wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2024 – 07:09 pm BST

PDB ID	:	8PPS
Title	:	Dimeric RbdA EAL, in apo state
Authors	:	Cordery, C.R.; Maly, M.; Walsh, M.A.; Tews, I.
Deposited on	:	2023-07-08
Resolution	:	2.30 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 (i) 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 (i)) were used in the production of this report:

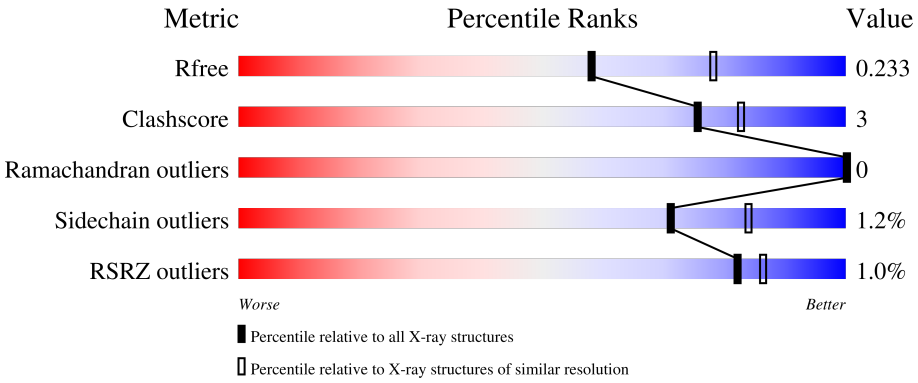
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	269	% 83% 10% 7%
1	B	269	% 87% 6% 6%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4235 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 EAL domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	S	0	2	0
			1994	1276	340	367	11			
1	B	252	Total	C	N	O	S	0	2	0
			2003	1281	342	369	11			

There are 40 discrepancies between the modelled and reference sequences:

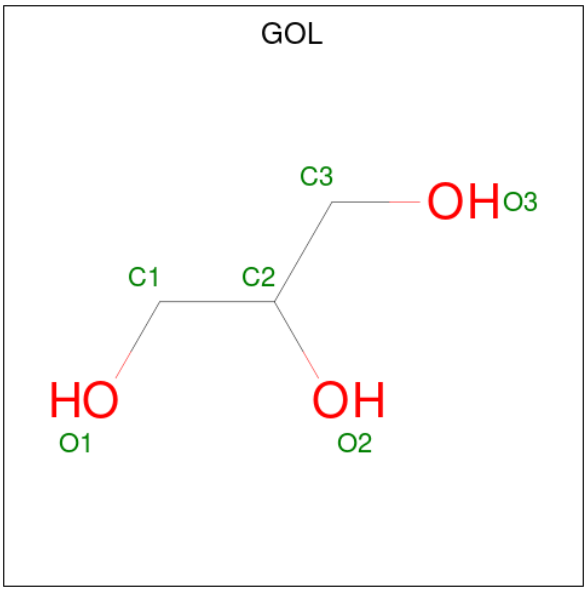
Chain	Residue	Modelled	Actual	Comment	Reference
A	529	MET	-	initiating methionine	UNP Q9I580
A	530	GLY	-	expression tag	UNP Q9I580
A	531	SER	-	expression tag	UNP Q9I580
A	532	SER	-	expression tag	UNP Q9I580
A	533	HIS	-	expression tag	UNP Q9I580
A	534	HIS	-	expression tag	UNP Q9I580
A	535	HIS	-	expression tag	UNP Q9I580
A	536	HIS	-	expression tag	UNP Q9I580
A	537	HIS	-	expression tag	UNP Q9I580
A	538	HIS	-	expression tag	UNP Q9I580
A	539	SER	-	expression tag	UNP Q9I580
A	540	SER	-	expression tag	UNP Q9I580
A	541	GLY	-	expression tag	UNP Q9I580
A	542	LEU	-	expression tag	UNP Q9I580
A	543	VAL	-	expression tag	UNP Q9I580
A	544	PRO	-	expression tag	UNP Q9I580
A	545	ARG	-	expression tag	UNP Q9I580
A	546	GLY	-	expression tag	UNP Q9I580
A	547	SER	-	expression tag	UNP Q9I580
A	548	HIS	-	expression tag	UNP Q9I580
B	529	MET	-	initiating methionine	UNP Q9I580
B	530	GLY	-	expression tag	UNP Q9I580
B	531	SER	-	expression tag	UNP Q9I580
B	532	SER	-	expression tag	UNP Q9I580
B	533	HIS	-	expression tag	UNP Q9I580

