



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2020 – 05:41 PM BST

PDB ID : 6LYF
Title : Crystal structure of the mouse endonuclease EndoG(H138A/Se-Met)
Authors : Park, K.H.; Woo, E.J.
Deposited on : 2020-02-14
Resolution : 2.80 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 : 2.13
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) : 2.13

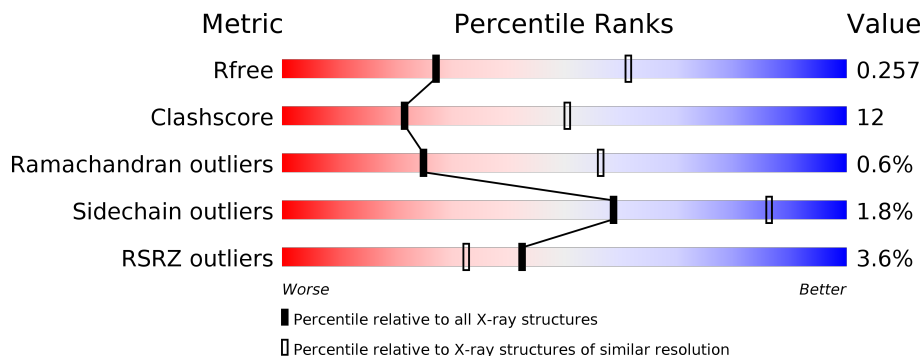
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 $\leq 5\%$. 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	249	 2% 70% 17% 12%
1	B	249	 7% 51% 32% 12%
1	C	249	 71% 16% 12%
1	D	249	 2% 68% 19% 10%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7194 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 Endonuclease G, mitochondrial.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	218	1745	1101	319	321	2	2	0	0	0
1	B	219	1746	1101	320	321	2	2	0	0	1
1	C	219	1745	1100	321	320	2	2	0	0	1
1	D	223	1774	1119	325	326	2	2	0	0	1

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	138	ALA	HIS	engineered mutation	UNP O08600
B	138	ALA	HIS	engineered mutation	UNP O08600
C	138	ALA	HIS	engineered mutation	UNP O08600
D	138	ALA	HIS	engineered mutation	UNP O08600

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mg	0	0
			1	1		
2	A	1	Total	Mg	0	0
			1	1		
2	D	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

- Molecule 3 is water.

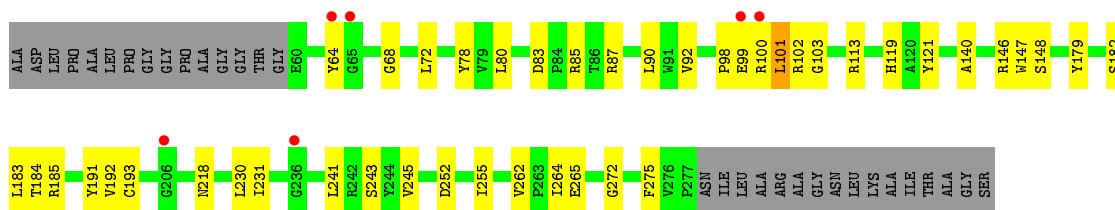
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	39	Total 39	O 39	0	0
3	B	34	Total 34	O 34	0	0
3	C	61	Total 61	O 61	0	0
3	D	46	Total 46	O 46	0	0

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.

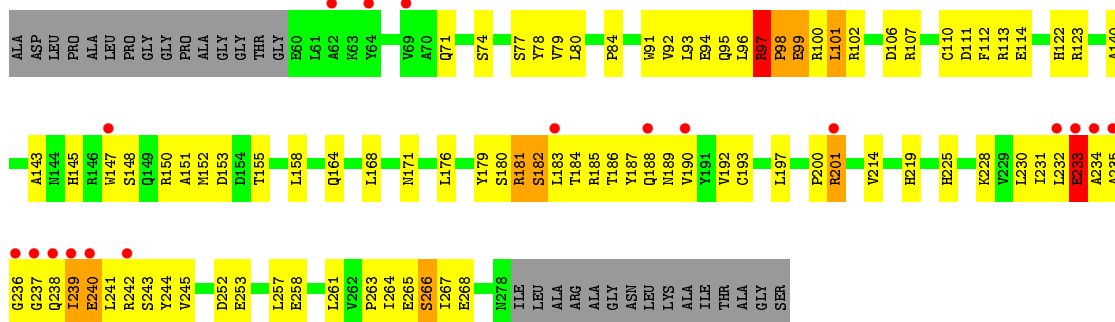
- Molecule 1: Endonuclease G, mitochondrial

