



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2023 – 04:23 PM EDT

PDB ID : 2O5J
Title : Crystal structure of the *T. thermophilus* RNAP polymerase elongation complex with the NTP substrate analog
Authors : Vassylyev, D.G.; Vassylyeva, 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

<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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35
buster-report : 1.1.7 (2018)
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) : 2.35

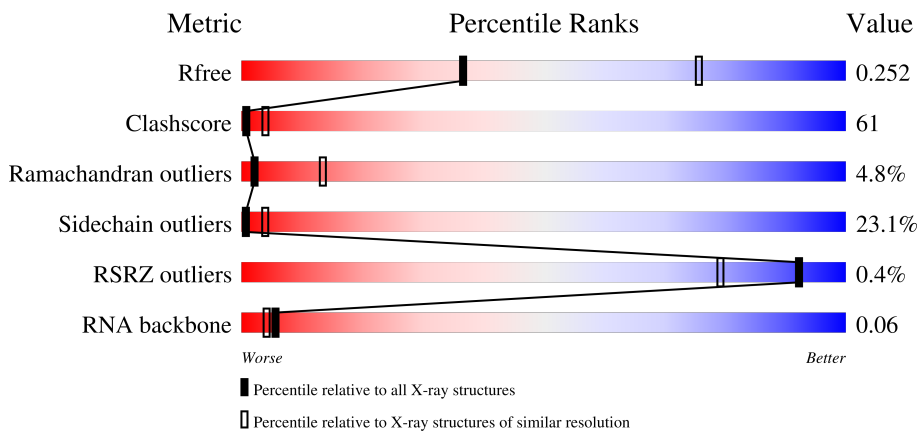
1 Overall quality at a glance

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.





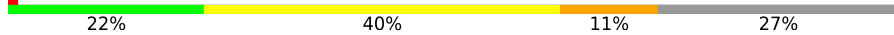
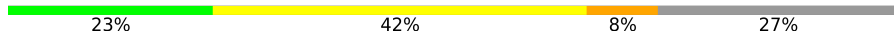
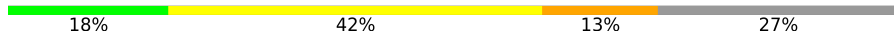
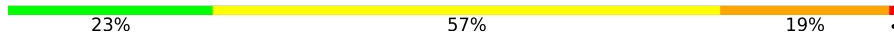
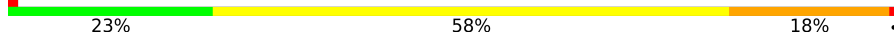
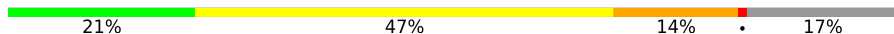
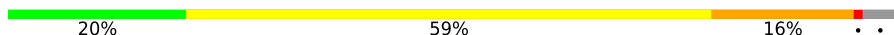
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

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

Mol	Chain	Length	Quality of chain
1	G	23	 17% 57% 22% .
1	X	23	 17% 61% 13% 9%
2	H	16	 19% 81%
2	Y	16	 31% 69%

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 i

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
			Total	C	N	O	P			
1	G	23	Total 467	C 220	N 80	O 144	P 23	0	0	0
1	X	23	Total 467	C 220	N 80	O 144	P 23	0	0	0

- 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
			Total	C	N	O	P			
2	H	16	Total 347	C 153	N 64	O 114	P 16	0	0	0
2	Y	16	Total 347	C 153	N 64	O 114	P 16	0	0	0

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

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

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

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

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).

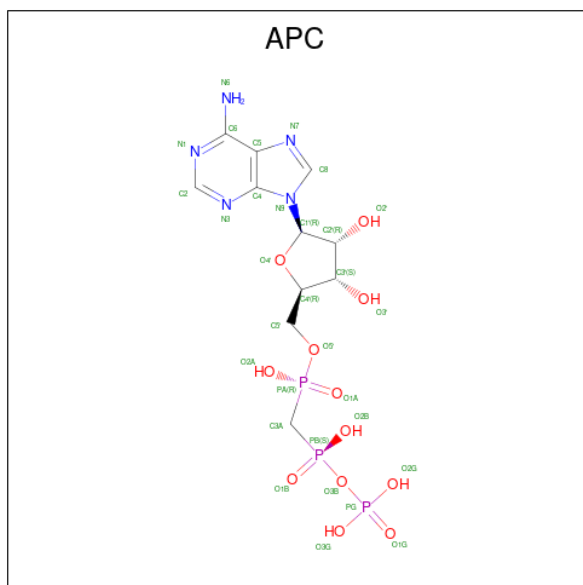
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Mg	0	0
			2	2		

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: C₁₁H₁₈N₅O₁₂P₃).



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

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	G	32	Total	O	0	0
			32	32		
11	H	37	Total	O	0	0
			37	37		
11	I	22	Total	O	0	0
			22	22		
11	X	43	Total	O	0	0
			43	43		
11	Y	30	Total	O	0	0
			30	30		
11	Z	30	Total	O	0	0
			30	30		

Continued on next page...

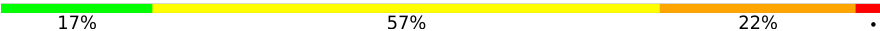
Continued from previous page...

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

3 Residue-property plots [i](#)

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: 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: 



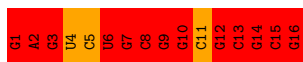
- Molecule 1: 5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3'

Chain X: 



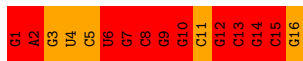
- Molecule 2: 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3',

Chain H: 

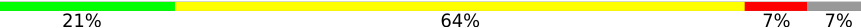


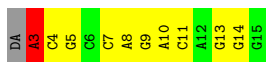
- Molecule 2: 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3',

Chain Y: 



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

Chain I: 



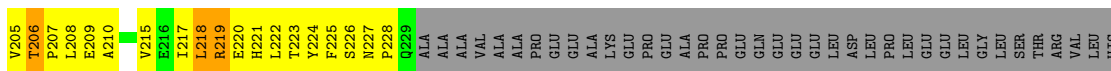
- Molecule 3: 5'-D(*AP*AP*CP*GP*CP*CP*AP*GP*AP*CP*AP*GP*GP*G)-3'

