



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2026 – 07:52 AM UTC

PDB ID : 9KGJ / pdb_00009kgj
Title : Discovery of an orally bioavailable reversible covalent SARS-CoV-2 Mpro inhibitor with pan-coronavirus activity
Authors : Baburajendran, N.
Deposited on : 2024-11-08
Resolution : 1.37 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 ⓘ 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

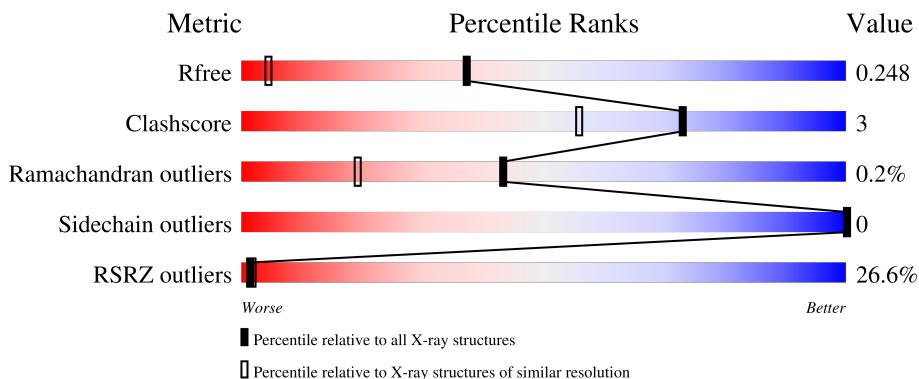
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.37 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}	180053	4403 (1.40-1.36)
Clashscore	190562	4528 (1.40-1.36)
Ramachandran outliers	187476	4459 (1.40-1.36)
Sidechain outliers	187428	4458 (1.40-1.36)
RSRZ outliers	180081	4399 (1.40-1.36)

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 $\leq 5\%$. 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	311	<div> <div>29%</div> <div>88%</div> <div>7%</div> <div>.</div> </div>
1	B	311	<div> <div>23%</div> <div>89%</div> <div>7%</div> <div>.</div> </div>
2	C	4	<div> <div>25%</div> <div>50%</div> <div>25%</div> </div>
2	D	4	<div> <div>25%</div> <div>75%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5094 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 3C-like proteinase nsp5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	298	Total	C	N	O	S	0	0	0
			2308	1460	393	433	22			
1	B	299	Total	C	N	O	S	0	0	0
			2298	1459	384	433	22			

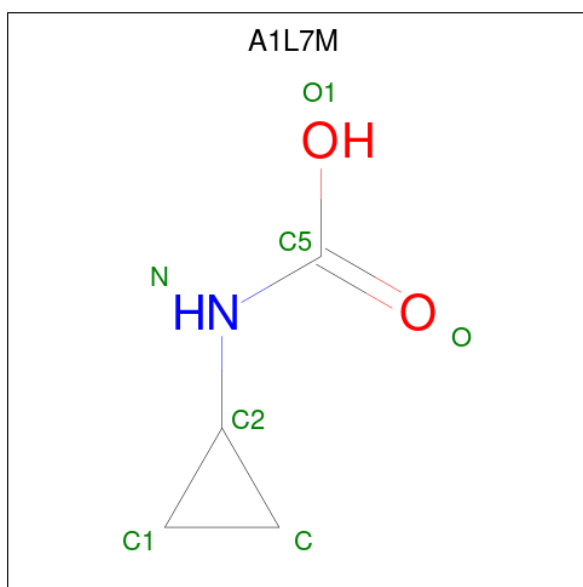
There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP P0DTD1
A	-3	PRO	-	expression tag	UNP P0DTD1
A	-2	LEU	-	expression tag	UNP P0DTD1
A	-1	GLY	-	expression tag	UNP P0DTD1
A	0	SER	-	expression tag	UNP P0DTD1
B	-4	GLY	-	expression tag	UNP P0DTD1
B	-3	PRO	-	expression tag	UNP P0DTD1
B	-2	LEU	-	expression tag	UNP P0DTD1
B	-1	GLY	-	expression tag	UNP P0DTD1
B	0	SER	-	expression tag	UNP P0DTD1

- Molecule 2 is a protein called (S)-3-((S)-3-cyclobutyl-2-((R)-2-cyclopentyl-N-methyl-2-pival amidoacetamido)propanamido)-N-cyclopropyl-2-oxo-4-((S)-2-oxopiperidin-3-yl)butanamide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	4	Total	C	N	O	0	0	0
			37	28	4	5			
2	D	4	Total	C	N	O	0	0	0
			37	28	4	5			

