



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:33 AM GMT

PDB ID : 2R7X  
Title : Crystal Structure of Rotavirus SA11 VP1/RNA (UGUGACC)/GTP complex  
Authors : Lu, X.; Harrison, S.C.; Tao, Y.J.; Patton, J.T.; Nibert, M.L.  
Deposited on : 2007-09-10  
Resolution : 2.80 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](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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

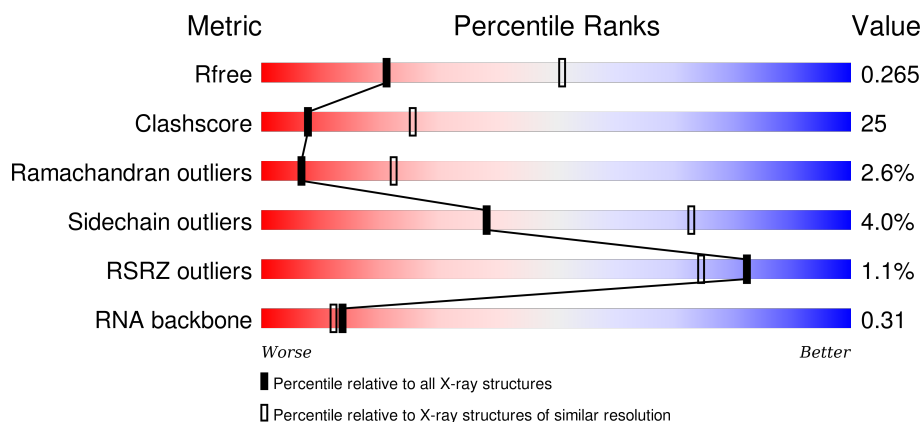
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)
RNA backbone	2183	1091 (3.20-2.40)

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	X	7	<div> <div></div> <div>29%71%</div> </div>
1	Y	7	<div> <div></div> <div>29%71%</div> </div>
2	A	1095	<div> <div>%</div> <div>53%41%..</div> </div>
2	B	1095	<div> <div>%</div> <div>52%42%..</div> </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
3	GTP	A	1111	-	-	-	X
3	GTP	B	1111	-	-	-	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 17744 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.

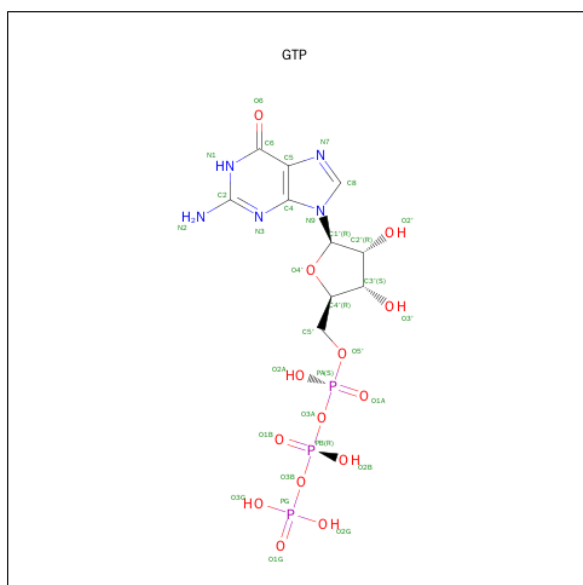
- Molecule 1 is a RNA chain called RNA (5'-R(\*UP\*GP\*UP\*GP\*AP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	X	7	Total	C	N	O	P	0	0	0
			145	66	25	48	6			
1	Y	7	Total	C	N	O	P	0	0	0
			145	66	25	48	6			

- Molecule 2 is a protein called RNA-dependent RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	1073	Total	C	N	O	S	0	0	0
			8695	5576	1447	1634	38			
2	B	1073	Total	C	N	O	S	0	0	0
			8695	5576	1447	1634	38			

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>14</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			32	10	5	14	3		
3	B	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

### 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(\*UP\*GP\*UP\*GP\*AP\*CP\*C)-3')

Chain X: 

