



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2026 – 10:06 PM UTC

PDB ID : 6UPY / pdb_00006upy
Title : RNA polymerase II elongation complex with 5-guanidinohydantoin lesion in state 2E
Authors : Oh, J.; Wang, D.
Deposited on : 2019-10-18
Resolution : 3.40 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

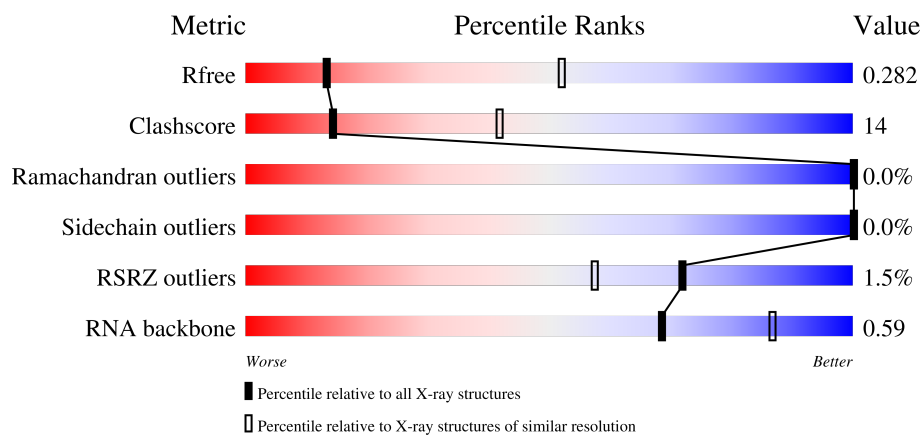
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1001 (3.44-3.36)
Clashscore	190562	1022 (3.44-3.36)
Ramachandran outliers	187476	1012 (3.44-3.36)
Sidechain outliers	187428	1012 (3.44-3.36)
RSRZ outliers	180081	1001 (3.44-3.36)
RNA backbone	3983	1157 (3.80-3.00)

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	67% 33%
2	T	29	38% 45% 14%
3	N	18	22% 61% 17%
4	A	1733	54% 25% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	
6	C	318	
7	E	215	
8	F	155	
9	H	146	
10	I	122	
11	J	70	
12	K	120	
13	L	70	

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 29013 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	25	Total	C	N	O	P	0	0	0
			500	239	78	158	25			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	15	Total	C	N	O	P	0	0	0
			317	148	71	83	15			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1384	Total	C	N	O	S	0	0	0
			10851	6845	1898	2048	60			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1109	Total	C	N	O	S	0	0	0
			8790	5566	1538	1633	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
			1740	1104	307	318	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	117	Total	C	N	O	S	0	0	0
			945	581	172	182	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
			337	208	66	59	4			

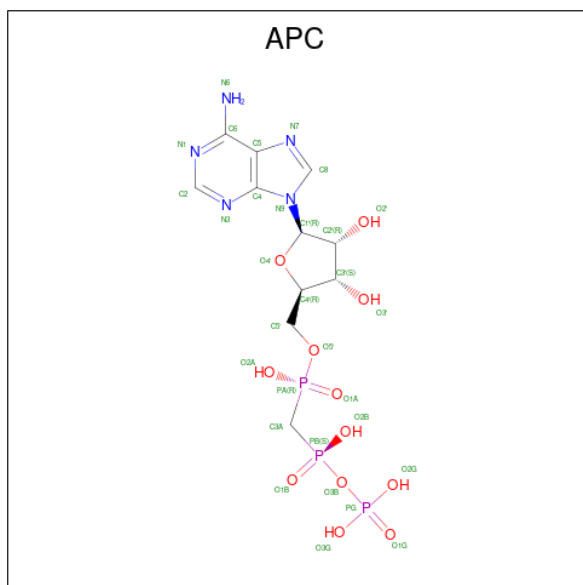
- Molecule 14 is ZINC ION (CCD ID: ZN) (formula: Zn).

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.

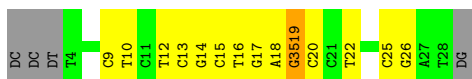
- Molecule 1: RNA

Chain R: 

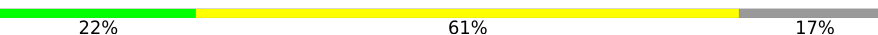


- Molecule 2: Template strand DNA

Chain T: 



