



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 02:46 am BST

PDB ID : 3AOI
Title : RNA polymerase-Gfh1 complex (Crystal type 2)
Authors : Tagami, S.; Sekine, S.; Kumarevel, T.; Yamamoto, M.; Yokoyama, S.; RIKEN
Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2010-09-30
Resolution : 4.30 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

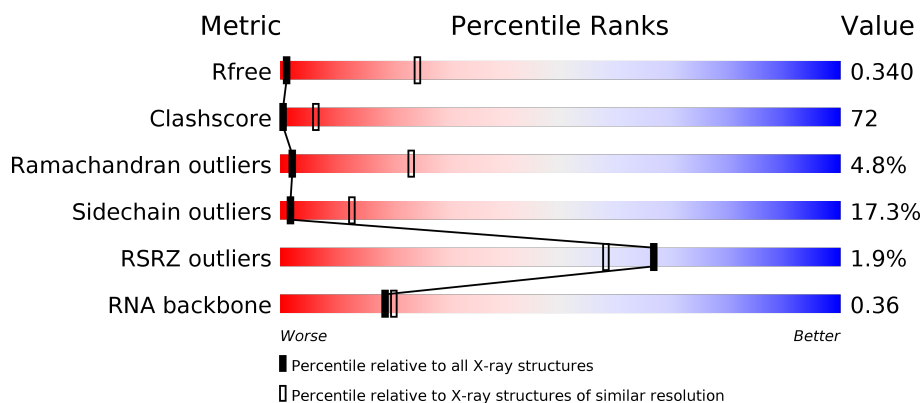
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1014 (4.80-3.80)
Clashscore	141614	1077 (4.80-3.80)
Ramachandran outliers	138981	1029 (4.80-3.80)
Sidechain outliers	138945	1012 (4.80-3.80)
RSRZ outliers	127900	1075 (4.90-3.70)
RNA backbone	3102	1058 (5.60-3.00)

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 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	315	<div> <div>19%</div> <div>44%</div> <div>8%</div> <div>29%</div> </div>
1	B	315	<div> <div>21%</div> <div>42%</div> <div>8%</div> <div>29%</div> </div>
1	F	315	<div> <div>%</div> <div>20%</div> <div>43%</div> <div>9%</div> <div>29%</div> </div>
1	G	315	<div> <div>22%</div> <div>39%</div> <div>10%</div> <div>29%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	315	
1	L	315	
2	C	1119	
2	H	1119	
2	M	1119	
3	D	1524	
3	I	1524	
3	N	1524	
4	E	99	
4	J	99	
4	O	99	
5	P	27	
5	R	27	
5	T	27	
6	Q	32	
6	S	32	
6	U	32	
7	X	156	
7	Y	156	
7	Z	156	

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 73646 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	223	Total	C	N	O	S	0	0	0
			1759	1123	306	328	2			
1	B	223	Total	C	N	O	S	0	0	0
			1759	1123	306	328	2			
1	F	223	Total	C	N	O	S	0	0	0
			1759	1123	306	328	2			
1	G	224	Total	C	N	O	S	0	0	0
			1764	1126	307	329	2			
1	K	225	Total	C	N	O	S	0	0	0
			1769	1129	308	330	2			
1	L	223	Total	C	N	O	S	0	0	0
			1759	1123	306	328	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1083	Total	C	N	O	S	0	0	0
			8550	5413	1525	1588	24			
2	H	1080	Total	C	N	O	S	0	0	0
			8524	5395	1521	1584	24			
2	M	1084	Total	C	N	O	S	0	0	0
			8555	5413	1528	1590	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1316	Total	C	N	O	S	0	0	0
			10384	6574	1840	1941	29			
3	I	1262	Total	C	N	O	S	0	0	0
			9965	6314	1765	1858	28			
3	N	1327	Total	C	N	O	S	0	0	0
			10475	6634	1852	1961	28			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	93	Total	C	N	O	S	0	0	0
			754	481	131	138	4			
4	J	93	Total	C	N	O	S	0	0	0
			754	481	131	138	4			
4	O	93	Total	C	N	O	S	0	0	0
			754	481	131	138	4			

- Molecule 5 is a DNA chain called DNA (5'-D(*GP*GP*TP*CP*TP*GP*TP*AP*TP*CP*AP*CP*GP*AP*GP*CP*CP*A*CP*CP*GP*CP*CP*GP*CP*AP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	P	7	Total	C	N	O	P	0	0	0
			136	65	25	40	6			
5	R	6	Total	C	N	O	P	0	0	0
			119	57	24	33	5			
5	T	5	Total	C	N	O	P	0	0	0
			98	47	19	28	4			

- Molecule 6 is a RNA chain called RNA (5'-R(*CP*CP*CP*GP*GP*AP*AP*GP*AP*UP*CP*AP*UP*CP*UP*UP*CP*CP*GP*GP*GP*GP*GP*AP*U*GP*CP*GP*GP*CP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	Q	7	Total	C	N	O	P	0	0	0
			152	68	31	47	6			
6	S	8	Total	C	N	O	P	0	0	0
			172	77	33	55	7			
6	U	7	Total	C	N	O	P	0	0	0
			152	68	31	47	6			

- Molecule 7 is a protein called Anti-cleavage anti-GreA transcription factor Gfh1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	X	154	Total	C	N	O	S	0	0	0
			1189	730	212	243	4			
7	Y	152	Total	C	N	O	S	0	0	0
			1169	719	207	239	4			
7	Z	152	Total	C	N	O	S	0	0	0
			1169	719	207	239	4			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	1	Total 1	Zn 1	0	0
8	D	1	Total 1	Zn 1	0	0
8	N	1	Total 1	Zn 1	0	0

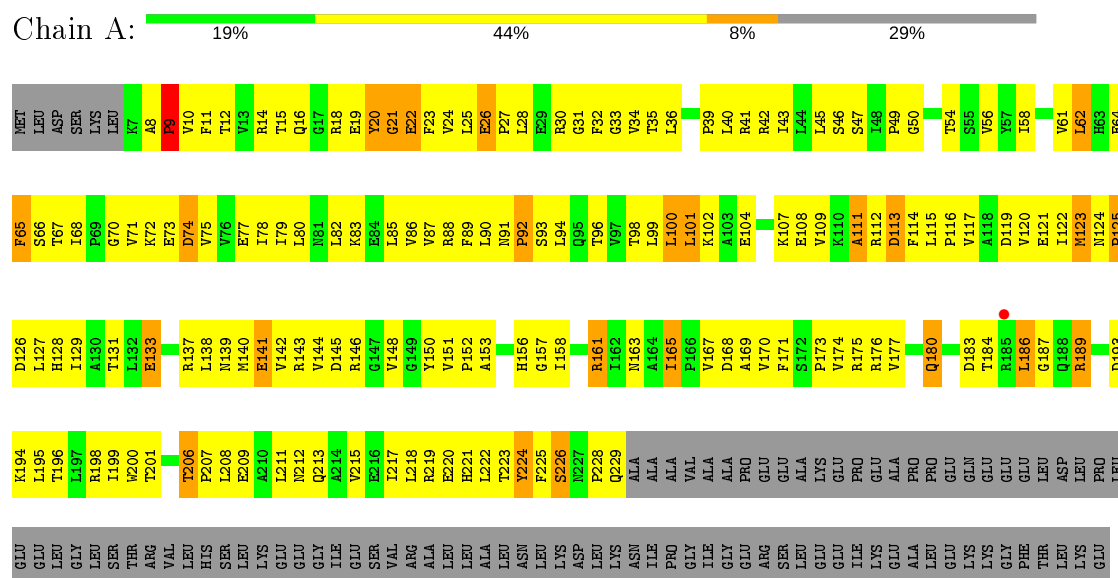
- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	1	Total 1	Mg 1	0	0
9	S	1	Total 1	Mg 1	0	0
9	N	1	Total 1	Mg 1	0	0

