ELECTRONIC SUPPLEMENTARY INFORMATION

Manuscript title: **Synthesis, structural properties, electrophilic substitution reactions and DFT computational studies of calix[3]benzofurans**

Author(s): Md. Monarul Islam,[a,b] Thamina Akther,[a] Yusuke Ikejiri,[a] Taisuke Matsumoto,[c] Junji Tanaka,[c] Shofiur Rahman,[d] Paris E. Georghiou,[d] David L. Hughes,[e] Carl Redshaw[f] and Takehiko Yamato*[a]

[a]*Department of Applied Chemistry, Faculty of Science and Engineering, Saga University, Honjo-machi 1, Saga-shi, Saga 840-8502, Japan. Fax: (internat.) + 81(0)952/28-8548; E-mail: yamatot@cc.saga-u.ac.jp*
[b]*Chemical Research Division, Bangladesh Council of Scientific and Industrial Research(BCSIR), Dhanmondi, Dhaka-1205, Bangladesh.*
[c]*Institute of Material Chemistry and Engineering, Kyushu University, 6-1, Kasugakoen, Kasuga 816-8580, Japan*
[d]*Department of Chemistry, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador A1B 3X7, Canada*
[e]*School of Chemistry, University of East Anglia, Norwich, NR4 7TJ, UK*
[f]*Department of Chemistry, The University of Hull, Cottingham Road, Hull, Yorkshire, HU6 7RX, UK.*

Table of Contents

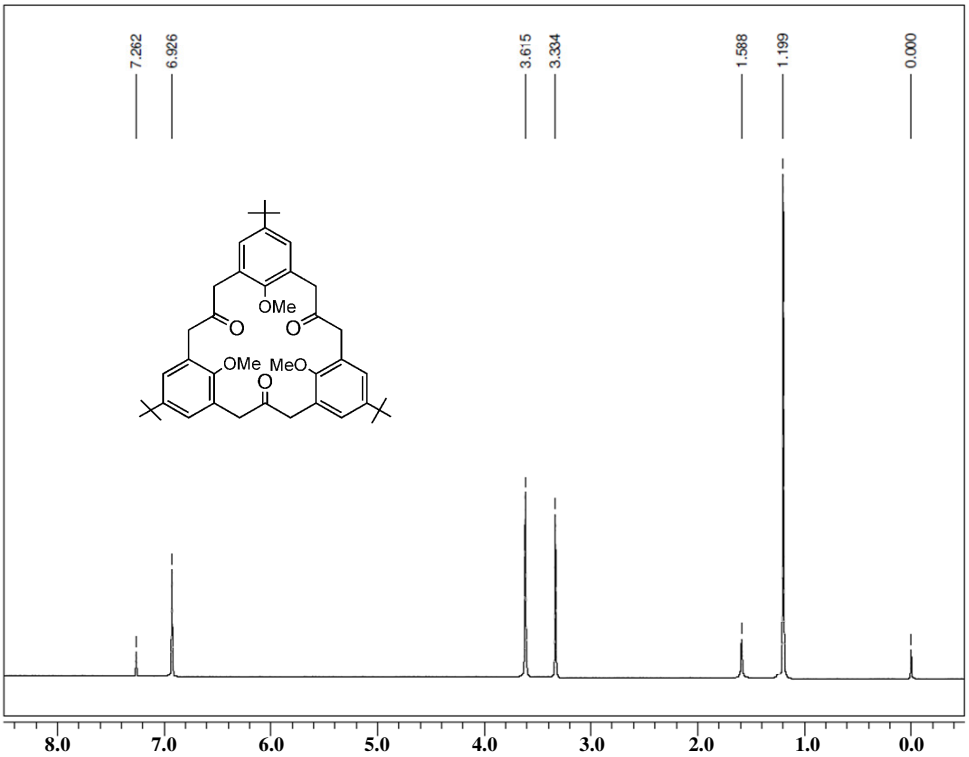


Figure S1. ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **2a.**

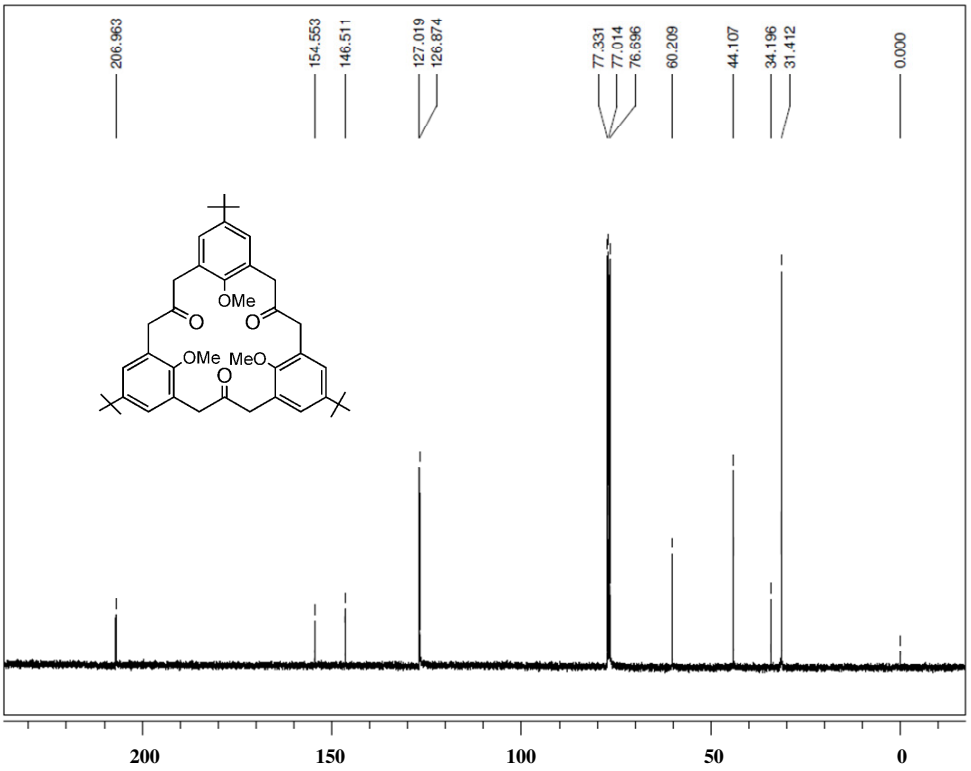


Figure S2: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **2a.**

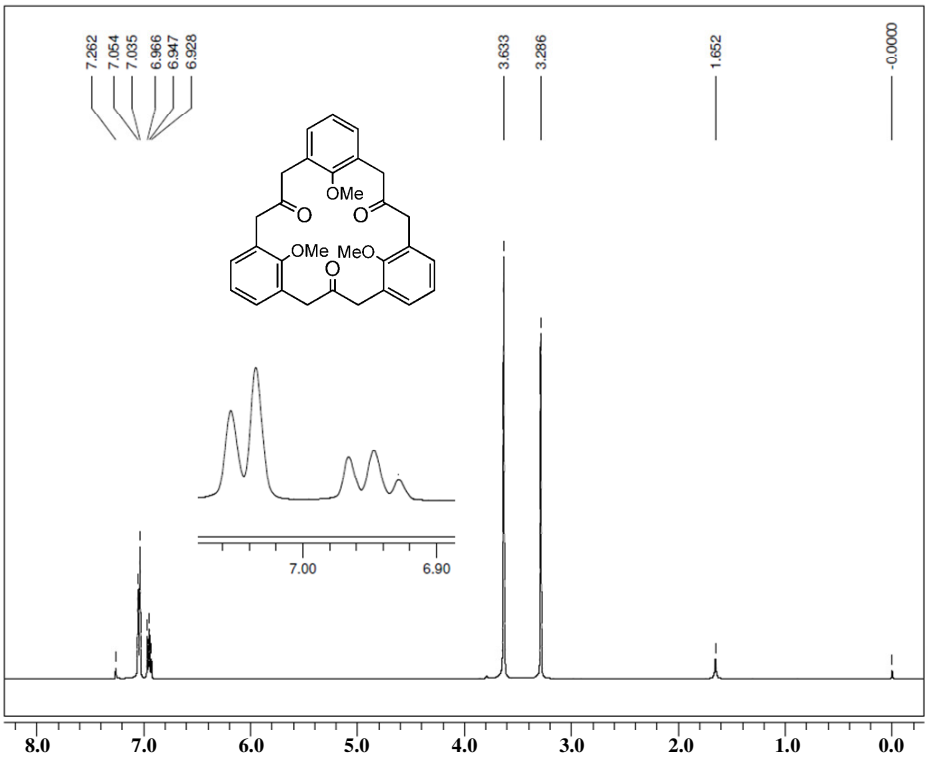


Figure S3: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **2b.**

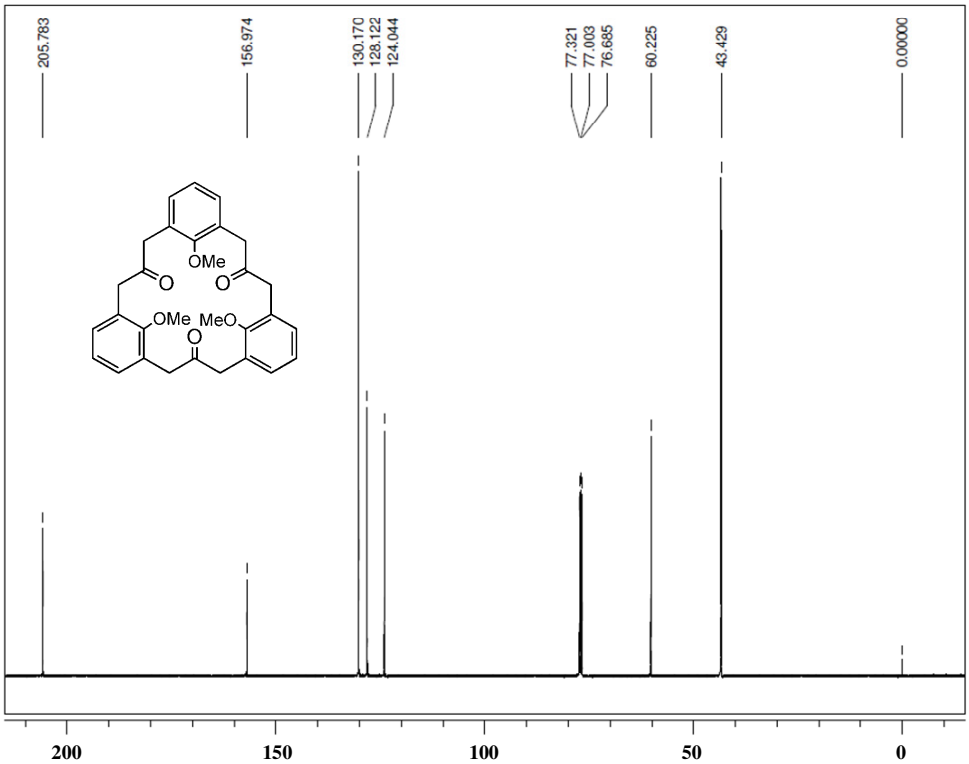


Figure S4: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **2b**.

