

# wwPDB X-ray Structure Validation Summary Report (i)

Sep 6, 2023 – 10:44 am BST

PDB ID : 8OK6

Title : Variant Surface Glycoprotein VSG11 two monomers Authors : Gkeka, A.; Vlachou, E.P.; Zeelen, J.P.; Stebbins, C.E.

Deposited on : 2023-03-27

Resolution : 1.75 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 (1)) were used in the production of this report:

MolProbity: 4.02b-467

Mogul : 1.8.4, CSD as541be (2020)

Xtriage (Phenix) : 1.13

EDS : 2.35

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)

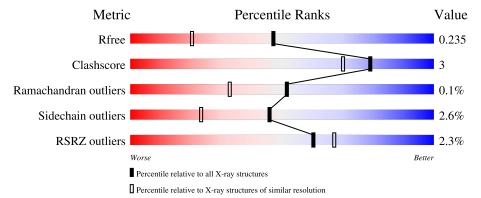
Validation Pipeline (wwPDB-VP) : 2.35

## 1 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 1.75 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\text{Entries}) \end{array}$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries,\ resolution\ range(\AA)}) \end{array}$
$R_{free}$	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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	415	79%	9% • 11%
1	С	415	81%	7% 11%
2	В	6	67%	33%
3	D	5	100%	



## 2 Entry composition (i)

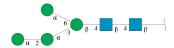
There are 6 unique types of molecules in this entry. The entry contains 6278 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Variant surface glycoprotein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	A	368	Total 2781	C 1716	N 496	O 556	S 13	0	2	0
1	С	368	Total 2811	C 1735	N 504	O 559	S 13	0	5	0

• Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	В	6	Total 72	C 40	N 2	O 30	0	0	0

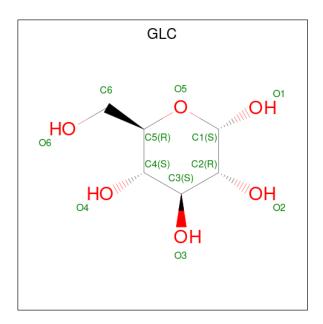
• Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	5	Total 61	C 34	N 2	O 25	0	0	0

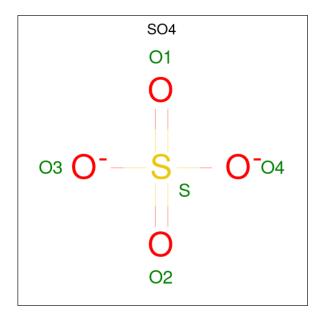
• Molecule 4 is alpha-D-glucopyranose (three-letter code: GLC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>) (labeled as "Ligand of Interest" by depositor).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 11 6 5	0	0
4	С	1	Total C O 11 6 5	0	0

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



Mo	ıl	Chain	Residues	Atom	S	ZeroOcc	AltConf
5		A	1	Total O 5 4	S 1	0	0

Continued on next page...



 $Continued\ from\ previous\ page...$ 

Mol	Chain	$oxed{ \mathbf{Residues} \ }$	Ato	oms		ZeroOcc	AltConf	
F	Λ	1	Total	О	S	0	0	
5	A	1	5	4	1	0	0	
5	A	1	Total	О	S	0	0	
)	A	1	5	4	1	U	U	
5	A	1	Total	О	S	0	0	
	Λ	1	5	4	1	U	U	
5	A	1	Total	О	S	0	0	
	Λ	1	5	4	1	U	U	
5	A	1	Total	О	S	0	0	
	11	1	5	4	1	O	0	
5	A	1	Total	О	S	0	0	
	71	1	5	4	1	0	0	
5	A	1	Total	Ο	S	0	0	
	11	1	5	4	1	Ŭ	Ů	
5	С	1	Total	О	S	0	0	
			5	4	1	Ü		
5	С	1	Total	O	S	0	0	
		_	5	4	1	0		
5	С	1	Total	O	S	0	0	
		_	5	4	1			
5	С	1	Total	O	S	0	0	
			5	4	1			
5	С	1	Total	0	S	0	0	
			5	4	1			
5	С	1	Total	0	S	0	0	
			5	4	1			
5	С	1	Total	O 4	S	0	0	
			5	4	1			
5	С	1	Total	0	S	0	0	
			5	4	1			
5	С	1	Total	O	S	0	0	
			5	4	1			

### • Molecule 6 is water.

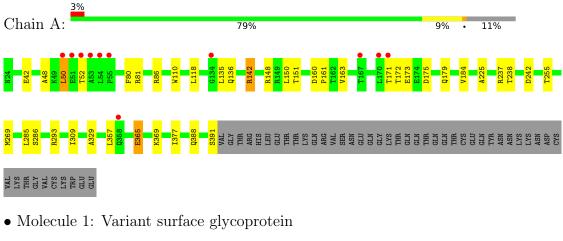
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	178	Total O 178 178	0	0
6	C	268	Total O 268 268	0	0

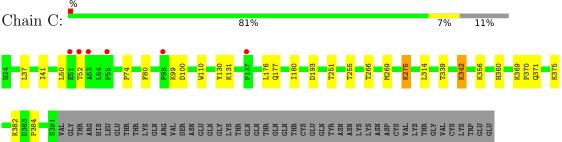


## 3 Residue-property plots (i)

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: Variant surface glycoprotein





 $\bullet$  Molecule 2: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



 $\bullet \ \, Molecule \ 3: \ alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)] beta-D-mannopyranose-(1-6)] beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-$ 

Chain D: 100%







## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants	76.24Å 76.24Å 212.26Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $120.00^{\circ}$	Depositor
Resolution (Å)	48.27 - 1.75	Depositor
rtesolution (A)	48.27 - 1.75	EDS
% Data completeness	99.8 (48.27-1.75)	Depositor
(in resolution range)	99.8 (48.27-1.75)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.06 (at 1.75Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
D D.	0.201 , 0.235	Depositor
$R, R_{free}$	0.201 , $0.235$	DCC
$R_{free}$ test set	3646  reflections  (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.7	Xtriage
Anisotropy	0.103	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.37 , 52.7	EDS
L-test for twinning <sup>2</sup>	$< L >=0.42, < L^2>=0.25$	Xtriage
Estimated twinning fraction	0.055 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6278	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The analyses of the Patterson function reveals a significant off-origin peak that is 32.40 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.4596e-04.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: GLC, NAG, BMA, MAN, 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		
IVIOI	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	0.45	0/2825	0.62	0/3836	
1	С	0.45	0/2855	0.63	0/3873	
All	All	0.45	0/5680	0.62	0/7709	

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 maintenain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	С	0	1
All	All	0	5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	148	ARG	Sidechain
1	A	237	ARG	Mainchain
1	A	238[A]	THR	Mainchain
1	С	275[A]	LYS	Mainchain



### 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	2781	0	2745	18	2
1	С	2811	0	2786	16	1
2	В	72	0	61	1	0
3	D	61	0	52	0	0
4	A	11	0	10	0	0
4	С	11	0	10	0	0
5	A	40	0	0	0	0
5	С	45	0	0	0	1
6	A	178	0	0	0	0
6	C	268	0	0	2	0
All	All	6278	0	5664	32	3

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.

The worst 5 of 32 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{aligned}  ext{Interatomic} \  ext{distance} \ ( ext{Å}) \end{aligned}$	$egin{aligned} \operatorname{Clash} \ \operatorname{overlap}\ (\begin{subarray}{c} \begin{subarray}{c} \begi$
1:A:172:THR:HG23	1:A:175:ASP:H	1.58	0.67
1:A:50:LEU:H	1:A:173:GLU:HG3	1.63	0.64
1:A:285:LEU:HD13	1:A:357:LEU:HD11	1.85	0.58
1:C:371:GLN:HG2	1:C:375:LYS:HE2	1.86	0.58
1:A:135:LEU:O	1:C:74:PRO:HG2	2.04	0.58

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	$egin{aligned}  ext{Interatomic} \  ext{distance} & ( ext{Å}) \end{aligned}$	Clash overlap (Å)
1:A:365:GLU:OE2	1:A:365:GLU:OE2[2_545]	1.43	0.77
1:A:365:GLU:CD	1:A:365:GLU:OE2[2_545]	2.07	0.13
1:C:251:THR:OG1	5:C:503:SO4:O1[2_655]	2.11	0.09



### 5.3 Torsion angles (i)

#### 5.3.1 Protein backbone (i)

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	Percenti	iles
1	A	368/415 (89%)	349 (95%)	18 (5%)	1 (0%)	41 2	2
1	C	371/415~(89%)	359 (97%)	12 (3%)	0	100 1	00
All	All	739/830 (89%)	708 (96%)	30 (4%)	1 (0%)	51 3	3

#### All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	THR

#### 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	n Analysed Rotameric Outliers		Percentiles	
1	A	293/338 (87%)	283 (97%)	10 (3%)	37 14
1	С	297/338~(88%)	291 (98%)	6 (2%)	55 34
All	All	590/676 (87%)	574 (97%)	16 (3%)	46 22

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	С	275[B]	LYS
1	С	275[A]	LYS
1	A	369	LYS
1	С	193	ASP
1	A	365	GLU



Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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)

