



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 21, 2024 – 07:20 AM EDT

PDB ID : 2B8K  
Title : 12-subunit RNA Polymerase II  
Authors : Meyer, P.A.; Ye, P.; Zhang, M.; Suh, M.H.; Fu, J.  
Deposited on : 2005-10-07  
Resolution : 4.15 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

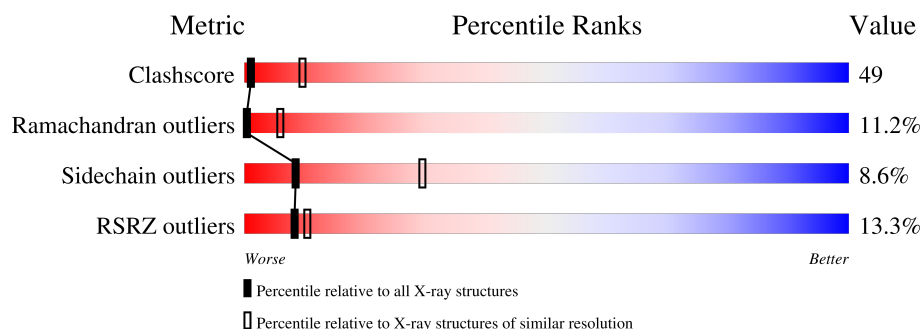
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.15 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))
Clashscore	180529	1048 (4.50-3.82)
Ramachandran outliers	177936	1231 (4.52-3.80)
Sidechain outliers	177891	1216 (4.52-3.80)
RSRZ outliers	164620	1239 (4.52-3.80)

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

Mol	Chain	Length	Quality of chain
1	A	1733	
2	B	1224	
3	C	318	
4	D	221	
5	E	215	
6	F	155	
7	G	215	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	146	<div><div></div><div>23%32%49%10%9%</div></div>
9	I	122	<div><div></div><div>22%40%43%13%. .</div></div>
10	J	70	<div><div></div><div>4%19%51%23%7%</div></div>
11	K	120	<div><div></div><div>8%41%47%8%. .</div></div>
12	L	70	<div><div></div><div>21%13%34%19%34%</div></div>

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 31040 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	1416	Total	C	N	O	S	0	0	0
			11140	7021	1946	2111	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1114	Total	C	N	O	S	0	0	0
			8800	5573	1540	1633	54			

- 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 polymerase II 32 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	182	Total	C	N	O	S	0	0	0
			1373	851	243	277	2			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1339	861	222	248	8			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	172	SER	-	expression tag	UNP P34087
G	173	HIS	-	expression tag	UNP P34087
G	174	GLU	-	expression tag	UNP P34087
G	175	LYS	-	expression tag	UNP P34087
G	176	ARG	-	expression tag	UNP P34087
G	177	ARG	-	expression tag	UNP P34087
G	178	TRP	-	expression tag	UNP P34087
G	179	LYS	-	expression tag	UNP P34087
G	180	LYS	-	expression tag	UNP P34087
G	181	ASN	-	expression tag	UNP P34087
G	182	PHE	-	expression tag	UNP P34087
G	183	ILE	-	expression tag	UNP P34087
G	184	ALA	-	expression tag	UNP P34087
G	185	VAL	-	expression tag	UNP P34087
G	186	SER	-	expression tag	UNP P34087
G	187	ALA	-	expression tag	UNP P34087
G	188	ALA	-	expression tag	UNP P34087
G	189	ASN	-	expression tag	UNP P34087
G	190	ARG	-	expression tag	UNP P34087
G	191	PHE	-	expression tag	UNP P34087
G	192	LYS	-	expression tag	UNP P34087
G	193	LYS	-	expression tag	UNP P34087
G	194	ILE	-	expression tag	UNP P34087
G	195	SER	-	expression tag	UNP P34087
G	196	SER	-	expression tag	UNP P34087
G	197	SER	-	expression tag	UNP P34087
G	198	GLY	-	expression tag	UNP P34087
G	199	ALA	-	expression tag	UNP P34087
G	200	LEU	-	expression tag	UNP P34087
G	201	ASP	-	expression tag	UNP P34087
G	202	TYR	-	expression tag	UNP P34087

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	203	ASP	-	expression tag	UNP P34087
G	204	ILE	-	expression tag	UNP P34087
G	205	PRO	-	expression tag	UNP P34087
G	206	THR	-	expression tag	UNP P34087
G	207	THR	-	expression tag	UNP P34087
G	208	ALA	-	expression tag	UNP P34087
G	209	SER	-	expression tag	UNP P34087
G	210	GLU	-	expression tag	UNP P34087
G	211	ASN	-	expression tag	UNP P34087
G	212	LEU	-	expression tag	UNP P34087
G	213	TYR	-	expression tag	UNP P34087
G	214	PHE	-	expression tag	UNP P34087
G	215	GLN	-	expression tag	UNP P34087

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

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 9.

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/III subunit 10.

