



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:39 PM GMT

PDB ID : 1L9U  
Title : THERMUS AQUATICUS RNA POLYMERASE HOLOENZYME AT 4 Å  
RESOLUTION  
Authors : Murakami, K.S.; Masuda, S.; Darst, S.A.  
Deposited on : 2002-03-26  
Resolution : 4.00 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

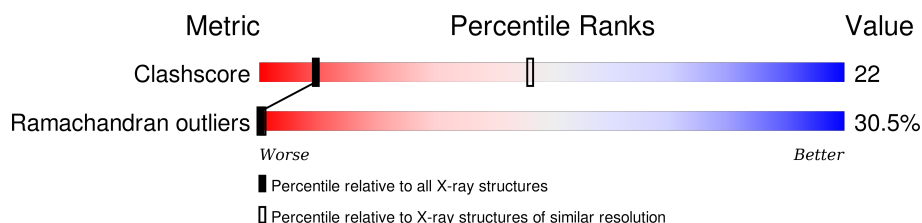
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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	102246	1052 (4.40-3.60)
Ramachandran outliers	100387	1005 (4.40-3.60)

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  $\geq 3$ , 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  $\leq 5\%$ . 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	314	
1	B	314	
1	J	314	
1	K	314	
2	C	1118	
2	L	1118	
3	D	1524	
3	M	1524	
4	E	99	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	N	99	<div><div></div><div>47%</div><div>44%</div><div>• 7%</div></div>
5	H	332	<div><div></div><div>75%</div><div>21%</div><div>• •</div></div>
5	Q	332	<div><div></div><div>75%</div><div>21%</div><div>• •</div></div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 18756 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 RNA POLYMERASE, ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
1	A	224	Total	C	N	0	0	0
			672	448	224			
1	B	220	Total	C	N	0	0	0
			660	440	220			
1	J	224	Total	C	N	0	0	0
			672	448	224			
1	K	220	Total	C	N	0	0	0
			660	440	220			

- Molecule 2 is a protein called RNA POLYMERASE, BETA SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	1084	Total	C	N	0	0	0
			3252	2168	1084			
2	L	1084	Total	C	N	0	0	0
			3252	2168	1084			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	GLU	DELETION	UNP Q9KWU7
L	?	-	GLU	DELETION	UNP Q9KWU7

- Molecule 3 is a protein called RNA POLYMERASE, BETA-PRIME SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	D	1183	Total	C	N	0	0	0
			3549	2366	1183			
3	M	1183	Total	C	N	0	0	0
			3549	2366	1183			

- Molecule 4 is a protein called RNA POLYMERASE, OMEGA SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	E	92	Total	C	N	0	0	0
			276	184	92			
4	N	92	Total	C	N	0	0	0
			276	184	92			

- Molecule 5 is a protein called SIGMA FACTOR SIGA.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
5	H	322	Total	C	N	0	0	0
			966	644	322			
5	Q	322	Total	C	N	0	0	0
			966	644	322			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	?	-	ILE	DELETION	UNP Q9EZJ8
H	?	-	GLN	DELETION	UNP Q9EZJ8
H	?	-	LYS	DELETION	UNP Q9EZJ8
H	?	-	ILE	DELETION	UNP Q9EZJ8
H	?	-	PRO	DELETION	UNP Q9EZJ8
H	?	-	GLY	DELETION	UNP Q9EZJ8
H	?	-	LEU	DELETION	UNP Q9EZJ8
H	?	-	LYS	DELETION	UNP Q9EZJ8
H	?	-	GLU	DELETION	UNP Q9EZJ8
H	?	-	LYS	DELETION	UNP Q9EZJ8
H	?	-	PRO	DELETION	UNP Q9EZJ8
H	?	-	ASP	DELETION	UNP Q9EZJ8
H	?	-	PRO	DELETION	UNP Q9EZJ8
H	?	-	LYS	DELETION	UNP Q9EZJ8
H	?	-	THR	DELETION	UNP Q9EZJ8
Q	?	-	ILE	DELETION	UNP Q9EZJ8
Q	?	-	GLN	DELETION	UNP Q9EZJ8
Q	?	-	LYS	DELETION	UNP Q9EZJ8
Q	?	-	ILE	DELETION	UNP Q9EZJ8
Q	?	-	PRO	DELETION	UNP Q9EZJ8
Q	?	-	GLY	DELETION	UNP Q9EZJ8
Q	?	-	LEU	DELETION	UNP Q9EZJ8
Q	?	-	LYS	DELETION	UNP Q9EZJ8
Q	?	-	GLU	DELETION	UNP Q9EZJ8
Q	?	-	LYS	DELETION	UNP Q9EZJ8
Q	?	-	PRO	DELETION	UNP Q9EZJ8
Q	?	-	ASP	DELETION	UNP Q9EZJ8

