



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 4, 2025 – 10:09 AM EDT

PDB ID : 9N5E / pdb_00009n5e
Title : RNA polymerase II elongation complex with 8-oxoG at +1 site, AMPCPP in E-site
Authors : Oh, J.; Wang, D.
Deposited on : 2025-02-04
Resolution : 3.75 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 Phenix2.0rc1
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
buster-report : 1.1.7 (2018)
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.45.1

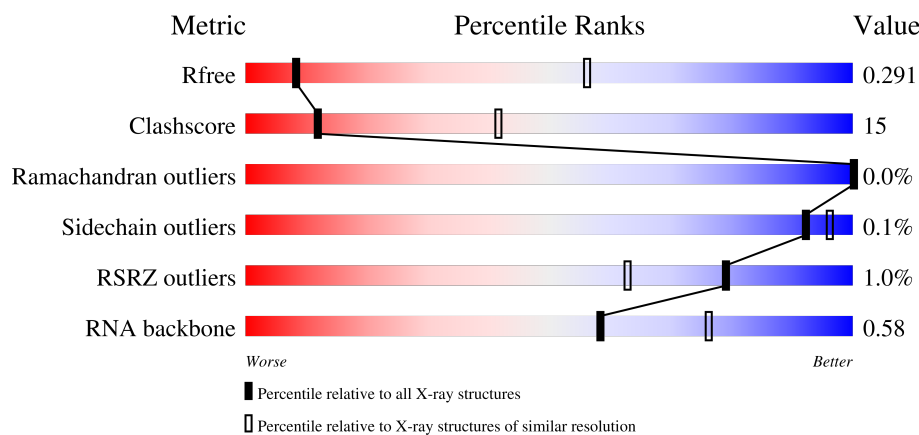
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.75 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}	164625	1256 (3.92-3.60)
Clashscore	180529	1321 (3.92-3.60)
Ramachandran outliers	177936	1293 (3.92-3.60)
Sidechain outliers	177891	1288 (3.92-3.60)
RSRZ outliers	164620	1256 (3.92-3.60)
RNA backbone	3690	1130 (4.52-3.00)

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	R	9	<div> <div>11%</div> <div>89%</div> </div>
2	T	29	<div> <div>7%</div> <div>72%</div> <div>•</div> <div>17%</div> </div>
3	N	18	<div> <div>22%</div> <div>50%</div> <div>28%</div> </div>
4	A	1733	<div> <div>%</div> <div>53%</div> <div>27%</div> <div>20%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	 % 63% 29% 8%
6	C	318	 % 51% 33% 16%
7	E	215	 % 67% 32% 1%
8	F	155	 % 34% 22% 45%
9	H	146	 % 58% 33% 9%
10	I	122	 % 68% 29% 3%
11	J	70	 3% 54% 39% 7%
12	K	120	 % 68% 28% 5%
13	L	70	 % 43% 19% 39%

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 28950 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 RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	9	Total	C	N	O	P	0	0	0
			199	88	40	62	9			

- Molecule 2 is a DNA chain called Template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	24	Total	C	N	O	P	0	0	0
			481	230	76	151	24			

- Molecule 3 is a DNA chain called Non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	13	Total	C	N	O	P	0	0	0
			275	128	61	73	13			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1385	Total	C	N	O	S	0	0	0
			10807	6821	1883	2043	60			

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1121	Total	C	N	O	S	0	0	0
			8843	5598	1547	1645	53			

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	213	Total	C	N	O	S	0	0	0
			1734	1102	304	317	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	86	Total	C	N	O	S	0	0	0
			684	437	115	129	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total	C	N	O	S	0	0	0
			1058	667	176	211	4			

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	118	Total	C	N	O	S	0	0	0
			946	582	170	184	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	43	Total	C	N	O	S	0	0	0
			331	205	63	59	4			

