



wwPDB X-ray Structure Validation Summary Report ⓘ

May 20, 2024 – 07:06 PM EDT

PDB ID : 5D4D
Title : Crystal structure of Thermus thermophilus product complex for transcription initiation with NAD and CTP
Authors : Zhang, Y.; Ebright, R.H.
Deposited on : 2015-08-07
Resolution : 3.00 Å(reported)

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

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

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)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
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)	:	2.36.2

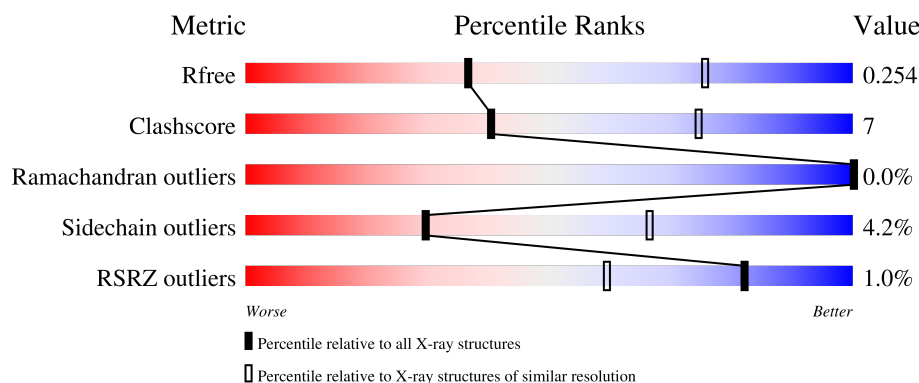
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	
2	C	1119	

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	R	19	
7	H	27	
7	S	27	

2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 57349 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	229	Total	C	N	O	S	0	0	0
			1797	1147	313	335	2			
1	B	222	Total	C	N	O	S	0	0	0
			1750	1120	303	325	2			
1	K	228	Total	C	N	O	S	0	0	0
			1792	1144	312	334	2			
1	L	222	Total	C	N	O	S	0	0	0
			1750	1120	303	325	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1108	Total	C	N	O	S	0	0	0
			8747	5536	1561	1626	24			
2	M	1091	Total	C	N	O	S	0	0	0
			8611	5449	1539	1600	23			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1485	Total	C	N	O	S	0	0	0
			11730	7435	2066	2194	35			
3	N	1483	Total	C	N	O	S	0	0	0
			11716	7427	2064	2190	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
			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	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			
5	P	316	Total	C	N	O	S	0	0	0
			2574	1624	466	480	4			

There are 42 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
F	46	THR	ALA	conflict	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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	-6	LEU	-	expression tag	UNP Q72L95
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
P	46	THR	ALA	conflict	UNP Q72L95

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	16	Total	C	N	O	P	0	0	0
			328	157	62	94	15			
6	R	16	Total	C	N	O	P	0	0	0
			328	157	62	94	15			

- Molecule 7 is a DNA chain called DNA (27-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	21	Total	C	N	O	P	0	0	0
			435	207	87	121	20			
7	S	21	Total	C	N	O	P	0	0	0
			434	206	87	121	20			

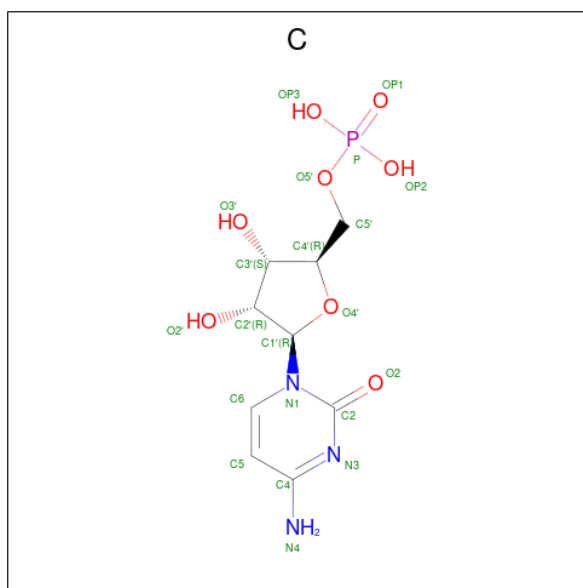
- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

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

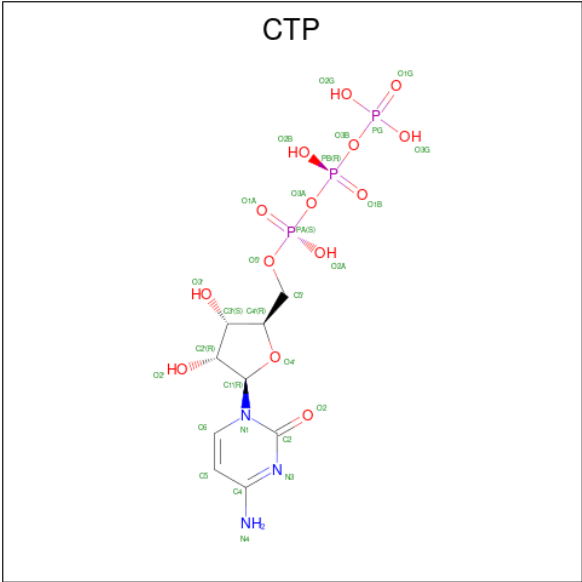
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Zn	0	0
			2	2		
9	N	2	Total	Zn	0	0
			2	2		

- Molecule 10 is CYTIDINE-5'-MONOPHOSPHATE (three-letter code: C) (formula: C₉H₁₄N₃O₈P).



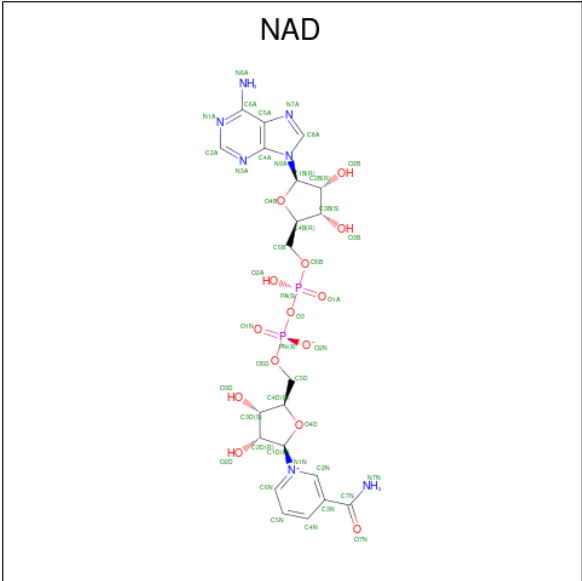
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
10	D	1	Total	C	N	O	P	0	0
			20	9	3	7	1		

- Molecule 11 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: C₉H₁₆N₃O₁₄P₃).



