W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report (i)

Oct 10, 2023 – 06:46 AM EDT

PDB ID	:	7KK2
Title	:	Structure of the catalytic domain of PARP1
Authors	:	Gajiwala, K.S.; Ryan, K.
Deposited on	:	2020-10-27
Resolution	:	1.70 Å(reported)

The following versions of software and data (see references (i)) 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	:	2.35.1
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)	:	2.35.1

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 1.70 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.
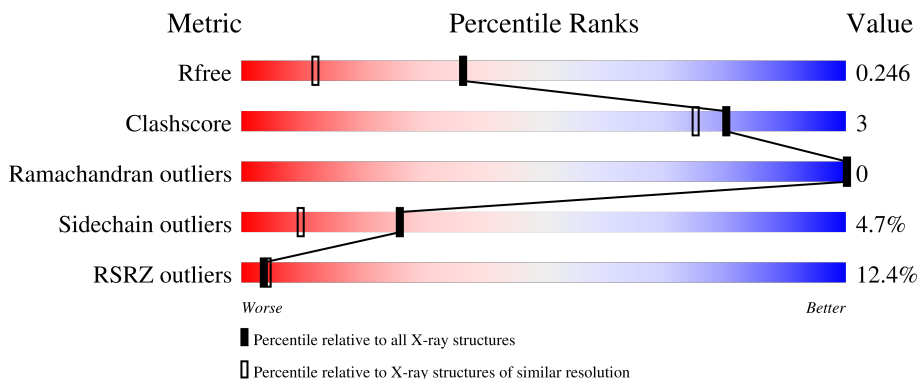
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	352	11% / 89% / 11%
1	B	352	14% / 89% / 9% ··

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 6052 atoms, of which 0 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 Poly [ADP-ribose] polymerase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	351	Total	C	N	O	S	0	0	0
			2759	1755	466	527	11			
1	B	349	Total	C	N	O	S	0	0	0
			2746	1748	464	523	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	660	GLY	-	expression tag	UNP P09874
A	661	SER	-	expression tag	UNP P09874
B	660	GLY	-	expression tag	UNP P09874
B	661	SER	-	expression tag	UNP P09874

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	276	Total 276	O 276	0	0
3	B	236	Total 236	O 236	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Poly [ADP-ribose] polymerase 1

Chain A:

• Molecule 1: Poly [ADP-ribose] polymerase 1

Chain B:
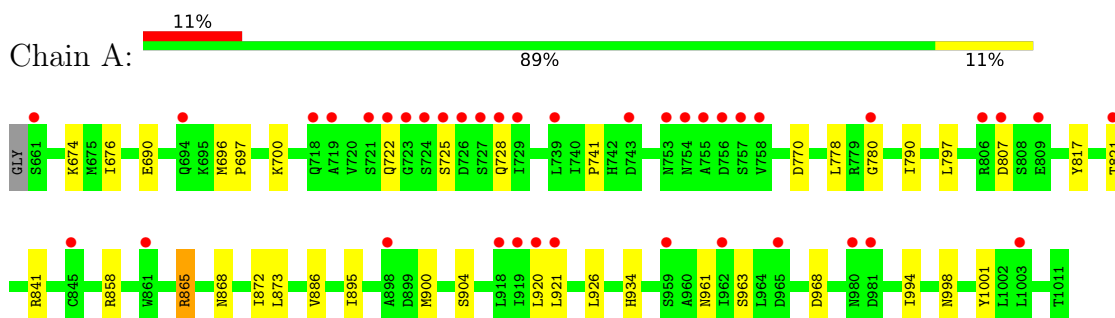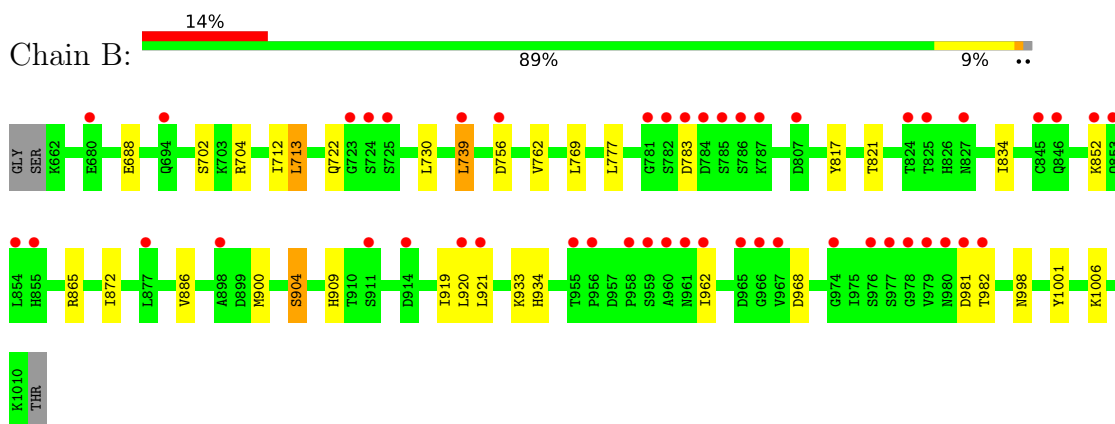
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	48.06Å 91.64Å 162.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 1.70	Depositor
	24.58 – 1.69	EDS
% Data completeness	72.2 (25.00-1.70)	Depositor
(in resolution range)	72.2 (24.58-1.69)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.46 (at 1.70Å)	Xtriage
Refinement program	BUSTER 2.11.7 (3-OCT-2019)	Depositor
R, R_{free}	0.204 , 0.238	Depositor
	0.210 , 0.246	DCC
R_{free} test set	2833 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	27.9	Xtriage
Anisotropy	0.067	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 41.6	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6052	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.42	0/2811	0.60	0/3794
1	B	0.39	0/2798	0.57	0/3776
All	All	0.41	0/5609	0.59	0/7570

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	2759	0	2800	20	0
1	B	2746	0	2788	12	0
2	A	25	0	0	0	0
2	B	10	0	0	0	0
3	A	276	0	0	2	0
3	B	236	0	0	1	0
All	All	6052	0	5588	30	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 3.

All (30) 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:780:GLY:O	3:A:9101:HOH:O	2.17	0.59
1:A:841:ARG:HD2	1:A:873:LEU:O	2.04	0.57
1:A:696:MET:HB2	1:A:741:PRO:HG2	1.87	0.55
1:A:886:VAL:HG11	1:B:934:HIS:CD2	2.42	0.54
1:A:770:ASP:HB2	1:A:868:ASN:ND2	2.23	0.53
1:A:770:ASP:HB2	1:A:868:ASN:HD22	1.74	0.52
1:B:834:ILE:HD11	1:B:1006:LYS:HB2	1.93	0.50
1:B:821:THR:HB	1:B:900:MET:HA	1.95	0.49
1:A:821:THR:HB	1:A:900:MET:HA	1.95	0.49
1:B:921:LEU:HB2	1:B:1001:TYR:HB2	1.95	0.49
1:A:725:SER:OG	1:A:728:GLN:NE2	2.46	0.49
1:A:676:ILE:HD11	1:A:797:LEU:HD11	1.95	0.48
1:A:858:ARG:HG3	1:A:926:LEU:HD12	1.95	0.48
1:A:934:HIS:HD2	3:B:9113:HOH:O	1.96	0.47
1:B:904:SER:HB3	1:B:919:ILE:HD11	1.96	0.47
1:A:865:ARG:HG2	3:A:9307:HOH:O	2.15	0.46
1:A:921:LEU:HB2	1:A:1001:TYR:HB2	1.97	0.46
1:A:697:PRO:HD2	1:A:700:LYS:HB2	1.98	0.45
1:A:674:LYS:HG3	1:A:790:ILE:CD1	2.46	0.45
1:A:934:HIS:CD2	1:B:886:VAL:HG11	2.52	0.45
1:A:895:ILE:HD11	1:A:994:ILE:HG22	1.99	0.44
1:B:712:ILE:HD12	1:B:739:LEU:HD12	2.00	0.44
1:B:872:ILE:HG21	1:B:920:LEU:HD11	1.99	0.43
1:A:674:LYS:HG3	1:A:790:ILE:HD11	2.01	0.43
1:A:865:ARG:HG2	1:A:865:ARG:H	1.69	0.42
1:B:865:ARG:CZ	1:B:909:HIS:CD2	3.03	0.42
1:B:933:LYS:HB3	1:B:982:THR:HG22	2.01	0.42
1:A:872:ILE:HG21	1:A:920:LEU:HD11	2.02	0.41
1:B:865:ARG:NE	1:B:909:HIS:CD2	2.88	0.41
1:B:713:LEU:HD23	1:B:762:VAL:HG22	2.02	0.41

