



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 04:50 AM GMT

PDB ID : 2O5J
Title : Crystal structure of the T. thermophilus RNAP polymerase elongation complex with the NTP substrate analog
Authors : Vassilyev, D.G.; Vassilyeva, M.N.
Deposited on : 2006-12-06
Resolution : 3.00 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

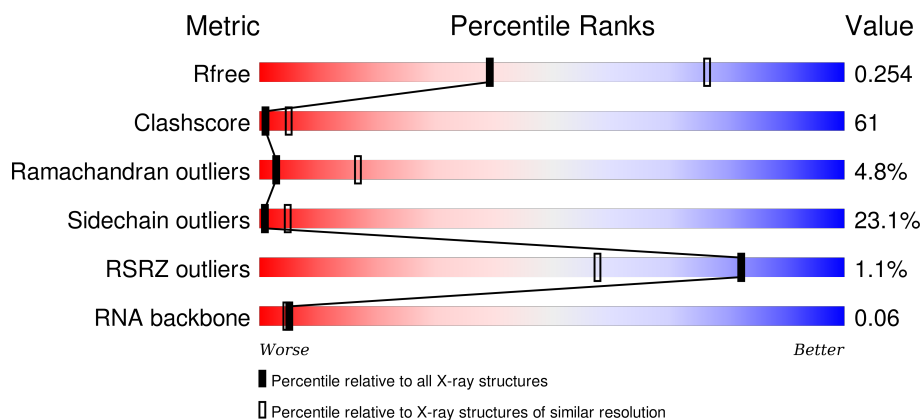
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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.



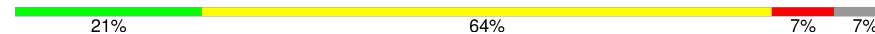


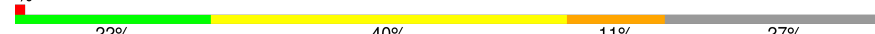




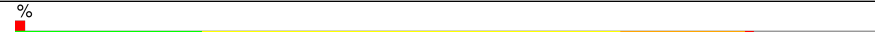

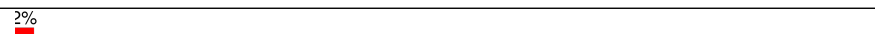

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)
RNA backbone	2183	1036 (3.40-2.60)

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. 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	G	23	<div> <div>17%</div> <div>57%</div> <div>22%</div> <div>.</div> </div>
1	X	23	<div> <div>17%</div> <div>61%</div> <div>13%</div> <div>9%</div> </div>
2	H	16	<div> <div>19%</div> <div>81%</div> </div>
2	Y	16	<div> <div>31%</div> <div>69%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	I	14	
3	Z	14	
4	A	315	
4	B	315	
4	K	315	
4	L	315	
5	C	1119	
5	M	1119	
6	D	1524	
6	N	1524	
7	E	99	
7	O	99	

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 51213 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(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*GP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	23	Total	C	N	O	P	0	0	0
			467	220	80	144	23			
1	X	23	Total	C	N	O	P	0	0	0
			467	220	80	144	23			

- Molecule 2 is a RNA chain called 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	16	Total	C	N	O	P	0	0	0
			347	153	64	114	16			
2	Y	16	Total	C	N	O	P	0	0	0
			347	153	64	114	16			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	13	Total	C	N	O	P	0	0	0
			270	126	57	74	13			
3	Z	13	Total	C	N	O	P	0	0	0
			270	126	57	74	13			

- Molecule 4 is a protein called DNA-directed RNA polymerase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
4	B	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	K	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
4	L	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			

- Molecule 5 is a protein called DNA-directed RNA polymerase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			
5	M	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			

- Molecule 6 is a protein called DNA-directed RNA polymerase beta' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	1264	Total	C	N	O	S	0	0	0
			9960	6302	1773	1852	33			
6	N	1264	Total	C	N	O	S	0	0	0
			9960	6302	1773	1852	33			

- Molecule 7 is a protein called DNA-directed RNA polymerase omega chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	95	Total	C	N	O	S	0	0	0
			770	491	133	142	4			
7	O	95	Total	C	N	O	S	0	0	0
			770	491	133	142	4			

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	2	Total	Zn	0	0
			2	2		
8	N	2	Total	Zn	0	0
			2	2		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

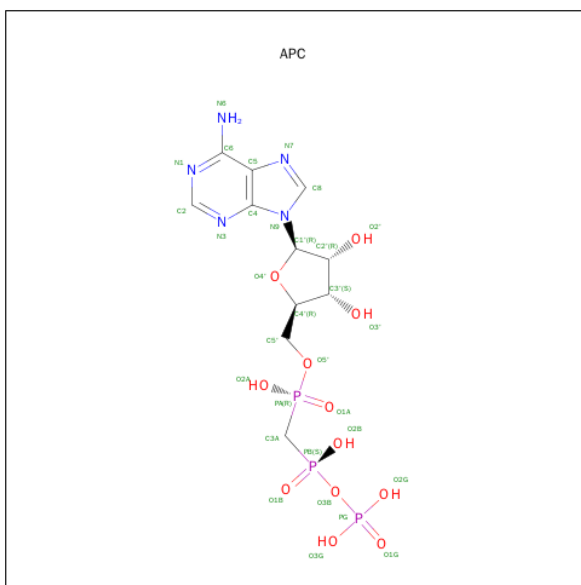
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Mg	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	N	2	Total	Mg	0	0
			2	2		

