



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 11:32 AM UTC

PDB ID : 5X22 / pdb_00005x22
Title : Crystal structure of Thermus thermophilus transcription initiation complex with GpA and CMPcPP
Authors : Zhang, Y.; Ebright, R.
Deposited on : 2017-01-29
Resolution : 3.35 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
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
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

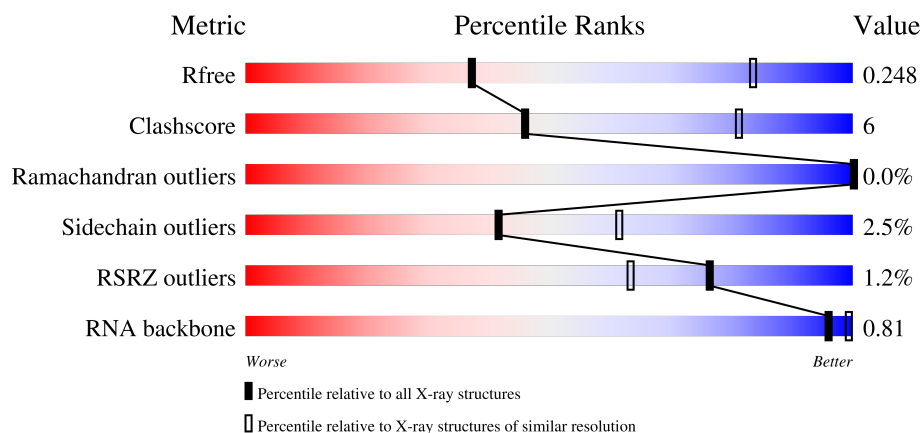
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.35 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	1099 (3.40-3.32)
Clashscore	190562	1116 (3.40-3.32)
Ramachandran outliers	187476	1101 (3.40-3.32)
Sidechain outliers	187428	1101 (3.40-3.32)
RSRZ outliers	180081	1099 (3.40-3.32)
RNA backbone	3983	1110 (3.72-3.00)

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	315	
1	B	315	
1	K	315	
1	L	315	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	1119	
2	M	1119	
3	D	1524	
3	N	1524	
4	E	99	
4	O	99	
5	F	443	
5	P	443	
6	G	21	
6	Q	21	
7	H	27	
7	R	27	
8	I	2	
8	S	2	

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 56863 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	231	Total	C	N	O	S	0	0	0
			1809	1155	315	337	2			
1	B	223	Total	C	N	O	S	0	0	0
			1758	1124	305	327	2			
1	K	231	Total	C	N	O	S	0	0	0
			1809	1155	315	337	2			
1	L	227	Total	C	N	O	S	0	0	0
			1789	1143	310	334	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1112	Total	C	N	O	S	0	0	0
			8774	5550	1565	1635	24			
2	M	1112	Total	C	N	O	S	0	0	0
			8774	5550	1565	1635	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1438	Total	C	N	O	S	0	0	0
			11299	7157	2001	2106	35			
3	N	1500	Total	C	N	O	S	0	0	0
			11834	7499	2086	2213	36			

- 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
			758	483	132	139	4			
4	O	94	Total	C	N	O	S	0	0	0
			758	483	132	139	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	327	Total	C	N	O	S	0	0	0
			2649	1671	479	495	4			
5	P	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	initiating methionine	UNP Q72L95
F	-18	GLY	-	expression tag	UNP Q72L95
F	-17	SER	-	expression tag	UNP Q72L95
F	-16	SER	-	expression tag	UNP Q72L95
F	-15	HIS	-	expression tag	UNP Q72L95
F	-14	HIS	-	expression tag	UNP Q72L95
F	-13	HIS	-	expression tag	UNP Q72L95
F	-12	HIS	-	expression tag	UNP Q72L95
F	-11	HIS	-	expression tag	UNP Q72L95
F	-10	HIS	-	expression tag	UNP Q72L95
F	-9	SER	-	expression tag	UNP Q72L95
F	-8	SER	-	expression tag	UNP Q72L95
F	-7	GLY	-	expression tag	UNP Q72L95
F	-6	LEU	-	expression tag	UNP Q72L95
F	-5	VAL	-	expression tag	UNP Q72L95
F	-4	PRO	-	expression tag	UNP Q72L95
F	-3	ARG	-	expression tag	UNP Q72L95
F	-2	GLY	-	expression tag	UNP Q72L95
F	-1	SER	-	expression tag	UNP Q72L95
F	0	HIS	-	expression tag	UNP Q72L95
P	-19	MET	-	initiating methionine	UNP Q72L95
P	-18	GLY	-	expression tag	UNP Q72L95
P	-17	SER	-	expression tag	UNP Q72L95
P	-16	SER	-	expression tag	UNP Q72L95
P	-15	HIS	-	expression tag	UNP Q72L95
P	-14	HIS	-	expression tag	UNP Q72L95
P	-13	HIS	-	expression tag	UNP Q72L95
P	-12	HIS	-	expression tag	UNP Q72L95
P	-11	HIS	-	expression tag	UNP Q72L95
P	-10	HIS	-	expression tag	UNP Q72L95
P	-9	SER	-	expression tag	UNP Q72L95
P	-8	SER	-	expression tag	UNP Q72L95
P	-7	GLY	-	expression tag	UNP Q72L95
P	-6	LEU	-	expression tag	UNP Q72L95

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	-5	VAL	-	expression tag	UNP Q72L95
P	-4	PRO	-	expression tag	UNP Q72L95
P	-3	ARG	-	expression tag	UNP Q72L95
P	-2	GLY	-	expression tag	UNP Q72L95
P	-1	SER	-	expression tag	UNP Q72L95
P	0	HIS	-	expression tag	UNP Q72L95

- Molecule 6 is a DNA chain called promoter DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	19	Total	C	N	O	P	0	0	0
			386	184	71	113	18			
6	Q	19	Total	C	N	O	P	0	0	0
			386	184	71	113	18			

- Molecule 7 is a DNA chain called promoter DNA nontemplate strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	27	Total	C	N	O	P	0	0	0
			560	266	109	159	26			
7	R	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			

- Molecule 8 is a RNA chain called RNA (5'-R(*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	2	Total	C	N	O	P	0	0	0
			42	20	10	11	1			
8	S	2	Total	C	N	O	P	0	0	0
			42	20	10	11	1			

- Molecule 9 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	1	Total	Mg	0	0
			1	1		
9	D	4	Total	Mg	0	1
			5	5		
9	F	1	Total	Mg	0	0
			1	1		
9	L	1	Total	Mg	0	0
			1	1		

Continued on next page...

