Full wwPDB X-ray Structure Validation Report ⓘ

May 19, 2022 – 02:50 pm BST

PDB ID	:	7PD2
Title	:	Crystal structure of the substrate-free radical SAM tyrosine lyase ThiH (2-iminoacetate synthase) from Thermosinus carboxydivorans
Authors	:	Amara, P.; Saragaglia, C.; Mouesca, J.-M.; Martin, L.; Nicolet, Y.
Deposited on	:	2021-08-04
Resolution	:	1.99 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 following versions of software and data (see references ⓘ) were used in the production of this report:

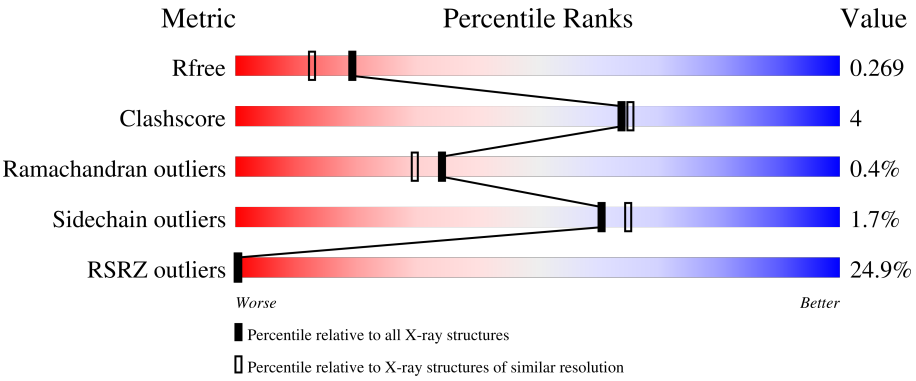
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

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 1.99 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	377	23% 86% 11%
1	B	377	26% 87% 10%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6180 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.

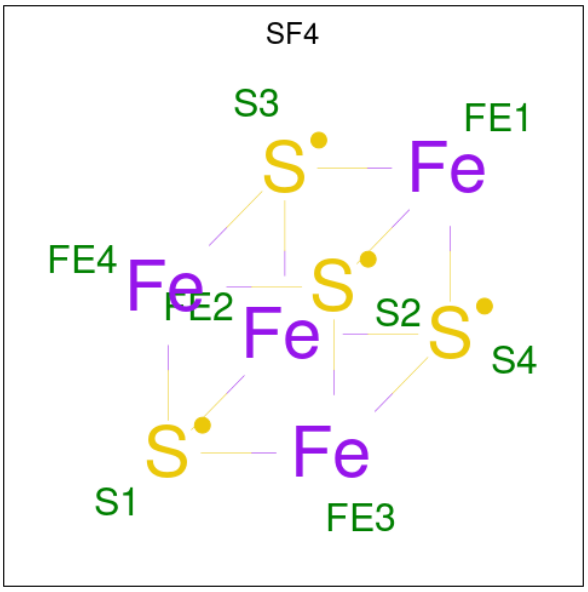
- Molecule 1 is a protein called Thiazole biosynthesis protein ThiH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	367	Total	C	N	O	S	0	6	0
			2933	1878	509	531	15			
1	B	365	Total	C	N	O	S	0	2	0
			2882	1850	497	519	16			

There are 22 discrepancies between the modelled and reference sequences:

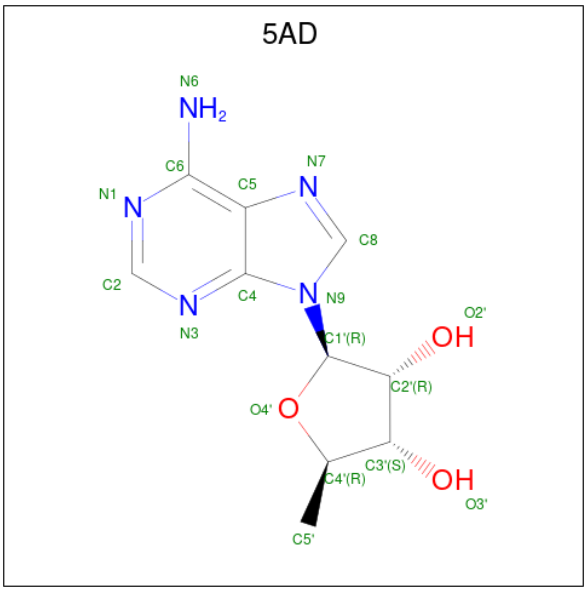
Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	initiating methionine	UNP A1HPQ5
A	-8	TRP	-	expression tag	UNP A1HPQ5
A	-7	SER	-	expression tag	UNP A1HPQ5
A	-6	HIS	-	expression tag	UNP A1HPQ5
A	-5	PRO	-	expression tag	UNP A1HPQ5
A	-4	GLN	-	expression tag	UNP A1HPQ5
A	-3	PHE	-	expression tag	UNP A1HPQ5
A	-2	GLU	-	expression tag	UNP A1HPQ5
A	-1	LYS	-	expression tag	UNP A1HPQ5
A	0	ALA	-	expression tag	UNP A1HPQ5
A	1	SER	-	expression tag	UNP A1HPQ5
B	-9	MET	-	initiating methionine	UNP A1HPQ5
B	-8	TRP	-	expression tag	UNP A1HPQ5
B	-7	SER	-	expression tag	UNP A1HPQ5
B	-6	HIS	-	expression tag	UNP A1HPQ5
B	-5	PRO	-	expression tag	UNP A1HPQ5
B	-4	GLN	-	expression tag	UNP A1HPQ5
B	-3	PHE	-	expression tag	UNP A1HPQ5
B	-2	GLU	-	expression tag	UNP A1HPQ5
B	-1	LYS	-	expression tag	UNP A1HPQ5
B	0	ALA	-	expression tag	UNP A1HPQ5
B	1	SER	-	expression tag	UNP A1HPQ5

- Molecule 2 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Fe S 8 4 4	0	0
2	B	1	Total Fe S 8 4 4	0	0

- Molecule 3 is 5'-DEOXYADENOSINE (three-letter code: 5AD) (formula: $C_{10}H_{13}N_5O_3$) (labeled as "Ligand of Interest" by depositor).



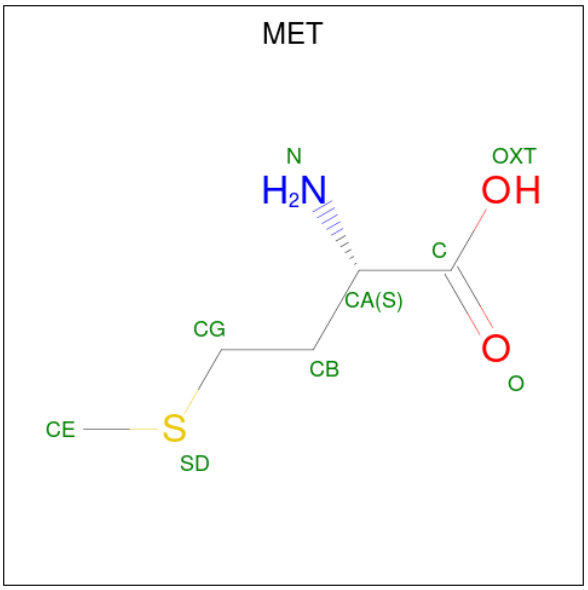
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 18 10 5 3	0	0

Continued on next page...

Continued from previous page...

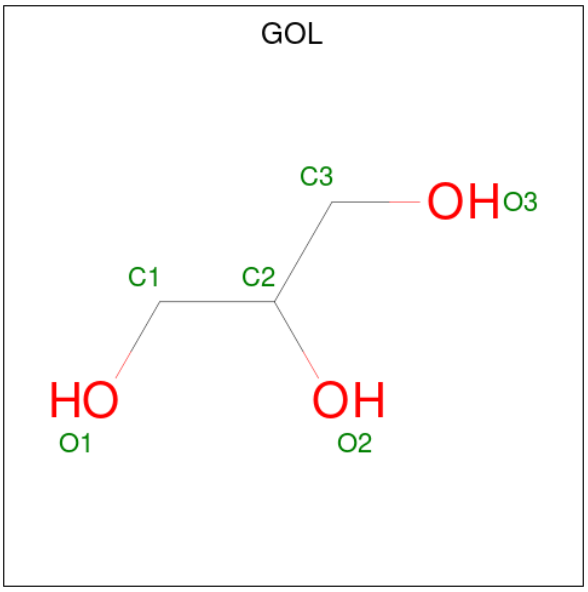
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			18	10	5	3		

- Molecule 4 is METHIONINE (three-letter code: MET) (formula: $C_5H_{11}NO_2S$) (labeled as "Ligand of Interest" by depositor).



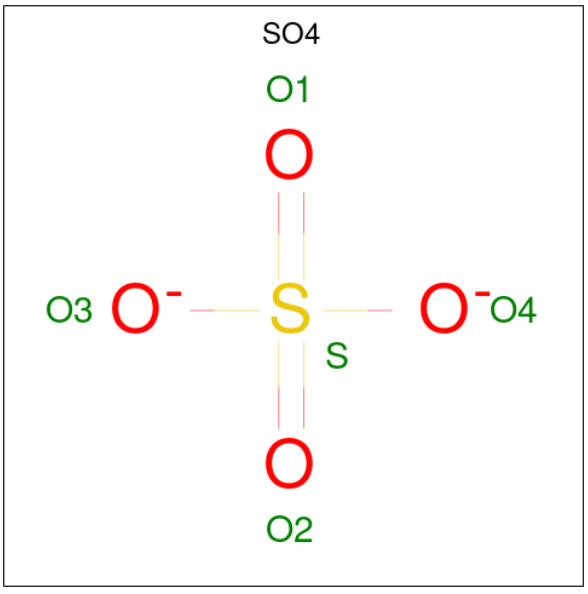
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			9	5	1	2	1		
4	B	1	Total	C	N	O	S	0	0
			9	5	1	2	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	B	1	Total C O 6 3 3	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 7 is water.

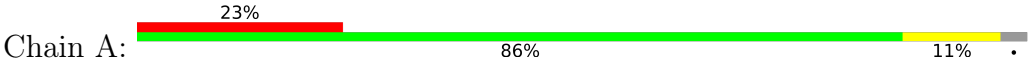
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	130	Total O 130 130	0	0
7	B	117	Total O 117 117	0	0

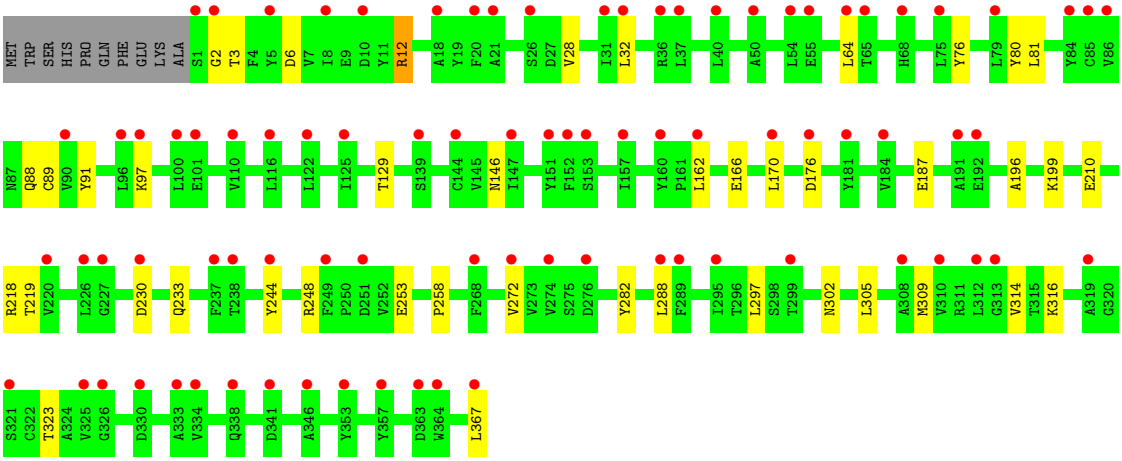
3 Residue-property plots (i)

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: Thiazole biosynthesis protein ThiH

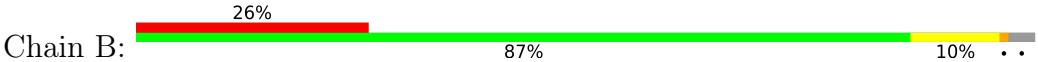
Chain A:

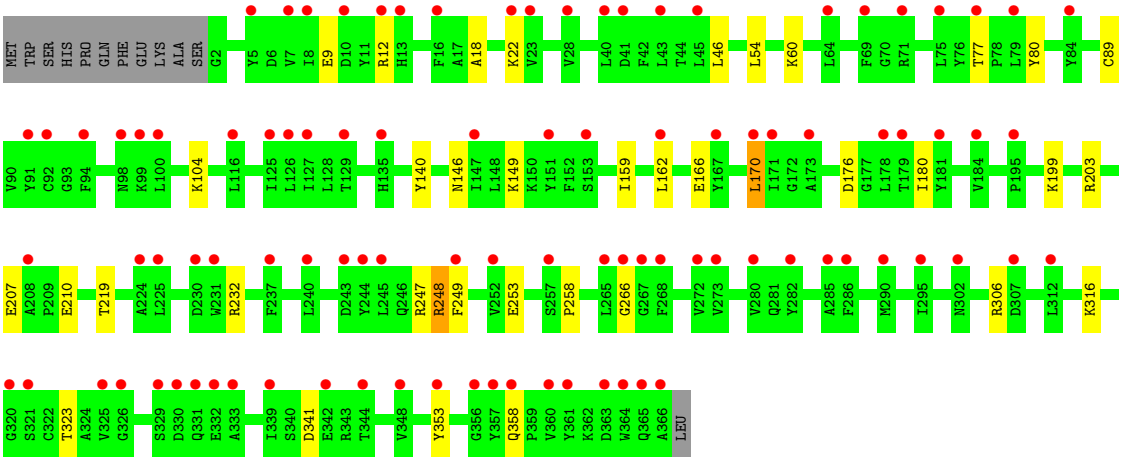




• Molecule 1: Thiazole biosynthesis protein ThiH

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	175.09Å 49.33Å 85.21Å 90.00° 96.91° 90.00°	Depositor
Resolution (Å)	47.46 – 1.99 47.46 – 1.99	Depositor EDS
% Data completeness (in resolution range)	97.5 (47.46-1.99) 97.5 (47.46-1.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.21	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.230 , 0.269 0.230 , 0.269	Depositor DCC
R_{free} test set	2435 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	39.7	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6180	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.90% 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: SF4, SO4, GOL, 5AD

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.28	0/3017	0.43	0/4085
1	B	0.28	0/2951	0.43	0/3997
All	All	0.28	0/5968	0.43	0/8082

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	2933	0	2920	25	0
1	B	2882	0	2849	24	0
2	A	8	0	0	0	0
2	B	8	0	0	0	0
3	A	18	0	13	1	0
3	B	18	0	13	0	0
4	A	9	0	8	0	0
4	B	9	0	8	0	0
5	A	12	0	16	1	0
5	B	6	0	8	1	0
6	A	15	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	15	0	0	0	0
7	A	130	0	0	4	0
7	B	117	0	0	7	0
All	All	6180	0	5835	49	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 4.

