



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 11:11 pm GMT

PDB ID : 2PPB
Title : Crystal structure of the T. thermophilus RNAP polymerase elongation complex with the ntp substrate analog and antibiotic streptolydigin
Authors : Vassylyev, D.G.; Vassylyeva, M.N.; Artsimovitch, I.; Landick, R.
Deposited on : 2007-04-28
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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

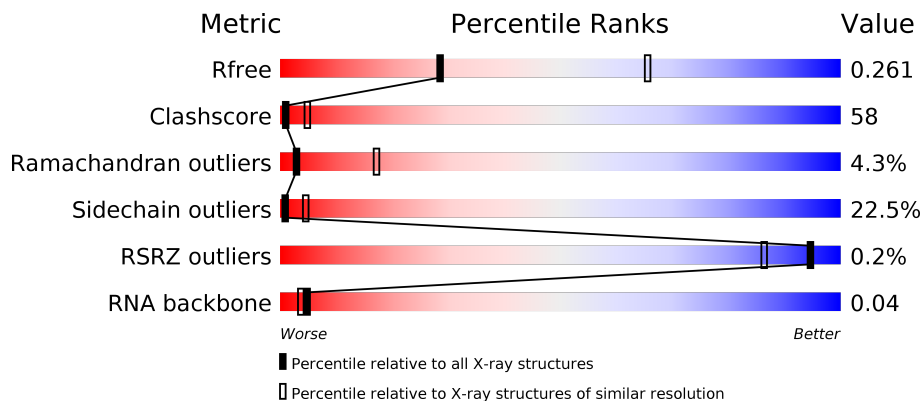
1 Overall quality at a glance

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}	111664	1851 (3.00-3.00)
Clashscore	122126	2167 (3.00-3.00)
Ramachandran outliers	120053	2101 (3.00-3.00)
Sidechain outliers	120020	2104 (3.00-3.00)
RSRZ outliers	108989	1751 (3.00-3.00)
RNA backbone	2636	1017 (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 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>22%</div> <div>48%</div> <div>26%</div> <div>.</div> </div>
1	X	23	<div> <div>17%</div> <div>52%</div> <div>30%</div> </div>
2	H	16	<div> <div>19%</div> <div>81%</div> </div>
2	Y	16	<div> <div>25%</div> <div>75%</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 12 unique types of molecules in this entry. The entry contains 51962 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 DNA (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 RNA (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 DNA (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			

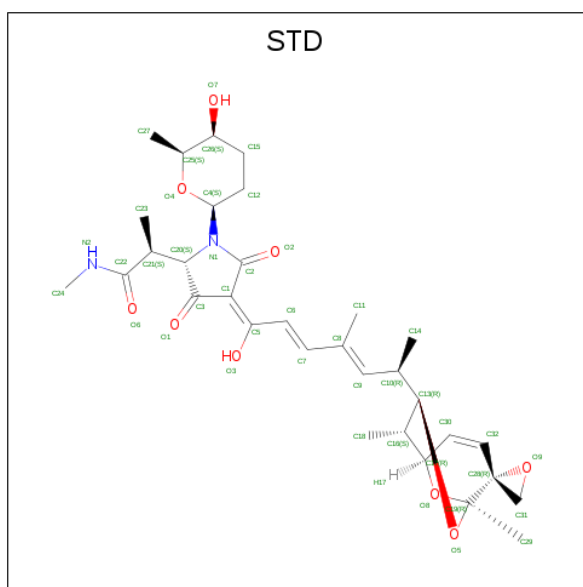
- Molecule 6 is a protein called DNA-directed RNA polymerase beta' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	1314	Total	C	N	O	S	0	0	0
			10373	6565	1838	1937	33			
6	N	1314	Total	C	N	O	S	0	0	0
			10373	6565	1838	1937	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 STREPTOLYDIGIN (three-letter code: STD) (formula: C₃₂H₄₄N₂O₉).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	D	1	Total	C	N	O	0	0
			43	32	2	9		
8	N	1	Total	C	N	O	0	0
			43	32	2	9		

