wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:01 PM GMT

PDB ID	:	4QIW
Title	:	Crystal structure of euryarchaeal RNA polymerase from Thermococcus kodakarensis
Authors	:	Jun, S.-H.; Murakami, K.S.
Deposited on	:	2014-06-02
Resolution	:	3.50 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:

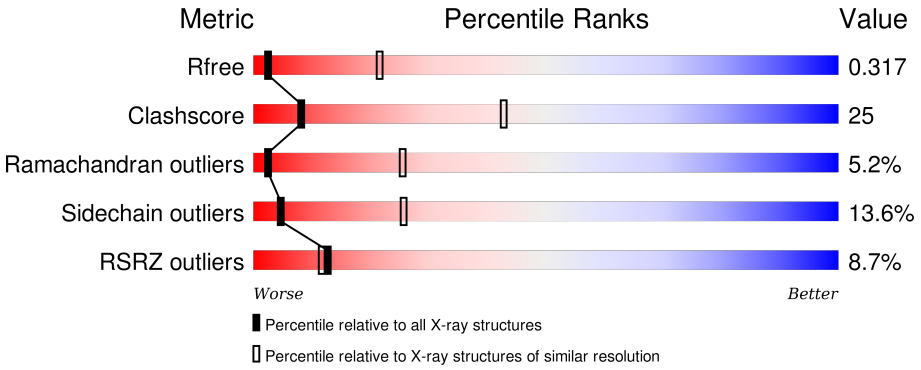
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.50 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	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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	A	906	9% 52% 35% 7% • 5%
1	I	906	9% 50% 36% 8% • 5%
2	B	1123	6% 48% 37% 9% • 5%
2	J	1123	4% 47% 39% 9% • 5%
3	C	391	19% 39% 41% 13% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	M	391	19% 38% 43% 12% 6%
4	D	259	10% 71% 25%
4	O	259	5% 70% 27%
5	E	190	13% 63% 28% 5%
5	Q	190	12% 65% 27% 5%
6	F	122	12% 74% 23%
6	R	122	5% 72% 26%
7	H	82	4% 40% 39% 13% 7%
7	S	82	5% 39% 41% 12% 7%
8	K	57	65% 25% 9%
8	T	57	5% 56% 35% 7%
9	L	100	3% 61% 26% 6%
9	U	100	12% 63% 25% 6%
10	N	65	2% 48% 37% 12%
10	V	65	3% 54% 34% 9%
11	P	49	10% 31% 51% 14%
11	W	49	6% 41% 37% 6% 14%

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
12	MG	I	1101	-	-	-	X
13	ZN	J	1301	-	-	-	X

2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 51069 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	863	Total	C	N	O	S	0	0	0
			6891	4357	1221	1275	38			
1	I	862	Total	C	N	O	S	0	0	0
			6875	4347	1220	1270	38			

- Molecule 2 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1069	Total	C	N	O	S	0	0	0
			8536	5396	1520	1584	36			
2	J	1069	Total	C	N	O	S	0	0	0
			8536	5396	1520	1584	36			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit A".

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	369	Total	C	N	O	S	0	0	0
			2882	1819	498	555	10			
3	M	369	Total	C	N	O	S	0	0	0
			2879	1816	498	555	10			

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	258	Total	C	N	O	S	0	0	0
			2066	1330	341	390	5			
4	O	258	Total	C	N	O	S	0	0	0
			2066	1330	341	390	5			

- Molecule 5 is a protein called DNA-directed RNA polymerase, subunit E'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	181	Total	C	N	O	S	0	0	0
			1465	939	250	267	9			
5	Q	181	Total	C	N	O	S	0	0	0
			1465	939	250	267	9			

- Molecule 6 is a protein called DNA-directed RNA polymerase, subunit F.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	122	Total	C	N	O	S	0	0	0
			1020	654	169	193	4			
6	R	122	Total	C	N	O	S	0	0	0
			1020	654	169	193	4			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	115	ILE	-	EXPRESSION TAG	UNP Q5JI52
F	116	ASP	-	EXPRESSION TAG	UNP Q5JI52
F	117	GLU	-	EXPRESSION TAG	UNP Q5JI52
F	118	TYR	-	EXPRESSION TAG	UNP Q5JI52
F	119	ARG	-	EXPRESSION TAG	UNP Q5JI52
F	120	PRO	-	EXPRESSION TAG	UNP Q5JI52
F	121	LEU	-	EXPRESSION TAG	UNP Q5JI52
F	122	GLU	-	EXPRESSION TAG	UNP Q5JI52
R	115	ILE	-	EXPRESSION TAG	UNP Q5JI52
R	116	ASP	-	EXPRESSION TAG	UNP Q5JI52
R	117	GLU	-	EXPRESSION TAG	UNP Q5JI52
R	118	TYR	-	EXPRESSION TAG	UNP Q5JI52
R	119	ARG	-	EXPRESSION TAG	UNP Q5JI52
R	120	PRO	-	EXPRESSION TAG	UNP Q5JI52
R	121	LEU	-	EXPRESSION TAG	UNP Q5JI52
R	122	GLU	-	EXPRESSION TAG	UNP Q5JI52

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	H	76	Total	C	N	O	0	0	0
			627	408	105	114			
7	S	76	Total	C	N	O	0	0	0
			627	408	105	114			

- Molecule 8 is a protein called DNA-directed RNA polymerase subunit K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	K	56	Total	C	N	O	S	0	0	0
			433	284	75	73	1			
8	T	56	Total	C	N	O	S	0	0	0
			433	284	75	73	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	L	94	Total	C	N	O	S	0	0	0
			775	493	134	146	2			
9	U	94	Total	C	N	O	S	0	0	0
			775	493	134	146	2			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	95	HIS	-	EXPRESSION TAG	UNP Q5JE88
L	96	HIS	-	EXPRESSION TAG	UNP Q5JE88
L	97	HIS	-	EXPRESSION TAG	UNP Q5JE88
L	98	HIS	-	EXPRESSION TAG	UNP Q5JE88
L	99	HIS	-	EXPRESSION TAG	UNP Q5JE88
L	100	HIS	-	EXPRESSION TAG	UNP Q5JE88
U	95	HIS	-	EXPRESSION TAG	UNP Q5JE88
U	96	HIS	-	EXPRESSION TAG	UNP Q5JE88
U	97	HIS	-	EXPRESSION TAG	UNP Q5JE88
U	98	HIS	-	EXPRESSION TAG	UNP Q5JE88
U	99	HIS	-	EXPRESSION TAG	UNP Q5JE88
U	100	HIS	-	EXPRESSION TAG	UNP Q5JE88

- Molecule 10 is a protein called DNA-directed RNA polymerase subunit N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	N	63	Total	C	N	O	S	0	0	0
			510	326	87	91	6			
10	V	63	Total	C	N	O	S	0	0	0
			510	326	87	91	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	P	42	Total	C	N	O	S	0	0	0
			329	206	65	54	4			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	W	42	Total	C	N	O	S	0	0	0
			329	206	65	54	4			

- Molecule 12 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	I	2	Total	Mg	0	0
			2	2		
12	A	2	Total	Mg	0	0
			2	2		

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	P	1	Total	Zn	0	0
			1	1		
13	J	1	Total	Zn	0	0
			1	1		
13	B	1	Total	Zn	0	0
			1	1		
13	I	2	Total	Zn	0	0
			2	2		
13	V	1	Total	Zn	0	0
			1	1		
13	W	1	Total	Zn	0	0
			1	1		
13	A	2	Total	Zn	0	0
			2	2		
13	N	1	Total	Zn	0	0
			1	1		

- Molecule 14 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total	O	0	0
			1	1		
14	B	1	Total	O	0	0
			1	1		
14	C	1	Total	O	0	0
			1	1		
14	I	1	Total	O	0	0
			1	1		

Continued on next page...

