



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 03:02 am BST

PDB ID : 5WCU
Title : Crystal structure of 167 bp nucleosome bound to the globular domain of linker histone H5
Authors : Jiang, J.S.; Zhou, B.R.
Deposited on : 2017-07-02
Resolution : 5.53 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

<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	:	2.11
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

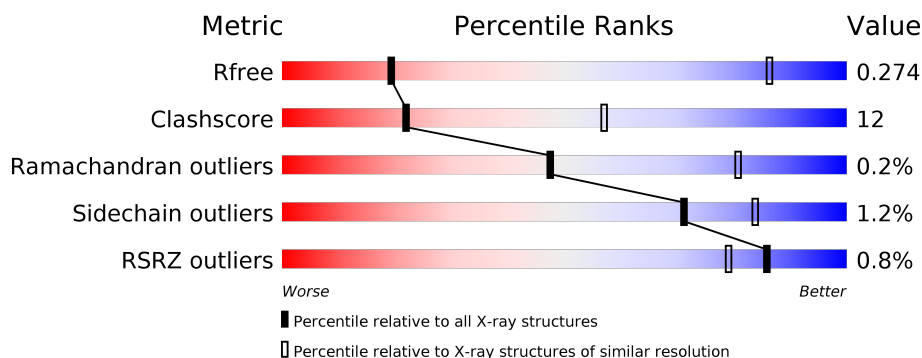
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.53 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.







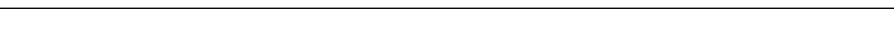





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1000 (7.20-3.84)
Clashscore	141614	1011 (7.12-3.90)
Ramachandran outliers	138981	1015 (7.20-3.82)
Sidechain outliers	138945	1191 (7.20-3.80)
RSRZ outliers	127900	1024 (7.32-3.76)

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 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	A	98	<div> <div>2%</div> <div> <div></div> <div>76%</div> <div>24%</div> </div> </div>
1	E	98	<div> <div>68%</div> <div>31%</div> <div></div> </div>
1	K	98	<div> <div>72%</div> <div>28%</div> <div></div> </div>
1	O	98	<div> <div>72%</div> <div>27%</div> <div></div> </div>
2	B	82	<div> <div>2%</div> <div>73%</div> <div>24%</div> <div></div> </div>
2	F	82	<div> <div>%</div> <div>73%</div> <div>27%</div> <div></div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	82	 73% 24% .
2	P	82	 73% 27% .
3	C	104	 64% 36% .
3	G	104	 79% 20% .
3	M	104	 81% 19% .
3	Q	104	 71% 28% .
4	D	94	 76% 24% .
4	H	94	 79% 18% ..
4	N	94	 79% 21% .
4	R	94	 80% 18% ..
5	I	167	 50% 49% .
5	S	167	 60% 38% .
6	J	167	 53% 47% .
6	T	167	 59% 41% .
7	U	76	 53% 43% .
7	V	76	 66% 30% .

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 26619 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	98	Total	C	N	O	S	0	0	0
			797	500	154	140	3			
1	E	97	Total	C	N	O	S	0	0	0
			801	505	155	138	3			
1	K	98	Total	C	N	O	S	0	0	0
			807	508	156	140	3			
1	O	97	Total	C	N	O	S	0	0	0
			801	505	155	138	3			

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	80	Total	C	N	O	S	0	0	0
			638	401	125	111	1			
2	F	82	Total	C	N	O	S	0	0	0
			653	412	127	113	1			
2	L	80	Total	C	N	O	S	0	0	0
			638	401	125	111	1			
2	P	82	Total	C	N	O	S	0	0	0
			651	411	127	112	1			

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	104	Total	C	N	O	S	0	0	0
			796	502	155	138	1			
3	G	103	Total	C	N	O	S	0	0	0
			786	494	153	138	1			
3	M	104	Total	C	N	O	S	0	0	0
			798	503	155	139	1			
3	Q	103	Total	C	N	O	S	0	0	0
			789	497	153	138	1			

