



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:23 PM GMT

PDB ID : 3QT1  
Title : RNA polymerase II variant containing A Chimeric RPB9-C11 subunit  
Authors : Ruan, W.; Lehmann, E.; Thomm, M.; Kostrewa, D.; Cramer, P.  
Deposited on : 2011-02-22  
Resolution : 4.30 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

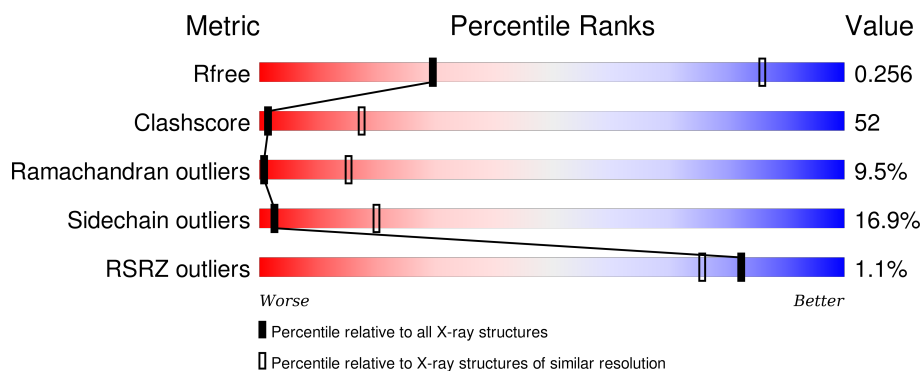
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.30 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))
$R_{free}$	91344	1059 (5.00-3.60)
Clashscore	102246	1166 (5.00-3.60)
Ramachandran outliers	100387	1106 (5.00-3.60)
Sidechain outliers	100360	1089 (5.00-3.60)
RSRZ outliers	91569	1062 (5.00-3.60)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	<div> <div>23%</div> <div>44%</div> <div>14%</div> <div>•</div> <div>18%</div> </div>
2	B	1224	<div> <div>26%</div> <div>49%</div> <div>15%</div> <div>•</div> <div>10%</div> </div>
3	C	318	<div> <div>26%</div> <div>44%</div> <div>13%</div> <div>16%</div> </div>
4	D	219	<div> <div>21%</div> <div>45%</div> <div>14%</div> <div>•</div> <div>19%</div> </div>
5	E	215	<div> <div>27%</div> <div>60%</div> <div>11%</div> <div>•</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	F	155	
7	G	171	
8	H	146	
9	I	133	
10	J	70	
11	K	120	
12	L	70	

## 2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 30535 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
1	A	1416	Total	C	N	O	S	0	0	0
			11143	7021	1949	2111	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1103	Total	C	N	O	S	0	0	0
			8770	5554	1535	1626	55			

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

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

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	178	Total	C	N	O	S	0	0	0
			1434	887	257	288	2			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- 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	134	Total	C	N	O	S	0	0	0
			1076	677	182	213	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9, DNA-directed RNA polymerase III subunit RPC10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	47	Total	C	N	O	S	0	0	0
			398	246	72	75	5			

- 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	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	J	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