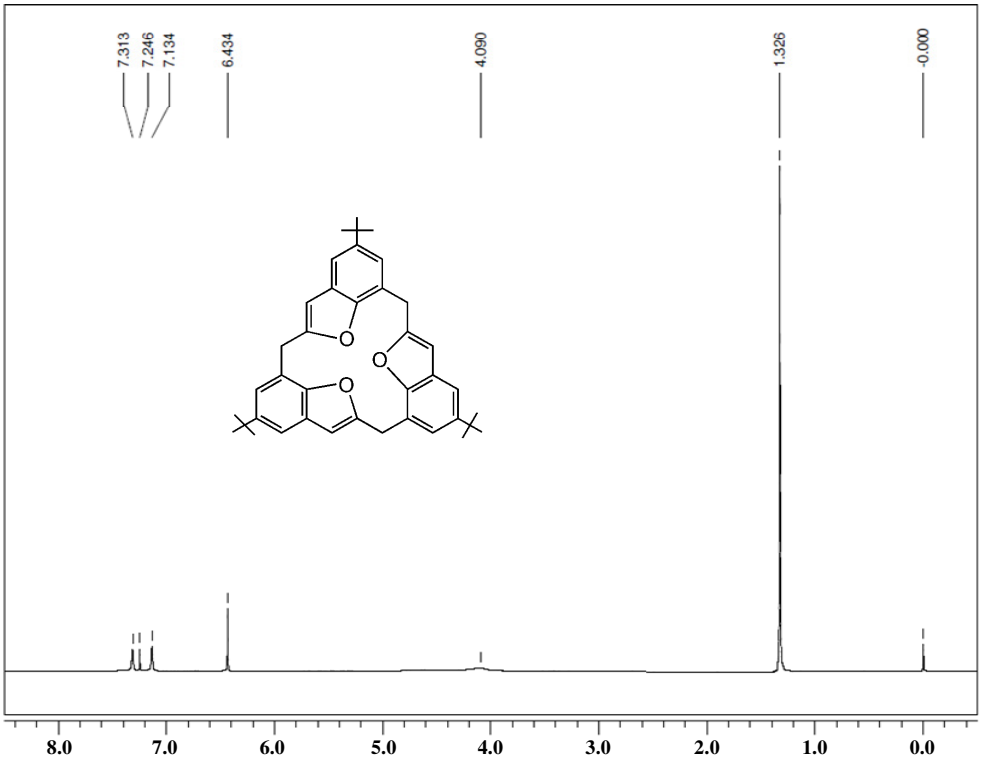


Figure S5: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **4a.**

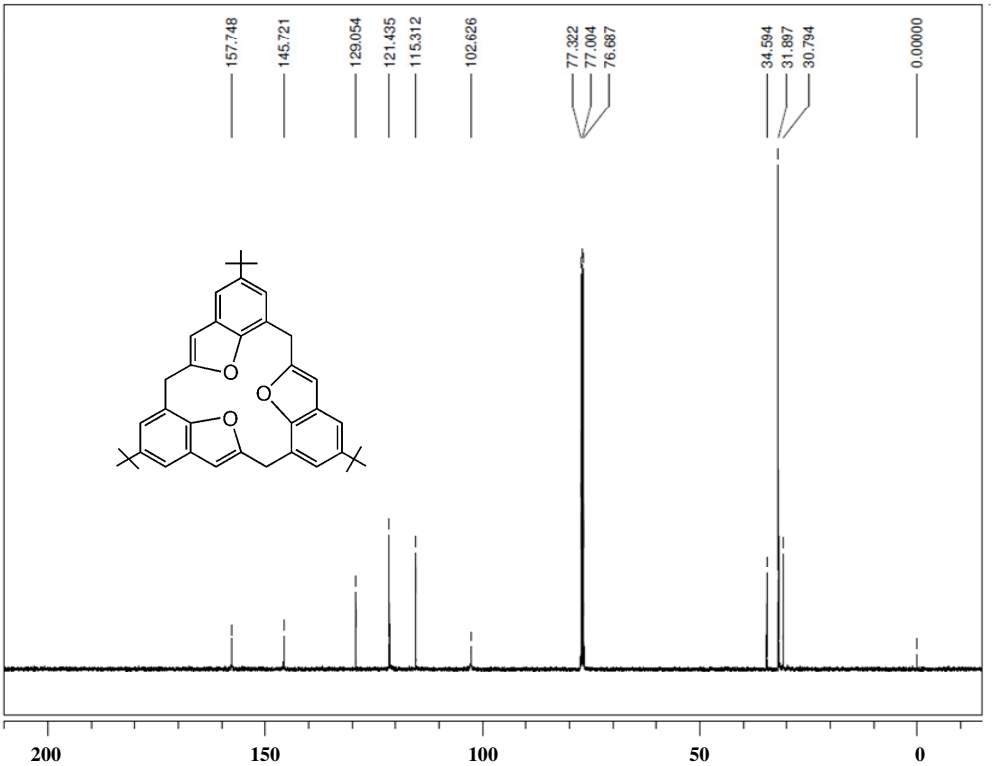


Figure S6: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **4a.**

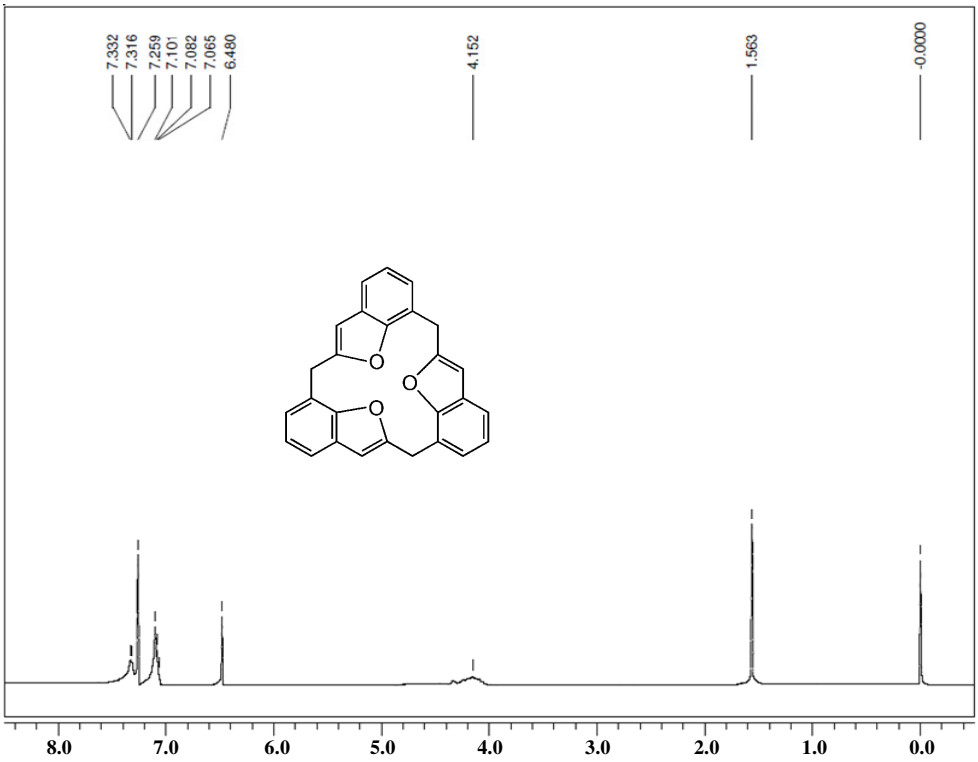


Figure S7: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **5a.**

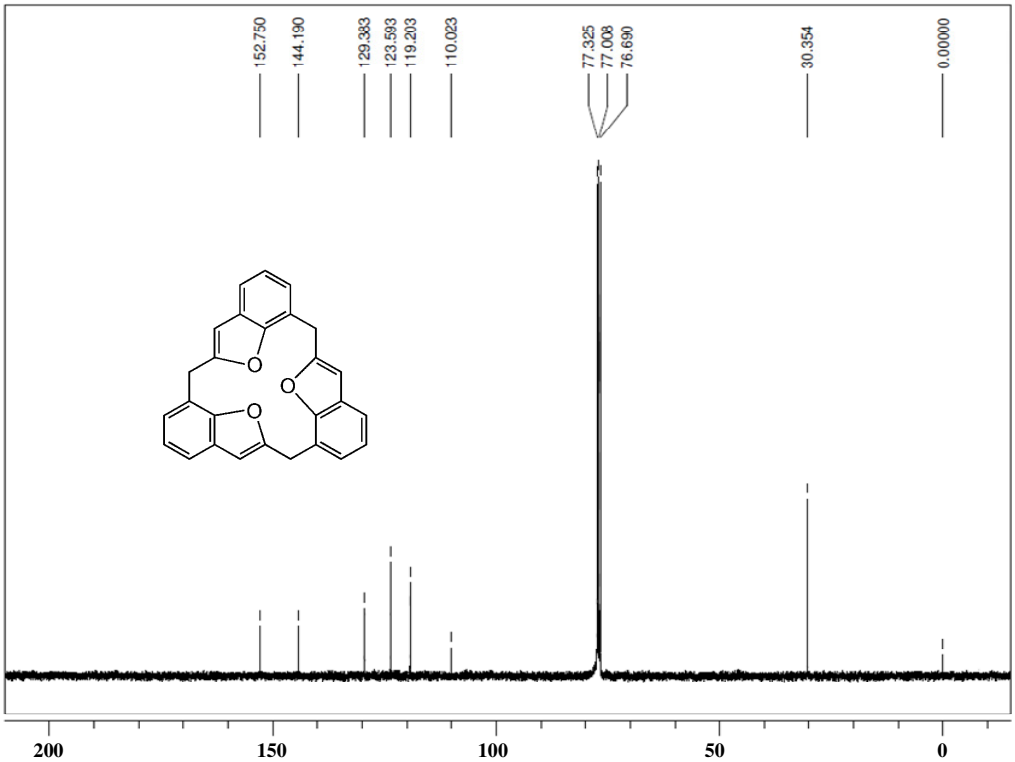


Figure S8: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **5a.**

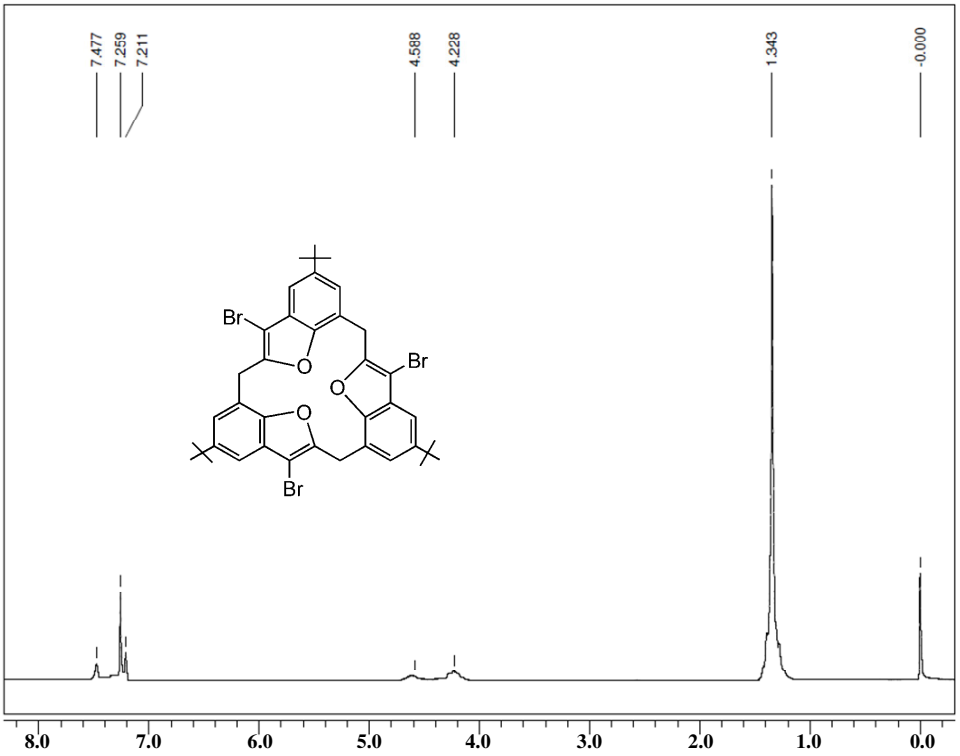


Figure S9: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **4b.**

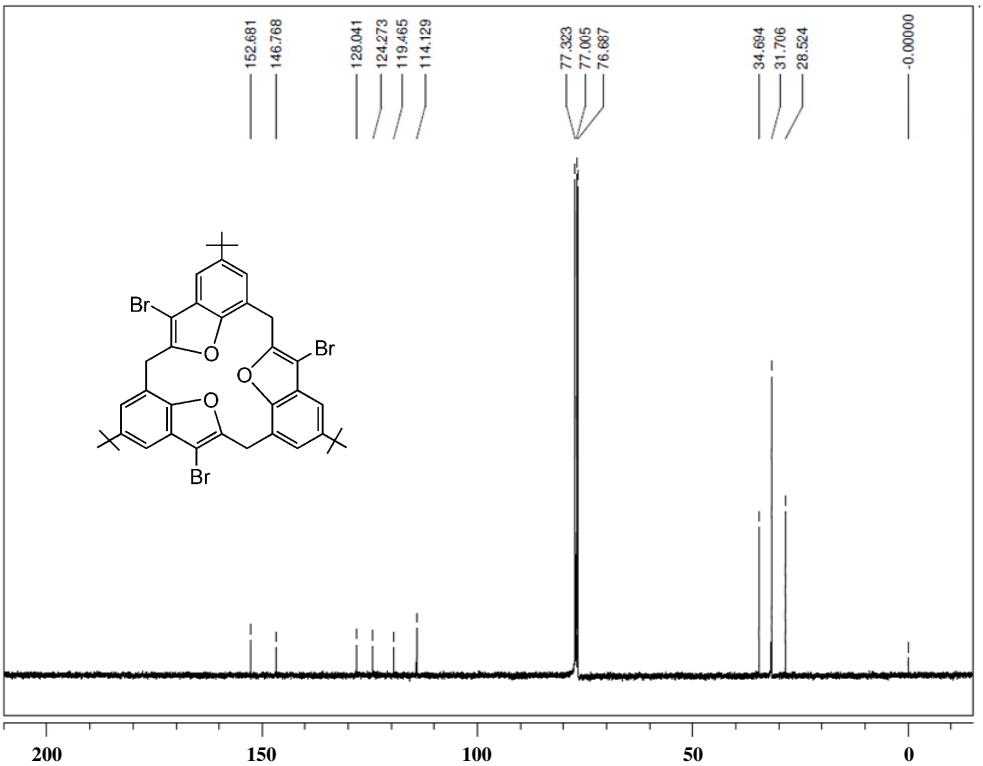


Figure S10: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **4b.**

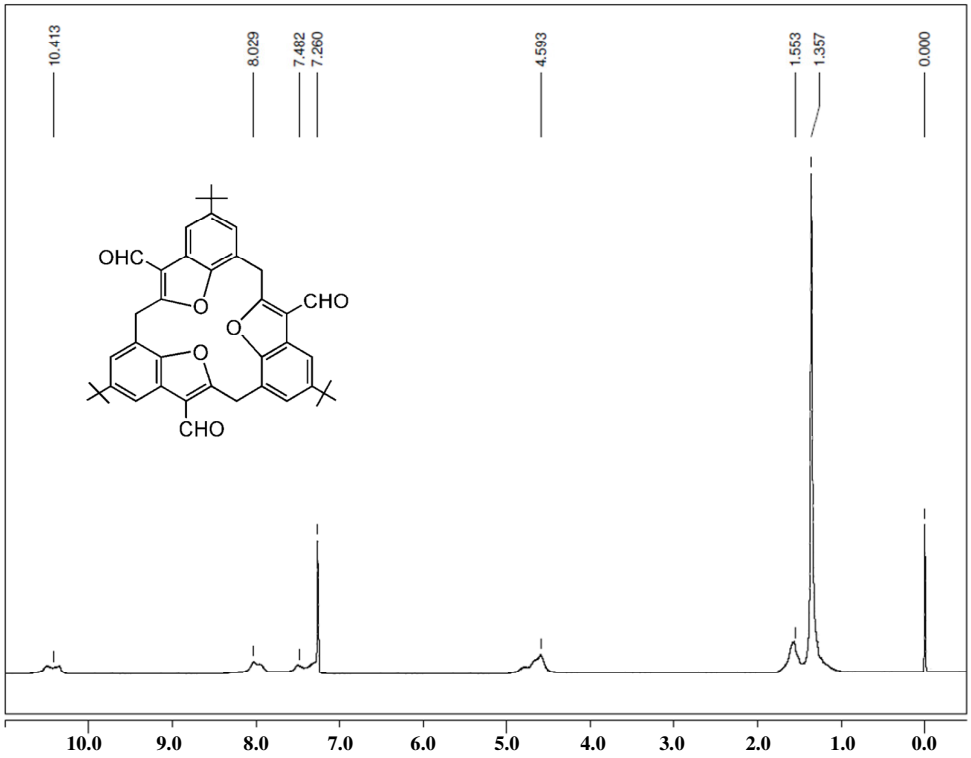


Figure S11: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **4c.**

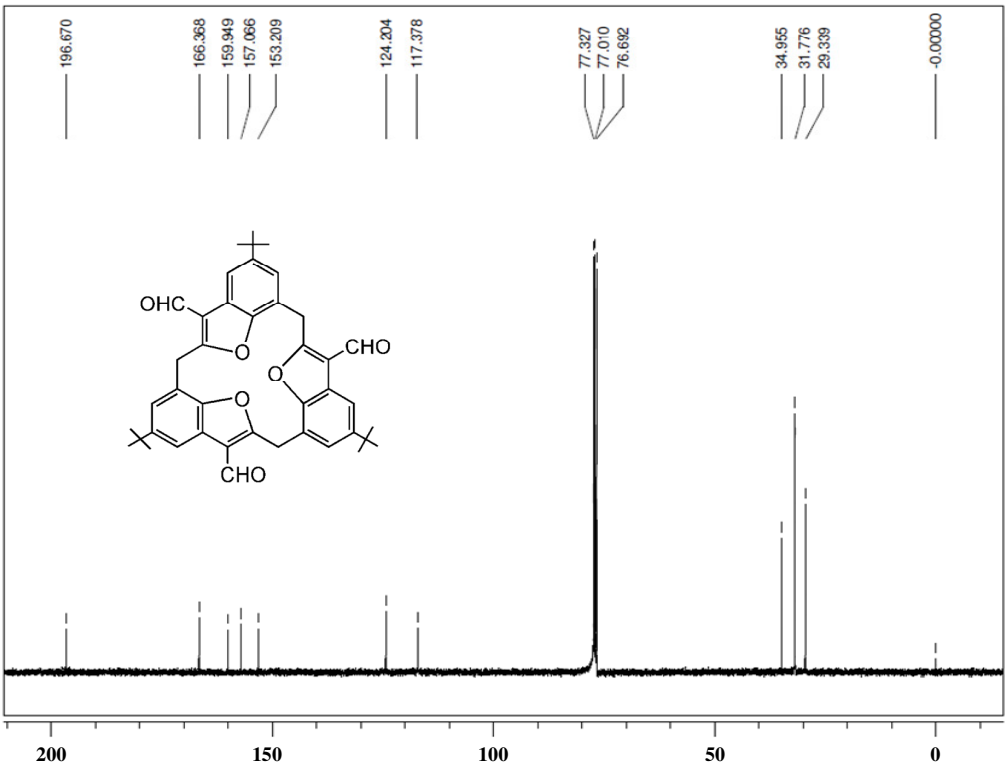


