



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:24 AM GMT

PDB ID : 3HOU
Title : Complete RNA polymerase II elongation complex I with a T-U mismatch
Authors : Sydow, J.F.; Brueckner, F.; Cheung, A.C.M.; Damsma, G.E.; Dengl, S.;
Lehmann, E.; Vassilyev, D.; Cramer, P.
Deposited on : 2009-06-03
Resolution : 3.20 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

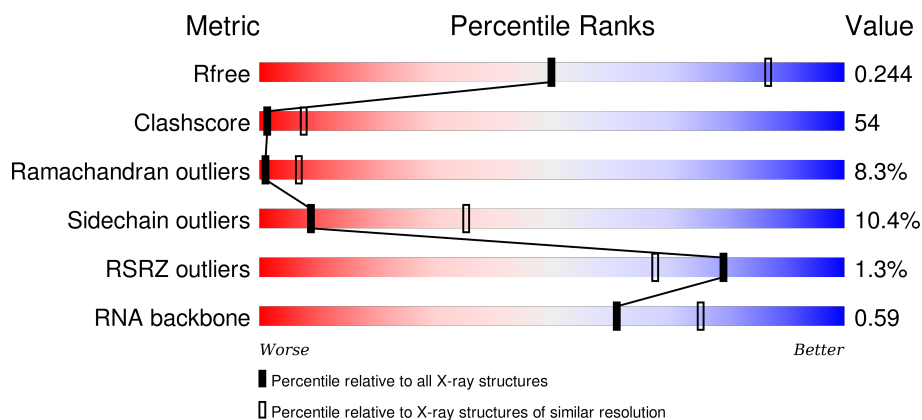
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)
RNA backbone	2183	1079 (3.70-2.70)

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	A	1733	
1	M	1733	
2	B	1224	
2	N	1224	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	318	
3	O	318	
4	D	221	
4	P	221	
5	E	215	
5	Q	215	
6	F	155	
6	R	155	
7	G	171	
7	S	171	
8	H	146	
8	T	146	
9	I	122	
9	U	122	
10	J	70	
10	V	70	
11	K	120	
11	W	120	
12	L	70	
12	X	70	
13	1	26	
13	4	26	
14	2	13	
14	5	13	
15	3	17	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
15	6	17	<div><div></div><div>6%</div><div>29%</div><div>35%</div><div>35%</div></div>

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 63664 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 II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1416	Total	C	N	O	S	0	0	0
			11143	7021	1949	2111	62			
1	M	1416	Total	C	N	O	S	0	0	0
			11143	7021	1949	2111	62			

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1104	Total	C	N	O	S	0	0	0
			8779	5560	1537	1627	55			
2	N	1104	Total	C	N	O	S	0	0	0
			8779	5560	1537	1627	55			

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			
3	O	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	178	Total	C	N	O	S	0	0	0
			1434	887	257	288	2			
4	P	178	Total	C	N	O	S	0	0	0
			1434	887	257	288	2			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			
5	Q	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			
6	R	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			
7	S	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	134	Total	C	N	O	S	0	0	0
			1076	677	182	213	4			
8	T	134	Total	C	N	O	S	0	0	0
			1076	677	182	213	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			
9	U	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	V	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			
11	W	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			
12	X	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

- Molecule 13 is a DNA chain called 5'-D(*AP*GP*CP*TP*CP*A*AP*GP*TP*AP*GP*TP*TP*AP*TP*GP*CP*CP*(BRU)P*GP*GP*TP*CP*AP*TP*T)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
13	1	18	Total	Br	C	N	O	P	0	0	0
			368	1	176	66	108	17			
13	4	18	Total	Br	C	N	O	P	0	0	0
			368	1	176	66	108	17			

- Molecule 14 is a DNA chain called 5'-D(*A*AP*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	2	6	Total	C	N	O	P	0	0	0
			117	58	20	34	5			
14	5	6	Total	C	N	O	P	0	0	0
			117	58	20	34	5			

- Molecule 15 is a RNA chain called 5'-R(*UP*GP*CP*AP*UP*U*UP*CP*GP*AP*CP*CP*AP*GP*GP*CP*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	3	11	Total	C	N	O	P	0	0	0
			230	104	41	75	10			
15	6	11	Total	C	N	O	P	0	0	0
			230	104	41	75	10			

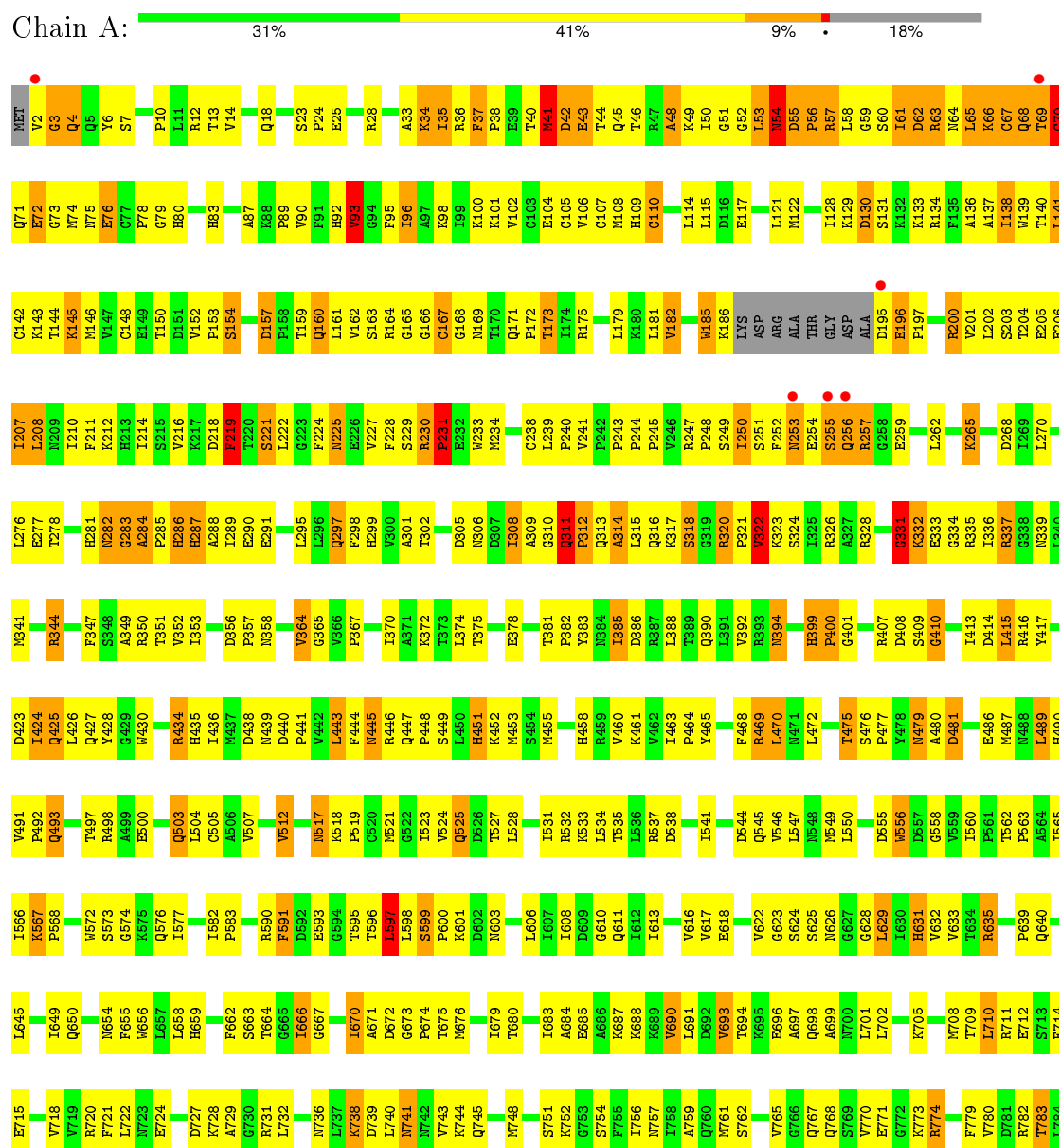
- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	J	1	Total	Zn	0	0
			1	1		
16	B	1	Total	Zn	0	0
			1	1		
16	I	2	Total	Zn	0	0
			2	2		
16	C	1	Total	Zn	0	0
			1	1		
16	V	1	Total	Zn	0	0
			1	1		
16	A	2	Total	Zn	0	0
			2	2		
16	N	1	Total	Zn	0	0
			1	1		
16	U	2	Total	Zn	0	0
			2	2		
16	X	1	Total	Zn	0	0
			1	1		
16	O	1	Total	Zn	0	0
			1	1		
16	L	1	Total	Zn	0	0
			1	1		
16	M	2	Total	Zn	0	0
			2	2		