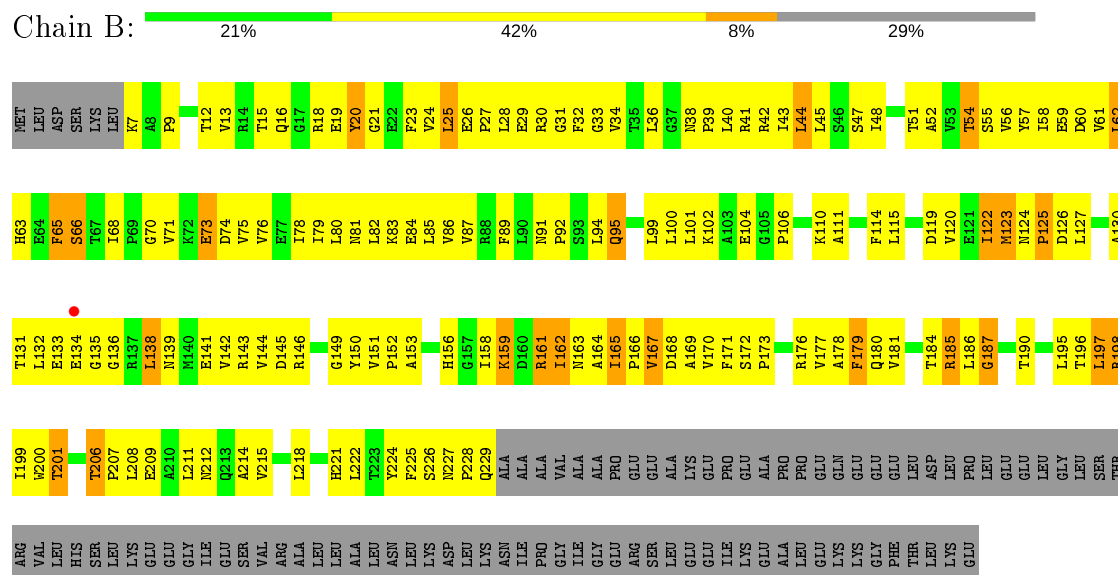
3 Residue-property plots

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.

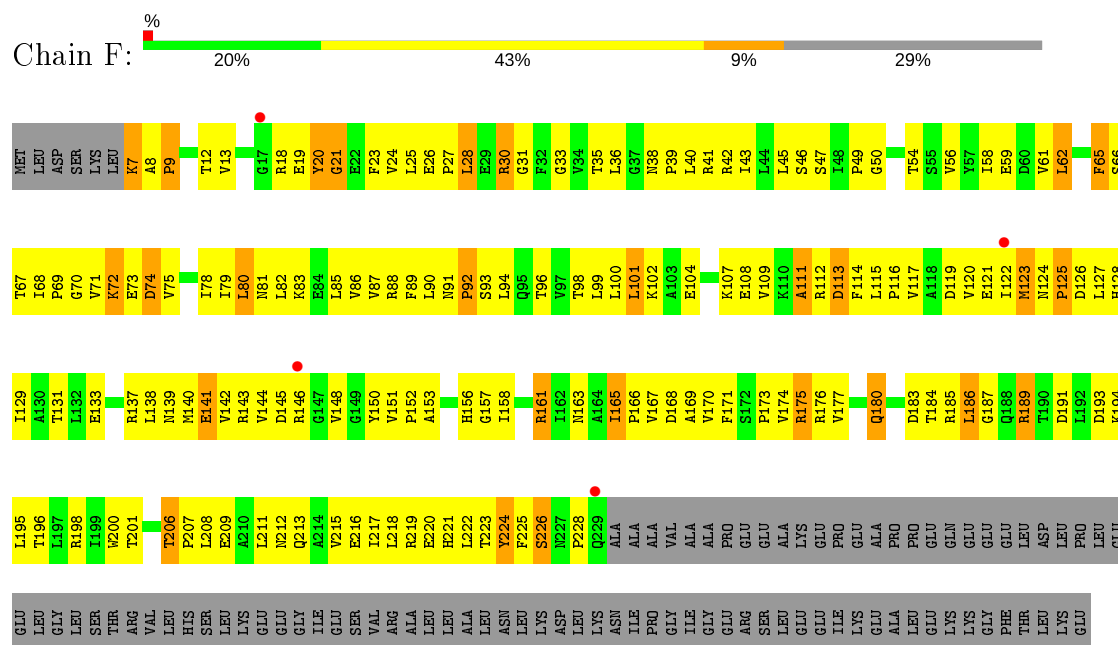
• Molecule 1: DNA-directed RNA polymerase subunit alpha

