



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 02:02 am BST

PDB ID : 4A93  
Title : RNA Polymerase II elongation complex containing a CPD Lesion  
Authors : Walmacq, C.; Cheung, A.C.M.; Kireeva, M.L.; Lubkowska, L.; Ye, C.; Gotte, D.; Strathern, J.N.; Carell, T.; Cramer, P.; Kashlev, M.  
Deposited on : 2011-11-23  
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](mailto: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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

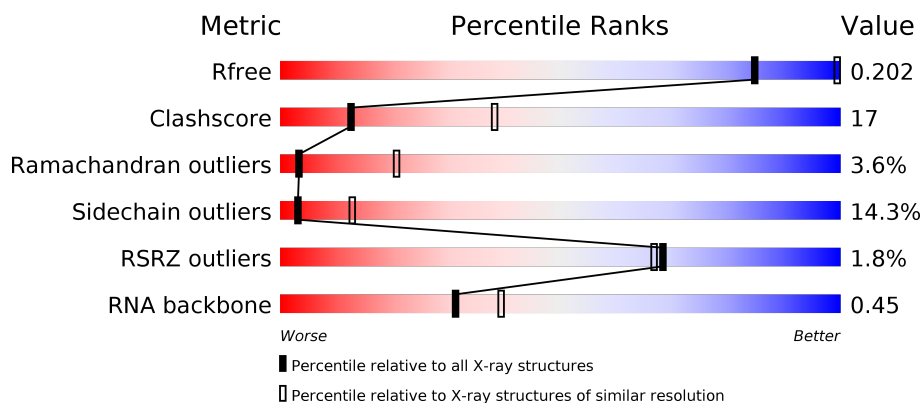
# 1 Overall quality at a glance

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	1732	<div> <div>2%</div> <div> <div></div> <div>45%</div> <div>30%</div> <div>6%</div> <div>18%</div> </div> </div>
2	B	1224	<div> <div>2%</div> <div> <div></div> <div>54%</div> <div>30%</div> <div>7%</div> <div>9%</div> </div> </div>
3	C	318	<div> <div></div> <div> <div></div> <div>52%</div> <div>27%</div> <div>5%</div> <div>16%</div> </div> </div>
4	D	221	<div> <div></div> <div> <div></div> <div>42%</div> <div>27%</div> <div>10%</div> <div>19%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	E	215	
6	F	155	
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	N	14	
14	P	12	
15	T	25	

## 2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 32105 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	1422	Total	C	N	O	S	0	0	0
			11174	7037	1954	2121	62			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1113	Total	C	N	O	S	0	0	0
			8839	5597	1548	1639	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 RPABC 1.

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 RPABC 2.

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

- Molecule 7 is a protein called RPB7, 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 RPABC 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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	115	Total	C	N	O	S	0	0	1
			920	590	157	171	2			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

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 a DNA chain called 5'-D(\*TP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*AP\*GP\*CP\*TP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	12	Total	C	N	O	P	0	0	0
			247	118	44	73	12			

- Molecule 14 is a RNA chain called 5'-R(\*UP\*UP\*CP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*AP\*AP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	P	9	Total	C	N	O	P	0	0	0
			197	88	41	59	9			

- Molecule 15 is a DNA chain called 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TT\*TP\*TP\*TP\*CP\*C BRUP\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	T	23	Total	Br	C	N	O	P	0	0
			485	1	234	81	146	23		

