wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2023 – 02:25 PM EDT

PDB ID	:	3H0G
Title	:	RNA Polymerase II from Schizosaccharomyces pombe
Authors	:	Spahr, H.; Calero, G.; Bushnell, D.A.; Kornberg, R.D.
Deposited on	:	2009-04-09
Resolution	:	3.65 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 (i) 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 (i)) were used in the production of this report:

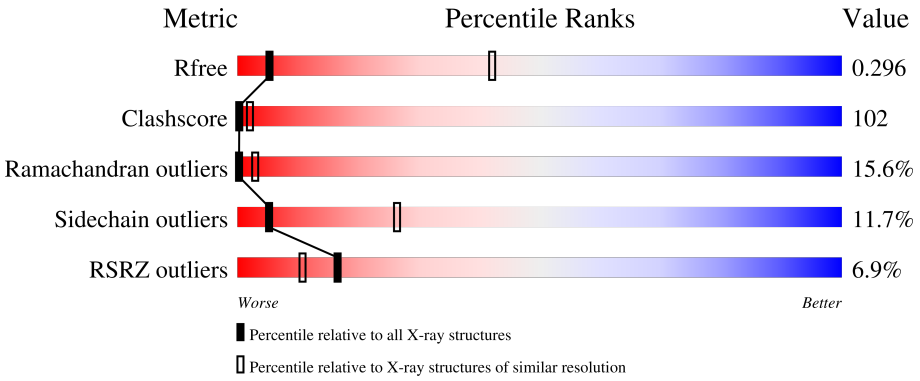
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.65 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.



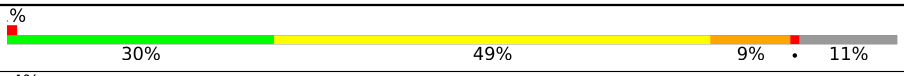
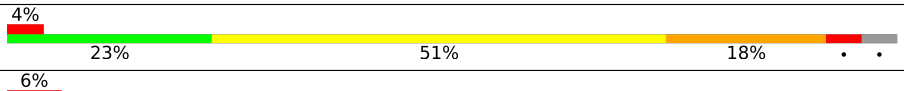
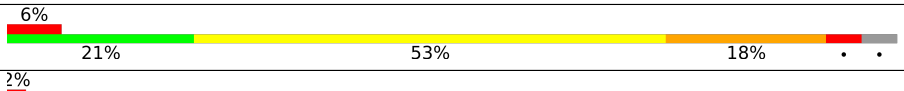
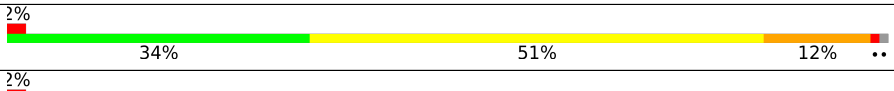
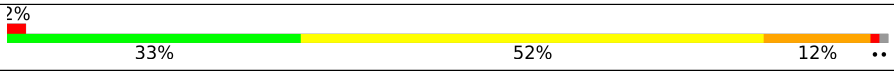
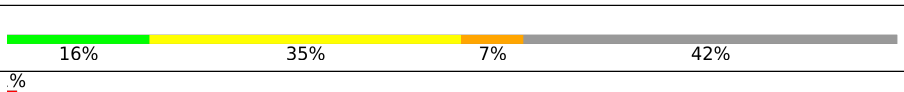
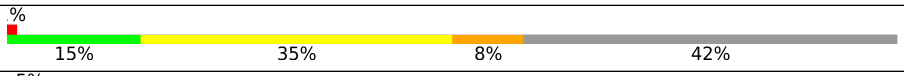
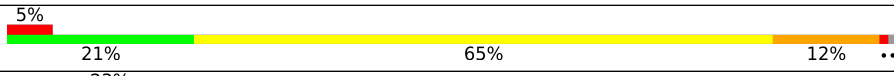
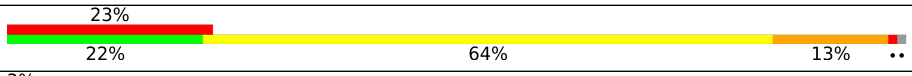
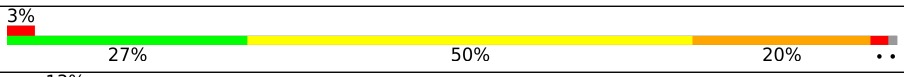
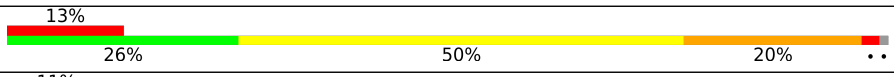
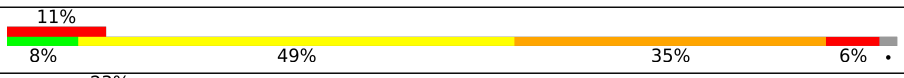
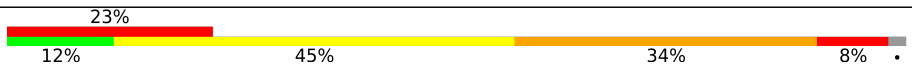
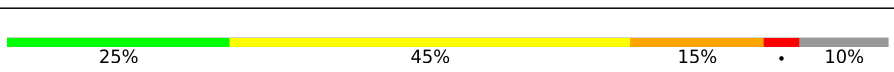
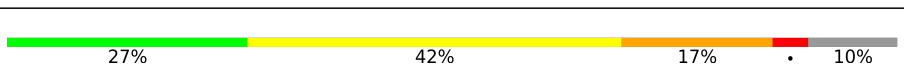
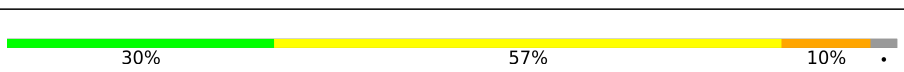
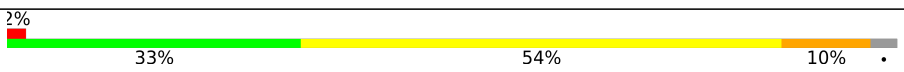
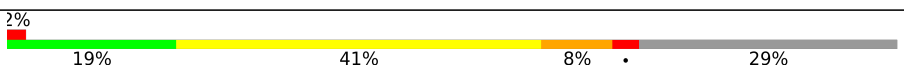
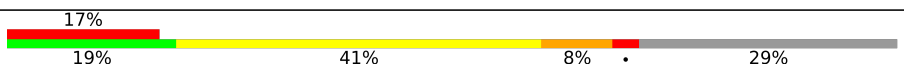
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)

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 <=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	1752	5% 19% 45% 18% • 15%
1	M	1752	10% 19% 45% 17% • 16%
2	B	1210	2% 25% 54% 14% • 5%
2	N	1210	10% 25% 54% 14% • 5%
3	C	297	30% 48% 10% • 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	O	297	
4	D	135	
4	P	135	
5	E	210	
5	Q	210	
6	F	142	
6	R	142	
7	G	172	
7	S	172	
8	H	125	
8	T	125	
9	I	113	
9	U	113	
10	J	71	
10	V	71	
11	K	123	
11	W	123	
12	L	63	
12	X	63	

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
14	MG	A	2458	-	-	-	X

2 Entry composition (i)

There are 14 unique types of molecules in this entry. The entry contains 62870 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	1496	Total	C	N	O	S	0	0	0
			11802	7415	2071	2246	70			
1	M	1476	Total	C	N	O	S	0	0	0
			11666	7334	2047	2216	69			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1150	Total	C	N	O	S	0	0	0
			9180	5772	1630	1716	62			
2	N	1150	Total	C	N	O	S	0	0	0
			9180	5772	1630	1716	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	263	Total	C	N	O	S	0	0	0
			2088	1315	355	406	12			
3	O	263	Total	C	N	O	S	0	0	0
			2088	1315	355	406	12			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit rpb4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	130	Total	C	N	O	S	0	0	0
			1036	649	176	205	6			
4	P	130	Total	C	N	O	S	0	0	0
			1036	649	176	205	6			

- 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	207	Total	C	N	O	S	0	0	0
			1663	1050	301	306	6			
5	Q	207	Total	C	N	O	S	0	0	0
			1663	1050	301	306	6			

- 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	83	Total	C	N	O	S	0	0	0
			656	416	112	125	3			
6	R	83	Total	C	N	O	S	0	0	0
			656	416	112	125	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit rpb7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	170	Total	C	N	O	S	0	0	0
			1330	860	217	247	6			
7	S	170	Total	C	N	O	S	0	0	0
			1330	860	217	247	6			

- 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	124	Total	C	N	O	S	0	0	0
			996	631	167	195	3			
8	T	124	Total	C	N	O	S	0	0	0
			996	631	167	195	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	111	Total	C	N	O	S	0	0	0
			902	551	164	176	11			
9	U	111	Total	C	N	O	S	0	0	0
			902	551	164	176	11			

- 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	64	Total	C	N	O	S	0	0	0
			518	330	87	94	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	V	64	Total	C	N	O	S	0	0	0
			518	330	87	94	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	119	Total	C	N	O	S	0	0	0
			955	608	159	182	6			
11	W	119	Total	C	N	O	S	0	0	0
			955	608	159	182	6			

- 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	45	Total	C	N	O	S	0	0	0
			368	225	74	61	8			
12	X	45	Total	C	N	O	S	0	0	0
			368	225	74	61	8			

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	V	1	Total Zn 1 1	0	0
13	X	1	Total Zn 1 1	0	0

- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

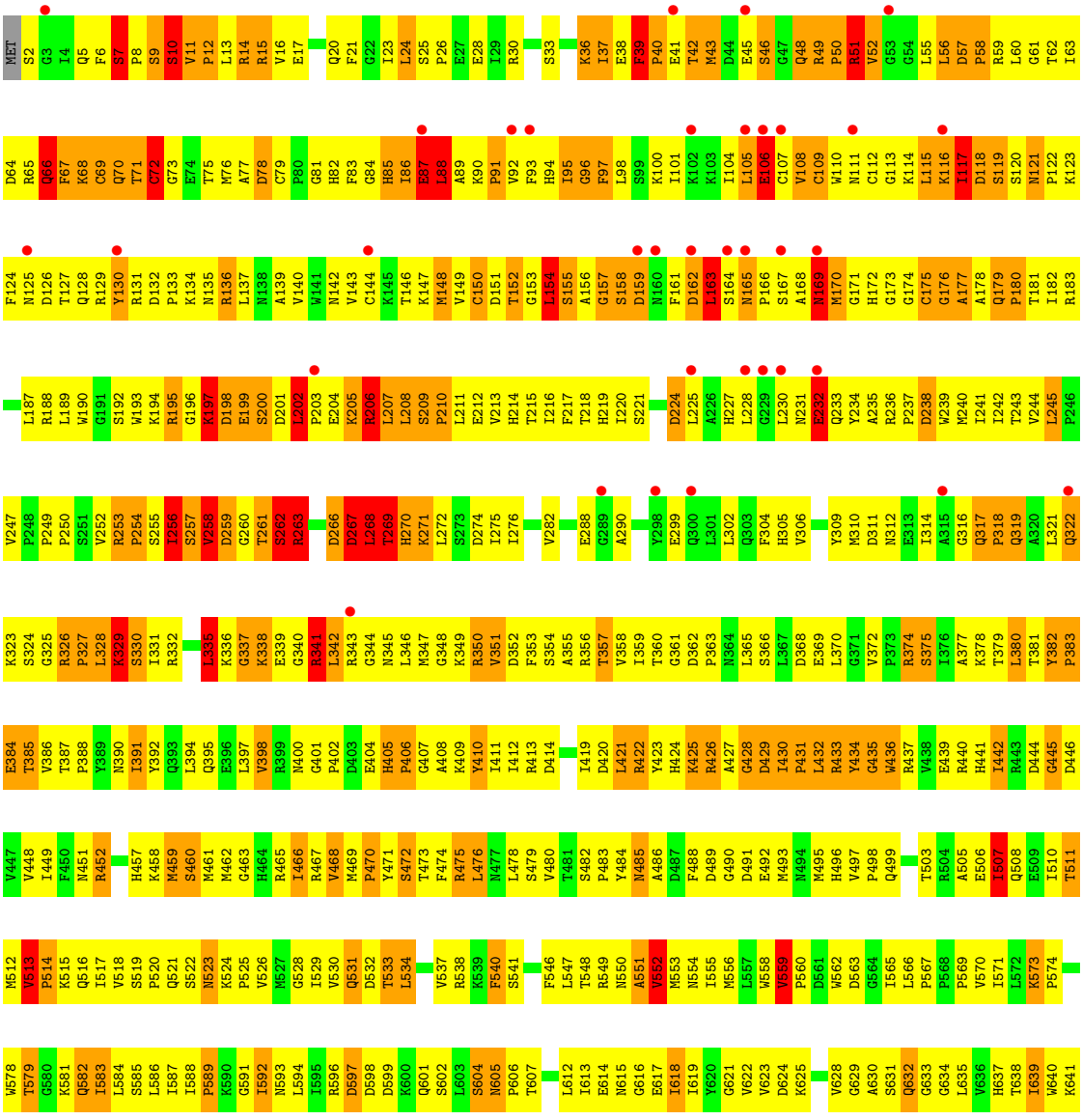
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	A	1	Total Mg 1 1	0	0
14	M	1	Total Mg 1 1	0	0

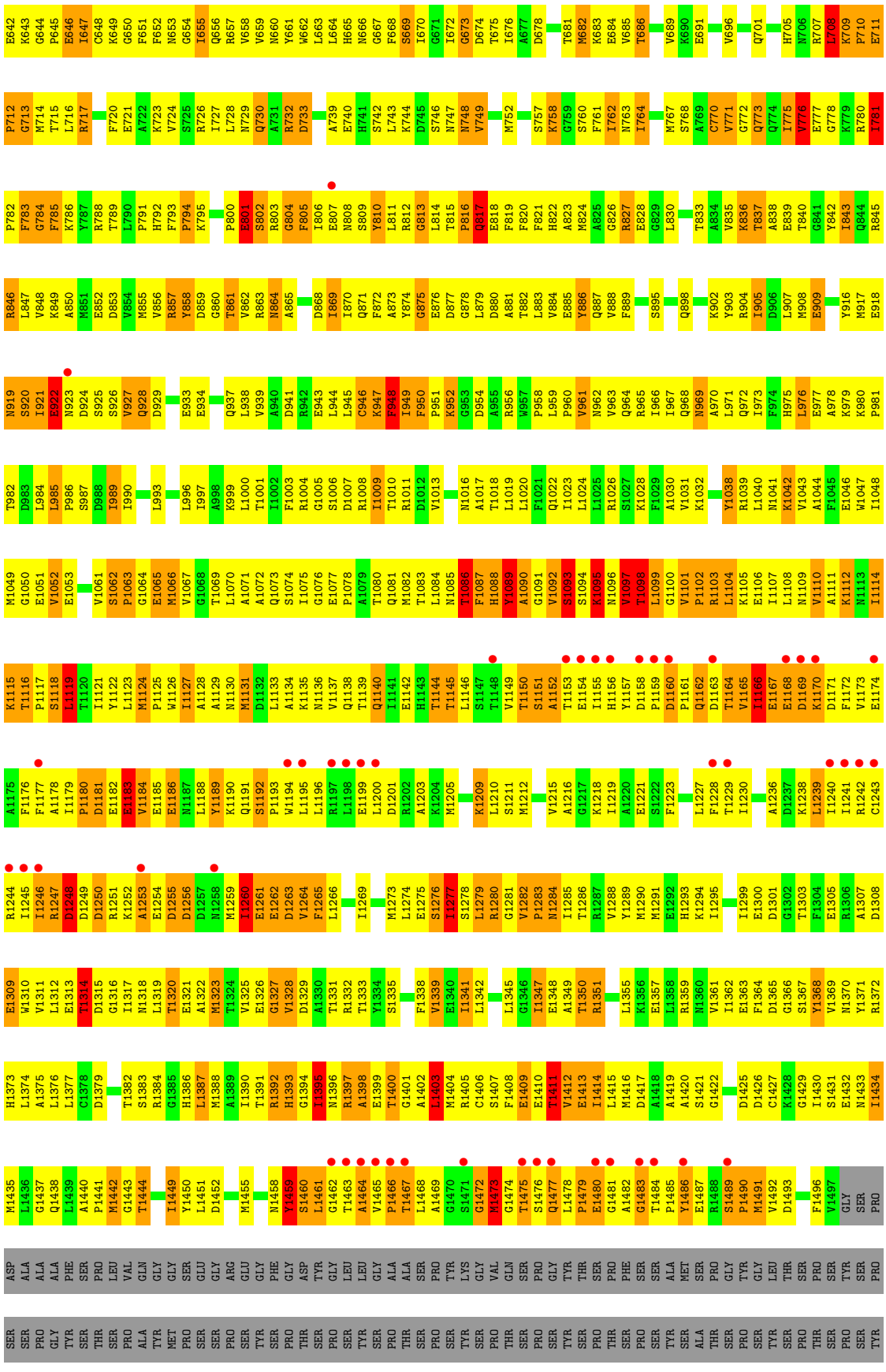
3 Residue-property plots (i)

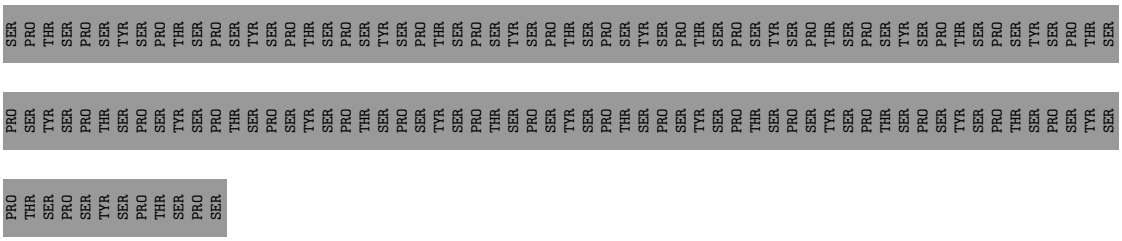
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

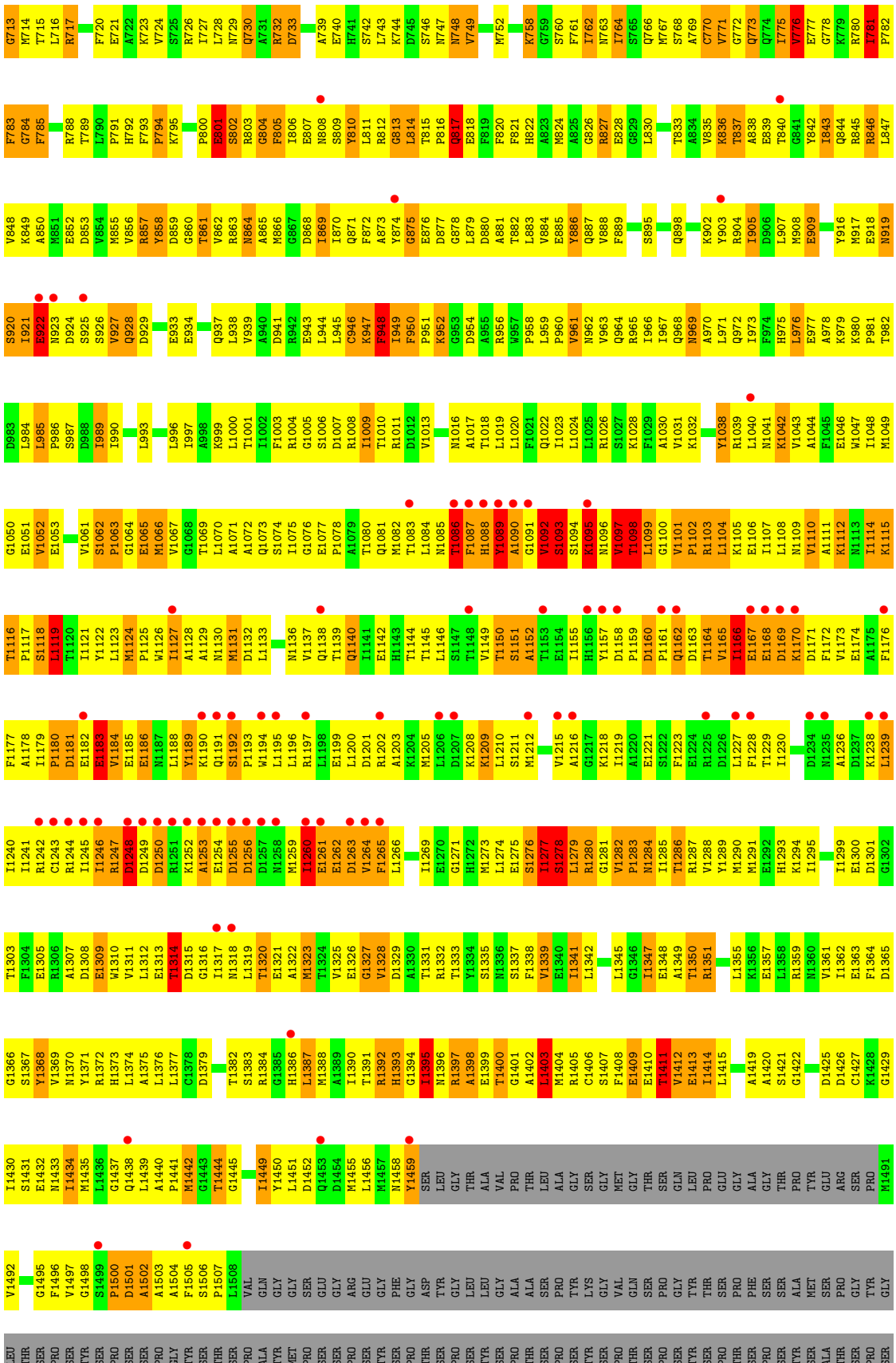
- Molecule 1: DNA-directed RNA polymerase II subunit rpb1

