



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 07:33 am BST

PDB ID : 5IP7
Title : Structure of RNA Polymerase II-Tfg1 peptide complex
Authors : Plaschka, C.; Hantsche, M.; Dienemann, C.; Burzinski, C.; Pitzko, J.; Cramer, P.
Deposited on : 2016-03-09
Resolution : 3.52 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 : 2.11
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

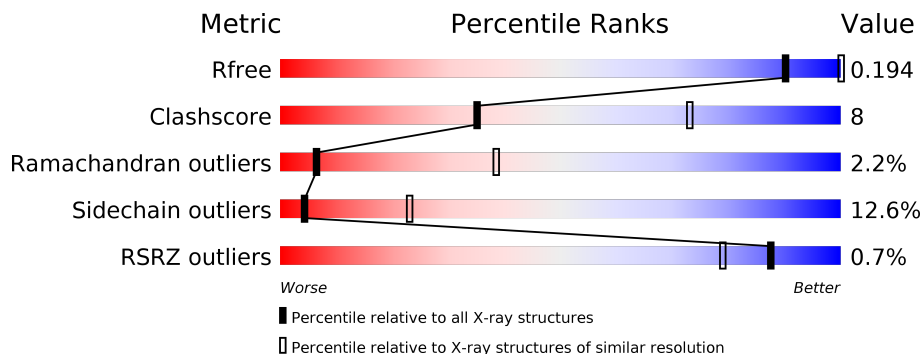
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.52 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	1161 (3.60-3.44)
Clashscore	141614	1244 (3.60-3.44)
Ramachandran outliers	138981	1206 (3.60-3.44)
Sidechain outliers	138945	1207 (3.60-3.44)
RSRZ outliers	127900	1080 (3.60-3.44)

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 ≥ 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	
2	B	1223	
3	C	266	
4	D	221	
5	E	214	
6	F	87	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	171	
8	H	145	
9	I	119	
10	J	65	
11	K	115	
12	L	46	
13	Q	15	

2 Entry composition [i](#)

There are 15 unique types of molecules in this entry. The entry contains 31339 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	1421	11190	7052	1957	2119	62	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	1118	8876	5620	1557	1644	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 polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	180	1440	890	256	291	3	0	0	0

- Molecule 5 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			
5	E	214	1752	1111	309	321	11	0	0	0

- Molecule 6 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			
6	F	87	705	451	119	132	3	0	0	0

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

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

- Molecule 8 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			
8	H	133	1068	673	180	211	4	0	0	0

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

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

- Molecule 10 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			
10	J	65	532	339	93	94	6	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	115	920	590	157	171	2	0	0	1

- Molecule 12 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			
12	L	46	363	224	72	63	4	0	0	0

- Molecule 13 is a protein called PHE-ILE-LYS-ARG-ASP-ARG-MET-ARG-ARG-ASN-PHE-LEU-ARG-MET-ARG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	Q	15	78	47	15	15	1	0	0	0

- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

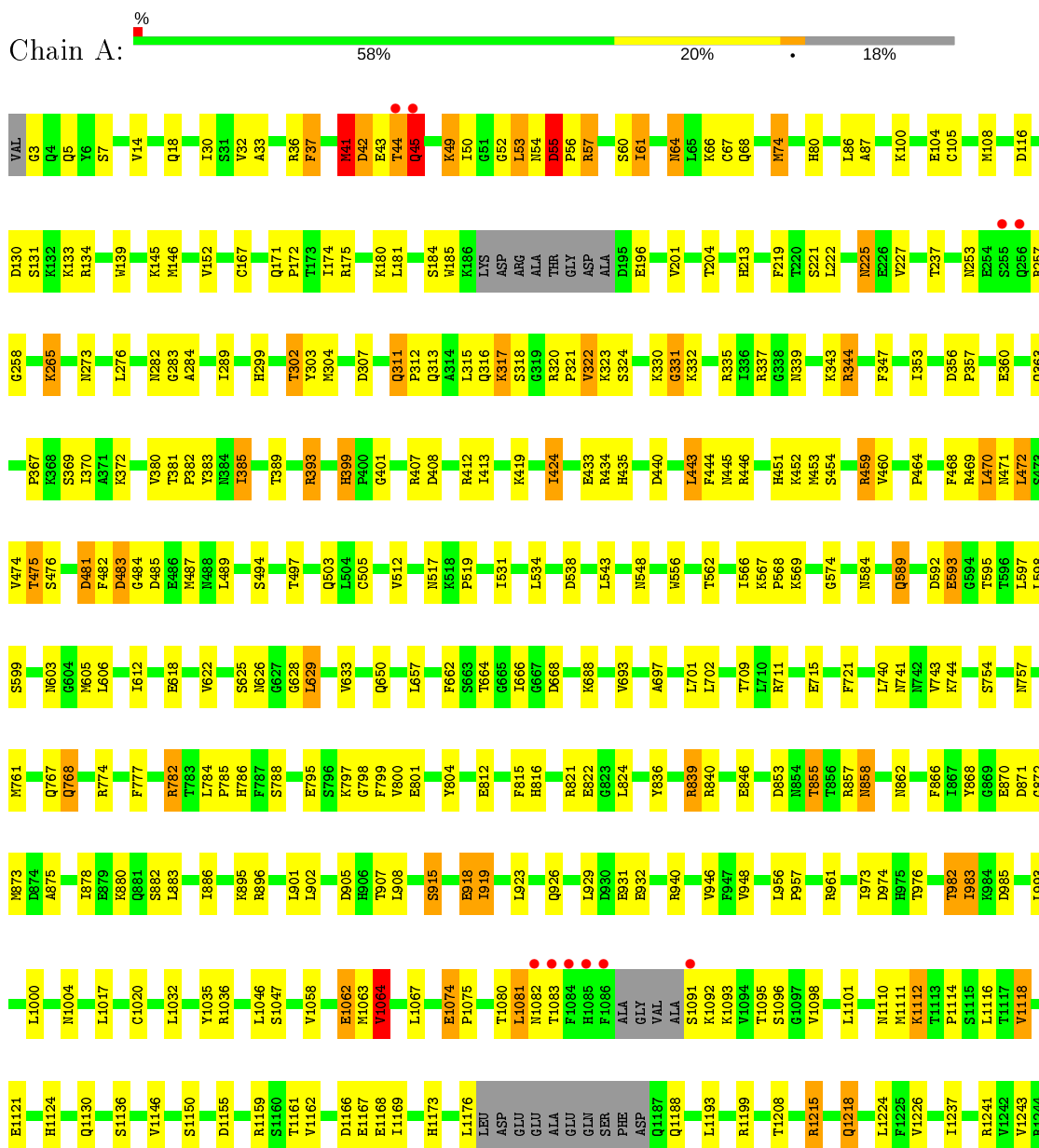
- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

