



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 5, 2022 – 04:20 PM EST

PDB ID : 7RIY  
Title : RNA polymerase II elongation complex with hairpin polyamide Py-Im 1, scaffold 2 soaked with UTP  
Authors : Oh, J.; Dervan, P.B.; Wang, D.  
Deposited on : 2021-07-20  
Resolution : 3.70 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 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 : 2.25  
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) : 2.25

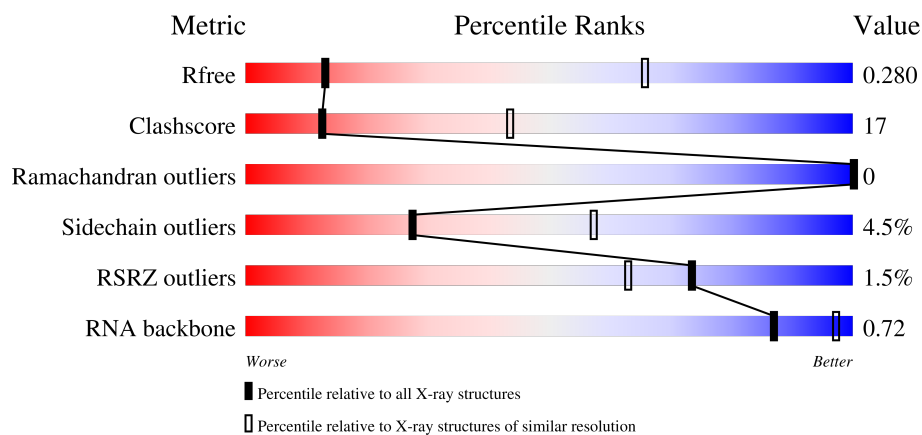
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)
RNA backbone	3102	1027 (4.40-3.00)

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	11	
2	T	30	
3	N	20	
4	A	1733	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	B	1224	 % 57% 33% • 8%
6	C	318	 50% 32% • 16%
7	E	215	 3% 64% 33% ••
8	F	155	 % 37% 17% • 45%
9	H	146	 2% 56% 32% • 9%
10	I	122	 52% 39% 5% •
11	J	70	 43% 49% • 7%
12	K	120	 65% 29% • 5%
13	L	70	 3% 26% 33% • 39%

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 29168 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	11	Total	C	N	O	P	0	0	0
			235	106	45	74	10			

- Molecule 2 is a DNA chain called Template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	26	Total	C	N	O	P	0	0	0
			525	252	84	163	26			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	14	Total	C	N	O	P	0	0	0
			293	138	63	78	14			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1384	Total	C	N	O	S	0	0	0
			10828	6831	1896	2041	60			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1123	Total	C	N	O	S	0	0	0
			8859	5607	1552	1647	53			

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	212	Total	C	N	O	S	0	0	0
			1731	1100	305	315	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	86	Total	C	N	O	S	0	0	0
			684	437	115	129	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total	C	N	O	S	0	0	0
			1064	670	179	211	4			

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	118	Total	C	N	O	S	0	0	0
			952	585	173	184	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

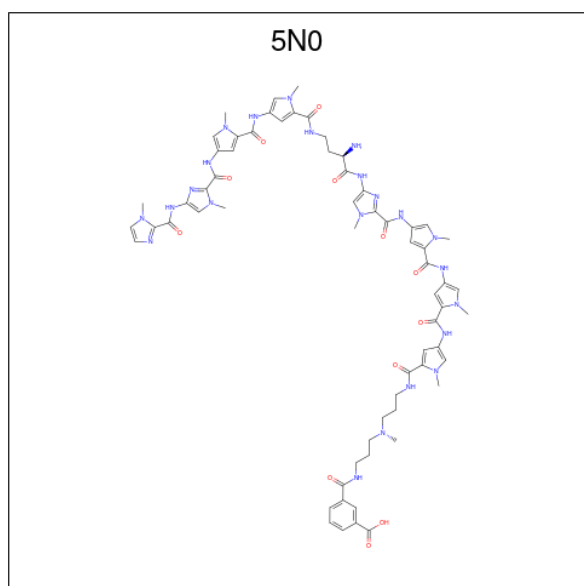
- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	43	Total	C	N	O	S	0	0	0
			337	208	66	59	4			

- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	R	1	Total	Mg	0	0
			1	1		

- Molecule 15 is 3-({3-[(3-[(4-[(4-[(2R)-2-amino-4-[(1-methyl-4-[(1-methyl-4-[(1-methyl-4-[(1-methyl-1H-imidazole-2-carbonyl)amino]-1H-imidazole-2-carbonyl}amino)-1H-pyrrole-2-carbonyl]amino}-1H-pyrrole-2-carbonyl)amino]butanoyl}amino)-1-methyl-1H-imidazole-2-carbonyl]amino}-1-methyl-1H-pyrrole-2-carbonyl)amino]-1-methyl-1H-pyrrole-2-carbonyl]amino}propyl)(methyl)amino]propyl}carbamoyl)benzoic acid (three-letter code: 5N0) (formula: C<sub>64</sub>H<sub>75</sub>N<sub>23</sub>O<sub>12</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
15	T	1	Total	C	N	O	0	0
			99	64	23	12		

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	L	1	Total	Zn	0	0
			1	1		

### 3 Residue-property plots

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

#### • Molecule 1: RNA

Chain R: 



