



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 12:16 AM UTC

PDB ID : 6TYF / pdb_00006tyf
Title : Crystal structure of MTB sigma L transcription initiation complex with 6 nt long RNA primer
Authors : Molodtsov, V.; Ebright, R.H.
Deposited on : 2019-08-08
Resolution : 3.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Xtrriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

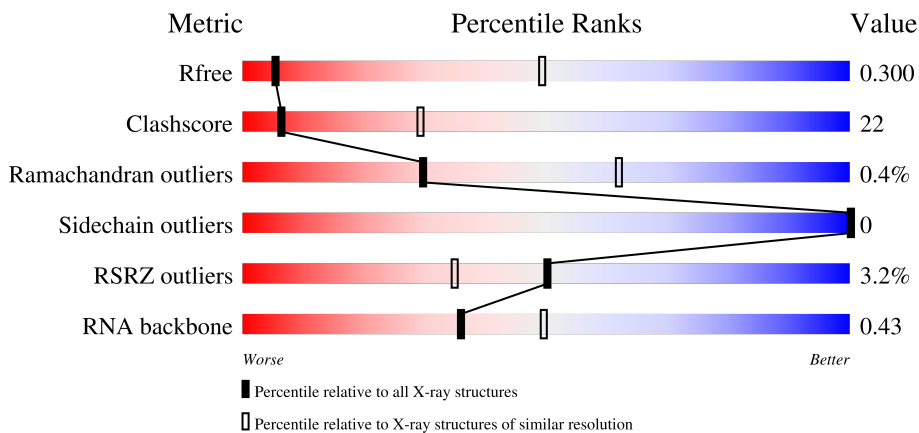
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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.



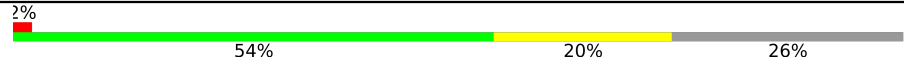
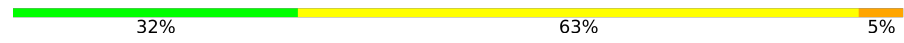
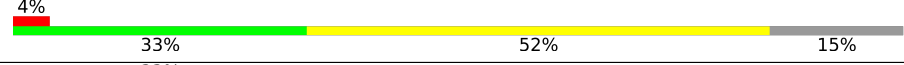


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1065 (3.96-3.64)
Clashscore	190562	1012 (3.94-3.66)
Ramachandran outliers	187476	1048 (3.96-3.64)
Sidechain outliers	187428	1043 (3.96-3.64)
RSRZ outliers	180081	1064 (3.96-3.64)
RNA backbone	3983	1007 (4.50-3.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	347	 3% 44% 21% 35%
1	B	347	 3% 35% 31% 33%
2	C	1178	 2% 59% 35% 5%
3	D	1316	 3% 60% 35% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	110	 <p>2% 54% 20% 26%</p>
5	G	19	 <p>32% 63% 5%</p>
6	H	27	 <p>4% 33% 52% 15%</p>
7	I	6	 <p>33% 50% 33% 17%</p>
8	F	177	 <p>8% 79% 19%</p>

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 25308 atoms, of which 372 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
			Total	C	N	O	S			
1	A	225	Total	C	N	O	S	0	0	0
			1716	1080	296	338	2			
1	B	232	Total	C	N	O	S	0	0	0
			1732	1093	296	341	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	1114	Total	C	N	O	S	0	0	0
			8643	5411	1512	1681	39			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	1262	Total	C	N	O	S	0	0	0
			9872	6182	1790	1860	40			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	E	81	Total	C	N	O	0	0	0
			630	403	106	121			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
5	G	19	Total	C	H	N	O	P	0	0	0
			504	186	113	75	112	18			

- Molecule 6 is a DNA chain called DNA (5'-D(P*GP*TP*GP*TP*CP*AP*GP*TP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				P
6	H	23	734	225	259	87	140	23	0	0	0

