



wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 03:11 pm BST

PDB ID : 2R7Z
Title : Cisplatin lesion containing RNA polymerase II elongation complex
Authors : Damsma, G.E.; Alt, A.; Brueckner, F.; Carell, T.; Cramer, P.
Deposited on : 2007-09-10
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

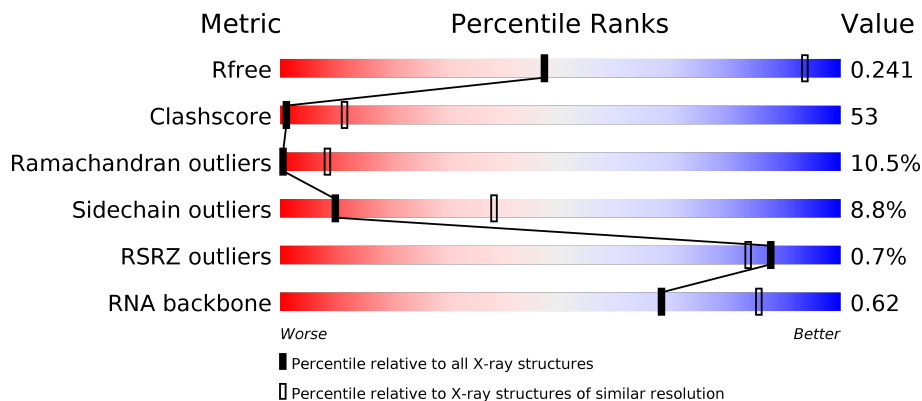
1 Overall quality at a glance

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	T	17	 65% 29% 6%
2	N	7	 71% 29%
3	P	10	 20% 70% 10%
4	A	1733	 26% 45% 10% 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	<p>%</p> <p>27% 52% 11% • 9%</p>
6	C	318	<p>23% 48% 11% • 16%</p>
7	D	221	<p>28% 38% 11% • 20%</p>
8	E	215	<p>39% 55% 5%</p>
9	F	155	<p>19% 28% 6% • 46%</p>
10	G	171	<p>37% 54% 9%</p>
11	H	146	<p>27% 52% 12% 9%</p>
12	I	122	<p>2%</p> <p>34% 52% 11% ••</p>
13	J	70	<p>29% 47% 16% • 7%</p>
14	K	120	<p>%</p> <p>28% 57% 11% 5%</p>
15	L	70	<p>11% 39% 13% • 34%</p>

2 Entry composition [i](#)

There are 18 unique types of molecules in this entry. The entry contains 31804 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 DNA chain called 5'-D(*TP*AP*CP*TP*TP*GUP*CP*CP*CP*TP*CP*CP*TP*CP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	T	17	336	163	53	104	16	0	0	0

- Molecule 2 is a DNA chain called 5'-D(*CP*AP*AP*GP*TP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	N	7	143	69	30	38	6	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	P	10	216	97	41	69	9	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	A	1416	11140	7021	1946	2111	62	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	1108	8810	5580	1541	1634	55	0	0	0

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

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	D	177	1427	882	256	287	2	0	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	E	214	1752	1111	309	321	11	0	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

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

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB7.

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

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

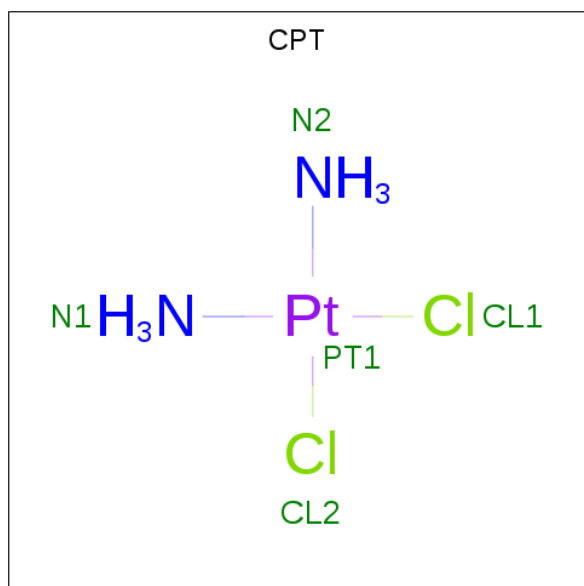
- Molecule 14 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 15 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