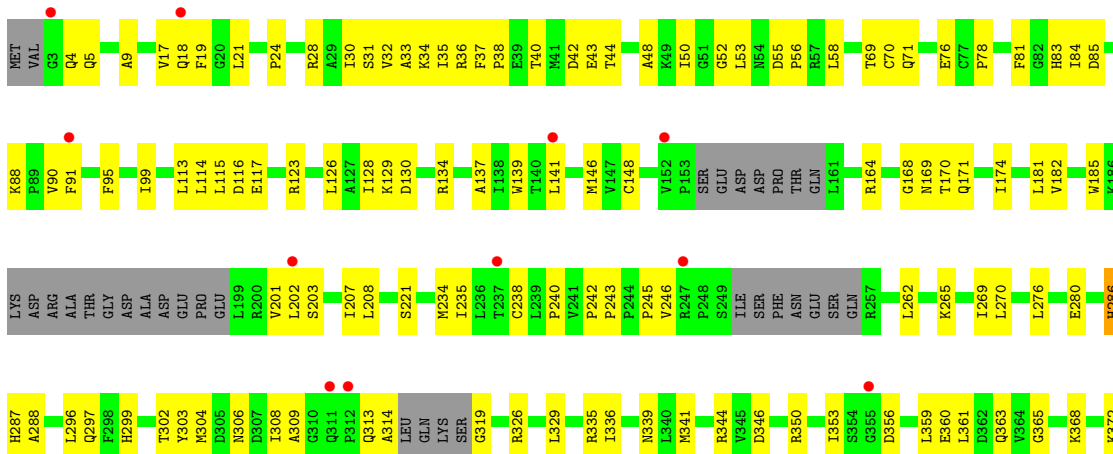
- Molecule 3: Non-template strand DNA

Chain N: 



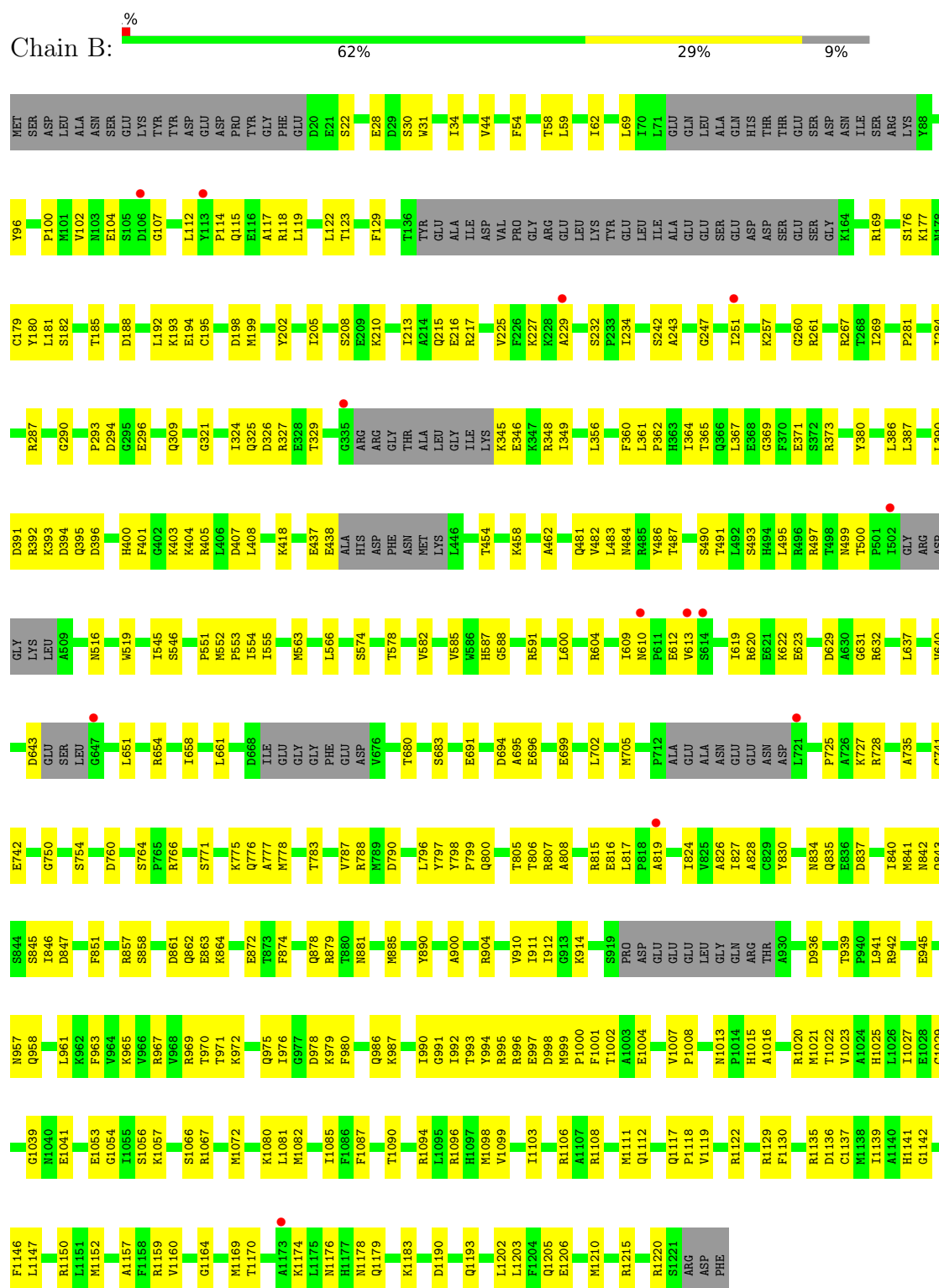
- Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain A: 

