molbank

Supporting information for

4-(2-(5-(2-(*tert*-Butoxycarbonyl)hydrazinecarbonyl)-2-methylthiophen-3-yl)cyclopent-1-enyl)-5-methylthiophene-2-carboxylic Acid

Marija Matković

Division of Organic Chemistry and Biochemistry, Laboratory for Biomolecular Interactions and Spectroscopy, Ruđer Bošković Institute, Bijenička 54 10000 Zagreb, Croatia
Correspondence: Marija.Matkovic@irb.hr

Content:

Figure S1: ^{1}H NMR (300 MHz, (CD$_3$)$_2$CO) spectrum of **2** isolated by extraction. Blue arrows depict traces of EtOAc.

SpinWorks 3: Compound 2
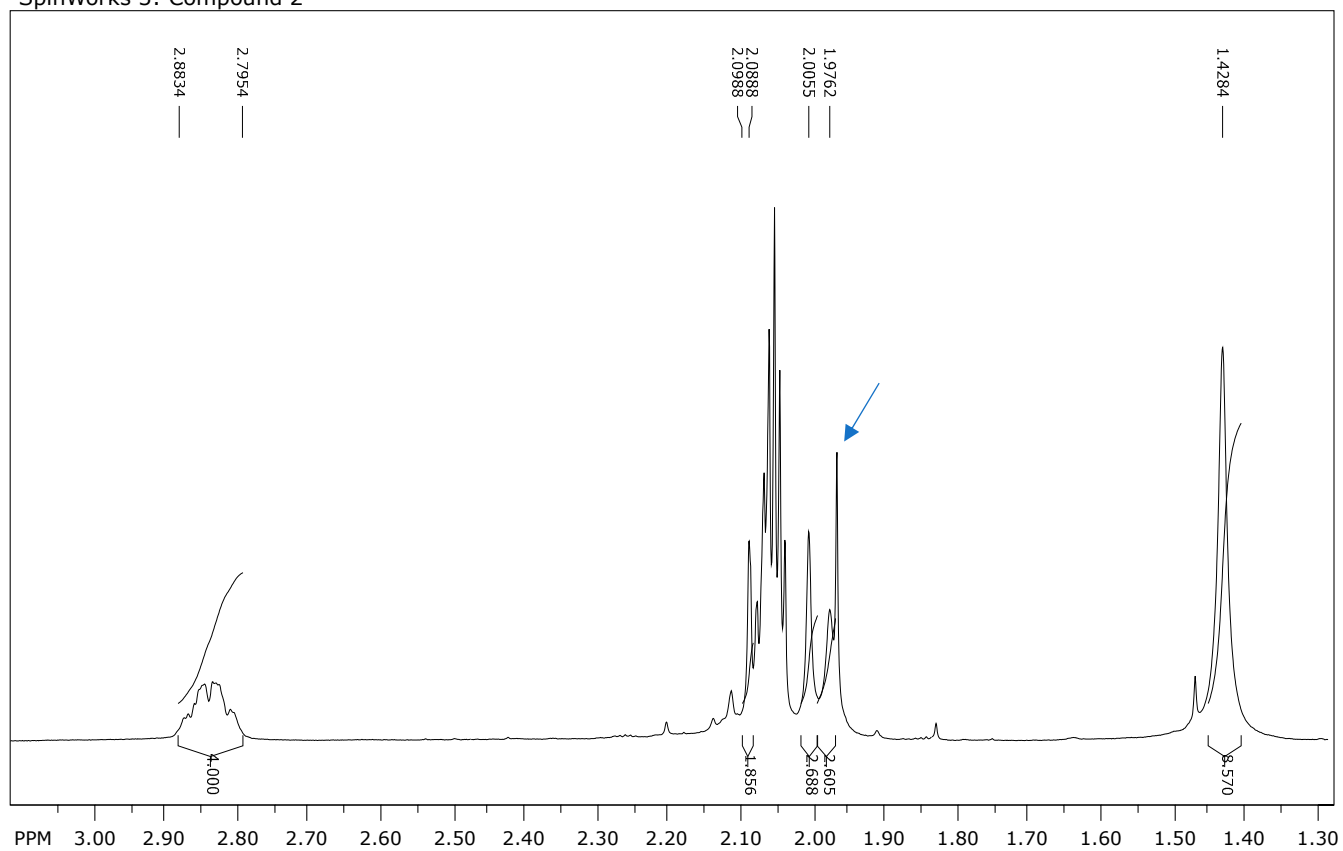
Figure S2: ^{1}H NMR (300 MHz, (CD$_3$)$_2$CO) spectrum of **2** (zoomed) isolated by exstraction. Blue arrow depicts traces of EtOAc.

SpinWorks 3: Compound 2
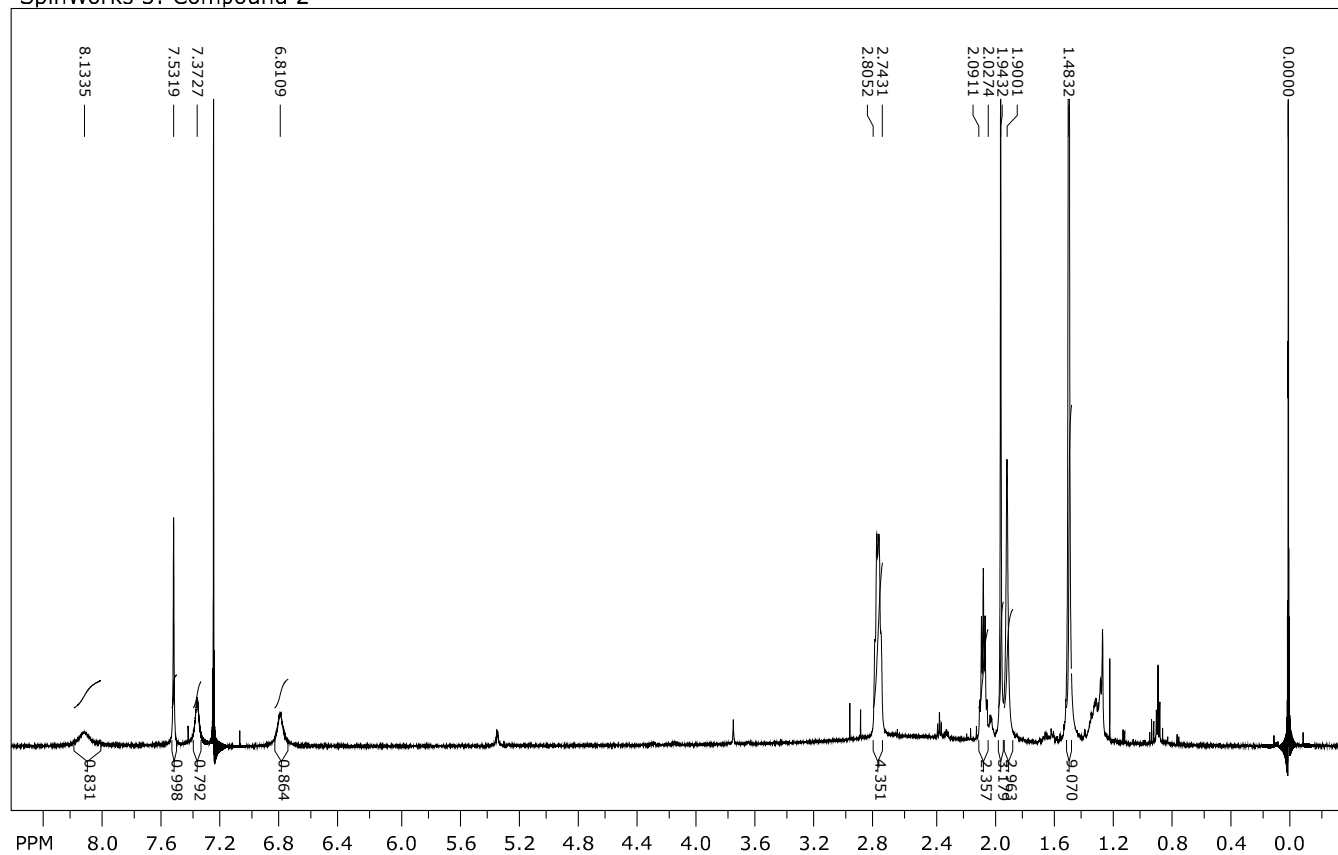
Figure S3: ^{1}H NMR (600 MHz, CDCl$_3$) spectrum of **2** isolated by preparative TLC (normal phase, 7% MeOH-CH$_2$Cl$_2$). From 0.8-1.4 ppm some contaminants are recognizable.

SpinWorks 3: Compound 2

PPM   150   140   130   120   110   100   90   80   70   60   50   40   30   20   10   0
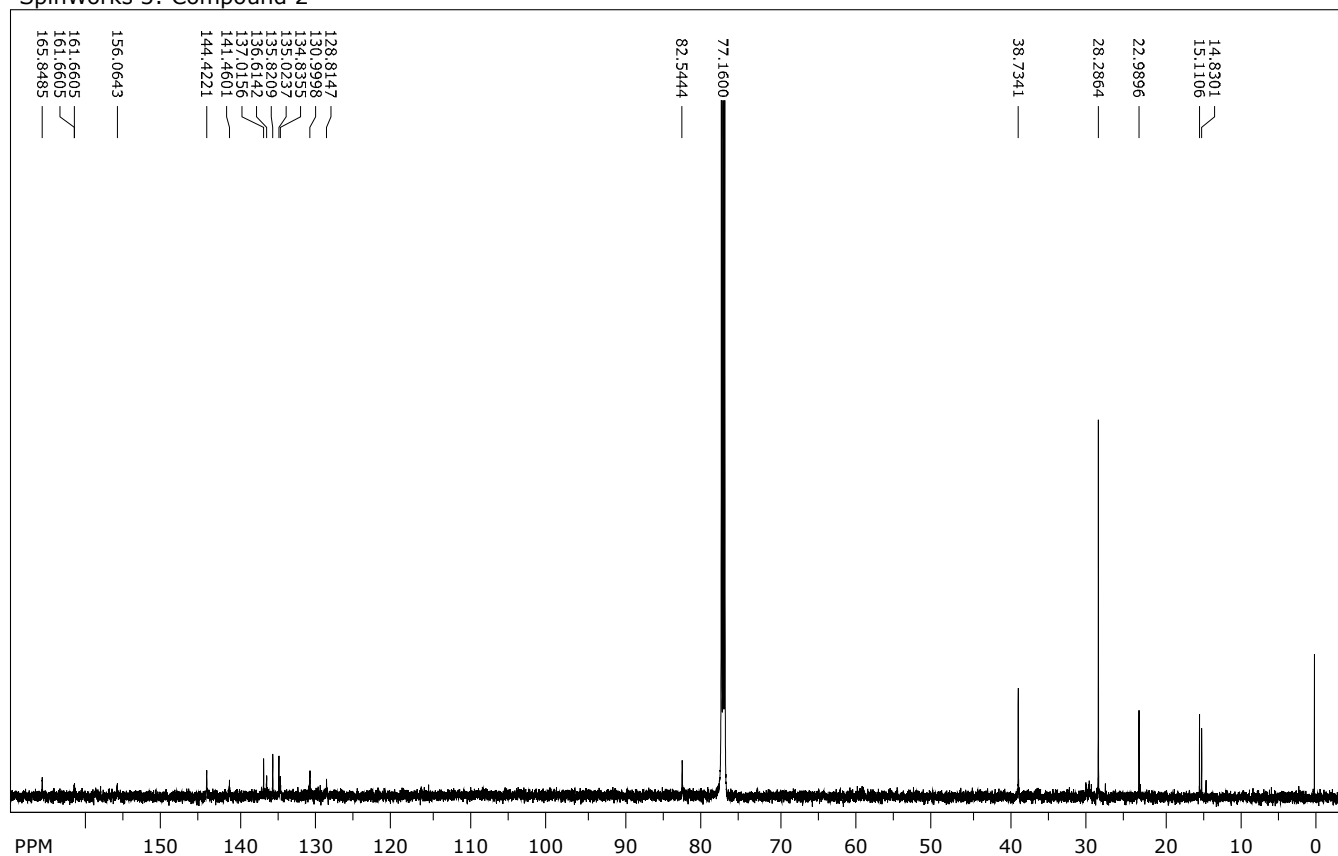
Figure S4: ^{13}C NMR (150 MHz, CDCl$_3$, COMPLETE) spectrum of **2** isolated by preparative TLC (normal phase, 7 % MeOH-CH$_2$Cl$_2$)

SpinWorks 3: Compound 2
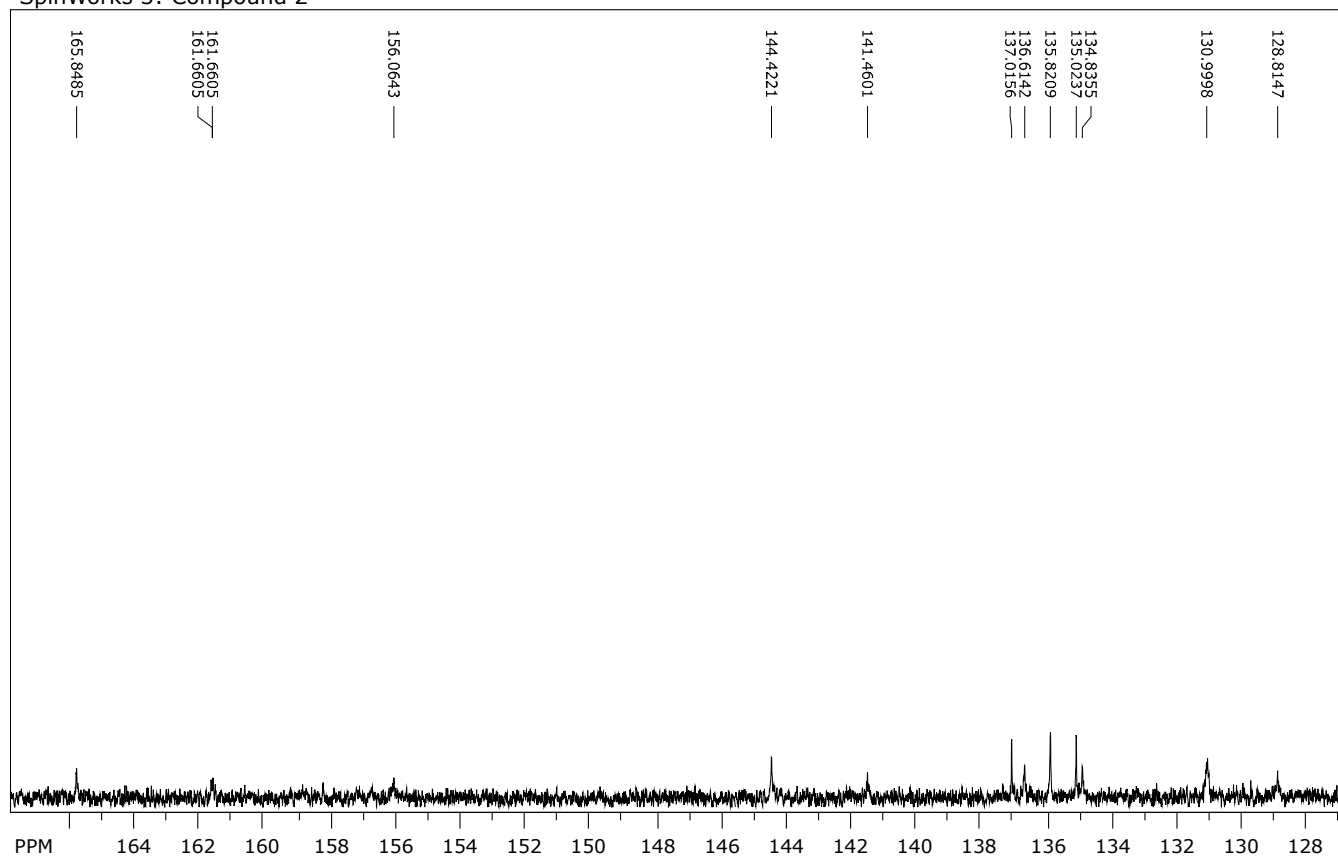
Figure S5: ^{13}C NMR (150 MHz, CDCl$_3$, COMPLETE) spectrum of **2** (zoomed) isolated by preparative TLC (normal phase, 7% MeOH-CH$_2$Cl$_2$). This section represents only 12 different signals. this is beacuse atoms Cd and Ce as well as Cb and Cc (FigureS6) are present in not enough different chemical environment to be explicitly depicted as different signals (FigureS22).