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 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	115	Total	C	N	O	S	0	0	1
			920	590	157	171	2			

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

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

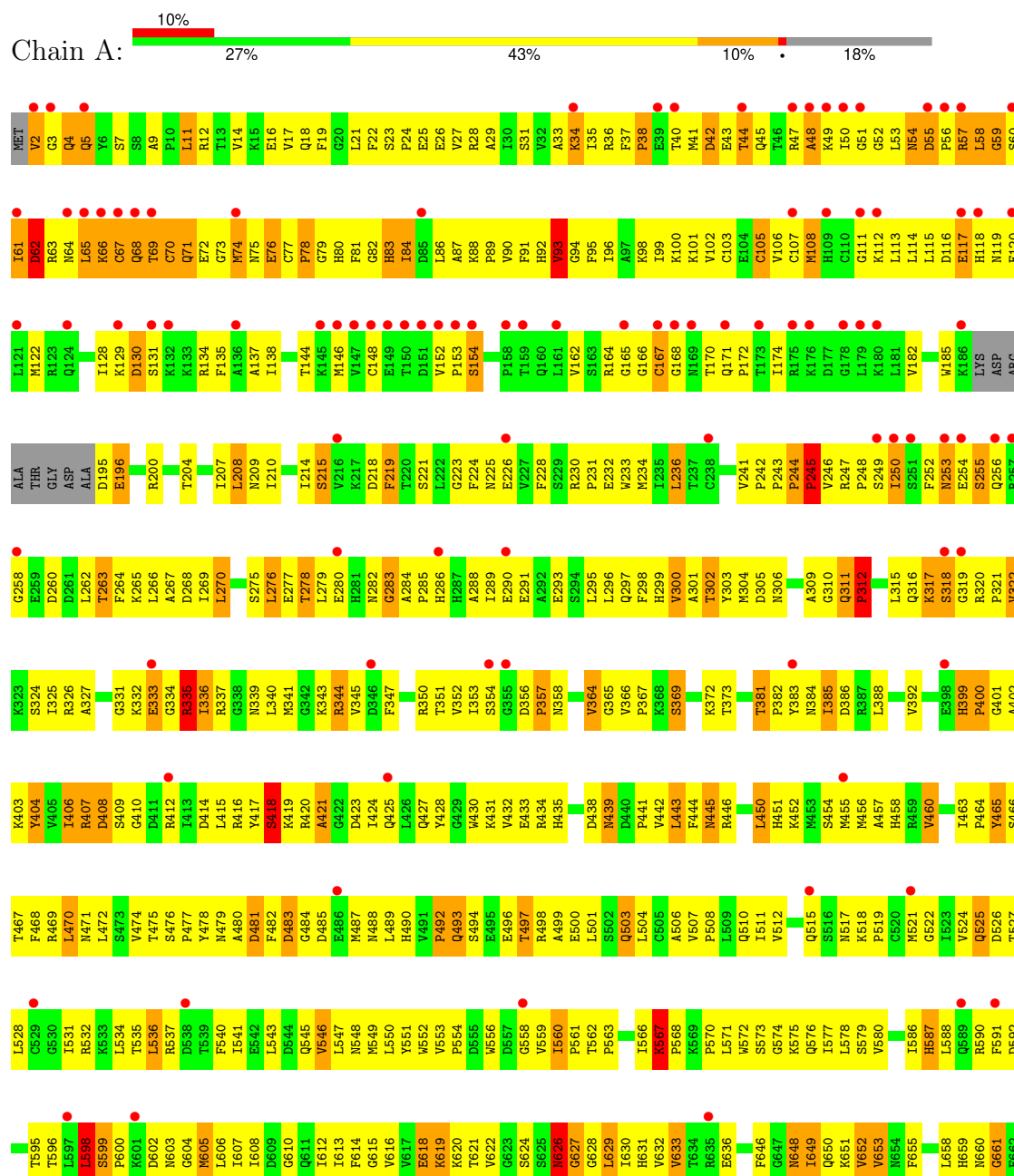
- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

