



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2023 – 08:20 AM EDT

PDB ID : 3C1B
Title : The effect of H3 K79 dimethylation and H4 K20 trimethylation on nucleosome and chromatin structure
Authors : Lu, X.; Simon, M.; Chodaparambil, J.; Hansen, J.; Shokat, K.; Luger, K.
Deposited on : 2008-01-22
Resolution : 2.20 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35
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) : 2.35

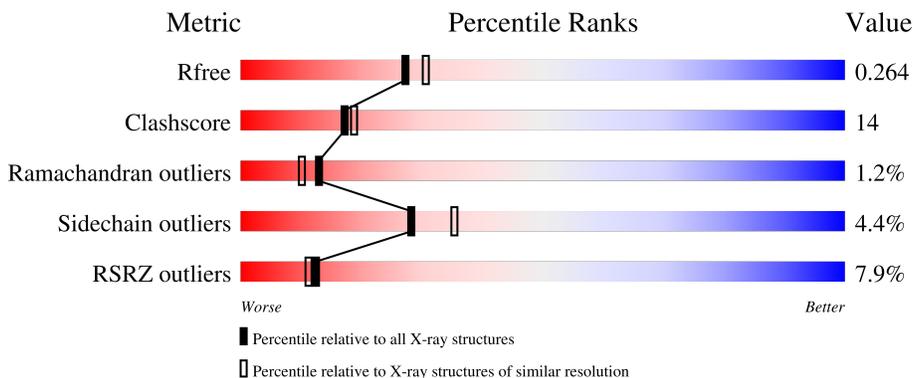
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 ≥ 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	135	 8% 56% 13% 2% 27%
1	E	135	 2% 60% 12% 2% 27%
2	B	102	 2% 63% 15% 2% 22%
2	F	102	 8% 69% 11% 5% 15%
3	C	129	 3% 67% 11% 2% 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	129	<p>1% 61% 16% 20%</p>
4	D	125	<p>4% 64% 10% 24%</p>
4	H	125	<p>2% 60% 13% 25%</p>
5	I	146	<p>21% 51% 49%</p>
5	J	146	<p>18% 55% 45%</p>

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 12828 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-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	98	808	509	156	140	3	0	0	0
1	E	99	817	515	158	141	3	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	421	ALA	VAL	conflict	UNP P02302
A	426	ARG	LYS	conflict	UNP P02302
A	428	SER	CYS	conflict	UNP P02302
A	486	SER	ARG	conflict	UNP P02302
E	621	ALA	VAL	conflict	UNP P02302
E	626	ARG	LYS	conflict	UNP P02302
E	628	SER	CYS	conflict	UNP P02302
E	686	SER	ARG	conflict	UNP P02302

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	80	638	401	125	111	1	0	0	0
2	F	87	706	444	142	118	2	0	0	0

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	106	818	516	160	142	0	0	0
3	G	103	795	501	155	139	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	899	ARG	GLY	conflict	UNP P06897
C	923	SER	ALA	conflict	UNP P06897
C	926	THR	ALA	conflict	UNP P06897
G	1099	ARG	GLY	conflict	UNP P06897
G	1123	SER	ALA	conflict	UNP P06897
G	1126	THR	ALA	conflict	UNP P06897

- Molecule 4 is a protein called Histone 2, H2bf.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	95	Total	C	N	O	S	0	0	0
			744	469	134	139	2			
4	H	94	Total	C	N	O	S	0	0	0
			735	463	132	138	2			

- Molecule 5 is a DNA chain called Palindromic 146bp Human Alpha satellite DNA.

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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	70	Total	O	0	0
			70	70		
6	B	48	Total	O	0	0
			48	48		
6	C	84	Total	O	0	0
			84	84		
6	D	60	Total	O	0	0
			60	60		
6	E	87	Total	O	0	0
			87	87		
6	F	61	Total	O	0	0
			61	61		
6	G	69	Total	O	0	0
			69	69		
6	H	43	Total	O	0	0
			43	43		

Continued on next page...

