wwPDB EM Map/Model Validation Report (i)

Jul 21, 2016 – 04:03 PM EDT

PDB ID	:	5G5L
EMDB ID:	:	EMD-3439
Title	:	RNA polymerase I-Rrn3 complex at 4.8 A resolution
Authors	:	Engel, C.; Plitzko, J.; Cramer, P.
Deposited on	:	2016-05-26
Resolution	:	4.80 Å(reported)
Based on PDB ID	:	4C2M

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

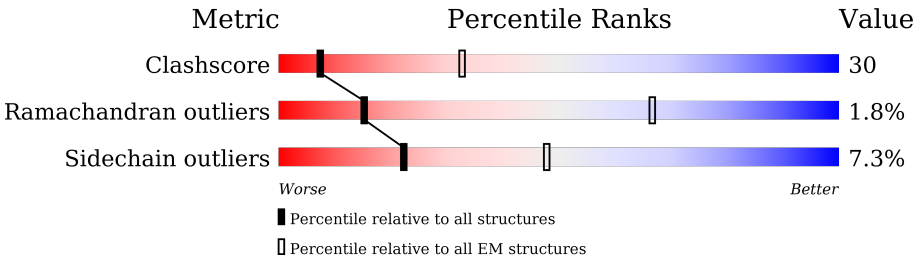
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027790

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.80 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.



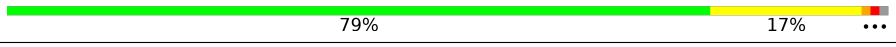
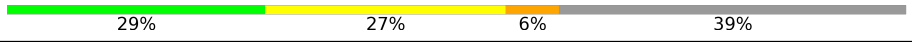
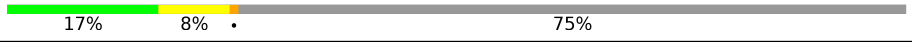
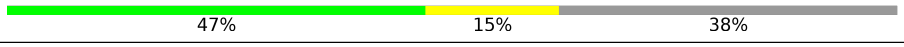
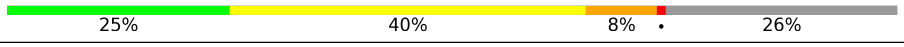
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1664	63% 23% • 11%
2	B	1203	69% 24% • •
3	C	335	61% 26% • • 9%
4	D	137	23% 18% •• 58%
5	E	215	77% 18% • •
6	F	155	53% 10% • 35%
7	G	326	39% 17% • 41%
8	H	146	72% 16% • 10%
9	I	125	54% 30% • 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	70	
11	K	142	
12	L	70	
13	M	415	
14	N	233	
15	O	627	

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 37349 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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					AltConf	Trace
			Total	C	N	O	S		
1	A	1480	11686	7384	2030	2211	61	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1174	9327	5899	1635	1743	50	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	305	2423	1539	416	460	8	0	0

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	D	58	459	289	78	92	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	212	1735	1102	306	316	11	0	0

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	100	Total	C	N	O	S	0	0
			823	522	144	154	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	193	Total	C	N	O	S	0	0
			1520	982	259	274	5		

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	131	Total	C	N	O	S	0	0
			1052	664	176	208	4		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	107	Total	C	N	O	S	0	0
			820	511	138	162	9		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	69	Total	C	N	O	S	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					AltConf	Trace
11	K	101	Total	C	N	O	S	0	0
			793	496	130	162	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	43	Total	C	N	O	S	0	0
			340	211	66	59	4		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
13	M	105	Total	C	N	O	0	0
			833	528	138	167		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	145	Total	C	N	O	S	0	0
			1151	735	188	224	4		

- Molecule 15 is a protein called RNA POLYMERASE I-SPECIFIC TRANSCRIPTION INITIATION FACTOR RRN3.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	463	Total	C	N	O	S	0	0
			3811	2473	623	694	21		