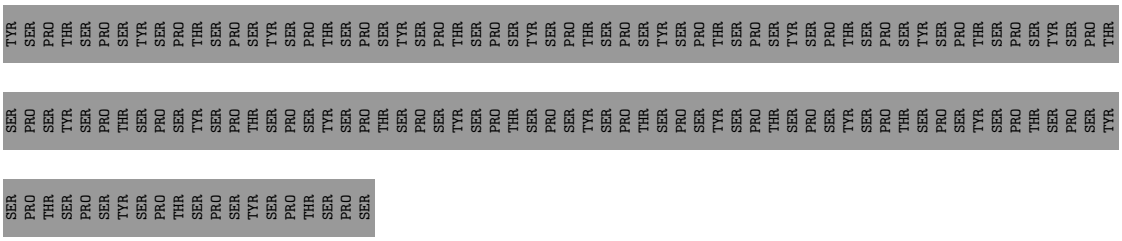
Chain A:





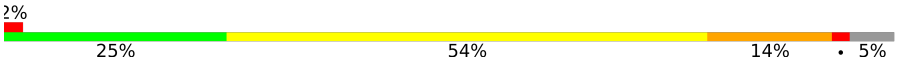

SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR

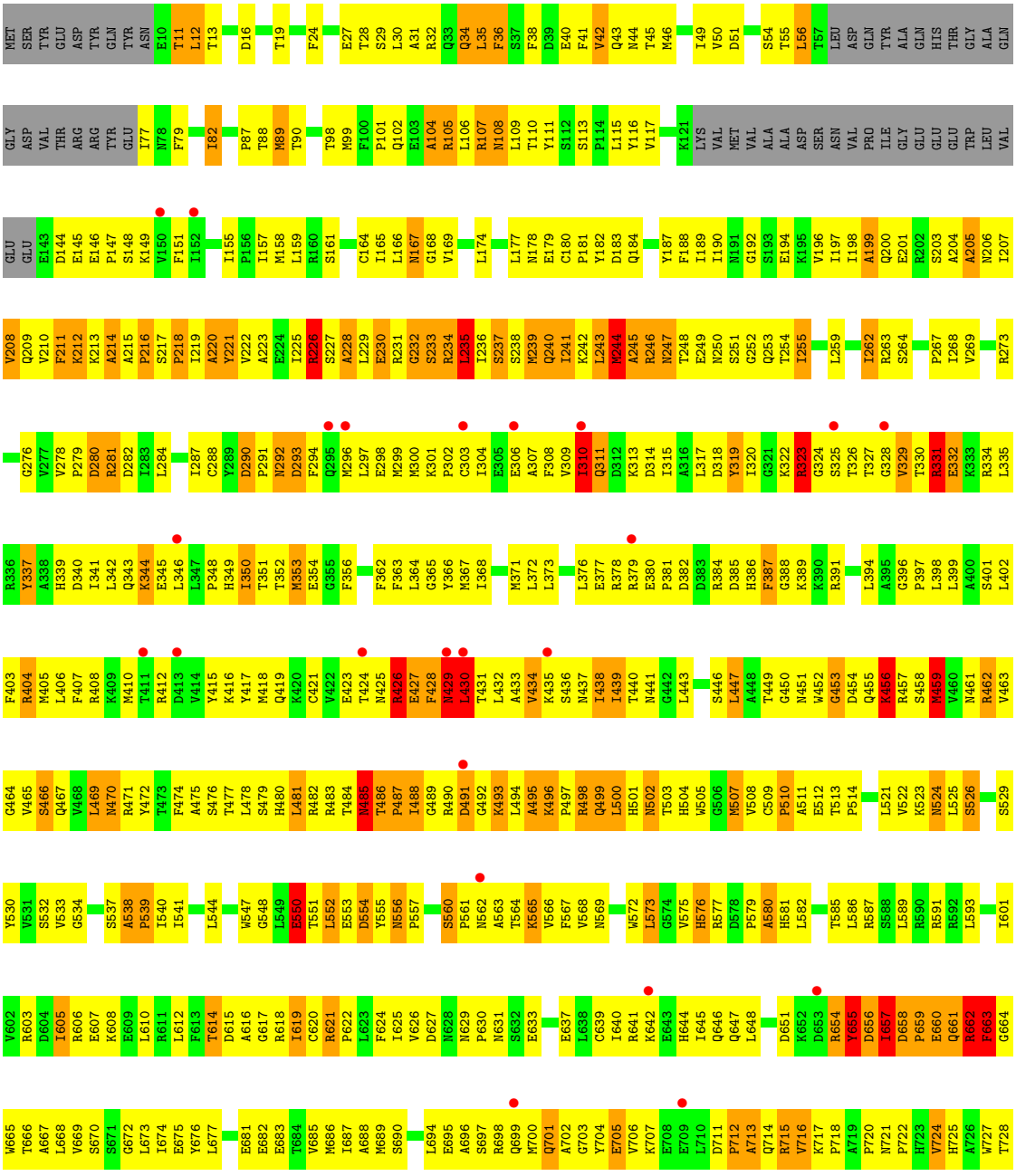


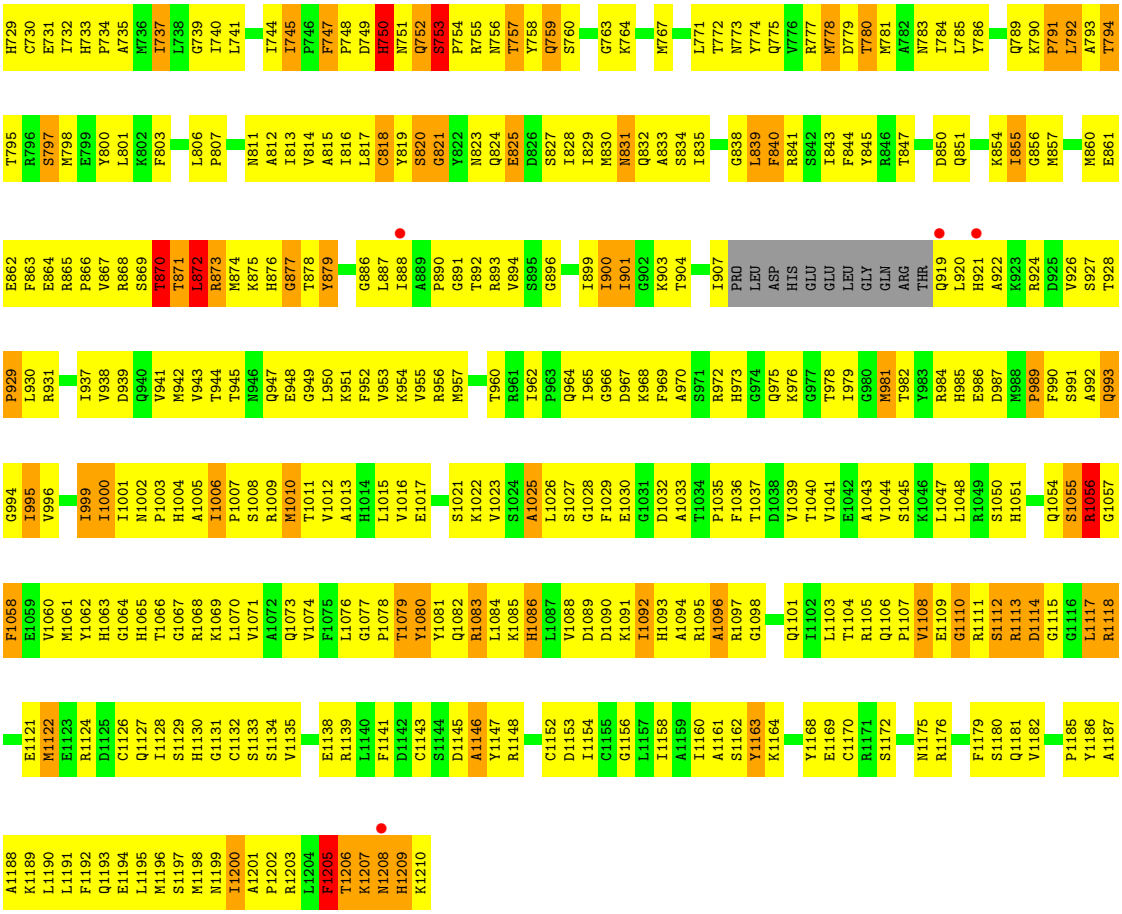


• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

Chain B:

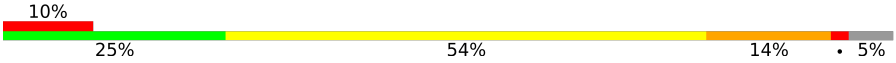


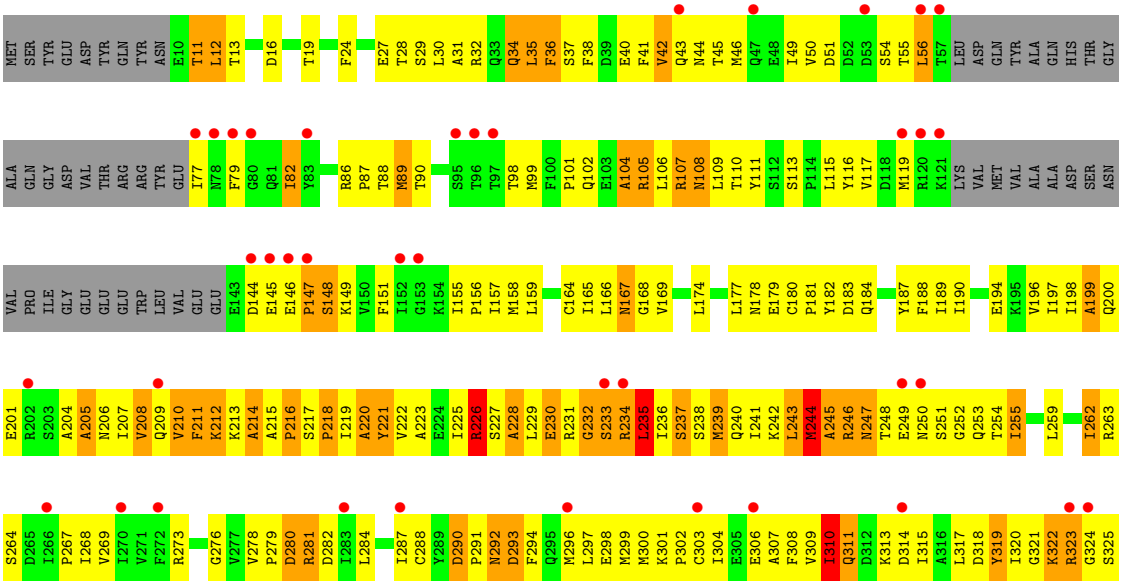


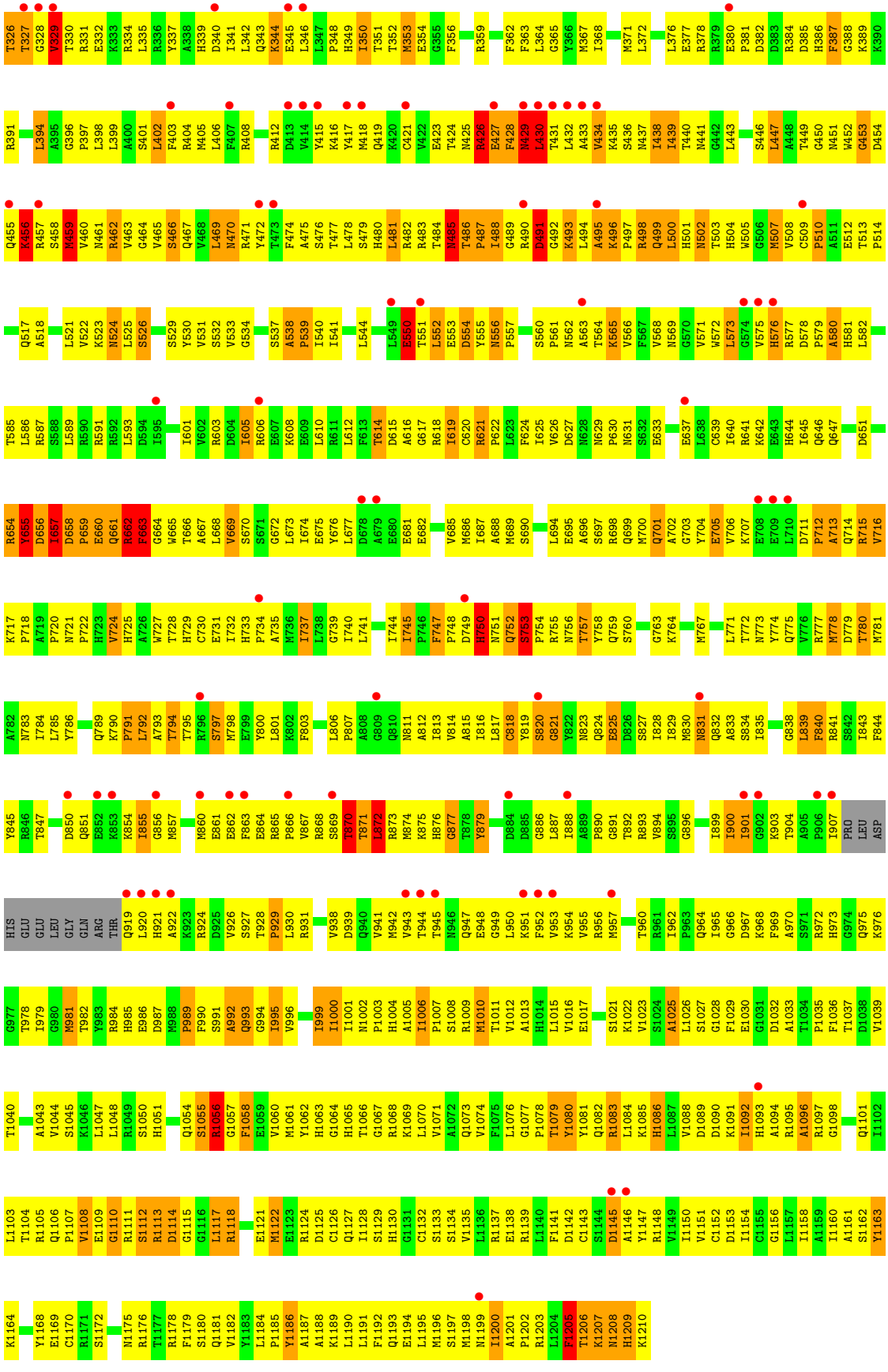


• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

Chain N:

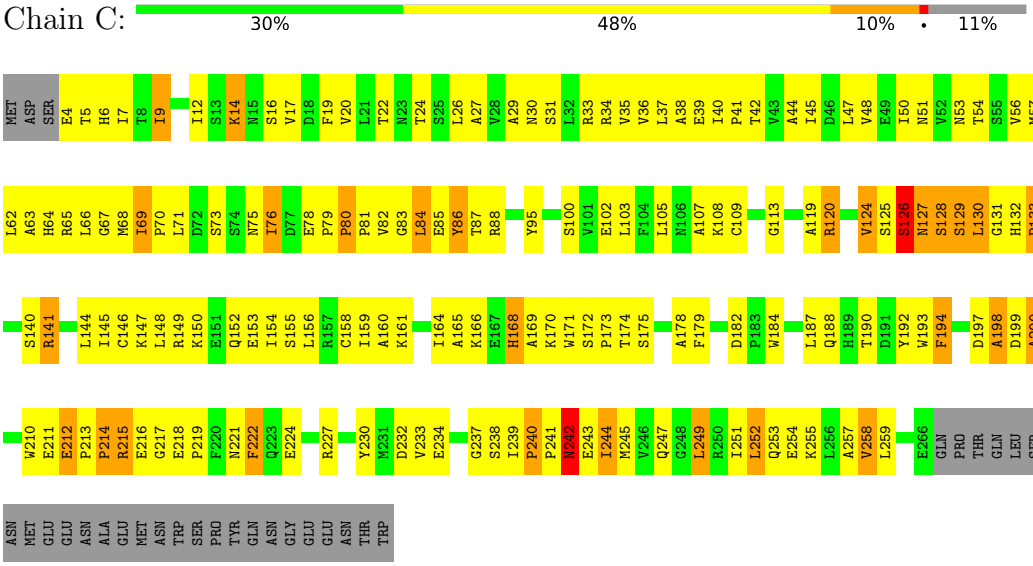





T326 T327 G328 V329 T330 R331 E332 K333 R334 L335 R336 Y337 A338 H339 D340 I341 L342 Q343 K344 E345 L346 L347 P348 H349 I350 T351 T352 M353 E354 G355 F356 R359 F362 F363 L364 G365 Y366 M367 I368 M371 L372 L376 E377 R378 R379 E380 P381 D382 D383 R384 D385 H386 F387 G388 K389 K390
R391 L394 A396 G396 P397 L398 L399 A400 S401 L402 F403 R404 M405 L406 F407 R408 R412 D413 V414 Y415 K416 Y417 M418 Q419 K420 C421 V422 E423 T424 N425 R426 E427 F428 N429 L430 T431 L432 A433 V434 K435 S436 N437 I438 I439 T440 N441 G442 L443 S446 L447 A448 T449 G450 N451 W452 G453 D454
Q455 K456 R457 S458 M459 V460 N461 R462 V463 G464 V465 S466 Q467 V468 L469 N470 R471 Y472 T473 F474 A475 S476 T477 L478 S479 H480 L481 R482 R483 T484 N485 T486 P487 I488 G489 R490 D491 G492 K493 L494 A495 K496 P497 R498 Q499 L500 H501 N502 T503 H504 W505 G506 M507 V508 C509 P510 A511 E512 T513 P514
Q517 A518 L521 V522 K523 N524 L525 S526 S529 Y530 V531 S532 V533 G534 S537 A538 P539 I540 I541 L544 L549 E550 T551 L552 E553 D554 Y555 N556 P557 S560 P561 N562 A563 T564 K565 V566 F567 V568 N569 G570 V571 W572 L573 G574 V575 H576 R577 D578 P579 A580 H581 L582
T585 L586 R587 S588 L589 R590 R591 R592 L593 D594 I595 I601 V602 R603 D604 I605 R606 E607 K608 E609 L610 R611 L612 F613 T614 D615 A616 G617 R618 I619 C620 R621 P622 L623 F624 I625 V626 D627 H628 N629 P630 M631 S632 E633 E637 L638 G639 I640 R641 K642 E643 H644 I645 Q646 Q647 D651
R654 Y655 D656 I657 D658 P659 E660 Q661 R662 F663 G664 W665 T666 A667 L668 V669 S670 S671 G672 L673 I674 E675 Y676 L677 D678 A679 E680 E681 E682 V685 M686 I687 A688 M689 S690 L694 E695 A696 S697 N629 P630 M631 S632 E633 E637 L638 G639 I640 R641 K642 E643 H644 I645 Q646 Q647 D651
K717 P718 A719 P720 N721 P722 H723 V724 H725 A726 W727 T728 H729 C730 E731 I732 H733 P734 A735 M736 I737 L738 G739 I740 L741 I744 I745 P746 F747 P748 D749 H750 N751 Q752 S753 P754 R755 N756 T757 Y758 Q759 S760 G763 K764 M767 L771 T772 N773 Y774 Q775 V776 R777 M778 D779 T780 M781
A782 N783 I784 L785 Y786 Q789 K790 P791 L792 A793 T794 T795 R796 S797 M798 E799 Y800 L801 K802 F803 L806 P807 A808 G809 Q810 N811 A812 I813 V814 A815 I816 L817 C818 Y819 S820 G821 Y822 N823 Q824 E825 D826 S827 I828 I829 M830 N831 Q832 A833 S834 I835 G838 L839 F840 R841 S842 I843 F844
Y845 R846 T847 D850 Q851 E852 K853 K854 I855 G856 M857 M860 E861 E862 F863 E864 R865 P866 V867 R868 S869 T870 T871 L872 R873 M874 K875 H876 G877 T878 Y879 D884 D885 G886 L887 I888 A889 P890 G891 T892 R893 V894 S895 G896 I899 I900 I901 G902 K903 T904 A905 P906 I907 PRO LEU ASP
HIS GLU GLU LEU GLY GLN ARG THR Q919 L920 H921 A922 K923 R924 D925 V926 S927 T928 P929 L930 R931 V938 D939 Q940 V941 M942 V943 T944 T945 N946 Q947 E948 G949 L950 K951 F952 V953 K954 V955 R956 M957 T960 R961 I962 P963 Q964 I965 G966 D967 K968 F969 A970 S971 R972 H973 G974 Q975 K976
G977 T978 I979 G980 M981 T982 Y983 R984 H985 E986 D987 W988 P989 F990 S991 A992 Q993 G994 I995 V996 I999 I1000 I1001 N1002 P1003 H1004 A1005 I1006 P1007 S1008 R1009 M1010 T1011 V1012 A1013 H1014 L1015 V1016 E1017 S1021 K1022 V1023 S1024 A1025 L1026 S1027 G1028 F1029 E1030 G1031 D1032 A1033 T1034 P1035 F1036 T1037 D1038 V1039
T1040 A1043 V1044 S1045 K1046 L1047 L1048 R1049 S1050 H1051 Q1054 S1055 R1056 G1057 F1058 E1059 V1060 M1061 Y1062 H1063 G1064 H1065 T1066 G1067 R1068 K1069 L1070 V1071 A1072 Q1073 V1074 F1075 L1076 G1077 P1078 T1079 Y1080 Y1081 Q1082 R1083 L1084 K1085 H1086 L1087 V1088 D1089 D1090 K1091 I1092 H1093 A1094 R1095 A1096 R1097 G1098 Q1101 I1102
L1103 T1104 R1105 Q1106 P1107 V1108 E1109 G1110 R1111 S1112 R1113 D1114 G1115 G1116 L1117 R1118 E1121 M1122 E1123 R1124 D1125 C1126 Q1127 I1128 S1129 H1130 G1131 C1132 S1133 S1134 V1135 L1136 R1137 E1138 R1139 L1140 F1141 D1142 C1143 S1144 D1145 A1146 Y1147 R1148 V1149 I1150 V1151 C1152 D1153 I1154 C1155 G1156 L1157 I1158 A1159 I1160 A1161 S1162 Y1163
K1164 Y1168 E1169 C1170 R1171 S1172 N1175 R1176 T1177 R1178 F1179 S1180 Q1181 V1182 Y1183 L1184 P1185 Y1186 A1187 A1188 K1189 L1190 L1191 F1192 Q1193 E1194 L1195 M1196 S1197 M1198 N1199 I1200 A1201 P1202 R1203 L1204 F1205 T1206 K1207 N1208 H1209 K1210

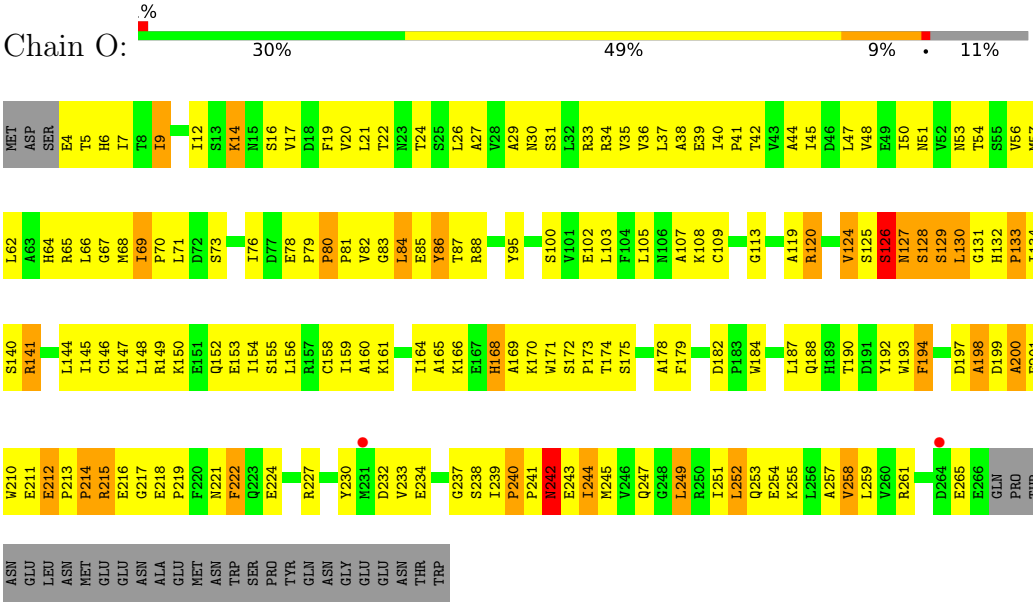
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C:



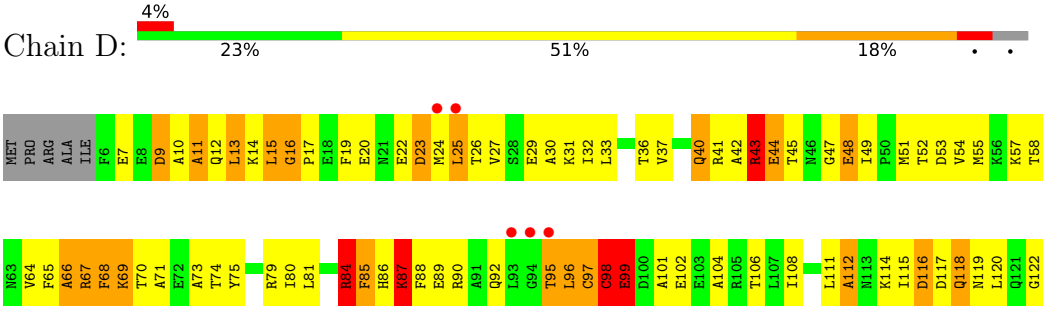
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain O:

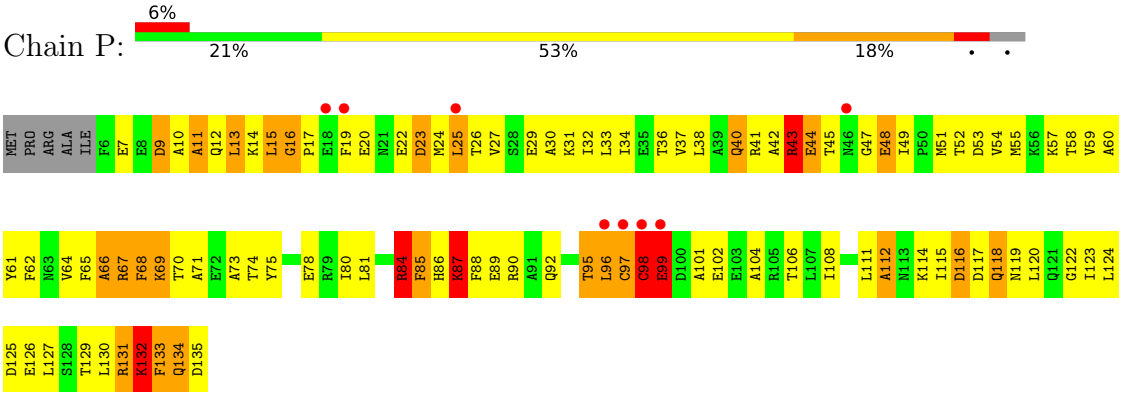


• Molecule 4: DNA-directed RNA polymerase II subunit rpb4

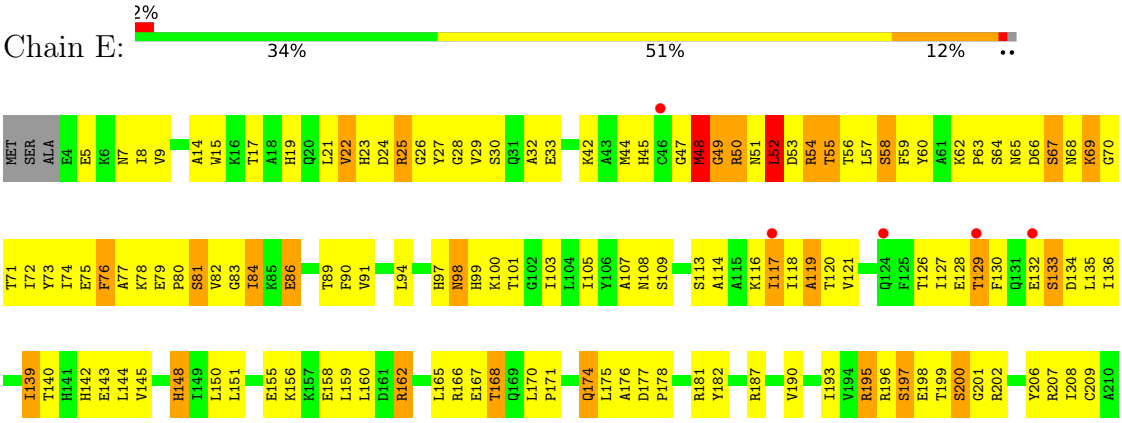
Chain D:



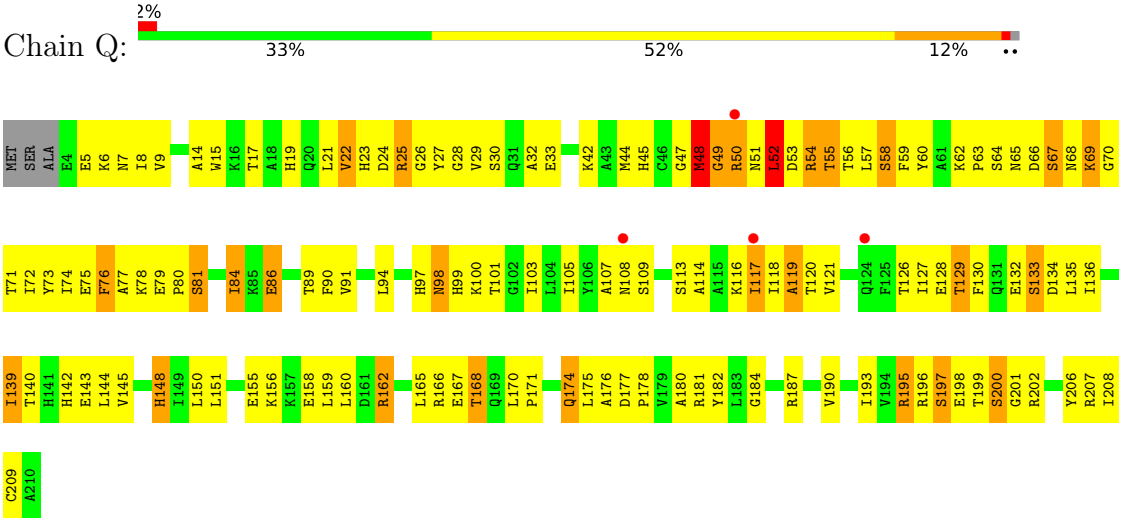
• Molecule 4: DNA-directed RNA polymerase II subunit rpb4



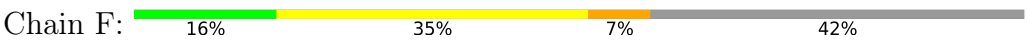
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

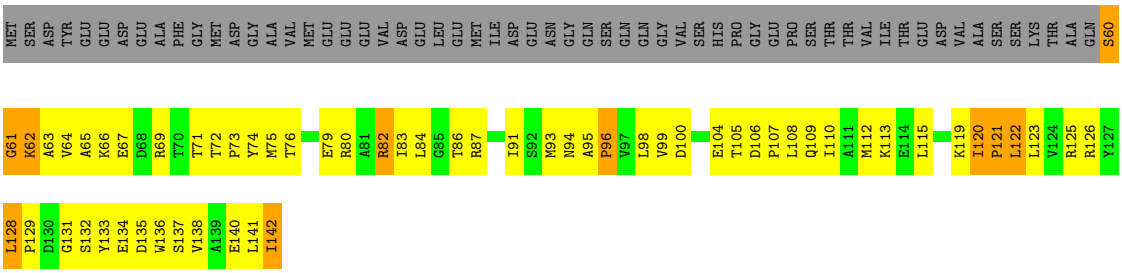


• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

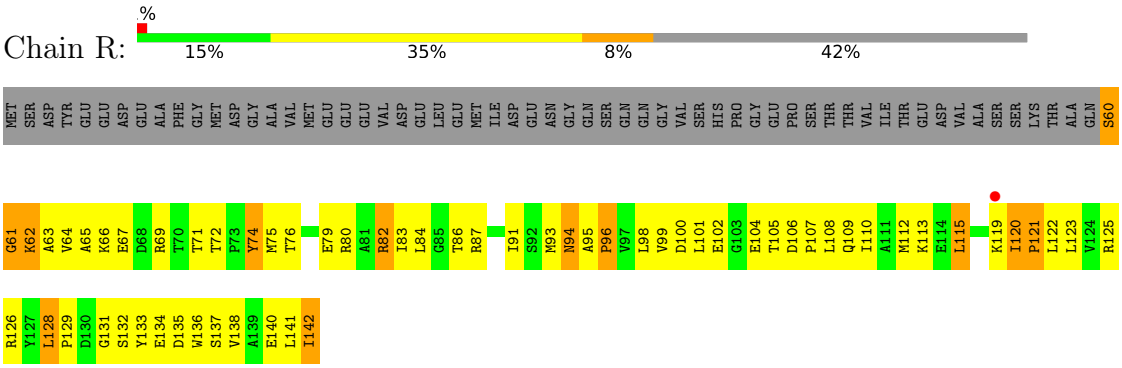


• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

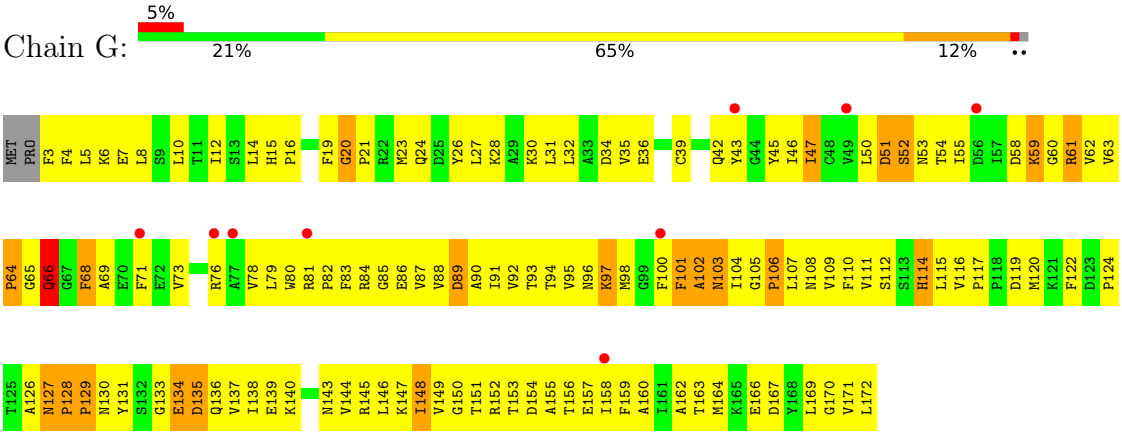




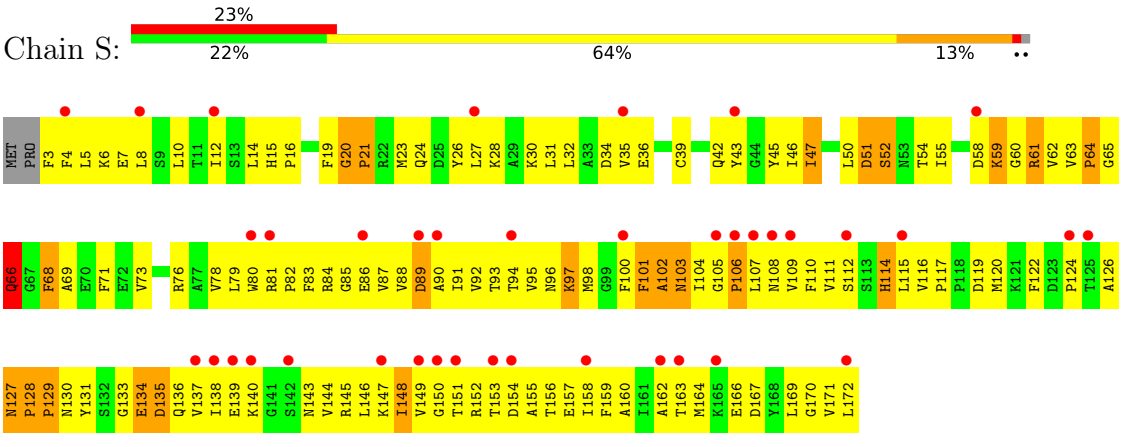
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



• Molecule 7: DNA-directed RNA polymerase II subunit rpb7

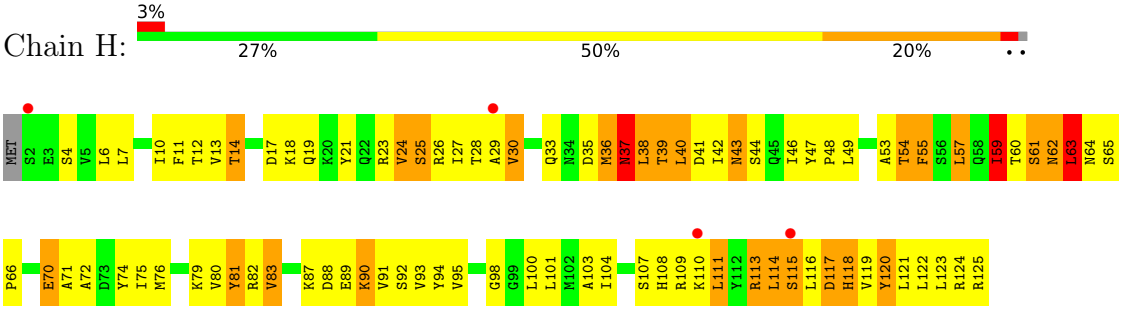


• Molecule 7: DNA-directed RNA polymerase II subunit rpb7



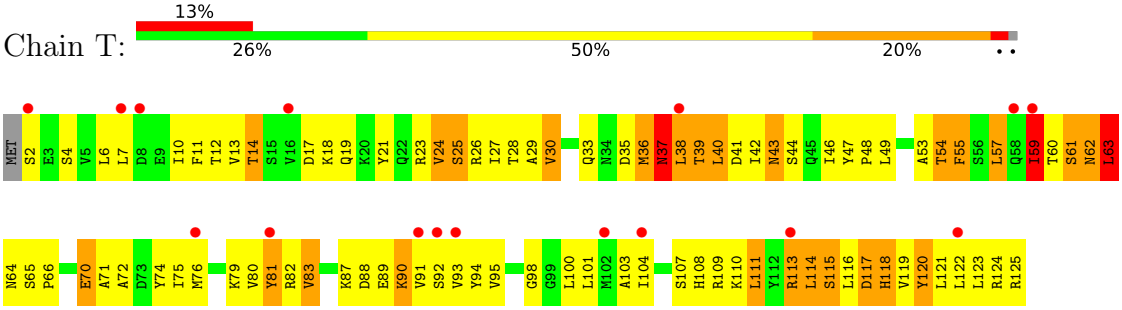
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



