



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 12, 2026 – 05:48 PM UTC

PDB ID : 1TWG / pdb_00001twg
Title : RNA polymerase II complexed with CTP
Authors : Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.
Deposited on : 2004-06-30
Resolution : 3.30 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

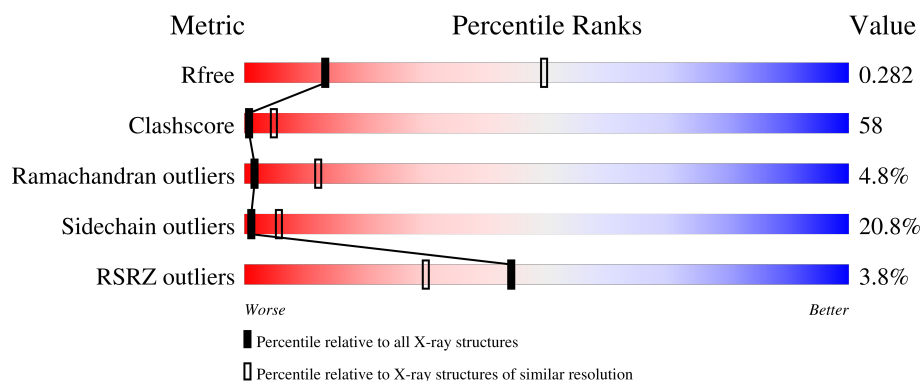
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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}	180053	1169 (3.32-3.28)
Clashscore	190562	1209 (3.32-3.28)
Ramachandran outliers	187476	1188 (3.32-3.28)
Sidechain outliers	187428	1187 (3.32-3.28)
RSRZ outliers	180081	1169 (3.32-3.28)

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	A	1733	3% 34% 30% 12% 22%
2	B	1224	3% 35% 37% 15% 11%
3	C	318	28% 38% 16% 16%
4	E	215	6% 32% 44% 22%
5	F	155	% 24% 21% 8% 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	H	146	
7	I	122	
8	J	70	
9	K	120	
10	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
12	ZN	J	3001	-	-	X	-
12	ZN	L	3005	-	-	X	-
13	CTP	B	3008	X	-	X	-

2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 27731 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1349	Total	C	N	O	S	0	0	0
			10606	6692	1839	2017	58			

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1091	Total	C	N	O	S	0	0	0
			8690	5511	1516	1610	53			

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	215	Total	C	N	O	S	0	0	0
			1760	1116	310	322	12			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	83	Total	C	N	O	S	0	0	0
			670	428	114	125	3			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	121	Total	C	N	O	S	0	0	0
			990	610	181	188	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	J	64	Total	C	N	O	S	0	0	0
			525	334	92	93	6			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

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

- Molecule 11 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	2	Total	Mn	0	0
			2	2		

- Molecule 12 is ZINC ION (CCD ID: ZN) (formula: Zn).

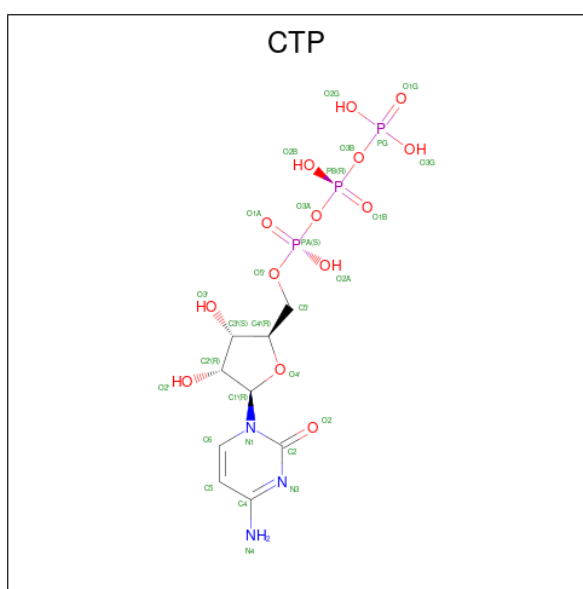
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	2	Total	Zn	0	0
			2	2		
12	B	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	C	1	Total	Zn	0	0
			1	1		
12	I	2	Total	Zn	0	0
			2	2		
12	J	1	Total	Zn	0	0
			1	1		
12	L	1	Total	Zn	0	0
			1	1		

- Molecule 13 is CYTIDINE-5'-TRIPHOSPHATE (CCD ID: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
13	B	1	Total	C	N	O	P	0	0
			29	9	3	14	3		

