wwPDB EM Validation Summary Report ⓘ

Nov 6, 2024 – 05:57 AM EST

PDB ID	:	7ML2
EMDB ID	:	EMD-23906
Title	:	RNA polymerase II pre-initiation complex (PIC3)
Authors	:	Yang, C.; Fujiwara, R.; Kim, H.J.; Gorbea Colon, J.J.; Steimle, S.; Garcia, B.A.; Murakami, K.
Deposited on	:	2021-04-27
Resolution	:	3.40 Å(reported)
Based on initial model	:	5OQJ

This is a wwPDB EM 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/EMValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

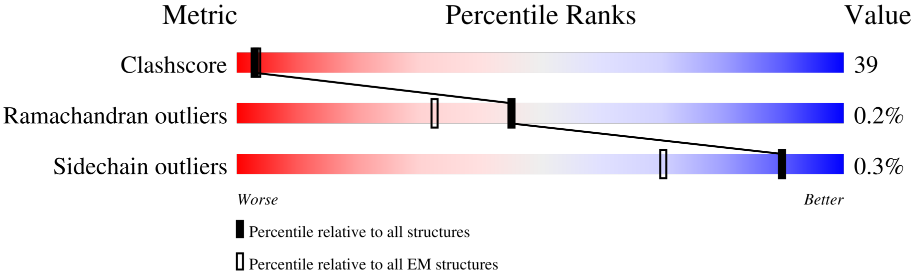
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

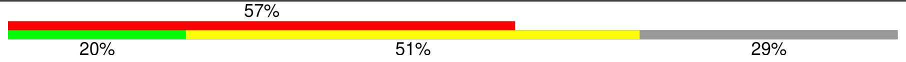
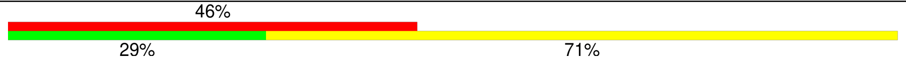
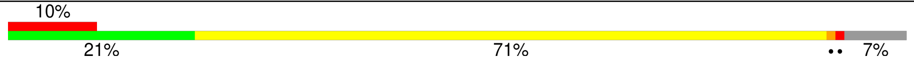
The reported resolution of this entry is 3.40 Å.

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.



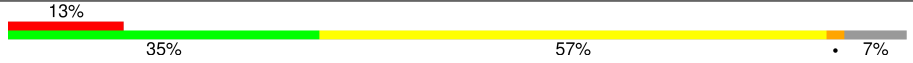
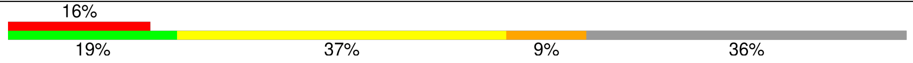
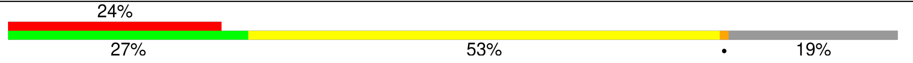
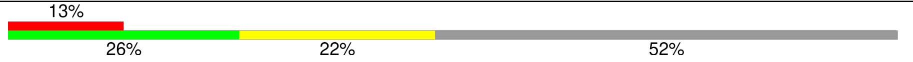
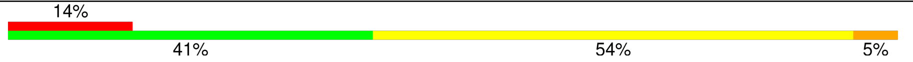
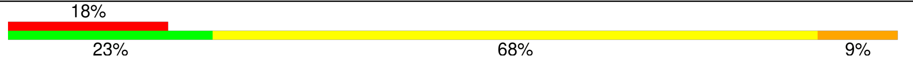
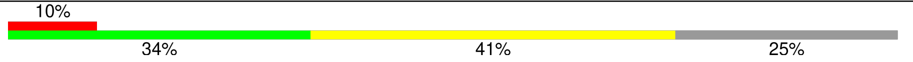
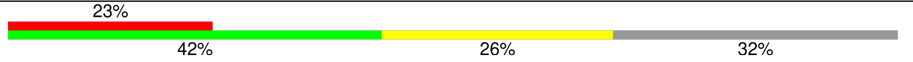
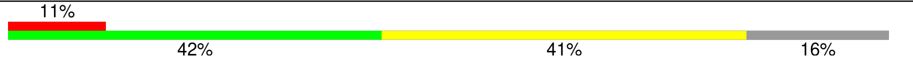
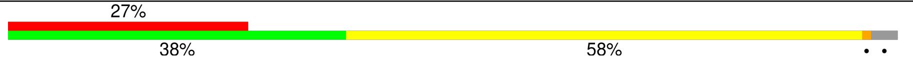
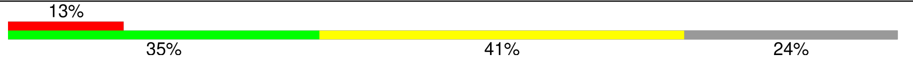
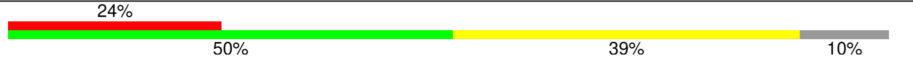
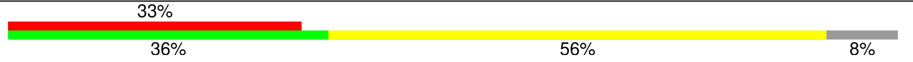
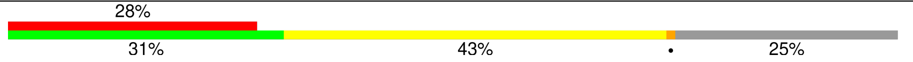
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1733	
2	B	1224	
3	C	318	
4	D	221	
5	E	215	
6	F	155	
7	G	171	
8	H	146	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	M	345	
14	Q	735	
15	R	400	
16	U	286	
17	V	122	
18	W	482	
19	X	328	
20	T	56	
21	N	56	
22	O	240	
23	1	542	
24	4	338	
25	0	778	
26	6	461	
27	2	513	
28	5	72	
29	7	843	
30	3	321	

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
33	SF4	0	801	-	-	X	-

2 Entry composition ⓘ

There are 33 unique types of molecules in this entry. The entry contains 64550 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1398	10997	6931	1927	2078	61	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1152	9178	5807	1608	1708	55	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	262	2061	1299	343	406	13	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	157	1253	779	220	252	2	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	213	1744	1107	308	318	11	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	83	670	428	114	125	3	0	0