- Molecule 10 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
10	D	1	Total	C	N	O	P	0	0
			31	11	5	12	3		
10	N	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	106	Total	O	0	0
			106	106		
11	B	82	Total	O	0	0
			82	82		
11	C	482	Total	O	0	0
			482	482		
11	D	506	Total	O	0	0
			506	506		
11	E	60	Total	O	0	0
			60	60		
11	G	32	Total	O	0	0
			32	32		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	H	37	Total 37	O 37	0	0
11	I	22	Total 22	O 22	0	0
11	K	86	Total 86	O 86	0	0
11	L	104	Total 104	O 104	0	0
11	M	483	Total 483	O 483	0	0
11	N	491	Total 491	O 491	0	0
11	O	39	Total 39	O 39	0	0
11	X	43	Total 43	O 43	0	0
11	Y	30	Total 30	O 30	0	0
11	Z	30	Total 30	O 30	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 errors 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(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3'

Chain G: 



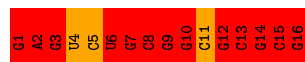
- Molecule 1: 5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3'

Chain X: 



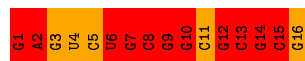
- Molecule 2: 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3',

Chain H: 



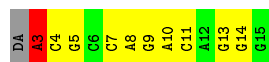
- Molecule 2: 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3',

Chain Y: 



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

Chain I: 

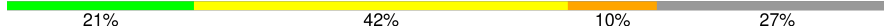


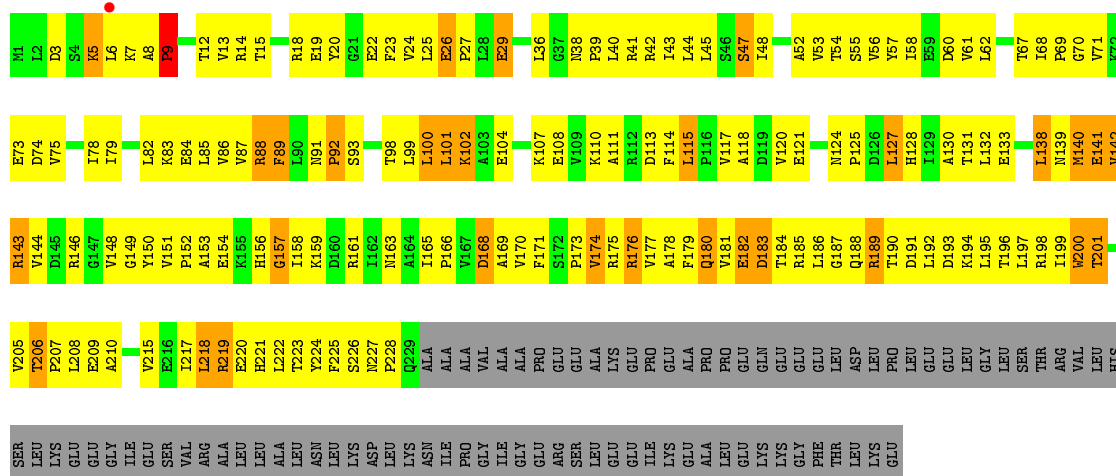
- Molecule 3: 5'-D(*AP*AP*CP*GP*CP*CP*AP*GP*AP*CP*AP*GP*GP*G)-3'

Chain Z: 




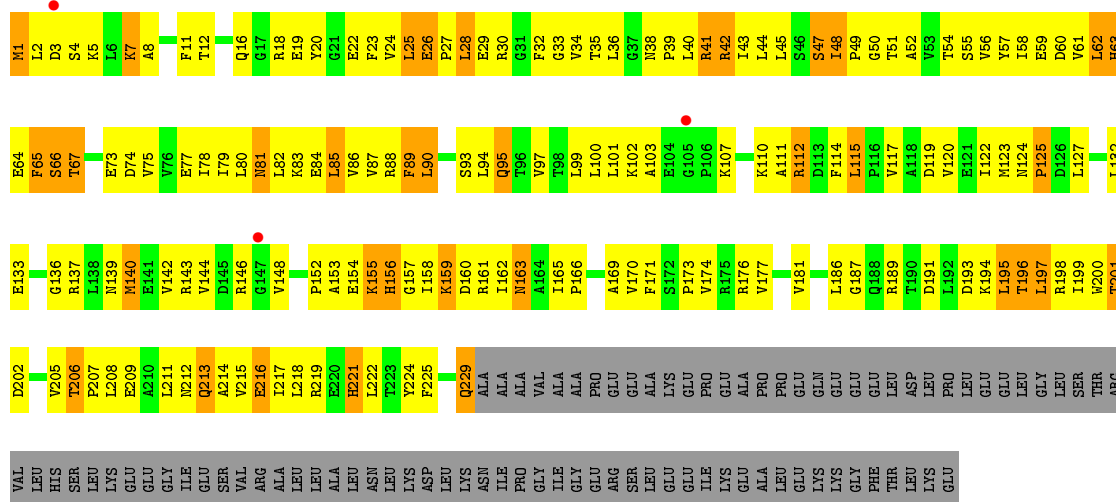
- Molecule 4: DNA-directed RNA polymerase alpha chain

Chain A: 




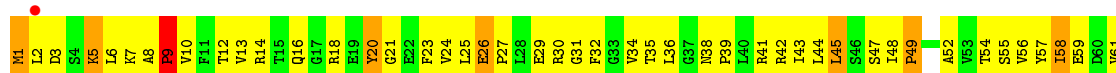
- Molecule 4: DNA-directed RNA polymerase alpha chain

