



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 19, 2023 – 07:49 AM EDT

PDB ID : 2NVQ
Title : RNA Polymerase II Elongation Complex in 150 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 : 2.90 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity : 4.02b-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

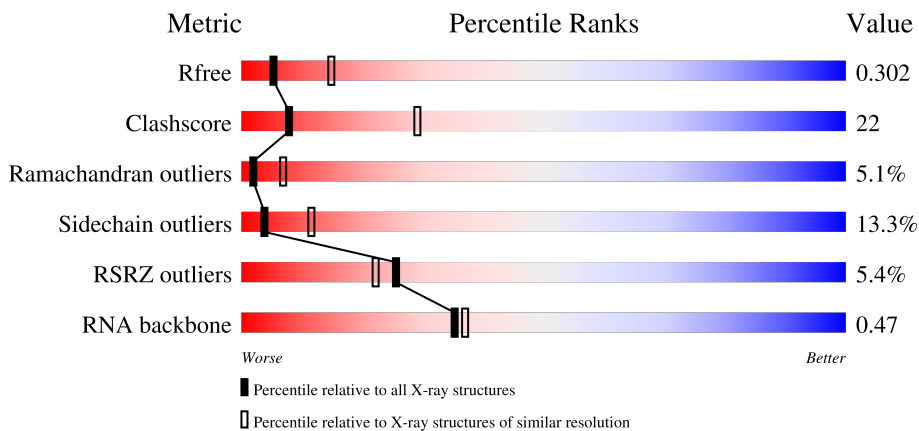
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)
RNA backbone	3102	1007 (3.16-2.64)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	R	10	
2	T	28	
3	N	14	
4	A	1733	

Continued on next page...

Continued from previous page...

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

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
14	DUT	T	29[B]	-	-	X	-
16	MG	A	2002[A]	-	-	-	X
16	MG	A	2002[B]	-	-	-	X

2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 29425 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 28-MER DNA template strand.

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

- Molecule 3 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			
3	N	14	284	137	49	85	13	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	1405	11043	6965	1936	2081	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	8861	5610	1549	1647	55	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	85	688	439	116	130	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	133	1068	673	180	211	4	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.

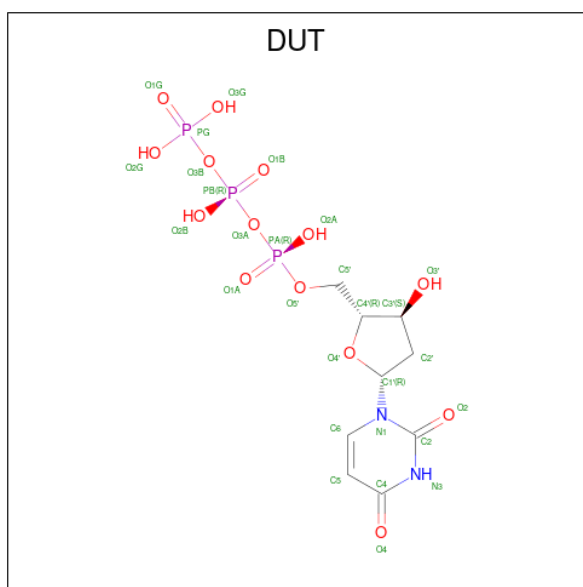
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	K	114	919	590	156	171	2	0	0	0

- 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
			Total	C	N	O	S			
13	L	46	363	224	72	63	4	0	0	0

- Molecule 14 is DEOXYURIDINE-5'-TRIPHOSPHATE (three-letter code: DUT) (formula: $C_9H_{15}N_2O_{14}P_3$).



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

- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	2	Total 3	Mg 3	0	1

3 Residue-property plots [i](#)

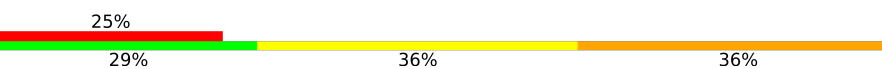
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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'

Chain R: 



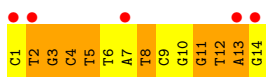
- Molecule 2: 28-MER DNA template strand

Chain T: 



