



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 3, 2024 – 07:24 AM EST

PDB ID : 6BLO
Title : Pol II elongation complex with an abasic lesion at i+1 position
Authors : Wang, W.; Wang, D.
Deposited on : 2017-11-10
Resolution : 3.40 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 : 2.36
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) : 2.36

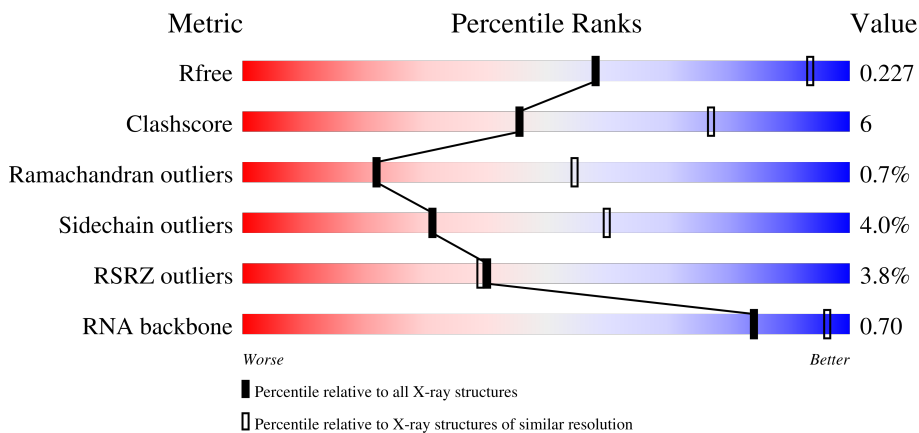
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)
RNA backbone	3102	1006 (3.84-2.96)

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	1733	 4% 62% 16% • 21%
2	B	1224	 1% 71% 17% • 10%
3	C	318	 65% 17% • 16%
4	E	215	 12% 87% 11% ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	F	155	<p>3% 43% 10% 46%</p>
6	H	146	<p>14% 71% 16% 11%</p>
7	I	122	<p>4% 83% 11% 6%</p>
8	J	70	<p>74% 16% 7%</p>
9	K	120	<p>82% 12% 5%</p>
10	L	70	<p>9% 50% 13% 37%</p>
11	T	29	<p>3% 21% 14% 62%</p>
12	R	8	<p>88% 12%</p>

2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 28206 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-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1372	10784	6802	1887	2034	61	0	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1097	8726	5526	1530	1615	55	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	266	2095	1317	348	417	13	0	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	213	1744	1107	308	318	11	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	F	84	679	434	115	127	3	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	H	130	1043	660	173	206	4	0	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	I	115	935	575	170	180	10	0	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	L	44	351	217	70	60	4	0	0	0

- Molecule 11 is a DNA chain called DNA (5'-D(P*AP*(3DR)P*CP*TP*CP*TP*CP*GP*A P*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
11	T	11	214	102	35	66	11	0	0	0

- Molecule 12 is a RNA chain called RNA (5'-R(P*AP*UP*CP*GP*AP*GP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
12	R	8	175	78	35	54	8	0	0	0

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	2	Total	Zn	0	0
			2	2		

Continued on next page...

Continued from previous page...

