wwPDB X-ray Structure Validation Summary Report (i)

May 24, 2020 – 07:43 pm BST

PDB ID	:	4BY1
Title	:	elongating RNA Polymerase II-Bye1 TLD complex soaked with AMPCPP
Authors	:	Kinkelin, K.; Wozniak, G.G.; Rothbart, S.B.; Lidschreiber, M.; Strahl, B.D.; Cramer, P.
Deposited on	:	2013-07-17
Resolution	:	3.60 Å(reported)

The following versions of software and data (see references (i)) 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	:	2.11
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)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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.
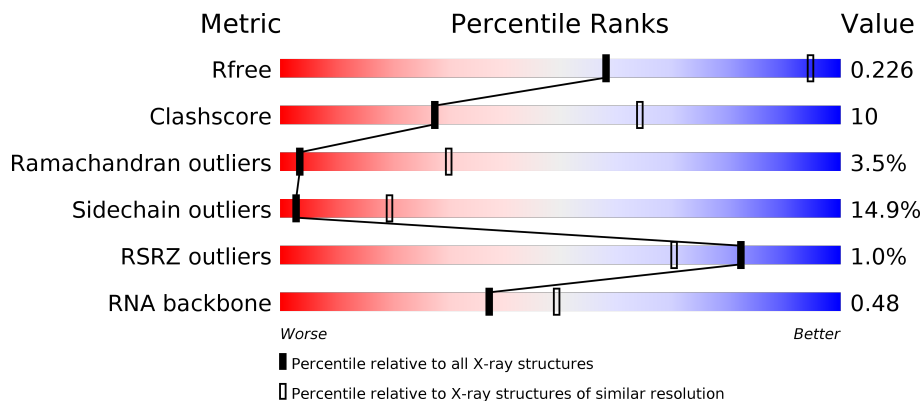
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)
RNA backbone	3102	1017 (4.20-3.00)

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 on 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 <=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	1733	53% 24% 5% 18%
2	B	1224	% 61% 25% • 10%
3	C	318	55% 26% • 16%
4	D	221	57% 22% • 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	215	71% 24% 5%
6	F	155	32% 21% 45%
7	G	171	60% 33% 7%
8	H	146	53% 31% 7% 9%
9	I	122	76% 19%
10	J	70	37% 43% 11% 7%
11	K	120	64% 29%
12	L	70	30% 26% 10% 34%
13	N	14	50% 21% 7% 21%
14	P	11	36% 27% 9% 18% 9%
15	T	26	58% 27% 8% 8%
16	X	146	42% 34% 21%

2 Entry composition ⓘ

There are 19 unique types of molecules in this entry. The entry contains 33026 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 II SUBUNIT RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1427	Total	C	N	O	S	0	0	0
			11230	7072	1960	2136	62			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1106	Total	C	N	O	S	0	0	0
			8793	5569	1538	1631	55			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	177	Total	C	N	O	S	0	0	0
			1356	840	241	273	2			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUB-UNIT RPABC 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUB-UNIT RPABC 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	85	Total 688	C 439	N 116	O 130	S 3	0	0	0

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total 1340	C 861	N 222	O 249	S 8	0	0	0

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUB-UNIT RPABC 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	133	Total 1068	C 673	N 180	O 211	S 4	0	0	0

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	119	Total 971	C 596	N 179	O 186	S 10	0	0	0

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	115	Total 920	C 590	N 157	O 171	S 2	0	0	1

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total 363	C 224	N 72	O 63	S 4	0	0	0

- Molecule 13 is a DNA chain called 5'-D(*AP*AP*AP*GP*TP*AP*CP*TP*TP*GP*AP*G

P*CP*TP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	11	Total 229	C 109	N 44	O 65	P 11	0	0	0

- Molecule 14 is a RNA chain called 5'-R(*UP*UP*CP*GP*AP*CP*CP*AP*GP*GP*AP)-3 '.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	P	10	Total 215	C 96	N 41	O 68	P 10	0	0	0

- Molecule 15 is a DNA chain called 5'-D(*AP*GP*CP*TP*CP*AP*AP*GP*TP*AP*CP*T P*TP*AP *TP*TP*CP*CP*BRUP*GP*GP*TP*CP*AP*AP*T)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
15	T	24	Total 481	Br 1	C 230	N 75	O 151	P 24	0	0	0

- Molecule 16 is a protein called TRANSCRIPTION FACTOR BYE1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	X	116	Total 953	C 611	N 158	O 181	S 3	0	0	0

- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 18 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 19 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: $C_{11}H_{18}N_5O_{12}P_3$).

