



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:04 am GMT

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](mailto: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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

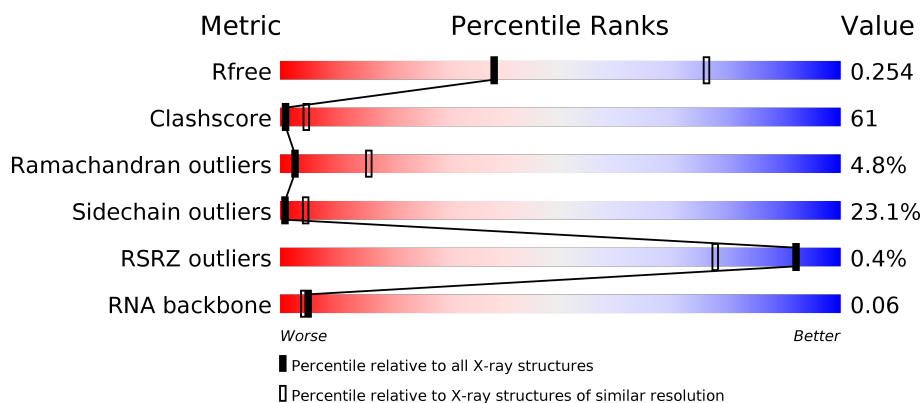
# 1 Overall quality at a glance

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.



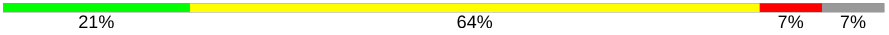
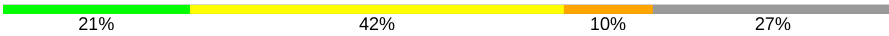
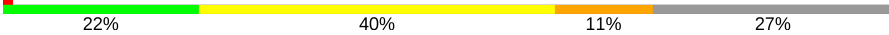
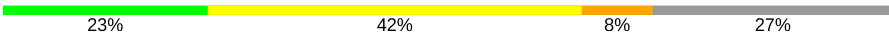
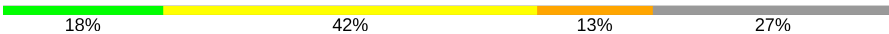
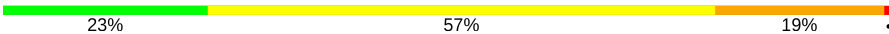
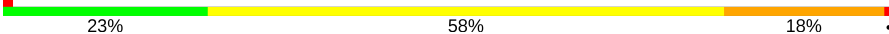
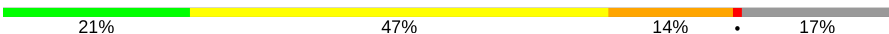
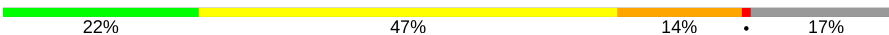
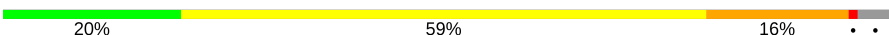
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	1692 (3.00-3.00)
Clashscore	112137	2037 (3.00-3.00)
Ramachandran outliers	110173	1973 (3.00-3.00)
Sidechain outliers	110143	1976 (3.00-3.00)
RSRZ outliers	101464	1716 (3.00-3.00)
RNA backbone	2435	1007 (3.34-2.66)

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  $\geq 3$ , 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>25%</div> <div>75%</div> </div>
2	Y	16	<div> <div>38%</div> <div>63%</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).

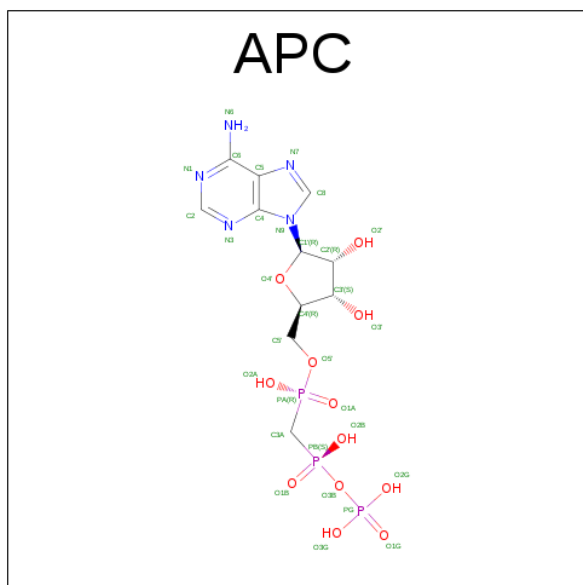
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Mg	0	0
			2	2		