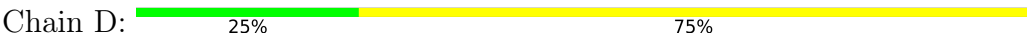
- Molecule 3 is cyclopropylcarbamic acid (CCD ID: A1L7M) (formula: C₄H₇NO₂) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	N	O	0	0
			6	4	1	1		
3	D	1	Total	C	N	O	0	0
			6	4	1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	190	Total	O	0	0
			190	190		
4	B	209	Total	O	0	0
			209	209		
4	C	1	Total	O	0	0
			1	1		
4	D	2	Total	O	0	0
			2	2		



X1	A1E6K2
	A1L7L3
	A1L7K4

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.97Å 99.70Å 59.83Å 90.00° 108.76° 90.00°	Depositor
Resolution (Å)	46.96 – 1.37 46.96 – 1.37	Depositor EDS
% Data completeness (in resolution range)	98.5 (46.96-1.37) 98.5 (46.96-1.37)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	23.57 (at 1.37Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.215 , 0.248 0.216 , 0.248	Depositor DCC
R_{free} test set	1999 reflections (1.55%)	wwPDB-VP
Wilson B-factor (Å ²)	15.0	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 29.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5094	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.81% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: A1L7M, PIV, A1EAK, A1L7L, A1L7K

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	0.49	0/2359	0.72	2/3207 (0.1%)
1	B	0.50	0/2350	0.67	0/3198
All	All	0.50	0/4709	0.70	2/6405 (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

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	222	ARG	CB-CG-CD	-8.14	92.58	111.30
1	A	222	ARG	CD-NE-CZ	-6.78	114.90	124.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	222	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	2308	0	2257	16	0
1	B	2298	0	2229	18	0
2	C	37	0	9	1	0
2	D	37	0	9	0	0
3	C	6	0	0	0	0
3	D	6	0	0	0	0
4	A	190	0	0	1	0
4	B	209	0	0	0	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0
All	All	5094	0	4504	29	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 29 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:165:MET:HG2	2:C:3:A1L7L:CD2	2.14	0.77
1:A:110:GLN:HG3	4:A:513:HOH:O	2.01	0.61
1:A:4:ARG:HH12	1:B:5:LYS:NZ	2.00	0.60
1:A:106:ILE:HG12	1:A:110:GLN:HB2	1.85	0.58
1:B:188:ARG:CZ	1:B:190:THR:HG21	2.34	0.57

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	296/311 (95%)	287 (97%)	8 (3%)	1 (0%)	36 16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	297/311 (96%)	291 (98%)	6 (2%)	0	100	100
All	All	593/622 (95%)	578 (98%)	14 (2%)	1 (0%)	43	19

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	275	GLY

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	257/266 (97%)	257 (100%)	0	100	100
1	B	254/266 (96%)	254 (100%)	0	100	100
All	All	511/532 (96%)	511 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	19	GLN
1	B	214	ASN
1	B	274	ASN
1	B	273	GLN
1	A	119	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues are modelled in this entry.

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	PIV	D	1	2	5,5,6	1.00	0	7,7,9	0.67	0
2	A1EAK	D	2	2	8,9,10	2.49	5 (62%)	6,11,13	0.64	0
2	A1EAK	C	2	2	8,9,10	2.56	6 (75%)	6,11,13	0.40	0
2	PIV	C	1	2	5,5,6	1.05	0	7,7,9	0.68	0
2	A1L7K	C	4	2,3,1	10,12,13	4.39	7 (70%)	10,15,17	1.61	2 (20%)
2	A1L7L	C	3	2	9,10,11	1.20	2 (22%)	9,12,14	2.00	3 (33%)
2	A1L7K	D	4	2,3,1	10,12,13	4.71	6 (60%)	10,15,17	1.86	3 (30%)
2	A1L7L	D	3	2	9,10,11	0.86	0	9,12,14	1.45	1 (11%)

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	PIV	D	1	2	-	0/3/3/6	-
2	A1EAK	D	2	2	-	1/3/13/15	0/1/1/1
2	A1EAK	C	2	2	-	2/3/13/15	0/1/1/1
2	PIV	C	1	2	-	1/3/3/6	-
2	A1L7K	C	4	2,3,1	-	1/5/17/19	0/1/1/1
2	A1L7L	C	3	2	-	1/5/14/16	0/1/1/1
2	A1L7K	D	4	2,3,1	-	1/5/17/19	0/1/1/1
2	A1L7L	D	3	2	-	1/5/14/16	0/1/1/1

