



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 21, 2023 – 12:13 PM EST

PDB ID : 5X22
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 : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : 2.36
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

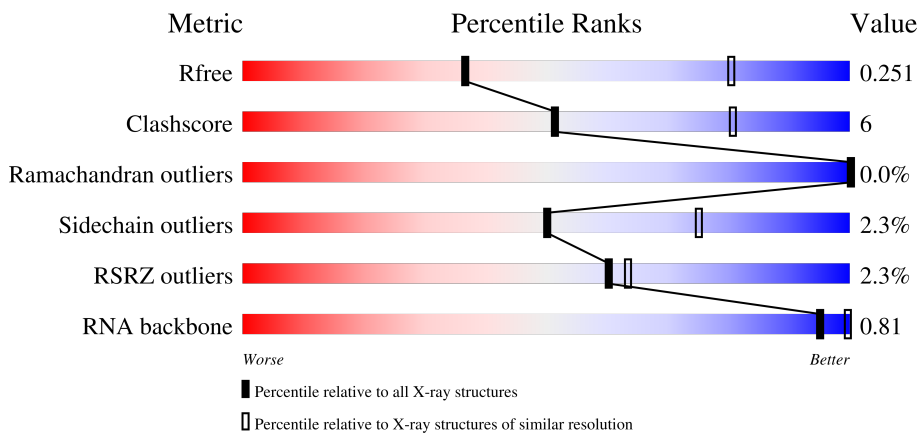
1 Overall quality at a glance

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}	130704	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)
RNA backbone	3102	1023 (3.80-2.92)

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	 2% 64% 9% 27%
1	B	315	 56% 14% 29%
1	K	315	 61% 12% 27%
1	L	315	 59% 12% 28%

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	

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
9	MG	B	2001	-	-	-	X

2 Entry composition [i](#)

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
			Total	C	N	O	S			
1	A	231	Total 1809	C 1155	N 315	O 337	S 2	0	0	0
1	B	223	Total 1758	C 1124	N 305	O 327	S 2	0	0	0
1	K	231	Total 1809	C 1155	N 315	O 337	S 2	0	0	0
1	L	227	Total 1789	C 1143	N 310	O 334	S 2	0	0	0

- 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	1112	Total 8774	C 5550	N 1565	O 1635	S 24	0	0	0
2	M	1112	Total 8774	C 5550	N 1565	O 1635	S 24	0	0	0

- 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	1438	Total 11299	C 7157	N 2001	O 2106	S 35	0	0	0
3	N	1500	Total 11834	C 7499	N 2086	O 2213	S 36	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	94	Total 758	C 483	N 132	O 139	S 4	0	0	0
4	O	94	Total 758	C 483	N 132	O 139	S 4	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
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 (three-letter code: 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...

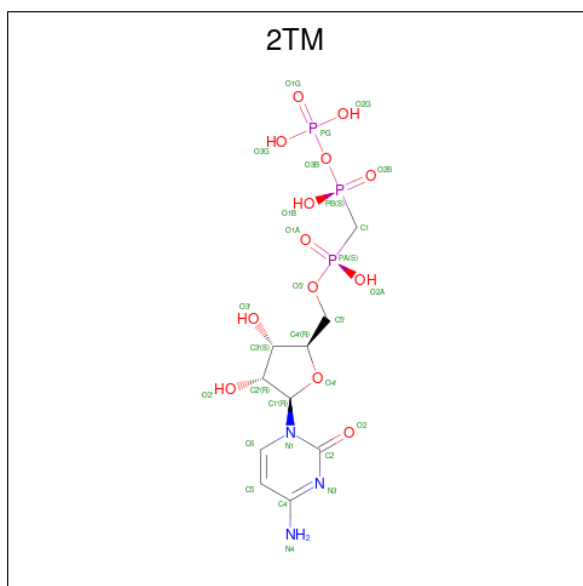
Continued from previous page...

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

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

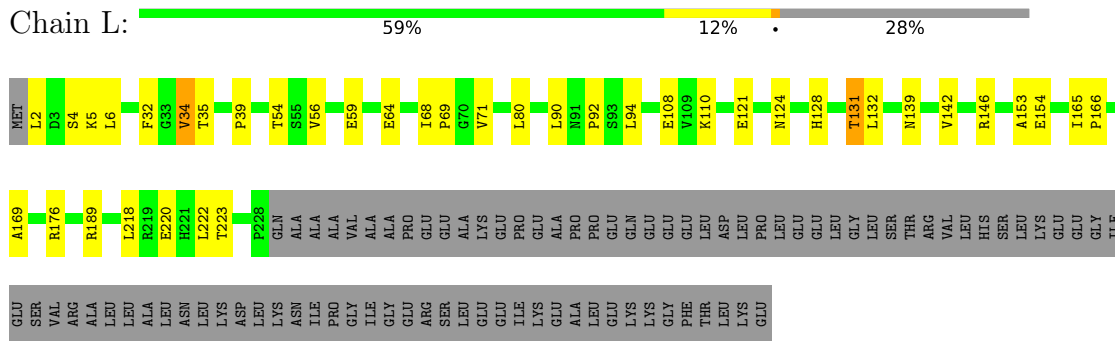
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	D	2	Total Zn 2 2	0	0
10	N	2	Total Zn 2 2	0	0

- Molecule 11 is 5'-O-[(S)-hydroxy{[(S)-hydroxy(phosphonoxy)phosphoryl]methyl}phosphoryl]cytidine (three-letter code: 2TM) (formula: C₁₀H₁₈N₃O₁₃P₃).

