



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 04:11 am BST

PDB ID : 3X1V  
Title : Crystal structure of nucleosome core particle in the presence of histone variant involved in reprogramming  
Authors : Sivaraman, P.; Kumarevel, T.S.  
Deposited on : 2014-11-28  
Resolution : 2.92 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  
Xtrriage (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

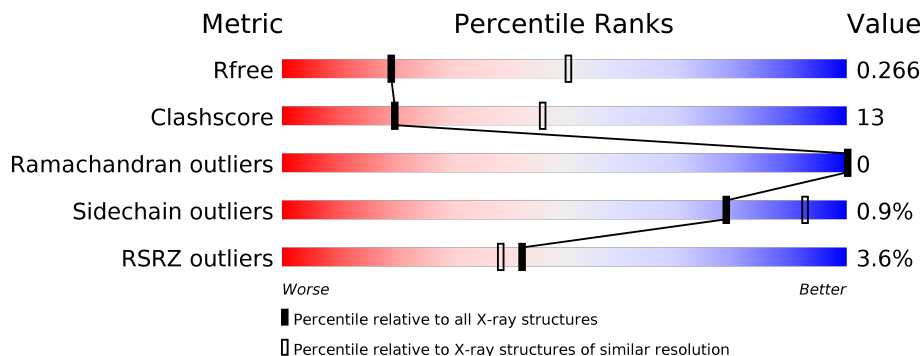
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.92 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	2307 (2.94-2.90)
Clashscore	141614	2531 (2.94-2.90)
Ramachandran outliers	138981	2462 (2.94-2.90)
Sidechain outliers	138945	2464 (2.94-2.90)
RSRZ outliers	127900	2248 (2.94-2.90)

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 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	I	146	
1	J	146	
2	A	135	
2	E	135	
3	B	102	
3	F	102	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	C	129	
4	G	129	
5	D	126	
5	H	126	

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
6	MN	J	301	-	-	-	X
7	CL	D	202	-	-	X	-

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 12197 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 DNA chain called DNA (146-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	I	146	2990	1431	540	874	145	0	0	0
1	J	146	2990	1431	540	874	145	0	0	0

- Molecule 2 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	98	807	508	156	139	4	0	0	0
2	E	99	816	514	158	140	4	0	0	0

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	79	627	395	121	110	1	0	0	0
3	F	87	703	442	142	118	1	0	0	0

- Molecule 4 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	C	106	819	517	160	142	0	0	0
4	G	109	844	532	167	145	0	0	0

- Molecule 5 is a protein called Histone H2B type 1-A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	99	Total	C	N	O	S	0	1	0
			797	499	149	146	3			
5	H	94	Total	C	N	O	S	0	1	0
			746	469	133	141	3			

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	G	1	Total	Mn	0	0
			1	1		
6	J	6	Total	Mn	0	0
			6	6		
6	I	11	Total	Mn	0	0
			11	11		
6	D	1	Total	Mn	0	0
			1	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	J	2	Total	Cl	0	0
			2	2		
7	I	2	Total	Cl	0	0
			2	2		
7	D	1	Total	Cl	0	0
			1	1		
7	B	1	Total	Cl	0	0
			1	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	12	Total	O	0	0
			12	12		
8	J	5	Total	O	0	0
			5	5		
8	A	3	Total	O	0	0
			3	3		
8	C	2	Total	O	0	0
			2	2		
8	D	2	Total	O	0	0
			2	2		

*Continued on next page...*

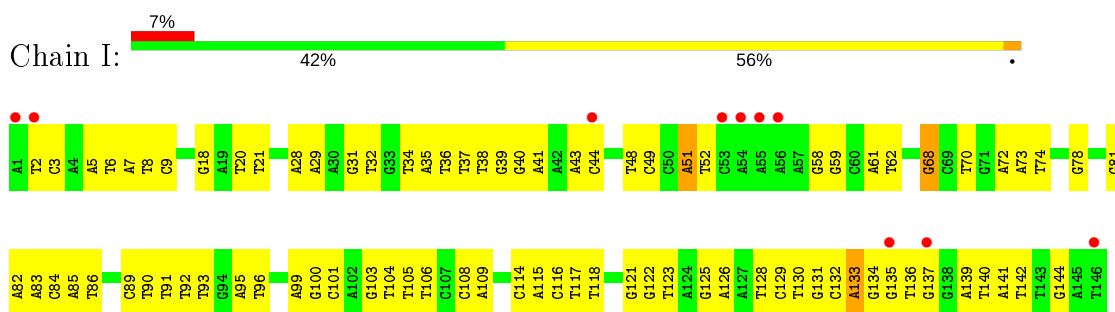
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
8	E	4	Total O 4 4	0	0
8	F	4	Total O 4 4	0	0
8	G	1	Total O 1 1	0	0