Figure S12: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **4c.**

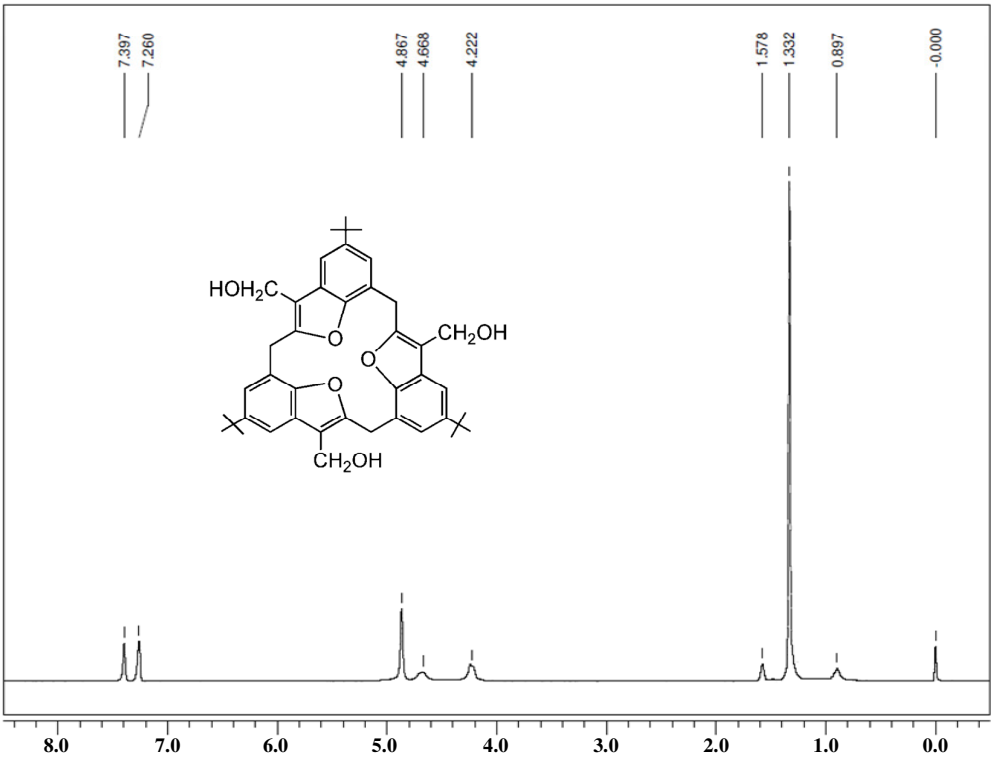


Figure S13: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **4d.**

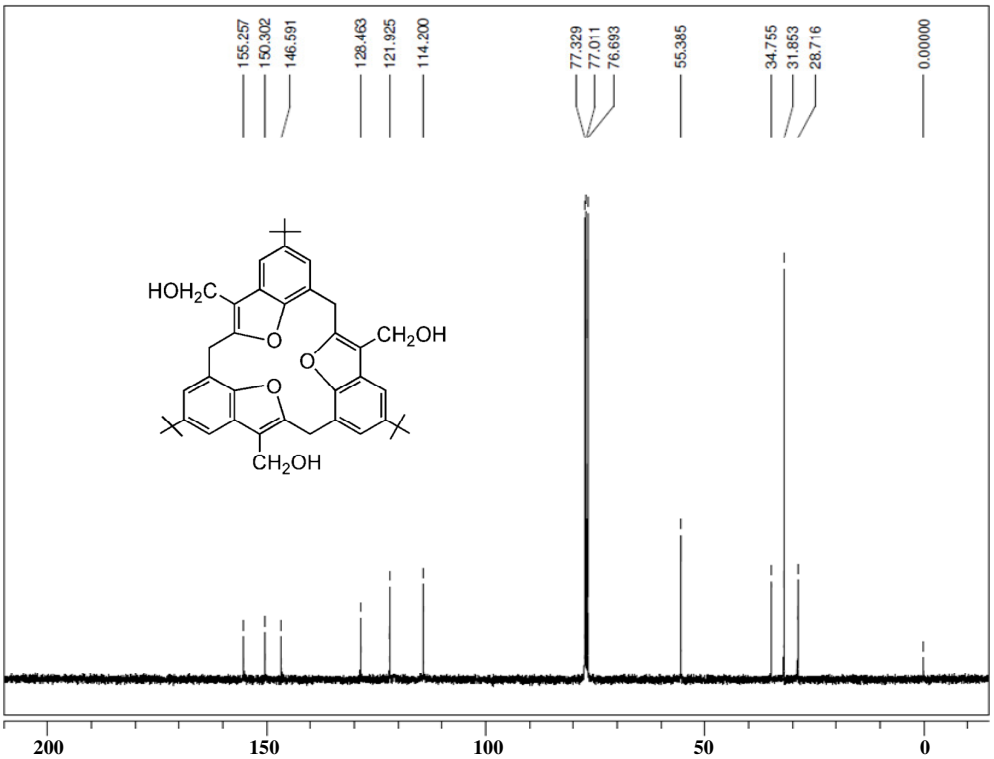


Figure S14: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) for the compound **4d.**

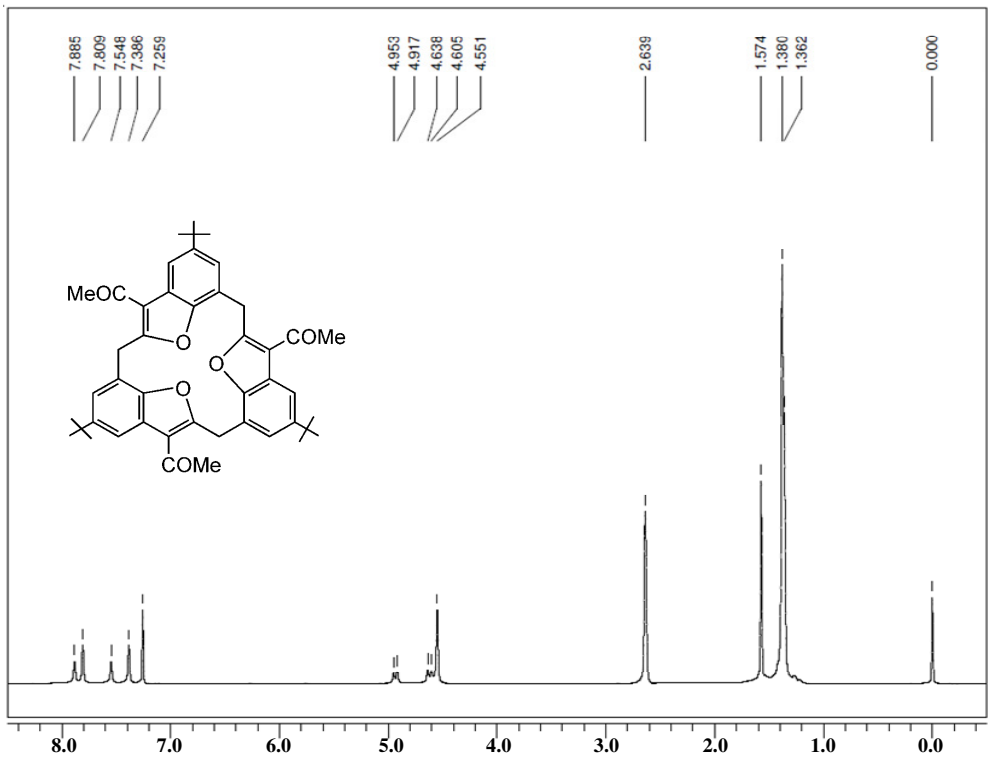


Figure S15: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **4e.**

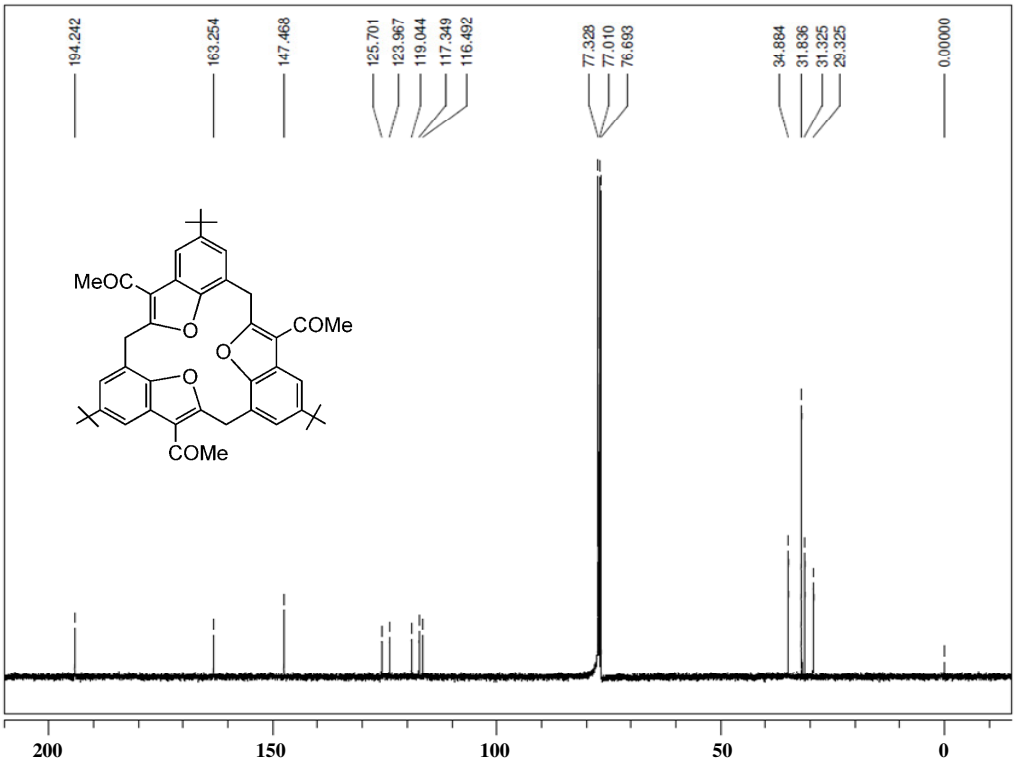


Figure S16: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) for the compound **4e.**

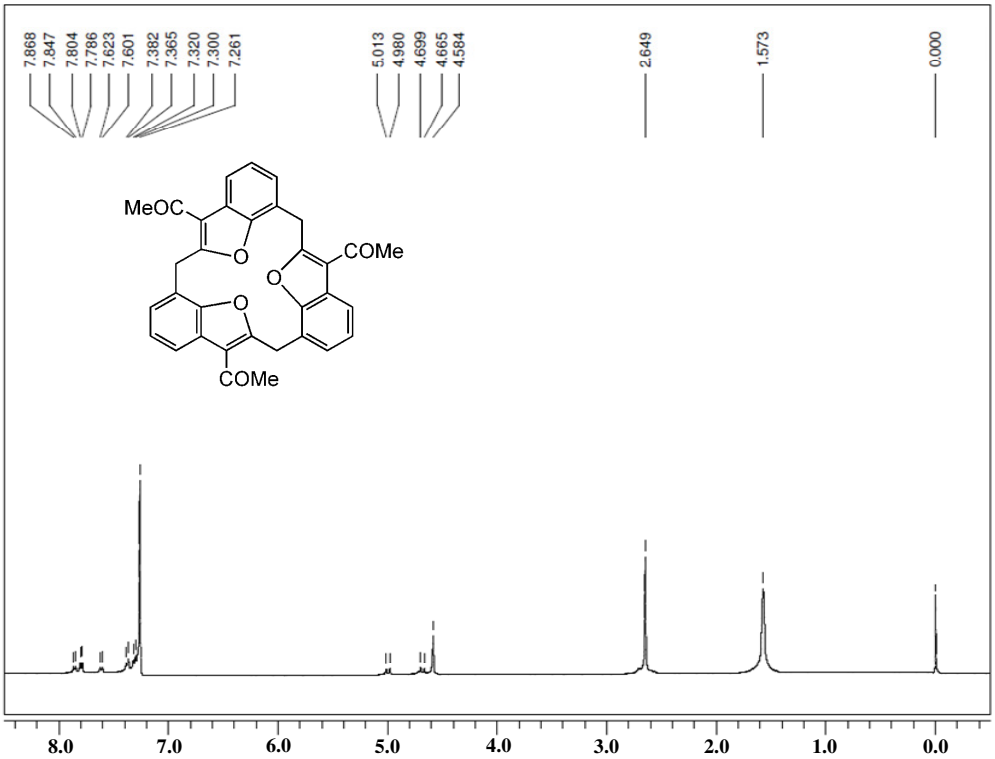


Figure S17: ^{1}H-NMR spectrum (400 MHz, 298 K, *$CDCl_3$) of the compound **5b**.

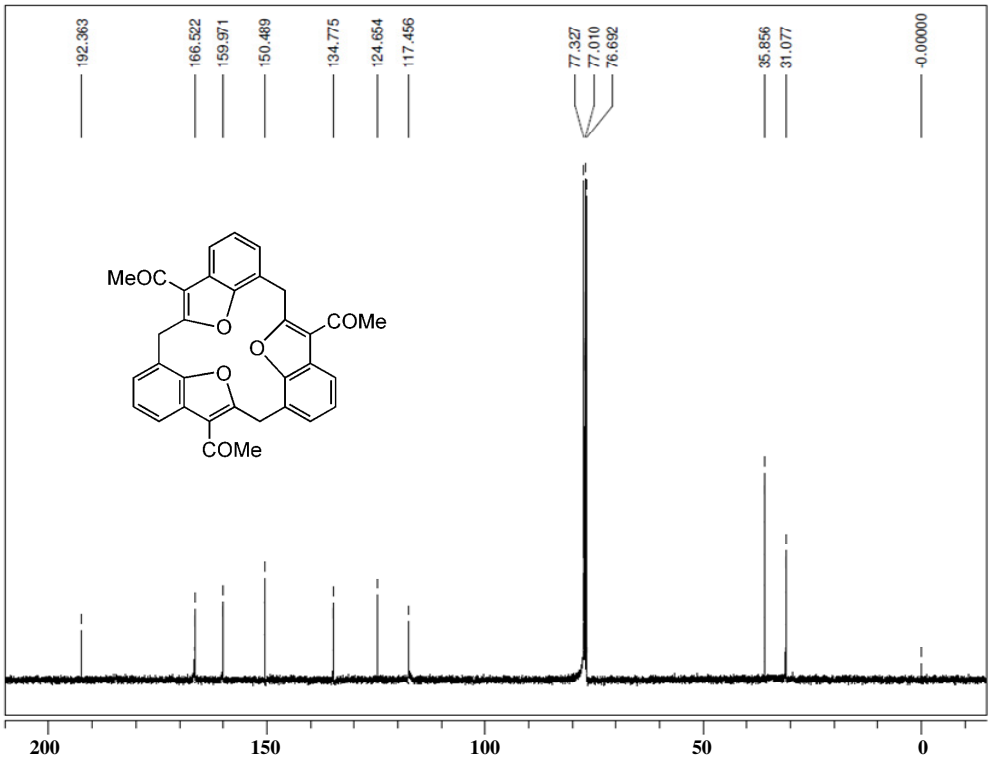


Figure S18: ^{13}C-NMR spectrum (100 MHz, 298 K, *$CDCl_3$) of the compound **5b**.