SpinWorks 3: Compound 2
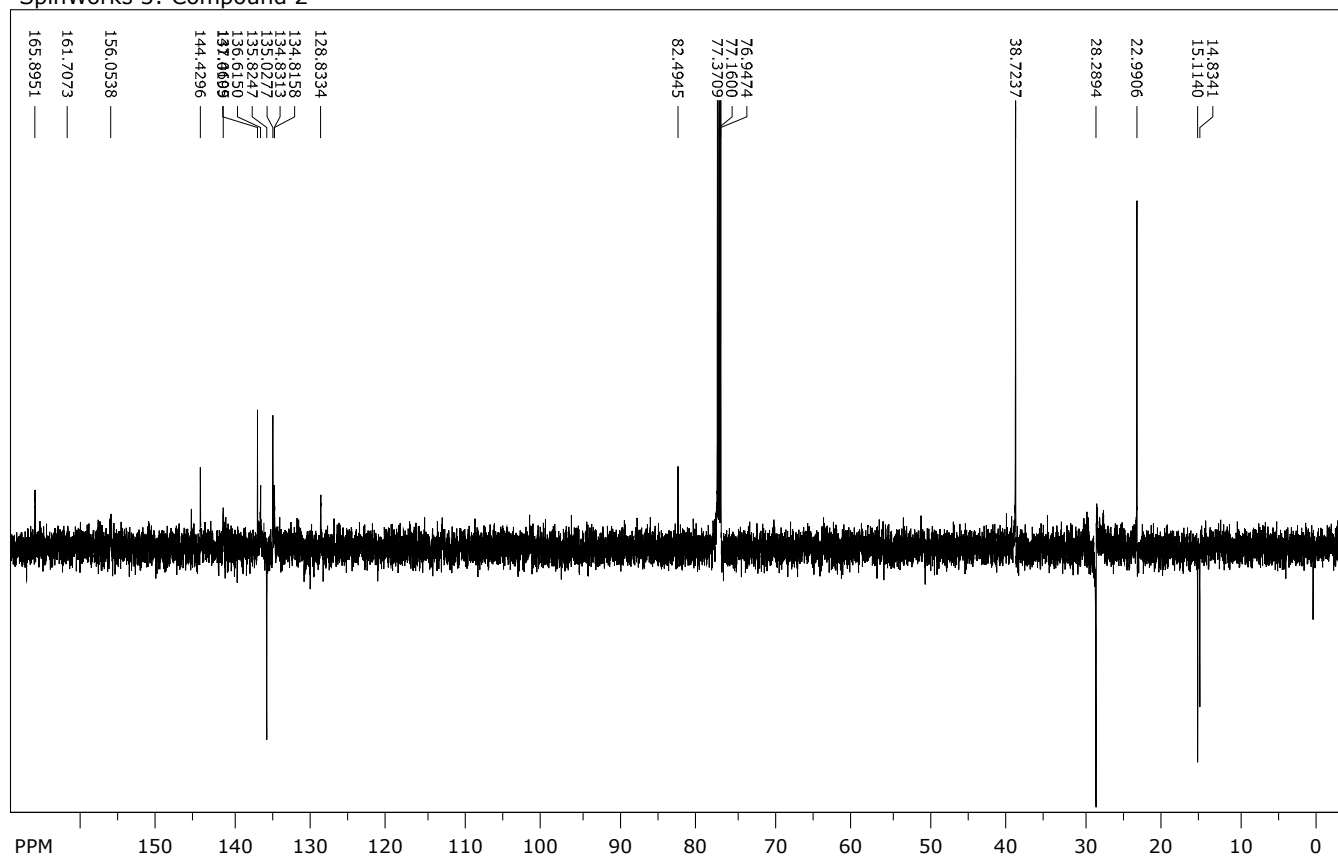
Figure S6: ^{13}C NMR (150 MHz, CDCl$_3$, APT) spectrum of **2** isolated by preparative TLC (normal phase, 7 % MeOH-CH$_2$Cl$_2$)

SpinWorks 3: Compound 2
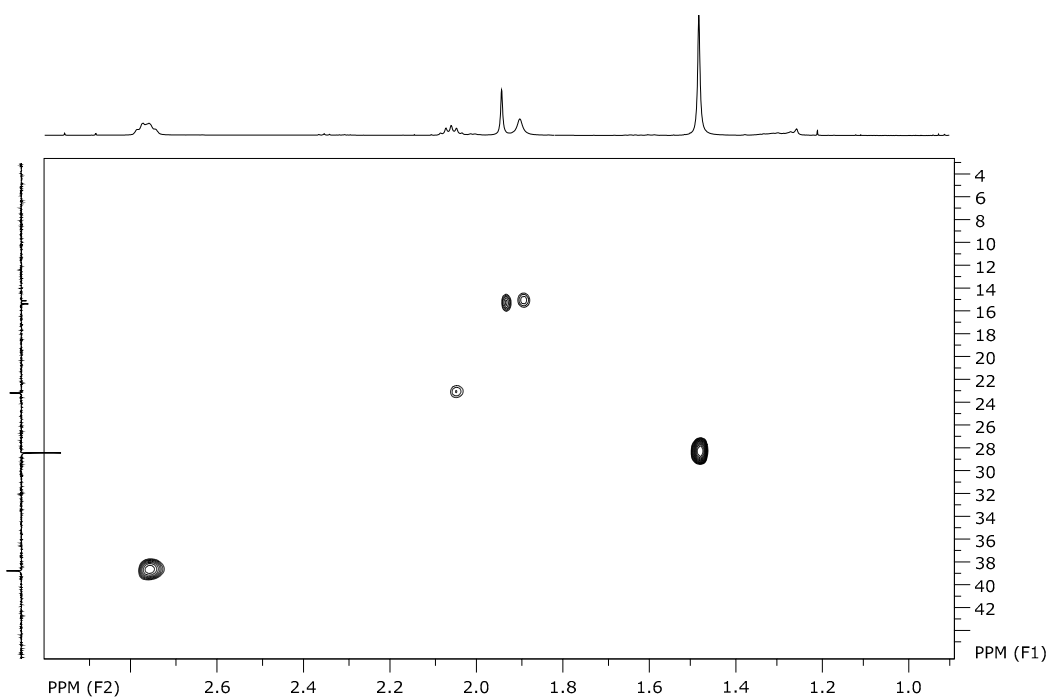
Figure S7a

SpinWorks 3: Compound 2
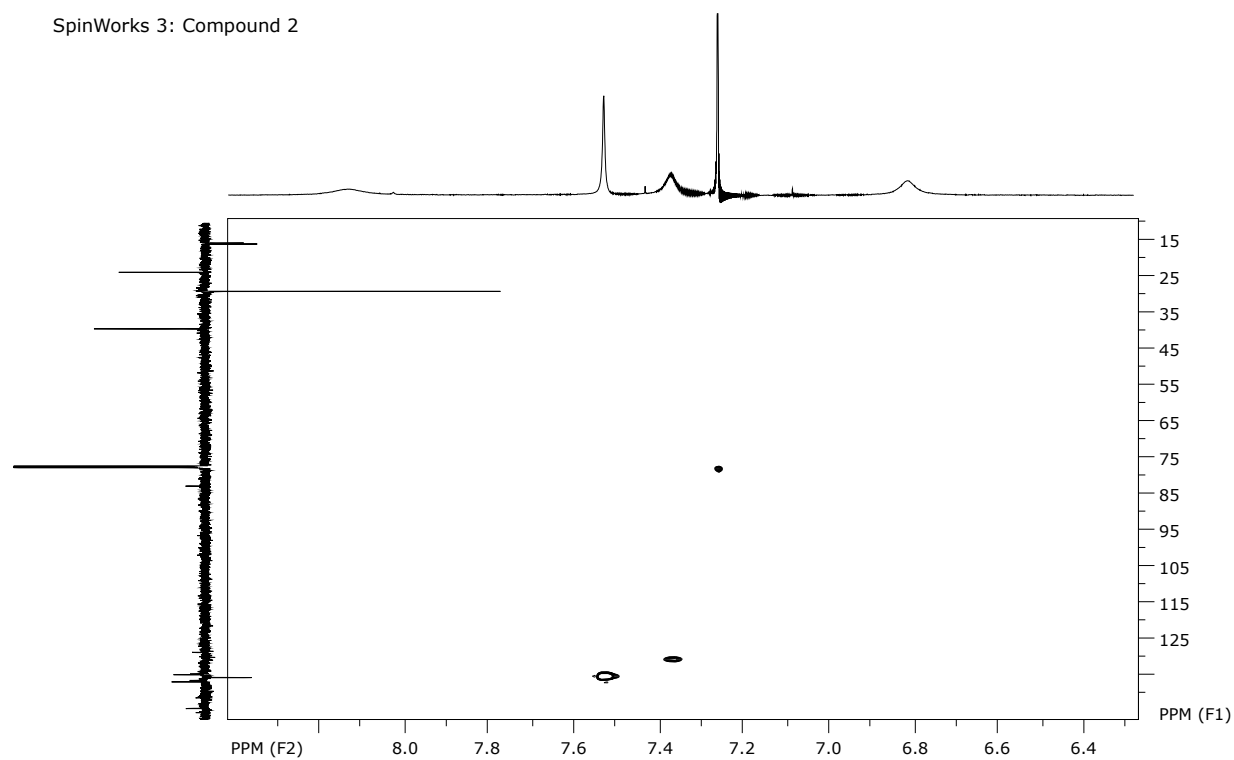
Figure S7b

Figure S7 (a-b): HSQC spectrum (600 MHz, CDCl$_3$) of compound **2**: a) up to 2.8 ppm (^{1}HNMR) b) from 2.8 ppm to 10 ppm (^{1}HNMR)

SpinWorks 3: Compound raw 3
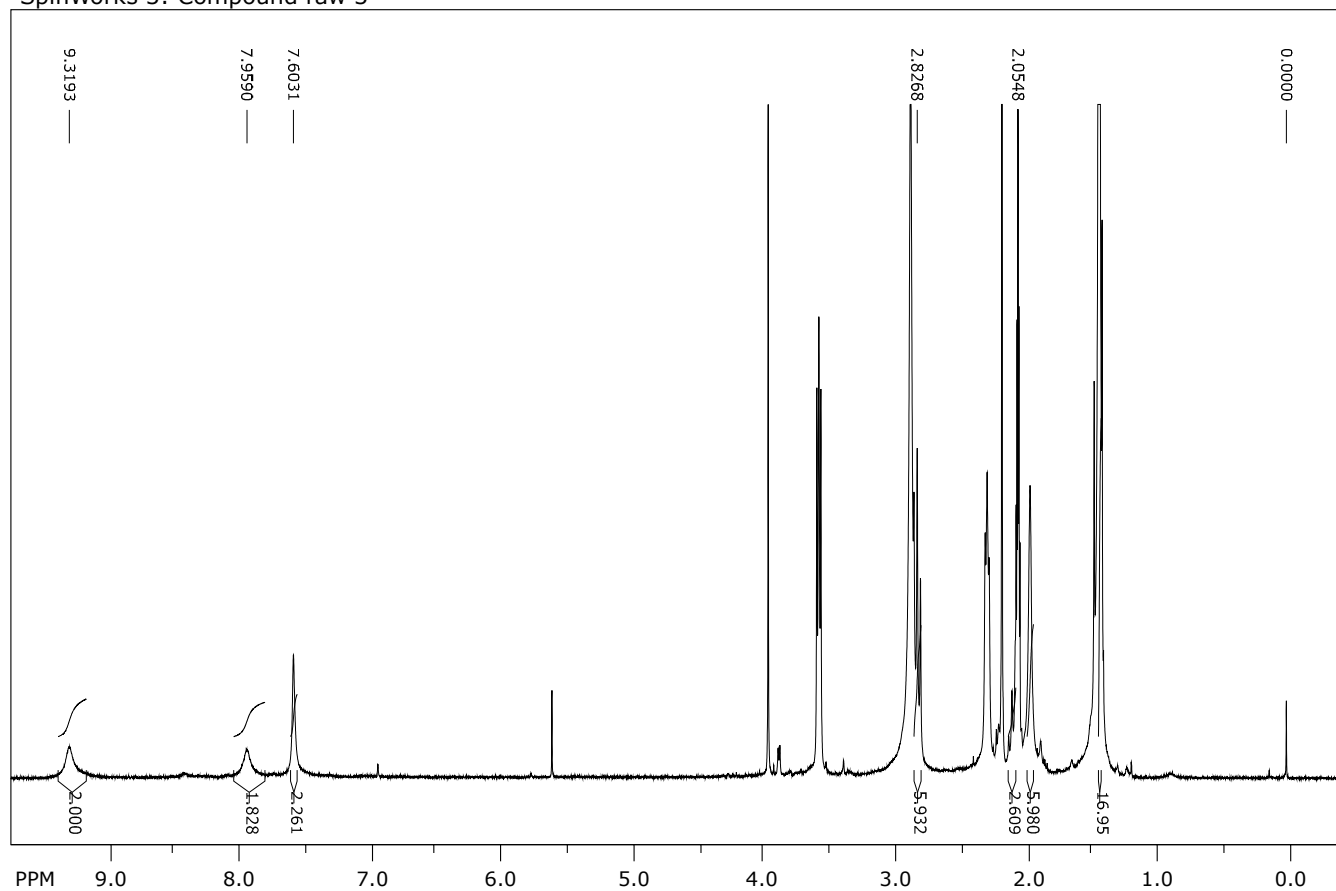
Figure S8: ^{1}H NMR spectrum (300 MHz, (CD$_3$)$_2$CO) of raw **3** obtained by extraction (O1 organic phase, see experimental and Appendix A).