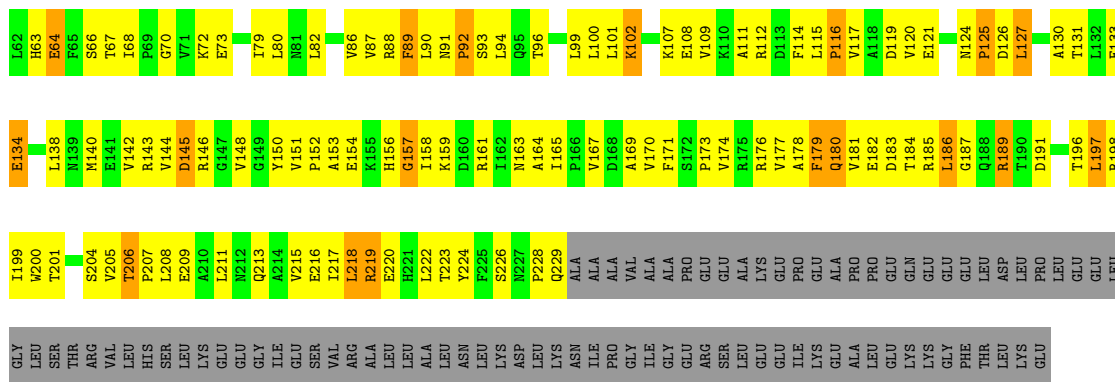
Chain Z: 29% 64% 7%



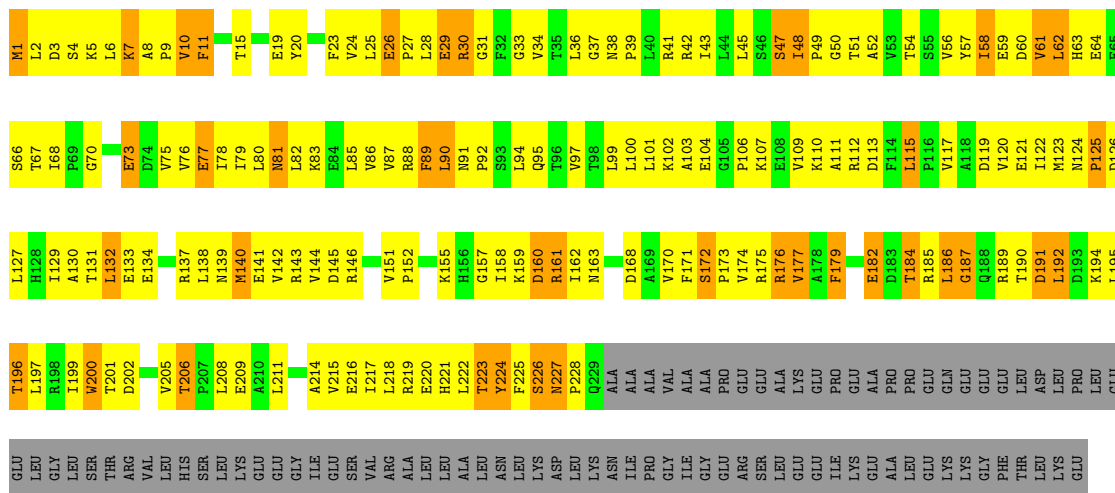
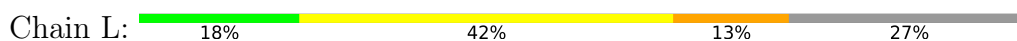
• Molecule 4: DNA-directed RNA polymerase alpha chain