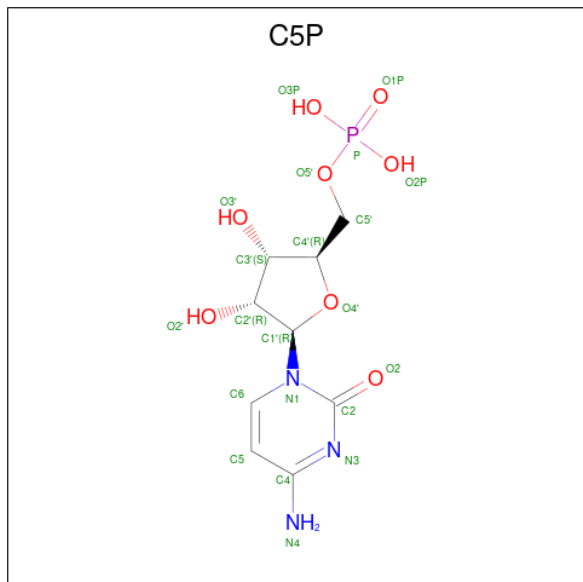
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	D	1	Total	O	P		0	0
			9	7	2			
11	M	1	Total	O	P		0	0
			9	7	2			

- Molecule 12 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C₂₁H₂₇N₇O₁₄P₂).



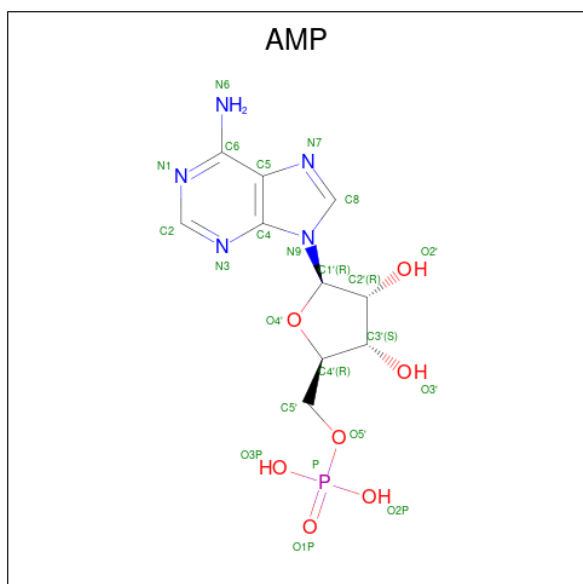
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
12	G	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 13 is CYTIDINE-5'-MONOPHOSPHATE (three-letter code: C5P) (formula: $C_9H_{14}N_3O_8P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
13	N	1	Total	C	N	O	P	0	0
			20	9	3	7	1		

- Molecule 14 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: $C_{10}H_{14}N_5O_7P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
14	R	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

- Molecule 15 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	23	Total 23	O 23	0	0
15	B	16	Total 16	O 16	0	0
15	C	196	Total 196	O 196	0	0
15	D	243	Total 243	O 243	0	0
15	E	16	Total 16	O 16	0	0
15	F	45	Total 45	O 45	0	0
15	G	9	Total 9	O 9	0	0
15	H	4	Total 4	O 4	0	0
15	K	14	Total 14	O 14	0	0
15	L	13	Total 13	O 13	0	0
15	M	97	Total 97	O 97	0	0
15	N	187	Total 187	O 187	0	0
15	O	12	Total 12	O 12	0	0
15	P	14	Total 14	O 14	0	0
15	R	3	Total 3	O 3	0	0
15	S	3	Total 3	O 3	0	0


[illegible]

- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain L: 52% 17% . 30%

[illegible]

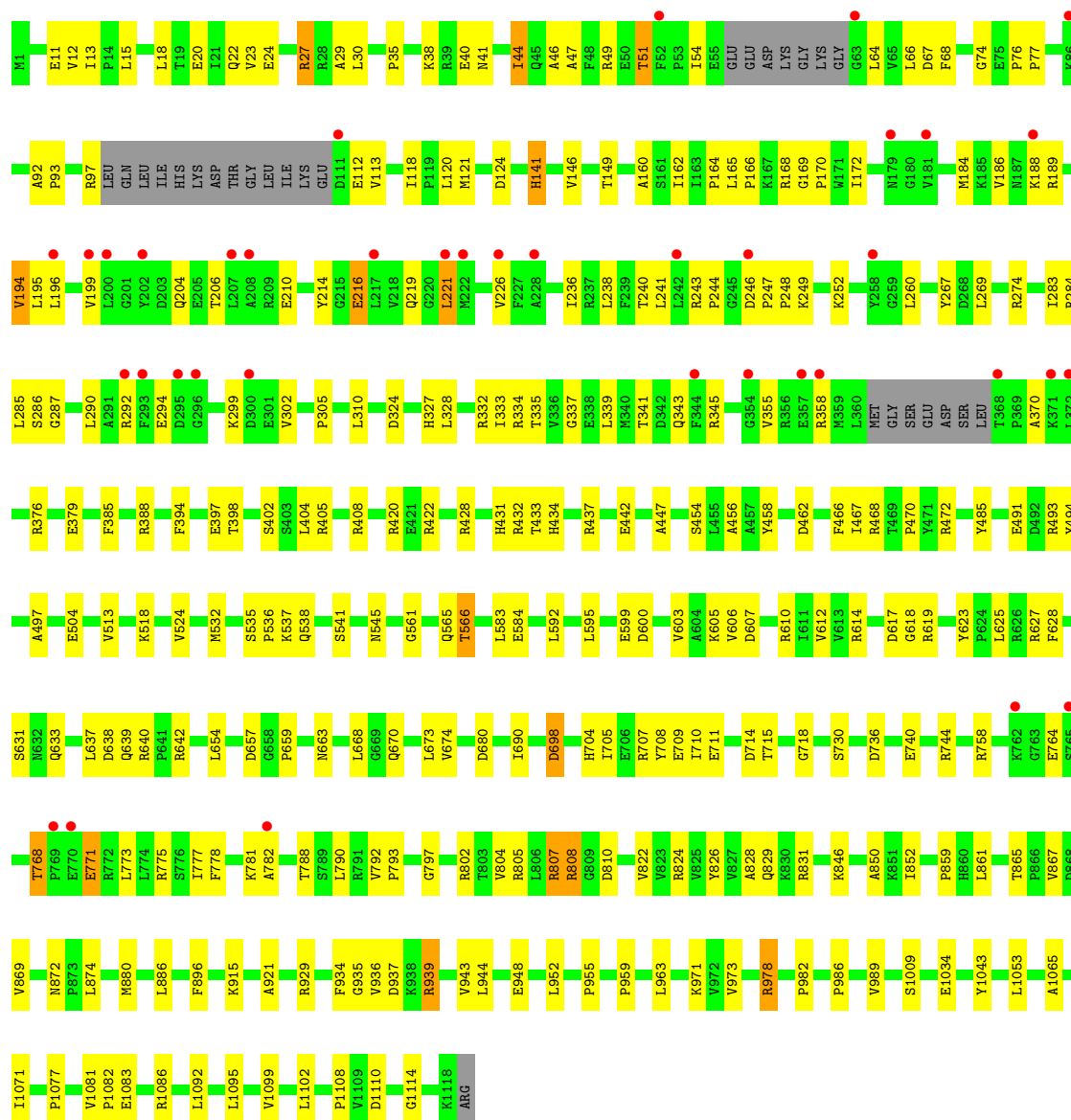
- Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:  80% 18%