SpinWorks 3: Compound 3 raw
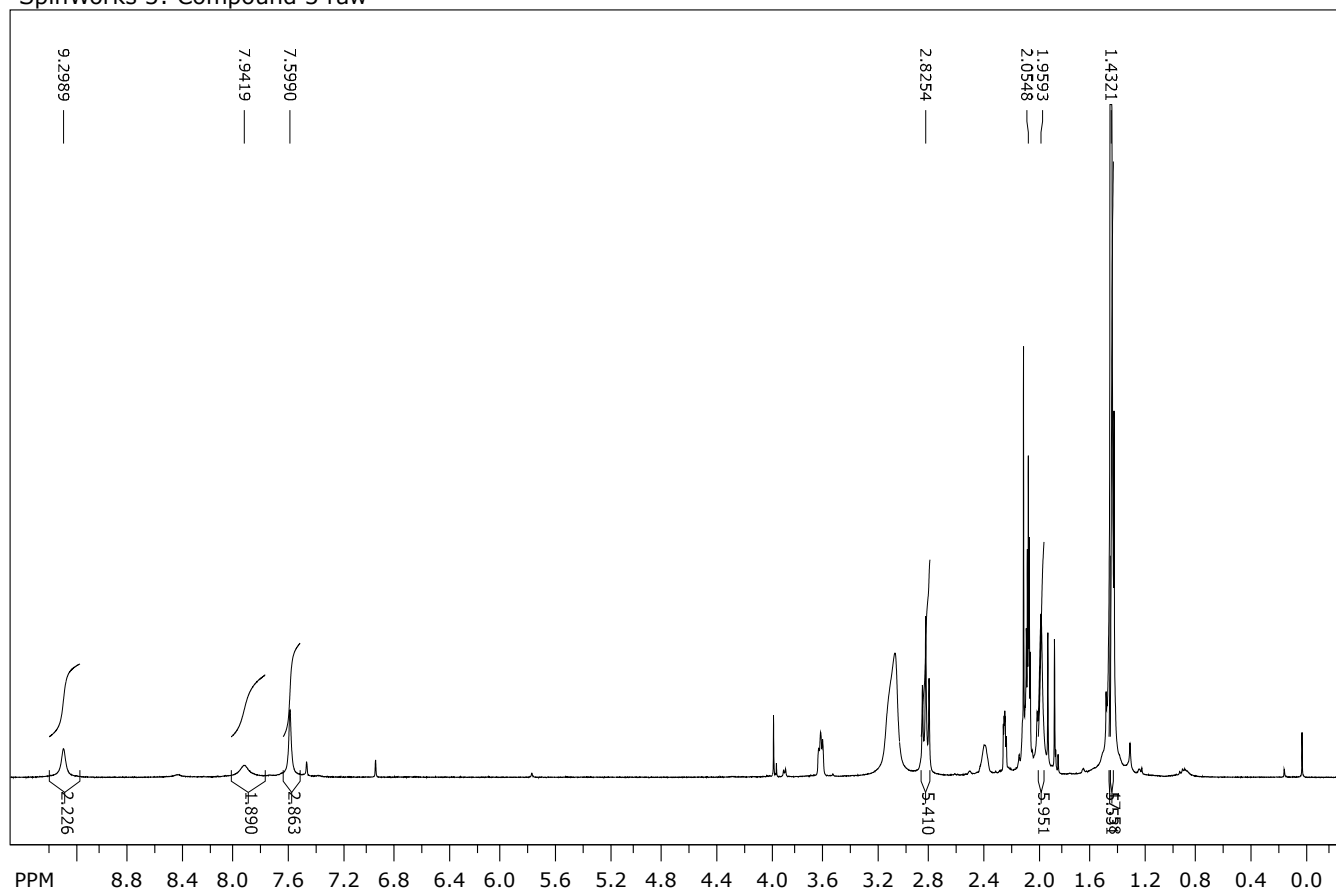
Figure S9: ^{1}H NMR spectrum (300 MHz, (CD$_3$)$_2$CO) of raw **3** obtained by extraction (O2 organic phase, see experimental and Appendix A).

SpinWorks 3: Mixture 2 + 3
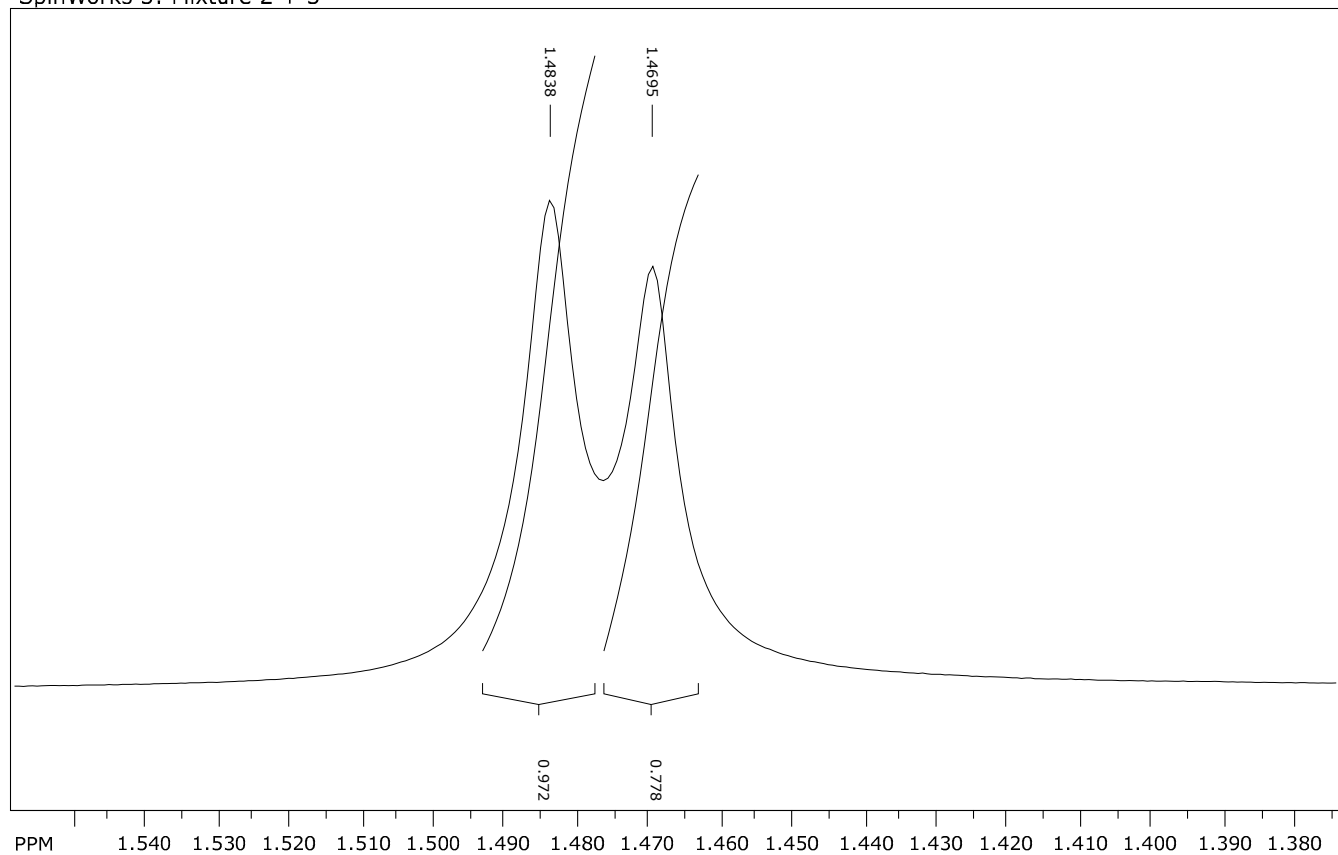
Figure S10: ^{1}H NMR spectrum (300 MHz, CDCl$_3$) of mixture **2 + 3** (10:4, Boc protecting group signals are shown) further purified.

Figure S11: Preparative TLC (7 % MeOH/CH$_2$Cl$_2$,) used for isolation of compounds from mixture **2+3**; the stripes are located (visualized) by exposing the edge part of the plate to UV irradiation (254 nm, 2-3 min)-the irradiated area becomes "dark pink" due to DAE unit cyclisation on spot. The unirradiated stripes of SiO$_2$ are separated acoording to "dark pink" regions.

Figure S12: Analytical TLC (7 % MeOH/CH$_2$Cl$_2$); left—mixture od compounds (**2+3**); right—compound **2** after purification by preparative TLC (7 % MeOH/CH$_2$Cl$_2$); the spots where vizualised by exposing the analy. plate to UV irradiation (UV lamp, 254 nm, 4 min.)

SpinWorks 3: Compound 3

Figure S13: ^{1}H NMR (600 MHz, (CD$_3$)$_2$CO) spectrum of **3** isolated by consecutive preparative TLC (normal phase, EtOAc:petroleter = 1:1, 5 % MeOH/CH$_2$Cl$_2$)

SpinWorks 3: Compound 3

206.7812
206.9359

162.7672
157.2813
142.0189
138.1788
136.4126
134.9639
131.2479

81.2486

39.9188
31.1077
30.9850
30.8559
30.7280
30.6000
30.4723
30.4408
30.3445
30.2156
29.1798
29.0714
24.2420
15.4618

PPM   220   200   180   160   140   120   100   80   60   40   20   0
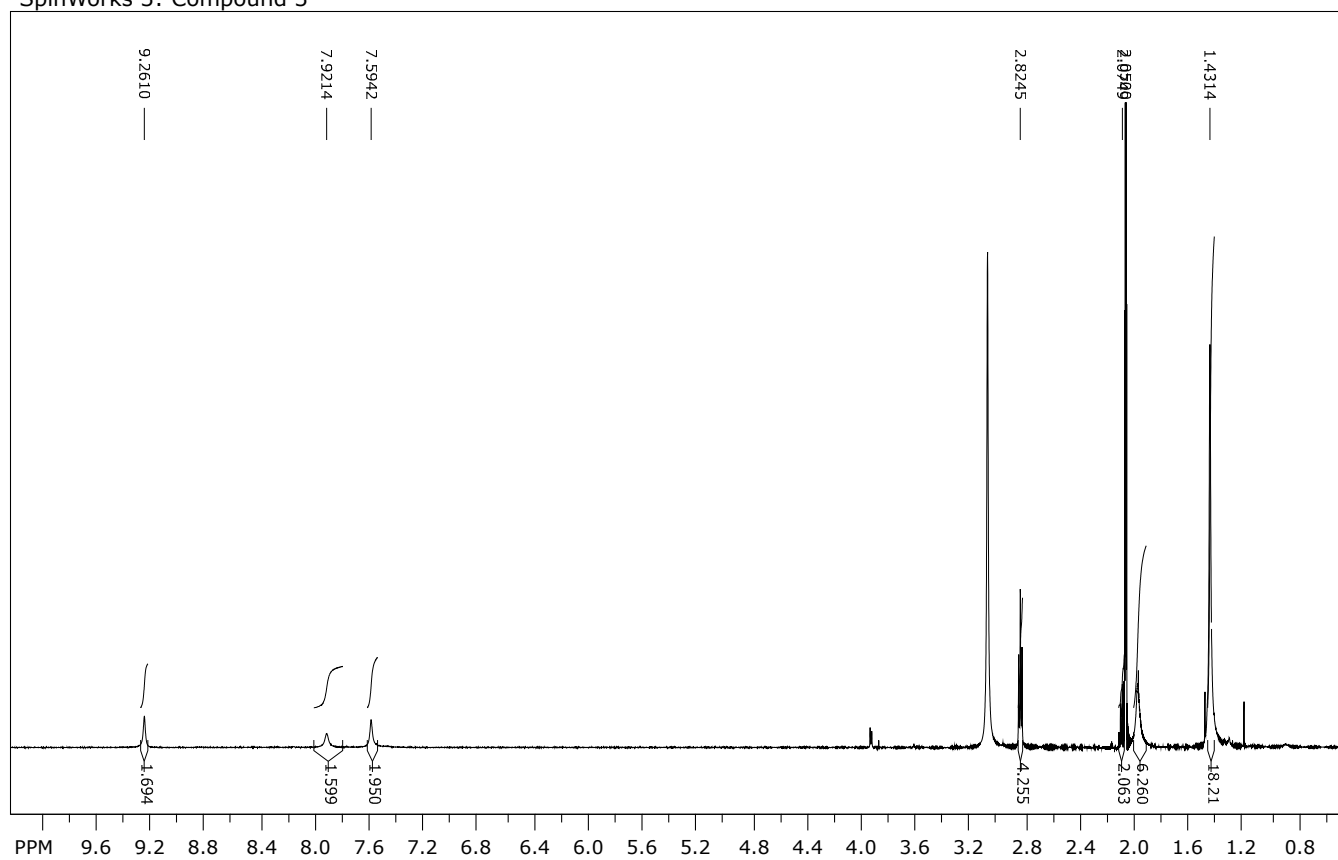
Figure S14: ¹³C NMR (150 MHz, (CD₃)₂CO, APT) spectrum of **3** isolated by consecutive preparative TLC (normal phase, EtOAc:petroleter = 1:1, 5 % MeOH/CH₂Cl₂
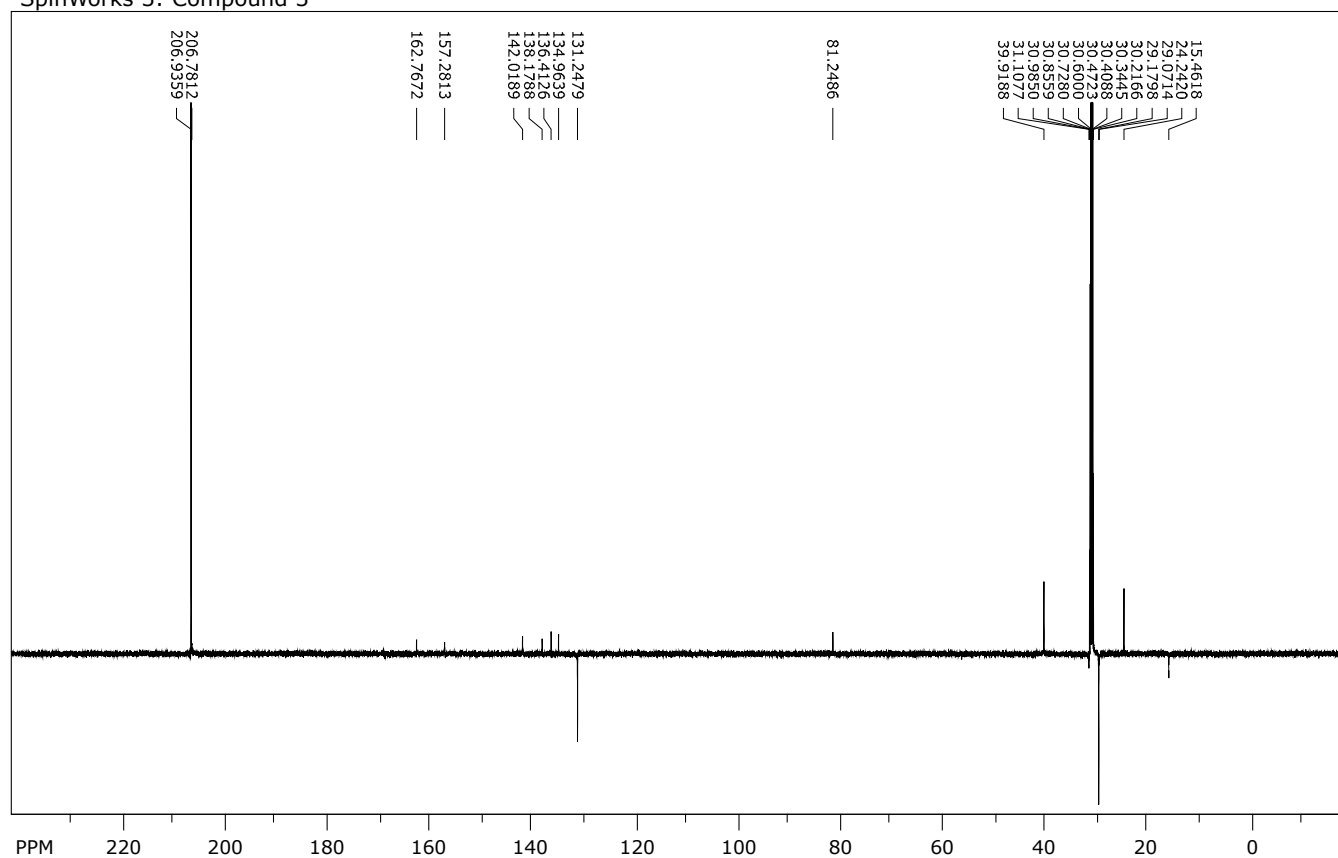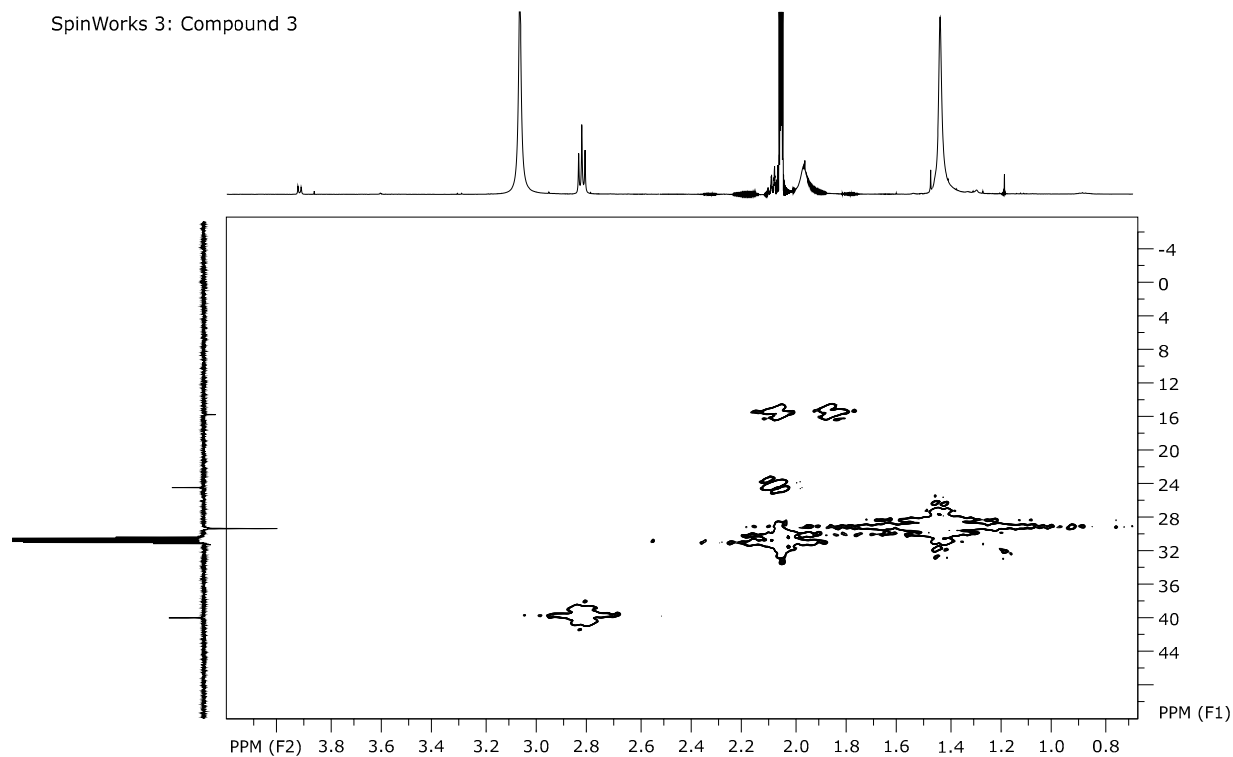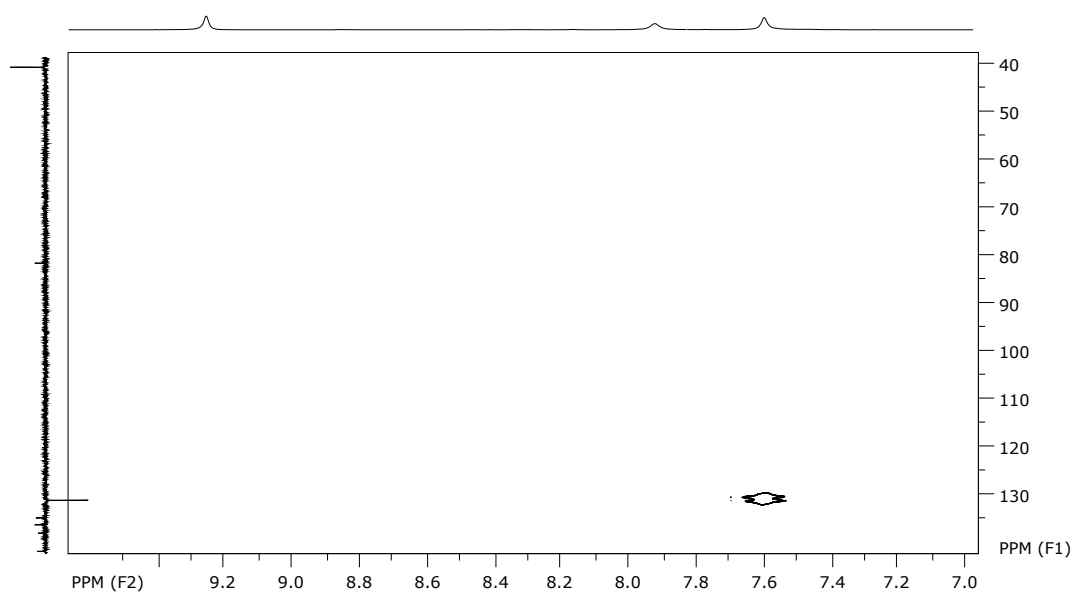
SpinWorks 3: Compound 3