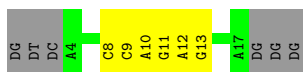
#### • Molecule 2: Template strand DNA

Chain T: 



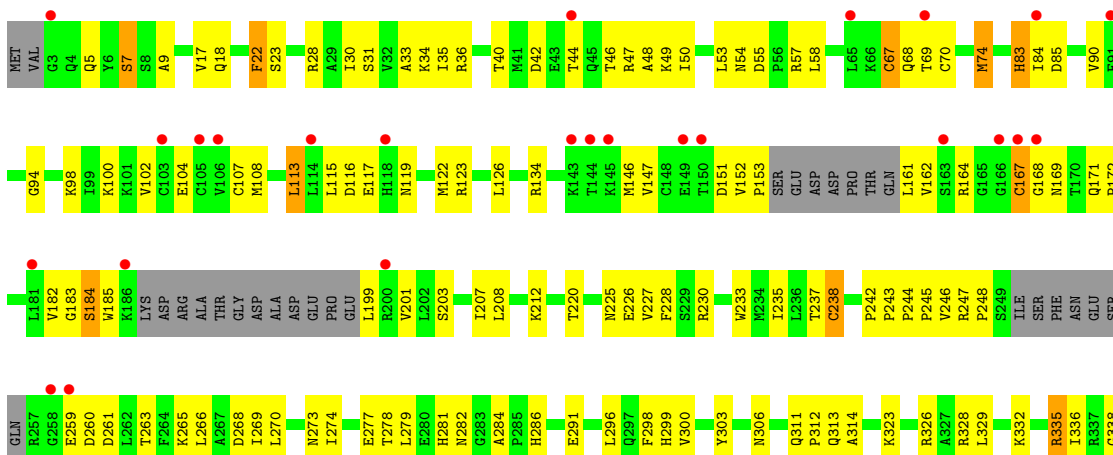
#### • Molecule 3: Non-template strand DNA

Chain N: 



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

Chain A: 

