



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 06:45 am BST

PDB ID : 2JA5
Title : CPD lesion containing RNA Polymerase II elongation complex A
Authors : Brueckner, F.; Hennecke, U.; Carell, T.; Cramer, P.
Deposited on : 2006-11-23
Resolution : 3.80 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 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 : 2.11
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) : 2.11

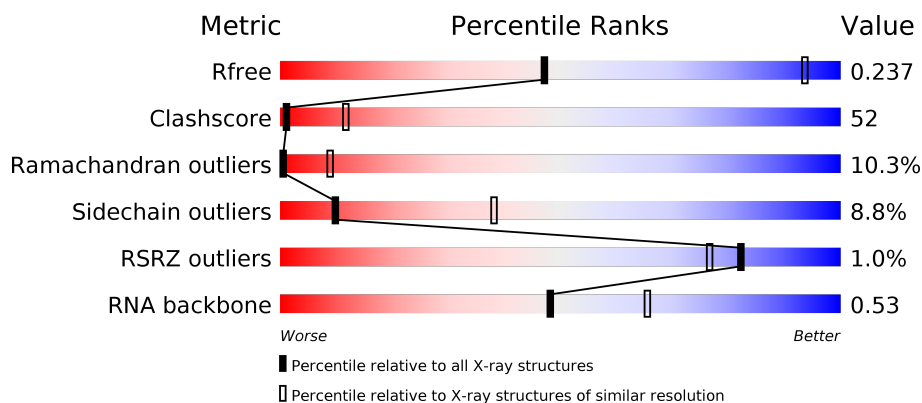
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.80 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	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)
RNA backbone	3102	1036 (4.60-3.00)

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 on 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	<div> <div>26%</div> <div>45%</div> <div>10%</div> <div>•</div> <div>18%</div> </div>
2	B	1224	<div> <div>27%</div> <div>52%</div> <div>11%</div> <div>•</div> <div>9%</div> </div>
3	C	318	<div> <div>25%</div> <div>49%</div> <div>9%</div> <div>•</div> <div>16%</div> </div>
4	D	221	<div> <div>25%</div> <div>41%</div> <div>12%</div> <div>•</div> <div>20%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	215	<div><div></div><div>46%49%5%</div></div>
6	F	155	<div><div></div><div>17%34%6%44%</div></div>
7	G	171	<div><div></div><div>%39%52%9%</div></div>
8	H	146	<div><div></div><div>3%29%53%10%8%</div></div>
9	I	122	<div><div></div><div>%34%47%13%5%</div></div>
10	J	70	<div><div></div><div>20%50%20%7%</div></div>
11	K	120	<div><div></div><div>2%33%49%11%5%</div></div>
12	L	70	<div><div></div><div>4%17%37%10%34%</div></div>
13	P	11	<div><div></div><div>9%9%73%9%9%</div></div>
14	T	25	<div><div></div><div>36%8%56%</div></div>

2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 31660 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	1421	Total	C	N	O	S	0	0	0
			11186	7048	1958	2118	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1115	Total	C	N	O	S	0	0	0
			8866	5614	1553	1644	55			

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	177	Total	C	N	O	S	0	0	0
			1427	882	256	287	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	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	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	135	Total	C	N	O	S	0	0	0
			1084	683	183	214	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	116	Total	C	N	O	S	0	0	0
			944	581	172	181	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
			364	224	72	64	4			

- Molecule 13 is a RNA chain called 5'-R(*UP*UP*CP*GP*AP*CP*CP*AP*GP*GP*AP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	P	10	Total	C	N	O	P	0	0	0
			212	96	41	66	9			

- Molecule 14 is a DNA chain called 5'-D(*AP*GP*CP*TP*CP*AP*AP*GP*TP*AP *CP*

TP*TTP*TP*TP*CP*CP*BRUP*GP*GP*TP*CP*AP*TP*T)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
14	T	11	Total	Br	C	N	O	P	0	0	0
			219	1	106	34	68	10			