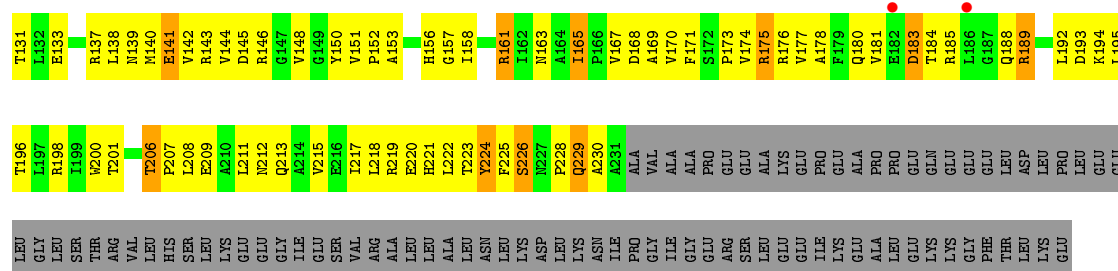


• Molecule 1: DNA-directed RNA polymerase subunit alpha



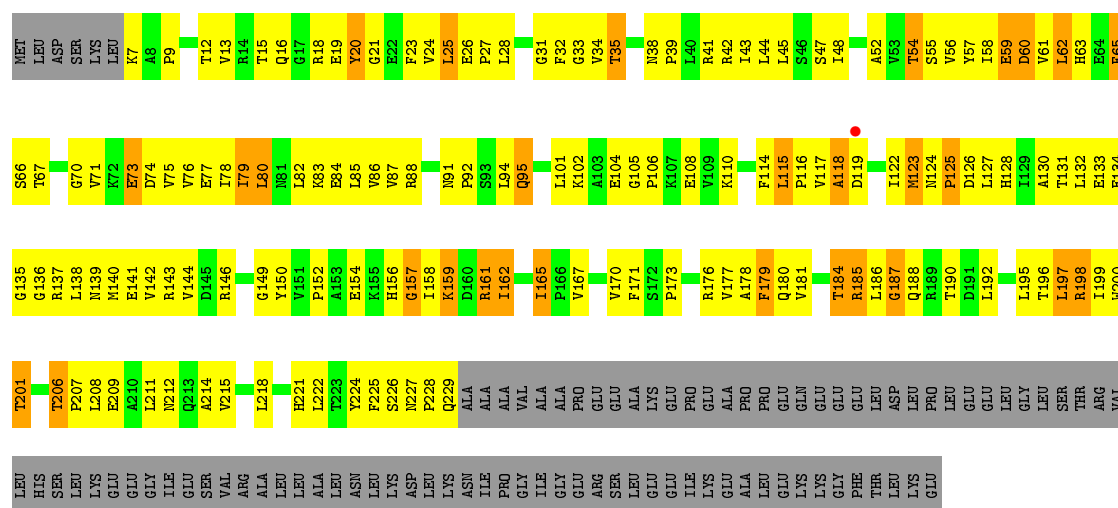
• Molecule 1: DNA-directed RNA polymerase subunit alpha





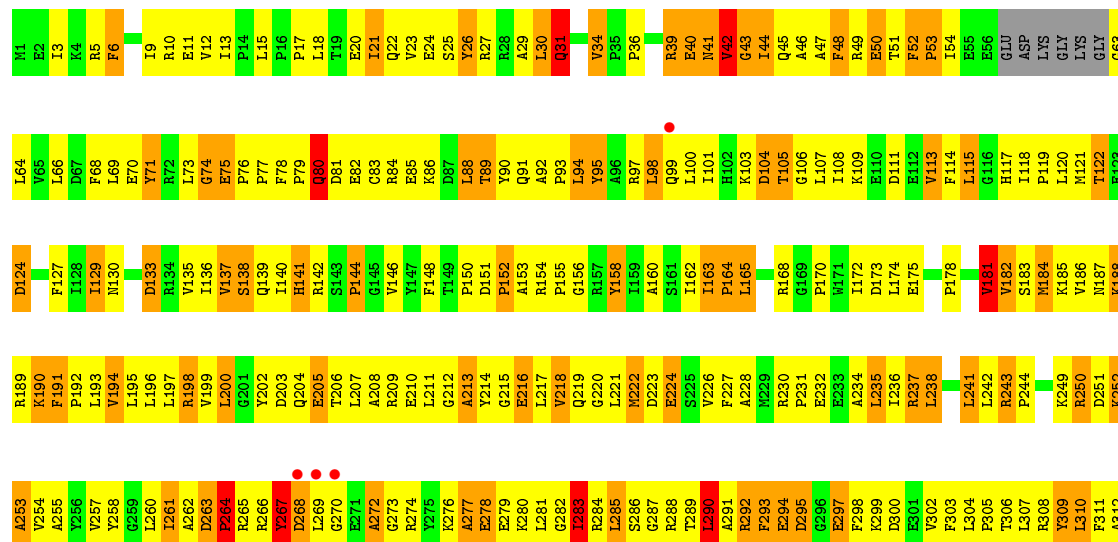
• Molecule 1: DNA-directed RNA polymerase subunit alpha

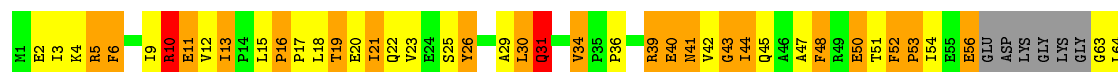
Chain L: 22% 39% 9% 29%



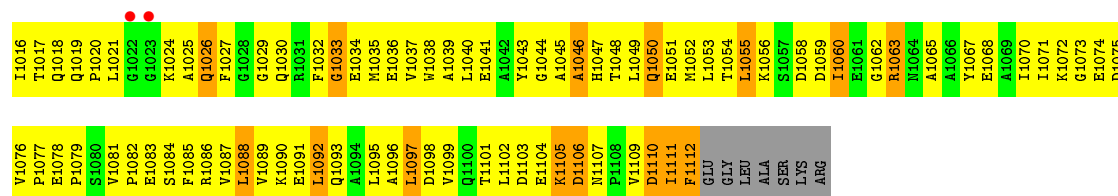
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C: 25% 55% 16%