Continued on next page...

Continued from previous page...

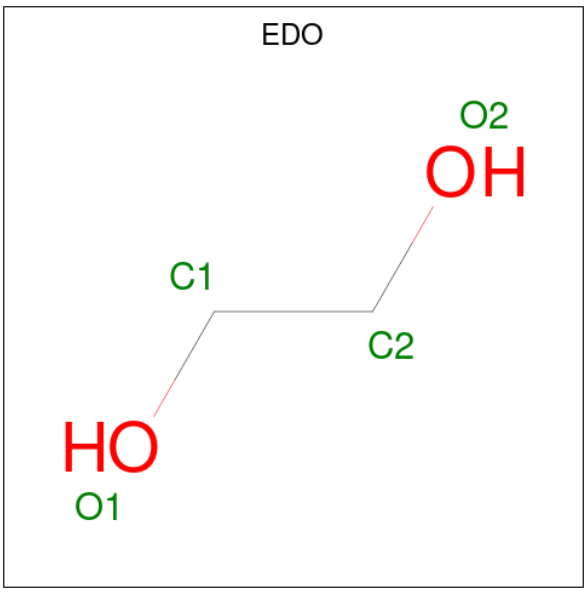
Chain	Residue	Modelled	Actual	Comment	Reference
B	534	HIS	-	expression tag	UNP Q9I580
B	535	HIS	-	expression tag	UNP Q9I580
B	536	HIS	-	expression tag	UNP Q9I580
B	537	HIS	-	expression tag	UNP Q9I580
B	538	HIS	-	expression tag	UNP Q9I580
B	539	SER	-	expression tag	UNP Q9I580
B	540	SER	-	expression tag	UNP Q9I580
B	541	GLY	-	expression tag	UNP Q9I580
B	542	LEU	-	expression tag	UNP Q9I580
B	543	VAL	-	expression tag	UNP Q9I580
B	544	PRO	-	expression tag	UNP Q9I580
B	545	ARG	-	expression tag	UNP Q9I580
B	546	GLY	-	expression tag	UNP Q9I580
B	547	SER	-	expression tag	UNP Q9I580
B	548	HIS	-	expression tag	UNP Q9I580

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



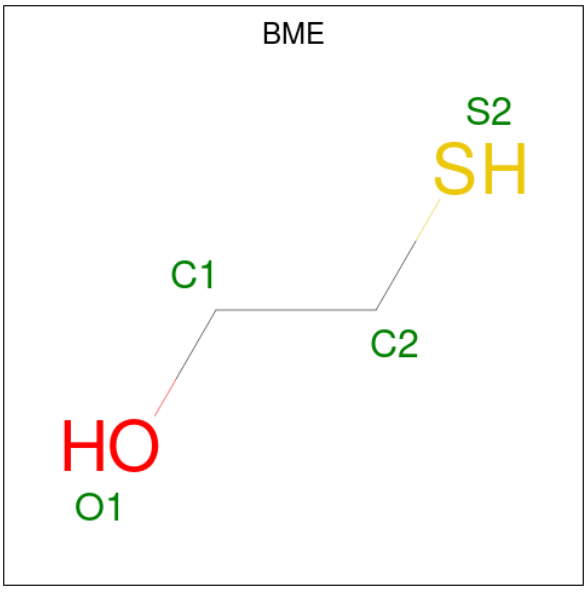
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 4 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C_2H_6OS).



