



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 03:08 PM GMT

PDB ID : 3S1R  
Title : RNA Polymerase II Initiation Complex with a 5-nt 3'-deoxy RNA soaked with GTP  
Authors : Liu, X.; Bushnell, D.A.; Silva, D.A.; Huang, X.; Kornberg, R.D.  
Deposited on : 2011-05-16  
Resolution : 3.20 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

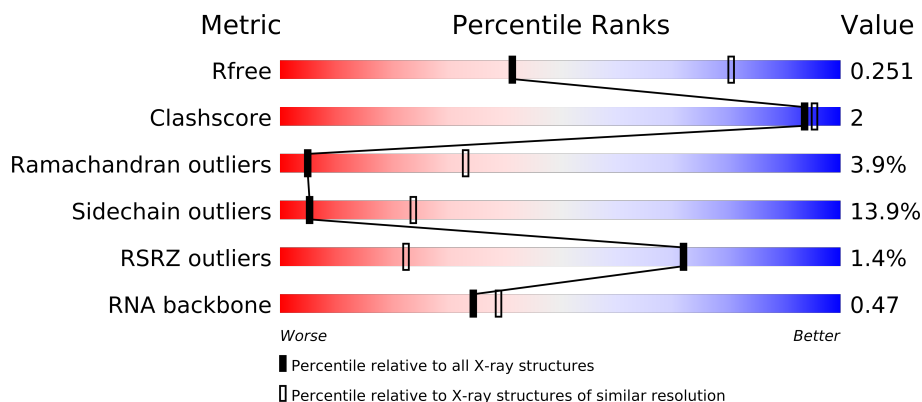
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.20 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))
$R_{free}$	66092	1824 (3.30-3.10)
Clashscore	79885	1078 (3.26-3.14)
Ramachandran outliers	78287	1059 (3.26-3.14)
Sidechain outliers	78261	1058 (3.26-3.14)
RSRZ outliers	66119	1825 (3.30-3.10)
RNA backbone	1838	1002 (3.72-2.68)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1733	
2	B	1224	
3	C	318	
4	E	215	
5	F	155	
6	H	146	
7	I	122	
8	J	70	
9	K	120	
10	L	70	
11	R	5	
12	T	29	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
13	GTP	R	100	-	X

## 2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 28601 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1405	Total	C	N	O	S	0	0	0
			11043	6965	1936	2081	61			

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1114	Total	C	N	O	S	0	0	0
			8861	5610	1549	1647	55			

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	85	Total	C	N	O	S	0	0	0
			688	439	116	130	3			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

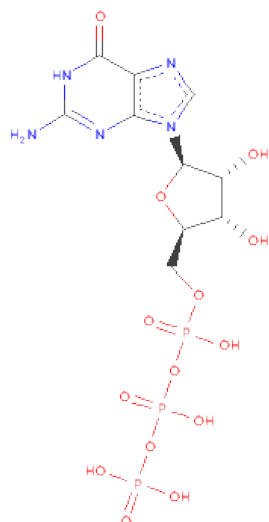
- Molecule 11 is a RNA chain called RNA (5'-R(\*AP\*GP\*AP\*GP\*G\*)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	R	5	Total	C	N	O	P	0	0	0
			109	50	25	30	4			

- Molecule 12 is a DNA chain called DNA (5'-D(\*CP\*TP\*AP\*CP\*CP\*GP\*AP\*TP\*AP\*AP\*GP\*CP\*AP\*GP\*AP\*CP\*GP\*AP\*TP\*CP\*CP\*TP\*CP\*TP\*CP\*GP\*AP\*TP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	T	8	Total	C	N	O	P	0	0	0
			159	76	26	49	8			

- Molecule 13 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
13	R	1	Total	C	N	O	P	0	0
			32	10	5	14	3		

- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	J	1	Total	Zn	0	0
			1	1		
14	B	1	Total	Zn	0	0
			1	1		
14	I	2	Total	Zn	0	0
			2	2		
14	C	1	Total	Zn	0	0
			1	1		
14	A	2	Total	Zn	0	0
			2	2		
14	L	1	Total	Zn	0	0
			1	1		

