Electronic Supplementary Information

Crystalline Supramolecular Organic Frameworks via Hydrogen-bonding between Nucleobases

Miguel Martín-Arroyo,[a] Javier Castells-Gil,[b] Nerea Bilbao,[a] Neyvis Almora-Barrios,[b] Carlos Martí-Gastaldo,*[b] David González-Rodríguez*[,a,c]

[a] Departamento de Química Orgánica, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain.

[b] Instituto de Ciencia Molecular, Universitat de Valencia, Catedrático José Beltrán 2, 46980 Paterna, Spain.

[c] Institute for Advanced Research in Chemical Sciences (IAdChem), 28049 Madrid, Spain

Table of contents

SO. Material and Methods	2
S1. NMR studies	15
S2. UV/visible spectroscopy studies	16
S3. Crystallization process	17
S4. SEM images	18
S5. FT-IR analysis	19
S6. Powder XRD analysis	20
S7. Theoretical study	21
S8. Gas absorption measurement	27
S9. Thermogravimetric analysis	28

S0. Materials and methods

General Methods

Starting materials. Chemicals were purchased from commercial suppliers and used without further purification. Solid hygroscopic reagents were dried in a vacuum oven before use. Reaction solvents were thoroughly dried before use using standard methods.

Column chromatography was carried out on silica gel Merck-60 (230-400 mesh, 60 Å), and TLC on aluminum sheets precoated with silica gel 60 F254 (Merck).

NMR spectra were recorded with a BRUKER AVANCE-II (300 MHz) instrument and BRUKER DRX 500 MHz. The temperature was actively controlled at 298 K. Chemical shifts are measured in ppm using the signals of the deuterated solvent as the internal standard [$CDCl_3$ calibrated at 7.26 ppm (^1H) and 76.0 ppm (^{13}C), DMSO-d6 calibrated at 2.50 ppm (^1H) and 39.5 ppm (^{13}C)].

HR-MS spectra were determined on a Waters GCT Agilent Technologies 6890N, MAXIS II, ABSciex QSTAR, Bruker ULTRAFLEX III or Waters VG AutoSpec equipment, with the conditions indicated in each case. The data are in mass units m/z.

FT-IR spectra were recorded with a JASCO Cary 630 FT-IR equipment, equipped with a diamond ATR.

Absorption and fluorescence spectra were recorded with a JASCO V-660 equipment and a JASCO FP-8600 equipment, respectively. The temperature was controlled using a JASCO Peltier thermostatted cell holder with a range of 263–383 K, adjustable temperature slope, and accuracy of ± 0.1 K.

Scanning Electron Microscopy: particle morphologies and dimensions were studied with a Hitachi S-4800 scanning electron microscope at an accelerating voltage of 20 kV, over metalized samples with a mixture of gold and palladium for 90 seconds.

Gas adsorption measurements were recorded on a Micromeritics 3Flex apparatus. The samples were degassed at 80 °C under dynamic vacuum (10^{-6} torr) overnight before recording the adsorption isotherms.

Powder XRD patterns were collected in a PANalytical X'Pert PRO diffractometer using copper radiation (Cu Kα = 1.5418 A) with an X'Celerator detector, operating at 40 mA and 45 kV. Profiles were collected in the $3° < 2\theta < 40°$ range with a step size of 0.013°. Synchrotron XRD data were collected at ALBA Synchrotron (BL04-MSPD). The sample was mounted in a borosilicate glass capillary (0.7 mm) which were subsequently mounted and aligned on the Powder diffraction station for data collection at room temperature (λ = 0.82629125 Å) with a Mythen detector. LeBail refinements were carried out with the FullProf Suite Software Package.[1]

Synthesis and characterization

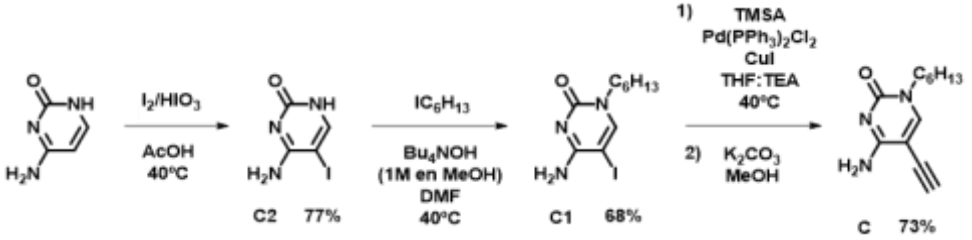
The synthesis of compounds **C2**,[2] **C1**,[2] **G4**[3] and **G3**[2] was previously described in the literature. The synthesis of compounds **C**,[2] **G2**,[2] **G1**[2] and **B1**[4] was modified slightly with respect to the methods described in the literature.

Standard Procedure A (*Sonogashira coupling with TMSA*). A dry THF/NEt_3 or DMF/NEt_3(4:1) solvent mixture was subjected to deoxygenation by three freeze-pump-thaw cycles with argon. Then, this solvent was added over the system containing the corresponding halogenated nucleobase, $Pd(PPh_3)_2Cl_2$ and CuI. The mixture was stirred at room temperature during a few minutes. Then, trimethylsilylacetylene (TMSA) was added dropwise. The reaction is stirred under argon at a given temperature and for a period of time (indicated in each case) until completion, which was monitored by TLC.

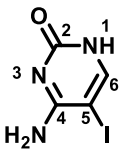
Standard Procedure B (*Sonogashira coupling between central block and nucleobase*). A dry DMF/NEt_3(4:1) solvent mixture was subjected to deoxygenation by three freeze-pump-thaw cycles with argon. Then, this solvent was added over the system containing the corresponding halogenated central block and ethynyl-nucleobase derivative, $Pd(PPh_3)_2Cl_2$ and CuI. The reaction is stirred under argon at a given temperature and for a period of time (indicated in each case) until completion, which was monitored by TLC.

Standard Procedure C (*Sonogashira coupling between central block and nucleobase with controlled addition*). A dry DMF/NEt_3(4:1) solvent mixture was subjected to deoxygenation by three freeze-pump-thaw cycles with argon. Then, this solvent was added over the system containing the corresponding halogenated central block, $Pd(PPh_3)_2Cl_2$ and CuI and to another one containing the corresponding ethynyl-nucleobase derivative. The solution with the ethynyl-nucleobase derivative is added over the other one over the time and at the temperature indicated. The reaction is stirred under argon at a given temperature and for a period of time (indicated in each case) until completion, which was monitored by TLC.

· Cytosine derivative synthesis

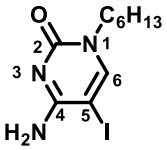


***Scheme S1**. Synthetic route to the cytosine derivative **C**.*



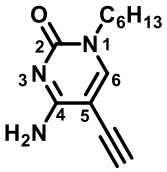
C2 was synthesized according to the literature[2] using cytosine (10.0 g, 90.0 mmol), I_2 (34.3 g, 135.1 mmol) and HIO_3 (22.2 g, 126.2 mmol) in glacial AcOH (300 mL), obtaining **C2** as a white solid (16.38 g, 77%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 10.71 (bs, 1H, N^1*H*), 7.74 (s, 1H, *H*^{6}), 7.50 (bs, 1H, C^{4}NH-*H*), 6.46 (bs, 1H, C^{4}NH-*H*).



C1 was synthesized according to the literature[2] using **C2** (6.0 g, 21.1 mmol), Bu_4NOH (1 M in MeOH, 31.0 mL, 31.0 mmol) and 6-iodohexane (3.4 mL, 42.3 mmol) in dry DMF (80 mL), obtaining **C1** as a yellowish solid (5.53g, 68%).

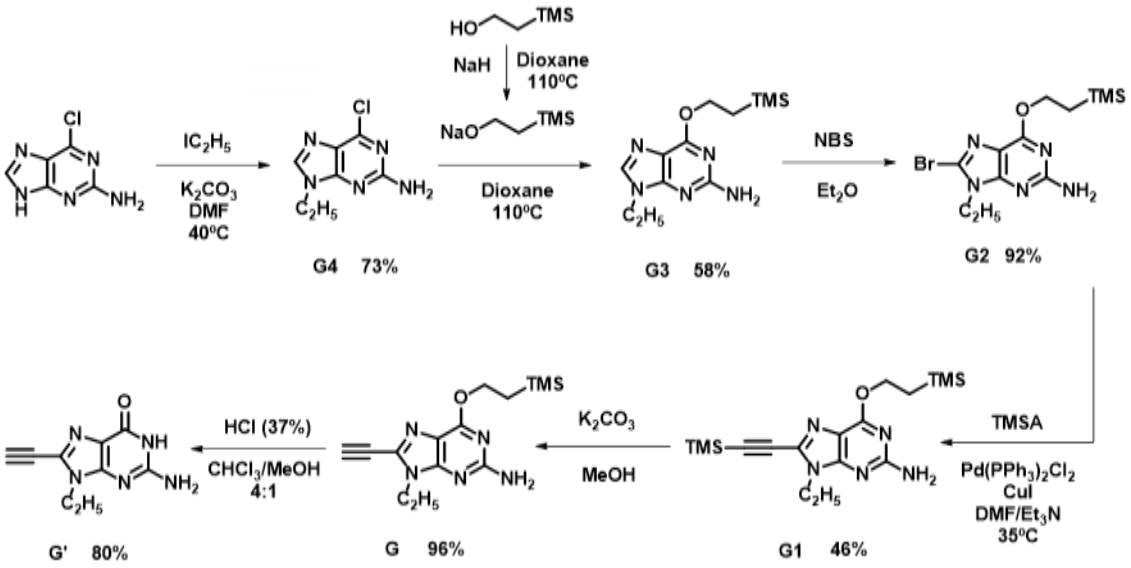
^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm) = 7.57 (s, 1H, H^6), 6.89 (bs, 1H, C^4NH-*H*), 5.36 (bs, 1H, C^4NH-*H*), 3.76 (t, *J* = 7.5 Hz, 2H, $N^1CH_2(CH_2)_4CH_3$), 1.70 (m, 2H, $N^1CH_2CH_2(CH_2)_3CH_3$), 1.30 (m, 6H, $N^1CH_2CH_2(CH_2)_3CH_3$), 0.88 (t, *J* = 7.0 Hz, 3H, $N^1CH_2(CH_2)_4CH_3$).



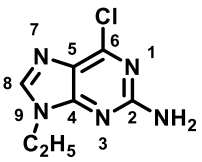
C was synthesized modifying the literature procedure.[2] **C** was synthesized following the **Standard Procedure A** using **C1** (3.65 g, 11.4 mmol), TMSA (3.7 mL, 26.2 mmol), Pd $(PPh_3)_2Cl_2$ (165 mg, 0.24 mmol) and CuI (21.3 mg, 0.11 mmol) in THF/NEt_3 (75 mL). The mixture was stirred at 40°C overnight. Then, the solution was filtered over a celite plug and the solvent was eliminated under reduced pressure. Then, K_2CO_3 (3.06g, 22.1 mmol) and MeOH (50 mL) were added and the mixture was stirred at rt for 1 hour. After that, $CHCl_3$ (70 mL) was added, the solution was filtered over a celite plug and the solvent was eliminated under reduced pressure. Finally, the crude was purified by flash column chromatography in silica gel using $CHCl_3$/MeOH mixture (20:1) as eluent, obtaining **C** as a light beige solid (1.87 g, 73%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.06 (s, 1H, H^6), 7.54 (bs, 1H, C^4NH-*H*), 6.67 (bs, 1H, C^4NH-*H*), 4.33 (s, 1H, C^5CCH), 3.65 (t, *J* = 7.2 Hz, 2H, $N^1CH_2(CH_2)_4CH_3$), 1.55 (m, 2H, $N^1CH_2CH_2(CH_2)_3CH_3$), 1.24 (m, 6H, $N^1CH_2CH_2(CH_2)_3CH_3$), 0.85 (t, *J* = 6.7 Hz, 3H, $N^1CH_2(CH_2)_4CH_3$).

· *Guanosine derivative synthesis*

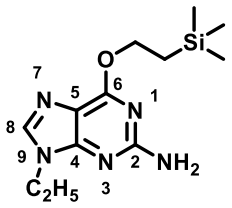


Scheme S2*. Synthetic route to the guanosine derivative* ***G****.*



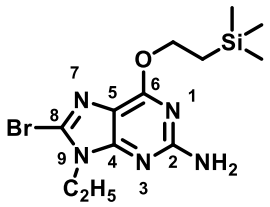
G4 was synthesized according to the literature[3] using 2-amine-4-chloropurine (10.0 g, 59.0 mmol), K_2CO_3 (9.80 g, 70.9 mmol) and iodoethane (5.6 mL, 70.0 mmol) in dry DMF (200 mL), obtaining **G4** as a yellowish solid (8.54 g, 73%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.15 (s, 1H, H^8), 6.89 (bs, 2H, C^2NH_2), 4.07 (q, J = 7.3 Hz, 2H, $N^9CH_2CH_3$), 1.36 (t, J = 7.3 Hz, 3H, $N^9CH_2CH_3$).



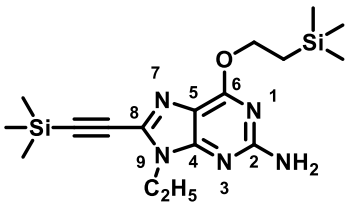
G3 was synthesized according to the literature[2] using NaH (60%, 3.72 g, 93.0 mmol) and 2-(trimethylsilyl)ethanol (13.5 mL, 96.6 mmol) in dry dioxane (150 mL) in the first step and **G4** (6.12 g, 30.5 mmol) in dry dioxane (100 mL) in the second step, obtaining **G3** as a brown solid (4.99 g, 58%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 7.85 (s, 1H, H^8), 6.32 (bs, 2H, C^2NH_2), 4.49 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 4.02 (q, J = 7.2 Hz, 2H, $N^9CH_2CH_3$), 1.34 (t, J = 7.2 Hz, 3H, $N^9CH_2CH_3$), 1.12 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 0.07 (s, 9H, $C^6OCH_2CH_2Si(CH_3)_3$).