11 monosaccharides 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	Вс	ond leng	ths	В	ond ang	les
MIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	В	1	1,2	14,14,15	0.42	0	17,19,21	0.60	0
2	NAG	В	2	2	14,14,15	0.41	0	17,19,21	0.57	0
2	BMA	В	3	2	11,11,12	0.39	0	15,15,17	0.88	1 (6%)
2	MAN	В	4	2	11,11,12	0.40	0	15,15,17	0.58	0
2	MAN	В	5	2	11,11,12	0.25	0	15,15,17	0.53	0
2	MAN	В	6	2	11,11,12	0.31	0	15,15,17	0.56	0
3	NAG	D	1	1,3	14,14,15	0.37	0	17,19,21	0.57	0
3	NAG	D	2	3	14,14,15	0.42	0	17,19,21	0.64	0
3	BMA	D	3	3	11,11,12	0.35	0	15,15,17	0.65	0
3	MAN	D	4	3	11,11,12	0.34	0	15,15,17	0.44	0
3	MAN	D	5	3	11,11,12	0.25	0	15,15,17	0.62	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
2	NAG	В	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	В	2	2	-	0/6/23/26	0/1/1/1
2	BMA	В	3	2	-	1/2/19/22	0/1/1/1
2	MAN	В	4	2	-	0/2/19/22	0/1/1/1
2	MAN	В	5	2	-	2/2/19/22	0/1/1/1
2	MAN	В	6	2	-	0/2/19/22	0/1/1/1
3	NAG	D	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	D	2	3	-	0/6/23/26	0/1/1/1
3	BMA	D	3	3	-	0/2/19/22	0/1/1/1
3	MAN	D	4	3	-	1/2/19/22	0/1/1/1
3	MAN	D	5	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	В	3	BMA	C1-C2-C3	2.71	113.00	109.67

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	В	5	MAN	O5-C5-C6-O6
3	D	4	MAN	O5-C5-C6-O6
2	В	3	BMA	C4-C5-C6-O6
2	В	5	MAN	C4-C5-C6-O6

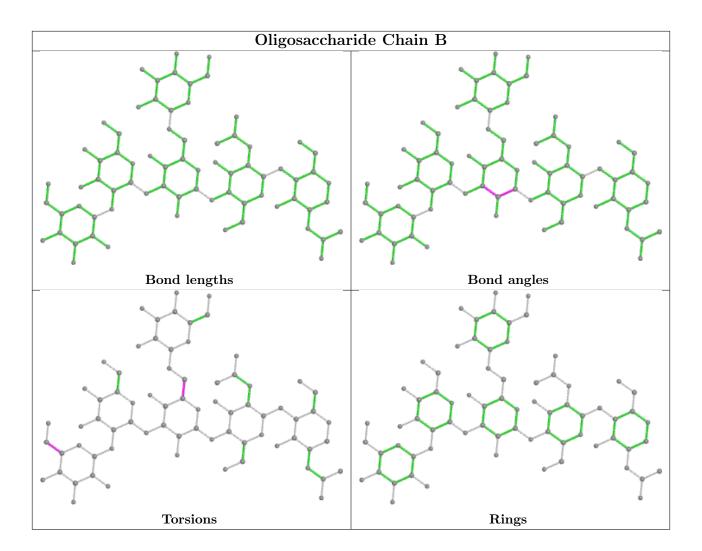
There are no ring outliers.

1 monomer is involved in 1 short contact:

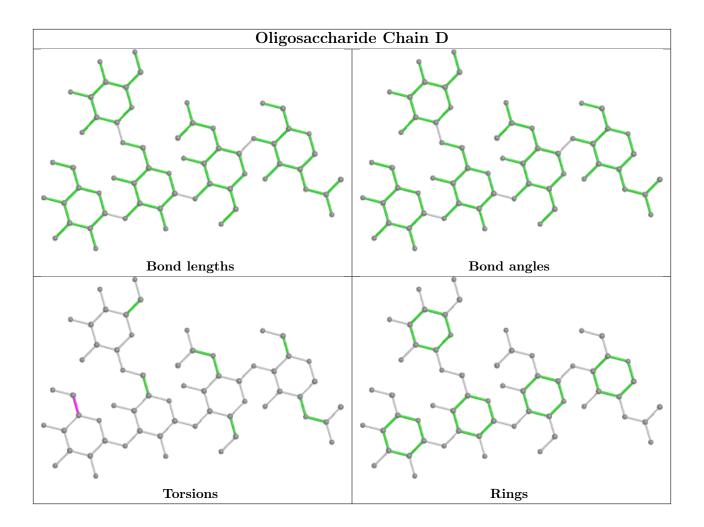
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	В	1	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.









### 5.6 Ligand geometry (i)

19 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	Tuno	Chain	Res	Link	Bond lengths			Bond angles		
MIOI	Type				Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	С	503	-	4,4,4	0.42	0	6,6,6	0.07	0
5	SO4	A	509	-	4,4,4	0.15	0	6,6,6	0.08	0
5	SO4	A	503	-	4,4,4	0.46	0	6,6,6	0.07	0
4	GLC	С	501	1	11,11,12	0.56	0	15,15,17	0.61	0
5	SO4	A	502	-	4,4,4	0.16	0	6,6,6	0.10	0
5	SO4	A	506	-	4,4,4	0.14	0	6,6,6	0.05	0
5	SO4	A	505	-	4,4,4	0.15	0	6,6,6	0.05	0



Mal	Trino	Chain	Dag	T inle	Bo	Bond lengths			Bond angles		
Mol	Type	Chain	Res	Link	Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2	
5	SO4	A	508	-	4,4,4	0.15	0	6,6,6	0.10	0	
5	SO4	A	504	-	4,4,4	0.13	0	6,6,6	0.05	0	
5	SO4	С	509	-	4,4,4	0.14	0	6,6,6	0.05	0	
5	SO4	A	507	-	4,4,4	0.14	0	6,6,6	0.05	0	
5	SO4	С	508	-	4,4,4	0.13	0	6,6,6	0.14	0	
5	SO4	С	502	-	4,4,4	0.46	0	6,6,6	0.07	0	
5	SO4	С	506	-	4,4,4	0.12	0	6,6,6	0.10	0	
5	SO4	С	505	-	4,4,4	0.17	0	6,6,6	0.07	0	
4	GLC	A	501	1	11,11,12	0.45	0	15,15,17	0.84	0	
5	SO4	С	504	-	4,4,4	0.18	0	6,6,6	0.20	0	
5	SO4	С	510	-	4,4,4	0.12	0	6,6,6	0.06	0	
5	SO4	С	507	-	4,4,4	0.13	0	6,6,6	0.20	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.

	$\mathbf{Mol}$	$\mathbf{Type}$	Chain	$\operatorname{Res}$	Link	Chirals	Torsions	Rings
	4	GLC	A	501	1	-	0/2/19/22	0/1/1/1
ſ	4	GLC	С	501	1	-	0/2/19/22	0/1/1/1

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.

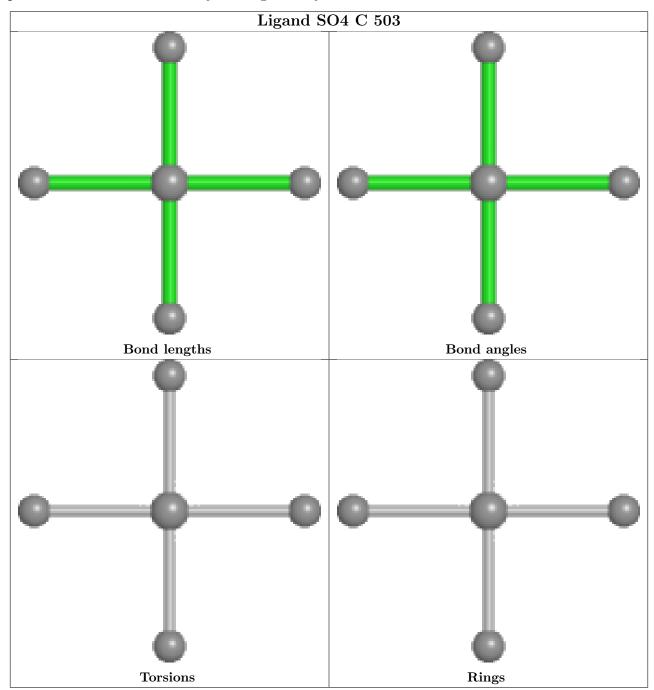
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	С	503	SO4	0	1

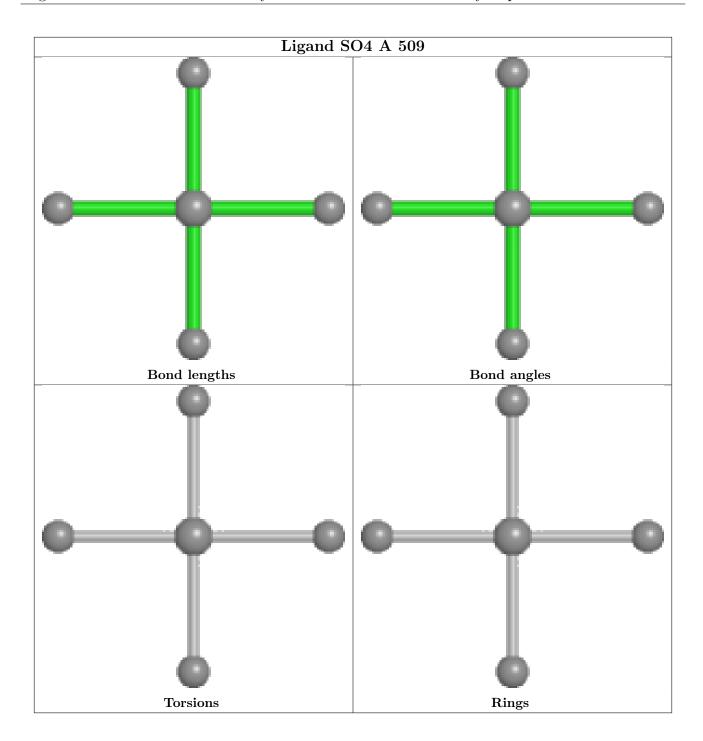
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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