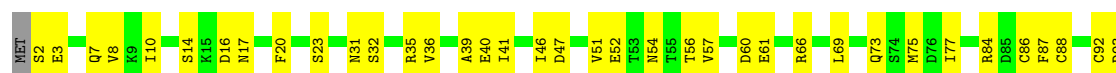


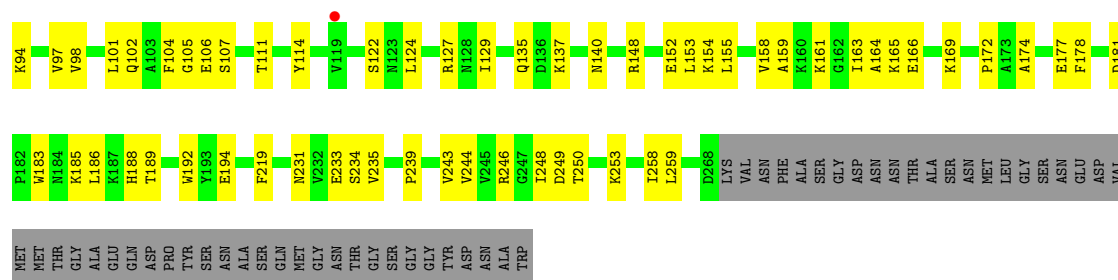
THR	SER	PRO	GLY	GLN	T1376	M1267	K1132	A1041	E932	Y836	V718	D592	F468	T373
SER	PRO	GLY	GLY	ASP	S1383	M1271	L1133	A1051	L936	I837	R731	T596	R469	L374
GLY	PRO	ALA	GLY	GLY	V1384	T1271	R1134		R939	L841	N736	L597	R470	T381
TYR	THR	VAL	VAL	THR	F1389	L1273	R1135	R1055	D940	K843	L737	L598	L472	Y382
PRO	SER	THR	THR	PRO				P1060	R930	Y844	K738	S599	N384	
PRO	PRO	SER	PRO	TYR	S1392	V1276	I1138	G1061	L943	D844	D739	P600	P477	R387
TYR	TYR	GLY	SER	SER	M1393	E1277	T1161	E1062	R944	E845	K744	R601	L388	L391
PRO	PRO	ASN	ASN	ASN	V1162	N1278	V1063	M1063	E945	E846	L740	D602		P386
ALA	PRO	GLY	GLY	GLY	I1163	I1279	V1064	V1065	V946	I848	N741		V491	H399
TYR	TYR	VAL	SER	SER	P1164		G1065	V1066	D949	Y852	N742	1607	D483	
SER	SER	GLY	GLY	LEU	I1165	P1292	V1067	L1067		D853	K745	1608		L391
PRO	PRO	LEU	D1166		D1166		L1068		P955	N854	M746	G610	M487	P386
ALA	PRO	VAL	I1169		I1169	T1295	A1068		L956	T855	M746	1612	V491	H399
GLN	GLY	ASN	L1176		L1176	E1297	I1072	G1073	P957	T856	F755	1613	P400	
ASP	ASP	ASP	LEU	ASP	ASP	E1303	E1074	E1075	V958	N857		1614		K403
LEU	LEU	LEU	ASP	LEU	ASP		P1075	P1075	R961	N858	T758		S502	
ASP	ASP	ASP	GLU	ASP	ASP	E1306	A1076	A1076	R962	S859	A759	V617	S503	I406
VAL	VAL	VAL	GLU	VAL	GLU	L1306	T1077	T1077	G861	G861	M761	1630	A506	R407
LYS	LYS	LYS	ALA	LYS	ALA	E1307	Q1078	Q1078	F971	M761	S762		V507	D408
THR	THR	THR	GLY	THR	GLY	T1308	M1079	T1080	H972	F866			P508	
PRO	PRO	LEU	GLN	LEU	GLN	N1312	T1080	T1081	I973	M873	V765	R635		I413
PRO	PRO	MET	SER	MET	SER		SER	L1081		D874		E636		D414
THR	THR	THR	PHE	THR	PHE	E1315	ASP	ASN	T976	A875	Q767		S513	
THR	THR	THR	ASP	THR	ASP	Q1187	ASP	THR	K977	A876	F646			L416
PRO	PRO	PRO	SER	SER	SER	V1316		PHE	S978	H877				R416
PRO	PRO	ALA	LEU	LEU	LEU	M1317	HIS	PHE	S979	E878	R774	Q650	K518	I424