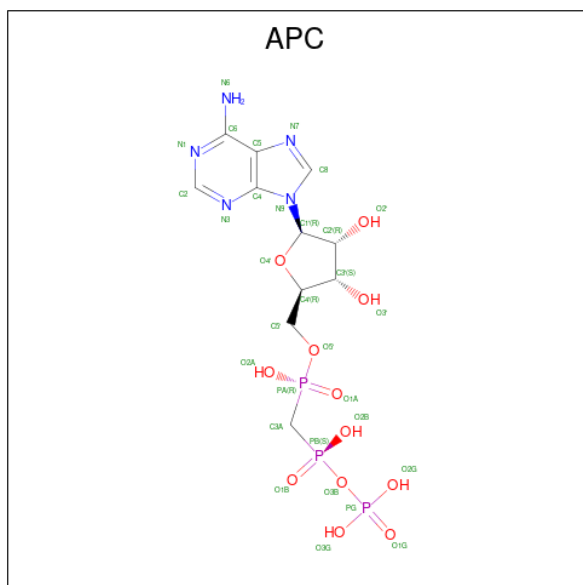
- Molecule 14 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total	Zn	0	0
			2	2		
14	B	1	Total	Zn	0	0
			1	1		
14	C	1	Total	Zn	0	0
			1	1		
14	I	2	Total	Zn	0	0
			2	2		
14	J	1	Total	Zn	0	0
			1	1		
14	L	1	Total	Zn	0	0
			1	1		

- Molecule 15 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	1	Total	Mg	0	0
			1	1		

- Molecule 16 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (CCD ID: APC) (formula: C₁₁H₁₈N₅O₁₂P₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	B	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

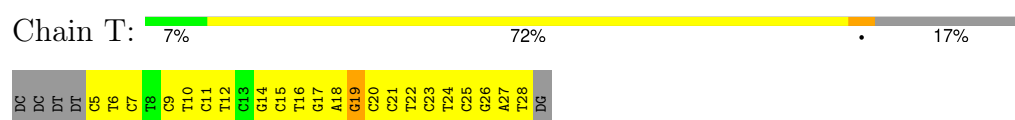
3 Residue-property plots

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.