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
Q	?	-	PRO	DELETION	UNP Q9EZJ8
Q	?	-	LYS	DELETION	UNP Q9EZJ8
Q	?	-	THR	DELETION	UNP Q9EZJ8

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	D	1	Total Mg 1 1	0	0
6	M	1	Total Mg 1 1	0	0

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

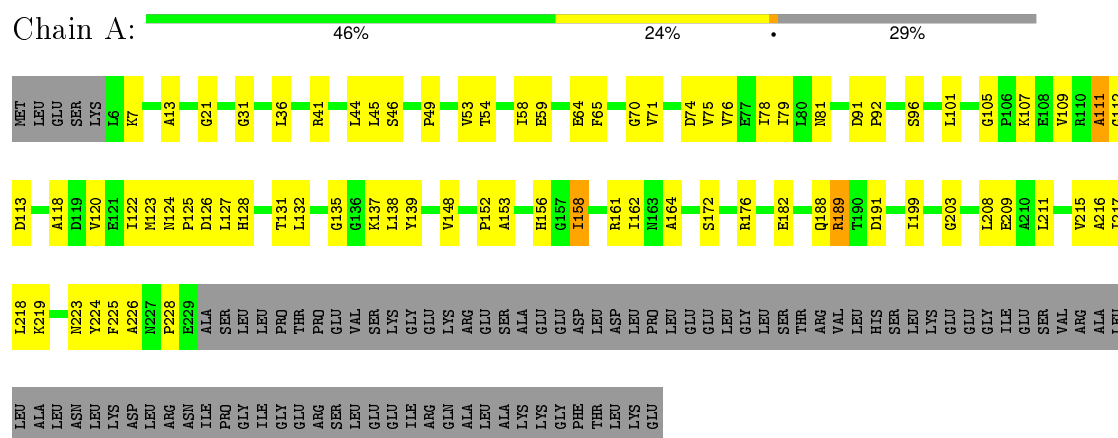
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	2	Total Zn 2 2	0	0
7	M	2	Total Zn 2 2	0	0

### 3 Residue-property plots

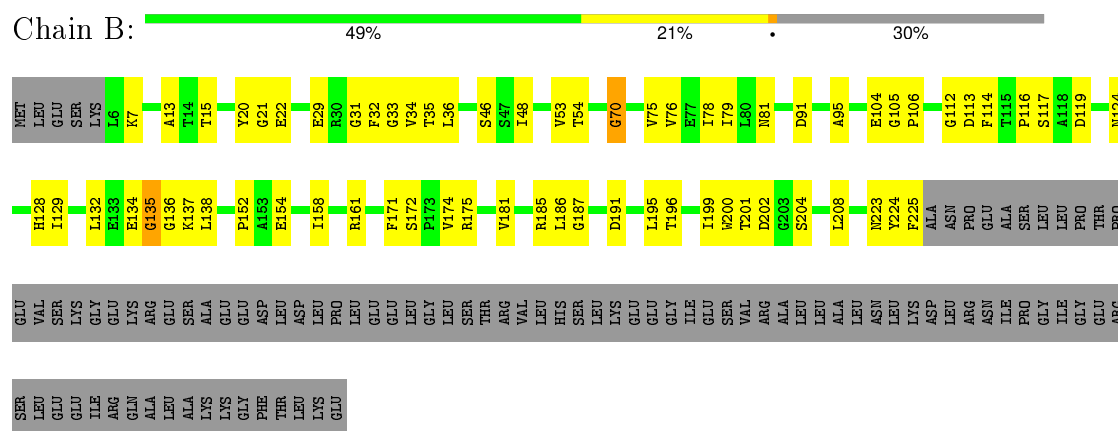
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

Note EDS was not executed.

