Full wwPDB X-ray Structure Validation Report (i)

Oct 15, 2024 – 10:03 AM EDT

PDB ID	:	9DYT
Title	:	Acanthamoeba Polyphaga Mimivirus R699
Deposited on	:	2024-10-14
Resolution	:	1.80 Å(reported)

This wwPDB validation report is for manuscript review

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

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 (1)) were used in the production of this report:

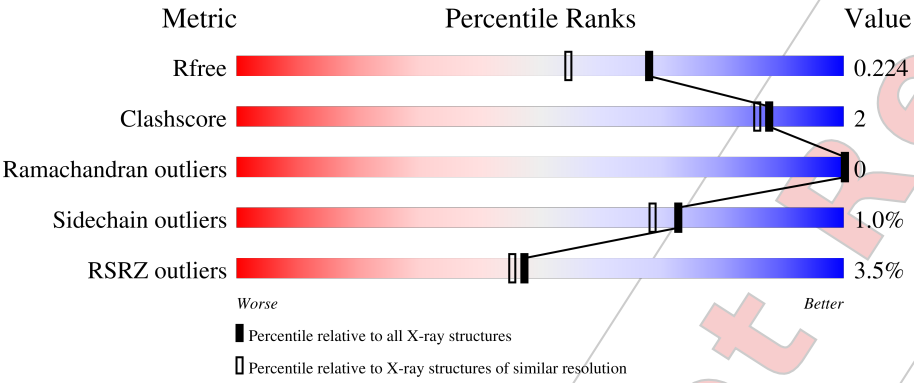
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.80 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}	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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 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	458	2% 91% 5%
1	B	458	4% 90% 5% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7789 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 R699.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	442	Total	C	N	O	S	0	6	0
			3677	2368	608	680	21			
1	B	433	Total	C	N	O	S	0	4	0
			3604	2328	591	663	22			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q5UNV6
A	-1	PRO	-	expression tag	UNP Q5UNV6
A	0	GLY	-	expression tag	UNP Q5UNV6
A	1	SER	-	expression tag	UNP Q5UNV6
B	-2	GLY	-	expression tag	UNP Q5UNV6
B	-1	PRO	-	expression tag	UNP Q5UNV6
B	0	GLY	-	expression tag	UNP Q5UNV6
B	1	SER	-	expression tag	UNP Q5UNV6

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	265	Total	O	0	0
			265	265		

Continued on next page...

Continued from previous page...

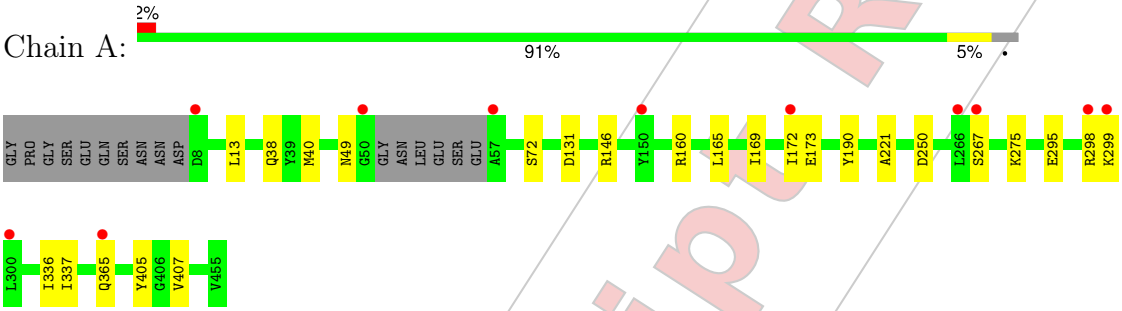
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	241	Total O 241 241	0	0

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.

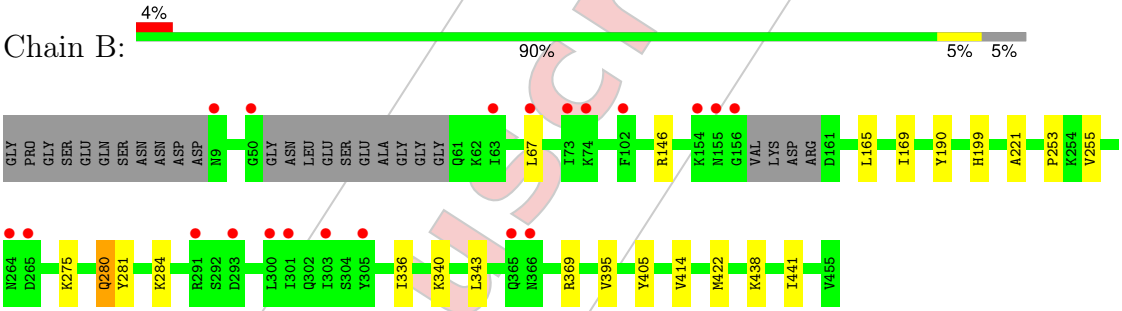
- Molecule 1: R699

Chain A:



- Molecule 1: R699

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	69.82Å 121.70Å 72.69Å 90.00° 118.59° 90.00°	Depositor
Resolution (Å)	61.38 – 1.80 61.38 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.7 (61.38-1.80) 98.7 (61.38-1.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.190 , 0.220 0.193 , 0.224	Depositor DCC
R_{free} test set	4986 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	26.5	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 34.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h-l,k,h 0.000 for l,k,-h-l 0.017 for h,-k,-h-l 0.013 for -h-l,-k,l 0.019 for l,-k,h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7789	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.18% 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: MN

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	0.41	0/3775	0.79	2/5094 (0.0%)
1	B	0.40	0/3695	0.78	0/4987
All	All	0.41	0/7470	0.79	2/10081 (0.0%)

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	1
All	All	0	2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	250[A]	ASP	CB-CA-C	6.65	123.69	110.40
1	A	250[B]	ASP	CB-CA-C	6.65	123.69	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	405	TYR	Peptide
1	B	405	TYR	Peptide

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	3677	0	3620	14	0
1	B	3604	0	3550	13	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	265	0	0	0	0
3	B	241	0	0	1	0
All	All	7789	0	7170	27	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 2.

All (27) 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:49:ASN:HB3	1:A:160:ARG:HH21	1.50	0.76
1:A:13:LEU:HD22	1:A:40:MET:HG2	1.76	0.68
1:A:275:LYS:HG3	1:A:336:ILE:HG23	1.75	0.68
1:A:49:ASN:HB3	1:A:160:ARG:NH2	2.09	0.68
1:B:422:MET:HE1	1:B:438:LYS:HE2	1.78	0.66
1:B:275:LYS:HG3	1:B:336:ILE:HG23	1.83	0.61
1:A:365[A]:GLN:HA	1:A:365[A]:GLN:OE1	2.01	0.60
1:A:13:LEU:HD22	1:A:40:MET:CG	2.38	0.53
1:B:280:GLN:HG2	1:B:340:LYS:HB3	1.91	0.52
1:A:190:TYR:HA	1:A:221:ALA:HB2	1.91	0.51
1:A:13:LEU:CD2	1:A:40:MET:HG2	2.40	0.51
1:B:199:HIS:HD2	3:B:694:HOH:O	1.94	0.51
1:A:172:ILE:HG13	1:A:173:GLU:HG3	1.93	0.50
1:B:190:TYR:HA	1:B:221:ALA:HB2	1.94	0.50
1:B:281:TYR:O	1:B:284:LYS:HE2	2.12	0.49
1:A:337:ILE:HD11	1:A:407:VAL:HG21	1.94	0.49
1:B:369:ARG:HB2	1:B:369:ARG:HH11	1.77	0.48
1:A:295:GLU:HA	1:A:298[A]:ARG:HG2	1.96	0.47
1:B:255:VAL:HG11	1:B:343:LEU:HD21	1.96	0.46
1:B:369:ARG:HH11	1:B:369:ARG:CB	2.29	0.46
1:A:295:GLU:HG3	1:A:298[A]:ARG:CZ	2.47	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:165:LEU:O	1:B:169:ILE:HG12	2.19	0.42
1:B:414:VAL:HG13	1:B:441:ILE:HD13	2.02	0.41
1:A:165:LEU:O	1:A:169:ILE:HG12	2.20	0.41
1:A:295:GLU:HG2	1:A:299:LYS:HD3	2.03	0.41
1:B:253:PRO:HD2	1:B:281:TYR:OH	2.21	0.40
1:B:67:LEU:HD23	1:B:67:LEU:HA	1.95	0.40

There are no symmetry-related clashes.

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	444/458 (97%)	434 (98%)	10 (2%)	0	100	100
1	B	431/458 (94%)	425 (99%)	6 (1%)	0	100	100
All	All	875/916 (96%)	859 (98%)	16 (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	408/415 (98%)	403 (99%)	5 (1%)	67	62
1	B	401/415 (97%)	397 (99%)	4 (1%)	73	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	809/830 (98%)	800 (99%)	9 (1%)	73 65

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	GLN
1	A	72	SER
1	A	131	ASP
1	A	146	ARG
1	A	267	SER
1	B	146	ARG
1	B	280	GLN
1	B	395[A]	VAL
1	B	395[B]	VAL

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	296	GLN
1	B	199	HIS
1	B	302	GLN
1	B	365	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

