wwPDB X-ray Structure Validation Summary Report (i)

Mar 18, 2024 – 11:02 AM JST

PDB ID	:	5XOG
Title	:	RNA Polymerase II elongation complex bound with Spt5 KOW5 and Elf1
Authors	:	Ehara, H.; Shirouzu, M.; Sekine, S.
Deposited on	:	2017-05-28
Resolution	:	3.00 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

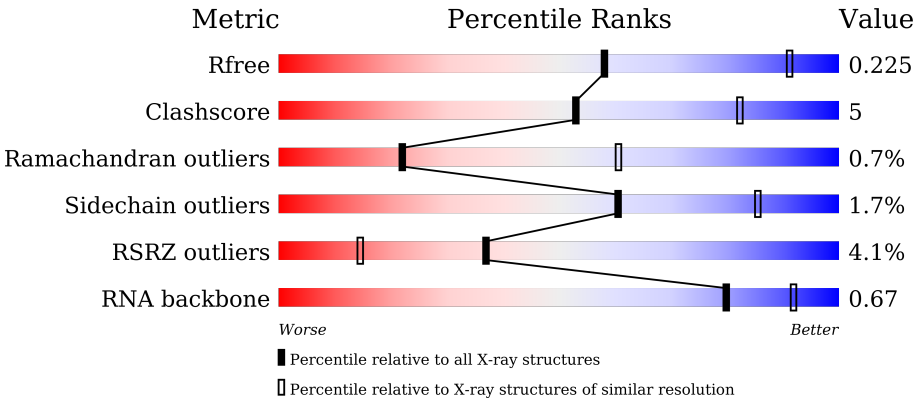
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.00 Å.

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



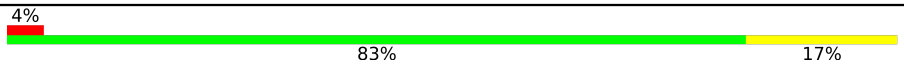
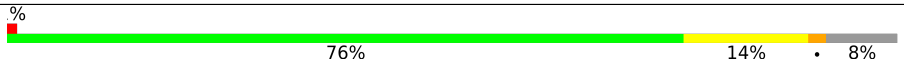
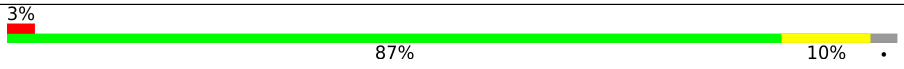
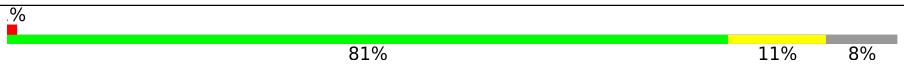
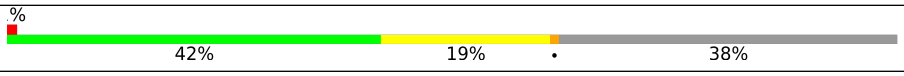
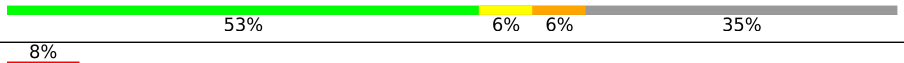
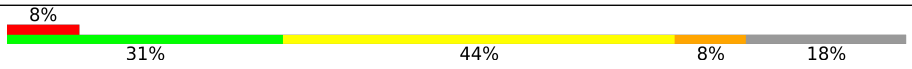
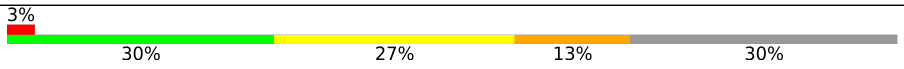
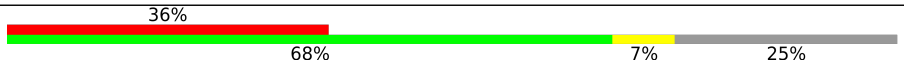
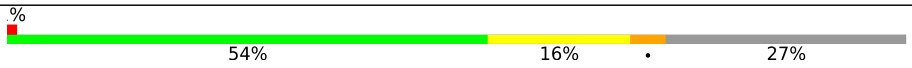
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

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 <=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	1743	3% 73% 8% • 19%
2	B	1227	5% 80% 13% • 5%
3	C	304	% 76% 10% 13%
4	D	186	6% 80% 10% •• 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	214	
6	F	155	
7	G	171	
8	H	145	
9	I	115	
10	J	72	
11	K	118	
12	L	72	
13	P	17	
14	T	39	
15	N	30	
16	M	85	
17	W	83	

2 Entry composition (i)

There are 20 unique types of molecules in this entry. The entry contains 33711 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1414	Total	C	N	O	S	0	0	0
			11139	7025	1941	2103	70			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1161	Total	C	N	O	S	0	0	0
			9261	5835	1636	1732	58			

- Molecule 3 is a protein called RNA polymerase II third largest subunit B44, part of central core.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	263	Total	C	N	O	S	0	0	0
			2098	1319	354	413	12			

- Molecule 4 is a protein called RNA polymerase II subunit B32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	168	Total	C	N	O	S	0	0	0
			1314	812	237	263	2			

- Molecule 5 is a protein called RNA polymerase subunit ABC27, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	213	Total	C	N	O	S	0	0	0
			1740	1094	312	324	10			

- Molecule 6 is a protein called RNA polymerase subunit ABC23, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	84	Total	C	N	O	S	0	0	0
			677	429	114	131	3			

- Molecule 7 is a protein called RNA polymerase II subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1324	858	214	247	5			