*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 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: 



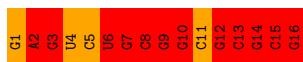
- Molecule 1: 5'-D(P\*CP\*CP\*CP\*TP\*GP\*TP\*CP\*TP\*GP\*GP\*CP\*GP\*TP\*TP\*CP\*GP\*CP\*GP\*CP\*GP\*CP\*CP\*G)-3'

Chain X: 



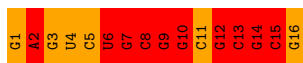
- Molecule 2: 5'-R(P\*GP\*AP\*GP\*UP\*CP\*UP\*GP\*CP\*GP\*GP\*CP\*GP\*CP\*GP\*CP\*G)-3',

Chain H: 



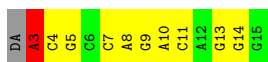
- Molecule 2: 5'-R(P\*GP\*AP\*GP\*UP\*CP\*UP\*GP\*CP\*GP\*GP\*CP\*GP\*CP\*GP\*CP\*G)-3',

Chain Y: 



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

Chain I: 



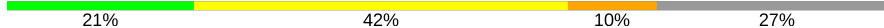
- Molecule 3: 5'-D(\*AP\*AP\*CP\*GP\*CP\*CP\*AP\*GP\*AP\*CP\*AP\*GP\*GP\*G)-3'