VAL	VAL	VAL	LEU	LEU	LEU	T1318	PHE	PHE	D980	E879	T775	K651		
ASP	ASP	ASP	ALA	ASP	ALA	L1195	ALA	ALA	L981	Q881	A776	V652	I523	V430
PRO	PRO	PRO	VAL	PRO	VAL	M1202	VAL	VAL	T982		F777	V653	V524	K431
GLY	GLY	GLY	G1213	GLY	G1213	G1213	SER	SER	D985	I886	F779	G661	Q526	H435
ASP	ASP	ASP	V1212	ASP	V1212	I1327			I986	G887	D781	S663	T527	I436
ALA	ALA	ALA	G1214	ALA	G1214	E1328	K1092	K1093	L997	G888		D668	I531	D440
MET	MET	MET	E1214	MET	E1214	T1329	K1094	V1094	G1002	S889	F799	T669	R532	P441
GLY	GLY	GLY	I1216	GLY	I1216				K1003	D890	V800	1670	K533	V442
GLY	GLY	GLY	K1217	GLY	K1217	S1331	G1097		Q1008	R896	N802		L534	L443
THR	THR	THR	Q1218	THR	Q1218	T1219			Q1011	Y897	S803	T675	T535	F444
SER	SER	SER	F1332	SER	F1332		N1106	N1106	Q1011	R898	R806	N676	R537	N445
PRO	PRO	PRO	G1340	PRO	G1340		V1107	V1107	R1012	V899	G807			R446
THR	THR	THR	T1341	THR	T1341		A1108	A1108	R1012	D900	L808	1679	D544	S449
PRO	PRO	PRO	E1342	PRO	E1342		K1109	K1109	R1012	L901			Q545	L450
GLY	GLY	GLY	R1239	GLY	R1239		M1110	M1110	T1016	N902		1683	Q545	H451
TYR	TYR	TYR	R1345	TYR	R1345		M1111	M1111	T1016	L902	E812		M549	K452
ALA	ALA	ALA		ALA			K1112	K1112	L1021	N903				M453
PRO	PRO	PRO	I1356	PRO	I1356		T1113	T1113		H906	H816	V690	T566	M453
ASP	ASP	ASP		ASP					R1025	T907				S454
TYR	TYR	TYR	V1363	TYR	V1363		L1116	L1116	R1025					M455
GLY	GLY	GLY	N1364	GLY	N1364		T1117	T1117	R1029	L913	R821	A699	G574	M456
LYS	LYS	LYS	Y1365	LYS	Y1365		V1118	V1118	R1029	E914		N700	K575	A457
THR	THR	THR	R1366	THR	R1366		Y1119	Y1119	R1030	S915	I825	L701		
THR	THR	THR	H1367	THR	H1367		L1120	L1120	V1031				S579	V460
PRO	PRO	PRO	M1368	PRO	M1368		E1121	E1121	L1032	I919	K830	L710	G585	Y465
ILE	ILE	ILE	A1369	ILE	A1369		P1122	P1122	L1032	L920	T834	R711	S466	
GLY	GLY	GLY	L1370	GLY	L1370		H1124	H1124	R1036				S587	T467
ASP	ASP	ASP		ASP										
GLY	GLY	GLY		GLY										