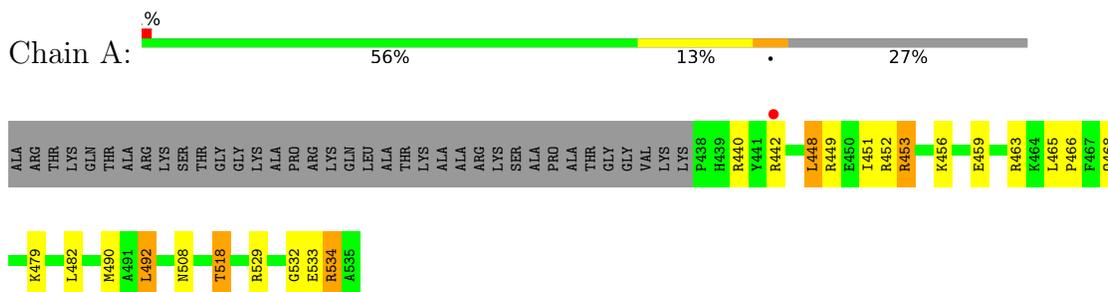
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	143	Total 143	O 143	0	0
6	J	122	Total 122	O 122	0	0

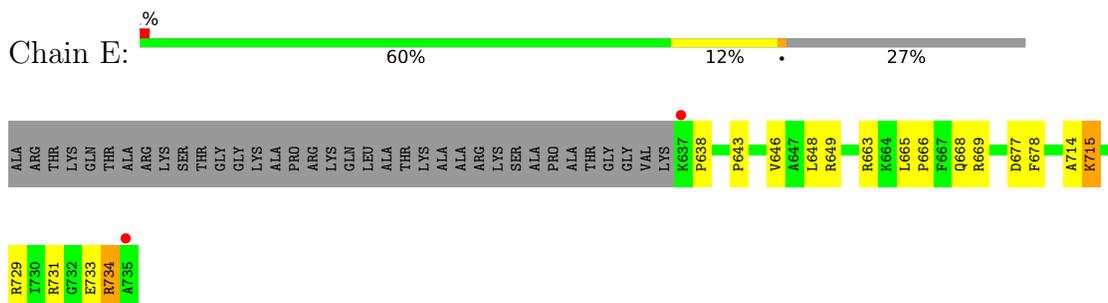
3 Residue-property plots [i](#)

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.

