WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2024 – 02:54 pm BST

PDB ID	:	9EVN
Title	:	Plasmodium falciparum apical membrane antigen FVO
Authors	:	Bentley, G.A.; Saul, F.A.; Vulliez-Le Normand, B.
Deposited on	:	2024-03-31
Resolution	:	2.00 Å(reported)

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

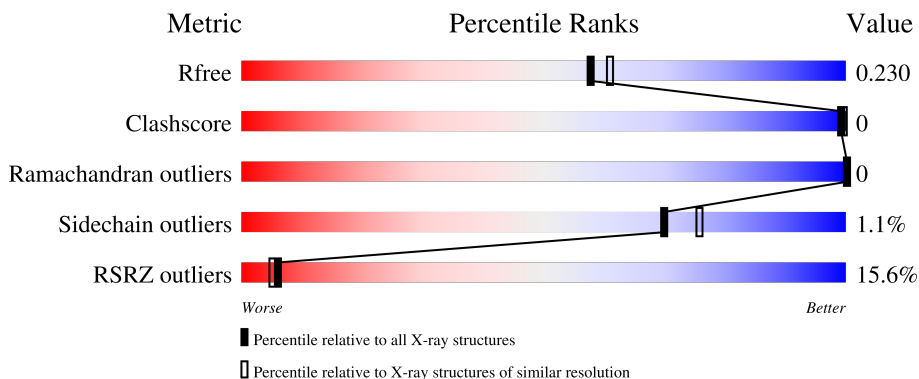|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.00 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		0
Ramachandran outliers		0
Sidechain outliers		1.1%
RSRZ outliers		15.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

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	349	15% / 92% / 7%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2879 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 Apical membrane antigen 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	326	Total 2621	C 1664	N 437	O 503	S 17	0	3	0

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	GLU	-	expression tag	UNP A0A075DBS4
A	103	ALA	-	expression tag	UNP A0A075DBS4
A	104	GLU	-	expression tag	UNP A0A075DBS4
A	105	PHE	-	expression tag	UNP A0A075DBS4
A	162	LYS	ASN	engineered mutation	UNP A0A075DBS4
A	288	VAL	THR	engineered mutation	UNP A0A075DBS4
A	373	ASP	SER	engineered mutation	UNP A0A075DBS4
A	422	ASP	ASN	engineered mutation	UNP A0A075DBS4
A	423	LYS	SER	engineered mutation	UNP A0A075DBS4
A	443	VAL	-	expression tag	UNP A0A075DBS4
A	444	ASP	-	expression tag	UNP A0A075DBS4
A	445	HIS	-	expression tag	UNP A0A075DBS4
A	446	HIS	-	expression tag	UNP A0A075DBS4
A	447	HIS	-	expression tag	UNP A0A075DBS4
A	448	HIS	-	expression tag	UNP A0A075DBS4
A	449	HIS	-	expression tag	UNP A0A075DBS4
A	450	HIS	-	expression tag	UNP A0A075DBS4

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	256	Total 256	O 256	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: Apical membrane antigen 1
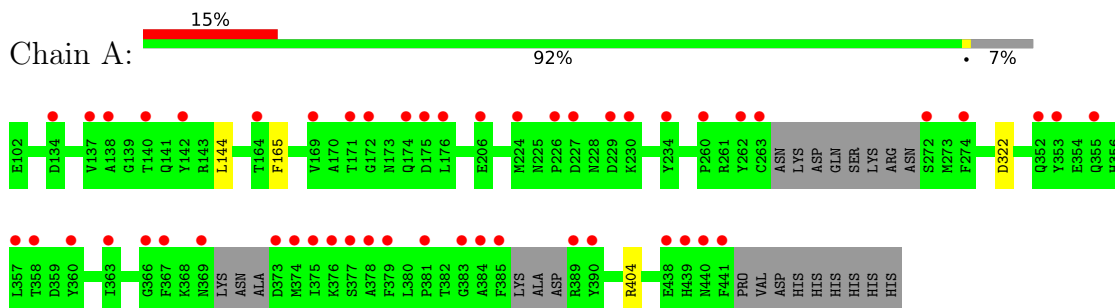
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	38.28Å 94.22Å 96.20Å	Depositor
a, b, c, α, β, γ	90.00° 90.00° 90.00°	
Resolution (Å)	47.11 – 2.00	Depositor
	47.11 – 2.00	EDS
% Data completeness	99.3 (47.11-2.00)	Depositor
(in resolution range)	99.3 (47.11-2.00)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.49 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0415	Depositor
R, R_{free}	0.177 , 0.219	Depositor
	0.187 , 0.230	DCC
R_{free} test set	1225 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	25.7	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 58.0	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for -h,l,k	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2879	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.97% 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: CL

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	0.46	0/2696	0.71	0/3646

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	2621	0	2467	1	0
2	A	2	0	0	0	0
3	A	256	0	0	0	0
All	All	2879	0	2467	1	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 0.

