



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2020 – 06:43 PM BST

PDB ID : 6UQ2
Title : RNA polymerase II elongation complex with dG in state 1
Authors : Oh, J.; Wang, D.
Deposited on : 2019-10-18
Resolution : 3.20 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13
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)	:	2.13

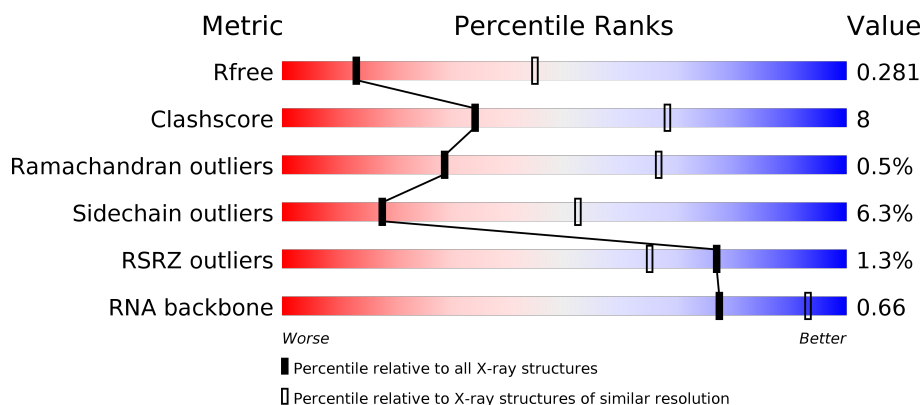
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)
RNA backbone	3102	1010 (3.50-2.90)

2 Entry composition [i](#)

There are 15 unique types of molecules in this entry. The entry contains 29048 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	9	Total	C	N	O	P	0	0	0
			195	88	40	59	8			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	26	Total	C	N	O	P	0	0	0
			520	250	80	164	26			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	15	Total	C	N	O	P	0	0	0
			317	148	71	83	15			

- 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 (three-letter code: ZN) (formula: Zn).

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

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	165.69Å 223.32Å 193.52Å 90.00° 99.37° 90.00°	Depositor
Resolution (Å)	49.75 – 3.20 49.76 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.75-3.20) 99.9 (49.76-3.20)	Depositor EDS
R_{merge}	0.33	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.13 _2998	Depositor
R, R_{free}	0.230 , 0.280 0.230 , 0.281	Depositor DCC
R_{free} test set	1801 reflections (1.58%)	wwPDB-VP
Wilson B-factor (Å ²)	80.5	Xtriage
Anisotropy	0.566	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 43.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	29048	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.57% 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

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

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	0.32	0/219	0.88	0/341
2	T	0.56	0/577	1.07	1/886 (0.1%)
3	N	0.57	0/359	0.94	1/553 (0.2%)
4	A	0.30	0/11020	0.59	6/14907 (0.0%)
5	B	0.31	0/9030	0.55	3/12186 (0.0%)
6	C	0.30	0/2139	0.56	0/2899
7	E	0.32	0/1767	0.60	1/2378 (0.0%)
8	F	0.28	0/696	0.55	0/943
9	H	0.33	0/1082	0.66	0/1466
10	I	0.32	0/970	0.58	0/1308
11	J	0.30	0/541	0.60	0/727
12	K	0.32	0/937	0.63	0/1265
13	L	0.32	0/339	0.57	0/450
All	All	0.32	0/29676	0.60	12/40309 (0.0%)

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1150	ARG	NE-CZ-NH2	-6.48	117.06	120.30
4	A	472	LEU	CB-CG-CD2	-6.33	100.25	111.00
5	B	1150	ARG	CA-CB-CG	-6.27	99.60	113.40
4	A	550	LEU	CA-CB-CG	6.09	129.30	115.30
4	A	702	LEU	CB-CG-CD1	6.02	121.24	111.00

There are no chirality outliers.

There are no planarity outliers.

4.2 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	R	195	0	99	0	0
2	T	520	0	297	6	0
3	N	317	0	166	3	0
4	A	10828	0	10876	203	0
5	B	8859	0	8816	153	0
6	C	2101	0	2056	38	0
7	E	1731	0	1758	37	0
8	F	684	0	692	16	0
9	H	1064	0	1029	22	0
10	I	952	0	897	29	0
11	J	532	0	542	13	0
12	K	919	0	929	12	0
13	L	337	0	352	7	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	29048	0	28509	477	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 8.

The worst 5 of 477 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:B:29:ASP:OD2	5:B:655:LYS:NZ	1.65	1.27
10:I:59:VAL:O	10:I:62:ILE:HD13	1.54	1.06
10:I:59:VAL:H	10:I:62:ILE:CD1	1.68	1.05
7:E:55:ARG:NH2	7:E:137:GLU:OE1	2.01	0.93
10:I:59:VAL:H	10:I:62:ILE:HD11	1.35	0.90

There are no symmetry-related clashes.

