

Supporting Information File

The effect of the intramolecular C-H...O interactions on the conformational preferences of bis-arylsulfones - 5-HT₆ receptor antagonists and beyond

Justyna Kalinowska-Tłuścik,^{*a} Jakub Staroń,^b Anna Krawczuk,^a Stefan Mordalski,^b Dawid Warszycki,^b Grzegorz Satała,^b Adam S. Hogendorf,^{a,b} and Andrzej J. Bojarski^b

[†] Department of Crystal Chemistry and Crystal Physic, Faculty of Chemistry, Jagiellonian University, Gronostajowa 2, 30-387 Kraków, Poland;

[‡] Department of Medicinal Chemistry, Institute of Pharmacology Polish Academy of Sciences, Smętna 12, 31-343 Kraków, Poland

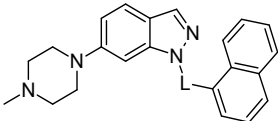
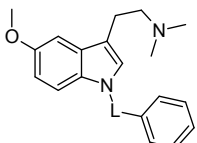
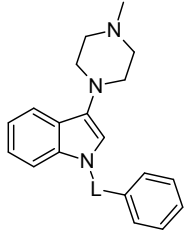
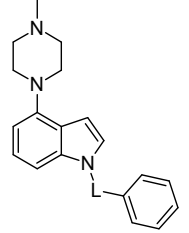
Contents

S1 Binding affinity determination.....	1
S1.1 In vitro pharmacology	3
S1.2 Radioligand binding assays.....	4
S1.3 Example of ligand's geometry in the binding site of 5HT ₆ R model and 2D binding mode representations.....	5
S2. X-ray structure analysis	7
S2.1 General information and additional applied refinement procedures	7
S2.2 Crystal structure description of compounds 1a, 1b and 1c.....	9
S2.3 Crystal structure description of compounds 2a, 2b and 2c.....	18
S2.4 Crystal structure description of compounds 3a and 3c.....	26
S2.5 Crystal structure description of compounds 4a, 4b and 4c.....	35
S2.6 Comparison of selected geometric parameters for the presented structures that can be crucial for binding to 5HT ₆ R	47
S3. Mass spectroscopy data for the dissolved crystal 3a - confirmation of bromination at the C3 position of the indole ring.....	49
S4. Cambridge Crystallographic Database statistical analysis	51
S.4.1 Distribution of the Cg1...Cg2 distances and PLN1-PLN2 angles for N-arylsulfonyl and N-benzyl deposited within the CSD.....	52
S.4.2 Distribution of the torsion angle corresponding to C12-C11-S10(or C10)-N1 angles for benzenesulfonyl benzene and benzylbenzene derivatives deposited within the CSD.....	55
S5. Theoretical calculations	57
S6. Protein Data Bank analysis	67
S7. Abbreviations.....	70
S8. References	71

S1 Binding affinity determination

The bioisosteric groups were tested to check the binding affinities towards 5-HT_{1A}, 5-HT₆, 5-HT₇ and D₂ receptors. The results (**Table S1**) revealed the highest binding affinity among all sulfonyl derivatives to 5-HT₆ receptor, indicating a selective binding profile. Substitution of the sulfonyl group with a carbonyl linker resulted in a great loss of affinity towards 5-HT₆R (approx. 50-2000-fold increase in K_i values in comparison to the sulfonyl analogues; an exception is compound **2b** with a 4-fold increase). By contrast, introduction of the methylene moiety in place of -SO₂- caused only a moderate drop in affinity, suggesting that the tetrahedral geometry of the linker may be an important feature of 5-HT₆ receptor ligands. The rotational freedom of the methylene linker allows the adaptation of the similar conformations observed for sulfonyl derivatives. However, the decrease in affinity suggests additional structural effects, making compounds **1a**, **2a**, **3a** and **4a** the best ligands of the 5-HT₆ receptor.

Table S1. Molecular structures of the investigated compounds and binding affinities towards 5-HT_{1a}R, 5-HT₆R, 5-HT₇R and D₂R

General structure	No.	L	5-HT ₆ R K _i [nM]	5-HT _{1a} R K _i [nM]	5-HT ₇ R K _i [nM]	D ₂ R K _i [nM]
	1a	-SO ₂ -	1 ± 0.2	n.d.	1 780 ± 93	973 ± 81
	1b	-CO-	1280 ± 231	n.d.	10 210 ± 1418	3 117 ± 465
	1c	-CH ₂ -	51 ± 7	1618 ± 246	3121 ± 422	1251 ± 214
	2a	-SO ₂ -	11 ± 2	2 921 ± 411	13 910 ± 2164	352 ± 47
	2b	-CO-	44 ± 8	686 ± 52	4 391 ± 643	2 591 ± 321
	2c	-CH ₂ -	23 ± 5	379 ± 17	937 ± 72	327 ± 58
	3a	-SO ₂ -	4 ± 1	1 464 ± 183	38 590 ± 5486	532 ± 63
	3b	-CO-	187 ± 22	3 990 ± 571	15 490 ± 1732	162 ± 9
	3c	-CH ₂ -	18 ± 3	8 422 ± 724	8 612 ± 756	35 ± 4
	4a	-SO ₂ -	1 ± 0.3	n.d.	5 462 ± 824	31 ± 2
	4b	-CO-	2 204 ± 189	n.d.	5 267 ± 655	88 ± 11
	4c	-CH ₂ -	24 ± 5	n.d.	5 531 ± 452	305 ± 26

S1.1 In vitro pharmacology

Cell culture and preparation of cell membranes

HEK293 cells with stable expression of human serotonin 5-HT_{1A}R, 5-HT₆R, 5-HT_{7b}R or dopamine D_{2L}R (prepared using Lipofectamine 2000) were maintained at 37°C in a humidified atmosphere with 5% CO₂ and grown in Dulbecco's Modified Eagle's Medium containing 10% dialyzed fetal bovine serum and 500 µg/mL G418 sulfate. For the membrane preparations, cells were subcultured in 10-cm-diameter dishes, grown to 90% confluence, washed twice with phosphate-buffered saline (PBS) prewarmed to 37°C and pelleted by centrifugation (200 g) in PBS containing 0.1 mM EDTA and 1 mM dithiothreitol. Prior to membrane preparations, the pellets were stored at -80°C.

S1.2 Radioligand binding assays

Cell pellets were thawed and homogenized in 20 volumes of assay buffer using an Ultra Turrax tissue homogenizer and centrifuged twice at 35 000 g for 20 min at 4°C, with 15-min incubations at 37°C in between rounds of centrifugation. The composition of the assay buffers was as follows: for 5-HT_{1A}R: 50 mM Tris-HCl, 0.1 mM EDTA, 4 mM MgCl₂, 10 µM pargyline and, 0.1% ascorbate; for 5-HT₆R: 50 mM Tris-HCl, 0.5 mM EDTA and 4 mM MgCl₂; for 5-HT_{7b}R: 50 mM Tris-HCl, 4 mM MgCl₂, 10 µM pargyline and 0.1% ascorbate; for dopamine D_{2L}R: 50 mM Tris-HCl, 1 mM EDTA, 4 mM MgCl₂, 120 mM NaCl, 5 mM KCl, 1.5 mM CaCl₂ and 0.1% ascorbate.

All assays were incubated in a total volume of 200 µL in 96-well microtiter plates for 1 h at 37°C, excluding 5-HT_{1A}R, which was incubated at room temperature for 1 h. The equilibration process was terminated by rapid filtration through Unifilter plates with a 96-well cell harvester, and the radioactivity retained on the filters was quantified on a Microbeta plate reader.

For displacement studies, the assay samples contained the following radioligands: 1.5 nM [3H]-8-OH-DPAT (187 C_i/mmol) for 5-HT_{1A}R; 2 nM [3H]-LSD (85.2 C_i/mmol) for 5-HT₆R; 0.6 nM [3H]-5-CT (39.2 C_i/mmol) for 5-HT_{7b}R; and [3H]-Raclopride (74.4 C_i/mmol) for D_{2L}R.

Non-specific binding was defined with 10 µM of 5-HT in the 5-HT_{1A}R and 5-HT_{7b}R binding experiments, whereas 10 µM methiothepine or 1 µM of (+)butaclamol was used in the 5-HT₆R and D_{2L}R assays, respectively. Each compound was tested in triplicate at 7-8 concentrations (10⁻¹¹-10⁻⁴ M). The inhibition constants (K_i) were calculated using the Cheng-Prusoff equation.¹ Results are expressed as the means of at least two separate experiments.

Initial screening experiments were performed under the same conditions at two compound concentrations: 10⁻⁶ and 10⁻⁷ M.

S1.3 Example of ligand's geometry in the binding site of 5HT₆R model and 2D binding mode representations

Figures for this section were generated using Maestro (Schrödinger Release 2018-1: Maestro, Schrödinger, LLC, New York, NY, 2018.). Presented results are yet unpublished data.

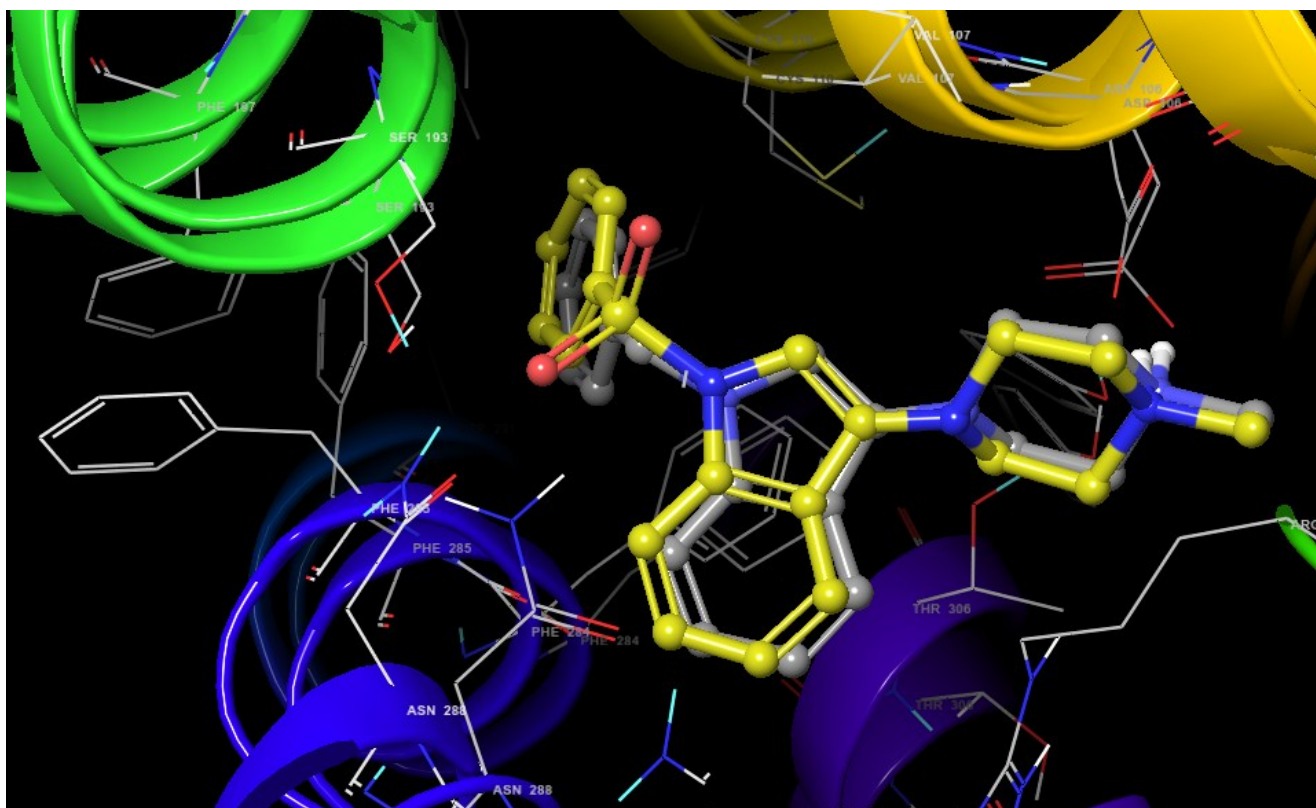


Figure 1.3.1 Predicted poses of compounds 3a and 3c (as an example) in the binding site of the 5HT₆R model (unpublished data). Both compounds exhibit a binding mode that corresponds with binding modes published in the literature, i.e. protonated nitrogen atom forms a charge-assisted hydrogen bond with the aspartic acid D3.32, the sulfonyl group lies in the proximity of asparagine N6.55 and serine S5.43, whereas phenyl ring forms interaction with the phenylalanine F6.51 and F6.52 cluster.

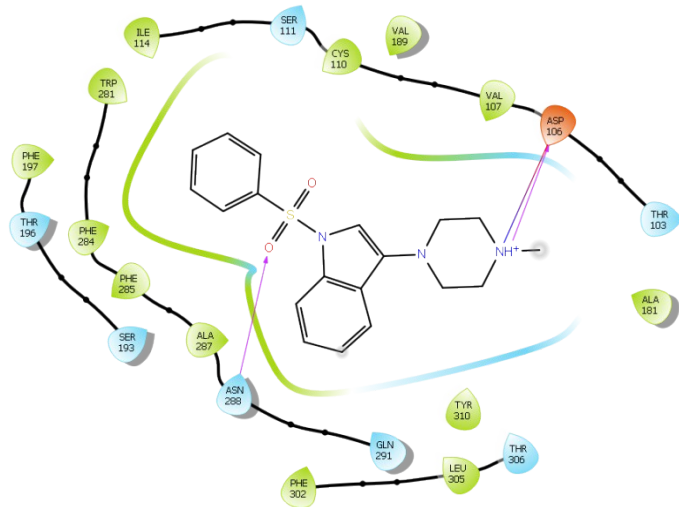


Figure 1.3.2 Predicted 2D binding mode for compound 3a in the 5HT6R model

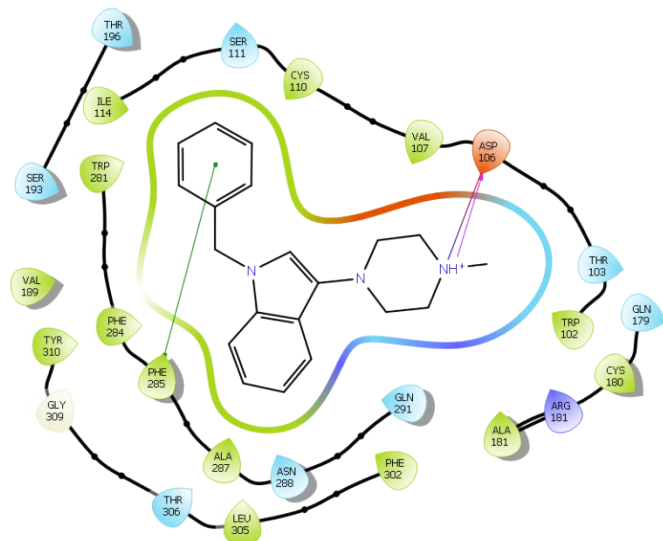


Figure 1.3.3 Predicted 2D binding mode for compound 3a in the 5HT6R model

S2. X-ray structure analysis

S2.1 General information and additional applied refinement procedures

Compounds **1a**², **2a**³, **3a**⁴ and **4a**⁵ are confirmed arylsulfonyl ligands of 5-HT₆R, antagonistic activity that has been previously reported. Bioisosteric analogues of these compounds were synthesized with carbonyl (**1b-4b**) and methylene (**1c-4c**) linker between two aromatic systems. For 11 out of 12 compounds, crystals allowing X-ray experiments were obtained. Compound **3b** did not provide sufficient single crystals, even after several re-crystallization trials.

The X-ray diffraction data were collected with SuperNova (Rigaku - Oxford Diffraction) equipment. The phase problem was solved by direct methods using SHELXS-2013/1.⁶ The model parameters were refined by full-matrix least-squares on F^2 using SHELXL-2014/6.⁶ For room and low-temperature experiments, the positions of hydrogen atoms attached to carbon atoms were calculated in the model and set at C-H = 0.93 Å and 0.95 Å (aromatic), and C-H = 0.97 Å and 0.99 Å (methylene), C-H = 0.96 Å and 0.98 Å (methyl), respectively. All mentioned hydrogen atoms were refined using the riding model with the isotropic displacement parameter $U_{iso} = 1.2 U_{eq}$ and $U_{iso} = 1.5 U_{eq}$ (methyl groups only) of the parent atom. H atoms attached to nitrogen and oxygen atoms were localized on the difference Fourier map.

Due to the observed disorder of the solvent molecules in structure **1a**, the positions of missing hydrogen atoms were determined with CALC-OH⁷ by combined geometric and force-field calculations based on hydrogen bonding interactions. For structures **1a**, **3c** and **4b**, DFIX and/or DANG restraints were applied to the solvent bond lengths and angles to improve its geometry.

The asymmetric unit of **4a** consists of two molecules, of which one exhibits a positional disorder within the sulfonyl group. The final refined site occupancy for the alternative positions is approximately 0.6 : 0.4. Additionally, for the mentioned disordered molecule, a positive electron density near C3 of the indole ring was observed on the difference Fourier map. Further investigation confirmed, that a

partial/unexpected bromination occurred during the crystallization process (the crystallization process was performed in a working area where bromination is frequently performed). It is well established in the literature that the most reactive position of indole for electrophilic substitution is position C3⁸⁻¹⁰, thus a substitution occurred in the presence of even a small amount of brominating agent. The partial bromination was confirmed by mass spectrometry - see section **S3** of the Supplementary Materials).

Compound **1a** crystallized as a free base. The remaining compounds crystallized in the salt form as hydrochlorides (**1b**, **2a**, **2b**) or oxalates (**1c**, **2c**, **3a**, **3c**, **4a-4c**), with a positive charge located on the most basic nitrogen atom of the piperazine or non-rigid tertiary amine group (**2a-2c**). Despite very similar molecular structures, all compounds exhibit different crystal architectures due to the presence of solvent molecules and different types of counter ions, which serve as donors/acceptors of strong hydrogen bonds.

S2.2 Crystal structure description of compounds **1a**, **1b** and **1c**

Crystal structure data and refinement parameters for **1a**, **1b** and **1c** are shown in **Table S2.2**. Asymmetric units of crystal structures **1a**, **1b** and **1c** consist of one molecule of the indazole derivative (**Figure 2.2.1**).

Compound **1a** crystallizes as a free base, assisted by water molecules. The solvent molecules exhibit strong disorder and occupy channels in the [010] direction, forming a chain *via* O-H...O hydrogen bonds (**Figure 2.2.2**). The water molecules interact with the indazole derivative as hydrogen bond donors, with acceptors localized on the most basic nitrogen atoms N74 and N2.

Compound **1b** crystallizes in a hydrochloride salt form, with two co-crystallizing water molecules. This compound exhibits the most distinguished molecular conformation, with indazole placed in the less favorable axial position of the piperazine ring. Among all the structures presented herein and in previously studied molecular geometries of different arylpiperazine derivatives ^{e.g.11}, the most common position of the aromatic moiety with respect to arylpiperazine has been shown to be equatorial.

The negative charge of the chloride ion in **1b** is balanced by the protonated nitrogen atom N74 in the piperazine ring. Surprisingly, the ions do not interact directly, but they interact indirectly by water molecule serving as donor and acceptor of hydrogen bonds N74⁺-H74...O1w and O1w-H...Cl⁻, respectively. In the case of arylpiperazine salts, a charge-assisted hydrogen bond is usually expected to be observed in the crystal structure. The solvent molecules in the crystal lattice are additionally engaged in the strong hydrogen bonds O-H...O and with the two above-mentioned interactions, creating a ribbon-like motif in the [100] direction (**Figure 2.2.3**).

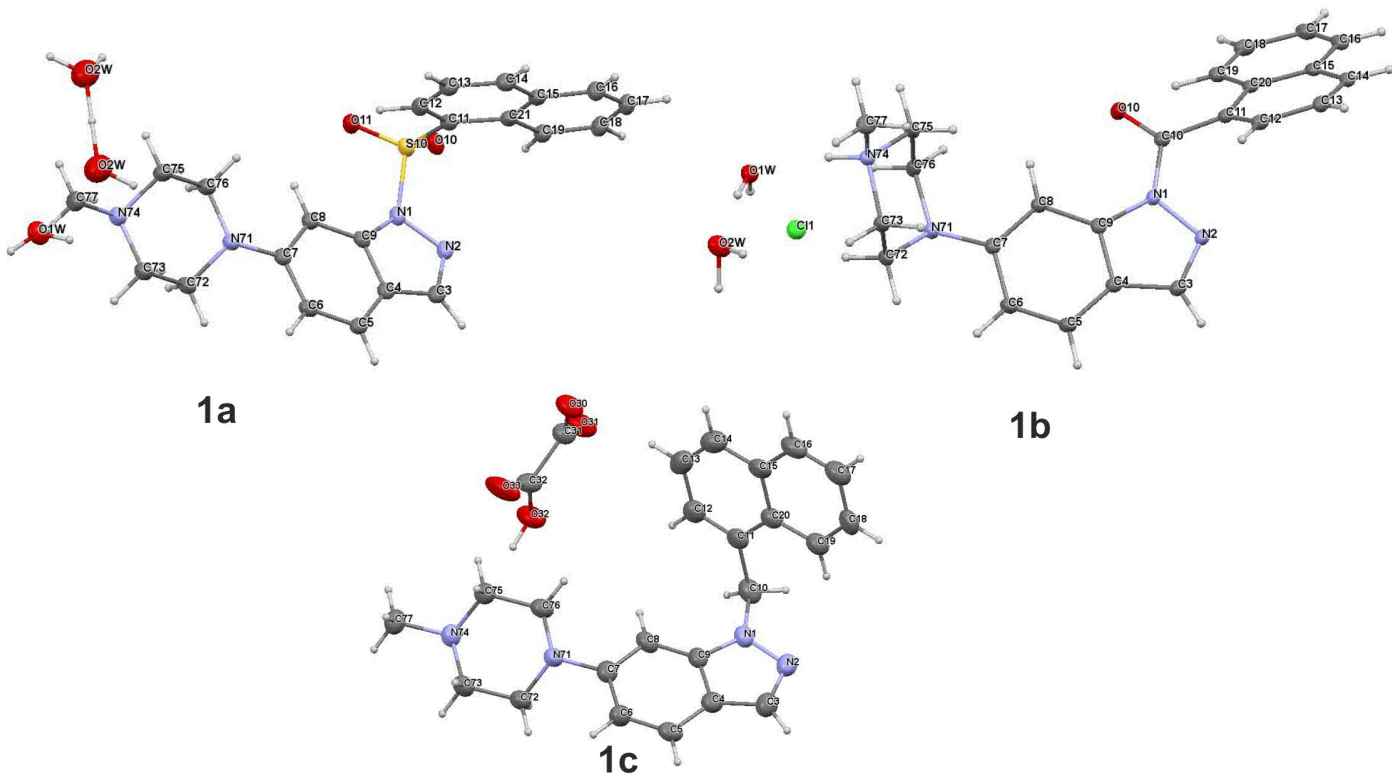


Figure 2.2.1 Asymmetric units in crystal structures of compounds 1a, 1b and 1c. Displacement ellipsoids of non-hydrogen atoms are drawn at the 30% probability level. H atoms are presented as small spheres with an arbitrary radius.

Table S2.2 Crystal data and refinement parameters for compounds **1a-1c**

Identification code	1a	1b	1c
Chemical formula	C ₂₂ H ₂₂ N ₄ O ₂ S, 2(H ₂ O)	C ₂₃ H ₂₃ N ₄ O ⁺ , Cl ⁻ , 2(H ₂ O)	C ₂₃ H ₂₅ N ₄ ⁺ , C ₂ HO ₄ ⁻
Formula mass	442.53	442.93	446.50
Crystal data			
Crystal system	Triclinic	Orthorombic	Monoclinic
Space group	$P\bar{1}$	<i>Pbca</i>	<i>Ia</i>
Unit cell dimensions	a = 9.1583(6) Å b = 10.1568(8) Å c = 12.0609(9) Å α=101.718(6)° β=100.729(6)° γ=100.020(6)°	a = 7.5146(2) Å b = 18.6958(3) Å c = 30.8781(5) Å α = 90° β = 90° γ = 90°	a = 8.5126(19) Å b = 22.934(16) Å c = 11.454(4) Å α = 90° β = 100.85(3)° γ = 90°
Unit cell volume [Å ³]	1052.83(14)	4338.11(15)	2196.2(18)
Z	2	8	4
D _{calc} [g/cm ³]	1.396	1.356	1.350
Absorption coefficient [mm ⁻¹]	1.686	1.831	0.093
F(000)	468	1872	944
Crystal size [mm ³]	0.2 x 0.2 x 0.1	0.4 x 0.2 x 0.05	0.2 x 0.15 x 0.05
Data collection			
Temperature [K]	120(2)	121(2)	130(1)
Radiation type	CuKα	CuKα	MoKα
θ range [°]	3.846° to 71.213°	4.731 to 72.087°	3.222° to 28.682°
Index ranges	-11 ≤ h ≤ 11, -12 ≤ k ≤ 12, -14 ≤ l ≤ 14	-8 ≤ h ≤ 9, -22 ≤ k ≤ 22, -37 ≤ l ≤ 37	-11 ≤ h ≤ 11, -30 ≤ k ≤ 30, -15 ≤ l ≤ 15
Reflections collected	14911	63365	13886
Independent reflections	4025 [R(int) = 0.0611]	4202 [R(int) = 0.0573]	5103 [R(int) = 0.1750]
Completeness [%]	99.9 (θ = 67.7°)	100 (θ = 67.7°)	99.8 (θ = 25.2°)
Refinement			
Data/restraints/parameters	4025 / 11 / 286	4202 / 0 / 301	5103 / 2 / 307
Goodness-of-fit on F ²	1.049	1.139	0.935
Final R indices [I>2σ(I)]	R1=0.0567, wR2= 0.1509	R1= 0.0504, wR2= 0.1239	R1=0.0789, wR2= 0.1682
R indices (all data)	R1=0.0803, wR2=0.1688	R1=0.0563, wR2=0.1284	R1=0.2077, wR2=0.2356
Δρ _{max} /Δρ _{min} [e.Å ⁻³]	0.68 and -0.57	0.42 and -0.32	0.27 and -0.22

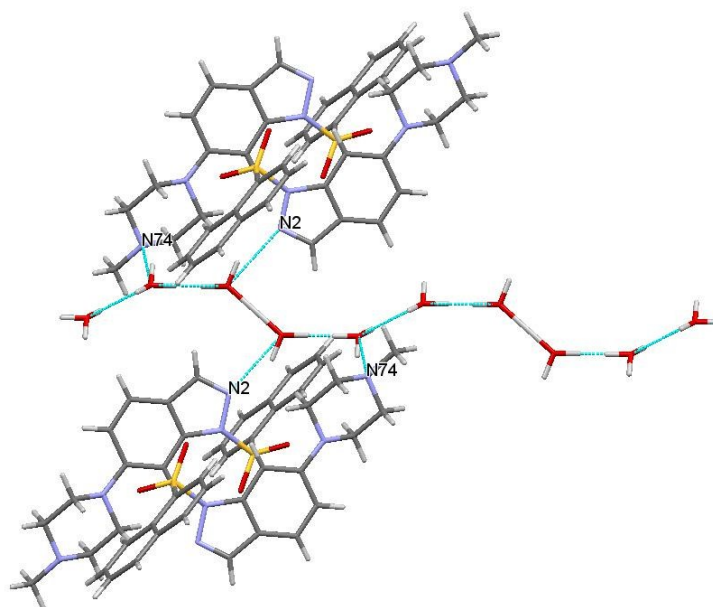


Figure 2.2.2 Water chains and hydrogen bond systems in the crystal structure of compound 1a. View along [100]. Hydrogen bonds are shown as a cyan dashed lines.