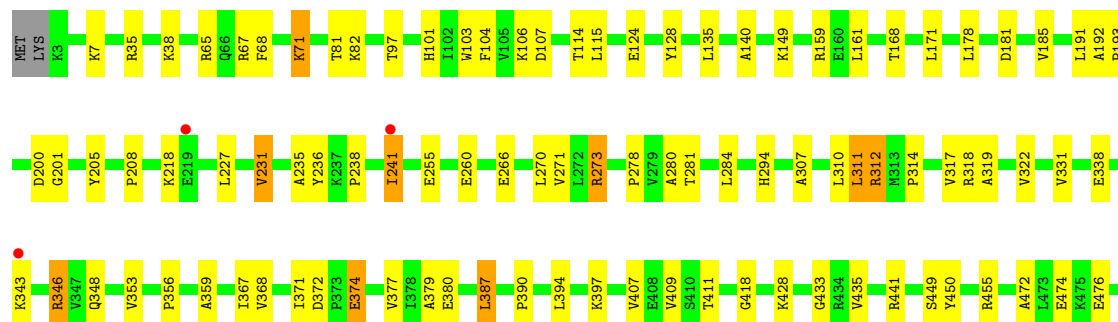
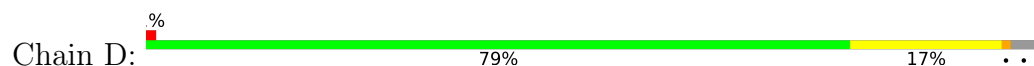
R978	R907	Q633	D429	Y267	H1
E984	D810	L637	G436	D268	H141
H1006	L315	D638	A447	K280	V146
K1024	K816	R642	I452	R284	A160
Q1030	V819	V645	T453	L285	S161
R1031	V822	L655	S454	D300	I162
V1037	Q829	A656	D462	V309	L165
W1038	Q870	Q670	R468	D324	R168
A1042	P854	V674	R472	H527	I172
L1049	P862	F684	V483	L328	D173
Q1050	V869	T690	T489	K330	E175
T1054	N872	E693	V513	Q343	V176
L1055	P873	K696	V524	F344	E177
K1056	L374	K696	V524	L351	P178
A1065	N881	F699	Q538	R358	A46
A1066	Q884	Y708	N545	K359	E56
Y1067	T888	E711	T566	L367	GLU
P1077	T902	L717	Q575	GLU	ASP
S1080	E916	P727	L583	ASP	LYS
K1090	Q920	D738	L595	L372	LYS
L1102	A921	Y596	A597	R209	GLY
D1110	F922	G757	E598	E210	G63
G1114	K930	G763	E599	L211	L64
S1117	F934	E766	D600	A380	E70
K1118	Q935	P767	K605	E384	G74
ARG	V936	T768	V606	F386	E75
	D937	D607	D607	E397	P77
	K938	E771		S392	D87
	R939	R783	R610	M222	L217
	L944	T788	I611	D223	L218
	R945	S789	V612	K395	Q219
	R946	T789	G618	D396	A92
	L952	L790	R619	E397	F93
	P959	R791	R619	R405	R97
		V792	R626	R409	
	K964	P793	R627	L413	E110
	Y975	G797	F628	R420	T118
	D976		R630	E421	P119
			S631	L261	L120
			V823	R428	D133
					P124

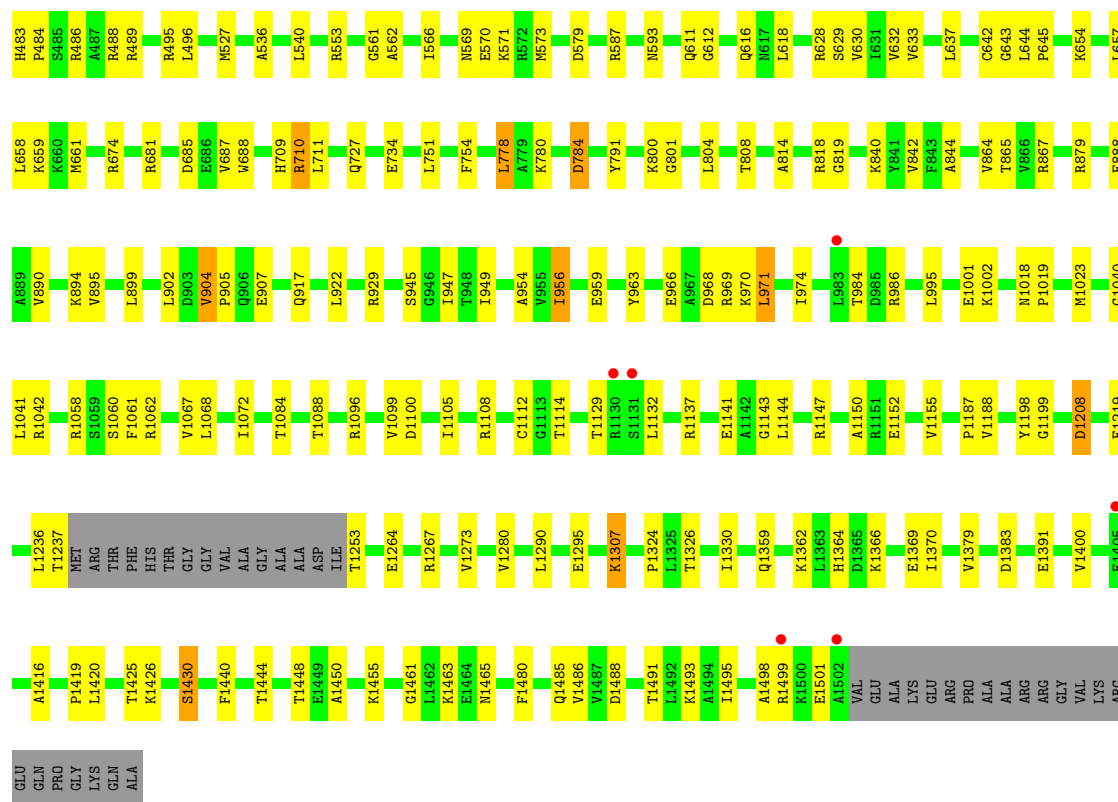
- Molecule 2: DNA-directed RNA polymerase subunit beta