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

- Molecule 16 is Cisplatin (three-letter code: CPT) (formula: $\text{Cl}_2\text{H}_6\text{N}_2\text{Pt}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
16	T	1	Total	N	Pt	0	0
			3	2	1		

- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	J	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

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

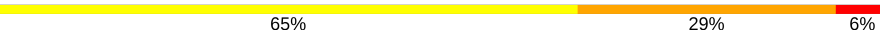
- Molecule 18 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

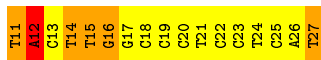
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	1	Total 1	Mg 1	0	0

3 Residue-property plots [i](#)


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: 5'-D(*TP*AP*CP*TP*TP*GUP*CP*CP*CP*TP*CP*CP*TP*CP*AP*T)-3'

Chain T: 




- Molecule 2: 5'-D(*CP*AP*AP*GP*TP*AP*G)-3'

Chain N: 



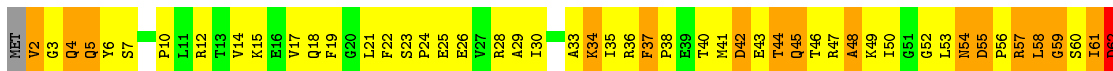
- Molecule 3: 5'-R(*UP*UP*UP*GP*AP*GP*GP*AP*GP*G)-3'

Chain P: 



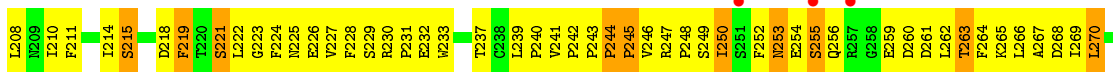
- Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain A: 

