



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 19, 2023 – 06:58 AM EDT

PDB ID : 2NVX  
Title : RNA polymerase II elongation complex in 5 mM Mg+2 with 2'-dUTP  
Authors : Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.  
Deposited on : 2006-11-13  
Resolution : 3.60 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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

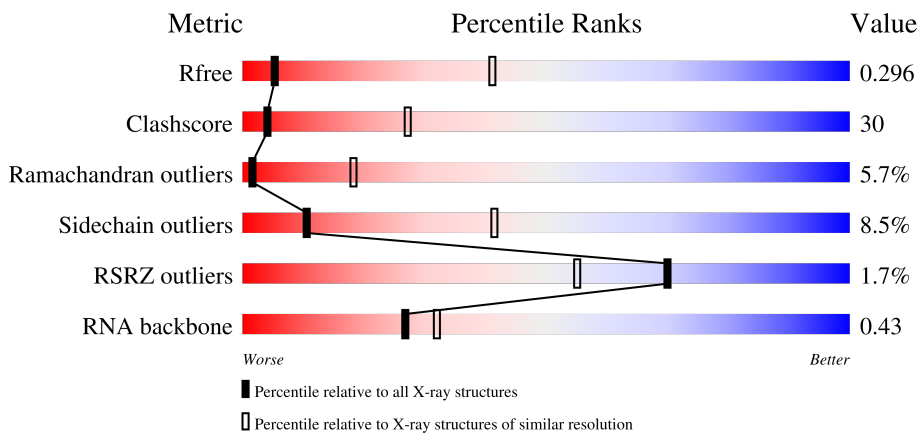
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	10	 20% 70% 10%
2	N	14	 21% 93% 7%
3	T	28	 25% 46% 32% 21%
4	A	1733	 44% 29% 6% 19%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	B	1224	<p>%</p> <p>48% 35% 7% 9%</p>
6	C	318	<p>48% 30% 5% 16%</p>
7	E	215	<p>2%</p> <p>72% 27%</p>
8	F	155	<p>38% 15% 43%</p>
9	H	146	<p>3%</p> <p>55% 31% 6% 8%</p>
10	I	122	<p>%</p> <p>71% 24%</p>
11	J	70	<p>53% 33% 7% 7%</p>
12	K	120	<p>67% 27% 5%</p>
13	L	70	<p>3%</p> <p>40% 21% 34%</p>

## 2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 29436 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 5'-R(\*AP\*UP\*CP\*GP\*AP\*GP\*AP\*GP\*GP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	R	10	216	98	45	64	9	0	0	0

- Molecule 2 is a DNA chain called 5'-D(\*CP\*TP\*GP\*CP\*TP\*TP\*AP\*TP\*CP\*GP\*GP\*TP\*AP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	N	14	284	137	49	85	13	0	0	0

- Molecule 3 is a DNA chain called 28-MER DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	T	28	566	271	104	164	27	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	A	1402	11028	6950	1934	2083	61	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	1114	8856	5605	1553	1644	54	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	F	88	712	455	120	134	3	0	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	H	134	1076	678	181	212	5	0	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit 9.

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

- Molecule 11 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

tide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

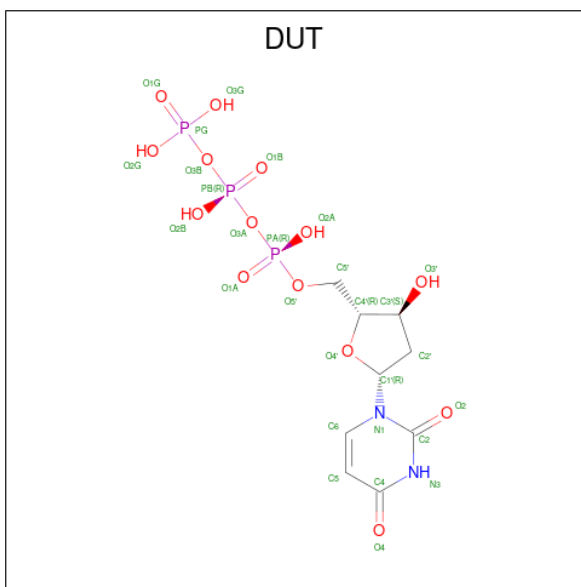
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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 (three-letter code: MG) (formula: Mg).

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

- Molecule 16 is DEOXYURIDINE-5'-TRIPHOSPHATE (three-letter code: DUT) (formula: C<sub>9</sub>H<sub>15</sub>N<sub>2</sub>O<sub>14</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
16	B	1	56	18	4	28	6	0	1

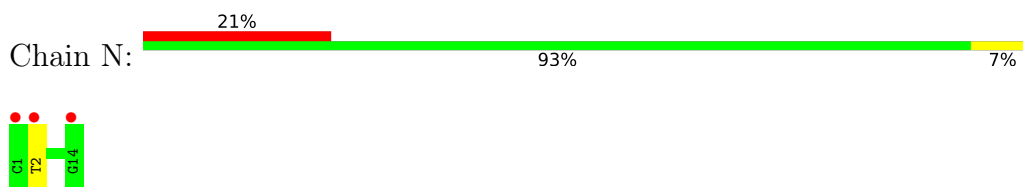
### 3 Residue-property plots [i](#)

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.