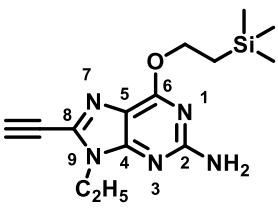
G2 was synthesized modifying the literature procedure.[2] In a round bottom flask equipped with a magnetic stirrer NBS (2.95 g, 16.6 mmol) was added in portions over 1 hour over a solution of **G3** (4.19 g, 15.0 mmol) in Et_2O (130 mL). After that, the solution was filtered over a celite plug and the solvent was eliminated under reduced pressure. Finally, the crude was purified by flash column chromatography in silica gel using Cy/AcOEt mixture (2:1) as eluent, obtaining **G2** as a white solid (4.94 g, 92%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 6.52 (bs, 2H, C^2NH_2), 4.48 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 4.02 (q, J = 7.2 Hz, 2H, $N^9CH_2CH_3$), 1.26 (t, J = 7.2 Hz, 3H, $N^9CH_2CH_3$), 1.12 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 0.07 (s, 9H, $C^6OCH_2CH_2Si(CH_3)_3$).



G1 was synthesized modifying the literature procedure.[2] **G1** was synthesized following the **Standard Procedure A** using **G2** (1.50 g, 5.83 mmol), TMSA (3.2 mL, 20.9 mmol), Pd $(PPh_3)_2Cl_2$ (88.2 mg, 0.13 mmol) and CuI (8.0 mg, 0.042 mmol) in DMF/NEt_3 (20 mL). The mixture was stirred at 35°C overnight. Then, the solution was filtered over a celite plug and the solvent was eliminated under reduced pressure. Finally, the crude was purified by flash column chromatography in silica gel using Cy/AcOEt mixture (3:1) as eluent, obtaining **G1** as a brown oil (0.74 g, 46%).

^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm) = 4.83 (bs, 2H, C^2NH_2), 4.55 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 4.18 (q, J = 7.3 Hz, 2H, $N^9CH_2CH_3$), 1.41 (t, J = 7.2 Hz, 3H, $N^9CH_2CH_3$), 1.21 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 0.28 (s, 9H, $C^8CCSi(CH_3)_3$), 0.08 (s, 9H, $C^6OCH_2CH_2Si(CH_3)_3$).

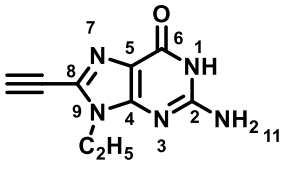


G. In a round bottom flask equipped with a magnetic stirrer K_2CO_3 (0.70 g, 5.03 mmol) was added over a solution of **G1** (0.63 g, 1.68 mmol) in MeOH (50 mL). The mixture was stirred at rt for 1 hour. After that, $CHCl_3$ (50 mL) was added, the solution was filtered over a celite plug and the solvent was eliminated under reduced pressure. Finally, the crude was purified by flash column chromatography in silica gel using Cy/AcOEt mixture (2:1) as eluent, obtaining **G** as a beige solid (0.49 g, 96%).

^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm) = 4.86 (bs, 2H, C^2NH_2), 4.56 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 4.20 (q, *J* = 7.2 Hz, 2H, $N^9CH_2CH_3$), 3.37 (s, 1H, C^8CCH), 1.42 (t, *J* = 7.2 Hz, 3H, $N^9CH_2CH_3$), 1.22 (m, 2H, $C^6OCH_2CH_2Si(CH_3)_3$), 0.08 (s, 9H, $C^6OCH_2CH_2Si(CH_3)_3$).

^{13}C NMR (75 MHz, $CDCl_3$) δ (ppm) = 160.5, 159.0, 152.1, 130.9, 114.7, 81.1, 72.3, 64.1, 37.3, 16.6, 14.0, -2.4.

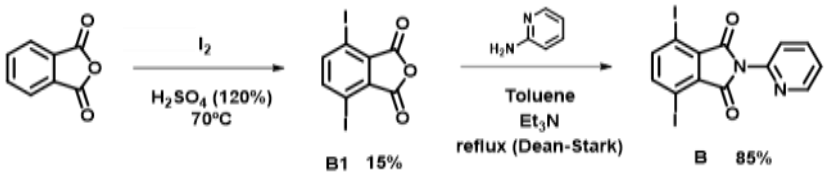
HRMS (ESI+): calculated for $C_{14}H_{22}N_5OSi$ $[M+H]^+$: 304.1588. Found: 304.1589 $[M+H]^+$.



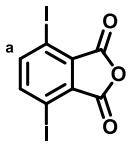
G'. In a round bottom flask equipped with a magnetic stirrer aqueous solution of HCl (37%, 10 drops) was added over a solution of **G** (200 mg, 0.659 mmol) in a $CHCl_3$/MeOH mixture (4:1, 20 mL). The mixture was stirred at rt overnight. Then, an aqueous solution of saturated $NaHCO_3$ (15 drops) was added, and the solvent was eliminated under reduced pressure. Finally, the solid was filtered and washed with water, MeCN and $CHCl_3$, obtaining **G'** as a beige solid (107 mg, 80%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 10.71 (bs, 1H, N^1H), 6.61 (bs, 2H, C^2NH_2), 4.71 (s, 1H, C^8CCH), 4.01 (q, *J* = 7.0 Hz, 2H, $N^9CH_2CH_3$), 1.27 (t, *J* = 7.2 Hz, 3H, $N^9CH_2CH_3$).

· *Central block synthesis*

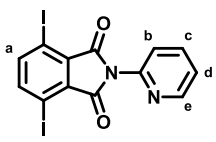


Scheme S3. *Synthetic route to the central block **B**.*



B1 was synthesized modifying the literature procedure.[4] In a round bottom flask equipped with a magnetic stirrer I_2 (34.4 g, 135.5 mmol) was added over a solution of phthalic anhydride (10.0 g, 67.6 mmol) in fuming H_2SO_4 (120%, 50 mL). The reaction was stirred at 70°C for 24 hours. Then, the solution was poured over an ice (500 mL) and the solid filtered and washed with water, an aqueous saturated solution of $NaHCO_3$, an aqueous saturated solution of $N_2S_2O_3$, and water, successively. The solid was extracted with CH_2Cl_2, and the organic solvent was dried over $MgSO_4$, filtered, and eliminated over reduced pressure. The crude was purified by crystallization in glacial AcOH, obtaining **B** as a yellow solid (4.06 g, 15%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.01 (s, 2H, H^a).



B. In a round bottom flask equipped with a magnetic stirrer, 2-aminopiridine (1.15 g, 12.2 mmol) was added to a solution of **B1** (4.00 g, 10.0 mmol) in toluene (100 mL) and NEt_3 (2.0 mL). The mixture was stirred at reflux with a Dean-Stark apparatus for 24 hours. Then, the solvent was eliminated under reduced pressure and the crude was purified by flash column chromatography

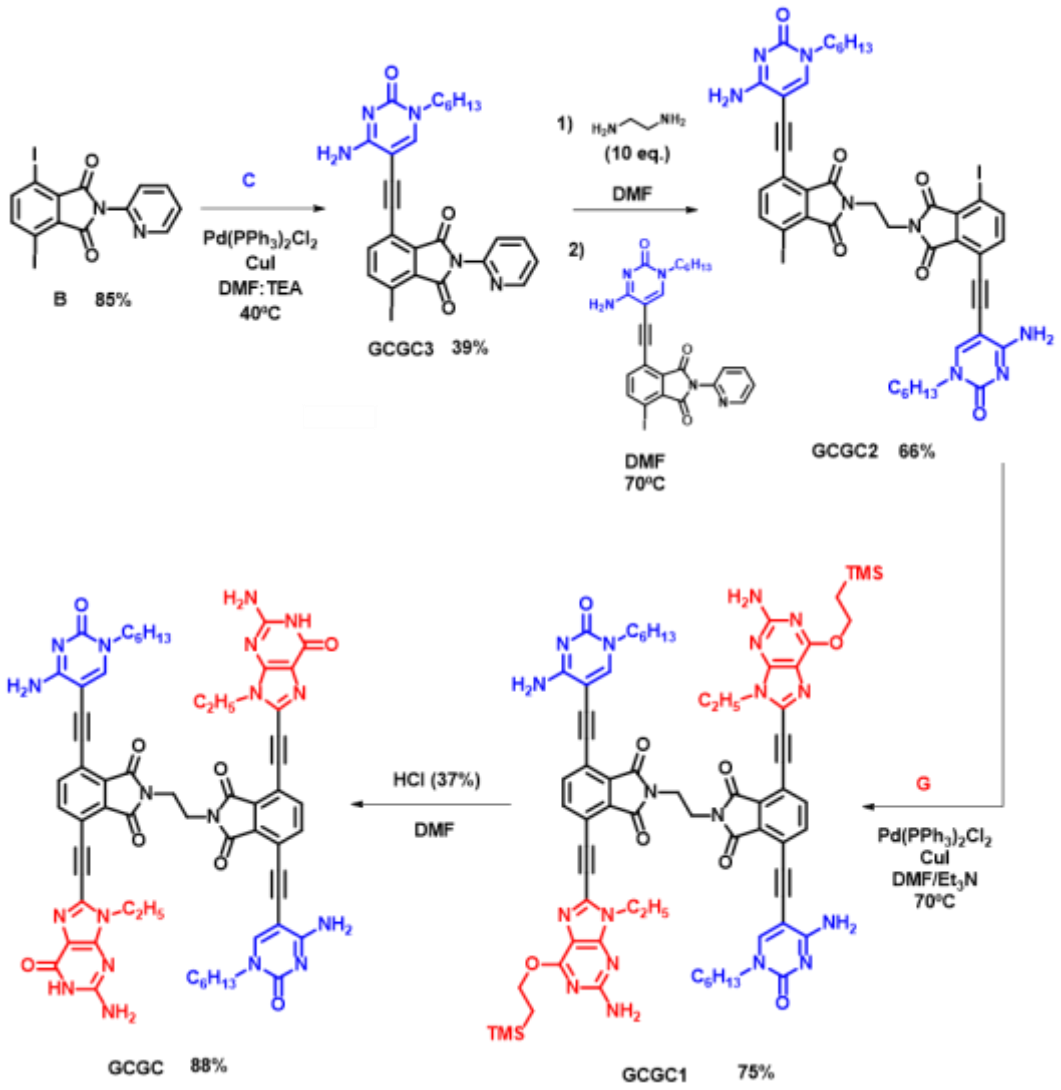
in silica gel using toluene/AcOEt mixture (10:1) as eluent, obtaining **B** as a beige solid (4.45 g, 93%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.65 (ddd, J = 4.8 Hz, J' = 2.0 Hz, J'' = 1.0 Hz, 1H, H^e), 8.06 (td, J = 7.8 Hz, J' = 1.9 Hz, 1H, H^c), 7.97 (s, 2H, H^a), 7.55 (m, 2H, H^b, H^d).

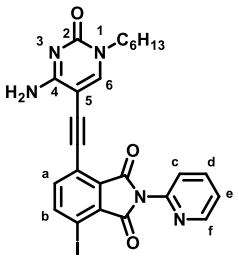
^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm) = 163.7, 149.4, 145.9, 145.7, 138.7, 133.4, 124.3, 123.1, 90.6.

HRMS (ESI+): calculated for $C_{13}H_7I_2N_2O_2$ $[M+H]^+$: 476.8591. Found: 476.8589 $[M+H]^+$.

· *Fused monomer synthesis*



Scheme S4. *Synthetic route to the fused monomer* ***GCGC****.*

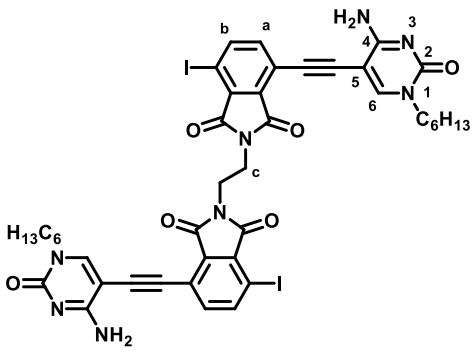


GCGC3 was synthesized following the **Standard Procedure B** using **B** (1.63 g, 3.42 mmol), **C** (251 mg, 1.14 mmol), $Pd(PPh_3)_2Cl_2$ (24.3 mg, 0.0035 mmol) and CuI (3.4 mg, 0.018 mmol) in DMF/NEt_3 (100 mL). The mixture was stirred at 40°C overnight. Then, the solvent was eliminated under reduced pressure and the crude was purified by flash column chromatography in silica gel using $CHCl_3$/MeOH mixture (100:1) as eluent, obtaining **GCGC3** as a yellow solid (253 mg, 39%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.67 (ddd, J = 4.9, J' = 1.9, J'' = 0.8 Hz, 1H, H^f), 8.32 (d, J = 8.2 Hz, 1H, H^b), 8.31 (s, 1H, H^6), 8.08 (td, J = 7.7, J' = 1.8 Hz, 1H, H^d), 7.94 (bs, 1H, C^4NH-H), 7.58 (m, 3H, H^a, H^c, H^e), 7.02 (bs, 1H, C^4NH-H), 3.73 (t, J = 7.3 Hz, 2H, $N^1CH_2(CH_2)_4CH_3$), 1.59 (m, 2H, $N^1CH_2CH_2(CH_2)_3CH_3$), 1.26 (m, 6H, $N^1CH_2CH_2(CH_2)_3CH_3$), 0.85 (t, J = 6.7 Hz, 3H, $N^1CH_2(CH_2)_4CH_3$).

^{13}C NMR (75 MHz, DMSO-d_6) δ (ppm) = 164.8, 164.6, 164.3, 153.8, 150.9, 149.5, 145.6, 145.5, 138.8, 136.4, 132.4, 131.0, 124.4, 123.3, 118.9, 90.9, 90.7, 89.9, 87.4, 49.3, 30.8, 28.5, 25.5, 21.9, 13.8.