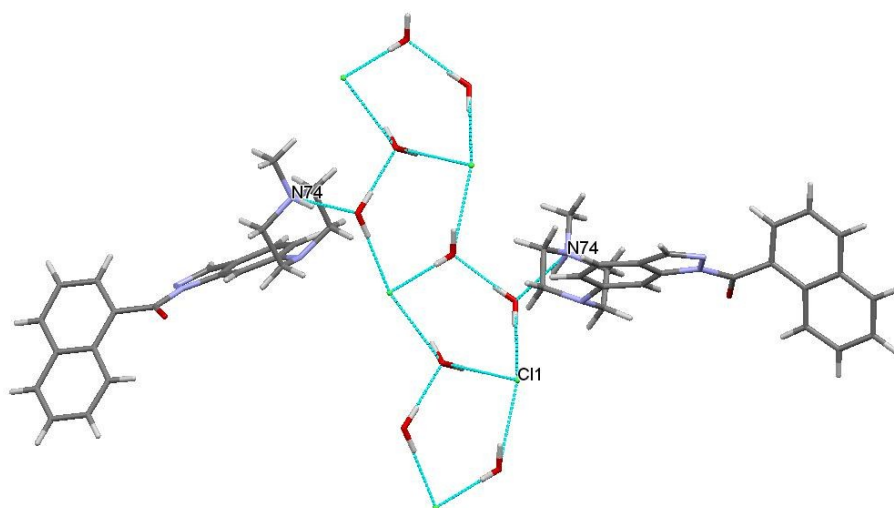


Figure 2.2.3 The ribbon-like motifs of the hydrogen bond systems in the crystal structure of compound 1b. The ribbon is elongated in the [100] direction. View along [001]. Hydrogen bonds are shown as cyan dashed lines.

Crystallization trials of compound **1c** resulted in low-quality crystals. However, due to a structural comparison in the structure-activity relationship study, we decided to include the data in this paper. This compound crystallizes as oxalate, with a positive charge localized on the protonated N74 atom of the

piperazine moiety. The oxalate anion forms a chain *via* the O32-H32...O30 hydrogen bond, with elongation in the [001] direction. The cation is a donor of the bifurcated hydrogen bond, with O30 and O33 as acceptors of the interaction (**Figure 2.2.4**).

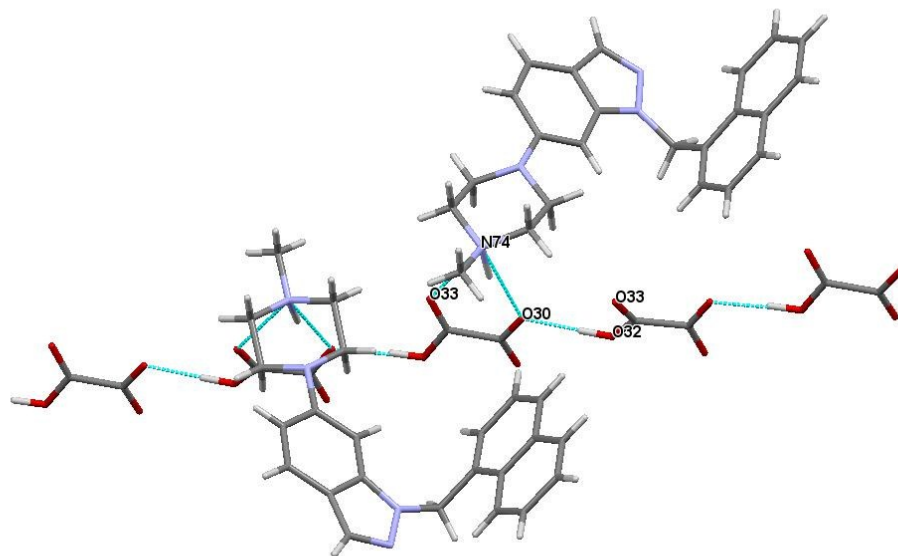


Figure 2.2.4 The hydrogen bond system in the crystal structure of compound **1c**. The oxalate chain is elongated in the [001] direction. Hydrogen bonds are shown as cyan dashed lines.

The change in linker type strongly influences the molecular geometry. Superposition with respect to the indazole ring, performed for a single molecule of **1a**, **1b** and **1c** extracted from the crystal structure, is shown in **Figure 2.2.5**.

The crystal packing schemes of **1a-1c** are shown in **Figures 2.2.6-8**.

Additionally, the crystal structures of compounds **1a-1c** are stabilized by several weak hydrogen bonds C-H...A (where **A** is N, O and/or Cl). The geometric parameters of strong and weak hydrogen bonds are shown in **Tables S2.2.1-3**.

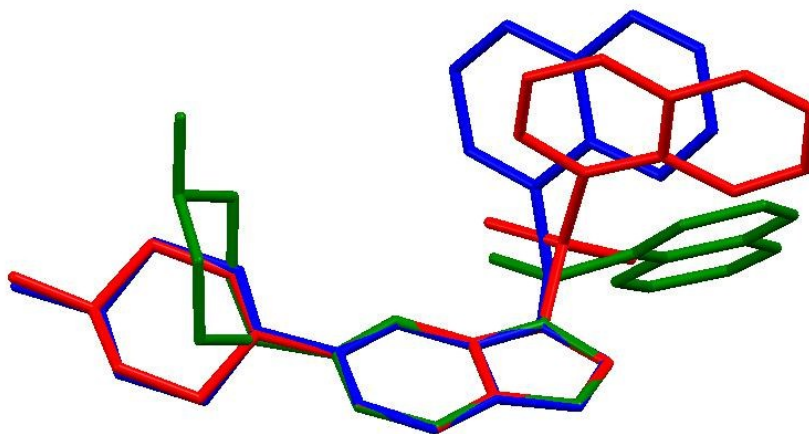


Figure 2.2.5 Superposition of compounds 1a (red), 1b (green) and 1c (blue) with respect to the indazole ring, showing the different molecular conformations observed in the crystal structure as a consequence of the different linkers between the naphthyl and indazole moieties.

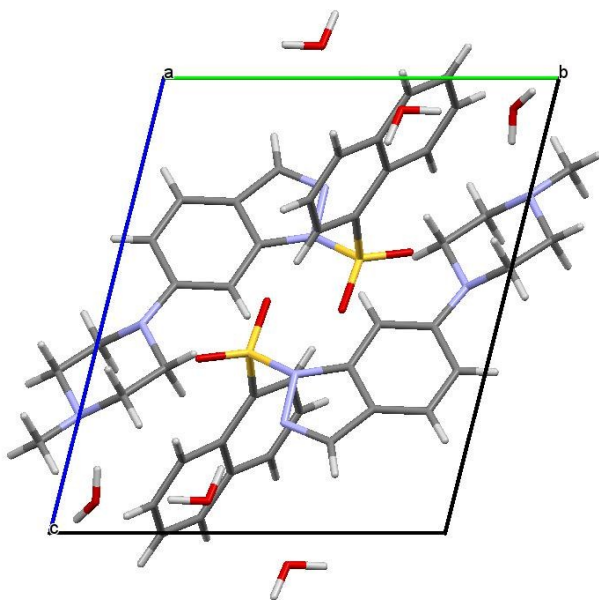


Figure 2.2.6 Crystal packing of 1a along [100]

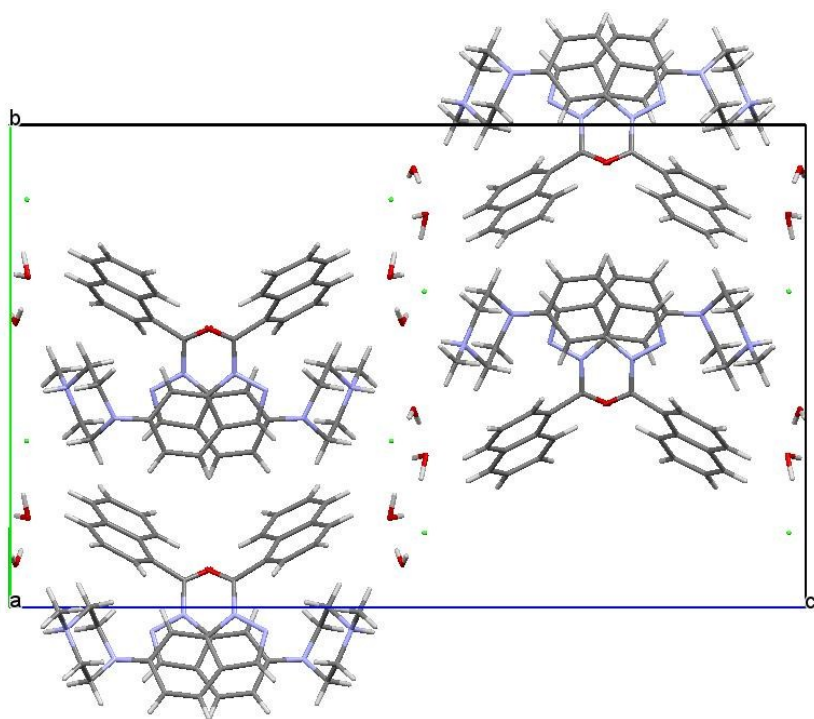


Figure 2.2.7 Crystal packing of 1b along [100]

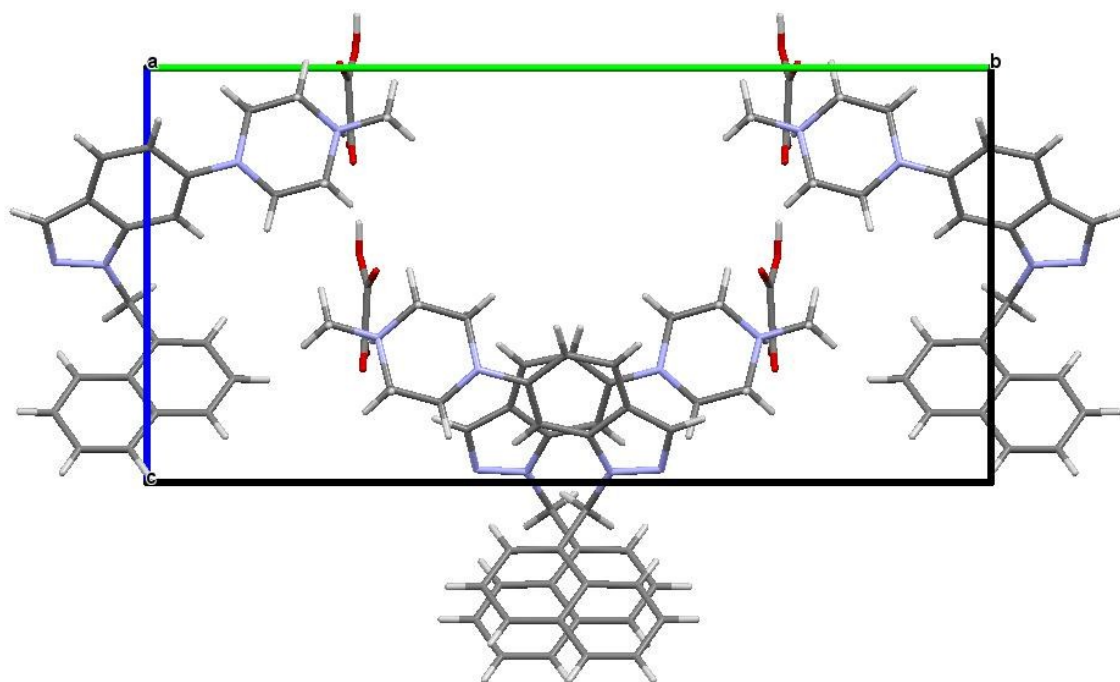


Figure 2.2.6 Crystal packing of 1c along [100]

Table S2.2.1 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **1a** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(8)-H(8)...O(11)	0.95	2.44	3.004(3)	117.9
C(72)-H(72A)...O(10)#1	0.99	2.55	3.490(4)	159.5
C(72)-H(72B)...O(10)#2	0.99	2.63	3.332(4)	128.2
O(2W)-H(3W)...O(1W)	0.93(2)	1.93(4)	2.835(4)	164(12)
O(2W)-H(4W)...O(2W)#3	1.345(4)	1.345(4)	2.691(7)	180.0(3)
O(2W)-H(5W)...N(2)#2	0.936(19)	2.06(3)	2.910(4)	150(5)
O(1W)-H(1W)...N(74)	0.829(2)	2.244(3)	2.915(3)	138.27(18)
O(1W)-H(6W)...O(1W)#4	0.796(2)	2.186(2)	2.905(5)	150.53(13)
O(1W)-H(2W)...O(2W)	0.914(2)	1.956(4)	2.835(4)	160.95(18)

Symmetry transformations used to generate equivalent atoms:

#1 x,y+1,z #2 -x,-y+1,-z+1 #3 -x,-y+1,-z #4 -x,-y+2,-z

Table S2.2.2 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **1b** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(8)-H(8)...O(10)	0.95	2.48	2.974(3)	112.6
C(72)-H(72B)...Cl(1)#1	0.99	2.78	3.713(3)	156.5
C(73)-H(73B)...N(2)#2	0.99	2.66	3.595(3)	157.2
C(75)-H(75B)...O(1W)#3	0.99	2.56	3.311(3)	132.5
C(77)-H(77B)...Cl(1)#4	0.98	2.88	3.779(3)	153.1
N(74)-H(74)...O(1W)	0.95(3)	1.84(3)	2.762(3)	164(3)
O(1W)-H(2W)...Cl(1)#1	0.87(4)	2.30(4)	3.158(2)	173(3)
O(2W)-H(4W)...Cl(1)#4	0.89(5)	2.32(5)	3.154(2)	155(4)
O(1W)-H(1W)...O(2W)	0.82(4)	1.89(4)	2.705(3)	174(4)
O(2W)-H(3W)...Cl(1)#5	0.96(5)	2.21(5)	3.155(2)	165(4)
C(8)-H(8)...O(10)	0.95	2.48	2.974(3)	112.6
C(72)-H(72B)...Cl(1)#1	0.99	2.78	3.713(3)	156.5
C(73)-H(73B)...N(2)#2	0.99	2.66	3.595(3)	157.2
C(75)-H(75B)...O(1W)#3	0.99	2.56	3.311(3)	132.5
C(77)-H(77B)...Cl(1)#4	0.98	2.88	3.779(3)	153.1
N(74)-H(74)...O(1W)	0.95(3)	1.84(3)	2.762(3)	164(3)

O(1W)-H(2W)...Cl(1)#1	0.87(4)	2.30(4)	3.158(2)	173(3)
O(2W)-H(4W)...Cl(1)#4	0.89(5)	2.32(5)	3.154(2)	155(4)
O(1W)-H(1W)...O(2W)	0.82(4)	1.89(4)	2.705(3)	174(4)
O(2W)-H(3W)...Cl(1)#5	0.96(5)	2.21(5)	3.155(2)	165(4)

Symmetry transformations used to generate equivalent atoms:

#1 $-x+1/2, -y, z-1/2$ #2 $x-1/2, y, -z+1/2$ #3 $-x, -y, -z$ #4 $-x-1/2, -y, z-1/2$ #5 $-x, y-1/2, -z+1/2$

Table S2.2.3 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **1c** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(75)-H(75A)...O(31)#1	0.97	2.55	3.406(11)	146.6
C(75)-H(75B)...O(33)	0.97	2.33	3.225(12)	152.9
C(77)-H(77C)...O(31)#1	0.96	2.46	3.337(11)	151.3
C(3)-H(3)...O(31)#2	0.93	2.59	3.434(13)	151.5
N(74)-H(74)...O(30)#3	1.14(8)	1.79(8)	2.888(10)	159(6)
N(74)-H(74)...O(33)#3	1.14(8)	2.22(7)	2.920(9)	117(5)
O(32)-H(32)...O(30)#3	0.92(14)	1.73(14)	2.643(9)	175(12)

Symmetry transformations used to generate equivalent atoms:

#1 $x-1, -y+1/2, z-1/2$ #2 $x-1/2, y-1/2, z-1/2$ #3 $x+0, -y+1/2, z-1/2$

S2.3 Crystal structure description of compounds **2a**, **2b** and **2c**

Compounds **2a-2c** crystallizes in the salt form as hydrochlorides (**2a** and **2b**) and oxalate (**2c**). Crystal data and refinement parameters are shown in **Table 2.3**. The asymmetric units (presented in **Figure 2.3.1**) consist of a single, positively charged molecule of the compound, accompanied by the anion. For structures **2a-2c**, the positive charge is located on the less rigid, protonated tertiary amine (N33 atom). In the crystal lattice of compounds **2a** and **2b** a charge-assisted hydrogen bond of the $N^+-H\dots Cl^-$ type is observed (**Figures 2.3.2-3**), which is a discrete hydrogen bond motif. In structure **2c** there are two independent oxalic acid molecules, from which one is fully deprotonated, holding the double-negative charge. This anion is an acceptor of the charge-assisted bifurcated hydrogen bond with donor located on the positively charged N33 of two neighboring molecules (**Figure 2.3.4**). The additional oxalic acid molecule is involved in two equal $O-H\dots O$ hydrogen bonds, with acceptors on the deprotonated carboxylic oxygen atoms. This hydrogen bond motif is expanded along the [010] direction and forms the oxalate channels in the crystal structure (**Figure 2.3.4**).

The comparison of different conformations of single molecules of **2a**, **2b** and **2c** extracted from the crystal structure is shown in **Figure 2.3.5** as a superposition with respect to the indole ring.

The crystal packing schemes of structures **2a-2c** are shown in **Figures 2.3.6-8**.

The crystal structures of compounds **2a-2c** are further stabilized by several weak hydrogen bonds $C-H\dots A$ (where **A** are O, S and/or Cl). The geometric parameters of strong and weak hydrogen bonds are shown in **Tables S2.3.1-3**.

Table S2.3 Crystal data and refinement parameters for compounds **2a-2c**

Identification code	2a	2b	2c
Chemical formula	C ₁₉ H ₂₃ N ₂ O ₃ S ⁺ , Cl ⁻	C ₂₀ H ₂₃ N ₂ O ₂ ⁺ , Cl ⁻	C ₂₀ H ₂₅ N ₂ O ⁺ , C ₂ HO ₄ ⁻
Formula mass	394.90	358.85	398.45
Crystal data			
Crystal system	Monoclinic	Triclinic	Triclinic
Space group	P2 ₁ /c	p $\bar{1}$	p $\bar{1}$
Unit cell dimensions	a = 10.7820(2) Å b = 7.2978(1) Å c = 24.7359(4) Å α = 90° β = 91.136(1)° γ = 90°	a = 7.1893(4) Å b = 11.0732(6) Å c = 12.4016(9) Å α = 101.763(5)° β = 96.088(6)° γ = 105.598(5)°	a = 9.1897(4) Å b = 11.1766(6) Å c = 11.9074(7) Å α = 114.544(5)° β = 92.599(4)° γ = 112.821(5)°
Unit cell volume [Å ³]	1945.96(5)	917.3(1)	993.84(10)
Z	4	2	2
D _{calc} [g/cm ³]	1.348	1.299	1.331
Absorption coefficient [mm ⁻¹]	0.325	0.224	0.095
F(000)	832	380	424
Crystal size [mm ³]	0.5 x 0.3 x 0.3	0.5 x 0.5 x 0.1	0.3 x 0.2 x 0.15
Data collection			
Temperature [K]	293(2)	298(2)	293(2)
Radiation type	MoKα	MoKα	MoKα
θ range [°]	2.910° to 28.646°	3.047° to 28.540°	3.419° to 28.708°
Index ranges	-14 ≤ h ≤ 14, -9 ≤ k ≤ 9, -32 ≤ l ≤ 33	-8 ≤ h ≤ 9, -14 ≤ k ≤ 14, -15 ≤ l ≤ 16	-11 ≤ h ≤ 12, -14 ≤ k ≤ 14, -15 ≤ l ≤ 15
Reflections collected	26277	6563	13992
Independent reflections	4748 [R(int)=0.0295]	3963 [R(int)=0.0220]	4623 [R(int)=0.0199]
Completeness [%]	99.9 (θ = 25.2°)	97.7 (θ = 25.2°)	99.7 (θ = 25.2°)
Refinement			
Data/restraints/parameters	4748 / 0 / 239	3963 / 0 / 230	4623 / 0 / 270
Goodness-of-fit on F ²	1.046	1.046	1.067
Final R indices [I > 2σ(I)]	R1=0.0394, wR2= 0.0956	R1=0.0484, wR2= 0.1101	R1=0.0460, wR2= 0.1116
R indices (all data)	R1=0.0539, wR2=0.1054	R1=0.0693, wR2=0.1250	R1=0.0616, wR2=0.1208
Δρ _{max} /Δρ _{min} [e.Å ⁻³]	0.21 and -0.35	0.19 and -0.30	0.27 and -0.24

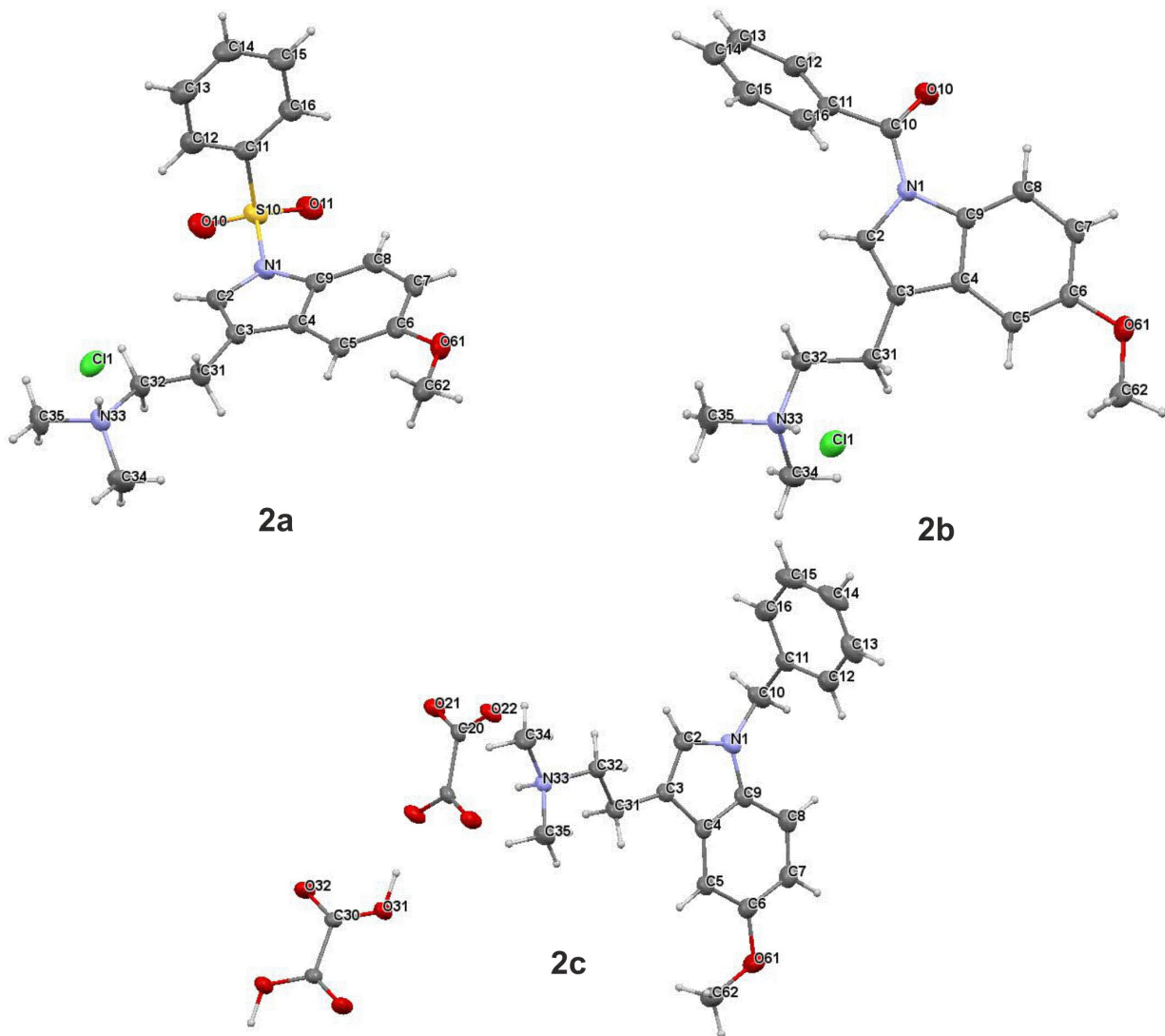


Figure 2.3.1 Asymmetric units in the crystal structures of compounds 2a, 2b and 2c. For structure 2c, the asymmetric unit (the only atoms with labels) consists of one indole derivative, a half oxalic acid molecule and half oxalate anion. Displacement ellipsoids of non-hydrogen atoms are drawn at the 30% probability level. H atoms are presented as small spheres with an arbitrary radius.

