



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 12, 2021 – 04:36 PM EDT

PDB ID : 1ZLA  
Title : X-ray Structure of a Kaposi's sarcoma herpesvirus LANA peptide bound to the nucleosomal core  
Authors : Chodaparambil, J.V.; Barbera, A.J.; Kaye, K.M.; Luger, K.  
Deposited on : 2005-05-05  
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

<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	:	2.23.2
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)	:	2.23.2

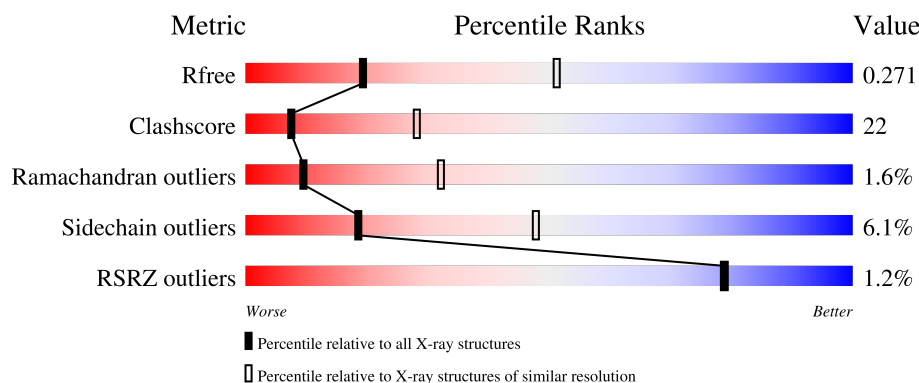
# 1 Overall quality at a glance

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))
$R_{free}$	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (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 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	I	146	<div> <div>43%</div> <div>56%</div> <div>.</div> </div>
1	J	146	<div> <div>40%</div> <div>59%</div> <div>.</div> </div>
2	A	135	<div> <div>42%</div> <div>27%</div> <div>.</div> <div>27%</div> </div>
2	E	135	<div> <div>2%</div> <div>52%</div> <div>18%</div> <div>.</div> <div>27%</div> </div>
3	B	102	<div> <div>50%</div> <div>27%</div> <div>23%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	102	<div><div></div><div>58%22%•18%</div></div>
4	C	129	<div><div>2%</div><div></div><div>52%26%5%17%</div></div>
4	G	129	<div><div>2%</div><div></div><div>43%33%5%•18%</div></div>
5	D	125	<div><div>%</div><div></div><div>45%30%26%</div></div>
5	H	125	<div><div></div><div>42%30%••26%</div></div>
6	K	22	<div><div>5%</div><div></div><div>23%5%32%5%36%</div></div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12168 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 Palindromic 146bp Human alpha-Satellite DNA fragment.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	146	Total	C	N	O	P	0	0	0
			2990	1430	541	874	145			
1	J	146	Total	C	N	O	P	0	0	0
			2990	1430	541	874	145			

- Molecule 2 is a protein called histone H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	98	Total	C	N	O	S	0	0	0
			811	511	158	139	3			
2	E	98	Total	C	N	O	S	0	0	0
			811	511	158	139	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	511	ALA	GLY	engineered mutation	GB 288992
A	518	HIS	THR	engineered mutation	GB 288992
E	711	ALA	GLY	engineered mutation	GB 288992
E	718	HIS	THR	engineered mutation	GB 288992

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	79	Total	C	N	O	S	0	0	0
			627	395	121	110	1			
3	F	84	Total	C	N	O	S	0	0	0
			673	424	133	115	1			

- Molecule 4 is a protein called Xenopus laevis-like histone H2A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	C	107	Total	C	N	O	0	0	0
			825	520	161	144			
4	G	106	Total	C	N	O	0	0	0
			818	516	160	142			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	869	TRP	ALA	engineered mutation	GB 30268540
C	870	GLU	ALA	engineered mutation	GB 30268540
C	926	THR	ALA	engineered mutation	GB 30268540
G	1069	TRP	ALA	engineered mutation	GB 30268540
G	1070	GLU	ALA	engineered mutation	GB 30268540
G	1126	THR	ALA	engineered mutation	GB 30268540

- Molecule 5 is a protein called histone H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	93	Total	C	N	O	S	0	0	0
			729	459	131	137	2			
5	H	93	Total	C	N	O	S	0	0	0
			729	459	131	137	2			