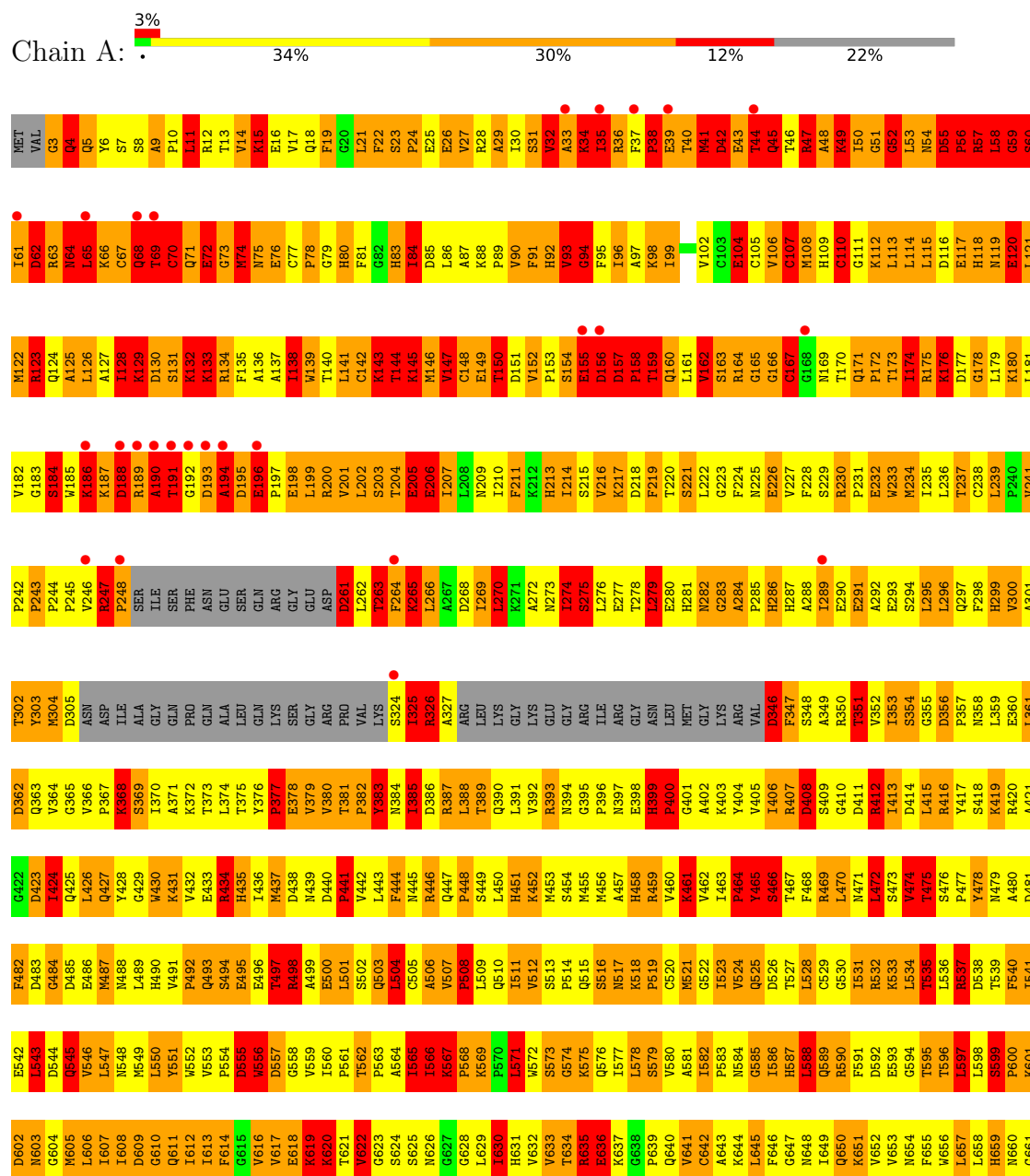
- Molecule 14 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total	O	0	0
			2	2		
14	B	3	Total	O	0	0
			3	3		

3 Residue-property plots

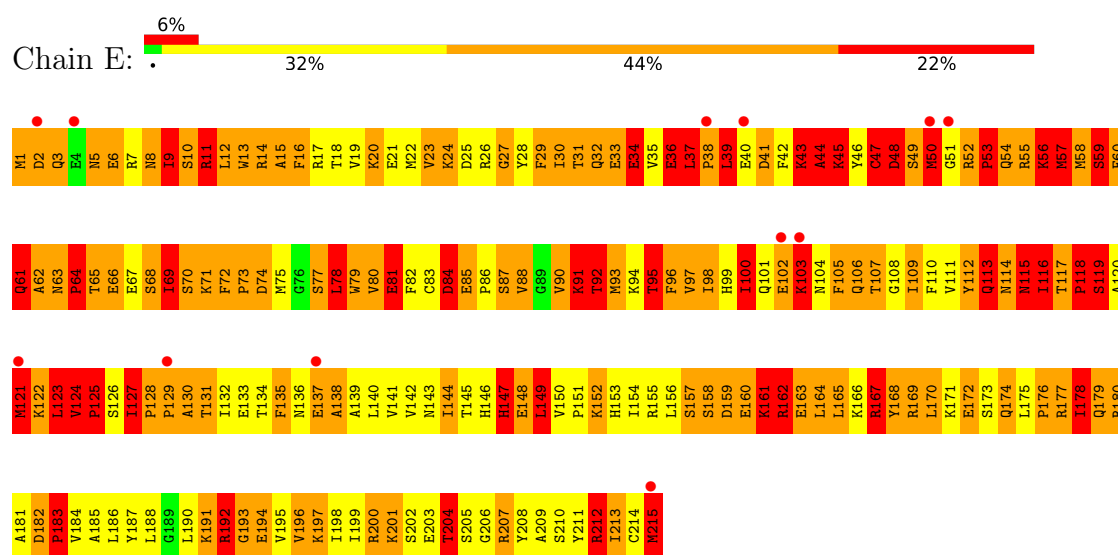
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

