



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 10:36 pm GMT

PDB ID : 5B2J
Title : Human nucleosome containing CpG methylated DNA
Authors : Fujii, Y.; Wakamori, M.; Umehara, T.; Yokoyama, S.
Deposited on : 2016-01-18
Resolution : 2.60 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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

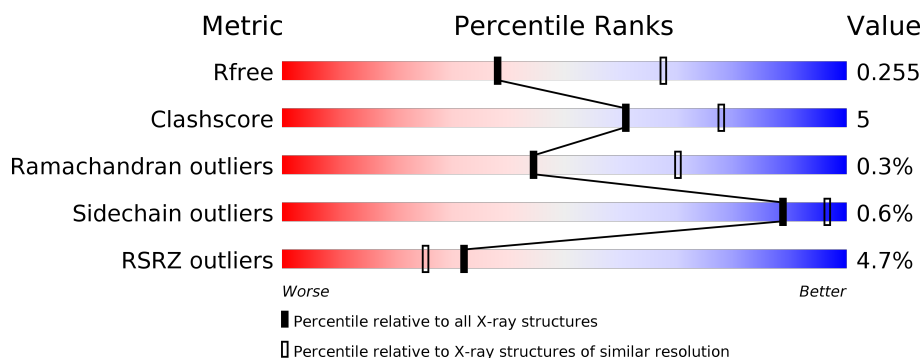
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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}	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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	139	<div> <div>4%</div> <div> <div></div> <div>62%</div> <div>9%</div> <div>29%</div> </div> </div>
1	E	139	<div> <div>%</div> <div> <div></div> <div>63%</div> <div>8%</div> <div>29%</div> </div> </div>
2	B	106	<div> <div>4%</div> <div> <div></div> <div>67%</div> <div>10%</div> <div>23%</div> </div> </div>
2	F	106	<div> <div>3%</div> <div> <div></div> <div>70%</div> <div>10%</div> <div>20%</div> </div> </div>
3	C	133	<div> <div></div> <div> <div></div> <div>74%</div> <div>6%</div> <div>20%</div> </div> </div>
3	G	133	<div> <div>4%</div> <div> <div></div> <div>71%</div> <div>8%</div> <div>21%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	129	
4	H	129	
5	I	146	
5	J	146	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	MN	E	301	-	-	-	X
6	MN	I	101	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12171 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 H3.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	99	Total	C	N	O	S	0	0	0
			815	514	158	139	4			
1	E	99	Total	C	N	O	S	0	0	0
			816	514	158	140	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P68431
A	-2	SER	-	expression tag	UNP P68431
A	-1	HIS	-	expression tag	UNP P68431
E	-3	GLY	-	expression tag	UNP P68431
E	-2	SER	-	expression tag	UNP P68431
E	-1	HIS	-	expression tag	UNP P68431

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	82	Total	C	N	O	S	0	0	0
			653	412	127	113	1			
2	F	85	Total	C	N	O	S	0	0	0
			683	430	136	116	1			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-3	GLY	-	expression tag	UNP P62805
B	-2	SER	-	expression tag	UNP P62805
B	-1	HIS	-	expression tag	UNP P62805
F	-3	GLY	-	expression tag	UNP P62805
F	-2	SER	-	expression tag	UNP P62805
F	-1	HIS	-	expression tag	UNP P62805

- Molecule 3 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	106	Total	C	N	O	0	0	0
			819	517	160	142			
3	G	105	Total	C	N	O	0	0	0
			810	511	158	141			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP P04908
C	-2	SER	-	expression tag	UNP P04908
C	-1	HIS	-	expression tag	UNP P04908
G	-3	GLY	-	expression tag	UNP P04908
G	-2	SER	-	expression tag	UNP P04908
G	-1	HIS	-	expression tag	UNP P04908

- Molecule 4 is a protein called Histone H2B type 1-J.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	97	Total	C	N	O	S	0	0	0
			765	480	142	141	2			
4	H	94	Total	C	N	O	S	0	0	0
			736	462	134	138	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-6	GLY	-	expression tag	UNP P06899
D	-5	SER	-	expression tag	UNP P06899
D	-4	HIS	-	expression tag	UNP P06899
H	-6	GLY	-	expression tag	UNP P06899
H	-5	SER	-	expression tag	UNP P06899
H	-4	HIS	-	expression tag	UNP P06899

- 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
			2994	1436	539	874	145			
5	J	146	Total	C	N	O	P	0	0	0
			2994	1436	539	874	145			

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	3	Total 3	Mn 3	0	0
6	I	3	Total 3	Mn 3	0	0
6	E	1	Total 1	Mn 1	0	0