- Molecule 7 is a RNA chain called RNA (5'-R(P*CP*CP*UP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
7	I	6	125	56	21	42	6	0	0	0

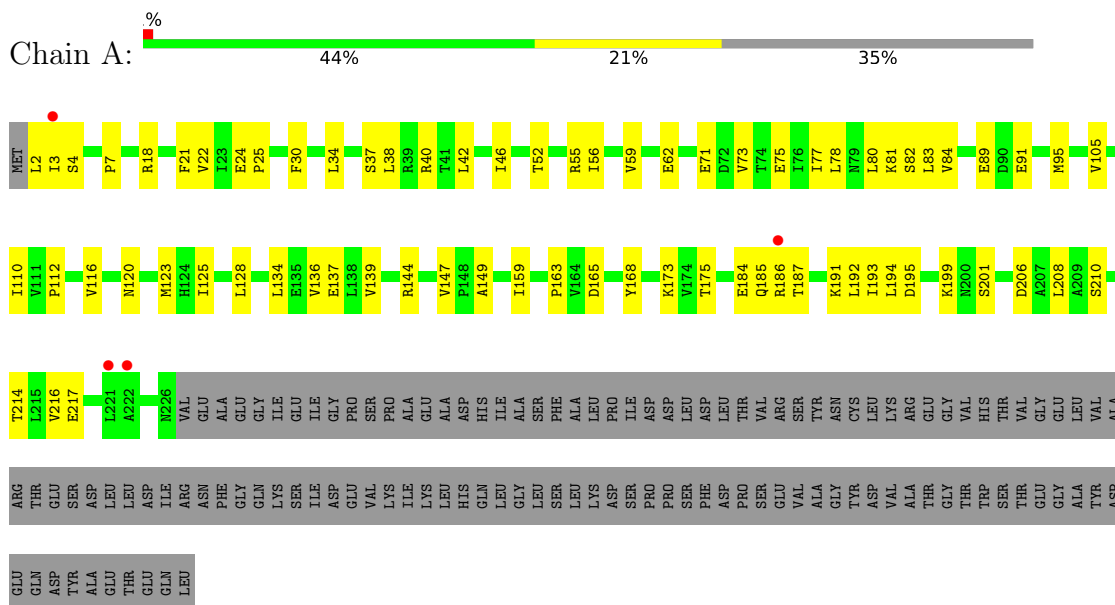
- Molecule 8 is a protein called RNA polymerase sigma factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	F	174	1352	840	256	254	2	0	0	0

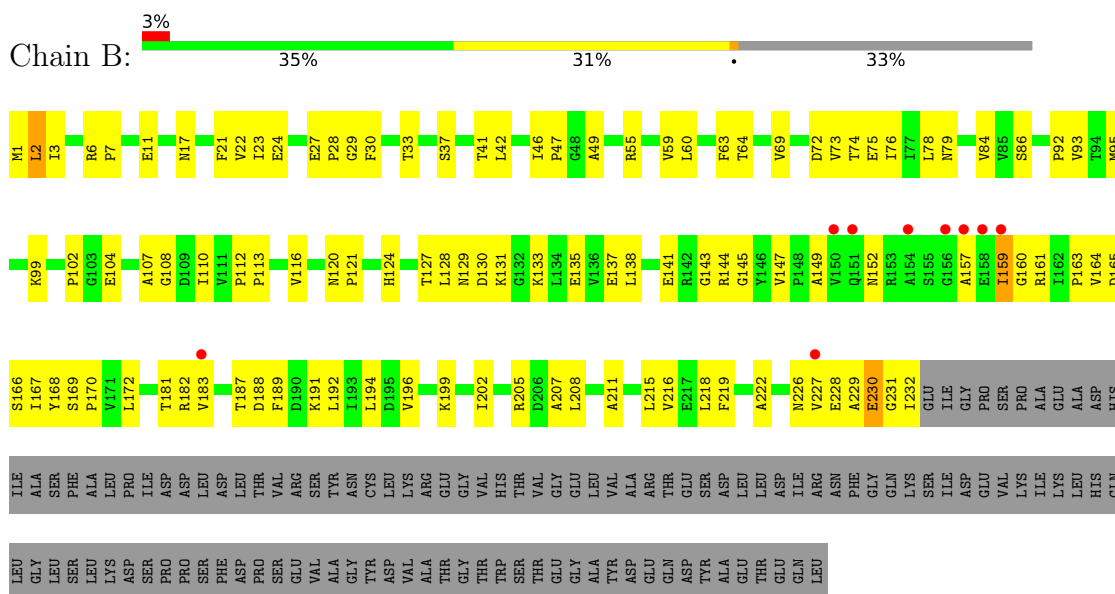
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