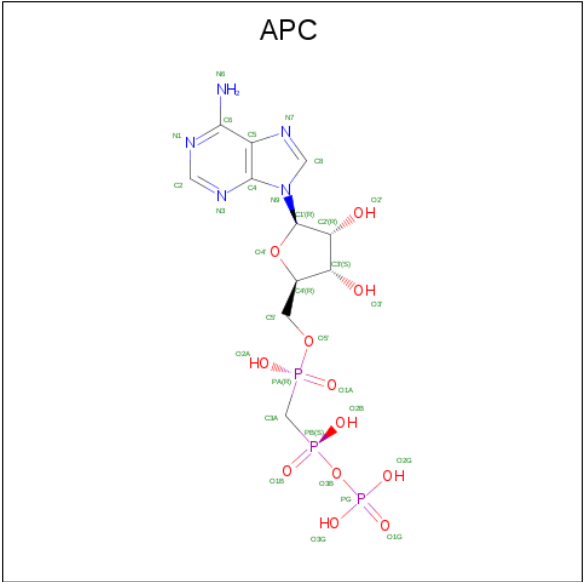
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	2	Total	Mg	0	0
			2	2		
10	N	2	Total	Mg	0	0
			2	2		

- Molecule 11 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: C₁₁H₁₈N₅O₁₂P₃).



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

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	G	39	Total	O	0	0
			39	39		
12	H	22	Total	O	0	0
			22	22		
12	I	31	Total	O	0	0
			31	31		
12	X	31	Total	O	0	0
			31	31		
12	Y	26	Total	O	0	0
			26	26		
12	Z	18	Total	O	0	0
			18	18		
12	A	78	Total	O	0	0
			78	78		
12	B	117	Total	O	0	0
			117	117		
12	C	408	Total	O	0	0
			408	408		
12	D	531	Total	O	0	0
			531	531		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	E	34	Total 34	O 34	0	0
12	K	81	Total 81	O 81	0	0
12	L	95	Total 95	O 95	0	0
12	M	396	Total 396	O 396	0	0
12	N	510	Total 510	O 510	0	0
12	O	53	Total 53	O 53	0	0

3 Residue-property plots

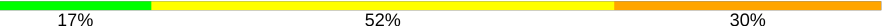
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

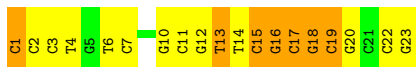
- Molecule 1: DNA (5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3')

Chain G: 



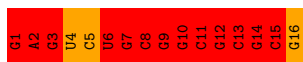
- Molecule 1: DNA (5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3')

Chain X: 



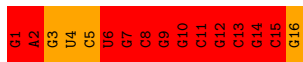
- Molecule 2: RNA (5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3')

Chain H: 



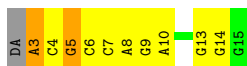
- Molecule 2: RNA (5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3')

Chain Y: 



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

Chain I: 