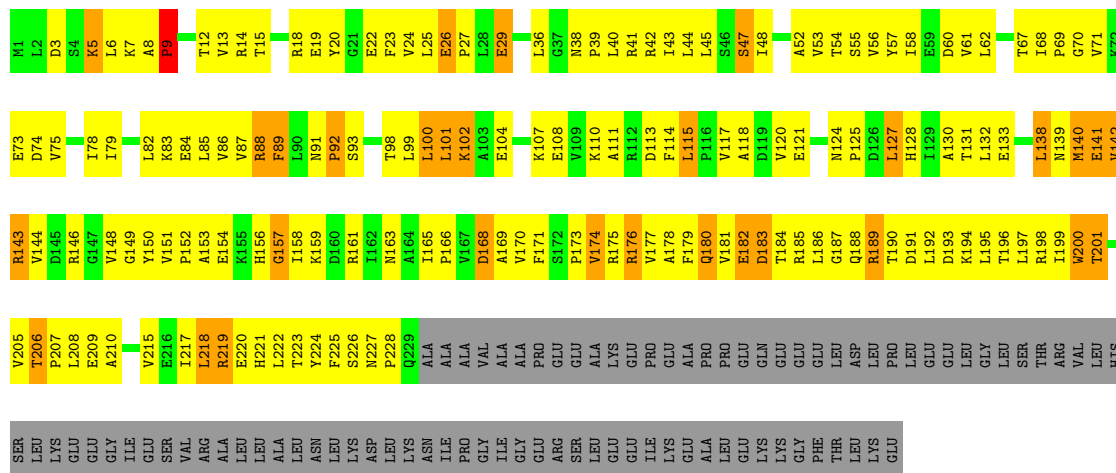


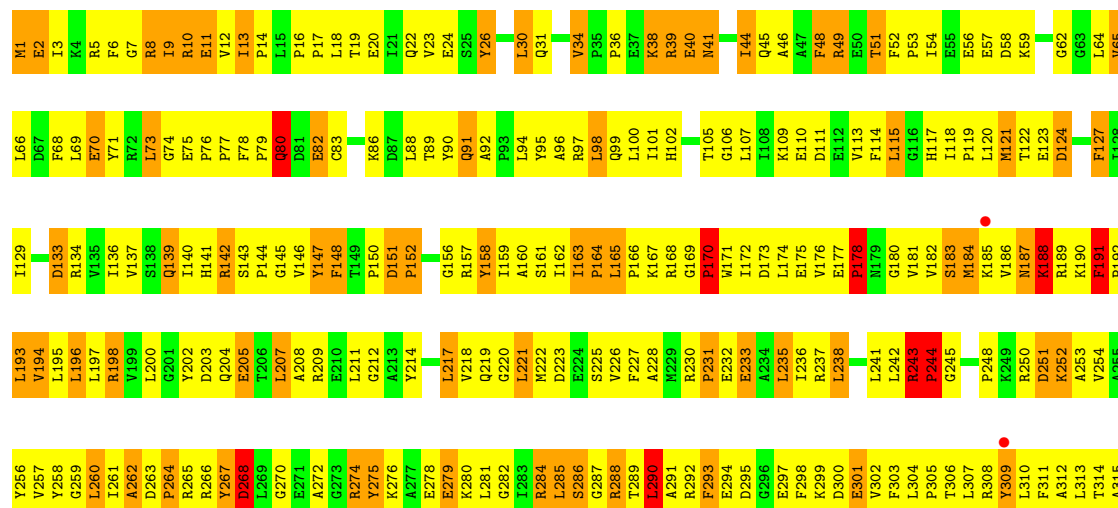
Chain Z: 

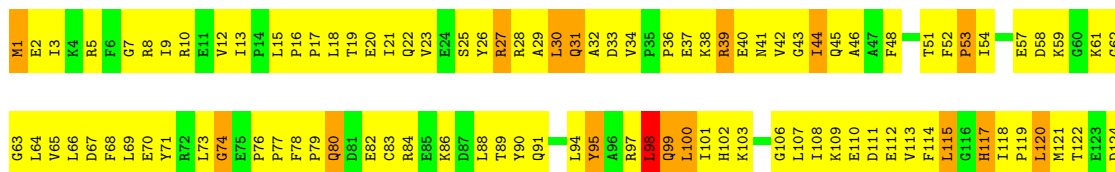


• Molecule 4: DNA-directed RNA polymerase alpha chain

Chain A: 

