wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2023 – 03:06 AM EDT

PDB ID	:	6X5M
Title	:	Crystal structure of a stabilized PAN ENE bimolecular triplex with a GC-clamped polyA tail, in complex with Fab-BL-3,6.
Authors	:	Swain, M.; Li, M.; Wlodawer, A.; Le Grice, S.F.J.
Deposited on	:	2020-05-26
Resolution	:	2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) 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 (i)) were used in the production of this report:

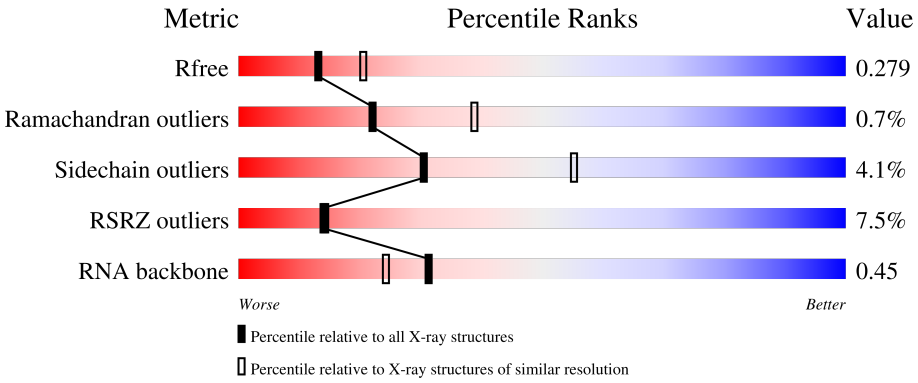
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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.



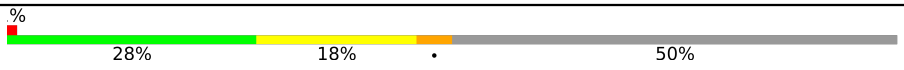
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)
RNA backbone	3102	1008 (2.84-2.16)

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	L	215	15% 94% 6%
1	l	215	7% 94% 6%
2	H	225	5% 95% .
2	h	225	5% 96% .
3	R	78	4% 29% 15% 5% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	r	78	
4	A	12	
4	a	12	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8881 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 Light chain Fab BL-3 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	l	215	Total	C	N	O	S	2	0	0
			1643	1025	275	337	6			
1	L	215	Total	C	N	O	S	2	0	0
			1643	1025	275	337	6			

- Molecule 2 is a protein called Heavy chain Fab Bl-3 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	h	225	Total	C	N	O	S	0	0	0
			1678	1055	287	330	6			
2	H	225	Total	C	N	O	S	0	0	0
			1678	1055	287	330	6			

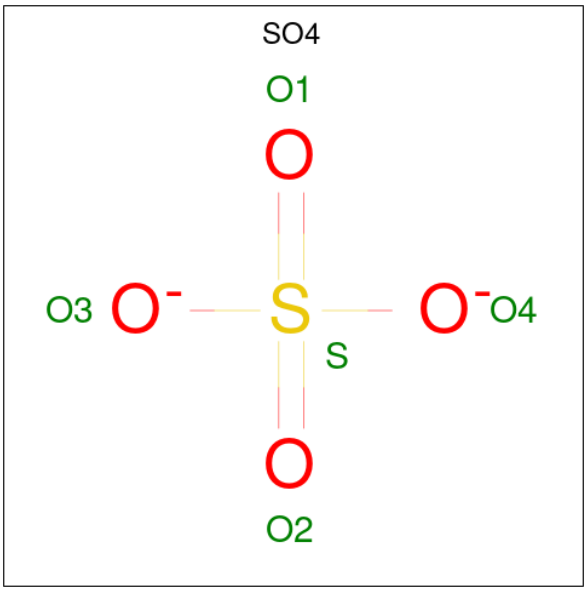
- Molecule 3 is a RNA chain called ggPAN RNA (39-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	39	Total	C	N	O	P	0	0	0
			818	366	133	280	39			
3	r	39	Total	C	N	O	P	0	0	0
			818	366	133	280	39			