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

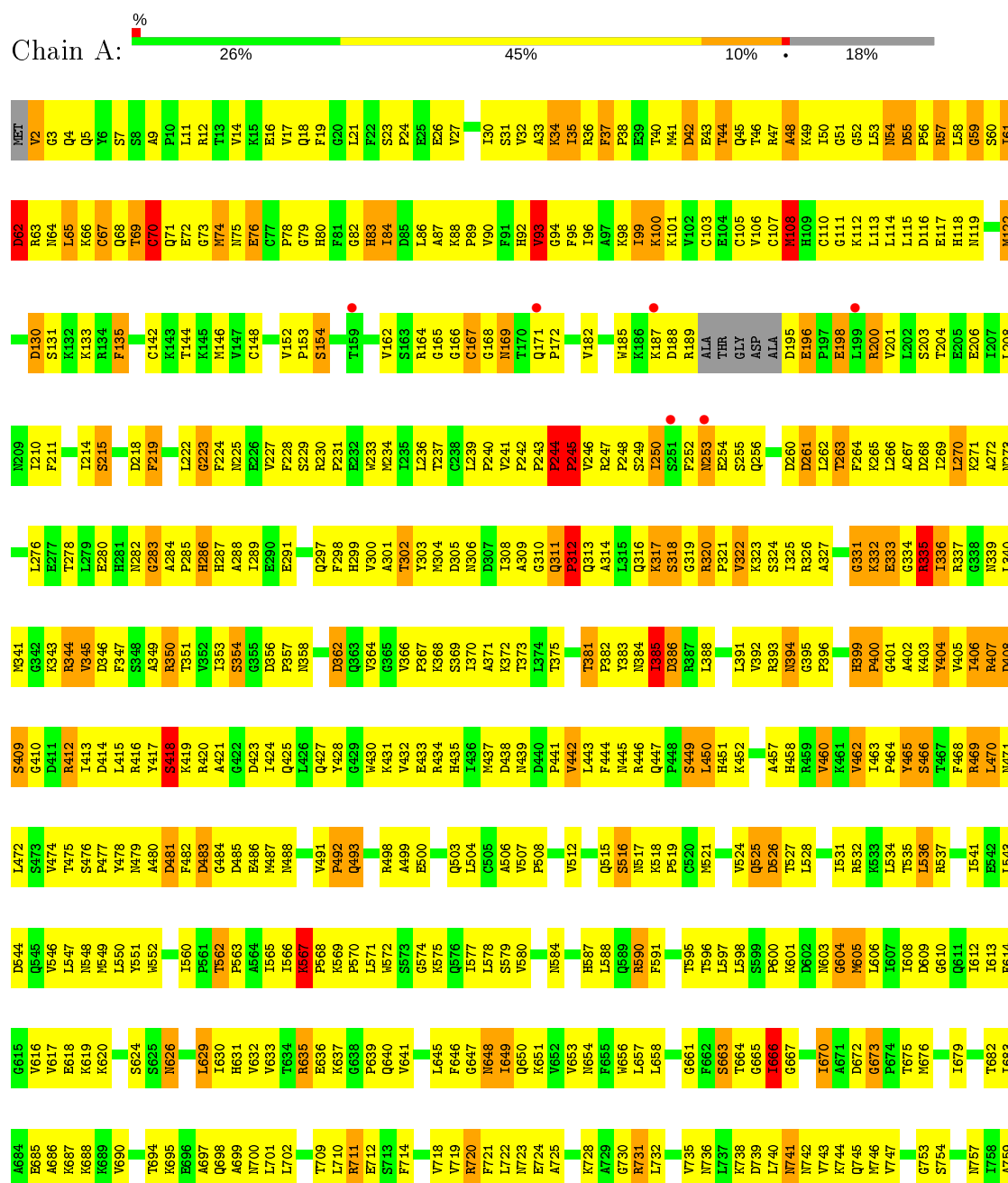
- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	8	Total	Zn	0	0
			8	8		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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

