



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 4, 2023 – 10:51 PM EDT

PDB ID : 6VVT
Title : Crystal structure of a Mycobacterium smegmatis transcription initiation complex with Rifampicin-resistant RNA polymerase and antibiotic Sorangicin
Authors : Lilic, M.; Darst, S.A.; Campbell, E.A.
Deposited on : 2020-02-18
Resolution : 2.90 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : **FAILED**
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : **FAILED**
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

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 2.90 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 23757 atoms, of which 10 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 RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	J	84	672	421	122	127	2	0	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	222	1641	1040	277	321	3	0	0	0
2	B	225	1616	1018	283	313	2	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	883	6474	4060	1130	1255	29	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	447	LEU	SER	variant	UNP P60281

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	1200	9053	5669	1623	1722	39	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	E	82	613	388	106	119	0	0	0

- Molecule 6 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	303	2393	1501	433	452	7	0	0	0

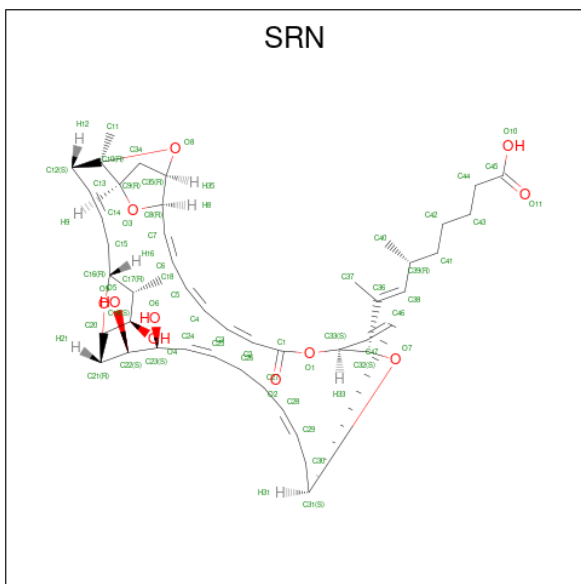
- Molecule 7 is a DNA chain called DNA (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
7	O	31	635	306	114	185	30	0	0	0

- Molecule 8 is a DNA chain called DNA (26-MER).

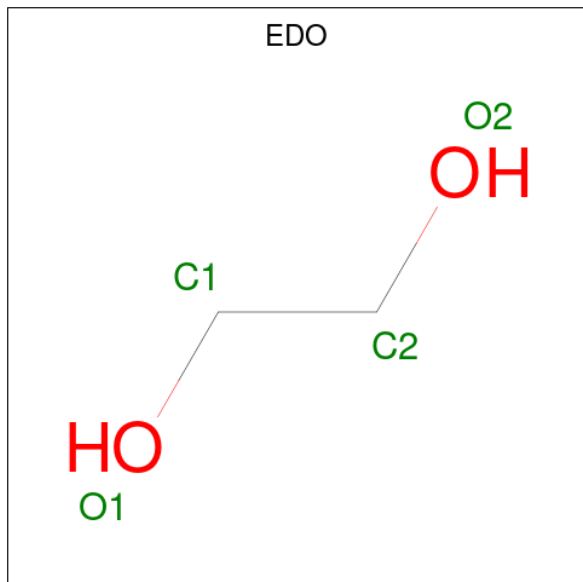
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
8	P	26	526	254	94	153	25	0	0	0

- Molecule 9 is SORANGICIN A (three-letter code: SRN) (formula: $C_{47}H_{66}O_{11}$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
9	C	1	58	47	11	0	0

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
10	C	1	10	2	6	2	0	0

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
11	D	2	2	2	0	0

- Molecule 12 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	D	1	Total O S 5 4 1	0	0
12	D	1	Total O S 5 4 1	0	0
12	D	1	Total O S 5 4 1	0	0
12	D	1	Total O S 5 4 1	0	0
12	D	1	Total O S 5 4 1	0	0
12	F	1	Total O S 5 4 1	0	0
12	F	1	Total O S 5 4 1	0	0
12	F	1	Total O S 5 4 1	0	0
12	F	1	Total O S 5 4 1	0	0

- Molecule 13 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	J	1	Total O 1 1	0	0
13	B	1	Total O 1 1	0	0
13	C	2	Total H O 4 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	H	O		
13	D	8	10	2	8	0	0
13	F	1	1	1		0	0
13	O	1	1	1		0	0
13	P	1	1	1		0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	129.22Å 162.21Å 136.29Å 90.00° 111.36° 90.00°	Depositor
Resolution (Å)	49.75 – 2.90	Depositor
% Data completeness (in resolution range)	82.6 (49.75-2.90)	Depositor
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.99 (at 2.91Å)	Xtrriage
Refinement program	PHENIX v0	Depositor
R, R_{free}	0.240 , 0.275	Depositor
Wilson B-factor (Å ²)	101.7	Xtrriage
Anisotropy	0.176	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	23757	wwPDB-VP
Average B, all atoms (Å ²)	96.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

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

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

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 2 are monoatomic - leaving 11 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
12	SO4	D	2007	-	4,4,4	0.14	0	6,6,6	0.05	0
12	SO4	F	502	-	4,4,4	0.14	0	6,6,6	0.05	0
9	SRN	C	1201	-	60,62,62	5.11	31 (51%)	62,84,84	1.60	11 (17%)
12	SO4	F	501	-	4,4,4	0.13	0	6,6,6	0.09	0
12	SO4	D	2003	-	4,4,4	0.13	0	6,6,6	0.09	0
10	EDO	C	1202	-	3,3,3	0.45	0	2,2,2	0.31	0
12	SO4	D	2004	-	4,4,4	0.14	0	6,6,6	0.05	0
12	SO4	F	504	-	4,4,4	0.13	0	6,6,6	0.05	0
12	SO4	D	2005	-	4,4,4	0.14	0	6,6,6	0.05	0
12	SO4	D	2006	-	4,4,4	0.14	0	6,6,6	0.05	0
12	SO4	F	503	-	4,4,4	0.14	0	6,6,6	0.06	0

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
10	EDO	C	1202	-	-	0/1/1/1	-
9	SRN	C	1201	-	-	8/52/105/105	0/4/5/5

