



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 27, 2026 – 10:10 AM EDT

PDB ID : 9PVU / pdb_00009pvu
Title : RNA polymerase II elongation complex with dC at +1 site, 8-oxo-GMP added.
Authors : Hou, P.; Oh, J.; Wang, D.
Deposited on : 2025-08-03
Resolution : 3.31 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
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.015 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

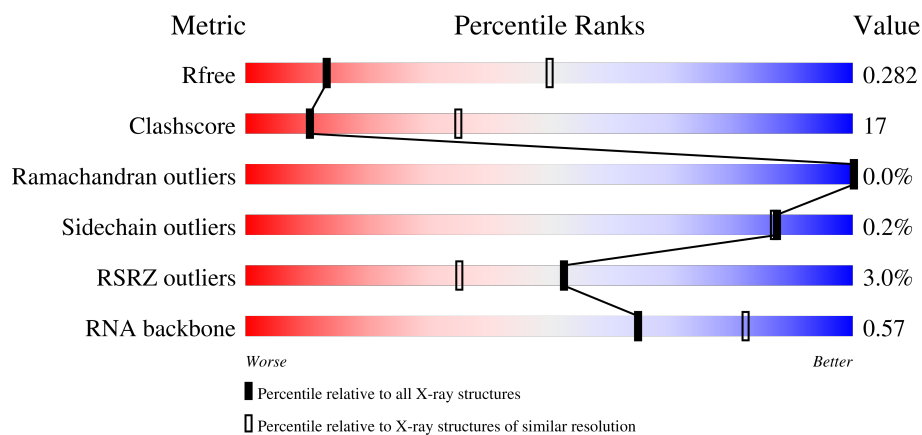
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.31 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	1153 (3.34-3.30)
Clashscore	190562	1193 (3.34-3.30)
Ramachandran outliers	187476	1172 (3.34-3.30)
Sidechain outliers	187428	1171 (3.34-3.30)
RSRZ outliers	180081	1153 (3.34-3.30)
RNA backbone	3983	1067 (3.64-3.00)

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	R	10	30% 60% 10%
2	T	29	45% 41% 14%
3	N	18	33% 44% 22%
4	A	1733	3% 51% 28% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	3% 58% 34% 8%
6	C	318	% 61% 23% 16%
7	E	215	4% 61% 37% .
8	F	155	% 37% 18% 45%
9	H	146	% 53% 38% 9%
10	I	122	2% 66% 30% ..
11	J	70	% 51% 40% 7%
12	K	120	4% 67% 28% 5%
13	L	70	7% 44% 17% 39%

2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 29028 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 RNA chain called RNA with 8-oxo-guanine added to 3' end.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			219	98	45	67	9			

- Molecule 2 is a DNA chain called Template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	25	Total	C	N	O	P	0	0	0
			497	239	76	157	25			

- Molecule 3 is a DNA chain called Non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	14	Total	C	N	O	P	0	0	0
			296	138	66	78	14			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1384	Total	C	N	O	S	0	0	0
			10828	6831	1896	2041	60			

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1123	Total	C	N	O	S	0	0	0
			8859	5607	1552	1647	53			

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	212	Total	C	N	O	S	0	0	0
			1731	1100	305	315	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	86	Total	C	N	O	S	0	0	0
			684	437	115	129	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total	C	N	O	S	0	0	0
			1064	670	179	211	4			

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	118	Total	C	N	O	S	0	0	0
			952	585	173	184	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	43	Total	C	N	O	S	0	0	0
			337	208	66	59	4			

- Molecule 14 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total 2	Zn 2	0	0
14	B	1	Total 1	Zn 1	0	0
14	C	1	Total 1	Zn 1	0	0
14	I	2	Total 2	Zn 2	0	0
14	J	1	Total 1	Zn 1	0	0
14	L	1	Total 1	Zn 1	0	0

- Molecule 15 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	1	Total 1	Mg 1	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: RNA with 8-oxo-guanine added to 3' end



