



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 12:18 PM BST

PDB ID : 1ZBB
Title : Structure of the 4_601_167 Tetranucleosome
Authors : Schalch, T.; Duda, S.; Sargent, D.F.; Richmond, T.J.
Deposited on : 2005-04-08
Resolution : 9.00 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

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	:	2.13
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)	:	2.13

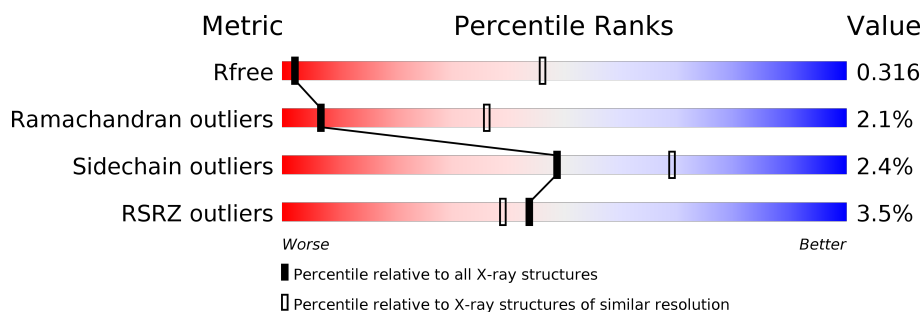
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 9.00 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}	130704	1005 (11.50-3.90)
Ramachandran outliers	138981	1003 (11.50-3.90)
Sidechain outliers	138945	1003 (11.50-3.86)
RSRZ outliers	127900	1004 (9.50-3.80)

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	I	347	<div> <div>3%</div> <div>97%</div> <div>.</div> </div>
2	J	347	<div> <div>3%</div> <div>99%</div> <div>.</div> </div>
3	A	135	<div> <div>70%</div> <div>28%</div> <div>.</div> </div>
3	E	135	<div> <div>70%</div> <div>28%</div> <div>.</div> </div>
3	a	135	<div> <div>70%</div> <div>28%</div> <div>.</div> </div>
3	e	135	<div> <div>70%</div> <div>28%</div> <div>.</div> </div>
4	B	102	<div> <div>4%</div> <div>77%</div> <div>22%</div> <div>.</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	102	
4	b	102	
4	f	102	
5	C	129	
5	G	129	
5	c	129	
5	g	129	
6	D	125	
6	H	125	
6	d	125	
6	h	125	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 26851 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 STRAND 1 (ARBITRARY MODEL SEQUENCE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	347	Total	C	N	O	P	0	0	0
			7111	3394	1292	2079	346			

- Molecule 2 is a DNA chain called DNA STRAND 2 (ARBITRARY MODEL SEQUENCE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	347	Total	C	N	O	P	0	0	0
			7110	3394	1289	2081	346			

- Molecule 3 is a protein called HISTONE H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
3	E	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
3	a	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			
3	e	97	Total	C	N	O	S	0	0	0
			801	504	155	139	3			

There are 4 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
a	102	ALA	GLY	CONFLICT	UNP P84233
e	102	ALA	GLY	CONFLICT	UNP P84233

- Molecule 4 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	80	Total	C	N	O	S	0	0	0
			634	400	122	111	1			
4	F	94	Total	C	N	O	S	0	0	0
			750	469	154	126	1			
4	b	80	Total	C	N	O	S	0	0	0
			634	400	122	111	1			
4	f	94	Total	C	N	O	S	0	0	0
			750	469	154	126	1			

- Molecule 5 is a protein called Histone H2A.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	111	Total	C	N	O		0	0	0
			858	539	169	150				
5	G	109	Total	C	N	O		0	0	0
			843	531	164	148				
5	c	111	Total	C	N	O		0	0	0
			858	539	169	150				
5	g	109	Total	C	N	O		0	0	0
			843	531	164	148				

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	99	ARG	GLY	CONFLICT	UNP P06897
C	123	SER	ALA	CONFLICT	UNP P06897
G	99	ARG	GLY	CONFLICT	UNP P06897
G	123	SER	ALA	CONFLICT	UNP P06897
c	99	ARG	GLY	CONFLICT	UNP P06897
c	123	SER	ALA	CONFLICT	UNP P06897
g	99	ARG	GLY	CONFLICT	UNP P06897
g	123	SER	ALA	CONFLICT	UNP P06897

