



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 06:34 PM UTC

PDB ID : 3HOY / pdb_00003hoy
Title : Complete RNA polymerase II elongation complex VI
Authors : Sydow, J.F.; Brueckner, F.; Cheung, A.C.M.; Damsma, G.E.; Dengl, S.;
Lehmann, E.; Vassylyev, D.; Cramer, P.
Deposited on : 2009-06-03
Resolution : 3.40 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
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

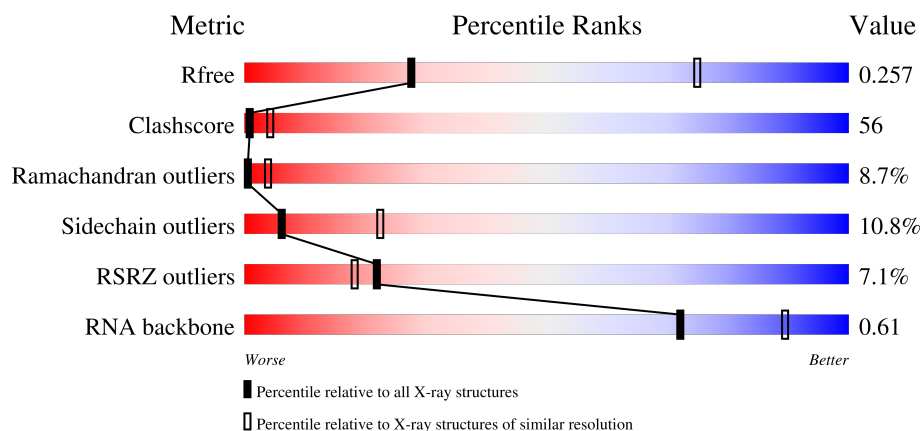
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	1001 (3.44-3.36)
Clashscore	190562	1022 (3.44-3.36)
Ramachandran outliers	187476	1012 (3.44-3.36)
Sidechain outliers	187428	1012 (3.44-3.36)
RSRZ outliers	180081	1001 (3.44-3.36)
RNA backbone	3983	1157 (3.80-3.00)

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 ≥ 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	4% 24% 43% 13% 18%
2	B	1224	7% 25% 50% 14% 10%
3	C	347	% 20% 43% 13% 23%
4	D	221	11% 28% 38% 15% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	215	
6	F	155	
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	T	41	
14	N	41	
15	P	20	

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
13	BRU	T	20	-	-	X	-

2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 31803 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 subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1419	Total	C	N	O	S	0	0	0
			11166	7036	1953	2115	62			

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1105	Total	C	N	O	S	0	0	0
			8786	5564	1541	1627	54			

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-28	MET	-	expression tag	UNP P16370
C	-27	GLY	-	expression tag	UNP P16370
C	-26	SER	-	expression tag	UNP P16370
C	-25	HIS	-	expression tag	UNP P16370
C	-24	HIS	-	expression tag	UNP P16370
C	-23	HIS	-	expression tag	UNP P16370
C	-22	HIS	-	expression tag	UNP P16370
C	-21	HIS	-	expression tag	UNP P16370
C	-20	HIS	-	expression tag	UNP P16370
C	-19	SER	-	expression tag	UNP P16370
C	-18	ASN	-	expression tag	UNP P16370
C	-17	SER	-	expression tag	UNP P16370
C	-16	GLY	-	expression tag	UNP P16370
C	-15	LEU	-	expression tag	UNP P16370

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-14	ASN	-	expression tag	UNP P16370
C	-13	ASP	-	expression tag	UNP P16370
C	-12	ILE	-	expression tag	UNP P16370
C	-11	PHE	-	expression tag	UNP P16370
C	-10	GLU	-	expression tag	UNP P16370
C	-9	ALA	-	expression tag	UNP P16370
C	-8	GLN	-	expression tag	UNP P16370
C	-7	LYS	-	expression tag	UNP P16370
C	-6	ILE	-	expression tag	UNP P16370
C	-5	GLU	-	expression tag	UNP P16370
C	-4	TRP	-	expression tag	UNP P16370
C	-3	HIS	-	expression tag	UNP P16370
C	-2	GLU	-	expression tag	UNP P16370
C	-1	ASP	-	expression tag	UNP P16370
C	0	THR	-	expression tag	UNP P16370
C	1	GLY	-	expression tag	UNP P16370

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	178	Total	C	N	O	S	0	0	0
			1365	845	242	276	2			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	213	Total	C	N	O	S	0	0	0
			1744	1107	308	318	11			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

- Molecule 13 is a DNA chain called 5'-D(*CP*CP*AP*AP*GP*CP*TP*CP*AP*AP*G*TP*AP*CP*TP*TP*AP*CP*GP*CP*CP*(BRU)P*GP*GP*TP*CP*AP*TP*TP*AP*CP*TP*AP*GP*TP*AP*CP*TP*GP*CP*C)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
13	T	19	Total	Br	C	N	O	P	8	0	0
			382	1	184	64	115	18			

