



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 07:49 PM GMT

PDB ID : 2CW0
Title : Crystal structure of Thermus thermophilus RNA polymerase holoenzyme at 3.3 angstroms resolution
Authors : Tuske, S.; Sarafianos, S.G.; Wang, X.; Hudson, B.; Sineva, E.; Mukhopadhyay, J.; Birktoft, J.J.; Leroy, O.; Ismail, S.; Clark Jr., A.D.; Dharia, C.; Napoli, A.; Laptenko, O.; Lee, J.; Borukhov, S.; Ebright, R.H.; Arnold, E.
Deposited on : 2005-06-15
Resolution : 3.30 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

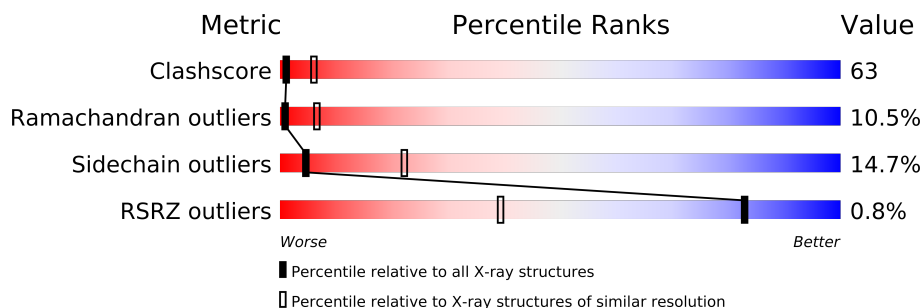
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.30 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))
Clashscore	79885	1696 (3.40-3.20)
Ramachandran outliers	78287	1664 (3.40-3.20)
Sidechain outliers	78261	1662 (3.40-3.20)
RSRZ outliers	66119	1342 (3.40-3.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	315	
1	B	315	
1	K	315	
1	L	315	
2	C	1119	
2	M	1119	
3	D	1524	
3	N	1524	
4	E	99	
4	O	99	
5	F	423	
5	P	423	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	ZN	D	1525	-	X
6	ZN	N	1525	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 53962 atoms, of which 0 are hydrogen and 0 are deuterium.

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 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	B	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	K	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	L	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			
2	M	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			

- Molecule 3 is a protein called DNA-directed RNA polymerase beta' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1392	Total	C	N	O	S	0	0	0
			10975	6953	1941	2048	33			
3	N	1392	Total	C	N	O	S	0	0	0
			10975	6953	1941	2048	33			

- Molecule 4 is a protein called RNA polymerase omega chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			
4	O	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			

- Molecule 5 is a protein called RNA polymerase sigma factor rpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	345	Total	C	N	O	S	0	0	0
			2793	1762	504	523	4			
5	P	345	Total	C	N	O	S	0	0	0
			2793	1762	504	523	4			

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	2	Total	Zn	0	0
			2	2		
6	N	2	Total	Zn	0	0
			2	2		

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

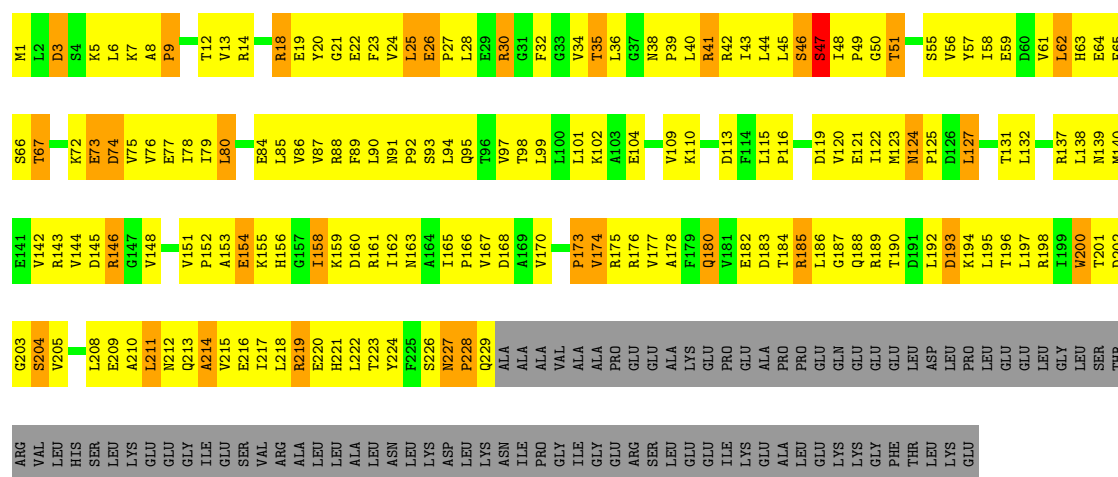
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	1	Total	Mg	0	0
			1	1		
7	N	1	Total	Mg	0	0
			1	1		

3 Residue-property plots

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.

