



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 07:52 PM BST

PDB ID : 6TQ4  
Title : Crystal structure of the Orexin-1 receptor in complex with Compound 16  
Authors : Rappas, M.; Ali, A.; Bennett, K.A.; Brown, J.D.; Bucknell, S.J.; Congreve, M.; Cooke, R.M.; Cseke, G.; de Graaf, C.; Dore, A.S.; Errey, J.C.; Jazayeri, A.; Marshall, F.H.; Mason, J.S.; Mould, R.; Patel, J.C.; Tehan, B.G.; Weir, M.; Christopher, J.A.  
Deposited on : 2019-12-16  
Resolution : 2.30 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

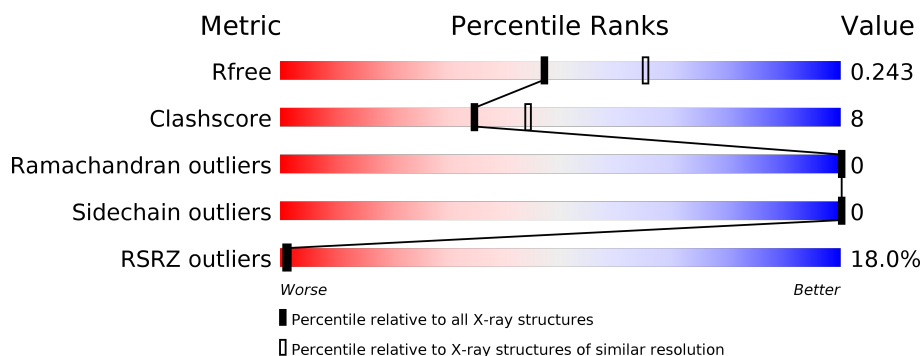
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	368	<div> <div>14%</div> <div>64%</div> <div>17%</div> <div>19%</div> </div>
1	B	368	<div> <div>15%</div> <div>68%</div> <div>14%</div> <div>19%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	SOG	B	408	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5217 atoms, of which 46 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Orexin receptor type 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total	C	N	O	S	0	0	0
			2388	1589	393	389	17			
1	B	299	Total	C	N	O	S	0	0	0
			2374	1576	397	384	17			

There are 52 discrepancies between the modelled and reference sequences:

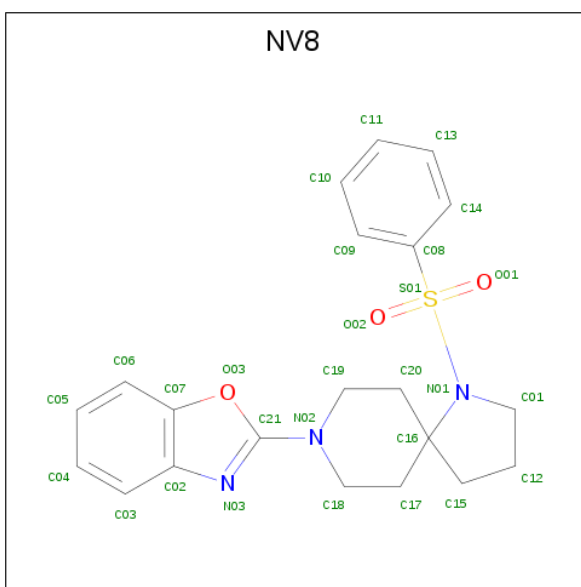
Chain	Residue	Modelled	Actual	Comment	Reference
A	25	ALA	-	expression tag	UNP O43613
A	26	ALA	-	expression tag	UNP O43613
A	27	SER	-	expression tag	UNP O43613
A	46	ALA	GLU	engineered mutation	UNP O43613
A	85	LEU	ILE	engineered mutation	UNP O43613
A	95	ALA	VAL	engineered mutation	UNP O43613
A	162	LEU	ARG	engineered mutation	UNP O43613
A	194	ALA	ASN	engineered mutation	UNP O43613
A	198	ALA	LEU	engineered mutation	UNP O43613
A	211	ALA	TYR	engineered mutation	UNP O43613
A	304	VAL	LEU	engineered mutation	UNP O43613
A	339	ALA	CYS	engineered mutation	UNP O43613
A	375	TRP	CYS	engineered mutation	UNP O43613
A	376	TRP	CYS	engineered mutation	UNP O43613
A	381	ALA	-	expression tag	UNP O43613
A	382	ALA	-	expression tag	UNP O43613
A	383	ALA	-	expression tag	UNP O43613
A	384	HIS	-	expression tag	UNP O43613
A	385	HIS	-	expression tag	UNP O43613
A	386	HIS	-	expression tag	UNP O43613
A	387	HIS	-	expression tag	UNP O43613
A	388	HIS	-	expression tag	UNP O43613
A	389	HIS	-	expression tag	UNP O43613
A	390	HIS	-	expression tag	UNP O43613
A	391	HIS	-	expression tag	UNP O43613

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	392	HIS	-	expression tag	UNP O43613
B	25	ALA	-	expression tag	UNP O43613
B	26	ALA	-	expression tag	UNP O43613
B	27	SER	-	expression tag	UNP O43613
B	46	ALA	GLU	engineered mutation	UNP O43613
B	85	LEU	ILE	engineered mutation	UNP O43613
B	95	ALA	VAL	engineered mutation	UNP O43613
B	162	LEU	ARG	engineered mutation	UNP O43613
B	194	ALA	ASN	engineered mutation	UNP O43613
B	198	ALA	LEU	engineered mutation	UNP O43613
B	211	ALA	TYR	engineered mutation	UNP O43613
B	304	VAL	LEU	engineered mutation	UNP O43613
B	339	ALA	CYS	engineered mutation	UNP O43613
B	375	TRP	CYS	engineered mutation	UNP O43613
B	376	TRP	CYS	engineered mutation	UNP O43613
B	381	ALA	-	expression tag	UNP O43613
B	382	ALA	-	expression tag	UNP O43613
B	383	ALA	-	expression tag	UNP O43613
B	384	HIS	-	expression tag	UNP O43613
B	385	HIS	-	expression tag	UNP O43613
B	386	HIS	-	expression tag	UNP O43613
B	387	HIS	-	expression tag	UNP O43613
B	388	HIS	-	expression tag	UNP O43613
B	389	HIS	-	expression tag	UNP O43613
B	390	HIS	-	expression tag	UNP O43613
B	391	HIS	-	expression tag	UNP O43613
B	392	HIS	-	expression tag	UNP O43613

- Molecule 2 is 2-[1-(phenylsulfonyl)-1,8-diazaspiro[4.5]decan-8-yl]-1,3-benzoxazole (three-letter code: NV8) (formula: C<sub>21</sub>H<sub>23</sub>N<sub>3</sub>O<sub>3</sub>S) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			51	21	23	3	3	1		
2	B	1	Total	C	H	N	O	S	0	0
			51	21	23	3	3	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



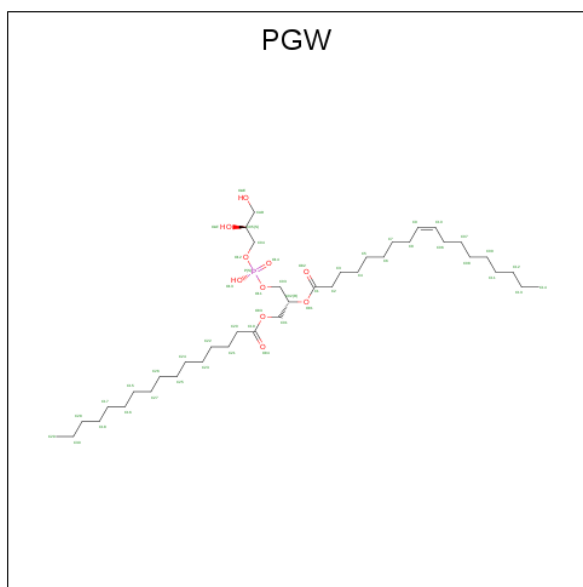
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	O S	0	0
			5	4 1		
3	A	1	Total	O S	0	0
			5	4 1		

*Continued on next page...*

Continued from previous page...

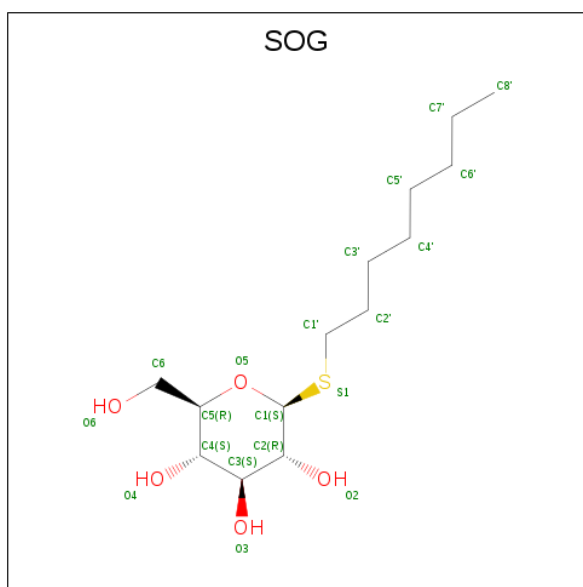
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is (1R)-2-{[(S)-{[(2S)-2,3-dihydroxypropyl]oxy}(hydroxy)phosphoryl]oxy}-1-[(hexadecanoyloxy)methyl]ethyl (9Z)-octadec-9-enoate (three-letter code: PGW) (formula: C<sub>40</sub>H<sub>77</sub>O<sub>10</sub>P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	P	0	0
			51	40	10	1		
4	B	1	Total	C	O	P	0	0
			51	40	10	1		

- Molecule 5 is octyl 1-thio-beta-D-glucopyranoside (three-letter code: SOG) (formula: C<sub>14</sub>H<sub>28</sub>O<sub>5</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O S 20 14 5 1	0	0
5	A	1	Total C O S 20 14 5 1	0	0
5	A	1	Total C O S 20 14 5 1	0	0
5	A	1	Total C O S 13 7 5 1	0	0
5	A	1	Total C O S 14 8 5 1	0	0
5	A	1	Total C O S 20 14 5 1	0	0
5	A	1	Total C S 5 4 1	0	0
5	A	1	Total C 4 4	0	0
5	B	1	Total C O S 14 8 5 1	0	0
5	B	1	Total C O S 13 7 5 1	0	0
5	B	1	Total C S 5 4 1	0	0
5	B	1	Total C O S 20 14 5 1	0	0
5	B	1	Total C O S 20 14 5 1	0	0

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Na 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	27	Total O 27 27	0	0
7	B	15	Total O 15 15	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.34Å 145.89Å 71.34Å 90.00° 112.33° 90.00°	Depositor
Resolution (Å)	34.69 – 2.30 34.69 – 2.30	Depositor EDS
% Data completeness (in resolution range)	77.1 (34.69-2.30) 87.1 (34.69-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.00 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, $R_{free}$	0.209 , 0.236 0.219 , 0.243	Depositor DCC
$R_{free}$ test set	2465 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.0	Xtriage
Anisotropy	0.275	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 68.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.030 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.03% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, PGW, SO4, SOG, NV8

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	A	0.28	0/2455	0.44	0/3344
1	B	0.29	0/2440	0.46	1/3328 (0.0%)
All	All	0.28	0/4895	0.45	1/6672 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	151	PRO	N-CA-C	5.42	126.19	112.10

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2388	0	2434	50	0
1	B	2374	0	2445	35	0
2	A	28	23	0	1	0
2	B	28	23	0	1	0
3	A	10	0	0	0	0
3	B	10	0	0	1	0
4	A	51	0	76	8	0
4	B	51	0	76	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	116	0	149	4	0
5	B	72	0	86	3	0
6	A	1	0	0	0	0
7	A	27	0	0	1	0
7	B	15	0	0	1	0
All	All	5171	46	5266	86	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