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	136	Total 1089	C 686	N 184	O 215	S 4	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	116	Total 944	C 581	N 172	O 181	S 10	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerases II subunit RPABC5.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	112	Total 904	C 580	N 154	O 168	S 2	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerases II subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	45	Total 358	C 221	N 71	O 62	S 4	0	0

- Molecule 13 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	279	Total 2175	C 1382	N 373	O 403	S 17	0	0

- Molecule 14 is a protein called Transcription initiation factor IIF subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	Q	148	1144	733	195	212	4	0	0

- Molecule 15 is a protein called Transcription initiation factor IIF subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	R	190	1303	812	238	246	7	0	0

- Molecule 16 is a protein called Transcription initiation factor IIA large subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	U	46	383	242	67	71	3	0	0

- Molecule 17 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	V	49	381	241	63	74	3	0	0

- Molecule 18 is a protein called Transcription initiation factor IIE subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	W	191	1469	932	254	277	6	0	0

- Molecule 19 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	X	156	984	608	180	192	4	0	0

- Molecule 20 is a DNA chain called template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
20	T	56	1140	550	188	346	56	0	0

- Molecule 21 is a DNA chain called non-template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	N	56	Total	C	N	O	P	0	0
			1156	552	222	326	56		

- Molecule 22 is a protein called BJ4_G0004860.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	O	180	Total	C	N	O	S	0	0
			1416	921	242	247	6		

- Molecule 23 is a protein called Tfb1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	1	367	Total	C	N	O	S	0	0
			2411	1536	438	430	7		

- Molecule 24 is a protein called Tfb4.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	4	284	Total	C	N	O	S	0	0
			2041	1310	343	376	12		

- Molecule 25 is a protein called General transcription and DNA repair factor IIH helicase subunit XPD.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	0	754	Total	C	N	O	S	0	0
			6108	3891	1032	1147	38		

- Molecule 26 is a protein called General transcription and DNA repair factor IIH.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	6	351	Total	C	N	O	S	0	0
			2527	1590	454	456	27		

- Molecule 27 is a protein called RNA polymerase II transcription factor B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	2	460	Total	C	N	O	S	0	0
			3011	1856	562	584	9		

- Molecule 28 is a protein called General transcription and DNA repair factor IIH subunit TFB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	5	66	Total	C	N	O	S	0	0
			498	314	89	93	2		

- Molecule 29 is a protein called General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	7	634	Total	C	N	O	S	0	0
			4447	2722	827	874	24		

- Molecule 30 is a protein called BJ4_G0050160.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	3	138	Total	C	N	O	S	0	0
			860	533	160	160	7		

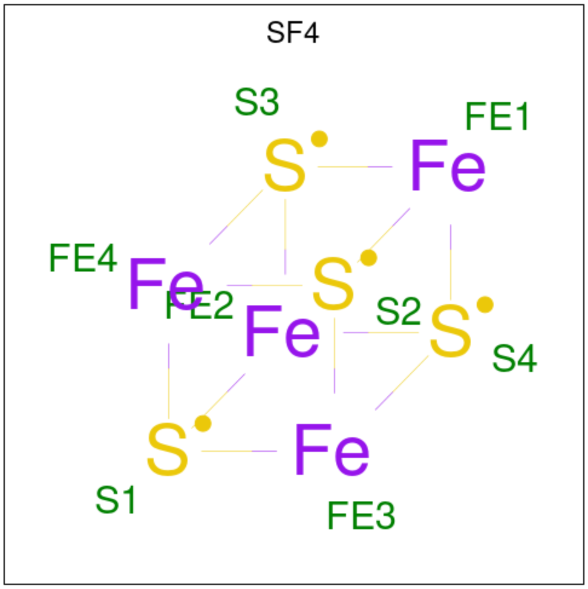
- Molecule 31 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
31	A	2	Total	Zn	0
			2	2	
31	B	1	Total	Zn	0
			1	1	
31	C	1	Total	Zn	0
			1	1	
31	I	2	Total	Zn	0
			2	2	
31	J	1	Total	Zn	0
			1	1	
31	L	1	Total	Zn	0
			1	1	
31	M	1	Total	Zn	0
			1	1	
31	W	1	Total	Zn	0
			1	1	
31	4	1	Total	Zn	0
			1	1	
31	6	4	Total	Zn	0
			4	4	
31	3	2	Total	Zn	0
			2	2	

- Molecule 32 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
32	A	1	1	1	0

- Molecule 33 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



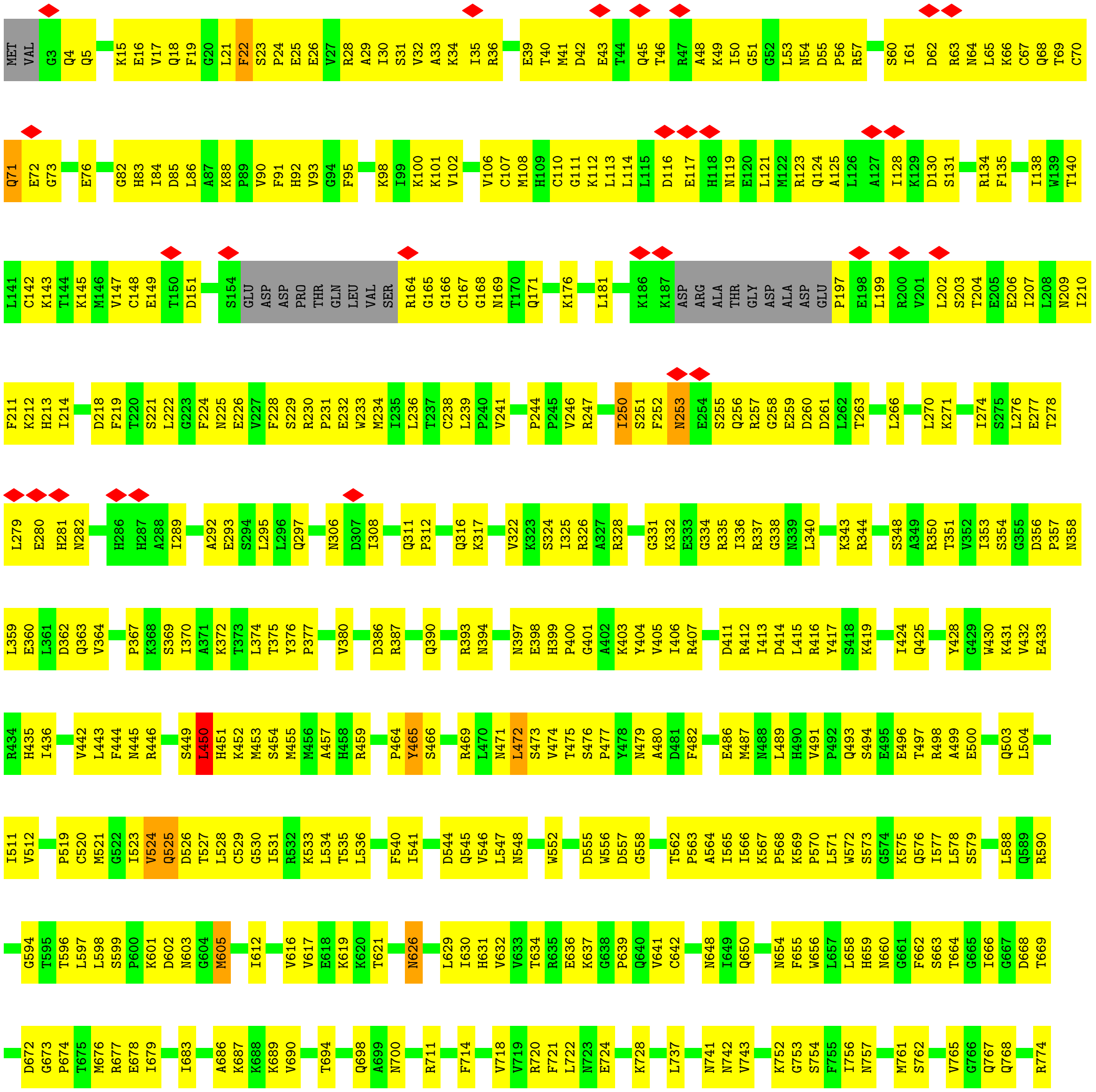
Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
33	0	1	8	4	4	0