- Molecule 14 is a DNA chain called 5'-D(*CP*CP*GP*GP*CP*AP*GP*TP*AP*CP*TP*AP*GP*TP*AP*AP*AP*CP*TP*AP*GP*TP*AP*TP*T*GP*AP*AP*AP*GP*TP*AP*CP*TP*TP*GP*AP*GP*CP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	7	Total	C	N	O	P	11	0	0
			145	70	32	37	6			

- Molecule 15 is a RNA chain called 5'-R(*UP*AP*UP*AP*UP*GP*CP*A*UP*AP*AP*AP*GP*AP*CP*CP*AP*GP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	P	10	Total	C	N	O	P	0	0	0
			213	97	43	64	9			

- Molecule 16 is ZINC ION (CCD ID: 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		
16	L	1	Total	Zn	0	0
			1	1		

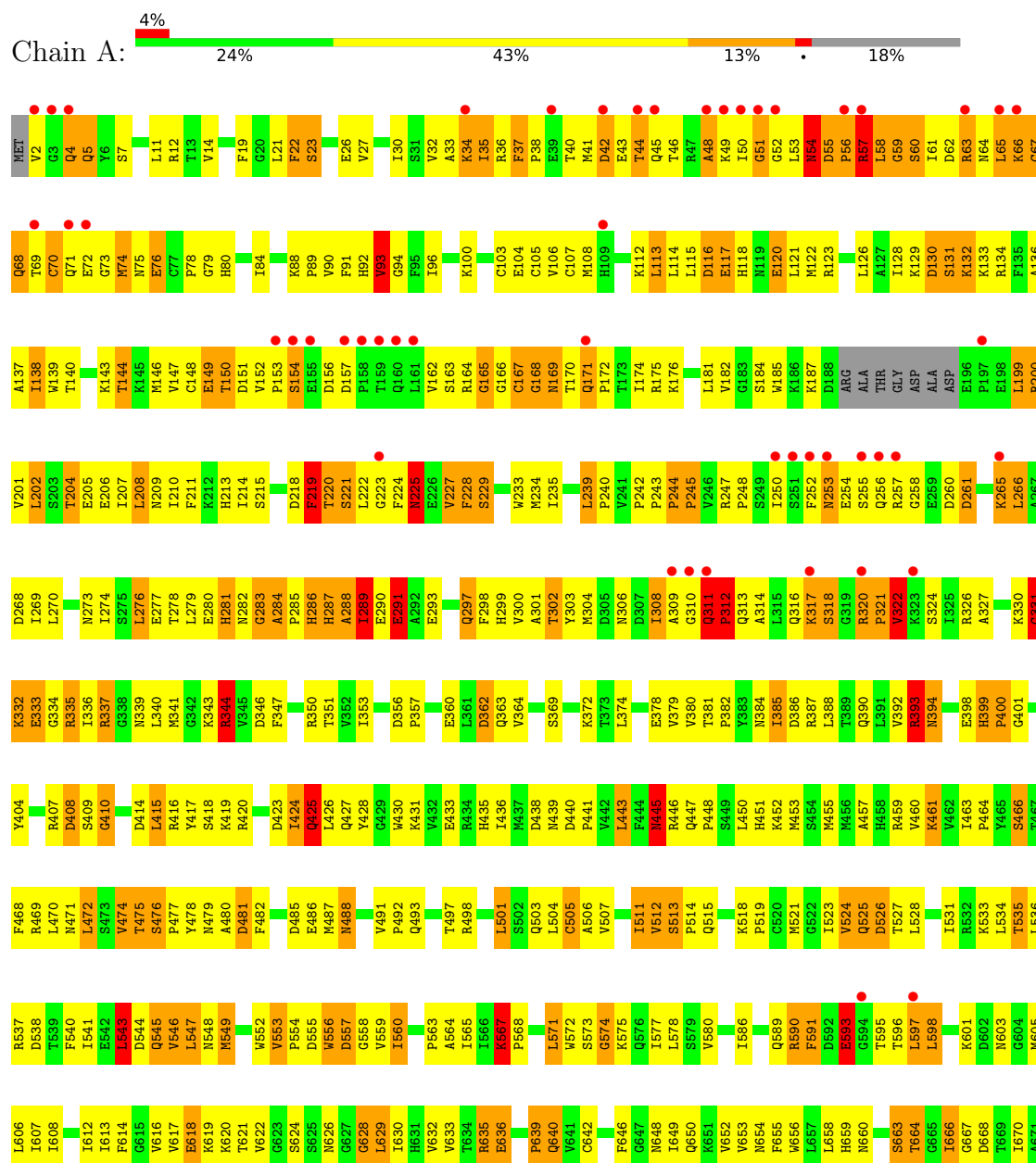
- Molecule 17 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	1	Total	Mg	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.