a)

SpinWorks 3: Compound 3

b)

Figure S15 (a-b): HMQC spectrum (600 MHz, CDCl$_3$) of compound **3**: a) up to 4.0 ppm (^{1}HNMR) b) from 7.0 ppm to 9.6 ppm (^{1}HNMR)
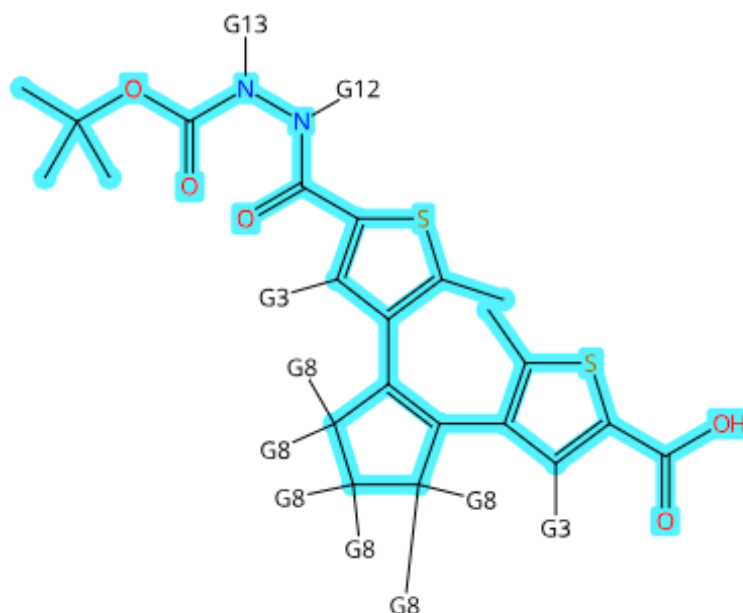
WO2014127919, Patent claim 6

Figure S16: Patent Markush (WO2014127919 A1) found (SciFinder) resembling target compound **2**. To the best of author's knowledge compound **2** has not been synthesized.
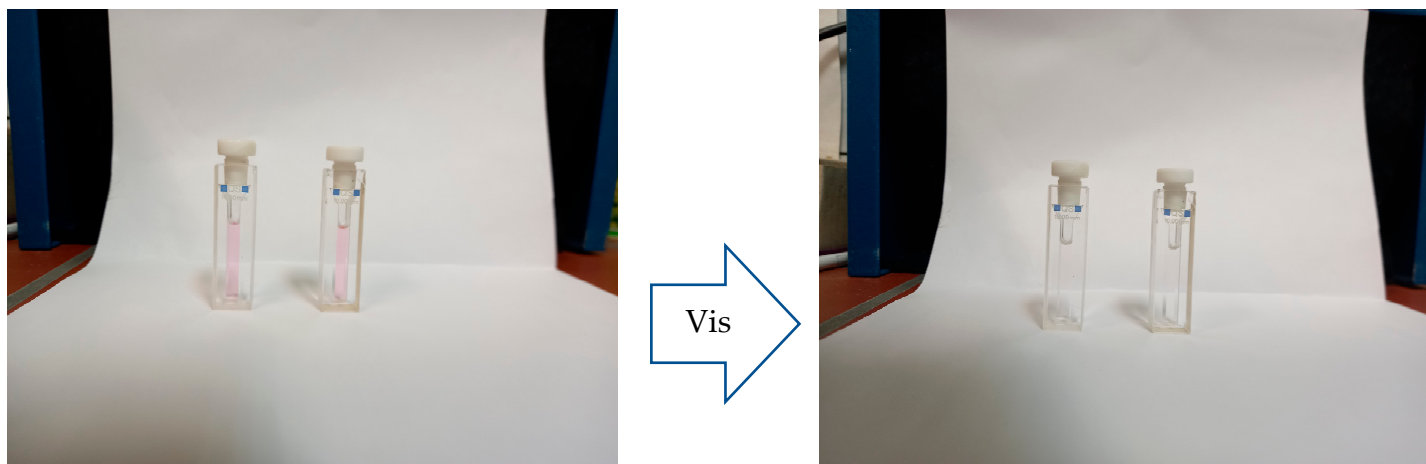

Figure S17: Left: Solutions of compounds **2** and **3** after exposing to UV light (254 nm, 8 x 8W, 19s) turn form coluorless to pink.

      Right: The pink colour of UV irradiated solutions turns to colourless after exposing to Vis light (room light, discharge lamps, PHILIPS master 2 x TL-D 36W/840 and OSRAM 3 x L36W/765, 1h47min) ($c = 10^{-4}$ moldm^{-3}, CH$_3$CN).
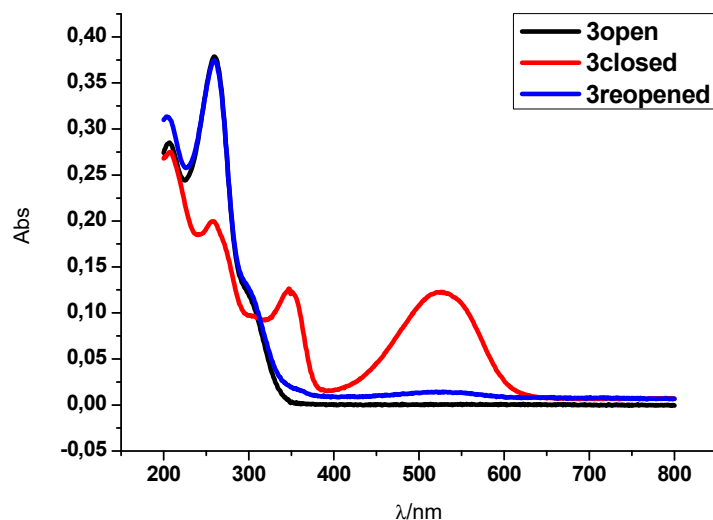
Figure S18: The photo-switch process of compound **3** (—) detected by UV-Vis absorption. The closed form (—) was induced using UV light (254 nm, 8 x 8W, 19s) and the reopened form (—) using visible light (room light, 1h47min). The closed form has a characteristic λ_{max} in the visible light region of spectra (λ_{max} = 528 nm); (c(**3**) = 10^{-4} mol dm^{-3}, CH$_3$CN)
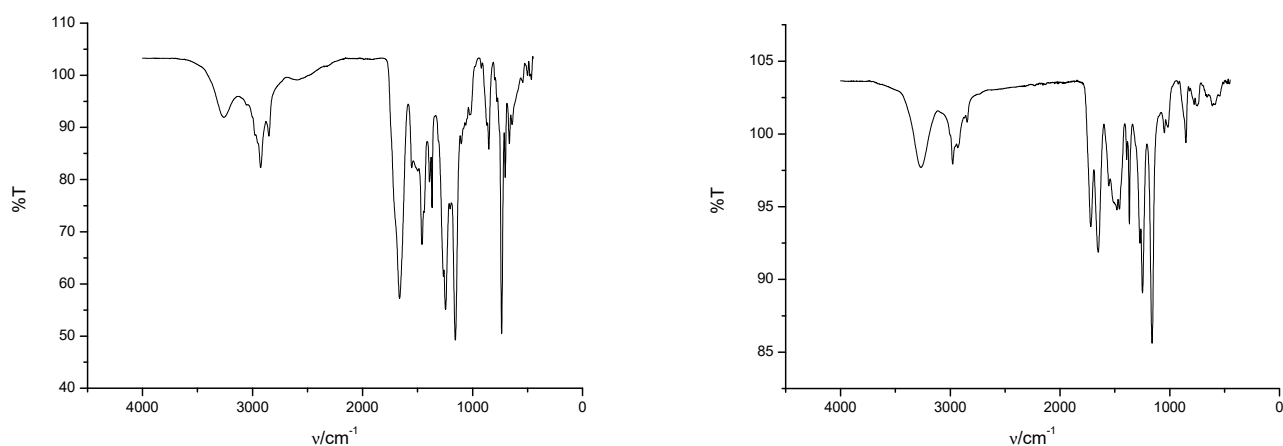

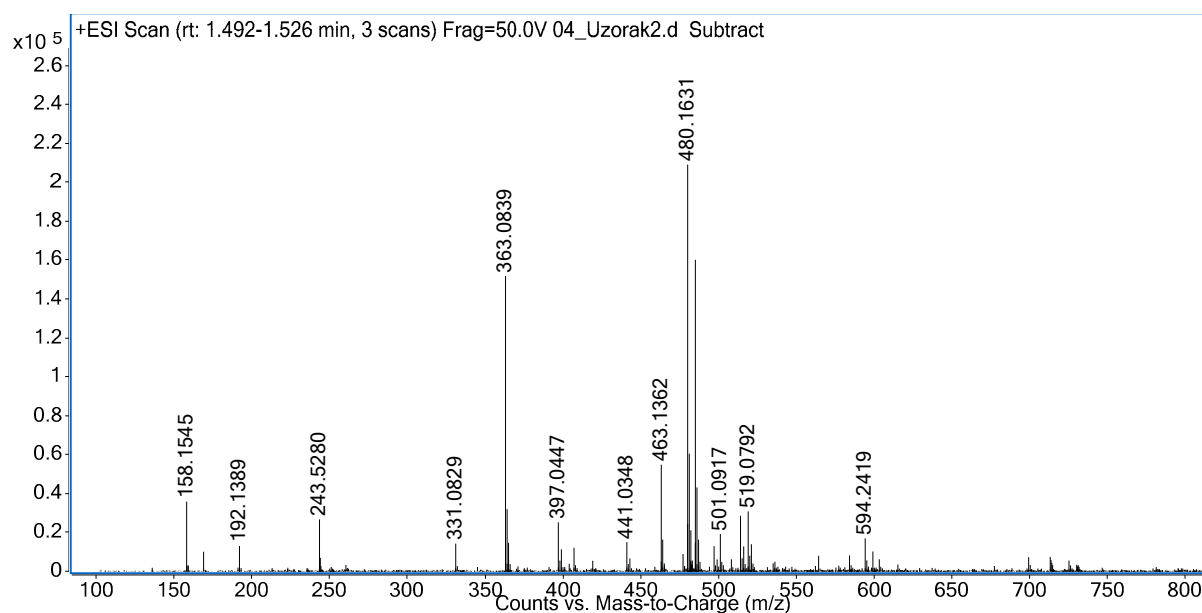Figure S19: IR spectrum of compound **2** (left) and compound **3** (right). (Instrument Perkin Elmer FT-IR UATR).