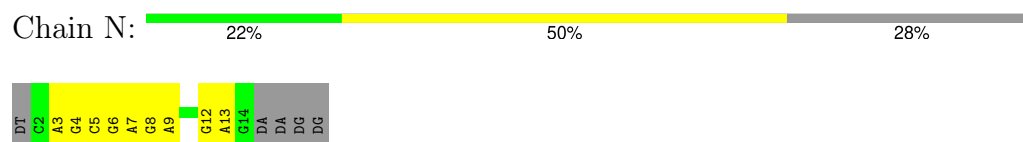
• Molecule 1: RNA



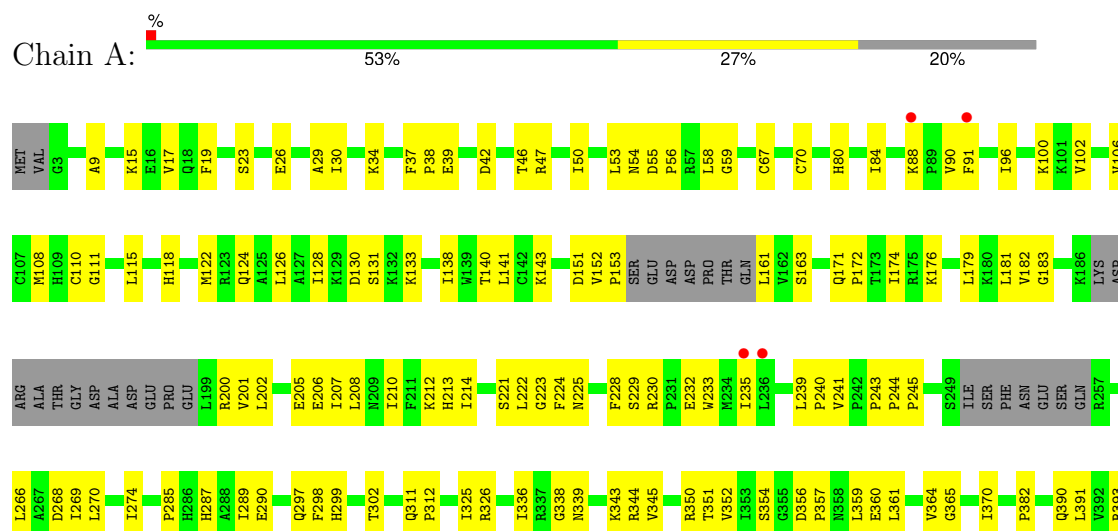
• Molecule 2: Template strand DNA



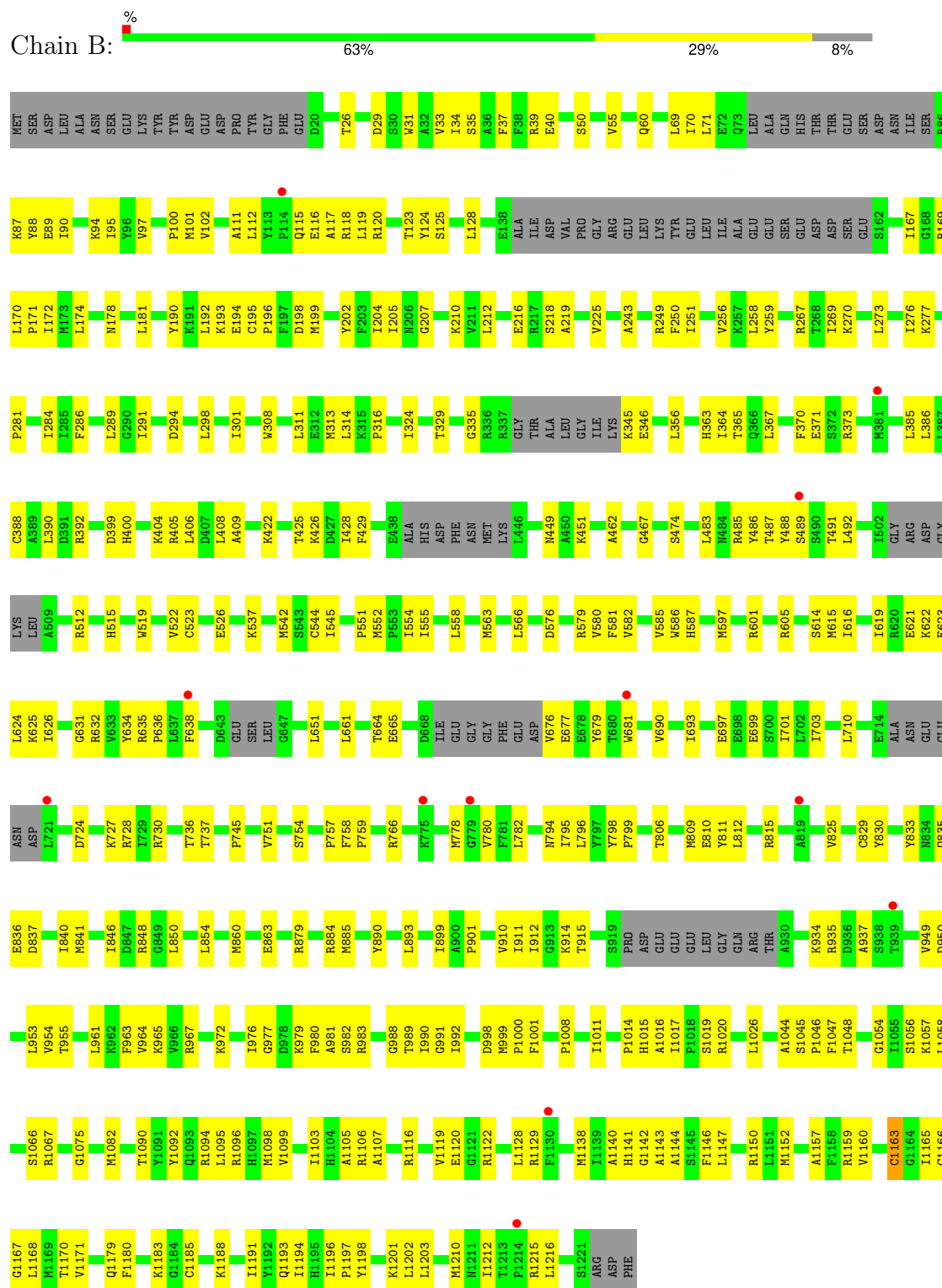
• Molecule 3: Non-template strand DNA



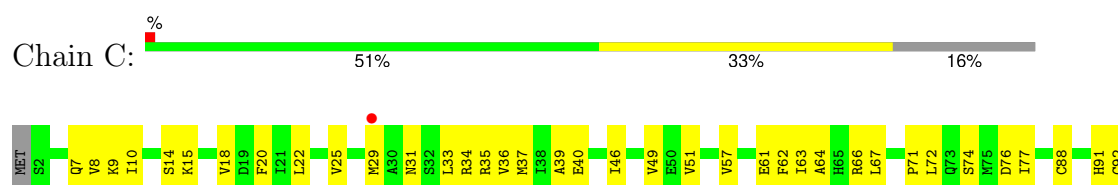
• Molecule 4: DNA-directed RNA polymerase II subunit RPB1

