Full wwPDB X-ray Structure Validation Report ⓘ

Oct 15, 2023 – 03:23 AM EDT

PDB ID	:	8D71
Title	:	Human Ago2 bound to miR122(21nt)
Authors	:	Xiao, Y.; MacRae, I.
Deposited on	:	2022-06-07
Resolution	:	2.50 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:

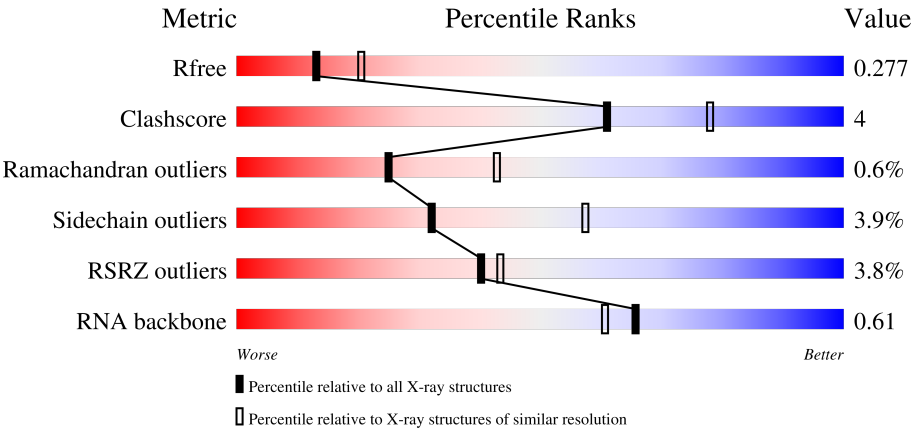
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)
RNA backbone	3102	1008 (2.84-2.16)

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	859	4% 81% 12% • 6%
2	B	21	43% 10% 48%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13443 atoms, of which 6663 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 Protein argonaute-2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	810	13054	4132	6557	1171	1153	41	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	387	ASP	SER	conflict	UNP Q9UKV8
A	824	ALA	SER	conflict	UNP Q9UKV8
A	828	ASP	SER	conflict	UNP Q9UKV8
A	831	ASP	SER	conflict	UNP Q9UKV8
A	834	ALA	SER	conflict	UNP Q9UKV8

- Molecule 2 is a RNA chain called miR-122-21nt.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
2	B	11	325	95	106	35	78	11	0	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
3	A	1	1	1	0	0

- Molecule 4 is water.

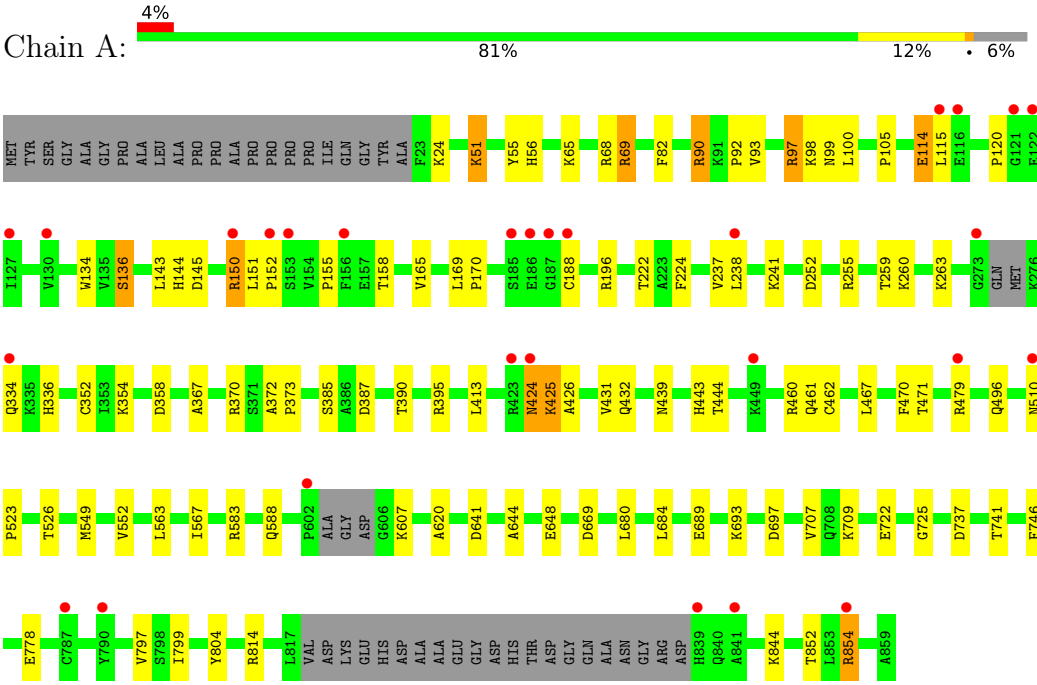
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	60	60	60	0	0
4	B	3	Total	O	0	0
			3	3		

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.

