**, Constant-current d*I*/d*V* maps acquired with a CO-tip and corresponding DFT-calculated maps (tip-sample height = 5 Å) at the energy corresponding to the distinct frontier orbitals. HOMO – 1 (V_b = −0.87 V, I_t = 250 pA), HOMO (V_b = −0.56 V, I_t = 250 pA), LUMO (V_b = 0.25 V, I_t = 250 pA), and LUMO + 1 (V_b = 0.77 V, I_t = 250 pA); HOMO – 1 (V_b = −0.80 V, I_t = 250 pA), HOMO (V_b = −0.5 V, I_t = 250 pA), Soliton (V_b = −0.10 V, I_t = 70 pA), LUMO (V_b = 0.60 V, I_t = 250 pA), and LUMO + 1 (V_b = 1.10 V, I_t = 250 pA); with V_{rms} = 20 mV.

12. Enlarged high resolution spectra of few odd-membered polymer 3.

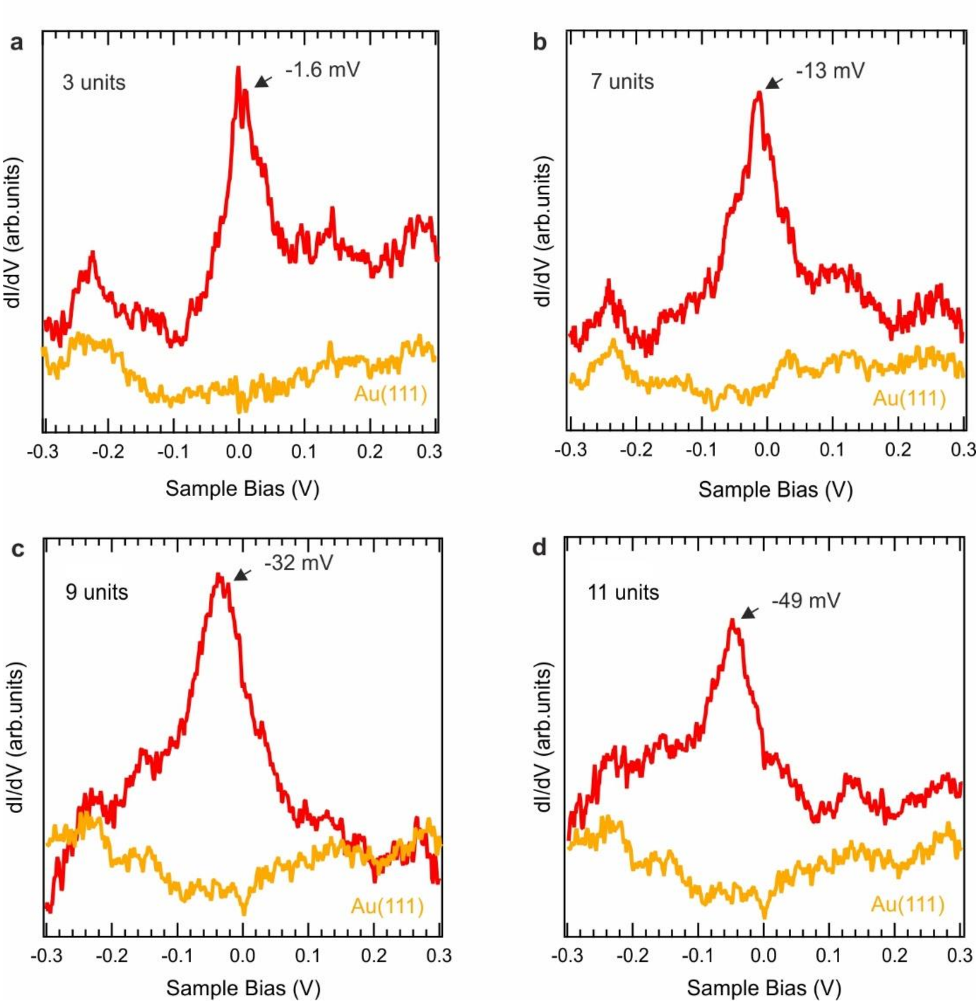


Figure S20 | Enlarged high resolution spectra of a few odd-membered oligo naphthoindenylidene polymers. a-d, d*I*/d*V* spectra of selected polymers of different lengths. Set points for STS: $V_b = -0.3$ V, $I_t = 500$ pA, $V_{rms} = 2$ mV. The position of the soliton peak, which is located close to the Fermi energy, is denoted by an arrow in the graphs.

13. Experimental d^2*I*/d*V*^{2} point spectroscopy using nickelocene (NiCp2)-functionalized tips on a 9-membered and 8-membered polymer 3.