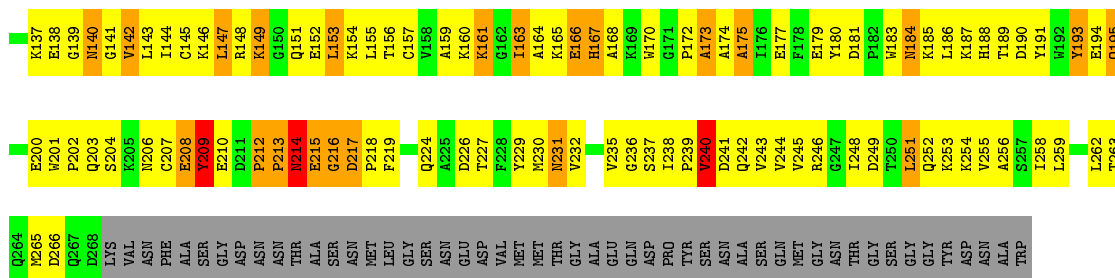






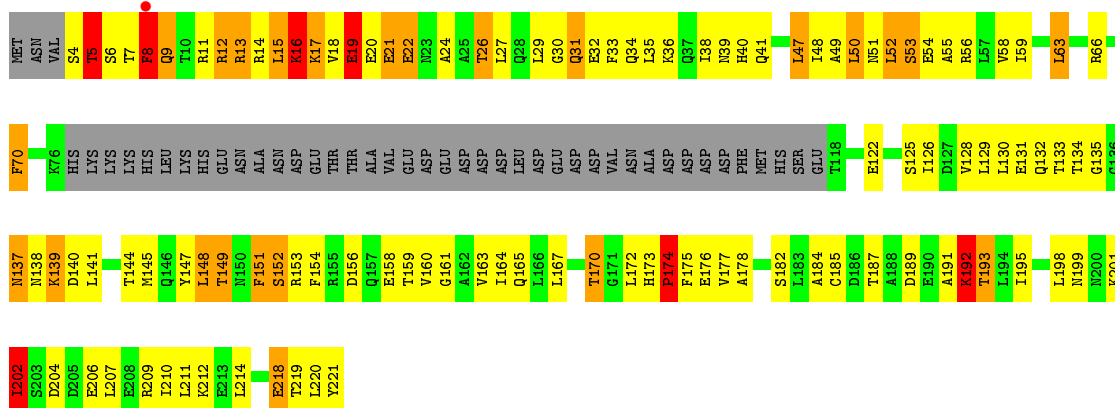


D1223	L1261	ALA	R821	V747	T675	L606	L536	R469	D408	N339	L276
L1224	V1162	SER	I825	S751	M676	F607	R537	L470	S409	L340	E277
F1225	I1163	K1092	I826	K752	F679	I608	I541	N471	G410	M341	L278
V1226	P1164	R1093	D826	S753	T680	D609	E542	L472	D411	G342	L279
I1227	P1165	R1030	T827	G754	E681	I613	L543	S473	R412	K343	E280
V1228	D1166	I1031	A828	S755	T682	F614	D544	V474	I413	K344	E281
S1229	E1167	L1032	V829	N757	T683	G615	Q945	S475	D414	V345	N282
L1236	E1168	Q1033	K830	I757	E684	V616	V546	S476	R415	D346	G283
L1237	I1169	A966	T831	M761	A684	V616	V546	S477	R416	F347	A284
I1238	L1170	Q1036	E832	S762	E685	V617	L547	N478	R417	S348	P285
R1239	Q1171	R1036	E833	G763	A686	E618	N548	N479	S418	P286	E286
C1240	L1172	T1037	T834	A764	D692	K619	M549	A460	K419	E287	E287
R1241	E1173	L1038	G835	C764	V693	K620	M550	A461	R420	R351	E288
V1242	S1174	K1039	V836	V765	V694	T621	Y551	F462	K420	V352	L289
V1243	S1175	Q1040	I837	G767	T694	V622	W552	F462	D423	E290	E290
ARG	L1176	F1042	R839	Q768	K695	G623	W556	E466	I424	E291	E291
PRO	LEU	F1043	R840	Q768	A697	S624	W557	M467	Q426	G355	E292
LYS	ASP	D1043	R841	R774	Q698	S625	G558	M468	Q427	G356	E293
LEU	GLU	D1044	L841	I775	K698	G627	V559	L469	Y428	P357	Q297
LEU	GLU	V1045	V842	I775	A699	E628	I560	H490	G429	E360	F298
LEU	GLU	L1046	R843	F779	N700	L629	V560	V491	W430	L361	H299
ASP	ALA	S1047	A844	V780	L701	I630	T562	P492	W431	V300	H300
ALA	GLU	N1048	L845	V781	L702	H631	P563	Q493	K431	A301	A301
ALA	GLN	I1049	E846	D781	H631	H632	A564	S494	V432	V366	T302
GLN	SER	E1050	D847	R782	T709	V632	I565	E495	E433	P367	Y303
THR	SER	E1050	E848	T783	L710	V633	I566	E496	R434	R368	M304
PHE	PHE	H851	M849	L784	R711	T634	K567	E497	H435	R369	D305
ASP	ASP	Q1052	V850	F785	E712	R635	W569	M467	I436	D306	D306
ASP	GLU	V1058	R851	H786	S713	P639	P568	R498	M437	N306	N306
ASP	GLU	V1059	E852	F787	F714	P640	K569	A499	D438	K372	D307
ASP	GLU	P1060	L920	S788	E715	P640	P570	E500	M439	I308	I308
ALA	GLU	E1062	G921	K789	D716	V641	I571	E500	N440	A309	A309
GLN	SER	E1062	D922	I789	N717	C642	M672	Q603	P441	G310	G310
PHE	PHE	H851	S915	T789	V718	F646	S573	L504	V442	P312	P312
ASP	ASP	Q1052	G916	F794	R720	G647	G574	L504	L443	Q313	Q313
ASP	GLU	V1063	V916	E795	R720	G647	K575	V507	F444	A314	A314
ASP	GLU	V1064	S917	S796	F721	M648	Q576	I511	N445	L315	L315
ASP	GLU	V1065	L923	E796	L722	I649	I577	V512	R446	Q316	Q316
ASP	GLU	V1066	E932	G798	N723	K651	S579	S513	Q447	K317	K317
ASP	GLU	V1067	E933	F799	E724	G650	V580	P514	S449	I385	I385
ASP	GLU	V1068	R1001	R800	A725	M654	M584	Q515	L450	G319	G319
ASP	GLU	V1069	G1002	E801	D727	F655	I588	S516	H451	R320	R320
ASP	GLU	V1070	N1003	N802	K728	G655	L588	M517	K452	E321	E321
ASP	GLU	V1071	E1005	N802	A729	G655	L588	K518	M452	V322	V322
ASP	GLU	V1072	E1006	S803	R730	G655	Q589	P519	S454	K323	K323
ASP	GLU	V1073	I1007	L804	A729	G655	R590	C520	M455	S324	S324
ASP	GLU	V1074	Q1008	L805	R731	G655	R590	M521	M456	I325	I325
ASP	GLU	V1075	N1009	R806	G807	G661	F591	G522	A457	R326	R326
ASP	GLU	V1076	R940	E870	L732	F662	D592	G522	H458	A327	A327
ASP	GLU	V1077	F942	D871	N736	S663	T595	V524	R459	P396	P396
ASP	GLU	V1078	L943	T809	L737	T664	T596	Q525	V460	G331	G331
ASP	GLU	V1079	L943	F810	K738	G665	T596	D526	K461	H399	H399
ASP	GLU	V1080	V946	Q811	D739	I666	L597	B526	V462	P400	P400
ASP	GLU	V1081	F947	E812	L740	G667	L598	T527	V462	E333	E333
ASP	GLU	V1082	F947	F813	N741	G667	S599	L528	I463	G401	G401
ASP	GLU	V1083	N953	F814	N741	G667	P600	I528	I463	G334	G334
ASP	GLU	V1084	N954	F815	N741	G667	A671	I528	I463	G334	G334
ASP	GLU	V1085	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1086	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1087	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1088	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1089	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1090	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1091	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1092	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1093	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1094	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1095	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1096	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1097	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1098	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1099	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1100	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1101	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1102	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1103	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1104	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1105	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1106	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1107	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1108	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1109	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1110	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1111	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1112	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1113	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1114	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1115	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1116	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1117	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1118	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1119	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1120	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1121	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1122	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1123	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1124	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1125	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1126	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1127	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1128	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1129	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1130	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1131	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1132	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1133	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1134	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1135	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1136	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1137	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1138	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1139	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1140	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1141	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1142	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1143	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1144	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1145	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1146	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1147	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1148	F955	R816	N741	G667	D672	I528	I463	G334	G334
ASP	GLU	V1149									