3 Residue-property plots

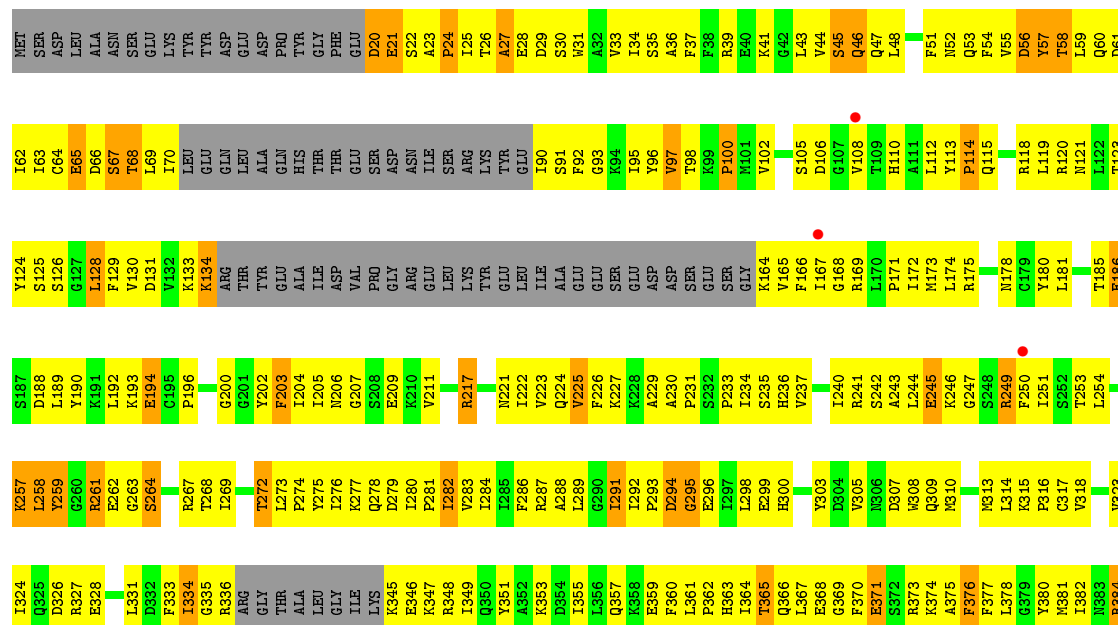
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

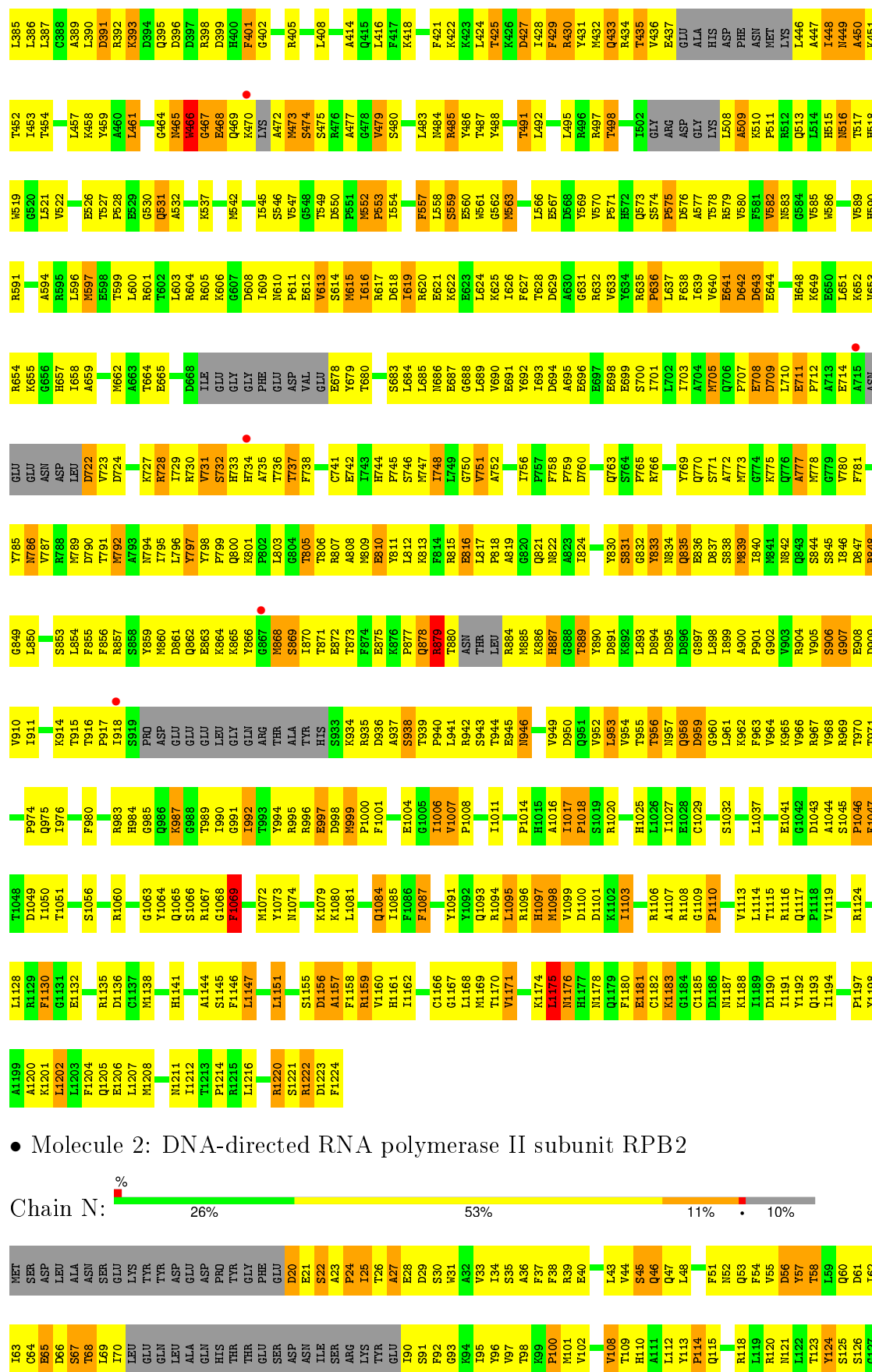
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1