### 3 Residue-property plots

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

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

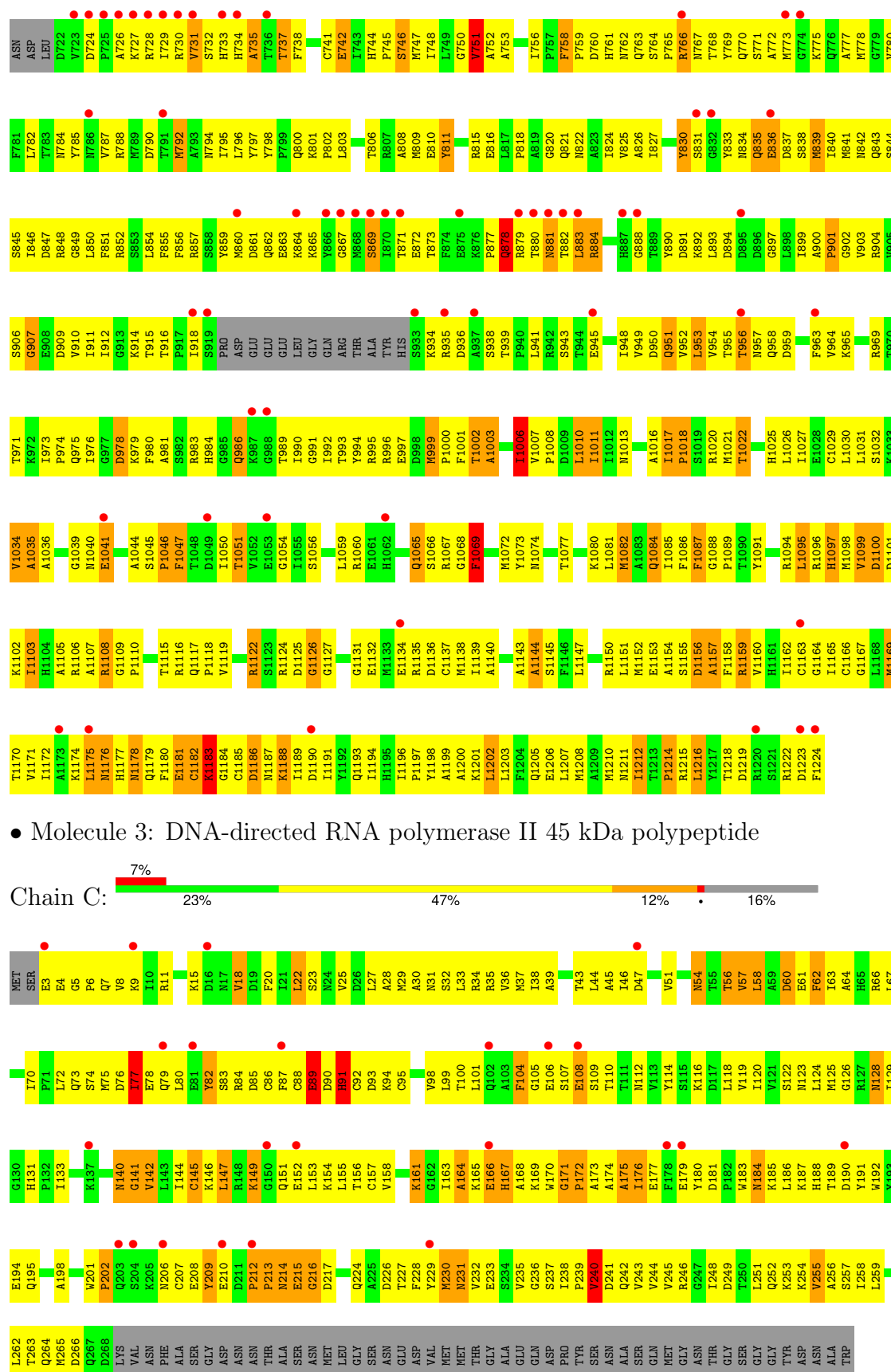




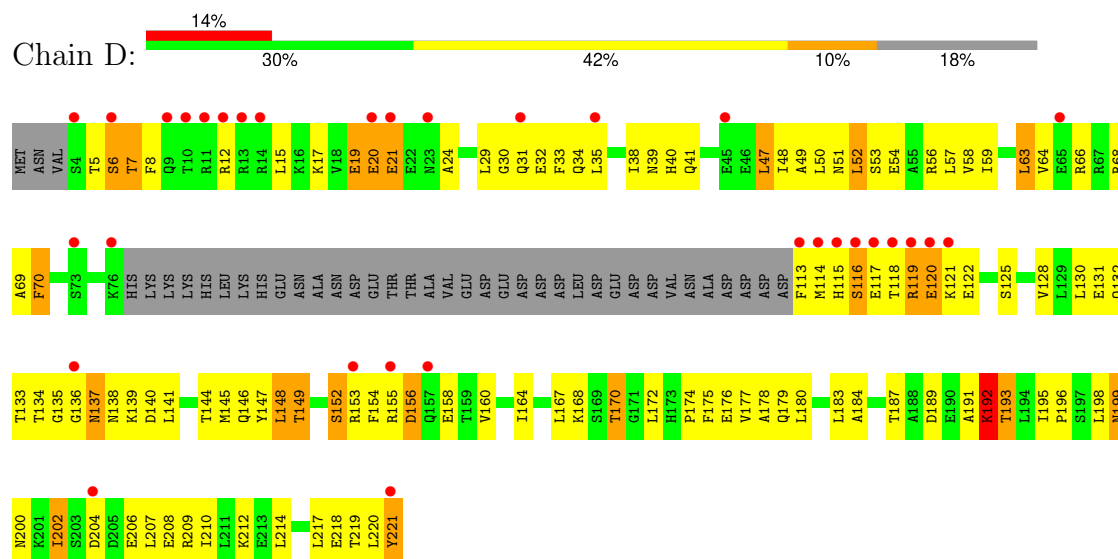
TYR	GLN	F1402	M1336	E1269	M1202	H1137	P1075	Q1008	Q935	S803	W736	S663
	ASP	E1403	V1337	M1270	M1203	I1138	A1076	M1009	Q869	S804	L936	T664
GLY	GLY	E1404	V1338	I1271	M1204	T1077	T1077	A1010	E870	L905	L737	G665
ALA	GLY	T1405	L1339	T1272	K1205	H1140	Q1078	Q1011	K938	R906	D739	E666
PRO	VAL	T1406	G1340	L1273	D1206	T1141	M1079	D939	D939	G807	L740	G667
THR	THR	E1407	I1341	R1274	I1207	T1147	T1080	D1013	R940	L808	W741	D668
SER	PRO	I1408	E1342		T1208	T1147	T1081	A1014	K941	T809	W742	T669
PRO	TYR	L1409	A1343		M1209	I1148	ASN	V1015	D875	P910	W743	T670
GLY	SER	E1410	R1344	I1280	Q1211	S1150	PHE	T1016	A876	Q811	K744	A671
ASN	ASN	R1281	R1345	R1281	Q1211	S1150	THR	T1017	H877	Q812	Q745	D672
GLY	GLY	A1412	A1346	V1282	V1212	E1151	HIS	F1018	H878	F813	W746	G673
VAL	SER	G1413	A1347	V1283	G1213	I1152	PHE	C1019	E879	F814	W747	P674
SER	GLY	A1414	L1348	M1284	E1214	Y1153	ALA	C1020	K880	F815	W748	T675
SER	LEU	S1415	V1349		R1215	Y1154	GLY	L1021	Q881	H816	A749	M676
PRO	VAL	K1350	K1350	D1288	Q1218	D1155	VAL	L1022	S882	A817	G750	
GLY	ASN	E1417	E1351	R1289	P1156	D1156	ALA	R1023	L883	M818	S751	W779
PHE	ALA	L1418	V1352	K1290	T1219	D1157	SER	S1024	D884	G819	K752	
SER	ASP	D1419		V1291	F1220	P1158	K1092		T885	G820	G753	T682
PRO	LEU	D1420	V1355	P1292	K1221	R1159	K1093	A1027	V858	R821	S754	I683
THR	ASP	C1421		S1293	M1222	S1160	T1094	T1028	R959	E822	F755	A684
SER	VAL	R1422		T1294	D1223	T1161	T1095	R1029	R961	G823	W756	E685
PRO	LYS	G1423		T1295	L1224	V1162	S1096	R1030	R962	L824	W757	A686
THR	ASP	V1424		G1296	F1225	I1163	G1097	I1031	I963	I825		
TYR	GLY	L1363	V1363	E1297	V1226	P1164	V1098	L1032	I964	D826	W760	V693
SER	LEU	M1364	M1364	Y1298	I1227	E1165	P1099	Q1053	K895	T827	W761	T694
PRO	MET	L1430	V1365	Y1299	M1228	D1166	E1100	V1035	R896	A828	S762	K695
THR	PHE	G1431	R1366	K1300		E1167	L1101	R1036	R967	R829	A763	E696
SER	SER	K1432	K1367		D1233	E1168		T1037	Q968	R830	G764	A697