All (31) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	C	1201	SRN	O3-C8	19.93	1.80	1.44
9	C	1201	SRN	O3-C9	16.27	1.81	1.45
9	C	1201	SRN	C34-C9	-11.11	1.28	1.52
9	C	1201	SRN	C3-C2	9.61	1.59	1.34
9	C	1201	SRN	C34-C35	-8.87	1.33	1.52
9	C	1201	SRN	C4-C5	8.26	1.57	1.36
9	C	1201	SRN	C13-C14	6.77	1.60	1.31
9	C	1201	SRN	C24-C25	6.22	1.58	1.31
9	C	1201	SRN	C6-C7	6.04	1.57	1.33
9	C	1201	SRN	C8-C7	-5.91	1.30	1.50
9	C	1201	SRN	C12-C13	-5.58	1.31	1.50
9	C	1201	SRN	O7-C32	5.56	1.52	1.43
9	C	1201	SRN	C10-C9	-5.23	1.43	1.53
9	C	1201	SRN	C28-C29	5.16	1.61	1.31

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	C	1201	SRN	C3-C4	4.86	1.58	1.44
9	C	1201	SRN	C5-C6	4.75	1.57	1.44
9	C	1201	SRN	O8-C12	4.31	1.51	1.43
9	C	1201	SRN	C35-C8	-4.17	1.46	1.53
9	C	1201	SRN	C11-C10	-4.17	1.44	1.53
9	C	1201	SRN	O1-C1	3.93	1.42	1.34
9	C	1201	SRN	O1-C33	3.81	1.52	1.46
9	C	1201	SRN	C46-C47	3.79	1.44	1.33
9	C	1201	SRN	C30-C29	3.70	1.61	1.50
9	C	1201	SRN	C2-C1	3.51	1.56	1.48
9	C	1201	SRN	O7-C31	3.44	1.53	1.44
9	C	1201	SRN	C33-C47	-3.40	1.44	1.49
9	C	1201	SRN	C23-C24	3.26	1.55	1.50
9	C	1201	SRN	C15-C14	2.73	1.58	1.50
9	C	1201	SRN	C31-C46	2.48	1.54	1.50
9	C	1201	SRN	C22-C21	2.39	1.56	1.52
9	C	1201	SRN	O8-C35	2.07	1.49	1.44

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	C	1201	SRN	O9-C21-C20	4.37	114.47	109.94
9	C	1201	SRN	C8-C7-C6	-4.01	117.75	125.61
9	C	1201	SRN	C6-C5-C4	-3.30	117.46	124.81
9	C	1201	SRN	C9-C34-C35	3.17	112.89	103.73
9	C	1201	SRN	C37-C36-C32	2.84	120.58	115.68
9	C	1201	SRN	C32-C33-C47	2.79	113.48	110.64
9	C	1201	SRN	C15-C16-C17	-2.59	107.65	113.67
9	C	1201	SRN	C43-C44-C45	-2.21	108.91	114.47
9	C	1201	SRN	C30-C29-C28	-2.04	118.81	126.40
9	C	1201	SRN	C12-C13-C14	-2.01	118.79	124.60
9	C	1201	SRN	C11-C10-C9	-2.00	108.26	111.96

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	C	1201	SRN	C25-C26-C27-C28
9	C	1201	SRN	C14-C15-C16-O9
9	C	1201	SRN	C42-C43-C44-C45
9	C	1201	SRN	C39-C41-C42-C43
9	C	1201	SRN	C14-C15-C16-C17

Continued on next page...

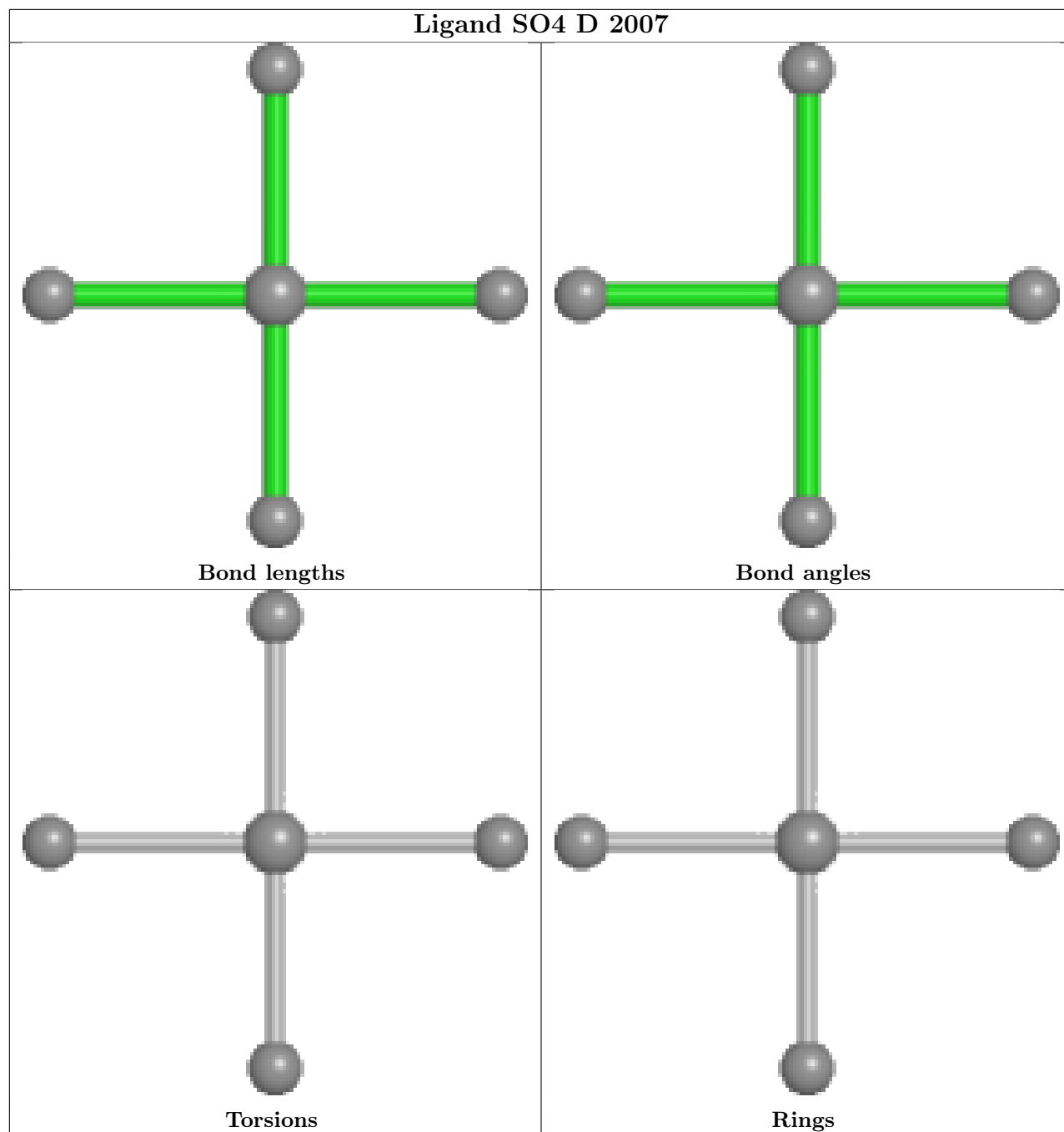
Continued from previous page...