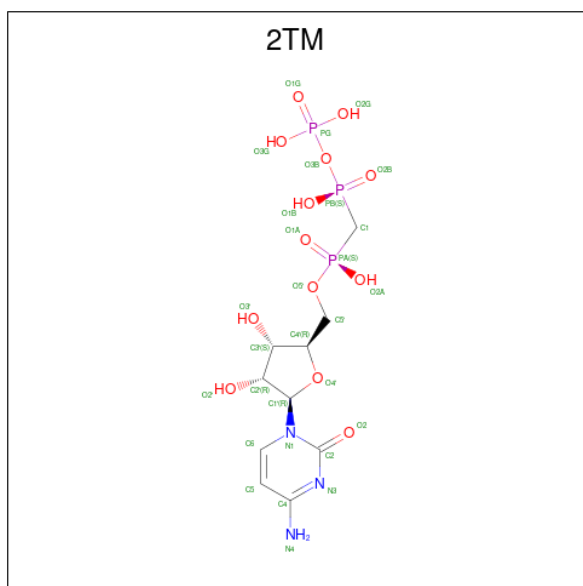
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	N	4	Total	Mg	0	1
			5	5		
9	P	1	Total	Mg	0	0
			1	1		

- Molecule 10 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 11 is 5'-O-[(S)-hydroxy{[(S)-hydroxy(phosphonooxy)phosphoryl]methyl}phosphoryl]cytidine (CCD ID: 2TM) (formula: C₁₀H₁₈N₃O₁₃P₃).

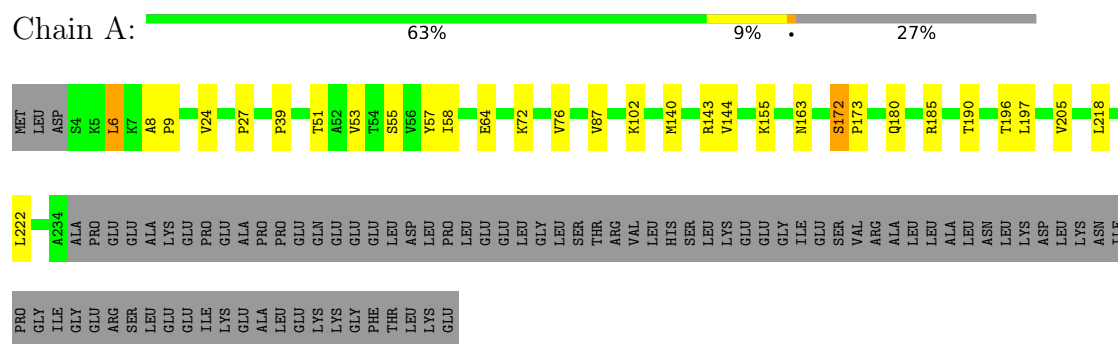


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
11	D	1	Total	C	N	O	P	0	1
			58	20	6	26	6		
11	N	1	Total	C	N	O	P	0	1
			58	20	6	26	6		

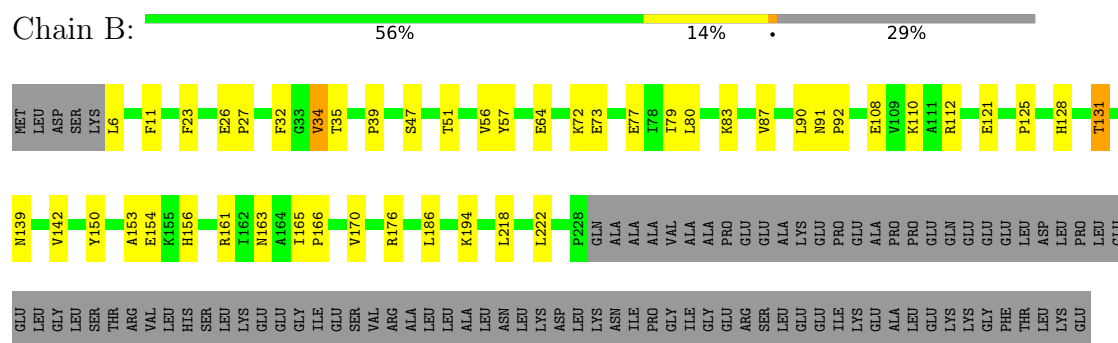
3 Residue-property plots [i](#)

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.

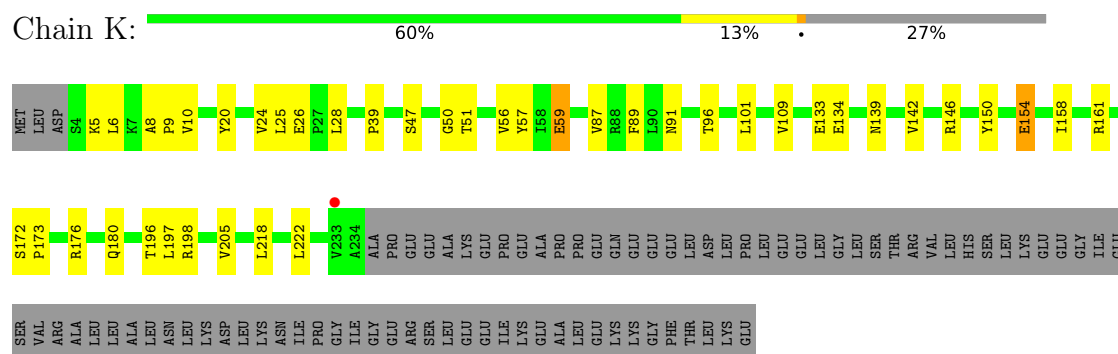
• Molecule 1: DNA-directed RNA polymerase subunit alpha



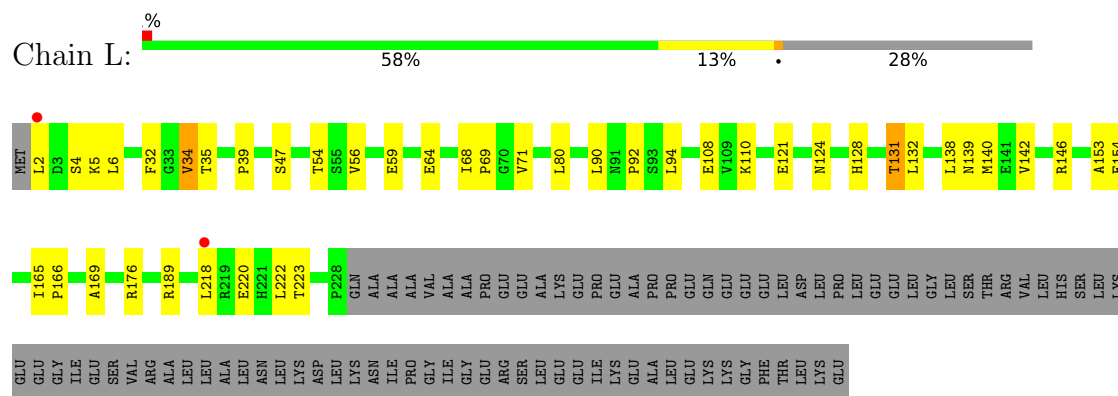
• Molecule 1: DNA-directed RNA polymerase subunit alpha



