



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 28, 2016 – 02:54 PM EST

PDB ID : 5DHC  
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 : 1.55 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20028320  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028320

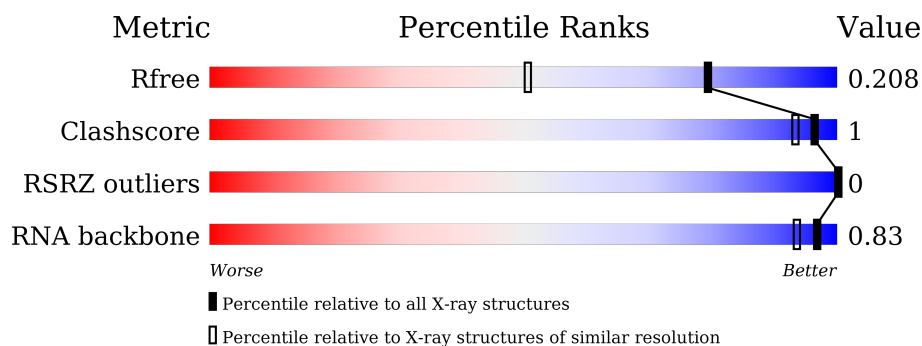
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.55 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))
$R_{free}$	91344	1665 (1.58-1.54)
Clashscore	102246	1014 (1.56-1.56)
RSRZ outliers	91569	1668 (1.58-1.54)
RNA backbone	2183	1046 (2.70-0.62)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	14	<div> <div style="width: 79%; background-color: green;"></div> <div style="width: 21%; background-color: yellow;"></div> </div> <div>79% 21%</div>
1	B	14	<div> <div style="width: 79%; background-color: green;"></div> <div style="width: 21%; background-color: yellow;"></div> </div> <div>79% 21%</div>
1	C	14	<div> <div style="width: 79%; background-color: green;"></div> <div style="width: 21%; background-color: yellow;"></div> </div> <div>79% 21%</div>
1	D	14	<div> <div style="width: 86%; background-color: green;"></div> <div style="width: 14%; background-color: yellow;"></div> </div> <div>86% 14%</div>

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
2	MG	A	102	-	-	-	X
2	MG	C	102	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1679 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.

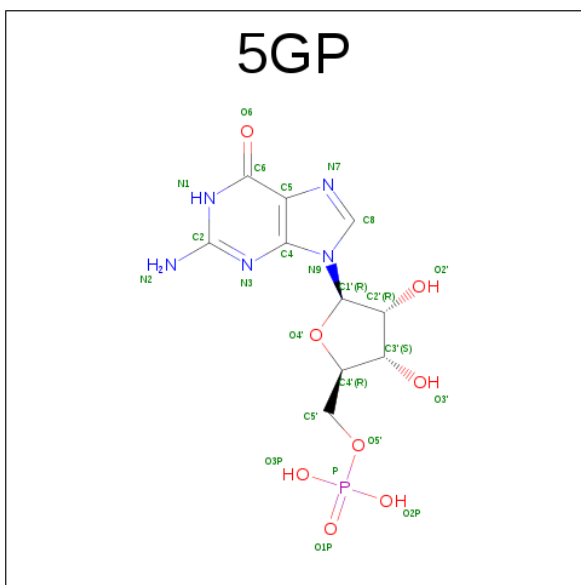
- Molecule 1 is a RNA chain called RNA (5'-R\*(LCC)P\*(LCC)P\*(LCA)P\*(LCG)P\*AP\*CP\*UP\*UP\*AP\*AP\*GP\*UP\*CP\*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	14	Total	C	N	O	P	0	0	0
			297	138	50	96	13			
1	B	14	Total	C	N	O	P	0	0	0
			297	138	50	96	13			
1	C	14	Total	C	N	O	P	0	0	0
			297	138	50	96	13			
1	D	14	Total	C	N	O	P	0	0	0
			297	138	50	96	13			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	4	Total	Mg	0	0
			4	4		
2	A	3	Total	Mg	0	0
			3	3		
2	D	4	Total	Mg	0	0
			4	4		
2	C	4	Total	Mg	0	0
			4	4		

- Molecule 3 is GUANOSINE-5'-MONOPHOSPHATE (three-letter code: 5GP) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>5</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	A	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	B	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	B	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	C	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	C	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	D	1	Total 24	C 10	N 5	O 8	P 1	0	0
3	D	1	Total 24	C 10	N 5	O 8	P 1	0	0


- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	66	Total O 66 66	0	0
4	B	73	Total O 73 73	0	0
4	C	69	Total O 69 69	0	0
4	D	76	Total O 76 76	0	0

### 3 Residue-property plots


These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: RNA (5'-R\*(LCC)P\*(LCC)P\*(LCA)P\*(LCG)P\*AP\*CP\*UP\*UP\*AP\*AP\*GP\*UP\*CP\*U)-3')

Chain A:  79% 21%




- Molecule 1: RNA (5'-R\*(LCC)P\*(LCC)P\*(LCA)P\*(LCG)P\*AP\*CP\*UP\*UP\*AP\*AP\*GP\*UP\*CP\*U)-3')

Chain B:  79% 21%



- Molecule 1: RNA (5'-R\*(LCC)P\*(LCC)P\*(LCA)P\*(LCG)P\*AP\*CP\*UP\*UP\*AP\*AP\*GP\*UP\*CP\*U)-3')

Chain C:  79% 21%



- Molecule 1: RNA (5'-R\*(LCC)P\*(LCC)P\*(LCA)P\*(LCG)P\*AP\*CP\*UP\*UP\*AP\*AP\*GP\*UP\*CP\*U)-3')

Chain D:  86% 14%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	43.11Å 43.11Å 83.25Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 1.55 41.62 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-1.55) 99.9 (41.62-1.55)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.31 (at 1.55Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, $R_{free}$	0.170 , 0.195 0.177 , 0.208	Depositor DCC
$R_{free}$ test set	1300 reflections (5.45%)	DCC
Wilson B-factor (Å <sup>2</sup> )	14.8	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 38.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	0.021 for -h,-k,l 0.478 for h,-h-k,-l 0.022 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1679	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.71% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: LCG, MG, 5GP, LCC, LCA

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	1.17	1/232 (0.4%)	1.02	0/358
1	B	1.15	2/232 (0.9%)	1.07	0/358
1	C	1.15	1/232 (0.4%)	1.04	1/358 (0.3%)
1	D	1.14	1/232 (0.4%)	1.03	0/358
All	All	1.15	5/928 (0.5%)	1.04	1/1432 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	7	U	O3'-P	-8.08	1.51	1.61
1	A	7	U	O3'-P	-7.95	1.51	1.61
1	B	10	A	O3'-P	-6.36	1.53	1.61
1	D	10	A	O3'-P	-5.99	1.53	1.61
1	B	7	U	O3'-P	-5.79	1.54	1.61

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	10	A	O5'-P-OP2	-5.04	101.17	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	H(model)	H(added)	Clashes	Symm-Clashes
1	A	297	0	155	1	0
1	B	297	0	155	0	0
1	C	297	0	155	0	0
1	D	297	0	155	0	0
2	A	3	0	0	0	0
2	B	4	0	0	0	0
2	C	4	0	0	1	0
2	D	4	0	0	0	0
3	A	48	0	24	0	0
3	B	48	0	24	1	0
3	C	48	0	24	1	0
3	D	48	0	24	0	0
4	A	66	0	0	0	0
4	B	73	0	0	0	0
4	C	69	0	0	1	0
4	D	76	0	0	0	0
All	All	1679	0	716	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 1.

All (3) 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:106:5GP:H8	3:C:106:5GP:H5'2	1.76	0.66
2:C:104:MG:MG	4:C:268:HOH:O	1.45	0.58
1:A:14:U:H2'	3:B:105:5GP:H4'	2.01	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 ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	10/14 (71%)	0	1 (10%)
1	B	10/14 (71%)	0	1 (10%)
1	C	10/14 (71%)	0	1 (10%)
1	D	10/14 (71%)	0	1 (10%)
All	All	40/56 (71%)	0	4 (10%)

There are no RNA backbone outliers to report.

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	2	LCC
1	B	2	LCC
1	C	2	LCC
1	D	2	LCC

