

wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 06:15 am BST

PDB ID : 5DHB

Title : Cooperativity and Downstream Binding in RNA Replication Authors : Zhang, W.; Fahrenbach, A.C.; Tam, C.P.; Szostak, J.W.

Deposited on : 2015-08-30

Resolution : 1.80 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 following versions of software and data (see references (1)) 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)

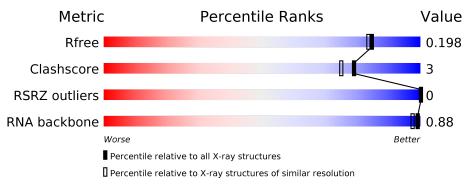
Validation Pipeline (wwPDB-VP) : 2.11

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 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)
RNA backbone	3102	1060 (2.40-1.20)

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	7	71%	29%					
1	В	7	71%	29%					
1	С	7	57%	43%					
1	D	7	57%	43%					



2 Entry composition (i)

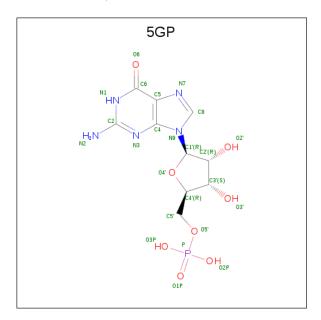
There are 3 unique types of molecules in this entry. The entry contains 773 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 RNA chain called RNA (5'-R(*(LCC)P*(TLN)P*(LCG)P*UP*AP*CP*A)-3').

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Λ	7	Total	С	N	О	Р	0	0	0
1	A	1	149	71	25	47	6			
1	B	7	Total	С	N	О	Р	0	0	0
1	Ъ	4	149	71	25	47	6			
1	С	C 7	Total	С	N	О	Р	0	0	0
1			149	71	25	47	6	0		U
1	1 D	D 7	Total	С	N	О	Р	0	0	0
	(149	71	25	47	6	U	U	U	

• Molecule 2 is GUANOSINE-5'-MONOPHOSPHATE (three-letter code: 5GP) (formula: $C_{10}H_{14}N_5O_8P$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
2	Δ	1	Total	С	N	О	Р	0	0
	Λ	1	24	10	5	8	1	U	

Continued on next page...



Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
2	D	1	Total	С	N	Ο	Р	0	0	
	В	1	24	10	5	8	1	U	0	
2	2 C	С	C 1	Total	С	N	О	Р	0	0
		1	24	10	5	8	1	0		
9	D	1	Total	С	N	О	Р	0	0	
		1	24	10	5	8	1	U	U	

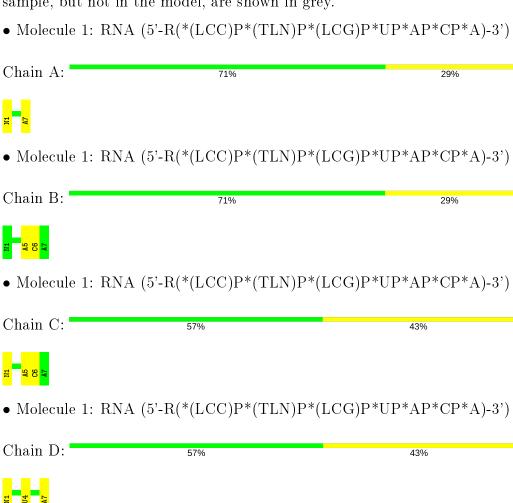
• Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	23	Total O 23 23	0	0
3	В	18	Total O 18 18	0	0
3	С	18	Total O 18 18	0	0
3	D	22	Total O 22 22	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.





