



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 04:50 PM GMT

PDB ID : 4G7H  
Title : Crystal structure of Thermus thermophilus transcription initiation complex  
Authors : Zhang, Y.; Ebright, R.H.  
Deposited on : 2012-07-20  
Resolution : 2.90 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) : 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

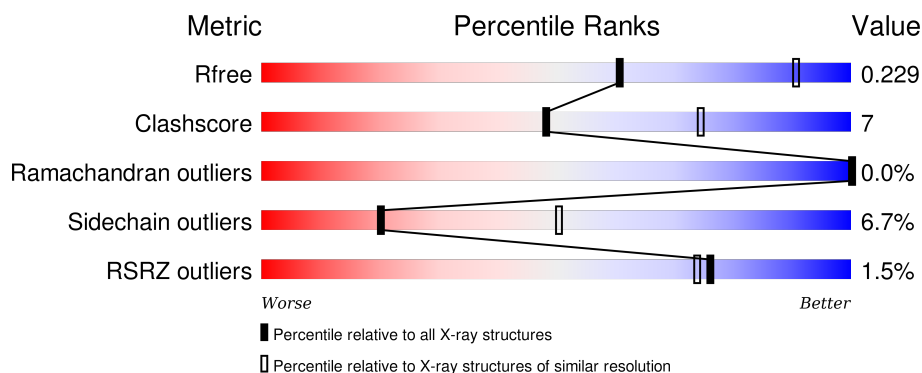
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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.

Mol	Chain	Length	Quality of chain
1	A	315	<div> <div></div> <div>54%15%•28%</div> </div>
1	B	315	<div> <div>%</div> <div>57%11%•30%</div> </div>
1	K	315	<div> <div></div> <div>53%17%•28%</div> </div>
1	L	315	<div> <div>%</div> <div>59%11%•29%</div> </div>
2	C	1119	<div> <div>%</div> <div>77%20%••</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	1119	
3	D	1524	
3	N	1524	
4	E	99	
4	O	99	
5	F	443	
5	P	443	
6	G	19	
6	Q	19	
7	H	27	
7	R	27	

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
8	MG	B	2001	-	-	-	X
8	MG	D	2004	-	-	-	X
8	MG	K	1001	-	-	-	X
8	MG	N	2004	-	-	-	X
9	ZN	D	2001	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 57420 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	B	222	Total	C	N	O	S	0	0	0
			1750	1118	304	326	2			
1	K	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	L	225	Total	C	N	O	S	0	0	0
			1773	1133	308	330	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1111	Total	C	N	O	S	0	0	0
			8770	5548	1564	1634	24			
2	M	1111	Total	C	N	O	S	0	0	0
			8770	5548	1564	1634	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1486	Total	C	N	O	S	0	0	0
			11738	7441	2067	2195	35			
3	N	1486	Total	C	N	O	S	0	0	0
			11738	7441	2067	2195	35			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	94	Total	C	N	O	S	0	0	0
			761	486	132	139	4			
4	O	94	Total	C	N	O	S	0	0	0
			761	486	132	139	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			
5	P	347	Total	C	N	O	S	0	0	0
			2814	1774	510	526	4			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	EXPRESSION TAG	UNP Q5SKW1
F	-18	GLY	-	EXPRESSION TAG	UNP Q5SKW1
F	-17	SER	-	EXPRESSION TAG	UNP Q5SKW1
F	-16	SER	-	EXPRESSION TAG	UNP Q5SKW1
F	-15	HIS	-	EXPRESSION TAG	UNP Q5SKW1
F	-14	HIS	-	EXPRESSION TAG	UNP Q5SKW1
F	-13	HIS	-	EXPRESSION TAG	UNP Q5SKW1
F	-12	HIS	-	EXPRESSION TAG	UNP Q5SKW1
F	-11	HIS	-	EXPRESSION TAG	UNP Q5SKW1
F	-10	HIS	-	EXPRESSION TAG	UNP Q5SKW1
F	-9	SER	-	EXPRESSION TAG	UNP Q5SKW1
F	-8	SER	-	EXPRESSION TAG	UNP Q5SKW1
F	-7	GLY	-	EXPRESSION TAG	UNP Q5SKW1
F	-6	LEU	-	EXPRESSION TAG	UNP Q5SKW1
F	-5	VAL	-	EXPRESSION TAG	UNP Q5SKW1
F	-4	PRO	-	EXPRESSION TAG	UNP Q5SKW1
F	-3	ARG	-	EXPRESSION TAG	UNP Q5SKW1
F	-2	GLY	-	EXPRESSION TAG	UNP Q5SKW1
F	-1	SER	-	EXPRESSION TAG	UNP Q5SKW1
F	0	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-19	MET	-	EXPRESSION TAG	UNP Q5SKW1
P	-18	GLY	-	EXPRESSION TAG	UNP Q5SKW1
P	-17	SER	-	EXPRESSION TAG	UNP Q5SKW1
P	-16	SER	-	EXPRESSION TAG	UNP Q5SKW1
P	-15	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-14	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-13	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-12	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-11	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-10	HIS	-	EXPRESSION TAG	UNP Q5SKW1
P	-9	SER	-	EXPRESSION TAG	UNP Q5SKW1
P	-8	SER	-	EXPRESSION TAG	UNP Q5SKW1
P	-7	GLY	-	EXPRESSION TAG	UNP Q5SKW1
P	-6	LEU	-	EXPRESSION TAG	UNP Q5SKW1