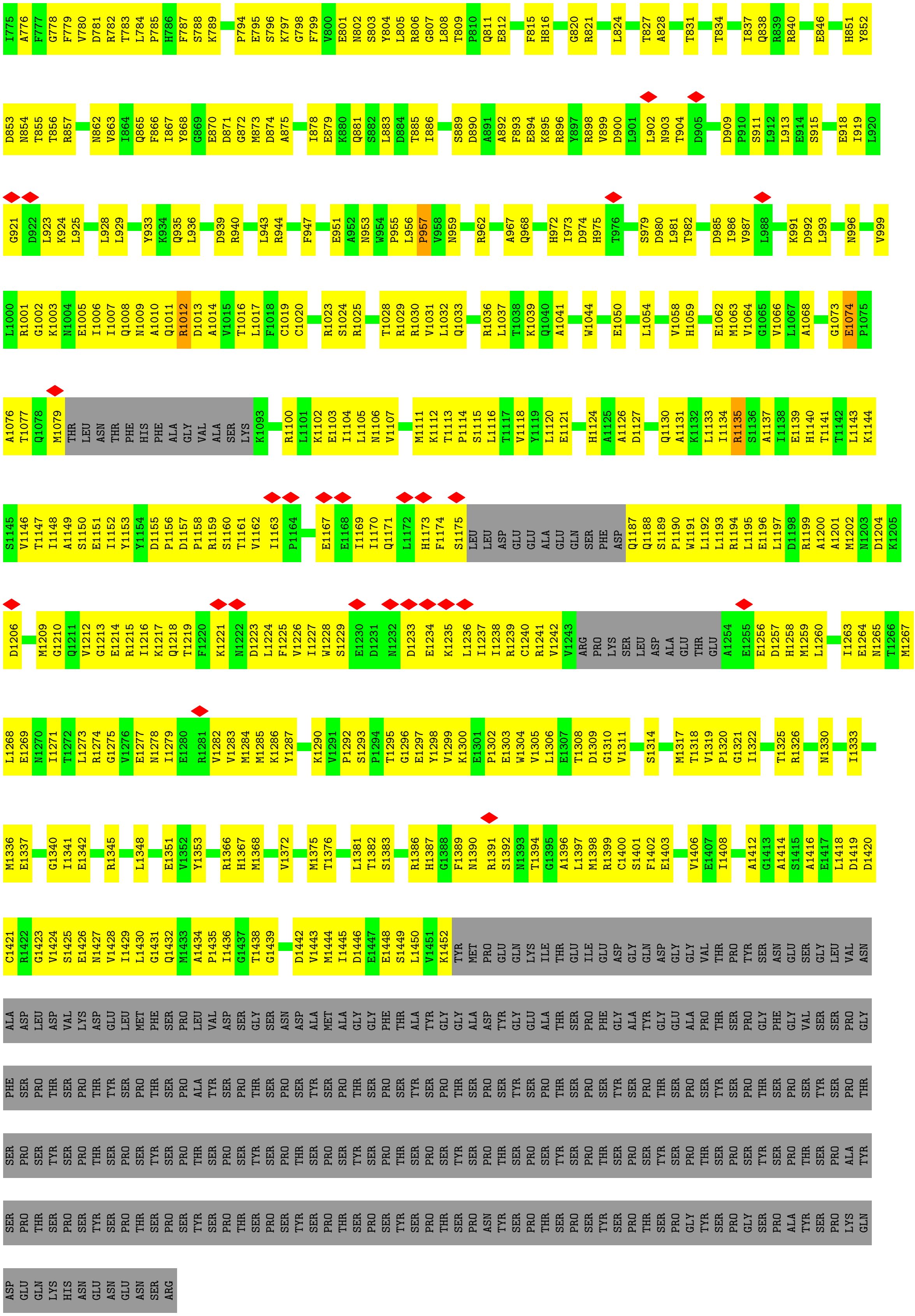
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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.