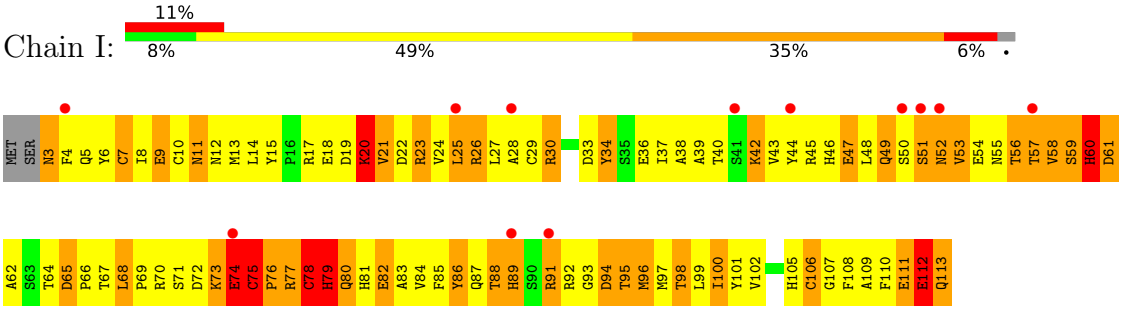
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain T:



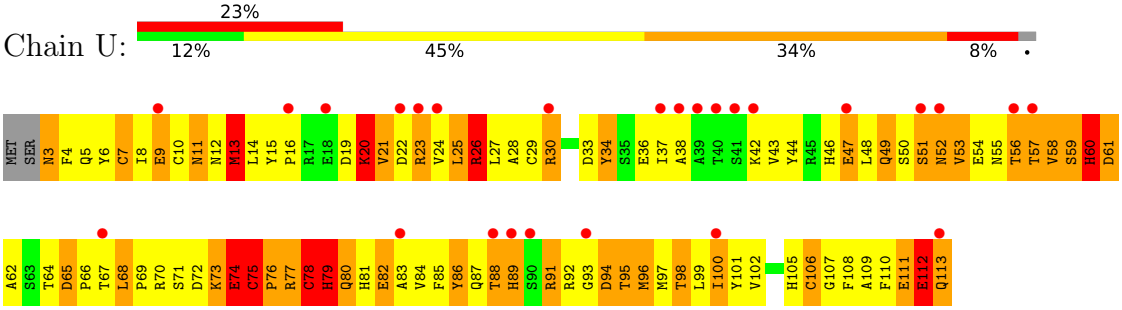
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:



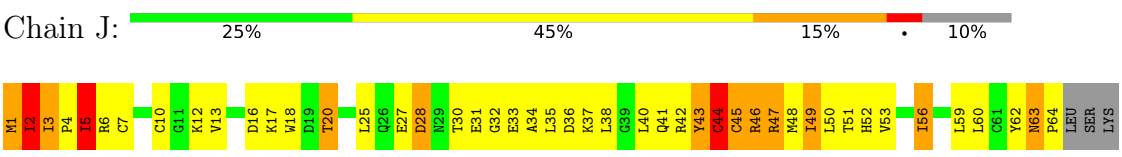
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain U:



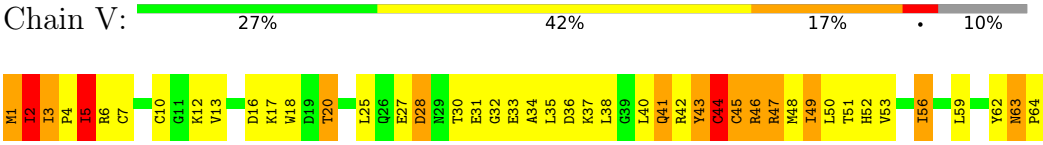
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



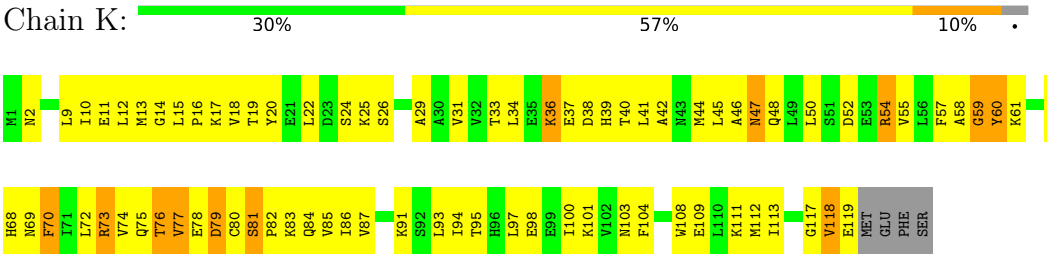
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain V:



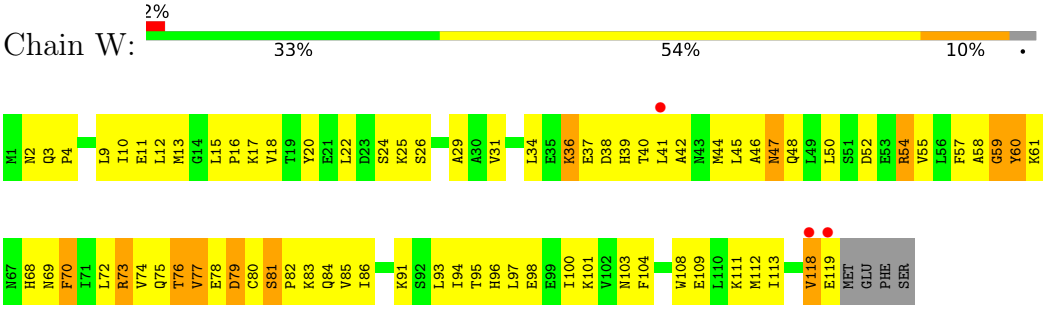
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain K:



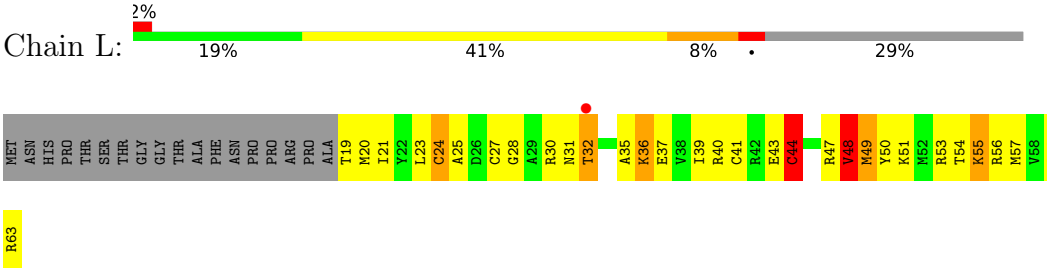
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain W:



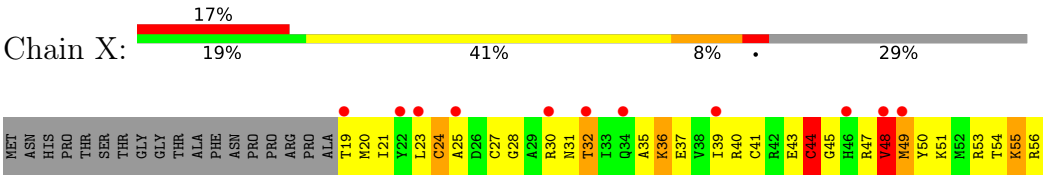
• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

Chain L:



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

Chain X:



R63

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	163.03Å 202.68Å 391.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.70 – 3.65 48.69 – 3.65	Depositor EDS
% Data completeness (in resolution range)	92.8 (48.70-3.65) 96.2 (48.69-3.65)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 3.67Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.297 , 0.321 0.292 , 0.296	Depositor DCC
R_{free} test set	6932 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	125.6	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 139.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	62870	wwPDB-VP
Average B, all atoms (Å^2)	203.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.41	0/12026	0.57	1/16260 (0.0%)
1	M	0.42	0/11887	0.56	1/16069 (0.0%)
2	B	0.44	0/9360	0.60	5/12643 (0.0%)
2	N	0.43	0/9360	0.59	5/12643 (0.0%)
3	C	0.44	0/2135	0.60	0/2904
3	O	0.44	0/2135	0.59	0/2904
4	D	0.24	0/1049	0.38	0/1412
4	P	0.24	0/1049	0.38	0/1412
5	E	0.38	0/1695	0.60	0/2287
5	Q	0.39	0/1695	0.60	0/2287
6	F	0.50	0/666	0.67	0/901
6	R	0.50	0/666	0.67	0/901
7	G	0.26	0/1361	0.57	3/1847 (0.2%)
7	S	0.26	0/1361	0.57	3/1847 (0.2%)
8	H	0.42	0/1010	0.65	0/1363
8	T	0.42	0/1010	0.65	0/1363
9	I	0.22	0/921	0.37	0/1246
9	U	0.24	0/921	0.37	0/1246
10	J	0.57	0/526	0.77	0/709
10	V	0.57	0/526	0.76	0/709
11	K	0.47	0/972	0.61	0/1317
11	W	0.47	0/972	0.61	0/1317
12	L	0.36	0/371	0.57	0/491
12	X	0.37	0/371	0.57	0/491
All	All	0.41	0/64045	0.58	18/86569 (0.0%)

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	992	ALA	N-CA-C	-5.71	95.59	111.00
2	N	992	ALA	N-CA-C	-5.69	95.65	111.00
7	G	167	ASP	CB-CG-OD2	5.38	123.14	118.30
7	S	167	ASP	CB-CG-OD2	5.35	123.12	118.30
2	N	711	ASP	CB-CG-OD2	5.30	123.07	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11802	0	11784	3283	0
1	M	11666	0	11647	3212	0
2	B	9180	0	9163	1630	0
2	N	9180	0	9164	1653	0
3	C	2088	0	2045	267	0
3	O	2088	0	2045	269	0
4	D	1036	0	1025	349	0
4	P	1036	0	1025	318	0
5	E	1663	0	1684	205	0
5	Q	1663	0	1684	209	0
6	F	656	0	679	75	0
6	R	656	0	679	82	0
7	G	1330	0	1329	424	0
7	S	1330	0	1329	425	0
8	H	996	0	1006	168	0
8	T	996	0	1006	178	0
9	I	902	0	840	282	0
9	U	902	0	839	268	0
10	J	518	0	529	90	0
10	V	518	0	529	90	0
11	K	955	0	968	120	0
11	W	955	0	968	113	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	L	368	0	380	38	0
12	X	368	0	380	38	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
13	M	2	0	0	0	0
13	N	1	0	0	0	0
13	O	1	0	0	0	0
13	U	2	0	0	0	0
13	V	1	0	0	0	0
13	X	1	0	0	0	0
14	A	1	0	0	0	0
14	M	1	0	0	0	0
All	All	62870	0	62727	12749	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 102.

