



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 30, 2026 – 09:40 AM UTC

PDB ID : 4YLN / pdb_00004yln
Title : E. coli Transcription Initiation Complex - 17-bp spacer and 4-nt RNA
Authors : Zuo, Y.; Steitz, T.A.
Deposited on : 2015-03-05
Resolution : 5.50 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

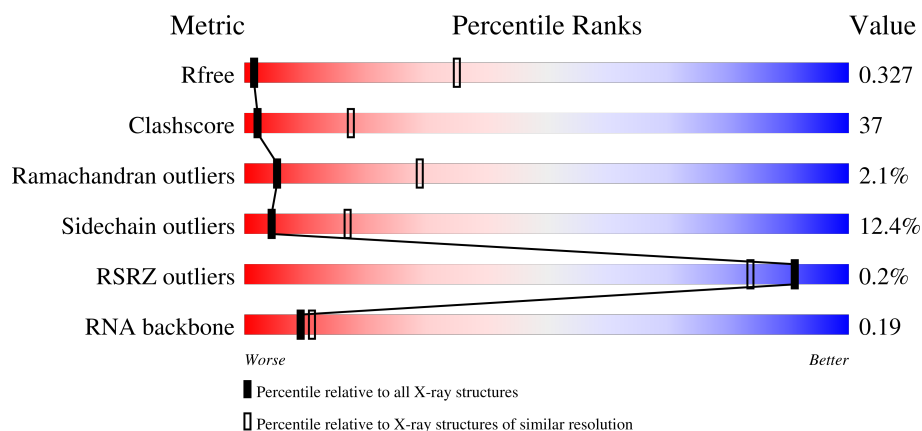
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 5.50 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}	180053	1088 (7.00-4.00)
Clashscore	190562	1147 (7.00-4.00)
Ramachandran outliers	187476	1012 (7.02-3.98)
Sidechain outliers	187428	1023 (7.04-3.96)
RSRZ outliers	180081	1081 (7.00-4.00)
RNA backbone	3983	1052 (7.80-3.10)

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	A	242	
1	B	242	
1	G	242	
1	H	242	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	242	
1	N	242	
2	C	1342	
2	I	1342	
2	O	1342	
3	D	1407	
3	J	1407	
3	P	1407	
4	E	90	
4	K	90	
4	Q	90	
5	F	628	
5	L	628	
5	R	628	
6	1	49	
6	4	49	
6	7	49	
7	2	49	
7	5	49	
7	8	49	
8	3	4	
8	6	4	
8	9	4	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
9	ZN	J	1502	-	-	X	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 94608 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 protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	B	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			
1	G	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	H	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			
1	M	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	N	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	ALA	-	expression tag	UNP A7ZSI4
A	-5	HIS	-	expression tag	UNP A7ZSI4
A	-4	HIS	-	expression tag	UNP A7ZSI4
A	-3	HIS	-	expression tag	UNP A7ZSI4
A	-2	HIS	-	expression tag	UNP A7ZSI4
A	-1	HIS	-	expression tag	UNP A7ZSI4
A	0	HIS	-	expression tag	UNP A7ZSI4
B	-6	ALA	-	expression tag	UNP A7ZSI4
B	-5	HIS	-	expression tag	UNP A7ZSI4
B	-4	HIS	-	expression tag	UNP A7ZSI4
B	-3	HIS	-	expression tag	UNP A7ZSI4
B	-2	HIS	-	expression tag	UNP A7ZSI4
B	-1	HIS	-	expression tag	UNP A7ZSI4
B	0	HIS	-	expression tag	UNP A7ZSI4
G	-6	ALA	-	expression tag	UNP A7ZSI4
G	-5	HIS	-	expression tag	UNP A7ZSI4
G	-4	HIS	-	expression tag	UNP A7ZSI4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-3	HIS	-	expression tag	UNP A7ZSI4
G	-2	HIS	-	expression tag	UNP A7ZSI4
G	-1	HIS	-	expression tag	UNP A7ZSI4
G	0	HIS	-	expression tag	UNP A7ZSI4
H	-6	ALA	-	expression tag	UNP A7ZSI4
H	-5	HIS	-	expression tag	UNP A7ZSI4
H	-4	HIS	-	expression tag	UNP A7ZSI4
H	-3	HIS	-	expression tag	UNP A7ZSI4
H	-2	HIS	-	expression tag	UNP A7ZSI4
H	-1	HIS	-	expression tag	UNP A7ZSI4
H	0	HIS	-	expression tag	UNP A7ZSI4
M	-6	ALA	-	expression tag	UNP A7ZSI4
M	-5	HIS	-	expression tag	UNP A7ZSI4
M	-4	HIS	-	expression tag	UNP A7ZSI4
M	-3	HIS	-	expression tag	UNP A7ZSI4
M	-2	HIS	-	expression tag	UNP A7ZSI4
M	-1	HIS	-	expression tag	UNP A7ZSI4
M	0	HIS	-	expression tag	UNP A7ZSI4
N	-6	ALA	-	expression tag	UNP A7ZSI4
N	-5	HIS	-	expression tag	UNP A7ZSI4
N	-4	HIS	-	expression tag	UNP A7ZSI4
N	-3	HIS	-	expression tag	UNP A7ZSI4
N	-2	HIS	-	expression tag	UNP A7ZSI4
N	-1	HIS	-	expression tag	UNP A7ZSI4
N	0	HIS	-	expression tag	UNP A7ZSI4

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			
2	I	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			
2	O	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			
3	P	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	K	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	Q	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			
5	L	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			
5	R	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-14	MET	-	expression tag	UNP P00579
F	-13	ARG	-	expression tag	UNP P00579
F	-12	GLY	-	expression tag	UNP P00579
F	-11	SER	-	expression tag	UNP P00579
F	-10	HIS	-	expression tag	UNP P00579
F	-9	HIS	-	expression tag	UNP P00579
F	-8	HIS	-	expression tag	UNP P00579
F	-7	HIS	-	expression tag	UNP P00579
F	-6	HIS	-	expression tag	UNP P00579
F	-5	HIS	-	expression tag	UNP P00579
F	-4	THR	-	expression tag	UNP P00579
F	-3	ASP	-	expression tag	UNP P00579
F	-2	GLN	-	expression tag	UNP P00579
F	-1	PHE	-	expression tag	UNP P00579

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	THR	-	expression tag	UNP P00579
L	-14	MET	-	expression tag	UNP P00579
L	-13	ARG	-	expression tag	UNP P00579
L	-12	GLY	-	expression tag	UNP P00579
L	-11	SER	-	expression tag	UNP P00579
L	-10	HIS	-	expression tag	UNP P00579
L	-9	HIS	-	expression tag	UNP P00579
L	-8	HIS	-	expression tag	UNP P00579
L	-7	HIS	-	expression tag	UNP P00579
L	-6	HIS	-	expression tag	UNP P00579
L	-5	HIS	-	expression tag	UNP P00579
L	-4	THR	-	expression tag	UNP P00579
L	-3	ASP	-	expression tag	UNP P00579
L	-2	GLN	-	expression tag	UNP P00579
L	-1	PHE	-	expression tag	UNP P00579
L	0	THR	-	expression tag	UNP P00579
R	-14	MET	-	expression tag	UNP P00579
R	-13	ARG	-	expression tag	UNP P00579
R	-12	GLY	-	expression tag	UNP P00579
R	-11	SER	-	expression tag	UNP P00579
R	-10	HIS	-	expression tag	UNP P00579
R	-9	HIS	-	expression tag	UNP P00579
R	-8	HIS	-	expression tag	UNP P00579
R	-7	HIS	-	expression tag	UNP P00579
R	-6	HIS	-	expression tag	UNP P00579
R	-5	HIS	-	expression tag	UNP P00579
R	-4	THR	-	expression tag	UNP P00579
R	-3	ASP	-	expression tag	UNP P00579
R	-2	GLN	-	expression tag	UNP P00579
R	-1	PHE	-	expression tag	UNP P00579
R	0	THR	-	expression tag	UNP P00579

- Molecule 6 is a DNA chain called NT strand DNA (49-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	1	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			
6	4	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			
6	7	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			

- Molecule 7 is a DNA chain called T strand DNA (49-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	2	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			
7	5	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			
7	8	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			