● Molecule 5: DNA-directed RNA polymerase II subunit RPB2

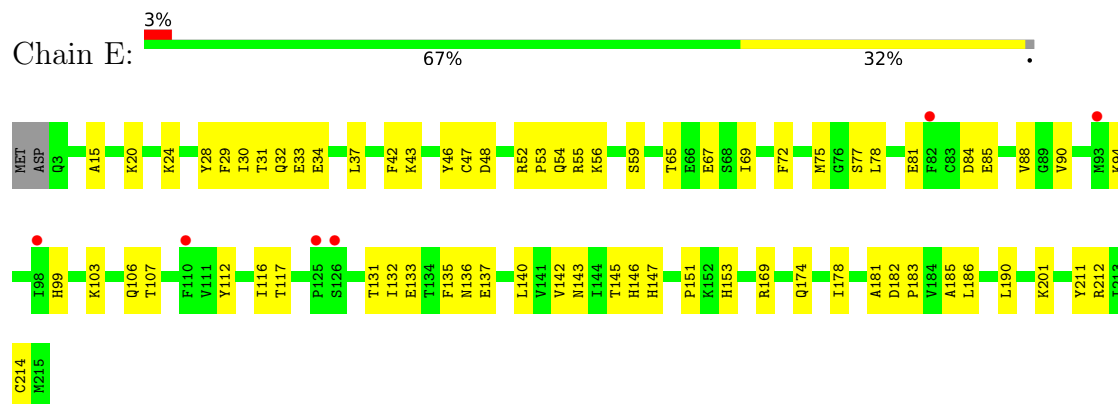


• Molecule 6: DNA-directed RNA polymerase II subunit RPB3

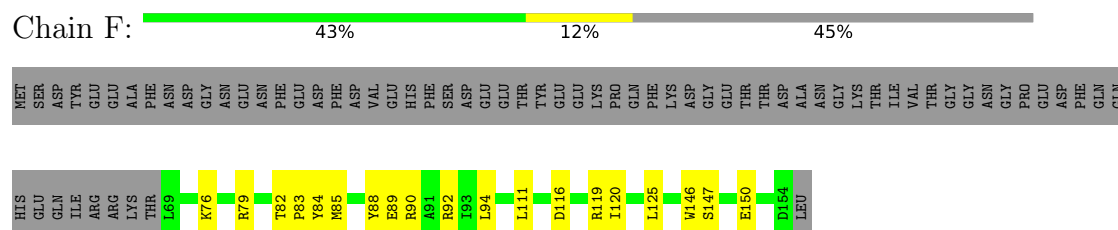




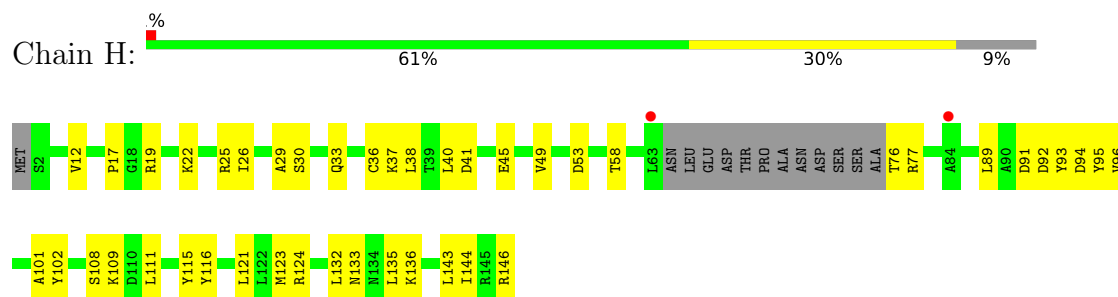
- Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1



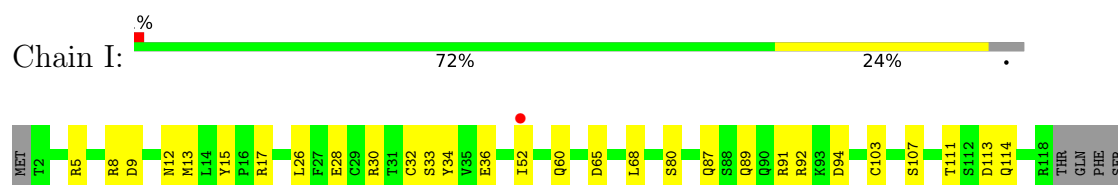
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2



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



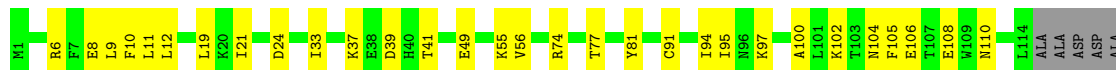
- Molecule 10: DNA-directed RNA polymerase II subunit RPB9



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

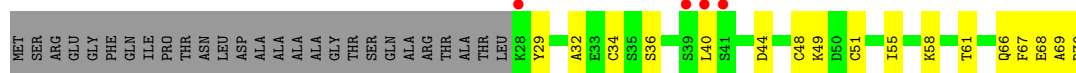


- Molecule 12: DNA-directed RNA polymerase II subunit RPB11



PHE

- 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	168.54Å 223.12Å 193.06Å 90.00° 100.92° 90.00°	Depositor
Resolution (Å)	49.22 – 3.40 49.22 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.22-3.40) 99.4 (49.22-3.40)	Depositor EDS
R_{merge}	0.45	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.38 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.227 , 0.282 0.229 , 0.282	Depositor DCC
R_{free} test set	2000 reflections (1.89%)	wwPDB-VP
Wilson B-factor (Å ²)	89.9	Xtriage
Anisotropy	0.531	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 90.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	29013	wwPDB-VP
Average B, all atoms (Å ²)	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, G35, APC, 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	R	0.11	0/223	0.23	0/345
2	T	0.20	0/529	0.46	0/809
3	N	0.22	0/359	0.38	0/553
4	A	0.14	0/11042	0.39	2/14932 (0.0%)
5	B	0.13	0/8961	0.33	0/12090
6	C	0.12	0/2139	0.32	0/2899
7	E	0.13	0/1776	0.35	0/2391
8	F	0.12	0/696	0.33	0/943
9	H	0.13	0/1076	0.36	0/1459
10	I	0.13	0/963	0.38	0/1298
11	J	0.14	0/541	0.32	0/727
12	K	0.11	0/937	0.29	0/1265
13	L	0.21	0/339	0.61	0/450
All	All	0.14	0/29581	0.36	2/40161 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
4	A	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1097	GLY	CA-C-N	5.06	123.43	120.24
4	A	1097	GLY	C-N-CA	5.06	123.43	120.24