All (1) 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:404[A]:ARG:HH11	1:A:404[A]:ARG:HG3	1.86	0.41

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	321/349 (92%)	307 (96%)	14 (4%)	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	283/308 (92%)	280 (99%)	3 (1%)	70	76

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	144	LEU
1	A	165	PHE
1	A	322	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 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	326/349 (93%)	0.60	51 (15%)	6	5	20, 40, 114, 141	3 (0%)

All (51) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	367	PHE	5.1
1	A	385	PHE	4.3
1	A	137	VAL	4.3
1	A	363	ILE	4.2
1	A	441	PHE	4.1
1	A	352	GLN	4.0
1	A	390	TYR	3.6
1	A	176	LEU	3.6
1	A	142	TYR	3.5
1	A	375	ILE	3.5
1	A	263	CYS	3.4
1	A	355	GLN	3.3
1	A	171	THR	3.3
1	A	440	ASN	3.3
1	A	379	PHE	3.2
1	A	357	LEU	3.2
1	A	376	LYS	3.2
1	A	381	PRO	3.1
1	A	389	ARG	3.1
1	A	169	VAL	3.0
1	A	262	TYR	2.9
1	A	384	ALA	2.8
1	A	374	MET	2.8
1	A	358	THR	2.7
1	A	226	PRO	2.6
1	A	360	TYR	2.6
1	A	438	GLU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	175	ASP	2.6
1	A	174	GLN	2.6
1	A	373	ASP	2.5
1	A	138	ALA	2.5
1	A	383	GLY	2.5
1	A	377	SER	2.4
1	A	140	THR	2.4
1	A	230	LYS	2.4
1	A	378	ALA	2.4
1	A	227	ASP	2.3
1	A	439	HIS	2.2
1	A	234	TYR	2.2
1	A	353	TYR	2.2
1	A	224	MET	2.2
1	A	260	PRO	2.2
1	A	206	GLU	2.1
1	A	229	ASP	2.1
1	A	272	SER	2.1
1	A	172	GLY	2.1
1	A	164	THR	2.1
1	A	134	ASP	2.1
1	A	366	GLY	2.0
1	A	274	PHE	2.0
1	A	369	ASN	2.0

