wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2023 – 05:11 PM EDT

PDB ID	:	7RIQ
Title	:	RNA polymerase II elongation complex scaffold 1 without polyamide
Authors	:	Oh, J.; Dervan, P.B.; Wang, D.
Deposited on	:	2021-07-20
Resolution	:	3.00 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 (i) 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 (i)) were used in the production of this report:

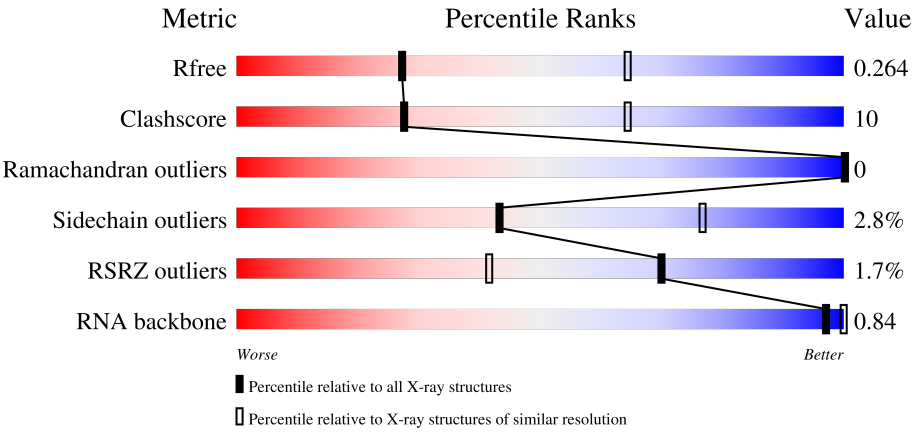
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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.



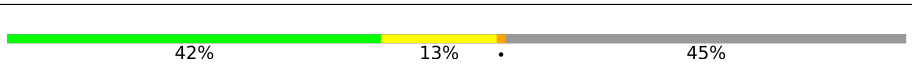
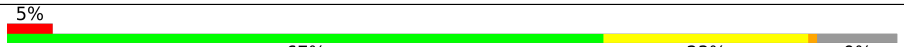
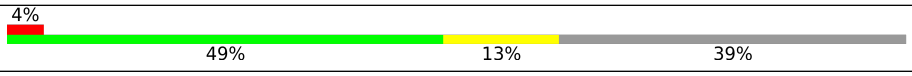
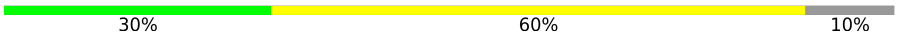
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

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 <=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	9	33% 67%
2	T	30	7% 80% 7% 7%
3	A	1733	2% 60% 20% 19%
4	B	1224	% 69% 22% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	C	318	
6	E	215	
7	F	155	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	N	20	

2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 29255 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	28	Total	C	N	O	P	0	0	0
			560	270	90	173	27			

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	1397	Total	C	N	O	S	0	0	0
			10933	6892	1913	2068	60			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	1124	Total	C	N	O	S	0	0	0
			8862	5609	1552	1648	53			

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

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

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

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

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

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

- Molecule 13 is a DNA chain called Non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	18	Total	C	N	O	P	0	0	0
			376	177	78	103	18			

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

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

- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

3 Residue-property plots (i)

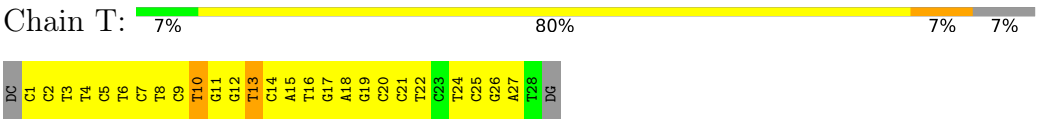
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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

Chain R:

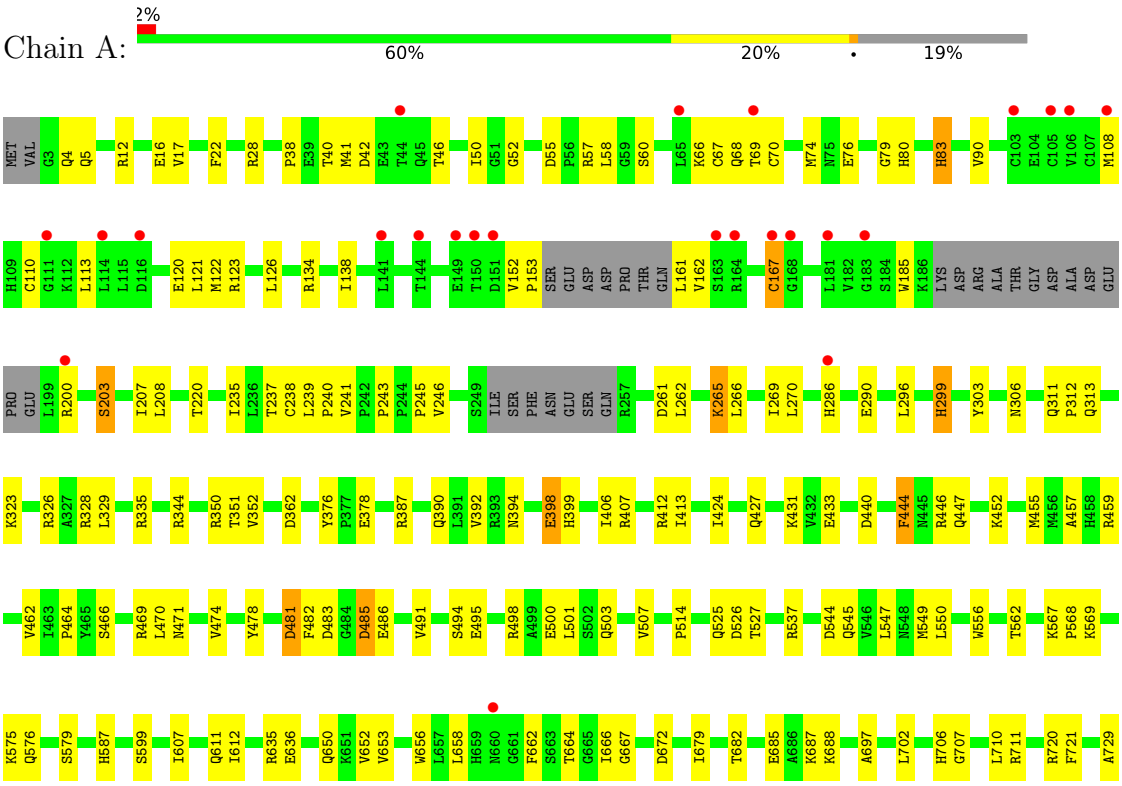
• Molecule 2: Template strand DNA

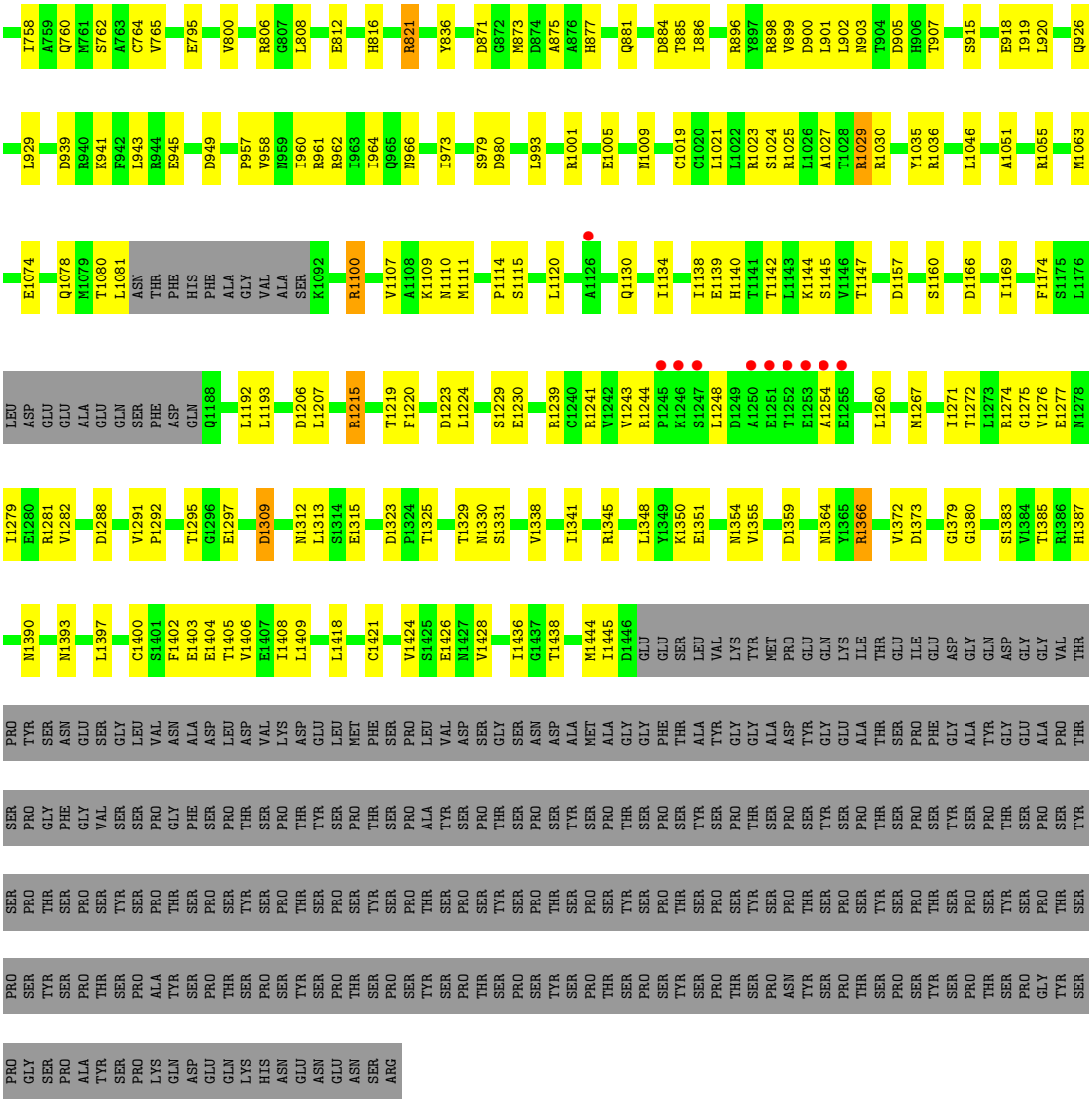