- Molecule 6 is a protein called latent nuclear antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	K	14	Total	C	N	O	S	0	0	0
			99	58	23	17	1			

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	I	16	Total	O	0	0
			16	16		
7	J	11	Total	O	0	0
			11	11		
7	A	2	Total	O	0	0
			2	2		
7	B	1	Total	O	0	0
			1	1		
7	C	6	Total	O	0	0
			6	6		

*Continued on next page...*

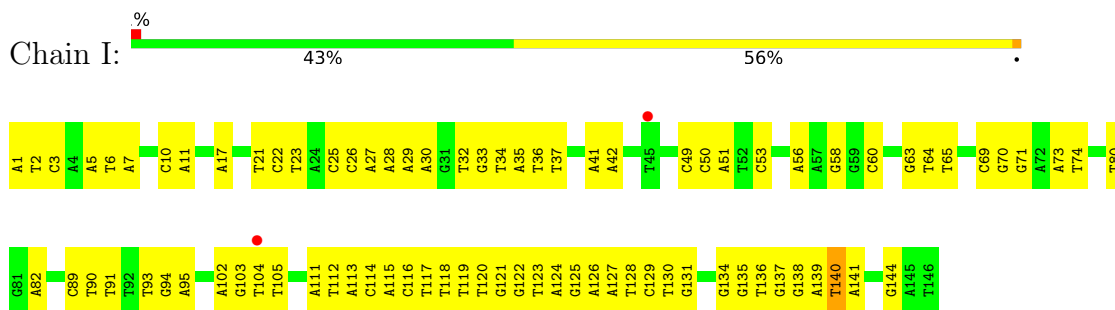
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	8	Total 8	O 8	0	0
7	E	4	Total 4	O 4	0	0
7	F	6	Total 6	O 6	0	0
7	G	7	Total 7	O 7	0	0
7	H	5	Total 5	O 5	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.

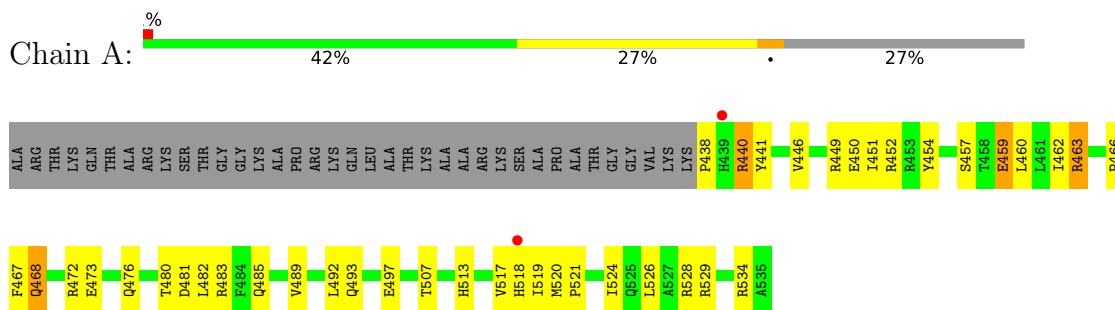
- Molecule 1: Palindromic 146bp Human alpha-Satellite DNA fragment



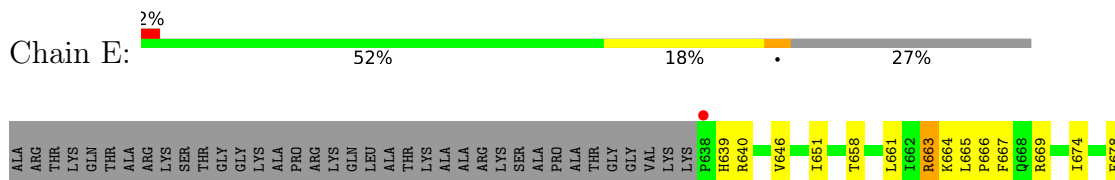
- Molecule 1: Palindromic 146bp Human alpha-Satellite DNA fragment