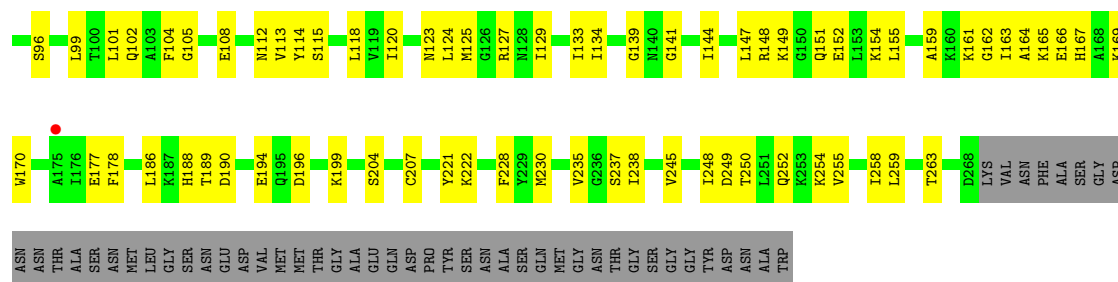


THR	VAL	LEU	C1421	V1316	L1197	A1108	F1018	L925	K797	L702	H587	V491	P386
SER	ASP	VAL	E1426	M1317	D1198	M1111	L1021	Q926	G798	T703	H587	E500	P386
PRO	SER	ASP	I1322	I1322	K1205	K1112	A1027	E932	F799	A704	R590	E501	P400
THR	GLY	SER	M1433	R1326	V1212	P1113	T1028	L936	E812	T709	T596	V507	Y404
PRO	ASN	ASN	A1434	I1326	I1216	S1115	R1029	D939	F813	L710	L597	P508	W406
THR	SER	ASP	T1438	T1329	I1216	T1117	R1030	R940	F814	R711	L598	I406	R407
SER	THR	ALA	F1441	N1330	F1220	L1118	E1034	F942	F815	E712	P600	V512	I406
PRO	PRO	ALA	F1332	F1332	F1220	Y1119	Y1035	F942	F815	E712	P600	V512	R407
THR	THR	ALA	F1332	F1332	F1220	Y1119	Y1035	F942	F815	E712	P600	V512	R407
THR	THR	GLY	D1446	I1335	L1224	L1120	R1036	L943	M819	K728	G604	S513	R412
SER	GLY	GLY	GLY	I1335	L1224	L1120	R1036	L943	M819	K728	G604	S513	R412
PRO	PHE	PHE	GLY	I1335	L1224	L1120	R1036	L943	M819	K728	G604	S513	R412
THR	THR	THR	THR	L1339	I1227	H1124	W1044	W954	G820	L732	L606	N517	D414
SER	SER	ALA	G1340	G1340	S1229	Q1128	L1046	P955	R821	A733	L606	K518	D414
PRO	SER	VAL	I1341	I1341	S1229	Q1128	L1046	P955	R821	A733	L606	K518	D414
THR	GLY	GLY	R1345	R1345	I1237	E1129	I1049	P957	I825	L737	F614	V524	P441
PRO	GLY	GLY	R1345	R1345	I1237	E1129	I1049	P957	I825	L737	F614	V524	P441
SER	ALA	ALA	R1345	R1345	I1237	E1129	I1049	P957	I825	L737	F614	V524	P441
PRO	ASP	PRO	L1348	L1348	R1239	M1131	A1056	I960	G835	N741	G615	Q525	V442
THR	THR	THR	L1348	L1348	R1239	M1131	A1056	I960	G835	N741	G615	Q525	V442
PRO	THR	GLY	E1351	E1351	R1241	L1133	A1056	I960	G835	N741	G615	Q525	V442
SER	GLY	GLY	V1242	V1242	R1241	L1133	A1056	I960	G835	N741	G615	Q525	V442
PRO	GLY	GLY	V1352	V1352	R1241	L1133	A1056	I960	G835	N741	G615	Q525	V442
THR	THR	ALA	ARG	ARG	V1243	R1135	R1055	I964	L845	N742	E618	L528	R446
PRO	ALA	ALA	ARG	ARG	V1243	R1135	R1055	I964	L845	N742	E618	L528	R446
THR	THR	THR	PRO	PRO	PRO	A1137	V1058	P965	I848	Q745	V622	I531	Q447
SER	SER	THR	THR	THR	THR	A1137	V1058	P965	I848	Q745	V622	I531	Q447
PRO	SER	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	PRO	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	PRO	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
THR	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
SER	THR	GLY	GLY	GLY	GLY	I1138	M1063	P966	M849	V747	R635	R532	P448
PRO	THR	GLY											