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
P	-5	VAL	-	EXPRESSION TAG	UNP Q5SKW1
P	-4	PRO	-	EXPRESSION TAG	UNP Q5SKW1
P	-3	ARG	-	EXPRESSION TAG	UNP Q5SKW1
P	-2	GLY	-	EXPRESSION TAG	UNP Q5SKW1
P	-1	SER	-	EXPRESSION TAG	UNP Q5SKW1
P	0	HIS	-	EXPRESSION TAG	UNP Q5SKW1

- Molecule 6 is a DNA chain called 5'-D(\*CP\*CP\*T\*GP\*CP\*AP\*TP\*CP\*CP\*GP\*TP\*GP\*AP\*GP\*TP\*CP\*GP\*AP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	16	Total	C	N	O	P	0	0	0
			328	156	63	94	15			
6	Q	16	Total	C	N	O	P	0	0	0
			328	156	63	94	15			

- Molecule 7 is a DNA chain called 5'-D(\*TP\*AP\*TP\*AP\*AP\*TP\*GP\*GP\*GP\*AP\*GP\*C P\*TP\*GP\*TP\*CP\*AP\*CP\*GP\*GP\*AP\*TP\*GP\*CP\*AP\*GP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			
7	R	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	P	1	Total	Mg	0	0
			1	1		
8	D	3	Total	Mg	0	0
			3	3		
8	K	1	Total	Mg	0	0
			1	1		
8	B	1	Total	Mg	0	0
			1	1		
8	N	3	Total	Mg	0	0
			3	3		
8	F	1	Total	Mg	0	0
			1	1		

