wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 03:26 PM GMT

PDB ID	:	4C2M
Title	:	Structure of RNA polymerase I at 2.8 A resolution
Authors	:	Engel, C.; Sainsbury, S.; Cheung, A.C.; Kostrewa, D.; Cramer, P.
Deposited on	:	2013-08-19
Resolution	:	2.80 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

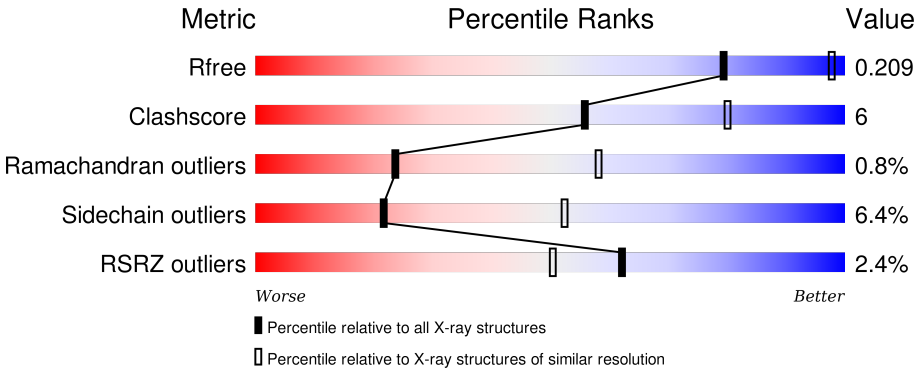
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk26765
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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. 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	1	70	7% 41% 19% • 37%
1	L	70	3% 41% 19% • 37%
2	2	415	3% 19% 5% 75%
2	M	415	% 20% 6% 74%
3	3	233	5% 50% 12% 38%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	N	233	2% 49% 13% 38%
4	4	326	% 14% 83%
4	G	326	9% 47% 12% 41%
4	O	326	2% 13% 84%
4	V	326	2% 49% 11% 40%
5	A	1664	% 74% 16% 9%
5	P	1664	% 73% 16% 9%
6	B	1203	2% 78% 17%
6	Q	1203	2% 76% 18%
7	C	335	% 73% 16% 9%
7	R	335	2% 72% 17% 9%
8	D	137	31% 10% 58%
8	S	137	% 31% 10% 57%
9	E	215	5% 80% 16%
9	T	215	3% 78% 18%
10	F	155	59% 5% 35%
10	U	155	59% 5% 35%
11	H	146	75% 14% 10%
11	W	146	% 78% 11% 10%
12	I	125	3% 73% 26%
12	X	125	3% 70% 24% 5%
13	J	70	74% 20%
13	Y	70	70% 24%
14	K	142	% 54% 15% 29%
14	Z	142	3% 49% 20% 30%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
16	SO4	B	2204	-	-	-	X

2 Entry composition (i)

There are 17 unique types of molecules in this entry. The entry contains 69107 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 POLYMERASES I, II, AND III SUBUNIT RPABC 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1	44	Total	C	N	O	S	0	0	0
			352	217	70	61	4			
1	L	44	Total	C	N	O	S	0	0	0
			352	217	70	61	4			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA49.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	2	103	Total	C	N	O	0	0	0
			814	517	134	163			
2	M	108	Total	C	N	O	0	0	0
			856	543	142	171			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA34.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	3	145	Total	C	N	O	S	0	0	0
			1152	735	189	224	4			
3	N	145	Total	C	N	O	S	0	0	0
			1151	735	188	224	4			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	4	54	Total	C	N	O		0	0	0
			430	262	69	99				
4	G	193	Total	C	N	O	S	0	0	0
			1526	985	262	274	5			
4	O	52	Total	C	N	O		0	0	0
			413	253	64	96				
4	V	197	Total	C	N	O	S	0	0	0
			1557	1001	266	285	5			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA190.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	A	1521	Total	C	N	O	S	0	0	0
			12019	7579	2088	2290	62			
5	P	1518	Total	C	N	O	S	0	0	0
			12000	7567	2085	2286	62			

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA135.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	B	1182	Total	C	N	O	S	0	0	0
			9386	5934	1648	1753	51			
6	Q	1164	Total	C	N	O	S	0	0	0
			9261	5862	1623	1725	51			

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC1.

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

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA14.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	D	58	Total	C	N	O	0	0	0
			459	289	78	92			
8	S	59	Total	C	N	O	0	0	0
			467	293	79	95			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

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

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

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

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3.

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

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	I	124	Total	C	N	O	S	0	0	0
			943	584	160	190	9			
12	X	119	Total	C	N	O	S	0	0	0
			900	557	152	182	9			

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	J	69	Total	C	N	O	S	0	0	0
			569	362	101	100	6			
13	Y	69	Total	C	N	O	S	0	0	0
			569	362	101	100	6			

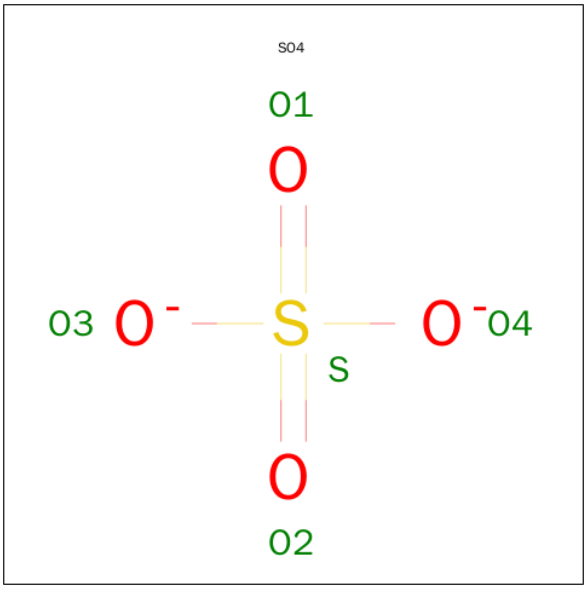
- Molecule 14 is a protein called DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	K	101	Total	C	N	O	S	0	0	0
			793	496	130	162	5			
14	Z	100	Total	C	N	O	S	0	0	0
			786	491	129	161	5			

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

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

- Molecule 16 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	B	1	Total O S 5 4 1	0	0
16	Q	1	Total O S 5 4 1	0	0

- Molecule 17 is water.

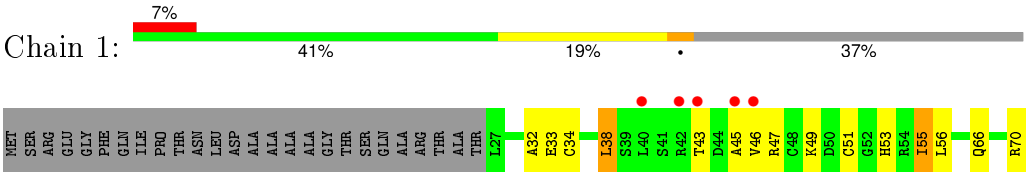
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	3	1	Total O 1 1	0	0
17	A	96	Total O 96 96	0	0
17	B	62	Total O 62 62	0	0
17	C	1	Total O 1 1	0	0
17	D	3	Total O 3 3	0	0
17	E	4	Total O 4 4	0	0
17	F	4	Total O 4 4	0	0
17	G	1	Total O 1 1	0	0
17	H	10	Total O 10 10	0	0
17	I	3	Total O 3 3	0	0
17	N	3	Total O 3 3	0	0
17	O	1	Total O 1 1	0	0
17	P	43	Total O 43 43	0	0
17	Q	15	Total O 15 15	0	0
17	R	1	Total O 1 1	0	0
17	S	1	Total O 1 1	0	0
17	T	4	Total O 4 4	0	0
17	U	4	Total O 4 4	0	0
17	V	2	Total O 2 2	0	0
17	W	1	Total O 1 1	0	0
17	X	2	Total O 2 2	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 errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

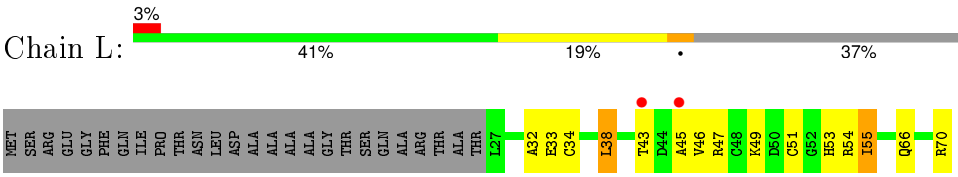
• Molecule 1: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain 1:



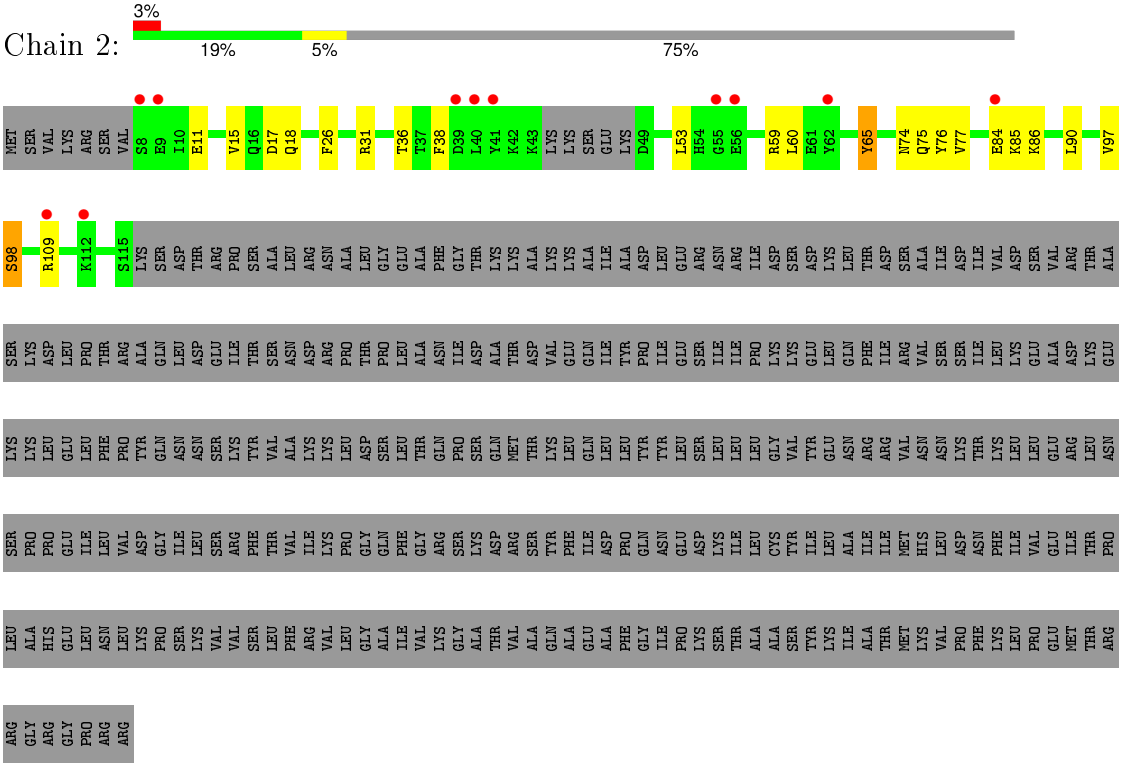
• Molecule 1: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain L:



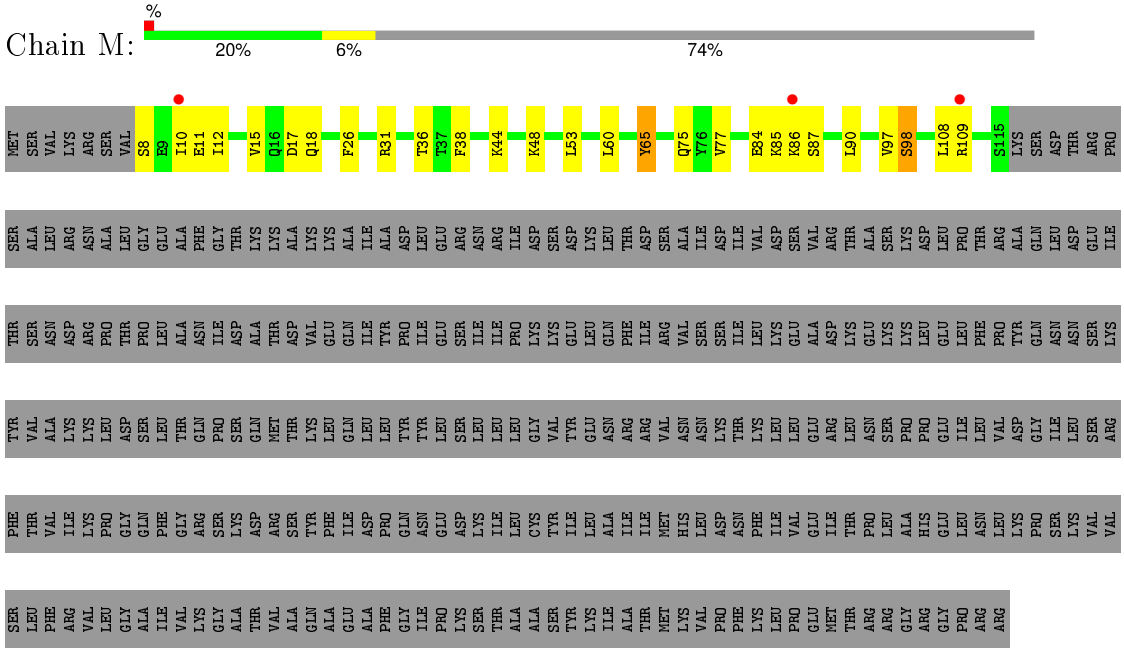
• Molecule 2: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA49

Chain 2:



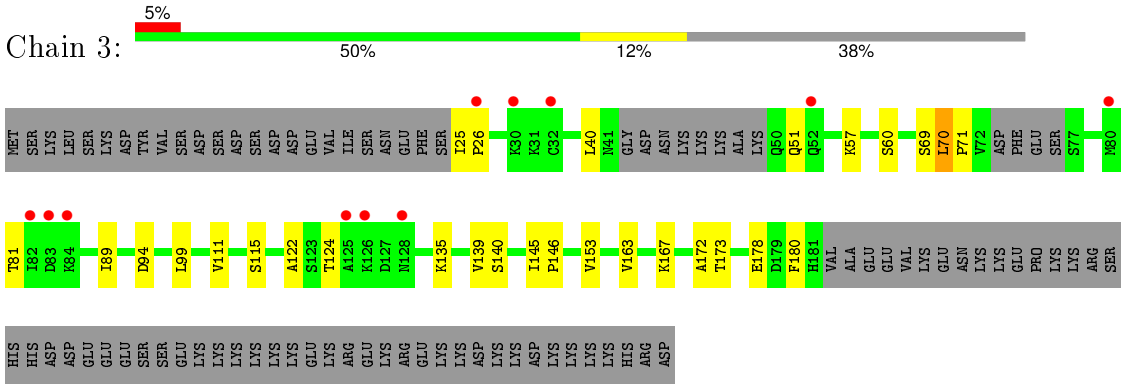
• Molecule 2: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA49

Chain M:



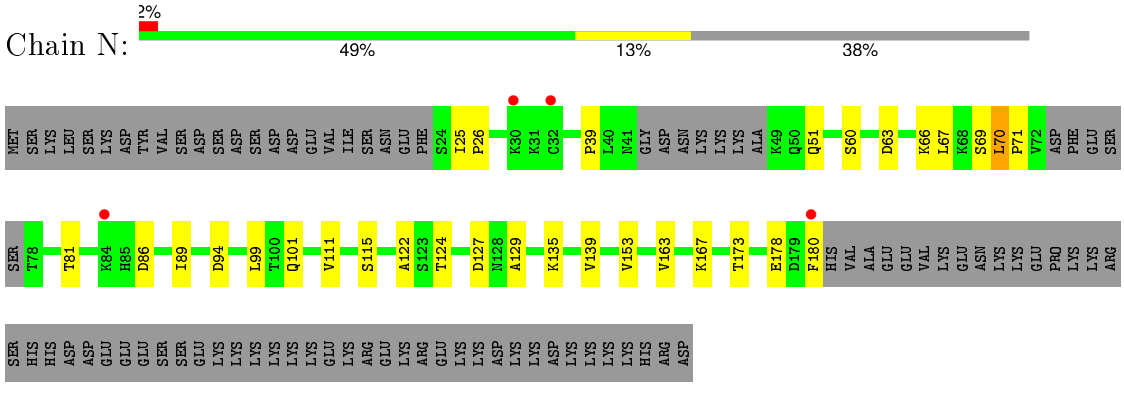
• Molecule 3: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA34

Chain 3:



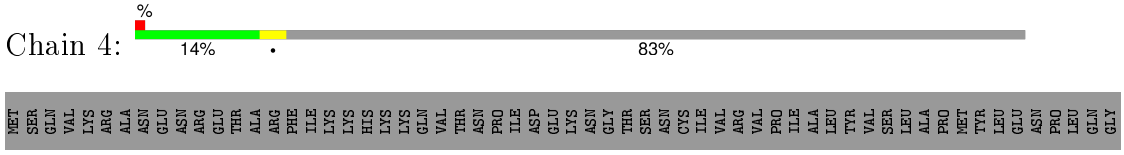
• Molecule 3: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA34