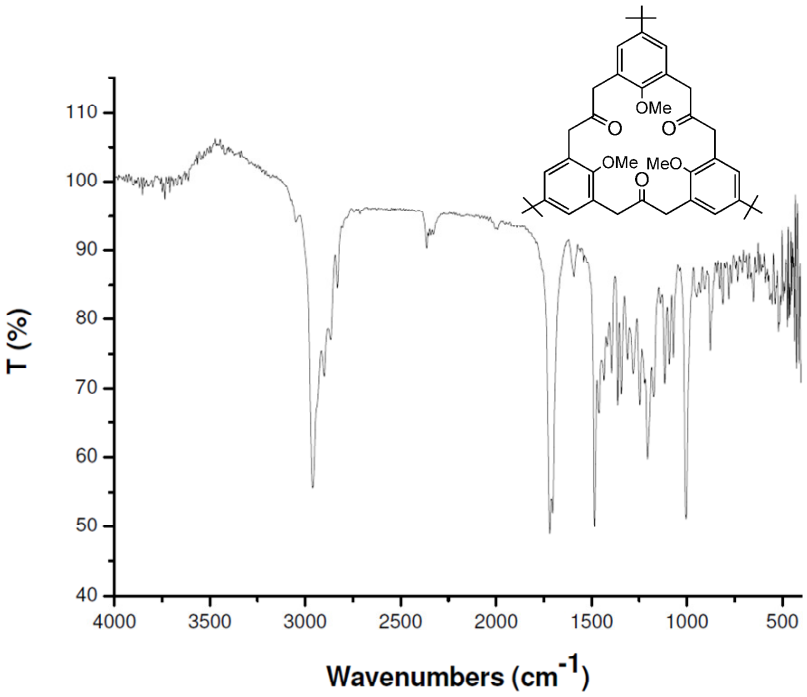


Figure S19: FT-IR spectrum of the compound **2a.**

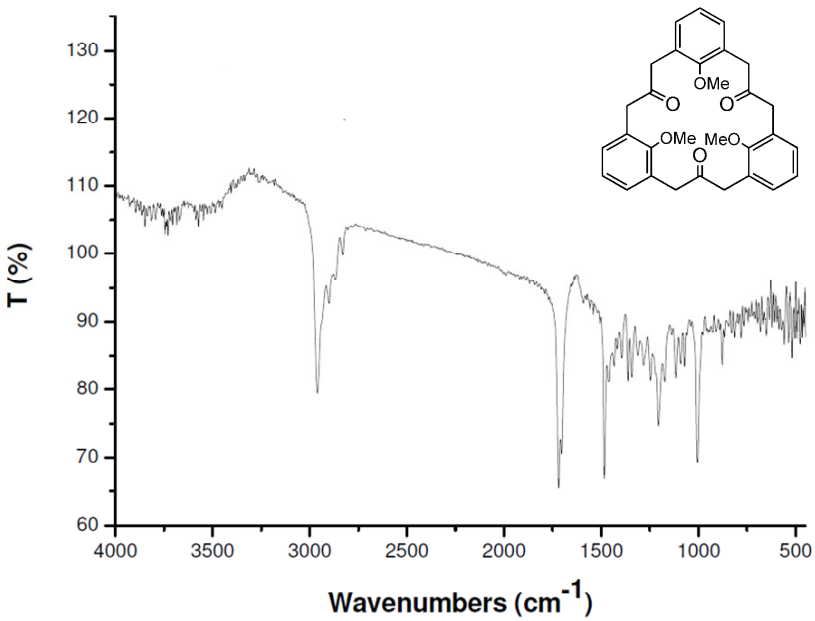


Figure S20: FT-IR spectrum of the compound **2b.**

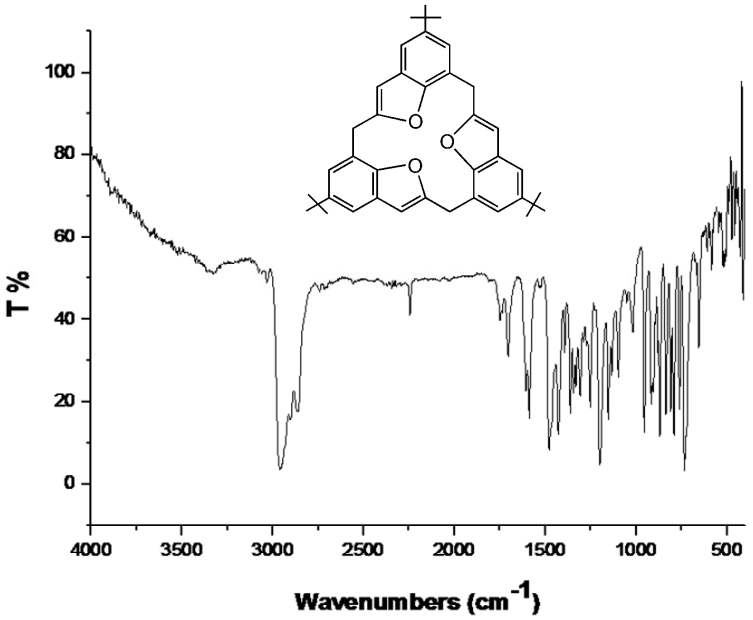


Figure S21: FT-IR spectrum of the compound **4a.**

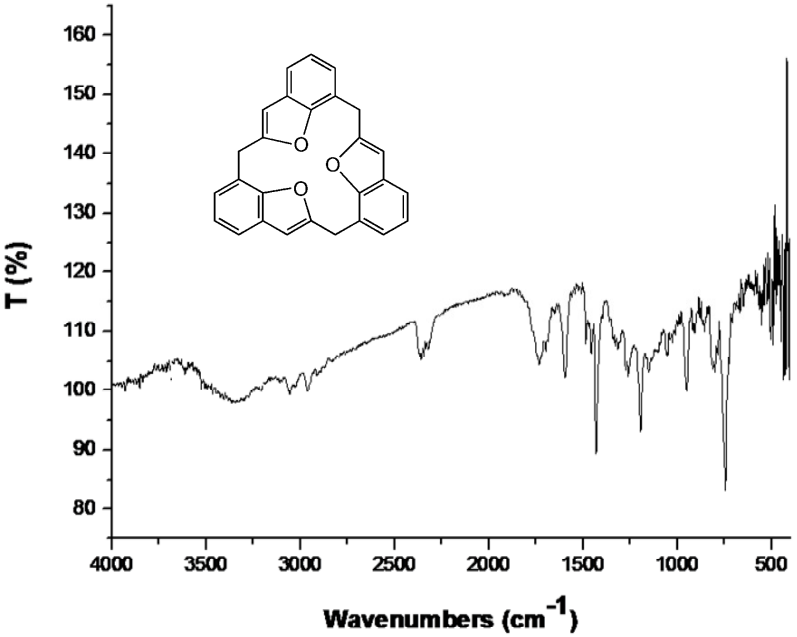


Figure S22: FT-IR spectrum of the compound **5a.**

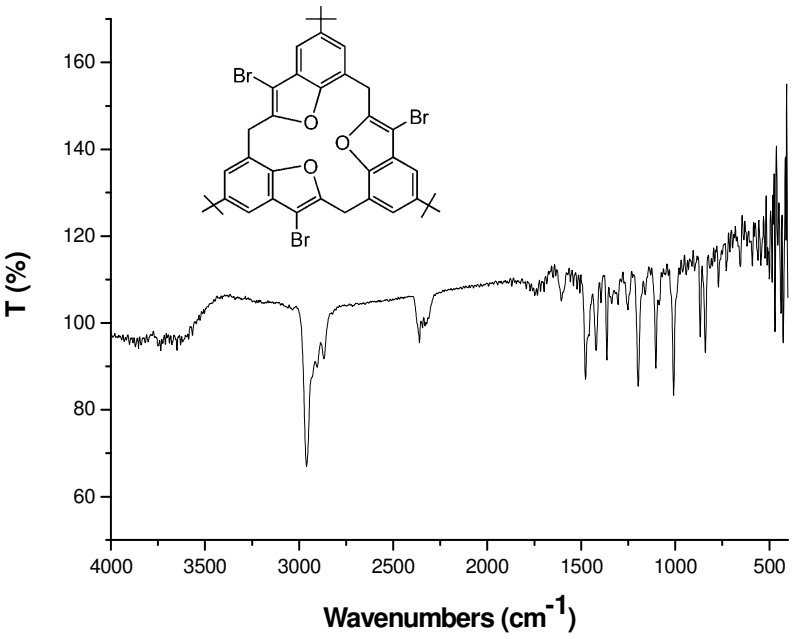


Figure S23: FT-IR spectrum of the compound **4b.**

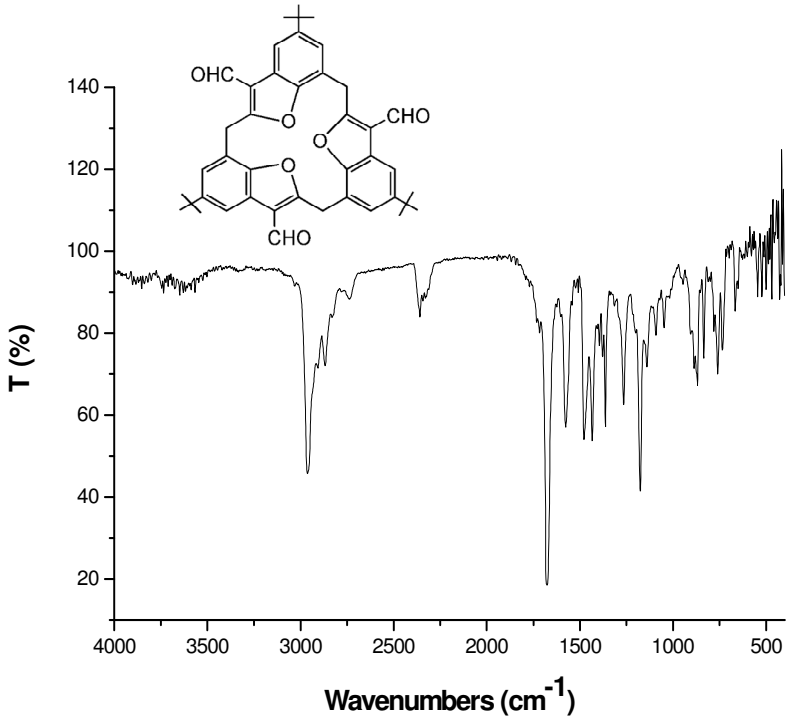


Figure S24: FT-IR spectrum of the compound **4c.**

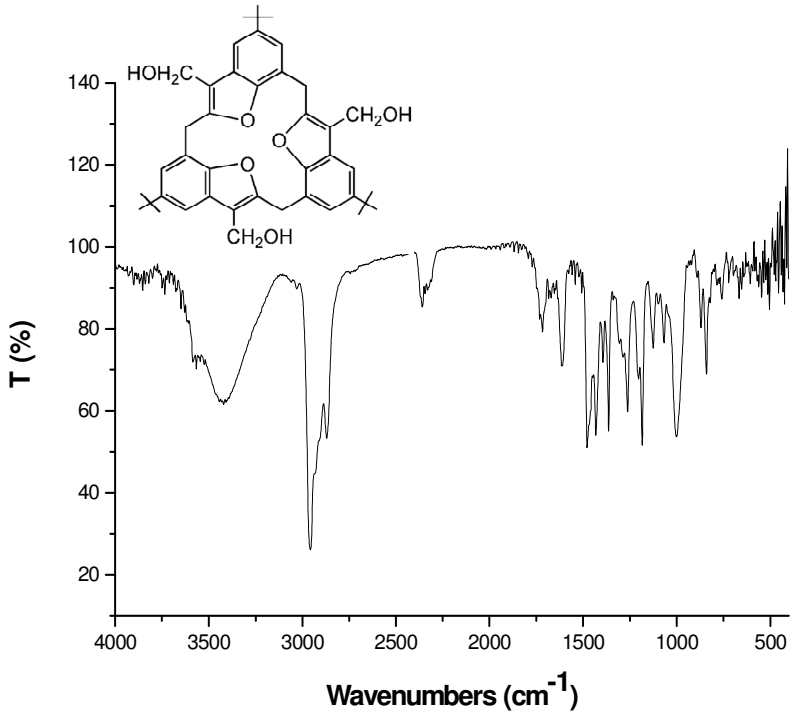


Figure S25: FT-IR spectrum of the compound **4d.**