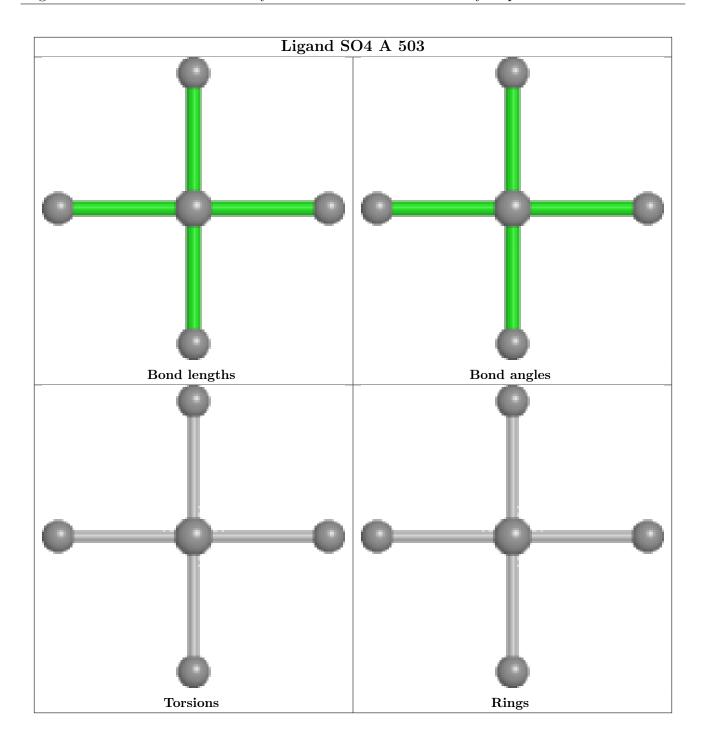
any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



