



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 21, 2021 – 06:08 PM BST

PDB ID : 6ZWW  
Title : Crystal structure of E. coli RNA helicase HrpA in complex with RNA  
Authors : Grass, L.M.; Wollenhaupt, J.; Barthel, T.; Loll, B.; Wahl, M.C.  
Deposited on : 2020-07-29  
Resolution : 3.16 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.20  
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) : 2.20

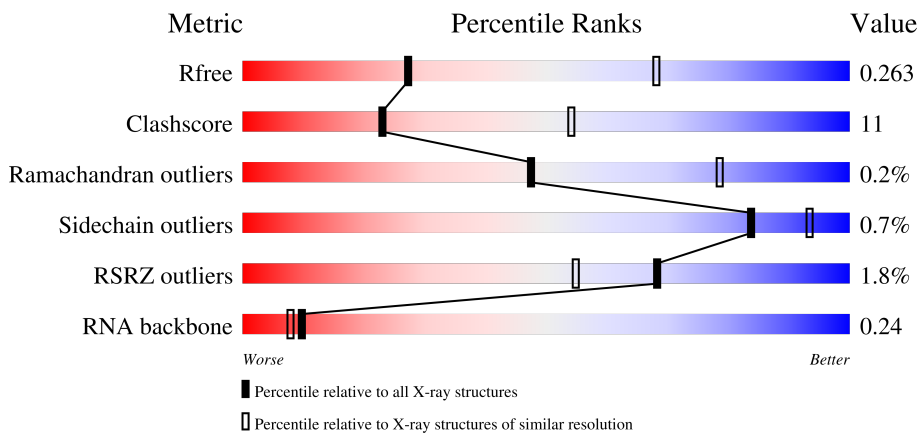
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.16 Å.

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




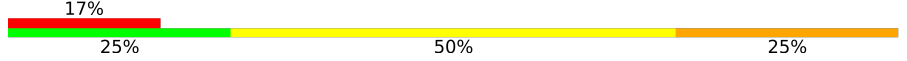
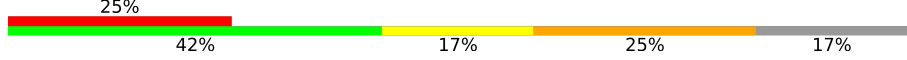
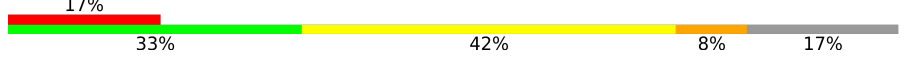
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)
RSRZ outliers	127900	1616 (3.20-3.12)
RNA backbone	3102	1073 (3.50-2.82)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	785	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 67%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 28%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">67%      28%      5%</p>
1	C	785	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 69%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 27%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">69%      27%      5%</p>
1	E	785	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 66%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 24%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 7%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">69%      24%      7%</p>
1	G	785	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 70%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 23%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 6%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">70%      23%      6%</p>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	B	12	
2	D	12	
2	F	12	
2	H	12	

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	CA	A	801	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 24593 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called ATP-dependent RNA helicase HrpA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	744	Total	C	N	O	S	0	0	0
			5961	3749	1091	1100	21			
1	C	749	Total	C	N	O	S	0	0	0
			5997	3771	1095	1110	21			
1	E	731	Total	C	N	O	S	0	0	0
			5862	3689	1070	1082	21			
1	G	735	Total	C	N	O	S	0	0	0
			5891	3707	1075	1088	21			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP A0A167DLY9
A	0	ALA	-	expression tag	UNP A0A167DLY9
C	-1	GLY	-	expression tag	UNP A0A167DLY9
C	0	ALA	-	expression tag	UNP A0A167DLY9
E	-1	GLY	-	expression tag	UNP A0A167DLY9
E	0	ALA	-	expression tag	UNP A0A167DLY9
G	-1	GLY	-	expression tag	UNP A0A167DLY9
G	0	ALA	-	expression tag	UNP A0A167DLY9