- Molecule 4 is a RNA chain called ggcA9 RNA (12-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	12	Total	C	N	O	P	0	0	0
			264	119	58	75	12			
4	a	12	Total	C	N	O	P	0	0	0
			264	119	58	75	12			

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	l	1	Total O S 5 4 1	0	0
5	L	1	Total O S 5 4 1	0	0

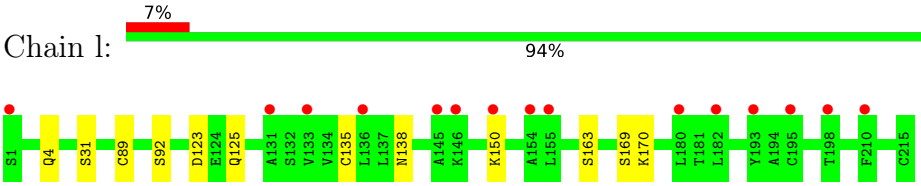
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	l	12	Total O 12 12	0	0
6	h	16	Total O 16 16	0	0
6	H	14	Total O 14 14	0	0
6	L	7	Total O 7 7	0	0
6	R	6	Total O 6 6	0	0
6	A	1	Total O 1 1	0	0
6	r	8	Total O 8 8	0	0
6	a	1	Total O 1 1	0	0

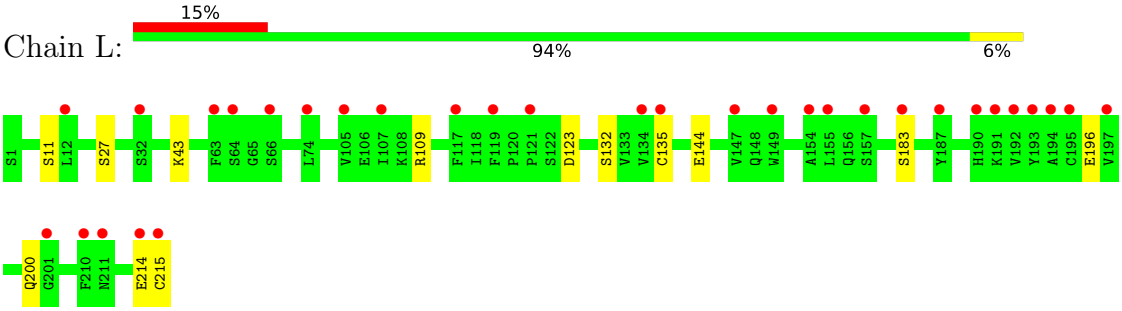
3 Residue-property plots (i)

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.