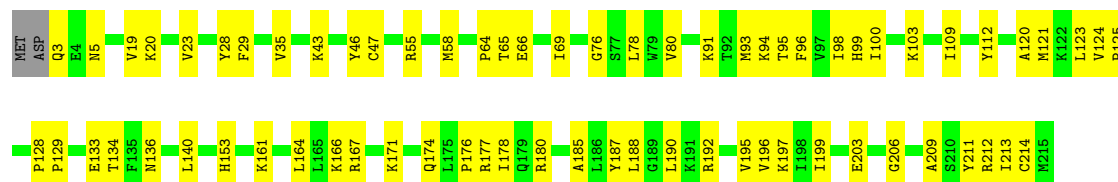
- Molecule 6: DNA-directed RNA polymerase II subunit RPB3





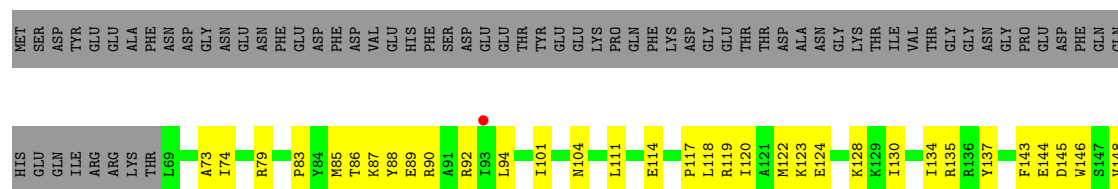
- Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 67% 32%



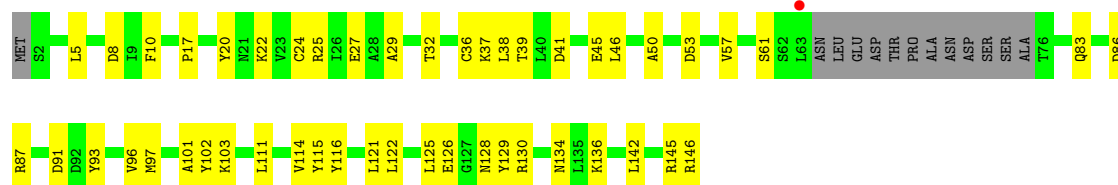
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 34% 22% 45%



- Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC3

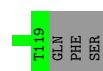
Chain H: 58% 33% 9%



- Molecule 10: DNA-directed RNA polymerase II subunit RPB9

Chain I: 68% 29%

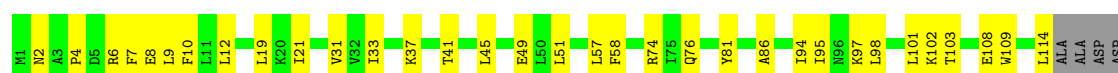




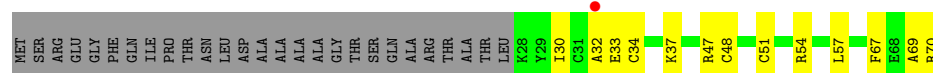
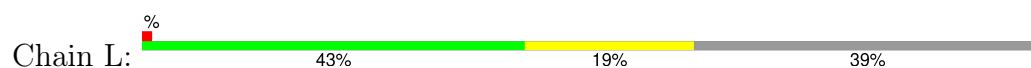
- Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC5



- Molecule 12: DNA-directed RNA polymerase II subunit RPB11



- Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC4



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	164.72Å 222.93Å 193.28Å 90.00° 99.20° 90.00°	Depositor
Resolution (Å)	48.96 – 3.75 48.96 – 3.75	Depositor EDS
% Data completeness (in resolution range)	98.9 (48.96-3.75) 98.8 (48.96-3.75)	Depositor EDS
R_{merge}	0.57	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.25 (at 3.77Å)	Xtriage
Refinement program	PHENIX (1.21.2_5419: ???)	Depositor
R, R_{free}	0.248 , 0.291 0.249 , 0.291	Depositor DCC
R_{free} test set	2000 reflections (2.51%)	wwPDB-VP
Wilson B-factor (Å ²)	109.2	Xtriage
Anisotropy	0.589	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 118.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	28950	wwPDB-VP
Average B, all atoms (Å ²)	141.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 8OG, ZN, APC