- Molecule 8 is a protein called RNA polymerase subunit ABC14.5, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	133	Total	C	N	O	S	0	0	0
			1052	671	169	208	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	111	Total	C	N	O	S	0	0	0
			917	565	161	180	11			

- Molecule 10 is a protein called RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	66	Total	C	N	O	S	0	0	0
			545	349	95	95	6			

- Molecule 11 is a protein called RNA polymerase II subunit B12.5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	113	Total	C	N	O	S	0	0	0
			932	599	160	169	4			

- Molecule 12 is a protein called RNA polymerase subunit ABC10-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	45	Total	C	N	O	S	0	0	0
			359	221	72	61	5			

- Molecule 13 is a RNA chain called RNA (5'-R(*UP*UP*UP*UP*UP*UP*UP*AP*UP*CP

*GP*AP*GP*AP*GP*GP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	P	11	Total	C	N	O	P	0	0	0
			238	106	44	77	11			

- Molecule 14 is a DNA chain called DNA (39-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	T	32	Total	C	N	O	P	0	0	0
			648	309	117	190	32			

- Molecule 15 is a DNA chain called DNA (30-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	N	21	Total	C	N	O	P	0	0	0
			433	206	76	130	21			

- Molecule 16 is a protein called Transcription elongation factor 1 homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	M	64	Total	C	N	O	S	0	0	0
			507	318	85	98	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	-2	GLY	-	expression tag	UNP A0A1B2JER8
M	-1	PRO	-	expression tag	UNP A0A1B2JER8
M	0	GLY	-	expression tag	UNP A0A1B2JER8
M	53	GLY	ASN	engineered mutation	UNP A0A1B2JER8
M	54	GLN	LEU	engineered mutation	UNP A0A1B2JER8
M	55	ARG	SER	engineered mutation	UNP A0A1B2JER8

- Molecule 17 is a protein called Spt4/5 complex component.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	W	61	Total	C	N	O	S	0	0	0
			486	311	86	88	1			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	733	GLY	-	expression tag	UNP F2QUC3
W	734	PRO	-	expression tag	UNP F2QUC3
W	735	GLY	-	expression tag	UNP F2QUC3

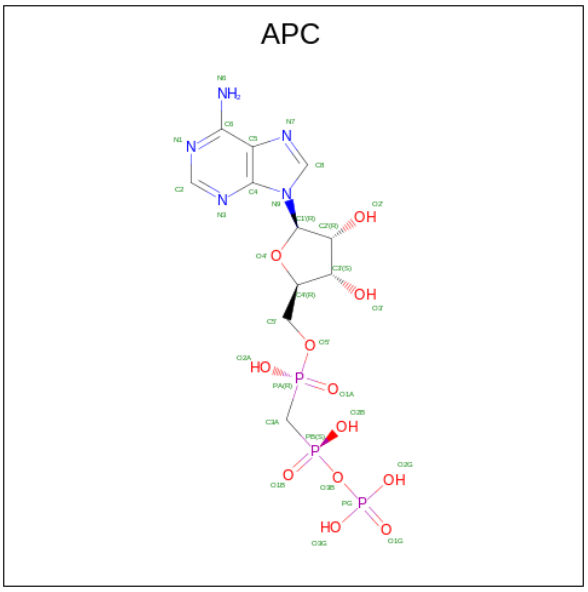
- Molecule 18 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 19 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 20 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



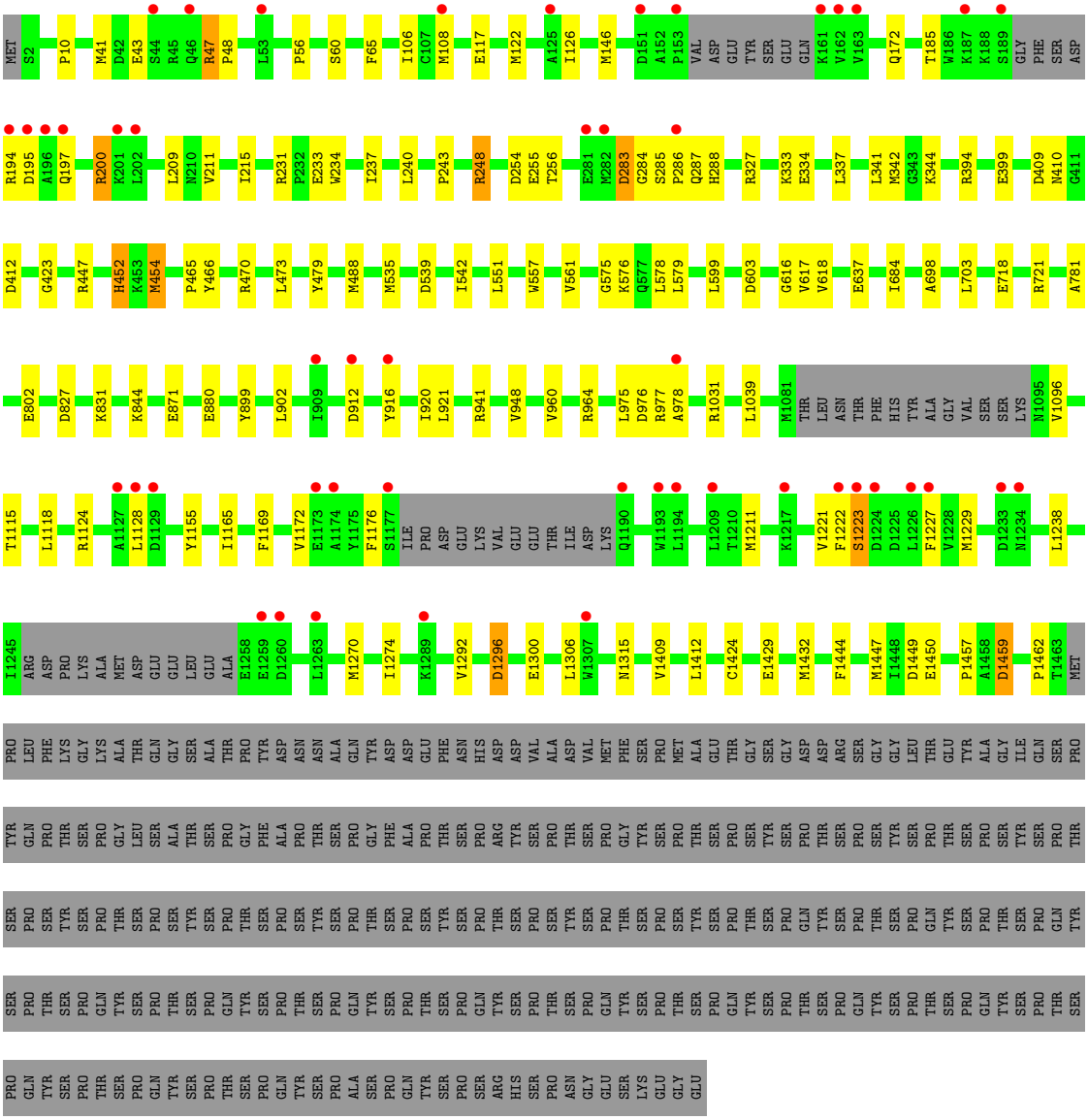
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
20	A	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, 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.

