Supplementary Information:

Covalent C–N Bond Formation Through a Surface Catalyzed Thermal Cyclodehydrogenation

Ilya Piskun,[†,#] Raymond Blackwell,[†,#] Joaquim Jornet-Somoza,[‡,◊] Fangzhou Zhao,[⊥,§] Angel Rubio,[‡,ʃ] Steven G. Louie,[⊥,§] Felix R. Fischer[†,§,¶,*]

[†]Department of Chemistry, University of California, Berkeley, CA 94720, U.S.A.

[‡]Nano-Bio Spectroscopy Group and ETSF, Universidad del País Vasco UPV/EHU, Avenida de Tolosa 72, E-20018 Donostia, Spain

[◊]Max Planck Institute for the Structure and Dynamics of Matter, Bldg. 99, Luruper Chaussee 149, 22761 Hamburg, Germany

[⊥]Department of Physics, University of California, Berkeley, CA 94720, U.S.A.

[ʃ]Center for Computational Quantum Physics (CCQ), The Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, U.S.A.

[§]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, U.S.A.

[¶]Kavli Energy NanoSciences Institute at the University of California Berkeley and the Lawrence Berkeley National Laboratory, Berkeley, California 94720, U.S.A.

Geometry optimization and activation energy barrier calculations

Geometry optimizations for reactant, intermediates and product species of **2a** molecule are computed using the all-electron FHI-aims code[1] at the PBE+VdW level and zeroth order regular approximation (ZORA) scalar relativistic correction. We used van der Waals scheme proposed by Tkatchenko *et al.*.[2] The selected Au(111) surface is modeled using ASE code[3] with four layers of 7 by 7 Au atoms and a *fcc* unit cell parameter of 2.885 Å. The final model consists of 15762 electrons distributed over 260 atoms (64 atoms from the adsorbed molecule **2a** and 196 Au atoms) in periodic lattice with a 25 Å vacuum spacing between the Au surface. We use standard basis set recommended by the FHI-aims authors: a tight defaults for C, N, and H atoms, and light defaults set for the Au atoms. This basis set produces 9804 Kohn-Sham orbitals for each spin component. The system geometries are relaxed until reaching a minimum force threshold of 0.01 eV $Å^{-2}$, keeping the unit cell fixed. During the geometry optimization and transition state search of the two inner layers are frozen in order to reduce the computational cost and simulate the bulk behavior of the surface. The minimum energy path (MEP) for each of the proposed mechanisms are obtained using the string method (SM)[4,5] as implemented in the *aimsChain* code provided by FHI-Aims. For each reaction mechanism, the initial and final structures are interpolated to generate six intermediates, and the MEP is reconstructed until reaching a residual force smaller than 0.2 eV $Å^{-2}$, which is sufficient for energy barrier evaluation.
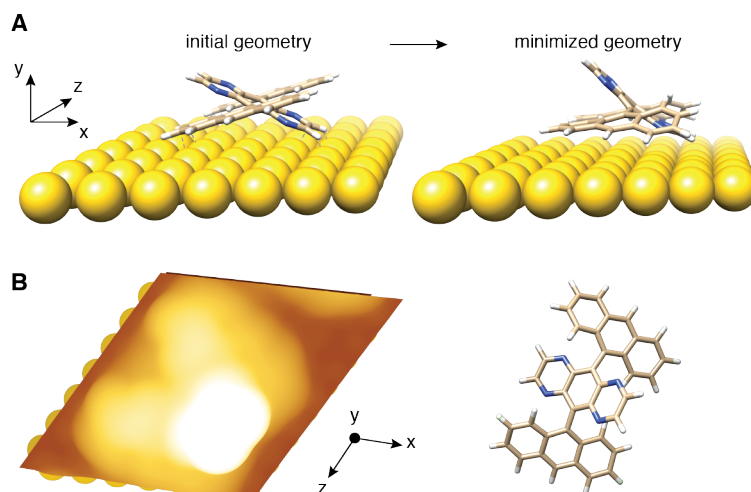
Figure S1. Optimization of adsorption geometry of **2a**. (A) Comparison of starting geometry (left) and final geometry (right) for **2a**. Starting geometry is placed 3.5 Å from the surface, central torsion angle between anthracene and pyrazino[2,3-*g*]quinoxaline moieties is 40°, four Au layers are used in calculation. (B) Simulated STM image of **2a**.
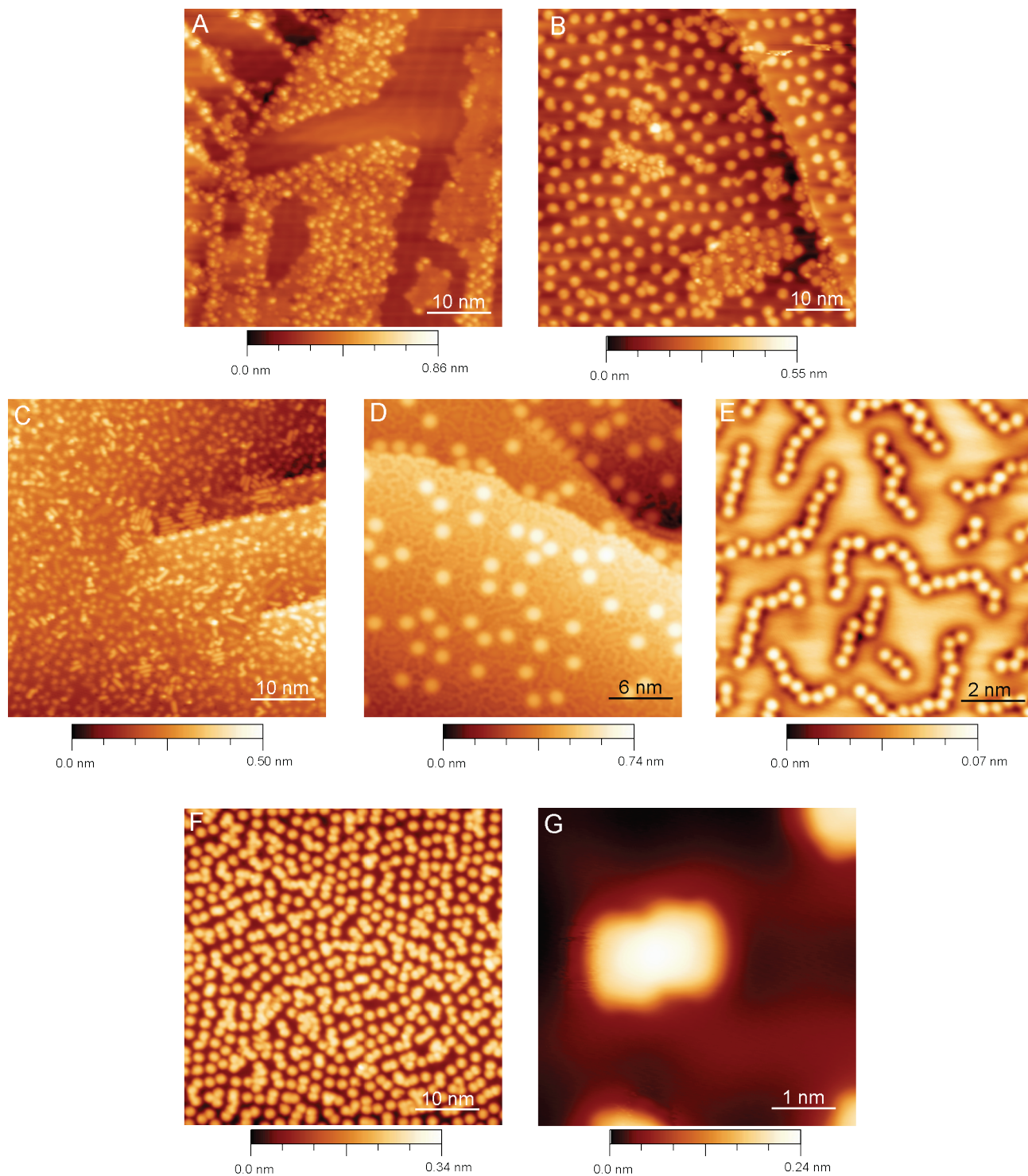
Figure S2. SPM topographic images of **2b,c** on Au(111) and Ag(111). (A) SPM topographic image of a high-coverage deposition of **2c** on a Au(111) surface. (B) SPM topographic image obtained after annealing sample in (A) to 100 °C. The predominant molecular species on the surface is the cyclized product **1**. The closely packed spherical structure below the layer of molecules is atomic iodine adsorbed to the Au(111) surface. (C) SPM topographic image of a high-coverage deposition of **2b** on a Ag(111) surface. (D) SPM topographic image

obtained after annealing sample in (C) to 100 °C. The predominant molecular species on the surface is the cyclized product **1**. The closely packed spherical structure below the layer of molecules is atomic bromine adsorbed to the Ag(111) surface (panel E provides zoom-in of underlying bromine). (F) SPM topographic image of a high-coverage deposition of **2b** on a Au(111) surface followed by annealing to 250 °C. The predominant molecular species on the surface is the cyclized product **1**. The few fused species observed on the surface can be attributed to random C–H bond dissociation processes that lead the formation of dimers and trimers. No structures that extend via the halogenated positions can be identified on the surface observed. (G) SPM topographic image of a representative dimer structure that can be found upon annealing of **2b** on Au(111) to 250 °C, which is attributed to random lateral fusion of **1**. ((A)–(G)): $V_s = 50$ mV, $I_t = 20$ pA)