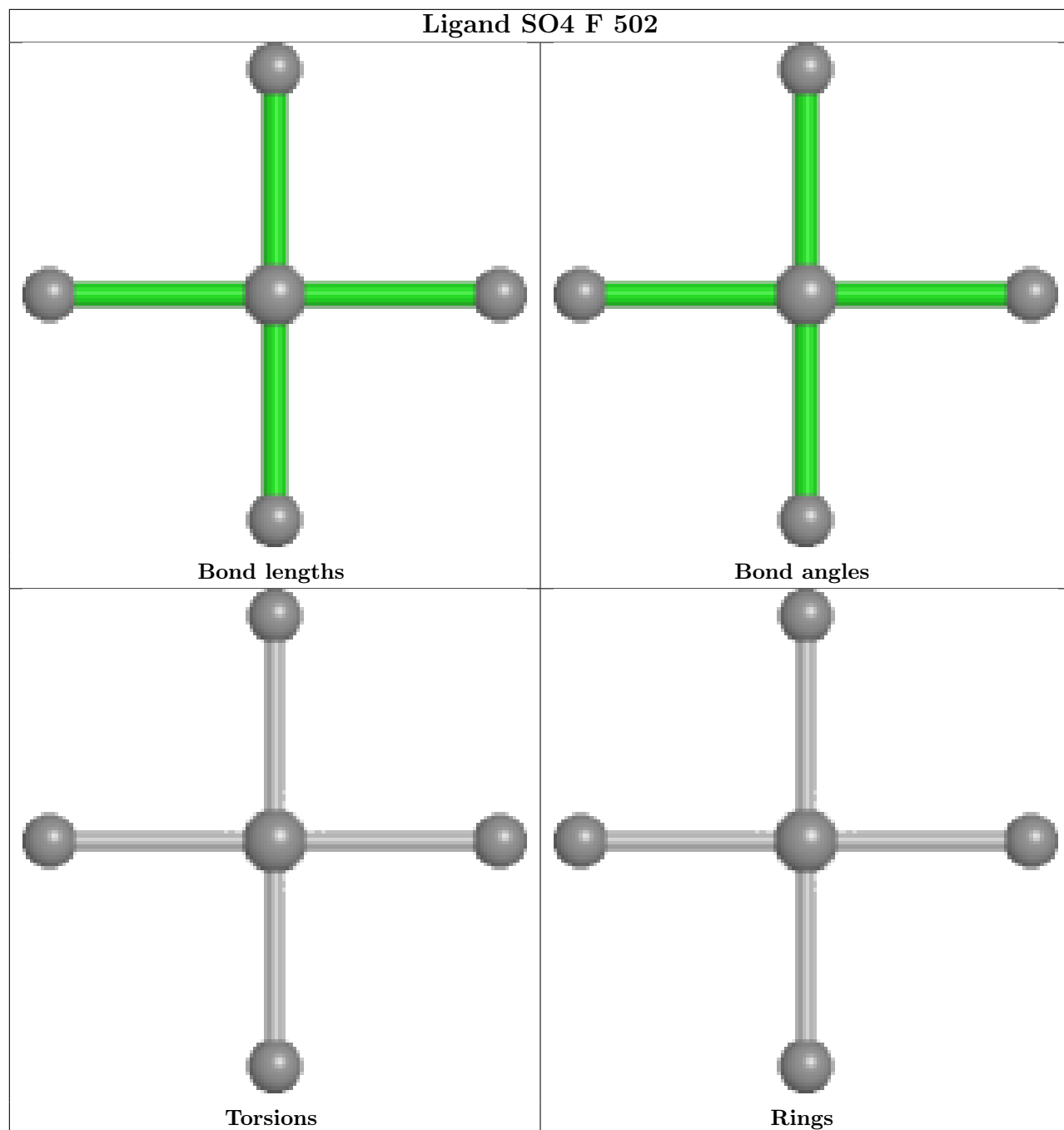
Mol	Chain	Res	Type	Atoms
9	C	1201	SRN	O2-C1-O1-C33
9	C	1201	SRN	C43-C44-C45-O10
9	C	1201	SRN	C43-C44-C45-O11