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

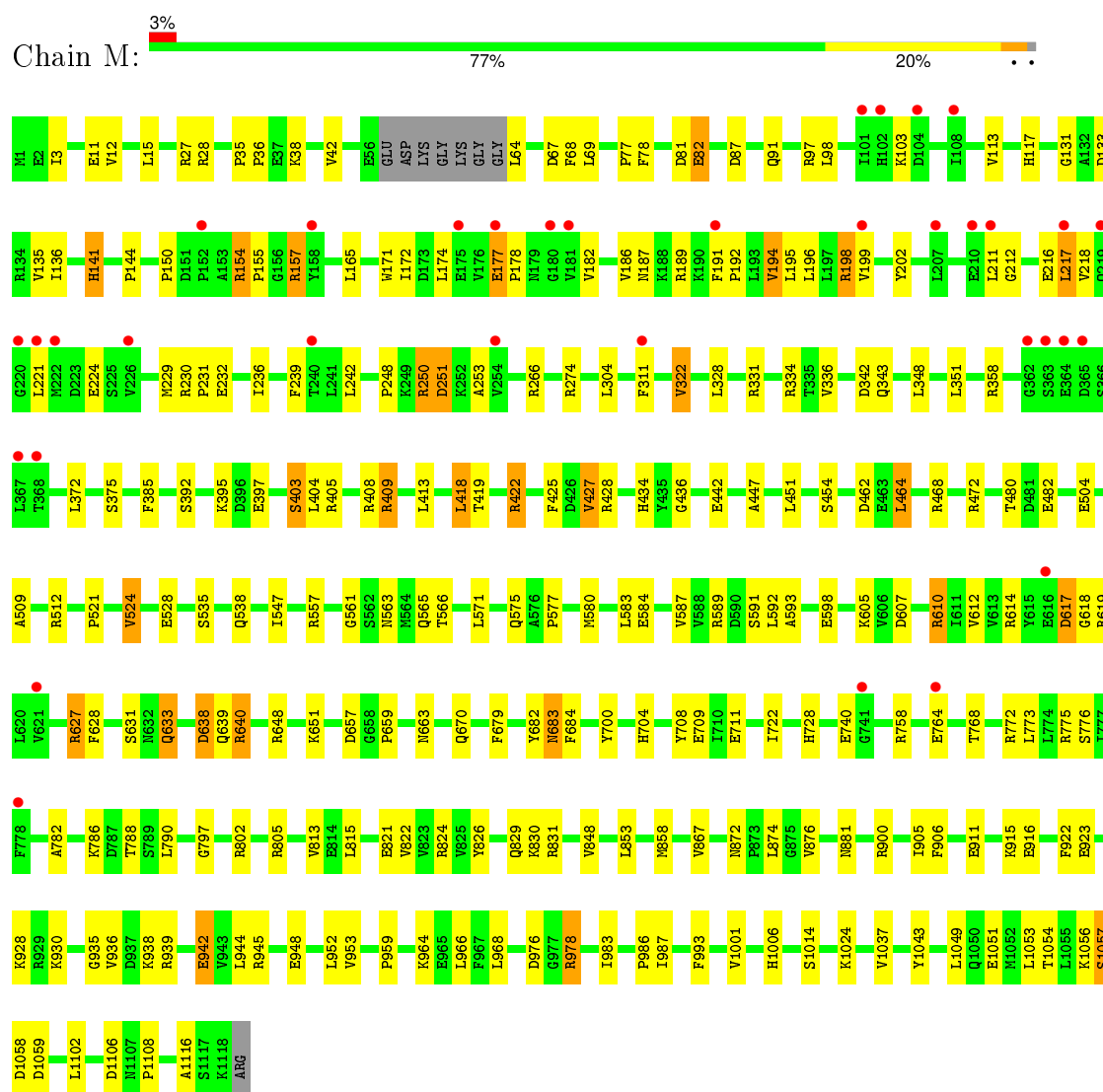
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	16	Total 16	O 16	0	0
10	B	4	Total 4	O 4	0	0
10	C	104	Total 104	O 104	0	0
10	D	138	Total 138	O 138	0	0
10	E	5	Total 5	O 5	0	0
10	F	32	Total 32	O 32	0	0
10	G	7	Total 7	O 7	0	0
10	H	6	Total 6	O 6	0	0
10	K	12	Total 12	O 12	0	0
10	L	8	Total 8	O 8	0	0
10	M	58	Total 58	O 58	0	0
10	N	89	Total 89	O 89	0	0
10	O	6	Total 6	O 6	0	0
10	P	21	Total 21	O 21	0	0
10	Q	3	Total 3	O 3	0	0
10	R	5	Total 5	O 5	0	0