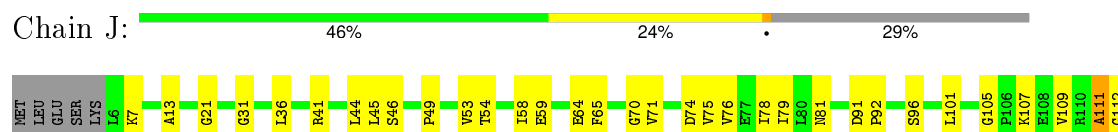
#### • Molecule 1: RNA POLYMERASE, ALPHA SUBUNIT

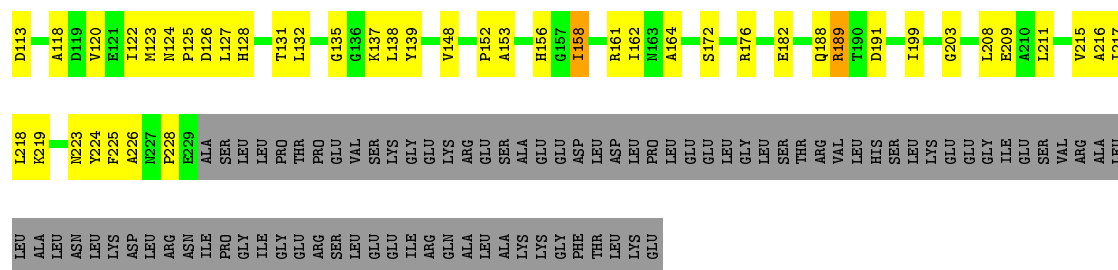


#### • Molecule 1: RNA POLYMERASE, ALPHA SUBUNIT

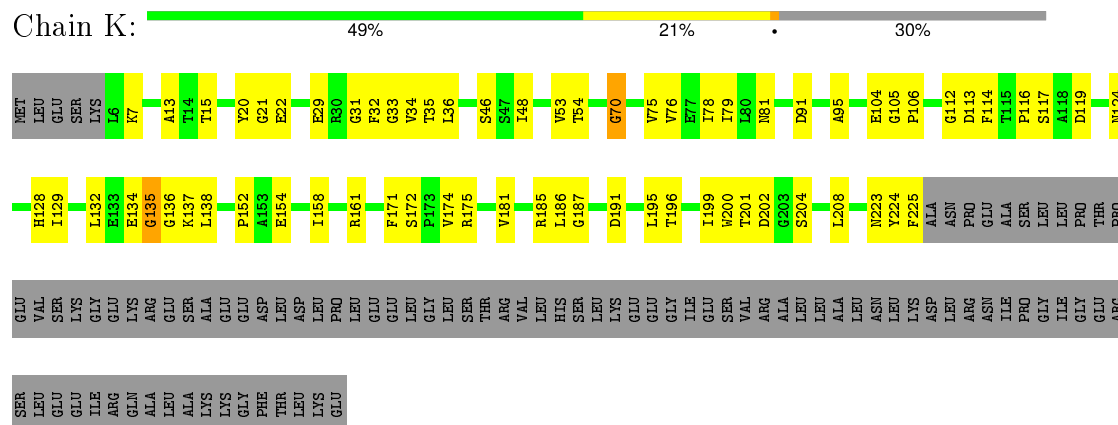


