



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 9, 2020 – 05:18 PM GMT

PDB ID : 6T7J
Title : As-isolated Ni-free crystal structure of carbon monoxide dehydrogenase from Thermococcus sp. AM4 produced without CooC maturase
Authors : Dobbek, H.; Jeoung, J.-H.
Deposited on : 2019-10-22
Resolution : 2.43 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.14.6
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) : 2.14.6

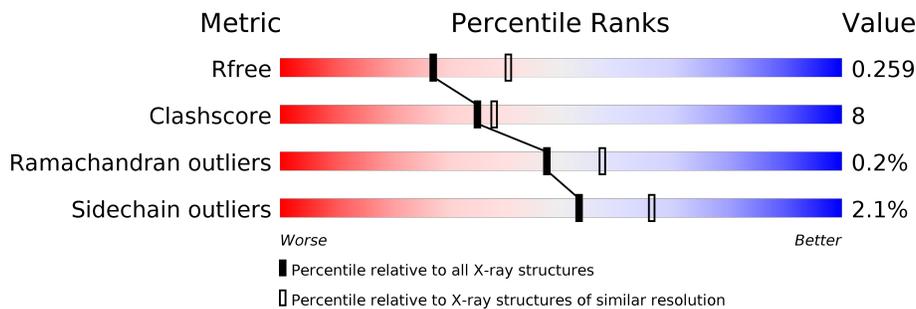
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.43 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	1564 (2.46-2.42)
Clashscore	141614	1631 (2.46-2.42)
Ramachandran outliers	138981	1617 (2.46-2.42)
Sidechain outliers	138945	1617 (2.46-2.42)

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 on 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain	
1	A	632	84%	16%
1	B	632	78%	22%
1	D	632	76%	22%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	FES	A	702	-	-	X	-

2 Entry composition [i](#)

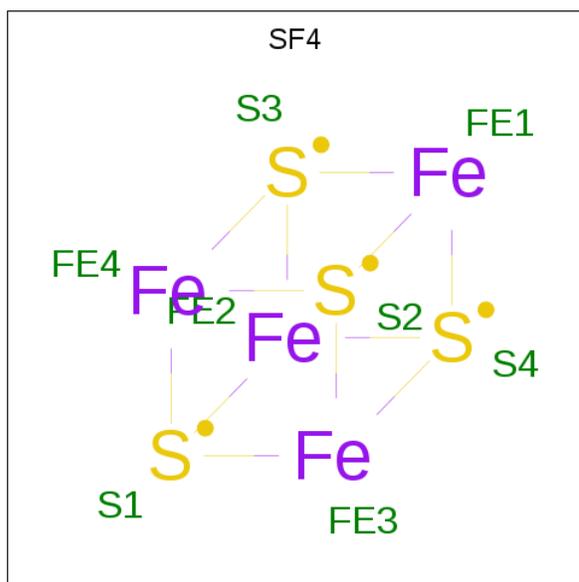
There are 9 unique types of molecules in this entry. The entry contains 14591 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Carbon monoxide dehydrogenase.

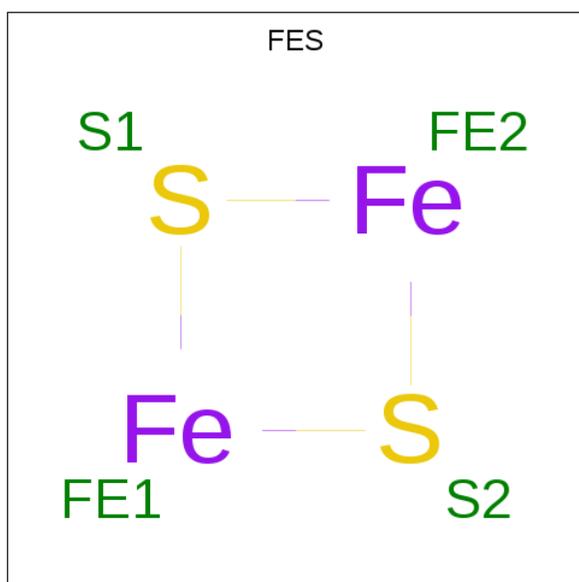
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	632	Total 4768	C 3035	N 819	O 877	S 37	0	0	0
1	A	631	Total 4763	C 3032	N 818	O 876	S 37	0	1	0
1	D	629	Total 4749	C 3024	N 816	O 872	S 37	0	1	0

- Molecule 2 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



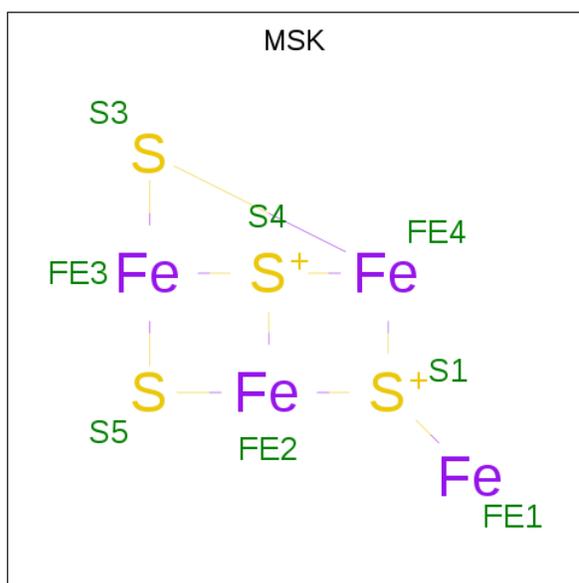
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
2	B	1	Total 8	Fe 4	S 4	0	0
2	A	1	Total 8	Fe 4	S 4	0	0
2	D	1	Total 8	Fe 4	S 4	0	0

- Molecule 3 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe₂S₂).