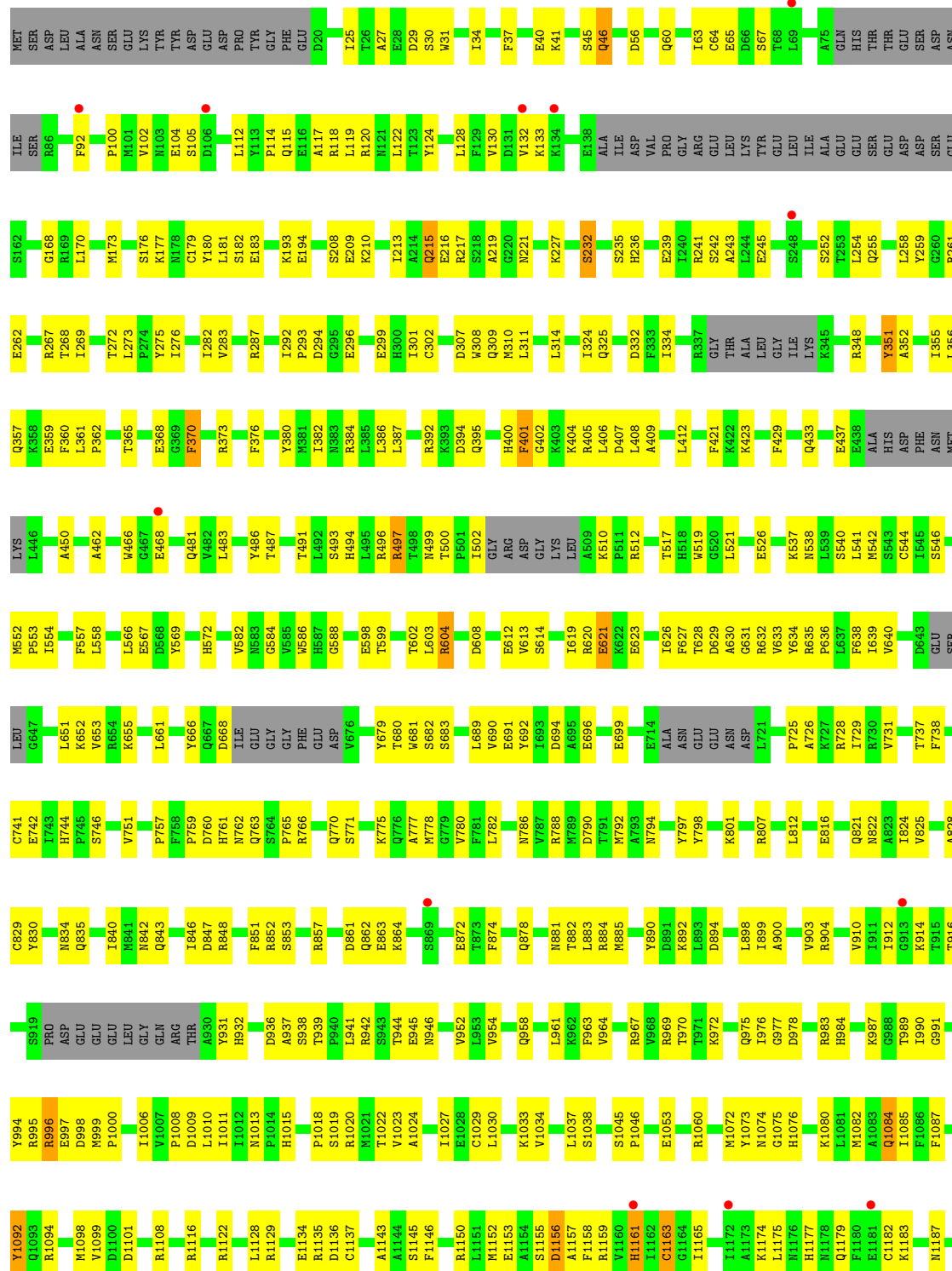


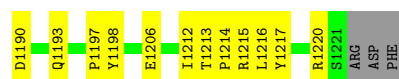


PRO	THR	PRO	THR	PRO	GLY	GLU	C1400	T1308	S1229	R1135	V1044	W954	T848	N717	1607	K533	P481	K343
THR	PRO	THR	PRO	THR	VAL	SER	S1401	M1312	E1230	I1138	V1045	P955	D853	N718	1608	L534	V442	K344
ALA	PRO	PRO	ALA	PRO	VAL	LEU	F1402	E1315	D1233	E1139	L1046	P956	D854	V719	6610	R537	F443	V345
THR	THR	THR	VAL	THR	GLY	ASN	E1403	E1316	K1235	H1140	A1051	N958	T855	F721	1612	D538	N445	D346
SER	SER	SER	V1406	SER	ALA	ASP	V1407	V1316	L1236	L1143	R1055	N959	T856	K738	1613	I541	R446	F347
PRO	PRO	PRO	F1410	PRO	PHE	ASP	F1411	V1319	I1237	T1147	P1060	N960	R857	Q745	F614	E542	Q447	R350
THR	THR	THR	E1411	THR	LEU	ASP	E1412	V1322	I1238	T1147	P1061	N961	R858	N741	6618	E543	T351	V352
SER	SER	SER	A1412	SER	VAL	ASP	A1413	D1323	C1239	E1151	G1082	N962	L860	N742	6619	D544	S449	
THR	THR	THR	E1417	THR	VAL	ASP	E1418	P1324	R1241	I1152	G1083	N963	F866	Q746	6622	M556	M455	D356
THR	THR	THR	D1420	THR	GLU	ASP	D1421	R1326	V1242	Y1153	G1084	Q969	F867	M746	6623	M557	M456	E360
PRO	PRO	PRO	C1421	PRO	GLU	ASP	C1422	I1327	V1243	P1156	V1086	Q970	E870	I759	6628	M558	M457	E361
THR	THR	THR	R1422	THR	THR	ASP	R1423	Y1328	ARG	E1165	L1087	N971	E871	S762	6629	L550	M458	D362
PRO	PRO	PRO	G1423	PRO	THR	ASP	G1424	T1329	PRO	D1166	A1088	N972	E872	S763	6630	L551	M459	D363
PRO	PRO	PRO	E1426	PRO	ALA	ASP	E1427	N1330	LEU	E1167	S1071	N973	A875	V765	6631	M556	M460	V364
THR	THR	THR	N1427	THR	GLU	ASP	N1428	S1331	ASP	E1168	I1072	N974	E876	V766	6632	D557	M461	G365
SER	SER	SER	I1436	SER	VAL	ASP	I1437	D1334	ALA	I1169	L1081	N975	E877	Q769	6633	G558	M462	V366
THR	THR	THR	G1437	THR	GLU	ASP	G1438	V1338	GLU	Q1171	ASN	N976	E878	E771	6636	V559	M463	K368
PRO	PRO	PRO	T1438	PRO	THR	ASP	T1439	L1339	THR	E1170	THR	N977	E879	E772	6637	M560	M464	I370
PRO	PRO	PRO	G1439	PRO	GLU	ASP	G1440	G1340	GLU	E1171	PHE	N978	E880	E773	6638	G559	M465	G366
SER	SER	SER	A1440	SER	ALA	ASP	A1441	I1341	GLU	S1175	HIS	N979	E881	I775	6639	D561	M466	V367
THR	THR	THR	F1441	THR	GLU	ASP	F1442	T1342	GLU	L1176	THR	N980	E882	I776	6640	P562	M467	K371
SER	SER	SER	I1443	SER	VAL	ASP	I1444	V1343	ALA	E1177	ALA	N981	E883	E795	6641	M562	M468	K372
THR	THR	THR	M1445	THR	GLY	ASP	M1446	E1351	L1260	ASP	GLY	N982	E884	E800	6642	G567	M469	Y376
PRO	PRO	PRO	D1446	PRO	GLY	ASP	D1447	V1352	GLU	GLU	GLY	N983	E885	E801	6643	M568	M470	P377
PRO	PRO	PRO	GLU	PRO	THR	ASP	GLU	V1353	ALA	GLU	SER	N984	E886	E802	6644	M569	M471	E378
THR	THR	THR	LEU	THR	ALA	ASP	LEU	E1354	L1261	ASP	GLN	N985	E887	E803	6645	M570	M472	V379
PRO	PRO	PRO	GLY	PRO	GLY	ASP	GLY	V1355	I1262	ALA	GLU	N986	E888	E804	6646	M571	M473	V380
THR	THR	THR	THR	THR	THR	ASP	THR	A1356	M1267	GLN	R1001	N987	E889	E805	6647	M572	M474	Y383
PRO	PRO	PRO	VAL	PRO	VAL	ASP	VAL	E1357	I1271	SER	R1002	N988	E890	E806	6648	M573	M475	D386
THR	THR	THR	LYS	THR	GLY	ASP	LYS	D1358	T1272	PHE	R1003	N989	E891	E807	6649	M574	M476	R387
SER	SER	SER	TYR	THR	GLY	ASP	TYR	D1359	L1273	ASP	R1004	N990	E892	E808	6650	M575	M477	L388
PRO	PRO	PRO	MET	PRO	GLY	ASP	MET	N1364	E1274	GLN	Q1008	N991	E893	E809	6651	M576	M478	
ASN	ASN	ASN	PRO	ASN	ALA	ASP	PRO	Y1365	R1275	Q1188	N1106	N992	E894	E810	6652	M577	M479	D386
TYR	TYR	TYR	GLU	TYR	GLN	ASP	GLU	R1366	V1276	P1190	N1107	N993	E895	E811	6653	M578	M480	R387
PRO	PRO	PRO	GLN	PRO	GLY	ASP	GLY	E1367	E1277	W1191	T907	N994	E896	E812	6654	M579	M481	L388
THR	THR	THR	LYS	THR	ILE	ASP	LYS	L1370	M1278	L1192	T1016	N995	E897	E813	6655	M580	M482	V392
SER	SER	SER	ILE	SER	THR	THR	THR	D1373	E1279	L1193	T1017	N996	E898	E814	6656	M581	M483	P396
PRO	PRO	PRO	GLU	PRO	THR	THR	GLU	G1380	R1281	R1194	L1021	N997	E899	E815	6657	M582	M484	K403
THR	THR	THR	THR	THR	THR	THR	THR	S1383	K1286	D1204	L1022	N998	E900	E816	6658	M583	M485	D407
PRO	PRO	PRO	GLY	PRO	THR	THR	GLY	V1384	D1205	L1204	R1023	N999	E901	E817	6659	M584	M486	D408
THR	THR	THR	ASP	THR	GLY	THR	ASP	T1385	D1206	L1205	R1024	N1000	E902	E818	6660	M585	M487	S409
SER	SER	SER	GLN	SER	GLY	THR	GLN	R1386	V1291	P1292	R1025	N1001	E903	E819	6661	M586	M488	