PRO	PRO	M1433	M1368	E1303	E1234	I1169	L1105	L1037	R969	R831	W765	Q698
ALA	LEU	A1434	A1369	V1304	K1235	T1170	M1106	T1038	D900		G766	A699
VAL	VAL	F1435	L1370	V1305	L1236	Q1171	T1107	K1039	L901	T834	Q767	W700
SER	ASP	L1436	L1371	L1306	T1237	L1172	A1108	Q1040	L902	G835	Q768	L701
THR	SER	G1437	T1372	E1307	T1238	H1173	K1109	F1041	N903	Y836		
GLY	GLY	T1438	D1373	T1308	R1239	F1174	M1110	D1043	T904			
SER	SER	G1439	V1374	D1309	G1240	S1175	M1111	D1043	D905	R840	W774	K705
PRO	ASN	A1440	M1375		R1241	L1176	K1112	W1044	R806		I775	
SER	ASP	F1441	T1376	M1312	V1242	LEU	T1113	V1045	K977	V842		
TYR	ALA	D1442	T1377	L1313	V1243	ASP	P1114		T907	L941		
SER	MET	V1443	Q1378	S1314	ARG	GLU	S1115	N1048	P978	L908	F779	T709
PRO	ALA	M1444	G1379	E1315	PRO	GLU	L1116	I1049	S979	D909	W780	R711
THR	GLY	T1445	G1380	V1316	LYS	ALA	T1117	D980	P910	D980	D781	E712
SER	GLY	D1446	L1381	M1317	SER	GLU	V1118	E846	S911	E846	R782	S713
PRO	PHE	E1447		T1318	LEU	GLN	Y1119	L912	L913	D847	T783	F714
SER	THR		V1384	V1319	ASP	SER	L1120	L913	E914	I848	L784	W718
TYR	ALA	L1450	T1385	P1320	ALA	PHE	E1121	F1052	K984	E914	P785	V719
SER	TYR	L1451	R1386	G1321	GLU	ASP	P1122	R1055	D985	D853	H786	R720
PRO	GLY	K1452	H1387	I1322	THR	Q1187	G1123	S1056	I986	N854	S788	F721
THR	GLY	V1453	G1388	D1323	GLU	Q1188	H1124	V1057	I919	T855	K789	L722
SER	ALA	M1454	F1389	P1324	A1254	S1189	A1125	V1058	L920	T856	D790	R723
PRO	ASP	F1455	M1390	R1326	E1255	P1190	I1126	H059	D922	R857	W724	
SER	TYR	GLU		T1327	E1256	W1191	D1127	P1060	L923	N858	D791	A725
TYR	GLY	L1328	V1327	I1327	D1257	L1192	Q1128	V999	S859	K924	P794	
SER	GLY	L1394	T1394	V1328	L1193	L1193	F1129	L925	L860	L1000	G795	K728
PRO	ALA	G1395	G1395	T1329	R1194	R1194	Q1130	R1001	G926	G361	S796	A729
THR	THR	A1396	A1396	M1330	L1195	L1195	K1132	K1003	N962	G730	K797	R731
SER	SER	L1397	L1397	S1331	E1196	E1196	A1132	N963	L929	W793	G798	L732
PRO	GLU	F1332	M1398	T1265	L1197	L1197	L1134	I864	N1004	L799	W800	A733
SER	PHE	R1399	R1399	T1333	D1198	D1198	T1134	Q365	E932	G365	E801	R734
TYR	GLY	D1334	G1400	M1267	M1267	G1073	R1135	I1006	Y933	F866	W801	
ALA	GLY	S1401	L1395	T1295	L1269	E1404	S1126	K094	K094	T827	W802	W725



