wwPDB X-ray Structure Validation Summary Report (i)

Dec 2, 2023 – 01:00 pm GMT

PDB ID	:	2WB1
Title	:	The complete structure of the archaeal 13-subunit DNA-directed RNA Polymerase
Authors	:	Korkhin, Y.; Unligil, U.M.; Littlefield, O.; Nelson, P.J.; Stuart, D.I.; Sigler, P.B.; Bell, S.D.; Abrescia, N.G.A.
Deposited on	:	2009-02-19
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 (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 (i)) were used in the production of this report:

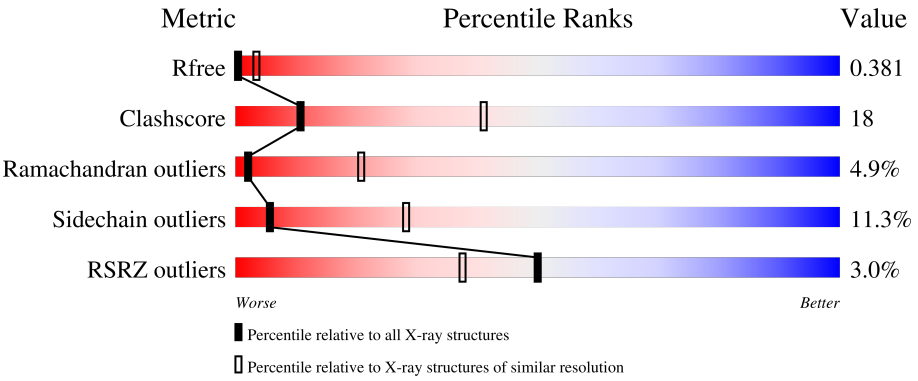
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.52 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 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	880	52% 36% 7%
1	W	880	51% 37% 7%
2	B	1131	52% 37% 7%
2	R	1131	51% 37% 7%
3	C	395	4% 47% 39% 6% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Y	395	2% 45% 40% 7% 7%
4	D	265	6% 60% 35%
4	S	265	5% 61% 33%
5	E	180	12% 63% 30%
5	T	180	7% 65% 28%
6	F	113	10% 63% 15% 19%
6	U	113	9% 58% 19% 19%
7	G	132	3% 52% 30% 14%
7	V	132	3% 53% 29% 14%
8	H	84	51% 30% 7% 12%
8	Z	84	48% 33% 7% 12%
9	I	95	% 52% 31% 14%
9	K	95	% 48% 35% 14%
10	J	104	5% 23% 11% 66%
10	Q	104	2% 19% 12% 68%
11	L	92	2% 72% 23% 5%
11	M	92	% 71% 25%
12	N	66	5% 45% 45% 6%
12	O	66	2% 47% 44% 6%
13	P	48	2% 44% 44% 10%
13	X	48	2% 63% 25% 10%

2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 52739 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 RPO1N SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	841	Total	C	N	O	S	0	0	0
			6691	4256	1183	1226	26			
1	W	841	Total	C	N	O	S	0	0	0
			6691	4256	1183	1226	26			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE RPO2 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1090	Total	C	N	O	S	0	0	0
			8652	5484	1534	1605	29			
2	R	1090	Total	C	N	O	S	0	0	0
			8652	5484	1534	1605	29			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE RPO1C SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	367	Total	C	N	O	S	0	0	0
			2833	1797	481	547	8			
3	Y	367	Total	C	N	O	S	0	0	0
			2833	1797	481	547	8			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASERPO3 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	260	Total	C	N	O	S	0	0	0
			2071	1332	334	392	13			
4	S	260	Total	C	N	O	S	0	0	0
			2071	1332	334	392	13			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE RPO7 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	174	Total	C	N	O	S	0	0	0
			1384	893	232	255	4			
5	T	174	Total	C	N	O	S	0	0	0
			1384	893	232	255	4			

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE RPO4 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	91	Total	C	N	O	S	0	0	1
			702	439	115	145	3			
6	U	91	Total	C	N	O	S	0	0	1
			702	439	115	145	3			

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE RPO8 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	113	Total	C	N	O	S	0	0	0
			901	572	152	173	4			
7	V	113	Total	C	N	O	S	0	0	0
			901	572	152	173	4			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE RPO5 SUBUNIT.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	H	74	Total	C	N	O	0	0	0
			609	396	108	105			
8	Z	74	Total	C	N	O	0	0	0
			609	396	108	105			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE RPO6 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	82	Total	C	N	O	S	0	0	0
			658	420	121	116	1			
9	K	82	Total	C	N	O	S	0	0	0
			658	420	121	116	1			

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASE RPO13 SUBUNIT.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
10	J	35	Total	C	N	O	0	0	0
			301	186	53	62			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	Q	33	Total	C	N	O	S	0	0	1
			274	172	45	56	1			

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE RPO11 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	L	92	Total	C	N	O	S	0	0	1
			708	454	115	137	2			
11	M	92	Total	C	N	O	S	0	0	1
			708	454	115	137	2			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE RPO10 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	N	64	Total	C	N	O	S	0	0	0
			514	327	93	87	7			
12	O	64	Total	C	N	O	S	0	0	0
			514	327	93	87	7			

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASE RPO12 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	P	43	Total	C	N	O	S	0	0	0
			346	230	58	53	5			
13	X	43	Total	C	N	O	S	0	0	0
			346	230	58	53	5			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total	Zn	0	0
			2	2		
14	B	1	Total	Zn	0	0
			1	1		
14	N	1	Total	Zn	0	0
			1	1		
14	O	1	Total	Zn	0	0
			1	1		
14	P	1	Total	Zn	0	0
			1	1		
14	R	1	Total	Zn	0	0
			1	1		

Continued on next page...

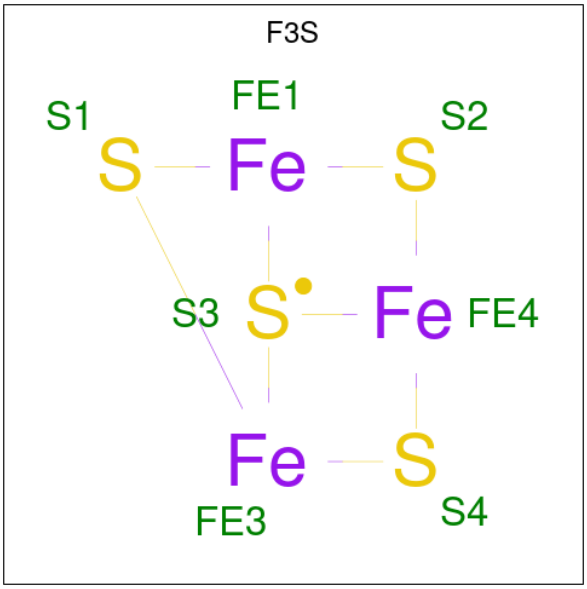
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	W	2	Total Zn 2 2	0	0
14	X	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
15	W	1	Total Mg 1 1	0	0

- Molecule 16 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).