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

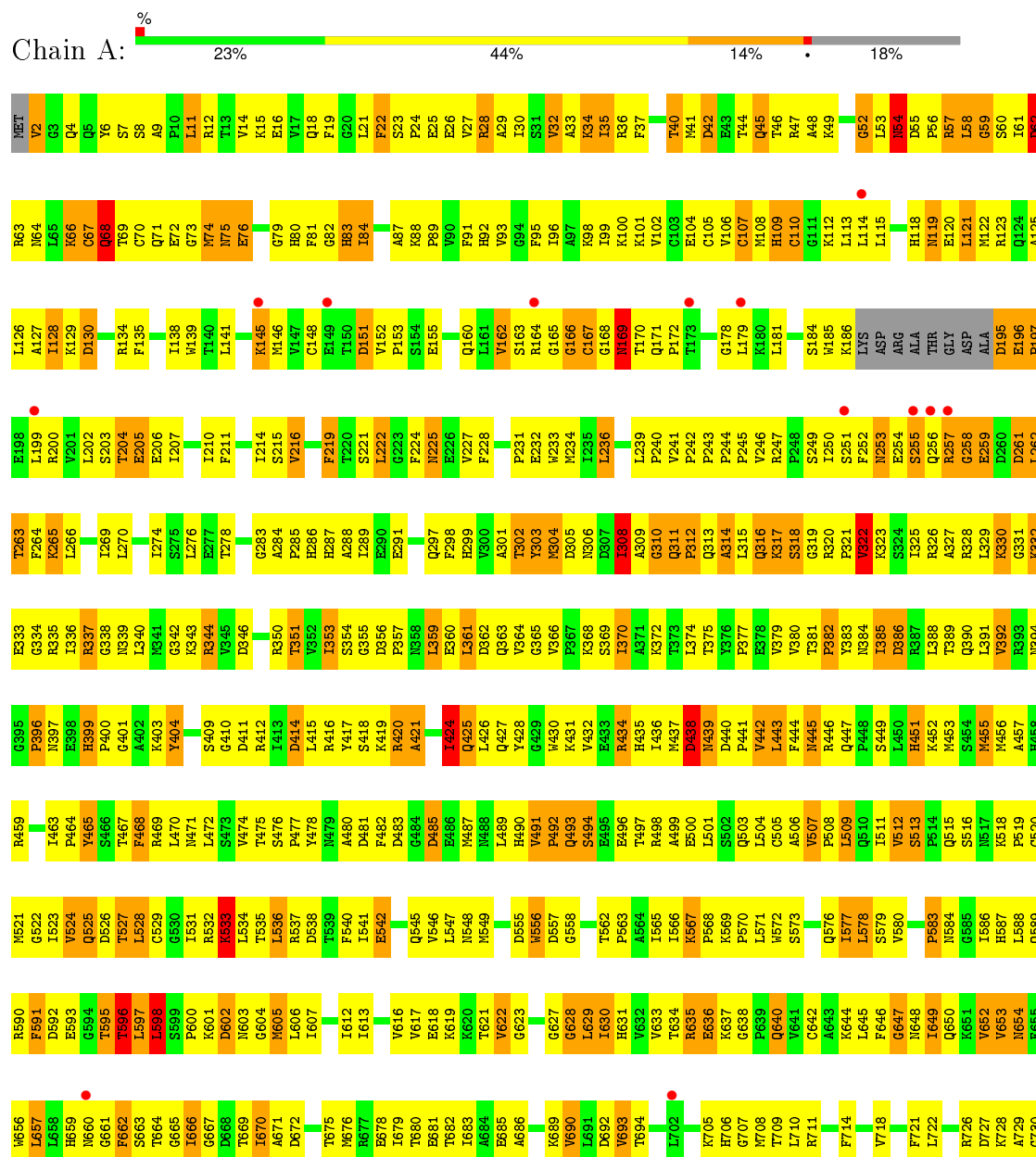
- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