- Molecule 8 is a RNA chain called RNA (5'-D(*(GTP))-R(P*AP*GP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	3	4	Total	C	N	O	P	0	0	0
			97	39	17	35	6			
8	6	4	Total	C	N	O	P	0	0	0
			97	39	17	35	6			
8	9	4	Total	C	N	O	P	0	0	0
			97	39	17	35	6			

- Molecule 9 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

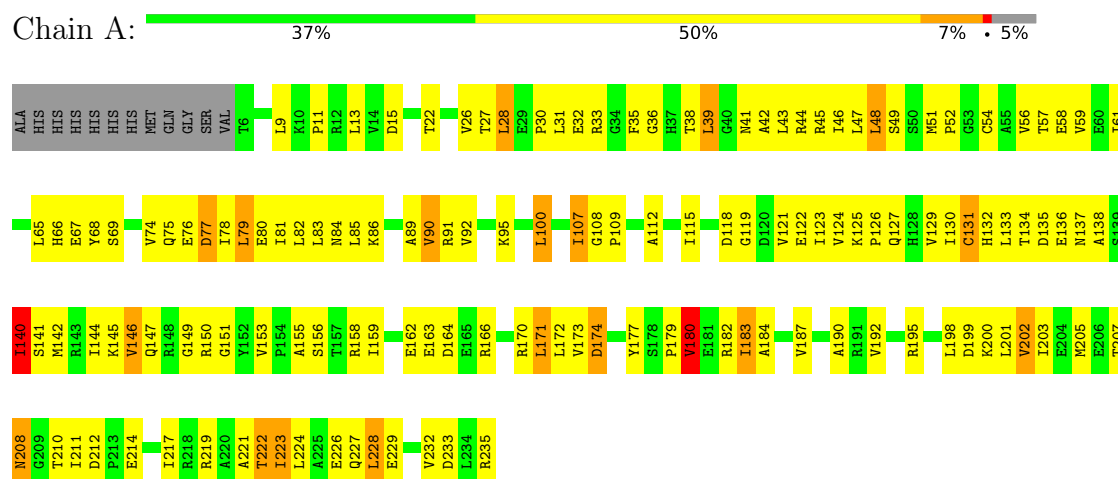
- Molecule 10 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	1	Total	Mg	0	0
			1	1		
10	6	1	Total	Mg	0	0
			1	1		
10	P	1	Total	Mg	0	0
			1	1		

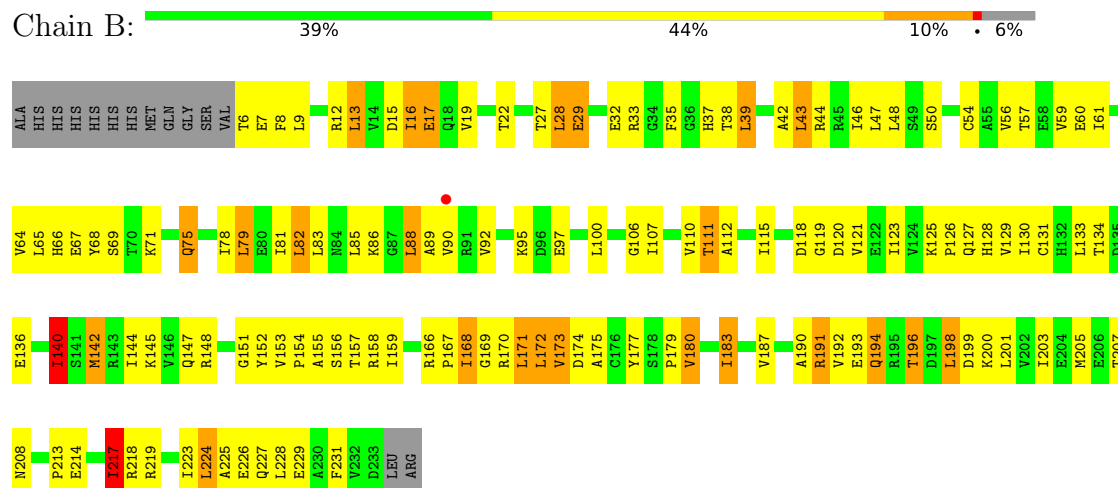
3 Residue-property plots

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: DNA-directed RNA polymerase subunit alpha

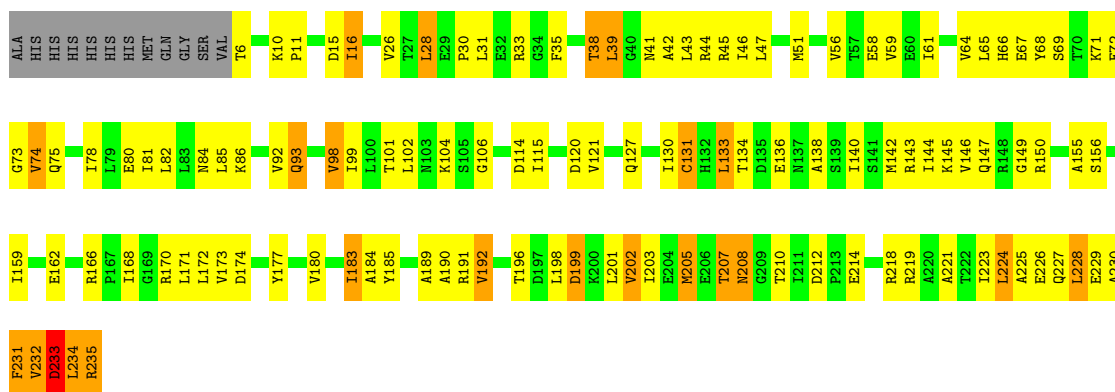


• Molecule 1: DNA-directed RNA polymerase subunit alpha



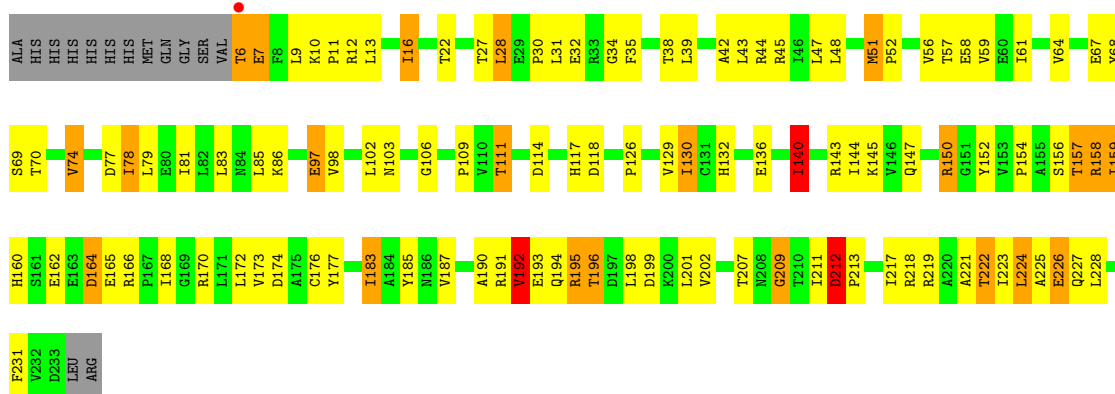
• Molecule 1: DNA-directed RNA polymerase subunit alpha





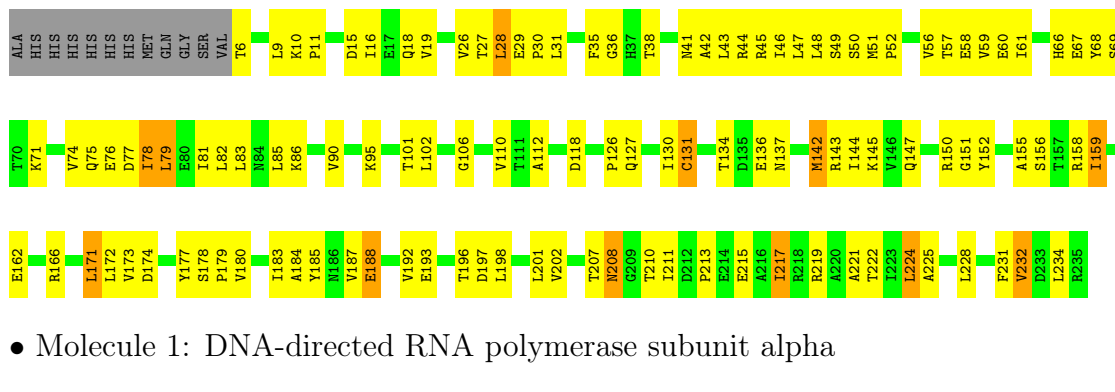
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain H: 47% 37% 9% 6%