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

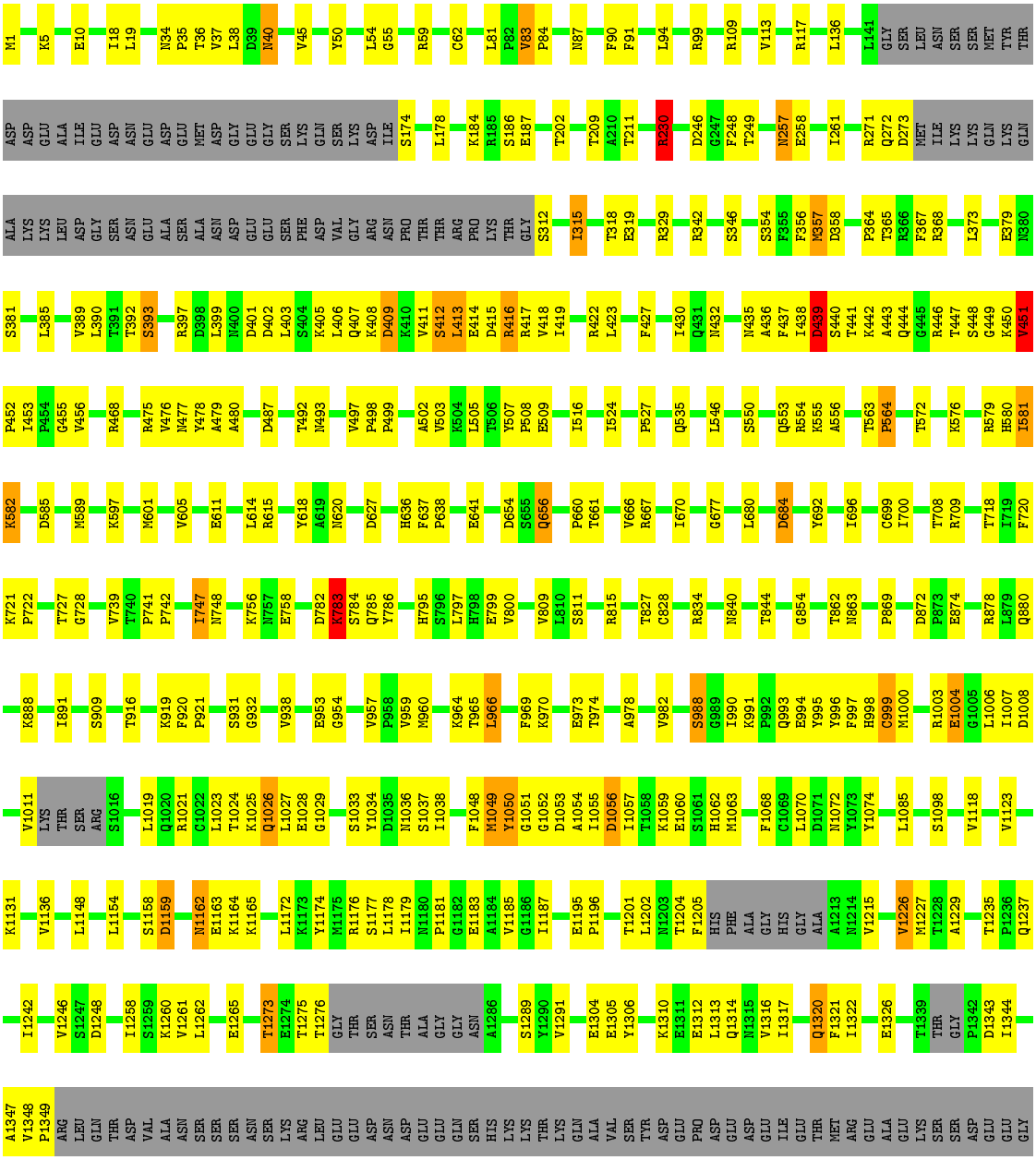
Mol	Chain	Residues	Atoms		AltConf
16	B	1	Total	Zn	0
			1	1	
16	A	2	Total	Zn	0
			2	2	
16	L	1	Total	Zn	0
			1	1	
16	J	1	Total	Zn	0
			1	1	
16	I	2	Total	Zn	0
			2	2	

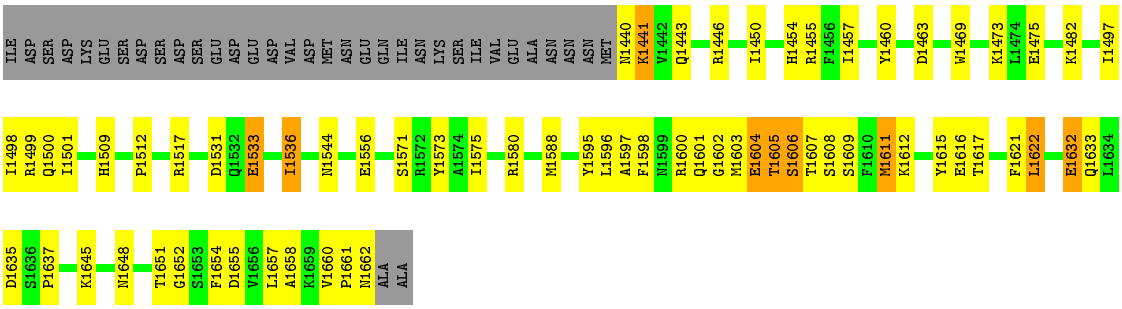
3 Residue-property plots (i)