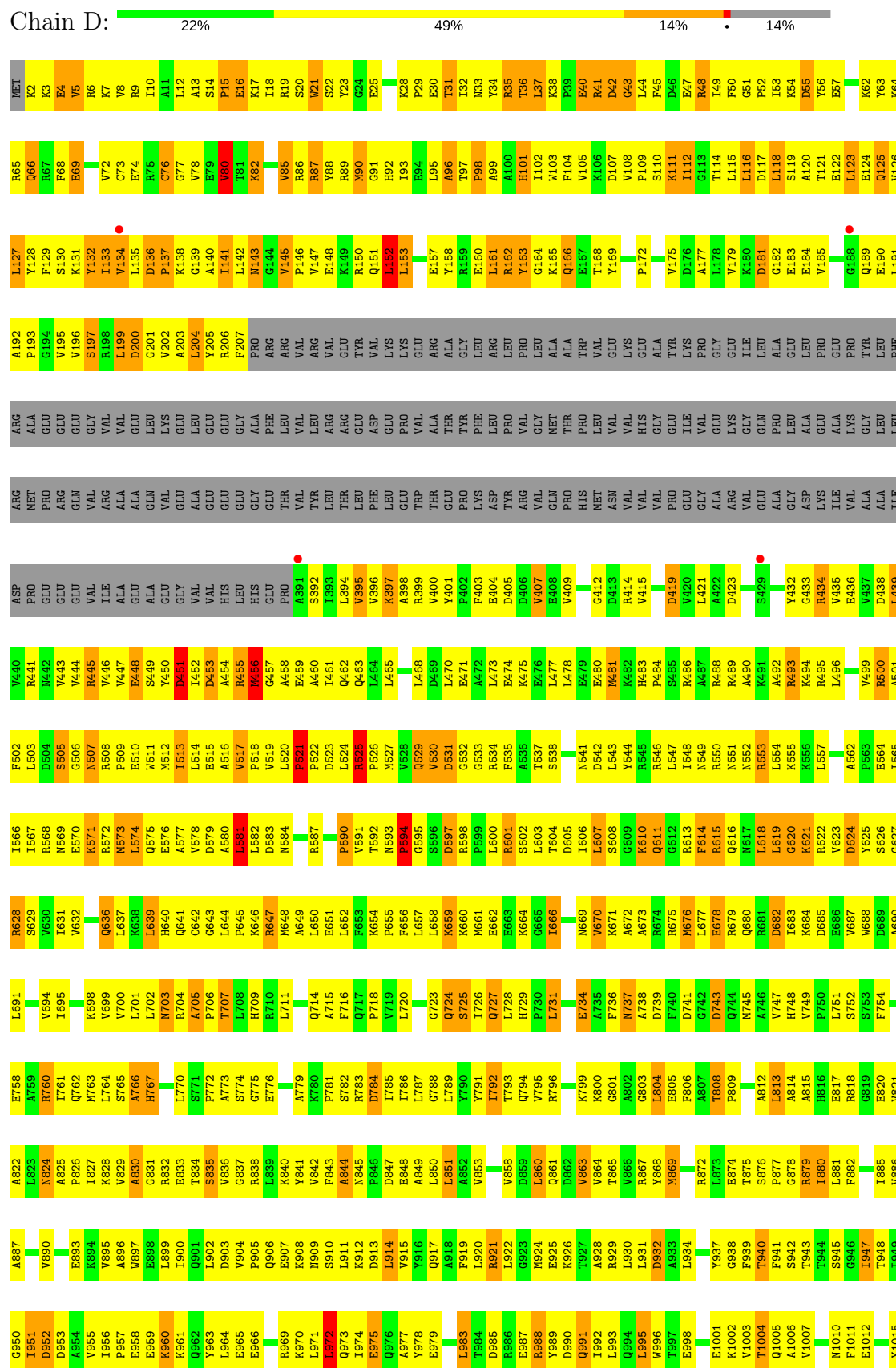




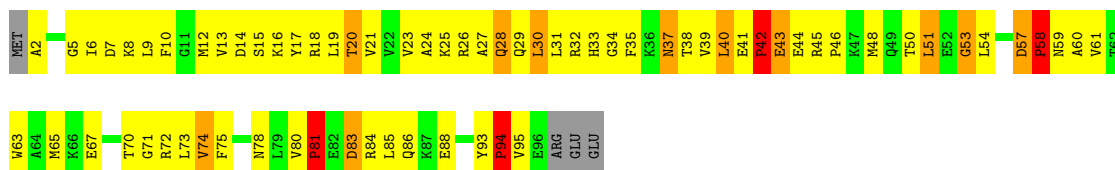


P1082	P1020	E960	L886	V825	S765	S702	Q639	V569	G505	T443	R383	E321	K252	R188	S126	L63
E1083	L1021	E961	E887	Y826	E766	I703	R640	P570	N506	P444	E384	E321	K252	R189	F127	L64
S1084	G1023	Q962	T888	Y827	P767	H704	P641	L571	R507	A447	F386	V322	Y258	R190	I128	V65
F1086	G1024	L963	H889	A828	T768	I705	R642	L572	L508	A448	F387	V323	G259	F191	N129	L66
V1087	A1025	E965	L890	K829	P769	E706	V643	A574	A509	N448	S387	D324	L260	P192	M130	D67
L1088	G1026	P966	L896	K830	E770	R707	V644	Q575	A510	I449	R388	I325	L261	L193	G131	F68
V1089	F1027	P967	L897	R831	E771	Y708	V645	Q576	E511	G450	S389	D326	A262	V194	A132	L69
L1090	G1028	L968	L898	K832	R772	E709	G646	A576	R512	L461	Q390	H327	D263	L196	D133	E70
E1091	G1029	Q969	Q899	Q834	L774	A712	R648	V578	V514	I462	L391	L328	P264	R197	V135	R72
L1092	G1030	G970	R900	V835	R775	E713	Q649	V579	A515	T463	Q392	G329	R265	L198	W136	L73
Q1093	R1031	K971	Y901	G836	S776	R713	R650	M580	R516	L455	F394	N330	R266	R198	L136	L73
A1094	D1032	V972	L892	D837	L777	D714	R651	T581	R517	A456	F394	R331	Y267	V199	G74	G74
L1095	G1033	V973	P904	K838	F778	T715	G652	G582	K518	A457	D396	I333	D268	L200	S138	F76
A1096	E1034	L974	I905	L839	G779	K716	D653	L583	G519	Y458	E397	I334	L269	Q139	Q139	P76
L1097	L1035	Y975	F906	A840	E780	L717	L654	E584	S525	A459	E397	T335	G270	I140	I140	P77
D1098	E1036	D976	D907	R841	K781	G718	L655	E585	P521	R460	N399	V336	A272	Q204	H141	F78
V1099	G1037	G977	G908	H842	A782	F719	L655	R586	V522	V461	P400	G337	G273	Q205	P79	P79
Q1100	V1038	R978	H843	H843	R783	E720	P659	D580	I523	D462	L401	E338	R274	T206	S143	D80
T1101	A1039	T979	Q844	Q844	D784	R721	A660	D580	V524	E463	S402	L339	K280	L207	P144	D81