DFT and GW calculations and discussion of STS data

The electronic structure of **1** in the gas phase was calculated using density functional theory (DFT) for the Kohn-Sham orbital energies within the local density approximation (LDA), and the GW approximation for the physical quasiparticle excitations which incorporates the electron self-energy in a many-electron Green function approach. [6,7] The HOMO-LUMO gap in the LDA level is 1.3 eV, while the quasiparticle gap between HOMO and LUMO is increased to 4.3 eV (Supporting Information Figure S3). Scanning tunneling spectroscopy (STS) was used to measure the experimental electronic structure of **1** adsorbed on Au(111). A representative differential conductance (dI/dV) point spectra—acquired along the armchair edge near the nitrogen dopant—is displayed in Figure 2a. Here, we define the orbital energies by the peak positions in STS point spectra. Figure 2a shows two clear features at –0.6 eV for the occupied states and 1.6 eV for the unoccupied states when compared to an Au(111) reference. We used dI/dV mapping to determine the spatial distribution of the local density of states (LDOS) at these energies and compared the experimental maps to the calculated DOS in order to assign these peaks (Figure 2B–C). The calculated energy difference between LUMO and LUMO+1 state in LDA is only 80 meV and they are even closer (50 meV) in GW calculation, so we assign the shoulder in Figure 2A at ~1.6 eV as a superposition of the LUMO/LUMO+1 states (Figure 2C). The dI/dV map measured at 1.6 eV matches well with the calculated superposition of LDOS of these two states. The wavefunction of HOMO–1 has the *s*-like symmetry and the wavefunction of HOMO has s-like symmetry along long axis and it has less nodal planes, so HOMO and HOMO-1 tend to have larger overlap and have stronger hybridization with the gold substrate wavefunction which is also *s*-like. The stronger hybridization is expected to strongly broaden the HOMO and HOMO–1 states which make them harder to resolve in STS. The wavefunction of HOMO-2 has the *p*-like symmetry and many nodal planes, so the hybridization with surface is weaker. The measured dI/dV map at –0.6 eV matches well with the calculated LDOS of HOMO–2 in the gas phase, and the measurement still shows a broadening of ~0.2 eV for this state from the interaction with the gold substrate. The energy difference between the LUMO and HOMO–2 in the gas phase calculation in LDA is ~1.7 eV, which is smaller than the measured 2.2 eV. In the gas phase GW calculation the quasiparticle energy difference is enlarged to ~4.8 eV because of the incorporation of electron self-energy. However, the gold substrate that the molecule is adsorbed on increases the screening of the Coulomb interaction between electrons to a large extent compared to that of the gas phase.[8] Previous studies have shown the reduction of quasiparticle energy gap of benzene molecules adsorbed on metallic surface such graphite and gold by over 30% (~3.2 eV) comparing to the quasiparticle energy gap in the gas phase,[8-9] so we expect a drastic decrease of the quasiparticle energy difference between the LUMO and HOMO–2 when we fully consider the screening effect of the gold substrate. The decrease in energy gap is estimated using an image charge model to be 2.4 eV with an estimation that the molecule is ~3.0 Å above the substrate. Thus, the energy difference between the two measured states compared to that of GW calculation is due to the screening from the gold substrate.
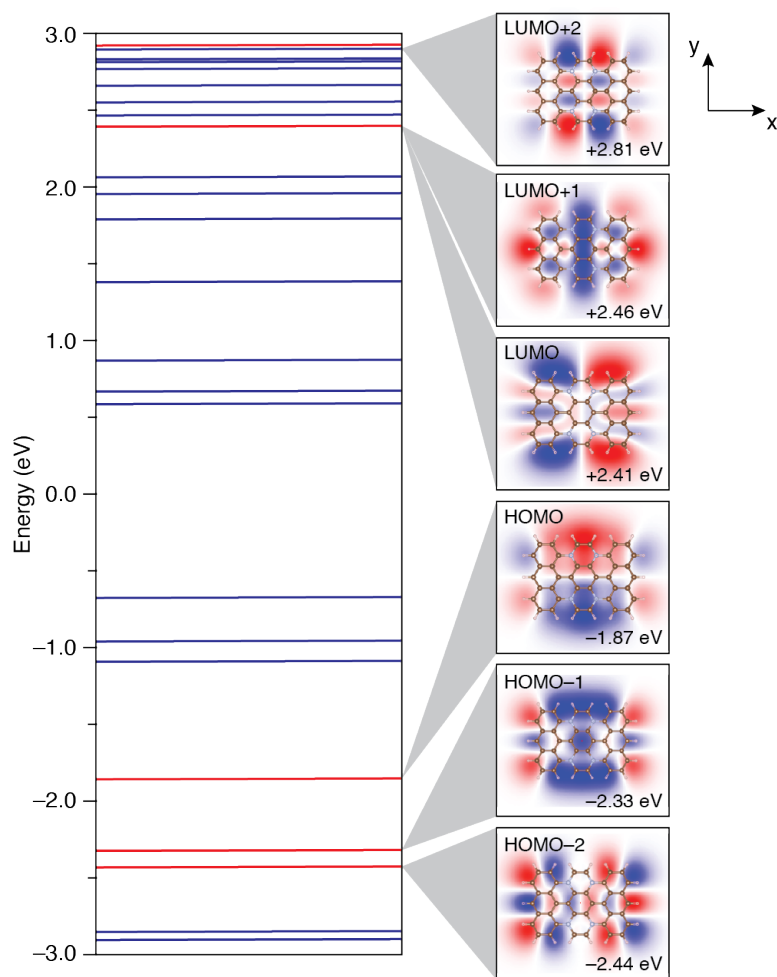
Figure S3. Gas phase DFT (blue) and GW (red) energy levels and the orbital wavefunctions of **1**. Red/blue color in the wavefunction plot show the plus/minus sign of the wavefunction. The values of energy levels marked below the wavefunction plots are the GW energy levels, referred to an arbitrary zero of energy in the middle of the HOMO-LUMO gap.
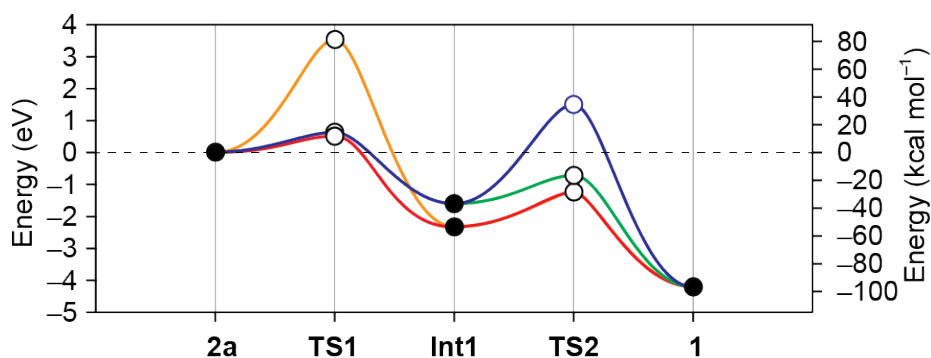
Figure S4. Energy diagram comparing all four plausible reaction mechanisms that lead from **2a** to **1**. Legend: orange - Figure S5 ($\Delta E_1^{\ddagger}$ = 3.524 eV), blue - Figure S6 (($\Delta E_1^{\ddagger}$ = 0.66 eV, $\Delta E_2^{\ddagger}$ = 3.097 eV), red - Figure 3, main text ($\Delta E_1^{\ddagger}$= 0.56 eV), green - Figure S7 ($\Delta E_1^{\ddagger}$= 0.66 eV).