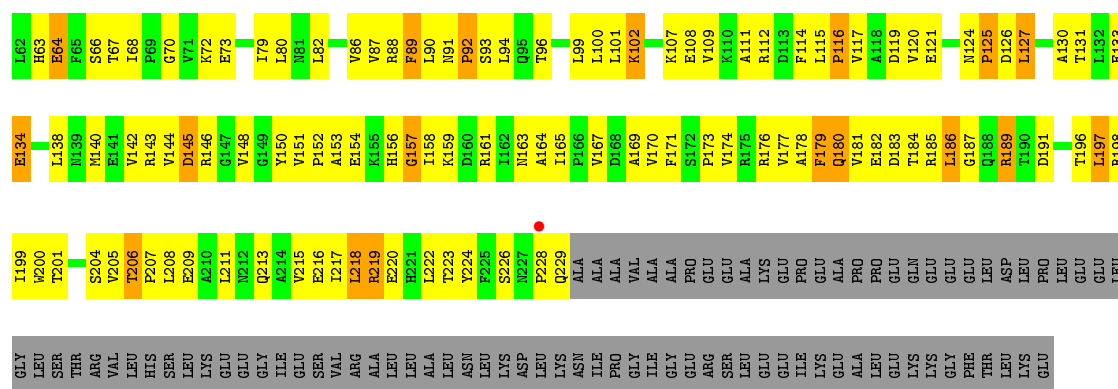
Chain B: 



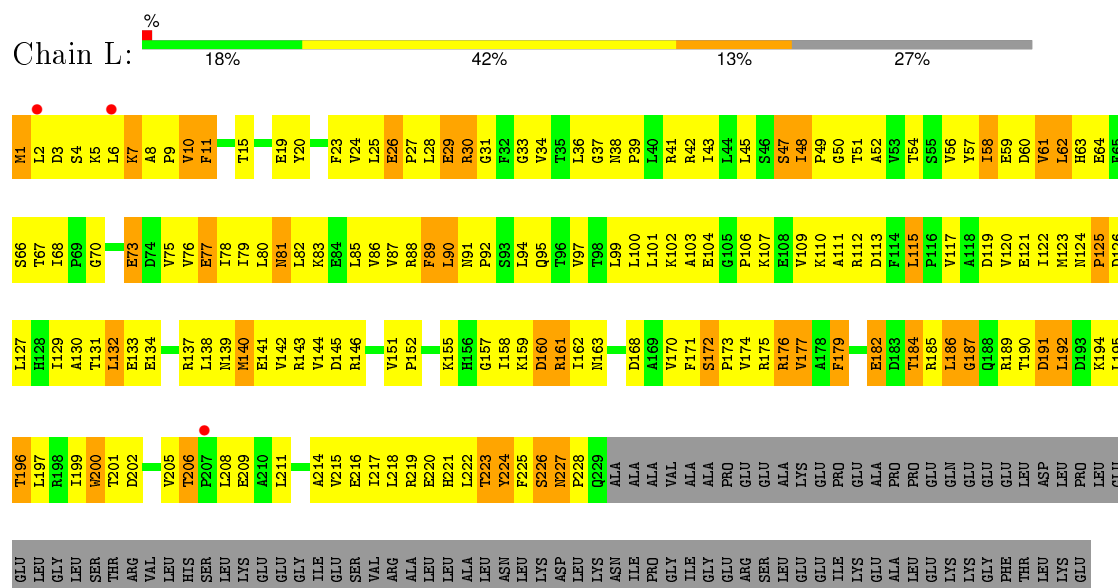
- Molecule 4: DNA-directed RNA polymerase alpha chain

Chain K: 