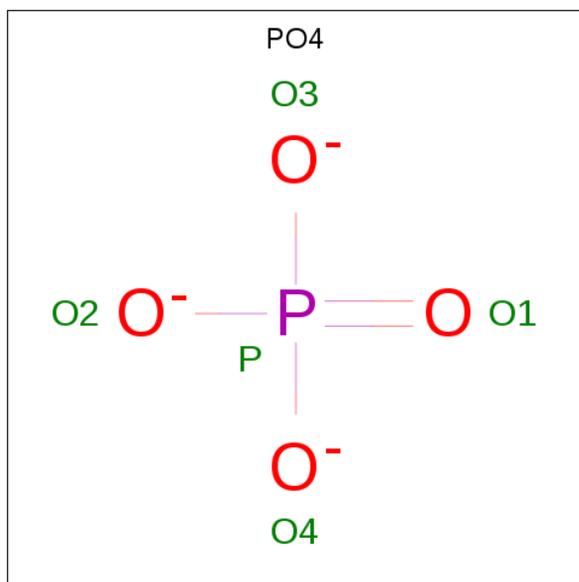
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	Fe	S	0	0
			4	2	2		
3	A	1	Total	Fe	S	0	0
			4	2	2		
3	D	1	Total	Fe	S	0	0
			4	2	2		

- Molecule 4 is Broken Fe₄S₄ cluster (three-letter code: MSK) (formula: Fe₄S₄) (labeled as "Ligand of Interest" by author).



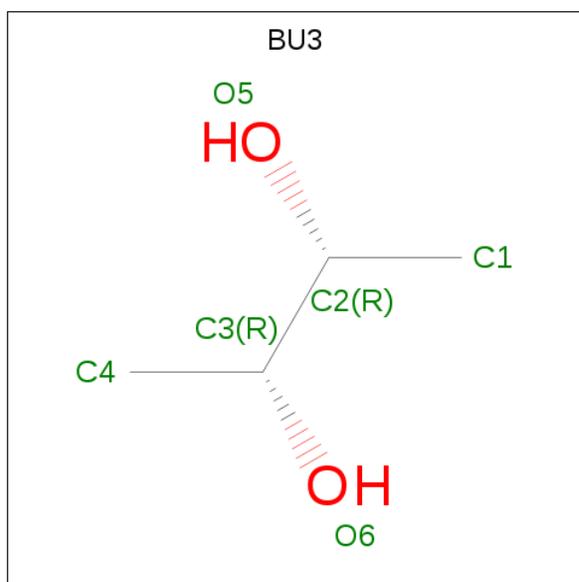
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	Fe	S	0	0
			8	4	4		
4	A	1	Total	Fe	S	0	0
			8	4	4		
4	D	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



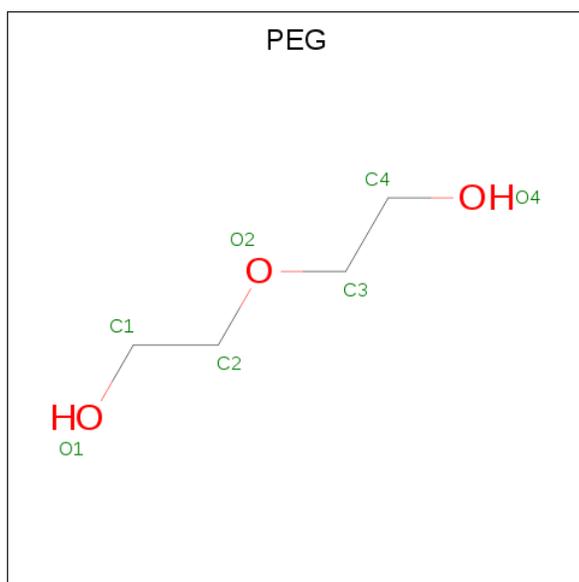
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	P	0	0
			5	4	1		
5	A	1	Total	O	P	0	0
			5	4	1		
5	D	1	Total	O	P	0	0
			5	4	1		
5	D	1	Total	O	P	0	0
			5	4	1		

- Molecule 6 is (R,R)-2,3-BUTANEDIOL (three-letter code: BU3) (formula: C₄H₁₀O₂).



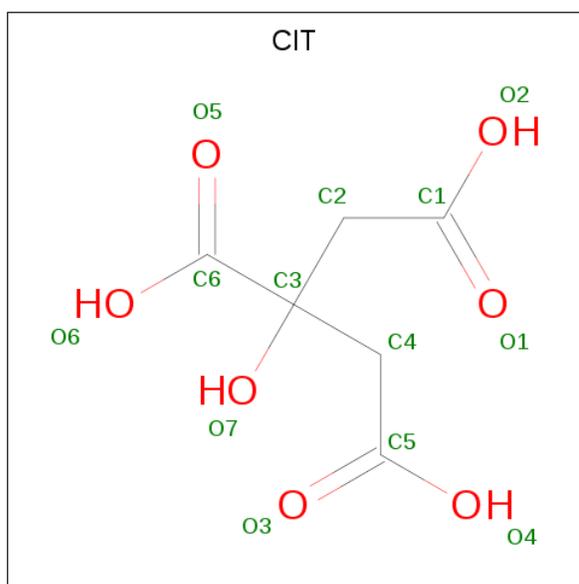
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			6	4	2		
6	B	1	Total	C	O	0	0
			6	4	2		
6	B	1	Total	C	O	0	0
			6	4	2		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 13 6 7	0	0