The worst 5 of 26 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	4	A1L7K	CD2-NE2	11.37	1.52	1.33
2	C	4	A1L7K	CD2-NE2	10.25	1.50	1.33
2	D	4	A1L7K	O-C	6.25	1.43	1.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	4	A1L7K	O-C	5.81	1.42	1.20
2	C	4	A1L7K	CG-CD2	4.87	1.57	1.50

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	3	A1L7L	CG-CB-CA	-4.93	107.89	114.52
2	D	3	A1L7L	C01-N-CA	3.40	123.88	113.70
2	D	4	A1L7K	CF-NE2-CD2	-3.00	117.66	125.79
2	D	4	A1L7K	CB-CG-CD2	-2.84	105.54	111.03
2	C	4	A1L7K	CB-CG-CD2	-2.84	105.54	111.03

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	A1EAK	C-CA-CB-C15
2	C	3	A1L7L	O-C-CA-CB
2	D	3	A1L7L	CA-CB-CG-CD2
2	C	4	A1L7K	O-C-CA-CB
2	D	4	A1L7K	O-C-CA-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	3	A1L7L	1	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

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	$\# Z > 2$
3	A1L7M	D	101	2	6,6,7	2.26	3 (50%)	6,7,9	2.40	2 (33%)
3	A1L7M	C	101	2	6,6,7	2.26	3 (50%)	6,7,9	1.59	2 (33%)

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
3	A1L7M	D	101	2	-	0/1/5/6	0/1/1/1
3	A1L7M	C	101	2	-	0/1/5/6	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	101	A1L7M	C5-N	4.28	1.47	1.33
3	D	101	A1L7M	C5-N	4.12	1.46	1.33
3	C	101	A1L7M	C1-C2	2.27	1.54	1.48
3	D	101	A1L7M	C-C2	2.20	1.53	1.48
3	D	101	A1L7M	C1-C2	2.20	1.53	1.48

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	101	A1L7M	C-C2-N	-4.93	112.00	118.56
3	D	101	A1L7M	O-C5-N	-3.00	117.58	125.32
3	C	101	A1L7M	C1-C2-N	-2.59	115.11	118.56
3	C	101	A1L7M	C-C2-N	-2.37	115.41	118.56

There are no chirality outliers.