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	1106	ASN	Peptide
4	A	524	VAL	Peptide
4	A	566	ILE	Peptide

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	R	199	0	98	7	0
2	T	500	0	287	16	0
3	N	317	0	166	11	0
4	A	10851	0	10916	348	0
5	B	8790	0	8781	272	0
6	C	2101	0	2056	84	0
7	E	1740	0	1761	54	0
8	F	684	0	692	14	0
9	H	1058	0	1018	34	0
10	I	945	0	890	27	0
11	J	532	0	543	32	0
12	K	919	0	929	24	0
13	L	337	0	352	13	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	14	2	0
All	All	29013	0	28503	804	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 14.

The worst 5 of 804 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)
16:B:1301:APC:O4'	16:B:1301:APC:C1'	1.63	1.25
9:H:108:SER:HB2	9:H:111:LEU:HD13	1.21	1.15
4:A:451:HIS:CE1	4:A:1074:GLU:HG3	1.88	1.07
4:A:443:LEU:HD12	5:B:1146:PHE:CZ	2.01	0.95
9:H:109:LYS:O	9:H:111:LEU:HD12	1.77	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	1368/1733 (79%)	1298 (95%)	69 (5%)	1 (0%)	48	78
5	B	1089/1224 (89%)	1043 (96%)	46 (4%)	0	100	100
6	C	265/318 (83%)	259 (98%)	6 (2%)	0	100	100
7	E	211/215 (98%)	204 (97%)	7 (3%)	0	100	100
8	F	84/155 (54%)	82 (98%)	2 (2%)	0	100	100
9	H	129/146 (88%)	123 (95%)	6 (5%)	0	100	100
10	I	115/122 (94%)	110 (96%)	5 (4%)	0	100	100
11	J	63/70 (90%)	63 (100%)	0	0	100	100
12	K	112/120 (93%)	108 (96%)	4 (4%)	0	100	100
13	L	41/70 (59%)	36 (88%)	5 (12%)	0	100	100
All	All	3477/4173 (83%)	3326 (96%)	150 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	286	HIS

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	1200/1520 (79%)	1200 (100%)	0	100	100
5	B	955/1061 (90%)	954 (100%)	1 (0%)	88	89
6	C	235/274 (86%)	235 (100%)	0	100	100
7	E	194/197 (98%)	194 (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	37/57 (65%)	37 (100%)	0	100	100
All	All	3077/3657 (84%)	3076 (100%)	1 (0%)	100	100

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	827	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	309	GLN
12	K	112	GLN
5	B	800	GLN
9	H	131	ASN
5	B	590	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/9 (88%)	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](#)

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	G35	T	19	2	18,23,24	4.83	14 (77%)	18,33,36	1.93	3 (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
2	G35	T	19	2	-	3/11/41/42	0/2/2/2

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	19	G35	C2-N3	9.53	1.49	1.33
2	T	19	G35	O4'-C4'	7.93	1.62	1.45
2	T	19	G35	C3'-C4'	-6.74	1.35	1.53
2	T	19	G35	C8-N9	6.56	1.46	1.37
2	T	19	G35	C5-N7	6.40	1.46	1.37

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	19	G35	O4'-C1'-N9	6.07	115.85	108.65

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	19	G35	C4'-O4'-C1'	-3.83	100.40	109.51
2	T	19	G35	O4'-C1'-C2'	-3.18	100.30	106.25

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	T	19	G35	C3'-C4'-C5'-O5'
2	T	19	G35	C4'-C5'-O5'-P
2	T	19	G35	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	T	19	G35	2	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	1301	-	29,33,33	3.38	9 (31%)	44,52,52	2.11	13 (29%)

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	1301	-	-	9/19/38/38	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	1301	APC	O4'-C1'	9.44	1.63	1.42
16	B	1301	APC	PB-O3B	8.25	1.67	1.58
16	B	1301	APC	C2'-C1'	-7.03	1.31	1.53
16	B	1301	APC	O4'-C4'	-6.48	1.30	1.45
16	B	1301	APC	O2'-C2'	4.39	1.53	1.43

