



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 07:10 AM UTC

PDB ID : 5W2M / pdb_00005w2m
Title : APOBEC3F Catalytic Domain Complex with a Single-Stranded DNA
Authors : Fang, Y.; Xiao, X.; Li, S.-X.; Wolfe, A.; Chen, X.S.
Deposited on : 2017-06-06
Resolution : 3.70 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

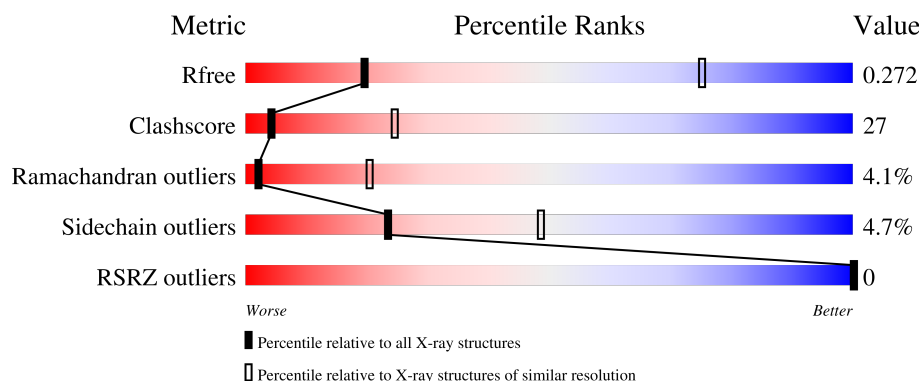
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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}	180053	1131 (3.80-3.60)
Clashscore	190562	1171 (3.80-3.60)
Ramachandran outliers	187476	1129 (3.80-3.60)
Sidechain outliers	187428	1126 (3.80-3.60)
RSRZ outliers	180081	1130 (3.80-3.60)

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 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	184	54% 41% 5%
1	B	184	53% 42% 5%
1	C	184	49% 43% 7% .
1	D	184	47% 45% 8% .
1	J	184	53% 42% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	184	54%41%5%
1	L	184	47%46%7%
1	M	184	47%46%7%
2	E	10	10%80%10%
2	N	10	10%90%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 12926 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 DNA dC->dU-editing enzyme APOBEC-3F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	B	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	C	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	D	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	J	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	K	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	L	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			
1	M	184	Total	C	N	O	S	0	0	0
			1565	1017	255	282	11			

- Molecule 2 is a DNA chain called DNA (5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	10	Total	C	N	O	P	0	0	0
			201	100	20	71	10			
2	N	10	Total	C	N	O	P	0	0	0
			201	100	20	71	10			

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		

Continued on next page...

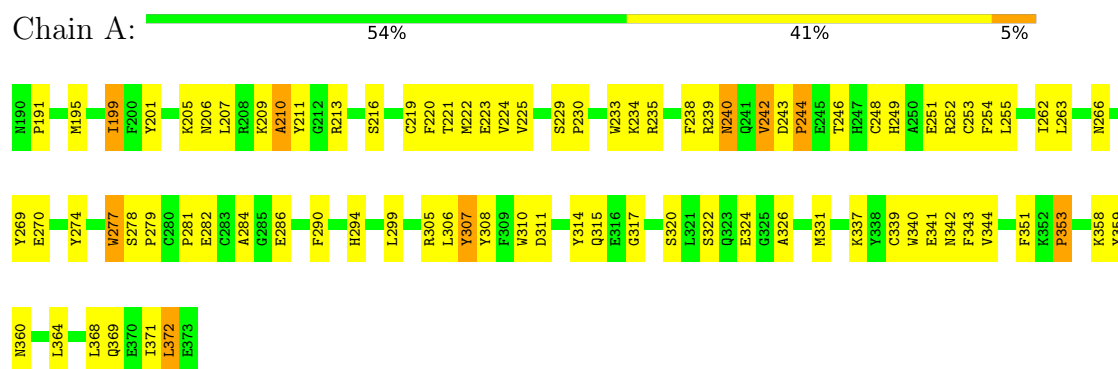
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	J	1	Total 1	Zn 1	0	0
3	K	1	Total 1	Zn 1	0	0