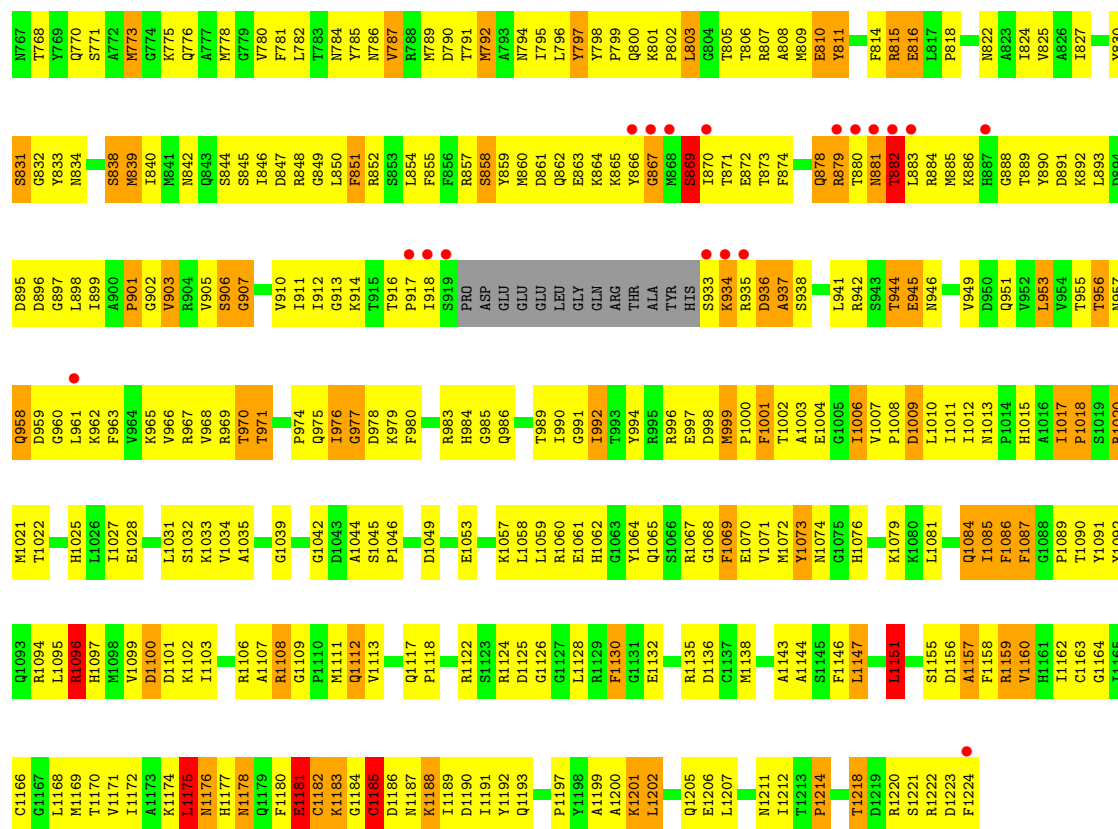
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1