All (86) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:233:MET:HB3	4:A:404:PGW:H2A	1.38	1.05
1:A:233:MET:HE1	4:A:404:PGW:H23A	1.38	1.04
1:A:71:VAL:HG21	1:A:86:VAL:HG21	1.51	0.93
1:B:234:ALA:HB2	4:B:404:PGW:H7A	1.54	0.90
1:A:150:HIS:HE1	5:B:409:SOG:H2'1	1.44	0.83
1:A:318:ASN:ND2	2:A:401:NV8:O01	2.16	0.78
1:A:83:TYR:HE2	1:A:158:ALA:HB1	1.52	0.74
1:A:150:HIS:CE1	5:B:409:SOG:H2'1	2.24	0.72
1:A:71:VAL:HG21	1:A:86:VAL:CG2	2.20	0.72
1:B:83:TYR:HE2	1:B:158:ALA:HB1	1.57	0.69
1:A:83:TYR:CE2	1:A:158:ALA:HB1	2.26	0.69
1:B:138:SER:HB3	4:B:404:PGW:H07	1.76	0.68
1:B:190:PRO:HG2	1:B:191:GLU:OE1	1.94	0.68
1:B:83:TYR:CE2	1:B:158:ALA:HB1	2.31	0.65
1:A:233:MET:CB	4:A:404:PGW:H2A	2.23	0.65
1:B:318:ASN:ND2	2:B:401:NV8:O01	2.30	0.64
1:A:228:LEU:HD23	1:A:231:MET:CE	2.28	0.64
1:B:363:GLY:N	3:B:403:SO4:O2	2.34	0.61
1:B:156:SER:HB3	5:B:405:SOG:S1	2.41	0.60
1:B:153:LEU:N	1:B:153:LEU:HD23	2.18	0.58
1:A:36:TRP:HA	1:A:40:LEU:HB2	1.85	0.58
1:A:234:ALA:HB2	4:A:404:PGW:H7A	1.87	0.57
1:B:250:THR:CG2	1:B:286:GLU:HG3	2.35	0.57
1:A:228:LEU:HD23	1:A:231:MET:HE3	1.87	0.55
1:A:167:GLY:O	1:A:171:VAL:HG23	2.07	0.54
1:B:290:MET:O	1:B:294:ARG:HG3	2.07	0.53
1:A:82:ASN:O	1:A:86:VAL:HG23	2.07	0.53
1:A:139:PHE:HZ	5:A:406:SOG:H5'2	1.74	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:404:PGW:H02	4:B:404:PGW:H04A	1.92	0.52
1:B:191:GLU:OE1	1:B:191:GLU:N	2.43	0.52
1:A:293:ARG:HA	1:A:296:THR:HG22	1.92	0.52
1:A:228:LEU:HD23	1:A:231:MET:HE1	1.92	0.51
1:B:92:ASP:HB3	1:B:351:SER:HB3	1.92	0.51
1:A:70:ALA:HB1	1:A:77:MET:HE1	1.93	0.50
1:B:251:THR:HG22	1:B:285:ALA:H	1.77	0.50
1:A:333:ARG:HD3	1:A:337:TYR:OH	2.12	0.50
1:A:77:MET:SD	1:A:364:LYS:HB3	2.52	0.50
1:B:50:ILE:HG12	1:B:108:ILE:HD11	1.95	0.49
1:A:179:GLN:O	1:A:183:MET:HG2	2.13	0.49
1:B:297:ALA:O	1:B:301:MET:HG3	2.13	0.49
1:B:250:THR:HG23	1:B:286:GLU:HG3	1.95	0.49
5:A:412:SOG:H8'1	7:A:525:HOH:O	2.13	0.48
1:A:184:GLU:HG3	1:A:205:ARG:HE	1.78	0.48
1:A:333:ARG:HG2	1:A:337:TYR:CE2	2.50	0.47
1:A:353:ALA:O	1:A:357:ILE:HG13	2.14	0.47
1:A:233:MET:CE	4:A:404:PGW:H21A	2.45	0.47
1:A:321:LYS:NZ	5:A:412:SOG:H8'2	2.30	0.46
1:A:50:ILE:O	1:A:54:VAL:HG23	2.15	0.46
1:A:100:LEU:HB3	1:A:101:PRO:HD3	1.97	0.46
1:A:63:GLY:O	1:A:67:VAL:HG23	2.16	0.46
1:B:126:GLN:O	1:B:130:VAL:HG23	2.16	0.45
1:A:28:GLU:O	1:A:32:LEU:HD12	2.16	0.45
1:B:100:LEU:HD23	1:B:100:LEU:C	2.36	0.45
1:B:62:VAL:O	1:B:66:LEU:HG	2.16	0.45
1:A:164:SER:O	1:A:168:ILE:HG13	2.16	0.45
1:A:233:MET:HE1	4:A:404:PGW:C23	2.29	0.45
1:A:312:LEU:HB3	1:A:313:PRO:HD3	1.98	0.45
1:A:211:ALA:N	1:A:212:PRO:HD2	2.31	0.45
1:A:62:VAL:O	1:A:66:LEU:HG	2.16	0.45
1:B:316:VAL:O	1:B:320:LEU:HG	2.17	0.45
5:A:406:SOG:H4'1	1:B:171:VAL:HG21	1.98	0.44
1:A:299:MET:O	1:A:303:VAL:HG23	2.17	0.44
1:A:215:TYR:HB2	4:B:404:PGW:H24	1.99	0.44
1:B:312:LEU:HB3	1:B:313:PRO:HD3	1.99	0.44
1:B:63:GLY:O	1:B:67:VAL:HG23	2.17	0.44
1:B:159:ARG:NH2	7:B:501:HOH:O	2.45	0.44
1:B:251:THR:HG22	1:B:252:SER:N	2.33	0.43
1:A:77:MET:HE2	1:A:368:GLN:HG3	2.00	0.43
1:A:227:PRO:HD3	4:A:404:PGW:H14B	2.01	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:100:LEU:HD12	1:A:348:TYR:HB3	1.99	0.43
1:B:124:TYR:HB2	1:B:180:ALA:HB2	2.00	0.43
1:A:287:VAL:HG22	1:A:287:VAL:O	2.19	0.42
1:A:221:ILE:HD12	1:B:221:ILE:HG13	2.01	0.42
1:B:250:THR:HG22	1:B:286:GLU:HG3	2.01	0.42
1:A:300:LEU:O	1:A:304:VAL:HG23	2.19	0.42
1:B:179:GLN:O	1:B:183:MET:HG2	2.19	0.42
1:B:211:ALA:N	1:B:212:PRO:HD2	2.35	0.42
1:A:297:ALA:O	1:A:301:MET:HG3	2.20	0.41
4:A:404:PGW:O02	1:B:177:VAL:HG12	2.20	0.41
1:A:89:SER:O	1:A:93:VAL:HG23	2.21	0.41
1:A:227:PRO:O	1:A:231:MET:HG3	2.21	0.41
1:A:41:TYR:HB3	1:A:42:PRO:HD3	2.03	0.41
1:B:100:LEU:HB3	1:B:101:PRO:HD3	2.02	0.41
1:B:187:SER:OG	1:B:192:LEU:HB3	2.21	0.41
1:A:28:GLU:HG2	1:A:32:LEU:HD11	2.02	0.40
1:B:119:CYS:O	1:B:123:PRO:HG2	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	291/368 (79%)	288 (99%)	3 (1%)	0	100	100
1	B	295/368 (80%)	292 (99%)	3 (1%)	0	100	100
All	All	586/736 (80%)	580 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	251/305 (82%)	251 (100%)	0	100	100
1	B	249/305 (82%)	249 (100%)	0	100	100
All	All	500/610 (82%)	500 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	150	HIS
1	A	318	ASN
1	B	318	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 1 is monoatomic - leaving 21 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The



Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NV8	B	401	-	27,32,32	1.76	3 (11%)	31,48,48	1.48	3 (9%)
3	SO4	B	402	-	4,4,4	0.15	0	6,6,6	0.06	0
5	SOG	B	408	-	20,20,20	0.90	1 (5%)	24,25,25	0.70	0
5	SOG	A	406	-	20,20,20	1.02	2 (10%)	24,25,25	0.45	0
4	PGW	A	404	-	50,50,50	0.93	2 (4%)	53,56,56	1.11	4 (7%)
5	SOG	B	407	-	4,4,20	0.65	0	3,3,25	0.78	0
3	SO4	A	403	-	4,4,4	0.14	0	6,6,6	0.05	0
5	SOG	B	406	-	12,13,20	0.53	0	16,18,25	0.64	0
3	SO4	B	403	-	4,4,4	0.13	0	6,6,6	0.08	0
4	PGW	B	404	-	50,50,50	0.92	2 (4%)	53,56,56	1.08	4 (7%)
5	SOG	B	405	-	14,14,20	0.79	1 (7%)	18,19,25	0.77	0
5	SOG	A	408	-	12,13,20	0.55	0	16,18,25	0.71	0
5	SOG	A	407	-	20,20,20	0.98	2 (10%)	24,25,25	1.10	2 (8%)
2	NV8	A	401	-	27,32,32	1.85	2 (7%)	31,48,48	1.47	5 (16%)
5	SOG	A	412	-	3,3,20	0.35	0	2,2,25	0.63	0
3	SO4	A	402	-	4,4,4	0.14	0	6,6,6	0.05	0
5	SOG	B	409	-	20,20,20	1.03	2 (10%)	24,25,25	0.67	0
5	SOG	A	409	-	14,14,20	0.82	1 (7%)	18,19,25	0.56	0
5	SOG	A	410	-	20,20,20	1.00	2 (10%)	24,25,25	0.75	0
5	SOG	A	411	-	4,4,20	0.67	0	3,3,25	0.85	0
5	SOG	A	405	-	20,20,20	1.07	2 (10%)	24,25,25	0.84	1 (4%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	SOG	A	412	-	-	0/1/1/31	-
5	SOG	B	408	-	-	1/11/31/31	0/1/1/1
5	SOG	A	406	-	-	2/11/31/31	0/1/1/1
5	SOG	B	406	-	-	0/4/24/31	0/1/1/1
5	SOG	A	411	-	-	0/2/2/31	-
4	PGW	A	404	-	-	13/55/55/55	-
5	SOG	B	407	-	-	0/2/2/31	-

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PGW	B	404	-	-	19/55/55/55	-
5	SOG	B	405	-	-	1/5/25/31	0/1/1/1
5	SOG	A	408	-	-	2/4/24/31	0/1/1/1
5	SOG	A	407	-	-	4/11/31/31	0/1/1/1
2	NV8	A	401	-	-	0/12/41/41	0/5/5/5
2	NV8	B	401	-	-	0/12/41/41	0/5/5/5
5	SOG	A	410	-	-	5/11/31/31	0/1/1/1
5	SOG	B	409	-	-	1/11/31/31	0/1/1/1
5	SOG	A	409	-	-	1/5/25/31	0/1/1/1
5	SOG	A	405	-	-	4/11/31/31	0/1/1/1

All (22) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	NV8	S01-N01	8.15	1.77	1.63
2	B	401	NV8	S01-N01	7.58	1.76	1.63
4	A	404	PGW	O03-C19	4.40	1.46	1.33
4	B	404	PGW	O03-C19	4.22	1.45	1.33
4	B	404	PGW	O01-C1	4.09	1.45	1.34
4	A	404	PGW	O01-C1	4.05	1.45	1.34
5	A	405	SOG	C1'-S1	-3.83	1.76	1.81
5	B	409	SOG	C1'-S1	-3.72	1.76	1.81
5	A	406	SOG	C1'-S1	-3.67	1.76	1.81
5	A	410	SOG	C1'-S1	-3.59	1.76	1.81
5	A	407	SOG	C1'-S1	-3.55	1.76	1.81
5	B	408	SOG	C1'-S1	-3.23	1.77	1.81
2	A	401	NV8	C21-N03	-3.02	1.31	1.35
2	B	401	NV8	C21-N03	-2.74	1.31	1.35
5	A	405	SOG	C1-S1	-2.55	1.76	1.80
5	B	409	SOG	C1-S1	-2.48	1.76	1.80
5	A	409	SOG	C1-S1	-2.48	1.76	1.80
5	A	406	SOG	C1-S1	-2.44	1.77	1.80
5	A	410	SOG	C1-S1	-2.40	1.77	1.80
5	B	405	SOG	C1-S1	-2.34	1.77	1.80
5	A	407	SOG	C1-S1	-2.10	1.77	1.80
2	B	401	NV8	C08-S01	2.07	1.79	1.76