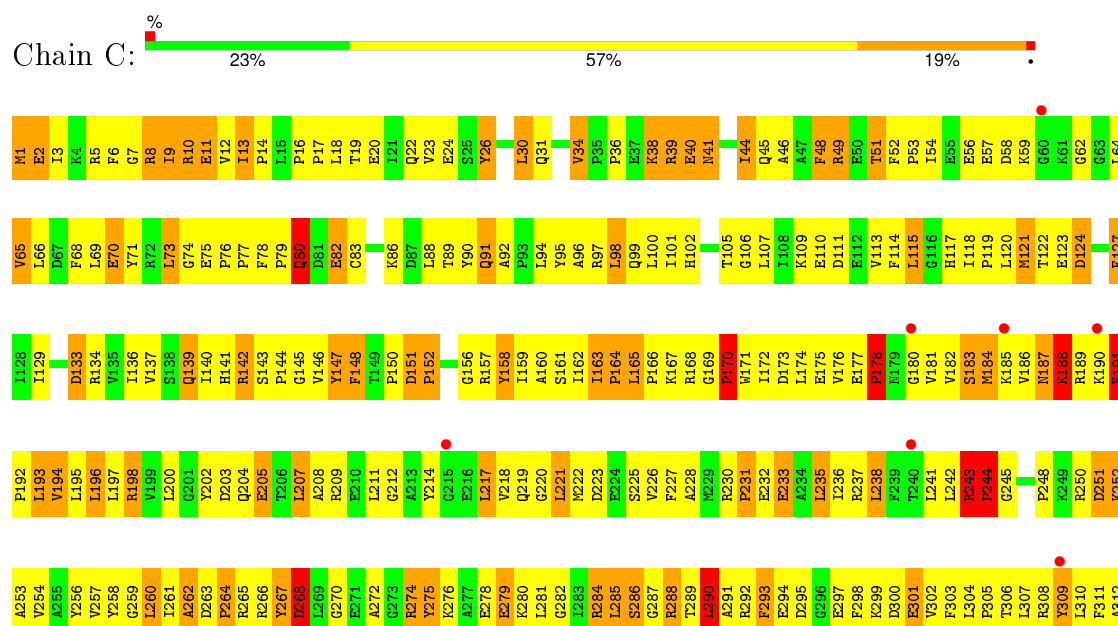


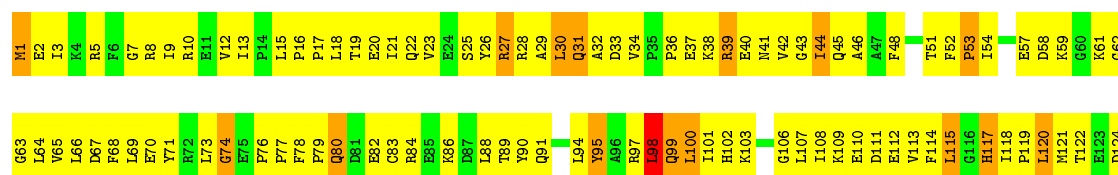


• Molecule 4: DNA-directed RNA polymerase alpha chain



• Molecule 5: DNA-directed RNA polymerase beta chain

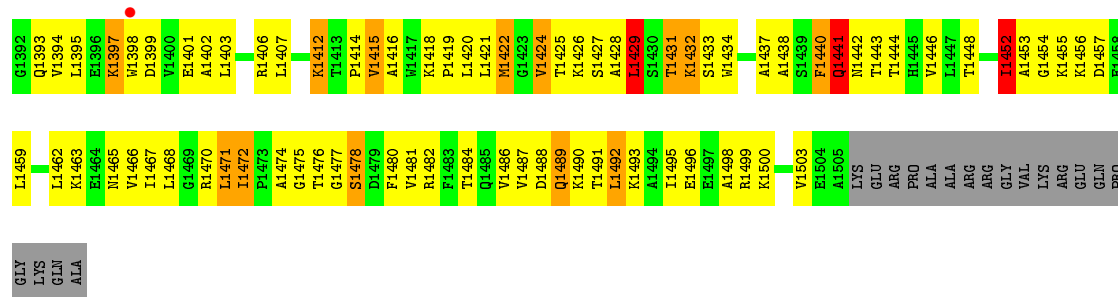




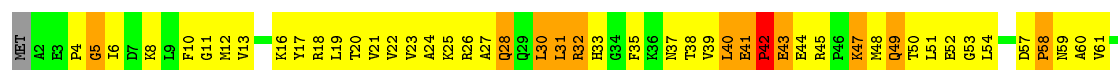
A1096	F1032	K971	Y001	L839	S776	E711	O647	G582	R518	I452	A381	P318	A253	V186	G125
L1097	G1033	V972	I902	A840	I777	A712	R648	L583	G519	T453	A382	G319	V254	H187	S126
D1098	E1034	S903	I903	H841	F778	R713	V649	E584	P521	T454	R383	E321	A255	R188	F127
T1101	M1035	L974	P904	R842	G779	D714	R650	E585	P522	G545	R384	H322	V256	R189	I128
L1102	E1036	Y975	I905	H843	K781	K716	K651	E586	V523	L455	E384	V322	V257	F191	N129
D1103	V1037	Y976	F906	G844	K782	K717	G652	V587	I523	A456	F385	D323	Y258	F192	N130
E1104	G1038	G977	D907	H845	A782	L717	D653	V588	V524	A457	F386	D324	G259	L193	G131
K1105	A1039	R978	G908	K846	R783	G718	D654	R589	S525	Y458	S387	I325	L261	V194	A132
D1106	L1040	T979	A909	R847	R784	E719	L655	P526	P526	A459	R388	D326	A262	L195	D133
N1107	Y1043	G980	K910	V848	D785	E720	A656	E527	E527	R460	S392	H327	A263	L196	V135
P1108	G1044	E981	E911	H849	K786	R721	D657	A594	V529	D462	F393	G329	D263	L197	I136
V1109	A1045	K851	P912	A850	D787	I726	G658	L595	P530	E463	F394	N330	R265	R198	V137
D1110	A1046	K851	E913	H851	T788	P727	P659	Y596	P531	L464	K395	R331	R266	V199	S138
I1111	H1047	R985	I914	L852	S789	F727	A660	A597	F581	G465	D396	R332	Y267	L200	Q139
F1112	P986	K915	E916	L853	R791	R728	S661	E598	M532	F466	E397	I333	D268	G270	H141
E1113	T1048	R987	R917	H855	V792	S730	N663	D600	V534	I467	T398	R334	L269	D203	R142
G1114	L1049	V988	L918	G856	F793	E731	G664	G601	S535	R468	T399	T335	G271	Q204	R143
L1115	Q1050	R988	E919	F793	F794	A732	F685	E602	P536	T469	P400	V336	E271	E205	S143
R1119	E1051	V989	A919	F794	F794	A733	L666	E603	K537	P470	L401	G337	A272	T236	P144
	T1054	G990	Q920	M858	G798	A734	L667	V603	Q538	Y471	A402	G273	G272	L207	G145
	K1055	Q991	F926	P859	I799	L734	A657	A604	Q538	R472	S403	G273	R274	L207	P146
	L1056	M992	G927	H860	V900	R735	L668	K605	V539	R473	L404	M340	Y275	A208	Y147
	K1057	F993	K928	L861	V800	D736	G669	V606	F540	T487	P415	R289	E224	R209	F148
	S1057	K994	K929	P863	R801	L737	G670	D607	S541	T488	L418	R353	S225	E210	F149
	D1058	M995	R929	D863	R802	D738	M671	G608	V542	T489	T419	G354	Y226	L211	T149
	K1059	K996	K930	G864	T803	E739	V672	M609	M543	E490	R493	V355	R292	G212	P150
	I1060	L997	G931	T865	V804	E740	L673	R610	L546	T480	R408	R345	R287	E216	P151
