



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:43 AM GMT

PDB ID : 3MVD  
Title : Crystal structure of the chromatin factor RCC1 in complex with the nucleosome core particle  
Authors : Makde, R.D.; England, J.R.; Yennawar, H.P.; Tan, S.  
Deposited on : 2010-05-04  
Resolution : 2.90 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

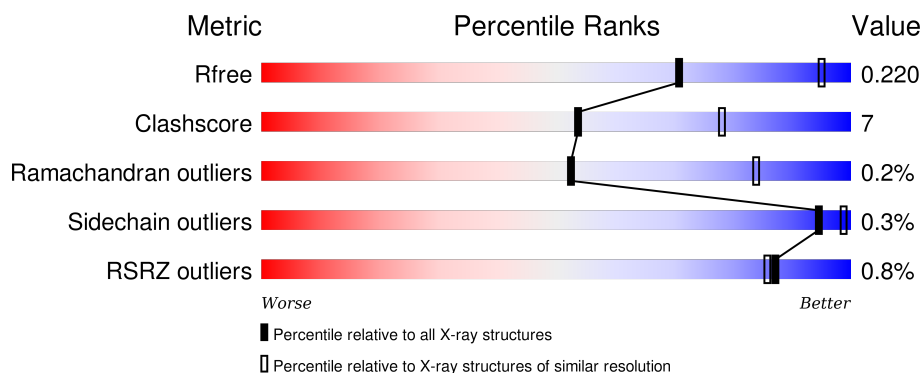
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	A	135	<div> <div> <div></div> <div>65%</div> <div>7%</div> <div>27%</div> </div> </div>
1	E	135	<div> <div> <div></div> <div>66%</div> <div>.</div> <div>30%</div> </div> </div>
2	B	102	<div> <div>2%</div> <div> <div></div> <div>70%</div> <div>11%</div> <div>20%</div> </div> </div>
2	F	102	<div> <div> <div></div> <div>73%</div> <div>12%</div> <div>16%</div> </div> </div>
3	C	129	<div> <div> <div></div> <div>72%</div> <div>11%</div> <div>17%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	129	 71%12%17%
4	D	122	 66%11%24%
4	H	122	 65%11%24%
5	I	147	 3%60%39%. .
6	J	147	 3%60%38%.. ..
7	K	423	 76%16%8%
7	L	423	 79%14%7%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 17743 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 Histone H3.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	S	0	0	0
			801	506	153	139	3			
1	E	95	Total	C	N	O	S	0	0	0
			779	492	148	136	3			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	ALA	GLY	SEE REMARK 999	UNP P84233
E	102	ALA	GLY	SEE REMARK 999	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	82	Total	C	N	O	S	0	0	0
			653	413	127	112	1			
2	F	86	Total	C	N	O	S	0	0	0
			672	424	130	117	1			

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	107	Total	C	N	O		0	0	0
			811	510	158	143				
3	G	107	Total	C	N	O		0	0	0
			815	513	159	143				

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	93	Total	C	N	O	S	0	0	0
			718	451	128	137	2			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	93	Total	C	N	O	S	0	0	0
			726	457	130	137	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	29	THR	SER	SEE REMARK 999	UNP P02281
H	29	THR	SER	SEE REMARK 999	UNP P02281

- Molecule 5 is a DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	146	Total	C	N	O	P	0	0	0
			2975	1413	540	876	146			

- Molecule 6 is a DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	146	Total	C	N	O	P	0	0	0
			3011	1425	564	876	146			

- Molecule 7 is a protein called Regulator of chromosome condensation.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	K	391	Total	C	N	O	S	0	0	0
			2893	1802	516	559	16			
7	L	394	Total	C	N	O	S	0	0	0
			2889	1798	515	560	16			

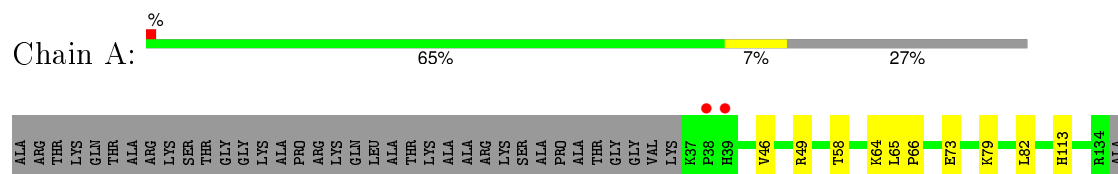
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	0	GLY	-	EXPRESSION TAG	UNP P25171
K	1	SER	-	EXPRESSION TAG	UNP P25171
L	0	GLY	-	EXPRESSION TAG	UNP P25171
L	1	SER	-	EXPRESSION TAG	UNP P25171