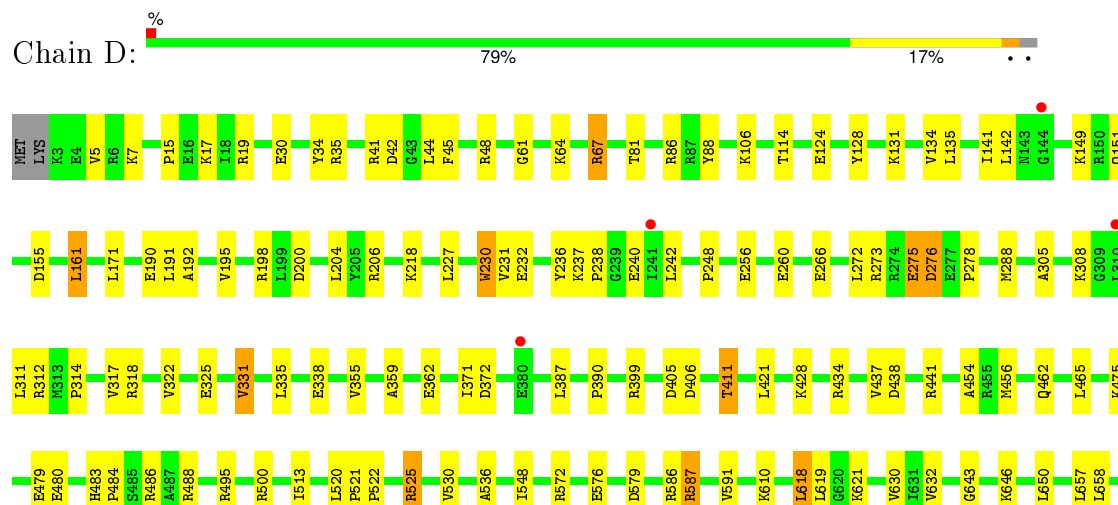


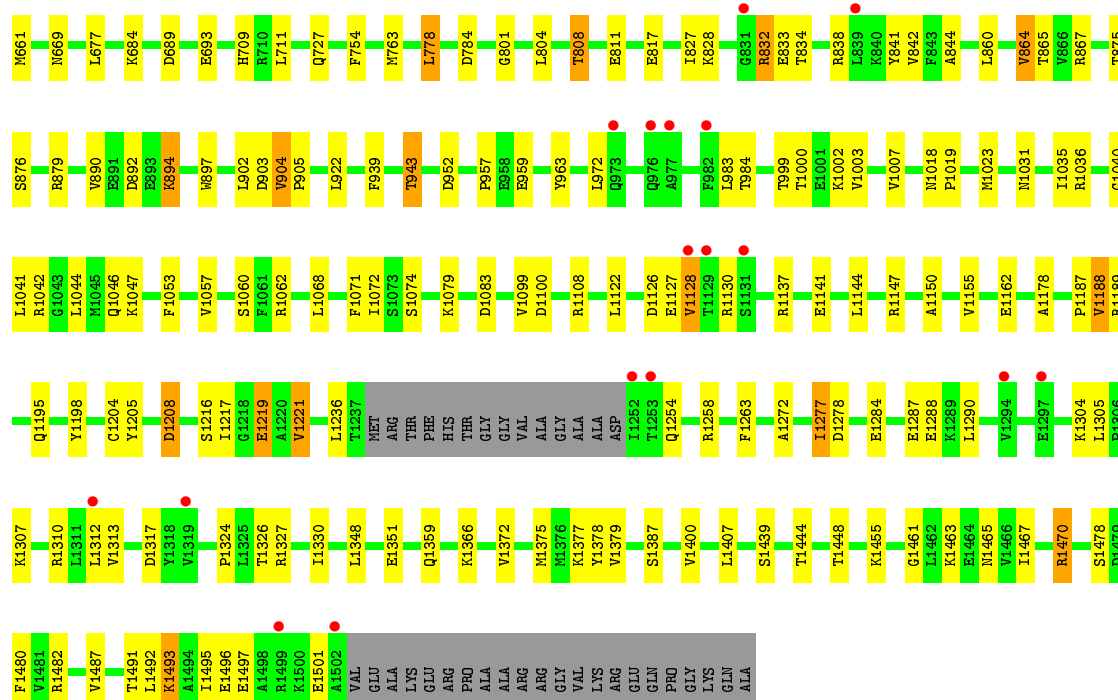




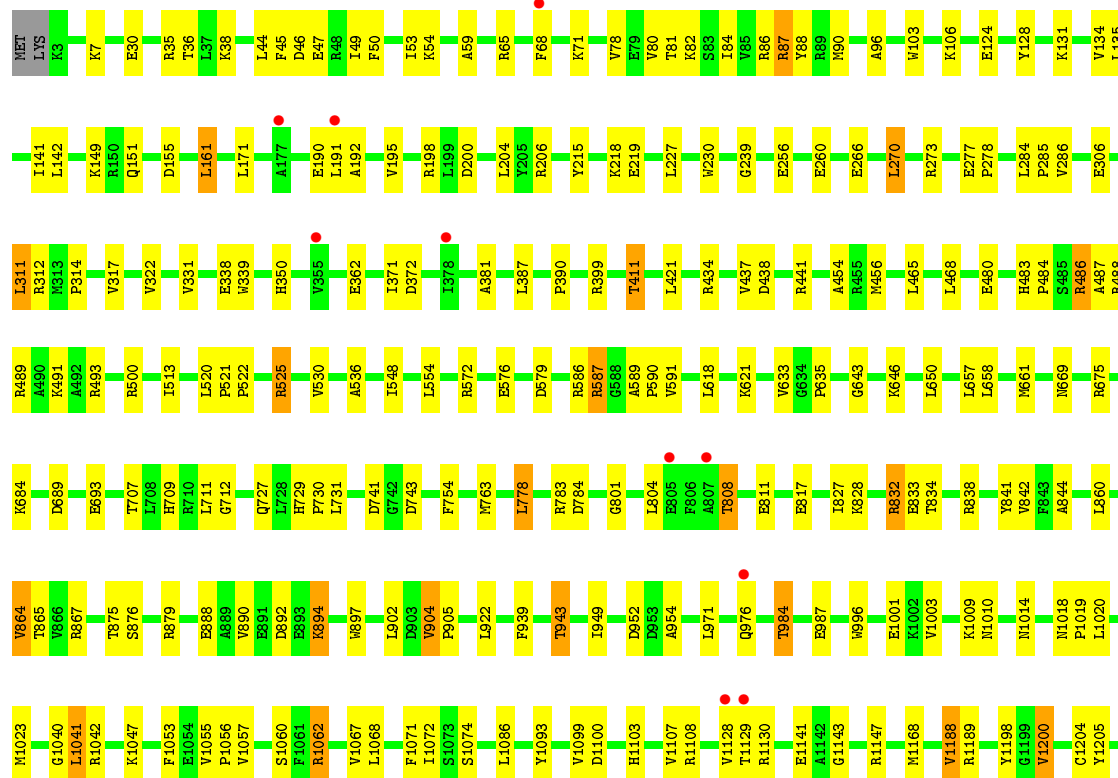
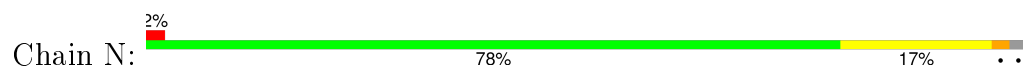


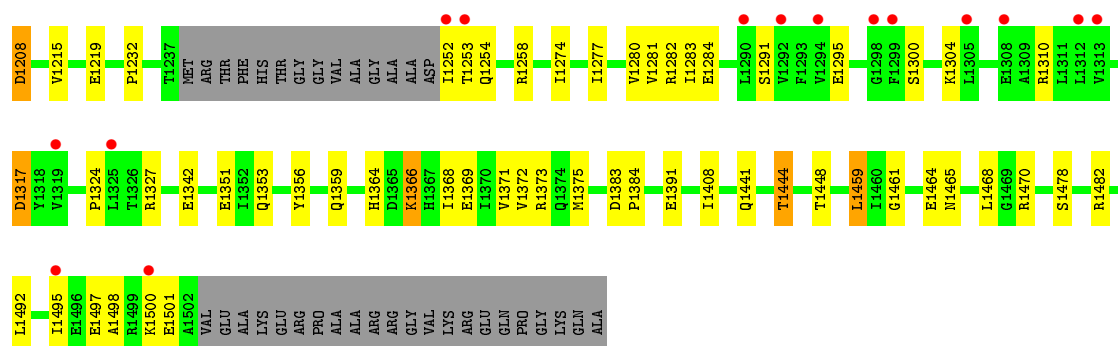
- Molecule 3: DNA-directed RNA polymerase subunit beta'