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	349/352 (99%)	346 (99%)	3 (1%)	0	100	100
1	B	347/352 (99%)	344 (99%)	3 (1%)	0	100	100
All	All	696/704 (99%)	690 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2　Protein sidechains ⓘ

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	309/309 (100%)	298 (96%)	11 (4%)	35	16
1	B	307/309 (99%)	289 (94%)	18 (6%)	19	6
All	All	616/618 (100%)	587 (95%)	29 (5%)	26	10

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	690	GLU
1	A	722	GLN
1	A	778	LEU
1	A	807	ASP
1	A	817	TYR
1	A	865	ARG
1	A	904	SER
1	A	961	ASN
1	A	963	SER
1	A	968	ASP
1	A	998	ASN
1	B	688	GLU
1	B	702	SER
1	B	704	ARG
1	B	713	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	722	GLN
1	B	730	LEU
1	B	739	LEU
1	B	756	ASP
1	B	769	LEU
1	B	777	LEU
1	B	783	ASP
1	B	817	TYR
1	B	852	LYS
1	B	904	SER
1	B	962	ILE
1	B	968	ASP
1	B	981	ASP
1	B	998	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	728	GLN
1	A	759	GLN
1	A	767	ASN
1	A	868	ASN
1	A	961	ASN
1	A	998	ASN
1	B	722	GLN
1	B	734	ASN
1	B	868	ASN
1	B	909	HIS
1	B	961	ASN
1	B	998	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

7 ligands are modelled in this entry.

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	SO4	A	9002	-	4,4,4	0.15	0	6,6,6	0.07	0
2	SO4	A	9004	-	4,4,4	0.24	0	6,6,6	0.16	0
2	SO4	B	9001	-	4,4,4	0.20	0	6,6,6	0.15	0
2	SO4	A	9001	-	4,4,4	0.25	0	6,6,6	0.37	0
2	SO4	B	9002	-	4,4,4	0.16	0	6,6,6	0.14	0
2	SO4	A	9003	-	4,4,4	0.17	0	6,6,6	0.08	0
2	SO4	A	9005	-	4,4,4	0.21	0	6,6,6	0.16	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

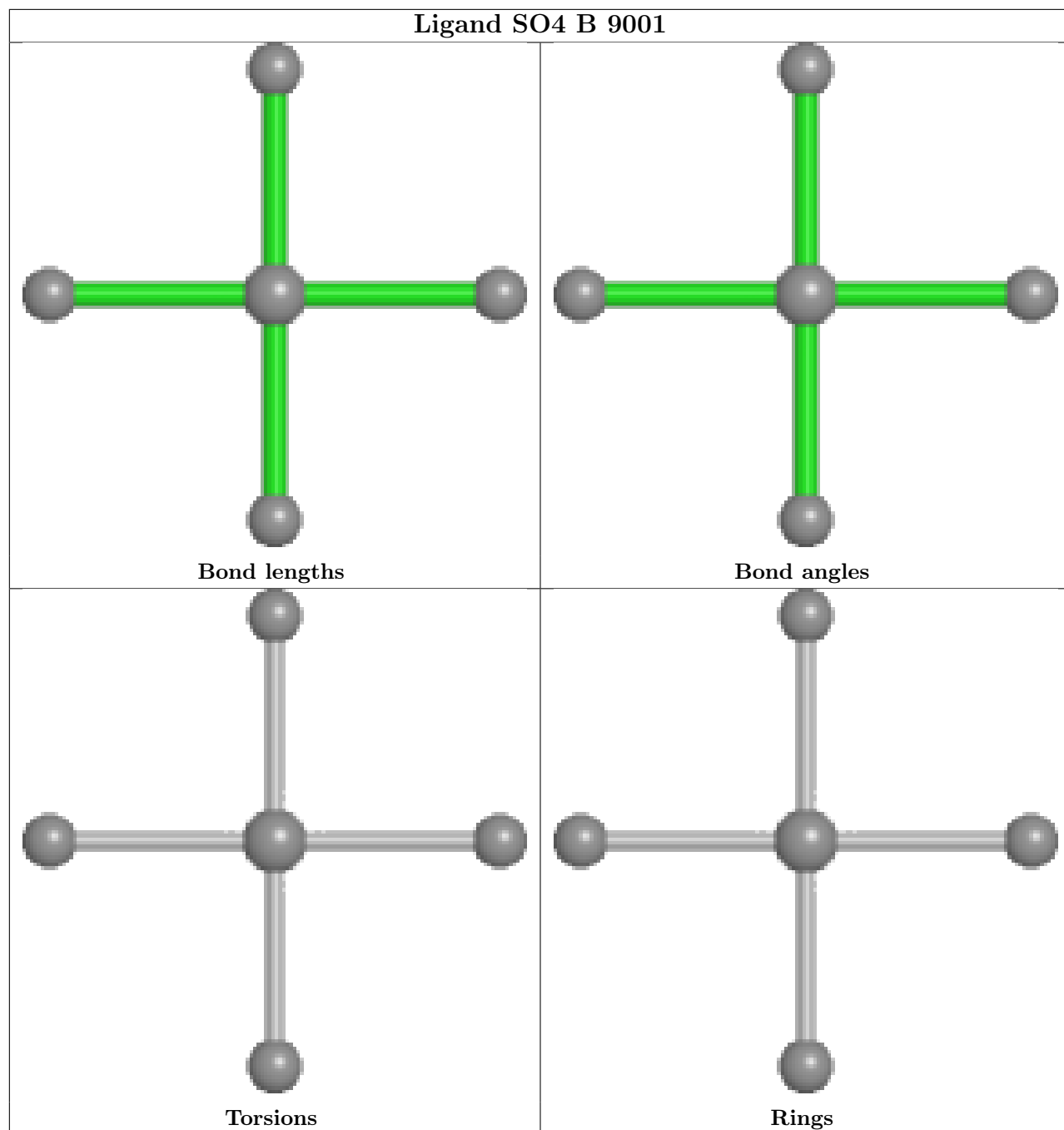There are no torsion outliers.

There are no ring outliers.