All (49) 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:B:266:GLY:H	5:B:404:GOL:H31	1.53	0.73
1:B:353:TYR:OH	7:B:501:HOH:O	2.09	0.71
1:A:210:GLU:OE1	1:A:248:ARG:NH2	2.27	0.68
1:B:210:GLU:OE1	1:B:248:ARG:NH1	2.25	0.67
1:B:176:ASP:OD2	7:B:502:HOH:O	2.14	0.66
1:B:46:LEU:O	1:B:232:ARG:NH1	2.29	0.65
1:A:367:LEU:OXT	7:A:501:HOH:O	2.15	0.64
1:A:219:THR:HG22	1:A:253:GLU:HB2	1.83	0.60
1:A:302:ASN:ND2	7:A:508:HOH:O	2.33	0.60
1:A:146:ASN:OD1	7:A:502:HOH:O	2.17	0.59
1:B:146:ASN:HA	1:B:149:LYS:HE2	1.85	0.59
1:B:18:ALA:O	1:B:22:LYS:HG2	2.04	0.57
1:B:247:ARG:NE	7:B:506:HOH:O	2.38	0.57
1:A:12:ARG:NH1	1:A:187:GLU:OE2	2.38	0.57
1:B:219:THR:HG22	1:B:253:GLU:HB2	1.88	0.55
1:A:302:ASN:ND2	7:A:504:HOH:O	2.27	0.55
1:B:358:GLN:NE2	7:B:501:HOH:O	2.25	0.52
1:A:297:LEU:HD21	1:A:309:MET:HE3	1.93	0.50
1:A:76:TYR:CD1	1:A:316:LYS:HE2	2.47	0.50
1:A:305:LEU:HG	1:A:309:MET:HE1	1.93	0.50
1:B:146:ASN:ND2	7:B:505:HOH:O	2.44	0.49
1:B:248:ARG:HG2	1:B:249:PHE:CE2	2.48	0.49
1:A:89:CYS:SG	1:A:199:LYS:HD3	2.52	0.48
1:A:88:GLN:HG3	1:A:196:ALA:HB3	1.96	0.47
1:B:306:ARG:NH1	1:B:341:ASP:OD2	2.44	0.47
1:B:60:LYS:NZ	7:B:514:HOH:O	2.47	0.47
1:A:80:TYR:CE1	1:A:323:THR:HA	2.51	0.46
1:A:233:GLN:HE22	5:A:405:GOL:H31	1.82	0.45
1:A:81:LEU:HB2	1:A:129:THR:HG22	1.99	0.44
1:B:89:CYS:SG	1:B:199:LYS:HD3	2.58	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:28:VAL:O	1:A:32:LEU:HG	2.18	0.44
1:A:162:LEU:HB3	1:A:166:GLU:HB2	2.00	0.44
1:B:203:ARG:NH1	1:B:207:GLU:OE2	2.50	0.44
1:A:2:GLY:N	1:A:6:ASP:OD2	2.51	0.43
1:A:64:LEU:HD12	1:A:288[B]:LEU:HD23	2.00	0.43
1:A:305:LEU:HG	1:A:309:MET:CE	2.49	0.43
1:B:80:TYR:CE1	1:B:323:THR:HA	2.54	0.42
1:B:162:LEU:HB3	1:B:166:GLU:HB2	2.01	0.42
1:A:297:LEU:HD22	1:A:314:VAL:HG21	2.02	0.42
1:A:91:TYR:O	3:A:402:5AD:N6	2.43	0.42
1:B:316:LYS:NZ	7:B:520:HOH:O	2.53	0.42
1:A:88:GLN:HE22	1:A:97:LYS:HE2	1.85	0.41
1:A:244:TYR:OH	1:A:248:ARG:NH2	2.51	0.41
1:B:170:LEU:HD12	1:B:170:LEU:HA	1.89	0.41
1:A:3:THR:HA	1:A:248:ARG:HH22	1.86	0.41
1:B:54:LEU:HD23	1:B:54:LEU:HA	1.89	0.41
1:B:9:GLU:HG2	1:B:12:ARG:NH1	2.36	0.40
1:B:104:LYS:HE3	1:B:140:TYR:CD1	2.56	0.40
1:B:159:ILE:HD11	1:B:162:LEU:HD11	2.04	0.40

There are no symmetry-related clashes.

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	Percentiles	
1	A	371/377 (98%)	363 (98%)	6 (2%)	2 (0%)	29	23
1	B	365/377 (97%)	358 (98%)	6 (2%)	1 (0%)	41	37
All	All	736/754 (98%)	721 (98%)	12 (2%)	3 (0%)	34	30

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	258	PRO
1	B	258	PRO
1	A	230	ASP

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	306/320 (96%)	300 (98%)	6 (2%)	55	58
1	B	295/320 (92%)	291 (99%)	4 (1%)	67	72
All	All	601/640 (94%)	591 (98%)	10 (2%)	60	65

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	ARG
1	A	170	LEU
1	A	176	ASP
1	A	218	ARG
1	A	272	VAL
1	A	282	TYR
1	B	77	THR
1	B	170	LEU
1	B	180	ILE
1	B	248	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	233	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

15 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	$\#\|Z\| > 2$
4	MET	A	403	2	4,8,8	0.19	0	2,9,9	0.14	0
6	SO4	B	406	-	4,4,4	0.13	0	6,6,6	0.07	0
4	MET	B	403	2	4,8,8	0.25	0	2,9,9	0.04	0
6	SO4	A	408	-	4,4,4	0.15	0	6,6,6	0.12	0
6	SO4	B	405	-	4,4,4	0.13	0	6,6,6	0.14	0
3	5AD	B	402	-	17,20,20	4.22	8 (47%)	15,30,30	1.84	4 (26%)
6	SO4	B	407	-	4,4,4	0.13	0	6,6,6	0.08	0
2	SF4	B	401	1,4	0,12,12	-	-	-		
5	GOL	A	405	-	5,5,5	0.89	0	5,5,5	0.97	0
2	SF4	A	401	1,4	0,12,12	-	-	-		
5	GOL	A	404	-	5,5,5	0.91	0	5,5,5	1.00	0
3	5AD	A	402	-	17,20,20	4.19	8 (47%)	15,30,30	1.94	4 (26%)
6	SO4	A	406	-	4,4,4	0.18	0	6,6,6	0.12	0
5	GOL	B	404	-	5,5,5	0.85	0	5,5,5	0.98	0
6	SO4	A	407	-	4,4,4	0.13	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
4	MET	A	403	2	-	1/4/8/8	-
4	MET	B	403	2	-	1/4/8/8	-
3	5AD	B	402	-	-	0/0/20/20	0/3/3/3
2	SF4	B	401	1,4	-	-	0/6/5/5
5	GOL	A	405	-	-	1/4/4/4	-
2	SF4	A	401	1,4	-	-	0/6/5/5
5	GOL	A	404	-	-	0/4/4/4	-
3	5AD	A	402	-	-	0/0/20/20	0/3/3/3
5	GOL	B	404	-	-	2/4/4/4	-

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	5AD	C3'-C2'	-10.44	1.24	1.53
3	A	402	5AD	C3'-C2'	-10.44	1.24	1.53
3	B	402	5AD	C3'-C4'	8.29	1.64	1.52
3	A	402	5AD	C3'-C4'	8.26	1.64	1.52
3	B	402	5AD	O4'-C1'	6.92	1.50	1.41
3	A	402	5AD	O4'-C1'	6.89	1.50	1.41
3	A	402	5AD	O4'-C4'	-5.34	1.28	1.44
3	B	402	5AD	O4'-C4'	-5.32	1.28	1.44
3	B	402	5AD	O2'-C2'	4.09	1.52	1.43
3	A	402	5AD	O2'-C2'	3.98	1.52	1.43
3	B	402	5AD	C6-N6	3.21	1.45	1.34
3	A	402	5AD	C6-N6	3.13	1.45	1.34
3	B	402	5AD	C2'-C1'	3.04	1.58	1.53
3	A	402	5AD	C2'-C1'	2.86	1.58	1.53
3	B	402	5AD	O3'-C3'	2.68	1.49	1.43
3	A	402	5AD	O3'-C3'	2.55	1.49	1.43

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	402	5AD	N3-C2-N1	-4.52	121.62	128.68
3	A	402	5AD	N3-C2-N1	-4.38	121.83	128.68
3	A	402	5AD	C5'-C4'-C3'	-4.33	111.15	115.70
3	B	402	5AD	C5'-C4'-C3'	-3.57	111.95	115.70
3	A	402	5AD	C3'-C2'-C1'	3.03	105.53	100.98
3	B	402	5AD	C3'-C2'-C1'	2.69	105.03	100.98
3	A	402	5AD	C4-C5-N7	-2.14	107.17	109.40
3	B	402	5AD	C4-C5-N7	-2.03	107.28	109.40

There are no chirality outliers.

All (5) torsion outliers are listed below:

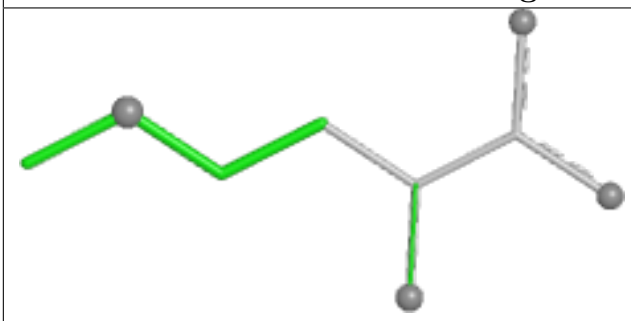
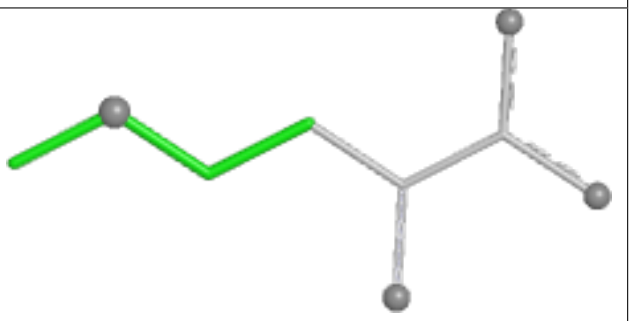
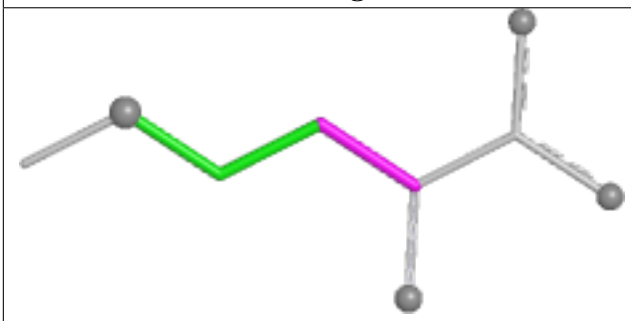
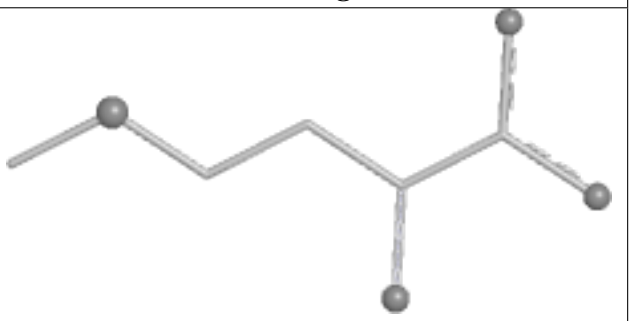
Mol	Chain	Res	Type	Atoms
5	B	404	GOL	O1-C1-C2-C3
5	A	405	GOL	O1-C1-C2-C3
5	B	404	GOL	O1-C1-C2-O2
4	A	403	MET	C-CA-CB-CG
4	B	403	MET	C-CA-CB-CG

There are no ring outliers.