APC
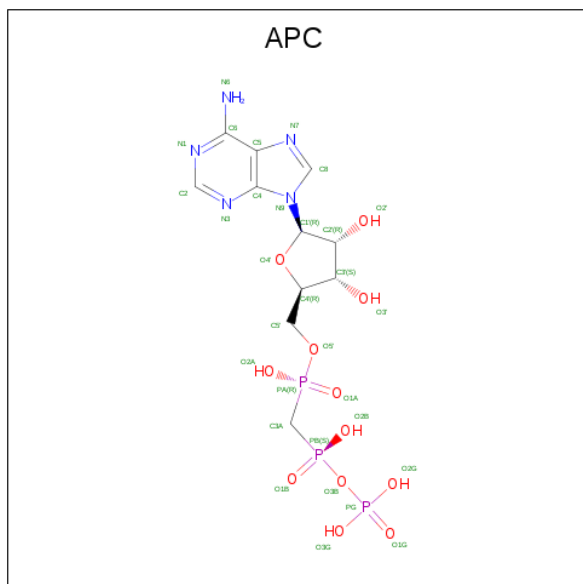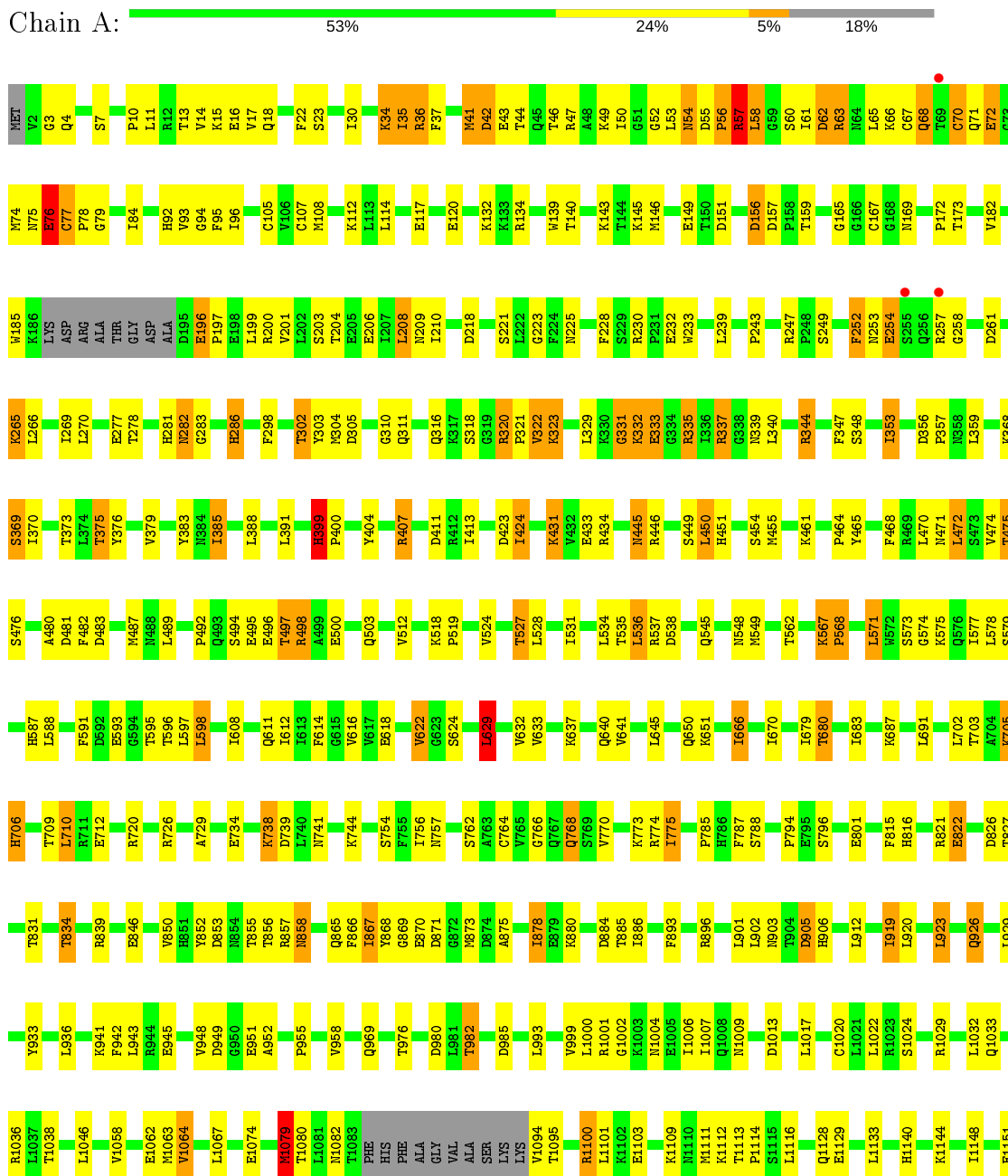
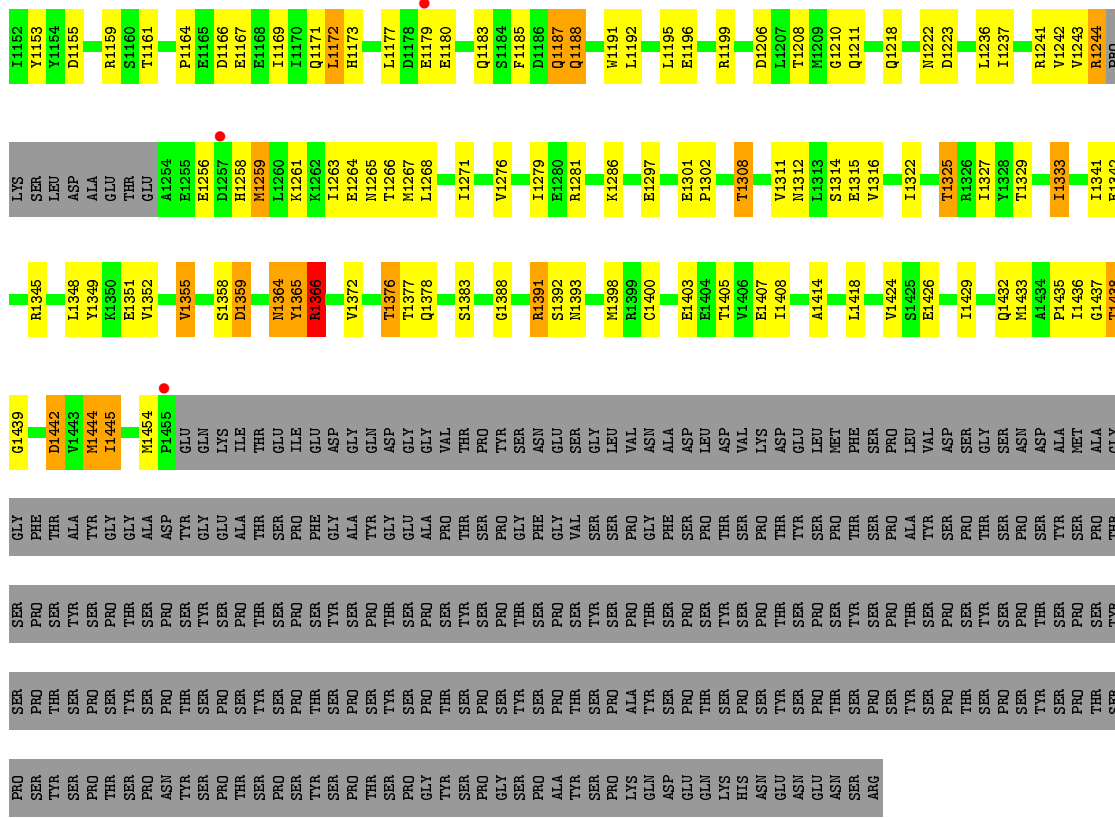| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
19	P	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