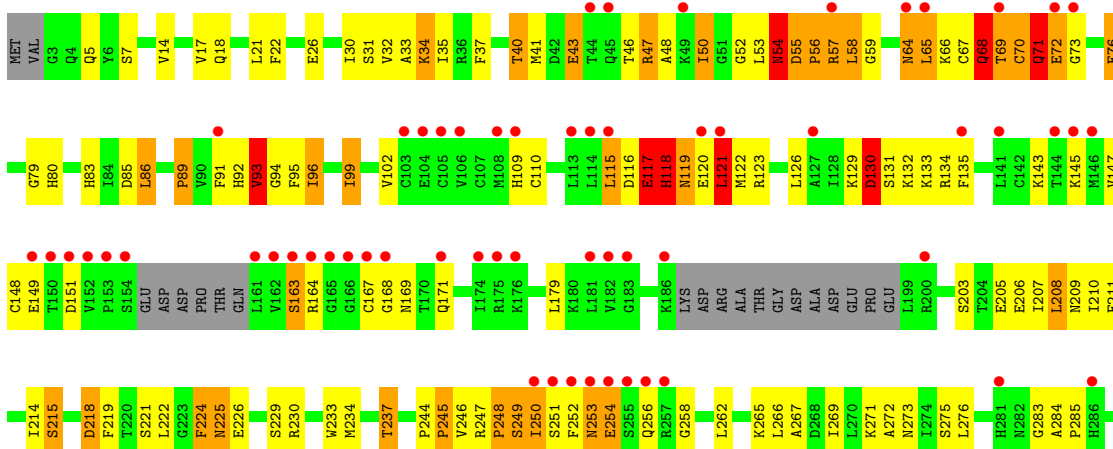
- Molecule 3: 5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3'

Chain N: 

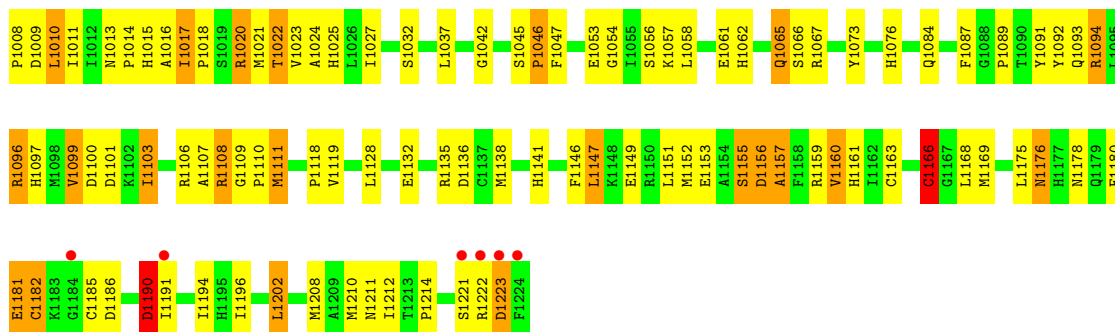


- Molecule 4: DNA-directed RNA polymerase II largest subunit

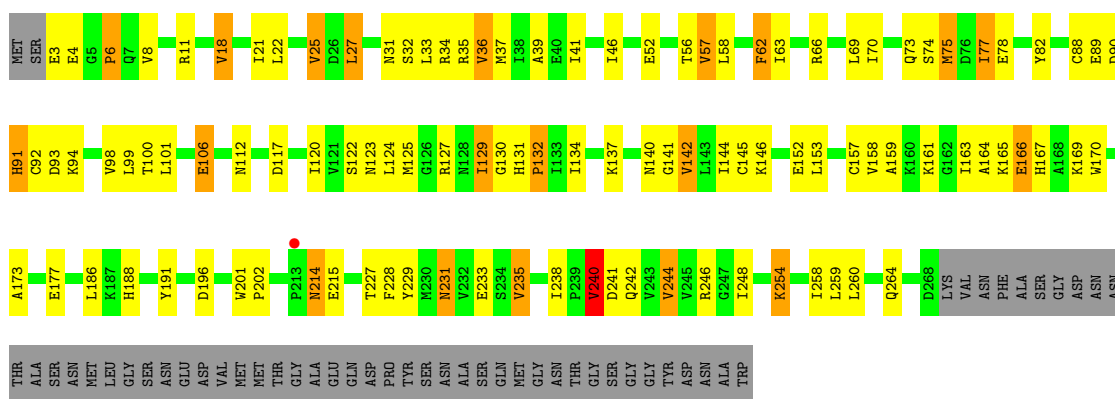
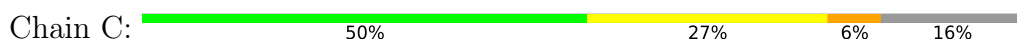
Chain A: 