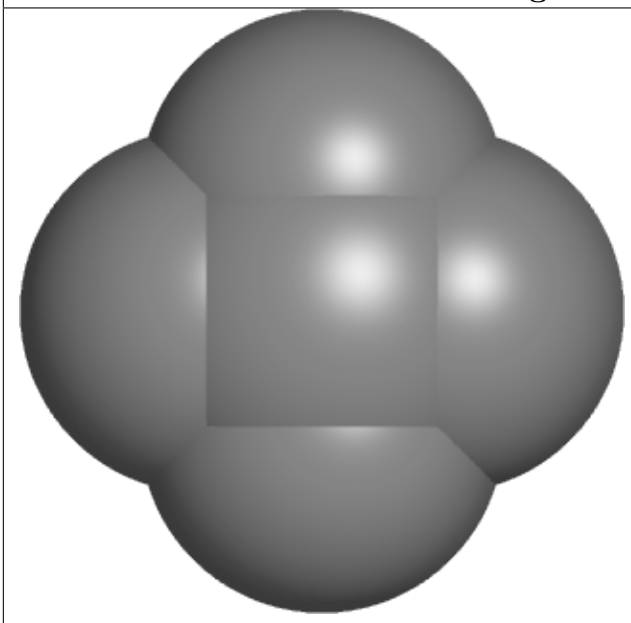
3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	405	GOL	1	0
3	A	402	5AD	1	0
5	B	404	GOL	1	0

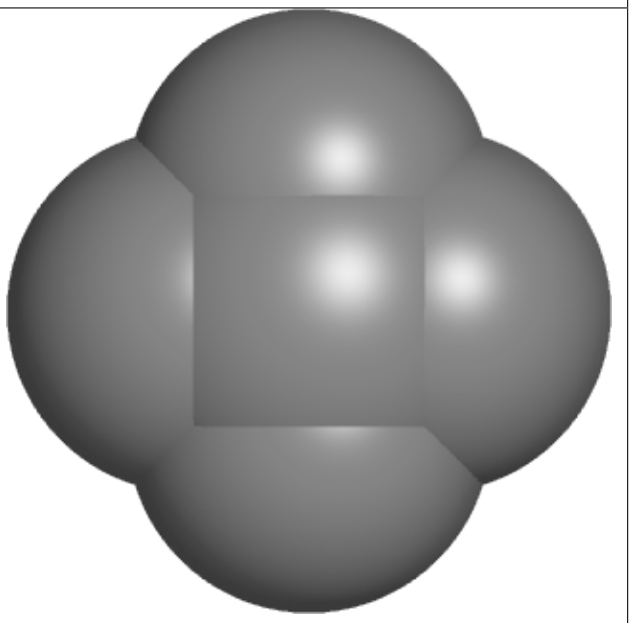
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 MET A 403	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

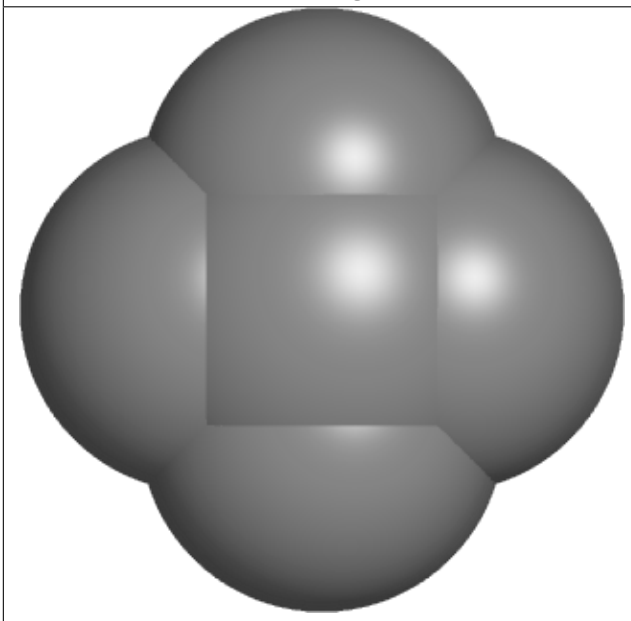
Ligand SO4 B 406



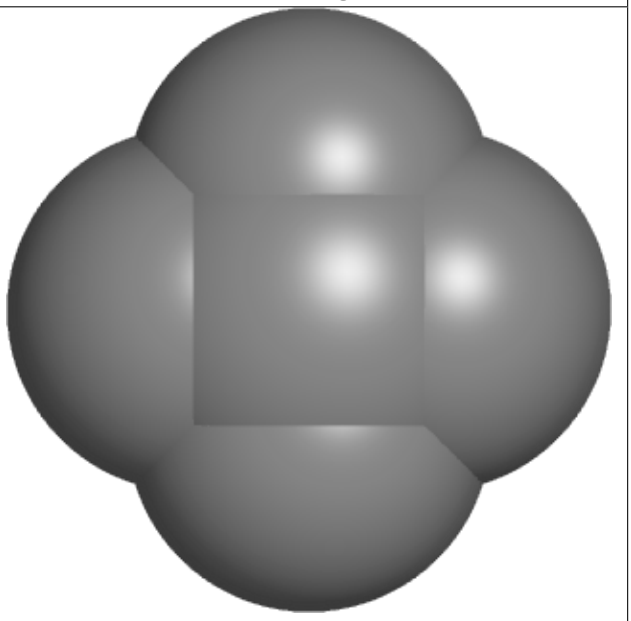
Bond lengths



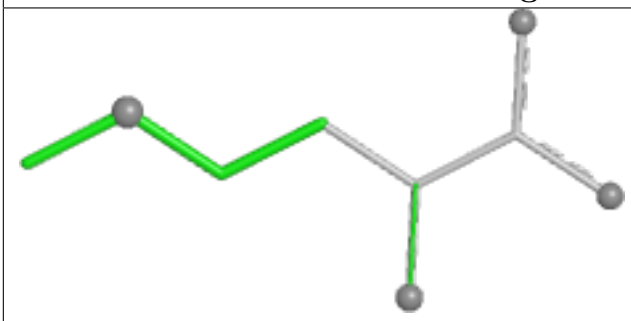
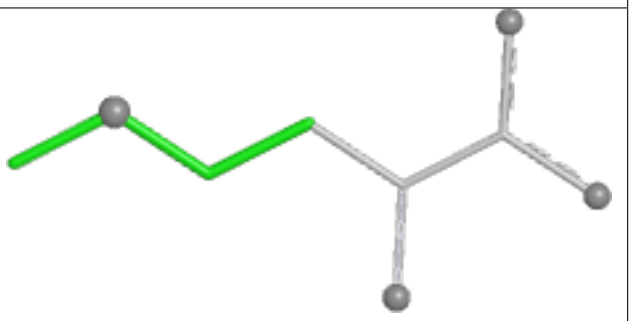
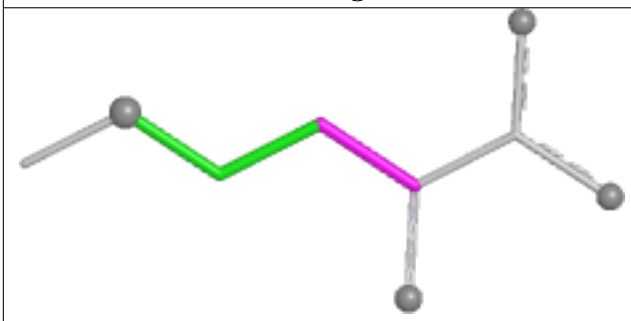
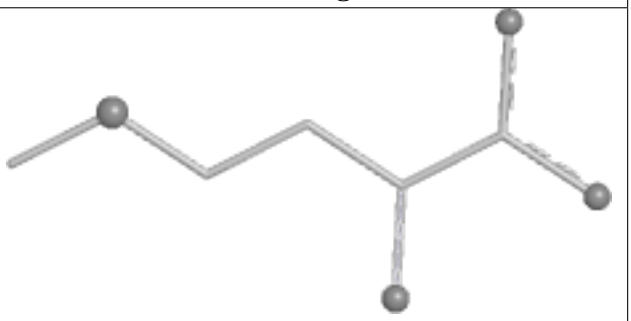
Bond angles