Chain M:  3% 72% 24%

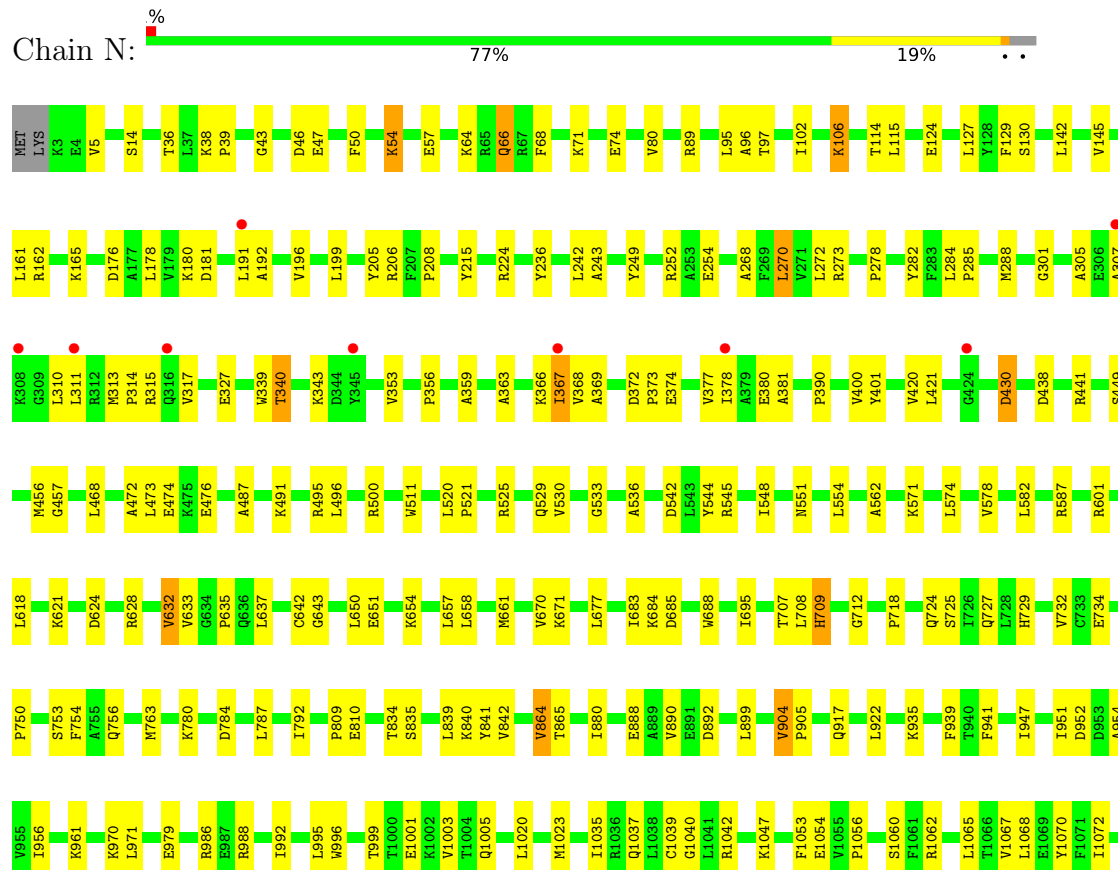


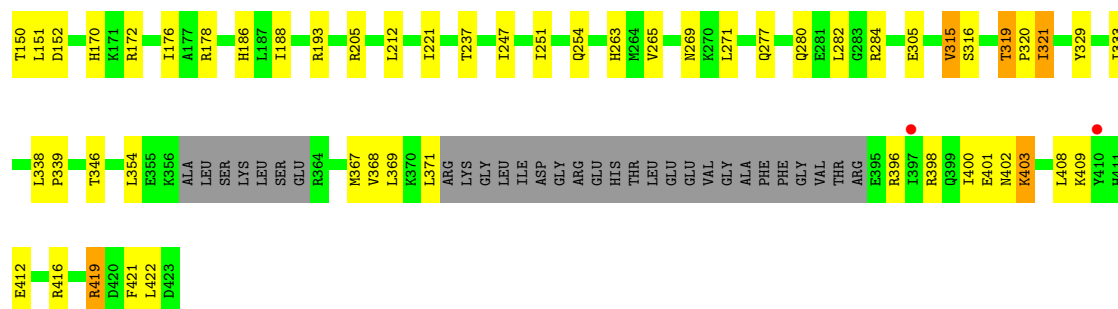
• Molecule 3: DNA-directed RNA polymerase subunit beta'



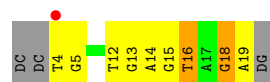


• Molecule 3: DNA-directed RNA polymerase subunit beta'





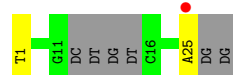
- Molecule 6: DNA (5'-D(*CP*C*TP*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*AP*GP*AP*G)-3')



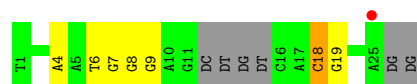
- Molecule 6: DNA (5'-D(*CP*C*TP*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*AP*GP*AP*G)-3')



- Molecule 7: DNA (27-MER)



- Molecule 7: DNA (27-MER)



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	185.00Å 103.51Å 296.30Å 90.00° 98.27° 90.00°	Depositor
Resolution (Å)	42.17 – 3.00 49.42 – 2.99	Depositor EDS
% Data completeness (in resolution range)	97.5 (42.17-3.00) 97.7 (49.42-2.99)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.12	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.78 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.8-1069	Depositor
R, R_{free}	0.201 , 0.254 0.203 , 0.254	Depositor DCC
R_{free} test set	10942 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	48.8	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 31.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	57349	wwPDB-VP
Average B, all atoms (Å ²)	31.0	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 42.42 % 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 2.0439e-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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, NAD, C5P, MG, CTP, AMP

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	0.29	0/1829	0.53	0/2487
1	B	0.28	0/1781	0.50	0/2420
1	K	0.27	0/1824	0.49	0/2480
1	L	0.28	0/1781	0.50	0/2420
2	C	0.32	0/8913	0.51	0/12053
2	M	0.29	0/8775	0.49	0/11867
3	D	0.31	0/11936	0.50	1/16138 (0.0%)
3	N	0.30	0/11922	0.49	0/16119
4	E	0.31	0/772	0.50	0/1040
4	O	0.28	0/772	0.47	0/1040
5	F	0.29	0/2852	0.46	0/3837
5	P	0.28	0/2614	0.46	0/3516
6	G	0.66	1/368 (0.3%)	1.15	2/567 (0.4%)
6	R	0.53	0/368	1.08	2/567 (0.4%)
7	H	0.57	0/489	1.14	1/752 (0.1%)
7	S	0.55	0/488	1.15	2/750 (0.3%)
All	All	0.31	1/57484 (0.0%)	0.53	8/78053 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	18	DG	O3'-P	-7.08	1.52	1.61

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	16	DT	O4'-C4'-C3'	-7.75	101.35	106.00
7	S	18	DC	O4'-C1'-N1	6.36	112.45	108.00
6	R	16	DT	O4'-C4'-C3'	-5.75	102.20	104.50
6	G	14	DA	O4'-C4'-C3'	-5.62	102.25	104.50
6	R	14	DA	O4'-C4'-C3'	-5.49	102.30	104.50

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	1797	0	1849	33	0
1	B	1750	0	1802	29	0
1	K	1792	0	1844	39	0
1	L	1750	0	1802	41	0
2	C	8747	0	8858	121	0
2	M	8611	0	8710	169	0
3	D	11730	0	11960	158	0
3	N	11716	0	11949	174	1
4	E	758	0	770	14	0
4	O	758	0	770	7	0
5	F	2807	0	2882	41	0
5	P	2574	0	2643	45	1
6	G	328	0	182	7	0
6	R	328	0	182	4	0
7	H	435	0	238	2	0
7	S	434	0	235	10	0
8	B	1	0	0	0	0
8	D	3	0	0	0	0
8	F	1	0	0	0	0
8	L	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	D	20	0	11	0	0
11	D	9	0	0	0	0
11	M	9	0	0	0	0
12	G	44	0	24	1	0
13	N	20	0	11	3	0
14	R	23	0	11	0	0
15	A	23	0	0	0	0
15	B	16	0	0	0	0
15	C	196	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
15	D	243	0	0	7	0
15	E	16	0	0	0	0
15	F	45	0	0	1	0
15	G	9	0	0	0	0
15	H	4	0	0	0	0
15	K	14	0	0	0	0
15	L	13	0	0	1	0
15	M	97	0	0	4	0
15	N	187	0	0	2	0
15	O	12	0	0	0	0
15	P	14	0	0	1	0
15	R	3	0	0	0	0
15	S	3	0	0	0	0
All	All	57349	0	56733	808	1

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 808 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)
2:M:428:ARG:NH2	2:M:447:ALA:O	2.07	0.88
3:N:562:ALA:O	5:P:140:ARG:NH1	2.10	0.84
3:D:124:GLU:OE2	3:D:587:ARG:NH2	2.11	0.84
2:M:674:VAL:HG12	2:M:869:VAL:HB	1.63	0.81
2:C:787:ASP:OD2	2:C:791:ARG:NH2	2.15	0.78