HRMS (ESI+): calculated for $C_{25}H_{23}IN_5O_3$ $[M+H]^+$: 568.0840. Found: 568.0857 $[M+H]^+$.

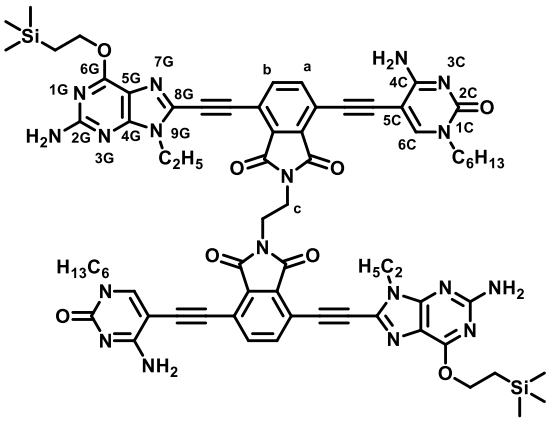


GCGC2. In a round bottom flask equipped with a magnetic stirrer, ethylenediamine (0.38 mL, 5.68 mmol) was added over a solution of **GCGC3** (300 mg, 0.529 mmol) in DMF (15 mL). The mixture was stirred at rt for 1 hour. Then the solvent was eliminated under reduced pressure, DMF was added (10 mL) and eliminated again under reduced pressure. The cycle was repeated 3 times. After that, the solid was redissolved in DMF (15 mL) and more **GCGC3** (330 mg, 0.582 mmol) was added. The mixture was stirred at 70°C overnight. Finally, the solvent was eliminated under reduced pressure and the crude was purified by flash column chromatography in silica gel using $CHCl_3$/MeOH mixture (30:1) as eluent, obtaining **GCGC2** as a yellow solid (353 mg, 66%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.28 (s, 2H, H^6), 8.18 (d, J = 8.2 Hz, 2H, H^b), 7.94 (bs, 2H, C^4NH-H), 7.49 (d, J = 8.2 Hz, 2H, H^a), 7.01 (bs, 2H, C^4NH-H), 3.84 (s, 4H, H^c), 3.74 (t, J = 7.3 Hz, 4H, $N^1CH_2(CH_2)_4CH_3$), 1.61 (m, 4H, $N^1CH_2CH_2(CH_2)_3CH_3$), 1.27 (m, 12H, $N^1CH_2CH_2(CH_2)_3CH_3$), 0.86 (t, J = 6.7 Hz, 6H, $N^1CH_2(CH_2)_4CH_3$).

^{13}C NMR (DMSO-d_6). Due to the extremely low solubility of this compound, a ^{13}C NMR spectrum of sufficient quality could not be acquired even after 24 h in a 500 MHz instrument.

HRMS (ESI+): calculated for $C_{42}H_{41}N_8O_6I_2$ $[M+H]^+$: 1007.1233. Found: 1007.1244 $[M+H]^+$.

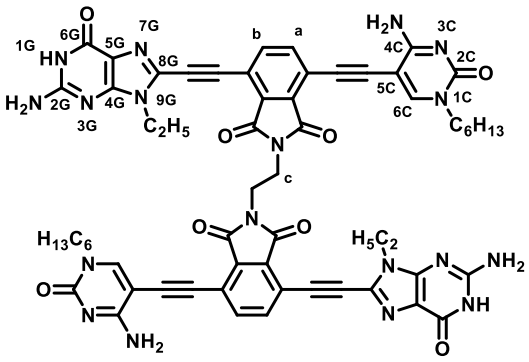


GCGC1 was synthesized following the **Standard Procedure C** using **GCGC2** (253 g, 0.251 mmol), Pd $(PPh_3)_2Cl_2$ (18.3 mg, 0.026 mmol) and CuI (2.1 mg, 0.011 mmol) in DMF/NEt_3 (10 mL) and **G** (168 mg, 0.555 mmol) in DMF/NEt_3 (10 mL). The addition was carried out for 4 hours at 50°C and the mixture was stirred at 70°C overnight. Then, the solvent was eliminated under reduced pressure and the crude was purified by flash column chromatography in silica gel using $CHCl_3$/MeOH mixture (25:1) as eluent, obtaining **GCGC1** as a red solid (255 mg, 75%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 8.30 (s, 2H, H^{6C}), 7.95 (bs, 2H, C^{4C}NH-*H*), 7.93 (d, *J* = 8.1 Hz, 2H, H^b), 7.80 (d, *J* = 8.2 Hz, 2H, H^a), 7.01 (bs, 2H, C^{4C}NH-*H*), 6.71 (bs, 4H, $C^{2G}NH_2$), 4.53 (t, *J* = 8.1 Hz, 4H, $C^{6G}OCH_2CH_2Si(CH_3)_3$), 4.14 (q, *J* = 6.8 Hz, 4H, $N^{9G}CH_2CH_3$), 3.92 (s, 4H, H^c), 3.75 (t, *J* = 6.0 Hz, 4H, $N^{1C}CH_2(CH_2)_4CH_3$), 1.62 (m, 4H, $N^{1C}CH_2CH_2(CH_2)_3CH_3$), 1.26 (m, 12H, $N^{1C}CH_2CH_2(CH_2)_3CH_3$), 1.14 (m, 10H, $N^{9G}CH_2CH_3$, $C^{6G}OCH_2CH_2Si(CH_3)_3$), 0.86 (t, *J* = 5.9 Hz, 6H, $N^{1C}CH_2(CH_2)_4CH_3$), 0.09 (s, 18H, $C^{6G}OCH_2CH_2Si(CH_3)_3$).

^{13}C NMR (DMSO-d_6). Due to the extremely low solubility of this compound, a ^{13}C NMR spectrum of sufficient quality could not be acquired even after 24 h in a 500 MHz instrument.

HRMS (MALDI, Dithranol): calculated for $C_{70}H_{80}N_{18}NaO_8Si_2$ $[M+Na]^+$: 1379.5837. Found: 1379.5885 $[M+Na]^+$.

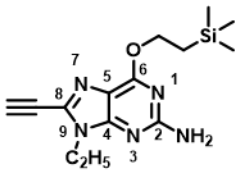


GCGC. In a round bottom flask equipped with a magnetic stirrer an aqueous solution of HCl (37%, 3 drops) was added over a solution of **GCGC1** (10.0 mg, 0.007 mmol) in DMF (3 mL). The mixture was stirred at rt for 48 hours. Then, an aqueous saturated $NaHCO_3$ solution (5 drops) was added and the mixture was stirred for 10 minutes. After that, the solid was filtered and washed with water, DMF, MeOH and CH_2Cl_2, obtaining **GCGC** as an orange solid (7.5 mg, 88%).

^{1}H NMR (300 MHz, DMSO-d_6) δ (ppm) = 10.55 (bs, 2H, $N^{1G}H$), 8.19 (s, 2H, H^{6C}), 7.97 (bs, 2H, C^{4C}NH-*H*), 7.87 (d, *J* = 8.2 Hz, 2H, H^b), 7.76 (d, *J* = 8.2 Hz, 2H, H^a), 7.13 (bs, 2H, C^{4C}NH-*H*), 6.66 (bs, 4H, $C^{2G}NH_2$), 4.14 (q, *J* = 7.2 Hz, 4H, $N^{9G}CH_2CH_3$), 3.97 (s, 4H, H^c), 3.78 (t, *J* = 7.2 Hz, 4H, $N^{1C}CH_2(CH_2)_4CH_3$), 1.67 (m, 4H, $N^{1C}CH_2CH_2(CH_2)_3CH_3$), 1.31 (m, 12H, $N^{1C}CH_2CH_2(CH_2)_3CH_3$), 1.25 (t, *J* = 7.1 Hz, 6H, $N^{9G}CH_2CH_3$), 0.89 (t, *J* = 6.4 Hz, 6H, $N^{1C}CH_2(CH_2)_4CH_3$).

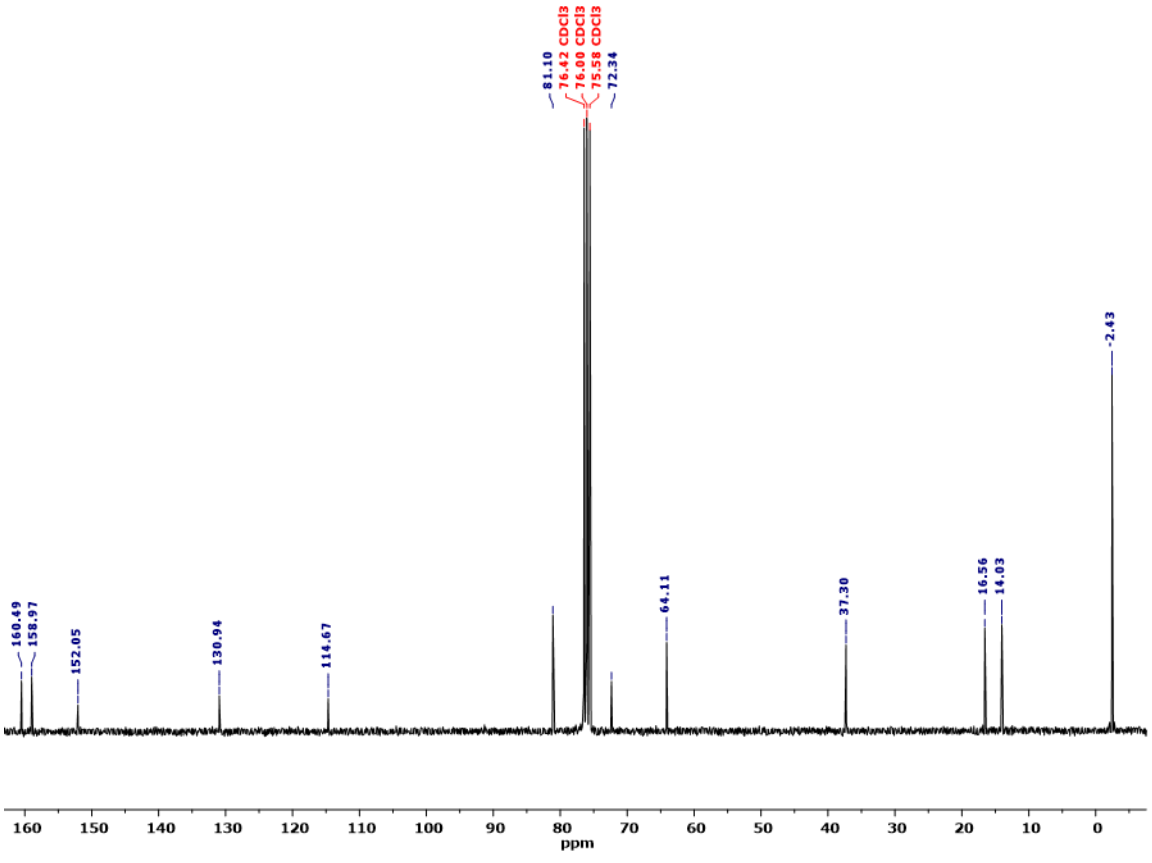
^{13}C NMR (DMSO-d_6). Due to the extremely low solubility of this compound, a ^{13}C NMR spectrum of sufficient quality could not be acquired even after 24 h in a 500 MHz instrument.

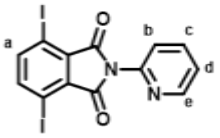
HRMS (MALDI, Dithranol): calculated for $C_{60}H_{57}N_{18}O_8$ $[M+H]^+$: 1157.4601. Found: 1157.4597 $[M+H]^+$.