U1101  
G1102  
U1103  
G1104  
A1105  
C1106  
G1107

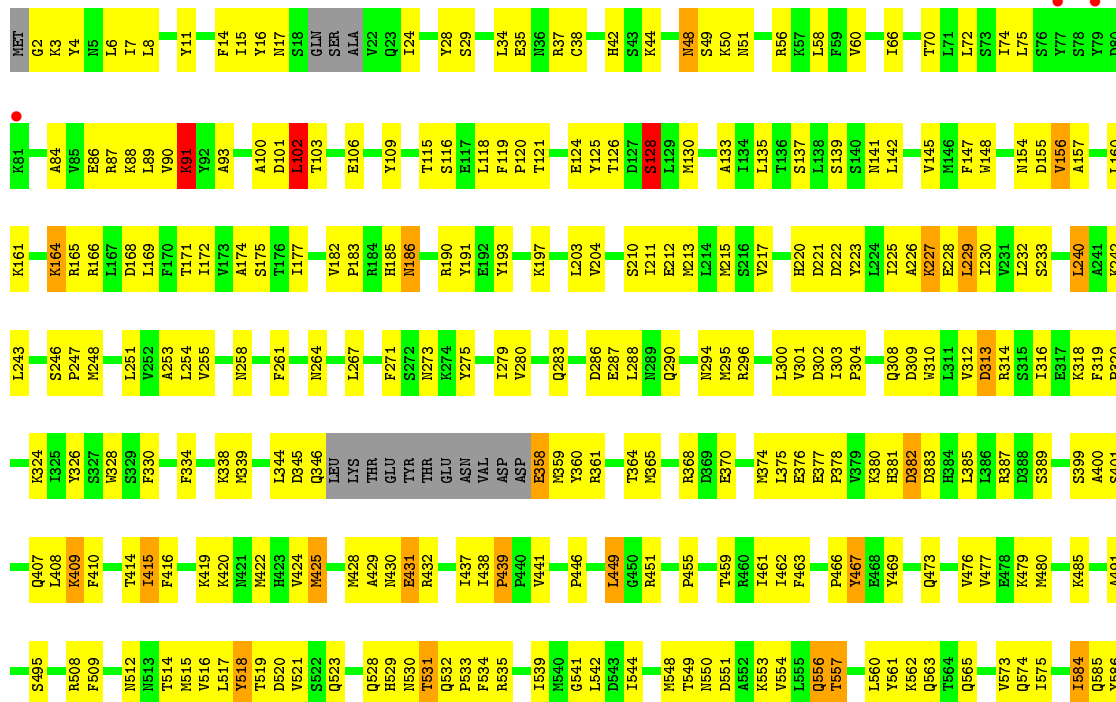
- Molecule 1: RNA (5'-R(\*UP\*GP\*UP\*GP\*AP\*CP\*C)-3')

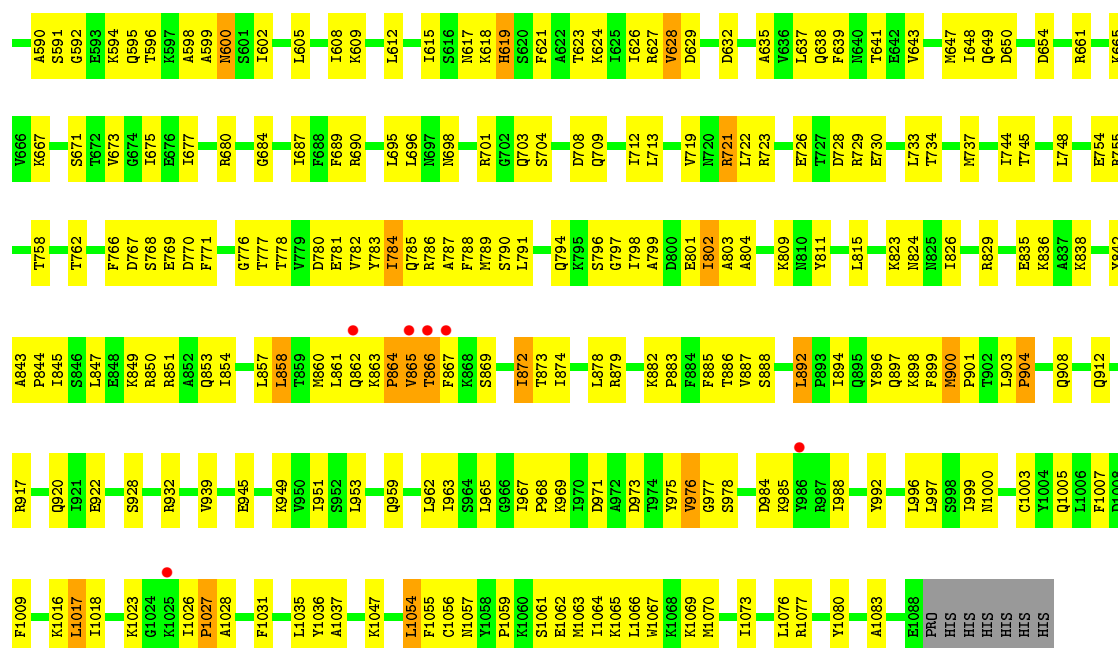
Chain Y: 