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	404	PGW	O01-C1-C2	4.66	121.55	111.50

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	401	NV8	C20-C19-N02	4.61	118.50	110.34
4	B	404	PGW	O01-C1-C2	4.41	121.00	111.50
2	A	401	NV8	C20-C19-N02	3.48	116.50	110.34
2	A	401	NV8	C17-C18-N02	3.38	116.33	110.34
2	A	401	NV8	O01-S01-C08	-3.23	103.96	108.05
2	A	401	NV8	C12-C15-C16	3.11	108.64	104.58
4	A	404	PGW	C02-O01-C1	-3.07	110.23	117.79
5	A	407	SOG	C1-C2-C3	2.90	116.31	110.59
2	B	401	NV8	C12-C15-C16	2.83	108.26	104.58
2	A	401	NV8	O02-S01-N01	2.66	112.86	107.74
4	B	404	PGW	C02-O01-C1	-2.65	111.26	117.79
2	B	401	NV8	C17-C18-N02	2.64	115.01	110.34
4	A	404	PGW	O03-C19-C20	2.54	119.87	111.91
4	B	404	PGW	O03-C19-C20	2.51	119.80	111.91
5	A	407	SOG	O5-C1-C2	2.38	113.31	110.31
5	A	405	SOG	C1'-S1-C1	-2.24	95.89	100.09
4	B	404	PGW	O01-C1-O02	-2.22	118.34	123.70
4	A	404	PGW	O01-C1-O02	-2.04	118.78	123.70

There are no chirality outliers.

All (53) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	404	PGW	C03-O11-P-O14
4	B	404	PGW	C03-O11-P-O14
5	A	410	SOG	O5-C1-S1-C1'
5	A	407	SOG	C2-C1-S1-C1'
5	A	407	SOG	O5-C1-S1-C1'
5	A	405	SOG	C2-C1-S1-C1'
5	A	405	SOG	O5-C1-S1-C1'
4	A	404	PGW	C1-C2-C3-C4
5	A	406	SOG	S1-C1'-C2'-C3'
5	A	408	SOG	C4-C5-C6-O6
4	B	404	PGW	C03-O11-P-O12
5	A	410	SOG	C1'-C2'-C3'-C4'
5	A	408	SOG	O5-C5-C6-O6
4	B	404	PGW	C5-C6-C7-C8
5	A	405	SOG	C4'-C5'-C6'-C7'
5	A	409	SOG	O5-C5-C6-O6
4	A	404	PGW	C16-C15-C27-C26
4	B	404	PGW	C2-C3-C4-C5
4	A	404	PGW	C19-C20-C21-C22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	A	405	SOG	C2'-C3'-C4'-C5'
5	B	409	SOG	O5-C5-C6-O6
4	A	404	PGW	C4-C5-C6-C7
4	B	404	PGW	C24-C25-C26-C27
4	A	404	PGW	C08-C09-C11-C12
4	A	404	PGW	C3-C4-C5-C6
5	A	406	SOG	O5-C5-C6-O6
5	B	405	SOG	O5-C5-C6-O6
5	A	407	SOG	O5-C5-C6-O6
4	B	404	PGW	C08-C09-C11-C12
4	B	404	PGW	C2-C1-O01-C02
5	B	408	SOG	O5-C5-C6-O6
4	B	404	PGW	C20-C19-O03-C01
4	A	404	PGW	O03-C01-C02-C03
5	A	407	SOG	S1-C1'-C2'-C3'
4	B	404	PGW	O02-C1-O01-C02
5	A	410	SOG	C2'-C1'-S1-C1
4	B	404	PGW	C11-C12-C13-C14
4	B	404	PGW	O04-C19-O03-C01
4	A	404	PGW	C03-O11-P-O12
4	B	404	PGW	C03-O11-P-O13
4	B	404	PGW	C23-C24-C25-C26
5	A	410	SOG	C2-C1-S1-C1'
4	A	404	PGW	O03-C01-C02-O01
5	A	410	SOG	C4-C5-C6-O6
4	A	404	PGW	C7-C8-C9-C10
4	B	404	PGW	C25-C26-C27-C15
4	B	404	PGW	C7-C8-C9-C10
4	B	404	PGW	C01-C02-C03-O11
4	A	404	PGW	C2-C3-C4-C5
4	A	404	PGW	C25-C26-C27-C15
4	B	404	PGW	O01-C1-C2-C3
4	B	404	PGW	O02-C1-C2-C3
4	B	404	PGW	C15-C16-C17-C18

There are no ring outliers.

9 monomers are involved in 22 short contacts:

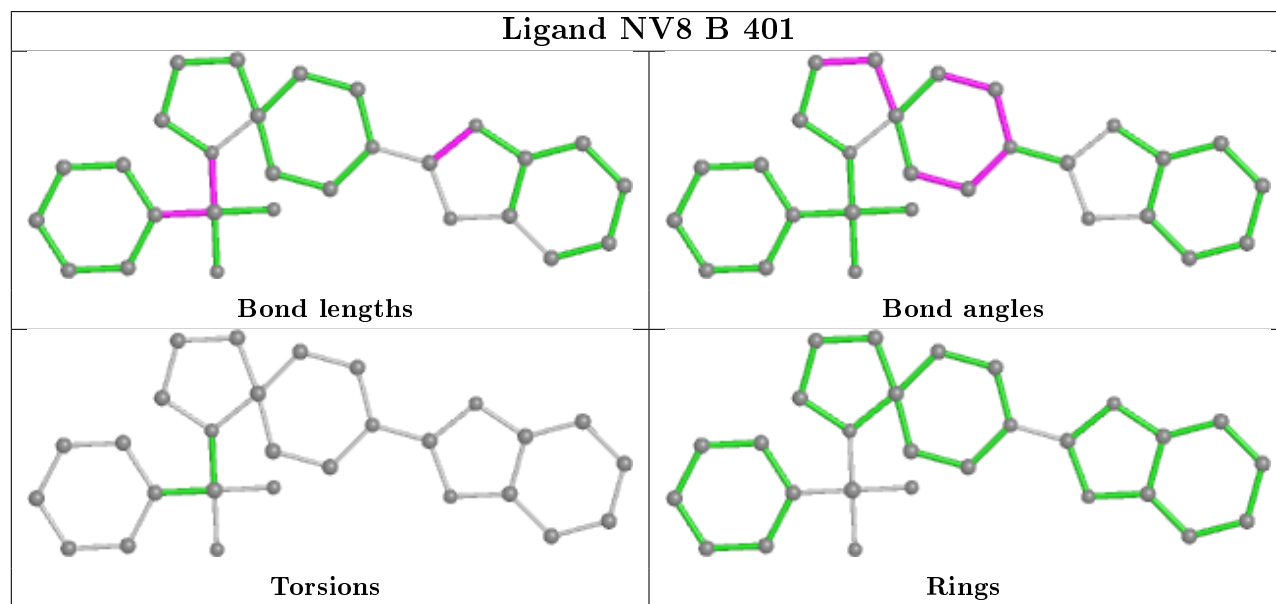
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	NV8	1	0
5	A	406	SOG	2	0
4	A	404	PGW	8	0

*Continued on next page...*

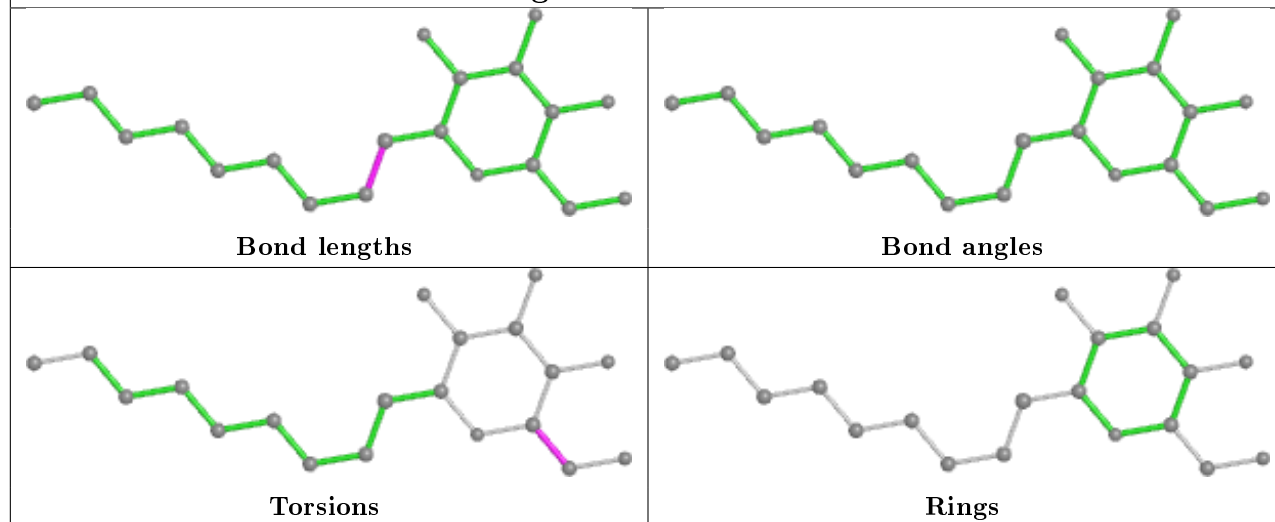
*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	403	SO4	1	0
4	B	404	PGW	4	0
5	B	405	SOG	1	0
2	A	401	NV8	1	0
5	A	412	SOG	2	0
5	B	409	SOG	2	0

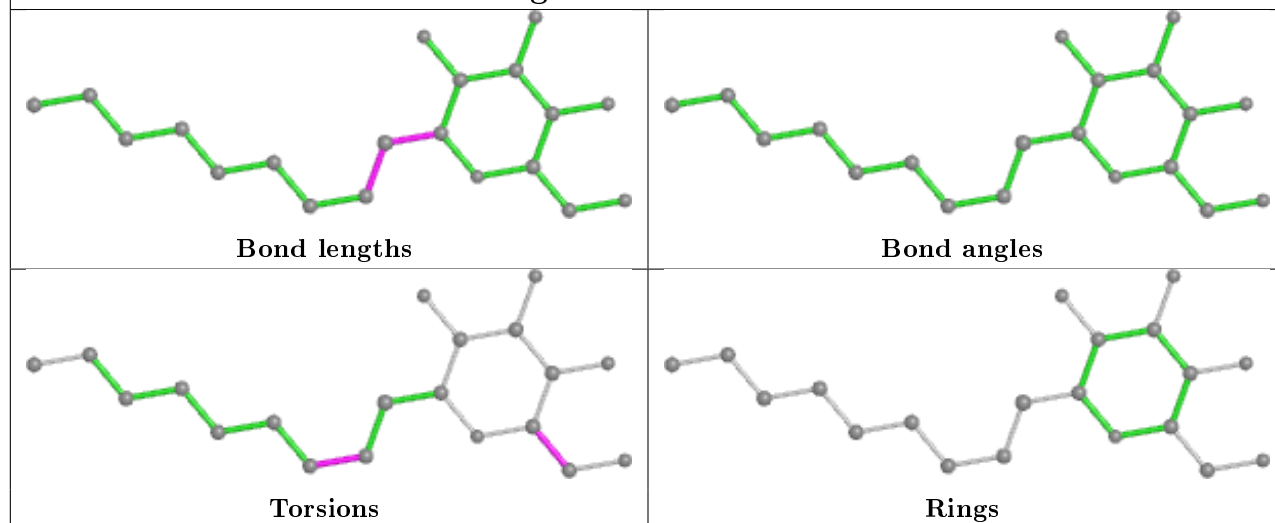
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