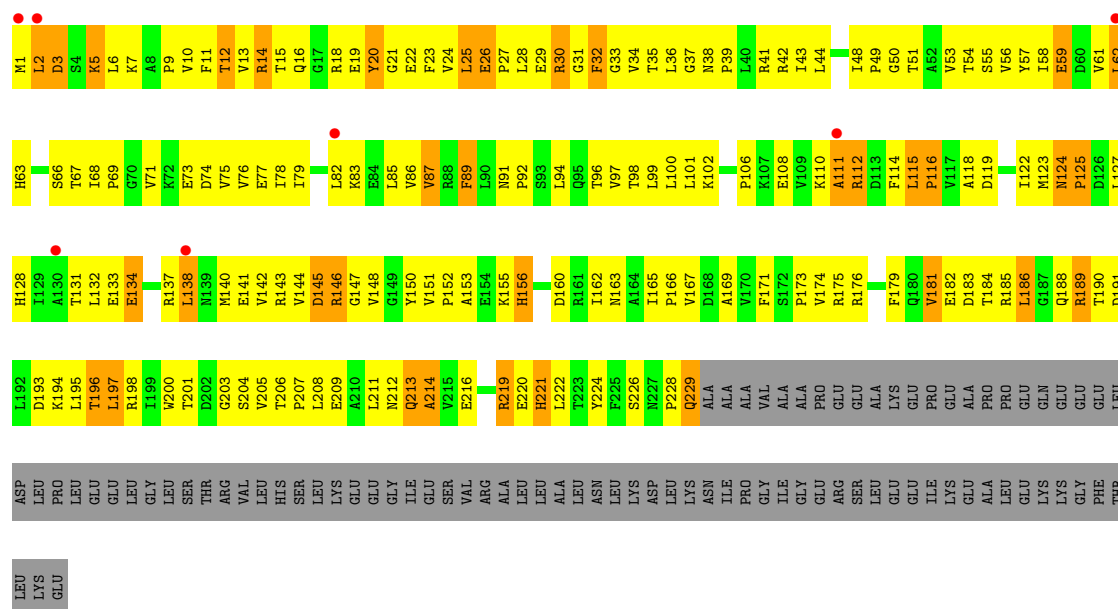
- Molecule 1: DNA-directed RNA polymerase alpha chain

Chain A:



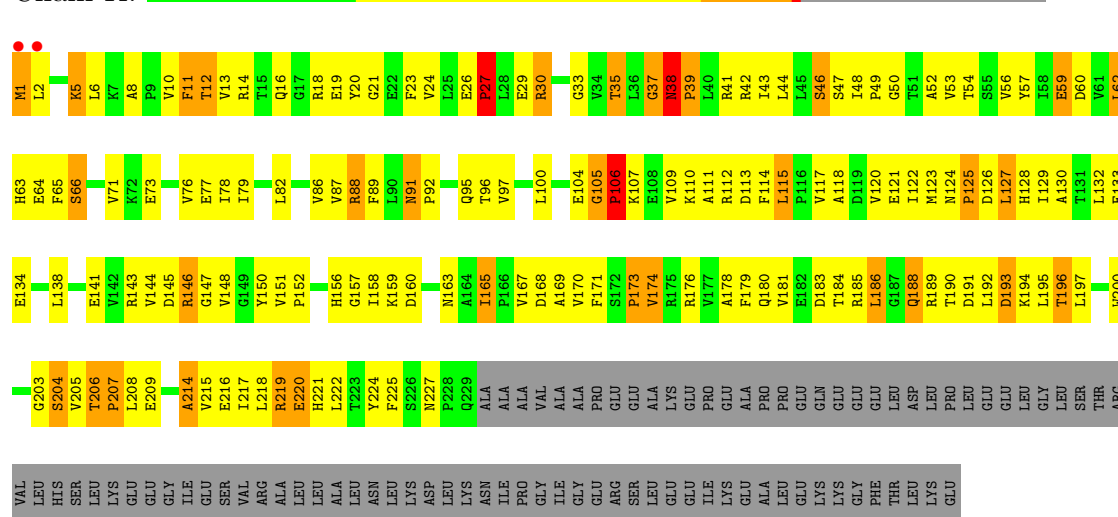
- Molecule 1: DNA-directed RNA polymerase alpha chain