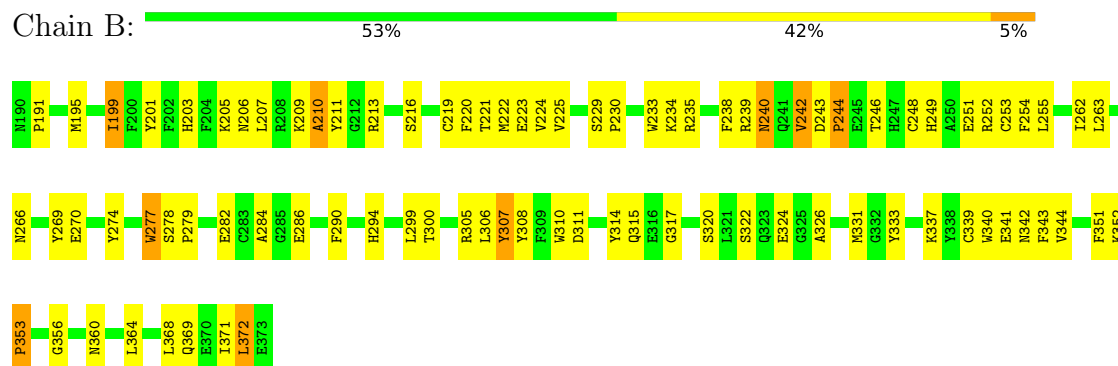
3 Residue-property plots

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 dC->dU-editing enzyme APOBEC-3F



- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F



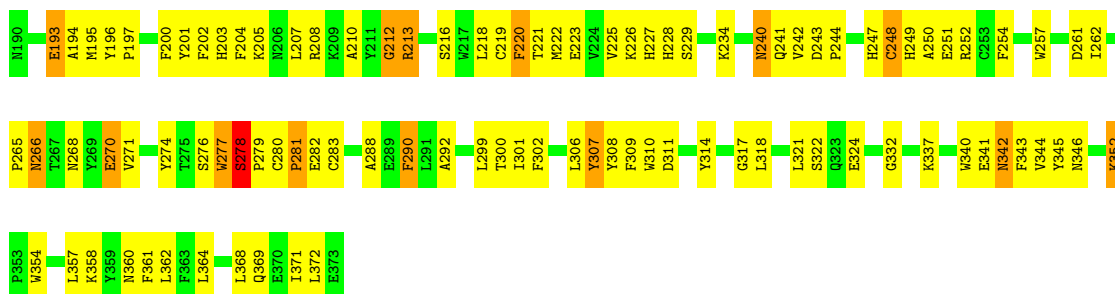
- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F





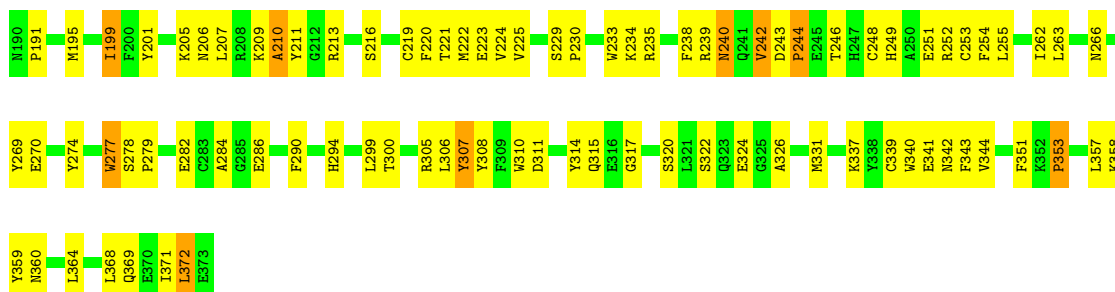
- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F