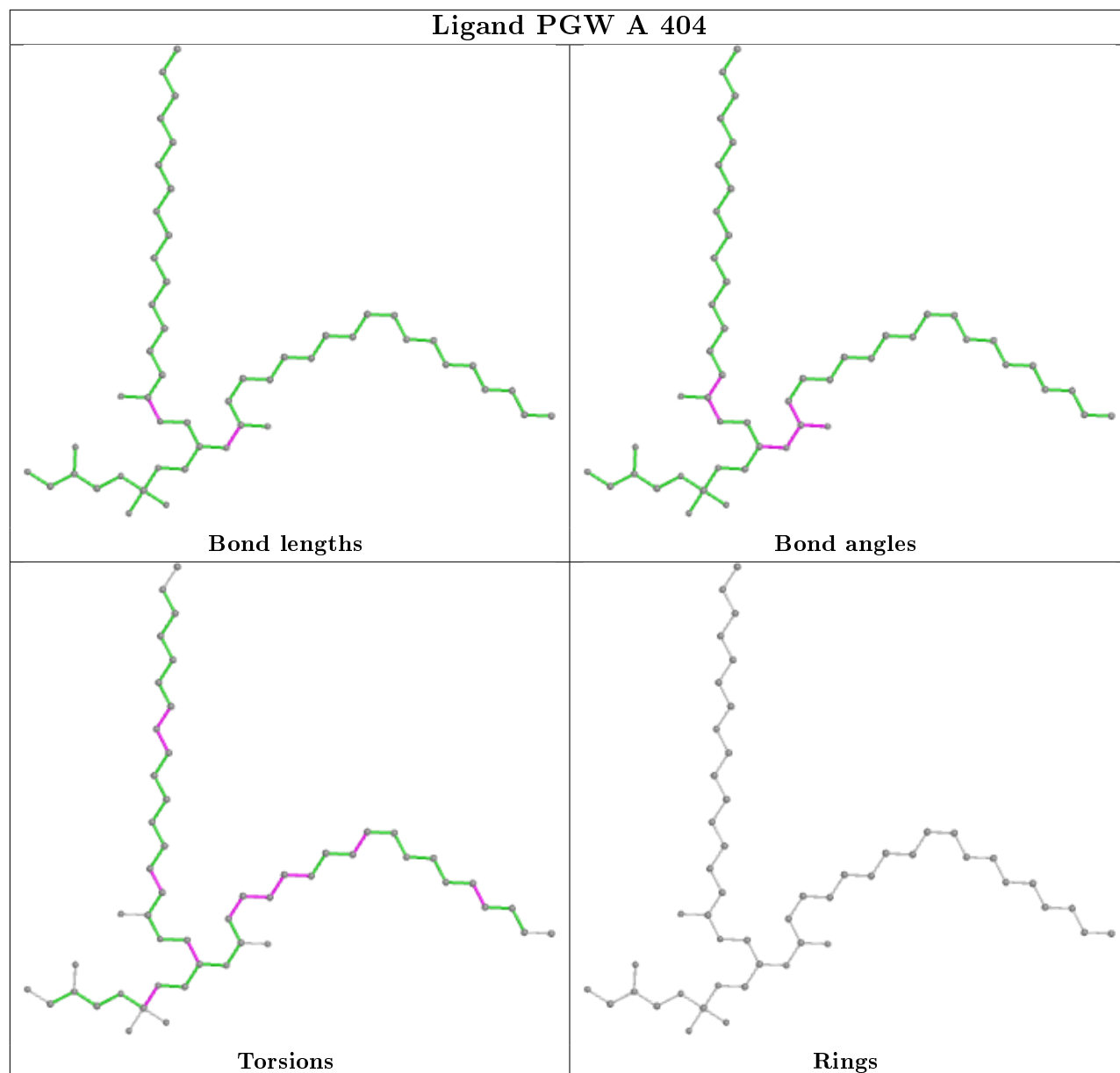


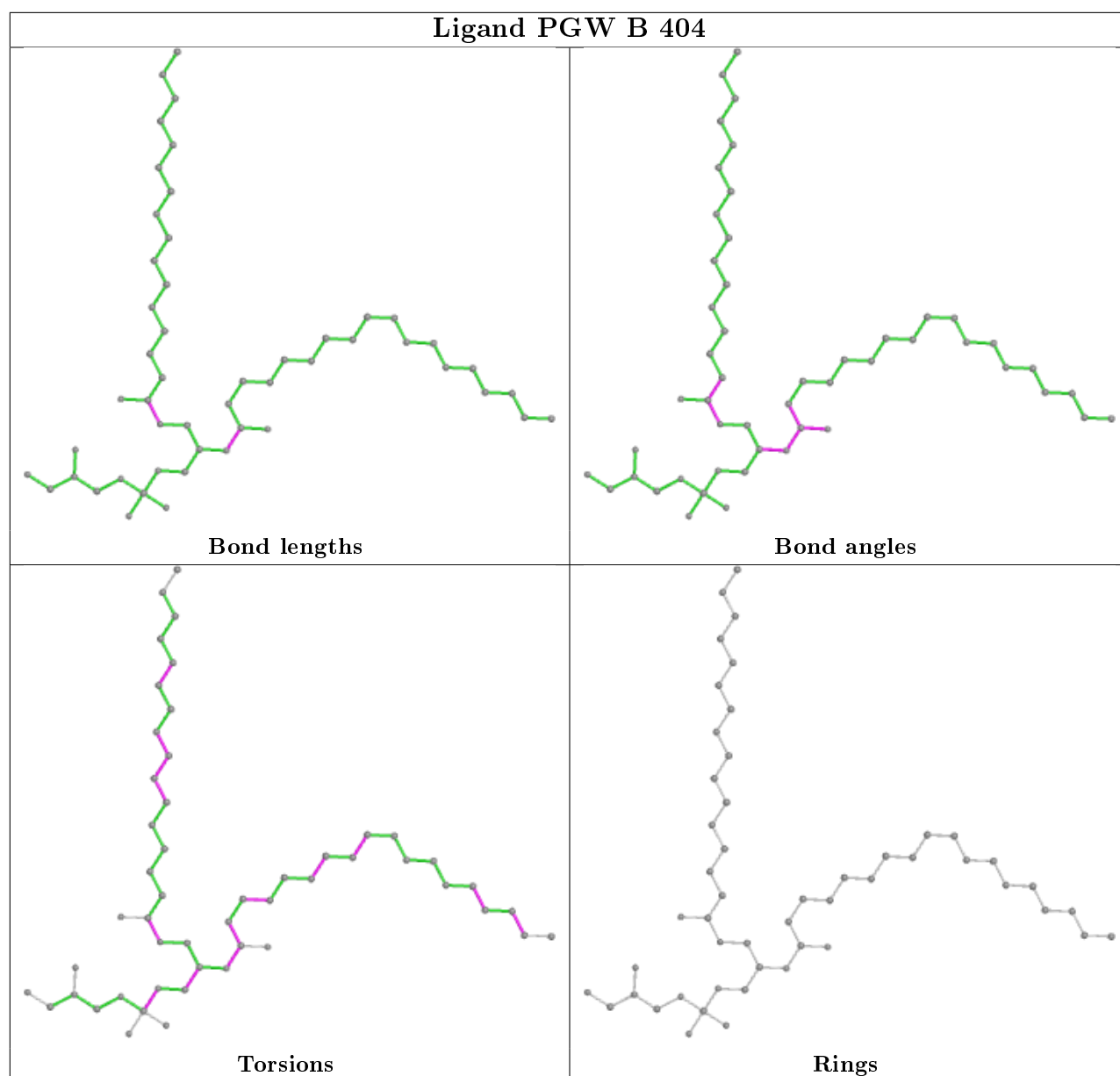
## Ligand SOG B 408



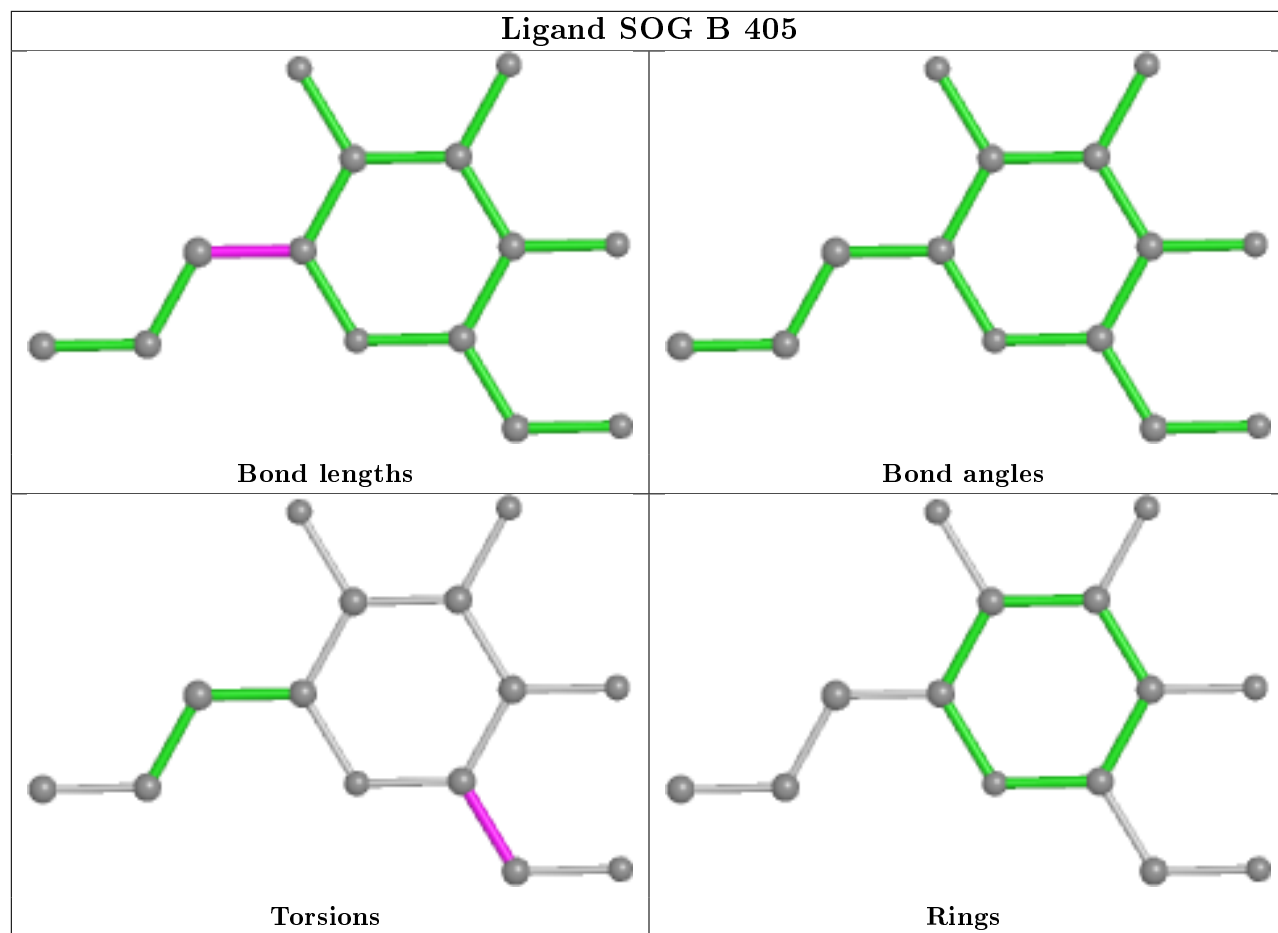
## Ligand SOG A 406



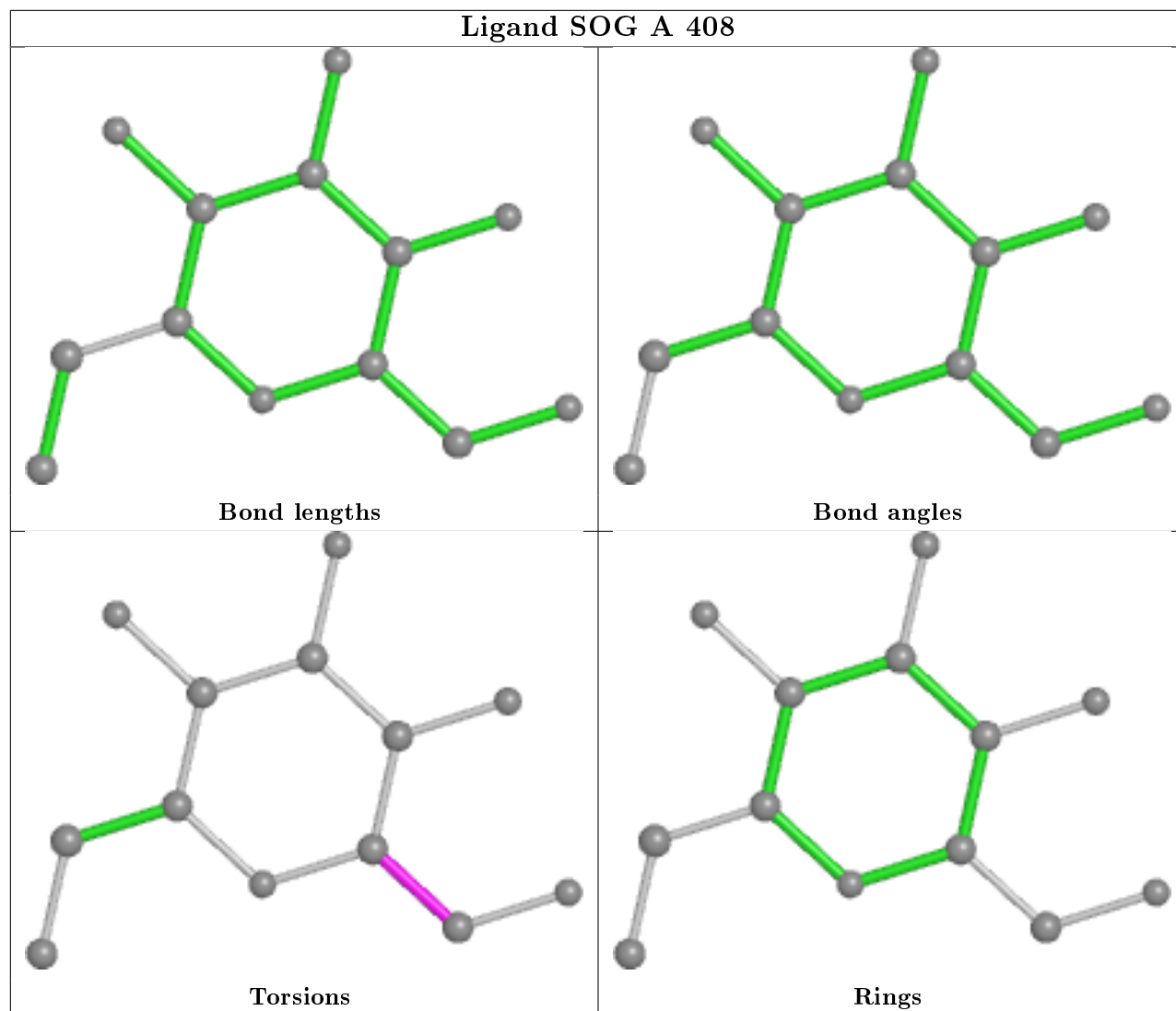




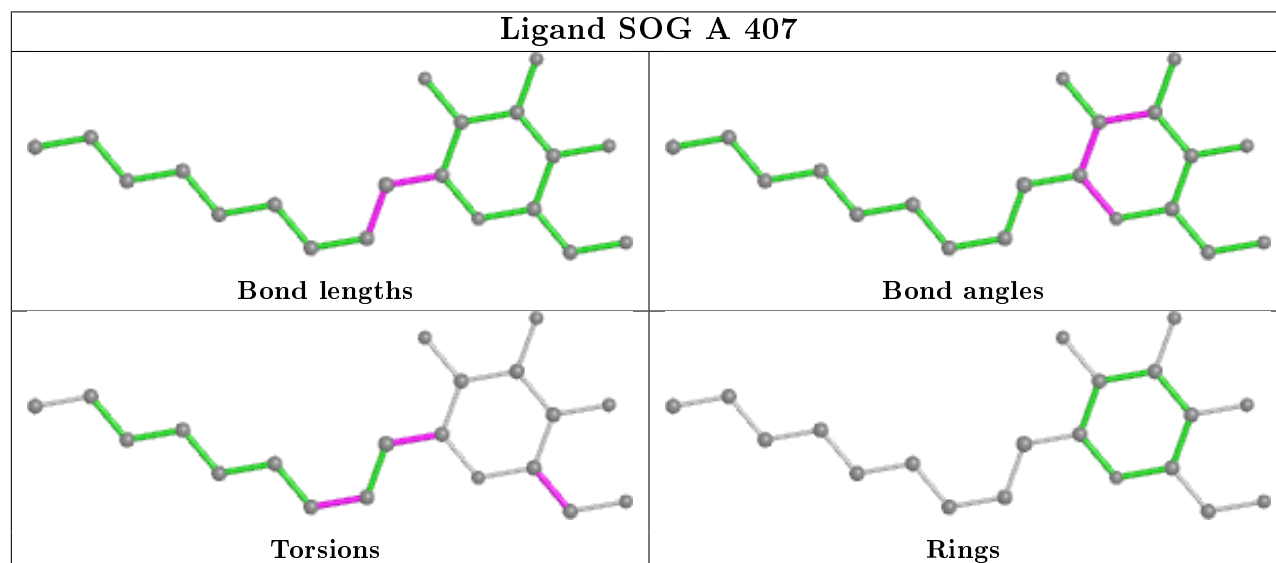


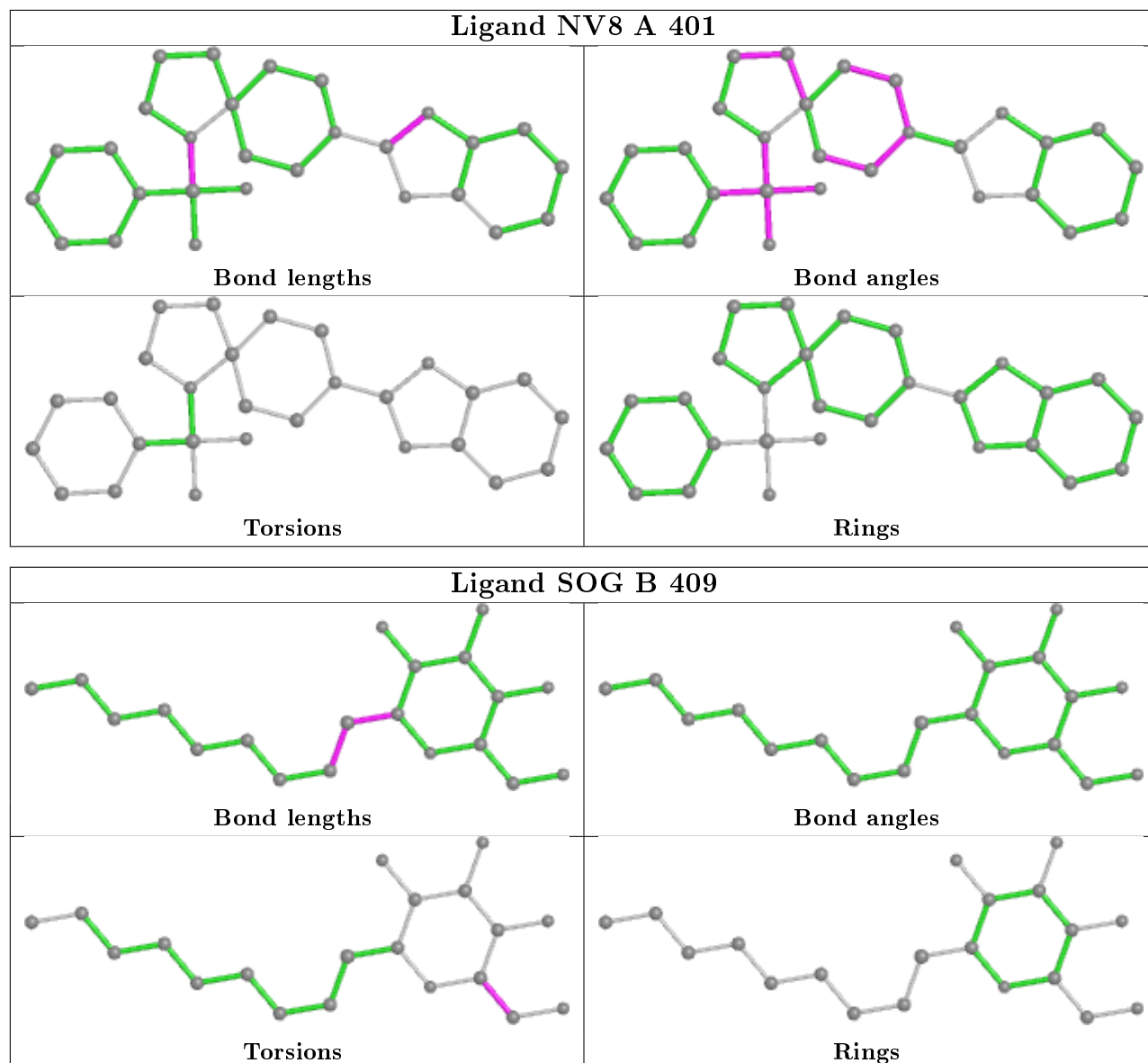


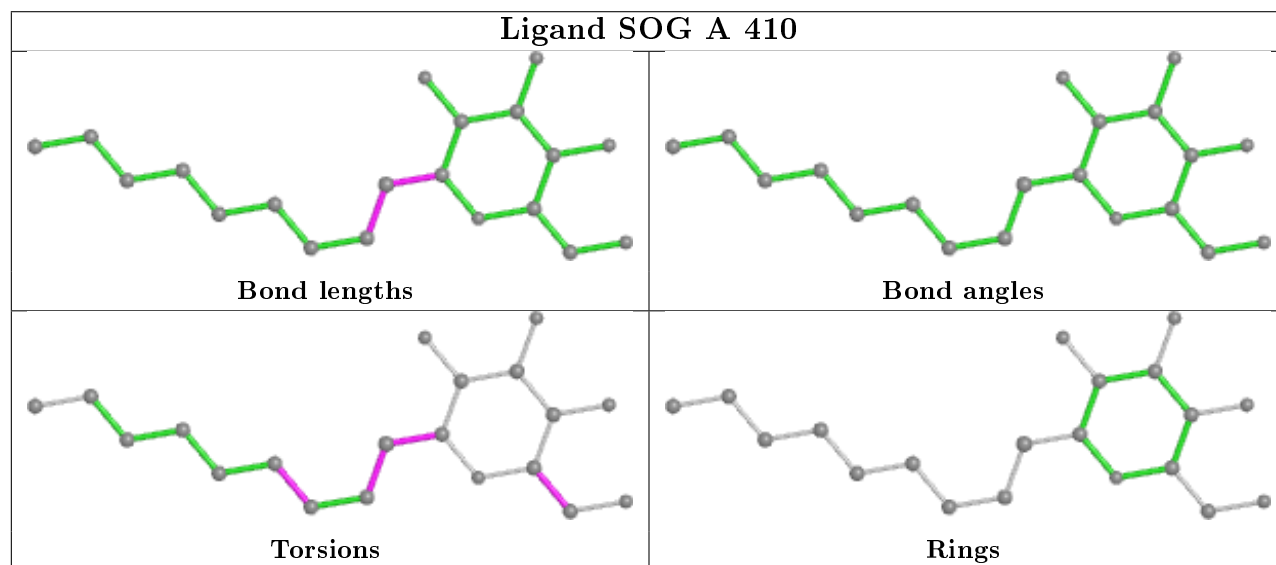
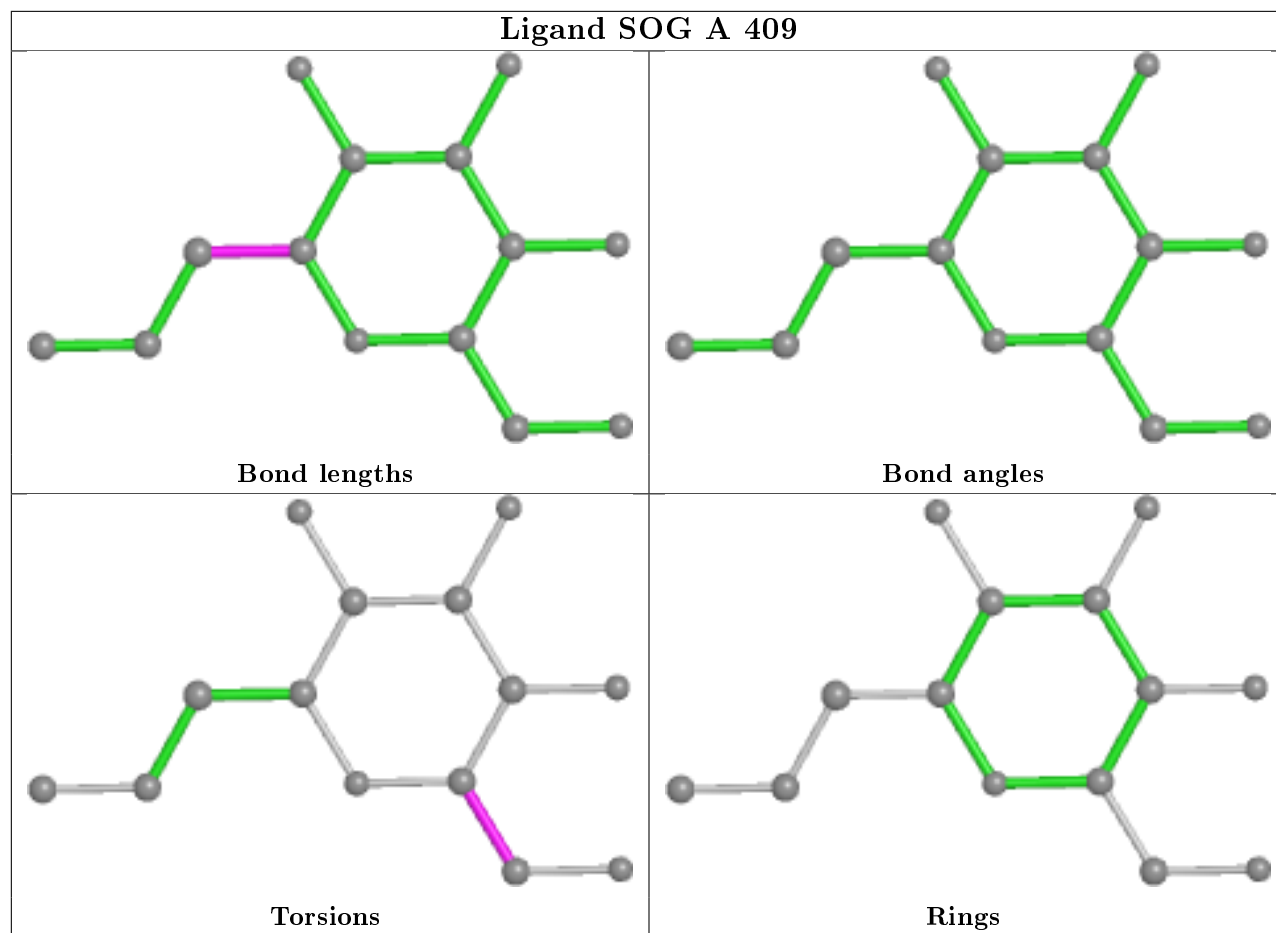
## Ligand SOG A 408

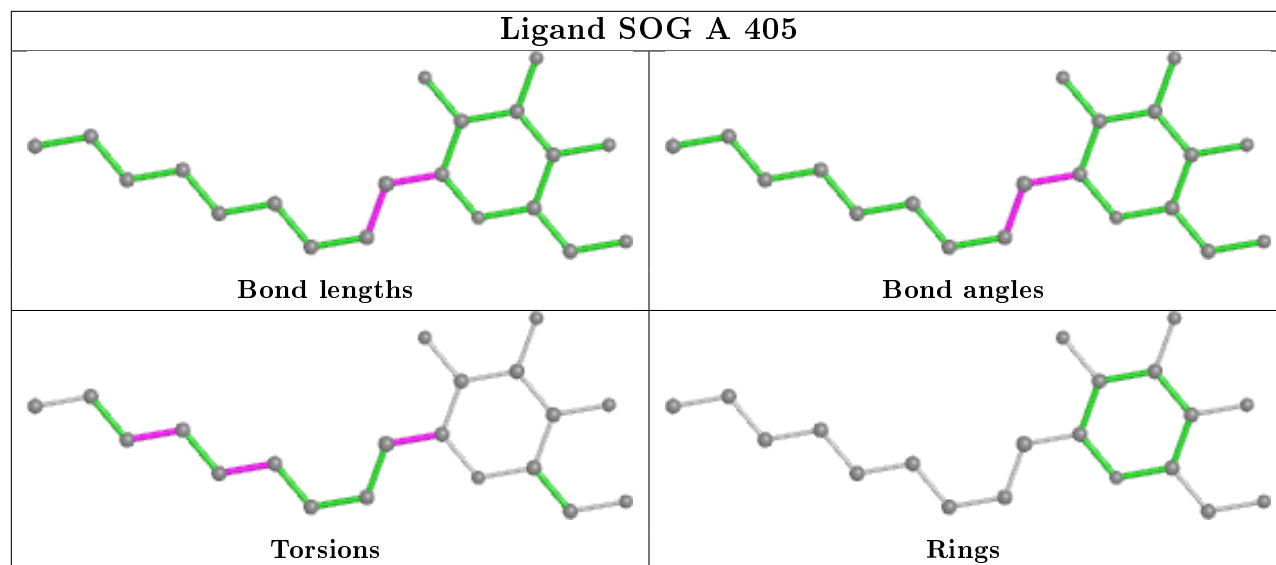


## Ligand SOG A 407









## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	297/368 (80%)	0.75	50 (16%) ⓘ ⓘ	44, 74, 125, 178	0
1	B	299/368 (81%)	1.04	57 (19%) ⓘ ⓘ	47, 73, 133, 171	0
All	All	596/736 (80%)	0.89	107 (17%) ⓘ ⓘ	44, 73, 130, 178	0

All (107) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	45	TYR	13.1
1	B	47	TRP	11.6
1	B	245	ARG	10.8
1	B	376	TRP	9.8
1	B	375	TRP	9.1
1	B	246	GLN	8.9
1	B	288	LYS	8.5
1	B	329	GLN	8.3
1	B	378	PRO	6.7
1	B	373	PHE	6.7
1	A	188	VAL	6.0
1	A	114	PHE	6.0
1	B	377	LEU	5.7
1	B	287	VAL	5.4
1	A	373	PHE	5.3
1	B	247	ILE	5.3
1	B	48	VAL	5.3
1	A	375	TRP	5.3
1	B	46	ALA	4.9
1	A	113	LEU	4.9
1	A	291	ARG	4.7
1	A	199	PHE	4.6
1	B	248	PRO	4.6
1	A	69	LEU	4.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	291	ARG	4.5
1	B	190	PRO	4.5
1	A	329	GLN	4.5