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

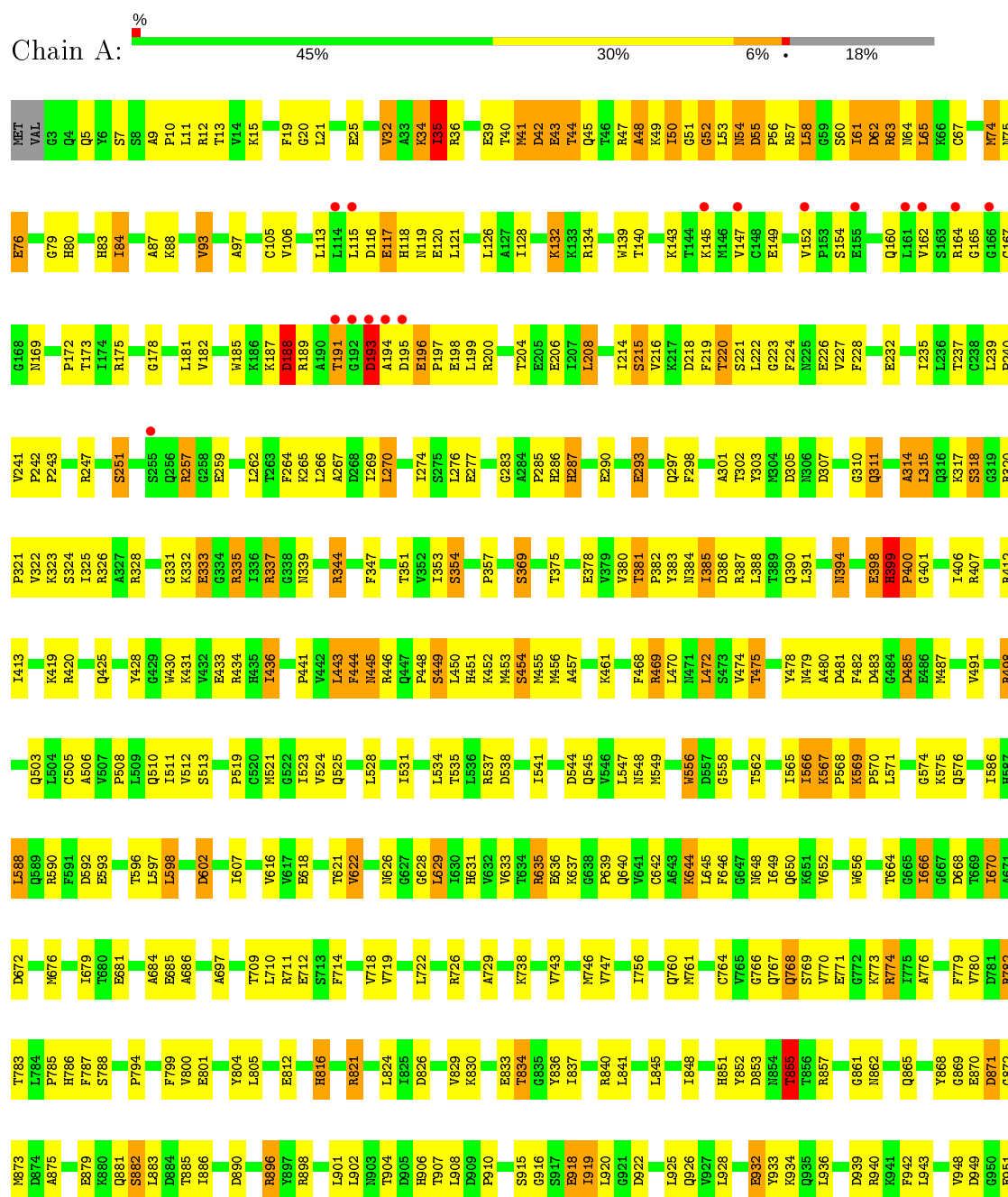
- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