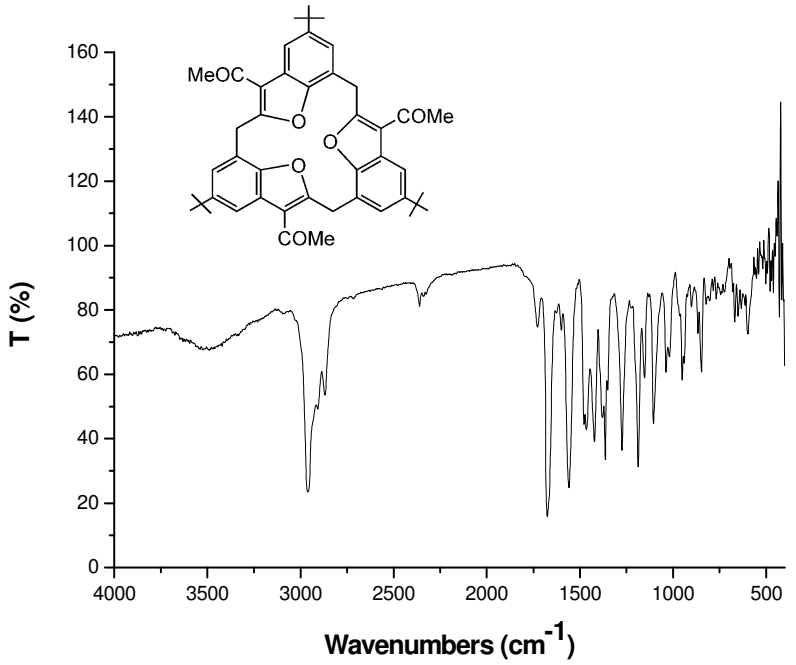


Figure S26: FT-IR spectrum of the compound **4e.**

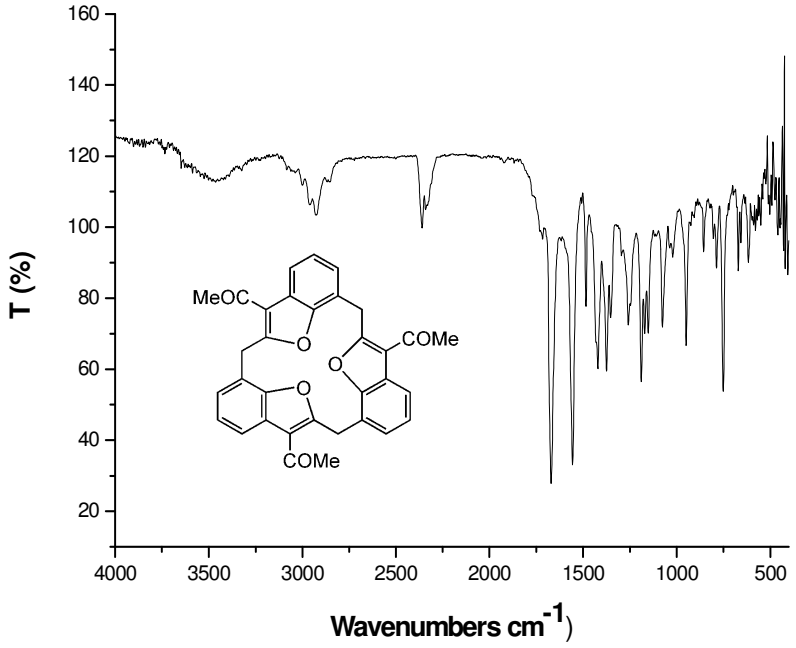


Figure S27: FT-IR spectrum of the compound **5b.**

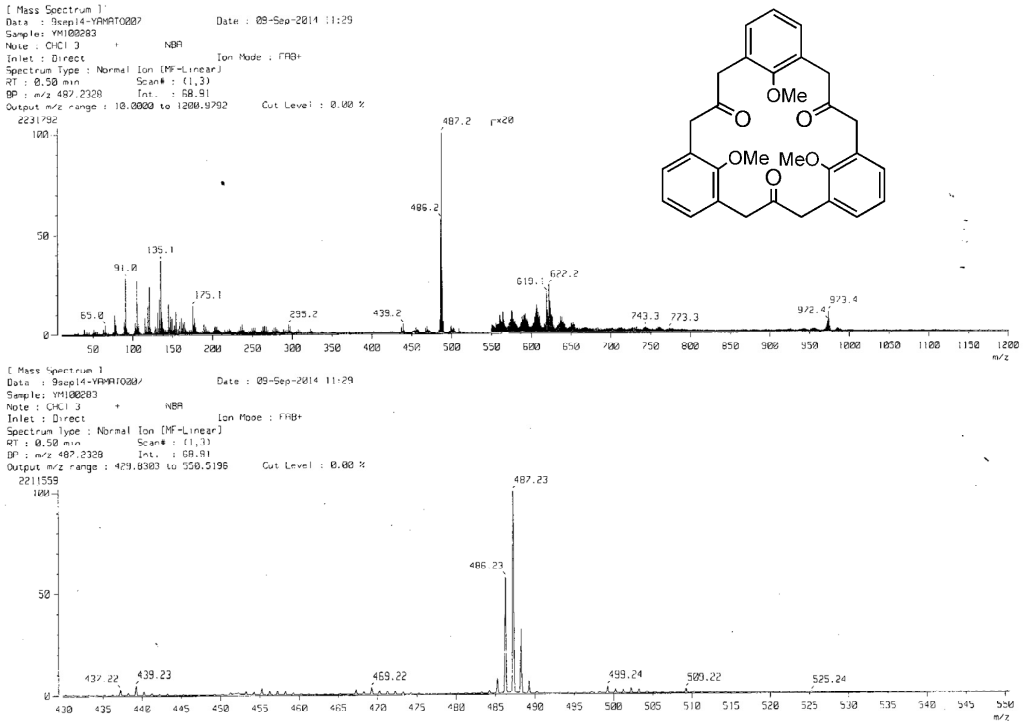


Figure S28: Mass-spectrum of the compound **2b**.

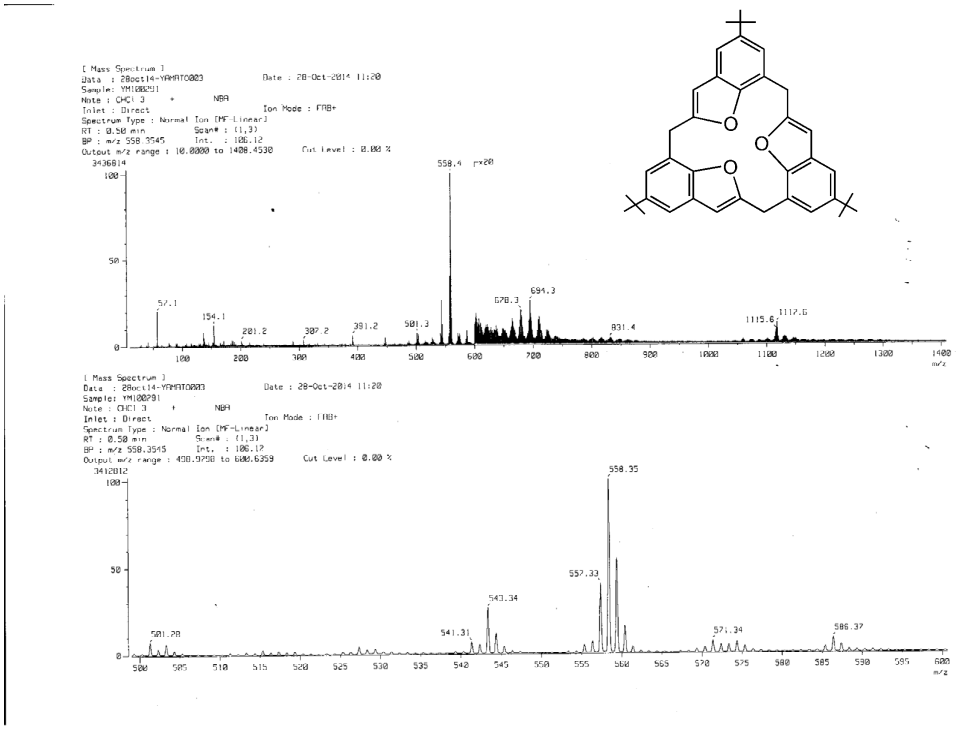


Figure S29: Mass-spectrum of the compound **4a**.

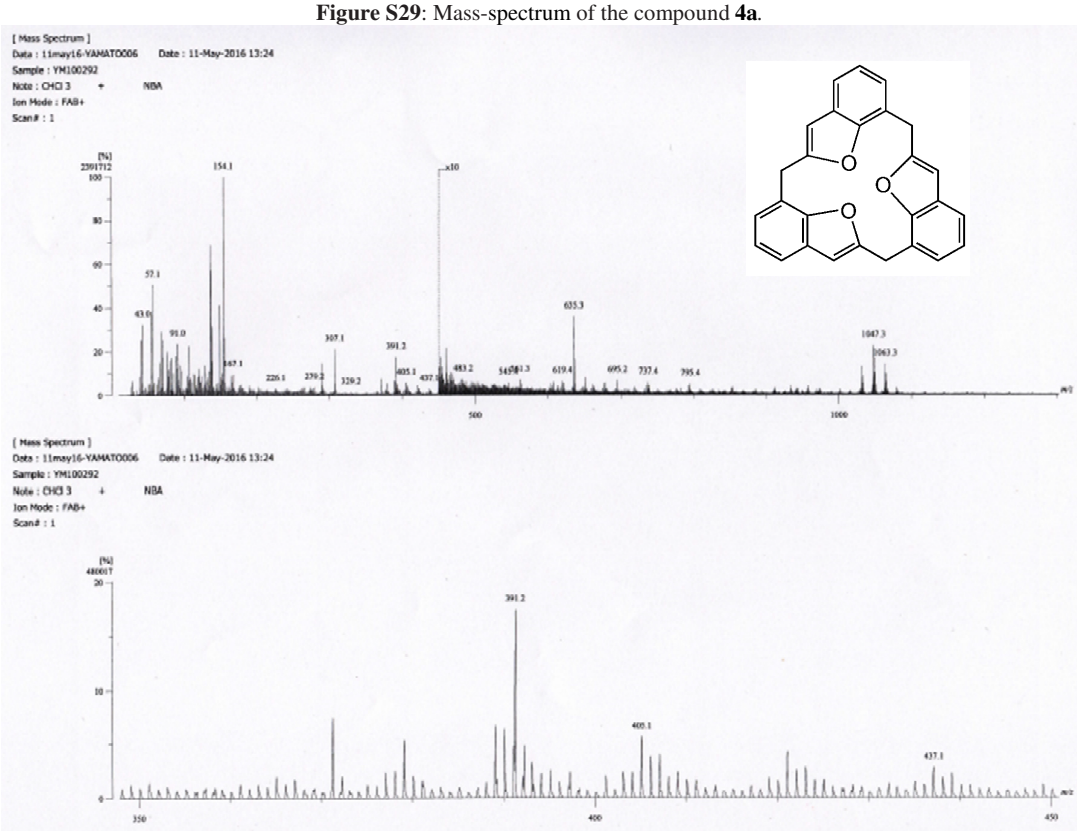


Figure S30: High resolution Mass-spectrum of the compound **5a**.

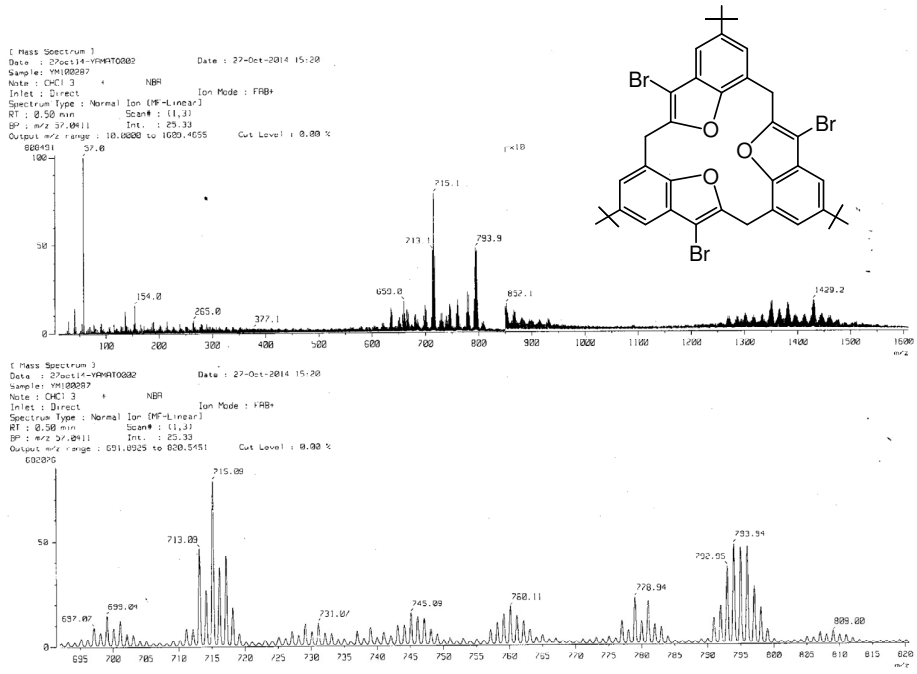


Figure S31: Mass-spectrum of the compound **4b**.

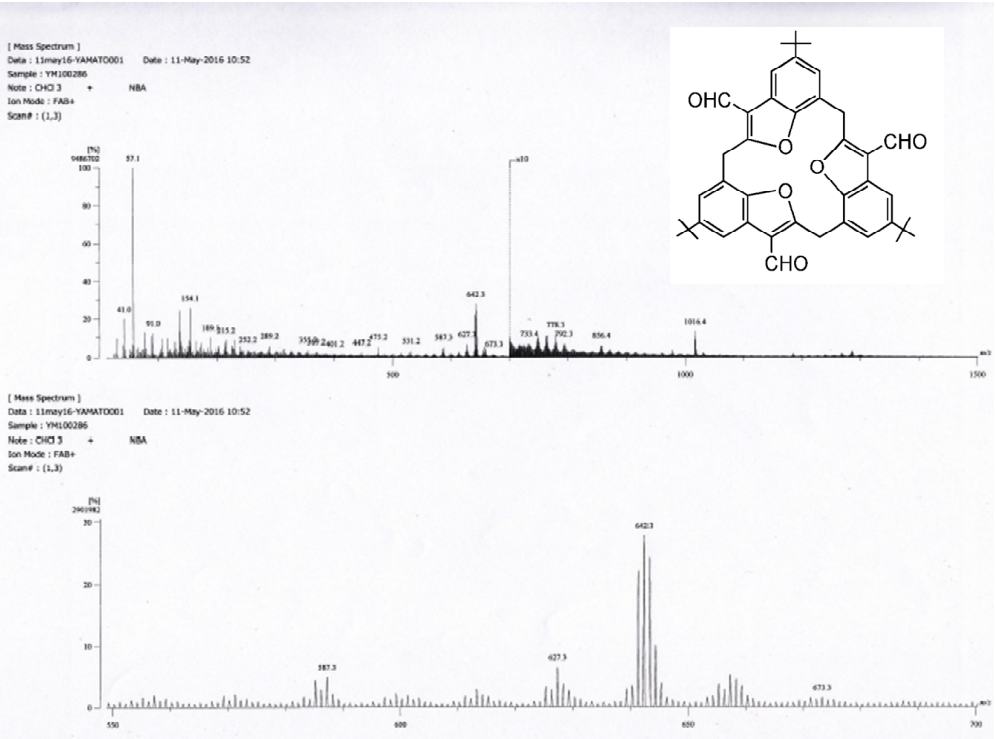


Figure S32: High resolution Mass-spectrum of the compound **4c**.