Figure S20: HRMS scan of **2** ($C_{22}H_{26}N_2O_5S_2$); Characteristic ionic species found: [M+H]+: 463,1361; [M+NH4]+: 480,1627; [M+Na]+: 485,1181; [M+K]+: 501,0920
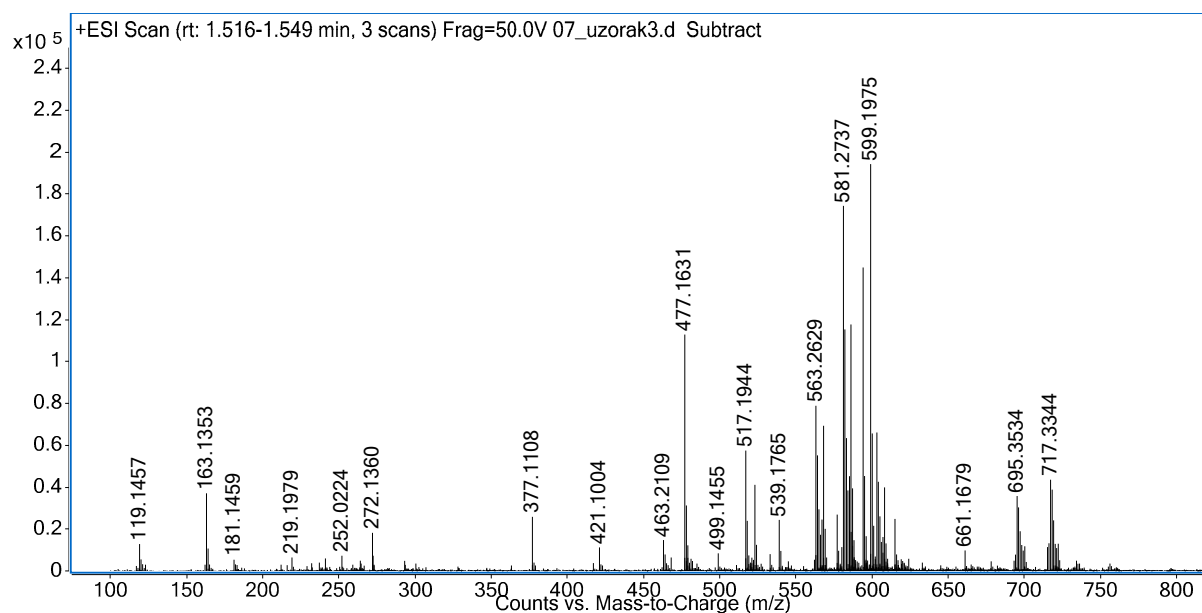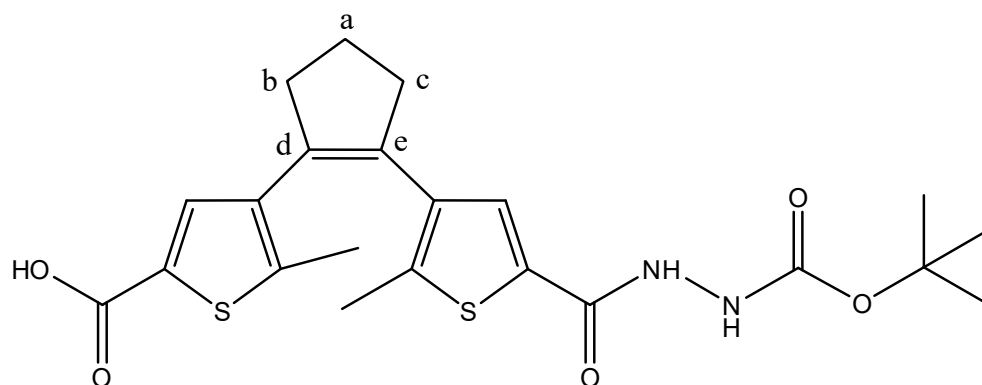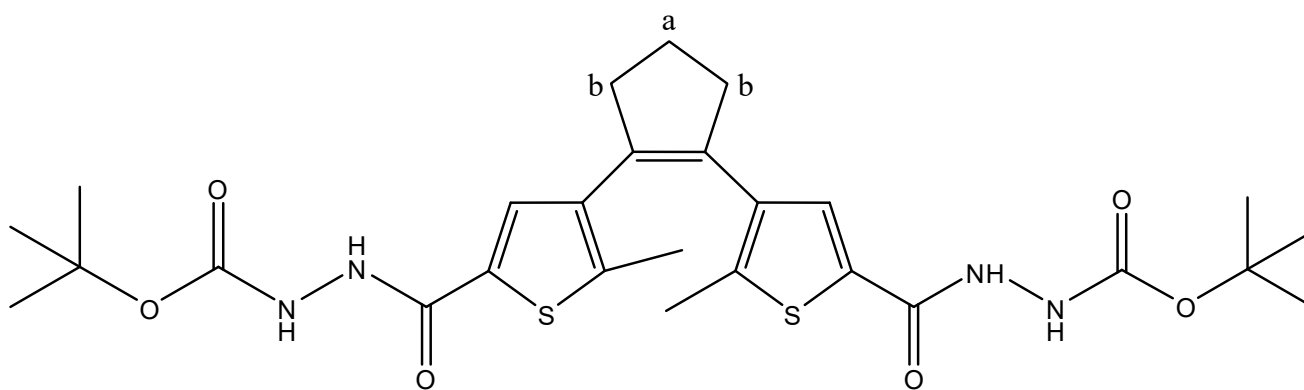

Figure S21: HRMS scan of **3**; ($C_{27}H_{36}N_4O_6S_2$); Characteristic ionic species found: [M+H]+: 577,2155; [M+NH4]+: 594,2420; [M+Na]+: 599,1974; [M+K]+: 615,1713.

Compound: **2**

Chemical Formula: $C_{22}H_{26}N_2O_5S_2$

Molecular Weight: 462,58

Compound: **3**

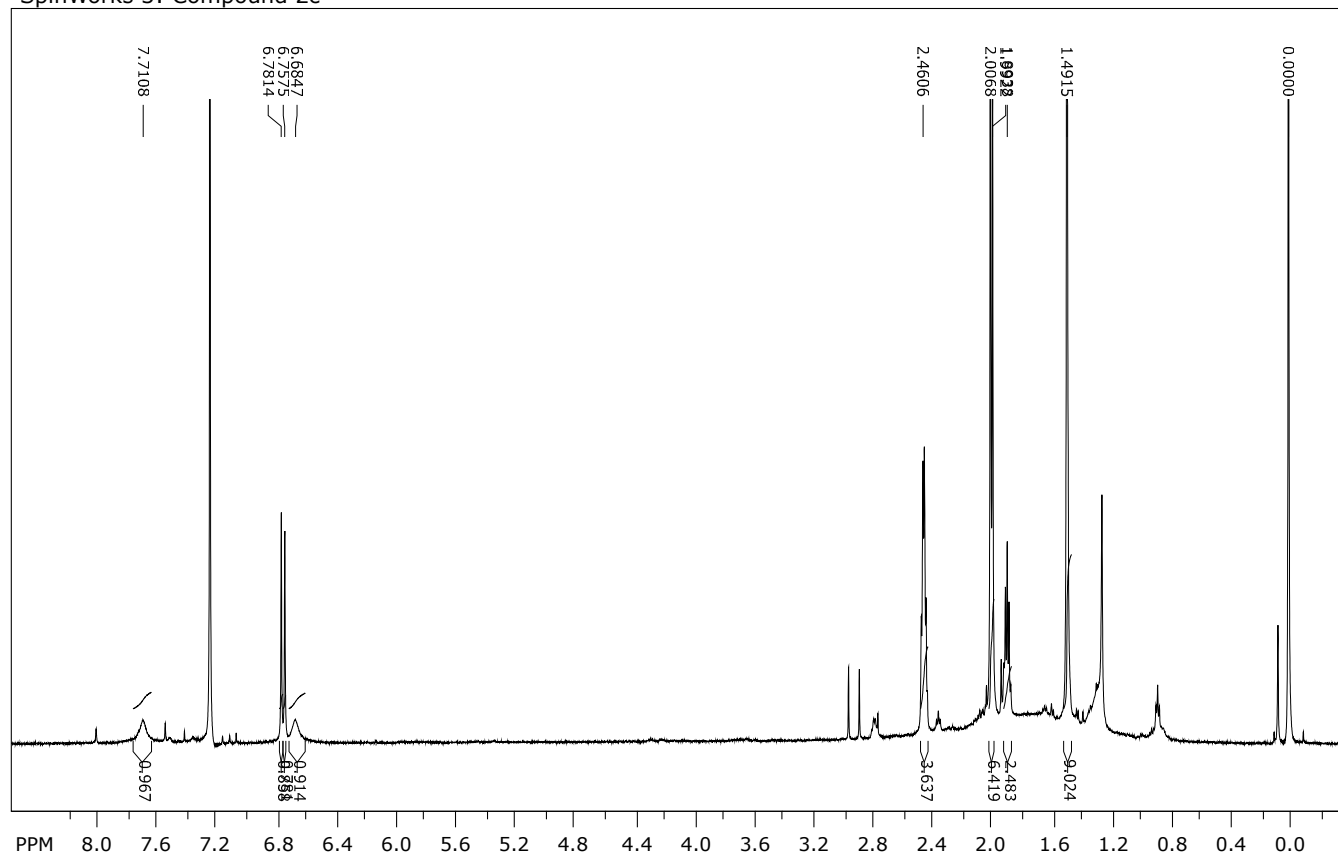Chemical Formula: $C_{27}H_{36}N_4O_6S_2$

Molecular Weight: 576,73

Figure S22: Molecular structures of compound **2** (up) and compound **3** (down). Different C atoms in the cyclopentene ring are marked with a-e.

SpinWorks 3: Compound 2c

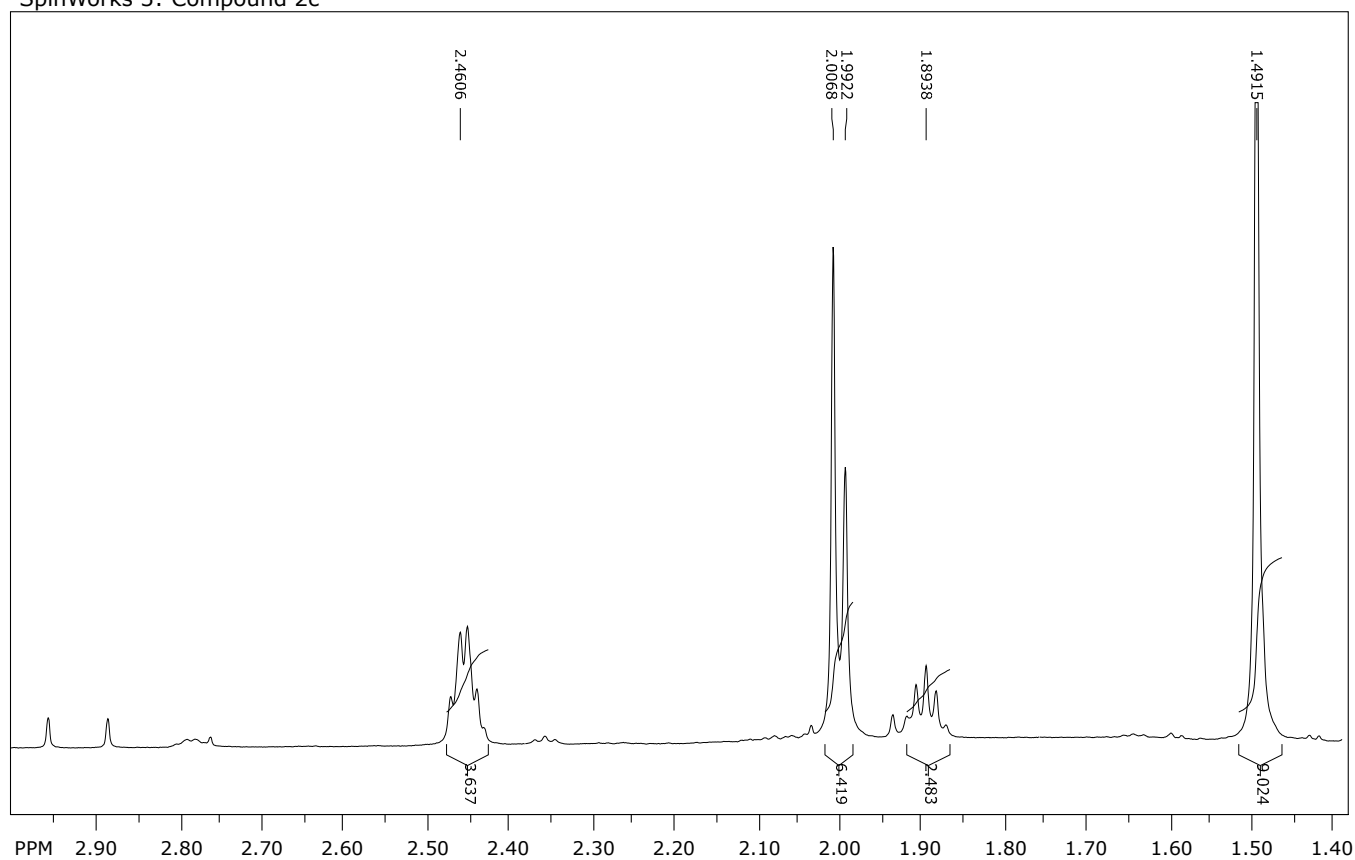Figure S23: ¹H NMR (600 MHz, CDCl₃) spectrum of **2c**

SpinWorks 3: Compound 2c
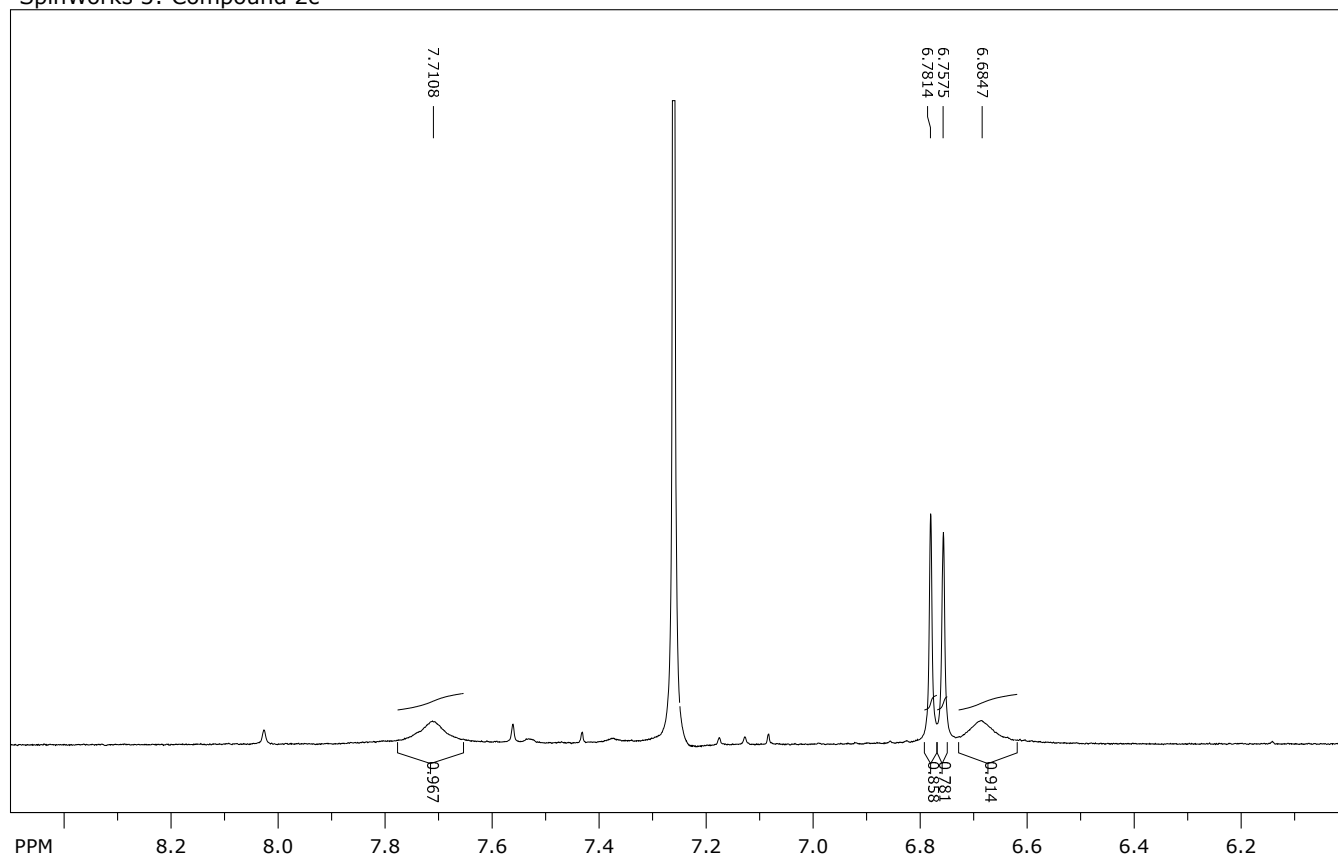
Figure S24: ^{1}H NMR (600 MHz, CDCl$_3$) spectrum (zoomed)of **2c.**