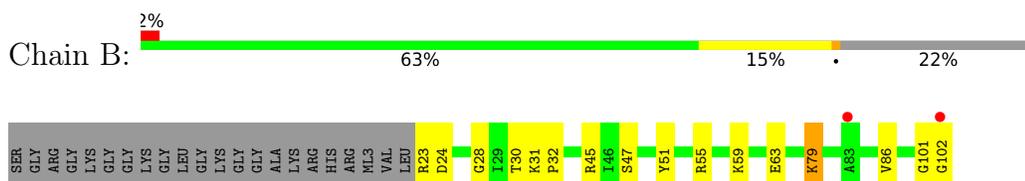
- Molecule 1: Histone H3-like



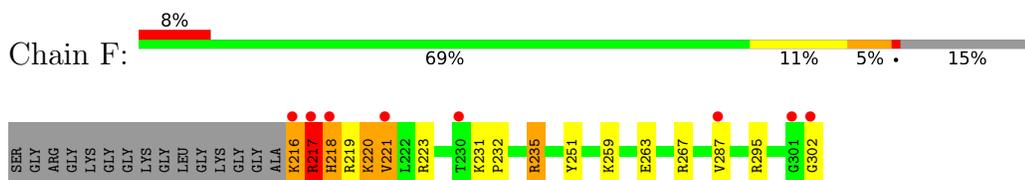
- Molecule 1: Histone H3-like



- Molecule 2: Histone H4

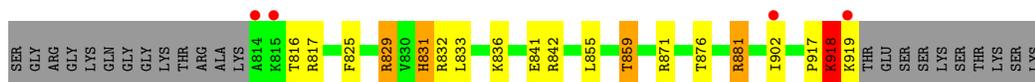


- Molecule 2: Histone H4



- Molecule 3: Histone H2A type 1





• Molecule 3: Histone H2A type 1



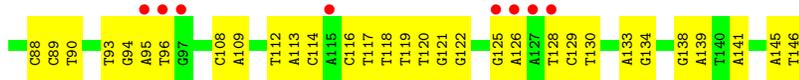
• Molecule 4: Histone 2, H2bf



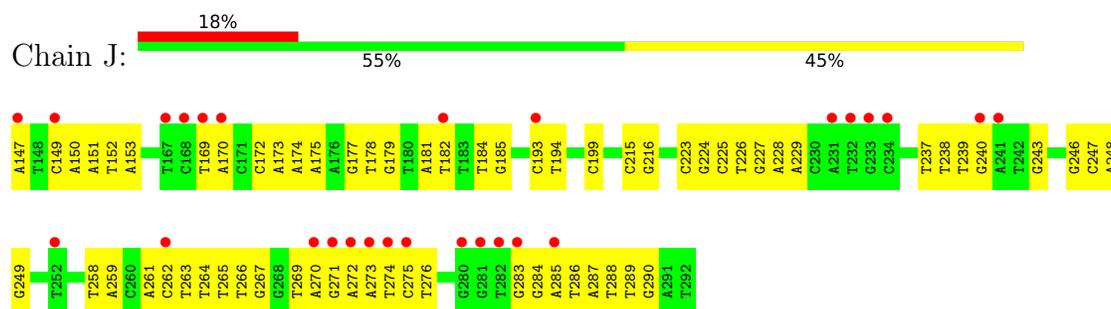
• Molecule 4: Histone 2, H2bf



• Molecule 5: Palindromic 146bp Human Alpha satellite DNA



• Molecule 5: Palindromic 146bp Human Alpha satellite DNA



4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.67Å 109.83Å 181.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.00 – 2.20 42.91 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (43.00-2.20) 96.5 (42.91-2.20)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.99 (at 2.20Å)	Xtrriage
Refinement program	CNS	Depositor
R, R_{free}	0.240 , 0.270 0.230 , 0.264	Depositor DCC
R_{free} test set	5365 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	42.3	Xtrriage
Anisotropy	0.297	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 66.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.009 for k,h,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	12828	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% 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

Bond lengths and bond angles in the following residue types are not validated in this section: ML3

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	0.36	0/820	0.63	0/1099
1	E	0.44	0/829	0.68	0/1111
2	B	0.37	0/645	0.62	0/862
2	F	0.50	0/701	0.79	2/934 (0.2%)
3	C	0.40	0/828	0.63	0/1117
3	G	0.32	0/805	0.55	0/1088
4	D	0.38	0/755	0.67	1/1015 (0.1%)
4	H	0.36	0/746	0.64	1/1004 (0.1%)
5	I	0.27	0/3354	0.68	0/5175
5	J	0.29	0/3354	0.68	0/5175
All	All	0.34	0/12837	0.67	4/18580 (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
5	I	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1500	PRO	CA-N-CD	-8.90	99.04	111.50
4	D	1300	PRO	CA-N-CD	-8.11	100.15	111.50
2	F	217	ARG	N-CA-C	6.31	128.03	111.00
2	F	216	LYS	C-N-CA	-6.08	106.50	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	I	88	DC	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	A	808	0	846	37	0
1	E	817	0	858	20	0
2	B	638	0	676	14	0
2	F	706	0	758	26	0
3	C	818	0	877	22	0
3	G	795	0	846	35	0
4	D	744	0	773	11	0
4	H	735	0	760	25	0
5	I	2990	0	1651	80	0
5	J	2990	0	1651	85	0
6	A	70	0	0	2	0
6	B	48	0	0	0	0
6	C	84	0	0	4	0
6	D	60	0	0	3	0
6	E	87	0	0	3	0
6	F	61	0	0	3	0
6	G	69	0	0	7	0
6	H	43	0	0	3	0
6	I	143	0	0	8	0
6	J	122	0	0	8	0
All	All	12828	0	9696	301	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 14.