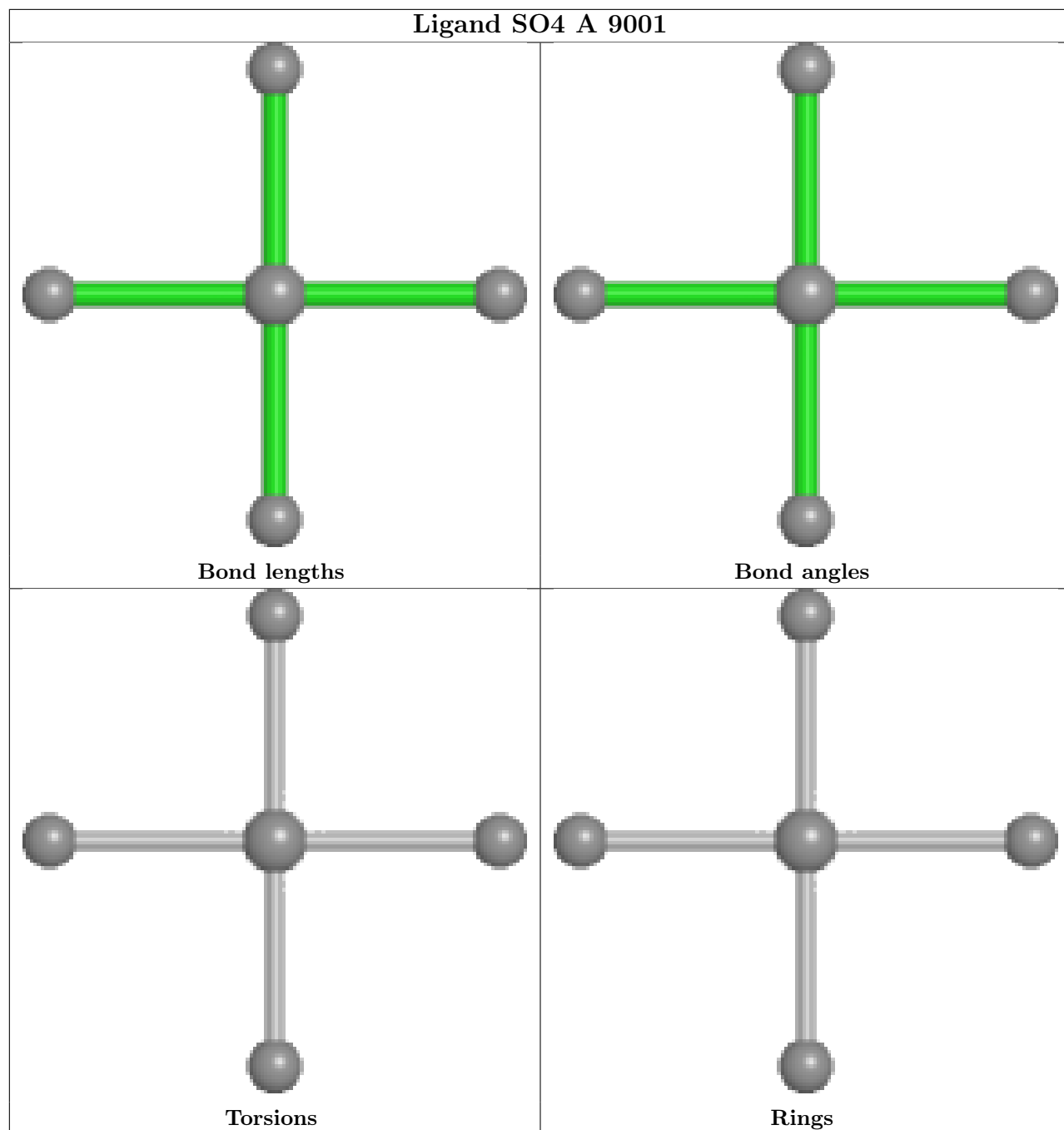No monomer is involved in short contacts.