L1102	L1040	G980	P912	L785	I722	H728	S661	D580	S525	E463	S402	L339	K280	L207	G145	E82
D1103	E1041	E981	E913	G847	T723	L729	E662	A594	P526	L464	S402	L339	K280	L207	G145	C83
L1104	A1042	P982	I914	V848	D787	R724	N663	L596	E527	G465	L404	T341	A277	Q209	Y147	K86
K1105	Y1043	L983	K915	V849	T788	D725	G664	Y596	E528	F466	R405	D342	E278	E210	F148	P93
D1106	G1044	E984	E916	A850	S789	I726	F665	A597	V529	R468	L407	F344	G279	L211	T149	D87
N1107	A1045	G985	L917	K851	L790	P727	L666	E598	E530	T469	R408	R345	K280	G212	P150	L88
P1108	L1046	P986	L918	L852	R791	H728	A667	E599	F531	P470	R409	R346	L281	L207	D151	T89
V1109	H1047	L987	A919	L853	V792	L729	L668	D600	N532	Y471	L410	G347	G282	Q145	P152	Y90
L1110	T1048	V988	Q920	P854	P793	S730	G669	G601	D534	R472	S411	L348	T289	Q219	G156	Q91
I1111	L1049	V989	A921	V855	P794	E731	Q670	E602	V533	R473	A412	A349	L285	Q220	R157	P93
F1112	Q1050	G990	F922	D856	G795	A732	N671	V603	S535	V474	L413	R350	G286	L221	Y158	L94
E1113	E1051	Q991	E923	D857	E796	A733	V672	A604	P536	Q475	G414	L351	G287	M222	I159	Y95
G1114	M1052	P992	V924	M858	G797	L734	L673	R605	K537	G477	P415	A352	R288	Q223	A160	A96
L1115	L1053	F993	Y925	P859	G798	R735	V674	V606	Q538	G477	G416	A352	R288	E224	S161	R97
A1116	A1105	L994	F926	H860	I799	D736	A675	D607	Q538	V478	L418	V355	L290	S225	I162	L98
S1117	L1056	K996	G927	L861	V800	L737	L676	G608	V542	V479	L418	E357	R292	V226	P163	Q99
K1118	K1056	L997	K928	D862	R801	D738	N677	N609	N545	T480	T419	R358	R292	F227	P164	L100
L1119	S1057	L997	R929	D863	R802	E739	P673	R610	L546	E481	R420	R358	F293	A228	L165	I101
	D1058	T998	K930	G864	T803	E740	P679	L611	T547	E482	E421	M359	G296	M229	P166	H102
	D1059	H999	G931	T865	V804	V743	D680	V612	F548	Y485	A423	G362	K299	R230	R168	K103