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

16 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$
1	LCC	A	1	1	17,21,25	2.43	5 (29%)	17,33,40	1.84	5 (29%)
1	LCC	A	2	1	16,24,25	2.43	5 (31%)	17,37,40	1.77	4 (23%)
1	LCA	A	3	1	19,26,27	1.22	2 (10%)	16,40,43	1.75	4 (25%)
1	LCG	A	4	1	19,27,28	1.67	2 (10%)	19,42,45	2.69	4 (21%)
1	LCC	B	1	1	17,21,25	2.21	5 (29%)	17,33,40	1.48	3 (17%)
1	LCC	B	2	1	16,24,25	2.24	5 (31%)	17,37,40	1.50	5 (29%)
1	LCA	B	3	1	19,26,27	1.17	1 (5%)	16,40,43	2.09	2 (12%)
1	LCG	B	4	1	19,27,28	2.29	7 (36%)	19,42,45	2.73	6 (31%)
1	LCC	C	1	1	17,21,25	2.55	6 (35%)	17,33,40	2.36	5 (29%)
1	LCC	C	2	1	16,24,25	2.36	5 (31%)	17,37,40	1.77	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	LCA	C	3	1	19,26,27	1.23	3 (15%)	16,40,43	2.12	5 (31%)
1	LCG	C	4	1	19,27,28	1.52	4 (21%)	19,42,45	2.37	6 (31%)
1	LCC	D	1	1	17,21,25	2.30	4 (23%)	17,33,40	1.64	2 (11%)
1	LCC	D	2	1	16,24,25	2.56	7 (43%)	17,37,40	1.71	5 (29%)
1	LCA	D	3	1	19,26,27	1.12	1 (5%)	16,40,43	2.11	3 (18%)
1	LCG	D	4	1	19,27,28	2.14	7 (36%)	19,42,45	2.83	5 (26%)

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	LCC	A	1	1	-	0/3/32/36	0/1/3/3
1	LCC	A	2	1	-	0/4/35/36	0/1/3/3
1	LCA	A	3	1	-	0/4/35/36	0/2/4/4
1	LCG	A	4	1	-	0/4/35/36	0/2/4/4
1	LCC	B	1	1	-	0/3/32/36	0/1/3/3
1	LCC	B	2	1	-	0/4/35/36	0/1/3/3
1	LCA	B	3	1	-	0/4/35/36	0/2/4/4
1	LCG	B	4	1	-	0/4/35/36	0/2/4/4
1	LCC	C	1	1	-	0/3/32/36	0/1/3/3
1	LCC	C	2	1	-	0/4/35/36	0/1/3/3
1	LCA	C	3	1	-	0/4/35/36	0/2/4/4
1	LCG	C	4	1	-	0/4/35/36	0/2/4/4
1	LCC	D	1	1	-	0/3/32/36	0/1/3/3
1	LCC	D	2	1	-	0/4/35/36	0/1/3/3
1	LCA	D	3	1	-	0/4/35/36	0/2/4/4
1	LCG	D	4	1	-	0/4/35/36	0/2/4/4