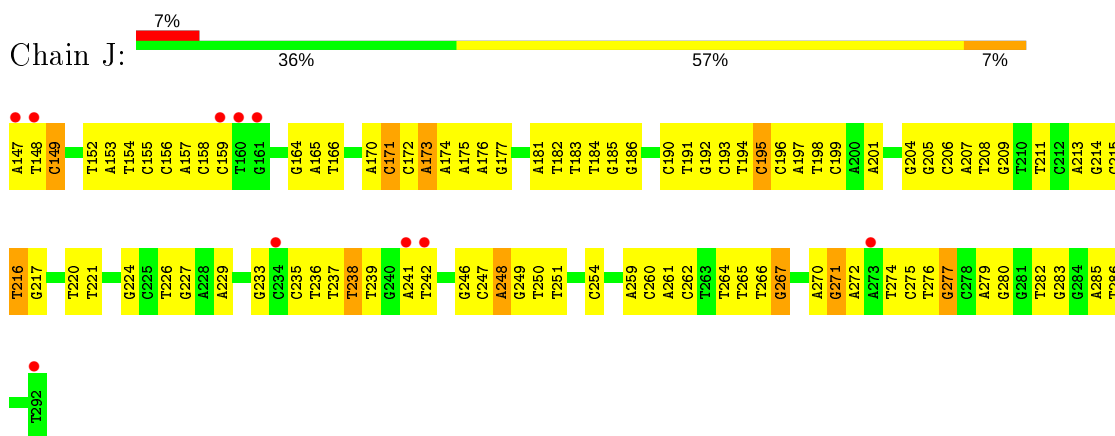
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

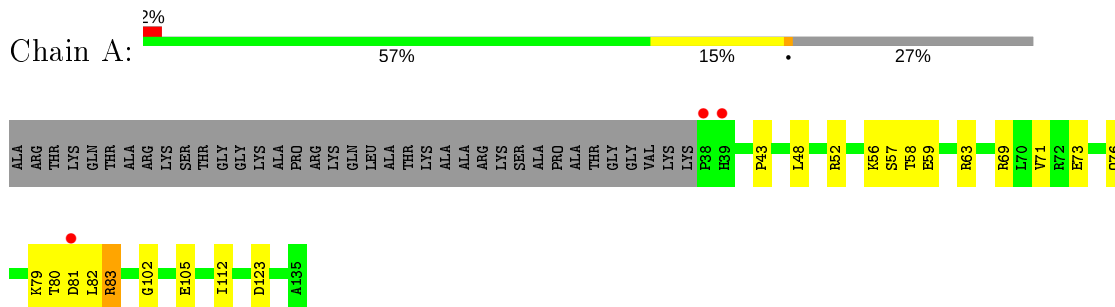
- Molecule 1: DNA (146-MER)



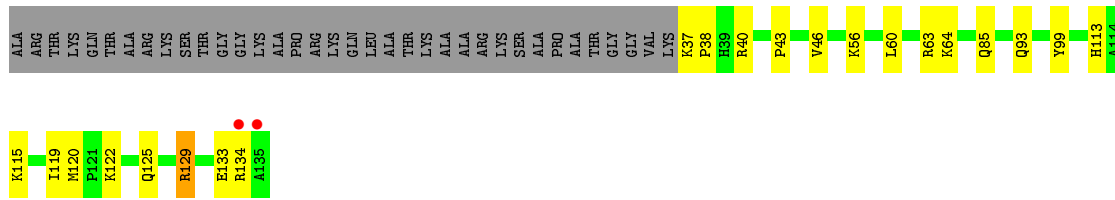
- Molecule 1: DNA (146-MER)



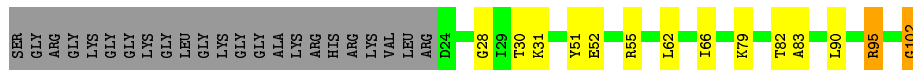
- Molecule 2: Histone H3.1



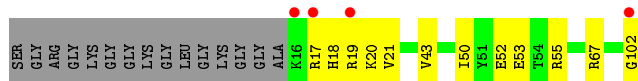
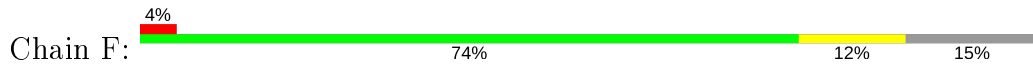
- Molecule 2: Histone H3.1