• Molecule 1: DNA-directed RNA polymerase subunit alpha



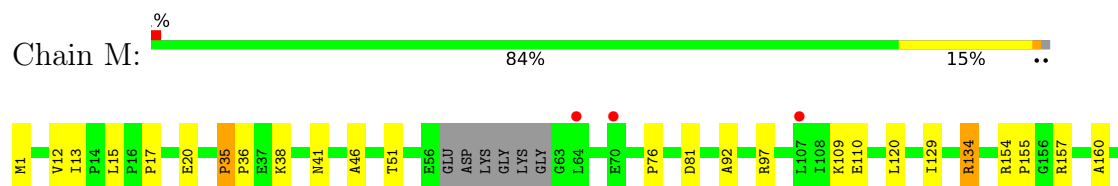
• Molecule 1: DNA-directed RNA polymerase subunit alpha

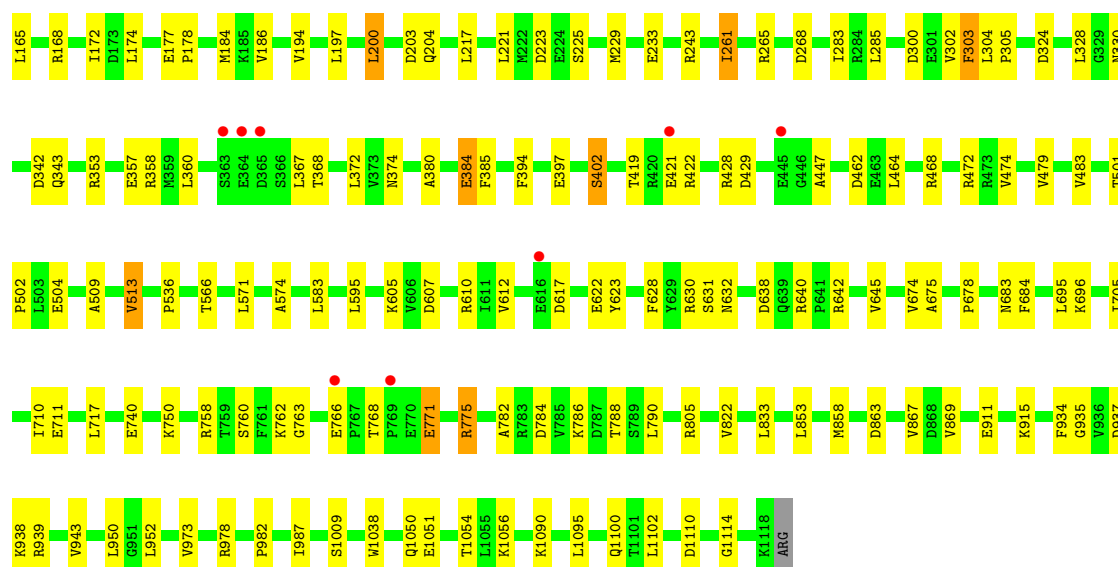


• Molecule 2: DNA-directed RNA polymerase subunit beta

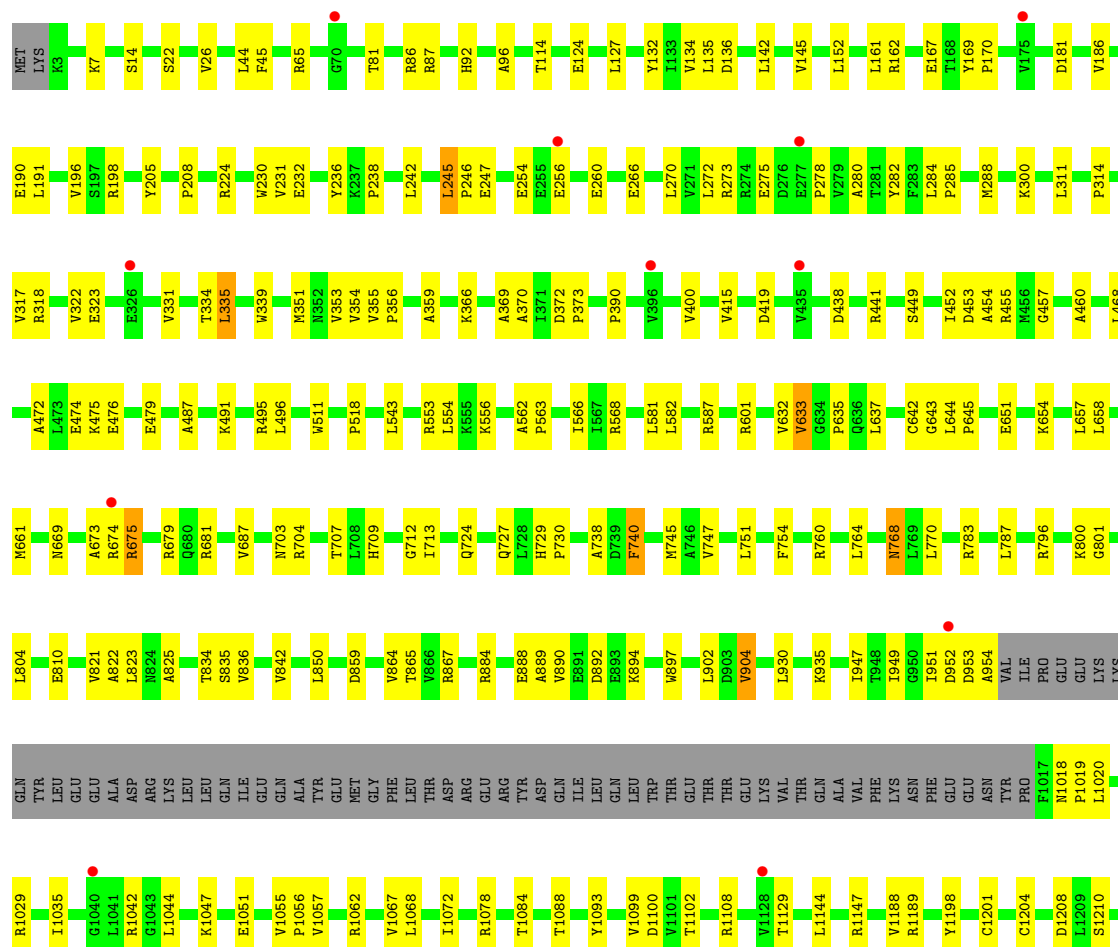
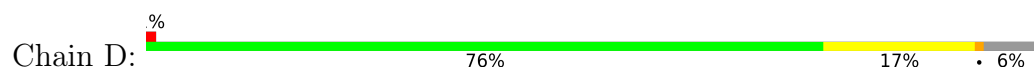


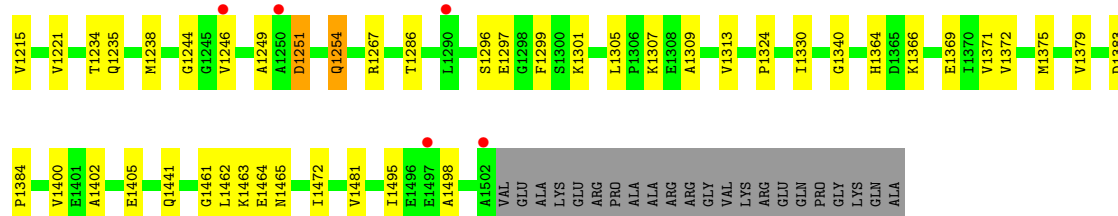
• Molecule 2: DNA-directed RNA polymerase subunit beta