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 then 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.
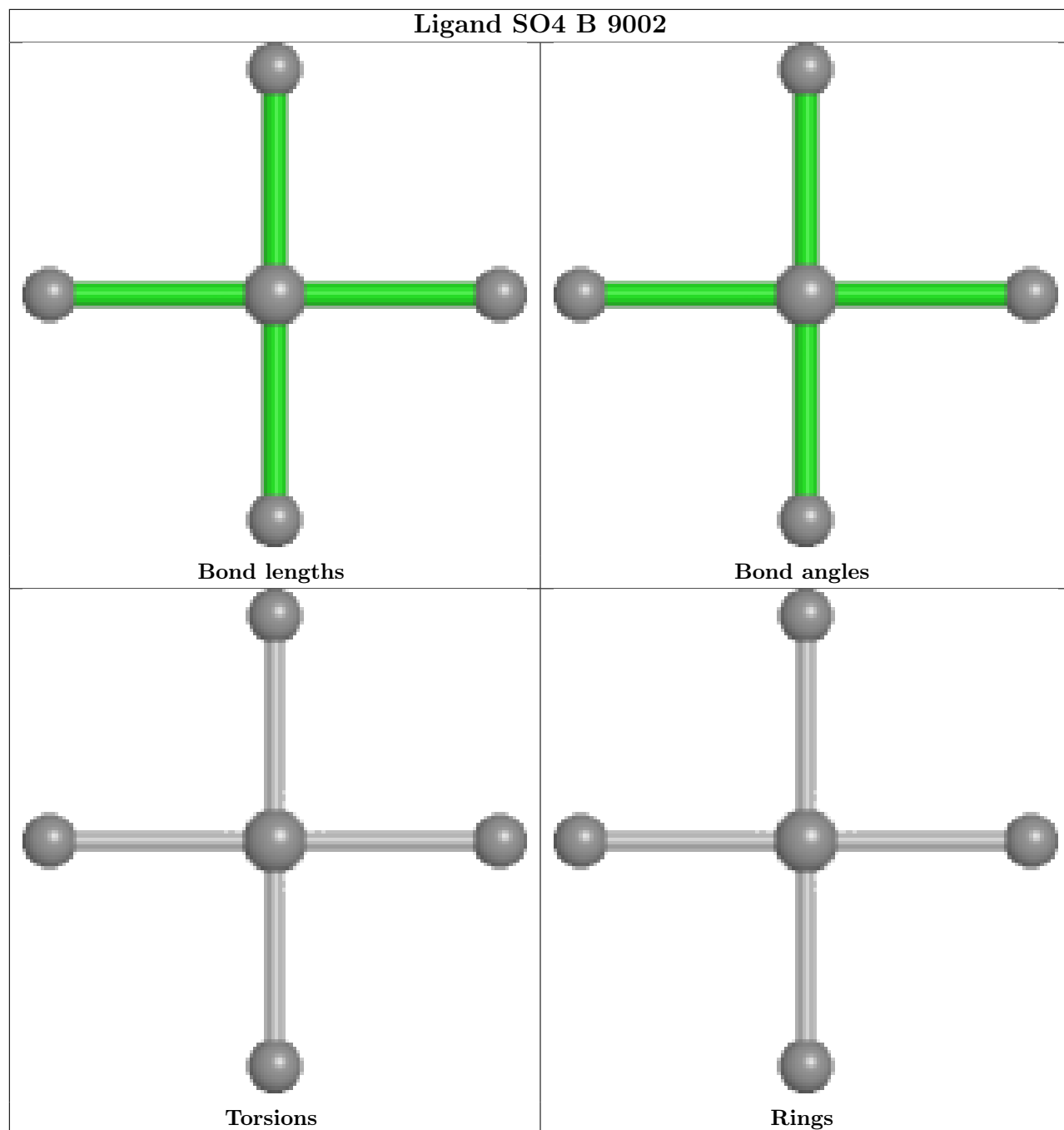
The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.
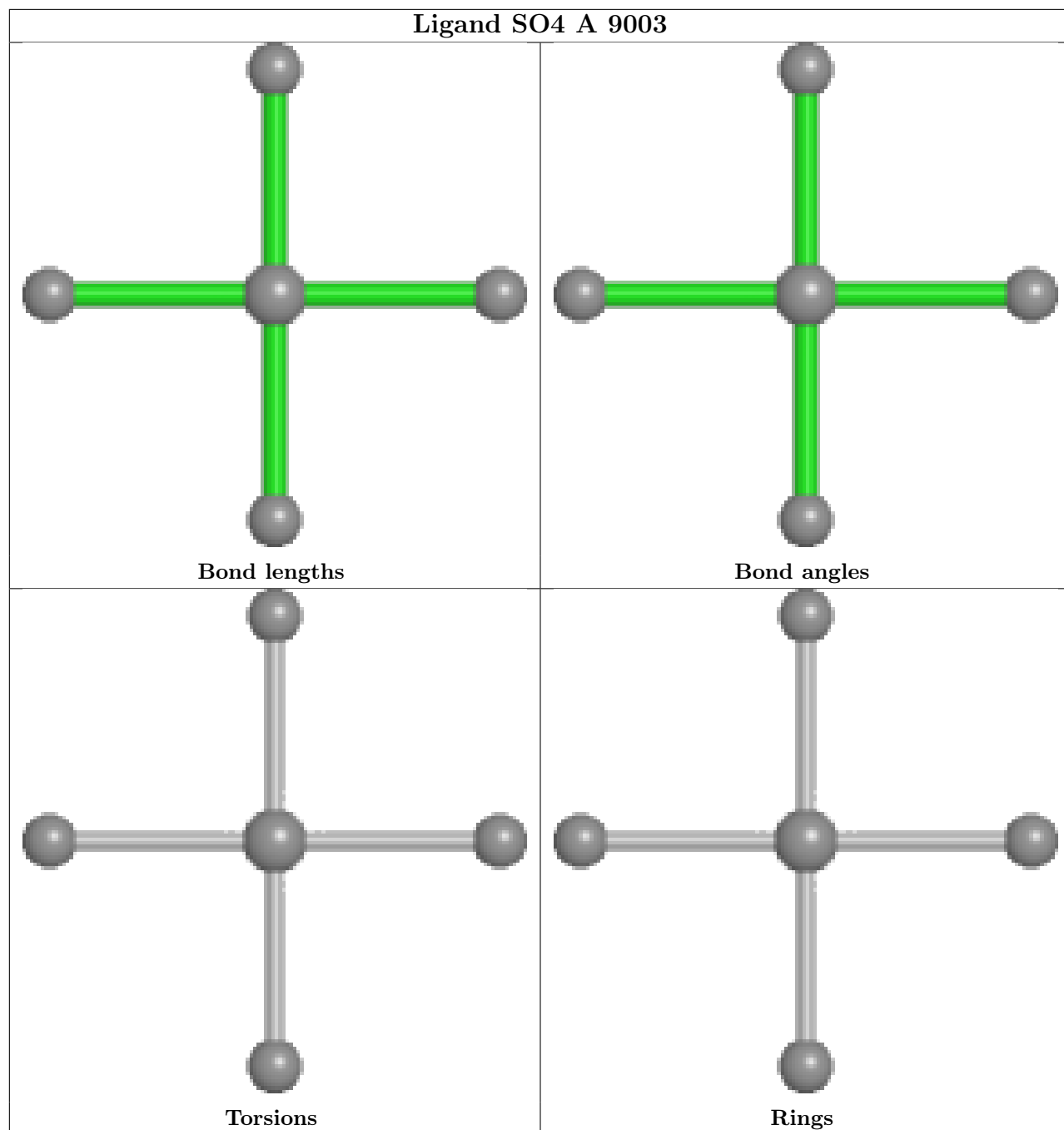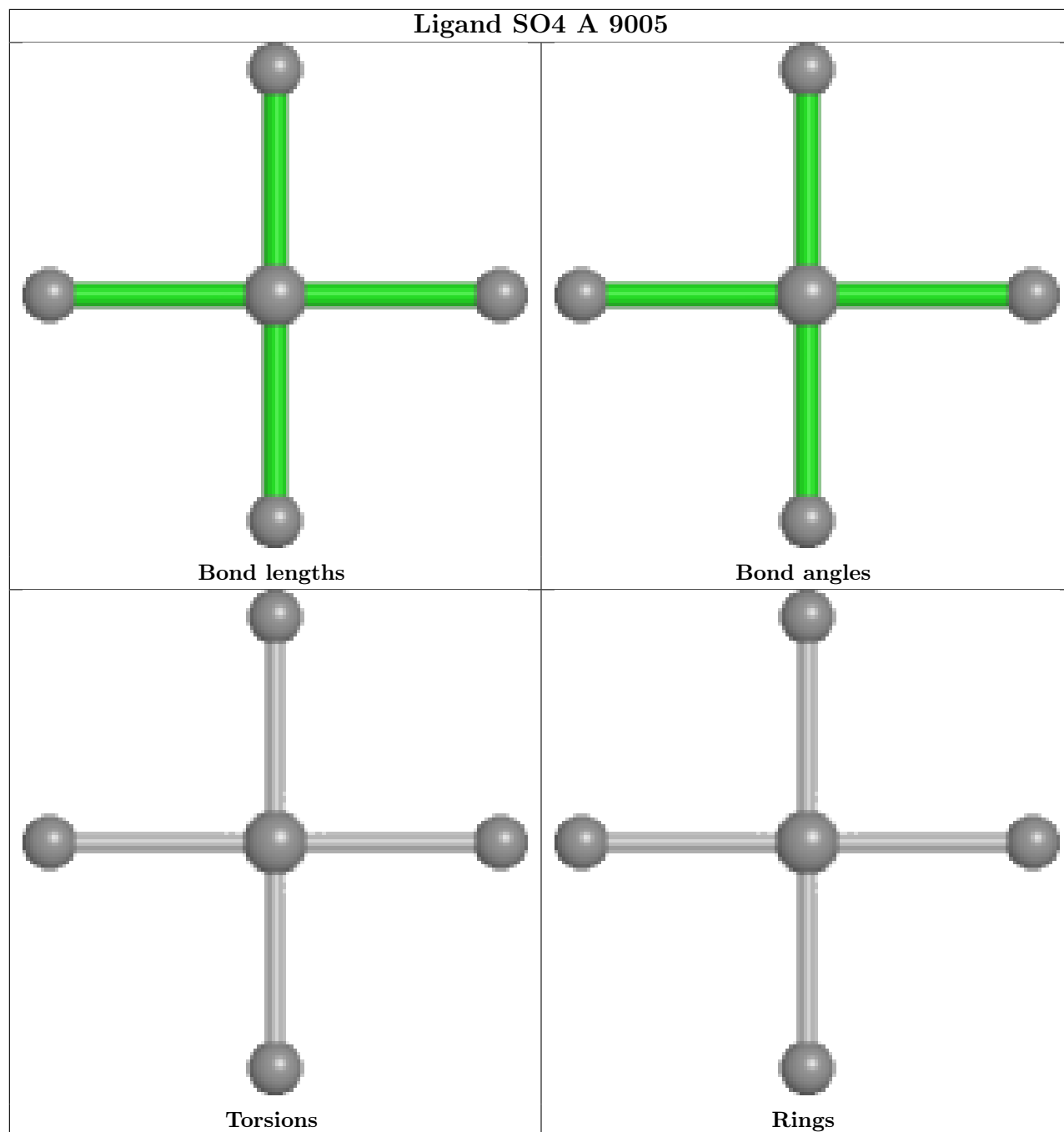

