



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 18, 2023 – 10:18 am BST

PDB ID : 8CM7
Title : W-formate dehydrogenase M405A from *Desulfovibrio vulgaris*
Authors : Vilela-Alves, G.; Mota, C.; Oliveira, A.R.; Manuel, R.R.; Pereira, I.C.; Romao, M.J.
Deposited on : 2023-02-17
Resolution : 2.12 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

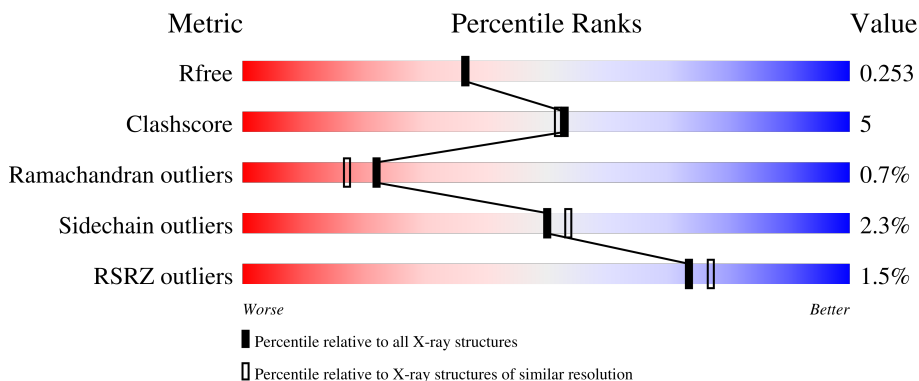
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.12 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

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 $\leq 5\%$. 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	1013	 83% 12% 6%
2	B	215	 80% 17% 2%

2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 9524 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 Formate dehydrogenase, alpha subunit, selenocysteine-containing.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	956	7496	4779	1309	1368	40	0	1	0

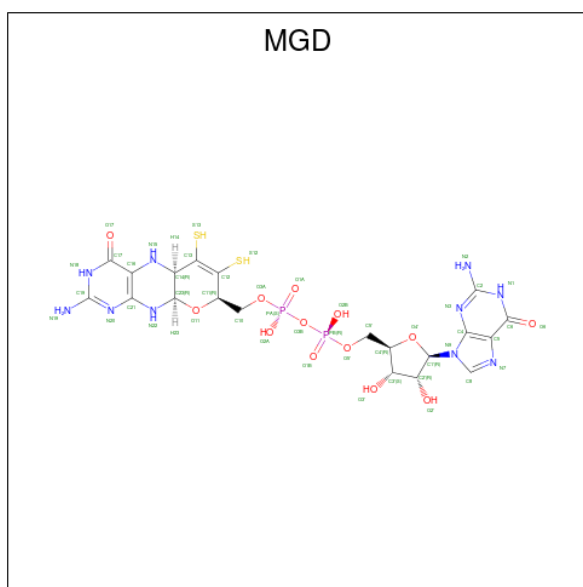
There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	405	ALA	MET	engineered mutation	UNP Q72EJ1
A	1006	TRP	-	expression tag	UNP Q72EJ1
A	1007	SER	-	expression tag	UNP Q72EJ1
A	1008	HIS	-	expression tag	UNP Q72EJ1
A	1009	PRO	-	expression tag	UNP Q72EJ1
A	1010	GLN	-	expression tag	UNP Q72EJ1
A	1011	PHE	-	expression tag	UNP Q72EJ1
A	1012	GLU	-	expression tag	UNP Q72EJ1
A	1013	LYS	-	expression tag	UNP Q72EJ1

- Molecule 2 is a protein called Formate dehydrogenase, beta subunit, putative.

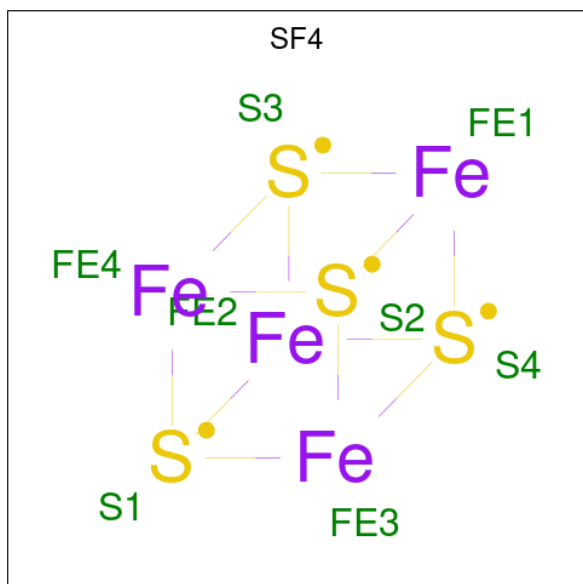
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	214	1676	1050	292	318	16	0	1	0

- Molecule 3 is 2-AMINO-5,6-DIMERCAPTO-7-METHYL-3,7,8A,9-TETRAHYDRO-8-OXA-1,3,9,10-TETRAAZA-ANTHRACEN-4-ONE GUANOSINE DINUCLEOTIDE (three-letter code: MGD) (formula: C₂₀H₂₆N₁₀O₁₃P₂S₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
3	A	1	47	20	10	13	2	2	0	0
3	A	1	47	20	10	13	2	2	0	0

- Molecule 4 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4) (labeled as "Ligand of Interest" by depositor).



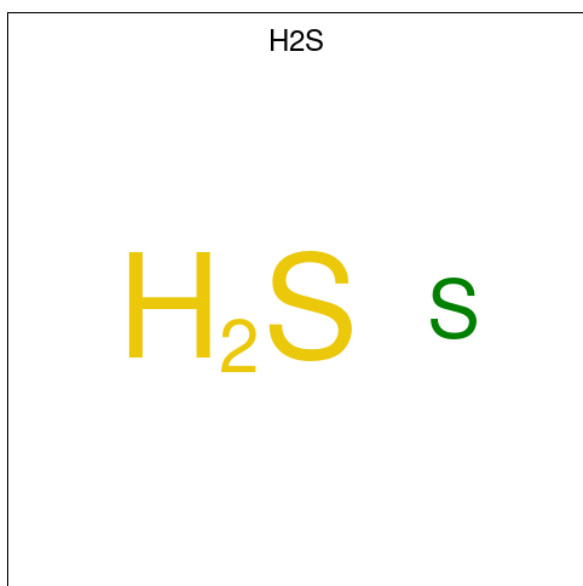
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	
			Total	Fe			S
4	A	1	8	4	4	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 5 is HYDROSULFURIC ACID (three-letter code: H2S) (formula: H₂S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	S	0	0
			1	1		

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

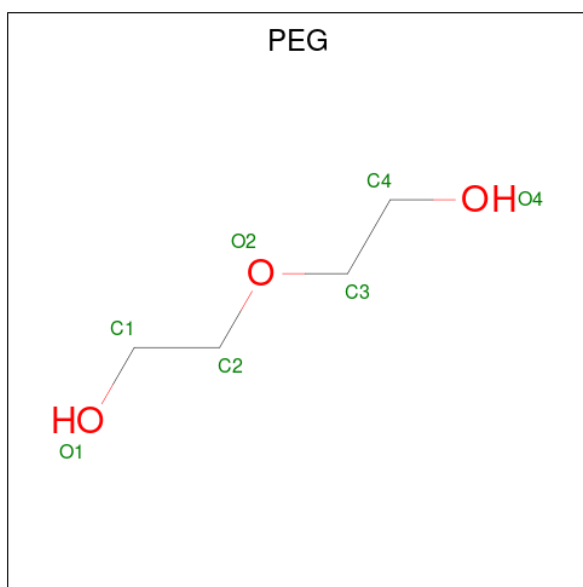
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	W	0	0
			1	1		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



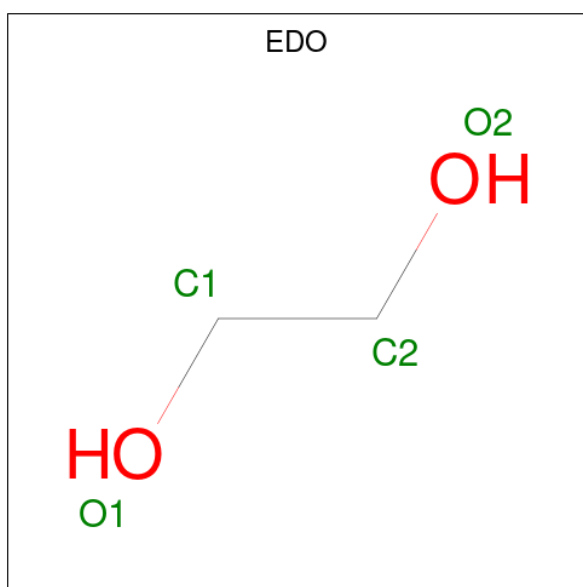
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			7	4	3		

Continued on next page...

Continued from previous page...