- Molecule 4 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	94	Total	C	N	O	S	0	0	0
			743	468	135	138	2			
4	H	93	Total	C	N	O	S	0	0	0
			728	460	131	135	2			
4	N	94	Total	C	N	O	S	0	0	0
			743	468	135	138	2			
4	R	93	Total	C	N	O	S	0	0	0
			732	462	131	137	2			

- Molecule 5 is a DNA chain called DNA (167-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	164	Total	C	N	O	P	0	0	0
			3343	1587	609	983	164			
5	S	164	Total	C	N	O	P	0	0	0
			3343	1587	609	983	164			

- Molecule 6 is a DNA chain called DNA (167-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	167	Total	C	N	O	P	0	0	0
			3441	1629	642	1003	167			
6	T	167	Total	C	N	O	P	0	0	0
			3441	1629	642	1003	167			

- Molecule 7 is a protein called Histone H5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	U	76	Total	C	N	O	S	0	0	0
			575	358	113	103	1			
7	V	76	Total	C	N	O	S	0	0	0
			575	358	113	103	1			

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: Histone H3



- Molecule 1: Histone H3



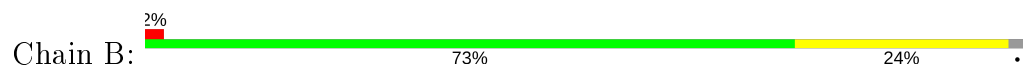
- Molecule 1: Histone H3



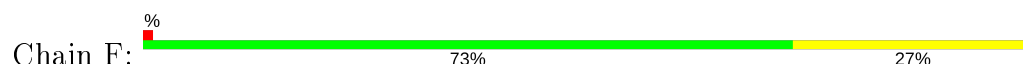
- Molecule 1: Histone H3



- Molecule 2: Histone H4



- Molecule 2: Histone H4





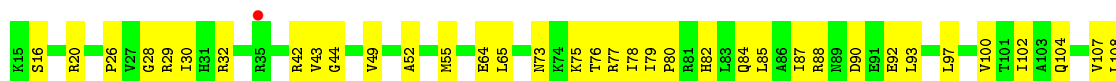
- Molecule 2: Histone H4



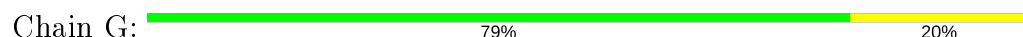
- Molecule 2: Histone H4



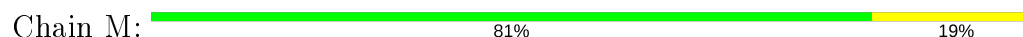
- Molecule 3: Histone H2A



- Molecule 3: Histone H2A




- Molecule 3: Histone H2A

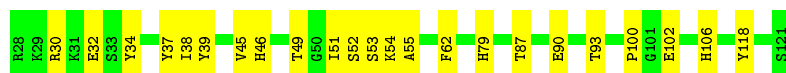


- Molecule 3: Histone H2A




- Molecule 4: Histone H2B

Chain D:  76% 24%




- Molecule 4: Histone H2B

Chain H:  79% 18% ..



- Molecule 4: Histone H2B

Chain N:  79% 21%



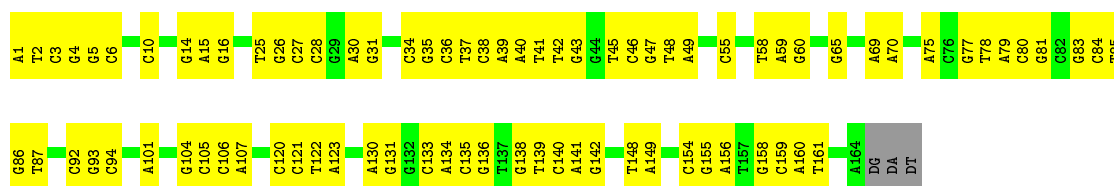
- Molecule 4: Histone H2B

Chain R:  80% 18% ..