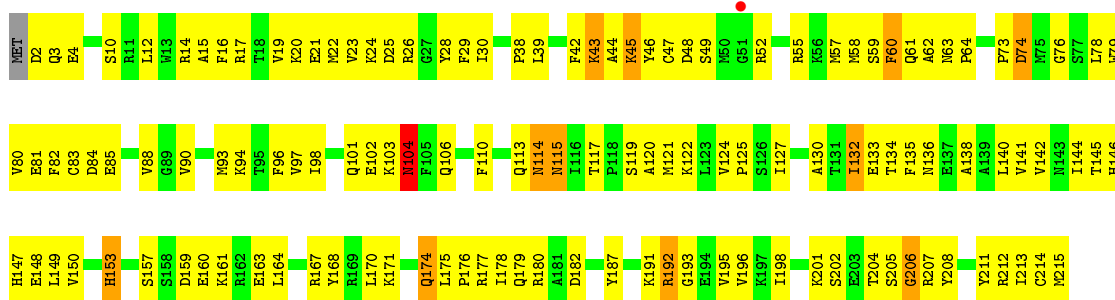
- Molecule 7: DNA-directed RNA polymerase II subunit RPB4

Chain D: 28% 38% 11% 20%



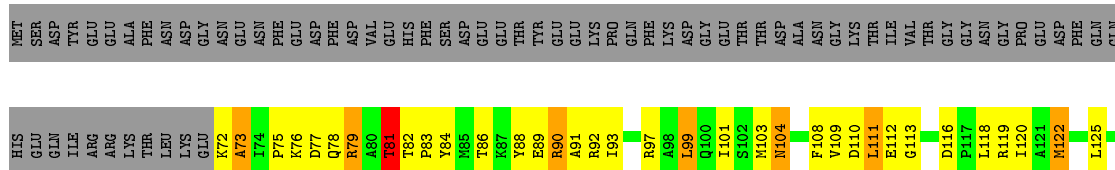
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC1

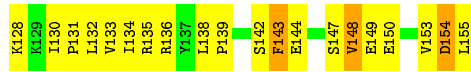
Chain E: 39% 55% 5%



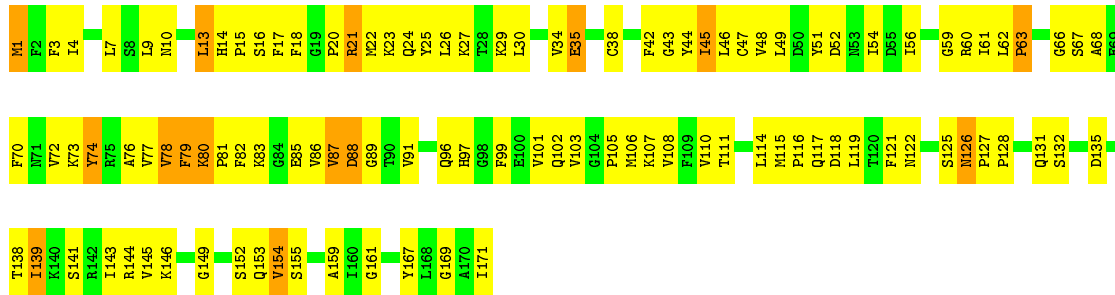
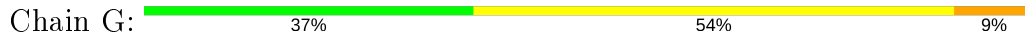
- Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 19% 28% 6% 46%

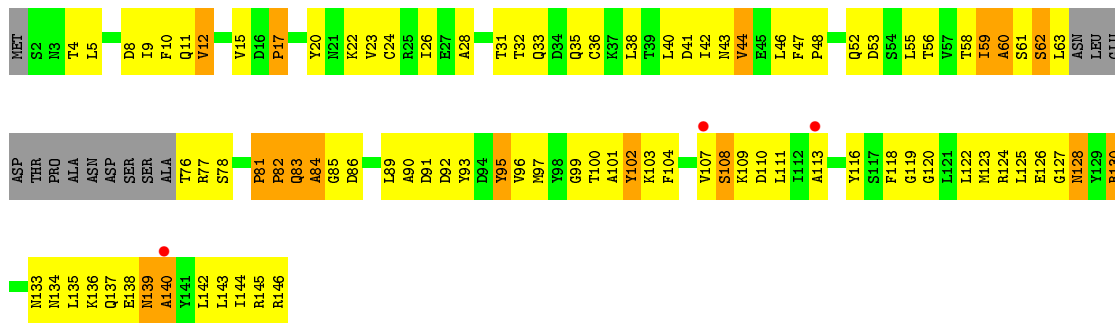




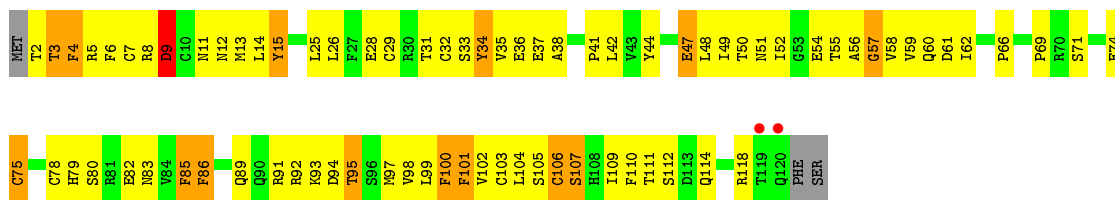
• Molecule 10: DNA-directed RNA polymerase II subunit RPB7