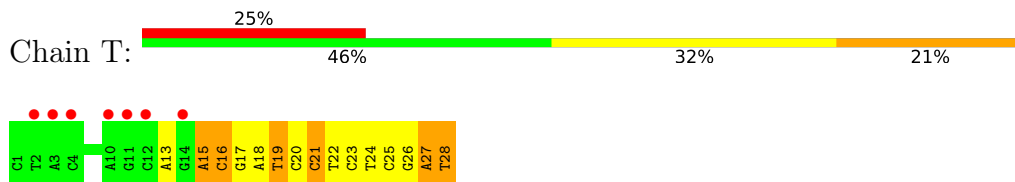
- Molecule 1: 5'-R(\*AP\*UP\*CP\*GP\*AP\*GP\*AP\*GP\*GP\*A)-3'



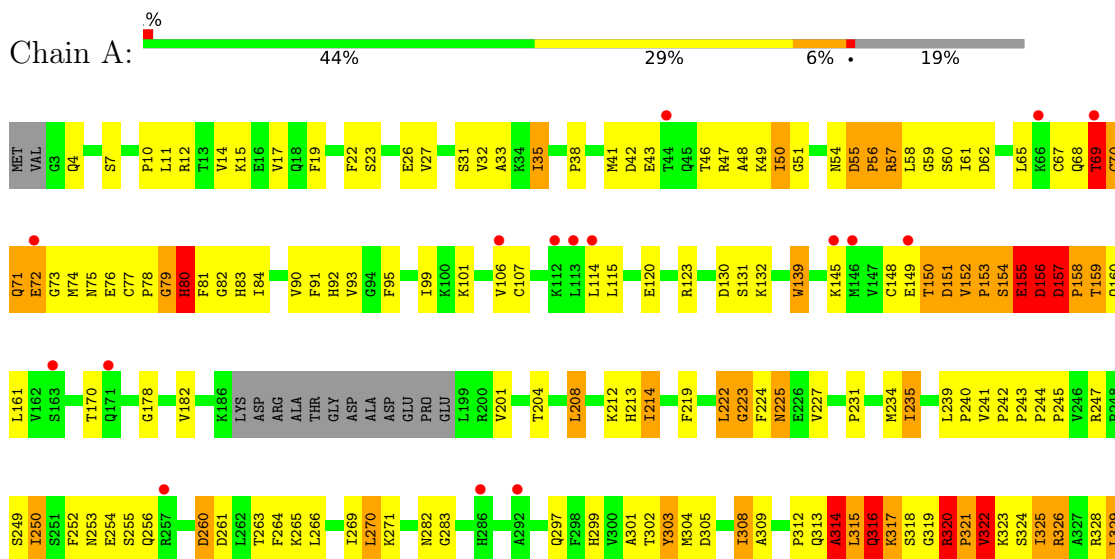
- Molecule 2: 5'-D(\*CP\*TP\*GP\*CP\*TP\*TP\*AP\*TP\*CP\*GP\*GP\*TP\*AP\*G)-3'