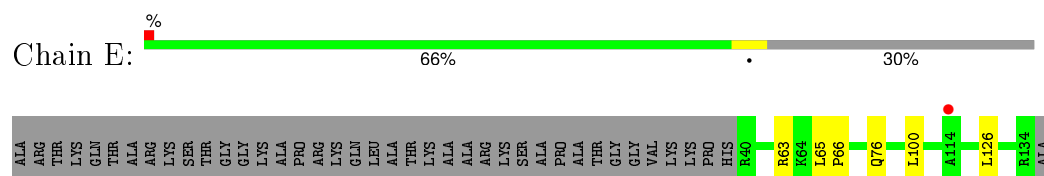
### 3 Residue-property plots [i](#)

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.

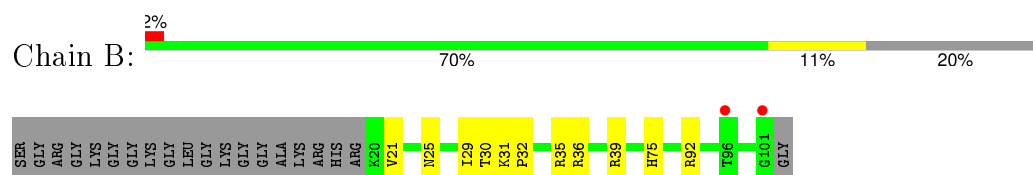
- Molecule 1: Histone H3.2



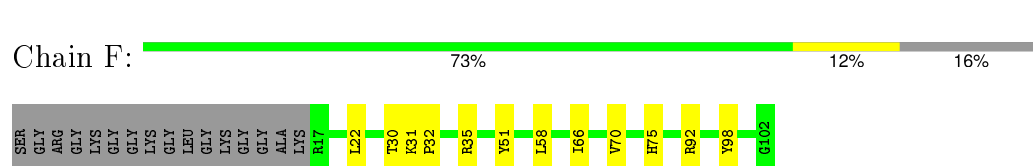
- Molecule 1: Histone H3.2



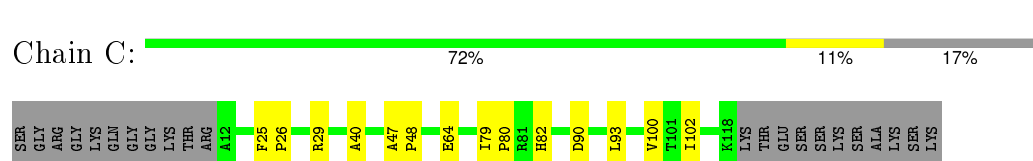
- Molecule 2: Histone H4



- Molecule 2: Histone H4



- Molecule 3: Histone H2A

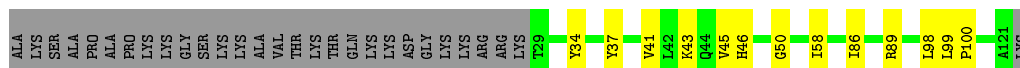


- Molecule 3: Histone H2A

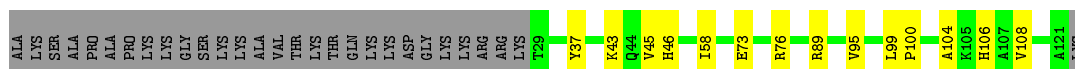




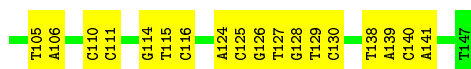
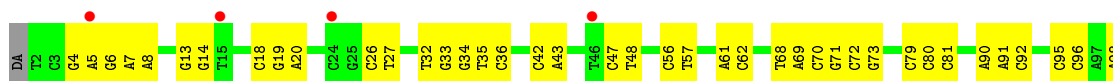
- Molecule 4: Histone H2B 1.1