THR	THR	THR	THR	THR	THR	THR	THR	R1387	G1296	G1297	R1026	N1002	E904	E820	6662	M587	M489	R407
PRO	PRO	PRO	GLY	PRO	GLY	THR	GLY	H1388	E1297	I1216	S1024	N1003	E905	E821	6663	M588	M490	D408
GLY	GLY	GLY	VAL	GLY	ALA	THR	VAL	F1389	E1298	I1217	R1027	N1004	E906	E822	6664	M589	M491	S409
THR	THR	THR	THR	THR	THR	THR	THR	M1390	E1303	Q1218	L1028	N1005	E907	E823	6665	M590	M492	I424
PRO	PRO	PRO	THR	PRO	THR	THR	THR	L1397	W1304	Q1219	T1029	N1006	E908	E824	6666	M591	M493	Q427
GLY	GLY	GLY	THR	GLY	THR	THR	THR	M1398	V1305	I1224	R1030	N1007	E909	E825	6667	M592	M494	V432
SER	SER	SER	THR	SER	THR	THR	THR	L1399	L1306	L1225	R1031	N1008	E910	E826	6668	M593	M495	E433
PRO	PRO	PRO	THR	PRO	THR	THR	THR	R1399	E1307		V1031	N1009	E911	E827	6669	M594	M496	R434
GLY	GLY	GLY	THR	GLY	THR	THR	THR				V1032	N1010	E912	E828	6670	M595	M497	H435
SER	SER	SER	THR	SER	THR	THR	THR				V1033	N1011	E913	E829	6671	M596	M498	I436
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1034	N1012	E914	E830	6672	M597	M499	D440
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1035	N1013	E915	E831	6673	M598	M500	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1036	N1014	E916	E832	6674	M599	M501	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1037	N1015	E917	E833	6675	M600	M502	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1038	N1016	E918	E834	6676	M601	M503	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1039	N1017	E919	E835	6677	M602	M504	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1040	N1018	E920	E836	6678	M603	M505	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1041	N1019	E921	E837	6679	M604	M506	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1042	N1020	E922	E838	6680	M605	M507	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1043	N1021	E923	E839	6681	M606	M508	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1044	N1022	E924	E840	6682	M607	M509	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1045	N1023	E925	E841	6683	M608	M510	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1046	N1024	E926	E842	6684	M609	M511	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1047	N1025	E927	E843	6685	M610	M512	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1048	N1026	E928	E844	6686	M611	M513	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1049	N1027	E929	E845	6687	M612	M514	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1050	N1028	E930	E846	6688	M613	M515	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1051	N1029	E931	E847	6689	M614	M516	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1052	N1030	E932	E848	6690	M615	M517	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1053	N1031	E933	E849	6691	M616	M518	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1054	N1032	E934	E850	6692	M617	M519	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1055	N1033	E935	E851	6693	M618	M520	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1056	N1034	E936	E852	6694	M619	M521	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1057	N1035	E937	E853	6695	M620	M522	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1058	N1036	E938	E854	6696	M621	M523	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1059	N1037	E939	E855	6697	M622	M524	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1060	N1038	E940	E856	6698	M623	M525	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1061	N1039	E941	E857	6699	M624	M526	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1062	N1040	E942	E858	6700	M625	M527	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1063	N1041	E943	E859	6701	M626	M528	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1064	N1042	E944	E860	6702	M627	M529	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1065	N1043	E945	E861	6703	M628	M530	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1066	N1044	E946	E862	6704	M629	M531	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1067	N1045	E947	E863	6705	M630	M532	
PRO	PRO	PRO	THR	PRO	THR	THR	THR				V1068	N1046	E948	E864	6706	M631	M533	