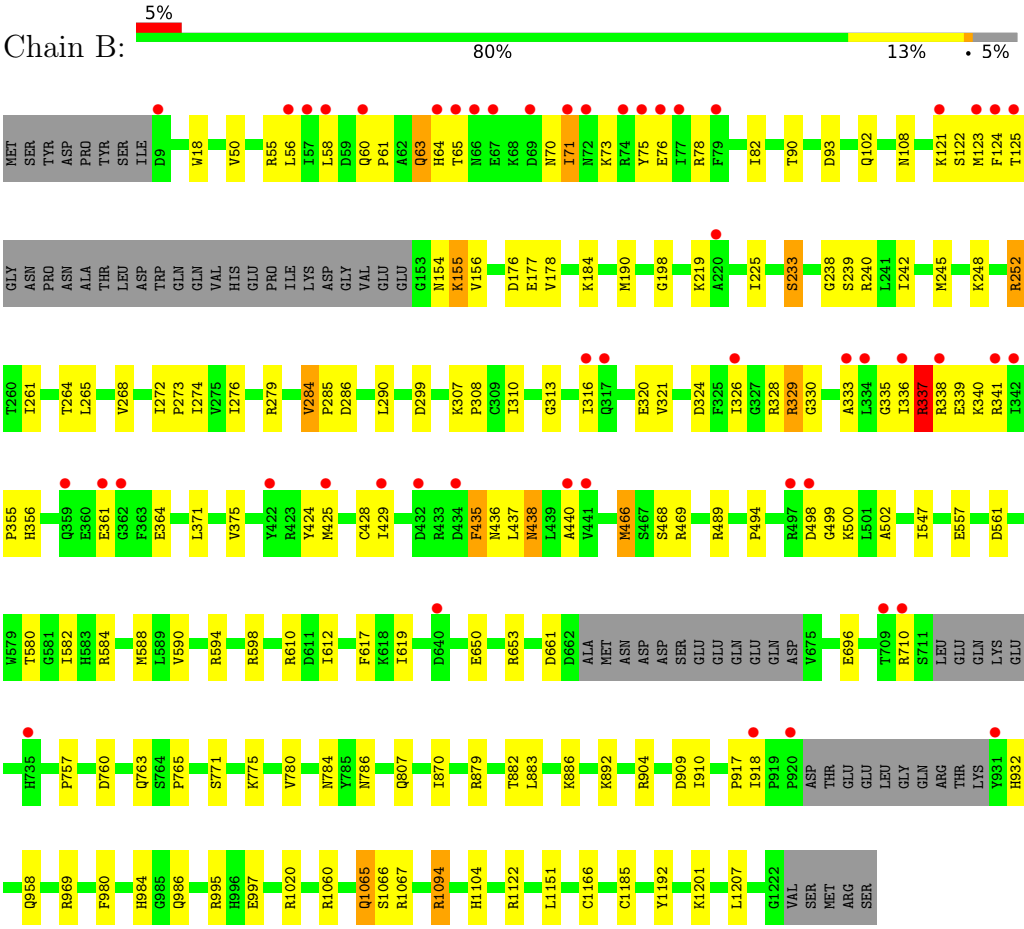
• Molecule 1: DNA-directed RNA polymerase subunit

Chain A:



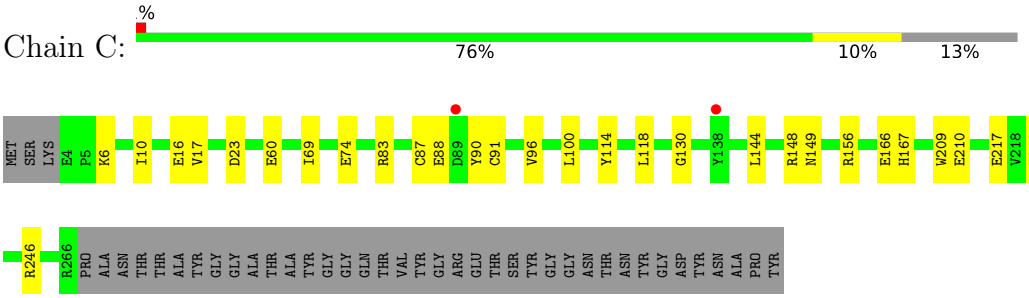
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain B:



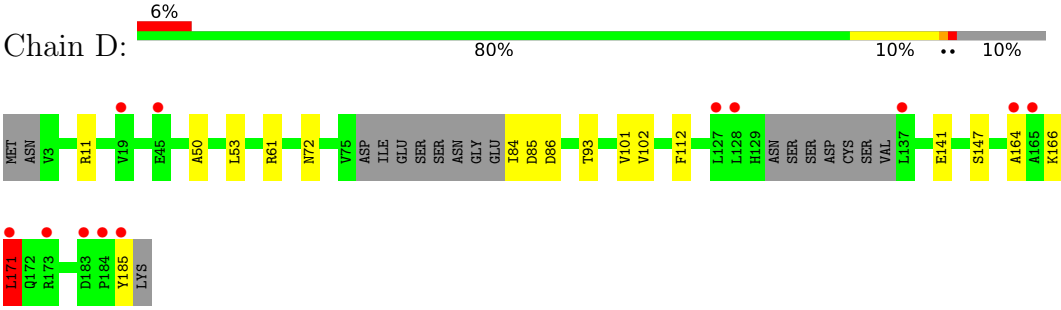
• Molecule 3: RNA polymerase II third largest subunit B44, part of central core

Chain C:



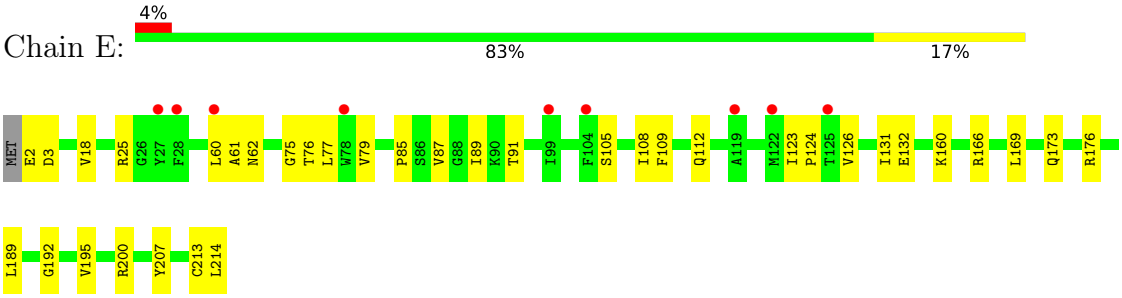
• Molecule 4: RNA polymerase II subunit B32

Chain D:



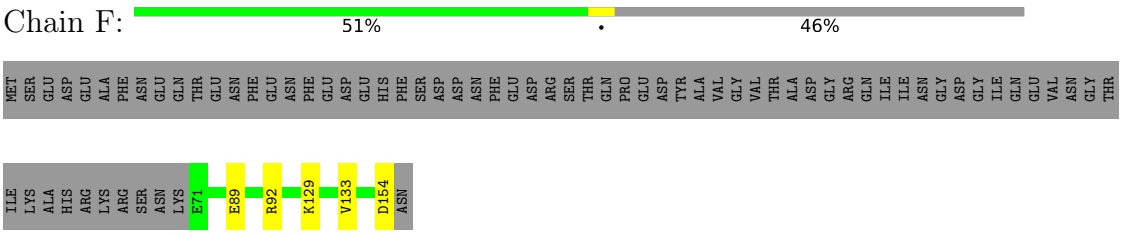
• Molecule 5: RNA polymerase subunit ABC27, common to RNA polymerases I, II, and III

Chain E:



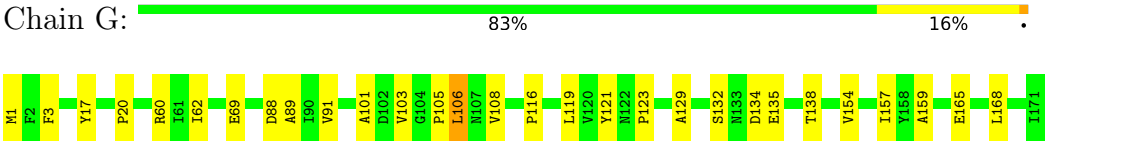
• Molecule 6: RNA polymerase subunit ABC23, common to RNA polymerases I, II, and III

Chain F:



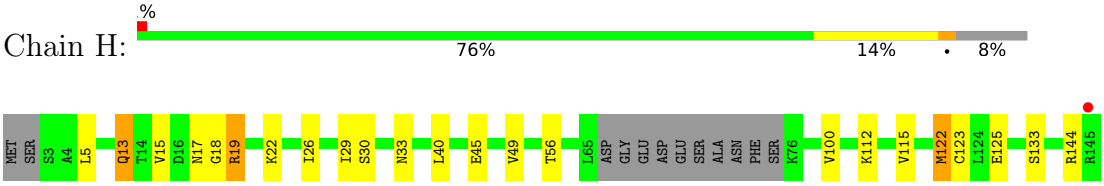
• Molecule 7: RNA polymerase II subunit

Chain G:



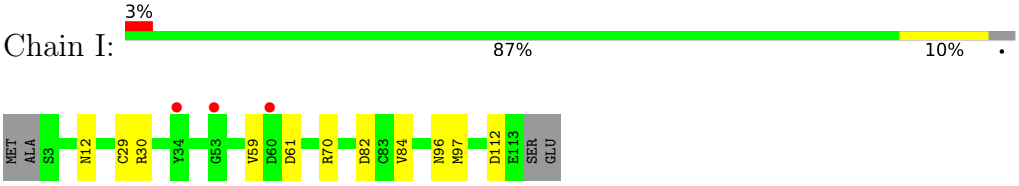
• Molecule 8: RNA polymerase subunit ABC14.5, common to RNA polymerases I, II, and III