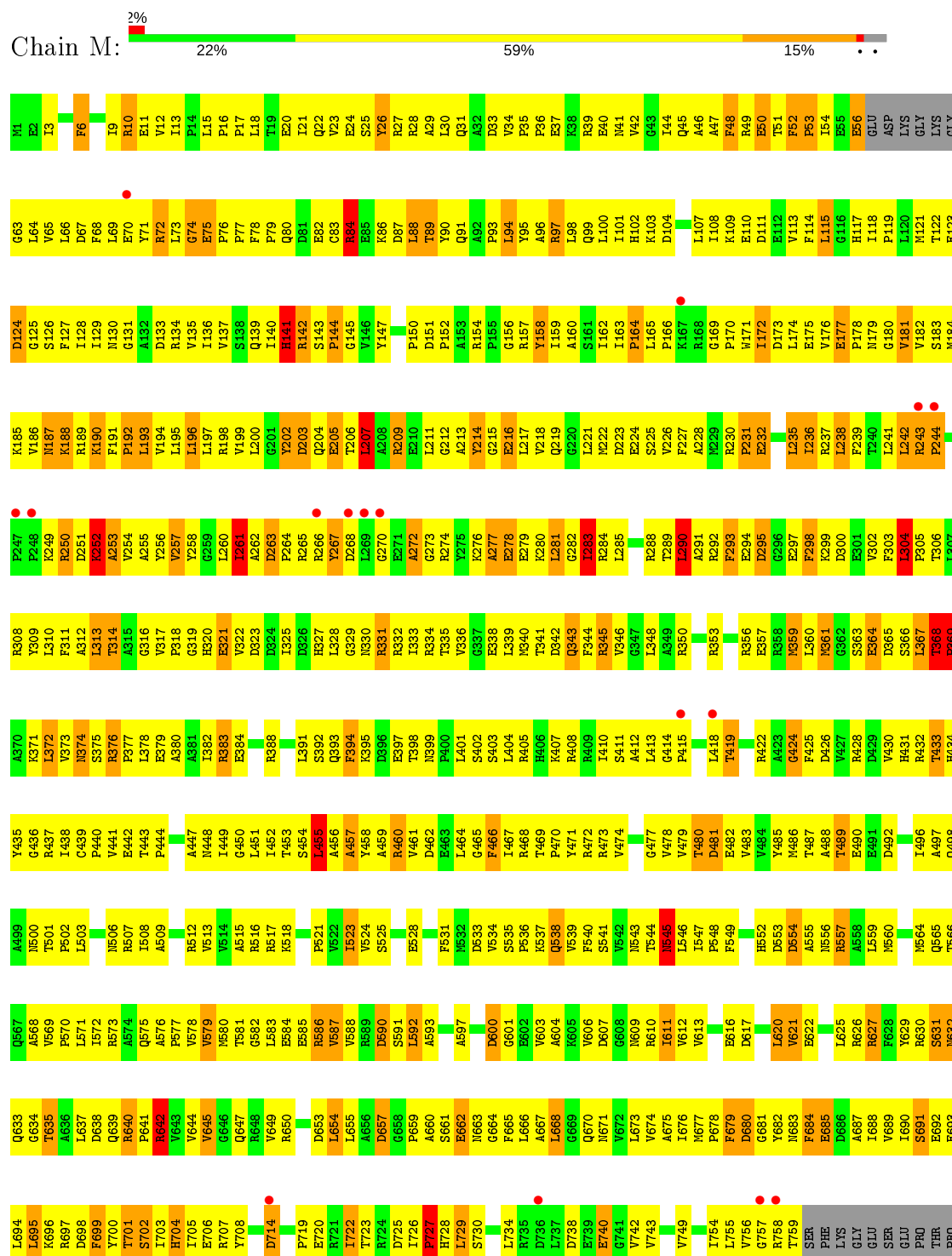


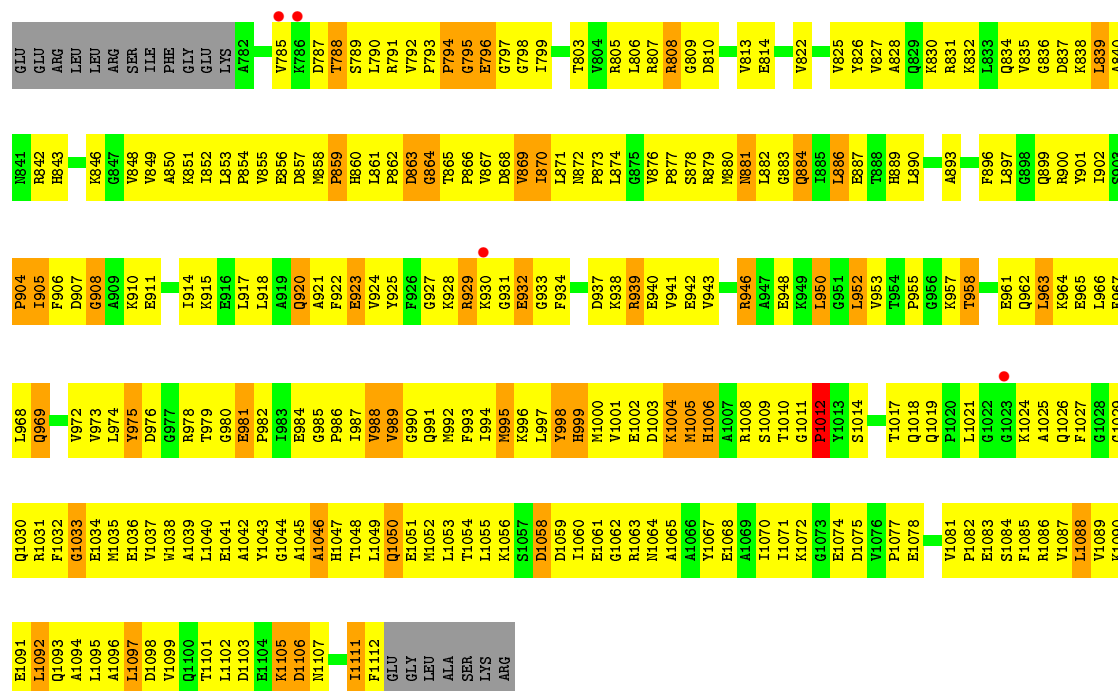




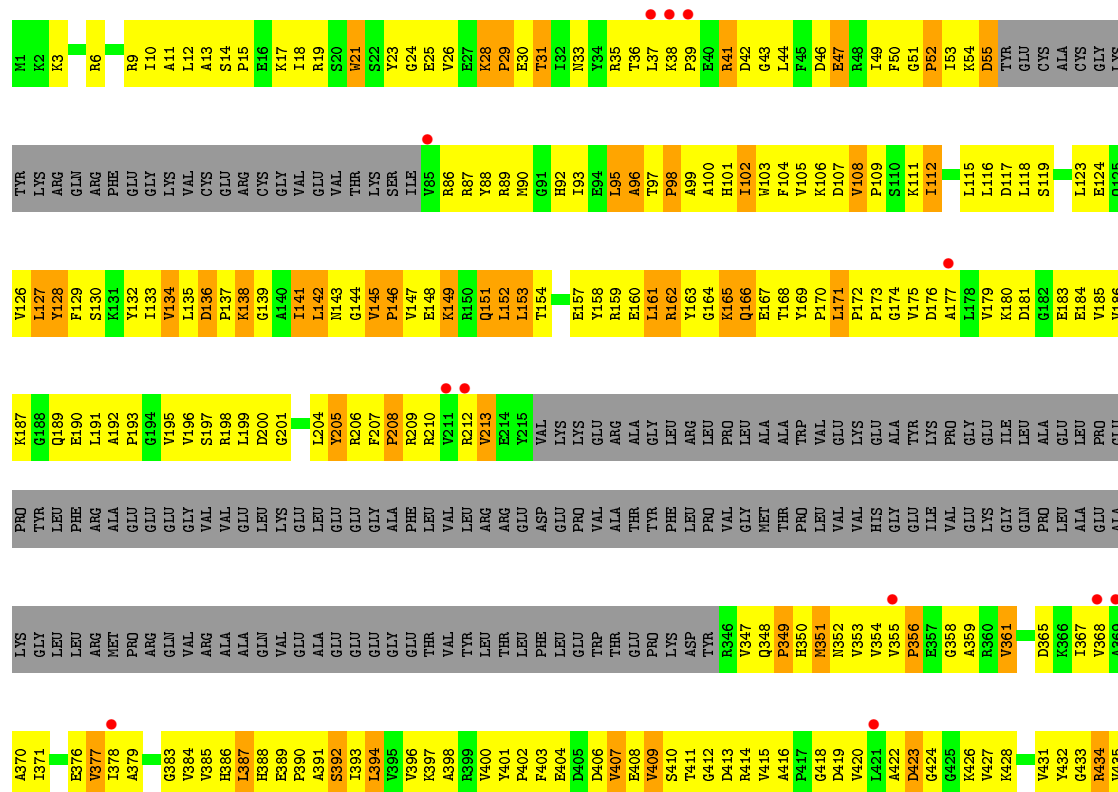


• Molecule 2: DNA-directed RNA polymerase subunit beta