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

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		
9	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 10 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	2	Total	Cl	0	0
			2	2		

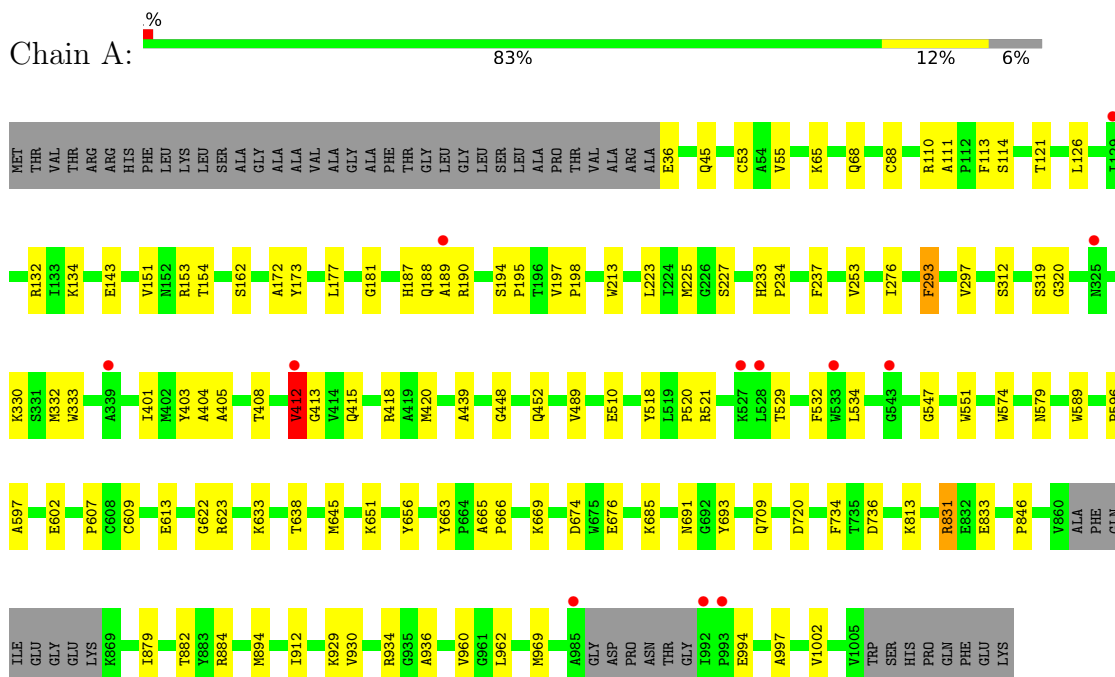
- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	137	Total 137	O 137	0	0
11	B	30	Total 30	O 30	0	0

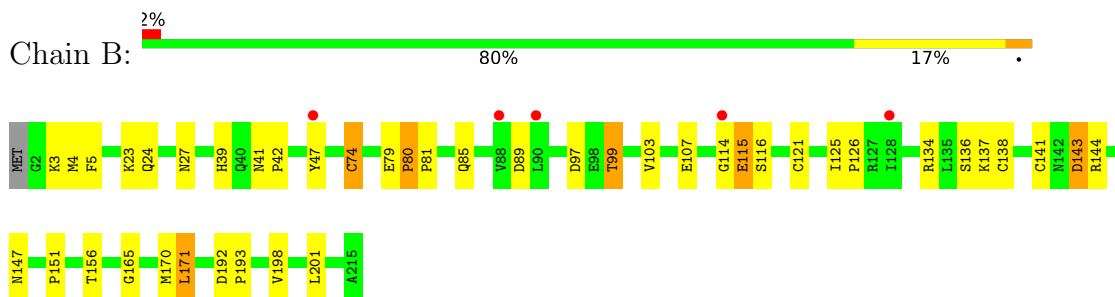
3 Residue-property plots [i](#)

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: Formate dehydrogenase, alpha subunit, selenocysteine-containing



- Molecule 2: Formate dehydrogenase, beta subunit, putative



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	64.87Å 123.83Å 149.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	75.06 – 2.12 74.95 – 2.12	Depositor EDS
% Data completeness (in resolution range)	68.4 (75.06-2.12) 68.4 (74.95-2.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 2.12Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.197 , 0.252 0.204 , 0.253	Depositor DCC
R_{free} test set	2417 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	37.6	Xtrriage
Anisotropy	0.090	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 41.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	9524	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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: W, PEG, GOL, CL, SEC, H2S, SF4, MGD, EDO

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	0.69	0/7696	0.83	0/10438
2	B	0.70	0/1712	0.88	0/2320
All	All	0.69	0/9408	0.84	0/12758

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 [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	7496	0	7338	69	0
2	B	1676	0	1641	32	0
3	A	94	0	46	4	0
4	A	8	0	0	0	0
4	B	24	0	0	0	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0
7	A	18	0	24	2	0
8	A	14	0	20	0	0
8	B	7	0	10	1	0
9	A	16	0	24	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	A	2	0	0	0	0
11	A	137	0	0	2	0
11	B	30	0	0	0	0
All	All	9524	0	9103	100	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 5.

All (100) 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)
2:B:79:GLU:O	2:B:81:PRO:HD3	1.80	0.82
1:A:172:ALA:HB3	1:A:645:MET:HE1	1.65	0.79
1:A:676:GLU:O	1:A:685:LYS:NZ	2.15	0.76
1:A:113:PHE:HD1	1:A:597:ALA:O	1.82	0.62
2:B:4:MET:SD	2:B:171:LEU:HD22	2.40	0.61
1:A:882:THR:HA	1:A:962:LEU:O	2.01	0.60
1:A:162:SER:HB2	1:A:551:TRP:O	2.02	0.59
2:B:143:ASP:OD1	2:B:144:ARG:N	2.30	0.59
1:A:330:LYS:HG2	1:A:333:TRP:CZ3	2.37	0.59
1:A:172:ALA:HB3	1:A:645:MET:CE	2.34	0.58
1:A:36:GLU:N	2:B:147:ASN:HD21	2.02	0.58
1:A:113:PHE:CD1	1:A:597:ALA:O	2.57	0.58
1:A:153:ARG:NH1	1:A:521:ARG:O	2.37	0.57
1:A:173:TYR:CE2	1:A:177:LEU:HD11	2.42	0.55
2:B:23:LYS:O	2:B:27:ASN:N	2.40	0.54
1:A:190:ARG:HA	1:A:194:SER:HB3	1.90	0.54
2:B:85:GLN:HG2	2:B:116:SER:O	2.09	0.52
1:A:276:ILE:HD12	1:A:412:VAL:HG12	1.91	0.52
1:A:894:MET:SD	3:A:1102:MGD:H2'	2.49	0.52
2:B:107:GLU:H	2:B:107:GLU:CD	2.12	0.52
1:A:111:ALA:O	1:A:114:SER:HB3	2.10	0.52
1:A:930:VAL:HG22	1:A:1002:VAL:HG23	1.92	0.51
1:A:194:SER:N	1:A:195:PRO:CD	2.75	0.50
1:A:622:GLY:O	1:A:623:ARG:HB2	2.11	0.50
2:B:5:PHE:CD1	2:B:198:VAL:HG22	2.46	0.50
1:A:401:ILE:HG21	1:A:420:MET:HG3	1.92	0.50
1:A:833:GLU:OE1	1:A:833:GLU:N	2.41	0.49
2:B:4:MET:SD	2:B:171:LEU:CD2	3.00	0.49
1:A:110:ARG:HG3	1:A:589:TRP:CE2	2.48	0.48
2:B:47[B]:TYR:CE1	2:B:136:SER:HB3	2.48	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:121:THR:HG22	11:A:1310:HOH:O	2.13	0.48
1:A:929:LYS:HE2	1:A:936:ALA:HB1	1.94	0.48
2:B:47[B]:TYR:HA	2:B:136:SER:OG	2.14	0.48
1:A:188:GLN:HB3	3:A:1102:MGD:S12	2.54	0.47
1:A:312:SER:HA	7:A:1111:GOL:H12	1.96	0.47
1:A:529:THR:HA	1:A:532:PHE:CD2	2.50	0.47
1:A:151:VAL:O	1:A:520:PRO:HA	2.15	0.46
1:A:225:MET:HG3	1:A:253:VAL:HB	1.97	0.46
1:A:609:CYS:HB2	1:A:613:GLU:OE1	2.15	0.46
2:B:121:CYS:SG	2:B:126:PRO:HD3	2.54	0.46
1:A:233:HIS:N	1:A:234:PRO:HD3	2.31	0.46
1:A:665:ALA:N	1:A:666:PRO:HD2	2.30	0.46
1:A:401:ILE:HG21	1:A:420:MET:CG	2.45	0.46
2:B:137:LYS:HG3	8:B:304:PEG:H42	1.98	0.46
2:B:171:LEU:HD13	2:B:171:LEU:HA	1.83	0.46
1:A:320:GLY:CA	1:A:332:MET:HG3	2.46	0.46
1:A:173:TYR:CZ	1:A:177:LEU:HD11	2.51	0.46
2:B:47[A]:TYR:HA	2:B:136:SER:HG	1.81	0.45
2:B:47[A]:TYR:HA	2:B:136:SER:OG	2.16	0.45
1:A:405:ALA:HA	1:A:408:THR:OG1	2.17	0.45
2:B:201:LEU:C	2:B:201:LEU:HD23	2.37	0.45
1:A:518:TYR:HB3	1:A:666:PRO:HB3	1.98	0.45
1:A:884:ARG:HH22	3:A:1102:MGD:H15	1.63	0.45
1:A:912:ILE:HG22	1:A:960:VAL:HG12	1.98	0.45
1:A:114:SER:O	1:A:596:PRO:HB3	2.17	0.45
1:A:121:THR:CG2	11:A:1310:HOH:O	2.64	0.45
1:A:720:ASP:OD1	1:A:720:ASP:C	2.56	0.45
1:A:312:SER:OG	1:A:319:SER:HB3	2.16	0.45
2:B:114:GLY:O	2:B:115:GLU:HB2	2.16	0.45
1:A:709:GLN:HG3	1:A:736:ASP:OD2	2.17	0.45
2:B:107:GLU:HB3	2:B:134:ARG:HH21	1.82	0.45
2:B:114:GLY:O	2:B:115:GLU:CB	2.64	0.45
1:A:181:GLY:HA2	1:A:520:PRO:HG3	1.99	0.44
2:B:47[B]:TYR:HA	2:B:136:SER:HG	1.82	0.44
1:A:197:VAL:HB	1:A:198:PRO:HD3	2.00	0.44
1:A:418:ARG:HD3	1:A:831:ARG:HD2	1.99	0.44
2:B:141:CYS:HB2	2:B:151:PRO:HB3	2.00	0.44
1:A:674:ASP:OD2	1:A:693:TYR:OH	2.31	0.43
1:A:132:ARG:NH2	1:A:602:GLU:OE2	2.37	0.43
1:A:534:LEU:HD12	1:A:534:LEU:HA	1.84	0.43
1:A:126:LEU:HB3	1:A:651:LYS:HD2	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:320:GLY:HA3	1:A:332:MET:HG3	2.01	0.43
1:A:401:ILE:HG12	1:A:420:MET:HG2	2.01	0.43
1:A:403:TYR:CD1	1:A:403:TYR:C	2.92	0.43
1:A:489:VAL:HG22	1:A:813:LYS:HD2	2.00	0.43
1:A:415:GLN:HE22	1:A:831:ARG:HG2	1.84	0.43
1:A:55:VAL:HA	1:A:213:TRP:CE2	2.54	0.42
1:A:134:LYS:HG3	1:A:656:TYR:CZ	2.55	0.42
2:B:79:GLU:O	2:B:81:PRO:CD	2.59	0.42
1:A:187:HIS:NE2	1:A:189:ALA:HB2	2.34	0.42
1:A:223:LEU:HD11	1:A:225:MET:HE2	2.01	0.42
2:B:3:LYS:HA	2:B:165:GLY:O	2.20	0.42
1:A:404:ALA:HB3	3:A:1101:MGD:O2A	2.19	0.42
2:B:41:ASN:HA	2:B:42:PRO:C	2.40	0.42
1:A:293:PHE:O	1:A:297:VAL:HG23	2.20	0.41
2:B:80:PRO:HG3	2:B:103:VAL:HG21	2.01	0.41
1:A:846:PRO:HG3	1:A:879:ILE:HD11	2.02	0.41
1:A:110:ARG:HG3	1:A:589:TRP:CZ2	2.55	0.41
1:A:510:GLU:OE2	1:A:669:LYS:NZ	2.53	0.41
1:A:547:GLY:HA3	1:A:574:TRP:CZ2	2.55	0.41
2:B:125:ILE:HB	2:B:126:PRO:HD3	2.02	0.41
1:A:607:PRO:HB2	1:A:638:THR:HG22	2.02	0.41
2:B:74:CYS:HB2	2:B:125:ILE:HG21	2.02	0.41
7:A:1106:GOL:H31	2:B:39:HIS:CE1	2.56	0.41
1:A:65:LYS:HB2	1:A:68:GLN:HB2	2.02	0.40
1:A:448:GLY:O	1:A:452:GLN:HG2	2.21	0.40
2:B:4:MET:O	2:B:170:MET:HG3	2.22	0.40
2:B:97:ASP:OD1	2:B:99:THR:N	2.54	0.40
2:B:192:ASP:N	2:B:193:PRO:CD	2.84	0.40
1:A:994:GLU:HG3	1:A:997:ALA:O	2.21	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	950/1013 (94%)	900 (95%)	46 (5%)	4 (0%)	34	32
2	B	213/215 (99%)	195 (92%)	14 (7%)	4 (2%)	8	3
All	All	1163/1228 (95%)	1095 (94%)	60 (5%)	8 (1%)	22	17