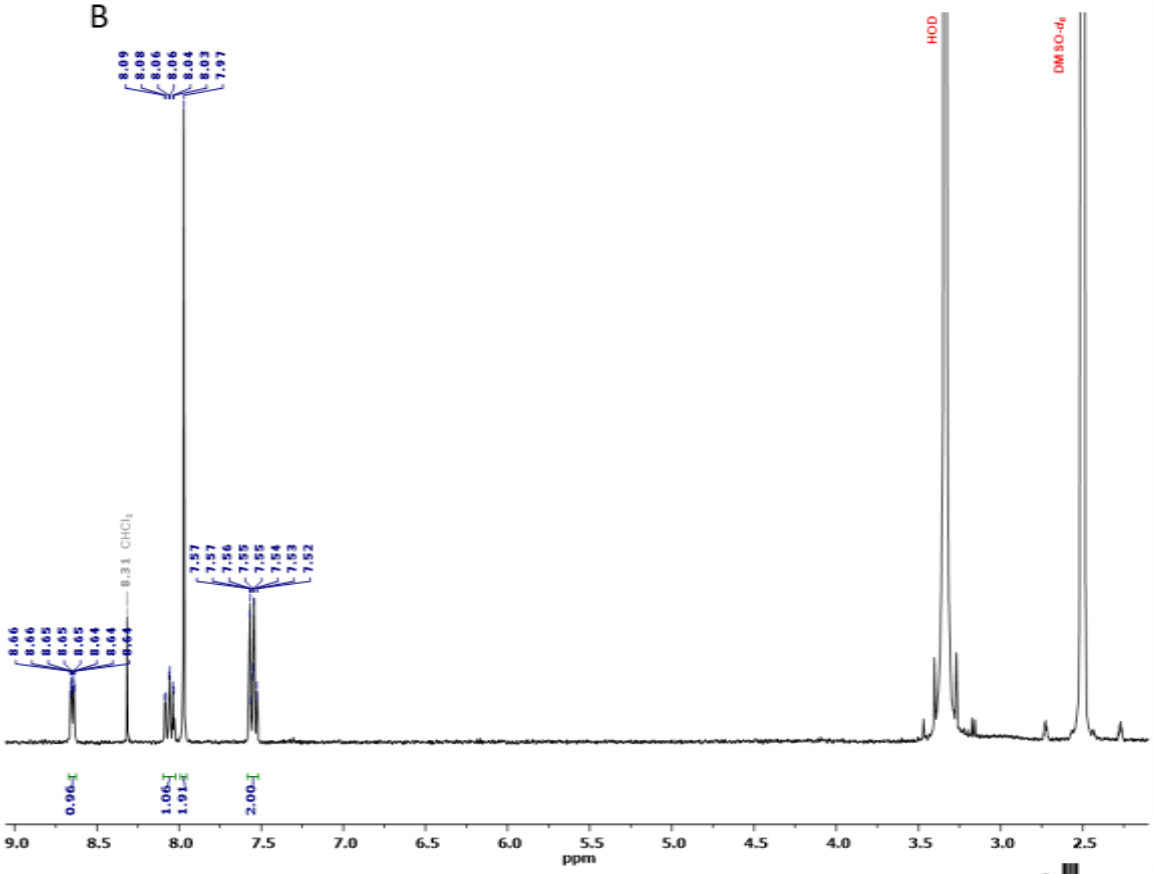
G

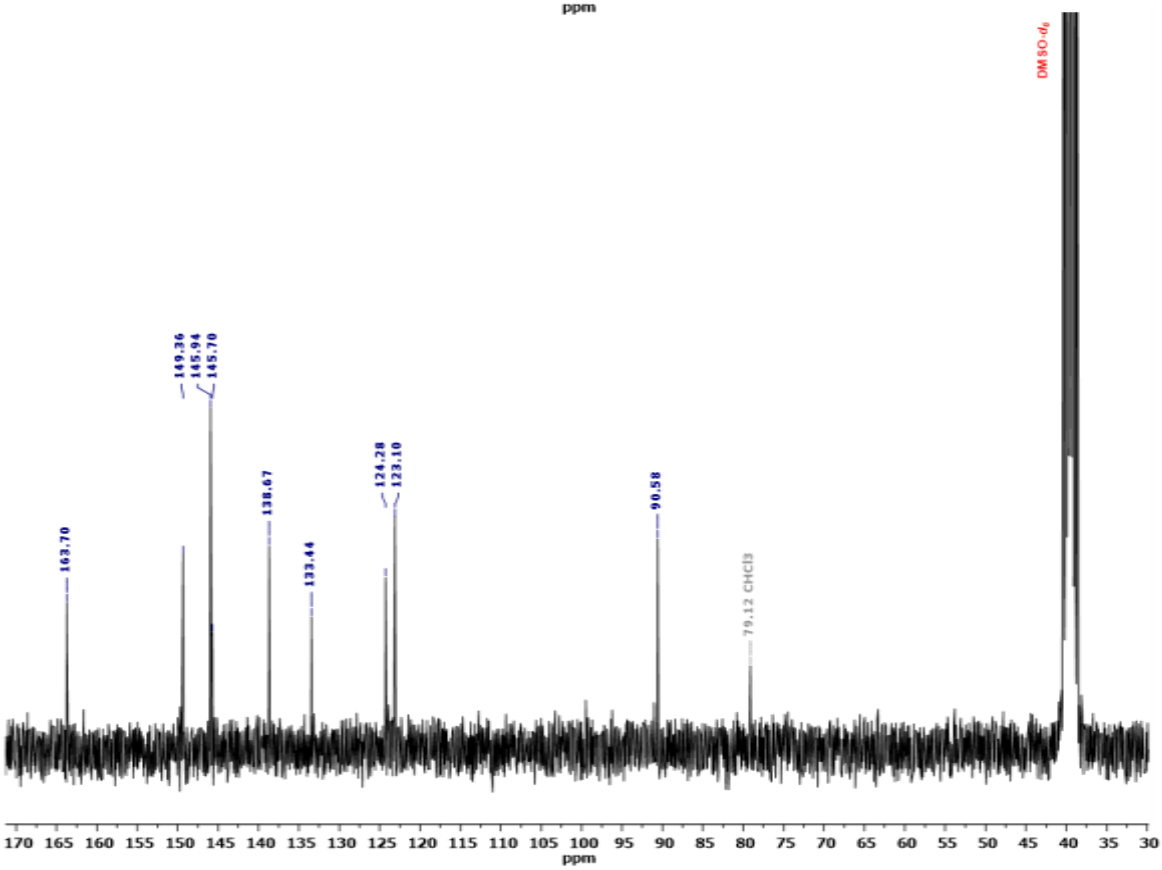
0.08
7.26 CDCl3
4.91
4.59
4.56
4.53
4.23
4.21
4.18
4.16
3.37
1.44
1.41
1.39
1.25
1.22
1.22
1.19
1.78
2.01
2.00
0.91
3.00
2.09
8.73
ppm

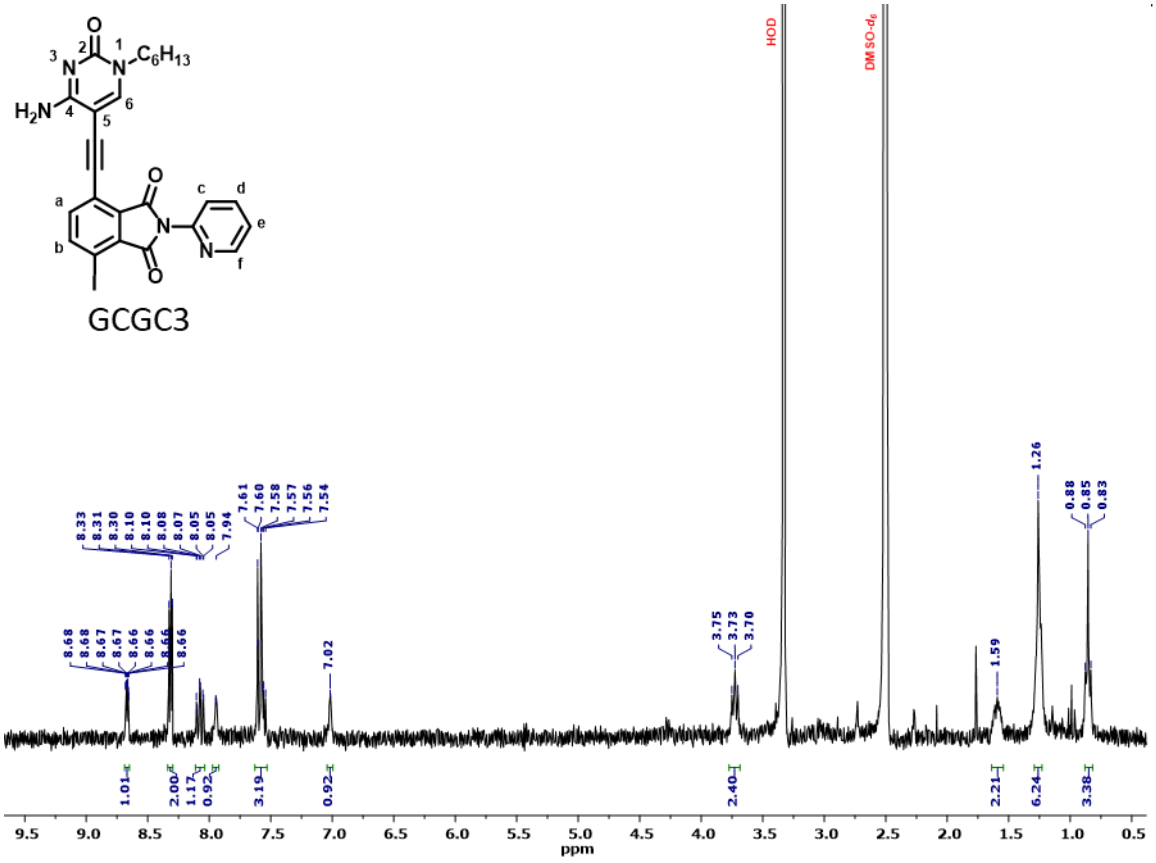

81.10
76.42 CDCl3
76.00 CDCl3
75.58 CDCl3
72.34
-2.43
160.49
158.97
152.05
130.94
114.67
64.11
37.30
16.56
14.03
ppm