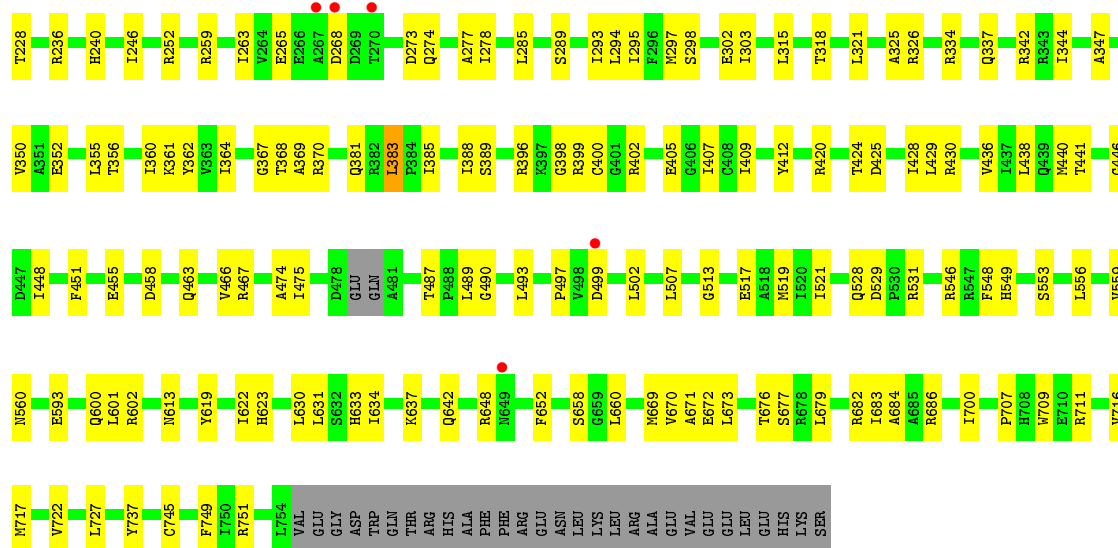
- Molecule 2 is a RNA chain called ssRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	12	Total	C	N	O	P	0	0	0
			240	108	24	96	12			
2	D	12	Total	C	N	O	P	0	0	0
			240	108	24	96	12			
2	F	10	Total	C	N	O	P	0	0	0
			200	90	20	80	10			
2	H	10	Total	C	N	O	P	0	0	0
			200	90	20	80	10			

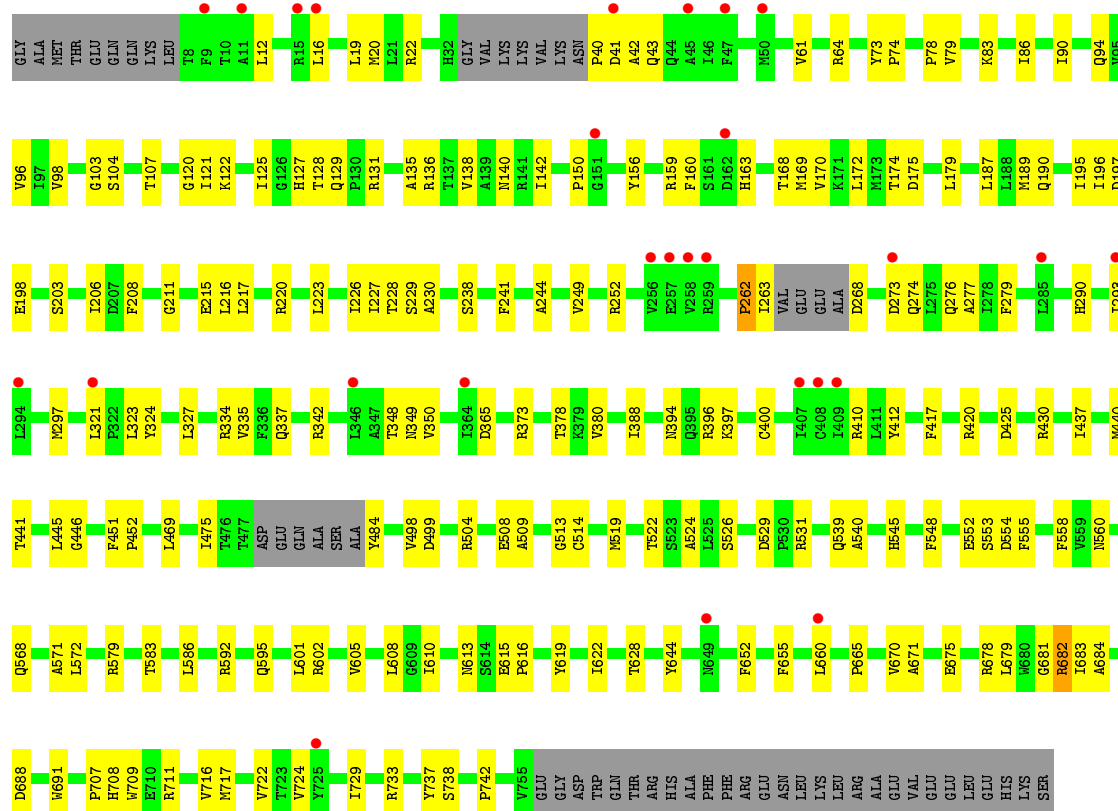
- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Ca 1	0	0
3	G	1	Total 1	Ca 1	0	0