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

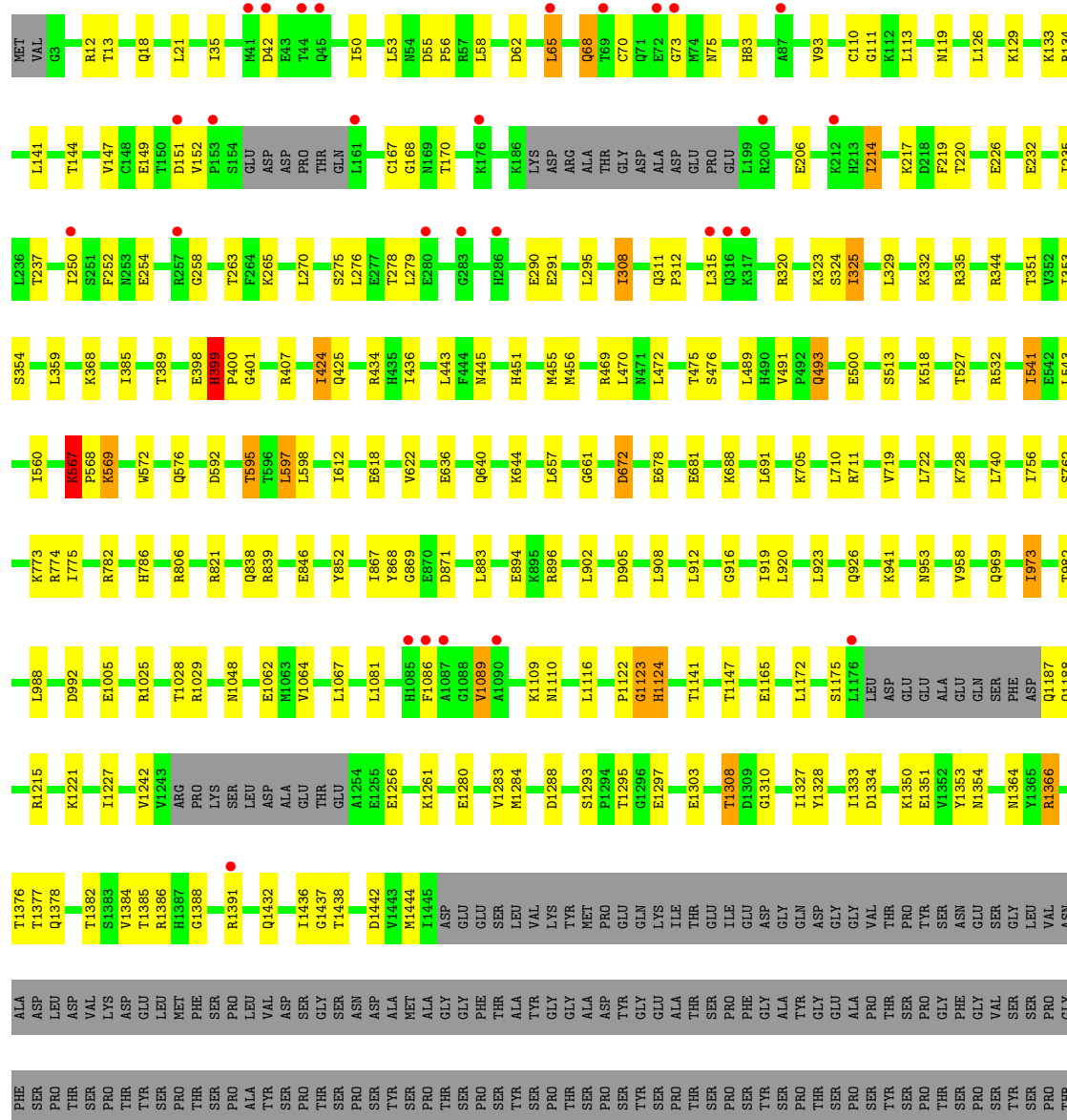
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	1	Total	Mg	0	0
			1	1		

### 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: DNA-directed RNA polymerase II subunit RPB1

Chain A: 