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants	51.25	Danasitan
a, b, c, α , β , γ	90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.38 - 1.80	Depositor
Resolution (A)	38.76 - 1.80	EDS
% Data completeness	98.1 (44.38-1.80)	Depositor
(in resolution range)	98.1 (38.76-1.80)	EDS
R_{merge}	0.08	Depositor
R_{sum}	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.26 (at 1.81Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.187 , 0.192	Depositor
It, It free	0.190 , 0.198	DCC
R_{free} test set	622 reflections (6.01%)	wwPDB-VP
Wilson B-factor (Å ²)	32.9	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.26, 42.3	EDS
L-test for twinning ²	$< L >=0.53, < L^2>=0.37$	Xtriage
	0.489 for -h,-k,l	
Estimated twinning fraction	0.067 for h,-h-k,-l	Xtriage
	0.065 for -k,-h,-l	
F_o, F_c correlation	0.97	EDS
Total number of atoms	773	wwPDB-VP
Average B, all atoms (\mathring{A}^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 9.59% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: LCG, 5GP, LCC, TLN

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	Bon	nd lengths	Bond angles		
MIOI	Chain	RMSZ	$RMSZ \mid \# Z > 5$		# Z >5	
1	A	0.78	0/93	0.91	$1/142 \ (0.7\%)$	
1	В	0.72	0/93	0.81	0/142	
1	С	0.72	0/93	0.86	0/142	
1	D	0.79	1/93 (1.1%)	0.90	1/142 (0.7%)	
All	All	0.75	1/372~(0.3%)	0.87	$2/568 \ (0.4\%)$	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(\mathbf{\mathring{A}})$	$\operatorname{Ideal}(ext{\AA})$
1	D	4	U	O3'-P	5.18	1.67	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^o)$
1	D	7	A	O5'-P-OP2	-6.35	99.99	105.70
1	A	7	A	O5'-P-OP2	-5.00	101.20	105.70

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	$\mathbf{H}(\mathbf{model})$	H(added)	Clashes	Symm-Clashes
1	A	149	0	80	0	0

Continued on next page...



$\alpha \cdots \tau$	e	•	
Continued	trom	mraniaone	maaa
-	110116	predidus	puyc

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	В	149	0	80	1	0
1	С	149	0	80	2	0
1	D	149	0	80	0	0
2	A	24	0	12	0	0
2	В	24	0	12	0	0
2	С	24	0	12	0	0
2	D	24	0	12	0	0
3	A	23	0	0	0	0
3	В	18	0	0	0	0
3	С	18	0	0	0	0
3	D	22	0	0	0	0
All	All	773	0	368	3	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.

All (3) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$egin{array}{ll} ext{Interatomic} \ ext{distance} \ (ext{\AA}) \end{array}$	Clash overlap (Å)	
1:B:5:A:H2'	1:B:6:C:O4'	2.19	0.43	
1:C:1:LCC:H41	1:C:1:LCC:H5M1	1.75	0.42	
1:C:5:A:H2'	1:C:6:C:O4'	2.19	0.42	

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein molecules in this entry.

5.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

5.3.3 RNA (i)



Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	4/7 (57%)	0	1 (25%)
1	В	3/7 (42%)	0	0
1	С	3/7 (42%)	0	0
1	D	4/7 (57%)	0	1 (25%)
All	All	14/28 (50%)	0	2 (14%)

There are no RNA backbone outliers to report.

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1	LCC
1	D	1	LCC

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

12 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).