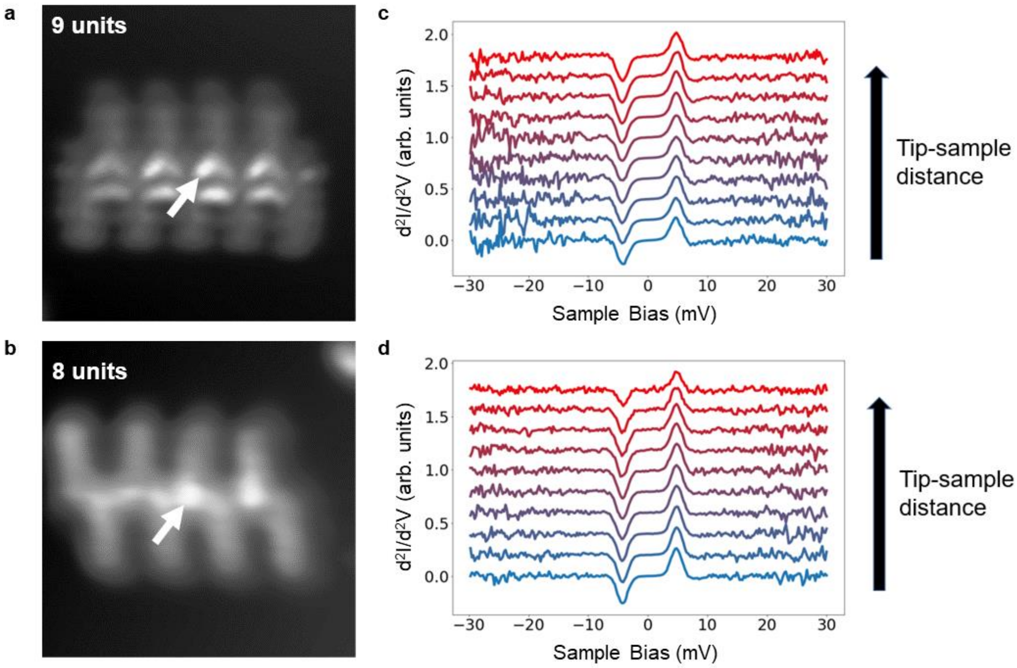


Figure S21 | Experimental d^2*I*/d*V*^{2} point spectroscopy using nickelocene (NiCp2)-functionalized tips of a 9-membered and 8-membered naphthoindenylidene polymer. a,b, Constant-height ultra-high resolution STM images of a 9-membered and 8-membered naphthoindenylidene polymer, respectively. **c**,**d**, A series of d^2*I*/d*V*^{2} spectra was obtained using a scanning-probe tip terminated with NiCp2 on a 9-membered and 8-membered naphthoindenylidene polymer at different tip–sample distances. To take the spectra the tip was brought closer to the sample about 1.5-2 Å from the minima of the AFM frequency shift profile with the NiCp2 profile. Tip-sample distance was increased in steps of 20 pm. The white arrows indicate the position of the acquired point spectroscopy in **a**,**b**. The spectra remain unaltered as the tip-sample distance varies for both odd- and even-membered polymers, showing absence of the magnetic exchange interaction between nickelocene and polymer[8]. Therefore, in the case of 9-membered polymer, it indicates that a radical soliton state (S=1/2) is quenched by charge transfer between the polymer and substrate.

14. Theoretical modelling of the adsorption geometry and electronic structure of an odd-membered polymer 3o on Au(111).

To provide more insight into the interaction between the even-membered polymer and the metallic Au(111) substrate, we carried out the total energy DFT calculations as well as a single impurity Anderson model describing charge transfer between a localized singly occupied soliton state and wide metallic surface band. In particular, we will demonstrate that the absence of a radical soliton state in the odd-membered polymers on Au(111) surface is caused by charge transfer from the molecular soliton state to the metallic substrate inducing the mixed-valence regime.