#### • Molecule 1: RNA POLYMERASE, ALPHA SUBUNIT

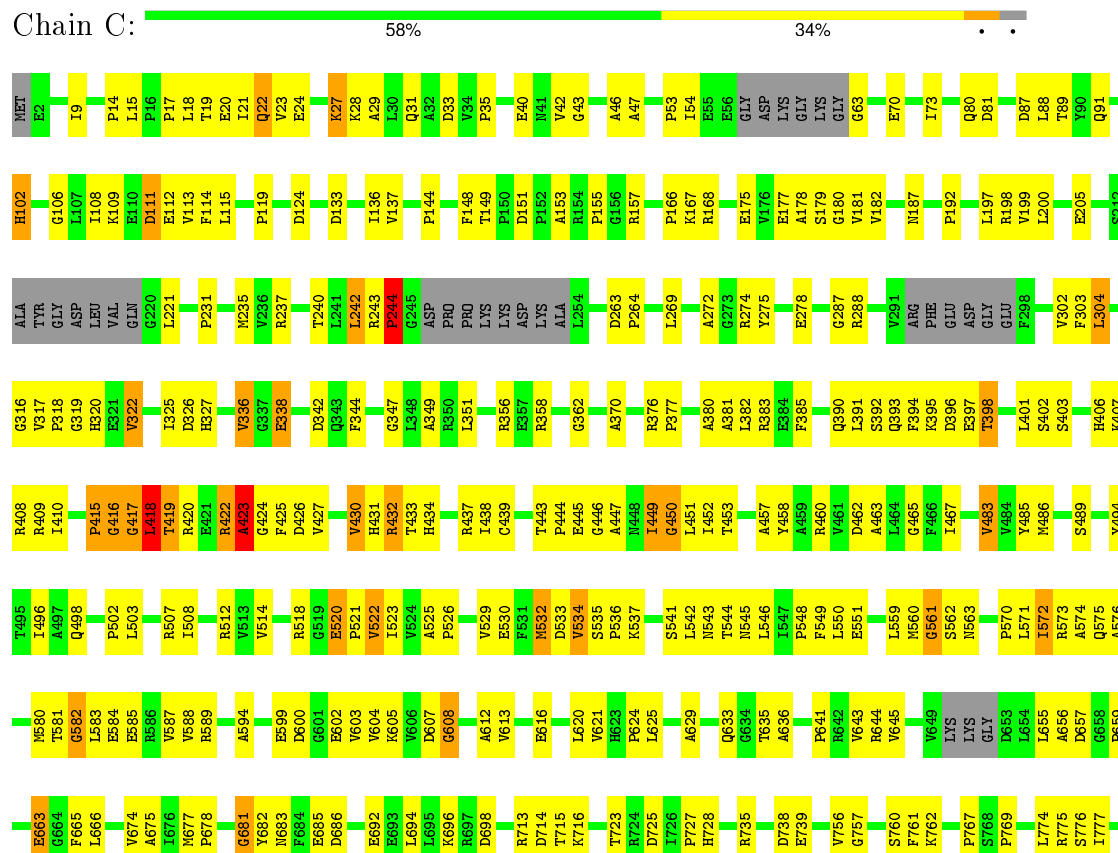




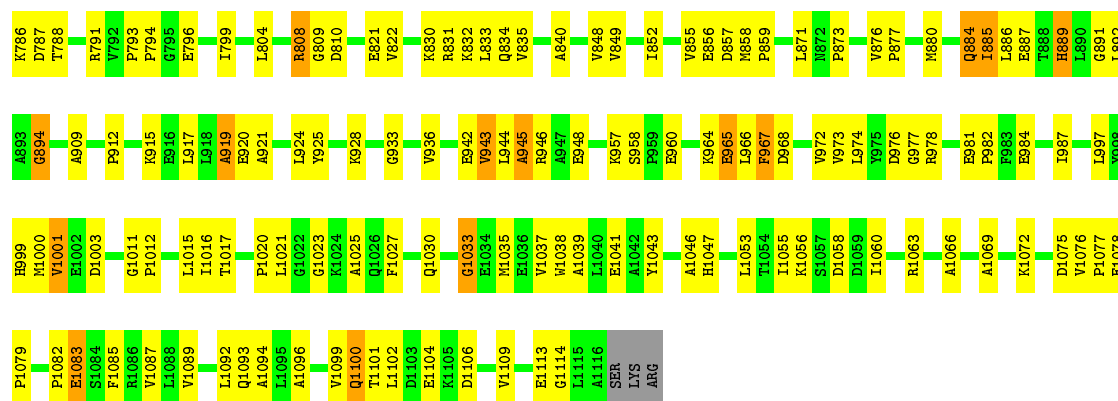
### • Molecule 1: RNA POLYMERASE, ALPHA SUBUNIT



