



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:01 PM GMT

PDB ID : 4NFT
Title : Crystal structure of human lnkH2B-h2A.Z-Anp32e
Authors : Shan, S.; Pan, L.; Mao, Z.; Wang, W.; Sun, J.; Dong, Q.; Liang, X.; Ding, X.;
Chen, S.; Dai, L.; Zhang, Z.; Zhu, B.; Zhou, Z.
Deposited on : 2013-11-01
Resolution : 2.61 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
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

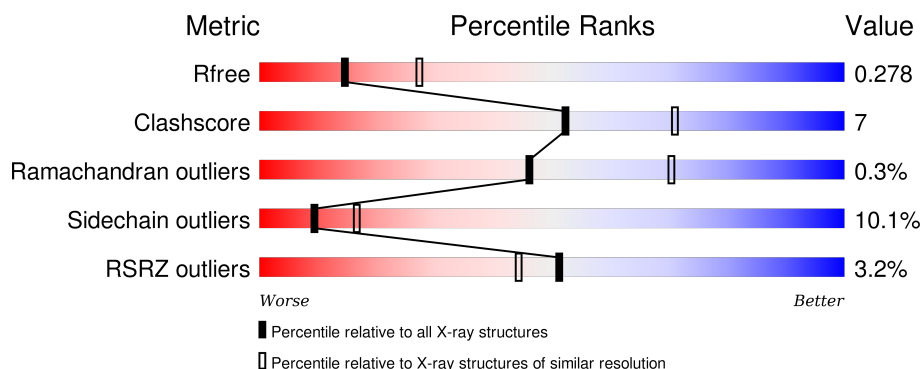
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.61 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	2700 (2.64-2.60)
Clashscore	102246	3065 (2.64-2.60)
Ramachandran outliers	100387	3015 (2.64-2.60)
Sidechain outliers	100360	3015 (2.64-2.60)
RSRZ outliers	91569	2706 (2.64-2.60)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	209	
1	B	209	
1	C	209	
1	D	209	
2	E	52	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	52	 A horizontal bar chart showing the quality of chain F. The bar is divided into three segments: green (12%), yellow (8%), and orange (79%). The green segment is the largest, followed by the orange segment, and the yellow segment is the smallest. The percentages are labeled below the segments: 12%, 8%, and 79%.

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5932 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 H2B type 2-E, Histone H2A.Z.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	184	Total	C	N	O	S	0	1	0
			1433	898	269	263	3			
1	B	184	Total	C	N	O	S	0	1	0
			1433	898	269	263	3			
1	C	178	Total	C	N	O	S	0	0	0
			1371	859	255	255	2			
1	D	178	Total	C	N	O	S	0	0	0
			1371	859	255	255	2			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP Q16778
A	-1	SER	-	EXPRESSION TAG	UNP Q16778
A	0	HIS	-	EXPRESSION TAG	UNP Q16778
A	1	MET	-	EXPRESSION TAG	UNP Q16778
A	194	SER	-	EXPRESSION TAG	UNP P0C0S5
A	195	GLY	-	EXPRESSION TAG	UNP P0C0S5
A	196	SER	-	EXPRESSION TAG	UNP P0C0S5
A	197	GLY	-	EXPRESSION TAG	UNP P0C0S5
A	198	GLY	-	EXPRESSION TAG	UNP P0C0S5
A	199	SER	-	EXPRESSION TAG	UNP P0C0S5
A	200	GLY	-	EXPRESSION TAG	UNP P0C0S5
A	201	GLY	-	EXPRESSION TAG	UNP P0C0S5
A	202	GLY	-	EXPRESSION TAG	UNP P0C0S5
A	203	LEU	-	EXPRESSION TAG	UNP P0C0S5
A	204	VAL	-	EXPRESSION TAG	UNP P0C0S5
A	205	PRO	-	EXPRESSION TAG	UNP P0C0S5
A	206	ARG	-	EXPRESSION TAG	UNP P0C0S5
B	-2	GLY	-	EXPRESSION TAG	UNP Q16778
B	-1	SER	-	EXPRESSION TAG	UNP Q16778
B	0	HIS	-	EXPRESSION TAG	UNP Q16778
B	1	MET	-	EXPRESSION TAG	UNP Q16778