- Molecule 3: 28-MER DNA template strand



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

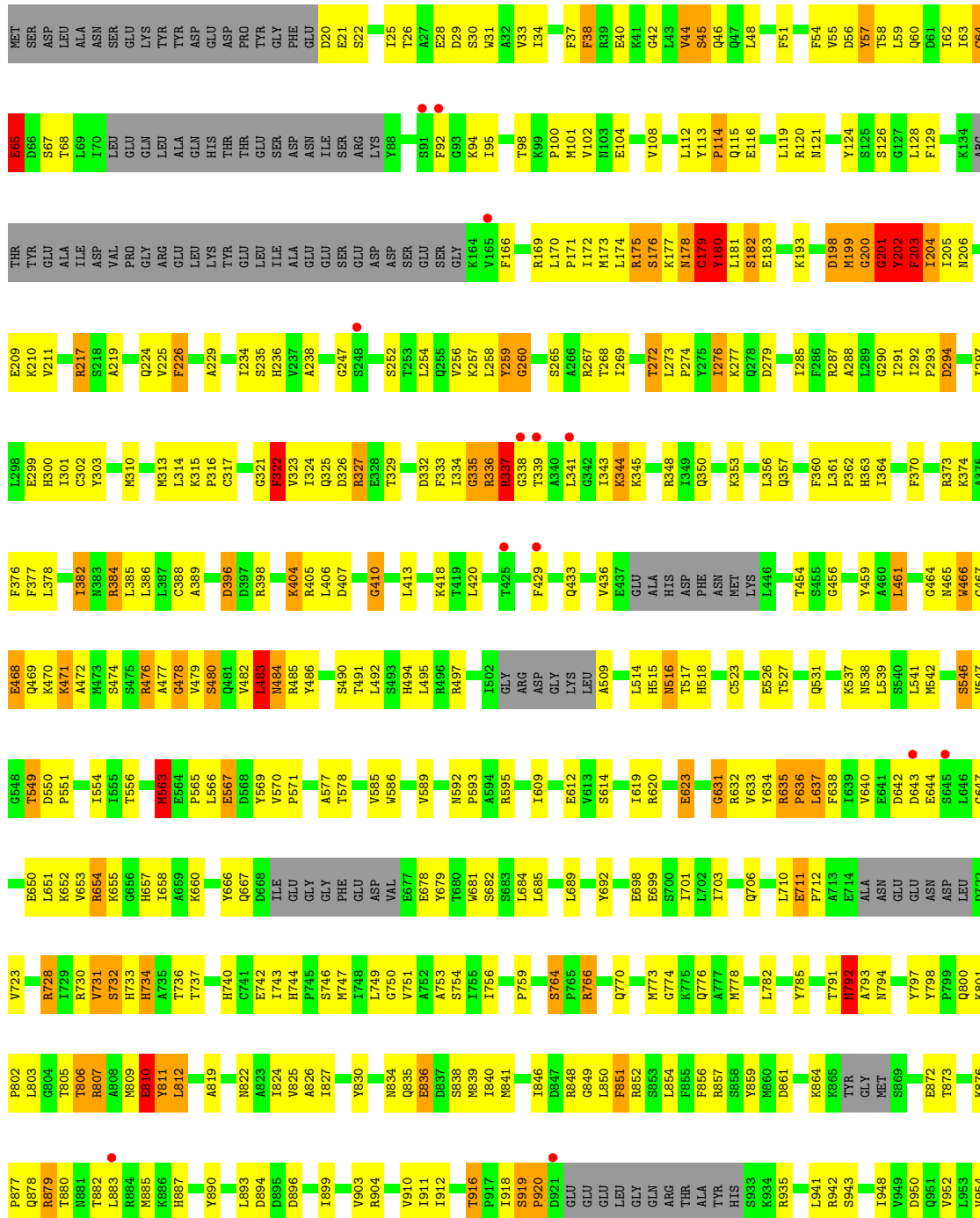


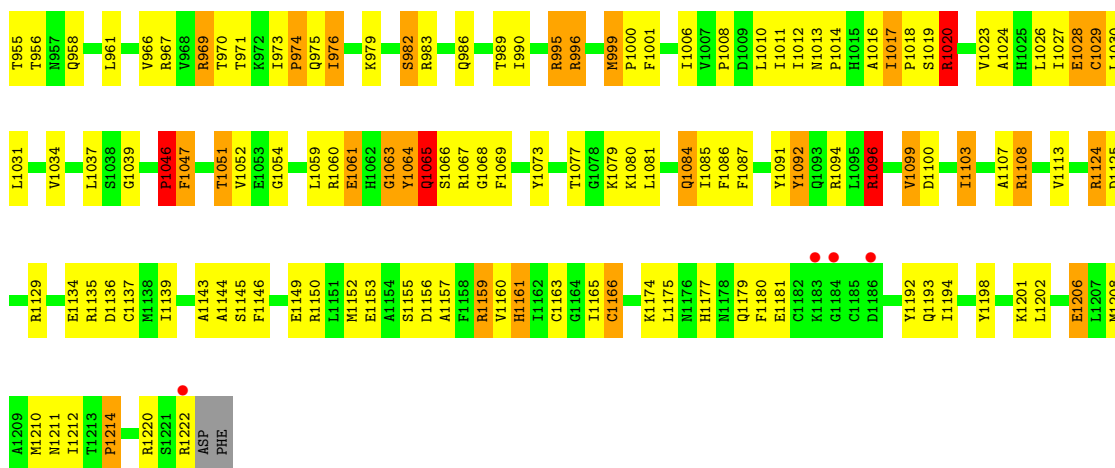


SER	VAL	SER	C1400	Y1298	I1227	R1135	T1028	E918	V680	V765	V653	G574	E495	L415	K332
THR	SER	GLY	S1401	V1299	W1228	S1136	R1029	I919	R651	R774	T664	K575	E496	R416	E333
PRO	SER	VAL	F1402	K1300	A1137	I1138	V1030	L920	R852	I775	G685	Q576	E499	Y417	G334
THR	GLY	ASN	V1406	L1305	S1229	K1144	E1230	G921	D853	A776	I666	L578	R498	S418	R395
SER	PHE	ALA	V1407	L1306	D1231	M1292	M1292	D922	R854	F777	G667	S579	A499	D423	I336
PRO	SER	ASP	V1408	E1307	L1037	R1036	R1036	Q926	R856	G778	D688	S579	R337	L424	R337
THR	THR	LEU	L1409	T1308	L1037	L1037	L1037	L929	R857	F779	T689	I586	E500	Q425	R344
THR	THR	ASP	F1410	D1309	L1046	R858	L1046	L929	R858	W780	D672	H587	L504	L426	R344
SER	SER	VAL	V1411	K1235	S1047	S859	S1047	Y933	S859	D781	L588	O505	C505	F347	F347
PRO	PRO	LYS	A1412	G1310	V1065	Y1152	V1065	Y933	R860	R782	G589	A506	A506	K431	K431
THR	THR	ASP	L1418	L1236	M1048	Y1154	M1048	D939	G661	T783	R590	V507	V507	V432	V432
PRO	THR	GLU	D1419	D1155	L1084	D1155	L1084	R940	R662	L784	F591	P508	P508	E433	E433
SER	SER	LEU	V1424	T1161	V1084	T1161	V1084	V946	V663	H786	I679	I511	I511	P441	P441
THR	THR	PHE	V1424	V1162	V1088	V1162	V1088	L943	R865	D790	T682	N517	N517	V442	V442
SER	SER	SER	V1428	I1163	R1059	I1163	R1059	V946	F866	D790	T682	L598	L598	L443	L443
THR	ALA	PRO	V1432	R1289	P1060	R1289	P1060	F947	R868	S793	E685	R446	R446	R446	R446
PRO	VAL	LEU	Q1432	L1289	M1083	E1167	M1083	E951	G869	P794	R689	Q447	Q447	Q447	Q447
PRO	SER	ASP	M1433	F1324	V1064	E1168	V1064	E951	E870	P794	R689	P448	P448	P448	P448
THR	SER	SER	A1434	A1325	A952	I1169	A952	N953	D871	S796	V690	S449	S449	S449	S449
THR	THR	GLY	P1435	ALA	V1066	I1170	V1066	N953	G872	K797	R797	L450	L450	L450	L450
SER	SER	SER	I1436	GLU	L1067	Q1171	L1067	W954	R873	G798	T694	R451	R451	R451	R451
PRO	PRO	ASN	G1437	THR	A1088	L1172	A1088	P955	R874	F798	T703	D526	D526	D526	D526
THR	SER	ASP	F1441	GLU	A1089	H1173	A1089	L956	A875	W800	T703	T527	T527	T527	T527
SER	ALA	ALA	V1445	A1254	E1074	F1174	E1074	P957	A876	E901	G707	L528	L528	S454	S454
PRO	SER	MET	I1445	H1258	S1175	S1175	S1175	V958	R878	N802	M708	C529	C529	M455	M455
PRO	PRO	PRO	I1445	H1258	L1176	L1176	L1176	R961	E879	N802	M708	G530	G530	M456	M456
THR	THR	THR	ASP	L1263	M1079	LEU	M1079	R961	E879	L805	T709	H457	H457	A457	A457
PRO	SER	GLY	GLU	E1263	T1080	ASP	T1080	R962	E879	R806	H458	K533	K533	H458	H458
PRO	THR	PHE	GLU	E1264	L1081	GLU	L1081	I963	R883	G807	R459	L534	L534	R459	R459
THR	SER	THR	LEU	M1265	ASN	GLU	ASN	Q969	D884	L808	V617	T539	T539	V462	V462
PRO	THR	ALA	LEU	T1266	THR	ALA	THR	P810	R886	E618	E463	F540	F540	P464	P464
PRO	THR	VAL	VAL	E1269	PHE	GLU	PHE	N717	R886	P810	L463	F465	F465	L463	L463