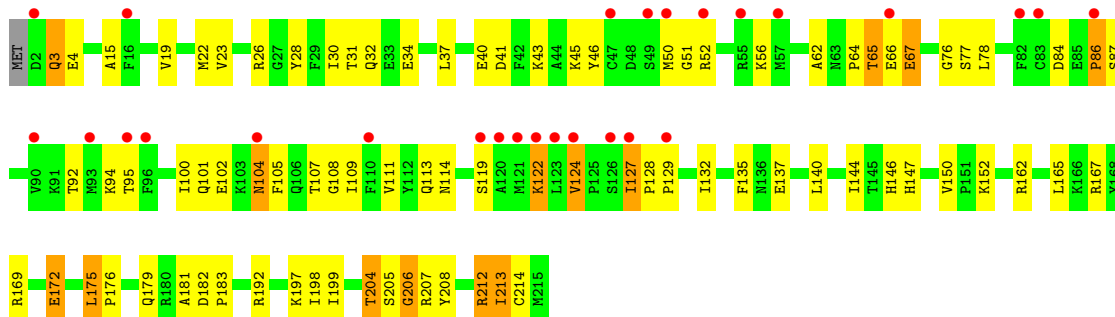
SER	Y1349	T1438	SER	L943	D853	V747	K651	T535	M456	E290
PRO	M1354	F1441	ASP	R944	M854	M748	N654	I541	A457	E291
SER	V1355	M1063	ALA	E946	T855	S754	G655	I541	A371	A292
SER	I1356	V1064	THR	V946	R857	F754	M656	Q545	K372	E293
PRO	D1359	T1077	GLU	D949	M858	I756	L657	V546	E378	L296
THR	M1364	T1080	A1294	V958	Q855	M757	L658	M549	V379	Q297
PRO	V1365	L1081	L1260	M959	F866	Q760	H659	R466	V390	F298
SER	R1366	M1082	K1261	I960	I867	M761	G661	R469	H299	H299
SER	D1370	F1083	E1264	R961	T868	C764	I666	L470	V381	V300
PRO	L1371	H1084	M1267	R962	Q869	I768	G667	N471	V383	A301
THR	V1372	F1085	L1174	Q965	E870	Q768	D668	L472	N384	T302
PRO	D1373	H1087	H1173	M966	A875	S769	I669	V473	R387	M304
PRO	L1376	G1088	S1175	Q969	T885	E771	D672	T475	E398	D305
PRO	G1388	A1090	L1176	I973	I886	R774	G673	I476	N306	N306
THR	R1381	S1091	ASP	I973	G887	I775	P674	T477	D307	D307
SER	T1382	K1092	GLU	R977	R896	F779	M676	S476	P400	I308
PRO	S1383	M1093	ALA	P978	M902	M677	G677	V477	G401	A309
PRO	V1384	T1095	ALA	S979	Y897	G484	R677	P477	A402	Q311
SER	V1394	M1095	GLN	D980	R898	L784	E678	Y478	Y404	Q312
THR	G1388	L1105	GLN	R981	V898	P785	I679	I406	I406	Q313
PRO	R1391	L1109	PHE	T982	D900	H786	K688	R407	D408	A314
PRO	S1392	M1110	ASP	I983	L901	S793	L691	S409	S409	Q316
PRO	T1393	M1111	GLU	K984	N903	P794	D692	G410	G410	A317
PRO	T1394	K1112	GLU	L988	T904	K797	G694	D411	D411	S318
THR	M1398	P1113	ALA	D992	D905	I898	T694	R412	R412	G319
PRO	R1399	P1114	ALA	R992	H906	V800	T694	D413	D413	R320
PRO	C1400	S1115	L1192	R1001	T907	E901	E696	D414	D414	P321
THR	F1402	T1117	L1193	G1002	L908	E901	E696	R415	R415	K323
PRO	E1403	V1118	L1195	K1003	N909	N802	E500	R416	R416	I325
PRO	E1404	G1123	D1206	M1004	P910	L808	L702	S418	S418	R326
THR	V1406	H1124	Q1211	E1005	L913	E812	T704	D423	D423	K332
PRO	E1407	D1127	Q1215	Q1008	G916	R821	H706	I424	I424	E333
PRO	F1410	Q1128	R1215	A1014	T919	E822	G707	L426	L426	R335
PRO	D1419	E1129	Q1218	V1015	L920	T709	M708	E433	E433	M341
THR	C1421	A1130	Q1218	T1016	G921	L825	L710	R434	R434	K343
THR	R1422	A1131	K1221	L1017	D922	D826	R711	H435	H435	R344
PRO	D1426	K1132	K1221	F1018	L923	T827	E715	I436	I436	V345
PRO	N1427	L1133	I1227	C1019	Q926	A828	R620	D440	D440	T351
PRO	L1430	I1134	I1227	C1019	V829	V829	T621	L443	L443	V352
PRO	S1426	I1138	E1234	R1025	L929	K830	E833	F444	F444	I353
PRO	E1426	T1142	K1235	R1029	E932	E933	V842	N445	N445	L359
PRO	N1427	V1146	R1239	R1030	L936	V842	V842	R446	R446	E360
PRO	V1428	T1147	R1239	L1037	V937	L845	L845	I523	I523	L359
PRO	L1430	I1148	V1243	L1037	K938	E846	H451	L450	L450	Q363
THR	P1435	I1148	ARG	D1043	D939	D847	V743	K452	K452	V364
PRO	I1436	Y1154	PRO	W1044	R940	I848	K744	H453	H453	G365
PRO	G1437	P1156	LYS	V1058	K941	I948	Q745	M455	M455	V366
THR			SER	H1059	F942	Y852	M746			P367