Chain A: 21% 42% 10% 27%

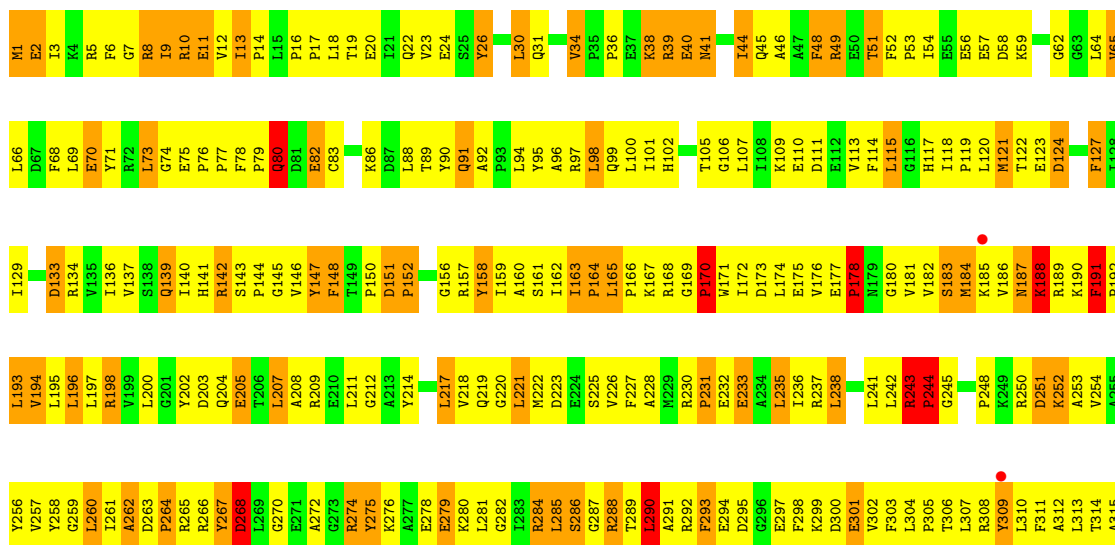


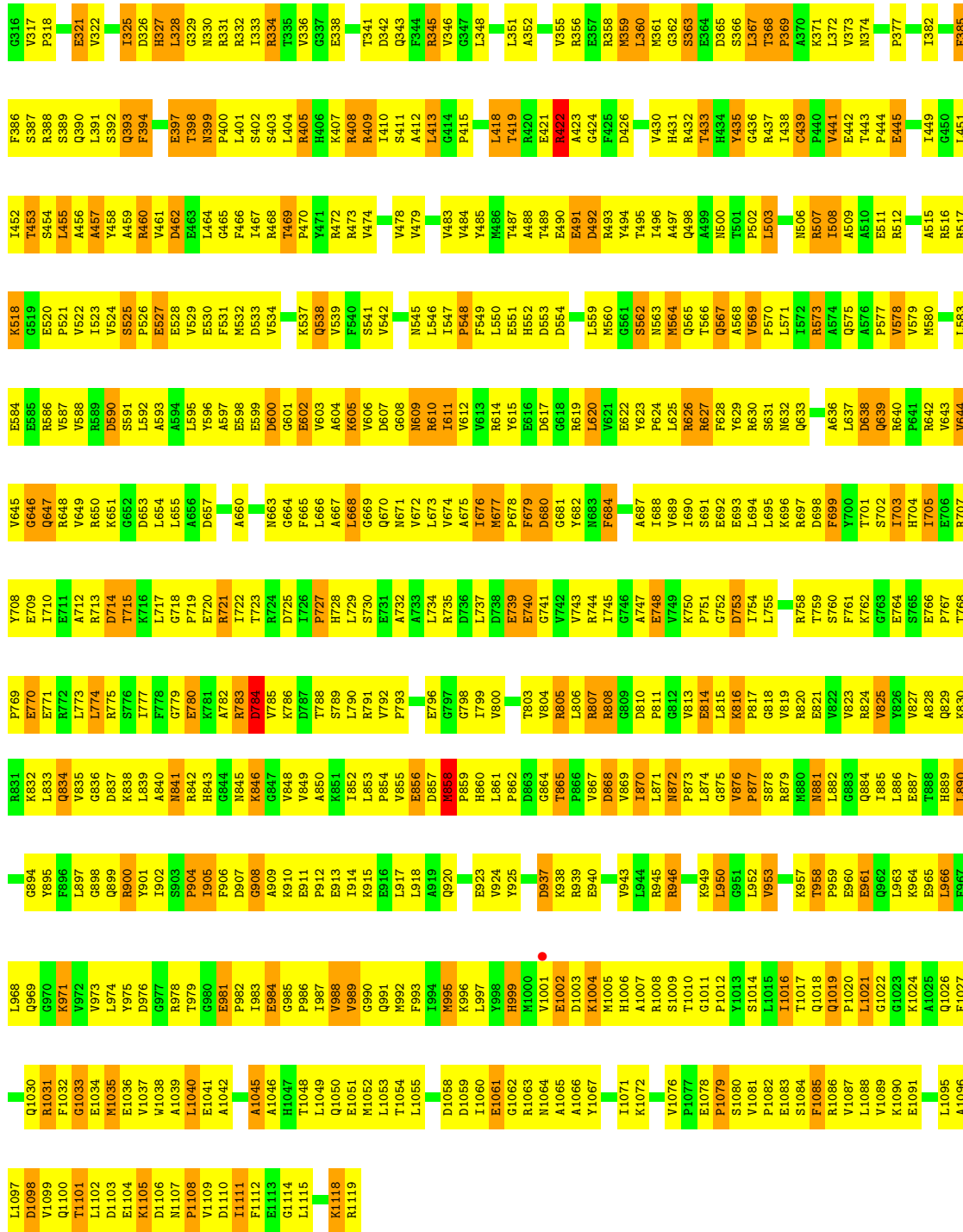


• Molecule 4: DNA-directed RNA polymerase alpha chain



• Molecule 5: DNA-directed RNA polymerase beta chain





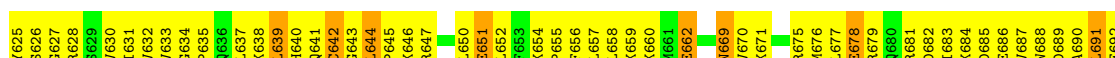
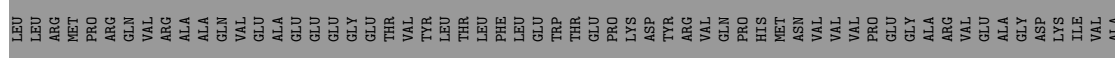
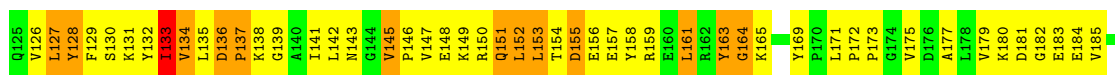
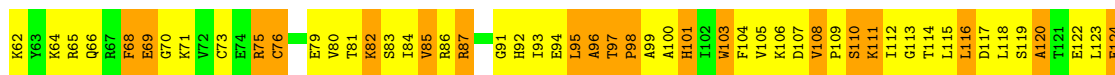
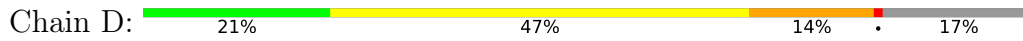
• Molecule 5: DNA-directed RNA polymerase beta chain