THR	THR	LYS	THR	M1270	HIS	GLN	HIS	H972	G887	O811	V718	Y465	Y465	Y465	Y465
THR	THR	TYR	TYR	E1271	SER	GLN	SER	H972	G888	E812	F721	S466	S466	S466	S466
SER	ALA	ALA	MET	L1271	ALA	PHE	ALA	L881	S889	F814	F721	S466	S466	S466	S466
PRO	ASP	ASP	PRO	T1272	GLY	ASP	GLY	T982	D890	F814	D727	S624	S624	S624	S624
THR	THR	TYR	THR	Q1187	VAL	Q1187	VAL	I983	K895	F815	D727	F468	F468	F468	F468
THR	GLY	GLN	GLN	R1274	ALA	ALA	ALA	I986	R896	R821	R731	D544	D544	D544	D544
THR	GLY	LYS	LYS	L1193	SER	L1193	SER	I986	R896	R821	R731	Q545	Q545	Q545	Q545
PRO	PRO	ILE	ILE	R1194	K1092	R1194	K1092	L956	Y697	R821	R731	L470	L470	L470	L470
THR	THR	THR	THR	L1195	K1093	L1195	K1093	L983	R898	I825	N736	V546	V546	V546	V546
SER	SER	GLU	GLU	E1196	V1094	E1196	V1094	L983	R898	I825	N736	L547	L547	L547	L547
PRO	PRO	ILE	ILE	L1197	V1094	L1197	V1094	Q1011	V899	D826	L737	R546	R546	R546	R546
THR	THR	GLU	GLU	D1198	V1098	D1198	V1098	L1011	D900	K830	L740	M549	M549	M549	M549
THR	THR	ASP	ASP	R1199	P1099	R1199	P1099	V1015	L901	R830	L740	V553	V553	V553	V553
PRO	PRO	ALA	ALA	M1284	R1100	M1284	R1100	V1015	L902	R830	L740	P563	P563	P563	P563
THR	THR	GLY	GLY	Y1287	L1101	L1207	L1101	L1017	N903	R839	K744	A584	A584	A584	A584
SER	SER	GLU	GLU	D1288	L1105	T1207	L1105	F1018	T907	R839	K744	I585	I585	I585	I585
PRO	PRO	ALA	ALA	M1209	L1105	M1209	L1105	C1019	L908	R840	K744	I585	I585	I585	I585
PRO	PRO	VAL	VAL	V1291	L1116	M1209	L1116	C1020	D909	R840	K744	I585	I585	I585	I585
THR	THR	THR	THR	V1292	T1117	L1215	T1117	L1021	P910	L841	S754	K567	K567	K567	K567
PRO	PRO	SER	SER	S1293	V1118	L1216	V1118	L1022	S911	V842	F755	P568	P568	P568	P568
PRO	PRO	THR	THR	F1294	L1118	L1216	L1118	R1023	R912	L845	N757	F486	F486	F486	F486
THR	THR	TYR	TYR	T1295	Q1128	T1219	Q1128	S1024	L913	E846	I758	K569	K569	K569	K569
THR	THR	ASN	ASN	G1296	I1134	F1220	I1134	R1025	E914	D847	I758	P570	P570	P570	P570
PRO	PRO	GLU	GLU	K1221	I1134	K1221	I1134	L1226	S915	R848	M761	L571	L571	L571	L571
PRO	PRO	GLU	GLU	E1297	I1134	K1221	I1134	A1027	S915	R848	M761	L489	L489	L489	L489
PRO	PRO	GLU	GLU	E1297	I1134	K1221	I1134	A1027	S915	R848	M761	L489	L489	L489	L489

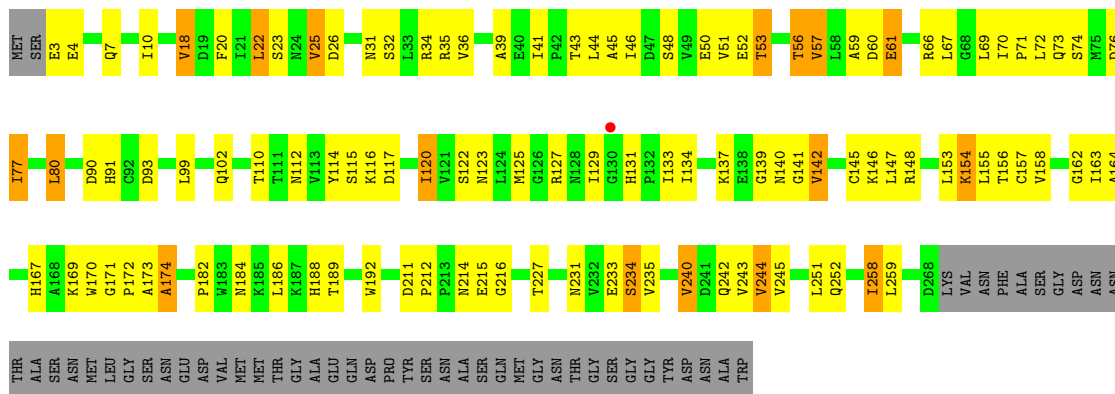
THR	THR
SER	SER
PRO	PRO
ALA	ALA
LEU	LEU
GLN	GLN
ASP	ASP
PRO	PRO
THR	THR
GLY	GLY
LYS	LYS
SER	SER
ASN	ASN
HIS	HIS
ASN	ASN
GLU	GLU
ASN	ASN
GLU	GLU
ASP	ASP
PRO	PRO
THR	THR
ASN	ASN
SER	SER
PRO	PRO
THR	THR
SER	SER
PRO	PRO
THR	THR
TYR	TYR
SER	SER
E28	E28
D29	D29
S30	S30
K94	K94
W31	W31
V165	V165
A32	A32
V33	V33
I34	I34
F37	F37
M101	M101
V102	V102
F38	F38
R39	R39
E40	E40
K41	K41
G42	G42
V44	V44
L43	L43
L112	L112
Y113	Y113
S45	S45
Q46	Q46
Q47	Q47
L48	L48
F51	F51
F54	F54
N55	N55
D56	D56
Y57	Y57
G200	G200
S125	S125
T58	T58
L289	L289
G290	G290
L291	L291
I292	I292
L128	L128
F129	F129
K134	K134
ARG	ARG