### • Molecule 2: RNA POLYMERASE, BETA SUBUNIT

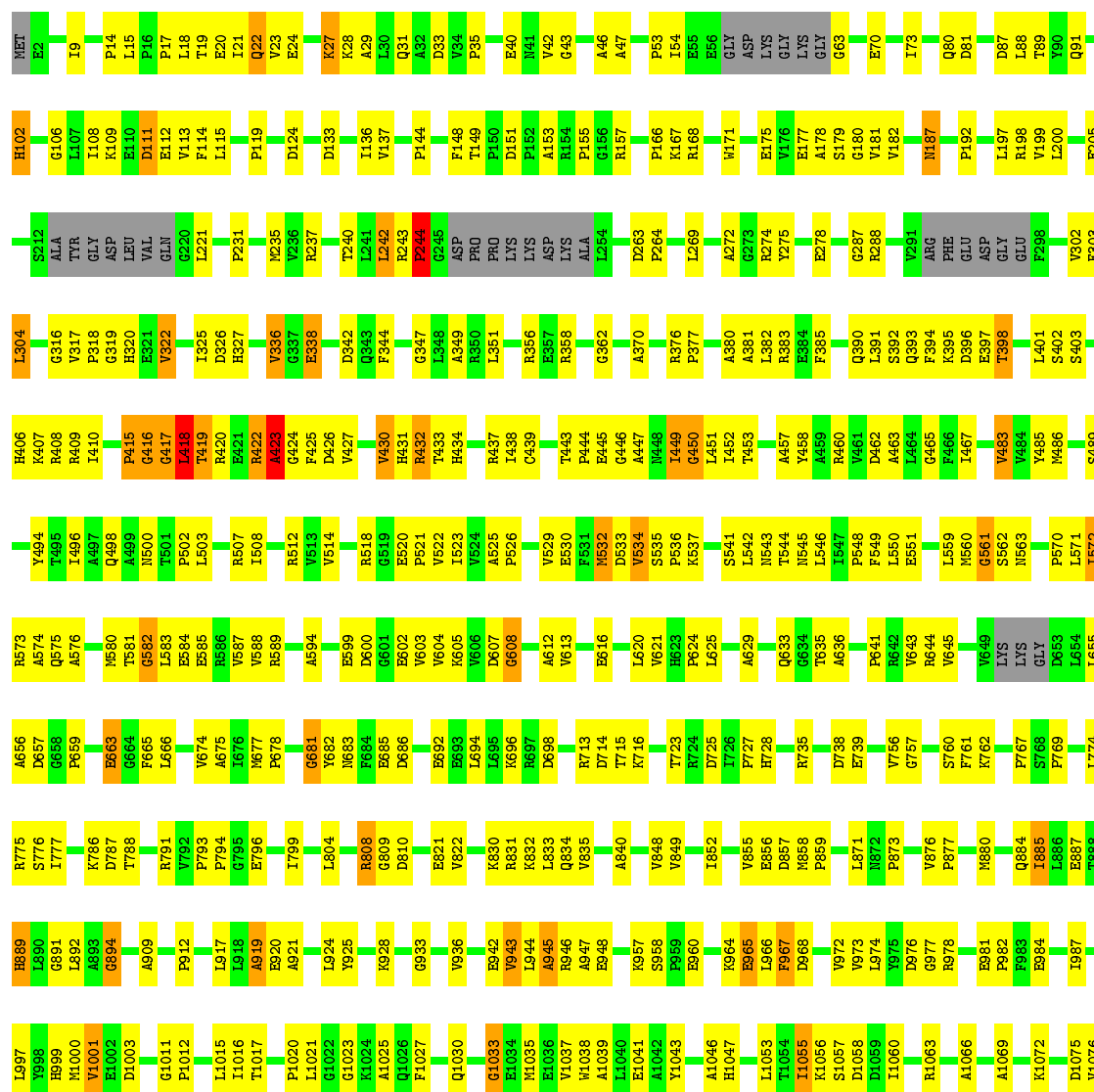