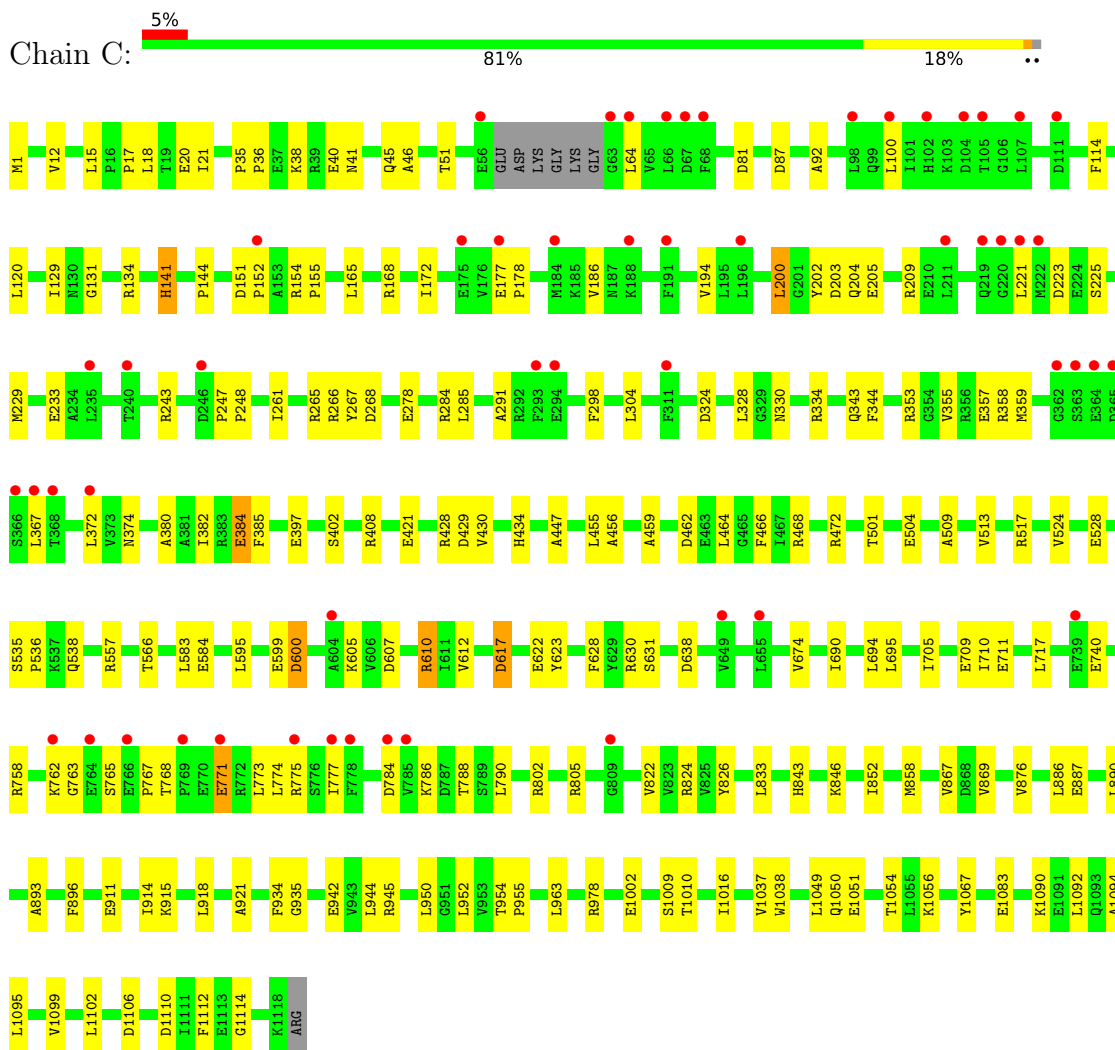


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	D	1	Total C N O P 58 20 6 26 6	0	1
11	N	1	Total C N O P 58 20 6 26 6	0	1

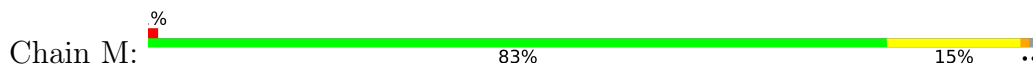
● Molecule 1: DNA-directed RNA polymerase subunit alpha