Chain A: 



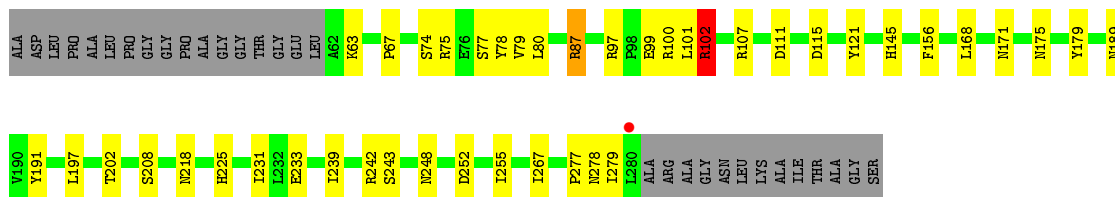
- Molecule 1: Endonuclease G, mitochondrial

Chain B: 

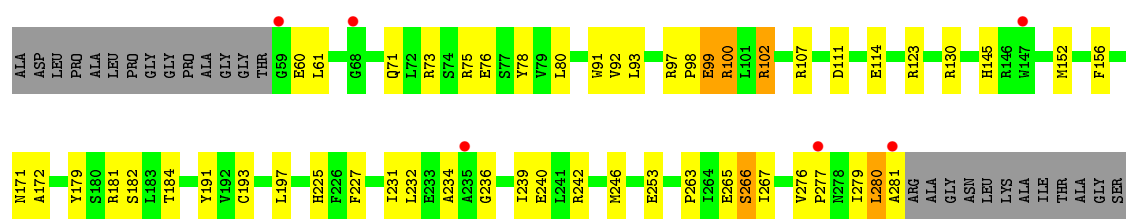


- Molecule 1: Endonuclease G, mitochondrial

Chain C: 



- Molecule 1: Endonuclease G, mitochondrial



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	72.20Å 81.88Å 88.66Å 90.00° 97.59° 90.00°	Depositor
Resolution (Å)	48.12 – 2.80 48.12 – 2.67	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.12-2.80) 94.5 (48.12-2.67)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.08 (at 2.69Å)	Xtrriage
Refinement program	PHENIX 1.16_3549	Depositor
R, R_{free}	0.199 , 0.257 0.199 , 0.257	Depositor DCC
R_{free} test set	1398 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	36.1	Xtrriage
Anisotropy	0.151	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 47.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7194	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.48% 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: MG

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	0.27	0/1787	0.47	0/2426
1	B	0.55	4/1788 (0.2%)	0.86	9/2428 (0.4%)
1	C	0.29	0/1787	0.58	3/2427 (0.1%)
1	D	0.36	1/1816 (0.1%)	0.60	2/2466 (0.1%)
All	All	0.38	5/7178 (0.1%)	0.64	14/9747 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	5
1	C	0	1
1	D	0	1
All	All	0	8

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	110	CYS	CB-SG	-8.32	1.68	1.82
1	B	96	LEU	C-N	-6.82	1.18	1.34
1	D	76	GLU	CG-CD	-6.16	1.42	1.51
1	B	201	ARG	CG-CD	5.75	1.66	1.51
1	B	233	GLU	CD-OE2	-5.47	1.19	1.25

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	242	ARG	NE-CZ-NH2	-10.85	114.87	120.30
1	C	102	ARG	CG-CD-NE	9.67	132.11	111.80
1	B	242	ARG	NE-CZ-NH1	8.97	124.79	120.30
1	B	97	ARG	NE-CZ-NH2	-8.29	116.16	120.30
1	B	242	ARG	CD-NE-CZ	7.73	134.42	123.60

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	101	LEU	Peptide
1	B	200	PRO	Peptide
1	B	233	GLU	Peptide
1	B	237	GLY	Peptide
1	B	99	GLU	Peptide

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	1745	0	1698	28	0
1	B	1746	0	1697	77	1
1	C	1745	0	1698	25	1
1	D	1774	0	1729	42	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	39	0	0	2	0
3	B	34	0	0	2	0
3	C	61	0	0	0	0
3	D	46	0	0	1	0
All	All	7194	0	6822	165	1

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 12.