• Molecule 1: DNA-directed RNA polymerase subunit alpha

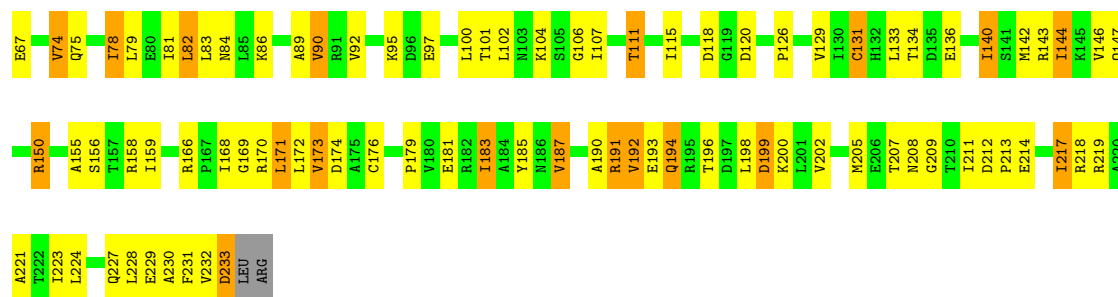
Chain M: 47% 43% 5% 5%



• Molecule 1: DNA-directed RNA polymerase subunit alpha

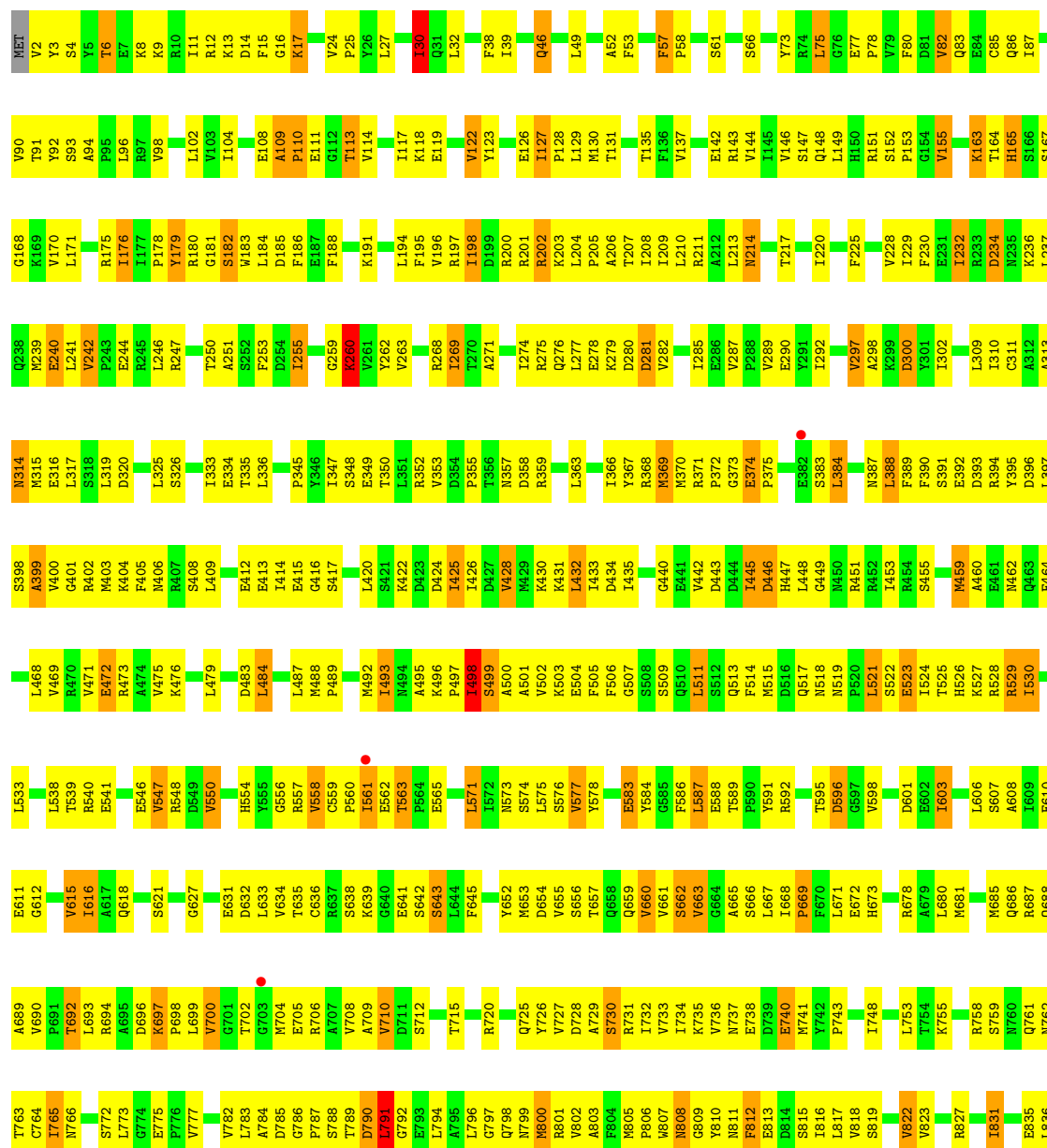
Chain N: 45% 40% 10% 6%





• Molecule 2: DNA-directed RNA polymerase subunit beta

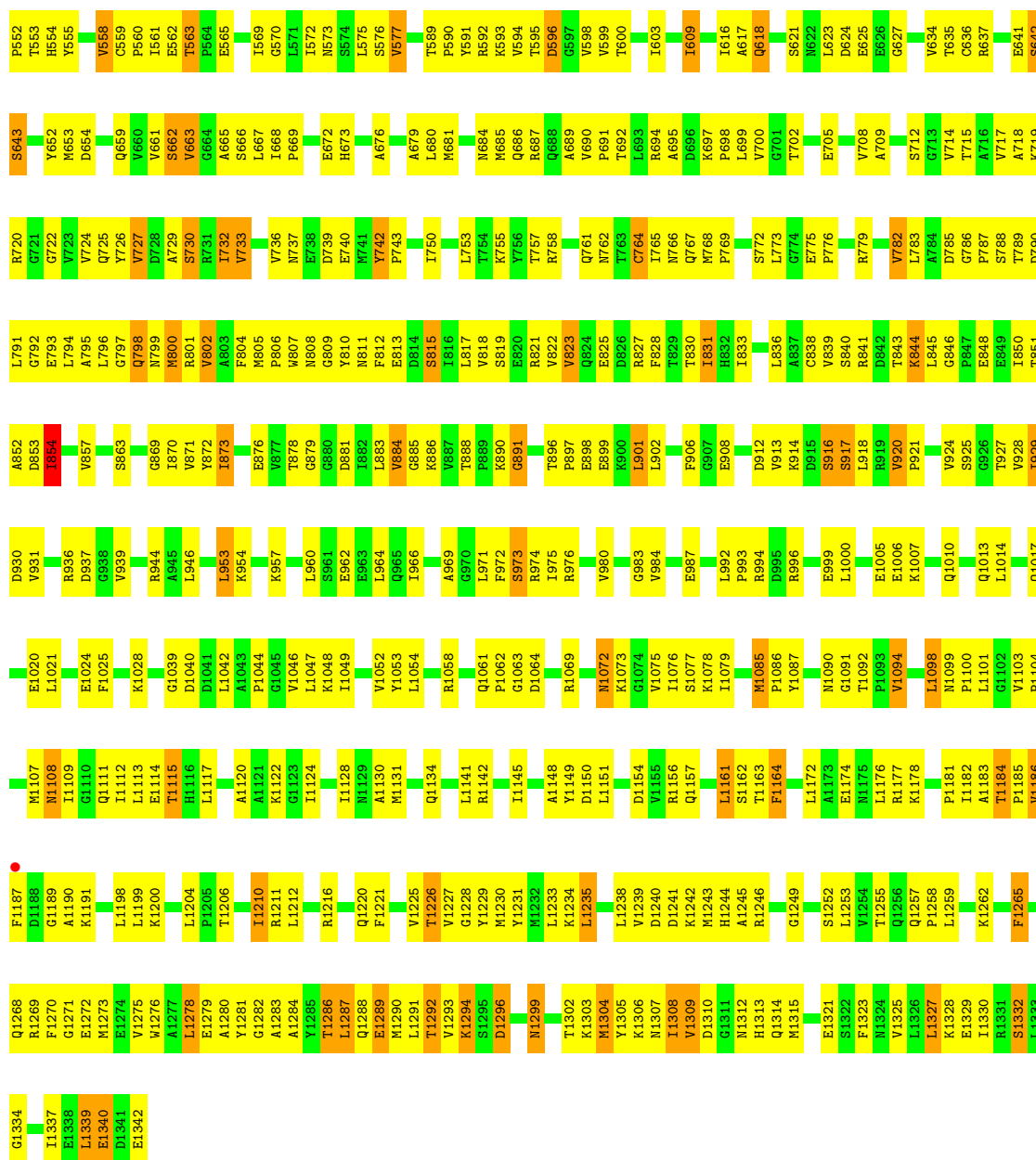
Chain C: 44% 46% 9%





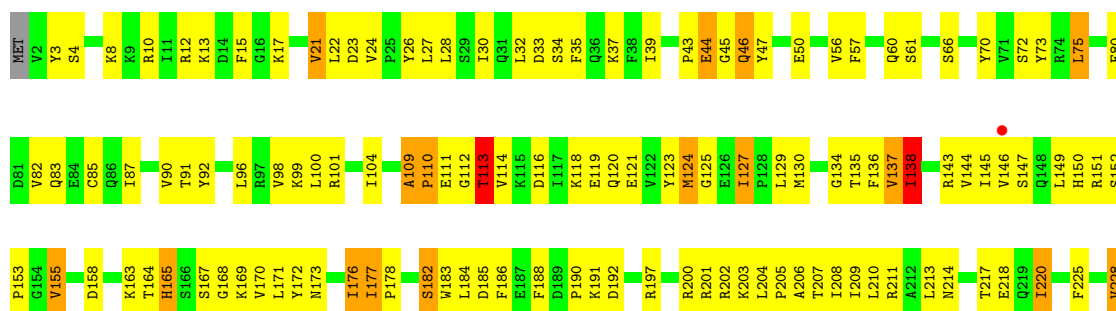
Category	Percentage
Very bad	46%
Bad	46%
Good	7%