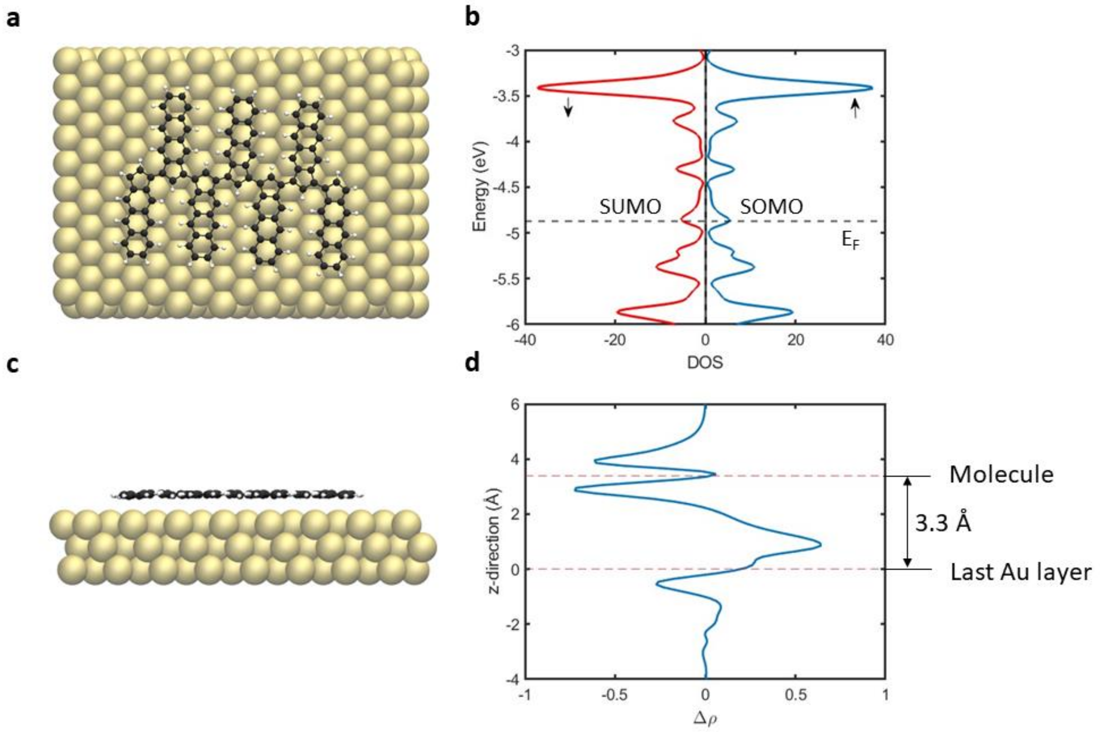


Figure S22. Theoretical modelling of the adsorption geometry and electronic structure of an odd-membered polymer 3o on Au(111). a, Molecule relaxed on the surface from top view. **b**, Density of States projected on Carbon atoms for the polymer on the surface, spin resolved. SOMO and SUMO levels are indicated as well as Fermi level of the Au(111) surface (dashed line). **c**, Molecule relaxed on surface from sideview. **d**, Charge density transfer integrated and projected on z-direction showing substantial charge tranfer .

We have performed the total energy DFT calculations using GGA-PBE (Perdew-Burke-Ernzerhof) exchange-correlation functional[6], including scalar relativistic corrections to the kinetic energy of all atoms at the zeroth-order regular approximation with the packet AIMS[5] for a seven-membered oligo-naphthoindenylidene polymer on Au(111) surface. The metallic Au(111) surface was presented by a periodic slab consisting of three Au layers with 480 atoms, see Figure S22a and S22c, where only the last Au layer was kept fixed. The atomic structure was fully optimized until the residual forces were below the threshold of 10^{-3} eV/ Å. The ground state is obtained under self-consistency criteria at differences below 10^{-6} eV for the total energy and 10^{-3} eV for the sum of the Kohn-Sham eigenvalues. The Brillouin zone was sampled with the Γ-point only. According to the total energy DFT calculations, the polymer adopts a planar configuration with the optimal height of approximately 3.3 Å above the Au(111) surface, see Figure S22c.

Figure S22b shows the density of states (DOS) projected on carbon atoms, while Figure S22d displays an analysis of charge transfer across molecule/metal interface. The position of SOMO/SUMO frontiers orbitals is indicated together with the Fermi energy of the surface (dashed line). We observe that SOMO and SUMO orbitals become degenerated and pinned at the Fermi level, which causes the mixed valence regime. Consequently, the presence of the valence regime suppresses the net spin of the soliton state.

In addition, we carried out a numerical analysis of the single impurity Anderson model (SIAM) to understand the emergence of the mixed-valence regime accompanied by the suppression of the net spin related to the soliton orbital of even-membered polymers.