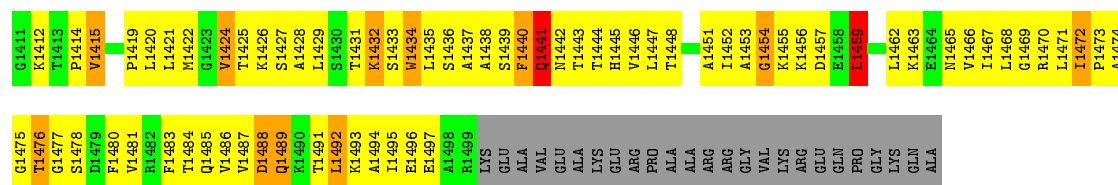




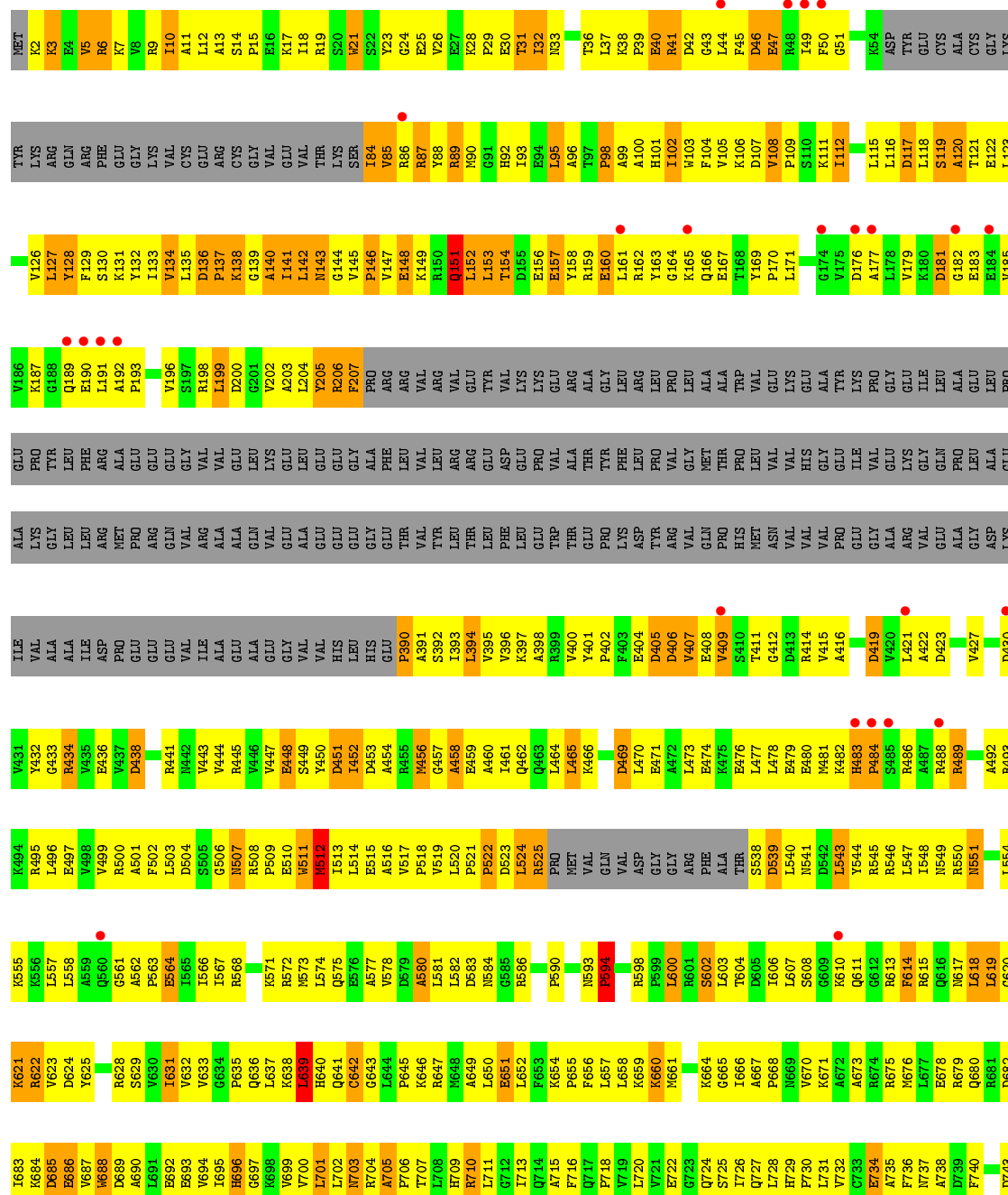
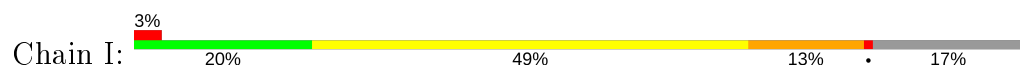
• Molecule 3: DNA-directed RNA polymerase subunit beta'