4.3 Torsion angles

4.3.1 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%)	1276 (93%)	85 (6%)	9 (1%)	22	61
5	B	1103/1224 (90%)	1040 (94%)	60 (5%)	3 (0%)	41	74
6	C	265/318 (83%)	246 (93%)	19 (7%)	0	100	100
7	E	210/215 (98%)	192 (91%)	16 (8%)	2 (1%)	15	54
8	F	84/155 (54%)	75 (89%)	9 (11%)	0	100	100
9	H	129/146 (88%)	116 (90%)	12 (9%)	1 (1%)	19	58
10	I	116/122 (95%)	107 (92%)	9 (8%)	0	100	100
11	J	63/70 (90%)	60 (95%)	3 (5%)	0	100	100
12	K	112/120 (93%)	106 (95%)	5 (4%)	1 (1%)	17	56
13	L	41/70 (59%)	36 (88%)	5 (12%)	0	100	100
All	All	3493/4173 (84%)	3254 (93%)	223 (6%)	16 (0%)	29	67

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
12	K	28	PRO
4	A	55	ASP
4	A	704	ALA
4	A	1036	ARG
4	A	1242	VAL

4.3.2 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%)	1116 (94%)	78 (6%)	17	51
5	B	955/1061 (90%)	900 (94%)	55 (6%)	20	55
6	C	235/274 (86%)	226 (96%)	9 (4%)	33	67
7	E	193/197 (98%)	181 (94%)	12 (6%)	18	53
8	F	73/137 (53%)	69 (94%)	4 (6%)	21	57
9	H	116/128 (91%)	106 (91%)	10 (9%)	10	38
10	I	110/116 (95%)	99 (90%)	11 (10%)	7	30
11	J	60/65 (92%)	58 (97%)	2 (3%)	38	71
12	K	99/102 (97%)	91 (92%)	8 (8%)	11	42
13	L	37/57 (65%)	33 (89%)	4 (11%)	6	27
All	All	3072/3657 (84%)	2879 (94%)	193 (6%)	18	52

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	250	PHE
5	B	662	MET
10	I	88	SER
5	B	264	SER
5	B	437	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
5	B	103	ASN
5	B	835	GLN
5	B	1193	GLN
5	B	1211	ASN

4.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/9 (88%)	2 (25%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	4	G
1	R	9	G

There are no RNA pucker outliers to report.

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

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 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.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data ⓘ

5.1 Protein, DNA and RNA chains ⓘ

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	9/9 (100%)	-0.51	0 100 100	85, 102, 180, 188	0
2	T	26/29 (89%)	-0.22	2 (7%) 13 7	85, 159, 230, 258	0
3	N	15/18 (83%)	0.17	0 100 100	181, 217, 240, 254	0
4	A	1384/1733 (79%)	-0.14	24 (1%) 70 57	46, 92, 174, 257	0
5	B	1123/1224 (91%)	-0.23	3 (0%) 94 92	37, 74, 140, 220	0
6	C	267/318 (83%)	-0.35	0 100 100	46, 78, 123, 146	0
7	E	212/215 (98%)	-0.06	7 (3%) 46 30	74, 117, 180, 247	0
8	F	86/155 (55%)	-0.42	1 (1%) 79 67	62, 91, 136, 185	0
9	H	133/146 (91%)	0.07	5 (3%) 40 26	80, 123, 169, 233	0
10	I	118/122 (96%)	-0.31	0 100 100	65, 99, 138, 176	0
11	J	65/70 (92%)	-0.40	0 100 100	49, 69, 104, 129	0
12	K	114/120 (95%)	-0.30	0 100 100	50, 84, 121, 133	0
13	L	43/70 (61%)	0.36	3 (6%) 16 9	64, 141, 209, 232	0
All	All	3595/4229 (85%)	-0.19	45 (1%) 77 65	37, 88, 168, 258	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	69	THR	8.0
4	A	144	THR	6.1
4	A	183	GLY	5.6
4	A	141	LEU	5.5
4	A	182	VAL	5.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.4 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(\AA^2)	Q<0.9
14	ZN	I	201	1/1	0.86	0.10	100,100,100,100	0
14	ZN	L	101	1/1	0.88	0.07	208,208,208,208	0
14	ZN	A	1801	1/1	0.88	0.08	204,204,204,204	0
15	MG	A	1803	1/1	0.92	0.06	81,81,81,81	0
14	ZN	J	101	1/1	0.95	0.25	74,74,74,74	0
14	ZN	B	1301	1/1	0.96	0.09	152,152,152,152	0
14	ZN	C	401	1/1	0.96	0.06	69,69,69,69	0
14	ZN	A	1802	1/1	0.97	0.11	128,128,128,128	0
14	ZN	I	202	1/1	0.99	0.14	101,101,101,101	0

5.5 Other polymers [i](#)

There are no such residues in this entry.