Combined view of all four plausible mechanisms reveals two likely (red, green) and two unlikely (orange, blue) mechanistic pathways. In the case of two plausible mechanisms, both are a permutation of two fundamental steps: (*1*) concerted formation of two C–N bonds at the physisorbed site of pyrazino[2,3-g]quinoxaline followed by a 1,2-suprafacial hydrogen atom shift (exemplified in Figure S7, **Int1** to **1**) and (*2*) direct transfer of a *peri*-hydrogen atom combined with approach of the pyrazino[2,3-g]quinoxaline wing from above the plane of anthracenes (exemplified in Figure S7, **2a** to **Int1**). The mechanism in red (Figure 3, main text) is (*1*) followed by (*2*), and the slightly more energetically demanding mechanism (Figure S7) is (*2*) followed by (*1*). We hypothesize that both of these are feasible under the reaction conditions.

The defining feature of (*1*) is that the hydrogen that is adsorbed to the surface during the cyclodehydrogenation is not the original hydrogen from the *peri*-position. The hydrogens of interest are color-coded and can be followed in Figure S7 – note that the original *peri*-hydrogen atom (red) ends up in the final product. In (*2*), the *peri*-hydrogen is adsorbed directly onto the gold surface.

A common high-energy transition state in two unlikely mechanisms (orange, Figure S5 and blue, Figure S6) is (*3*) the direct surface adsorption of hydrogen atom from *peri*-position combined with cyclization of physisorbed side pyrazino[2,3-g]quinoxaline wing (Figure S5, **2a** to **Int1**). Pathway in Figure S5 has an activation barrier $\Delta E^{\ddagger}$ = 3.524 eV, which is too high to be feasible under the experimental annealing conditions (T = 75 – 100 °C, t = 15 min). The pathway in Figure S6 is different from the other three – here, the second cyclization step actually has the highest energy transition state (same one that is found in (*3*)). Here, $\Delta E_2^{\ddagger}$ = 3.097 eV, which is still very large. We have never experimentally observed the intermediate where only one side has cyclized, which would be reasonable to expect from the potential well that results along this pathway; thus this mechanism can also be excluded.
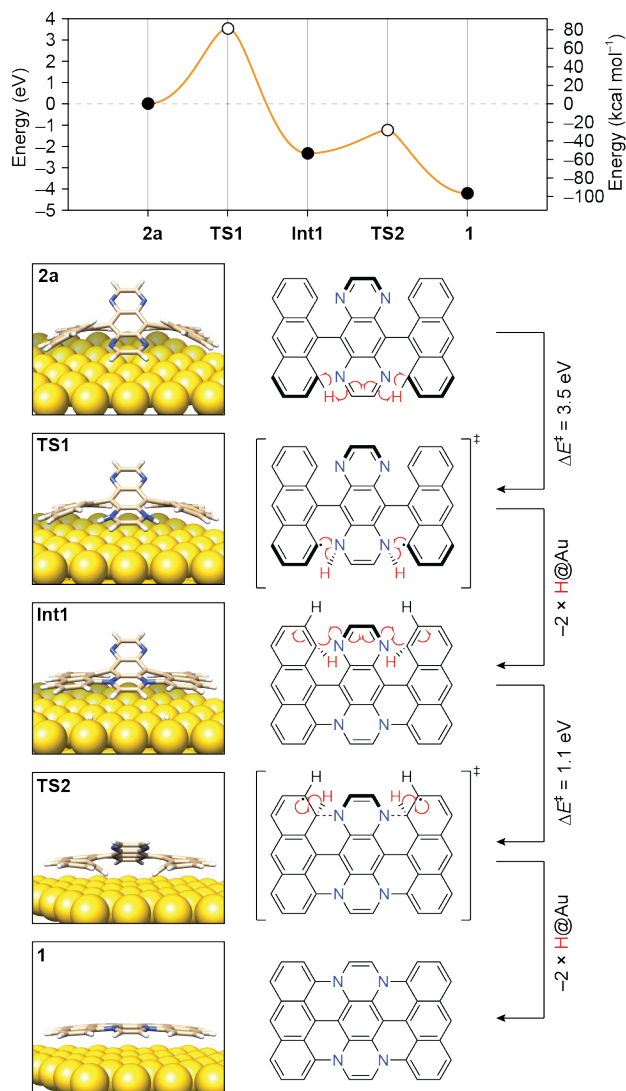
Figure S5. Calculated energy diagram for the stepwise cyclodehydrogenation of **2a** to **1** – (*3*) followed by (*1*). The graph shows the *ab-initio* energy landscape for intermediates and transition states along the reaction pathway from **2a** to **1**. Calculated activation energies are shown next to the reaction arrows.