L1348	R1282	L1221	L1160	P1032	L964	V895	G831	H767	A705	L638	P563	V499	E436
L1283	L1283	V1221	L1161	Q1033	D968	A896	R832	L770	P706	H640	E564		V437
E1284	E1284	I1222	E1162	I1034	D969	W897	E833	L770	T707	Q641	E564	F502	D438
L1285	L1285	I1223	E1163	I1035	L1098	E898	T834	S771	L708	Q642	E570	L503	R441
K1286	T1286	V1224	R1164	Q1036	L099	L899	S835	P772	H709	G643	R571	D504	R442
L1287	L1287	A1225	R1165	L1037	L971	I900	S836	A773	R710	L644	R572	S505	K443
E1288	E1288	V1101	Y1165	L1038	L972	Q901	G837	S774	L711	P845	R573	G506	V443
L1289	L1289	T1102	L1166	C1039	Q973	L902	R838		G712	R646	L574	N507	V444
L1290	L1290	H1103	S1167	G1040	L974	D903	L839	L778	I713	R647	E575	R508	R445
A1291	A1291	M1168	E1168	L1041	E975	Q906	R840	A779	Q714	W648	E576	P509	V446
L1292	L1292	D1169	T1105	R1042	Q976	E907	K941	K780	A715	AG49	E577	E510	V447
L1293	L1293	V1106	D1170	G1043	A977	R908	V842	P781	F716	L650	Y578	N511	E448
E1295	E1295	H1171	H1171	L1044	Y978	K908	E843	S782	Q717	R651	D579	R512	S449
L1296	L1296	E1109	L1173	M1045	E979	N909	A844	R783	P718	L652	A580	I513	Y450
S1297	S1297	E1110	L1174	Q1046	N980	S910	P845	D784	V719	F653	L581	L514	D451
L1298	L1298	D1111	F982	K1047	G981	I911	R847	I786	L726	R655	D583	E516	D453
F1299	F1299	C1112	L983	P1048	L983	L914	D847	L787	E722	P656	N584	A517	A454
S1300	S1300	G1113	T984	G1050	T984	V915	A849	G788	Q723	L657	G585	P518	R455
K1301	K1301	T1114	E987	E1051	E987	A918	L850	L789	Q724	L659	R586	V519	R456
E1302	E1302	E1177	T1052	T1052	R988	F919	R872	Y790	S725	L659	R587	L520	G457
L1303	L1303	G1180	F1083	F1083	R989	F920	A852	I791	L726	R660	G588	P521	A458
K1304	K1304	E1182	E1054	E1054	L998	L921	V853	I792	Q727	R661	A589	D523	A460
L1305	L1305	G1183	V1055	V1055	D990	R921	R857	T793	L728	K664	P590	L524	I461
V1306	V1306	P1120	P1056	V1057	Q981	L922	I857	Q794	H729	G665	R593	R525	Q462
Q1374	Q1374	L1122	L1122	V1057	I982	G923	V858	R796	P730	L666	P594	P526	Q463
M1375	M1375	F1123	Q1124	R1058	Q994	R924	L860	L731	V732		VAL	GLN	L464
R1376	R1376	E1186	Q1124	S1059	L995	R925	R861	K799	E734	G671	L600	VAL	L465
K1377	K1377	P1187	P1124	S1060	L995	K926	Q861	R800	A735	A672	L603	ASP	D469
L1378	L1378	V1188	D1126	F1061	W996	T927	V863	G801	F736	A673	T604	GLY	L470
V1379	V1379	R1189	E1127	R1062	T999	A928	R864	A802	A737	R674	P605	GLY	E471
E1380	E1380	S1190	V1128	E1063	L1000	R929	T865	G803	N737	R675	L606	ARG	A472
D1315	D1315	P1191	T1129	G1064	E1001	L930	R866	E805	D738	W676	S608	PHE	L473
L1316	L1316	L1192	Q1254	L1065	K1002	D932	R867	E805	F740	L677		ALA	E474
D1382	D1382	T1193	S1131	T1066	K1002	D932	R867	F806			R613	THR	K475
P1384	P1384	G1194	L1132	V1067	V1003	A933	Y868	A807	D743		F614	S538	E476
G1385	G1385	Q1195	R1133	L1068	V1007	L934	M869	R808	Q744	D681	D539	D539	L477
E1320	E1320	T1196	L1134	E1069	F1008	Y937	R872	P809	M745	L683	N541	N541	E478
A1321	A1321	R1197	K1136	Y1070	F1008	G938		E810	A746	K684			E479
L1325	L1325	Y1198	R1136	F1071	F1011	F939	T875	E811	V747	D685	L618		E480
L1326	L1326	G1199	A1138	I1072		T940	S876	A812	H748	D686	L619		E481
F1263	F1263	C1201	A1139	S1074	H1014	S942	G878	A814	W749	R687	G620		K481
E1264	E1264	Q1202	E1141	R1078	Y1015	T943	G879	A815	P750	R688	K621		K482
A1265	A1265	K1203	E1142		P1016	T943	L880	H816		D689	R622		P483
L1266	L1266	G1204	G1143	G1081	F1017	T943	L881	E817		A690	V623		P484
R1267	R1267	Y1205	G1143	A1082	M1018	S945	F882	R818	F754	L691	D624		S485
P1268	P1268	L1144	L1144	D1083	P1019	G946	F883	R819	A755	E692	R624		R486
K1269	K1269	Y1145	Y1145	T1084	L1020	I947	A883	G819	A756	E693	V625		A487
		G1146	G1146	T1084	V1021	T948	R884	E820	Q756	V694	R628		R488
		R1147	R1147	A1085	V1022	I949	L885	V821	A757	L695	S629		R489
		V1148	V1148	L1086	M1023	G950	V886	A822	E758	R696	R552		P489
		M1211	L1149	T1087	A1024	I951	A887	L823	A759		R551		A490
P1341	P1341	A1150	Q1025	T1088	Q1025	D952	E888	N824	A760		L631		K491
E1342	E1342	R1151	L1026	A1089	S1026	D953	A889	A825	I761		V632		A492
L1277	L1277	E1152	G1027	A1090	G1027	I956	R890	P826	Q762				R493
D1278	D1278	V1153	A1028	S1091	A1028	P957	E891	I827	M763		P635		R494
G1279	G1279	E1154	R1029	G1092	R1029	R957	D892	R828	L764		Q636		R495
L1281	L1281	V1155	G1030	Y1093	G1030	E993	E893	V829	S765		L637		R496
		E1219	L1156	L1094	N1031	Y963	K894	A830	A766		R638		



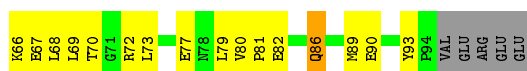
• Molecule 3: DNA-directed RNA polymerase subunit beta'



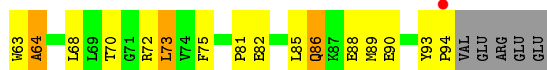
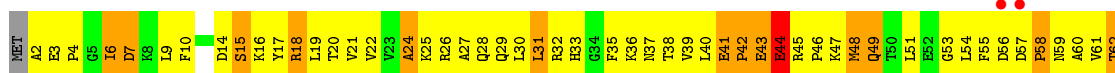




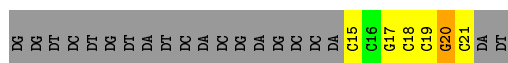




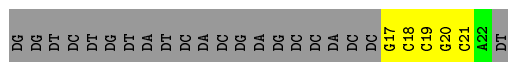
- Molecule 4: DNA-directed RNA polymerase subunit omega



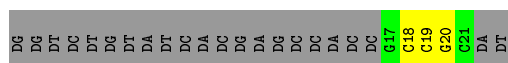
- Molecule 5: DNA (5'-D(*GP*GP*TP*CP*TP*GP*TP*AP*TP*CP*AP*CP*GP*AP*GP*CP*CP*A*CP*CP*GP*CP*CP*GP*CP*AP*T)-3')



- Molecule 5: DNA (5'-D(*GP*GP*TP*CP*TP*GP*TP*AP*TP*CP*AP*CP*GP*AP*GP*CP*CP*A*CP*CP*GP*CP*CP*GP*CP*AP*T)-3')



- Molecule 5: DNA (5'-D(*GP*GP*TP*CP*TP*GP*TP*AP*TP*CP*AP*CP*GP*AP*GP*CP*CP*A*CP*CP*GP*CP*CP*GP*CP*AP*T)-3')