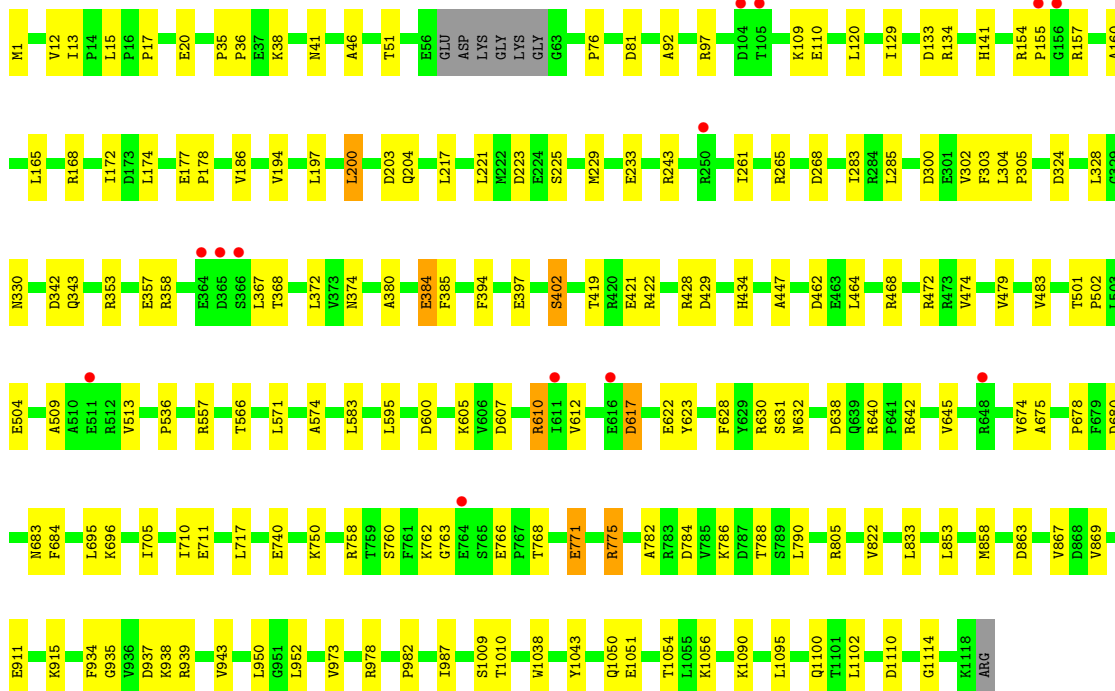


● Molecule 2: DNA-directed RNA polymerase subunit beta

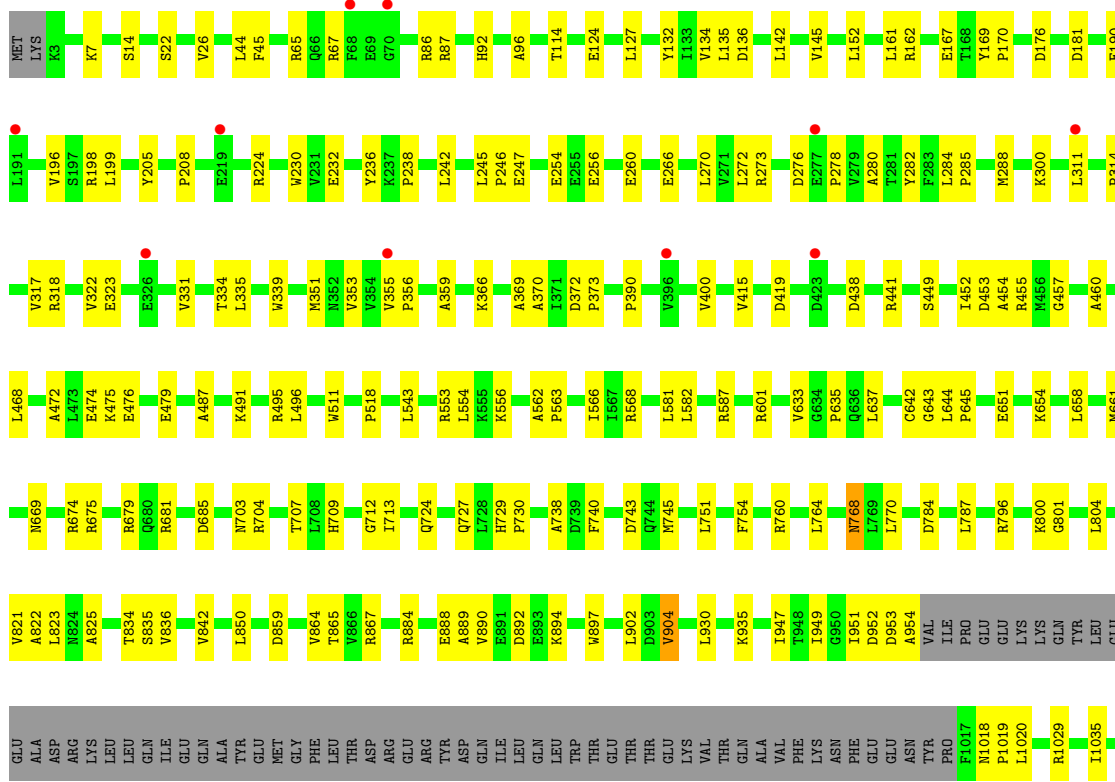
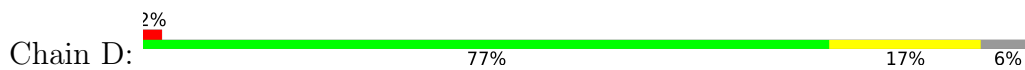


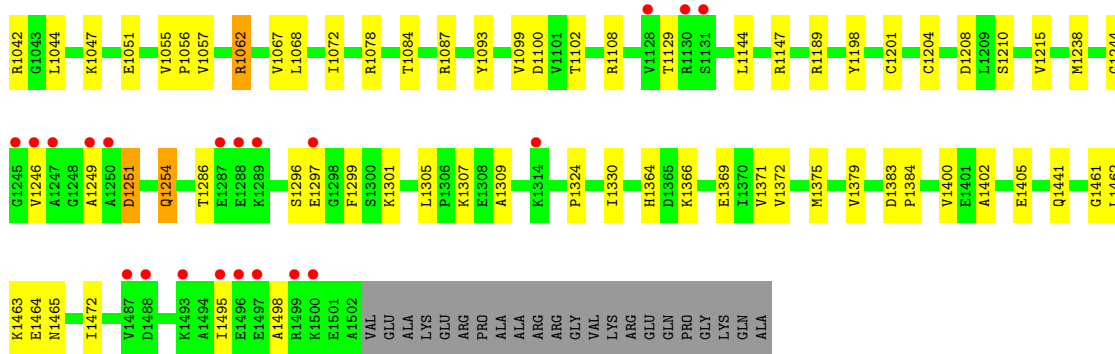
● Molecule 2: DNA-directed RNA polymerase subunit beta



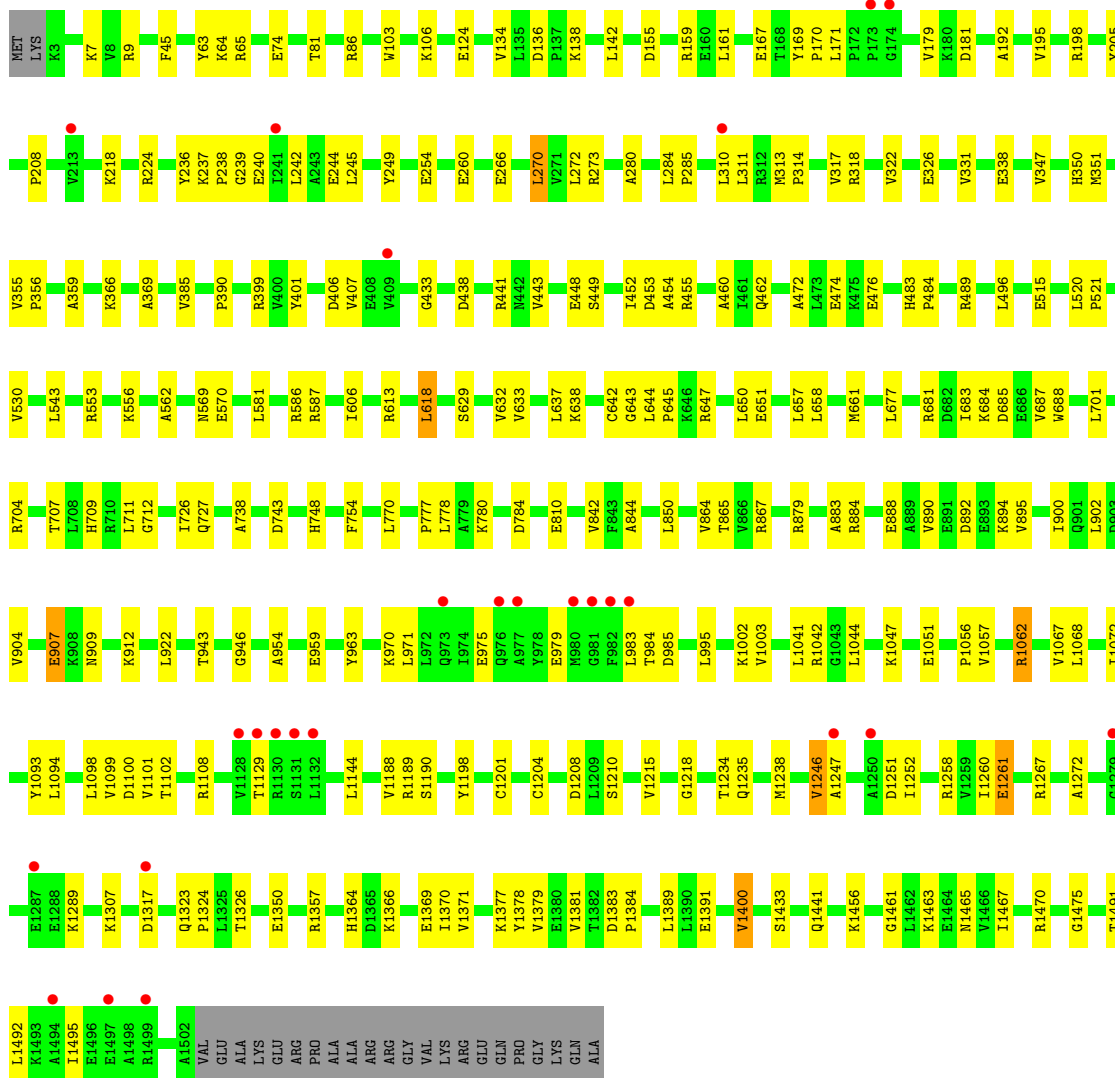
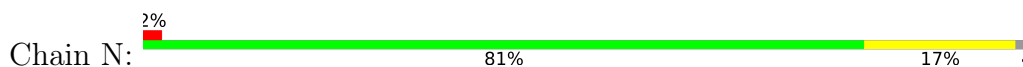