U1101  
G1102  
U1103  
G1104  
A1105  
C1106  
G1107

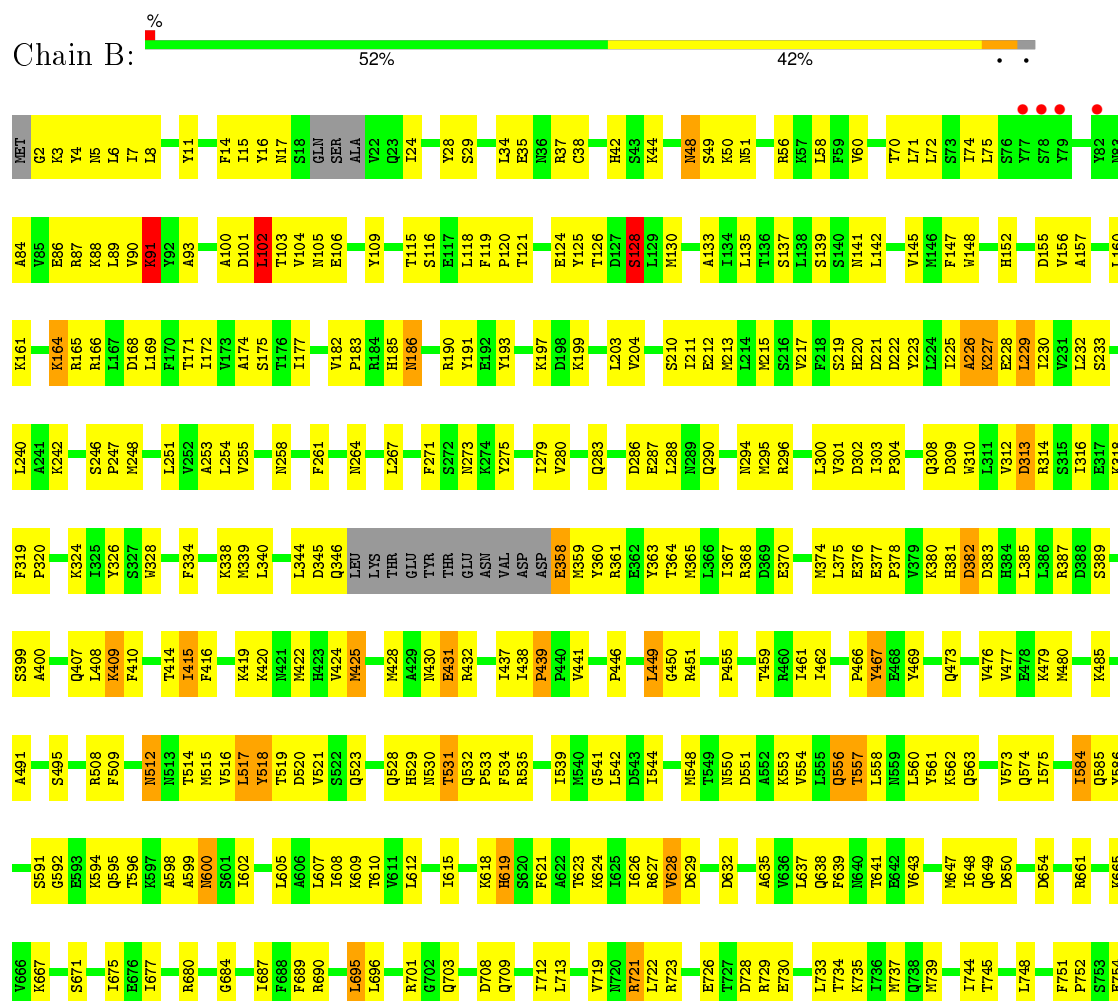
- Molecule 2: RNA-dependent RNA polymerase