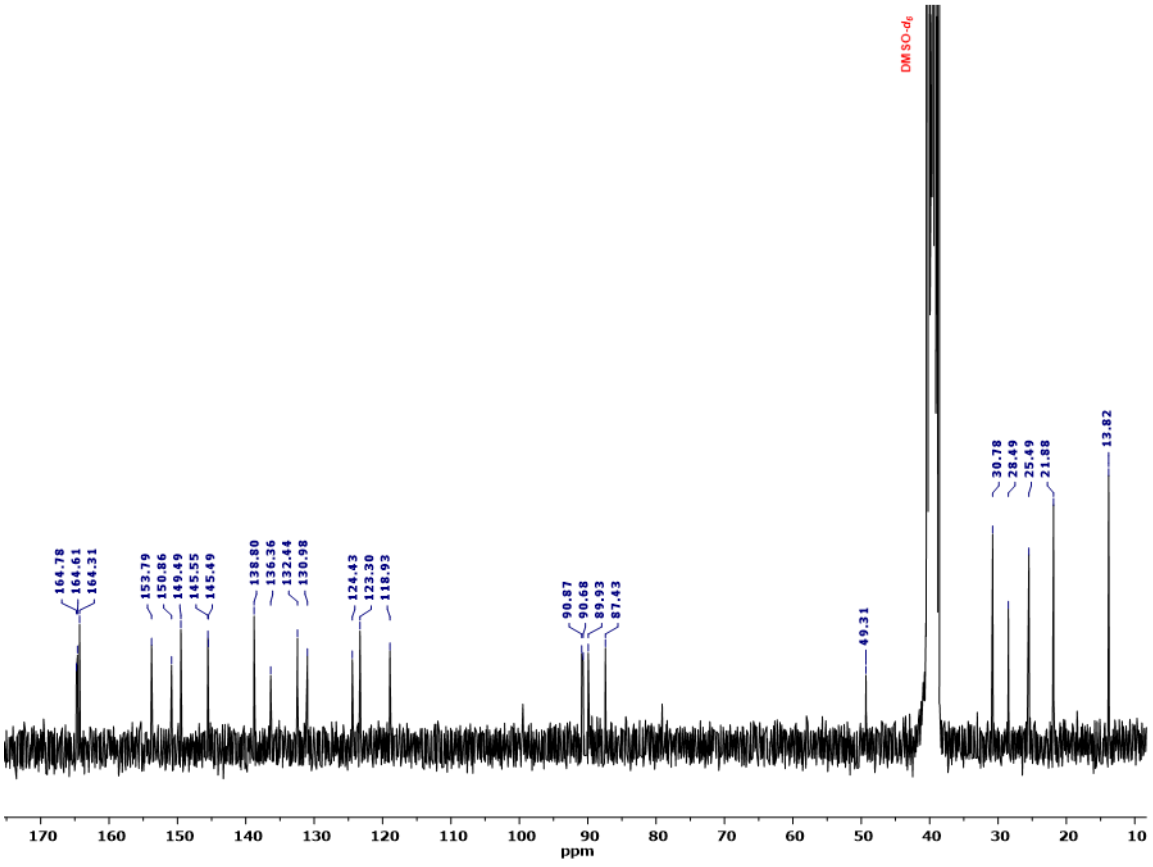

a
b
c
d
e
I
I
O
O
N
N

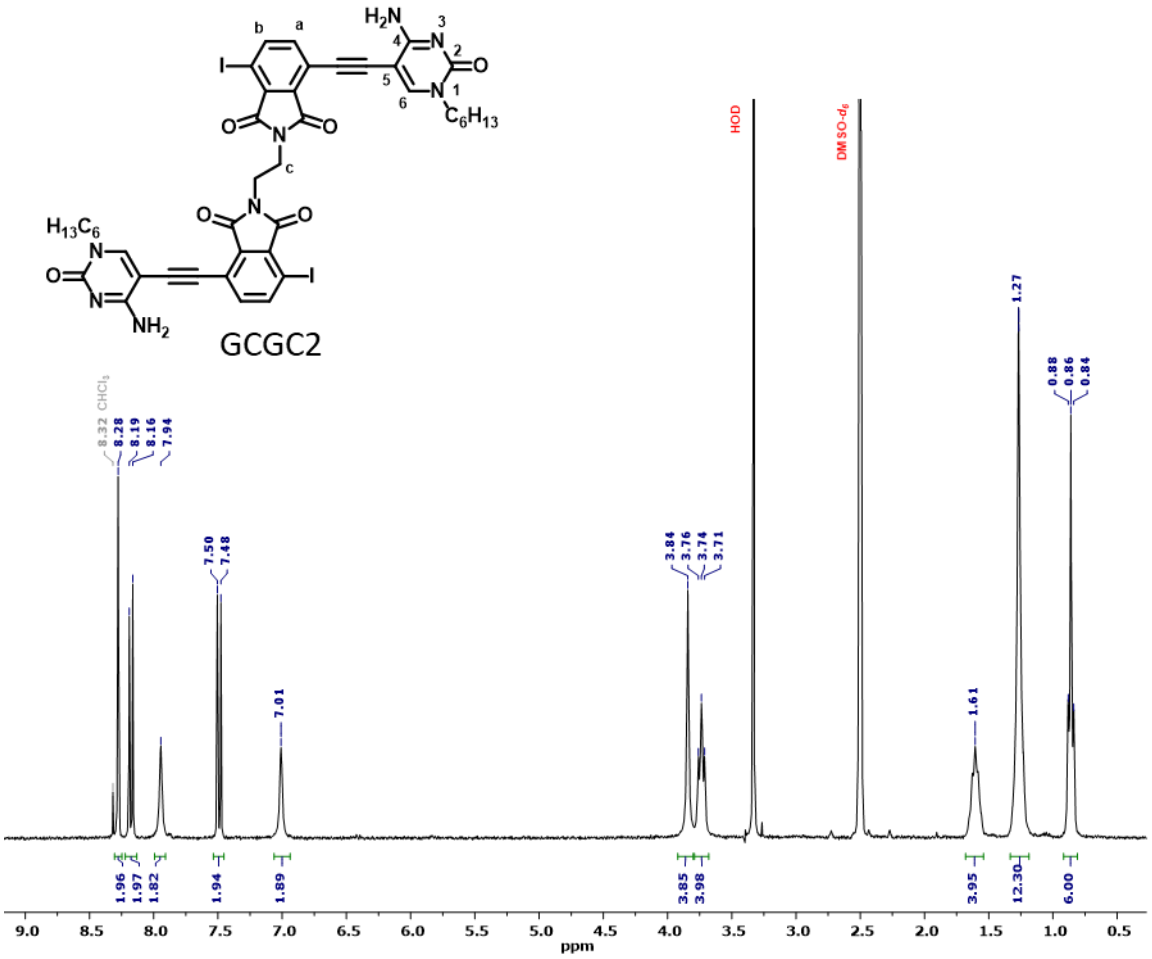
B

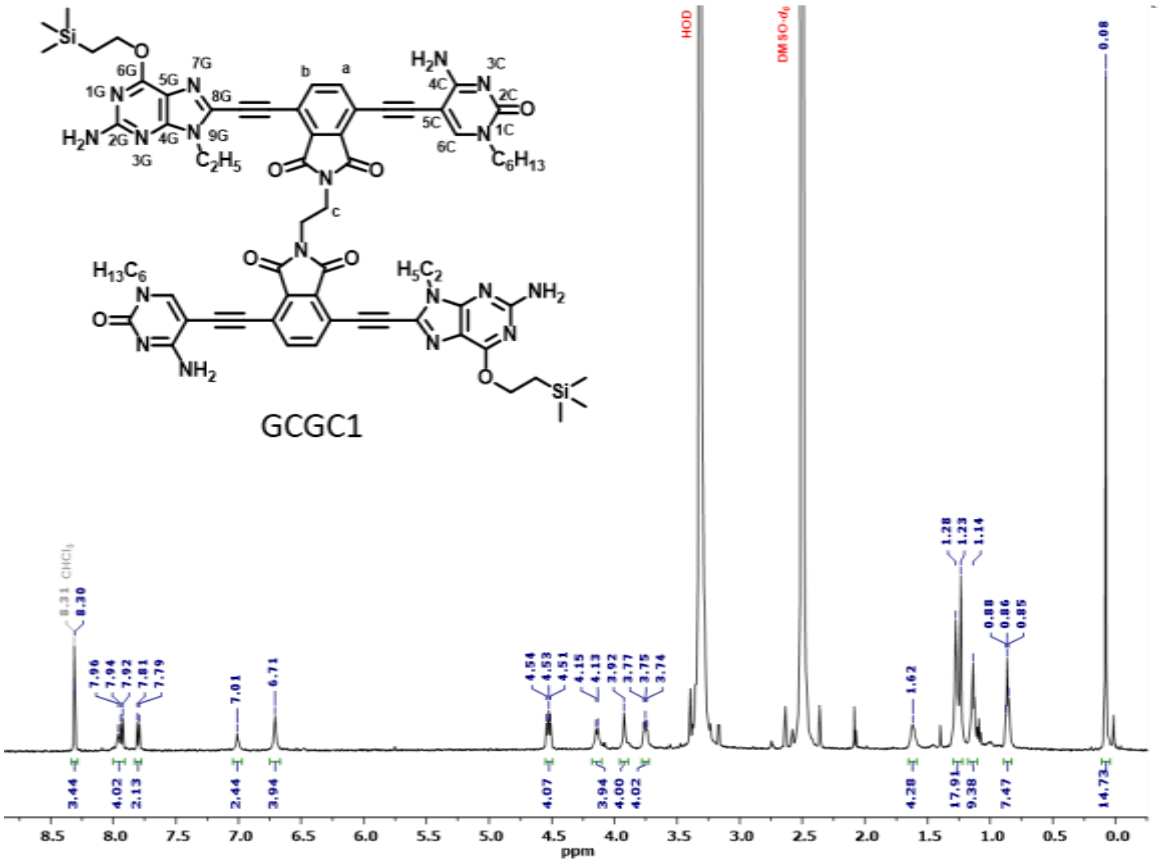

8.69
8.68
8.66
8.66
8.64
8.64
8.63
7.97
8.31 CHCl3
7.57
7.57
7.56
7.55
7.55
7.54
7.53
7.52
8.66
8.66
8.65
8.65
8.65
8.64
8.64
8.64
HOD
DMSO-d6
0.96
1.06
1.91
2.00
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
ppm

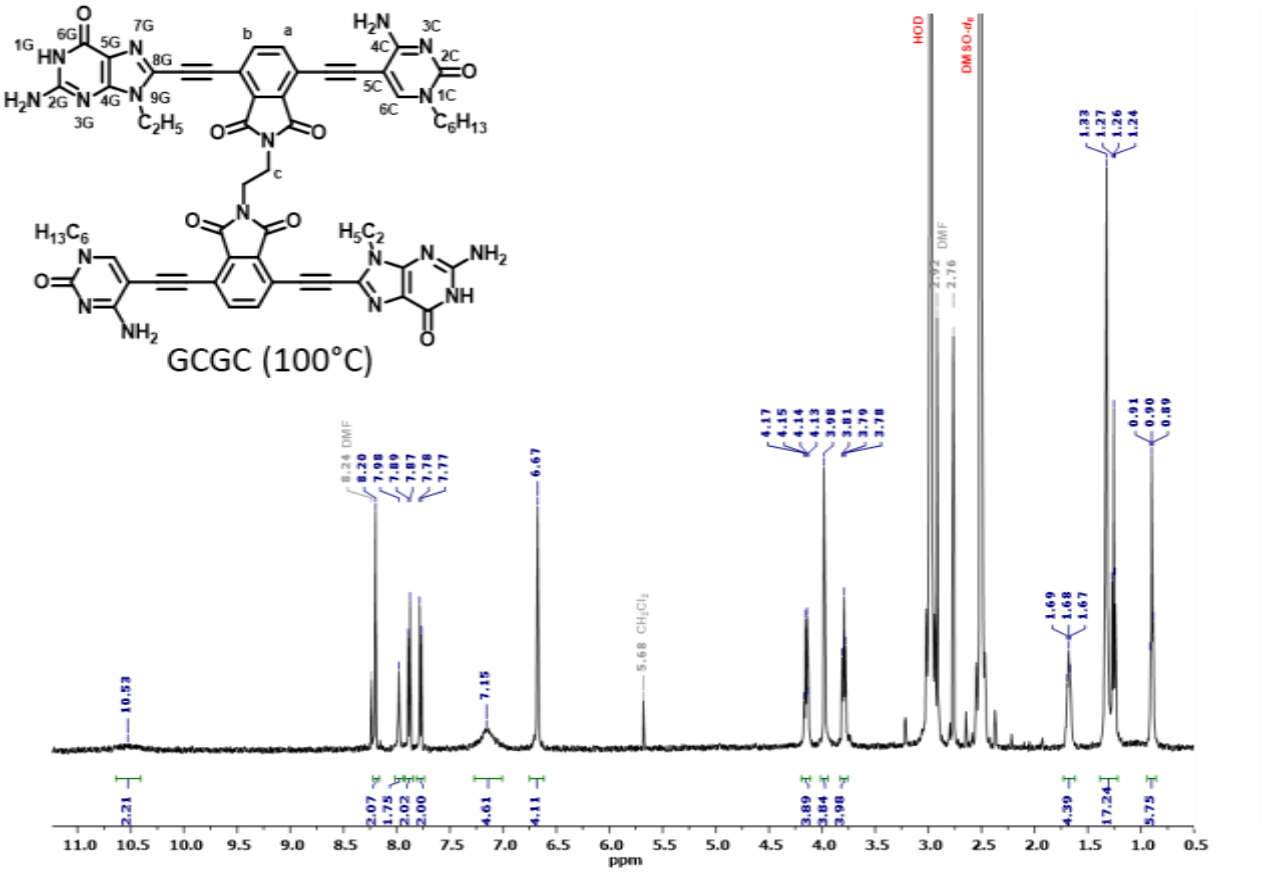

DMSO-d6
149.36
145.94
145.70
138.67
133.44
124.28
123.10
90.58
79.12 CHCl3
163.70
170
165
160
155
150
145
140
135
130
125
120
115
110
105
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
ppm


GCGC3
HOD
DMSO-d6
8.68
8.66
7.61
7.60
7.58
7.57
7.56
7.54
7.02
3.75
3.73
3.70
1.59
1.26
0.88
0.85
0.83
1.01
2.00
1.17
0.92
3.19
0.92
2.40
2.21
6.24
3.38
ppm


DMSO-d6
164.78
164.61
164.31
153.79
150.86
149.49
145.55
145.49
138.80
136.36
132.44
130.98
124.43
123.30
118.93
90.87
90.68
89.93
87.43
49.31
30.78
28.49
25.49
21.88
13.82
ppm


GCGC2
HOD
DMSO-d6
8.32 CHCl3
8.28
8.19
8.16
7.94
7.50
7.48
7.01
3.84
3.76
3.74
3.71
1.61
1.27
0.88
0.86
0.84
1.96
1.97
1.82
1.94
1.89
3.85
3.98
3.95
12.30
6.00
ppm


GCGC1
HOD
DMSO-d6
0.08
8.31 CHCl3
8.30
7.96
7.94
7.92
7.81
7.79
7.01
6.71
4.54
4.53
4.51
4.15
4.13
3.92
3.77
3.75
3.74
1.62
1.28
1.23
1.14
0.88
0.86
0.85
3.44
4.02
2.13
2.44
3.94
4.07
3.94
4.00
4.02
4.28
17.91
9.38
7.47
14.73
ppm


GCGC (100°C)
HOD
DMSO-d6
ppm

S1. NMR studies

At low temperatures the ^{1}H NMR spectra of **GCGC** presented just broad signals due to the strong aggregation and precipitation of the fused monomers. With the increase of temperature, the signals of the different protons of the non-aggregated **GCGC** molecule become visible. The typical hydrogen bonded signals could not be seen because the concentration of aggregates with a small enough size to be soluble is too low.

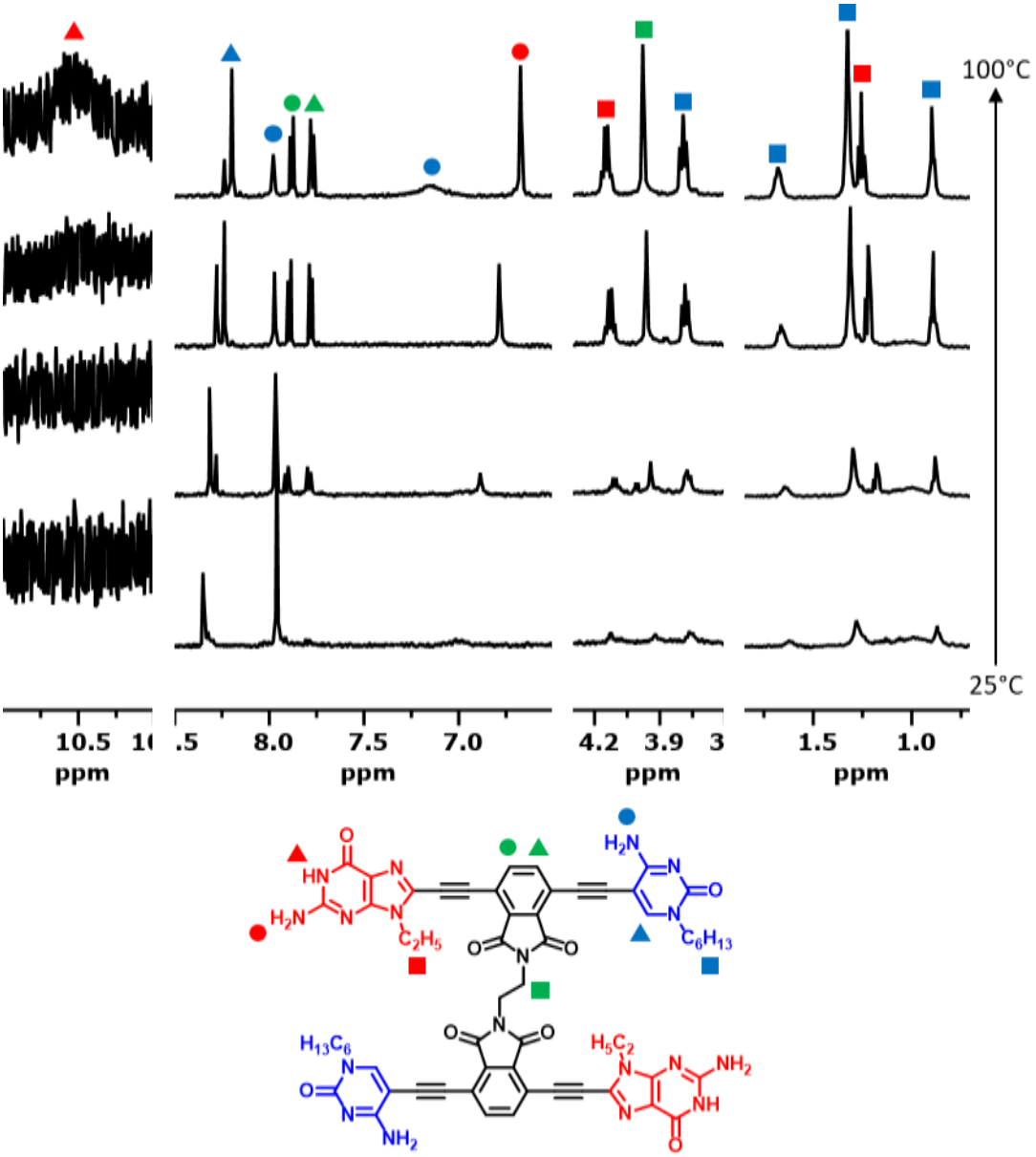


Figure S1. *^{1}H NMR spectra of the fused monomer **GCGC** in DMSO-d_6 at different temperatures.*

S2. UV/visible spectroscopy studies

The absorption spectra show a maximum at 400 nm independently of the concentration. The decrease of the absorption in the most concentrated sample is due to the presence of undissolved material, as seen in the picture by the naked eye.