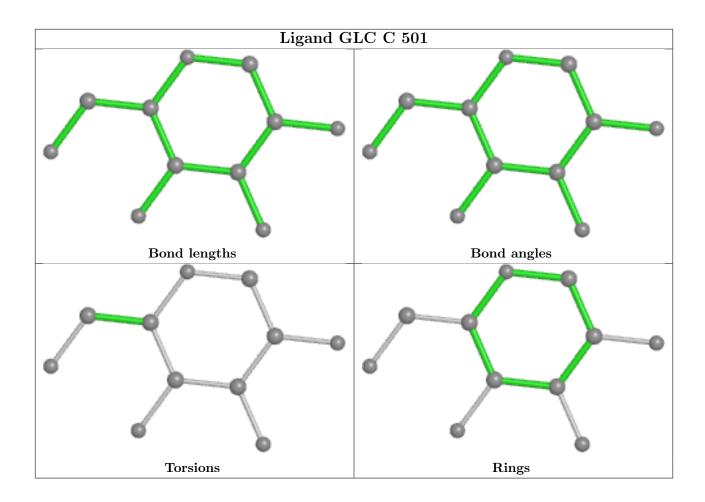




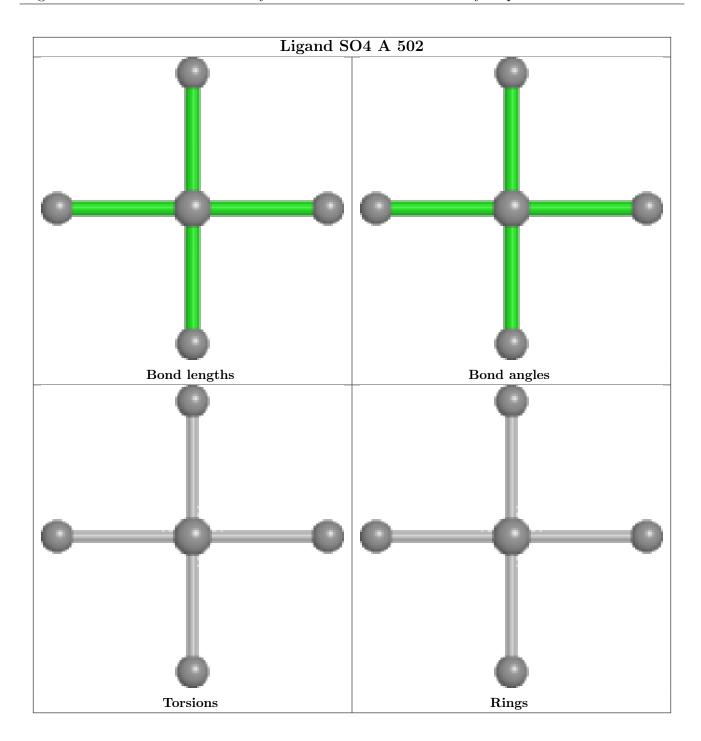




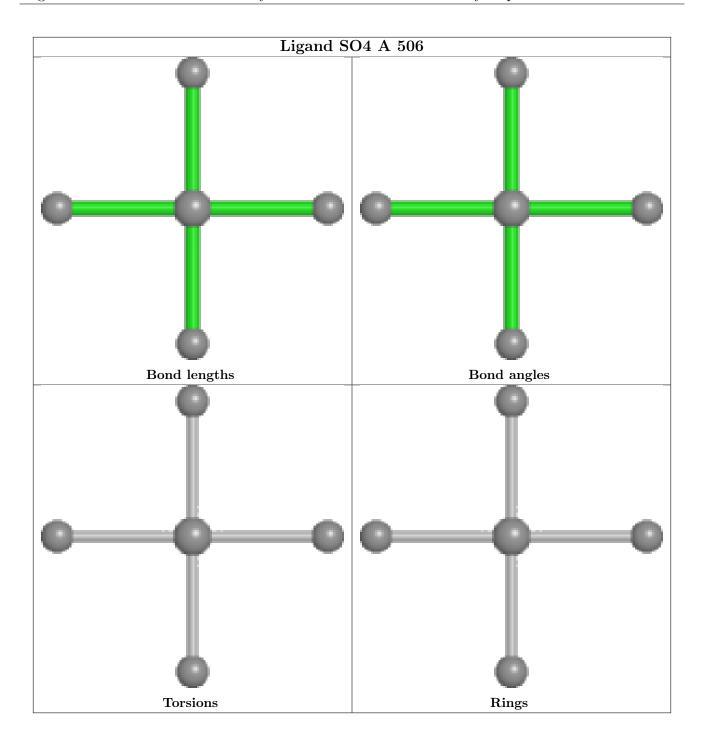




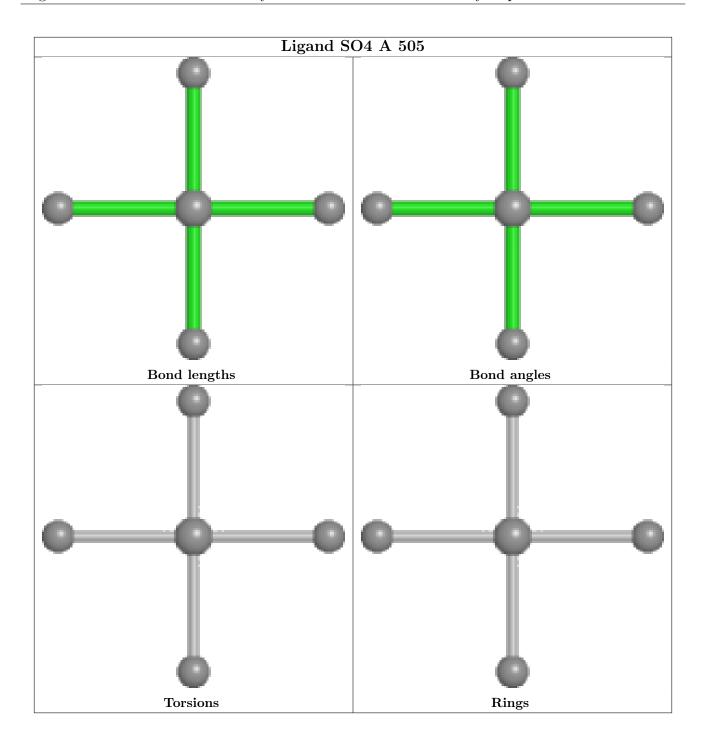




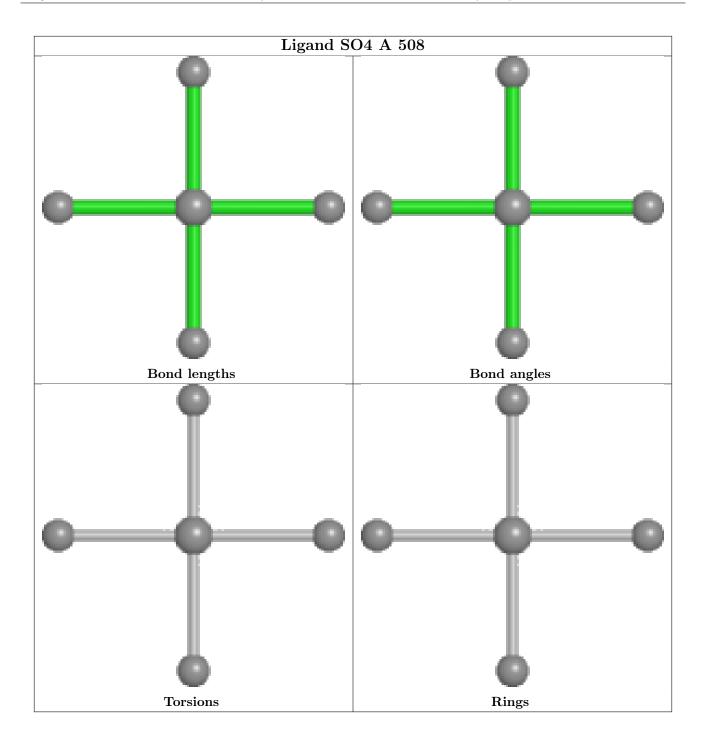




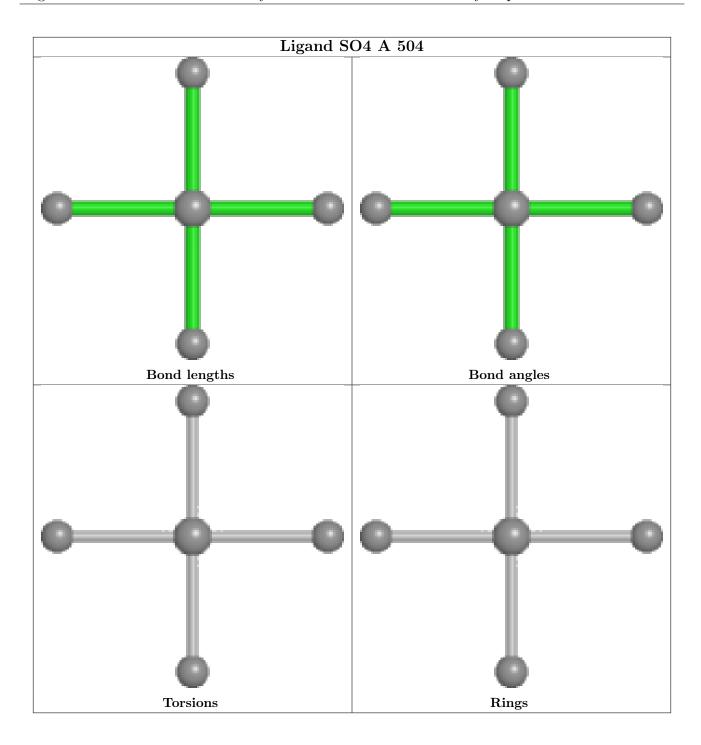




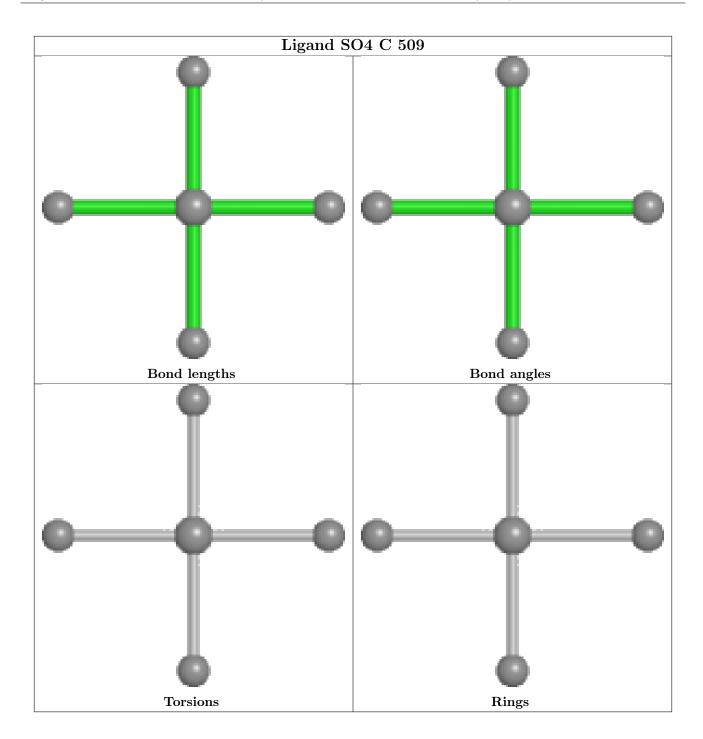




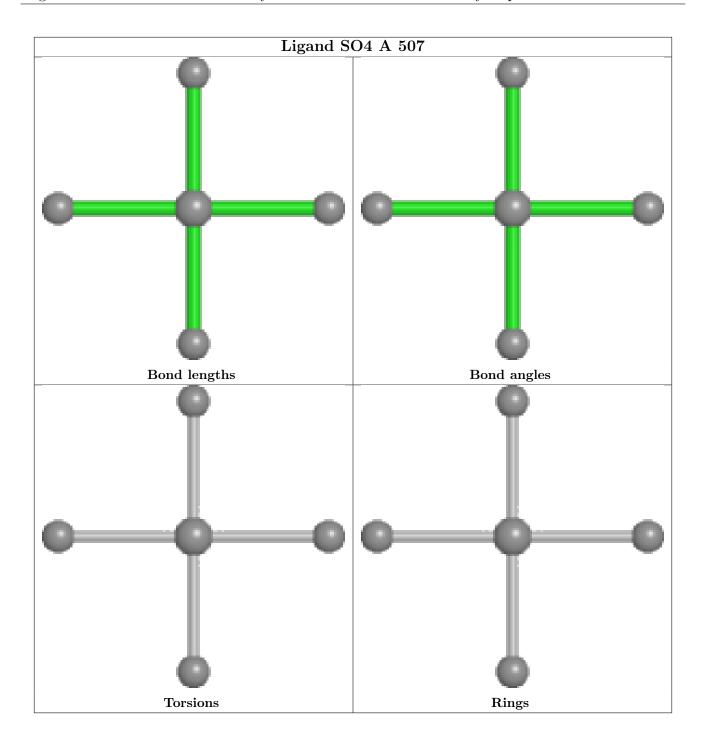




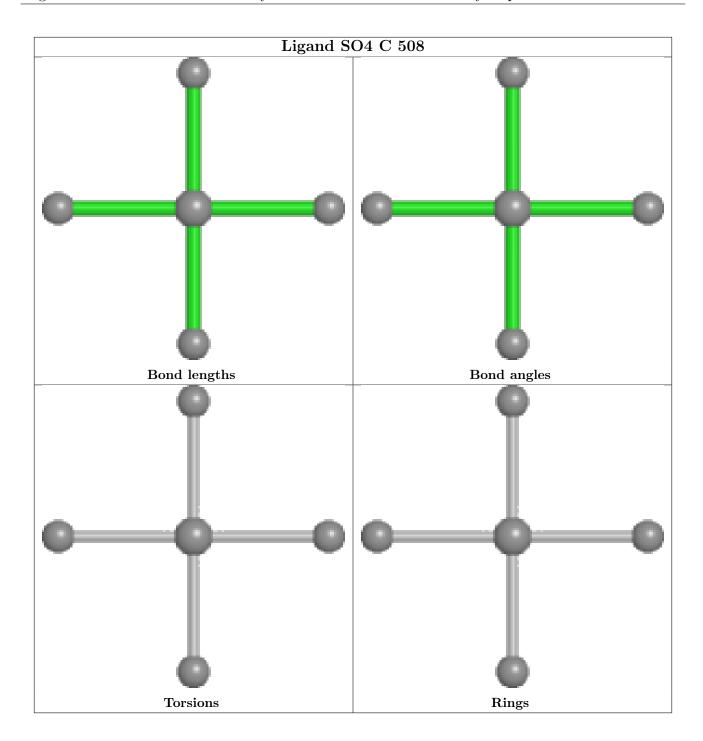




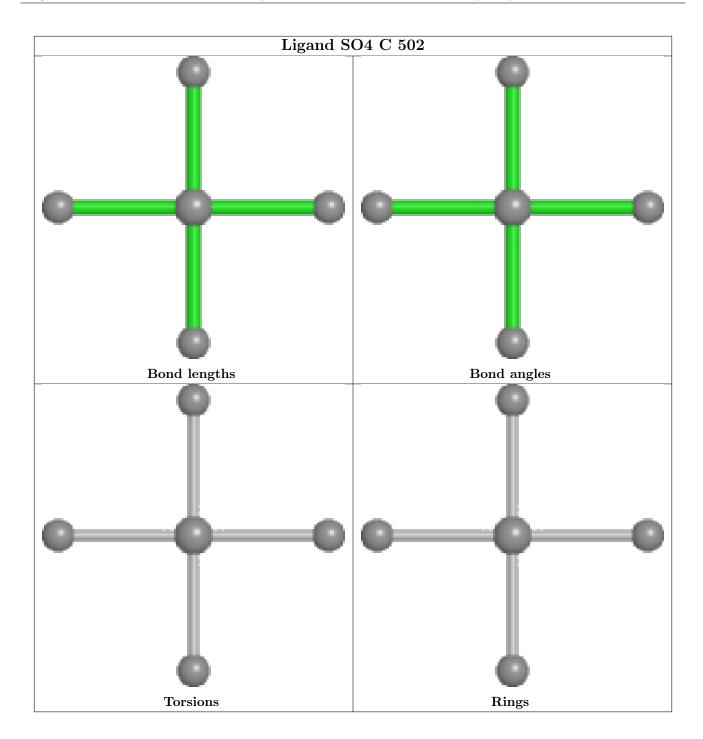




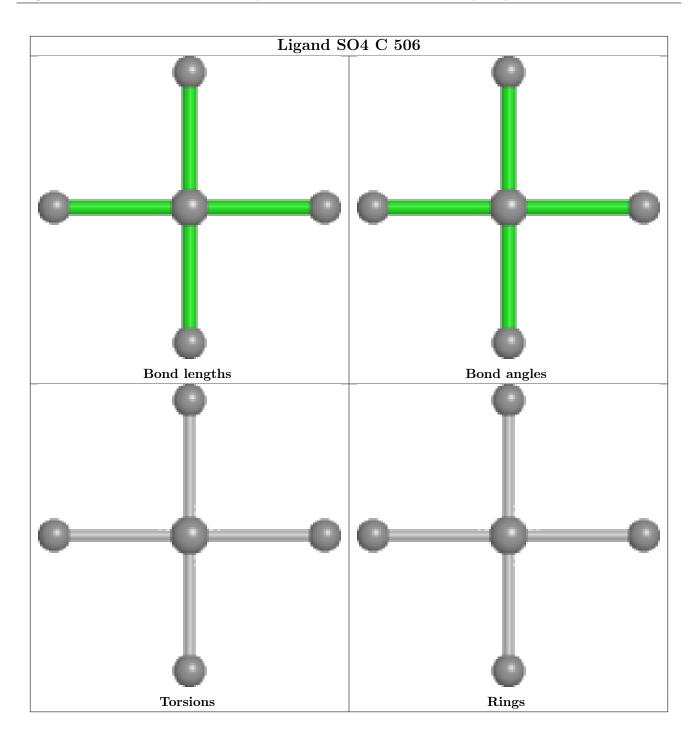




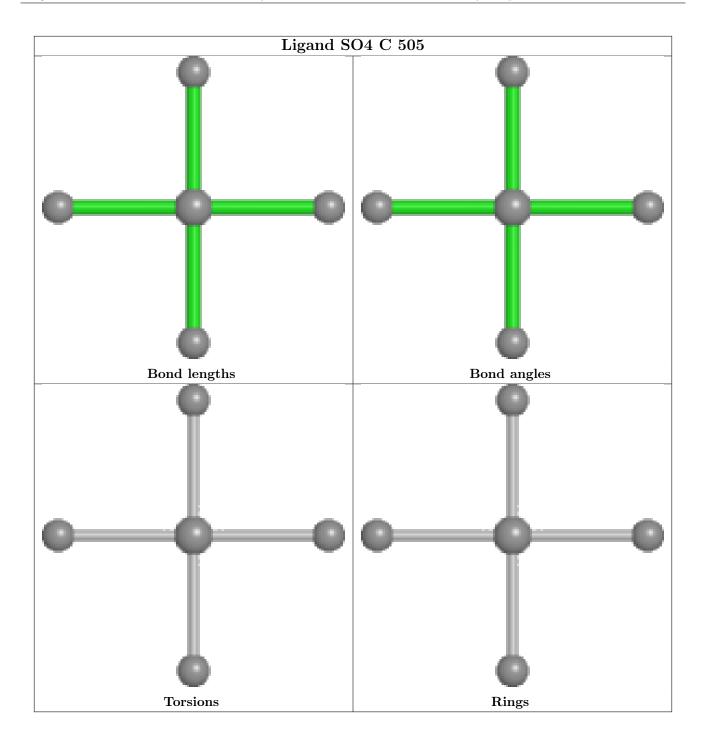




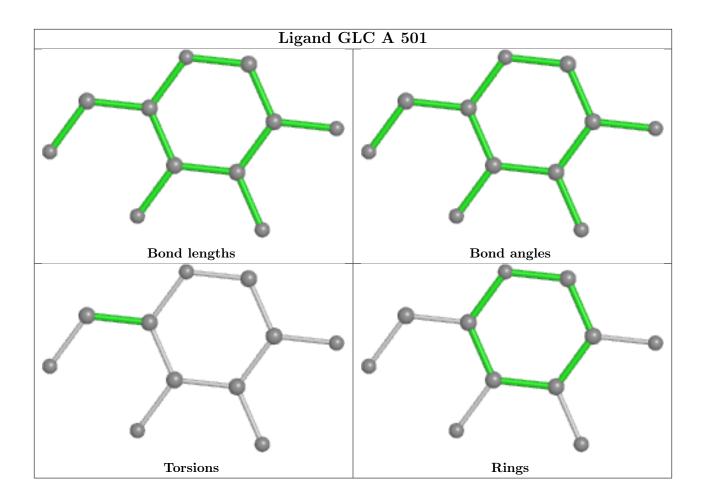




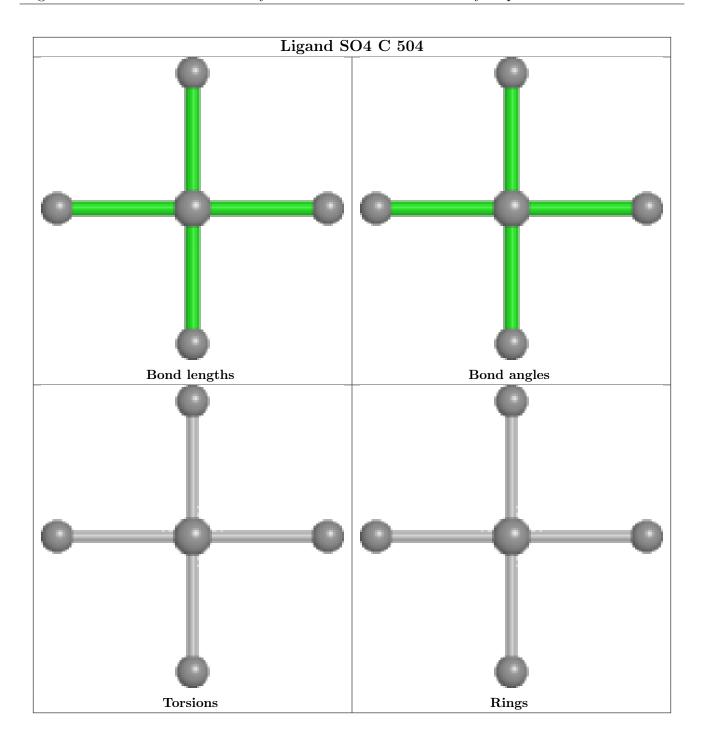




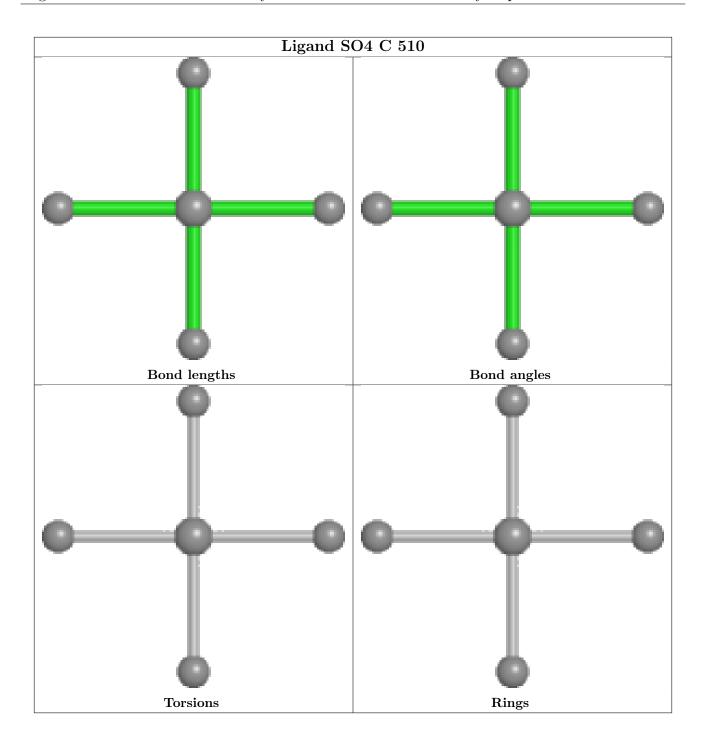




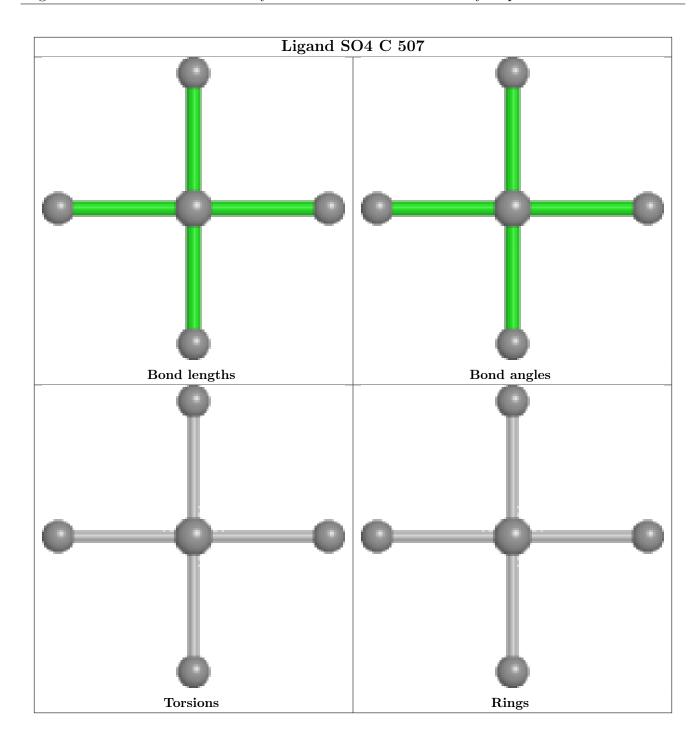












## 5.7 Other polymers (i)

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 (i)

