



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 26, 2023 – 01:12 PM EDT

PDB ID : 3FKI  
Title : 12-Subunit RNA Polymerase II Refined with Zn-SAD data  
Authors : Meyer, P.A.; Ye, P.; Suh, M.H.; Zhang, M.; Fu, J.  
Deposited on : 2008-12-16  
Resolution : 3.88 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

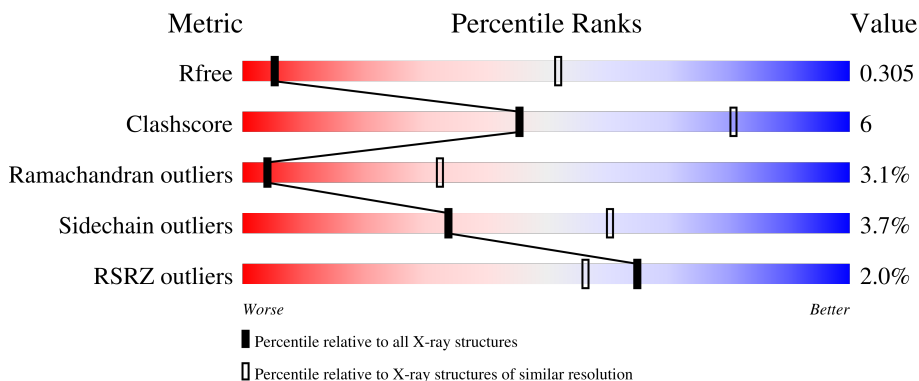
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.88 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 (4.12-3.64)
Clashscore	141614	1045 (4.10-3.66)
Ramachandran outliers	138981	1008 (4.10-3.66)
Sidechain outliers	138945	1001 (4.10-3.66)
RSRZ outliers	127900	1213 (4.16-3.60)

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	1733	 2% 66% 15% 18%
2	B	1224	 2% 73% 17% 8%
3	C	318	 2% 65% 16% 16%
4	D	221	 2% 66% 13% 19%
5	E	215	 2% 88% 12%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	F	155	
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	

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
14	MG	A	1736	-	-	-	X

## 2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 31412 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	1427	11225	7069	1964	2130	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	1125	8947	5662	1568	1661	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	267	2104	1323	350	418	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	1436	887	257	290	2	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

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 RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	87	704	451	119	131	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 RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	136	1089	685	184	215	5	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 RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	66	540	345	94	95	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	115	924	593	157	172	2	0	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	47	371	228	73	66	4	0	0	0

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	2	Total	Zn	0	0
			2	2		
13	B	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	C	1	Total 1	Zn 1	0	0
13	I	2	Total 2	Zn 2	0	0
13	J	1	Total 1	Zn 1	0	0
13	L	1	Total 1	Zn 1	0	0

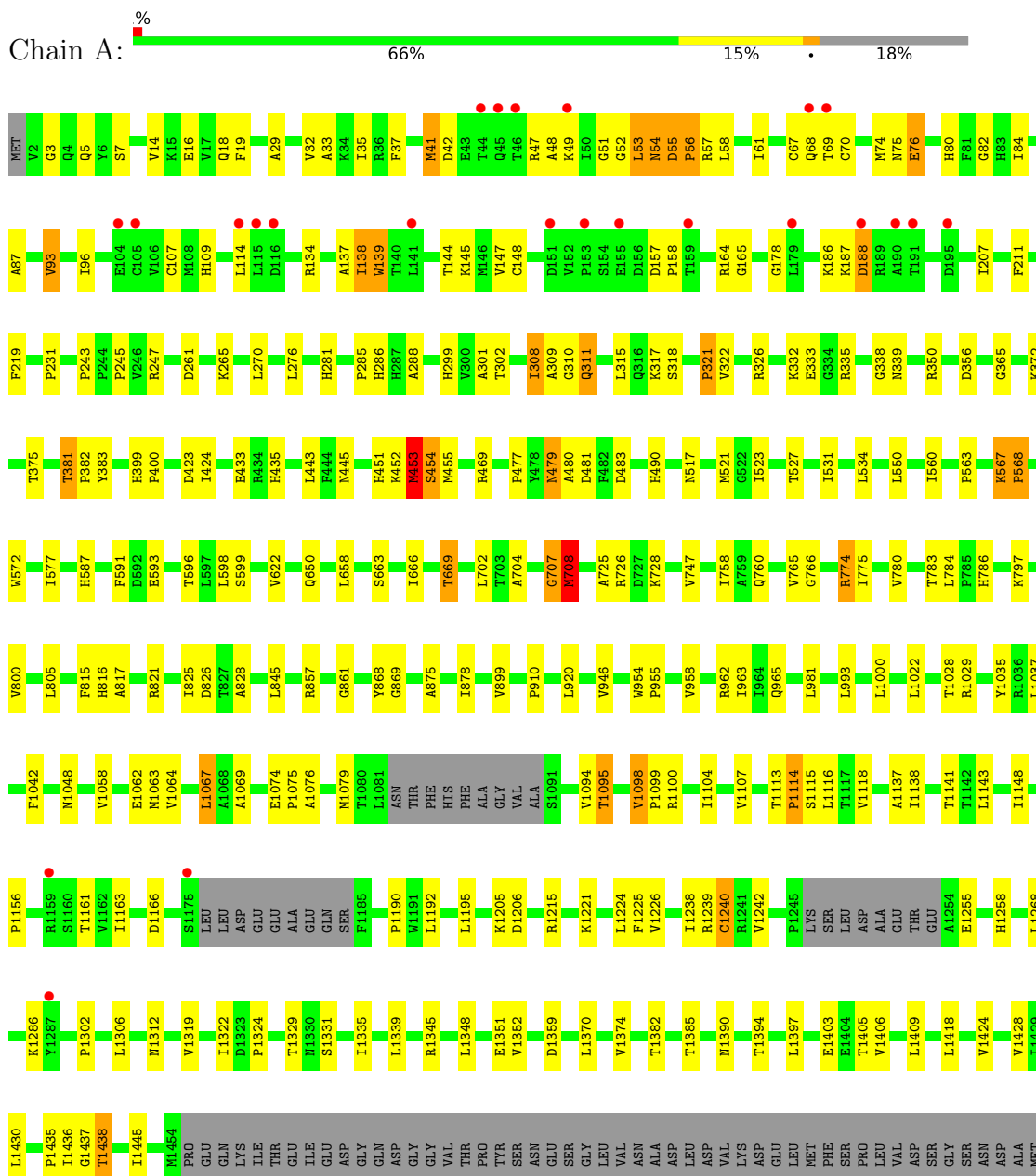
- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