Figure S6. Calculated energy diagram for the stepwise cyclodehydrogenation of **2a** to **1** – (*2*) followed by (*3*). The graph shows the *ab-initio* energy landscape for intermediates and transition states along the reaction pathway from **2a** to **1**. Calculated activation energies are shown next to the reaction arrows.
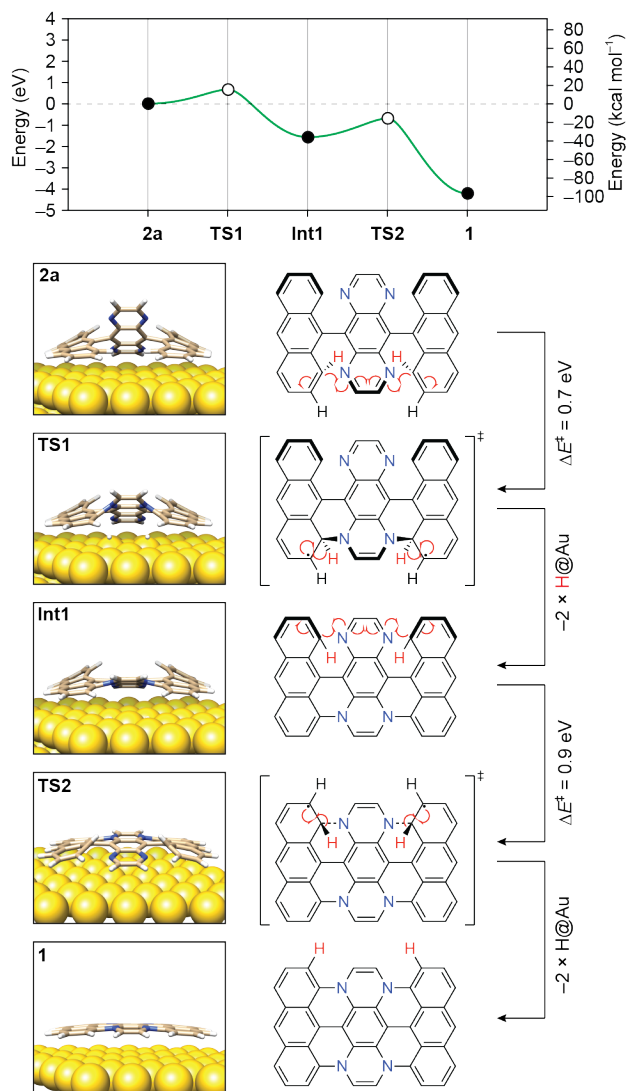
Figure S7. Calculated energy diagram for the stepwise cyclodehydrogenation of **2a** to **1** – (2) followed by (*1*). The graph shows the *ab-initio* energy landscape for intermediates and transition states along the reaction pathway from **2a** to **1**. Calculated activation energies are shown next to the reaction arrows.
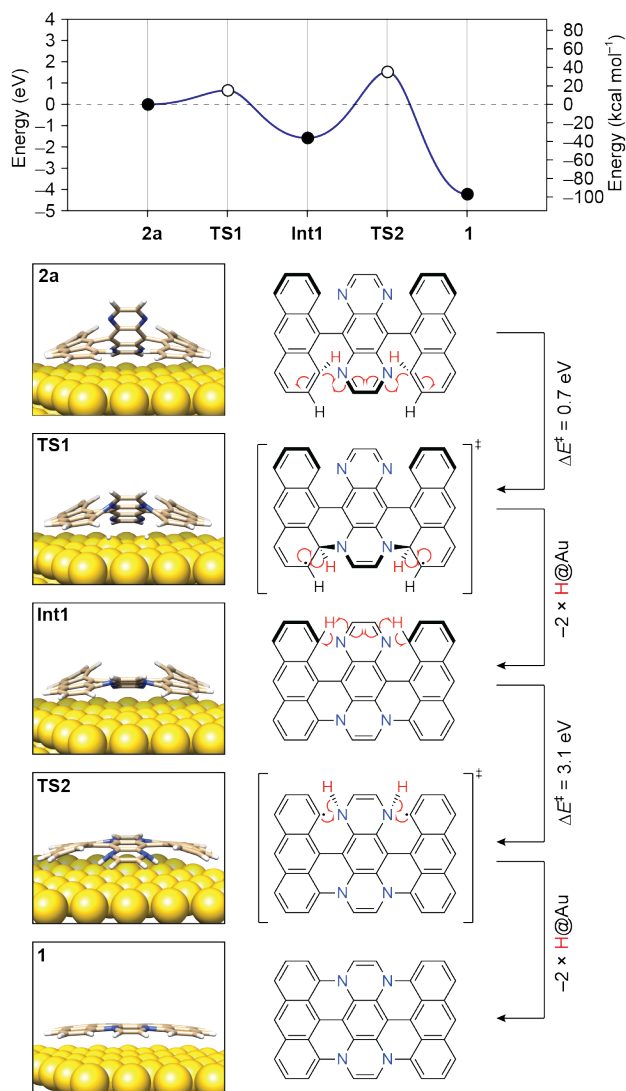
Figure S8. Calculated energy diagram for the concerted (left) and stepwise (right) cyclodehydrogenation step that leads from of **2a** to **Int2**. The graph shows the *ab-initio* energy landscape for intermediates and transition states along the reaction pathway from **2a** to **Int1**. Calculated activation energies are shown next to the reaction arrows.
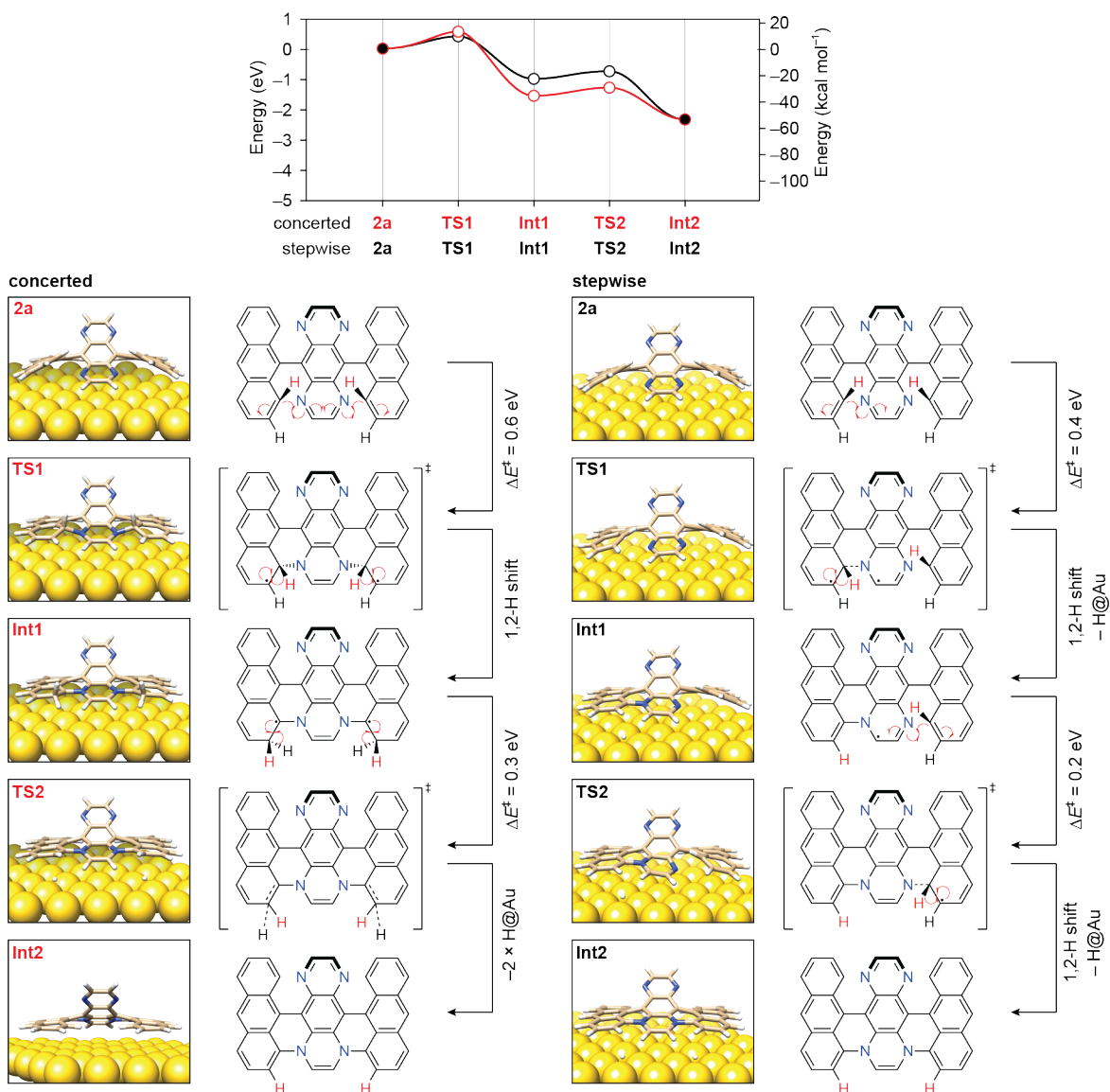
Materials and General Methods

Unless otherwise stated, all manipulations of air and/or moisture sensitive compounds were carried out in oven-dried glassware, under an atmosphere of nitrogen. All solvents and reagents were purchased from Alfa Aesar, Spectrum Chemicals, Acros Organics, TCI America, and Sigma-Aldrich and were used as received unless otherwise noted. Organic solvents were dried by passing through a column of alumina and were degassed by vigorous bubbling of N_2 through the solvent for 20 min. Flash column chromatography was performed on SiliCycle silica gel (particle size 40–63 µm). Thin layer chromatography was carried out using SiliCycle silica gel 60 Å F-254 precoated plates (0.25 mm thick) and visualized by UV absorption. All ^{1}H and ^{13}C NMR spectra were recorded on Bruker AV-600, NEO-500, and AV-500 MHz spectrometers, and are referenced to residual solvent peaks (CDCl$_3$ ^{1}H NMR = 7.26 ppm, ^{13}C NMR = 77.16 ppm; d2-DCM ^{1}H NMR = 5.32 ppm, ^{13}C NMR = 53.84 ppm.) ESI mass spectrometry was performed on a Finnigan LTQFT (Thermo) spectrometer in positive ionization mode. MALDI mass spectrometry was performed on a Voyager-DE PRO (Applied Biosystems Voyager System 6322) in positive mode using a matrix of dithranol or 2,5-dihydroxybenzoic acid.

Synthetic Procedures

Pyrazino[2,3-g]quinoxaline (**4**) A 200 mL round bottom flask with a reflux condenser was charged with 2,5-dihydroxy-1,4-benzoquinone **3** (4.5 g, 32.2 mmol) in H$_2$O (100 mL). The reaction mixture is heated to 95 °C and stirred for 5 min before ethylenediamine (10 mL) is added to the reaction mixture. The reaction mixture is stirred at 95 °C for 1h, cooled rapidly in a 0 °C ice bath, and the precipitate filtered to yield a crude brown solid. A 100 mL round bottom flask with reflux condenser is charged with the crude brown solid (1 g), 10% palladium on carbon (1.5 g) in nitrobenzene (30 mL). The reaction mixture is heated at 213 °C for 21 h. The solution is cooled to 25 ºC, filtered over a medium-fritted glass funnel, and concentrated on a rotary evaporator. Column chromatography (SiO$_2$, EtOAc, R_f = 0.4) yielded pyrazino[2,3-g]quinoxaline (175 mg, 0.96 mmol, 10%) as a colorless solid. ^{1}H NMR (500 MHz, CDCl$_3$) δ = 9.02 (s, 4H), 9.01 (s, 2H) ppm; ^{13}C{^{1}H} NMR (125 MHz, CDCl$_3$) δ = 147.2, 141.6, 130.3 ppm.

5,10-dibromopyrazino[2,3-g]quinoxaline (**5**) A 20 mL glass vial was charged with pyrazino[2,3-g]quinoxaline **4** (175 mg, 0.96 mmol) and *N*-bromosuccinimide (400 mg, 2.24 mmol) under an atmosphere of N_2. Degassed DMF (10 mL) is added, and the reaction mixture is stirred at 60 °C for 16 h. Upon cooling H$_2$O (10 mL) is added, and the reaction mixture is filtered and the precipitate washed with of H$_2$O (100 mL), and cold CH$_2$Cl$_2$ (100 mL) to yield **5** (218 mg, 0.64 mmol, 67%) as a yellow solid. ^{1}H NMR (600 MHz, CDCl$_3$) δ = 9.17 (s, 4H) ppm. The solubility of **5** in common deuterated solvents is insufficient to acquire a ^{13}C{^{1}H} NMR spectrum.