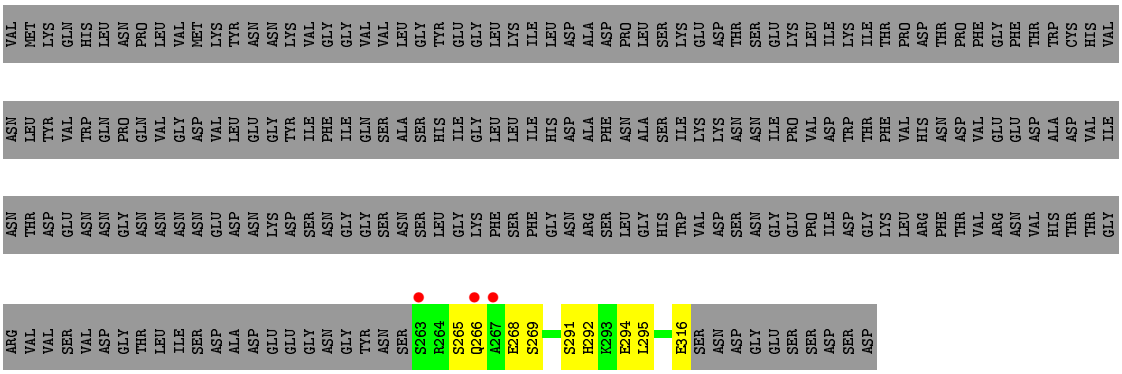
Chain N:



• Molecule 4: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43

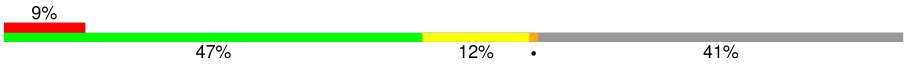
Chain 4:

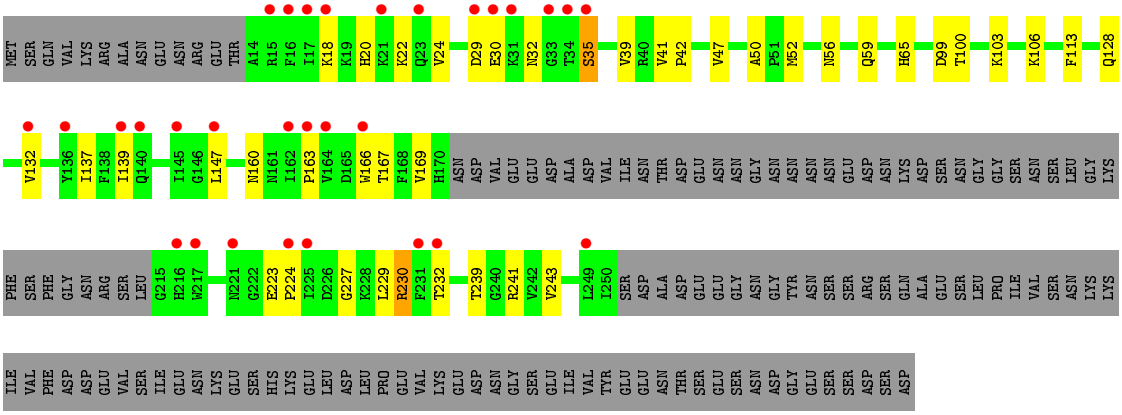




• Molecule 4: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43

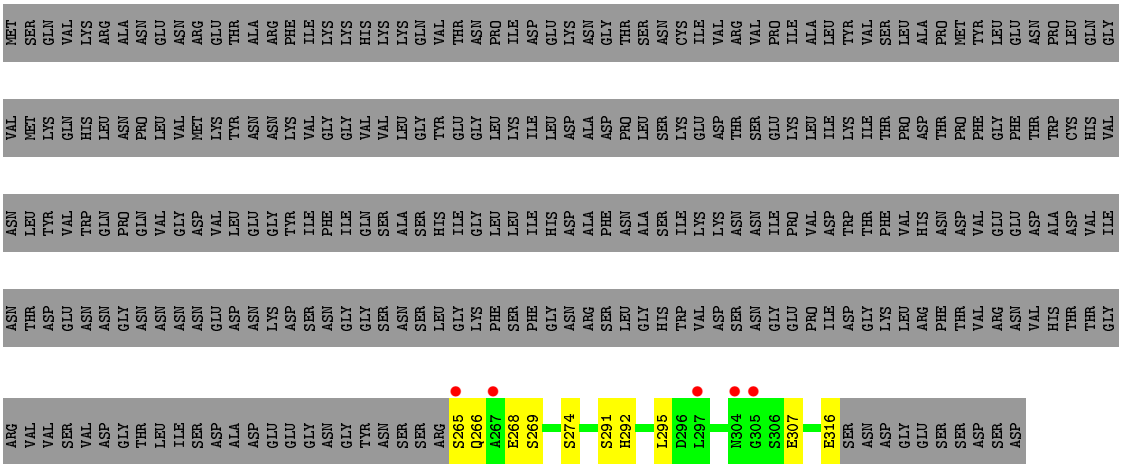
Chain G:





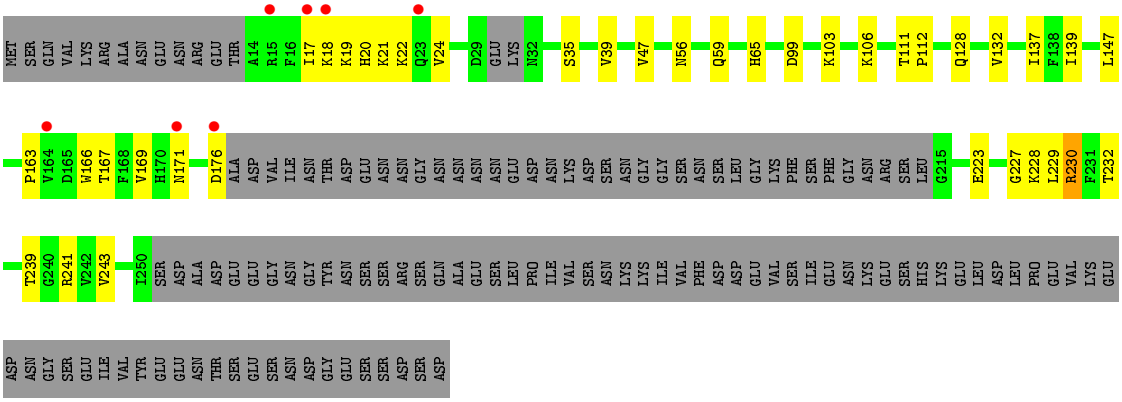
• Molecule 4: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43

Chain O:



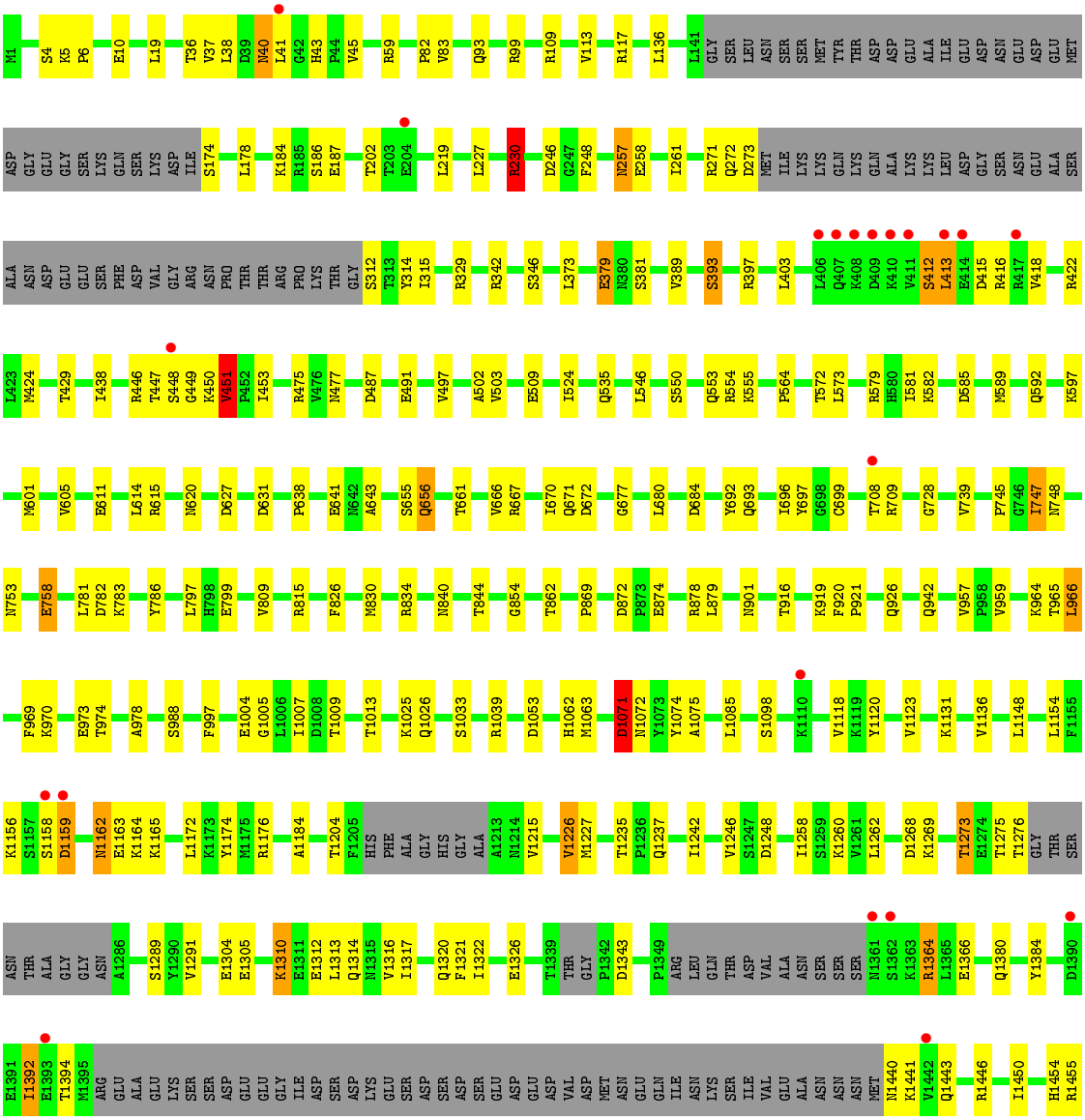
• Molecule 4: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA43