### 3 Residue-property plots [i](#)

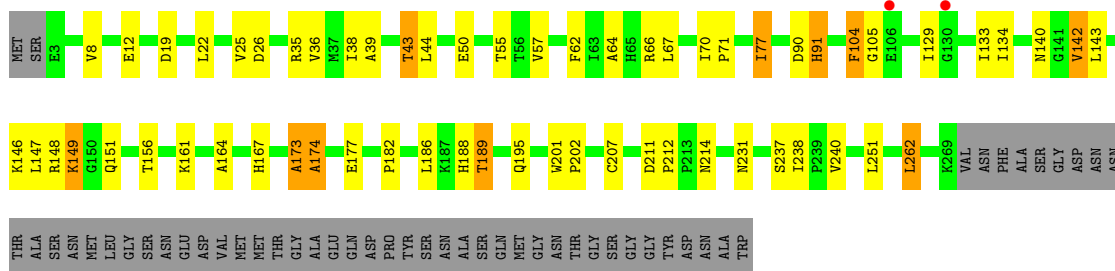
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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

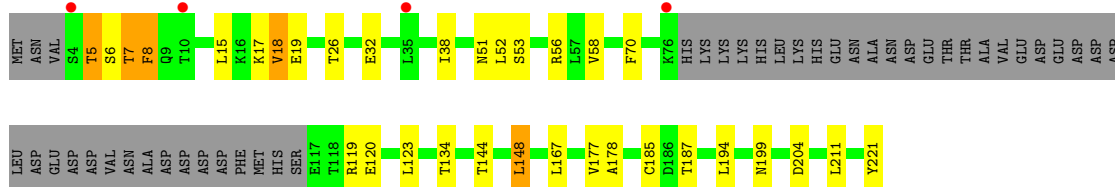




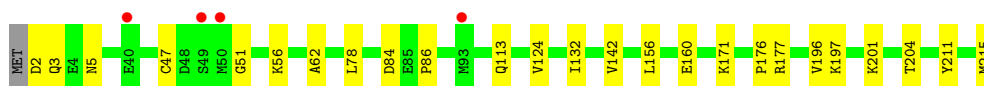
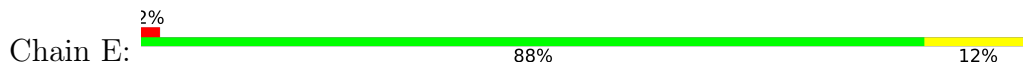