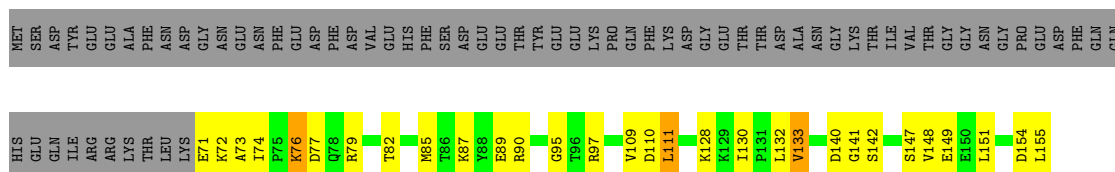
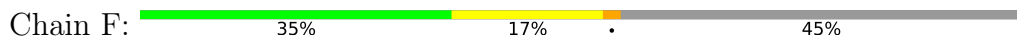
● Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide



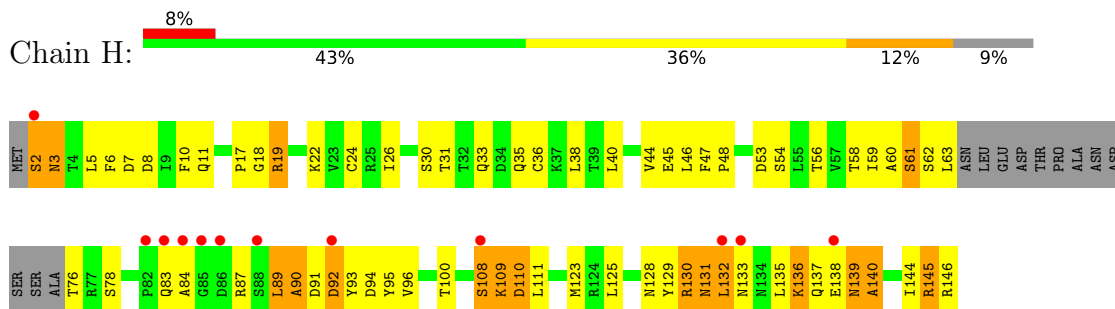
● Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide



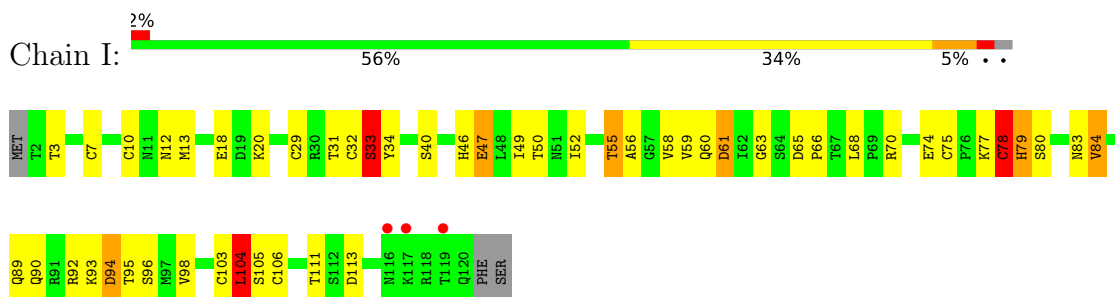
● Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



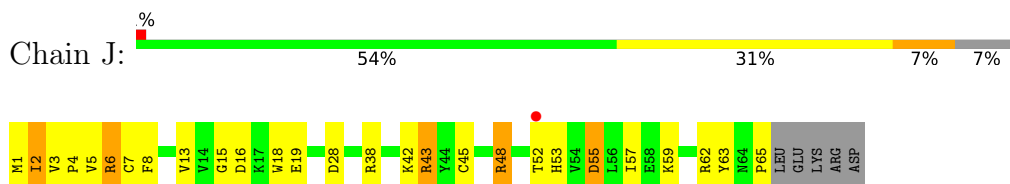
- Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



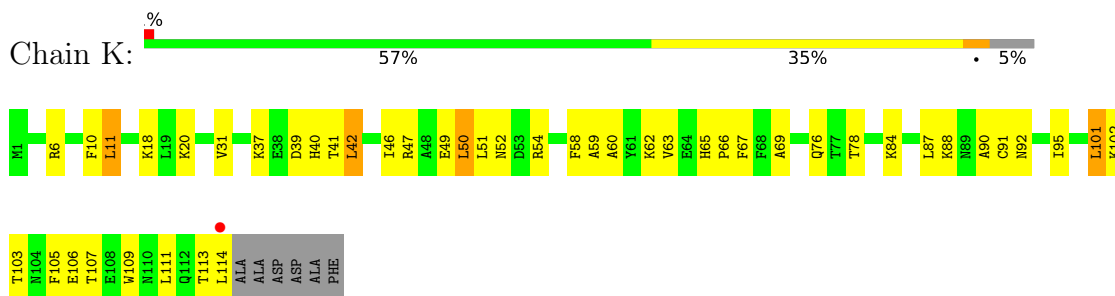
- Molecule 10: DNA-directed RNA polymerase II subunit 9