Chain B:



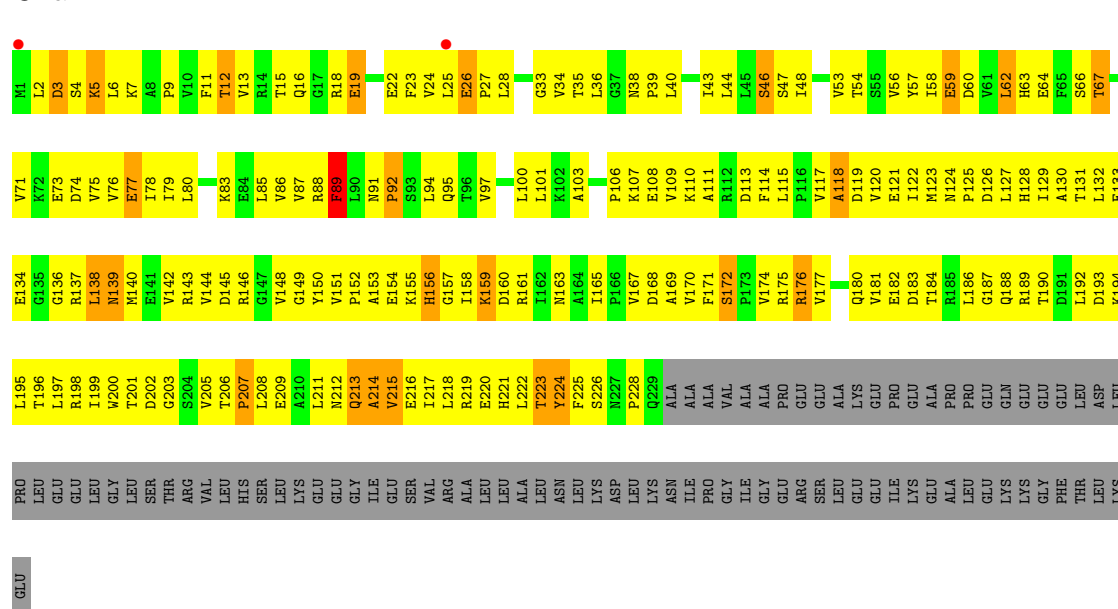
- Molecule 1: DNA-directed RNA polymerase alpha chain

Chain K:



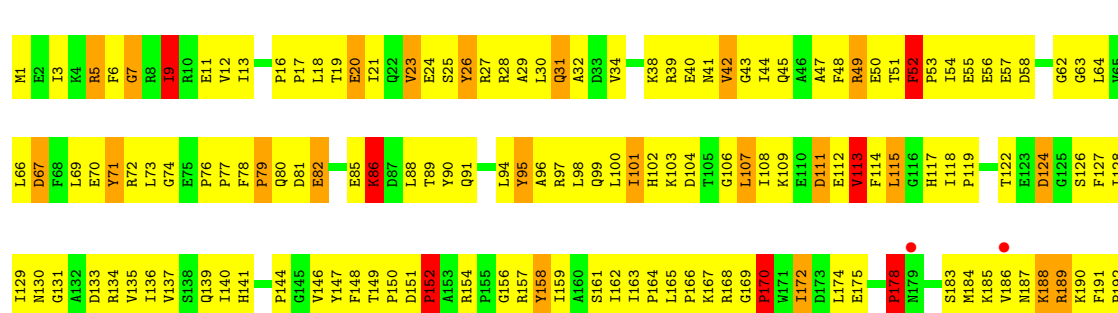
- Molecule 1: DNA-directed RNA polymerase alpha chain