Continued from previous page...

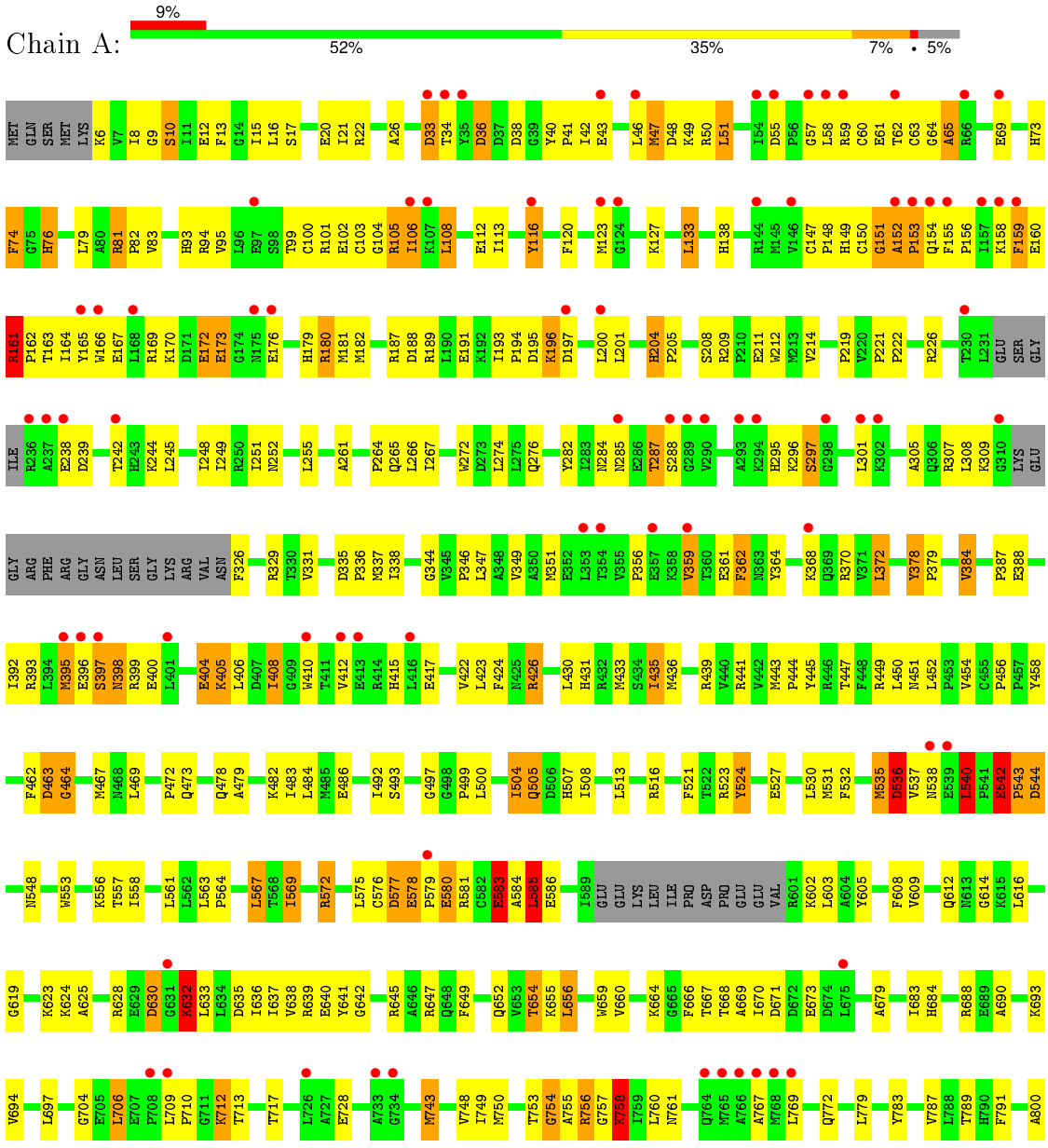
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	J	1	Total O 1 1	0	0
14	M	1	Total O 1 1	0	0

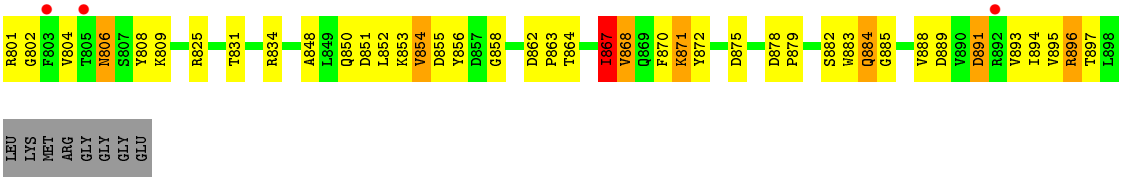
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 polymerase