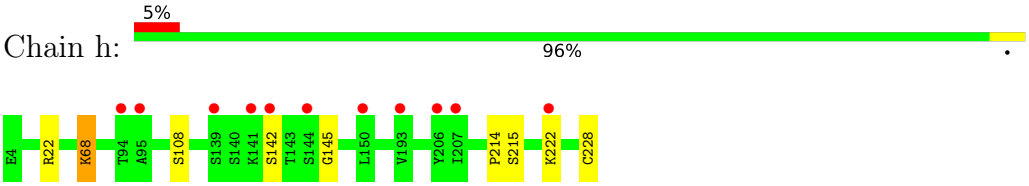
• Molecule 1: Light chain Fab BL-3 6



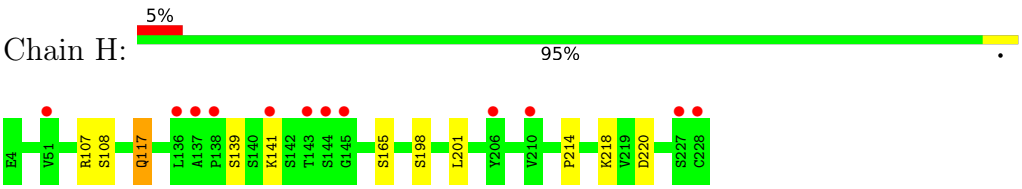
• Molecule 1: Light chain Fab BL-3 6



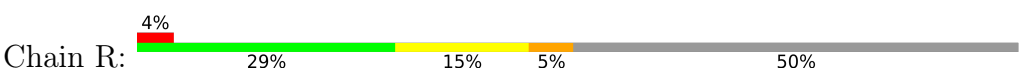
• Molecule 2: Heavy chain Fab Bl-3 6

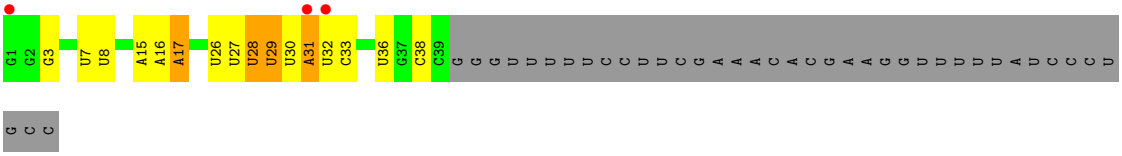


• Molecule 2: Heavy chain Fab Bl-3 6

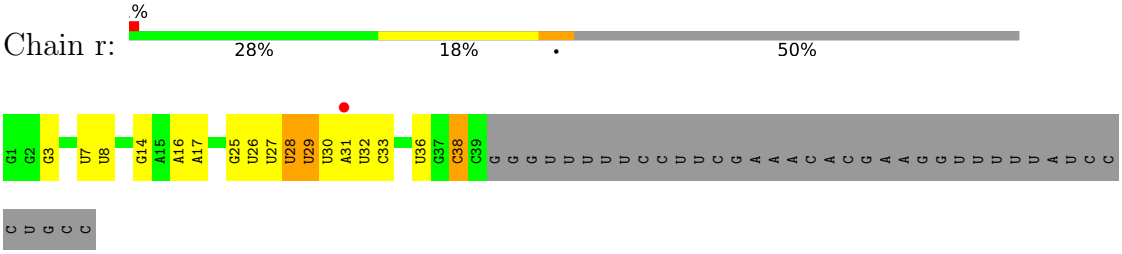


• Molecule 3: ggPAN RNA (39-MER)

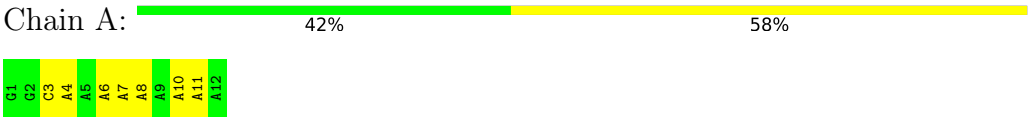




• Molecule 3: ggPAN RNA (39-MER)



• Molecule 4: ggcA9 RNA (12-MER)



• Molecule 4: ggcA9 RNA (12-MER)

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	43.13Å 76.26Å 110.66Å 71.90° 88.59° 86.58°	Depositor
Resolution (Å)	38.04 – 2.50 38.04 – 2.50	Depositor EDS
% Data completeness (in resolution range)	89.5 (38.04-2.50) 89.5 (38.04-2.50)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.02 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.16_3549, PHENIX 1.16_3549	Depositor
R, R_{free}	0.207 , 0.279 0.207 , 0.279	Depositor DCC
R_{free} test set	1299 reflections (3.14%)	wwPDB-VP
Wilson B-factor (Å^2)	77.1	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 76.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,-k+l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8881	wwPDB-VP
Average B, all atoms (Å^2)	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.01% 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	L	0.52	0/1678	0.84	2/2277 (0.1%)
1	l	0.51	0/1678	0.74	1/2277 (0.0%)
2	H	0.52	0/1719	0.79	3/2339 (0.1%)
2	h	0.57	0/1719	0.73	0/2339
3	R	0.58	0/910	1.21	6/1413 (0.4%)
3	r	0.64	0/910	1.38	10/1413 (0.7%)
4	A	0.53	0/298	1.11	0/463
4	a	0.47	0/298	1.08	0/463
All	All	0.55	0/9210	0.94	22/12984 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	L	0	1
1	l	0	1
2	H	0	1
2	h	0	2
All	All	0	5

There are no bond length outliers.