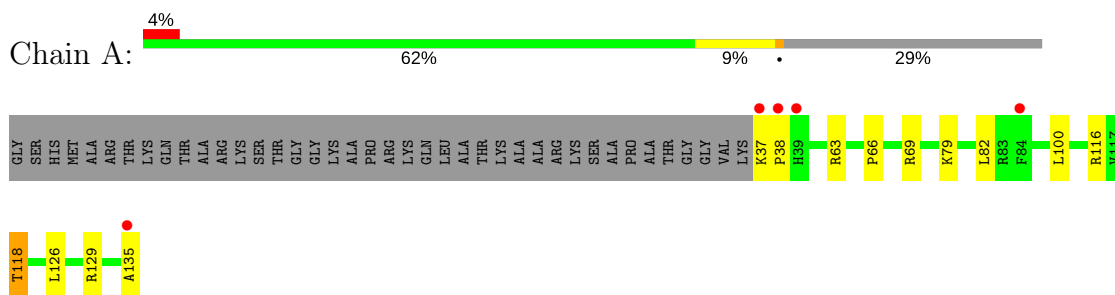
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	7	Total 7	O 7	0	0
7	B	4	Total 4	O 4	0	0
7	C	7	Total 7	O 7	0	0
7	D	8	Total 8	O 8	0	0
7	E	19	Total 19	O 19	0	0
7	F	15	Total 15	O 15	0	0
7	G	6	Total 6	O 6	0	0
7	H	2	Total 2	O 2	0	0
7	I	5	Total 5	O 5	0	0
7	J	6	Total 6	O 6	0	0

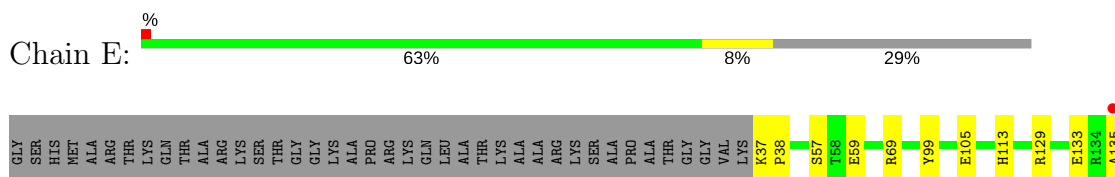
3 Residue-property plots [i](#)

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.

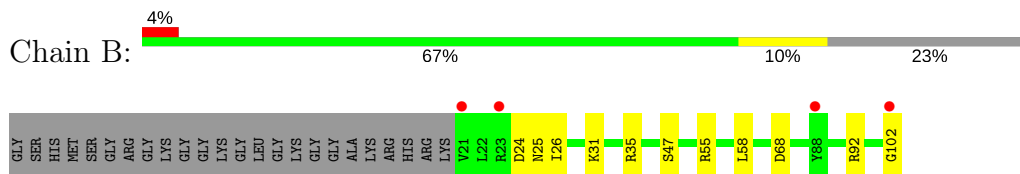
- Molecule 1: Histone H3.1



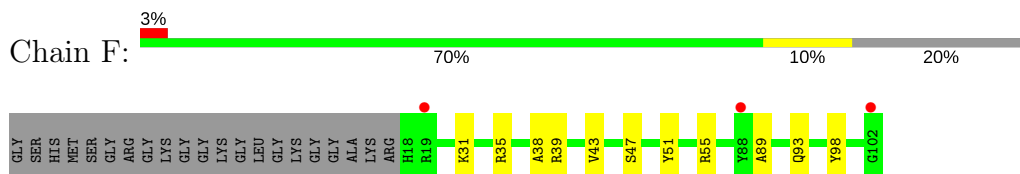
- Molecule 1: Histone H3.1



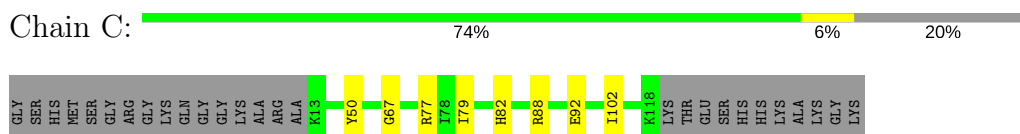
- Molecule 2: Histone H4



- Molecule 2: Histone H4



- Molecule 3: Histone H2A type 1-B/E



- Chain G:
-
- 4% 71% 8% 21%
- GLY SER HIS MET SER ARG GLY GLY LYS GLN GLY LYS ALA ARG ALA LYS K14 K15 T16 P26 R29 R32 G37 N38 Y39 V43 I62 I79 H82 I87 L93 L97 K118 LYS THR GLU SER HIS HIS LYS ALA LYS GLY