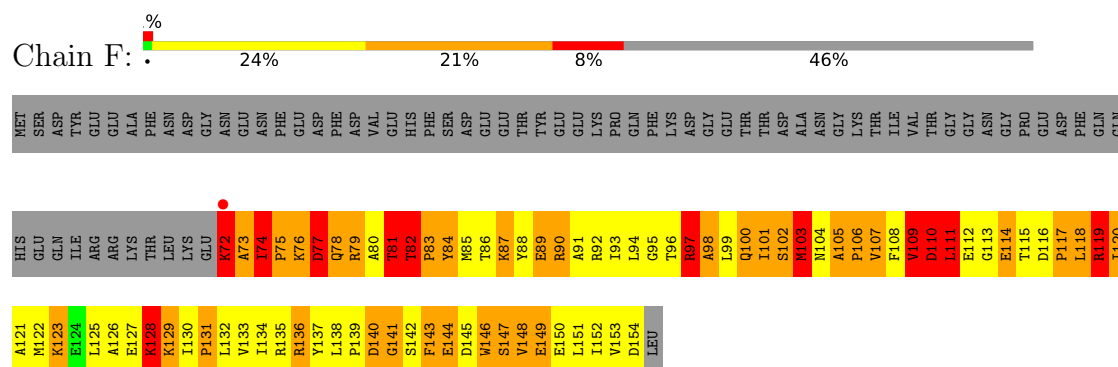


ASP	D1442	T1382	I1322	K1262	M1202	T1142	ASN	L1022	R962	L902	V842	R782	L722	P662
ALA	V1443	S1383	D1323	I1263	N1203	L1143	THR	R1023	I963	R903	R643	T783	N723	S663
MET	M1444	V1384	P1324	E1264	D1204	K1144	PHE	S1024	I964	T904	A844	L784	E724	T664
ALA	T1445	T1385	P1325	N1265	K1205	H1145	HIS	R1025	Q965	D905	R1025	P785	A725	G665
GLY	D1446	R1326	R1326	T1266	D1206	V1146	PHE	L1026	R966	H906	E846	H786	R726	I666
GLY	E1447	I1327	I1327	M1267	L1207	T1147	ALA	A1027	A967	T907	D847	F787	D727	G667
PHE	E1448	X1328	X1328	L1268	T1208	I1148	GLY	T1028	Q968	L908	T848	S788	K728	D668
THR	S1449	E1269	T1329	E1269	M1209	A1149	VAL	R1029	Q969	D909	R849	K789	A729	T669
ALA	T1450	N1270	M1330	E1270	G1210	S1150	ALA	R1030	T970	P910	H850	D790	G730	T670
TYR	VAL	ASN	S1331	I1271	Q1211	E1151	SER	L1031	F971	S911	H851	D791	R731	A671
GLY	LYS	SER	F1332	T1272	V1212	I1152	THR	L1032	H972	L912	R852	F792	L732	D672
GLY	TVR	ASN	I1333	L1273	G1213	Y1153	K1093	Q1033	I973	L913	D853	S793	A733	G673
ALA	MET	THR	D1334	R1274	E1214	Y1154	V1094	E1034	D974	E914	H854	F794	E734	P674
ALA	PRO	GLY	I1335	G1275	R1215	D1155	T1095	Y1035	H975	S915	T855	E795	V735	T675
ASP	M1336	E1276	M1336	V1277	I1216	P1156	S1096	Y1036	T976	G916	R856	S796	N736	M676
GLY	E1337	E1277	E1337	D1278	K1217	I1157	G1097	L1037	K977	S917	R857	K797	L737	R677
GLY	L1338	N1278	V1338	N1278	Q1218	P1158	V1098	L1038	P978	E918	R858	G798	K738	E678
ALA	L1339	I1279	L1339	I1279	T1219	R1159	P1099	K1039	S979	I919	S859	F799	D739	I679
THR	G1340	E1280	G1340	E1280	F1220	S1160	R1100	Q1040	D980	L920	L860	V800	L740	T680
PRO	I1341	R1281	I1341	R1281	K1221	T1161	L1101	A1041	L981	G921	G861	E801	R741	S681
SER	A1342	V1282	A1342	V1282	N1222	V1162	K1102	F1042	I982	D922	R862	H802	N742	T682
PHE	A1343	V1283	A1343	V1283	D1223	I1163	E1103	D1043	I983	L923	V863	S803	F743	T683
GLY	G1344	M1284	G1344	M1284	L1224	P1164	T1104	W1044	K984	K924	R864	H804	K744	A684
ALA	R1345	M1285	R1345	M1285	F1225	E1165	L1105	V1045	D985	L925	Q865	L805	Q745	E685
ALA	A1346	K1286	A1346	K1286	V1226	D1166	N1106	L1046	I986	D926	F866	H806	M746	A686
TYR	A1347	Y1287	A1347	Y1287	I1227	E1167	V1107	S1047	V987	V927	T867	G807	V747	K687
GLY	I1348	D1288	I1348	D1288	W1228	E1168	A1108	L1048	L988	L928	R868	L808	M748	K688
ALA	L1349	R1289	L1349	R1289	S1229	I1169	K1109	I1049	G989	L929	C869	T809	A749	K689
PRO	K1350	K1290	K1350	K1290	E1230	I1170	M1110	E1050	V990	D930	E870	G810	G750	V690
THR	E1351	V1291	E1351	V1291	D1231	Q1171	M1111	A1051	K991	E931	D871	Q811	S751	L691
SER	P1352	P1292	P1352	P1292	M1232	L1172	K1112	O1052	D992	E932	G872	E812	K752	D692
PRO	G1413	S1293	G1413	S1293	D1233	H1173	T1113	F1053	L993	V933	K873	F813	G753	V693
GLY	M1354	P1294	M1354	P1294	E1234	PHE	P1114	L1054	Q994	K934	D874	F814	S754	T694
PHE	S1415	T1295	V1355	T1295	K1235	SER	S1115	R1055	E995	Q935	A875	H815	F755	K695
GLY	A1416	G1296	I1356	G1296	L1236	L1176	T1117	V1057	N996	L936	R876	H816	T756	E696
VAL	E1417	E1297	A1357	E1297	I1237	LEU	T1118	V1058	L997	V937	H877	A817	N757	A697
SER	L1418	Y1298	S1358	Y1298	I1238	ASP	Y1119	V1059	L998	K938	T878	M818	T758	G698
SER	D1359	K1300	D1359	K1300	R1239	GLU	L1120	H1059	V999	D939	E879	G819	A759	A699
PRO	G1360	E1301	G1360	E1301	C1240	GLU	E1121	P1060	L1000	R940	K890	G820	Q760	N700
PHE	A1421	P1302	V1362	P1302	R1241	ALA	P1122	G1061	R1001	K941	Q881	E821	M761	L701
SER	V1422	E1303	V1363	E1303	V1242	GLU	P1122	E1062	G1002	F942	S882	E822	S762	T702
PRO	G1423	G1297	G1423	G1297	V1243	GLN	G1123	M1063	K1003	L943	L883	C823	A763	T703
LEU	V1424	M1304	M1364	V1304	ARG	SER	H1124	V1064	N1004	R944	D884	L824	G764	A704
THR	S1425	Y1305	Y1365	Y1305	PRO	PHE	A1125	G1065	E1005	E945	T885	L825	V765	K705
SER	E1426	L1306	R1366	L1306	LYS	ASP	A1126	V1066	I1006	V946	I886	D826	G766	R706
PRO	L1427	H1307	H1367	H1307	SER	R1187	D1127	L1067	I1007	F947	G887	T827	Q767	G707
THR	V1428	T1308	M1368	T1308	LEU	Q1188	Q1128	A1068	Q1008	V948	C888	A828	Q768	N708
TYR	A1369	D1309	A1369	D1309	ASP	S1189	E1129	A1069	N1009	D949	S889	V829	S769	T709
LEU	L1370	G1310	L1370	G1310	ALA	P1190	Q1130	Q1070	A1010	G950	D890	K830	V770	L710
PRO	MET	V1311	L1371	V1311	GLU	W1191	A1131	S1071	Q1011	E951	A891	T831	E771	R711
THR	PHE	M1312	V1372	M1312	THR	L1192	K1132	I1072	R1012	A952	A892	A832	G772	E712
SER	D1373	L1313	D1373	L1313	GLU	L1193	L1133	G1073	D1013	R953	F893	E833	G773	S713
PRO	A1434	S1314	V1374	S1314	R1194	R1194	L1134	E1074	A1014	W954	E894	T834	R774	F714
ALA	P1435	E1255	M1375	E1255	L1195	L1195	R1135	P1075	V1015	P955	K895	G835	T775	E715
TYR	I1436	V1316	T1376	V1316	E1256	E1196	S1136	A1076	T1016	L956	R896	Y836	A776	D716
SER	T1377	M1317	T1377	M1317	L1257	L1197	A1137	L1077	L1017	P957	X897	T837	F777	M717
PRO	Q1378	H1318	Q1378	H1318	H1258	D1198	T1138	Q1078	F1018	V958	R898	Q838	G778	V718
THR	G1379	V1319	G1379	V1319	H1259	R1199	E1139	M1079	C1019	R959	W899	R839	F779	V719
SER	G1380	L1260	G1380	L1260	A1200	I1200	H140	T1080	T1080	P960	R900	R840	V780	R720
PRO	ASN	F1441	L1381	G1321	K1261	A1201	T141	L1081	L1021	R961	L901	L841	D781	F721