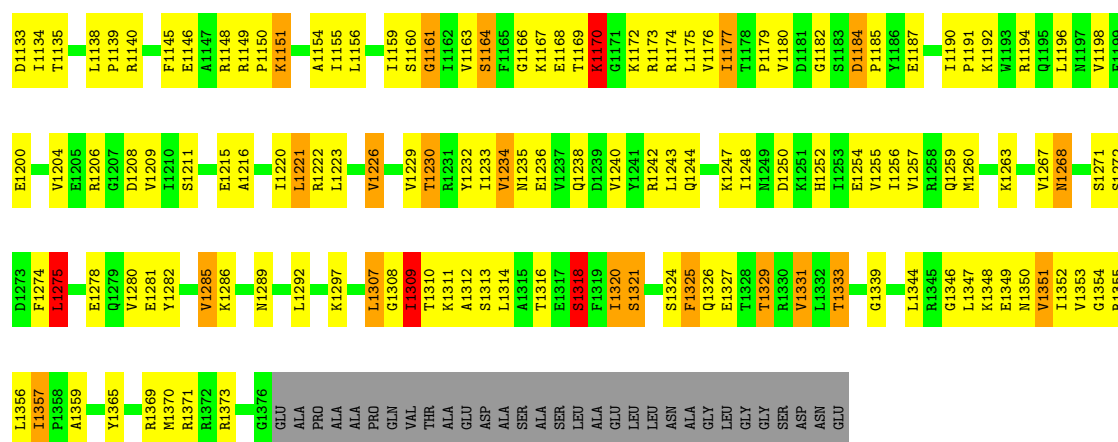


● Molecule 2: DNA-directed RNA polymerase subunit beta

Chain O: 48% 44% 8%

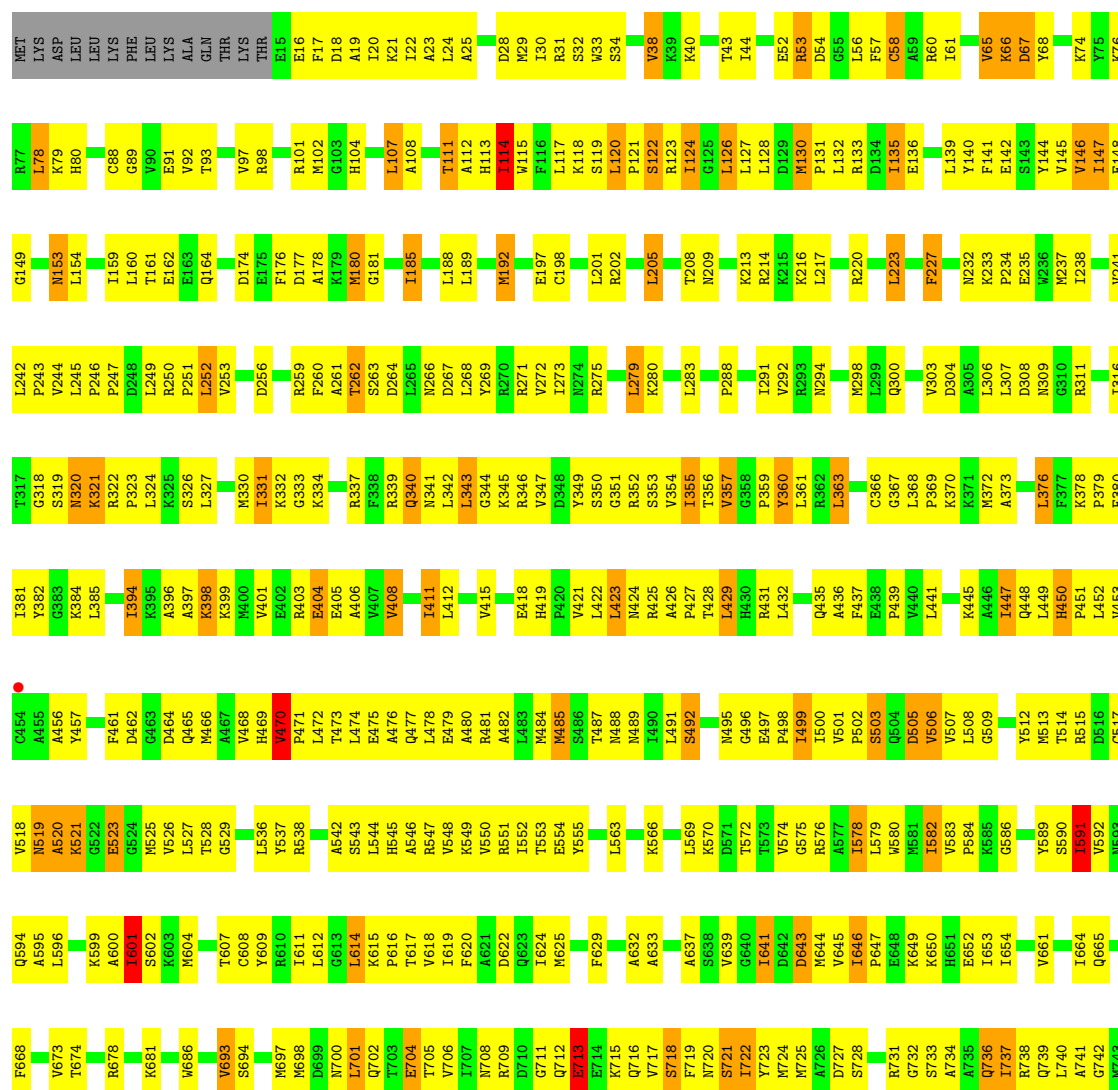


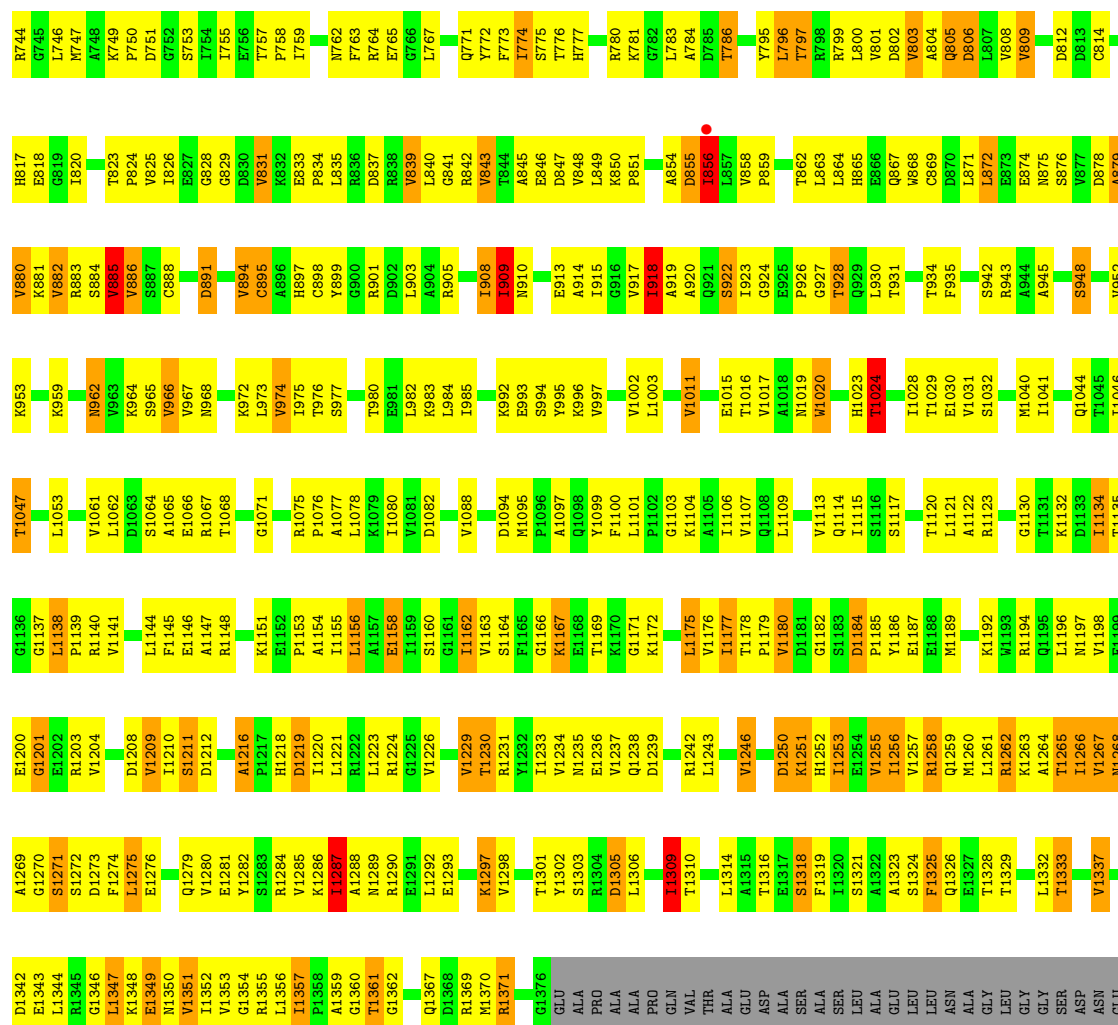
V966	R883	D806	R738	L807	Q739	S884	S1057	V967	S884	S1058	L807	Q739	S884	S1057	V966
L973	V885	L807	L740	V808	L740	V885	L1059	L972	V885	L1058	L740	Q739	S884	S1057	V967
V974	C888	E811	G742	T810	A741	V886	V1060	V973	V887	V1059	G742	A741	V886	V1058	V974
L975	C888	E811	M743	D812	R744	Q665	V1061	L975	R744	V1060	M743	R744	Q665	V1061	L975
S977	D891	D813	G745	D813	G745	F668	D1062	T976	F668	D1063	G745	R514	S977	D1062	S977
T980	F892	C814	L746	G815	L746	G671	S1064	T976	L746	S1063	L746	R514	S977	D1062	T980
V981	G893	T816	A748	T816	A748	V673	E1065	V980	V673	E1066	A748	R514	V980	E1065	V981
L982	C895	H817	K749	H817	K749	T674	R1067	L982	T674	R1067	K749	L452	L982	R1067	L982
	C898	G819	P750	G819	P750	M604	G1071		M604	G1071	P750	L385		G1071	
	A904	M822	S753	M822	S753	K603	L1074	D986	K603	L1074	S753	E386	R101	L1074	D986
E987	I909	P824	T755	P824	T755	E686	P1076	E987	E686	P1076	T755	E386	M102	P1076	E987
T991	N910	I826	P758	I826	P758	M697	A1077	T991	M697	A1077	P758	E386	E183	A1077	T991
S994	E913	G828	T760	G828	T760	M698	K1079	S994	M698	K1079	T760	E386	E183	K1079	S994
V997	A914	G829	A761	G829	A761	M699	I1080	V997	M699	I1080	A761	E386	E183	I1080	V997
F998	I915	N762	N762	N762	N762	N700	V1081	F998	N700	V1081	N762	E386	E183	V1081	F998
	I918	D830	N762	D830	N762	L612			L612		N762	E386	E183		
Y999	A919	V831	F763	V831	F763	G613	G1085	Y999	G613	G1086	F763	E386	E183	G1085	Y999
V1002	E833	K832	R764	E833	R764	T702	D1087	V1002	T702	D1087	R764	E386	E183	D1087	V1002
L1003	P834	L835	N768	P834	L835	E704	V1088	L1003	E704	V1088	N768	E386	E183	V1088	L1003
V1011	R924	R836	V768	R924	R836	T705	I1090	V1011	T705	I1090	V768	E386	E183	I1090	V1011
	E925	D837	L770	E925	D837	R616	P1091		R616	P1091	L770	E386	E183	P1091	
V1017	P926	R838	Q771	P926	R838	N708	D1094	V1017	N708	D1094	Q771	E386	E183	D1094	V1017