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	1301	APC	N3-C2-N1	-5.66	120.02	128.58
16	B	1301	APC	N6-C6-N1	-4.98	107.28	118.38
16	B	1301	APC	C5-C4-N3	-4.87	120.01	126.72
16	B	1301	APC	N9-C8-N7	-4.28	107.86	113.94
16	B	1301	APC	C5-C6-N6	3.84	132.80	123.29

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	B	1301	APC	PB-C3A-PA-O1A
16	B	1301	APC	PB-C3A-PA-O2A
16	B	1301	APC	PB-C3A-PA-O5'
16	B	1301	APC	O4'-C4'-C5'-O5'
16	B	1301	APC	C3'-C4'-C5'-O5'

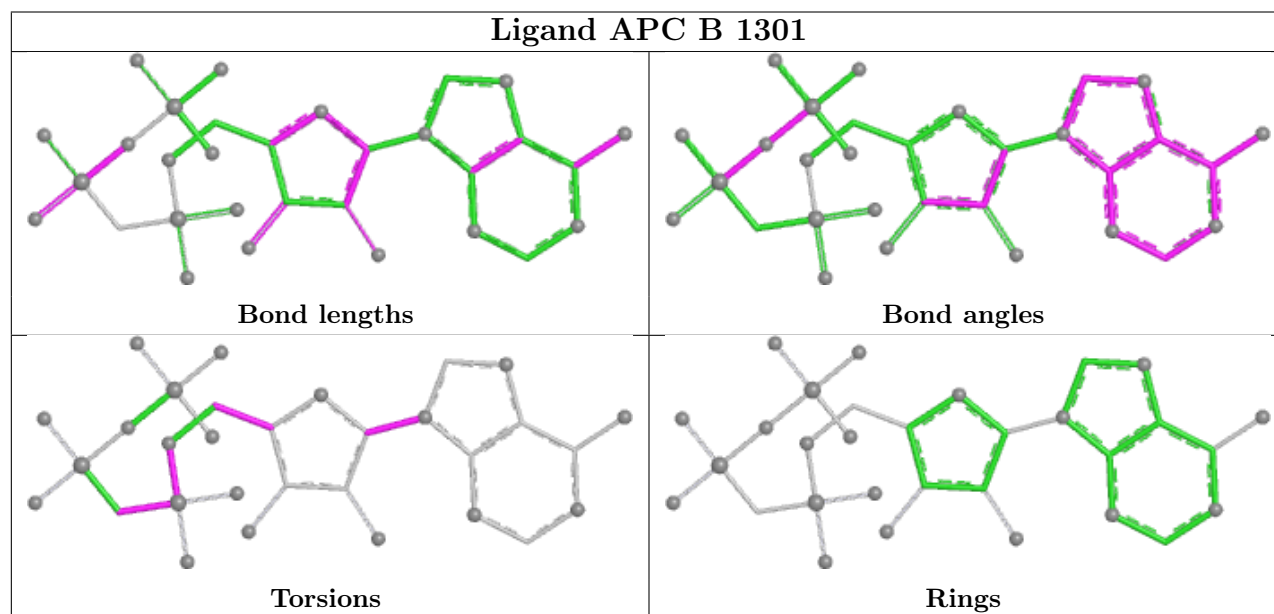
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	1301	APC	2	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	R	9/9 (100%)	-0.49	0 100 100	89, 97, 145, 188	0
2	T	24/29 (82%)	-0.08	0 100 100	86, 171, 219, 235	0
3	N	15/18 (83%)	0.16	0 100 100	183, 203, 251, 254	0
4	A	1384/1733 (79%)	0.04	25 (1%) 67 53	51, 104, 184, 265	0
5	B	1109/1224 (90%)	-0.00	13 (1%) 76 63	32, 85, 153, 231	0
6	C	267/318 (83%)	-0.15	1 (0%) 88 81	56, 92, 138, 180	0
7	E	213/215 (99%)	0.15	6 (2%) 55 41	85, 140, 236, 292	0
8	F	86/155 (55%)	-0.17	0 100 100	66, 107, 149, 209	0
9	H	133/146 (91%)	0.05	2 (1%) 72 57	82, 125, 183, 224	0
10	I	117/122 (95%)	-0.08	1 (0%) 81 68	73, 106, 144, 167	0
11	J	65/70 (92%)	-0.09	1 (1%) 72 57	42, 73, 127, 145	0
12	K	114/120 (95%)	-0.30	0 100 100	62, 93, 131, 180	0
13	L	43/70 (61%)	0.71	4 (9%) 14 14	58, 146, 220, 268	0
All	All	3579/4229 (84%)	0.01	53 (1%) 72 57	32, 100, 186, 292	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	B	647	GLY	4.6
4	A	91	PHE	4.5
5	B	335	GLY	3.9
7	E	98	ILE	3.7
5	B	721	LEU	3.1

