Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 04:33 pm BST

PDB ID	:	6EVK
Title	:	Crystal structure of bat influenza A/H17N10 polymerase with viral RNA promoter and cap analogue m7GTP
Authors	:	Pflug, A.; Cusack, S.
Deposited on	:	2017-11-01
Resolution	:	2.90 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 following versions of software and data (see references ⓘ) were used in the production of this report:

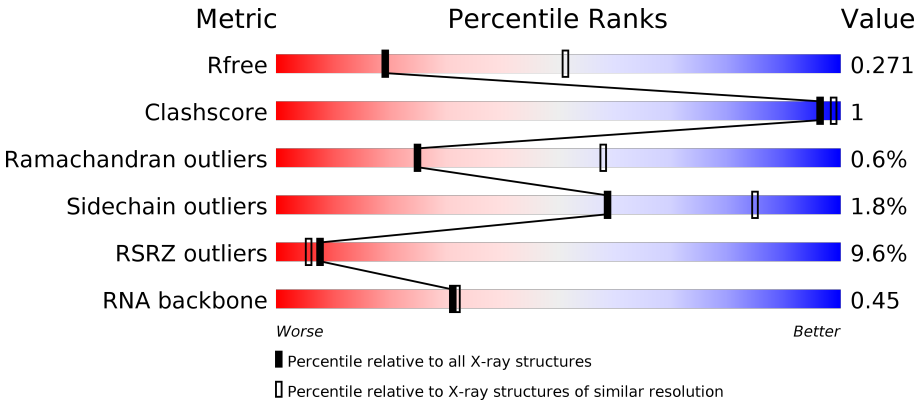
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)
RNA backbone	3102	1007 (3.16-2.64)

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 on 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	738	18% 90% 5% 5%
2	B	776	3% 90% 6%
3	C	809	6% 87% 10%
4	R	18	67% 6% 28%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	V	16	6% 69% 31%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	PO4	A	804	-	-	-	X
7	PO4	C	903	-	-	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 18166 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 Polymerase acidic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	699	Total	C	N	O	S	0	0	0
			5692	3617	959	1079	37			

There are 25 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	GLY	-	expression tag	UNP H6QM92
A	-12	SER	-	expression tag	UNP H6QM92
A	-11	HIS	-	expression tag	UNP H6QM92
A	-10	HIS	-	expression tag	UNP H6QM92
A	-9	HIS	-	expression tag	UNP H6QM92
A	-8	HIS	-	expression tag	UNP H6QM92
A	-7	HIS	-	expression tag	UNP H6QM92
A	-6	HIS	-	expression tag	UNP H6QM92
A	-5	HIS	-	expression tag	UNP H6QM92
A	-4	HIS	-	expression tag	UNP H6QM92
A	-3	GLY	-	expression tag	UNP H6QM92
A	-2	SER	-	expression tag	UNP H6QM92
A	-1	GLY	-	expression tag	UNP H6QM92
A	0	SER	-	expression tag	UNP H6QM92
A	714	GLY	-	expression tag	UNP H6QM92
A	715	SER	-	expression tag	UNP H6QM92
A	716	GLY	-	expression tag	UNP H6QM92
A	717	SER	-	expression tag	UNP H6QM92
A	718	GLY	-	expression tag	UNP H6QM92
A	719	GLU	-	expression tag	UNP H6QM92
A	720	ASN	-	expression tag	UNP H6QM92
A	721	LEU	-	expression tag	UNP H6QM92
A	722	TYR	-	expression tag	UNP H6QM92
A	723	PHE	-	expression tag	UNP H6QM92
A	724	GLN	-	expression tag	UNP H6QM92

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	748	Total	C	N	O	S	0	0	0
			5968	3754	1058	1116	40			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	GLY	-	expression tag	UNP H6QM91
B	-7	SER	-	expression tag	UNP H6QM91
B	-6	GLY	-	expression tag	UNP H6QM91
B	-5	SER	-	expression tag	UNP H6QM91
B	-4	GLY	-	expression tag	UNP H6QM91
B	-3	SER	-	expression tag	UNP H6QM91
B	-2	GLY	-	expression tag	UNP H6QM91
B	-1	SER	-	expression tag	UNP H6QM91
B	0	GLY	-	expression tag	UNP H6QM91
B	757	GLY	-	expression tag	UNP H6QM91
B	758	SER	-	expression tag	UNP H6QM91
B	759	GLY	-	expression tag	UNP H6QM91
B	760	SER	-	expression tag	UNP H6QM91
B	761	GLY	-	expression tag	UNP H6QM91
B	762	GLU	-	expression tag	UNP H6QM91
B	763	ASN	-	expression tag	UNP H6QM91
B	764	LEU	-	expression tag	UNP H6QM91
B	765	TYR	-	expression tag	UNP H6QM91
B	766	PHE	-	expression tag	UNP H6QM91
B	767	GLN	-	expression tag	UNP H6QM91

- Molecule 3 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	727	Total	C	N	O	S	0	0	0
			5774	3644	1021	1077	32			