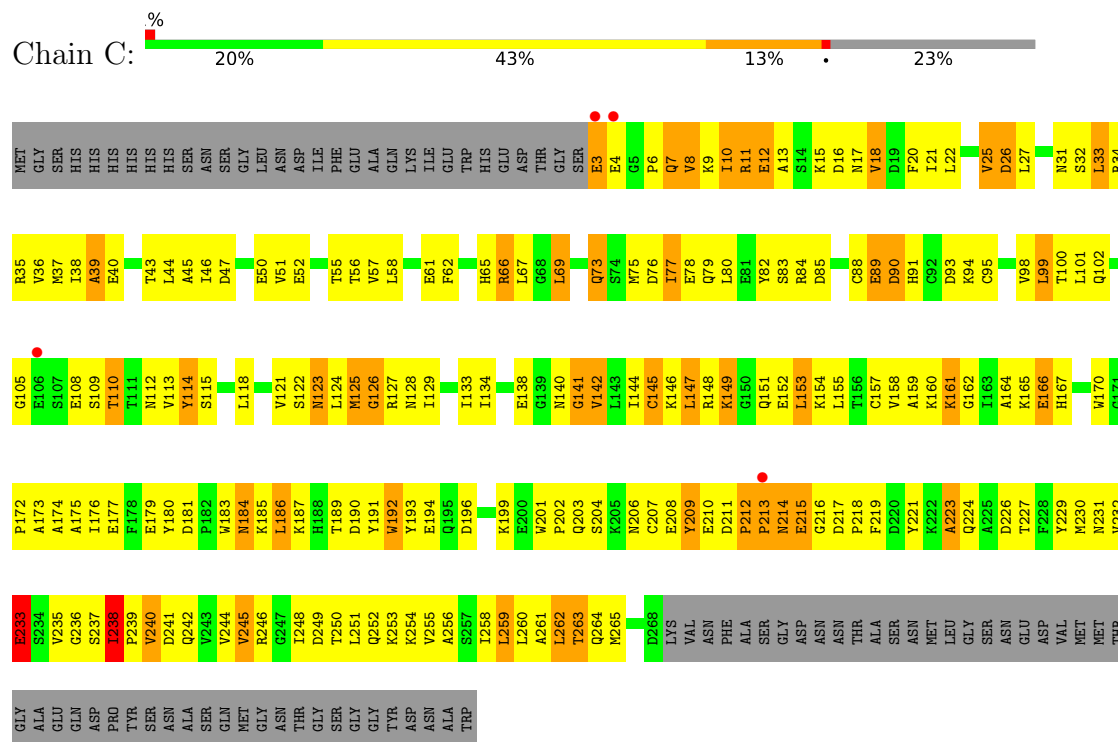
SER	ALA	M1454	M1390	I1322	L1260	R1199	E1139	E1074	I1007	L936	G885	N802	V735	D672
PRO	ASP	P1465	M1393	D1323	K1261	A1200	H1140	P1075	Q1008	L937	F866	S803	N736	G673
SER	TYR	GLU	M1394	P1324	K1262	A1201	T1141	A1076	M1009	V937	F867	S804	N737	P674
TYR	GLY	GLN	T1325	T1326	M1202	M1201	T1142		M1010	K941	F868	L805	K738	T675
SER	GLU	LVS	R1326	R1327	N1203	N1203	L1143	M1079	Q1011	L943	G869	R806	D739	M676
PRO	ALA	ILE	I1327	I1327	N1265	D1204	L1144	T1080			E870	L808	L740	R677
THR	THR	THR	Y1328	Y1328	T1266	K1205	S1145	L1081	A1014		D871	L809	N741	E678
SER	SER	GLU	T1329	T1329	M1267	D1206	V1146	ASN	V1015	V946	G872	T809	N742	I679
PRO	PRO	ILE	N1330	N1330	T1207	T1207	T1147	THR	T1016	F947	M873	P810	F743	
SER	PHE	GLU	S1331	S1331	T1208	T1208	I1148	PHE	L1017	D948	D874	Q811	K744	T682
TYR	GLY	ASP	F1332	F1332	M1209	M1209	A1149	HIS	F1018	V948	A875	E812	Q745	I683
SER	ALA	GLY	I1333	I1333	G1210	G1210	S1150	PHE	C1019	G950	A876	F813	M746	A684
PRO	GLY	GLN	E1404	E1404	G1211	G1211	E1151	ALA	C1020		H877	F814	V747	E685
THR	GLY	ASP	T1405	T1405	V1212	V1212	I1152	GLY	L1021	V954	L878	F815	M748	A686
SER	GLY	GLY	E1337	E1337	G1213	G1213	Y1153	VAL	S1024	P955	E879	H816	A749	K687
PRO	ALA	GLY	V1276	V1276	E1214	E1214	Y1154	ALA	R1025	L956	K880	A817	G750	K688
PRO	THR	VAL	G1277	G1277	R1215	R1215	D1155	S1091	L1032	P957	M818	G751	S751	K689
TYR	PRO	THR	G1339	G1339	R1216	R1216	P1156	K1092	L1033	V958	G888	G819	K752	V690
TYR	SER	THR	I1341	I1341	K1217	K1217	D1157	K1093	E1034		L883	G820	G753	L691
PRO	PRO	TYR	E1342	E1342	Q1218	Q1218	P1158	V1094	Y1035	R961	T885	R821	S754	D692
THR	GLY	SER	R1345	R1345	T1219	T1219	P1159	T1095	R1029	K962	K866	E822	F755	V693
SER	PHE	ASN	V1282	V1282	F1220	F1220	S1096	S1096	V1031	R963	T887	I825	T756	T694
PRO	GLY	GLU	M1283	M1283	K1221	K1221	L1161	G1097	L1032	I964	G888	T826	N757	K695
SER	VAL	SER	M1284	M1284	N1222	N1222	V1162	V1098	Q1033	Q965	S889	G827	I758	E696
TYR	SER	GLY	L1348	L1348	D1223	D1223	I1163	R1100	Y1035	A967	R886	A828	A759	A697
SER	SER	LEU	Y1349	Y1349	F1225	F1225	E1165	L1101	R1036	Q968	H887	R829		Q698
PRO	PRO	VAL	K1350	K1350	V1226	V1226	L1166	K1102	L1037	Q969	R888	R830	A763	A699
THR	GLY	ASN	E1351	E1351	K1227	K1227	E1167	E1103	T1038	K772	T831	T831	C764	N700