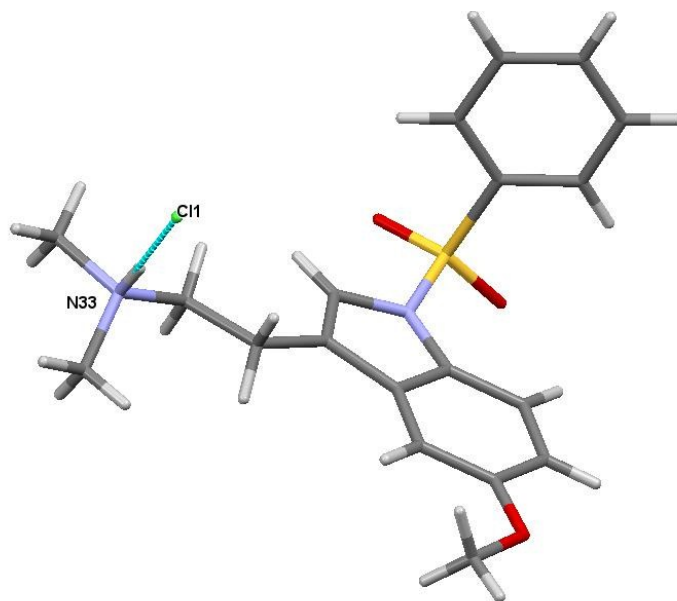


Figure 2.3.2 The discrete motif of the $N^+H \dots Cl^-$ hydrogen bond in the crystal structure of compound 2a. Hydrogen bond is shown as a cyan dashed line.

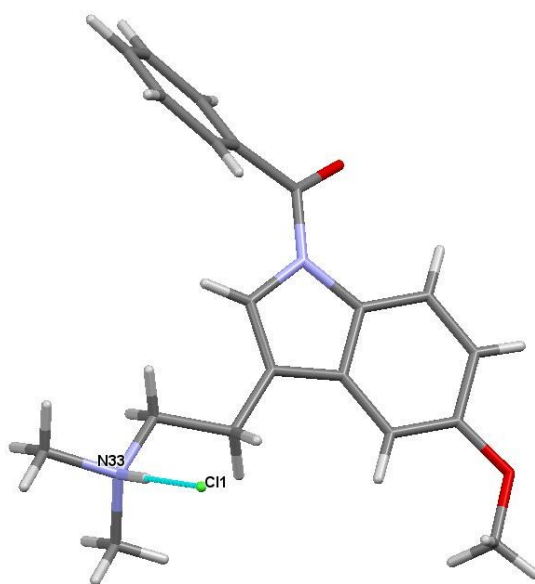


Figure 2.3.2 The discrete motif of the $N^+H \dots Cl^-$ hydrogen bond in the crystal structure of compound 2b. The hydrogen bond is shown as a cyan dashed line.

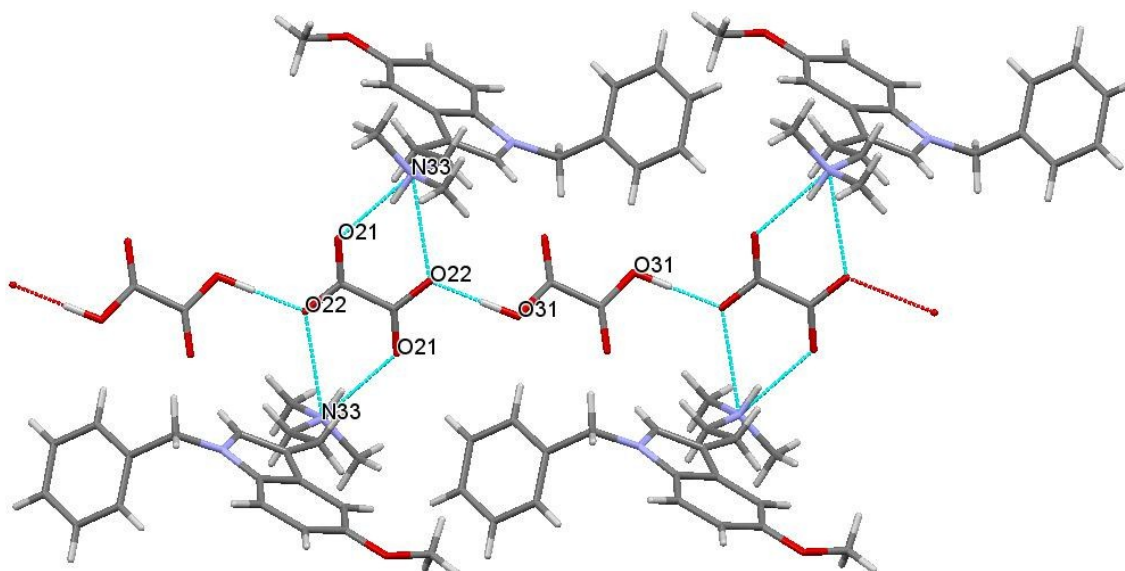


Figure 2.3.4 The bifurcated hydrogen bond formed between protonated nitrogen atoms N33 of the indole derivative 2c and the double-negative oxalate anion. The oxalate anion and oxalic acid molecule form a chain *via* O-H...O hydrogen bonds that is elongated in the [010] direction. Hydrogen bonds are shown as cyan dashed lines.

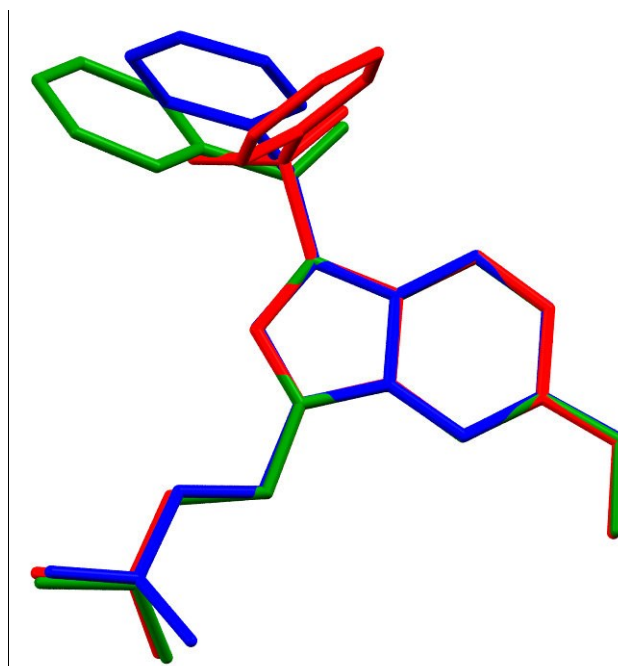


Figure 2.3.5 Superposition of 2a (red), 2b (green) and 2c (blue) with respect to the indole ring, showing the different molecular conformations observed in the crystal structure as a consequence of the different linker type.

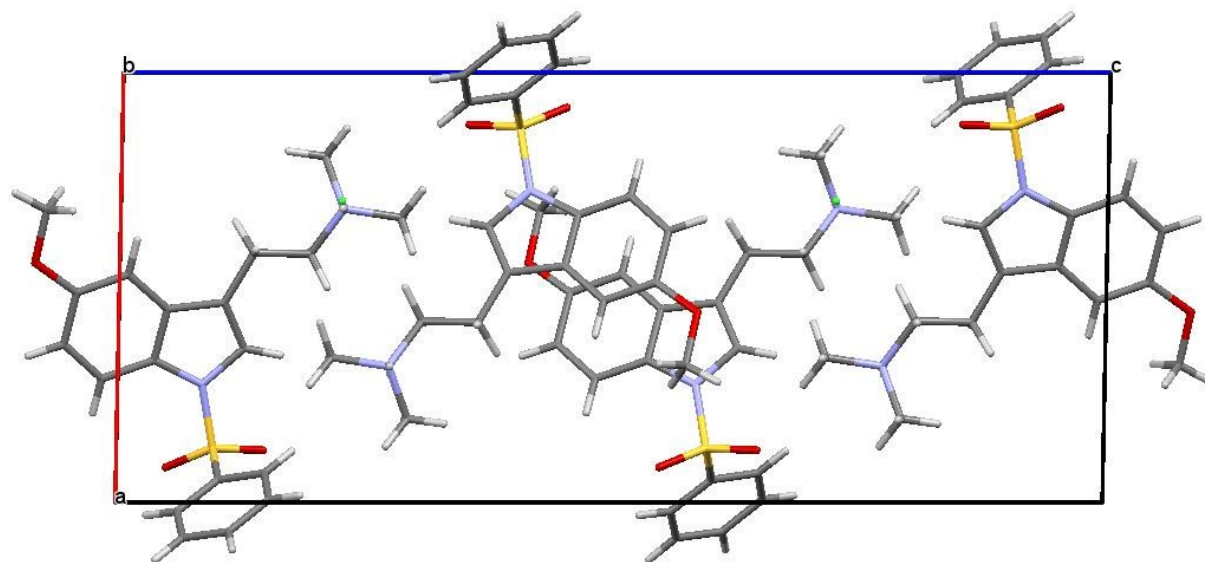


Figure 2.3.6 Crystal packing of 2a along [010]

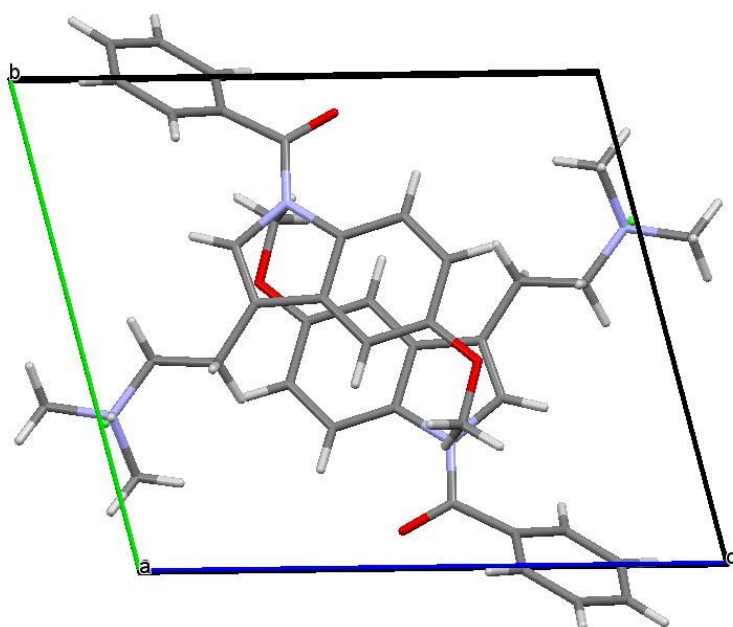


Figure 2.3.7 Crystal packing of 2b along [100]

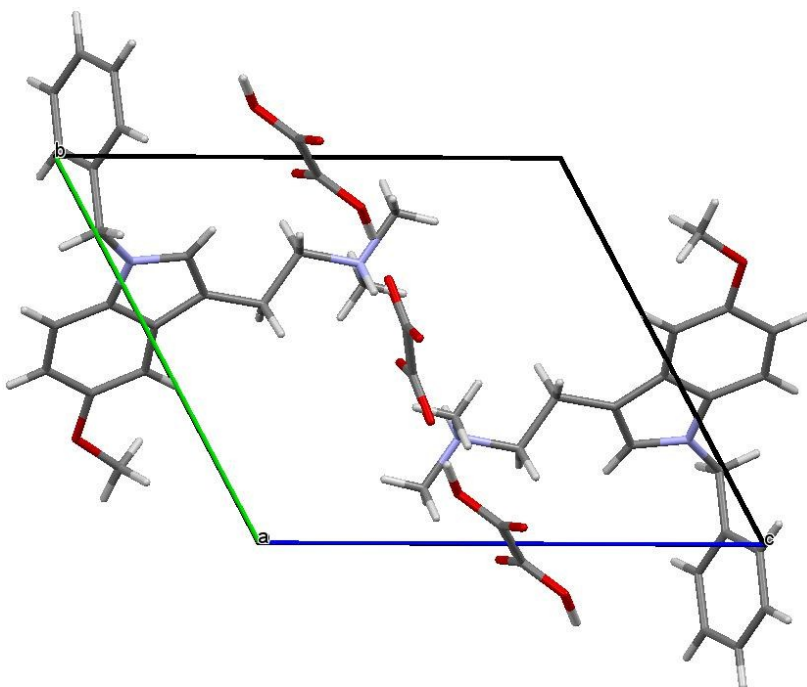


Figure 2.3.8 Crystal packing of **2c** along [100]

Table S2.3.1 Geometrical parameters of the strong and weak hydrogen bonds observed in crystal structure **2a** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(32)-H(32B)...Cl(1)#1	0.97	2.70	3.6140(16)	157.0
C(8)-H(8)...O(11)	0.93	2.58	3.115(2)	117.3
C(2)-H(2)...Cl(1)#1	0.93	2.83	3.7224(17)	161.0
C(12)-H(12)...O(11)#2	0.93	2.49	3.416(2)	172.9
C(62)-H(62B)...S(10)#3	0.96	3.00	3.796(2)	141.6
C(35)-H(35A)...O(10)#4	0.96	2.57	3.506(2)	165.3
N(33)-H(33)...Cl(1)	0.93(2)	2.16(2)	3.0887(15)	174.9(17)

Symmetry transformations used to generate equivalent atoms:

#1 $-x+1, y+1/2, -z+1/2$ #2 $-x, -y+1, -z+1$ #3 $-x+1, -y+1, -z+1$ #4 $-x+1, y-1/2, -z+1/2$

Table S2.3.2 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **2b** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(32)-H(32A)...Cl(1)#1	0.97	2.72	3.666(2)	165.3
C(35)-H(35C)...Cl(1)#1	0.96	2.97	3.834(3)	151.2
C(34)-H(34C)...Cl(1)#2	0.96	2.97	3.862(3)	155.9
C(8)-H(8)...O(10)	0.93	2.44	2.940(3)	113.6
N(33)-H(33)...Cl(1)	0.96(2)	2.07(3)	3.032(2)	177(2)

Symmetry transformations used to generate equivalent atoms:

#1 -x,-y+1,-z #2 x+1,y,z

Table S2.3.3 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **2c** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(31)-H(31)...O(22)	1.01(3)	1.55(3)	2.5560(16)	179(3)
C(32)-H(32A)...O(32)#2	0.97	2.44	3.3893(19)	165.3
C(2)-H(2)...O(32)#2	0.93	2.57	3.485(2)	166.8
C(34)-H(34B)...O(22)#2	0.96	2.61	3.132(2)	114.8
N(33)-H(6)...O(22)#2	0.908(19)	2.391(18)	2.9829(18)	122.9(14)
N(33)-H(6)...O(21)	0.908(19)	1.86(2)	2.7304(18)	159.6(17)
O(31)-H(31)...O(22)	1.01(3)	1.55(3)	2.5560(16)	179(3)
C(32)-H(32A)...O(32)#2	0.97	2.44	3.3893(19)	165.3
C(2)-H(2)...O(32)#2	0.93	2.57	3.485(2)	166.8
C(34)-H(34B)...O(22)#2	0.96	2.61	3.132(2)	114.8
N(33)-H(6)...O(22)#2	0.908(19)	2.391(18)	2.9829(18)	122.9(14)
N(33)-H(6)...O(21)	0.908(19)	1.86(2)	2.7304(18)	159.6(17)

Symmetry transformations used to generate equivalent atoms:

#1 -x,-y,-z+1 #2 -x,-y+1,-z+1

S2.4 Crystal structure description of compounds **3a** and **3c**

Compounds **3a** and **3c** crystallize as oxalate salts. Crystal data and refinement parameters are shown in **Table 2.4**. The asymmetric unit of compound **3a** (**Figure 2.4.1**) consists of a single molecule of the indole derivative and single-deprotonated oxalic acid molecule. Molecules of the arylsulfonyl derivative are protonated, with a positive charge localized on the N34 atom. The donor-bifurcated charge-assisted hydrogen bond is formed between the protonated N34 and two oxygen atoms of carboxylic groups of oxalates. In structure **3a**, the oxalate molecules interact with each other by the strong hydrogen bond O-H...O, creating channels along the [001] direction (**Figure 2.4.2**). The crystal structure **3c** has two independent cations of the benzyl indole derivative in asymmetric units assisted by the double anion of oxalate and one oxalic acid molecule. Additionally, a methanol molecule, as a co-crystallizing solvent, is present in the crystal lattice. The deprotonated oxalate interacts with two molecules of compound **3c** by donor-bifurcated, charge-assisted hydrogen bonds N34(A)⁺-H...O. Oxalic acid molecules interact with oxalate by O-H...O hydrogen bonds, forming polar channels that expand along [100]. The methanol molecule stabilizes the crystal lattice by an additional O-H...O hydrogen bond as a donor, with the carbonyl oxygen atom of the oxalic acid acting as an acceptor (**Figure 2.4.3**).

The indole ring superposition of the two independent molecules of compound **3c** shows differences in the spatial orientations of the piperazine and benzene rings (**Figure 2.4.4**), resulting from the diverse neighborhood of these fragments in the crystal environment and thus leading to several weak interactions that stabilize the structure. Due to the observed conformational differences, molecules **3c** and **3cA** are considered two separate and possible conformers.

The superposition of single molecules of **3a**, **3c** and **3cA** extracted from the crystal structure with respect to the indole ring is shown in **Figure 2.4.5**.

The crystal packing schemes of structures **3a** and **3c** are shown in **Figures 2.4.6-7**. The crystal structures of compounds **3a** and **3c** are further stabilized by several weak C-H...O hydrogen bonds. The geometric parameters of strong and weak hydrogen bonds are shown in **Tables S2.4.1-2**.

Table S2.4 Crystal data and refinement parameters for compounds **3a** and **3c**

Identification code	3a	3c
Chemical formula	C ₁₉ H ₂₂ N ₃ O ₂ S ⁺ , C ₂ HO ₄ ⁻	2(C ₂₀ H ₂₄ N ₃) ⁺ , C ₂ O ₄ ²⁻ , C ₂ H ₂ O ₄ , CH ₃ OH
Formula mass	445.48	822.94
Crystal data		
Crystal system	Monoclinic	Monoclinic
Space group	P2 ₁ /c	P2 ₁ /n
Unit cell dimensions	a = 8.1025(2) Å b = 25.7392(4) Å c = 10.5439(2) Å α = 90° β = 109.355(2)° γ = 90°	a = 11.0646(2) Å b = 25.5172(4) Å c = 15.8131(2) Å α = 90° β = 107.833(2)° γ = 90°
Unit cell volume [Å ³]	2074.67(8)	4250.12(12)
Z	4	4
D _{calc} [g/cm ³]	1.426	1.286
Absorption coefficient [mm ⁻¹]	0.201	0.090
F(000)	936	1752
Crystal size [mm ³]	0.5 x 0.4 x 0.1	0.5 x 0.2 x 0.2
Data collection		
Temperature [K]	120(2)	120(2)
Radiation type	MoKα	MoKα
θ range [°]	3.099° to 28.714°	3.078° to 28.714°
Index ranges	-10 ≤ h ≤ 10, -34 ≤ k ≤ 34, -14 ≤ l ≤ 13	-14 ≤ h ≤ 14, -33 ≤ k ≤ 33, -21 ≤ l ≤ 20
Reflections collected	28278	58649
Independent reflections	5061 [R(int) = 0.0387]	10446 [R(int) = 0.0388]
Completeness [%]	99.8 (θ = 25.2°)	99.7 (θ = 25.0°)
Refinement		
Data/restraints/parameters	5061 / 0 / 288	10446 / 2 / 561
Goodness-of-fit on F ²	1.081	1.038
Final R indices [I > 2σ(I)]	R1=0.0383, wR2=0.0906	R1=0.0462, wR2=0.1037
R indices (all data)	R1=0.0517, wR2=0.0997	R1=0.0653, wR2=0.1141

$\Delta\rho_{\max}/\Delta\rho_{\min}$ [$e\cdot\text{\AA}^{-3}$]	0.37 and -0.42	0.41 and -0.26
---	----------------	----------------

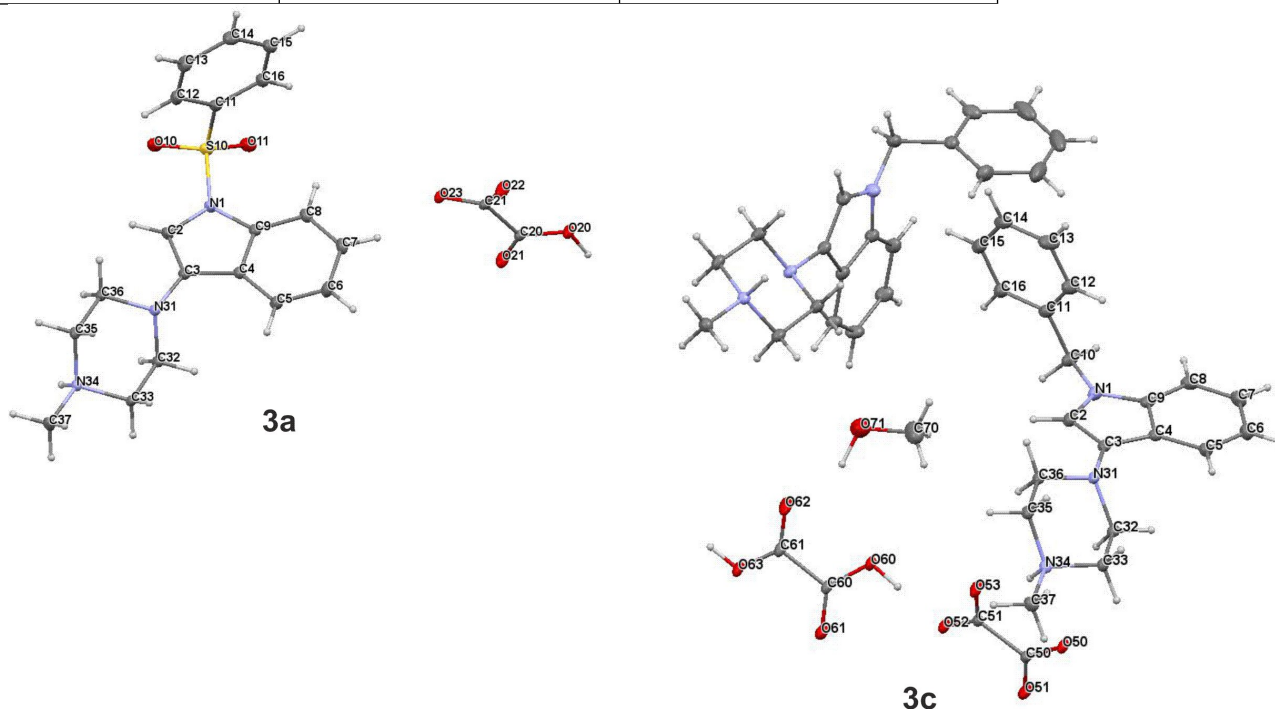


Figure 2.4.1 Asymmetric units in the crystal structures of 3a and 3c. For structure 3c, the asymmetric unit consists of two molecules of the indole derivative (only one molecule is labeled for the figure clarity - labels of second molecule have corresponding numbers additionally indicated with letter A), two oxalic acid molecules, from which only one is fully deprotonated, with double-negative charge. Additionally, a methanol molecule was observed in the crystal structure of compound 3c. Displacement ellipsoids of non-hydrogen atoms are drawn at the 30% probability level. H atoms are presented as small spheres with an arbitrary radius.

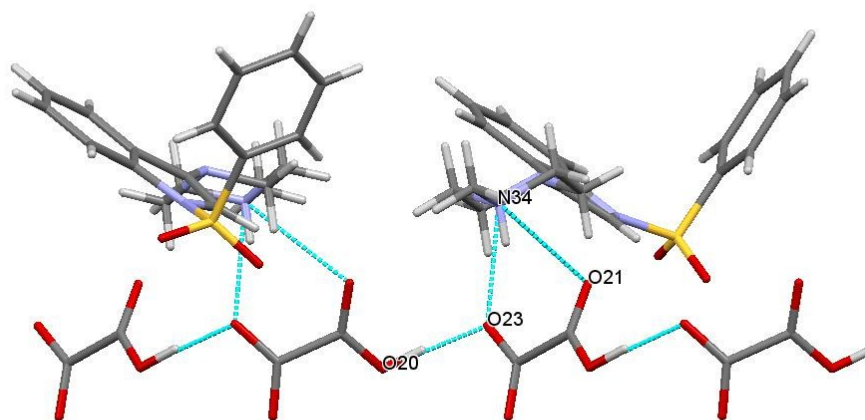


Figure 2.4.2 The bifurcated hydrogen bond in the crystal structure of 3a, formed between protonated N34 and the oxalate anion. The oxalate anions interact with each other *via* O-H...O hydrogen bonds,

forming wavy-chain motif that is elongated in the [001] direction. Hydrogen bonds are shown as cyan dashed lines.