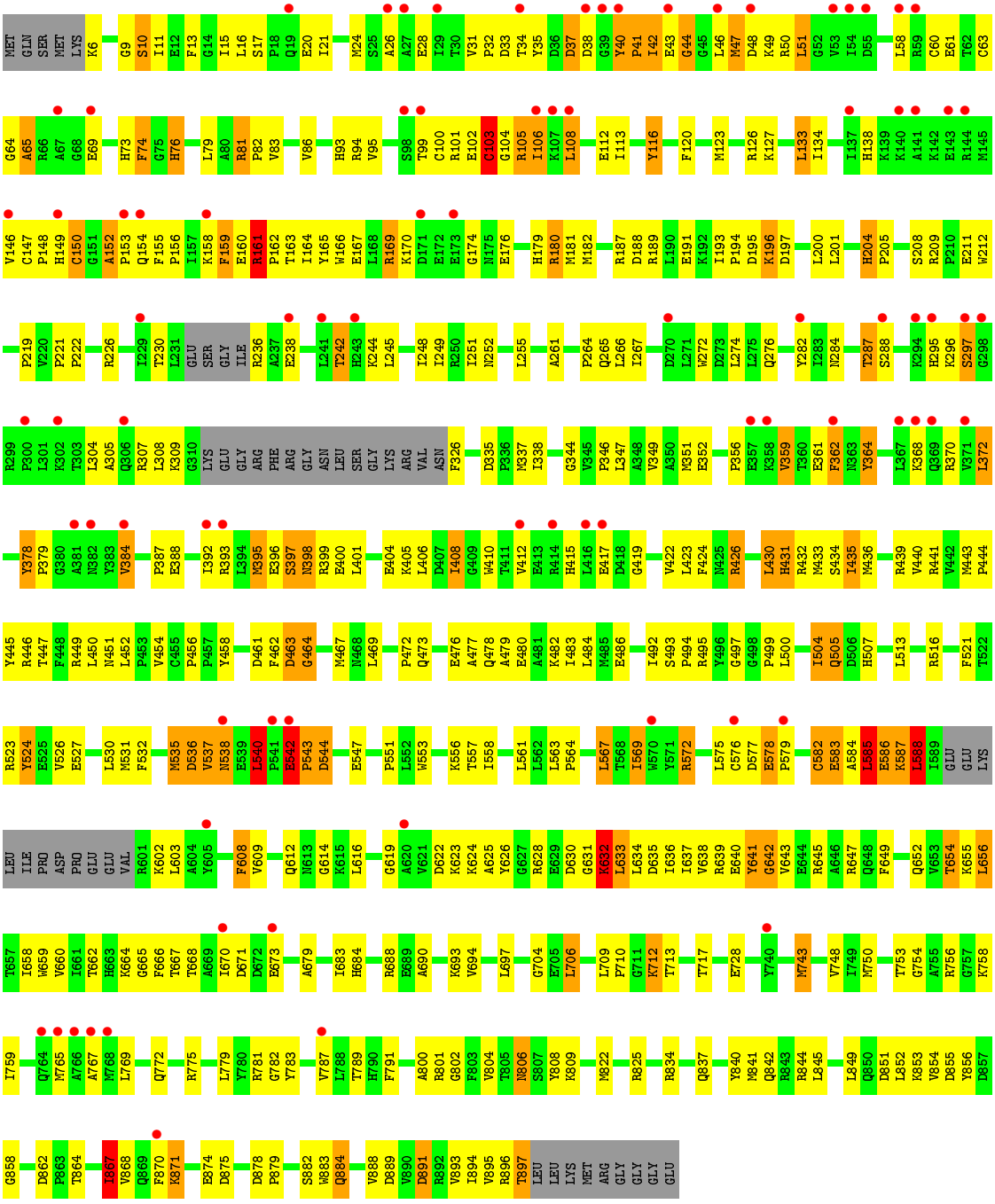
Chain A:





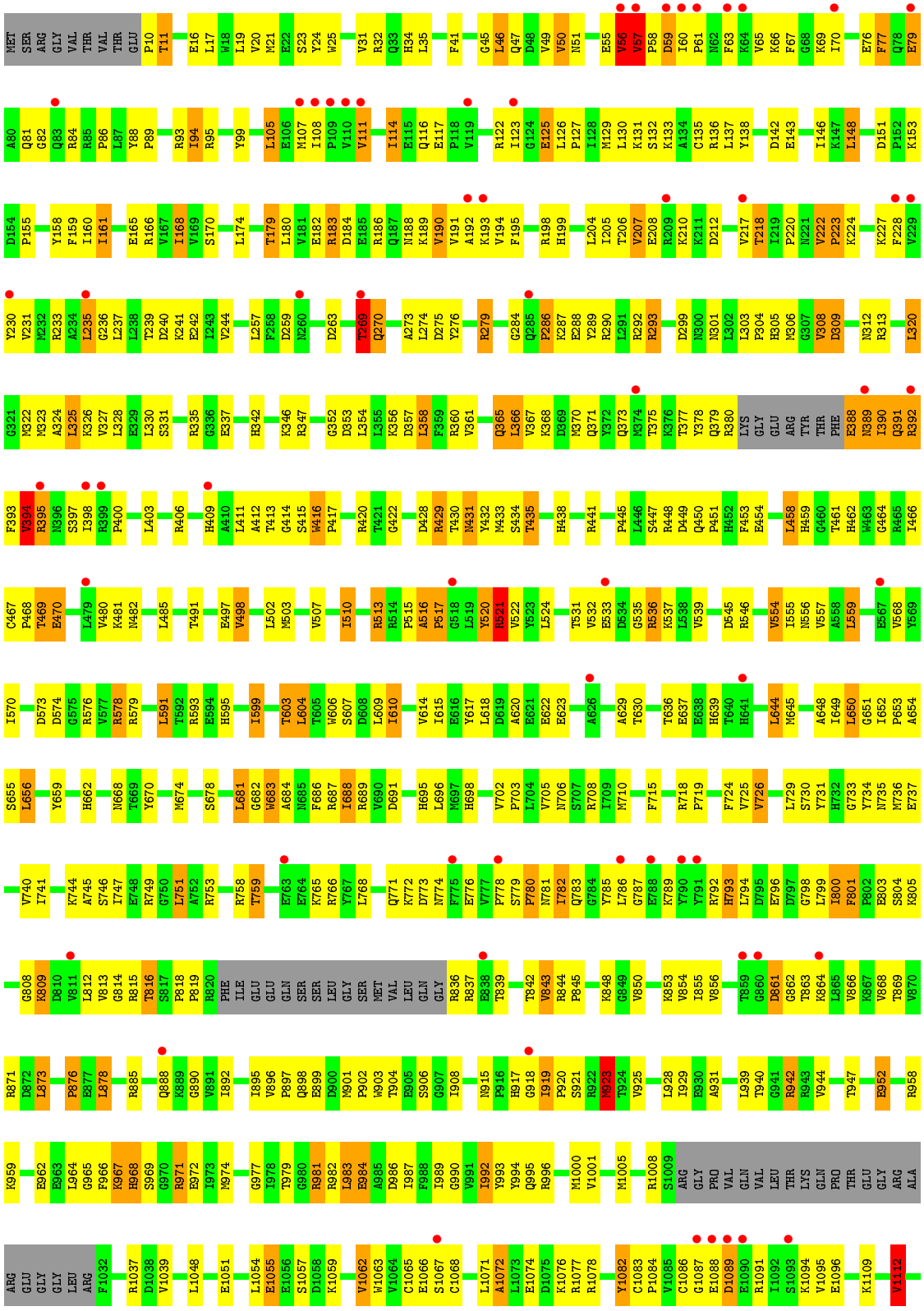
• Molecule 1: DNA-directed RNA polymerase

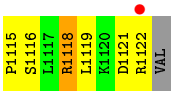
Chain I:



• Molecule 2: DNA-directed RNA polymerase

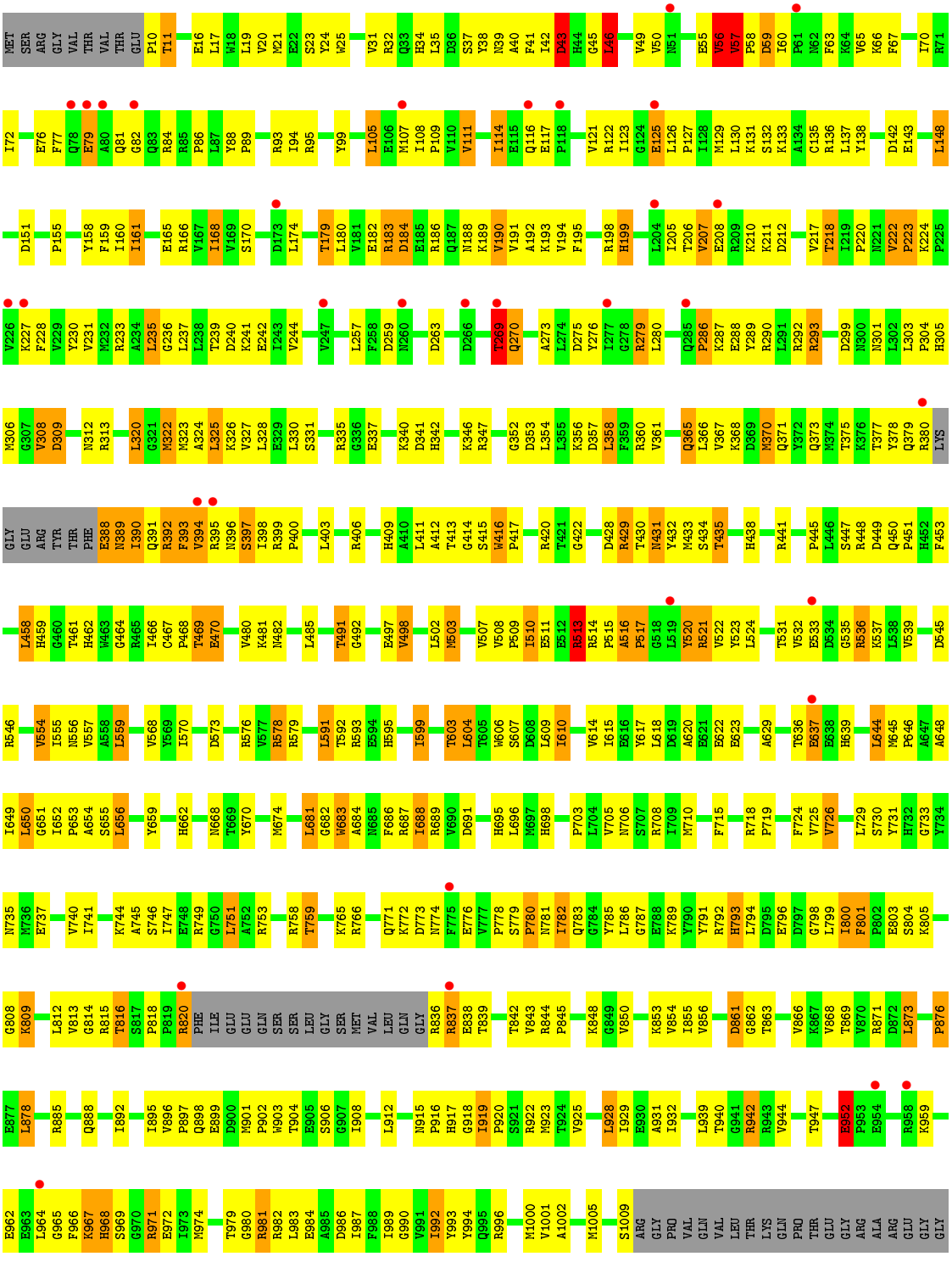
Chain B:

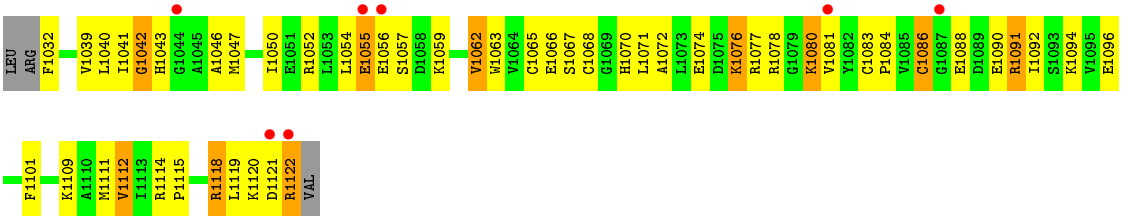




- Molecule 2: DNA-directed RNA polymerase

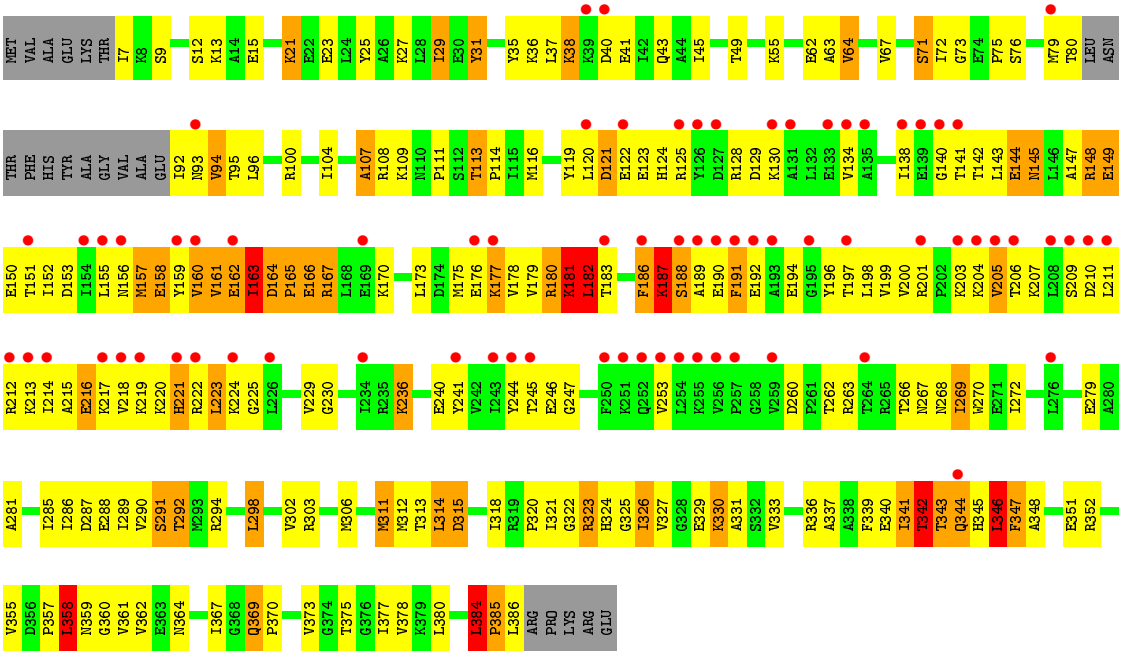
Chain J:





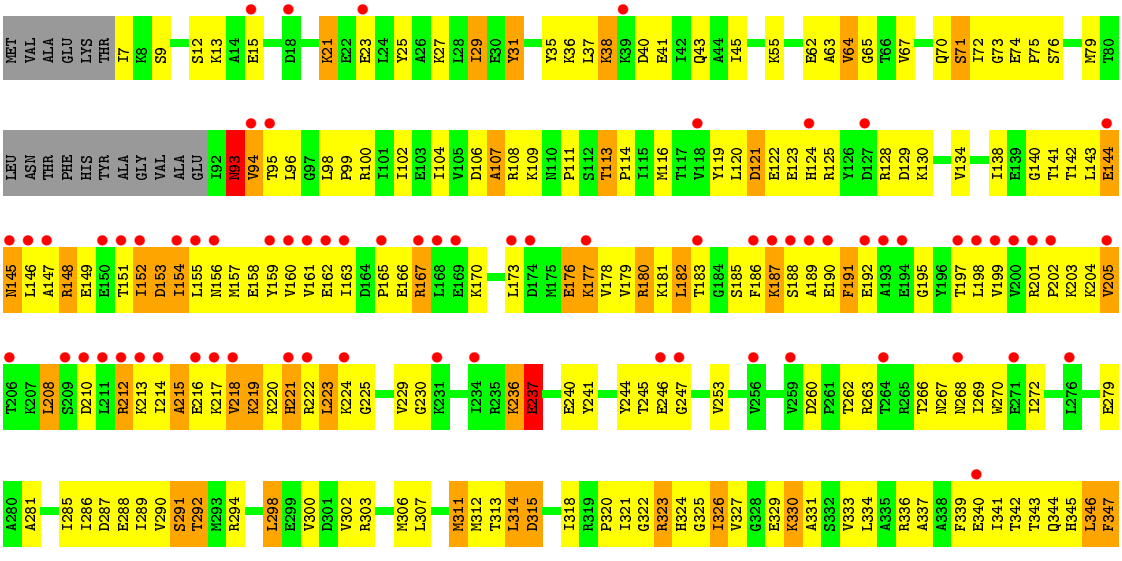
• Molecule 3: DNA-directed RNA polymerase subunit A"

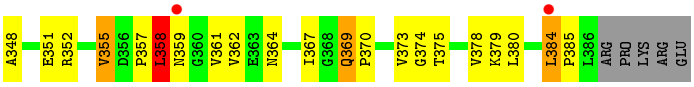
Chain C:



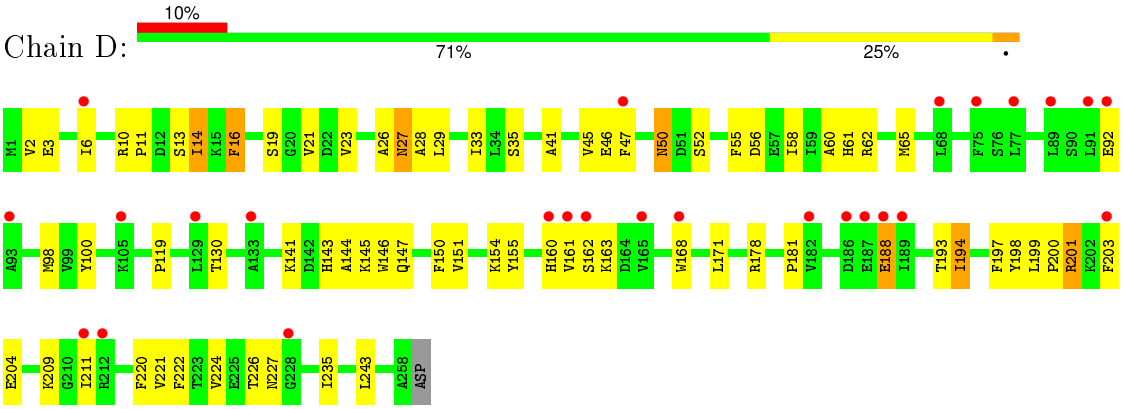
• Molecule 3: DNA-directed RNA polymerase subunit A"

Chain M:

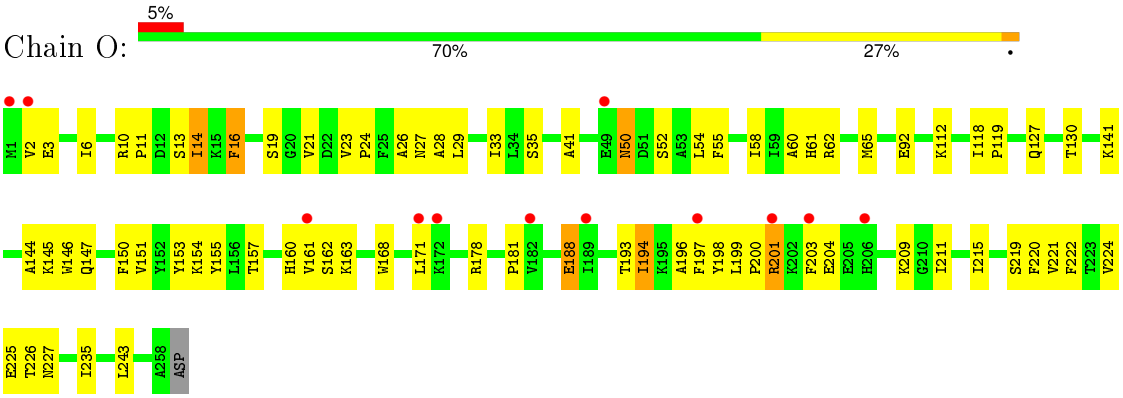




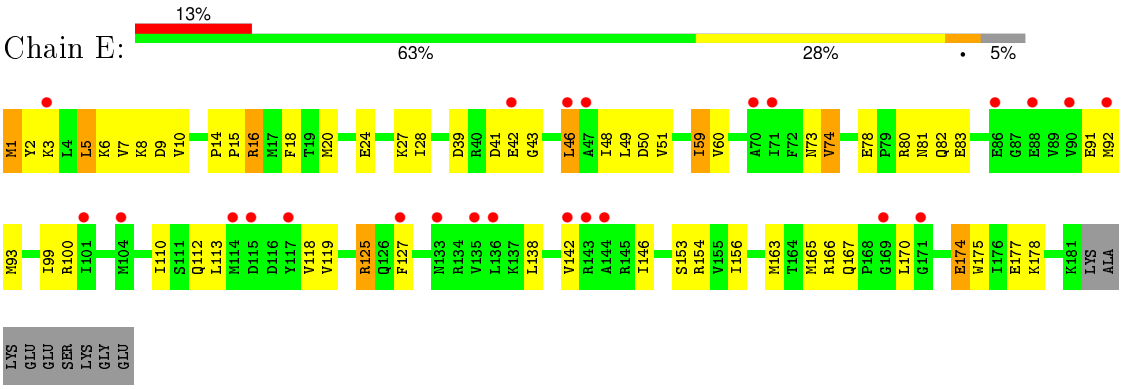
• Molecule 4: DNA-directed RNA polymerase subunit D