The emission spectra show an important decrease in the intensity when the concentration increases because of the presence of small aggregates. In the normalized spectra a shift to the blue from 538 to 458 nm can be observed as the sample is diluted.

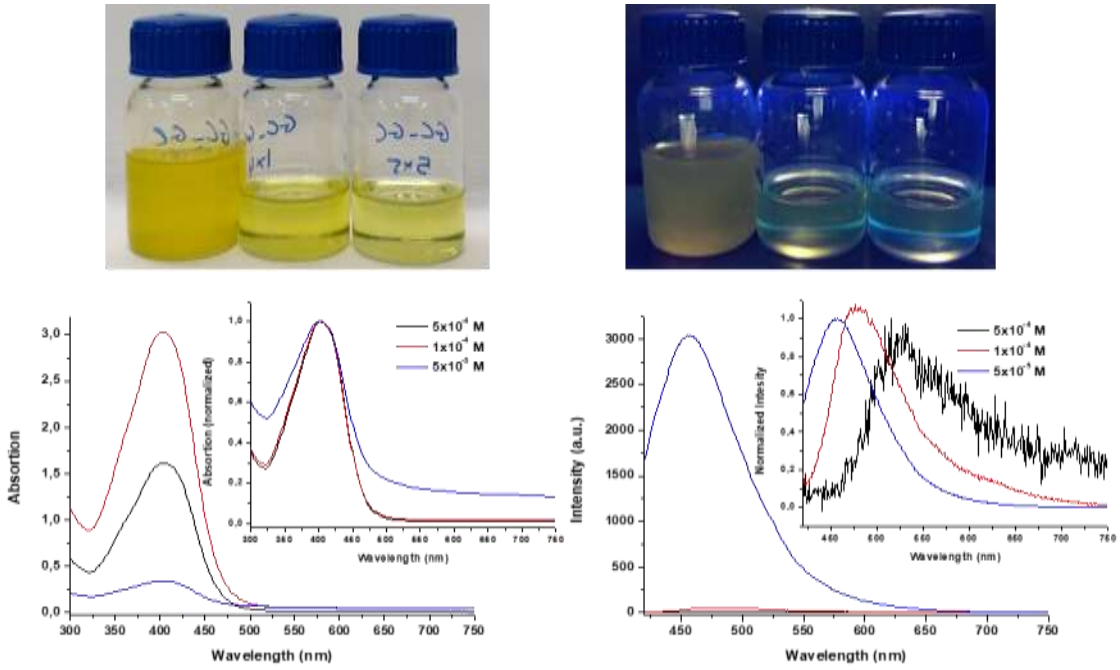


Figure S2. *a) Samples of **GCGC** at different concentrations (5x10^{-4} M, 1x10^{-4} M and 5x10^{-5} M from left to right) in DMSO; b) absorption spectra of **GCGC** in DMSO at 90°C; c) emission spectra of **GCGC** in DMSO at 90°C (the insets show the normalized spectra).*

S3. Crystallization process

For the crystallization protocol, we followed a similar approach than the one described in other publications.[5] Because of the extremely low solubility, we used the protected fused monomer **GCGC1** and deprotected the carbonyl groups of the guanosine fragment *in situ*. In this way the presence of fused monomer and its assembly is carried out in a controlled way, producing (theoretically) materials with a better crystallinity.

For the preparation of the samples, fused monomer **GCGC1** was dissolved in DMF down to a concentration of $1x10^{-2}$ M. Then, in a vial with the previous solution (0.8 mL), a solution of HCl (0.1 M in DMF, 0.2 mL) was added. After that, different quantities of cosolvent ($CHCl_3$, MeOH, MeCN) and DMF (to complete a volume of 2.5 mL) were added, the vial was sealed and treated thermally. Finally, the supernatant was extracted, and the crystals were washed with MeOH several times.

The best results were obtained using 0.5 mL of MeOH as cosolvent and annealing the sample at 65°C for 12 hours (25°C to 65°C at 0.5°C/min, 65°C for 12 hours, 65°C to 25°C at 0.2°C/min).

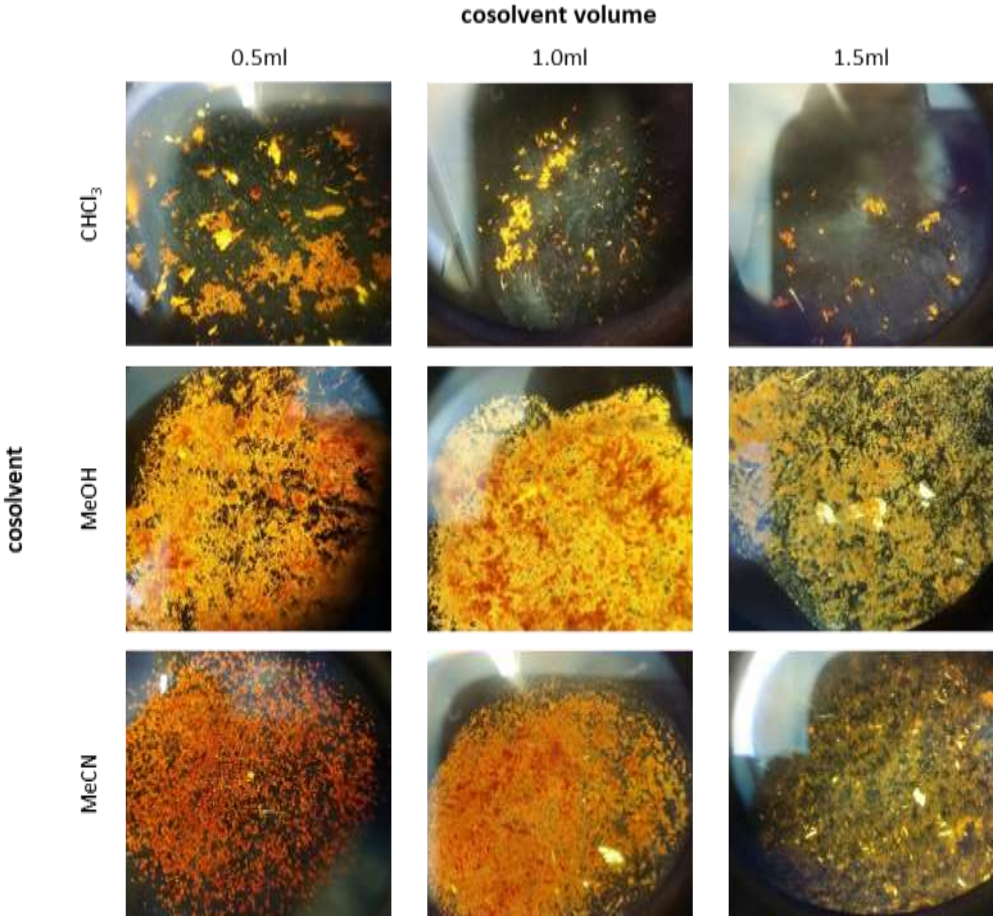


Figure S3. *Crystals observed at the optic microscope using polarized light.*

S4. SEM images

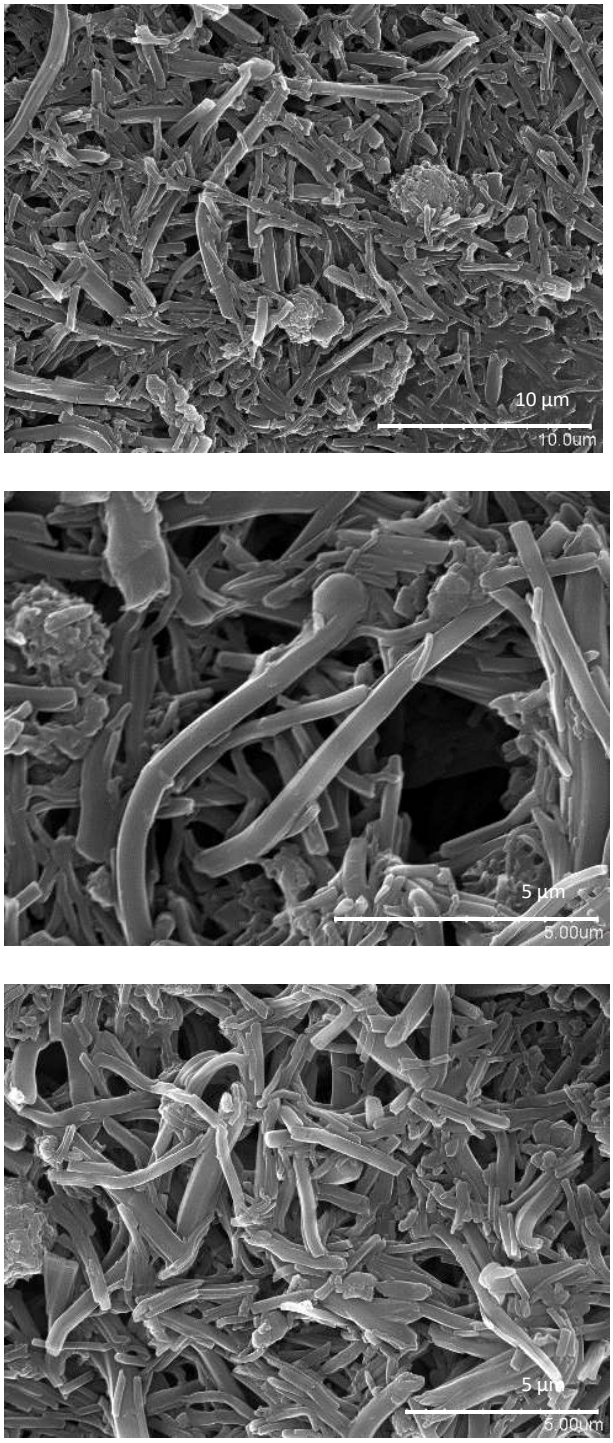


Figure S4. *SEM images of the* ***GCGC*** *crystals obtained using 0.5 mL of MeOH as cosolvent.*

S5. FT-IR analysis

The FT-IR spectrum of the crystals (in red in Figure S5) was compared with a 1:1 mixture of both **G** and **C** nucleobase derivatives. In the 3000-3500 cm^{-1} region a band centered in 3125 cm^{-1} can be seen while, in the 1500-1800 cm^{-1} region, a couple of bands at 1650 and 1640 cm^{-1} can be detected. Those bands are not present in the nucleobases alone or in the protected fused monomer, but are present in the 1:1 nucleobase mixture, so they could be assigned to N-H and C=O vibrations of H-bonded species.

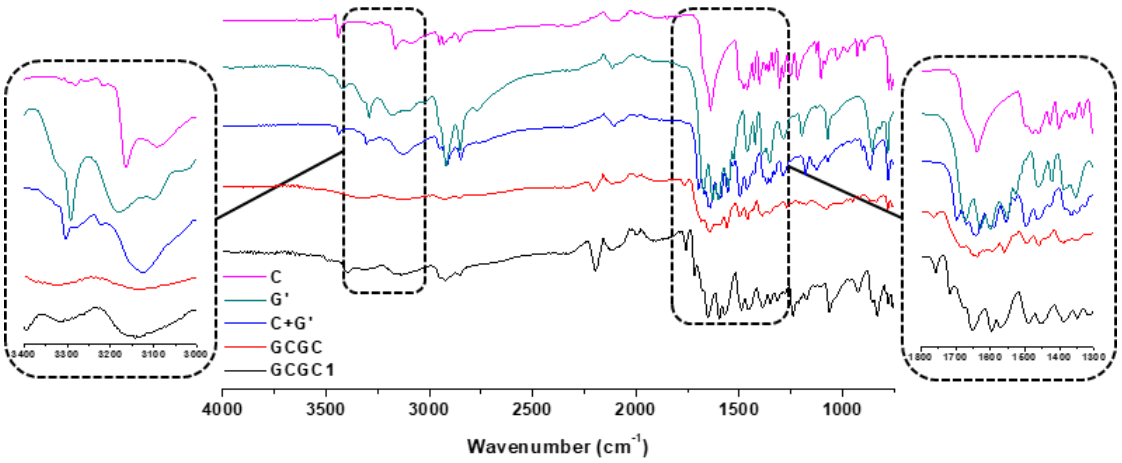


Figure S5. *FT-IR spectra of (top to down) cytosine derivative **C**, guanosine derivative **G**, 1:1 mixture of **G** and **C** (prepared in solution and then evaporated), fused monomer **GCGC** and protected fused monomer **GCGC1**.*