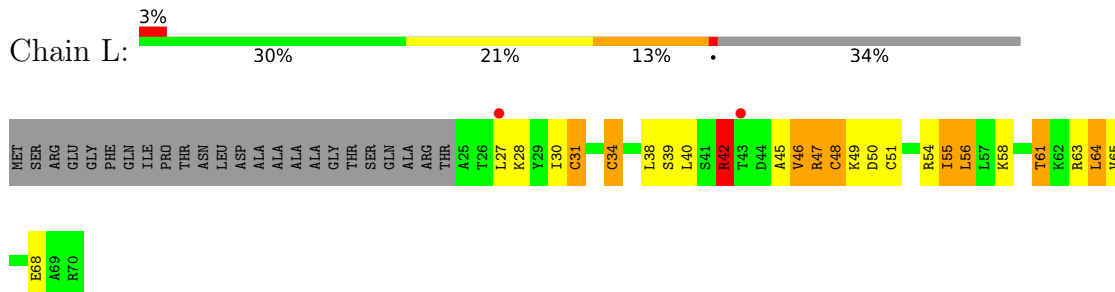
- Molecule 11: DNA-directed RNA polymerases I/II/III subunit 10



- 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, α , β , γ	170.65Å 222.76Å 195.28Å 90.00° 101.31° 90.00°	Depositor
Resolution (Å)	48.45 – 2.90 48.14 – 2.90	Depositor EDS
% Data completeness (in resolution range)	92.7 (48.45-2.90) 92.7 (48.14-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 2.91Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.229 , 0.283 0.263 , 0.302	Depositor DCC
R_{free} test set	4429 reflections (3.03%)	wwPDB-VP
Wilson B-factor (Å ²)	59.7	Xtrriage
Anisotropy	0.202	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 28.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	29425	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.67% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG, 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	1.30	1/243 (0.4%)	2.35	19/378 (5.0%)
2	T	1.23	0/634	1.97	30/975 (3.1%)
3	N	1.30	0/317	2.06	26/488 (5.3%)
4	A	0.75	0/11241	0.79	1/15199 (0.0%)
5	B	0.86	2/9033 (0.0%)	0.87	6/12181 (0.0%)
6	C	0.83	2/2133 (0.1%)	0.84	0/2891
7	E	0.65	0/1788	0.69	0/2406
8	F	0.69	0/700	0.74	0/945
9	H	0.67	0/1086	0.79	0/1470
10	I	0.78	1/989 (0.1%)	0.85	2/1331 (0.2%)
11	J	0.89	0/541	0.88	0/727
12	K	0.78	0/937	0.82	0/1265
13	L	0.93	1/365 (0.3%)	1.08	1/485 (0.2%)
All	All	0.81	7/30007 (0.0%)	0.92	85/40741 (0.2%)

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	11
5	B	0	8
6	C	0	2
All	All	0	21

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	L	34	CYS	CB-SG	-6.03	1.72	1.82
6	C	88	CYS	CB-SG	-5.95	1.72	1.81

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	1	A	N9-C4	5.93	1.41	1.37
5	B	529	GLU	CG-CD	5.91	1.60	1.51
5	B	1087	PHE	CE2-CZ	5.37	1.47	1.37

The worst 5 of 85 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	18	DA	O4'-C4'-C3'	-10.86	99.48	106.00
2	T	16	DC	O4'-C1'-N1	9.82	114.87	108.00
1	R	8	G	C4'-C3'-C2'	-9.14	93.46	102.60
3	N	1	DC	O4'-C1'-N1	8.83	114.18	108.00
1	R	8	G	O4'-C1'-N9	-8.56	101.35	108.20

There are no chirality outliers.

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	115	LEU	Peptide
4	A	117	GLU	Peptide
4	A	218	ASP	Peptide
4	A	297	GLN	Peptide
4	A	85	ASP	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	4	0
2	T	566	0	316	21	0
3	N	284	0	161	8	0
4	A	11043	0	11133	534	0
5	B	8861	0	8884	458	0
6	C	2095	0	2051	91	0
7	E	1752	0	1776	56	0
8	F	688	0	707	19	0
9	H	1068	0	1040	52	0
10	I	971	0	927	32	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	38	0
12	K	919	0	929	43	0
13	L	363	0	386	14	0
14	T	56	0	22	16	0
15	A	2	0	0	0	0
15	B	1	0	0	0	0
15	C	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	L	1	0	0	0	0
16	A	3	0	0	0	0
All	All	29425	0	28983	1261	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 22.