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	117	GLN	CA-CB-CG	8.82	132.81	113.40
3	r	28	U	O4'-C1'-N1	8.15	114.72	108.20
3	r	29	U	O5'-P-OP1	-8.10	98.41	105.70
2	H	201	LEU	CA-CB-CG	7.65	132.90	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	r	16	A	O4'-C1'-N9	-7.10	102.52	108.20

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	117	GLN	Sidechain
1	L	123	ASP	Peptide
2	h	142	SER	Peptide
2	h	68	LYS	Peptide
1	l	123	ASP	Peptide

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	L	213/215 (99%)	200 (94%)	13 (6%)	0	100	100
1	l	213/215 (99%)	202 (95%)	11 (5%)	0	100	100
2	H	223/225 (99%)	207 (93%)	13 (6%)	3 (1%)	12	21
2	h	223/225 (99%)	207 (93%)	13 (6%)	3 (1%)	12	21
All	All	872/880 (99%)	816 (94%)	50 (6%)	6 (1%)	22	39

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	h	145	GLY
2	H	141	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	h	108	SER
2	H	108	SER
2	H	214	PRO

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	L	190/190 (100%)	180 (95%)	10 (5%)	22	43
1	l	190/190 (100%)	180 (95%)	10 (5%)	22	43
2	H	186/186 (100%)	180 (97%)	6 (3%)	39	65
2	h	186/186 (100%)	181 (97%)	5 (3%)	44	71
All	All	752/752 (100%)	721 (96%)	31 (4%)	30	55

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	h	228	CYS
1	L	183	SER
2	H	165	SER
1	L	214	GLU
1	L	109	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	167	GLN
1	L	39	GLN
2	H	183	GLN
2	H	176	HIS
2	H	211	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	38/78 (48%)	15 (39%)	1 (2%)
3	r	38/78 (48%)	14 (36%)	0
4	A	11/12 (91%)	7 (63%)	0
4	a	11/12 (91%)	6 (54%)	0
All	All	98/180 (54%)	42 (42%)	1 (1%)

5 of 42 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	3	G
3	R	7	U
3	R	8	U
3	R	15	A
3	R	17	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	R	28	U

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)

2 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	$\#\|Z\| > 2$
5	SO4	L	301	-	4,4,4	0.17	0	6,6,6	0.14	0
5	SO4	l	301	-	4,4,4	0.16	0	6,6,6	0.43	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

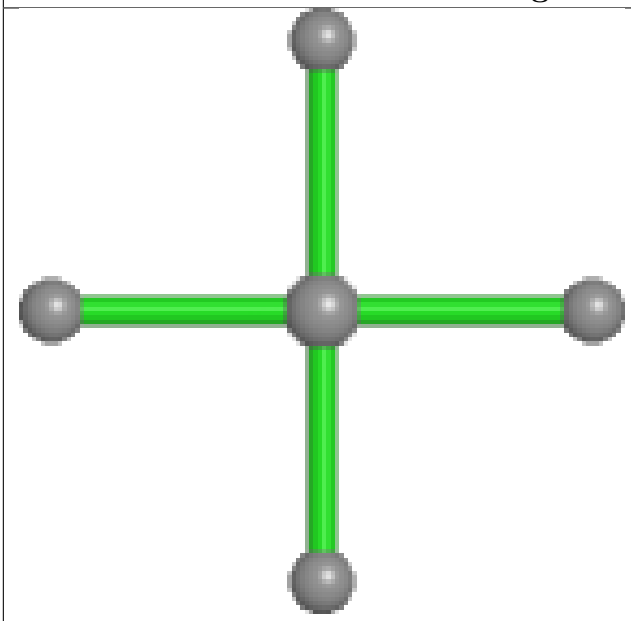
There are no torsion outliers.