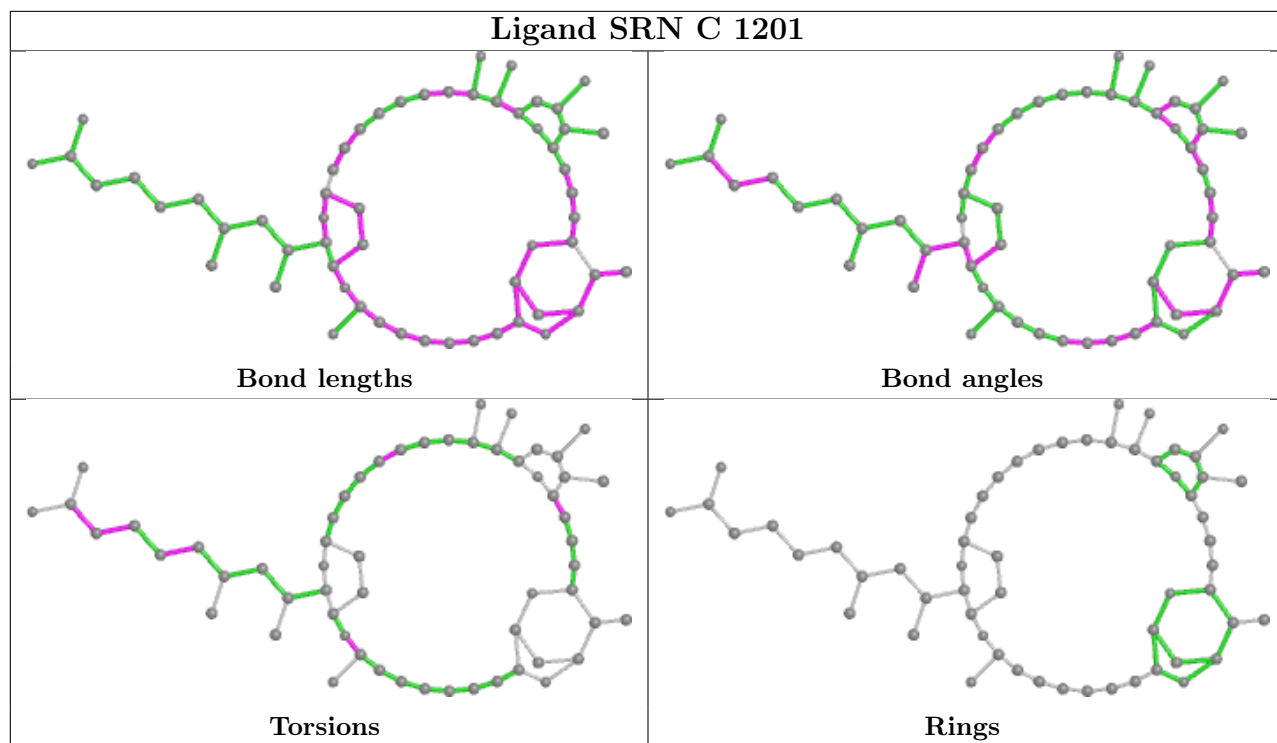
There are no ring outliers.

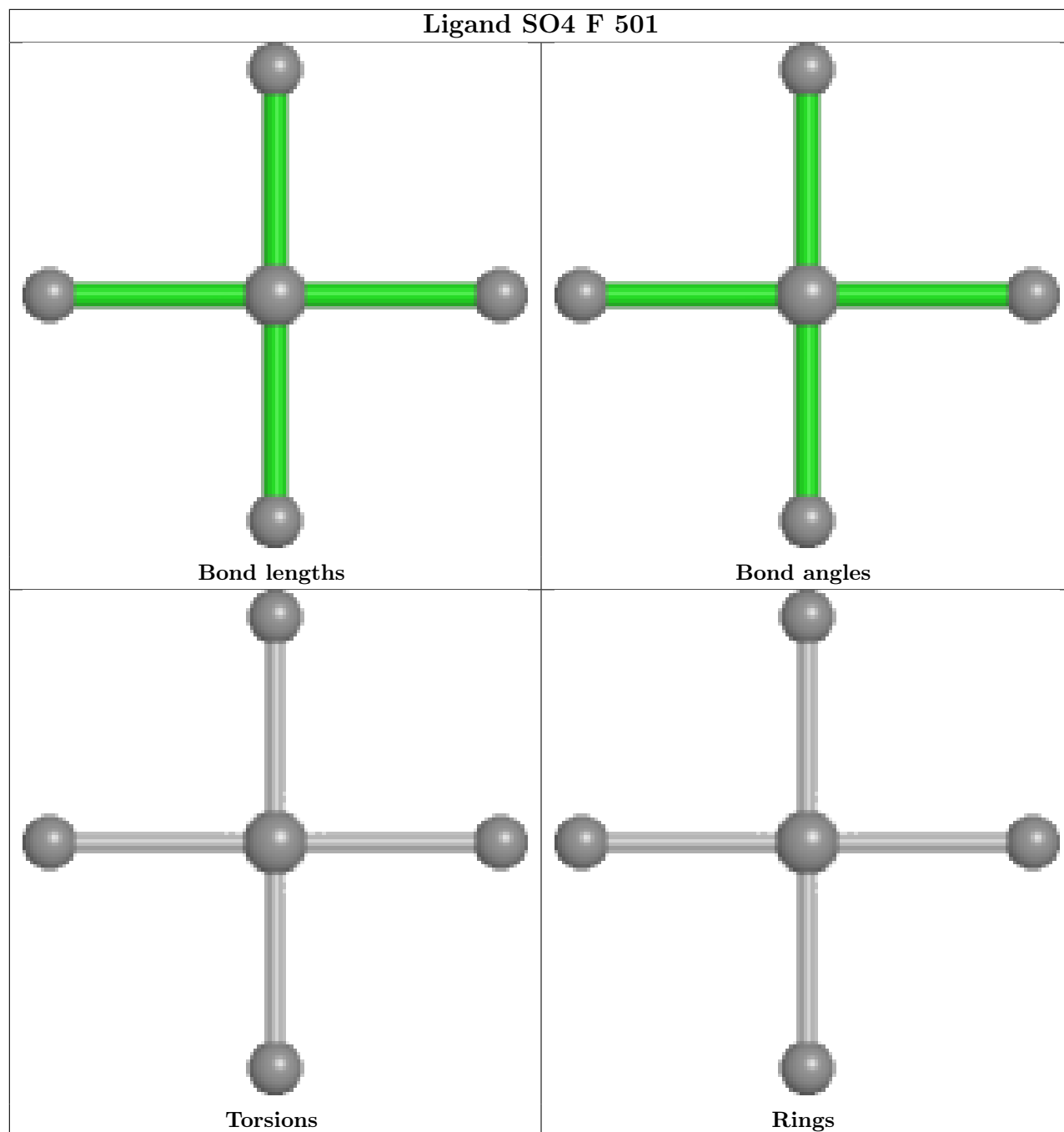
No monomer is involved in short contacts.