• Molecule 3: DNA-directed RNA polymerase subunit beta'

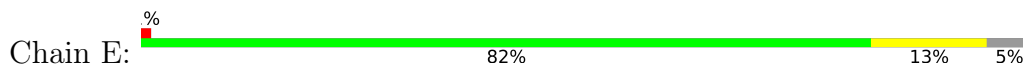




• Molecule 3: DNA-directed RNA polymerase subunit beta'

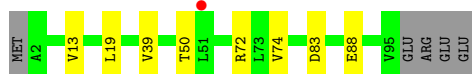
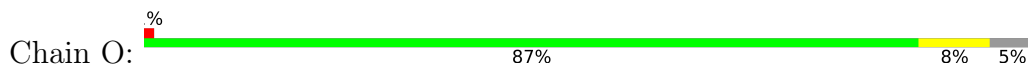


• Molecule 4: DNA-directed RNA polymerase subunit omega

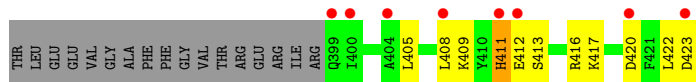
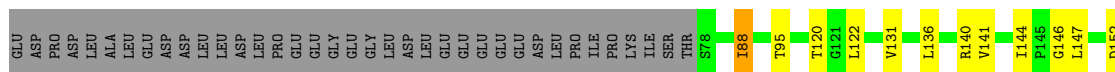
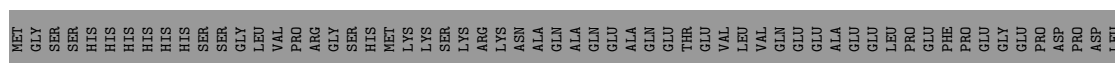




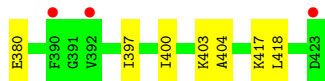
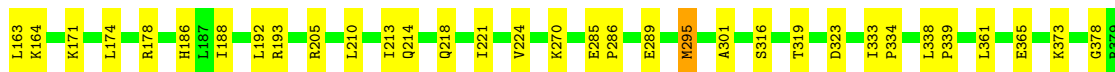
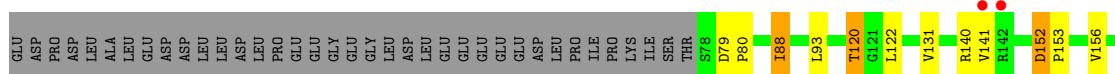
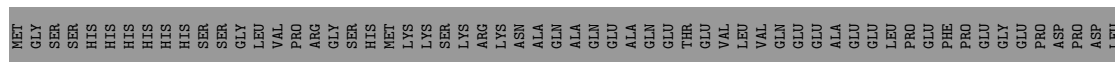
• Molecule 4: DNA-directed RNA polymerase subunit omega



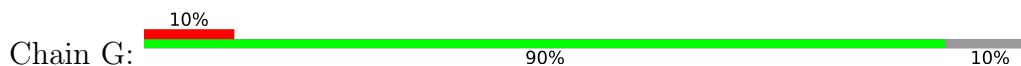
• Molecule 5: RNA polymerase sigma factor SigA

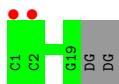


• Molecule 5: RNA polymerase sigma factor SigA

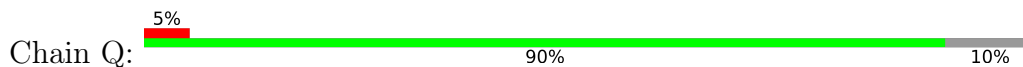


• Molecule 6: promoter DNA template strand

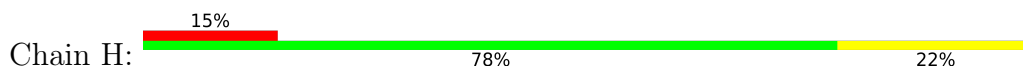




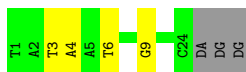
- Molecule 6: promoter DNA template strand



- 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	186.41Å 104.32Å 297.29Å 90.00° 98.36° 90.00°	Depositor
Resolution (Å)	49.02 – 3.35 49.25 – 3.33	Depositor EDS
% Data completeness (in resolution range)	89.3 (49.02-3.35) 89.4 (49.25-3.33)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.59 (at 3.33Å)	Xtrriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.208 , 0.250 0.208 , 0.251	Depositor DCC
R_{free} test set	7396 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	66.5	Xtrriage
Anisotropy	0.043	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 20.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	56863	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtrriage'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: 2TM, 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	0.22	0/1841	0.43	0/2504
1	B	0.24	0/1790	0.45	0/2435
1	K	0.22	0/1841	0.43	0/2504
1	L	0.24	0/1821	0.45	0/2476
2	C	0.22	0/8941	0.43	0/12092
2	M	0.22	0/8941	0.43	0/12092
3	D	0.22	0/11496	0.44	0/15543
3	N	0.23	0/12043	0.44	0/16284
4	E	0.21	0/772	0.39	0/1040
4	O	0.22	0/772	0.39	0/1040
5	F	0.23	0/2690	0.41	0/3618
5	P	0.22	0/2852	0.39	0/3837
6	G	0.44	0/432	0.97	0/665
6	Q	0.43	0/432	0.98	0/665
7	H	0.46	0/630	1.07	0/973
7	R	0.45	0/556	1.09	0/858
8	I	0.21	0/47	0.56	0/72
8	S	0.15	0/47	0.53	0/72
All	All	0.24	0/57944	0.47	0/78770

There are no bond length outliers.