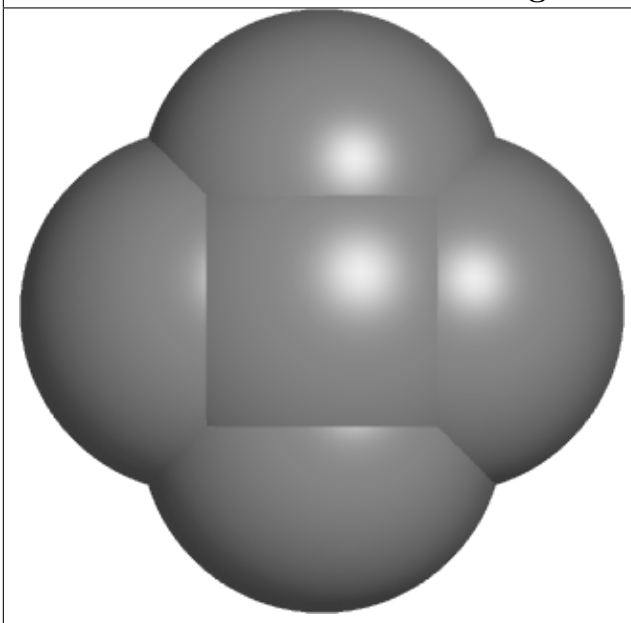
Torsions



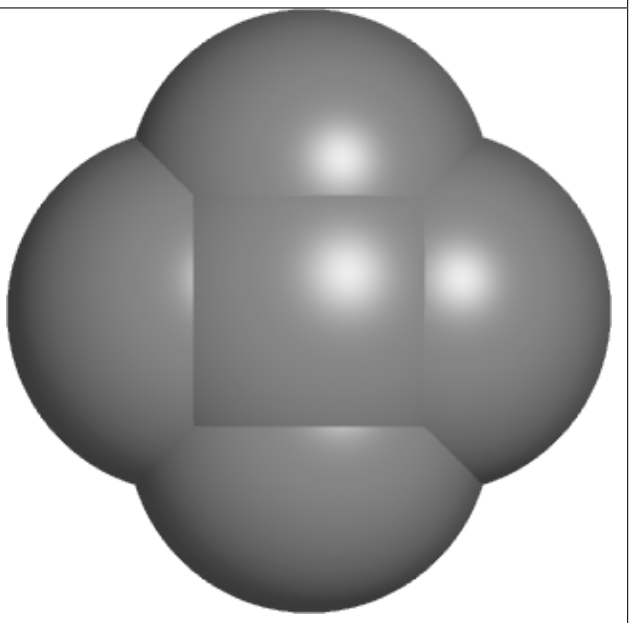
Rings

Ligand MET B 403	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

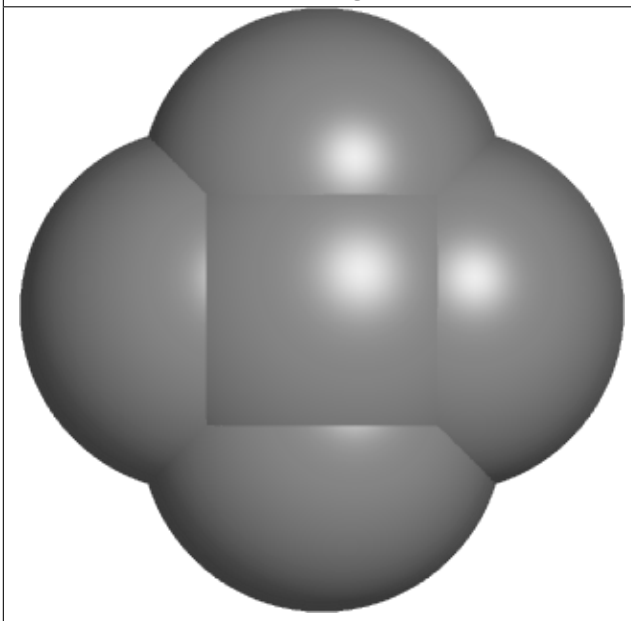
Ligand SO4 A 408



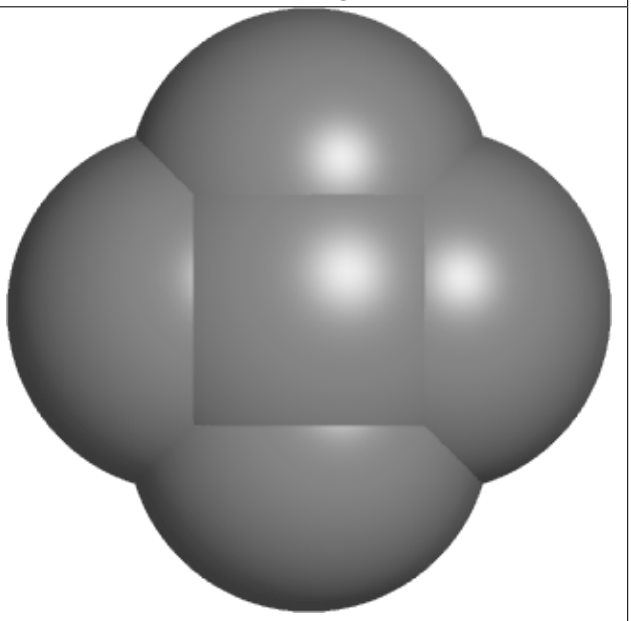
Bond lengths



Bond angles

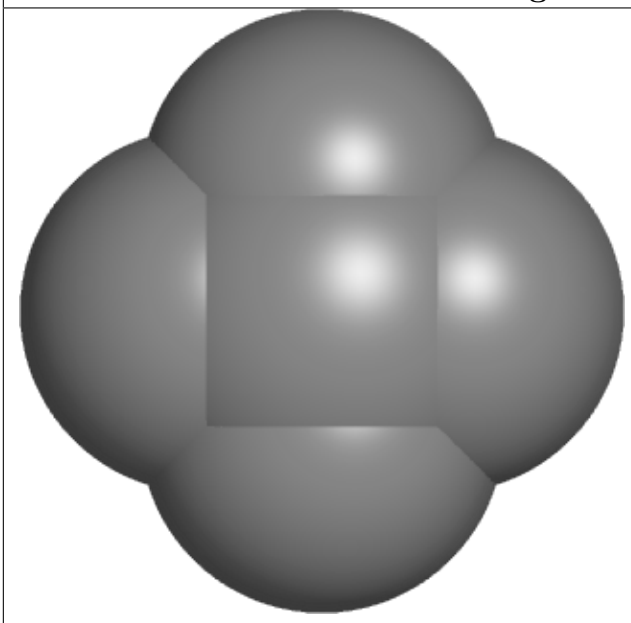


Torsions

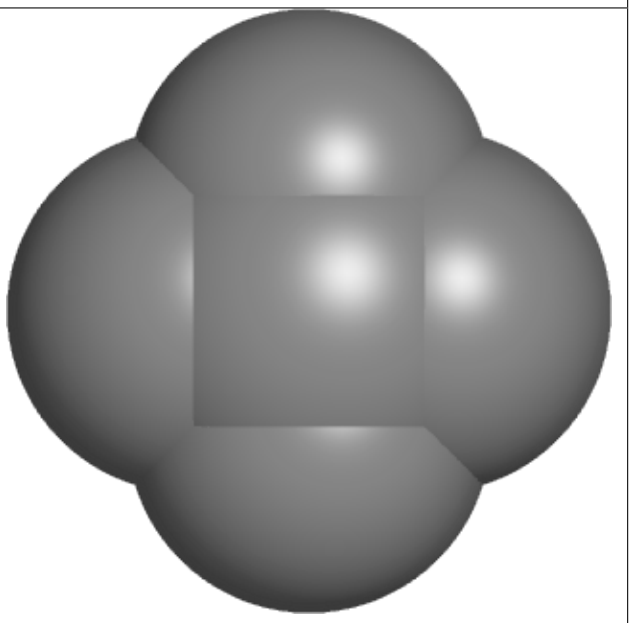


Rings

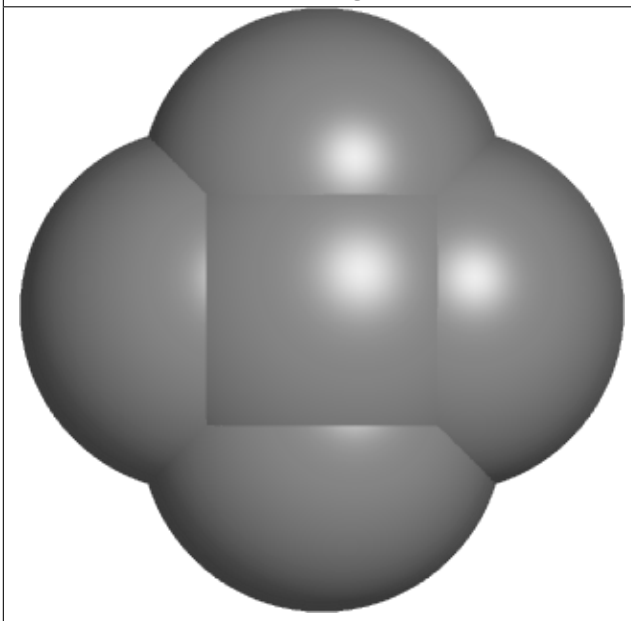
Ligand SO4 B 405



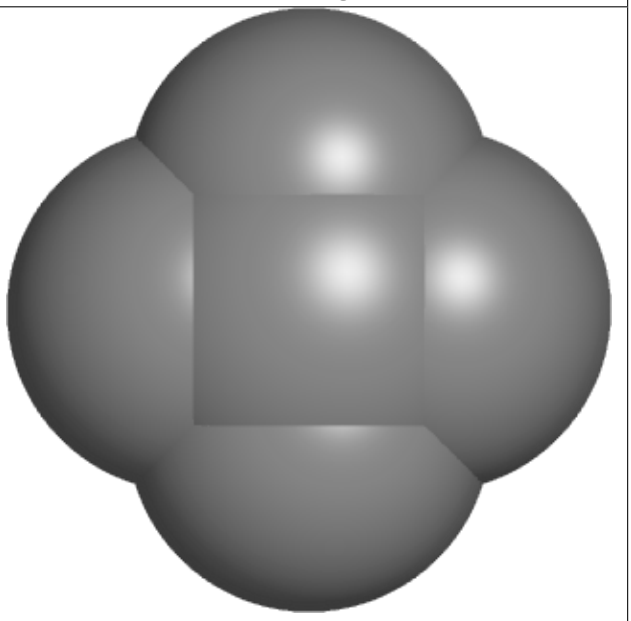
Bond lengths



Bond angles

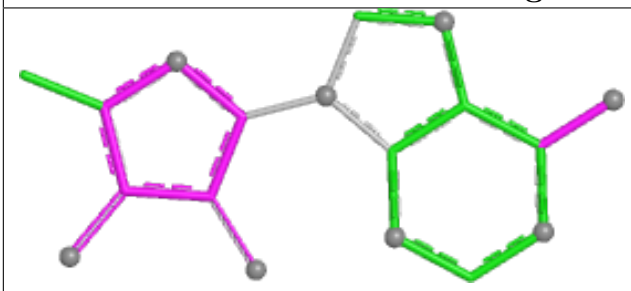


Torsions

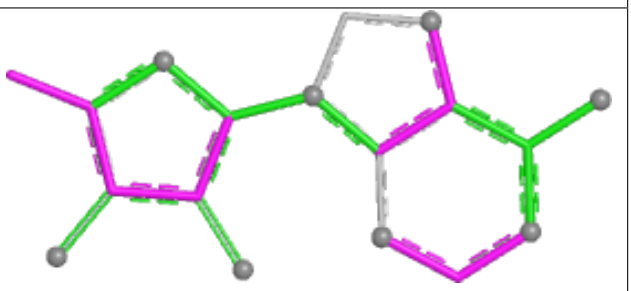


Rings

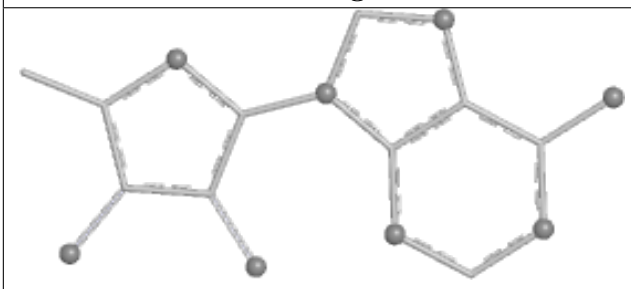
Ligand 5AD B 402



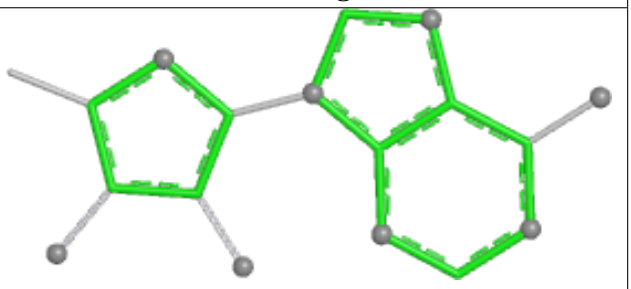
Bond lengths



Bond angles

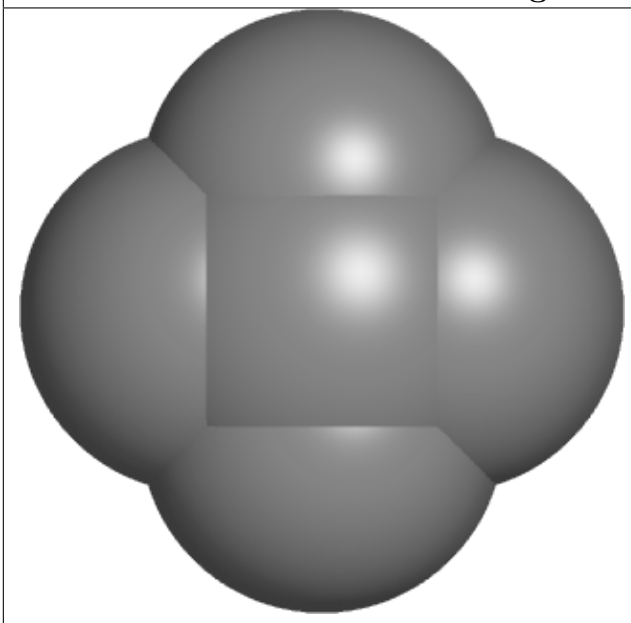


Torsions

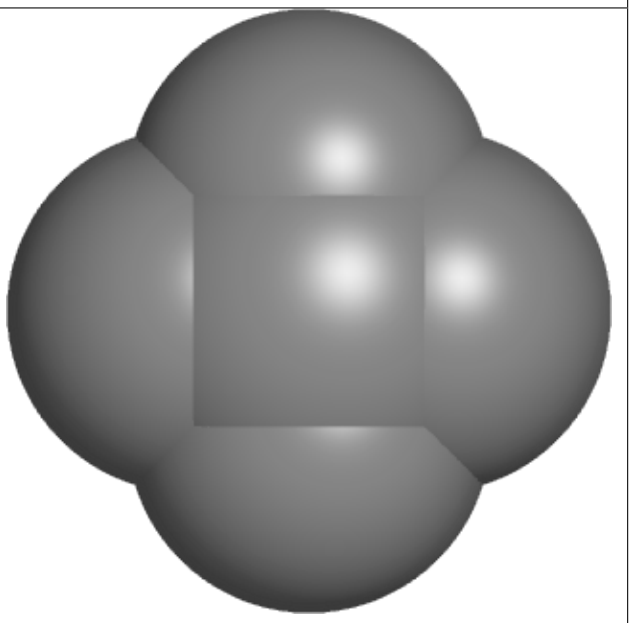


Rings

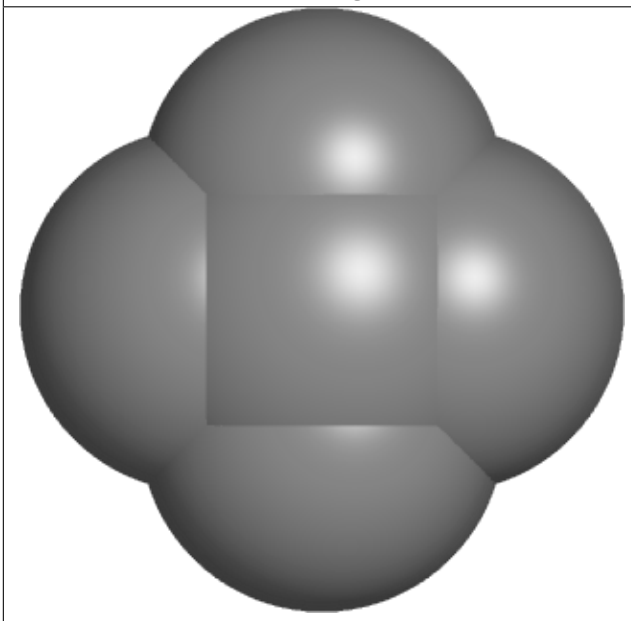
Ligand SO4 B 407



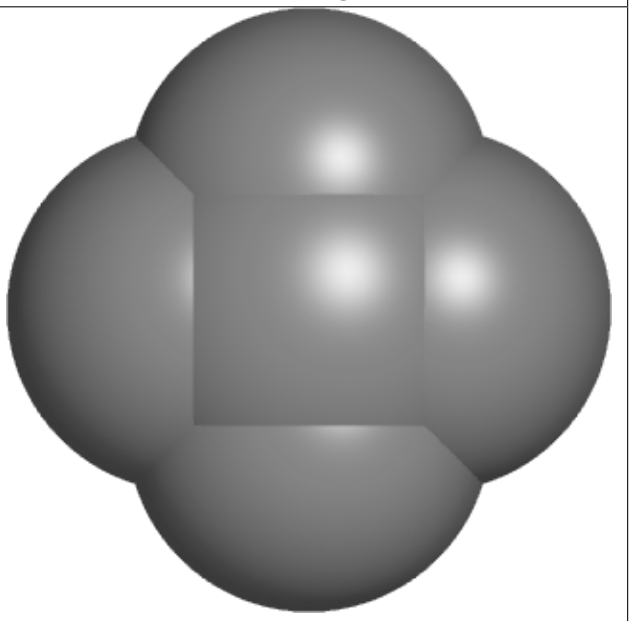
Bond lengths



Bond angles

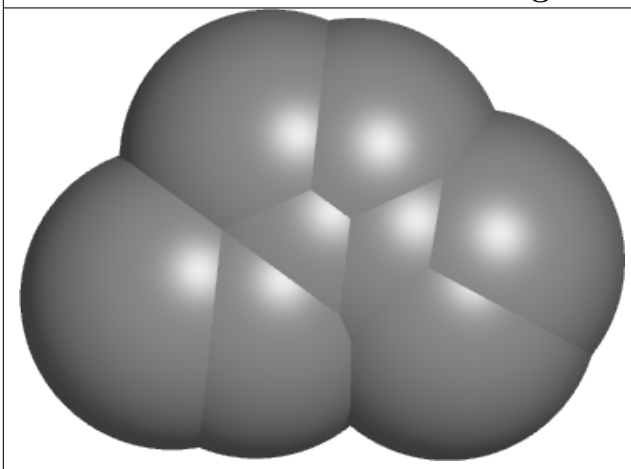


Torsions

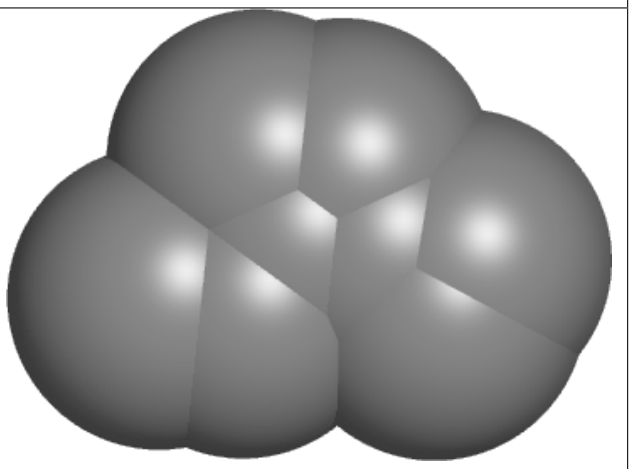


Rings

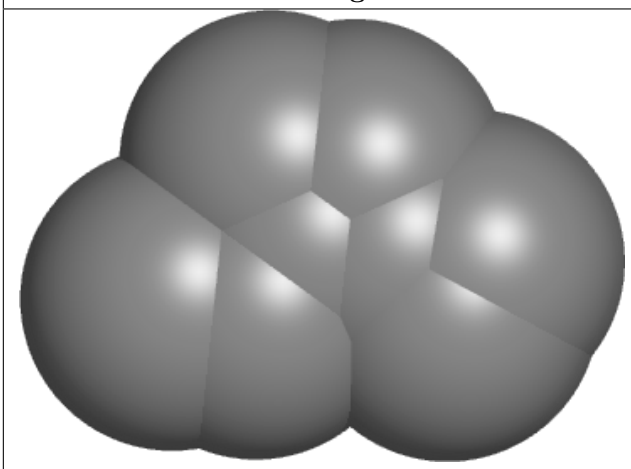
Ligand SF4 B 401



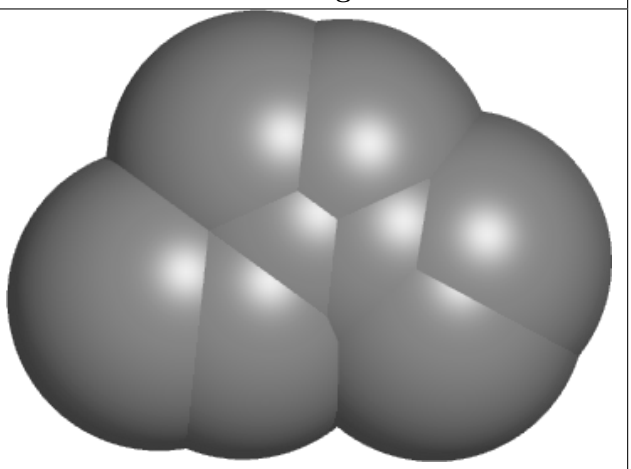
Bond lengths



Bond angles

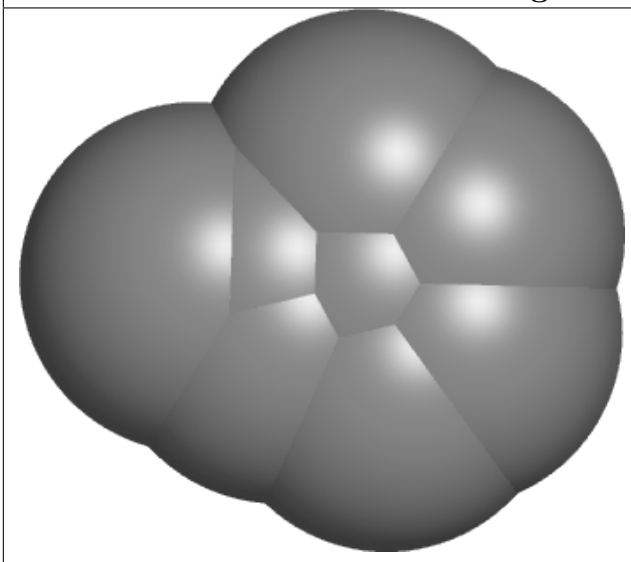


Torsions

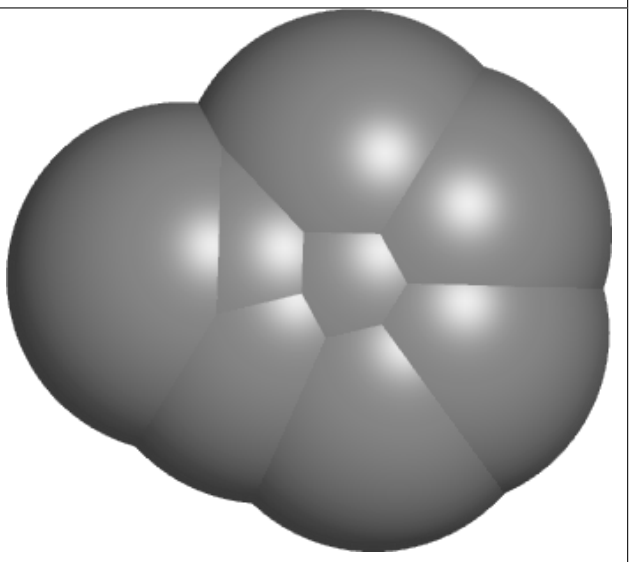


Rings

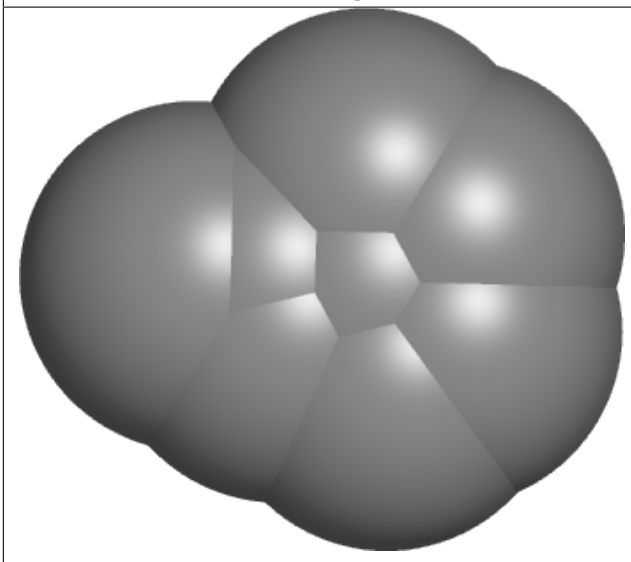
Ligand SF4 A 401



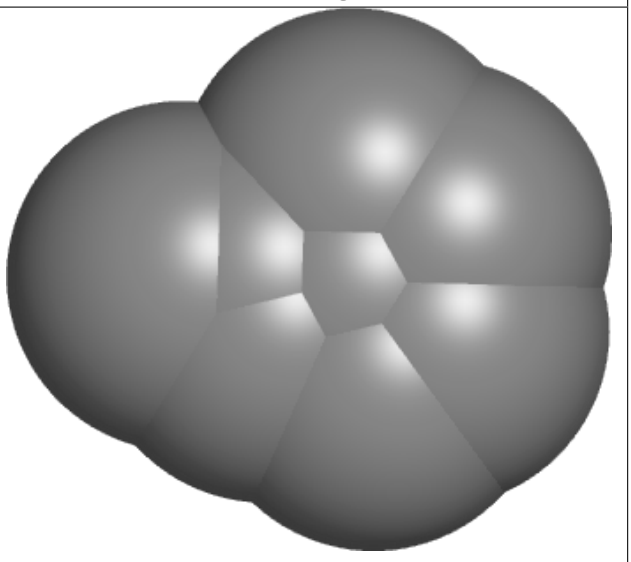
Bond lengths



Bond angles

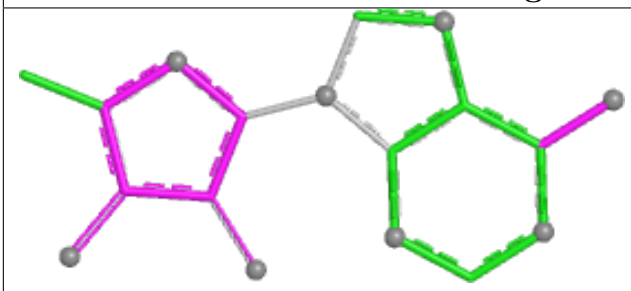


Torsions

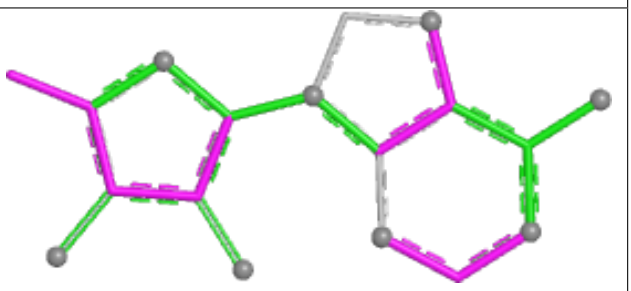


Rings

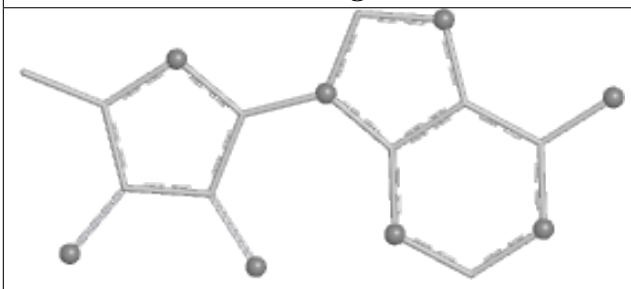
Ligand 5AD A 402



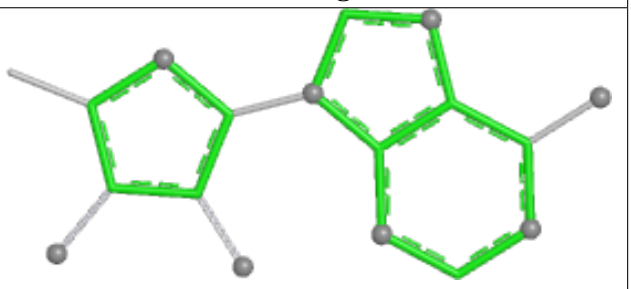
Bond lengths



Bond angles

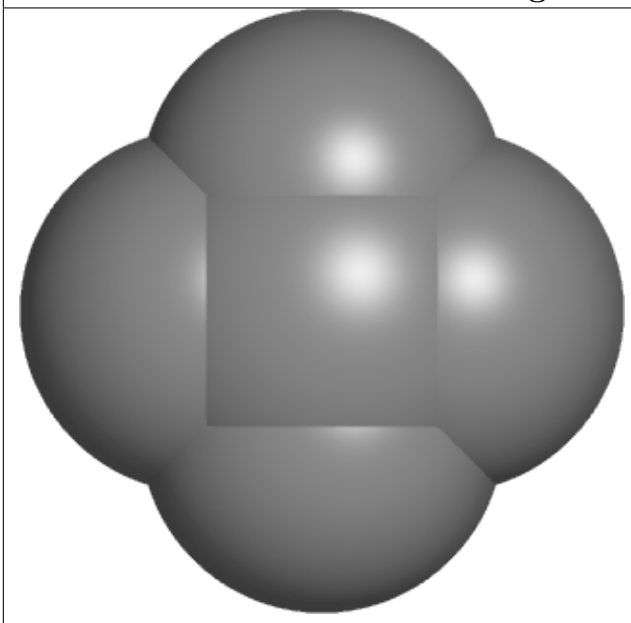


Torsions

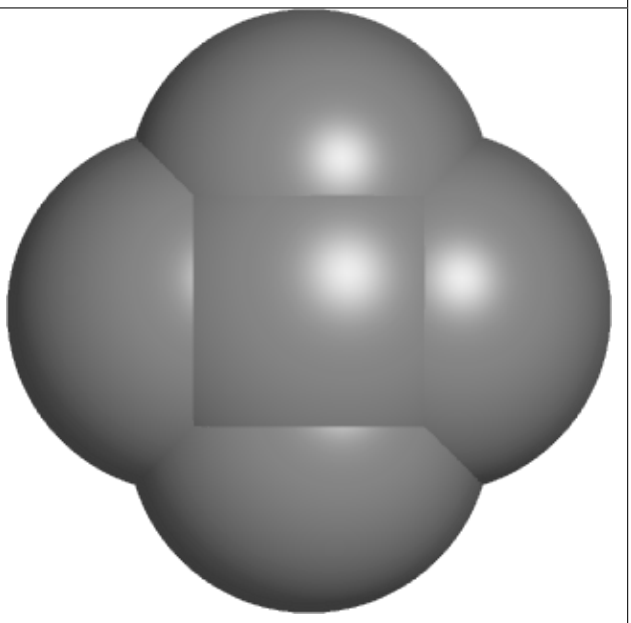


Rings

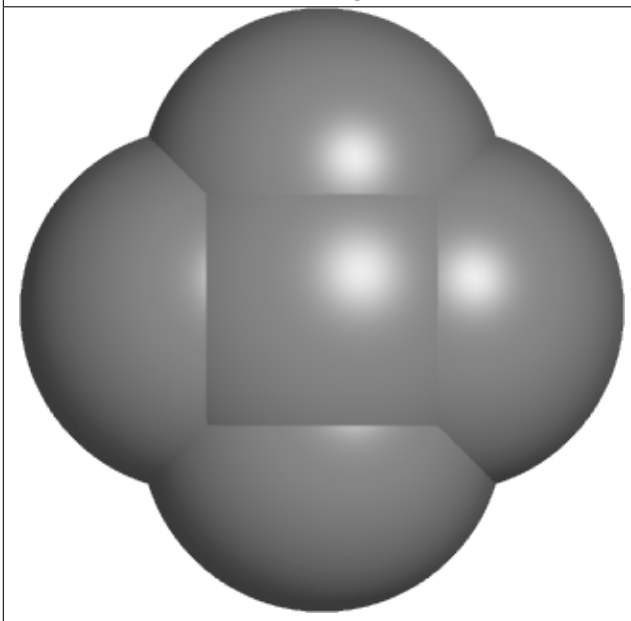
Ligand SO4 A 406



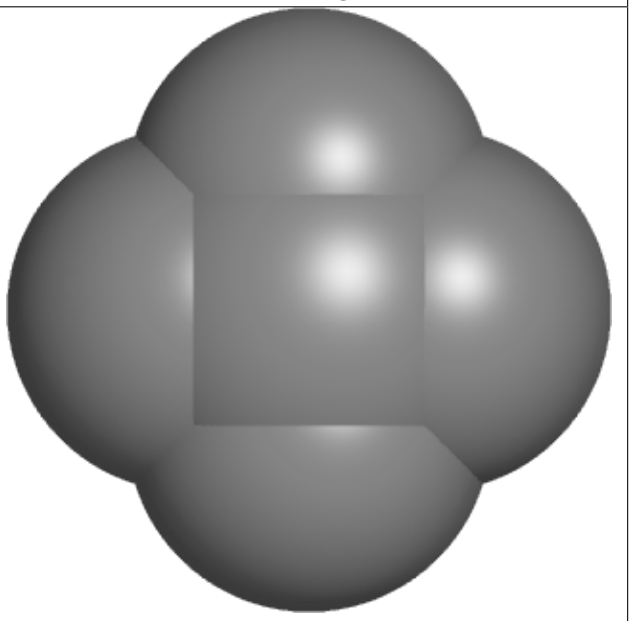
Bond lengths



Bond angles

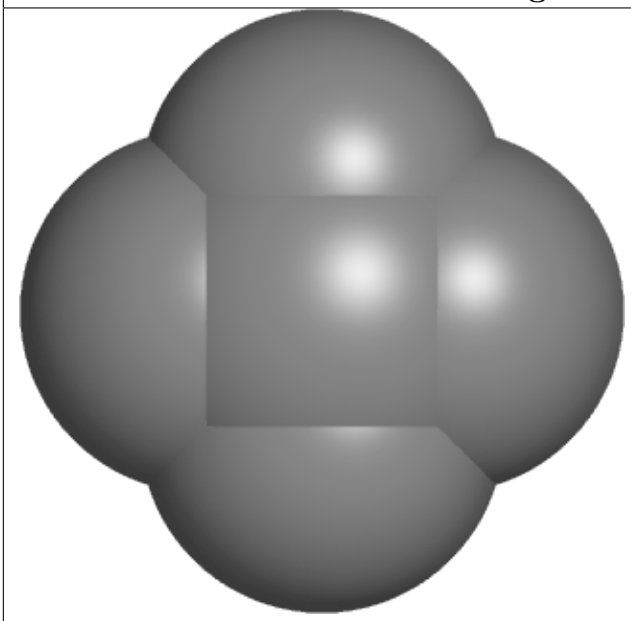


Torsions

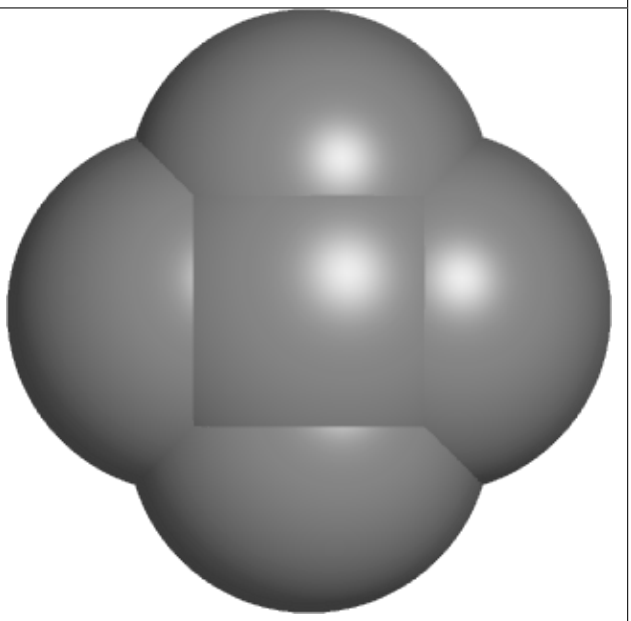


Rings

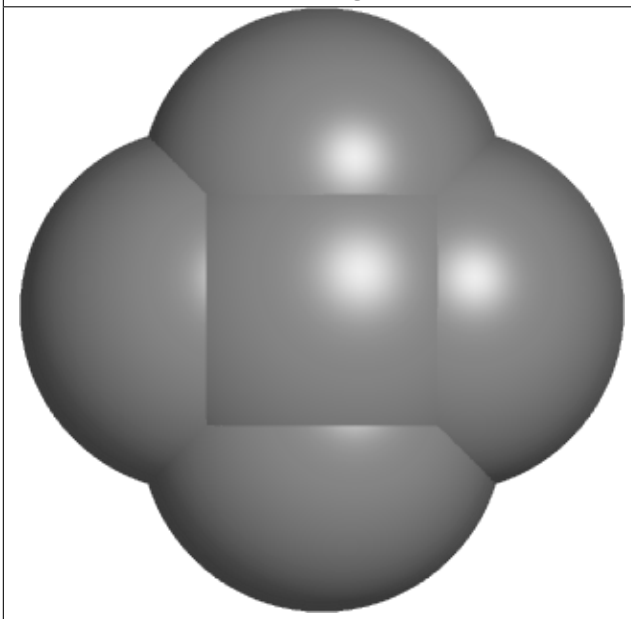
Ligand SO4 A 407



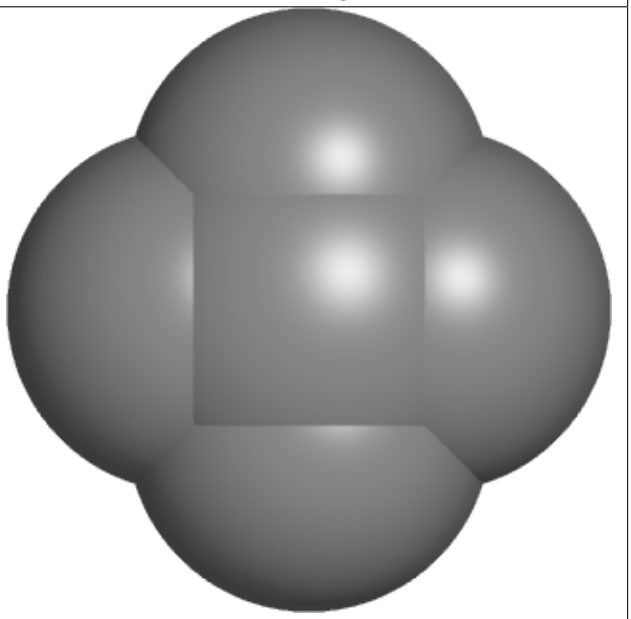
Bond lengths