V1020	G927	V839	T772	G927	V839	F620	P1096	V1020	F620	P1096	T772	E386	E183	P1096	V1020
D1021	T928	L840	F773	T928	L840	D710	V1097	D1021	D710	V1097	F773	E386	E183	V1097	D1021
P1022	Q929	G841	I774	Q929	G841	D622	A1097	P1022	D622	A1097	I774	E386	E183	A1097	P1022
H1023	L830	R842	S775	L830	R842	G712	Q1098	H1023	G712	Q1098	S775	E386	E183	Q1098	H1023
T1024	M932	E946	H777	M932	E946	E713	F1100	T1024	E713	F1100	H777	E386	E183	F1100	T1024
M1025	R933	D847	R777	R933	D847	M625	L1101	M1025	M625	L1101	R777	E386	E183	L1101	M1025
P1026	V848	L849	R780	V848	L849	T627		P1026	T627		R780	E386	E183		
V1027	F935	L849	R780	F935	L849	V717		V1027	V717		R780	E386	E183		
I1028	R836			R836		S718	K1104	I1028	S718	K1104		E386	E183		
T1029	I937	A854	A787	T1029	I937	A630	A1105	T1029	A630	A1105	A787	E386	E183	A1105	T1029
E1030		L788	R788	E1030		N720	V1106	E1030	N720	V1106	R788	E386	E183	V1106	E1030
V1031	A941	I856	K789	V1031	A941	S721	V1107	V1031	S721	V1107	K789	E386	E183	V1107	V1031
S1032			T790	S1032		T722		S1032	T722		T790	E386	E183		
	A946	R860		A946	R860	Y723	E1110		Y723	E1110		E386	E183		
V1035	E947		S793	V1035	E947	M724	Q1114	V1035	M724	Q1114	S793	E386	E183	Q1114	V1035
R1036	S948	L863	G794	R1036	S948	A726		R1036	A726		G794	E386	E183		
S949	I950	E873	Y795	S949	I950	R796	S1117	S949	R796	S1117	Y795	E386	E183	S1117	S949
M1040				M1040		T797	G1118		T797	G1118		E386	E183		
Q1044	I958	N875	R798	Q1044	I958	S874	D1119	Q1044	S874	D1119	R798	E386	E183	D1119	Q1044
T1045	K859	S876	R799	T1045	K859	G640	T1120	T1045	G640	T1120	R799	E386	E183	T1120	T1045
I1046	L980	V877	L800	I1046	L980	I641	A1121	I1046	I641	A1121	L800	E386	E183	A1121	I1046
												E386	E183		
Q1049	N962	A879	D802	Q1049	N962	V645	L1122	Q1049	V645	L1122	D802	E386	E183	L1122	Q1049
L1123	V963	R880	A795	L1123	V963	I646	I1124	L1123	I646	I1124	A795	E386	E183	I1124	L1123
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		
												E386	E183		