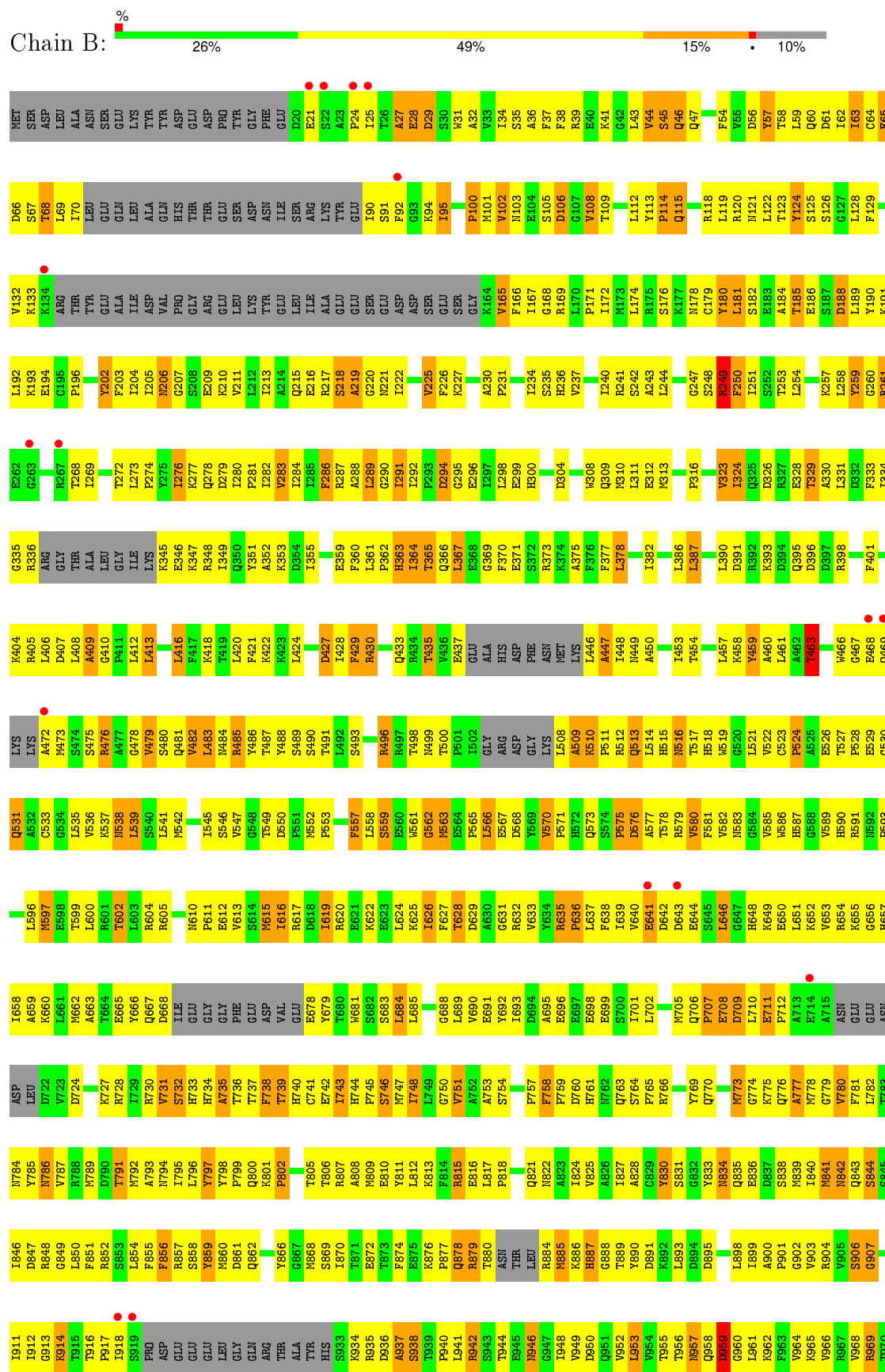
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