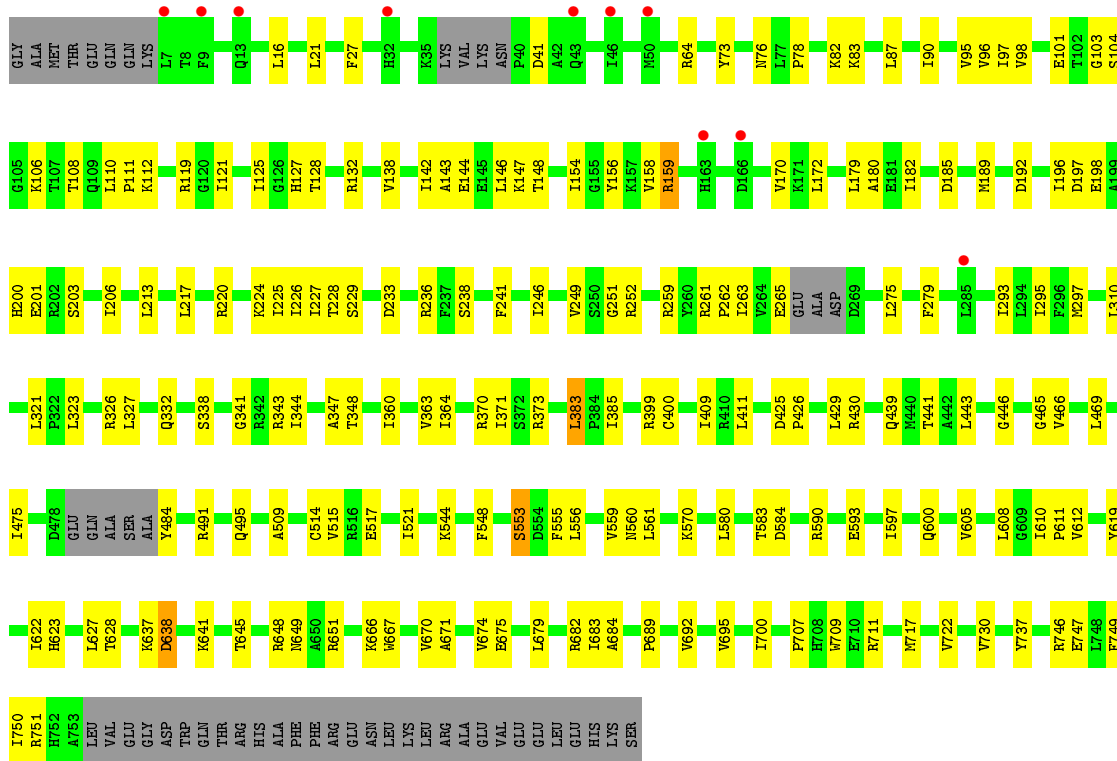


● Molecule 1: ATP-dependent RNA helicase HrpA



● Molecule 1: ATP-dependent RNA helicase HrpA

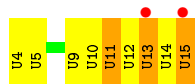




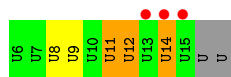
• Molecule 2: ssRNA



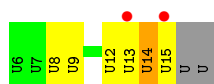
• Molecule 2: ssRNA



• Molecule 2: ssRNA



• Molecule 2: ssRNA





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.85Å 126.20Å 179.25Å 90.00° 110.76° 90.00°	Depositor
Resolution (Å)	49.40 – 3.16 49.40 – 3.16	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.40-3.16) 99.1 (49.40-3.16)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.98 (at 3.19Å)	Xtrriage
Refinement program	PHENIX 1.19.1_4122, PHENIX 1.19.1_4122	Depositor
R, $R_{free}$	0.226 , 0.265 0.223 , 0.263	Depositor DCC
$R_{free}$ test set	2096 reflections (2.67%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	131.7	Xtrriage
Anisotropy	0.339	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 83.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	24593	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	144.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% 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:  
CA

