



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 04:46 am BST

PDB ID : 4BBR  
Title : Structure of RNA polymerase II-TFIIB complex  
Authors : Sainsbury, S.; Niesser, J.; Cramer, P.  
Deposited on : 2012-09-27  
Resolution : 3.40 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 : 2.11  
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

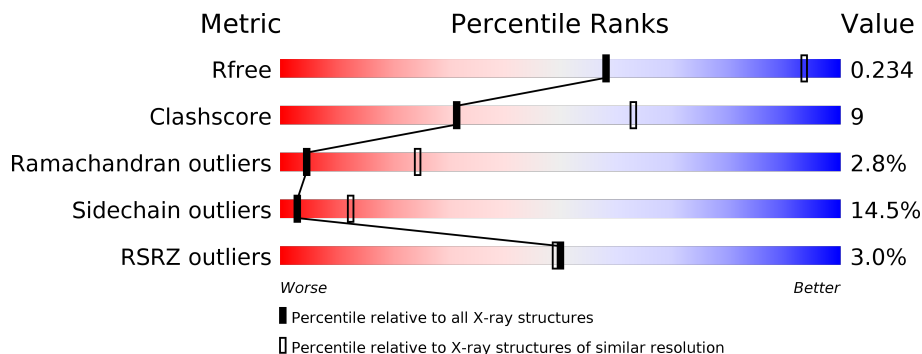
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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.








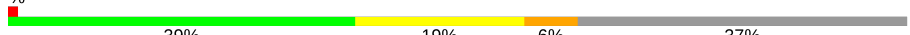
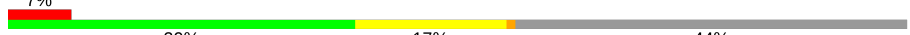
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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 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	1733	
2	B	1224	
3	C	318	
4	D	221	
5	E	215	
6	F	155	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	G	171	 69% 25% 6% .
8	H	146	 5% 60% 27% . . 8%
9	I	122	 6% 74% 22% . .
10	J	70	 50% 33% 7% . 7%
11	K	120	 64% 29% . 5%
12	L	70	 % 39% 19% 6% 37%
13	M	345	 7% 39% 17% . 44%

## 2 Entry composition [i](#)

There are 15 unique types of molecules in this entry. The entry contains 32800 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 DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1414	11123	7007	1945	2109	62	0	0	0

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1150	9095	5751	1598	1690	56	0	0	0

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

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

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	178	1434	887	257	288	2	0	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

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

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	87	705	451	119	132	3	0	0	0

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	171	1340	861	222	249	8	0	0	0

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	134	1076	677	182	213	4	0	0	0

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

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

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

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

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	44	351	217	70	60	4	0	0	0

- Molecule 13 is a protein called TRANSCRIPTION INITIATION FACTOR IIB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	193	Total	C	N	O	S	0	0	0
			1396	862	246	274	14			

- 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		
14	M	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	2	Total	Mg	0	0
			2	2		