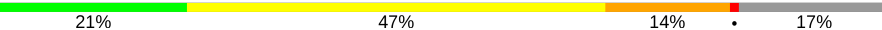




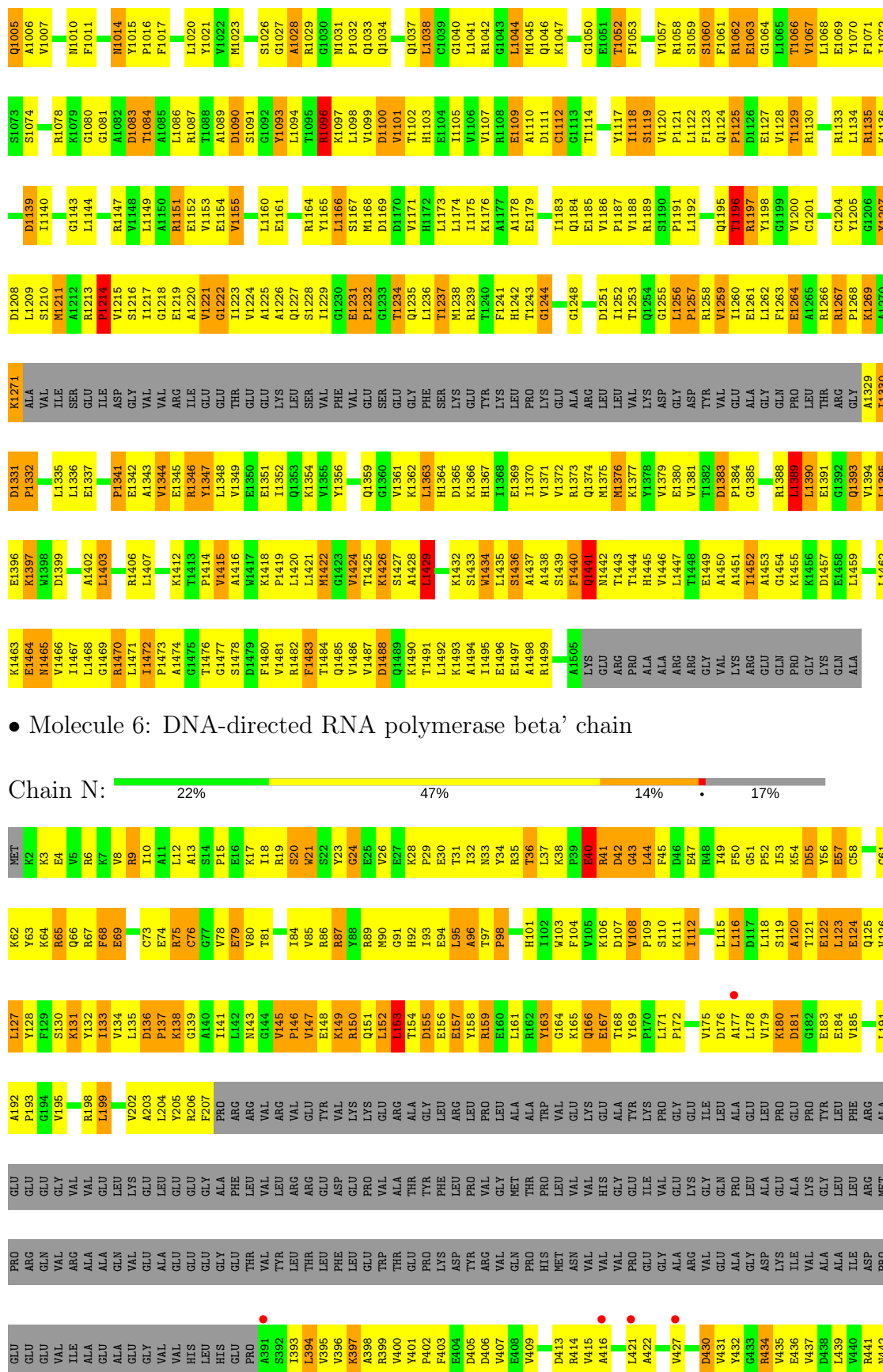


A1096	F1032	K971	Y901	L839	S776	E711	G647	G582	K518	I452	A381	P318	A253	V186	G125
L1097	G1033	Y972	I902	A940	I777	A712	R648	L583	G519	I453	I382	G319	V254	N187	S126
D1098	E1034	Y973	S903	N841	F778	R713	V649	E584	E520	T454	R383	E321	A255	R188	F127
T1101	M1035	L974	P904	R842	G779	D714	R650	E585	P521	T455	R382	E320	V256	R189	I128
L1102	E1036	Y975	F905	G844	E780	T715	K651	E586	I522	L455	F385	D323	V257	K190	I129
D1103	M1038	D976	I906	H843	K781	K716	G652	V587	I523	A456	F386	D324	G259	F191	N130
E1104	H1039	G977	G908	N845	A782	L717	D653	V588	V524	A457	F387	I325	L260	P192	G131
K1105	A1040	R978	K808	K846	R783	G718	L655	R889	S525	A458	R388	I326	L261	L193	A132
D1106	L1041	G980	R979	V848	D784	E720	A656	A594	P526	A459	R388	I327	A262	L194	D133
N1107	E1042	K785	K786	V849	K786	R721	D657	A594	E528	V461	S392	L328	D263	L196	R134
Y1043	P912	D787	D787	A850	D787	I726	G658	L595	V529	D462	F393	G329	P264	L197	I136
F1044	K913	T788	T788	K851	T788	R727	A660	Y596	E530	E463	F394	R330	R265	R198	I137
A1046	I914	S789	S789	I852	S789	R728	A660	A597	F531	L464	K395	R331	R266	V199	S138
I1111	K915	L790	L790	L853	L790	H728	S661	E598	N532	G465	D396	R332	Y267	L200	Q139
F1112	E916	R791	R791	P854	R791	L729	E662	E599	D533	I467	E397	R333	D268	A208	I140
E1113	L1049	V792	V792	V855	V792	S730	G663	D600	V534	R472	T398	R334	L269	Z204	H141
G1114	Q1050	F793	F793	E856	F793	E731	N664	G601	S535	R468	N399	R335	L270	E205	R142
L1115	E1051	F794	F794	D857	F794	A732	F665	E602	P536	T469	P400	V336	E271	E206	S143
T1054	T1054	G990	G920	M858	G990	A733	L666	V603	K537	P470	L401	G337	A272	T206	P144
K1055	L1055	F926	F926	H860	F926	L734	A667	Q538	Q538	Y471	S402	E338	G273	Z206	G145
K1056	F927	R800	R800	H861	R800	D735	L668	K605	V539	R472	S403	L339	R274	A208	Y147
S1057	E928	V801	V801	P862	V801	E740	Q670	V606	F540	R473	L404	K340	Y275	E210	Y148
D1058	K928	R802	R802	G863	R802	L737	M671	D607	S541	V474	H405	T341	K276	L211	F148
D1059	K930	T803	T803	G864	T803	E739	V672	G608	V542	V478	H406	D342	A272	L211	T149
I1060	G931	R804	R804	T865	R804	E740	L673	R610	N543	V479	K407	Q343	K280	G212	P150
E1061	Y938	R805	R805	P866	R805	V741	V674	L546	L546	T480	R408	R344	L281	E216	D151