All (69) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	2	LCC	O4'-C4'	-5.70	1.37	1.45
1	A	2	LCC	C5M-C5	-5.27	1.40	1.51
1	C	2	LCC	C5M-C5	-5.20	1.40	1.51
1	A	4	LCG	C4'-C3'	-5.19	1.47	1.53
1	B	2	LCC	O4'-C4'	-5.16	1.38	1.45
1	D	1	LCC	C5M-C5	-4.85	1.41	1.51
1	B	1	LCC	C5M-C5	-4.79	1.41	1.51
1	C	1	LCC	C5M-C5	-4.70	1.41	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1	LCC	C5M-C5	-4.43	1.42	1.51
1	A	2	LCC	C2'-C3'	-4.03	1.42	1.53
1	D	1	LCC	C4'-C3'	-3.92	1.48	1.53
1	C	2	LCC	C2'-C3'	-3.82	1.43	1.53
1	B	4	LCG	C4'-C3'	-3.78	1.49	1.53
1	B	2	LCC	C5M-C5	-3.68	1.43	1.51
1	D	4	LCG	C4'-C3'	-3.43	1.49	1.53
1	D	4	LCG	C5-C4	-3.37	1.32	1.40
1	C	1	LCC	C2'-C3'	-3.29	1.44	1.53
1	B	1	LCC	C4'-C3'	-3.29	1.49	1.53
1	D	2	LCC	C5M-C5	-3.19	1.44	1.51
1	D	2	LCC	O2'-C2'	-3.12	1.36	1.43
1	D	2	LCC	O4'-C1'	-2.82	1.38	1.42
1	D	4	LCG	C6'-C4'	-2.74	1.47	1.53
1	B	4	LCG	C6-C5	-2.64	1.35	1.41
1	D	4	LCG	C6-C5	-2.62	1.35	1.41
1	C	2	LCC	O4'-C4'	-2.60	1.42	1.45
1	C	1	LCC	C2-N3	-2.53	1.32	1.38
1	B	2	LCC	O4'-C1'	-2.47	1.38	1.42
1	A	2	LCC	O4'-C4'	-2.44	1.42	1.45
1	B	4	LCG	C5-C4	-2.42	1.35	1.40
1	A	1	LCC	C2'-C3'	-2.39	1.47	1.53
1	C	1	LCC	C4'-C3'	-2.37	1.50	1.53
1	D	2	LCC	C5-C4	-2.26	1.37	1.41
1	C	3	LCA	C4'-C3'	-2.22	1.51	1.53
1	D	2	LCC	C4'-C3'	-2.18	1.51	1.53
1	B	1	LCC	C2-N3	-2.18	1.33	1.38
1	B	2	LCC	O2'-C2'	-2.12	1.38	1.43
1	C	2	LCC	C4'-C3'	-2.10	1.51	1.53
1	A	1	LCC	C2-N3	-2.09	1.33	1.38
1	C	4	LCG	C4'-C3'	-2.02	1.51	1.53
1	D	3	LCA	C4'-C3'	-2.01	1.51	1.53
1	C	4	LCG	C2'-C1'	2.01	1.58	1.53
1	D	4	LCG	C2-N1	2.07	1.39	1.35
1	C	3	LCA	C2-N3	2.27	1.36	1.32
1	A	3	LCA	C2-N3	2.31	1.36	1.32
1	B	3	LCA	C5-C4	2.34	1.45	1.40
1	C	3	LCA	C5-C4	2.38	1.45	1.40
1	B	1	LCC	O3'-C3'	2.50	1.47	1.42
1	A	2	LCC	O3'-C3'	2.56	1.47	1.42
1	C	4	LCG	C2-N1	2.57	1.40	1.35
1	A	3	LCA	C5-C4	3.02	1.47	1.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	1	LCC	O3'-C3'	3.07	1.48	1.42
1	A	4	LCG	O4'-C4'	3.23	1.50	1.45
1	C	4	LCG	O4'-C4'	3.31	1.50	1.45
1	B	4	LCG	C6-N1	3.56	1.39	1.33
1	B	2	LCC	O2'-C6'	3.70	1.50	1.44
1	B	4	LCG	O4'-C4'	3.71	1.51	1.45
1	B	4	LCG	O4'-C1'	3.83	1.48	1.42
1	D	4	LCG	O4'-C1'	3.83	1.48	1.42
1	C	2	LCC	O2'-C6'	4.52	1.52	1.44
1	A	1	LCC	O4'-C1'	4.62	1.49	1.42
1	D	4	LCG	O2'-C6'	4.65	1.52	1.44
1	A	2	LCC	O2'-C6'	4.66	1.52	1.44
1	D	2	LCC	O2'-C6'	4.89	1.53	1.44
1	B	4	LCG	O2'-C6'	4.95	1.53	1.44
1	C	1	LCC	O2'-C6'	4.98	1.53	1.44
1	B	1	LCC	O2'-C6'	5.07	1.53	1.44
1	D	1	LCC	O2'-C6'	5.56	1.54	1.44
1	C	1	LCC	O4'-C1'	5.62	1.51	1.42
1	A	1	LCC	O2'-C6'	6.49	1.56	1.44