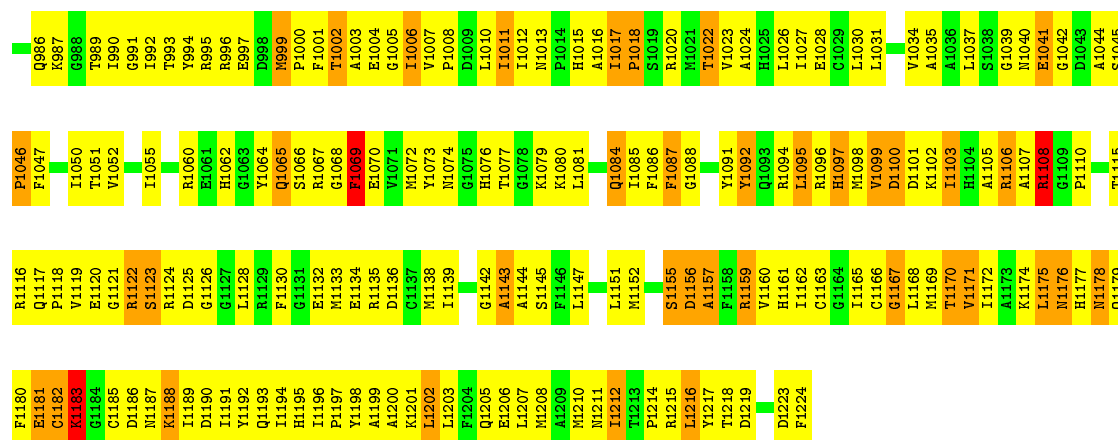


Q760	K895	Q968	L1037	N1106	LEU	LYS	V1316	L1381	V1443	ALA	TYR	SER
M761	K896	Q969	T1038	V1107	ASP	SER	M1317	L1382	M1444	MET	PRO	PRO
S762	K897	A1108	K1039	A1109	GLU	LEU	T1318	T1385	T1445	ALA	THR	THR
A763	R898	T970	Q1040	M1110	ALA	ALA	V1319	T1386	D1446	GLY	THR	THR
C764	V899	F971	Q1043	N1111	ALA	ALA	P1320	H1387	E1447	PHE	PRO	PRO
V765	D900	D974	D1043	K1112	GLN	THR	D1323	H1388	E1448	THR	SER	THR
Q766	L901	H975	M1048	T1113	SER	GLU	P1324	F1389	S1449	ALA	TYR	TYR
Q767	L902	T976	I1049	P1114	PHE	GLU	T1325	M1390	V1450	ALA	TYR	TYR
Q768	N903	S979	Q1052	S1115	ASP	ASP	R1326	M1393	K1452	GLY	PRO	PRO
S769	T904	D980	F1053	L1116	ASP	E1255	I1327	T1394	V1453	GLY	THR	THR
V770	R839	F1053	L1117	T1117	ASP	E1256	Y1328	G1395	M1454	ALA	PRO	PRO
E774	R840	L981	L1054	T1118	P1188	D1257	T1329	A1396	M1455	ASP	PRO	PRO
I775	L841	T907	R1055	Y1119	P1190	M1259	M1330	L1397		THR	TYR	TYR
	V842	T982	R1056	L1120	M1191	L1260	F1332	M1398		GLY	TYR	TYR
	K843	T983	V1057	L1121	L1192	K1261	L1339	C1400		GLN	PRO	PRO
	R844	R984	P1058	P1123	L1193	E1264	D1334	R1399		ILE	THR	THR
	L845	N986	H1059	H1124	L1197	N1265	I1335	S1401		GLU	SER	SER
	E846	V987	P1060	H1124	D1198	T1266	M1336	F1402		ILE	PRO	PRO
	R782	L988	G1061	D1127	D1198	L1268	E1337	E1403		GLY	PRO	PRO
	T783	D992	E1062	Q1130	M1202	L1268	V1338	E1404		ALA	TYR	TYR
	L784	L993	V1064				L1339	T1405		GLY	PRO	PRO
	H785	N996		I1134	K1205	I1271	G1340	V1406		GLN	PRO	PRO
	H786	L997	L1067	R1135	D1206	R1274	I1341	E1407		ASP	GLY	GLY
	H787	L998			L1207		E1342	I1408		GLY	SER	SER
	S788	Q926	Q1070	I1138	T1208		A1343	L1409		GLY	PRO	PRO
	K789	G926	S1071	I1138	M1209		G1344	F1410		VAL	PRO	PRO
			I1072	G1073	G1210		R1345	E1411		THR	TYR	TYR
			K1002	G1074	Q1211		L1348	A1412		PRO	PRO	PRO
			L929	E1004	V1212		L1349	G1413		TYR	PRO	PRO
				E1005	G1213		K1350	A1414		GLY	THR	THR
				I1006	E1214		R1351	S1415		PHE	SER	SER
				I1007	L1215		V1352	E1416		GLY	PRO	PRO
				Q1008	T1216			E1417		VAL	SER	SER
				Q1009	K1217			L1418		SER	TYR	TYR
				A1010	Q1218		V1355	D1419		SER	SER	SER
				Q1011	T1219		I1356	D1420		VAL	PRO	PRO
				R1012	F1220		D1359	C1421		ASN	THR	THR
				D1013	K1221			R1422		ALA	SER	SER
				A1014	M1222			G1423		ASP	PRO	PRO
				V1015	D1223		Y1362	V1424		LEU	SER	SER
				T1016	L1224			S1425		ASP	TYR	TYR
				L1017	F1225		M1363	E1426		VAL	SER	SER
				F1018	V1226		Y1365	H1427		LYS	PRO	PRO
				C1019	R1159		R1366	V1428		ASP	THR	THR
				L1020	S1160		H1367	L1430		GLU	SER	SER
				ALA	T1161		M1368	G1431		LEU	MET	MET
				VAL	V1162		A1369	Q1432		THR	PRO	PRO
				SER	I1163		L1370	M1433		SER	TYR	TYR
				K1092	P1164		L1371	A1434		PRO	SER	SER
				K1093	D1165		D1372	M1435		PRO	PRO	PRO
				V1094	D1166		D1373	I1436		ALA	TYR	TYR
				T1095	R1239		V1374	G1437		VAL	SER	SER
				S1096	C1240		M1375	T1376		ASP	PRO	PRO
				G1097	R1241		V1377	G1438		GLY	THR	THR
				V1098	Q1171		T1377	A1440		GLY	TYR	TYR
				P1099	F1174		Q1378	F1441		SER	PRO	PRO
				R1100	L1176		G1380	D1442		ASN	PRO	PRO
				Q1093	L1176					ASP	TYR	TYR