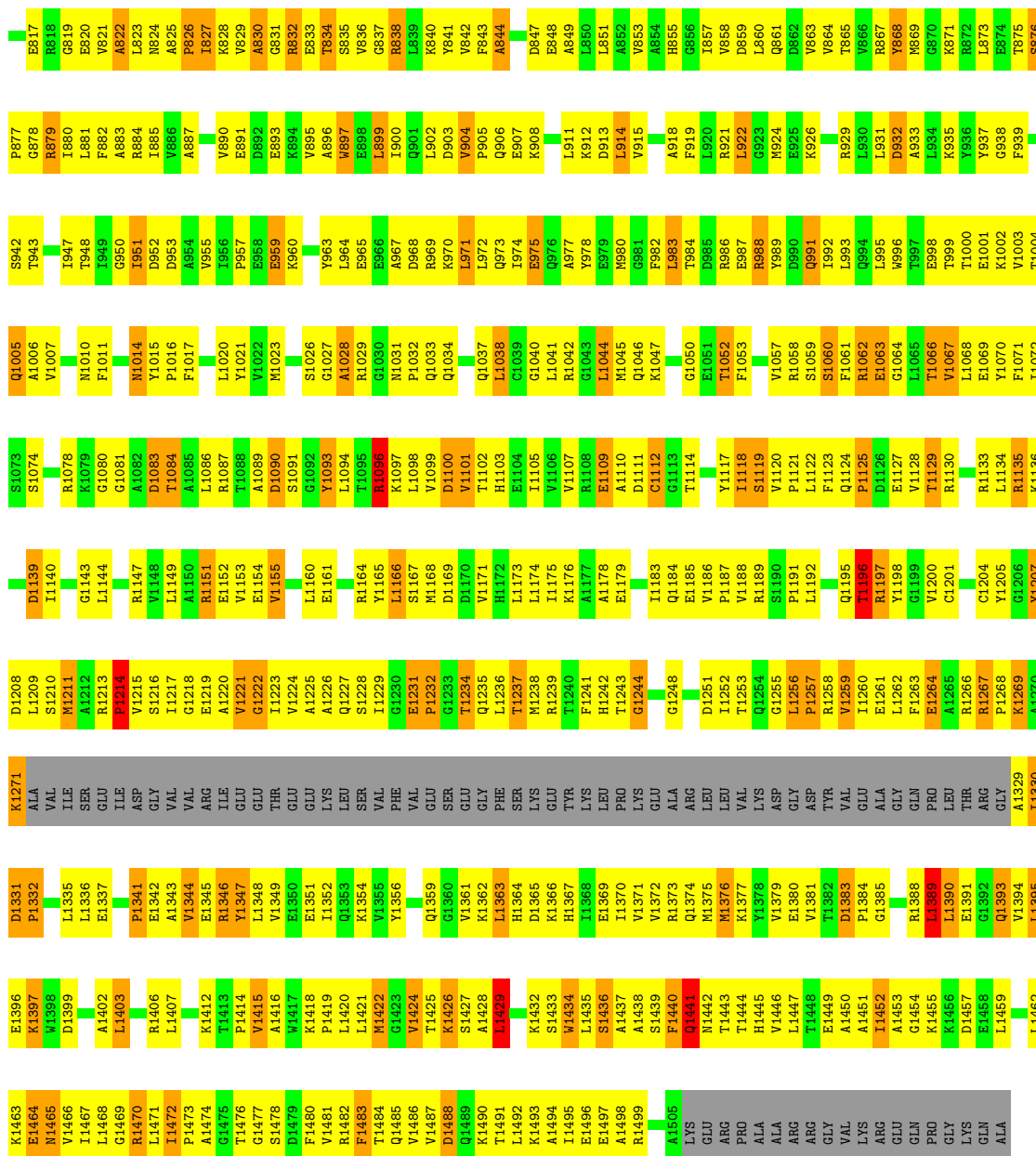


K971	G991	L839	S776	E711	Q647	G582	K518	I452	A381	P318	A283	V186	G125	G63
V972	G992	A840	I777	A712	R648	L583	G519	T452	A381	G319	V254	N187	S126	L64
L973	G993	M841	F778	R713	V649	E584	E520	T453	I382	H320	A255	R188	F127	V65
L974	G994	R842	G779	D714	R650	E585	P521	S454	R383	E321	Y256	R189	I128	L66
L975	G995	H843	E780	T715	G651	V522	V523	L455	E384	K190	V257	K190	I129	D67
L976	G996	G844	K781	K716	G652	V524	V525	A456	F385	F191	V258	F191	I130	F68
L977	G997	N845	A782	L717	D653	V526	S525	A457	F386	D324	G259	P192	G131	L69
L978	G998	K846	L654	G718	L654	S526	S526	A458	S387	L325	L260	P193	A132	E70
L979	G999	R847	L655	P719	L655	E527	E527	A459	R388	D326	D133	V194	D133	Y71
L980	G1000	V848	D656	E720	A656	E528	E528	R460	R460	L195	R134	L195	R134	R72
L981	G1001	V849	D657	R721	D657	E529	E529	V461	S392	L328	A262	L196	V135	L73
L982	G1002	A850	G658	E722	E658	V529	V529	D462	D462	G328	P264	L197	V136	G74
L983	G1003	K851	P659	F531	F531	E530	E530	E463	F394	R330	R265	R198	E75	E75
L984	G1004	L852	A660	F532	A660	M532	M532	L464	K395	R331	R266	V199	S138	P76
L985	G1005	S861	S661	M533	S661	D533	D533	G465	D396	R332	Y267	L200	Q139	P77
L986	G1006	R662	E662	D534	E662	D534	D534	F466	F466	I333	D268	I140	I140	P78
L987	G1007	N663	N663	S534	N663	V534	V534	R467	T398	R334	L269	D203	H141	P79
L988	G1008	G664	G664	F535	G664	S535	S535	R468	T399	R335	G270	Q204	R142	Q80
L989	G1009	E665	E665	P536	E665	G601	G601	T469	N399	R336	E271	A208	Y146	R84
L990	G1010	D666	D666	V537	D666	E602	E602	P470	P400	P400	V336	E205	S143	K86
L991	G1011	M667	M667	K537	K537	V603	V603	P471	L401	G337	A272	T286	P144	E82
L992	G1012	A667	A667	Q538	Q538	A604	A604	Y471	S402	E338	G273	G145	G145	C83
L993	G1013	R668	R668	V539	V539	K605	K605	R472	S403	L339	R274	A208	Y146	R84
L994	G1014	G669	G669	F540	F540	F606	F606	R473	L404	M340	Y275	Y147	Y147	E85
L995	G1015	D670	D670	S541	S541	D807	D807	V474	R405	T341	K276	E210	F148	K86
L996	G1016	M671	M671	V542	V542	G608	G608	V478	H406	D342	L211	T149	T149	D87
L997	G1017	N672	N672	N643	N643	M809	M809	V478	K407	Q343	K280	P150	P150	L88
L998	G1018	L673	L673	L546	L546	I610	I610	T480	R408	R344	L281	D151	P152	T89
L999	G1019	R674	R674	E547	E547	I611	I611	T480	R409	F345	G282	E216	G282	Y90
M000	G1020	A675	A675	I548	I548	V612	V612	V483	I410	V346	L283	A153	A153	Q91
M001	G1021	I676	I676	P548	P548	G613	G613	V484	S411	G347	K284	V218	V218	Q91