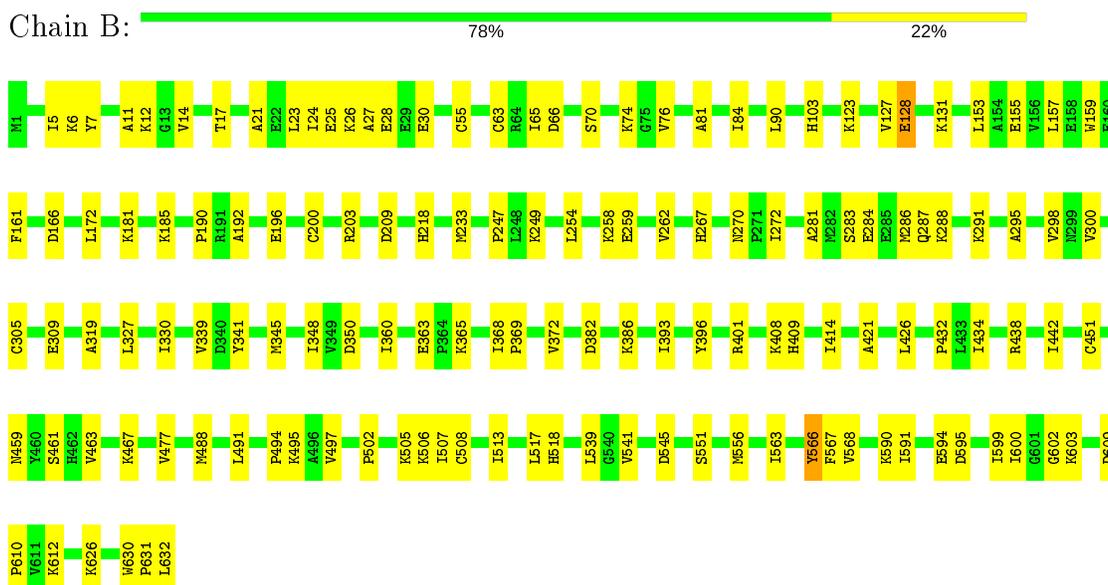
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	75	Total O 75 75	0	0
9	A	63	Total O 63 63	0	0
9	D	55	Total O 55 55	0	0

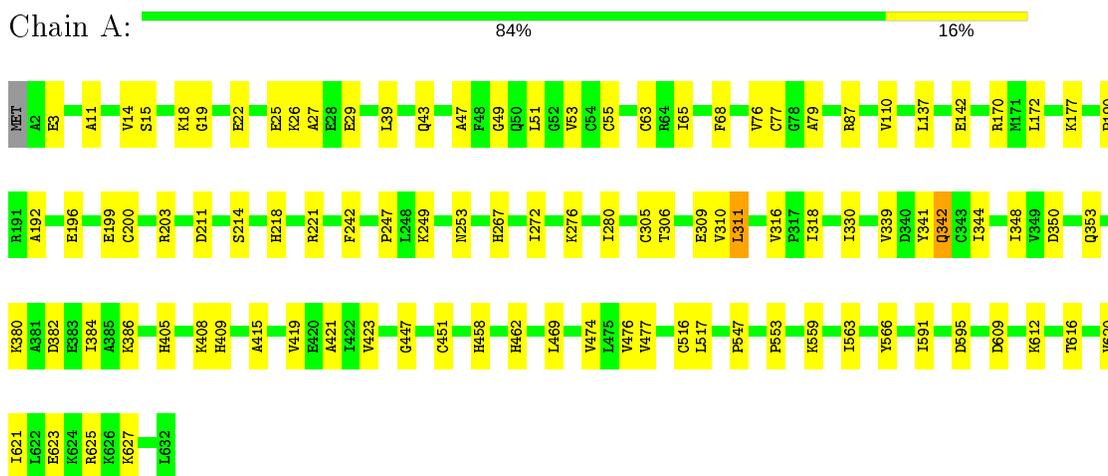
3 Residue-property plots [i](#)

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. 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. 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: Carbon monoxide dehydrogenase