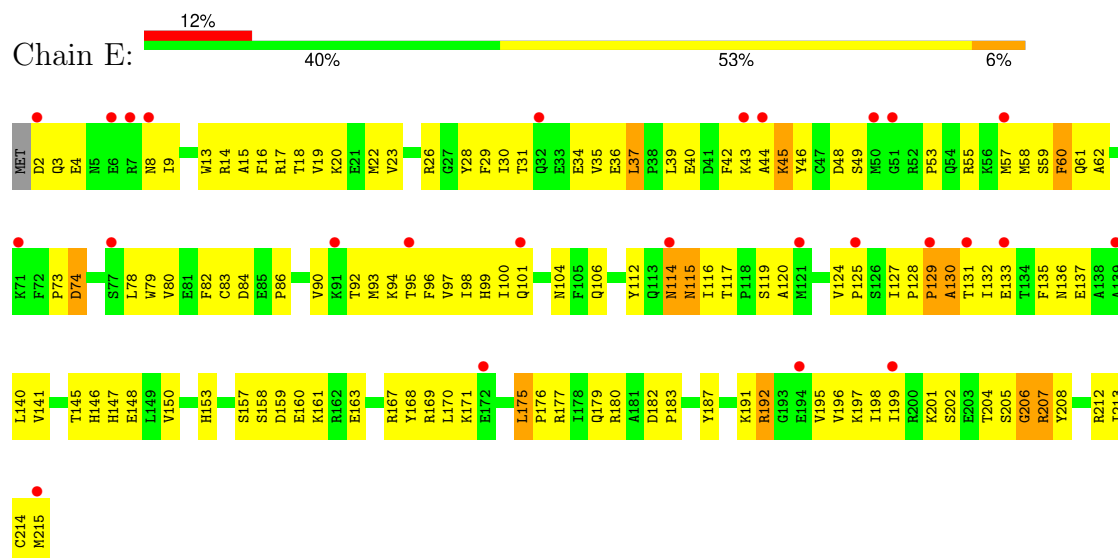
L658	L596	T527	M465	Q395	A330	G260	L192	L128	G64	MET
L661	L600	E529	W466	D396	L331	R261	K193	F129	E65	SER
M662	R601	G530	G467	D399	D332	E262	E194	F130	D66	ASP
A663	R602	G469	E468	D399	F333	G263	C195	P136	S67	LEU
T664	L603	G533	Q469	H400	T334	S264	K133	K133	T68	ALA
E665	L604	G536	K470	F401	G335	S265	F197	K134	L69	ASN
F666	R605	N538	K471	K404	ARG	A266	D198	ARG	I70	SER
Q667	K606	L539	M472	R405	ARG	R267	M199	THR	LEU	GLU
D668	G607	M642	M473	R405	GLY	T268	G200	GLY	LVS	TYR
TLE	D608	M642	S474	L408	THR	I269	G201	GLU	GLN	TYR
GLU	I609	R476	S475	A409	ALA	K270	Y202	ALA	LEU	TYR
GLY	N610	A477	R476	A409	LEU	L273	F203	ILE	ALA	ASP
GLY	P611	G478	G477	GLY	GLY	L273	I204	ASP	GLN	GLU
PHE	E612	G548	G478	P411	ILE	L276	I205	VAL	HIS	ASP
GLU	V613	T549	V479	L412	LVS	I276	N206	PRO	THR	PRO
ASP	S614	D550	S480	L413	L413	K277	G207	GLY	THR	TYR
VAL	M615	P551	L483	A414	K345	Q278	K210	ARG	GLU	GLY
GLU	I616	M552	M484	Q415	F346	D279	K210	GLU	SER	PHE
E678	R617	P553	R485	K418	Y351	I280	V211	LEU	ASP	GLU
F679	D618	I554	Y486	T419	L355	I282	V223	LVS	ASN	D20
T680	I619	T491	T487	L428	L355	V283	Q224	THR	ILE	E21
M681	R620	T556	Y488	F421	L366	I284	Y225	LEU	SER	S22
S682	E621	F557	S489	F429	Q357	I285	K217	ILE	ARG	A23
S683	K622	L568	S490	L424	K358	F286	K217	ILE	LVS	P24
L684	E623	S559	T491	D427	E359	R287	V223	ALA	TYR	I25
L685	L624	E560	L492	F360	L361	A288	Q224	GLU	GLU	T26
M686	V661	S561	S493	L428	L361	I291	Y225	SER	I90	E28
E687	K625	G562	H494	F429	P362	I292	F226	GLU	K94	D29
S688	F627	M563	L495	R430	H363	I292	K227	ASP	I95	
L689	T628	D628	R496	Y431	L364	D293	K228	ASP	I96	
V690	E629	L566	R497	T435	T365	D294	A229	SER	I97	
E691	R630	E567	T498	V436	Q366	G295	A230	GLU	T98	
V692	G631	D568	F437	V436	L367	E296	P231	SER	K99	
L693	R632	Y569	P501	E437	F368	I297	S232	SER	P100	
D694	V633	V670	I562	E438	G369	L298	P233	GLY	M101	
A695	G634	P571	GLY	A439	F370	E299	I234	K164	V102	
E696	R635	H572	ARG	H440	E371	H300	S235	I165	M103	
E697	P636	Q573	ASP	D441	S372	G296	H236	I167	E104	
L698	L637	S574	GLY	F442	R373	V305	V237	G168	G42	
E699	F638	P575	LVS	M443	R373	K306	A238	R169	V108	
S700	I639	D576	LEU	M444	F376	D307	E239	L170	K169	
L701	V640	A577	A509	R445	F377	R308	E240	L170	T109	
T702	E641	T578	L446	R445	L378	Q309	R241	P171	H110	
I703	D642	R579	P511	L446	G379	M310	M173	I172	A111	
A704	D643	V580	R512	A447	G379	M310	M173	I172	L112	
M705	E644	F581	Q513	A450	Y380	L311	L244	I174	Y113	
Q706	S645	V682	L514	R450	X381	E312	E245	R175	P114</	