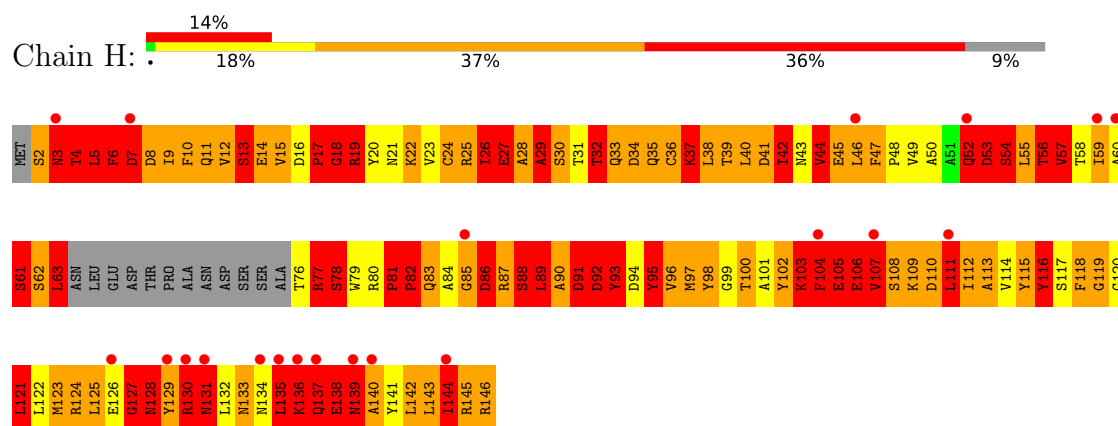




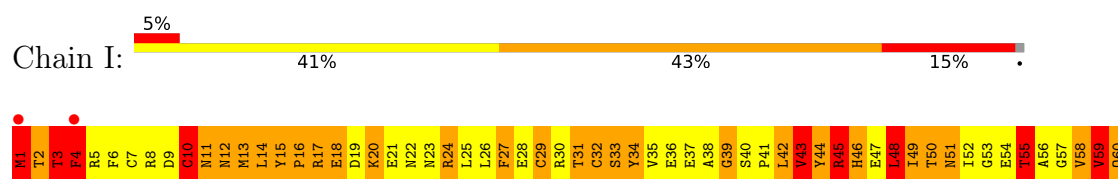
- Molecule 5: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