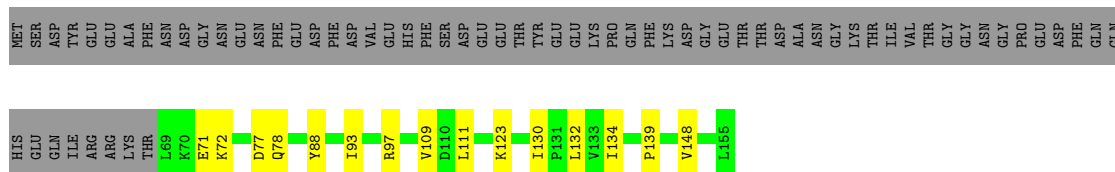
- Molecule 4: DNA-directed RNA polymerase II subunit RPB4



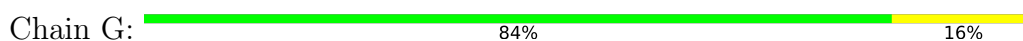
- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1



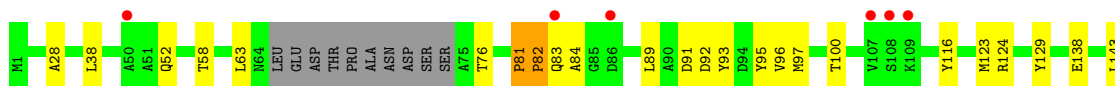
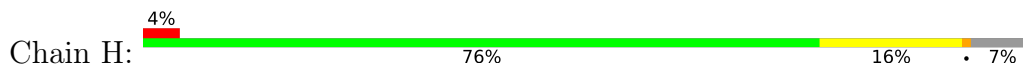
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

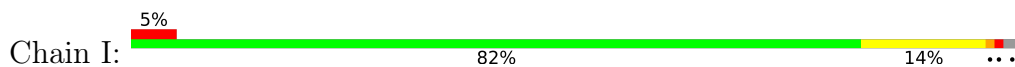


- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3





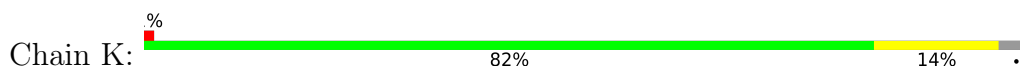
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



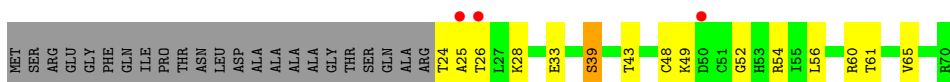
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



- Molecule 11: DNA-directed RNA polymerase II subunit RPB11



- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.58Å 391.54Å 280.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.22 – 3.88 50.07 – 3.88	Depositor EDS
% Data completeness (in resolution range)	95.5 (20.22-3.88) 95.5 (50.07-3.88)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 3.88Å)	Xtriage
Refinement program	REFMAC 5D	Depositor
R, $R_{free}$	0.282 , 0.301 0.284 , 0.305	Depositor DCC
$R_{free}$ test set	5355 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	94.3	Xtriage
Anisotropy	0.422	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 21.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.36$ , $\langle L^2 \rangle = 0.19$	Xtriage
Estimated twinning fraction	0.098 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.105 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	31412	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.87% 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

### 5.1 Standard geometry

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	0.30	0/11426	0.47	0/15451
2	B	0.31	0/9122	0.48	0/12300
3	C	0.29	0/2142	0.45	0/2902
4	D	0.31	0/1446	0.49	0/1937
5	E	0.31	0/1788	0.44	0/2406
6	F	0.28	0/716	0.47	0/967
7	G	0.31	0/1368	0.45	0/1844
8	H	0.33	0/1107	0.48	0/1498
9	I	0.33	0/989	0.48	0/1331
10	J	0.30	0/549	0.49	0/738
11	K	0.31	0/942	0.47	0/1272
12	L	0.34	0/373	0.54	0/495
All	All	0.31	0/31968	0.47	0/43141

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	508	LEU	Peptide

## 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	11225	0	11293	170	0
2	B	8947	0	8975	135	0
3	C	2104	0	2064	33	0
4	D	1436	0	1457	20	0
5	E	1752	0	1776	14	0
6	F	704	0	731	13	0
7	G	1340	0	1357	17	0
8	H	1089	0	1063	16	0
9	I	971	0	932	8	0
10	J	540	0	554	11	0
11	K	924	0	934	12	0
12	L	371	0	394	4	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
14	A	1	0	0	0	0
All	All	31412	0	31530	401	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 6.