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:N:327:GLU:OE2	5:P:263:HIS:NE2[2_545]	2.16	0.04

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	227/315 (72%)	224 (99%)	3 (1%)	0	100	100
1	B	218/315 (69%)	214 (98%)	4 (2%)	0	100	100
1	K	226/315 (72%)	221 (98%)	5 (2%)	0	100	100
1	L	218/315 (69%)	213 (98%)	5 (2%)	0	100	100
2	C	1102/1119 (98%)	1076 (98%)	26 (2%)	0	100	100
2	M	1083/1119 (97%)	1050 (97%)	33 (3%)	0	100	100
3	D	1481/1524 (97%)	1452 (98%)	29 (2%)	0	100	100
3	N	1479/1524 (97%)	1449 (98%)	28 (2%)	2 (0%)	51	85
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	344/443 (78%)	339 (98%)	5 (2%)	0	100	100
5	P	310/443 (70%)	304 (98%)	6 (2%)	0	100	100
All	All	6872/7630 (90%)	6722 (98%)	148 (2%)	2 (0%)	100	100

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	N	1131	SER
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%)	192 (96%)	7 (4%)	36	71
1	B	195/273 (71%)	189 (97%)	6 (3%)	40	75
1	K	199/273 (73%)	193 (97%)	6 (3%)	41	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	L	195/273 (71%)	189 (97%)	6 (3%)	40	75
2	C	933/941 (99%)	899 (96%)	34 (4%)	35	70
2	M	917/941 (97%)	876 (96%)	41 (4%)	27	64
3	D	1252/1279 (98%)	1196 (96%)	56 (4%)	27	64
3	N	1251/1279 (98%)	1195 (96%)	56 (4%)	27	64
4	E	82/88 (93%)	79 (96%)	3 (4%)	34	70
4	O	82/88 (93%)	77 (94%)	5 (6%)	18	53
5	F	301/388 (78%)	293 (97%)	8 (3%)	44	77
5	P	277/388 (71%)	258 (93%)	19 (7%)	15	48
All	All	5883/6484 (91%)	5636 (96%)	247 (4%)	30	66

5 of 247 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	10	VAL
3	N	1497	GLU
2	M	221	LEU
3	N	1487	VAL
5	P	271	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
3	N	1172	HIS
3	N	1195	GLN
5	P	280	GLN
3	D	1195	GLN
3	D	1184	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 [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 14 are monoatomic - leaving 6 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	$\# Z > 2$
14	AMP	R	101	13	22,25,25	2.20	8 (36%)	25,38,38	1.27	2 (8%)
11	CTP	D	2006	8	6,8,30	1.70	1 (16%)	13,13,47	1.42	1 (7%)
11	CTP	M	1201	8	6,8,30	1.07	1 (16%)	13,13,47	1.23	1 (7%)
12	NAD	G	101	10	42,48,48	2.16	11 (26%)	50,73,73	1.30	6 (12%)
13	C5P	N	2005	14,8	18,21,22	0.31	0	26,30,33	0.34	0
10	C	D	2005	12,8	18,21,22	0.30	0	26,30,33	0.38	0

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
14	AMP	R	101	13	-	5/6/26/26	0/3/3/3
11	CTP	D	2006	8	-	0/6/6/38	-
11	CTP	M	1201	8	-	0/6/6/38	-
12	NAD	G	101	10	-	9/26/62/62	0/5/5/5
13	C5P	N	2005	14,8	-	0/7/25/26	0/2/2/2
10	C	D	2005	12,8	-	0/7/25/26	0/2/2/2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	G	101	NAD	C7N-N7N	5.92	1.44	1.33
14	R	101	AMP	O4'-C1'	5.77	1.49	1.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	G	101	NAD	O4B-C1B	5.67	1.49	1.41
12	G	101	NAD	O4D-C1D	5.33	1.48	1.41
14	R	101	AMP	C2-N3	4.66	1.39	1.32

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	R	101	AMP	N3-C2-N1	-4.65	121.41	128.68
12	G	101	NAD	N3A-C2A-N1A	-4.61	121.47	128.68
11	D	2006	CTP	PB-O3A-PA	-3.76	119.92	132.83
12	G	101	NAD	PN-O3-PA	-3.75	119.96	132.83
11	M	1201	CTP	PB-O3A-PA	-3.14	122.04	132.83