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	R	0.14	0/223	0.32	0/345
2	T	0.21	0/507	0.46	0/775
3	N	0.19	0/311	0.34	0/479
4	A	0.16	0/10999	0.35	0/14884
5	B	0.14	0/9014	0.34	0/12165
6	C	0.13	0/2139	0.32	0/2899
7	E	0.14	0/1770	0.34	0/2383
8	F	0.12	0/696	0.32	0/943
9	H	0.15	0/1076	0.41	0/1459
10	I	0.14	0/964	0.35	0/1301
11	J	0.12	0/541	0.28	0/727
12	K	0.14	0/937	0.35	0/1265
13	L	0.15	0/333	0.35	0/443
All	All	0.15	0/29510	0.35	0/40068

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 [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	R	199	0	98	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	T	481	0	273	21	0
3	N	275	0	144	7	0
4	A	10807	0	10831	398	0
5	B	8843	0	8798	277	0
6	C	2101	0	2057	80	0
7	E	1734	0	1755	54	0
8	F	684	0	692	30	0
9	H	1058	0	1018	37	0
10	I	946	0	888	29	0
11	J	532	0	544	23	0
12	K	919	0	929	29	0
13	L	331	0	343	12	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	1	0	0	0	0
16	B	31	0	13	0	0
All	All	28950	0	28383	888	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 15.

The worst 5 of 888 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)
4:A:70:CYS:SG	4:A:80:HIS:CE1	2.46	1.08
9:H:36:CYS:HA	9:H:126:GLU:O	1.54	1.07
8:F:135:ARG:NH2	8:F:143:PHE:CE2	2.31	0.97
10:I:74:GLU:HB3	10:I:81:ARG:HD3	1.57	0.87
4:A:469:ARG:NH1	5:B:991:GLY:O	2.09	0.85

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	A	1371/1733 (79%)	1335 (97%)	35 (3%)	1 (0%)	48	79
5	B	1101/1224 (90%)	1076 (98%)	25 (2%)	0	100	100
6	C	265/318 (83%)	264 (100%)	1 (0%)	0	100	100
7	E	211/215 (98%)	207 (98%)	4 (2%)	0	100	100
8	F	84/155 (54%)	83 (99%)	1 (1%)	0	100	100
9	H	129/146 (88%)	127 (98%)	2 (2%)	0	100	100
10	I	116/122 (95%)	113 (97%)	3 (3%)	0	100	100
11	J	63/70 (90%)	63 (100%)	0	0	100	100
12	K	112/120 (93%)	109 (97%)	3 (3%)	0	100	100
13	L	41/70 (59%)	39 (95%)	2 (5%)	0	100	100
All	All	3493/4173 (84%)	3416 (98%)	76 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	958	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	
4	A	1190/1520 (78%)	1188 (100%)	2 (0%)	92	95
5	B	954/1061 (90%)	953 (100%)	1 (0%)	92	96

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	C	235/274 (86%)	235 (100%)	0	100	100
7	E	193/197 (98%)	193 (100%)	0	100	100
8	F	73/137 (53%)	73 (100%)	0	100	100
9	H	115/128 (90%)	115 (100%)	0	100	100
10	I	109/116 (94%)	109 (100%)	0	100	100
11	J	60/65 (92%)	60 (100%)	0	100	100
12	K	99/102 (97%)	99 (100%)	0	100	100
13	L	36/57 (63%)	36 (100%)	0	100	100
All	All	3064/3657 (84%)	3061 (100%)	3 (0%)	92	96

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	67	CYS
4	A	110	CYS
5	B	1163	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	465	ASN
5	B	975	GLN
12	K	96	ASN
5	B	686	ASN
5	B	1112	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/9 (88%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	9	G

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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
2	8OG	T	19	2	22,25,26	4.14	18 (81%)	26,37,40	1.82	7 (26%)

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
2	8OG	T	19	2	-	2/7/21/22	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	19	8OG	C8-N7	7.27	1.51	1.38
2	T	19	8OG	C8-N9	6.22	1.51	1.40
2	T	19	8OG	C2-N3	6.17	1.48	1.33
2	T	19	8OG	C4-N3	5.95	1.47	1.34
2	T	19	8OG	C2-N2	4.97	1.45	1.34

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	19	8OG	C2-N3-C4	4.74	120.46	112.30
2	T	19	8OG	O6-C6-C5	-3.14	119.70	127.26
2	T	19	8OG	N9-C4-N3	3.05	129.94	126.13
2	T	19	8OG	C2-N1-C6	-2.97	119.72	125.11
2	T	19	8OG	C5-C6-N1	2.81	119.85	112.13