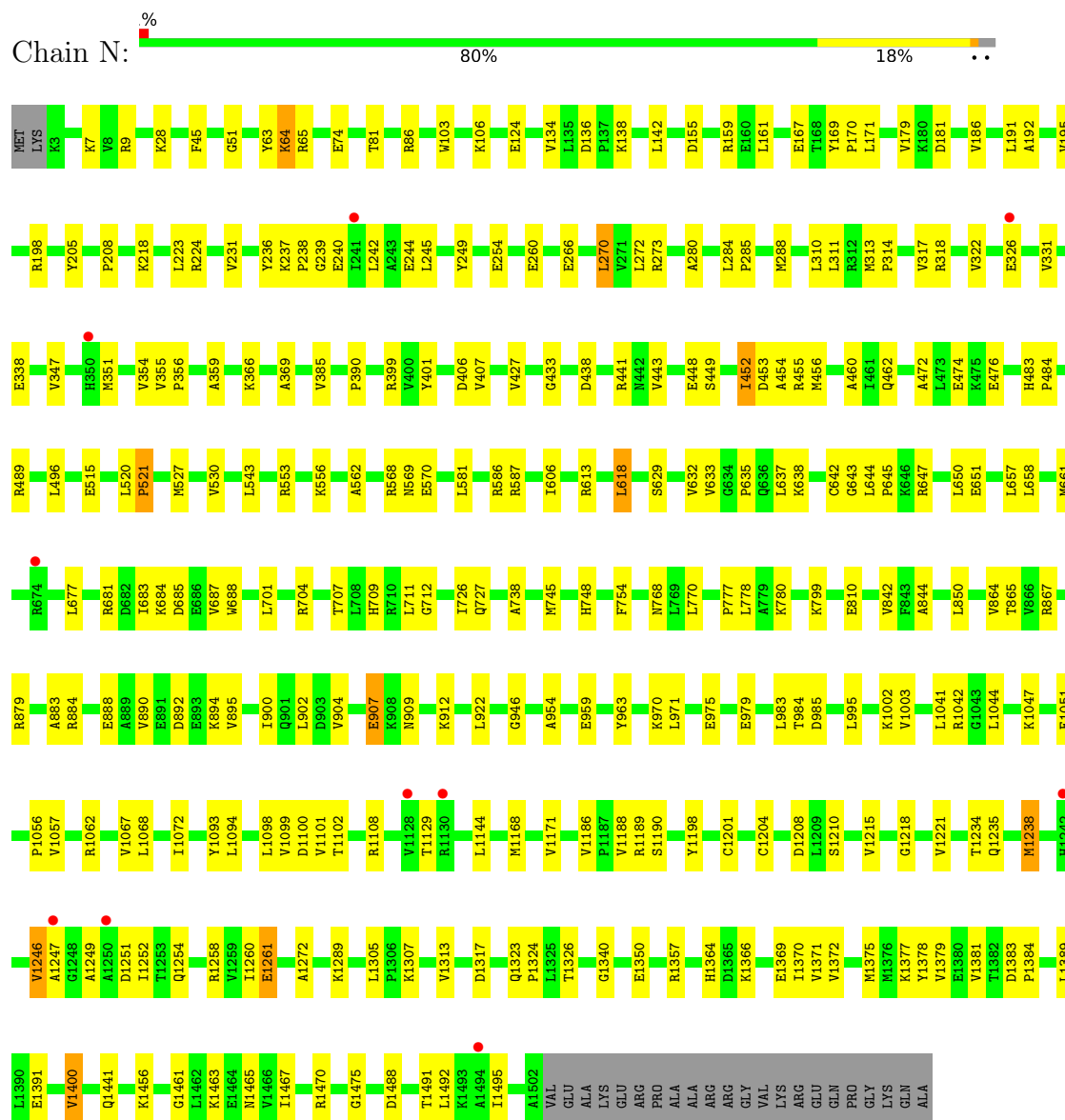


• Molecule 3: DNA-directed RNA polymerase subunit beta'





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

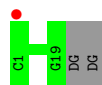
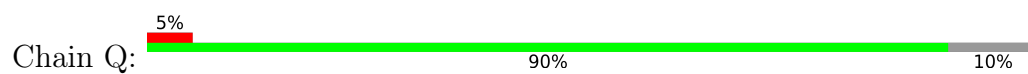


• Molecule 4: DNA-directed RNA polymerase subunit omega

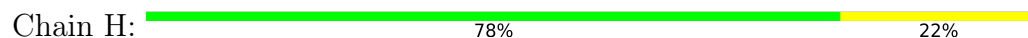
Chain E: 80% 15% 5%



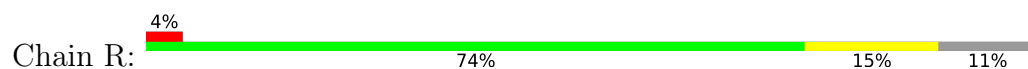
- Molecule 6: promoter DNA template strand



- Molecule 7: promoter DNA nontemplate strand



- Molecule 7: promoter DNA nontemplate strand



- Molecule 8: RNA (5'-R(*GP*A)-3')



- Molecule 8: RNA (5'-R(*GP*A)-3')



There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	186.41 Å 104.32 Å 297.29 Å 90.00° 98.36° 90.00°	Depositor
Resolution (Å)	49.02 – 3.35 49.02 – 3.35	Depositor EDS
% Data completeness (in resolution range)	89.3 (49.02-3.35) 90.6 (49.02-3.35)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.59 (at 3.33 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.208 , 0.250 0.207 , 0.248	Depositor DCC
R_{free} test set	7396 reflections (4.45%)	wwPDB-VP
Wilson B-factor (Å ²)	66.5	Xtriage
Anisotropy	0.043	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 23.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	56863	wwPDB-VP
Average B, all atoms (Å ²)	41.0	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 36.02 % 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 5.3307e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 2TM, ZN

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	0.28	0/1841	0.78	2/2504 (0.1%)
1	B	0.30	0/1790	0.80	1/2435 (0.0%)
1	K	0.28	0/1841	0.79	5/2504 (0.2%)
1	L	0.30	0/1821	0.80	1/2476 (0.0%)
2	C	0.29	0/8941	0.76	4/12092 (0.0%)
2	M	0.29	0/8941	0.76	3/12092 (0.0%)
3	D	0.30	0/11496	0.78	11/15543 (0.1%)
3	N	0.30	0/12043	0.78	13/16284 (0.1%)
4	E	0.26	0/772	0.71	0/1040
4	O	0.28	0/772	0.72	0/1040
5	F	0.29	0/2690	0.74	0/3618
5	P	0.28	0/2852	0.72	0/3837
6	G	0.09	0/432	0.48	0/665
6	Q	0.10	0/432	0.48	0/665
7	H	0.10	0/630	0.54	0/973
7	R	0.10	0/556	0.56	0/858
8	I	0.07	0/47	0.19	0/72
8	S	0.05	0/47	0.14	0/72
All	All	0.29	0/57944	0.76	40/78770 (0.1%)

There are no bond length outliers.