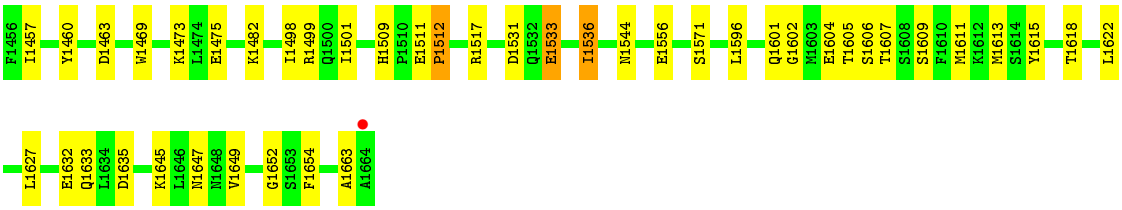
Chain V:



• Molecule 5: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA190

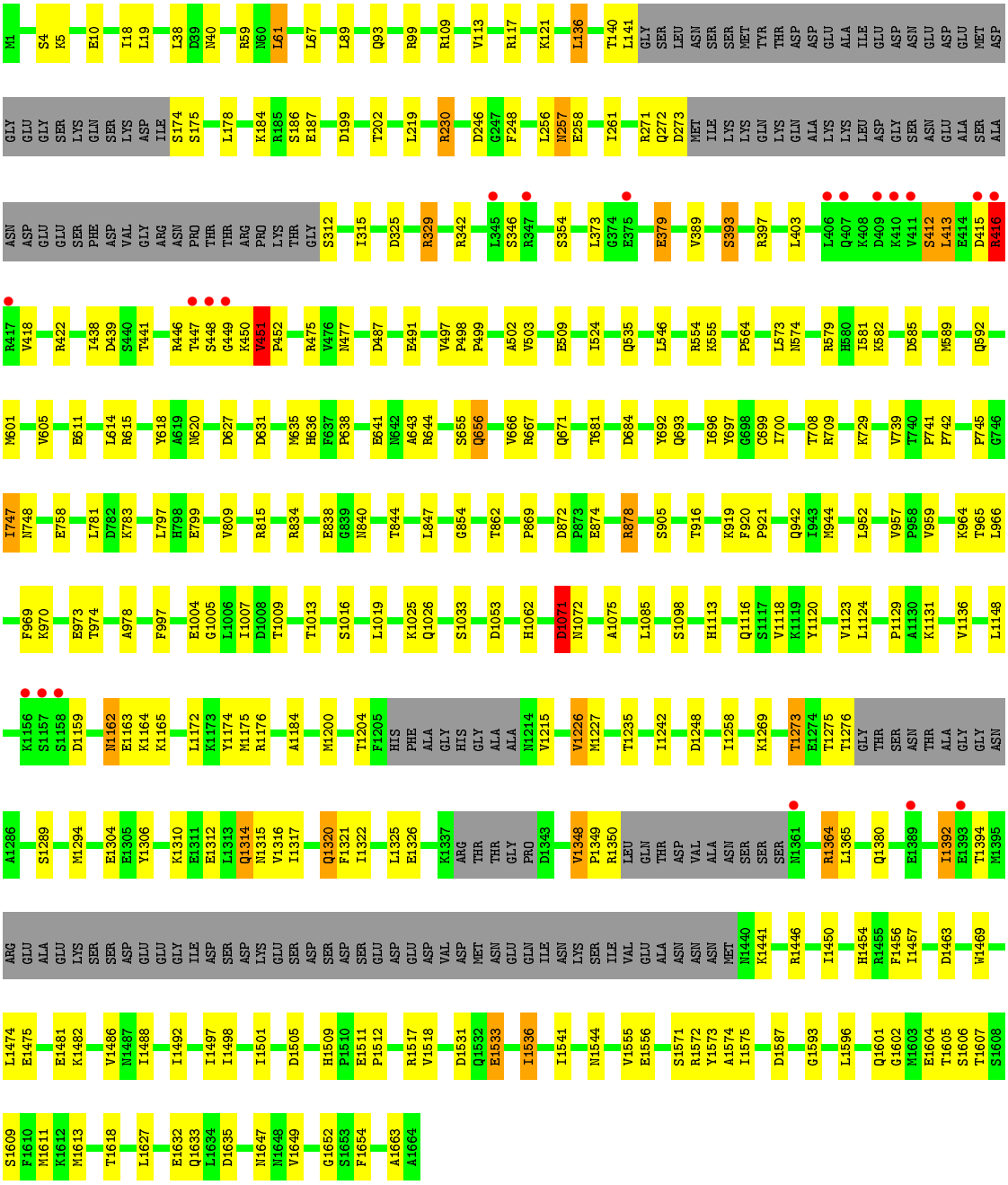
Chain A:





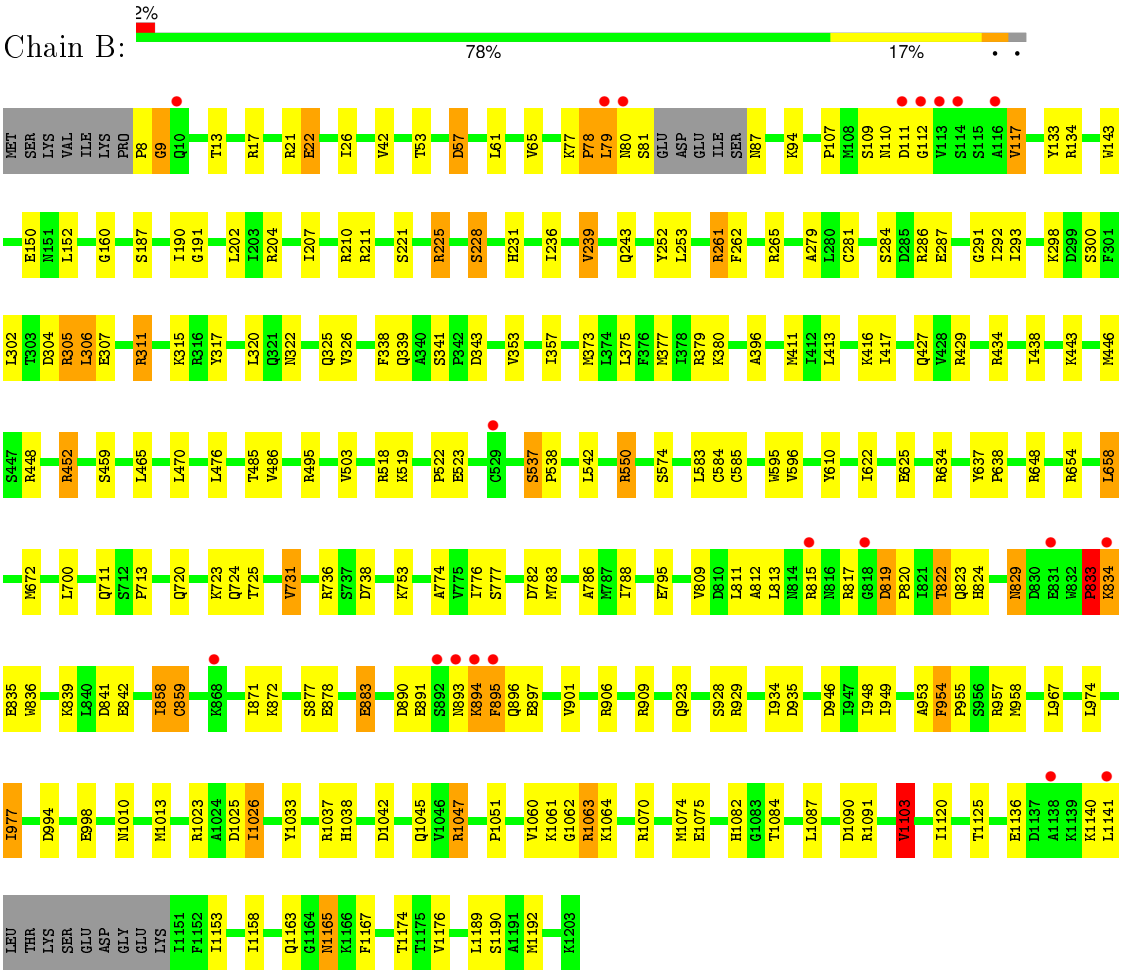
• Molecule 5: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA190