The worst 5 of 1261 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:647:GLY:HA3	5:B:648:HIS:CB	1.51	1.34
5:B:800:GLN:HB3	11:J:52:THR:CG2	1.58	1.32
4:A:1364:ASN:ND2	4:A:1366:ARG:HG2	1.51	1.25
4:A:672:ASP:OD2	4:A:736:ASN:ND2	1.73	1.21
4:A:116:ASP:CB	4:A:117:GLU:HB2	1.73	1.17

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	1395/1733 (80%)	1153 (83%)	158 (11%)	84 (6%)	1 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	B	1096/1224 (90%)	938 (86%)	108 (10%)	50 (5%)	2	9
6	C	264/318 (83%)	229 (87%)	27 (10%)	8 (3%)	4	17
7	E	212/215 (99%)	177 (84%)	28 (13%)	7 (3%)	4	15
8	F	83/155 (54%)	71 (86%)	11 (13%)	1 (1%)	13	40
9	H	129/146 (88%)	94 (73%)	22 (17%)	13 (10%)	0	1
10	I	117/122 (96%)	94 (80%)	14 (12%)	9 (8%)	1	2
11	J	63/70 (90%)	58 (92%)	2 (3%)	3 (5%)	2	8
12	K	112/120 (93%)	104 (93%)	8 (7%)	0	100	100
13	L	44/70 (63%)	28 (64%)	10 (23%)	6 (14%)	0	0
All	All	3515/4173 (84%)	2946 (84%)	388 (11%)	181 (5%)	2	7

5 of 181 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	54	ASN
4	A	55	ASP
4	A	56	PRO
4	A	65	LEU
4	A	69	THR

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%)	1052 (86%)	173 (14%)	3	10
5	B	967/1061 (91%)	851 (88%)	116 (12%)	5	15
6	C	234/274 (85%)	205 (88%)	29 (12%)	4	14
7	E	196/197 (100%)	172 (88%)	24 (12%)	5	15
8	F	75/137 (55%)	69 (92%)	6 (8%)	12	33
9	H	117/128 (91%)	96 (82%)	21 (18%)	2	5
10	I	113/116 (97%)	96 (85%)	17 (15%)	3	9

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	J	60/65 (92%)	52 (87%)	8 (13%)	4	11
12	K	99/102 (97%)	90 (91%)	9 (9%)	9	28
13	L	40/57 (70%)	27 (68%)	13 (32%)	0	0
All	All	3126/3657 (86%)	2710 (87%)	416 (13%)	4	11

5 of 416 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	731	VAL
5	B	1196	ILE
12	K	50	LEU
5	B	807	ARG
5	B	983	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 89 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	794	ASN
6	C	167	HIS
5	B	887	HIS
5	B	1076	HIS
6	C	242	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/10 (80%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	U

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 13 ligands modelled in this entry, 11 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
14	DUT	T	29[A]	16	25,29,29	1.11	1 (4%)	37,45,45	1.79	7 (18%)
14	DUT	T	29[B]	16	25,29,29	1.12	1 (4%)	37,45,45	3.09	14 (37%)

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
14	DUT	T	29[A]	16	-	3/22/34/34	0/2/2/2
14	DUT	T	29[B]	16	-	4/22/34/34	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	29[B]	DUT	C2-N1	-2.49	1.34	1.38
14	T	29[A]	DUT	C2-N1	-2.49	1.34	1.38

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	T	29[B]	DUT	O4'-C4'-C5'	11.48	147.16	109.37
14	T	29[B]	DUT	O5'-C5'-C4'	7.74	135.64	108.99
14	T	29[B]	DUT	O2G-PG-O3B	-5.77	85.27	104.64
14	T	29[A]	DUT	C4-N3-C2	-5.27	119.63	126.58

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	T	29[A]	DUT	N3-C2-N1	4.50	120.86	114.89