1	A	108	ILE	4.4
1	A	330	ALA	4.4
1	A	331	SER	4.3
1	A	239	PHE	4.3
1	B	72	TRP	4.2
1	B	372	ALA	4.0
1	A	288	LYS	4.0
1	A	130	VAL	3.8
1	A	372	ALA	3.8
1	A	116	HIS	3.7
1	B	130	VAL	3.6
1	B	285	ALA	3.6
1	B	337	TYR	3.6
1	A	360	PHE	3.5
1	B	88	LEU	3.5
1	B	189	LEU	3.5
1	A	128	VAL	3.5
1	B	75	HIS	3.4
1	B	244	GLY	3.4
1	B	133	ALA	3.4
1	A	105	LEU	3.3
1	B	85	LEU	3.3
1	B	132	VAL	3.3
1	B	327	PHE	3.3
1	A	243	TRP	3.2
1	B	328	ARG	3.2
1	B	76	HIS	3.2
1	B	129	SER	3.2
1	A	152	LEU	3.2
1	A	361	LEU	3.2
1	B	136	THR	3.1
1	B	108	ILE	3.0
1	A	374	SER	2.9
1	B	370	LYS	2.9
1	B	134	VAL	2.9
1	B	333	ARG	2.9
1	A	172	SER	2.8
1	B	242	LEU	2.8
1	B	243	TRP	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	242	LEU	2.8
1	A	298	LYS	2.8
1	B	289	GLN	2.8
1	B	155	LYS	2.8
1	B	239	PHE	2.8
1	A	115	GLY	2.7
1	A	171	VAL	2.7
1	B	152	LEU	2.7
1	A	137	LEU	2.7
1	B	374	SER	2.7
1	B	91	ALA	2.6
1	A	166	LEU	2.6
1	A	45	TYR	2.6
1	A	132	VAL	2.5
1	A	324	PHE	2.5
1	B	137	LEU	2.5
1	A	88	LEU	2.5
1	B	131	SER	2.5
1	A	127	ALA	2.4
1	B	252	SER	2.4
1	A	109	THR	2.4
1	A	47	TRP	2.4
1	A	117	ALA	2.4
1	B	128	VAL	2.3
1	A	72	TRP	2.3
1	A	236	PHE	2.2
1	B	172	SER	2.2
1	A	129	SER	2.2
1	B	292	ALA	2.2
1	B	135	LEU	2.2
1	A	240	ARG	2.1
1	A	292	ALA	2.1
1	A	131	SER	2.1
1	A	83	TYR	2.1
1	B	240	ARG	2.1
1	A	139	PHE	2.1