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	412	VAL
2	B	115	GLU
1	A	413	GLY
1	A	439	ALA
2	B	89	ASP
2	B	143	ASP
1	A	663	TYR
2	B	80	PRO

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	778/818 (95%)	762 (98%)	16 (2%)	53	57
2	B	186/186 (100%)	180 (97%)	6 (3%)	39	40
All	All	964/1004 (96%)	942 (98%)	22 (2%)	50	53

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	GLN
1	A	53	CYS
1	A	88	CYS
1	A	143	GLU
1	A	154	THR
1	A	227	SER
1	A	237	PHE
1	A	293	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	412	VAL
1	A	579	ASN
1	A	633	LYS
1	A	691	ASN
1	A	734	PHE
1	A	831	ARG
1	A	934	ARG
1	A	969	MET
2	B	24	GLN
2	B	74	CYS
2	B	99	THR
2	B	138	CYS
2	B	156	THR
2	B	171	LEU

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](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 1 is modelled with single atom and 3 are monoatomic - leaving 16 for Mogul analysis.

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$
7	GOL	A	1111	-	5,5,5	0.11	0	5,5,5	0.39	0
4	SF4	A	1103	1	0,12,12	-	-	-	-	-
9	EDO	A	1108	-	3,3,3	0.31	0	2,2,2	0.36	0
9	EDO	A	1110	-	3,3,3	0.04	0	2,2,2	0.11	0
3	MGD	A	1102	6	41,52,52	0.82	1 (2%)	40,81,81	1.36	5 (12%)
8	PEG	A	1113	-	6,6,6	0.33	0	5,5,5	0.26	0
4	SF4	B	302	2	0,12,12	-	-	-	-	-
8	PEG	B	304	-	6,6,6	0.36	0	5,5,5	0.26	0
4	SF4	B	301	2	0,12,12	-	-	-	-	-
9	EDO	A	1112	-	3,3,3	0.11	0	2,2,2	0.05	0
8	PEG	A	1107	-	6,6,6	0.28	0	5,5,5	0.32	0
9	EDO	A	1114	-	3,3,3	0.11	0	2,2,2	0.13	0
7	GOL	A	1109	-	5,5,5	0.18	0	5,5,5	0.51	0
7	GOL	A	1106	-	5,5,5	0.11	0	5,5,5	0.24	0
4	SF4	B	303	2	0,12,12	-	-	-	-	-
3	MGD	A	1101	6	41,52,52	0.78	1 (2%)	40,81,81	1.09	3 (7%)

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
7	GOL	A	1111	-	-	2/4/4/4	-
9	EDO	A	1108	-	-	0/1/1/1	-
4	SF4	A	1103	1	-	-	0/6/5/5
4	SF4	B	303	2	-	-	0/6/5/5
9	EDO	A	1110	-	-	1/1/1/1	-
3	MGD	A	1102	6	-	7/18/66/66	0/6/6/6
9	EDO	A	1112	-	-	0/1/1/1	-
4	SF4	B	302	2	-	-	0/6/5/5
8	PEG	B	304	-	-	1/4/4/4	-
4	SF4	B	301	2	-	-	0/6/5/5
8	PEG	A	1107	-	-	2/4/4/4	-
9	EDO	A	1114	-	-	1/1/1/1	-
7	GOL	A	1109	-	-	4/4/4/4	-
7	GOL	A	1106	-	-	2/4/4/4	-
8	PEG	A	1113	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MGD	A	1101	6	-	1/18/66/66	0/6/6/6

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1102	MGD	C5-C6	-2.30	1.42	1.47
3	A	1101	MGD	C5-C6	-2.07	1.43	1.47

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1102	MGD	O11-C23-C14	5.33	112.52	108.96
3	A	1102	MGD	C19-N20-C21	2.85	118.57	113.43
3	A	1101	MGD	C19-N20-C21	2.85	118.57	113.43
3	A	1101	MGD	O6-C6-C5	2.79	129.81	124.37
3	A	1102	MGD	C17-C16-N15	2.58	123.69	116.76
3	A	1102	MGD	O6-C6-C5	2.19	128.64	124.37
3	A	1101	MGD	O4'-C1'-C2'	-2.06	103.92	106.93
3	A	1102	MGD	O4'-C1'-C2'	-2.04	103.95	106.93

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1102	MGD	C10-O3A-PA-O1A
3	A	1102	MGD	C4'-C5'-O5'-PB
7	A	1106	GOL	O1-C1-C2-O2
7	A	1106	GOL	O1-C1-C2-C3
7	A	1109	GOL	O1-C1-C2-C3
7	A	1111	GOL	O1-C1-C2-C3
3	A	1102	MGD	O4'-C4'-C5'-O5'
3	A	1102	MGD	C3'-C4'-C5'-O5'
8	B	304	PEG	O1-C1-C2-O2
8	A	1113	PEG	O2-C3-C4-O4
7	A	1109	GOL	O1-C1-C2-O2
9	A	1110	EDO	O1-C1-C2-O2
7	A	1111	GOL	O1-C1-C2-O2
3	A	1101	MGD	PA-O3B-PB-O5'
3	A	1102	MGD	PA-O3B-PB-O5'
8	A	1107	PEG	O1-C1-C2-O2
3	A	1102	MGD	C10-O3A-PA-O3B

Continued on next page...

Continued from previous page...