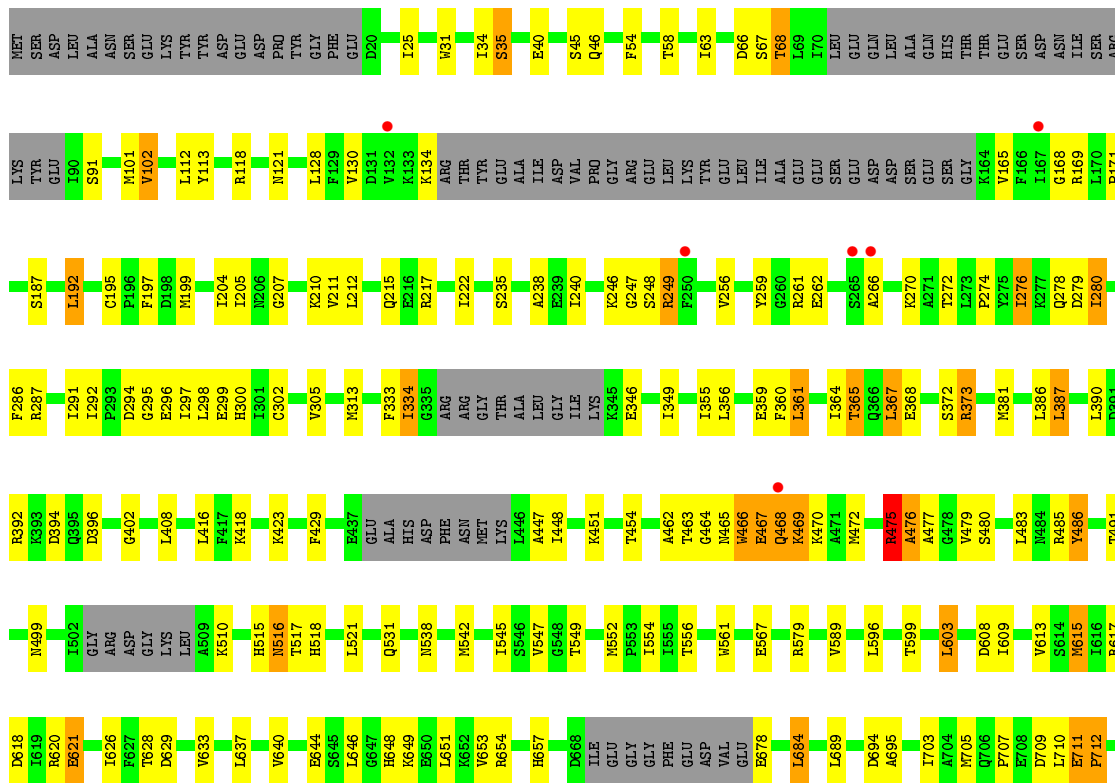
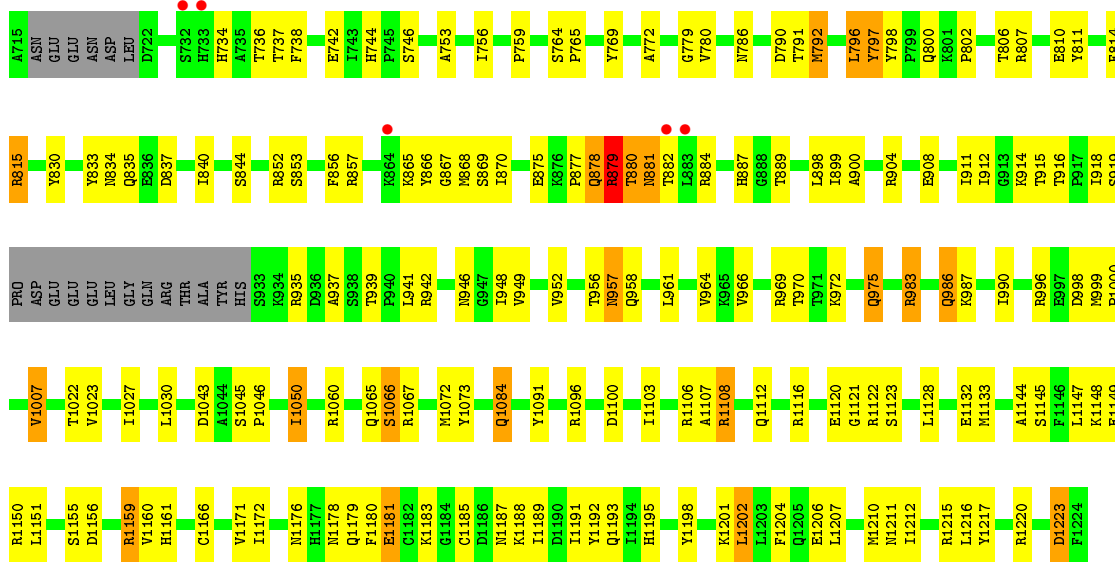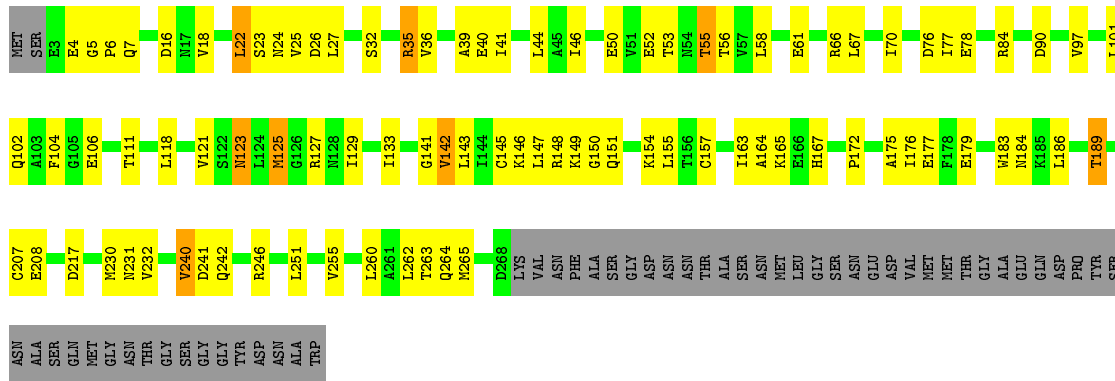- Molecule 1: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1