All (68) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	4	LCG	C5-C6-N1	-7.06	114.29	123.52
1	B	3	LCA	N3-C2-N1	-6.81	123.52	128.87
1	D	3	LCA	N3-C2-N1	-6.13	124.05	128.87
1	A	4	LCG	C5-C6-N1	-6.06	115.59	123.52
1	B	4	LCG	C5-C6-N1	-6.01	115.67	123.52
1	B	4	LCG	N3-C2-N1	-5.45	120.14	127.56
1	A	4	LCG	N3-C2-N1	-5.44	120.15	127.56
1	C	1	LCC	O4'-C4'-C3'	-5.34	97.02	101.76
1	C	4	LCG	C5-C6-N1	-5.34	116.55	123.52
1	C	4	LCG	N3-C2-N1	-4.95	120.82	127.56
1	C	3	LCA	N3-C2-N1	-4.91	125.01	128.87
1	A	3	LCA	C1'-N9-C4	-4.80	121.44	126.81
1	D	4	LCG	N3-C2-N1	-4.60	121.29	127.56
1	C	3	LCA	C1'-N9-C4	-4.31	121.99	126.81
1	B	1	LCC	O5'-C5'-C4'	-3.70	103.38	111.56
1	A	1	LCC	O4'-C4'-C3'	-3.50	98.65	101.76
1	C	1	LCC	O5'-C5'-C4'	-3.38	104.07	111.56
1	A	1	LCC	O5'-C5'-C4'	-3.27	104.33	111.56
1	D	1	LCC	O5'-C5'-C4'	-3.23	104.41	111.56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	2	LCC	O5'-C5'-C4'	-3.09	101.97	107.54
1	A	2	LCC	O5'-C5'-C4'	-2.97	102.19	107.54
1	C	2	LCC	O5'-C5'-C4'	-2.90	102.31	107.54
1	C	3	LCA	O4'-C1'-N9	-2.72	104.78	108.96
1	C	4	LCG	O4'-C1'-N9	-2.69	104.83	108.96
1	D	3	LCA	C1'-N9-C4	-2.66	123.83	126.81
1	D	4	LCG	C4'-O4'-C1'	-2.63	102.63	107.40
1	A	3	LCA	N3-C2-N1	-2.50	126.91	128.87
1	B	4	LCG	C4'-O4'-C1'	-2.50	102.87	107.40
1	D	4	LCG	O4'-C1'-N9	-2.45	105.20	108.96
1	D	2	LCC	O3'-C3'-C4'	-2.43	106.30	112.22
1	B	1	LCC	O4'-C4'-C3'	-2.42	99.61	101.76
1	C	4	LCG	C4'-O4'-C1'	-2.36	103.13	107.40
1	B	4	LCG	O4'-C1'-N9	-2.35	105.34	108.96
1	A	3	LCA	O4'-C1'-N9	-2.29	105.44	108.96
1	B	3	LCA	C1'-N9-C4	-2.25	124.29	126.81
1	A	2	LCC	O3'-C3'-C4'	-2.25	106.75	112.22
1	B	2	LCC	O3'-C3'-C4'	-2.24	106.78	112.22
1	D	3	LCA	O4'-C1'-N9	-2.23	105.53	108.96
1	C	1	LCC	C4'-O4'-C1'	-2.10	103.59	107.40
1	B	2	LCC	O5'-C5'-C4'	-2.09	103.78	107.54
1	A	1	LCC	C6'-O2'-C2'	-2.06	104.03	107.88
1	C	3	LCA	O4'-C4'-C3'	-2.06	99.93	101.76
1	D	2	LCC	O4'-C1'-N1	2.04	112.10	108.96
1	B	2	LCC	N4-C4-N3	2.15	120.08	116.92
1	B	2	LCC	C5M-C5-C4	2.20	123.80	121.47
1	C	4	LCG	N2-C2-N3	2.21	121.86	117.72
1	A	3	LCA	N6-C6-N1	2.26	122.31	118.52
1	A	1	LCC	N4-C4-N3	2.30	120.29	116.92
1	B	4	LCG	C2'-C1'-N9	2.31	120.08	113.48
1	C	1	LCC	N4-C4-N3	2.39	120.42	116.92
1	C	2	LCC	O4'-C4'-C5'	2.44	114.59	109.16
1	C	2	LCC	O4'-C1'-N1	2.48	112.77	108.96
1	A	4	LCG	C2'-C1'-N9	2.59	120.87	113.48
1	A	2	LCC	O4'-C4'-C5'	2.73	115.23	109.16
1	C	3	LCA	N6-C6-N1	2.85	123.30	118.52
1	D	2	LCC	C5M-C5-C4	2.91	124.55	121.47
1	B	1	LCC	O4'-C4'-C5'	2.95	114.90	109.15
1	B	2	LCC	O4'-C4'-C5'	2.98	115.79	109.16
1	D	2	LCC	O4'-C4'-C5'	3.21	116.30	109.16
1	A	1	LCC	C5M-C5-C4	3.77	125.46	121.47
1	C	2	LCC	C5M-C5-C4	3.98	125.68	121.47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	1	LCC	O4'-C4'-C5'	4.13	117.19	109.15
1	A	2	LCC	C5M-C5-C4	4.25	125.97	121.47
1	C	4	LCG	C6-N1-C2	4.96	121.70	115.88
1	C	1	LCC	C5M-C5-C4	5.68	127.47	121.47
1	B	4	LCG	C6-N1-C2	6.22	123.17	115.88
1	A	4	LCG	C6-N1-C2	6.68	123.70	115.88
1	D	4	LCG	C6-N1-C2	6.73	123.76	115.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 23 ligands modelled in this entry, 15 are monoatomic - leaving 8 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
3	5GP	A	104	2	22,26,26	2.19	6 (27%)	25,40,40	1.74	5 (20%)
3	5GP	A	105	-	22,26,26	2.28	6 (27%)	25,40,40	2.77	7 (28%)
3	5GP	B	105	-	22,26,26	2.30	8 (36%)	25,40,40	2.24	8 (32%)
3	5GP	B	106	-	22,26,26	1.98	7 (31%)	25,40,40	2.63	8 (32%)
3	5GP	C	105	2	22,26,26	2.24	6 (27%)	25,40,40	1.66	4 (16%)
3	5GP	C	106	-	22,26,26	2.41	6 (27%)	25,40,40	2.83	6 (24%)
3	5GP	D	105	-	22,26,26	2.04	9 (40%)	25,40,40	1.90	8 (32%)
3	5GP	D	106	-	22,26,26	2.14	7 (31%)	25,40,40	2.73	6 (24%)