- Molecule 6: RNA (5'-R(*CP*CP*CP*GP*GP*AP*AP*GP*AP*UP*CP*AP*UP*CP*UP*UP*CP*CP*GP*GP*GP*GP*AP*U*GP*CP*GP*GP*CP*GP*G)-3')



- Molecule 6: RNA (5'-R(*CP*CP*CP*GP*GP*AP*AP*GP*AP*UP*CP*AP*UP*CP*UP*UP*CP*CP*GP*GP*GP*GP*AP*U*GP*CP*GP*GP*CP*GP*G)-3')



F149
R150
V151
V152
A153
I154
H155
G156

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	189.23Å 264.51Å 193.93Å 90.00° 116.68° 90.00°	Depositor
Resolution (Å)	47.57 – 4.30 47.57 – 4.30	Depositor EDS
% Data completeness (in resolution range)	97.6 (47.57-4.30) 97.7 (47.57-4.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.68 (at 4.29Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.317 , 0.338 0.320 , 0.340	Depositor DCC
R_{free} test set	3392 reflections (3.01%)	wwPDB-VP
Wilson B-factor (Å ²)	128.7	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 107.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.054 for l,-k,h	Xtriage
F_o, F_c correlation	0.81	EDS
Total number of atoms	73646	wwPDB-VP
Average B, all atoms (Å ²)	160.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.76% 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: 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	A	0.48	0/1791	0.76	0/2436
1	B	0.47	0/1791	0.71	1/2436 (0.0%)
1	F	0.48	0/1791	0.75	0/2436
1	G	0.47	0/1796	0.70	1/2443 (0.0%)
1	K	0.50	0/1801	0.76	0/2450
1	L	0.48	0/1791	0.70	1/2436 (0.0%)
2	C	0.49	0/8713	0.78	5/11785 (0.0%)
2	H	0.51	0/8686	0.78	3/11750 (0.0%)
2	M	0.51	0/8717	0.81	6/11792 (0.1%)
3	D	0.50	0/10559	0.77	5/14272 (0.0%)
3	I	0.50	0/10131	0.79	6/13685 (0.0%)
3	N	0.49	0/10653	0.79	8/14403 (0.1%)
4	E	0.47	0/768	0.72	1/1035 (0.1%)
4	J	0.47	0/768	0.73	1/1035 (0.1%)
4	O	0.49	0/768	0.77	1/1035 (0.1%)
5	P	0.89	0/151	1.70	3/230 (1.3%)
5	R	0.81	0/133	1.04	0/203
5	T	0.84	0/109	0.89	0/166
6	Q	1.05	0/170	1.04	0/265
6	S	1.02	0/192	0.92	0/299
6	U	0.94	0/170	0.97	2/265 (0.8%)
7	X	0.56	1/1198 (0.1%)	0.70	1/1608 (0.1%)
7	Y	0.64	1/1178 (0.1%)	0.66	0/1582
7	Z	0.65	1/1178 (0.1%)	0.68	1/1582 (0.1%)
All	All	0.51	3/75003 (0.0%)	0.78	46/101629 (0.0%)

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
2	C	0	5
2	H	0	1
2	M	0	3
3	D	0	2
3	I	0	4
3	N	0	5
6	Q	0	1
6	S	0	1
All	All	0	22

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	Z	156	GLY	C-OXT	16.27	1.54	1.23
7	Y	156	GLY	C-OXT	16.19	1.54	1.23
7	X	156	GLY	C-OXT	11.06	1.44	1.23

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	P	20	DG	N9-C1'-C2'	9.98	131.56	112.60
5	P	20	DG	O4'-C1'-C2'	8.43	112.64	105.90
3	I	1209	LEU	N-CA-C	-8.04	89.28	111.00
3	N	1209	LEU	N-CA-C	-7.97	89.49	111.00
3	N	142	LEU	CA-CB-CG	7.29	132.07	115.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	267	TYR	Sidechain
2	C	589	ARG	Sidechain
2	C	642	ARG	Sidechain
2	C	71	TYR	Sidechain
2	C	735	ARG	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	A	1759	0	1805	209	3
1	B	1759	0	1805	263	2
1	F	1759	0	1805	195	3
1	G	1764	0	1810	253	3
1	K	1769	0	1815	202	0
1	L	1759	0	1805	220	0
2	C	8550	0	8654	1412	1
2	H	8524	0	8626	1521	0
2	M	8555	0	8658	1519	1
3	D	10384	0	10615	1752	3
3	I	9965	0	10206	1707	1
3	N	10475	0	10699	1791	3
4	E	754	0	769	94	0
4	J	754	0	769	116	0
4	O	754	0	769	111	0
5	P	136	0	79	3	0
5	R	119	0	68	12	0
5	T	98	0	57	8	0
6	Q	152	0	78	10	0
6	S	172	0	88	12	0
6	U	152	0	79	12	0
7	X	1189	0	1205	141	0
7	Y	1169	0	1186	151	0
7	Z	1169	0	1186	146	0
8	D	1	0	0	0	0
8	I	1	0	0	0	0
8	N	1	0	0	0	0
9	D	1	0	0	0	0
9	N	1	0	0	0	0
9	S	1	0	0	0	0
All	All	73646	0	74636	10729	10

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 72.

The worst 5 of 10729 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:N:50:PHE:CD2	3:N:522:PRO:HD3	1.20	1.62
2:H:182:VAL:HG11	2:H:193:LEU:CD2	1.09	1.56
2:H:1090:LYS:HE3	3:I:90:MET:SD	1.45	1.55
2:H:182:VAL:CG1	2:H:193:LEU:HD21	1.13	1.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:I:783:ARG:NH1	7:Y:41:ASP:CB	1.67	1.53

The worst 5 of 10 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)
1:A:100:LEU:CD1	1:F:59:GLU:OE1[2_455]	1.51	0.69
3:D:1182:GLU:OE2	1:G:112:ARG:NE[1_655]	1.61	0.59
1:B:162:ILE:CG1	3:N:976:GLN:NE2[1_655]	1.90	0.30
1:A:100:LEU:CD1	1:F:59:GLU:CD[2_455]	1.97	0.23
2:C:223:ASP:O	3:N:562:ALA:O[2_444]	2.02	0.18