- Chain D:
-
- | Amino Acid | Percentage |
|------------|------------|
| GLY | 2% |
| SER | 64% |
| HIS | 0% |
| MET | 0% |
| PRO | 0% |
| GLU | 0% |
| PRO | 0% |
| ALA | 0% |
| LYS | 11% |
| SER | 0% |
| ALA | 0% |
| PRO | 0% |
| ALA | 0% |
| PRO | 0% |
| LYS | 0% |
| LYS | 0% |
| SER | 0% |
| GLY | 0% |
| LYS | 0% |
| LYS | 0% |
| VAL | 0% |
| THR | 0% |
| LYS | 0% |
| ALA | 0% |
| ALA | 0% |
| GLN | 0% |
| LYS | 0% |
| LYS | 0% |
| ASP | 0% |
| GLY | 0% |
| LYS | 0% |
| LYS | 0% |
| R26 | 0% |
| R27 | 0% |
| R30 | 0% |
| K43 | 0% |
| Q44 | 0% |
| V45 | 0% |
| H46 | 0% |
| K54 | 0% |
| I58 | 0% |
| D65 | 0% |
| V95 | 0% |
| L98 | 0% |
| L99 | 0% |
| P100 | 0% |
| G101 | 0% |
| E102 | 0% |
| T112 | 0% |
| T116 | 0% |

- Chain H:
-
- | Residue | Type |
|---------|--------|
| GLY | Green |
| SER | Green |
| HIS | Green |
| MET | Green |
| PRO | Green |
| GLU | Green |
| PRO | Green |
| ALA | Green |
| LYS | Green |
| SER | Green |
| PRO | Green |
| PRO | Green |
| PRO | Green |
| LYS | Green |
| GLY | Green |
| SER | Green |
| LYS | Green |
| LYS | Green |
| ALA | Green |
| VAL | Green |
| THR | Green |
| LYS | Green |
| ALA | Green |
| GLN | Green |
| LYS | Green |
| LYS | Green |
| ASP | Green |
| GLY | Green |
| LYS | Green |
| LYS | Green |
| ARG | Green |
| LYS | Green |
| R28 | Red |
| S29 | Green |
| R30 | Red |
| K31 | Green |
| E32 | Yellow |
| S33 | Yellow |
| Y36 | Red |
| Y37 | Yellow |
| V38 | Yellow |
| Y39 | Yellow |
| K43 | Yellow |
| N64 | Grey |
| D65 | Grey |
| E68 | Grey |
| G101 | Grey |
| A121 | Green |

- Chain I:
-
- | Segment Color | Percentage |
|---------------|------------|
| Red | 8% |
| Green | 68% |
| Yellow | 32% |
- Below the bar chart, a sequence of labels is shown, each with a colored dot above it:
- A-72 (Red dot)
 - A-69 (Green dot)
 - A-68 (Green dot)
 - C-63 (Green dot)
 - A-62 (Green dot)
 - T-59 (Red dot)
 - G-58 (Green dot)
 - C-57 (Green dot)
 - C-56 (Green dot)
 - T-50 (Red dot)
 - A-49 (Red dot)
 - C-48 (Red dot)
 - C-47 (Red dot)
 - A-46 (Green dot)
 - A-45 (Green dot)
 - A-44 (Green dot)
 - A-43 (Green dot)
 - T-35 (Green dot)
 - G-34 (Green dot)
 - A-30 (Red dot)
 - C-29 (Red dot)
 - T-28 (Red dot)
 - A-24 (Green dot)
 - A-23 (Green dot)
 - G-20 (Green dot)
 - A-19 (Green dot)
 - A-18 (Red dot)
 - A-17 (Red dot)
 - A-16 (Green dot)
 - G-15 (Green dot)
 - G-14 (Green dot)
 - C-13 (Green dot)
 - A-12 (Green dot)
 - T-3 (Green dot)
 - G-2 (Green dot)
 - T7 (Green dot)
 - G8 (Green dot)
 - C15 (Green dot)
 - C16 (Green dot)
 - T17 (Green dot)
 - A22 (Green dot)
 - G23 (Green dot)
 - A29 (Green dot)
 - G30 (Green dot)
 - T31 (Green dot)
 - G34 (Green dot)