Chain A:

I1152 Y1153 Y1154 D1155 R1159 S1160 T1161 P1164 E1165 D1166 E1167 I1168 I1169 I1170 Q1171 L1172 H1173 L1177 D1178 E1179 E1180 Q1183 S1184 F1185 D1186 Q1187 Q1188 W1191 L1192 L1195 E1196 R1199 D1206 L1207 T1208 M1209 G1210 Q1211 Q1218 N1222 D1223 L1236 I1237 R1241 V1242 V1243 R1244 PRO

LYS SER LEU ASP ALA GLU THR GLU A1254 E1255 E1256 D1257 H1258 M1259 L1260 K1261 K1262 I1263 E1264 N1265 T1266 M1267 L1268 I1271 V1276 I1279 E1280 R1281 K1286 E1297 E1301 P1302 T1308 V1311 N1312 L1313 S1314 E1315 V1316 I1322 T1325 R1326 Y1327 Y1328 T1329 I1333 I1341 E1342

R1345 L1348 Y1349 K1350 E1351 V1352 V1355 S1358 D1359 N1364 Y1365 R1366 V1372 T1376 T1377 Q1378 S1383 G1388 R1391 S1392 N1393 M1398 R1399 C1400 E1403 E1404 T1405 V1406 E1407 I1408 A1414 L1418 V1424 S1425 E1426 I1429 Q1432 M1433 A1434 P1435 I1436 G1437 T1438

G1439 D1442 V1443 M1444 I1445 M1454 P1455 GLU GLN LYS THR GLU GLU ASP GLY GLN ALA ASP GLY VAL THR PRO TYR SER ASN GLU GLY LEU VAL ASN ALA ASP LEU LEU VAL LYS ASP GLU MET PHE SER PRO VAL ASP SER GLY ASN ASP MET ALA GLY

GLY PHE THR ALA TYR GLY GLY ALA ASP GLY TYR GLU ALA THR GLU GLU SER PHE GLY ALA TYR GLY GLU ALA PRO SER PRO GLY PHE VAL SER GLY PHE PRO SER PRO SER TYR SER THR SER PRO ALA TYR PRO THR SER PRO SER TYR SER PRO THR

SER PRO SER TYR PRO SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER TYR

SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER PRO SER TYR SER PRO SER

PRO SER TYR SER PRO THR SER PRO ASN TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO GLY TYR SER PRO ALA TYR SER PRO LYS GLN ASP GLU GLN LYS HIS ASN GLU ASN GLU ASN SER ARG

● **Molecule 2: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2**

%

Chain B:

61%   25%   ●   10%

MET SER ASP ALA ASN SER GLU LYS TYR TYR ASP GLU ASP PRO TYR GLY PHE GLU D20 I25 W31 I34 S35 E40 S45 Q46 F54 T58 I63 D66 S67 T68 L69 I70 LEU GLU GLN LEU ALA GLN HIS THR THR GLU ASP ASN ILE SER ARG

LYS TYR GLU I90 S91 M101 V102 L112 Y113 R118 N121 L128 F129 V130 D131 V132 K133 K134 ARG THR TYR GLU ALA ILE ASP VAL PRO GLY ARG GLU GLU LYS GLU ILE ALA GLU GLU ILE ILE TYR LEU LEU SER GLU ASP ASP SER GLU LEU SER GLY K164 V165 F166 I167 G168 R169 L170 P171

S187 L192 C195 P196 F197 D198 M199 I204 I205 N206 G207 K210 V211 L212 Q215 E216 R217 I222 S235 A238 E239 I240 K246 G247 R249 F250 V256 Y259 G260 R261 E262 S265 A266 K270 A271 T272 L273 P274 Y275 I276 K277 D278 D279 I280

F286 R287 I291 I292 P293 D294 G295 E296 I297 L298 E299 H300 I301 C302 V305 M313 F333 I334 G335 ARG GLY THR ALA LEU GLY ILE LYS K345 E346 I349 I355 L356 E359 F360 L361 I364 T365 Q366 L367 E368 S372 R373 M381 L386 L387 L390 D391

R392 K393 D394 Q395 D396 G402 L408 L416 F417 K418 K423 F429 E437 GLU ALA HIS ASP PHE ASN MET LYS L446 A447 I448 K451 T454 A462 T463 G464 N465 W466 E467 Q468 K469 K470 A471 M472 E475 A476 A477 G478 V479 S480 L483 N484 R485 Y486 T491

N499 I502 GLY ARG ASP GLY LYS LEU A509 K510 H515 N516 T517 H518 L521 Q531 N538 M542 P553 I554 I555 T556 W561 E567 R579 V589 L596 T599 L603 D608 I609 V613 S614 M615 I616 R617

D618 I619 R620 E621 I626 F627 T628 D629 V633 L637 V640 E644 S645 L646 G647 H648 K649 E650 L651 K652 V653 R654 H657 D668 ILE GLU GLY GLY PHE GLU ASP VAL GLU E678 L684 L689 D694 A695 I703 A704 M705 Q706 P707 E708 D709 L710 E711 F712

WORLDWIDE
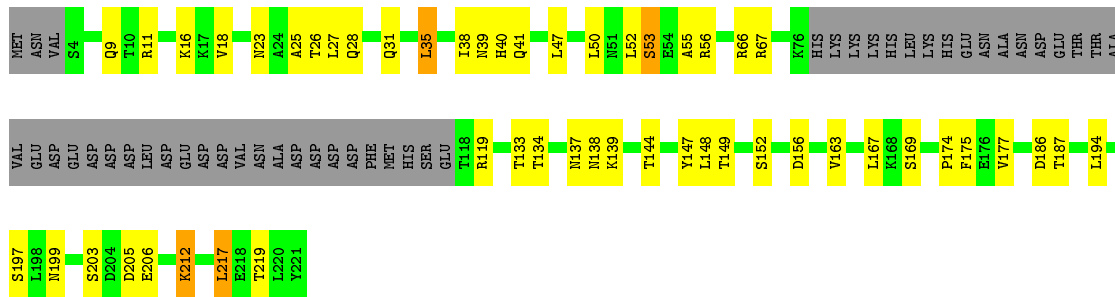wwPDB
PROTEIN DATA BANK

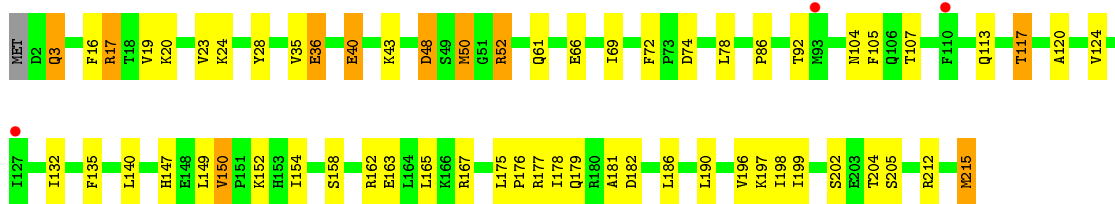• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