	E1061	Y998	E932	P866	R805	G741	V674	I611	J547	V483	R409	R345	G282	E216	P152
	K1062	H999	G932	V867	L806	V742	A675	V612	P548	Y483	R410	G346	L283	L217	A153
	R1063	M1000	V936	D868	R807	V743	L676	V613	F549	Y484	S411	G347	R284	V218	R154
	N1064	V1001	D937	V869	R808	R744	M677	R614	L550	Y485	A412	L348	L285	G155	P155
	A1065	E1002	K938	L870	G809	I745	P678	Y615	E551	M486	G414	A349	S286	L221	G156
	A1066	D1003	R939	L871	D810	E748	F679	E616	H552	T487	P415	R350	R288	N222	G157
	Y1067	K1004	E940	P873	G812	V749	D680	D617	D583	T488	L418	A352	T289	D223	Y158
	I1070	H1006	E942	L874	V813	K750	F684	R619	D584	T489	L418	R353	L290	E224	A160
	I1071	A1007	V943	G875	E814	P751	E685	L620	B557	E490	T419	G354	A291	S225	S161
	E1074	K1008	R946	V876	L815	G752	D686	V621	B557	R493	R422	V355	R292	F227	I162
	D1075	S1009	A947	P877	K816	D753	A637	E622	M560	Y494	A423	R358	F293	R230	I163
	V1076	T1010	G947	S878	P817	I754	L688	Y623	G561	T495	G424	M359	E294	E294	P164
	P1077	G1011	E948	R879	G818	L755	V689	P624	O561	I496	R432	L360	D295	P231	L165
	E1078	P1012	K949	M880	V819	V756	L690	L625	M564	A497	D426	N361	G296	E232	E175
	P1079	S1013	N881	R820	R820	G757	S691	R626	O564	Q498	V427	G362	G296	E233	P166
	G1080	E1014	V953	L882	E821	R758	E692	R627	O565	Q498	V427	G362	G296	E233	P166
	V1081	L1015	T954	G853	V822	T759	E693	F628	T566	P502	V430	S363	D300	A284	G169
	P1082	I1016	P955	K884	V823	K762	L694	Y629	Q567	P502	V430	E364	E301	L285	P170
	E1083	T1017	G956	L885	R824	G763	L695	R630	A568	L503	H431	D365	F303	F239	I172
	S1084	Q1018	K957	L886	V827	E764	K696	S631	P569	N506	T433	L367	F303	T240	D173
	F1085	Q1019	T958	E887	V827	S765	R697	N632	P570	N506	T433	L367	L304	L241	L174
	R1086	P1020	P959	T888	A828	S765	D698	Q633	L571	R507	H434	T368	P305	L242	E175
	V1087	L1021	E960	H889	Q829	E766	F699	Q633	L572	I508	Y435	P389	T306	R243	V176
	L1088	G1022	E961	R891	K830	T767	V700	L637	R573	A509	G436	A370	L307	E177	E177
	V1089	K1023	Q962	L892	R831	T768	T701	R640	A574	A510	R437	K371	R308	G245	P178
	K1090	E1025	L963	A893	K832	P769	S702	R640	Q575	E511	L438	L372	Y309	D246	N179
	E1091	Q1026	P966	G894	L833	E770	H704	R642	A576	R512	G439	V373	L310	P247	V181
	L1092	F1027	P967	L897	Q834	R772	L705	V843	V578	V514	V441	N374	F311	P248	V181
	Q1093	Q1030	L968	G898	V835	R773	E706	V644	V579	V514	V441	S375	A312	K249	V182
	A1094	Q999	Q969	G836	G836	L773	E706	V644	V579	A515	E442	R376	L313	R250	S183
	L1095	R1031	G970	D837	D837	R773	E706	V644	V579	A515	E442	R376	L313	R250	S183
				K838	K838	R775	I710	G646	T581	R517	P444	L378	T314	K252	K185