There are no bond angle outliers.

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	23	0
1	L	1789	0	1841	25	0
2	C	8774	0	8877	127	0
2	M	8774	0	8877	106	0
3	D	11299	0	11543	146	0
3	N	11834	0	12061	149	0
4	E	758	0	770	8	0
4	O	758	0	770	5	0
5	F	2649	0	2728	35	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	633	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 633 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

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
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%)	51	82
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%)	197 (98%)	3 (2%)	65 82
1	B	196/273 (72%)	190 (97%)	6 (3%)	40 69
1	K	200/273 (73%)	194 (97%)	6 (3%)	41 69
1	L	200/273 (73%)	195 (98%)	5 (2%)	47 73
2	C	936/941 (100%)	922 (98%)	14 (2%)	65 82
2	M	936/941 (100%)	916 (98%)	20 (2%)	53 77
3	D	1203/1279 (94%)	1174 (98%)	29 (2%)	49 74
3	N	1261/1279 (99%)	1234 (98%)	27 (2%)	53 77
4	E	82/88 (93%)	81 (99%)	1 (1%)	71 85
4	O	82/88 (93%)	81 (99%)	1 (1%)	71 85
5	F	285/387 (74%)	273 (96%)	12 (4%)	30 60
5	P	301/387 (78%)	292 (97%)	9 (3%)	41 69
All	All	5882/6482 (91%)	5749 (98%)	133 (2%)	50 75

5 of 133 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	1235	GLN
3	N	1307	LYS
5	P	295	MET
3	D	1254	GLN
3	D	1251	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	L	38	ASN
3	N	640	HIS
5	P	214	GLN
3	N	569	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	N	737	ASN

5.3.3 RNA [i](#)

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 [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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[B]	9	27,30,30	1.59	6 (22%)	39,47,47	1.22	4 (10%)
11	2TM	N	2007[A]	9	27,30,30	1.86	8 (29%)	39,47,47	1.24	4 (10%)
11	2TM	D	2007[A]	9	27,30,30	1.59	6 (22%)	39,47,47	1.15	4 (10%)
11	2TM	N	2007[B]	9	27,30,30	1.85	10 (37%)	39,47,47	1.18	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[B]	9	-	4/19/38/38	0/2/2/2
11	2TM	N	2007[A]	9	-	6/19/38/38	0/2/2/2
11	2TM	D	2007[A]	9	-	2/19/38/38	0/2/2/2
11	2TM	N	2007[B]	9	-	4/19/38/38	0/2/2/2

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	D	2007[A]	2TM	C4-N4	4.21	1.43	1.33
11	N	2007[A]	2TM	C4-N4	4.21	1.43	1.33
11	N	2007[B]	2TM	C4-N4	4.18	1.43	1.33
11	D	2007[B]	2TM	C4-N4	4.18	1.43	1.33
11	N	2007[A]	2TM	PA-O1A	2.98	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	-3.71	119.55	132.62
11	D	2007[B]	2TM	PB-O3B-PG	-3.62	119.86	132.62
11	N	2007[B]	2TM	PB-O3B-PG	-3.52	120.20	132.62
11	D	2007[A]	2TM	PB-O3B-PG	-3.39	120.67	132.62
11	N	2007[A]	2TM	O2-C2-N3	-2.85	117.70	122.33

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

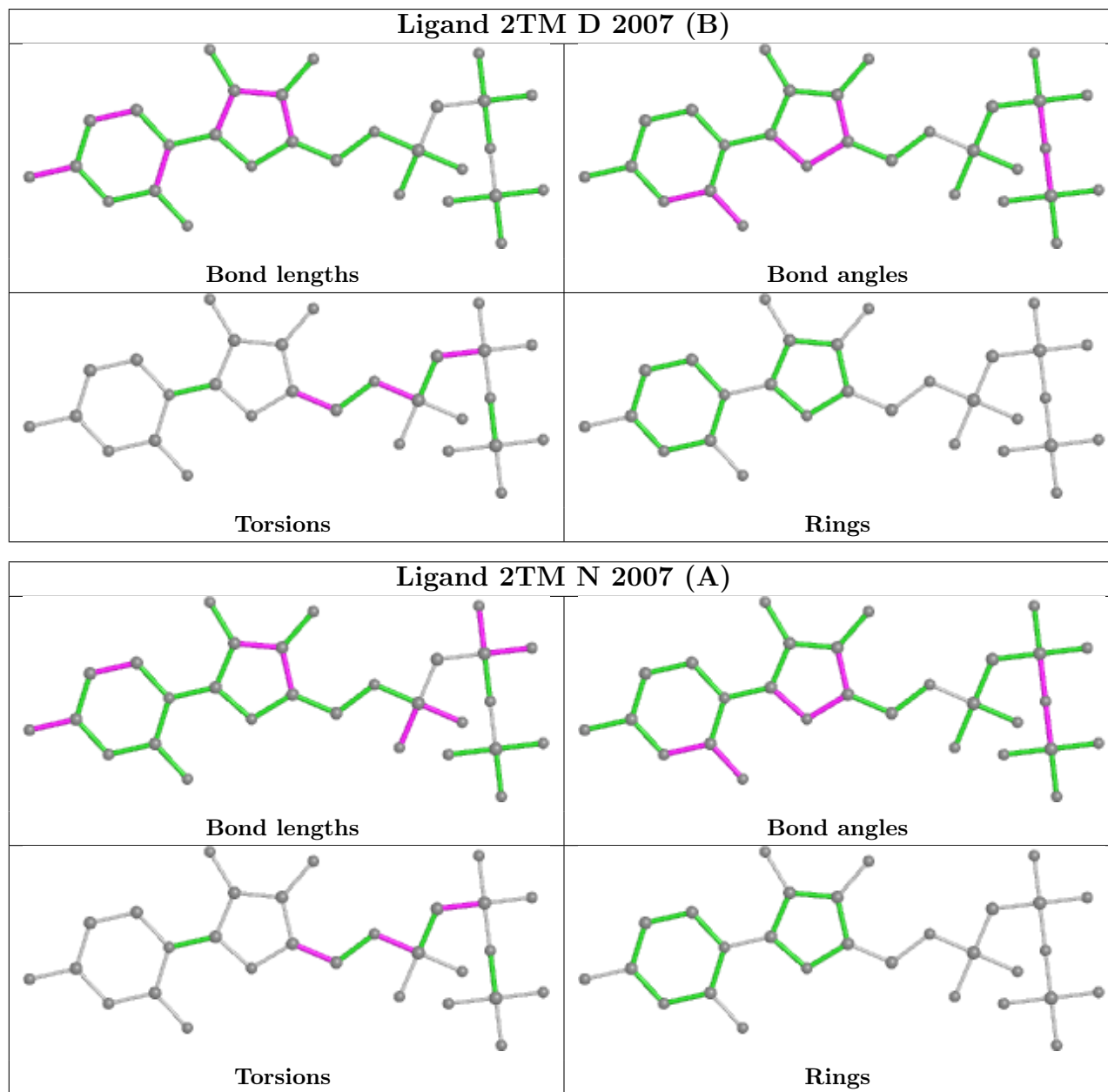
Mol	Chain	Res	Type	Atoms
11	D	2007[A]	2TM	PA-C1-PB-O3B
11	D	2007[A]	2TM	PA-C1-PB-O2B
11	D	2007[B]	2TM	C5'-O5'-PA-O2A
11	D	2007[B]	2TM	O4'-C4'-C5'-O5'
11	D	2007[B]	2TM	C3'-C4'-C5'-O5'