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

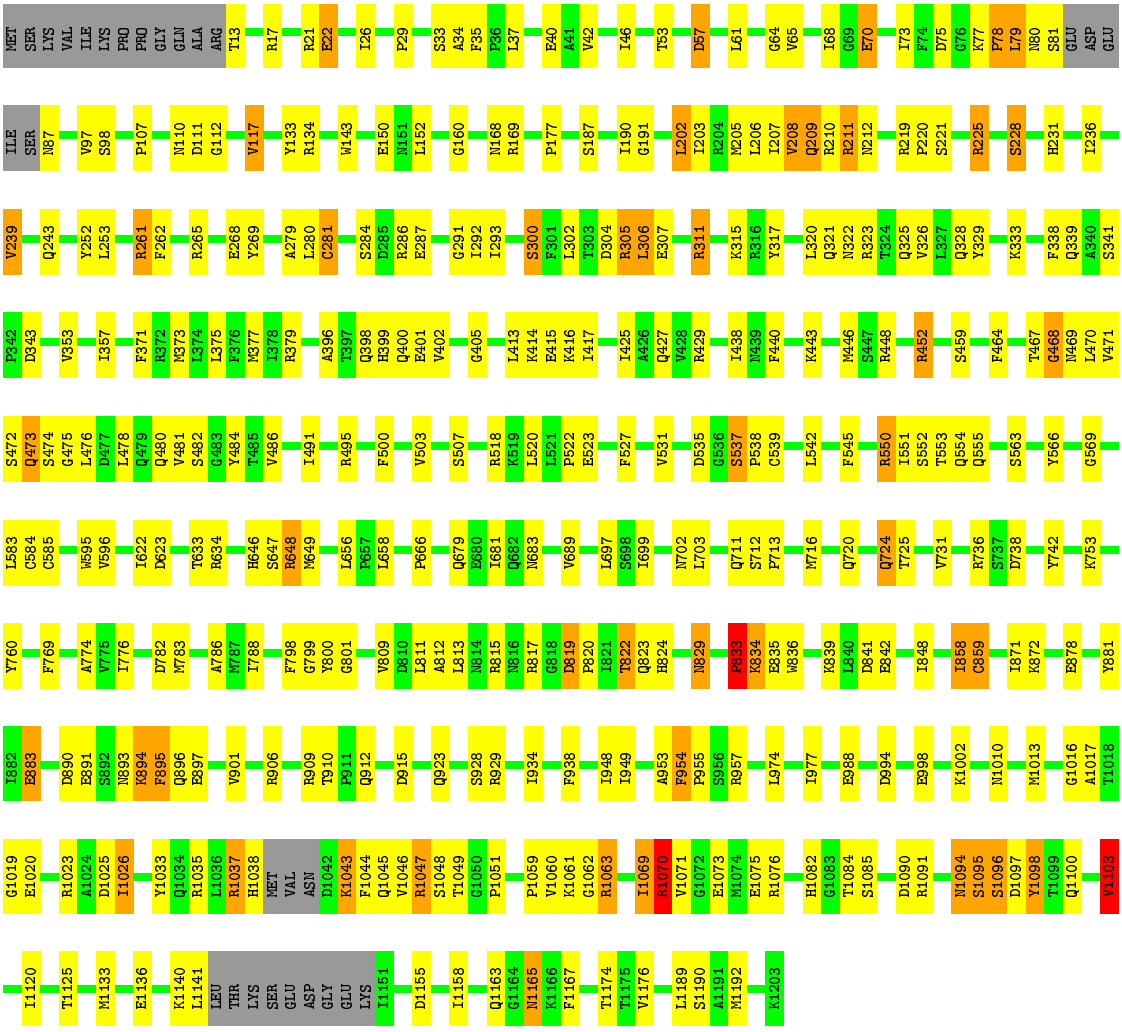
Chain A: 63% 23% • 11%





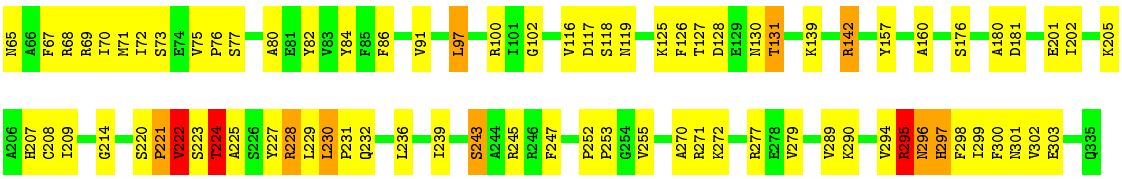
• Molecule 2: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA135

Chain B:



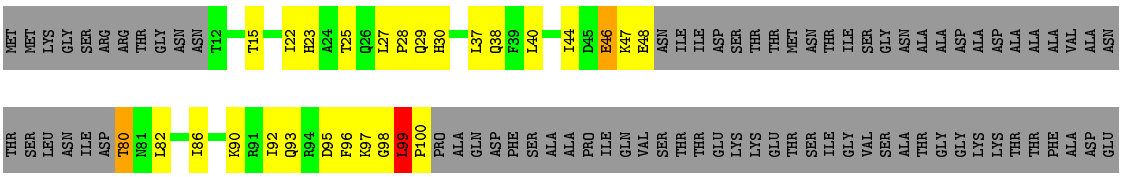
• Molecule 3: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC1

Chain C:



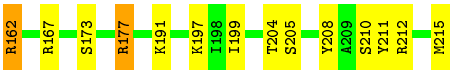
• Molecule 4: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA14

Chain D:



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

Chain E:

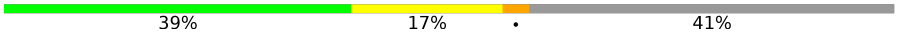


• Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

Chain F:

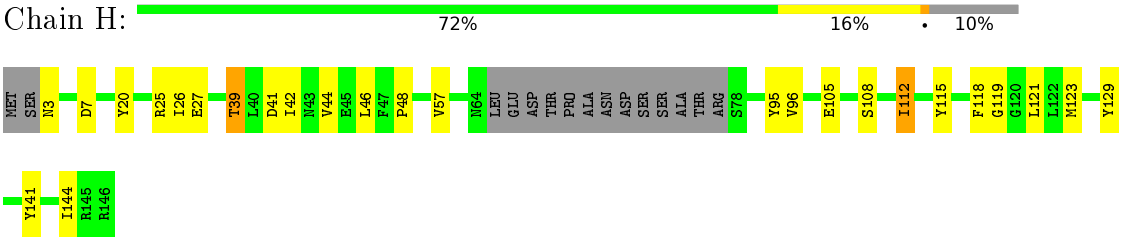
• Molecule 7: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43

Chain G:



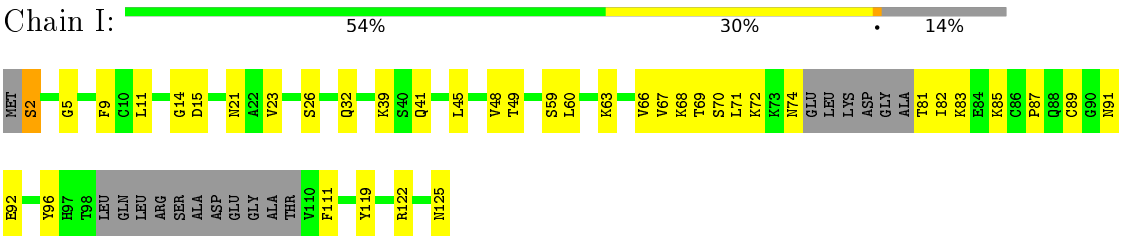
• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H:



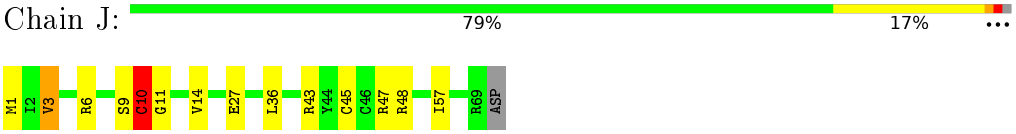
• Molecule 9: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA12

Chain I:



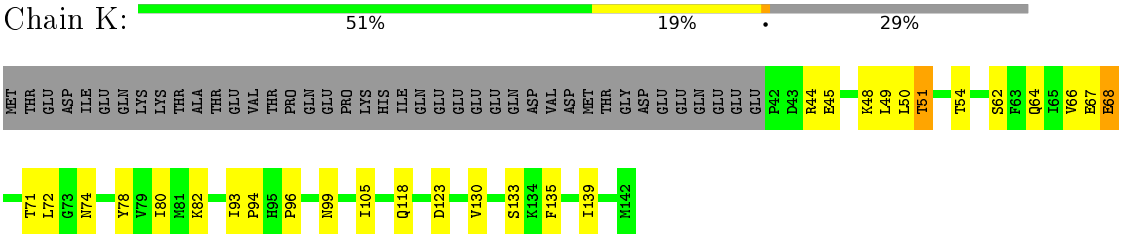
• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

Chain J:



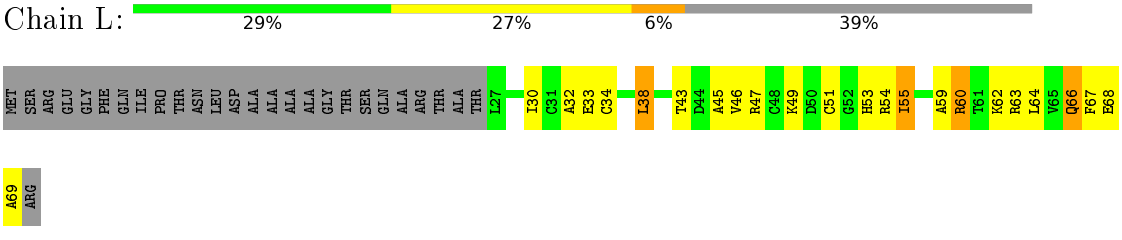
• Molecule 11: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC2

Chain K:



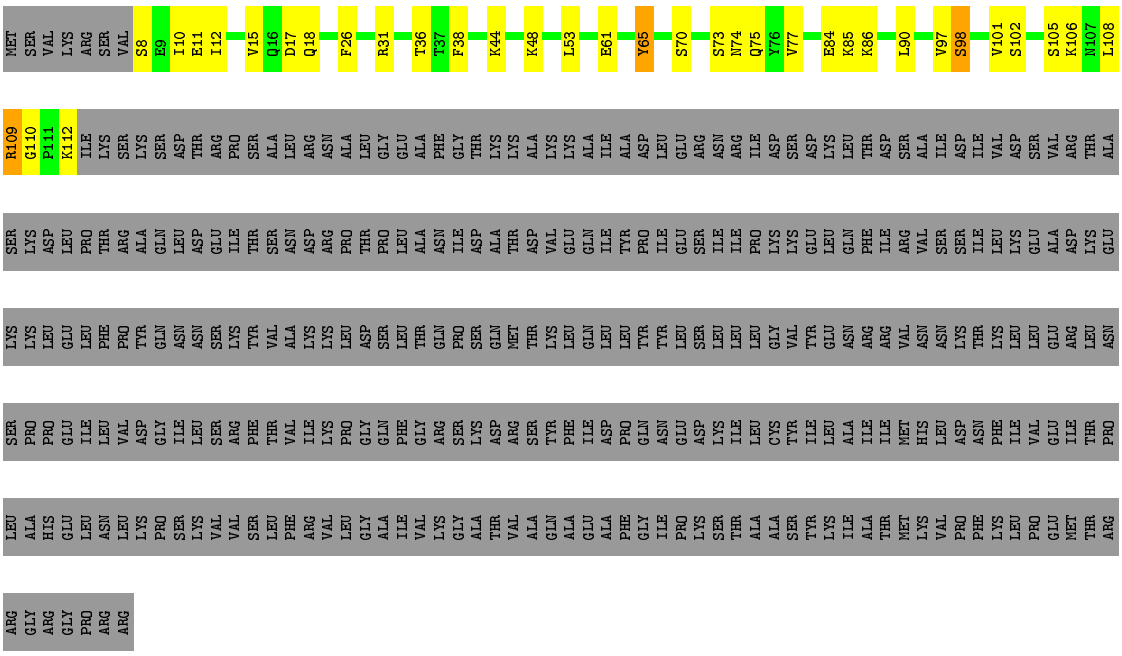
• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain L:



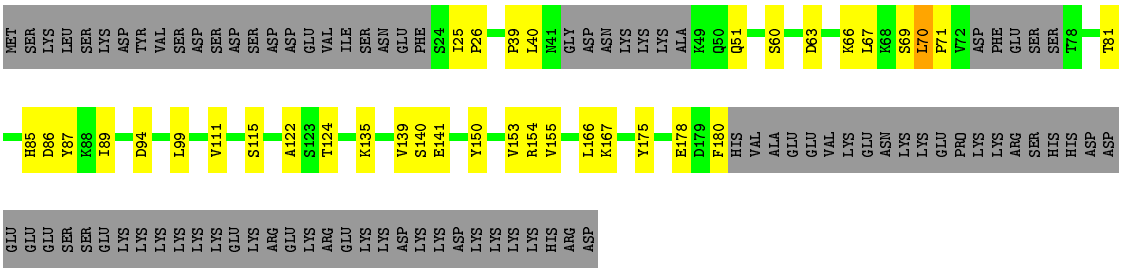
• Molecule 13: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA49

Chain M:



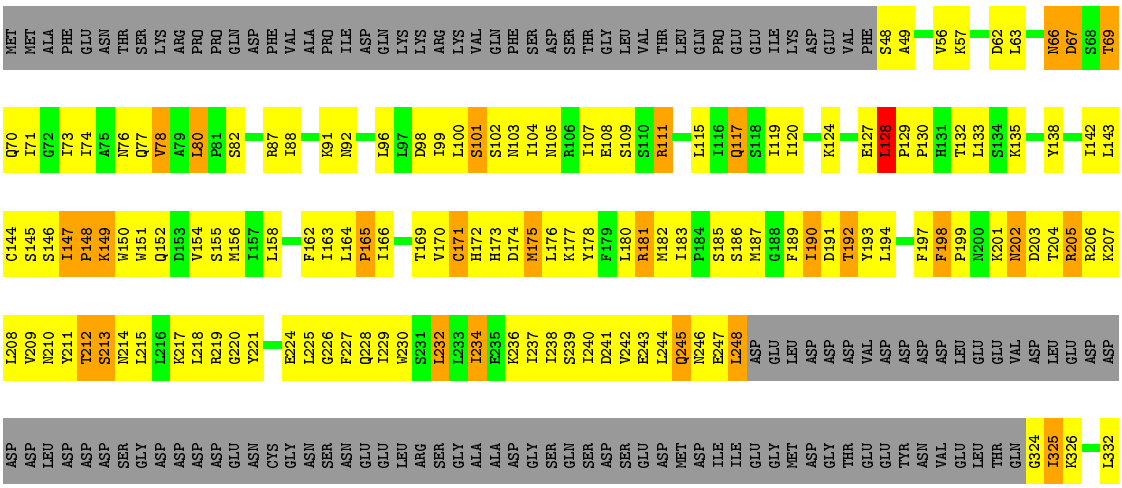
• Molecule 14: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA34

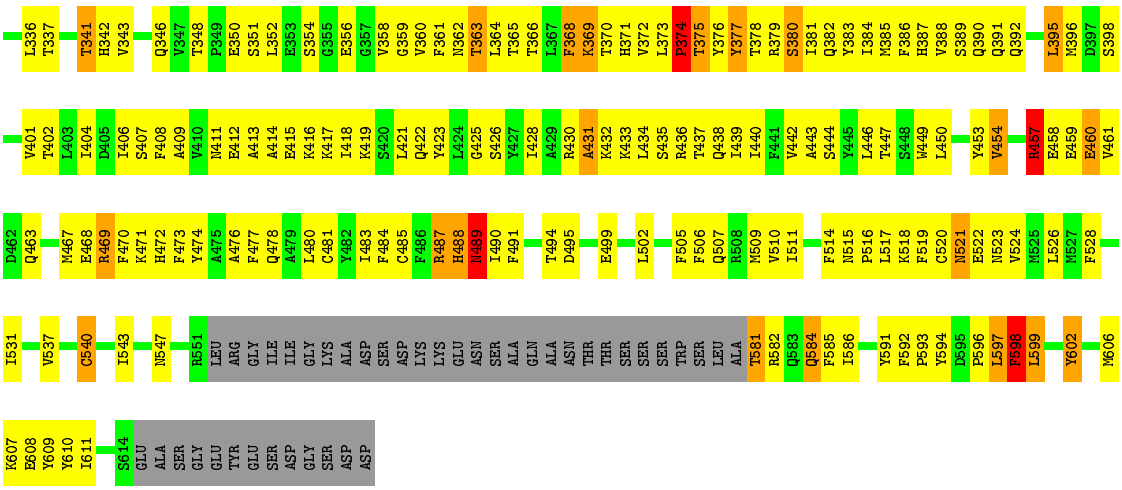
Chain N:



• Molecule 15: RNA POLYMERASE I-SPECIFIC TRANSCRIPTION INITIATION FACTOR RRN3

Chain O:




L336 T337 T341 H342 V343 Q346 V347 T348 P349 E350 S351 L352 E353 S354 G355 E356 G357 V358 G359 V360 F361 N362 T363 L364 T365 T366 L367 F368 K369 T370 H371 V372 L373 P374 T375 Y376 Y377 T378 R379 S380 I381 Q382 Y383 I384 M385 F386 H387 V388 S389 Q390 Q391 Q392 L395 H396 D397 S398
V401 T402 L403 I404 D405 I406 S407 F408 A409 V410 N411 E412 A413 A414 E415 K416 K417 I418 K419 S420 L421 Q422 Y423 L424 G425 S426 Y427 I428 A429 R430 A431 K432 K433 L434 S435 R436 T437 Q438 I439 I440 F441 V442 A443 S444 Y445 L446 T447 S448 W449 L450 Y453 V454 R457 E458 E459 E460 V461
D462 Q463 M467 E468 R469 F470 K471 H472 F473 Y474 A475 A476 F477 Q478 A479 L480 C481 Y482 I483 F484 C485 F486 R487 H488 N489 I490 F491 T494 D495 E499 L502 F505 F506 Q507 R508 M509 V510 I511 F514 N515 P516 L517 K518 F519 C520 N521 E522 N523 V524 M525 L526 M527 F528
I531 V537 C540 I543 N547 R551 LEU ARG GLY ILE ILE GLY LYS ALA ASP SER ASP LYS LYS GLU ASN SER ALA GLN ALA ASN THR THR SER SER SER TRP SER LEU ALA T581 R582 Q583 Q584 F585 I586 Y591 F592 P593 Y594 D595 P596 L597 F598 L599 Y602 M606
K607 E608 Y609 Y610 I611 S614 GLU ALA SER GLY GLU TYR GLU SER ASP GLY SER ASP ASP

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	RELION	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	3600	Depositor
Magnification	37000	Depositor
Image detector	GATAN K2 SUMMIT (4K X 4K)	Depositor

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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.48	0/11900	0.72	22/16073 (0.1%)
10	J	0.57	1/578 (0.2%)	0.59	0/775
11	K	0.46	0/804	0.79	3/1083 (0.3%)
12	L	0.38	0/342	0.55	0/454
13	M	0.42	0/849	0.56	1/1140 (0.1%)
14	N	0.40	0/1172	0.54	0/1580
15	O	0.38	1/3897 (0.0%)	0.58	3/5268 (0.1%)
2	B	0.49	1/9533 (0.0%)	0.77	25/12884 (0.2%)
3	C	0.43	0/2475	0.68	5/3354 (0.1%)
4	D	0.40	0/465	0.59	0/630
5	E	0.40	0/1771	0.66	3/2383 (0.1%)
6	F	0.45	0/838	0.58	0/1129
7	G	0.39	0/1558	0.60	3/2120 (0.1%)
8	H	0.42	0/1070	0.61	0/1449
9	I	0.42	0/831	0.57	0/1117
All	All	0.45	3/38083 (0.0%)	0.69	65/51439 (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	1
15	O	0	5
All	All	0	6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	J	10	CYS	CB-SG	7.64	1.95	1.82

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	O	198	PHE	C-N	-7.16	1.20	1.34
2	B	859	CYS	CB-SG	-6.12	1.71	1.82

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1023	ARG	NE-CZ-NH2	-13.90	113.35	120.30
2	B	452	ARG	NE-CZ-NH2	-13.35	113.63	120.30
1	A	397	ARG	NE-CZ-NH1	13.22	126.91	120.30
1	A	329	ARG	NE-CZ-NH2	-13.12	113.74	120.30
2	B	448	ARG	NE-CZ-NH2	-12.97	113.81	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1343	ASP	Peptide
15	O	374	PRO	Peptide
15	O	375	THR	Peptide
15	O	411	ASN	Peptide
15	O	598	PHE	Peptide

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	11686	0	11770	682	0
2	B	9327	0	9214	482	0
3	C	2423	0	2409	283	0
4	D	459	0	461	103	0
5	E	1735	0	1764	40	0
6	F	823	0	840	64	0
7	G	1520	0	1529	162	0
8	H	1052	0	1021	15	0
9	I	820	0	805	71	0
10	J	569	0	585	6	0
11	K	793	0	790	36	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	L	340	0	361	47	0
13	M	833	0	826	32	0
14	N	1151	0	1169	44	0
15	O	3811	0	3800	754	0
16	A	2	0	0	0	0
16	B	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
All	All	37349	0	37344	2255	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 30.