Chain H:



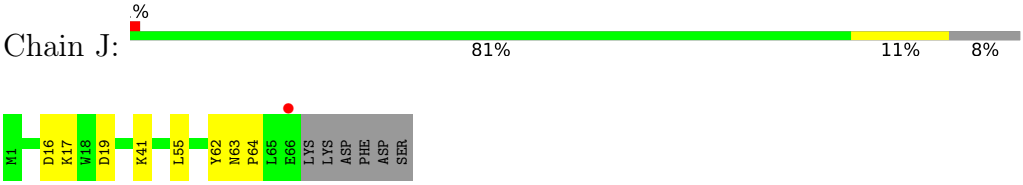
• Molecule 9: DNA-directed RNA polymerase subunit

Chain I:



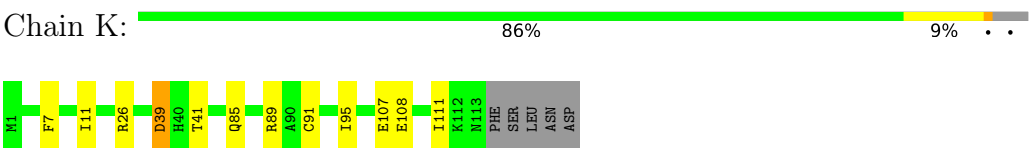
• Molecule 10: RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III

Chain J:



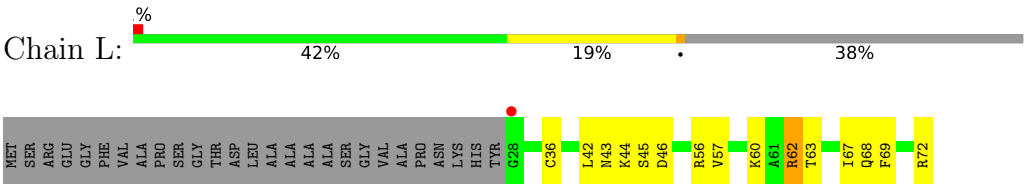
• Molecule 11: RNA polymerase II subunit B12.5

Chain K:



• Molecule 12: RNA polymerase subunit ABC10-alpha

Chain L:

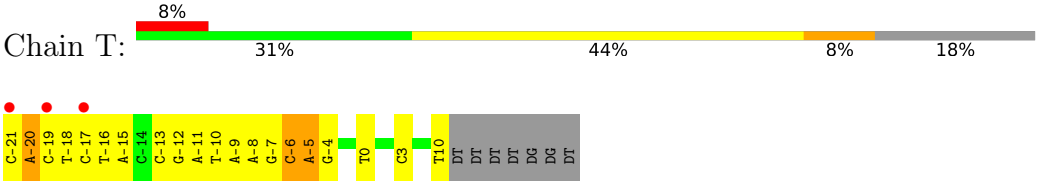


• Molecule 13: RNA (5'-R(*UP*UP*UP*UP*UP*UP*UP*AP*UP*CP*GP*AP*GP*AP*GP*GP*U)-3')

Chain P:

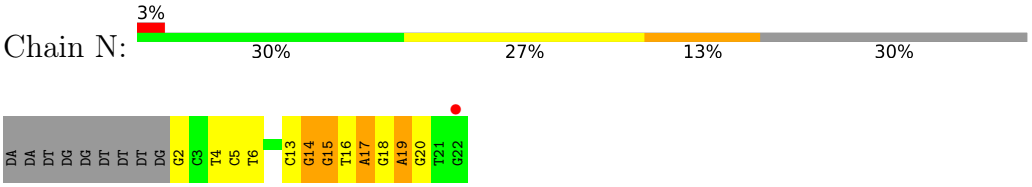
• Molecule 14: DNA (39-MER)

Chain T:



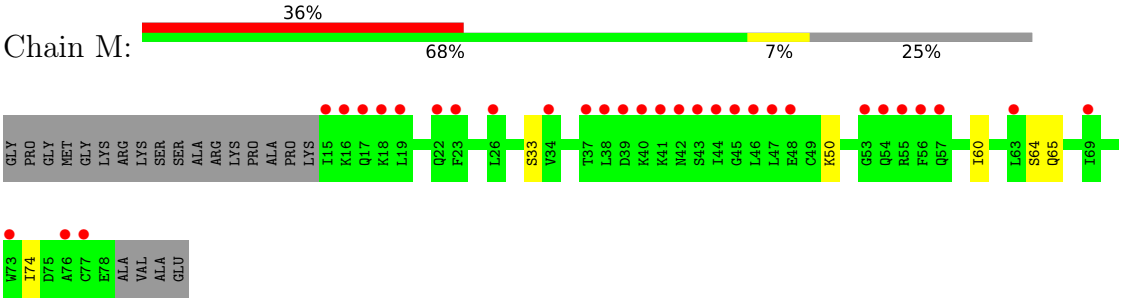
• Molecule 15: DNA (30-MER)

Chain N:



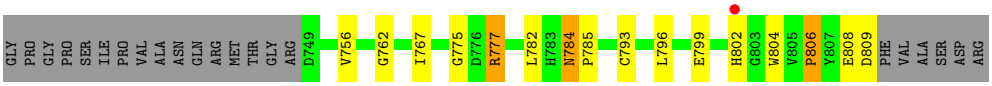
• Molecule 16: Transcription elongation factor 1 homolog

Chain M:



• Molecule 17: Spt4/5 complex component

Chain W:


GLY
PRO
GLY
PRO
SER
ILE
PRO
VAL
ALA
ASN
GLN
ARG
MET
THR
GLY
ARG
D749
V756
G762
I767
G775
D776
R777
L782
H783
N784
P785
C793
L796
E799
H802
G803
W804
V805
P806
Y807
E808
D809
PHE
VAL
ALA
SER
ASP
ARG

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	155.25Å 159.91Å 268.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.55 – 3.00 49.55 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.55-3.00) 99.9 (49.55-3.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.195 , 0.225 0.195 , 0.225	Depositor DCC
R_{free} test set	1538 reflections (1.15%)	wwPDB-VP
Wilson B-factor (Å^2)	84.8	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 64.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	33711	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.08% 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: ZN, MG, 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	A	0.31	0/11345	0.52	0/15331
2	B	0.31	0/9441	0.56	1/12732 (0.0%)
3	C	0.31	0/2139	0.58	1/2895 (0.0%)
4	D	0.28	0/1326	0.55	0/1788
5	E	0.29	0/1772	0.49	0/2385
6	F	0.30	0/687	0.54	0/931
7	G	0.32	0/1353	0.61	0/1837
8	H	0.32	0/1069	0.56	0/1444
9	I	0.27	0/934	0.51	0/1257
10	J	0.33	0/554	0.55	0/742
11	K	0.31	0/953	0.56	0/1291
12	L	0.38	0/365	0.70	0/484
13	P	0.38	0/266	0.89	0/413
14	T	1.01	2/725 (0.3%)	1.13	4/1114 (0.4%)
15	N	0.85	0/484	1.22	4/746 (0.5%)
16	M	0.25	0/515	0.45	0/694
17	W	0.34	0/498	0.54	0/674
All	All	0.35	2/34426 (0.0%)	0.59	10/46758 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	10	DT	C3'-O3'	7.31	1.53	1.44
14	T	10	DT	C1'-N1	6.21	1.57	1.49

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
14	T	-5	DA	O4'-C1'-N9	7.39	113.17	108.00
14	T	-20	DA	O4'-C1'-N9	-7.03	103.08	108.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	210	GLU	C-N-CA	-6.55	105.33	121.70
15	N	14	DG	O5'-P-OP1	6.48	118.48	110.70
2	B	499	GLY	N-CA-C	-6.38	97.14	113.10

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	A	11139	0	11168	93	0
2	B	9261	0	9265	119	0
3	C	2098	0	2057	20	0
4	D	1314	0	1314	15	0
5	E	1740	0	1754	25	0
6	F	677	0	693	2	0
7	G	1324	0	1342	20	0
8	H	1052	0	1050	13	0
9	I	917	0	866	6	0
10	J	545	0	560	6	0
11	K	932	0	944	7	0
12	L	359	0	358	10	0
13	P	238	0	118	1	0
14	T	648	0	360	21	0
15	N	433	0	239	23	0
16	M	507	0	498	9	0
17	W	486	0	491	10	0
18	A	2	0	0	0	0
18	B	1	0	0	0	0
18	C	1	0	0	0	0
18	I	2	0	0	0	0
18	J	1	0	0	0	0
18	L	1	0	0	0	0
18	M	1	0	0	0	0
19	A	1	0	0	0	0
20	A	31	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	33711	0	33090	354	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 5.

The worst 5 of 354 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)
14:T:-7:DG:H2"	14:T:-6:DC:H5"	1.43	1.00
15:N:15:DG:H2"	15:N:16:DT:H5"	1.46	0.94
2:B:904:ARG:NH1	17:W:782:LEU:O	2.02	0.92
2:B:55:ARG:HD2	2:B:76:GLU:OE2	1.75	0.87
2:B:121:LYS:HD2	2:B:435:PHE:HE1	1.38	0.86

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	1402/1743 (80%)	1347 (96%)	48 (3%)	7 (0%)	29 68
2	B	1151/1227 (94%)	1083 (94%)	57 (5%)	11 (1%)	15 53
3	C	261/304 (86%)	253 (97%)	7 (3%)	1 (0%)	34 72
4	D	162/186 (87%)	148 (91%)	11 (7%)	3 (2%)	8 36
5	E	211/214 (99%)	207 (98%)	4 (2%)	0	100 100
6	F	82/155 (53%)	80 (98%)	2 (2%)	0	100 100
7	G	169/171 (99%)	159 (94%)	8 (5%)	2 (1%)	13 48
8	H	129/145 (89%)	125 (97%)	3 (2%)	1 (1%)	19 57
9	I	109/115 (95%)	101 (93%)	8 (7%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
10	J	64/72 (89%)	61 (95%)	2 (3%)	1 (2%)	9 40
11	K	111/118 (94%)	110 (99%)	1 (1%)	0	100 100
12	L	43/72 (60%)	42 (98%)	0	1 (2%)	6 30
16	M	62/85 (73%)	59 (95%)	3 (5%)	0	100 100
17	W	59/83 (71%)	49 (83%)	9 (15%)	1 (2%)	9 39
All	All	4015/4690 (86%)	3824 (95%)	163 (4%)	28 (1%)	22 60

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	255	GLU
1	A	287	GLN
2	B	61	PRO
2	B	155	LYS
2	B	239	SER