Chain P:



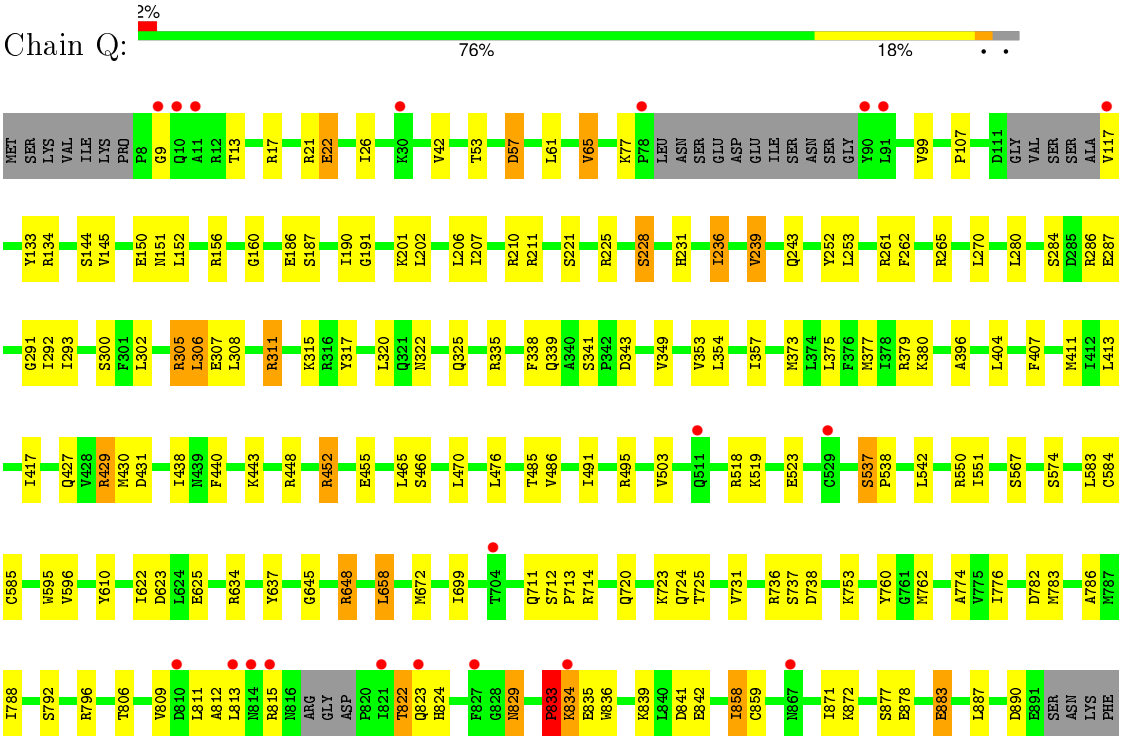
• Molecule 6: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA135

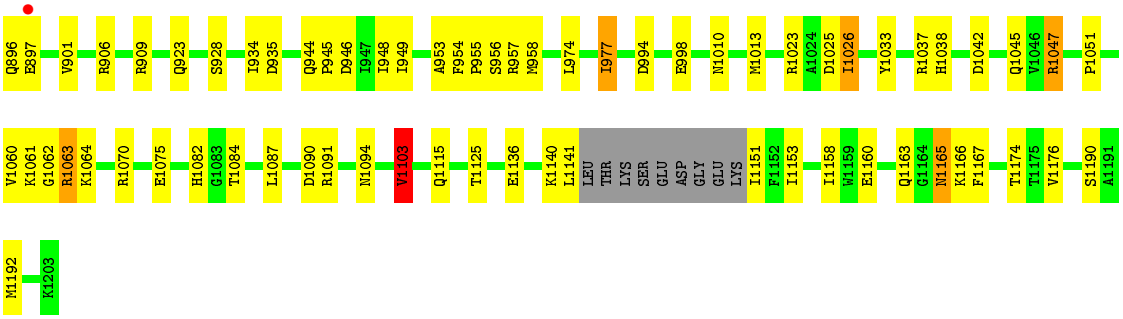
Chain B:



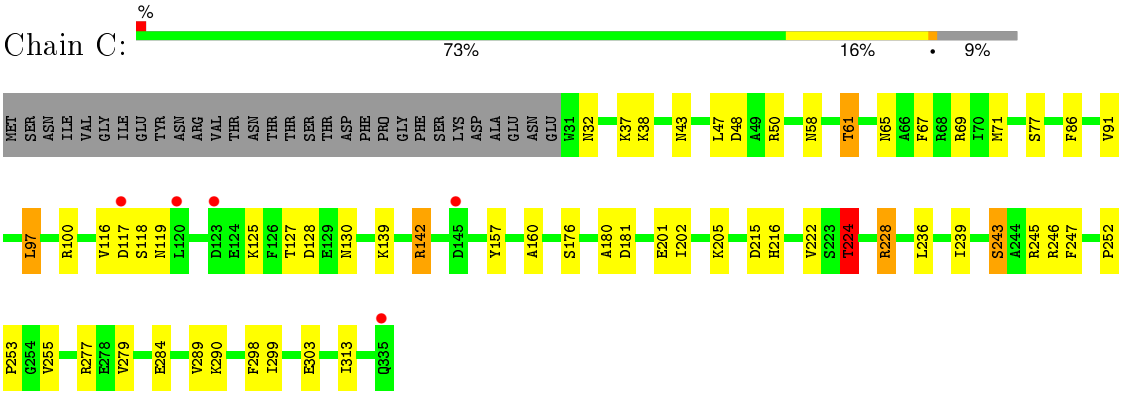
• Molecule 6: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA135

Chain Q:

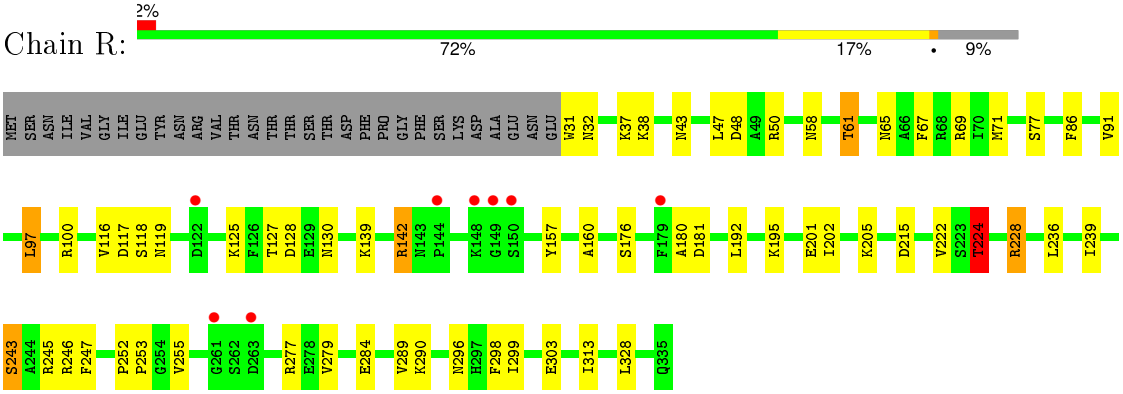




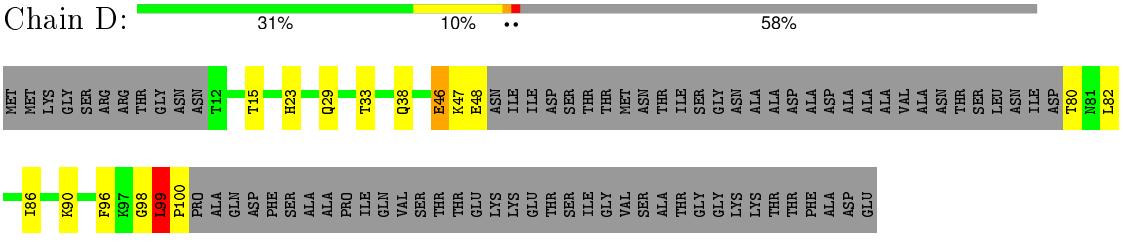
• Molecule 7: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC1