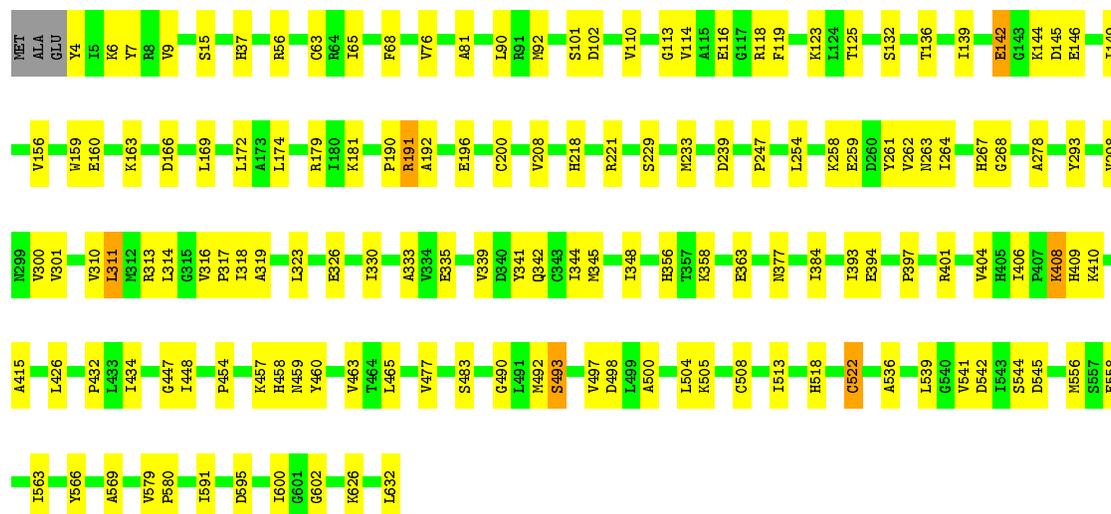


- Molecule 1: Carbon monoxide dehydrogenase



- Molecule 1: Carbon monoxide dehydrogenase





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	268.14Å 64.06Å 100.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.32 – 2.43 47.32 – 2.34	Depositor EDS
% Data completeness (in resolution range)	99.3 (47.32-2.43) 99.3 (47.32-2.34)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.13 (at 2.34Å)	Xtrriage
Refinement program	PHENIX 1.16-3549	Depositor
R, R_{free}	(Not available) , (Not available) 0.205 , 0.259	Depositor DCC
R_{free} test set	2100 reflections (2.84%)	wwPDB-VP
Wilson B-factor (Å ²)	49.8	Xtrriage
Anisotropy	0.057	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 43.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	14591	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: BU3, SF4, MSK, FES, CIT, PEG, PO4

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	0/4857	0.45	0/6577
1	B	0.31	1/4858 (0.0%)	0.46	0/6576
1	D	0.59	4/4843 (0.1%)	0.46	1/6558 (0.0%)
All	All	0.42	5/14558 (0.0%)	0.45	1/19711 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	408	LYS	CE-NZ	34.59	2.35	1.49
1	D	142	GLU	CD-OE2	-5.81	1.19	1.25
1	B	128	GLU	CD-OE2	-5.46	1.19	1.25
1	D	142	GLU	CG-CD	-5.08	1.44	1.51
1	D	408	LYS	CD-CE	-5.04	1.38	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	408	LYS	CD-CE-NZ	5.27	123.81	111.70

There are no chirality outliers.

There are no planarity outliers.

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	4763	0	4875	63	1
1	B	4768	0	4881	87	2
1	D	4749	0	4864	89	5
2	A	8	0	0	0	0
2	B	8	0	0	0	0
2	D	8	0	0	0	0
3	A	4	0	0	0	2
3	B	4	0	0	0	0
3	D	4	0	0	0	0
4	A	8	0	0	1	0
4	B	8	0	0	1	0
4	D	8	0	0	0	0
5	A	5	0	0	0	0
5	B	5	0	0	1	0
5	D	10	0	0	1	0
6	B	18	0	30	2	0
7	B	7	0	10	1	0
8	A	13	0	5	2	0
9	A	63	0	0	2	0
9	B	75	0	0	3	0
9	D	55	0	0	2	0
All	All	14591	0	14665	237	7

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 8.

The worst 5 of 237 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:D:408:LYS:CE	1:D:408:LYS:NZ	2.35	0.90
1:B:259:GLU:OE2	1:A:408:LYS:HD2	1.72	0.89
1:D:408:LYS:HD3	1:D:408:LYS:O	1.76	0.84
1:B:284:GLU:OE2	1:B:288:LYS:HD2	1.79	0.83
1:A:63:CYS:SG	1:A:76:VAL:HG23	2.19	0.82

The worst 5 of 7 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	Interatomic distance (Å)	Clash overlap (Å)
1:D:408:LYS:NZ	1:D:408:LYS:NZ[2_555]	1.76	0.44
1:B:74:LYS:NZ	1:B:350:ASP:OD2[2_565]	1.78	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:56:ARG:NH1	3:A:702:FES:S2[3_546]	2.00	0.20
1:D:408:LYS:CE	1:D:408:LYS:CE[2_555]	2.07	0.13
1:B:128:GLU:OE2	1:D:6:LYS:O[1_554]	2.17	0.03

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	630/632 (100%)	604 (96%)	25 (4%)	1 (0%)	47 57
1	B	629/632 (100%)	605 (96%)	24 (4%)	0	100 100
1	D	628/632 (99%)	605 (96%)	21 (3%)	2 (0%)	41 49
All	All	1887/1896 (100%)	1814 (96%)	70 (4%)	3 (0%)	47 57

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	342	GLN
1	D	342	GLN
1	D	208	VAL

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	505/505 (100%)	501 (99%)	4 (1%)	81 88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	505/505 (100%)	492 (97%)	13 (3%)	46	58
1	D	504/505 (100%)	489 (97%)	15 (3%)	41	53
All	All	1514/1515 (100%)	1482 (98%)	32 (2%)	53	66

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	177	LYS
1	D	68	PHE
1	D	522[A]	CYS
1	A	341	TYR
1	D	101	SER

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

Mol	Chain	Res	Type
1	B	218	HIS

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](#)

18 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond

length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	PO4	B	704	-	4,4,4	0.87	0	6,6,6	0.47	0
3	FES	D	702	1,3	0,4,4	0.00	-	-	-	-
5	PO4	D	704	-	4,4,4	0.89	0	6,6,6	0.44	0
4	MSK	A	703	1	0,10,10	0.00	-	-	-	-
2	SF4	B	701	1	0,12,12	0.00	-	-	-	-
5	PO4	A	704	-	4,4,4	0.90	0	6,6,6	0.44	0
6	BU3	B	707	-	4,5,5	0.35	0	6,6,6	0.25	0
3	FES	A	702	1,3	0,4,4	0.00	-	-	-	-
2	SF4	A	701	1	0,12,12	0.00	-	-	-	-
8	CIT	A	705	-	3,12,12	1.27	0	3,17,17	1.70	0
5	PO4	D	705	-	4,4,4	0.90	0	6,6,6	0.43	0
3	FES	B	702	1	0,4,4	0.00	-	-	-	-
6	BU3	B	706	-	4,5,5	0.32	0	6,6,6	0.27	0
4	MSK	D	703	1	0,10,10	0.00	-	-	-	-
4	MSK	B	703	1	0,10,10	0.00	-	-	-	-
6	BU3	B	705	-	4,5,5	0.28	0	6,6,6	0.28	0
7	PEG	B	708	-	6,6,6	0.45	0	5,5,5	0.25	0
2	SF4	D	701	1	0,12,12	0.00	-	-	-	-

