wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 29, 2024 – 12:20 PM JST

PDB ID : 7YPD
Title : Discovery and characterization of a new carbonyl reductase from Rhodotorula toluroides reducing fluoroketones, and X-ray analysis of the variant by rational engineering
Authors : Watanabe, Y.; Asano, Y.; Hibi, M.
Deposited on : 2022-08-03
Resolution : 1.27 Å(reported)

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

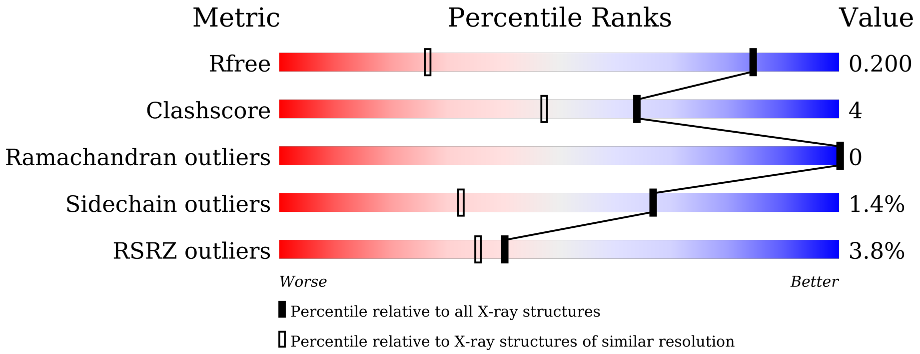
MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

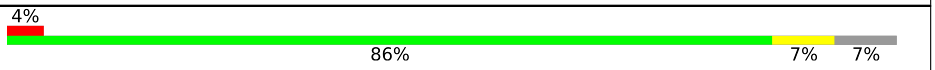
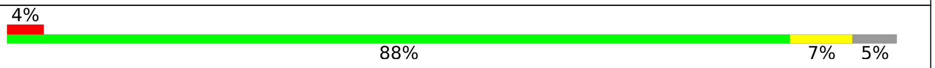
The reported resolution of this entry is 1.27 Å.

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	1850 (1.30-1.26)
Clashscore	141614	1926 (1.30-1.26)
Ramachandran outliers	138981	1860 (1.30-1.26)
Sidechain outliers	138945	1859 (1.30-1.26)
RSRZ outliers	127900	1807 (1.30-1.26)

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	256	
1	B	256	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4258 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 Carbonyl reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	243	Total 1852	C 1188	N 309	O 352	S 3	0	8	0
1	A	237	Total 1844	C 1179	N 310	O 350	S 5	0	11	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	54	ALA	GLU	conflict	UNP A0A061BJP9
B	57	ALA	LYS	conflict	UNP A0A061BJP9
B	58	ALA	LYS	conflict	UNP A0A061BJP9
B	166A	ALA	-	insertion	UNP A0A061BJP9
A	54	ALA	GLU	conflict	UNP A0A061BJP9
A	57	ALA	LYS	conflict	UNP A0A061BJP9
A	58	ALA	LYS	conflict	UNP A0A061BJP9
A	166A	ALA	-	insertion	UNP A0A061BJP9

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

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	315	Total 315	O 315	0	0

Continued on next page...

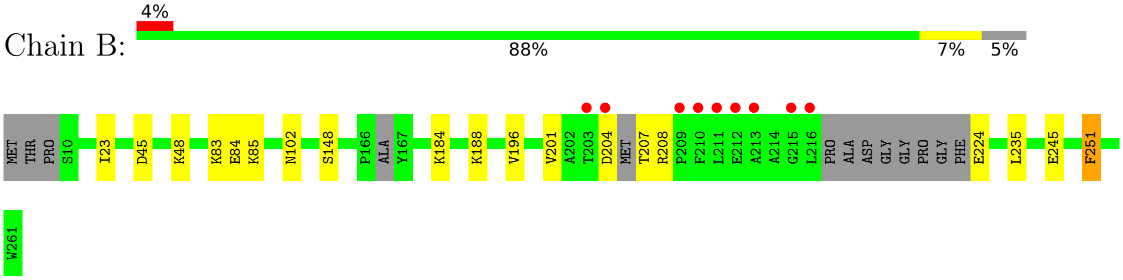
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	244	Total 244	O 244	0	0