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	5GP	A	104	2	-	0/6/26/26	0/3/3/3
3	5GP	A	105	-	-	0/6/26/26	0/3/3/3
3	5GP	B	105	-	-	0/6/26/26	0/3/3/3
3	5GP	B	106	-	-	0/6/26/26	0/3/3/3
3	5GP	C	105	2	-	0/6/26/26	0/3/3/3
3	5GP	C	106	-	-	0/6/26/26	0/3/3/3
3	5GP	D	105	-	-	0/6/26/26	0/3/3/3
3	5GP	D	106	-	-	0/6/26/26	0/3/3/3

All (55) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	105	5GP	C6-C5	-4.42	1.32	1.41
3	D	105	5GP	C6-C5	-3.99	1.33	1.41
3	D	106	5GP	C6-C5	-3.92	1.33	1.41
3	B	106	5GP	C6-C5	-3.63	1.33	1.41
3	A	105	5GP	C6-C5	-3.51	1.34	1.41
3	C	106	5GP	C6-C5	-3.44	1.34	1.41
3	A	104	5GP	C6-C5	-3.22	1.34	1.41
3	C	105	5GP	C6-C5	-3.14	1.34	1.41
3	C	105	5GP	C2'-C1'	-2.96	1.48	1.53
3	A	104	5GP	C2'-C3'	-2.90	1.45	1.53
3	D	106	5GP	C2-N2	-2.80	1.28	1.34
3	C	105	5GP	C2'-C3'	-2.70	1.46	1.53
3	D	106	5GP	C2'-C3'	-2.68	1.46	1.53
3	B	105	5GP	C8-N7	-2.63	1.29	1.34
3	A	104	5GP	C2'-C1'	-2.56	1.49	1.53
3	B	106	5GP	C2'-C3'	-2.51	1.46	1.53
3	D	105	5GP	C8-N7	-2.35	1.30	1.34
3	B	106	5GP	C5-C4	-2.23	1.35	1.40
3	B	106	5GP	C2-N2	-2.06	1.29	1.34
3	D	105	5GP	C2'-C3'	-2.02	1.48	1.53
3	A	104	5GP	C5-C4	-2.02	1.35	1.40
3	C	106	5GP	C2-N1	2.03	1.39	1.35
3	C	106	5GP	C6-N1	2.26	1.37	1.33
3	D	105	5GP	P-O5'	2.26	1.66	1.59
3	C	105	5GP	P-O2P	2.32	1.62	1.54
3	D	105	5GP	O4'-C1'	2.33	1.44	1.41
3	A	105	5GP	P-O2P	2.36	1.62	1.54

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	105	5GP	P-O2P	2.39	1.63	1.54
3	B	105	5GP	O3'-C3'	2.57	1.49	1.43
3	D	106	5GP	C2'-C1'	2.67	1.57	1.53
3	D	105	5GP	C2-N1	2.68	1.40	1.35
3	B	105	5GP	C2-N1	2.71	1.40	1.35
3	C	105	5GP	O4'-C4'	2.74	1.51	1.45
3	A	105	5GP	C6-N1	2.74	1.38	1.33
3	D	106	5GP	C6-N1	2.78	1.38	1.33
3	B	105	5GP	C3'-C4'	2.88	1.60	1.53
3	A	105	5GP	C2-N1	2.93	1.40	1.35
3	D	105	5GP	O3'-C3'	2.99	1.50	1.43
3	B	106	5GP	C6-N1	3.06	1.38	1.33
3	C	106	5GP	O4'-C4'	3.45	1.52	1.45
3	C	106	5GP	P-O1P	3.45	1.61	1.50
3	B	105	5GP	C6-N1	3.74	1.39	1.33
3	B	105	5GP	P-O1P	3.96	1.63	1.50
3	D	106	5GP	O4'-C1'	4.02	1.47	1.41
3	B	106	5GP	P-O1P	4.06	1.63	1.50
3	A	104	5GP	P-O2P	4.18	1.69	1.54
3	B	105	5GP	O2'-C2'	4.20	1.52	1.43
3	D	105	5GP	C6-N1	4.41	1.41	1.33
3	D	106	5GP	P-O1P	4.88	1.66	1.50
3	B	106	5GP	O4'-C1'	4.91	1.48	1.41
3	A	105	5GP	P-O5'	5.37	1.75	1.59
3	A	105	5GP	O4'-C1'	6.08	1.49	1.41
3	A	104	5GP	C6-N1	6.84	1.45	1.33
3	C	106	5GP	O4'-C1'	7.60	1.52	1.41
3	C	105	5GP	C6-N1	7.80	1.47	1.33