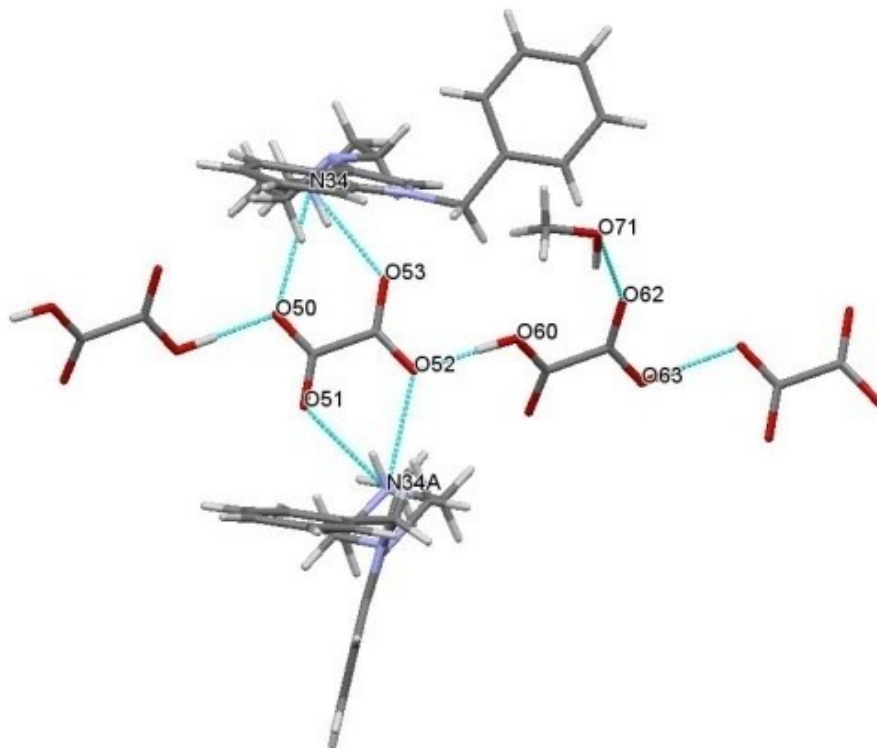


Figure 2.4.3 The strong hydrogen bond motifs observed in crystal structure 3c. The positively charged N34 and N34A serve as donors of the bifurcated hydrogen bond with oxygen atoms of the fully deprotonated oxalate anion serve as acceptors. The oxalate anion and oxalic acid molecule form a chain *via* O-H...O hydrogen bonds that is elongated in the [100] direction. The co-crystallizing methanol molecule is a donor of the hydrogen bond with oxygen atom of the oxalic acid carbonyl group as an acceptor. Hydrogen bonds are shown as cyan dashed lines.

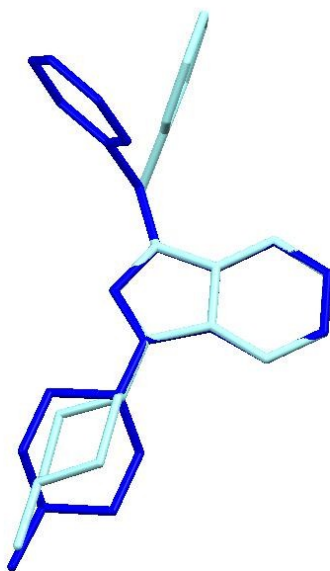


Figure 2.4.4 Superposition of the two alternative conformers observed in the asymmetric with respect to the indole ring (3c (blue) and 3cA (light blue)).

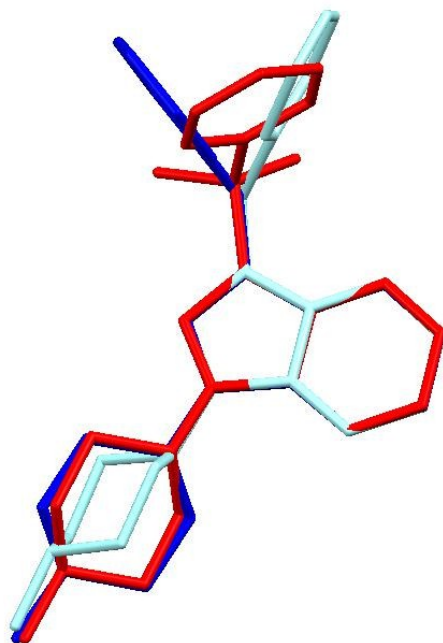


Figure 2.4.5 Superposition of molecules 3a (red), 3c (blue) and 3cA (light blue) with respect to the indole ring, showing the different conformations observed in the crystal structure as a consequence of the different linkers/conformers.

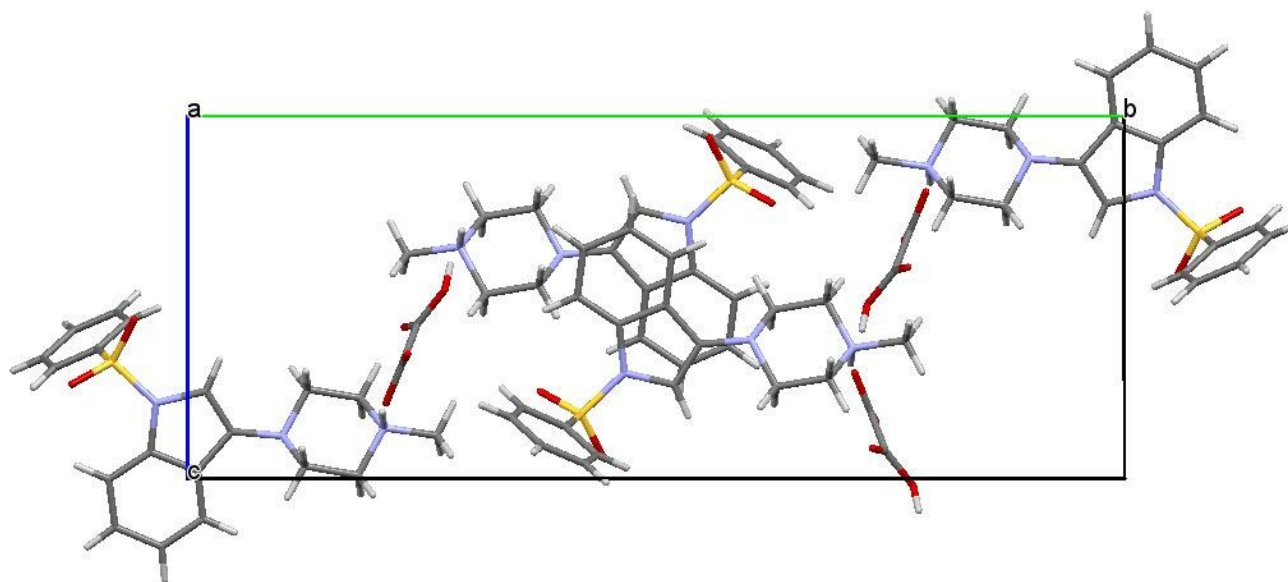


Figure 2.4.6 Crystal packing of 3a along [100]

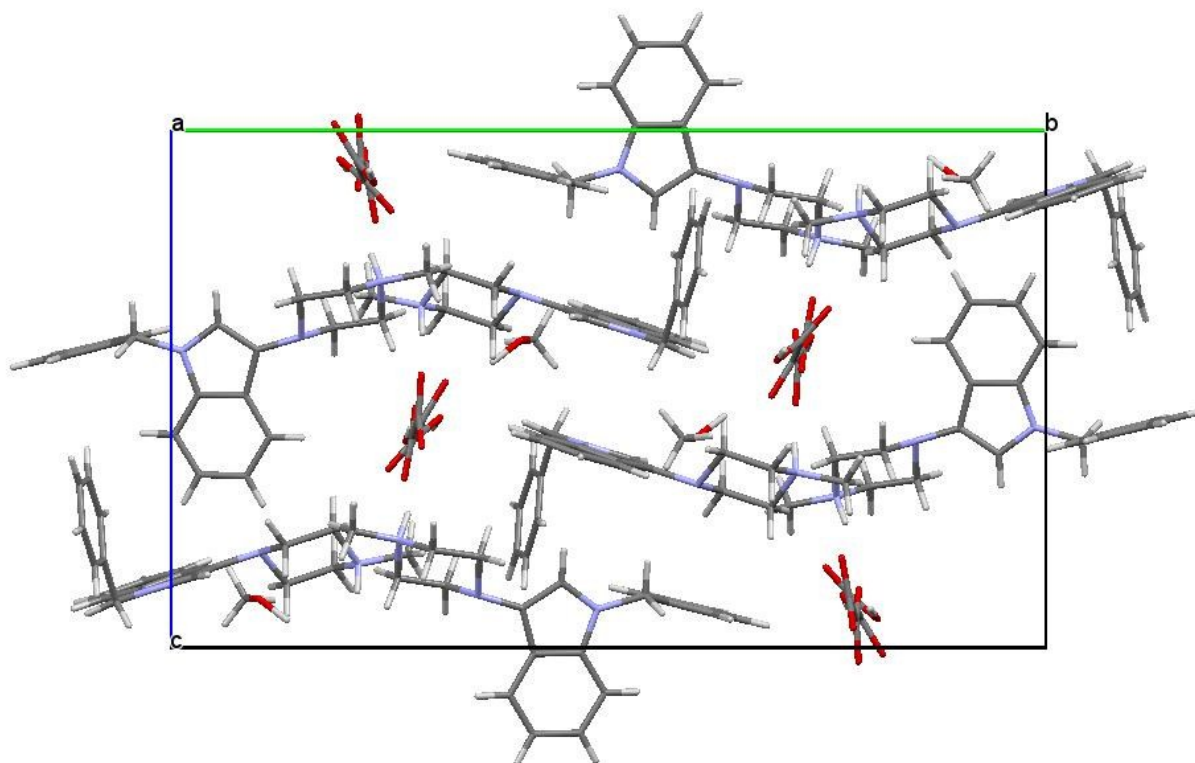


Figure 2.4.7 Crystal packing of **3c** along [100]

Table S2.4.1 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **3a** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(33)-H(33A)...O(21)#1	0.99	2.62	3.5273(19)	151.7
C(2)-H(2)...O(10)#2	0.95	2.65	3.454(2)	142.4
C(35)-H(35B)...O(20)#3	0.99	2.59	3.5276(18)	157.4
C(35)-H(35A)...O(21)#4	0.99	2.48	3.0448(18)	116.2
C(8)-H(8)...O(11)	0.95	2.55	3.096(2)	117.2
C(37)-H(37C)...O(22)#3	0.98	2.44	3.247(2)	139.0
C(37)-H(37A)...O(11)#5	0.98	2.58	3.2138(18)	122.1
C(12)-H(12)...O(10)#2	0.95	2.65	3.596(2)	177.8
N(34)-H(34)...O(23)#4	0.88(2)	1.91(2)	2.7422(16)	157.5(18)
N(34)-H(34)...O(21)#4	0.88(2)	2.414(19)	2.9676(17)	121.1(15)

O(20)-H(20)...O(23)#6	0.96(3)	1.59(3)	2.5493(14)	174(2)
C(33)-H(33A)...O(21)#1	0.99	2.62	3.5273(19)	151.7
C(2)-H(2)...O(10)#2	0.95	2.65	3.454(2)	142.4
C(35)-H(35B)...O(20)#3	0.99	2.59	3.5276(18)	157.4
C(35)-H(35A)...O(21)#4	0.99	2.48	3.0448(18)	116.2
C(8)-H(8)...O(11)	0.95	2.55	3.096(2)	117.2
C(37)-H(37C)...O(22)#3	0.98	2.44	3.247(2)	139.0
C(37)-H(37A)...O(11)#5	0.98	2.58	3.2138(18)	122.1
C(12)-H(12)...O(10)#2	0.95	2.65	3.596(2)	177.8
N(34)-H(34)...O(23)#4	0.88(2)	1.91(2)	2.7422(16)	157.5(18)
N(34)-H(34)...O(21)#4	0.88(2)	2.414(19)	2.9676(17)	121.1(15)
O(20)-H(20)...O(23)#6	0.96(3)	1.59(3)	2.5493(14)	174(2)

Symmetry transformations used to generate equivalent atoms:

#1 -x+1,y+1/2,-z+1/2 #2 -x+1,-y+1,-z+2 #3 -x,y+1/2,-z+1/2

#4 -x+1,-y+1,-z+1 #5 -x+1,y+1/2,-z+3/2 #6 x,-y+1/2,z-1/2

Table S2.4.2 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **3c** (molecule 3c' label system is corresponding to the one in molecule 3c, with A later after atom number) [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(34)-H(34)...O(53)	0.896(18)	1.874(18)	2.7251(16)	157.7(15)
N(34)-H(34)...O(50)	0.896(18)	2.456(17)	3.0668(16)	125.7(13)
O(60)-H(60)...O(52)	0.914(10)	1.576(10)	2.4894(13)	177(3)
O(63)-H(63)...O(50)#1	0.915(10)	1.556(10)	2.4698(14)	177(3)
O(71)-H(71)...O(62)	1.05(4)	1.82(4)	2.788(2)	152(3)
N(34A)-H(4AA)...O(52)#2	0.901(18)	2.168(18)	2.8247(16)	129.2(14)
N(34A)-H(4AA)...O(51)#2	0.901(18)	1.962(18)	2.7543(16)	145.7(15)
C(33)-H(33A)...O(62)#3	0.99	2.40	3.2089(18)	138.3
C(33)-H(33A)...O(50)	0.99	2.59	3.0750(17)	110.4
C(33)-H(33B)...O(61)#4	0.99	2.61	3.4344(18)	140.9
C(35)-H(35A)...O(51)#5	0.99	2.63	3.5260(17)	150.1
C(37)-H(37B)...O(61)#4	0.98	2.50	3.3720(19)	148.7
C(37)-H(37B)...O(63)#4	0.98	2.53	3.2936(19)	135.2
C(33A)-H(3AA)...O(60)#5	0.99	2.45	3.3698(18)	153.7
C(33A)-H(3AB)...O(53)#5	0.99	2.61	3.1989(17)	118.4

C(37A)-H(7AA)...O(61)#2	0.98	2.65	3.5099(19)	146.6
C(37A)-H(7AC)...O(62)#5	0.98	2.60	3.552(2)	162.6
C(35A)-H(5AA)...O(61)#2	0.99	2.65	3.5315(19)	148.5
N(34)-H(34)...O(53)	0.896(18)	1.874(18)	2.7251(16)	157.7(15)
N(34)-H(34)...O(50)	0.896(18)	2.456(17)	3.0668(16)	125.7(13)
O(60)-H(60)...O(52)	0.914(10)	1.576(10)	2.4894(13)	177(3)
O(63)-H(63)...O(50)#1	0.915(10)	1.556(10)	2.4698(14)	177(3)
O(71)-H(71)...O(62)	1.05(4)	1.82(4)	2.788(2)	152(3)
N(34A)-H(4AA)...O(52)#2	0.901(18)	2.168(18)	2.8247(16)	129.2(14)
N(34A)-H(4AA)...O(51)#2	0.901(18)	1.962(18)	2.7543(16)	145.7(15)
C(33)-H(33A)...O(62)#3	0.99	2.40	3.2089(18)	138.3
C(33)-H(33A)...O(50)	0.99	2.59	3.0750(17)	110.4
C(33)-H(33B)...O(61)#4	0.99	2.61	3.4344(18)	140.9
C(35)-H(35A)...O(51)#5	0.99	2.63	3.5260(17)	150.1
C(37)-H(37B)...O(61)#4	0.98	2.50	3.3720(19)	148.7
C(37)-H(37B)...O(63)#4	0.98	2.53	3.2936(19)	135.2
C(33A)-H(3AA)...O(60)#5	0.99	2.45	3.3698(18)	153.7
C(33A)-H(3AB)...O(53)#5	0.99	2.61	3.1989(17)	118.4
C(37A)-H(7AA)...O(61)#2	0.98	2.65	3.5099(19)	146.6
C(37A)-H(7AC)...O(62)#5	0.98	2.60	3.552(2)	162.6
C(35A)-H(5AA)...O(61)#2	0.99	2.65	3.5315(19)	148.5

Symmetry transformations used to generate equivalent atoms:

#1 $x-1,y,z$ #2 $x-1,y,z-1$ #3 $x+1,y,z$ #4 $x+1/2,-y+1/2,z-1/2$

#5 $x-1/2,-y+1/2,z-1/2$

S2.5 Crystal structure description of compounds 4a, 4b and 4c

Compounds **4a**, **4b** and **4c** crystallize as oxalates. Crystal data and refinement parameters are shown in **Table 2.5**. The asymmetric unit (**Figure 2.5.1**) of compound **4a** consists of two molecules of the arylsulfonyl derivative, protonated on N54 and N54A atoms. Superposition of both independent molecules with respect to the indole ring exhibits different orientation of the piperazine moiety (**Figure 2.5.2**). The sulfonyl group in one molecule is positionally disordered (site occupancies of 63% and 37% for both alternative positions), what is related to the unexpected bromination of the C3 atom of the indole ring. This bromination was discovered by a detailed inspection of the difference Fourier map and further confirmed by MS (mass spectroscopy) performed for dissolved crystals (see Section **S3**). The bromine atom occupies positions of approximately only 4% in the crystal structure.

In the asymmetric unit, two oxalate counter ions interact with each other by O-H...O hydrogen bonds, forming bands that occupy channels in the [010] direction. Hydrogen H70 is positioned between of O62 and O70 of the two interacting oxalates, suggesting a plausible delocalization of this proton. However, one oxalate is a double acceptor of the donor-bifurcated N54⁺-H...O61/O63 and N54A⁺-H...O60/O62 (**Figure 2.5.3**). This accumulation of positive charge around one oxalate suggests its full deprotonation and a double-negative charge of this molecule.

Table S2.5 Crystal data and refinement parameters for compounds **4a-4c**

Identification code	4a	4b	4c
Chemical formula	C ₁₉ H _{21.92} Br _{0.08} N ₃ O ₂ S ⁺ , C ₁₉ H ₂₂ N ₃ O ₂ ⁺ , C ₂ O ₄ ²⁻ , C ₂ H ₂ O ₄ ,	4(C ₂₀ H ₂₂ N ₃ O) ⁺ , C ₂ O ₄ ²⁻ , 2(C ₂ HO ₄) ⁻ , C ₂ H ₂ O ₄ , 2(H ₂ O)	C ₂₀ H ₂₄ N ₃ ⁺ , C ₂ HO ₄ ⁻
Formula mass	897.28	1673.76	395.45
Crystal data			
Crystal system	Monoclinic	Monoclinic	Monoclinic
Space group	P2 ₁ /n	P2 ₁	P2 ₁ /n
Unit cell dimensions	a = 17.3674(3) Å b = 11.1504(1) Å c = 23.6192(3) Å α = 90° β = 109.303(2)° γ = 90°	a = 7.3613(2) Å b = 39.3723(9) Å c = 14.2823(4) Å α = 90° β = 98.867(3)° γ = 90°	a = 11.9809(4) Å b = 11.1869(3) Å c = 14.9549(5) Å α = 90° β = 98.115(3)° γ = 90°
Unit cell volume [Å ³]	4316.81(11)	4089.99(19)	1984.32(11)
Z	4	2	4
D _{calc} [g/cm ³]	1.381	1.359	1.324
Absorption coefficient [mm ⁻¹]	0.267	0.099	0.092
F(000)	1883	1768	840
Crystal size [mm ³]	0.3 x 0.2 x 0.1	0.5 x 0.3 x 0.2	0.3 x 0.3 x 0.3
Data collection			
Temperature [K]	120(2)	120(2)	120(2)
Radiation type	MoKα	MoKα	MoKα
θ range [°]	2.975° to 28.772°	2.848° to 29.536°	2.971° to 29.586°
Index ranges	-23 ≤ h ≤ 23, -14 ≤ k ≤ 14, -31 ≤ l ≤ 31	-9 ≤ h ≤ 10, -54 ≤ k ≤ 53, -19 ≤ l ≤ 19	-16 ≤ h ≤ 16, -15 ≤ k ≤ 14, -20 ≤ l ≤ 20
Reflections collected	58928	57362	28581
Independent reflections	10491 [R(int)=0.0538]	20842 [R(int)=0.0558]	5258 [R(int)=0.0798]
Completeness [%]	99.8 (θ = 25.2°)	99.8 (θ = 25.0°)	99.8 (θ = 25.2°)
Refinement			
Data/restraints/parameters	10491 / 0 / 618	20842 / 2 / 1149	5258 / 0 / 270
Goodness-of-fit on F ²	1.054	1.028	1.085
Final R indices [I > 2σ(I)]	R1=0.0509, wR2= 0.0985	R1=0.0576, wR2=0.1068	R1=0.0569, wR2=0.1035
R indices (all data)	R1=0.0877, wR2= 0.1147	R1=0.0942, wR2=0.1230	R1=0.1107, wR2=0.1289

$\Delta\rho_{\max}/\Delta\rho_{\min}$ [$e\cdot\text{\AA}^{-3}$]	0.67 and -0.45	0.22 and -0.24	0.26 and -0.33
---	----------------	----------------	----------------

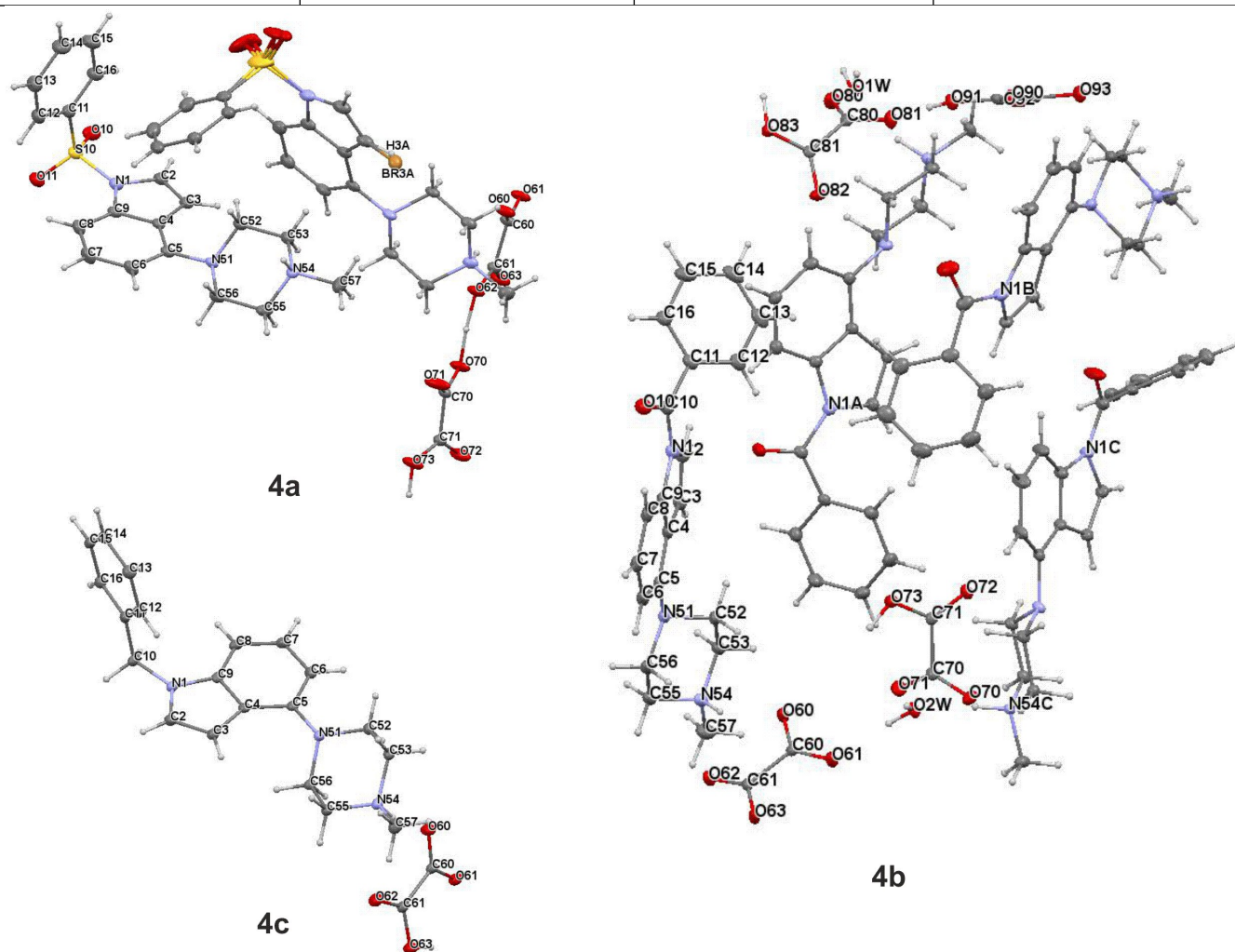


Figure 2.5.1 Asymmetric units in the crystal structures of compounds 4a, 4b and 4c. For structure 4a, the asymmetric unit consists of two indole derivative and two oxalate molecules. Within one molecule of 4a, there is positional disorder of the sulfonyl group, and partial bromination was observed at the indole C3 position. The asymmetric unit of 4b is complicated and consists of four molecules that are protonated at the N54 atom, one oxalic acid, two single-deprotonated oxalates (charge 1- each) and one fully deprotonated oxalate with a total charge of 2-. Additionally, two water molecules as a co-crystallizing solvent are observed in the crystal structure of compound 4b. For structures 4a and 4b only one molecule is labeled for figure clarity. Displacement ellipsoids of non-hydrogen atoms are drawn at the 30% probability level. H atoms are presented as small spheres with an arbitrary radius.



Figure 2.5.2 Superposition with respect to the indole ring of two alternative conformers observed in the asymmetric unit of the crystal structure 4a (4a (red) and partially brominated 4aA (yellow)).