• Molecule 3: DNA-directed RNA polymerase subunit beta'

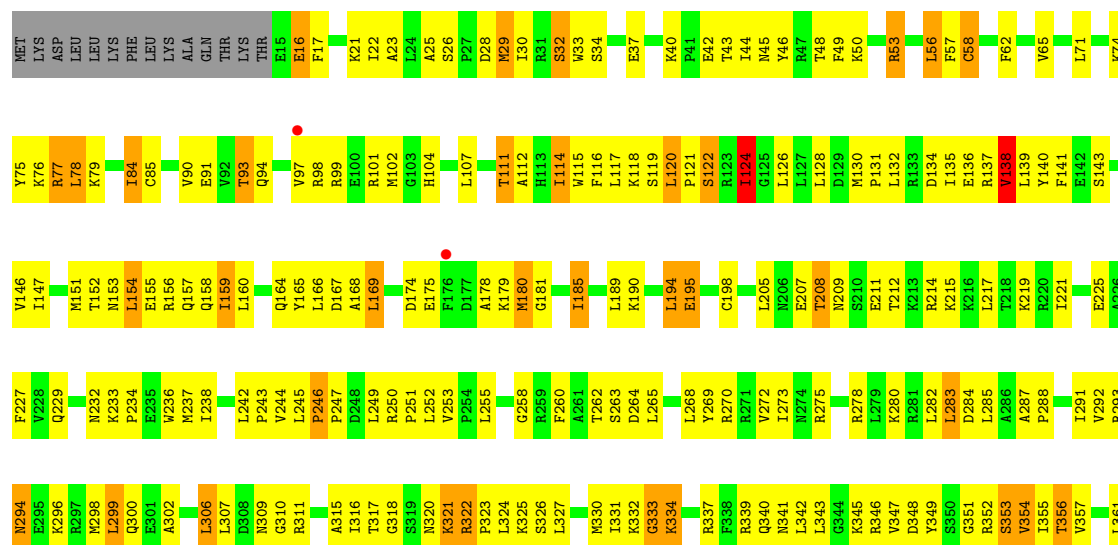
Chain J: 39% 46% 10% ..



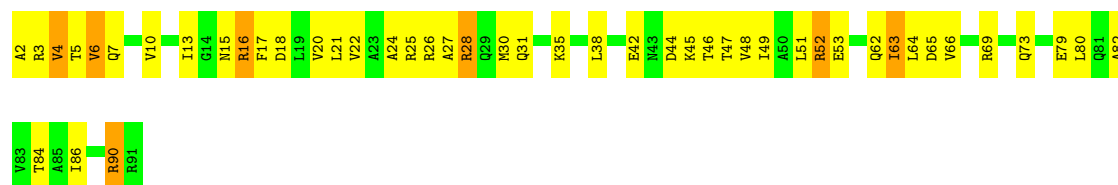


● Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain P: 43% 42% 10% ..







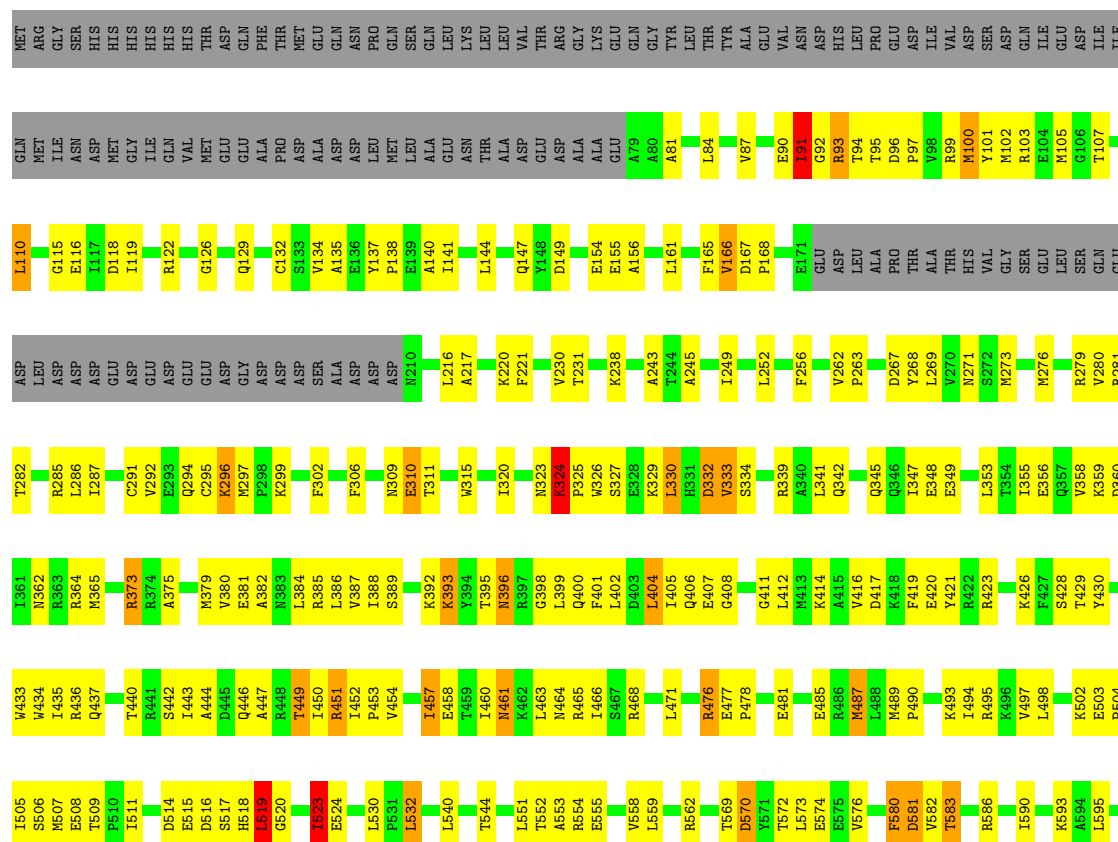
- Molecule 4: DNA-directed RNA polymerase subunit omega