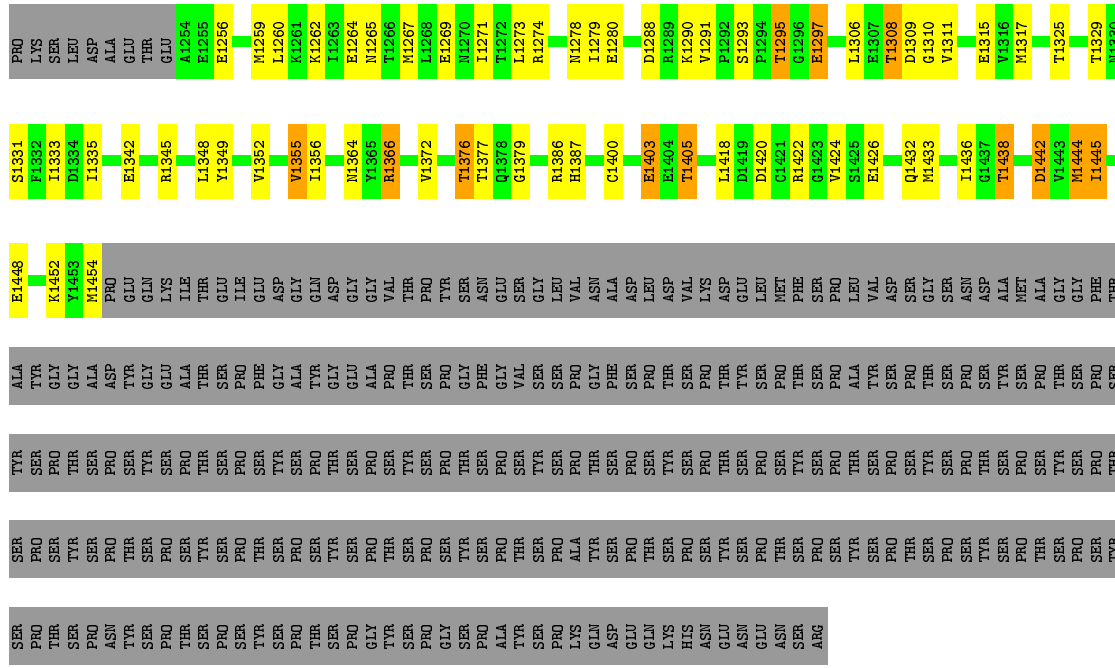
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	A	1	Total Mg 1 1	0	0

3 Residue-property plots [i](#)

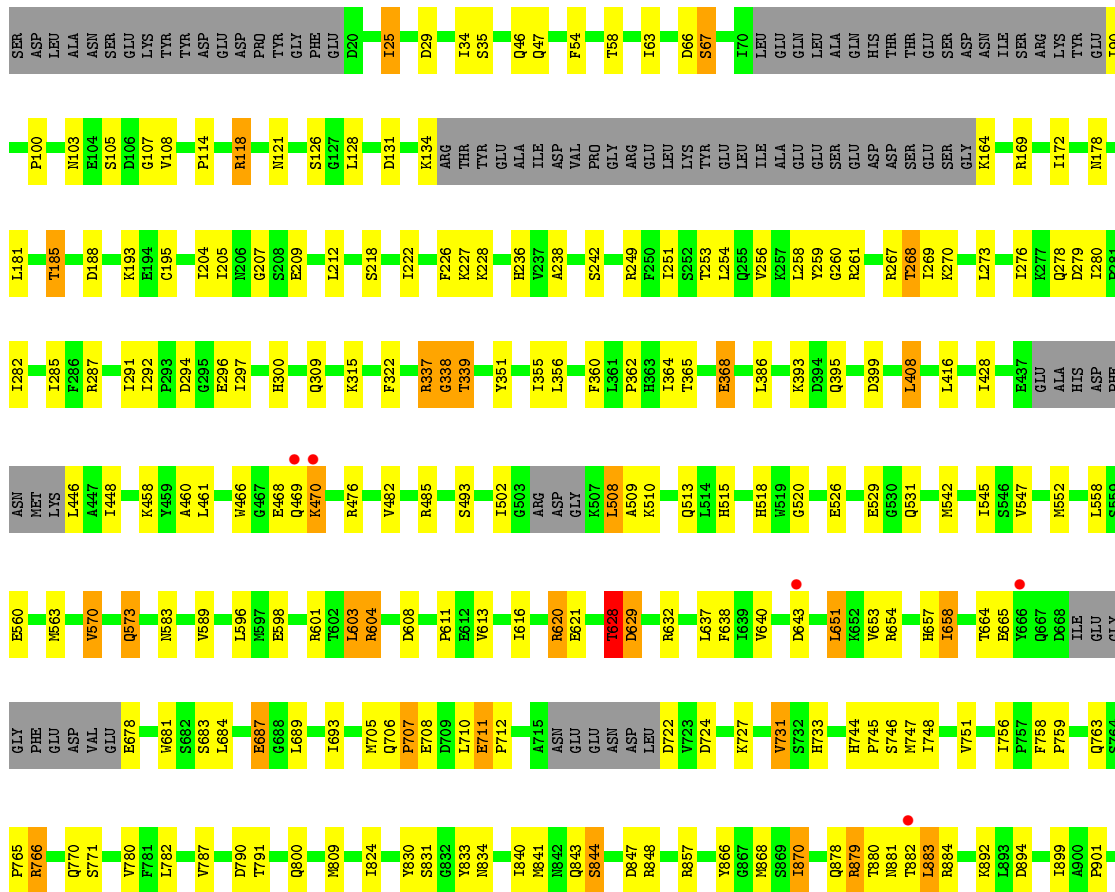
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