### 6.1 Protein, DNA and RNA chains (i)

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	$\langle { m RSRZ} \rangle$	# RSRZ > 2		$\mathbf{OWAB}(\mathbf{\mathring{A}}^2)$	Q < 0.9
1	A	368/415 (88%)	0.10	11 (2%) 50	56	25, 41, 76, 110	0
1	С	368/415 (88%)	-0.10	6 (1%) 72 7	79	25, 36, 58, 84	0
All	All	736/830 (88%)	0.00	17 (2%) 60	67	25, 38, 67, 110	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	54	LEU	5.9
1	A	55	PRO	4.0
1	A	50	LEU	3.4
1	С	53	ALA	3.3
1	С	52	THR	3.2

### 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)

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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
2	MAN	В	5	11/12	0.81	0.29	77,82,89,91	0
3	MAN	D	4	11/12	0.83	0.17	56,60,69,73	0
3	BMA	D	3	11/12	0.89	0.12	34,45,48,49	0
2	MAN	В	4	11/12	0.89	0.11	49,54,70,71	0
2	MAN	В	6	11/12	0.91	0.12	49,56,64,64	0

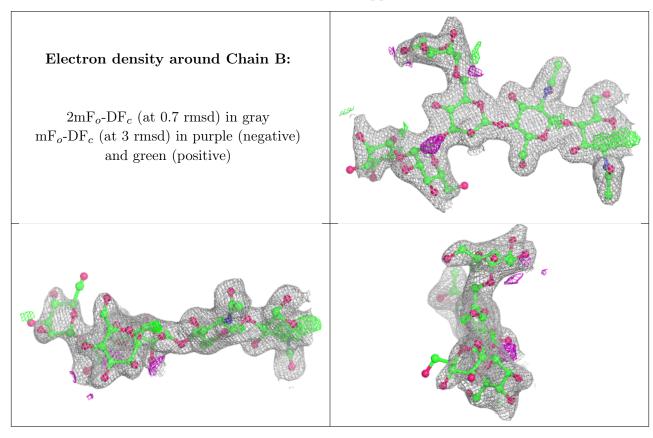
Continued on next page...