Chain A: 





• Molecule 2: RNA-dependent RNA polymerase



HIS	Q1005	Q908	K838	R755
HIS	L1006	Q912	Y842	T758
	F1007	Q917	A843	
	D1008	R917	P844	T762
	F1009		I845	
	K1016	Q920	S846	F766
	L1017	I921	L847	D767
	I1018	E922	E848	S768
	R1019		K849	E769
	I1020	S928	R850	D770
	P1021		R851	F771
	F1022	R932	A852	
	K1023		Q853	G776
	G1024	Y939	I854	T777
	K1025			T778
	I1026	E945	L857	V779
	P1027		L858	D780
	A1028	K949	T859	E781
		V950	H860	V782
	L1035	I951	L861	
	Y1036	S952	Q862	I784
	A1037	L953	K863	Q785
			P864	R786
	K1047	I958	V865	A787
		Q959	T866	F788
	L1054		R867	M789
	F1055	L962	K868	S790
	C1056	S964	S869	L791
		L965		S792
	P1059	G966	I872	S793
	S1061	I967	T873	Q794
	E1062	P968	I874	K795
	M1063	K969		S796
	I1064	I970	L878	G797
	K1065	D971	R879	I798
	L1066	A972	D880	A799
	M1067	D973	I881	D800
	K1068	T974	K882	E801
	K1069	Y975	P883	I802
	M1070	V976	F884	A803
		G977	F885	A804
		S978	T886	S805
	I1073		V887	
	L1076	D984	S888	K809
	R1077	K985		M810
			P893	Y811
	Y1080	I988	I894	
	A1083		Q895	L815
	F1086	Y992	Y896	K823
	Q1087	L996	Q897	N824
	E1088	L997	K898	N825
	PRO	S998	F899	I826
HIS	HIS	I999	M900	
HIS	HIS	M1000	P901	R829
HIS	HIS	C1003	T902	
HIS	HIS	Y1004	L903	E835
			P904	K836
				A837



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.00Å 143.77Å 112.83Å 90.00° 90.65° 90.00°	Depositor
Resolution (Å)	30.00 – 2.80 35.94 – 2.80	Depositor EDS
% Data completeness (in resolution range)	83.1 (30.00-2.80) 80.7 (35.94-2.80)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.39 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.233 , 0.277 0.224 , 0.265	Depositor DCC
$R_{free}$ test set	4650 reflections (10.20%)	DCC
Wilson B-factor (Å <sup>2</sup> )	55.4	Xtriage
Anisotropy	0.239	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 17.5	EDS
Estimated twinning fraction	0.358 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 57725 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	17744	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.50% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GTP

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	X	0.52	0/161	0.97	0/249
1	Y	0.46	0/161	0.96	0/249
2	A	0.41	0/8866	0.63	0/11985
2	B	0.40	0/8866	0.62	0/11985
All	All	0.41	0/18054	0.63	0/24468

There are no bond length outliers.

There are no bond angle outliers.

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	X	145	0	77	16	0
1	Y	145	0	77	16	0
2	A	8695	0	8782	433	0
2	B	8695	0	8782	436	0
3	A	32	0	12	2	0
3	B	32	0	12	2	0
All	All	17744	0	17742	884	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 25.