Chain D: 47% 45% 8%



- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F

Chain J: 53% 42% 5%



- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F

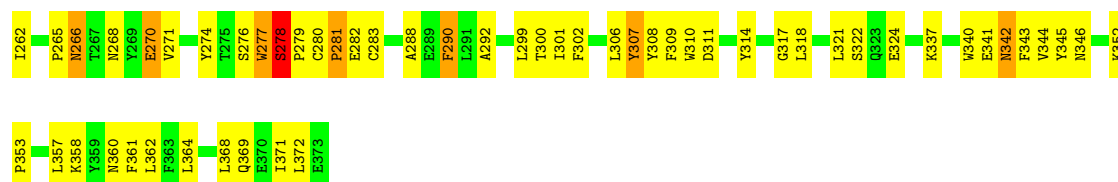
Chain K: 54% 41% 5%



- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F

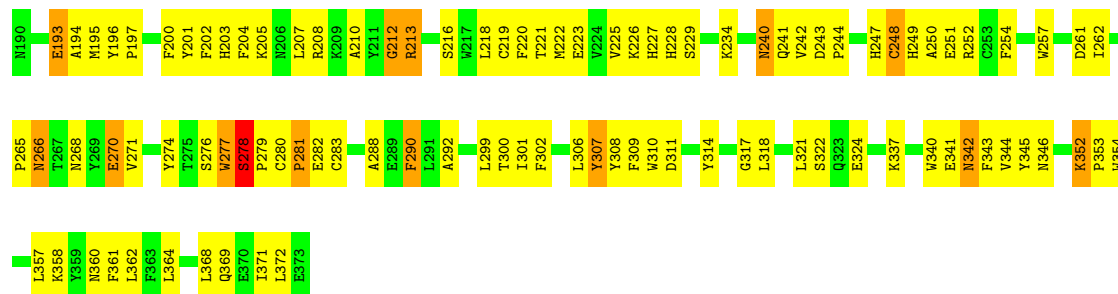
Chain L: 47% 46% 7%





- Molecule 1: DNA dC->dU-editing enzyme APOBEC-3F

Chain M: 47% 46% 7% .



- Molecule 2: DNA (5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*TP*T)-3')

Chain E: 10% 80% 10%



- Molecule 2: DNA (5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*TP*T)-3')

Chain N: 10% 90%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	68.19Å 108.49Å 108.55Å 79.60° 71.80° 71.75°	Depositor
Resolution (Å)	50.00 – 3.70 50.00 – 3.70	Depositor EDS
% Data completeness (in resolution range)	96.4 (50.00-3.70) 95.7 (50.00-3.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 3.67Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.250 , 0.267 0.258 , 0.272	Depositor DCC
R_{free} test set	2866 reflections (9.84%)	wwPDB-VP
Wilson B-factor (Å ²)	87.7	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 55.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.327 for h,h-k,h-l 0.307 for -h,-l,-k 0.317 for -h,-h+l,-h+k	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12926	wwPDB-VP
Average B, all atoms (Å ²)	107.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.56% 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: ZN

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	0.39	0/1623	0.80	1/2204 (0.0%)
1	B	0.39	0/1623	0.80	1/2204 (0.0%)
1	C	0.41	0/1623	0.84	5/2204 (0.2%)
1	D	0.41	0/1623	0.85	8/2204 (0.4%)
1	J	0.38	0/1623	0.80	1/2204 (0.0%)
1	K	0.39	0/1623	0.80	1/2204 (0.0%)
1	L	0.41	0/1623	0.85	6/2204 (0.3%)
1	M	0.40	0/1623	0.85	6/2204 (0.3%)
2	E	0.41	0/220	0.91	1/336 (0.3%)
2	N	0.41	0/220	0.91	0/336
All	All	0.40	0/13424	0.83	30/18304 (0.2%)

There are no bond length outliers.

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	239	ARG	N-CA-C	-6.66	100.66	110.52
1	J	239	ARG	N-CA-C	-6.63	100.71	110.52
1	B	239	ARG	N-CA-C	-6.33	101.15	110.52
1	A	239	ARG	N-CA-C	-6.28	101.23	110.52
1	M	290	PHE	N-CA-C	-5.69	103.85	112.04

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	1565	0	1443	77	0
1	B	1565	0	1443	81	0
1	C	1565	0	1443	82	0
1	D	1565	0	1443	89	0
1	J	1565	0	1443	75	0
1	K	1565	0	1443	81	0
1	L	1565	0	1443	89	0
1	M	1565	0	1443	91	0
2	E	201	0	121	21	0
2	N	201	0	121	22	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	J	1	0	0	0	0
3	K	1	0	0	0	0
All	All	12926	0	11786	670	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 27.