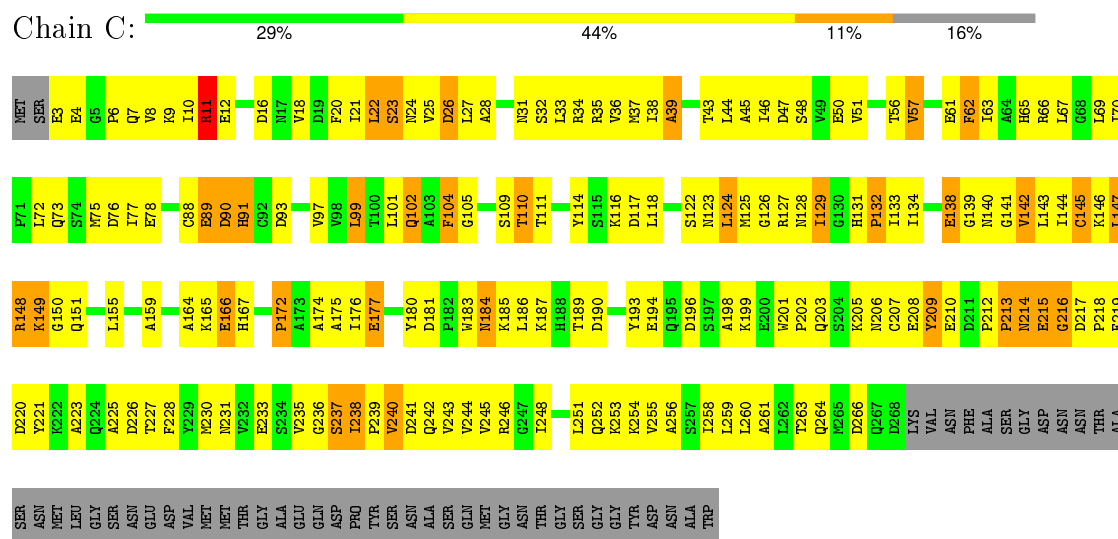




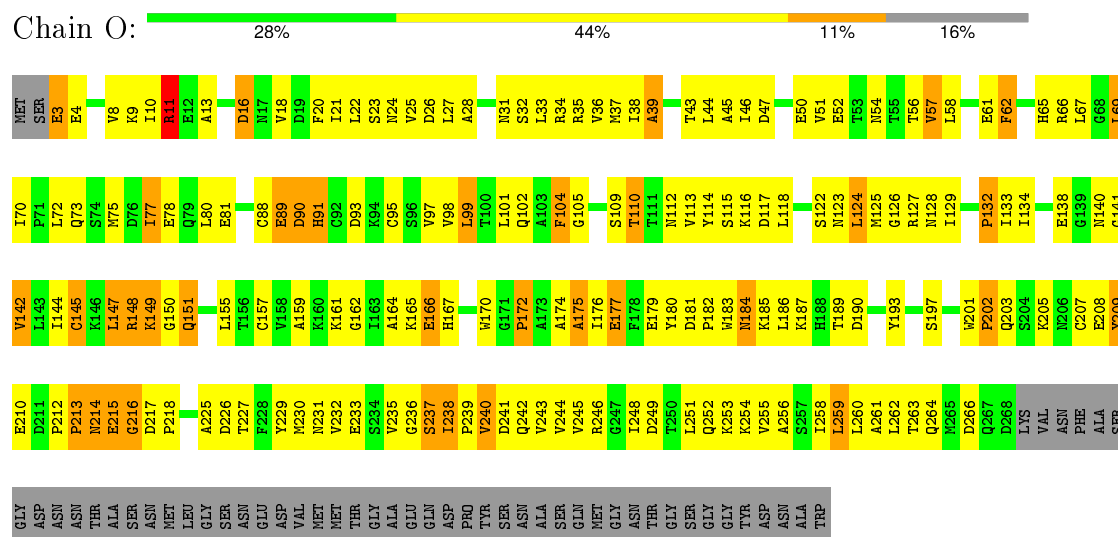


G1109	K1033	K965	P901	K841	A777	P712	B550	B590	G520	T454	L387	G325	L258	L192	L128
P1110	V1034	V966	G902	M842	K778	A713	L651	B591	L521	T457	C388	D326	Y259	K193	F129
V1113	L1037	R967	V905	K843	G779	E714	K652	B592	V522	L458	L390	R327	R261	A389	D131
L1114	S1038	R968	S906	S844	F781	A715	P593	A594	C523	Y459	D391	D332	E262	G195	
T1115	G1039	R969	G907	R846	L782	GLU	K654	B595	A524	A460	R392	D332	G263	P196	
R1116	N1040	T971	E908	D847	T793	GLU	G656	L596	A525	L461	K393	F333	S264	K134	
P1117	E1041	D909	D909	R848	K784	ASN	K657	B597	E527		D394	I394		Y202	ARG
P1118	G1042	R849	V910	K849	Y785	ASP	L658	B598	P528	G464	D397	G335	I269	F203	THR
V1119	D1043	K975	I911	L850	M786	LEU	K659	T599	E529	N465	D396	G336		T204	TTR
E1120	A1044	I976	F851	R851	V787	D722	L600	L500	G530	N466	R398	R336		I205	GLU
G1121	S1045	R852	K914	R788	K788	V723	L661	B601	Q631	G467	R398	GLY	T272	N206	ALA
S1122	P1046	R853	T915	R789	V723	GLU	M662	R602	Q632	E468	D399	THR	L273	G207	IIE
S1123	F1047	L854	T916	D790	D724	D724	A683	L603	A632	Q469	H400	ALA	Y275	E209	VAL
R1124	T1048	F855	P917	T791	K727	K727	T664	L604	I539	K470	F401	LEU	I276	K210	PRO
F1129	D1049	R856	I918	M792	R728	R728	B605	R604		L470	G402	GLY	K277	V211	GLY
F1130	T1051	R857	S919	A793	V731	V731	K606	R605	M542	A472	R405	IIE	D278	L212	ARG
G1131	S1056	K958	ASP	S858	N794	S732	D668	D608	I545	N473	K345	LVS	I280	E216	LEU
E1132	R1060	R987	GLU	M860	L795	H733	GLU	T609	S546	S474	L406	K346	I281	R217	LVS
R1135	G1066	R988	GLU	D861	L796	H734	GLY	T610	S547	S475	D407	E346	I282	S218	TTR
D1136	F1069	R989	GLU	Q862	V798	H735	GLY	P611	G548	R476	L408	R348	I283	A219	GLU
C1137	G1063	G991	LEU	E863	F798	T736	PHE	B612	T549	A477		I349	I284	G220	LEU
H1141	Y1065	I992	GLY	K864	Q800	T737	GLU	V613	D550	N479	L416	G350	I285	N221	IIE
A1144	Q1065	R993	ARG	K865	K801	F738	ASP	S614	P551	S480	F417	G351	F286	I222	ALA
S1145	S1066	R994	THR	R866	F802		VAL	R615	M552		T419	A382	K287	V223	GLU
D1156	R1080	R995	ALA	M868	L803	C741	GLU	T616	P553	L483	L420	K353	A288	Q224	GLU
A1157	F1107	R996	ALA	R869	G804	E742	B678	R617	I554	N484	F421	D354	I289	V225	SER
F1158	L1085	R997	HIS	T805	I743	H744	D618	D618	I555	R485	K422	I355	F226	G290	GLU
V1160	F1086	R998	ASP	R807	T806	H745	T680	T619	T556	Y486	K423	I356	I291	K227	ASP
H1161	Y1071	R999	ASP	R871	R807	P745	M681	R620	F557	T487	L424	D357	I292	K228	ASP
R1150	M1072	P1000	K934	E872	A808	S746	S682	B621	L558	Y488	T425	K358	P293	A229	SER
L1151	Y1073	F1001	R935	T873	R809	M747	S683	B622	S559	S489	K426	E359	D294	A230	GLU
S1155	K1079	E1004	S938	K876	E810	L749	L684	E523	M561	T491	F429	P362	E296	G295	SER
D1156	K1080	R999	T939	P877	L812	G750	M686	K625	G562	L492	R430	H363	I297	P231	GLY
A1157	Q1084	R999	P940	Q878	R815	A752	B687	T626	M563	L495	Y431	I364	I298	S235	K164
F1159	L1085	R999	L941	R879	E816	I756	L689	T628	L566	R496	M432	T366	E299	R236	F166
V1160	F1086	P1008	R942	ASN	L817	P757	V690	D629	Y569	R497	Q433	Q366	G168	V237	G168
H1161	F1087	D1009	S943	THR	F818	F758	B691	B630	P570	T498	R434	E368	A238	A238	R169
L1165	Y1091	I1011	E945	LEU	Q821	P759	Y692	B632	H572	I502	V436	G369	I240	E239	P171
C1166	A1016	A1016	N946	R884	B822	D760	D694	B633	Q573	GLY	E437	F370	V305	D304	L170
G1167	L1017	P1018	D950	M885	A823	H761	A695	B634	P571	ARG	GLU	E371	R241	R241	L172
L1168	R1096	S1019	Q951	K886	B824	M762	B696	R635	S574	ASP	ALA	S372	S242	S242	M173
M1169	R1097	G888	V953	R887	V825	Q763	B697	R636	P575	GLY	HIS	R373	A243	A243	L174
T1171	M1098	R889	L953	T889	A826	S764	B698	L637	D576	LVS	ASP	R374	L244	L244	R175
V1171	M1098	R890	V954	Y890	Y830	P765	B699	T639	A577	L508	PHE	A375	E245	E245	
I1172	V1099	D891	T955	D891	Y831	R766	S700	B640	T578	A509	ASN	F376	K246	K246	N178
A1173	D1100	R892	T956	R892	S832	H767	I701	V640	R579	K510	MET	F377	E312	G247	C179
K1174	H1101	H1025	N957	L893	T768	T768	B641	V640	V580	P511	LVS	L378	M313	S248	Y180
L1175	K1102	L1026	Q958	R894	Y769	Q770	D642	B642	F581	B512	L446	G379	L314	R249	L181
H1176	H1103	D895	Q958	R894	Y769	Q770	D642	B642	F581	B512	L446	G379	L314	R249	T185
N1177	H1104	E1028	D959	R895	Y771	S771	Q706	B644	N583	Q513	I448	N381	S252	S252	E186
N1178	A1105	C1029	L861	R896	A772	A772	Q706	B644	N583	Q513	I448	N381	S252	S252	S187
Q1179	L1106	L1030	K962	R897	D837	M773	E708	L646	V585	B515	M449	R383	D820	D820	D188
F1180	A1107	L1031	R899	L898	S838	G774	D709	B647	M586	B517	K451	R384	V323	Q255	L189
E1181	L1108	S1032	V964	A900	T840	Q776	E711	K649	V589	B519	I453	L396	I324	K257	K191