The worst 5 of 165 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:99:GLU:OE1	1:D:102:ARG:HG3	1.31	1.29
1:D:99:GLU:OE1	1:D:102:ARG:CG	2.04	1.05
1:C:107:ARG:HE	1:C:111:ASP:HB3	1.38	0.89
1:A:101:LEU:HA	1:A:147:TRP:HB2	1.66	0.77
1:B:78:TYR:HE1	1:B:80:LEU:HD23	1.48	0.76

All (1) 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:B:252:ASP:OD1	1:C:102:ARG:NH2[1_465]	2.12	0.08

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	216/249 (87%)	207 (96%)	9 (4%)	0	100	100
1	B	217/249 (87%)	202 (93%)	13 (6%)	2 (1%)	17	46
1	C	217/249 (87%)	207 (95%)	8 (4%)	2 (1%)	17	46
1	D	221/249 (89%)	211 (96%)	9 (4%)	1 (0%)	29	61
All	All	871/996 (87%)	827 (95%)	39 (4%)	5 (1%)	25	56

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	240	GLU
1	C	279	ILE
1	D	280	LEU
1	B	98	PRO
1	C	277	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	183/198 (92%)	181 (99%)	2 (1%)	73	92
1	B	183/198 (92%)	178 (97%)	5 (3%)	44	78
1	C	183/198 (92%)	180 (98%)	3 (2%)	62	88
1	D	186/198 (94%)	183 (98%)	3 (2%)	62	88
All	All	735/792 (93%)	722 (98%)	13 (2%)	59	86

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	182	SER
1	B	266	SER
1	D	102	ARG
1	B	181	ARG
1	C	102	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	119	HIS
1	B	145	HIS
1	C	71	GLN
1	C	145	HIS
1	D	95	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](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	96:LEU	C	97:ARG	N	1.18

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	216/249 (86%)	-0.15	6 (2%) 53 43	24, 44, 91, 135	0
1	B	217/249 (87%)	0.30	18 (8%) 11 6	30, 60, 126, 142	0
1	C	217/249 (87%)	-0.35	1 (0%) 91 88	18, 32, 71, 109	0
1	D	221/249 (88%)	-0.21	6 (2%) 54 44	18, 37, 95, 127	0
All	All	871/996 (87%)	-0.10	31 (3%) 42 32	18, 42, 101, 142	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	234	ALA	6.9
1	B	242	ARG	4.1
1	B	240	GLU	4.0
1	D	277	PRO	3.6
1	C	280	LEU	3.5

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](#)