Chain D: ■ 21% ■ 47% ■ 14% ■ 17%

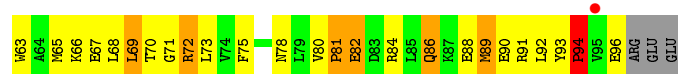
GLY	P1268	G1204	I1140	Y1015	T948	L881	A814	P750	P689	Y625	A562	Y499	L439	PRO
A1329	K1269	Y1205	E1141	P1016	I949	F882	A815	L751	A690	S625	P563	R500	Y440	GLU
I1330	A1270	G1206	E1142	F1017	G950	A883	B816	F754	L691	G627	S564	R501	R441	GLU
D1331	K1271	G1143	G1143		I951	A884	E817		E692	R628	R564	F502	N442	GLU
	ALA	D1208	L1144	L1020	D952	I885	R818	A756	E693	S629	I566	L503	V443	VAL
Q1334	VAL	L1409	Y1021	Y1021	D953	I886	R819	Q756	G694	V630	I567	G506	V444	ILE
L1335	ILE	S1210	G1092	G1092	A954	A887	E820	A757	L695	I631	R568	R507	R445	ALA
L1336	SER	M1211	M1023	M1023	R955	E888	B821	E758	R696	V632	R569	R508	R446	GLU
L1337	GLU	A1085	A1024	A1024	I956	A889	A822	A759	V699	G634	R570	P509	V447	ALA
A1338	ILE	R1087	L1086	L1086	P957	A890	L823	R761	V700	G635	K571	E510	S449	GLU
R1339	ASP	P1214	A1150	Q1026	E958	R824	R824	R760	R700	R636	K572	P511	S448	GLY
G1340	GLY	R1151	G1027	G1027	E959	K894	A825	Q762	L702	I637	M573	M511	Y450	VAL
P1341	VAL	L1152	A1028	A1028	K960	K894	R826	M763	L703	L637	L574	M512	D451	VAL
E1342	VAL	V1153	R1029	R1029	K961	V895	I827	L764	N703	R538	Q575	I513	I452	HIS
A1343	ARG	E1154	G1030	G1030	Q962	A896	R828	S765	R704	L639	E576	L514	D453	LEU
V1344	ILE	V1155	N1031	N1031	Q963	R897	R829		A705	H640	A577	E515	D454	HIS
E1345	GLU	L1156	G1093	G1093	L964	E998	A830	F772	T706	Q641	V578	A516	R455	GLU
R1346	GLU	G1157	Q1032	Q1032	E965	L899	G831	L769	T707	G542	D579	V517	R456	PRO
Y1347	THR	V1158	Q1034	Q1034	E966	I900	R832	L770	L708	G643	A580	P518	G457	
L1348	GLU	R1159	I1035	I1035	A967	Q901	E833	S771	H709	L644	L581	V519	A458	
E1350	GLU	L1160	R1036	R1036	D968	L902	T834	P772	R710	P645	L582	L520	E459	
E1351	LYS	E1161	Q1037	Q1037	R969	D903	S835	A773	L711	K646	D583	P521	A460	
I1352	LEU	E1162	L1038	L1038	K970	V904	R836	S774	G712	R647	N584	P522	I461	
K1353	SER	G1163	G1039	G1039	L971	P905	G837	G775	L713	M648	G585	D523	Q462	
K1354	VAL	R1164	L1041	L1041	L972	Q906	R838	E776	G714		R586	L524	Q463	
V1355	PHE	Y1165	R1042	R1042	Q973	E907	L839	P777	A715		R587	R525	L464	
V1356	VAL	L1166	G1043	G1043	I974	K908	R840	L778	F716	L652	G988	P526	L465	
R1357	GLU	T1167	L1044	L1044	E975		Y841	A779	F717	G653	A589	N527	K466	
	SER	L1168	M1045	M1045	Q876			L780	P718	K654	P590	V528	E467	
V1361	GLY	D1170	Q1046	Q1046	A977			R781	V719	G529	P584	Q529	P402	
K1362	PHE	V1171	K1047	K1047	F982	L914	R846	S782	R721	L657	G585	D531	D469	
K1363	SER	R1172	P1048	P1048	L983	V915	D847	R783	G723	L588	S596	G532	E471	
K1364	LYS	L1174	S1049	S1049	T984	Q917	E848	I785	G723	K659	D597	G533	A472	
D1365	GLU	C1112	G1050	G1050	D985	Q917		R786	Q724	R598	R598	R534	L473	
K1366	TYR	H1113	E1051	E1051	E987	A918	L851	L787	S725	N661	P599		E474	
H1367	LYS	T1114	T1052	T1052	E987	F919	A852	G788	L726		L600	T537	K475	
I1368	LEU	T1115	F1061	F1061	R988	L820	V853	L789	Q727		R601	S538	E476	
E1369	PRO	N1116	E1054	E1054	I989	R921		T790	L728	I666	S602	S538	L477	
I1370	LYS	Y1117	V1055	V1055	D990	L922	T857	Y791	H729	A667	L603	D539	L478	
V1371	GLU	I1118	P1056	P1056	Q991	G923	W858	I792	P730	P668	L603	L540	R414	
V1372	ALA	E1185	V1057	V1057	I992	M924	D859	T932	L731	N669	T604	N541	V415	
R1373	ARG	V1186	R1058	R1058	L993	E925	L860	Q794	L732	V670	D605	D542	E480	
Q1374	LEU	P1187	S1059	S1059	Q994	K926	Q861	V795	G733	K671	L606	L543	M481	
M1375	LEU	V1188	S1060	S1060	L995		D862	R796	E734	A672	L607	Y544	K482	
M1376	VAL	F1123	F1061	F1061	W996		R863	R797	A735	R674		R545	R483	
L1377	LYS	Q1124	R1062	R1062	T997	L830	W864	E798	F736	R675	Q611	L547	S485	
K1378	ASP	P1125	E1063	E1063	E998			R799	T737	A676	G612	T548	S485	
V1379	GLY	L1126	G1064	G1064	T999	L834	Y868	K800	A738	L677	R613	N549	R486	
E1380	ASP	T1193	L1065	L1065	T1000	K935			D739	E878	F614	R550	R488	
V1381	TYR	G1194	T1066	T1066	E1001				F740	R679	R615	N551	R489	
T1382	VAL	Q1195	V1067	V1067	K1002	Y937	R872	L804	D741	Q880	Q616	R552	A490	
D1383	GLU	T1196	L1068	L1068	V1003	G938	L873	E905	G742	R681	N617	R553	V431	
P1384	ALA	R1197	E1069	E1069	F939		E874	R606	D743	D682	L618	L554	Y432	
G1385	GLY	L1134	Y1070	Y1070	V1007	T943	T875	A807	G744	I683	L619		R493	
	GLN	R1135	F1071	F1071	F1008		S876	T808	K745	K684	G620	L557	K494	
R1388	PRO	K1136	I1072	I1072	K1009	T944	P877	P809	A746	D685	K621	L558	R495	
L1389	LEU	R1137	S1074	S1074	N1010	S945	G878	E910	V747	E586	R622	A559	E436	
L1390	THR	K1202	G946	G946	F1011		H879	V637	H748	V637	V623	Q560	L496	
E1391	ARG	K1203	H1075	H1075		I947	T890	L813	V749	W688	D624	G561	V498	