The worst 5 of 884 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)
2:A:865:VAL:HG13	2:A:866:THR:H	1.14	1.08
2:B:865:VAL:HG13	2:B:866:THR:H	1.14	1.08
2:B:520:ASP:HB3	2:B:667:LYS:HG2	1.45	0.98
2:A:520:ASP:HB3	2:A:667:LYS:HG2	1.46	0.97
2:B:186:ASN:ND2	2:B:190:ARG:H	1.65	0.95

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	
2	A	1067/1095 (97%)	915 (86%)	124 (12%)	28 (3%)	7	22
2	B	1067/1095 (97%)	919 (86%)	120 (11%)	28 (3%)	7	22
All	All	2134/2190 (97%)	1834 (86%)	244 (11%)	56 (3%)	7	22

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	866	THR
2	A	976	VAL
2	A	978	SER
2	B	866	THR
2	B	976	VAL

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	A	974/996 (98%)	935 (96%)	39 (4%)	38	73
2	B	974/996 (98%)	935 (96%)	39 (4%)	38	73
All	All	1948/1992 (98%)	1870 (96%)	78 (4%)	38	73

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	900	MET
2	B	121	THR
2	B	872	ILE
2	A	922	GLU
2	B	17	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	840	ASN
2	B	143	ASN
2	B	760	ASN
2	A	959	GLN
2	B	152	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	X	7/7 (100%)	4 (57%)	3 (42%)
1	Y	7/7 (100%)	4 (57%)	3 (42%)
All	All	14/14 (100%)	8 (57%)	6 (42%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	X	1102	G
1	X	1103	U
1	X	1104	G
1	X	1105	A
1	Y	1102	G

5 of 6 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	X	1103	U
1	Y	1103	U
1	Y	1101	U
1	X	1102	G
1	Y	1102	G

## 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](#)

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	GTP	A	1111	-	25,34,34	1.74	7 (28%)	34,54,54	4.63	16 (47%)
3	GTP	B	1111	-	25,34,34	1.79	7 (28%)	34,54,54	4.64	16 (47%)

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	GTP	A	1111	-	-	0/18/38/38	0/3/3/3
3	GTP	B	1111	-	-	0/18/38/38	0/3/3/3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1111	GTP	PG-O1G	-4.00	1.38	1.51
3	A	1111	GTP	PG-O1G	-3.70	1.39	1.51
3	B	1111	GTP	C8-N7	-3.41	1.28	1.34
3	B	1111	GTP	O4'-C1'	-3.12	1.37	1.41
3	A	1111	GTP	C8-N7	-3.05	1.28	1.34

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1111	GTP	O3G-PG-O1G	-14.08	65.25	110.58
3	A	1111	GTP	O3G-PG-O1G	-14.05	65.36	110.58
3	A	1111	GTP	O2G-PG-O1G	-11.42	73.83	110.58
3	B	1111	GTP	O2G-PG-O1G	-10.99	75.20	110.58
3	A	1111	GTP	PA-O3A-PB	-9.37	106.41	132.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1111	GTP	2	0
3	B	1111	GTP	2	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	X	7/7 (100%)	0.00	0 100 100	47, 60, 73, 77	0
1	Y	7/7 (100%)	-0.02	0 100 100	47, 59, 72, 77	0
2	A	1073/1095 (97%)	-0.24	9 (0%) 87 81	15, 49, 79, 120	0
2	B	1073/1095 (97%)	-0.19	15 (1%) 78 69	17, 50, 80, 121	0
All	All	2160/2204 (98%)	-0.22	24 (1%) 82 74	15, 50, 79, 121	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	79	TYR	5.0
2	A	865	VAL	4.9
2	B	79	TYR	4.1
2	B	867	PHE	3.9
2	A	866	THR	3.8

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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
3	GTP	A	1111	32/32	0.46	0.36	2.74	107,137,145,145	0
3	GTP	B	1111	32/32	0.43	0.41	2.49	112,142,150,150	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.