● Molecule 5: DNA-directed RNA polymerase II 140 kDa polypeptide

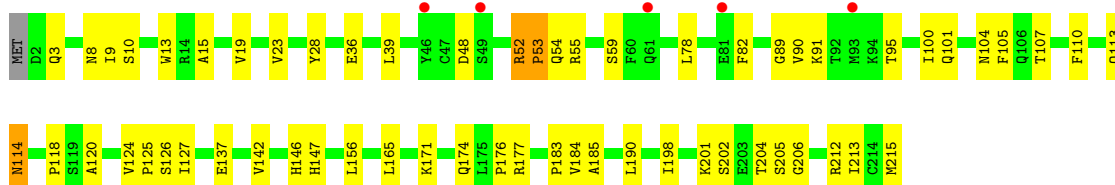
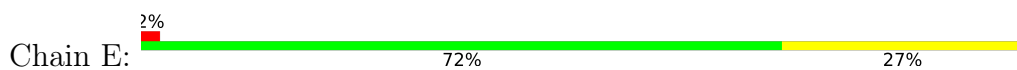




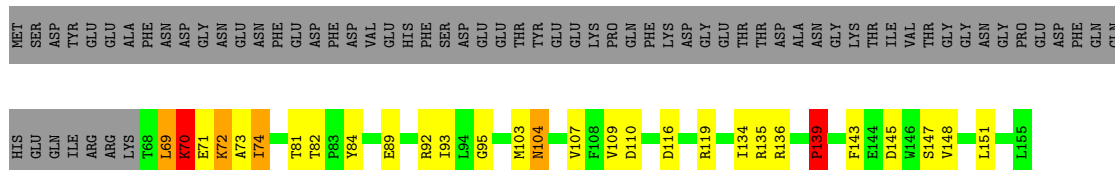
• Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide



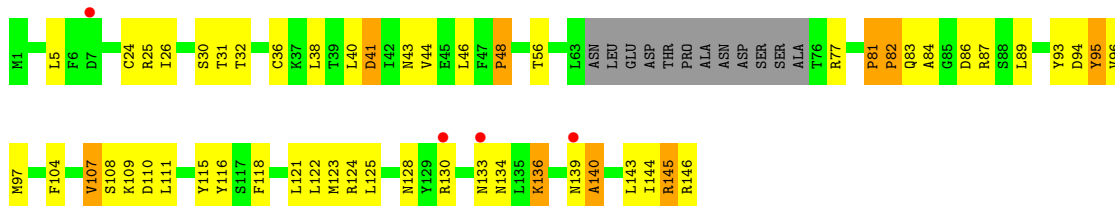
• Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide



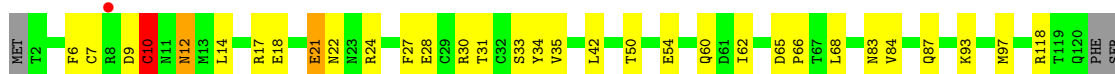
• Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



- Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



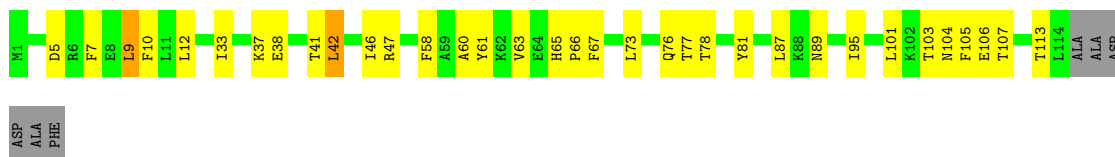
- Molecule 10: DNA-directed RNA polymerase II subunit 9



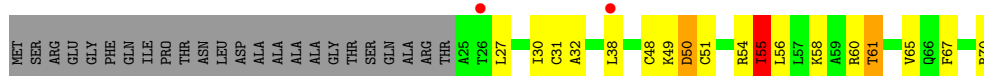
- Molecule 11: DNA-directed RNA polymerases I/II/III subunit 10



- Molecule 12: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 13: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.72Å 222.41Å 193.07Å 90.00° 101.29° 90.00°	Depositor
Resolution (Å)	50.00 – 3.60 44.25 – 3.60	Depositor EDS
% Data completeness (in resolution range)	91.5 (50.00-3.60) 91.6 (44.25-3.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.22 (at 3.57Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.285 , 0.304 0.282 , 0.296	Depositor DCC
$R_{free}$ test set	3721 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	104.2	Xtrriage
Anisotropy	0.228	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 59.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	29436	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	109.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

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

## 5 Model quality i

### 5.1 Standard geometry i

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

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	1.16	1/243 (0.4%)	1.75	6/378 (1.6%)
2	N	0.74	0/317	1.27	0/488
3	T	1.18	4/634 (0.6%)	1.76	18/975 (1.8%)
4	A	0.68	8/11224 (0.1%)	0.67	2/15176 (0.0%)
5	B	0.82	11/9027 (0.1%)	0.73	3/12172 (0.0%)
6	C	0.71	0/2133	0.69	0/2891
7	E	0.58	1/1788 (0.1%)	0.61	0/2406
8	F	0.60	0/724	0.72	0/977
9	H	0.57	0/1094	0.65	0/1480
10	I	0.75	2/989 (0.2%)	0.68	0/1331
11	J	0.69	0/541	0.68	0/727
12	K	0.66	0/937	0.65	0/1265
13	L	0.70	0/366	0.78	0/485
All	All	0.74	27/30017 (0.1%)	0.76	29/40751 (0.1%)