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

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 6 is water.

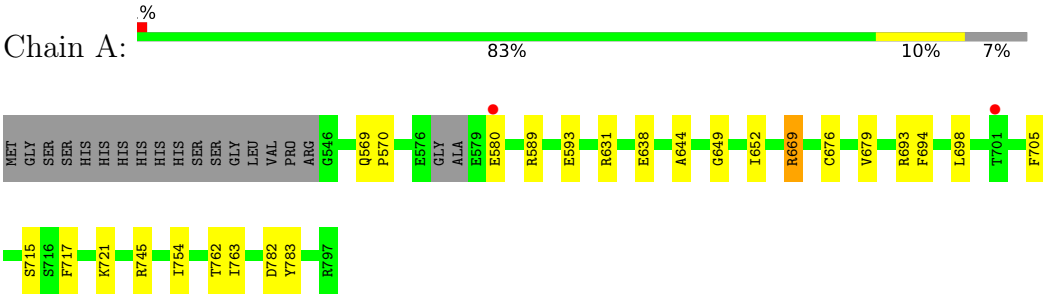
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	104	Total O 104 104	0	0
6	B	84	Total O 84 84	0	0

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.

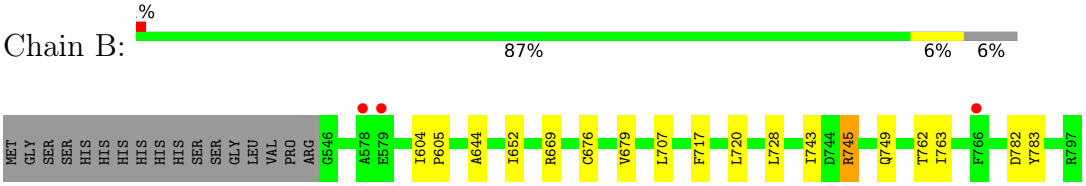
• Molecule 1: EAL domain-containing protein

Chain A:



• Molecule 1: EAL domain-containing protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	64.40Å 65.72Å 172.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.04 – 2.30 46.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	93.5 (46.04-2.30) 93.5 (46.00-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0415	Depositor
R, R_{free}	0.184 , 0.232 0.187 , 0.233	Depositor DCC
R_{free} test set	1495 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	52.7	Xtriage
Anisotropy	0.104	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 38.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.053 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4235	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.35% 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: EDO, MG, GOL, BME

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	0.52	1/2035 (0.0%)	0.78	0/2759
1	B	0.47	0/2045	0.79	0/2773
All	All	0.49	1/4080 (0.0%)	0.79	0/5532

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	2
1	B	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	593	GLU	CD-OE1	5.93	1.32	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	669	ARG	Sidechain
1	A	745	ARG	Sidechain
1	B	669	ARG	Sidechain
1	B	745[A]	ARG	Sidechain

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	1994	0	1990	16	0
1	B	2003	0	1997	10	0
2	A	18	0	24	0	0
2	B	6	0	8	0	0
3	A	8	0	12	0	0
3	B	8	0	11	0	0
4	A	4	0	6	2	0
4	B	4	0	5	1	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	104	0	0	0	0
6	B	84	0	0	0	0
All	All	4235	0	4053	25	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 3.

The worst 5 of 25 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:711:CYS:SG	4:A:806:BME:S2	2.45	0.96
1:A:589[B]:ARG:HG3	1:A:589[B]:ARG:HH11	1.51	0.73
1:A:631:ARG:NH2	1:A:638[B]:GLU:O	2.21	0.71
1:B:652:ILE:HD11	1:B:679:VAL:HG11	1.73	0.70
1:A:631:ARG:NH2	1:A:638[A]:GLU:O	2.23	0.70

There are no symmetry-related clashes.