N / L 1	TD.	Cl '	D	т. 1	Во	nd leng	$_{ m ths}$	В	ond ang	les
Mol	Type	Chain	Res	Link	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
1	TLN	D	2	1	16,24,25	1.87	3 (18%)	19,37,40	2.22	3 (15%)
1	LCC	D	1	1	16,21,25	2.92	7 (43%)	24,33,40	2.45	8 (33%)
1	LCC	A	1	1	16,21,25	2.69	7 (43%)	24,33,40	2.44	8 (33%)
1	LCC	С	1	1	16,21,25	2.40	5 (31%)	24,33,40	1.94	6 (25%)
1	TLN	С	2	1	16,24,25	1.87	6 (37%)	19,37,40	2.37	5 (26%)
1	TLN	В	2	1	16,24,25	1.78	5 (31%)	19,37,40	2.27	5 (26%)
1	LCG	С	3	1	19,27,28	3.12	7 (36%)	23,42,45	2.19	8 (34%)
1	LCG	A	3	1	19,27,28	2.52	6 (31%)	23,42,45	2.13	7 (30%)
1	LCG	В	3	1	19,27,28	2.81	7 (36%)	23,42,45	1.98	7 (30%)
1	LCG	D	3	1	19,27,28	2.50	7 (36%)	23,42,45	2.06	7 (30%)
1	TLN	A	2	1	16,24,25	1.92	3 (18%)	19,37,40	2.11	2 (10%)
1	LCC	В	1	1	16,21,25	2.21	4 (25%)	24,33,40	1.92	8 (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
1	TLN	D	2	1	-	0/6/35/36	0/4/3/3
1	LCC	D	1	1	-	0/5/32/36	0/4/3/3
1	LCC	A	1	1	-	0/5/32/36	0/4/3/3
1	LCC	С	1	1	-	0/5/32/36	0/4/3/3
1	TLN	С	2	1	-	0/6/35/36	0/4/3/3
1	TLN	В	2	1	-	0/6/35/36	0/4/3/3
1	LCG	С	3	1	-	0/4/35/36	0/5/4/4
1	LCG	A	3	1	-	0/4/35/36	0/5/4/4
1	LCG	В	3	1	-	0/4/35/36	0/5/4/4
1	LCG	D	3	1	-	0/4/35/36	0/5/4/4
1	TLN	A	2	1	-	0/6/35/36	0/4/3/3
1	LCC	В	1	1	_	0/5/32/36	0/4/3/3

The worst 5 of 67 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	$\operatorname{Ideal}(ext{\AA})$
1	С	3	LCG	C4'-C3'	-7.99	1.44	1.53
1	В	3	LCG	C4'-C3'	-6.83	1.45	1.53
1	С	3	LCG	O2'-C6'	6.25	1.55	1.43
1	С	3	LCG	O4'-C1'	5.98	1.51	1.42
1	D	1	LCC	O2'-C6'	5.91	1.54	1.43

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^o)$
1	D	2	TLN	C4-N3-C2	8.12	122.00	115.14
1	A	2	TLN	C4-N3-C2	7.78	121.71	115.14
1	С	2	TLN	C4-N3-C2	7.75	121.69	115.14
1	В	2	TLN	C4-N3-C2	7.17	121.20	115.14
1	D	1	LCC	O4'-C4'-C5'	6.37	120.50	108.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	С	1	LCC	1	0

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

4 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	Mol Type		Res	Link	Вс	nd leng	ths	Bond angles		
MIGI	туре	Chain	nes	LIIIK	Counts	RMSZ	$\mid \# Z > 2$	Counts	RMSZ	# Z >2
2	5GP	В	101	-	22,26,26	2.20	7 (31%)	27,40,40	2.13	11 (40%)
2	5GP	D	101	-	22,26,26	2.01	5 (22%)	27,40,40	2.11	9 (33%)
2	5GP	A	101	-	22,26,26	2.01	6 (27%)	27,40,40	2.11	10 (37%)
2	5GP	С	101	-	22,26,26	2.15	7 (31%)	27,40,40	2.23	8 (29%)

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	5GP	В	101	_	_	0/6/26/26	0/3/3/3
2	5GP	D	101	_	_	0/6/26/26	0/3/3/3
2	5GP	A	101	-	-	0/6/26/26	0/3/3/3
2	5GP	С	101	-	-	0/6/26/26	0/3/3/3

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	$Ideal(\AA)$
2	D	101	5GP	C6-N1	6.00	1.43	1.33
2	В	101	5GP	C6-N1	5.12	1.42	1.33