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





• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E: 79% 16% 5%



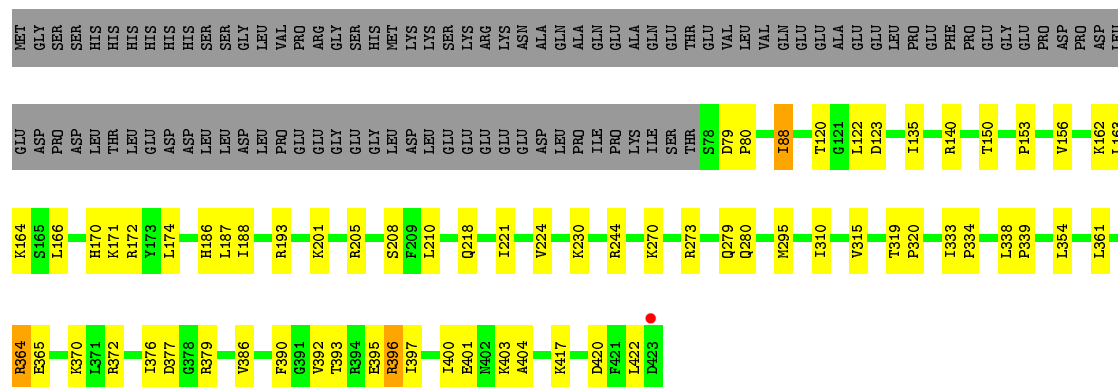
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain O: 77% 18% 5%