1	A	153	LEU	2.0
1	B	191	GLU	2.0
1	A	175	ILE	2.0
1	A	133	ALA	2.0
1	B	335	ALA	2.0



## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

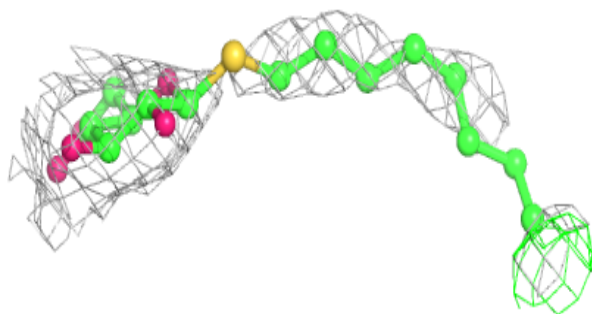
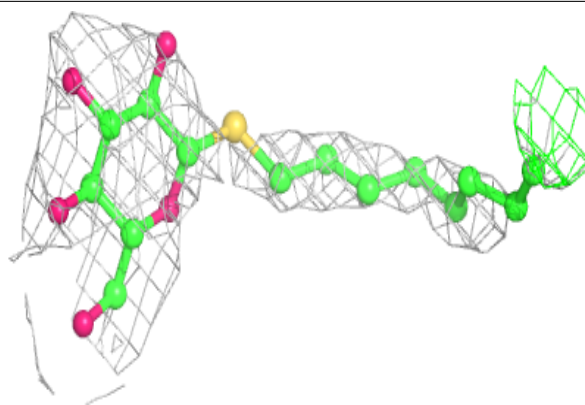
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SOG	A	411	5/20	0.39	0.33	107,112,139,157	0
5	SOG	B	408	20/20	0.62	0.45	102,141,151,156	0
5	SOG	B	409	20/20	0.66	0.35	122,127,138,140	0
2	NV8	B	401	28/28	0.70	0.26	80,132,160,162	0
5	SOG	B	405	14/20	0.74	0.36	160,165,167,168	0
5	SOG	A	407	20/20	0.74	0.36	71,126,140,141	0
5	SOG	A	406	20/20	0.75	0.29	69,125,133,137	0
5	SOG	A	408	13/20	0.76	0.36	135,145,149,149	0
5	SOG	A	409	14/20	0.81	0.15	136,148,152,155	0
4	PGW	B	404	51/51	0.82	0.23	44,79,147,155	0
4	PGW	A	404	51/51	0.83	0.21	49,77,145,148	0
5	SOG	A	410	20/20	0.85	0.17	86,113,128,128	0
5	SOG	A	412	4/20	0.86	0.20	69,72,74,82	0
3	SO4	A	403	5/5	0.86	0.18	124,126,127,133	0
5	SOG	B	406	13/20	0.87	0.33	104,124,127,129	0
2	NV8	A	401	28/28	0.87	0.15	81,115,143,145	0
5	SOG	B	407	5/20	0.91	0.11	90,95,99,105	0
3	SO4	B	403	5/5	0.91	0.12	106,112,120,124	0
6	NA	A	413	1/1	0.93	0.54	82,82,82,82	0
5	SOG	A	405	20/20	0.94	0.23	73,110,123,123	0
3	SO4	A	402	5/5	0.96	0.10	95,97,104,108	0
3	SO4	B	402	5/5	0.98	0.07	88,96,101,111	0