F1062	H999	L806	L806	V867	L806	E742	A675	I547	R345	R409	R409	R345	G282	E216	P152
N1064	M1000	R807	R807	D868	R807	V743	L676	V613	V483	L410	L217	V346	I283	L121	A153
A1065	E1002	V808	V808	I870	V808	R744	M677	R614	V484	S411	G347	G347	R284	V218	R154
K1066	D1003	L871	L871	L870	D1003	I745	P678	Y615	L550	Y485	L413	L349	S286	L221	G156
Y1067	A1066	R810	R810	L871	A1066	E749	F679	E616	M486	G414	P415	R350	G287	M222	G156
I1070	H1006	V813	V813	L874	H1006	V749	D680	G618	H552	A488	L418	R352	R288	D223	Y158
I1071	A1007	E814	E814	G875	A1007	K750	F684	R619	D554	T489	T419	R353	L290	S225	I159
E1074	S1009	L815	L815	V876	S1009	G752	D686	L620	D554	E490	L418	G354	A291	V226	S161
D1075	T1010	R816	R816	P877	T1010	D753	A687	E622	R557	R493	R422	V355	R292	F227	I162
V1076	G1011	P817	P817	S878	G1011	L754	L688	Y623	M560	Y494	A423	R358	F293	R230	I163
P1077	P1012	V819	V819	R879	P1012	L756	V689	P624	G561	T495	G424	M359	E294	P231	L163
E1078	Y1013	R820	R820	N881	Y1013	G757	S691	L626	M564	A497	F425	L360	D295	E232	P166
P1079	S1014	E821	E821	L882	P1079	R758	E692	R626	Q565	Q498	D426	M361	G296	E233	P166
S1080	L1015	R822	R822	G883	S1080	T759	E693	F628	T566	Q498	V427	G362	K289	E234	G169
P1082	T1016	R824	R824	Q884	P1082	K762	L694	Y629	Q567	P502	V430	S363	D300	L235	G169
E1083	Q1018	V827	V827	L885	E1083	E764	L694	S631	L566	L503	H430	E364	E301	L235	P170
S1084	Q1019	A828	A828	L886	S1084	S765	K696	N632	P570	N506	R432	D365	V302	F239	I172
F1085	P1020	Q829	Q829	L887	F1085	E766	R697	Q633	L571	R507	R434	S366	F303	T240	I173
A1086	L1021	R830	R830	T888	A1086	R766	F699	Q633	I572	I508	H434	T368	P305	L242	E175
V1087	G1022	R831	R831	H889	V1087	E767	T700	L837	I572	R507	Y435	P369	T306	R243	V176
L1088	Q1023	K832	K832	L892	L1088	T768	T701	L837	R573	A509	G436	A370	L307	P244	E177
V1089	K1024	R833	R833	A893	V1089	S702	S702	R640	A574	A510	R437	K371	R308	F239	P178
E1091	A1025	L834	L834	G894	E1091	E770	H703	P641	Q575	E511	I438	G245	N178	D246	N178
L1092	F1027	Q834	Q834	L897	L1092	R772	I705	V643	P577	V513	C439	V373	L310	G180	D246
Q1093	Q1030	G836	G836	G998	Q1093	L773	E706	V644	V578	V513	V441	M374	F311	P248	V181
A1094	Q1094	D837	D837	Q999	A1094	L774	E706	V644	V579	A515	E442	R376	A312	K249	V182
L1095	L1095	K838	K838	R900	L1095	R775	I710	G646	M580	R516	T443	P377	T314	D251	M184
															K252