● Molecule 2: DNA-directed RNA polymerase II subunit RPB2

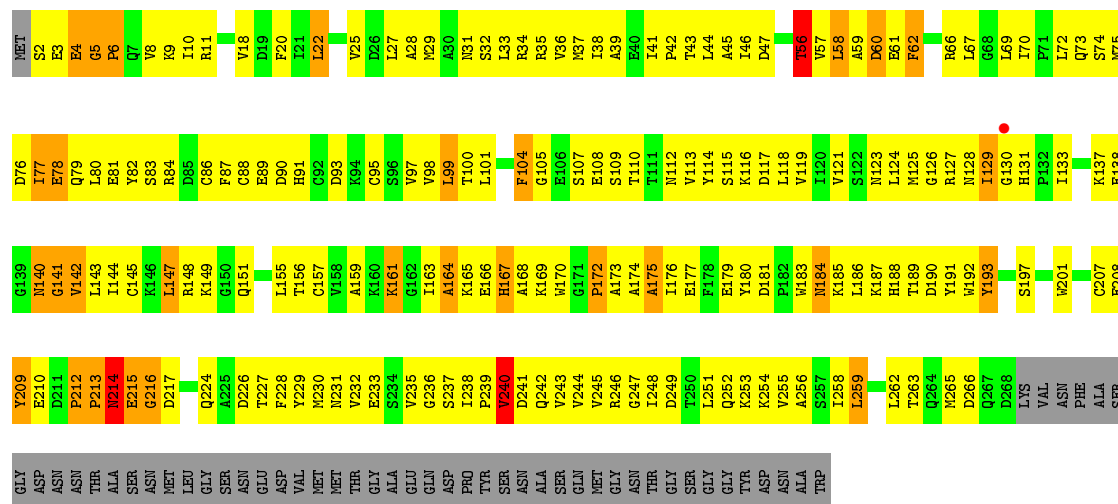


GLU	Y859	Y798	V731	ILE	T602	E529	A460	K393	E328	L258	D188	E65	MET
GLU	M860	P799	S732	GLU	L603	G530	L461	D894	T329	Y259	L189	D66	SER
LEU	Q800	Q801	H733	GLY	R604	A532	T463	Q395	A330	G260	Y190	S67	ASP
GLY	Q862	K801	H734	PHE	R605	A532	T463	D396	L331	R261	K191	LEU	ASP
GLN	E863	P802		GLU	G607	K637	M465	D397	D332	E262	V132	I70	ALA
ARG	K864		T737	ASP	D608	N638	M466	R398	F333	G263	K193	LEU	ASN
THR	K865	T805	F738	VAL	D609	N639	M467	D399	T334	S264	E194	GLU	SER
ALA	Y866	R807	T739	GLU	M610	L540	E468	H400	G336	C195	E194	GLN	GLU
THR	G867	A808	H740		P611	S540	E469	F401	R336	A266	P196	LEU	THR
HIS	M868	A809	C741	GLU	B612	M542	K470	K402	G338	R267	F197	ALA	LYS
ASP	S869	N809	E742	THR	P613	M542	K471	K403	G338	R268	D198	GLN	TYR
GLU	I870	E810	I743	GLU	V614	V547	A472	K404	T339	I269	M199	GLN	ASP
GLY	K871	Y811	H744	ASP	M615	G548	M473	R405	A340	L273	I204	THR	GLU
THR	E872	L812	P745	VAL	R616	T549	S474	L408	T343	P274	I205	THR	ASP
ALA	T873	R813	S746	GLU	R617	D550	S475	A409	K344	VAL	I205	GLU	PRO
THR	F874	R814	M747		D618	P551	S475	A409	Y275	PRO	M206	GLU	TYR
GLU		R815	I748	GLU	P619	M552	V479	P411	G410	ARG	G207	ASP	GLY
GLN	K875	E816	L749	GLU	M620	P553	S480	P412	K347	GLU	K210	ASN	F18
ASP	R876	L817	G750	THR	E621	I554	Q481	L412	K347	LEU	V211	ILE	E19
GLU	S877	P818	V751	ASP	R622	I555	Q482	L413	R346	LEU	L212	SER	D20
GLY	T878	A819	A752	VAL	E623	T556	M483	A414		ARG	L213	ARG	
THR	S879	G820	A753	GLU	E624	F557	M484	T419	Y351	TYR	L214	LYS	A23
ALA	T880	B821		GLU	L625	L558	Y486	L420	I355	GLU	Q215	GLU	P24
GLN	L881	N822	I756	THR	R626	S559	Y486	F421	I356	LEU	E216	LEU	I25
GLY	T882	G823	P757	ASP	P627	E560	T487		Q357	ILE	R217	GLU	T26
THR	L883	A824	F758	GLU	D628	M561	S490	L424	R358	ALA		ALA	A27
ASP	K884	E825	P759	GLU	T629	G562	T491	T425	K359	GLU	G220	GLU	E28
GLU	S885	R826	D760	GLU	D630	M563	T491	K426	F360	GLU	D221	GLU	D29
GLY	T886	H827	H761	GLU	P633	P571		D427	L361	SER	I222	SER	S30
THR	L887	N828	N762	GLU	V634	Y569	R496	I428	P362	ASP	V223	ASP	V33
ASP	K888	G829	Q763	GLU	R635	Y570	R497	F429	R363	ASP	Q224	ASP	I34
GLU	S889	E830	S764	GLU	P636	P571	T498	R430	I364	SER	V225	SER	S35
GLY	T890	L831	P765	GLU	L637	P571	M499	Q433	T365	GLU	F226	GLU	A36
THR	L891	G832	R766	GLU	F638	P575	I502	M432	K366	SER	E296	SER	F37
GLN	K892	N833	N767	GLU	E639	D576	GLY	Q433	L367	GLY	K228	GLY	F38
ASP	L893	E834	I768	GLU	V640	T578	ARG	R434	E368	GLY	A229	K164	R39
GLU	D894	R835	Q770	GLU	B641	Y577	ASP		G369	V165	A230	V165	E40
GLY	S895	D837	A771	GLU	D642	T578	ASP	E437	F370	F166	P231	F166	K41
THR	K896	E708	A772	GLU	B643	V580	GLY	GLU	E371	G107	D106	G107	K41
ASP	L897	D643		GLU	E644	F581	LYS	ALA	S372	G168	V108	L43	L43
GLN	S898	B644		GLU	B644	V582	LEU	HIS	R373	T109	T109	V44	V44
GLY	K899	K649		GLU	K649	N583	ALA	ASP	R373	S235	H110	H110	S45
THR	L899	E650		GLU	E650	G584	ALA	PHE	F376	P171	A111	A111	Q46
ASP	S899	L651		GLU	L651	V585	P511	ASN	F377	I172	L112	L112	Q47
GLU	T899	M652		GLU	M652	H586	R512	MET	L378	M173	Y113	Y113	L48
GLY	K900	G653		GLU	G653	H587	Q513	LYS	G379	M174	P114	P114	F51
THR	L901	R654		GLU	R654	G588	L514	L446	Y380	R175	Q115	Q115	N52
ASP	S902	K655		GLU	K655	V589	H515	A447	N381	R241	S176	S176	Q53
GLN	T903	G656		GLU	G656	H590	N516	I448	N382	S242	K177	K177	F54
GLY	K904	H657		GLU	H657	R591	T517	M449	N383	A243	M178	M178	F54
THR	L905	E659		GLU	E659	N592	H518	A450	N384	L244	C179	C179	V55
ASP	S906	A659		GLU	A659	P593	M519	K451	L385	E245	Y180	Y180	D66
GLU	T907	G660		GLU	G660	A594	G520	T452	L386	P316	E246	E246	N121
GLY	K908	L661		GLU	L661	A594	L521	I453	N387	C317	G247	G247	T58
THR	L909	E665		GLU	E665	M597	T454	C383	F322	E163	S182	S182	T58
ASP	S910	T666		GLU	T666	E598	A526	K389	V323	A184	I124	I124	D61
GLN	T911	D668		GLU	D668	T599	E526	L457	I324	T185	Y124	Y124	I62
GLY	K912			GLU		L600	T527	K458	R327	E186	S126	S126	I63
THR	L913			GLU		R601	P528	Y459					C64