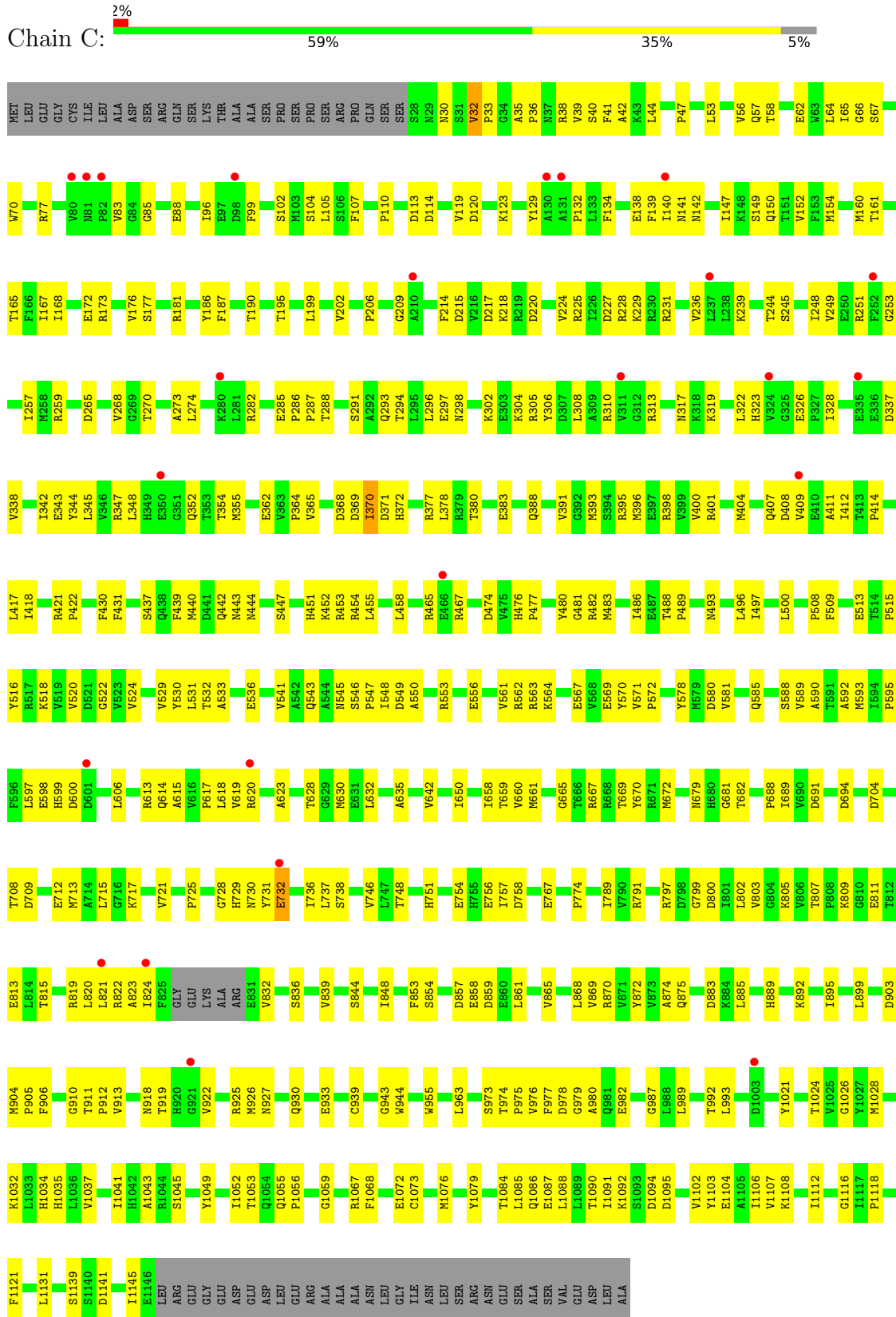
- Molecule 1: DNA-directed RNA polymerase subunit alpha