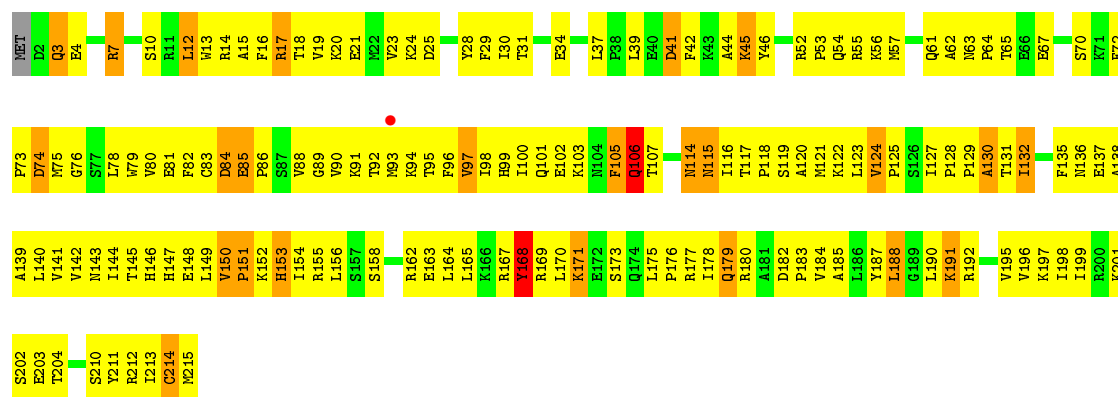
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1



- Molecule 2: DNA-directed RNA polymerase II subunit RPB2

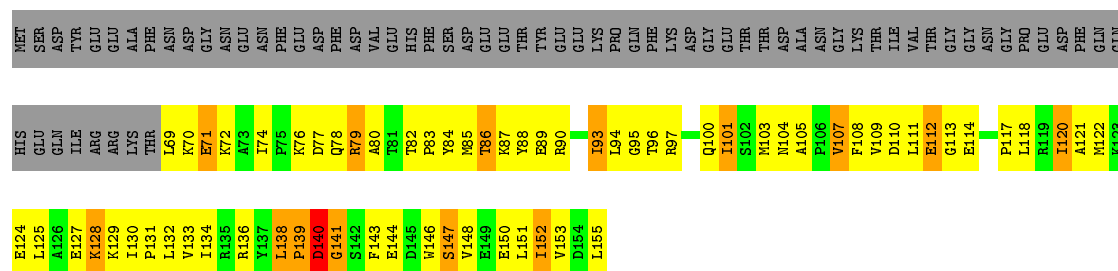






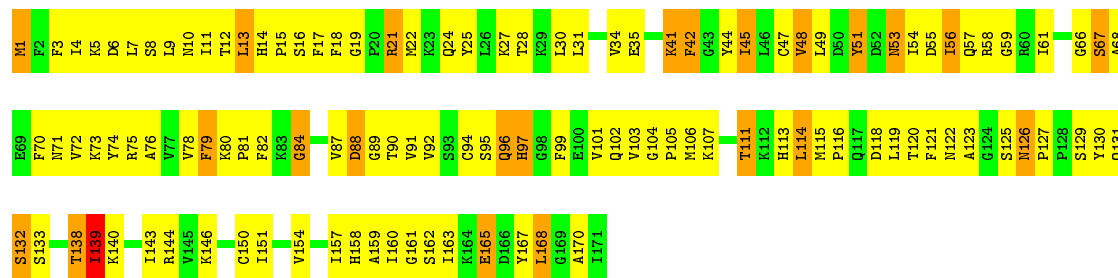
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 13% 34% 9% 44%



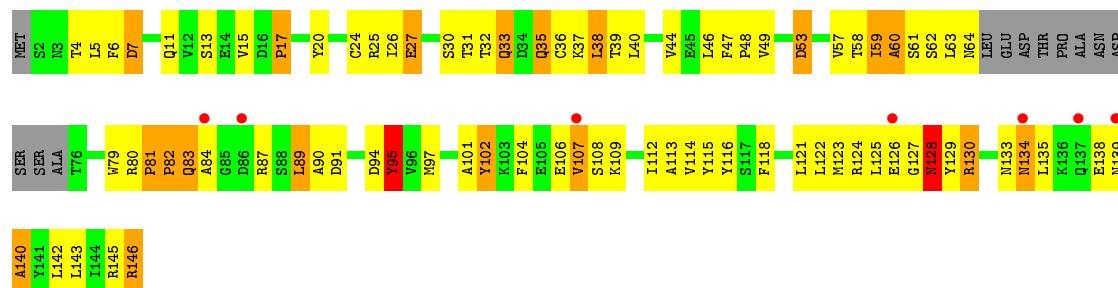
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G: 32% 54% 13%