There are 49 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-8	GLY	-	expression tag	UNP H6QM90
C	-7	SER	-	expression tag	UNP H6QM90
C	-6	GLY	-	expression tag	UNP H6QM90
C	-5	SER	-	expression tag	UNP H6QM90
C	-4	GLY	-	expression tag	UNP H6QM90
C	-3	SER	-	expression tag	UNP H6QM90
C	-2	GLY	-	expression tag	UNP H6QM90

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	SER	-	expression tag	UNP H6QM90
C	0	GLY	-	expression tag	UNP H6QM90
C	761	GLY	-	expression tag	UNP H6QM90
C	762	TRP	-	expression tag	UNP H6QM90
C	763	SER	-	expression tag	UNP H6QM90
C	764	HIS	-	expression tag	UNP H6QM90
C	765	PRO	-	expression tag	UNP H6QM90
C	766	GLN	-	expression tag	UNP H6QM90
C	767	PHE	-	expression tag	UNP H6QM90
C	768	GLU	-	expression tag	UNP H6QM90
C	769	LYS	-	expression tag	UNP H6QM90
C	770	GLY	-	expression tag	UNP H6QM90
C	771	GLY	-	expression tag	UNP H6QM90
C	772	GLY	-	expression tag	UNP H6QM90
C	773	SER	-	expression tag	UNP H6QM90
C	774	GLY	-	expression tag	UNP H6QM90
C	775	GLY	-	expression tag	UNP H6QM90
C	776	GLY	-	expression tag	UNP H6QM90
C	777	SER	-	expression tag	UNP H6QM90
C	778	GLY	-	expression tag	UNP H6QM90
C	779	GLY	-	expression tag	UNP H6QM90
C	780	SER	-	expression tag	UNP H6QM90
C	781	ALA	-	expression tag	UNP H6QM90
C	782	TRP	-	expression tag	UNP H6QM90
C	783	SER	-	expression tag	UNP H6QM90
C	784	HIS	-	expression tag	UNP H6QM90
C	785	PRO	-	expression tag	UNP H6QM90
C	786	GLN	-	expression tag	UNP H6QM90
C	787	PHE	-	expression tag	UNP H6QM90
C	788	GLU	-	expression tag	UNP H6QM90
C	789	LYS	-	expression tag	UNP H6QM90
C	790	GLY	-	expression tag	UNP H6QM90
C	791	ARG	-	expression tag	UNP H6QM90
C	792	SER	-	expression tag	UNP H6QM90
C	793	GLY	-	expression tag	UNP H6QM90
C	794	GLY	-	expression tag	UNP H6QM90
C	795	GLU	-	expression tag	UNP H6QM90
C	796	ASN	-	expression tag	UNP H6QM90
C	797	LEU	-	expression tag	UNP H6QM90
C	798	TYR	-	expression tag	UNP H6QM90
C	799	PHE	-	expression tag	UNP H6QM90
C	800	GLN	-	expression tag	UNP H6QM90

- Molecule 4 is a RNA chain called RNA (5'-R(*UP*AP*UP*AP*CP*CP*UP*CP*UP*GP*CP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	R	13	Total	C	N	O	P	0	0	0
			247	111	37	87	12			

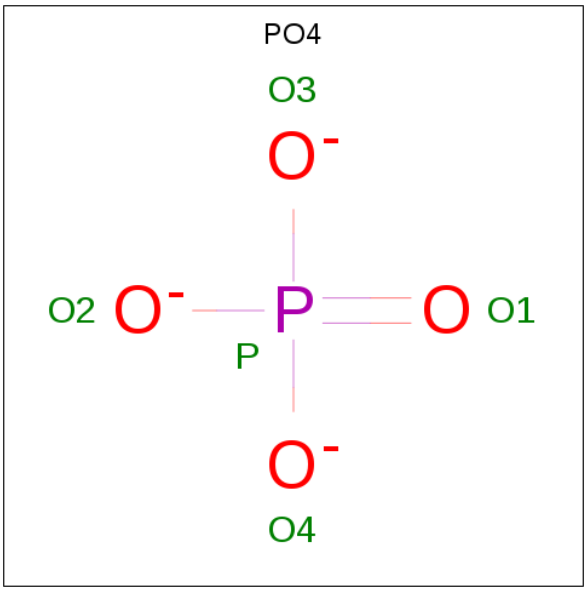
- Molecule 5 is a RNA chain called RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*GP*AP*GP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	V	16	Total	C	N	O	P	0	0	0
			353	157	72	108	16			

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Zn	0	0
			1	1		

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



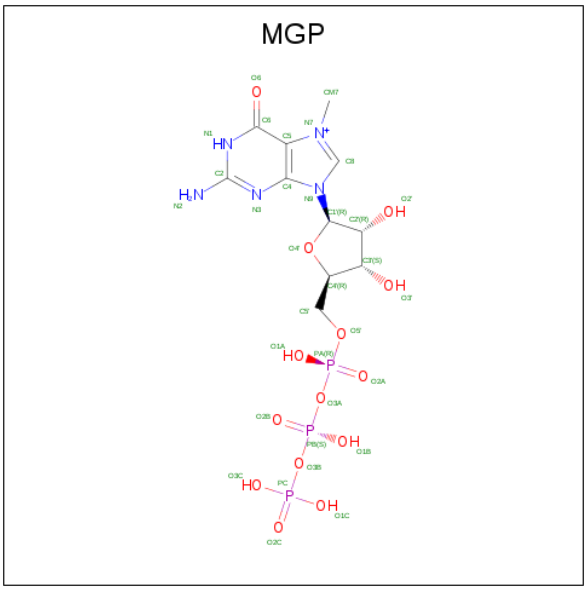
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	O	P	0	0
			5	4	1		
7	A	1	Total	O	P	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O P 5 4 1	0	0
7	A	1	Total O P 5 4 1	0	0
7	B	1	Total O P 5 4 1	0	0
7	B	1	Total O P 5 4 1	0	0
7	B	1	Total O P 5 4 1	0	0
7	B	1	Total O P 5 4 1	0	0
7	B	1	Total O P 5 4 1	0	0
7	B	1	Total O P 5 4 1	0	0
7	C	1	Total O P 5 4 1	0	0
7	C	1	Total O P 5 4 1	0	0