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

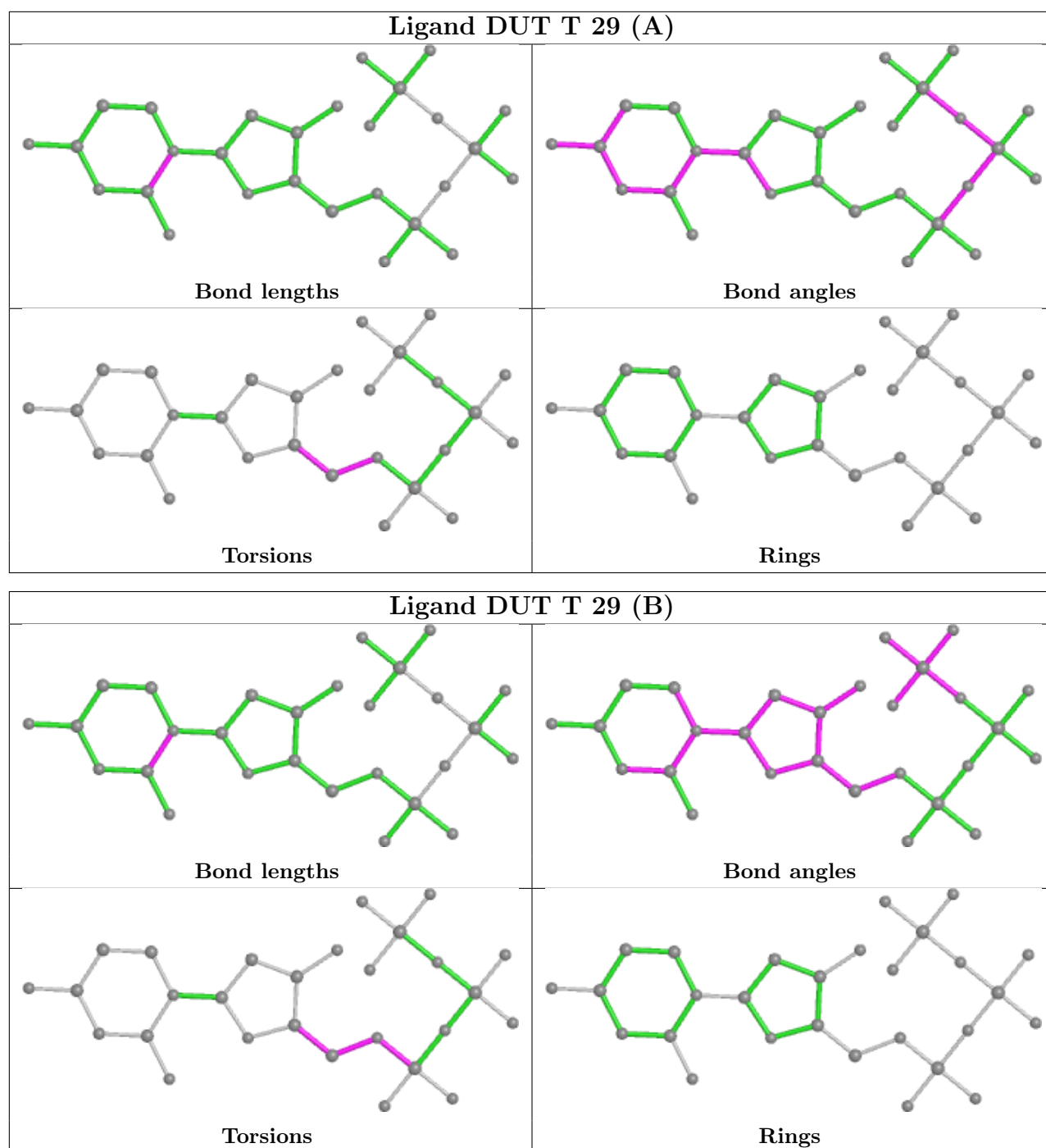
Mol	Chain	Res	Type	Atoms
14	T	29[A]	DUT	C3'-C4'-C5'-O5'
14	T	29[B]	DUT	C4'-C5'-O5'-PA
14	T	29[B]	DUT	C5'-O5'-PA-O2A
14	T	29[B]	DUT	C5'-O5'-PA-O3A
14	T	29[A]	DUT	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	T	29[A]	DUT	3	0
14	T	29[B]	DUT	13	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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	R	10/10 (100%)	-0.51	0 100 100	30, 50, 93, 109	0
2	T	28/28 (100%)	0.59	7 (25%) 0 0	40, 108, 139, 145	0
3	N	14/14 (100%)	1.35	5 (35%) 0 0	73, 119, 144, 149	0
4	A	1405/1733 (81%)	0.34	88 (6%) 20 16	24, 56, 77, 131	0
5	B	1114/1224 (91%)	0.37	48 (4%) 35 31	28, 55, 84, 104	0
6	C	266/318 (83%)	-0.02	1 (0%) 92 93	40, 54, 73, 99	0
7	E	214/215 (99%)	0.58	27 (12%) 3 3	43, 64, 88, 91	0
8	F	85/155 (54%)	-0.11	0 100 100	44, 59, 81, 90	0
9	H	133/146 (91%)	0.58	12 (9%) 9 7	52, 66, 88, 94	0
10	I	119/122 (97%)	0.06	3 (2%) 57 55	37, 51, 84, 96	0
11	J	65/70 (92%)	0.17	1 (1%) 73 73	44, 56, 72, 81	0
12	K	114/120 (95%)	0.06	1 (0%) 84 84	45, 57, 73, 82	0
13	L	46/70 (65%)	0.45	2 (4%) 35 31	57, 87, 98, 101	0
All	All	3613/4225 (85%)	0.32	195 (5%) 25 22	24, 57, 85, 149	0

The worst 5 of 195 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	1087	ALA	7.8
4	A	1082	ASN	7.8
5	B	250	PHE	7.4
5	B	869	SER	7.4
4	A	1089	VAL	6.7

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](#)