### 3 Residue-property plots

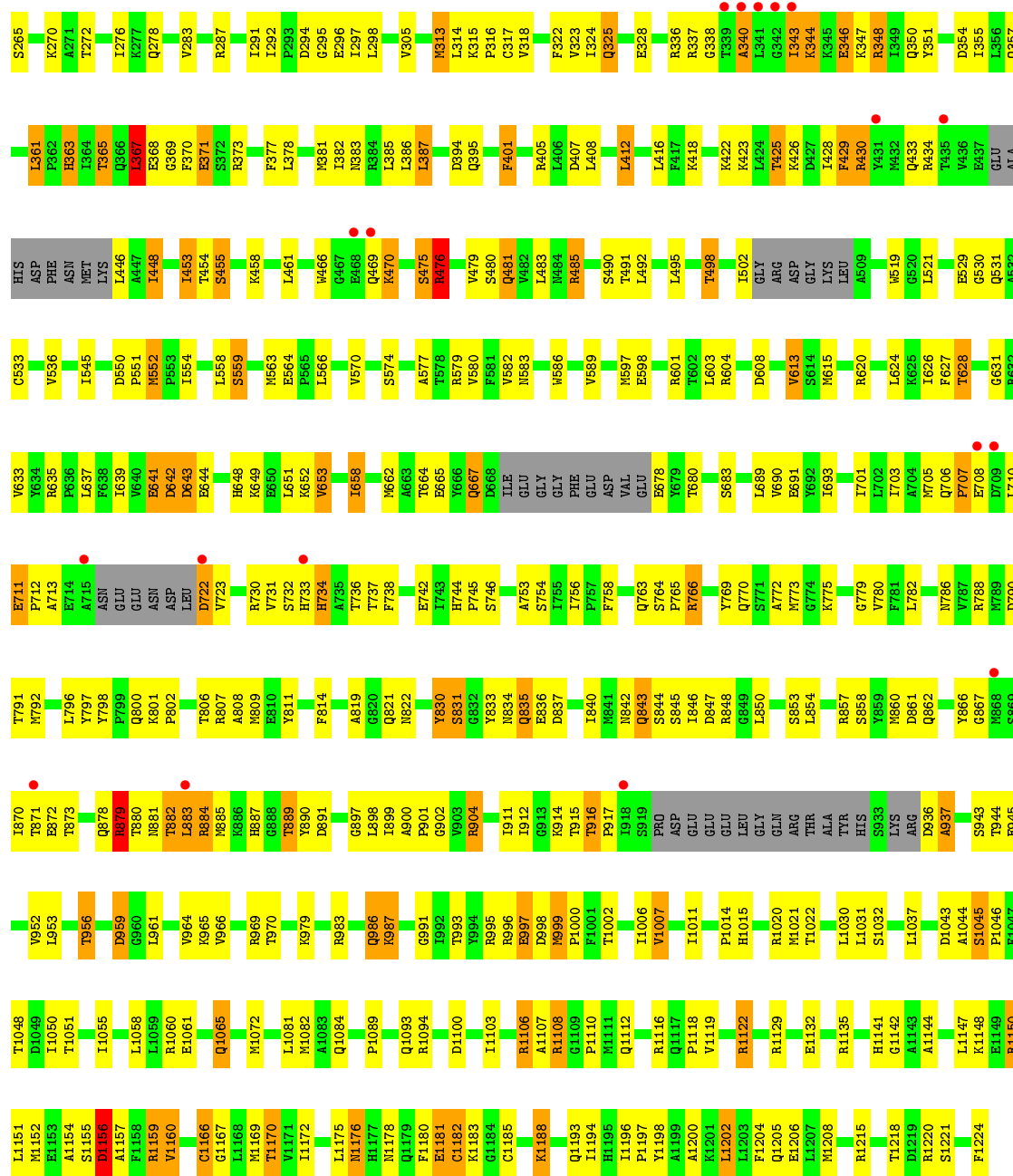
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

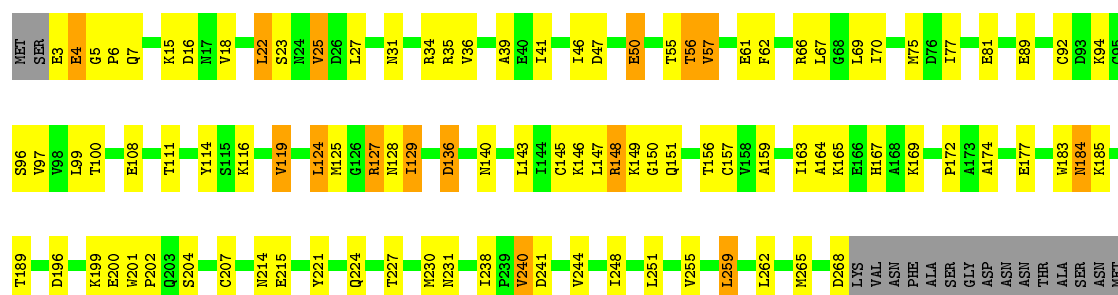








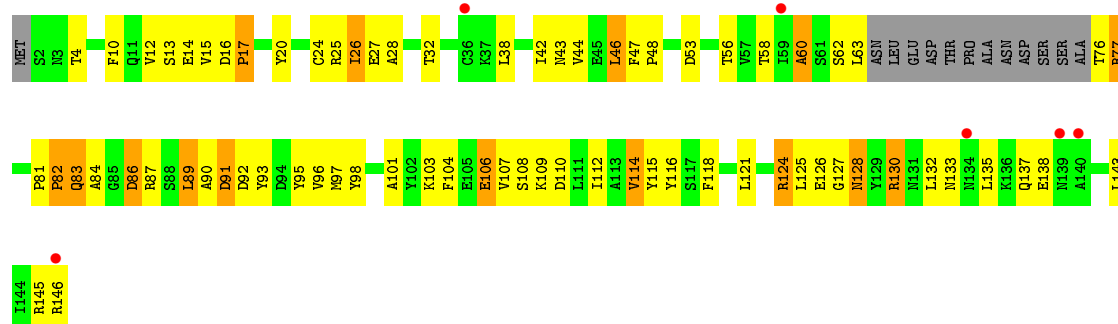
Chain C:



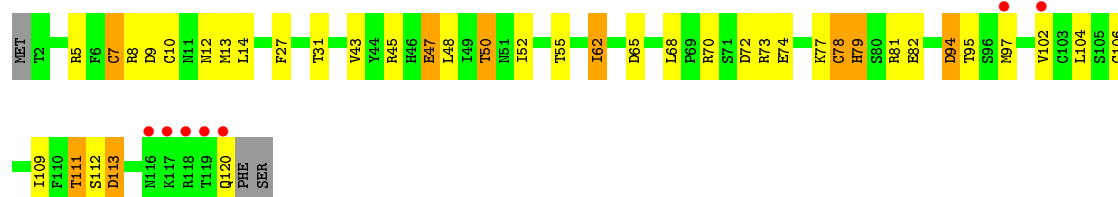




• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3



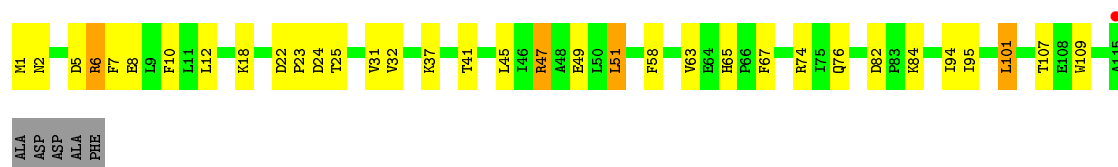
• Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9



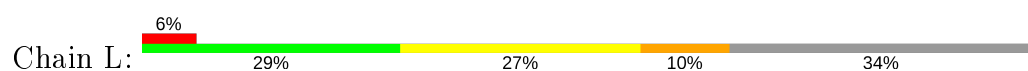
• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5



• Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11



• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4





- |    |    |    |    |    |    |    |    |    |     |     |     |     |     |
|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|
| DT | DA | A3 | G4 | T5 | A6 | C7 | T8 | T9 | G10 | A11 | G12 | C13 | T14 |
|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|

- |   |   |   |    |    |    |    |    |  |     |     |
|---|---|---|----|----|----|----|----|--|-----|-----|
| U | U | C | G3 | A4 | C5 | C6 | A7 |  | A10 | A11 |
|---|---|---|----|----|----|----|----|--|-----|-----|

- |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |    |    |
|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|
| A5 | G6 | C7 | T8 | C9 | A10 | A11 | G12 | T13 | A14 | C15 | T16 | N17 | T18 | T19 | G23 | G24 | T25 | C26 | A27 | DT | DT |
|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	221.64Å 391.52Å 281.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	54.09 – 3.40 54.09 – 3.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (54.09-3.40) 99.9 (54.09-3.40)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.25 (at 3.40Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.165 , 0.199 0.171 , 0.202	Depositor DCC
$R_{free}$ test set	3298 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	103.5	Xtriage
Anisotropy	0.405	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 100.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.025 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.029 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	32105	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	117.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% 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, 0.333 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, BRU, TT

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	0.53	0/11374	0.75	1/15383 (0.0%)
2	B	0.50	0/9010	0.73	3/12149 (0.0%)
3	C	0.49	0/2133	0.67	0/2891
4	D	0.62	2/1444 (0.1%)	0.76	1/1935 (0.1%)
5	E	0.44	0/1788	0.63	0/2406
6	F	0.64	0/691	0.88	1/933 (0.1%)
7	G	0.51	0/1368	0.71	0/1844
8	H	0.39	0/1086	0.63	0/1470
9	I	0.41	0/989	0.63	0/1331
10	J	0.50	0/541	0.79	2/727 (0.3%)
11	K	0.52	0/938	0.68	0/1267
12	L	0.51	0/365	0.83	0/485
13	N	0.94	0/276	1.92	11/424 (2.6%)
14	P	0.69	0/221	1.21	0/343
15	T	0.76	0/475	1.52	12/725 (1.7%)
All	All	0.52	2/32699 (0.0%)	0.77	31/44313 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
4	D	0	1
All	All	0	6

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	187	THR	CB-CG2	9.38	1.83	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	187	THR	CA-CB	8.33	1.75	1.53