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
3	FES	D	702	1,3	-	-	0/1/1/1
4	MSK	A	703	1	-	-	0/3/3/3
6	BU3	B	707	-	-	4/4/4/4	-
3	FES	A	702	1,3	-	-	0/1/1/1
2	SF4	A	701	1	-	-	0/6/5/5
8	CIT	A	705	-	-	2/6/16/16	-
3	FES	B	702	1	-	-	0/1/1/1
6	BU3	B	706	-	-	0/4/4/4	-
4	MSK	B	703	1	-	-	0/3/3/3
4	MSK	D	703	1	-	-	0/3/3/3
2	SF4	B	701	1	-	-	0/6/5/5
6	BU3	B	705	-	-	0/4/4/4	-
7	PEG	B	708	-	-	3/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SF4	D	701	1	-	-	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

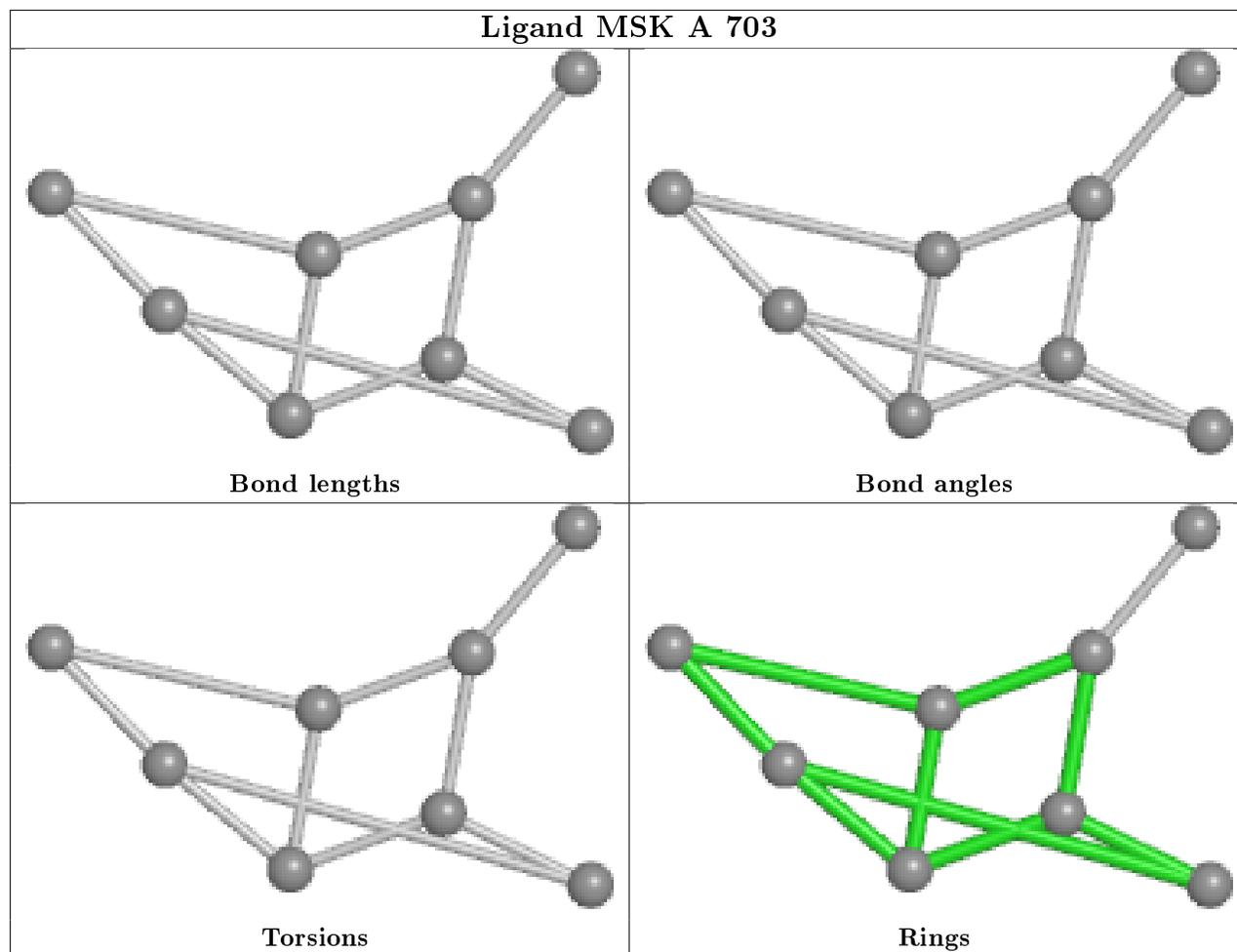
Mol	Chain	Res	Type	Atoms
6	B	707	BU3	O5-C2-C3-O6
6	B	707	BU3	C1-C2-C3-O6
6	B	707	BU3	O5-C2-C3-C4
6	B	707	BU3	C1-C2-C3-C4
7	B	708	PEG	O2-C3-C4-O4