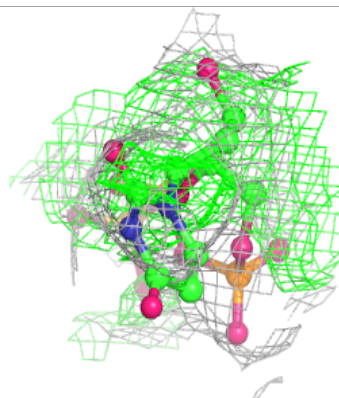
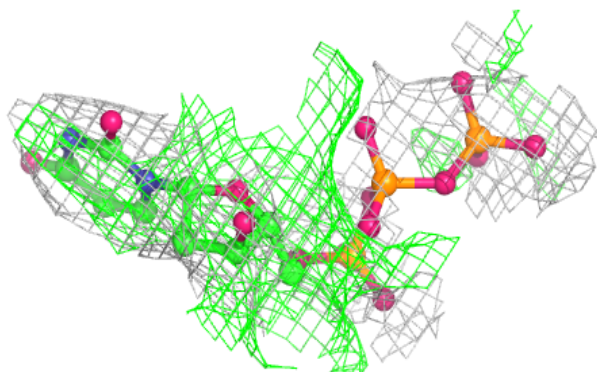
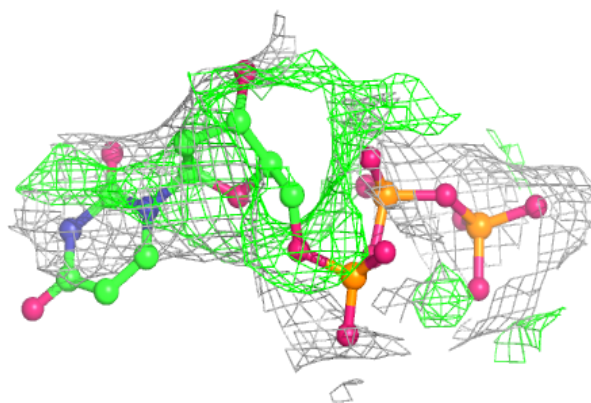
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
16	MG	A	2002[A]	1/1	0.72	0.60	43,43,43,43	1
16	MG	A	2002[B]	1/1	0.72	0.60	33,33,33,33	1
14	DUT	T	29[A]	28/28	0.73	0.40	76,80,97,98	28
14	DUT	T	29[B]	28/28	0.73	0.40	70,73,79,80	28
15	ZN	A	1734	1/1	0.77	0.11	84,84,84,84	0
15	ZN	I	203	1/1	0.89	0.17	83,83,83,83	0
16	MG	A	2001	1/1	0.95	0.29	46,46,46,46	0
15	ZN	B	1307	1/1	0.96	0.05	80,80,80,80	0
15	ZN	C	319	1/1	0.96	0.13	48,48,48,48	0
15	ZN	I	204	1/1	0.97	0.07	57,57,57,57	0
15	ZN	A	1735	1/1	0.98	0.07	78,78,78,78	0
15	ZN	J	101	1/1	0.98	0.20	50,50,50,50	0
15	ZN	L	105	1/1	0.98	0.06	91,91,91,91	0

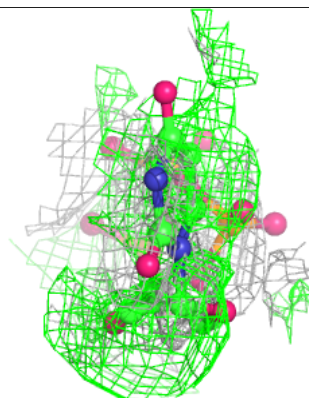
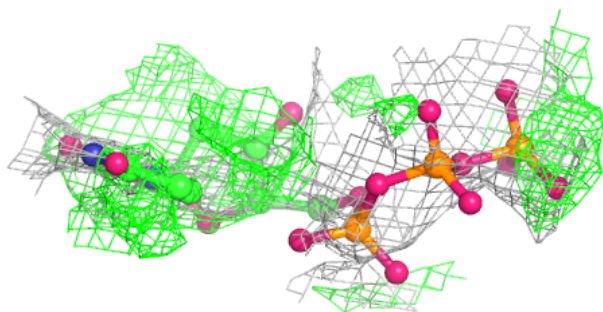
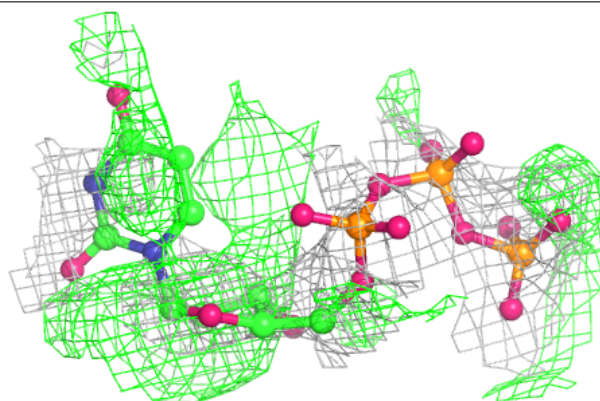
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DUT T 29 (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 T 29 (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 [i](#)

There are no such residues in this entry.