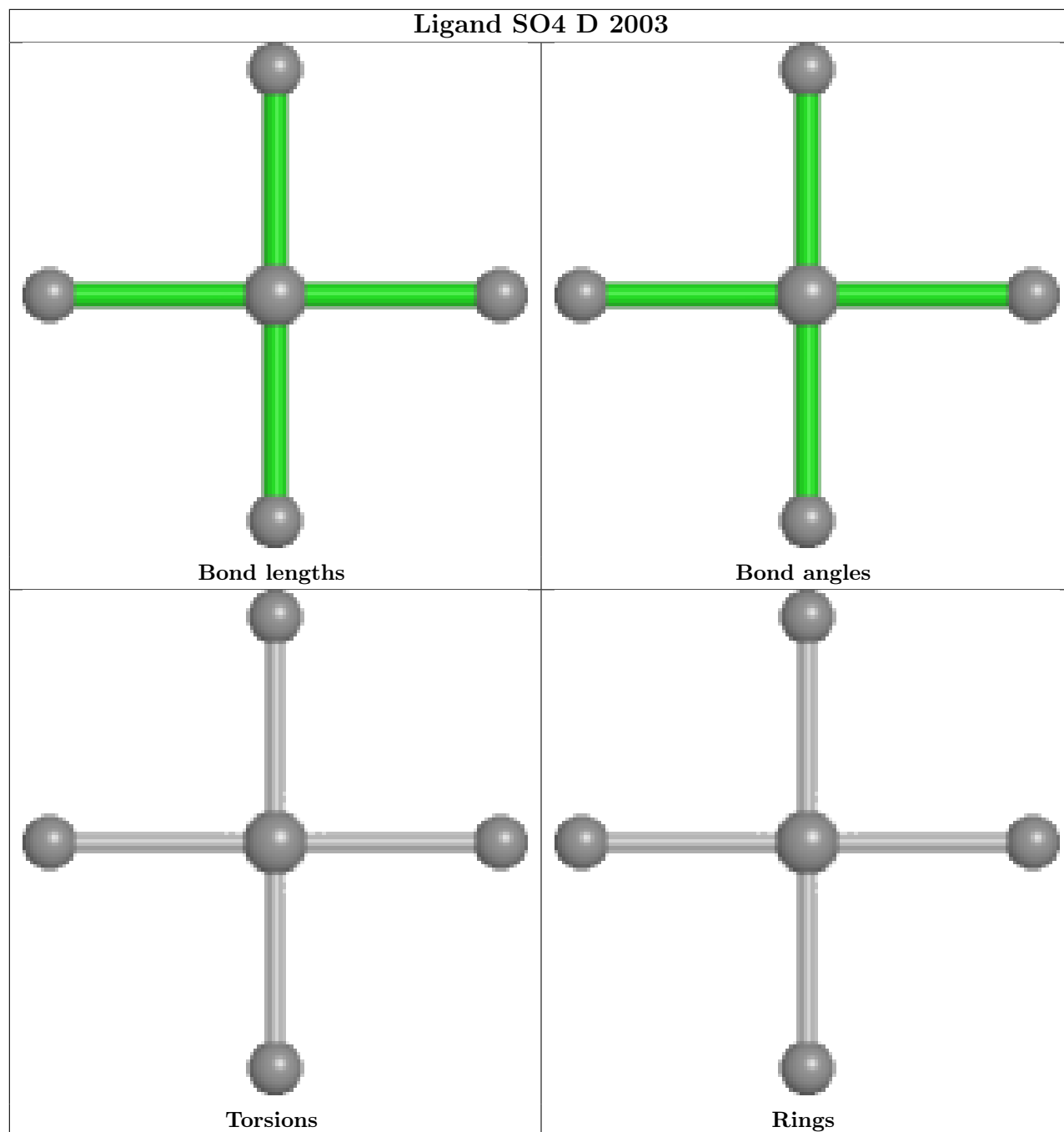
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