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	1227/1528 (80%)	1213 (99%)	14 (1%)	73 90
2	B	1016/1077 (94%)	990 (97%)	26 (3%)	46 78
3	C	236/264 (89%)	235 (100%)	1 (0%)	91 97
4	D	143/160 (89%)	139 (97%)	4 (3%)	43 77
5	E	196/197 (100%)	195 (100%)	1 (0%)	88 96
6	F	75/137 (55%)	72 (96%)	3 (4%)	31 68
7	G	148/148 (100%)	147 (99%)	1 (1%)	84 94
8	H	120/130 (92%)	117 (98%)	3 (2%)	47 79
9	I	106/109 (97%)	105 (99%)	1 (1%)	78 92
10	J	60/66 (91%)	60 (100%)	0	100 100
11	K	104/109 (95%)	103 (99%)	1 (1%)	76 91

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
12	L	38/56 (68%)	36 (95%)	2 (5%)	22 58
16	M	60/74 (81%)	60 (100%)	0	100 100
17	W	54/71 (76%)	49 (91%)	5 (9%)	9 33
All	All	3583/4126 (87%)	3521 (98%)	62 (2%)	60 85

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	435	PHE
12	L	36	CYS
2	B	786	ASN
11	K	39	ASP
17	W	793	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	197	GLN
1	A	452	HIS
2	B	154	ASN
16	M	65	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	P	10/17 (58%)	1 (10%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	P	1	A

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 10 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	$\#\|Z\| > 2$
20	APC	A	1804	19	27,33,33	1.73	5 (18%)	31,52,52	1.55	4 (12%)

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
20	APC	A	1804	19	-	8/15/38/38	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
20	A	1804	APC	PA-O5'	5.53	1.65	1.57
20	A	1804	APC	O3'-C3'	-3.72	1.34	1.43
20	A	1804	APC	PB-O2B	-2.81	1.49	1.56
20	A	1804	APC	C8-N7	-2.63	1.30	1.34
20	A	1804	APC	O4'-C1'	-2.35	1.37	1.41

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
20	A	1804	APC	O1A-PA-C3A	5.22	122.87	109.07
20	A	1804	APC	PB-O3B-PG	3.75	145.82	132.62
20	A	1804	APC	O2B-PB-O1B	2.30	117.75	110.07
20	A	1804	APC	C2'-C3'-C4'	-2.11	98.54	102.64

There are no chirality outliers.

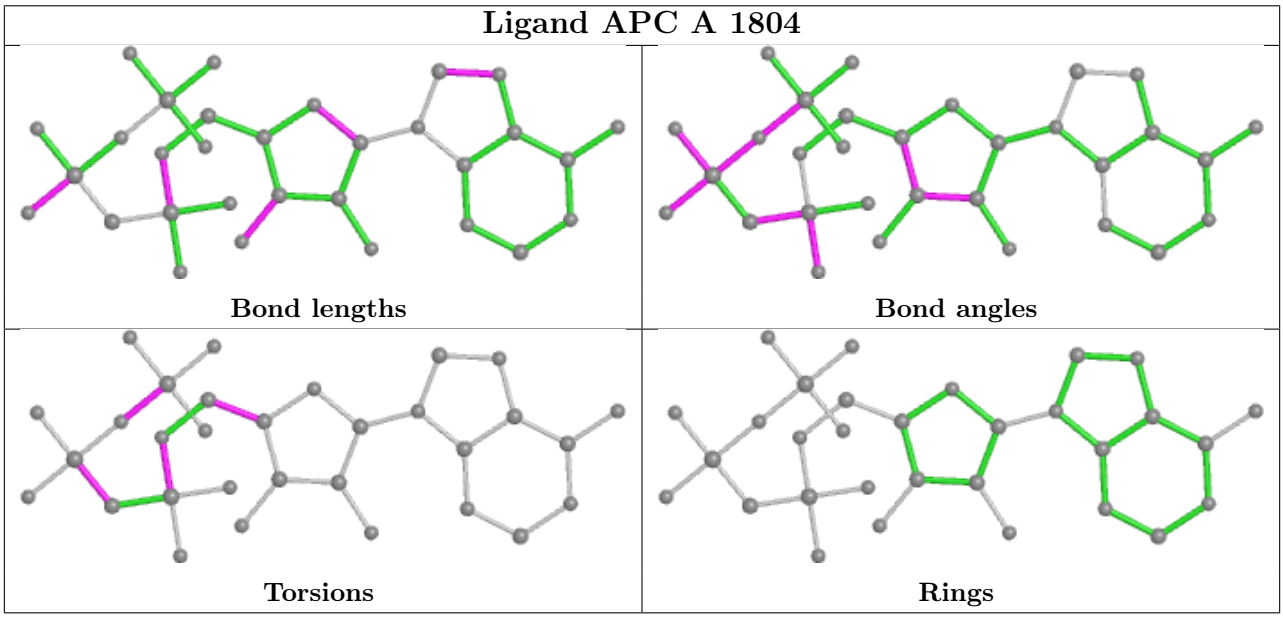
5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
20	A	1804	APC	PA-C3A-PB-O1B
20	A	1804	APC	PA-C3A-PB-O2B
20	A	1804	APC	PA-C3A-PB-O3B
20	A	1804	APC	C5'-O5'-PA-O1A
20	A	1804	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.