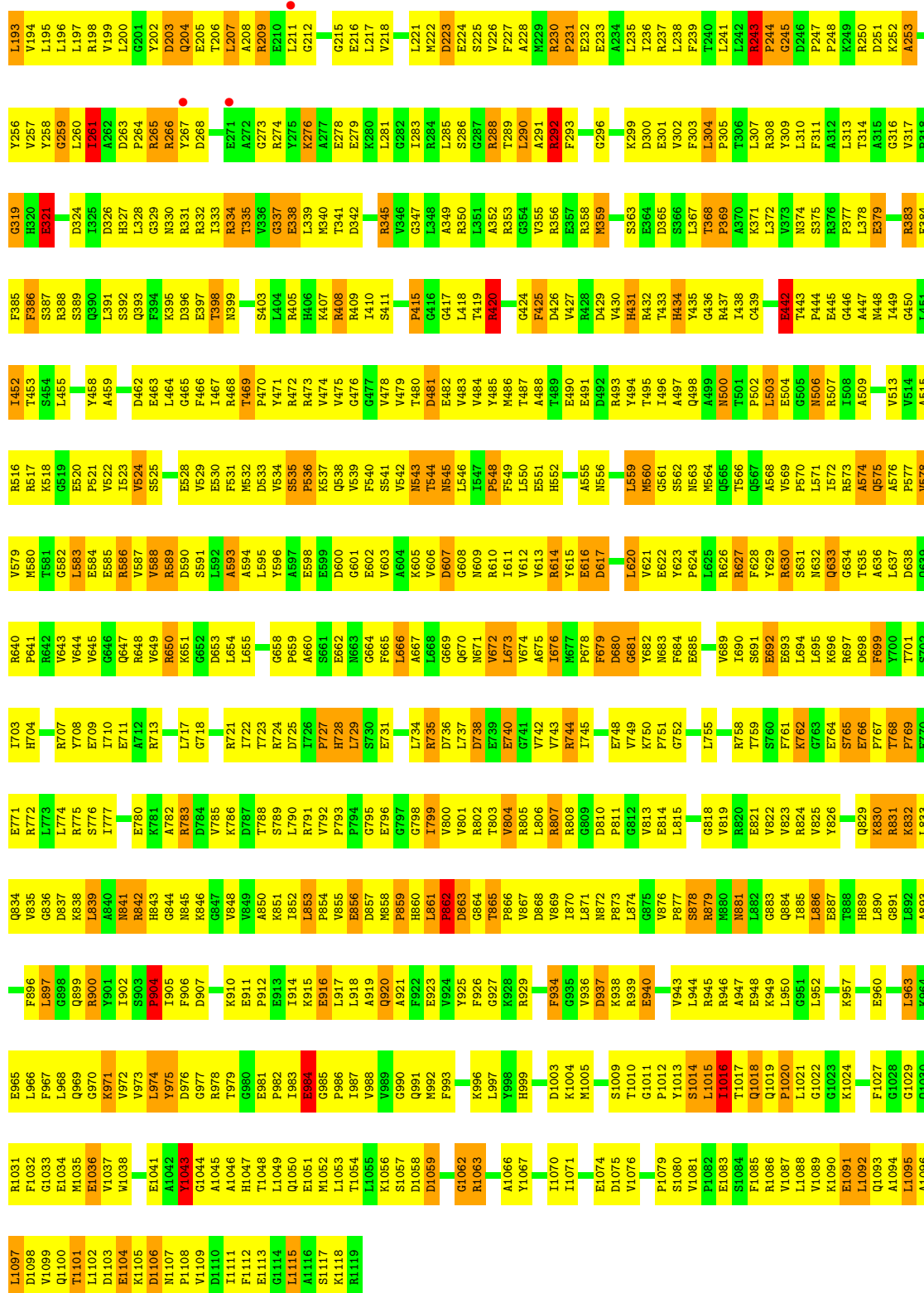
Chain L:



- Molecule 2: DNA-directed RNA polymerase beta chain

Chain C:

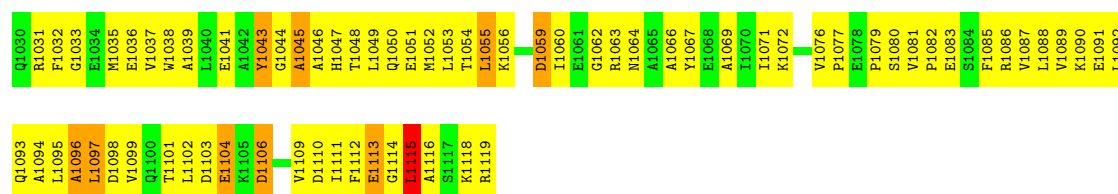




● Molecule 2: DNA-directed RNA polymerase beta chain

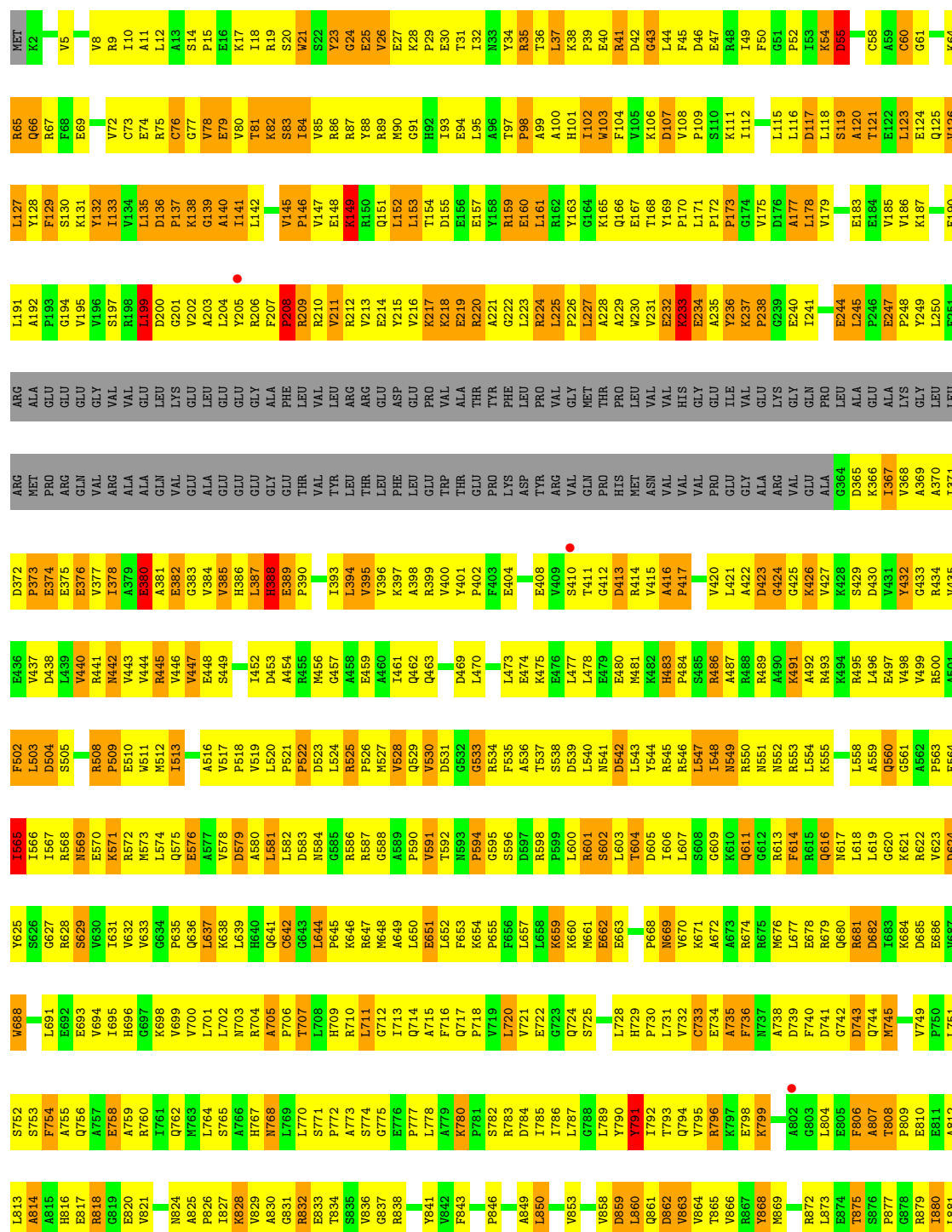
Chain M:





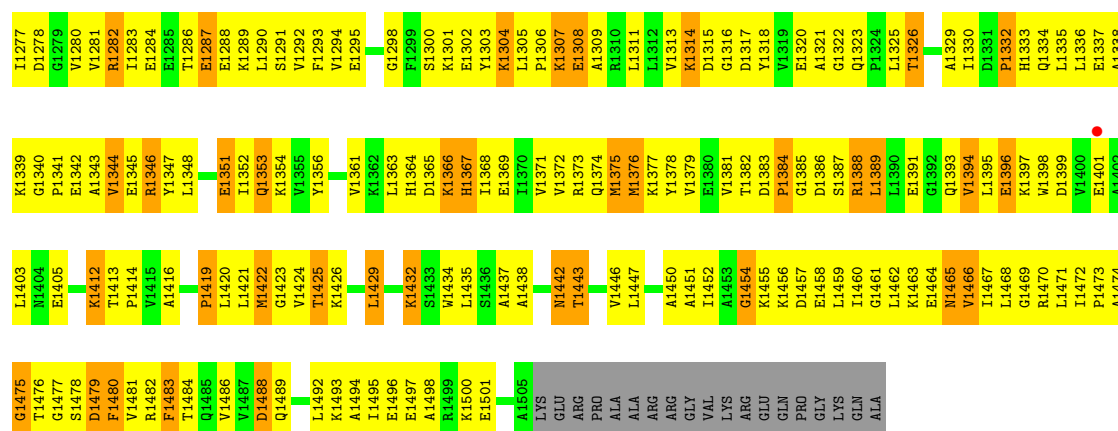
• Molecule 3: DNA-directed RNA polymerase beta' chain

Chain D:



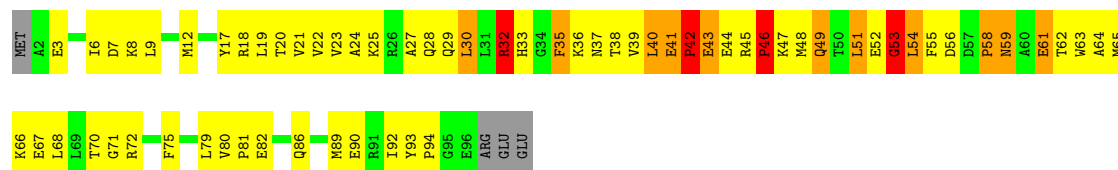






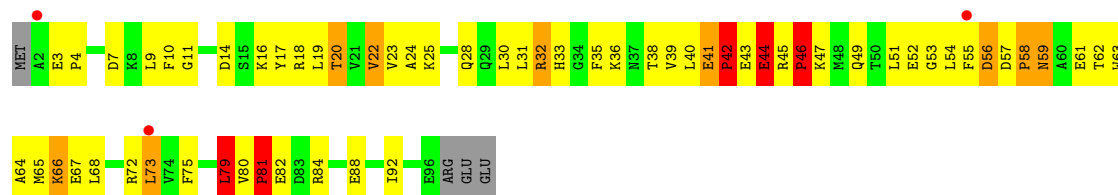
• Molecule 4: RNA polymerase omega chain

Chain E:



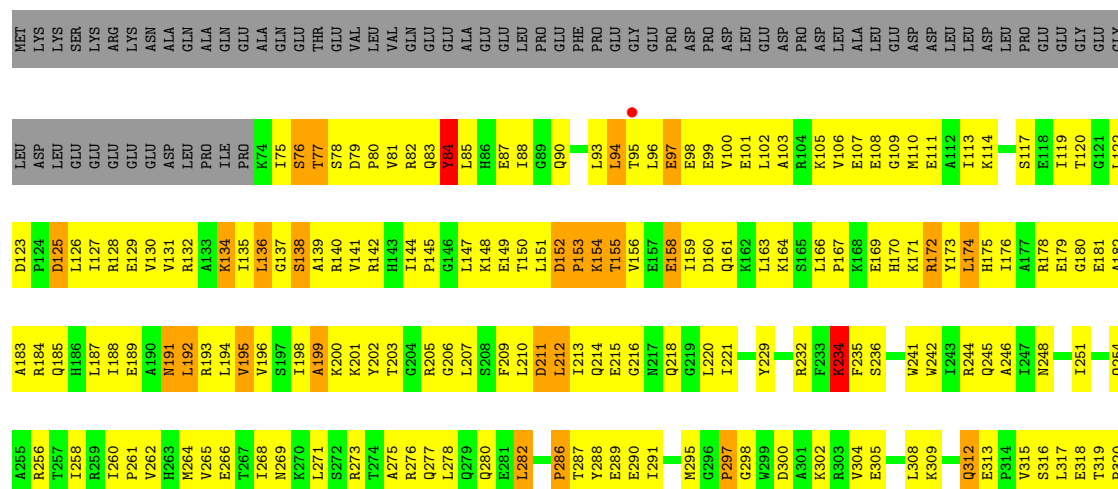
• Molecule 4: RNA polymerase omega chain