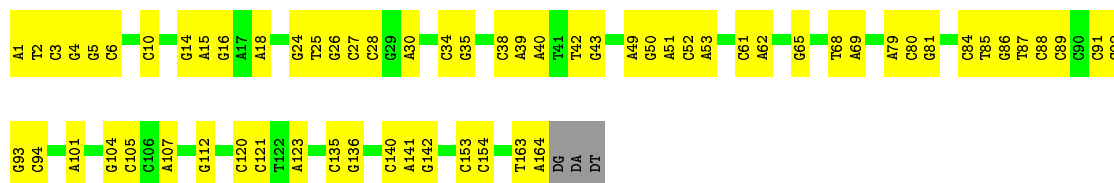
- Molecule 5: DNA (167-MER)

Chain I:  50% 49% .



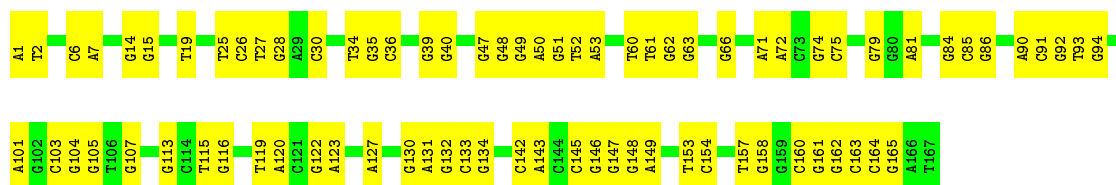
- Molecule 5: DNA (167-MER)

Chain S:  60% 38% .

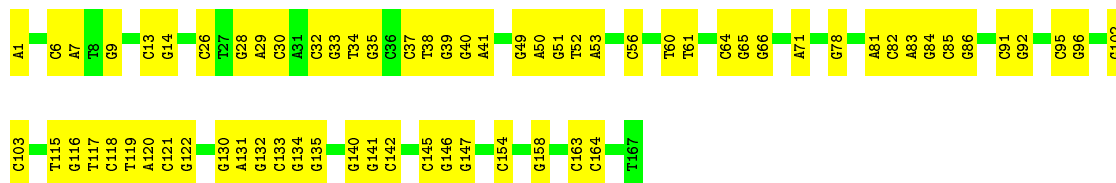


- Molecule 6: DNA (167-MER)

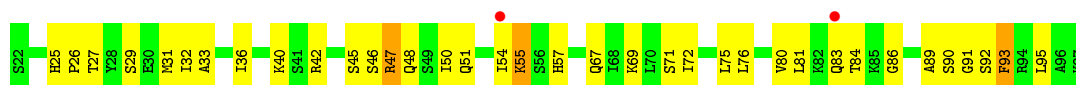
Chain J:  53% 47%



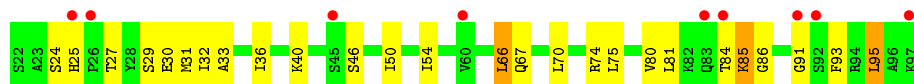
• Molecule 6: DNA (167-MER)