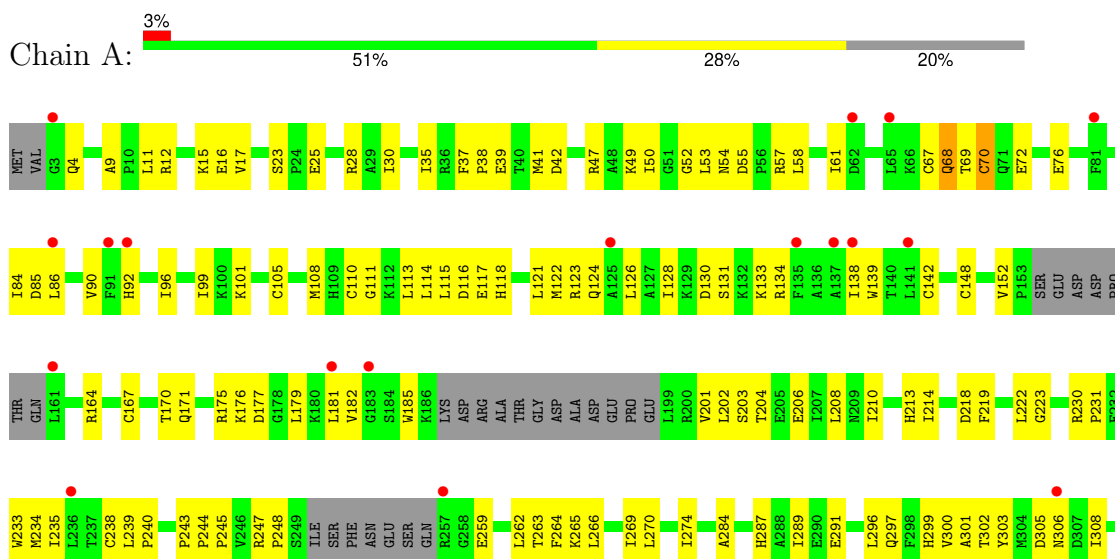
- Molecule 2: Template strand DNA



- Molecule 3: Non-template strand DNA



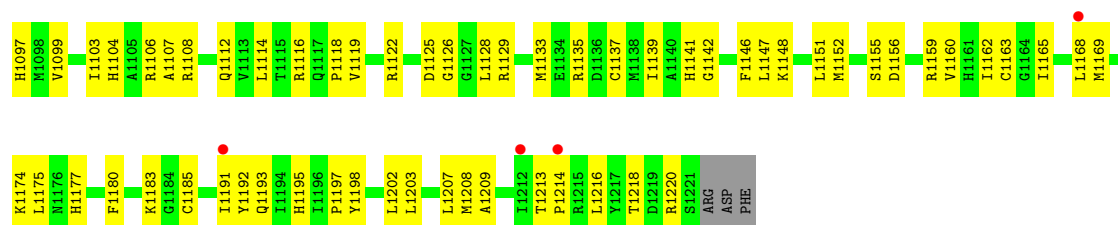
- Molecule 4: DNA-directed RNA polymerase II subunit RPB1