- Molecule 2: histone H3

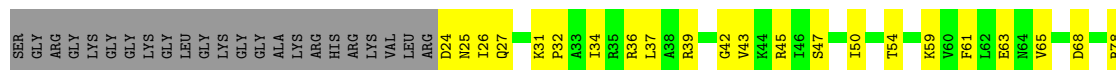


- Molecule 2: histone H3

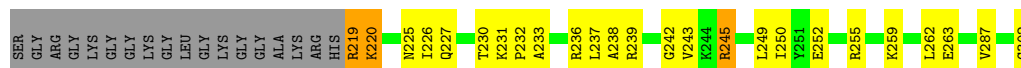




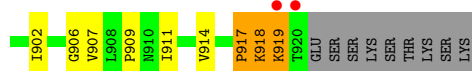
• Molecule 3: Histone H4



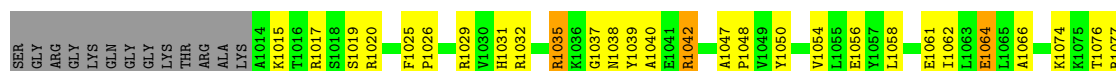
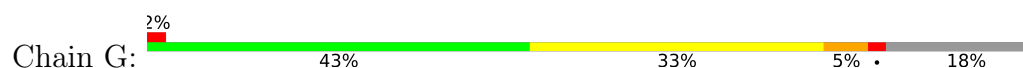
• Molecule 3: Histone H4



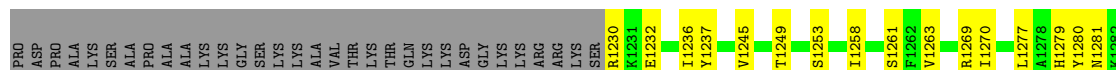
• Molecule 4: *Xenopus laevis*-like histone H2A



• Molecule 4: *Xenopus laevis*-like histone H2A



• Molecule 5: histone H2B



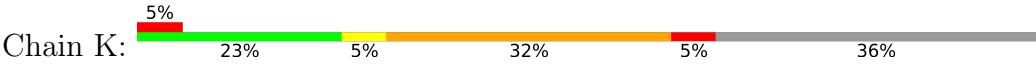




● Molecule 5: histone H2B



● Molecule 6: latent nuclear antigen



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.16Å 109.60Å 182.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.90 42.31 – 2.88	Depositor EDS
% Data completeness (in resolution range)	90.8 (50.00-2.90) 90.3 (42.31-2.88)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.18 (at 2.90Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.222 , 0.277 0.212 , 0.271	Depositor DCC
$R_{free}$ test set	2208 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.5	Xtriage
Anisotropy	0.264	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 60.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.023 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12168	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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

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	0.35	0/3354	0.69	0/5175
1	J	0.39	1/3354 (0.0%)	0.71	0/5175
2	A	0.47	0/824	0.81	1/1104 (0.1%)
2	E	0.74	1/824 (0.1%)	0.73	0/1104
3	B	0.41	0/634	0.67	0/848
3	F	0.58	0/680	1.07	5/908 (0.6%)
4	C	0.46	0/835	0.86	1/1127 (0.1%)
4	G	0.40	0/828	0.83	1/1117 (0.1%)
5	D	0.45	0/740	0.65	0/994
5	H	0.68	1/740 (0.1%)	0.70	0/994
6	K	0.56	0/100	1.69	1/131 (0.8%)
All	All	0.46	3/12913 (0.0%)	0.76	9/18677 (0.0%)

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
1	I	0	1
1	J	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	735	ALA	C-OXT	16.60	1.54	1.23
5	H	1522	LYS	C-OXT	14.23	1.50	1.23
1	J	199	DC	P-OP1	-8.92	1.33	1.49

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	917	PRO	CA-N-CD	-15.82	89.36	111.50
3	F	219	ARG	CA-C-N	-15.64	82.78	117.20
2	A	438	PRO	CA-N-CD	-13.27	92.92	111.50
6	K	17	PRO	CA-N-CD	-13.15	93.09	111.50
3	F	219	ARG	O-C-N	11.37	140.89	122.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	140	DT	Sidechain
1	J	147	DA	Sidechain

## 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	1651	99	0
1	J	2990	0	1651	107	0
2	A	811	0	846	62	0
2	E	811	0	846	48	0
3	B	627	0	663	40	0
3	F	673	0	722	38	1
4	C	825	0	878	42	0
4	G	818	0	871	65	0
5	D	729	0	753	40	1
5	H	729	0	753	41	0
6	K	99	0	104	29	0
7	A	2	0	0	5	0
7	B	1	0	0	0	0
7	C	6	0	0	1	0
7	D	8	0	0	2	0
7	E	4	0	0	1	0
7	F	6	0	0	2	0
7	G	7	0	0	0	0
7	H	5	0	0	0	0
7	I	16	0	0	2	0
7	J	11	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	12168	0	9738	482	1

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 22.