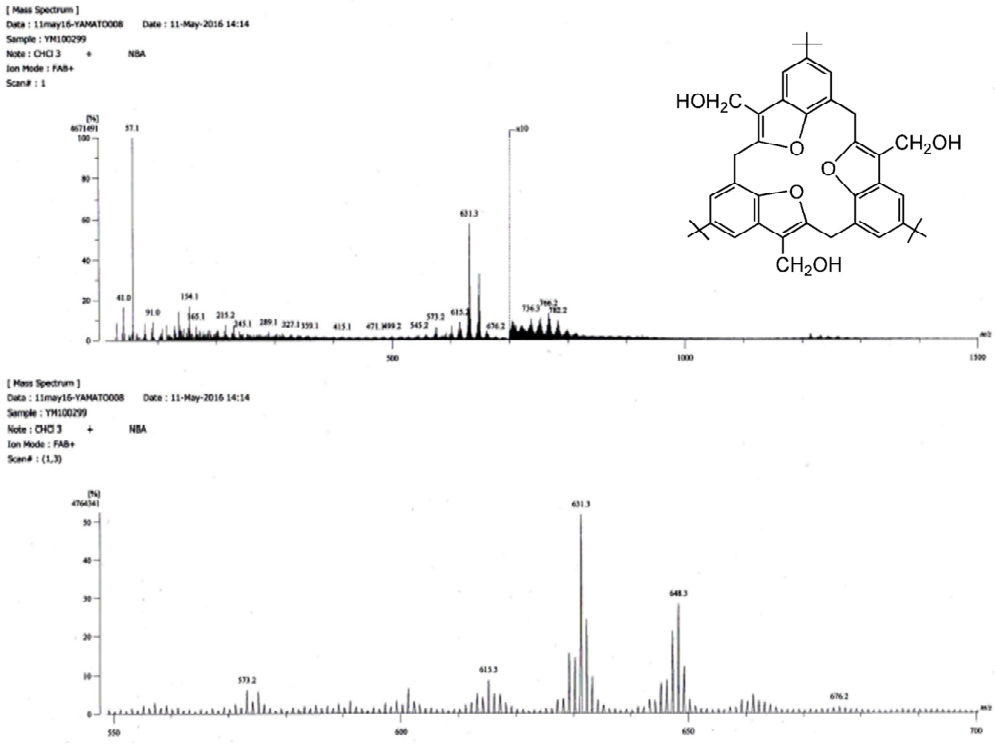


Figure S33: High resolution Mass-spectrum of the compound **4d**.

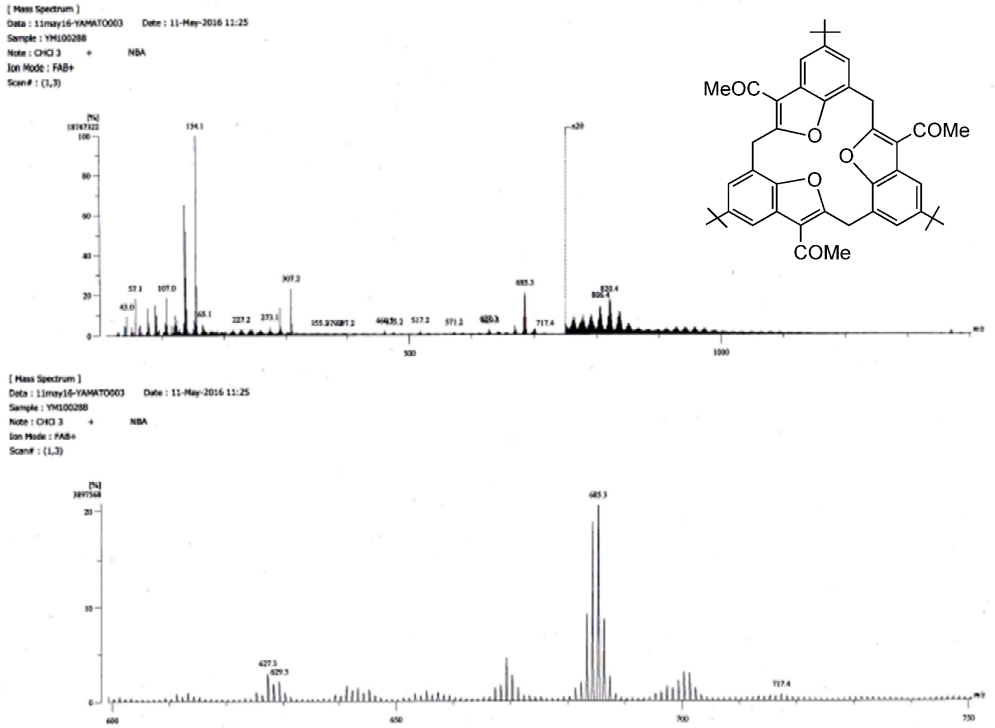


Figure S34: High resolution Mass-spectrum of the compound **4e**.

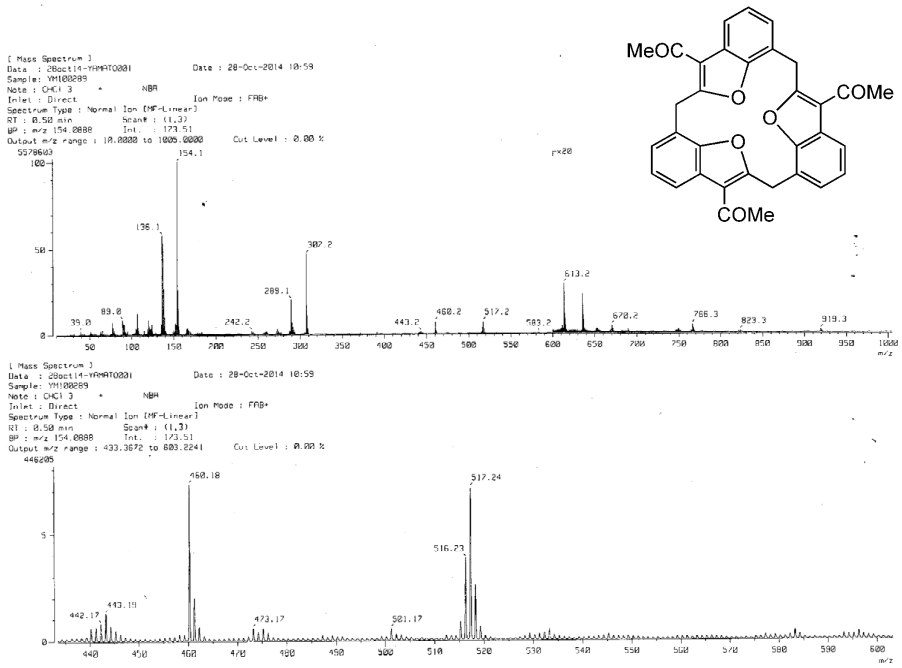


Figure S35: Mass-spectrum of the compound **5b**.

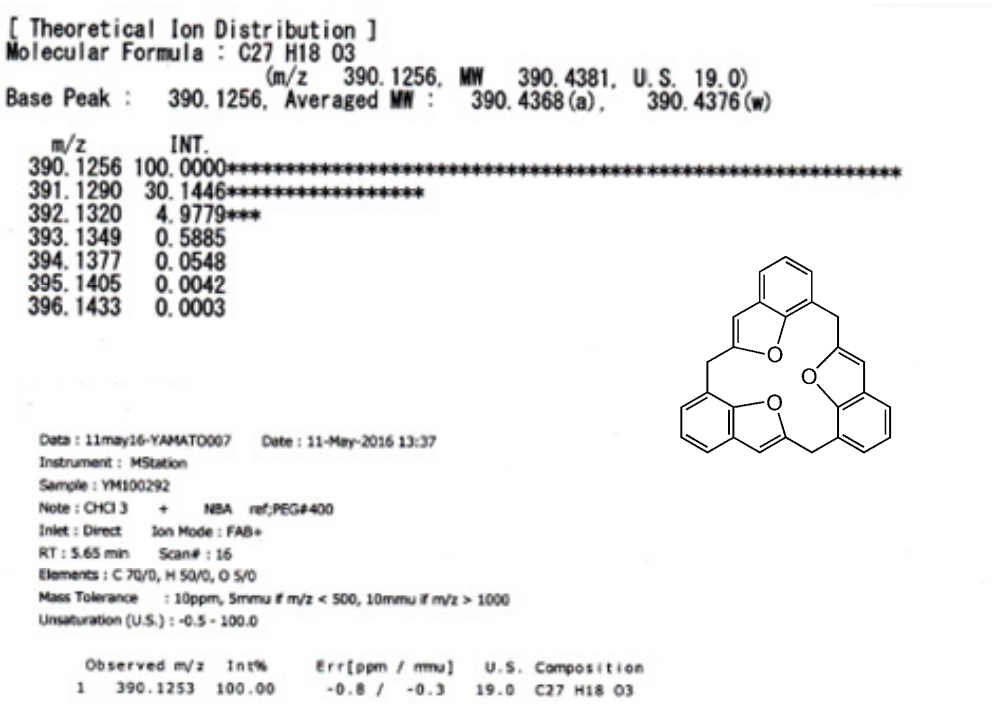
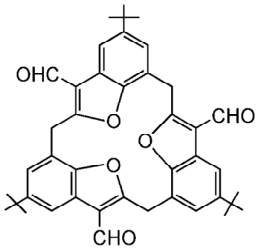


Figure S36: **HRMS** result of the compound **5a**.

[Theoretical Ion Distribution]
Molecular Formula : C42 H42 O6
(m/z 642.2981, MW 642.7919, U.S. 22.0)
Base Peak : 642.2981, Averaged MW : 642.7880(a), 642.7888(w)

m/z	INT.
642.2981	100.0000
643.3015	46.9424
644.3046	11.9610
645.3076	2.1682
646.3104	0.3101
647.3133	0.0369
648.3160	0.0038
649.3188	0.0003



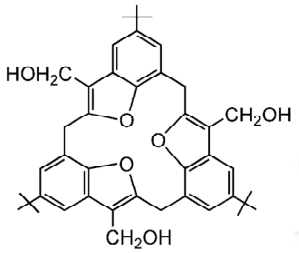
Data : 11may16-YAMATO002 Date : 11-May-2016 11:00
Instrument : MStation
Sample : YM100286
Note : CHCl 3 + NBA ref;PEG#600
Inlet : Direct Ion Mode : FAB+
RT : 0.00 min Scan# : 1
Elements : C 70/0, H 70/0, O 8/0
Mass Tolerance : 10ppm, 5mmu if m/z < 500, 10mmu if m/z > 1000
Unsaturation (U.S.) : -0.5 - 100.0

	Observed m/z	Int%	Err[ppm / mmu]	U.S.	Composition
1	642.2994	93.59	+2.0 / +1.3	22.0	C42 H42 O6

Figure S37: **HRMS** result of the compound **4c**.

[Theoretical Ion Distribution]
Molecular Formula : C42 H48 O6
(m/z 648.3451, MW 648.8395, U.S. 19.0)
Base Peak : 648.3451, Averaged MW : 648.8350(a), 648.8358(w)

m/z	INT.
648.3451	100.0000
649.3484	46.9424
650.3516	11.9610
651.3545	2.1682
652.3574	0.3101
653.3602	0.0369
654.3630	0.0038
655.3657	0.0003



Data : 11may16-YAMATO009 Date : 11-May-2016 14:23
Instrument : MStation
Sample : YM100299
Note : CHCl 3 + NBA ref;PEG#600
Inlet : Direct Ion Mode : FAB+
RT : 0.25 min Scan# : 2
Elements : C 70/0, H 70/0, O 8/0
Mass Tolerance : 10ppm, 5mmu if m/z < 500, 10mmu if m/z > 1000
Unsaturation (U.S.) : -0.5 - 100.0

	Observed m/z	Int%	Err[ppm / mmu]	U.S.	Composition
1	648.3472	61.54	+3.3 / +2.1	19.0	C42 H48 O6

Figure S38: **HRMS** result of the compound **4d**.

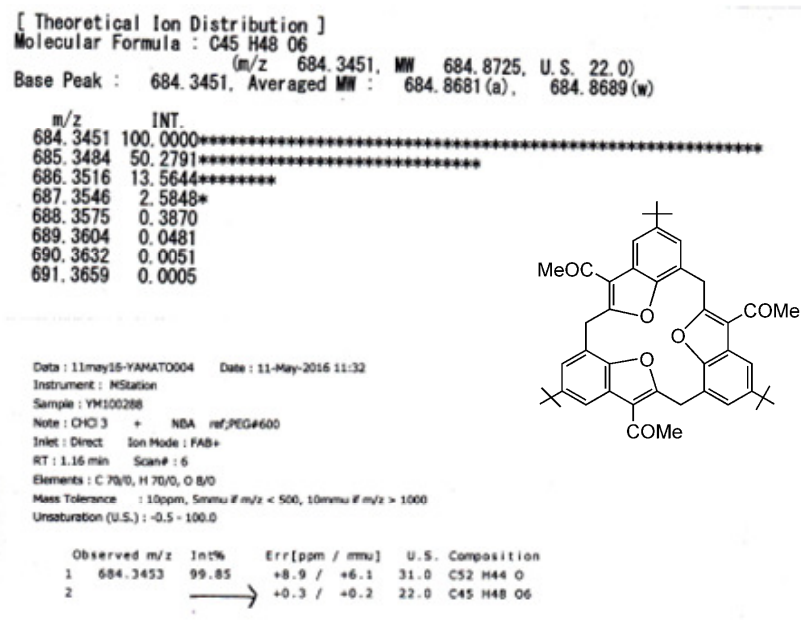


Figure S39: **HRMS** result of the compound **4e**.

X-ray crystallography

Table S1: Summary of crystal data for **4a** and **4b**

Table S1 Summary of crystal data for **4a** and **4b**.[a,b]

Parameter	**4a**	**4b**
Empirical formula	C_{39} H_{42} O_3	C_{39} H_{39} Br_3O_3, $CHCl_3$, *ca* 6(CH_4O)
Formula weight [g mol^{-1}]	558.76	1107.05
Crystal system	trigonal	trigonal
Space group	*R-3*	*R-3*
A [Å]	19.3942(14)	15.4391(12)
B [Å]	19.3942(14)	15.4391(12)
C [Å]	16.5448(6)	31.782 (4)
α [°]	90.0000	90.0000
β [°]	90.0000	90.0000
γ [°]	120.0000	120.0000
Volume [Å^3]	5389.4(6)	6560.8(13)
Z	6	6
Density, calcd [g m^{-3}]	1.033	1.681
Temperature [K]	123	140
Unique reflns	2193	1888
Obsd reflns	1984	1293
Parameters	127	182
R_{int}	0.0399	0.076
R[I>2σ(I)][a]	0.0647	0.072
wR[I>2σ(I)][b]	0.1455	0.138
GOF on F^2	1.154	1.119

[a] Conventional *R* on F_{hkl}: $\Sigma||F_o| - |F_c||/\sigma|F_o|$. [b] Weighted *R* on $|F_{hkl}|^2$: $\Sigma[w(F_o^2 - F_c^2)^2]/\Sigma[w(F_o^2)^2]^{1/2}$

Crystal structure analysis of the tris-benzofuran calixarene, compound 4a:[1–4]

Crystal data: $C_{39}H_{42}O_3$. M = 558.76. Trigonal, space group R-3 (no. 148; hexagonal axes), a = b = 19.3942(14), c = 16.5448(6) Å, α = β = 90, γ = 120 °, V = 5389.4(6) Å^3. Z = 6, Dc = 1.033 g cm^{-3}, F(000) = 3408, T = 123(1) K, μ(Cu-Kα) = 4.96 cm^{-1}, λ(Mo-Kα) = 1.5418 Å.