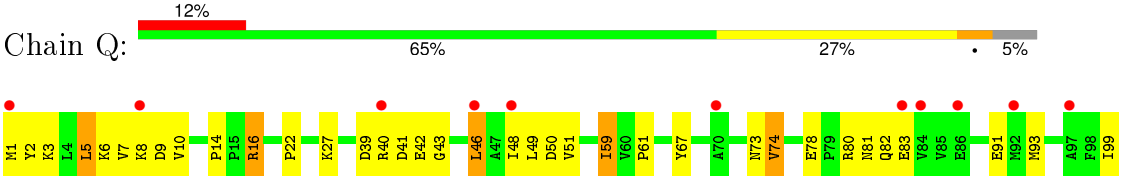
• Molecule 4: DNA-directed RNA polymerase subunit D

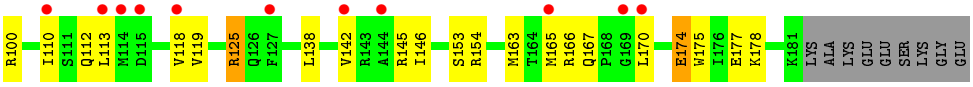


• Molecule 5: DNA-directed RNA polymerase, subunit E'

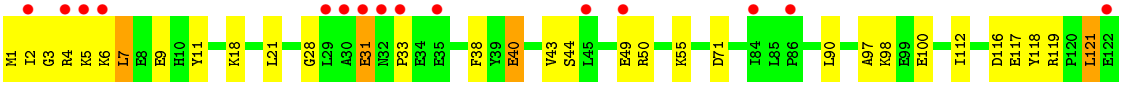


• Molecule 5: DNA-directed RNA polymerase, subunit E'

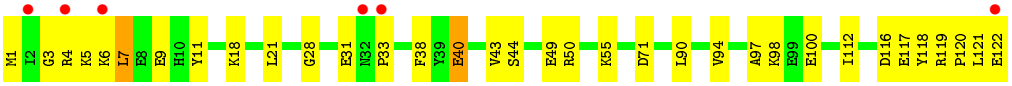




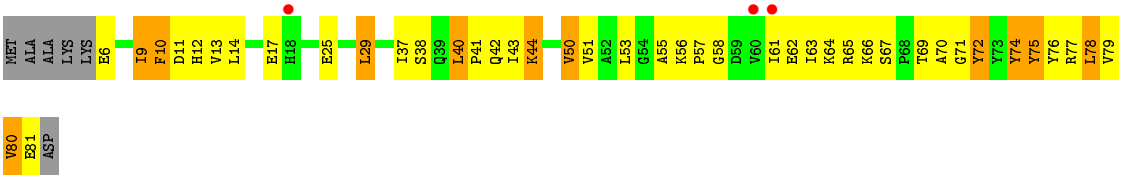
• Molecule 6: DNA-directed RNA polymerase, subunit F



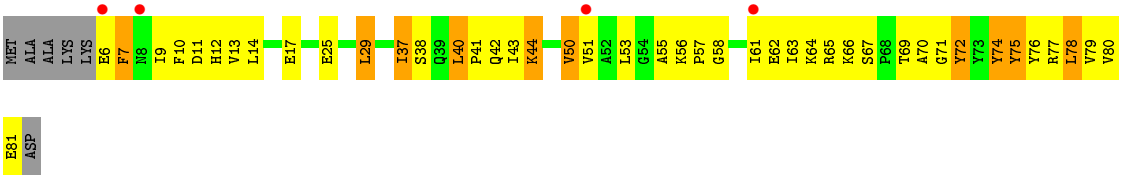
• Molecule 6: DNA-directed RNA polymerase, subunit F



• Molecule 7: DNA-directed RNA polymerase subunit H

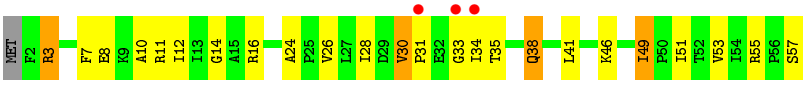


• Molecule 7: DNA-directed RNA polymerase subunit H

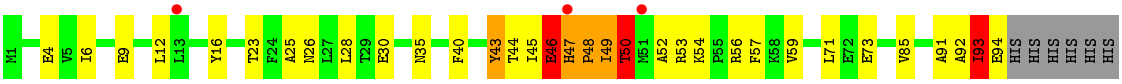


• Molecule 8: DNA-directed RNA polymerase subunit K

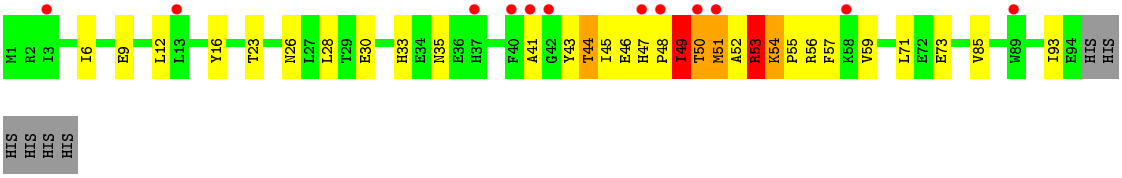
• Molecule 8: DNA-directed RNA polymerase subunit K



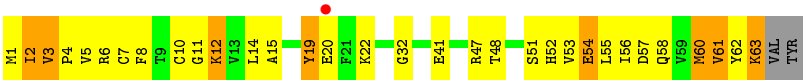
• Molecule 9: DNA-directed RNA polymerase subunit L



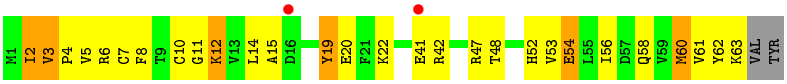
• Molecule 9: DNA-directed RNA polymerase subunit L