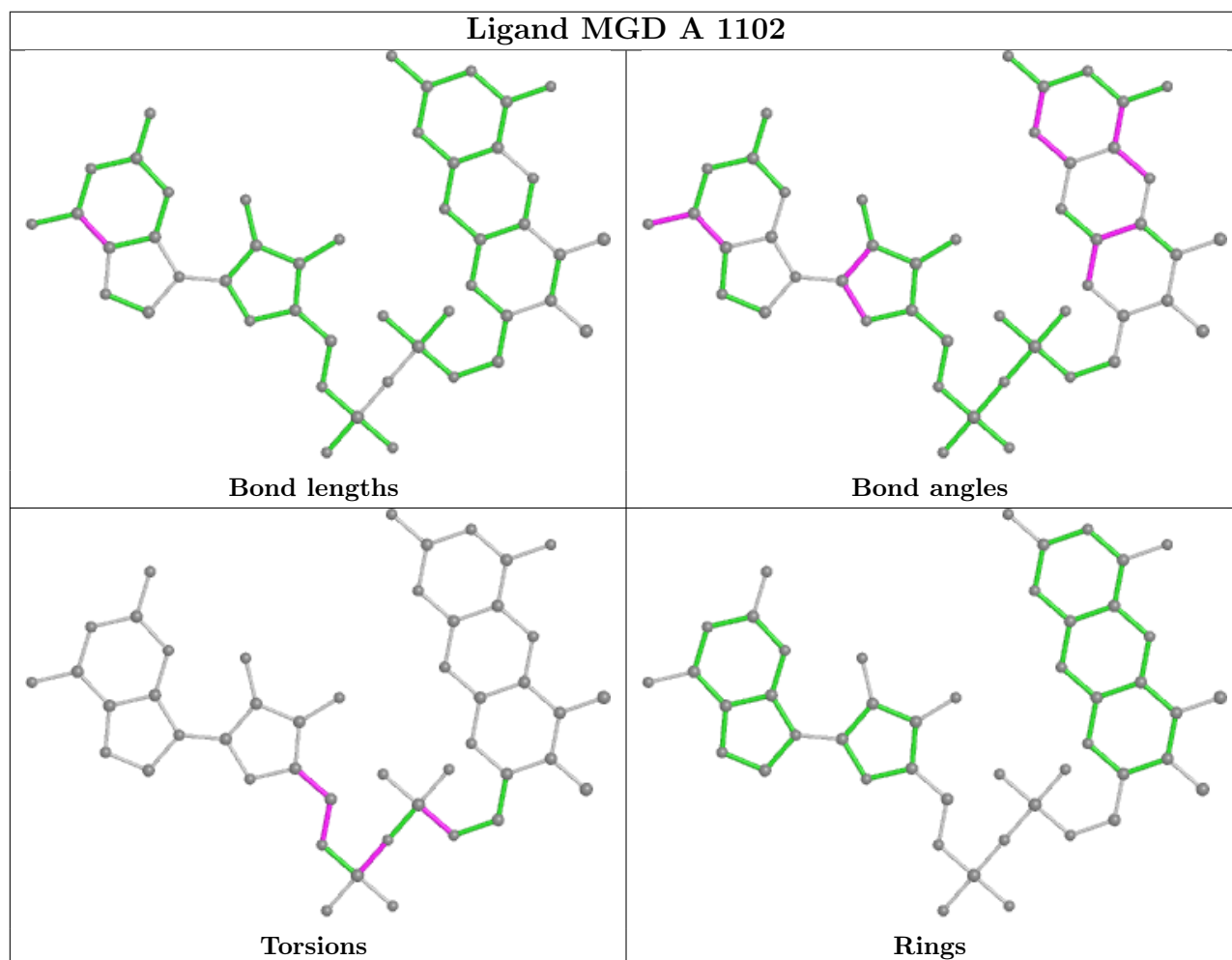
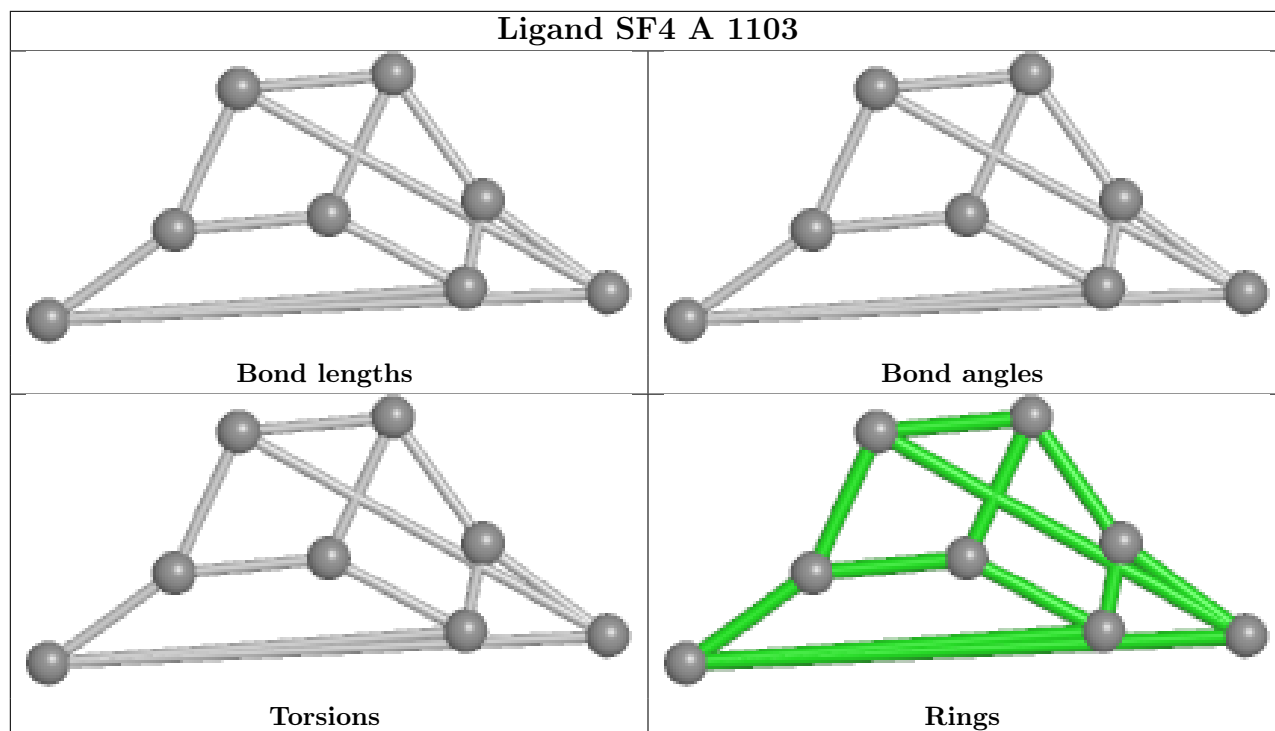
Mol	Chain	Res	Type	Atoms
3	A	1102	MGD	C10-O3A-PA-O2A
9	A	1114	EDO	O1-C1-C2-O2
8	A	1113	PEG	C4-C3-O2-C2
7	A	1109	GOL	O2-C2-C3-O3
7	A	1109	GOL	C1-C2-C3-O3
8	A	1107	PEG	C4-C3-O2-C2

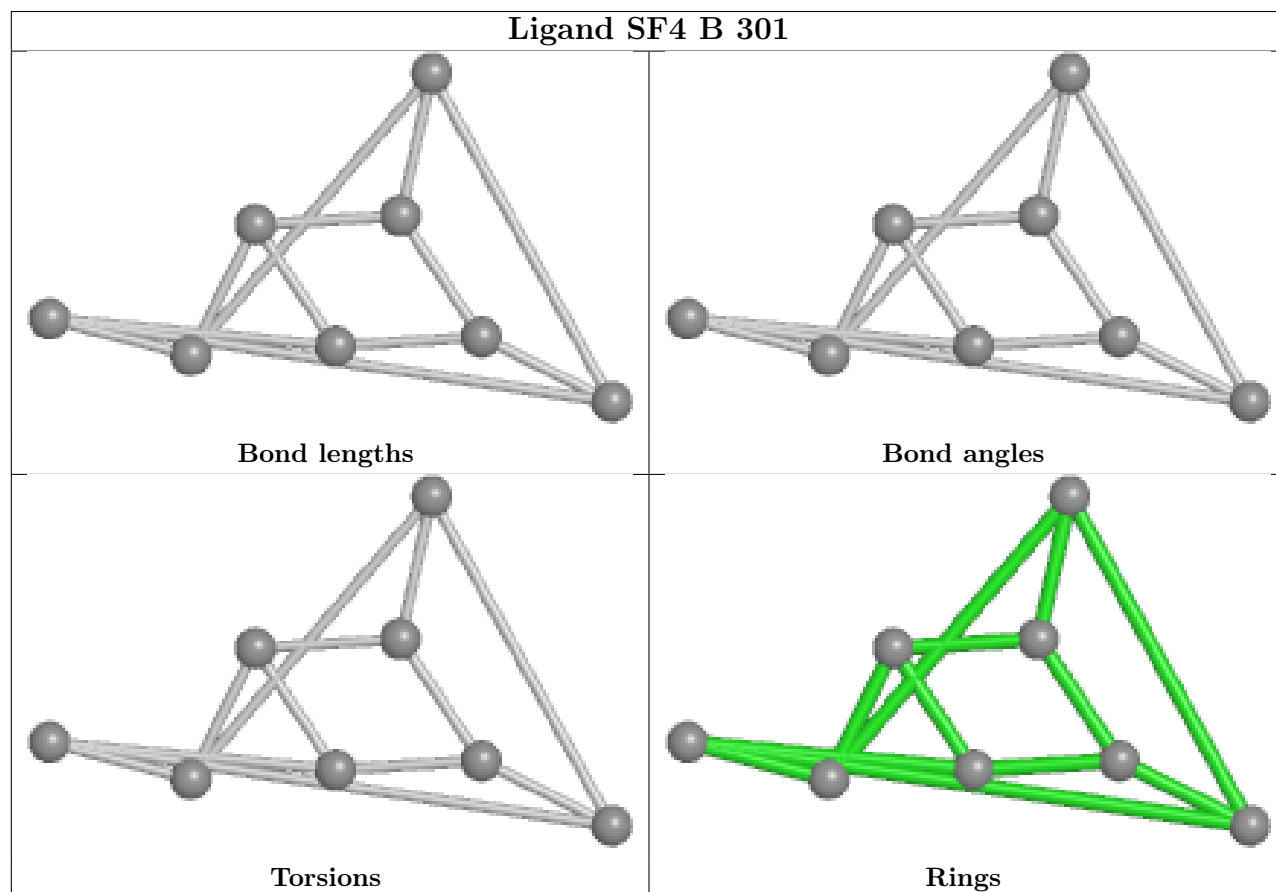
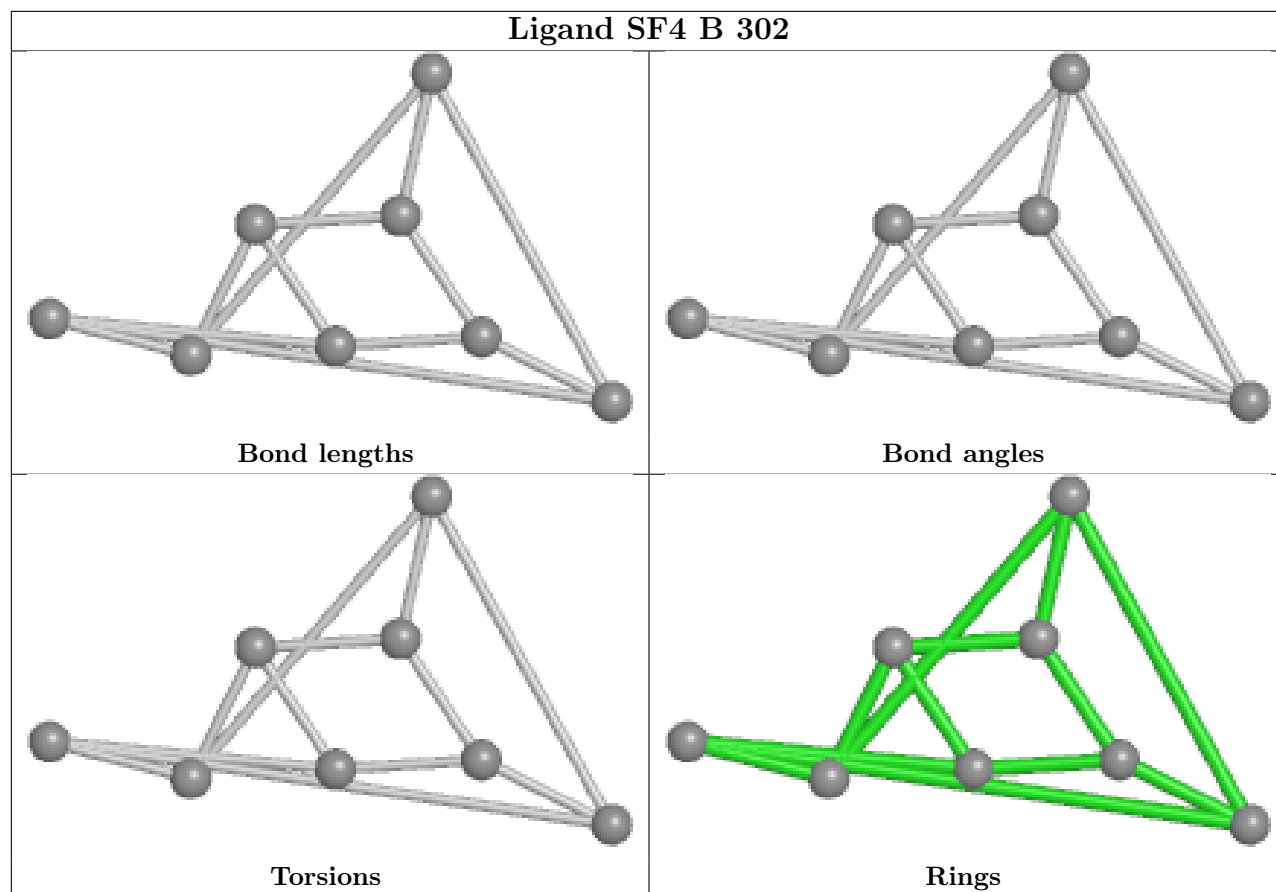
There are no ring outliers.