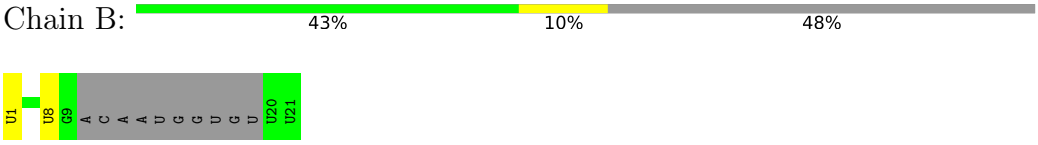
- Molecule 1: Protein argonaute-2

Chain A:



- Molecule 2: miR-122-21nt

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	63.13Å 107.73Å 68.82Å 90.00° 106.89° 90.00°	Depositor
Resolution (Å)	39.20 – 2.50 39.20 – 1.86	Depositor EDS
% Data completeness (in resolution range)	98.8 (39.20-2.50) 61.3 (39.20-1.86)	Depositor EDS
R_{merge}	0.02	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.80 (at 1.85Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.223 , 0.280 0.222 , 0.277	Depositor DCC
R_{free} test set	2000 reflections (4.06%)	wwPDB-VP
Wilson B-factor (Å^2)	38.7	Xtriage
Anisotropy	0.330	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 45.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	13443	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% 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: 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	0.37	1/6650 (0.0%)	0.53	0/8997
2	B	0.88	1/242 (0.4%)	0.75	0/372
All	All	0.40	2/6892 (0.0%)	0.54	0/9369

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	U	OP3-P	-10.28	1.48	1.61
1	A	114	GLU	CB-CG	-5.88	1.41	1.52

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	6497	6557	6557	60	3
2	B	219	106	106	1	0
3	A	1	0	0	0	0
4	A	60	0	0	12	0
4	B	3	0	0	0	0
All	All	6780	6663	6663	60	3