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	T	19	8OG	O4'-C4'-C5'-O5'
2	T	19	8OG	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	T	19	8OG	3	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 9 are monoatomic - leaving 1 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
16	APC	B	1302	-	25,33,33	4.43	9 (36%)	30,52,52	2.65	5 (16%)

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
16	APC	B	1302	-	-	8/15/38/38	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	1302	APC	O4'-C1'	17.15	1.63	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	1302	APC	PB-O3B	8.91	1.68	1.58
16	B	1302	APC	O4'-C4'	-6.46	1.30	1.45
16	B	1302	APC	O2'-C2'	4.44	1.53	1.43
16	B	1302	APC	PB-O2B	-3.26	1.48	1.56

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	1302	APC	C5-C6-N6	8.22	132.84	120.31
16	B	1302	APC	C4'-O4'-C1'	-6.91	103.59	109.92
16	B	1302	APC	N3-C2-N1	-6.34	120.07	128.67
16	B	1302	APC	N6-C6-N1	-5.25	107.11	118.33
16	B	1302	APC	O4'-C1'-N9	2.31	111.81	108.75

There are no chirality outliers.

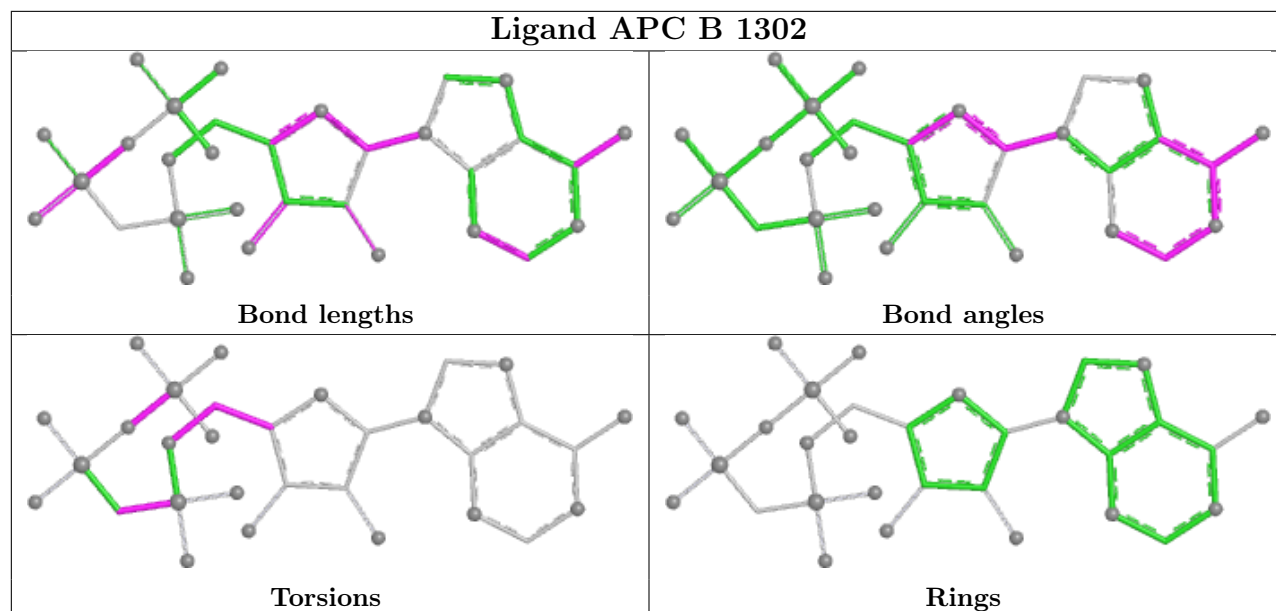
5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	B	1302	APC	PB-O3B-PG-O2G
16	B	1302	APC	PB-O3B-PG-O3G
16	B	1302	APC	PB-C3A-PA-O1A
16	B	1302	APC	PB-C3A-PA-O2A
16	B	1302	APC	PB-C3A-PA-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	R	9/9 (100%)	-0.17	0 100 100	102, 111, 218, 279	0
2	T	23/29 (79%)	-0.01	0 100 100	118, 197, 302, 324	0
3	N	13/18 (72%)	-0.17	0 100 100	184, 217, 310, 339	0
4	A	1385/1733 (79%)	-0.16	16 (1%) 76 58	80, 135, 219, 308	0
5	B	1121/1224 (91%)	-0.13	12 (1%) 77 61	61, 118, 192, 305	0
6	C	267/318 (83%)	-0.26	2 (0%) 84 70	64, 122, 177, 220	0
7	E	213/215 (99%)	-0.33	0 100 100	110, 161, 264, 339	0
8	F	86/155 (55%)	-0.42	1 (1%) 76 58	109, 131, 191, 236	0
9	H	133/146 (91%)	-0.28	1 (0%) 82 67	114, 158, 231, 346	0
10	I	118/122 (96%)	-0.25	0 100 100	111, 149, 204, 280	0
11	J	65/70 (92%)	-0.13	2 (3%) 51 38	70, 112, 160, 176	0
12	K	114/120 (95%)	-0.39	0 100 100	96, 128, 186, 235	0
13	L	43/70 (61%)	-0.05	1 (2%) 61 45	89, 209, 283, 320	0
All	All	3590/4229 (84%)	-0.19	35 (0%) 79 63	61, 131, 221, 346	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	B	779	GLY	4.0
5	B	681	TRP	3.8
6	C	29	MET	3.3
5	B	381	MET	2.6
11	J	44	TYR	2.6