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

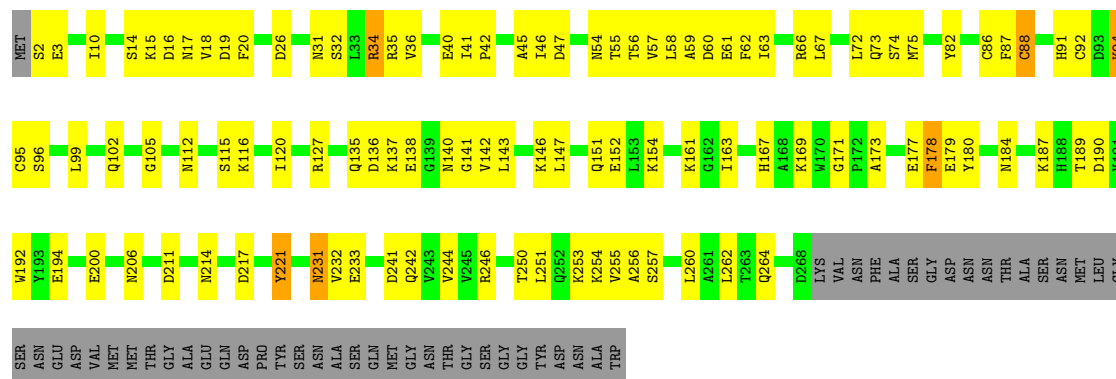
● Molecule 5: DNA-directed RNA polymerase II subunit RPB2





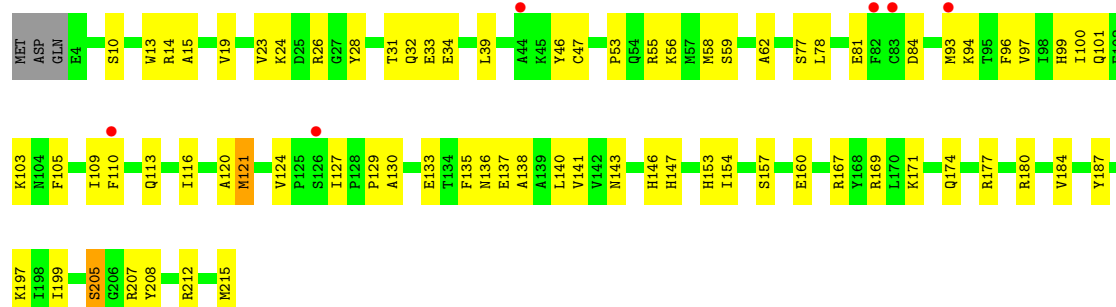
• Molecule 6: DNA-directed RNA polymerase II subunit RPB3

Chain C: 50% 32% 16%



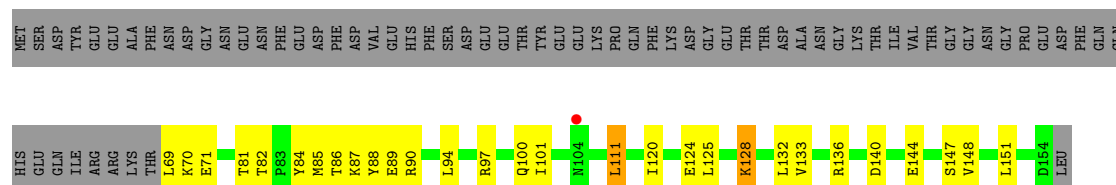
• Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 3% 64% 33% 2%



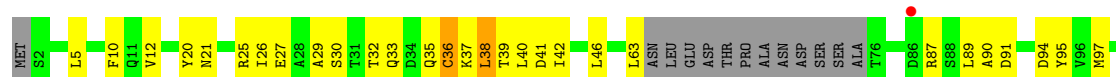
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 37% 17% 45%



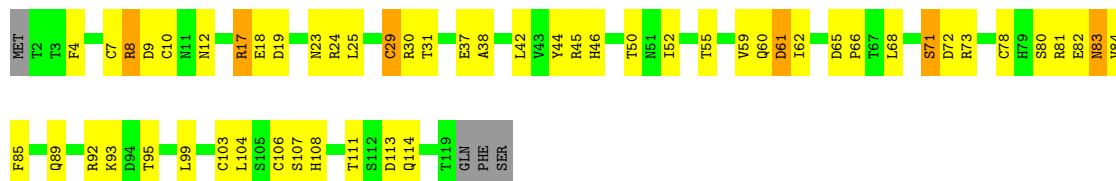
• Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 2% 56% 32% 9%