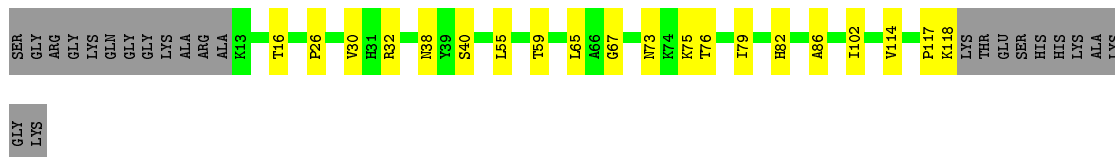
• Molecule 3: Histone H4



• Molecule 3: Histone H4



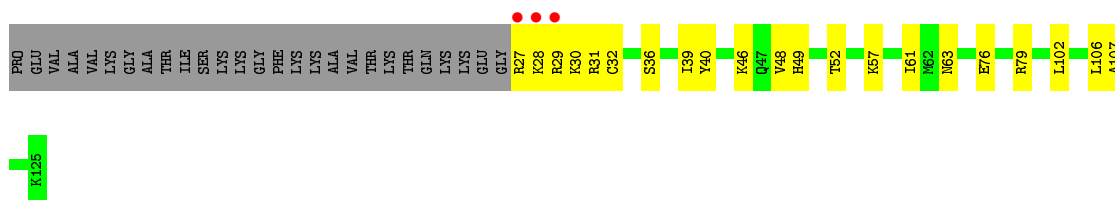
• Molecule 4: Histone H2A type 1-B/E



• Molecule 4: Histone H2A type 1-B/E



• Molecule 5: Histone H2B type 1-A







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.33Å 108.47Å 168.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.81 – 2.92 39.81 – 2.92	Depositor EDS
% Data completeness (in resolution range)	94.3 (39.81-2.92) 94.0 (39.81-2.92)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.60 (at 2.90Å)	Xtriage
Refinement program	REFMAC, PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.191 , 0.259 0.203 , 0.266	Depositor DCC
$R_{free}$ test set	1940 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.8	Xtriage
Anisotropy	0.277	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 47.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12197	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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: MN, CL

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	I	0.90	3/3354 (0.1%)	1.07	8/5175 (0.2%)
1	J	1.00	8/3354 (0.2%)	1.14	11/5175 (0.2%)
2	A	0.61	0/819	0.77	1/1097 (0.1%)
2	E	0.65	0/828	0.77	0/1109
3	B	0.61	0/634	0.71	1/848 (0.1%)
3	F	0.77	0/711	0.79	1/948 (0.1%)
4	C	0.57	0/829	0.70	0/1118
4	G	0.50	0/854	0.66	0/1150
5	D	0.62	0/813	0.83	2/1087 (0.2%)
5	H	0.56	0/762	0.64	0/1023
All	All	0.81	11/12958 (0.1%)	0.96	24/18730 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	277	DG	C3'-O3'	-6.84	1.35	1.44
1	J	174	DA	C3'-O3'	-6.49	1.35	1.44
1	J	173	DA	C3'-O3'	-5.94	1.36	1.44
1	J	248	DA	C3'-O3'	-5.80	1.36	1.44
1	J	213	DA	C3'-O3'	-5.70	1.36	1.44

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	271	DG	O5'-P-OP1	-17.57	89.62	110.70
5	D	27	ARG	CB-CA-C	-9.83	90.73	110.40
1	J	191	DT	O4'-C1'-N1	7.43	113.20	108.00
3	B	102	GLY	N-CA-C	7.12	130.91	113.10
1	J	171	DC	O5'-P-OP2	-6.90	99.49	105.70

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	I	2990	0	1652	88	0
1	J	2990	0	1652	98	0
2	A	807	0	844	24	0
2	E	816	0	856	21	0
3	B	627	0	663	9	0
3	F	703	0	755	10	0
4	C	819	0	879	14	0
4	G	844	0	910	18	0
5	D	797	0	833	25	0
5	H	746	0	768	24	0
6	D	1	0	0	1	0
6	G	1	0	0	0	0
6	I	11	0	0	0	0
6	J	6	0	0	0	0
7	B	1	0	0	0	0
7	D	1	0	0	2	0
7	I	2	0	0	0	0
7	J	2	0	0	1	0
8	A	3	0	0	0	0
8	C	2	0	0	0	0
8	D	2	0	0	0	0
8	E	4	0	0	0	0
8	F	4	0	0	0	0
8	G	1	0	0	0	0
8	I	12	0	0	3	0
8	J	5	0	0	0	0
All	All	12197	0	9812	271	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 13.