A colorless prism crystal of $C_{39}H_{42}O_3$ having approximate dimensions of 0.350 x 0.250 x 0.200 mm was mounted on a glass fiber. All intensity measurements were made on a Rigaku R-AXIS RAPID diffractometer using graphite monochromated Cu-Kα radiation. The data were collected at a temperature of -150 ± 1°C to a maximum 2θ value of 136.4°. Of the 20857 reflections collected, 2193 were unique (Rint = 0.0399) and 1984 were 'observed'; equivalent reflections were merged. The linear absorption coefficient, μ, for Cu-Kα radiation is 4.956 cm^{-1}. An empirical absorption correction was applied which resulted in transmission factors ranging from 0.641 to 0.906. The data were corrected for Lorentz and polarization effects. The structure was solved by direct methods[1] and expanded using Fourier techniques. The non-hydrogen atoms were refined anisotropically. Hydrogen atoms were refined using the riding model. The final cycle of full-matrix least-squares refinement on F^2 was based on all 2193 reflections and 127 variable parameters and converged with R_1 = 0.071 and wR_2 = 0.146; for the observed data, R_1 = 0.065. The goodness of fit was 1.15. The maximum and minimum peaks on the final difference Fourier map corresponded to 0.36 and -0.33 eÅ^{-3}, respectively.

Neutral atom scattering factors were taken from International Tables for Crystallography (IT), Vol. C, Table 6.1.1.4.[2] All calculations were performed using the CrystalStructure[3] crystallographic software package except for refinement, which was performed using SHELXL2013.[4]

Crystal structure analysis of a tris-benzofuran calixarene/$CHCl_3$/methanol complex, compound 4b:[5–8]

Crystal data: $C_{39}H_{39}Br_3O_3$, $CHCl_3$, *ca* 6(CH_4O). M = 1107.1. Trigonal, space group R-3 (no. 148; hexagonal axes), a = b = 15.4391(12), c = 31.782(4) Å, α = β = 90, γ = 120 °, V = 6560.8(13) Å^3. Z = 6, Dc = 1.681 g cm^{-3}, F(000) = 3408, T = 140(1) K, μ(Mo-Kα) = 30.1 cm^{-1}, λ(Mo-Kα) = 0.71073 Å.

Crystals are colorless, cubic blocks. One, *ca* 0.43 x 0.37 x 0.25 mm, was mounted in oil on a glass fiber and fixed in the cold nitrogen stream on an Oxford Diffraction Xcalibur-3 CCD diffractometer equipped with Mo-Kα radiation and graphite monochromator. Intensity data were measured by thin-slice ω- and φ-scans. Total no. of reflections recorded, to θ_{max} = 22.5°, was 13024 of which 1888 were unique (Rint = 0.076); 1293 were 'observed' with $I > 2\sigma_I$.

Data were processed using the CrysAlis-CCD and -RED (1) programs. The structure was determined by the direct methods routines in the SHELXS program (2A) and refined by full-matrix least-squares methods, on F^2's, in SHELXL (2B). The analysis shows the calixarene molecule lying around a threefold symmetry axis with, on one side, a $CHCl_3$ solvent molecule (also on the symmetry axis) and, on the other side, a complex ring structure, presumably of a disordered array of methanol molecules; in this region, 50 atoms have been refined as isotropic carbon atoms, mostly with site occupancies of 0.5, about a point of *-3* symmetry. In the calixarene and chloroform molecules, the non-hydrogen atoms were refined with anisotropic thermal parameters; hydrogen atoms were included in idealized positions and their Uiso values were set to ride on the Ueq values of the parent carbon atoms. At the conclusion of the refinement, wR_2 = 0.149 and R_1 = 0.114 (2B) for all 1888 reflections weighted $w = [\sigma^2(F_o^2) + (0.0368P)^2 + 104.34P]^{-1}$ with $P = (F_o^2 + 2F_c^2)/3$; for the 'observed' data only, R_1 = 0.072.

In the final difference map, the highest peak (*ca* 0.37 eÅ^{-3}) was close to Br(13).

Scattering factors for neutral atoms were taken from reference.[7] Computer programs used in this analysis have been noted above, and were run through WinGX (4) on a Dell Optiplex 755 PC at the University of East Anglia.

General description for the DFT computational study:[9,10]

Density functional theory (DFT) computational studies were carried out to determine the geometry-optimized energies of compounds **4a**-**e** and **5a**-**b**. The starting structures were generated with the initial geometries based upon the X-ray structures of **4a** and 4**b** and from the presumed structures of **4c**–**4d** (derived from *cone*-**4a** and *cone*-4**b**) and **5a**–**b** using *SpartanPro'10* with the MMFF94 method.[9] The individual geometry-optimized structures of these molecules were first conducted in the gas phase and then in solvent (chloroform) with the B3LYP/6-31G(d) basis set using Gaussian-*09.*[10] The results are summarized in Tables S2 and S3 for both *cone* and *saddle* conformations for compounds **4a**–**e**, **5a**–**b** (Figures S36 to S42). The results presented in Table S2 show that **4a**–**e**, **5a**–**b** were energetically more-favoured in solvent $CHCl_3$ than in the gas phase. The results presented in Table S3 of the synthesized calix[3]benzofurans and their derivatives, **4a**–**e**, suggest that the *saddle* conformers are more stable than the *cone* isomers. The results presented in Tables S2 & S3 show that among the calix[3]benzofurans, **4b** is the energetically most-favoured (in both the solvent and gas-phase) and the order is as follows: **4b>4e>4d>4c >4a>5b >5a** in both the solvent and gas phase. So by introducing the different groups at the furan moieties, the derivatives become energetically more favored over the corresponding calix[3]benzofuran according to the increasing size of groups (i.e. COMe > CH_2OH > CHO) except for **4b**. In the case of **4b**, there may be two factors influencing the stability: bromine is electronegative in nature and has greater electron-density due to multiple lone-pairs of electrons.

The DFT optimized B3LYP/6-31G(d) energies of these two conformers imply that the *saddle* conformers of **4a** and **4b**, which are -4 and -35 $kJmol^{-1}$, are therefore more stable than the *cone* conformers in the solvent, similar to what was computed in gas phase (Table S2). On the other hand, for the *tert*-butyl group analogues, calix[3]benzofuran **5a** and its derivative **5b**, the *saddle* conformers are energetically less stable than the *cone* conformers by 4 and 10 $kJmol^{-1}$ in the gas phase, and by 5 and 7 $kJmol^{-1}$ in solvent (Table S3), respectively. Similarly, *saddle*-**4c**, **4d**, **4e** are energetically more stable by -12, -20 and -48 $kJmol^{-1}$ than *cone*-**4c**, **4d**, **4e** in the gas phase, respectively (Table S3).

Table S2. Geometry optimization energies using B3LYP/6-31G(d)($\Delta E=E_{chlorofom}-E_{gas\text{-}phase}$).

	Cone			***Saddle***		
	Gas phase	Chloroform	ΔE	Gas phase	Chloroform	ΔE
Compound	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}	kJ mol-^{1}
4a	-4560873	-4560891	-17	-4560878	-4560895	-17
5a	-3322273	-3322291	-18	-3322269	-3322286	-17
4b	-24812139	-24812152	-13	-24812173	-24812188	-15
4c	-5453483	-5453506	-23	-5453495	-5453524	-28
4d	-5462889	-5462924	-35	-5462910	-5462940	-31
4e	-5763178	-5763204	-26	-5763226	-5763254	-28
5b	-4524625	-4524650	-25	-4524615	-4524643	-28

Table S3. Geometry optimization energies using B3LYP/6-31G(d)($\Delta E = E_{Saddle} - E_{Cone}$).

Compound	Gas-phase			Chloroform		
	Cone	*Saddle*	ΔE	*Cone*	*Saddle*	ΔE
	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}	kJ mol^{-1}
4a	-4560873	-4560878	-4	-4560891	-4560895	-4
5a	-3322273	-3322269	4	-3322291	-3322286	5
4b	-24812139	-24812173	-34	-24812152	-24812188	-35
4c	-5453483	-5453495	-12	-5453506	-5453524	-18
4d	-5462889	-5462910	-20	-5462924	-5462940	-16
4e	-5763178	-5763226	-48	-5763204	-5763254	-50
5b	-4524625	-4524615	10	-4524650	-4524643	7

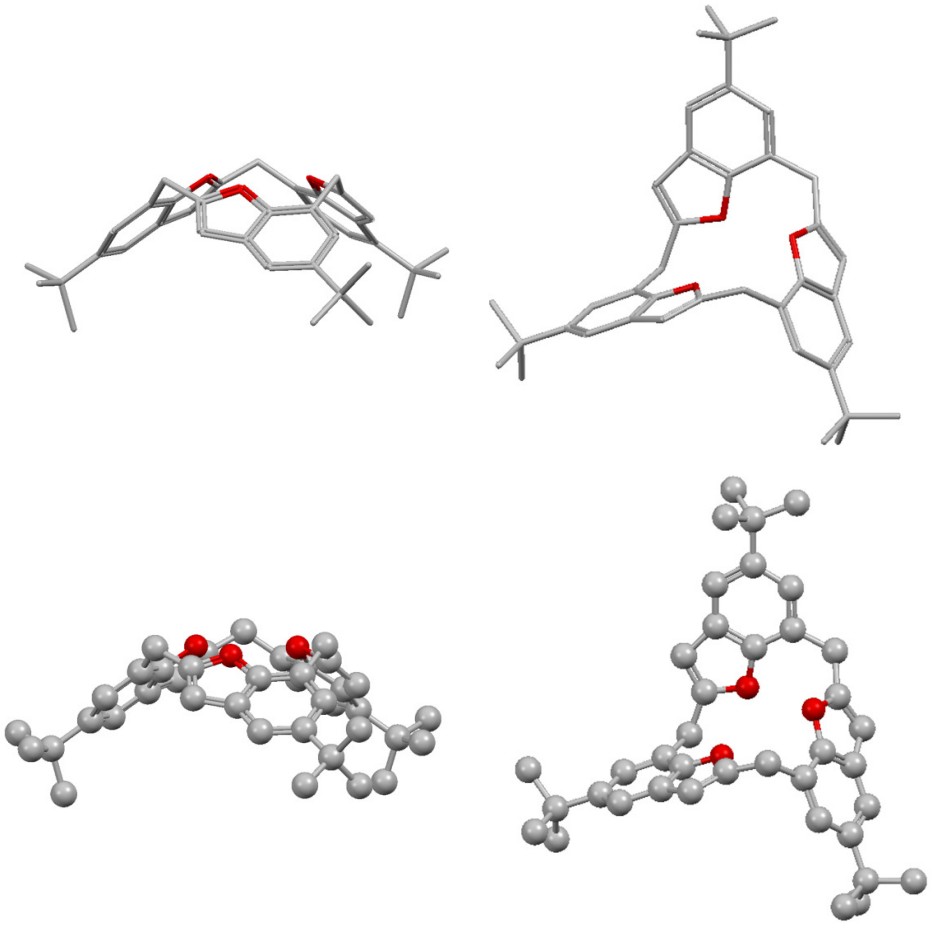


Figure S40. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **4a** *cone* (Ellipsoid); *Top Right*: **4a** *saddle* (Ellipsoid). *Bottom Left*: **4a** *cone* (ball-and-stick) and *Bottom Right*: **4a** *saddle* (ball-and-stick). Colour code: carbon = dark grey and oxygen atom = red. All hydrogens are omitted for clarity.

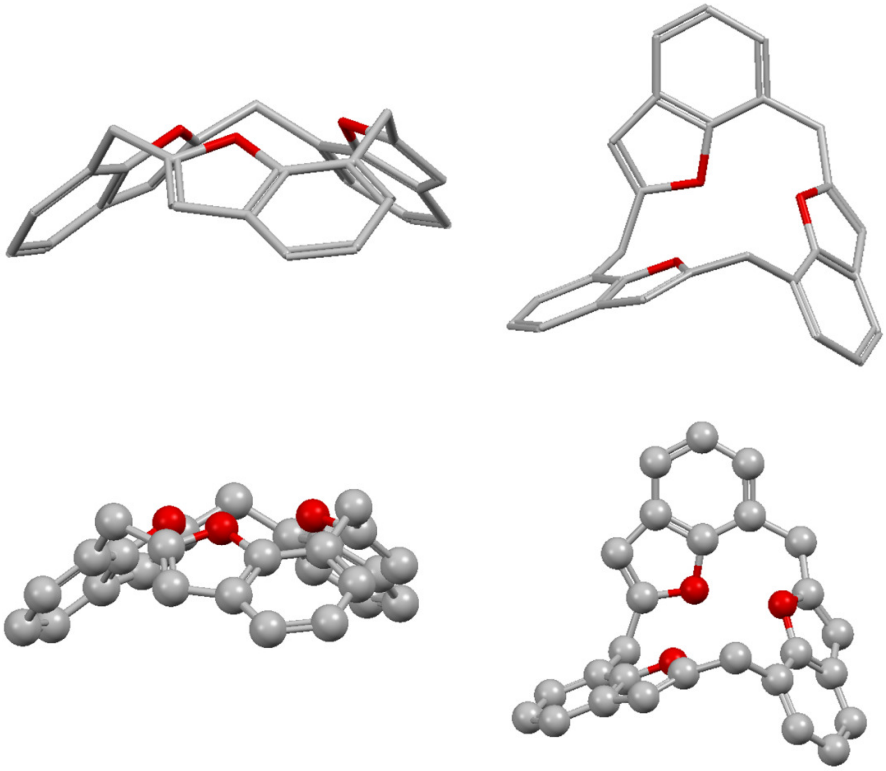


Figure S41. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **5a** *cone* (Ellipsoid); *Top Right*: **5a** *saddle* (Ellipsoid). *Bottom Left*: **5a** *cone* (ball-and-stick) and *Bottom Right*: **5a** *cone* (ball-and-stick). Colour code: carbon = dark grey and oxygen atom = red. All hydrogens are omitted for clarity.