M002	G1022	M677	M677	F549	F549	R614	R614	V485	A412	L348	L285	P155	P155	L94
M003	G1023	E678	E678	L550	L550	V615	V615	M486	L413	A349	S286	G156	G156	Y95
M004	G1024	E679	E679	E551	E551	E616	E616	M487	F414	R350	G287	N222	N222	A96
M005	G1025	D680	D680	D552	D552	G617	G617	T487	P415	A352	T288	D223	D223	R97
M006	G1026	F681	F681	D554	D554	R619	R619	T488	A488	R288	E224	E224	E224	L98
M007	G1027	E682	E682	D555	D555	L620	L620	E490	L418	R353	L290	S225	S225	Q99
M008	G1028	D686	D686	R557	R557	V621	V621	E490	T419	G354	A291	V226	V226	L100
M009	G1029	A687	A687	E622	E622	E622	E622	R493	R422	V355	R292	F227	F227	I101
M010	G1030	I688	I688	Y623	Y623	P624	P624	Y494	A423	R358	F293	I163	I163	H102
M011	G1031	V689	V689	G561	G561	L825	L825	T495	G424	M359	E294	R230	P164	K103
M012	G1032	I690	I690	M564	M564	L826	L826	I496	F425	L360	D295	P231	L165	G106
M013	G1033	S691	S691	Q498	Q498	R626	R626	A497	D426	M361	G296	E232	P166	L107
M014	G1034	E692	E692	Q565	Q565	F627	F627	Q498	V427	G362	K299	E233	P166	L108
M015	G1035	E693	E693	T566	T566	F628	F628	P502	V430	S363	D300	A234	G169	K109
M016	G1036	L694	L694	Q567	Q567	Y629	Y629	L503	H430	E364	E301	L235	P170	L109
M017	G1037	L695	L695	A568	A568	R630	R630	E502	H431	D365	F302	G245	W171	E110
M018	G1038	K696	K696	V569	V569	S631	S631	L503	H431	D366	F303	G245	W171	E110
M019	G1039	E697	E697	N632	N632	G633	G633	M506	T433	L367	F303	G246	I172	D111
M020	G1040	D698	D698	L571	L571	Q633	Q633	R507	H434	T368	L241	D246	L173	E112
M021	G1041	F699	F699	I572	I572	L637	L637	I508	Y435	P369	L242	P247	L174	E113
M022	G1042	T701	T701	R573	R573	Y700	Y700	A509	G436	A370	T306	R244	V176	F114
M023	G1043	R831	R831	A574	A574	K371	K371	A510	R437	L371	E306	G247	E177	L115
M024	G1044	K832	K832	Q675	Q675	E511	E511	E510	I438	R308	F309	G248	P178	H117
M025	G1045	L833	L833	A576	A576	P641	P641	R512	I438	L372	Y309	D246	P178	H117
M026	G1046	H704	H704	F577	F577	I703	I703	R512	C439	V373	G311	P247	G180	H118
M027	G1047	I705	I705	V513	V513	R642	R642	V514	P440	M374	F311	P248	G180	H119
M028	G1048	V706	V706	V514	V514	V643	V643	V515	V441	S375	A312	K249	V181	P120
M029	G1049	E706	E706	V515	V515	V644	V644	A515	V442	S376	A312	K249	V181	M121
M030	G1050	L773	L773	M580	M580	V645	V645	R516	T443	R376	L313	R250	S183	M122
M031	G1051	R775	R775	V645	V645	G646	G646	R517	P444	E442	T314	D251	M184	E123
M032	G1052	I710	I710	T581	T581							K252	K185	D124