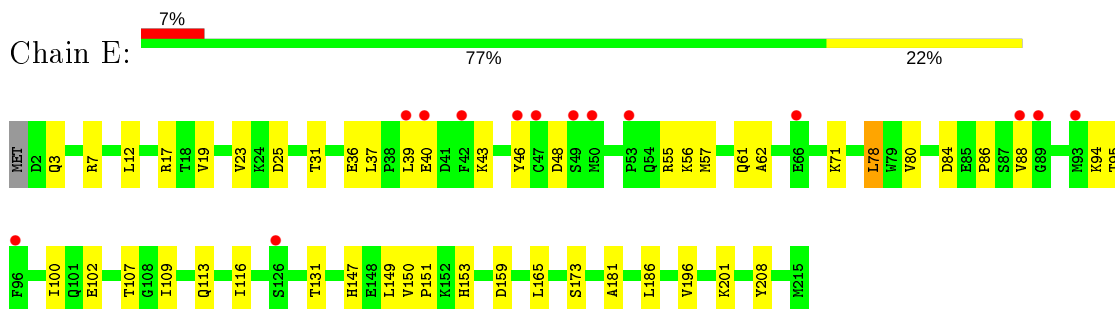




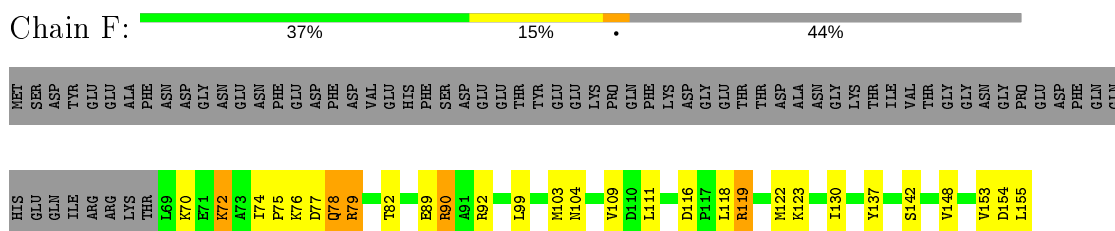




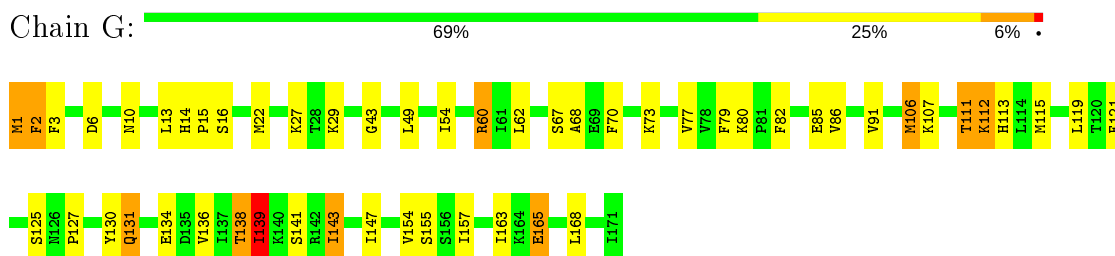
- Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1



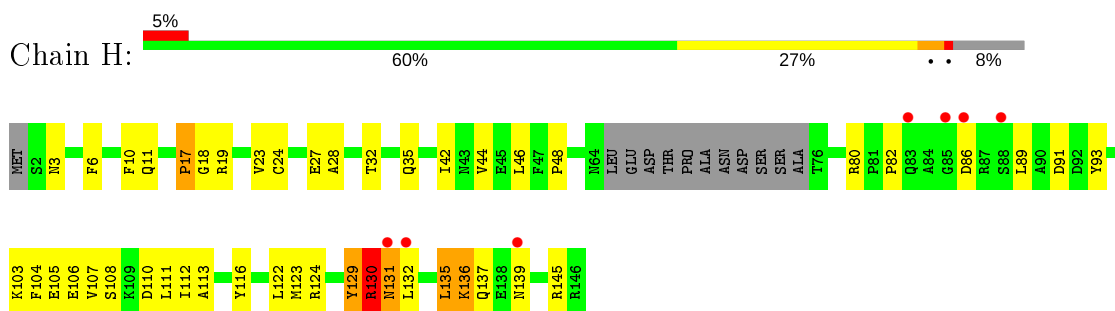
- Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2



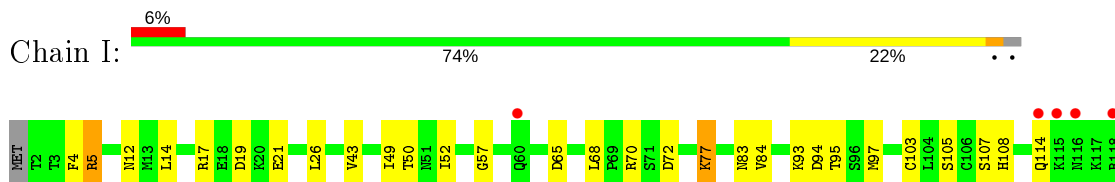
- Molecule 7: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7



- Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

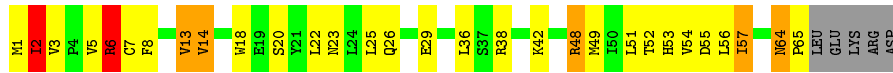


- Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

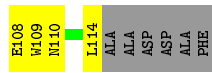




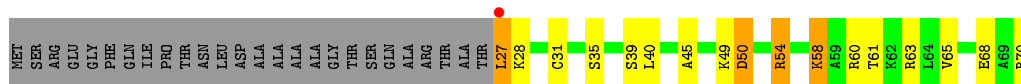
● Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5