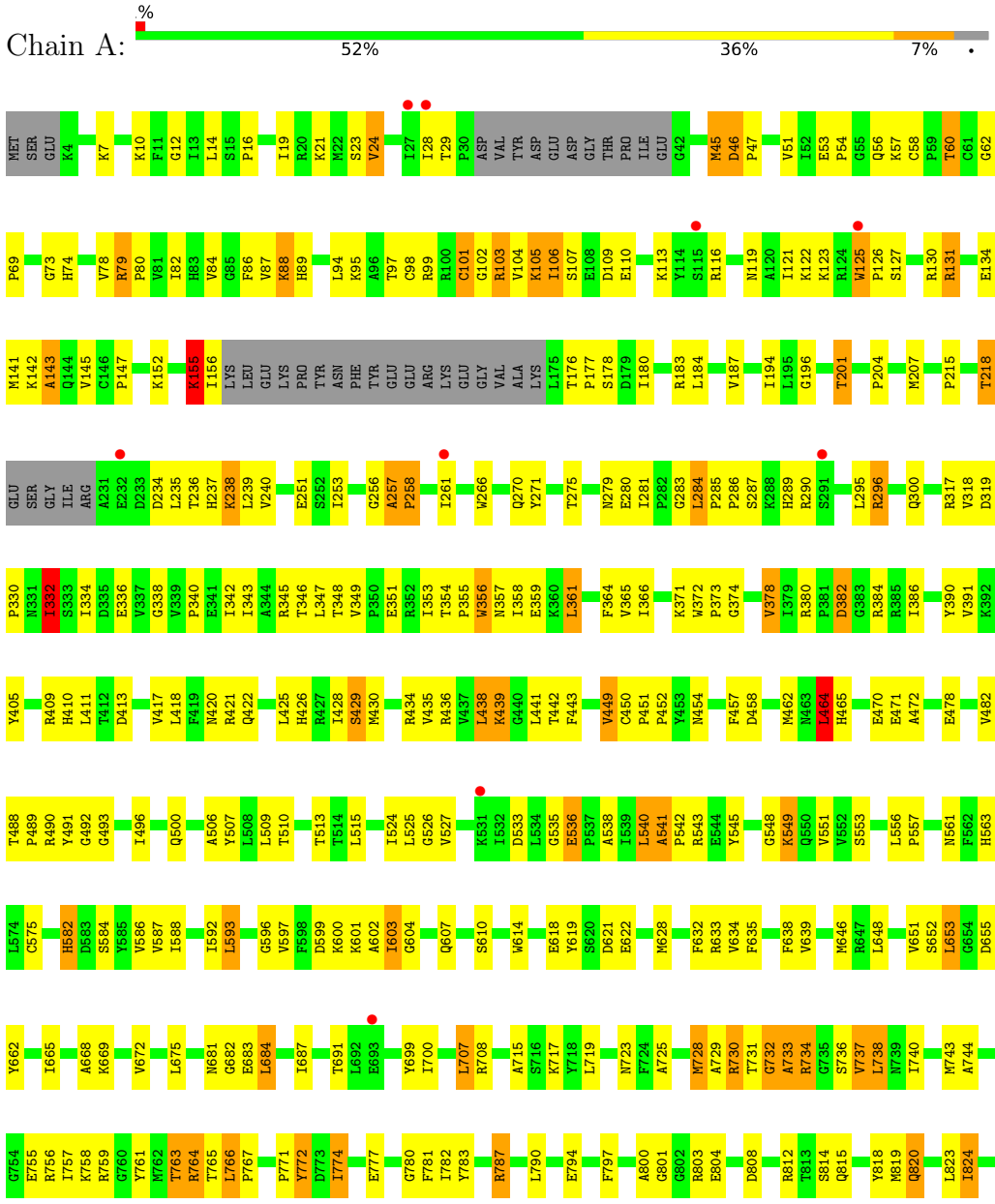
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	D	1	Total Fe S 7 3 4	0	0
16	S	1	Total Fe S 7 3 4	0	0

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.

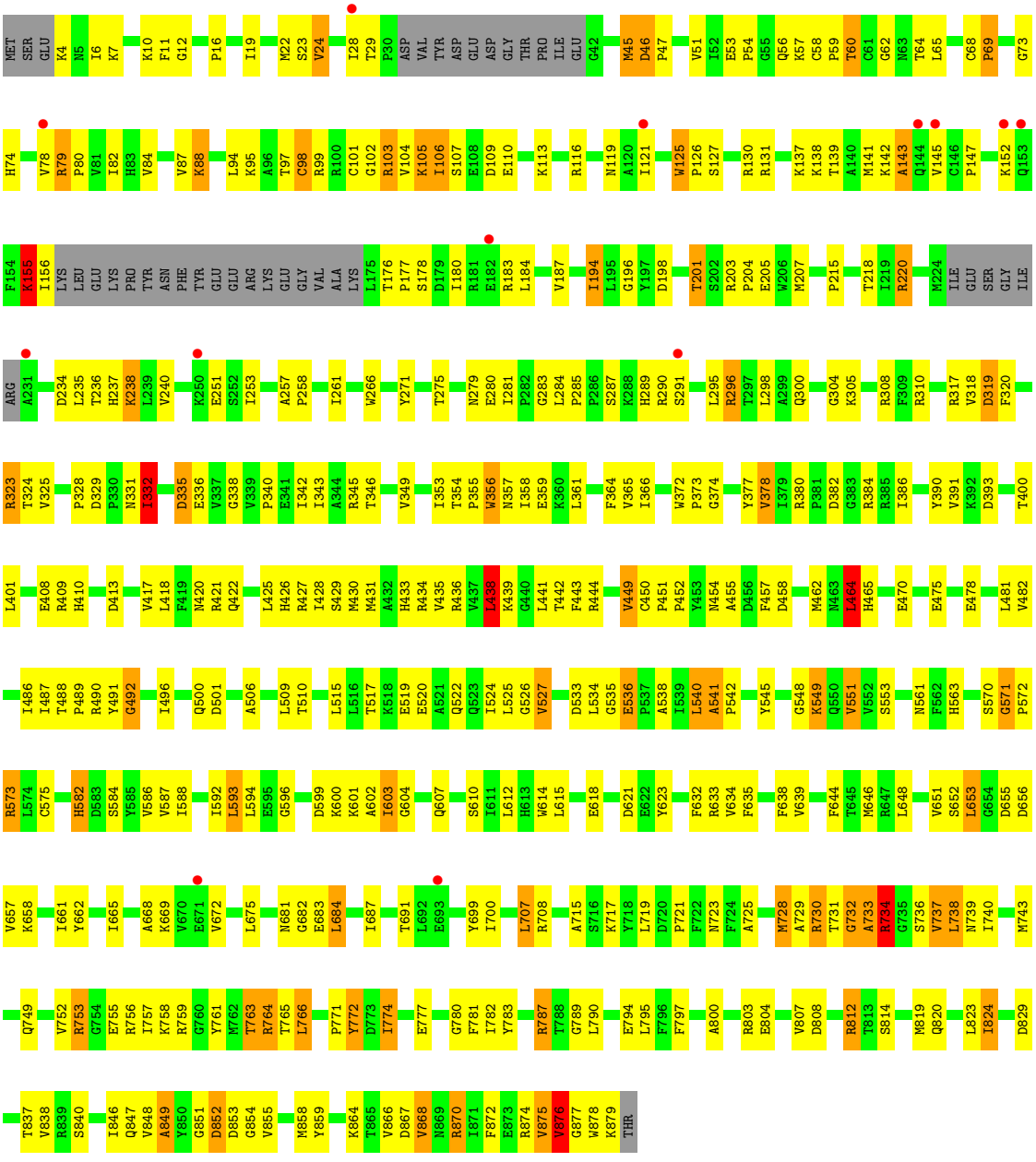
• Molecule 1: DNA-DIRECTED RNA POLYMERASE RPO1N SUBUNIT

Chain A:



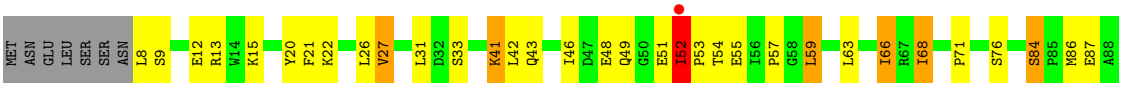
• Molecule 1: DNA-DIRECTED RNA POLYMERASE RPO1N SUBUNIT