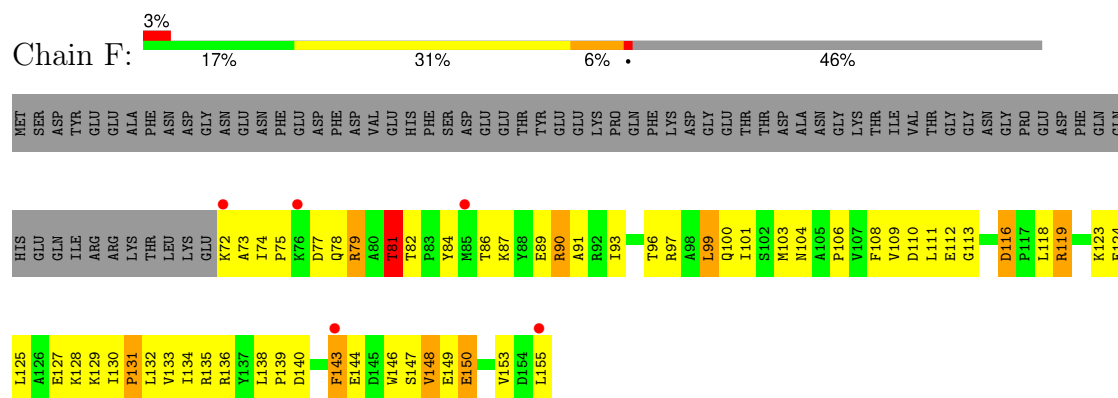
- Molecule 4: DNA-directed RNA polymerase II 32 kDa polypeptide



- Molecule 5: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide



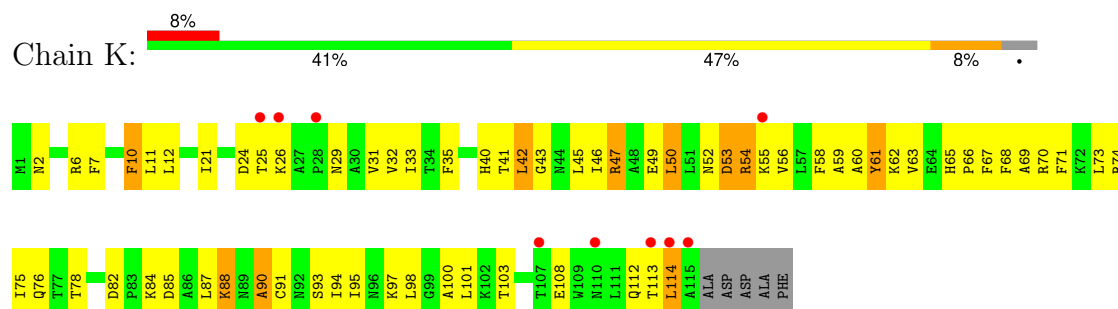
- Molecule 6: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