The worst 5 of 670 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:K:333:TYR:OH	2:N:908:DT:H72	1.70	0.90
1:K:351:PHE:HB3	2:N:908:DT:H71	1.56	0.87
1:B:207:LEU:HB2	1:B:210:ALA:HB2	1.57	0.87
1:K:207:LEU:HB2	1:K:210:ALA:HB2	1.57	0.87
1:A:207:LEU:HB2	1:A:210:ALA:HB2	1.58	0.85

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	182/184 (99%)	141 (78%)	33 (18%)	8 (4%)	2	19
1	B	182/184 (99%)	141 (78%)	33 (18%)	8 (4%)	2	19
1	C	182/184 (99%)	142 (78%)	33 (18%)	7 (4%)	2	22
1	D	182/184 (99%)	142 (78%)	33 (18%)	7 (4%)	2	22
1	J	182/184 (99%)	141 (78%)	33 (18%)	8 (4%)	2	19
1	K	182/184 (99%)	141 (78%)	33 (18%)	8 (4%)	2	19
1	L	182/184 (99%)	142 (78%)	33 (18%)	7 (4%)	2	22
1	M	182/184 (99%)	142 (78%)	33 (18%)	7 (4%)	2	22
All	All	1456/1472 (99%)	1132 (78%)	264 (18%)	60 (4%)	2	20

5 of 60 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	210	ALA
1	A	353	PRO
1	B	210	ALA
1	B	353	PRO
1	C	208	ARG

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	169/169 (100%)	163 (96%)	6 (4%)	31	54
1	B	169/169 (100%)	163 (96%)	6 (4%)	31	54
1	C	169/169 (100%)	159 (94%)	10 (6%)	18	44
1	D	169/169 (100%)	159 (94%)	10 (6%)	18	44
1	J	169/169 (100%)	163 (96%)	6 (4%)	31	54
1	K	169/169 (100%)	163 (96%)	6 (4%)	31	54
1	L	169/169 (100%)	159 (94%)	10 (6%)	18	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	M	169/169 (100%)	159 (94%)	10 (6%)	18	44
All	All	1352/1352 (100%)	1288 (95%)	64 (5%)	23	48

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	240	ASN
1	M	270	GLU
1	D	213	ARG
1	D	197	PRO
1	M	277	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 73 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	266	ASN
1	M	360	ASN
1	L	315	GLN
1	M	240	ASN
1	C	342	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 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	184/184 (100%)	-1.15	0 100 100	80, 103, 141, 157	0
1	B	184/184 (100%)	-1.11	0 100 100	79, 104, 142, 157	0
1	C	184/184 (100%)	-1.09	0 100 100	81, 105, 139, 155	0
1	D	184/184 (100%)	-1.09	0 100 100	77, 105, 140, 154	0
1	J	184/184 (100%)	-1.12	0 100 100	80, 103, 143, 157	0
1	K	184/184 (100%)	-1.17	0 100 100	78, 104, 143, 157	0
1	L	184/184 (100%)	-1.07	0 100 100	79, 105, 139, 154	0
1	M	184/184 (100%)	-1.15	0 100 100	79, 105, 139, 155	0
2	E	10/10 (100%)	-0.57	0 100 100	82, 100, 156, 177	0
2	N	10/10 (100%)	-0.61	0 100 100	83, 102, 156, 177	0
All	All	1492/1492 (100%)	-1.11	0 100 100	77, 105, 143, 177	0

There are no RSRZ outliers to report.

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 oligosaccharides 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. 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(Å ²)	Q<0.9
3	ZN	A	801	1/1	0.98	0.08	225,225,225,225	0
3	ZN	B	801	1/1	0.99	0.04	221,221,221,221	0
3	ZN	J	801	1/1	0.99	0.03	225,225,225,225	0
3	ZN	K	801	1/1	0.99	0.07	221,221,221,221	0

6.5 Other polymers ⓘ

There are no such residues in this entry.