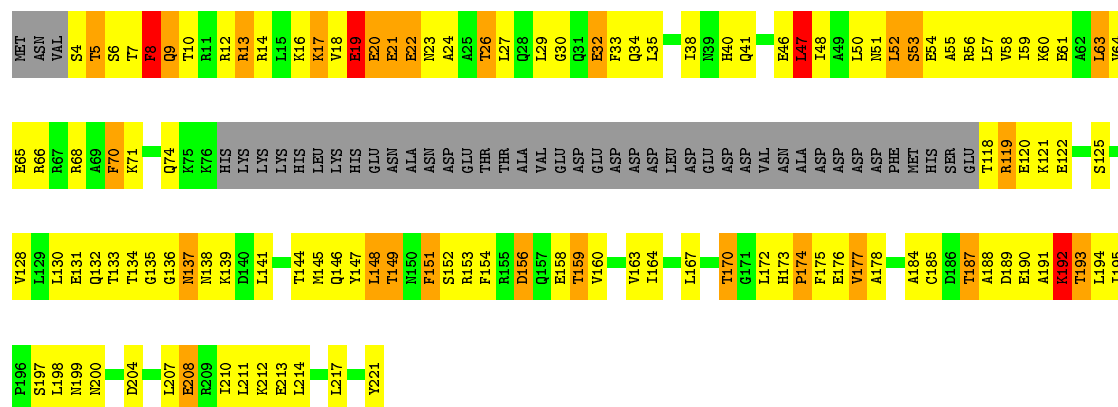
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C: 25% 49% 9% 16%

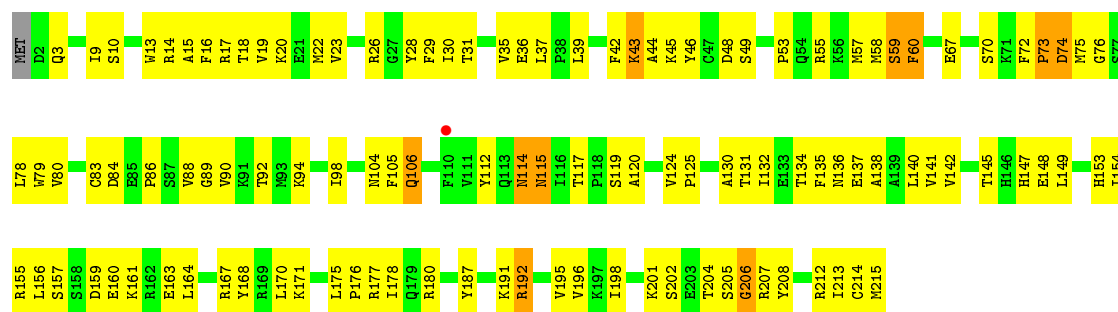


• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D: 25% 41% 12% 20%

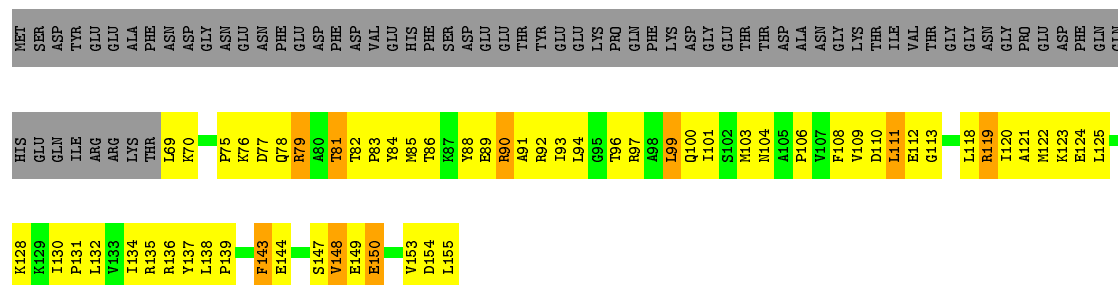


Chain E: 