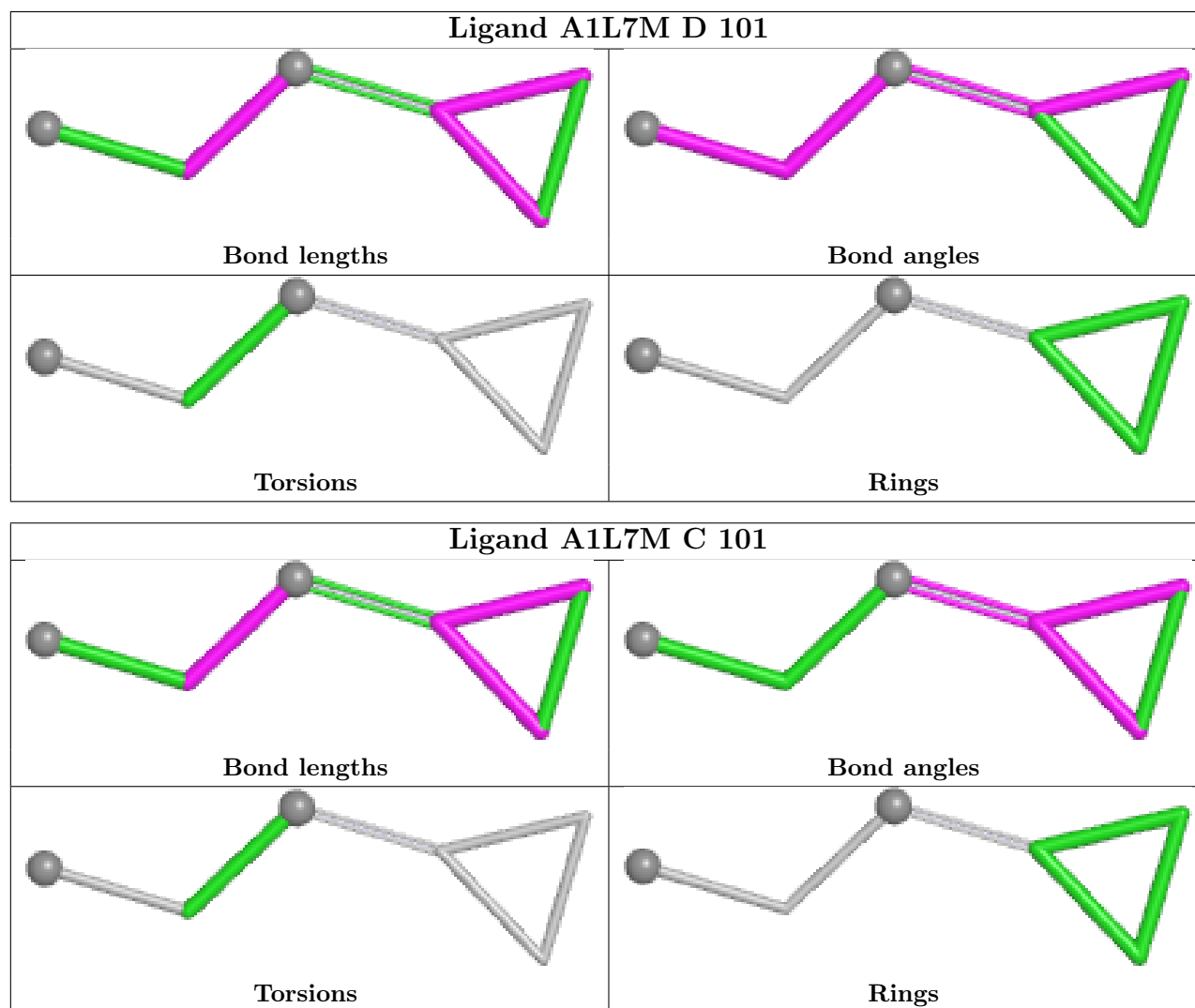
There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	298/311 (95%)	1.52	89 (29%) 1 1	13, 25, 47, 80	0
1	B	299/311 (96%)	1.17	70 (23%) 2 2	12, 21, 46, 78	0
2	C	0/4	-	-	-	-
2	D	0/4	-	-	-	-
All	All	597/630 (94%)	1.34	159 (26%) 1 2	12, 23, 47, 80	0

The worst 5 of 159 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	223	PHE	8.6
1	B	3	PHE	7.5
1	A	275	GLY	6.7
1	B	300	CYS	6.3
1	B	275	GLY	6.2

6.2 Non-standard residues in protein, DNA, RNA chains [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, 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	PIV	D	1	6/7	0.85	0.14	29,30,33,35	0
2	A1EAK	C	2	9/10	0.89	0.16	20,25,39,39	0
2	A1L7L	D	3	10/11	0.91	0.12	20,25,36,38	0
2	A1EAK	D	2	9/10	0.92	0.11	19,26,34,40	0
2	A1L7K	D	4	12/13	0.92	0.10	20,23,26,26	0
2	PIV	C	1	6/7	0.93	0.10	23,26,30,31	0
2	A1L7K	C	4	12/13	0.94	0.08	16,20,22,23	0
2	A1L7L	C	3	10/11	0.94	0.11	20,23,34,36	0