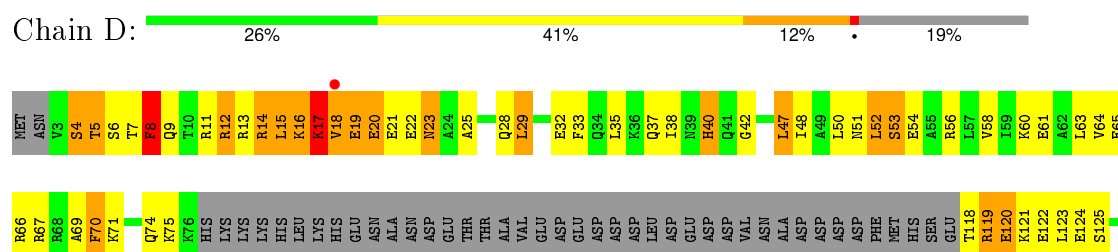
- Molecule 3: DNA-directed RNA polymerase II subunit RPB3

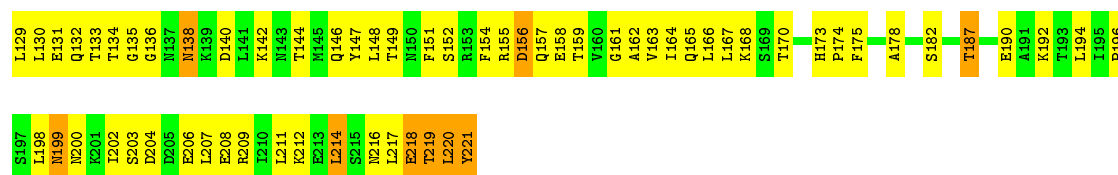


- Molecule 3: DNA-directed RNA polymerase II subunit RPB3

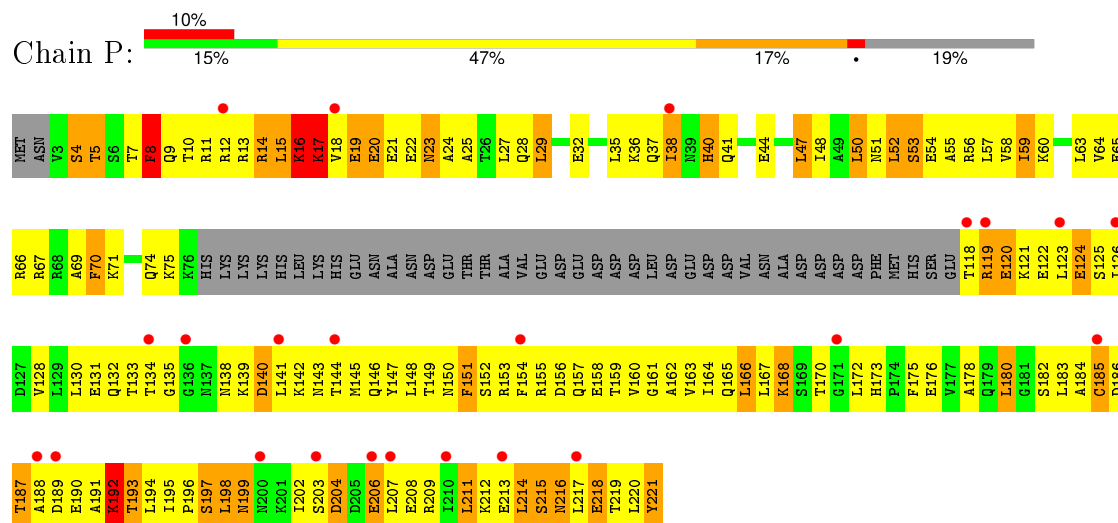


- Molecule 4: DNA-directed RNA polymerase II subunit RPB4

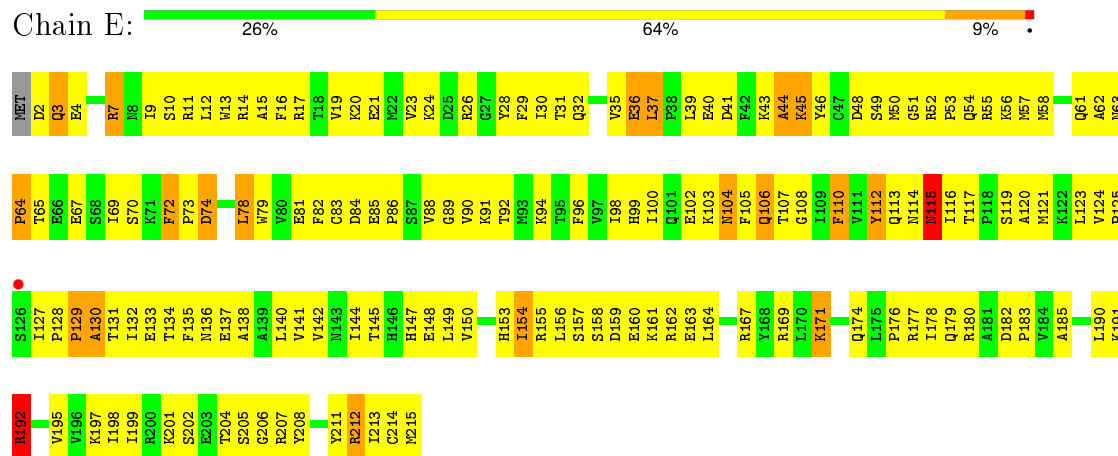




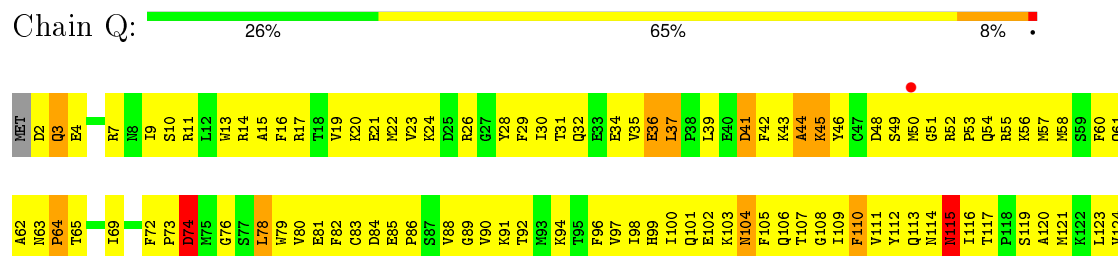
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

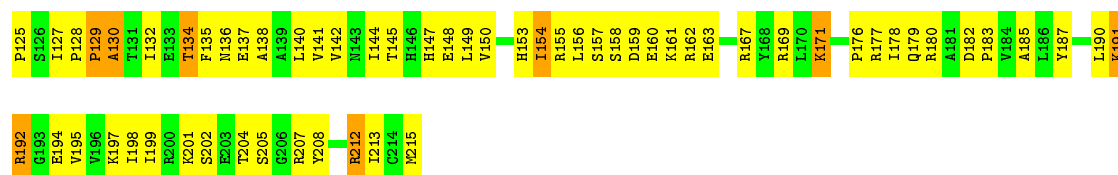


• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