Chain C:
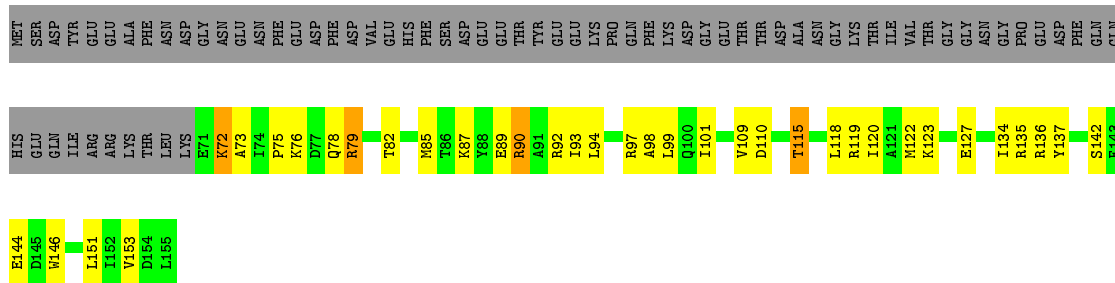

• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4

Chain D:

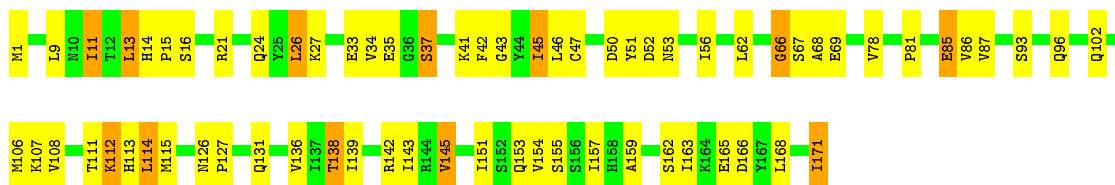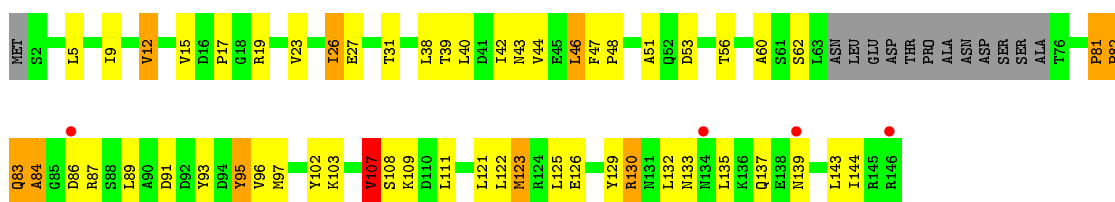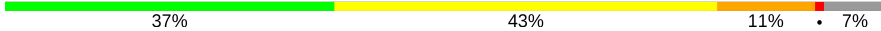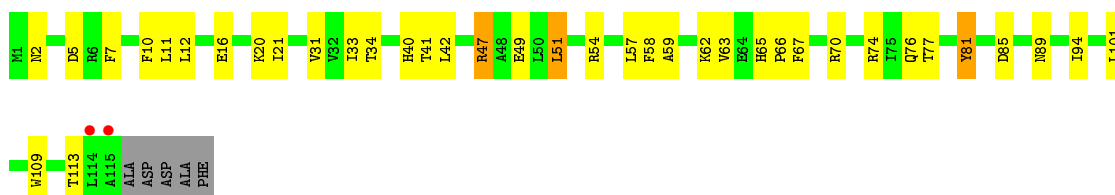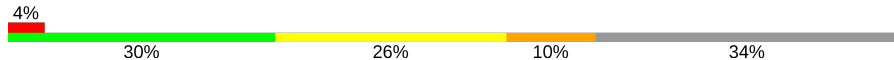• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1

Chain E:

Chain F residue diagram (MET D2 Q3 F16 R17 T18 V19 K20 V23 K24 Y28 V35 E36 E40 K43 D48 S49 M50 G51 R52 Q61 E66 I69 F72 P73 D74 L78 P86 T92 M93 N104 F105 Q106 T107 F110 Q113 T117 A120 V124 ... I127 I132 F135 L140 H147 E148 L149 V150 P151 K152 H153 I154 S158 R162 E163 L164 L165 K166 R167 L175 P176 R177 I178 Q179 R180 A181 D182 L186 L190 V196 K197 I198 I199 S203 E203 T204 S205 R212 M215)

● Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

Chain F: 32%   21%   ·   45%

● Molecule 7: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7

Chain G: 60%   33%   7%

● Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H: 3%   53%   31%   7%   ·   9%

● Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

Chain I: 3%   76%   19%   · ·

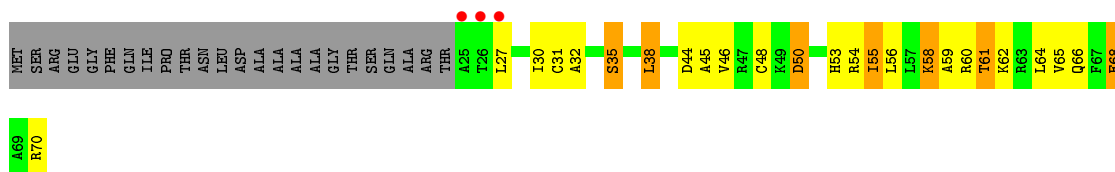● Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

Chain J:

• Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11
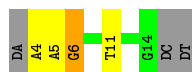
Chain K:

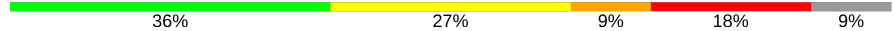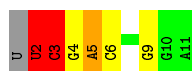• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain L:

• Molecule 13: 5'-D(*AP*AP*AP*GP*TP*AP*CP*TP*TP*GP*AP*GP*CP*TP)-3'

Chain N:

• Molecule 14: 5'-R(*UP*UP*CP*GP*AP*CP*CP*AP*GP*GP*AP)-3'

Chain P:

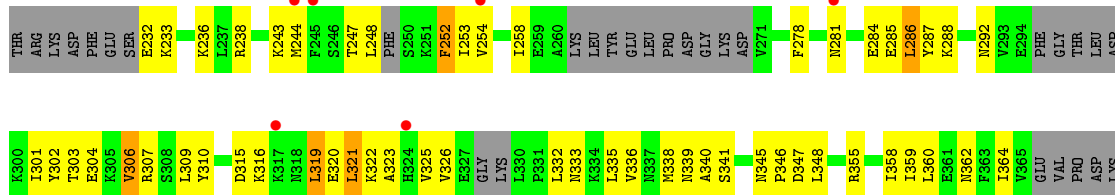• Molecule 15: 5'-D(*AP*GP*CP*TP*CP*AP*AP*GP*TP*AP*CP*TP*TP*AP *TP*TP*CP* CP*BRUP*GP*GP*TP*CP*AP*AP*T)-3'

Chain T:

• Molecule 16: TRANSCRIPTION FACTOR BYE1

Chain X:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	222.24Å 391.58Å 281.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.95 – 3.60 48.95 – 3.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.95-3.60) 100.0 (48.95-3.60)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.04 (at 3.57Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.175 , 0.206 0.195 , 0.226	Depositor DCC
R_{free} test set	2784 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å^2)	106.1	Xtriage
Anisotropy	0.211	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 113.2	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.019 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.027 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	33026	wwPDB-VP
Average B, all atoms (Å^2)	125.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.80% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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, APC, ZN, BRU

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	0.54	0/11431	0.81	9/15462 (0.1%)
2	B	0.50	0/8963	0.77	1/12085 (0.0%)
3	C	0.48	0/2133	0.75	0/2891
4	D	0.51	0/1365	0.80	2/1837 (0.1%)
5	E	0.49	0/1788	0.71	0/2406
6	F	0.58	0/700	0.80	0/945
7	G	0.50	0/1368	0.79	0/1844
8	H	0.48	0/1086	0.80	0/1470
9	I	0.47	0/989	0.72	0/1331
10	J	0.52	0/541	0.85	0/727
11	K	0.45	0/938	0.68	0/1267
12	L	0.59	0/365	0.97	0/485
13	N	1.09	0/257	1.07	1/395 (0.3%)
14	P	1.35	1/240 (0.4%)	1.04	1/372 (0.3%)
15	T	1.25	1/512 (0.2%)	1.06	0/783
16	X	0.54	0/963	0.77	0/1287
All	All	0.55	2/33639 (0.0%)	0.80	14/45587 (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
14	P	0	1
15	T	0	3
All	All	0	4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	P	2	U	C3'-O3'	6.44	1.51	1.42
15	T	27	DA	C3'-O3'	5.33	1.50	1.44

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	399	HIS	N-CA-CB	7.16	123.48	110.60
1	A	56	PRO	C-N-CA	6.69	138.42	121.70
4	D	25	ALA	C-N-CA	5.76	136.11	121.70
1	A	3	GLY	C-N-CA	5.59	135.67	121.70
1	A	223	GLY	C-N-CA	5.46	135.35	121.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
14	P	3	C	Sidechain
15	T	18	DT	Sidechain
15	T	23	DG	Sidechain
15	T	28	DT	Sidechain

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	A	11230	0	11280	284	0
2	B	8793	0	8824	166	0
3	C	2095	0	2051	49	0
4	D	1356	0	1319	19	0
5	E	1752	0	1776	40	0
6	F	688	0	707	22	0
7	G	1340	0	1357	35	0
8	H	1068	0	1040	38	0
9	I	971	0	927	8	0
10	J	532	0	542	24	0
11	K	920	0	929	24	0
12	L	363	0	386	9	0
13	N	229	0	125	3	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	P	215	0	109	5	0
15	T	481	0	264	7	0
16	X	953	0	969	23	0
17	A	2	0	0	0	0
17	B	1	0	0	0	0
17	C	1	0	0	0	0
17	I	2	0	0	0	0
17	J	1	0	0	0	0
17	L	1	0	0	0	0
18	A	1	0	0	0	0
19	P	31	0	14	0	0
All	All	33026	0	32619	656	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 10.

The worst 5 of 656 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:1172:ILE:CG1	2:B:1172:ILE:CD1	1.76	1.58
1:A:84:ILE:CG1	1:A:84:ILE:CD1	1.75	1.56
1:A:853:ASP:OD1	1:A:855:THR:HG22	1.55	1.04
1:A:1187:GLN:HA	1:A:1188:GLN:HB2	1.42	0.98
1:A:14:VAL:H	1:A:1432:GLN:HE22	1.17	0.93

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1419/1733 (82%)	1223 (86%)	142 (10%)	54 (4%)	3	27

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	1086/1224 (89%)	949 (87%)	99 (9%)	38 (4%)	3	30
3	C	264/318 (83%)	229 (87%)	31 (12%)	4 (2%)	10	47
4	D	173/221 (78%)	155 (90%)	11 (6%)	7 (4%)	3	26
5	E	212/215 (99%)	199 (94%)	11 (5%)	2 (1%)	17	57
6	F	83/155 (54%)	75 (90%)	7 (8%)	1 (1%)	13	51
7	G	169/171 (99%)	153 (90%)	14 (8%)	2 (1%)	13	51
8	H	129/146 (88%)	102 (79%)	18 (14%)	9 (7%)	1	14
9	I	117/122 (96%)	93 (80%)	23 (20%)	1 (1%)	17	57
10	J	63/70 (90%)	51 (81%)	6 (10%)	6 (10%)	0	8
11	K	113/120 (94%)	106 (94%)	6 (5%)	1 (1%)	17	57
12	L	44/70 (63%)	33 (75%)	3 (7%)	8 (18%)	0	2
16	X	106/146 (73%)	90 (85%)	9 (8%)	7 (7%)	1	16
All	All	3978/4711 (84%)	3458 (87%)	380 (10%)	140 (4%)	3	30