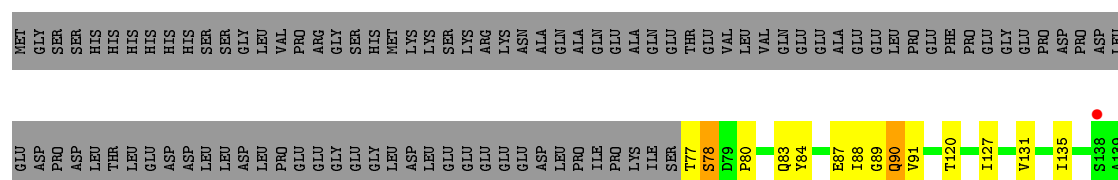
• Molecule 5: RNA polymerase sigma factor

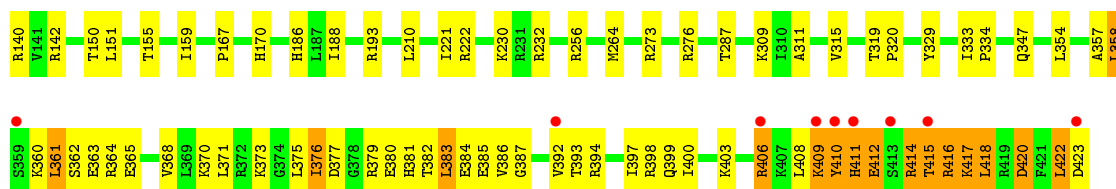
Chain F: 63% 15% 22%



• Molecule 5: RNA polymerase sigma factor

Chain P: 2% 58% 16% 22%





- Molecule 6: 5'-D(\*CP\*CP\*T\*GP\*CP\*AP\*TP\*CP\*CP\*GP\*TP\*GP\*AP\*GP\*TP\*CP\*GP\*AP\*G)-3'

Chain G: 68% 16% 16%



- Molecule 6: 5'-D(\*CP\*CP\*T\*GP\*CP\*AP\*TP\*CP\*CP\*GP\*TP\*GP\*AP\*GP\*TP\*CP\*GP\*AP\*G)-3'

Chain Q: 63% 21% 16%



- Molecule 7: 5'-D(\*TP\*AP\*TP\*AP\*AP\*TP\*GP\*GP\*GP\*AP\*GP\*CP\*TP\*GP\*TP\*CP\*AP\*CP\*GP\*GP\*AP\*TP\*GP\*CP\*AP\*GP\*G)-3'

Chain H: 33% 44% 11% 11%



- Molecule 7: 5'-D(\*TP\*AP\*TP\*AP\*AP\*TP\*GP\*GP\*GP\*AP\*GP\*CP\*TP\*GP\*TP\*CP\*AP\*CP\*GP\*GP\*AP\*TP\*GP\*CP\*AP\*GP\*G)-3'

Chain R: 44% 37% 7% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	185.56Å 104.57Å 297.55Å 90.00° 98.32° 90.00°	Depositor
Resolution (Å)	49.27 – 2.90 49.90 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.27-2.90) 99.6 (49.90-2.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, $R_{free}$	0.188 , 0.226 0.191 , 0.229	Depositor DCC
$R_{free}$ test set	2191 reflections (0.88%)	DCC
Wilson B-factor (Å <sup>2</sup> )	61.5	Xtriage
Anisotropy	0.653	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 46.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	3 of 249334 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	57420	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	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 25.73 % 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 3.0020e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 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	0.32	0/1814	0.59	2/2466 (0.1%)
1	B	0.30	0/1782	0.52	0/2424
1	K	0.29	0/1814	0.67	5/2466 (0.2%)
1	L	0.29	0/1805	0.63	3/2454 (0.1%)
2	C	0.33	0/8937	0.54	4/12087 (0.0%)
2	M	0.32	0/8937	0.54	3/12087 (0.0%)
3	D	0.32	0/11944	0.50	0/16149
3	N	0.31	0/11944	0.50	1/16149 (0.0%)
4	E	0.28	0/775	0.45	0/1045
4	O	0.30	0/775	0.44	0/1045
5	F	0.30	0/2852	0.47	0/3837
5	P	0.30	0/2859	0.50	0/3847
6	G	0.60	0/368	1.26	3/567 (0.5%)
6	Q	0.54	0/368	1.27	4/567 (0.7%)
7	H	0.59	0/556	1.33	6/858 (0.7%)
7	R	0.58	0/556	1.35	7/858 (0.8%)
All	All	0.33	0/58086	0.57	38/78906 (0.0%)

There are no bond length outliers.