There are no ring outliers.

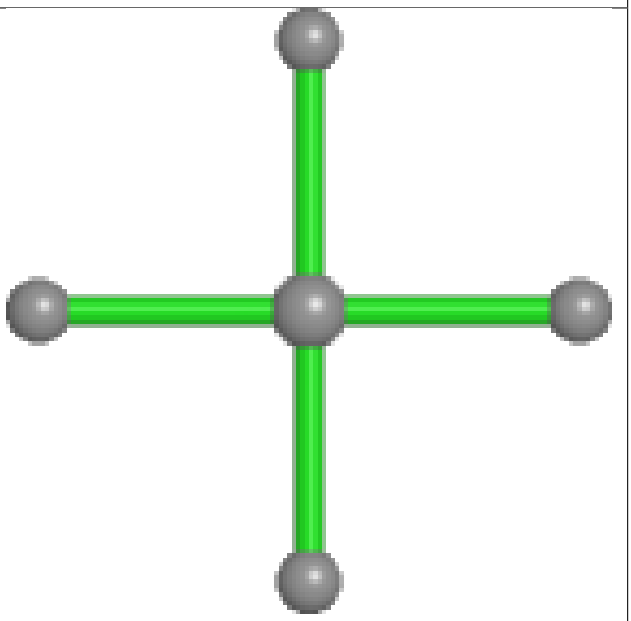
No monomer is involved in short contacts.

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.