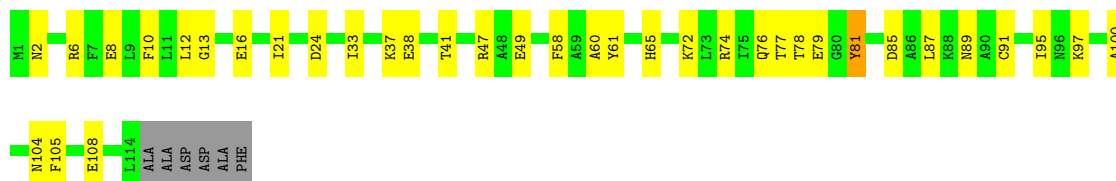
- Molecule 10: DNA-directed RNA polymerase II subunit RPB9



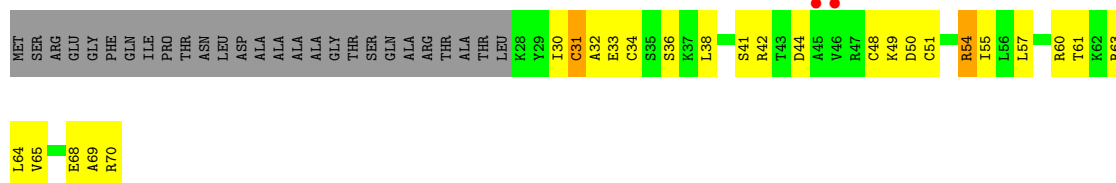
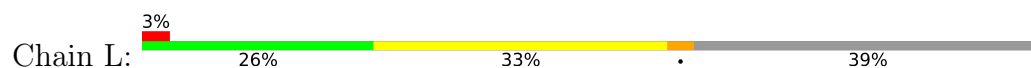
- Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC5



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



- Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC4



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	167.78Å 222.76Å 192.96Å 90.00° 100.54° 90.00°	Depositor
Resolution (Å)	49.53 – 3.70 49.53 – 3.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.53-3.70) 99.9 (49.53-3.70)	Depositor EDS
$R_{merge}$	0.52	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.30 (at 3.67Å)	Xtriage
Refinement program	PHENIX 1.13	Depositor
R, $R_{free}$	0.232 , 0.281 0.232 , 0.280	Depositor DCC
$R_{free}$ test set	1888 reflections (2.55%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	106.3	Xtriage
Anisotropy	0.319	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 55.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	29168	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	119.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% 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, 5N0

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 5$	RMSZ	# $ Z  > 5$
1	R	0.34	0/263	1.12	0/409
2	T	0.76	0/584	1.09	1/898 (0.1%)
3	N	0.57	0/331	0.79	0/509
4	A	0.29	0/11020	0.48	0/14907
5	B	0.29	0/9030	0.48	0/12186
6	C	0.30	0/2139	0.49	0/2899
7	E	0.28	0/1767	0.46	0/2378
8	F	0.29	0/696	0.48	0/943
9	H	0.28	0/1082	0.51	0/1466
10	I	0.31	0/970	0.50	0/1308
11	J	0.29	0/541	0.48	0/727
12	K	0.30	0/937	0.48	0/1265
13	L	0.28	0/339	0.56	0/450
All	All	0.31	0/29699	0.52	1/40345 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	T	14	DC	O4'-C1'-N1	5.48	111.84	108.00

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	R	235	0	121	7	0
2	T	525	0	297	32	0
3	N	293	0	156	9	0
4	A	10828	0	10876	386	0
5	B	8859	0	8816	329	0
6	C	2101	0	2056	87	0
7	E	1731	0	1758	50	0
8	F	684	0	692	28	0
9	H	1064	0	1029	38	0
10	I	952	0	897	42	0
11	J	532	0	542	33	0
12	K	919	0	929	30	0
13	L	337	0	352	16	0
14	R	1	0	0	0	0
15	T	99	0	0	5	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
All	All	29168	0	28521	963	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 17.

The worst 5 of 963 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)
4:A:446:ARG:NH2	4:A:480:ALA:HA	1.53	1.23
4:A:446:ARG:HH21	4:A:480:ALA:CA	1.76	0.98
4:A:446:ARG:HH21	4:A:480:ALA:HA	0.79	0.94
11:J:10:CYS:SG	11:J:43:ARG:NH2	2.46	0.89
4:A:1224:LEU:HD11	4:A:1240:CYS:HB3	1.57	0.84

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	A	1370/1733 (79%)	1264 (92%)	106 (8%)	0	100	100
5	B	1103/1224 (90%)	1030 (93%)	73 (7%)	0	100	100
6	C	265/318 (83%)	245 (92%)	20 (8%)	0	100	100
7	E	210/215 (98%)	196 (93%)	14 (7%)	0	100	100
8	F	84/155 (54%)	78 (93%)	6 (7%)	0	100	100
9	H	129/146 (88%)	119 (92%)	10 (8%)	0	100	100
10	I	116/122 (95%)	104 (90%)	12 (10%)	0	100	100
11	J	63/70 (90%)	57 (90%)	6 (10%)	0	100	100
12	K	112/120 (93%)	107 (96%)	5 (4%)	0	100	100
13	L	41/70 (59%)	39 (95%)	2 (5%)	0	100	100
All	All	3493/4173 (84%)	3239 (93%)	254 (7%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	1194/1520 (79%)	1137 (95%)	57 (5%)	25	56
5	B	955/1061 (90%)	919 (96%)	36 (4%)	33	61
6	C	235/274 (86%)	225 (96%)	10 (4%)	29	58
7	E	193/197 (98%)	188 (97%)	5 (3%)	46	69