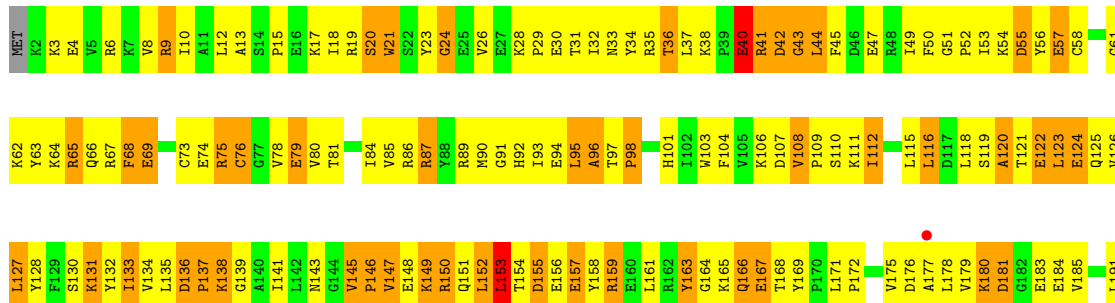


● Molecule 6: DNA-directed RNA polymerase beta' chain





● Molecule 6: DNA-directed RNA polymerase beta' chain



Y1145	G1081	L1020	D952	I1885	R818	A755	F693	S629	I566	L503	V443	GLU	PRO
G1146	A1082	Y1021	D953	V886	G819	Q756	V694	V630	I567	L506	V444	GLU	ARG
R1147	D1083	Y1022	A954	A887	A757	A757	I695	I631	R568	O507	R445	GLU	GLN
V1148	T1084	M1023	V955	E888	V821	E758	H696	V632	V507	O508	V446	VAL	VAL
L1149	A1085	L1024	I956	A889	A822	A759	A699	V633	R508	P509	V447	ARG	ALA
A1150	L1086	Q1025	P957	V890	L823	R760	V699	G634	K571	E510	E448	VAL	ALA
E1151	R1087	S1026	E958	E958	N824	I761	V700	P635	R572	E511	S449	ALA	ALA
V1152	T1088	G1027	E959	E893	A825	Q762	L702	Q636	H573	N512	V456	LEU	GLN
V1153	A1089	A1028	K960	K894	R826	R763	L637	L637	L574	D451	D451	VAL	VAL
E1154	D1090	K961	K961	V895	I827	E764	N703	R638	O575	L452	L452	GLY	GLY
V1155	S1091	G1030	Q962	A896	K828	S765	R704	L639	E576	D453	D453	VAL	ALA
L1156	G1092	M1031	V963	H897	V829	A774	A705	H640	A577	A454	A454	VAL	GLU
G1157	Y1093	P1032	I964	E898	H830	G776	P706	O641	V578	R465	R465	HIS	GLU
V1158	L1094	Q1033	E965	L899	G831	L769	T707	C642	R579	N517	H456	LEU	GLU
L1159	F1095	Q1034	E966	I900	R832	L770	L708	G643	A580	P518	G457	HIS	GLY
L1160	R1096	I1035	A967	Q901	E833	S771	H709	L644	L581	V519	G458	GLU	GLY
E1161	K1097	L1036	D968	L902	R834	A772	R710	P645	L582	L520	E459	THR	THR
E1162	L1098	Q1037	R969	D903	S835	A773	L711	K646	D583	P521	A460	VAL	VAL
G1163	V1099	L1038	K970	V904	V836	S774	G712	R647	H584	P522	L461	LEU	ARG
R1164	D1100	C1089	L971	P905	G837	G775	I713	H648	G585	D523	Q462	LEU	ARG
Y1165	V1101	Q1040	R972	Q906	R838	E776	Q714	E651	R586	L524	Q463	THR	THR
L1166	T1102	L1041	K973	K908	L839	P777	A715	L652	R587	R525	L464	LEU	LEU
S1167	R1042	R1042	E907	K908	K840	L778	F716	L652	L464	P526	L465	ASP	VAL
M1168	G1043	G1043	E975	V904	Y841	A779	Q717	F653	R588	K397	R466	GLU	GLY
D1169	V1105	L1044	O976	A918	L851	K780	P718	K654	P590	V528	E467	LEU	LYS
D1170	V1107	M1045	A977	L912	A844	P781	V719	P655	R599	H528	L468	PRO	LYS
V1171	R1108	Q1046	G987	D913	N845	S782	L720	F656	P594	V530	D469	VAL	GLU
L1174	E1109	L1047	F982	L914	P846	R783	V721	L657	D531	D531	L470	ALA	ALA
A1177	A1110	P1048	L983	V915	D847	D784	E722	L658	S596	O532	E471	GLY	GLY
A1178	D1111	S1049	T984	V916	E948	I785	G723	L659	S597	G533	A472	PHE	LEU
A1179	G1112	G1050	D985	Q917	R885	L786	Q724	K660	R598	B534	L473	ASP	ARG
I1183	G1113	E1051	E986	A918	L851	L787	S725	H661	P599	E474	D405	PRO	LEU
Q1184	T1114	I1052	E987	F919	A852	G788	I726	L600	L600	K475	D406	PRO	PRO
Y1186	M1116	F1053	R988	L920	V853	L789	Q727	I666	R601	S538	E476	VAL	GLY
V1187	N1117	D990	G989	L921	G989	Y790	L728	A667	S602	D539	L477	GLN	ALA
P1188	Y1118	V1055	G923	G923	I857	I791	H729	P668	L603	L540	L478	VAL	ALA
V1189	S1119	P1056	N924	N924	D859	T793	P730	N669	T604	N541	E479	PRO	ALA
V1190	V1120	R1058	E925	E925	L860	Q794	L731	V670	D605	D542	E480	TRP	TRP
L1191	P1121	S1059	L994	K926	Q861	Q795	V732	K671	L606	L543	R481	VAL	GLU
S1192	L1122	R1060	L995	K926	D862	R796	C733	A672	R545	Y544	K482	VAL	VAL
P1125	F1123	S1061	W996	R929	V863	K797	A735	A673	K610	R546	P484	VAL	ALA
L1192	Q1124	R1062	E998	L930	V664	E798	F736	R675	O611	S485	R486	PRO	ALA
C1194	T1125	E1063	E998	L934	Y868	K789	N737	L677	G612	I548	A422	LEU	LYS
T1128	V1128	L1065	T999	K935	K871	K800	A738	E678	R613	N549	A487	VAL	PRO
T1129	R1130	T1066	E1001	Y936	K871	G803	D739	E678	F614	R550	R488	ALA	GLY
R1197	G1134	V1067	K1002	Y937	R872	L804	F740	R679	R615	N551	R489	ARG	GLY
Y1198	L1134	L1068	G938	G938	L873	E805	D741	Q680	Q616	N552	A490	VAL	ILE
G1199	R1135	E1069	F939	F939	E874	F806	G742	R681	N617	B553	K491	GLU	GLN
V1200	K1136	F1070	T943	T943	T875	A807	Q744	I683	L618	L554	A492	ALA	PRO
C1201	K1137	I1072	T944	T944	S876	T908	N745	G620	G620	L557	R494	ASP	ALA
Q1202	A1138	S1073	N1010	N1010	P877	P809	A746	D685	K621	L558	R495	LYS	GLU
C1203	D1139	S1074	G946	G946	R879	E810	V747	E686	R622	A559	E436	ILE	ALA
C1204	L1140	H1075	I947	I947	L880	L813	H748	V687	V623	O560	E497	VAL	PRO
Y1205	E1141	R1078	T948	T948	L881	A814	V750	R688	D624	O561	V498	ALA	TYR
L1206	A1142	F1017	I949	I949	F882	A815	L751	A690	S626	H562	V499	ALA	LEU
G1143	G1143	P1017	G950	G950	A883	H816	L691	L691	G627	P563	R500	ILE	ARG
D1208	L1144	G1080	I951	I951	R884	E817	F754	E692	R628	I565	F502	PRO	ALA

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)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.225 , 0.257 0.227 , 0.252	Depositor DCC
R_{free} test set	11219 reflections (5.70%)	wwPDB-VP
Wilson B-factor (Å ²)	64.2	Xtrriage
Anisotropy	0.049	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 136.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	0.146 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	51213	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtrriage'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, 0.333 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, MG, APC

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 [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	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%)	4	23
4	B	227/315 (72%)	206 (91%)	15 (7%)	6 (3%)	5	27
4	K	227/315 (72%)	208 (92%)	13 (6%)	6 (3%)	5	27
4	L	227/315 (72%)	206 (91%)	15 (7%)	6 (3%)	5	27
5	C	1117/1119 (100%)	919 (82%)	136 (12%)	62 (6%)	2	10
5	M	1117/1119 (100%)	923 (83%)	133 (12%)	61 (6%)	2	10
6	D	1258/1524 (82%)	1051 (84%)	149 (12%)	58 (5%)	2	14
6	N	1258/1524 (82%)	1058 (84%)	140 (11%)	60 (5%)	2	13
7	E	93/99 (94%)	76 (82%)	11 (12%)	6 (6%)	1	7
7	O	93/99 (94%)	74 (80%)	12 (13%)	7 (8%)	1	5
All	All	5844/6744 (87%)	4927 (84%)	638 (11%)	279 (5%)	2	13