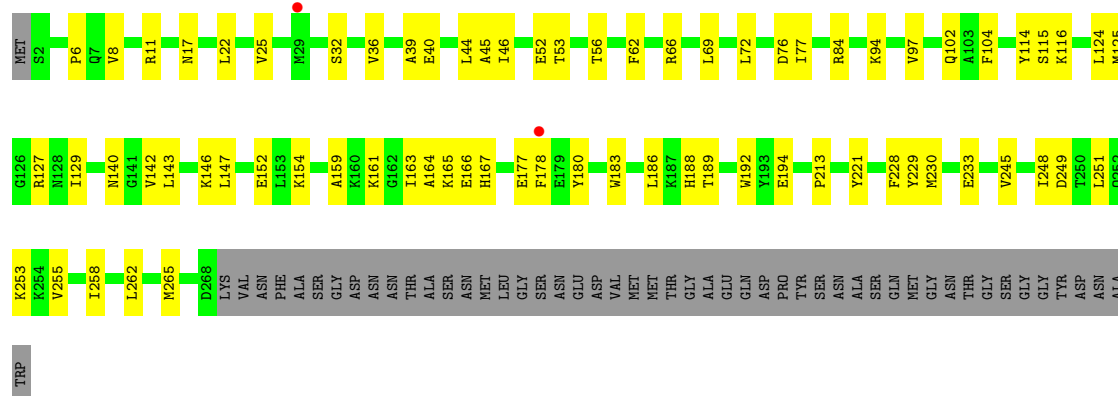


- Molecule 5: DNA-directed RNA polymerase II subunit RPB2





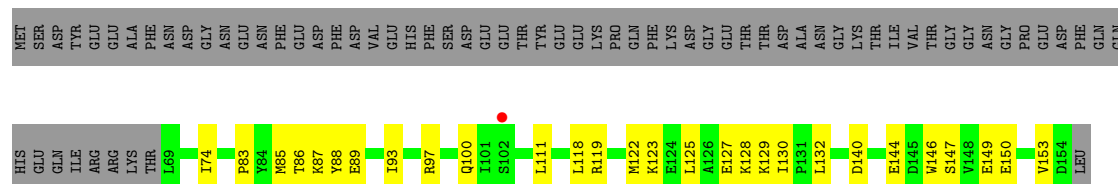
• Molecule 6: DNA-directed RNA polymerase II subunit RPB3



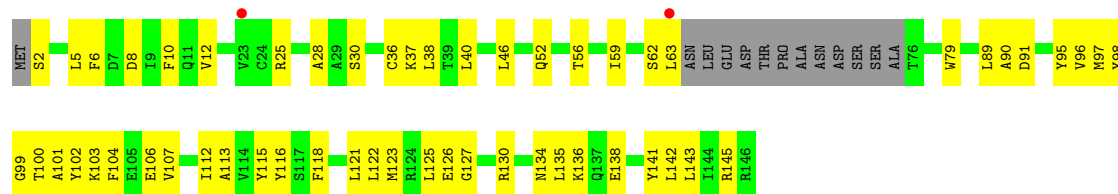
• Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1



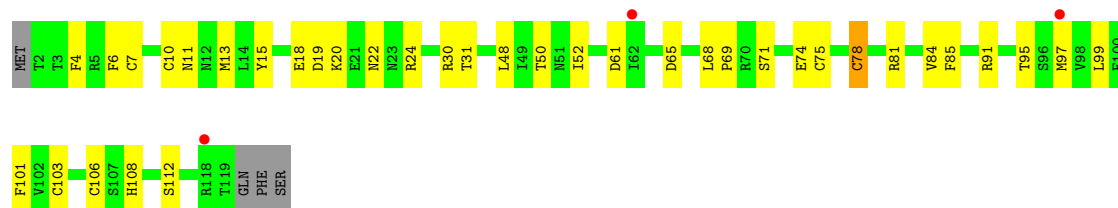
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2



• Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC3



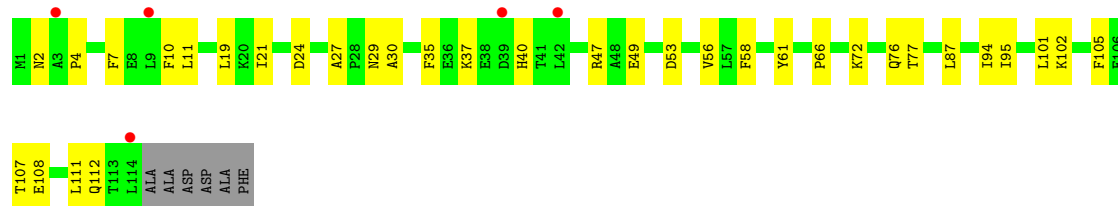
• Molecule 10: DNA-directed RNA polymerase II subunit RPB9