6.2 Non-standard residues in protein, DNA, RNA chains [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
2	8OG	T	19	23/24	0.88	0.08	106,122,143,171	0

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

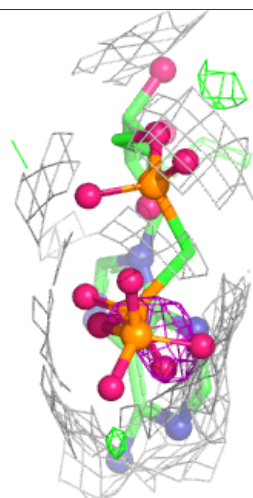
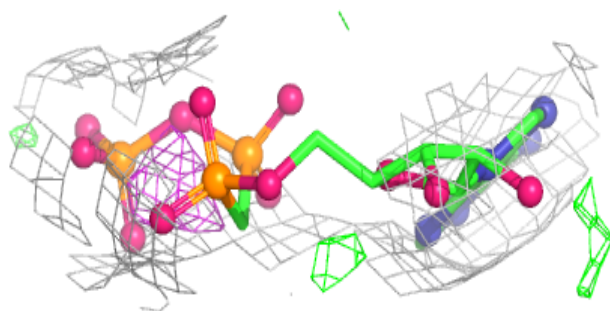
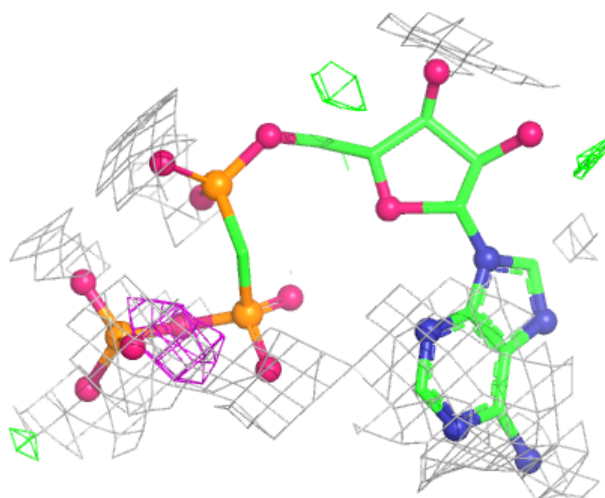
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
16	APC	B	1302	31/31	0.68	0.11	77,157,174,175	0
14	ZN	L	101	1/1	0.82	0.11	306,306,306,306	0
14	ZN	A	1802	1/1	0.91	0.08	187,187,187,187	0
14	ZN	A	1801	1/1	0.92	0.06	282,282,282,282	0
14	ZN	I	202	1/1	0.93	0.06	163,163,163,163	0
14	ZN	I	201	1/1	0.97	0.05	121,121,121,121	0
14	ZN	C	401	1/1	0.97	0.09	200,200,200,200	0
14	ZN	B	1301	1/1	0.98	0.03	230,230,230,230	0
15	MG	A	1803	1/1	0.98	0.06	79,79,79,79	0
14	ZN	J	101	1/1	0.98	0.04	82,82,82,82	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 APC B 1302:

$2mF_o-DF_c$ (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.