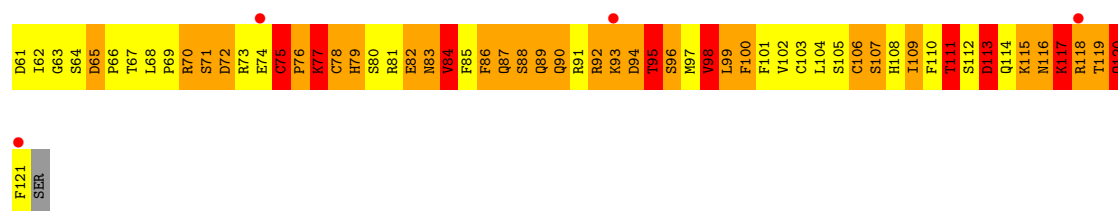


- Molecule 6: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

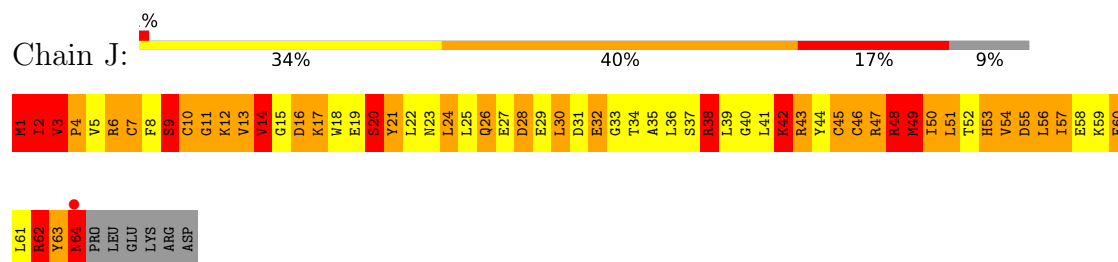


- Molecule 7: DNA-directed RNA polymerase II 14.2 kDa polypeptide

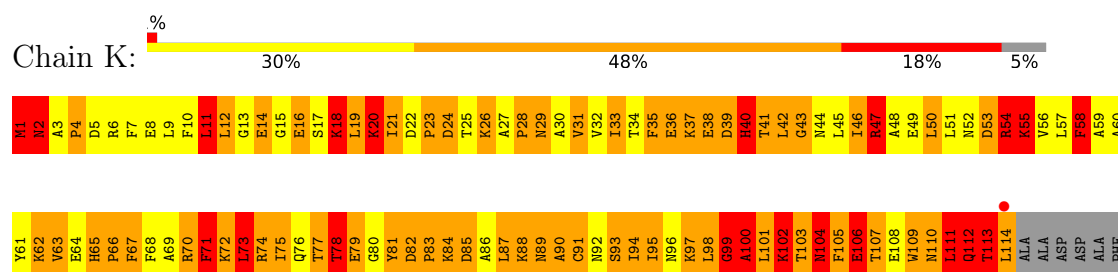




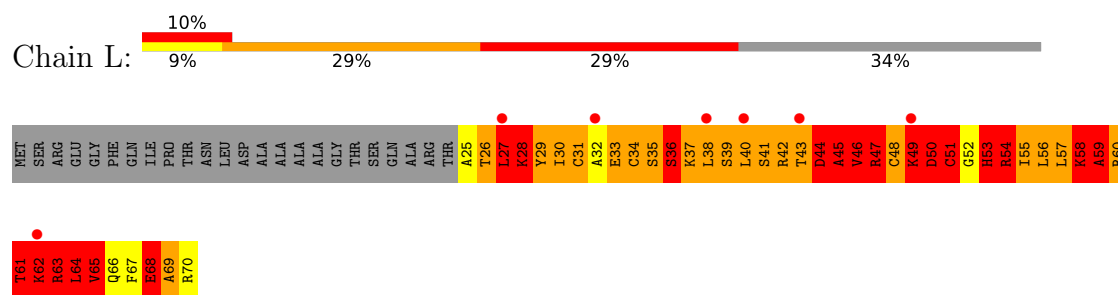
- Molecule 8: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide



- Molecule 9: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 10: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	123.00Å 223.00Å 374.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.30 40.00 – 3.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.30) 93.1 (40.00-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 3.32Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.263 , 0.281 0.227 , 0.282	Depositor DCC
R_{free} test set	2324 reflections (2.99%)	wwPDB-VP
Wilson B-factor (Å ²)	53.7	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 57.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	27731	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.88% 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