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, 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	A	367/377 (97%)	1.56	85 (23%)	0	0	30, 42, 56, 79	1 (0%)
1	B	365/377 (96%)	1.57	97 (26%)	0	0	30, 42, 59, 74	1 (0%)
All	All	732/754 (97%)	1.57	182 (24%)	0	0	30, 42, 57, 79	2 (0%)

All (182) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	SER	9.8
1	B	366	ALA	6.3
1	B	330	ASP	6.3
1	A	2	GLY	5.9
1	B	267	GLY	5.7
1	B	100	LEU	5.5
1	A	84	TYR	5.2
1	A	333	ALA	4.9
1	A	330	ASP	4.8
1	B	273	VAL	4.7
1	B	10	ASP	4.5
1	A	125[A]	ILE	4.5
1	A	312	LEU	4.5
1	A	170	LEU	4.3
1	A	36	ARG	4.2
1	B	364	TRP	4.2
1	B	342	GLU	4.1
1	A	299	THR	3.9
1	A	8	ILE	3.9
1	B	99	LYS	3.9
1	B	302	ASN	3.9
1	A	100	LEU	3.9
1	A	334	VAL	3.8
1	A	367	LEU	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	184	VAL	3.7
1	B	28	VAL	3.7
1	B	8	ILE	3.6
1	B	94	PHE	3.6
1	B	147	ILE	3.5
1	A	153	SER	3.5
1	B	257	SER	3.5
1	A	116[A]	LEU	3.4
1	B	252	VAL	3.4
1	B	307	ASP	3.4
1	A	288[A]	LEU	3.4
1	B	12	ARG	3.4
1	B	321	SER	3.4
1	B	365	GLN	3.4
1	A	31	ILE	3.3
1	B	13	HIS	3.3
1	B	40	LEU	3.2
1	B	91	TYR	3.2
1	A	326	GLY	3.2
1	B	272	VAL	3.2
1	A	122	LEU	3.1
1	B	84	TYR	3.1
1	B	268	PHE	3.1
1	A	276	ASP	3.1
1	A	184	VAL	3.1
1	A	54	LEU	3.1
1	A	86	VAL	3.0
1	B	237	PHE	3.0
1	A	181	TYR	3.0
1	B	361	TYR	3.0
1	A	192	GLU	3.0
1	A	37	LEU	3.0
1	A	79	LEU	3.0
1	A	244	TYR	3.0
1	B	358	GLN	3.0
1	B	77	THR	2.9
1	B	231	TRP	2.9
1	B	230	ASP	2.9
1	B	116	LEU	2.9
1	B	333	ALA	2.9
1	B	129	THR	2.8
1	A	32	LEU	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	162	LEU	2.8
1	B	195	PRO	2.8
1	B	135	HIS	2.8
1	A	191	ALA	2.8
1	B	224	ALA	2.8
1	A	5	TYR	2.8
1	A	295	ILE	2.8
1	B	69	PHE	2.7
1	B	92	CYS	2.7
1	B	167	TYR	2.7
1	B	326	GLY	2.7
1	B	325	VAL	2.7
1	A	26	SER	2.7
1	A	96	LEU	2.7
1	B	208	ALA	2.7
1	A	321	SER	2.7
1	A	230	ASP	2.7
1	B	353	TYR	2.6
1	B	344	THR	2.6
1	B	290	MET	2.6
1	B	285	ALA	2.6
1	B	357	TYR	2.6
1	A	10	ASP	2.6
1	A	147	ILE	2.6
1	A	220[A]	VAL	2.6
1	A	251	ASP	2.6
1	A	18	ALA	2.5
1	B	312	LEU	2.5
1	A	272	VAL	2.5
1	A	139	SER	2.5
1	A	226	LEU	2.5
1	B	178	LEU	2.5
1	B	332	GLU	2.5
1	B	41	ASP	2.5
1	B	181	TYR	2.5
1	B	320	GLY	2.5
1	B	173	ALA	2.5
1	B	360	VAL	2.5
1	B	266	GLY	2.4
1	B	125	ILE	2.4
1	A	75	LEU	2.4
1	A	176	ASP	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	319	ALA	2.4