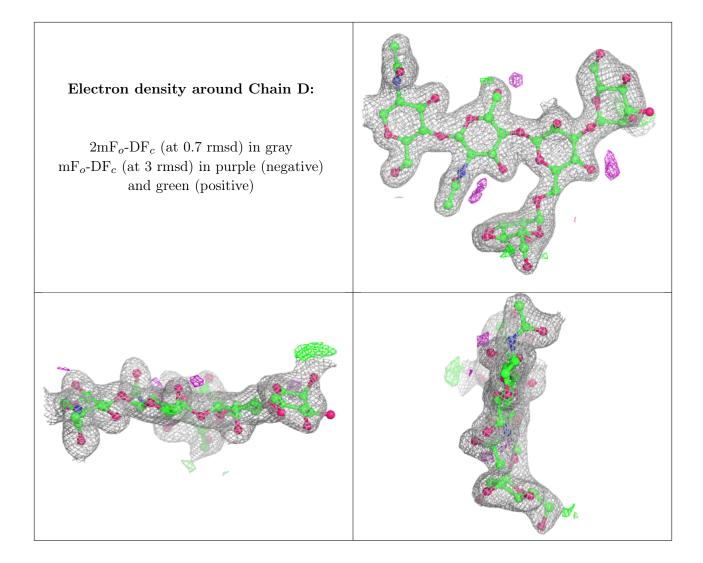
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
2	BMA	В	3	11/12	0.92	0.10	34,39,52,52	0
3	NAG	D	1	14/15	0.94	0.08	28,29,32,32	0
3	NAG	D	2	14/15	0.94	0.09	30,32,43,43	0
2	NAG	В	2	14/15	0.95	0.10	30,32,40,43	0
3	MAN	D	5	11/12	0.95	0.11	39,49,56,60	0
2	NAG	В	1	14/15	0.96	0.10	26,31,34,35	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.







# 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
5	SO4	A	504	5/5	0.70	0.37	66,75,77,100	5
5	SO4	С	503	5/5	0.80	0.17	31,41,45,49	5
5	SO4	A	507	5/5	0.81	0.24	68,81,102,106	0
5	SO4	A	508	5/5	0.84	0.17	61,65,94,96	0
5	SO4	С	505	5/5	0.86	0.18	41,44,46,56	5
5	SO4	С	510	5/5	0.90	0.14	61,68,73,83	0
5	SO4	A	505	5/5	0.91	0.17	61,61,69,69	5
5	SO4	A	509	5/5	0.91	0.32	64,69,83,83	0
5	SO4	С	502	5/5	0.91	0.14	42,49,51,52	5

Continued on next page...



Continued from previous page...

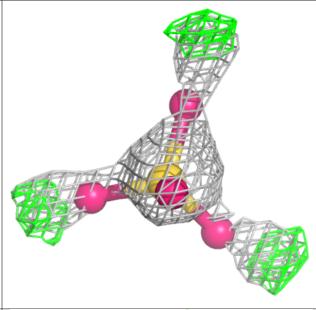
Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
5	SO4	С	509	5/5	0.92	0.20	72,74,76,85	5
4	GLC	A	501	11/12	0.92	0.08	32,38,40,41	0
5	SO4	С	506	5/5	0.93	0.10	38,43,57,60	5
5	SO4	A	502	5/5	0.94	0.11	43,44,52,52	5
5	SO4	A	503	5/5	0.96	0.09	39,40,47,57	5
4	GLC	С	501	11/12	0.96	0.07	26,30,34,35	0
5	SO4	С	504	5/5	0.97	0.09	43,46,53,57	5
5	SO4	С	507	5/5	0.97	0.09	44,49,49,51	5
5	SO4	A	506	5/5	0.98	0.10	80,84,84,85	5
5	SO4	С	508	5/5	0.98	0.07	33,41,43,54	5

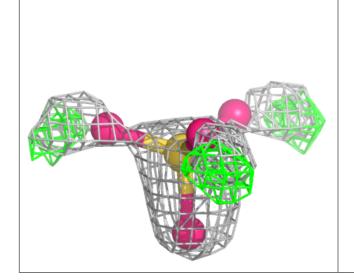
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

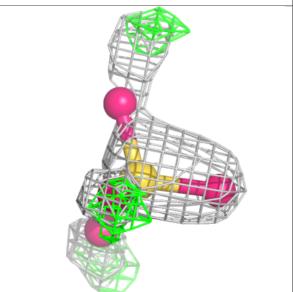


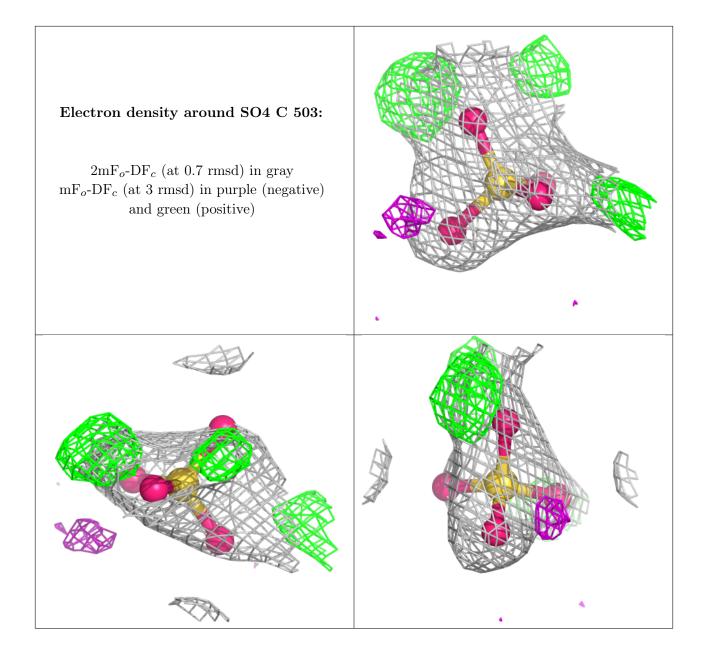
# Electron density around SO4 A 504: $2 \mathrm{mF}_o\text{-}\mathrm{DF}_c \ (\mathrm{at}\ 0.7\ \mathrm{rmsd}) \ \mathrm{in}\ \mathrm{gray}$

 ${
m mF}_o{
m -DF}_c$  (at 3 rmsd) in purple (negative) and green (positive)



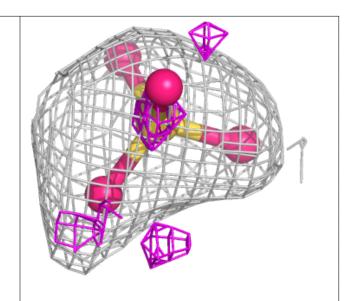


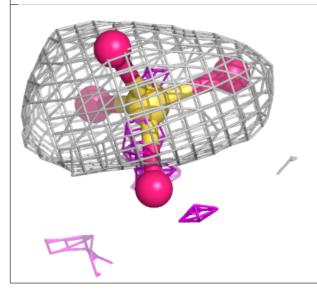


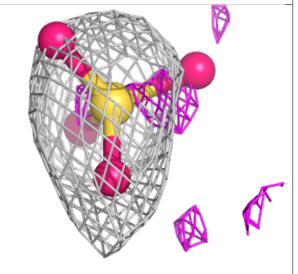




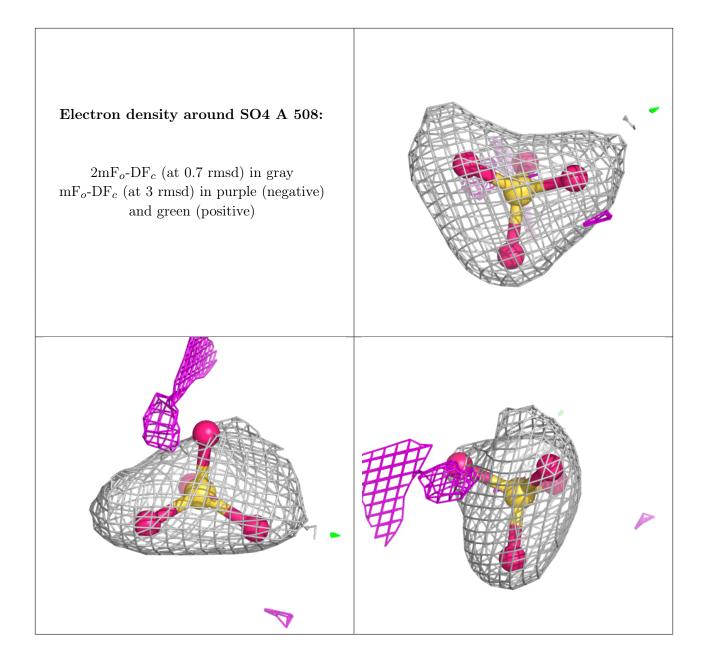
# Electron density around SO4 A 507:



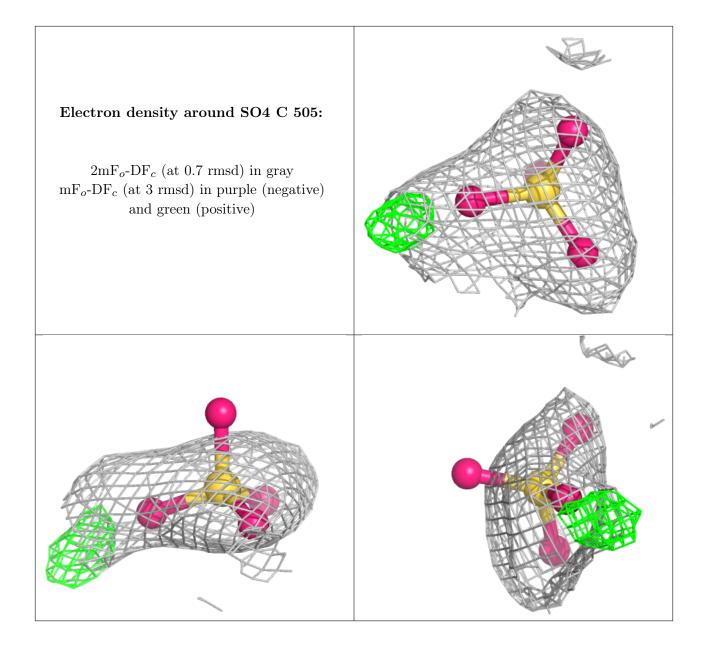








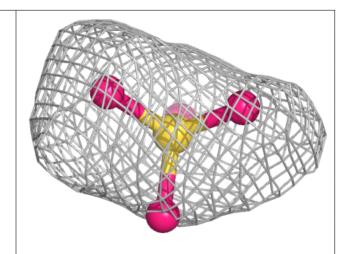


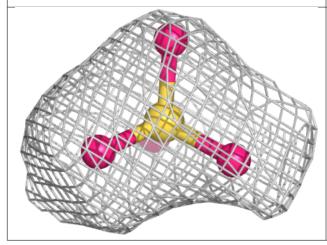


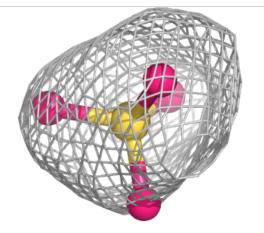


## Electron density around SO4 C 510:

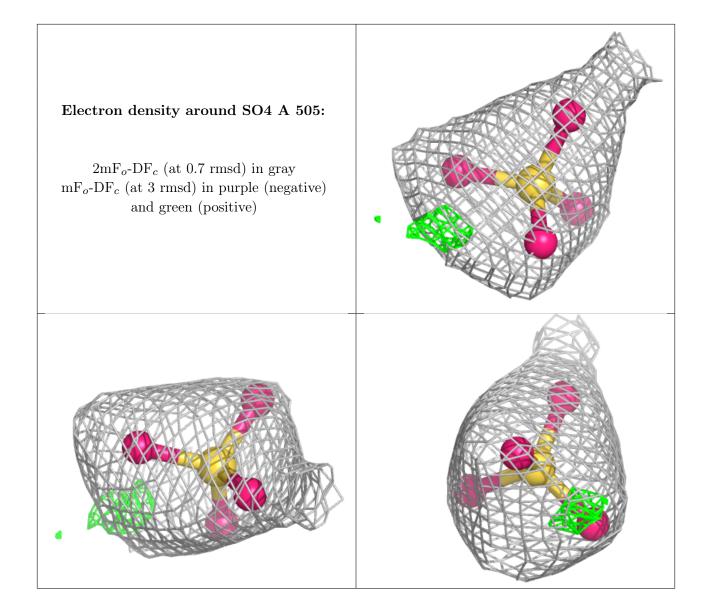
 $2 \text{mF}_o\text{-DF}_c$  (at 0.7 rmsd) in gray  $\text{mF}_o\text{-DF}_c$  (at 3 rmsd) in purple (negative) and green (positive)







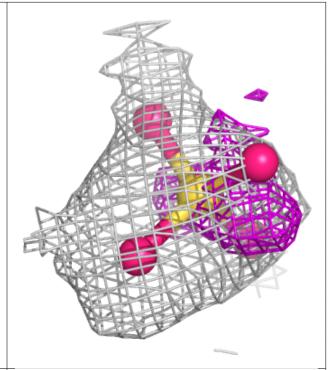


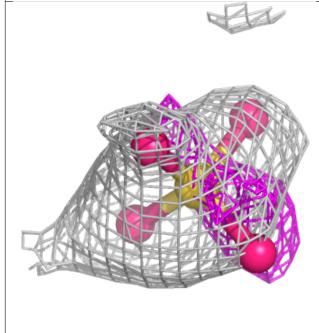


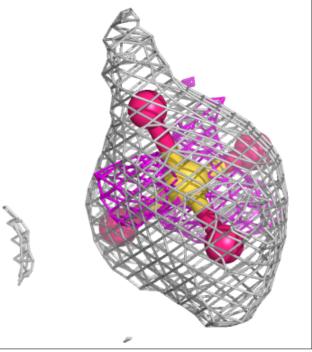


## Electron density around SO4 A 509:

 $2 \mathrm{mF}_o\text{-}\mathrm{DF}_c$  (at 0.7 rmsd) in gray  $\mathrm{mF}_o\text{-}\mathrm{DF}_c$  (at 3 rmsd) in purple (negative) and green (positive)

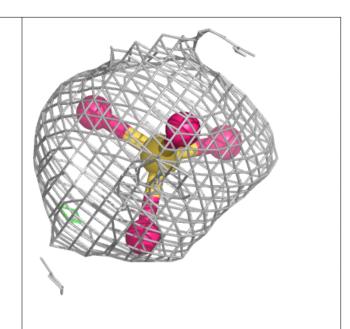


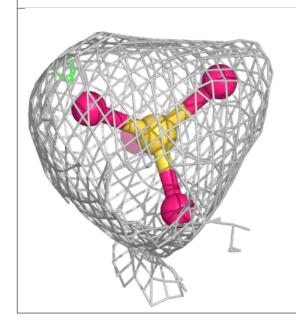


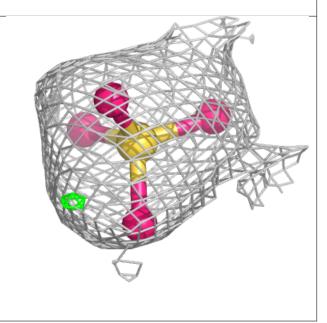




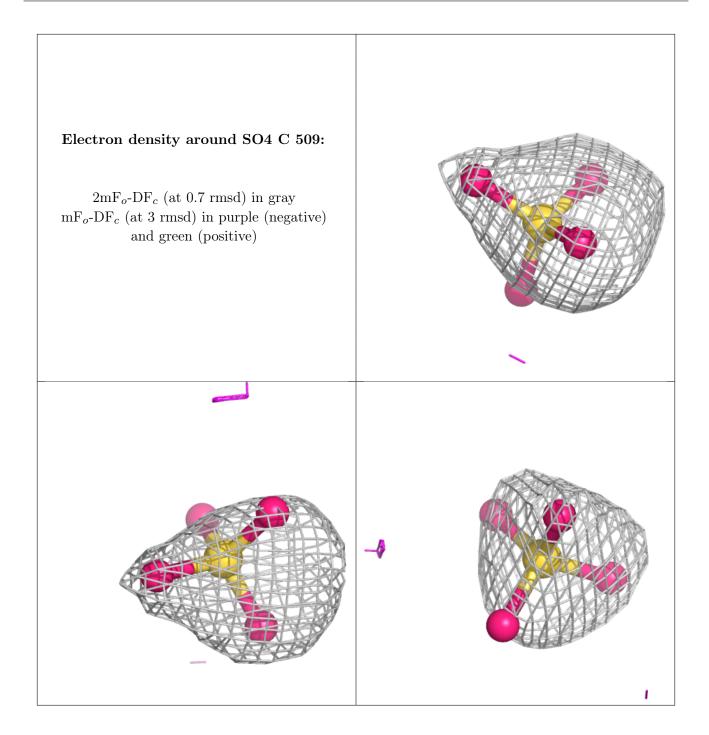
## Electron density around SO4 C 502:



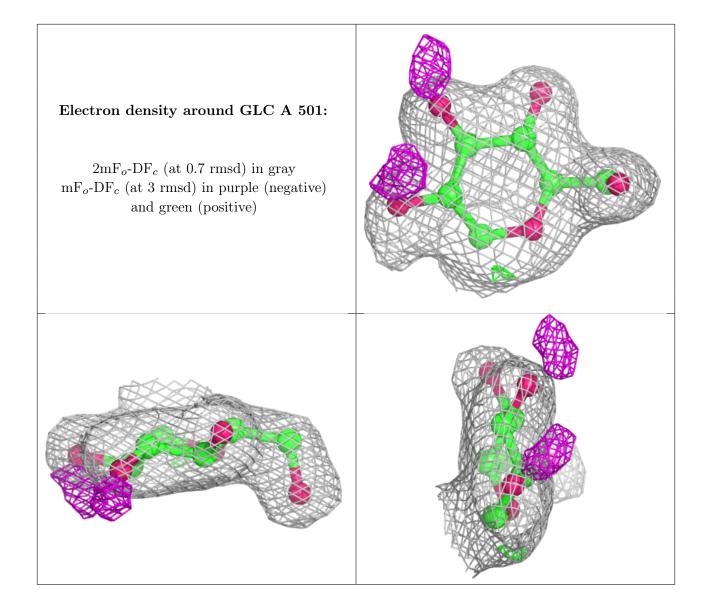






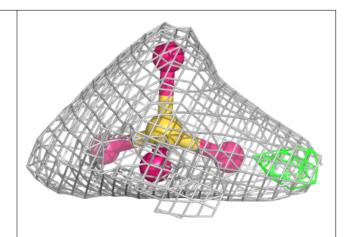


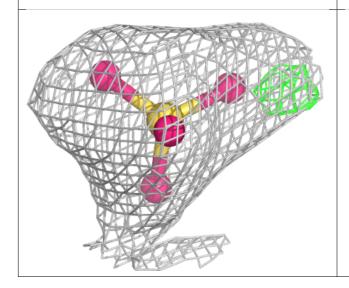


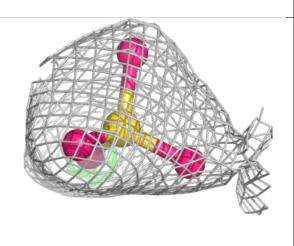




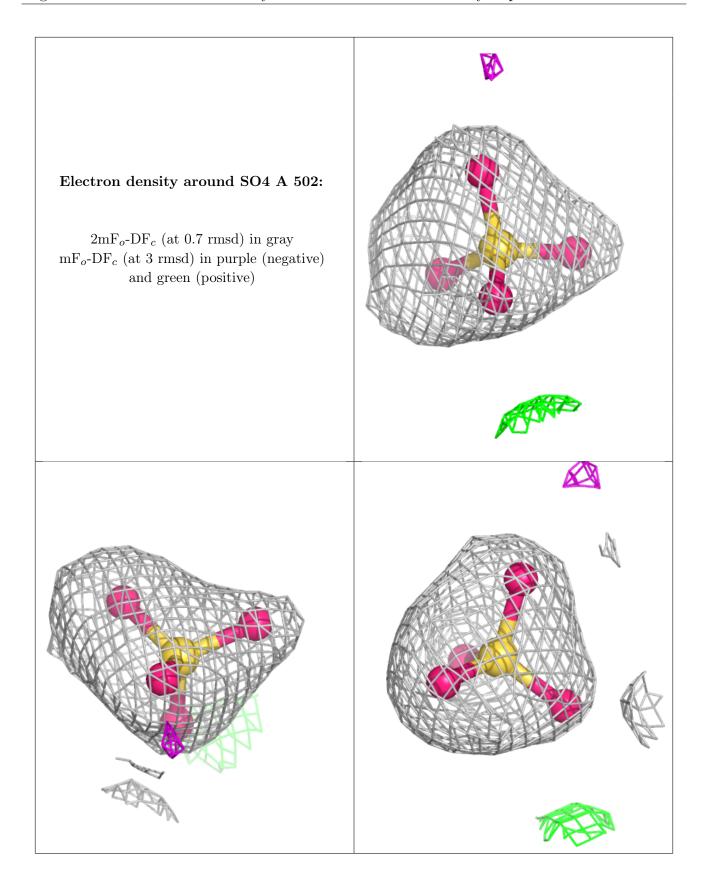
# Electron density around SO4 C 506:





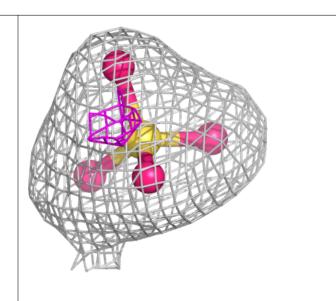


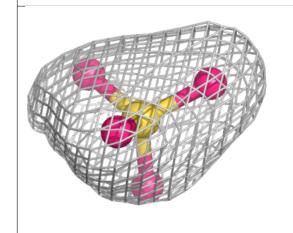




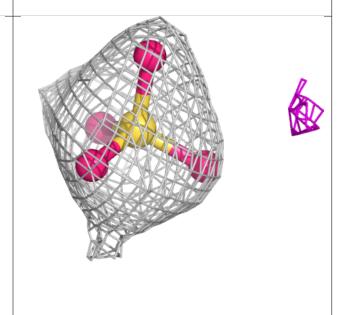


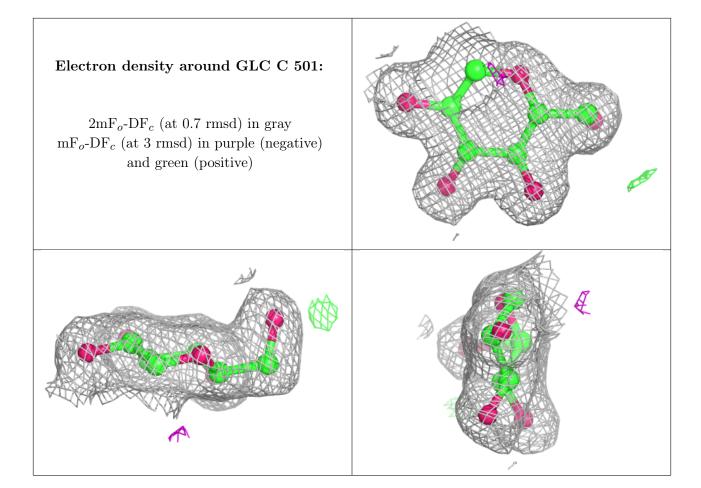
#### Electron density around SO4 A 503:





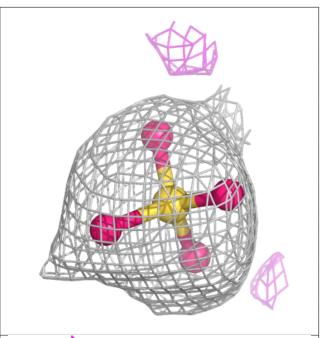


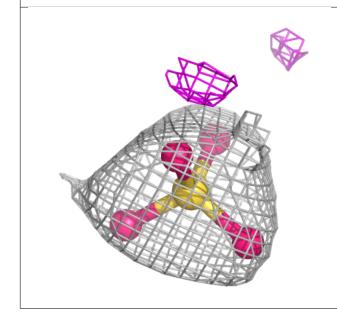


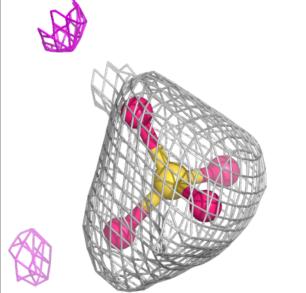




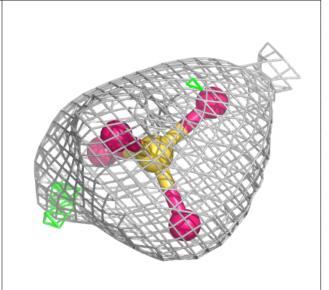
#### Electron density around SO4 C 504:

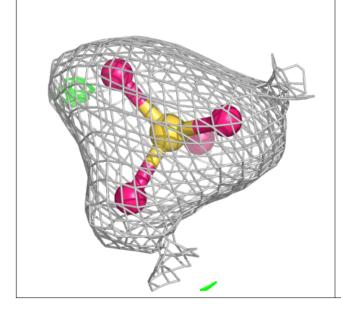


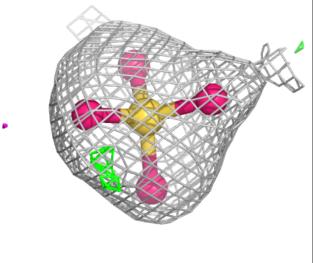




#### Electron density around SO4 C 507:



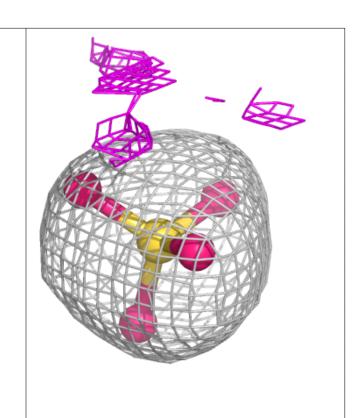


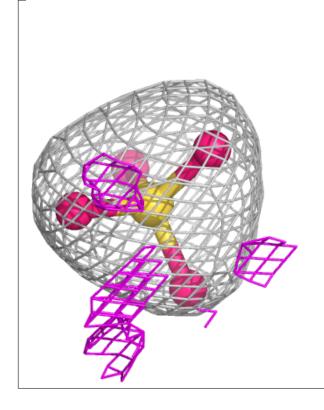


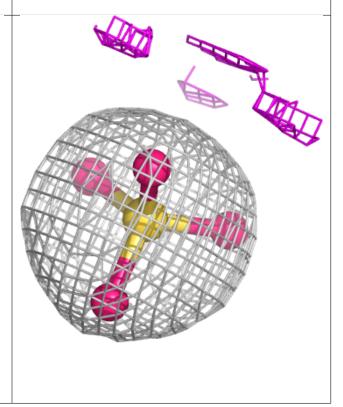


# Electron density around SO4 A 506:

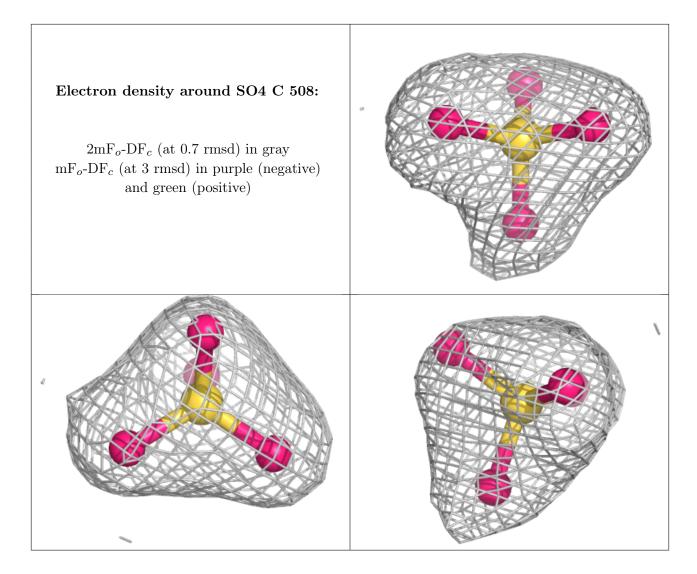
 $2 \mathrm{mF}_o\text{-}\mathrm{DF}_c$  (at 0.7 rmsd) in gray  $\mathrm{mF}_o\text{-}\mathrm{DF}_c$  (at 3 rmsd) in purple (negative) and green (positive)











# 6.5 Other polymers (i)

There are no such residues in this entry.