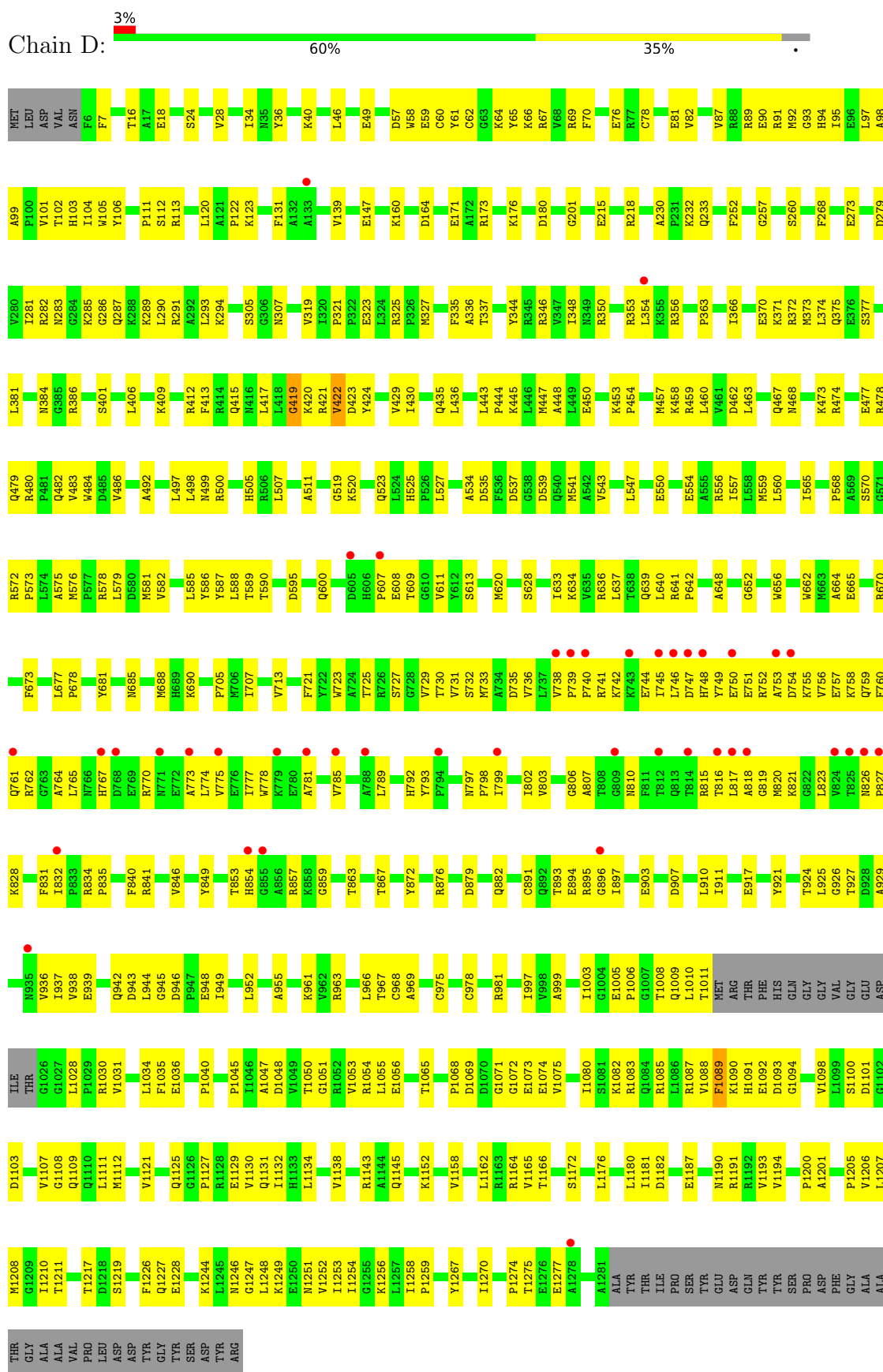
- Molecule 1: DNA-directed RNA polymerase subunit alpha