1	A	90	VAL	2.4
1	B	162	LEU	2.4
1	A	151	TYR	2.4
1	A	363	ASP	2.4
1	A	289	PHE	2.4
1	B	153	SER	2.4
1	B	329	SER	2.4
1	A	325	VAL	2.4
1	B	22	LYS	2.4
1	A	50	ALA	2.4
1	A	157	ILE	2.4
1	B	43	LEU	2.3
1	B	280	VAL	2.3
1	A	227	GLY	2.3
1	B	265	LEU	2.3
1	B	286	PHE	2.3
1	B	339	ILE	2.3
1	B	348	VAL	2.3
1	B	79	LEU	2.3
1	B	363	ASP	2.3
1	A	21	ALA	2.2
1	A	364	TRP	2.2
1	B	295	ILE	2.2
1	A	357	TYR	2.2
1	A	97	LYS	2.2
1	B	64	LEU	2.2
1	B	170	LEU	2.2
1	A	353	TYR	2.2
1	A	249	PHE	2.2
1	B	23	VAL	2.2
1	A	101	GLU	2.2
1	B	151	TYR	2.2
1	A	268	PHE	2.2
1	A	237	PHE	2.2
1	B	126	LEU	2.1
1	B	245	LEU	2.1
1	A	160	TYR	2.1
1	A	152	PHE	2.1
1	A	85	CYS	2.1
1	A	64	LEU	2.1
1	A	308	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	68	HIS	2.1
1	A	310	VAL	2.1
1	B	243	ASP	2.1
1	A	338	GLN	2.1
1	B	331	GLN	2.1
1	B	98	ASN	2.1
1	B	282	TYR	2.1
1	B	356	GLY	2.1
1	A	110	VAL	2.1
1	A	238	THR	2.1
1	B	75	LEU	2.1
1	B	240	LEU	2.1
1	B	5	TYR	2.1
1	A	20	PHE	2.1
1	B	7	VAL	2.1
1	B	127	ILE	2.1
1	B	171	ILE	2.1
1	A	65	THR	2.0
1	A	341	ASP	2.0
1	B	244	TYR	2.0
1	A	274	VAL	2.0
1	B	16	PHE	2.0
1	A	346	ALA	2.0
1	B	71	ARG	2.0
1	B	225	LEU	2.0
1	B	179	THR	2.0
1	A	313	GLY	2.0
1	B	249	PHE	2.0
1	A	40	LEU	2.0
1	B	45	LEU	2.0
1	A	144	CYS	2.0
1	A	55	GLU	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

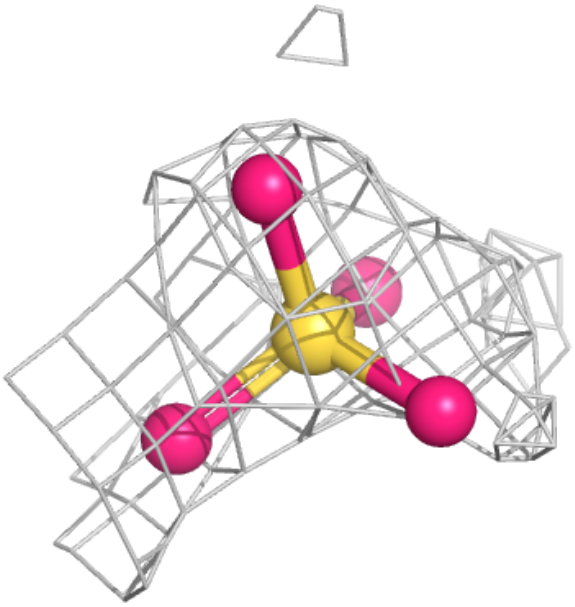
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

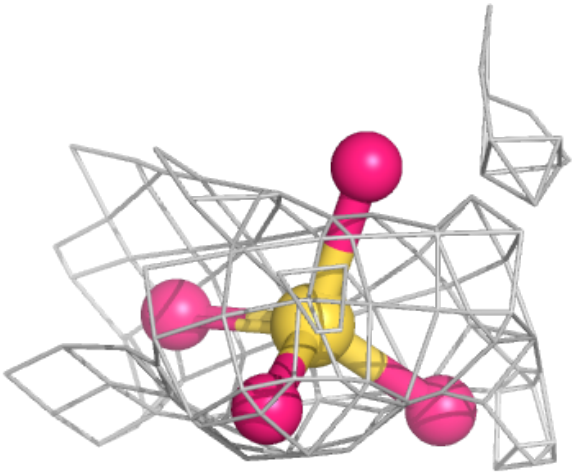
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	SO4	A	407	5/5	0.59	0.36	101,101,101,101	0
5	GOL	A	405	6/6	0.63	0.25	60,61,61,62	0
5	GOL	B	404	6/6	0.65	0.26	58,59,60,60	0
6	SO4	B	407	5/5	0.71	0.19	88,89,89,89	0
2	SF4	B	401	8/8	0.74	0.11	32,33,36,38	0
3	5AD	A	402	18/18	0.75	0.19	40,42,47,47	0
5	GOL	A	404	6/6	0.76	0.23	65,65,66,66	0
2	SF4	A	401	8/8	0.78	0.13	31,35,39,39	0
3	5AD	B	402	18/18	0.80	0.17	37,38,42,42	0
4	MET	A	403	9/9	0.83	0.24	54,58,60,63	0
4	MET	B	403	9/9	0.88	0.22	45,49,51,62	0
6	SO4	B	406	5/5	0.92	0.18	62,63,63,64	0
6	SO4	A	408	5/5	0.93	0.19	66,66,66,67	0
6	SO4	B	405	5/5	0.94	0.19	45,45,46,46	0
6	SO4	A	406	5/5	0.97	0.13	45,45,45,46	0

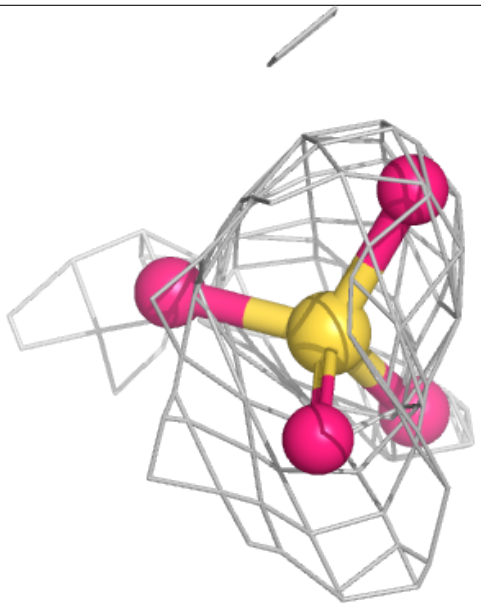
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around SO4 A 407:

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

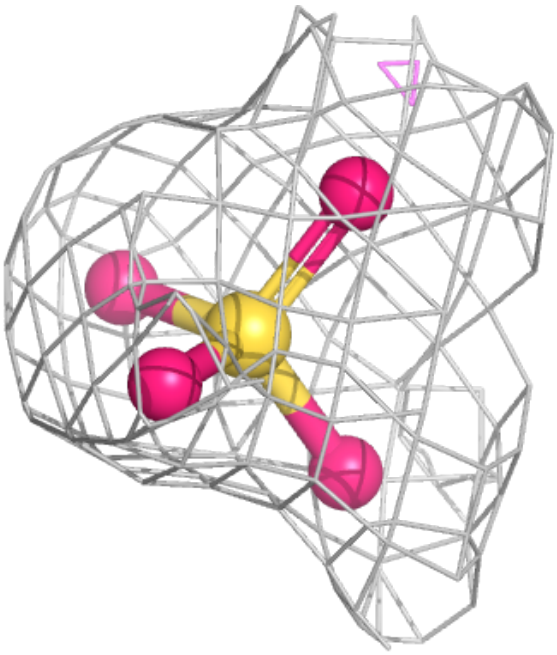


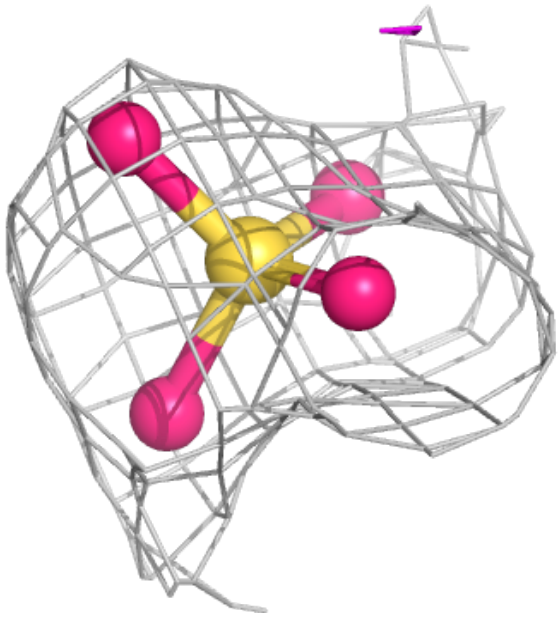


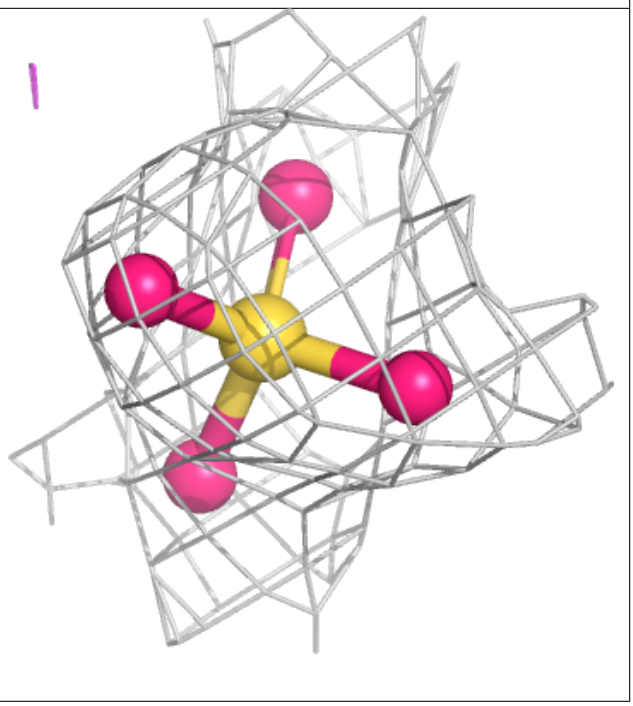


Electron density around SO4 B 407:

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

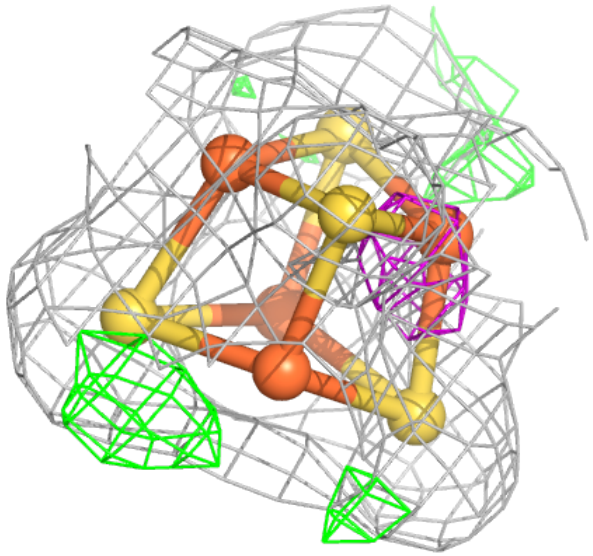


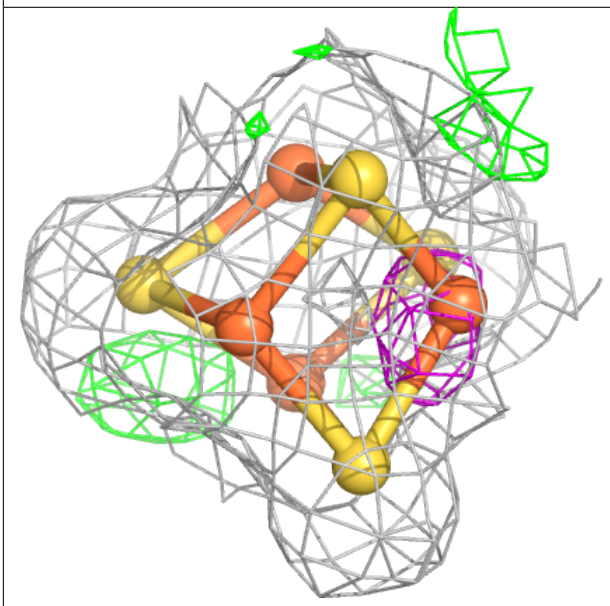


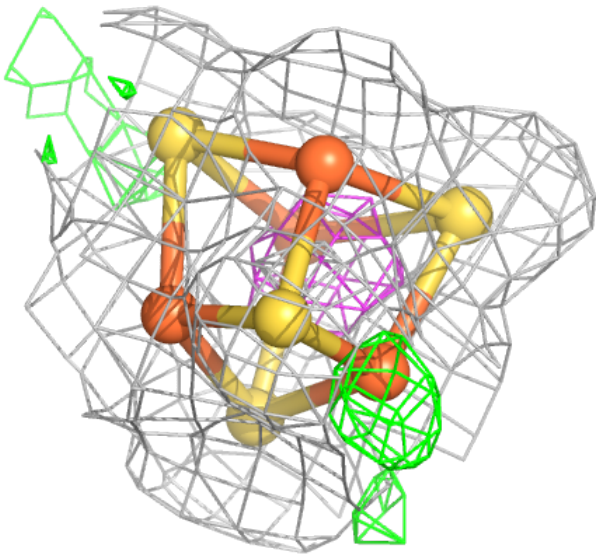


Electron density around SF4 B 401:

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

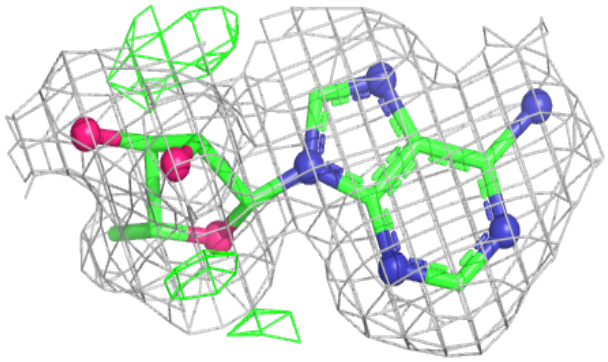


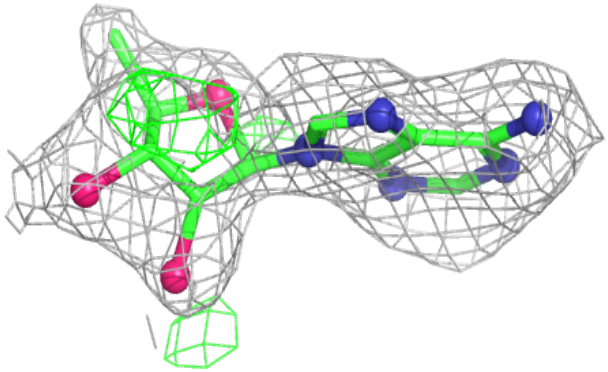


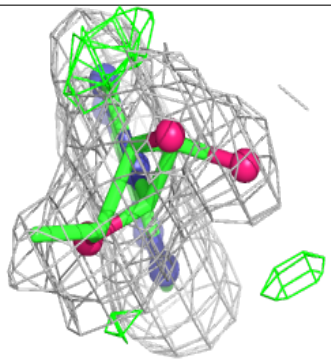


Electron density around 5AD A 402:

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

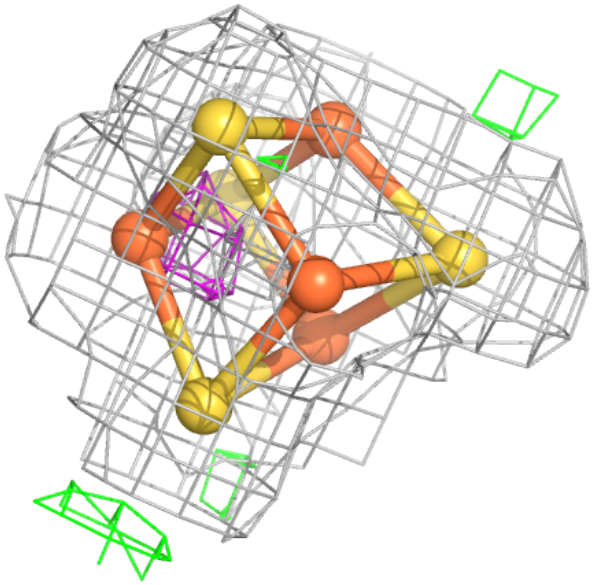


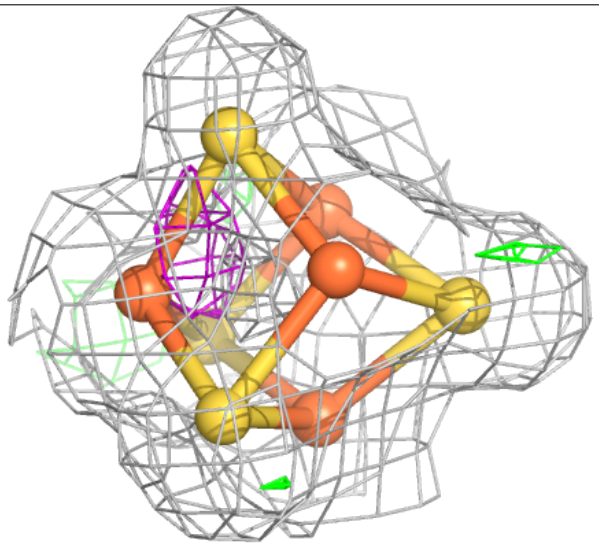


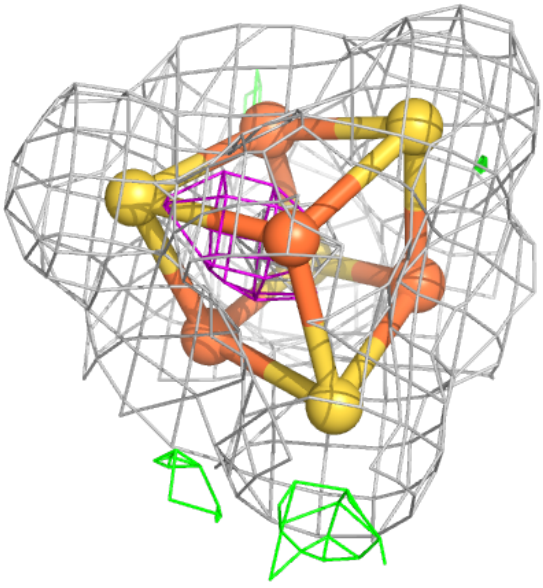


Electron density around SF4 A 401:

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

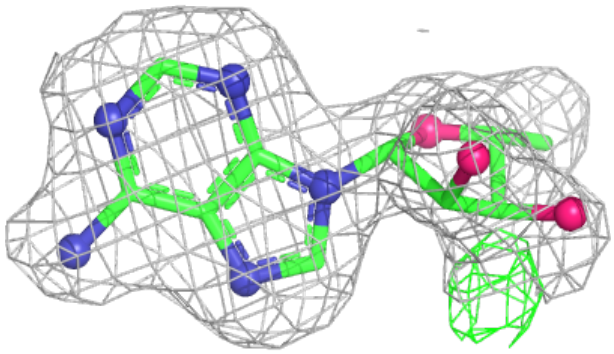


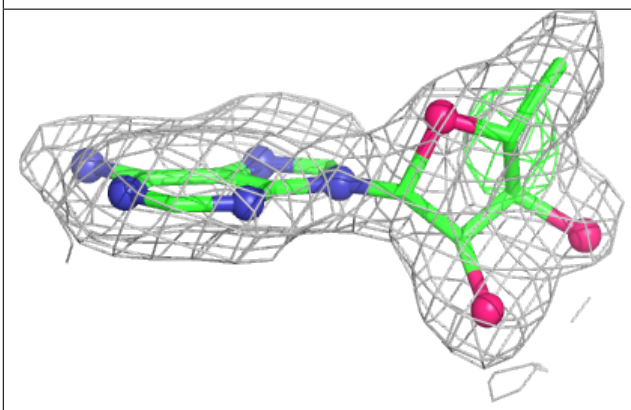


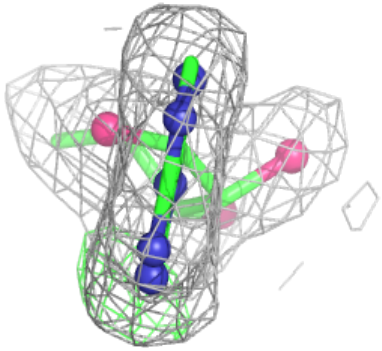


Electron density around 5AD B 402:

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

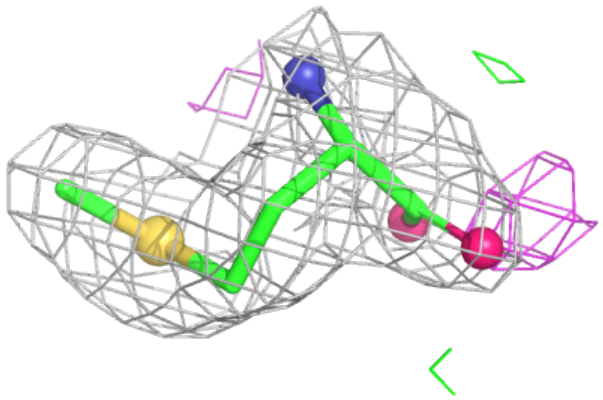


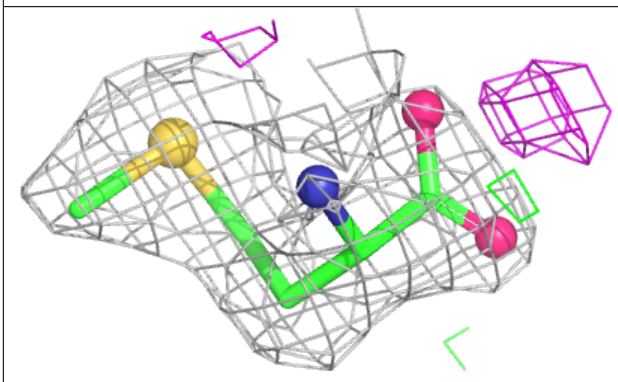


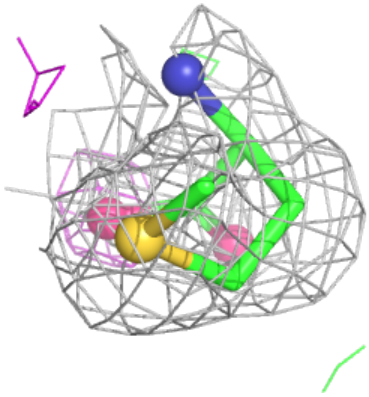


Electron density around MET A 403:

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

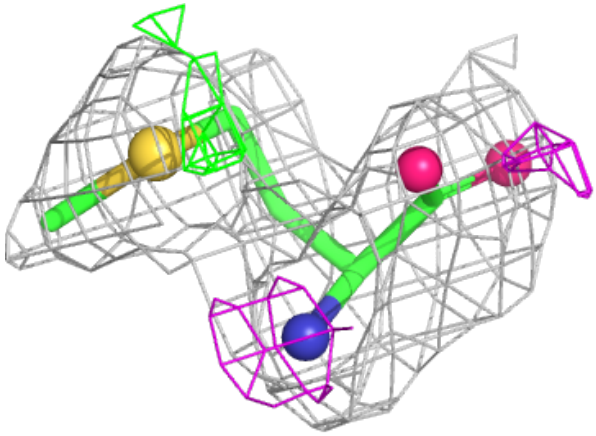


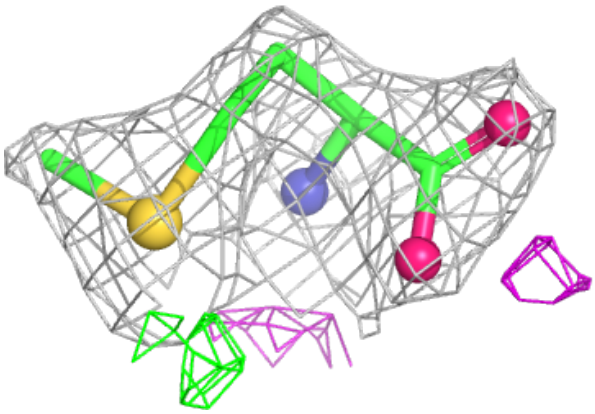


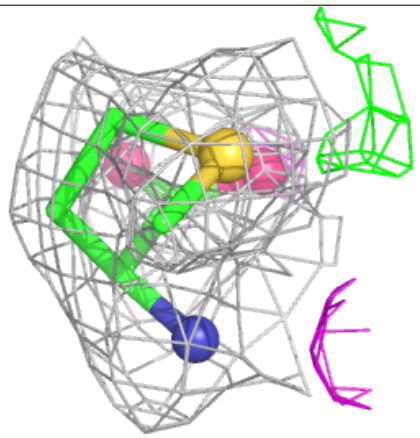


Electron density around MET B 403:

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

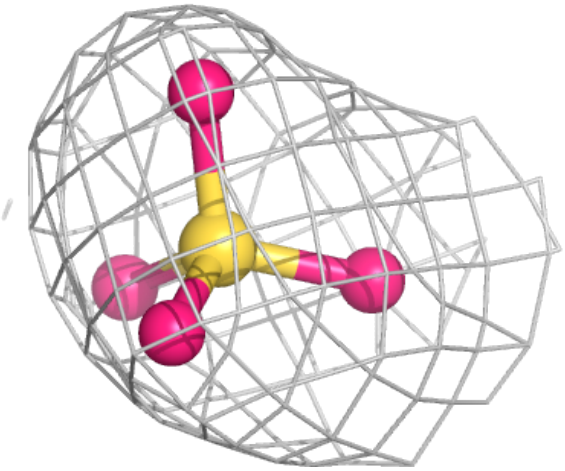


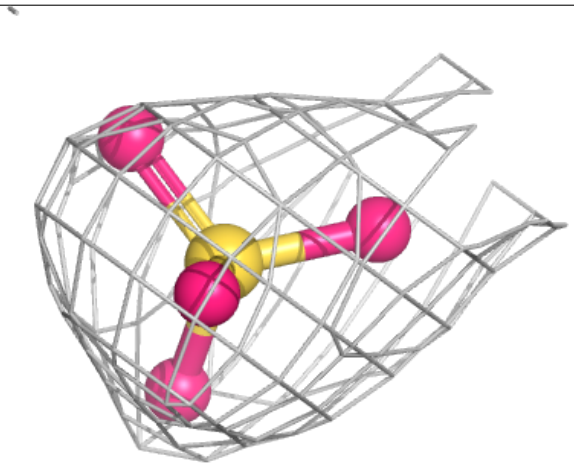


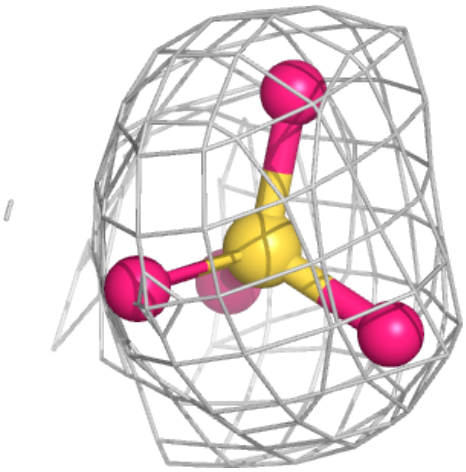


Electron density around SO4 B 406:

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

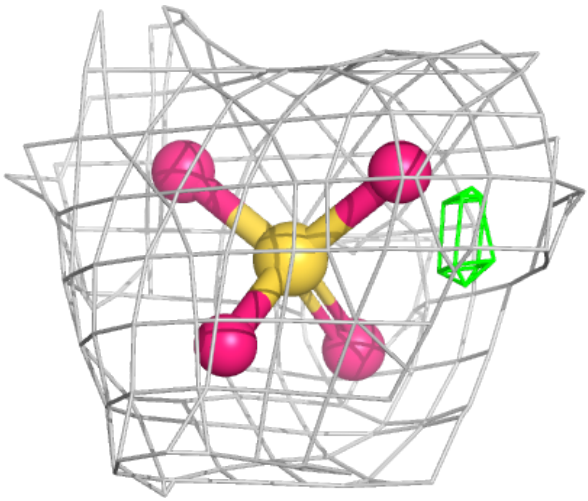


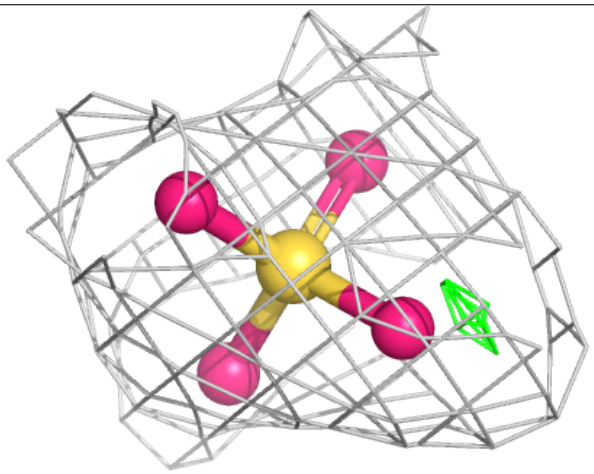


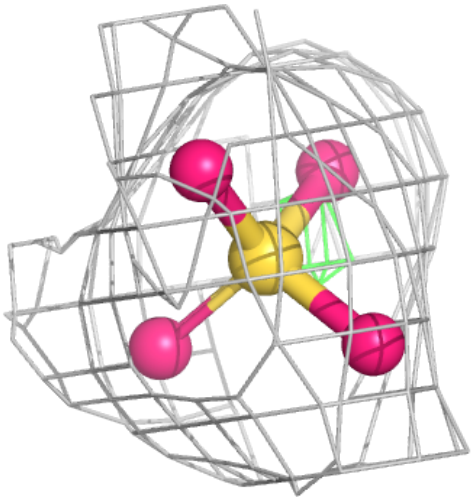


Electron density around SO4 A 408:

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

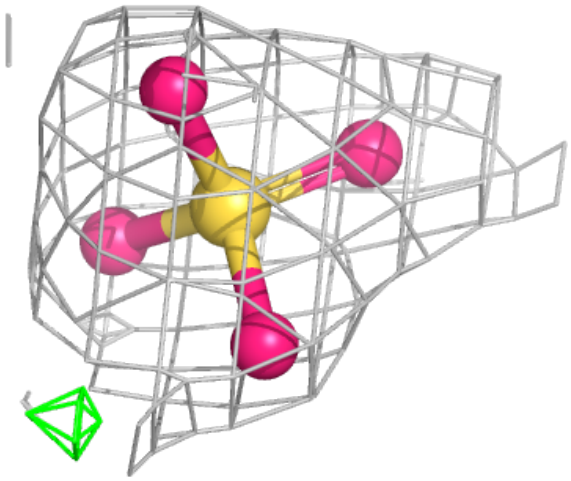


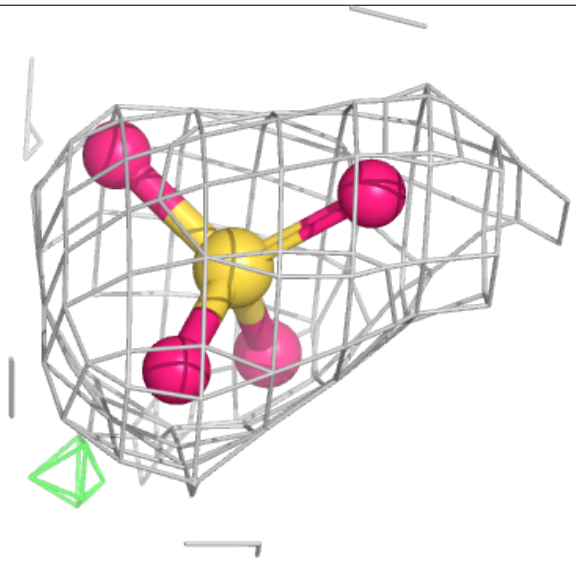


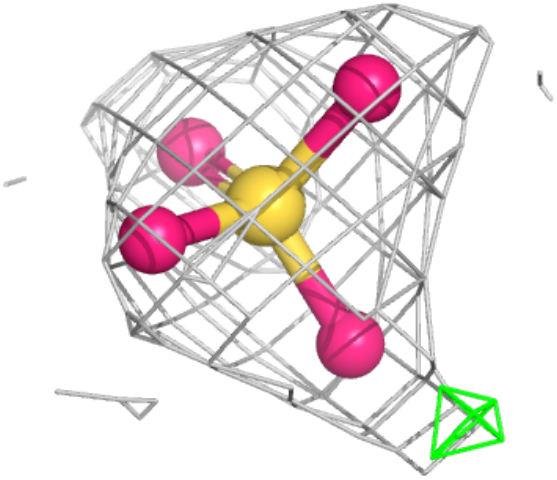


Electron density around SO4 B 405:

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

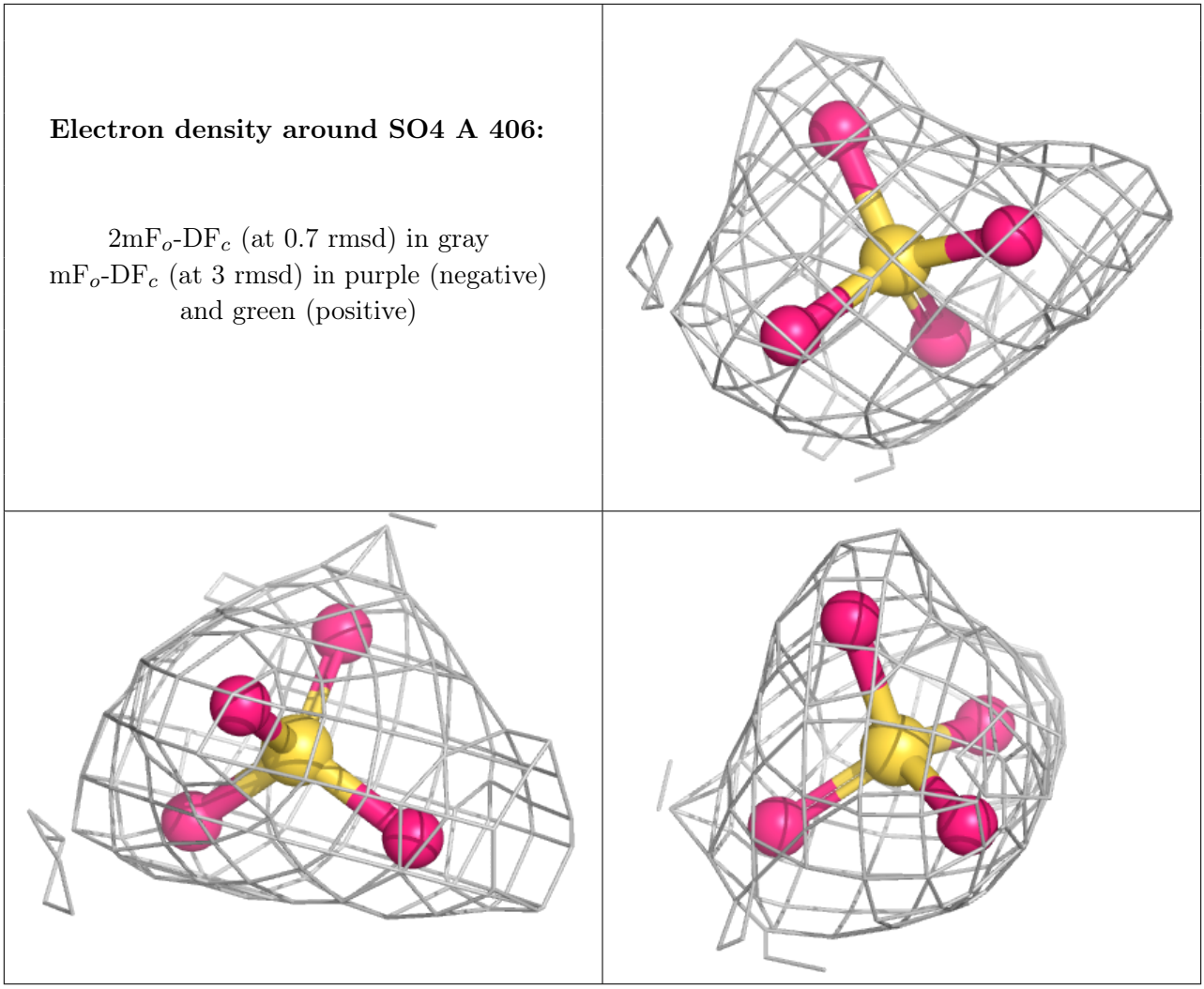






Electron density around SO4 A 406:

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



6.5 Other polymers

There are no such residues in this entry.