5.1 Standard geometry

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	5.56	3559/10792 (33.0%)	3.44	1967/14601 (13.5%)
2	B	5.59	2907/8860 (32.8%)	3.42	1557/11945 (13.0%)
3	C	5.64	694/2133 (32.5%)	3.50	390/2891 (13.5%)
4	E	5.54	563/1796 (31.3%)	3.40	321/2416 (13.3%)
5	F	5.17	200/682 (29.3%)	3.37	125/922 (13.6%)
6	H	5.49	336/1086 (30.9%)	3.28	185/1470 (12.6%)
7	I	6.06	372/1009 (36.9%)	3.74	221/1357 (16.3%)
8	J	5.25	170/533 (31.9%)	3.62	109/715 (15.2%)
9	K	5.52	327/937 (34.9%)	3.60	195/1265 (15.4%)
10	L	6.08	142/366 (38.8%)	3.58	72/485 (14.8%)
All	All	5.58	9270/28194 (32.9%)	3.45	5142/38067 (13.5%)

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
1	A	0	69
2	B	0	59
3	C	0	19
4	E	0	9
6	H	0	15
7	I	0	9
8	J	0	2
9	K	0	5
10	L	0	4
All	All	0	191

The worst 5 of 9270 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	57	MET	SD-CE	43.75	2.88	1.79
2	B	101	MET	SD-CE	33.55	2.63	1.79
2	B	870	ILE	CA-CB	32.73	2.00	1.53
4	E	162	ARG	NE-CZ	31.22	1.67	1.33
2	B	552	MET	SD-CE	31.13	2.57	1.79

The worst 5 of 5142 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	61	ILE	N-CA-C	-21.85	86.24	110.05
1	A	774	ARG	NE-CZ-NH1	21.33	142.83	121.50
1	A	466	SER	N-CA-C	19.18	134.25	112.57
3	C	240	VAL	N-CA-C	19.18	128.45	110.42
2	B	995	ARG	NE-CZ-NH2	-18.87	102.22	119.20

There are no chirality outliers.

5 of 191 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	44	THR	Peptide
1	A	52	GLY	Peptide
1	A	60	SER	Peptide
1	A	65	LEU	Peptide
1	A	74	MET	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	A	10606	0	10667	1137	0
2	B	8690	0	8714	1044	2
3	C	2095	0	2053	276	0
4	E	1760	0	1788	208	1
5	F	670	0	690	74	0
6	H	1068	0	1040	226	0
7	I	990	0	949	108	0
8	J	525	0	538	63	0
9	K	919	0	927	131	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	L	364	0	389	107	0
11	A	2	0	0	0	0
12	A	2	0	0	1	0
12	B	1	0	0	0	0
12	C	1	0	0	1	0
12	I	2	0	0	0	0
12	J	1	0	0	2	0
12	L	1	0	0	2	0
13	B	29	0	11	13	0
14	A	2	0	0	0	0
14	B	3	0	0	7	0
All	All	27731	0	27766	3195	2

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 58.

The worst 5 of 3195 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:226:PHE:CD1	2:B:226:PHE:CE1	1.77	1.67
2:B:370:PHE:CD1	2:B:370:PHE:CE1	1.78	1.67
3:C:165:LYS:CG	3:C:165:LYS:CD	1.74	1.65
2:B:18:PHE:CB	2:B:18:PHE:CG	1.79	1.65
4:E:79:TRP:CB	4:E:79:TRP:CG	1.78	1.65

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:434:ARG:O	2:B:465:ASN:ND2[2_655]	2.11	0.09
2:B:1223:ASP:OD1	4:E:1:MET:CE[8_455]	2.17	0.03

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	
1	A	1332/1733 (77%)	1160 (87%)	113 (8%)	59 (4%)	2	13
2	B	1071/1224 (88%)	897 (84%)	131 (12%)	43 (4%)	2	15
3	C	264/318 (83%)	218 (83%)	36 (14%)	10 (4%)	2	16
4	E	213/215 (99%)	180 (84%)	20 (9%)	13 (6%)	1	9
5	F	81/155 (52%)	71 (88%)	8 (10%)	2 (2%)	4	23
6	H	129/146 (88%)	85 (66%)	21 (16%)	23 (18%)	0	0
7	I	119/122 (98%)	108 (91%)	8 (7%)	3 (2%)	4	23
8	J	62/70 (89%)	58 (94%)	4 (6%)	0	100	100
9	K	112/120 (93%)	98 (88%)	10 (9%)	4 (4%)	2	17
10	L	44/70 (63%)	26 (59%)	11 (25%)	7 (16%)	0	1
All	All	3427/4173 (82%)	2901 (85%)	362 (11%)	164 (5%)	2	12

5 of 164 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ILE
1	A	43	GLU
1	A	45	GLN
1	A	47	ARG
1	A	56	PRO

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	
1	A	1181/1520 (78%)	963 (82%)	218 (18%)	1	7
2	B	947/1061 (89%)	777 (82%)	170 (18%)	2	8
3	C	234/274 (85%)	184 (79%)	50 (21%)	1	5
4	E	197/197 (100%)	142 (72%)	55 (28%)	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	F	73/137 (53%)	61 (84%)	12 (16%)	2	11
6	H	117/128 (91%)	67 (57%)	50 (43%)	0	0
7	I	115/116 (99%)	92 (80%)	23 (20%)	1	6
8	J	59/65 (91%)	45 (76%)	14 (24%)	1	4
9	K	99/102 (97%)	77 (78%)	22 (22%)	1	5
10	L	40/57 (70%)	17 (42%)	23 (58%)	0	0
All	All	3062/3657 (84%)	2425 (79%)	637 (21%)	1	6