SER	PHE	ALA	Y1353	Y1353	E1289	E1289	E1168	L1104	T1038	T970	G766	A832	V765	L701
PRO	SER	ASP	K1290	K1290	V1228	V1228	E1169	L1105	Q1040	F971	D900	F833	G766	
SER	PRO	LEU	V1291	V1291	E1229	E1229	I1170	N1106	E1041	D974	L902	T834	Q767	K705
TYR	THR	ASP	P1292	P1292	D1230	D1230	Q1171	K1109	F1042	H975	N903	G835	Q768	H706
SER	SER	VAL	S1293	S1293	D1231	D1231	H1172	M1111	D1043	T976	T904	R836	S769	G707
PRO	PRO	LVS	T1294	T1294	M1232	M1232	H1173	N1110	P1044	K977	D905	I837	V770	M708
THR	THR	ASP	T1295	T1295	D1233	D1233	F1174	K1112	V1045	P978	H906	Q838	E771	T709
SER	TYR	GLU	G1296	G1296	L1175	L1175	S1175	T1113	L1046	S979	K773	R839	K773	R711
PRO	PRO	LEU	E1297	E1297	L1176	L1176	L1176	T1113	S1047	D980	P910	R840	R774	E712
SER	PRO	MET	Y1298	Y1298	L1236	L1236	LEU	P1114	M1048	L841		L841		E713
TYR	THR	PHE	V1299	V1299	I1238	I1238	ASP	S1115	I1049	T982	L913	V842	F779	S713
SER	SER	SER	K1300	K1300	R1239	R1239	GLU	L1116	E1050	K984	E914	K843	V780	F714
PRO	PRO	PRO	E1301	E1301	C1240	C1240	GLU	T1117	A1051	I983	S915	A844	D781	E715
THR	ALA	LEU	P1302	P1302	R1241	R1241	ALA	Y1118	Q1052	D985	G916	L845	R782	D716
SER	TYR	VAL	E1303	E1303	V1242	V1242	GLU	Y1119	F1053	L986		E846	T783	N717
PRO	PRO	GLU	V1243	V1243	GLN	GLN	GLU	L1120	L1054	V987		D847	L784	V718
SER	PRO	SER	R1244	R1244	PRO	PRO	SER	P1122	R1055	L988		V850	P785	V719
TYR	THR	LVS	LYS	LYS	LYS	LYS	PHE	P1122	S1056	G989	D922	H851	H786	R721
THR	SER	LEU	ASP	ASP	LEU	LEU	ASP	G1123	V1057	V990	L923	Y852	F787	L722
SER	TYR	ALA	ASP	ASP	ALA	ALA	Q1187	H1124	H1058	Q994	K924	D853	S788	N723
PRO	PRO	MET	V1310	V1310	ASP	ASP	Q1188	A1125	H1059		L925	M854	K789	E724
SER	PRO	ALA	N1312	N1312	GLU	GLU	S1189	A1126	P1060	L998	Q926	M854	Y792	A725
TYR	THR	GLY	L1313	L1313	THR	THR	P1190	D1127	F1061	V999	V927	T856	F793	D727
SER	SER	GLY	S1314	S1314	GLU	GLU	W1191	Q1128	E1062	L1000	L928	R857	P794	K728
PRO	PRO	PHE	E1315	E1315	ALA	ALA	L1192	Q1130	V1064	R1001	L929	N858	E795	A729
THR	THR	THR	V1316	V1316	E1254	E1254	R1194		G1002	G1002	D930	S859	S796	G730
SER	TYR	ALA	E1255	E1255	R1256	R1256	L1194	I1134	K1003		E931	L860	K797	R731
PRO	PRO	ALA	E1256	E1256	E1257	E1257	L1195	I1137	N1004		E932	L863	V800	L732
SER	TYR	GLY	T1318	T1318	H1258	H1258	E1196	A1137	I1072	E1005	Y933	V863	A733	E734
PRO	PRO	GLY	V1319	V1319	M1259	M1259	L1197	I1138	G1073	I1006		I864		