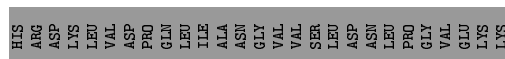
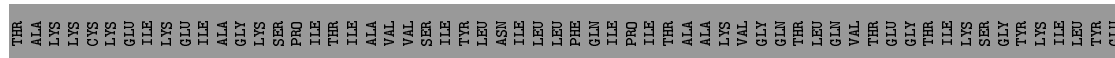
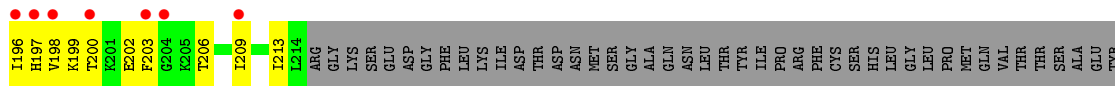
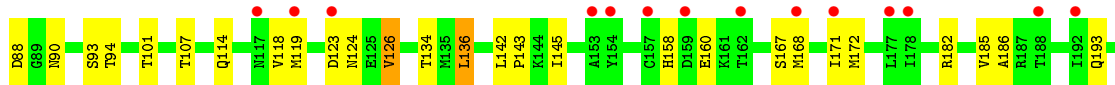
● Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11



● Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4



● Molecule 13: TRANSCRIPTION INITIATION FACTOR IIB



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.44Å 386.76Å 254.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.41 – 3.40 49.41 – 3.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.41-3.40) 100.0 (49.41-3.40)	Depositor EDS
$R_{merge}$	0.38	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.02 (at 3.40Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, $R_{free}$	0.174 , 0.211 0.195 , 0.234	Depositor DCC
$R_{free}$ test set	2995 reflections (2.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.4	Xtriage
Anisotropy	0.347	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 86.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.025 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.029 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	32800	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.97% 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: 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	0.52	2/11322 (0.0%)	0.81	6/15312 (0.0%)
2	B	0.50	0/9271	0.80	2/12505 (0.0%)
3	C	0.53	0/2133	0.81	0/2891
4	D	0.55	0/1444	0.88	1/1935 (0.1%)
5	E	0.45	0/1788	0.71	0/2406
6	F	0.53	0/717	0.78	0/967
7	G	0.52	0/1368	0.82	0/1844
8	H	0.49	0/1094	0.81	0/1481
9	I	0.44	0/989	0.77	0/1331
10	J	0.61	0/541	0.91	0/727
11	K	0.50	0/937	0.70	0/1265
12	L	0.52	0/353	0.89	0/468
13	M	0.50	0/1413	0.85	0/1916
All	All	0.51	2/33370 (0.0%)	0.81	9/45048 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	867	ILE	CG1-CD1	6.57	1.95	1.50
1	A	57	ARG	CA-C	5.52	1.67	1.52

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	470	LYS	C-N-CA	6.89	138.91	121.70
1	A	56	PRO	C-N-CA	6.71	138.47	121.70
1	A	399	HIS	N-CA-CB	5.53	120.55	110.60
1	A	1403	GLU	N-CA-C	5.40	125.59	111.00
1	A	332	LYS	N-CA-C	-5.21	96.93	111.00

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	11123	0	11184	240	0
2	B	9095	0	9055	180	0
3	C	2095	0	2051	43	0
4	D	1434	0	1460	26	0
5	E	1752	0	1776	18	0
6	F	705	0	731	19	0
7	G	1340	0	1357	34	0
8	H	1076	0	1046	23	0
9	I	971	0	927	12	0
10	J	532	0	542	27	0
11	K	919	0	929	26	0
12	L	351	0	374	13	0
13	M	1396	0	1312	19	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
14	M	1	0	0	0	0
15	A	2	0	0	0	0
All	All	32800	0	32744	592	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 9.

The worst 5 of 592 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:867:ILE:CD1	1:A:867:ILE:CG1	1.95	1.41
2:B:515:HIS:HD2	2:B:517:THR:H	1.04	0.93
11:K:65:HIS:HD2	11:K:67:PHE:H	1.09	0.92

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:955:THR:HG22	2:B:956:THR:H	1.33	0.90
12:L:61:THR:HB	12:L:63:ARG:HG3	1.53	0.90

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	1404/1733 (81%)	1248 (89%)	107 (8%)	49 (4%)	3	21
2	B	1134/1224 (93%)	1007 (89%)	94 (8%)	33 (3%)	4	24
3	C	264/318 (83%)	239 (90%)	22 (8%)	3 (1%)	14	44
4	D	174/221 (79%)	152 (87%)	16 (9%)	6 (3%)	3	21
5	E	212/215 (99%)	196 (92%)	14 (7%)	2 (1%)	17	49
6	F	85/155 (55%)	79 (93%)	4 (5%)	2 (2%)	6	28
7	G	169/171 (99%)	151 (89%)	14 (8%)	4 (2%)	6	28
8	H	130/146 (89%)	110 (85%)	14 (11%)	6 (5%)	2	15
9	I	117/122 (96%)	96 (82%)	19 (16%)	2 (2%)	9	34
10	J	63/70 (90%)	54 (86%)	6 (10%)	3 (5%)	2	15
11	K	112/120 (93%)	107 (96%)	4 (4%)	1 (1%)	17	49
12	L	42/70 (60%)	28 (67%)	12 (29%)	2 (5%)	2	15
13	M	191/345 (55%)	176 (92%)	12 (6%)	3 (2%)	9	34
All	All	4097/4910 (83%)	3643 (89%)	338 (8%)	116 (3%)	5	24