6.3 Carbohydrates [i](#)

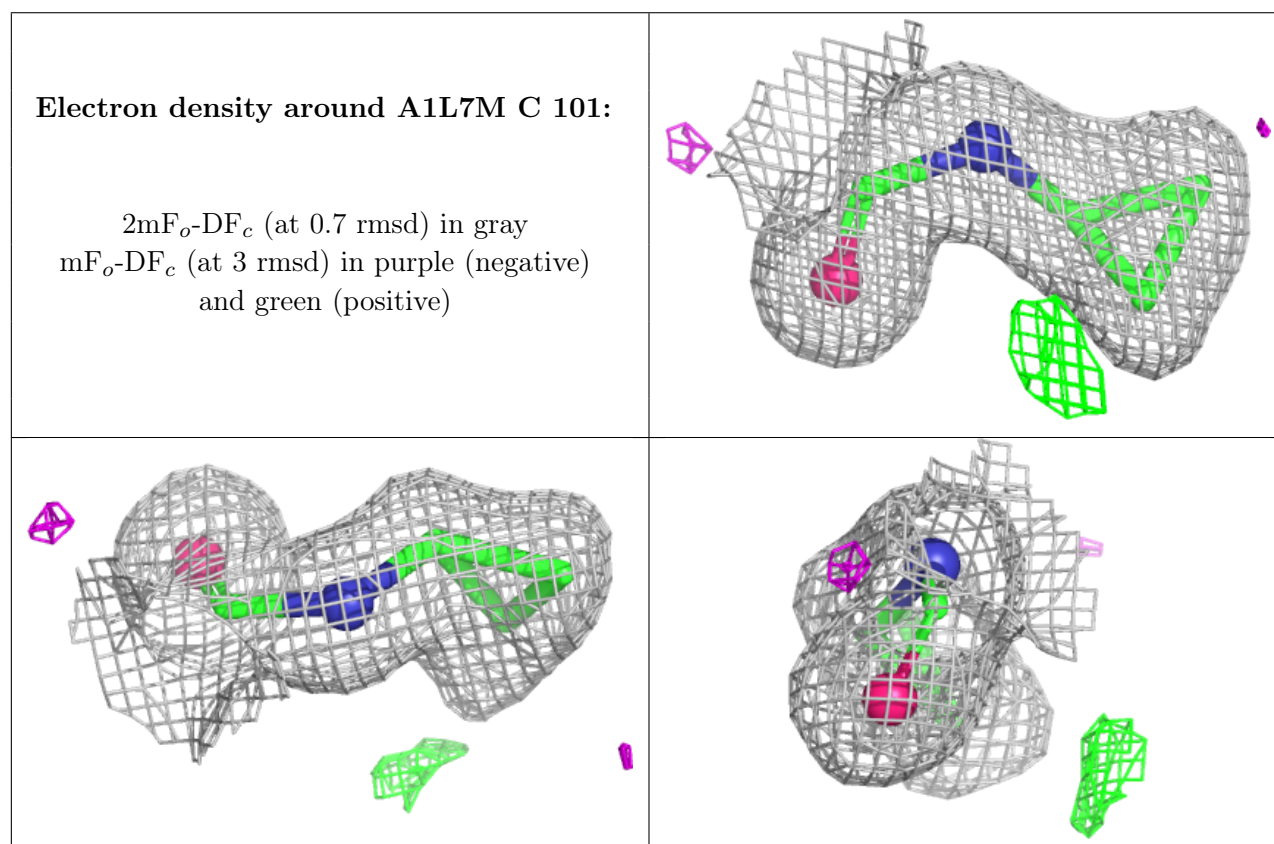
There are no oligosaccharides in this entry.

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, 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.

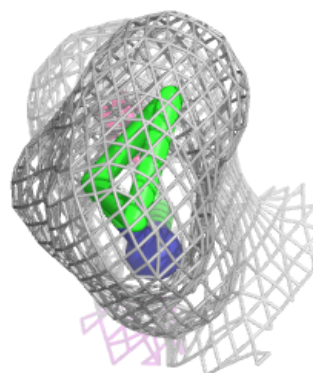
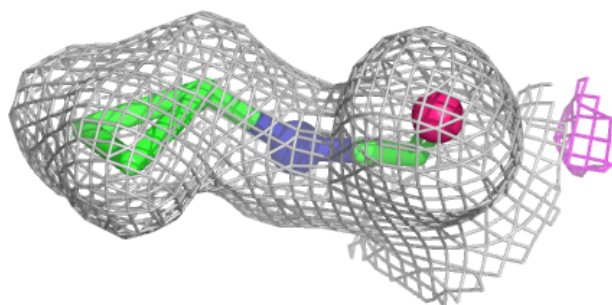
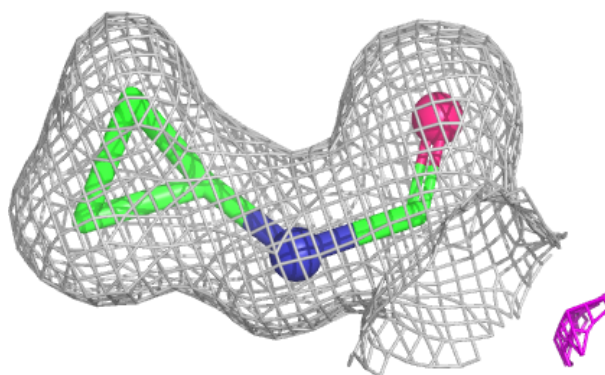
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	A1L7M	C	101	6/7	0.93	0.10	20,27,29,31	0
3	A1L7M	D	101	6/7	0.97	0.07	18,20,21,22	0

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 A1L7M D 101:

$2mF_o-DF_c$ (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.