The worst 5 of 12749 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)
1:M:1161:PRO:HG2	1:M:1190:LYS:CG	1.29	1.62
1:A:1161:PRO:HG2	1:A:1190:LYS:CG	1.29	1.58
1:M:1161:PRO:CG	1:M:1190:LYS:HG2	1.33	1.57
1:M:1091:GLY:HA3	1:M:1092:VAL:CG1	1.35	1.56
1:M:267:ASP:CB	1:M:268:LEU:HB2	1.35	1.54

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	1494/1752 (85%)	939 (63%)	280 (19%)	275 (18%)	0	1
1	M	1472/1752 (84%)	932 (63%)	281 (19%)	259 (18%)	0	2
2	B	1142/1210 (94%)	735 (64%)	245 (22%)	162 (14%)	0	3
2	N	1142/1210 (94%)	738 (65%)	243 (21%)	161 (14%)	0	3
3	C	261/297 (88%)	178 (68%)	56 (22%)	27 (10%)	0	7
3	O	261/297 (88%)	178 (68%)	56 (22%)	27 (10%)	0	7
4	D	128/135 (95%)	86 (67%)	20 (16%)	22 (17%)	0	2
4	P	128/135 (95%)	86 (67%)	20 (16%)	22 (17%)	0	2
5	E	205/210 (98%)	137 (67%)	43 (21%)	25 (12%)	0	5
5	Q	205/210 (98%)	137 (67%)	43 (21%)	25 (12%)	0	5
6	F	81/142 (57%)	58 (72%)	15 (18%)	8 (10%)	0	8
6	R	81/142 (57%)	57 (70%)	16 (20%)	8 (10%)	0	8
7	G	168/172 (98%)	128 (76%)	24 (14%)	16 (10%)	0	8
7	S	168/172 (98%)	128 (76%)	24 (14%)	16 (10%)	0	8
8	H	122/125 (98%)	73 (60%)	27 (22%)	22 (18%)	0	1
8	T	122/125 (98%)	73 (60%)	27 (22%)	22 (18%)	0	1
9	I	109/113 (96%)	46 (42%)	27 (25%)	36 (33%)	0	0
9	U	109/113 (96%)	45 (41%)	25 (23%)	39 (36%)	0	0
10	J	62/71 (87%)	41 (66%)	15 (24%)	6 (10%)	0	8
10	V	62/71 (87%)	41 (66%)	15 (24%)	6 (10%)	0	8
11	K	117/123 (95%)	80 (68%)	24 (20%)	13 (11%)	0	5
11	W	117/123 (95%)	80 (68%)	24 (20%)	13 (11%)	0	5
12	L	43/63 (68%)	23 (54%)	12 (28%)	8 (19%)	0	1
12	X	43/63 (68%)	23 (54%)	12 (28%)	8 (19%)	0	1
All	All	7842/8826 (89%)	5042 (64%)	1574 (20%)	1226 (16%)	0	3

5 of 1226 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	SER
1	A	12	PRO
1	A	37	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	42	THR
1	A	50	PRO

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	1301/1536 (85%)	1113 (86%)	188 (14%)	3 18
1	M	1286/1536 (84%)	1097 (85%)	189 (15%)	3 18
2	B	1012/1064 (95%)	914 (90%)	98 (10%)	8 33
2	N	1012/1064 (95%)	914 (90%)	98 (10%)	8 33
3	C	236/267 (88%)	220 (93%)	16 (7%)	16 47
3	O	236/267 (88%)	220 (93%)	16 (7%)	16 47
4	D	111/115 (96%)	94 (85%)	17 (15%)	2 17
4	P	111/115 (96%)	94 (85%)	17 (15%)	2 17
5	E	182/184 (99%)	169 (93%)	13 (7%)	14 45
5	Q	182/184 (99%)	169 (93%)	13 (7%)	14 45
6	F	71/121 (59%)	64 (90%)	7 (10%)	8 32
6	R	71/121 (59%)	64 (90%)	7 (10%)	8 32
7	G	146/148 (99%)	139 (95%)	7 (5%)	25 56
7	S	146/148 (99%)	139 (95%)	7 (5%)	25 56
8	H	113/114 (99%)	99 (88%)	14 (12%)	4 23
8	T	113/114 (99%)	99 (88%)	14 (12%)	4 23
9	I	103/105 (98%)	80 (78%)	23 (22%)	1 6
9	U	103/105 (98%)	80 (78%)	23 (22%)	1 6
10	J	59/66 (89%)	46 (78%)	13 (22%)	1 6
10	V	59/66 (89%)	46 (78%)	13 (22%)	1 6
11	K	109/113 (96%)	104 (95%)	5 (5%)	27 57
11	W	109/113 (96%)	103 (94%)	6 (6%)	21 53

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
12	L	39/53 (74%)	34 (87%)	5 (13%)	4 22
12	X	39/53 (74%)	34 (87%)	5 (13%)	4 22
All	All	6949/7772 (89%)	6135 (88%)	814 (12%)	5 26

5 of 814 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	326	ARG
1	M	1278	SER
10	V	46	ARG
1	M	430	ILE
1	M	322	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 225 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	82	HIS
11	W	39	HIS
1	M	898	GLN
10	V	52	HIS
5	Q	19	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 18 ligands modelled in this entry, 18 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, 95^{th} 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	1496/1752 (85%)	0.16	84 (5%) 24 16	51, 150, 385, 628	0
1	M	1476/1752 (84%)	0.53	174 (11%) 4 3	83, 199, 449, 672	0
2	B	1150/1210 (95%)	0.05	27 (2%) 60 46	69, 142, 314, 580	0
2	N	1150/1210 (95%)	0.55	118 (10%) 6 4	85, 222, 419, 618	0
3	C	263/297 (88%)	-0.18	0 100 100	80, 126, 252, 477	0
3	O	263/297 (88%)	0.08	2 (0%) 86 77	125, 190, 350, 511	0
4	D	130/135 (96%)	0.32	5 (3%) 40 28	128, 240, 360, 596	0
4	P	130/135 (96%)	0.55	8 (6%) 20 13	152, 287, 414, 545	0
5	E	207/210 (98%)	-0.15	5 (2%) 59 45	70, 164, 293, 516	0
5	Q	207/210 (98%)	-0.05	4 (1%) 66 53	119, 181, 373, 501	0
6	F	83/142 (58%)	-0.49	0 100 100	70, 97, 170, 232	0
6	R	83/142 (58%)	-0.35	1 (1%) 79 68	111, 126, 200, 335	0
7	G	170/172 (98%)	0.35	9 (5%) 26 18	92, 199, 323, 514	0
7	S	170/172 (98%)	1.07	39 (22%) 0 0	122, 240, 388, 494	0
8	H	124/125 (99%)	0.14	4 (3%) 47 34	82, 138, 298, 408	0
8	T	124/125 (99%)	0.60	16 (12%) 3 3	115, 197, 344, 429	0
9	I	111/113 (98%)	0.50	12 (10%) 5 4	98, 239, 367, 549	0
9	U	111/113 (98%)	1.63	26 (23%) 0 0	151, 330, 482, 585	0
10	J	64/71 (90%)	-0.20	0 100 100	85, 108, 220, 304	0
10	V	64/71 (90%)	0.04	0 100 100	137, 203, 326, 373	0
11	K	119/123 (96%)	-0.34	0 100 100	60, 126, 221, 319	0
11	W	119/123 (96%)	0.15	3 (2%) 57 43	76, 168, 295, 470	0
12	L	45/63 (71%)	-0.02	1 (2%) 62 48	93, 177, 299, 481	0
12	X	45/63 (71%)	1.00	11 (24%) 0 0	153, 277, 425, 580	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	7904/8826 (89%)	0.29	549 (6%)	16	11	51, 181, 397, 672	0

The worst 5 of 549 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
9	U	41	SER	21.8
9	U	40	THR	16.2
9	U	39	ALA	14.8
11	W	119	GLU	14.5
1	A	107	CYS	12.1

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
14	MG	M	2458	1/1	0.27	0.39	197,197,197,197	0
14	MG	A	2458	1/1	0.54	1.81	109,109,109,109	0
13	ZN	M	2456	1/1	0.63	0.26	432,432,432,432	0
13	ZN	A	2456	1/1	0.68	0.06	245,245,245,245	0
13	ZN	O	1269	1/1	0.70	0.11	189,189,189,189	0
13	ZN	N	2225	1/1	0.80	0.22	180,180,180,180	0
13	ZN	U	1122	1/1	0.80	0.06	240,240,240,240	0
13	ZN	M	2457	1/1	0.83	0.14	172,172,172,172	0
13	ZN	X	1071	1/1	0.83	0.05	195,195,195,195	0
13	ZN	I	1122	1/1	0.83	0.10	126,126,126,126	0
13	ZN	I	1121	1/1	0.83	0.06	160,160,160,160	0
13	ZN	L	1071	1/1	0.87	0.09	123,123,123,123	0
13	ZN	U	1121	1/1	0.90	0.06	284,284,284,284	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
13	ZN	C	1269	1/1	0.91	0.07	122,122,122,122	0
13	ZN	B	2225	1/1	0.94	0.17	149,149,149,149	0
13	ZN	A	2457	1/1	0.94	0.20	122,122,122,122	0
13	ZN	V	1066	1/1	0.96	0.20	155,155,155,155	0
13	ZN	J	1066	1/1	0.97	0.22	87,87,87,87	0

6.5 Other polymers (i)

There are no such residues in this entry.