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

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

Mol	Chain	Res	Type
6	N	101	HIS
7	O	89	MET
6	N	199	LEU
6	N	98	PRO
6	N	913	ASP

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

Mol	Chain	Res	Type
5	M	872	ASN
6	N	976	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	M	1050	GLN
6	N	617	ASN
6	N	1465	ASN

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	Y	13	C
2	Y	12	G
2	Y	1	G
2	Y	9	G
2	H	15	C

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 monosaccharides 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	N	4999	9	27,33,33	1.25	4 (14%)	31,52,52	1.93	7 (22%)
10	APC	D	3999	9	27,33,33	1.39	3 (11%)	31,52,52	1.97	8 (25%)

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	N	4999	9	-	8/15/38/38	0/3/3/3
10	APC	D	3999	9	-	8/15/38/38	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	-3.54	1.48	1.56
10	D	3999	APC	PB-O3B	3.53	1.62	1.58
10	D	3999	APC	PA-O2A	-3.13	1.49	1.56
10	N	4999	APC	PB-O3B	3.13	1.61	1.58
10	N	4999	APC	PB-O2B	-3.02	1.49	1.56

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	N	4999	APC	C1'-N9-C4	-6.01	116.08	126.64
10	D	3999	APC	C1'-N9-C4	-5.65	116.72	126.64
10	D	3999	APC	C5-C6-N6	4.52	127.21	120.35
10	N	4999	APC	C5-C6-N6	4.29	126.87	120.35
10	N	4999	APC	PB-O3B-PG	-3.73	119.47	132.62

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

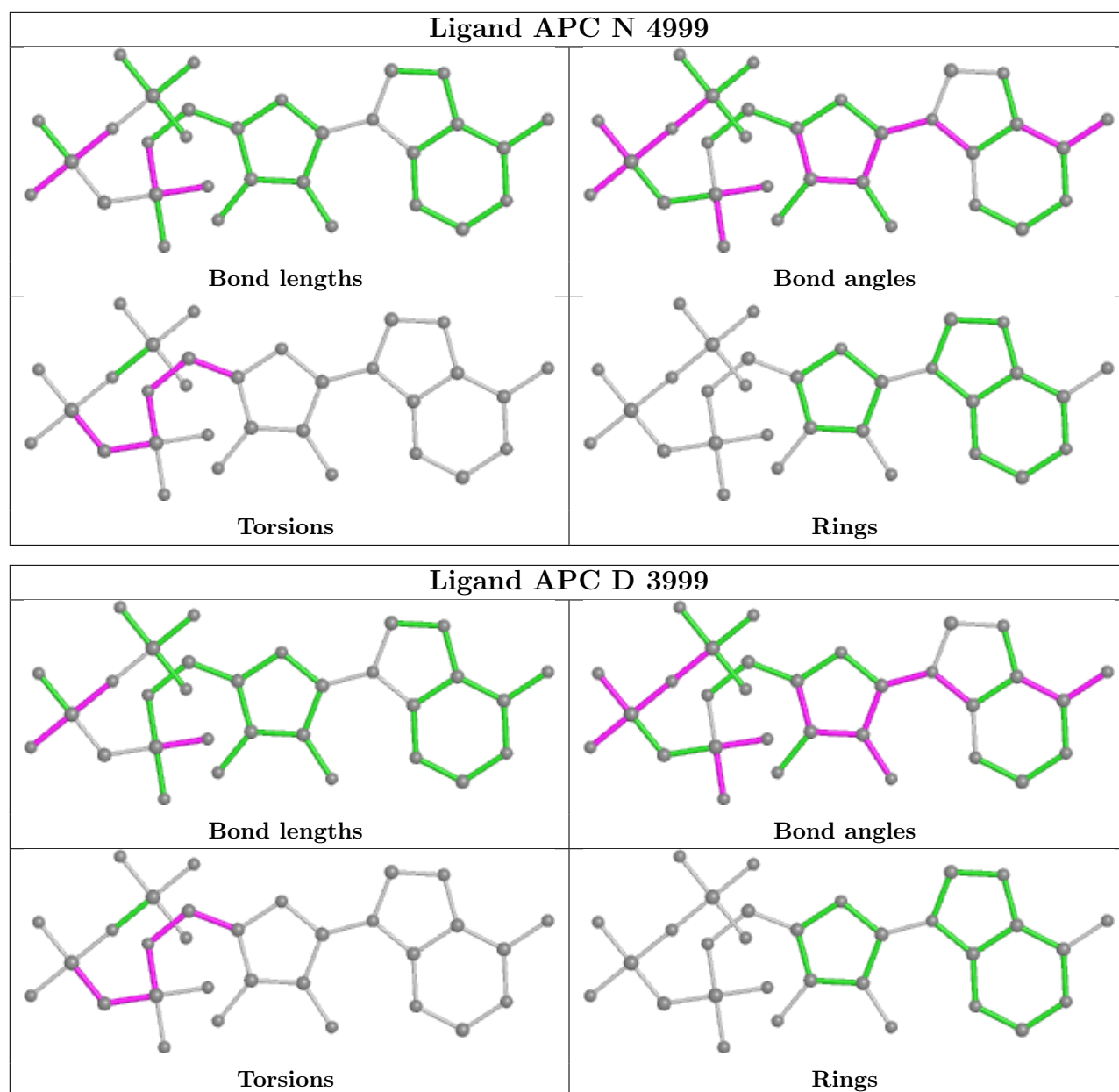
Mol	Chain	Res	Type	Atoms
10	D	3999	APC	PA-C3A-PB-O1B
10	D	3999	APC	PA-C3A-PB-O3B
10	D	3999	APC	PB-C3A-PA-O1A
10	D	3999	APC	PB-C3A-PA-O5'
10	D	3999	APC	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 4 short contacts:

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

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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%)	-0.73	0 100 100	22, 46, 76, 79	0
1	X	23/23 (100%)	-0.67	0 100 100	21, 38, 71, 90	0
2	H	16/16 (100%)	-0.58	0 100 100	36, 56, 100, 102	0
2	Y	16/16 (100%)	-0.62	0 100 100	21, 73, 101, 103	0
3	I	13/14 (92%)	-0.89	0 100 100	49, 65, 74, 85	0
3	Z	13/14 (92%)	-0.86	0 100 100	54, 67, 80, 81	0
4	A	229/315 (72%)	-0.62	0 100 100	41, 66, 86, 94	0
4	B	229/315 (72%)	-0.55	2 (0%) 84 63	43, 70, 85, 101	0
4	K	229/315 (72%)	-0.61	1 (0%) 92 79	41, 65, 81, 93	0
4	L	229/315 (72%)	-0.60	0 100 100	38, 68, 80, 90	0
5	C	1119/1119 (100%)	-0.60	3 (0%) 94 84	26, 64, 86, 101	0
5	M	1119/1119 (100%)	-0.58	6 (0%) 91 75	25, 65, 89, 109	0
6	D	1264/1524 (82%)	-0.65	5 (0%) 92 79	23, 61, 83, 100	0
6	N	1264/1524 (82%)	-0.66	5 (0%) 92 79	26, 60, 84, 100	0
7	E	95/99 (95%)	-0.66	0 100 100	35, 59, 74, 82	0
7	O	95/99 (95%)	-0.67	0 100 100	45, 66, 80, 84	0
All	All	5976/6850 (87%)	-0.62	22 (0%) 92 79	21, 64, 86, 109	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	D	416	ALA	4.9
4	B	147	GLY	4.3
6	D	391	ALA	3.5
6	N	427	VAL	3.2
6	D	192	ALA	3.1

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](#)

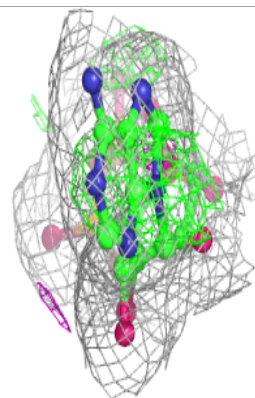
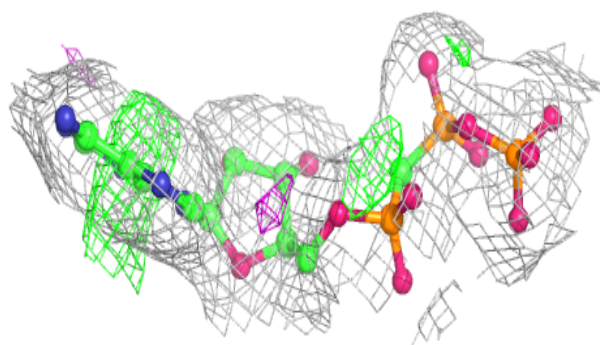
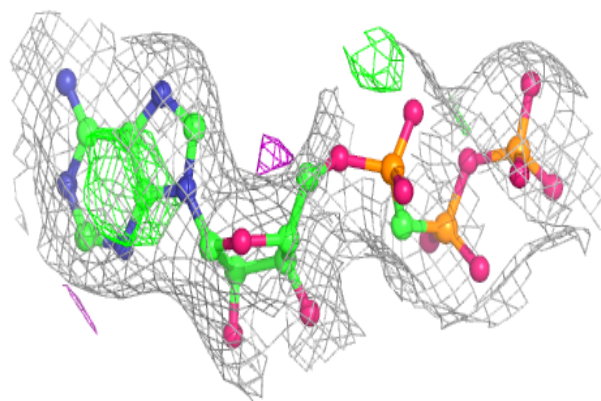
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	ZN	N	5058	1/1	0.97	0.09	66,66,66,66	0
10	APC	D	3999	31/31	0.97	0.16	41,49,52,54	0
9	MG	N	9002	1/1	0.98	0.09	23,23,23,23	0
10	APC	N	4999	31/31	0.98	0.16	45,49,51,54	0
9	MG	D	8002	1/1	0.99	0.09	25,25,25,25	0
9	MG	N	9001	1/1	0.99	0.11	21,21,21,21	0
8	ZN	D	6112	1/1	0.99	0.14	59,59,59,59	0
8	ZN	D	4058	1/1	0.99	0.09	66,66,66,66	0
9	MG	D	8001	1/1	0.99	0.10	23,23,23,23	0
8	ZN	N	7112	1/1	1.00	0.13	65,65,65,65	0

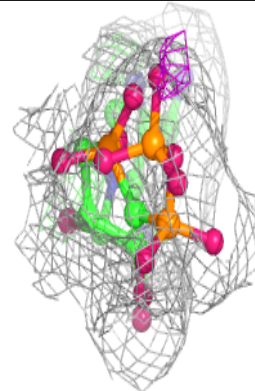
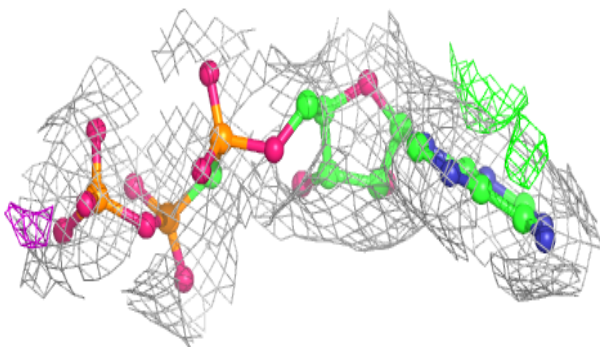
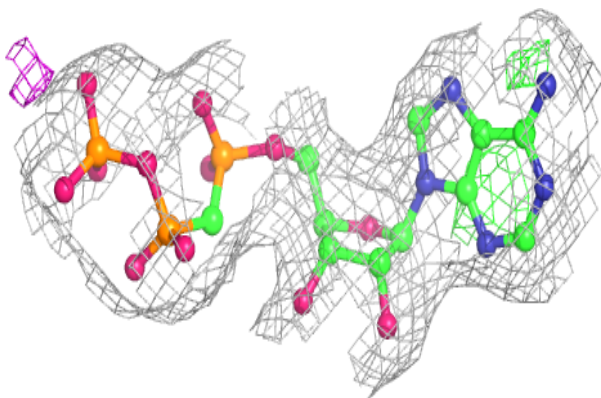
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around APC D 3999:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around APC N 4999:**

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers

There are no such residues in this entry.