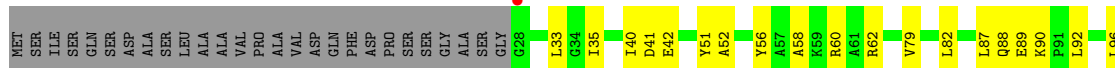
• Molecule 2: DNA-directed RNA polymerase subunit beta



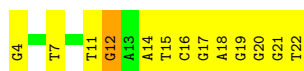
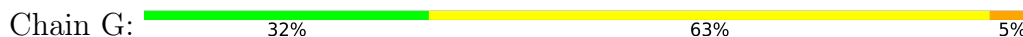
• Molecule 3: DNA-directed RNA polymerase subunit beta'



● Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 5: DNA (5'-D(*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*AP*TP*CP*GP*AP*GP*GP*GP*T)-3')



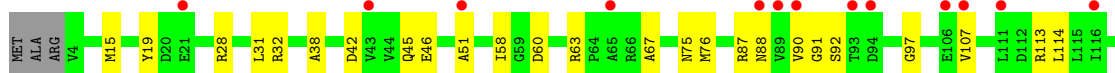
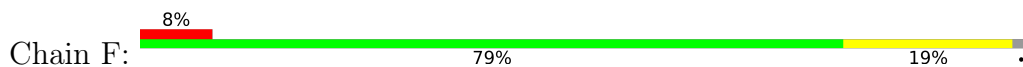
- Molecule 6: DNA (5'-D(P*GP*TP*GP*TP*CP*AP*GP*TP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*C)-3')



- Molecule 7: RNA (5'-R(P*CP*CP*UP*CP*GP*A)-3')



- Molecule 8: RNA polymerase sigma factor



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	126.43Å 161.78Å 213.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.26 – 3.80 49.26 – 3.80	Depositor EDS
% Data completeness (in resolution range)	96.8 (49.26-3.80) 96.7 (49.26-3.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.87 (at 3.77Å)	Xtrriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.256 , 0.300 0.257 , 0.300	Depositor DCC
R_{free} test set	2000 reflections (3.90%)	wwPDB-VP
Wilson B-factor (Å ²)	66.1	Xtrriage
Anisotropy	0.308	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 35.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	25308	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.25% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality i

5.1 Standard geometry i

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	0.25	0/1742	0.41	0/2370
1	B	0.24	0/1758	0.47	0/2397
2	C	0.25	0/8801	0.42	0/11933
3	D	0.29	1/10038 (0.0%)	0.50	4/13568 (0.0%)
4	E	0.23	0/643	0.42	0/877
5	G	0.39	0/439	0.63	1/677 (0.1%)
6	H	0.27	0/532	0.54	0/820
7	I	1.16	2/138 (1.4%)	1.38	5/212 (2.4%)
8	F	0.10	0/1374	0.28	0/1869
All	All	0.28	3/25465 (0.0%)	0.47	10/34723 (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
3	D	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	I	4	C	O3'-P	-8.21	1.48	1.61
7	I	3	C	O3'-P	-7.60	1.49	1.61
3	D	422	VAL	C-O	-5.20	1.17	1.24

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	I	4	C	C1'-C2'-O2'	-12.37	89.85	108.40
5	G	12	DG	C3'-C2'-C1'	-10.16	86.36	101.60
3	D	422	VAL	N-CA-C	9.71	124.24	109.20
7	I	4	C	C4'-C3'-O3'	-7.69	101.46	113.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	422	VAL	CB-CA-C	-7.53	100.42	111.19