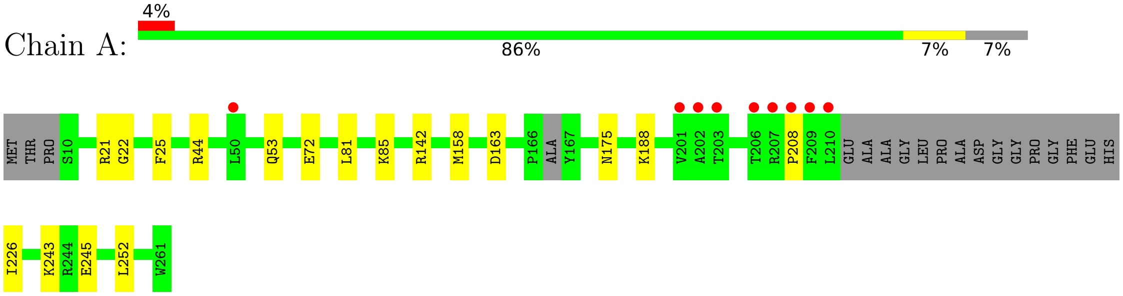
3 Residue-property plots

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: Carbonyl reductase



- Molecule 1: Carbonyl reductase



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.36Å 65.08Å 79.77Å 90.00° 99.83° 90.00°	Depositor
Resolution (Å)	40.56 – 1.27 40.56 – 1.27	Depositor EDS
% Data completeness (in resolution range)	92.3 (40.56-1.27) 92.3 (40.56-1.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	1.10 (at 1.27Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.174 , 0.200 0.174 , 0.200	Depositor DCC
R_{free} test set	5446 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	14.8	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 41.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4258	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.46% 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 (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	0.34	0/1884	0.62	0/2560
1	B	0.36	0/1888	0.63	0/2567
All	All	0.35	0/3772	0.62	0/5127

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	1844	0	1853	17	0
1	B	1852	0	1866	14	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
3	A	244	0	0	12	3
3	B	315	0	0	9	3
All	All	4258	0	3719	31	4

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.

The worst 5 of 31 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:A:142:ARG:NH1	3:A:402:HOH:O	2.07	0.88
1:A:81:LEU:HD11	1:A:85:LYS:HE2	1.55	0.87
1:A:22:GLY:O	3:A:401:HOH:O	2.05	0.75
1:A:188:LYS:NZ	3:A:406:HOH:O	2.23	0.71
1:A:158[B]:MET:HG3	3:A:410:HOH:O	1.90	0.70

All (4) 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)
3:B:676:HOH:O	3:A:625:HOH:O[2_446]	2.10	0.10
3:B:591:HOH:O	3:B:698:HOH:O[2_556]	2.18	0.02
3:B:663:HOH:O	3:A:527:HOH:O[2_446]	2.19	0.01
3:A:560:HOH:O	3:A:606:HOH:O[2_545]	2.19	0.01

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	245/256 (96%)	242 (99%)	3 (1%)	0	100	100
1	B	245/256 (96%)	241 (98%)	4 (2%)	0	100	100
All	All	490/512 (96%)	483 (99%)	7 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	196/195 (100%)	194 (99%)	2 (1%)	76	46
1	B	194/195 (100%)	191 (98%)	3 (2%)	65	30
All	All	390/390 (100%)	385 (99%)	5 (1%)	67	34

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	85	LYS
1	B	224	GLU
1	B	251	PHE
1	A	44	ARG
1	A	53	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	36	ASN
1	B	102	ASN
1	A	53	GLN
1	A	95	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 3 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 ⓘ

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, 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	237/256 (92%)	-0.03	9 (3%)	40	35	12, 20, 39, 54	0
1	B	243/256 (94%)	-0.11	9 (3%)	41	36	11, 16, 47, 81	0
All	All	480/512 (93%)	-0.07	18 (3%)	40	35	11, 18, 42, 81	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	210	PHE	5.9
1	B	216	LEU	5.2
1	B	203	THR	5.1
1	A	209	PHE	5.0
1	A	210	LEU	4.9