- Molecule 8 is 7-METHYL-GUANOSINE-5'-TRIPHOSPHATE (three-letter code: MGP) (formula: $C_{11}H_{19}N_5O_{14}P_3$) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	C	1	Total C N O P 33 11 5 14 3	0	0

- Molecule 9 is water.

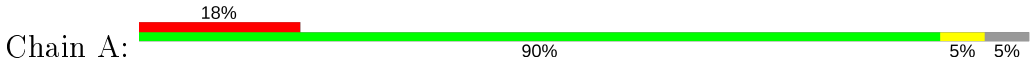
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	7	Total O 7 7	0	0
9	B	16	Total O 16 16	0	0
9	C	11	Total O 11 11	0	0
9	R	1	Total O 1 1	0	0
9	V	3	Total O 3 3	0	0

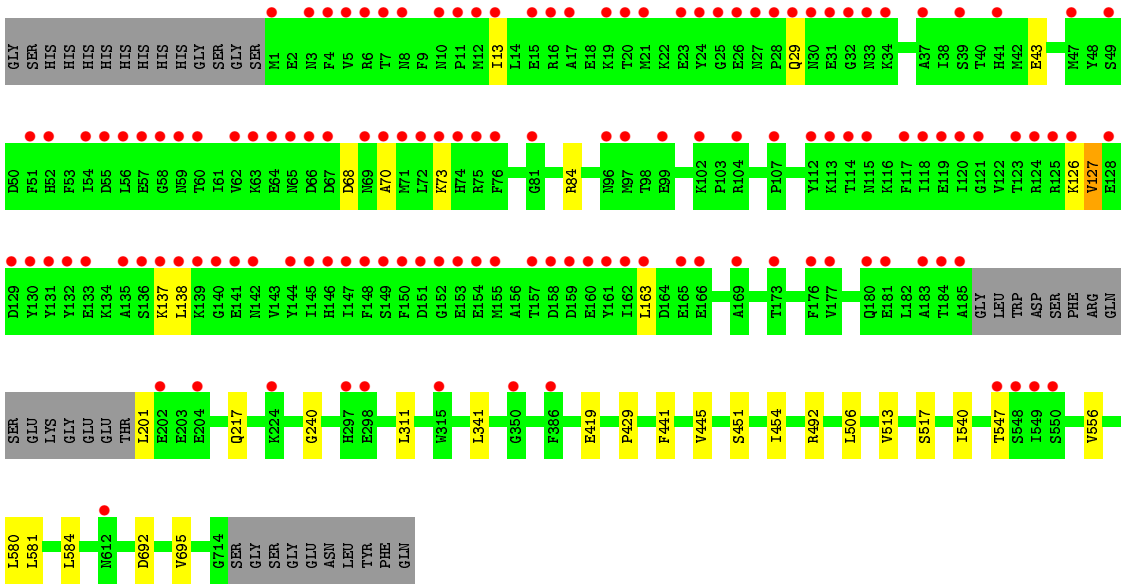
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: Polymerase acidic protein

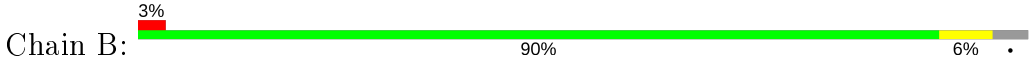
Chain A:

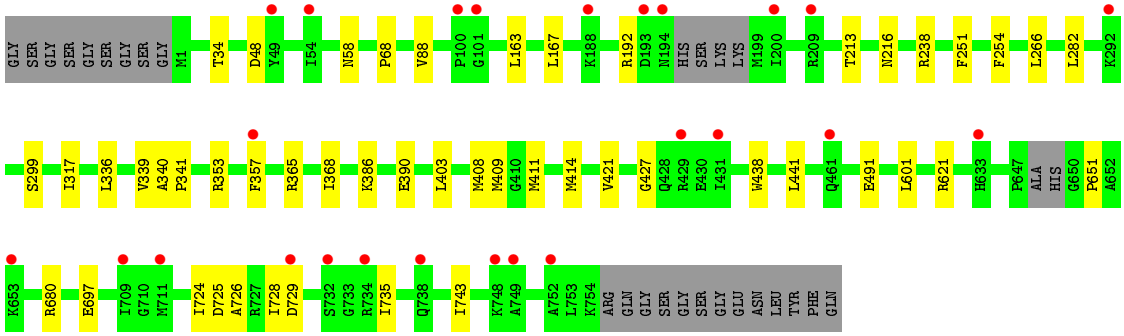




- Molecule 2: RNA-directed RNA polymerase catalytic subunit

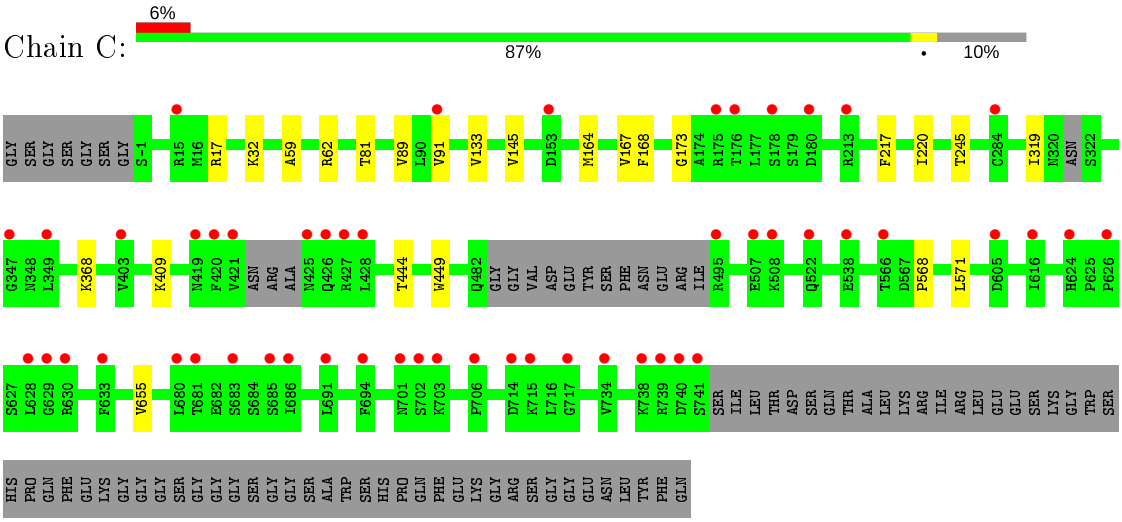
Chain B:





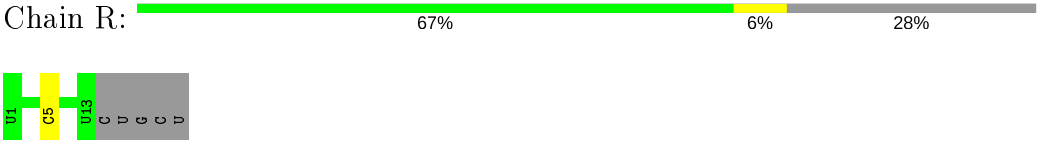
- Molecule 3: Polymerase basic protein 2

Chain C:



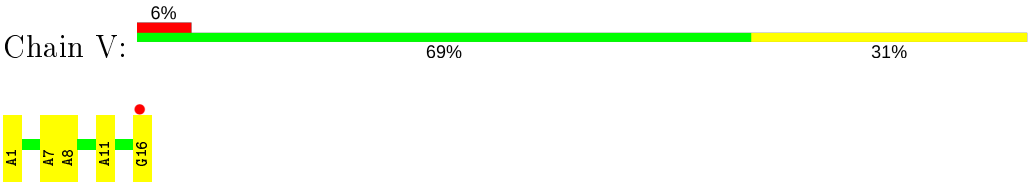
• Molecule 4: RNA (5'-R(*UP*AP*UP*AP*CP*CP*UP*CP*UP*GP*CP*UP*U)-3')

Chain R:



• Molecule 5: RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*GP*AP*GP*GP*G)-3')

Chain V:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	269.34Å 148.70Å 88.51Å 90.00° 98.17° 90.00°	Depositor
Resolution (Å)	50.00 – 2.90 19.98 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.6 (50.00-2.90) 99.9 (19.98-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	1.19 (at 2.88Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.234 , 0.273 0.236 , 0.271	Depositor DCC
R_{free} test set	3755 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	86.7	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 69.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	18166	wwPDB-VP
Average B, all atoms (Å^2)	118.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.92% 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: PO4, MGP, ZN

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	0.43	0/5812	0.59	0/7828
2	B	0.42	0/6082	0.62	0/8207
3	C	0.40	0/5871	0.60	0/7922
4	R	0.25	0/273	0.66	0/421
5	V	0.56	1/397 (0.3%)	0.68	0/617
All	All	0.42	1/18435 (0.0%)	0.61	0/24995

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	V	1	A	OP3-P	-10.16	1.49	1.61

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	5692	0	5613	11	0
2	B	5968	0	5980	18	0
3	C	5774	0	5900	7	0
4	R	247	0	128	0	0
5	V	353	0	175	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	1	0	0	0	0
7	A	20	0	0	0	0
7	B	30	0	0	0	0
7	C	10	0	0	0	0
8	C	33	0	15	0	0
9	A	7	0	0	0	0
9	B	16	0	0	0	0
9	C	11	0	0	0	0
9	R	1	0	0	0	0
9	V	3	0	0	0	0
All	All	18166	0	17811	34	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 1.