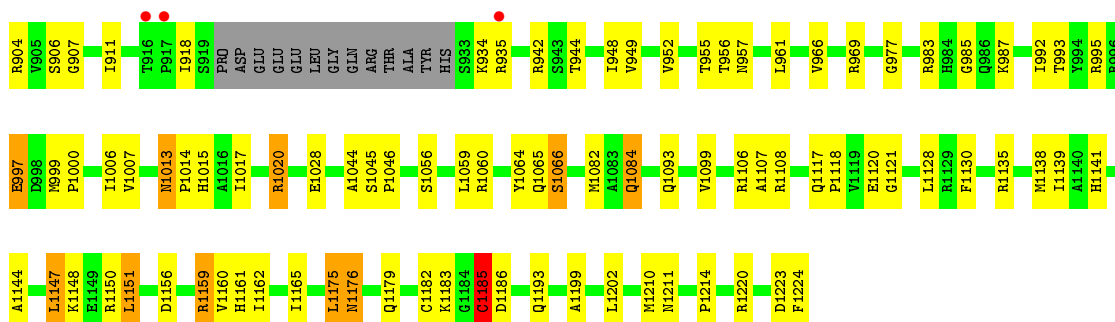
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1





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





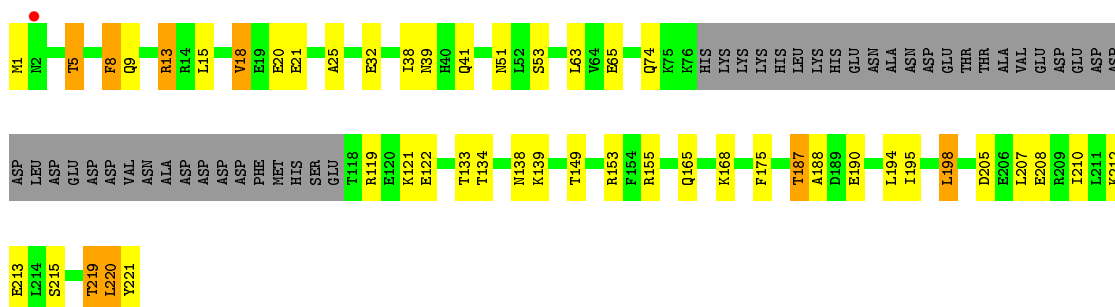
- Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C: 70% 23% 6%



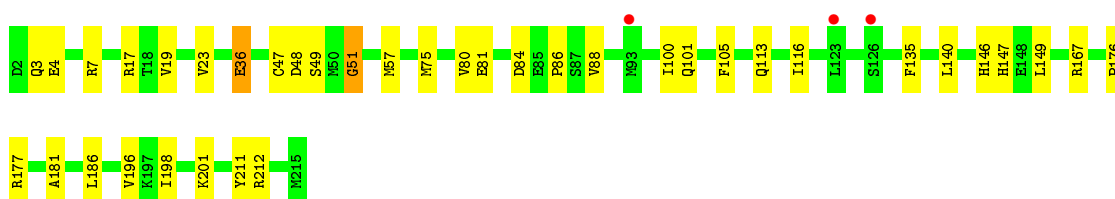
- Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D: 60% 18% 19%



- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 82% 17%



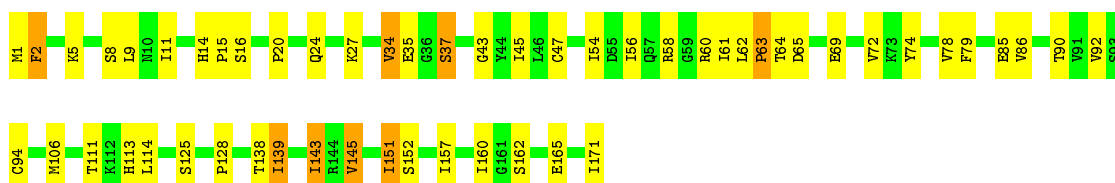
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:  71% 24% 5%



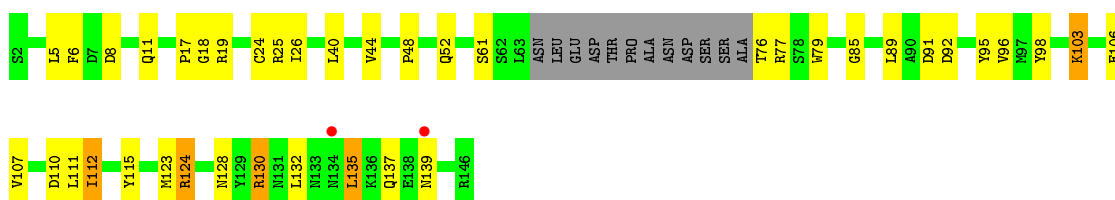
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G:  68% 27% 5%



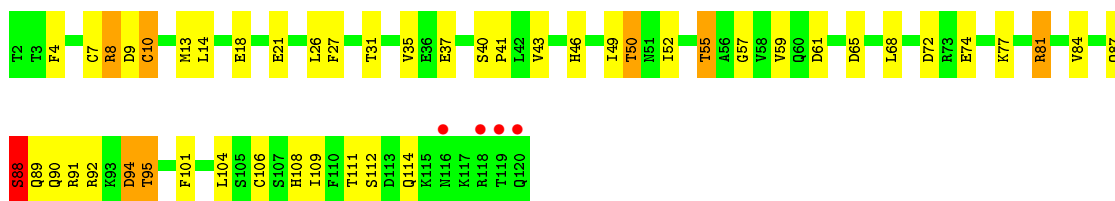
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:  64% 24% 8%



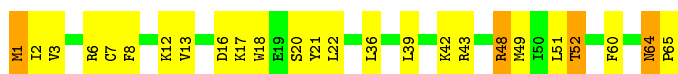
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:  60% 34% 6% 3%




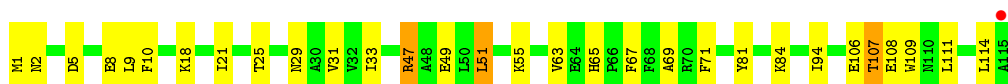
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:  62% 32% 6%



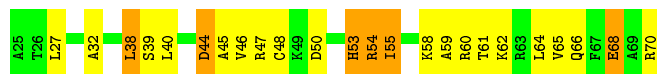
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain K:  74% 23%