Chain B: 7% 25% 50% 14% 10%

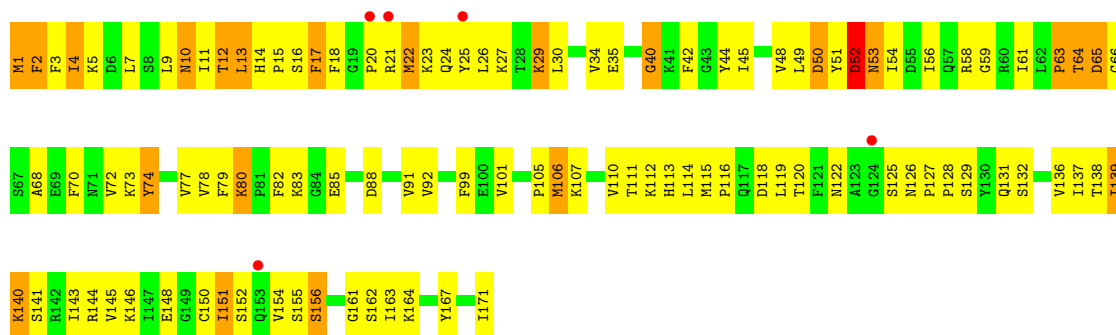




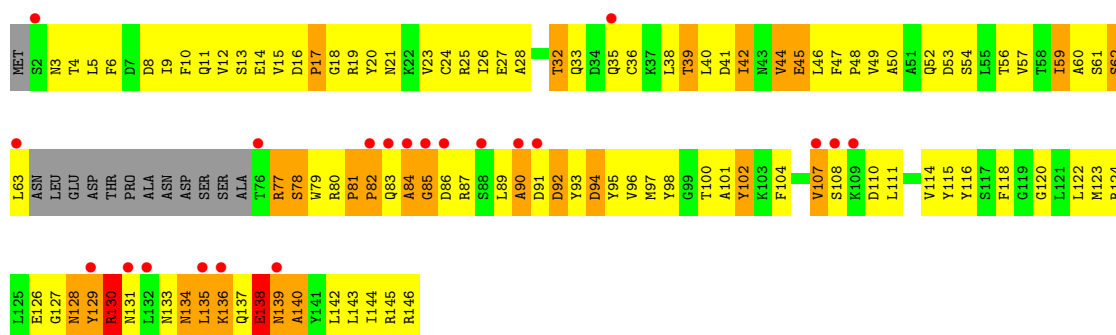
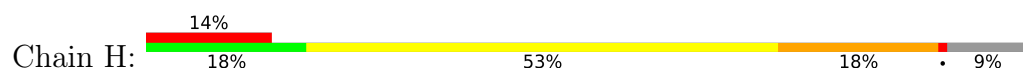
- Molecule 3: DNA-directed RNA polymerase II subunit RPB3



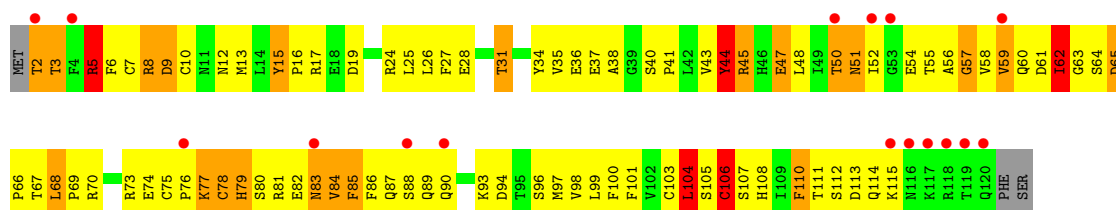
- Molecule 4: DNA-directed RNA polymerase II subunit RPB4



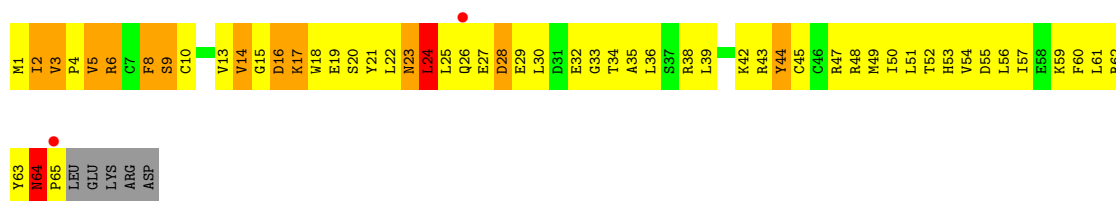
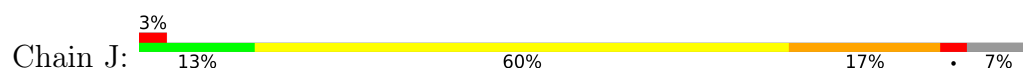
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

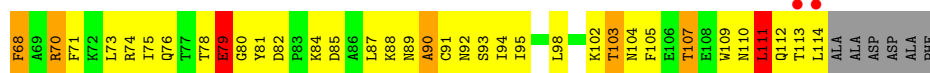


• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

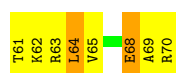
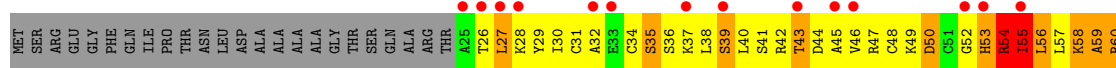


• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

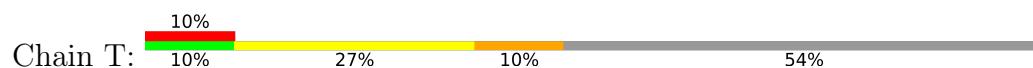




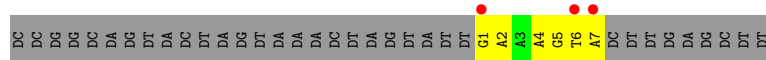
- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



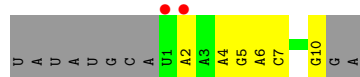
- Molecule 13: 5'-D(*CP*CP*AP*AP*GP*CP*TP*CP*AP*AP*G*TP*AP*CP*TP*TP*AP*C
P*GP*CP*CP*(BRU)P*GP*GP*TP*CP*AP*TP*TP*AP*CP*TP*AP*GP*TP*AP*CP*TP*
GP*CP*C)-3'