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	194	SER	-	EXPRESSION TAG	UNP P0C0S5
B	195	GLY	-	EXPRESSION TAG	UNP P0C0S5
B	196	SER	-	EXPRESSION TAG	UNP P0C0S5
B	197	GLY	-	EXPRESSION TAG	UNP P0C0S5
B	198	GLY	-	EXPRESSION TAG	UNP P0C0S5
B	199	SER	-	EXPRESSION TAG	UNP P0C0S5
B	200	GLY	-	EXPRESSION TAG	UNP P0C0S5
B	201	GLY	-	EXPRESSION TAG	UNP P0C0S5
B	202	GLY	-	EXPRESSION TAG	UNP P0C0S5
B	203	LEU	-	EXPRESSION TAG	UNP P0C0S5
B	204	VAL	-	EXPRESSION TAG	UNP P0C0S5
B	205	PRO	-	EXPRESSION TAG	UNP P0C0S5
B	206	ARG	-	EXPRESSION TAG	UNP P0C0S5
C	-2	GLY	-	EXPRESSION TAG	UNP Q16778
C	-1	SER	-	EXPRESSION TAG	UNP Q16778
C	0	HIS	-	EXPRESSION TAG	UNP Q16778
C	1	MET	-	EXPRESSION TAG	UNP Q16778
C	194	SER	-	EXPRESSION TAG	UNP P0C0S5
C	195	GLY	-	EXPRESSION TAG	UNP P0C0S5
C	196	SER	-	EXPRESSION TAG	UNP P0C0S5
C	197	GLY	-	EXPRESSION TAG	UNP P0C0S5
C	198	GLY	-	EXPRESSION TAG	UNP P0C0S5
C	199	SER	-	EXPRESSION TAG	UNP P0C0S5
C	200	GLY	-	EXPRESSION TAG	UNP P0C0S5
C	201	GLY	-	EXPRESSION TAG	UNP P0C0S5
C	202	GLY	-	EXPRESSION TAG	UNP P0C0S5
C	203	LEU	-	EXPRESSION TAG	UNP P0C0S5
C	204	VAL	-	EXPRESSION TAG	UNP P0C0S5
C	205	PRO	-	EXPRESSION TAG	UNP P0C0S5
C	206	ARG	-	EXPRESSION TAG	UNP P0C0S5
D	-2	GLY	-	EXPRESSION TAG	UNP Q16778
D	-1	SER	-	EXPRESSION TAG	UNP Q16778
D	0	HIS	-	EXPRESSION TAG	UNP Q16778
D	1	MET	-	EXPRESSION TAG	UNP Q16778
D	194	SER	-	EXPRESSION TAG	UNP P0C0S5
D	195	GLY	-	EXPRESSION TAG	UNP P0C0S5
D	196	SER	-	EXPRESSION TAG	UNP P0C0S5
D	197	GLY	-	EXPRESSION TAG	UNP P0C0S5
D	198	GLY	-	EXPRESSION TAG	UNP P0C0S5
D	199	SER	-	EXPRESSION TAG	UNP P0C0S5
D	200	GLY	-	EXPRESSION TAG	UNP P0C0S5
D	201	GLY	-	EXPRESSION TAG	UNP P0C0S5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	202	GLY	-	EXPRESSION TAG	UNP P0C0S5
D	203	LEU	-	EXPRESSION TAG	UNP P0C0S5
D	204	VAL	-	EXPRESSION TAG	UNP P0C0S5
D	205	PRO	-	EXPRESSION TAG	UNP P0C0S5
D	206	ARG	-	EXPRESSION TAG	UNP P0C0S5

- Molecule 2 is a protein called Acidic leucine-rich nuclear phosphoprotein 32 family member E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	11	Total 89	C 57	N 12	O 19	S 1	0	0	0
2	F	11	Total 89	C 57	N 12	O 19	S 1	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	-27	GLY	-	EXPRESSION TAG	UNP Q9BTT0
E	-26	SER	-	EXPRESSION TAG	UNP Q9BTT0
E	-25	HIS	-	EXPRESSION TAG	UNP Q9BTT0
E	-24	MET	-	EXPRESSION TAG	UNP Q9BTT0
F	-27	GLY	-	EXPRESSION TAG	UNP Q9BTT0
F	-26	SER	-	EXPRESSION TAG	UNP Q9BTT0
F	-25	HIS	-	EXPRESSION TAG	UNP Q9BTT0
F	-24	MET	-	EXPRESSION TAG	UNP Q9BTT0