SpinWorks 3: Compound 2c

Figure S25: ¹H NMR (600 MHz, CDCl₃) spectrum (zoomed)of **2c.**
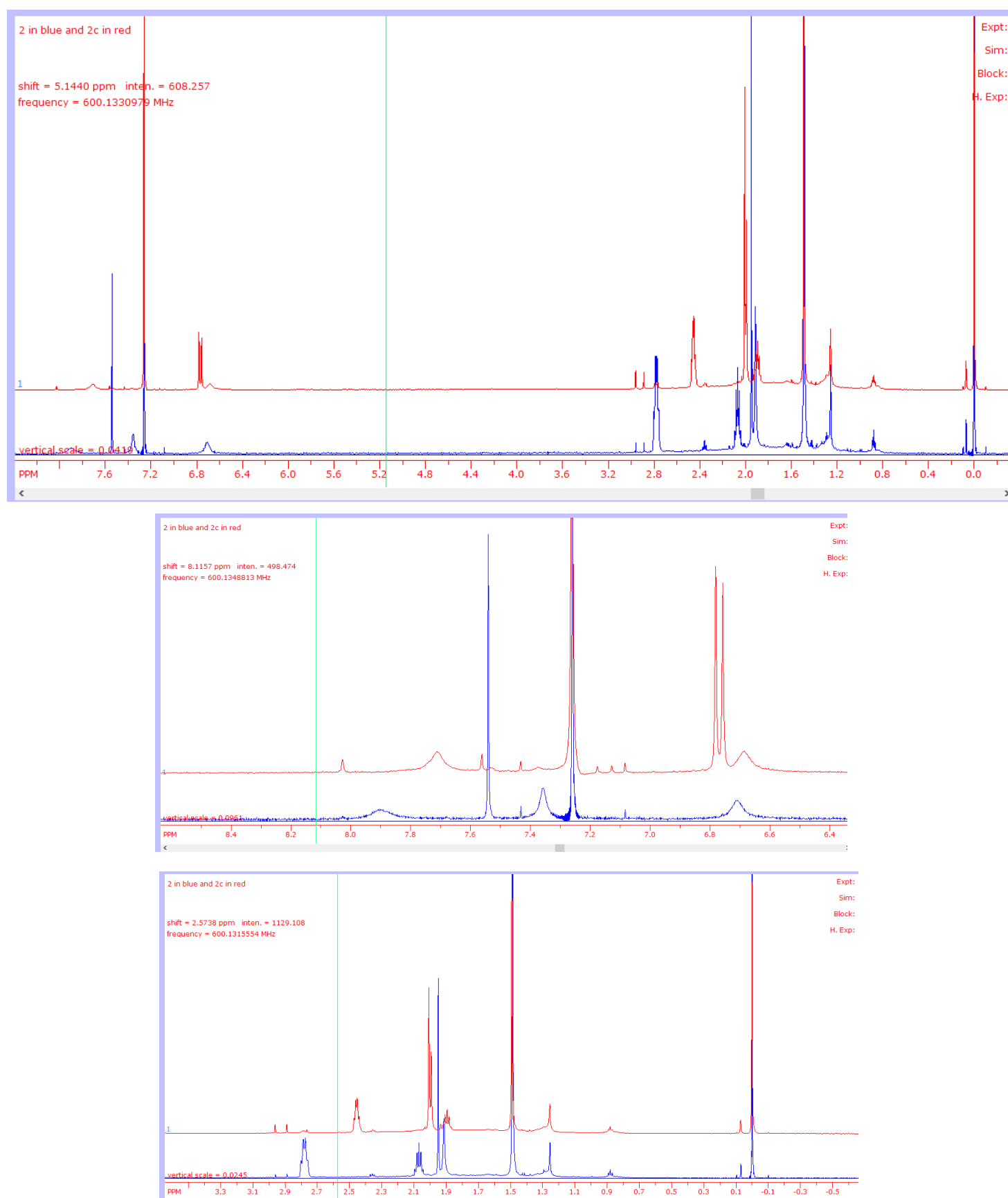
Figure S26 : Overlapped ¹H spectra (600 MHz) of open form **2** (blue) and closed form **2c** (red): up (complete spectra), midle (8.2-6.4 ppm), down (3.1-0.0 ppm).
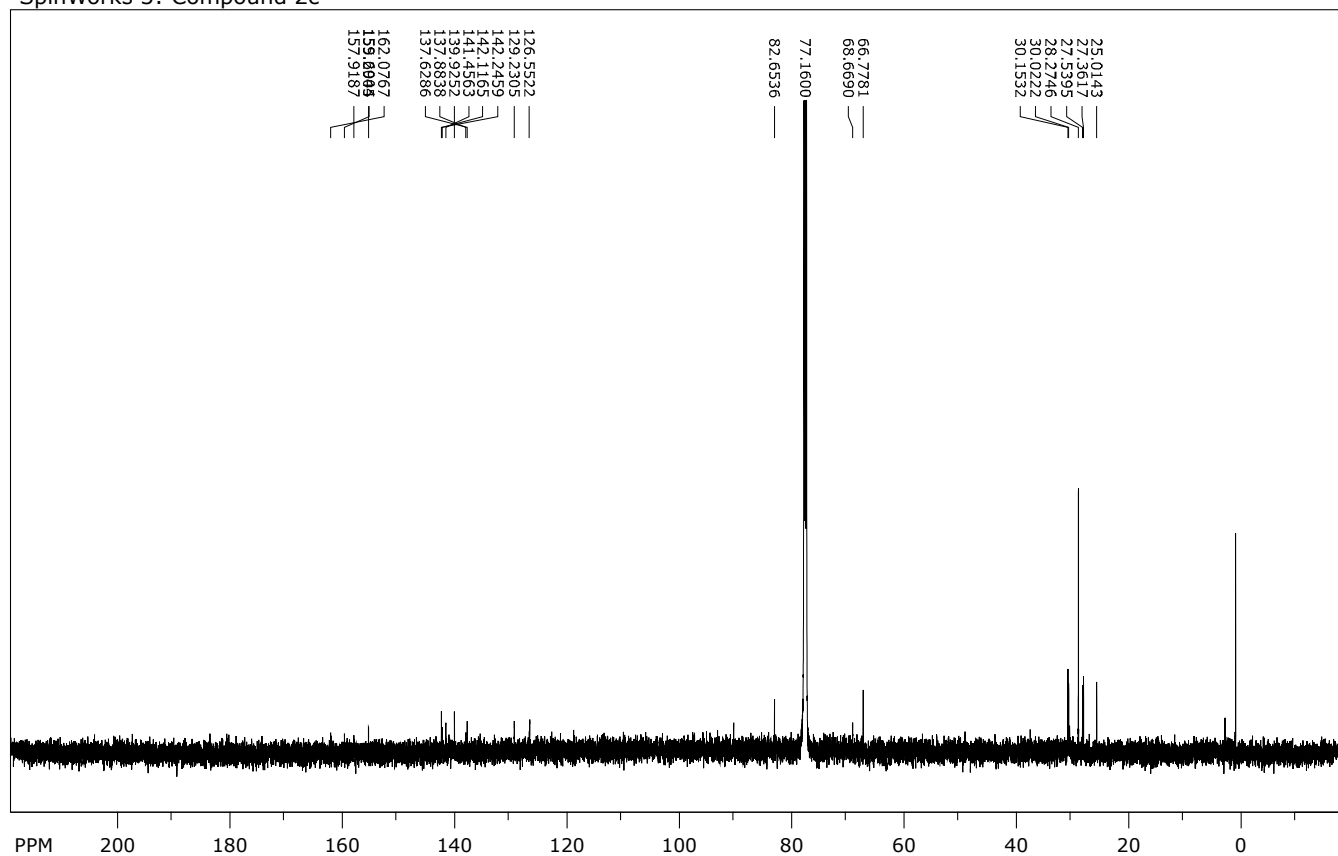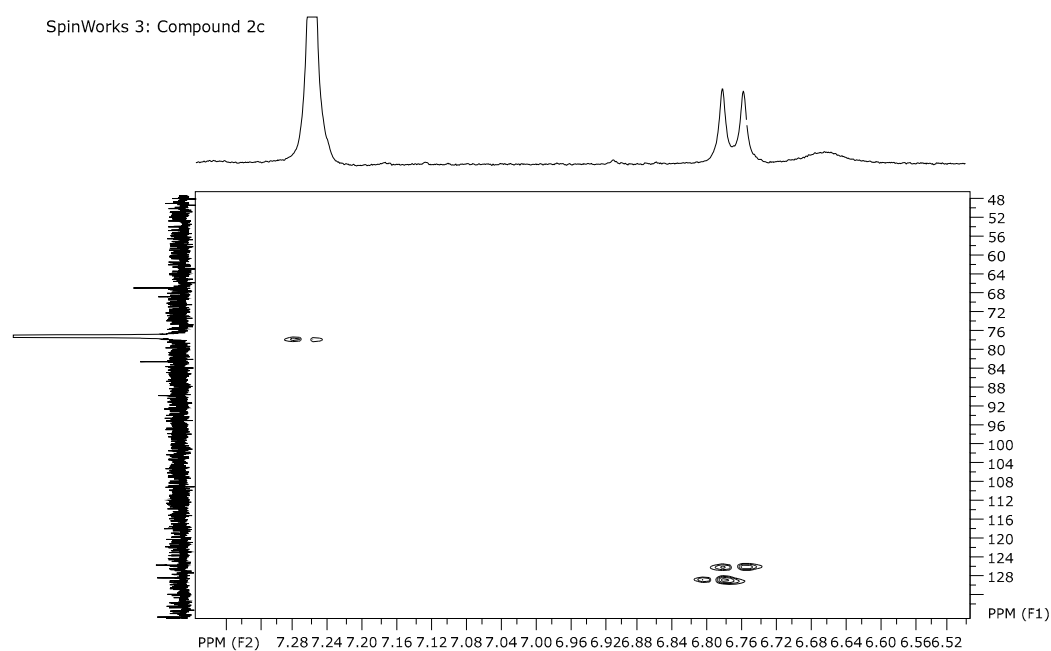
SpinWorks 3: Compound 2c

162.0767
159.8005
159.8095
157.9187
139.9252
141.4563
141.1165
142.2459
137.8838
137.62286
126.5522
129.2305
82.6536
77.1600
66.7781
68.6690
30.0222
28.2746
27.5395
30.1532
27.3617
25.0143

PPM   200   180   160   140   120   100   80   60   40   20   0

Figure S27: ¹³C NMR (75 MHz, CDCl₃, COMPLETE) spectrum of **2c:** *to be precisely determined,* δ 129.2 (CH), 126.6 (CH), 82.7 (CBoc), <u>68.7 (Cquart),</u> 66.8 (Cquart), 30.2 and 30.0 (CH₂(b,c)), 28.3 ((CH₃)₃), 27.5 (CH₃), 27.4 (CH₃), 25.0 (CH₂(a)); (to be corrected)
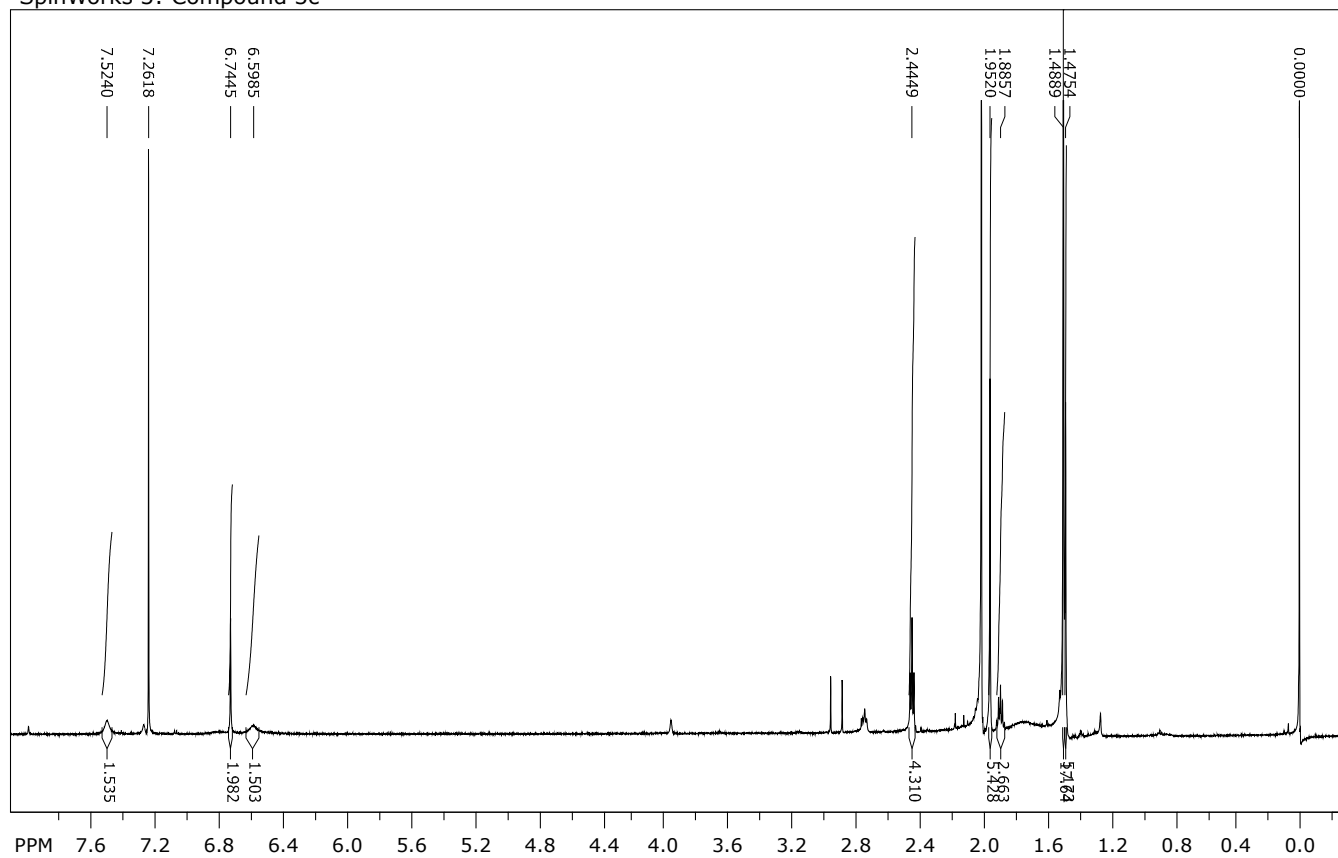
a)
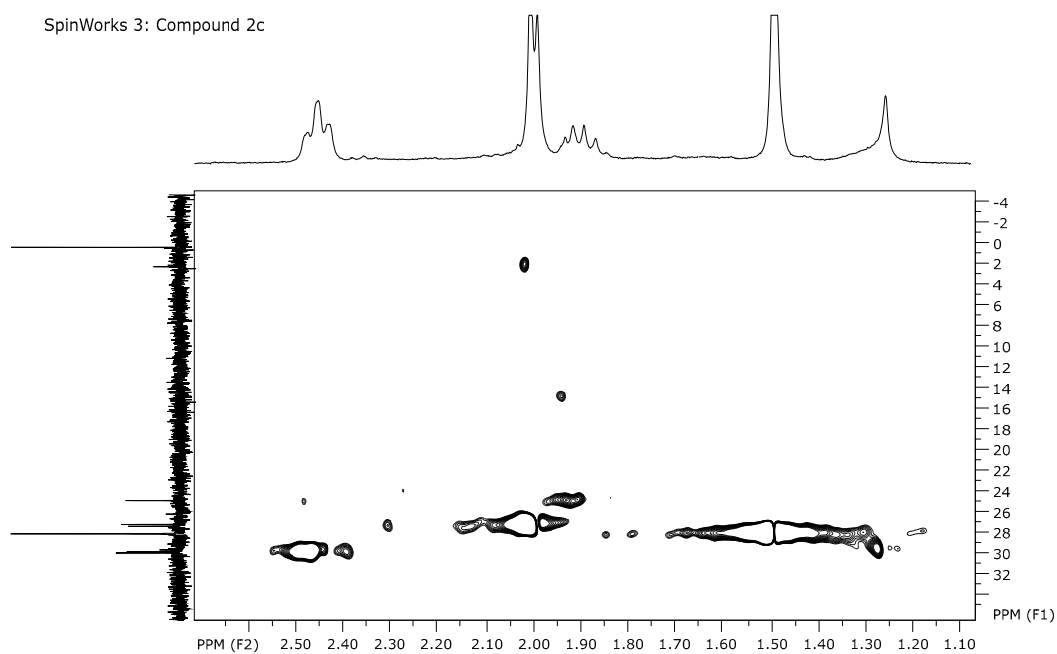
b)