- Molecule 6 is a protein called Histone H2B.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	107	Total	C	N	O	S	0	0	0
			843	528	158	155	2			
6	H	99	Total	C	N	O	S	0	0	0
			785	493	146	144	2			
6	d	107	Total	C	N	O	S	0	0	0
			843	528	158	155	2			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	h	99	Total	C	N	O	S	0	0	0
			785	493	146	144	2			

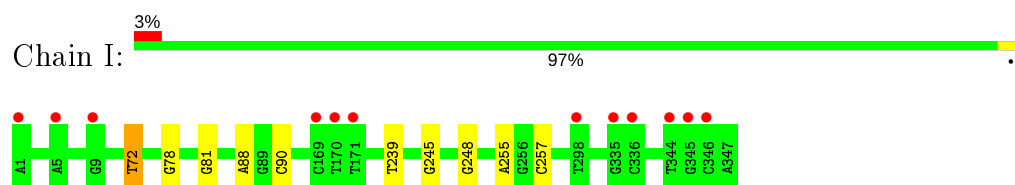
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	29	THR	SER	CONFLICT	UNP P02281
H	29	THR	SER	CONFLICT	UNP P02281
d	29	THR	SER	CONFLICT	UNP P02281
h	29	THR	SER	CONFLICT	UNP P02281

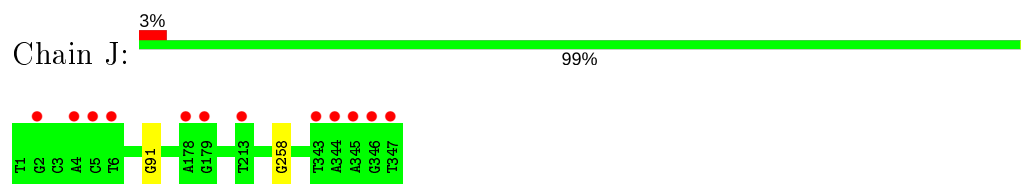
3 Residue-property plots [i](#)

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.

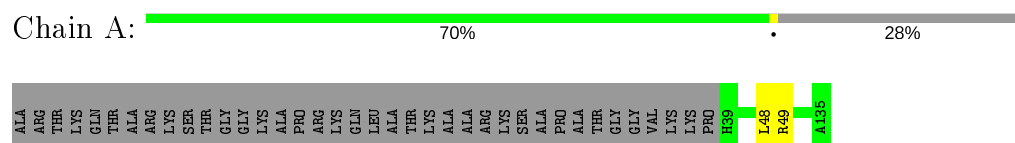
- Molecule 1: DNA STRAND 1 (ARBITRARY MODEL SEQUENCE)



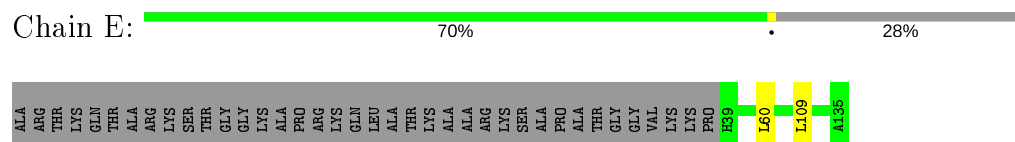
- Molecule 2: DNA STRAND 2 (ARBITRARY MODEL SEQUENCE)



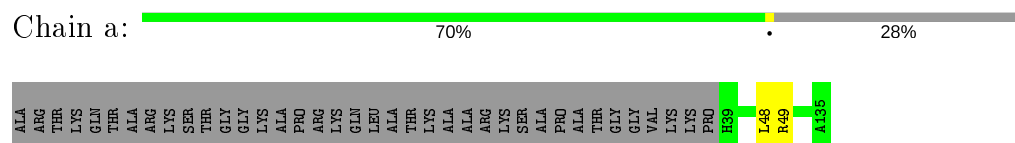
- Molecule 3: HISTONE H3



- Molecule 3: HISTONE H3

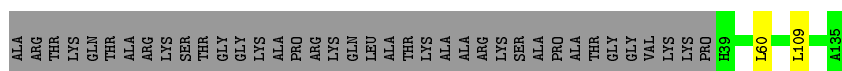


- Molecule 3: HISTONE H3

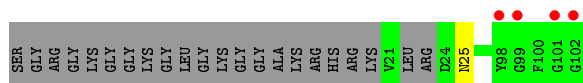
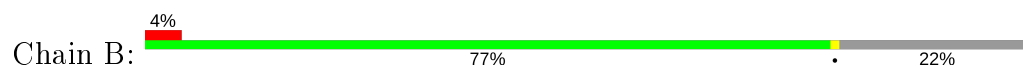