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	N	3	DA	O4'-C1'-N9	11.17	115.82	108.00
15	T	18	DT	O4'-C1'-N1	11.08	115.76	108.00
13	N	3	DA	O4'-C1'-C2'	-11.01	97.09	105.90
13	N	3	DA	C1'-O4'-C4'	-9.53	100.57	110.10
4	D	187	THR	CA-CB-CG2	7.69	123.17	112.40

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	311	GLN	Peptide
1	A	34	LYS	Peptide
1	A	399	HIS	Peptide
1	A	55	ASP	Peptide
1	A	63	ARG	Peptide

## 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	11174	0	11233	432	0
2	B	8839	0	8876	286	0
3	C	2095	0	2051	72	0
4	D	1434	0	1460	64	0
5	E	1752	0	1776	53	0
6	F	679	0	701	30	0
7	G	1340	0	1357	42	0
8	H	1068	0	1040	50	0
9	I	971	0	927	23	0
10	J	532	0	542	24	0
11	K	920	0	929	28	0
12	L	363	0	386	14	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	N	247	0	137	14	0
14	P	197	0	99	28	0
15	T	485	0	273	22	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
17	A	1	0	0	0	0
All	All	32105	0	31787	1066	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 1066 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:D:187:THR:CA	4:D:187:THR:CB	1.75	1.60
4:D:187:THR:CG2	4:D:187:THR:CB	1.83	1.55
14:P:11:A:C8	14:P:11:A:H5''	1.77	1.18
1:A:1100:ARG:HH21	1:A:1351:GLU:HG2	1.11	1.14
14:P:11:A:C8	14:P:11:A:C5'	2.30	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	1414/1732 (82%)	1209 (86%)	145 (10%)	60 (4%)	3	18
2	B	1096/1224 (90%)	945 (86%)	120 (11%)	31 (3%)	5	24

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	264/318 (83%)	237 (90%)	23 (9%)	4 (2%)	10	36
4	D	174/221 (79%)	146 (84%)	15 (9%)	13 (8%)	1	7
5	E	212/215 (99%)	193 (91%)	14 (7%)	5 (2%)	6	28
6	F	82/155 (53%)	74 (90%)	7 (8%)	1 (1%)	13	41
7	G	169/171 (99%)	153 (90%)	13 (8%)	3 (2%)	8	32
8	H	129/146 (88%)	102 (79%)	19 (15%)	8 (6%)	1	10
9	I	117/122 (96%)	100 (86%)	12 (10%)	5 (4%)	2	17
10	J	63/70 (90%)	56 (89%)	4 (6%)	3 (5%)	2	15
11	K	113/120 (94%)	104 (92%)	9 (8%)	0	100	100
12	L	44/70 (63%)	31 (70%)	6 (14%)	7 (16%)	0	0
All	All	3877/4564 (85%)	3350 (86%)	387 (10%)	140 (4%)	3	21

5 of 140 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	GLU
1	A	48	ALA
1	A	57	ARG
1	A	58	LEU
1	A	74	MET

### 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	1240/1519 (82%)	1057 (85%)	183 (15%)	3	12
2	B	964/1061 (91%)	816 (85%)	148 (15%)	2	11
3	C	234/274 (85%)	206 (88%)	28 (12%)	5	19
4	D	160/200 (80%)	136 (85%)	24 (15%)	3	12
5	E	196/197 (100%)	178 (91%)	18 (9%)	9	31
6	F	74/137 (54%)	65 (88%)	9 (12%)	5	18

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	G	152/152 (100%)	128 (84%)	24 (16%)	2	10
8	H	117/128 (91%)	100 (86%)	17 (14%)	3	12
9	I	113/116 (97%)	101 (89%)	12 (11%)	6	24
10	J	60/65 (92%)	53 (88%)	7 (12%)	5	20
11	K	99/102 (97%)	88 (89%)	11 (11%)	6	22
12	L	40/57 (70%)	28 (70%)	12 (30%)	0	1
All	All	3449/4008 (86%)	2956 (86%)	493 (14%)	3	13

5 of 493 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	423	LYS
2	B	843	GLN
9	I	43	VAL
2	B	453	ILE
2	B	620	ARG

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

Mol	Chain	Res	Type
2	B	395	GLN
4	D	37	GLN
2	B	481	GLN
2	B	103	ASN
2	B	986	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	9/12 (75%)	2 (22%)	2 (22%)