- Molecule 14: 5'-D(*CP*CP*GP*GP*CP*AP*GP*TP*AP*CP*TP*AP*GP*TP*AP*AP*AP*
CP*TP*AP*GP*TP*AP*TP*T*GP*AP*AP*AP*GP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T
P*T)-3'



- Molecule 15: 5'-R(*UP*AP*UP*AP*UP*GP*CP*A*UP*AP*AP*AP*GP*AP*CP*CP*AP*G
P*GP*A)-3'



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	222.14Å 392.69Å 282.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.40 50.00 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-3.40) 99.9 (50.00-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.88 (at 3.40Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.216 , 0.254 0.216 , 0.257	Depositor DCC
R_{free} test set	3325 reflections (1.98%)	wwPDB-VP
Wilson B-factor (Å ²)	89.1	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 88.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.013 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.018 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	31803	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.04% 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, MG, BRU

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z > 5$	RMSZ	# $ Z > 5$
1	A	0.65	1/11365 (0.0%)	1.17	111/15367 (0.7%)
2	B	0.62	2/8957 (0.0%)	1.11	66/12078 (0.5%)
3	C	0.67	0/2133	1.18	13/2891 (0.4%)
4	D	0.60	0/1374	1.14	10/1849 (0.5%)
5	E	0.54	0/1780	1.13	15/2395 (0.6%)
6	F	0.74	0/717	1.30	7/967 (0.7%)
7	G	0.65	0/1368	1.15	11/1844 (0.6%)
8	H	0.50	0/1086	1.10	8/1470 (0.5%)
9	I	0.51	0/989	1.05	9/1331 (0.7%)
10	J	0.63	0/541	1.21	8/727 (1.1%)
11	K	0.60	0/937	1.02	4/1265 (0.3%)
12	L	0.65	0/365	1.14	2/485 (0.4%)
13	T	0.38	0/403	0.88	1/617 (0.2%)
14	N	0.42	0/164	0.73	0/252
15	P	0.43	0/239	0.79	0/371
All	All	0.62	3/32418 (0.0%)	1.13	265/43909 (0.6%)

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
13	T	0	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1170	THR	CA-CB	-5.37	1.45	1.53
1	A	55	ASP	CA-C	-5.24	1.46	1.52
2	B	968	VAL	CA-CB	-5.01	1.48	1.54

The worst 5 of 265 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	26	THR	N-CA-C	-15.91	89.31	113.02
6	F	71	GLU	N-CA-C	-13.74	95.85	113.17
1	A	452	LYS	N-CA-C	-13.65	96.48	111.36
2	B	1185	CYS	N-CA-C	-11.62	97.54	114.39
5	E	171	LYS	N-CA-C	-10.86	94.32	109.95

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
13	T	18	DC	Sidechain
13	T	19	DC	Sidechain
13	T	21	DG	Sidechain

5.2 Too-close contacts ⓘ

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	11166	0	11248	1247	0
2	B	8786	0	8819	1130	0
3	C	2095	0	2051	295	0
4	D	1365	0	1325	125	0
5	E	1744	0	1772	203	0
6	F	705	0	731	75	0
7	G	1340	0	1357	160	0
8	H	1068	0	1040	181	0
9	I	971	0	929	134	0
10	J	532	0	542	87	0
11	K	919	0	929	106	0
12	L	363	0	388	74	0
13	T	382	0	215	31	0
14	N	145	0	80	9	0
15	P	213	0	111	4	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
17	A	1	0	0	0	0
All	All	31803	0	31537	3535	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 56.

The worst 5 of 3535 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
14:N:5:DG:H2''	14:N:6:DT:H71	1.17	1.15
3:C:57:VAL:HG11	10:J:60:PHE:HB3	1.22	1.14
2:B:345:LYS:HE2	2:B:349:ILE:HD11	1.30	1.14
1:A:868:TYR:CE1	1:A:1064:VAL:HG11	1.83	1.14
3:C:189:THR:HG22	3:C:190:ASP:H	1.11	1.14

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1409/1733 (81%)	1031 (73%)	264 (19%)	114 (8%)	1	4
2	B	1087/1224 (89%)	798 (73%)	191 (18%)	98 (9%)	0	3
3	C	264/347 (76%)	192 (73%)	49 (19%)	23 (9%)	0	4
4	D	174/221 (79%)	121 (70%)	35 (20%)	18 (10%)	0	2
5	E	211/215 (98%)	155 (74%)	41 (19%)	15 (7%)	1	6
6	F	85/155 (55%)	72 (85%)	10 (12%)	3 (4%)	3	16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	G	169/171 (99%)	143 (85%)	20 (12%)	6 (4%)	2	16
8	H	129/146 (88%)	83 (64%)	26 (20%)	20 (16%)	0	0
9	I	117/122 (96%)	78 (67%)	27 (23%)	12 (10%)	0	2
10	J	63/70 (90%)	39 (62%)	15 (24%)	9 (14%)	0	0
11	K	112/120 (93%)	91 (81%)	15 (13%)	6 (5%)	1	10
12	L	44/70 (63%)	18 (41%)	14 (32%)	12 (27%)	0	0
All	All	3864/4594 (84%)	2821 (73%)	707 (18%)	336 (9%)	0	4