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

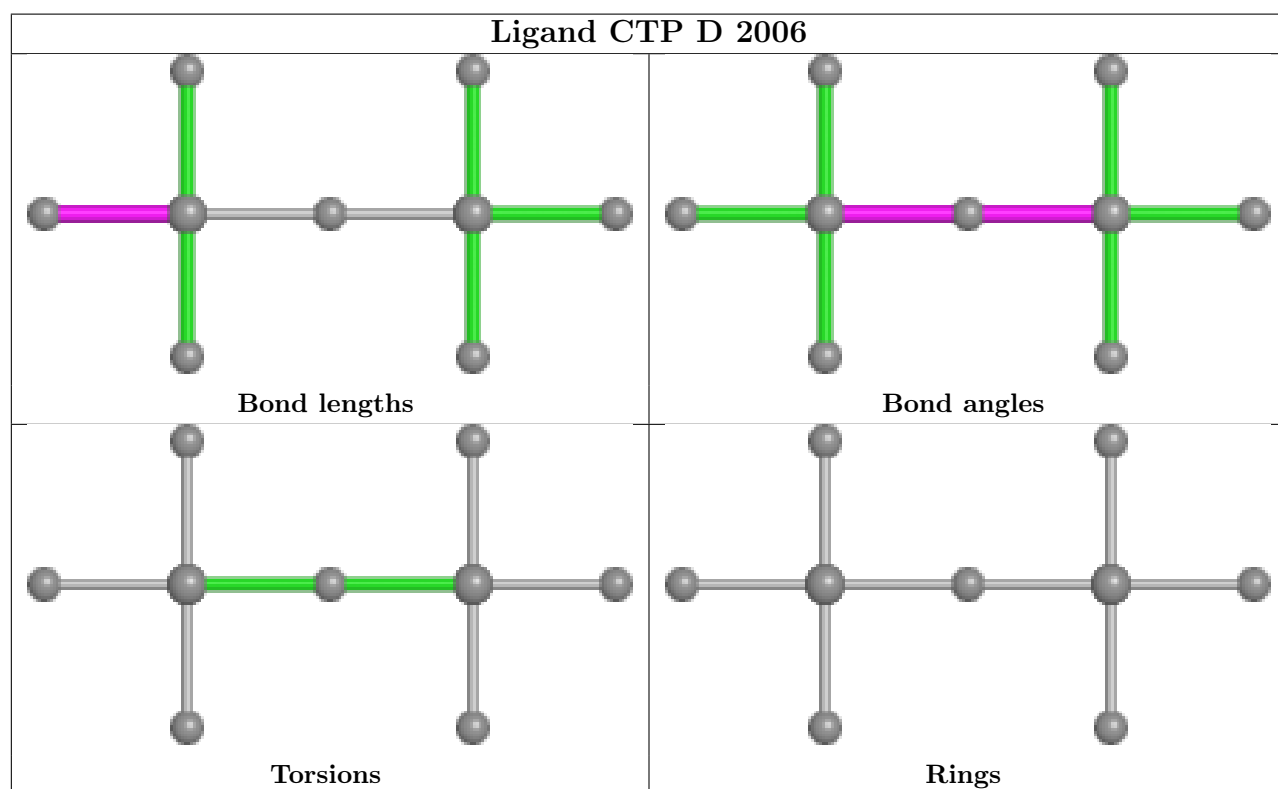
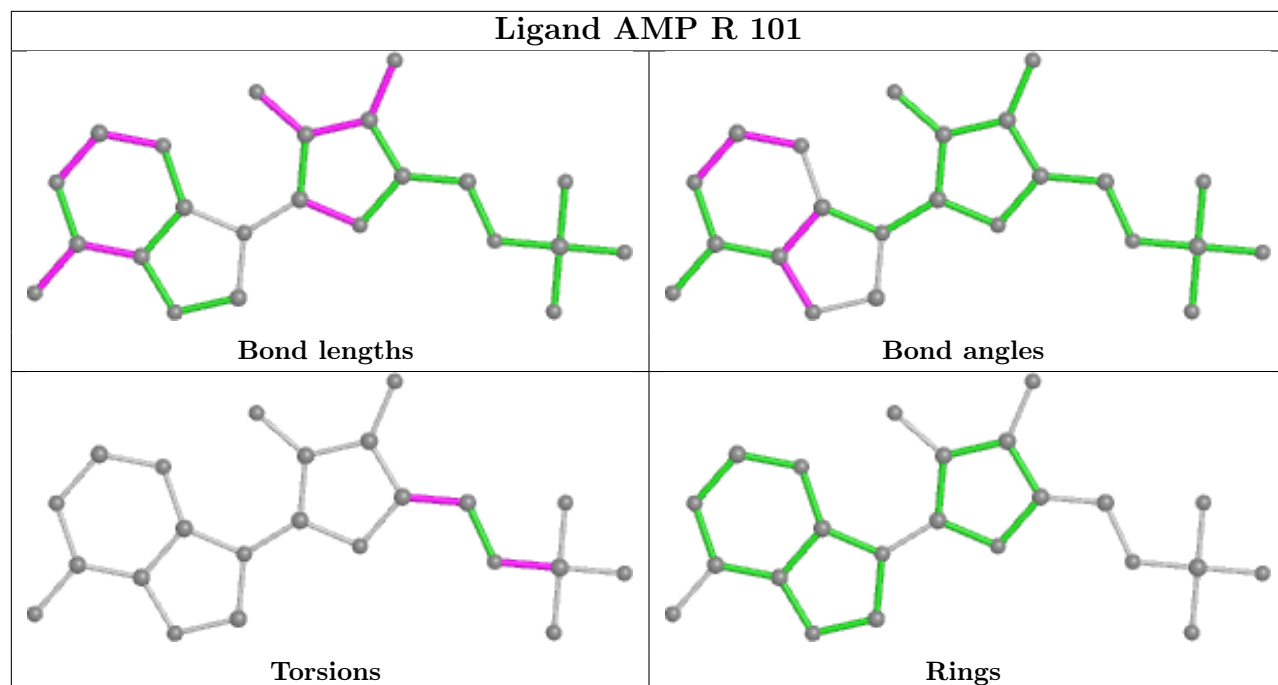
Mol	Chain	Res	Type	Atoms
12	G	101	NAD	C5B-O5B-PA-O3
12	G	101	NAD	PN-O3-PA-O5B
12	G	101	NAD	O4B-C4B-C5B-O5B
12	G	101	NAD	C5D-O5D-PN-O3
12	G	101	NAD	C5D-O5D-PN-O2N

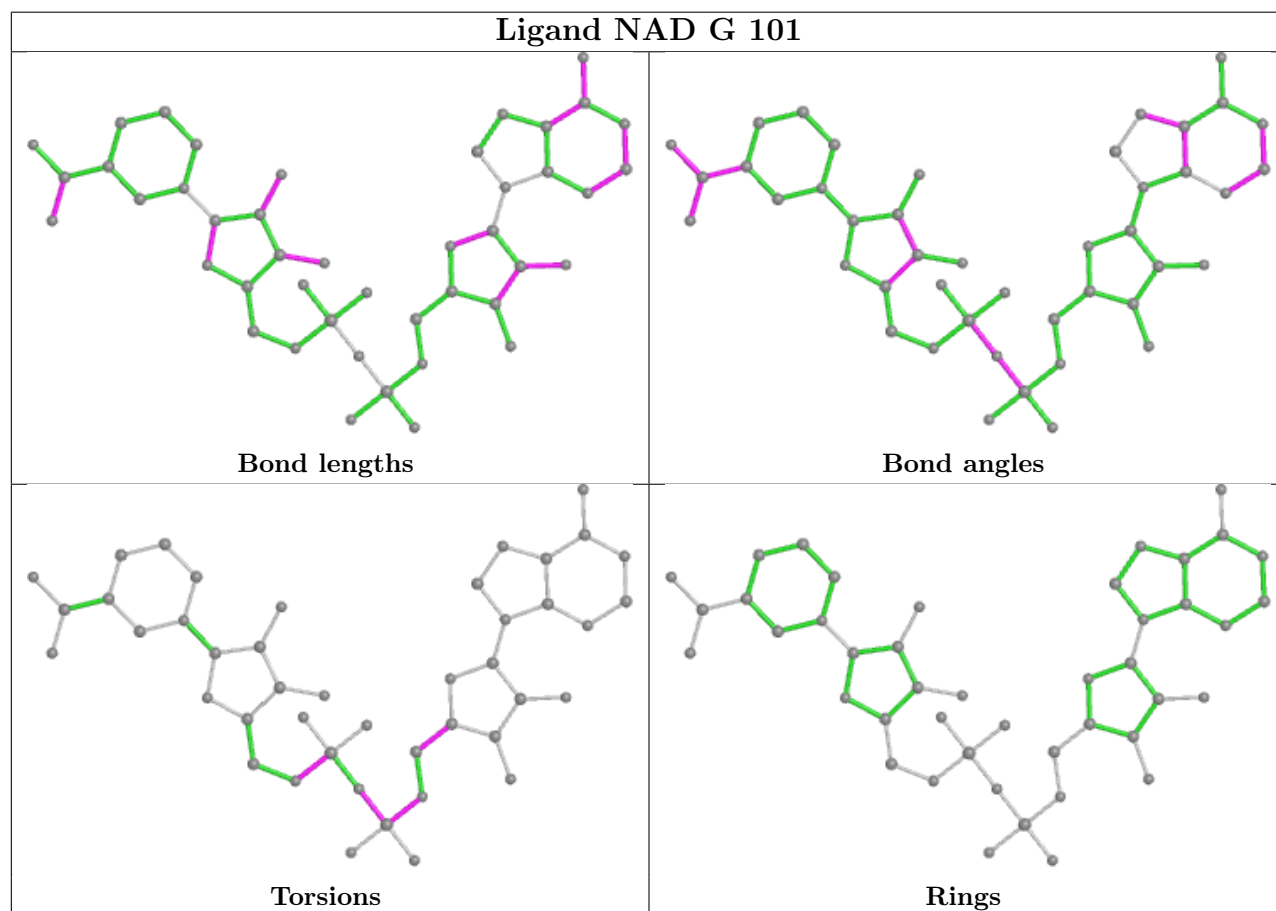
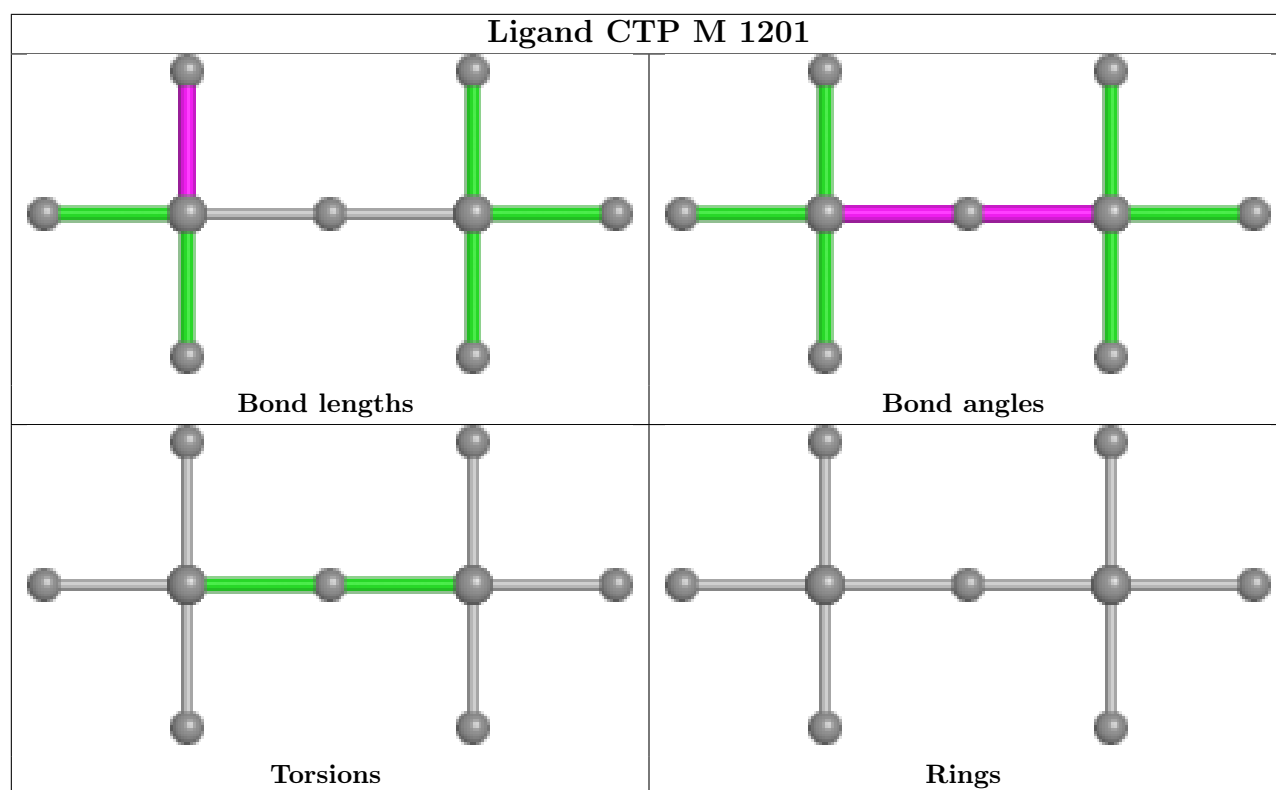
There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	G	101	NAD	1	0
13	N	2005	C5P	3	0

