



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 01:25 PM GMT

PDB ID : 3TU4  
Title : Crystal structure of the Sir3 BAH domain in complex with a nucleosome core particle.  
Authors : Armache, K.-J.; Garlick, J.D.; Canzio, D.; Narlikar, G.J.; Kingston, R.E.  
Deposited on : 2011-09-15  
Resolution : 3.00 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

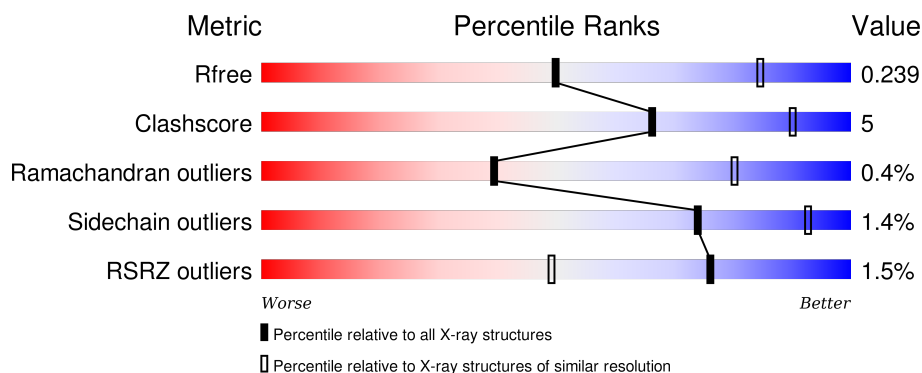
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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>69%</div> <div>27%</div> </div>
1	E	135	<div> <div>61%</div> <div>10%</div> <div>30%</div> </div>
2	B	102	<div> <div>80%</div> <div>7%</div> <div>13%</div> </div>
2	F	102	<div> <div>82%</div> <div>6%</div> <div>12%</div> </div>
3	C	129	<div> <div>75%</div> <div>8%</div> <div>17%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	129	 2% 76% 7% 17%
4	D	122	 66% 11% 24%
4	H	122	 65% 11% 24%
5	I	147	 76% 22% ..
6	J	147	 79% 18% ..
7	K	229	 4% 77% 14% • 7%
7	L	229	 3% 75% 16% • 7%

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 15660 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	CONFLICT	UNP P84233
E	102	ALA	GLY	CONFLICT	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	89	Total	C	N	O	S	0	0	0
			711	447	144	119	1			
2	F	90	Total	C	N	O	S	0	0	0
			716	449	145	121	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	CONFLICT	UNP P02281
H	29	THR	SER	CONFLICT	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	1414	539	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	1426	563	876	146			

- Molecule 7 is a protein called Regulatory protein SIR3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	K	214	Total	C	N	O	S	0	0	0
			1802	1159	304	337	2			
7	L	213	Total	C	N	O	S	0	0	0
			1795	1154	303	336	2			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	205	ASN	ASP	ENGINEERED MUTATION	UNP P06701
K	214	SER	-	EXPRESSION TAG	UNP P06701
K	215	VAL	-	EXPRESSION TAG	UNP P06701
K	216	ASP	-	EXPRESSION TAG	UNP P06701
K	217	LYS	-	EXPRESSION TAG	UNP P06701
K	218	LEU	-	EXPRESSION TAG	UNP P06701
K	219	ALA	-	EXPRESSION TAG	UNP P06701
K	220	ALA	-	EXPRESSION TAG	UNP P06701
K	221	ALA	-	EXPRESSION TAG	UNP P06701