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	245	LEU	CA-C-N	7.45	127.12	119.82
3	D	245	LEU	C-N-CA	7.45	127.12	119.82
3	D	1340	GLY	CA-C-N	7.23	126.76	119.24
3	D	1340	GLY	C-N-CA	7.23	126.76	119.24
3	N	1340	GLY	CA-C-N	6.49	126.72	119.32

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	1809	0	1863	18	0
1	B	1758	0	1808	29	0
1	K	1809	0	1863	24	0
1	L	1789	0	1841	26	0
2	C	8774	0	8877	131	0
2	M	8774	0	8877	106	0
3	D	11299	0	11543	151	0
3	N	11834	0	12061	156	0
4	E	758	0	770	9	0
4	O	758	0	770	6	0
5	F	2649	0	2728	37	0
5	P	2807	0	2882	30	0
6	G	386	0	215	0	0
6	Q	386	0	215	0	0
7	H	560	0	305	5	0
7	R	495	0	272	3	0
8	I	42	0	23	1	0
8	S	42	0	23	0	0
9	B	1	0	0	0	0
9	D	5	0	0	0	0
9	F	1	0	0	0	0
9	L	1	0	0	0	0
9	N	5	0	0	0	0
9	P	1	0	0	0	0
10	D	2	0	0	0	0
10	N	2	0	0	0	0
11	D	58	0	28	0	0
11	N	58	0	28	0	0
All	All	56863	0	56992	656	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 6.

The worst 5 of 656 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:1495:ILE:HD13	4:E:80:VAL:HG21	1.69	0.75
2:M:168:ARG:HD3	2:M:268:ASP:HB3	1.70	0.73
2:C:194:VAL:HG22	2:C:221:LEU:HD12	1.70	0.72
3:D:1108:ARG:NH2	3:D:1198:TYR:O	2.22	0.72
2:C:165:LEU:HB2	2:C:168:ARG:HG3	1.71	0.72

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	229/315 (73%)	225 (98%)	4 (2%)	0	100	100
1	B	221/315 (70%)	218 (99%)	3 (1%)	0	100	100
1	K	229/315 (73%)	225 (98%)	4 (2%)	0	100	100
1	L	225/315 (71%)	222 (99%)	3 (1%)	0	100	100
2	C	1108/1119 (99%)	1081 (98%)	27 (2%)	0	100	100
2	M	1108/1119 (99%)	1078 (97%)	30 (3%)	0	100	100
3	D	1434/1524 (94%)	1378 (96%)	55 (4%)	1 (0%)	48	76
3	N	1498/1524 (98%)	1447 (97%)	51 (3%)	0	100	100
4	E	92/99 (93%)	90 (98%)	2 (2%)	0	100	100
4	O	92/99 (93%)	90 (98%)	2 (2%)	0	100	100
5	F	323/443 (73%)	317 (98%)	6 (2%)	0	100	100
5	P	344/443 (78%)	339 (98%)	5 (2%)	0	100	100
All	All	6903/7630 (90%)	6710 (97%)	192 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	904	VAL

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	200/273 (73%)	196 (98%)	4 (2%)	48	65
1	B	196/273 (72%)	190 (97%)	6 (3%)	35	59
1	K	200/273 (73%)	193 (96%)	7 (4%)	32	57
1	L	200/273 (73%)	195 (98%)	5 (2%)	42	61
2	C	936/941 (100%)	919 (98%)	17 (2%)	51	66
2	M	936/941 (100%)	920 (98%)	16 (2%)	53	67
3	D	1203/1279 (94%)	1168 (97%)	35 (3%)	37	60
3	N	1261/1279 (99%)	1230 (98%)	31 (2%)	42	61
4	E	82/88 (93%)	81 (99%)	1 (1%)	63	72
4	O	82/88 (93%)	81 (99%)	1 (1%)	63	72
5	F	285/387 (74%)	271 (95%)	14 (5%)	22	49
5	P	301/387 (78%)	290 (96%)	11 (4%)	30	55
All	All	5882/6482 (91%)	5734 (98%)	148 (2%)	42	61

5 of 148 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	632	VAL
5	P	310	ILE
3	N	768	ASN
3	N	1261	GLU
3	D	894	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
3	N	737	ASN
5	P	218	GLN
3	N	768	ASN
3	N	1323	GLN
3	D	1242	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	I	1/2 (50%)	0	0
8	S	1/2 (50%)	0	0
All	All	2/4 (50%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 22 ligands modelled in this entry, 18 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
11	2TM	D	2007[A]	9	26,30,30	1.59	6 (23%)	39,47,47	1.15	4 (10%)
11	2TM	N	2007[A]	9	26,30,30	1.87	10 (38%)	39,47,47	1.24	4 (10%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
11	2TM	N	2007[B]	9	26,30,30	1.87	10 (38%)	39,47,47	1.18	4 (10%)
11	2TM	D	2007[B]	9	26,30,30	1.59	6 (23%)	39,47,47	1.22	4 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
11	2TM	D	2007[A]	9	-	3/19/38/38	0/2/2/2
11	2TM	N	2007[A]	9	-	6/19/38/38	0/2/2/2
11	2TM	N	2007[B]	9	-	4/19/38/38	0/2/2/2
11	2TM	D	2007[B]	9	-	4/19/38/38	0/2/2/2

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	D	2007[A]	2TM	C4-N4	4.10	1.43	1.33
11	N	2007[A]	2TM	C4-N4	4.09	1.43	1.33
11	N	2007[B]	2TM	C4-N4	4.06	1.43	1.33
11	D	2007[B]	2TM	C4-N4	4.06	1.43	1.33
11	N	2007[A]	2TM	PA-O1A	3.05	1.58	1.51

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	N	2007[A]	2TM	PB-O3B-PG	-3.60	119.55	132.45
11	D	2007[B]	2TM	PB-O3B-PG	-3.51	119.86	132.45
11	N	2007[B]	2TM	PB-O3B-PG	-3.42	120.20	132.45
11	D	2007[A]	2TM	PB-O3B-PG	-3.29	120.67	132.45
11	N	2007[A]	2TM	O2-C2-N3	-2.94	117.70	122.33

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	D	2007[A]	2TM	PA-C1-PB-O3B
11	D	2007[A]	2TM	PA-C1-PB-O2B
11	D	2007[B]	2TM	C5'-O5'-PA-O2A
11	D	2007[B]	2TM	O4'-C4'-C5'-O5'

Continued on next page...