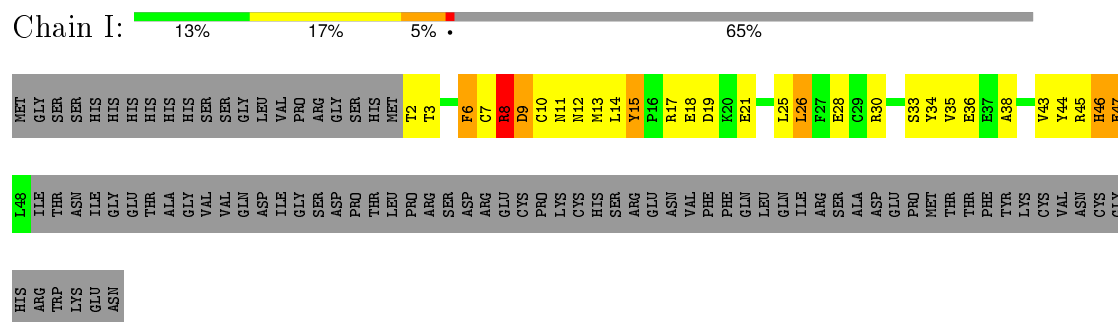


- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

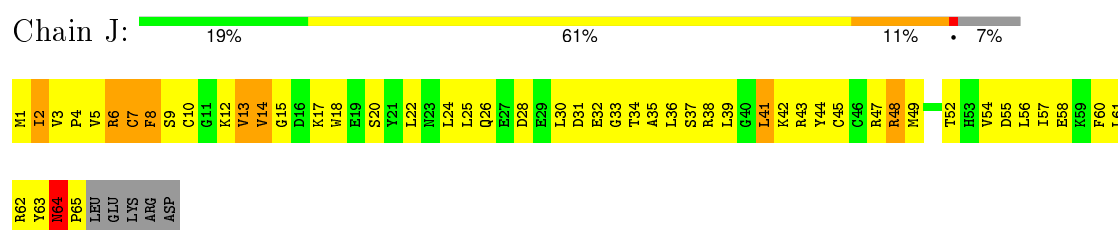
Chain H: 5% 35% 42% 13% 8%



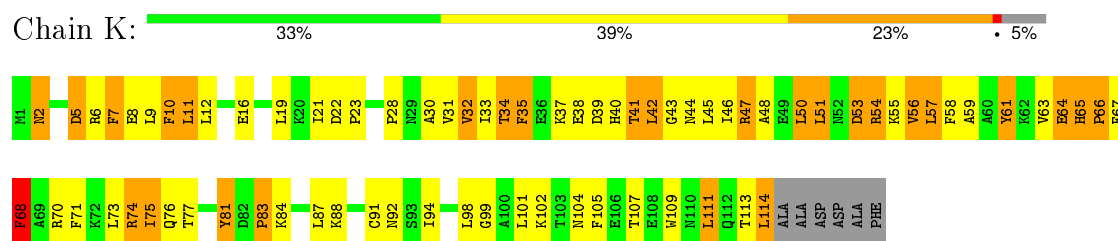
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9, DNA-directed RNA polymerase III subunit RPC10