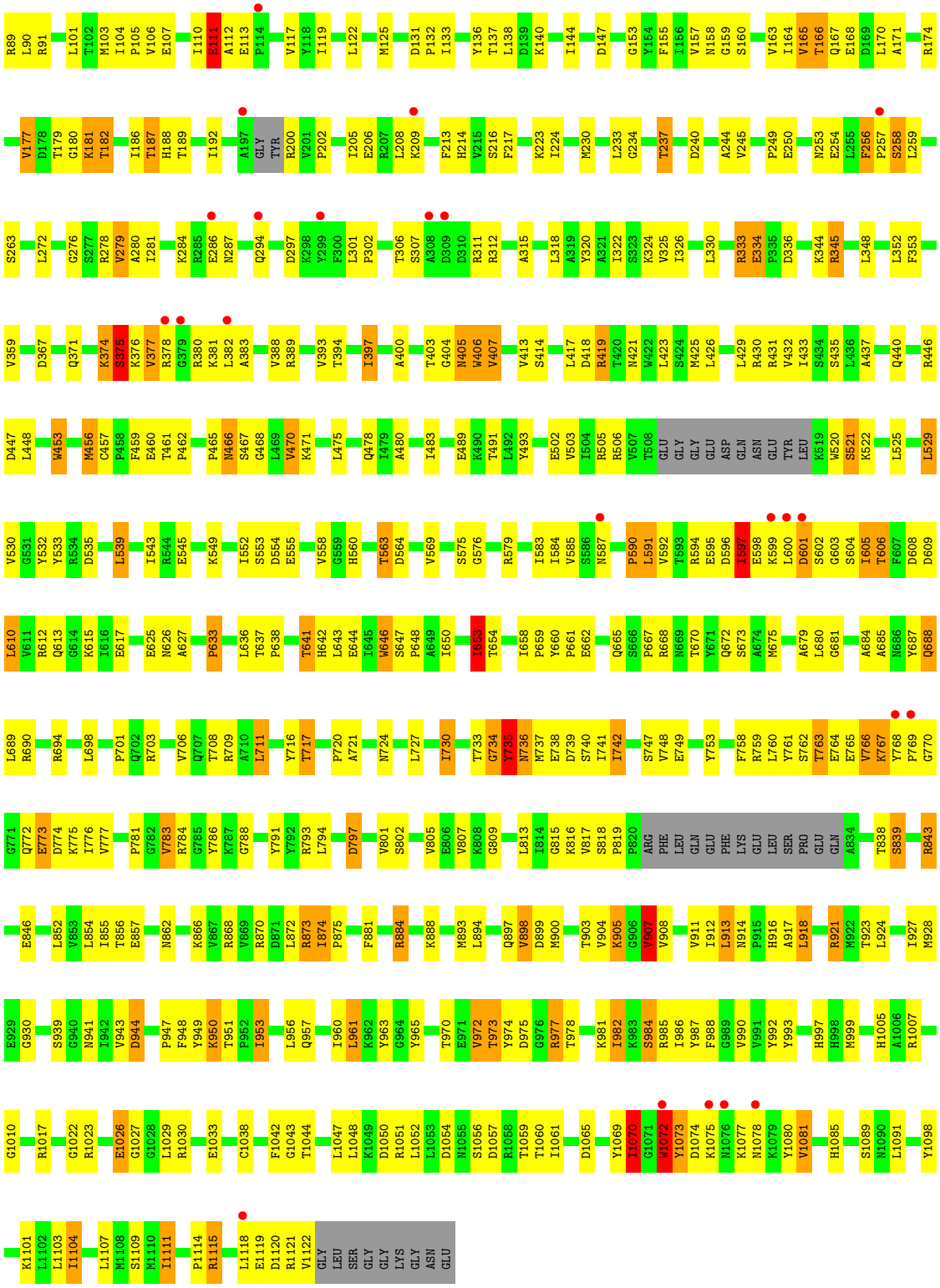
Chain W:



• Molecule 2: DNA-DIRECTED RNA POLYMERASE RPO2 SUBUNIT

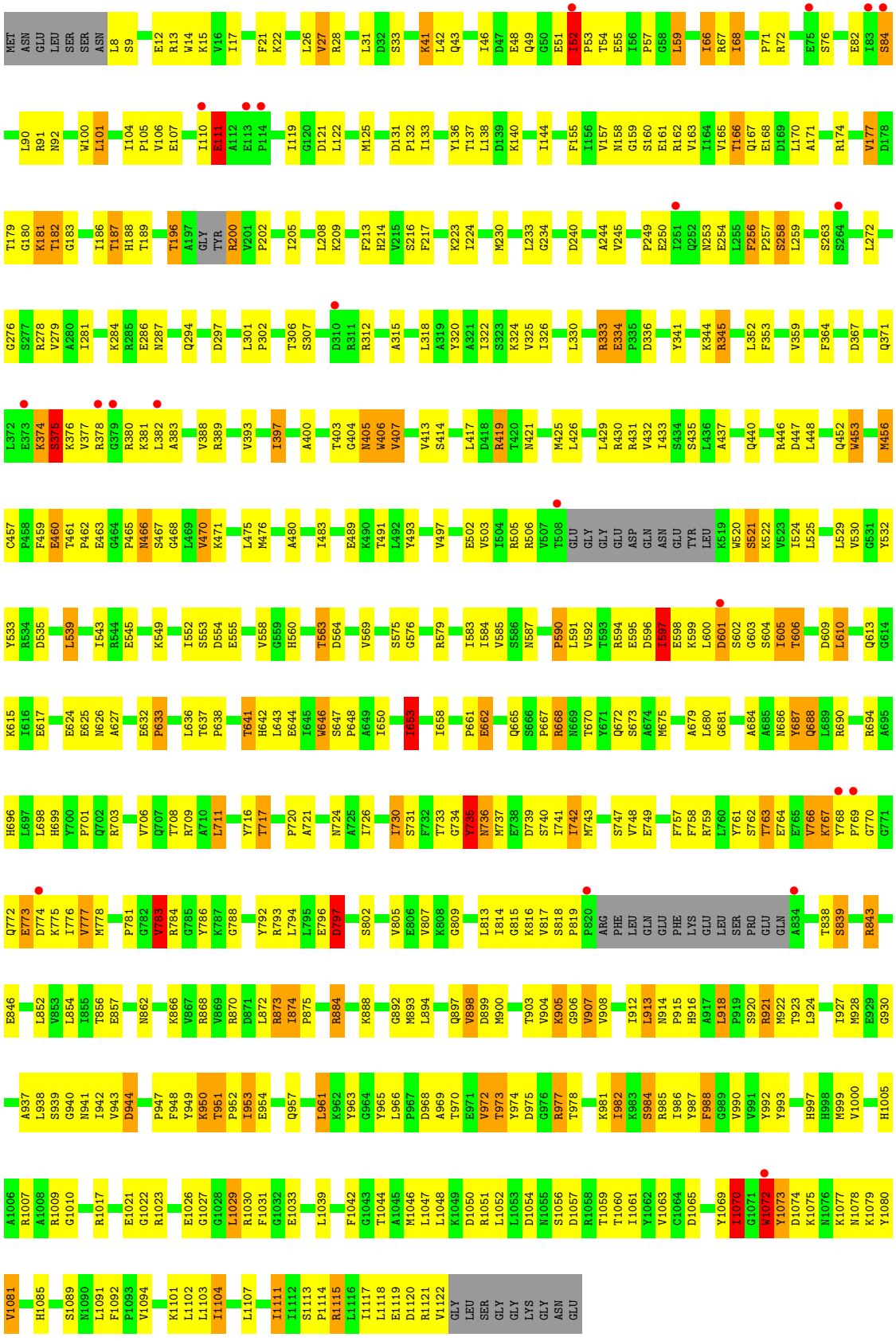
Chain B:





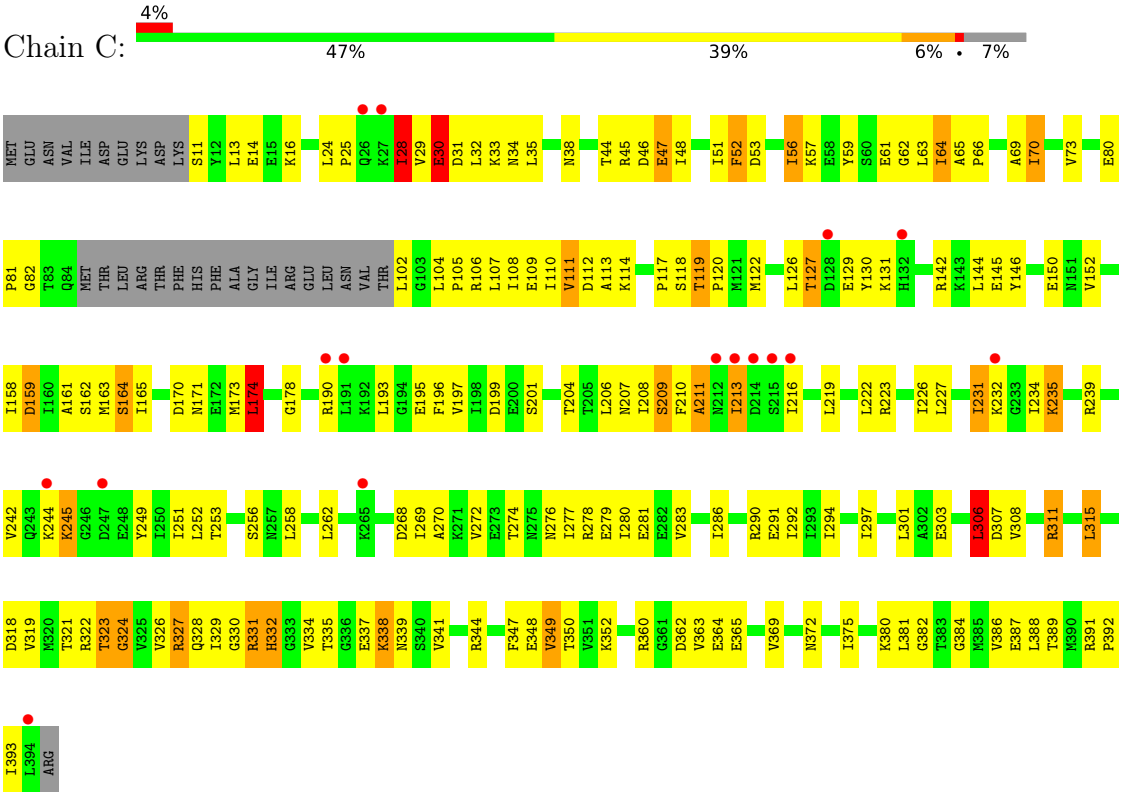
• Molecule 2: DNA-DIRECTED RNA POLYMERASE RPO2 SUBUNIT

Chain R:



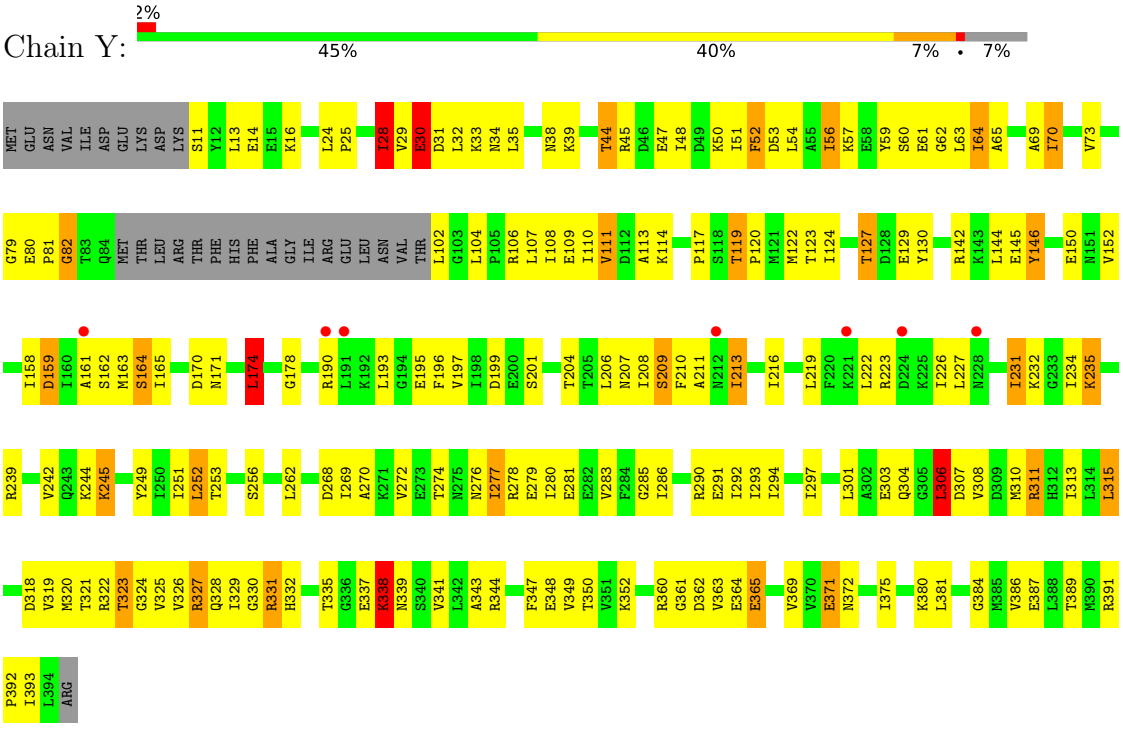
• Molecule 3: DNA-DIRECTED RNA POLYMERASE RPO1C SUBUNIT

Chain C:



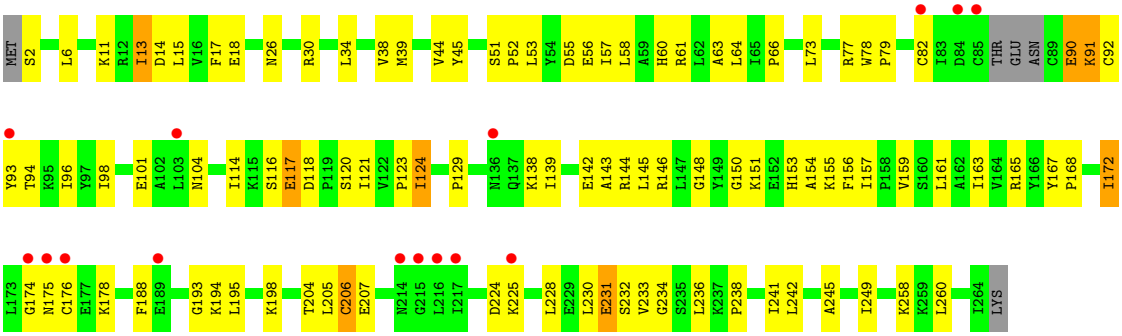
• Molecule 3: DNA-DIRECTED RNA POLYMERASE RPO1C SUBUNIT

Chain Y:



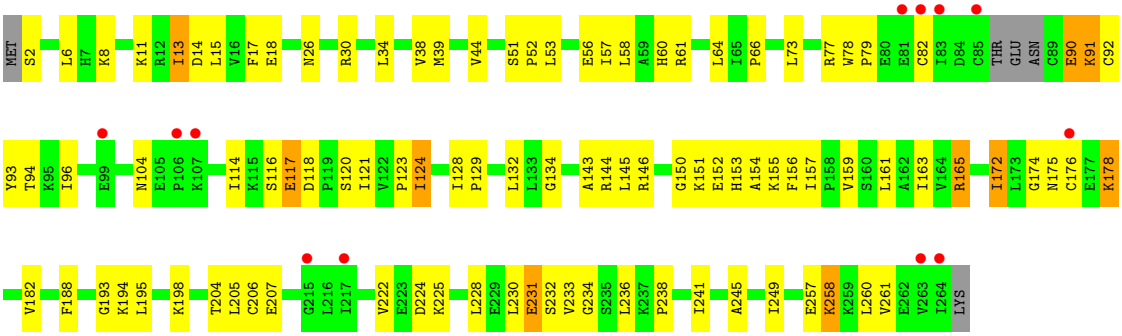
• Molecule 4: DNA-DIRECTED RNA POLYMERASERPO3 SUBUNIT

Chain D:



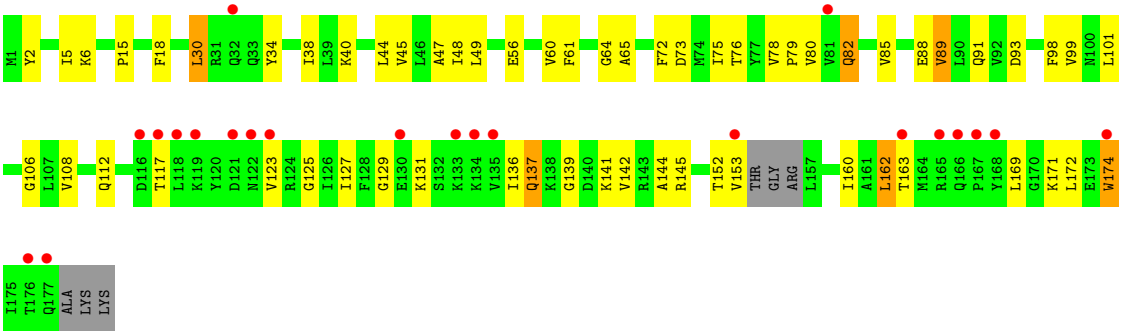
• Molecule 4: DNA-DIRECTED RNA POLYMERASERPO3 SUBUNIT

Chain S:



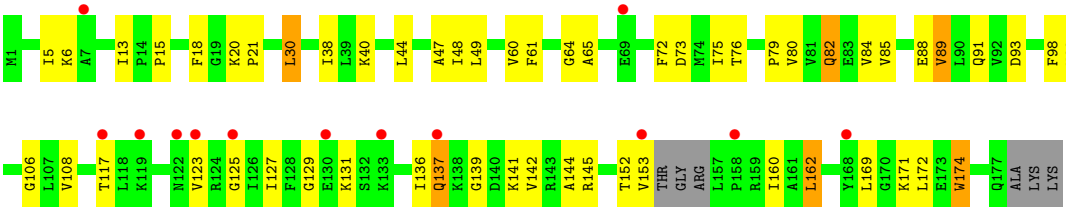
• Molecule 5: DNA-DIRECTED RNA POLYMERASE RPO7 SUBUNIT

Chain E:



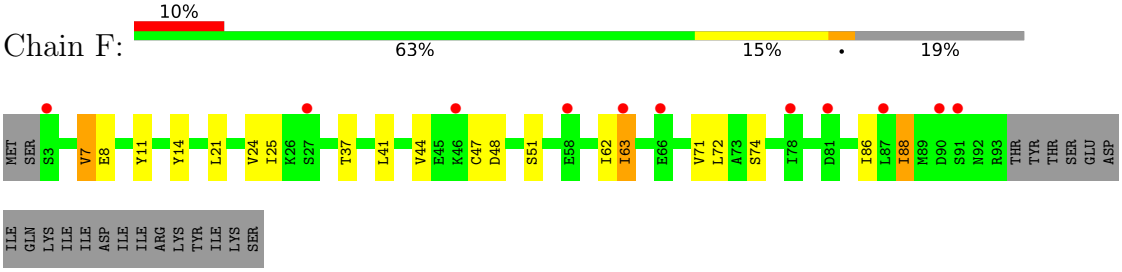
• Molecule 5: DNA-DIRECTED RNA POLYMERASE RPO7 SUBUNIT

Chain T:



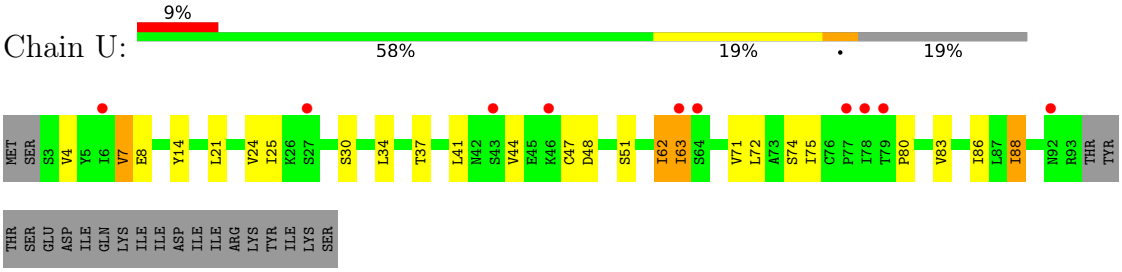
• Molecule 6: DNA-DIRECTED RNA POLYMERASE RPO4 SUBUNIT

Chain F:



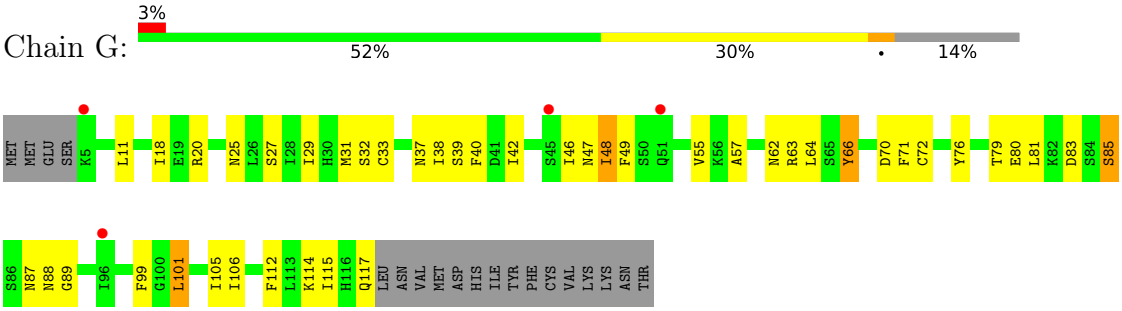
• Molecule 6: DNA-DIRECTED RNA POLYMERASE RPO4 SUBUNIT

Chain U:



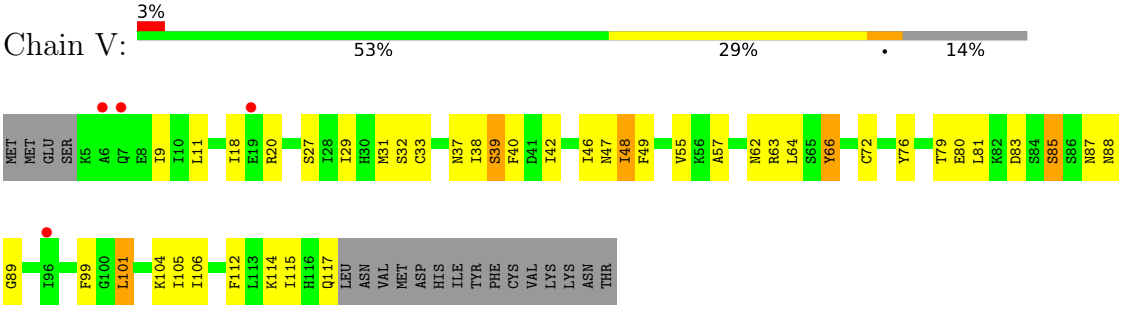
• Molecule 7: DNA-DIRECTED RNA POLYMERASE RPO8 SUBUNIT