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

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

6.3 Carbohydrates ⓘ

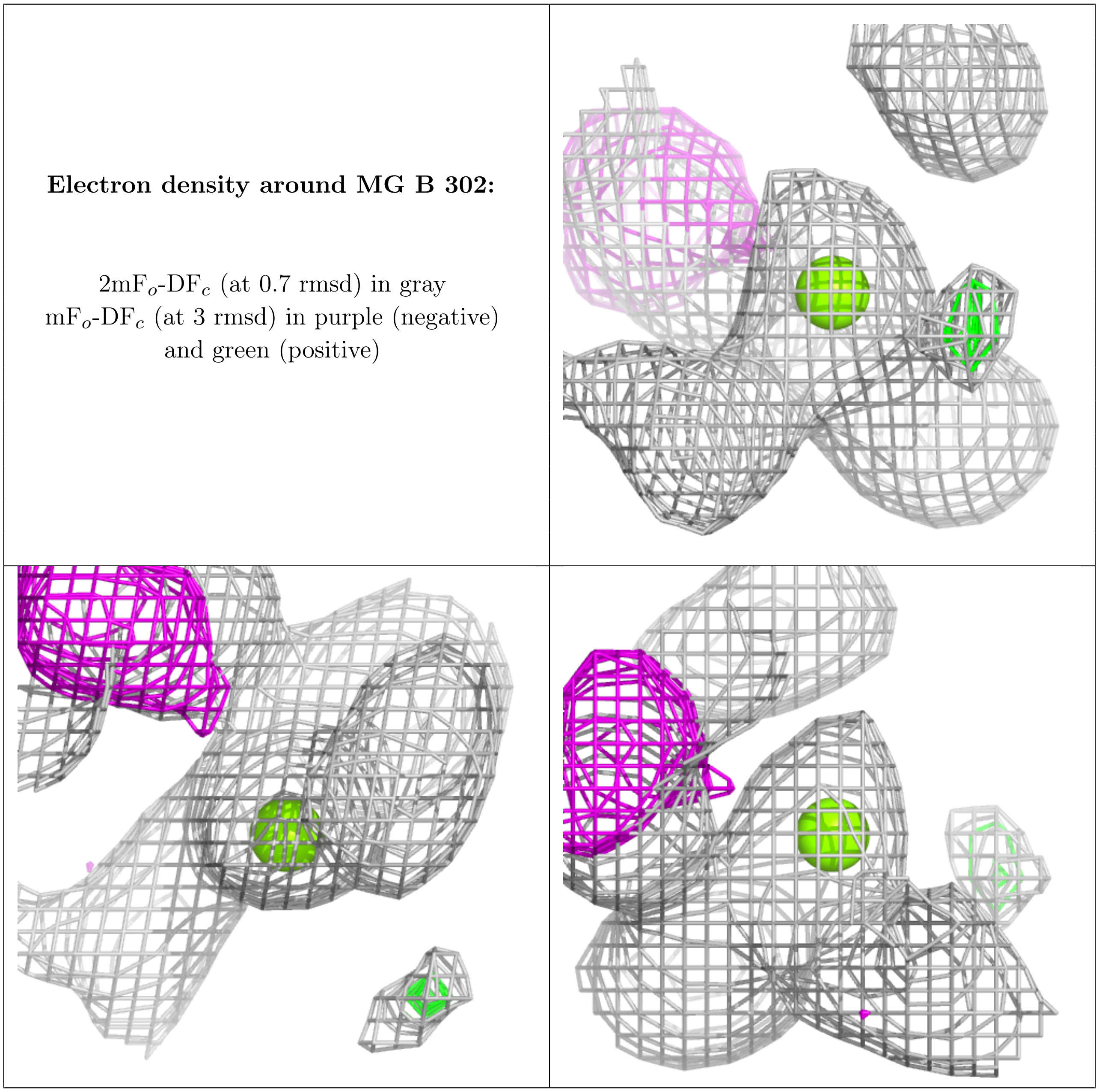
There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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.