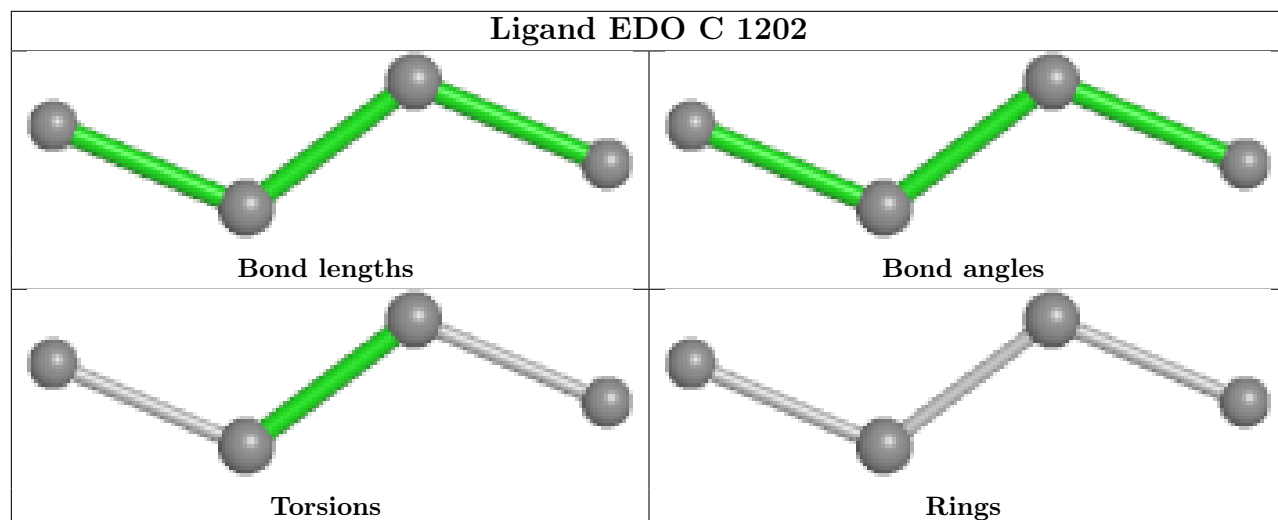


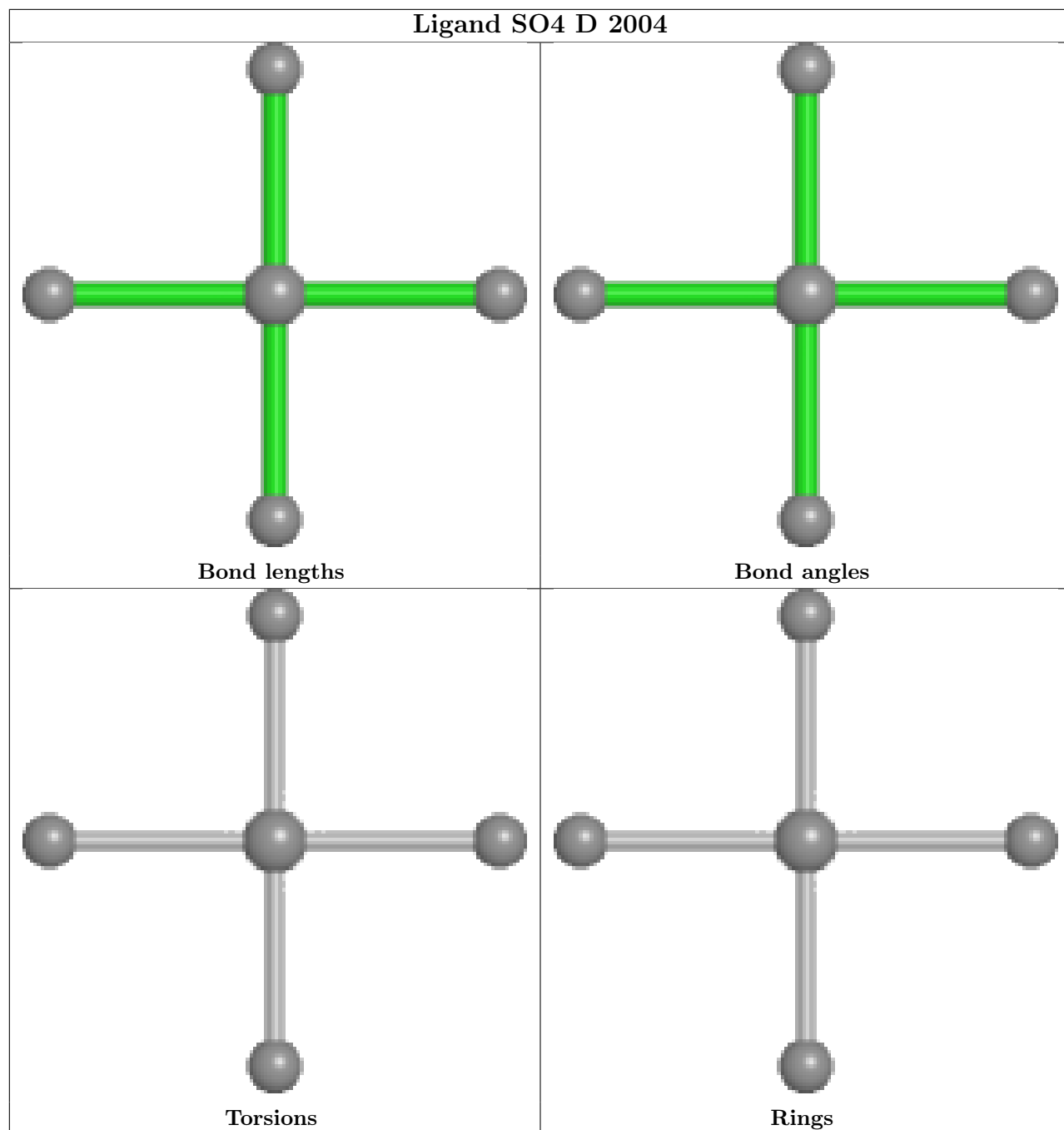


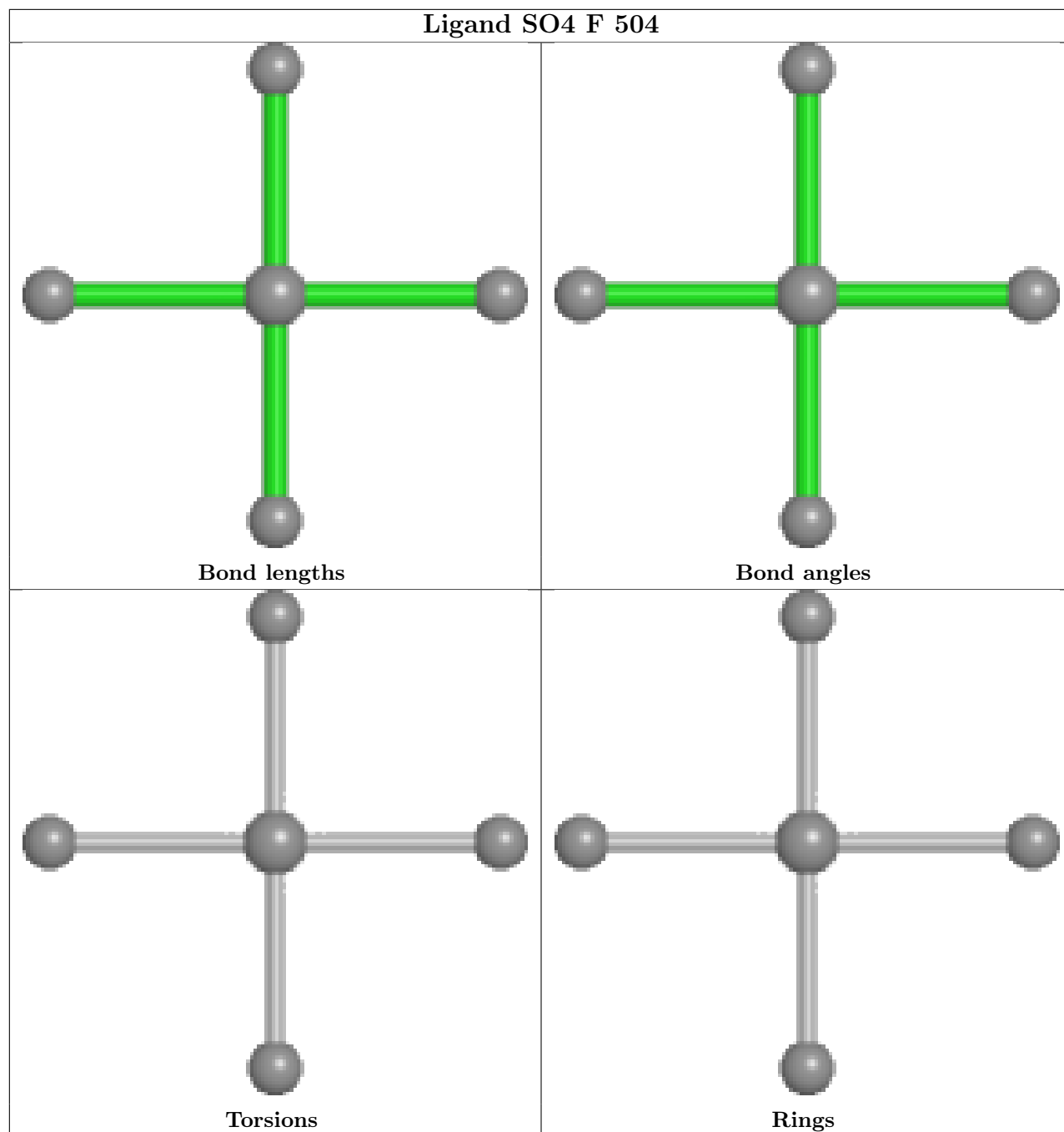