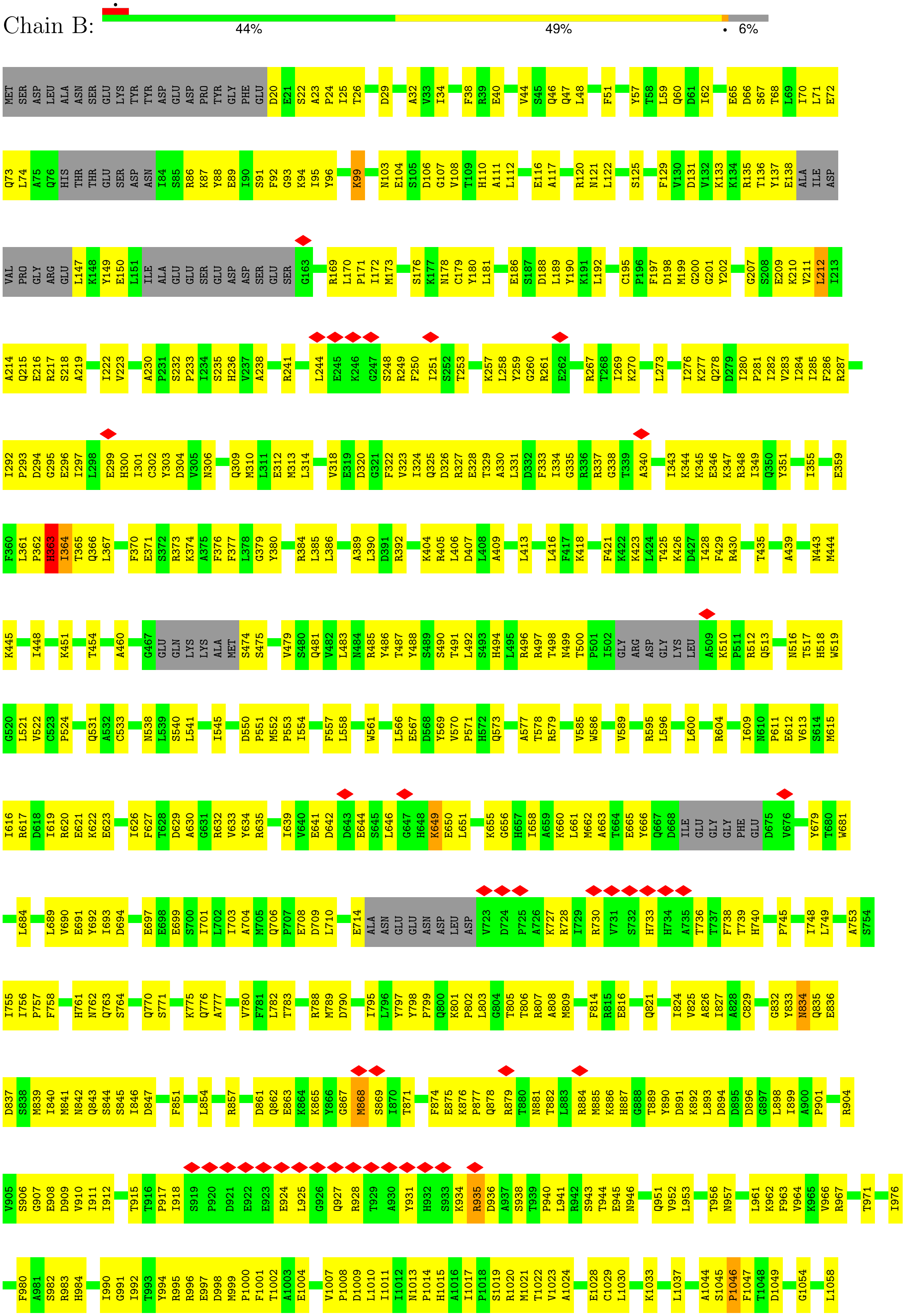
- Molecule 1: DNA-directed RNA polymerase subunit

Chain A:



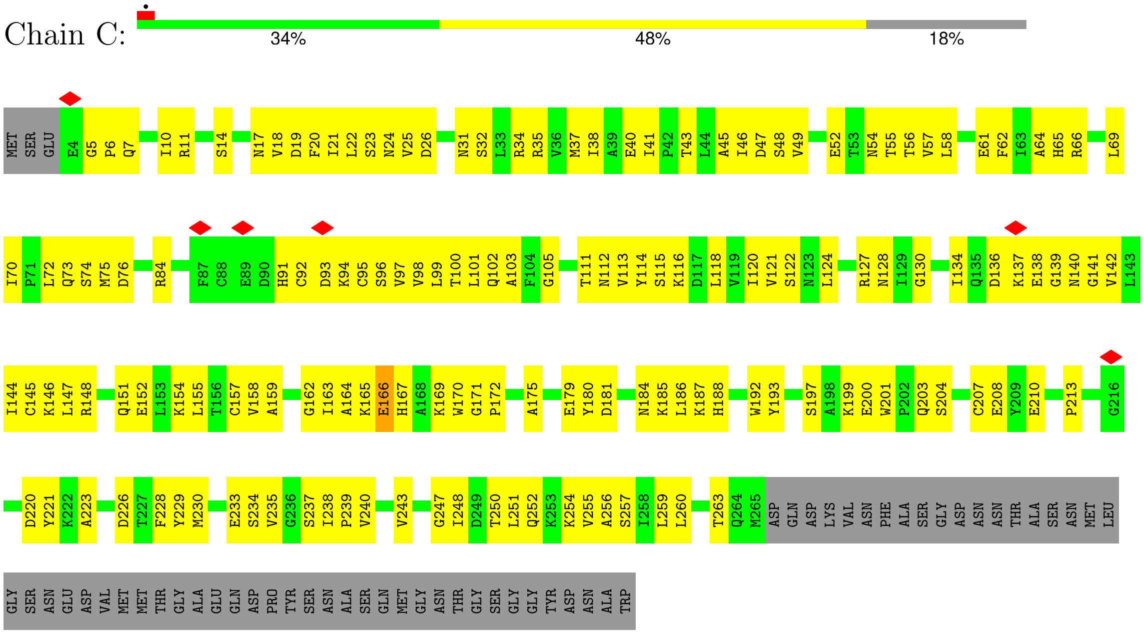


- Molecule 2: DNA-directed RNA polymerase subunit beta


Chain B:
44%
49%
6%
MET SER ASP LEU ALA ASN SER GLU LYS TYR TYR ASP GLU ASP PRO TYR GLY PHE GLU D20 E21 S22 A23 P24 I25 T26 D29 A32 V33 I34 F38 R39 E40 V44 S45 Q46 Q47 L48 F51 Y57 T58 L59 Q60 D61 I62 E65 D66 S67 T68 L69 I70 L71 E72
Q73 L74 A75 Q76 HIS THR THR GLU SER ASP ASN I84 S85 R86 K87 Y88 E89 I90 S91 F92 G93 K94 I95 Y96 K99 N103 E104 S105 D106 G107 V108 T109 H110 A111 L112 E116 A117 R120 N121 L122 S125 F129 V130 D131 V132 K133 K134 R135 T136 Y137 E138 ALA ILE ASP
VAL PRO GLY ARG GLU L147 K148 Y149 E150 L151 ILE ALA GLU GLU SER GLU ASP ASP SER GLU SER G163 R169 L170 P171 I172 M173 S176 K177 N178 C179 Y180 L181 E186 S187 D188 L189 Y190 K191 L192 C195 P196 F197 D198 M199 G200 G201 Y202 G207 S208 E209 K210 V211 L212 I213
A214 Q215 E216 R217 S218 A219 I222 V223 A230 P231 S232 P233 I234 S235 H236 V237 A238 R241 L244 E245 K246 G247 S248 R249 F250 I251 S252 T253 K257 L258 Y259 G260 R261 E262 R267 T268 I269 K270 L273 I276 K277 Q278 D279 I280 P281 I282 V283 I284 I285 F286 R287
I292 P293 D294 G295 E296 I297 L298 E299 H300 I301 C302 Y303 D304 V305 N306 Q309 M310 L311 E312 M313 L314 V318 E319 D320 G321 F322 V323 I324 Q325 D326 R327 E328 T329 A330 L331 D332 F333 I334 G335 R336 R337 G338 T339 A340 I343 K344 K345 E346 K347 R348 I349 Q350 Y351 I355 E359
F360 L361 P362 H363 I364 T365 Q366 L367 F370 E371 S372 R373 K374 A375 F376 F377 L378 G379 Y380 R384 L385 L386 A389 L390 D391 R392 K404 R405 L406 D407 L408 A409 L413 L416 F417 K418 F421 K422 K423 L424 T425 K426 D427 I428 F429 R430 T435 A439 N443 M444
K445 I448 K451 T454 A460 G467 GLU GLN LYS LYS ALA MET S474 S475 V479 S480 Q481 V482 L483 N484 R485 Y486 T487 Y488 S489 S490 T491 L492 S493 H494 L495 R496 R497 T498 N499 T500 P501 I502 GLY ARG ASP GLY LYS LEU A509 K510 P511 R512 Q513 N516 T517 H518 W519
G520 L521 V522 C523 P524 Q531 A532 C533 N538 L539 S540 L541 I545 D550 P551 M552 P553 I554 F557 L558 W561 L566 E567 D568 Y569 V570 P571 H572 Q573 A577 T578 R579 V585 W586 V589 R595 L596 L600 R604 I609 N610 P611 E612 V613 S614 M615
I616 R617 D618 I619 R620 E621 K622 E623 I626 F627 T628 D629 A630 G631 R632 V633 Y634 R635 I639 V640 E641 D642 D643 E644 S645 L646 G647 H648 K649 E650 L651 K655 G656 H657 I658 A659 K660 L661 M662 A663 T664 E665 Y666 Q667 D668 ILE GLU GLY GLY PHE GLU D675 V676 Y679 T680 W681
L684 L689 V690 E691 Y692 I693 D694 E697 E698 E699 S700 I701 L702 I703 A704 M705 Q706 P707 E708 D709 L710 E714 ALA ASN GLU GLU ASN ASP LEU ASP V723 D724 P725 A726 K727 R728 I729 R730 V731 S732 H733 H734 A735 T736 T737 F738 T739 H740 P745 I748 L749 A753 S754
I755 I756 P757 F758 H761 N762 Q763 S764 Q770 S771 K775 Q776 A777 V780 F781 L782 T783 R788 M789 D790 I795 L796 Y797 Y798 P799 Q800 K801 P802 L803 G804 T805 T806 R807 A808 M809 F814 R815 E816 Q821 I824 V825 A826 I827 A828 C829 G832 Y833 N834 Q835 E836
D837 S838 M839 I840 M841 N842 Q843 S844 S845 I846 D847 F851 L854 R857 D861 Q862 E863 K864 K865 Y866 G867 M868 S869 I870 T871 F874 E875 K876 P877 Q878 R879 T880 N881 T882 L883 R884 M885 K886 H887 G888 T889 Y890 D891 K892 L893 D894 D895 D896 G897 L898 I899 A900 P901 R904
V905 S906 G907 E908 D909 V910 I911 I912 T915 T916 P917 I918 S919 P920 D921 E922 E923 E924 L925 G926 Q927 R928 T929 A930 Y931 H932 S933 K934 R935 D936 A937 S938 T939 P940 L941 R942 S943 T944 E945 N946 Q951 V952 L953 T956 N957 L961 K962 F963 V964 K965 V966 R967 T971 I976
F980 A981 S982 R983 H984 I990 G991 I992 T993 Y994 R995 R996 E997 D998 M999 P1000 F1001 T1002 A1003 E1004 V1007 P1008 D1009 L1010 I1011 I1012 N1013 P1014 H1015 A1016 I1017 P1018 S1019 R1020 M1021 T1022 V1023 A1024 E1028 C1029 L1030 K1033 L1037 A1044 S1045 P1046 F1047 T1048 D1049 G1054 L1058

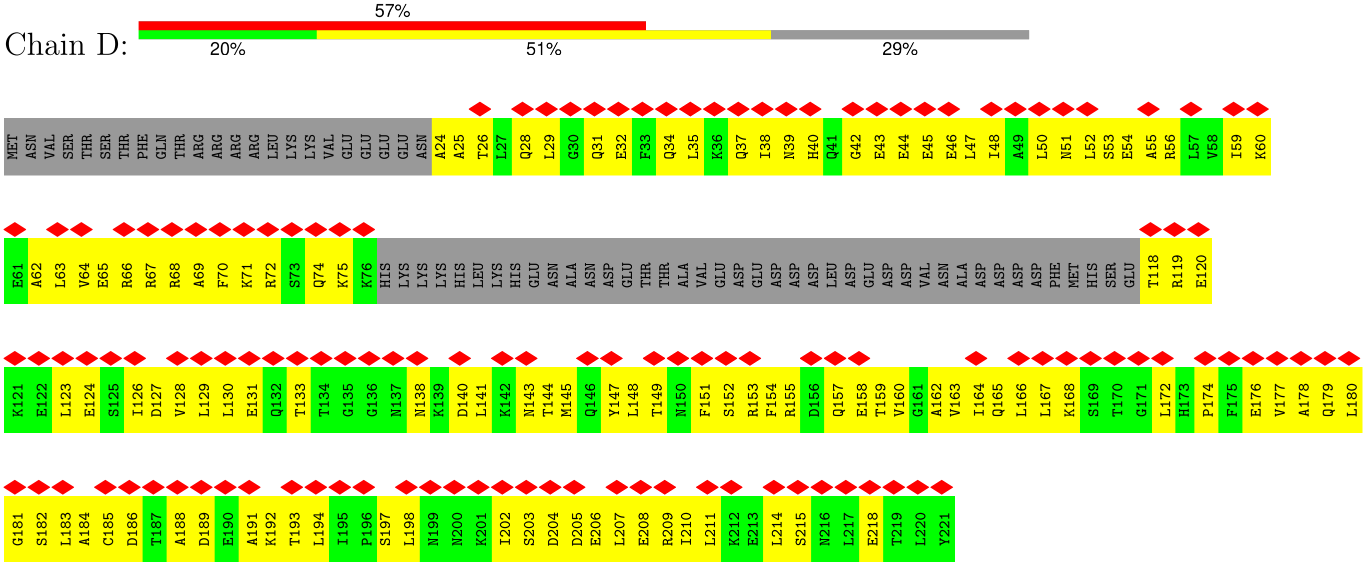
- Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C:



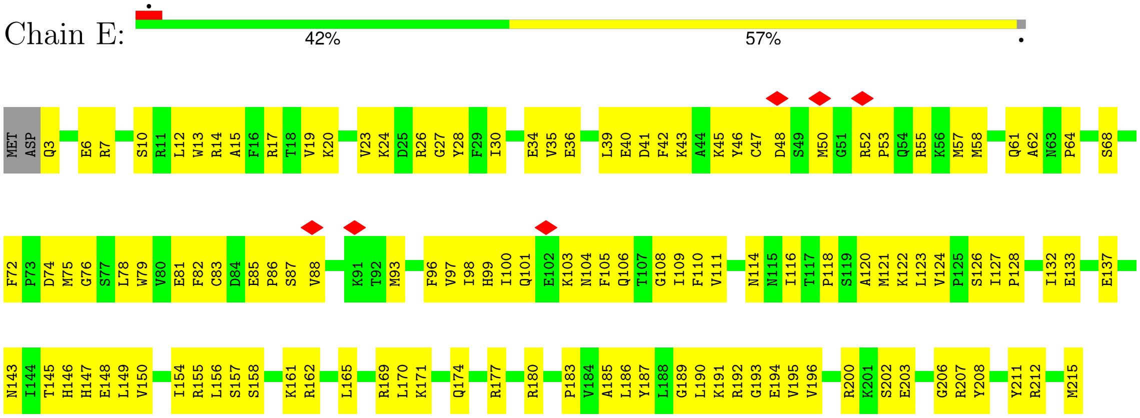
- Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D:



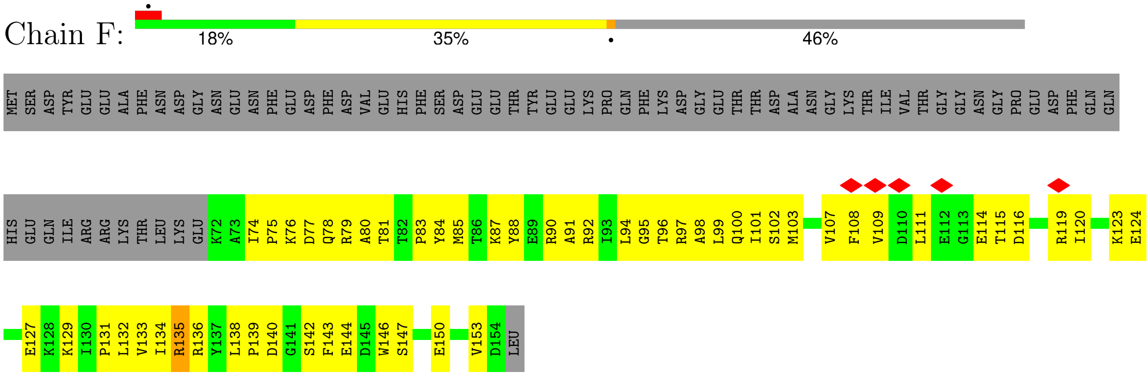
- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



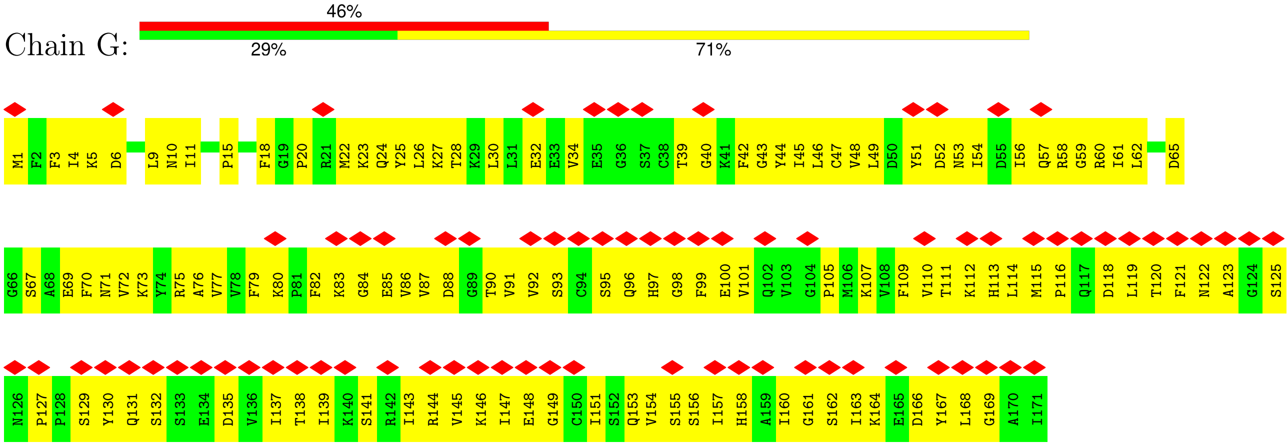
- Molecule 6: DNA-directed RNA polymerases I,II,and III subunit RPABC2

Chain F:



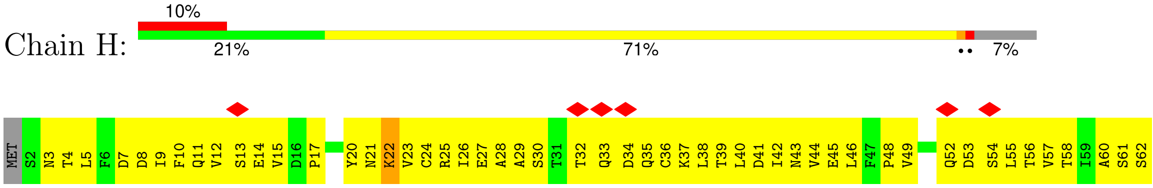
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

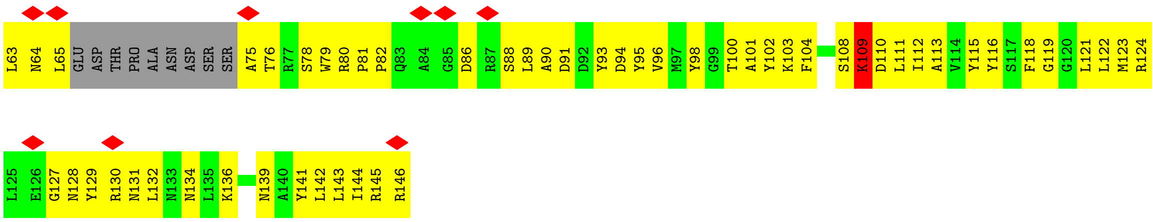
Chain G:



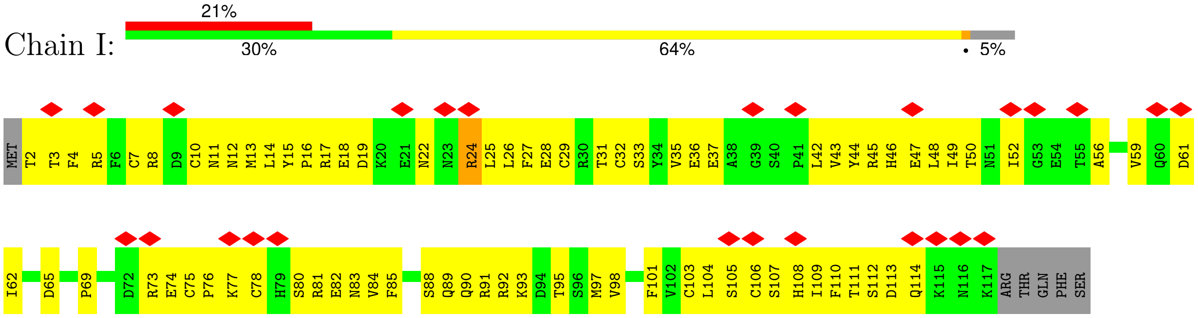
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:

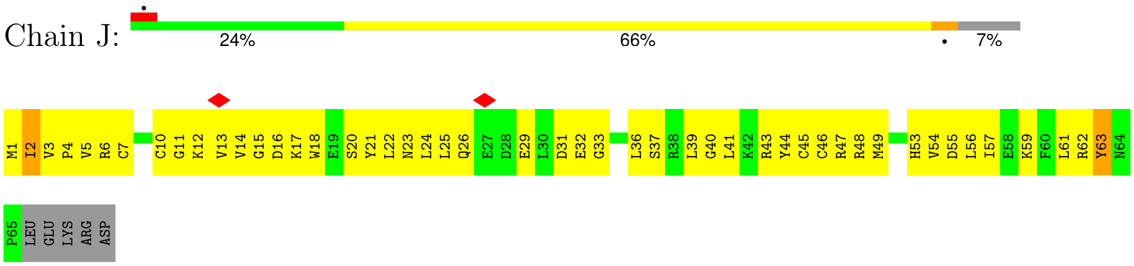




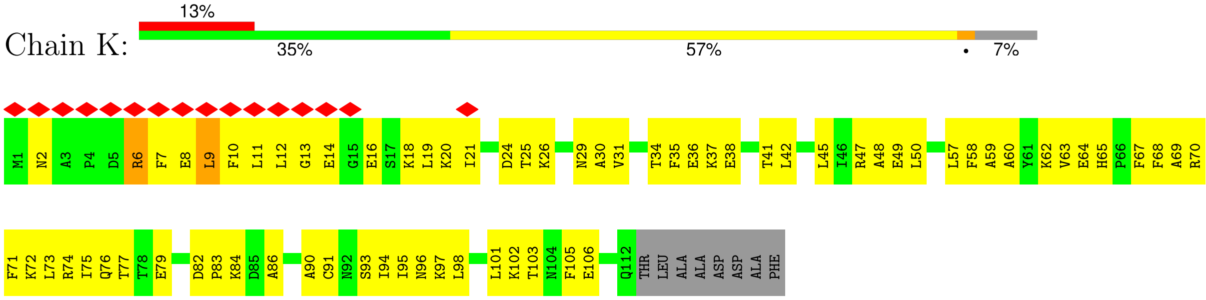
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



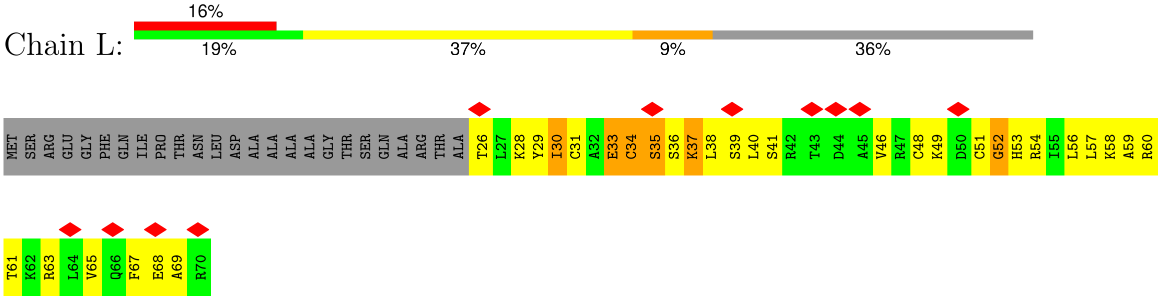
- Molecule 10: DNA-directed RNA polymerases II subunit RPABC5