5,10-di(anthracen-9-yl)pyrazino[2,3-g]quinoxaline (**2a**) A 20 mL glass vial was charged with 5,10-dibromopyrazino[2,3-g]quinoxaline **5** (50 mg, 0.147 mmol), 2-(anthracen-9-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (89.5 mg, 0.295 mmol), Pd(PPh$_3$)$_4$ (20 mg, 0.0147 mmol), Cs$_2$CO$_3$ (280 mg, 0.870 mmol), and toluene (8 mL) under an atmosphere of N_2. The vial is sealed and heated to 120 ºC for 48 h. Upon cooling to 25

ºC, the reaction mixture is concentrated on a rotary evaporator. The solid residue is suspended in CH_2Cl_2 and filtered. The filter cake is resuspended in H_2O (15 mL), sonicated, filtered and washed with H_2O (200 mL). The solid is resuspended in CH_2Cl_2, sonicated, filtered and washed with CH_2Cl_2 and EtOAc until the eluent is colorless. Collection of filter cake yielded **2a** (40 mg, 0.076 mmol, 52%) as a red solid. FTMS (MALDI-TOF): m/z: $[C_{38}H_{21}N_4]^+$, calcd. 534.1844; found 534.1730. The solubility of **2a** in common deuterated solvents is insufficient to acquire ^1H or $^{13}C\{^1H\}$ NMR spectra.

5,10-bis(10-bromoanthracen-9-yl)pyrazino[2,3-g]quinoxaline (**2b**) A 4 mL glass vial was charged with 5,10-di(anthracen-9-yl)pyrazino[2,3-*g*]quinoxaline **2a** (20 mg, 0.037 mmol), and *N*-bromosuccinimide (10.8 mg, 0.077 mmol). Trifluoroacetic acid (1 mL) is added, and the reaction mixture is stirred for 90 min. The reaction mixture is poured into cold aqueous 1M NaOH (15 mL), and the precipitate is isolated by filtration. Column chromatography (SiO_2, CH_2Cl_2, R_f = 0.85) yielded **2b** (15 mg, 0.021 mmol, 57%) as an orange solid. ^1H NMR (600 MHz, CD_2Cl_2) δ = 8.75 (d, J = 8.8 Hz, 4H), 8.69 (s, 4H), 7.66 (dd, J = 8.8, 6.5 Hz, 4H), 7.35 (dd, J = 8.7, 6.5 Hz, 4H), 7.26 (d, J = 8.7 Hz, 4H) ppm; $^{13}C\{^1H\}$ NMR (151 MHz, CD_2Cl_2) δ = 147.7, 141.4, 139.0, 132.6, 132.4, 130.7, 128.7, 127.55, 127.4, 126.6, 124.3 ppm; FTMS (ESI-TOF): m/z $[C_{38}H_{21}N_4Br_2]^+$, calcd. $C_{38}H_{21}N_4Br_2$ 692.0040; found 692.0034.

5,10-bis(10-iodoanthracen-9-yl)pyrazino[2,3-g]quinoxaline (**2c**) A 20 mL glass vial was charged with **2b** (30.0 mg, 0.043 mmol), bis(neopentylglycolato)diboron (59 mg, 0.260 mmol), $Pd(OAc)_2$ (2.0 mg, 0.009 mmol), potassium acetate (51 mg 0.52 mmol), and DMF (4 mL) under and atmosphere of N_2. The vial is sealed and heated to 80 ºC for 16 h. The crude mixture was diluted with CH_2Cl_2 (40 mL) and washed with a saturated aqueous NaCl solution. Column chromatography (1:1 to 0:1 hexanes/CH_2Cl_2) yields a crude mixture of mono- and diborylated intermediates that were subjected to the iodination without further purification. A 4 mL glass vial was charged with the crude mixture of mono- and diborylated intermediates (20 mg), CuI (20 mg) and KI (40 mg), in a mixture of THF (1 mL), MeOH (0.2 mL), and H_2O (0.1 mL). The vial is sealed, and the reaction mixture is heated to 80 ºC for 16 h. The crude reaction mixture is concentrated onto Celite on a rotary evaporator. Column chromatography (SiO_2, CH_2Cl_2, R_f = 0.85) yielded **2c** (2.2 mg, 7% yield over 2 steps) as an orange solid. ^1H NMR (600 MHz, $CDCl_3$) δ = 8.76–8.71 (m, 8H), 7.66–7.60 (m, 4H), 7.38–7.32 (m, 4H), 7.19 (d, J = 8.6 Hz, 4H) ppm; $^{13}C\{^1H\}$ NMR (151 MHz, $CDCl_3$) δ = 147.4, 141.1, 139.1, 134.6, 133.7, 132.9, 132.2, 127.6, 127.1, 126.4, 108.3 ppm; FTMS (ESI-TOF): m/z $[C_{38}H_{21}N_4I_2]^+$, calcd. $C_{38}H_{21}N_4I_2$ 786.9862; found 786.9850.
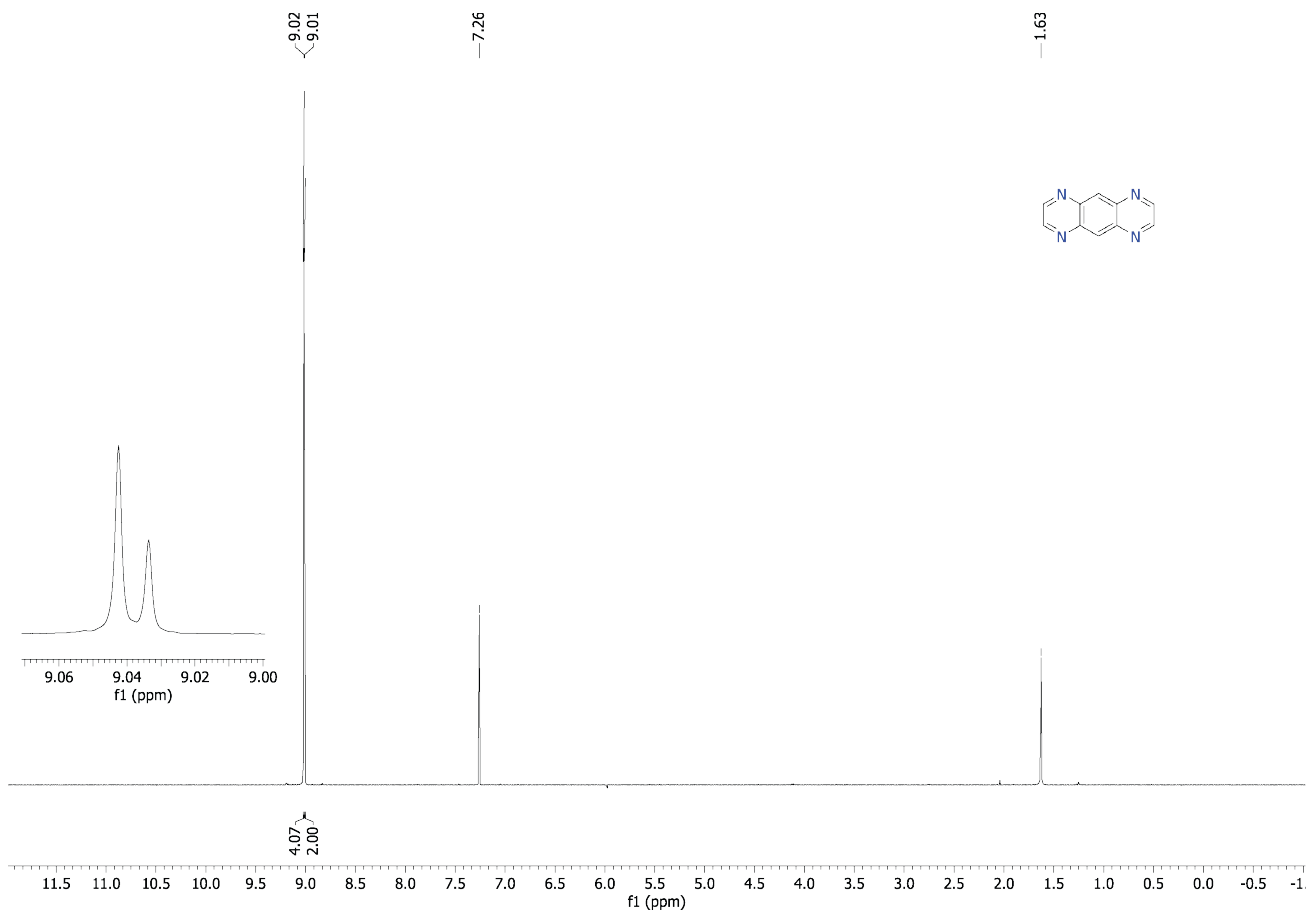
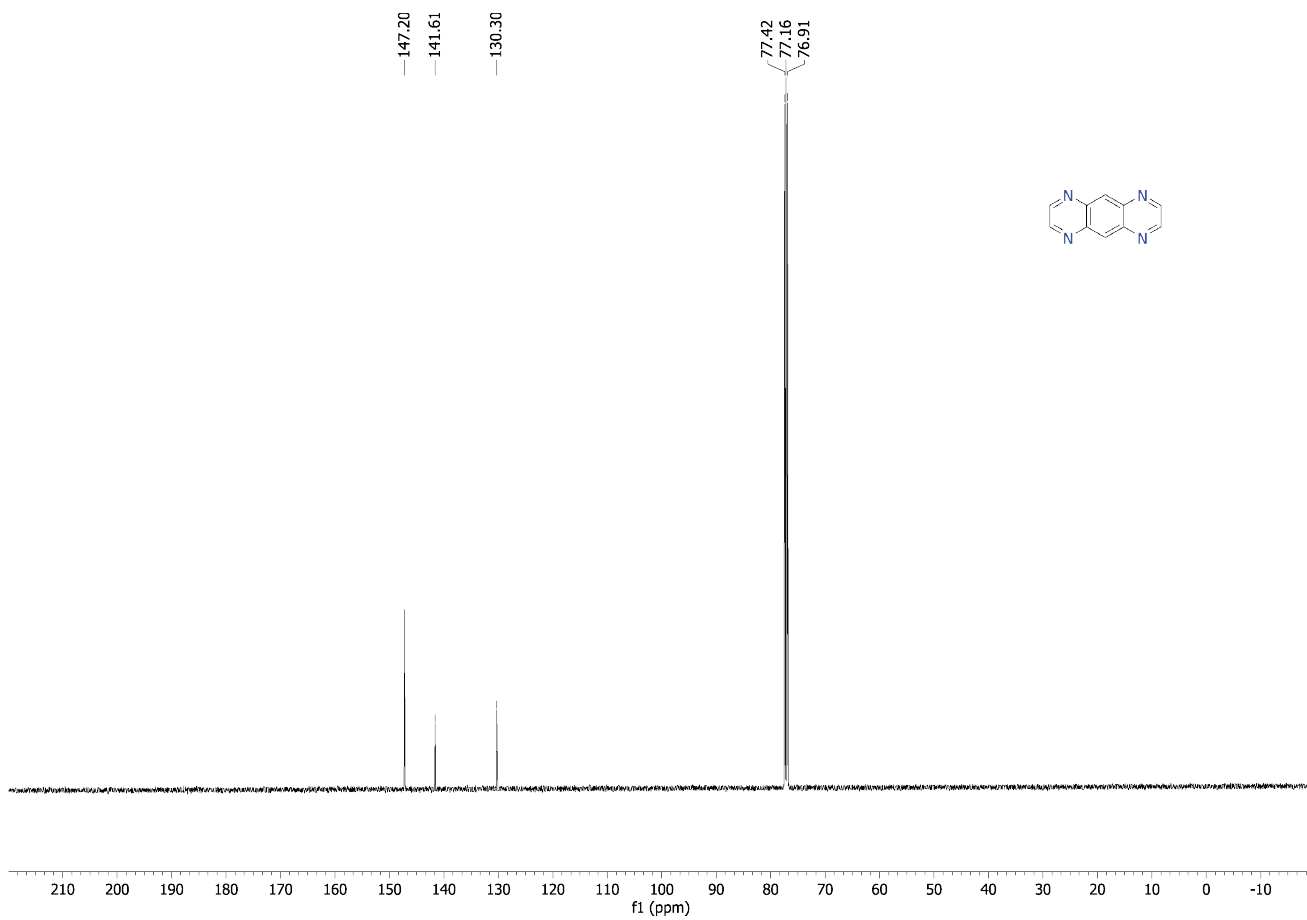**Figure S9.** ^{1}H NMR (500 MHz, CDCl$_3$) spectrum of pyrazino[2,3-*g*]quinoxaline (**4**).