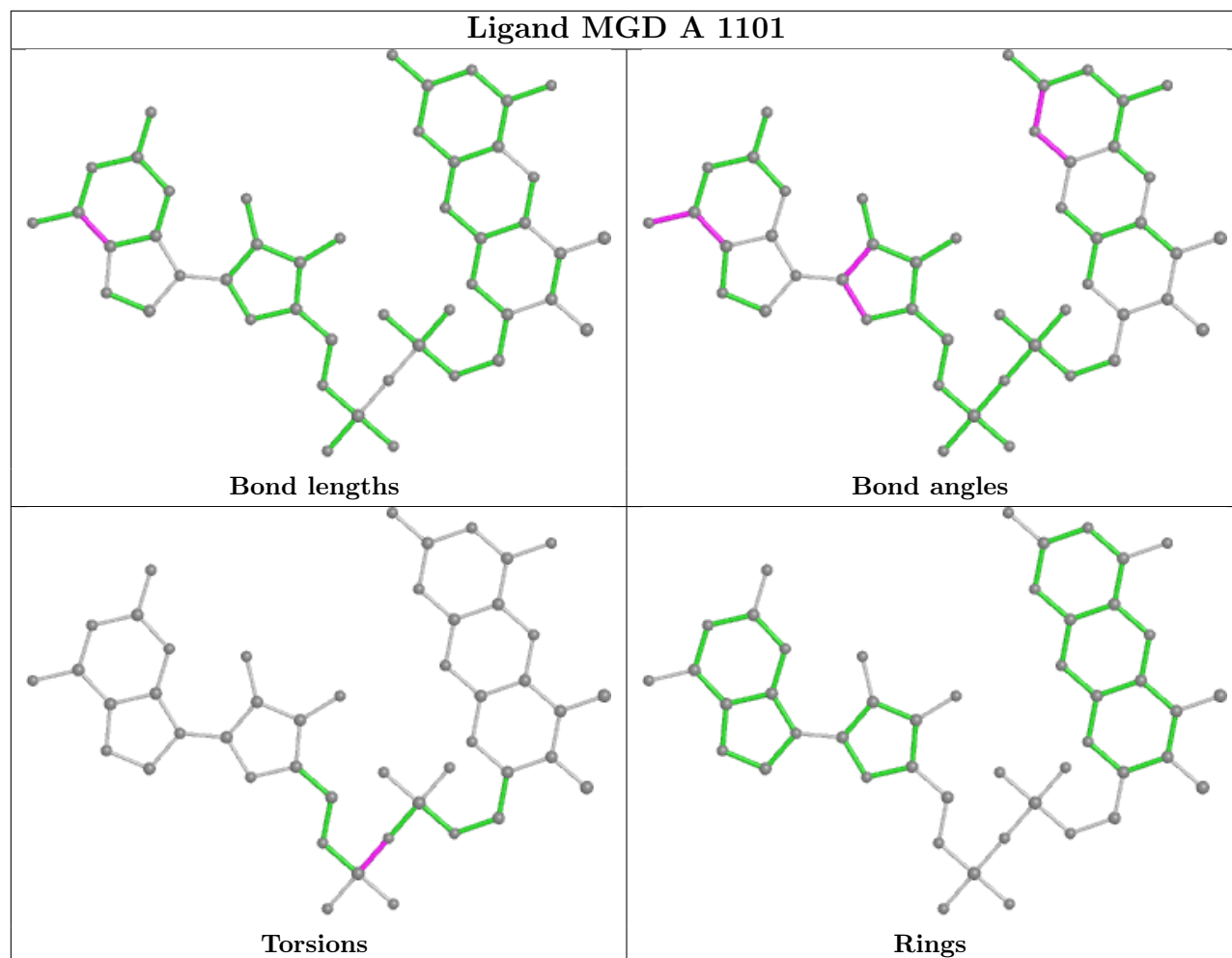
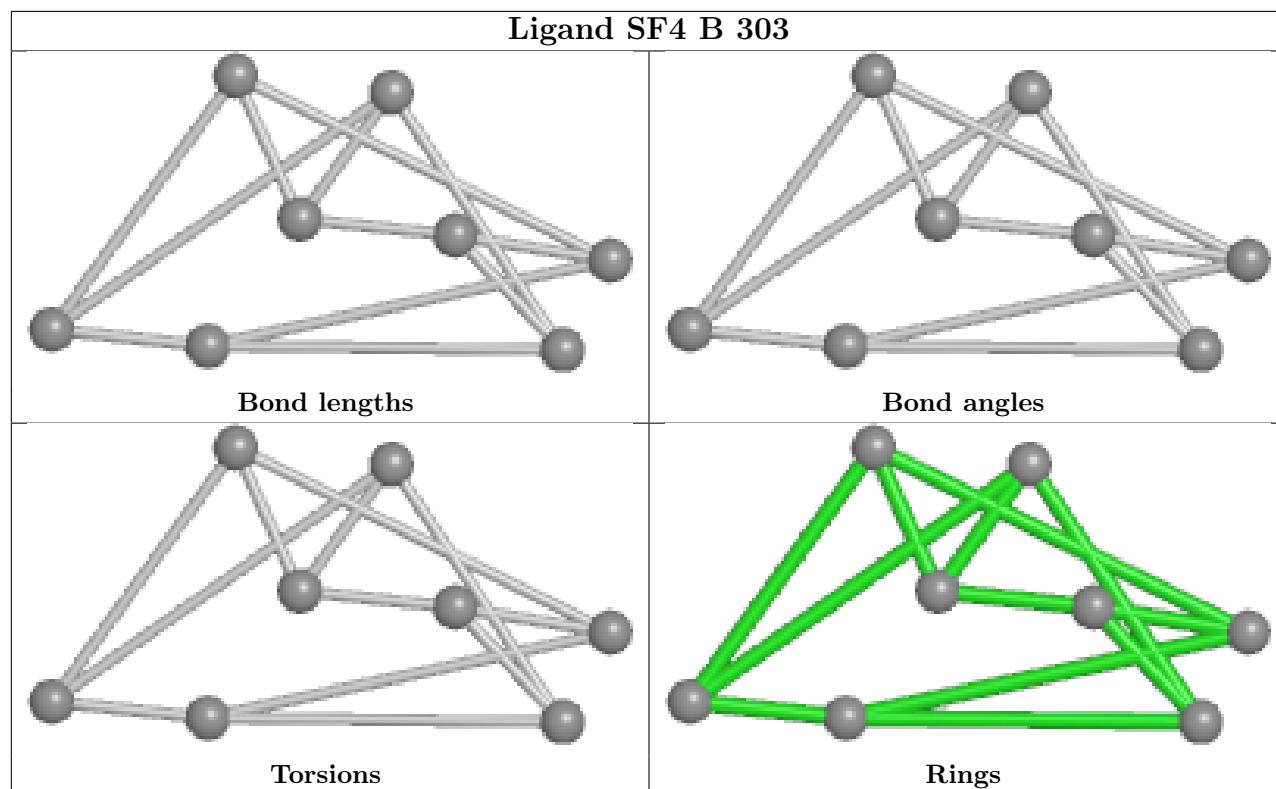
5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1111	GOL	1	0
3	A	1102	MGD	3	0
8	B	304	PEG	1	0
7	A	1106	GOL	1	0
3	A	1101	MGD	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](#)