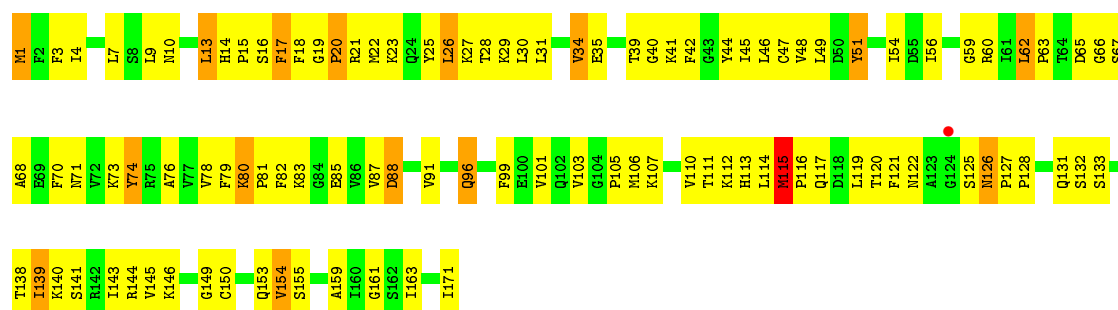
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 



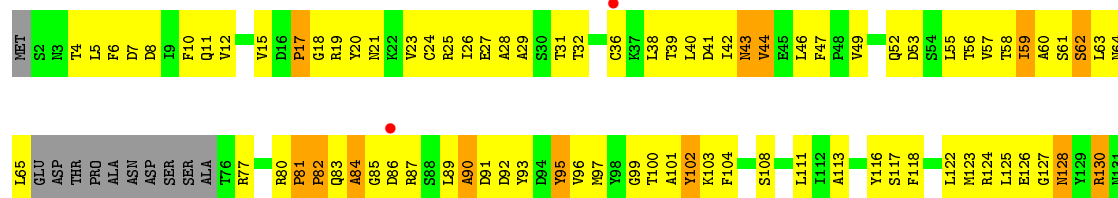
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

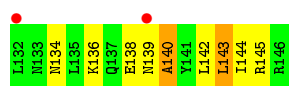
Chain G: 



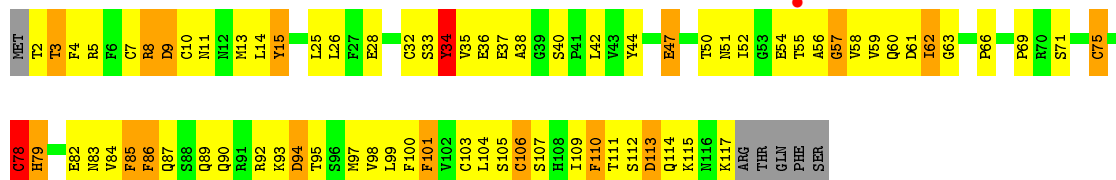
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 

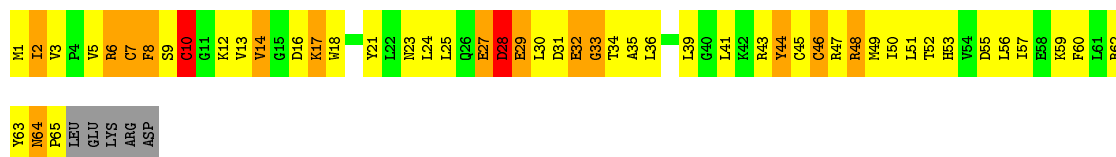
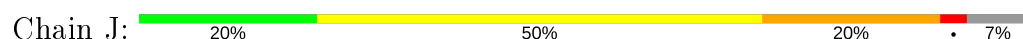




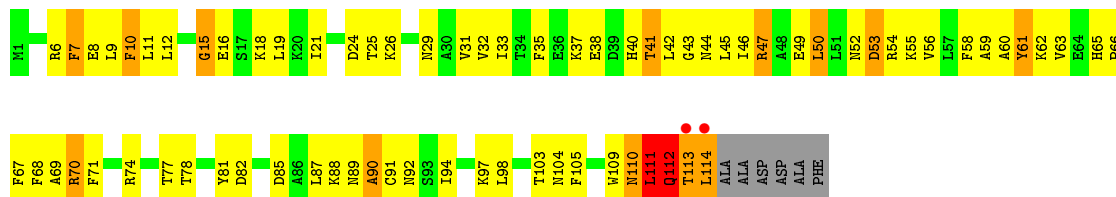
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



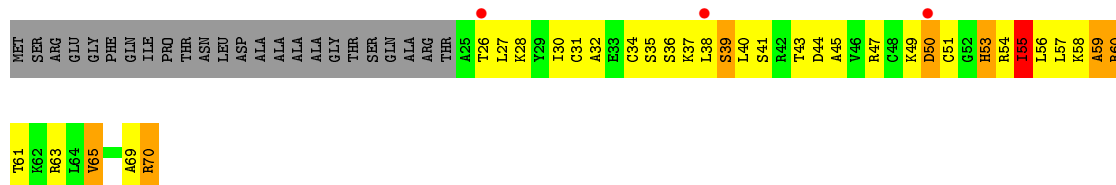
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