The worst 5 of 271 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:I:122:DG:H5'	5:H:33:ARG:NH2	1.35	1.36
1:I:122:DG:C5'	5:H:33:ARG:NH2	2.01	1.23
5:D:29:ARG:NH2	5:D:31:ARG:NH2	2.05	1.04
1:I:122:DG:C5'	5:H:33:ARG:HH21	1.75	0.96
6:D:201:MN:MN	7:D:202:CL:CL	1.64	0.93

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	
2	A	96/135 (71%)	96 (100%)	0	0	100	100
2	E	97/135 (72%)	97 (100%)	0	0	100	100
3	B	77/102 (76%)	75 (97%)	2 (3%)	0	100	100
3	F	85/102 (83%)	84 (99%)	1 (1%)	0	100	100
4	C	104/129 (81%)	102 (98%)	2 (2%)	0	100	100
4	G	107/129 (83%)	106 (99%)	1 (1%)	0	100	100
5	D	98/126 (78%)	97 (99%)	1 (1%)	0	100	100
5	H	93/126 (74%)	91 (98%)	2 (2%)	0	100	100
All	All	757/984 (77%)	748 (99%)	9 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	
2	A	85/110 (77%)	85 (100%)	0	100	100
2	E	86/110 (78%)	84 (98%)	2 (2%)	50	79
3	B	64/78 (82%)	63 (98%)	1 (2%)	62	85
3	F	72/78 (92%)	72 (100%)	0	100	100
4	C	84/99 (85%)	83 (99%)	1 (1%)	71	90
4	G	86/99 (87%)	86 (100%)	0	100	100
5	D	89/109 (82%)	89 (100%)	0	100	100
5	H	84/109 (77%)	82 (98%)	2 (2%)	49	78
All	All	650/792 (82%)	644 (99%)	6 (1%)	78	92

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	115	LYS
5	H	56	SER
2	E	129	ARG
4	C	114	VAL
5	H	48	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
4	C	38	ASN
3	F	18	HIS

### 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 25 ligands modelled in this entry, 25 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	I	146/146 (100%)	0.27	10 (6%) 17 14	20, 61, 103, 115	0
1	J	146/146 (100%)	0.38	10 (6%) 17 14	26, 60, 102, 125	0
2	A	98/135 (72%)	-0.18	3 (3%) 49 45	13, 32, 69, 85	0
2	E	99/135 (73%)	-0.40	2 (2%) 65 64	8, 16, 36, 57	0
3	B	79/102 (77%)	-0.33	0 100 100	20, 30, 55, 74	0
3	F	87/102 (85%)	-0.32	4 (4%) 32 29	8, 14, 37, 84	0
4	C	106/129 (82%)	-0.43	0 100 100	9, 21, 42, 60	0
4	G	109/129 (84%)	-0.25	4 (3%) 41 38	12, 28, 70, 82	0
5	D	99/126 (78%)	-0.17	3 (3%) 50 46	11, 24, 77, 91	0
5	H	94/126 (74%)	-0.25	2 (2%) 63 62	10, 25, 58, 100	1 (1%)
All	All	1063/1276 (83%)	-0.12	38 (3%) 42 39	8, 30, 85, 125	1 (0%)

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	38	PRO	5.1
4	G	13	LYS	5.0
1	J	242	DT	4.3
1	J	148	DT	4.2
1	J	147	DA	4.1

### 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(Å <sup>2</sup> )	Q<0.9
6	MN	J	301	1/1	0.21	0.75	120,120,120,120	0
6	MN	I	203	1/1	0.77	0.09	113,113,113,113	0
7	CL	I	213	1/1	0.77	0.17	82,82,82,82	0
6	MN	I	210	1/1	0.80	0.23	83,83,83,83	0
6	MN	I	211	1/1	0.86	0.38	137,137,137,137	0
7	CL	B	201	1/1	0.86	0.15	54,54,54,54	0
6	MN	I	206	1/1	0.87	0.09	106,106,106,106	0
6	MN	I	205	1/1	0.88	0.09	81,81,81,81	0
6	MN	I	208	1/1	0.91	0.10	63,63,63,63	0
6	MN	I	204	1/1	0.91	0.06	55,55,55,55	0
6	MN	I	209	1/1	0.91	0.19	98,98,98,98	0
6	MN	J	302	1/1	0.91	0.16	65,65,65,65	0
7	CL	J	307	1/1	0.93	0.16	50,50,50,50	0
7	CL	I	212	1/1	0.93	0.24	64,64,64,64	0
7	CL	J	308	1/1	0.94	0.10	37,37,37,37	0
6	MN	I	201	1/1	0.95	0.07	66,66,66,66	0
6	MN	G	201	1/1	0.95	0.07	43,43,43,43	0
6	MN	J	305	1/1	0.96	0.05	68,68,68,68	0
6	MN	J	306	1/1	0.96	0.12	77,77,77,77	0
6	MN	J	304	1/1	0.96	0.08	49,49,49,49	0
6	MN	I	207	1/1	0.96	0.09	44,44,44,44	0
6	MN	J	303	1/1	0.97	0.16	68,68,68,68	0
6	MN	I	202	1/1	0.98	0.20	74,74,74,74	0
7	CL	D	202	1/1	0.99	0.15	33,33,33,33	0
6	MN	D	201	1/1	0.99	0.16	29,29,29,29	0

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

There are no such residues in this entry.