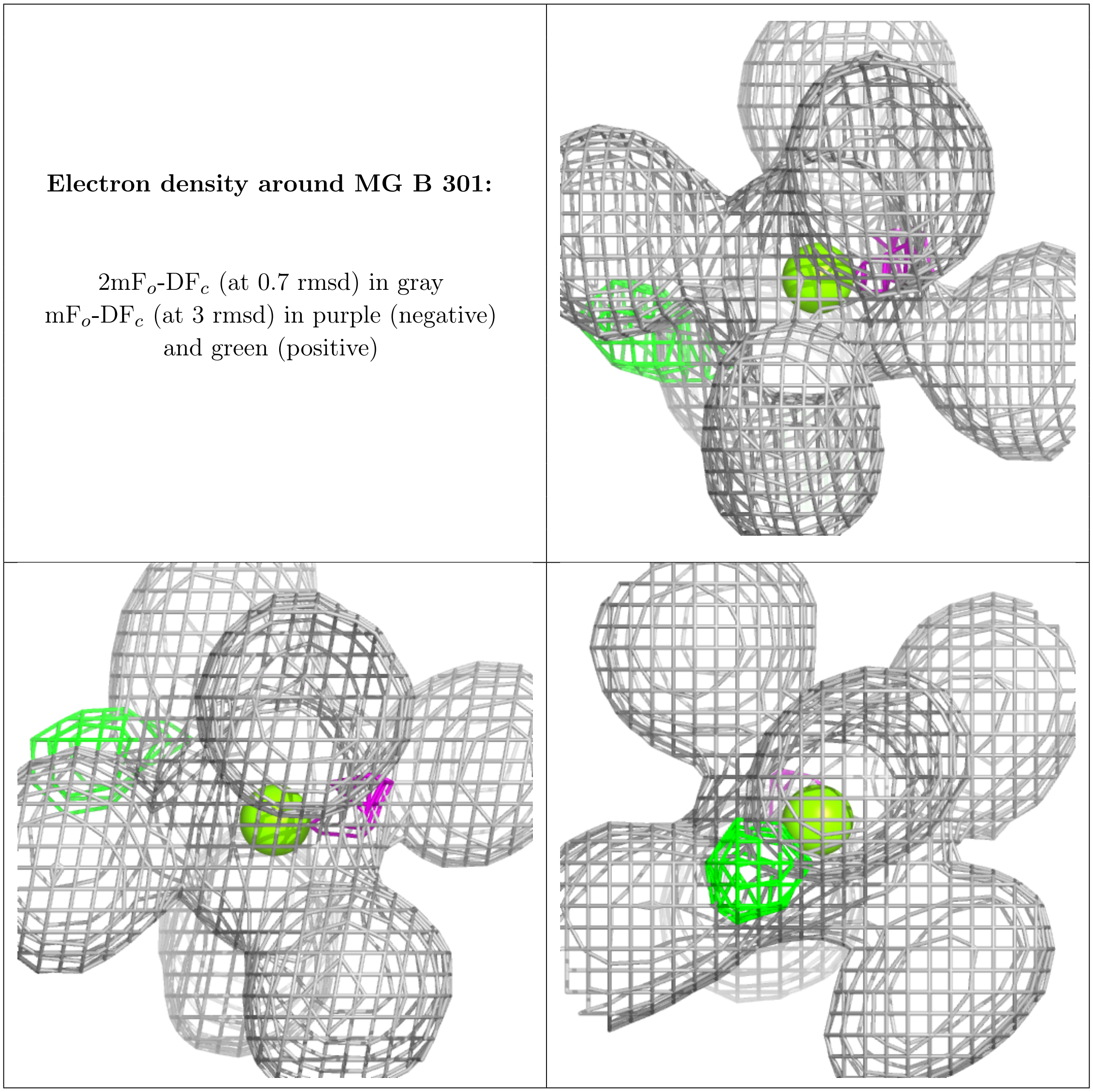
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	MG	B	302	1/1	0.93	0.09	48,48,48,48	0
2	MG	B	301	1/1	0.98	0.03	20,20,20,20	0
2	MG	A	301	1/1	0.99	0.15	33,33,33,33	0

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.