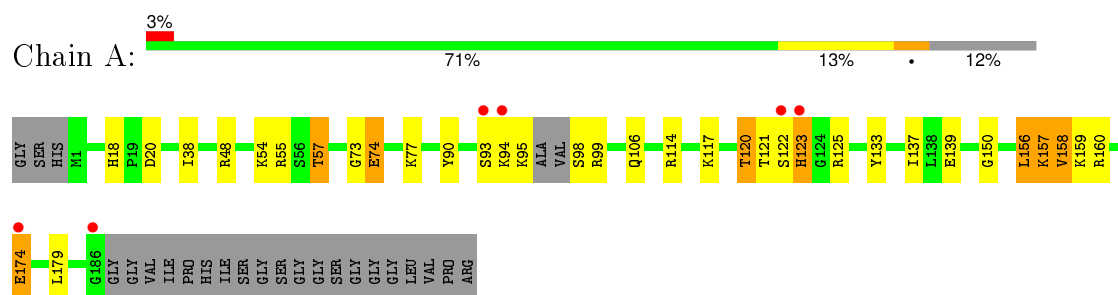
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	38	Total	O	0	0
			38	38		
3	B	36	Total	O	0	0
			36	36		
3	C	32	Total	O	0	0
			32	32		
3	D	34	Total	O	0	0
			34	34		
3	E	3	Total	O	0	0
			3	3		
3	F	3	Total	O	0	0
			3	3		

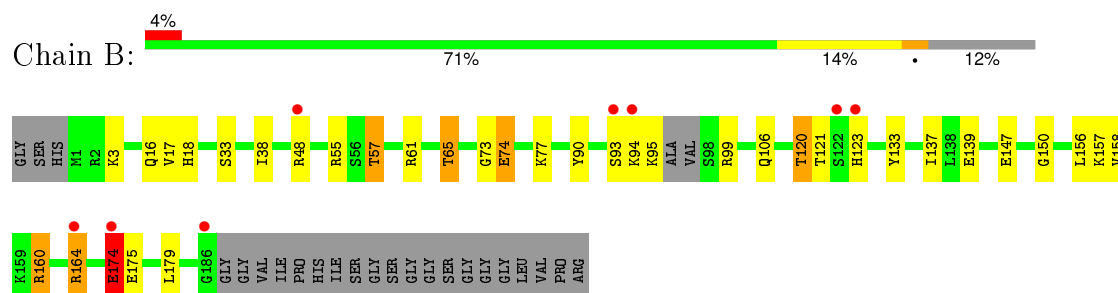
3 Residue-property plots [i](#)

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 ($\text{RSRZ} > 2$). 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.

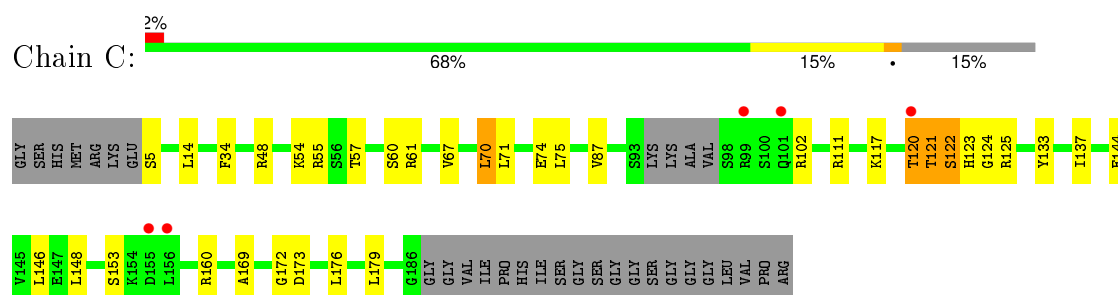
- Molecule 1: Histone H2B type 2-E, Histone H2A.Z



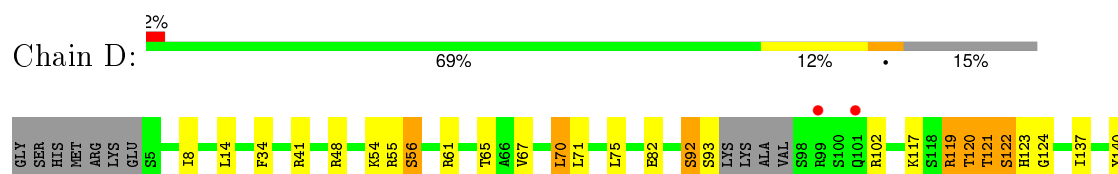
- Molecule 1: Histone H2B type 2-E, Histone H2A.Z