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 (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	442/458 (96%)	-0.03	11 (2%)	58	57	10, 30, 51, 83	6 (1%)
1	B	433/458 (94%)	0.06	20 (4%)	38	35	10, 30, 60, 97	4 (0%)
All	All	875/916 (95%)	0.01	31 (3%)	47	45	10, 30, 57, 97	10 (1%)

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	365	GLN	3.6
1	A	172	ILE	3.4
1	B	301	ILE	2.9
1	B	74	LYS	2.9
1	A	57	ALA	2.9
1	B	303	ILE	2.8
1	A	8	ASP	2.8
1	B	300	LEU	2.8
1	B	156	GLY	2.6
1	A	365[A]	GLN	2.6
1	B	63	ILE	2.6
1	B	50	GLY	2.5
1	A	298[A]	ARG	2.5
1	B	102	PHE	2.5
1	B	305	TYR	2.5
1	B	366	ASN	2.5
1	A	150	TYR	2.4
1	A	50	GLY	2.4
1	B	155	ASN	2.4
1	B	73	ILE	2.2
1	B	291	ARG	2.2
1	A	299	LYS	2.2
1	A	266	LEU	2.2
1	A	300	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	9	ASN	2.2
1	B	264	ASN	2.2
1	B	293	ASP	2.2
1	B	265	ASP	2.2
1	B	154	LYS	2.1
1	A	267	SER	2.1
1	B	67	LEU	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)

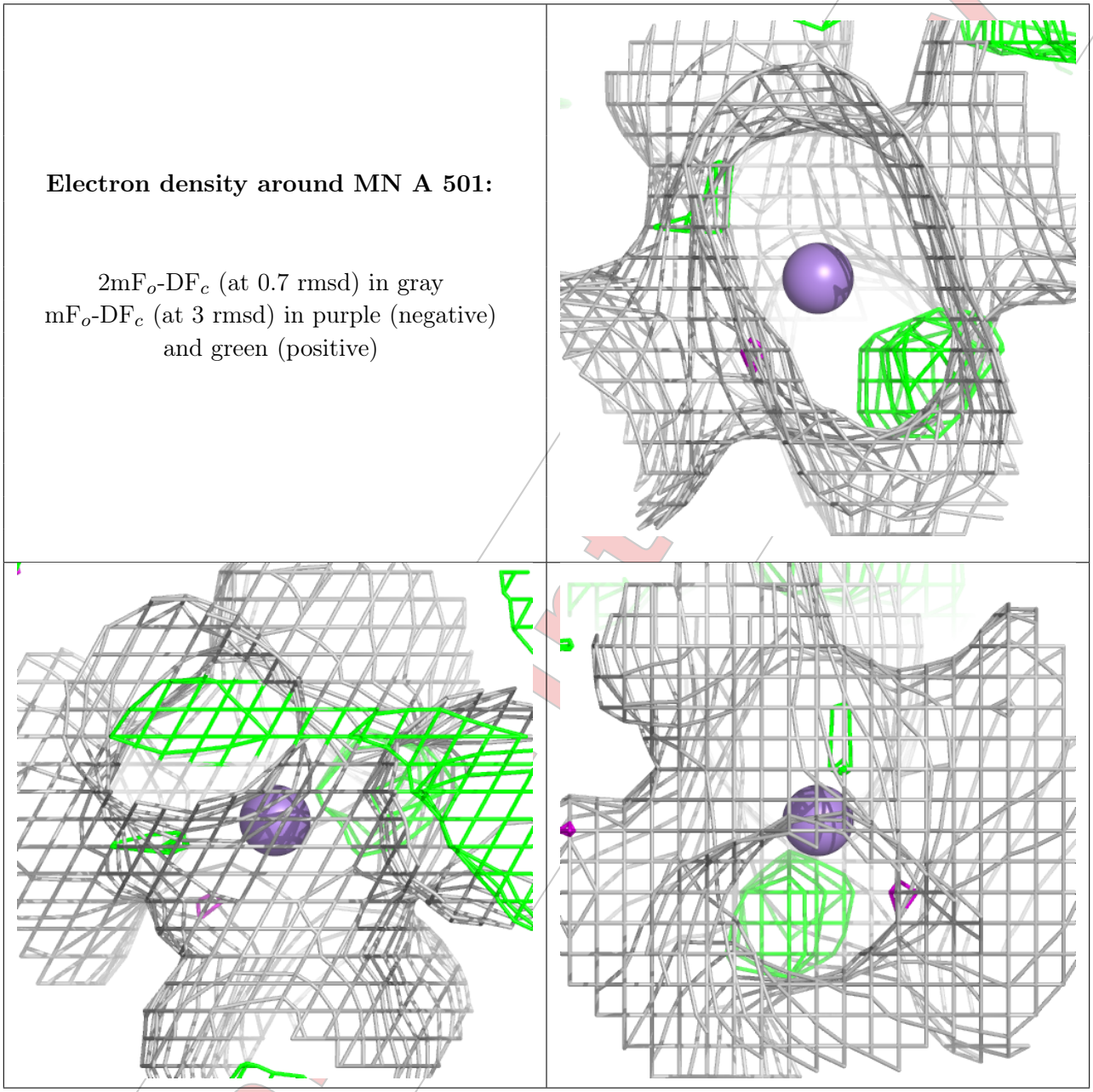
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	MN	A	501	1/1	0.99	0.02	25,25,25,25	0
2	MN	B	501	1/1	1.00	0.01	22,22,22,22	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.