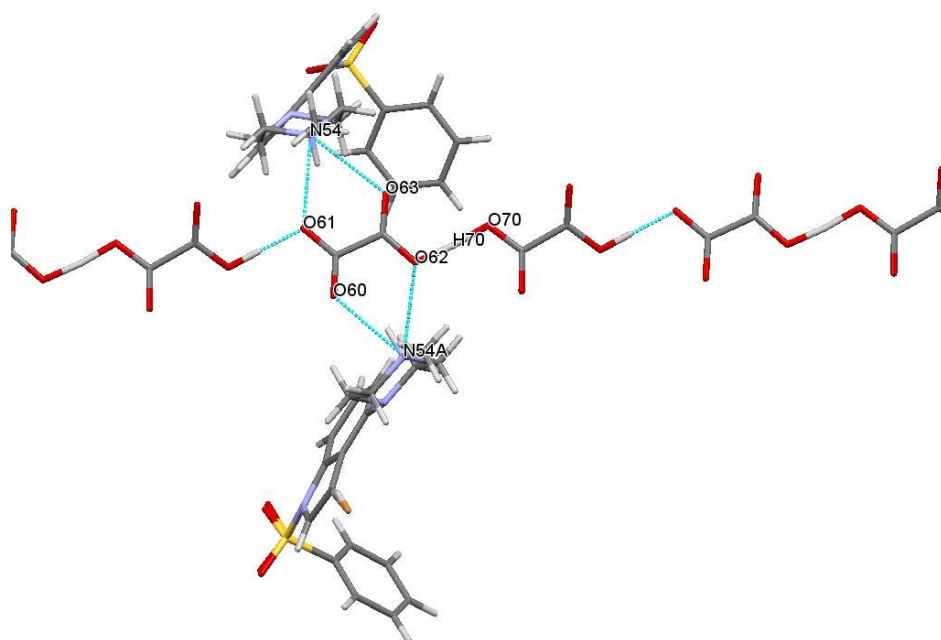


Figure 2.5.3 The strong hydrogen bond motifs observed in crystal structure 4a. The positively charged N54 and N54A serve as bifurcated donors of the hydrogen bond and oxygen atoms of the fully deprotonated oxalate anion serve as acceptors. The oxalate anion and oxalic acid molecule form a chain *via* O-H...O hydrogen bonds that is elongated in the [010] direction.

The most complicated asymmetric unit is observed for the crystal structure of **4b**, in which four protonated molecules of the main compound are observed (**Figure 2.5.1**). In contrast to the above-described structure **4a**, in this case all molecules are in a similar conformation (superposition of all independent molecules with respect to indole ring is shown in **Figure 2.3.4**). Excluding cationic molecules of the indole derivative, the asymmetric units contain two water and four oxalate molecules. Careful inspection of the difference Fourier map and hydrogen bond network demonstrated that oxalic acid molecules are not equivalent in this structure. One of them is fully deprotonated, carrying a double-negative charge, two possess a single negative charge and one molecule is neutral. The different protonation states of oxalic acid leads to reduced symmetry and a greater number of molecules in the asymmetric unit. Similar to structure **4a**, positively charged N54⁺ and N54B⁺ are bifurcated donors of the hydrogen bonds, while N54A⁺ and N54C⁺ interact indirectly with oxalate, by water molecules (**Figure 2.5.5**). Due to multiple donors and acceptors, the hydrogen bond framework is complicated, forming multiple ring motives (**Figure 2.5.6 (left)**). These ring motives form a wavy net perpendicular to the [010] axis (**Figure 2.5.6 (right)**).

The asymmetric unit of **4c** consists of one molecule of the indole derivative with a positive charge localized on the N54 atom of the piperazine moiety, accompanied by an oxalate anion. The strongest interaction observed in this crystal structure is the charge-assisted hydrogen bond with the bifurcated donor. Additionally, two oxalate anions related *via* the inversion center interact by O-H...O hydrogen bonds, forming dimers in the crystal structure (**Figure 2.5.7**).

The different conformations of a single molecules of **4a**, **4b** and **4c** extracted from the crystal structure, depending on the linker type are shown in **Figure 2.5.8** as superpositions with respect to the indole ring.

The crystal packing schemes of structures **4a-4c** are shown in **Figures 2.5.9-11**.

The crystal structures of compounds **4a-4c** are further stabilized by several weak hydrogen bonds C-H...A (where **A** are O, N and Br). The geometric parameters of strong and weak hydrogen bonds are shown in **Tables S2.5.1-3**.

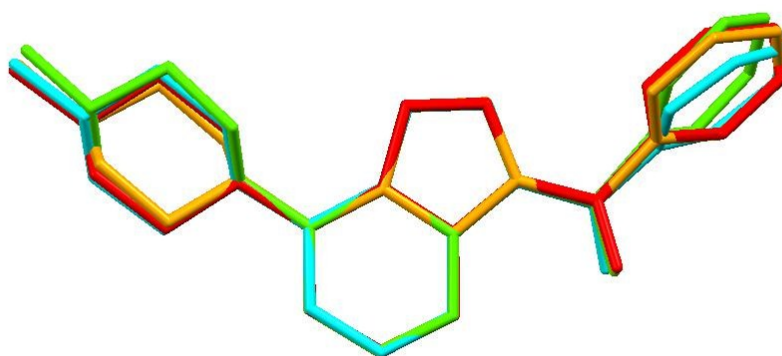


Figure 2.5.4 Superposition of all independent molecules in the asymmetric unit of structure **4b** with respect to the indole ring.

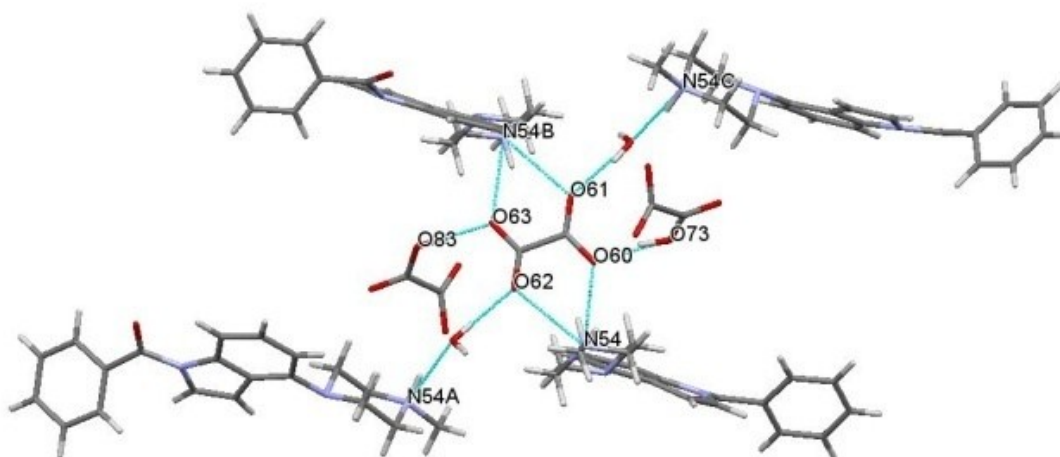


Figure 2.5.5 Charge-assisted hydrogen bonds motifs observed in crystal structure **4b**.

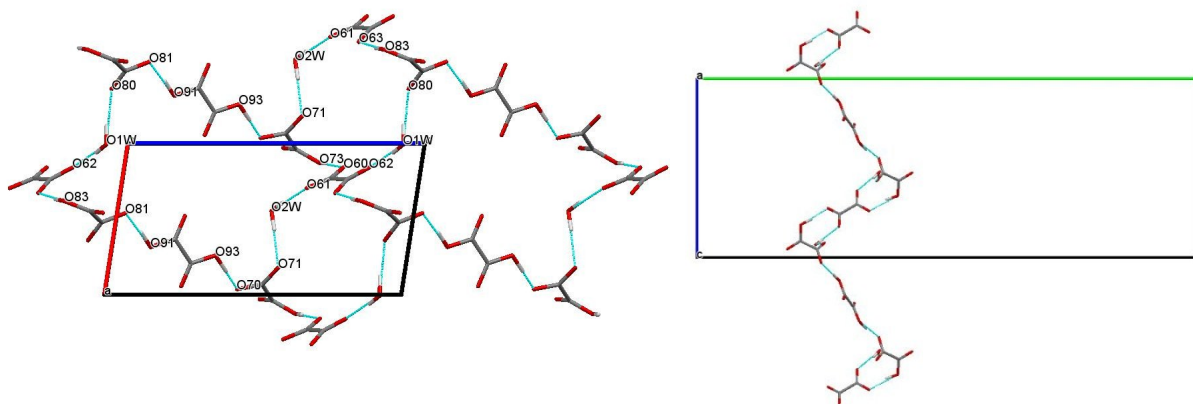


Figure 2.5.6 Complicated hydrogen bond ring motifs observed in crystal structure 4b, forming a net perpendicular to the [010] axis (the left is the view along [010], the right is the view along [100] showing the wave of the net).

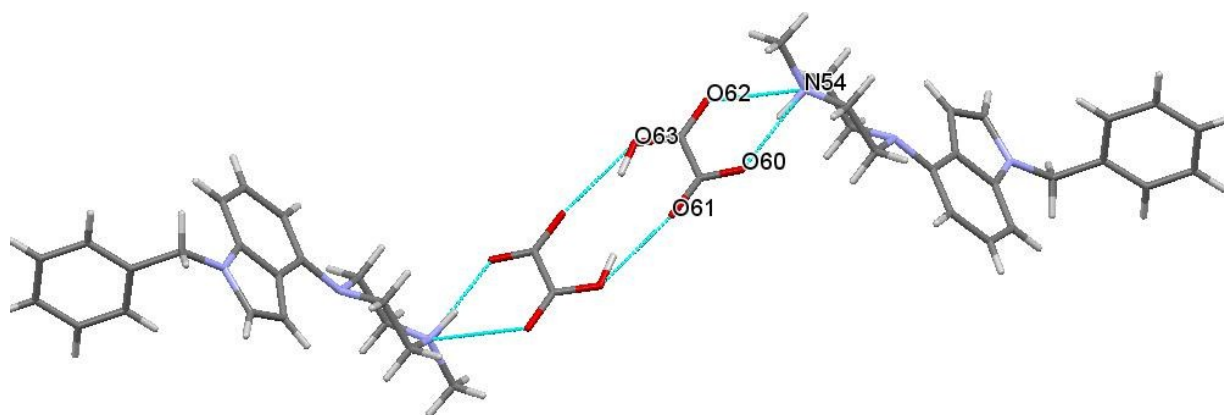


Figure 2.5.7 Centrosymmetric dimers observed in the crystal structure of 4c. The positively charged N54 serves as bifurcated donor of the hydrogen bond, with O62 and O60 oxygen atoms of the oxalate molecule serving as acceptors.

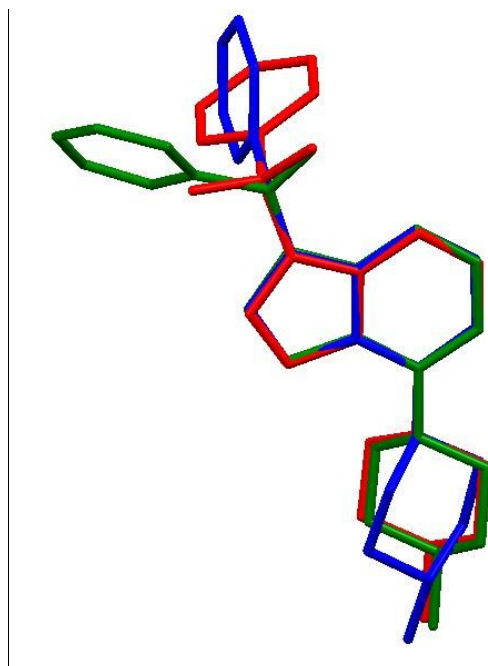


Figure 2.5.8 Superposition of compounds 4a (red), 4b (green) and 4c (blue) with respect to the indole ring, showing the different molecular conformations observed in the crystal structure, as a consequence of the different linker types.

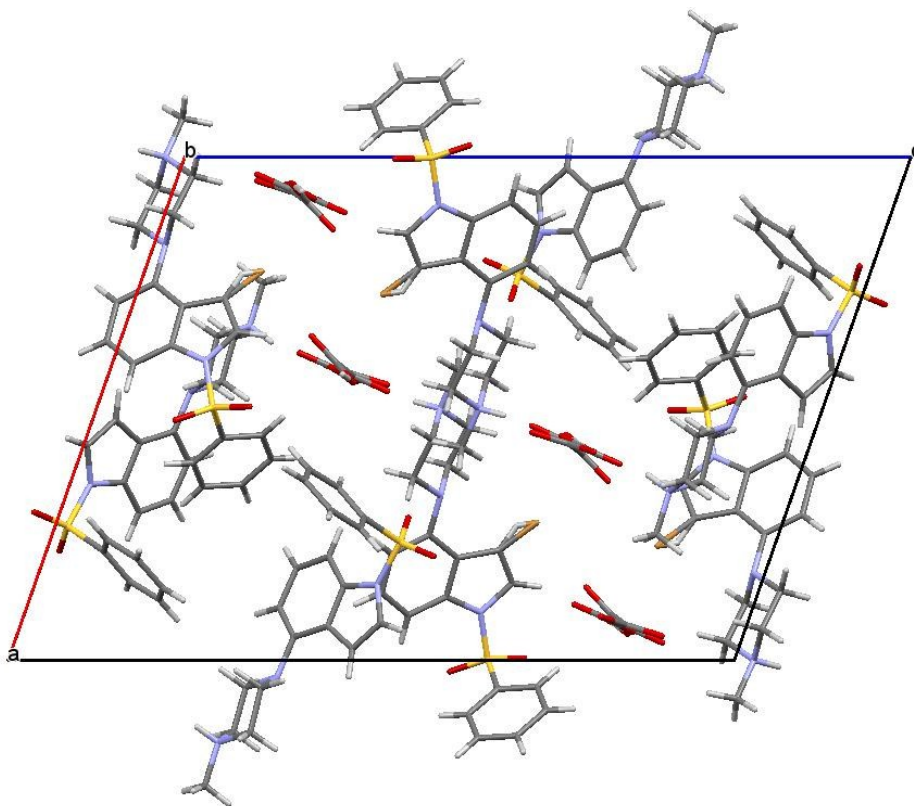


Figure 2.5.9 Crystal packing of 4a along [010]

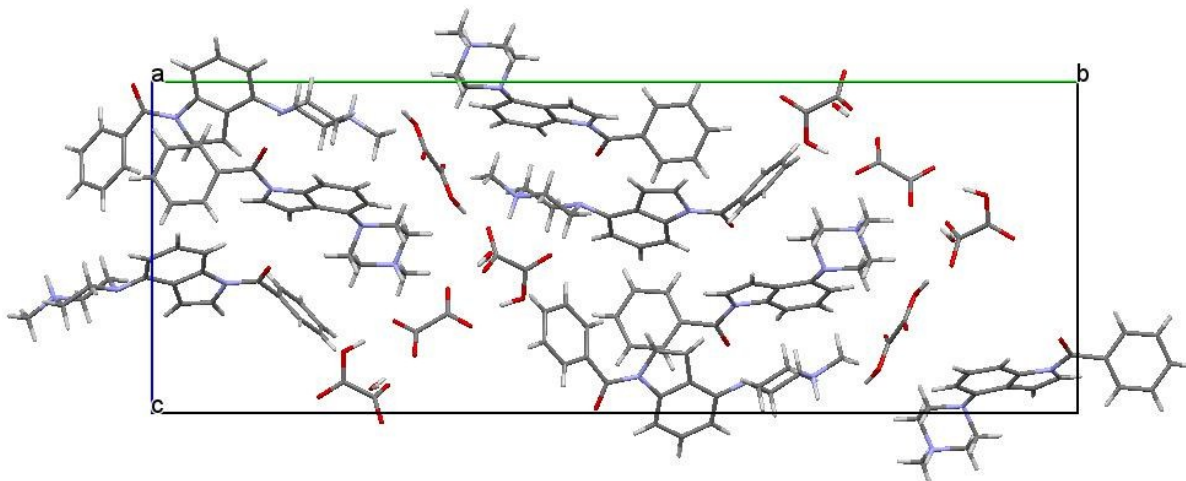


Figure 2.5.10 Crystal packing of 4b along [100]

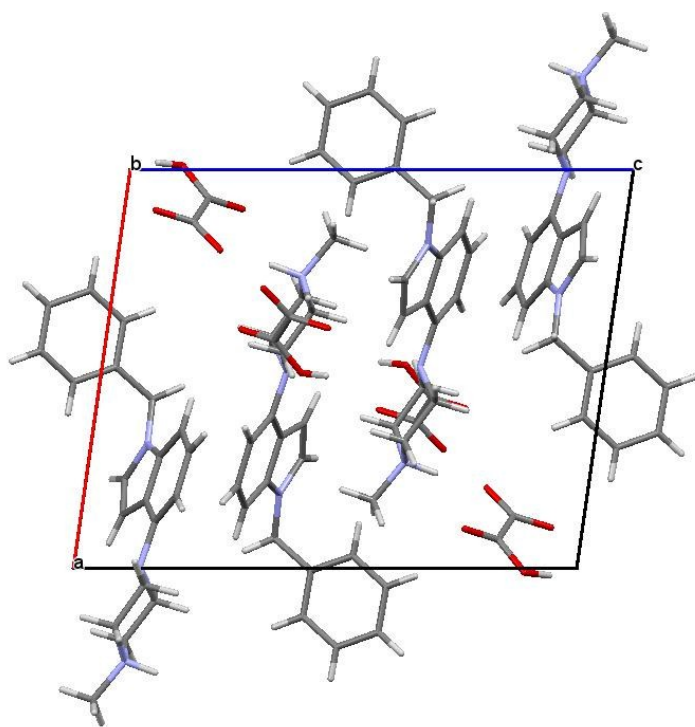


Figure 2.5.11 Crystal packing of 4c along [010]

Table S2.5.1 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **4a** (corresponding molecules of the asymmetric units have label system non and A letter after atom number) [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(56)-H(56A)...O(10A)#1	0.99	2.59	3.307(14)	129.6
C(56)-H(56A)...O(10B)#1	0.99	2.55	3.32(3)	134.5
C(55)-H(55A)...O(72)#2	0.99	2.34	3.308(3)	166.1
C(55)-H(55B)...O(10)#3	0.99	2.42	3.285(3)	144.9
C(2)-H(2)...O(11A)#4	0.95	2.65	3.568(6)	162.7
C(8)-H(8)...O(11)	0.95	2.55	3.102(3)	117.4
C(53)-H(53B)...O(63)#5	0.99	2.48	2.967(3)	110.2
C(52A)-H(2AA)...Br(3A)	0.99	2.75	3.431(3)	126.2
C(52A)-H(2AB)...O(60)	0.99	2.66	3.327(3)	125.1
C(53A)-H(3AB)...O(73)#6	0.99	2.33	3.251(3)	154.1
C(56A)-H(6AB)...O(11)#3	0.99	2.54	3.433(3)	150.6
C(12A)-H(2P)...O(70)#5	0.95	2.47	3.422(3)	174.9
C(55A)-H(5AA)...O(71)#6	0.99	2.51	3.371(3)	144.9
C(55A)-H(5AB)...O(71)	0.99	2.39	3.344(3)	162.1
C(8A)-H(8A)...O(11A)	0.95	2.52	3.060(9)	116.0
C(8A)-H(8A)...O(11B)	0.95	2.39	3.030(18)	124.5
C(57A)-H(7AC)...O(71)#6	0.98	2.45	3.310(3)	146.9
C(2A)-H(2A)...O(70)#5	0.95	2.43	3.230(3)	141.5
C(2A)-H(2A)...O(72)#5	0.95	2.52	3.382(3)	150.2
O(62)-H(70)...O(71)	1.24(4)	2.46(3)	3.296(2)	122(2)
O(70)-H(70)...O(63)	1.21(4)	2.54(3)	3.198(2)	112(2)
O(73)-H(73)...O(61)#1	1.01(4)	1.53(4)	2.537(2)	176(3)
O(73)-H(73)...O(60)#1	1.01(4)	2.63(3)	3.277(2)	122(2)
N(54A)-H(54A)...O(62)	0.92(2)	2.21(2)	2.868(2)	128.0(19)
N(54A)-H(54A)...O(60)	0.92(2)	1.93(2)	2.761(2)	150(2)
N(54)-H(54)...O(61)#5	0.87(2)	2.00(2)	2.795(2)	151(2)
N(54)-H(54)...O(63)#5	0.87(2)	2.24(2)	2.848(2)	126.7(19)

Symmetry transformations used to generate equivalent atoms:

#1 $x, y+1, z$ #2 $x+1/2, -y+5/2, z+1/2$ #3 $-x+1, -y+2, -z$
#4 $-x+1, -y+1, -z$ #5 $x+1/2, -y+3/2, z+1/2$ #6 $-x, -y+2, -z$

Table S2.5.2 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **4b** (corresponding molecules of the asymmetric units have label system non and A, B or C letters after atom number) [\AA and $^\circ$].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(8)-H(8)...O(10)	0.95	2.40	2.908(5)	113.5
C(3B)-H(3B)...O(10C)	0.95	2.48	3.210(5)	134.0
C(3)-H(3)...O(10A)#1	0.95	2.46	3.325(5)	151.1
C(52)-H(52A)...O(60)	0.99	2.56	3.256(5)	126.9
C(55C)-H(55CA)...O(72)#1	0.99	2.52	3.188(5)	124.6
C(53C)-H(3CB)...O(93)#2	0.99	2.54	3.503(5)	163.7
C(56A)-H(6AA)...O(82)	0.99	2.48	3.107(5)	120.9
C(56A)-H(6AB)...O(10B)#1	0.99	2.59	3.461(5)	147.1
C(56B)-H(6BA)...O(92)#3	0.99	2.57	3.508(5)	157.5
C(52A)-H(2AA)...O(82)#3	0.99	2.56	3.143(4)	117.5
C(8B)-H(8B)...O(10B)	0.95	2.43	2.898(6)	110.3
C(57A)-H(7AB)...O(91)#3	0.98	2.66	3.576(5)	155.7
C(57A)-H(7AA)...O(81)	0.98	2.59	3.377(5)	138.0
C(57A)-H(7AA)...O(90)	0.98	2.43	3.226(5)	137.7
C(8A)-H(8A)...O(10A)	0.95	2.36	2.886(5)	114.7
C(57C)-H(7CC)...O(90)#2	0.98	2.55	3.450(5)	153.0
C(57C)-H(7CB)...O(92)#4	0.98	2.35	3.074(5)	130.3
C(57C)-H(7CB)...O(70)#1	0.98	2.56	3.385(5)	142.2
C(6)-H(6)...O(91)#5	0.95	2.60	3.415(5)	144.6
C(8C)-H(8C)...O(10C)	0.95	2.38	2.893(5)	113.6
C(6C)-H(6C)...O(72)	0.95	2.46	3.406(5)	171.7
C(57)-H(57B)...O(80)#6	0.98	2.50	3.367(5)	147.8
N(54B)-H(54B)...O(63)#7	0.96(5)	1.94(5)	2.772(4)	144(4)
N(54B)-H(54B)...O(61)#7	0.96(5)	2.12(4)	2.857(4)	133(4)
N(54C)-H(54C)...O(2W)	0.95(4)	1.75(4)	2.692(4)	168(4)
O(1W)-H(2W)...O(62)#8	0.83(6)	1.89(6)	2.719(4)	174(6)
N(54)-H(54)...O(62)	0.88(4)	2.29(4)	2.929(4)	130(4)
N(54)-H(54)...O(60)	0.88(4)	1.94(4)	2.739(4)	151(4)
O(2W)-H(4W)...O(61)	0.84(6)	1.88(6)	2.714(4)	170(5)
N(54A)-H(54A)...O(1W)	0.95(5)	1.80(5)	2.702(4)	158(4)
O(91)-H(91)...O(81)	0.97(3)	1.53(3)	2.496(4)	173(5)

O(2W)-H(3W)...O(71)	0.91(6)	1.85(6)	2.758(4)	170(5)
O(93)-H(93)...O(70)#7	1.04(6)	1.47(6)	2.490(4)	167(5)
O(1W)-H(1W)...O(80)#3	0.94(7)	1.84(7)	2.771(4)	173(6)
O(73)-H(73)...O(60)#3	0.95(5)	1.62(5)	2.539(4)	163(5)
O(83)-H(83)...O(63)#8	0.94(6)	1.58(6)	2.519(4)	173(6)

Symmetry transformations used to generate equivalent atoms:

#1 $x-1, y, z$ #2 $-x+1, y-1/2, -z+1$ #3 $x+1, y, z$

#4 $-x, y-1/2, -z+1$ #5 $-x+1, y-1/2, -z+2$ #6 $-x, y-1/2, -z+2$

#7 $-x+1, y+1/2, -z+1$ #8 $-x+1, y+1/2, -z+2$

Table S2.5.3 Geometric parameters of the strong and weak hydrogen bonds observed in crystal structure **4c** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
C(55)-H(55B)...O(62)	0.99	2.52	3.170(2)	123.2
C(8)-H(8)...O(62)#1	0.95	2.60	3.510(2)	160.9
C(53)-H(53B)...N(51)#2	0.99	2.64	3.604(3)	163.9
C(3)-H(3)...O(61)#3	0.95	2.47	3.322(2)	149.1
C(57)-H(57B)...O(62)	0.98	2.64	3.256(3)	121.1
N(54)-H(54)...O(60)	0.98(2)	1.76(2)	2.714(2)	166(2)
N(54)-H(54)...O(62)	0.98(2)	2.51(2)	3.055(2)	115.2(16)
O(63)-H(63)...O(61)	0.92(3)	2.17(2)	2.6791(19)	114.5(19)
O(63)-H(63)...O(61)#4	0.92(3)	1.91(3)	2.697(2)	143(2)
C(55)-H(55B)...O(62)	0.99	2.52	3.170(2)	123.2
C(8)-H(8)...O(62)#1	0.95	2.60	3.510(2)	160.9
C(53)-H(53B)...N(51)#2	0.99	2.64	3.604(3)	163.9
C(3)-H(3)...O(61)#3	0.95	2.47	3.322(2)	149.1
C(57)-H(57B)...O(62)	0.98	2.64	3.256(3)	121.1
N(54)-H(54)...O(60)	0.98(2)	1.76(2)	2.714(2)	166(2)
N(54)-H(54)...O(62)	0.98(2)	2.51(2)	3.055(2)	115.2(16)
O(63)-H(63)...O(61)	0.92(3)	2.17(2)	2.6791(19)	114.5(19)
O(63)-H(63)...O(61)#4	0.92(3)	1.91(3)	2.697(2)	143(2)

Symmetry transformations used to generate equivalent atoms:

#1 $x+1, y, z$ #2 $-x+1, -y+1, -z+1$ #3 $-x+1/2, y+1/2, -z+1/2$

#4 $-x, -y+1, -z$

S2.6 Comparison of selected geometric parameters for the presented structures that can be crucial for binding to 5HT₆R

Cg1 and Cg2 are centroids that are assigned to the six-membered ring of the indole/indazole and the six-membered phenyl/naphthyl ring directly bound to the linker atom, respectively. For most active sulfonyl derivatives, the Cg1...Cg2 distance remains in the range from 5.1-5.5 Å.