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	955/1013 (94%)	0.30	12 (1%) 77 80	18, 39, 65, 109	0
2	B	214/215 (99%)	0.42	5 (2%) 60 65	23, 40, 65, 76	1 (0%)
All	All	1169/1228 (95%)	0.32	17 (1%) 73 77	18, 39, 65, 109	1 (0%)

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	88	VAL	3.8
1	A	992	ILE	3.2
1	A	189	ALA	2.8
1	A	528	LEU	2.8
1	A	325	ASN	2.7
1	A	339	ALA	2.7
1	A	993	PRO	2.6
1	A	533	TRP	2.5
1	A	543	GLY	2.5
1	A	412	VAL	2.5
2	B	114	GLY	2.5
1	A	985	ALA	2.4
1	A	527	LYS	2.3
2	B	128	ILE	2.3
2	B	47[A]	TYR	2.2
1	A	129	ILE	2.2
2	B	90	LEU	2.1

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](#)

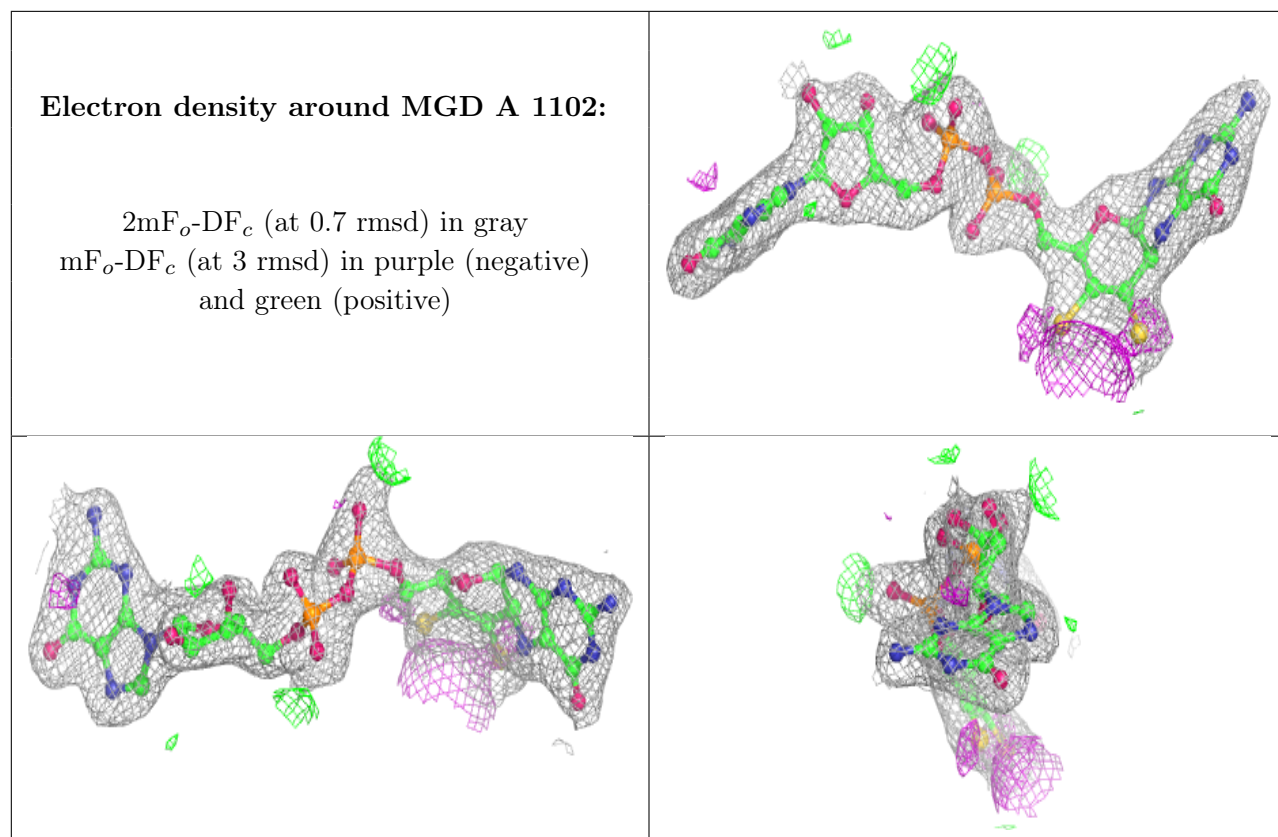
There are no monosaccharides in this entry.

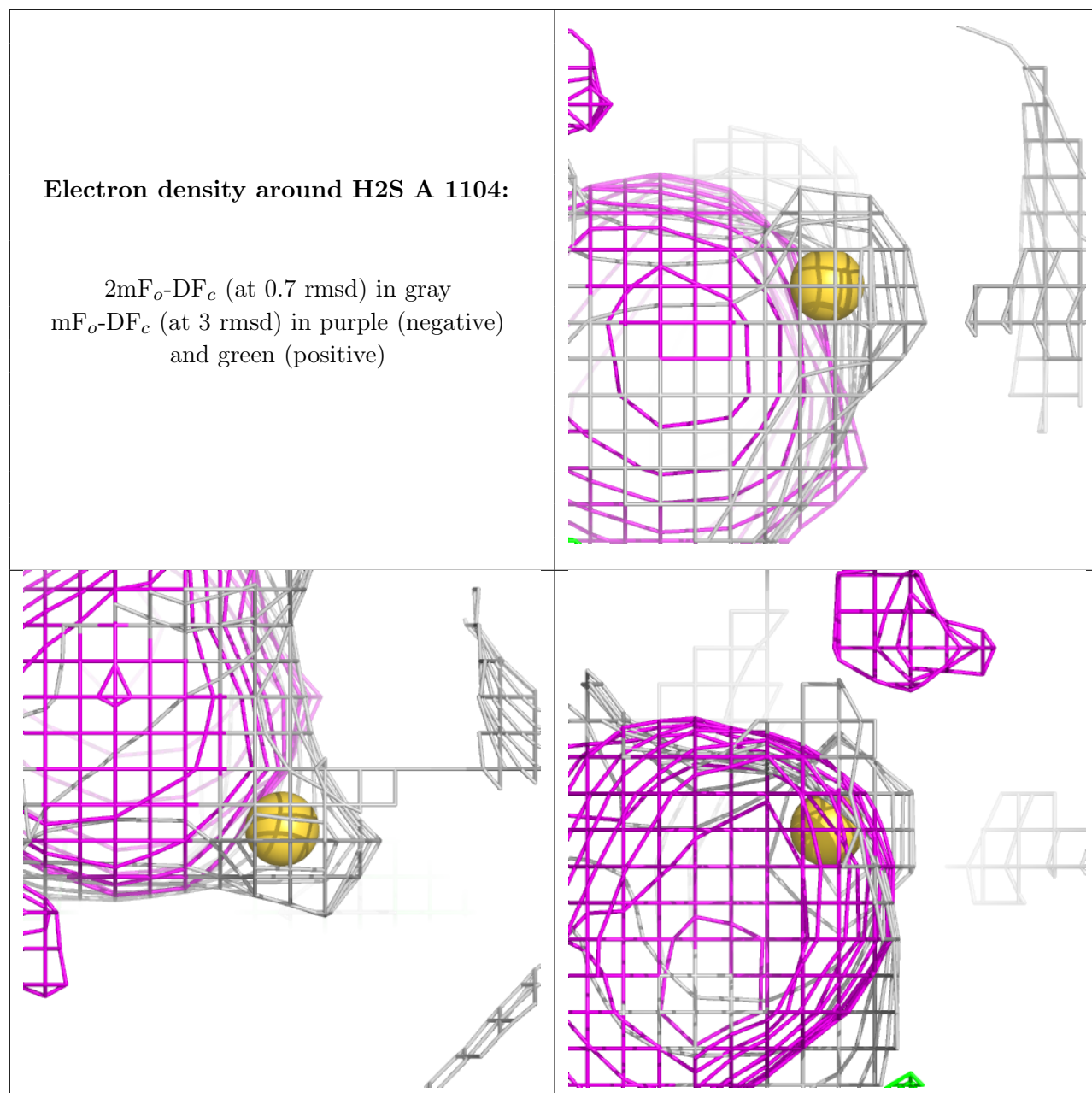
6.4 Ligands [i](#)