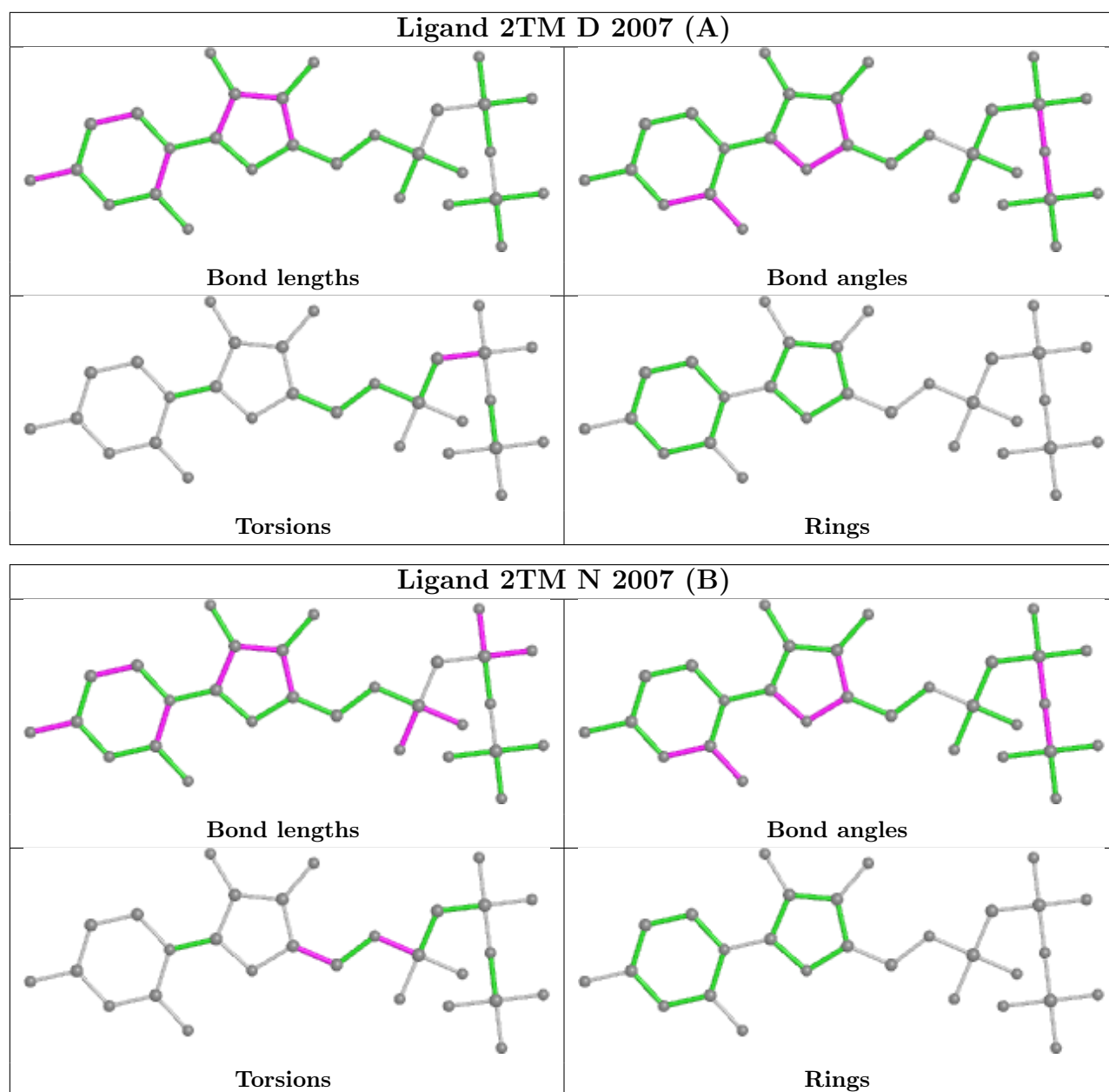
There are no ring outliers.