In the model, we consider a single molecular orbital (SOMO) representing the soliton state coupled with an effective chain comprising 5 connecting sites. This model will be represented by a Hamiltonian of the form

$$\hat{H} = \hat{H}_{imp} + \hat{H}_{chain} + \hat{H}_{int} .$$

Here the impurity is nothing but, $\hat{H}_{imp} = U\, \hat{d}^{\dagger}{}_{\uparrow}\hat{d}^{\dagger}{}_{\downarrow}\hat{d}_{\downarrow}\hat{d}_{\uparrow} + \Delta \sum_{\sigma} \hat{d}^{\dagger}{}_{\sigma}\hat{d}_{\sigma}$, with U being the coulombic repulsion in the SOMO orbital, $U \sim 2.5\, eV$ in our case. $\hat{d}^{\dagger}{}_{\sigma}$ and $\hat{d}_{\sigma}$ are the creation and annihilation operators in the level with spin σ, and Δ represents the energy level of the impurity. The other terms read,

$$\hat{H}_{chain} = -T \sum_{m=1,\sigma}^{N-1} \hat{c}^{\dagger}{}_{m+1,\sigma}\hat{c}_{m,\sigma} + \varepsilon \sum_{m=1,\sigma}^{N} \hat{c}^{\dagger}{}_{m,\sigma}\hat{c}_{m,\sigma} ,$$

$$\hat{H}_{int} = -t \sum_{\sigma} \hat{c}^{\dagger}{}_{1,\sigma}\, \hat{d}_{\sigma} + h.c.$$

Hamiltonian $\hat{H}_{chain}$ represents the chain by a one-electron Hamiltonian parametrized by a hopping between sites T and an on-site energy ε for each site. This last term characterizes the Fermi energy of the metal. For simplicity, we shall consider it equal to zero so that the energy distance between the impurity level and metal will be determined by variable Δ. Hamiltonian $\hat{H}_{int}$ describes the interaction of the impurity level with the chain. It is represented by a hopping *t* between the impurity level and the first site of the chain, as illustrated in Figure S23.

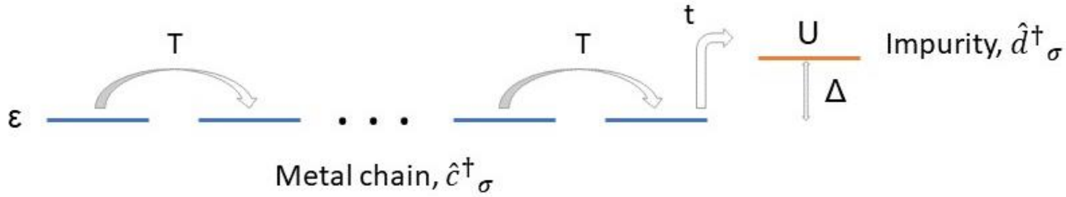


Figure S23 | Scheme of the singly impurity Anderson model. It is depicted the coupling between the metal, represented by a linear chain, and the impurity level with an on-site repulsion U at energy Δ respect to the chain.

Next, we perform a parametric study of this model to account for the different regimes solving in each situation the total Hamiltonian by exact diagonalization. For the sake of simplicity, we will consider a fixed inter-site hopping *T* = 0.2 eV. This value has shown to be good enough to describe a metallic behavior for the finite chain.

Figure S24 displays a variation of the occupancy (N) and the net spin (S^2) of the impurity level for its different alignments with respect to the Fermi level of the metallic chain. We varied the parameter Δ in the range of -100 meV to 100 meV, mimicking the movement of the impurity level across the Fermi level. Over this range, we see how different couplings affect the population of the impurity as well as its spin value and how they are perfectly correlated. For t = 0.00, as was expected, there is only one electron at the impurity level, and the spin value corresponds to a pure radical duplet state ($S^2 = 0.75$). Switching on the interaction t, the population starts to change. When the impurity level is below the Fermi energy, the mixing increases with the coupling t, and it becomes stronger the closer the impurity level approaches the Fermi level of the metallic surface, i.e. $\Delta \rightarrow 0$. Also, a change in t is much more significant close to the limit $\Delta = 0$ rather than far from the Fermi level. Once the impurity level is pinned near the Fermi level, charge fluctuations in/out of the metal keep half occupancy in the impurity, promoting the mixed valence regime. Finally, moving towards positive values Δ, at some point, we can recover the spin as a good quantum number again, even despite relatively large interaction t. However, in this case, $S^2 = 0$ and N = 0, i.e., the impurity discharges.