- Chain J:
-
- 10% 66% 34%
- A-73 T-72 A-70 T-66 C-65 C-64 A-63 T-60 G-59 C-58 C-57 A-56 G-55 A-50 C-49 C-48 A-47 A-46 A-45 G-41 T-40 A-39 T-36 G-35 G-34 A-33 A-32 A-31 C-30 C-14 A-13 T0 T1 T6 G7 A8 G13 T17 A21 G22 T23 T24 G29 C34 A35
- A37 T38 A41 C42 T43 T44 T45 T46 G47 A52 A53 C54 G57 C58 A59 C60 G61 G64 A65 T66 T72

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	106.54Å 109.67Å 181.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.90 – 2.60 48.76 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.90-2.60) 89.4 (48.76-2.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.13	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.197 , 0.252 0.198 , 0.255	Depositor DCC
R_{free} test set	1765 reflections (2.69%)	DCC
Wilson B-factor (Å ²)	52.2	Xtriage
Anisotropy	0.308	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 58.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.033 for k,h,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	12171	wwPDB-VP
Average B, all atoms (Å ²)	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 5CM, MN

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	0.26	0/827	0.41	0/1109
1	E	0.26	0/828	0.43	0/1109
2	B	0.23	0/660	0.43	0/883
2	F	0.25	0/691	0.45	0/923
3	C	0.25	0/829	0.43	0/1118
3	G	0.24	0/820	0.41	0/1107
4	D	0.26	0/776	0.41	0/1040
4	H	0.24	0/747	0.42	0/1004
5	I	0.52	0/3266	0.95	0/5031
5	J	0.53	0/3266	0.96	0/5031
All	All	0.42	0/12710	0.76	0/18355

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 [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	815	0	856	10	0
1	E	816	0	856	9	0
2	B	653	0	696	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	683	0	729	7	0
3	C	819	0	879	6	0
3	G	810	0	866	7	0
4	D	765	0	797	11	0
4	H	736	0	758	9	0
5	I	2994	0	1661	34	0
5	J	2994	0	1661	38	0
6	E	1	0	0	0	0
6	I	3	0	0	0	0
6	J	3	0	0	0	0
7	A	7	0	0	0	0
7	B	4	0	0	0	0
7	C	7	0	0	0	0
7	D	8	0	0	0	0
7	E	19	0	0	1	0
7	F	15	0	0	0	0
7	G	6	0	0	0	0
7	H	2	0	0	0	0
7	I	5	0	0	1	0
7	J	6	0	0	2	0
All	All	12171	0	9759	110	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 110 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)
5:I:29:DA:OP2	7:I:201:HOH:O	2.00	0.79
1:E:69:ARG:NH2	5:I:17:DT:OP2	2.16	0.79
5:J:-14:DC:H2''	5:J:-13:DA:C8	2.25	0.72
2:B:92:ARG:HH21	4:D:98:LEU:HD23	1.53	0.72
4:H:30:ARG:NH2	5:J:-45:DA:OP1	2.25	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	97/139 (70%)	97 (100%)	0	0	100	100
1	E	97/139 (70%)	96 (99%)	1 (1%)	0	100	100
2	B	80/106 (76%)	78 (98%)	2 (2%)	0	100	100
2	F	83/106 (78%)	80 (96%)	3 (4%)	0	100	100
3	C	104/133 (78%)	103 (99%)	1 (1%)	0	100	100
3	G	103/133 (77%)	101 (98%)	2 (2%)	0	100	100
4	D	95/129 (74%)	94 (99%)	0	1 (1%)	17	35
4	H	92/129 (71%)	90 (98%)	1 (1%)	1 (1%)	17	35
All	All	751/1014 (74%)	739 (98%)	10 (1%)	2 (0%)	44	70

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	H	101	GLY
4	D	101	GLY

5.3.2 Protein sidechains [i](#)

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	86/113 (76%)	85 (99%)	1 (1%)	75	91
1	E	86/113 (76%)	86 (100%)	0	100	100
2	B	67/81 (83%)	66 (98%)	1 (2%)	70	88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	70/81 (86%)	69 (99%)	1 (1%)	71	89
3	C	84/102 (82%)	83 (99%)	1 (1%)	75	91
3	G	83/102 (81%)	83 (100%)	0	100	100
4	D	83/107 (78%)	83 (100%)	0	100	100
4	H	80/107 (75%)	80 (100%)	0	100	100
All	All	639/806 (79%)	635 (99%)	4 (1%)	89	96

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	118	THR
2	B	47	SER
3	C	50	TYR
2	F	47	SER