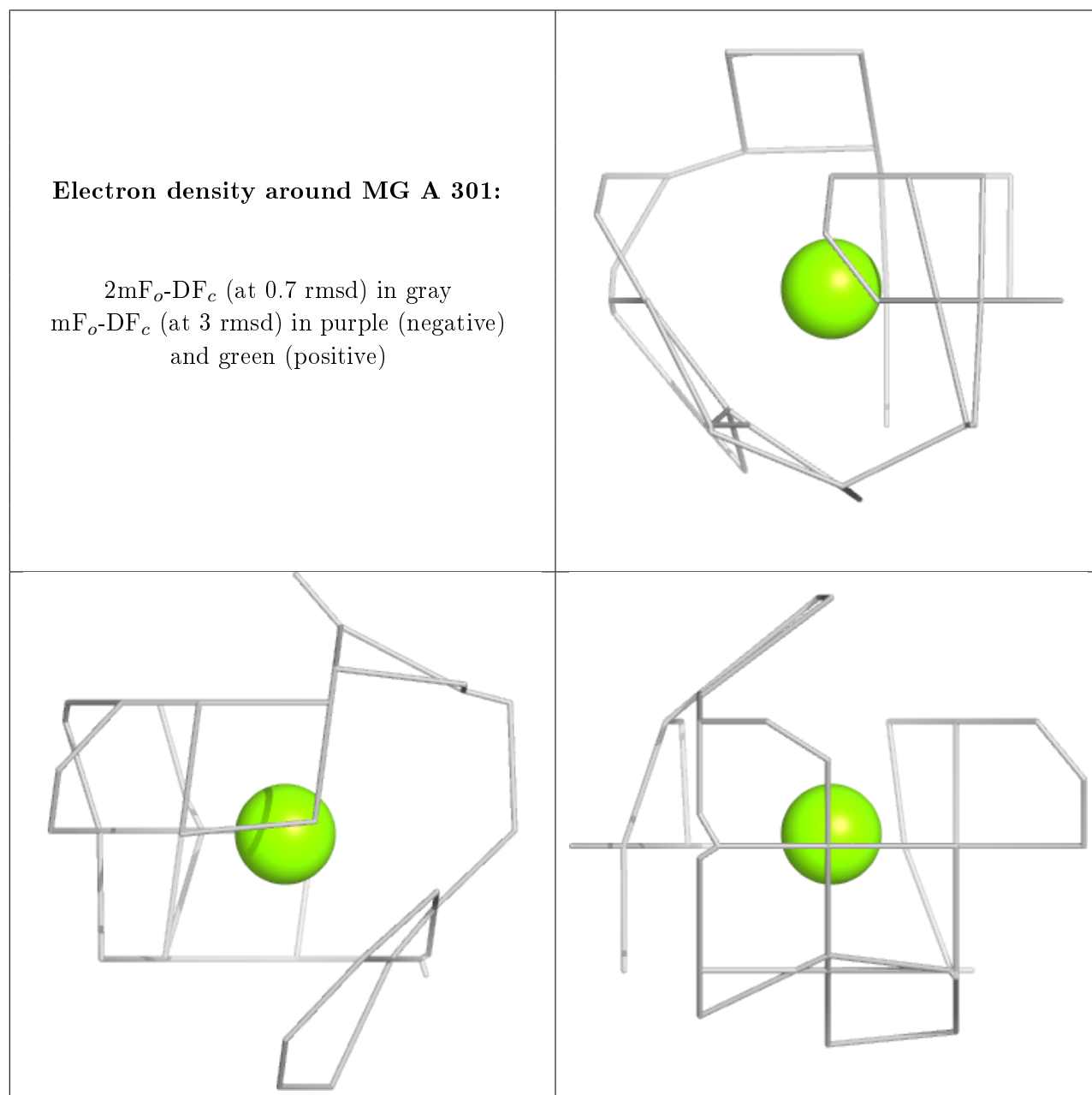
There are no monosaccharides in this entry.

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.

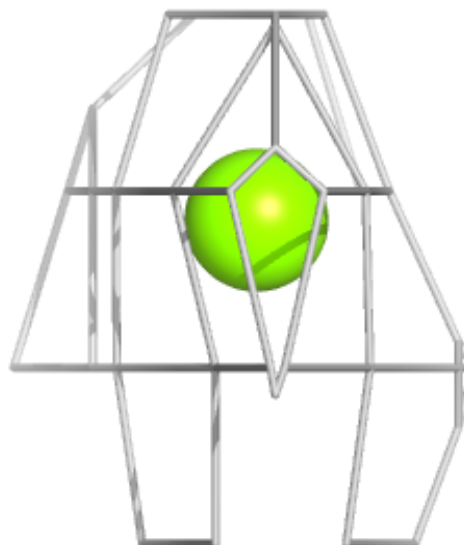
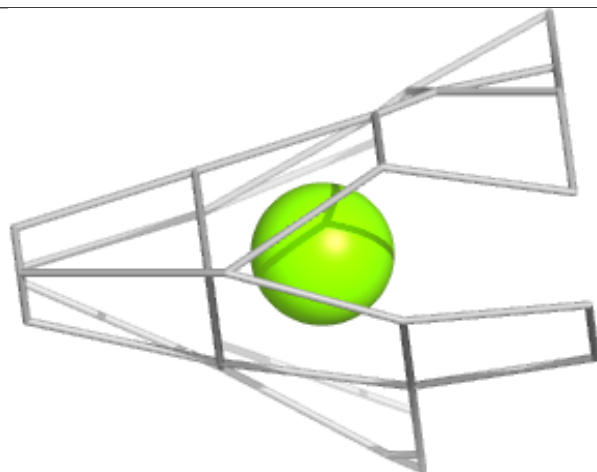
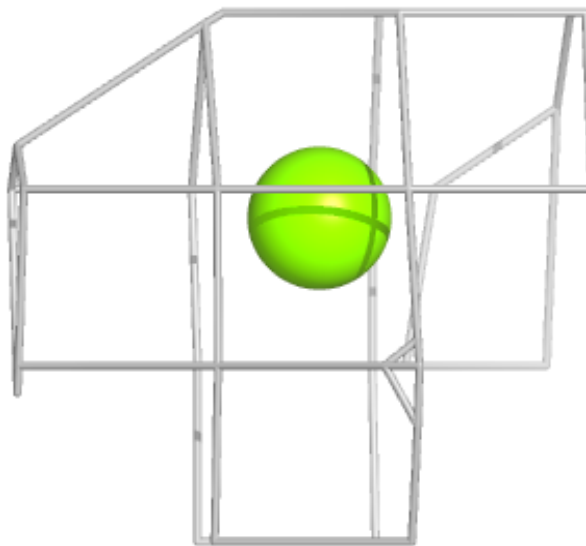
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	MG	A	301	1/1	0.85	0.14	69,69,69,69	0
2	MG	D	301	1/1	0.87	0.20	56,56,56,56	0
2	MG	C	301	1/1	0.91	0.14	58,58,58,58	0
2	MG	B	301	1/1	0.97	0.13	73,73,73,73	0