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	$\operatorname{Ideal}(ext{\AA})$
2	A	101	5GP	C6-N1	4.81	1.41	1.33
2	A	101	5GP	O4'-C1'	4.80	1.47	1.41
2	С	101	5GP	C6-N1	4.73	1.41	1.33

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^o)$
2	С	101	5GP	N3-C2-N1	-5.66	119.68	127.22
2	С	101	5GP	C5-C6-N1	-5.34	116.12	123.43
2	В	101	5GP	N3-C2-N1	-5.29	120.17	127.22
2	A	101	5GP	C1'-N9-C4	-4.99	117.88	126.64
2	A	101	5GP	N3-C2-N1	-4.90	120.69	127.22

There are no chirality outliers.

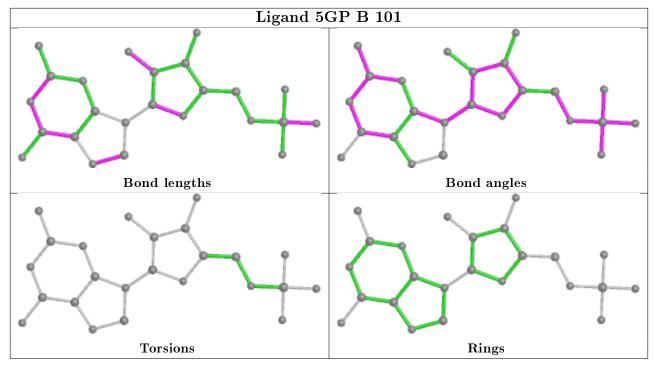
There are no torsion outliers.

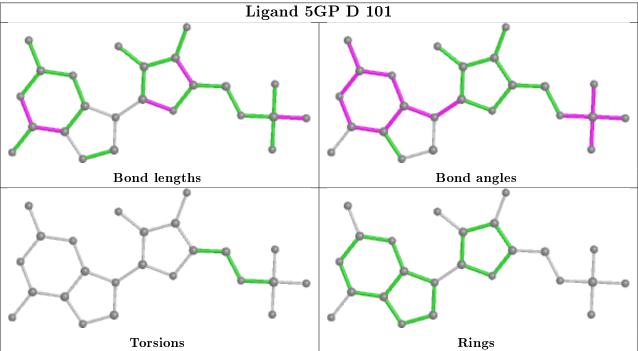
There are no ring outliers.

No monomer is involved in short contacts.

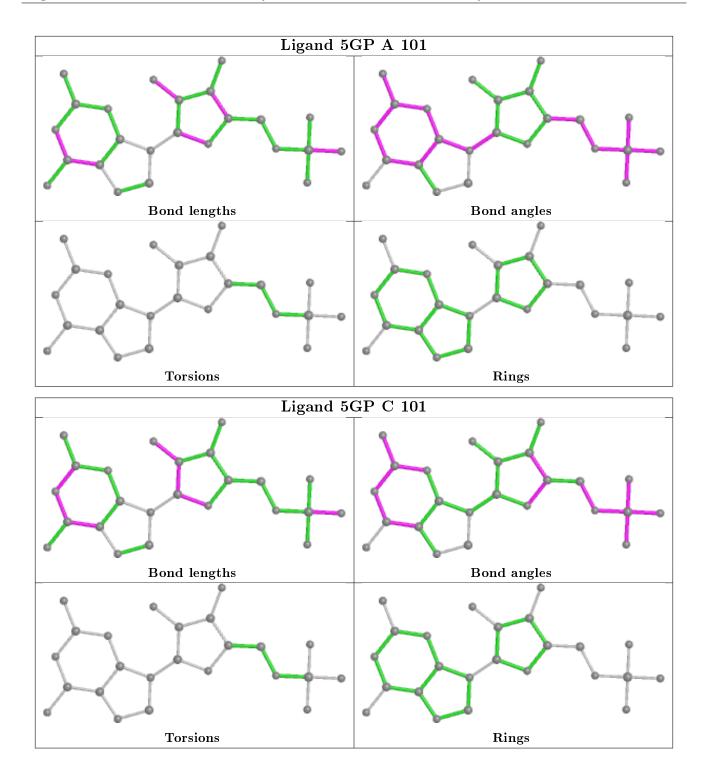
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 (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 $>$	$\#RSRZ{>}2$		Z>2	$OWAB(A^2)$	Q < 0.9
1	A	4/7~(57%)	-1.19	0	100	100	28, 31, 39, 42	0
1	В	4/7~(57%)	-1.18	0	100	100	28, 30, 33, 35	0
1	С	4/7~(57%)	-1.19	0	100	100	27, 30, 33, 35	0
1	D	4/7 (57%)	-1.17	0	100	100	28, 31, 39, 41	0
All	All	16/28~(57%)	-1.18	0	100	100	27, 31, 41, 42	0