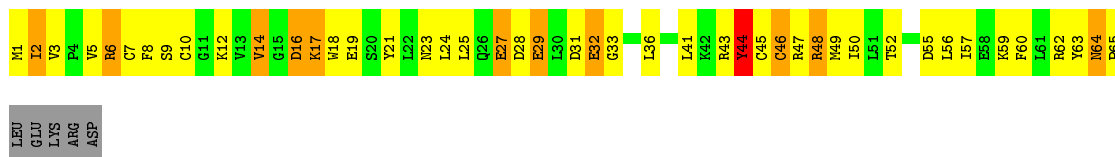
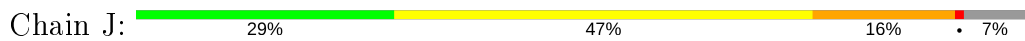
• Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC3



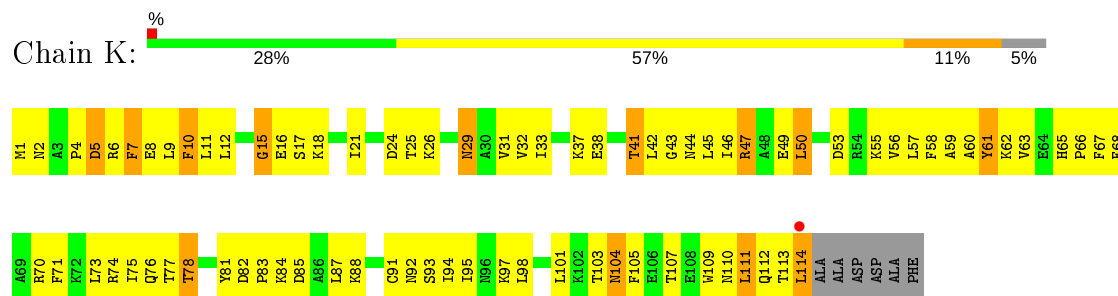
• Molecule 12: DNA-directed RNA polymerase II subunit RPB9



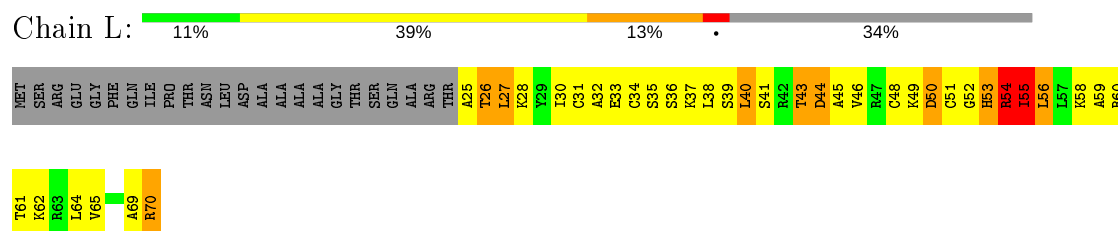
• Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC5



- Molecule 14: DNA-directed RNA polymerase II subunit RPB11



- Molecule 15: DNA-directed RNA polymerases I, II, and III subunit RPABC4



4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	222.06Å 393.12Å 283.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.80 49.14 – 3.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-3.80) 99.8 (49.14-3.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.95 (at 3.77Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.215 , 0.240 0.216 , 0.241	Depositor DCC
R_{free} test set	2410 reflections (1.98%)	wwPDB-VP
Wilson B-factor (Å ²)	106.6	Xtriage
Anisotropy	0.442	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 59.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.045 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.045 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	31804	wwPDB-VP
Average B, all atoms (Å ²)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.84% 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 i

5.1 Standard geometry i

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

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	T	1.41	5/373 (1.3%)	1.79	11/572 (1.9%)
2	N	1.38	0/161	1.10	0/247
3	P	0.95	0/242	0.98	0/377
4	A	0.43	0/11339	0.72	4/15334 (0.0%)
5	B	0.42	0/8981	0.68	0/12108
6	C	0.44	0/2133	0.72	0/2891
7	D	0.42	0/1437	0.69	1/1925 (0.1%)
8	E	0.41	0/1788	0.65	0/2406
9	F	0.49	0/691	0.77	0/933
10	G	0.45	0/1368	0.72	0/1844
11	H	0.38	0/1086	0.65	0/1470
12	I	0.36	0/989	0.65	0/1331
13	J	0.47	0/541	0.74	0/727
14	K	0.45	0/937	0.68	0/1265
15	L	0.47	0/366	0.71	0/485
All	All	0.46	5/32432 (0.0%)	0.73	16/43915 (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	T	0	2
2	N	0	2
6	C	0	1
13	J	0	1
All	All	0	6