5 of 336 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ARG
1	A	59	GLY
1	A	67	CYS
1	A	70	CYS
1	A	74	MET

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	1242/1520 (82%)	1092 (88%)	150 (12%)	5	19
2	B	958/1061 (90%)	868 (91%)	90 (9%)	8	29
3	C	234/299 (78%)	210 (90%)	24 (10%)	7	24
4	D	141/200 (70%)	122 (86%)	19 (14%)	4	15
5	E	195/197 (99%)	175 (90%)	20 (10%)	7	24
6	F	77/137 (56%)	70 (91%)	7 (9%)	9	30
7	G	152/152 (100%)	138 (91%)	14 (9%)	8	29
8	H	117/128 (91%)	109 (93%)	8 (7%)	14	40
9	I	113/116 (97%)	95 (84%)	18 (16%)	2	10
10	J	60/65 (92%)	55 (92%)	5 (8%)	10	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	K	99/102 (97%)	90 (91%)	9 (9%)	9	30
12	L	40/57 (70%)	35 (88%)	5 (12%)	4	18
All	All	3428/4034 (85%)	3059 (89%)	369 (11%)	6	22

5 of 369 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1151	LEU
5	E	57	MET
2	B	1202	LEU
3	C	245	VAL
5	E	204	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 113 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	951	GLN
11	K	44	ASN
3	C	65	HIS
11	K	29	ASN
7	G	102	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	P	9/20 (45%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	BRU	T	20	15,13	18,21,22	3.94	1 (5%)	25,30,33	1.17	2 (8%)

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	BRU	T	20	15,13	-	0/7/21/22	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	T	20	BRU	BR-C5	-16.59	1.50	1.88

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	T	20	BRU	C6-C5-C4	-2.80	117.83	120.67
13	T	20	BRU	C2'-C1'-N1	-2.63	107.25	113.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	T	20	BRU	8	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	1419/1733 (81%)	0.22	75 (5%)	32 24	21, 68, 106, 122	0
2	B	1105/1224 (90%)	0.48	88 (7%)	18 16	28, 78, 111, 122	0
3	C	266/347 (76%)	0.09	4 (1%)	72 57	38, 66, 91, 115	0
4	D	178/221 (80%)	0.65	24 (13%)	7 8	46, 77, 106, 115	0
5	E	213/215 (99%)	0.70	20 (9%)	14 14	47, 91, 115, 121	0
6	F	87/155 (56%)	-0.25	2 (2%)	61 46	29, 49, 75, 86	0
7	G	171/171 (100%)	0.24	5 (2%)	53 39	45, 67, 96, 108	0
8	H	133/146 (91%)	1.11	21 (15%)	5 6	81, 98, 112, 118	0
9	I	119/122 (97%)	0.98	16 (13%)	7 8	64, 97, 116, 124	0
10	J	65/70 (92%)	0.07	2 (3%)	51 38	42, 63, 88, 99	0
11	K	114/120 (95%)	-0.11	2 (1%)	67 53	33, 70, 88, 101	0
12	L	46/70 (65%)	1.72	14 (30%)	1 2	53, 105, 121, 123	0
13	T	18/41 (43%)	1.71	4 (22%)	2 4	63, 103, 129, 134	1 (5%)
14	N	7/41 (17%)	2.37	3 (42%)	0 1	54, 116, 124, 126	1 (14%)
15	P	10/20 (50%)	1.22	2 (20%)	3 4	80, 99, 124, 127	0
All	All	3951/4696 (84%)	0.39	282 (7%)	22 18	21, 74, 111, 134	2 (0%)

The worst 5 of 282 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	883	LEU	8.0
1	A	3	GLY	7.9
1	A	1176	LEU	6.3
2	B	882	THR	6.1
14	N	7	DA	5.8

6.2 Non-standard residues in protein, DNA, RNA chains [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	BRU	T	20	20/21	0.70	0.20	75,80,83,84	0

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
17	MG	A	2458	1/1	0.89	0.14	100,100,100,100	0
16	ZN	I	1122	1/1	0.96	0.07	134,134,134,134	0
16	ZN	L	1071	1/1	0.97	0.06	119,119,119,119	0
16	ZN	A	2456	1/1	0.99	0.03	85,85,85,85	0
16	ZN	I	1121	1/1	1.00	0.02	75,75,75,75	0
16	ZN	A	2457	1/1	1.00	0.04	53,53,53,53	0
16	ZN	J	1066	1/1	1.00	0.03	64,64,64,64	0
16	ZN	B	2225	1/1	1.00	0.06	73,73,73,73	0
16	ZN	C	1269	1/1	1.00	0.05	65,65,65,65	0

6.5 Other polymers [i](#)

There are no such residues in this entry.