• Molecule 7: Histone H5



• Molecule 7: Histone H5



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	65.93Å 108.54Å 180.77Å 100.79° 90.08° 89.94°	Depositor
Resolution (Å)	48.62 – 5.53 48.62 – 5.53	Depositor EDS
% Data completeness (in resolution range)	97.6 (48.62-5.53) 97.4 (48.62-5.53)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.34 (at 5.39Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.189 , 0.238 0.229 , 0.274	Depositor DCC
R_{free} test set	1521 reflections (9.96%)	wwPDB-VP
Wilson B-factor (Å ²)	176.6	Xtriage
Anisotropy	0.803	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.22 , 91.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.337 for h,-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	26619	wwPDB-VP
Average B, all atoms (Å ²)	223.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.68% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	A	0.29	0/807	0.52	0/1081
1	E	0.31	0/813	0.53	0/1090
1	K	0.24	0/819	0.48	0/1097
1	O	0.26	0/813	0.50	0/1090
2	B	0.24	0/645	0.49	0/862
2	F	0.27	0/660	0.51	0/883
2	L	0.25	0/645	0.51	0/862
2	P	0.27	0/658	0.51	0/880
3	C	0.28	0/806	0.49	0/1087
3	G	0.27	0/796	0.54	0/1075
3	M	0.26	0/808	0.49	0/1090
3	Q	0.27	0/799	0.51	0/1079
4	D	0.30	0/754	0.51	0/1013
4	H	0.27	0/739	0.58	0/994
4	N	0.27	0/754	0.48	0/1013
4	R	0.30	0/743	0.50	0/999
5	I	0.62	0/3746	0.93	0/5774
5	S	0.46	0/3746	0.86	0/5774
6	J	0.66	0/3863	0.94	0/5965
6	T	0.46	0/3863	0.86	0/5965
7	U	0.24	0/582	0.42	0/775
7	V	0.30	0/582	0.46	0/775
All	All	0.45	0/28441	0.75	0/41223

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	797	0	824	22	0
1	E	801	0	839	32	0
1	K	807	0	844	27	0
1	O	801	0	839	31	0
2	B	638	0	676	16	0
2	F	653	0	696	26	0
2	L	638	0	676	17	0
2	P	651	0	691	23	0
3	C	796	0	843	44	0
3	G	786	0	826	23	0
3	M	798	0	848	20	0
3	Q	789	0	835	27	0
4	D	743	0	772	27	0
4	H	728	0	755	19	0
4	N	743	0	772	19	0
4	R	732	0	759	20	0
5	I	3343	0	1840	90	0
5	S	3343	0	1840	61	0
6	J	3441	0	1875	96	0
6	T	3441	0	1875	75	0
7	U	575	0	607	32	0
7	V	575	0	607	19	0
All	All	26619	0	21139	516	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 12.

The worst 5 of 516 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:A:41:TYR:HD2	6:J:94:DG:OP1	1.50	0.93
1:A:41:TYR:CD2	6:J:94:DG:OP1	2.34	0.80
1:E:43:PRO:HG2	6:J:79:DG:H5'	1.68	0.74
1:A:41:TYR:HD1	5:I:154:DC:H5'	1.54	0.73
1:A:41:TYR:HA	5:I:154:DC:H5''	1.73	0.70

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	96/98 (98%)	94 (98%)	2 (2%)	0	100	100
1	E	95/98 (97%)	91 (96%)	4 (4%)	0	100	100
1	K	96/98 (98%)	96 (100%)	0	0	100	100
1	O	95/98 (97%)	95 (100%)	0	0	100	100
2	B	78/82 (95%)	78 (100%)	0	0	100	100
2	F	80/82 (98%)	80 (100%)	0	0	100	100
2	L	78/82 (95%)	78 (100%)	0	0	100	100
2	P	80/82 (98%)	79 (99%)	1 (1%)	0	100	100
3	C	102/104 (98%)	101 (99%)	1 (1%)	0	100	100
3	G	101/104 (97%)	100 (99%)	1 (1%)	0	100	100
3	M	102/104 (98%)	99 (97%)	3 (3%)	0	100	100
3	Q	101/104 (97%)	100 (99%)	1 (1%)	0	100	100
4	D	92/94 (98%)	91 (99%)	1 (1%)	0	100	100
4	H	91/94 (97%)	88 (97%)	1 (1%)	2 (2%)	6	35
4	N	92/94 (98%)	90 (98%)	2 (2%)	0	100	100
4	R	91/94 (97%)	91 (100%)	0	0	100	100
7	U	74/76 (97%)	70 (95%)	3 (4%)	1 (1%)	11	45
7	V	74/76 (97%)	67 (90%)	7 (10%)	0	100	100
All	All	1618/1664 (97%)	1588 (98%)	27 (2%)	3 (0%)	47	81