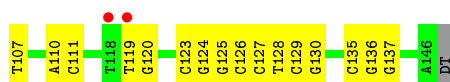
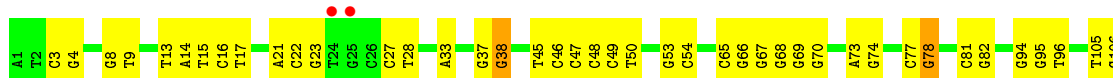
- Molecule 4: Histone H2B 1.1



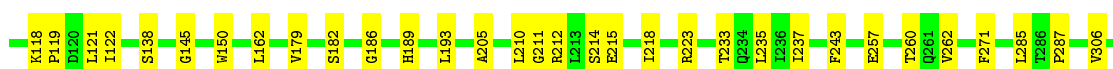
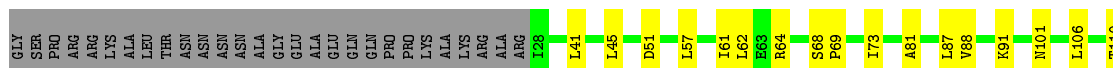
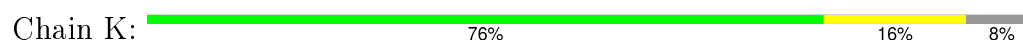
- Molecule 5: DNA (146-MER)



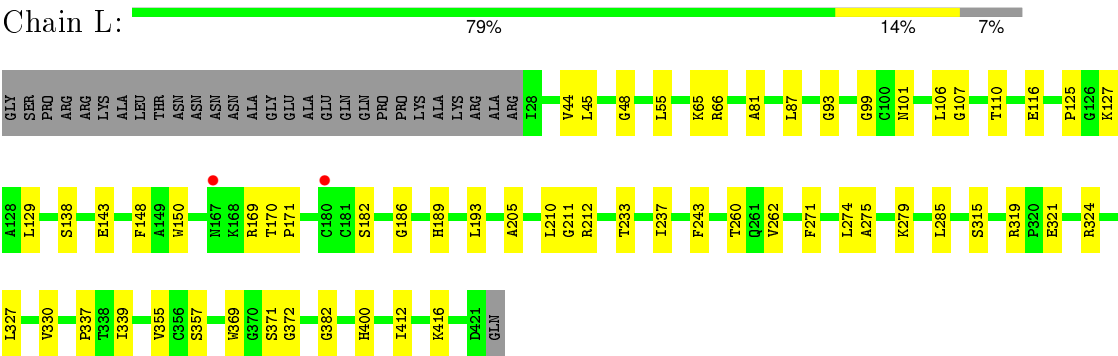
- Molecule 6: DNA (146-MER)



- Molecule 7: Regulator of chromosome condensation



- Molecule 7: Regulator of chromosome condensation





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.68Å 183.04Å 107.03Å 90.00° 101.52° 90.00°	Depositor
Resolution (Å)	34.70 – 2.90 48.99 – 2.89	Depositor EDS
% Data completeness (in resolution range)	94.1 (34.70-2.90) 98.3 (48.99-2.89)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.34 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6 _289)	Depositor
R, $R_{free}$	0.175 , 0.215 0.185 , 0.220	Depositor DCC
$R_{free}$ test set	2088 reflections (2.54%)	DCC
Wilson B-factor (Å <sup>2</sup> )	77.1	Xtriage
Anisotropy	0.434	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 46.6	EDS
Estimated twinning fraction	0.017 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 85000 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	17743	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.73% 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](#)

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	0.21	0/813	0.35	0/1093
1	E	0.20	0/789	0.36	0/1059
2	B	0.20	0/660	0.37	0/885
2	F	0.20	0/680	0.37	0/912
3	C	0.19	0/821	0.34	0/1112
3	G	0.19	0/825	0.35	0/1116
4	D	0.21	0/729	0.37	0/985
4	H	0.21	0/737	0.37	0/993
5	I	0.41	0/3333	1.01	0/5137
6	J	0.41	0/3381	1.03	3/5221 (0.1%)
7	K	0.20	0/2938	0.38	0/3975
7	L	0.20	0/2934	0.38	0/3976
All	All	0.29	0/18640	0.70	3/26464 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	J	78	DG	O4'-C4'-C3'	-8.00	101.20	106.00
6	J	38	DG	O4'-C4'-C3'	-6.26	102.00	104.50
6	J	78	DG	C4'-C3'-C2'	-5.35	98.28	103.10

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	801	0	831	9	0
1	E	779	0	815	5	0
2	B	653	0	695	10	0
2	F	672	0	698	8	0
3	C	811	0	849	11	0
3	G	815	0	860	14	0
4	D	718	0	725	12	0
4	H	726	0	747	11	0
5	I	2975	0	1639	37	0
6	J	3011	0	1639	43	0
7	K	2893	0	2871	44	0
7	L	2889	0	2831	35	0
All	All	17743	0	15200	215	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 7.