ANG PLN1-PLN2 is the angle between mean planes defined by atoms of both aromatic fragments (PLN1 and PLN2 for indole/indazole and phenyl/naphthyl rings, respectively), which is approximately in the perpendicular orientation (angle in the range from 80°-90°) for ligands with the best affinities among the presented compounds.

The torsion angle C12-C11-S10(or C10)-N1 for most active derivatives stays in the range from 70°-100°.

α is the angular deviation of the S10 or C10 atoms from the mean planes assigned for both aromatic systems (PLN1 and PLN2). ANG N1-S10(C10) /PLN1 is the external angle between the bond and the mean plane.

All parameters are presented in **Table 2.6** (bolded values are in ranges predicted for the best 5HT₆R ligands).

Table 2.6 Selected geometric parameters defined for the presented structure with a plausible strong impact on the ligand-receptor affinity

No.	A	<i>d</i> Cg1...Cg2 [Å]	ANG PLN1-PLN2 [°]	N1-S10(C10)- C11-C12*) [°]	<i>d</i> (S10 or C10) to PLN1 [Å]	<i>d</i> (S10 or C10) to PLN2 [Å]	ANG N1-S10(C10) /PLN1 [°]	Angular deviation of S10 or C10 of the PLN1 [°]
1a	-SO ₂ -	5.321	83.09(6)	85.5(2)	0.577(1)	0.013(1)	159.8	20.2
1b	-CO-	6.452	54.18(5)	45.5(3)	0.084(2)	0.146(2)	176.6	3.4
1c	-CH ₂ -	5.342	72.55(15)	110.7(1)	0.026(9)	0.034(9)	179.0	1.0
2a	-SO ₂ -	5.448	84.41(5)	74.8(1)	0.280(1)	0.064(1)	170.2	9.8
2b	-CO-	6.540	53.59(7)	139.0(2)	0.103(2)	0.098(2)	175.7	4.3
2c	-CH ₂ -	6.010	74.51(6)	67.9(2)	0.138(2)	0.008(2)	174.6	5.4
3a	-SO ₂ -	5.225	89.25(4)	74.4(1)	0.272(1)	0.027 (1)	170.5	9.5
3c	-CH ₂ -	5.889	83.11(4)	35.8(2)	0.069(2)	0.022(2)	177.3	2.7
3c-A	-CH ₂ -	5.163	82.22(5)	35.7(2)	0.161(2)	0.070 (2)	173.6	6.4
4a	-SO ₂ -	5.210	88.92(7)	93.6(2)	0.322(1)	0.061(1)	168.8	11.2
4a-A	-SO ₂ -	5.343	85.23(6)	71.9(6)	0.272(3)	0.038(3)	170.8	9.2
4b	-CO-	6.487	56.43(11)	43.7(5)	0.110 (4)	0.106(4)	175.5	4.5
4b-A	-CO-	6.496	56.03(10)	48.2(5)	0.038(4)	0.079(4)	178.4	1.6
4b-B	-CO-	6.485	72.98(12)	50.4(5)	0.367(4)	0.107(4)	164.9	15.1
4b-C	-CO-	6.474	68.27(11)	122.1(4)	0.238(4)	0.023(4)	170.2	9.8
4c	-CH ₂ -	5.412	86.80(4)	22.8(3)	0.099(2)	0.057(2)	176.1	3.9

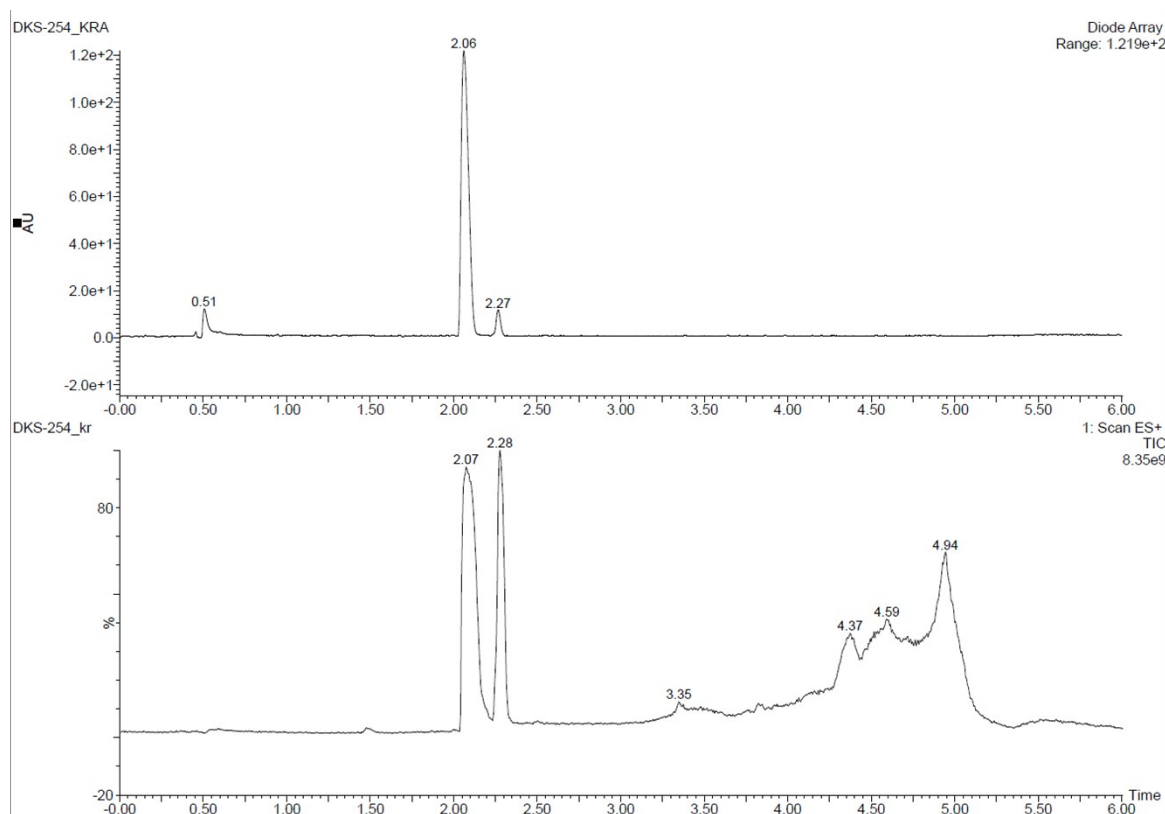
d - distance; ANG - angle; Cg1 and Cg2 - centroids of the six-membered ring of the indole/indazole and phenyl/naphthyl ring directly bound to sulfonyl linker, respectively; PLN1 and PLN2 -mean plane of the indole (indazole) moiety and the phenyl ring, respectively;

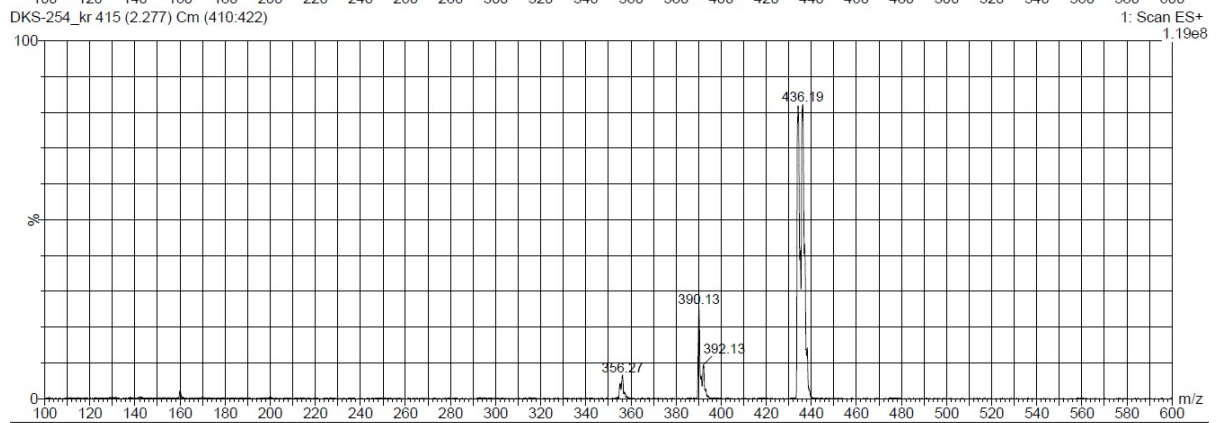
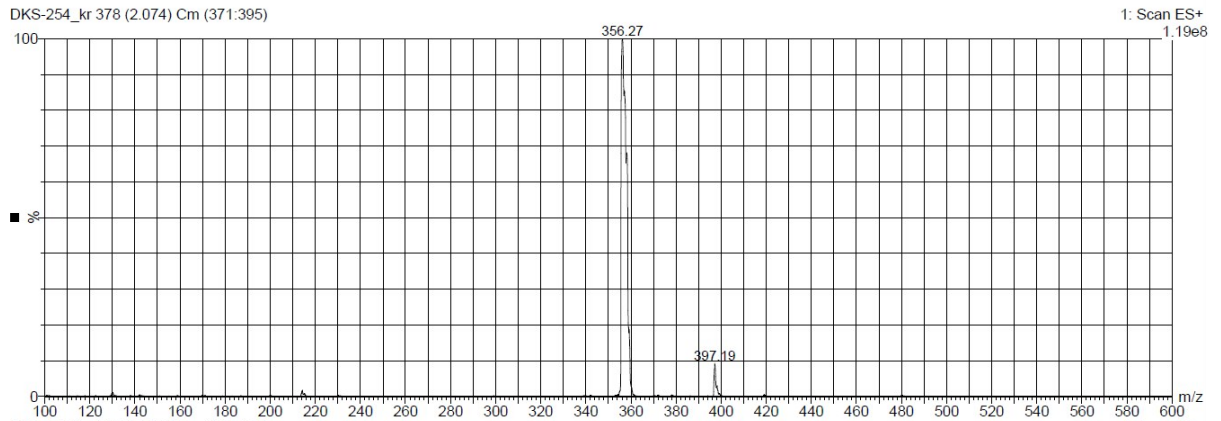
*) the numerical value of the torsion angle was kept in range (0;π);

S3. Mass spectroscopy data for the dissolved crystal 3a - confirmation of bromination at the C3 position of the indole ring

Mass spectra was recorded using a TQD Waters LC/MS spectrometer with the electrospray ionization method. Substrates and solvents were purchased from Sigma-Aldrich and the Apollo Scientific Company and used without further purification.

LC-MS analysis of the crystals of compound **4a** (working symbol DKS-254) revealed that it contain small amounts of a monobrominated derivative with a retention time 2.27 min (two peaks in the MS spectrum: 434.19 and 436.19).





S4. Cambridge Crystallographic Database statistical analysis

The statistical approach was applied to search for the conformational preferences of indole derivatives with N-arylsulfonyl and N-benzyl substituents. The search was performed using the Cambridge Structural Database (CSD Version 5.37 (November 2015)¹²) with the ConQuest 1.18 program.¹³ Searches were performed only for organic compounds with the R factor set as ≤ 0.05 . The search resulted in 265 and 129 structures for sulfonyl and methylene linkers, respectively. From the group of structures with a methylene linker, four were excluded (refcodes: GATEQ, TEJMIF, UJALOG and UJALUM) because they are macrocycles with a more rigid conformation of the benzene ring with respect to the rest of the molecule. H atoms are fixed as bound to carbon atoms C12-16 to retain hydrogen bond donor properties. An additional search with same parameter was performed for structures with H atoms fixed as bound to C2 and C8 atoms of the indole moiety to avoid steric hindrance effect. For structures with sulfonyl and methylene linkers, 27 and 37 structures were found, respectively. Among the chosen 27 structures of sulfonyl derivatives, three were silica derivatives and were excluded due to additional effects influencing the geometry of the molecule (refcodes: ATOYAJ, BUYVUY and QASXAK).

S.4.1 Distribution of the Cg1...Cg2 distances and PLN1-PLN2 angles for N-arylsulfonyl and N-benzyl deposited within the CSD

For all structures selected for the statistical study, three geometric parameters were selected:

- The torsion angle corresponding to C12-C11-S10(or C10)-N1 (histogram in the main manuscript, **Figure 4**)
- PLN1-PLN2 angles (**Figures S4.1.1** and **S4.1.3**)
- Cg1...Cg2 distances (**Figures S4.1.2** and **S4.1.4**)

The statistical distribution of Cg1...Cg2 distances and PLN1-PLN2 angles did not indicate differentiation between derivatives with sulfonyl and methylene linker. Only the torsion angle showed the conformational preferences and characterized the perpendicular orientation of the aromatic moieties of compounds with sulfonyl linkers.

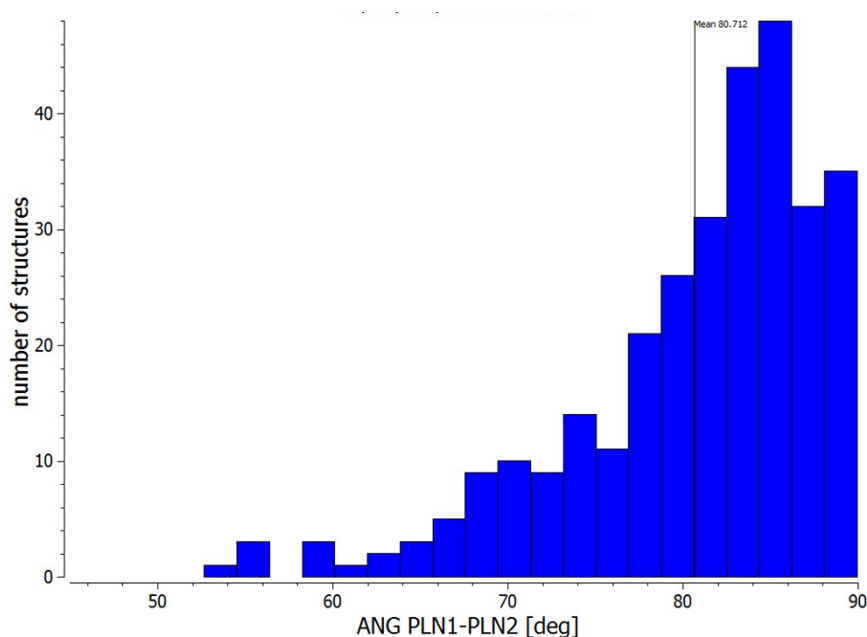


Figure 4.1.1 Distribution of angles between mean planes of both aromatic systems calculated for N-arylsulfonyl indole derivatives deposited in the CSD.

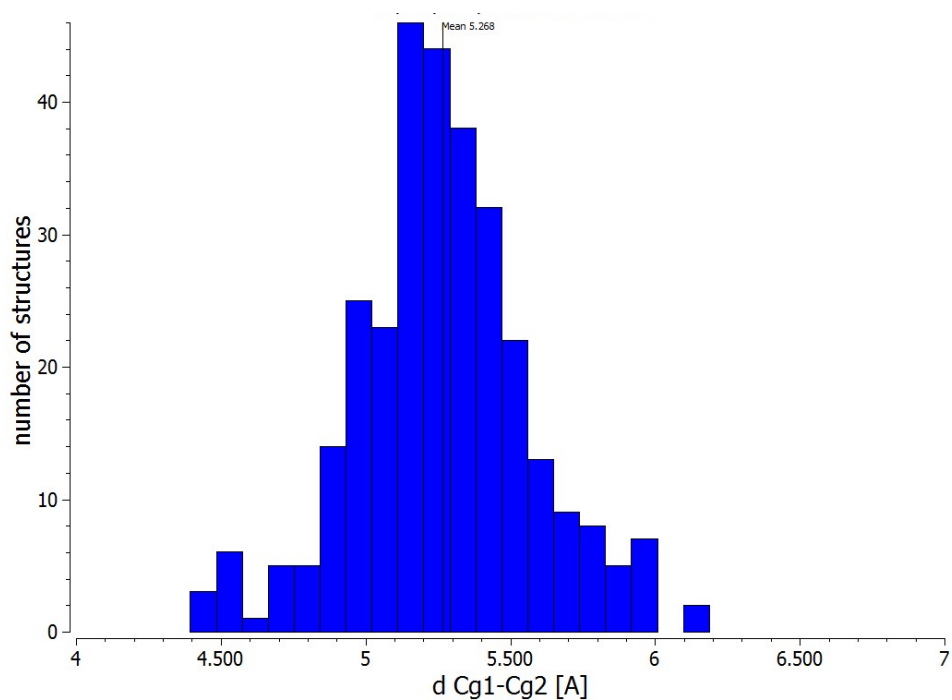


Figure 4.1.2 Distribution of centroid's distances, defined (as described in Section S2.6) for both aromatic systems of N-arylsulfonyl indole derivatives deposited in the CSD.

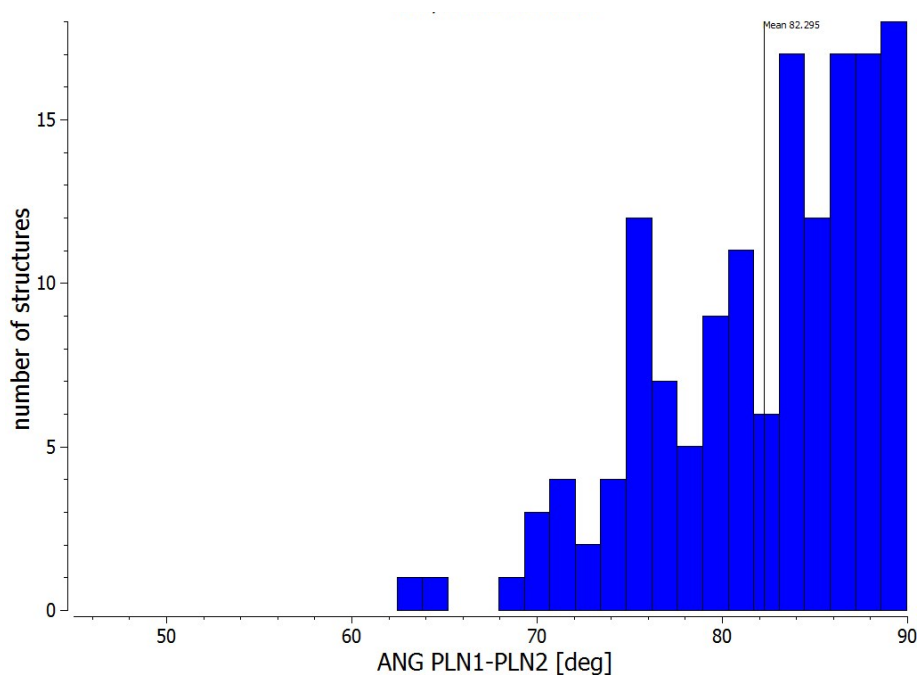


Figure 4.1.3 Distribution of angles between mean planes of both aromatic systems calculated for N-benzyl indole derivatives deposited in the CSD.

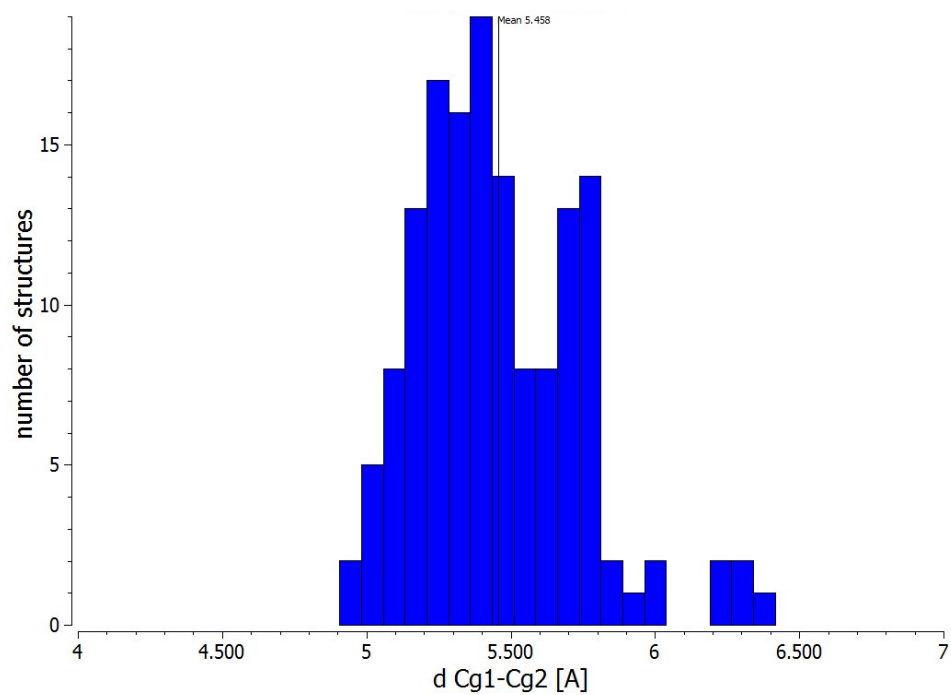


Figure 4.1.4 Distribution of the centroid's distances, defined (as described in Section S2.6) for both aromatic systems of N-benzyl indole derivatives deposited in the CSD.

S.4.2 Distribution of the torsion angle corresponding to C12-C11-S10(or C10)-N1 angles for benzenesulfonyl benzene and benzylbenzene derivatives deposited within the CSD

To determine whether the observed conformational preferences are a more general rule and characteristic for all bis-aromatic compounds with the sulfonyl linker, an additional search was performed. Derivatives of benzenesulfonyl benzene (98 structures, excluding macrocycles in number 10) and benzylbenzene (126 structures, excluding macrocycles in number 29) were recognized. During the search, the carbon atoms in proximity to the linkers had permanently fixed hydrogen positions to retain hydrogen bond donor properties.

Histograms presenting the statistical distribution of the torsion angle corresponding to C12-C11-S10(or C10)-C1 are shown in **Figure 4.2.1**. The obtained data confirm that the sulfonyl group serves as a linker to impact the observed conformational preferences of bis-aromatic compounds.