All (34) 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)
3:C:568:PRO:O	3:C:571:LEU:N	2.31	0.63
1:A:506:LEU:HD11	1:A:513:VAL:CG2	2.35	0.56
2:B:357:PHE:CE2	2:B:368:ILE:CD1	2.91	0.54
2:B:697:GLU:OE2	3:C:32:LYS:NZ	2.39	0.53
2:B:167:LEU:HD23	2:B:336:LEU:HD11	1.90	0.53
2:B:282:LEU:HG	2:B:441:LEU:HD22	1.91	0.53
3:C:167:VAL:HG23	3:C:168:PHE:CD2	2.44	0.53
1:A:13:ILE:HD11	1:A:43:GLU:HA	1.92	0.51
3:C:409:LYS:HD2	3:C:449:TRP:HA	1.94	0.50
2:B:724:ILE:HG13	2:B:728:ILE:HD13	1.94	0.49
3:C:59:ALA:HB3	3:C:91:VAL:HG21	1.94	0.49
1:A:540:ILE:HG13	1:A:556:VAL:HG23	1.94	0.49
1:A:445:VAL:HG11	1:A:580:LEU:HD22	1.95	0.47
2:B:251:PHE:CZ	2:B:339:VAL:HG11	2.49	0.47
2:B:251:PHE:CE1	2:B:339:VAL:HG11	2.49	0.47
2:B:340:ALA:HB3	2:B:341:PRO:HD3	1.98	0.46
2:B:427:GLY:HA3	2:B:438:TRP:CD1	2.51	0.46
2:B:34:THR:HG22	2:B:353:ARG:HB2	1.96	0.46
3:C:145:VAL:HG21	3:C:220:ILE:HD11	1.96	0.46
2:B:266:LEU:HD13	2:B:421:VAL:HG11	1.98	0.46
2:B:408:MET:HB3	2:B:411:MET:HE3	1.99	0.45
1:A:429:PRO:HB2	3:C:133:VAL:HG23	1.98	0.45
2:B:726:ALA:HB2	2:B:743:ILE:HG21	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:419:GLU:OE1	1:A:419:GLU:N	2.49	0.44
2:B:88:VAL:HG21	2:B:317:ILE:HD12	2.00	0.43
2:B:386:LYS:O	2:B:390:GLU:HG2	2.19	0.43
1:A:506:LEU:HD21	1:A:513:VAL:HG22	2.01	0.42
1:A:692:ASP:HB3	1:A:695:VAL:HG12	2.02	0.42
1:A:451:SER:HA	1:A:454:ILE:HD12	2.02	0.41
1:A:441:PHE:CG	1:A:584:LEU:HD22	2.55	0.41
2:B:729:ASP:HB3	2:B:735:ILE:HG13	2.02	0.41
1:A:419:GLU:HG2	1:A:419:GLU:O	2.21	0.41
2:B:163:LEU:HD23	2:B:254:PHE:CZ	2.56	0.41
2:B:213:THR:O	2:B:216:ASN:N	2.54	0.41

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	695/738 (94%)	634 (91%)	55 (8%)	6 (1%)	17	48
2	B	742/776 (96%)	692 (93%)	46 (6%)	4 (0%)	29	61
3	C	719/809 (89%)	664 (92%)	51 (7%)	4 (1%)	25	58
All	All	2156/2323 (93%)	1990 (92%)	152 (7%)	14 (1%)	25	58

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	126	LYS
1	A	137	LYS
2	B	414	MET
1	A	127	VAL
3	C	319	ILE
3	C	368	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	68	ASP
2	B	651	PRO
2	B	680	ARG
1	A	240	GLY
2	B	68	PRO
1	A	70	ALA
3	C	655	VAL
3	C	173	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	625/657 (95%)	611 (98%)	14 (2%)	52 81
2	B	657/676 (97%)	645 (98%)	12 (2%)	59 85
3	C	644/706 (91%)	636 (99%)	8 (1%)	71 91
All	All	1926/2039 (94%)	1892 (98%)	34 (2%)	59 85

All (34) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	GLN
1	A	73	LYS
1	A	84	ARG
1	A	127	VAL
1	A	138	LEU
1	A	163	LEU
1	A	201	LEU
1	A	217	GLN
1	A	311	LEU
1	A	341	LEU
1	A	492	ARG
1	A	517	SER
1	A	547	THR
1	A	581	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	48	ASP
2	B	58	ASN
2	B	192	ARG
2	B	238	ARG
2	B	299	SER
2	B	365	ARG
2	B	403	LEU
2	B	409	MET
2	B	491	GLU
2	B	601	LEU
2	B	621	ARG
2	B	725	ASP
3	C	17	ARG
3	C	62	ARG
3	C	81	THR
3	C	89	VAL
3	C	164	MET
3	C	217	PHE
3	C	245	THR
3	C	444	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	268	GLN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	R	11/18 (61%)	1 (9%)	0
5	V	15/16 (93%)	4 (26%)	0
All	All	26/34 (76%)	5 (19%)	0

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	R	5	C
5	V	7	A
5	V	8	A
5	V	11	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	V	16	G