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
8	F	73/137 (53%)	69 (94%)	4 (6%)	21	53
9	H	116/128 (91%)	109 (94%)	7 (6%)	19	50
10	I	110/116 (95%)	102 (93%)	8 (7%)	14	44
11	J	60/65 (92%)	58 (97%)	2 (3%)	38	64
12	K	99/102 (97%)	96 (97%)	3 (3%)	41	66
13	L	37/57 (65%)	32 (86%)	5 (14%)	4	21
All	All	3072/3657 (84%)	2935 (96%)	137 (4%)	27	57

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	H	102	TYR
10	I	4	PHE
12	K	81	TYR
4	A	1218	GLN
4	A	1215	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	1117	GLN
6	C	264	GLN
9	H	131	ASN
7	E	99	HIS
5	B	1177	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	10/11 (90%)	2 (20%)	1 (10%)

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	10	C
1	R	11	U

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	R	9	G

## 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 10 ligands modelled in this entry, 9 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
15	5N0	T	101	-	91,107,107	2.46	31 (34%)	91,153,153	1.58	10 (10%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
15	5N0	T	101	-	-	6/47/92/92	0/9/9/9

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	101	5N0	C59-C63	7.37	1.54	1.47
15	T	101	5N0	C49-N23	6.38	1.47	1.33
15	T	101	5N0	C56-N25	6.27	1.47	1.33
15	T	101	5N0	C22-N10	6.08	1.47	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	101	5N0	C26-N14	5.33	1.47	1.35

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	T	101	5N0	C4-C5-N3	6.97	121.26	113.69
15	T	101	5N0	C6-C10-N6	5.45	119.60	113.69
15	T	101	5N0	C29-N14-C26	-5.18	120.21	127.55
15	T	101	5N0	C24-C25-C26	-4.84	100.38	110.85
15	T	101	5N0	C19-N9-C17	2.55	111.36	108.65

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

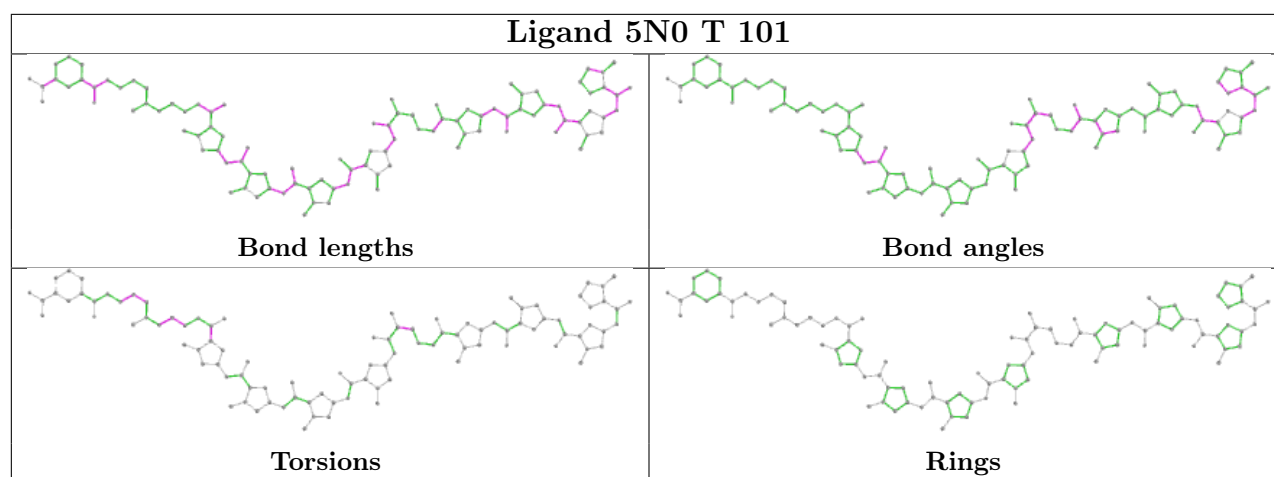
Mol	Chain	Res	Type	Atoms
15	T	101	5N0	C23-C24-C25-C26
15	T	101	5N0	C47-C44-C49-O9
15	T	101	5N0	C50-C51-C52-N24
15	T	101	5N0	C53-C54-C55-N25
15	T	101	5N0	N24-C53-C54-C55