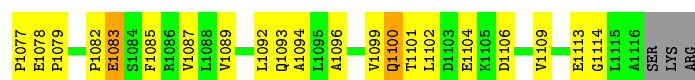




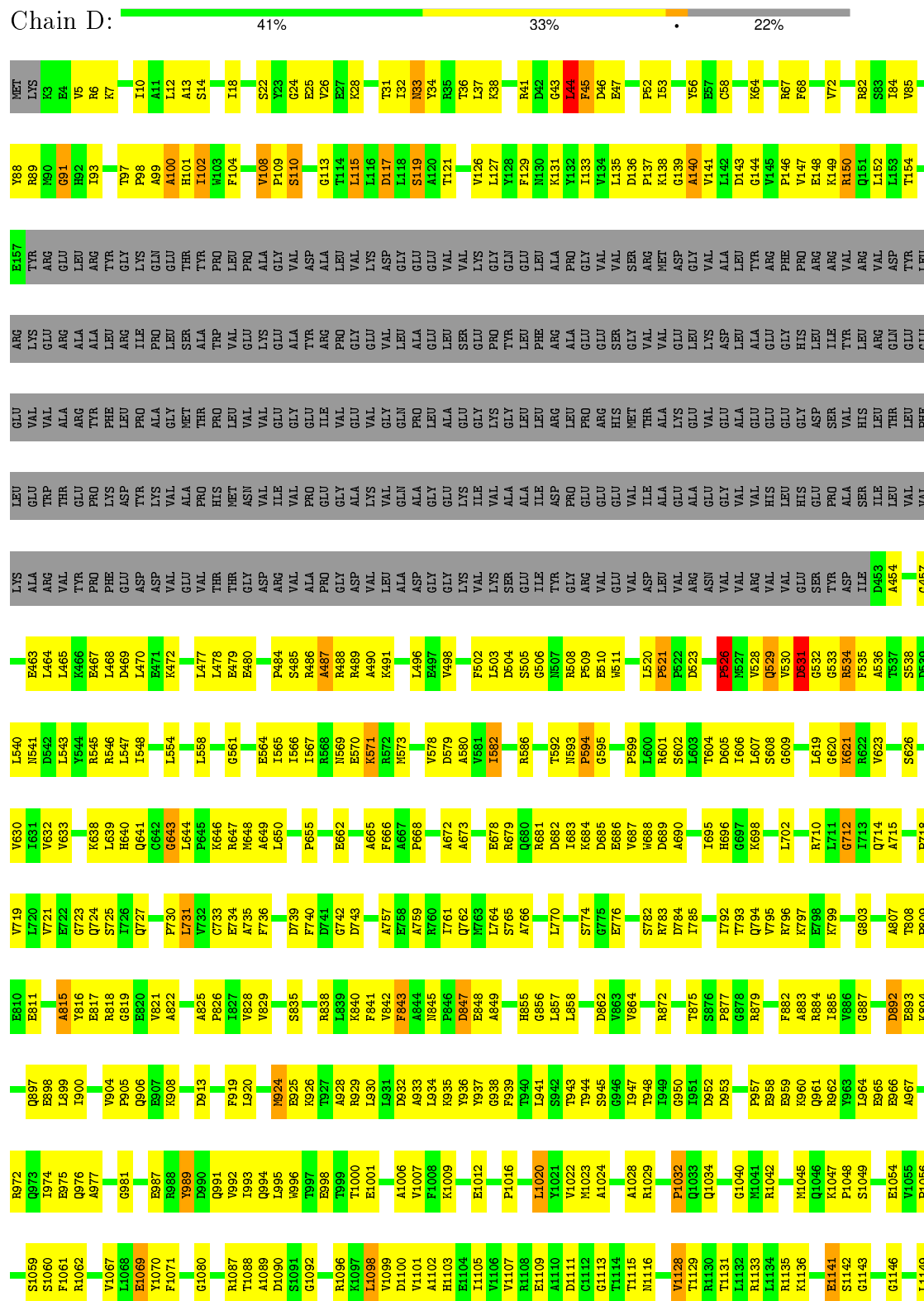
• Molecule 2: RNA POLYMERASE, BETA SUBUNIT