5 of 637 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	92	THR
7	I	98	VAL
4	E	144	ILE
4	E	91	LYS
6	H	42	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 99 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1062	HIS
3	C	203	GLN
2	B	1084	GLN
3	C	73	GLN
4	E	3	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 11 ligands modelled in this entry, 10 are monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
13	CTP	B	3008	11	29,30,30	2.74	6 (20%)	43,47,47	3.00	10 (23%)

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
13	CTP	B	3008	11	1/1/7/7	5/22/38/38	0/2/2/2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	B	3008	CTP	C2-N1	9.94	1.60	1.40
13	B	3008	CTP	C1'-N1	7.57	1.69	1.47
13	B	3008	CTP	C6-N1	4.64	1.49	1.38
13	B	3008	CTP	C6-C5	4.21	1.44	1.35
13	B	3008	CTP	C4-N3	2.21	1.39	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	B	3008	CTP	O4'-C1'-N1	11.39	134.17	108.36
13	B	3008	CTP	C1'-N1-C2	8.54	137.30	118.44
13	B	3008	CTP	C6-N1-C2	-7.63	107.58	120.46
13	B	3008	CTP	O2-C2-N3	-7.40	110.66	122.33
13	B	3008	CTP	O2-C2-N1	3.23	125.23	118.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
13	B	3008	CTP	C1'

All (5) torsion outliers are listed below:

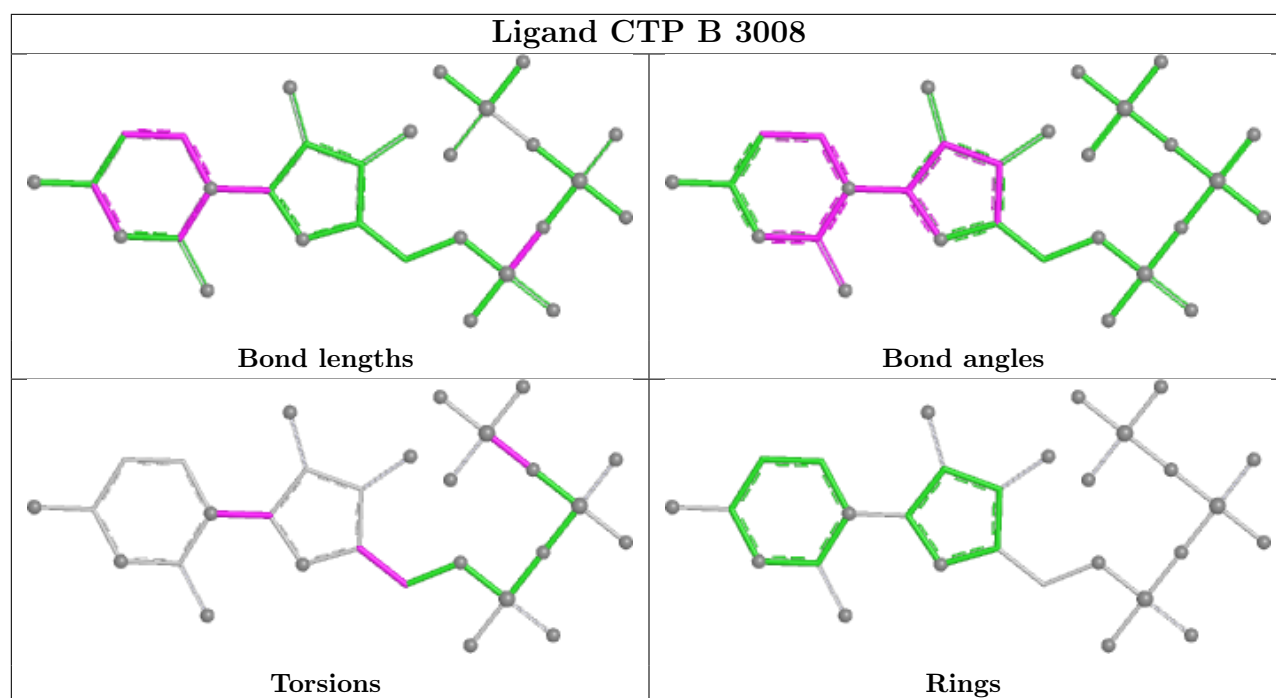
Mol	Chain	Res	Type	Atoms
13	B	3008	CTP	O4'-C4'-C5'-O5'
13	B	3008	CTP	C3'-C4'-C5'-O5'
13	B	3008	CTP	PB-O3B-PG-O2G
13	B	3008	CTP	C2'-C1'-N1-C2
13	B	3008	CTP	PB-O3B-PG-O3G

There are no ring outliers.

1 monomer is involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	B	3008	CTP	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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	13
2	B	13
9	K	3
4	E	2
6	H	1
8	J	1
7	I	1

The worst 5 of 34 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1066:VAL	C	1067:LEU	N	1.20
1	A	1106:ASN	C	1107:VAL	N	1.20
1	A	1191:TRP	C	1192:LEU	N	1.20
1	A	1328:TYR	C	1329:THR	N	1.20
1	B	224:GLN	C	225:VAL	N	1.20