5 of 116 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ARG

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	58	LEU
1	A	72	GLU
1	A	74	MET
1	A	254	GLU

### 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	1236/1520 (81%)	1063 (86%)	173 (14%)	3	13
2	B	980/1061 (92%)	835 (85%)	145 (15%)	3	12
3	C	234/274 (85%)	203 (87%)	31 (13%)	4	15
4	D	160/200 (80%)	129 (81%)	31 (19%)	1	4
5	E	196/197 (100%)	176 (90%)	20 (10%)	7	26
6	F	77/137 (56%)	69 (90%)	8 (10%)	7	25
7	G	152/152 (100%)	134 (88%)	18 (12%)	5	19
8	H	118/128 (92%)	98 (83%)	20 (17%)	2	8
9	I	113/116 (97%)	101 (89%)	12 (11%)	6	24
10	J	60/65 (92%)	48 (80%)	12 (20%)	1	3
11	K	99/102 (97%)	87 (88%)	12 (12%)	5	18
12	L	39/57 (68%)	31 (80%)	8 (20%)	1	3
13	M	142/299 (48%)	109 (77%)	33 (23%)	1	2
All	All	3606/4308 (84%)	3083 (86%)	523 (14%)	3	12

5 of 523 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	589	VAL
2	B	1060	ARG
12	L	50	ASP
2	B	616	ILE
2	B	835	GLN



Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	443	ASN
2	B	744	HIS
12	L	53	HIS
2	B	465	ASN
2	B	516	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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](#)

Of 11 ligands modelled in this entry, 11 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

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	1414/1733 (81%)	-0.08	26 (1%) 68 67	30, 65, 118, 185	0
2	B	1150/1224 (93%)	-0.10	44 (3%) 40 39	29, 69, 132, 191	0
3	C	266/318 (83%)	-0.16	2 (0%) 86 85	30, 53, 88, 140	0
4	D	178/221 (80%)	0.06	2 (1%) 80 79	43, 75, 120, 138	0
5	E	214/215 (99%)	0.32	14 (6%) 18 20	47, 101, 154, 173	0
6	F	87/155 (56%)	-0.28	0 100 100	32, 56, 88, 98	0
7	G	171/171 (100%)	-0.23	0 100 100	35, 56, 84, 104	0
8	H	134/146 (91%)	0.31	7 (5%) 27 27	55, 92, 134, 160	0
9	I	119/122 (97%)	0.21	7 (5%) 22 23	68, 98, 135, 166	0
10	J	65/70 (92%)	-0.44	0 100 100	35, 50, 76, 92	0
11	K	114/120 (95%)	-0.16	0 100 100	35, 56, 88, 100	0
12	L	44/70 (62%)	-0.04	1 (2%) 60 59	52, 90, 122, 136	0
13	M	193/345 (55%)	0.68	23 (11%) 4 5	52, 101, 140, 154	0
All	All	4149/4910 (84%)	-0.03	126 (3%) 50 49	29, 68, 132, 191	0

The worst 5 of 126 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	250	PHE	5.1
2	B	709	ASP	4.7
12	L	27	LEU	4.6
9	I	119	THR	4.2
1	A	421	ALA	4.2

### 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 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<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( $\text{\AA}^2$ )	Q<0.9
15	MG	A	2459	1/1	0.83	0.44	80,80,80,80	0
15	MG	A	2458	1/1	0.87	0.29	75,75,75,75	0
14	ZN	L	1071	1/1	0.98	0.06	98,98,98,98	0
14	ZN	I	1122	1/1	0.98	0.03	130,130,130,130	0
14	ZN	C	1269	1/1	0.99	0.07	64,64,64,64	0
14	ZN	A	2457	1/1	0.99	0.12	51,51,51,51	0
14	ZN	I	1121	1/1	0.99	0.08	78,78,78,78	0
14	ZN	M	1216	1/1	0.99	0.07	73,73,73,73	0
14	ZN	A	2456	1/1	0.99	0.06	91,91,91,91	0
14	ZN	B	2225	1/1	1.00	0.13	46,46,46,46	0
14	ZN	J	1066	1/1	1.00	0.15	46,46,46,46	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.