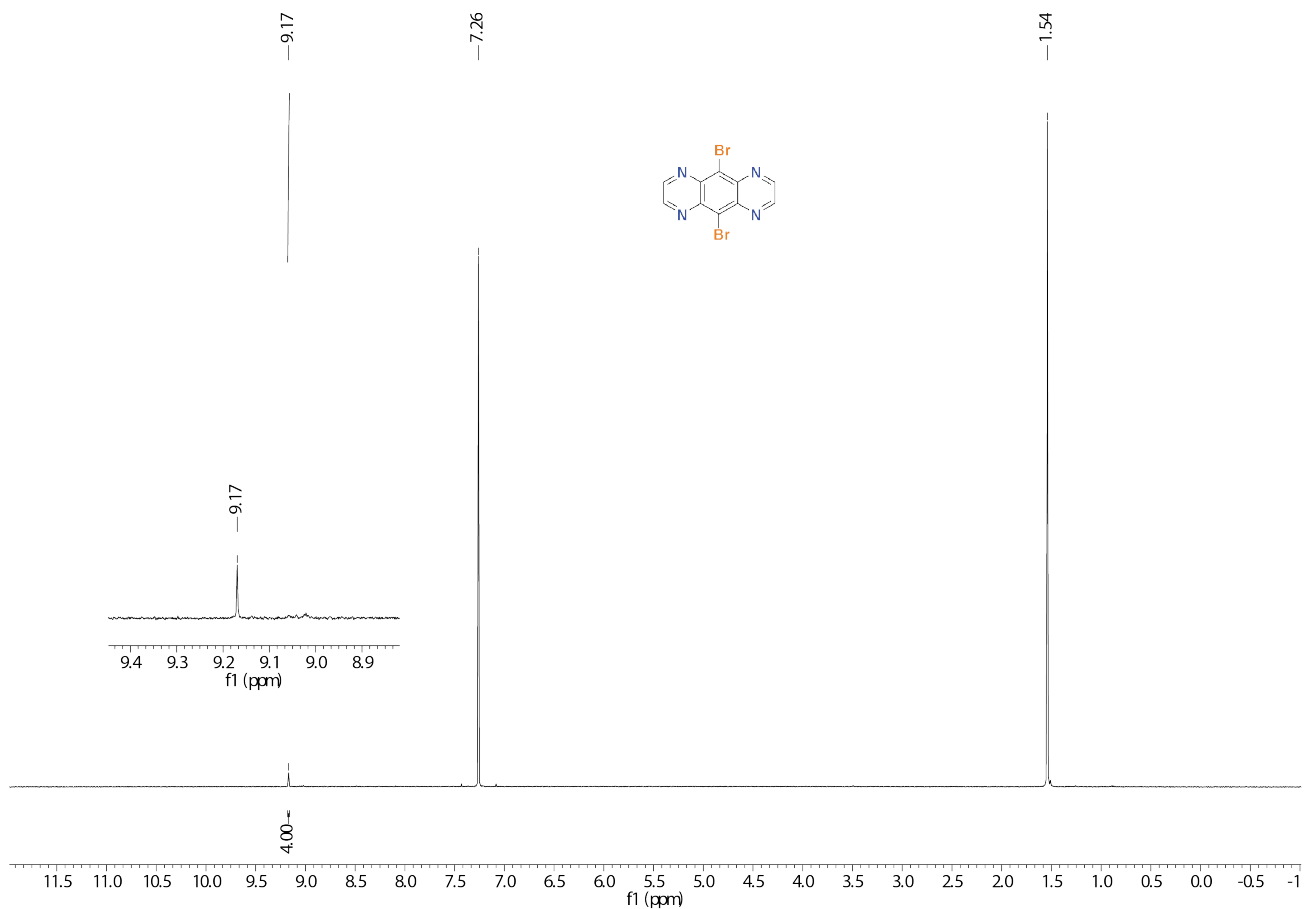**Figure S10.** $^{13}C\{^1H\}$ NMR (125 MHz, CDCl$_3$) spectrum of pyrazino[2,3-*g*]quinoxaline (**4**).