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	32
5	B	0	12
8	F	0	1
All	All	0	45

The worst 5 of 27 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	764	SER	CB-OG	13.86	1.60	1.42
5	B	490	SER	CB-OG	9.83	1.55	1.42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	879	GLU	CD-OE1	7.99	1.34	1.25
3	T	19	DT	C5-C7	7.74	1.54	1.50
5	B	404	LYS	CE-NZ	7.54	1.67	1.49

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	19	DT	C4-C5-C7	10.26	125.16	119.00
3	T	19	DT	C6-C5-C7	-9.91	116.95	122.90
3	T	19	DT	N3-C2-O2	-9.91	116.36	122.30
3	T	16	DC	O4'-C1'-N1	9.72	114.80	108.00
3	T	28	DT	O4'-C1'-N1	8.42	113.89	108.00

There are no chirality outliers.

5 of 45 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	153	PRO	Peptide
4	A	70	CYS	Peptide
4	A	71	GLN	Peptide
4	A	79	GLY	Peptide
4	A	80	HIS	Peptide

## 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	R	216	0	109	7	0
2	N	284	0	161	1	0
3	T	566	0	316	11	0
4	A	11028	0	11120	901	0
5	B	8856	0	8897	724	0
6	C	2095	0	2051	80	0
7	E	1752	0	1776	33	0
8	F	712	0	738	47	0
9	H	1076	0	1052	40	0
10	I	971	0	930	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	J	532	0	542	41	0
12	K	919	0	929	20	0
13	L	364	0	387	20	0
14	A	2	0	0	1	0
14	B	1	0	0	1	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	56	0	22	1	0
All	All	29436	0	29030	1745	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 30.

The worst 5 of 1745 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)
5:B:404:LYS:NZ	5:B:404:LYS:CE	1.67	1.52
4:A:1287:TYR:CD2	4:A:1305:VAL:HB	1.59	1.36
4:A:1229:SER:HB2	4:A:1236:LEU:CD1	1.59	1.33
4:A:1229:SER:CB	4:A:1236:LEU:HD12	1.64	1.27
5:B:203:PHE:CA	5:B:204:ILE:HD12	1.63	1.27

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
4	A	1392/1733 (80%)	1118 (80%)	182 (13%)	92 (7%)	<b>1</b> <b>16</b>

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	B	1096/1224 (90%)	897 (82%)	124 (11%)	75 (7%)	1	15
6	C	264/318 (83%)	229 (87%)	28 (11%)	7 (3%)	5	35
7	E	212/215 (99%)	190 (90%)	15 (7%)	7 (3%)	4	31
8	F	86/155 (56%)	75 (87%)	6 (7%)	5 (6%)	1	18
9	H	130/146 (89%)	105 (81%)	18 (14%)	7 (5%)	2	19
10	I	117/122 (96%)	96 (82%)	18 (15%)	3 (3%)	5	35
11	J	63/70 (90%)	53 (84%)	8 (13%)	2 (3%)	4	31
12	K	112/120 (93%)	107 (96%)	5 (4%)	0	100	100
13	L	44/70 (63%)	29 (66%)	13 (30%)	2 (4%)	2	23
All	All	3516/4173 (84%)	2899 (82%)	417 (12%)	200 (6%)	1	18

5 of 200 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	55	ASP
4	A	56	PRO
4	A	74	MET
4	A	157	ASP
4	A	158	PRO

### 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	
4	A	1225/1520 (81%)	1128 (92%)	97 (8%)	12	44
5	B	967/1061 (91%)	886 (92%)	81 (8%)	11	42
6	C	234/274 (85%)	211 (90%)	23 (10%)	8	36
7	E	196/197 (100%)	181 (92%)	15 (8%)	13	45
8	F	78/137 (57%)	74 (95%)	4 (5%)	24	58
9	H	118/128 (92%)	107 (91%)	11 (9%)	9	38
10	I	113/116 (97%)	103 (91%)	10 (9%)	10	40

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	J	60/65 (92%)	52 (87%)	8 (13%)	4	23
12	K	99/102 (97%)	87 (88%)	12 (12%)	5	26
13	L	40/57 (70%)	36 (90%)	4 (10%)	7	35
All	All	3130/3657 (86%)	2865 (92%)	265 (8%)	10	41

5 of 265 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	I	10	CYS
10	I	50	THR
12	K	113	THR
5	B	22	SER
4	A	1366	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 68 such sidechains are listed below:

Mol	Chain	Res	Type
7	E	101	GLN
7	E	114	ASN
10	I	90	GLN
5	B	121	ASN
4	A	1432	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/10 (80%)	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](#)

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, 9 are monoatomic - leaving 2 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	DUT	B	1308[A]	-	25,29,29	0.93	1 (4%)	37,45,45	1.59	6 (16%)
16	DUT	B	1308[B]	15	25,29,29	1.05	2 (8%)	37,45,45	1.73	7 (18%)

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	DUT	B	1308[A]	-	-	6/22/34/34	0/2/2/2
16	DUT	B	1308[B]	15	-	4/22/34/34	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	1308[A]	DUT	C6-C5	2.26	1.40	1.35
16	B	1308[B]	DUT	C2-N1	-2.20	1.34	1.38
16	B	1308[B]	DUT	C4-N3	-2.03	1.34	1.38