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	229/315 (72%)	-0.51	1 (0%) 92 79	7, 21, 48, 79	1 (0%)
1	B	222/315 (70%)	-0.44	0 100 100	6, 31, 64, 81	0
1	K	228/315 (72%)	-0.38	0 100 100	10, 31, 57, 71	1 (0%)
1	L	222/315 (70%)	-0.40	0 100 100	12, 36, 73, 103	0
2	C	1108/1119 (99%)	-0.45	5 (0%) 91 75	0, 14, 62, 89	3 (0%)
2	M	1091/1119 (97%)	-0.12	38 (3%) 44 18	1, 39, 97, 116	2 (0%)
3	D	1485/1524 (97%)	-0.36	9 (0%) 89 72	0, 18, 68, 102	4 (0%)
3	N	1483/1524 (97%)	-0.30	12 (0%) 86 65	0, 24, 77, 106	4 (0%)
4	E	94/99 (94%)	-0.55	1 (1%) 80 56	1, 13, 49, 67	0
4	O	94/99 (94%)	-0.40	0 100 100	4, 23, 62, 71	0
5	F	346/443 (78%)	-0.43	1 (0%) 94 84	3, 26, 68, 84	0
5	P	316/443 (71%)	-0.20	2 (0%) 89 72	17, 46, 91, 109	0
6	G	16/19 (84%)	-0.14	1 (6%) 20 6	16, 48, 112, 117	0
6	R	16/19 (84%)	-0.08	0 100 100	39, 65, 115, 122	0
7	H	21/27 (77%)	-0.18	1 (4%) 30 11	19, 56, 108, 124	0
7	S	21/27 (77%)	0.04	1 (4%) 30 11	33, 84, 121, 144	0
All	All	6992/7722 (90%)	-0.33	72 (1%) 82 59	0, 26, 78, 144	15 (0%)

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	M	188	LYS	5.0
2	M	200	LEU	4.4
2	M	196	LEU	4.4
2	M	63	GLY	3.8
2	M	368	THR	3.7

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](#)

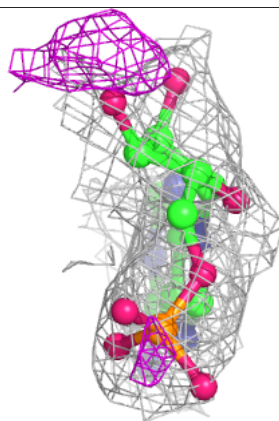
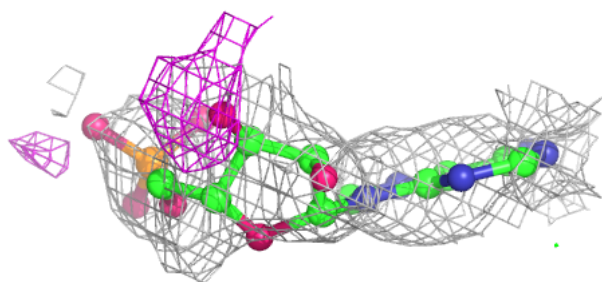
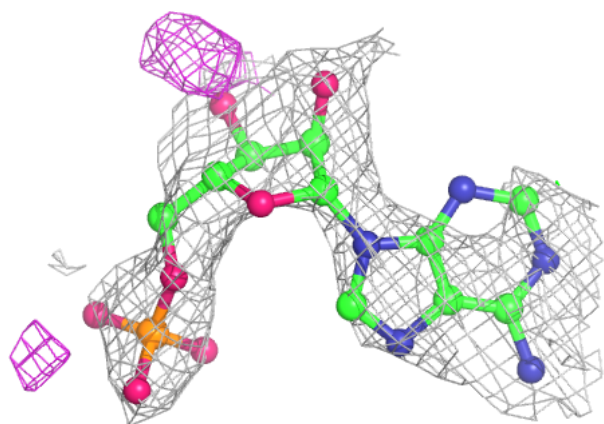
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
14	AMP	R	101	23/23	0.82	0.26	37,54,63,83	23
13	C5P	N	2005	20/21	0.88	0.23	22,46,53,59	20
8	MG	D	2003	1/1	0.88	0.23	1,1,1,1	0
11	CTP	M	1201	9/29	0.89	0.20	25,52,69,73	0
12	NAD	G	101	44/44	0.89	0.24	15,34,53,67	44
8	MG	P	2001	1/1	0.91	0.24	51,51,51,51	0
11	CTP	D	2006	9/29	0.91	0.20	21,41,62,80	0
8	MG	B	2001	1/1	0.92	0.20	17,17,17,17	0
8	MG	F	2001	1/1	0.93	0.06	25,25,25,25	0
8	MG	L	2001	1/1	0.93	0.10	41,41,41,41	0
8	MG	N	2003	1/1	0.93	0.26	6,6,6,6	0
10	C	D	2005	20/21	0.94	0.18	6,19,38,47	20
8	MG	N	2006	1/1	0.94	0.28	32,32,32,32	0
8	MG	D	2007	1/1	0.94	0.33	28,28,28,28	0
9	ZN	N	2002	1/1	0.96	0.07	89,89,89,89	0
8	MG	N	2004	1/1	0.98	0.07	27,27,27,27	0
9	ZN	D	2001	1/1	0.99	0.14	5,5,5,5	0
9	ZN	D	2002	1/1	0.99	0.05	40,40,40,40	0
8	MG	D	2004	1/1	0.99	0.18	15,15,15,15	0
9	ZN	N	2001	1/1	1.00	0.15	7,7,7,7	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.