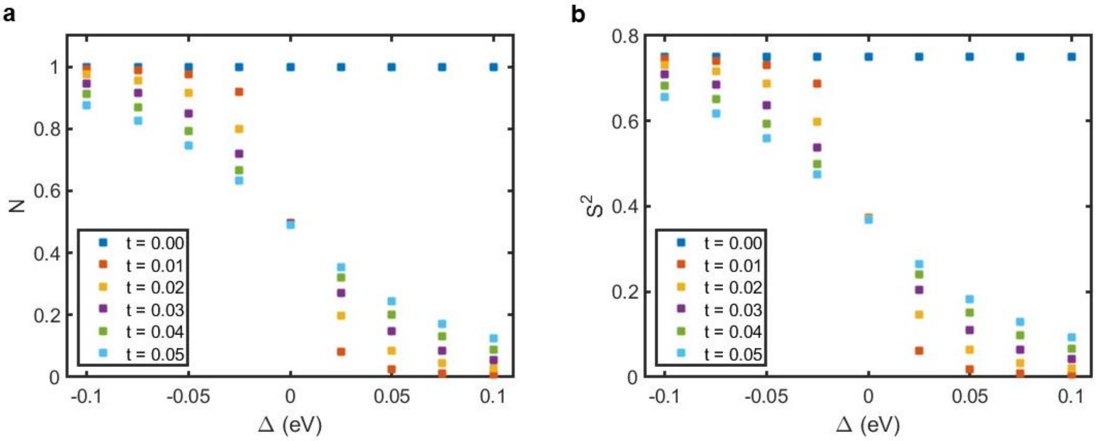


Figure S24 | Variation of occupancy and net spin of the impurity level. Number of electrons, **a**, and spin value **b**, in the molecule level as function of Δ for different couplings t and U = 2.5 eV.

Since the in-gap soliton state observed in the experiment appears near the Fermi level, we computed the DOS for the complete system chain plus an impurity. Figures S25a-c display calculated DOS for small negative values Δ approaching the Fermi level and a weak hybridization t = 0.03eV. Despite the weak coupling t, we observe the emergence of the mixed-valence regime accompanied by the charge transfer (N<1) and loss of radical character ($S^2 < 0.75$) of the soliton state when it approaches the Fermi level. Figures S25d-f show the effect keeping constant a small negative value of Δ and increasing the coupling t, which causes a stronger hybridization of the soliton state with the substrate. Figure S25b represents a situation very similar to what is observed in the experiment, with the impurity spin already lost.

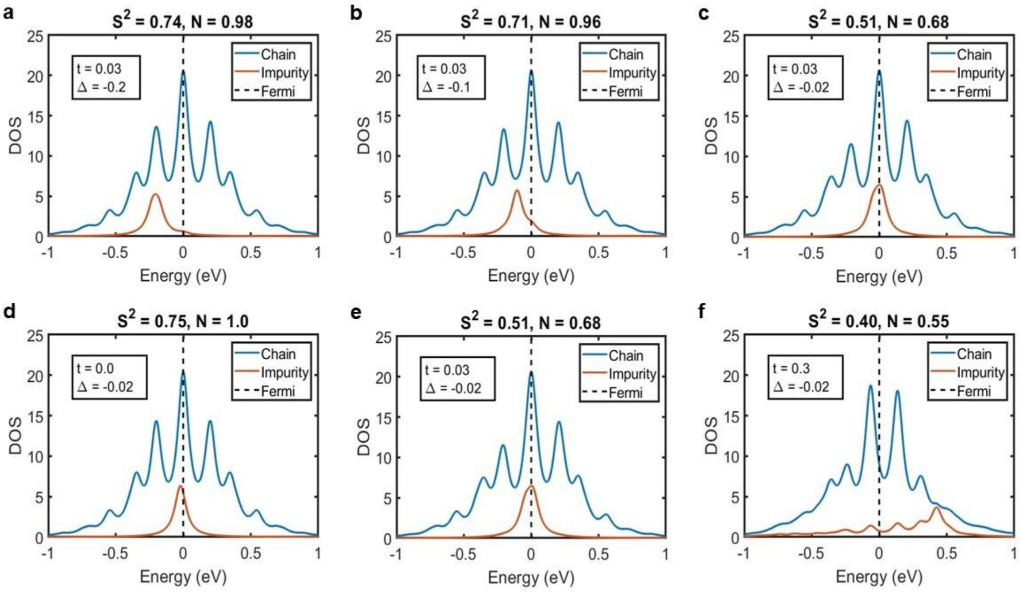


Figure S25 | Calculated Density of States of the impurity model for different alignments to the Fermi energy and interaction with the surface. Panels **a-c** display the calculated density of states (DOS) for different alignments Δ of the impurity level to the Fermi level (vertical dashed line) for constant coupling t = 0.03 eV between the soliton state and substrate as well as the occupancy (N) and the expected spin value (S^2) for the soliton state. The calculations demonstrate that approaching the soliton state close to the Fermi level causes the formation of the mixed-valence regime, accompanied by the charge transfer (N<1) and loss of radical character of the soliton state (S^2<<0.75). Panels **d-f** display how DOS and values N and S^2 change with increasing coupling t between the soliton state and substrate for the alignment Δ = - 0.02 eV of the soliton case to the Fermi level(see panel **c**). We observe that increasing coupling t mixes the soliton state with the electronic states of the substrate accompanied by the broadening of the soliton state (orange line) and charge transfer. Consequently, the radical character of the soliton state is quenched.