Chain T:



• Molecule 3: DNA-directed RNA polymerase II subunit RPB1

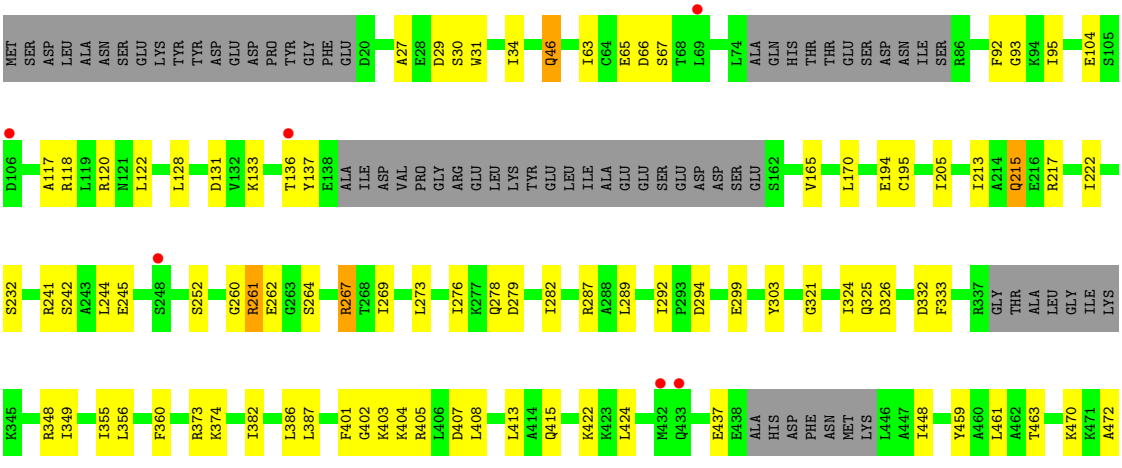
Chain A:

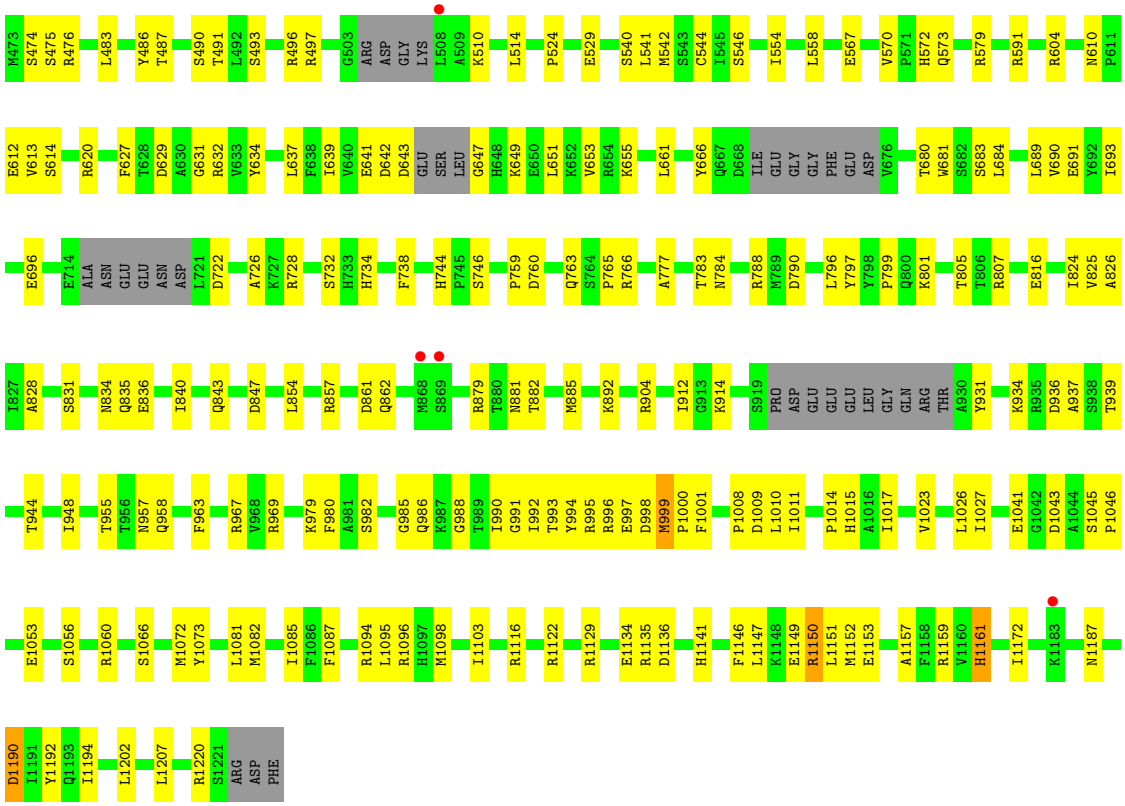




• Molecule 4: DNA-directed RNA polymerase II subunit RPB2

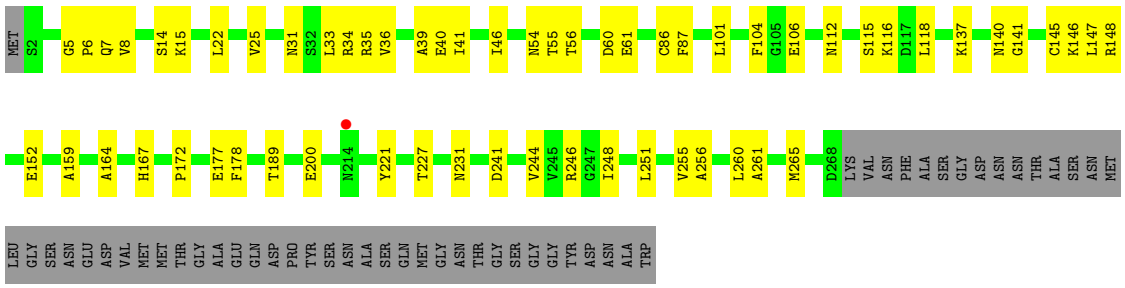
Chain B:





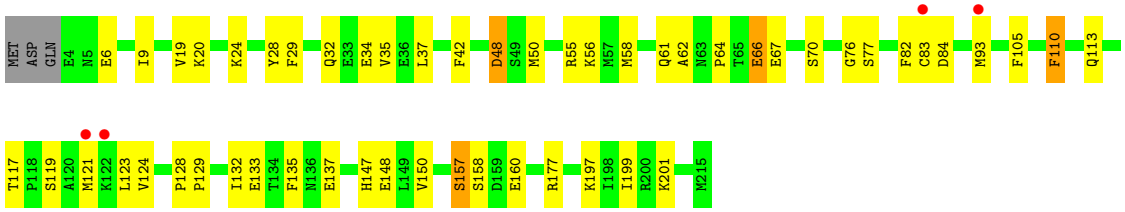
• Molecule 5: DNA-directed RNA polymerase II subunit RPB3

Chain C:



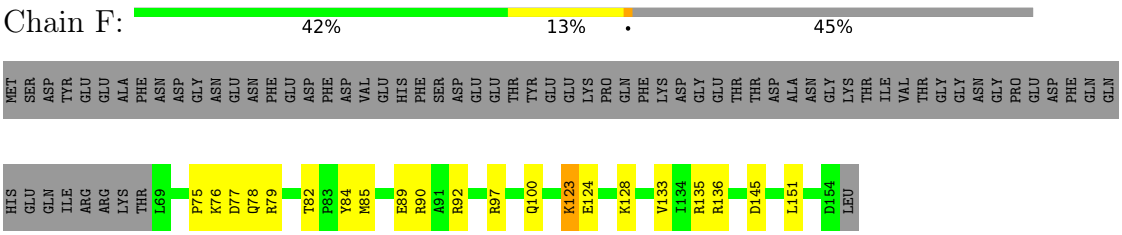
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



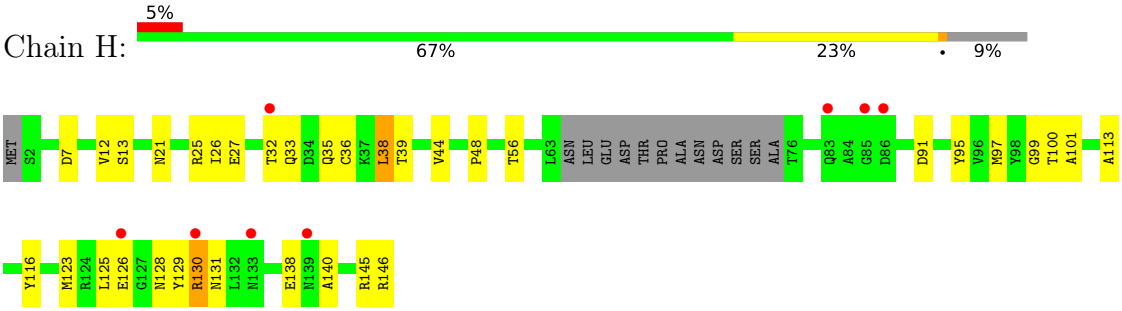
• Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:



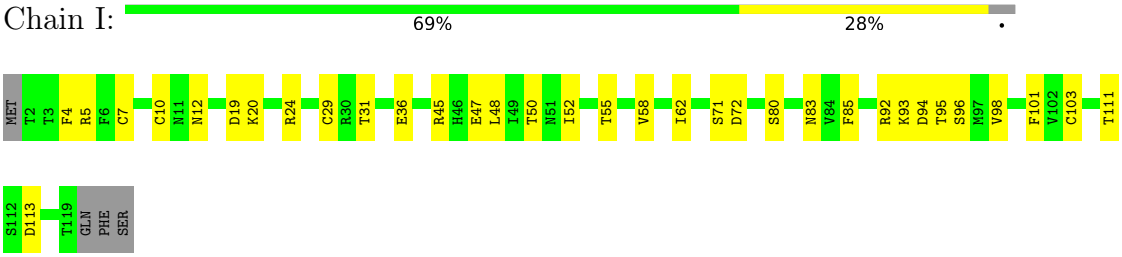
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



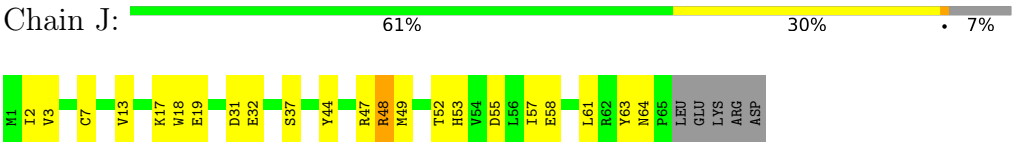
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:



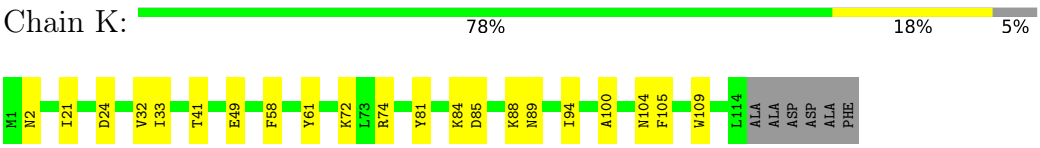
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



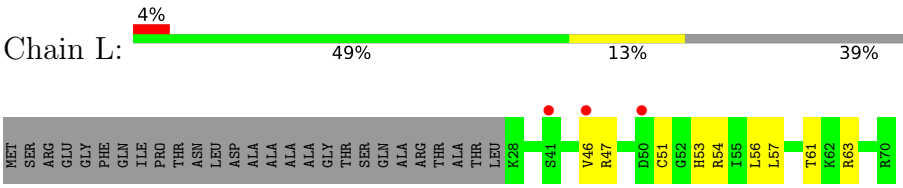
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain K:



• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

Chain L:



• Molecule 13: Non-template strand DNA

Chain N:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	167.85Å 223.47Å 193.03Å 90.00° 100.30° 90.00°	Depositor
Resolution (Å)	48.15 – 3.00 48.15 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.15-3.00) 99.8 (48.15-3.00)	Depositor EDS
R_{merge}	0.31	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.11 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.13	Depositor
R, R_{free}	0.213 , 0.264 0.213 , 0.264	Depositor DCC
R_{free} test set	1919 reflections (1.38%)	wwPDB-VP
Wilson B-factor (Å^2)	79.5	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 61.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	29255	wwPDB-VP
Average B, all atoms (Å^2)	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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, 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	0.25	0/219	0.82	0/341
2	T	0.59	0/623	1.06	2/958 (0.2%)
3	A	0.24	0/11127	0.43	1/15053 (0.0%)
4	B	0.24	0/9033	0.43	0/12191
5	C	0.25	0/2139	0.44	0/2899
6	E	0.24	0/1767	0.42	0/2378
7	F	0.23	0/696	0.42	0/943
8	H	0.25	0/1082	0.48	0/1466
9	I	0.26	0/970	0.44	0/1308
10	J	0.27	0/541	0.47	0/727
11	K	0.24	0/937	0.42	0/1265
12	L	0.25	0/339	0.50	0/450
13	N	0.56	0/424	0.89	0/653
All	All	0.26	0/29897	0.47	3/40632 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	398	GLU	C-N-CA	6.62	138.24	121.70
2	T	10	DT	N3-C4-O4	5.21	123.03	119.90
2	T	13	DT	N3-C4-O4	5.12	122.97	119.90

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	195	0	99	6	0
2	T	560	0	320	31	0
3	A	10933	0	10974	248	0
4	B	8862	0	8814	187	0
5	C	2101	0	2056	40	0
6	E	1731	0	1758	32	0
7	F	684	0	692	15	0
8	H	1064	0	1029	17	0
9	I	952	0	897	24	0
10	J	532	0	542	16	0
11	K	919	0	929	16	0
12	L	337	0	352	4	0
13	N	376	0	201	10	0
14	R	1	0	0	0	0
15	A	2	0	0	0	0
15	B	1	0	0	0	0
15	C	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	L	1	0	0	0	0
All	All	29255	0	28663	577	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 10.

The worst 5 of 577 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)
3:A:323:LYS:HZ1	3:A:328:ARG:NE	1.68	0.91
4:B:763:GLN:HG2	4:B:765:PRO:HD2	1.56	0.85
3:A:323:LYS:NZ	3:A:328:ARG:NE	2.26	0.83
13:N:10:DA:H2"	13:N:11:DG:H5"	1.62	0.81
3:A:1111:MET:HG3	3:A:1114:PRO:HG3	1.62	0.79