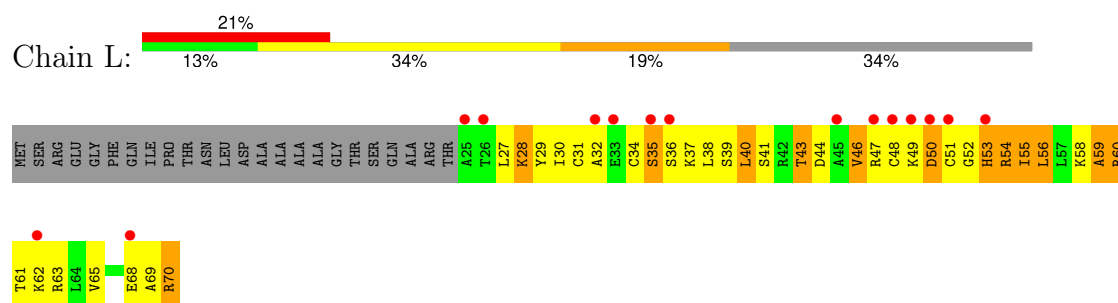
- Molecule 7: DNA-directed RNA polymerase II 19 kDa polypeptide



- Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.69Å 394.33Å 281.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	158.11 – 4.15 158.11 – 4.15	Depositor EDS
% Data completeness (in resolution range)	(Not available) (158.11-4.15) 83.4 (158.11-4.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 4.15Å)	Xtriage
Refinement program		Depositor
R, $R_{free}$	0.387 , (Not available) 0.306 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	118.6	Xtriage
Anisotropy	0.584	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 176.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.023 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.023 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.83	EDS
Total number of atoms	31040	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	169.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.01% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.49	3/11339 (0.0%)	0.75	9/15334 (0.1%)
2	B	0.53	6/8971 (0.1%)	0.97	25/12103 (0.2%)
3	C	0.52	0/2133	0.76	0/2891
4	D	0.49	0/1382	0.81	3/1862 (0.2%)
5	E	0.44	0/1788	0.64	0/2406
6	F	0.53	0/691	0.78	0/933
7	G	0.53	0/1367	0.74	0/1844
8	H	0.40	0/1086	0.66	0/1470
9	I	0.83	1/989 (0.1%)	0.94	3/1331 (0.2%)
10	J	0.54	0/541	0.90	1/727 (0.1%)
11	K	0.50	0/938	0.68	0/1267
12	L	0.55	0/365	0.79	0/485
All	All	0.52	10/31590 (0.0%)	0.82	41/42653 (0.1%)

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	2
2	B	1	10
3	C	0	1
9	I	0	1
All	All	1	14

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	39	GLY	C-N	-21.36	0.84	1.34
2	B	442	PHE	C-N	-8.69	1.14	1.34

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	446	LEU	N-CA	-8.12	1.30	1.46
2	B	439	ALA	C-N	7.40	1.51	1.34
1	A	1274	ARG	C-N	-6.45	1.21	1.33

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	475	SER	CB-CA-C	-51.20	12.82	110.10
9	I	39	GLY	O-C-N	-18.15	93.66	122.70
2	B	439	ALA	N-CA-CB	-16.49	87.01	110.10
2	B	442	PHE	C-N-CA	15.77	161.13	121.70
2	B	476	ARG	C-N-CA	-15.07	84.02	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	475	SER	CA

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	303	TYR	Sidechain
1	A	807	GLY	Mainchain
2	B	217	ARG	Mainchain
2	B	405	ARG	Mainchain
2	B	438	GLU	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	11140	0	11218	1169	0
2	B	8800	0	8777	947	0
3	C	2095	0	2053	241	0
4	D	1373	0	1312	144	0
5	E	1752	0	1776	149	0
6	F	679	0	701	94	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	1339	0	1357	145	0
8	H	1068	0	1040	107	0
9	I	971	0	929	91	0
10	J	532	0	542	93	0
11	K	920	0	929	83	0
12	L	363	0	387	47	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
All	All	31040	0	31021	3039	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 49.