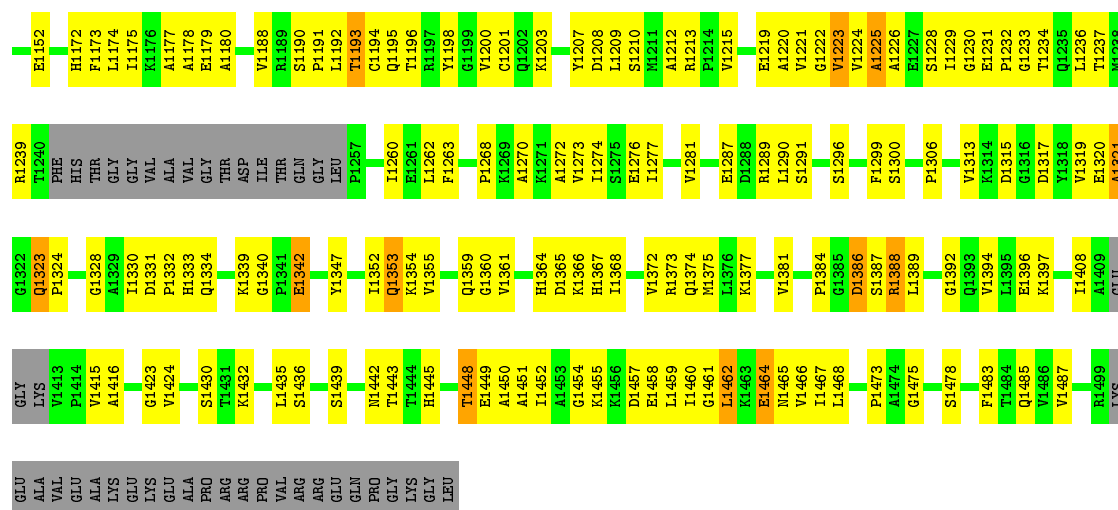
Chain L: 58% 35%



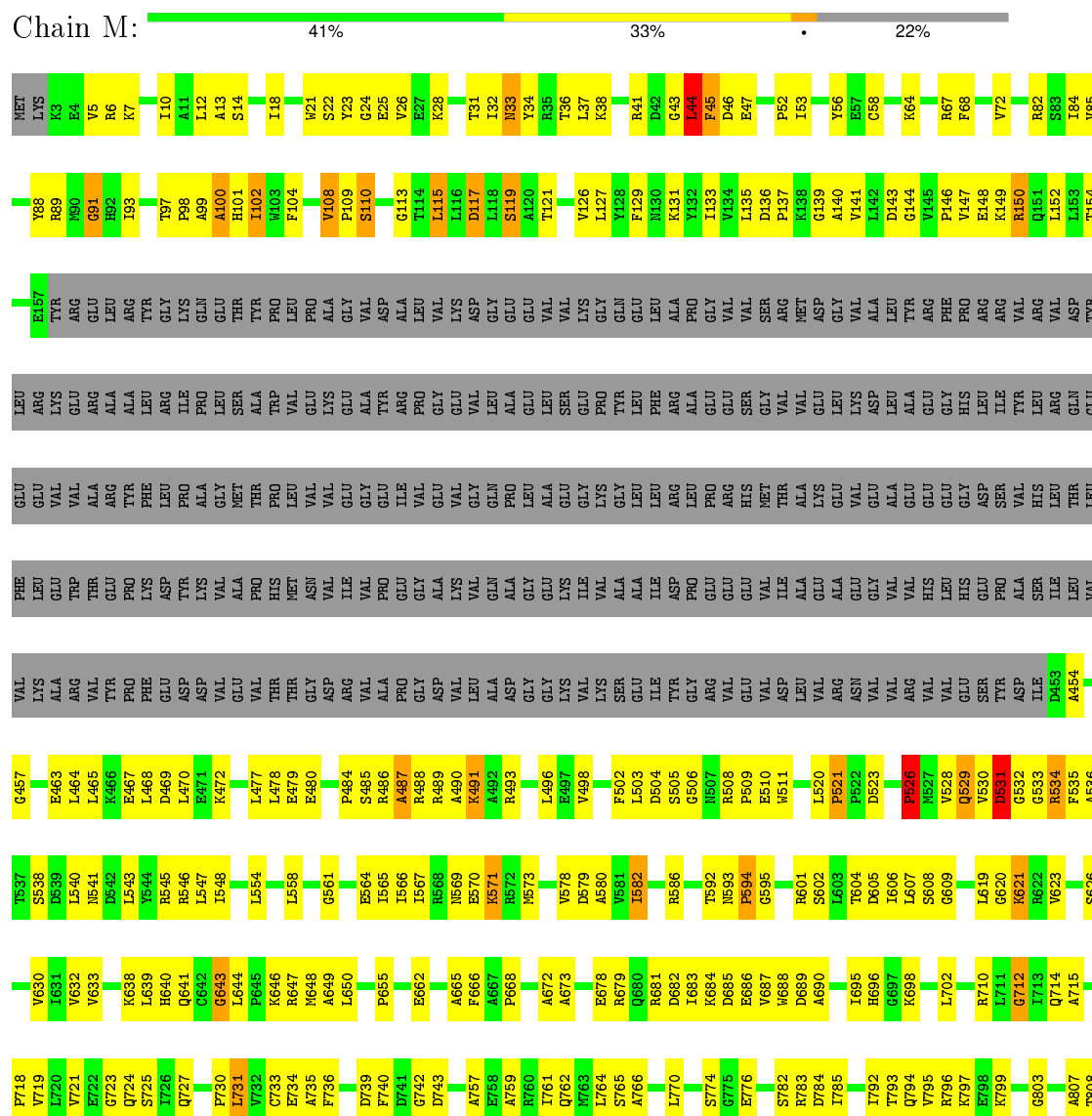


• Molecule 3: RNA POLYMERASE, BETA-PRIME SUBUNIT





• Molecule 3: RNA POLYMERASE, BETA-PRIME SUBUNIT







## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	155.00Å 271.20Å 155.30Å 90.00° 91.40° 90.00°	Depositor
Resolution (Å)	(Not available) – 4.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-4.00)	Depositor
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	?	Depositor
R, $R_{free}$	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	18756	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