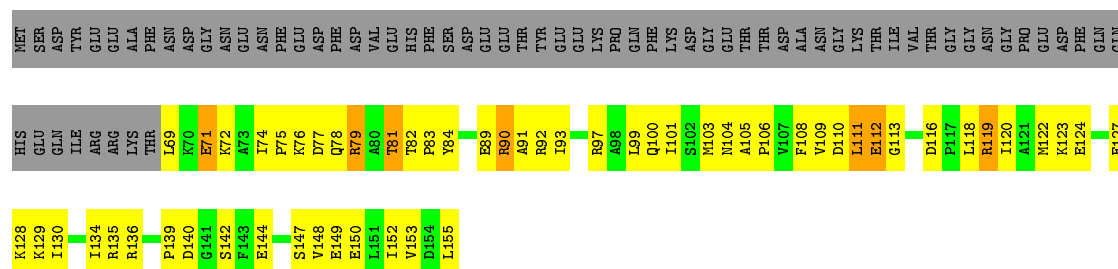


• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

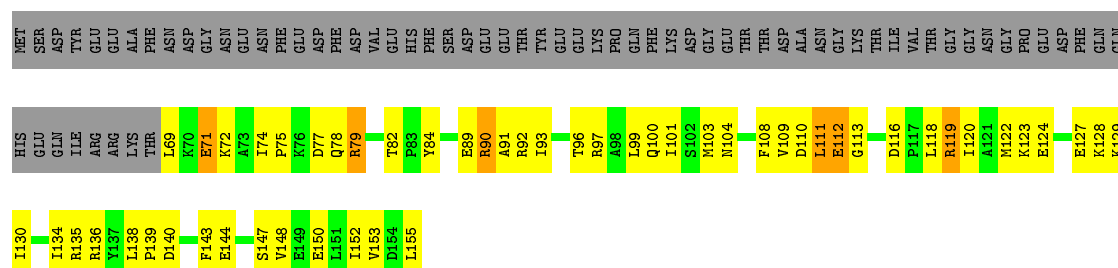




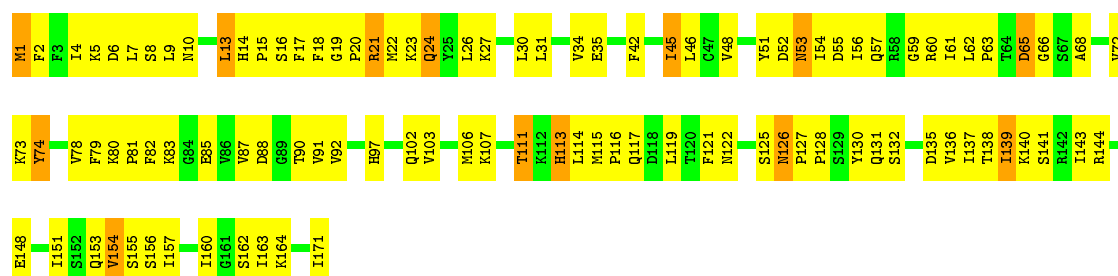
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



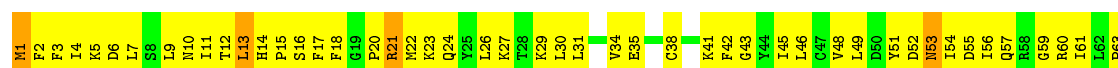
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

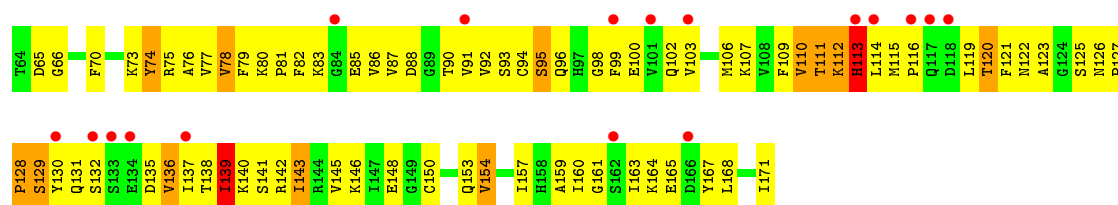


- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

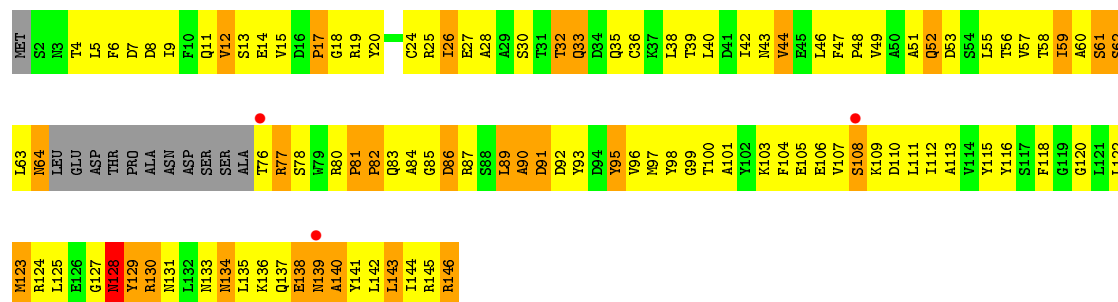
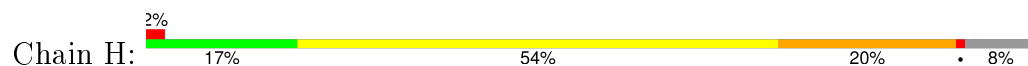


- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

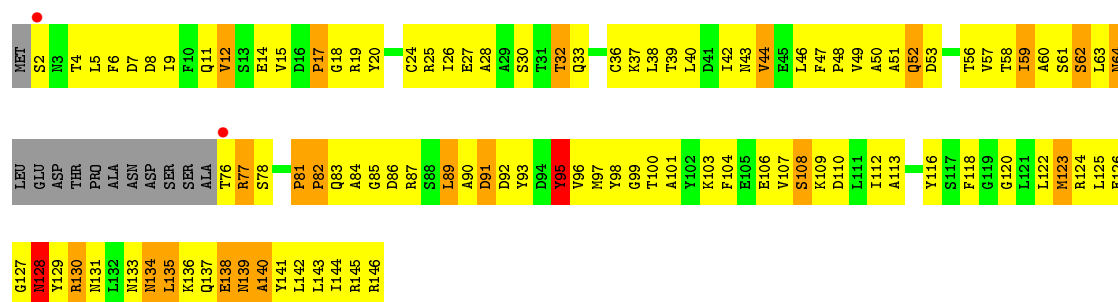




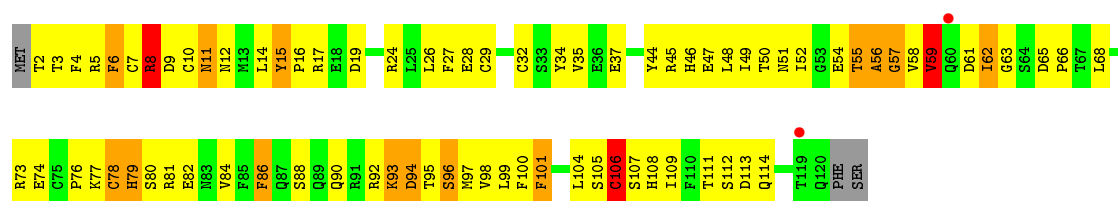
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

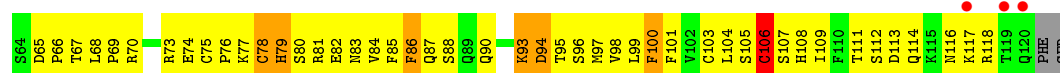


• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

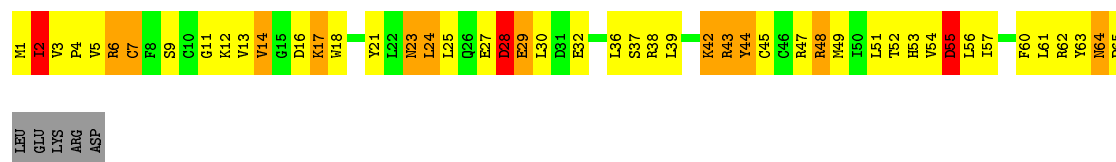


• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

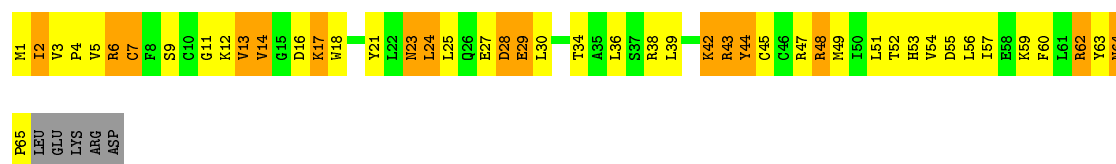




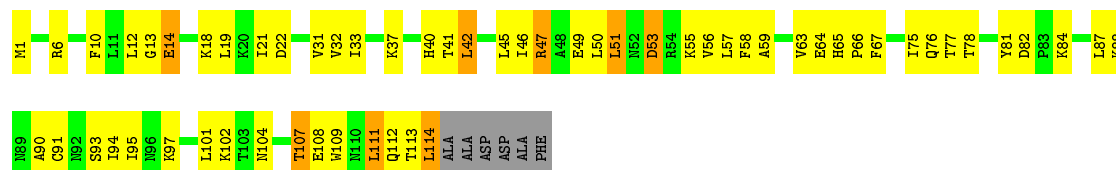
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



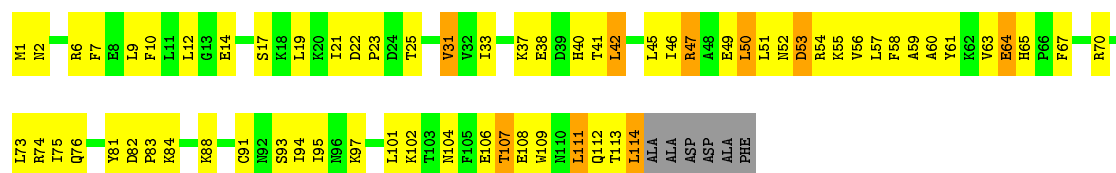
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



- Molecule 11: DNA-directed RNA polymerase II subunit RPB11

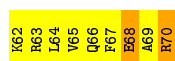


- Molecule 11: DNA-directed RNA polymerase II subunit RPB11

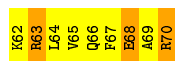
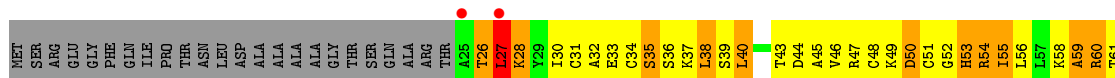


- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4





- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



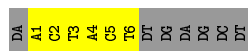
- Molecule 13: 5'-D(*AP*GP*CP*TP*CP*A*AP*GP*TP*AP*GP*TP*TP*AP*TP*GP*CP*C P*(BRU)P*GP*GP*TP*CP*AP*TP*T)-3'



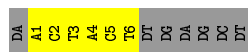
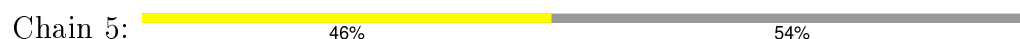
- Molecule 13: 5'-D(*AP*GP*CP*TP*CP*A*AP*GP*TP*AP*GP*TP*TP*AP*TP*GP*CP*C P*(BRU)P*GP*GP*TP*CP*AP*TP*T)-3'



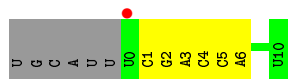
- Molecule 14: 5'-D(*A*AP*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'



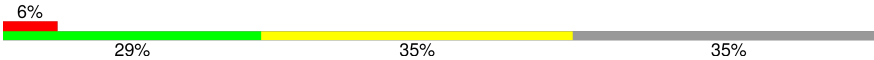
- Molecule 14: 5'-D(*A*AP*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'

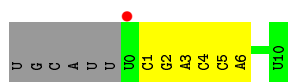


- Molecule 15: 5'-R(*UP*GP*CP*AP*UP*U*UP*CP*GP*AP*CP*CP*AP*GP*GP*CP*U)-3',



- Molecule 15: 5'-R(*UP*GP*CP*AP*UP*U*UP*CP*GP*AP*CP*CP*AP*GP*GP*CP*U)-3',

Chain 6: 



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	394.26Å 221.61Å 283.45Å 90.00° 90.90° 90.00°	Depositor
Resolution (Å)	40.00 – 3.20 38.25 – 3.02	Depositor EDS
% Data completeness (in resolution range)	95.6 (40.00-3.20) 85.7 (38.25-3.02)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 3.01Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.233 , 0.252 0.238 , 0.244	Depositor DCC
R_{free} test set	18782 reflections (5.20%)	DCC
Wilson B-factor (Å ²)	62.6	Xtriage
Anisotropy	0.223	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
Estimated twinning fraction	0.017 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.018 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.017 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.017 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.257 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	2 of 422548 reflections (0.000%)	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	63664	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.35% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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, BRU