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 1 is monoatomic - leaving 13 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	$\#\|Z\| > 2$
8	MGP	C	901	-	28,35,35	2.54	5 (17%)	33,56,56	1.76	7 (21%)
7	PO4	B	803	-	4,4,4	0.88	0	6,6,6	0.40	0
7	PO4	A	803	-	4,4,4	0.93	0	6,6,6	0.49	0
7	PO4	A	804	-	4,4,4	0.93	0	6,6,6	0.39	0
7	PO4	C	902	-	4,4,4	0.87	0	6,6,6	0.44	0
7	PO4	B	804	-	4,4,4	0.94	0	6,6,6	0.52	0
7	PO4	B	805	-	4,4,4	0.98	0	6,6,6	0.33	0
7	PO4	B	801	-	4,4,4	0.92	0	6,6,6	0.47	0
7	PO4	C	903	-	4,4,4	0.93	0	6,6,6	0.45	0
7	PO4	B	806	-	4,4,4	0.90	0	6,6,6	0.49	0
7	PO4	B	802	-	4,4,4	0.87	0	6,6,6	0.53	0
7	PO4	A	802	-	4,4,4	0.92	0	6,6,6	0.45	0
7	PO4	A	805	-	4,4,4	0.85	0	6,6,6	0.69	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
8	MGP	C	901	-	-	0/18/38/38	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	C	901	MGP	C8-N9	7.59	1.47	1.33
8	C	901	MGP	C8-N7	7.33	1.46	1.33
8	C	901	MGP	C5-C4	6.16	1.48	1.39
8	C	901	MGP	C6-C5	3.98	1.48	1.41
8	C	901	MGP	C2'-C1'	-2.10	1.50	1.53

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	C	901	MGP	C5-C6-N1	-5.17	116.36	123.43
8	C	901	MGP	C6-N1-C2	4.14	122.51	115.93
8	C	901	MGP	C2-N3-C4	3.60	119.47	115.36
8	C	901	MGP	N3-C2-N1	-2.58	123.79	127.22
8	C	901	MGP	CM7-N7-C8	-2.54	113.19	125.43
8	C	901	MGP	PB-O3B-PC	-2.33	124.84	132.83
8	C	901	MGP	C6-C5-C4	-2.10	118.79	120.80

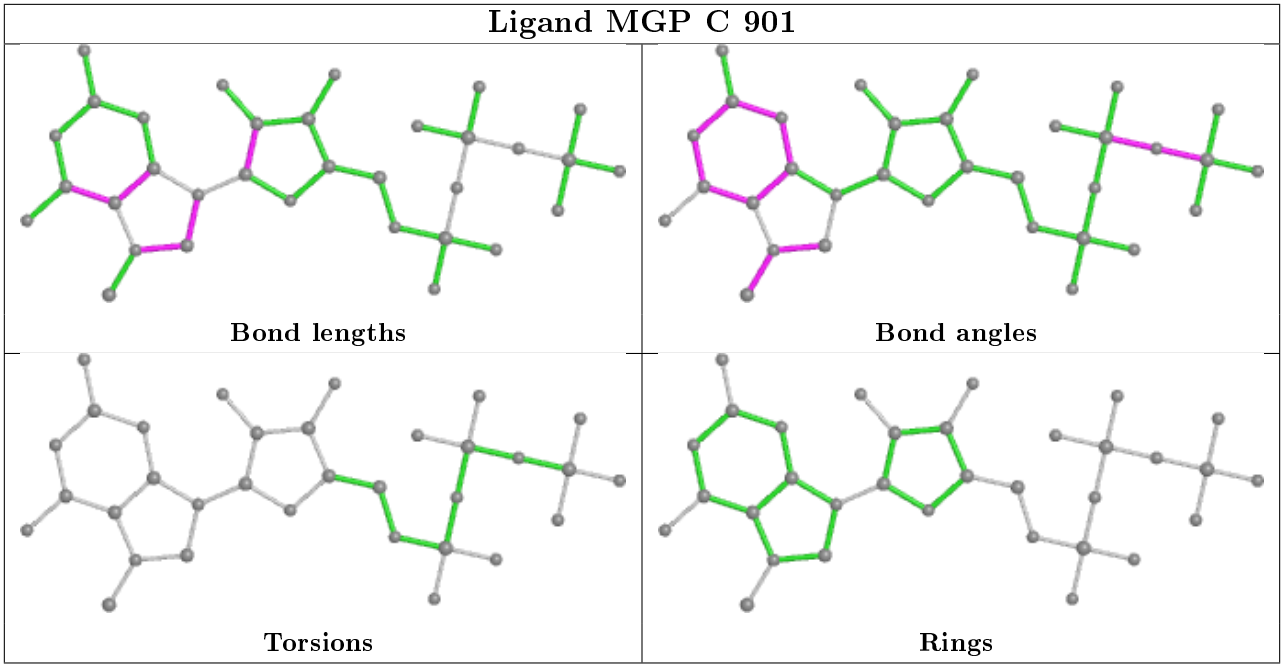
There are no chirality outliers.