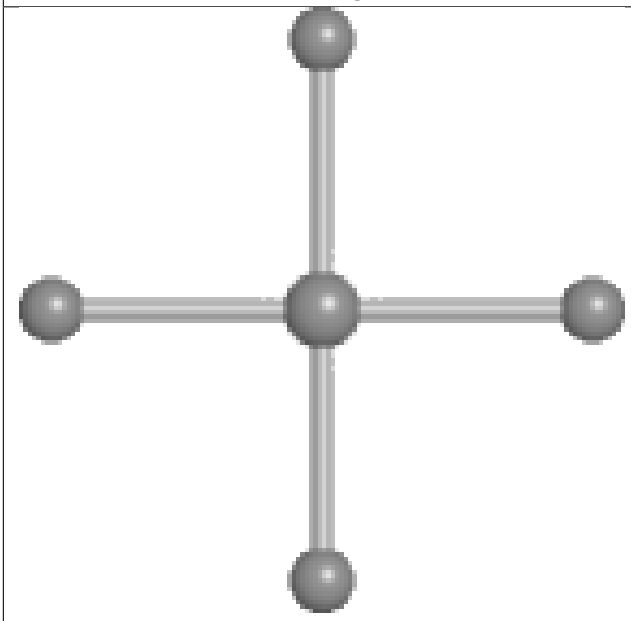
Ligand SO4 L 301



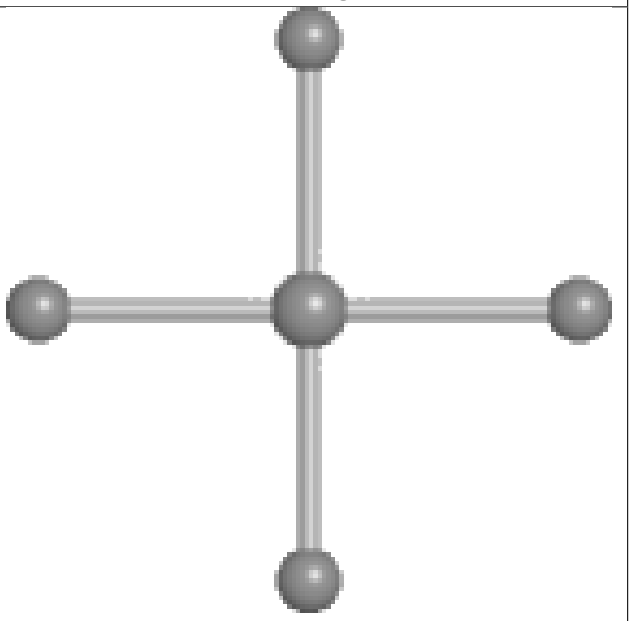
Bond lengths



Bond angles

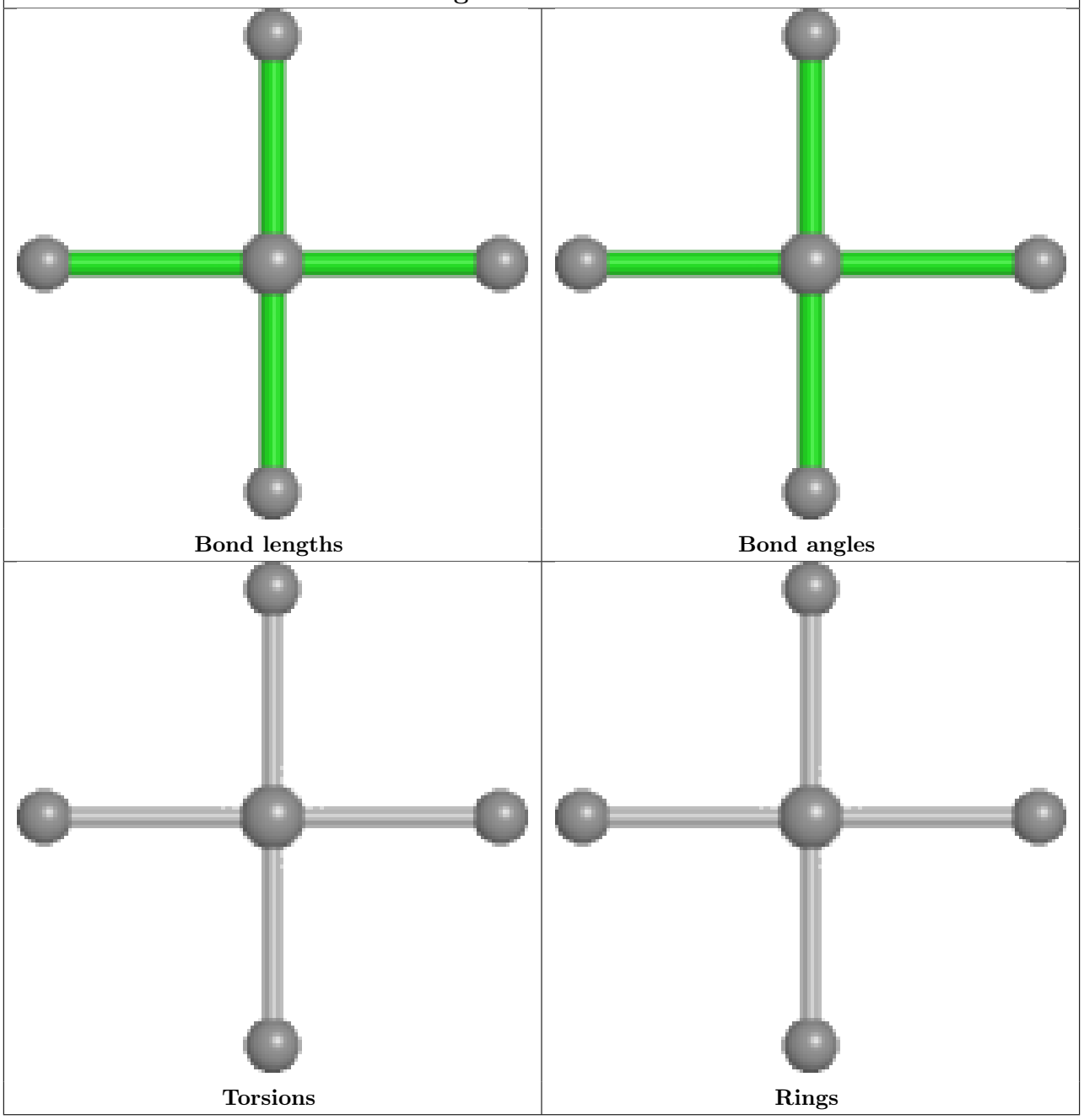


Torsions



Rings

Ligand SO4 l 301



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^{th} 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	L	215/215 (100%)	1.00	32 (14%) 2 2	84, 131, 202, 220	2 (0%)
1	l	215/215 (100%)	0.40	15 (6%) 16 16	54, 96, 171, 207	2 (0%)
2	H	225/225 (100%)	0.58	12 (5%) 26 28	57, 96, 186, 211	0
2	h	225/225 (100%)	0.51	11 (4%) 29 31	53, 81, 164, 218	0
3	R	39/78 (50%)	0.11	3 (7%) 13 13	91, 128, 170, 236	0
3	r	39/78 (50%)	-0.05	1 (2%) 56 59	69, 130, 175, 210	0