	E1060	M1000	E932	P866	R805	R744	G681	V613	P548	T487	G424	S363	D300	E232	G169	D104
	M1001	V1001	E932	P867	L806	R744	Y682	R614	F549	T487	G424	S363	D300	E232	G169	T105
	G1062	E1002	D937	D868	R807	A747	N683	Y615	L550	A488	D426	D365	E301	E234	W171	G106
	R1063	D1003	K938	V869	R808	E748	F684	E622	E551	T489	V427	S366	V302	A234	P170	L107
	K1064	K1004	R939	I870	G809	E748	E685	Y623	D553	E490	R428	L367	P303	E236	I172	L108
	A1065	M1005	E940	N871	P811	K750	A687	P624	D554	E491	D429	L367	L304	E237	D173	K109
	A1066	A1007	V941	N872	P812	P751	I688	L625	L625	D492	V430	P369	P305	E238	L174	E110
		L1007	E942	P873	G812	G752	V689	R626	R557	R493	H431	A370	T306	L238	E175	D111
	A1069	R1008	V943	L874	V813	D753	L690	R627	A558	Y494	R432	K371	T306	L241	V176	F114
	I1070	S1009	L944	G875	E814	D753	L690	R627	A558	Y494	R432	K371	T306	L241	E177	F115
	I1071	L1010	R945	V876	L815	I794	S691	F628	L559	T495	T433	L372	Y309	L242	P178	L116
	K1072	G1011	R946	P877	R816	L755	E692	Y629	N560	T495	H434	V373	Y309	R243	M179	G116
	G1073	P1012	L946	S878	P817	V756	E693	R630	G561	A497	Y435	L310	P344	R244	G180	H117
	E1074	L1013	L950	R879	G818	G757	L694	S631	S562	Q498	G436	S375	F311	G245	V181	I118
	D1075	S1014	G951	M880	V819	R758	L694	N632	N563	A499	R437	R376	A312	D246	P119	G182
	V1076	L1015	L952	N881	R820	E758	L694	N632	N563	A499	R437	R376	A312	D246	V182	L120
	P1077	I1016	V953	L882	E821	F761	R697	T635	N564	T501	T438	P377	L313	P247	S183	M121
	E1078	T1017	V953	L882	E821	F761	R697	T635	N564	T501	T438	P377	L313	P247	S183	M121
	L1079	Q1018	T958	G883	V822	K762	P699	A636	T566	P502	P440	L378	A315	P248	M184	
				Q884	V822	K762	P699	A636	T566	P502	P440	L378	A315	P248	M184	
						E764	T701	D638	A568	E504	V441	A380	P318	R249	V186	G125
														D251	M187	