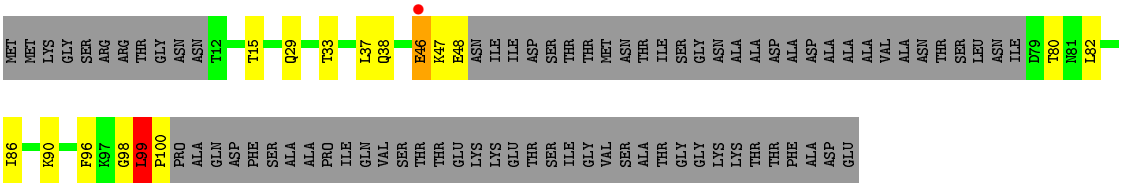
• Molecule 7: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC1



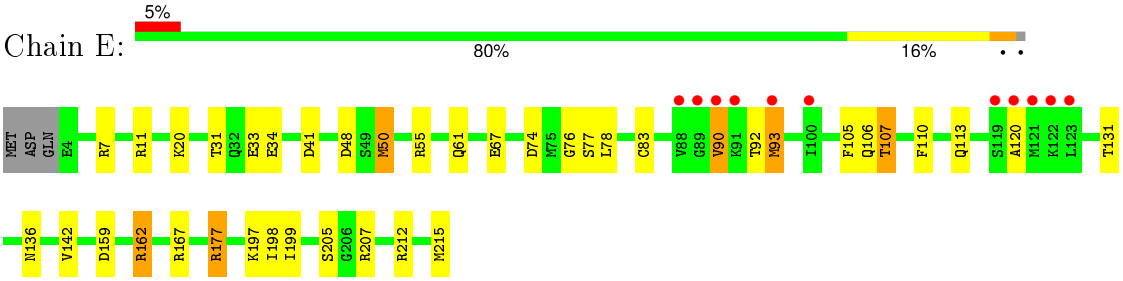
• Molecule 8: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA14



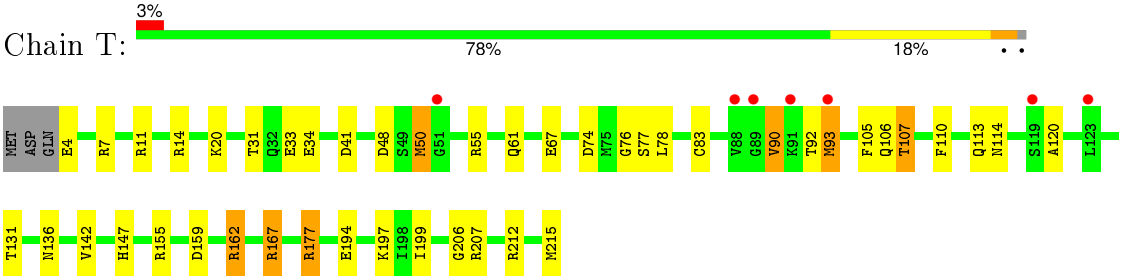
• Molecule 8: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA14



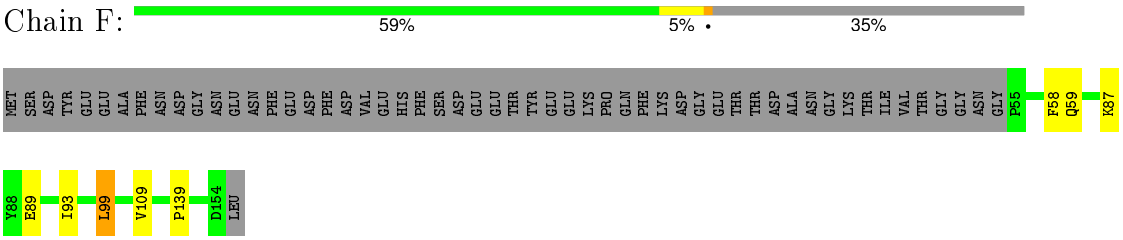
• Molecule 9: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1



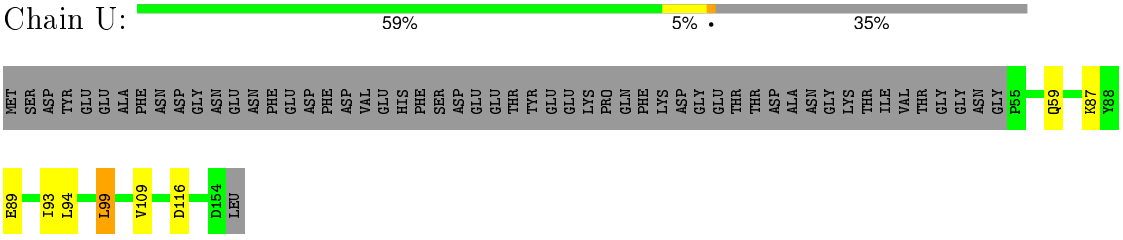
• Molecule 9: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1



• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

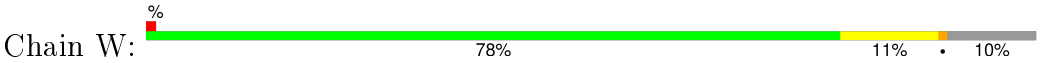


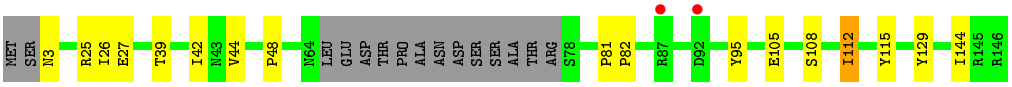
• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2



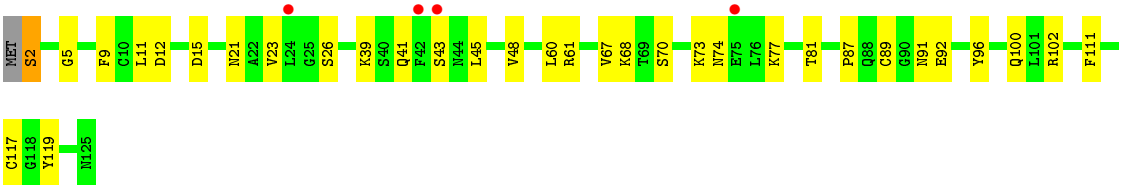
• Molecule 11: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

• Molecule 11: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

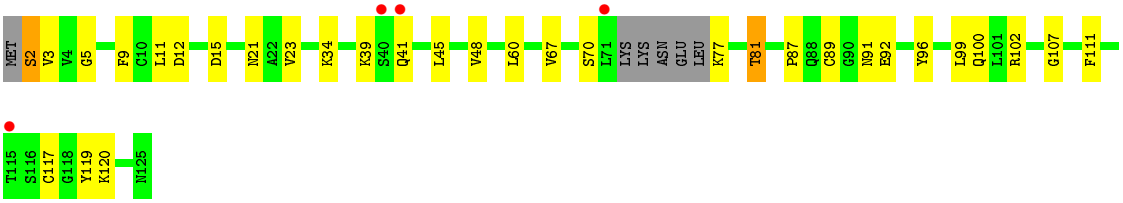




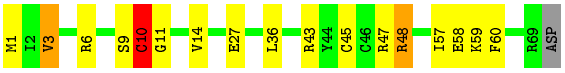
• Molecule 12: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA12