• Molecule 10: DNA-directed RNA polymerase subunit N

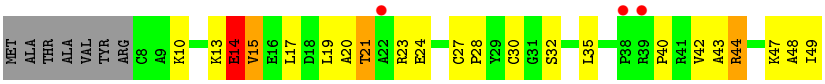


• Molecule 10: DNA-directed RNA polymerase subunit N



• Molecule 11: DNA-directed RNA polymerase subunit P

• Molecule 11: DNA-directed RNA polymerase subunit P



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	112.97Å 206.61Å 365.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.73 – 3.50 49.73 – 3.39	Depositor EDS
% Data completeness (in resolution range)	96.3 (49.73-3.50) 93.1 (49.73-3.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.26 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.3_1479)	Depositor
R, R_{free}	0.277 , 0.316 0.279 , 0.317	Depositor DCC
R_{free} test set	5185 reflections (5.03%)	DCC
Wilson B-factor (Å^2)	84.8	Xtriage
Anisotropy	0.078	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 105.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.31$, $< L^2 > = 0.15$	Xtriage
Outliers	1 of 113202 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.82	EDS
Total number of atoms	51069	wwPDB-VP
Average B, all atoms (Å^2)	163.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.84% 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, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.26	0/7028	0.50	1/9489 (0.0%)
1	I	0.29	2/7011 (0.0%)	0.52	1/9465 (0.0%)
2	B	0.29	1/8706 (0.0%)	0.53	3/11765 (0.0%)
2	J	0.28	1/8706 (0.0%)	0.52	3/11765 (0.0%)
3	C	0.27	0/2917	0.55	2/3936 (0.1%)
3	M	0.26	0/2914	0.53	1/3932 (0.0%)
4	D	0.27	0/2111	0.43	0/2858
4	O	0.28	0/2111	0.43	0/2858
5	E	0.23	0/1491	0.44	0/2008
5	Q	0.22	0/1491	0.43	0/2008
6	F	0.22	0/1040	0.40	0/1399
6	R	0.22	0/1040	0.40	0/1399
7	H	0.61	4/641 (0.6%)	0.57	0/866
7	S	0.60	4/641 (0.6%)	0.56	0/866
8	K	0.26	0/441	0.52	0/598
8	T	0.26	0/441	0.53	0/598
9	L	0.39	0/790	0.51	1/1066 (0.1%)
9	U	0.41	0/790	0.56	2/1066 (0.2%)
10	N	0.26	0/518	0.57	0/695
10	V	0.27	0/518	0.58	0/695
11	P	0.27	0/333	0.60	0/445
11	W	0.27	0/333	0.56	0/445
All	All	0.29	12/52012 (0.0%)	0.51	14/70222 (0.0%)

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

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	I	0	4
2	J	0	1
All	All	0	6

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	H	72	TYR	CE2-CZ	7.29	1.48	1.38
7	H	72	TYR	CE1-CZ	7.14	1.47	1.38
7	S	72	TYR	CE2-CZ	7.13	1.47	1.38
7	S	72	TYR	CG-CD1	7.11	1.48	1.39
1	I	150	CYS	CB-SG	-7.05	1.70	1.82

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	125	GLU	CA-CB-CG	7.21	129.26	113.40
3	C	164	ASP	C-N-CD	-6.97	105.27	120.60
1	I	150	CYS	CB-CA-C	-6.86	96.68	110.40
2	J	57	VAL	C-N-CD	-6.80	105.63	120.60
2	B	57	VAL	C-N-CD	-6.70	105.86	120.60

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	536	ASP	Peptide
1	I	582	CYS	Peptide
1	I	632	LYS	Peptide
1	I	641	TYR	Peptide
1	I	642	GLY	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	6891	0	6945	310	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	6875	0	6927	382	1
2	B	8536	0	8585	463	10
2	J	8536	0	8583	578	0
3	C	2882	0	2982	305	0
3	M	2879	0	2973	247	0
4	D	2066	0	2080	59	0
4	O	2066	0	2080	66	0
5	E	1465	0	1503	74	2
5	Q	1465	0	1503	68	3
6	F	1020	0	1024	41	12
6	R	1020	0	1024	48	7
7	H	627	0	642	29	0
7	S	627	0	642	34	0
8	K	433	0	466	16	0
8	T	433	0	466	19	0
9	L	775	0	770	51	0
9	U	775	0	770	67	0
10	N	510	0	523	43	0
10	V	510	0	523	26	0
11	P	329	0	356	27	5
11	W	329	0	355	15	0
12	A	2	0	0	0	0
12	I	2	0	0	0	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	N	1	0	0	0	0
13	P	1	0	0	1	0
13	V	1	0	0	0	0
13	W	1	0	0	0	0
14	A	1	0	0	0	0
14	B	1	0	0	1	0
14	C	1	0	0	0	0
14	I	1	0	0	2	0
14	J	1	0	0	0	0
14	M	1	0	0	3	0
All	All	51069	0	51722	2518	20

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 25.

The worst 5 of 2518 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:B:378:TYR:CE1	2:B:388:GLU:CA	1.79	1.63
1:I:444:PRO:CG	9:U:50:THR:HG23	1.32	1.58
1:I:444:PRO:HG3	9:U:50:THR:CG2	1.09	1.56
2:J:378:TYR:CE1	2:J:388:GLU:HG3	1.38	1.56
3:C:194:GLU:HB2	3:C:196:TYR:CE2	1.41	1.55

The worst 5 of 20 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)
6:F:98:LYS:NZ	5:Q:82:GLN:NE2[3_445]	1.00	1.20
2:B:61:PRO:C	6:F:31:GLU:OE1[1_655]	1.06	1.14
2:B:61:PRO:O	6:F:31:GLU:OE1[1_655]	1.11	1.09
2:B:61:PRO:CB	6:F:31:GLU:CD[1_655]	1.35	0.85
2:B:61:PRO:CA	6:F:31:GLU:OE1[1_655]	1.41	0.79

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	855/906 (94%)	689 (81%)	116 (14%)	50 (6%)	2 23
1	I	853/906 (94%)	692 (81%)	110 (13%)	51 (6%)	2 21
2	B	1061/1123 (94%)	872 (82%)	141 (13%)	48 (4%)	3 30
2	J	1061/1123 (94%)	864 (81%)	143 (14%)	54 (5%)	2 26
3	C	365/391 (93%)	258 (71%)	73 (20%)	34 (9%)	1 11
3	M	365/391 (93%)	266 (73%)	75 (20%)	24 (7%)	1 19
4	D	256/259 (99%)	241 (94%)	12 (5%)	3 (1%)	16 61
4	O	256/259 (99%)	242 (94%)	11 (4%)	3 (1%)	16 61
5	E	179/190 (94%)	161 (90%)	18 (10%)	0	100 100
5	Q	179/190 (94%)	161 (90%)	18 (10%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
6	F	120/122 (98%)	108 (90%)	8 (7%)	4 (3%)	5 39
6	R	120/122 (98%)	108 (90%)	8 (7%)	4 (3%)	5 39
7	H	74/82 (90%)	59 (80%)	10 (14%)	5 (7%)	1 19
7	S	74/82 (90%)	60 (81%)	10 (14%)	4 (5%)	2 25
8	K	54/57 (95%)	43 (80%)	7 (13%)	4 (7%)	1 16
8	T	54/57 (95%)	43 (80%)	7 (13%)	4 (7%)	1 16
9	L	92/100 (92%)	74 (80%)	12 (13%)	6 (6%)	1 20
9	U	92/100 (92%)	75 (82%)	10 (11%)	7 (8%)	1 15
10	N	61/65 (94%)	44 (72%)	11 (18%)	6 (10%)	1 10
10	V	61/65 (94%)	44 (72%)	11 (18%)	6 (10%)	1 10
11	P	40/49 (82%)	24 (60%)	10 (25%)	6 (15%)	0 3
11	W	40/49 (82%)	22 (55%)	11 (28%)	7 (18%)	0 2
All	All	6312/6688 (94%)	5150 (82%)	832 (13%)	330 (5%)	2 25