Ligand SO4 A 9002

Bond lengths | Bond angles

Torsions | Rings

Ligand SO4 A 9004

Bond lengths

Bond angles

Torsions

Rings

Ligand SO4 B 9001

Bond lengths

Bond angles

Torsions

Rings

Ligand SO4 A 9001

Bond lengths

Bond angles

Torsions

Rings

Ligand SO4 B 9002

Bond lengths

Bond angles

Torsions

Rings

Ligand SO4 A 9003

Bond lengths

Bond angles

Torsions

Rings

Ligand SO4 A 9005

Bond lengths

Bond angles

Torsions

Rings
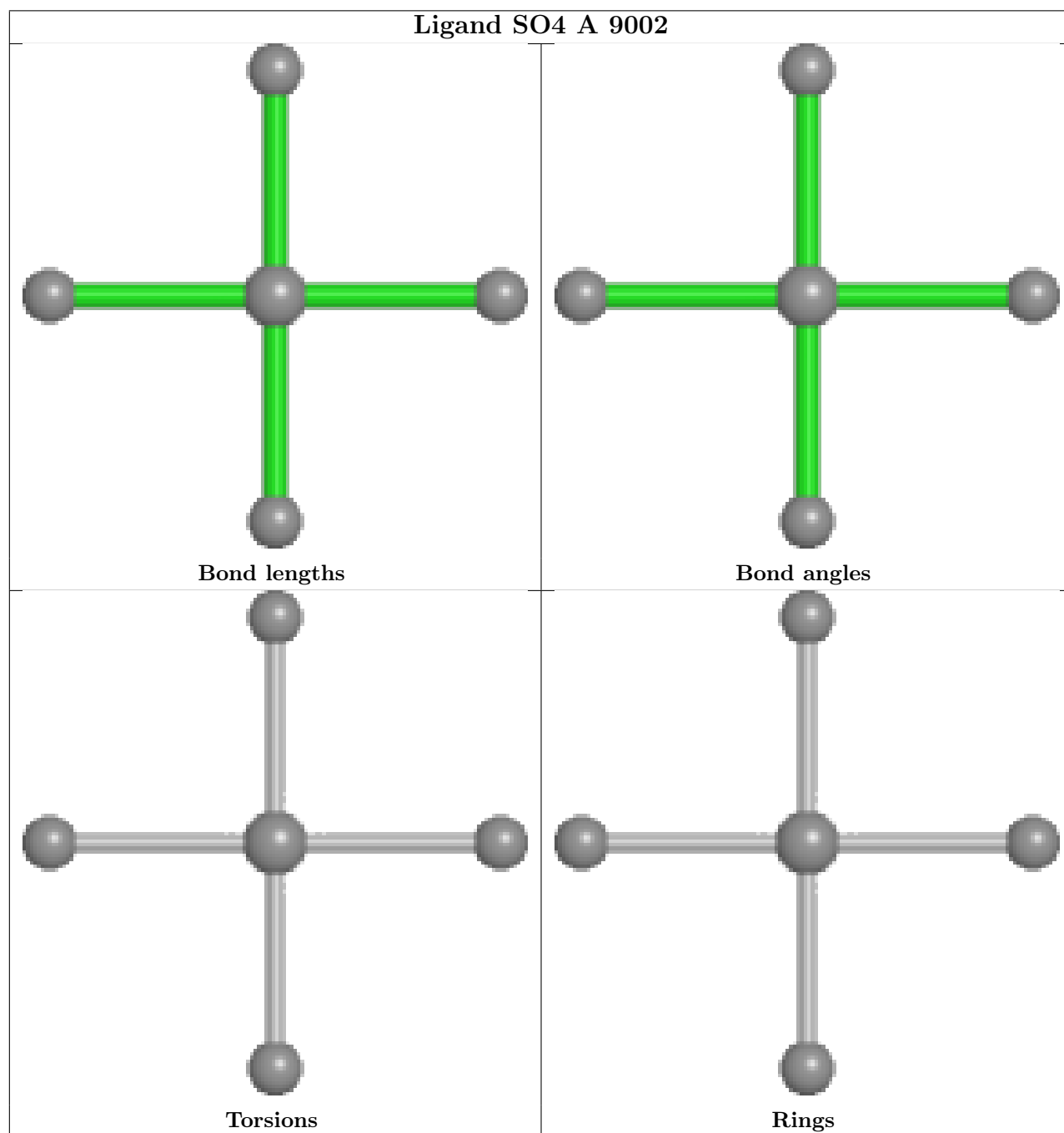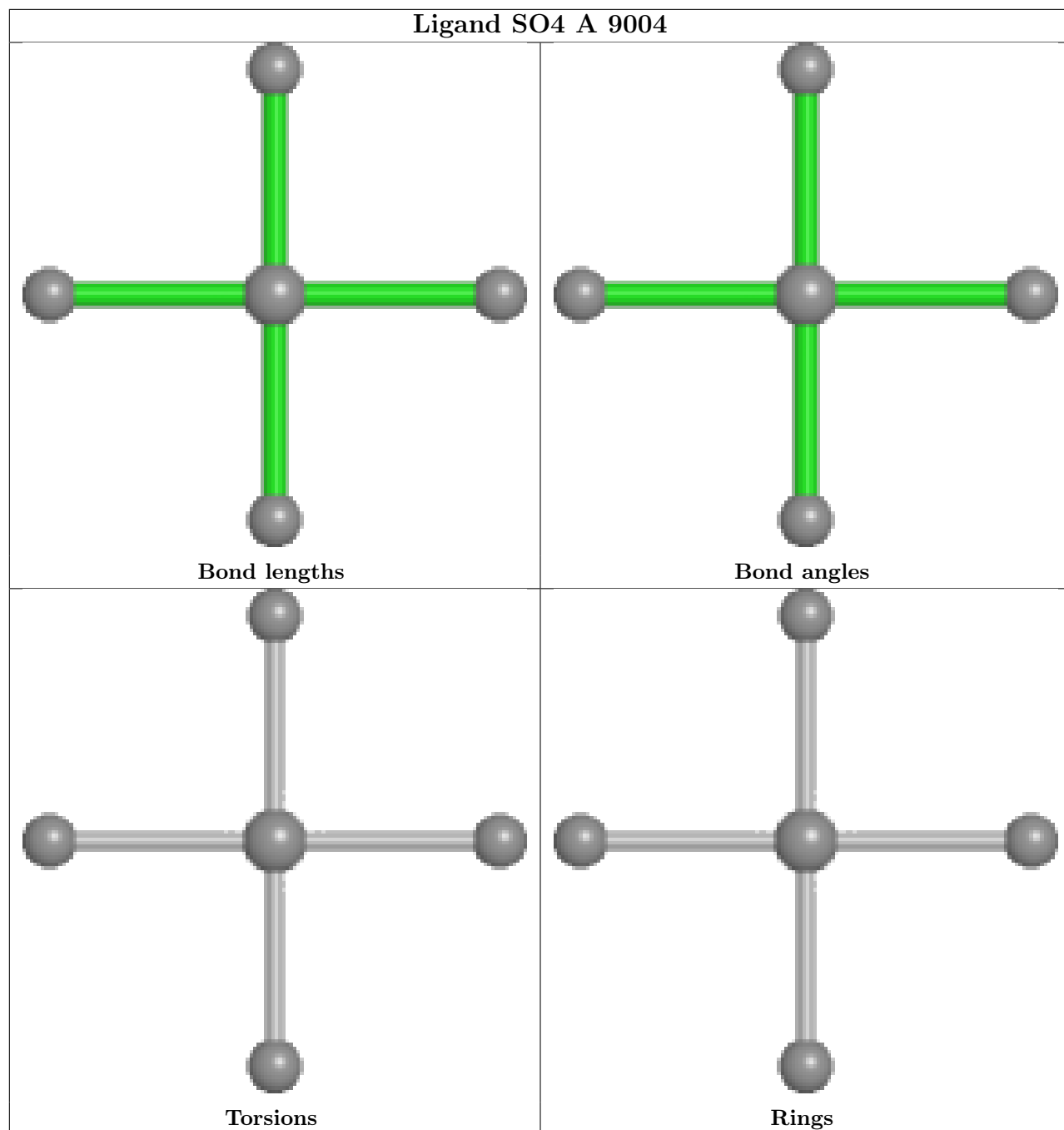
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th 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^2)	Q<0.9
1	A	351/352 (99%)	0.71	39 (11%)	5	6	15, 28, 50, 71	0
1	B	349/352 (99%)	0.79	48 (13%)	2	3	20, 31, 52, 64	0
All	All	700/704 (99%)	0.75	87 (12%)	4	4	15, 29, 52, 71	0