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

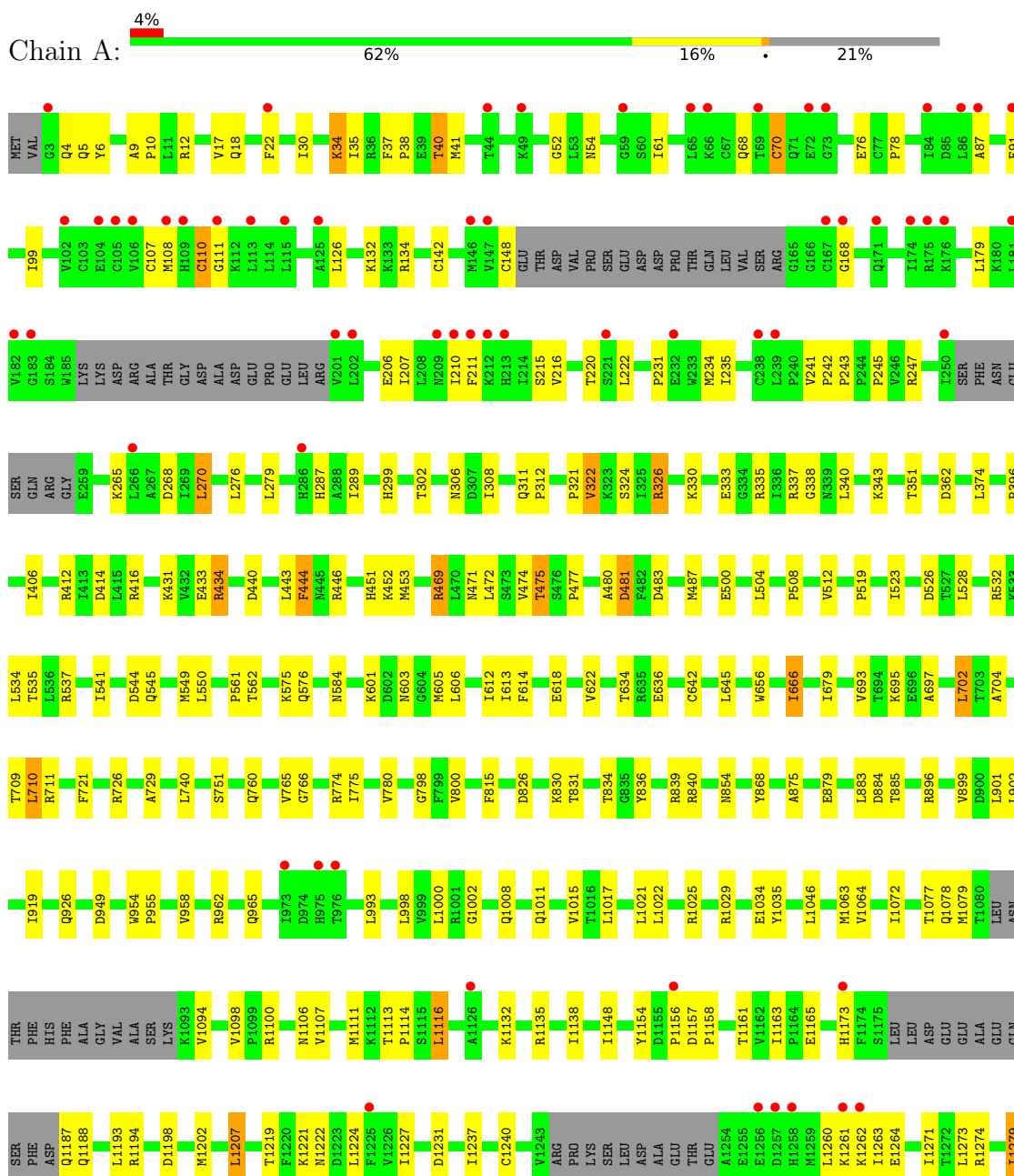
- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

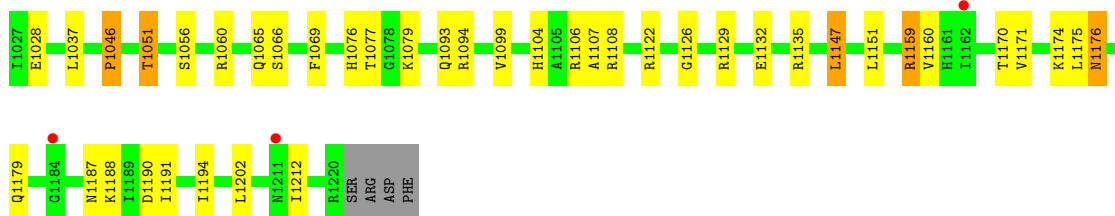
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total 1	Mg 1	0	0

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-directed RNA polymerase II subunit RPB1





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	168.55Å 222.28Å 193.77Å 90.00° 100.62° 90.00°	Depositor
Resolution (Å)	63.48 – 3.40 95.99 – 3.40	Depositor EDS
% Data completeness (in resolution range)	86.6 (63.48-3.40) 86.7 (95.99-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.63 (at 3.41Å)	Xtrriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, R_{free}	0.203 , 0.227 0.204 , 0.227	Depositor DCC
R_{free} test set	4035 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	68.7	Xtrriage
Anisotropy	0.022	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 50.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	28206	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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: MG, 3DR, 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	0.22	0/10975	0.38	0/14838
2	B	0.22	0/8896	0.39	0/11996
3	C	0.22	0/2133	0.38	0/2891
4	E	0.21	0/1780	0.36	0/2395
5	F	0.21	0/691	0.38	0/933
6	H	0.22	0/1060	0.39	0/1434
7	I	0.22	0/953	0.36	0/1284
8	J	0.22	0/541	0.38	0/727
9	K	0.22	0/937	0.36	0/1265
10	L	0.21	0/353	0.36	0/468
11	T	0.55	0/225	0.91	0/342
12	R	0.15	0/196	0.64	0/304
All	All	0.22	0/28740	0.39	0/38877

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

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	10784	0	10871	161	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	8726	0	8759	128	0
3	C	2095	0	2051	36	0
4	E	1744	0	1772	13	0
5	F	679	0	701	10	0
6	H	1043	0	1015	15	0
7	I	935	0	886	8	0
8	J	532	0	542	10	0
9	K	919	0	929	11	0
10	L	351	0	375	5	0
11	T	214	0	122	6	0
12	R	175	0	87	0	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
14	A	1	0	0	0	0
All	All	28206	0	28110	362	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 6.