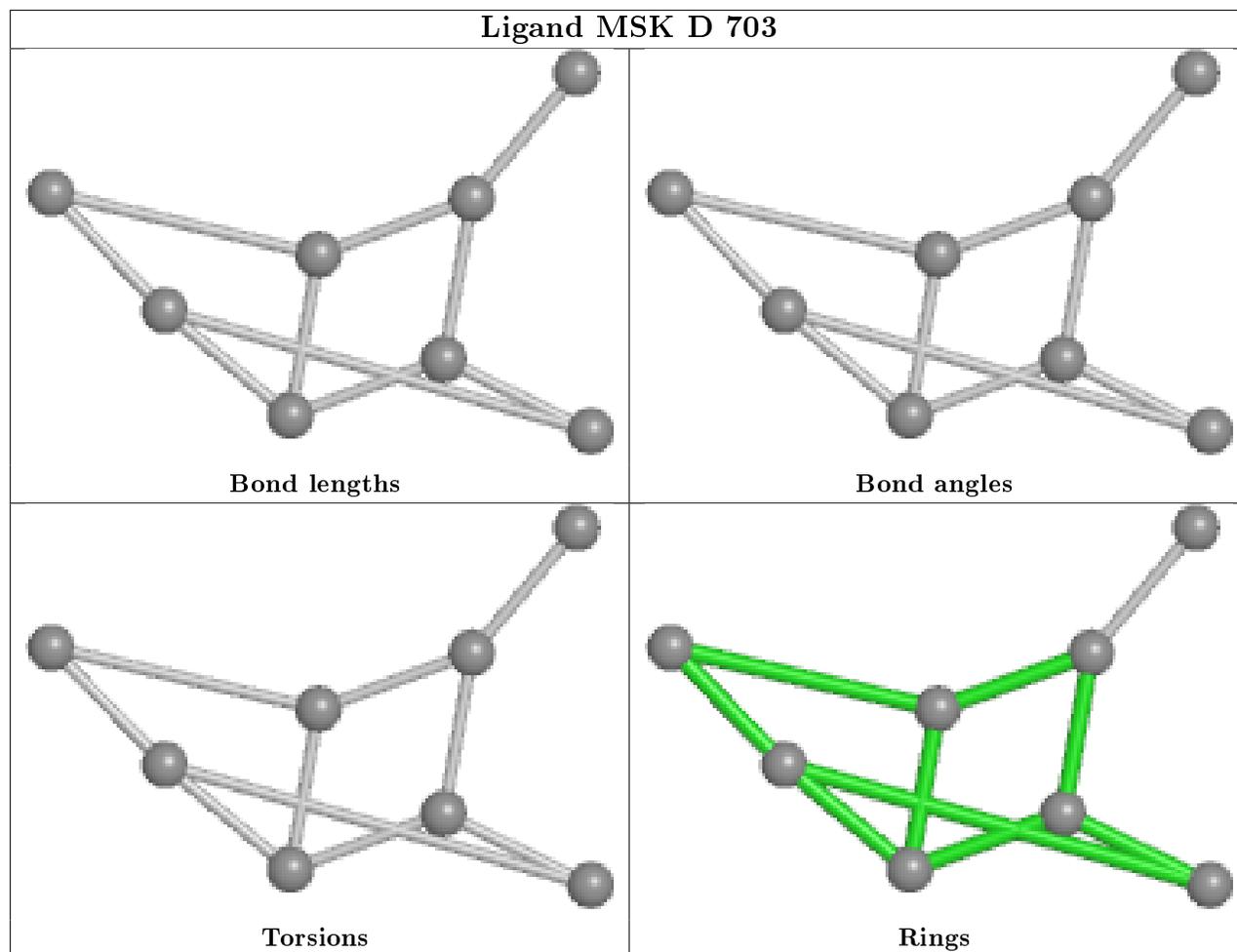
There are no ring outliers.

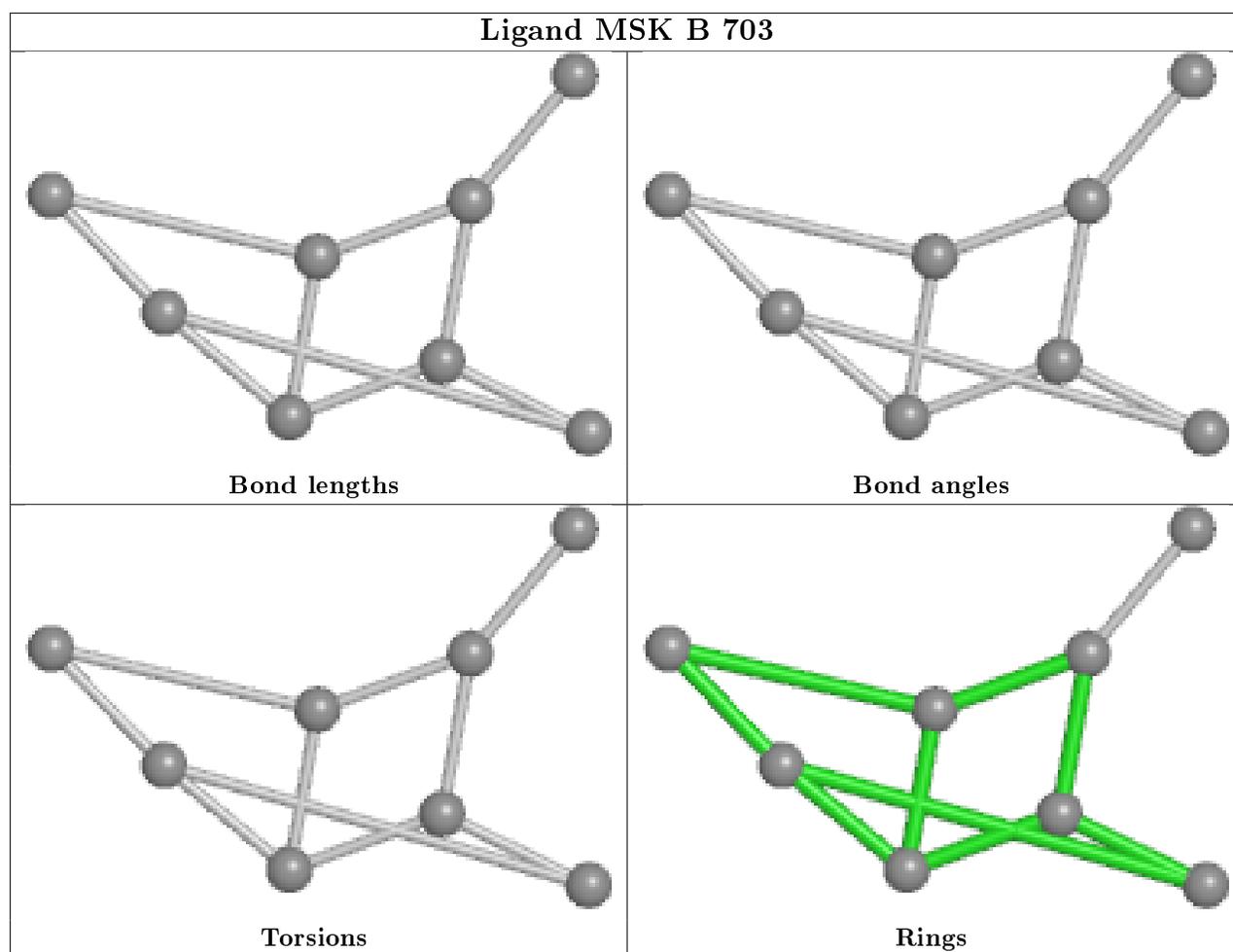
8 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	704	PO4	1	0
4	A	703	MSK	1	0
3	A	702	FES	0	2
8	A	705	CIT	2	0
5	D	705	PO4	1	0
4	B	703	MSK	1	0
6	B	705	BU3	2	0
7	B	708	PEG	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 than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	631:PRO	C	632:LEU	N	3.94