*Continued on next page...*

*Continued from previous page...*

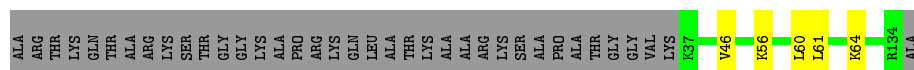
Chain	Residue	Modelled	Actual	Comment	Reference
K	222	LEU	-	EXPRESSION TAG	UNP P06701
K	223	GLU	-	EXPRESSION TAG	UNP P06701
K	224	HIS	-	EXPRESSION TAG	UNP P06701
K	225	HIS	-	EXPRESSION TAG	UNP P06701
K	226	HIS	-	EXPRESSION TAG	UNP P06701
K	227	HIS	-	EXPRESSION TAG	UNP P06701
K	228	HIS	-	EXPRESSION TAG	UNP P06701
K	229	HIS	-	EXPRESSION TAG	UNP P06701
L	205	ASN	ASP	ENGINEERED MUTATION	UNP P06701
L	214	SER	-	EXPRESSION TAG	UNP P06701
L	215	VAL	-	EXPRESSION TAG	UNP P06701
L	216	ASP	-	EXPRESSION TAG	UNP P06701
L	217	LYS	-	EXPRESSION TAG	UNP P06701
L	218	LEU	-	EXPRESSION TAG	UNP P06701
L	219	ALA	-	EXPRESSION TAG	UNP P06701
L	220	ALA	-	EXPRESSION TAG	UNP P06701
L	221	ALA	-	EXPRESSION TAG	UNP P06701
L	222	LEU	-	EXPRESSION TAG	UNP P06701
L	223	GLU	-	EXPRESSION TAG	UNP P06701
L	224	HIS	-	EXPRESSION TAG	UNP P06701
L	225	HIS	-	EXPRESSION TAG	UNP P06701
L	226	HIS	-	EXPRESSION TAG	UNP P06701
L	227	HIS	-	EXPRESSION TAG	UNP P06701
L	228	HIS	-	EXPRESSION TAG	UNP P06701
L	229	HIS	-	EXPRESSION TAG	UNP P06701

### 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.

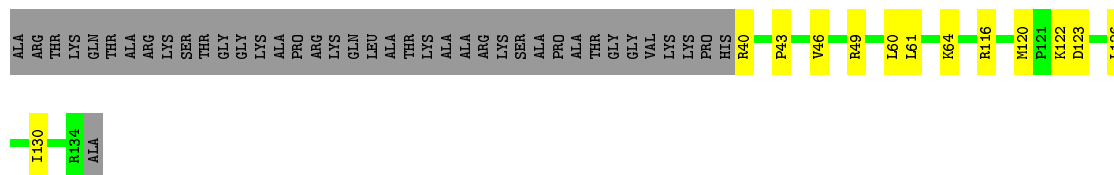
- Molecule 1: Histone H3.2

Chain A: 




- Molecule 1: Histone H3.2

Chain E: 




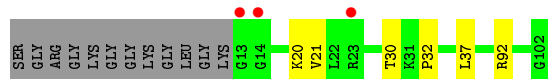
- Molecule 2: Histone H4

Chain B: 



- Molecule 2: Histone H4

Chain F: 

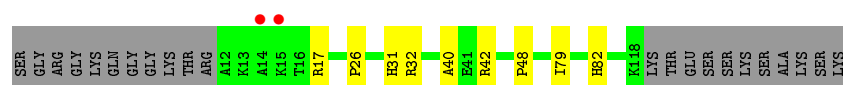
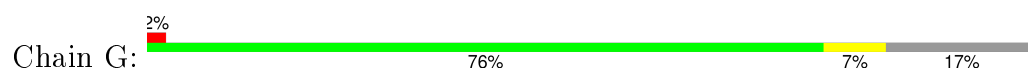


- Molecule 3: Histone H2A

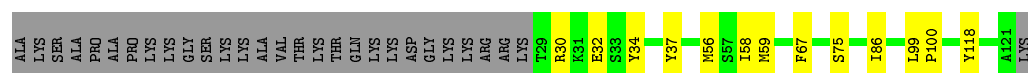
Chain C: 



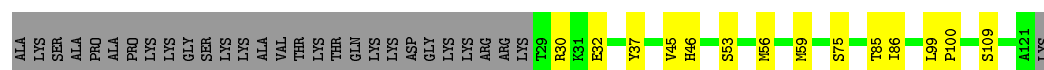
- Molecule 3: Histone H2A



- Molecule 4: Histone H2B 1.1



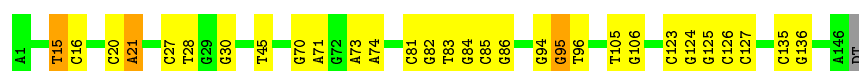
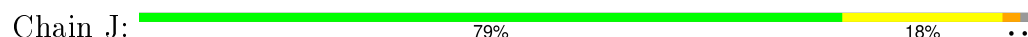
- Molecule 4: Histone H2B 1.1