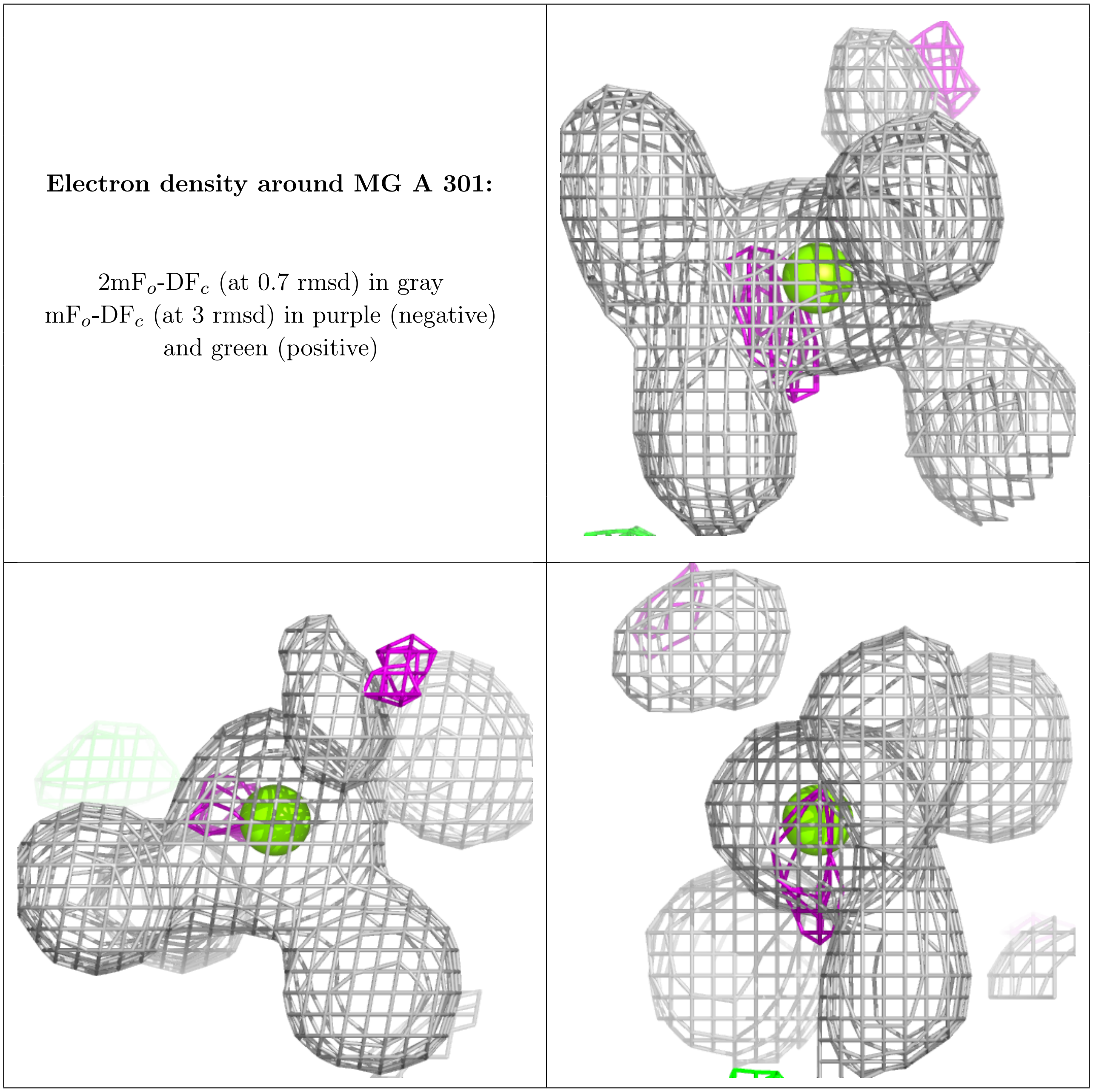
Electron density around MG B 302:

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 MG 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)



Electron density around MG A 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.