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	248/269 (92%)	243 (98%)	5 (2%)	0	100	100
1	B	252/269 (94%)	246 (98%)	6 (2%)	0	100	100
All	All	500/538 (93%)	489 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	214/227 (94%)	209 (98%)	5 (2%)	50	67
1	B	214/227 (94%)	214 (100%)	0	100	100
All	All	428/454 (94%)	423 (99%)	5 (1%)	71	84

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

Mol	Chain	Res	Type
1	A	580	GLU
1	A	669	ARG
1	A	693	ARG
1	A	705	PHE
1	A	715	SER

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

Mol	Chain	Res	Type
1	A	553	GLN
1	A	633	GLN
1	B	553	GLN
1	B	633	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 12 ligands modelled in this entry, 2 are monoatomic - leaving 10 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	$\#\|Z\| > 2$
2	GOL	A	802	-	5,5,5	0.26	0	5,5,5	0.62	0
4	BME	A	806	-	3,3,3	0.12	0	1,2,2	0.22	0
3	EDO	B	803	-	3,3,3	0.62	0	2,2,2	0.05	0
3	EDO	B	804	-	3,3,3	0.30	0	2,2,2	0.48	0
3	EDO	A	804	-	3,3,3	0.17	0	2,2,2	0.25	0
3	EDO	A	805	-	3,3,3	0.46	0	2,2,2	0.18	0
2	GOL	A	801	-	5,5,5	0.26	0	5,5,5	0.59	0
2	GOL	B	801	-	5,5,5	0.23	0	5,5,5	0.59	0
2	GOL	A	803	-	5,5,5	0.09	0	5,5,5	0.32	0
4	BME	B	802	-	3,3,3	0.24	0	1,2,2	0.39	0

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
2	GOL	A	802	-	-	4/4/4/4	-
4	BME	A	806	-	-	0/1/1/1	-
3	EDO	B	803	-	-	0/1/1/1	-
3	EDO	B	804	-	-	0/1/1/1	-
3	EDO	A	804	-	-	1/1/1/1	-
3	EDO	A	805	-	-	1/1/1/1	-
2	GOL	A	801	-	-	4/4/4/4	-
2	GOL	B	801	-	-	1/4/4/4	-
2	GOL	A	803	-	-	4/4/4/4	-
4	BME	B	802	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801	GOL	O1-C1-C2-O2
2	A	801	GOL	O1-C1-C2-C3
2	A	801	GOL	C1-C2-C3-O3
2	A	801	GOL	O2-C2-C3-O3
2	A	802	GOL	O1-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	806	BME	2	0
4	B	802	BME	1	0

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

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	250/269 (92%)	-0.29	2 (0%)	86	89	41, 56, 84, 131	0
1	B	252/269 (93%)	-0.30	3 (1%)	79	83	42, 60, 88, 158	0
All	All	502/538 (93%)	-0.30	5 (0%)	82	86	41, 58, 86, 158	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	579	GLU	5.2
1	B	578	ALA	2.8
1	A	701	THR	2.4
1	B	766	PHE	2.1
1	A	580	GLU	2.0