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 (60) 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:725:GLY:O	4:A:1001:HOH:O	1.89	0.91
1:A:709:LYS:NZ	4:A:1006:HOH:O	2.04	0.88
1:A:413:LEU:O	4:A:1002:HOH:O	1.93	0.85
1:A:697:ASP:OD2	4:A:1004:HOH:O	2.00	0.79
1:A:90:ARG:HA	1:A:90:ARG:HE	1.47	0.79
1:A:778:GLU:OE1	4:A:1005:HOH:O	2.04	0.75
1:A:641:ASP:OD1	4:A:1007:HOH:O	2.04	0.75
1:A:260:LYS:O	1:A:354:LYS:NZ	2.20	0.74
1:A:424:ASN:O	1:A:426:ALA:N	2.21	0.73
1:A:689:GLU:OE2	4:A:1008:HOH:O	2.07	0.72
1:A:82:PHE:CE1	1:A:115:LEU:HD22	2.25	0.71
1:A:722:GLU:OE1	4:A:1010:HOH:O	2.09	0.71
1:A:56:HIS:NE2	1:A:99:ASN:OD1	2.18	0.70
1:A:51:LYS:HZ1	1:A:144:HIS:HB2	1.59	0.67
1:A:90:ARG:HA	1:A:90:ARG:NE	2.11	0.66
1:A:367:ALA:O	1:A:370:ARG:NH2	2.27	0.64
1:A:439:ASN:N	4:A:1009:HOH:O	2.08	0.63
1:A:334:GLN:N	1:A:334:GLN:OE1	2.35	0.60
1:A:443:HIS:ND1	1:A:444:THR:HG23	2.19	0.57
1:A:462:CYS:SG	1:A:552:VAL:HG11	2.45	0.57
1:A:517:LEU:HD12	1:A:518:VAL:N	2.19	0.57
1:A:69:ARG:N	1:A:69:ARG:HD3	2.20	0.56
1:A:92:PRO:HB2	1:A:100:LEU:HD11	1.88	0.56
1:A:93:VAL:O	1:A:100:LEU:HD12	2.05	0.56
1:A:737:ASP:HA	1:A:741:THR:HG21	1.89	0.55
1:A:517:LEU:HD12	1:A:518:VAL:H	1.72	0.53
1:A:90:ARG:HE	1:A:90:ARG:CA	2.20	0.52
1:A:583:ARG:NH1	1:A:620:ALA:O	2.44	0.51
1:A:761:ARG:NH2	2:B:8:U:OP2	2.45	0.50
1:A:852:THR:C	4:A:1012:HOH:O	2.51	0.49
1:A:563:LEU:O	1:A:567:ILE:HG12	2.12	0.49
1:A:169:LEU:HB3	1:A:170:PRO:HD3	1.95	0.48
1:A:134:TRP:NE1	1:A:136:SER:O	2.39	0.47
1:A:470:PHE:CD2	1:A:552:VAL:HG22	2.49	0.47
1:A:145:ASP:O	1:A:150:ARG:HB3	2.16	0.46
1:A:432:GLN:N	4:A:1003:HOH:O	1.98	0.45
1:A:680:LEU:HD23	1:A:680:LEU:O	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:797:VAL:HG23	1:A:799:ILE:O	2.18	0.43
1:A:255:ARG:NH1	1:A:284:VAL:O	2.50	0.43
1:A:155:PRO:HB2	1:A:158:THR:HG22	2.00	0.43
1:A:68:ARG:HD3	1:A:97:ARG:NH1	2.34	0.43
1:A:854:ARG:N	4:A:1012:HOH:O	2.51	0.43
1:A:196:ARG:NH1	1:A:352:CYS:SG	2.86	0.43
1:A:372:ALA:HB3	1:A:373:PRO:HD3	2.01	0.42
1:A:424:ASN:O	1:A:425:LYS:C	2.56	0.42
1:A:222:THR:HG23	1:A:224:PHE:CE2	2.54	0.42
1:A:467:LEU:O	1:A:471:THR:OG1	2.32	0.42
1:A:237:VAL:HG12	1:A:238:LEU:N	2.35	0.41
1:A:680:LEU:HD12	1:A:768:LEU:HB3	2.01	0.41
1:A:55:TYR:CE1	1:A:105:PRO:HG3	2.55	0.41
1:A:680:LEU:O	1:A:684:LEU:HB3	2.20	0.41
1:A:431:VAL:HG12	1:A:432:GLN:HG3	2.03	0.41
1:A:93:VAL:HG21	1:A:165:VAL:HG22	2.02	0.41
1:A:143:LEU:HD12	1:A:158:THR:OG1	2.20	0.41
1:A:669:ASP:HA	1:A:707:VAL:HG13	2.03	0.41
1:A:259:THR:O	1:A:263:LYS:HB2	2.21	0.41
1:A:151:LEU:HD12	1:A:151:LEU:N	2.36	0.41
1:A:644:ALA:O	1:A:648:GLU:HG2	2.20	0.41
1:A:387:ASP:OD2	1:A:390:THR:OG1	2.30	0.40
1:A:461:GLN:OE1	1:A:461:GLN:N	2.53	0.40

All (3) 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:A:55:TYR:HH	1:A:588:GLN:HE21[2_646]	1.32	0.28
1:A:150:ARG:HH22	1:A:424:ASN:O[2_646]	1.46	0.14
1:A:98:LYS:NZ	1:A:510:ASN:OD1[2_645]	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	800/859 (93%)	752 (94%)	43 (5%)	5 (1%)	25 43

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	425	LYS
1	A	523	PRO
1	A	188	CYS
1	A	120	PRO
1	A	152	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	718/750 (96%)	690 (96%)	28 (4%)	32 57

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	24	LYS
1	A	51	LYS
1	A	65	LYS
1	A	69	ARG
1	A	90	ARG
1	A	97	ARG
1	A	114	GLU
1	A	136	SER
1	A	150	ARG
1	A	241	LYS
1	A	252	ASP
1	A	336	HIS
1	A	358	ASP
1	A	385	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	395	ARG
1	A	424	ASN
1	A	460	ARG
1	A	479	ARG
1	A	496	GLN
1	A	526	THR
1	A	549	MET
1	A	607	LYS
1	A	693	LYS
1	A	746	PHE
1	A	804	TYR
1	A	814	ARG
1	A	844	LYS
1	A	854	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	144	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	8/21 (38%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

Of 1 ligands modelled in this entry, 1 is 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, 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	810/859 (94%)	-0.01	31 (3%) 40 43	29, 59, 89, 104	0
2	B	11/21 (52%)	-0.26	0 100 100	41, 45, 84, 84	0
All	All	821/880 (93%)	-0.01	31 (3%) 40 43	29, 59, 89, 104	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	188	CYS	6.2
1	A	187	GLY	5.3
1	A	449	LYS	4.1
1	A	186	GLU	3.8
1	A	127	ILE	3.5
1	A	424	ASN	3.3
1	A	334	GLN	3.3
1	A	841	ALA	3.2
1	A	122	GLU	3.2
1	A	156	PHE	3.1
1	A	854	ARG	3.0
1	A	150	ARG	3.0
1	A	602	PRO	2.9
1	A	153	SER	2.9
1	A	124	LYS	2.9
1	A	510	ASN	2.9
1	A	238	LEU	2.8
1	A	273	GLY	2.7
1	A	423	ARG	2.7
1	A	787	CYS	2.6
1	A	790	TYR	2.6
1	A	130	VAL	2.4
1	A	125	ASP	2.4
1	A	479	ARG	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	839	HIS	2.2
1	A	115	LEU	2.2
1	A	121	GLY	2.2
1	A	116	GLU	2.1
1	A	185	SER	2.1
1	A	333	GLU	2.1
1	A	152	PRO	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)