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	1308[B]	DUT	N3-C2-N1	4.57	120.95	114.89
16	B	1308[B]	DUT	C4-N3-C2	-4.27	120.95	126.58
16	B	1308[A]	DUT	N3-C2-N1	4.14	120.39	114.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	1308[A]	DUT	C4-N3-C2	-3.90	121.44	126.58
16	B	1308[A]	DUT	O4-C4-C5	-3.46	119.07	125.16

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

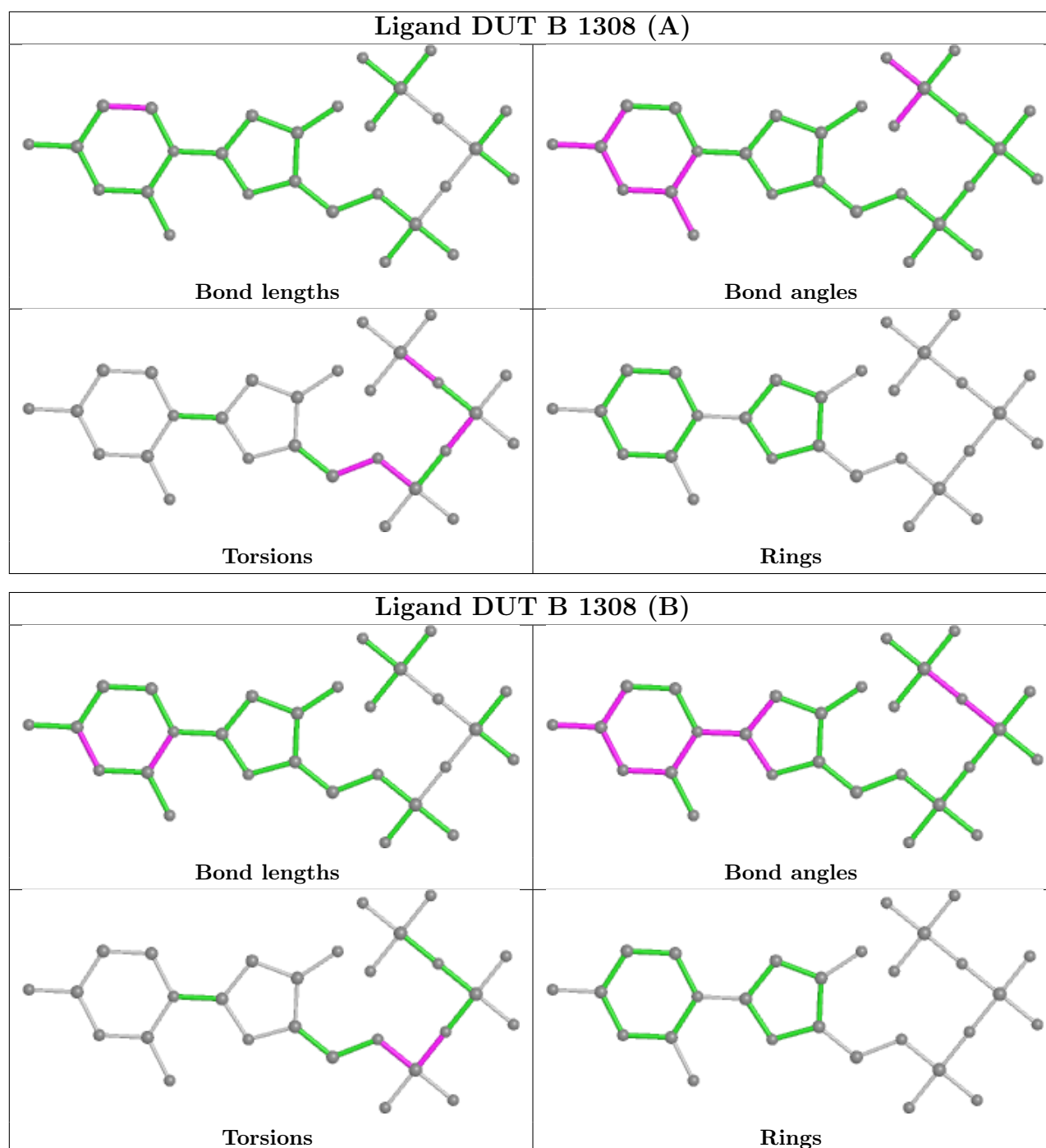
Mol	Chain	Res	Type	Atoms
16	B	1308[B]	DUT	C5'-O5'-PA-O1A
16	B	1308[B]	DUT	C5'-O5'-PA-O2A
16	B	1308[A]	DUT	PB-O3B-PG-O1G
16	B	1308[A]	DUT	C4'-C5'-O5'-PA
16	B	1308[A]	DUT	PB-O3B-PG-O2G

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	1308[A]	DUT	1	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	R	10/10 (100%)	-0.58	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	85, 109, 154, 170	0
2	N	14/14 (100%)	1.46	3 (21%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	178, 182, 198, 202	0
3	T	28/28 (100%)	0.74	7 (25%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	93, 173, 189, 192	0
4	A	1402/1733 (80%)	-0.04	21 (1%) <span style="border: 1px solid blue; padding: 2px;">73</span> <span style="border: 1px solid blue; padding: 2px;">60</span>	75, 103, 150, 163	0
5	B	1114/1224 (91%)	-0.05	17 (1%) <span style="border: 1px solid blue; padding: 2px;">73</span> <span style="border: 1px solid blue; padding: 2px;">60</span>	26, 102, 133, 143	0
6	C	266/318 (83%)	-0.23	1 (0%) <span style="border: 1px solid blue; padding: 2px;">92</span> <span style="border: 1px solid blue; padding: 2px;">86</span>	81, 100, 128, 144	0
7	E	214/215 (99%)	0.07	5 (2%) <span style="border: 1px solid blue; padding: 2px;">60</span> <span style="border: 1px solid red; padding: 2px;">44</span>	91, 131, 165, 167	0
8	F	88/155 (56%)	-0.12	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	84, 104, 130, 137	0
9	H	134/146 (91%)	0.12	4 (2%) <span style="border: 1px solid blue; padding: 2px;">50</span> <span style="border: 1px solid red; padding: 2px;">34</span>	100, 119, 152, 155	0
10	I	119/122 (97%)	-0.08	1 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">75</span>	87, 104, 122, 138	0
11	J	65/70 (92%)	-0.28	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	92, 102, 122, 125	0
12	K	114/120 (95%)	-0.28	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	82, 105, 119, 120	0
13	L	46/70 (65%)	0.31	2 (4%) <span style="border: 1px solid red; padding: 2px;">35</span> <span style="border: 1px solid red; padding: 2px;">22</span>	108, 145, 156, 157	0
All	All	3614/4225 (85%)	-0.04	61 (1%) <span style="border: 1px solid blue; padding: 2px;">70</span> <span style="border: 1px solid blue; padding: 2px;">55</span>	26, 104, 150, 202	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	69	THR	5.0
4	A	72	GLU	4.8
5	B	338	GLY	4.0
7	E	46	TYR	4.0
4	A	44	THR	3.9