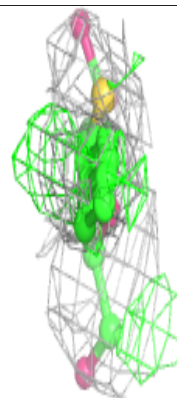
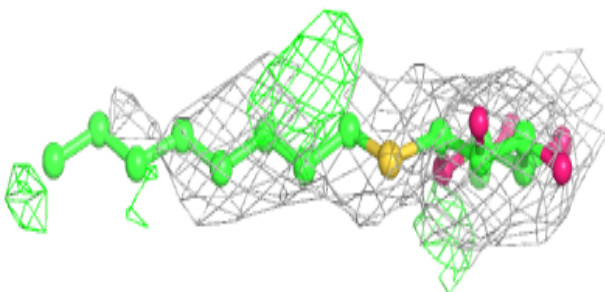
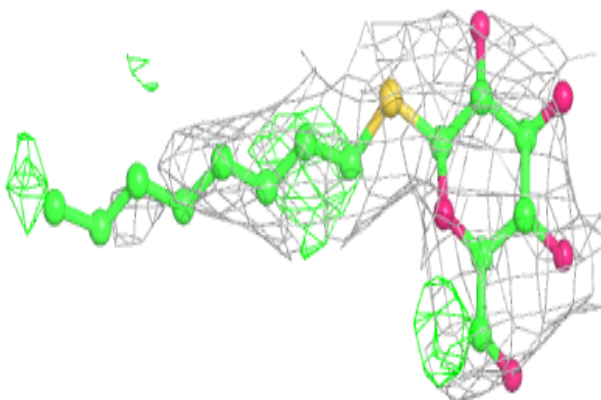
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around SOG B 408:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

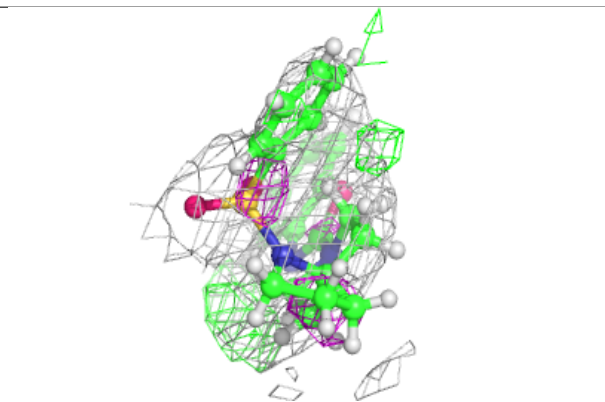
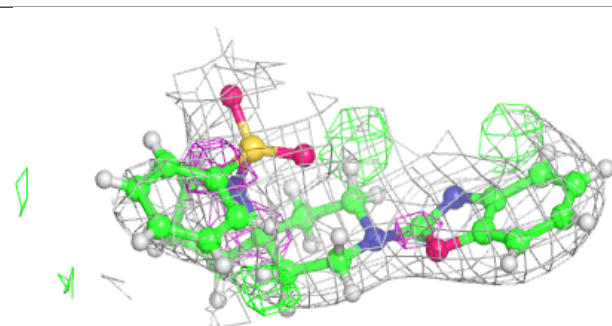
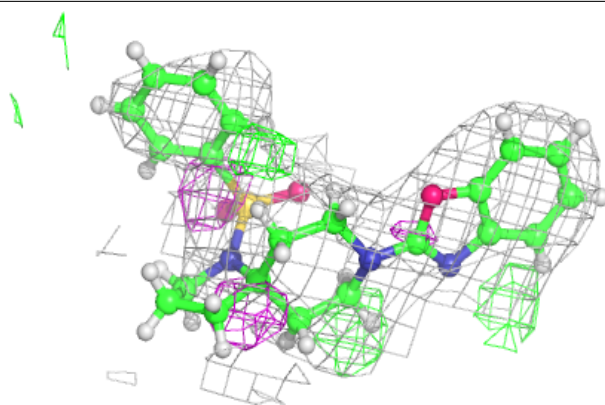
**Electron density around SOG B 409:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

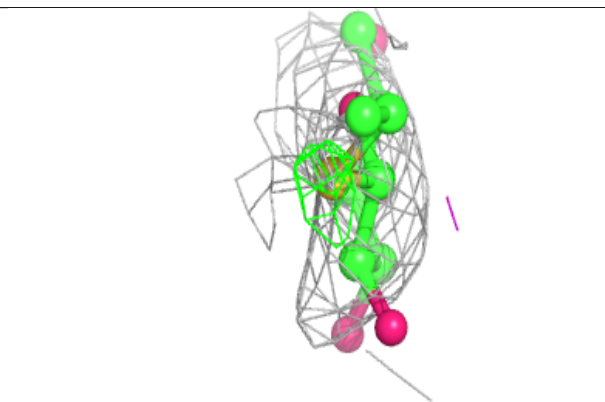
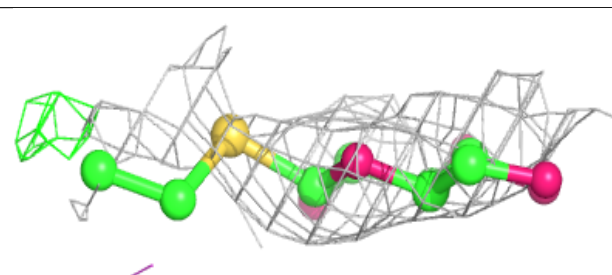
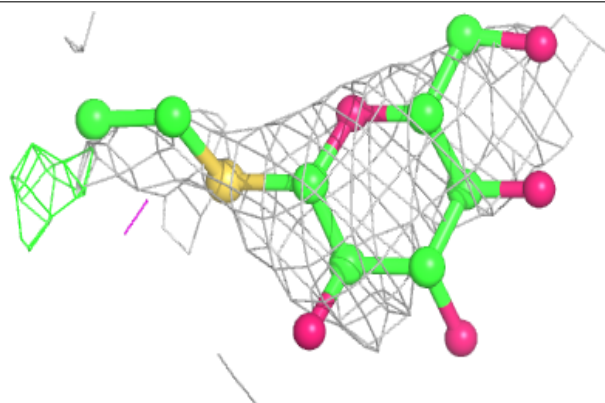


**Electron density around NV8 B 401:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

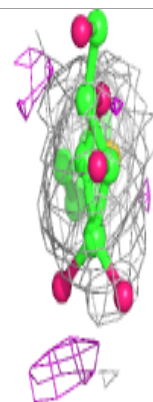
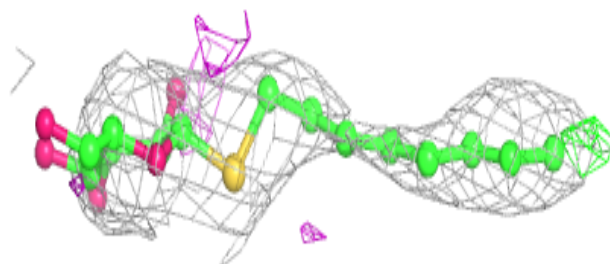
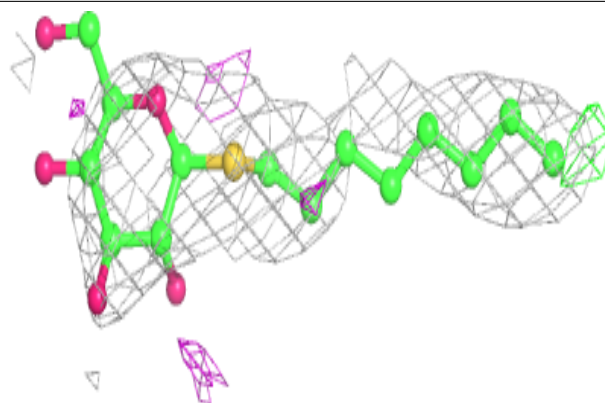
**Electron density around SOG B 405:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

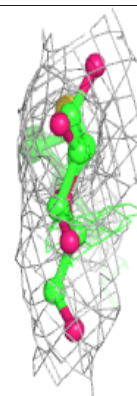
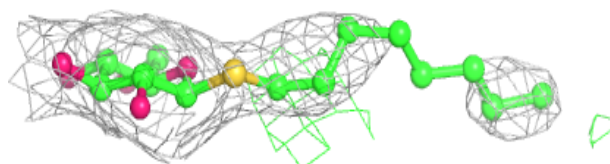
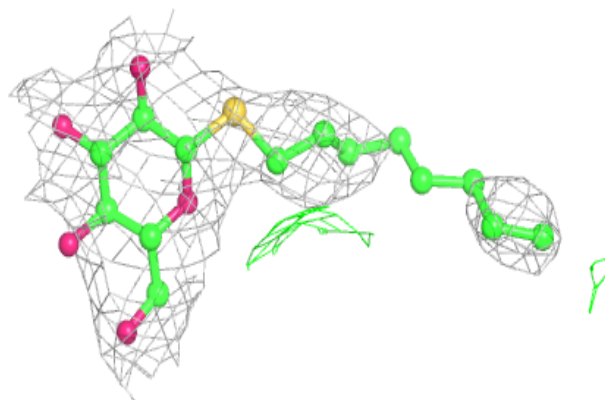