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%)	-0.05	5 (2%) 62 65	21, 42, 69, 126	0
1	B	223/315 (70%)	-0.20	0 100 100	21, 46, 73, 86	0
1	K	231/315 (73%)	-0.11	3 (1%) 77 80	19, 37, 69, 113	0
1	L	227/315 (72%)	-0.11	0 100 100	19, 46, 76, 120	0
2	C	1112/1119 (99%)	0.06	54 (4%) 29 31	7, 41, 85, 111	0
2	M	1112/1119 (99%)	-0.15	13 (1%) 79 82	5, 28, 75, 113	0
3	D	1438/1524 (94%)	-0.07	31 (2%) 62 65	5, 33, 88, 133	2 (0%)
3	N	1500/1524 (98%)	-0.08	26 (1%) 70 73	5, 33, 84, 134	2 (0%)
4	E	94/99 (94%)	0.04	1 (1%) 80 84	14, 31, 71, 83	0
4	O	94/99 (94%)	0.01	1 (1%) 80 84	14, 31, 70, 74	0
5	F	327/443 (73%)	0.06	14 (4%) 35 38	25, 51, 133, 153	0
5	P	346/443 (78%)	-0.11	5 (1%) 75 78	12, 35, 79, 98	0
6	G	19/21 (90%)	-0.01	2 (10%) 6 7	12, 48, 153, 159	0
6	Q	19/21 (90%)	-0.03	1 (5%) 26 28	9, 43, 154, 160	0
7	H	27/27 (100%)	0.13	4 (14%) 2 2	39, 56, 166, 174	0
7	R	24/27 (88%)	-0.25	0 100 100	27, 51, 113, 130	0
8	I	2/2 (100%)	-0.08	0 100 100	17, 17, 17, 18	0
8	S	2/2 (100%)	-0.25	0 100 100	14, 14, 14, 15	0
All	All	7028/7730 (90%)	-0.06	160 (2%) 60 63	5, 37, 85, 174	4 (0%)

The worst 5 of 160 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	F	400	ILE	6.1
2	M	365	ASP	5.1
3	D	1497	GLU	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	K	232	ALA	4.8
3	D	1130	ARG	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 monosaccharides 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.

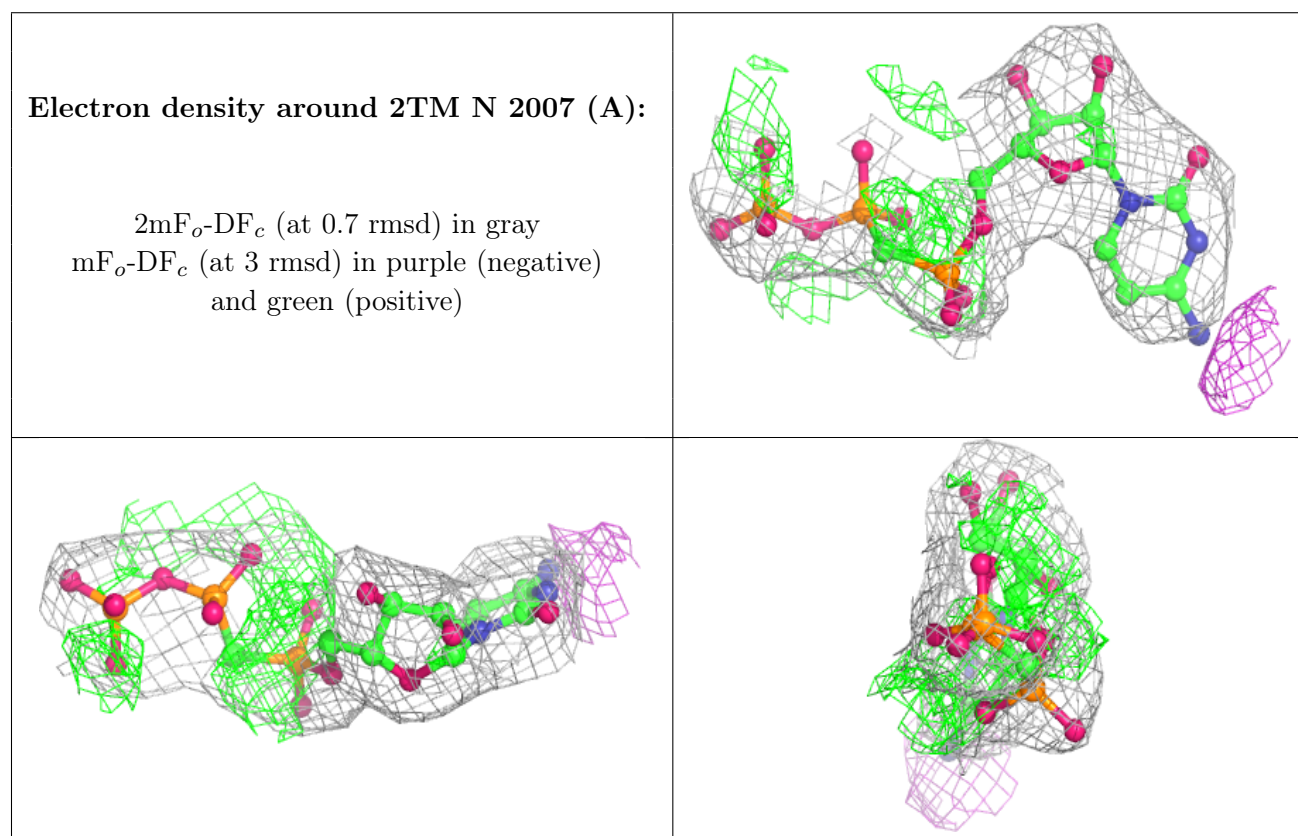
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
9	MG	B	2001	1/1	0.45	0.44	65,65,65,65	0
9	MG	L	2001	1/1	0.63	0.35	54,54,54,54	0
9	MG	N	2006	1/1	0.74	0.24	38,38,38,38	0
9	MG	D	2006	1/1	0.78	0.16	42,42,42,42	0
9	MG	F	2001	1/1	0.81	0.23	68,68,68,68	0
9	MG	N	2003	1/1	0.88	0.21	11,11,11,11	0
11	2TM	N	2007[A]	29/29	0.89	0.34	14,17,25,26	29
11	2TM	N	2007[B]	29/29	0.89	0.34	14,17,25,26	29
9	MG	P	2001	1/1	0.90	0.21	34,34,34,34	0
11	2TM	D	2007[A]	29/29	0.91	0.30	14,16,21,22	29
11	2TM	D	2007[B]	29/29	0.91	0.30	14,17,21,22	29
9	MG	D	2005	1/1	0.93	0.43	29,29,29,29	0
9	MG	N	2005	1/1	0.95	0.44	25,25,25,25	0
9	MG	D	2004[A]	1/1	0.96	0.35	22,22,22,22	1
10	ZN	D	2002	1/1	0.96	0.06	73,73,73,73	0
9	MG	N	2004[A]	1/1	0.96	0.45	25,25,25,25	1
9	MG	N	2004[B]	1/1	0.96	0.45	19,19,19,19	1
9	MG	D	2004[B]	1/1	0.96	0.35	18,18,18,18	1
9	MG	D	2003	1/1	0.96	0.15	13,13,13,13	0
10	ZN	N	2001	1/1	0.99	0.12	15,15,15,15	0
10	ZN	N	2002	1/1	0.99	0.10	25,25,25,25	0

Continued on next page...