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($\text{\AA}^2$)	Q<0.9
2	G35	T	19	22/23	0.80	0.16	138,155,172,184	0

6.3 Carbohydrates [i](#)

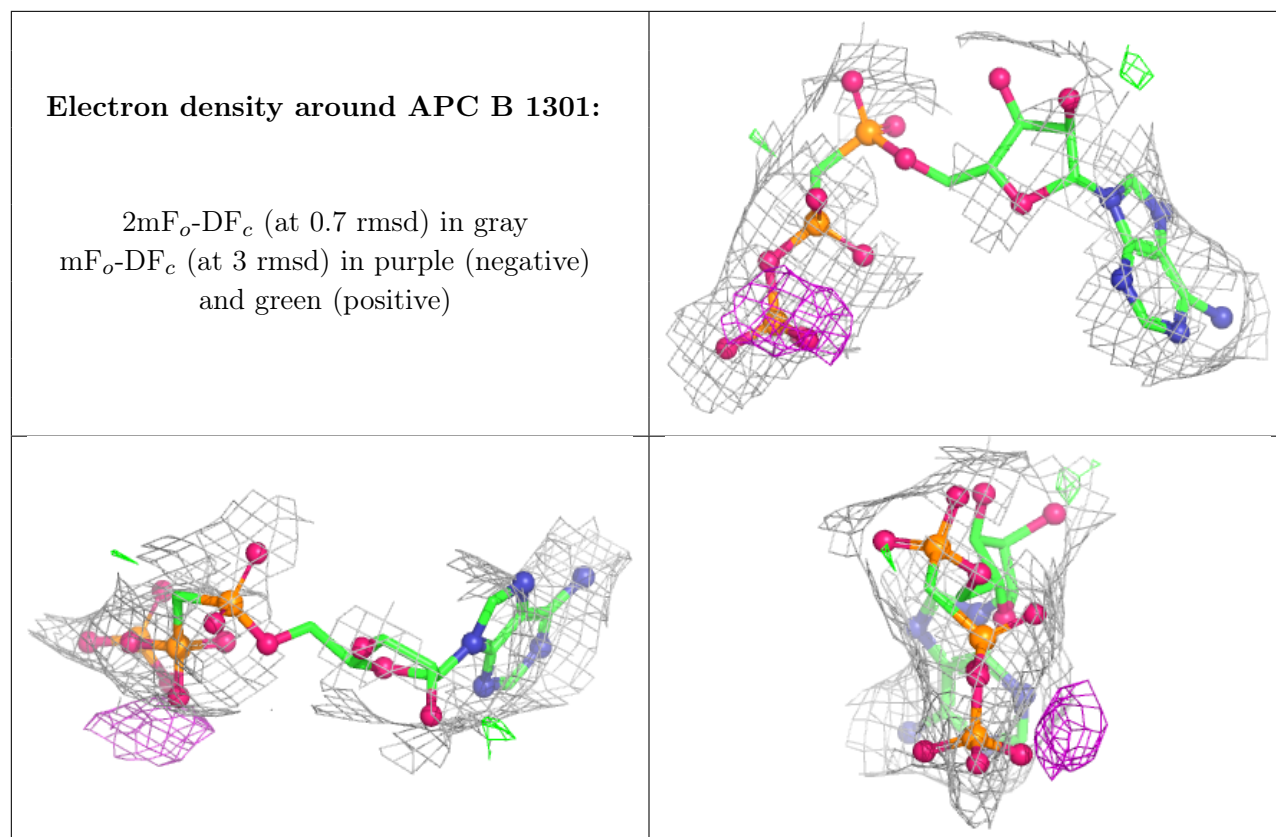
There are no oligosaccharides in this entry.

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($\text{\AA}^2$)	Q<0.9
16	APC	B	1301	31/31	0.70	0.14	131,160,243,251	0
14	ZN	A	1801	1/1	0.94	0.06	393,393,393,393	0
14	ZN	L	101	1/1	0.95	0.07	231,231,231,231	0
14	ZN	A	1802	1/1	0.98	0.04	102,102,102,102	0
14	ZN	I	201	1/1	0.99	0.06	110,110,110,110	0
14	ZN	B	1302	1/1	0.99	0.03	209,209,209,209	0
15	MG	A	1803	1/1	0.99	0.02	39,39,39,39	0
14	ZN	C	401	1/1	0.99	0.03	131,131,131,131	0
14	ZN	J	101	1/1	1.00	0.05	79,79,79,79	0
14	ZN	I	202	1/1	1.00	0.02	80,80,80,80	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.



6.5 Other polymers [i](#)

There are no such residues in this entry.