**Electron density around SOG A 407:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

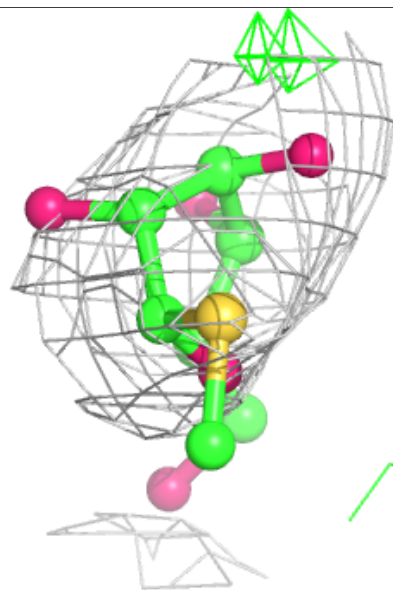
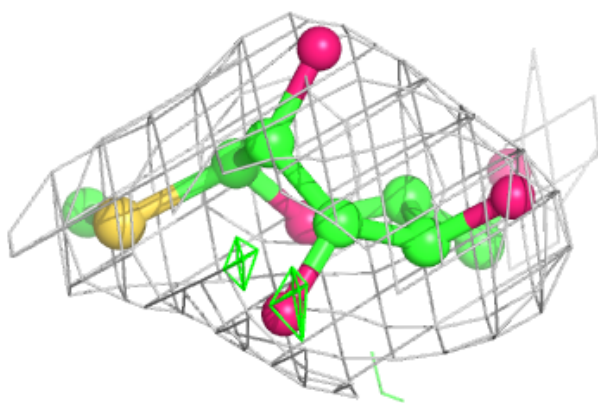
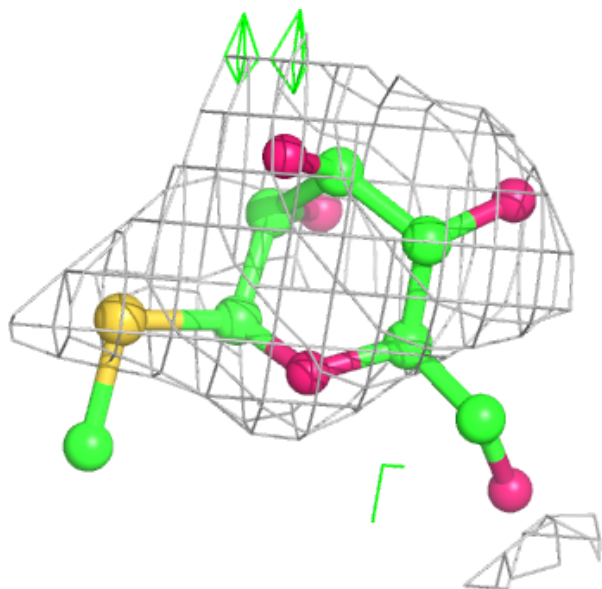
**Electron density around SOG A 406:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



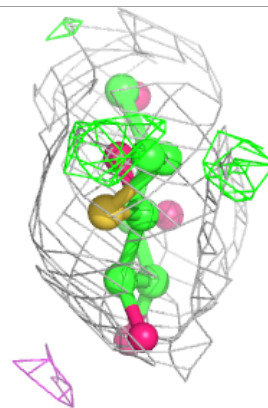
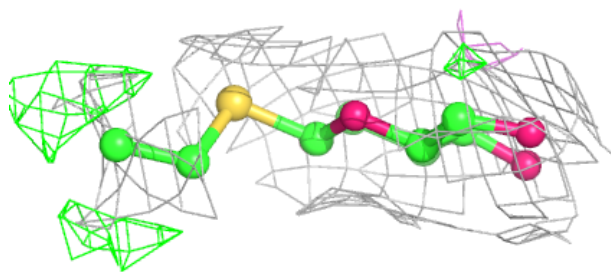
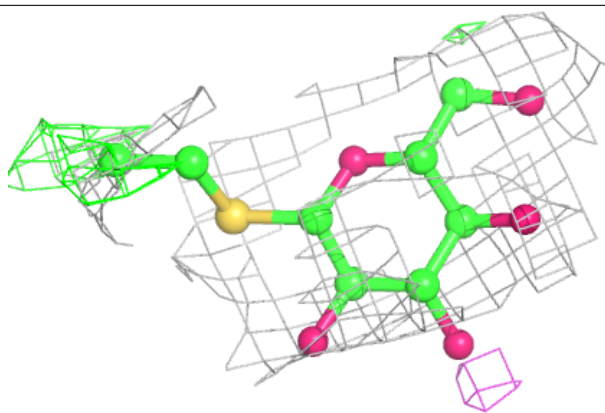
**Electron density around SOG A 408:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

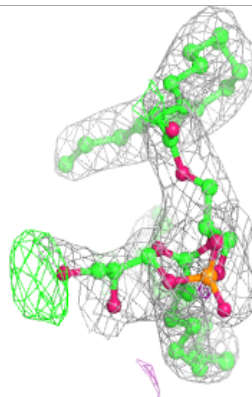
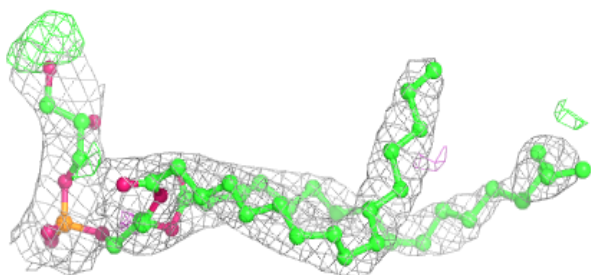
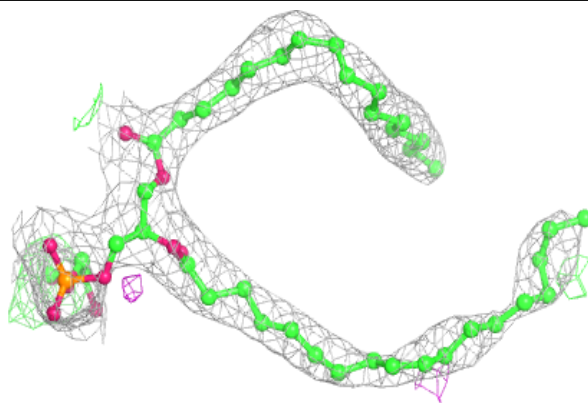


**Electron density around SOG A 409:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around PGW B 404:**

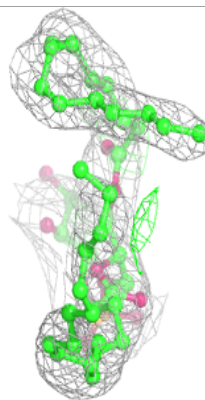
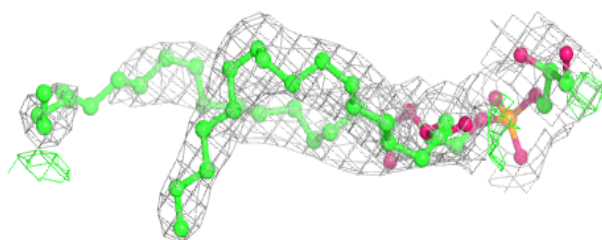
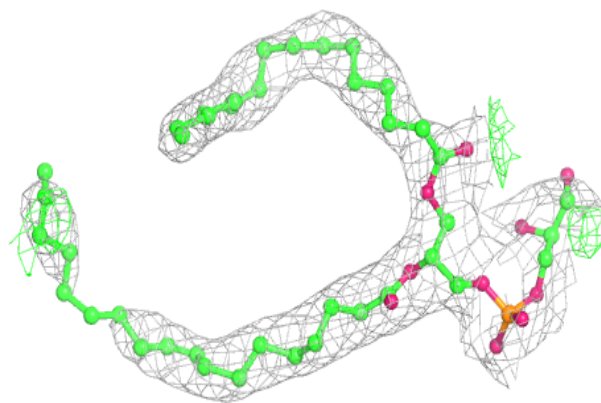
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



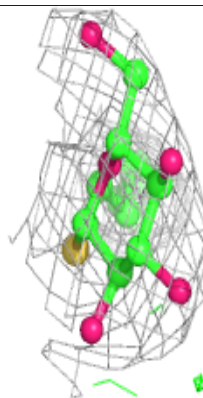
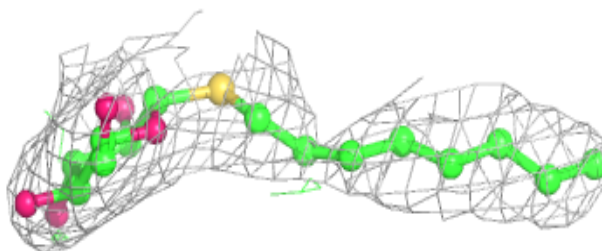
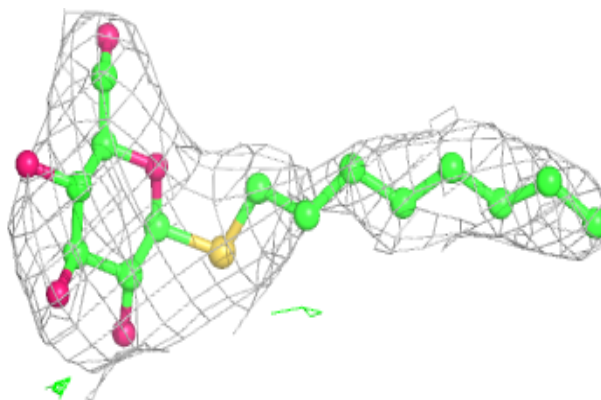


**Electron density around PGW A 404:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

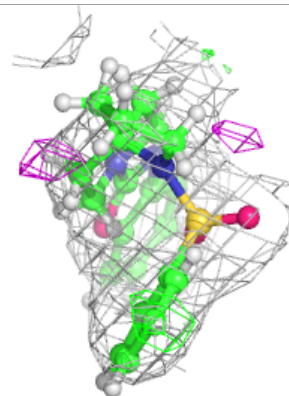
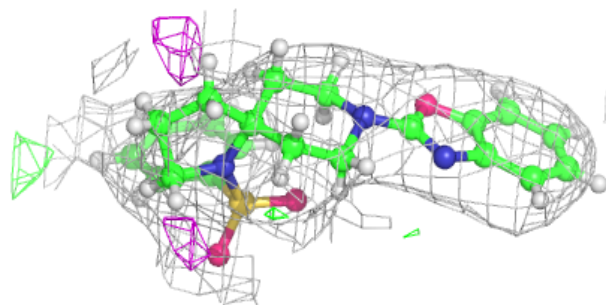
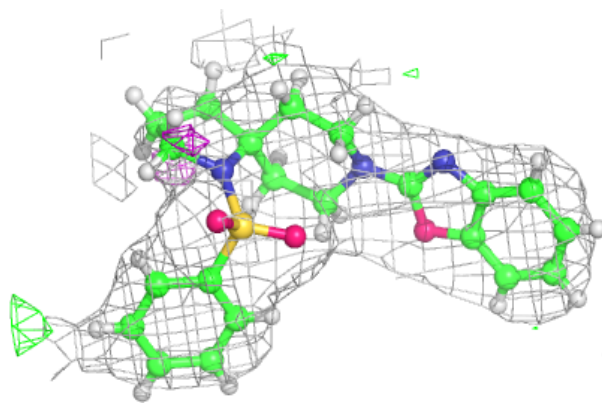
**Electron density around SOG A 410:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

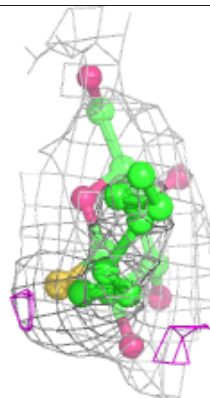
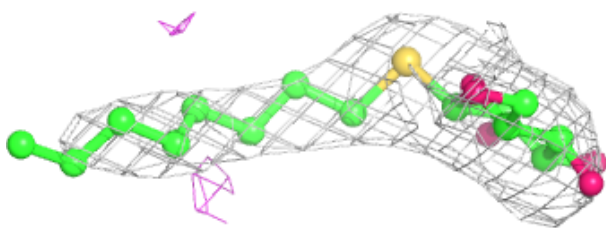
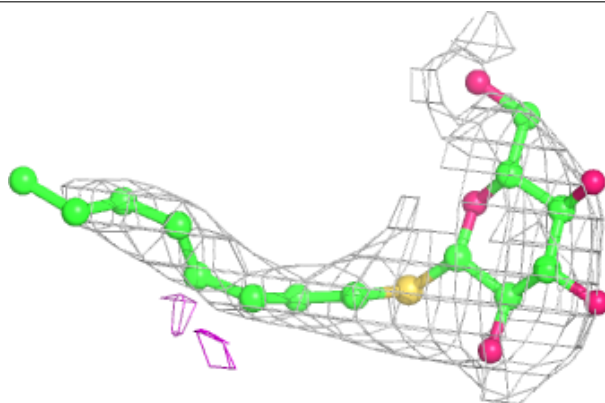


**Electron density around NV8 A 401:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around SOG A 405:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers

There are no such residues in this entry.