• Molecule 12: DNA-DIRECTED RNA POLYMERASE I SUBUNIT RPA12

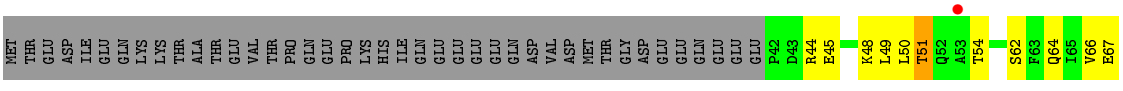


• Molecule 13: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5



• Molecule 13: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

• Molecule 14: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC2



• Molecule 14: DNA-DIRECTED RNA POLYMERASES I AND III SUBUNIT RPAC2

Chain Z:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	122.74Å 139.02Å 209.55Å 108.06° 95.40° 93.85°	Depositor
Resolution (Å)	39.29 – 2.80 39.98 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.9 (39.29-2.80) 89.3 (39.98-2.80)	Depositor EDS
R_{merge}	0.24	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.00 (at 2.81Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.169 , 0.210 0.166 , 0.209	Depositor DCC
R_{free} test set	6480 reflections (2.02%)	DCC
Wilson B-factor (Å^2)	65.7	Xtriage
Anisotropy	0.250	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 54.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Outliers	0 of 320921 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	69107	wwPDB-VP
Average B, all atoms (Å^2)	82.0	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, 0.375 respectively for untwinned datasets, and 0.333, 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, SO4

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	1	0.38	0/354	0.58	0/468
1	L	0.42	0/354	0.60	0/468
2	2	0.34	0/829	0.53	0/1114
2	M	0.40	0/872	0.55	0/1170
3	3	0.35	0/1174	0.52	0/1584
3	N	0.40	0/1172	0.55	0/1580
4	4	0.38	0/434	0.58	0/584
4	G	0.37	0/1564	0.66	3/2127 (0.1%)
4	O	0.42	0/417	0.60	0/562
4	V	0.38	0/1594	0.65	3/2168 (0.1%)
5	A	0.49	0/12236	0.75	25/16523 (0.2%)
5	P	0.45	0/12216	0.73	25/16495 (0.2%)
6	B	0.50	2/9594 (0.0%)	0.78	25/12967 (0.2%)
6	Q	0.43	0/9465	0.76	25/12789 (0.2%)
7	C	0.46	2/2475 (0.1%)	0.67	3/3354 (0.1%)
7	R	0.43	2/2475 (0.1%)	0.66	3/3354 (0.1%)
8	D	0.40	0/465	0.58	0/630
8	S	0.41	0/473	0.59	0/641
9	E	0.40	0/1771	0.66	3/2383 (0.1%)
9	T	0.41	0/1771	0.67	3/2383 (0.1%)
10	F	0.46	0/838	0.59	0/1129
10	U	0.43	0/838	0.59	0/1129
11	H	0.42	0/1070	0.61	0/1449
11	W	0.40	0/1070	0.60	0/1449
12	I	0.43	0/956	0.59	0/1288
12	X	0.42	0/912	0.58	0/1229
13	J	0.57	1/578 (0.2%)	0.59	0/775
13	Y	0.42	0/578	0.58	0/775
14	K	0.46	0/804	0.79	3/1083 (0.3%)
14	Z	0.42	0/796	0.75	3/1072 (0.3%)
All	All	0.45	7/70145 (0.0%)	0.71	124/94722 (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
5	A	0	2
5	P	0	1
All	All	0	3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	J	10	CYS	CB-SG	7.66	1.95	1.82
6	B	281	CYS	CB-SG	-6.93	1.70	1.82
7	C	58	ASN	CG-ND2	-6.88	1.15	1.32
7	R	58	ASN	CG-OD1	-6.67	1.09	1.24
7	C	58	ASN	CG-OD1	-6.47	1.09	1.24

The worst 5 of 124 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	Q	1023	ARG	NE-CZ-NH2	-14.08	113.26	120.30
6	Q	550	ARG	NE-CZ-NH2	-13.79	113.41	120.30
6	B	1023	ARG	NE-CZ-NH2	-13.75	113.43	120.30
5	P	397	ARG	NE-CZ-NH1	13.48	127.04	120.30
6	Q	452	ARG	NE-CZ-NH2	-13.35	113.62	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	A	1343	ASP	Peptide
5	A	781	LEU	Peptide
5	P	781	LEU	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	1	352	0	374	9	0
1	L	352	0	374	8	0
2	2	814	0	804	12	0
2	M	856	0	855	14	0
3	3	1152	0	1163	16	0
3	N	1151	0	1169	17	0
4	4	430	0	407	15	0
4	G	1526	0	1540	21	0
4	O	413	0	389	13	0
4	V	1557	0	1555	16	0
5	A	12019	0	12078	156	1
5	P	12000	0	12058	157	1
6	B	9386	0	9279	148	0
6	Q	9261	0	9161	153	0
7	C	2423	0	2412	33	0
7	R	2423	0	2412	39	0
8	D	459	0	462	7	0
8	S	467	0	466	7	0
9	E	1735	0	1764	19	0
9	T	1735	0	1764	21	0
10	F	823	0	841	5	0
10	U	823	0	841	6	0
11	H	1052	0	1021	10	0
11	W	1052	0	1021	7	0
12	I	943	0	929	20	0
12	X	900	0	879	22	0
13	J	569	0	585	12	0
13	Y	569	0	585	15	0
14	K	793	0	790	13	0
14	Z	786	0	782	18	0
15	1	1	0	0	0	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
15	P	2	0	0	0	0
15	Q	1	0	0	0	0
15	X	2	0	0	0	0
15	Y	1	0	0	0	0
16	B	5	0	0	1	0
16	Q	5	0	0	1	0
17	3	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	A	96	0	0	5	0
17	B	62	0	0	1	0
17	C	1	0	0	0	0
17	D	3	0	0	0	0
17	E	4	0	0	0	0
17	F	4	0	0	0	0
17	G	1	0	0	0	0
17	H	10	0	0	0	0
17	I	3	0	0	1	0
17	N	3	0	0	0	0
17	O	1	0	0	0	0
17	P	43	0	0	0	0
17	Q	15	0	0	0	0
17	R	1	0	0	0	0
17	S	1	0	0	0	0
17	T	4	0	0	0	0
17	U	4	0	0	0	0
17	V	2	0	0	0	0
17	W	1	0	0	0	0
17	X	2	0	0	0	0
All	All	69107	0	68760	848	1

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 6.

The worst 5 of 848 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)
5:A:620:ASN:OD1	5:A:667:ARG:NH2	2.01	0.93
6:B:713:PRO:HG3	12:I:100:GLN:HG3	1.48	0.93
6:Q:26:ILE:HG12	13:Y:58:GLU:HG2	1.51	0.93
5:A:1009:THR:HG21	12:I:102:ARG:H	1.38	0.88
14:Z:66:VAL:HG12	14:Z:67:GLU:HG2	1.55	0.88

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:1310:LYS:NZ	5:P:838:GLU:OE2[1_556]	2.07	0.13

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	1	42/70 (60%)	37 (88%)	3 (7%)	2 (5%)	3 9
1	L	42/70 (60%)	36 (86%)	4 (10%)	2 (5%)	3 9
2	2	99/415 (24%)	93 (94%)	4 (4%)	2 (2%)	9 30
2	M	106/415 (26%)	96 (91%)	8 (8%)	2 (2%)	10 32
3	3	139/233 (60%)	122 (88%)	15 (11%)	2 (1%)	14 42
3	N	139/233 (60%)	123 (88%)	13 (9%)	3 (2%)	8 28
4	4	52/326 (16%)	49 (94%)	3 (6%)	0	100 100
4	G	189/326 (58%)	174 (92%)	13 (7%)	2 (1%)	17 50
4	O	50/326 (15%)	46 (92%)	4 (8%)	0	100 100
4	V	191/326 (59%)	175 (92%)	15 (8%)	1 (0%)	34 69
5	A	1505/1664 (90%)	1434 (95%)	62 (4%)	9 (1%)	30 65
5	P	1502/1664 (90%)	1433 (95%)	59 (4%)	10 (1%)	26 62
6	B	1176/1203 (98%)	1116 (95%)	47 (4%)	13 (1%)	17 50
6	Q	1152/1203 (96%)	1104 (96%)	42 (4%)	6 (0%)	34 69
7	C	303/335 (90%)	288 (95%)	13 (4%)	2 (1%)	26 62
7	R	303/335 (90%)	289 (95%)	12 (4%)	2 (1%)	26 62
8	D	54/137 (39%)	50 (93%)	2 (4%)	2 (4%)	4 14
8	S	55/137 (40%)	51 (93%)	2 (4%)	2 (4%)	4 14
9	E	210/215 (98%)	198 (94%)	11 (5%)	1 (0%)	34 69
9	T	210/215 (98%)	199 (95%)	9 (4%)	2 (1%)	19 52
10	F	98/155 (63%)	96 (98%)	2 (2%)	0	100 100
10	U	98/155 (63%)	98 (100%)	0	0	100 100
11	H	127/146 (87%)	121 (95%)	6 (5%)	0	100 100
11	W	127/146 (87%)	119 (94%)	8 (6%)	0	100 100
12	I	122/125 (98%)	107 (88%)	12 (10%)	3 (2%)	7 24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
12	X	115/125 (92%)	102 (89%)	10 (9%)	3 (3%)	7 22
13	J	67/70 (96%)	63 (94%)	4 (6%)	0	100 100
13	Y	67/70 (96%)	63 (94%)	4 (6%)	0	100 100
14	K	99/142 (70%)	92 (93%)	7 (7%)	0	100 100
14	Z	98/142 (69%)	92 (94%)	6 (6%)	0	100 100
All	All	8537/11124 (77%)	8066 (94%)	400 (5%)	71 (1%)	24 58