The worst 5 of 3039 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:469:GLN:CA	2:B:474:SER:CB	1.80	1.51
4:D:119:ARG:N	4:D:121:LYS:HB2	1.46	1.30
4:D:113:PHE:CB	4:D:156:ASP:OD1	1.78	1.30
4:D:118:THR:HA	4:D:121:LYS:CB	1.64	1.27
2:B:435:THR:CG2	2:B:439:ALA:HB2	1.63	1.26

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	1406/1733 (81%)	949 (68%)	293 (21%)	164 (12%)	0	5
2	B	1096/1224 (90%)	744 (68%)	226 (21%)	126 (12%)	0	5
3	C	264/318 (83%)	159 (60%)	66 (25%)	39 (15%)	0	3
4	D	178/221 (80%)	124 (70%)	35 (20%)	19 (11%)	0	6
5	E	212/215 (99%)	147 (69%)	50 (24%)	15 (7%)	1	13
6	F	82/155 (53%)	64 (78%)	14 (17%)	4 (5%)	2	19
7	G	169/215 (79%)	131 (78%)	26 (15%)	12 (7%)	1	13
8	H	129/146 (88%)	84 (65%)	29 (22%)	16 (12%)	0	4
9	I	117/122 (96%)	79 (68%)	31 (26%)	7 (6%)	1	16
10	J	63/70 (90%)	37 (59%)	10 (16%)	16 (25%)	0	1
11	K	113/120 (94%)	89 (79%)	18 (16%)	6 (5%)	1	18
12	L	44/70 (63%)	19 (43%)	14 (32%)	11 (25%)	0	1
All	All	3873/4609 (84%)	2626 (68%)	812 (21%)	435 (11%)	0	6

5 of 435 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	48	ALA
1	A	54	ASN
1	A	55	ASP
1	A	57	ARG

### 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	1239/1520 (82%)	1135 (92%)	104 (8%)	9	30
2	B	952/1061 (90%)	865 (91%)	87 (9%)	7	26
3	C	234/274 (85%)	212 (91%)	22 (9%)	7	25
4	D	138/200 (69%)	122 (88%)	16 (12%)	4	19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	196/197 (100%)	187 (95%)	9 (5%)	23	47
6	F	74/137 (54%)	65 (88%)	9 (12%)	4	18
7	G	152/190 (80%)	142 (93%)	10 (7%)	14	37
8	H	117/128 (91%)	111 (95%)	6 (5%)	20	44
9	I	113/116 (97%)	99 (88%)	14 (12%)	4	18
10	J	60/65 (92%)	54 (90%)	6 (10%)	6	23
11	K	99/102 (97%)	92 (93%)	7 (7%)	12	35
12	L	40/57 (70%)	37 (92%)	3 (8%)	11	33
All	All	3414/4047 (84%)	3121 (91%)	293 (9%)	8	29

5 of 293 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	104	ASN
11	K	47	ARG
6	F	79	ARG
8	H	95	TYR
1	A	1400	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 77 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	167	HIS
9	I	12	ASN
3	C	252	GLN
5	E	104	ASN
11	K	65	HIS

### 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 [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 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
1	A	1
2	B	1
9	I	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	807:GLY	C	808:LEU	N	1.20
1	B	442:PHE	C	443:ASN	N	1.14
1	I	39:GLY	C	40:SER	N	0.84

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	1416/1733 (81%)	1.05	181 (12%)	9 11	96, 165, 229, 265	0
2	B	1114/1224 (91%)	1.03	157 (14%)	7 10	20, 162, 239, 270	0
3	C	266/318 (83%)	0.87	23 (8%)	18 17	114, 157, 215, 237	0
4	D	182/221 (82%)	1.25	32 (17%)	4 7	20, 170, 207, 228	0
5	E	214/215 (99%)	0.96	26 (12%)	10 12	111, 196, 241, 247	0
6	F	84/155 (54%)	0.81	5 (5%)	29 25	109, 143, 189, 209	0
7	G	171/215 (79%)	0.77	9 (5%)	33 28	119, 147, 177, 201	0
8	H	133/146 (91%)	1.38	34 (25%)	2 4	173, 210, 246, 255	0
9	I	119/122 (97%)	1.53	27 (22%)	3 5	104, 200, 236, 277	0
10	J	65/70 (92%)	0.72	3 (4%)	38 31	119, 156, 197, 204	0
11	K	115/120 (95%)	0.85	9 (7%)	20 19	119, 160, 191, 199	0
12	L	46/70 (65%)	1.78	15 (32%)	1 3	153, 214, 245, 253	0
All	All	3925/4609 (85%)	1.04	521 (13%)	8 10	20, 164, 234, 277	0

The worst 5 of 521 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	472	ALA	8.5
1	A	1081	LEU	7.3
4	D	116	SER	6.6
2	B	471	LYS	6.4
2	B	882	THR	6.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
13	ZN	B	1225	1/1	0.95	0.21	20,20,20,20	0
13	ZN	A	1735	1/1	0.96	0.21	20,20,20,20	0
13	ZN	A	1734	1/1	0.97	0.27	20,20,20,20	0
13	ZN	C	319	1/1	0.98	0.17	20,20,20,20	0
13	ZN	I	124	1/1	0.99	0.25	20,20,20,20	0
13	ZN	L	71	1/1	0.99	0.23	20,20,20,20	0
13	ZN	J	71	1/1	1.00	0.18	20,20,20,20	0
13	ZN	I	123	1/1	1.00	0.20	20,20,20,20	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.