- Molecule 3: HISTONE H3

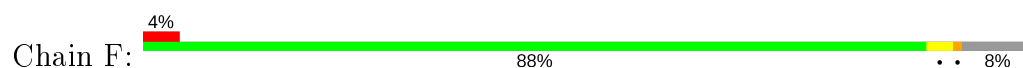




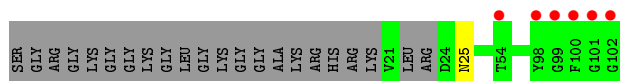
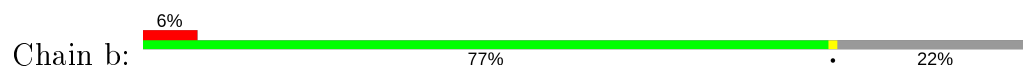
- Molecule 4: Histone H4



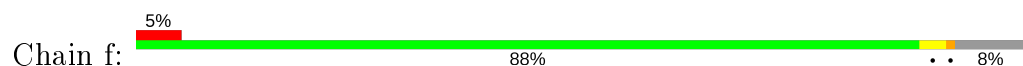
- Molecule 4: Histone H4



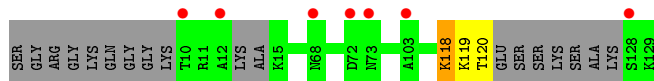
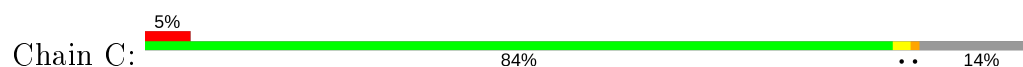
- Molecule 4: Histone H4



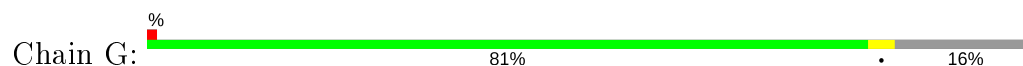
- Molecule 4: Histone H4



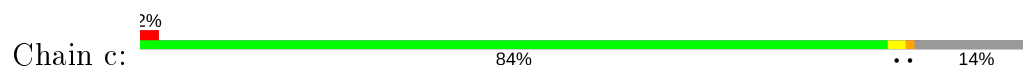
- Molecule 5: Histone H2A.1

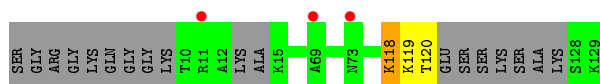


- Molecule 5: Histone H2A.1

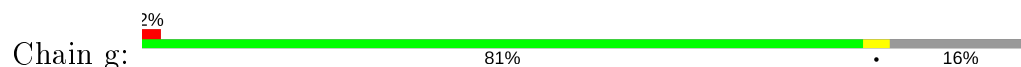


- Molecule 5: Histone H2A.1

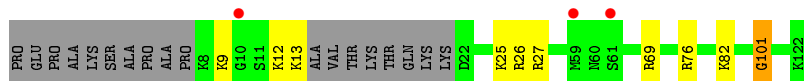
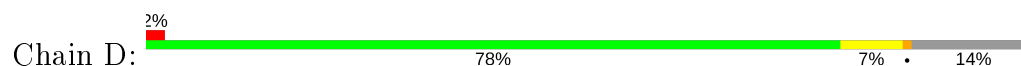




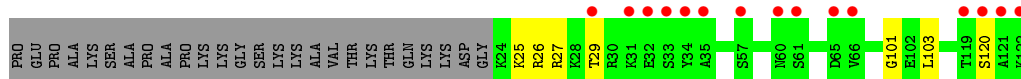
• Molecule 5: Histone H2A.1



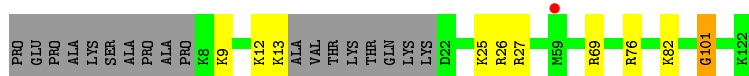
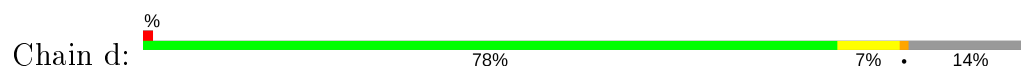
• Molecule 6: Histone H2B.1