5 of 71 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	A	1606	SER
6	B	111	ASP
6	B	895	PHE
7	C	224	THR
8	D	99	LEU

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	1	39/57 (68%)	36 (92%)	3 (8%)	16 41
1	L	39/57 (68%)	36 (92%)	3 (8%)	16 41
2	2	93/371 (25%)	85 (91%)	8 (9%)	13 36
2	M	98/371 (26%)	88 (90%)	10 (10%)	9 26
3	3	135/220 (61%)	128 (95%)	7 (5%)	29 62
3	N	135/220 (61%)	129 (96%)	6 (4%)	35 69
4	4	52/291 (18%)	51 (98%)	1 (2%)	65 91
4	G	171/291 (59%)	159 (93%)	12 (7%)	19 47
4	O	50/291 (17%)	49 (98%)	1 (2%)	63 90
4	V	175/291 (60%)	161 (92%)	14 (8%)	15 40
5	A	1345/1465 (92%)	1263 (94%)	82 (6%)	23 55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	P	1343/1465 (92%)	1261 (94%)	82 (6%)	23 55
6	B	1033/1053 (98%)	965 (93%)	68 (7%)	21 51
6	Q	1019/1053 (97%)	956 (94%)	63 (6%)	23 54
7	C	269/296 (91%)	253 (94%)	16 (6%)	24 57
7	R	269/296 (91%)	252 (94%)	17 (6%)	22 53
8	D	55/116 (47%)	49 (89%)	6 (11%)	8 23
8	S	56/116 (48%)	50 (89%)	6 (11%)	8 24
9	E	194/197 (98%)	180 (93%)	14 (7%)	18 45
9	T	194/197 (98%)	178 (92%)	16 (8%)	14 38
10	F	90/137 (66%)	86 (96%)	4 (4%)	35 69
10	U	90/137 (66%)	86 (96%)	4 (4%)	35 69
11	H	115/128 (90%)	111 (96%)	4 (4%)	43 77
11	W	115/128 (90%)	111 (96%)	4 (4%)	43 77
12	I	109/110 (99%)	102 (94%)	7 (6%)	22 52
12	X	104/110 (94%)	98 (94%)	6 (6%)	25 57
13	J	64/65 (98%)	57 (89%)	7 (11%)	8 23
13	Y	64/65 (98%)	57 (89%)	7 (11%)	8 23
14	K	91/130 (70%)	84 (92%)	7 (8%)	16 41
14	Z	90/130 (69%)	83 (92%)	7 (8%)	16 41
All	All	7696/9854 (78%)	7204 (94%)	492 (6%)	22 52

5 of 492 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	H	39	THR
5	P	257	ASN
4	V	18	LYS
12	I	74	ASN
2	M	17	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
5	A	43	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	A	1162	ASN
6	B	1163	GLN
5	P	1162	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 16 ligands modelled in this entry, 14 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
16	SO4	B	2204	-	4,4,4	0.30	0	6,6,6	0.43	0
16	SO4	Q	2204	-	4,4,4	0.30	0	6,6,6	0.43	0

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
16	SO4	B	2204	-	-	0/0/0/0	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
16	SO4	Q	2204	-	-	0/0/0/0	0/0/0/0

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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	2204	SO4	1	0
16	Q	2204	SO4	1	0

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^{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	1	44/70 (62%)	0.14	5 (11%) 7 3	79, 117, 154, 193	0
1	L	44/70 (62%)	-0.23	2 (4%) 37 26	65, 88, 134, 175	0
2	2	103/415 (24%)	0.55	11 (10%) 8 4	96, 128, 160, 171	0
2	M	108/415 (26%)	0.10	3 (2%) 56 44	66, 94, 134, 156	0
3	3	145/233 (62%)	0.23	11 (7%) 17 9	76, 121, 161, 187	0
3	N	145/233 (62%)	-0.15	4 (2%) 56 44	56, 93, 133, 160	0
4	4	54/326 (16%)	0.09	3 (5%) 28 18	74, 103, 143, 158	0
4	G	193/326 (59%)	0.36	30 (15%) 3 1	51, 107, 168, 189	0
4	O	52/326 (15%)	0.23	5 (9%) 10 5	64, 100, 145, 155	0
4	V	197/326 (60%)	-0.05	7 (3%) 46 34	57, 89, 140, 164	0
5	A	1521/1664 (91%)	-0.37	22 (1%) 78 69	40, 63, 117, 161	0
5	P	1518/1664 (91%)	-0.33	20 (1%) 79 71	48, 72, 123, 182	0
6	B	1182/1203 (98%)	-0.37	20 (1%) 73 63	37, 59, 108, 187	0
6	Q	1164/1203 (96%)	-0.23	21 (1%) 71 61	40, 83, 132, 165	0
7	C	305/335 (91%)	-0.23	5 (1%) 74 66	60, 81, 127, 148	0
7	R	305/335 (91%)	-0.12	8 (2%) 59 47	68, 95, 137, 151	0
8	D	58/137 (42%)	-0.20	0 100 100	56, 107, 136, 149	0
8	S	59/137 (43%)	-0.23	1 (1%) 73 63	62, 96, 128, 134	0
9	E	212/215 (98%)	-0.04	11 (5%) 31 20	49, 92, 149, 164	0
9	T	212/215 (98%)	-0.04	7 (3%) 50 38	57, 94, 147, 169	0
10	F	100/155 (64%)	-0.59	0 100 100	42, 60, 93, 129	0
10	U	100/155 (64%)	-0.56	0 100 100	51, 70, 107, 119	0
11	H	131/146 (89%)	-0.37	0 100 100	60, 80, 113, 125	0
11	W	131/146 (89%)	-0.25	2 (1%) 76 68	66, 91, 121, 133	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
12	I	124/125 (99%)	-0.19	4 (3%) 51 39	52, 81, 129, 155	0
12	X	119/125 (95%)	0.11	4 (3%) 49 36	65, 103, 152, 177	0
13	J	69/70 (98%)	-0.34	0 100 100	56, 68, 91, 106	0
13	Y	69/70 (98%)	-0.21	0 100 100	72, 85, 114, 128	0
14	K	101/142 (71%)	-0.35	1 (0%) 84 77	58, 72, 121, 148	0
14	Z	100/142 (70%)	-0.29	4 (4%) 42 30	63, 83, 133, 164	0
All	All	8665/11124 (77%)	-0.24	211 (2%) 62 50	37, 77, 136, 193	0

The worst 5 of 211 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	B	893	ASN	6.5
3	3	84	LYS	5.9
5	P	448	SER	5.8
1	1	43	THR	5.6
4	G	23	GLN	5.3

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(Å^2)	Q<0.9
16	SO4	B	2204	5/5	0.93	0.29	3.68	123,124,132,137	0
16	SO4	Q	2204	5/5	0.94	0.31	1.85	133,139,142,147	0
15	ZN	J	3001	1/1	0.99	0.21	0.55	547,547,547,547	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
15	ZN	Y	3001	1/1	0.99	0.19	0.11	489,489,489,489	0
15	ZN	Q	3001	1/1	0.99	0.11	-0.38	64,64,64,64	0
15	ZN	P	3002	1/1	0.99	0.11	-0.40	83,83,83,83	0
15	ZN	A	3002	1/1	1.00	0.11	-0.47	77,77,77,77	0
15	ZN	1	3001	1/1	0.99	0.10	-0.84	120,120,120,120	0
15	ZN	I	3001	1/1	0.99	0.11	-0.97	73,73,73,73	0
15	ZN	B	3001	1/1	1.00	0.10	-0.99	53,53,53,53	0
15	ZN	A	3001	1/1	0.99	0.10	-1.13	75,75,75,75	0
15	ZN	P	3001	1/1	1.00	0.09	-1.40	70,70,70,70	0
15	ZN	X	3001	1/1	0.99	0.08	-1.57	113,113,113,113	0
15	ZN	L	3001	1/1	1.00	0.07	-1.80	87,87,87,87	0
15	ZN	X	3002	1/1	0.97	0.04	-2.17	89,89,89,89	1
15	ZN	I	3002	1/1	0.98	0.06	-2.27	77,77,77,77	1

6.5 Other polymers

There are no such residues in this entry.