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	419	GLY	Mainchain

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	1716	0	1756	76	0
1	B	1732	0	1754	111	0
2	C	8643	0	8575	418	2
3	D	9872	0	9942	531	4
4	E	630	0	622	28	0
5	G	391	113	215	21	0
6	H	475	259	260	40	0
7	I	125	0	66	14	0
8	F	1352	0	1346	63	0
All	All	24936	372	24536	1087	4

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 1087 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)
1:A:40:ARG:HE	1:B:33:THR:HG22	1.14	1.12
3:D:111:PRO:HB3	6:H:22:DA:H5''	1.14	1.11
3:D:111:PRO:HB3	6:H:22:DA:C5'	1.82	1.09
2:C:1103:TYR:CE2	8:F:107:VAL:HG13	1.89	1.07
3:D:1055:LEU:H	3:D:1101:ASP:HB3	1.17	1.06

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:747:ASP:N	3:D:894:GLU:OE2[4_545]	1.99	0.21
3:D:754:ASP:O	3:D:939:GLU:OE2[4_545]	2.06	0.14
2:C:791:ARG:NE	3:D:171:GLU:OE2[3_554]	2.13	0.07
2:C:667:ARG:NH2	3:D:147:GLU:OE1[3_554]	2.17	0.03

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	223/347 (64%)	215 (96%)	7 (3%)	1 (0%)	30	62
1	B	230/347 (66%)	211 (92%)	15 (6%)	4 (2%)	7	34
2	C	1110/1178 (94%)	1055 (95%)	50 (4%)	5 (0%)	24	57
3	D	1258/1316 (96%)	1201 (96%)	55 (4%)	2 (0%)	43	73
4	E	79/110 (72%)	76 (96%)	3 (4%)	0	100	100
8	F	172/177 (97%)	167 (97%)	5 (3%)	0	100	100
All	All	3072/3475 (88%)	2925 (95%)	135 (4%)	12 (0%)	30	62

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	159	ILE
2	C	370	ILE
1	B	2	LEU
2	C	732	GLU
3	D	1089	PHE

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	194/297 (65%)	194 (100%)	0	100	100
1	B	191/297 (64%)	191 (100%)	0	100	100
2	C	945/998 (95%)	945 (100%)	0	100	100
3	D	1047/1095 (96%)	1047 (100%)	0	100	100
4	E	66/90 (73%)	66 (100%)	0	100	100
8	F	134/136 (98%)	134 (100%)	0	100	100
All	All	2577/2913 (88%)	2577 (100%)	0	100	100

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

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	467	GLN
3	D	748	HIS
3	D	564	ASN
3	D	767	HIS
2	C	941	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
7	I	5/6 (83%)	2 (40%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
7	I	4	C
7	I	8	A

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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](#)

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	225/347 (64%)	0.38	4 (1%) 67 46	38, 61, 104, 157	0
1	B	232/347 (66%)	0.50	9 (3%) 43 31	48, 74, 109, 168	0
2	C	1114/1178 (94%)	0.38	24 (2%) 62 44	24, 57, 132, 167	0
3	D	1262/1316 (95%)	0.47	43 (3%) 48 33	25, 66, 203, 265	0
4	E	81/110 (73%)	0.37	2 (2%) 58 40	51, 77, 133, 142	0
5	G	19/19 (100%)	0.99	0 100 100	38, 71, 140, 149	0
6	H	23/27 (85%)	1.03	1 (4%) 40 28	79, 146, 216, 227	0
7	I	6/6 (100%)	1.57	2 (33%) 1 2	38, 45, 67, 93	0
8	F	174/177 (98%)	1.04	15 (8%) 16 16	66, 123, 159, 174	0
All	All	3136/3527 (88%)	0.47	100 (3%) 50 35	24, 67, 151, 265	0

The worst 5 of 100 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	157	ALA	4.6
3	D	771	ASN	4.5
3	D	814	THR	4.4
1	B	150	VAL	4.4
1	A	186	ARG	4.4

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 oligosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.