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	T	16	DG	N3-C4	8.85	1.41	1.35
1	T	17	DG	N3-C4	7.68	1.40	1.35
1	T	17	DG	C2-N3	6.93	1.38	1.32
1	T	16	DG	N1-C2	5.71	1.42	1.37
1	T	14	DT	O3'-P	5.20	1.67	1.61

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	T	17	DG	O4'-C1'-N9	11.85	116.29	108.00
1	T	16	DG	O4'-C1'-N9	9.50	114.65	108.00
1	T	17	DG	N3-C2-N2	7.61	125.22	119.90
4	A	567	LYS	C-N-CD	6.05	141.11	128.40
1	T	16	DG	N1-C2-N3	-5.93	120.34	123.90

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	C	82	TYR	Sidechain
2	N	5	DT	Sidechain
2	N	6	DA	Sidechain
1	T	11	DT	Sidechain
1	T	12	DA	Sidechain

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	T	336	0	195	46	0
2	N	143	0	80	20	0
3	P	216	0	108	13	0
4	A	11140	0	11217	1294	0
5	B	8810	0	8847	1025	0
6	C	2095	0	2052	253	0
7	D	1427	0	1451	136	0
8	E	1752	0	1776	149	0
9	F	679	0	701	79	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	G	1340	0	1357	152	0
11	H	1068	0	1040	131	0
12	I	971	0	930	102	0
13	J	532	0	543	106	0
14	K	919	0	929	115	0
15	L	364	0	388	57	0
16	T	3	0	0	1	0
17	A	2	0	0	0	0
17	B	1	0	0	0	0
17	C	1	0	0	0	0
17	I	2	0	0	0	0
17	J	1	0	0	0	0
17	L	1	0	0	0	0
18	A	1	0	0	0	0
All	All	31804	0	31614	3347	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 53.

The worst 5 of 3347 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:H:100:THR:HG23	11:H:138:GLU:HA	1.26	1.16
10:G:138:THR:HG22	10:G:139:ILE:H	1.09	1.12
4:A:53:LEU:HD23	4:A:54:ASN:H	0.99	1.12
4:A:1094:VAL:HG13	4:A:1113:THR:HG21	1.32	1.11
5:B:510:LYS:HG3	5:B:511:PRO:HD3	1.12	1.11

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	
4	A	1406/1733 (81%)	988 (70%)	277 (20%)	141 (10%)	0	9
5	B	1090/1224 (89%)	754 (69%)	217 (20%)	119 (11%)	0	8
6	C	264/318 (83%)	181 (69%)	50 (19%)	33 (12%)	0	6
7	D	173/221 (78%)	125 (72%)	28 (16%)	20 (12%)	0	6
8	E	212/215 (99%)	157 (74%)	38 (18%)	17 (8%)	1	14
9	F	82/155 (53%)	62 (76%)	15 (18%)	5 (6%)	1	20
10	G	169/171 (99%)	131 (78%)	30 (18%)	8 (5%)	2	24
11	H	129/146 (88%)	90 (70%)	19 (15%)	20 (16%)	0	4
12	I	117/122 (96%)	84 (72%)	23 (20%)	10 (8%)	1	12
13	J	63/70 (90%)	36 (57%)	15 (24%)	12 (19%)	0	2
14	K	112/120 (93%)	86 (77%)	20 (18%)	6 (5%)	2	22
15	L	44/70 (63%)	17 (39%)	14 (32%)	13 (30%)	0	0
All	All	3861/4565 (85%)	2711 (70%)	746 (19%)	404 (10%)	0	8

5 of 404 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	42	ASP
4	A	44	THR
4	A	48	ALA
4	A	54	ASN
4	A	55	ASP