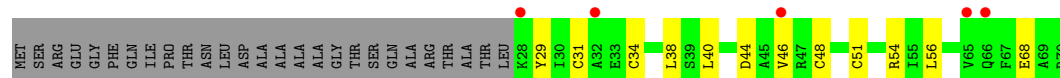
• Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC5



• Molecule 12: DNA-directed RNA polymerase II subunit RPB11



• Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC4



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	166.00Å 223.51Å 192.78Å 90.00° 100.26° 90.00°	Depositor
Resolution (Å)	49.32 – 3.31 49.32 – 3.31	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.32-3.31) 99.6 (49.32-3.31)	Depositor EDS
R_{merge}	0.48	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.10 (at 3.33Å)	Xtriage
Refinement program	PHENIX (1.21.2_5419: ???)	Depositor
R, R_{free}	0.243 , 0.281 0.244 , 0.282	Depositor DCC
R_{free} test set	2000 reflections (1.95%)	wwPDB-VP
Wilson B-factor (Å ²)	91.6	Xtriage
Anisotropy	0.407	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 104.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	29028	wwPDB-VP
Average B, all atoms (Å ²)	111.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% 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: MG, 8GM, 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	R	0.09	0/219	0.26	0/341
2	T	0.21	0/551	0.43	0/845
3	N	0.19	0/335	0.36	0/516
4	A	0.16	0/11020	0.35	1/14907 (0.0%)
5	B	0.14	0/9030	0.33	0/12186
6	C	0.12	0/2139	0.31	0/2899
7	E	0.15	0/1767	0.33	0/2378
8	F	0.16	0/696	0.32	0/943
9	H	0.16	0/1082	0.42	0/1466
10	I	0.15	0/970	0.34	0/1308
11	J	0.15	0/541	0.35	0/727
12	K	0.13	0/937	0.32	0/1265
13	L	0.26	0/339	0.41	0/450
All	All	0.15	0/29626	0.34	1/40231 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	A	525	GLN	CB-CA-C	-5.51	110.24	116.63

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	R	219	0	99	9	0
2	T	497	0	285	14	0
3	N	296	0	155	11	0
4	A	10828	0	10875	424	0
5	B	8859	0	8816	345	0
6	C	2101	0	2056	60	0
7	E	1731	0	1758	58	0
8	F	684	0	692	20	0
9	H	1064	0	1029	52	0
10	I	952	0	897	35	0
11	J	532	0	542	21	0
12	K	919	0	929	30	0
13	L	337	0	352	14	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	1	0	0	0	0
All	All	29028	0	28485	967	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 17.

The worst 5 of 967 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)
4:A:392:VAL:HG11	4:A:424:ILE:CD1	1.70	1.21
4:A:392:VAL:CG1	4:A:424:ILE:HD11	1.79	1.11
4:A:666:ILE:HD13	5:B:1030:LEU:HD11	1.39	1.01
4:A:613:ILE:HG21	9:H:102:TYR:HB3	1.51	0.93
4:A:392:VAL:HG11	4:A:424:ILE:HD11	1.39	0.92

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	
4	A	1370/1733 (79%)	1342 (98%)	27 (2%)	1 (0%)	48	76
5	B	1103/1224 (90%)	1092 (99%)	11 (1%)	0	100	100
6	C	265/318 (83%)	264 (100%)	1 (0%)	0	100	100
7	E	210/215 (98%)	206 (98%)	4 (2%)	0	100	100
8	F	84/155 (54%)	82 (98%)	2 (2%)	0	100	100
9	H	129/146 (88%)	127 (98%)	2 (2%)	0	100	100
10	I	116/122 (95%)	114 (98%)	2 (2%)	0	100	100
11	J	63/70 (90%)	63 (100%)	0	0	100	100
12	K	112/120 (93%)	109 (97%)	3 (3%)	0	100	100
13	L	41/70 (59%)	40 (98%)	1 (2%)	0	100	100
All	All	3493/4173 (84%)	3439 (98%)	53 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	55	ASP