15. Spatial distribution of soliton in odd-membered polymers 3o.

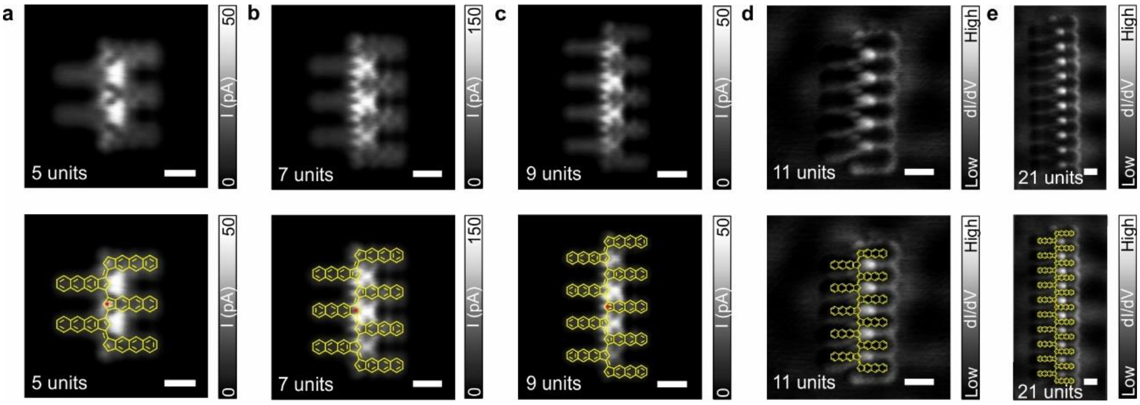
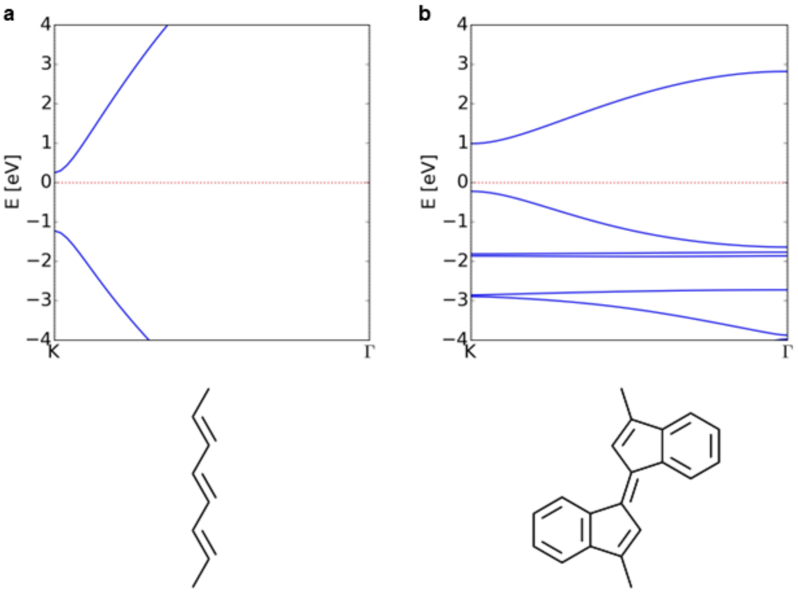


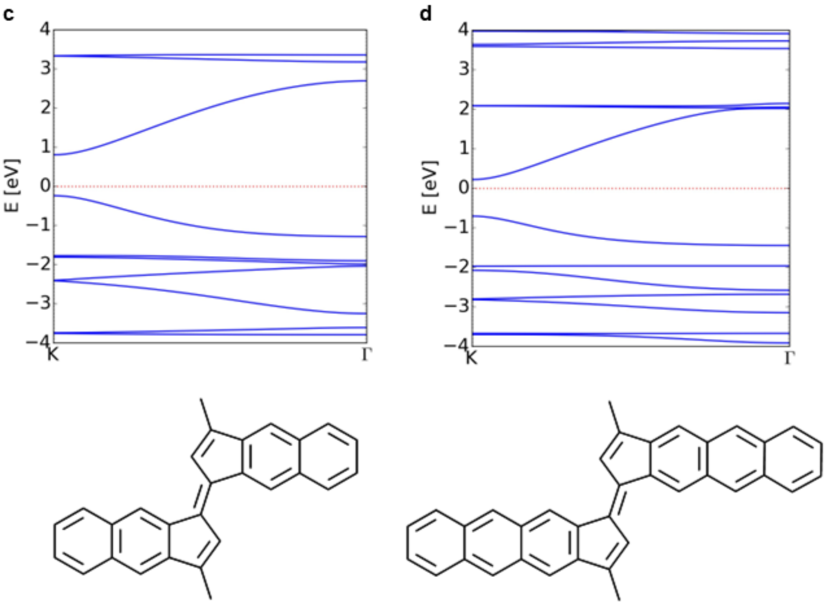
Figure S26 | Spatial distribution of soliton in odd-membered polymers 3o. a-c, High-resolution constant-height STM images of selected odd-membered polymers **3o** obtained using a CO-functionalized tip. Scanning parameters: V_b = 5 mV; I_t = 50 pA; with z_{offset} = 200 pm, 150 pm, 180 pm, for **a-c**, respectively. **d-e**, d*I*/d*V* maps acquired near Fermi for 11-membered polymer (V_b = -50 mV, I_t = 50 pA, V_{rms}= 20 mV) and a 21-membered polymer (V_b = -80 mV, I_t = 70 pA, V_{rms}= 20 mV), respectively. Scale bars = 0.5 nm. In the bottom panels we have superimposed the chemical sketch of the different polymers.