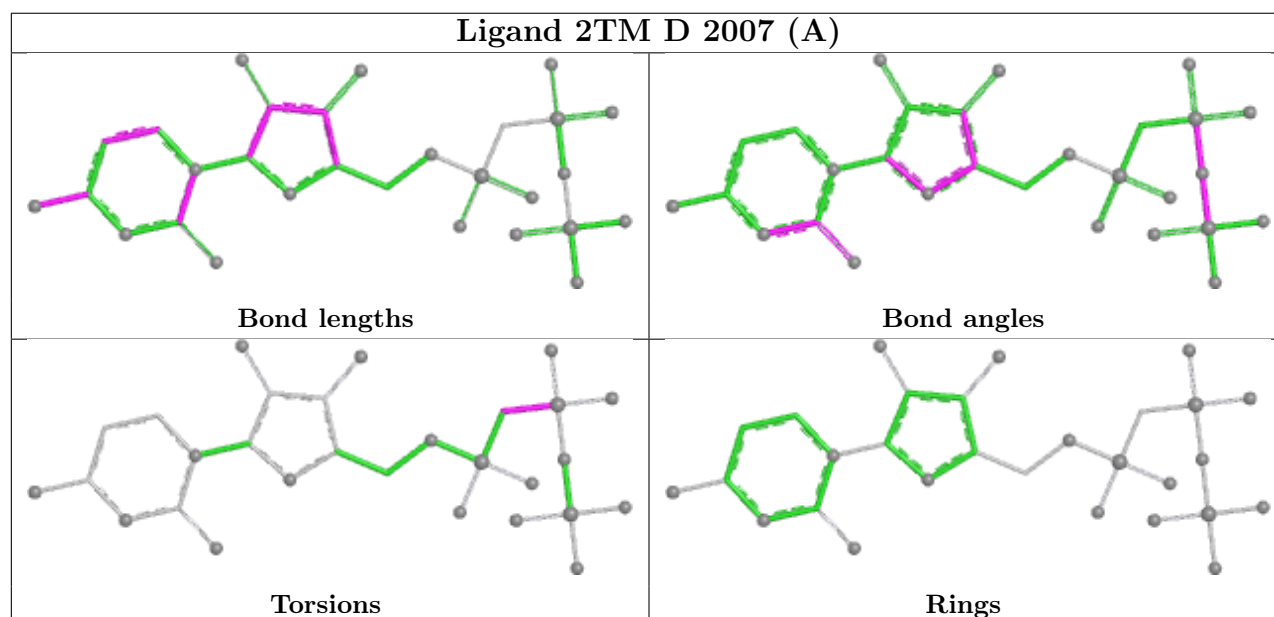
Continued from previous page...

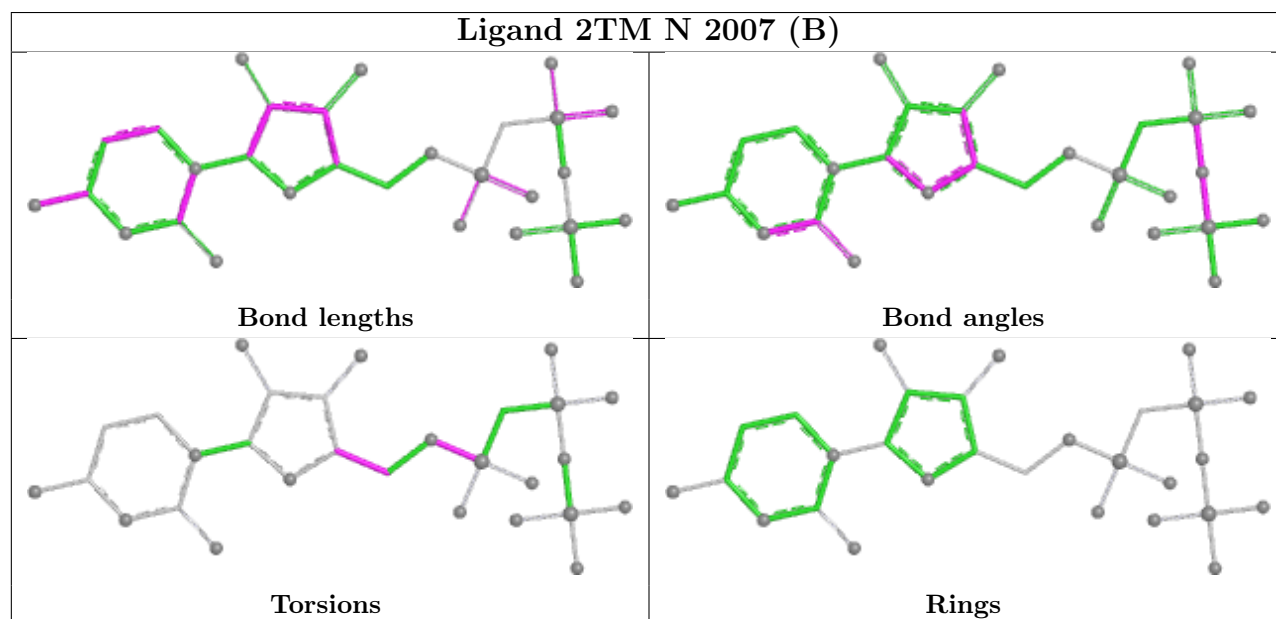
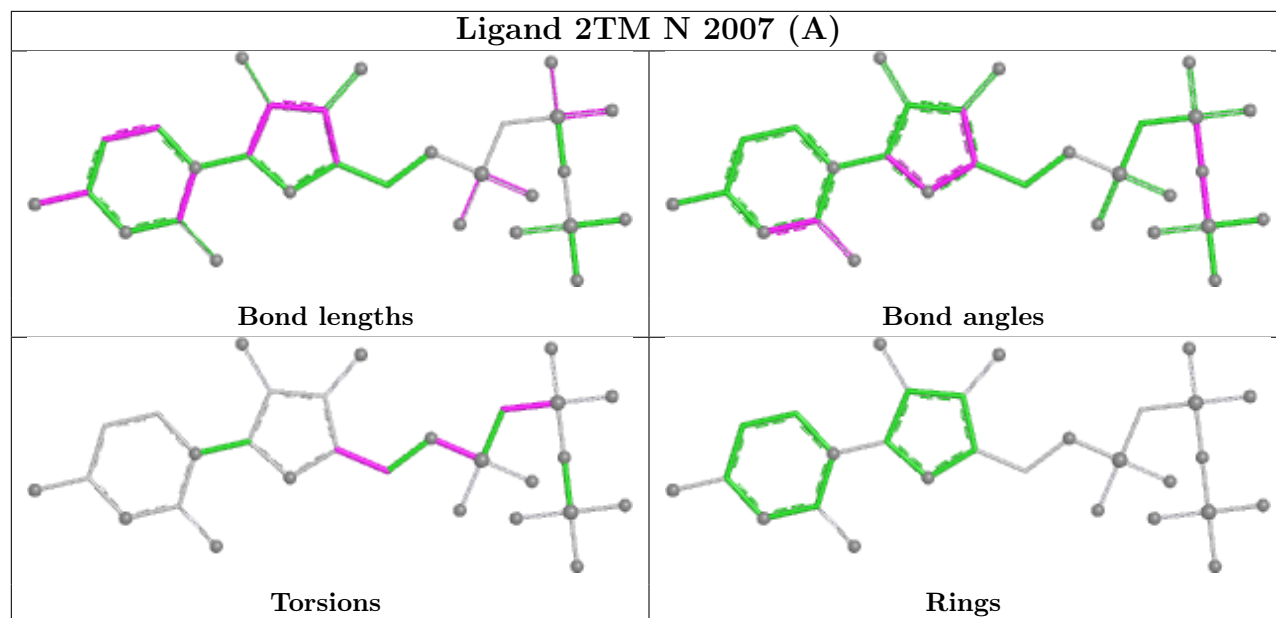
Mol	Chain	Res	Type	Atoms
11	N	2007[A]	2TM	C5'-O5'-PA-C1