S6. Powder DRX analysis

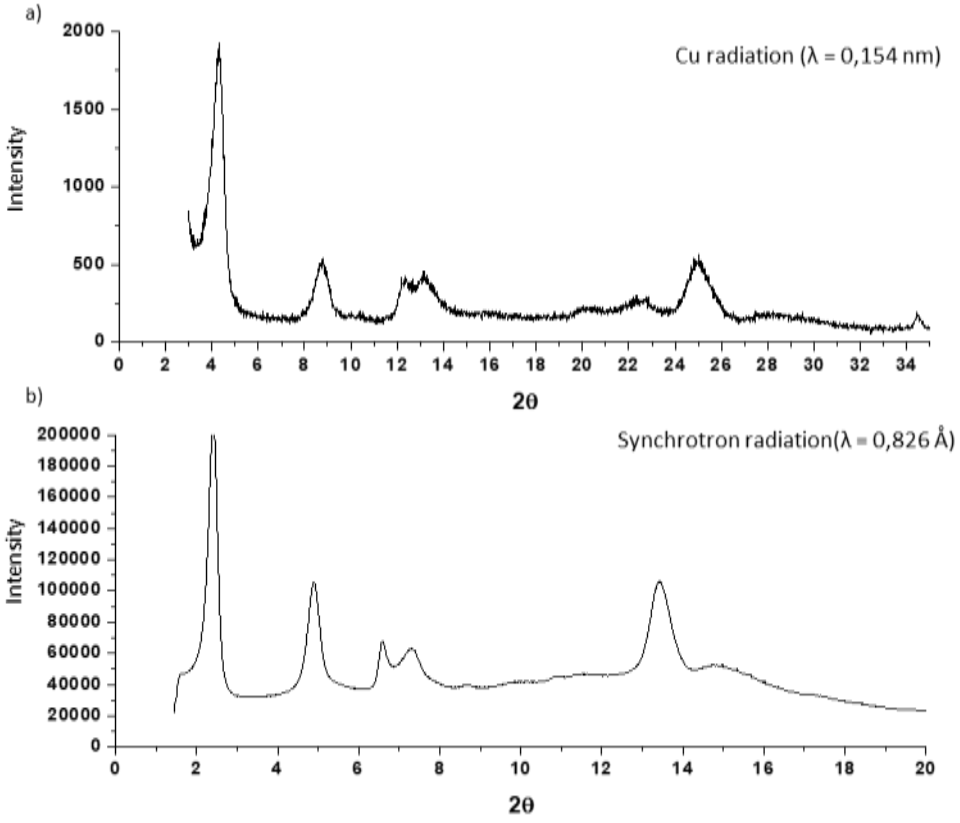


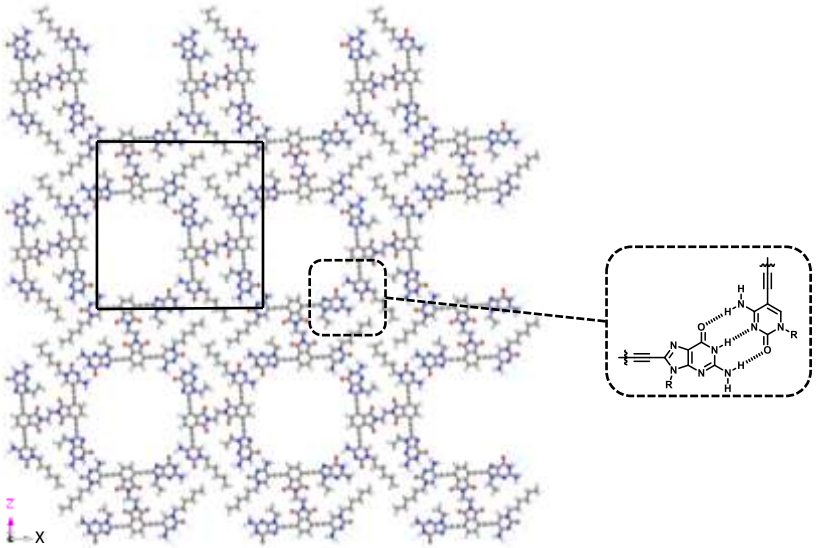
Figure S6. *PXRD diffractograms of* ***GCGC*** *crystals using Cu (up) and synchrotron (down) radiation.*

S7. Theoretical study

In order to elucidate the crystal structure, we followed a similar strategy to that followed by Wasielewski *et al.*[5] using *Materials Studio (MS) 2017 R2*. Given the presence of a peak corresponding to a possible pi-pi stacking (3.5 A) we delimited the study to 2D networks, for which we studied different types of packing as well as different H-bonding interactions between the nucleobases.

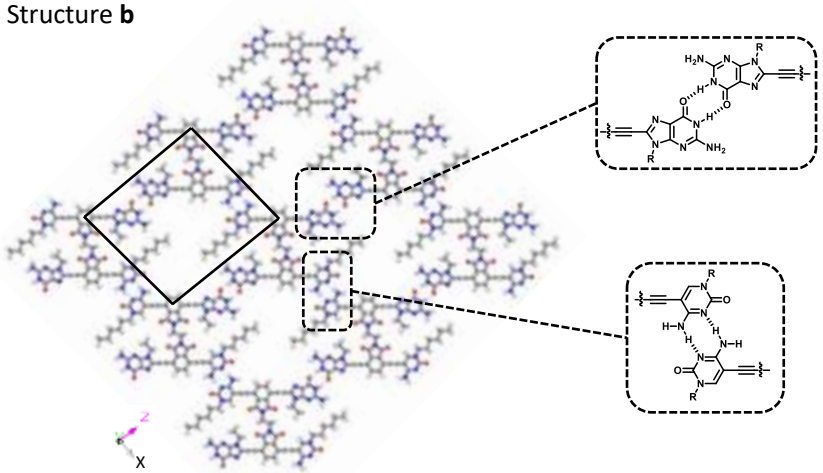
The first step was thus the construction of a 2D network. With this purpose, the fused monomer **GCGC** was constructed and optimized using the *Forcite* module with the DREIDING[6] force field where the charges of the atoms were obtained via Qeq calculations. Then, the monomer was assembled in different ways to obtain various 2D networks via hydrogen bond. Those networks were also optimized with the Forcite module to optimize short-range interactions such as hydrogen bonds. We studied three different possibilities (Figure S7A), the cyclic tetramer network (Figure S7Aa), one in which interactions between cytosines in one hand and guanines in the other hand are established (Figure S7Ab), and a third one with small cyclic tetramers similar to the G-quadruplex that forms a very compact network (Figure S7Ac). All 2D networks were undergone to a screening with different relative dispositions of the layers, always with a distance of 3.5 Å between them, because it was the distance estimated experimentally from the XRD pattern. We simulated the PDRX patterns of the three different possibilities in a simply ···AA··· stacking and in a ···AB··· stacking with a 5 Å translation (Figure S7B). Comparison of the simulated PXRD patterns with the experimental data allowed us to discard a cyclic tetramer H-bonding network due to the low similarity with the experimental data. We also discarded the third possibility (Figure S7Ac) given the high steric hindrance between the alkyl chains once the network is formed. In the second case, both simulations fitted with the experimental data. Therefore, we focused on this conformation to carry out our structural study.

Structure **a**



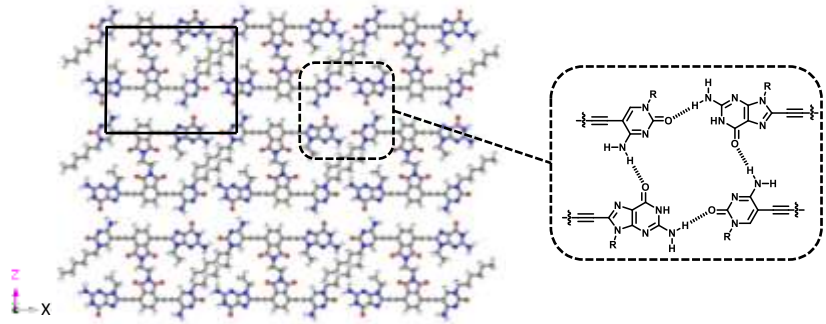
a = c = 33.5 Å
β = 90°

Structure **b**



a = 24.9 Å
c = 22.4 Å
β = 97°

Structure **c**



a = 18.0 Å
c = 22.6 Å
β = 91°

Figure S7A. *Different possibilities studied in the formation of a 2D network.*

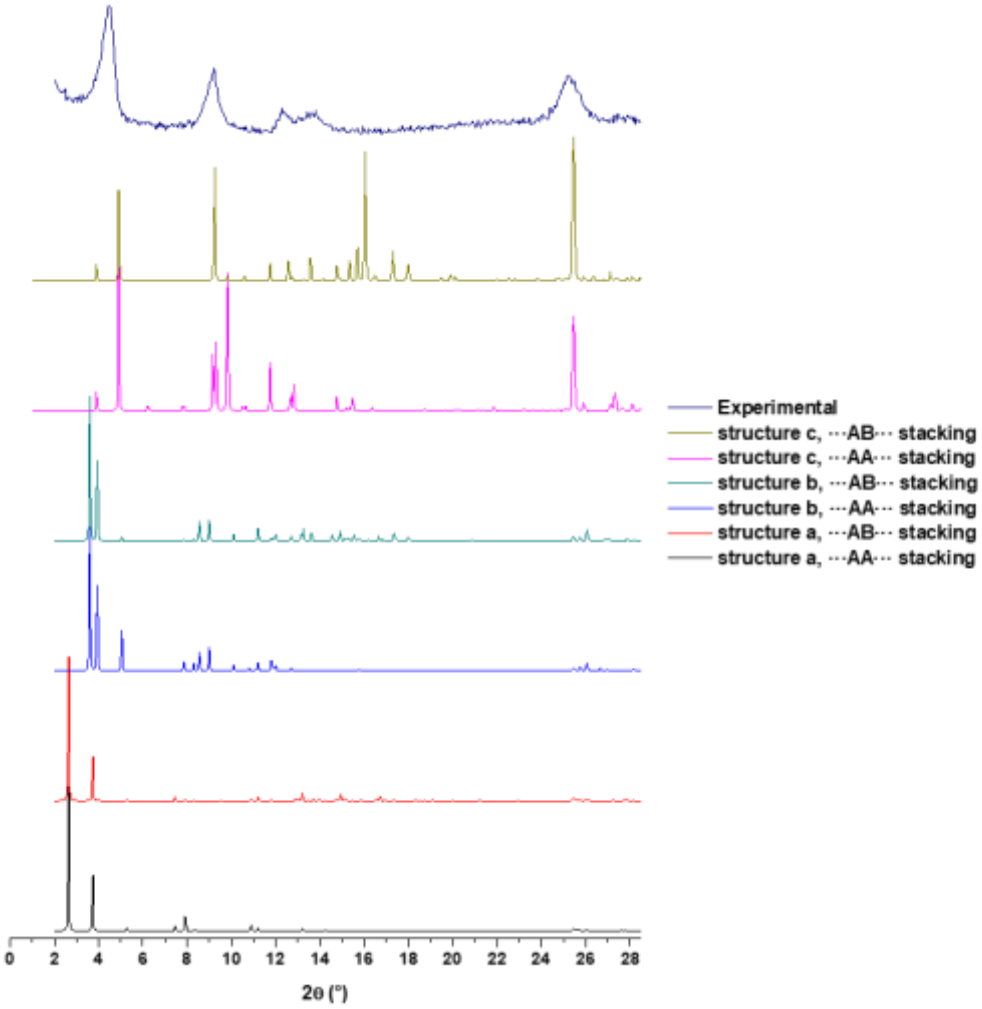


Figure S7B. Simulated diffractograms of the 3 proposed networks with a ···AA··· staking and with a 5 A side shift ···AB··· stacking.

In addition to the networks studied previously for the second option (Figure S7b), in the case of AB stacking, different shifts in the x and y axes were tried (Figure S7C). Systematic translations of 5, 10 and 15 Å allowed us to see the trend in the simulated XRD patterns. Thus, a big translation in the *c* axis was necessary to move the principal peak to the distance showed in the experimental one. In the other hand, the movement in the *a* axis led to minor improvements. Regarding the packing of the layers, different stacking possibilities were studied as a function of the number of layers as well as the formation of 1D channels along the *b* axis (Figure S7D). However, the simulated diffractograms of the latter case differed too much from the experimental data and, thus, no porous structures were further considered. Eventually, this information led us to a possible structure whose simulated PXRD was in accordance with the experimental data. In this structure, the second layer was shifted 12 and 18 Å in the *a* and *c* axes, respectively (Figure S7E).

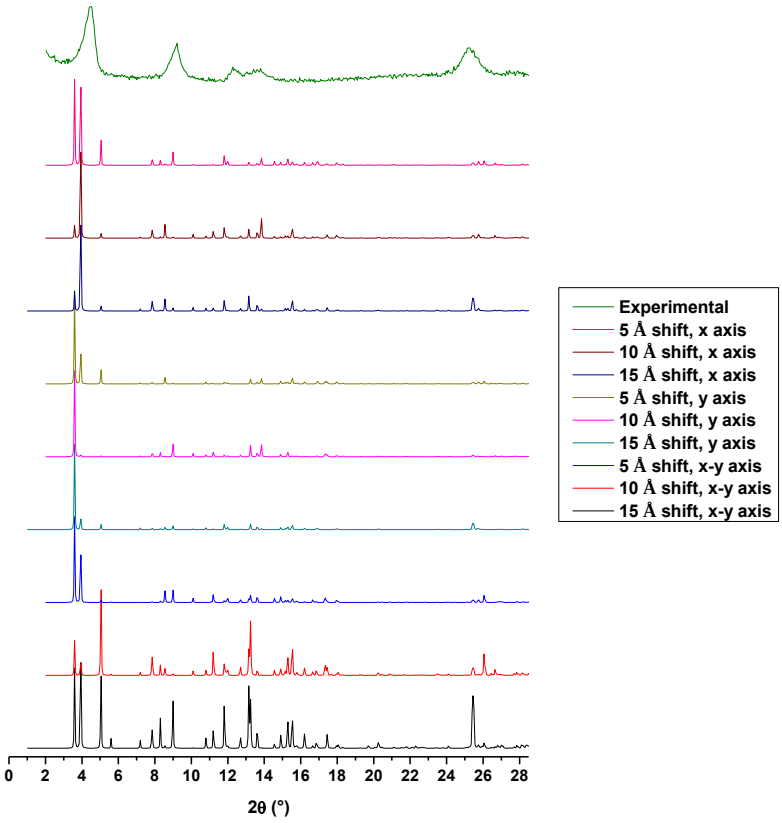


Figure S7C. Simulated diffractograms of the **b** network with a 5, 10 or 15 Å side shift ···AB··· stacking.