All (87) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	807	ASP	6.0
1	B	980	ASN	6.0
1	A	719	ALA	5.9
1	B	853	GLN	5.8
1	B	783	ASP	5.5
1	A	661	SER	5.4
1	A	723	GLY	5.2
1	B	785	SER	5.1
1	A	756	ASP	4.8
1	B	723	GLY	4.7
1	A	728	GLN	4.7
1	A	725	SER	4.7
1	A	726	ASP	4.6
1	A	724	SER	4.4
1	B	979	VAL	4.4
1	B	978	GLY	4.3
1	A	694	GLN	4.1
1	B	827	ASN	4.1
1	B	959	SER	4.1
1	B	982	THR	4.0
1	B	958	PRO	4.0
1	B	787	LYS	3.9
1	B	960	ALA	3.9
1	B	784	ASP	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	981	ASP	3.8
1	A	806	ARG	3.8
1	A	722	GLN	3.8
1	B	966	GLY	3.8
1	A	727	SER	3.7
1	A	758	VAL	3.6
1	A	920	LEU	3.5
1	B	694	GLN	3.4
1	B	920	LEU	3.3
1	B	846	GLN	3.2
1	B	911	SER	3.2
1	A	980	ASN	3.1
1	A	729	ILE	3.0
1	A	921	LEU	3.0
1	B	961	ASN	3.0
1	B	724	SER	3.0
1	B	756	ASP	2.9
1	A	754	ASN	2.9
1	B	786	SER	2.8
1	A	861	TRP	2.8
1	B	824	THR	2.8
1	A	721	SER	2.8
1	A	753	ASN	2.8
1	A	959	SER	2.8
1	A	981	ASP	2.8
1	A	739	LEU	2.7
1	A	743	ASP	2.7
1	B	914	ASP	2.7
1	B	855	HIS	2.6
1	B	825	THR	2.6
1	B	854	LEU	2.6
1	A	965	ASP	2.6
1	A	755	ALA	2.5
1	A	718	GLN	2.5
1	B	962	ILE	2.5
1	B	955	THR	2.5
1	A	757	SER	2.5
1	B	807	ASP	2.5
1	B	974	GLY	2.4
1	B	921	LEU	2.4
1	A	1003	LEU	2.4
1	A	780	GLY	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	976	SER	2.4
1	B	739	LEU	2.3
1	A	845	CYS	2.3
1	A	898	ALA	2.3
1	B	782	SER	2.3
1	A	962	ILE	2.2
1	B	956	PRO	2.2
1	B	965	ASP	2.2
1	B	967	VAL	2.2
1	B	977	SER	2.2
1	A	821	THR	2.2
1	B	898	ALA	2.2
1	B	725	SER	2.2
1	A	918	LEU	2.2
1	B	852	LYS	2.2
1	B	680	GLU	2.1
1	A	919	ILE	2.1
1	B	845	CYS	2.1
1	A	809	GLU	2.0
1	B	877	LEU	2.0