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

Chain D:  21% 47% 14% 17%

S942	P877	E817	S753	E693	Y625	L558	K494	Y432	ALA	LEU	LEU	Q189	Q125	K62	MET
T943	G878	R818	F754	V694	S626	A559	R495	G433	ILE	LEU	ARG	E190	G125	Y63	
	R879	E818	A755	L695	G627	Q560	L496	R434	ASP	ARG	ARG	L191	L127	K64	
T947	L880	G819	Q756	H696	R628	G561	E497	V435	PRO	MET	ALA	A192	Y128	R65	
T948	L881	E820	A757	G697	S629	A562	V498	E436	GLU	GLU	GLU	P193	F129	O66	
I949	F882	V821	E758	K698	V630	F863	V499	V437	GLU	ARG	GLU	G194	S130	R67	
G950	A883	A822	A759	V699	L631	E564	R500	D438	VAL	GLN	GLU	V195	K131	F68	
I951	R884	L823	R760	V700	V632	L565	A501	L439	VAL	VAL	GLU	V196	Y132	E69	
D952	L885	R824	I761	L702	G634	T566	F502	V440	ILE	ARG	VAL	S197	G70	R70	
D953	E886	A825	Q762	L701	L503	L667	F502	R441	ALA	VAL	VAL	R198	I133	K71	
A954	A887	R826	M763	N703	P635	R568	D504	N442	GLU	ALA	ALA	L198	L135	Y72	
V955		L827	L764	R704	Q636	N669	S505	V443	GLU	GLN	LEU		D136	C73	
	V890	K828	S765	A705	E570	E570	G506	V444	ALA	VAL	LEU		P137	C73	
P957	E891	V829	A766	P706	K638	K571	N507	R445	GLU	VAL	LYS	Y202	D136	E74	
E958	D892	A830	H767	T707	L639	R572	R508	V446	VAL	GLU	GLU	A203	K138	S14	
E959	E893	G831	N768	L708	H640	M573	P509	E447	VAL	GLU	GLU	Y205	G139	C76	
K960	K894	R832	L769	H709	Q641	L574	E510	E448	HIS	GLU	GLU	R206	A140		
	V895	E833	L770	R710	G642	Q575	W511	S449	LEU	GLU	GLY	F207	I141	E79	
	A896	T834	S771	L711	G643	E576	M512	Y450	HIS	GLU	ALA		L142	V80	
Y963	L897	S835	P772	G712	L644	A577	L513	D451	GLY	GLY	ALA	ARG	N143	T81	
E965	E898	V836	A773	L713	P645	V578	L514	L452	PRO	THR	VAL	ARG	G144	K82	
E966	L899	G837	S774	Q714	K646	D579	E515	D453	A391	VAL	VAL	ARG	W145	S83	
A967	I900	R838		A715	R647	A580	A516	A454	S392	TYR	LEU	ARG	P146	W85	
D968	Q901	L839	L778	F716		L581	V517	R455	L393	LEU	ARG	VAL	V147	R86	
R969	L902	K840	A779	Q717	L650	L582	P518	M456	L394	THR	ARG	GLU	E148	R87	
L970	K971	V841	K780	P718	E651	D583	V519	G457	V395	LEU	ARG	TYR	K149	V26	
L971	D904	H842	L781	V719	L652	N584	L520	A458	V396	PHE	ASP	VAL	Q151	G91	
L972	P905	F843	R783	L720	F653	G585	P521	E459	K397	LEU	GLU	LYS	L152	H92	
Q973	Q906	R844	R782	P721	K654	R586	P522	A460	A398	GLU	PRO	LYS	L153	T93	
I974	E907		D784	E722	P655		D523	L461	R399	TRP	VAL	GLU	T154	E94	
	K908	D847	L785	G723	F656	P590	L524	Q462	V400	THR	ALA	ARG	D155	L95	
E975		E848	L786	Q724	L657	L591	R525	Q463	Y401	GLU	THR	ALA	E156	I32	
Q976	L911	A849	L787	S725	L658	T592	L464	L465	P402	PRO	TYR	GLY	E157	T97	
E978	K912	L850	G788	L726	K659	N593	M527	L466	F403	PHE	LEU	LYS	Y158	P98	
E979	D913	L851	L789	Q727	K660	P594		K466		ASP	LEU	ARG	R159	A99	
M980	L914	A852	Y790	L728	K661	G595	V830	E467	D406	TYR	PRO	LEU	E160	A100	
	V915	R853	Y791	H729	E662		D531	L468	V407	ARG	VAL	PRO	E161	H101	
			L792	L730			G532	D469	E408	GLN	GLY	LEU	R162	T37	
F982	A918	R855	T793	P731	M669	L600	R601	L470	V409	VAL	MET	ALA	Y163	P39	
T984	F919	G856	Q794	V732	K670	S602	L603	E471	S410	PRO	THR	ALA	G164	F104	
D985	L920	L857	V795	C733	K671	S603	L603	A472	T411	HIS	PRO	TRP	K165	R41	
R986	R921	V858	R796	E734		T804		L473	G412	MET	LEU	VAL		D42	
E987	L922	D859	K797	A735	R675	D805	S538	E474	D413	ASN	VAL	GLY	Y169	K106	
R988	G923	L860	F798	F736	M676	L606	D539	K475	R414	VAL	VAL	LYS	P170	D107	
Y989	M924	Q861	K799	N737	L677	L607	L540	E476	V415	VAL	HIS	GLU	P171	P109	
D990	E925	D862	K900	A738	E678		N541	L477	A416	VAL	VAL	ALA	P172	S110	
Q991	K926	V863	G801	D739	R679	K610	D542	L478	P417	PRO	GLU	TYR	G173	K111	
I992		T864	A802	F740	Q680	Q611	L543	E479	D418	GLY	ILE	LYS	P174	I112	
	R929	T865	G803	D741	R681	G612	Y544	E480	C419	GLY	GLU	PRO	G175	G113	
L993	Q994	R866	L804	G742	D682	R613	R545	N481	D419	VAL	VAL	GLY	D176	T114	
Q994	L930	V867	E805	D743	L683	F614	R546	K482	L421	ARG	GLY	GLY	A177	L115	
L995	L931	R867	E806	Q744	K684	R615	L547	H483	A422	VAL	GLU	ILE	L178	L116	
V996	D932	V868	A807	Q745	D685	Q616	L548	P484	D423	GLU	GLN	LEU	V179	D117	
T997	A933	K669	F807	N745	E685	R616	N549	S485		ALA	PRO	ALA	K180	L118	
E998	L934	G870	T868	A746	E686	N617	N549	S485	K426	GLY	LEU	GLU	D181	S119	
	K935	R871	P809	V747	V687	L618	R550	R486	V427	GLY	LEU	GLU	E182	Y56	
T1000	Y936	R872	E810	H748	G688	L618		A487	K428	ASP	ALA	LEU	G181	T121	
E1001	Y937	L873	E811	V749	D689	G620	L554	R488	S429	ILE	GLU	PRO	E183	C58	
K1002	G938	E874	A812	P750	A690	K621	K555	R489	E132	TYR	ALA	GLU	E184	A59	
		T875	L813	L751	L691	R622	K556		D430	VAL	LYS	PRO	V185	C60	
I1003	F939	S876	A814	S752	E692		L557	R493	V431	ALA	GLY	TYR	E194	E124	