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	0.50	0/11342	0.77	8/15337 (0.1%)
1	M	0.50	0/11342	0.78	8/15337 (0.1%)
2	B	0.48	0/8948	0.74	1/12062 (0.0%)
2	N	0.48	1/8948 (0.0%)	0.74	1/12062 (0.0%)
3	C	0.49	0/2133	0.73	1/2891 (0.0%)
3	O	0.48	0/2133	0.74	1/2891 (0.0%)
4	D	0.44	0/1444	0.72	1/1935 (0.1%)
4	P	0.53	0/1444	0.85	5/1935 (0.3%)
5	E	0.46	0/1788	0.69	1/2406 (0.0%)
5	Q	0.46	0/1788	0.70	1/2406 (0.0%)
6	F	0.57	0/717	0.82	1/967 (0.1%)
6	R	0.56	0/717	0.82	1/967 (0.1%)
7	G	0.46	0/1368	0.75	1/1844 (0.1%)
7	S	0.57	0/1368	0.86	1/1844 (0.1%)
8	H	0.43	0/1094	0.71	0/1481
8	T	0.42	0/1094	0.72	0/1481
9	I	0.42	0/989	0.71	0/1331
9	U	0.45	0/989	0.71	0/1331
10	J	0.51	0/541	0.83	0/727
10	V	0.48	0/541	0.80	0/727
11	K	0.46	0/937	0.67	0/1265
11	W	0.48	0/937	0.68	0/1265
12	L	0.58	0/365	0.84	0/485
12	X	0.57	0/365	0.84	0/485
13	1	0.60	0/389	0.96	0/597
13	4	0.60	0/389	0.94	0/597
14	2	0.61	0/130	0.78	0/198
14	5	0.60	0/130	0.78	0/198
15	3	0.56	0/256	0.74	0/397
15	6	0.54	0/256	0.74	0/397
All	All	0.49	1/64882 (0.0%)	0.76	32/87846 (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	B	0	1
2	N	0	2
13	1	0	4
13	4	0	4
All	All	0	11

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	N	1137	CYS	CB-SG	-5.78	1.72	1.81

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	P	180	LEU	CA-CB-CG	-7.93	97.05	115.30
4	P	166	LEU	CA-CB-CG	6.53	130.32	115.30
1	A	56	PRO	N-CA-C	-6.37	95.55	112.10
1	M	56	PRO	N-CA-C	-6.35	95.59	112.10
3	C	39	ALA	N-CA-C	5.97	127.11	111.00

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
13	1	18	DA	Sidechain
13	1	19	DT	Sidechain
13	1	20	DG	Sidechain
13	1	21	DC	Sidechain
2	B	833	TYR	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	11143	0	11217	1159	0
1	M	11143	0	11217	1163	0
2	B	8779	0	8808	1066	0
2	N	8779	0	8808	1078	0
3	C	2095	0	2051	226	0
3	O	2095	0	2051	227	0
4	D	1434	0	1460	152	0
4	P	1434	0	1460	273	0
5	E	1752	0	1776	187	0
5	Q	1752	0	1776	202	0
6	F	705	0	731	85	0
6	R	705	0	731	75	0
7	G	1340	0	1357	145	0
7	S	1340	0	1357	205	0
8	H	1076	0	1046	171	0
8	T	1076	0	1046	154	0
9	I	971	0	929	117	0
9	U	971	0	929	126	0
10	J	532	0	542	97	0
10	V	532	0	542	95	0
11	K	919	0	929	81	0
11	W	919	0	929	84	0
12	L	363	0	387	87	0
12	X	363	0	387	84	0
13	1	368	0	203	27	0
13	4	368	0	203	27	0
14	2	117	0	70	13	0
14	5	117	0	70	11	0
15	3	230	0	121	8	0
15	6	230	0	121	8	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
16	M	2	0	0	0	0
16	N	1	0	0	0	0
16	O	1	0	0	0	0
16	U	2	0	0	0	0
16	V	1	0	0	0	0
16	X	1	0	0	0	0
All	All	63664	0	63254	6846	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 54.

The worst 5 of 6846 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:N:510:LYS:HG3	2:N:511:PRO:HD3	1.21	1.17
1:A:855:THR:HG21	1:A:857:ARG:HE	1.08	1.16
9:U:111:THR:HG22	9:U:113:ASP:H	1.05	1.15
5:Q:124:VAL:HG13	5:Q:132:ILE:HB	1.28	1.15
8:H:4:THR:HA	8:H:60:ALA:HB2	1.26	1.14

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1406/1733 (81%)	1075 (76%)	225 (16%)	106 (8%)	1	9
1	M	1406/1733 (81%)	1073 (76%)	228 (16%)	105 (8%)	1	9
2	B	1082/1224 (88%)	800 (74%)	186 (17%)	96 (9%)	1	5
2	N	1082/1224 (88%)	798 (74%)	186 (17%)	98 (9%)	1	5
3	C	264/318 (83%)	202 (76%)	41 (16%)	21 (8%)	1	7
3	O	264/318 (83%)	203 (77%)	42 (16%)	19 (7%)	1	10
4	D	174/221 (79%)	120 (69%)	37 (21%)	17 (10%)	1	4
4	P	174/221 (79%)	122 (70%)	34 (20%)	18 (10%)	1	4
5	E	212/215 (99%)	155 (73%)	41 (19%)	16 (8%)	1	9
5	Q	212/215 (99%)	159 (75%)	37 (18%)	16 (8%)	1	9
6	F	85/155 (55%)	72 (85%)	11 (13%)	2 (2%)	7	43
6	R	85/155 (55%)	72 (85%)	11 (13%)	2 (2%)	7	43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	G	169/171 (99%)	141 (83%)	23 (14%)	5 (3%)	5	35
7	S	169/171 (99%)	139 (82%)	23 (14%)	7 (4%)	3	27
8	H	130/146 (89%)	85 (65%)	25 (19%)	20 (15%)	0	1
8	T	130/146 (89%)	85 (65%)	25 (19%)	20 (15%)	0	1
9	I	117/122 (96%)	77 (66%)	28 (24%)	12 (10%)	1	4
9	U	117/122 (96%)	78 (67%)	28 (24%)	11 (9%)	1	5
10	J	63/70 (90%)	43 (68%)	9 (14%)	11 (18%)	0	0
10	V	63/70 (90%)	42 (67%)	10 (16%)	11 (18%)	0	0
11	K	112/120 (93%)	89 (80%)	20 (18%)	3 (3%)	6	39
11	W	112/120 (93%)	89 (80%)	19 (17%)	4 (4%)	4	30
12	L	44/70 (63%)	19 (43%)	15 (34%)	10 (23%)	0	0
12	X	44/70 (63%)	19 (43%)	15 (34%)	10 (23%)	0	0
All	All	7716/9130 (84%)	5757 (75%)	1319 (17%)	640 (8%)	1	7

5 of 640 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	43	GLU
1	A	57	ARG
1	A	62	ASP
1	A	63	ARG

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	1239/1520 (82%)	1116 (90%)	123 (10%)	10	38
1	M	1239/1520 (82%)	1107 (89%)	132 (11%)	8	34
2	B	958/1061 (90%)	860 (90%)	98 (10%)	9	36
2	N	958/1061 (90%)	853 (89%)	105 (11%)	8	33