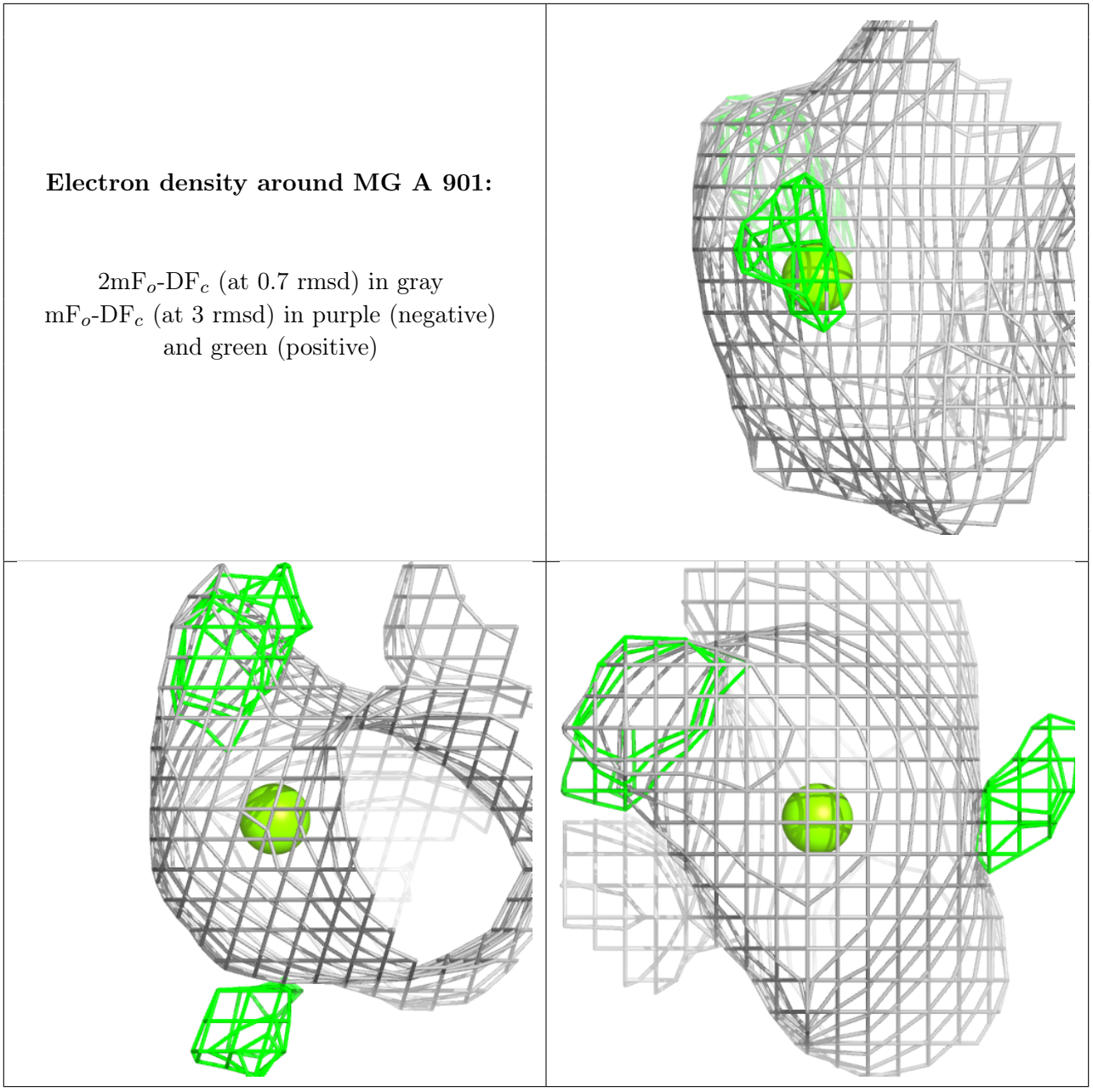
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	MG	A	901	1/1	0.92	0.23	49,49,49,49	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 A 901:

$2mF_o$-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.