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	0.28	0/6068	0.58	1/8202 (0.0%)
1	C	0.28	0/6104	0.59	2/8252 (0.0%)
1	E	0.27	0/5967	0.55	0/8066
1	G	0.28	0/5996	0.56	1/8104 (0.0%)
2	B	0.36	0/263	1.02	0/404
2	D	0.38	0/263	1.01	0/404
2	F	0.30	0/219	1.02	0/336
2	H	0.38	0/219	0.94	0/336
All	All	0.28	0/25099	0.60	4/34104 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	411	LEU	CA-CB-CG	5.29	127.47	115.30
1	G	383	LEU	CA-CB-CG	5.07	126.97	115.30
1	C	383	LEU	CA-CB-CG	5.07	126.95	115.30
1	C	216	LEU	CA-CB-CG	5.01	126.83	115.30

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	5961	0	6058	155	0
1	C	5997	0	6089	145	0
1	E	5862	0	5946	138	0
1	G	5891	0	5977	123	0
2	B	240	0	121	4	0
2	D	240	0	121	5	0
2	F	200	0	101	5	0
2	H	200	0	101	3	0
3	A	1	0	0	0	0
3	G	1	0	0	0	0
All	All	24593	0	24514	558	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 11.

The worst 5 of 558 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)
1:A:159:ARG:HH21	1:A:597:ILE:HG12	1.38	0.87
1:A:103:GLY:HA3	1:A:252:ARG:H	1.40	0.85
1:A:544:LYS:HG3	1:A:547:ARG:HH12	1.41	0.84
1:C:71:ILE:HD11	1:C:116:GLU:HB2	1.61	0.82
1:C:466:VAL:HG13	1:C:475:ILE:HD11	1.61	0.81

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	740/785 (94%)	708 (96%)	31 (4%)	1 (0%)	51 83
1	C	745/785 (95%)	723 (97%)	22 (3%)	0	100 100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	723/785 (92%)	697 (96%)	25 (4%)	1 (0%)	51	83
1	G	727/785 (93%)	699 (96%)	25 (3%)	3 (0%)	34	68
All	All	2935/3140 (94%)	2827 (96%)	103 (4%)	5 (0%)	47	78

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	338	SER
1	A	402	ARG
1	G	341	GLY
1	E	262	PRO
1	G	262	PRO

### 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	
1	A	645/680 (95%)	640 (99%)	5 (1%)	81	92
1	C	649/680 (95%)	643 (99%)	6 (1%)	78	91
1	E	635/680 (93%)	631 (99%)	4 (1%)	86	94
1	G	638/680 (94%)	634 (99%)	4 (1%)	86	94
All	All	2567/2720 (94%)	2548 (99%)	19 (1%)	84	93

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	737	TYR
1	G	638	ASP
1	G	737	TYR
1	G	553	SER
1	C	737	TYR

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

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	11/12 (91%)	6 (54%)	0
2	D	12/12 (100%)	7 (58%)	1 (8%)
2	F	9/12 (75%)	5 (55%)	0
2	H	9/12 (75%)	5 (55%)	0
All	All	41/48 (85%)	23 (56%)	1 (2%)

5 of 23 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	5	U
2	B	9	U
2	B	10	U
2	B	11	U
2	B	12	U

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	D	4	U

## 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 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	744/785 (94%)	-0.08	4 (0%) 91 86	80, 129, 181, 265	0
1	C	749/785 (95%)	-0.10	7 (0%) 84 75	85, 126, 180, 276	0
1	E	731/785 (93%)	0.09	27 (3%) 41 25	118, 157, 216, 280	0
1	G	735/785 (93%)	-0.05	10 (1%) 75 63	109, 142, 195, 282	0
2	B	12/12 (100%)	0.91	0 100 100	112, 149, 302, 316	0
2	D	12/12 (100%)	0.91	2 (16%) 1 1	120, 158, 272, 293	0
2	F	10/12 (83%)	0.76	3 (30%) 0 0	144, 161, 206, 220	0
2	H	10/12 (83%)	1.21	2 (20%) 1 0	133, 152, 213, 219	0
All	All	3003/3188 (94%)	-0.02	55 (1%) 68 55	80, 139, 198, 316	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	41	ASP	5.1
1	G	46	ILE	4.3
1	G	50	MET	4.2
1	E	9	PHE	4.0
2	H	15	U	3.9

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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<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	B-factors(Å <sup>2</sup> )	Q<0.9
3	CA	A	801	1/1	0.72	0.69	131,131,131,131	0
3	CA	G	801	1/1	0.92	0.49	122,122,122,122	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.