The worst 5 of 362 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:B:29:ASP:HB3	2:B:658:ILE:HG12	1.61	0.83
2:B:25:ILE:HD11	2:B:653:VAL:HB	1.62	0.82
2:B:653:VAL:HG22	2:B:689:LEU:HB3	1.65	0.79
1:A:35:ILE:HG13	1:A:241:VAL:HG21	1.70	0.73
3:C:70:ILE:HD11	3:C:144:ILE:HD11	1.72	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	1358/1733 (78%)	1249 (92%)	99 (7%)	10 (1%)	22	55
2	B	1077/1224 (88%)	989 (92%)	80 (7%)	8 (1%)	22	55
3	C	264/318 (83%)	248 (94%)	14 (5%)	2 (1%)	19	51
4	E	211/215 (98%)	198 (94%)	12 (6%)	1 (0%)	29	61
5	F	82/155 (53%)	76 (93%)	6 (7%)	0	100	100
6	H	124/146 (85%)	109 (88%)	13 (10%)	2 (2%)	9	34
7	I	113/122 (93%)	102 (90%)	11 (10%)	0	100	100
8	J	63/70 (90%)	59 (94%)	4 (6%)	0	100	100
9	K	112/120 (93%)	107 (96%)	4 (4%)	1 (1%)	17	49
10	L	42/70 (60%)	34 (81%)	7 (17%)	1 (2%)	6	28
All	All	3446/4173 (83%)	3171 (92%)	250 (7%)	25 (1%)	22	55

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	40	THR
1	A	110	CYS
1	A	322	VAL
1	A	1437	GLY
2	B	468	GLU

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	1196/1520 (79%)	1137 (95%)	59 (5%)	25	55
2	B	952/1061 (90%)	915 (96%)	37 (4%)	32	61
3	C	234/274 (85%)	227 (97%)	7 (3%)	41	68
4	E	195/197 (99%)	189 (97%)	6 (3%)	40	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	F	74/137 (54%)	70 (95%)	4 (5%)	22	52
6	H	114/128 (89%)	110 (96%)	4 (4%)	36	65
7	I	109/116 (94%)	107 (98%)	2 (2%)	59	79
8	J	60/65 (92%)	57 (95%)	3 (5%)	24	54
9	K	99/102 (97%)	97 (98%)	2 (2%)	55	77
10	L	39/57 (68%)	39 (100%)	0	100	100
All	All	3072/3657 (84%)	2948 (96%)	124 (4%)	31	60

5 of 124 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1400	CYS
5	F	93	ILE
2	B	570	VAL
5	F	72	LYS
7	I	111	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	762	ASN
2	B	984	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
12	R	7/8 (87%)	1 (14%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
12	R	8	G

There are no RNA pucker outliers to report.

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

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
11	3DR	T	20	11	8,11,12	1.53	1 (12%)	9,14,17	0.84	0

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
11	3DR	T	20	11	-	2/3/15/16	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	T	20	3DR	O4'-C4'	-2.42	1.40	1.44

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	T	20	3DR	O4'-C4'-C5'-O5'
11	T	20	3DR	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	T	20	3DR	3	0

5.5 Carbohydrates [i](#)

There are no monosaccharides 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	A	1372/1733 (79%)	0.31	61 (4%) 34 34	19, 76, 172, 296	0
2	B	1097/1224 (89%)	0.13	12 (1%) 80 79	19, 59, 133, 245	0
3	C	266/318 (83%)	0.11	1 (0%) 92 92	24, 60, 106, 189	0
4	E	213/215 (99%)	0.69	26 (12%) 4 5	44, 112, 190, 264	0
5	F	84/155 (54%)	0.03	1 (1%) 79 77	38, 75, 123, 159	0
6	H	130/146 (89%)	1.02	20 (15%) 2 2	58, 100, 155, 217	0
7	I	115/122 (94%)	0.33	5 (4%) 35 35	33, 78, 117, 175	0
8	J	65/70 (92%)	-0.07	0 100 100	26, 46, 100, 141	0
9	K	114/120 (95%)	0.01	0 100 100	31, 64, 108, 131	0
10	L	44/70 (62%)	0.66	6 (13%) 3 3	34, 114, 210, 228	0
11	T	10/29 (34%)	0.52	1 (10%) 7 8	70, 86, 170, 209	0
12	R	8/8 (100%)	0.19	0 100 100	62, 71, 139, 193	0
All	All	3518/4210 (83%)	0.27	133 (3%) 40 39	19, 70, 160, 296	0

The worst 5 of 133 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	H	139	ASN	5.8
2	B	643	ASP	5.7
4	E	93	MET	5.6
1	A	105	CYS	4.9
4	E	100	ILE	4.7

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(Å ²)	Q<0.9
11	3DR	T	20	11/12	0.88	0.25	143,150,185,188	0

6.3 Carbohydrates [i](#)

There are no monosaccharides 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
13	ZN	A	1801	1/1	0.65	0.07	188,188,188,188	0
13	ZN	A	1802	1/1	0.95	0.13	97,97,97,97	0
13	ZN	L	101	1/1	0.97	0.07	130,130,130,130	0
14	MG	A	1803	1/1	0.97	0.19	24,24,24,24	0
13	ZN	B	1301	1/1	0.98	0.11	99,99,99,99	0
13	ZN	C	401	1/1	0.98	0.20	53,53,53,53	0
13	ZN	I	201	1/1	0.99	0.12	71,71,71,71	0
13	ZN	I	202	1/1	0.99	0.17	55,55,55,55	0
13	ZN	J	101	1/1	1.00	0.16	42,42,42,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.