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	
4	A	1194/1520 (79%)	1191 (100%)	3 (0%)	86	85
5	B	955/1061 (90%)	954 (100%)	1 (0%)	88	89

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	C	235/274 (86%)	235 (100%)	0	100	100
7	E	193/197 (98%)	193 (100%)	0	100	100
8	F	73/137 (53%)	73 (100%)	0	100	100
9	H	116/128 (91%)	116 (100%)	0	100	100
10	I	110/116 (95%)	108 (98%)	2 (2%)	51	69
11	J	60/65 (92%)	59 (98%)	1 (2%)	53	70
12	K	99/102 (97%)	99 (100%)	0	100	100
13	L	37/57 (65%)	37 (100%)	0	100	100
All	All	3072/3657 (84%)	3065 (100%)	7 (0%)	87	87

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	1163	CYS
10	I	7	CYS
11	J	7	CYS
10	I	78	CYS
4	A	110	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	1427	ASN
6	C	54	ASN
5	B	325	GLN
11	J	64	ASN
5	B	1076	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/10 (80%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	9	G

There are no RNA pucker outliers to report.

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 9 ligands modelled in this entry, 9 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	R	9/10 (90%)	-0.15	0	100	100	88, 112, 155, 172	0
2	T	25/29 (86%)	0.38	0	100	100	101, 138, 217, 221	0
3	N	14/18 (77%)	0.37	0	100	100	159, 191, 249, 252	0
4	A	1384/1733 (79%)	0.34	47 (3%)	48	32	49, 105, 187, 260	0
5	B	1123/1224 (91%)	0.34	34 (3%)	52	35	49, 90, 160, 237	0
6	C	267/318 (83%)	0.12	2 (0%)	84	70	56, 91, 142, 187	0
7	E	212/215 (98%)	0.39	9 (4%)	40	26	64, 127, 199, 244	0
8	F	86/155 (55%)	0.13	1 (1%)	76	58	74, 101, 150, 174	0
9	H	133/146 (91%)	0.45	2 (1%)	72	53	93, 130, 179, 239	0
10	I	118/122 (96%)	0.30	3 (2%)	58	39	70, 114, 157, 195	0
11	J	65/70 (92%)	0.22	1 (1%)	72	53	60, 81, 121, 180	0
12	K	114/120 (95%)	0.16	5 (4%)	39	25	56, 97, 137, 154	0
13	L	43/70 (61%)	0.79	5 (11%)	9	8	72, 146, 237, 276	0
All	All	3593/4230 (84%)	0.32	109 (3%)	52	35	49, 102, 183, 276	0

The worst 5 of 109 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
7	E	112	TYR	5.7
4	A	1332	PHE	4.8
4	A	998	LEU	4.3
5	B	1214	PRO	4.3
4	A	666	ILE	4.1

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($\text{\AA}^2$)	Q<0.9
14	ZN	A	1801	1/1	0.77	0.10	321,321,321,321	0
14	ZN	L	101	1/1	0.83	0.12	363,363,363,363	0
15	MG	A	1803	1/1	0.87	0.17	182,182,182,182	0
14	ZN	B	1301	1/1	0.96	0.05	200,200,200,200	0
14	ZN	I	202	1/1	0.98	0.04	127,127,127,127	0
14	ZN	A	1802	1/1	0.99	0.03	145,145,145,145	0
14	ZN	C	401	1/1	0.99	0.06	114,114,114,114	0
14	ZN	I	201	1/1	0.99	0.04	125,125,125,125	0
14	ZN	J	101	1/1	1.00	0.04	86,86,86,86	0

6.5 Other polymers [i](#)

There are no such residues in this entry.