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

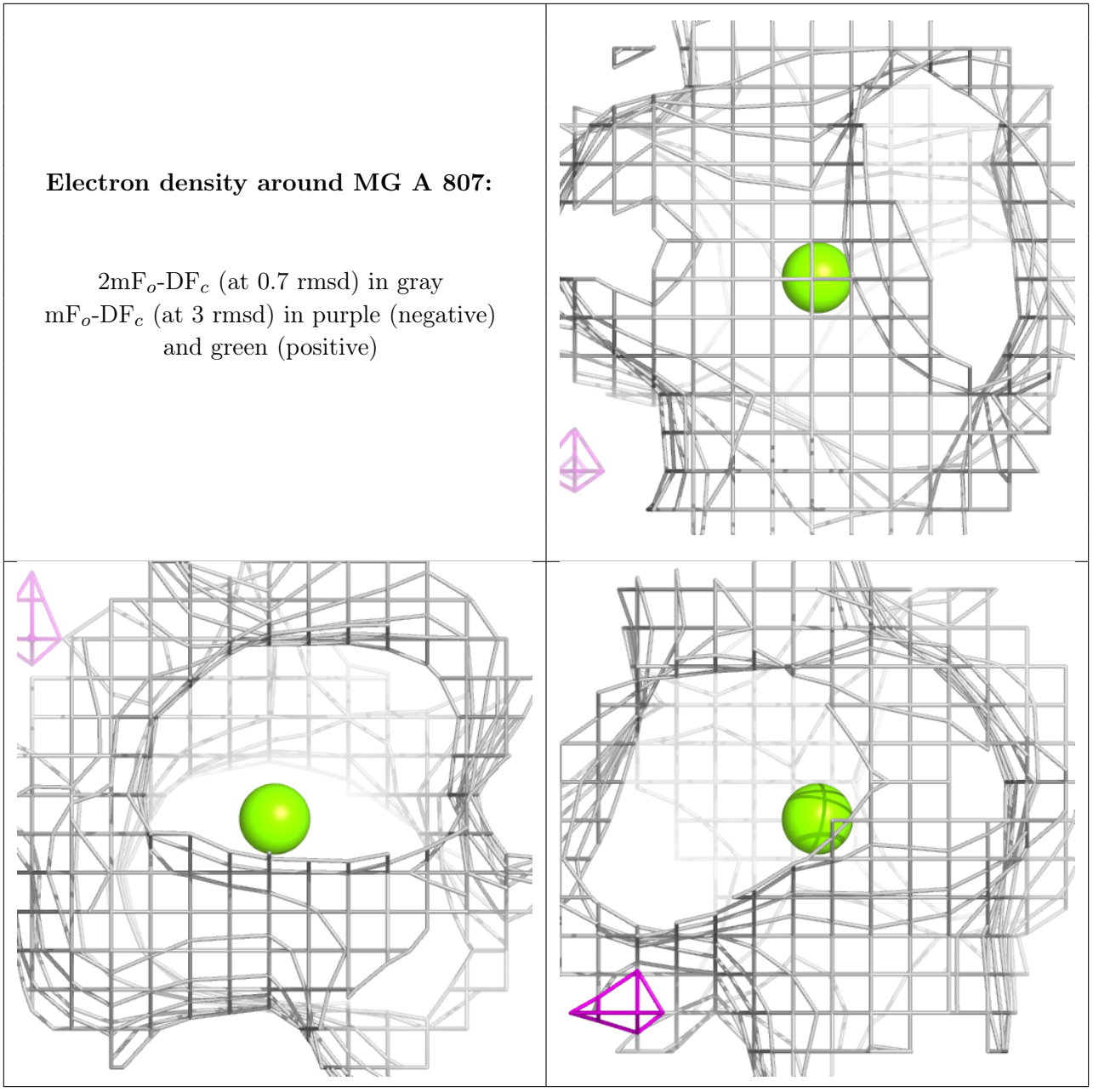
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	GOL	A	801	6/6	0.92	0.13	52,59,60,65	0
2	GOL	A	802	6/6	0.92	0.16	54,69,83,83	0
3	EDO	B	804	4/4	0.92	0.09	62,73,78,81	0
4	BME	A	806	4/4	0.93	0.11	86,87,92,94	0
2	GOL	A	803	6/6	0.94	0.17	71,83,85,91	0
3	EDO	A	804	4/4	0.94	0.10	76,82,82,84	0
2	GOL	B	801	6/6	0.95	0.19	60,64,66,72	0
4	BME	B	802	4/4	0.96	0.12	59,69,75,79	0
3	EDO	B	803	4/4	0.98	0.13	50,53,57,59	0
5	MG	A	807	1/1	0.98	0.10	46,46,46,46	0
3	EDO	A	805	4/4	0.99	0.16	46,50,56,56	0
5	MG	B	805	1/1	0.99	0.06	46,46,46,46	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.