5 of 140 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	35	ILE
1	A	57	ARG
1	A	58	LEU
1	A	76	GLU

5.3.2   Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1249/1520 (82%)	1042 (83%)	207 (17%)	2	15
2	B	960/1061 (90%)	827 (86%)	133 (14%)	3	22
3	C	234/274 (85%)	205 (88%)	29 (12%)	4	26
4	D	140/200 (70%)	123 (88%)	17 (12%)	5	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	196/197 (100%)	176 (90%)	20 (10%)	7	34
6	F	75/137 (55%)	64 (85%)	11 (15%)	3	20
7	G	152/152 (100%)	130 (86%)	22 (14%)	3	20
8	H	117/128 (91%)	100 (86%)	17 (14%)	3	20
9	I	113/116 (97%)	98 (87%)	15 (13%)	4	23
10	J	60/65 (92%)	46 (77%)	14 (23%)	1	5
11	K	99/102 (97%)	87 (88%)	12 (12%)	5	26
12	L	40/57 (70%)	28 (70%)	12 (30%)	0	2
16	X	107/136 (79%)	88 (82%)	19 (18%)	2	11
All	All	3542/4145 (86%)	3014 (85%)	528 (15%)	3	19

5 of 528 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	387	LEU
2	B	889	THR
11	K	47	ARG
2	B	468	GLN
2	B	629	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 75 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	366	GLN
2	B	744	HIS
9	I	12	ASN
2	B	383	ASN
2	B	516	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	10/11 (90%)	4 (40%)	1 (10%)

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	3	C
14	P	4	G
14	P	5	A
14	P	9	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	P	2	U

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	$\#|Z| > 2$
15	BRU	T	22	15,14	15,21,22	2.02	2 (13%)	17,30,33	2.21	3 (17%)

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
15	BRU	T	22	15,14	-	0/4/21/22	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	22	BRU	C4-C5	6.92	1.47	1.38
15	T	22	BRU	C2-N3	-2.60	1.33	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	T	22	BRU	C4-N3-C2	7.40	121.39	115.14
15	T	22	BRU	C5-C4-N3	-4.22	118.59	123.64
15	T	22	BRU	C2'-C1'-N1	-2.09	109.46	114.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	T	22	BRU	1	0

5.5 Carbohydrates (i)

There are no carbohydrates 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	#$|Z| > 2$
19	APC	P	12	18	27,33,33	1.74	7 (25%)	31,52,52	1.44	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
19	APC	P	12	18	-	6/15/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
19	P	12	APC	PA-O1A	4.37	1.62	1.51
19	P	12	APC	PA-O2A	-3.29	1.48	1.56
19	P	12	APC	PB-O3B	3.27	1.62	1.58
19	P	12	APC	PA-O5'	2.93	1.61	1.57
19	P	12	APC	C5-C4	2.61	1.47	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
19	P	12	APC	C4-C5-N7	-4.27	104.95	109.40
19	P	12	APC	N3-C2-N1	-3.49	123.23	128.68
19	P	12	APC	PG-O3B-PB	-2.45	123.99	132.62
19	P	12	APC	C5-C6-N6	2.27	123.80	120.35
19	P	12	APC	C3'-C2'-C1'	2.07	104.09	100.98

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
19	P	12	APC	PB-C3A-PA-O1A
19	P	12	APC	PB-C3A-PA-O2A
19	P	12	APC	PB-C3A-PA-O5'
19	P	12	APC	O4'-C4'-C5'-O5'
19	P	12	APC	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 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.

Ligand APC P 12

Bond lengths

Bond angles

Torsions

Rings
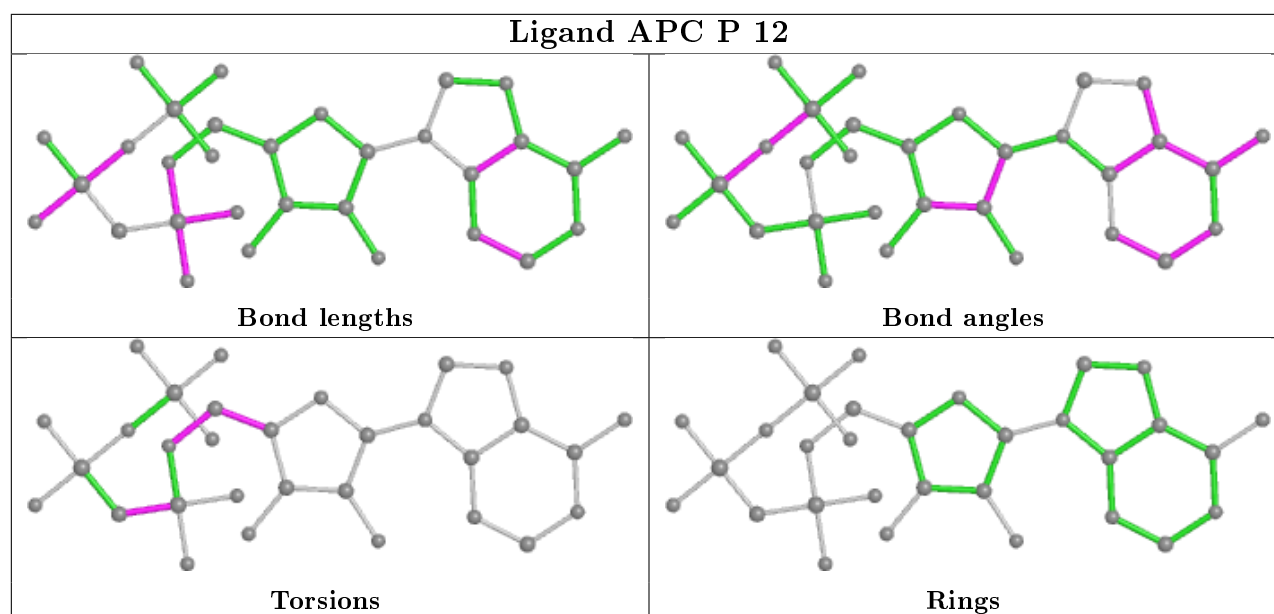
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	934:LYS	C	935:ARG	N	3.57