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	4	A
14	P	11	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	P	3	G
14	P	10	A

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

2 non-standard protein/DNA/RNA residues are 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
15	BRU	T	22	15,14	15,21,22	2.10	2 (13%)	17,30,33	2.23	3 (17%)
15	TT	T	17	15	40,43,44	1.27	5 (12%)	59,69,72	2.09	15 (25%)

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
15	BRU	T	22	15,14	-	2/4/21/22	0/2/2/2
15	TT	T	17	15	-	9/18/105/106	0/5/6/6

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	22	BRU	C4-C5	7.38	1.47	1.38
15	T	17	TT	C2-N3	-3.21	1.32	1.38
15	T	17	TT	C1R-N1T	3.04	1.49	1.45
15	T	17	TT	C4-N3	-2.98	1.32	1.37
15	T	17	TT	C2-N1	2.85	1.42	1.36

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	T	17	TT	O4R-C1R-N1T	8.05	118.19	108.65

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	T	17	TT	C4-N3-C2	-6.30	116.98	126.67
15	T	22	BRU	C4-N3-C2	6.20	120.37	115.14
15	T	22	BRU	C2'-C1'-N1	-4.71	103.41	114.27
15	T	17	TT	O4T-C4T-C5T	-4.03	119.66	122.88

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	T	22	BRU	C3'-C4'-C5'-O5'
15	T	22	BRU	O4'-C4'-C5'-O5'
15	T	17	TT	O4'-C4'-C5'-O5'
15	T	17	TT	C3'-C4'-C5'-O5'
15	T	17	TT	C2R-C1R-N1T-C6T

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	T	17	TT	5	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates 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<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	1422/1732 (82%)	-0.07	21 (1%) 73 72	56, 101, 162, 298	0
2	B	1113/1224 (90%)	0.04	24 (2%) 62 60	53, 113, 185, 237	0
3	C	266/318 (83%)	-0.14	0 100 100	72, 102, 141, 189	0
4	D	178/221 (80%)	0.03	0 100 100	80, 116, 180, 211	0
5	E	214/215 (99%)	0.05	8 (3%) 41 40	77, 137, 188, 195	0
6	F	84/155 (54%)	-0.36	0 100 100	59, 80, 109, 134	0
7	G	171/171 (100%)	-0.06	0 100 100	63, 100, 139, 166	0
8	H	133/146 (91%)	0.47	6 (4%) 33 33	106, 151, 194, 218	0
9	I	119/122 (97%)	0.10	7 (5%) 22 23	107, 141, 182, 219	0
10	J	65/70 (92%)	-0.19	0 100 100	77, 106, 146, 157	0
11	K	115/120 (95%)	-0.13	1 (0%) 84 83	66, 101, 137, 163	0
12	L	46/70 (65%)	0.19	4 (8%) 10 12	91, 154, 198, 202	0
13	N	12/14 (85%)	-0.19	0 100 100	164, 177, 260, 301	0
14	P	9/12 (75%)	-0.27	0 100 100	132, 150, 197, 204	0
15	T	21/25 (84%)	-0.20	1 (4%) 30 31	125, 177, 286, 307	0
All	All	3968/4615 (85%)	-0.02	72 (1%) 68 67	53, 110, 179, 307	0

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	883	LEU	5.3
9	I	119	THR	4.9
2	B	250	PHE	4.6
8	H	139	ASN	4.4
9	I	120	GLN	4.1

## 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, 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	B-factors(Å <sup>2</sup> )	Q<0.9
15	TT	T	17	38/39	0.88	0.24	112,230,240,242	0
15	BRU	T	22	20/21	0.91	0.14	161,193,214,215	0

## 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
17	MG	A	2458	1/1	0.89	0.09	124,124,124,124	0
16	ZN	I	1122	1/1	0.99	0.05	196,196,196,196	0
16	ZN	A	2456	1/1	0.99	0.05	127,127,127,127	0
16	ZN	I	1121	1/1	1.00	0.12	120,120,120,120	0
16	ZN	L	1071	1/1	1.00	0.07	187,187,187,187	0
16	ZN	J	1066	1/1	1.00	0.21	90,90,90,90	0
16	ZN	C	1269	1/1	1.00	0.13	75,75,75,75	0
16	ZN	B	2225	1/1	1.00	0.18	75,75,75,75	0
16	ZN	A	2457	1/1	1.00	0.14	70,70,70,70	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.