wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2024 – 01:45 pm BST

PDB ID	:	4C3H
Title	:	Structure of 14-subunit RNA polymerase I at 3.27 A resolution, crystal form C2-93
Authors	:	Fernandez-Tornero, C.; Moreno-Morcillo, M.; Rashid, U.J.; Taylor, N.M.I.; Ruiz, F.M.; Gruene, T.; Legrand, P.; Steuerwald, U.; Muller, C.W.
Deposited on	:	2013-08-24
Resolution	:	3.27 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 (1)) were used in the production of this report:

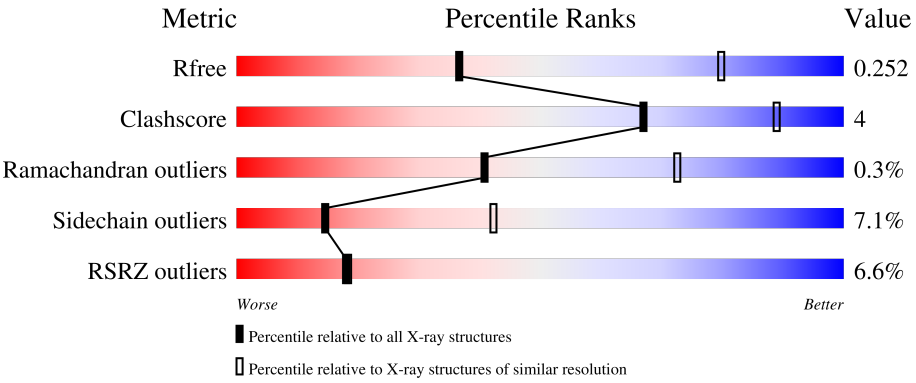
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.27 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.



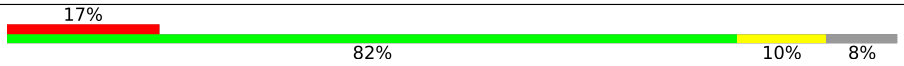
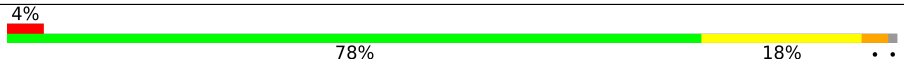
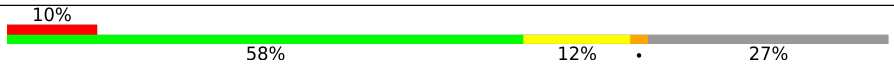
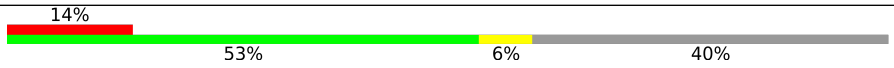
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1177 (3.32-3.24)
Clashscore	141614	1044 (3.30-3.26)
Ramachandran outliers	138981	1026 (3.30-3.26)
Sidechain outliers	138945	1025 (3.30-3.26)
RSRZ outliers	127900	1141 (3.32-3.24)

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	A	1664	3% 77% 14% · 8%
2	B	1203	5% 80% 16% ··
3	C	335	5% 74% 17% 9%
4	D	137	4% 36% 7% 57%
5	E	215	6% 92% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	F	155	
7	G	326	
8	H	146	
9	I	125	
10	J	70	
11	K	142	
12	L	70	
13	M	415	
14	N	233	

2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 34552 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 I SUBUNIT RPA190.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1523	Total	C	N	O	S	0	0	0
			12019	7577	2086	2292	64			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA135.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1176	Total	C	N	O	S	0	0	0
			9322	5898	1629	1745	50			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	304	Total	C	N	O	S	0	0	0
			2418	1536	414	460	8			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA14.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	59	Total	C	N	O	0	0	0
			466	292	80	94			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	12	SER	THR	conflict	UNP P50106

- 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	215	Total	C	N	O	S	0	0	0
			1759	1116	310	321	12			