Chain O:



• Molecule 5: RNA polymerase sigma factor rpoD

Chain F:



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	236.15Å 236.15Å 249.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 3.30 29.87 – 3.30	Depositor EDS
% Data completeness (in resolution range)	84.1 (30.00-3.30) 47.2 (29.87-3.30)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.71 (at 3.31Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.282 , 0.320 0.282 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	92.6	Xtriage
Anisotropy	0.321	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , -20.0	EDS
Estimated twinning fraction	0.499 for -h,-k,l 0.499 for h,-h-k,-l 0.086 for -k,-h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 115957 reflections	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	53962	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.25% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.46	0/1838	0.75	0/2498
1	B	0.36	0/1838	0.64	0/2498
1	K	0.46	0/1838	0.75	0/2498
1	L	0.39	0/1838	0.68	0/2498
2	C	0.45	0/8997	0.79	8/12164 (0.1%)
2	M	0.46	0/8997	0.79	8/12164 (0.1%)
3	D	0.48	0/11165	0.83	16/15088 (0.1%)
3	N	0.46	0/11165	0.81	15/15088 (0.1%)
4	E	0.42	0/783	0.80	3/1054 (0.3%)
4	O	0.42	0/783	0.80	1/1054 (0.1%)
5	F	0.40	0/2836	0.73	0/3812
5	P	0.41	0/2836	0.72	0/3812
All	All	0.45	0/54914	0.78	51/74228 (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
3	D	0	1
3	N	0	3
5	F	0	1
All	All	0	5

There are no bond length outliers.

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	1209	LEU	N-CA-C	-10.12	83.67	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	728	HIS	N-CA-C	7.66	131.69	111.00
2	C	728	HIS	N-CA-C	7.62	131.58	111.00
3	N	1209	LEU	N-CA-C	-7.26	91.39	111.00
2	M	319	GLY	N-CA-C	-7.22	95.05	113.10

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	132	TYR	Sidechain
5	F	84	TYR	Sidechain
3	N	1015	TYR	Sidechain
3	N	1318	TYR	Sidechain
3	N	132	TYR	Sidechain

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1806	0	1861	249	0
1	B	1806	0	1861	193	0
1	K	1806	0	1861	190	0
1	L	1806	0	1861	208	0
2	C	8829	0	8933	1143	0
2	M	8829	0	8933	1183	0
3	D	10975	0	11211	1723	0
3	N	10975	0	11210	1681	0
4	E	769	0	775	94	0
4	O	769	0	775	83	0
5	F	2793	0	2873	301	0
5	P	2793	0	2873	364	0
6	D	2	0	0	0	0
6	N	2	0	0	0	0
7	D	1	0	0	0	0
7	N	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	53962	0	55027	6830	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 63.

The worst 5 of 6830 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:D:1101:VAL:HG21	3:D:1424:VAL:HG22	1.21	1.20
2:C:1016:ILE:HD13	2:C:1016:ILE:H	1.06	1.16
3:D:907:GLU:HG2	3:D:1027:GLY:H	1.02	1.16
3:D:136:ASP:HB3	3:D:137:PRO:HD3	1.22	1.15
3:D:145:VAL:HG22	3:D:146:PRO:HD2	1.27	1.15