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.



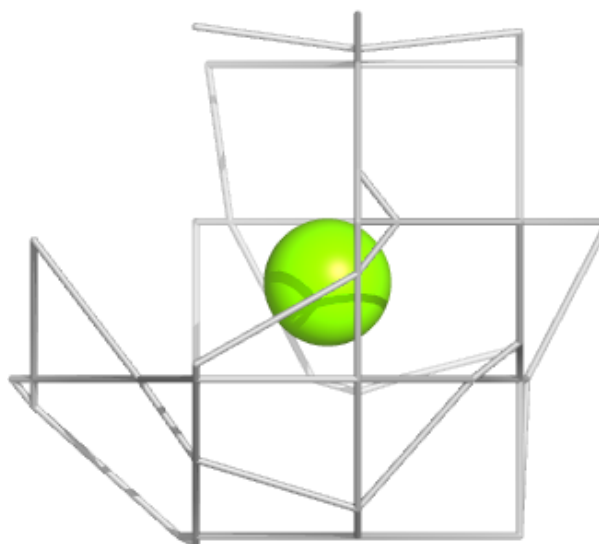
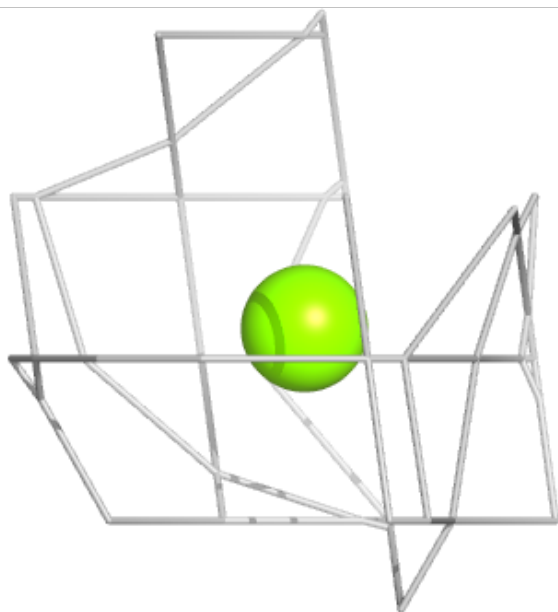
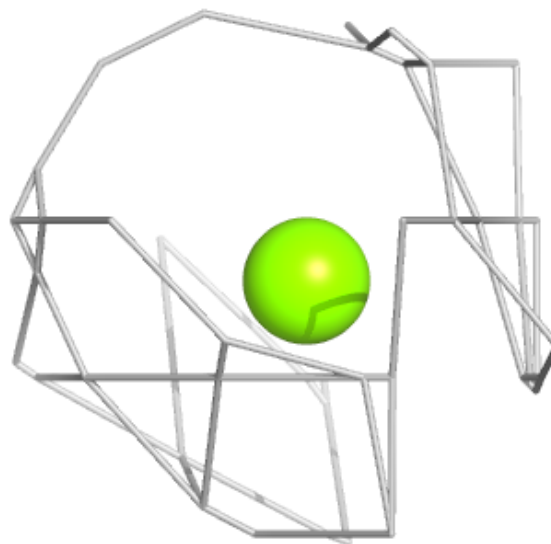
Electron density around MG D 301:

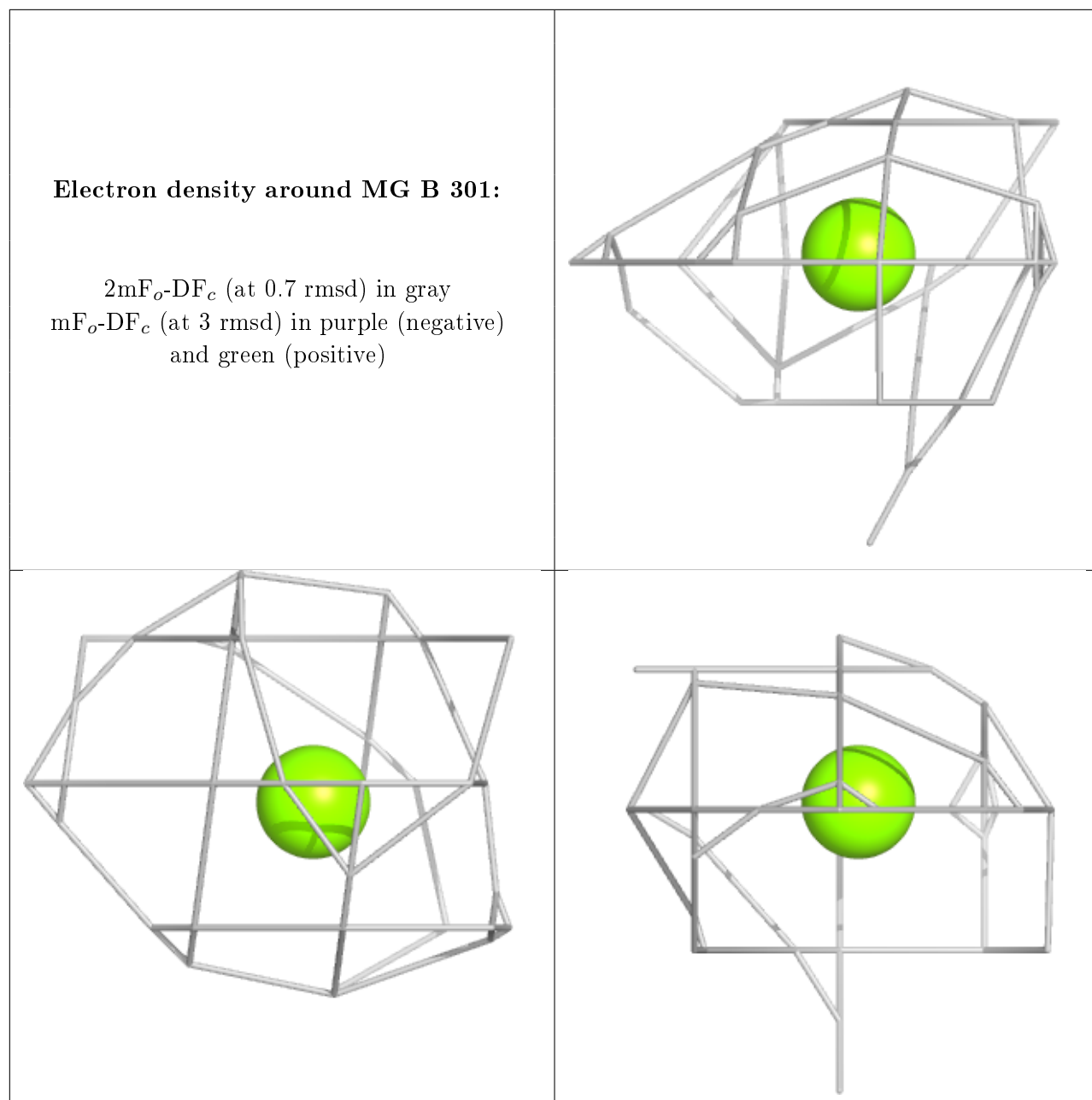
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around MG C 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 [i](#)

There are no such residues in this entry.