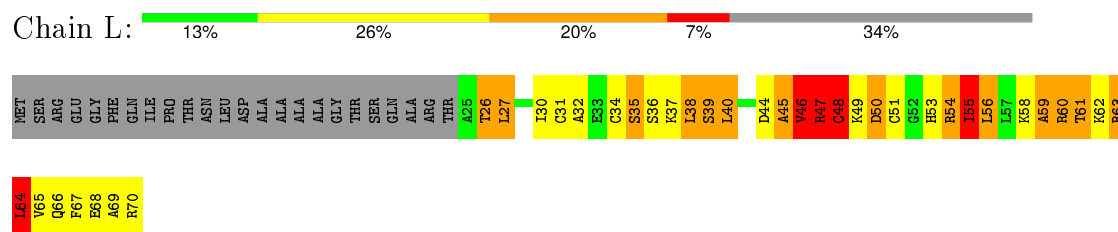
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.38Å 393.38Å 281.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.60 – 4.30 48.66 – 4.30	Depositor EDS
% Data completeness (in resolution range)	98.5 (48.60-4.30) 98.5 (48.66-4.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.67 (at 4.29Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, $R_{free}$	0.235 , 0.281 0.207 , 0.256	Depositor DCC
$R_{free}$ test set	2021 reflections (2.45%)	DCC
Wilson B-factor (Å <sup>2</sup> )	172.3	Xtriage
Anisotropy	0.553	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 119.5	EDS
Estimated twinning fraction	0.057 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.067 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 82532 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	30535	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	144.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.83% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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, 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	A	0.35	0/11342	0.58	0/15337
2	B	0.34	0/8939	0.56	0/12051
3	C	0.33	0/2133	0.56	0/2891
4	D	0.32	0/1444	0.52	0/1935
5	E	0.32	0/1788	0.54	0/2406
6	F	0.40	0/717	0.63	0/967
7	G	0.33	0/1368	0.55	0/1844
8	H	0.29	0/1094	0.50	0/1481
9	I	0.36	0/406	0.57	0/546
10	J	0.33	0/541	0.57	0/727
11	K	0.38	0/937	0.56	0/1265
12	L	0.36	0/365	0.64	0/485
All	All	0.34	0/31074	0.56	0/41935

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 [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	11143	0	11217	1243	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	8770	0	8795	984	0
3	C	2095	0	2051	223	0
4	D	1434	0	1460	163	0
5	E	1752	0	1776	172	0
6	F	705	0	731	101	0
7	G	1340	0	1357	153	0
8	H	1076	0	1046	97	0
9	I	398	0	370	37	0
10	J	532	0	542	92	0
11	K	919	0	929	113	0
12	L	363	0	386	65	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	1	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	30535	0	30660	3185	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 52.

The worst 5 of 3185 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:A:775:ILE:HB	1:A:797:LYS:O	1.37	1.23
1:A:567:LYS:HB2	1:A:568:PRO:HD2	1.21	1.17
2:B:763:GLN:HG2	2:B:765:PRO:HD2	1.26	1.15
9:I:25:LEU:HB3	9:I:38:ALA:HB2	1.17	1.14
2:B:577:ALA:HB1	2:B:589:VAL:HG11	1.30	1.13

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	1406/1733 (81%)	974 (69%)	284 (20%)	148 (10%)	1	12
2	B	1081/1224 (88%)	754 (70%)	232 (22%)	95 (9%)	1	17
3	C	264/318 (83%)	187 (71%)	48 (18%)	29 (11%)	0	11
4	D	174/219 (80%)	116 (67%)	43 (25%)	15 (9%)	1	17
5	E	212/215 (99%)	156 (74%)	41 (19%)	15 (7%)	1	23
6	F	85/155 (55%)	61 (72%)	19 (22%)	5 (6%)	2	28
7	G	169/171 (99%)	120 (71%)	43 (25%)	6 (4%)	4	40
8	H	130/146 (89%)	90 (69%)	29 (22%)	11 (8%)	1	17
9	I	45/133 (34%)	31 (69%)	10 (22%)	4 (9%)	1	17
10	J	63/70 (90%)	43 (68%)	14 (22%)	6 (10%)	1	15
11	K	112/120 (93%)	79 (70%)	24 (21%)	9 (8%)	1	19
12	L	44/70 (63%)	21 (48%)	8 (18%)	15 (34%)	0	0
All	All	3785/4574 (83%)	2632 (70%)	795 (21%)	358 (10%)	1	15