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	T	101	5N0	5	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	11/11 (100%)	0.04	0 100 100	116, 132, 191, 198	0
2	T	26/30 (86%)	-0.14	0 100 100	118, 225, 252, 258	0
3	N	14/20 (70%)	0.01	0 100 100	229, 244, 260, 264	0
4	A	1384/1733 (79%)	-0.08	29 (2%) 63 52	52, 113, 188, 236	0
5	B	1123/1224 (91%)	-0.13	12 (1%) 80 71	55, 99, 153, 196	0
6	C	267/318 (83%)	-0.30	0 100 100	67, 99, 136, 165	0
7	E	212/215 (98%)	-0.06	6 (2%) 53 40	96, 154, 209, 231	0
8	F	86/155 (55%)	-0.36	1 (1%) 79 69	82, 114, 147, 186	0
9	H	133/146 (91%)	0.17	3 (2%) 60 48	106, 137, 170, 202	0
10	I	118/122 (96%)	-0.25	0 100 100	83, 117, 150, 162	0
11	J	65/70 (92%)	-0.32	0 100 100	73, 93, 151, 161	0
12	K	114/120 (95%)	-0.18	0 100 100	78, 105, 131, 153	0
13	L	43/70 (61%)	0.10	2 (4%) 31 23	92, 158, 222, 258	0
All	All	3596/4234 (84%)	-0.12	53 (1%) 73 63	52, 111, 185, 264	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	144	THR	5.3
7	E	83	CYS	4.9
5	B	106	ASP	3.7
4	A	69	THR	3.7
7	E	110	PHE	3.4

### 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](#)

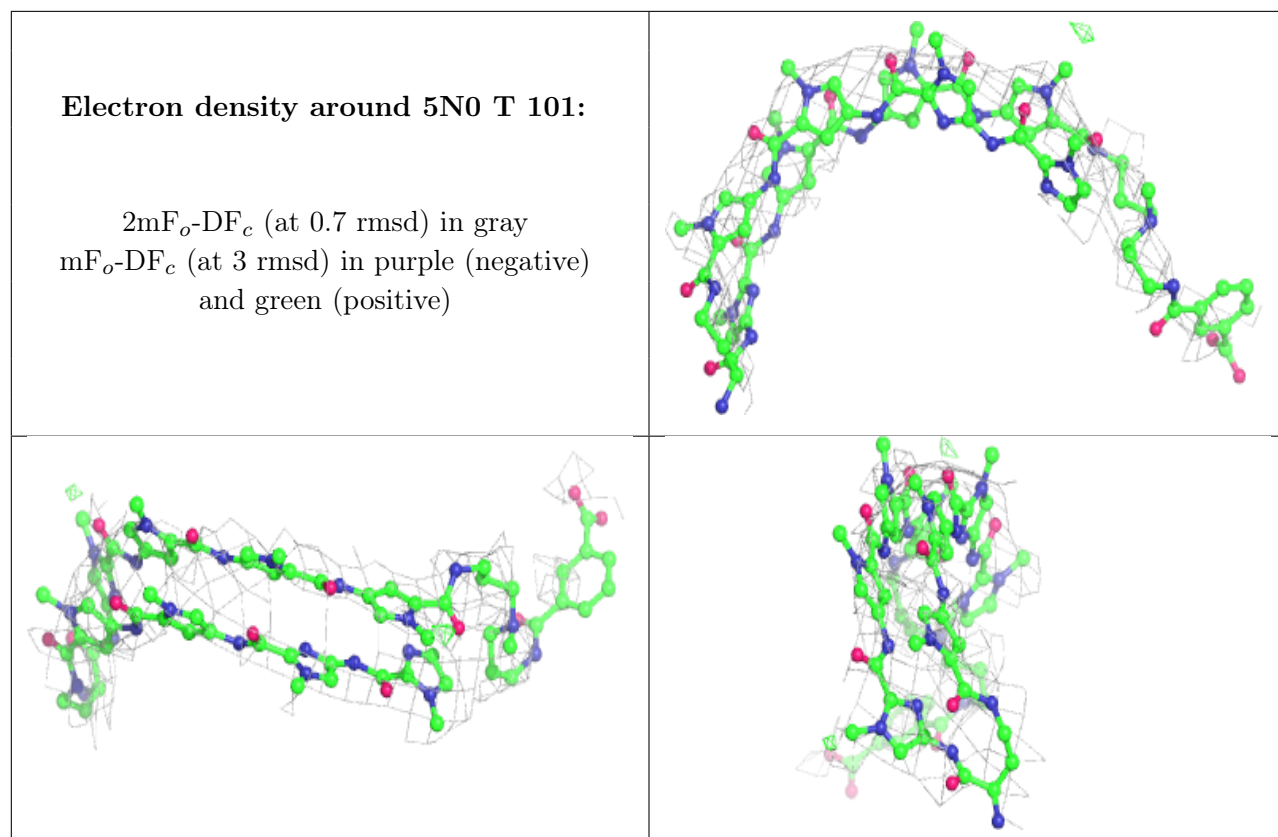
There are no monosaccharides in this entry.

### 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( $\text{\AA}^2$ )	Q<0.9
15	5N0	T	101	99/99	0.82	0.36	198,236,262,264	0
16	ZN	A	1802	1/1	0.88	0.10	167,167,167,167	0
16	ZN	B	1301	1/1	0.90	0.08	195,195,195,195	0
16	ZN	A	1801	1/1	0.91	0.17	242,242,242,242	0
14	MG	R	2001	1/1	0.92	0.35	141,141,141,141	0
16	ZN	L	101	1/1	0.92	0.06	182,182,182,182	0
16	ZN	J	101	1/1	0.96	0.22	90,90,90,90	0
16	ZN	C	401	1/1	0.98	0.10	108,108,108,108	0
16	ZN	I	201	1/1	0.99	0.12	99,99,99,99	0
16	ZN	I	202	1/1	1.00	0.12	110,110,110,110	0

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.