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







4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	155.38Å 155.38Å 496.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.00 39.78 – 3.00	Depositor EDS
% Data completeness (in resolution range)	87.8 (40.00-3.00) 82.1 (39.78-3.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 3.01Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.234 , 0.266 0.235 , 0.261	Depositor DCC
R_{free} test set	10938 reflections (5.70%)	wwPDB-VP
Wilson B-factor (Å ²)	58.4	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 114.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.147 for h,-k,-l	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	51962	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.89% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: STD, 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	0.90	1/520 (0.2%)	1.13	2/798 (0.3%)
1	X	0.97	1/520 (0.2%)	1.14	0/798
2	H	1.48	5/387 (1.3%)	2.79	37/601 (6.2%)
2	Y	1.46	2/387 (0.5%)	2.77	38/601 (6.3%)
3	I	0.81	0/304	1.22	3/467 (0.6%)
3	Z	0.76	0/304	1.10	1/467 (0.2%)
4	A	0.73	0/1838	0.79	2/2498 (0.1%)
4	B	0.73	0/1838	0.78	4/2498 (0.2%)
4	K	0.72	0/1838	0.82	3/2498 (0.1%)
4	L	0.76	0/1838	0.79	3/2498 (0.1%)
5	C	0.77	0/8997	0.89	15/12164 (0.1%)
5	M	0.79	2/8997 (0.0%)	0.90	14/12164 (0.1%)
6	D	0.82	12/10547 (0.1%)	0.93	21/14245 (0.1%)
6	N	0.81	7/10547 (0.1%)	0.90	16/14245 (0.1%)
7	E	0.77	1/784 (0.1%)	1.06	3/1057 (0.3%)
7	O	0.81	1/784 (0.1%)	1.07	4/1057 (0.4%)
All	All	0.81	32/50430 (0.1%)	0.97	166/68656 (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	7
2	H	0	2
2	Y	0	1
6	D	0	1
All	All	0	17

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	133	ILE	N-CA	11.89	1.70	1.46
6	D	132	TYR	CA-C	9.71	1.78	1.52
2	Y	1	G	C3'-O3'	8.60	1.54	1.42
2	H	1	G	OP3-P	-7.94	1.51	1.61
6	D	456	MET	N-CA	7.80	1.61	1.46

The worst 5 of 166 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1	G	N9-C1'-C2'	25.00	146.50	114.00
2	Y	1	G	N9-C1'-C2'	20.87	141.13	114.00
2	Y	1	G	P-O3'-C3'	19.40	142.98	119.70
2	H	1	G	P-O3'-C3'	19.01	142.51	119.70
2	Y	2	A	O4'-C1'-N9	16.04	121.03	108.20

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	1	DC	Sidechain
1	G	13	DT	Sidechain
1	G	15	DC	Sidechain
1	G	16	DG	Sidechain
1	G	17	DC	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	46	0
1	X	467	0	259	45	0
2	H	347	0	174	75	0
2	Y	347	0	174	61	0
3	I	270	0	144	18	0
3	Z	270	0	144	18	0
4	A	1806	0	1861	186	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	178	0
4	K	1806	0	1861	206	0
4	L	1806	0	1861	173	0
5	C	8829	0	8933	1078	0
5	M	8829	0	8933	1061	0
6	D	10373	0	10599	1472	0
6	N	10373	0	10599	1397	0
7	E	770	0	784	124	0
7	O	770	0	784	105	0
8	D	43	0	34	6	0
8	N	43	0	31	6	0
9	D	2	0	0	0	0
9	N	2	0	0	0	0
10	D	2	0	0	0	0
10	N	2	0	0	0	0
11	D	31	0	14	5	0
11	M	31	0	14	2	0
12	A	78	0	0	13	0
12	B	117	0	0	29	0
12	C	408	0	0	103	0
12	D	531	0	0	107	0
12	E	34	0	0	17	0
12	G	39	0	0	6	0
12	H	22	0	0	6	0
12	I	31	0	0	3	0
12	K	81	0	0	26	0
12	L	95	0	0	12	0
12	M	396	0	0	100	0
12	N	510	0	0	120	0
12	O	53	0	0	16	0
12	X	31	0	0	5	0
12	Y	26	0	0	3	0
12	Z	18	0	0	3	0
All	All	51962	0	49323	5743	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 58.