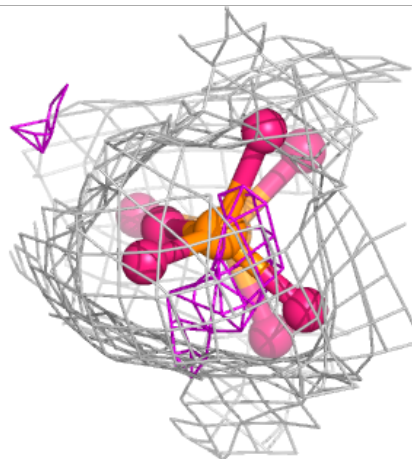
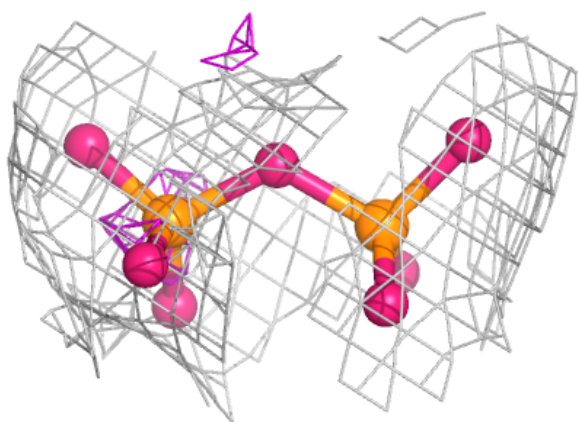
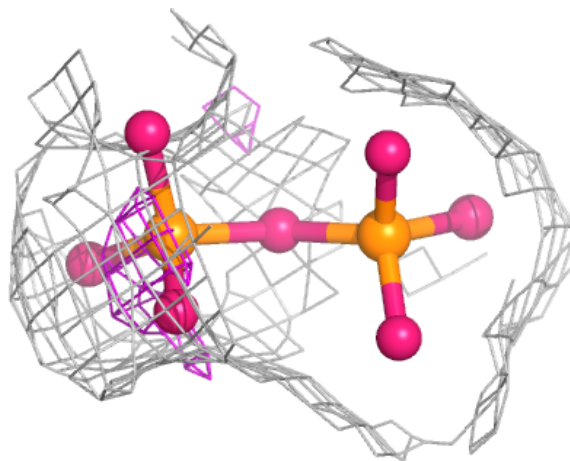
Electron density around AMP R 101:

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



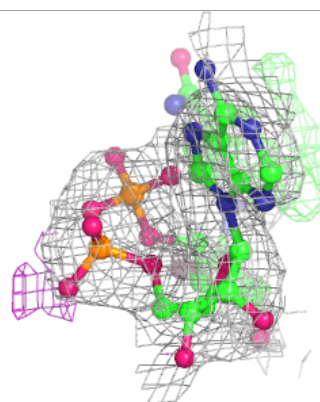
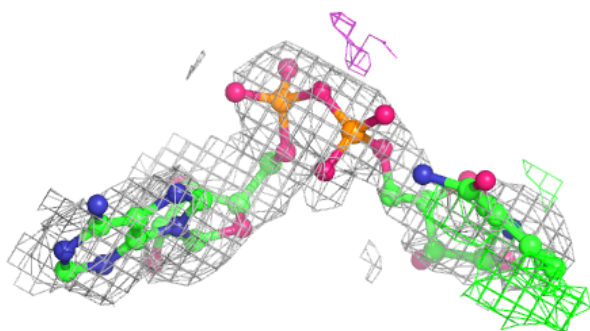
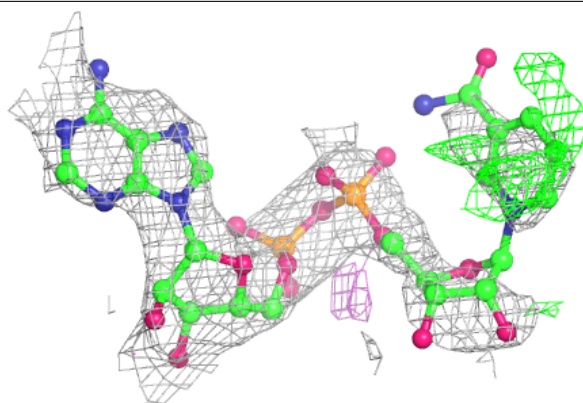
Electron density around CTP M 1201:

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

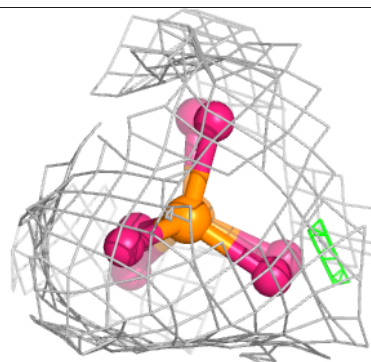
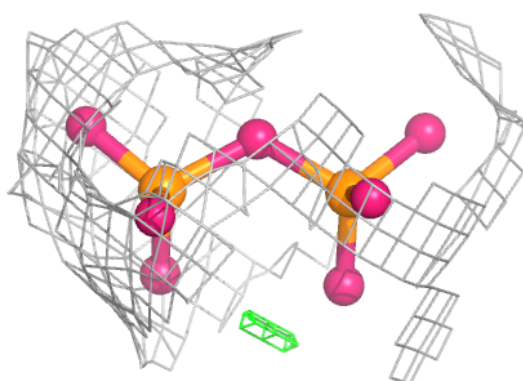
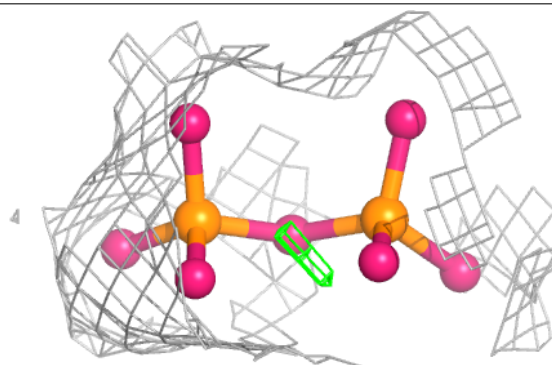


Electron density around NAD G 101:

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

**Electron density around CTP D 2006:**

$2mF_o-DF_c$ (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.