Figure S11. ^{1}H NMR (600 MHz, CDCl$_3$) spectrum of 5,10-dibromopyrazino[2,3-*g*]quinoxaline (**5**).
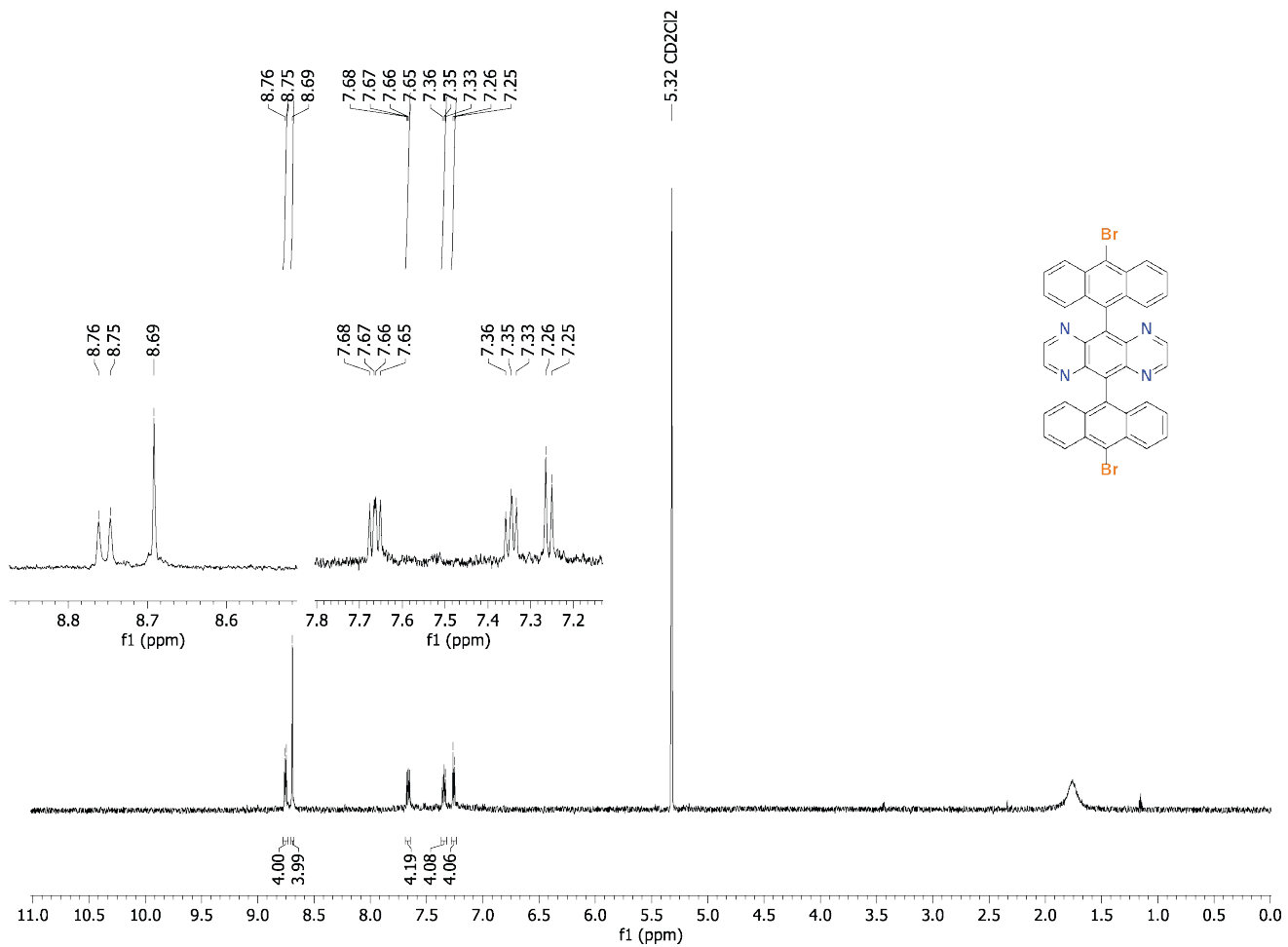
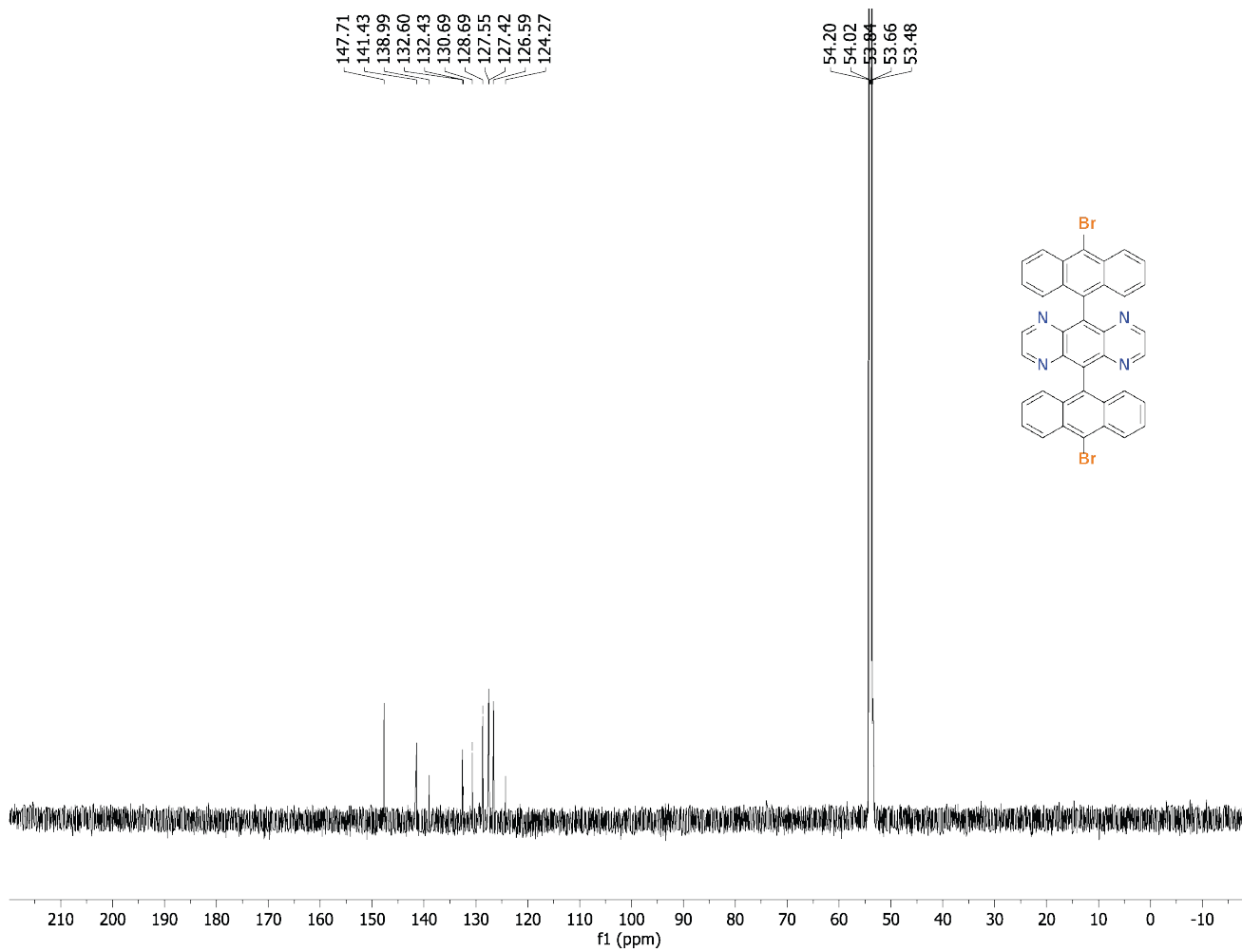**Figure S12.** ^{1}H NMR (600 MHz, CD$_2$Cl$_2$) spectrum of **2b**.