5 of 330 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	48	ASP
1	A	151	GLY
1	A	153	PRO
1	A	154	GLN
1	A	398	ASN

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	743/779 (95%)	639 (86%)	104 (14%)	4 24
1	I	740/779 (95%)	636 (86%)	104 (14%)	4 24
2	B	923/969 (95%)	788 (85%)	135 (15%)	4 22
2	J	923/969 (95%)	787 (85%)	136 (15%)	4 22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	C	314/334 (94%)	252 (80%)	62 (20%)	1 9
3	M	313/334 (94%)	250 (80%)	63 (20%)	1 9
4	D	227/228 (100%)	218 (96%)	9 (4%)	38 75
4	O	227/228 (100%)	218 (96%)	9 (4%)	38 75
5	E	160/167 (96%)	144 (90%)	16 (10%)	9 41
5	Q	160/167 (96%)	145 (91%)	15 (9%)	11 44
6	F	107/107 (100%)	99 (92%)	8 (8%)	17 55
6	R	107/107 (100%)	99 (92%)	8 (8%)	17 55
7	H	68/72 (94%)	51 (75%)	17 (25%)	1 4
7	S	68/72 (94%)	50 (74%)	18 (26%)	0 4
8	K	45/46 (98%)	38 (84%)	7 (16%)	3 20
8	T	45/46 (98%)	38 (84%)	7 (16%)	3 20
9	L	81/87 (93%)	77 (95%)	4 (5%)	31 70
9	U	81/87 (93%)	76 (94%)	5 (6%)	23 63
10	N	57/59 (97%)	50 (88%)	7 (12%)	6 29
10	V	57/59 (97%)	51 (90%)	6 (10%)	8 38
11	P	35/40 (88%)	30 (86%)	5 (14%)	4 24
11	W	35/40 (88%)	31 (89%)	4 (11%)	7 33
All	All	5516/5776 (96%)	4767 (86%)	749 (14%)	5 26

5 of 749 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	H	37	ILE
1	I	408	ILE
5	Q	93	MET
7	H	75	TYR
1	I	46	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
5	E	68	HIS
1	I	138	HIS
4	O	50	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
9	L	26	ASN
1	I	252	ASN

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 14 ligands modelled in this entry, 14 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.

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	863/906 (95%)	0.42	82 (9%) 10 10	41, 157, 264, 370	0
1	I	862/906 (95%)	0.44	83 (9%) 10 9	54, 163, 263, 381	0
2	B	1069/1123 (95%)	0.27	62 (5%) 26 21	46, 149, 253, 402	0
2	J	1069/1123 (95%)	0.12	40 (3%) 45 36	39, 139, 248, 360	0
3	C	369/391 (94%)	0.92	74 (20%) 1 2	68, 192, 319, 372	0
3	M	369/391 (94%)	1.02	73 (19%) 1 2	71, 193, 327, 419	0
4	D	258/259 (99%)	0.49	26 (10%) 9 9	71, 143, 235, 295	0
4	O	258/259 (99%)	0.14	12 (4%) 35 28	63, 136, 226, 319	0
5	E	181/190 (95%)	0.66	24 (13%) 4 5	69, 159, 250, 398	0
5	Q	181/190 (95%)	0.61	22 (12%) 5 6	90, 163, 244, 317	0
6	F	122/122 (100%)	0.58	15 (12%) 5 6	133, 165, 244, 447	0
6	R	122/122 (100%)	0.31	6 (4%) 33 25	142, 184, 257, 324	0
7	H	76/82 (92%)	0.20	3 (3%) 43 35	105, 157, 228, 268	0
7	S	76/82 (92%)	0.51	4 (5%) 30 23	114, 166, 248, 293	0
8	K	56/57 (98%)	0.08	0 100 100	88, 131, 196, 288	0
8	T	56/57 (98%)	0.16	3 (5%) 29 23	62, 119, 227, 330	0
9	L	94/100 (94%)	0.34	3 (3%) 51 42	78, 142, 221, 331	0
9	U	94/100 (94%)	0.54	12 (12%) 5 5	83, 149, 223, 474	0
10	N	63/65 (96%)	-0.03	1 (1%) 74 65	81, 141, 218, 281	0
10	V	63/65 (96%)	0.08	2 (3%) 51 42	69, 131, 196, 232	0
11	P	42/49 (85%)	0.50	5 (11%) 6 6	125, 173, 254, 294	0
11	W	42/49 (85%)	0.34	3 (7%) 19 15	90, 147, 201, 253	0
All	All	6385/6688 (95%)	0.40	555 (8%) 13 12	39, 156, 270, 474	0

The worst 5 of 555 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1089	ASP	14.0
3	M	200	VAL	12.6
3	M	340	GLU	11.1
5	E	92	MET	9.7
3	M	209	SER	9.7

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, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
13	ZN	J	1301	1/1	0.89	0.49	3.87	521,521,521,521	0
12	MG	I	1101	1/1	0.72	0.39	2.22	175,175,175,175	0
13	ZN	B	1201	1/1	0.93	0.28	1.00	364,364,364,364	0
13	ZN	V	1601	1/1	0.99	0.24	0.30	210,210,210,210	0
13	ZN	P	1501	1/1	0.72	0.26	-0.21	354,354,354,354	0
13	ZN	A	1003	1/1	0.59	0.18	-0.93	488,488,488,488	0
13	ZN	W	1701	1/1	0.97	0.08	-1.00	94,94,94,94	0
13	ZN	A	1002	1/1	0.88	0.20	-1.00	137,137,137,137	0
13	ZN	I	1102	1/1	0.89	0.10	-1.08	147,147,147,147	0
13	ZN	I	1103	1/1	0.78	0.17	-1.44	527,527,527,527	0
13	ZN	N	1401	1/1	0.74	0.20	-1.48	127,127,127,127	0
12	MG	I	1104	1/1	0.87	0.82	-	54,54,54,54	0
12	MG	A	1001	1/1	0.89	0.35	-	52,52,52,52	0
12	MG	A	1004	1/1	0.89	0.62	-	58,58,58,58	0

6.5 Other polymers (i)

There are no such residues in this entry.