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	
3	A	1385/1733 (80%)	1327 (96%)	58 (4%)	0	100	100
4	B	1104/1224 (90%)	1067 (97%)	37 (3%)	0	100	100
5	C	265/318 (83%)	260 (98%)	5 (2%)	0	100	100
6	E	210/215 (98%)	201 (96%)	9 (4%)	0	100	100
7	F	84/155 (54%)	84 (100%)	0	0	100	100
8	H	129/146 (88%)	113 (88%)	16 (12%)	0	100	100
9	I	116/122 (95%)	113 (97%)	3 (3%)	0	100	100
10	J	63/70 (90%)	62 (98%)	1 (2%)	0	100	100
11	K	112/120 (93%)	107 (96%)	5 (4%)	0	100	100
12	L	41/70 (59%)	40 (98%)	1 (2%)	0	100	100
All	All	3509/4173 (84%)	3374 (96%)	135 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	
3	A	1206/1520 (79%)	1168 (97%)	38 (3%)	39	74
4	B	955/1061 (90%)	928 (97%)	27 (3%)	43	77
5	C	235/274 (86%)	234 (100%)	1 (0%)	91	97
6	E	193/197 (98%)	187 (97%)	6 (3%)	40	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	F	73/137 (53%)	72 (99%)	1 (1%)	67 88
8	H	116/128 (91%)	110 (95%)	6 (5%)	23 59
9	I	110/116 (95%)	108 (98%)	2 (2%)	59 85
10	J	60/65 (92%)	58 (97%)	2 (3%)	38 73
11	K	99/102 (97%)	98 (99%)	1 (1%)	76 91
12	L	37/57 (65%)	36 (97%)	1 (3%)	44 77
All	All	3084/3657 (84%)	2999 (97%)	85 (3%)	43 77

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	722	ASP
6	E	157	SER
4	B	996	ARG
5	C	137	LYS
8	H	38	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	835	GLN
4	B	1205	GLN
6	E	147	HIS
3	A	1390	ASN
3	A	1312	ASN

5.3.3 RNA (i)

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

There are no RNA backbone outliers to report.

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 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

6.1 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, 95^{th} 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%)	-0.59	0 100 100	83, 94, 157, 160	0
2	T	28/30 (93%)	-0.28	0 100 100	75, 167, 234, 236	0
3	A	1397/1733 (80%)	-0.15	34 (2%) 59 30	45, 105, 184, 302	0
4	B	1124/1224 (91%)	-0.19	10 (0%) 84 63	44, 88, 156, 208	0
5	C	267/318 (83%)	-0.39	1 (0%) 92 79	53, 89, 145, 197	0
6	E	212/215 (98%)	-0.10	4 (1%) 66 37	78, 134, 194, 239	0
7	F	86/155 (55%)	-0.31	0 100 100	73, 104, 141, 206	0
8	H	133/146 (91%)	0.25	8 (6%) 21 7	96, 129, 194, 228	0
9	I	118/122 (96%)	-0.40	0 100 100	66, 109, 159, 217	0
10	J	65/70 (92%)	-0.28	0 100 100	63, 88, 128, 168	0
11	K	114/120 (95%)	-0.38	0 100 100	60, 96, 144, 161	0
12	L	43/70 (61%)	0.26	3 (6%) 16 5	76, 161, 223, 257	0
13	N	18/20 (90%)	0.01	0 100 100	144, 197, 246, 267	0
All	All	3614/4232 (85%)	-0.18	60 (1%) 70 41	44, 101, 180, 302	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	1252	THR	12.5
3	A	1254	ALA	9.3
3	A	1251	GLU	8.5
3	A	1250	ALA	6.2
3	A	1245	PRO	6.1

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

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

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, 95^{th} 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
15	ZN	A	1801	1/1	0.79	0.10	330,330,330,330	0
15	ZN	C	401	1/1	0.90	0.34	240,240,240,240	0
15	ZN	L	101	1/1	0.92	0.11	388,388,388,388	0
15	ZN	B	1301	1/1	0.93	0.13	155,155,155,155	0
15	ZN	J	101	1/1	0.94	0.23	83,83,83,83	0
15	ZN	I	202	1/1	0.94	0.13	97,97,97,97	0
14	MG	R	2001	1/1	0.96	0.20	127,127,127,127	0
15	ZN	I	201	1/1	0.98	0.13	113,113,113,113	0
15	ZN	A	1802	1/1	0.99	0.09	113,113,113,113	0

6.5 Other polymers (i)

There are no such residues in this entry.