Refcodes for the excluded macrocycles with a benzenesulfonyl benzene fragment are as follows: BALVOA, BALVOA01, CUMHUO, LAYGEY, LENKUK, MIKVEI, PAXTEN, PAXTEN01, RUYXIS, RUYXIS01

Refcodes for the excluded macrocycles with a benzylbenzene fragment are as follows: ACOHOP, AVILUM, AVIMEX, EBUQUN, EBUREY, EBURIC, ENEJEN, NISSOX, OGAYEB, QIYFUZ, RAYKAE, RIWLIS, SAQNIH, SAQNON, SAQPAB, SOPXUQ, TIVREV, TIVREV01, WEJJUQ, WICMUQ, WICMUQ01, WIGXUG, WUFPUI, WUFPUI01, WUFQAP, WUFQAP01, XARCID, XARCID01, ZEJHUR

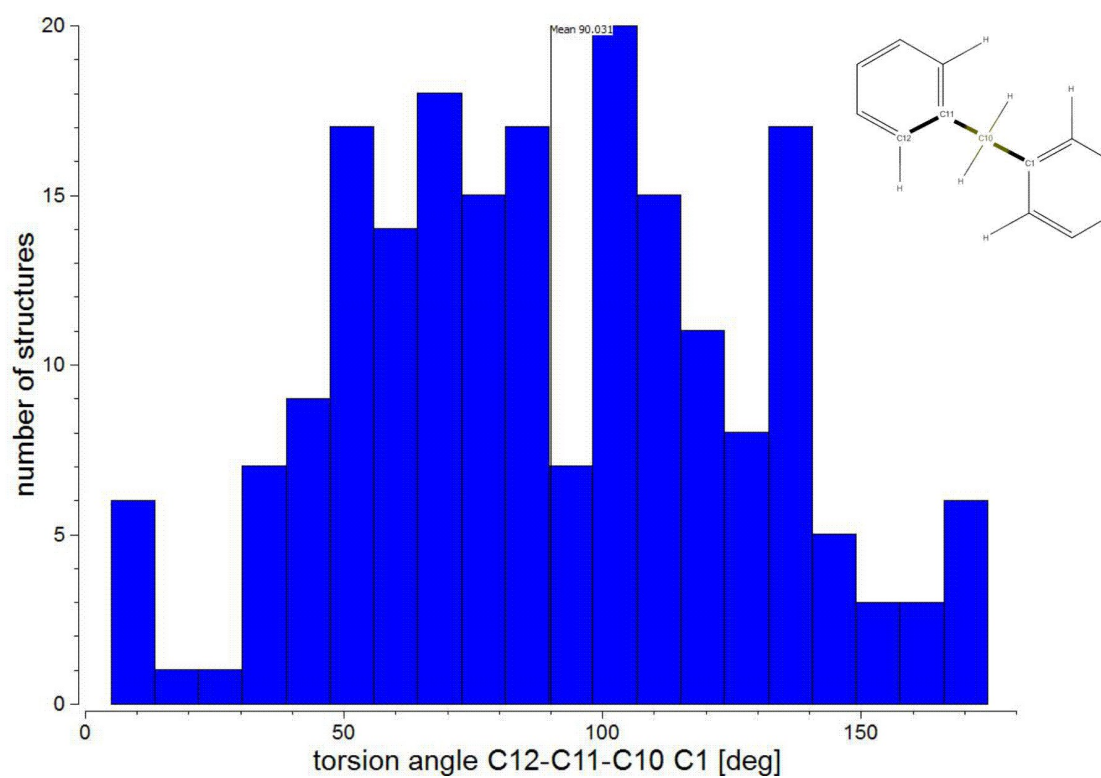
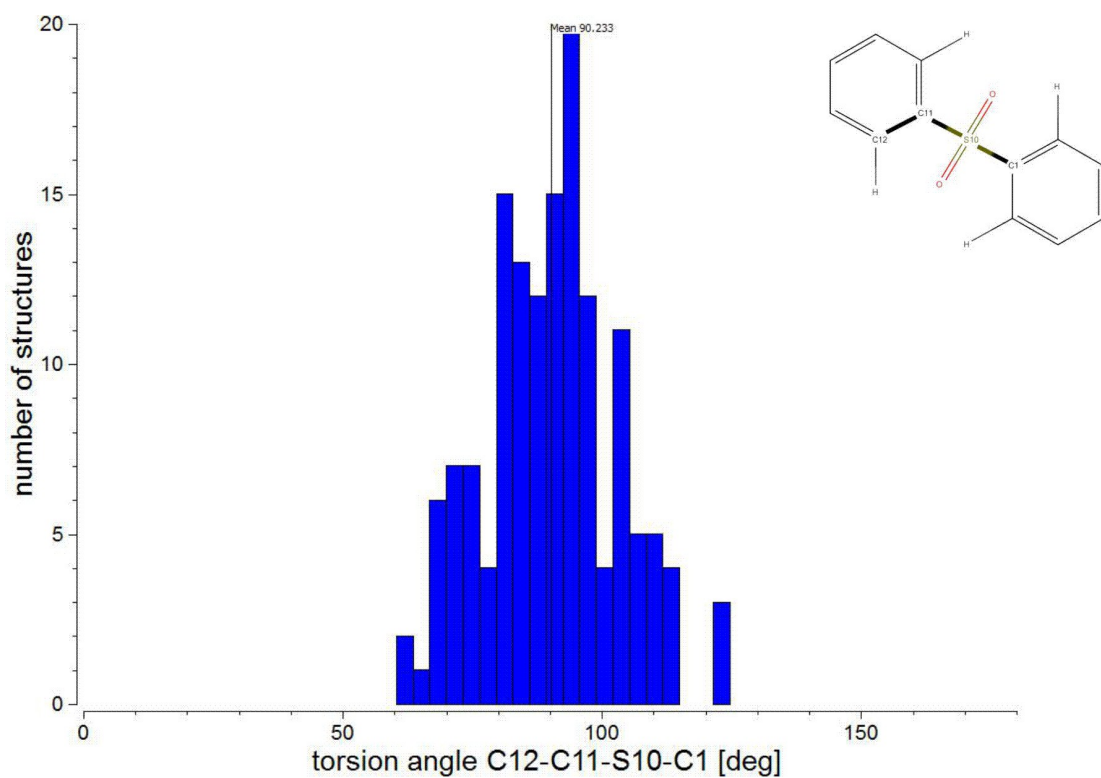


Figure 4.2.1 Distribution of the torsion angle corresponding to C12-C11-S10(or C10)-N1 angles for benzenesulfonyl benzene (top) and benzylbenzene (below) derivatives deposited in the CSD.

S5. Theoretical calculations

The AIMAll package¹⁴ was used to identify intramolecular interactions within the studied systems. To achieve this goal DFT calculations were performed for isolated molecules with the GAUSSIAN09 package at the B3LYP/6-311++G(2d,2p) level. Geometries were obtained from the X-ray diffraction data and kept frozen for calculations. Wave functions obtained in this manner were used to perform Bader's quantum theory of atoms (QTAIM)¹⁵ partitioning.

Analysis of the topology of the electron density $\rho(r)$ for the examined systems revealed characteristic C8-H8...O11 weak intramolecular interactions with non-zero charge densities and positive values of the Laplacian of electron density $\nabla^2\rho(r)$ at all bond critical points (BCPs) (**Figure 5.1**).

The relationship between the local kinetic energy density $G(r_{CP})$ and distance R_{ij} as well as the potential energy densities $V(r_{CP})$ and R_{ij} show exponential dependence¹⁶. Energetic criteria based on analyses of local potential and kinetic energy densities show that the hydrogen bonds in all molecules (**Table 5.1**) exhibit a closed shell character ($|V(r_{CP})|/G(r_{CP}) < 1$, $E(r_{CP})/\rho(r_{CP}) > 0$). In compound **1a**, an additional weak hydrogen bond between atom O10 of the sulfonyl group and the C19-H19 group of the side phenyl ring was recognized. The electron density and Laplacian values indicate that this hydrogen bond is relatively stronger in comparison to the others, potentially suggesting that it is responsible for the stabilization of the perpendicular orientation of the naphthyl group with respect to the rest of the molecule. As expected, it can be classified as a pure closed shell interaction (see **Table 5.1**).

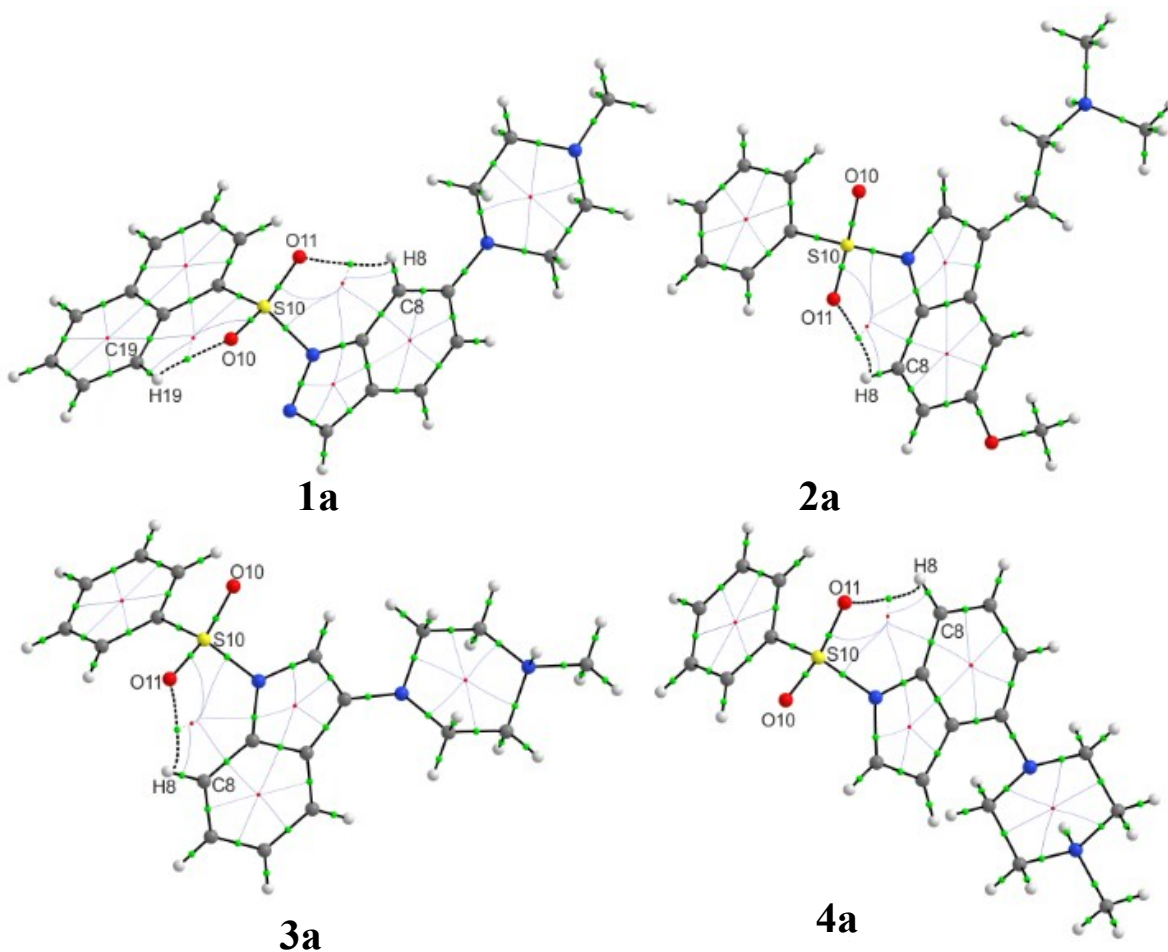


Figure 5.1 Molecular graphs of isolated molecules of 1a, 2a, 3a and 4a, showing the bond paths and bond critical points (BCP) in the studied systems. Small green and red spheres indicate the (3,-1) bond critical points and (3,+1) ring critical point in $\rho(r)$, respectively. Dashed bonds show weak intramolecular interactions of the C-H...O type. The presence of BCPs denotes the presence of an interaction between atoms.

The hydrogen bond energies derived from the proportionality $E_{\text{HB}} = \frac{1}{2} \cdot V(\text{rCP})^{16}$ and converted to kcal/mol units show that the energy of the C8-H8...O11 interaction varies in the range from -1.9 to -2,5 kcal/mol, whereas the energy for C19-H19...O10 is -3.5 kcal/mol (**Table 5.1**). This observation might suggest that the latter interaction is crucial for the perpendicular alignment of both aromatic systems.

Table 5.1 Topological analysis of intramolecular interactions in CP (3,-1)^a obtained for isolated molecules of compounds 1a, 2a, 3a and 4a.

		$\rho(r)$	$\nabla^2\rho(r)$	R_{ij} [Å]	d_1 [Å]	d_2 [Å]	λ_1	λ_2	λ_3	ε	$G(r_{CP})$	$V(r_{CP})$	E_{HB}	$E(r_{CP})$	$ V(r_{CP}) /G(r_{CP})$	$E(r_{CP})/\rho(r)$
1a	C8-H8...O11	0.011	0.044	2.545	1.114	1.431	-0.010	-0.007	0.061	0.47	0.009	-0.008	-2.510	0.002	0.89	0.18
2a	C8-H8...O11	0.009	0.036	2.718	1.509	1.209	-0.007	-0.004	0.047	0.81	0.007	-0.006	-1.883	0.002	0.86	0.22
3a	C8-H8...O11	0.010	0.038	2.677	1.497	1.180	-0.008	-0.005	0.051	0.69	0.008	-0.007	-2.196	0.002	0.88	0.20
4a	C8-H8...O11	0.010	0.037	2.675	1.175	1.500	-0.008	-0.005	0.050	0.67	0.008	-0.006	-1.883	0.002	0.75	0.20
1a	C19-H19...O10	0.016	0.067	2.287	0.950	1.337	-0.016	-0.014	0.097	0.18	0.014	-0.011	-3.451	0.003	0.79	0.19

^a $\rho(r)/e\text{Bohr}^{-3}$, $\nabla^2\rho(r)/e\text{Bohr}^{-5}$, d_1, d_2 – distance between BCP and atoms 1 and 2, respectively, ε – ellipticity, $G(r_{CP})$ and $V(r_{CP})$ are local kinetic and local potential energy density (hartree), respectively, $E_{HB} = \frac{1}{2} \cdot V(r_{CP})^{16}$ and converted into kcal/mol units, $E(r_{CP})$ is electronic local energy density (hartree)

Since QTAIM analysis does not recognize all expected weak non-covalent interactions, especially those between the phenyl or naphthalene group attached to S10 and the rest of the molecule, the NCI (non-covalent interaction) approach^{17,18} was used. The NCI analysis is based on the so called reduced electron density gradient (RDG) defined as:

$$s(r) = \frac{|\nabla\rho(r)|}{2(3\pi)^{1/3}\rho(r)^{4/3}}$$

where $\nabla\rho(r)$ is a gradient of electron density. This analysis enables visualization of regions in space involved in either attractive or repulsive interactions. If a non-covalent contact is present in the studied system, the characteristic spikes on scatterplots of $s(r)$ against $\rho(r)$ occur in low-gradient and low-density regions, which are absent when only covalent bonds are observed. Moreover, taking into account the sign of the second eigenvalue (λ_2) of the Hessian matrix of electron density provides information concerning whether the identified non-covalent interaction is stabilizing ($\lambda_2 < 0$) or destabilizing ($\lambda_2 > 0$). Therefore, the presence of a spike in the low-gradient, low-density area at a negative λ_2 indicates a stabilizing interaction such as a hydrogen bond. A smaller spike and slightly negative λ_2 indicates a weakly stabilizing interaction, and a spike that is associated with a positive λ_2 denotes the absence of non-covalent interaction. In many cases, if the peak representing NCI does

not reach $s(r)$ equal to zero, it is not associated with a critical point and therefore QTAIM is blind to this contact.

Using NCIPLOT¹⁹, we examined the isosurface of $s(r)$ (**Figure 5.2**) and plots of $s(r)$ versus $\rho(r)$ multiplied by the sign of λ_2 (**Figure 5.3**) for the studied systems. These plots were generated by evaluating the B3LYP 6-311G** density and reduced gradient on cuboid grids with a step size of 0.1 a.u. All the studied systems showed evidence of weak attractive non-covalent interactions between the indole or indazole rings and O11 atoms, which were absent when no sulfonyl group was present in the molecule (see **Figures 5.4** (QTAIM) and **5.5** (NCI) for compound **4c**). Moreover, the presence of the above interactions resulted in the identification of additional RDG domains corresponding to repulsive NCI with $\lambda_2 > 0$ [$\text{sign}(\lambda_2)\rho(r)$ above 0.07 a.u.] between the phenyl/naphthalene ring and indole/azaindole part of the molecule.

Furthermore, although the QTAIM analysis revealed the additional C19-H19...O10 hydrogen bond only in case of **1a**, from isosurfaces around O10 in **Figure 5.2**, it is clear that this kind of interaction is also present in other systems. In fact, the domain is smaller in **2a** and **3a** indicating more directional non-covalent interactions. It is worth mentioning here that NCI surfaces corresponding to **1a** and **4a** differ from these of **2a** and **3a** due to rotation of the indole/indazole along the N-S bond.

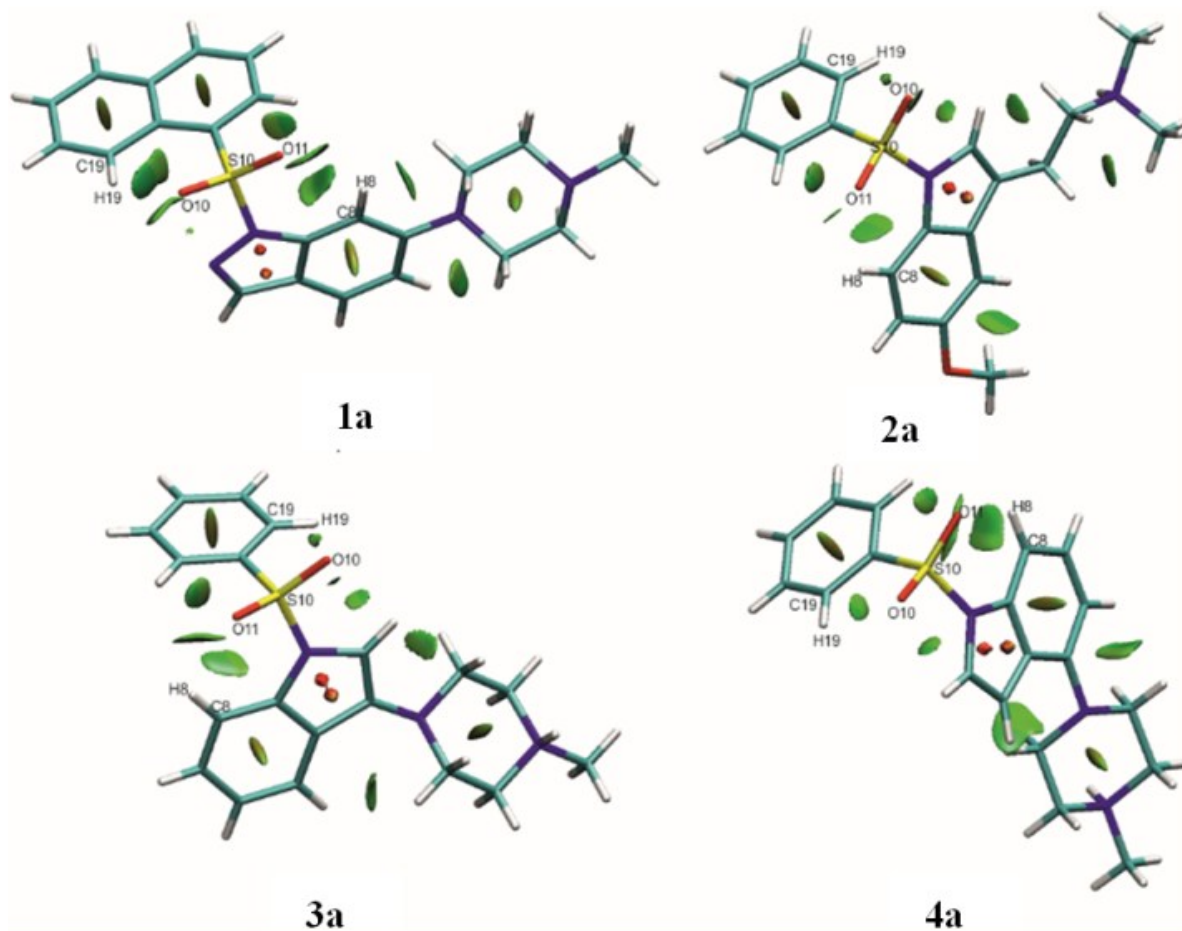


Figure 5.2 Gradient isosurfaces (RDG) representing intramolecular interactions in sulfonyl derivatives. The shape of domains is correlated to the strength of the interaction. Broad multiform domains indicate weak attractive or repulsive interactions, whereas small domains correspond to the strongest hydrogen bonds identified in the studied compounds. Delocalized electrons of aromatic groups are represented by the egg-shape domains. Gradient surfaces correspond to $s = 0.1$ a.u.

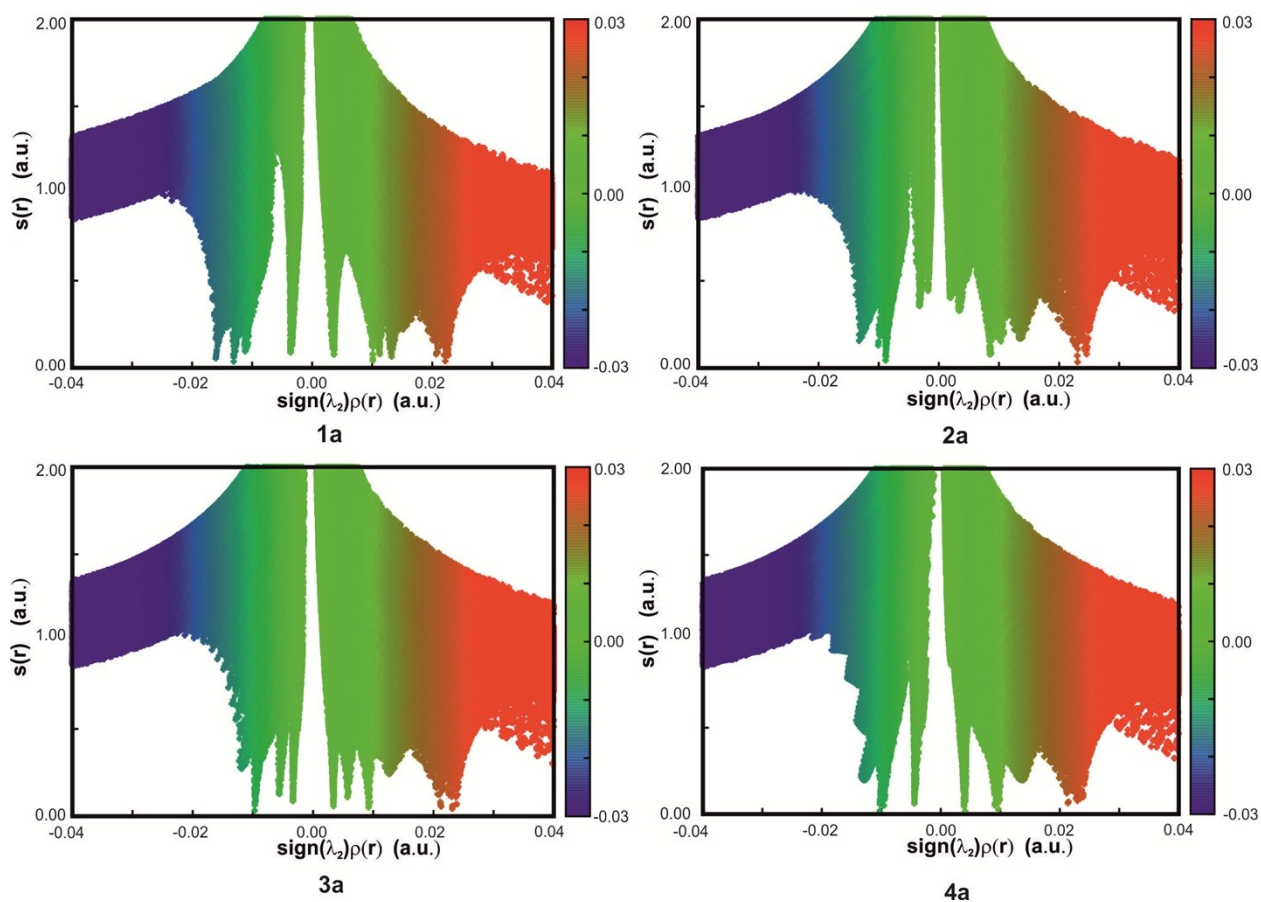


Figure 5.3 Plots of the reduced density gradient (RDG) versus the electron density multiplied by the sign of the second Hessian eigenvalue for isolated molecules of 1a, 2a, 3a and 4a. Blue regions, associated with a negative value of the second eigenvalue (λ_2) of the Hessian matrix of electron density, confirm the presence of intramolecular hydrogen bonds. Green spikes represent van der Waals interactions, whereas the red area is associated with positive values of λ_2 and confirm the existence of repulsive interactions within the studied molecules. The color scale correspond to $-0.04 < \rho < 0.04$ au.

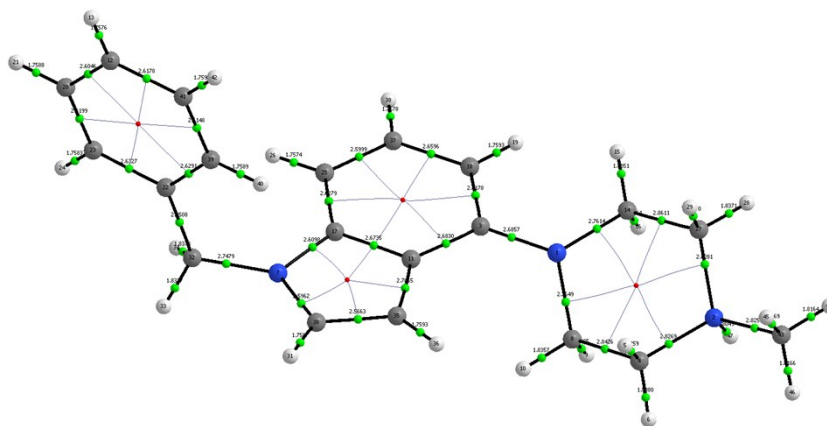


Figure 5.4 Molecular graph of the isolated molecule 4c (as an example), showing bond critical points (BCP) in the studied system. Green and red small spheres indicate the (3,-1) bond critical points and (3,+1) ring critical point in $\rho(r)$, respectively.