There are no RSRZ outliers to report.

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, 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	${f B\text{-factors}}({f \AA}^2)$	Q<0.9
1	LCC	D	1	19/23	0.97	0.05	28,31,37,41	0
1	LCC	A	1	19/23	0.97	0.06	29,31,36,41	0
1	LCC	С	1	19/23	0.97	0.05	30,32,37,42	0
1	TLN	D	2	22/23	0.98	0.05	25,28,33,39	0
1	TLN	С	2	22/23	0.98	0.05	31,35,38,44	0
1	TLN	В	2	22/23	0.98	0.05	31,34,40,44	0
1	LCG	С	3	24/25	0.98	0.05	29,35,40,43	0
1	LCG	A	3	24/25	0.98	0.05	23,27,30,31	0
1	LCG	В	3	24/25	0.98	0.05	28,35,40,44	0
1	LCG	D	3	24/25	0.98	0.05	23,27,30,30	0
1	TLN	A	2	22/23	0.98	0.05	25,28,32,39	0
1	LCC	В	1	19/23	0.98	0.04	30,32,38,42	0



6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

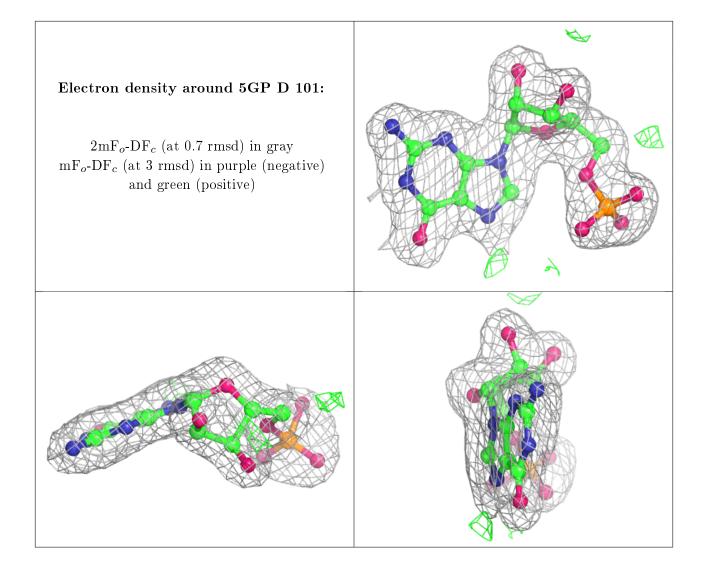
6.4 Ligands (i)

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.