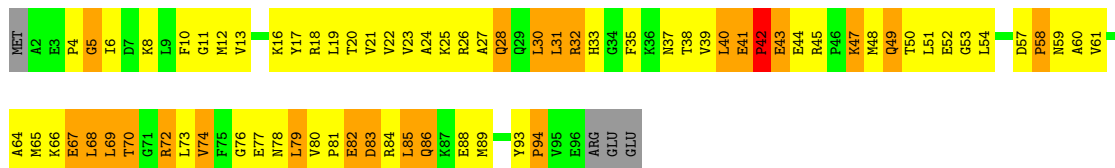






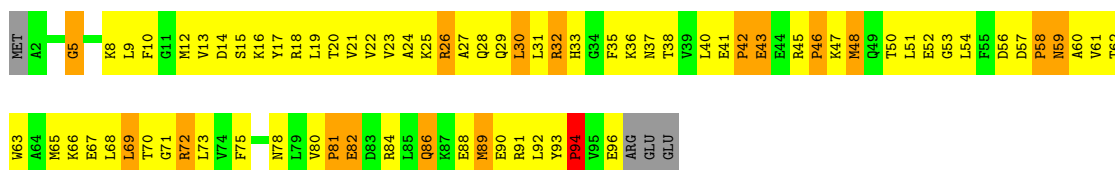
• Molecule 7: DNA-directed RNA polymerase omega chain