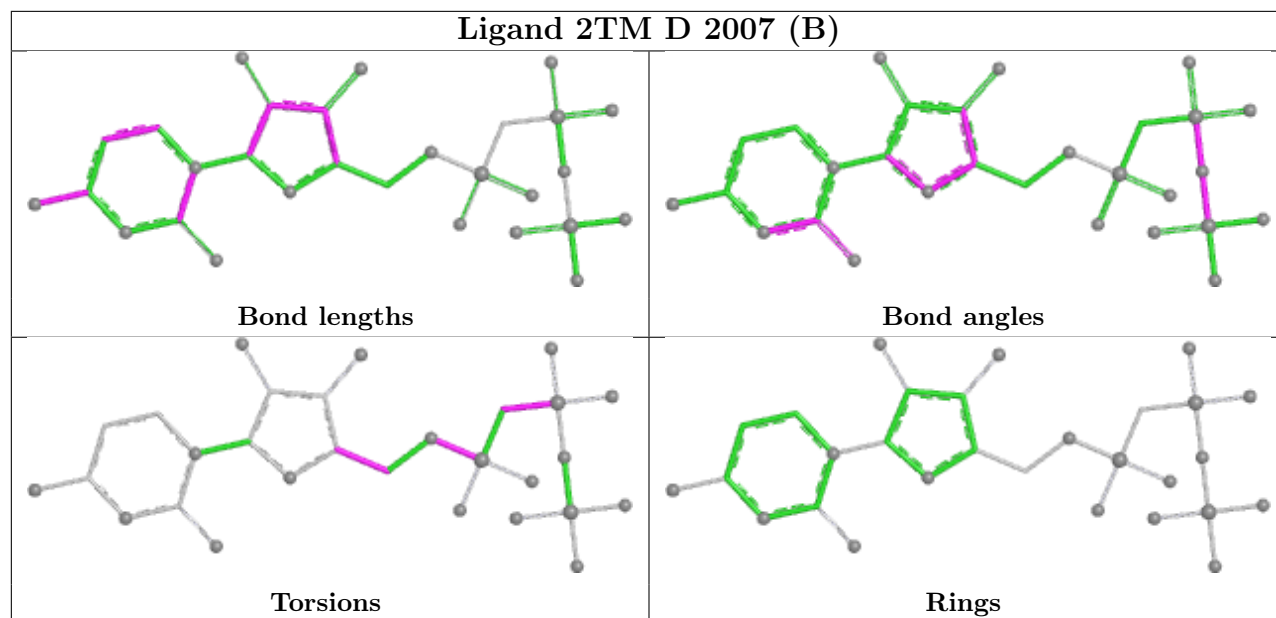
There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.







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 ⓘ

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	231/315 (73%)	-0.02	0 100 100	21, 42, 69, 126	0
1	B	223/315 (70%)	0.02	0 100 100	21, 46, 73, 86	0
1	K	231/315 (73%)	-0.06	1 (0%) 88 81	19, 37, 69, 113	0
1	L	227/315 (72%)	0.02	2 (0%) 81 69	19, 46, 76, 120	0
2	C	1112/1119 (99%)	0.11	17 (1%) 72 57	7, 41, 85, 111	0
2	M	1112/1119 (99%)	-0.17	11 (0%) 79 66	5, 28, 75, 113	0
3	D	1438/1524 (94%)	-0.05	16 (1%) 78 65	5, 33, 88, 133	2 (0%)
3	N	1500/1524 (98%)	-0.09	10 (0%) 84 73	5, 33, 84, 134	2 (0%)
4	E	94/99 (94%)	-0.16	0 100 100	14, 31, 71, 83	0
4	O	94/99 (94%)	-0.01	0 100 100	14, 31, 70, 74	0
5	F	327/443 (73%)	0.42	26 (7%) 18 15	25, 51, 133, 153	0
5	P	346/443 (78%)	-0.05	2 (0%) 85 75	12, 35, 79, 98	0
6	G	19/21 (90%)	0.10	0 100 100	12, 48, 153, 159	0
6	Q	19/21 (90%)	0.17	1 (5%) 32 23	9, 43, 154, 160	0
7	H	27/27 (100%)	0.30	0 100 100	39, 56, 166, 174	0
7	R	24/27 (88%)	0.19	1 (4%) 40 28	27, 51, 113, 130	0
8	I	2/2 (100%)	-0.67	0 100 100	17, 17, 17, 18	0
8	S	2/2 (100%)	-0.95	0 100 100	14, 14, 14, 15	0
All	All	7028/7730 (90%)	-0.02	87 (1%) 76 63	5, 37, 85, 174	4 (0%)

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	F	368	VAL	7.0
5	F	400	ILE	5.7
5	F	367	MET	5.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
5	F	404	ALA	4.6
5	F	408	LEU	4.5

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 [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. 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	RSCC	RSR	B-factors(Å ²)	Q<0.9
9	MG	N	2006	1/1	0.67	0.11	38,38,38,38	0
9	MG	B	2001	1/1	0.69	0.23	65,65,65,65	0
9	MG	L	2001	1/1	0.72	0.20	54,54,54,54	0
9	MG	D	2006	1/1	0.75	0.08	42,42,42,42	0
9	MG	F	2001	1/1	0.85	0.06	68,68,68,68	0
9	MG	N	2004[A]	1/1	0.88	0.18	25,25,25,25	1
9	MG	N	2004[B]	1/1	0.88	0.18	19,19,19,19	1
9	MG	D	2005	1/1	0.88	0.35	29,29,29,29	0
9	MG	N	2005	1/1	0.89	0.33	25,25,25,25	0
11	2TM	D	2007[A]	29/29	0.89	0.13	14,16,21,22	29
11	2TM	D	2007[B]	29/29	0.89	0.13	14,17,21,22	29
9	MG	P	2001	1/1	0.90	0.07	34,34,34,34	0
11	2TM	N	2007[A]	29/29	0.90	0.13	14,17,25,26	29
11	2TM	N	2007[B]	29/29	0.90	0.13	14,17,25,26	29
9	MG	N	2003	1/1	0.97	0.11	11,11,11,11	0
9	MG	D	2004[B]	1/1	0.98	0.11	18,18,18,18	1
9	MG	D	2003	1/1	0.98	0.04	13,13,13,13	0
9	MG	D	2004[A]	1/1	0.98	0.11	22,22,22,22	1
10	ZN	D	2002	1/1	0.99	0.04	73,73,73,73	0
10	ZN	N	2001	1/1	0.99	0.02	15,15,15,15	0
10	ZN	D	2001	1/1	1.00	0.01	10,10,10,10	0

Continued on next page...