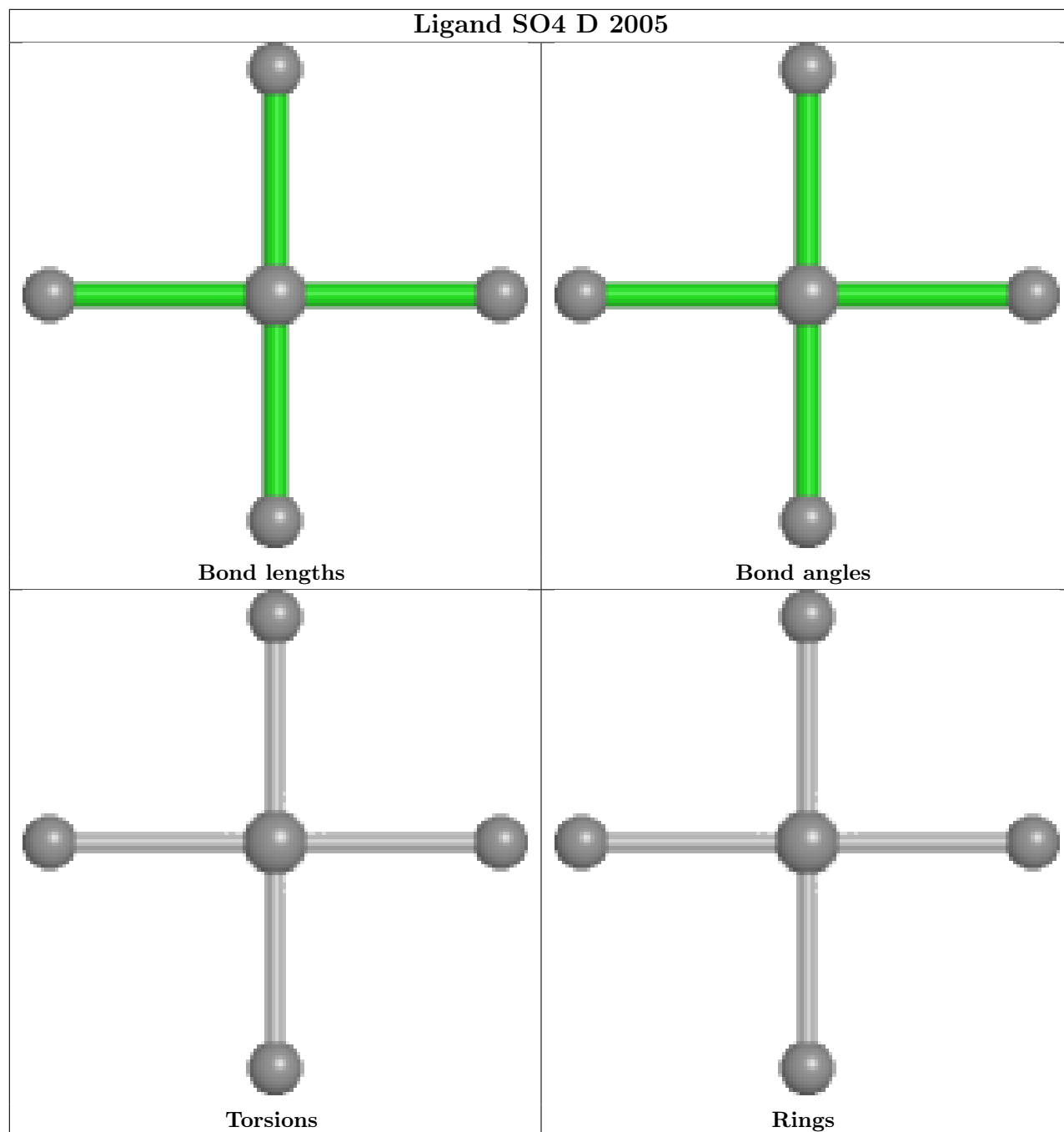


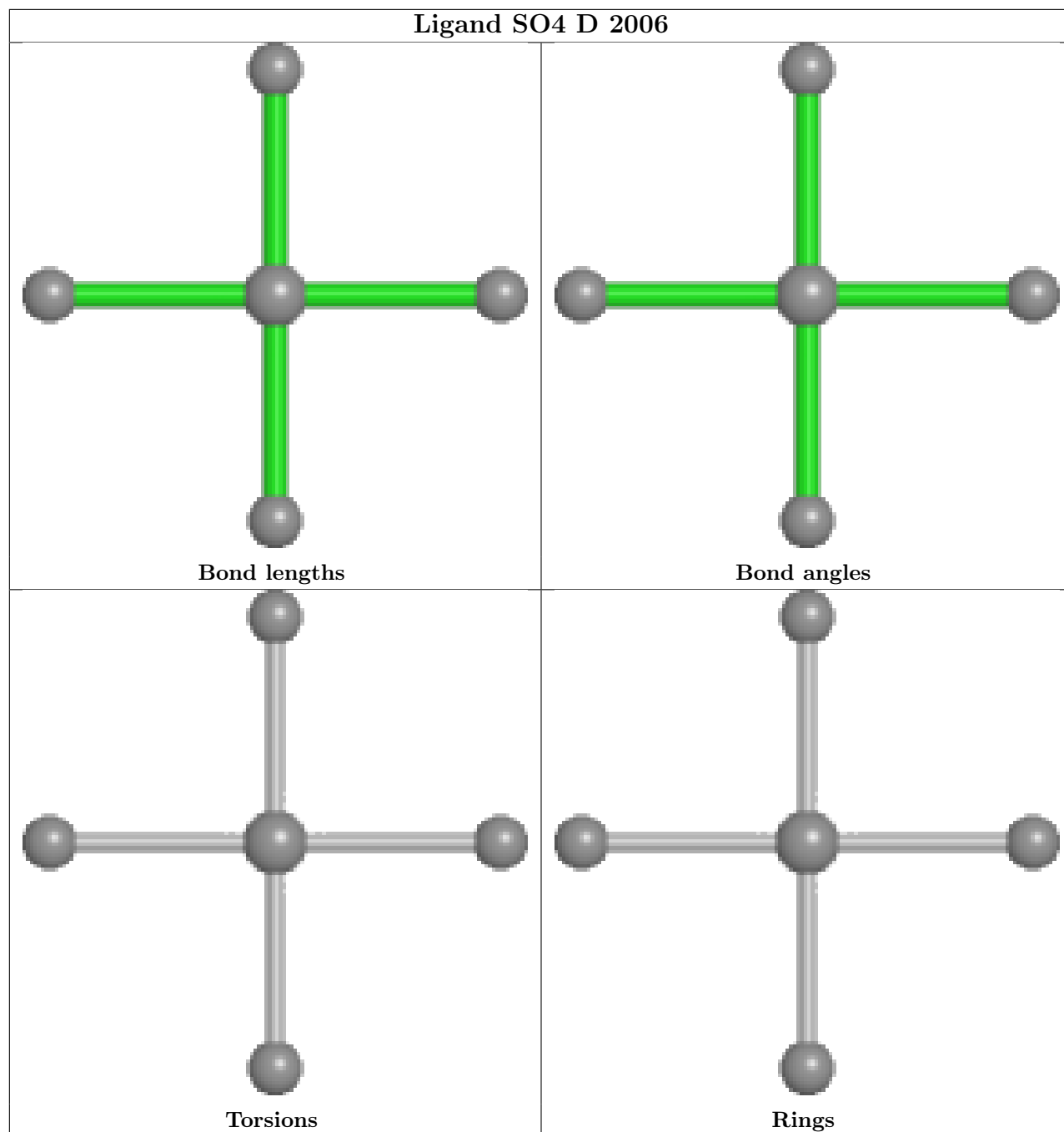