## 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	0.37	0/671	1.02	0/670
1	B	0.36	0/659	1.11	0/658
1	J	0.37	0/671	1.02	0/670
1	K	0.36	0/659	1.11	0/658
2	C	0.85	9/3246 (0.3%)	1.37	15/3240 (0.5%)
2	L	0.85	9/3246 (0.3%)	1.37	15/3240 (0.5%)
3	D	0.94	21/3545 (0.6%)	1.34	21/3541 (0.6%)
3	M	0.94	21/3545 (0.6%)	1.34	21/3541 (0.6%)
4	E	0.34	0/275	0.87	0/274
4	N	0.34	0/275	0.87	0/274
5	H	0.59	3/964 (0.3%)	0.85	4/962 (0.4%)
5	Q	0.59	3/964 (0.3%)	0.85	4/962 (0.4%)
All	All	0.80	66/18720 (0.4%)	1.26	80/18690 (0.4%)

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
2	C	0	8
2	L	0	8
3	D	0	7
3	M	0	7
All	All	0	30

The worst 5 of 66 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	244	PRO	C-N	-22.34	0.92	1.33
2	L	244	PRO	C-N	-22.31	0.92	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	91	GLY	C-N	-18.07	0.92	1.34
3	M	91	GLY	C-N	-18.07	0.92	1.34
3	D	531	ASP	C-N	-17.95	1.00	1.33

The worst 5 of 80 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	244	PRO	CA-C-N	24.93	166.06	116.20
2	C	244	PRO	CA-C-N	24.90	166.00	116.20
2	L	417	GLY	C-N-CA	20.66	173.35	121.70
2	C	417	GLY	C-N-CA	20.66	173.34	121.70
3	D	531	ASP	CA-C-N	19.67	155.53	116.20

There are no chirality outliers.

5 of 30 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	242	LEU	Peptide
2	C	244	PRO	Peptide
2	C	416	GLY	Peptide
2	C	417	GLY	Peptide
2	C	418	LEU	Peptide

## 5.2 Too-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 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	672	0	241	13	0
1	B	660	0	238	10	0
1	J	672	0	241	13	0
1	K	660	0	238	9	0
2	C	3252	0	1190	105	0
2	L	3252	0	1190	107	0
3	D	3549	0	1280	135	0
3	M	3549	0	1280	134	0
4	E	276	0	96	14	0
4	N	276	0	96	15	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	H	966	0	342	14	0
5	Q	966	0	342	14	0
6	D	1	0	0	0	0
6	M	1	0	0	0	0
7	D	2	0	0	0	0
7	M	2	0	0	0	0
All	All	18756	0	6774	571	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 22.

The worst 5 of 571 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:44:LEU:C	3:D:45:PHE:N	1.70	1.42
3:M:44:LEU:C	3:M:45:PHE:N	1.70	1.41
3:M:933:ALA:C	3:M:935:LYS:H	1.80	0.83
2:L:1033:GLY:HA2	3:M:620:GLY:HA2	1.60	0.82
2:L:542:LEU:C	2:L:544:THR:H	1.83	0.82

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	222/314 (71%)	95 (43%)	65 (29%)	62 (28%)	0	0
1	B	218/314 (69%)	110 (50%)	53 (24%)	55 (25%)	0	1
1	J	222/314 (71%)	95 (43%)	65 (29%)	62 (28%)	0	0
1	K	218/314 (69%)	110 (50%)	53 (24%)	55 (25%)	0	1
2	C	1072/1118 (96%)	439 (41%)	296 (28%)	337 (31%)	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	L	1072/1118 (96%)	439 (41%)	294 (27%)	339 (32%)	0	0
3	D	1175/1524 (77%)	456 (39%)	307 (26%)	412 (35%)	0	0
3	M	1175/1524 (77%)	456 (39%)	309 (26%)	410 (35%)	0	0
4	E	90/99 (91%)	40 (44%)	24 (27%)	26 (29%)	0	0
4	N	90/99 (91%)	40 (44%)	24 (27%)	26 (29%)	0	0
5	H	318/332 (96%)	209 (66%)	57 (18%)	52 (16%)	0	5
5	Q	318/332 (96%)	209 (66%)	57 (18%)	52 (16%)	0	5
All	All	6190/7402 (84%)	2698 (44%)	1604 (26%)	1888 (30%)	0	0

5 of 1888 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	59	GLU
1	A	64	GLU
1	A	76	VAL
1	A	125	PRO
1	A	137	LYS

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 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 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.

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.