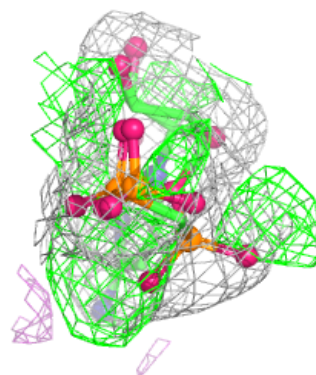
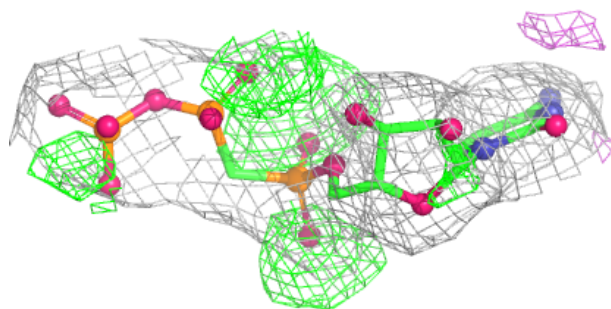
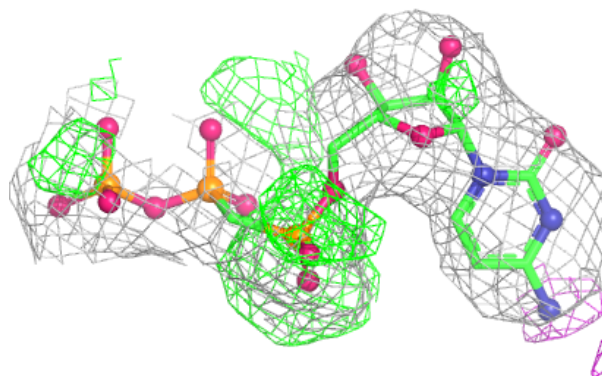
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
10	ZN	N	2002	1/1	1.00	0.01	25,25,25,25	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

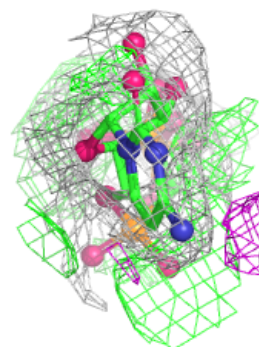
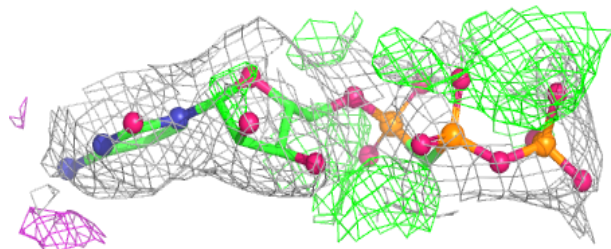
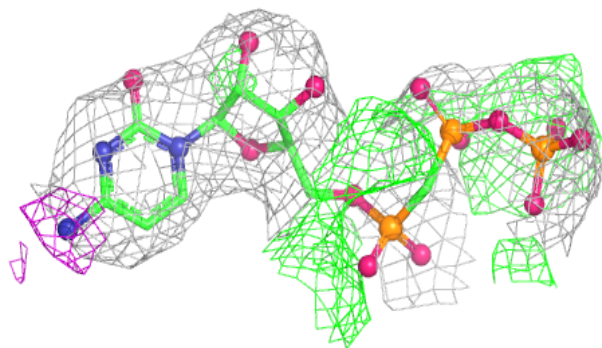
Electron density around 2TM D 2007 (A):

2mF_o-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

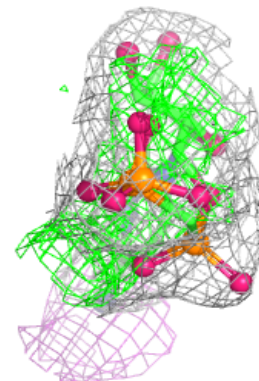
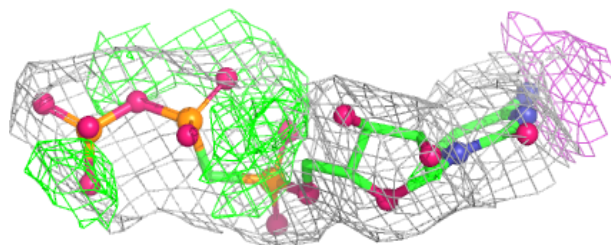
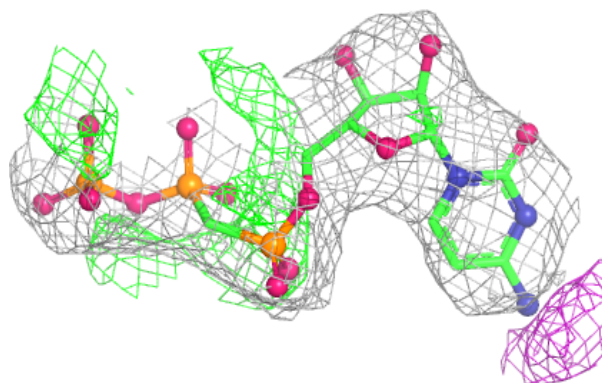


Electron density around 2TM D 2007 (B):

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

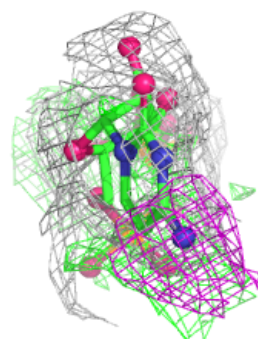
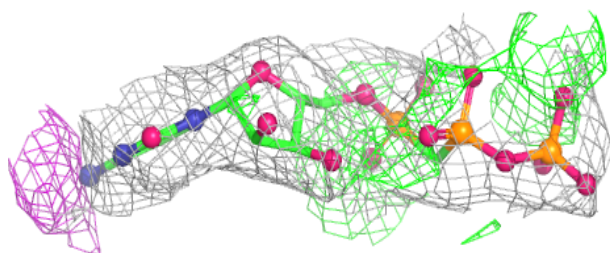
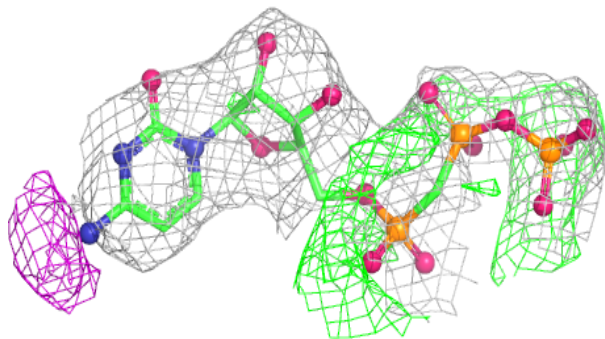
**Electron density around 2TM N 2007 (A):**

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around 2TM N 2007 (B):

$2mF_o - DF_c$ (at 0.7 rmsd) in gray
 $mF_o - DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers [i](#)

There are no such residues in this entry.