- Molecule 7: DNA-directed RNA polymerase omega chain



- Molecule 7: DNA-directed RNA polymerase omega chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	152.34Å 152.34Å 524.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.00 38.29 – 3.00	Depositor EDS
% Data completeness (in resolution range)	89.3 (40.00-3.00) 83.5 (38.29-3.00)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.51 (at 3.01Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.225 , 0.257 0.229 , 0.254	Depositor DCC
R_{free} test set	11219 reflections (5.70%)	DCC
Wilson B-factor (Å ²)	64.2	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 170.4	EDS
Estimated twinning fraction	0.146 for h,-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 196921 reflections	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	51213	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: APC, ZN, 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	G	1.04	1/520 (0.2%)	1.12	1/798 (0.1%)
1	X	1.12	2/520 (0.4%)	1.14	1/798 (0.1%)
2	H	1.39	3/387 (0.8%)	2.45	39/601 (6.5%)
2	Y	1.36	3/387 (0.8%)	2.44	33/601 (5.5%)
3	I	0.72	0/304	0.92	1/467 (0.2%)
3	Z	0.73	0/304	0.91	0/467
4	A	0.69	0/1838	0.76	0/2498
4	B	0.76	0/1838	0.76	2/2498 (0.1%)
4	K	0.73	0/1838	0.82	3/2498 (0.1%)
4	L	0.73	0/1838	0.78	4/2498 (0.2%)
5	C	0.79	1/8997 (0.0%)	0.89	17/12164 (0.1%)
5	M	0.78	1/8997 (0.0%)	0.90	17/12164 (0.1%)
6	D	0.79	1/10128 (0.0%)	0.91	18/13681 (0.1%)
6	N	0.79	2/10128 (0.0%)	0.89	22/13681 (0.2%)
7	E	0.83	1/784 (0.1%)	1.07	3/1057 (0.3%)
7	O	0.78	1/784 (0.1%)	1.07	3/1057 (0.3%)
All	All	0.80	16/49592 (0.0%)	0.95	164/67528 (0.2%)

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	G	0	6
1	X	0	5
2	H	0	2
2	Y	0	1
3	I	0	1
All	All	0	15

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	1	G	OP3-P	10.56	1.73	1.61
1	G	1	DC	OP3-P	-7.74	1.51	1.61
1	X	1	DC	OP3-P	-7.13	1.52	1.61
7	E	94	PRO	N-CA	6.34	1.58	1.47
5	C	439	CYS	CB-SG	-6.08	1.72	1.82

The worst 5 of 164 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Y	2	A	N9-C1'-C2'	-18.09	90.48	114.00
5	M	409	ARG	NE-CZ-NH1	15.09	127.85	120.30
2	Y	7	G	N9-C1'-C2'	-12.29	98.02	114.00
7	E	94	PRO	CA-N-CD	-11.28	95.71	111.50
2	H	1	G	N9-C1'-C2'	11.20	128.56	114.00

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	13	DT	Sidechain
1	G	15	DC	Sidechain
1	G	16	DG	Sidechain
1	G	17	DC	Sidechain
1	G	18	DG	Sidechain

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	467	0	259	43	0
1	X	467	0	259	43	0
2	H	347	0	174	58	0
2	Y	347	0	174	81	0
3	I	270	0	144	14	0
3	Z	270	0	144	12	0
4	A	1806	0	1861	169	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1806	0	1861	174	0
4	K	1806	0	1861	182	0
4	L	1806	0	1861	199	0
5	C	8829	0	8933	1208	0
5	M	8829	0	8933	1204	0
6	D	9960	0	10183	1379	0
6	N	9960	0	10183	1351	0
7	E	770	0	784	108	0
7	O	770	0	784	101	0
8	D	2	0	0	0	0
8	N	2	0	0	0	0
9	D	2	0	0	0	0
9	N	2	0	0	0	0
10	D	31	0	14	2	0
10	N	31	0	14	2	0
11	A	106	0	0	16	0
11	B	82	0	0	21	0
11	C	482	0	0	120	0
11	D	506	0	0	138	0
11	E	60	0	0	6	0
11	G	32	0	0	3	0
11	H	37	0	0	3	0
11	I	22	0	0	3	0
11	K	86	0	0	19	0
11	L	104	0	0	23	0
11	M	483	0	0	129	0
11	N	491	0	0	115	0
11	O	39	0	0	6	0
11	X	43	0	0	4	0
11	Y	30	0	0	6	0
11	Z	30	0	0	4	0
All	All	51213	0	48426	5871	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 61.

The worst 5 of 5871 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:Y:7:G:N1	5:M:1014:SER:HA	1.62	1.13
2:Y:16:G:H21	6:N:705:ALA:HB1	1.11	1.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:N:18:ILE:HG23	6:N:518:PRO:HG3	1.33	1.10
5:C:409:ARG:HA	5:C:454:SER:HA	1.27	1.10
6:N:1189:ARG:HB3	6:N:1204:CYS:HA	1.34	1.09

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	227/315 (72%)	206 (91%)	14 (6%)	7 (3%)	5	28
4	B	227/315 (72%)	206 (91%)	15 (7%)	6 (3%)	7	33
4	K	227/315 (72%)	208 (92%)	13 (6%)	6 (3%)	7	33
4	L	227/315 (72%)	206 (91%)	15 (7%)	6 (3%)	7	33
5	C	1117/1119 (100%)	919 (82%)	136 (12%)	62 (6%)	2	13
5	M	1117/1119 (100%)	923 (83%)	133 (12%)	61 (6%)	2	13
6	D	1258/1524 (82%)	1051 (84%)	149 (12%)	58 (5%)	3	18
6	N	1258/1524 (82%)	1058 (84%)	140 (11%)	60 (5%)	3	17
7	E	93/99 (94%)	76 (82%)	11 (12%)	6 (6%)	1	8
7	O	93/99 (94%)	74 (80%)	12 (13%)	7 (8%)	1	6
All	All	5844/6744 (87%)	4927 (84%)	638 (11%)	279 (5%)	3	17

5 of 279 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	29	GLU
4	A	187	GLY
4	B	29	GLU
4	B	187	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	C	178	PRO