5 of 358 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ILE
1	A	42	ASP
1	A	48	ALA
1	A	54	ASN
1	A	57	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	1239/1520 (82%)	1031 (83%)	208 (17%)	2	20
2	B	957/1061 (90%)	800 (84%)	157 (16%)	3	21

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	234/274 (85%)	200 (86%)	34 (14%)	4	26
4	D	160/198 (81%)	127 (79%)	33 (21%)	1	11
5	E	196/197 (100%)	170 (87%)	26 (13%)	5	30
6	F	77/137 (56%)	65 (84%)	12 (16%)	3	24
7	G	152/152 (100%)	124 (82%)	28 (18%)	2	15
8	H	118/128 (92%)	98 (83%)	20 (17%)	2	20
9	I	45/122 (37%)	35 (78%)	10 (22%)	1	10
10	J	60/65 (92%)	50 (83%)	10 (17%)	3	20
11	K	99/102 (97%)	76 (77%)	23 (23%)	1	8
12	L	40/57 (70%)	29 (72%)	11 (28%)	0	5
All	All	3377/4013 (84%)	2805 (83%)	572 (17%)	2	20

5 of 572 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	496	ARG
2	B	980	PHE
10	J	7	CYS
2	B	559	SER
2	B	748	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 119 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	60	GLN
2	B	761	HIS
7	G	113	HIS
2	B	178	ASN
2	B	516	ASN

### 5.3.3 RNA ⓘ

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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1416/1733 (81%)	-0.22	18 (1%) 79 71	31, 133, 223, 314	0
2	B	1103/1224 (90%)	-0.17	16 (1%) 76 67	23, 141, 230, 357	0
3	C	266/318 (83%)	-0.26	0 100 100	58, 139, 209, 281	0
4	D	178/219 (81%)	-0.30	0 100 100	64, 146, 212, 263	0
5	E	214/215 (99%)	-0.29	1 (0%) 91 88	71, 165, 246, 367	0
6	F	87/155 (56%)	-0.36	0 100 100	48, 95, 156, 221	0
7	G	171/171 (100%)	-0.29	0 100 100	43, 121, 199, 230	0
8	H	134/146 (91%)	0.27	7 (5%) 31 24	115, 185, 250, 406	0
9	I	47/133 (35%)	-0.04	0 100 100	93, 163, 204, 221	0
10	J	65/70 (92%)	-0.33	0 100 100	80, 147, 210, 261	0
11	K	114/120 (95%)	-0.31	0 100 100	63, 128, 194, 254	0
12	L	46/70 (65%)	-0.21	0 100 100	93, 159, 240, 306	0
All	All	3841/4574 (83%)	-0.21	42 (1%) 82 76	23, 140, 227, 406	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
8	H	139	ASN	4.5
2	B	134	LYS	4.4
1	A	164	ARG	3.8
1	A	114	LEU	3.7
1	A	257	ARG	3.7

### 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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
13	ZN	J	3001	1/1	0.97	0.29	0.39	153,153,153,153	0
13	ZN	B	3007	1/1	0.99	0.21	-0.50	116,116,116,116	0
13	ZN	I	3003	1/1	1.00	0.12	-0.66	133,133,133,133	0
13	ZN	C	3002	1/1	0.98	0.10	-0.93	129,129,129,129	0
13	ZN	A	3008	1/1	0.96	0.13	-0.97	97,97,97,97	0
13	ZN	A	3006	1/1	0.89	0.09	-1.54	274,274,274,274	0
13	ZN	L	3005	1/1	0.97	0.08	-	229,229,229,229	0
14	MG	A	3009	1/1	0.85	0.93	-	89,89,89,89	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.