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	1239/1520 (82%)	1124 (91%)	115 (9%)	9	35
5	B	962/1061 (91%)	887 (92%)	75 (8%)	12	42
6	C	234/274 (85%)	211 (90%)	23 (10%)	8	33
7	D	159/200 (80%)	129 (81%)	30 (19%)	1	10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	E	196/197 (100%)	190 (97%)	6 (3%)	40 65
9	F	74/137 (54%)	65 (88%)	9 (12%)	5 25
10	G	152/152 (100%)	139 (91%)	13 (9%)	10 40
11	H	117/128 (91%)	112 (96%)	5 (4%)	29 58
12	I	113/116 (97%)	104 (92%)	9 (8%)	12 42
13	J	60/65 (92%)	56 (93%)	4 (7%)	16 47
14	K	99/102 (97%)	89 (90%)	10 (10%)	7 32
15	L	40/57 (70%)	36 (90%)	4 (10%)	7 32
All	All	3445/4009 (86%)	3142 (91%)	303 (9%)	10 38

5 of 303 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	463	THR
5	B	978	ASP
12	I	75	CYS
5	B	498	THR
5	B	737	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 95 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	734	HIS
5	B	1179	GLN
12	I	60	GLN
5	B	763	GLN
5	B	1015	HIS

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	P	9/10 (90%)	1 (11%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	P	3	U

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
16	CPT	T	67	1	0,2,4	0.00	-	-		

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	T	67	CPT	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

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

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	T	17/17 (100%)	-0.41	0 100 100	110, 139, 169, 175	0
2	N	7/7 (100%)	0.36	0 100 100	148, 156, 169, 171	0
3	P	10/10 (100%)	-0.28	0 100 100	118, 128, 165, 172	0
4	A	1416/1733 (81%)	-0.32	7 (0%) 91 87	44, 103, 161, 199	0
5	B	1108/1224 (90%)	-0.26	13 (1%) 79 72	47, 116, 176, 199	0
6	C	266/318 (83%)	-0.40	0 100 100	64, 100, 144, 162	0
7	D	177/221 (80%)	-0.19	1 (0%) 89 85	70, 120, 177, 190	0
8	E	214/215 (99%)	-0.24	1 (0%) 91 87	79, 144, 185, 197	0
9	F	84/155 (54%)	-0.54	0 100 100	53, 81, 115, 132	0
10	G	171/171 (100%)	-0.31	0 100 100	82, 103, 141, 150	0
11	H	133/146 (91%)	0.15	3 (2%) 60 52	119, 149, 180, 188	0
12	I	119/122 (97%)	-0.10	2 (1%) 70 62	94, 150, 183, 199	0
13	J	65/70 (92%)	-0.60	0 100 100	71, 96, 131, 140	0
14	K	114/120 (95%)	-0.40	1 (0%) 84 79	64, 101, 130, 149	0
15	L	46/70 (65%)	-0.11	0 100 100	96, 154, 172, 178	0
All	All	3947/4599 (85%)	-0.28	28 (0%) 87 83	44, 111, 175, 199	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	1176	LEU	10.4
4	A	1175	SER	5.1
5	B	471	LYS	4.9
4	A	1455	PRO	3.7
5	B	470	LYS	3.7

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 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	CPT	T	67	3/5	0.96	0.23	131,131,131,132	0
18	MG	A	1736	1/1	0.97	0.21	58,58,58,58	0
17	ZN	A	1734	1/1	0.98	0.09	108,108,108,108	0
17	ZN	I	204	1/1	0.99	0.07	196,196,196,196	0
17	ZN	C	319	1/1	0.99	0.14	60,60,60,60	0
17	ZN	L	105	1/1	0.99	0.10	127,127,127,127	0
17	ZN	J	101	1/1	0.99	0.22	68,68,68,68	0
17	ZN	B	1307	1/1	0.99	0.19	71,71,71,71	0
17	ZN	I	203	1/1	0.99	0.14	113,113,113,113	0
17	ZN	A	1735	1/1	1.00	0.13	65,65,65,65	0

6.5 Other polymers [i](#)

There are no such residues in this entry.