4	A	12/12 (100%)	-0.08	0 100 100	140, 149, 162, 162	0
4	a	12/12 (100%)	-0.17	0 100 100	131, 139, 165, 168	0
All	All	982/1060 (92%)	0.55	74 (7%) 14 14	53, 109, 186, 236	4 (0%)

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	145	GLY	14.9
3	R	31	A	8.9
2	h	139	SER	8.3
2	h	142	SER	7.7
1	L	201	GLY	7.3

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 (i)

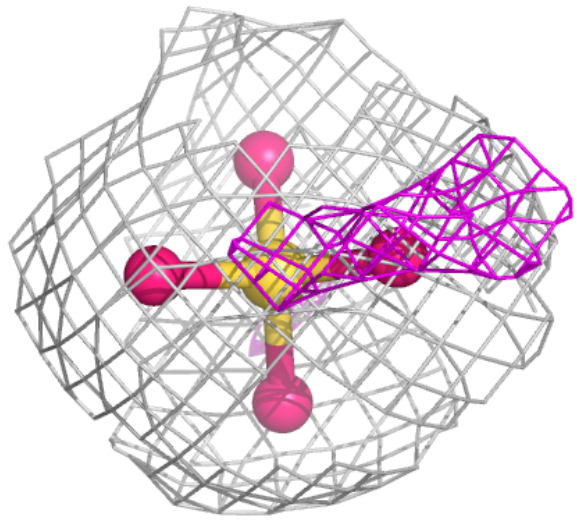
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

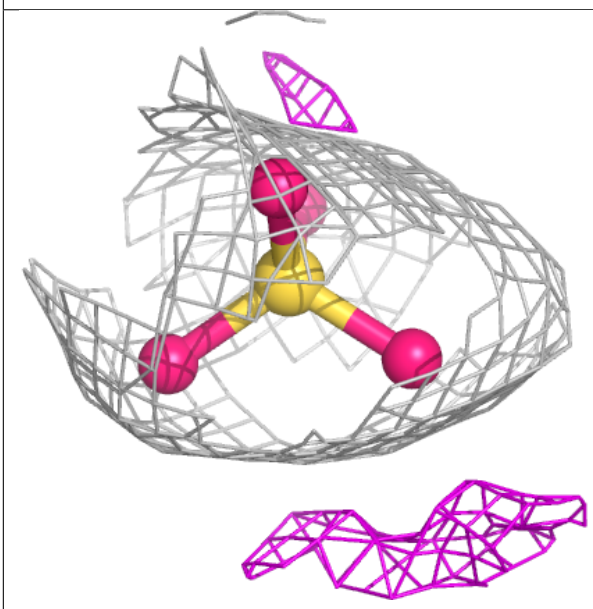
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	SO4	l	301	5/5	0.88	0.13	109,115,118,121	0
5	SO4	L	301	5/5	0.92	0.13	130,133,136,137	0