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	H	120	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	51	ILE
7	U	26	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	82/85 (96%)	81 (99%)	1 (1%)	71	84
1	E	85/85 (100%)	85 (100%)	0	100	100
1	K	85/85 (100%)	85 (100%)	0	100	100
1	O	85/85 (100%)	85 (100%)	0	100	100
2	B	65/67 (97%)	65 (100%)	0	100	100
2	F	67/67 (100%)	67 (100%)	0	100	100
2	L	65/67 (97%)	65 (100%)	0	100	100
2	P	66/67 (98%)	64 (97%)	2 (3%)	41	63
3	C	80/81 (99%)	80 (100%)	0	100	100
3	G	79/81 (98%)	78 (99%)	1 (1%)	69	82
3	M	81/81 (100%)	81 (100%)	0	100	100
3	Q	80/81 (99%)	79 (99%)	1 (1%)	69	82
4	D	81/81 (100%)	81 (100%)	0	100	100
4	H	79/81 (98%)	78 (99%)	1 (1%)	69	82
4	N	81/81 (100%)	81 (100%)	0	100	100
4	R	80/81 (99%)	79 (99%)	1 (1%)	69	82
7	U	59/59 (100%)	54 (92%)	5 (8%)	10	35
7	V	59/59 (100%)	55 (93%)	4 (7%)	16	42
All	All	1359/1374 (99%)	1343 (99%)	16 (1%)	71	84

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	U	40	LYS
7	U	47	ARG
7	V	40	LYS
4	R	46	HIS
7	V	66	LEU

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

Mol	Chain	Res	Type
3	G	68	ASN
4	N	46	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](#)

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, 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	98/98 (100%)	-0.34	2 (2%) 65 57	183, 206, 232, 250	0
1	E	97/98 (98%)	-0.50	0 100 100	174, 199, 227, 250	0
1	K	98/98 (100%)	-0.39	0 100 100	180, 200, 226, 240	0
1	O	97/98 (98%)	-0.50	0 100 100	190, 213, 230, 235	0
2	B	80/82 (97%)	-0.10	2 (2%) 57 49	178, 207, 224, 227	0
2	F	82/82 (100%)	-0.28	1 (1%) 79 71	176, 197, 212, 221	0
2	L	80/82 (97%)	-0.15	0 100 100	158, 196, 210, 220	0
2	P	82/82 (100%)	-0.07	1 (1%) 79 71	181, 205, 224, 236	0
3	C	104/104 (100%)	-0.26	2 (1%) 66 58	168, 189, 214, 228	0
3	G	103/104 (99%)	-0.33	0 100 100	168, 194, 219, 233	0
3	M	104/104 (100%)	-0.21	0 100 100	167, 192, 213, 219	0
3	Q	103/104 (99%)	-0.35	0 100 100	158, 193, 209, 218	0
4	D	94/94 (100%)	-0.42	0 100 100	168, 189, 208, 224	0
4	H	93/94 (98%)	-0.29	0 100 100	167, 189, 214, 229	0
4	N	94/94 (100%)	-0.38	0 100 100	166, 189, 207, 214	0
4	R	93/94 (98%)	-0.34	0 100 100	163, 190, 211, 224	0
5	I	164/167 (98%)	-0.67	0 100 100	214, 242, 271, 281	0
5	S	164/167 (98%)	-0.67	0 100 100	211, 241, 280, 290	0
6	J	167/167 (100%)	-0.68	0 100 100	207, 244, 273, 290	0
6	T	167/167 (100%)	-0.74	0 100 100	209, 242, 279, 290	0
7	U	76/76 (100%)	0.12	2 (2%) 56 47	196, 234, 252, 263	0
7	V	76/76 (100%)	0.45	9 (11%) 4 7	201, 230, 242, 254	14 (18%)
All	All	2316/2332 (99%)	-0.38	19 (0%) 86 79	158, 208, 260, 290	14 (0%)

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	98	TYR	3.2
7	V	91	GLY	3.1
7	U	83	GLN	2.8
3	C	35	ARG	2.8
2	F	44	LYS	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.