Chain G:



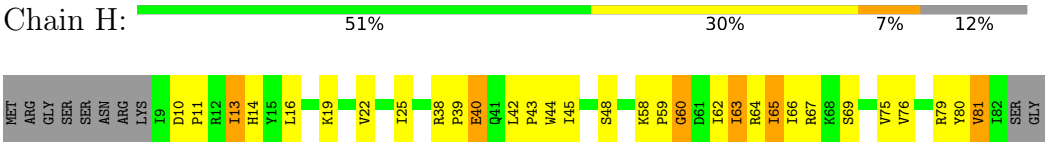
• Molecule 7: DNA-DIRECTED RNA POLYMERASE RPO8 SUBUNIT

Chain V:



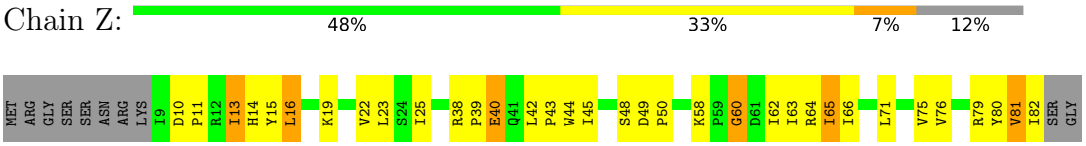
• Molecule 8: DNA-DIRECTED RNA POLYMERASE RPO5 SUBUNIT

Chain H:



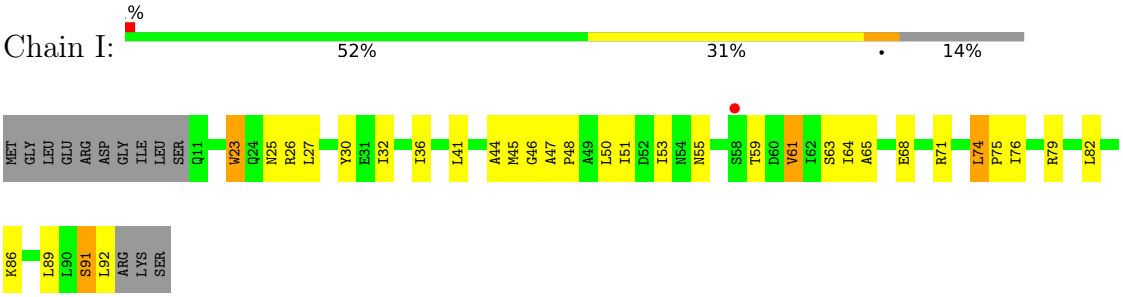
• Molecule 8: DNA-DIRECTED RNA POLYMERASE RPO5 SUBUNIT

Chain Z:



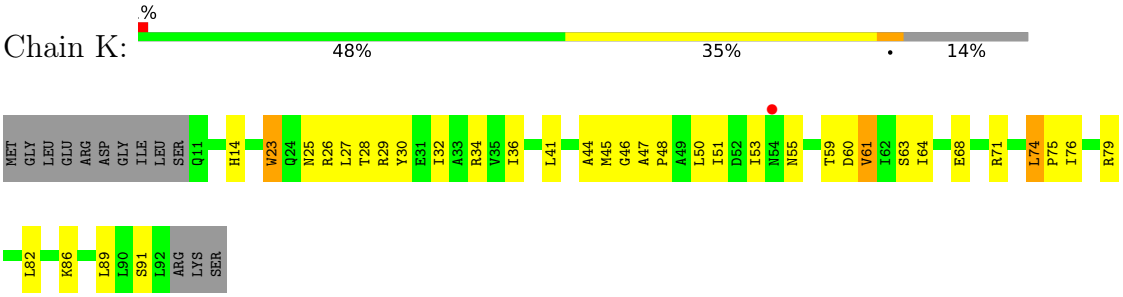
• Molecule 9: DNA-DIRECTED RNA POLYMERASE RPO6 SUBUNIT

Chain I:



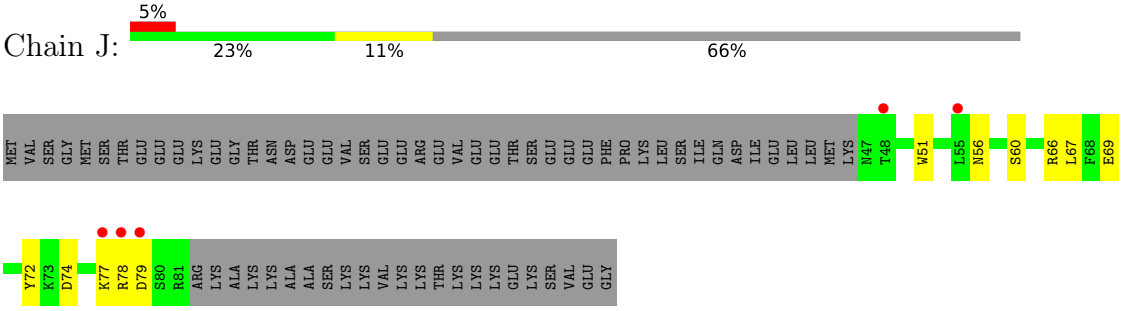
• Molecule 9: DNA-DIRECTED RNA POLYMERASE RPO6 SUBUNIT

Chain K:



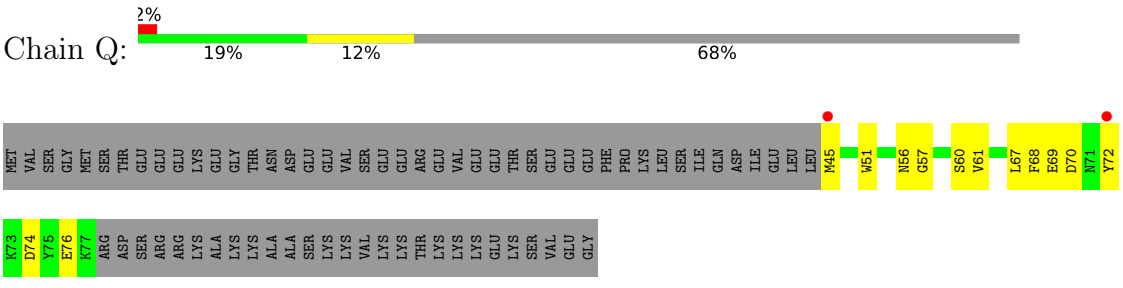
• Molecule 10: DNA-DIRECTED RNA POLYMERASE RPO13 SUBUNIT

Chain J:



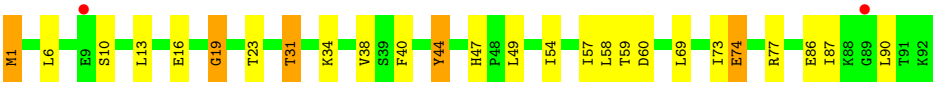
• Molecule 10: DNA-DIRECTED RNA POLYMERASE RPO13 SUBUNIT

Chain Q:

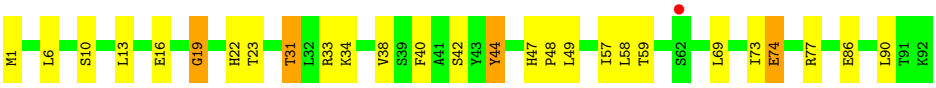


• Molecule 11: DNA-DIRECTED RNA POLYMERASE RPO11 SUBUNIT

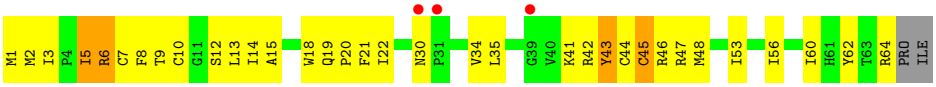
Chain L:



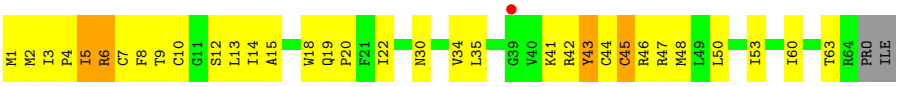
• Molecule 11: DNA-DIRECTED RNA POLYMERASE RPO11 SUBUNIT



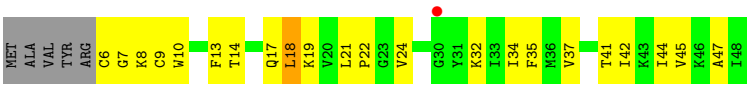
• Molecule 12: DNA-DIRECTED RNA POLYMERASE RPO10 SUBUNIT



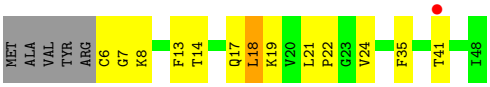
• Molecule 12: DNA-DIRECTED RNA POLYMERASE RPO10 SUBUNIT



• Molecule 13: DNA-DIRECTED RNA POLYMERASE RPO12 SUBUNIT



• Molecule 13: DNA-DIRECTED RNA POLYMERASE RPO12 SUBUNIT



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	196.09Å 211.87Å 238.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.20 – 3.52 30.21 – 3.52	Depositor EDS
% Data completeness (in resolution range)	92.4 (30.20-3.52) 92.4 (30.21-3.52)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.53 (at 3.56Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.333 , 0.384 0.332 , 0.381	Depositor DCC
R_{free} test set	5721 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	79.9	Xtriage
Anisotropy	0.651	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 47.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	52739	wwPDB-VP
Average B, all atoms (Å^2)	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.85 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.4742e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: F3S, MG, 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	0.44	0/6834	0.61	1/9247 (0.0%)
1	W	0.44	0/6834	0.63	2/9247 (0.0%)
2	B	0.45	0/8816	0.64	0/11926
2	R	0.45	0/8816	0.63	0/11926
3	C	0.44	0/2857	0.62	1/3847 (0.0%)
3	Y	0.47	0/2857	0.64	1/3847 (0.0%)
4	D	0.40	0/2106	0.57	0/2845
4	S	0.39	0/2106	0.57	1/2845 (0.0%)
5	E	0.40	0/1405	0.55	0/1899
5	T	0.42	0/1405	0.56	0/1899
6	F	0.40	0/710	0.51	0/963
6	U	0.40	0/710	0.51	0/963
7	G	0.47	0/913	0.57	0/1224
7	V	0.47	0/913	0.58	0/1224
8	H	0.43	0/623	0.62	0/845
8	Z	0.45	0/623	0.64	0/845
9	I	0.46	0/667	0.66	0/903
9	K	0.45	0/667	0.64	0/903
10	J	0.50	0/305	0.62	0/408
10	Q	0.51	0/278	0.58	0/373
11	L	0.43	0/718	0.56	0/970
11	M	0.43	0/718	0.55	0/970
12	N	0.44	0/524	0.58	0/706
12	O	0.45	0/524	0.57	0/706
13	P	0.52	0/354	0.66	0/475
13	X	0.48	0/354	0.65	0/475
All	All	0.44	0/53637	0.61	6/72481 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	Y	174	LEU	CA-CB-CG	6.97	131.33	115.30
3	C	174	LEU	CA-CB-CG	6.75	130.82	115.30
1	W	464	LEU	CA-CB-CG	6.63	130.55	115.30
1	A	464	LEU	CA-CB-CG	6.38	129.98	115.30
1	W	438	LEU	CA-CB-CG	5.22	127.30	115.30

There are no chirality outliers.

There are no planarity outliers.

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	6691	0	6756	296	0
1	W	6691	0	6760	327	0
2	B	8652	0	8795	352	0
2	R	8652	0	8796	376	0
3	C	2833	0	2992	132	0
3	Y	2833	0	2992	157	0
4	D	2071	0	2116	78	0
4	S	2071	0	2116	76	0
5	E	1384	0	1444	40	0
5	T	1384	0	1444	35	0
6	F	702	0	708	16	0
6	U	702	0	708	19	0
7	G	901	0	912	35	0
7	V	901	0	912	33	0
8	H	609	0	640	26	0
8	Z	609	0	640	33	0
9	I	658	0	692	25	0
9	K	658	0	692	31	0
10	J	301	0	284	8	0
10	Q	274	0	258	9	0
11	L	708	0	739	16	0
11	M	708	0	739	23	0
12	N	514	0	528	30	0
12	O	514	0	529	34	0
13	P	346	0	376	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	X	346	0	374	8	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	N	1	0	0	0	0
14	O	1	0	0	0	0
14	P	1	0	0	0	0
14	R	1	0	0	0	0
14	W	2	0	0	0	0
14	X	1	0	0	0	0
15	A	1	0	0	0	0
15	W	1	0	0	0	0
16	D	7	0	0	0	0
16	S	7	0	0	0	0
All	All	52739	0	53942	1946	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 18.

The worst 5 of 1946 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)
2:R:816:LYS:H	2:R:839:SER:HB3	1.13	1.12
2:B:856:THR:HG22	2:B:857:GLU:H	1.14	1.11
2:R:893:MET:HE2	2:R:894:LEU:H	1.10	1.10
2:R:856:THR:HG22	2:R:857:GLU:H	1.17	1.08
3:C:244:LYS:HA	3:C:245:LYS:HB3	1.32	1.07

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	833/880 (95%)	664 (80%)	130 (16%)	39 (5%)	2	22
1	W	833/880 (95%)	666 (80%)	127 (15%)	40 (5%)	2	21
2	B	1082/1131 (96%)	853 (79%)	164 (15%)	65 (6%)	1	17
2	R	1082/1131 (96%)	852 (79%)	160 (15%)	70 (6%)	1	15
3	C	363/395 (92%)	277 (76%)	66 (18%)	20 (6%)	2	19
3	Y	363/395 (92%)	276 (76%)	64 (18%)	23 (6%)	1	16
4	D	256/265 (97%)	226 (88%)	23 (9%)	7 (3%)	5	34
4	S	256/265 (97%)	226 (88%)	22 (9%)	8 (3%)	4	32
5	E	170/180 (94%)	148 (87%)	20 (12%)	2 (1%)	13	51
5	T	170/180 (94%)	148 (87%)	20 (12%)	2 (1%)	13	51
6	F	89/113 (79%)	78 (88%)	10 (11%)	1 (1%)	14	53
6	U	89/113 (79%)	79 (89%)	9 (10%)	1 (1%)	14	53
7	G	111/132 (84%)	92 (83%)	15 (14%)	4 (4%)	3	28
7	V	111/132 (84%)	92 (83%)	15 (14%)	4 (4%)	3	28
8	H	72/84 (86%)	60 (83%)	6 (8%)	6 (8%)	1	10
8	Z	72/84 (86%)	60 (83%)	7 (10%)	5 (7%)	1	14
9	I	80/95 (84%)	61 (76%)	15 (19%)	4 (5%)	2	20
9	K	80/95 (84%)	61 (76%)	15 (19%)	4 (5%)	2	20
10	J	33/104 (32%)	28 (85%)	5 (15%)	0	100	100
10	Q	31/104 (30%)	27 (87%)	4 (13%)	0	100	100
11	L	90/92 (98%)	80 (89%)	8 (9%)	2 (2%)	6	38
11	M	90/92 (98%)	81 (90%)	7 (8%)	2 (2%)	6	38
12	N	62/66 (94%)	42 (68%)	15 (24%)	5 (8%)	1	10
12	O	62/66 (94%)	42 (68%)	16 (26%)	4 (6%)	1	15
13	P	41/48 (85%)	34 (83%)	6 (15%)	1 (2%)	6	36
13	X	41/48 (85%)	35 (85%)	5 (12%)	1 (2%)	6	36
All	All	6562/7170 (92%)	5288 (81%)	954 (14%)	320 (5%)	2	21