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	202/273 (74%)	162 (80%)	40 (20%)	1	8
4	B	202/273 (74%)	159 (79%)	43 (21%)	1	6
4	K	202/273 (74%)	162 (80%)	40 (20%)	1	8
4	L	202/273 (74%)	150 (74%)	52 (26%)	0	3
5	C	941/941 (100%)	704 (75%)	237 (25%)	1	3
5	M	941/941 (100%)	713 (76%)	228 (24%)	1	4
6	D	1063/1279 (83%)	825 (78%)	238 (22%)	1	5
6	N	1063/1279 (83%)	833 (78%)	230 (22%)	1	6
7	E	84/88 (96%)	59 (70%)	25 (30%)	0	2
7	O	84/88 (96%)	68 (81%)	16 (19%)	2	10
All	All	4984/5708 (87%)	3835 (77%)	1149 (23%)	1	5

5 of 1149 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	D	1256	LEU
4	L	67	THR
6	N	1042	ARG
6	D	1365	ASP
4	K	1	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 154 such sidechains are listed below:

Mol	Chain	Res	Type
6	D	1103	HIS
4	K	156	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	N	906	GLN
6	D	1227	GLN
7	E	29	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	H	16/16 (100%)	10 (62%)	8 (50%)
2	Y	16/16 (100%)	10 (62%)	8 (50%)
All	All	32/32 (100%)	20 (62%)	16 (50%)

5 of 20 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	H	2	A
2	H	3	G
2	H	6	U
2	H	7	G
2	H	8	C

5 of 16 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	H	15	C
2	Y	1	G
2	Y	9	G
2	H	13	C
2	Y	12	G

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

Of 10 ligands modelled in this entry, 8 are monoatomic - leaving 2 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
10	APC	D	3999	9	25,33,33	1.41	3 (12%)	30,52,52	1.99	7 (23%)
10	APC	N	4999	9	25,33,33	1.30	4 (16%)	30,52,52	2.11	7 (23%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
10	APC	D	3999	9	-	0/15/38/38	0/3/3/3
10	APC	N	4999	9	-	0/15/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	D	3999	APC	PB-O2B	-3.40	1.48	1.56
10	D	3999	APC	PA-O2A	-3.01	1.49	1.56
10	N	4999	APC	PB-O2B	-2.90	1.49	1.56
10	N	4999	APC	PA-O2A	-2.68	1.49	1.56
10	N	4999	APC	PA-O5'	2.74	1.60	1.57

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	N	4999	APC	C1'-N9-C4	-7.20	116.08	126.94
10	D	3999	APC	C1'-N9-C4	-6.78	116.72	126.94
10	N	4999	APC	C2'-C1'-N9	-4.46	107.48	114.29
10	N	4999	APC	PG-O3B-PB	-3.94	119.47	132.67
10	D	3999	APC	PG-O3B-PB	-3.88	119.65	132.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
10	D	3999	APC	2	0
10	N	4999	APC	2	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [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	G	23/23 (100%)	-0.61	0 100 100	22, 46, 76, 79	0
1	X	23/23 (100%)	-0.68	0 100 100	21, 38, 71, 90	0
2	H	16/16 (100%)	-0.46	0 100 100	36, 56, 100, 102	0
2	Y	16/16 (100%)	-0.35	0 100 100	21, 73, 101, 103	0
3	I	13/14 (92%)	-0.75	0 100 100	49, 65, 74, 85	0
3	Z	13/14 (92%)	-0.51	0 100 100	54, 67, 80, 81	0
4	A	229/315 (72%)	-0.38	1 (0%) 93 80	41, 66, 86, 94	0
4	B	229/315 (72%)	-0.33	3 (1%) 79 53	43, 70, 85, 101	0
4	K	229/315 (72%)	-0.44	2 (0%) 85 64	41, 65, 81, 93	0
4	L	229/315 (72%)	-0.37	3 (1%) 79 53	38, 68, 80, 90	0
5	C	1119/1119 (100%)	-0.36	13 (1%) 81 55	26, 64, 86, 101	0
5	M	1119/1119 (100%)	-0.36	11 (0%) 84 60	25, 65, 89, 109	0
6	D	1264/1524 (82%)	-0.41	14 (1%) 82 58	23, 61, 83, 100	0
6	N	1264/1524 (82%)	-0.41	15 (1%) 81 55	26, 60, 84, 100	0
7	E	95/99 (95%)	-0.49	2 (2%) 67 36	35, 59, 74, 82	0
7	O	95/99 (95%)	-0.49	1 (1%) 82 58	45, 66, 80, 84	0
All	All	5976/6850 (87%)	-0.39	65 (1%) 82 58	21, 64, 86, 109	0

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	D	416	ALA	5.9
6	D	392	SER	4.6
6	N	427	VAL	4.4
6	N	391	ALA	4.4
4	B	147	GLY	4.2

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

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

6.3 Carbohydrates [i](#)

There are no 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
10	APC	N	4999	31/31	0.96	0.15	-0.47	45,49,51,54	0
10	APC	D	3999	31/31	0.96	0.15	-0.48	41,49,52,54	0
9	MG	D	8001	1/1	0.99	0.07	-1.58	23,23,23,23	0
9	MG	N	9001	1/1	0.96	0.07	-1.82	21,21,21,21	0
8	ZN	N	5058	1/1	0.95	0.04	-1.82	66,66,66,66	0
8	ZN	N	7112	1/1	0.99	0.04	-2.08	65,65,65,65	0
9	MG	D	8002	1/1	0.99	0.04	-2.40	25,25,25,25	0
8	ZN	D	4058	1/1	0.99	0.04	-2.87	66,66,66,66	0
8	ZN	D	6112	1/1	0.98	0.06	-3.14	59,59,59,59	0
9	MG	N	9002	1/1	0.98	0.06	-	23,23,23,23	0

6.5 Other polymers [i](#)

There are no such residues in this entry.