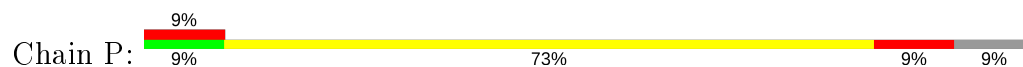
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11

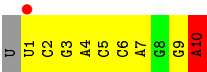


- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

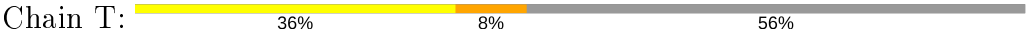


- Molecule 13: 5'-R(*UP*UP*CP*GP*AP*CP*CP*AP*GP*GP*AP)-3'





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



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	222.21Å 392.21Å 284.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.80 48.98 – 3.80	Depositor EDS
% Data completeness (in resolution range)	97.4 (50.00-3.80) 99.0 (48.98-3.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.79 (at 3.77Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.270 , 0.280 0.217 , 0.237	Depositor DCC
R_{free} test set	4681 reflections (1.98%)	wwPDB-VP
Wilson B-factor (Å ²)	99.7	Xtriage
Anisotropy	0.551	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 70.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.024 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.027 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	31660	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.95% 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: MG, ZN, 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	0.48	0/11385	0.73	1/15393 (0.0%)
2	B	0.46	0/9037	0.71	2/12181 (0.0%)
3	C	0.48	0/2138	0.72	0/2896
4	D	0.43	0/1437	0.68	1/1925 (0.1%)
5	E	0.43	0/1788	0.63	0/2406
6	F	0.55	0/716	0.77	0/964
7	G	0.48	0/1368	0.73	0/1844
8	H	0.40	0/1102	0.67	0/1492
9	I	0.41	0/962	0.68	0/1295
10	J	0.50	0/541	0.79	1/727 (0.1%)
11	K	0.90	6/937 (0.6%)	1.02	11/1265 (0.9%)
12	L	0.44	0/366	0.70	0/485
13	P	1.13	1/237 (0.4%)	1.22	2/368 (0.5%)
14	T	1.05	0/220	1.33	0/335
All	All	0.50	7/32234 (0.0%)	0.74	18/43576 (0.0%)

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	1
2	B	0	1
14	T	0	2
All	All	0	4

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	K	112	GLN	CA-C	9.87	1.78	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	K	113	THR	N-CA	9.16	1.64	1.46
11	K	112	GLN	CB-CG	9.05	1.76	1.52
11	K	112	GLN	N-CA	8.12	1.62	1.46
11	K	112	GLN	CG-CD	6.57	1.66	1.51

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	K	113	THR	N-CA-C	9.53	136.74	111.00
13	P	1	U	N1-C1'-C2'	9.37	126.18	114.00
11	K	112	GLN	N-CA-C	8.61	134.26	111.00
11	K	114	LEU	CB-CG-CD1	8.38	125.24	111.00
11	K	114	LEU	N-CA-C	7.93	132.41	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	303	TYR	Sidechain
2	B	486	TYR	Sidechain
14	T	19	DT	Sidechain
14	T	20	DC	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	11186	0	11266	1286	0
2	B	8866	0	8898	1020	0
3	C	2101	0	2055	267	0
4	D	1427	0	1451	141	0
5	E	1752	0	1776	127	0
6	F	705	0	730	84	0
7	G	1340	0	1357	161	0
8	H	1084	0	1057	123	0
9	I	944	0	899	101	0
10	J	532	0	542	98	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	K	919	0	929	109	0
12	L	364	0	386	43	0
13	P	212	0	109	20	0
14	T	219	0	125	31	0
15	A	1	0	0	0	0
16	A	8	0	0	0	0
All	All	31660	0	31580	3314	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 52.

The worst 5 of 3314 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)
11:K:112:GLN:CB	11:K:112:GLN:CG	1.77	1.62
11:K:112:GLN:CA	11:K:112:GLN:C	1.78	1.51
2:B:343:ILE:HG23	2:B:347:LYS:HB2	1.18	1.17
2:B:273:LEU:HB2	2:B:276:ILE:HD12	1.26	1.17
1:A:1445:ILE:H	1:A:1445:ILE:HD12	1.12	1.15

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	1410/1733 (81%)	975 (69%)	289 (20%)	146 (10%)	0	9
2	B	1096/1224 (90%)	781 (71%)	200 (18%)	115 (10%)	0	8
3	C	264/318 (83%)	171 (65%)	65 (25%)	28 (11%)	0	8
4	D	173/221 (78%)	124 (72%)	31 (18%)	18 (10%)	0	9
5	E	212/215 (99%)	154 (73%)	44 (21%)	14 (7%)	1	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	F	84/155 (54%)	69 (82%)	11 (13%)	4 (5%)	2	24
7	G	169/171 (99%)	131 (78%)	28 (17%)	10 (6%)	1	21
8	H	131/146 (90%)	82 (63%)	32 (24%)	17 (13%)	0	5
9	I	114/122 (93%)	80 (70%)	23 (20%)	11 (10%)	0	10
10	J	63/70 (90%)	37 (59%)	12 (19%)	14 (22%)	0	1
11	K	112/120 (93%)	85 (76%)	16 (14%)	11 (10%)	0	10
12	L	44/70 (63%)	18 (41%)	17 (39%)	9 (20%)	0	2
All	All	3872/4565 (85%)	2707 (70%)	768 (20%)	397 (10%)	0	9