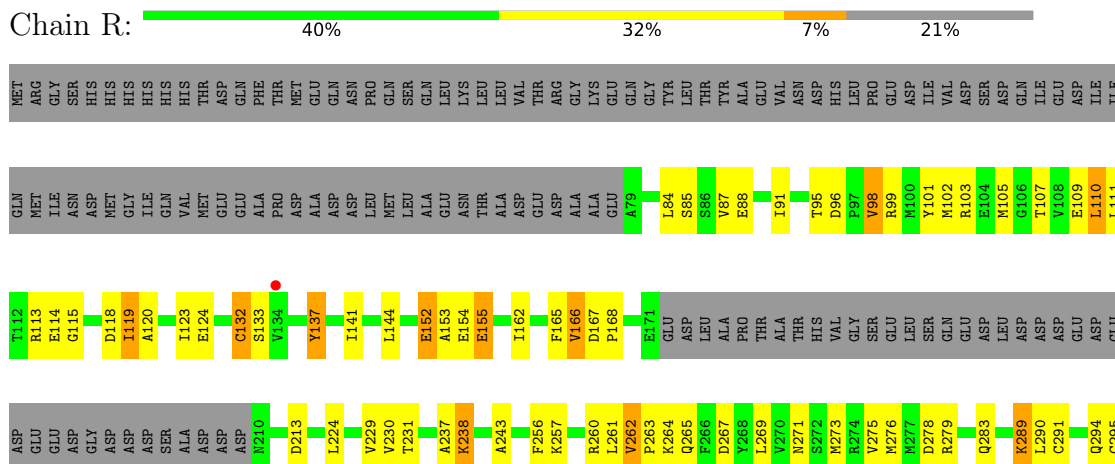


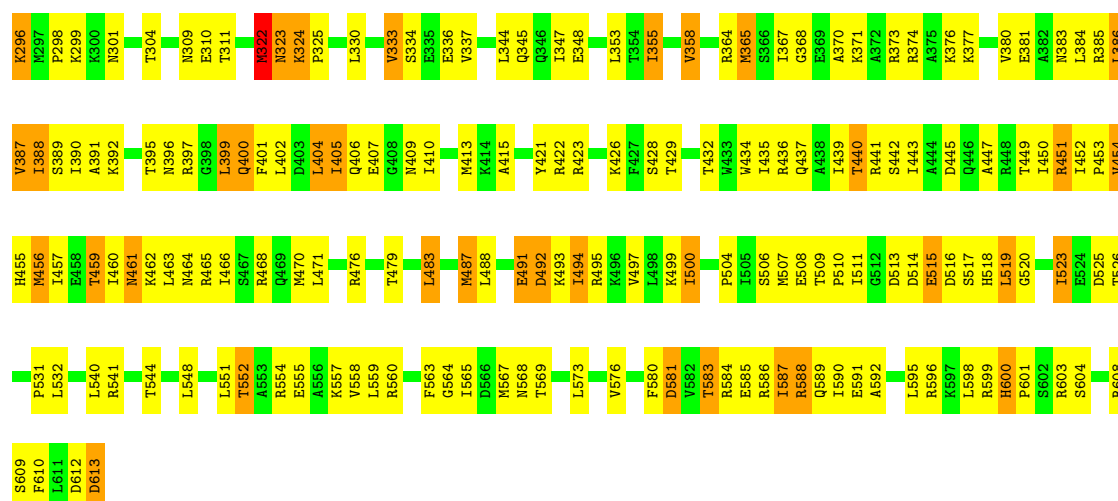
- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 5: RNA polymerase sigma factor RpoD







• Molecule 6: NT strand DNA (49-MER)



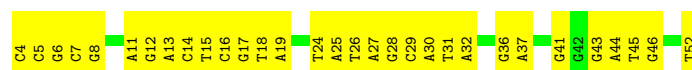
• Molecule 6: NT strand DNA (49-MER)



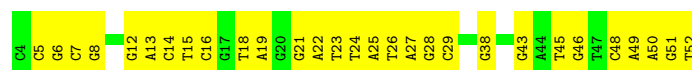
• Molecule 6: NT strand DNA (49-MER)



• Molecule 7: T strand DNA (49-MER)

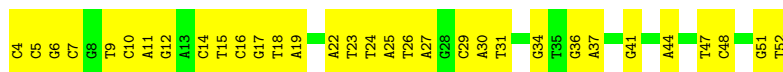


• Molecule 7: T strand DNA (49-MER)



- Molecule 7: T strand DNA (49-MER)

Chain 8:  35% 65%



- Molecule 8: RNA (5'-D(*(GTP))-R(P*AP*GP*U)-3')

Chain 3:  75% 25%



- Molecule 8: RNA (5'-D(*(GTP))-R(P*AP*GP*U)-3')

Chain 6:  50% 50%



- Molecule 8: RNA (5'-D(*(GTP))-R(P*AP*GP*U)-3')

Chain 9:  75% 25%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	237.40Å 206.05Å 248.69Å 90.00° 116.55° 90.00°	Depositor
Resolution (Å)	39.90 – 5.50 39.90 – 5.50	Depositor EDS
% Data completeness (in resolution range)	99.3 (39.90-5.50) 99.3 (39.90-5.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.20 (at 5.37Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.245 , 0.328 0.245 , 0.327	Depositor DCC
R_{free} test set	3459 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	268.1	Xtriage
Anisotropy	0.597	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 190.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	0.045 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	94608	wwPDB-VP
Average B, all atoms (Å ²)	219.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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: GTP, MG, ZN

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	A	0.82	0/1809	1.23	6/2450 (0.2%)
1	B	0.77	0/1789	1.18	3/2425 (0.1%)
1	G	0.80	0/1809	1.17	6/2450 (0.2%)
1	H	0.78	0/1789	1.20	7/2425 (0.3%)
1	M	0.75	0/1809	1.13	4/2450 (0.2%)
1	N	0.80	1/1789 (0.1%)	1.18	5/2425 (0.2%)
2	C	0.75	0/10745	1.17	42/14499 (0.3%)
2	I	0.73	0/10745	1.15	46/14499 (0.3%)
2	O	0.72	0/10745	1.12	35/14499 (0.2%)
3	D	0.77	1/10729 (0.0%)	1.14	37/14487 (0.3%)
3	J	0.82	1/10729 (0.0%)	1.19	52/14487 (0.4%)
3	P	0.77	2/10729 (0.0%)	1.15	55/14487 (0.4%)
4	E	0.71	0/710	1.08	1/956 (0.1%)
4	K	0.75	0/710	1.11	1/956 (0.1%)
4	Q	0.71	0/710	1.05	0/956
5	F	0.67	0/4076	1.05	9/5482 (0.2%)
5	L	0.71	0/4076	1.09	14/5482 (0.3%)
5	R	0.72	1/4076 (0.0%)	1.10	11/5482 (0.2%)
6	1	0.40	0/1115	0.67	0/1718
6	4	0.35	0/1112	0.64	0/1706
6	7	0.37	0/1114	0.64	0/1714
7	2	0.37	0/1134	0.62	0/1744
7	5	0.34	0/1134	0.59	0/1744
7	8	0.37	0/1136	0.60	0/1752
8	3	0.51	0/72	0.82	0/110
8	6	0.47	0/72	0.90	0/110
8	9	0.52	0/72	0.88	0/110
All	All	0.74	6/96535 (0.0%)	1.11	334/131605 (0.3%)

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
3	D	0	1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	1340	LYS	CB-CG	6.02	1.70	1.52
1	N	233	ASP	CG-OD1	5.58	1.35	1.25
3	P	1046	ILE	N-CA	5.48	1.52	1.46
5	R	322	MET	SD-CE	5.44	1.93	1.79
3	D	468	VAL	CA-C	-5.25	1.46	1.52

The worst 5 of 334 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	J	737	ILE	CB-CA-C	-12.69	95.64	111.88
3	D	737	ILE	CB-CA-C	-11.39	97.30	111.88
3	J	803	VAL	CB-CA-C	-11.18	97.38	112.02
3	D	774	ILE	CB-CA-C	-10.28	98.58	112.04
3	P	803	VAL	CB-CA-C	-9.86	99.11	112.02

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	671	GLY	Peptide

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	A	1787	0	1813	218	0
1	B	1767	0	1789	179	0
1	G	1787	0	1812	176	0
1	H	1767	0	1789	150	0
1	M	1787	0	1813	179	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	N	1767	0	1789	143	0
2	C	10576	0	10591	878	0
2	I	10576	0	10591	865	0
2	O	10576	0	10591	792	0
3	D	10568	0	10782	874	3
3	J	10568	0	10780	1095	2
3	P	10568	0	10780	924	0
4	E	708	0	719	42	0
4	K	708	0	719	49	0
4	Q	708	0	719	37	0
5	F	4022	0	4083	250	0
5	L	4022	0	4083	272	0
5	R	4022	0	4083	287	0
6	1	996	0	554	70	1
6	4	996	0	557	74	0
6	7	996	0	555	75	0
7	2	1012	0	556	62	0
7	5	1012	0	556	59	0
7	8	1012	0	554	64	0
8	3	97	0	44	8	0
8	6	97	0	44	8	0
8	9	97	0	44	6	0
9	D	2	0	0	0	0
9	J	2	0	0	2	0
9	P	2	0	0	0	0
10	6	1	0	0	0	0
10	D	1	0	0	0	0
10	P	1	0	0	0	0
All	All	94608	0	92790	6951	3

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 37.