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	221/315 (70%)	191 (86%)	22 (10%)	8 (4%)	3	28
1	B	221/315 (70%)	191 (86%)	27 (12%)	3 (1%)	11	47
1	F	221/315 (70%)	191 (86%)	20 (9%)	10 (4%)	2	24
1	G	222/315 (70%)	193 (87%)	26 (12%)	3 (1%)	11	47
1	K	223/315 (71%)	193 (86%)	21 (9%)	9 (4%)	3	26
1	L	221/315 (70%)	188 (85%)	28 (13%)	5 (2%)	6	37
2	C	1077/1119 (96%)	864 (80%)	150 (14%)	63 (6%)	1	20
2	H	1074/1119 (96%)	867 (81%)	145 (14%)	62 (6%)	1	20
2	M	1078/1119 (96%)	871 (81%)	149 (14%)	58 (5%)	2	22
3	D	1306/1524 (86%)	1062 (81%)	186 (14%)	58 (4%)	2	25
3	I	1252/1524 (82%)	1012 (81%)	177 (14%)	63 (5%)	2	23
3	N	1317/1524 (86%)	1052 (80%)	196 (15%)	69 (5%)	2	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	E	91/99 (92%)	67 (74%)	17 (19%)	7 (8%)	1	15
4	J	91/99 (92%)	70 (77%)	14 (15%)	7 (8%)	1	15
4	O	91/99 (92%)	68 (75%)	16 (18%)	7 (8%)	1	15
7	X	152/156 (97%)	132 (87%)	16 (10%)	4 (3%)	5	35
7	Y	150/156 (96%)	135 (90%)	12 (8%)	3 (2%)	7	40
7	Z	150/156 (96%)	131 (87%)	16 (11%)	3 (2%)	7	40
All	All	9158/10584 (86%)	7478 (82%)	1238 (14%)	442 (5%)	2	23

5 of 442 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	187	GLY
2	C	23	VAL
2	C	40	GLU
2	C	44	ILE
2	C	152	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	
1	A	196/273 (72%)	164 (84%)	32 (16%)	2	14
1	B	196/273 (72%)	166 (85%)	30 (15%)	2	16
1	F	196/273 (72%)	167 (85%)	29 (15%)	3	17
1	G	196/273 (72%)	164 (84%)	32 (16%)	2	14
1	K	196/273 (72%)	164 (84%)	32 (16%)	2	14
1	L	196/273 (72%)	164 (84%)	32 (16%)	2	14
2	C	912/941 (97%)	737 (81%)	175 (19%)	1	9
2	H	909/941 (97%)	736 (81%)	173 (19%)	1	9
2	M	912/941 (97%)	746 (82%)	166 (18%)	1	11
3	D	1113/1279 (87%)	928 (83%)	185 (17%)	2	14

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	I	1068/1279 (84%)	878 (82%)	190 (18%)	2	12
3	N	1124/1279 (88%)	937 (83%)	187 (17%)	2	14
4	E	82/88 (93%)	72 (88%)	10 (12%)	5	23
4	J	82/88 (93%)	72 (88%)	10 (12%)	5	23
4	O	82/88 (93%)	67 (82%)	15 (18%)	1	11
7	X	130/131 (99%)	115 (88%)	15 (12%)	5	24
7	Y	128/131 (98%)	104 (81%)	24 (19%)	1	10
7	Z	128/131 (98%)	110 (86%)	18 (14%)	3	19
All	All	7846/8955 (88%)	6491 (83%)	1355 (17%)	2	13

5 of 1355 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	691	SER
3	I	778	LEU
3	N	1297	GLU
2	H	870	ILE
3	I	138	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 252 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	834	GLN
3	I	973	GLN
3	N	1465	ASN
2	H	969	GLN
3	I	463	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	Q	6/32 (18%)	1 (16%)	0
6	S	7/32 (21%)	2 (28%)	0
6	U	6/32 (18%)	2 (33%)	0
All	All	19/96 (19%)	5 (26%)	0

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	Q	11	C
6	S	13	G
6	S	16	G
6	U	11	C
6	U	12	G

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 6 ligands modelled in this entry, 6 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](#)

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	A	223/315 (70%)	-0.18	1 (0%) 92 87	82, 143, 203, 231	0
1	B	223/315 (70%)	-0.25	1 (0%) 92 87	86, 150, 211, 231	0
1	F	223/315 (70%)	-0.04	4 (1%) 68 60	88, 145, 202, 231	0
1	G	224/315 (71%)	-0.22	0 100 100	74, 150, 204, 231	0
1	K	225/315 (71%)	0.03	4 (1%) 68 60	77, 152, 205, 231	0
1	L	223/315 (70%)	-0.24	1 (0%) 92 87	71, 146, 208, 231	0
2	C	1083/1119 (96%)	-0.17	11 (1%) 82 74	46, 150, 225, 231	0
2	H	1080/1119 (96%)	-0.18	15 (1%) 75 66	32, 148, 226, 231	0
2	M	1084/1119 (96%)	-0.07	20 (1%) 68 60	51, 150, 227, 231	0
3	D	1316/1524 (86%)	-0.08	25 (1%) 66 58	91, 161, 249, 285	0
3	I	1262/1524 (82%)	0.04	45 (3%) 42 34	91, 160, 246, 285	0
3	N	1327/1524 (87%)	-0.01	24 (1%) 68 60	91, 162, 250, 284	0
4	E	93/99 (93%)	0.06	2 (2%) 62 52	91, 182, 231, 231	0
4	J	93/99 (93%)	-0.01	3 (3%) 47 37	98, 176, 231, 231	0
4	O	93/99 (93%)	-0.06	3 (3%) 47 37	90, 172, 231, 231	0
5	P	7/27 (25%)	0.31	0 100 100	199, 200, 200, 200	0
5	R	6/27 (22%)	0.80	0 100 100	200, 200, 200, 200	0
5	T	5/27 (18%)	0.57	0 100 100	199, 200, 200, 200	0
6	Q	7/32 (21%)	0.94	0 100 100	195, 200, 200, 200	0
6	S	8/32 (25%)	2.18	4 (50%) 0 0	200, 200, 200, 200	0
6	U	7/32 (21%)	1.24	1 (14%) 2 3	194, 200, 200, 200	0
7	X	154/156 (98%)	0.11	4 (2%) 56 46	97, 175, 229, 231	0
7	Y	152/156 (97%)	0.11	6 (3%) 39 31	88, 174, 228, 231	0
7	Z	152/156 (97%)	0.18	6 (3%) 39 31	95, 172, 228, 231	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
All	All	9270/10761 (86%)	-0.07	180 (1%) 66 58	32, 156, 231, 285	0

The worst 5 of 180 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	E	56	ASP	6.1
3	I	1301	LYS	5.9
3	I	191	LEU	5.5
3	I	1292	VAL	5.3
2	C	270	GLY	4.9

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 carbohydrates 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
8	ZN	N	2001	1/1	0.91	0.07	115,115,115,115	0
8	ZN	D	2002	1/1	0.96	0.03	115,115,115,115	0
9	MG	N	2006	1/1	0.96	0.45	115,115,115,115	0
9	MG	S	2005	1/1	0.97	0.86	115,115,115,115	0
8	ZN	I	2003	1/1	0.99	0.09	115,115,115,115	0
9	MG	D	2004	1/1	0.99	0.11	115,115,115,115	0

6.5 Other polymers ⓘ

There are no such residues in this entry.