1	B	781	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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(Å^2)	Q<0.9
2	SO4	B	9002	5/5	0.91	0.18	79,79,79,80	0
2	SO4	A	9004	5/5	0.92	0.25	55,55,55,56	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	A	9005	5/5	0.94	0.20	58,59,59,59	0
2	SO4	A	9002	5/5	0.97	0.12	69,69,69,69	0
2	SO4	A	9003	5/5	0.98	0.17	65,65,65,65	0
2	SO4	A	9001	5/5	0.98	0.08	27,27,28,28	0
2	SO4	B	9001	5/5	0.99	0.06	37,37,37,37	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 SO4 B 9002:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)
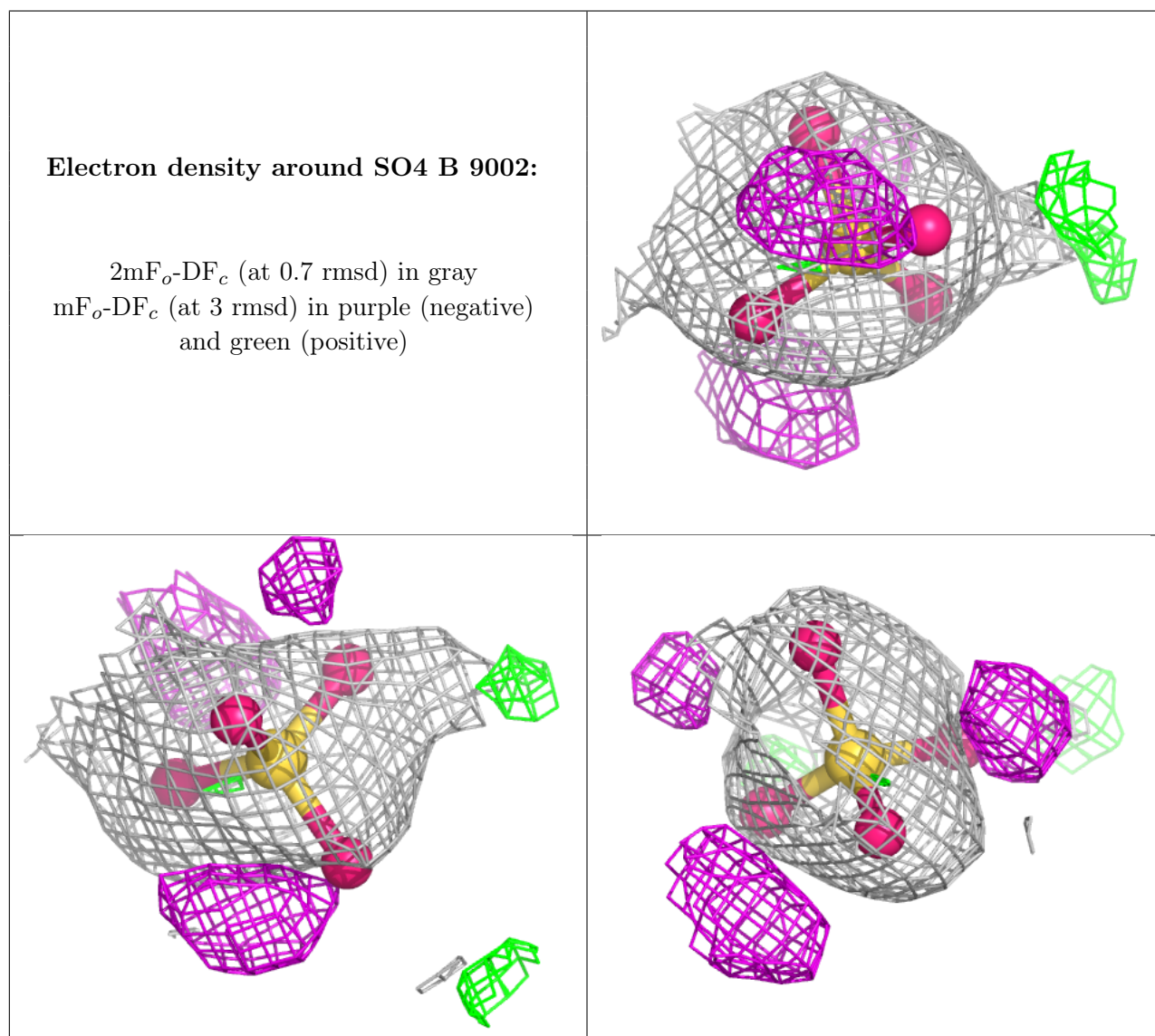
Electron density around SO4 A 9004:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
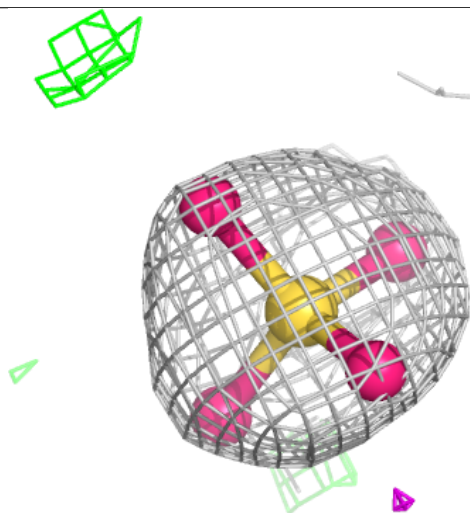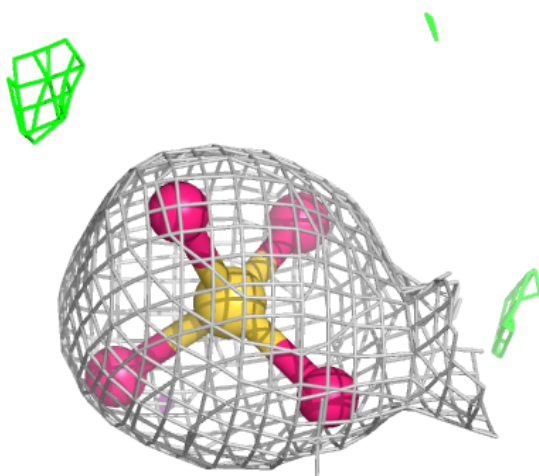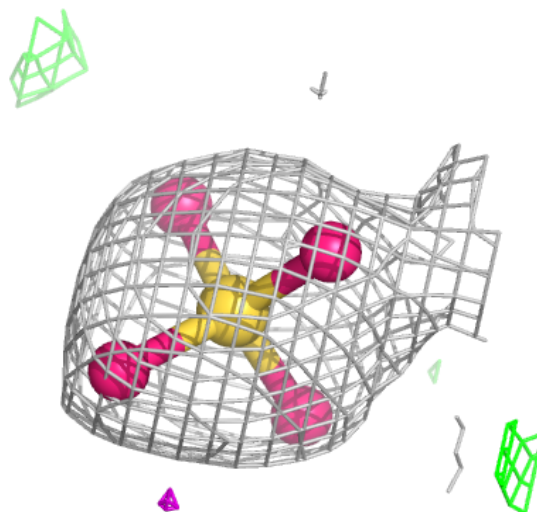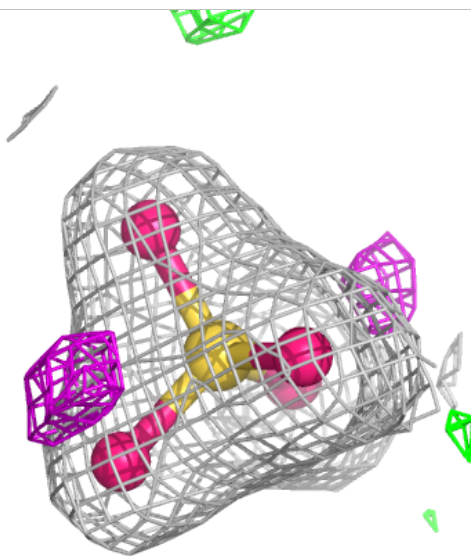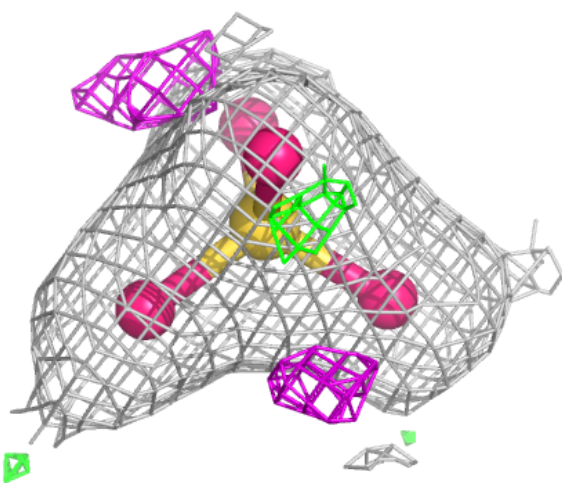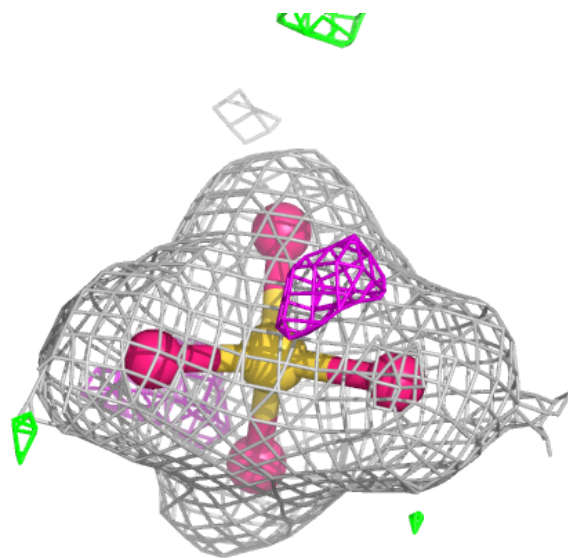and green (positive)
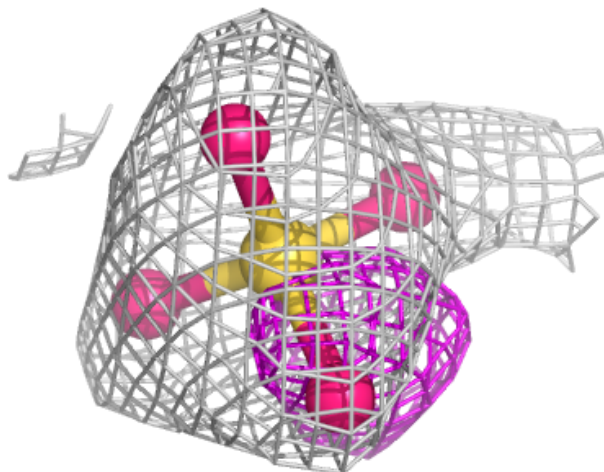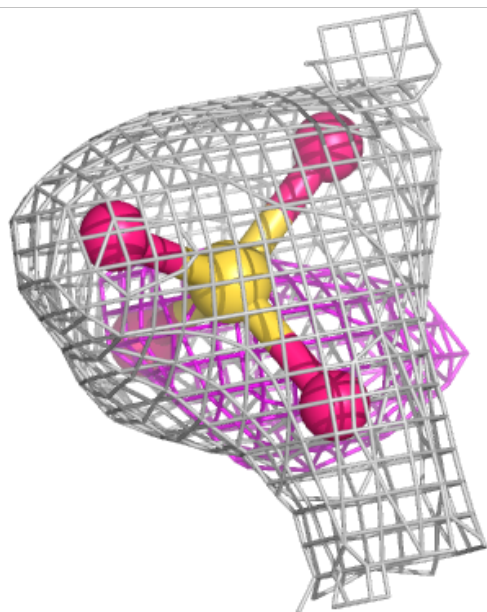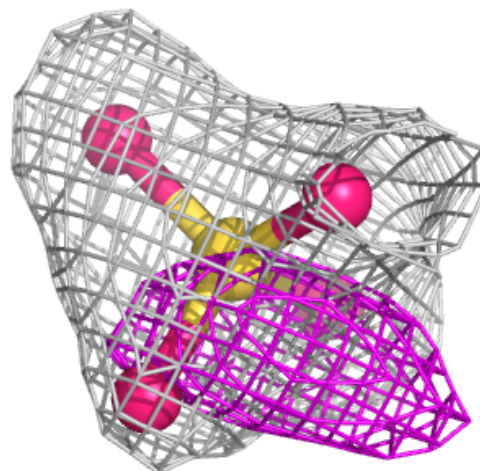
Electron density around SO4 A 9005:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
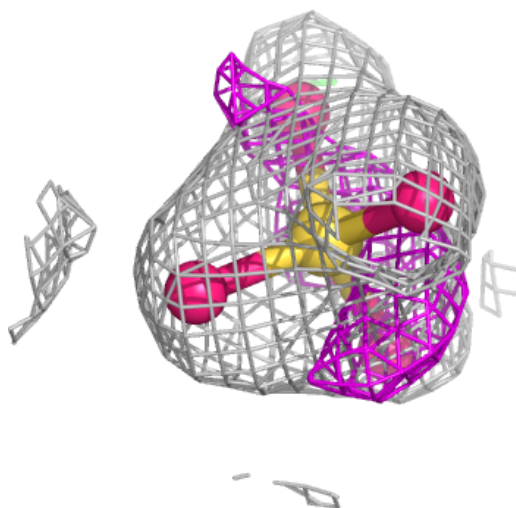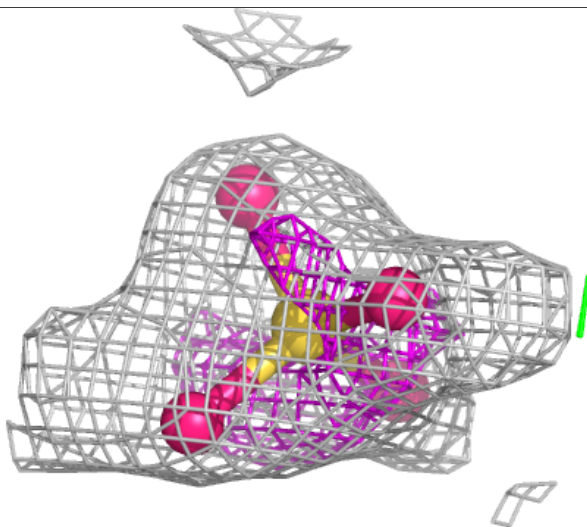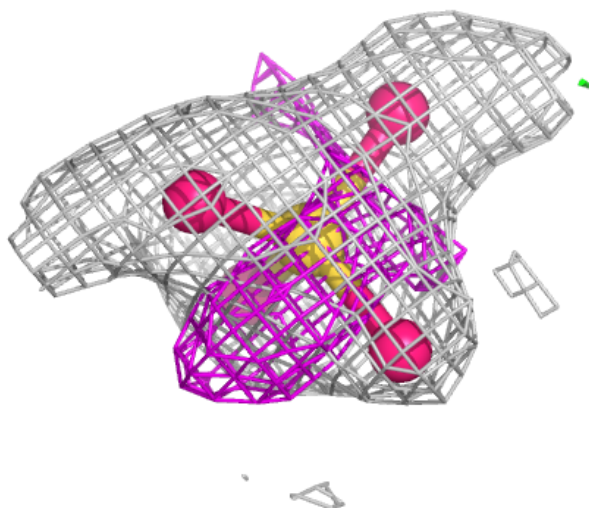
Electron density around SO4 A 9002:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
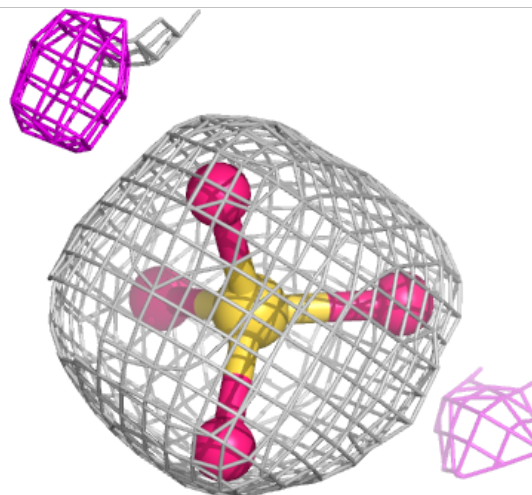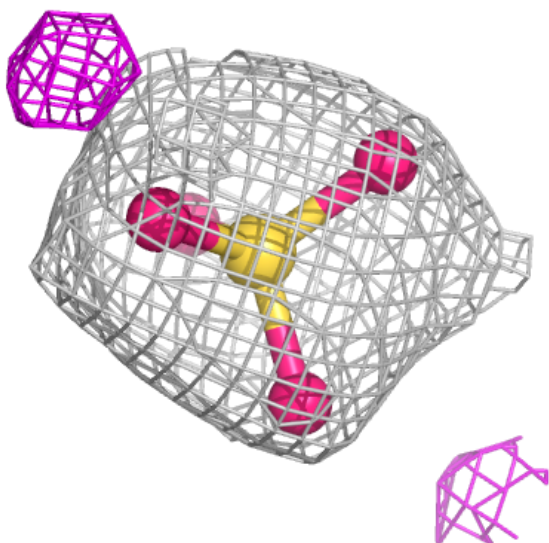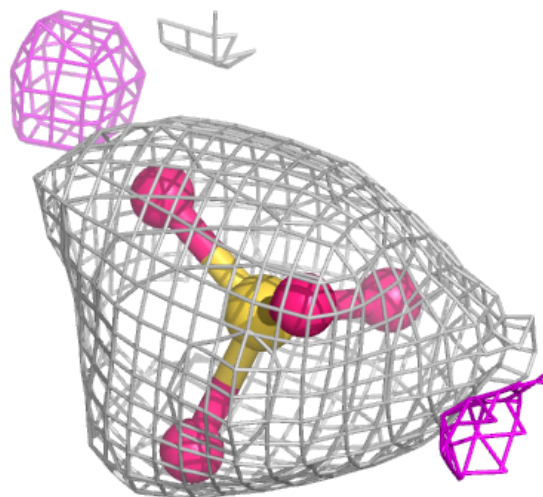
Electron density around SO4 A 9003:

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

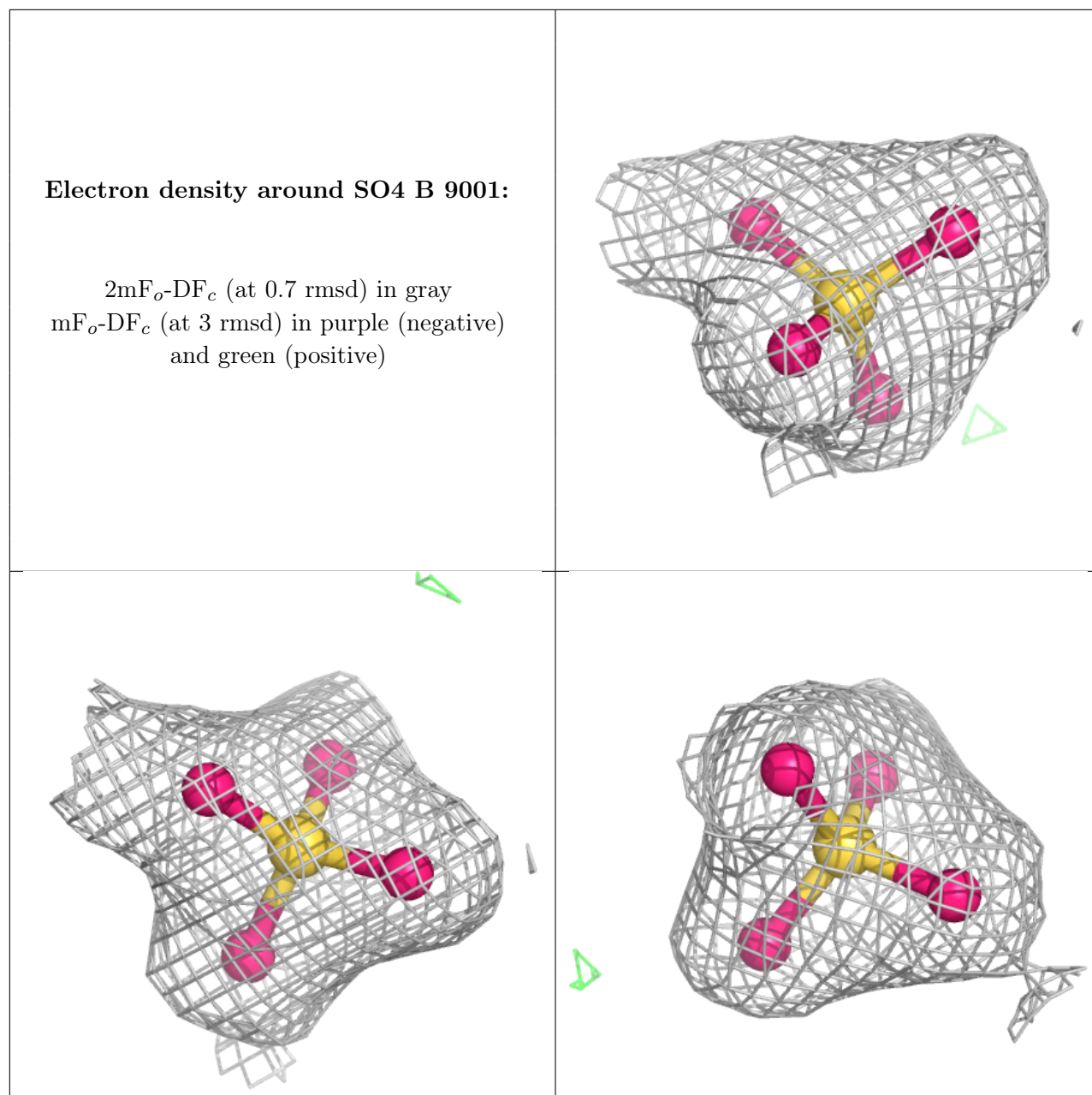**Electron density around SO4 A 9001:**

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

Electron density around SO4 B 9001:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.