The worst 5 of 482 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:G:1116:LEU:HB3	4:G:1117:PRO:HD2	1.26	1.15
1:J:261:DA:H2''	1:J:262:DC:H5''	1.34	1.08
1:J:174:DA:H2''	1:J:175:DA:H5''	1.34	1.06
6:K:9:ARG:HG3	6:K:9:ARG:HH11	1.22	1.04
5:H:1487:THR:O	5:H:1489:ARG:N	1.90	1.03

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:D:1302:GLU:OE2	3:F:219:ARG:NH2[3_544]	1.70	0.50

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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%)	90 (94%)	5 (5%)	1 (1%)	15	45
2	E	96/135 (71%)	91 (95%)	5 (5%)	0	100	100
3	B	77/102 (76%)	74 (96%)	3 (4%)	0	100	100
3	F	82/102 (80%)	78 (95%)	3 (4%)	1 (1%)	13	40
4	C	105/129 (81%)	95 (90%)	10 (10%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	G	104/129 (81%)	95 (91%)	6 (6%)	3 (3%)	4	18
5	D	91/125 (73%)	82 (90%)	8 (9%)	1 (1%)	14	42
5	H	91/125 (73%)	81 (89%)	6 (7%)	4 (4%)	2	10
6	K	12/22 (54%)	9 (75%)	1 (8%)	2 (17%)	0	0
All	All	754/1004 (75%)	695 (92%)	47 (6%)	12 (2%)	9	32

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	220	LYS
4	G	1117	PRO
5	H	1488	SER
5	H	1501	GLY
5	D	1301	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	
2	A	85/110 (77%)	81 (95%)	4 (5%)	26	59
2	E	85/110 (77%)	82 (96%)	3 (4%)	36	70
3	B	64/78 (82%)	64 (100%)	0	100	100
3	F	69/78 (88%)	67 (97%)	2 (3%)	42	76
4	C	85/104 (82%)	77 (91%)	8 (9%)	8	26
4	G	84/104 (81%)	73 (87%)	11 (13%)	4	12
5	D	79/104 (76%)	76 (96%)	3 (4%)	33	67
5	H	79/104 (76%)	77 (98%)	2 (2%)	47	78
6	K	10/16 (62%)	4 (40%)	6 (60%)	0	0
All	All	640/808 (79%)	601 (94%)	39 (6%)	18	48

5 of 39 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	1117	PRO
6	K	9	ARG
4	G	1118	LYS
5	H	1487	THR
6	K	14	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
4	G	1031	HIS
4	G	1112	GLN
5	H	1492	GLN
4	C	838	ASN
5	D	1292	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 monosaccharides 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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%)	-0.07	2 (1%) 75 75	43, 83, 142, 166	0
1	J	146/146 (100%)	-0.14	0 100 100	43, 86, 141, 181	0
2	A	98/135 (72%)	-0.30	2 (2%) 65 63	18, 35, 72, 113	0
2	E	98/135 (72%)	-0.34	3 (3%) 49 44	11, 30, 90, 161	0
3	B	79/102 (77%)	-0.47	0 100 100	12, 33, 60, 106	0
3	F	84/102 (82%)	-0.52	0 100 100	11, 28, 49, 121	0
4	C	107/129 (82%)	-0.32	2 (1%) 66 65	14, 31, 68, 194	0
4	G	106/129 (82%)	-0.37	2 (1%) 66 65	9, 38, 91, 178	0
5	D	93/125 (74%)	-0.35	1 (1%) 80 80	16, 34, 60, 127	0
5	H	93/125 (74%)	-0.35	0 100 100	17, 40, 80, 137	0
6	K	14/22 (63%)	0.61	1 (7%) 16 12	29, 58, 158, 177	0
All	All	1064/1296 (82%)	-0.29	13 (1%) 79 79	9, 41, 118, 194	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	C	920	THR	8.4
6	K	17	PRO	7.5
4	G	1119	LYS	6.3
4	C	919	LYS	5.3
4	G	1118	LYS	3.2

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.