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

Chain L:  48% 39% 13%



- Molecule 13: PHE-ILE-LYS-ARG-ASP-ARG-MET-ARG-ARG-ASN-PHE-LEU-ARG-MET-ARG

Chain Q:  93% 7%



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	221.22Å 392.73Å 282.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.93 – 3.52 48.70 – 3.52	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.93-3.52) 100.0 (48.70-3.52)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.57 (at 3.48Å)	Xtriage
Refinement program	BUSTER-TNT 2.10.2	Depositor
R, R_{free}	0.154 , 0.180 0.172 , 0.194	Depositor DCC
R_{free} test set	2983 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å ²)	111.0	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 101.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.007 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.008 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	31339	wwPDB-VP
Average B, all atoms (Å ²)	127.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.86% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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 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	0.55	4/11391 (0.0%)	0.81	10/15402 (0.1%)
2	B	0.50	0/9048	0.76	3/12200 (0.0%)
3	C	0.49	0/2133	0.78	1/2891 (0.0%)
4	D	0.52	0/1450	0.82	2/1945 (0.1%)
5	E	0.45	0/1788	0.71	0/2406
6	F	0.60	0/717	0.87	0/967
7	G	0.50	0/1368	0.76	0/1844
8	H	0.48	0/1086	0.75	0/1470
9	I	0.45	0/989	0.78	1/1331 (0.1%)
10	J	0.55	0/541	0.86	0/727
11	K	0.47	0/938	0.68	0/1267
12	L	0.51	0/365	0.87	0/485
13	Q	0.70	0/77	0.71	0/105
All	All	0.52	4/31891 (0.0%)	0.78	17/43040 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	57	ARG	CA-CB	5.50	1.66	1.53
1	A	57	ARG	CA-C	5.27	1.66	1.52
1	A	57	ARG	CG-CD	5.27	1.65	1.51
1	A	57	ARG	CB-CG	5.09	1.66	1.52

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	25	ALA	C-N-CA	7.60	140.70	121.70
1	A	41	MET	C-N-CA	7.17	139.63	121.70
1	A	399	HIS	N-CA-CB	7.15	123.47	110.60
1	A	57	ARG	CA-CB-CG	7.14	129.12	113.40
2	B	628	THR	C-N-CA	6.08	136.89	121.70

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	11190	0	11255	220	0
2	B	8876	0	8919	145	0
3	C	2095	0	2051	49	0
4	D	1440	0	1456	16	0
5	E	1752	0	1776	19	0
6	F	705	0	731	17	0
7	G	1340	0	1357	32	0
8	H	1068	0	1040	19	0
9	I	971	0	927	23	0
10	J	532	0	542	19	0
11	K	920	0	929	25	0
12	L	363	0	386	10	0
13	Q	78	0	36	0	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	31339	0	31405	517	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 517 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:30:ILE:CG1	1:A:30:ILE:CD1	1.75	1.56
1:A:56:PRO:HB2	1:A:68:GLN:HG3	1.37	1.06
1:A:1081:LEU:HB3	1:A:1082:ASN:HA	1.45	0.98

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:800:GLN:HB3	10:J:52:THR:HG23	1.48	0.92
1:A:855:THR:HG21	1:A:857:ARG:HE	1.34	0.92

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	1411/1732 (82%)	1269 (90%)	116 (8%)	26 (2%)	8	42
2	B	1102/1223 (90%)	996 (90%)	79 (7%)	27 (2%)	5	35
3	C	264/266 (99%)	243 (92%)	18 (7%)	3 (1%)	14	53
4	D	176/221 (80%)	155 (88%)	17 (10%)	4 (2%)	6	37
5	E	212/214 (99%)	198 (93%)	10 (5%)	4 (2%)	8	41
6	F	85/87 (98%)	79 (93%)	6 (7%)	0	100	100
7	G	169/171 (99%)	162 (96%)	5 (3%)	2 (1%)	13	51
8	H	129/145 (89%)	105 (81%)	17 (13%)	7 (5%)	2	19
9	I	117/119 (98%)	98 (84%)	17 (14%)	2 (2%)	9	44
10	J	63/65 (97%)	56 (89%)	4 (6%)	3 (5%)	2	21
11	K	113/115 (98%)	109 (96%)	4 (4%)	0	100	100
12	L	44/46 (96%)	33 (75%)	6 (14%)	5 (11%)	0	6
13	Q	13/15 (87%)	11 (85%)	1 (8%)	1 (8%)	1	11
All	All	3898/4419 (88%)	3514 (90%)	300 (8%)	84 (2%)	6	38