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 then 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

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	699/738 (94%)	0.99	134 (19%) 1 0	49, 106, 274, 327	0
2	B	748/776 (96%)	0.12	25 (3%) 46 41	56, 104, 151, 201	0
3	C	727/809 (89%)	0.35	52 (7%) 15 11	61, 117, 163, 193	0
4	R	13/18 (72%)	0.26	0 100 100	72, 89, 119, 126	0
5	V	16/16 (100%)	0.12	1 (6%) 20 16	77, 88, 107, 173	0
All	All	2203/2357 (93%)	0.48	212 (9%) 8 6	49, 108, 237, 327	0

All (212) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	71	MET	12.5
1	A	146	HIS	11.0
1	A	70	ALA	10.3
1	A	51	PHE	9.5
1	A	16	ARG	9.4
1	A	114	THR	9.3
1	A	21	MET	9.2
1	A	20	THR	9.0
1	A	126	LYS	9.0
1	A	25	GLY	8.9
1	A	549	ILE	8.9
1	A	176	PHE	8.8
1	A	129	ASP	8.7
1	A	24	TYR	8.7
1	A	121	GLY	7.9
1	A	147	ILE	7.9
1	A	26	GLU	7.5
1	A	142	ASN	7.4
1	A	162	ILE	7.2
1	A	75	ARG	7.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	184	THR	7.0
1	A	136	SER	6.9
1	A	72	LEU	6.6
1	A	59	ASN	6.5
1	A	13	ILE	6.2
1	A	30	ASN	6.1
1	A	33	ASN	6.1
1	A	181	GLU	6.1
1	A	315	TRP	6.0
1	A	66	ASP	6.0
1	A	115	ASN	6.0
1	A	140	GLY	5.9
1	A	157	THR	5.9
3	C	628	LEU	5.9
1	A	145	ILE	5.8
1	A	63	LYS	5.8
1	A	135	ALA	5.8
1	A	177	VAL	5.8
1	A	58	GLY	5.7
1	A	548	SER	5.7
1	A	132	TYR	5.5
1	A	55	ASP	5.5
1	A	159	ASP	5.5
1	A	60	THR	5.5
1	A	183	ALA	5.5
1	A	28	PRO	5.3
1	A	185	ALA	5.3
1	A	123	THR	5.1
3	C	715	LYS	5.1
1	A	49	SER	5.0
1	A	67	ASP	5.0
1	A	32	GLY	5.0
1	A	180	GLN	4.9
3	C	681	THR	4.9
1	A	27	ASN	4.8
3	C	703	LYS	4.8
3	C	741	SER	4.7
1	A	99	GLU	4.7
1	A	128	GLU	4.7
1	A	12	MET	4.7
1	A	130	TYR	4.6
1	A	65	ASN	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	144	TYR	4.5
1	A	57	GLU	4.5
1	A	118	ILE	4.5
1	A	550	SER	4.5
1	A	297	HIS	4.4
3	C	706	PRO	4.4
1	A	153	GLU	4.4
1	A	158	ASP	4.4
1	A	34	LYS	4.4
1	A	163	LEU	4.2
3	C	685	SER	4.2
1	A	165	GLU	4.2
1	A	173	THR	4.2
1	A	125	ARG	4.2
1	A	151	ASP	4.2
1	A	124	ARG	4.2
1	A	97	MET	4.1
1	A	149	SER	4.1
1	A	1	MET	4.1
1	A	547	THR	4.1
1	A	141	GLU	4.0
3	C	734	VAL	4.0
1	A	155	MET	4.0
1	A	139	LYS	4.0
3	C	714	ASP	4.0
3	C	495	ARG	4.0
2	B	752	ALA	4.0
1	A	119	GLU	3.9
1	A	137	LYS	3.9
2	B	633	HIS	3.9
3	C	428	LEU	3.9
1	A	96	ASN	3.9
1	A	54	ILE	3.8
1	A	7	THR	3.8
1	A	56	LEU	3.8
2	B	653	LYS	3.8
1	A	169	ALA	3.7
3	C	347	GLY	3.7
3	C	426	GLN	3.7
3	C	680	LEU	3.7
5	V	16	G	3.7
1	A	17	ALA	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	120	ILE	3.6
3	C	605	ASP	3.5
2	B	738	GLN	3.5
1	A	104	ARG	3.5
1	A	47	MET	3.5
1	A	64	GLU	3.5
1	A	19	LYS	3.5
1	A	113	LYS	3.5
1	A	76	PHE	3.5
3	C	633	PHE	3.4
1	A	15	GLU	3.4
1	A	161	TYR	3.4
3	C	419	ASN	3.4
3	C	175	ARG	3.4
3	C	740	ASP	3.3
3	C	15	ARG	3.3
3	C	739	ARG	3.3
1	A	298	GLU	3.3
2	B	709	ILE	3.2
2	B	200	ILE	3.2
3	C	176	THR	3.1
2	B	188	LYS	3.1
1	A	166	GLU	3.1
1	A	74	HIS	3.1
1	A	10	ASN	3.1
3	C	629	GLY	3.1
1	A	62	VAL	3.1
3	C	738	LYS	3.0
1	A	386	PHE	3.0
3	C	522	GLN	3.0
1	A	69	ASN	3.0
2	B	54	ILE	3.0
1	A	133	GLU	2.9
3	C	630	ARG	2.9
1	A	52	HIS	2.9
2	B	748	LYS	2.8
3	C	153	ASP	2.8
3	C	683	SER	2.8
3	C	180	ASP	2.8
1	A	3	ASN	2.8
3	C	178	SER	2.8
1	A	41	HIS	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	37	ALA	2.8
3	C	624	HIS	2.7
3	C	538	GLU	2.7
3	C	427	ARG	2.7
1	A	160	GLU	2.7
1	A	8	ASN	2.7
1	A	150	PHE	2.6
3	C	420	PHE	2.6
1	A	117	PHE	2.6
2	B	193	ASP	2.6
3	C	626	PRO	2.6
3	C	686	ILE	2.5
2	B	100	PRO	2.5
3	C	694	PHE	2.5
1	A	6	ARG	2.5
2	B	732	SER	2.5
1	A	612	ASN	2.5
1	A	11	PRO	2.5
3	C	425	ASN	2.5
3	C	421	VAL	2.5
1	A	81	GLY	2.5
1	A	202	GLU	2.4
1	A	224	LYS	2.4
3	C	566	THR	2.4
1	A	4	PHE	2.4
3	C	701	ASN	2.4
2	B	749	ALA	2.4
2	B	357	PHE	2.4
3	C	717	GLY	2.4
1	A	204	GLU	2.3
1	A	131	TYR	2.3
2	B	101	GLY	2.3
1	A	138	LEU	2.3
3	C	702	SER	2.3
3	C	403	VAL	2.3
1	A	31	GLU	2.3
1	A	107	PRO	2.3
2	B	292	LYS	2.3
1	A	5	VAL	2.3
2	B	49	TYR	2.2
1	A	152	GLY	2.2
2	B	209	ARG	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	616	ILE	2.2
1	A	29	GLN	2.2
1	A	23	GLU	2.2
1	A	350	GLY	2.2
3	C	508	LYS	2.2
3	C	691	LEU	2.2
2	B	734	ARG	2.2
1	A	154	GLU	2.1
1	A	73	LYS	2.1
3	C	91	VAL	2.1
1	A	148	PHE	2.1
3	C	507	GLU	2.1
1	A	39	SER	2.1
3	C	284	CYS	2.1
2	B	729	ASP	2.1
2	B	194	ASN	2.0
3	C	349	LEU	2.0
2	B	429	ARG	2.0
1	A	112	TYR	2.0
2	B	431	ILE	2.0
2	B	711	MET	2.0
2	B	461	GLN	2.0
3	C	213	ARG	2.0
1	A	102	LYS	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 carbohydrates in this entry.