The worst 5 of 6951 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)
3:D:608:CYS:SG	3:D:617:THR:HG22	1.31	1.67
3:D:501:VAL:CG1	3:D:502:PRO:HD2	1.33	1.55
3:J:349:TYR:O	3:J:470:VAL:HG23	1.24	1.30
3:D:645:VAL:CG2	3:D:701:LEU:HD13	1.59	1.30
5:L:573:LEU:HB2	7:5:46:DG:OP2	1.15	1.28

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:1169:THR:OG1	6:1:16:DA:OP1[2_657]	1.85	0.35
3:D:710:ASP:OD2	3:J:1282:TYR:OH[2_547]	1.93	0.27
3:D:710:ASP:CA	3:J:1302:TYR:OH[2_547]	2.07	0.13

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	
1	A	228/242 (94%)	214 (94%)	11 (5%)	3 (1%)	9	41
1	B	226/242 (93%)	204 (90%)	17 (8%)	5 (2%)	5	28
1	G	228/242 (94%)	209 (92%)	16 (7%)	3 (1%)	9	41
1	H	226/242 (93%)	207 (92%)	13 (6%)	6 (3%)	4	25
1	M	228/242 (94%)	214 (94%)	14 (6%)	0	100	100
1	N	226/242 (93%)	209 (92%)	14 (6%)	3 (1%)	9	41
2	C	1339/1342 (100%)	1218 (91%)	98 (7%)	23 (2%)	7	35
2	I	1339/1342 (100%)	1214 (91%)	105 (8%)	20 (2%)	8	39
2	O	1339/1342 (100%)	1234 (92%)	90 (7%)	15 (1%)	11	45
3	D	1360/1407 (97%)	1220 (90%)	109 (8%)	31 (2%)	5	27
3	J	1360/1407 (97%)	1227 (90%)	99 (7%)	34 (2%)	4	25
3	P	1360/1407 (97%)	1226 (90%)	99 (7%)	35 (3%)	4	25
4	E	88/90 (98%)	83 (94%)	5 (6%)	0	100	100
4	K	88/90 (98%)	84 (96%)	3 (3%)	1 (1%)	11	45
4	Q	88/90 (98%)	84 (96%)	4 (4%)	0	100	100
5	F	493/628 (78%)	444 (90%)	27 (6%)	22 (4%)	2	16
5	L	493/628 (78%)	447 (91%)	28 (6%)	18 (4%)	2	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	R	493/628 (78%)	449 (91%)	30 (6%)	14 (3%)	4	24
All	All	11202/11853 (94%)	10187 (91%)	782 (7%)	233 (2%)	5	29

5 of 233 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	164	ASP
1	B	118	ASP
2	C	110	PRO
2	C	214	ASN
2	C	247	ARG

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	
1	A	198/208 (95%)	172 (87%)	26 (13%)	4	16
1	B	196/208 (94%)	166 (85%)	30 (15%)	3	12
1	G	198/208 (95%)	171 (86%)	27 (14%)	3	14
1	H	196/208 (94%)	168 (86%)	28 (14%)	3	14
1	M	198/208 (95%)	175 (88%)	23 (12%)	5	18
1	N	196/208 (94%)	167 (85%)	29 (15%)	3	13
2	C	1156/1157 (100%)	1014 (88%)	142 (12%)	4	17
2	I	1156/1157 (100%)	1030 (89%)	126 (11%)	6	21
2	O	1156/1157 (100%)	1012 (88%)	144 (12%)	4	17
3	D	1135/1168 (97%)	998 (88%)	137 (12%)	5	18
3	J	1135/1168 (97%)	980 (86%)	155 (14%)	3	14
3	P	1135/1168 (97%)	983 (87%)	152 (13%)	4	15
4	E	74/74 (100%)	65 (88%)	9 (12%)	5	17
4	K	74/74 (100%)	66 (89%)	8 (11%)	6	21
4	Q	74/74 (100%)	66 (89%)	8 (11%)	6	21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	F	439/554 (79%)	403 (92%)	36 (8%)	10	30
5	L	439/554 (79%)	386 (88%)	53 (12%)	5	18
5	R	439/554 (79%)	383 (87%)	56 (13%)	4	16
All	All	9594/10107 (95%)	8405 (88%)	1189 (12%)	4	17

5 of 1189 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	O	863	SER
5	R	309	ASN
2	O	1134	GLN
2	O	857	VAL
3	P	617	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 174 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	208	ASN
3	P	266	ASN
1	N	137	ASN
2	O	658	GLN
3	P	450	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	3	2/4 (50%)	1 (50%)	0
8	6	3/4 (75%)	1 (33%)	1 (33%)
8	9	3/4 (75%)	1 (33%)	1 (33%)
All	All	8/12 (66%)	3 (37%)	2 (25%)

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	3	15	G
8	6	15	G
8	9	15	G

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	6	13	GTP
8	9	13	GTP

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	4	3
7	2	3
7	5	3
6	7	1
7	8	1

The worst 5 of 11 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	45:DT	O3'	46:DG	P	5.04
1	7	50:DT	O3'	51:DC	P	4.24
1	8	22:DA	O3'	23:DT	P	3.80
1	2	22:DA	O3'	23:DT	P	3.79
1	5	22:DA	O3'	23:DT	P	3.79

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	A	230/242 (95%)	-0.75	0 100 100	153, 175, 210, 235	0
1	B	228/242 (94%)	-0.70	1 (0%) 88 78	162, 194, 217, 238	0
1	G	230/242 (95%)	-0.60	0 100 100	157, 185, 216, 248	0
1	H	228/242 (94%)	-0.65	1 (0%) 88 78	160, 191, 229, 261	0
1	M	230/242 (95%)	-0.70	0 100 100	166, 200, 233, 252	0
1	N	228/242 (94%)	-0.48	0 100 100	186, 233, 258, 273	0
2	C	1341/1342 (99%)	-0.66	4 (0%) 90 81	119, 186, 244, 277	0
2	I	1341/1342 (99%)	-0.68	2 (0%) 92 87	130, 195, 278, 377	0
2	O	1341/1342 (99%)	-0.66	1 (0%) 92 87	144, 183, 235, 270	0
3	D	1362/1407 (96%)	-0.59	6 (0%) 88 78	128, 214, 296, 349	0
3	J	1362/1407 (96%)	-0.59	2 (0%) 92 87	132, 194, 280, 314	0
3	P	1362/1407 (96%)	-0.58	5 (0%) 88 78	148, 208, 292, 330	0
4	E	90/90 (100%)	-0.45	0 100 100	169, 206, 407, 461	0
4	K	90/90 (100%)	-0.48	0 100 100	144, 199, 394, 442	0
4	Q	90/90 (100%)	-0.62	0 100 100	167, 222, 416, 460	0
5	F	497/628 (79%)	-0.53	0 100 100	182, 294, 404, 418	0
5	L	497/628 (79%)	-0.50	1 (0%) 91 84	168, 262, 400, 406	0
5	R	497/628 (79%)	-0.52	1 (0%) 91 84	172, 259, 413, 444	0
6	1	49/49 (100%)	-0.37	0 100 100	201, 272, 311, 317	0
6	4	49/49 (100%)	-0.32	0 100 100	209, 264, 308, 350	0
6	7	49/49 (100%)	-0.23	0 100 100	211, 255, 278, 300	0
7	2	49/49 (100%)	-0.36	0 100 100	210, 278, 312, 343	0
7	5	49/49 (100%)	-0.34	0 100 100	195, 270, 339, 341	0
7	8	49/49 (100%)	-0.32	0 100 100	194, 260, 296, 335	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
8	3	3/4 (75%)	-0.74	0 100 100	255, 255, 281, 321	0
8	6	3/4 (75%)	-0.62	0 100 100	263, 263, 272, 282	0
8	9	3/4 (75%)	-0.47	0 100 100	262, 262, 277, 295	0
All	All	11547/12159 (94%)	-0.60	24 (0%) 91 84	119, 203, 358, 461	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	282	LEU	5.9
3	D	772	TYR	5.3
5	L	304	THR	3.5
3	P	971	GLY	3.4
1	B	90	VAL	2.8

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

6.4 Ligands ⓘ

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(Å ²)	Q<0.9
10	MG	6	101	1/1	0.96	0.23	189,189,189,189	0
9	ZN	D	1502	1/1	0.99	0.02	212,212,212,212	0
9	ZN	J	1501	1/1	0.99	0.06	200,200,200,200	0
9	ZN	D	1501	1/1	0.99	0.04	228,228,228,228	0
10	MG	P	1503	1/1	0.99	0.03	194,194,194,194	0
9	ZN	P	1502	1/1	1.00	0.06	187,187,187,187	0
10	MG	D	1503	1/1	1.00	0.02	176,176,176,176	0
9	ZN	J	1502	1/1	1.00	0.06	174,174,174,174	0
9	ZN	P	1501	1/1	1.00	0.02	214,214,214,214	0

6.5 Other polymers [i](#)

There are no such residues in this entry.