The worst 5 of 5743 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)
6:D:133:ILE:N	6:D:133:ILE:CA	1.70	1.55
6:D:132:TYR:C	6:D:132:TYR:CA	1.78	1.49
7:E:92:LEU:HD23	12:E:113:HOH:O	1.25	1.32
2:Y:2:A:OP2	6:N:671:LYS:NZ	1.72	1.21
6:D:87:ARG:HD3	6:D:524:LEU:HD11	1.30	1.12

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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%)	208 (92%)	16 (7%)	3 (1%)	13	49
4	B	227/315 (72%)	208 (92%)	15 (7%)	4 (2%)	9	40
4	K	227/315 (72%)	208 (92%)	16 (7%)	3 (1%)	13	49
4	L	227/315 (72%)	206 (91%)	18 (8%)	3 (1%)	13	49
5	C	1117/1119 (100%)	922 (82%)	136 (12%)	59 (5%)	2	13
5	M	1117/1119 (100%)	919 (82%)	137 (12%)	61 (6%)	2	12
6	D	1308/1524 (86%)	1104 (84%)	145 (11%)	59 (4%)	3	16
6	N	1308/1524 (86%)	1099 (84%)	158 (12%)	51 (4%)	3	19
7	E	93/99 (94%)	73 (78%)	13 (14%)	7 (8%)	1	5
7	O	93/99 (94%)	73 (78%)	12 (13%)	8 (9%)	1	4
All	All	5944/6744 (88%)	5020 (84%)	666 (11%)	258 (4%)	3	17

5 of 258 Ramachandran outliers are listed below:

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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	C	152	PRO
5	C	156	GLY

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%)	159 (79%)	43 (21%)	1	6
4	B	202/273 (74%)	162 (80%)	40 (20%)	1	8
4	K	202/273 (74%)	155 (77%)	47 (23%)	1	4
4	L	202/273 (74%)	153 (76%)	49 (24%)	1	3
5	C	941/941 (100%)	723 (77%)	218 (23%)	1	4
5	M	941/941 (100%)	714 (76%)	227 (24%)	1	3
6	D	1111/1279 (87%)	875 (79%)	236 (21%)	1	6
6	N	1111/1279 (87%)	863 (78%)	248 (22%)	1	5
7	E	84/88 (96%)	66 (79%)	18 (21%)	1	6
7	O	84/88 (96%)	67 (80%)	17 (20%)	1	7
All	All	5080/5708 (89%)	3937 (78%)	1143 (22%)	1	4

5 of 1143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	D	1345	GLU
4	L	181	VAL
6	N	1156	LEU
6	D	1465	ASN
4	K	143	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 151 such sidechains are listed below:

Mol	Chain	Res	Type
6	D	1441	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	M	91	GLN
6	N	1124	GLN
7	E	29	GLN
4	L	16	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	H	16/16 (100%)	11 (68%)	6 (37%)
2	Y	16/16 (100%)	11 (68%)	7 (43%)
All	All	32/32 (100%)	22 (68%)	13 (40%)

5 of 22 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 13 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	H	12	G
2	Y	1	G
2	Y	9	G
2	H	9	G
2	Y	8	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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
11	APC	D	5999	10	28,33,33	1.38	4 (14%)	27,52,52	1.64	6 (22%)
8	STD	D	7001	-	43,47,47	6.93	28 (65%)	46,73,73	2.66	12 (26%)
11	APC	M	6999	10	28,33,33	1.28	4 (14%)	27,52,52	1.63	6 (22%)
8	STD	N	8001	-	43,47,47	6.83	28 (65%)	46,73,73	2.66	11 (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
11	APC	D	5999	10	-	0/15/38/38	0/3/3/3
8	STD	D	7001	-	-	0/31/101/101	0/4/5/5
11	APC	M	6999	10	-	0/15/38/38	0/3/3/3
8	STD	N	8001	-	-	0/31/101/101	0/4/5/5