5 of 397 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	44	THR
1	A	48	ALA
1	A	57	ARG
1	A	62	ASP
1	A	65	LEU

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	1244/1520 (82%)	1129 (91%)	115 (9%)	9	35
2	B	967/1061 (91%)	884 (91%)	83 (9%)	10	40
3	C	235/274 (86%)	214 (91%)	21 (9%)	9	38
4	D	159/200 (80%)	135 (85%)	24 (15%)	3	18
5	E	196/197 (100%)	192 (98%)	4 (2%)	55	75
6	F	77/137 (56%)	69 (90%)	8 (10%)	7	30
7	G	152/152 (100%)	141 (93%)	11 (7%)	14	45
8	H	119/128 (93%)	113 (95%)	6 (5%)	24	55
9	I	110/116 (95%)	98 (89%)	12 (11%)	6	29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	J	60/65 (92%)	53 (88%)	7 (12%)	5	27
11	K	99/102 (97%)	89 (90%)	10 (10%)	7	32
12	L	40/57 (70%)	36 (90%)	4 (10%)	7	32
All	All	3458/4009 (86%)	3153 (91%)	305 (9%)	10	38

5 of 305 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	429	PHE
2	B	909	ASP
9	I	94	ASP
2	B	496	ARG
2	B	682	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	515	HIS
2	B	842	ASN
9	I	12	ASN
2	B	516	ASN
2	B	734	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	P	9/11 (81%)	1 (11%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	P	10	A

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	$\# Z > 2$
14	BRU	T	22	13,14	15,21,22	1.87	3 (20%)	17,30,33	3.96	3 (17%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
14	BRU	T	22	13,14	-	0/4/21/22	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	22	BRU	C4-C5	5.50	1.45	1.38
14	T	22	BRU	C4-N3	3.45	1.39	1.33
14	T	22	BRU	C6-C5	-2.07	1.34	1.39

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	T	22	BRU	C4-N3-C2	14.02	126.98	115.14
14	T	22	BRU	C5-C4-N3	-6.79	115.51	123.64
14	T	22	BRU	C5-C6-N1	2.93	123.75	119.97

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	T	22	BRU	3	0

5.5 Carbohydrates [i](#)

There are no carbohydrates 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	2
1	A	1
3	C	1
6	F	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	2:SER	C	3:GLU	N	4.08
1	B	18:PHE	C	19:GLU	N	3.80
1	F	69:LEU	C	70:LYS	N	3.44
1	A	1175:SER	C	1176:LEU	N	3.41
1	B	337:ARG	C	338:GLY	N	2.62

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	1421/1733 (81%)	-0.34	10 (0%) 87 83	12, 73, 146, 189	0
2	B	1115/1224 (91%)	-0.25	17 (1%) 73 66	12, 83, 154, 191	0
3	C	267/318 (83%)	-0.39	1 (0%) 92 89	30, 69, 125, 150	0
4	D	177/221 (80%)	-0.18	0 100 100	53, 104, 145, 162	0
5	E	214/215 (99%)	-0.19	1 (0%) 91 87	45, 126, 174, 178	0
6	F	87/155 (56%)	-0.54	0 100 100	17, 49, 91, 119	0
7	G	171/171 (100%)	-0.26	1 (0%) 89 85	53, 77, 115, 126	0
8	H	135/146 (92%)	0.20	4 (2%) 50 40	87, 128, 162, 171	0
9	I	116/122 (95%)	-0.05	1 (0%) 84 79	69, 122, 151, 173	0
10	J	65/70 (92%)	-0.63	0 100 100	35, 66, 109, 118	0
11	K	114/120 (95%)	-0.33	2 (1%) 68 61	33, 73, 102, 140	0
12	L	46/70 (65%)	0.33	3 (6%) 18 14	69, 141, 161, 169	0
13	P	10/11 (90%)	-0.05	1 (10%) 7 6	71, 85, 160, 163	0
14	T	10/25 (40%)	-0.36	0 100 100	72, 96, 137, 158	0
All	All	3948/4601 (85%)	-0.27	41 (1%) 82 76	12, 82, 154, 191	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
11	K	114	LEU	6.4
11	K	113	THR	6.4
2	B	882	THR	5.4
2	B	471	LYS	3.8
8	H	139	ASN	3.6

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(\AA^2)	Q<0.9
14	BRU	T	22	20/21	0.87	0.18	59,61,68,69	0

6.3 Carbohydrates [i](#)

There are no carbohydrates 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(\AA^2)	Q<0.9
16	ZN	A	2464	1/1	0.98	0.07	82,82,82,82	0
16	ZN	A	2461	1/1	0.98	0.08	163,163,163,163	0
16	ZN	A	2462	1/1	0.99	0.07	23,23,23,23	0
16	ZN	A	2459	1/1	0.99	0.04	115,115,115,115	0
16	ZN	A	2463	1/1	0.99	0.13	42,42,42,42	0
16	ZN	A	2460	1/1	0.99	0.14	69,69,69,69	0
16	ZN	A	2465	1/1	0.99	0.07	31,31,31,31	0
15	MG	A	2457	1/1	1.00	0.15	21,21,21,21	0
16	ZN	A	2458	1/1	1.00	0.14	51,51,51,51	0

6.5 Other polymers [i](#)

There are no such residues in this entry.