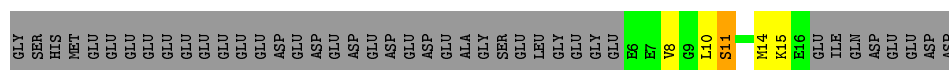


- Molecule 1: Histone H2B type 2-E, Histone H2A.Z



- Molecule 1: Histone H2B type 2-E, Histone H2A.Z





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	76.41Å 104.30Å 124.74Å 90.00° 99.15° 90.00°	Depositor
Resolution (Å)	39.80 – 2.61 48.02 – 2.61	Depositor EDS
% Data completeness (in resolution range)	98.0 (39.80-2.61) 99.1 (48.02-2.61)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.72 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.220 , 0.273 0.229 , 0.278	Depositor DCC
R_{free} test set	1493 reflections (5.39%)	DCC
Wilson B-factor (Å ²)	31.8	Xtriage
Anisotropy	0.657	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 49.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.27$	Xtriage
Outliers	2 of 29171 reflections (0.007%)	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5932	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 95.87 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 9.4087e-10. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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	0.58	0/1456	0.62	1/1956 (0.1%)
1	B	0.59	1/1456 (0.1%)	0.62	0/1956
1	C	0.52	0/1390	0.60	0/1872
1	D	0.54	0/1390	0.61	1/1872 (0.1%)
2	E	0.79	0/89	0.74	0/117
2	F	0.62	0/89	0.68	0/117
All	All	0.56	1/5870 (0.0%)	0.61	2/7890 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	174	GLU	CG-CD	5.13	1.59	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	92	SER	CB-CA-C	-5.23	100.16	110.10
1	A	123	HIS	N-CA-C	-5.07	97.30	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	1433	0	1497	26	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1433	0	1497	20	0
1	C	1371	0	1420	19	0
1	D	1371	0	1420	15	0
2	E	89	0	87	4	0
2	F	89	0	87	2	0
3	A	38	0	0	1	0
3	B	36	0	0	0	0
3	C	32	0	0	1	0
3	D	34	0	0	0	0
3	E	3	0	0	0	0
3	F	3	0	0	0	0
All	All	5932	0	6008	79	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 79 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:B:174:GLU:OE1	1:B:175:GLU:HG2	1.09	1.26
1:B:94:LYS:O	1:B:95:LYS:HB3	1.50	1.06
1:B:174:GLU:OE1	1:B:175:GLU:CG	2.05	1.04
1:A:121:THR:O	1:A:123:HIS:O	1.80	0.99
1:A:94:LYS:O	1:A:95:LYS:HG3	1.71	0.89

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	181/209 (87%)	178 (98%)	2 (1%)	1 (1%)	30 54

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	181/209 (87%)	177 (98%)	3 (2%)	1 (1%)	30	54
1	C	174/209 (83%)	168 (97%)	6 (3%)	0	100	100
1	D	174/209 (83%)	169 (97%)	5 (3%)	0	100	100
2	E	9/52 (17%)	9 (100%)	0	0	100	100
2	F	9/52 (17%)	9 (100%)	0	0	100	100
All	All	728/940 (77%)	710 (98%)	16 (2%)	2 (0%)	46	70

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	73	GLY
1	B	73	GLY

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	153/167 (92%)	142 (93%)	11 (7%)	18	34
1	B	153/167 (92%)	140 (92%)	13 (8%)	13	25
1	C	146/167 (87%)	132 (90%)	14 (10%)	10	19
1	D	146/167 (87%)	128 (88%)	18 (12%)	6	10
2	E	10/46 (22%)	7 (70%)	3 (30%)	0	0
2	F	10/46 (22%)	7 (70%)	3 (30%)	0	0
All	All	618/760 (81%)	556 (90%)	62 (10%)	9	17

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	74	GLU
1	C	123	HIS
2	E	11	SER
1	C	120	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	160	ARG

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

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, 95th 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 ²)	Q<0.9
1	A	184/209 (88%)	-0.21	6 (3%) 50 43	18, 29, 51, 76	0
1	B	184/209 (88%)	-0.20	8 (4%) 39 32	18, 29, 51, 76	0
1	C	178/209 (85%)	-0.24	5 (2%) 56 50	16, 28, 48, 64	0
1	D	178/209 (85%)	-0.23	4 (2%) 65 59	16, 28, 48, 64	0
2	E	11/52 (21%)	-0.21	1 (9%) 11 7	30, 38, 57, 60	0
2	F	11/52 (21%)	-0.26	0 100 100	29, 39, 58, 60	0
All	All	746/940 (79%)	-0.22	24 (3%) 51 44	16, 29, 51, 76	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	186	GLY	4.9
1	A	123	HIS	4.4
1	A	186	GLY	3.8
1	B	123	HIS	3.3
1	B	94	LYS	2.9

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 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.