MET	D2	Q3	R7	I30	L37	G51	R52	P53	Q54	R55	Q61	T65	M75	L78	V90	T95	H99	I100	N104	P125	S126	I127	V142	L156	S157	S158	R152	R169	T204	I213	C214	M215
-----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

- Chain F: 

HIS	GLN	GLN	ILE	ARG	ARG	LYS	THR	LEU	LYS	E71	E72	A73	Q78	T82	L99	D110	L111	E112	R119	E127	L138	I152	L155																															
MET	SER	ASP	TVR	GLU	GLU	ALA	PHE	ASN	ASP	GLY	ASN	GLU	ASN	PHE	ASP	VAL	GLU	HIS	PHE	SER	ASP	GLU	THR	TVR	GLU	GLY	LYS	PRO	GLN	PHE	LYS	ASP	GLY	GLY	THR	THR	ASP	ALA	ASN	GLY	LYS	THR	ILE	VAL	THR	GLY	GLY	ASN	GLY	PRO	GLU	ASP	PHE	GLN

- Chain H: 

The chart displays the frequency of various amino acids in 1000 random sequences. The x-axis represents the count of each amino acid, ranging from 0 to 1000. The y-axis lists the amino acids. The bars are color-coded: yellow for standard amino acids, orange for modified ones, and grey for non-standard ones. A red dot is placed above the bar for T76.

Amino Acid	Approximate Count	Category
Met	100	Standard
S2	100	Standard
Q11	100	Standard
V12	100	Standard
S13	100	Standard
G18	100	Standard
C24	100	Standard
R25	100	Standard
I26	100	Standard
D34	100	Standard
Q35	100	Standard
M43	100	Standard
F47	100	Standard
T56	100	Standard
L63	100	Standard
Asn	100	Standard
Leu	100	Standard
GlU	100	Standard
ASP	100	Standard
Thr	100	Standard
Pro	100	Standard
AlA	100	Standard
Asn	100	Standard
ASP	100	Standard
SER	100	Standard
SER	100	Standard
AlA	100	Standard
T76	100	Standard
R77	100	Standard
P82	100	Standard
Q83	100	Standard
A84	100	Standard
R87	100	Standard
S88	100	Standard
L89	100	Standard
A90	100	Standard
D91	100	Standard
D92	100	Standard
Y95	100	Standard
V107	100	Standard
S108	100	Standard
K109	100	Standard
D110	100	Standard
L111	100	Standard
R130	100	Standard
N131	100	Standard
L135	100	Standard
K136	100	Standard
N139	100	Standard
A140	100	Standard
R146	100	Standard

- Chain I:

MET	T2	N12	R17	T31	V35	E36	E37	D61	I62	R70	C75	C78	Q87	S88	Q89	Q90	R91	M97	K115	R118	T119	Q120	PHE	STR
-----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	-----	-----

- Chain J: 

- Chain K: 

M1		R6	L11	L12	G13	E14	K20	K26	V31	R54	V63	R74	I75	T78	Y81	D85	K97	L101	L114	ALA	ALA	ASP	ASP	ALA	PHE
----	--	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	-----	-----	-----	-----	-----	-----

- Chain L: 

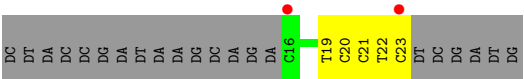
MET	ARG	SER	GLU	PHE	GLN	ILE	PRO	THR	ASN	LEU	ASP	ALA	ALA	ALA	ALA	GLY	THR	SER	GLN	ALA	ARG	THR						
A25	A26	T26	L27	K28	Z29	E30	E31	E32	E33	L40	D44	A45	V46	R47	C48	K49	D50	C51	F55	L56	L57	K58	A59	R62	K63	L64	V65	G66

- Chain R:



● Molecule 12: DNA (5'-D(\*CP\*TP\*AP\*CP\*CP\*GP\*AP\*TP\*AP\*AP\*GP\*CP\*AP\*GP\*AP\*CP\*GP\*AP\*TP\*CP\*CP\*TP\*CP\*TP\*CP\*GP\*AP\*TP\*G)-3')

Chain T:



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	160.67Å 221.32Å 193.21Å 90.00° 98.31° 90.00°	Depositor
Resolution (Å)	29.97 – 3.20 29.97 – 3.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (29.97-3.20) 99.4 (29.97-3.20)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.30 (at 3.18Å)	Xtriage
Refinement program	BUSTER 2.8.0	Depositor
R, $R_{free}$	0.177 , 0.226 0.201 , 0.251	Depositor DCC
$R_{free}$ test set	5435 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	74.7	Xtriage
Anisotropy	0.630	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 76.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	1 of 108691 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	28601	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, ZN, MG

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	0.50	0/11241	0.77	3/15199 (0.0%)
2	B	0.52	0/9033	0.80	2/12181 (0.0%)
3	C	0.49	0/2133	0.81	0/2891
4	E	0.45	0/1788	0.73	0/2406
5	F	0.51	0/700	0.76	0/945
6	H	0.50	0/1086	0.83	1/1470 (0.1%)
7	I	0.51	0/989	0.82	0/1331
8	J	0.55	0/541	0.88	0/727
9	K	0.45	0/937	0.71	0/1265
10	L	0.57	0/365	1.13	2/485 (0.4%)
11	R	0.93	0/123	1.64	0/191
12	T	1.29	0/176	1.87	5/268 (1.9%)
All	All	0.51	0/29112	0.81	13/39359 (0.0%)

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
12	T	22	DT	O4'-C1'-N1	8.14	113.70	108.00
2	B	647	GLY	C-N-CA	7.65	140.83	121.70
10	L	50	ASP	C-N-CA	7.62	140.74	121.70
12	T	23	DC	O4'-C1'-N1	6.07	112.25	108.00
1	A	1123	GLY	C-N-CA	6.02	136.75	121.70

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11043	0	23	16	0
2	B	8861	0	0	21	0
3	C	2095	0	0	6	0
4	E	1752	0	0	2	0
5	F	688	0	5	0	0
6	H	1068	0	0	4	0
7	I	971	0	0	2	0
8	J	532	0	0	4	0
9	K	919	0	0	1	0
10	L	363	0	0	1	0
11	R	109	0	0	0	0
12	T	159	0	0	0	0
13	R	32	0	0	0	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	1	0	0	0	0
All	All	28601	0	28	46	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

The worst 5 of 46 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:469:ARG:NH2	2:B:991:GLY:O	2.31	0.64
2:B:1009:ASP:OD2	8:J:48:ARG:NH2	2.31	0.63
1:A:399:HIS:O	1:A:401:GLY:N	2.32	0.63
2:B:788:ARG:NH1	2:B:790:ASP:OD1	2.38	0.57
2:B:848:ARG:NH2	2:B:996:ARG:NH1	2.54	0.55

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution. The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1395/1733 (80%)	1205 (86%)	137 (10%)	53 (4%)	5	34
2	B	1096/1224 (90%)	958 (87%)	90 (8%)	48 (4%)	4	29
3	C	264/318 (83%)	238 (90%)	19 (7%)	7 (3%)	8	46
4	E	212/215 (99%)	185 (87%)	22 (10%)	5 (2%)	9	51
5	F	83/155 (54%)	75 (90%)	6 (7%)	2 (2%)	9	51
6	H	129/146 (88%)	102 (79%)	18 (14%)	9 (7%)	2	13
7	I	117/122 (96%)	101 (86%)	14 (12%)	2 (2%)	14	62
8	J	63/70 (90%)	55 (87%)	7 (11%)	1 (2%)	14	63
9	K	112/120 (93%)	107 (96%)	5 (4%)	0	100	100
10	L	44/70 (63%)	27 (61%)	7 (16%)	10 (23%)	0	0
All	All	3515/4173 (84%)	3053 (87%)	325 (9%)	137 (4%)	5	33

5 of 137 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	55	ASP
1	A	65	LEU
1	A	119	ASN
1	A	399	HIS
1	A	424	ILE