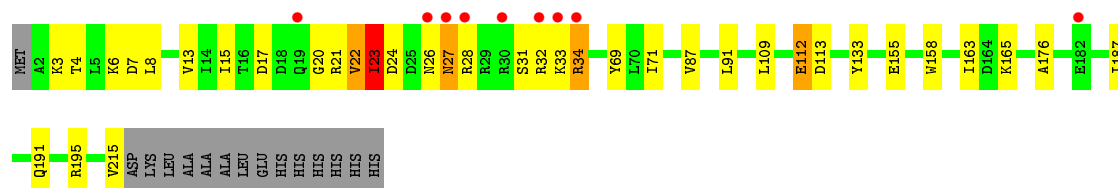
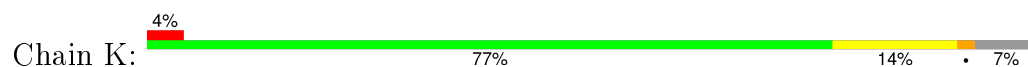
- Molecule 5: DNA (146-MER)



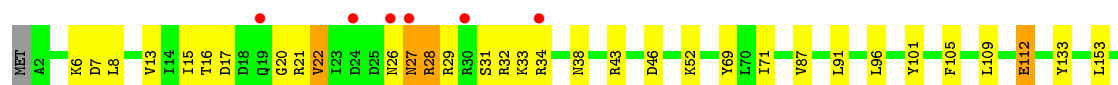
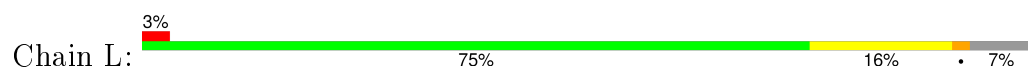
- Molecule 6: DNA (146-MER)



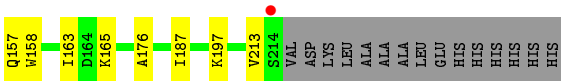
- Molecule 7: Regulatory protein SIR3



- Molecule 7: Regulatory protein SIR3







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.88Å 102.88Å 555.75Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.55 – 3.00 49.56 – 3.00	Depositor EDS
% Data completeness (in resolution range)	97.5 (44.55-3.00) 97.5 (49.56-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.25 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.192 , 0.241 0.188 , 0.239	Depositor DCC
$R_{free}$ test set	1765 reflections (2.74%)	DCC
Wilson B-factor (Å <sup>2</sup> )	78.8	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 67.3	EDS
Estimated twinning fraction	0.469 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 66104 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15660	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.14% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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	0.39	0/1093
1	E	0.22	0/789	0.39	0/1059
2	B	0.22	0/719	0.42	0/960
2	F	0.22	0/724	0.44	0/965
3	C	0.22	0/821	0.42	0/1112
3	G	0.21	0/825	0.41	0/1116
4	D	0.23	0/729	0.41	0/985
4	H	0.23	0/737	0.42	0/993
5	I	0.46	0/3333	1.11	5/5137 (0.1%)
6	J	0.46	0/3381	1.10	4/5221 (0.1%)
7	K	0.24	0/1842	0.48	1/2488 (0.0%)
7	L	0.23	0/1835	0.46	0/2478
All	All	0.34	0/16548	0.80	10/23607 (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
7	K	0	3
7	L	0	2
All	All	0	5

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	95	DC	O4'-C4'-C3'	-6.49	101.91	104.50
6	J	95	DG	O4'-C4'-C3'	-6.26	102.00	104.50
6	J	21	DA	C1'-O4'-C4'	-5.71	104.39	110.10
5	I	74	DT	C1'-O4'-C4'	-5.55	104.55	110.10
5	I	52	DA	C3'-C2'-C1'	-5.37	96.05	102.50

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	K	112	GLU	Peptide
7	K	23	ILE	Peptide
7	K	34	ARG	Peptide
7	L	112	GLU	Peptide
7	L	34	ARG	Peptide

## 5.2 Too-close contacts ⓘ

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	4	0
1	E	779	0	815	9	0
2	B	711	0	763	4	0
2	F	716	0	766	3	0
3	C	811	0	849	11	0
3	G	815	0	860	7	0
4	D	718	0	725	10	0
4	H	726	0	747	11	0
5	I	2975	0	1640	20	0
6	J	3011	0	1640	20	0
7	K	1802	0	1784	36	0
7	L	1795	0	1775	24	0
All	All	15660	0	13195	139	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 5.