The worst 5 of 401 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)
4:D:17:LYS:HA	4:D:18:VAL:HB	1.30	1.11
1:A:567:LYS:HB3	1:A:568:PRO:HD3	1.30	1.10
6:F:71:GLU:HA	6:F:72:LYS:HB3	1.08	1.07
2:B:471:LYS:HB3	2:B:472:ALA:HA	1.37	1.04
1:A:187:LYS:HB3	1:A:188:ASP:HA	1.36	1.03

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	1419/1733 (82%)	1228 (86%)	141 (10%)	50 (4%)	3	30
2	B	1111/1224 (91%)	957 (86%)	117 (10%)	37 (3%)	4	31
3	C	265/318 (83%)	236 (89%)	20 (8%)	9 (3%)	3	31
4	D	174/221 (79%)	148 (85%)	21 (12%)	5 (3%)	4	33
5	E	212/215 (99%)	193 (91%)	19 (9%)	0	100	100
6	F	85/155 (55%)	77 (91%)	8 (9%)	0	100	100
7	G	169/171 (99%)	146 (86%)	21 (12%)	2 (1%)	13	49
8	H	132/146 (90%)	106 (80%)	19 (14%)	7 (5%)	2	23
9	I	117/122 (96%)	99 (85%)	15 (13%)	3 (3%)	5	35
10	J	64/70 (91%)	57 (89%)	4 (6%)	3 (5%)	2	25
11	K	113/120 (94%)	110 (97%)	3 (3%)	0	100	100
12	L	45/70 (64%)	33 (73%)	7 (16%)	5 (11%)	0	8
All	All	3906/4565 (86%)	3390 (87%)	395 (10%)	121 (3%)	4	32

5 of 121 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	55	ASP
1	A	139	TRP
1	A	311	GLN
1	A	321	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	1247/1520 (82%)	1206 (97%)	41 (3%)	38	63
2	B	976/1061 (92%)	937 (96%)	39 (4%)	31	58
3	C	235/274 (86%)	221 (94%)	14 (6%)	19	49
4	D	160/200 (80%)	150 (94%)	10 (6%)	18	47
5	E	196/197 (100%)	194 (99%)	2 (1%)	76	86
6	F	77/137 (56%)	77 (100%)	0	100	100
7	G	152/152 (100%)	148 (97%)	4 (3%)	46	68
8	H	119/128 (93%)	117 (98%)	2 (2%)	60	78
9	I	113/116 (97%)	107 (95%)	6 (5%)	22	52
10	J	61/65 (94%)	57 (93%)	4 (7%)	16	46
11	K	99/102 (97%)	96 (97%)	3 (3%)	41	64
12	L	41/57 (72%)	37 (90%)	4 (10%)	8	31
All	All	3476/4009 (87%)	3347 (96%)	129 (4%)	34	60

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	75	CYS
10	J	46	CYS
2	B	294	ASP
2	B	261	ARG
10	J	51	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
9	I	12	ASN

### 5.3.3 RNA

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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 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	1427/1733 (82%)	0.12	24 (1%) 70 61	78, 111, 134, 158	0
2	B	1125/1224 (91%)	0.08	28 (2%) 57 47	79, 107, 133, 155	0
3	C	267/318 (83%)	0.00	2 (0%) 87 82	87, 103, 117, 127	0
4	D	178/221 (80%)	0.14	4 (2%) 62 52	114, 120, 129, 135	0
5	E	214/215 (99%)	0.17	4 (1%) 66 58	93, 122, 143, 152	0
6	F	87/155 (56%)	-0.10	0 100 100	89, 97, 104, 106	0
7	G	171/171 (100%)	0.25	0 100 100	113, 116, 127, 137	0
8	H	136/146 (93%)	0.40	6 (4%) 34 28	107, 126, 142, 151	0
9	I	119/122 (97%)	0.24	6 (5%) 28 24	109, 126, 143, 152	0
10	J	66/70 (94%)	-0.17	1 (1%) 73 65	91, 104, 117, 123	0
11	K	115/120 (95%)	-0.01	1 (0%) 84 78	80, 93, 117, 125	0
12	L	47/70 (67%)	0.26	3 (6%) 19 14	88, 112, 121, 127	0
All	All	3952/4565 (86%)	0.11	79 (1%) 65 56	78, 112, 135, 158	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	505	ASP	5.3
9	I	119	THR	4.5
2	B	443	ASN	4.2
8	H	108	SER	4.0
2	B	716	ASN	3.8

### 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( $\text{\AA}^2$ )	Q<0.9
14	MG	A	1736	1/1	0.75	0.84	114,114,114,114	0
13	ZN	A	1735	1/1	0.88	0.14	114,114,114,114	0
13	ZN	I	123	1/1	0.92	0.09	114,114,114,114	0
13	ZN	J	71	1/1	0.94	0.14	114,114,114,114	0
13	ZN	I	124	1/1	0.94	0.05	114,114,114,114	0
13	ZN	L	71	1/1	0.97	0.08	114,114,114,114	0
13	ZN	A	1734	1/1	0.97	0.07	114,114,114,114	0
13	ZN	C	319	1/1	0.99	0.12	114,114,114,114	0
13	ZN	B	1225	1/1	0.99	0.17	114,114,114,114	0

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

There are no such residues in this entry.