Figure S28 : HSQC spectrum (300 MHz, CDCl$_3$) of compound **2c**: a) from 6.52 ppm to 7.30 ppm (^{1}HNMR) b) up to 2.60 ppm (^{1}HNMR)

27 of 34

SpinWorks 3: Compound 3c

Figure S29: ¹H NMR (600 MHz, CDCl₃) spectrum of **3c** with traces of starting compound **3.**
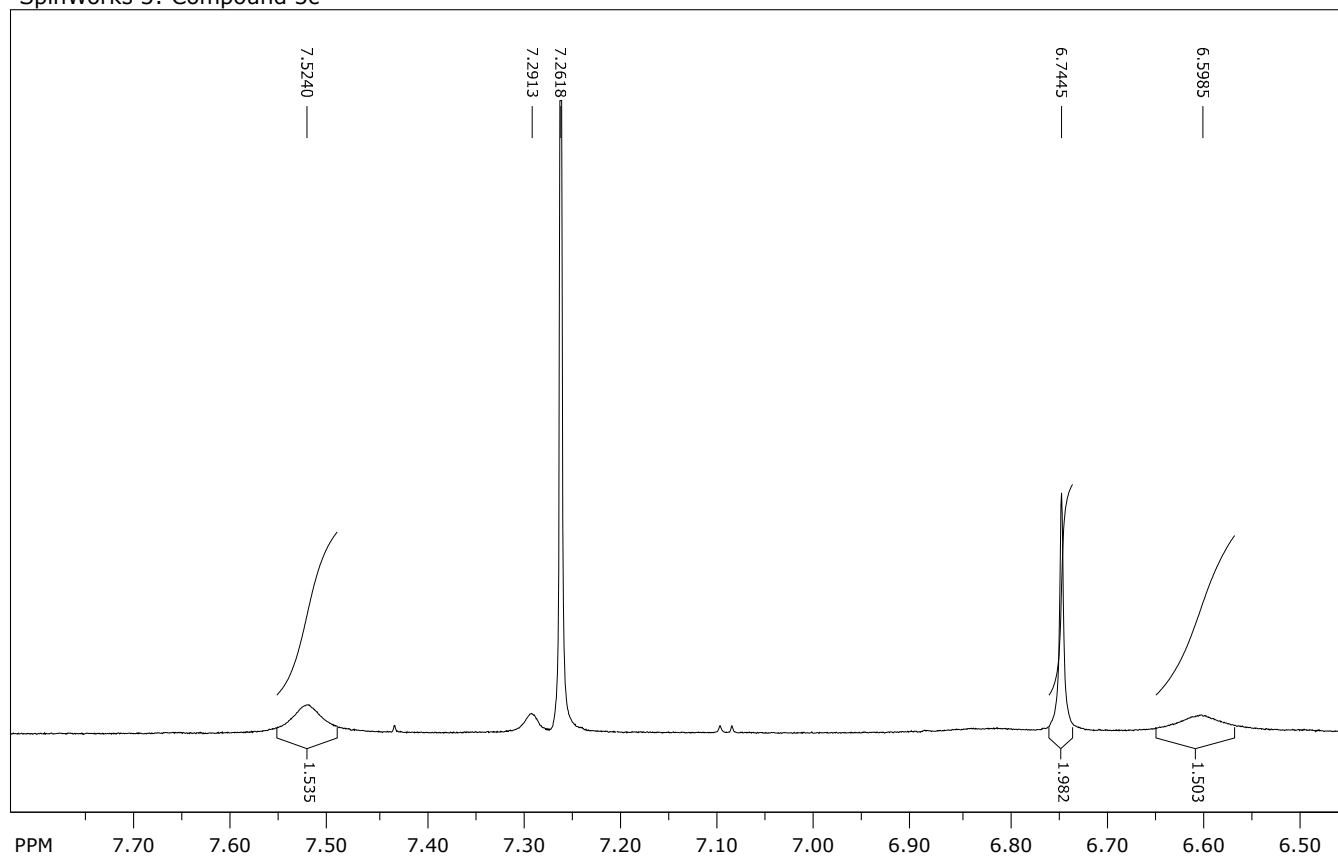
SpinWorks 3: Compound 3c
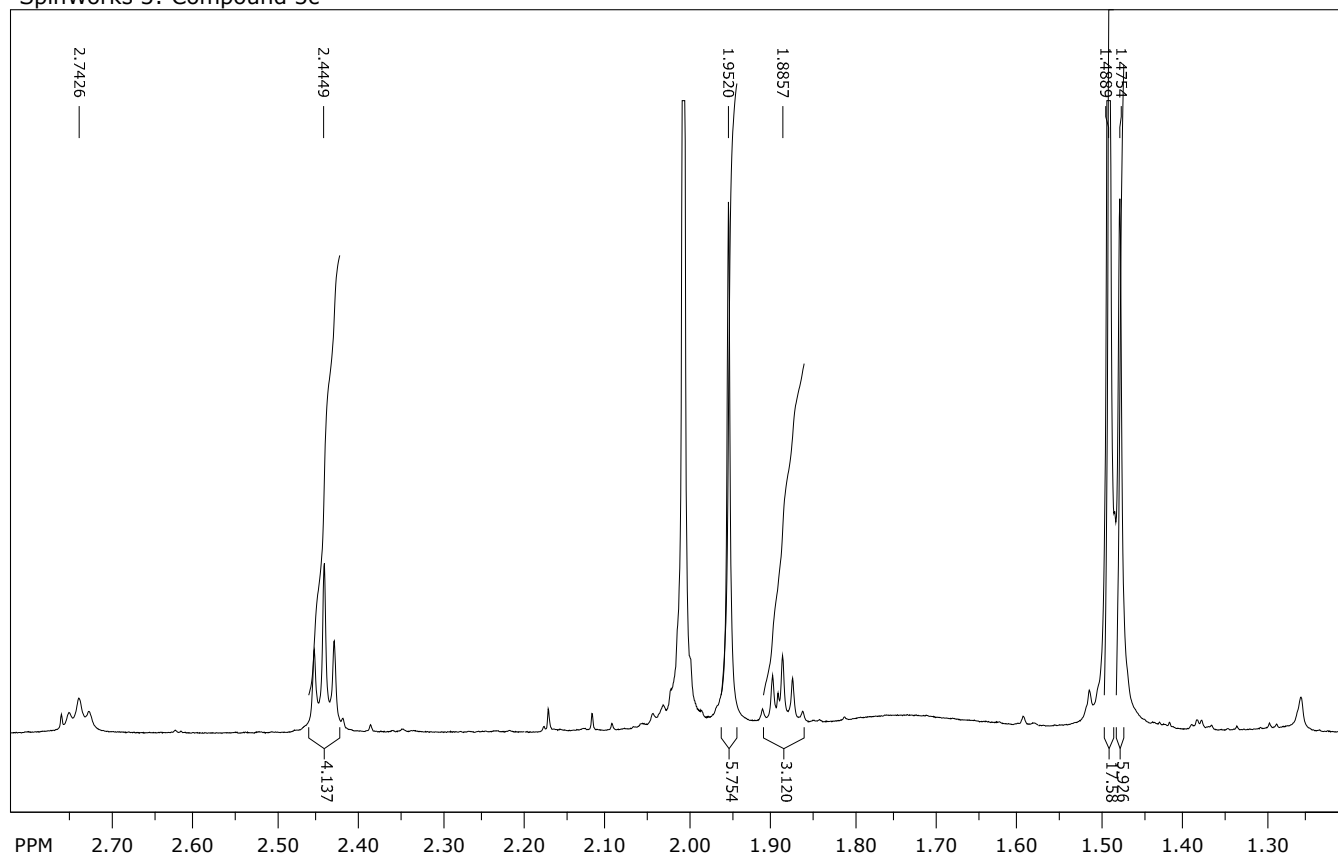

Figure S30: ¹H NMR (600 MHz, CDCl₃) spectrum (zoomed) of **3c** with traces of starting compound **3.**

SpinWorks 3: Compound 3c

Figure S31: ^{1}H NMR (600 MHz, CDCl$_3$) spectrum (zoomed) of **3c** with traces of starting compound **3.**

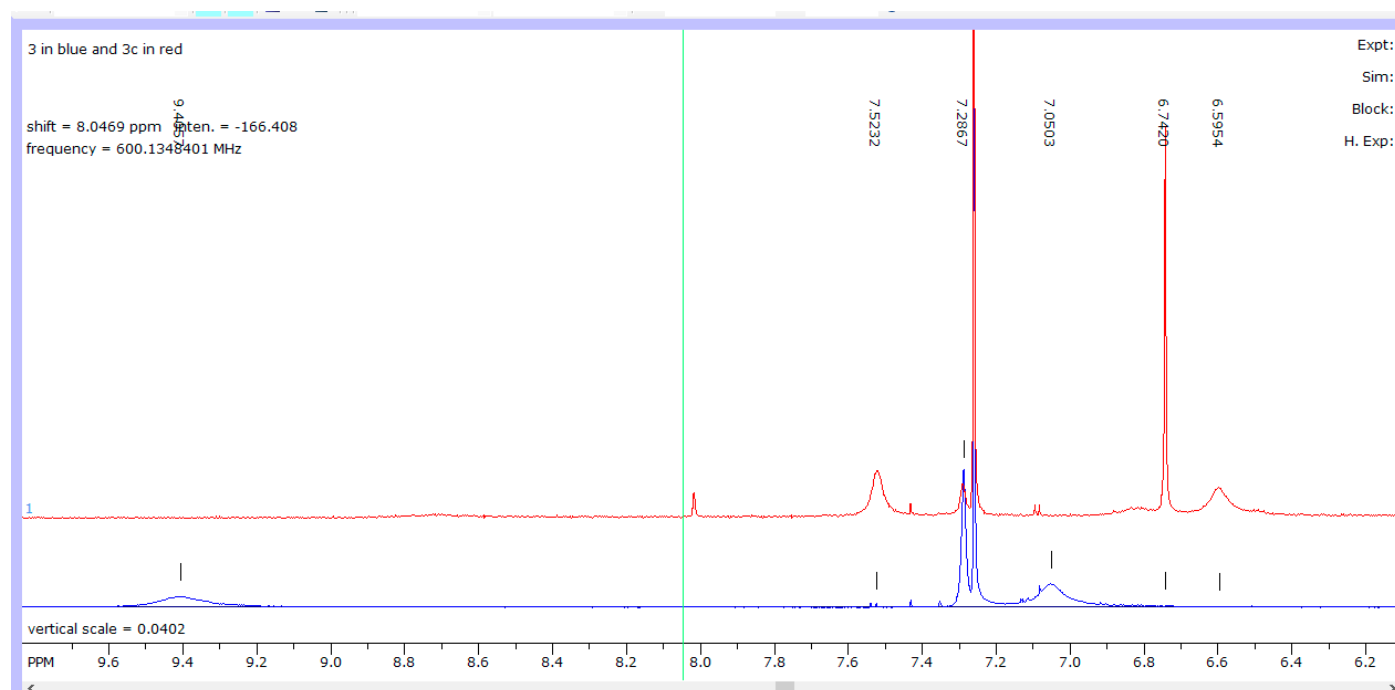Figure S32: Overlapped ¹H spectra (CDCl₃, 600 MHz) of open form **3** (blue) and closed form **3c** (red) (**3c** has traces of **3**).
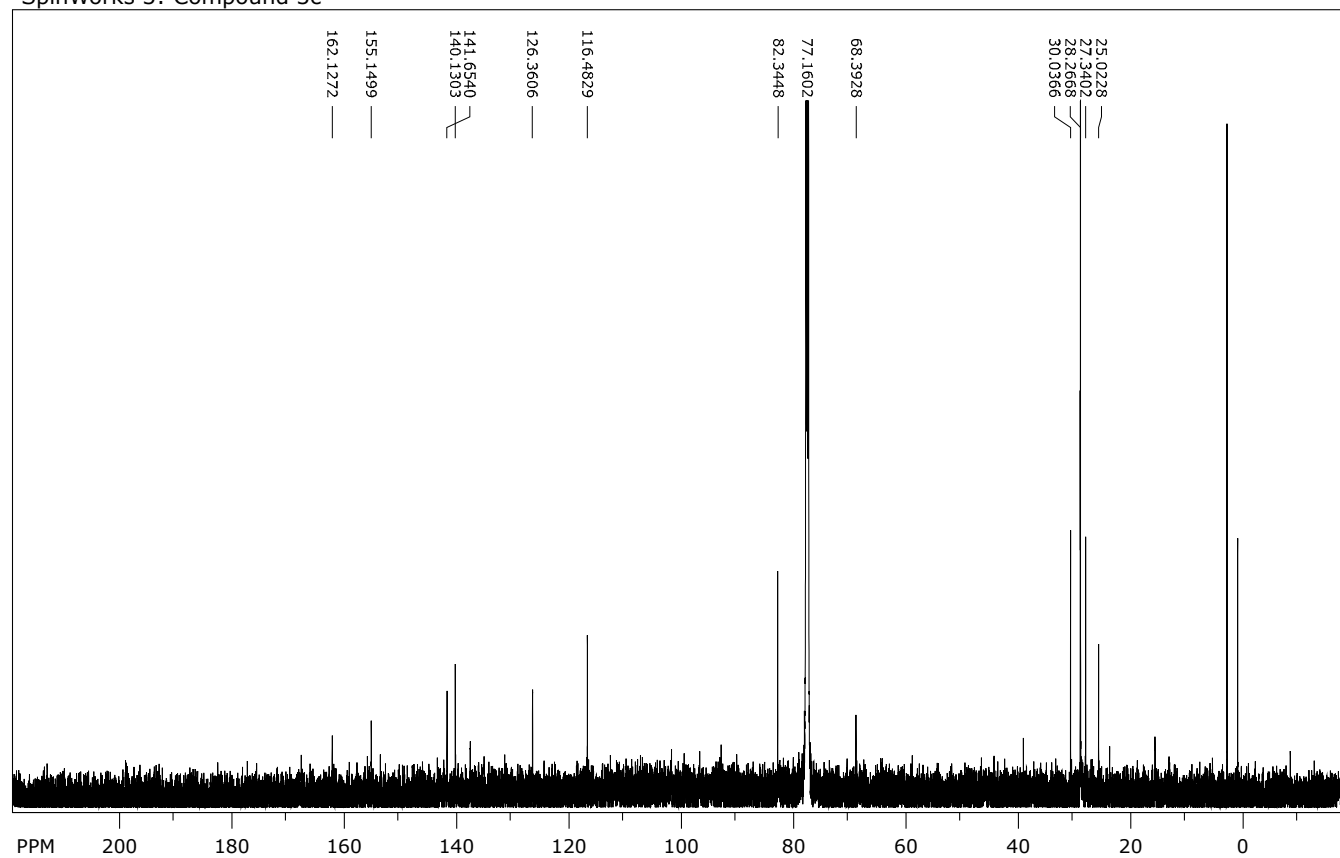
SpinWorks 3: Compound 3c