- 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	100	Total	C	N	O	S	0	0	0
			823	522	144	154	3			

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	259	Total	C	N	O	S	0	0	0
			2052	1301	348	398	5			

- 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	134	Total	C	N	O	S	0	0	0
			1072	676	181	211	4			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	124	Total	C	N	O	S	0	0	0
			942	584	160	189	9			

- 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	69	Total	C	N	O	S	0	0	0
			569	362	101	100	6			

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	103	Total	C	N	O	S	0	0	0
			810	506	132	167	5			

- 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	45	Total	C	N	O	S	0	0	0
			359	221	71	63	4			

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA49.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
13	M	105	Total	C	N	O	0	0	0
			831	528	137	166			

- Molecule 14 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA34.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	139	Total	C	N	O	S	0	0	0
			1103	706	179	214	4			

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

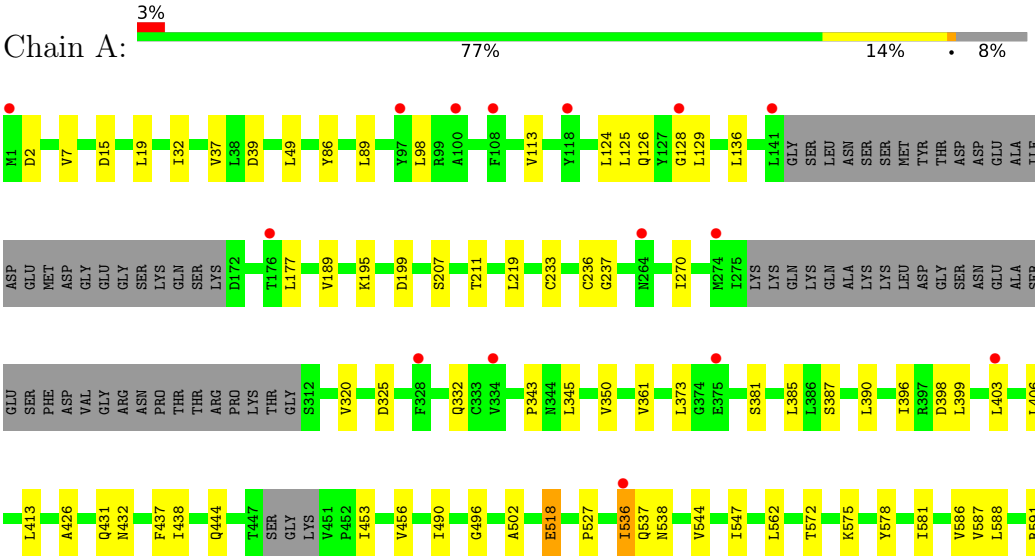
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	2	Total	Zn	0	0
			2	2		
15	B	1	Total	Zn	0	0
			1	1		
15	I	2	Total	Zn	0	0
			2	2		
15	J	1	Total	Zn	0	0
			1	1		
15	L	1	Total	Zn	0	0
			1	1		

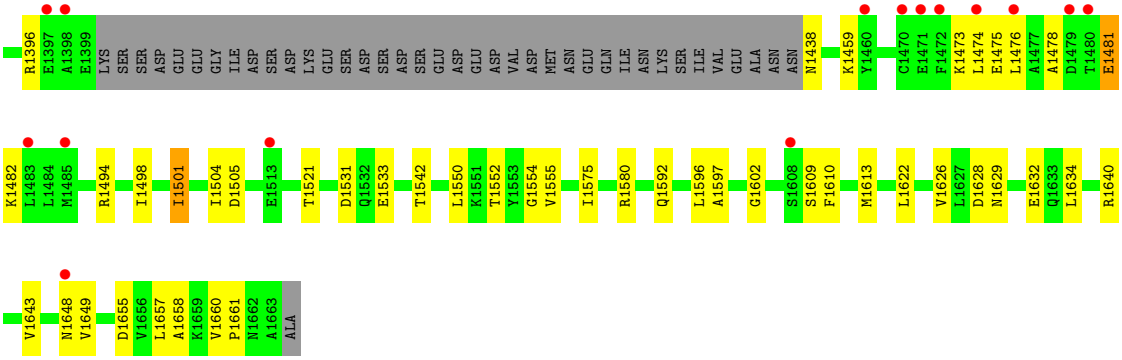
3 Residue-property plots (i)