All (52) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	106	5GP	C1'-N9-C4	-10.13	115.50	126.81
3	B	106	5GP	C1'-N9-C4	-9.20	116.54	126.81
3	A	105	5GP	N3-C2-N1	-6.99	118.04	127.56
3	C	106	5GP	C5-C6-N1	-6.75	114.70	123.52
3	C	106	5GP	N3-C2-N1	-6.61	118.56	127.56
3	B	105	5GP	C4'-O4'-C1'	-6.14	103.13	109.64
3	A	105	5GP	C5-C6-N1	-5.80	115.94	123.52
3	C	105	5GP	C5-C6-N1	-4.87	117.15	123.52
3	A	104	5GP	C5-C6-N1	-4.70	117.38	123.52
3	D	105	5GP	C5-C6-N1	-4.68	117.40	123.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	105	5GP	C5-C6-N1	-4.22	118.01	123.52
3	D	106	5GP	N3-C2-N1	-3.90	122.25	127.56
3	D	106	5GP	C5-C6-N1	-3.85	118.49	123.52
3	B	105	5GP	N3-C2-N1	-3.55	122.73	127.56
3	B	106	5GP	N3-C2-N1	-3.52	122.77	127.56
3	A	105	5GP	O4'-C1'-N9	-3.39	101.70	108.11
3	B	106	5GP	O5'-P-O1P	-3.26	98.90	107.08
3	D	105	5GP	N3-C2-N1	-3.20	123.20	127.56
3	D	106	5GP	C4'-O4'-C1'	-3.06	106.40	109.64
3	B	106	5GP	C5-C6-N1	-2.92	119.71	123.52
3	B	106	5GP	C4'-O4'-C1'	-2.90	106.56	109.64
3	A	105	5GP	C1'-N9-C4	-2.87	123.61	126.81
3	D	105	5GP	C4'-O4'-C1'	-2.80	106.68	109.64
3	D	105	5GP	O5'-C5'-C4'	-2.59	99.75	109.09
3	C	105	5GP	O4'-C4'-C3'	-2.48	100.13	105.16
3	C	106	5GP	O4'-C1'-N9	-2.37	103.63	108.11
3	C	105	5GP	O2P-P-O5'	-2.36	99.84	106.72
3	A	104	5GP	O4'-C4'-C3'	-2.33	100.44	105.16
3	B	105	5GP	O5'-P-O1P	-2.30	101.29	107.08
3	A	104	5GP	N3-C2-N1	-2.17	124.60	127.56
3	B	106	5GP	O5'-C5'-C4'	-2.14	101.36	109.09
3	C	106	5GP	C4'-O4'-C1'	-2.11	107.40	109.64
3	D	105	5GP	O2P-P-O5'	-2.04	100.76	106.72
3	B	105	5GP	C1'-N9-C4	2.24	129.31	126.81
3	D	105	5GP	C6-N1-C2	2.40	118.69	115.88
3	D	105	5GP	O3P-P-O1P	2.62	119.18	110.63
3	A	104	5GP	C6-N1-C2	2.65	118.99	115.88
3	A	105	5GP	N2-C2-N3	2.79	122.96	117.72
3	D	105	5GP	O4'-C1'-N9	2.81	113.41	108.11
3	B	105	5GP	C6-N1-C2	2.92	119.31	115.88
3	C	105	5GP	O3P-P-O1P	3.03	120.51	110.63
3	A	105	5GP	O3P-P-O1P	3.04	120.56	110.63
3	B	105	5GP	O3P-P-O2P	3.08	118.75	107.44
3	A	104	5GP	O3P-P-O1P	3.33	121.50	110.63
3	C	106	5GP	O3P-P-O2P	3.45	120.11	107.44
3	B	106	5GP	O3P-P-O2P	3.52	120.36	107.44
3	D	106	5GP	O3P-P-O2P	3.52	120.38	107.44
3	B	105	5GP	C2'-C1'-N9	3.70	123.37	113.47
3	D	106	5GP	C6-N1-C2	3.82	120.36	115.88
3	B	106	5GP	O4'-C1'-N9	4.58	116.77	108.11
3	A	105	5GP	C6-N1-C2	7.46	124.62	115.88
3	C	106	5GP	C6-N1-C2	8.26	125.56	115.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	105	5GP	1	0
3	C	106	5GP	1	0