Figure S13. $^{13}C\{^1H\}$ NMR (151 MHz, CD_2Cl_2) spectrum of **2b**.
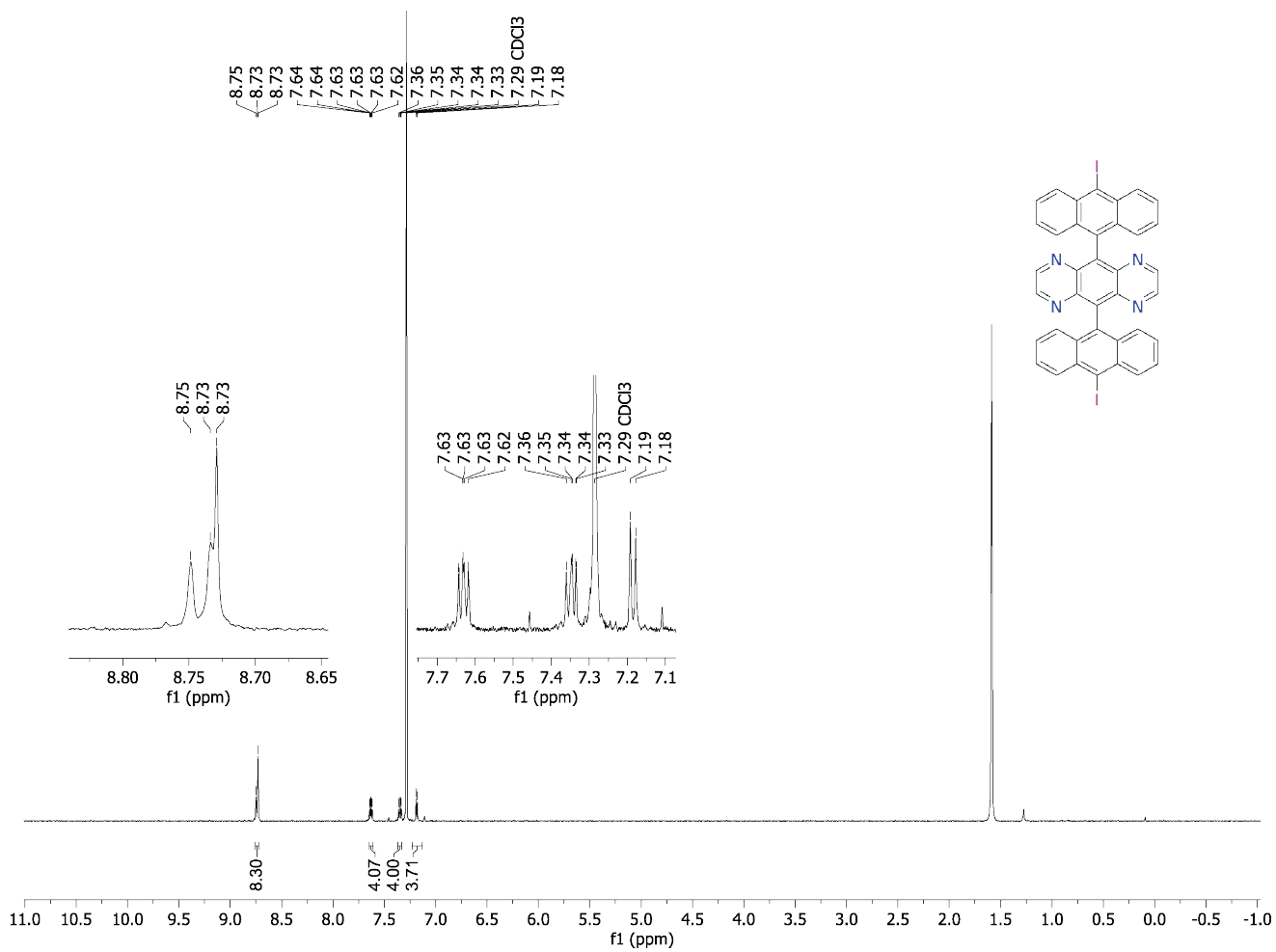
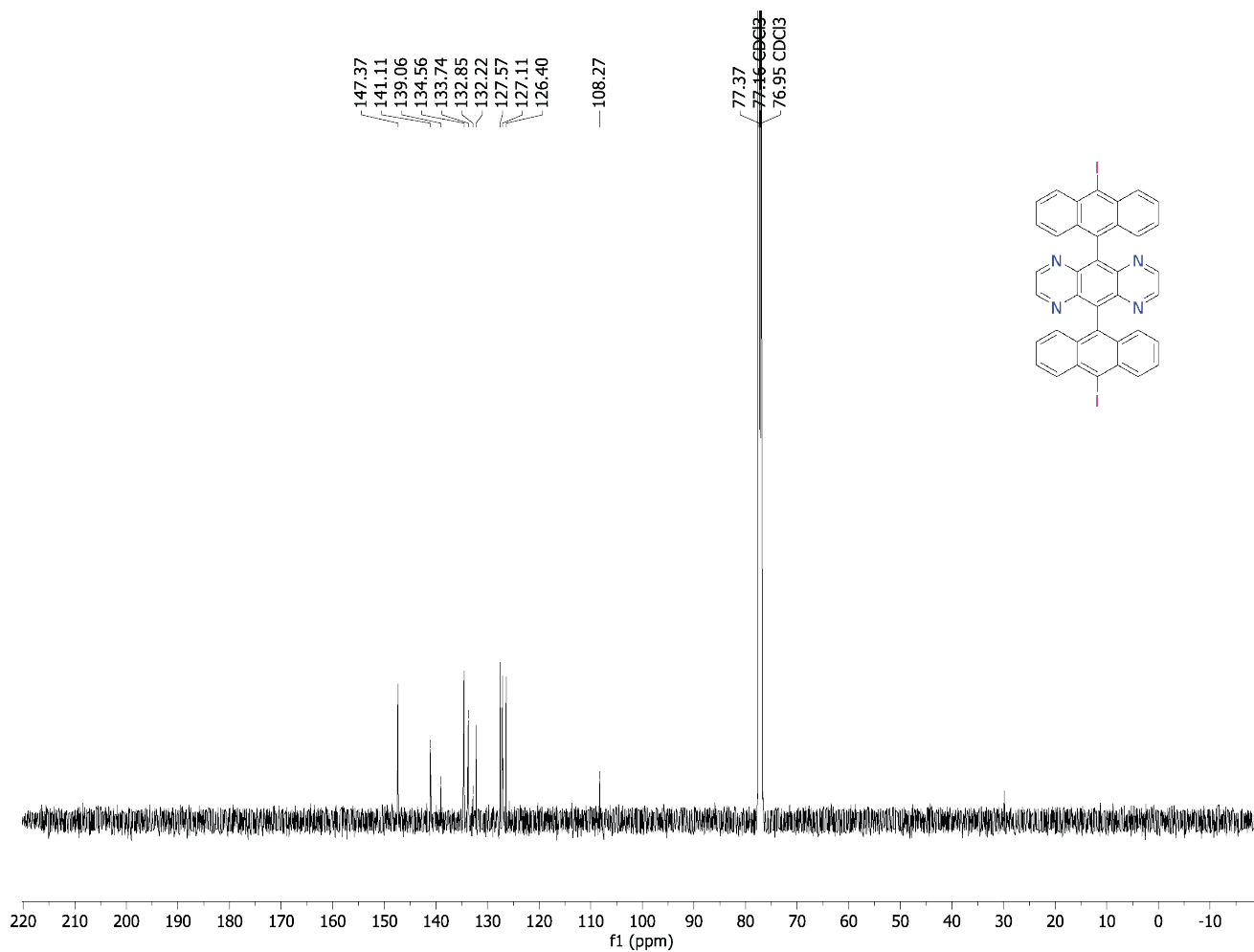**Figure S14.** ^{1}H NMR (600 MHz, CD$_2$Cl$_2$) spectrum of **2c**.

Figure S15. $^{13}C\{^1H\}$ NMR (151 MHz, CD$_2$Cl$_2$) spectrum of **2c**.

References

(1) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab Initio Molecular Simulations with Numeric Atom-Centered Orbitals. *Computer Physics Communications* **2009**, *180*, 2175–2196.

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

(3) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dulak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; Hermes, E. D.; Jennings, P. C.; Jensen, P. B.; Kermode, J.; Kitchin, J. R.; Kolsbjerg, E. L.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J. B.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiotz, J.; Schutt, O.; Strange, M.; Thygesen, K. S.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z. H.; Jacobsen, K. W. The Atomic Simulation Environment – A Python Library for Working with Atoms. *J. Phys-Condens Mat.* **2017**, *29*, 273002.

(4) Weinan, E.; Ren, W. Q.; Vanden-Eijnden, E. String Method for the Study of Rare Events. *Phys. Rev. B* **2002**, *66*, 052301.

(5) E, W. N.; Ren, W. Q.; Vanden-Eijnden, E., Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier-Crossing Events. *J. Chem. Phys.* **2007**, *126*, 164103.

(6) Hybertsen, M. S.; Louie, S. G. Electron Correlation in Semiconductors and Insulators – Band-Gaps and Quasi-Particle Energies. *Phys. Rev. B* **1986**, *34*, 5390–5413.

(7) Deslippe, J.; Samsonidze, G.; Strubbe, D. A.; Jain, M.; Cohen, M. L.; Louie, S. G. BerkeleyGW: A Massively Parallel Computer Package for the Calculation of the Quasiparticle and Optical Properties of Materials and Nanostructures. *Comp. Phys. Commun.* **2012**, *183*, 1269–1289.

(8) Neaton, J. B.; Hybertsen, M. S.; Louie, S. G. Renormalization of Molecular Electronic Levels at Metal-Molecule Interfaces. *Phys. Rev. Lett.* **2006**, *97*, 216405.

(9) Garcia-Lastra, J. M.; Rostgaard, C.; Rubio, A.; Thygesen, K. S. Polarization-Induced Renormalization of Molecular Levels at Metallic and Semiconducting Surfaces. *Phys. Rev. B* **2009**, *80*, 245427.