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

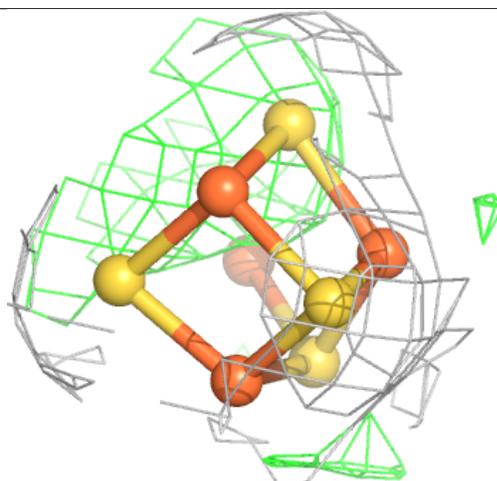
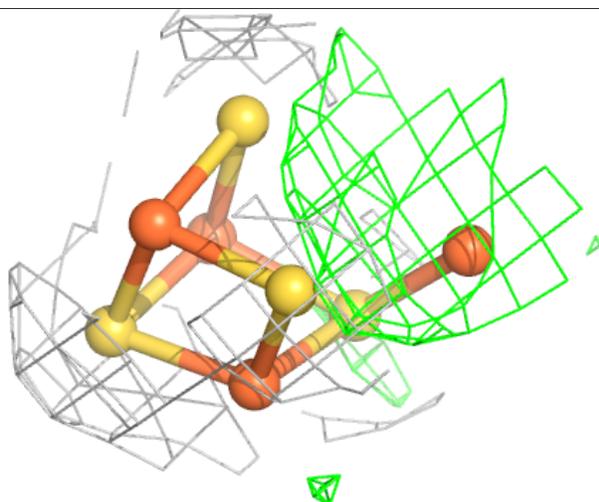
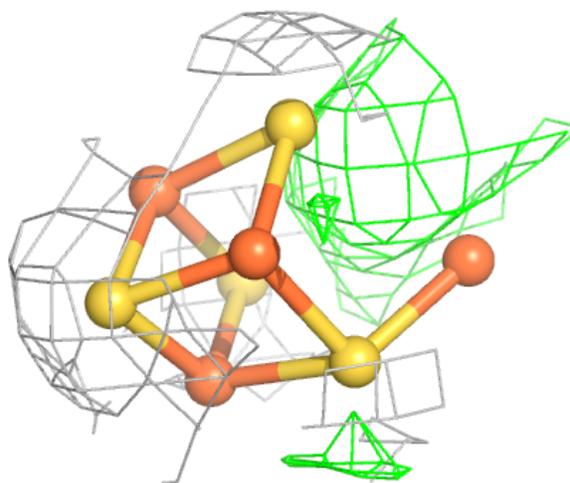
6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

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.