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	1427/1733 (82%)	-0.24	6 (0%)	92	86	58, 108, 175, 232	0
2	B	1106/1224 (90%)	-0.19	11 (0%)	82	70	65, 120, 181, 216	0
3	C	266/318 (83%)	-0.30	0	100	100	81, 109, 151, 180	0
4	D	177/221 (80%)	-0.24	0	100	100	85, 123, 173, 199	0
5	E	214/215 (99%)	-0.16	3 (1%)	75	61	83, 139, 189, 210	0
6	F	85/155 (54%)	-0.40	0	100	100	68, 91, 124, 144	0
7	G	171/171 (100%)	-0.08	0	100	100	79, 108, 151, 169	0
8	H	133/146 (91%)	0.23	4 (3%)	50	34	118, 153, 189, 226	0
9	I	119/122 (97%)	0.04	4 (3%)	45	30	117, 159, 198, 220	0
10	J	65/70 (92%)	-0.43	0	100	100	88, 105, 148, 160	0
11	K	115/120 (95%)	-0.24	2 (1%)	70	55	77, 106, 155, 173	0
12	L	46/70 (65%)	0.56	3 (6%)	18	11	105, 175, 195, 205	0
13	N	11/14 (78%)	-0.16	0	100	100	176, 194, 264, 267	0
14	P	10/11 (90%)	-0.32	0	100	100	85, 113, 169, 174	0
15	T	23/26 (88%)	-0.15	0	100	100	83, 148, 246, 258	0
16	X	116/146 (79%)	0.35	6 (5%)	27	17	163, 195, 212, 220	0
All	All	4084/4762 (85%)	-0.18	39 (0%)	82	70	58, 118, 189, 267	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	882	THR	6.0
8	H	86	ASP	4.4
12	L	25	ALA	4.4
11	K	115	ALA	4.0
12	L	26	THR	3.9

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
15	BRU	T	22	20/21	0.96	0.17	87,97,123,144	0

6.3 Carbohydrates (i)
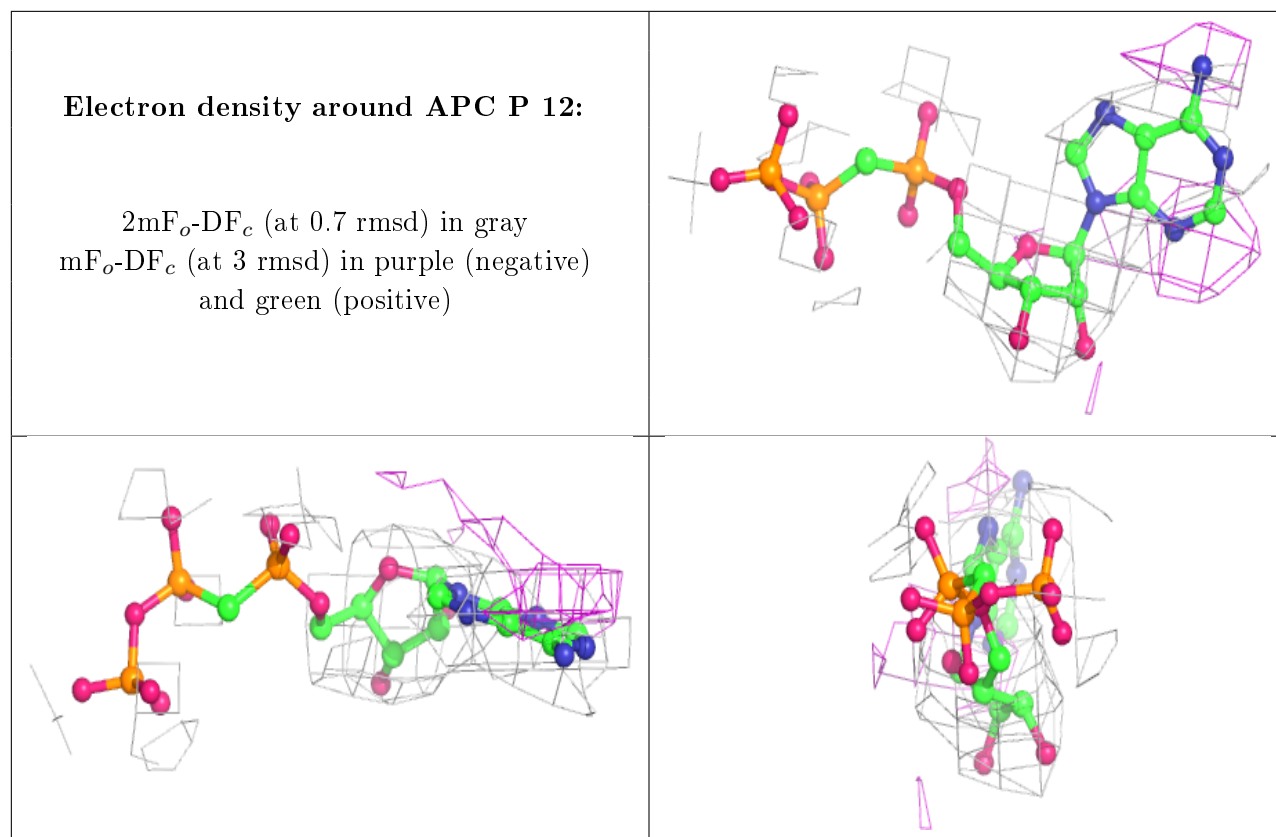
There are no carbohydrates in this entry.

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
19	APC	P	12	31/31	0.83	0.28	102,118,169,170	0
17	ZN	I	1122	1/1	0.94	0.04	234,234,234,234	0
18	MG	A	2458	1/1	0.99	0.12	58,58,58,58	0
17	ZN	C	1269	1/1	0.99	0.11	92,92,92,92	0
17	ZN	I	1121	1/1	0.99	0.11	129,129,129,129	0
17	ZN	L	1071	1/1	0.99	0.04	192,192,192,192	0
17	ZN	A	2457	1/1	0.99	0.16	77,77,77,77	0
17	ZN	A	2456	1/1	0.99	0.06	128,128,128,128	0
17	ZN	J	1066	1/1	1.00	0.20	85,85,85,85	0
17	ZN	B	2225	1/1	1.00	0.18	92,92,92,92	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 P 12:

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

6.5 Other polymers (i)

There are no such residues in this entry.