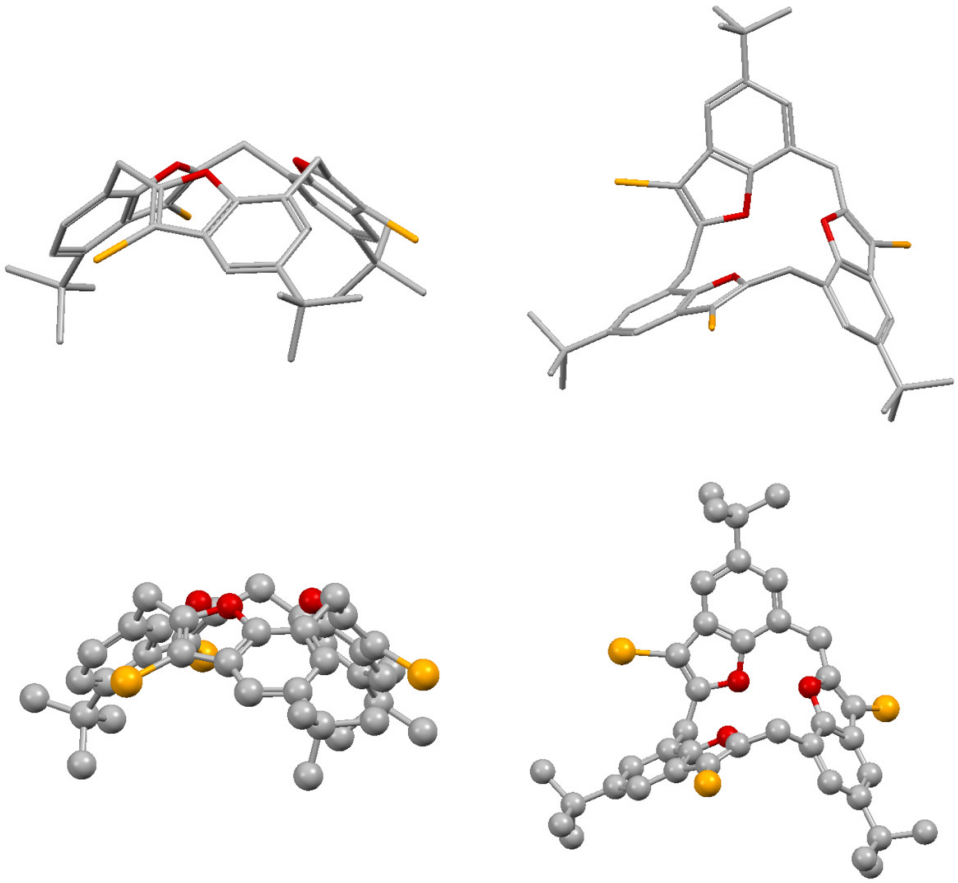


Figure S42. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **4b** *cone* (Ellipsoid); *Top Right*: **4b** *saddle* (Ellipsoid). *Bottom Left*: **4b** *cone* (ball-and-stick); and *Bottom Right*: 4b *saddle* (ball-and-stick). Colour code: bromide = orange, carbon = dark grey and oxygen atom = red. All hydrogens are omitted for clarity.

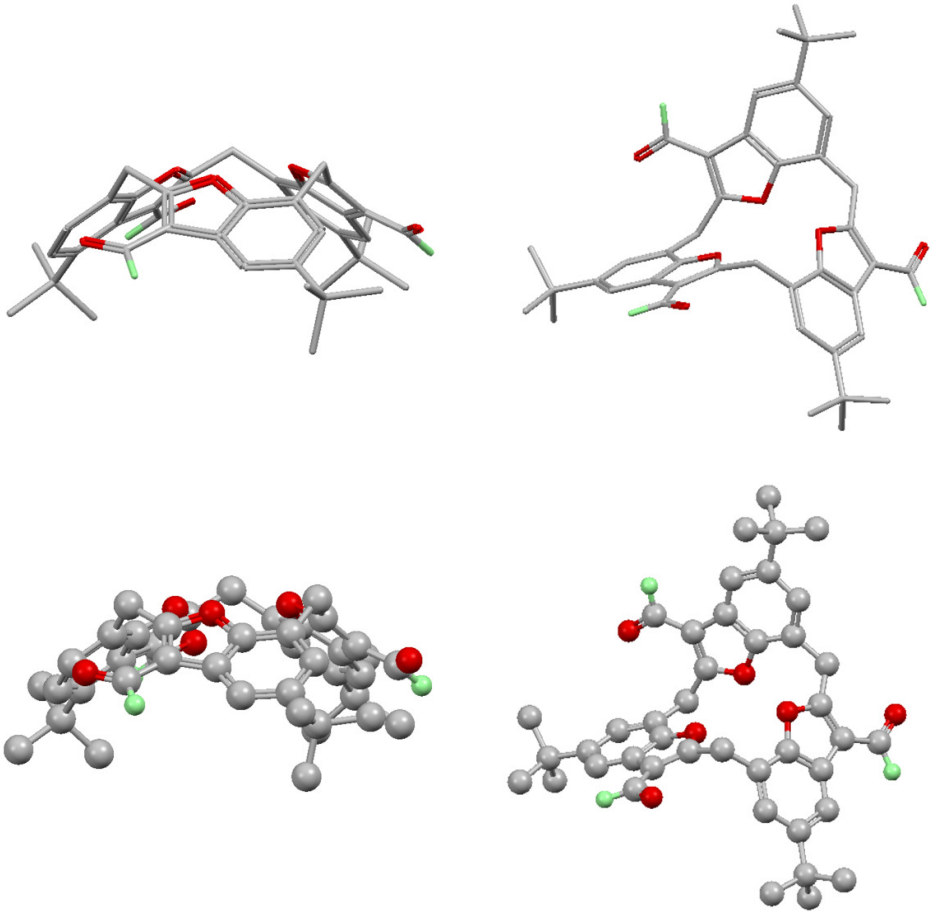


Figure S43. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **4c** *cone* (Ellipsoid); *Top Right*: **4c** *saddle* (Ellipsoid). *Bottom Left*:**4c***cone* (ball-and-stick); and *Bottom Right*: **4c** *saddle* (ball-and-stick).Colour code: carbon = dark grey and oxygen atom = red. All hydrogens except aldehyde hydrogen (light green) are omitted for clarity.

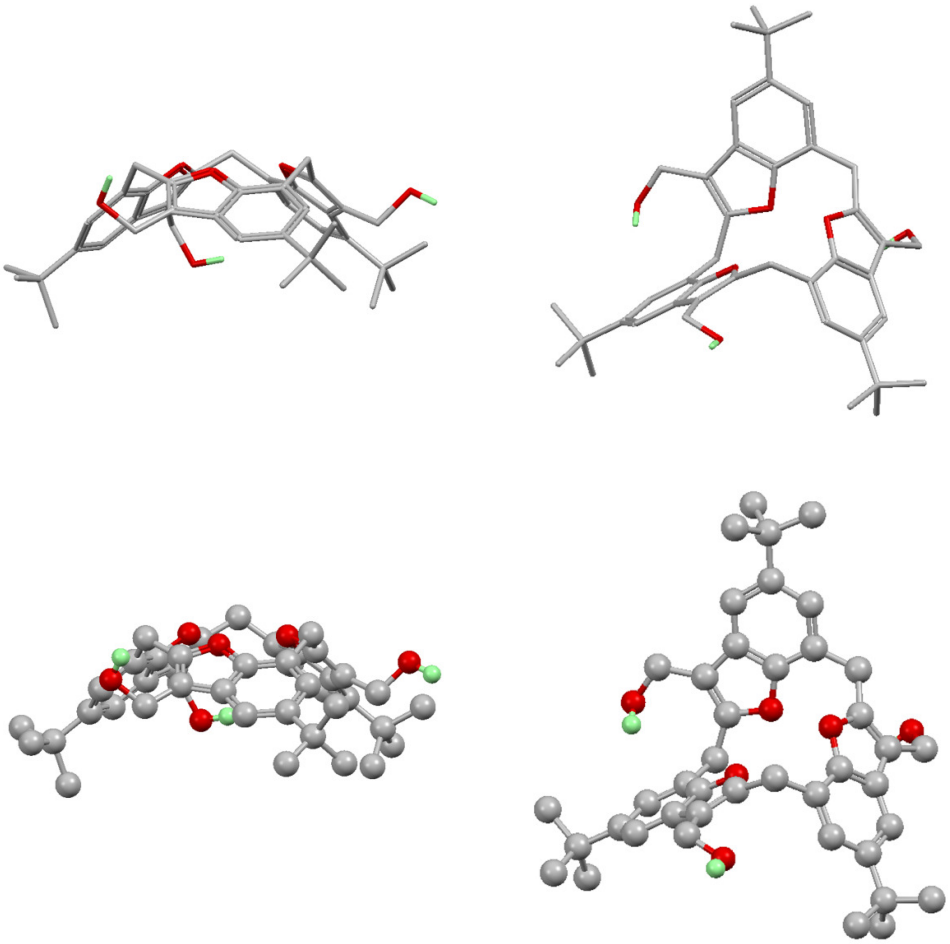


Figure S44. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **4d** *cone* (Ellipsoid); *Top Right*: **4d** *saddle* (Ellipsoid). *Bottom Left*:**4d** *cone* (ball-and-stick); and *Bottom Right*: **5c** *saddle* (ball-and-stick).Colour code: carbon = dark grey and oxygen atom = red. All hydrogens except hydroxyl hydrogen (light green) are omitted for clarity.

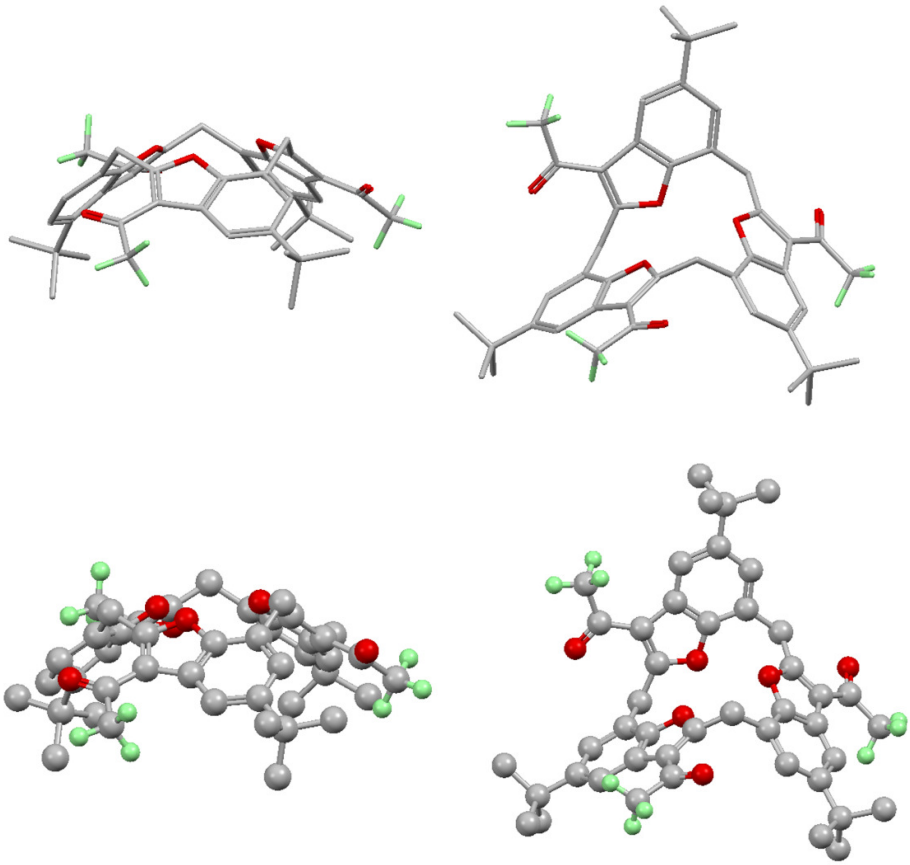


Figure S45. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **4e** *cone* (Ellipsoid); *Top Right*: **4e** *saddle* (Ellipsoid). *Bottom Left*: **4e** *cone* (ball-and-stick); and *Bottom Right*: **4e** *saddle* (ball-and-stick).Colour code: carbon = dark grey and oxygen atom = red. All hydrogens except carbonyl hydrogen (light green) are omitted for clarity.

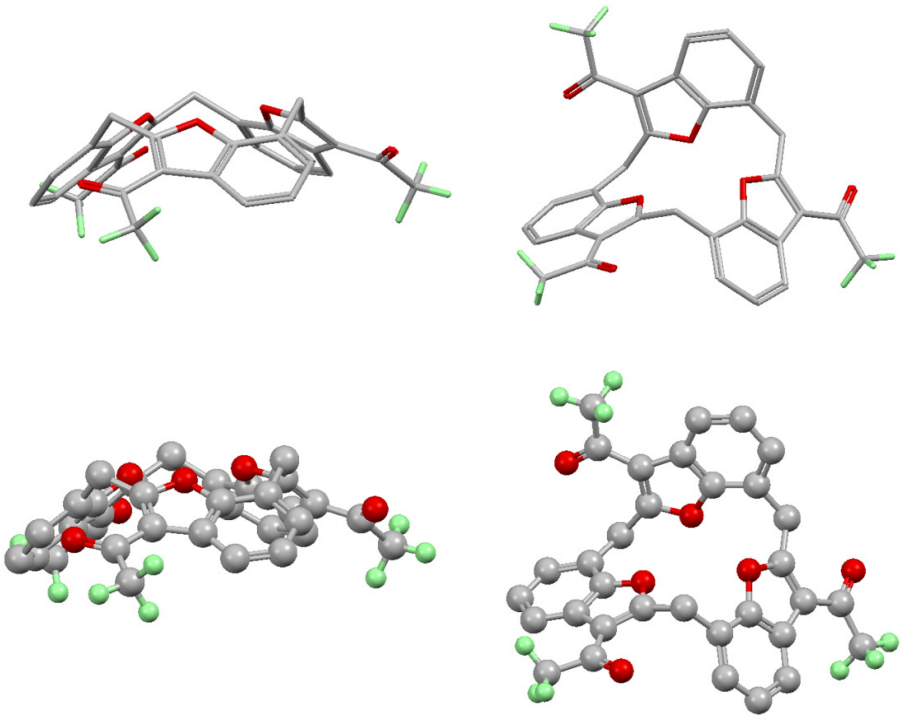


Figure S46. Geometry-optimized (in $CHCl_3$) structures of: *Top Left*: **5b***cone* (Ellipsoid); *Top Right*: **5b** *saddle* (Ellipsoid). *Bottom Left*: **5b** *cone* (ball-and-stick); and *Bottom Right*: **5b** *saddle* (ball-and-stick).Colour code: carbon = dark grey and oxygen atom = red. All hydrogens except aldehyde hydrogen (light green) are omitted for clarity.

References

1. SIR2008: Burla, M. C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. Siliqi, D. and Spagna R. (2007). *J. Appl. Cryst.***40**, 609–613.
2. *'International Tables for Crystallography'*, Vol. C (1992). Ed. A.J.C. Wilson, Kluwer Academic Publishers, Dordrecht, Netherlands, Table 6.1.1.4, pp. 572.
3. CrystalStructure 4.1: Crystal Structure Analysis Package, Rigaku Corporation (2000–2014). Tokyo 196–8666, Japan.
4. SHELXL2013: Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
5. Programs CrysAlis-CCD and -RED, Oxford Diffraction Ltd., Abingdon, UK (2005).
6. G. M. Sheldrick, SHELX-Programs for crystal structure determination (SHELXS-2013) and refinement (SHELXL-2014), *ActaCryst.* (2008) A**64**, 112–122 and (2015) C**71**, 3–8.
7. *'International Tables for X-ray Crystallography'*, Kluwer Academic Publishers, Dordrecht (1992). Vol. C, pp. 500, 219 and 193.
8. L. J. Farrugia, (2012) *J. Appl. Cryst.***45**, 849–854.
9. Initial molecular modeling calculations using the MMFF94 were performed using the PC Spartan'10 software from Wavefunction Inc., Irvine CA.
10. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.;

Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.: Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision D.01*; Gaussian, Inc., Wallingford CT, 2013.