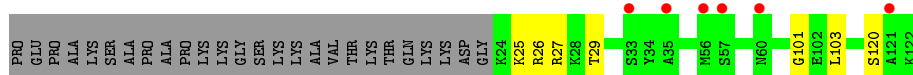
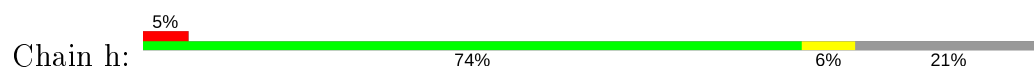
• Molecule 6: Histone H2B.1



• Molecule 6: Histone H2B.1



• Molecule 6: Histone H2B.1



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	127.67Å 168.44Å 237.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 9.00 137.32 – 9.00	Depositor EDS
% Data completeness (in resolution range)	97.0 (50.00-9.00) 97.0 (137.32-9.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	9.25 (at 8.44Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.386 , (Not available) 0.386 , 0.316	Depositor DCC
R_{free} test set	91 reflections (4.43%)	wwPDB-VP
Wilson B-factor (Å ²)	749.9	Xtriage
Anisotropy	0.482	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.02 , -10.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	26851	wwPDB-VP
Average B, all atoms (Å ²)	234.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.81% 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 ⓘ

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	I	0.46	0/7978	0.74	1/12312 (0.0%)
2	J	0.46	0/7976	0.73	0/12309
3	A	0.57	0/812	0.71	1/1088 (0.1%)
3	E	0.65	0/812	0.76	0/1088
3	a	0.57	0/812	0.71	1/1088 (0.1%)
3	e	0.65	0/812	0.75	0/1088
4	B	0.57	0/640	0.78	0/855
4	F	0.62	0/756	0.80	0/1001
4	b	0.57	0/640	0.78	0/855
4	f	0.62	0/756	0.80	0/1001
5	C	0.58	0/865	0.72	0/1161
5	G	0.46	0/853	0.67	0/1150
5	c	0.58	0/865	0.72	0/1161
5	g	0.46	0/853	0.67	0/1150
6	D	0.58	0/853	0.74	3/1135 (0.3%)
6	H	0.48	0/796	0.65	0/1065
6	d	0.58	0/853	0.74	3/1135 (0.3%)
6	h	0.48	0/796	0.65	0/1065
All	All	0.51	0/28728	0.73	9/41707 (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	10
2	J	0	2
All	All	0	12

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	d	76	ARG	NE-CZ-NH2	-5.35	117.62	120.30
6	D	69	ARG	NE-CZ-NH2	-5.32	117.64	120.30
3	A	49	ARG	NE-CZ-NH2	-5.27	117.66	120.30
6	D	76	ARG	NE-CZ-NH2	-5.22	117.69	120.30
3	a	49	ARG	NE-CZ-NH2	-5.20	117.70	120.30

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	72	DT	Sidechain
1	I	78	DG	Sidechain
1	I	81	DG	Sidechain
1	I	88	DA	Sidechain
1	I	90	DC	Sidechain

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	
3	A	95/135 (70%)	95 (100%)	0	0	100	100
3	E	95/135 (70%)	95 (100%)	0	0	100	100
3	a	95/135 (70%)	95 (100%)	0	0	100	100
3	e	95/135 (70%)	95 (100%)	0	0	100	100
4	B	77/102 (76%)	76 (99%)	0	1 (1%)	12	48
4	F	88/102 (86%)	83 (94%)	3 (3%)	2 (2%)	6	34
4	b	77/102 (76%)	76 (99%)	0	1 (1%)	12	48

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	f	88/102 (86%)	83 (94%)	3 (3%)	2 (2%)	6	34
5	C	105/129 (81%)	103 (98%)	1 (1%)	1 (1%)	15	55
5	G	107/129 (83%)	101 (94%)	4 (4%)	2 (2%)	8	38
5	c	105/129 (81%)	103 (98%)	1 (1%)	1 (1%)	15	55
5	g	107/129 (83%)	101 (94%)	4 (4%)	2 (2%)	8	38
6	D	103/125 (82%)	95 (92%)	3 (3%)	5 (5%)	2	20
6	H	97/125 (78%)	91 (94%)	1 (1%)	5 (5%)	2	19
6	d	103/125 (82%)	95 (92%)	3 (3%)	5 (5%)	2	20
6	h	97/125 (78%)	91 (94%)	1 (1%)	5 (5%)	2	19
All	All	1534/1964 (78%)	1478 (96%)	24 (2%)	32 (2%)	7	36