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	A	1349/1733 (77%)	0.13	44 (3%) 49 33	1, 24, 101, 195	0
2	B	1091/1224 (89%)	0.15	38 (3%) 47 31	1, 25, 114, 167	0
3	C	266/318 (83%)	0.09	0 100 100	1, 29, 77, 167	0
4	E	215/215 (100%)	0.51	12 (5%) 30 20	1, 38, 111, 135	0
5	F	83/155 (53%)	-0.16	1 (1%) 76 58	1, 22, 64, 86	0
6	H	133/146 (91%)	0.96	21 (15%) 5 4	22, 68, 127, 188	0
7	I	121/122 (99%)	0.37	6 (4%) 34 23	1, 31, 83, 100	0
8	J	64/70 (91%)	-0.03	1 (1%) 70 52	3, 22, 76, 91	0
9	K	114/120 (95%)	0.08	1 (0%) 81 65	3, 36, 73, 99	0
10	L	46/70 (65%)	1.01	7 (15%) 5 5	24, 73, 133, 137	0
All	All	3482/4173 (83%)	0.20	131 (3%) 44 30	1, 28, 108, 195	0

The worst 5 of 131 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	H	104	PHE	5.7
1	A	248	PRO	5.1
2	B	735	ALA	4.5
1	A	975	HIS	4.4
2	B	866	TYR	4.2

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](#)

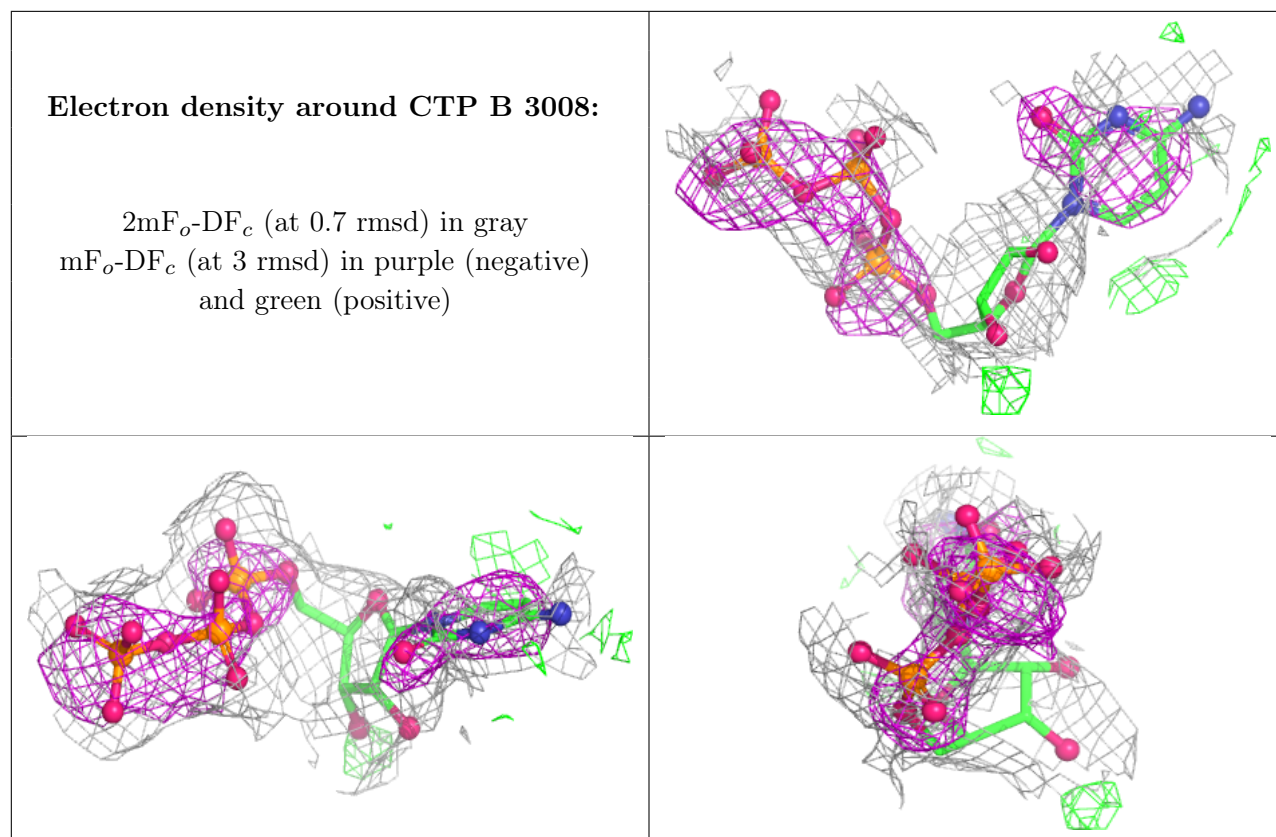
There are no oligosaccharides in this entry.

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($\text{\AA}^2$)	Q<0.9
13	CTP	B	3008	29/29	0.80	0.15	26,36,73,77	0
12	ZN	I	3004	1/1	0.88	0.20	49,49,49,49	0
12	ZN	L	3005	1/1	0.93	0.12	84,84,84,84	0
12	ZN	C	3002	1/1	0.97	0.12	35,35,35,35	0
11	MN	A	3010	1/1	0.97	0.05	29,29,29,29	0
12	ZN	A	3006	1/1	0.97	0.10	48,48,48,48	0
12	ZN	A	3008	1/1	0.97	0.06	87,87,87,87	0
12	ZN	I	3003	1/1	0.98	0.10	45,45,45,45	0
12	ZN	B	3007	1/1	0.99	0.10	37,37,37,37	0
11	MN	A	3009	1/1	0.99	0.04	15,15,15,15	0
12	ZN	J	3001	1/1	1.00	0.14	31,31,31,31	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.