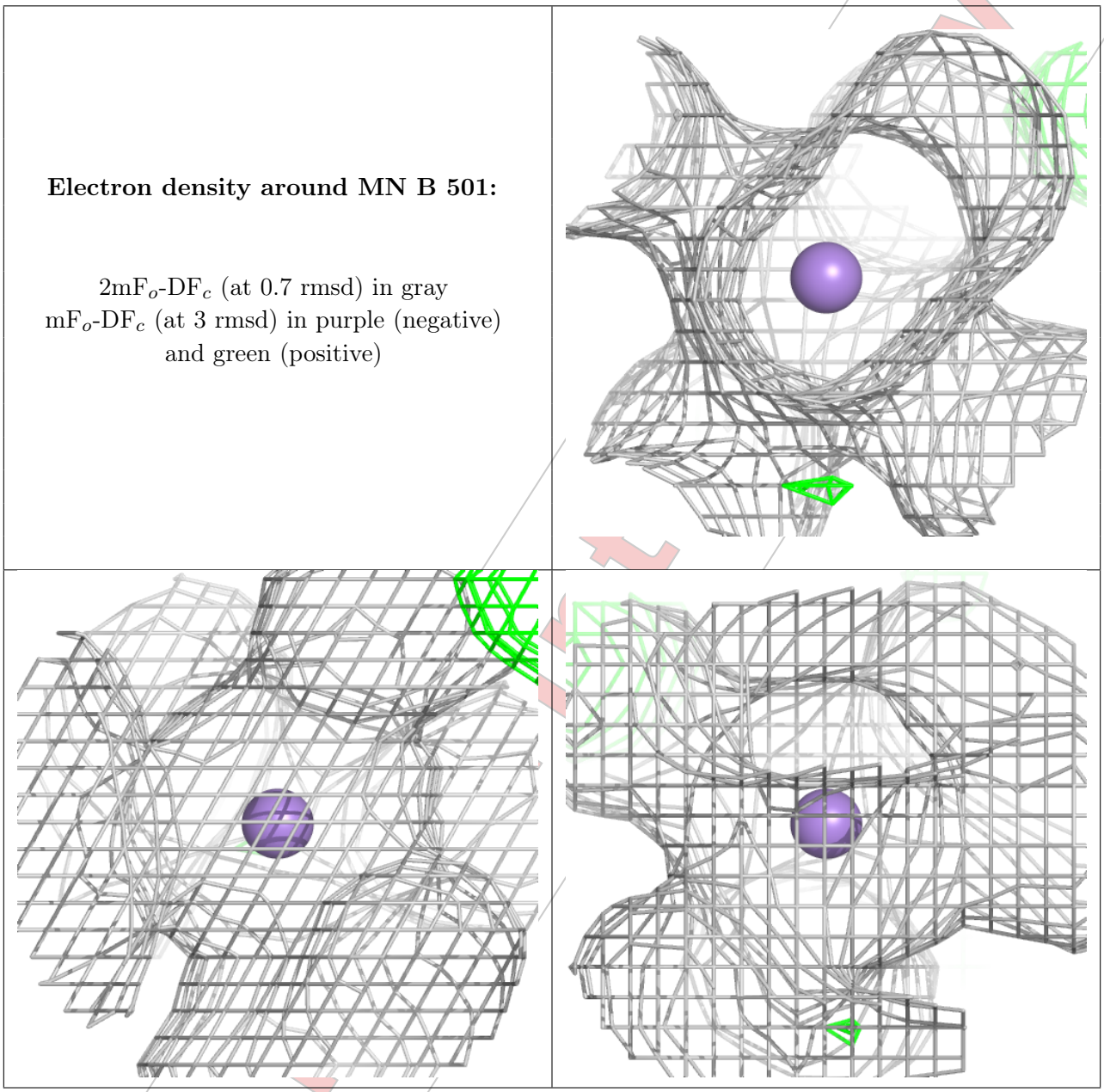
Electron density around MN A 501:

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



Electron density around MN B 501:

$2mF_o$-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.