16. Soliton delocalization

Here we discuss the localization of the soliton along the backbone chain for distinct archetype polymers. We compare the soliton localization, including the seminal polyacetylene polymer as well as and different derivatives of bi(acenoindenylidene) polymers with increasing monomer size, as indicated in Figure S27. In particular, we are interested in the dependence of the soliton delocalization as a function of the intrinsic band gap of the infinite polymers.



Polyacetylene chain (**10**)   Bi(indenylidene) polymer (**11**)



Bi(benzoindenylidene) (**12**)   Bi(naphthoindenylidene) (**3**)

Figure S27 | Band structures of polymers featuring polyacetylene longitudinal backbones. **a**, Polyacetylene chain (**10**) (E_g=1.47 eV); **b**, Indenylidene polymer (**11**) (E_g=1.21 eV); **c**, Benzoindenylidene polymer (**12**) (E_g=1.04 eV); **d**, Naphthoindenylidene polymer (**3**) (E_g=0.92 eV). Systems depicted in **b**-**d**, represent the unit cell as two monomers in "opposition", while in **a**, 4 units cells are plot.

Figure S27 displays the band structure of the selected polymers obtained from DFT calculations using FHI-aims code and PBE0 XC-functional[9]. We observe that polyacetylene features the largest band gap of 1.47 eV, while for cyclopentene-based polymers, the band gap decreases with an increasing number of attached benzene units until 0.92 eV in the last case. As shown in reference 10 this gap directly defines the delocalization length of the soliton formed in the finite odd-membered polymers.