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: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA190

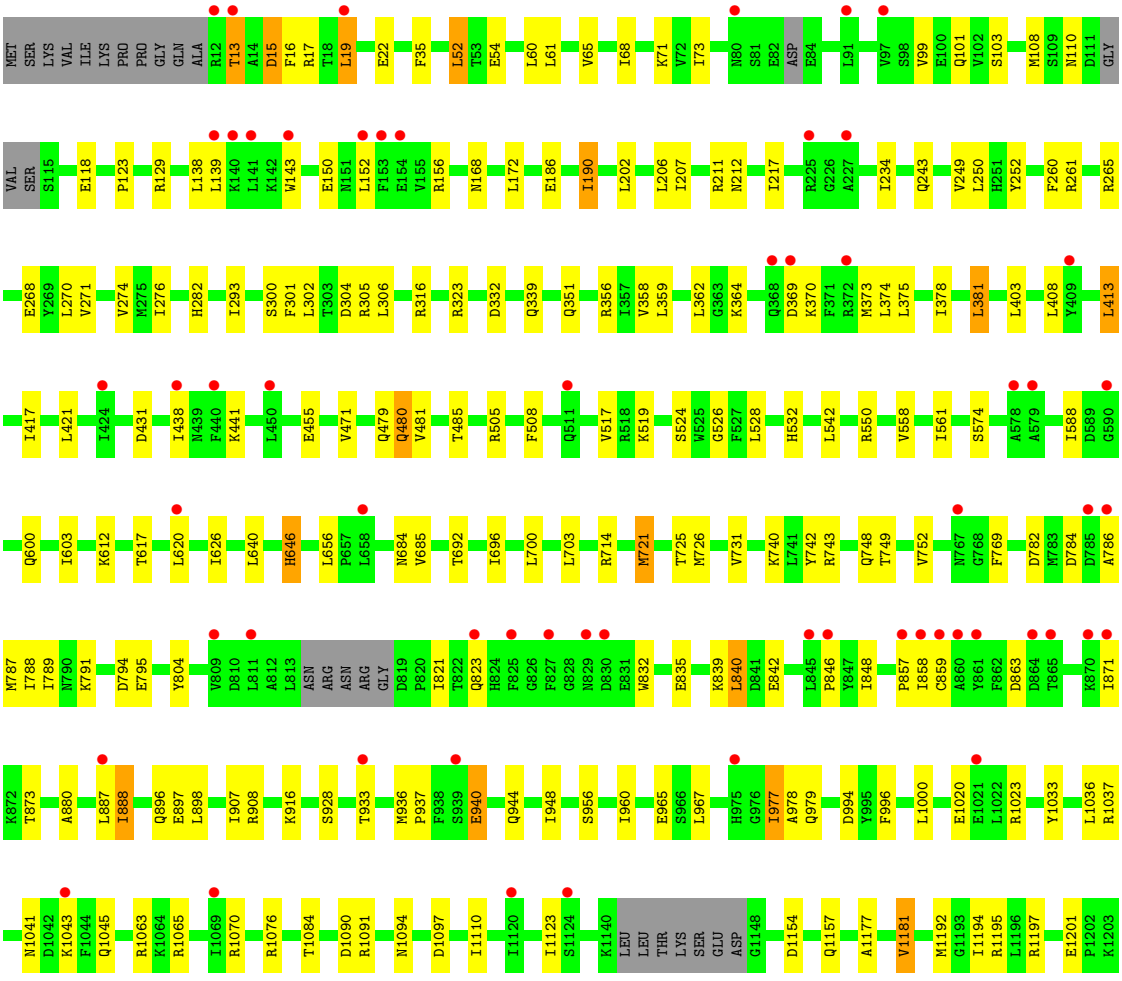
Chain A:





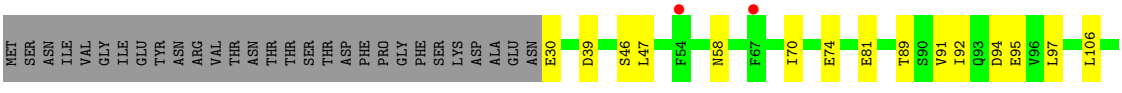
• Molecule 2: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA135

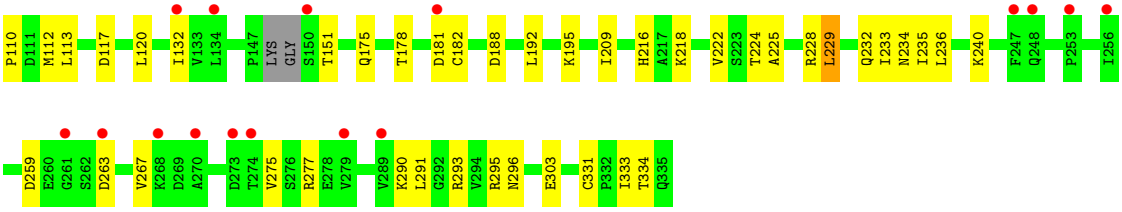
Chain B:



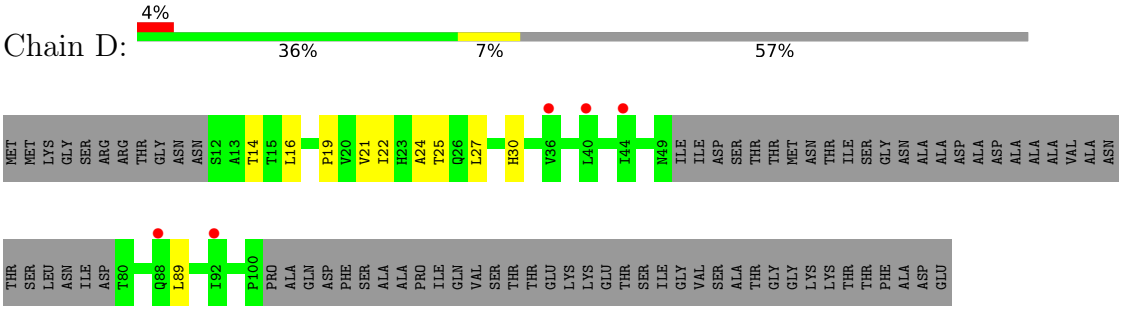
• Molecule 3: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC1

Chain C:

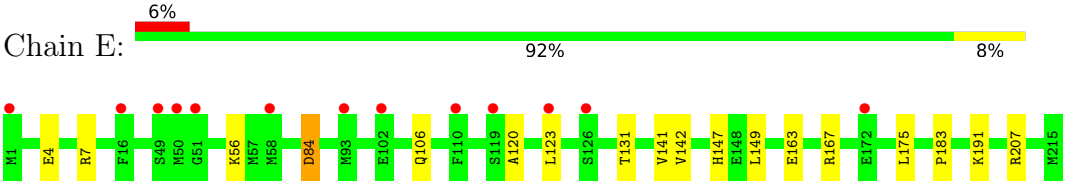




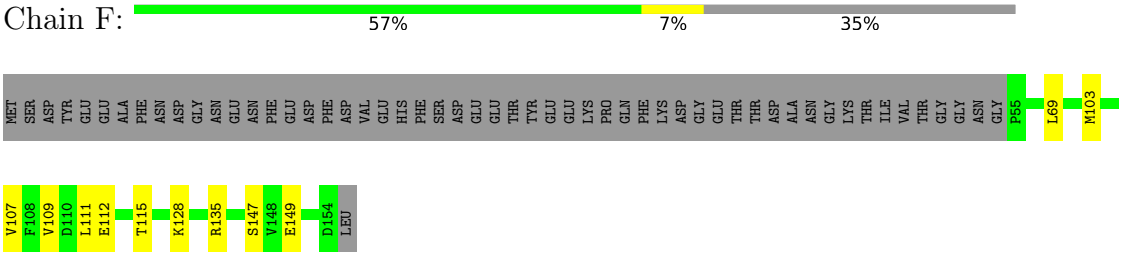
• Molecule 4: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA14



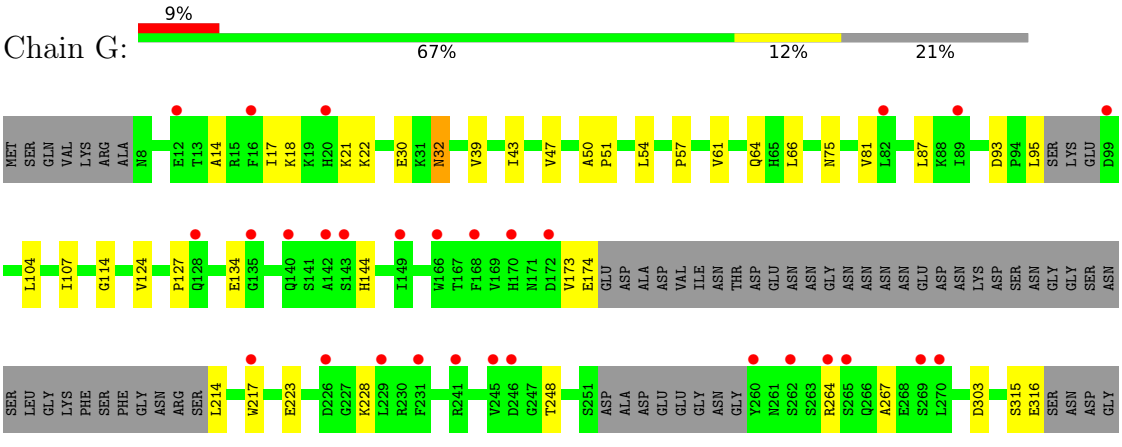
• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1



• Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

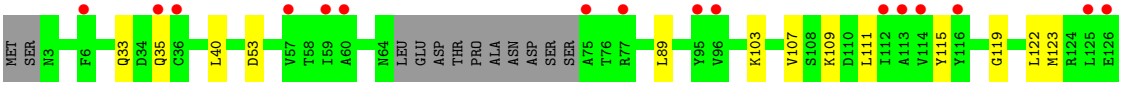


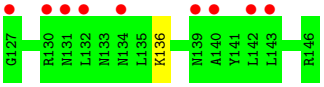
• Molecule 7: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43



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

Chain H:



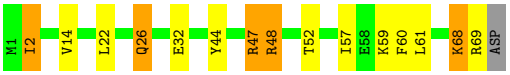


• Molecule 9: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA12

Chain I:

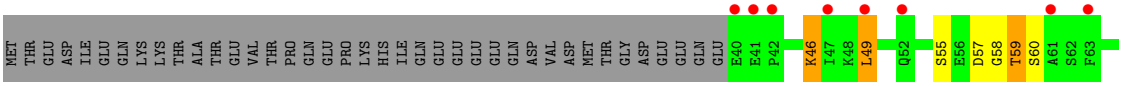
• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

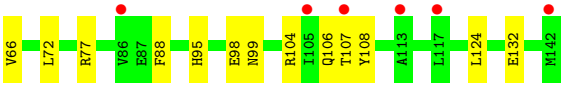
Chain J:



• Molecule 11: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC2

Chain K:



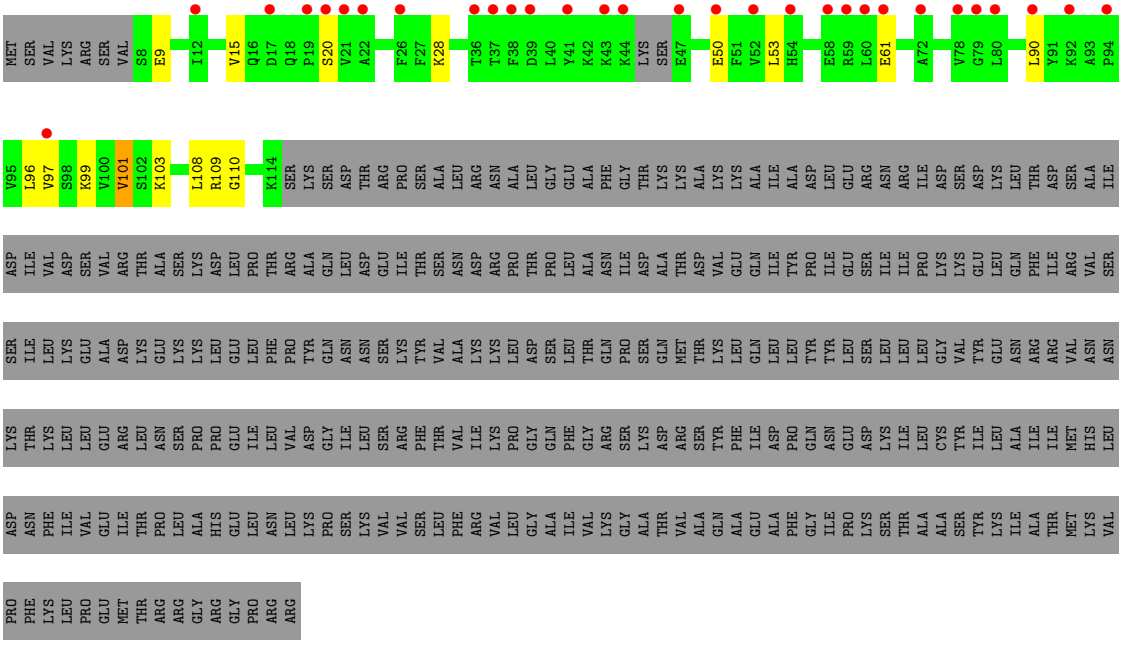


• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain L:

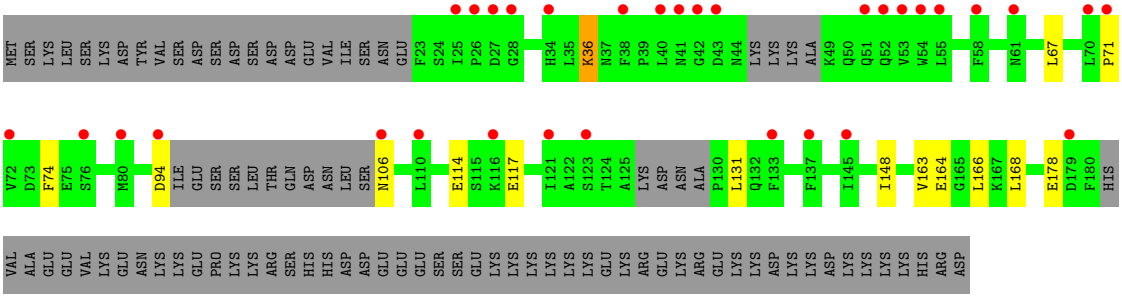
• Molecule 13: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA49

Chain M:



• Molecule 14: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA34

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	425.24Å 140.62Å 139.72Å 90.00° 93.35° 90.00°	Depositor
Resolution (Å)	47.88 – 3.27 47.78 – 3.27	Depositor EDS
% Data completeness (in resolution range)	97.7 (47.88-3.27) 97.7 (47.78-3.27)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 3.25Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.221 , 0.226 0.249 , 0.252	Depositor DCC
R_{free} test set	6209 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	135.5	Xtriage
Anisotropy	0.369	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 101.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	34552	wwPDB-VP
Average B, all atoms (Å^2)	171.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% 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: 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	A	0.41	0/12233	0.58	0/16523
2	B	0.41	0/9527	0.60	0/12879
3	C	0.40	0/2469	0.60	0/3347
4	D	0.39	0/472	0.52	0/639
5	E	0.41	0/1795	0.56	0/2416
6	F	0.40	0/838	0.54	0/1129
7	G	0.39	0/2094	0.58	0/2843
8	H	0.39	0/1090	0.57	0/1476
9	I	0.39	0/953	0.56	0/1282
10	J	0.41	0/578	0.62	0/775
11	K	0.39	0/821	0.59	0/1108
12	L	0.39	0/361	0.60	0/478
13	M	0.39	0/846	0.53	0/1136
14	N	0.39	0/1124	0.52	0/1512
All	All	0.40	0/35201	0.58	0/47543

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

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	12019	0	12051	127	0
2	B	9322	0	9187	92	0
3	C	2418	0	2401	21	0
4	D	466	0	466	3	0
5	E	1759	0	1788	8	0
6	F	823	0	841	6	0
7	G	2052	0	2016	15	0
8	H	1072	0	1042	6	0
9	I	942	0	928	14	0
10	J	569	0	585	12	0
11	K	810	0	801	12	0
12	L	359	0	381	1	0
13	M	831	0	820	10	0
14	N	1103	0	1106	6	0
15	A	2	0	0	0	0
15	B	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	34552	0	34413	272	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 4.

The worst 5 of 272 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)
11:K:59:THR:HG22	11:K:107:THR:OG1	1.46	1.13
11:K:60:SER:OG	11:K:106:GLN:HG2	1.62	0.99
1:A:1382:VAL:HA	2:B:1070:ARG:NH1	1.78	0.97
2:B:99:VAL:HG21	2:B:417:ILE:HD11	1.58	0.85
3:C:222:VAL:HG21	3:C:225:ALA:HB2	1.68	0.75

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	1503/1664 (90%)	1421 (94%)	78 (5%)	4 (0%)	41 72
2	B	1166/1203 (97%)	1095 (94%)	67 (6%)	4 (0%)	41 72
3	C	300/335 (90%)	283 (94%)	17 (6%)	0	100 100
4	D	55/137 (40%)	52 (94%)	3 (6%)	0	100 100
5	E	213/215 (99%)	202 (95%)	11 (5%)	0	100 100
6	F	98/155 (63%)	96 (98%)	2 (2%)	0	100 100
7	G	251/326 (77%)	232 (92%)	17 (7%)	2 (1%)	19 52
8	H	130/146 (89%)	119 (92%)	11 (8%)	0	100 100
9	I	118/125 (94%)	104 (88%)	12 (10%)	2 (2%)	9 37
10	J	67/70 (96%)	62 (92%)	5 (8%)	0	100 100
11	K	101/142 (71%)	95 (94%)	5 (5%)	1 (1%)	15 48
12	L	43/70 (61%)	40 (93%)	3 (7%)	0	100 100
13	M	101/415 (24%)	94 (93%)	7 (7%)	0	100 100
14	N	131/233 (56%)	122 (93%)	9 (7%)	0	100 100
All	All	4277/5236 (82%)	4017 (94%)	247 (6%)	13 (0%)	41 72

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1385	ASP
1	A	237	GLY
1	A	1389	GLU
2	B	1154	ASP
9	I	78	ASP