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	227/315 (72%)	167 (74%)	39 (17%)	21 (9%)	1	9
1	B	227/315 (72%)	177 (78%)	37 (16%)	13 (6%)	3	24
1	K	227/315 (72%)	161 (71%)	39 (17%)	27 (12%)	1	4
1	L	227/315 (72%)	166 (73%)	44 (19%)	17 (8%)	2	15
2	C	1117/1119 (100%)	781 (70%)	226 (20%)	110 (10%)	1	8
2	M	1117/1119 (100%)	769 (69%)	215 (19%)	133 (12%)	1	4
3	D	1388/1524 (91%)	941 (68%)	293 (21%)	154 (11%)	1	5
3	N	1388/1524 (91%)	907 (65%)	332 (24%)	149 (11%)	1	6
4	E	93/99 (94%)	67 (72%)	17 (18%)	9 (10%)	1	8
4	O	93/99 (94%)	59 (63%)	20 (22%)	14 (15%)	0	1
5	F	341/423 (81%)	241 (71%)	67 (20%)	33 (10%)	1	8
5	P	341/423 (81%)	249 (73%)	57 (17%)	35 (10%)	1	7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	6786/7590 (89%)	4685 (69%)	1386 (20%)	715 (10%)	1 6

5 of 715 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	47	SER
1	B	3	ASP
1	B	118	ALA
1	B	160	ASP
2	C	7	GLY

5.3.2 Protein sidechains ⓘ

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	202/273 (74%)	177 (88%)	25 (12%)	7 31
1	B	202/273 (74%)	172 (85%)	30 (15%)	4 21
1	K	202/273 (74%)	173 (86%)	29 (14%)	5 23
1	L	202/273 (74%)	182 (90%)	20 (10%)	11 44
2	C	941/941 (100%)	808 (86%)	133 (14%)	5 24
2	M	941/941 (100%)	805 (86%)	136 (14%)	5 23
3	D	1170/1279 (92%)	970 (83%)	200 (17%)	3 15
3	N	1170/1279 (92%)	980 (84%)	190 (16%)	3 17
4	E	83/87 (95%)	72 (87%)	11 (13%)	6 27
4	O	83/87 (95%)	70 (84%)	13 (16%)	4 19
5	F	300/370 (81%)	264 (88%)	36 (12%)	7 33
5	P	300/370 (81%)	269 (90%)	31 (10%)	10 41
All	All	5796/6446 (90%)	4942 (85%)	854 (15%)	4 22

5 of 854 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	1462	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	156	HIS
3	N	1326	THR
4	E	66	LYS
5	F	416	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 168 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1442	ASN
1	L	38	ASN
3	N	1441	GLN
3	D	1489	GLN
1	K	95	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 6 ligands modelled in this entry, 6 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.

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, 95th 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	229/315 (72%)	-0.14	0 100 100	41, 64, 87, 113	0
1	B	229/315 (72%)	0.03	7 (3%) 47 11	47, 121, 143, 143	0
1	K	229/315 (72%)	-0.09	2 (0%) 81 37	28, 63, 88, 111	0
1	L	229/315 (72%)	-0.07	2 (0%) 81 37	41, 79, 99, 117	0
2	C	1119/1119 (100%)	-0.10	5 (0%) 90 57	12, 67, 133, 143	0
2	M	1119/1119 (100%)	-0.09	7 (0%) 86 46	6, 71, 122, 133	0
3	D	1392/1524 (91%)	-0.06	10 (0%) 84 42	7, 60, 125, 143	0
3	N	1392/1524 (91%)	-0.03	14 (1%) 79 33	5, 65, 134, 143	0
4	E	95/99 (95%)	-0.08	0 100 100	54, 84, 100, 106	0
4	O	95/99 (95%)	-0.03	3 (3%) 45 11	46, 86, 117, 121	0
5	F	345/423 (81%)	-0.09	2 (0%) 86 46	48, 81, 113, 121	0
5	P	345/423 (81%)	-0.09	2 (0%) 86 46	47, 73, 110, 123	0
All	All	6818/7590 (89%)	-0.07	54 (0%) 83 39	5, 70, 129, 143	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	1247	ALA	8.5
2	C	211	LEU	6.4
3	D	1246	VAL	4.9
1	B	1	MET	4.8
3	N	1246	VAL	4.7

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	ZN	N	1525	1/1	0.31	3.84	108,108,108,108	0
6	ZN	D	1525	1/1	0.28	3.59	107,107,107,107	0
6	ZN	N	1526	1/1	0.27	1.76	72,72,72,72	0
6	ZN	D	1526	1/1	0.26	1.68	78,78,78,78	0
7	MG	N	1527	1/1	0.15	0.42	29,29,29,29	0
7	MG	D	1527	1/1	0.09	-7.21	19,19,19,19	0

6.5 Other polymers

There are no such residues in this entry.