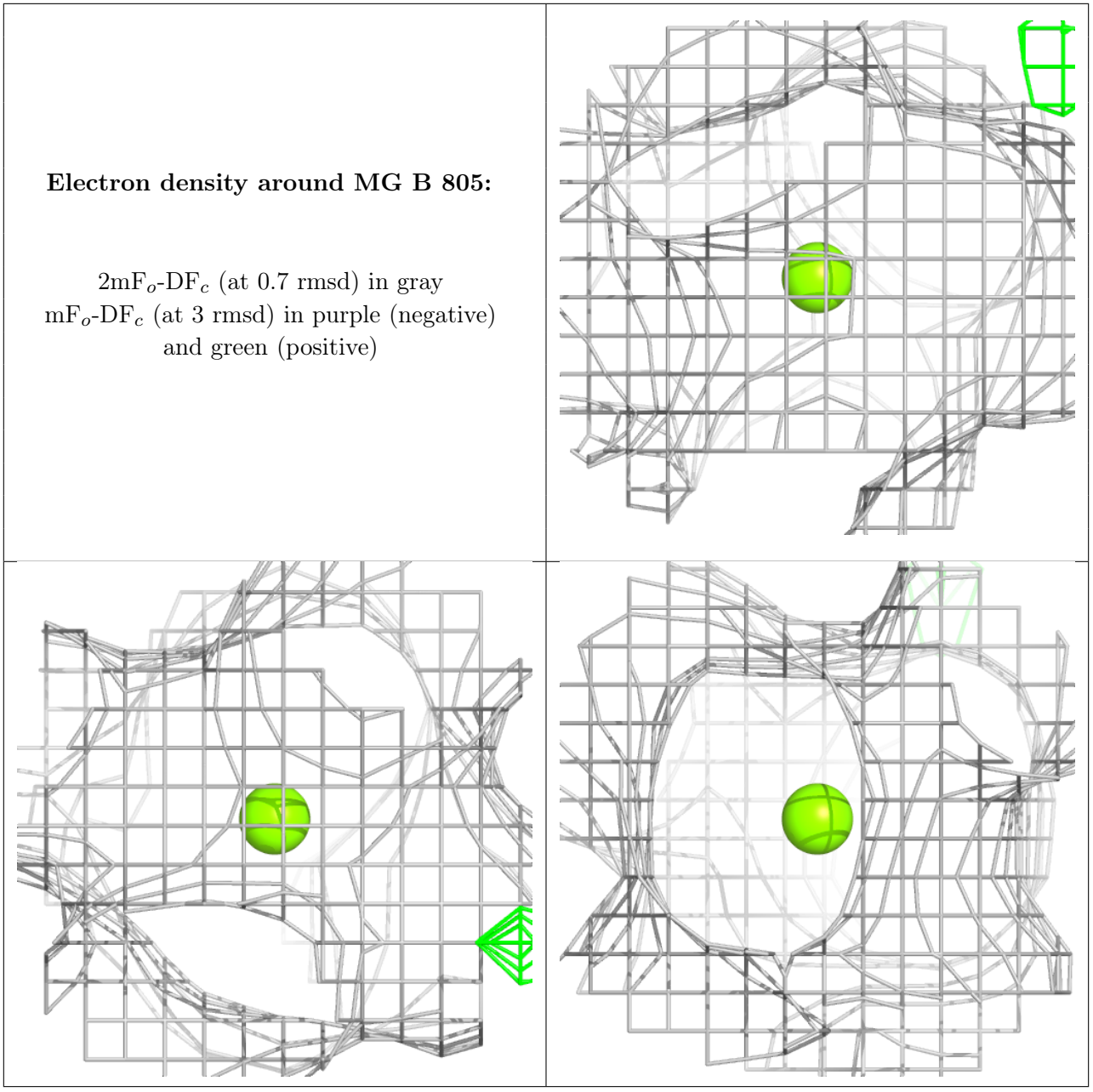
Electron density around MG A 807:

$2mF_o$-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 805:

$2mF_o$-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.