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	1343/1465 (92%)	1271 (95%)	72 (5%)	22 53
2	B	1024/1053 (97%)	927 (90%)	97 (10%)	8 29
3	C	269/296 (91%)	242 (90%)	27 (10%)	7 28
4	D	56/116 (48%)	50 (89%)	6 (11%)	6 25
5	E	197/197 (100%)	192 (98%)	5 (2%)	47 72
6	F	90/137 (66%)	86 (96%)	4 (4%)	28 59
7	G	234/291 (80%)	219 (94%)	15 (6%)	17 47
8	H	116/128 (91%)	111 (96%)	5 (4%)	29 59
9	I	109/110 (99%)	99 (91%)	10 (9%)	9 31
10	J	64/65 (98%)	56 (88%)	8 (12%)	4 19
11	K	93/130 (72%)	86 (92%)	7 (8%)	13 39
12	L	40/57 (70%)	37 (92%)	3 (8%)	13 39
13	M	94/371 (25%)	90 (96%)	4 (4%)	29 59
14	N	128/220 (58%)	117 (91%)	11 (9%)	10 34
All	All	3857/4636 (83%)	3583 (93%)	274 (7%)	14 42

5 of 274 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	24	LEU
9	I	73	LYS
13	M	101	VAL
2	B	268	GLU
2	B	234	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	361	HIS
8	H	35	GLN
2	B	896	GLN
7	G	150	HIS
11	K	106	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 7 ligands modelled in this entry, 7 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

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
9	I	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	I	65:SER	C	66:VAL	N	3.24
1	I	51:THR	C	52:ALA	N	2.96

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	A	1523/1664 (91%)	0.28	56 (3%) 41 39	110, 151, 210, 241	0
2	B	1176/1203 (97%)	0.46	59 (5%) 28 27	111, 155, 215, 239	0
3	C	304/335 (90%)	0.38	18 (5%) 22 22	147, 180, 216, 228	0
4	D	59/137 (43%)	0.11	5 (8%) 10 11	154, 230, 241, 253	0
5	E	215/215 (100%)	0.26	13 (6%) 21 21	131, 191, 235, 241	0
6	F	100/155 (64%)	0.00	0 100 100	121, 155, 195, 205	0
7	G	259/326 (79%)	0.56	29 (11%) 5 5	135, 213, 249, 261	0
8	H	134/146 (91%)	0.81	25 (18%) 1 1	153, 190, 223, 233	0
9	I	124/125 (99%)	0.18	5 (4%) 38 36	130, 171, 227, 234	0
10	J	69/70 (98%)	0.17	0 100 100	144, 160, 191, 218	0
11	K	103/142 (72%)	0.50	14 (13%) 3 3	133, 180, 208, 231	0
12	L	45/70 (64%)	-0.12	0 100 100	157, 192, 219, 224	0
13	M	105/415 (25%)	1.30	30 (28%) 0 0	214, 235, 251, 258	0
14	N	139/233 (59%)	1.13	32 (23%) 0 1	152, 236, 266, 275	0
All	All	4355/5236 (83%)	0.41	286 (6%) 18 18	110, 165, 235, 275	0

The worst 5 of 286 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
14	N	27	ASP	7.7
1	A	706	HIS	7.6
13	M	22	ALA	6.5
14	N	121	ILE	6.2
13	M	17	ASP	6.2

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 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	I	1126	1/1	0.95	0.16	195,195,195,195	0
15	ZN	A	2664	1/1	0.96	0.12	182,182,182,182	0
15	ZN	L	1071	1/1	0.97	0.14	178,178,178,178	0
15	ZN	A	2665	1/1	0.98	0.13	137,137,137,137	0
15	ZN	I	1127	1/1	0.98	0.12	136,136,136,136	0
15	ZN	B	2204	1/1	0.98	0.17	144,144,144,144	0
15	ZN	J	1070	1/1	1.00	0.26	148,148,148,148	0

6.5 Other polymers (i)

There are no such residues in this entry.