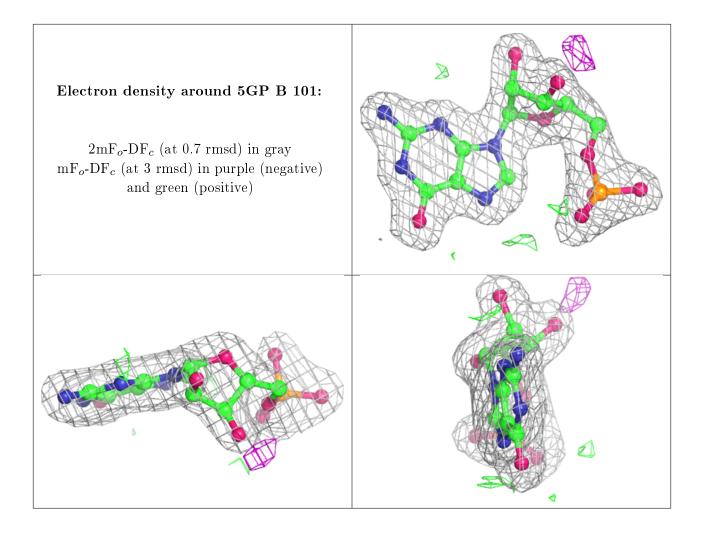
Mol	Type	Chain	${f Res}$	Atoms	RSCC	RSR	${f B-factors}({f A}^2)$	Q<0.9
2	5GP	D	101	24/24	0.96	0.06	30,37,67,68	0
2	5GP	В	101	24/24	0.97	0.06	25,31,49,57	0
2	5GP	A	101	24/24	0.97	0.06	31,36,66,69	0
2	5GP	С	101	24/24	0.98	0.06	26,32,48,57	0

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.

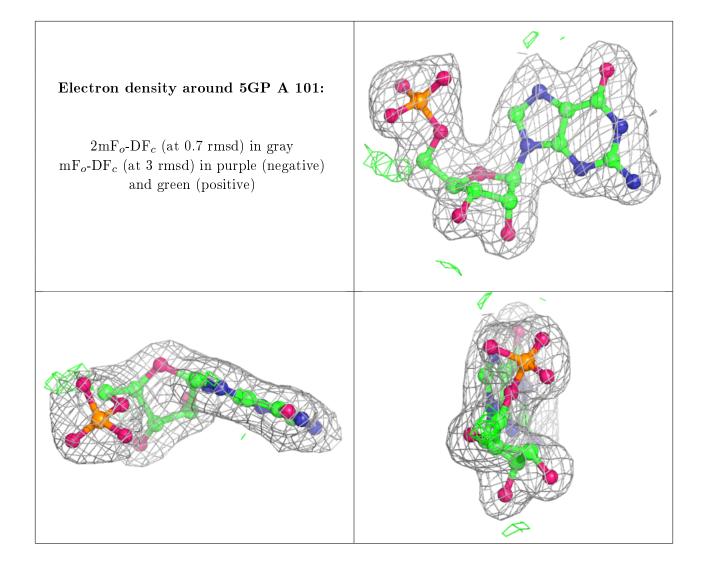




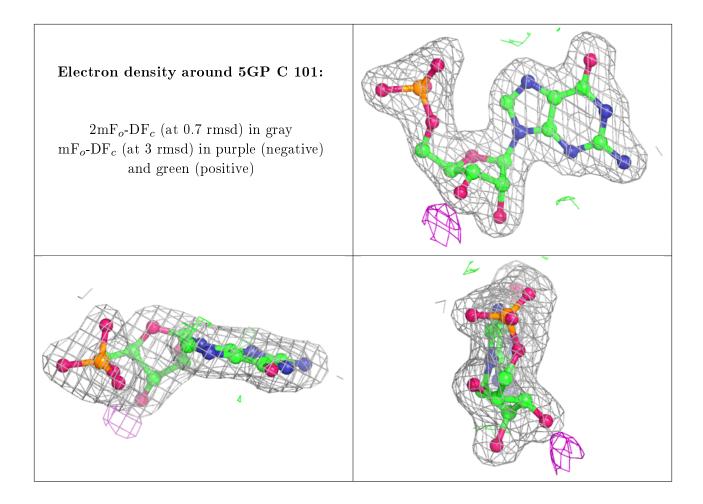












6.5 Other polymers (i)

There are no such residues in this entry.