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	234/274 (85%)	212 (91%)	22 (9%)	11	41
3	O	234/274 (85%)	207 (88%)	27 (12%)	7	30
4	D	160/200 (80%)	136 (85%)	24 (15%)	3	17
4	P	160/200 (80%)	127 (79%)	33 (21%)	1	7
5	E	196/197 (100%)	183 (93%)	13 (7%)	21	61
5	Q	196/197 (100%)	184 (94%)	12 (6%)	23	64
6	F	77/137 (56%)	71 (92%)	6 (8%)	16	53
6	R	77/137 (56%)	72 (94%)	5 (6%)	21	61
7	G	152/152 (100%)	140 (92%)	12 (8%)	15	53
7	S	152/152 (100%)	134 (88%)	18 (12%)	6	29
8	H	118/128 (92%)	105 (89%)	13 (11%)	8	33
8	T	118/128 (92%)	108 (92%)	10 (8%)	13	47
9	I	113/116 (97%)	101 (89%)	12 (11%)	8	34
9	U	113/116 (97%)	103 (91%)	10 (9%)	12	45
10	J	60/65 (92%)	51 (85%)	9 (15%)	3	17
10	V	60/65 (92%)	53 (88%)	7 (12%)	7	30
11	K	99/102 (97%)	94 (95%)	5 (5%)	29	70
11	W	99/102 (97%)	90 (91%)	9 (9%)	12	42
12	L	40/57 (70%)	34 (85%)	6 (15%)	3	17
12	X	40/57 (70%)	33 (82%)	7 (18%)	2	12
All	All	6892/8018 (86%)	6174 (90%)	718 (10%)	9	35

5 of 718 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	7	CYS
1	M	631	HIS
7	S	75	ARG
11	K	51	LEU
1	M	221	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 226 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	122	ASN
1	M	316	GLN
6	R	100	GLN
8	H	64	ASN
10	J	53	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	3	10/17 (58%)	0	0
15	6	10/17 (58%)	0	0
All	All	20/34 (58%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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
13	BRU	1	23	13,15	13,21,22	4.63	4 (30%)	16,30,33	4.03	3 (18%)
13	BRU	4	23	13,15	13,21,22	4.65	4 (30%)	16,30,33	4.06	3 (18%)

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
13	BRU	1	23	13,15	-	0/3/21/22	0/2/2/2
13	BRU	4	23	13,15	-	0/3/21/22	0/2/2/2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	4	23	BRU	BR-C5	-15.07	1.50	1.90
13	1	23	BRU	BR-C5	-14.98	1.50	1.90
13	1	23	BRU	C6-N1	2.87	1.39	1.35
13	4	23	BRU	C6-N1	2.88	1.39	1.35
13	4	23	BRU	C4-N3	3.43	1.39	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	4	23	BRU	C5-C4-N3	-8.15	115.30	124.00
13	1	23	BRU	C5-C4-N3	-8.09	115.36	124.00
13	4	23	BRU	C5-C6-N1	2.09	123.89	119.79
13	1	23	BRU	C5-C6-N1	2.11	123.93	119.79
13	1	23	BRU	C4-N3-C2	13.55	126.96	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	1	23	BRU	6	0
13	4	23	BRU	6	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 16 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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	1416/1733 (81%)	-0.07	8 (0%) 90 84	12, 52, 93, 119	0
1	M	1416/1733 (81%)	-0.06	14 (0%) 84 75	10, 53, 94, 123	0
2	B	1104/1224 (90%)	-0.02	8 (0%) 89 83	12, 62, 103, 120	0
2	N	1104/1224 (90%)	0.04	13 (1%) 81 69	16, 65, 104, 121	0
3	C	266/318 (83%)	-0.11	0 100 100	24, 52, 83, 100	0
3	O	266/318 (83%)	-0.12	0 100 100	25, 54, 85, 106	0
4	D	178/221 (80%)	-0.04	1 (0%) 90 84	36, 68, 100, 108	0
4	P	178/221 (80%)	0.81	23 (12%) 5 3	55, 85, 105, 113	0
5	E	214/215 (99%)	-0.01	1 (0%) 91 87	35, 80, 106, 114	0
5	Q	214/215 (99%)	0.10	1 (0%) 91 87	35, 82, 107, 119	0
6	F	87/155 (56%)	-0.28	0 100 100	13, 34, 62, 78	0
6	R	87/155 (56%)	-0.21	0 100 100	15, 34, 63, 76	0
7	G	171/171 (100%)	-0.08	0 100 100	37, 55, 85, 99	0
7	S	171/171 (100%)	0.75	17 (9%) 9 5	37, 69, 110, 116	0
8	H	134/146 (91%)	0.21	3 (2%) 65 50	60, 88, 105, 114	0
8	T	134/146 (91%)	0.18	2 (1%) 76 63	66, 89, 104, 116	0
9	I	119/122 (97%)	0.11	2 (1%) 73 60	47, 81, 102, 117	0
9	U	119/122 (97%)	0.11	3 (2%) 61 47	45, 84, 102, 119	0
10	J	65/70 (92%)	-0.16	0 100 100	23, 52, 74, 91	0
10	V	65/70 (92%)	-0.15	0 100 100	28, 53, 78, 91	0
11	K	114/120 (95%)	-0.26	0 100 100	23, 54, 72, 83	0
11	W	114/120 (95%)	-0.20	0 100 100	21, 54, 74, 84	0
12	L	46/70 (65%)	0.28	4 (8%) 13 7	37, 89, 107, 107	0
12	X	46/70 (65%)	0.29	2 (4%) 39 25	42, 93, 107, 108	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
13	1	17/26 (65%)	0.20	1 (5%) 26 14	47, 101, 140, 144	0
13	4	17/26 (65%)	0.26	0 100 100	50, 102, 139, 142	0
14	2	6/13 (46%)	0.32	0 100 100	114, 121, 127, 133	0
14	5	6/13 (46%)	0.34	0 100 100	114, 121, 129, 136	0
15	3	11/17 (64%)	0.17	1 (9%) 11 6	88, 93, 131, 133	0
15	6	11/17 (64%)	0.11	1 (9%) 11 6	88, 96, 130, 133	0
All	All	7896/9242 (85%)	0.01	105 (1%) 79 67	10, 61, 102, 144	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	734	HIS	5.1
4	P	188	ALA	5.0
7	S	137	ILE	4.9
4	P	185	CYS	4.6
7	S	116	PRO	4.4

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(Å ²)	Q<0.9
13	BRU	1	23	20/21	0.70	0.21	-	85,89,94,97	0
13	BRU	4	23	20/21	0.75	0.19	-	83,89,95,98	0

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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(Å ²)	Q<0.9
16	ZN	V	9998	1/1	0.99	0.25	1.91	52,52,52,52	0
16	ZN	J	9990	1/1	0.97	0.23	1.80	52,52,52,52	0
16	ZN	I	9988	1/1	0.98	0.23	1.35	65,65,65,65	0
16	ZN	C	9987	1/1	0.99	0.22	1.01	28,28,28,28	0
16	ZN	O	9995	1/1	0.98	0.23	0.86	43,43,43,43	0
16	ZN	U	9996	1/1	0.99	0.22	0.84	71,71,71,71	0
16	ZN	M	9992	1/1	0.95	0.22	0.48	74,74,74,74	0
16	ZN	N	9994	1/1	0.98	0.22	0.48	38,38,38,38	0
16	ZN	B	9986	1/1	0.99	0.24	0.45	33,33,33,33	0
16	ZN	A	9985	1/1	0.98	0.23	-0.21	40,40,40,40	0
16	ZN	A	9984	1/1	0.96	0.19	-0.29	70,70,70,70	0
16	ZN	M	9993	1/1	0.99	0.23	-0.61	37,37,37,37	0
16	ZN	I	9989	1/1	0.90	0.17	-0.66	117,117,117,117	0
16	ZN	U	9997	1/1	0.97	0.18	-0.81	119,119,119,119	0
16	ZN	L	9991	1/1	0.99	0.15	-0.93	90,90,90,90	0
16	ZN	X	9999	1/1	0.97	0.17	-1.08	103,103,103,103	0

6.5 Other polymers

There are no such residues in this entry.