The worst 5 of 139 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)
7:K:23:ILE:CG2	7:K:24:ASP:HA	1.25	1.59
7:K:23:ILE:HG22	7:K:24:ASP:CA	1.52	1.39
7:K:23:ILE:CG2	7:K:24:ASP:CA	2.06	1.31
7:K:15:ILE:CG1	7:K:23:ILE:HD11	1.76	1.15

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:K:15:ILE:CB	7:K:23:ILE:HD11	1.80	1.12

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	100	100
1	E	93/135 (69%)	93 (100%)	0	0	100	100
2	B	87/102 (85%)	87 (100%)	0	0	100	100
2	F	88/102 (86%)	86 (98%)	2 (2%)	0	100	100
3	C	105/129 (81%)	104 (99%)	1 (1%)	0	100	100
3	G	105/129 (81%)	104 (99%)	1 (1%)	0	100	100
4	D	91/122 (75%)	91 (100%)	0	0	100	100
4	H	91/122 (75%)	91 (100%)	0	0	100	100
7	K	212/229 (93%)	191 (90%)	18 (8%)	3 (1%)	14	51
7	L	211/229 (92%)	191 (90%)	18 (8%)	2 (1%)	21	64
All	All	1179/1434 (82%)	1133 (96%)	41 (4%)	5 (0%)	39	80

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	K	22	VAL
7	K	23	ILE
7	K	7	ASP
7	L	7	ASP
7	L	22	VAL

### 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%)	82 (100%)	0	100	100
2	B	72/78 (92%)	72 (100%)	0	100	100
2	F	72/78 (92%)	71 (99%)	1 (1%)	74	93
3	C	81/101 (80%)	80 (99%)	1 (1%)	78	94
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%)	78 (99%)	1 (1%)	76	93
7	K	201/213 (94%)	196 (98%)	5 (2%)	55	86
7	L	200/213 (94%)	194 (97%)	6 (3%)	48	83
All	All	1030/1208 (85%)	1016 (99%)	14 (1%)	74	93

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	K	155	GLU
7	K	215	VAL
7	L	153	LEU
7	K	113	ASP
7	L	28	ARG

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

Mol	Chain	Res	Type
3	C	73	ASN
3	G	31	HIS

### 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 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 ⓘ

### 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	A	98/135 (72%)	-0.05	0 100 100	49, 75, 125, 176	0
1	E	95/135 (70%)	-0.02	0 100 100	48, 75, 105, 155	0
2	B	89/102 (87%)	0.00	1 (1%) 82 58	45, 70, 118, 150	0
2	F	90/102 (88%)	0.07	3 (3%) 50 22	43, 69, 121, 149	0
3	C	107/129 (82%)	-0.15	1 (0%) 85 64	45, 69, 118, 140	0
3	G	107/129 (82%)	-0.12	2 (1%) 70 41	46, 69, 113, 136	0
4	D	93/122 (76%)	-0.24	0 100 100	43, 64, 109, 136	0
4	H	93/122 (76%)	-0.21	0 100 100	42, 65, 99, 128	0
5	I	146/147 (99%)	-0.52	0 100 100	87, 136, 191, 233	0
6	J	146/147 (99%)	-0.51	0 100 100	82, 140, 201, 233	0
7	K	214/229 (93%)	0.08	9 (4%) 40 16	48, 88, 178, 233	0
7	L	213/229 (93%)	0.12	7 (3%) 50 22	49, 91, 177, 242	0
All	All	1491/1728 (86%)	-0.12	23 (1%) 76 49	42, 83, 171, 242	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
7	L	26	ASN	4.9
7	L	24	ASP	4.7
3	G	14	ALA	4.3
7	L	30	ARG	4.0
7	L	19	GLN	3.9

### 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 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.