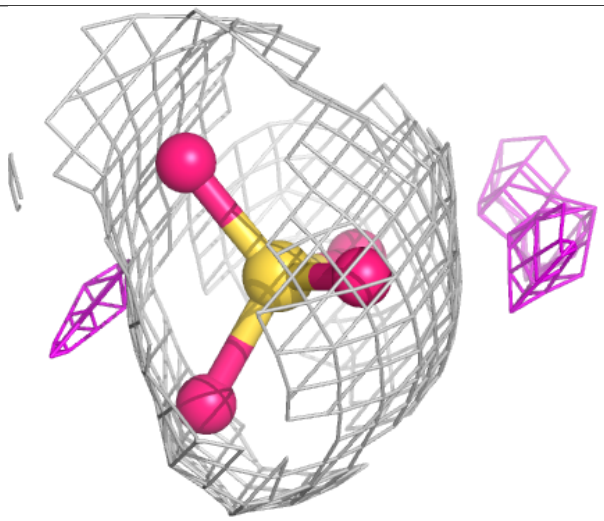
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around SO4 l 301:

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

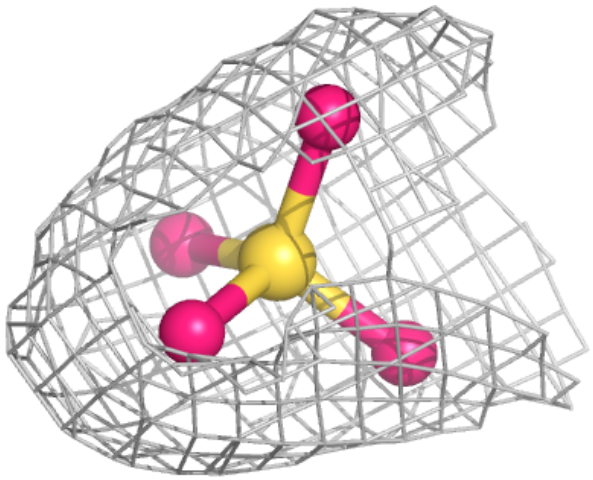


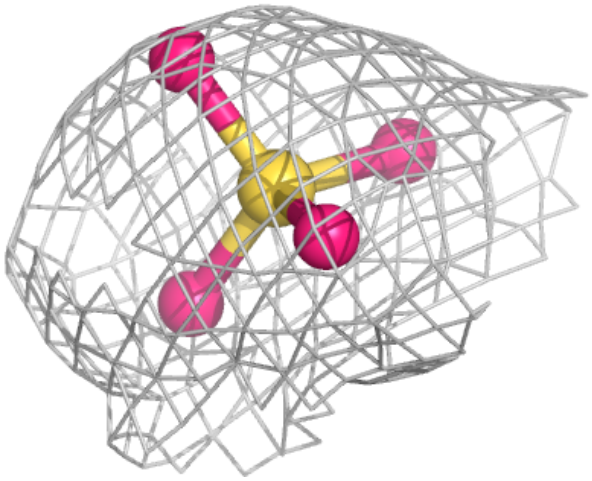


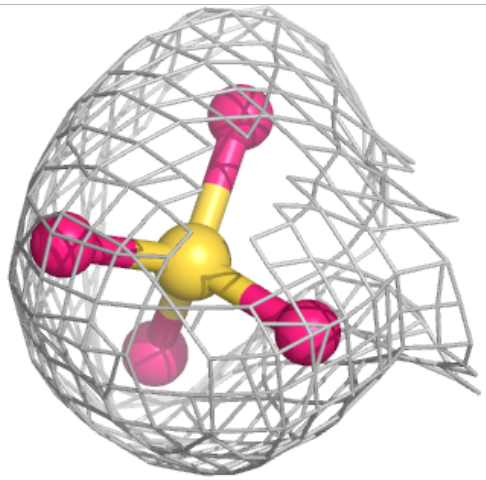


Electron density around SO4 L 301:

$2mF_o$-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.