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

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

8 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	5CM	I	-5	5	14,21,22	1.23	2 (14%)	18,30,33	0.63	0
5	5CM	I	-61	5	14,21,22	1.21	2 (14%)	18,30,33	0.72	0
5	5CM	I	5	5	14,21,22	1.12	2 (14%)	18,30,33	0.74	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	5CM	I	61	5	14,21,22	1.16	1 (7%)	18,30,33	0.65	0
5	5CM	J	-6	5	14,21,22	1.14	2 (14%)	18,30,33	0.67	1 (5%)
5	5CM	J	-62	5	14,21,22	1.23	2 (14%)	18,30,33	0.79	0
5	5CM	J	4	5	14,21,22	1.20	2 (14%)	18,30,33	0.77	0
5	5CM	J	60	5	14,21,22	1.15	2 (14%)	18,30,33	0.86	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	5CM	I	-5	5	-	0/3/21/22	0/2/2/2
5	5CM	I	-61	5	-	0/3/21/22	0/2/2/2
5	5CM	I	5	5	-	0/3/21/22	0/2/2/2
5	5CM	I	61	5	-	0/3/21/22	0/2/2/2
5	5CM	J	-6	5	-	0/3/21/22	0/2/2/2
5	5CM	J	-62	5	-	0/3/21/22	0/2/2/2
5	5CM	J	4	5	-	0/3/21/22	0/2/2/2
5	5CM	J	60	5	-	0/3/21/22	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	61	5CM	C6-C5	-2.34	1.33	1.40
5	J	-62	5CM	C6-C5	-2.30	1.34	1.40
5	I	-61	5CM	C6-C5	-2.30	1.34	1.40
5	J	4	5CM	C6-C5	-2.26	1.34	1.40
5	J	60	5CM	C6-C5	-2.22	1.34	1.40

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	J	60	5CM	C4'-O4'-C1'	-2.03	104.45	109.42
5	J	-6	5CM	C5A-C5-C6	2.01	122.69	118.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	I	61	5CM	1	0
5	J	60	5CM	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	99/139 (71%)	0.35	5 (5%) 29 22	49, 71, 100, 145	0
1	E	99/139 (71%)	0.18	1 (1%) 82 79	40, 56, 91, 120	0
2	B	82/106 (77%)	0.54	4 (4%) 30 24	51, 67, 98, 143	0
2	F	85/106 (80%)	0.47	3 (3%) 44 36	40, 55, 80, 113	0
3	C	106/133 (79%)	0.16	0 100 100	36, 62, 91, 129	0
3	G	105/133 (78%)	0.30	5 (4%) 31 24	52, 73, 104, 124	0
4	D	97/129 (75%)	0.39	3 (3%) 49 41	40, 63, 116, 140	0
4	H	94/129 (72%)	0.42	3 (3%) 48 40	50, 71, 103, 172	0
5	I	142/146 (97%)	0.43	11 (7%) 14 10	66, 124, 178, 201	0
5	J	142/146 (97%)	0.46	14 (9%) 8 5	68, 123, 177, 224	0
All	All	1051/1306 (80%)	0.37	49 (4%) 32 25	36, 72, 151, 224	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	26	ARG	5.7
5	I	-48	DC	4.7
5	I	-49	DA	4.3
3	G	15	LYS	4.3
2	B	21	VAL	4.2

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy

less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	5CM	I	-61	20/21	0.65	0.32	-	123,150,166,170	0
5	5CM	J	60	20/21	0.91	0.20	-	102,127,134,135	0
5	5CM	I	61	20/21	0.90	0.19	-	108,122,139,142	0
5	5CM	J	-6	20/21	0.94	0.17	-	76,91,114,119	0
5	5CM	I	5	20/21	0.94	0.16	-	78,98,110,115	0
5	5CM	J	-62	20/21	0.91	0.17	-	83,113,125,126	0
5	5CM	I	-5	20/21	0.94	0.16	-	76,91,108,108	0
5	5CM	J	4	20/21	0.93	0.23	-	96,107,132,133	0

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
6	MN	E	301	1/1	0.87	0.25	3.69	58,58,58,58	0
6	MN	I	101	1/1	0.92	0.18	2.81	113,113,113,113	0
6	MN	J	102	1/1	0.79	0.11	-	114,114,114,114	0
6	MN	I	102	1/1	0.66	0.15	-	134,134,134,134	0
6	MN	I	103	1/1	0.65	0.10	-	115,115,115,115	0
6	MN	J	101	1/1	0.75	0.09	-	140,140,140,140	0
6	MN	J	103	1/1	0.36	0.14	-	138,138,138,138	0

6.5 Other polymers [i](#)

There are no such residues in this entry.