Chain E: 22% 49% 23% . .



• Molecule 7: DNA-directed RNA polymerase omega chain

Chain O: 20% 59% 16% . .





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	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 <sup>2</sup> )	64.2	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 136.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.146 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	51213	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

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

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

## 5 Model quality

### 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	26
4	B	227/315 (72%)	206 (91%)	15 (7%)	6 (3%)	6	31
4	K	227/315 (72%)	208 (92%)	13 (6%)	6 (3%)	6	31
4	L	227/315 (72%)	206 (91%)	15 (7%)	6 (3%)	6	31
5	C	1117/1119 (100%)	919 (82%)	136 (12%)	62 (6%)	2	12
5	M	1117/1119 (100%)	923 (83%)	133 (12%)	61 (6%)	2	12
6	D	1258/1524 (82%)	1051 (84%)	149 (12%)	58 (5%)	3	16
6	N	1258/1524 (82%)	1058 (84%)	140 (11%)	60 (5%)	2	16
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%)	2	16

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%)	0	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	9
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	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	15/16 (93%)	10 (66%)	0
2	Y	15/16 (93%)	10 (66%)	0
All	All	30/32 (93%)	20 (66%)	0

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

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, 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	28,33,33	1.39	4 (14%)	28,52,52	1.89	7 (25%)
10	APC	N	4999	9	28,33,33	1.26	4 (14%)	28,52,52	1.88	6 (21%)

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 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	D	3999	APC	PB-O2B	-3.41	1.48	1.56
10	D	3999	APC	PA-O2A	-3.01	1.49	1.56
10	N	4999	APC	PB-O2B	-2.91	1.49	1.56
10	N	4999	APC	PA-O2A	-2.69	1.49	1.56
10	D	3999	APC	PG-O3B	2.03	1.63	1.60

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	N	4999	APC	C1'-N9-C4	-6.11	116.08	126.64
10	D	3999	APC	C1'-N9-C4	-5.74	116.72	126.64
10	N	4999	APC	PG-O3B-PB	-3.66	119.47	132.38
10	D	3999	APC	PG-O3B-PB	-3.61	119.65	132.38
10	D	3999	APC	O2'-C2'-C3'	-2.91	102.51	111.83

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

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
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, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	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 61	43, 70, 85, 101	0
4	K	229/315 (72%)	-0.62	1 (0%) 92 77	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%) 93 82	26, 64, 86, 101	0
5	M	1119/1119 (100%)	-0.58	6 (0%) 90 74	25, 65, 89, 109	0
6	D	1264/1524 (82%)	-0.65	5 (0%) 92 77	23, 61, 83, 100	0
6	N	1264/1524 (82%)	-0.66	5 (0%) 92 77	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 77	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 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, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	ZN	D	6112	1/1	0.99	0.14	0.22	59,59,59,59	0
10	APC	N	4999	31/31	0.98	0.16	0.22	45,49,51,54	0
10	APC	D	3999	31/31	0.97	0.16	0.21	41,49,52,54	0
8	ZN	N	7112	1/1	1.00	0.13	-0.80	65,65,65,65	0
8	ZN	N	5058	1/1	0.97	0.09	-1.18	66,66,66,66	0
9	MG	N	9001	1/1	0.99	0.11	-1.25	21,21,21,21	0
9	MG	D	8001	1/1	0.99	0.10	-1.42	23,23,23,23	0
9	MG	D	8002	1/1	0.99	0.09	-1.89	25,25,25,25	0
8	ZN	D	4058	1/1	0.99	0.09	-1.91	66,66,66,66	0
9	MG	N	9002	1/1	0.98	0.09	-	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.