### 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](#)

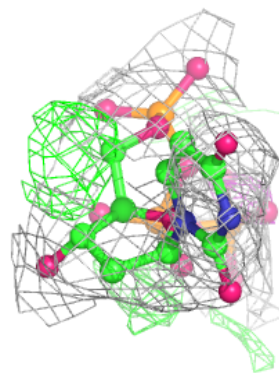
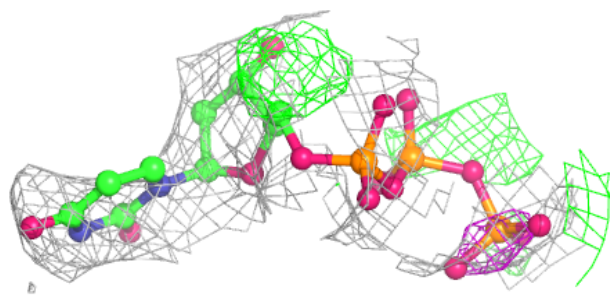
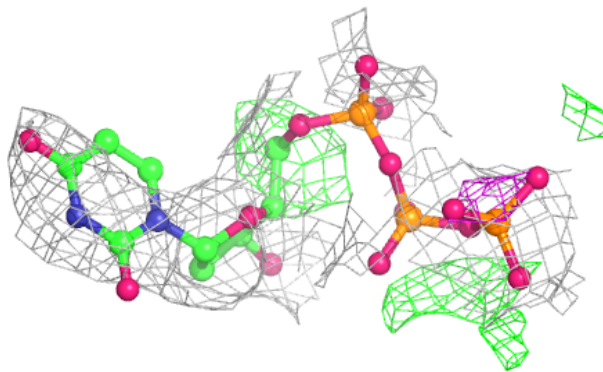
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
16	DUT	B	1308[A]	28/28	0.79	0.35	131,133,149,150	28
16	DUT	B	1308[B]	28/28	0.79	0.35	57,59,61,62	28
14	ZN	A	1734	1/1	0.89	0.05	159,159,159,159	0
14	ZN	I	203	1/1	0.93	0.07	97,97,97,97	0
14	ZN	B	1307	1/1	0.96	0.10	126,126,126,126	0
15	MG	A	2000	1/1	0.97	0.22	74,74,74,74	0
14	ZN	I	204	1/1	0.97	0.09	107,107,107,107	0
14	ZN	L	105	1/1	0.97	0.06	181,181,181,181	0
14	ZN	A	1735	1/1	0.98	0.09	142,142,142,142	0
14	ZN	J	101	1/1	0.98	0.12	149,149,149,149	0
14	ZN	C	319	1/1	0.99	0.06	104,104,104,104	0

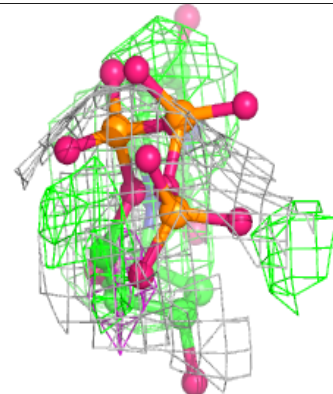
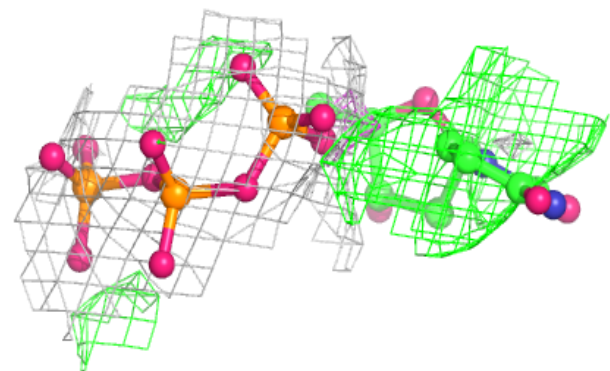
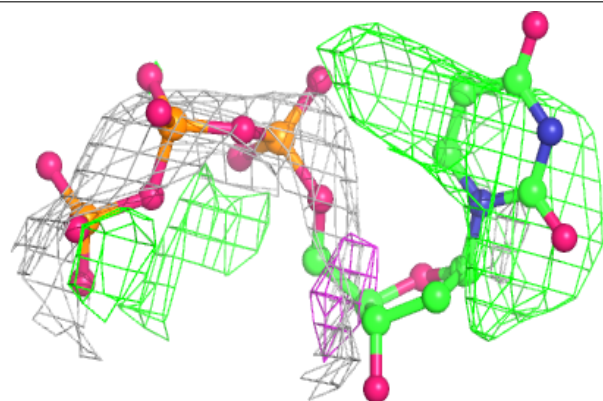
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around DUT B 1308 (A):**

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

**Electron density around DUT B 1308 (B):**

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





## 6.5 Other polymers

There are no such residues in this entry.