5 of 320 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	145	VAL
1	A	287	SER
1	A	378	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	732	GLY
1	A	734	ARG

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	729/766 (95%)	647 (89%)	82 (11%)	6 29
1	W	729/766 (95%)	645 (88%)	84 (12%)	5 27
2	B	941/975 (96%)	805 (86%)	136 (14%)	3 19
2	R	941/975 (96%)	806 (86%)	135 (14%)	3 20
3	C	315/341 (92%)	274 (87%)	41 (13%)	4 22
3	Y	315/341 (92%)	276 (88%)	39 (12%)	4 24
4	D	233/238 (98%)	220 (94%)	13 (6%)	21 55
4	S	233/238 (98%)	220 (94%)	13 (6%)	21 55
5	E	154/158 (98%)	139 (90%)	15 (10%)	8 34
5	T	154/158 (98%)	139 (90%)	15 (10%)	8 34
6	F	84/107 (78%)	78 (93%)	6 (7%)	14 47
6	U	84/107 (78%)	78 (93%)	6 (7%)	14 47
7	G	106/125 (85%)	98 (92%)	8 (8%)	13 44
7	V	106/125 (85%)	97 (92%)	9 (8%)	10 40
8	H	67/75 (89%)	61 (91%)	6 (9%)	9 38
8	Z	67/75 (89%)	61 (91%)	6 (9%)	9 38
9	I	72/83 (87%)	67 (93%)	5 (7%)	15 48
9	K	72/83 (87%)	67 (93%)	5 (7%)	15 48
10	J	33/96 (34%)	26 (79%)	7 (21%)	1 6
10	Q	30/96 (31%)	25 (83%)	5 (17%)	2 13
11	L	79/80 (99%)	74 (94%)	5 (6%)	18 51
11	M	79/80 (99%)	76 (96%)	3 (4%)	33 65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
12	N	58/60 (97%)	54 (93%)	4 (7%)	15 48
12	O	58/60 (97%)	53 (91%)	5 (9%)	10 39
13	P	39/43 (91%)	36 (92%)	3 (8%)	13 43
13	X	39/43 (91%)	37 (95%)	2 (5%)	24 58
All	All	5817/6294 (92%)	5159 (89%)	658 (11%)	6 28

5 of 658 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	R	984	SER
1	W	500	GLN
2	R	1072	TRP
2	R	982	ILE
7	V	47	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 71 such sidechains are listed below:

Mol	Chain	Res	Type
7	V	62	ASN
1	W	83	HIS
1	W	567	ASN
2	B	1078	ASN
2	B	994	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 12 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	F3S	S	1265	4	0,9,9	-	-	-		
16	F3S	D	1265	4	0,9,9	-	-	-		

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	F3S	S	1265	4	-	-	0/3/3/3
16	F3S	D	1265	4	-	-	0/3/3/3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{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	841/880 (95%)	0.01	9 (1%) 80 69	65, 91, 152, 185	0
1	W	841/880 (95%)	-0.03	13 (1%) 73 61	61, 87, 141, 171	0
2	B	1090/1131 (96%)	0.03	24 (2%) 62 48	64, 87, 129, 155	0
2	R	1090/1131 (96%)	0.04	22 (2%) 65 52	64, 89, 120, 137	0
3	C	367/395 (92%)	0.19	16 (4%) 34 25	82, 106, 147, 155	0
3	Y	367/395 (92%)	0.00	7 (1%) 66 53	62, 89, 145, 160	0
4	D	260/265 (98%)	0.29	15 (5%) 23 17	90, 110, 127, 142	0
4	S	260/265 (98%)	0.25	12 (4%) 32 24	88, 106, 126, 140	0
5	E	174/180 (96%)	0.66	22 (12%) 3 4	98, 132, 196, 214	0
5	T	174/180 (96%)	0.53	13 (7%) 14 12	92, 125, 174, 189	0
6	F	91/113 (80%)	0.74	11 (12%) 4 4	135, 163, 191, 204	0
6	U	91/113 (80%)	0.83	10 (10%) 5 5	131, 160, 202, 216	0
7	G	113/132 (85%)	0.44	4 (3%) 44 33	87, 114, 138, 155	0
7	V	113/132 (85%)	0.37	4 (3%) 44 33	81, 112, 136, 149	0
8	H	74/84 (88%)	-0.09	0 100 100	75, 96, 118, 145	0
8	Z	74/84 (88%)	0.01	0 100 100	72, 95, 114, 144	0
9	I	82/95 (86%)	-0.01	1 (1%) 79 67	61, 79, 103, 126	0
9	K	82/95 (86%)	0.01	1 (1%) 79 67	64, 82, 103, 121	0
10	J	35/104 (33%)	0.44	5 (14%) 2 2	112, 126, 147, 153	0
10	Q	33/104 (31%)	0.62	2 (6%) 21 16	129, 166, 225, 285	0
11	L	92/92 (100%)	0.06	2 (2%) 62 48	75, 97, 136, 181	0
11	M	92/92 (100%)	0.05	1 (1%) 80 69	68, 91, 141, 184	0
12	N	64/66 (96%)	0.20	3 (4%) 31 23	79, 96, 119, 136	0
12	O	64/66 (96%)	-0.11	1 (1%) 72 59	74, 89, 103, 111	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
13	P	43/48 (89%)	0.31	1 (2%) 60 46	80, 105, 127, 141	0
13	X	43/48 (89%)	0.28	1 (2%) 60 46	90, 106, 121, 125	0
All	All	6650/7170 (92%)	0.12	200 (3%) 50 37	61, 96, 153, 285	0

The worst 5 of 200 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	117	THR	6.5
5	E	116	ASP	5.9
11	L	9	GLU	4.9
6	U	78	ILE	4.8
3	C	212	ASN	4.6

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, 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(Å^2)	Q<0.9
14	ZN	W	1881	1/1	0.92	0.07	81,81,81,81	0
15	MG	W	1882	1/1	0.93	0.21	81,81,81,81	0
15	MG	A	1882	1/1	0.94	0.43	79,79,79,79	0
14	ZN	W	1880	1/1	0.94	0.07	81,81,81,81	0
14	ZN	A	1880	1/1	0.96	0.06	79,79,79,79	0
14	ZN	X	1049	1/1	0.96	0.07	91,91,91,91	0
14	ZN	R	2123	1/1	0.97	0.03	76,76,76,76	0
14	ZN	B	2123	1/1	0.97	0.06	75,75,75,75	0
16	F3S	S	1265	7/7	0.97	0.13	86,86,86,86	0
14	ZN	P	1049	1/1	0.98	0.07	87,87,87,87	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
16	F3S	D	1265	7/7	0.98	0.15	88,88,88,88	0
14	ZN	A	1881	1/1	0.98	0.04	79,79,79,79	0
14	ZN	O	1065	1/1	0.99	0.07	81,81,81,81	0
14	ZN	N	1065	1/1	0.99	0.11	83,83,83,83	0

6.5 Other polymers (i)

There are no such residues in this entry.