6.4 Ligands (i)

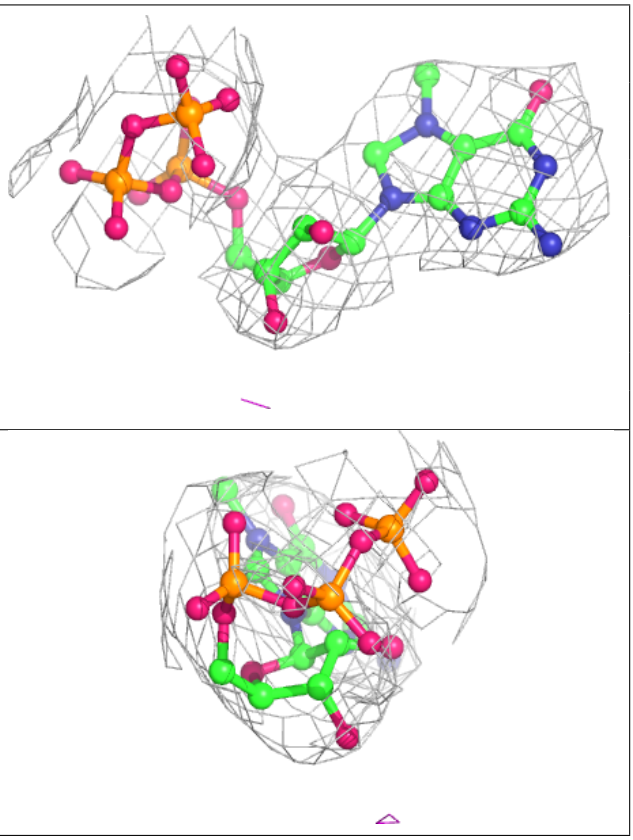
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
7	PO4	A	804	5/5	0.39	0.69	197,198,201,206	0
7	PO4	C	903	5/5	0.78	0.41	153,156,159,162	0
7	PO4	A	803	5/5	0.79	0.27	130,141,144,150	0
7	PO4	C	902	5/5	0.87	0.21	126,127,138,139	0
7	PO4	B	806	5/5	0.88	0.38	133,138,143,145	0
7	PO4	B	801	5/5	0.90	0.12	105,107,113,115	0
7	PO4	B	804	5/5	0.90	0.24	108,116,121,123	0
7	PO4	A	805	5/5	0.91	0.19	110,119,120,122	0
7	PO4	B	803	5/5	0.92	0.12	123,124,125,129	0
7	PO4	B	802	5/5	0.92	0.17	110,111,115,118	0
8	MGP	C	901	33/33	0.93	0.14	71,94,160,163	0
7	PO4	A	802	5/5	0.94	0.10	102,104,108,110	0
7	PO4	B	805	5/5	0.96	0.17	110,115,118,122	0
6	ZN	A	801	1/1	0.98	0.09	90,90,90,90	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 MGP C 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.