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

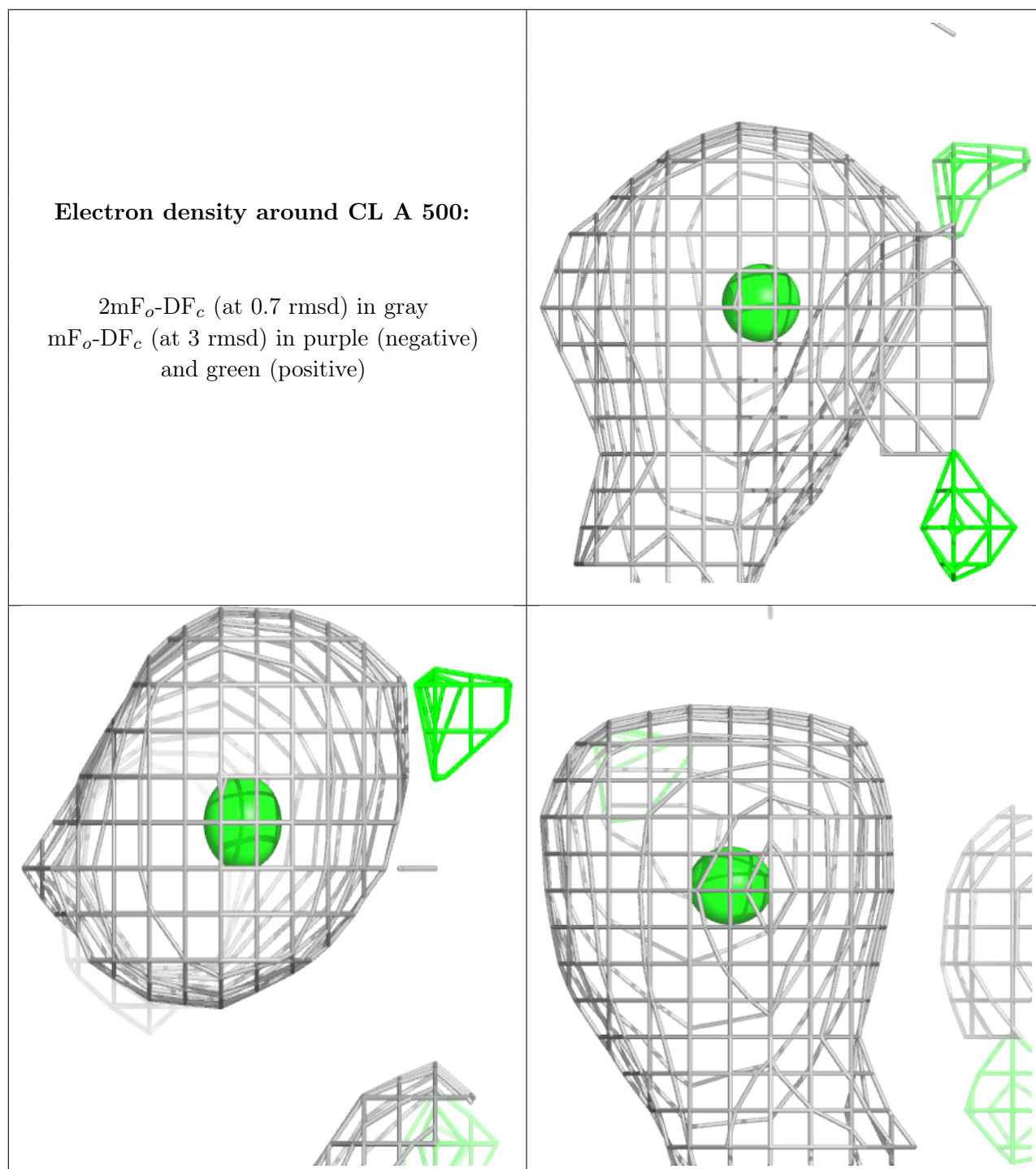There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.
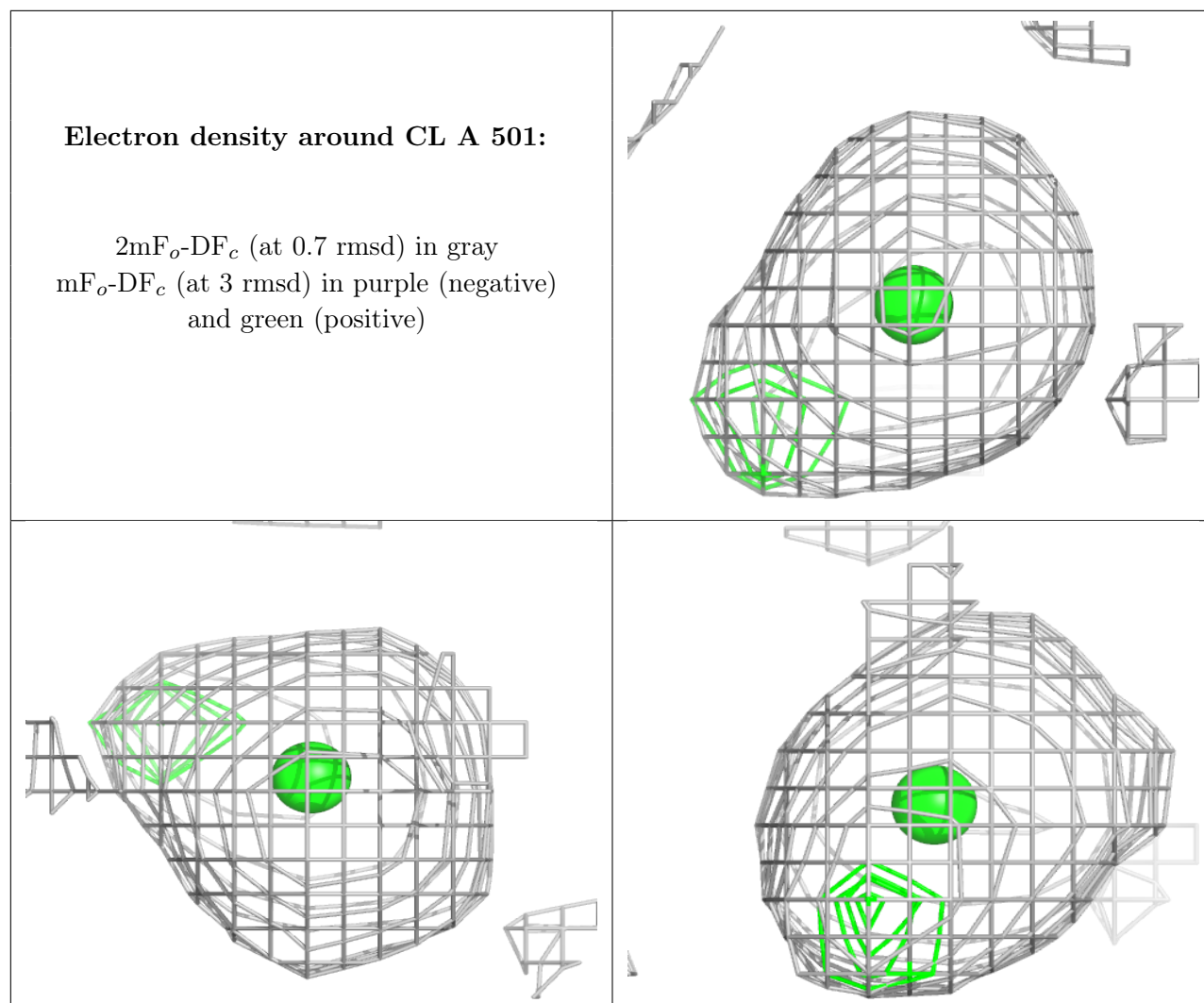
6.4 Ligands ⓘ

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, 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($\mathring{A}^2$)	Q<0.9
2	CL	A	500	1/1	0.94	0.13	73,73,73,73	0
2	CL	A	501	1/1	0.96	0.15	63,63,63,63	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 CL A 500:

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

Electron density around CL A 501:

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

6.5   Other polymers ⓘ

There are no such residues in this entry.