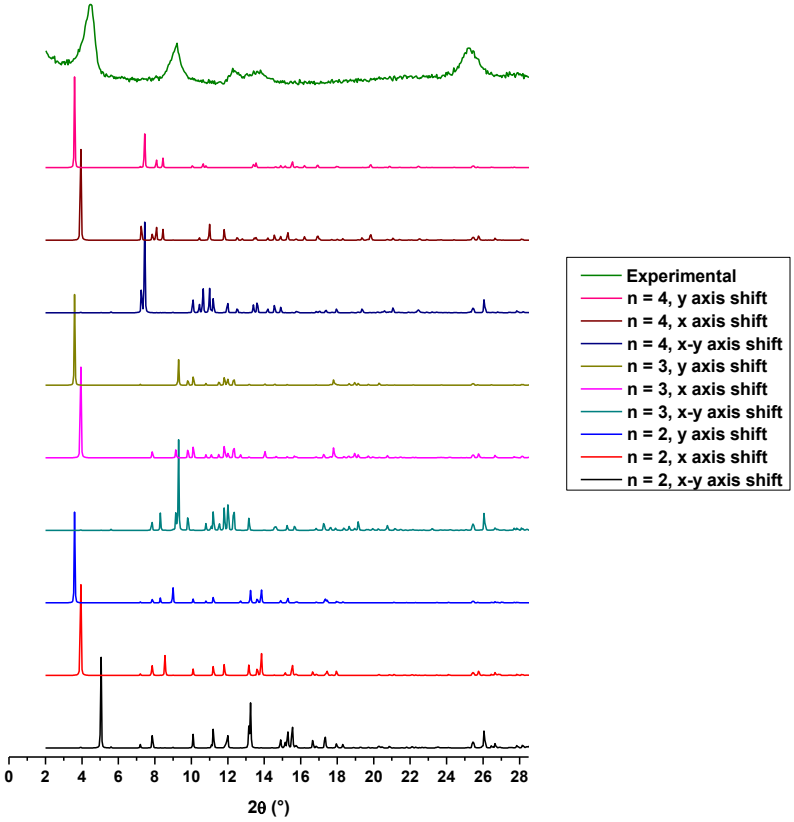


Figure S7D. Simulated diffractograms of the **b** network with shifts in function of the number of layers in the stacking.

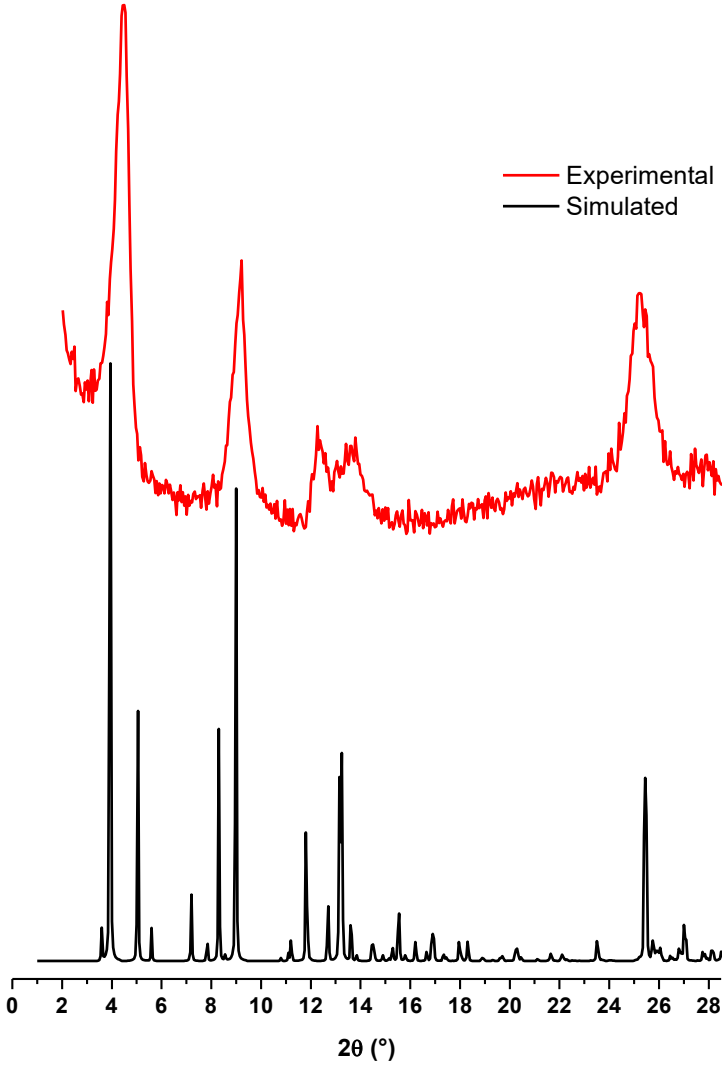


Figure S7E. *Comparison between the experimental PDRX diffractogram and the simulated from the proposed structure.*

Once we obtained a probable structure, we carried out a profile fitting of the PXRD data in order to evaluate the feasibility of the model. Whereas previous models yielded poor fittings, the profile fitting refinement of the model used in Figure S7E converged with excellent residual values for a monoclinic crystal system of cell parameters a = 22.1582 Å, b = 7.2195 Å, c = 19.3841 Å, β = 97.085°, V = 3077.3 Å^3, Rp = 2.23 %, Rwp = 2.55 % Rexp = 1.53 %, Chi2 = 2.76 (Figure S8). However, due to the broad diffraction peaks, it was not possible to estimate the most likely space group from the diffraction data and thus, the refinement was carried out in the triclinic *P1* space group, keeping $\alpha = \gamma = 90°$.

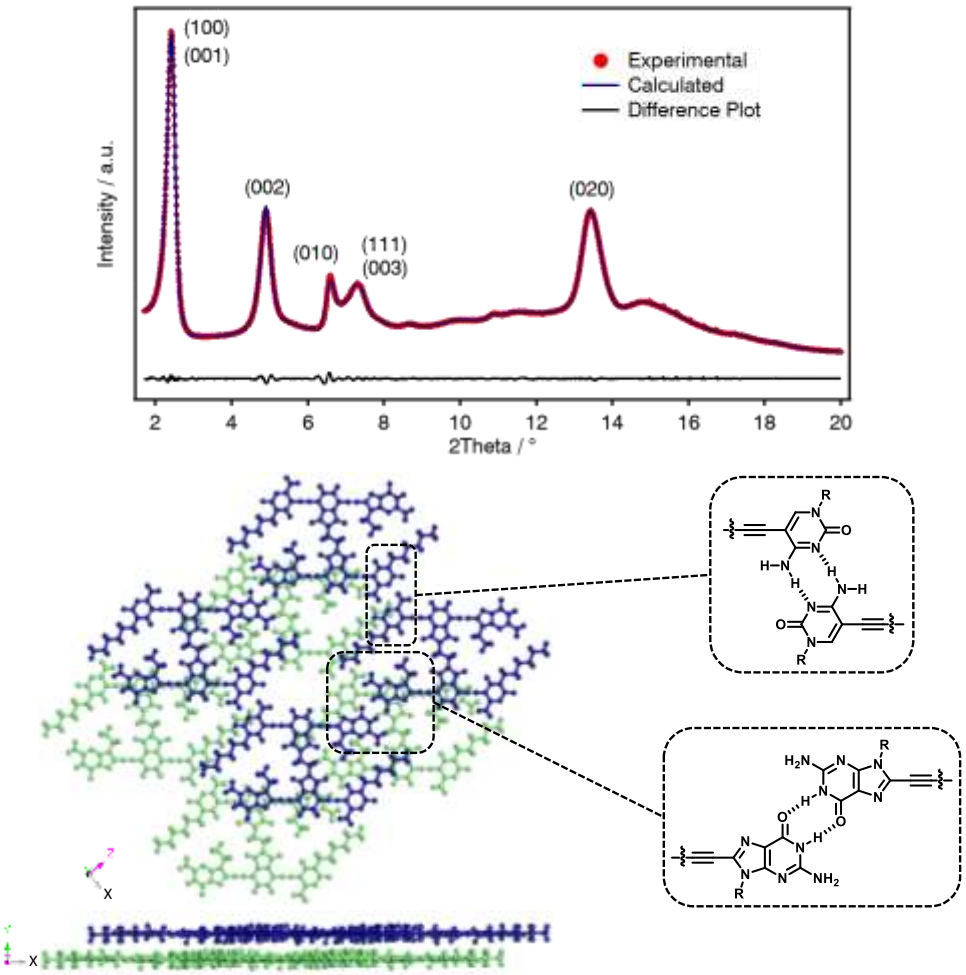


Figure S7F. *LeBail profile fitting of the experimental PDRX pattern (synchrotron radiation) (up) proposed structure with C:C and G:G interactions (down)*

S8. Gas absorption measurement

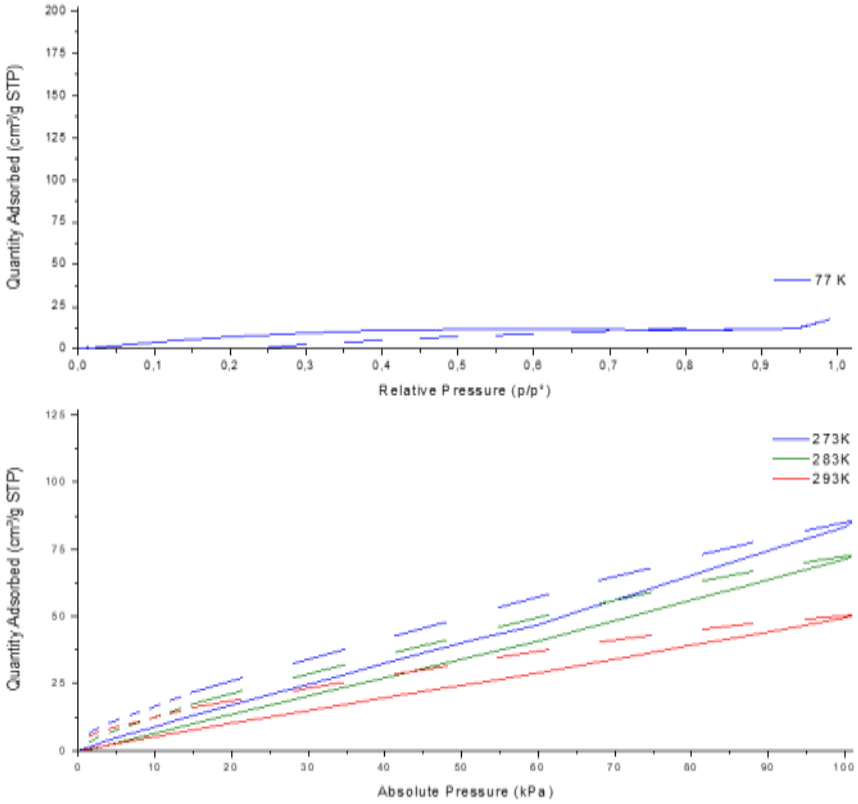


Figure S8. *Absorption isotherms of N2 (up) and CO_2 (down).*

S9. Thermogravimetric analysis

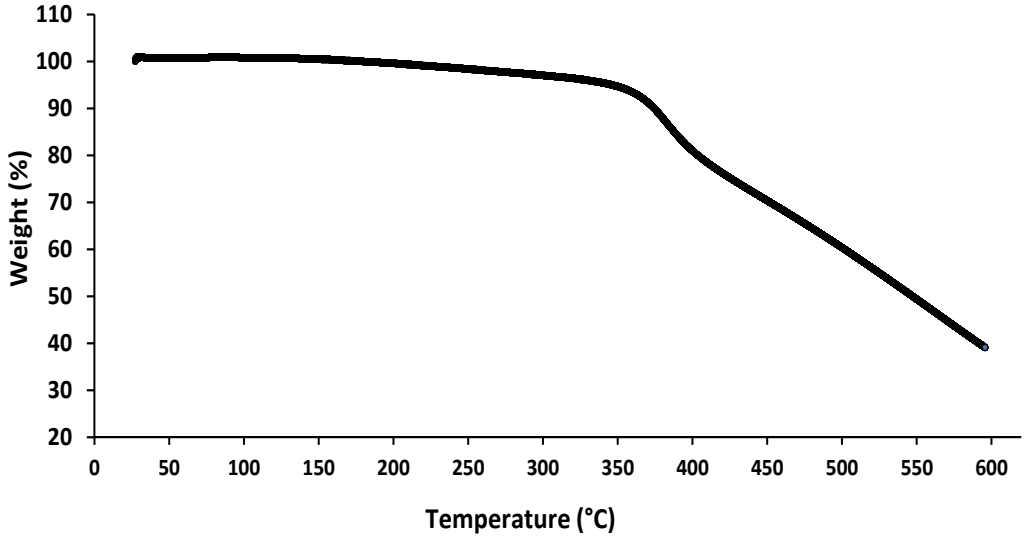


Figure S9. *TGA of the* ***GCGC*** *fused monomer under N_2 atmosphere.*

References

1 J. Rodríguez-Carvajal, *Phys. B Phys. Condens. Matter*, 1993, **192**, 55–69.

2 N. Bilbao, V. Vázquez-González, M. T. Aranda and D. González-Rodríguez, *European J. Org. Chem.*, 2015, **2015**, 7160–7175.

3 E. Camaioni, S. Costanzi, S. Vittori, R. Volpini, K. N. Klotz and G. Cristalli, *Bioorganic Med. Chem.*, 1998, **6**, 523–533.

4 P. Solanke, F. Bureš, O. Pytela, J. Kulhánek and Z. Padělková, *Synth.*, 2013, **45**, 3044–3051.

5 Y. L. Wu, N. E. Horwitz, K. S. Chen, D. A. Gomez-Gualdron, N. S. Luu, L. Ma, T. C. Wang, M. C. Hersam, J. T. Hupp, O. K. Farha, R. Q. Snurr and M. R. Wasielewski, *Nat. Chem.*, 2017, **9**, 466–472.

6 S. L. Mayo, B. D. Olafson and W. A. Goddard, *J. Phys. Chem.*, 1990, **94**, 8897–8909.