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, 95^{th} 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	1414/1743 (81%)	0.07	48 (3%) 45 19	39, 78, 142, 198	0
2	B	1161/1227 (94%)	0.15	57 (4%) 29 11	42, 74, 152, 336	0
3	C	263/304 (86%)	-0.17	2 (0%) 86 65	51, 72, 118, 147	0
4	D	168/186 (90%)	0.45	12 (7%) 16 5	66, 112, 166, 182	0
5	E	213/214 (99%)	0.14	9 (4%) 36 14	59, 115, 158, 164	0
6	F	84/155 (54%)	-0.38	0 100 100	46, 59, 94, 114	0
7	G	171/171 (100%)	0.02	0 100 100	54, 89, 141, 165	0
8	H	133/145 (91%)	0.15	1 (0%) 86 65	59, 85, 114, 131	0
9	I	111/115 (96%)	0.32	3 (2%) 54 26	80, 130, 146, 159	0
10	J	66/72 (91%)	-0.17	1 (1%) 73 46	55, 66, 105, 152	0
11	K	113/118 (95%)	-0.26	0 100 100	48, 71, 105, 140	0
12	L	45/72 (62%)	-0.17	1 (2%) 62 33	58, 84, 113, 128	0
13	P	11/17 (64%)	-0.20	0 100 100	48, 56, 95, 163	0
14	T	32/39 (82%)	0.30	3 (9%) 8 3	46, 149, 222, 228	0
15	N	21/30 (70%)	0.27	1 (4%) 30 11	121, 175, 234, 237	0
16	M	64/85 (75%)	2.11	31 (48%) 0 0	137, 162, 177, 186	0
17	W	61/83 (73%)	0.43	1 (1%) 72 44	72, 90, 120, 149	0
All	All	4131/4776 (86%)	0.12	170 (4%) 37 14	39, 81, 154, 336	0

The worst 5 of 170 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	253	GLU	10.1
2	B	65	THR	10.0
2	B	64	HIS	8.5
2	B	125	THR	8.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	194	ARG	7.7

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

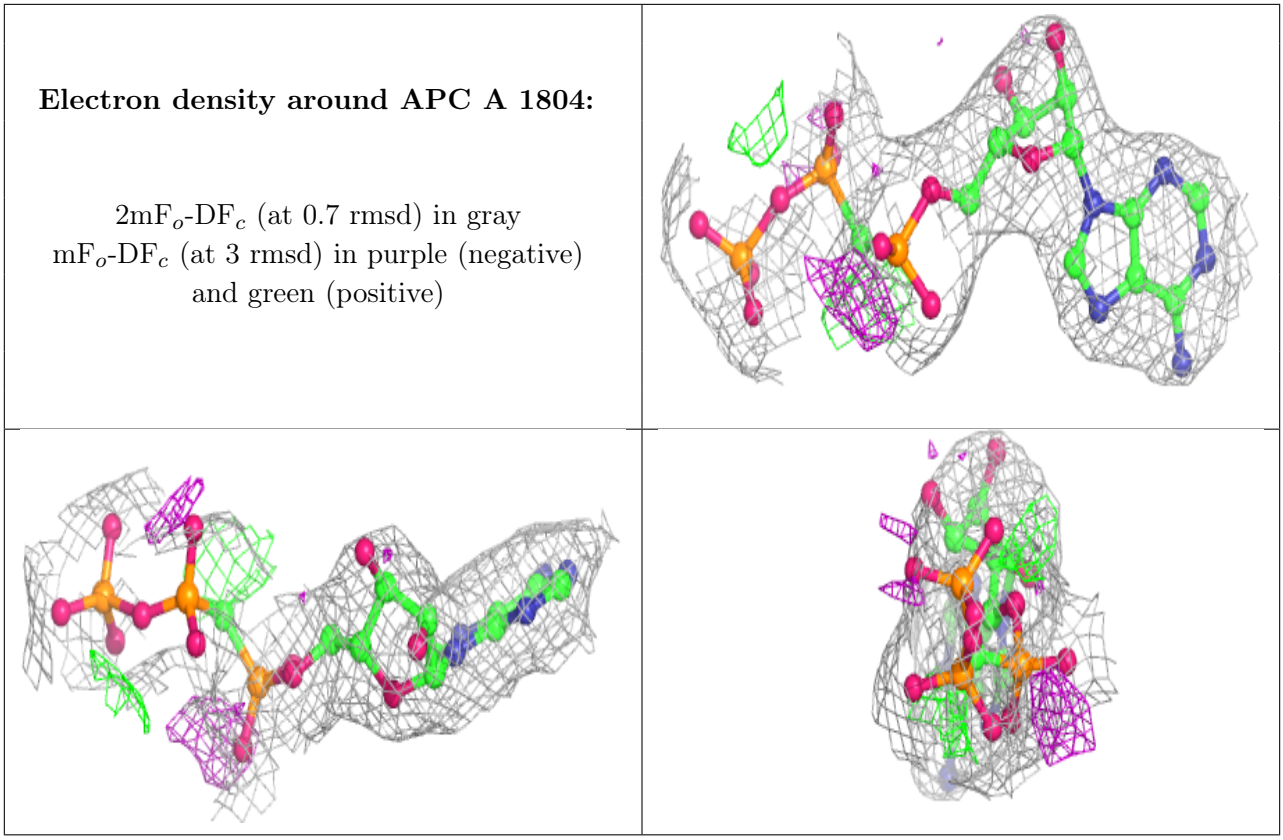
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
20	APC	A	1804	31/31	0.91	0.18	48,64,129,133	0
18	ZN	I	201	1/1	0.95	0.10	119,119,119,119	0
19	MG	A	1803	1/1	0.96	0.25	45,45,45,45	0
18	ZN	M	201	1/1	0.96	0.14	325,325,325,325	0
18	ZN	I	202	1/1	0.97	0.11	114,114,114,114	0
18	ZN	A	1802	1/1	0.97	0.10	99,99,99,99	0
18	ZN	A	1801	1/1	0.98	0.21	88,88,88,88	0
18	ZN	C	401	1/1	0.98	0.11	93,93,93,93	0
18	ZN	L	101	1/1	0.98	0.15	90,90,90,90	0
18	ZN	B	1301	1/1	0.99	0.25	91,91,91,91	0
18	ZN	J	101	1/1	0.99	0.22	63,63,63,63	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 A 1804:

$2mF_o$-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.