Continued from previous page...

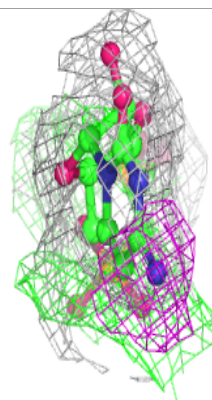
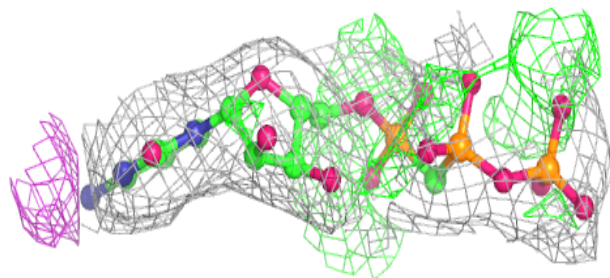
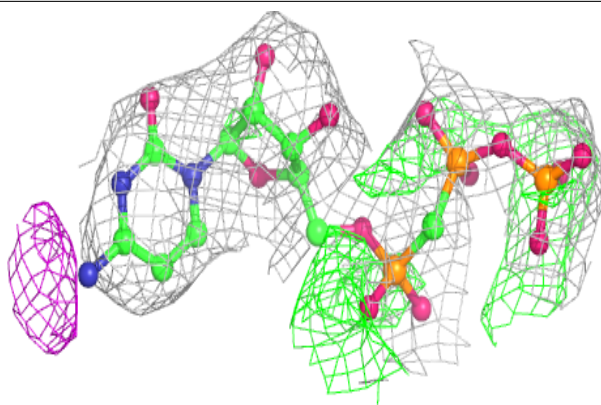
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
10	ZN	D	2001	1/1	1.00	0.15	10,10,10,10	0

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.

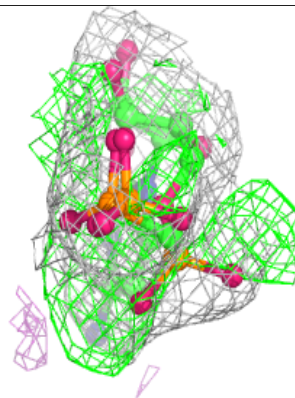
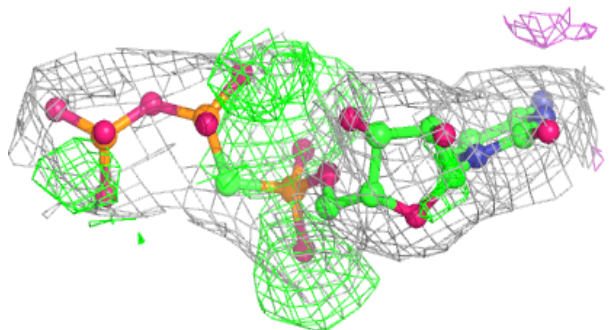
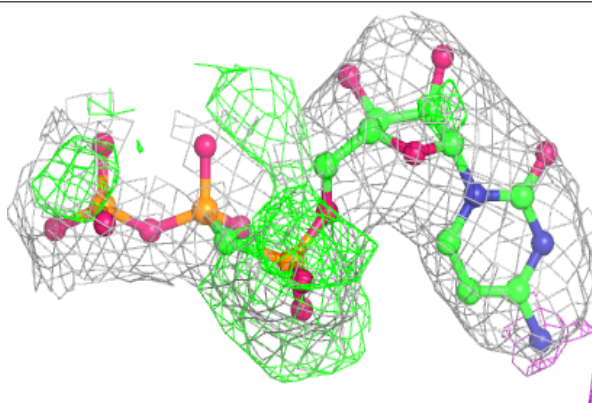


Electron density around 2TM N 2007 (B):

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

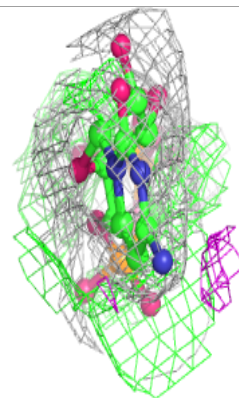
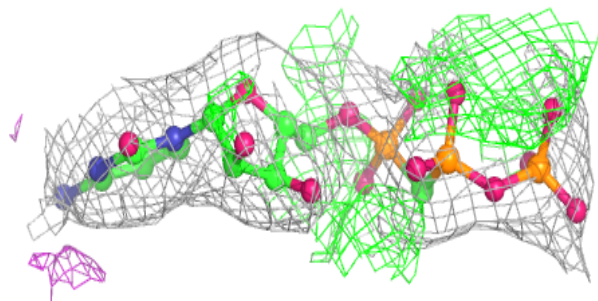
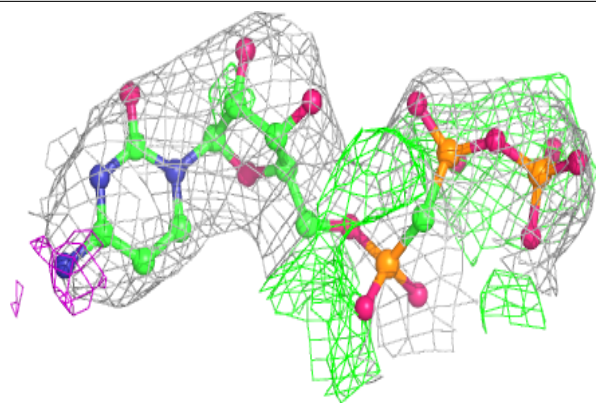
**Electron density around 2TM D 2007 (A):**

$2mF_o-DF_c$ (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)



6.5 Other polymers [i](#)

There are no such residues in this entry.