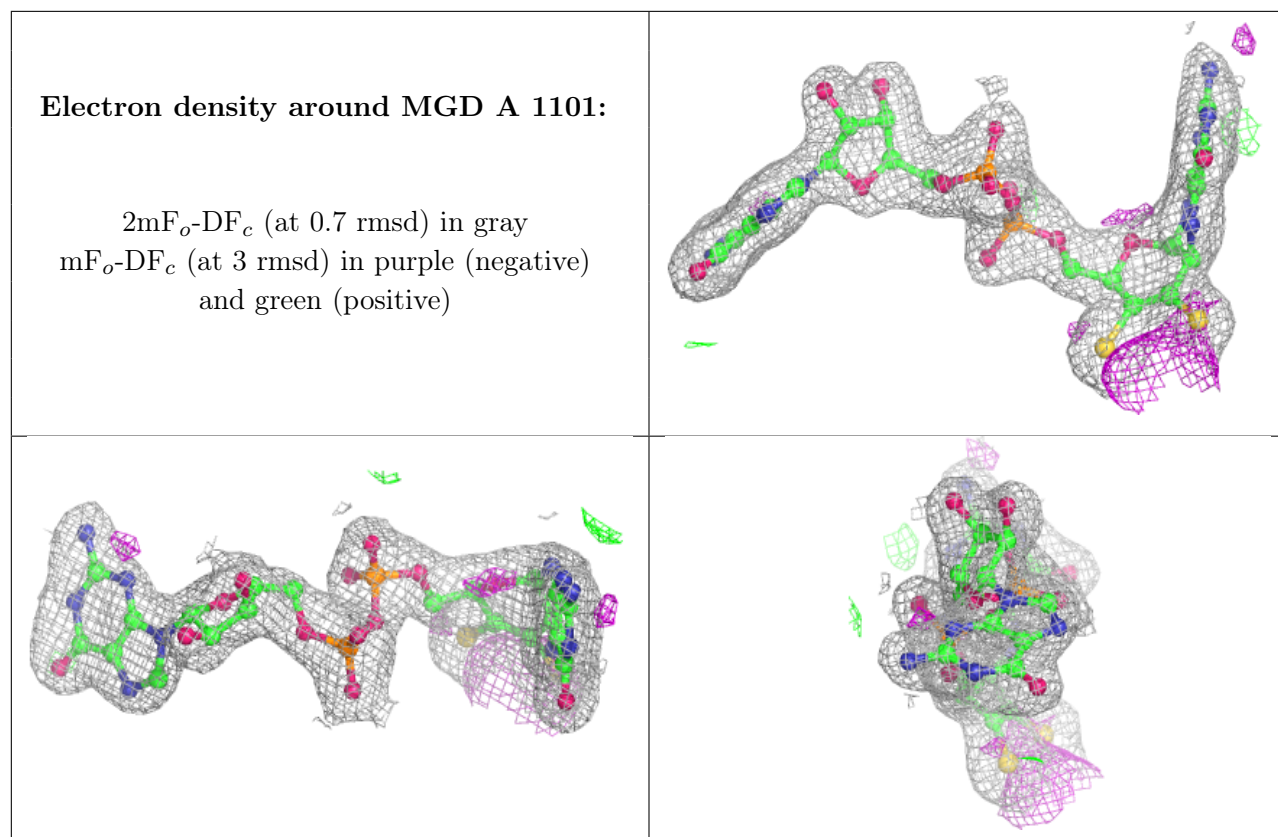
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
8	PEG	A	1113	7/7	0.78	0.20	50,54,58,62	0
9	EDO	A	1114	4/4	0.79	0.21	50,54,56,57	0
9	EDO	A	1110	4/4	0.81	0.17	53,54,59,65	0
10	CL	A	1116	1/1	0.84	0.19	73,73,73,73	0
7	GOL	A	1111	6/6	0.85	0.14	47,50,54,56	0
9	EDO	A	1112	4/4	0.89	0.14	60,61,61,62	0
8	PEG	A	1107	7/7	0.90	0.15	39,39,40,48	0
7	GOL	A	1109	6/6	0.91	0.14	42,50,53,53	0
8	PEG	B	304	7/7	0.91	0.19	25,38,47,52	0
7	GOL	A	1106	6/6	0.93	0.18	35,35,37,37	0
3	MGD	A	1102	47/47	0.94	0.14	31,41,48,60	0
10	CL	A	1115	1/1	0.94	0.17	70,70,70,70	0
5	H2S	A	1104	1/1	0.94	0.38	77,77,77,77	0
9	EDO	A	1108	4/4	0.95	0.21	38,38,41,41	0
3	MGD	A	1101	47/47	0.97	0.14	24,28,37,38	0
4	SF4	B	302	8/8	0.98	0.08	36,38,43,45	0
4	SF4	A	1103	8/8	0.99	0.08	22,25,26,26	0
4	SF4	B	303	8/8	0.99	0.06	35,36,39,39	0
4	SF4	B	301	8/8	0.99	0.08	25,26,27,28	0
6	W	A	1105	1/1	1.00	0.04	62,62,62,62	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.

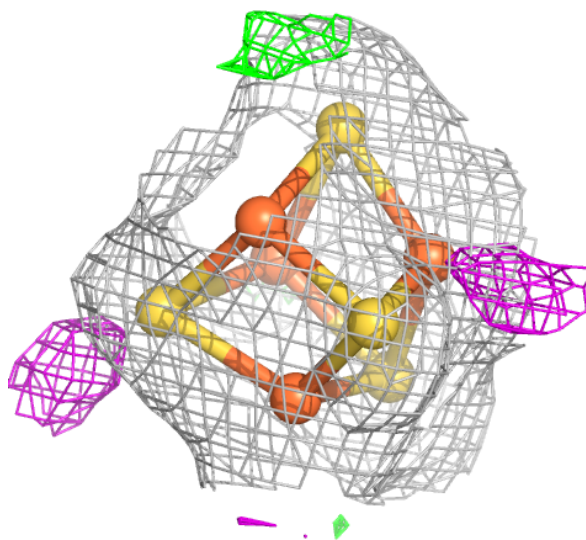
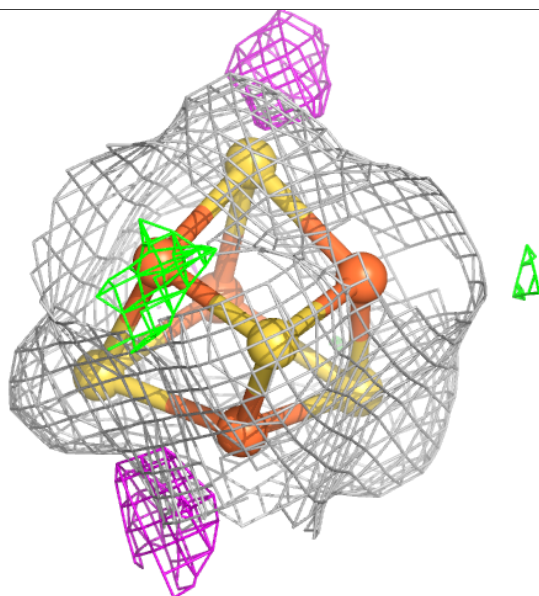
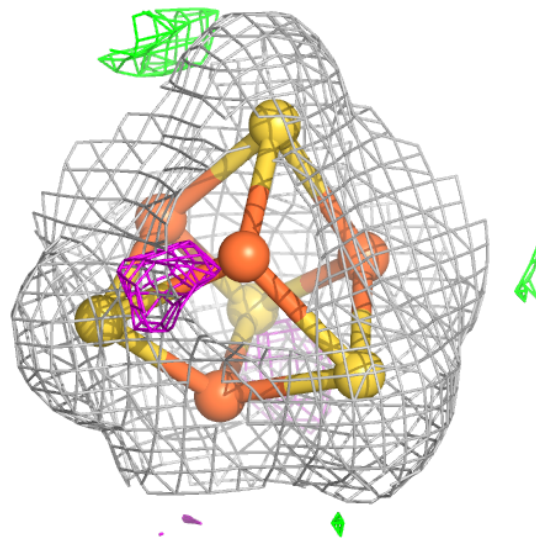






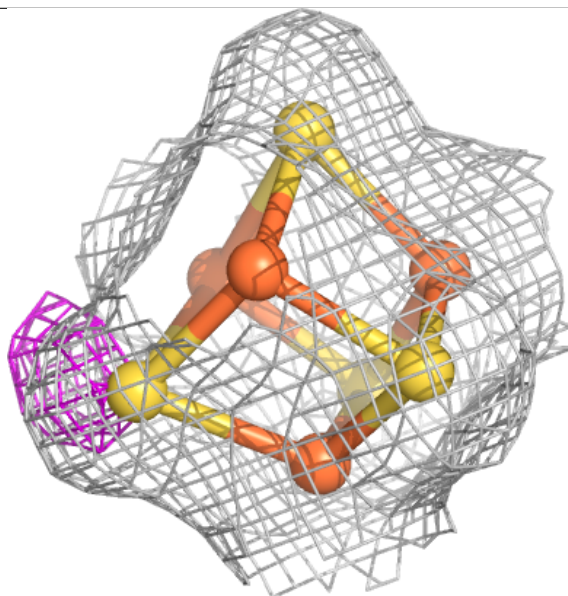
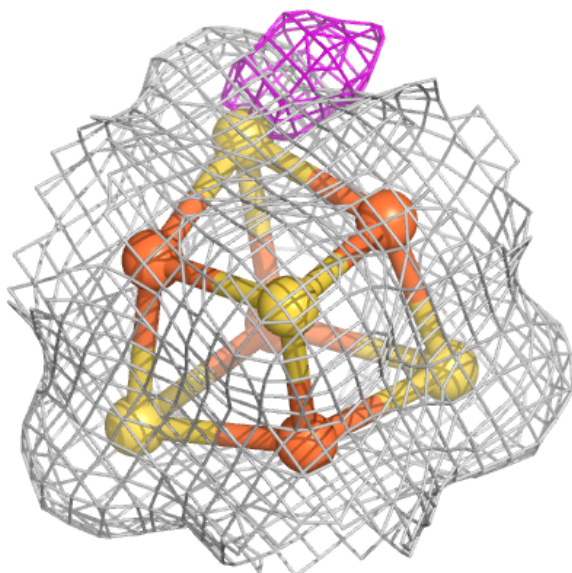
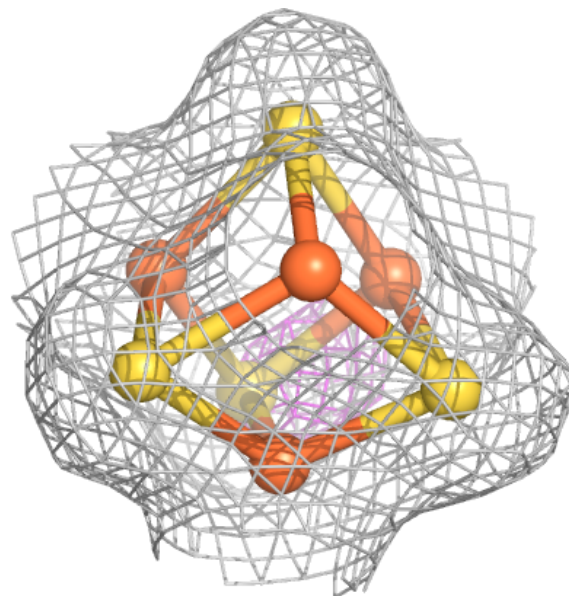
Electron density around SF4 B 302:

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



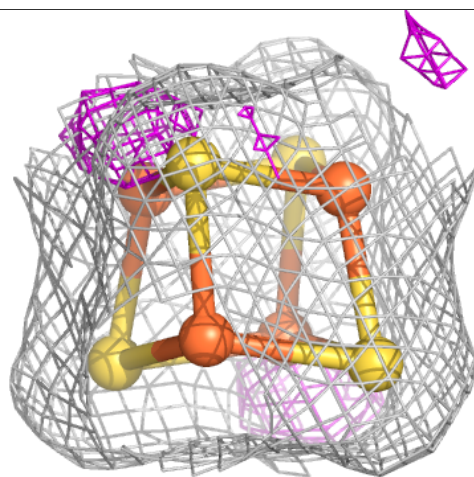
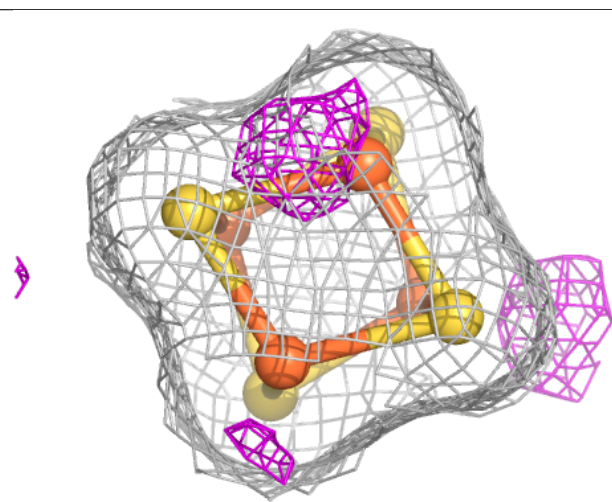
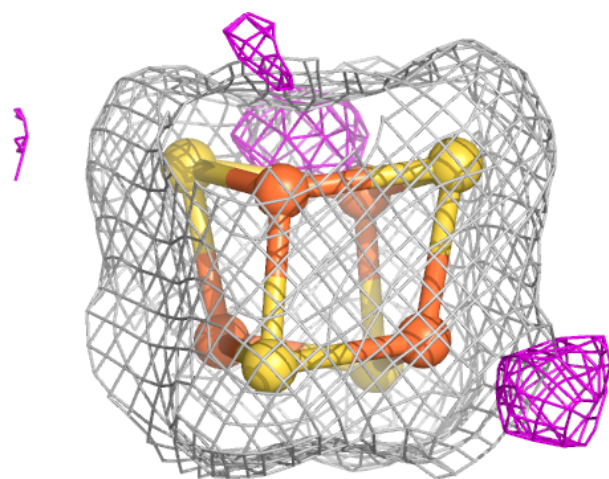
Electron density around SF4 A 1103:

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



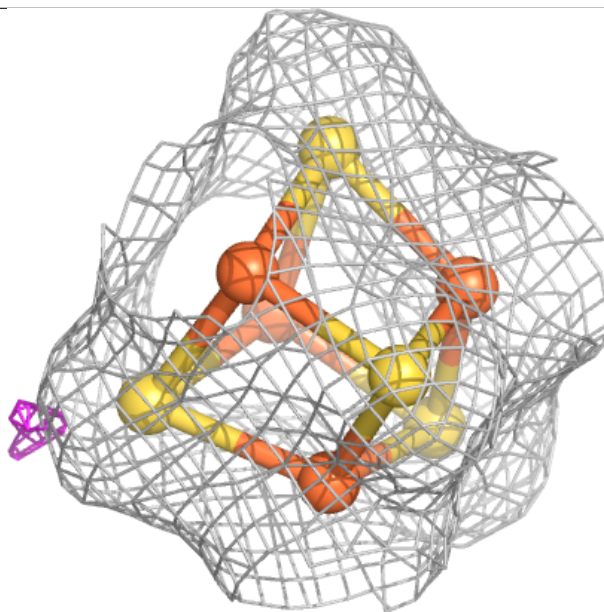
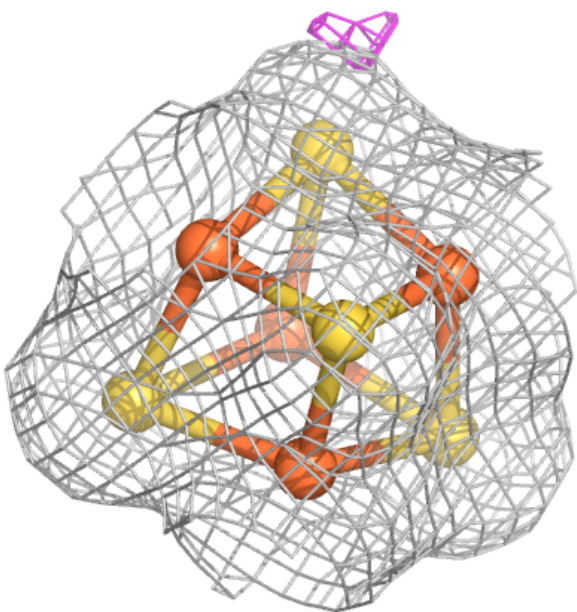
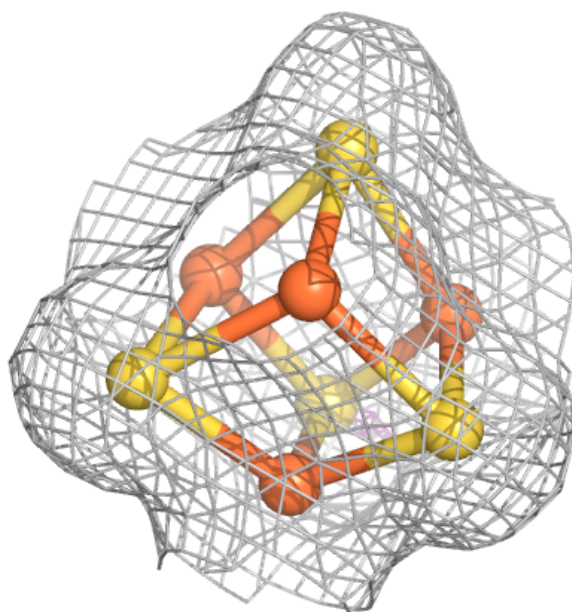
Electron density around SF4 B 303:

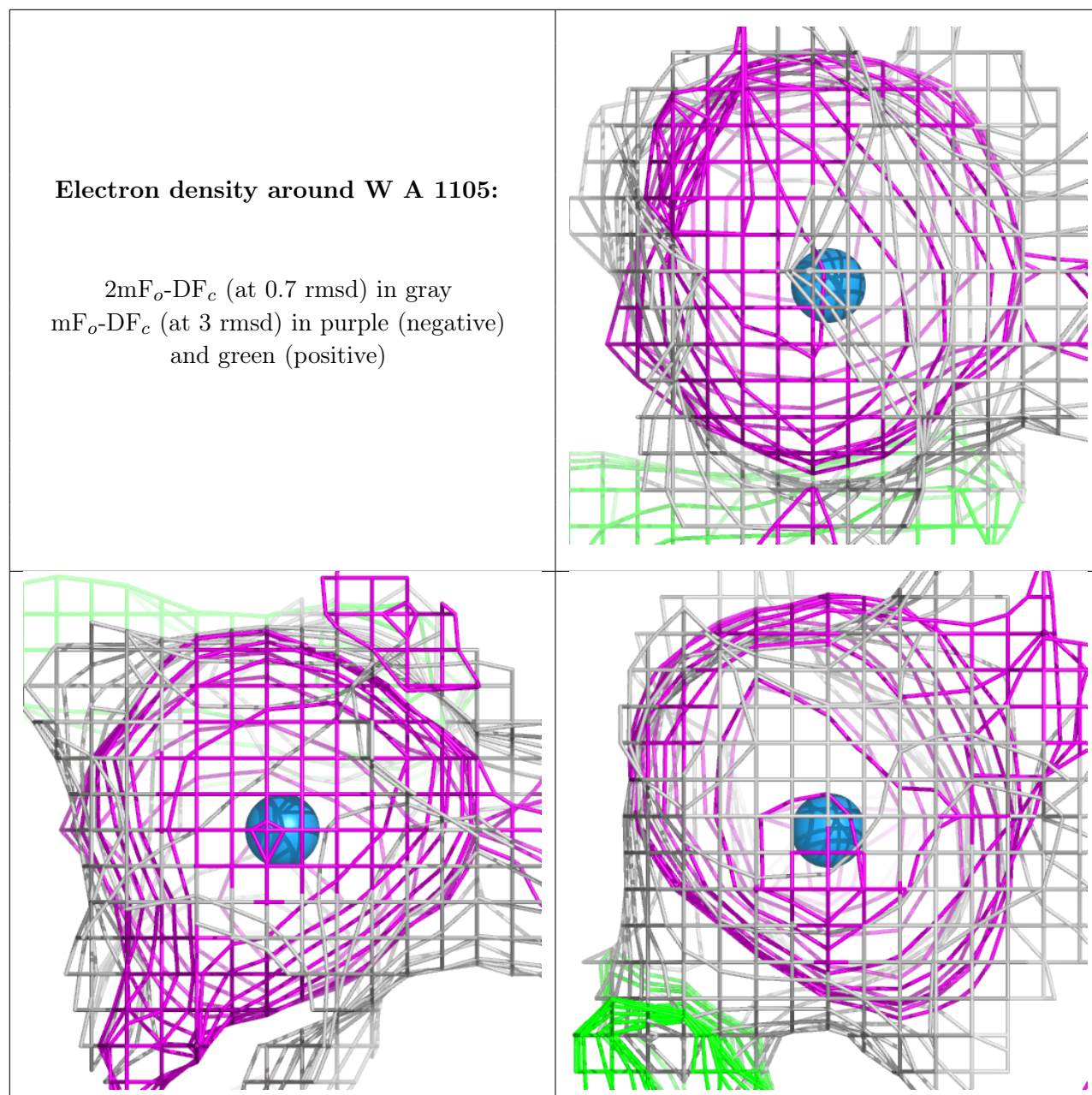
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around SF4 B 301:

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





6.5 Other polymers ⓘ

There are no such residues in this entry.