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	D	101	GLY
5	G	14	ALA
6	H	26	ARG
6	d	101	GLY
5	g	14	ALA

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	
3	A	84/110 (76%)	83 (99%)	1 (1%)	71	83
3	E	84/110 (76%)	82 (98%)	2 (2%)	49	69
3	a	84/110 (76%)	83 (99%)	1 (1%)	71	83
3	e	84/110 (76%)	82 (98%)	2 (2%)	49	69
4	B	65/78 (83%)	65 (100%)	0	100	100
4	F	76/78 (97%)	73 (96%)	3 (4%)	32	56
4	b	65/78 (83%)	65 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	f	76/78 (97%)	73 (96%)	3 (4%)	32	56
5	C	89/101 (88%)	86 (97%)	3 (3%)	37	60
5	G	87/101 (86%)	85 (98%)	2 (2%)	50	70
5	c	89/101 (88%)	86 (97%)	3 (3%)	37	60
5	g	87/101 (86%)	85 (98%)	2 (2%)	50	70
6	D	91/105 (87%)	88 (97%)	3 (3%)	38	61
6	H	85/105 (81%)	83 (98%)	2 (2%)	49	69
6	d	91/105 (87%)	88 (97%)	3 (3%)	38	61
6	h	85/105 (81%)	83 (98%)	2 (2%)	49	69
All	All	1322/1576 (84%)	1290 (98%)	32 (2%)	49	69

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	H	25	LYS
5	c	118	LYS
5	g	121	GLU
3	a	48	LEU
5	c	119	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
3	a	39	HIS
4	b	93	GLN
5	g	31	HIS
5	G	110	ASN
3	e	125	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
5	C	1
5	c	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	128:SER	C	129:LYS	N	5.58
1	c	128:SER	C	129:LYS	N	5.58

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	I	347/347 (100%)	0.06	12 (3%) 44 39	238, 238, 238, 238	347 (100%)
2	J	347/347 (100%)	0.07	12 (3%) 44 39	238, 238, 238, 238	347 (100%)
3	A	97/135 (71%)	-0.43	0 100 100	230, 230, 230, 230	97 (100%)
3	E	97/135 (71%)	-0.56	0 100 100	230, 230, 230, 230	97 (100%)
3	a	97/135 (71%)	-0.27	0 100 100	230, 230, 230, 230	97 (100%)
3	e	97/135 (71%)	-0.55	0 100 100	230, 230, 230, 230	97 (100%)
4	B	80/102 (78%)	0.10	4 (5%) 28 28	230, 230, 230, 230	80 (100%)
4	F	94/102 (92%)	-0.23	4 (4%) 35 32	230, 230, 230, 230	94 (100%)
4	b	80/102 (78%)	0.06	6 (7%) 14 15	230, 230, 230, 230	80 (100%)
4	f	94/102 (92%)	-0.15	5 (5%) 26 26	230, 230, 230, 230	94 (100%)
5	C	111/129 (86%)	0.17	7 (6%) 20 19	230, 230, 230, 230	111 (100%)
5	G	109/129 (84%)	-0.07	1 (0%) 84 77	230, 230, 230, 230	109 (100%)
5	c	111/129 (86%)	0.19	3 (2%) 54 48	230, 230, 230, 230	111 (100%)
5	g	109/129 (84%)	-0.11	2 (1%) 68 61	230, 230, 230, 230	109 (100%)
6	D	107/125 (85%)	-0.13	3 (2%) 53 46	230, 230, 230, 230	107 (100%)
6	H	99/125 (79%)	0.44	15 (15%) 2 5	230, 230, 230, 230	99 (100%)
6	d	107/125 (85%)	-0.18	1 (0%) 84 77	230, 230, 230, 230	107 (100%)
6	h	99/125 (79%)	0.23	6 (6%) 21 20	230, 230, 230, 230	99 (100%)
All	All	2282/2658 (85%)	-0.04	81 (3%) 44 39	230, 230, 238, 238	2282 (100%)

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	101	GLY	6.1
4	b	101	GLY	5.9
2	J	5	DC	5.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
6	H	121	ALA	5.6
6	H	29	THR	5.1

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.