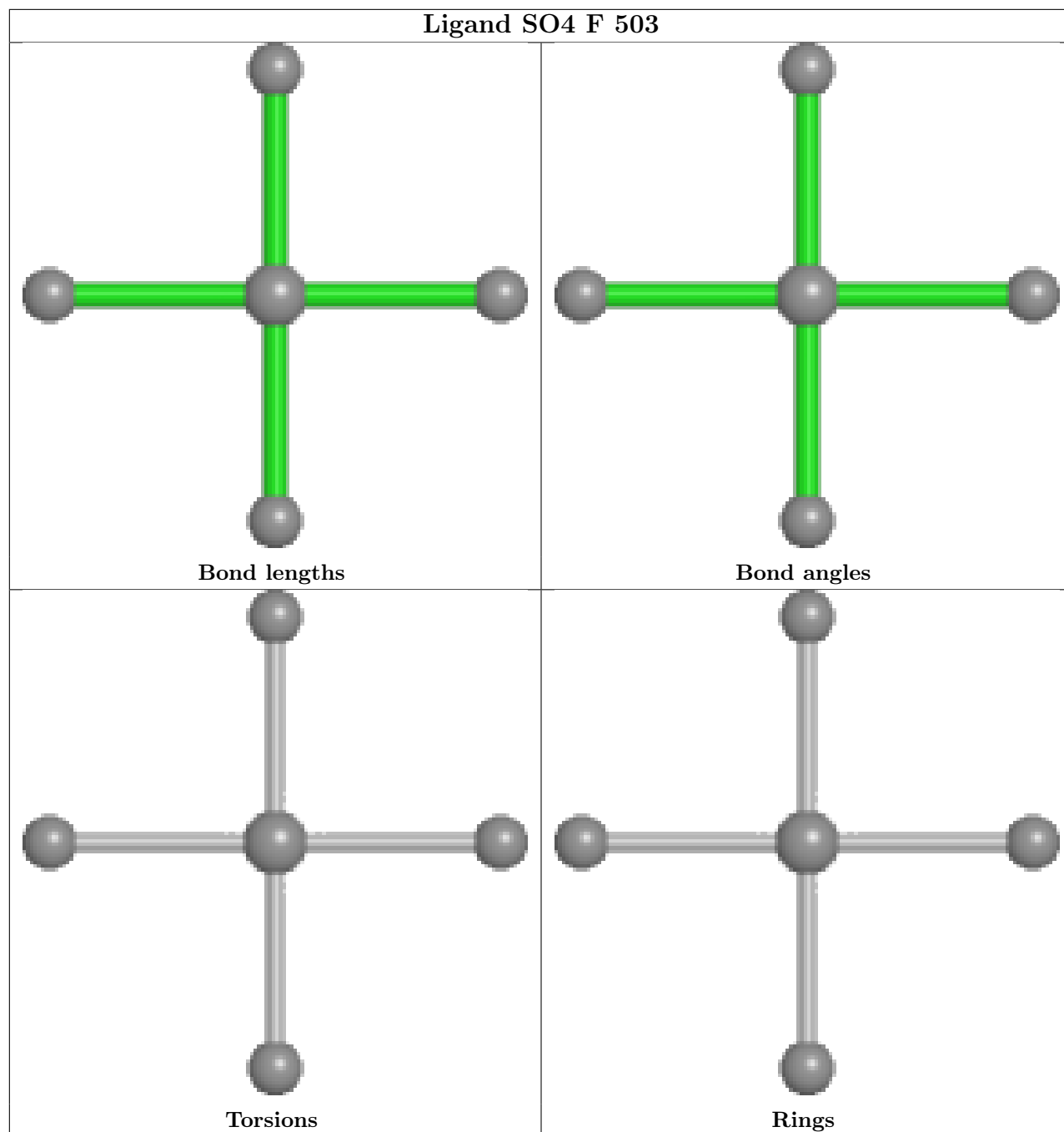












4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.