### 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	1225/1520 (81%)	1046 (85%)	179 (15%)	5	21

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	967/1061 (91%)	840 (87%)	127 (13%)	6	28
3	C	234/274 (85%)	208 (89%)	26 (11%)	9	36
4	E	196/197 (100%)	174 (89%)	22 (11%)	9	36
5	F	75/137 (55%)	64 (85%)	11 (15%)	4	21
6	H	117/128 (91%)	99 (85%)	18 (15%)	4	18
7	I	113/116 (97%)	101 (89%)	12 (11%)	10	38
8	J	60/65 (92%)	49 (82%)	11 (18%)	2	12
9	K	99/102 (97%)	83 (84%)	16 (16%)	3	15
10	L	40/57 (70%)	27 (68%)	13 (32%)	0	0
All	All	3126/3657 (86%)	2691 (86%)	435 (14%)	5	24

5 of 435 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	194	GLU
2	B	642	ASP
8	J	43	ARG
2	B	261	ARG
2	B	401	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
11	R	3/5 (60%)	1 (33%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
11	R	8	A

There are no RNA pucker outliers to report.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 9 are monoatomic - leaving 1 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$
13	GTP	R	100	-	34,34,34	1.59	7 (20%)	51,54,54	2.14	11 (21%)

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
13	GTP	R	100	-	-	0/22/38/38	0/1/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	R	100	GTP	PA-O3A	3.40	1.66	1.59
13	R	100	GTP	PB-O3A	3.29	1.65	1.59
13	R	100	GTP	C2-N2	3.08	1.37	1.32
13	R	100	GTP	PB-O3B	2.80	1.65	1.59
13	R	100	GTP	O4'-C1'	2.59	1.45	1.41

The worst 5 of 11 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	R	100	GTP	C6-C5-N7	7.83	135.19	134.14
13	R	100	GTP	PB-O3B-PG	-4.68	117.97	131.68
13	R	100	GTP	N3-C4-N9	4.60	133.66	126.91
13	R	100	GTP	O3A-PB-O3B	4.52	110.85	101.66
13	R	100	GTP	C2-N3-C4	4.29	121.12	115.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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<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	1405/1733 (81%)	-0.10	29 (2%) 60 15	49, 97, 192, 214	0
2	B	1114/1224 (91%)	-0.21	14 (1%) 74 24	48, 84, 152, 204	0
3	C	266/318 (83%)	-0.32	0 100 100	58, 81, 117, 175	0
4	E	214/215 (99%)	-0.17	0 100 100	69, 124, 172, 182	0
5	F	85/155 (54%)	-0.27	0 100 100	73, 103, 140, 156	0
6	H	133/146 (91%)	0.02	1 (0%) 83 35	97, 131, 163, 178	0
7	I	119/122 (97%)	-0.25	0 100 100	61, 101, 141, 156	0
8	J	65/70 (92%)	-0.30	0 100 100	55, 73, 107, 123	0
9	K	114/120 (95%)	-0.32	0 100 100	59, 91, 118, 133	0
10	L	46/70 (65%)	-0.06	1 (2%) 59 14	70, 116, 147, 156	0
11	R	5/5 (100%)	2.40	3 (60%) 0 0	202, 206, 209, 212	0
12	T	8/29 (27%)	1.12	2 (25%) 1 1	169, 178, 191, 196	0
All	All	3574/4207 (84%)	-0.16	50 (1%) 72 22	48, 94, 180, 214	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1176	LEU	4.9
2	B	1222	ARG	4.0
2	B	883	LEU	4.0
11	R	6	A	3.8
1	A	72	GLU	3.7

### 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
13	GTP	R	100	32/32	0.36	3.60	249,251,254,256	0
14	ZN	C	319	1/1	0.11	-0.62	89,89,89,89	0
14	ZN	J	101	1/1	0.17	-0.98	78,78,78,78	0
14	ZN	I	204	1/1	0.08	-1.13	82,82,82,82	0
14	ZN	B	1307	1/1	0.10	-1.21	183,183,183,183	0
14	ZN	I	203	1/1	0.06	-1.22	102,102,102,102	0
14	ZN	A	1735	1/1	0.16	-1.57	176,176,176,176	0
15	MG	A	2001	1/1	0.06	-2.19	60,60,60,60	0
14	ZN	L	105	1/1	0.05	-2.22	99,99,99,99	0
14	ZN	A	1734	1/1	0.06	-2.76	289,289,289,289	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.