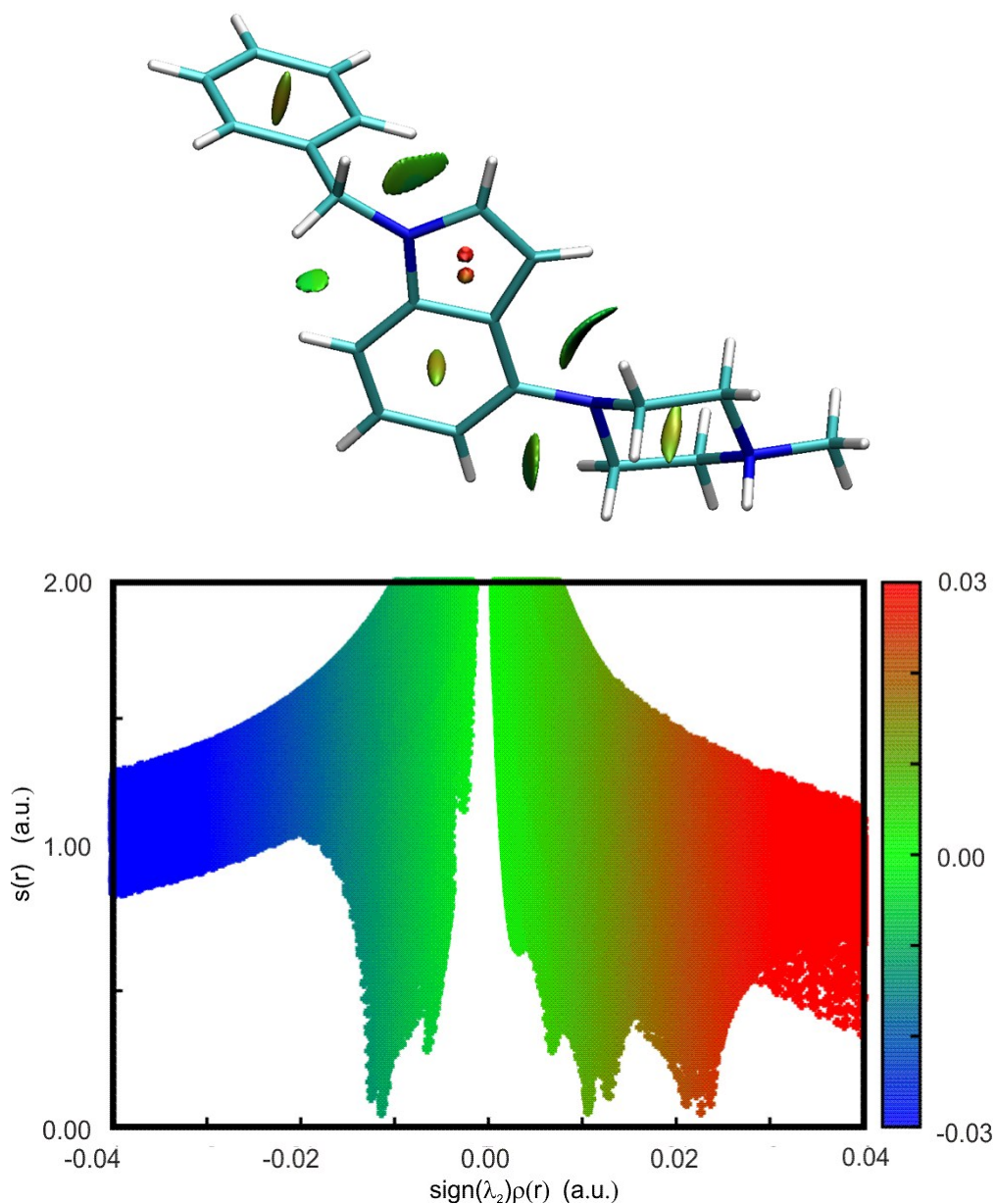


Figure 5.5 Gradient isosurfaces of $s(r)$ (0.1 a.u.) (top) and plots of the reduced density gradient versus the electron density multiplied by the sign of the second Hessian eigenvalue (below) for compound 4c.

Only in the case of compound **1c** an additional intramolecular C19-H19...N2 hydrogen bond has been recognized, as a consequence of the acceptor properties of the indazole nitrogen atom N2.

Figure 5.6, Table 5.2 (QTAIM) and Figure 5.7 (NCI) show the graphical and numerical interpretation of this intramoleculr interaction observed in **1c**.

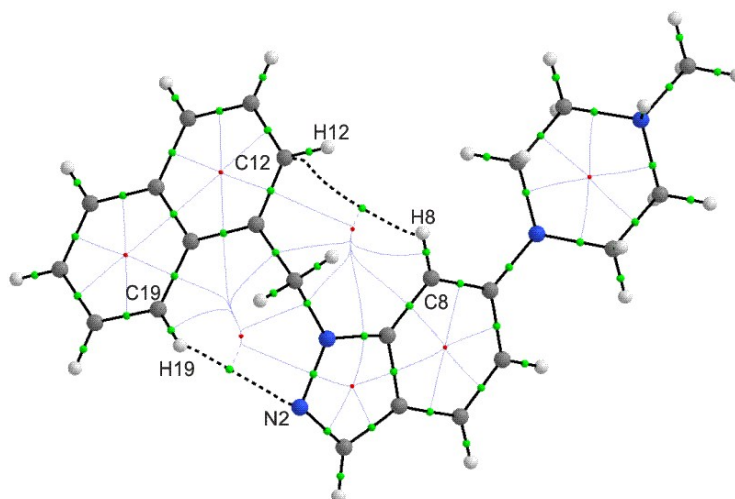


Figure 5.6 Molecular graphs of the isolated molecule 1c showing the bond paths and bond critical points (BCP). Small green and red spheres indicate the (3,-1) bond critical points and (3,+1) ring critical point in $\rho(r)$, respectively. Dashed bonds show weak intramolecular interactions of C19-H19...N2 with BCP between interacting atoms.

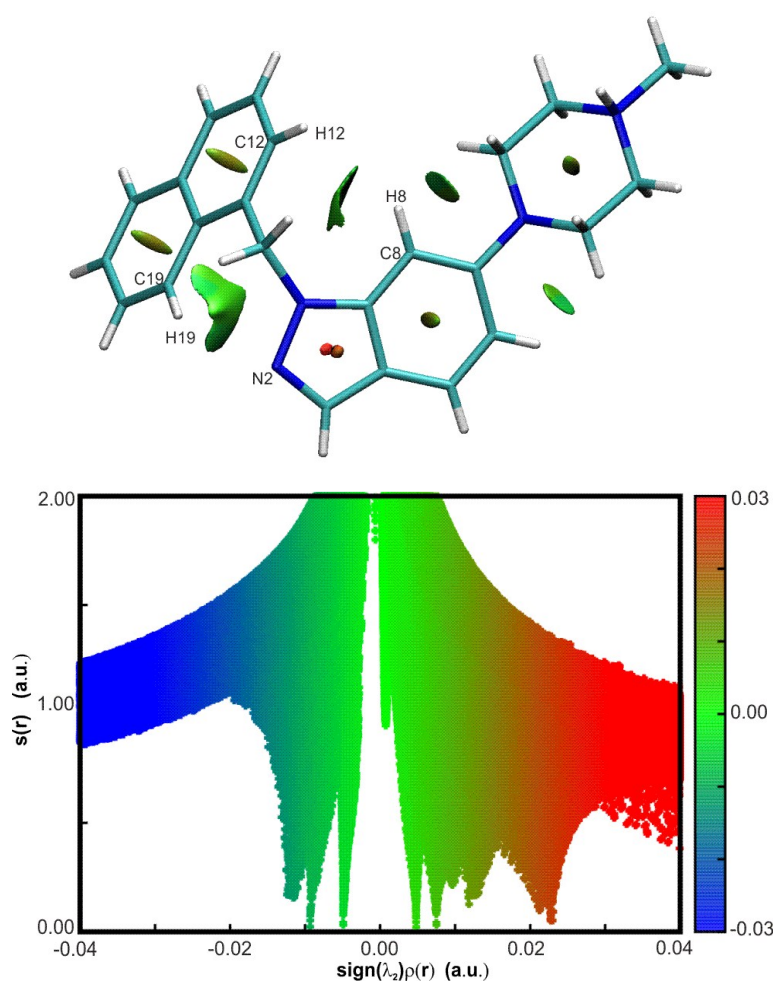


Figure 5.7 Gradient isosurfaces of $s(r)$ (0.1 a.u.) (top) indicating broad surface between H19 and N2 atoms that suggests a disperse interaction; plot of the reduced density gradient versus the electron density multiplied by the sign of the second Hessian eigenvalue (below) for compound 1c.

Table 5.2 Topological analysis of the intramolecular interaction in CP (3,-1)^a obtained for isolated molecule of 1c

	$\rho(r)$	$\nabla^2\rho(r)$	R_{ij} [Å]	d_1 [Å]	d_2 [Å]	λ_1	λ_2	λ_3	ε	$G(r_{CP})$	$V(r_{CP})$	E_{HB}	$E(r_{CP})$	$ V(r_{CP}) /G(r_{CP})$	$E(r_{CP})/\rho(r)$
C19-H19...N2	0.009	0.032	2.561	1.051	1.510	-0.008	-0.007	0.047	0.2	0.007	-0.005	-1.569	0.002	0.71	0.22

The unsubstituted 1-phenylsulfonyl-indole structure (CCDC ID: DUPTEN)²⁰ was also analyzed using the QTAIM approach in **Figure 5.8**. To identify spikes, an $s(r)$ vs. $\text{sign}(\lambda_2)\rho(r)$ plot was generated for DUPTEN (**Figure 5.9**) and analyzed.

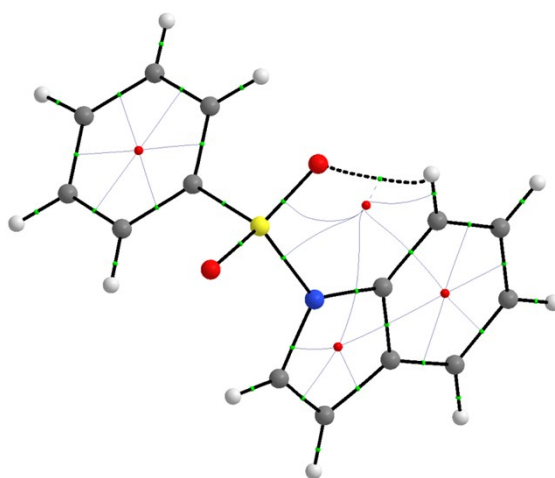


Figure 5.6 Molecular graphs of isolated molecule of DUPTEN showing bond paths and bond critical points (BCP). Small green and red spheres indicate the (3,-1) bond critical points and (3,+1) ring critical point in $\rho(r)$, respectively. Dashed bonds show the weak intramolecular interactions of C-H...O.

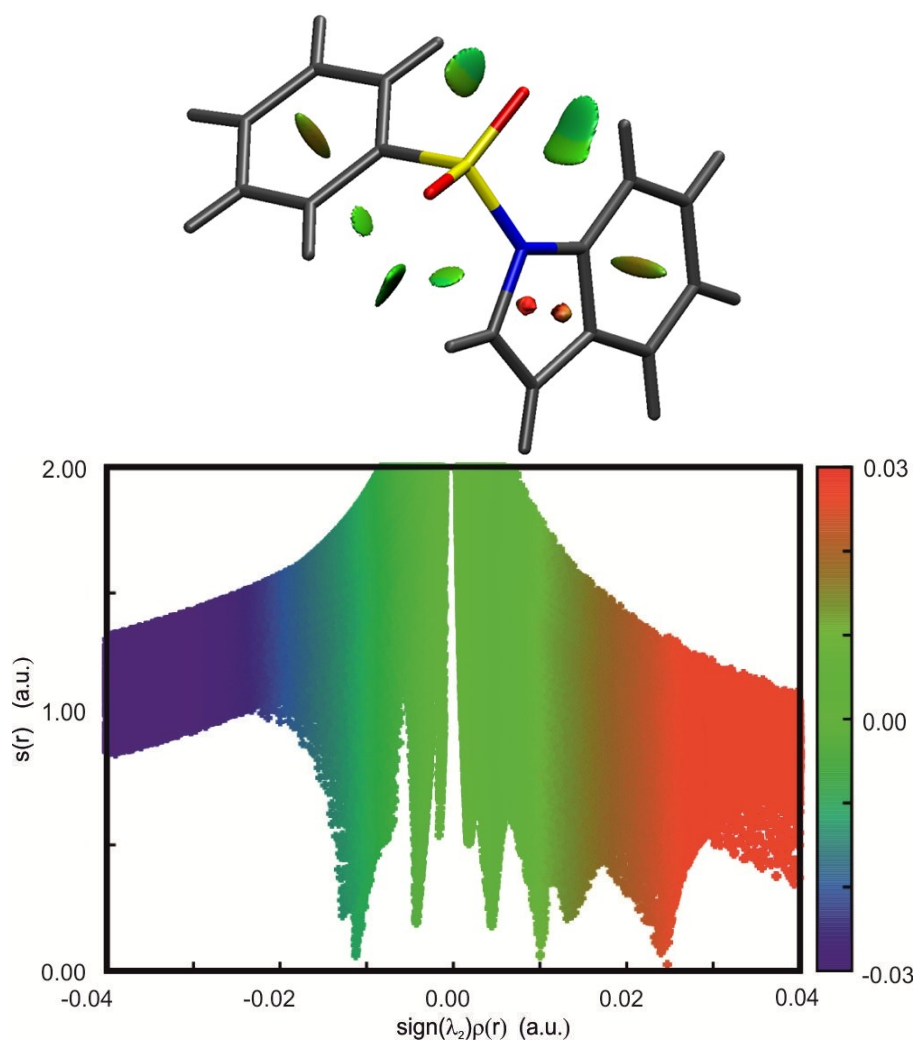
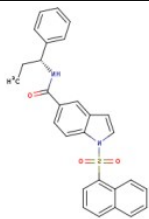
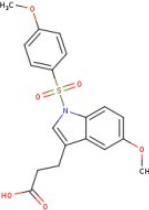


Figure 5.9 Gradient isosurfaces of $s(r)$ (0.1 a.u.) (top) calculated for DUPTEN and plots of the reduced density gradient versus the electron density multiplied by the sign of the second Hessian eigenvalue (below).

S6. Protein Data Bank analysis

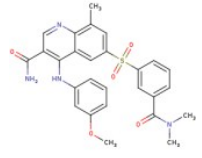
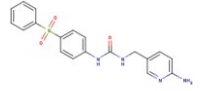
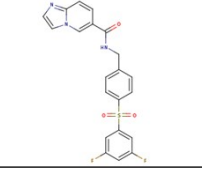
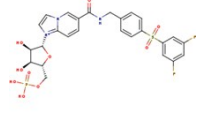
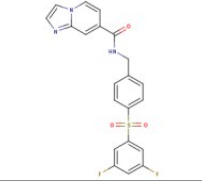
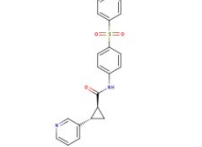
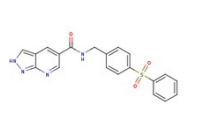
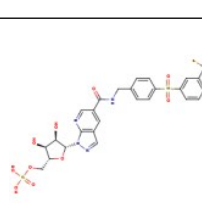
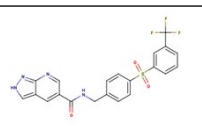
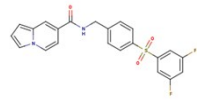
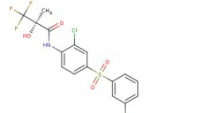
The Protein Data Bank²¹ (www.rcsb.org) was searched for ligands with the option of chemical substructure similarity. Two searches were performed to identify bis-aromatic sulfonyl derivatives in complexes with different macromolecular targets. Ligands containing arylsulfonyl indole (3 ligands in 4 structures) or diphenylsulfone (22 ligands in 25 structures) substructures were selected for further investigation. In the first set, a ligand with ID JCB was excluded due to an unavailable PDB file, reducing the number to 2 ligands in 4 structures (**Table 6.1**). In the second data set, ligand 6N0 (PDB ID: 5AKE) was excluded due to cyclization of the small molecule, introducing rigidity that affects the investigated torsion angle. Consequently, the number of selected ligands was reduced to 21 in 24 macromolecular structures (**Table 6.2**)

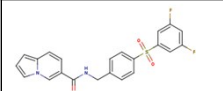
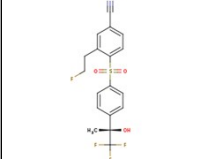
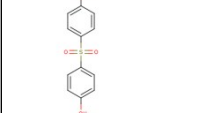
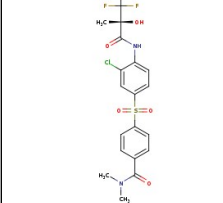
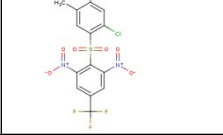
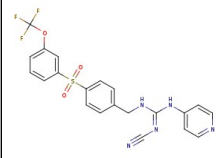
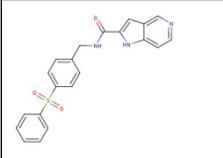
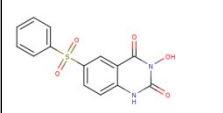
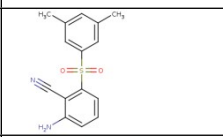
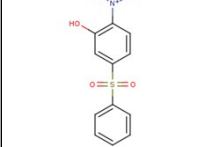
Table 6.1 Results of the PDB search for ligands containing arylsulfonyl indole fragment

Ligand ID	Ligand Formula	Structure	MW	Torsion Angle [°]	Ligand Name	PDB IDs
3E7	C ₂₈ H ₂₄ N ₂ O ₃ S		468.57	74.59	1-(naphthalen-1-ylsulfonyl)-N-[(1S)-1-phenylpropyl]-1H-indole-5-carboxamide	4R06
ET1	C ₁₉ H ₁₉ N O ₆ S		389.42	73.12 69.19 83.22	3-{5-methoxy-1-[(4-methoxyphenyl)sulfonyl]-1H-indol-3-yl}propanoic acid	3ET1, 3ET2, 3ET3

Torsion Angle corresponds to C12-C11-S10-N1, presented as the absolute value.

Table 6.2 Results of the PDB search for ligands containing diphenylsulfonyl fragment

Ligand ID	Ligand Formula	Structure	MW	Torsion Angle [°]	Ligand Name	PDB IDs
066	C ₂₇ H ₂₆ N ₄ O ₅ S		518.58	76.60	6-[[3-(dimethylcarbamoyl)phenyl]sulfonyl]-4-[[3-methoxyphenyl]amino]-8-methylquinoline-3-carboxamide	3GWT
1LJ	C ₁₉ H ₁₈ N ₄ O ₃ S		382.44	62.60	1-[[6-aminopyridin-3-yl]methyl]-3-[4-(phenylsulfonyl)phenyl]urea	4JNM
1QS	C ₂₁ H ₁₅ F ₂ N ₃ O ₃ S		427.42	65.66 64.15	n-[4-[(3,5-difluorophenyl)sulfonyl]benzyl]imidazo[1,2-a]pyridine-6-carboxamide	4KFO, 4O28
1XC	C ₂₆ H ₂₅ F ₂ N ₃ O ₁₀ P S		640.53	69.09 70.18	6-[[4-[(3,5-difluorophenyl)sulfonyl]benzyl]carbamoyl]-1-(5-o-phosphono-beta-d-ribofuranosyl)imidazo[1,2-a]pyridin-1-ium	4L4L, 4O16
1XD	C ₂₁ H ₁₅ F ₂ N ₃ O ₃ S		427.42	69.41	n-[4-[(3,5-difluorophenyl)sulfonyl]benzyl]imidazo[1,2-a]pyridine-7-carboxamide	4L4M
200	C ₂₁ H ₁₈ N ₂ O ₃ S		378.44	91.72	(1s,2s)-n-[4-(phenylsulfonyl)phenyl]-2-(pyridin-3-yl)cyclopropanecarboxamide	4LVG
20R	C ₂₀ H ₁₆ N ₄ O ₃ S		392.43	70.59	n-[4-(phenylsulfonyl)benzyl]-2h-pyrazolo[3,4-b]pyridine-5-carboxamide	4M6P
20T	C ₂₆ H ₂₄ F ₃ N ₄ O ₁₀ P S		672.52	64.62	1-(5-o-phosphono-beta-d-ribofuranosyl)-n-(4-[[3-(trifluoromethyl)phenyl]sulfonyl]benzyl)-1h-pyrazolo[3,4-b]pyridine-5-carboxamide	4M6Q
2P1	C ₂₁ H ₁₅ F ₃ N ₄ O ₃ S		460.43	64.76 66.25	n-[4-[[3-(trifluoromethyl)phenyl]sulfonyl]benzyl]-2h-pyrazolo[3,4-b]pyridine-5-carboxamide	4O13, 4O15
2QF	C ₂₂ H ₁₆ F ₂ N ₂ O ₃ S		426.44	67.72	n-[4-[(3,5-difluorophenyl)sulfonyl]benzyl]indolizine-7-carboxamide	4O10
2QJ	C ₁₆ H ₁₂ Br Cl F ₃ N O ₄ S		486.69	45.05	(2r)-n-[4-[(3-bromophenyl)sulfonyl]-2-chlorophenyl]-3,3,3-trifluoro-2-hydroxy-2-methylpropanamide	4NZ2

2RM	C22 H16 F2 N2 O3 S		426.44	70.34	n-[4-[(3,5-difluorophenyl)sulfonyl]benzyl]indolizine-6-carboxamide	4O0Z
3OQ	C18 H15 F4 N O3 S		401.38	73.60	3-(2-fluoroethyl)-4-[(4-[(2s)-1,1,1-trifluoro-2-hydroxypropan-2-yl]phenyl)sulfonyl]benzotrile	3OQ1
6JD	C12 H10 O4 S		250.27	64.85	4-(4-hydroxyphenyl)sulfonylphenol	5L4J
AZX	C19 H18 Cl F3 N2 O5 S		478.87	88.44	4-[(3-chloro-4-[(2r)-3,3,3-trifluoro-2-hydroxy-2-methylpropanoyl]amino)phenyl)sulfonyl]-n,n-dimethylbenzamide	2Q8G
D3F	C14 H7 Cl2 F3 N2 O6 S		459.18	22.32	2-[(2,4-dichloro-5-methylphenyl)sulfonyl]-1,3-dinitro-5-(trifluoromethyl)benzene	2GZ7
LTS	C21 H16 F3 N5 O3 S		475.44	79.33	2-cyano-1-pyridin-4-yl-3-(4-[(3-(trifluoromethoxy)phenyl)sulfonyl]benzyl)guanidine	4LTS
LWW	C21 H17 N3 O3 S		391.44	86.40	n-(4-(phenylsulfonyl)benzyl)-1h-pyrrolo[3,2-c]pyridine-2-carboxamide	4LWW
QID	C14 H10 N2 O5 S		318.30	75.97	3-hydroxy-6-(phenylsulfonyl)quinazoline-2,4(1h,3h)-dione	3QIO
SBN	C15 H14 N2 O2 S		286.35	69.82	2-amino-6-(3,5-dimethylphenyl)sulfonylbenzotrile	1JLQ
WDT	C12 H9 N O5 S		279.27	51.99	2-nitro-5-(phenylsulfonyl)phenol	4WDT

Torsion Angle corresponds to C12-C11-S10-C1, presented as the absolute value.

S7. Abbreviations

5-HT – 5-hydroxytryptamine,

K_i – binding constant,

S.D. – standard deviation,

PBS – phosphate buffer saline,

EDTA – ethylenediaminetetraacetic acid,

Tris – *tris*(hydroxymethyl)aminomethane,

8-OH-DPAT – (\pm)-8-Hydroxy-2-(dipropylamino)tetralin,

LSD – lysergic acid diethylamide,

5-CT – 5-carboxamidotryptamine,

S8. References

- 1 C. Yung-Chi and W. H. Prusoff, *Biochemical Pharmacology*, 1973, **22**, 3099–3108.
- 2 K. G. Liu, J. R. Lo, T. A. Comery, G. M. Zhang, J. Y. Zhang, D. M. Kowal, D. L. Smith, L. Di, E. H. Kerns, L. E. Schechter and A. J. Robichaud, *Bioorganic and Medicinal Chemistry Letters*, 2009, **19**, 2413–2415.
- 3 Y. Tsai, M. Dukat, A. Slassi, N. MacLean, L. Demchyshyn, J. E. Savage, B. L. Roth, S. Hufesein, M. Lee and R. A. Glennon, *Bioorganic and Medicinal Chemistry Letters*, 2000, **10**, 2295–2299.
- 4 R. V. S. Nirogi, A. D. Deshpande, R. Kambhampati, R. K. Badange, L. Kota, A. V. Daulatabad, A. K. Shinde, I. Ahmad, V. Kandikere, P. Jayarajan and P. K. Dubey, *Bioorganic and Medicinal Chemistry Letters*, 2011, **21**, 346–349.
- 5 M. Ahmed, M. A. Briggs, S. M. Bromidge, T. Buck, L. Campbell, N. J. Deeks, A. Garner, L. Gordon, D. W. Hamprecht, V. Holland, C. N. Johnson, A. D. Medhurst, D. J. Mitchell, S. F. Moss, J. Powles, J. T. Seal, T. O. Stean, G. Stemp, M. Thompson, B. Trail, N. Upton, K. Winborn and D. R. Witty, *Bioorganic and Medicinal Chemistry Letters*, 2005, **15**, 4867–4871.
- 6 G. M. Sheldrick, *Acta Cryst.*, 2008, **A64**, 112–122.
- 7 M. Nardelli, *J. Appl. Cryst.*, 1999, **32**, 563–571.
- 8 Y. Liu and G. W. Gribble, *Tetrahedron Letters*, 2000, **41**, 8717–8721.
- 9 S. C. Conway and G. W. Gribble, *Heterocycles*, 1990, **30**, 627–633.
- 10 D. A. Davis and G. W. Gribble, *Heterocycles*, 1992, **34**, 1613–1621.
- 11 G. Chłoń-Rzepa, P. Żmudzki, M. Pawłowski, A. Wesołowska, G. Satała, A. J. Bojarski, M. Jabłoński and J. Kalinowska-Tłuścik, *Journal of Molecular Structure*, 2014, **1067**, 243–251.
- 12 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*, 2016, **72**, 171–179.
- 13 I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson and R. Taylor, *Acta Crystallographica Section B: Structural Science*, 2002, **58**, 389–397.
- 14 T. A. Keith, 2013.
- 15 R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Clarendon Press, Oxford, UK, 1990.
- 16 E. Espinosa, E. Molins and C. Lecomte, *Chemical Physics Letters*, 1998, **285**, 170–173.
- 17 E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen and W. Yang, *Journal of the American Chemical Society*, 2010, **132**, 6498–6506.
- 18 J. Contreras-García, W. Yang and E. R. Johnson, *Journal of Physical Chemistry A*, 2011, **115**, 12983–12990.
- 19 J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J. P. Piquemal, D. N. Beratan and W. Yang, *Journal of Chemical Theory and Computation*, 2011, **7**, 625–632.
- 20 R. L. Beddoes, L. Dalton, J. A. Joule, O. S. Mills, J. Street and C. I. F. Watt, *J. CHEM. SOC. PERKIN TRANS. II*, 1986, 787–797.

- 21 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic acids research*, 2000, **28**, 235–242.