Figure S33: ¹³C NMR (150 MHz, CDCl₃, COMPLETE) spectrum of **3c:** 162.1, 155.1, 141.7, 140.1, 126.4, 116.5, 82.3, 68.4, 30.0, 28.3, 27.3, 25.0 (uncorrected)
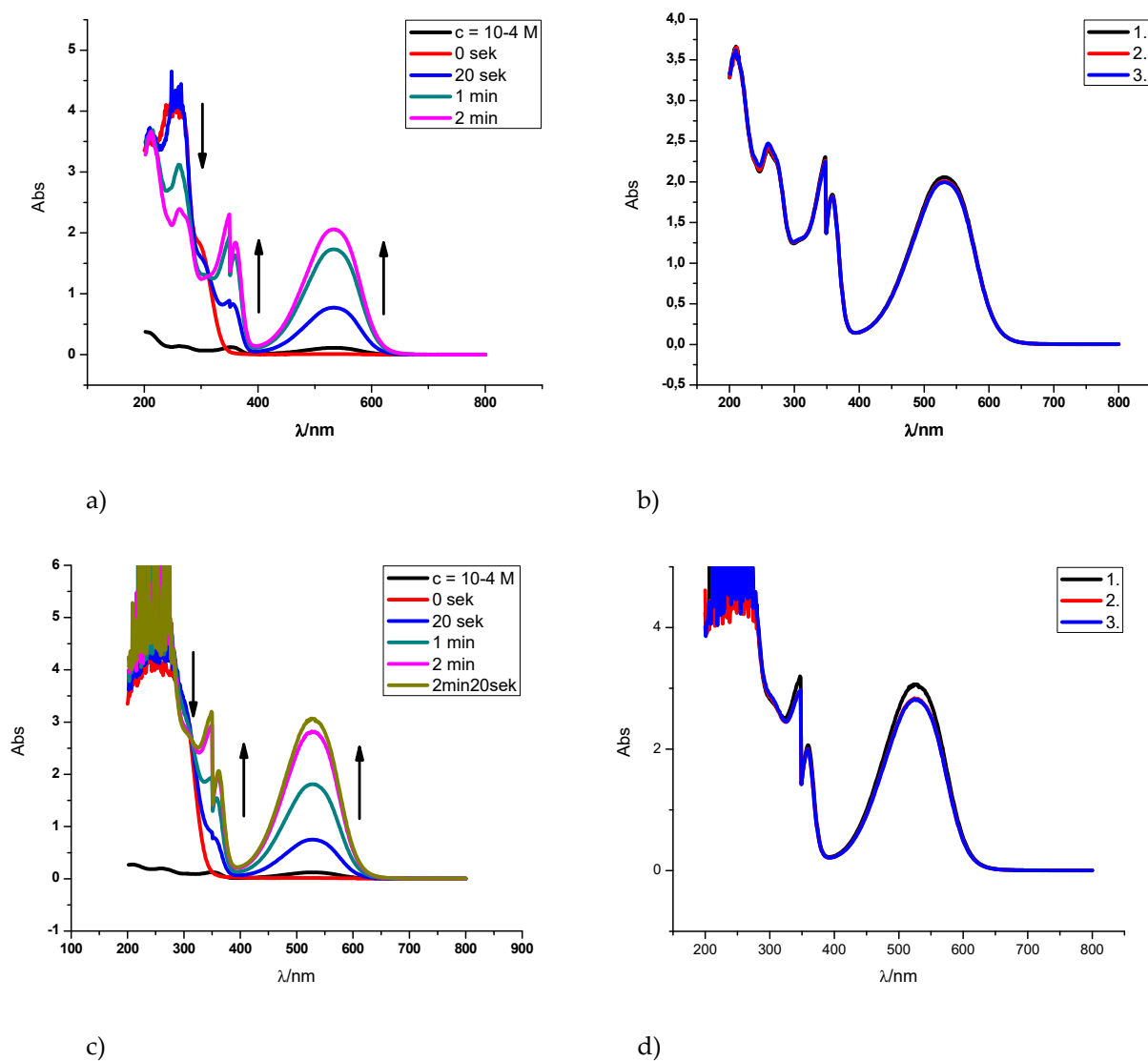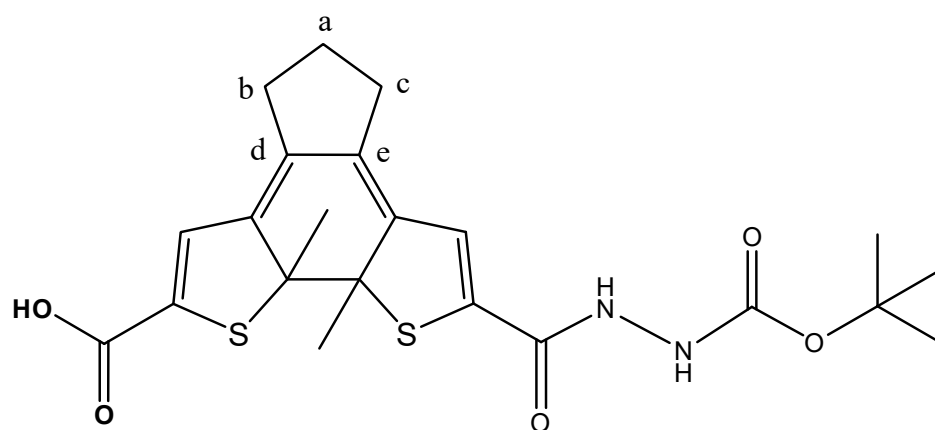
Figure S34 (a-d): Synthesis of **2c** and **3c** followed by UV spectroscopy; a) Accumulation of **2c** b) Cuvette (1.-3.) comparison for **2c; c**) Accumulation of **3c** d) Cuvette (1.-3.) comparison for **3c**
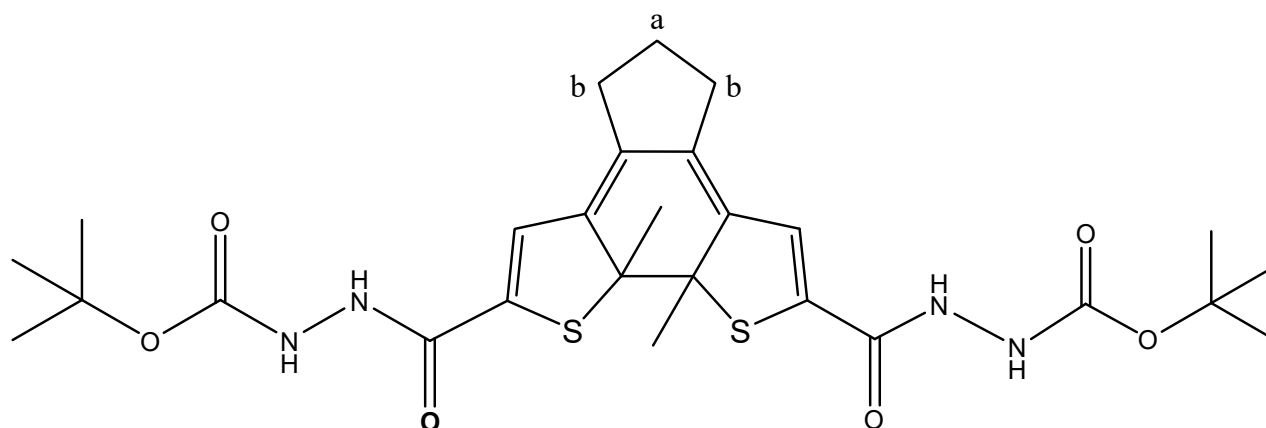
Compound: **2c**

Chemical Formula: $C_{22}H_{26}N_2O_5S_2$

Molecular Weight: 462,58

IUPAC name:

14-[N'-(*tert*-butoxycarbonyl)hydrazinecarbonyl]-1,2-dimethyl-3,15-dithiatetracyclo[10.3.0.0^{2,6}.0^{7,11}]pentadeca-4,6,11,13-tetraene-4-carboxylic acid
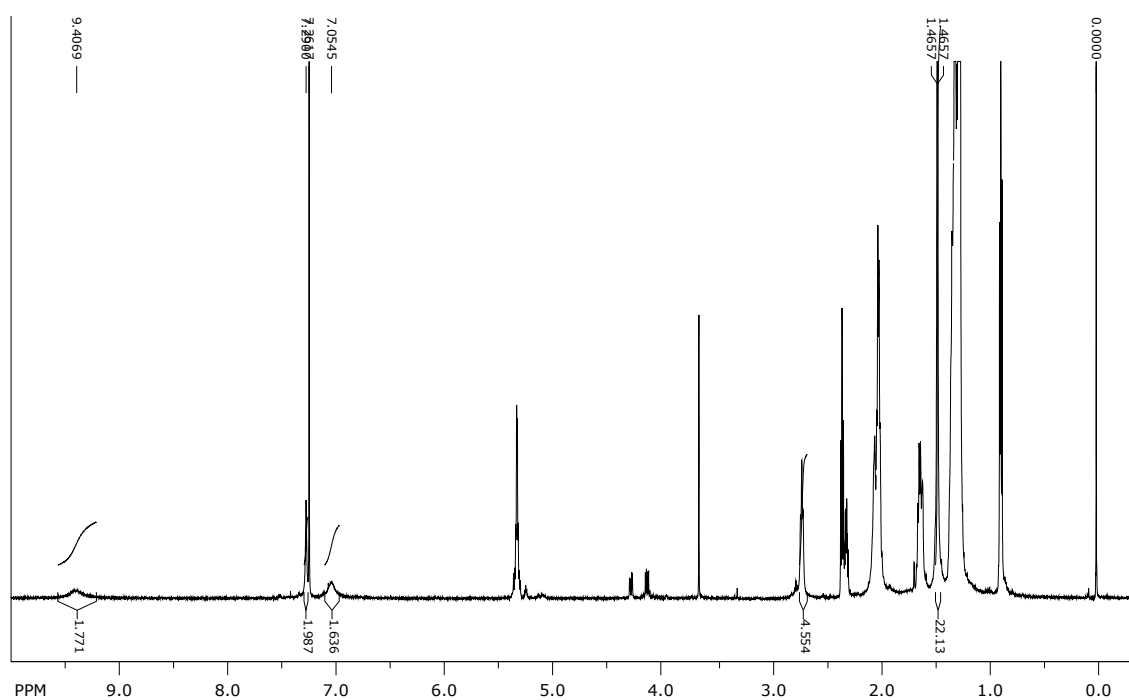

Compound: **3c**

Chemical Formula: $C_{27}H_{36}N_4O_6S_2$

Molecular Weight: 576,73

IUPAC name:

N'4,N'14-bis-(*tert*-butoxycarbonyl)-1,2-dimethyl-3,15- dithiatetracyclo[10.3.0.0^{2,6}.0^{7,11}]pentadeca-4,6,11,13-tetraene-4,14-dicarbohydrazide

Figure S35: Molecular structures of **2c** and **3c** presumed on the theory of DAE cyclisation and on NMR evidence. Different C atoms in the cyclopentane ring are marked with a-e.
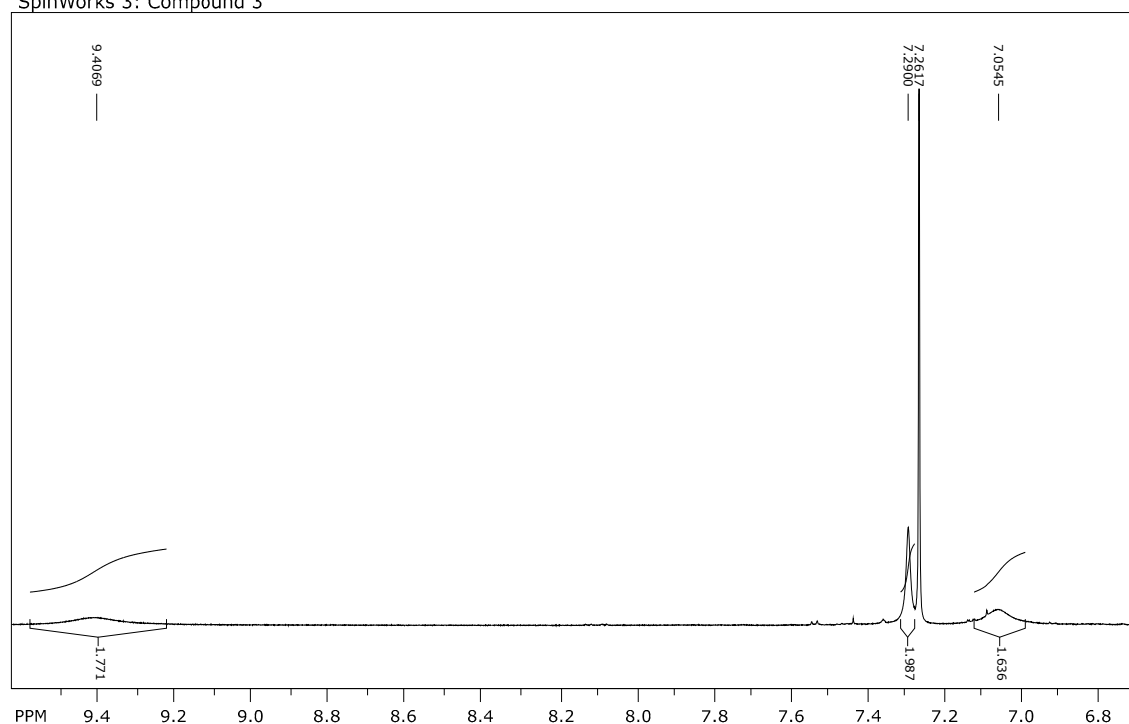
SpinWorks 3: Compound 3

Figure S36: ^{1}H NMR spectrum (600 MHz, CDCl$_3$) of **3** obtained as described in Figure S11-12; up whole spectra, down zoomed. Contaminants are present (lower ppm).