The worst 5 of 301 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)
2:F:218:HIS:CE1	5:J:199:DC:OP1	1.92	1.21

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:216:LYS:O	2:F:217:ARG:CB	1.90	1.14
2:F:216:LYS:O	2:F:217:ARG:CG	1.99	1.10
2:F:216:LYS:O	2:F:217:ARG:HB2	1.55	1.07
2:F:216:LYS:O	2:F:217:ARG:HG2	1.64	0.97

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%)	93 (97%)	2 (2%)	1 (1%)	15	14
1	E	97/135 (72%)	95 (98%)	1 (1%)	1 (1%)	15	14
2	B	78/102 (76%)	75 (96%)	3 (4%)	0	100	100
2	F	84/102 (82%)	82 (98%)	0	2 (2%)	6	3
3	C	104/129 (81%)	101 (97%)	2 (2%)	1 (1%)	15	14
3	G	101/129 (78%)	98 (97%)	2 (2%)	1 (1%)	15	14
4	D	93/125 (74%)	90 (97%)	2 (2%)	1 (1%)	14	12
4	H	92/125 (74%)	88 (96%)	2 (2%)	2 (2%)	6	4
All	All	745/982 (76%)	722 (97%)	14 (2%)	9 (1%)	13	10

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	534	ARG
4	D	1301	GLY
2	F	217	ARG
3	G	1117	PRO
4	H	1501	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	85/110 (77%)	80 (94%)	5 (6%)	19	23
1	E	86/110 (78%)	84 (98%)	2 (2%)	50	63
2	B	65/77 (84%)	62 (95%)	3 (5%)	27	34
2	F	71/77 (92%)	67 (94%)	4 (6%)	21	25
3	C	84/102 (82%)	78 (93%)	6 (7%)	14	16
3	G	82/102 (80%)	80 (98%)	2 (2%)	49	62
4	D	81/104 (78%)	78 (96%)	3 (4%)	34	43
4	H	80/104 (77%)	77 (96%)	3 (4%)	33	42
All	All	634/786 (81%)	606 (96%)	28 (4%)	28	35

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	1230	ARG
4	H	1500	PRO
1	E	715	LYS
3	G	1088	ARG
4	D	1300	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	1306	HIS
2	F	218	HIS
3	G	1073	ASN
3	G	1031	HIS
4	D	1279	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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
2	ML3	F	220	2	10,11,12	3.36	2 (20%)	10,14,16	1.08	1 (10%)

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
2	ML3	F	220	2	-	2/8/10/12	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	220	ML3	CD-SG	-7.70	1.52	1.81
2	F	220	ML3	CB-SG	-7.07	1.52	1.80

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	220	ML3	CB-SG-CD	3.00	111.27	102.27

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	220	ML3	SG-CD-CE-NZ
2	F	220	ML3	N-CA-CB-SG

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	220	ML3	2	0

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 [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	98/135 (72%)	0.34	1 (1%) 82 81	29, 41, 79, 106	0
1	E	99/135 (73%)	0.49	2 (2%) 65 63	22, 34, 61, 107	0
2	B	80/102 (78%)	0.44	2 (2%) 57 55	31, 41, 61, 117	0
2	F	86/102 (84%)	0.77	8 (9%) 8 7	25, 34, 63, 139	0
3	C	106/129 (82%)	0.41	4 (3%) 40 38	24, 36, 70, 142	0
3	G	103/129 (79%)	0.32	1 (0%) 82 81	31, 45, 75, 123	0
4	D	95/125 (76%)	0.58	5 (5%) 26 25	29, 38, 73, 96	0
4	H	94/125 (75%)	0.58	2 (2%) 63 61	32, 45, 79, 120	0
5	I	146/146 (100%)	1.02	31 (21%) 0 0	39, 88, 135, 174	0
5	J	146/146 (100%)	1.12	27 (18%) 1 1	41, 90, 142, 177	0
All	All	1053/1274 (82%)	0.65	83 (7%) 12 11	22, 45, 119, 177	0

The worst 5 of 83 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	J	272	DA	7.0
5	J	233	DG	6.5
5	J	273	DA	6.2
2	B	102	GLY	5.6
5	I	56	DA	5.5

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. 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	B-factors(\AA^2)	Q<0.9
2	ML3	F	220	12/13	0.78	0.25	58,139,139,139	0

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.