5 of 84 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	42	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	45	GLN
1	A	74	MET
1	A	593	GLU
1	A	1403	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	1244/1519 (82%)	1084 (87%)	160 (13%)	4	22
2	B	967/1060 (91%)	853 (88%)	114 (12%)	5	26
3	C	234/234 (100%)	205 (88%)	29 (12%)	4	24
4	D	160/200 (80%)	132 (82%)	28 (18%)	2	11
5	E	196/196 (100%)	187 (95%)	9 (5%)	27	61
6	F	77/77 (100%)	67 (87%)	10 (13%)	4	22
7	G	152/152 (100%)	132 (87%)	20 (13%)	4	22
8	H	117/127 (92%)	104 (89%)	13 (11%)	6	29
9	I	113/113 (100%)	93 (82%)	20 (18%)	2	11
10	J	60/60 (100%)	50 (83%)	10 (17%)	2	13
11	K	99/99 (100%)	89 (90%)	10 (10%)	7	33
12	L	40/40 (100%)	26 (65%)	14 (35%)	0	1
13	Q	1/15 (7%)	1 (100%)	0	100	100
All	All	3460/3892 (89%)	3023 (87%)	437 (13%)	4	23

5 of 437 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	529	GLU
2	B	997	GLU
10	J	1	MET
2	B	570	VAL
2	B	766	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	590	HIS
2	B	957	ASN
8	H	137	GLN
2	B	648	HIS
2	B	842	ASN

5.3.3 RNA [i](#)

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 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

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	1421/1732 (82%)	-0.28	10 (0%) 87 79	64, 112, 182, 240	0
2	B	1118/1223 (91%)	-0.19	8 (0%) 87 79	68, 126, 195, 228	0
3	C	266/266 (100%)	-0.26	0 100 100	80, 116, 160, 192	0
4	D	180/221 (81%)	-0.24	1 (0%) 89 81	97, 131, 188, 211	0
5	E	214/214 (100%)	-0.19	3 (1%) 75 62	91, 157, 206, 216	0
6	F	87/87 (100%)	-0.51	0 100 100	69, 92, 124, 138	0
7	G	171/171 (100%)	-0.25	0 100 100	83, 112, 149, 185	0
8	H	133/145 (91%)	0.07	2 (1%) 73 61	118, 160, 196, 204	0
9	I	119/119 (100%)	-0.16	4 (3%) 45 34	123, 153, 195, 219	0
10	J	65/65 (100%)	-0.35	0 100 100	89, 113, 155, 173	0
11	K	115/115 (100%)	-0.26	1 (0%) 84 73	79, 112, 160, 181	0
12	L	46/46 (100%)	-0.12	0 100 100	97, 155, 190, 204	0
13	Q	15/15 (100%)	0.38	0 100 100	132, 200, 214, 222	0
All	All	3950/4419 (89%)	-0.23	29 (0%) 87 79	64, 122, 191, 240	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	882	THR	7.3
1	A	1085	HIS	5.3
1	A	1082	ASN	5.0
1	A	44	THR	4.1
11	K	115	ALA	3.8

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

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. 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
15	MG	A	1803	1/1	0.96	0.44	62,62,62,62	0
14	ZN	L	101	1/1	0.98	0.09	151,151,151,151	0
14	ZN	A	1801	1/1	0.98	0.06	164,164,164,164	0
14	ZN	I	202	1/1	0.98	0.04	209,209,209,209	0
14	ZN	J	101	1/1	0.99	0.26	110,110,110,110	0
14	ZN	I	201	1/1	0.99	0.12	125,125,125,125	0
14	ZN	C	301	1/1	1.00	0.10	94,94,94,94	0
14	ZN	A	1802	1/1	1.00	0.15	97,97,97,97	0
14	ZN	B	1301	1/1	1.00	0.22	115,115,115,115	0

6.5 Other polymers [i](#)

There are no such residues in this entry.