The worst 5 of 64 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	7001	STD	O5-C19	-26.89	1.19	1.43
8	N	8001	STD	O5-C19	-26.23	1.19	1.43
8	D	7001	STD	C23-C21	-15.49	1.18	1.53
8	N	8001	STD	C23-C21	-14.86	1.20	1.53
8	N	8001	STD	C18-C16	-12.80	1.25	1.53

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	N	8001	STD	O8-C17-C30	-5.72	105.96	111.66
8	D	7001	STD	O8-C17-C30	-5.53	106.15	111.66

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	M	6999	APC	C1'-N9-C4	-4.41	119.01	126.64
11	D	5999	APC	C1'-N9-C4	-4.28	119.24	126.64
11	M	6999	APC	PG-O3B-PB	-4.06	118.17	132.40

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	D	5999	APC	5	0
8	D	7001	STD	6	0
11	M	6999	APC	2	0
8	N	8001	STD	6	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

6.1 Protein, DNA and RNA chains

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%)	-0.74	0 100 100	23, 43, 66, 69	0
1	X	23/23 (100%)	-0.73	0 100 100	9, 37, 77, 92	0
2	H	16/16 (100%)	-0.49	0 100 100	24, 52, 92, 93	0
2	Y	16/16 (100%)	-0.49	0 100 100	25, 43, 96, 99	0
3	I	13/14 (92%)	-0.79	0 100 100	39, 55, 76, 77	0
3	Z	13/14 (92%)	-0.87	0 100 100	50, 61, 75, 79	0
4	A	229/315 (72%)	-0.57	0 100 100	31, 58, 73, 77	0
4	B	229/315 (72%)	-0.57	1 (0%) 92 78	34, 62, 75, 83	0
4	K	229/315 (72%)	-0.57	0 100 100	30, 57, 71, 76	0
4	L	229/315 (72%)	-0.49	0 100 100	37, 62, 76, 87	0
5	C	1119/1119 (100%)	-0.62	1 (0%) 95 89	7, 54, 77, 90	0
5	M	1119/1119 (100%)	-0.61	2 (0%) 94 86	18, 54, 76, 90	0
6	D	1314/1524 (86%)	-0.56	5 (0%) 92 78	11, 56, 79, 89	0
6	N	1314/1524 (86%)	-0.57	3 (0%) 94 86	8, 56, 76, 91	0
7	E	95/99 (95%)	-0.68	0 100 100	42, 58, 67, 71	0
7	O	95/99 (95%)	-0.59	0 100 100	33, 59, 75, 80	0
All	All	6076/6850 (88%)	-0.58	12 (0%) 94 86	7, 56, 77, 99	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	N	416	ALA	4.2
5	C	1025	ALA	3.8
6	D	188	GLY	3.1
6	N	429	SER	2.9
6	D	391	ALA	2.7

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	STD	N	8001	43/43	0.96	0.17	14,32,53,55	0
8	STD	D	7001	43/43	0.96	0.17	11,24,27,28	0
9	ZN	D	7058	1/1	0.97	0.08	87,87,87,87	0
9	ZN	N	7158	1/1	0.97	0.06	70,70,70,70	0
11	APC	D	5999	31/31	0.97	0.15	30,38,64,65	0
10	MG	D	9001	1/1	0.98	0.08	22,22,22,22	0
11	APC	M	6999	31/31	0.98	0.14	35,45,57,58	0
10	MG	N	9004	1/1	0.99	0.09	27,27,27,27	0
9	ZN	N	8212	1/1	0.99	0.10	54,54,54,54	0
9	ZN	D	8112	1/1	0.99	0.07	58,58,58,58	0
10	MG	D	9002	1/1	0.99	0.16	25,25,25,25	0
10	MG	N	9003	1/1	0.99	0.09	21,21,21,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.