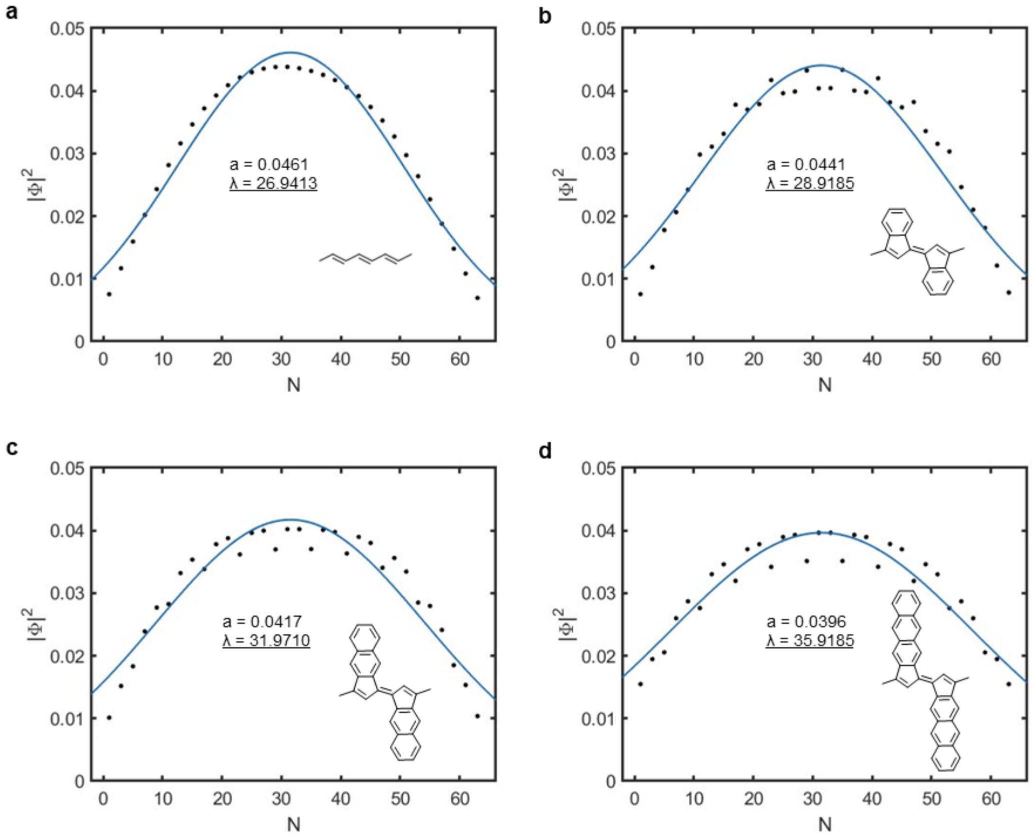


Figure S28 | Decay length of the soliton state for different polymers of increasing monomer size. The polymers consist of 63 carbon atoms in the backbone skeleton for **a**, Polyacetylene (10); **b**, Polymer **11**; **c**, Polymer **12**; **d**, Polymer **3**. Black dots represent values of square wave-function at carbon atoms. Blue line displays the optimal Gaussian fit. The optimal parameters are listed in the graph. The wavefunctions were normalized on the backbone and are localized in the carbon atoms.

Figure S28 illustrates the soliton localization of different polymers under the study consisting of the same number of backbone atoms. To compare directly the soliton localization, we fitted the squared value of the soliton wavefunction to a Gaussian function of the form $f(x) = a\ exp\left(\left(\frac{x-b}{\lambda}\right)^2\right)$, where the parameter λ represents the decay length which describes the delocalization of the soliton. It is worth to mention that the wavefunctions are localized at the carbon sites. Thus, the length of polymers is given in terms of C-C bond units, as well as the resulting localization length in Figure S29.

Figure S29 displays the decay length for the different polymers as a function of the band gap for an infinite system. The graph shows that the magnitude of the band gap is inversely proportional the delocalization length. For large band gaps we observe stronger localization of the soliton state, as predicted by Su et al[10].

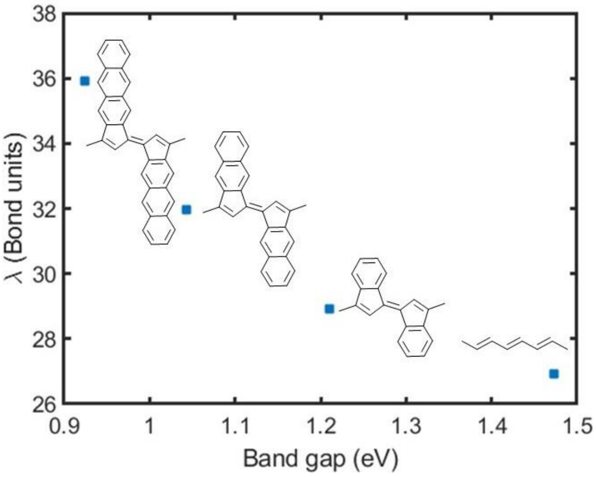


Figure S29 | Decay length of soliton versus energy gap. Graph displaying the decay length of selected polymers shown in Figure S28 as a function of the energy gap in the infinite system for polymers of the same size (63 backbone carbons, i.e. 21 monomers).

17. References

1 Yoshioka, E. & Miyabe, H. Insertion of arynes into the carbon–oxygen double bond of amides and its application into the sequential reactions. *Tetrahedron* **68**, 179-189, (2012).

2 Kelleghan, A. V. *et al.* Safety Assessment of Benzyne Generation from a Silyl Triflate Precursor. *Org. Lett.* **22**, 1665-1669, (2020).

3 Wan, C. *et al.* Benzyne Polyfunctionalization via a Tandem C–C σ-Bond Insertion and Photo-Nazarov Cyclization. *Org. Lett.* **24**, 7276-7281, (2022).

4 Hapala, P. *et al.* Mechanism of high-resolution STM/AFM imaging with functionalized tips. *Phys. Rev. B* **90**, 085421, (2014).

5 Blum, V. *et al.* Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **180**, 2175-2196, (2009).

6 Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **77**, 3865-3868, (1996).

7 Tkatchenko, A. & Scheffler, M. Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. *Phys. Rev. Lett.* **102**, 073005, (2009).

8 Verlhac, B. *et al.* Atomic-scale spin sensing with a single molecule at the apex of a scanning tunneling microscope. *Science* **366**, 623-627, (2019).

9 Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982-9985, (1996).

10 Su, W. P., Schrieffer, J. R. & Heeger, A. J. Soliton excitations in polyacetylene. *Phys. Rev. B* **22**, 2099-2111, 99 (1980).