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	422	ARG	NE-CZ-NH2	-13.77	113.42	120.30
2	M	422	ARG	NE-CZ-NH1	13.31	126.95	120.30
1	L	112	ARG	NE-CZ-NH1	-12.08	114.26	120.30
2	C	422	ARG	NE-CZ-NH1	-11.95	114.33	120.30
1	L	112	ARG	NE-CZ-NH2	11.77	126.19	120.30

There are no chirality outliers.

There are no planarity outliers.

## 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	1782	0	1834	41	0
1	B	1750	0	1797	29	0
1	K	1782	0	1834	35	0
1	L	1773	0	1826	25	0
2	C	8770	0	8874	146	0
2	M	8770	0	8874	158	0
3	D	11738	0	11971	157	0
3	N	11738	0	11971	158	0
4	E	761	0	778	11	0
4	O	761	0	778	12	0
5	F	2807	0	2882	41	0
5	P	2814	0	2889	85	0
6	G	328	0	181	1	0
6	Q	328	0	181	1	0
7	H	495	0	272	12	0
7	R	495	0	272	10	0
8	B	1	0	0	0	0
8	D	3	0	0	0	0
8	F	1	0	0	0	0
8	K	1	0	0	0	0
8	N	3	0	0	0	0
8	P	1	0	0	0	0
9	D	2	0	0	0	0
9	N	2	0	0	0	0
10	A	16	0	0	2	0
10	B	4	0	0	0	0
10	C	104	0	0	8	0
10	D	138	0	0	5	0
10	E	5	0	0	0	0
10	F	32	0	0	3	0
10	G	7	0	0	0	0
10	H	6	0	0	1	0
10	K	12	0	0	0	0
10	L	8	0	0	0	0
10	M	58	0	0	1	0
10	N	89	0	0	2	0
10	O	6	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	P	21	0	0	0	0
10	Q	3	0	0	0	0
10	R	5	0	0	0	0
All	All	57420	0	57214	835	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 7.

The worst 5 of 835 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:P:415:THR:HG22	5:P:416:ARG:CG	1.50	1.41
5:P:415:THR:CG2	5:P:416:ARG:HG3	1.51	1.40
2:M:172:ILE:HG13	2:M:186:VAL:HG22	1.19	1.14
5:P:415:THR:HG22	5:P:416:ARG:CD	1.89	1.01
2:C:627:ARG:NH1	2:C:638:ASP:OD2	1.92	1.01

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	224/315 (71%)	222 (99%)	2 (1%)	0	100	100
1	B	220/315 (70%)	215 (98%)	5 (2%)	0	100	100
1	K	224/315 (71%)	222 (99%)	2 (1%)	0	100	100
1	L	223/315 (71%)	218 (98%)	5 (2%)	0	100	100
2	C	1107/1119 (99%)	1086 (98%)	21 (2%)	0	100	100
2	M	1107/1119 (99%)	1080 (98%)	27 (2%)	0	100	100
3	D	1482/1524 (97%)	1455 (98%)	26 (2%)	1 (0%)	56	87

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	N	1482/1524 (97%)	1454 (98%)	27 (2%)	1 (0%)	56	87
4	E	92/99 (93%)	89 (97%)	2 (2%)	1 (1%)	17	51
4	O	92/99 (93%)	90 (98%)	2 (2%)	0	100	100
5	F	344/443 (78%)	339 (98%)	5 (2%)	0	100	100
5	P	345/443 (78%)	341 (99%)	4 (1%)	0	100	100
All	All	6942/7630 (91%)	6811 (98%)	128 (2%)	3 (0%)	100	100

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	E	94	PRO
3	D	530	VAL
3	N	530	VAL