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	10/14 (71%)	-1.01	0 100 100	13, 17, 20, 22	0
1	B	10/14 (71%)	-0.97	0 100 100	14, 16, 17, 21	0
1	C	10/14 (71%)	-1.02	0 100 100	13, 17, 20, 21	0
1	D	10/14 (71%)	-1.01	0 100 100	14, 16, 17, 20	0
All	All	40/56 (71%)	-1.00	0 100 100	13, 16, 21, 22	0

There are no RSRZ outliers to report.

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	LCG	C	4	24/25	0.99	0.04	-	12,13,15,16	0
1	LCC	B	2	22/23	0.97	0.05	-	19,21,25,28	0
1	LCA	C	3	23/24	0.98	0.04	-	13,14,17,22	0
1	LCC	B	1	19/23	0.92	0.10	-	26,31,38,41	0
1	LCC	A	1	19/23	0.94	0.07	-	16,22,31,37	0
1	LCC	D	1	19/23	0.93	0.10	-	25,31,39,44	0
1	LCC	C	2	22/23	0.96	0.06	-	15,18,24,26	0
1	LCC	A	2	22/23	0.96	0.06	-	15,18,25,26	0
1	LCA	B	3	23/24	0.97	0.05	-	15,16,20,24	0
1	LCC	C	1	19/23	0.95	0.06	-	16,22,30,37	0
1	LCG	D	4	24/25	0.99	0.04	-	14,15,16,17	0
1	LCG	B	4	24/25	0.98	0.05	-	14,15,16,17	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	LCC	D	2	22/23	0.96	0.05	-	18,21,26,28	0
1	LCA	D	3	23/24	0.97	0.04	-	15,16,21,25	0
1	LCG	A	4	24/25	0.99	0.04	-	11,13,15,15	0
1	LCA	A	3	23/24	0.98	0.04	-	13,15,17,23	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	MG	C	102	1/1	0.99	0.15	22.48	21,21,21,21	0
2	MG	A	102	1/1	0.99	0.14	20.89	21,21,21,21	0
2	MG	D	102	1/1	1.00	0.04	-	18,18,18,18	1
2	MG	B	102	1/1	1.00	0.02	-	17,17,17,17	1
3	5GP	B	105	24/24	0.92	0.11	-	21,39,104,110	0
2	MG	B	101	1/1	0.98	0.06	-	17,17,17,17	0
2	MG	D	101	1/1	0.99	0.05	-	18,18,18,18	0
2	MG	D	104	1/1	0.98	0.10	-	31,31,31,31	1
3	5GP	C	105	24/24	0.94	0.07	-	16,20,49,56	0
2	MG	A	101	1/1	0.99	0.06	-	15,15,15,15	0
2	MG	C	104	1/1	0.95	0.07	-	24,24,24,24	1
3	5GP	A	104	24/24	0.94	0.07	-	16,21,47,51	0
2	MG	C	103	1/1	0.99	0.17	-	14,14,14,14	1
2	MG	A	103	1/1	0.98	0.10	-	15,15,15,15	1
3	5GP	B	106	24/24	0.90	0.12	-	27,52,98,109	0
3	5GP	D	106	24/24	0.90	0.12	-	26,49,105,108	0
2	MG	C	101	1/1	0.99	0.04	-	15,15,15,15	0
2	MG	B	103	1/1	0.91	0.16	-	29,29,29,29	0
3	5GP	D	105	24/24	0.93	0.13	-	21,38,122,134	0
2	MG	D	103	1/1	0.98	0.20	-	29,29,29,29	0
2	MG	B	104	1/1	0.98	0.10	-	30,30,30,30	1
3	5GP	C	106	24/24	0.92	0.11	-	17,26,128,135	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	5GP	A	105	24/24	0.93	0.11	-	18,26,128,139	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.