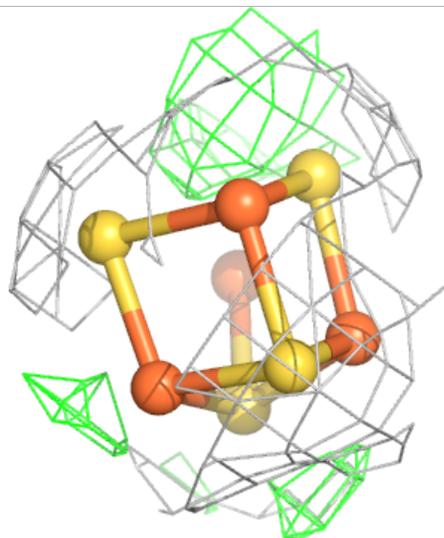
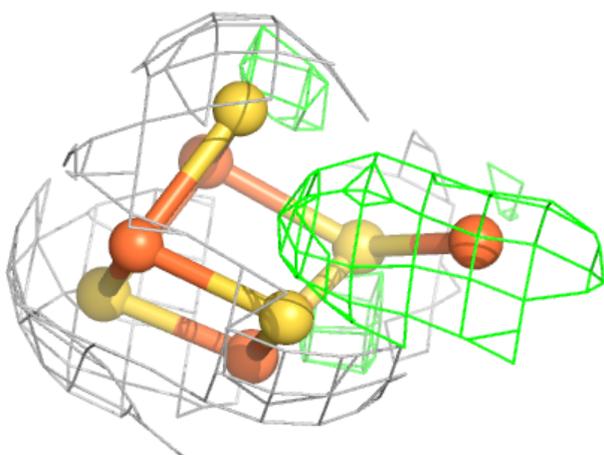
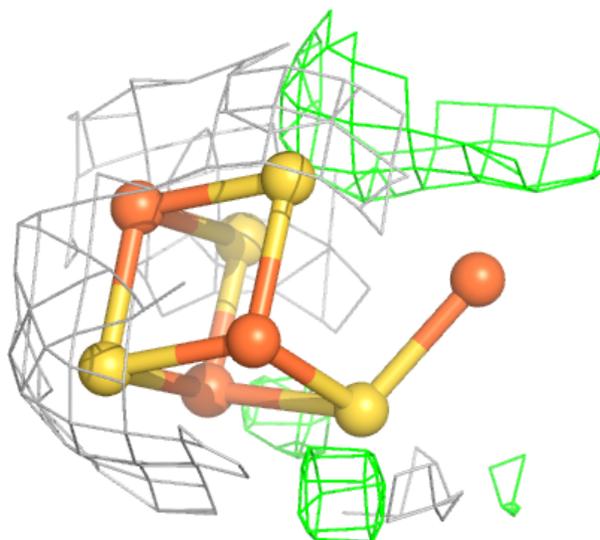
Electron density around MSK A 703:

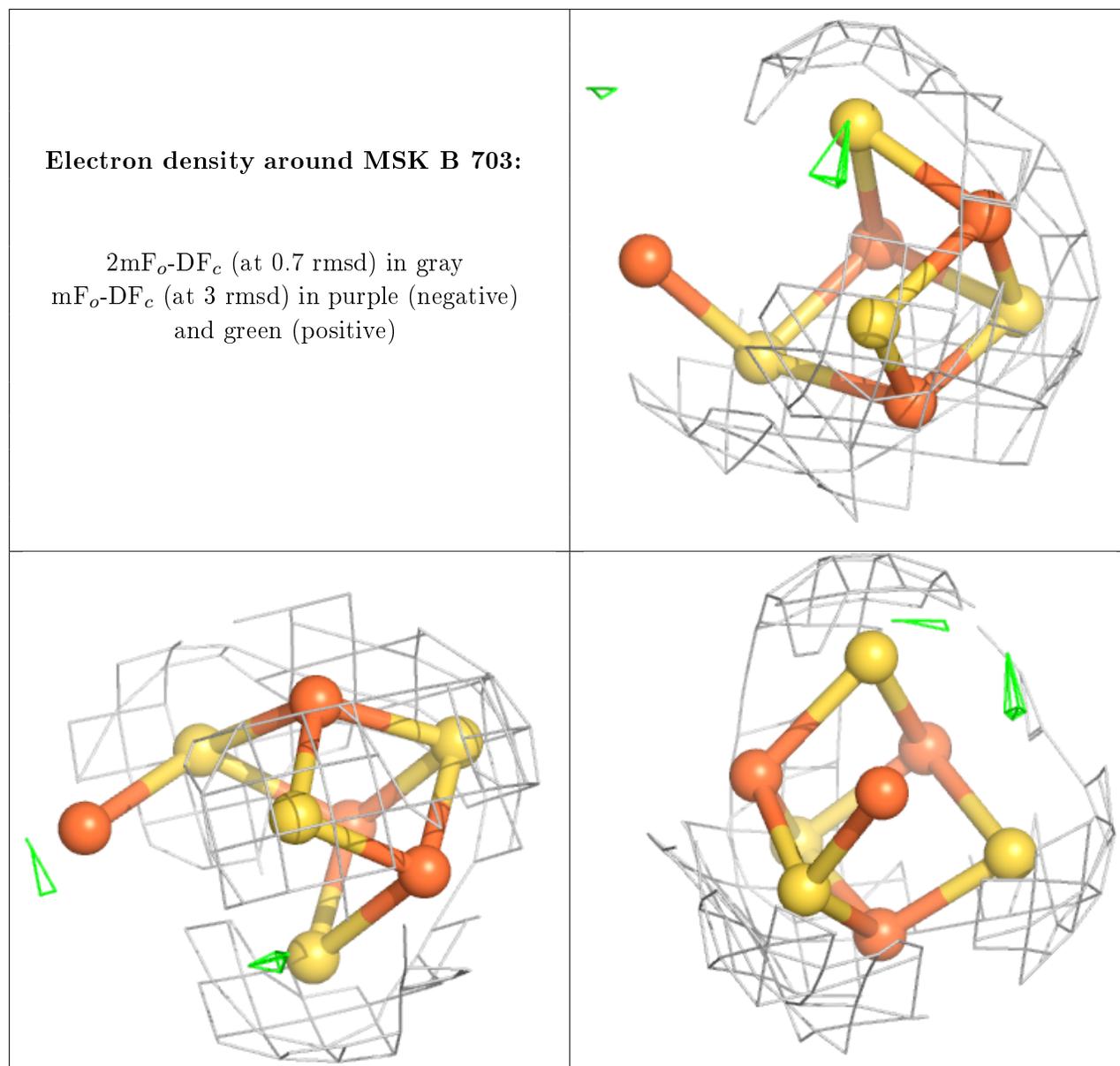
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around MSK D 703:

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





6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.