### 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	199/273 (73%)	187 (94%)	12 (6%)	24	57
1	B	195/273 (71%)	185 (95%)	10 (5%)	29	65
1	K	199/273 (73%)	188 (94%)	11 (6%)	27	61
1	L	198/273 (72%)	191 (96%)	7 (4%)	43	78
2	C	936/941 (100%)	869 (93%)	67 (7%)	18	46
2	M	936/941 (100%)	870 (93%)	66 (7%)	18	47
3	D	1253/1279 (98%)	1166 (93%)	87 (7%)	19	48
3	N	1253/1279 (98%)	1169 (93%)	84 (7%)	20	50
4	E	83/88 (94%)	82 (99%)	1 (1%)	78	94
4	O	83/88 (94%)	82 (99%)	1 (1%)	78	94
5	F	301/388 (78%)	285 (95%)	16 (5%)	28	63
5	P	302/388 (78%)	269 (89%)	33 (11%)	8	23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	5938/6484 (92%)	5543 (93%)	395 (7%)	20	50

5 of 395 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	123	ASP
2	M	174	LEU
5	P	150	THR
5	F	208	SER
1	K	186	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1172	HIS
3	D	1195	GLN
2	M	187	ASN
3	N	350	HIS
5	P	90	GLN

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	226/315 (71%)	-0.34	0 100 100	45, 63, 96, 118	0
1	B	222/315 (70%)	-0.20	2 (0%) 85 84	47, 77, 114, 137	0
1	K	226/315 (71%)	-0.26	0 100 100	49, 66, 97, 116	0
1	L	225/315 (71%)	-0.19	2 (0%) 85 84	51, 81, 118, 139	0
2	C	1111/1119 (99%)	-0.14	8 (0%) 89 88	32, 59, 111, 149	0
2	M	1111/1119 (99%)	-0.05	35 (3%) 51 43	34, 65, 129, 152	0
3	D	1486/1524 (97%)	-0.12	21 (1%) 78 76	29, 62, 120, 177	1 (0%)
3	N	1486/1524 (97%)	-0.14	25 (1%) 73 70	32, 64, 118, 178	1 (0%)
4	E	94/99 (94%)	-0.31	0 100 100	41, 64, 111, 121	0
4	O	94/99 (94%)	-0.38	0 100 100	45, 67, 112, 122	0
5	F	346/443 (78%)	-0.24	1 (0%) 94 94	37, 63, 108, 132	0
5	P	347/443 (78%)	-0.13	10 (2%) 55 49	42, 72, 139, 172	0
6	G	16/19 (84%)	-0.32	0 100 100	57, 84, 176, 182	0
6	Q	16/19 (84%)	-0.36	0 100 100	69, 93, 180, 182	0
7	H	24/27 (88%)	-0.35	0 100 100	54, 98, 145, 192	0
7	R	24/27 (88%)	-0.25	0 100 100	55, 112, 147, 199	0
All	All	7054/7722 (91%)	-0.15	104 (1%) 76 74	29, 65, 121, 199	2 (0%)

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	P	410	TYR	4.3
2	M	363	SER	4.2
5	P	411	HIS	4.1
3	D	144	GLY	4.0
3	N	1129	THR	3.9

## 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
8	MG	K	1001	1/1	0.90	0.46	18.86	67,67,67,67	0
8	MG	N	2004	1/1	0.91	0.45	4.04	66,66,66,66	0
8	MG	D	2004	1/1	0.96	0.48	3.60	66,66,66,66	0
9	ZN	D	2001	1/1	0.99	0.23	2.72	77,77,77,77	0
8	MG	B	2001	1/1	0.51	0.27	2.58	71,71,71,71	0
9	ZN	N	2001	1/1	0.99	0.14	-0.55	54,54,54,54	0
9	ZN	N	2002	1/1	0.99	0.09	-0.95	95,95,95,95	0
9	ZN	D	2002	1/1	0.99	0.10	-1.13	73,73,73,73	0
8	MG	F	2001	1/1	0.87	0.08	-2.38	63,63,63,63	0
8	MG	P	2001	1/1	0.93	0.09	-	71,71,71,71	0
8	MG	D	2003	1/1	0.97	0.30	-	32,32,32,32	0
8	MG	N	2003	1/1	0.98	0.37	-	38,38,38,38	0
8	MG	N	2005	1/1	0.92	0.15	-	67,67,67,67	0
8	MG	D	2005	1/1	0.88	0.13	-	71,71,71,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.