The worst 5 of 2255 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:75:VAL:HG11	3:C:221:PRO:CG	1.33	1.52
1:A:478:TYR:HA	2:B:1048:SER:CA	1.42	1.50
1:A:436:ALA:CB	1:A:443:ALA:HB2	1.43	1.46
1:A:83:VAL:HG21	1:A:427:PHE:CZ	1.50	1.46
15:O:458:GLU:HA	15:O:461:VAL:CG2	1.26	1.44

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 PDB entries followed by that with respect to all EM entries.

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	1464/1664 (88%)	1370 (94%)	82 (6%)	12 (1%)	24 69
2	B	1166/1203 (97%)	1086 (93%)	56 (5%)	24 (2%)	9 51

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	C	303/335 (90%)	278 (92%)	18 (6%)	7 (2%)	8 49
4	D	54/137 (39%)	50 (93%)	2 (4%)	2 (4%)	4 39
5	E	210/215 (98%)	197 (94%)	11 (5%)	2 (1%)	19 65
6	F	98/155 (63%)	94 (96%)	4 (4%)	0	100 100
7	G	189/326 (58%)	171 (90%)	13 (7%)	5 (3%)	7 46
8	H	127/146 (87%)	121 (95%)	6 (5%)	0	100 100
9	I	101/125 (81%)	89 (88%)	9 (9%)	3 (3%)	5 44
10	J	67/70 (96%)	63 (94%)	4 (6%)	0	100 100
11	K	99/142 (70%)	92 (93%)	7 (7%)	0	100 100
12	L	41/70 (59%)	32 (78%)	6 (15%)	3 (7%)	1 21
13	M	103/415 (25%)	93 (90%)	8 (8%)	2 (2%)	10 53
14	N	139/233 (60%)	123 (88%)	13 (9%)	3 (2%)	8 50
15	O	457/627 (73%)	400 (88%)	38 (8%)	19 (4%)	3 35
All	All	4618/5863 (79%)	4259 (92%)	277 (6%)	82 (2%)	15 54

5 of 82 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1606	SER
2	B	111	ASP
2	B	117	VAL
2	B	895	PHE
2	B	1069	ILE

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 PDB entries followed by that with respect to all EM entries.

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	1307/1465 (89%)	1221 (93%)	86 (7%)	21 59
2	B	1027/1053 (98%)	956 (93%)	71 (7%)	19 58
3	C	269/296 (91%)	249 (93%)	20 (7%)	17 55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	D	55/116 (47%)	49 (89%)	6 (11%)	8 37
5	E	194/197 (98%)	180 (93%)	14 (7%)	18 57
6	F	90/137 (66%)	86 (96%)	4 (4%)	35 70
7	G	170/291 (58%)	158 (93%)	12 (7%)	18 57
8	H	115/128 (90%)	111 (96%)	4 (4%)	43 76
9	I	97/110 (88%)	91 (94%)	6 (6%)	23 61
10	J	64/65 (98%)	57 (89%)	7 (11%)	8 37
11	K	91/130 (70%)	84 (92%)	7 (8%)	16 54
12	L	38/57 (67%)	34 (90%)	4 (10%)	8 39
13	M	95/371 (26%)	85 (90%)	10 (10%)	8 39
14	N	135/220 (61%)	129 (96%)	6 (4%)	35 70
15	O	427/576 (74%)	378 (88%)	49 (12%)	7 35
All	All	4174/5212 (80%)	3868 (93%)	306 (7%)	22 56

5 of 306 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	977	ILE
3	C	232	GLN
15	O	234	ILE
2	B	1037	ARG
2	B	1165	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	154	ASN
15	O	66	ASN
15	O	521	ASN
9	I	32	GLN
12	L	66	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 carbohydrates 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

There are no chain breaks in this entry.