The worst 5 of 215 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)
5:I:13:DG:H2''	5:I:14:DG:H5'	1.56	0.87
7:L:87:LEU:HD11	7:L:138:SER:HB3	1.60	0.83
7:L:260:THR:HG23	7:L:262:VAL:H	1.50	0.75
2:F:75:HIS:HA	4:H:89:ARG:HH12	1.52	0.74
5:I:98:DA:H5'	5:I:98:DA:C8	2.22	0.74

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	96/135 (71%)	95 (99%)	1 (1%)	0	<b>100</b> <b>100</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	93/135 (69%)	90 (97%)	3 (3%)	0	100	100
2	B	80/102 (78%)	78 (98%)	2 (2%)	0	100	100
2	F	84/102 (82%)	81 (96%)	3 (4%)	0	100	100
3	C	105/129 (81%)	102 (97%)	3 (3%)	0	100	100
3	G	105/129 (81%)	104 (99%)	1 (1%)	0	100	100
4	D	91/122 (75%)	90 (99%)	1 (1%)	0	100	100
4	H	91/122 (75%)	89 (98%)	2 (2%)	0	100	100
7	K	389/423 (92%)	370 (95%)	18 (5%)	1 (0%)	46	79
7	L	392/423 (93%)	370 (94%)	20 (5%)	2 (0%)	34	71
All	All	1526/1822 (84%)	1469 (96%)	54 (4%)	3 (0%)	52	84

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	L	279	LYS
7	K	337	PRO
7	L	337	PRO

### 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	84/110 (76%)	84 (100%)	0	100	100
1	E	82/110 (74%)	81 (99%)	1 (1%)	78	94
2	B	67/78 (86%)	66 (98%)	1 (2%)	72	92
2	F	67/78 (86%)	66 (98%)	1 (2%)	72	92
3	C	81/101 (80%)	81 (100%)	0	100	100
3	G	82/101 (81%)	82 (100%)	0	100	100
4	D	77/102 (76%)	77 (100%)	0	100	100
4	H	79/102 (78%)	79 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	K	308/342 (90%)	307 (100%)	1 (0%)	94	99
7	L	303/342 (89%)	303 (100%)	0	100	100
All	All	1230/1466 (84%)	1226 (100%)	4 (0%)	94	99

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	21	VAL
1	E	63	ARG
2	F	92	ARG
7	K	390	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	125	GLN
3	G	24	GLN
7	L	43	ASN
1	E	76	GLN
7	K	234	GLN

### 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](#)

There are no ligands in this entry.

## 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	98/135 (72%)	0.17	2 (2%) 68 64	42, 66, 97, 149	0
1	E	95/135 (70%)	0.06	1 (1%) 82 80	43, 61, 87, 110	0
2	B	82/102 (80%)	0.10	2 (2%) 62 57	41, 61, 97, 117	0
2	F	86/102 (84%)	0.01	0 100 100	37, 56, 102, 149	0
3	C	107/129 (82%)	0.11	0 100 100	42, 60, 94, 132	0
3	G	107/129 (82%)	0.08	0 100 100	41, 63, 94, 138	0
4	D	93/122 (76%)	-0.10	0 100 100	41, 59, 95, 127	0
4	H	93/122 (76%)	-0.12	0 100 100	41, 61, 94, 118	0
5	I	146/147 (99%)	0.42	4 (2%) 58 52	70, 117, 150, 195	0
6	J	146/147 (99%)	0.44	4 (2%) 58 52	70, 112, 152, 183	0
7	K	391/423 (92%)	-0.12	0 100 100	47, 73, 107, 134	0
7	L	394/423 (93%)	-0.09	2 (0%) 91 90	50, 84, 118, 160	0
All	All	1838/2116 (86%)	0.04	15 (0%) 87 86	37, 74, 131, 195	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	114	ALA	3.5
6	J	118	DT	3.1
7	L	180	CYS	3.1
5	I	15	DT	2.7
7	L	167	ASN	2.7

### 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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.