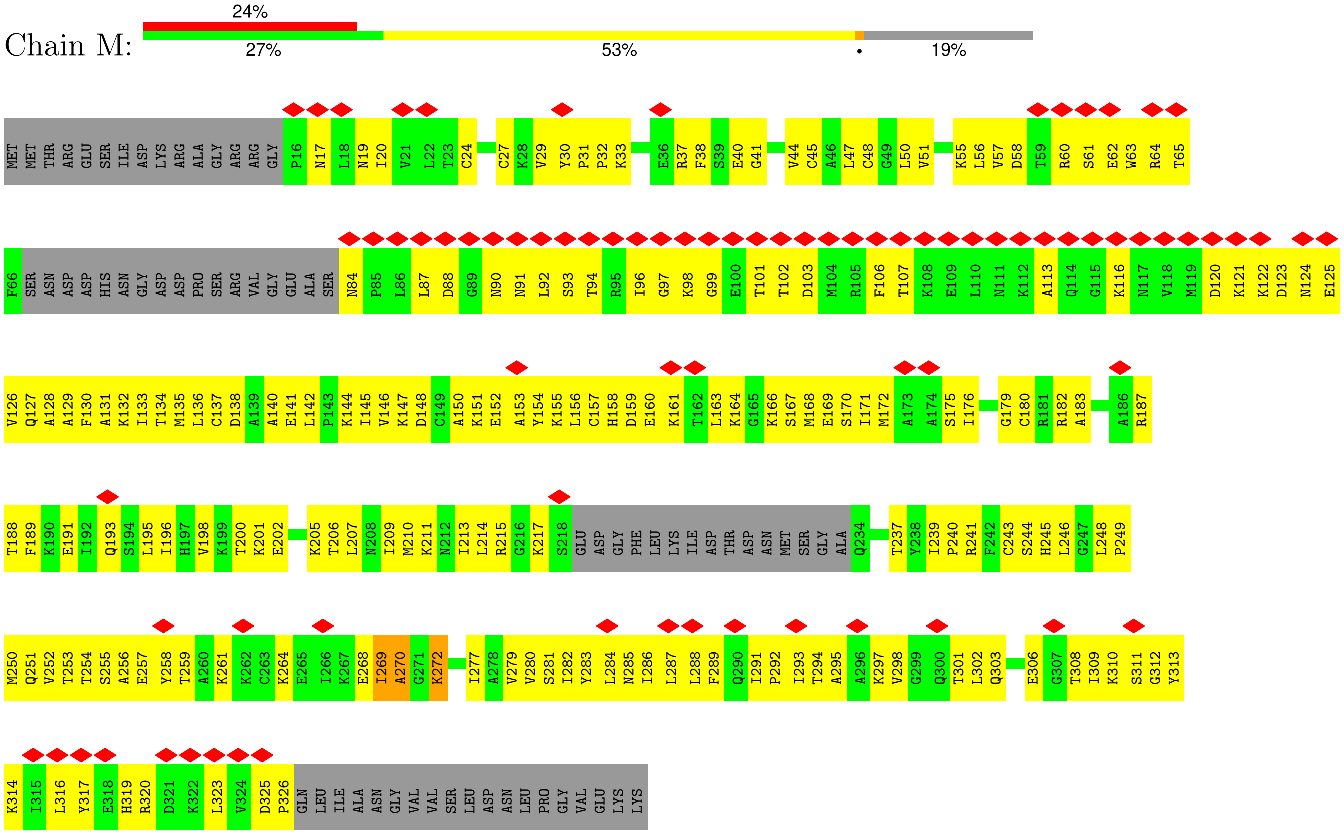
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11



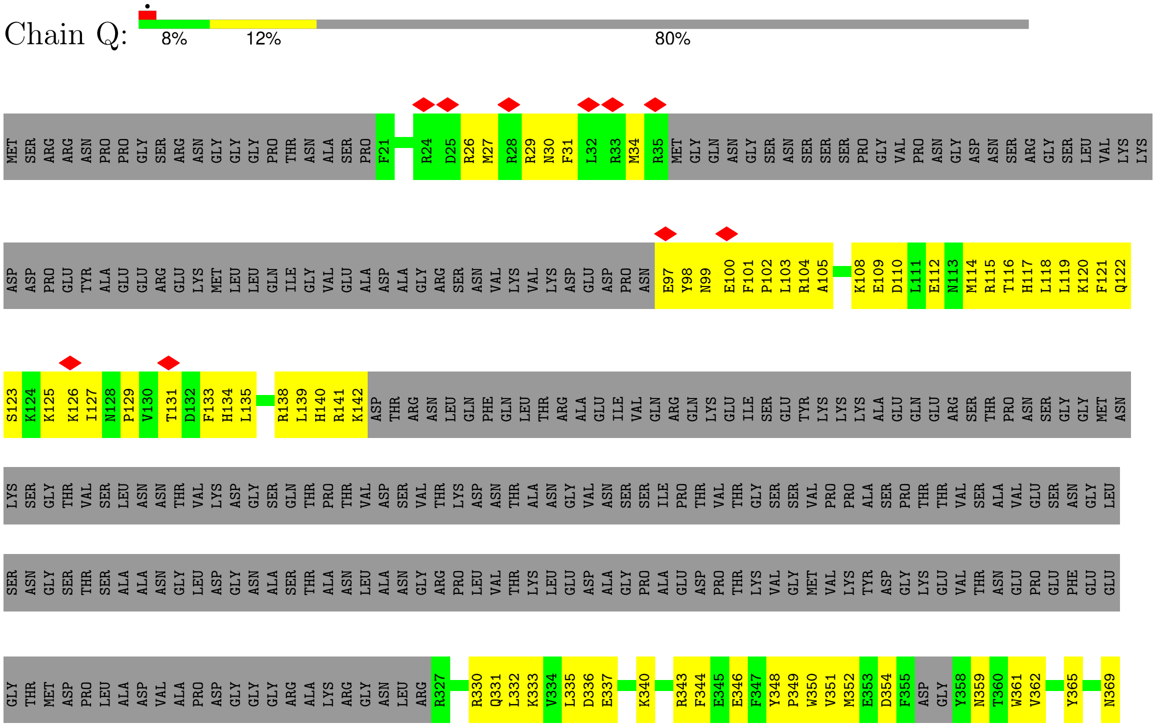
- Molecule 12: DNA-directed RNA polymerases II subunit RPABC4

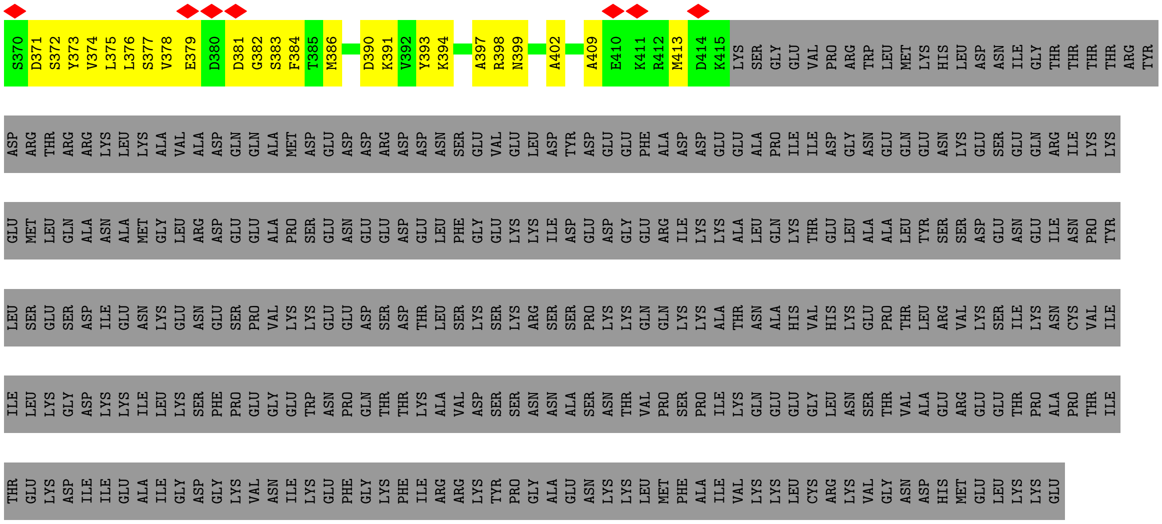


- Molecule 13: Transcription initiation factor IIB



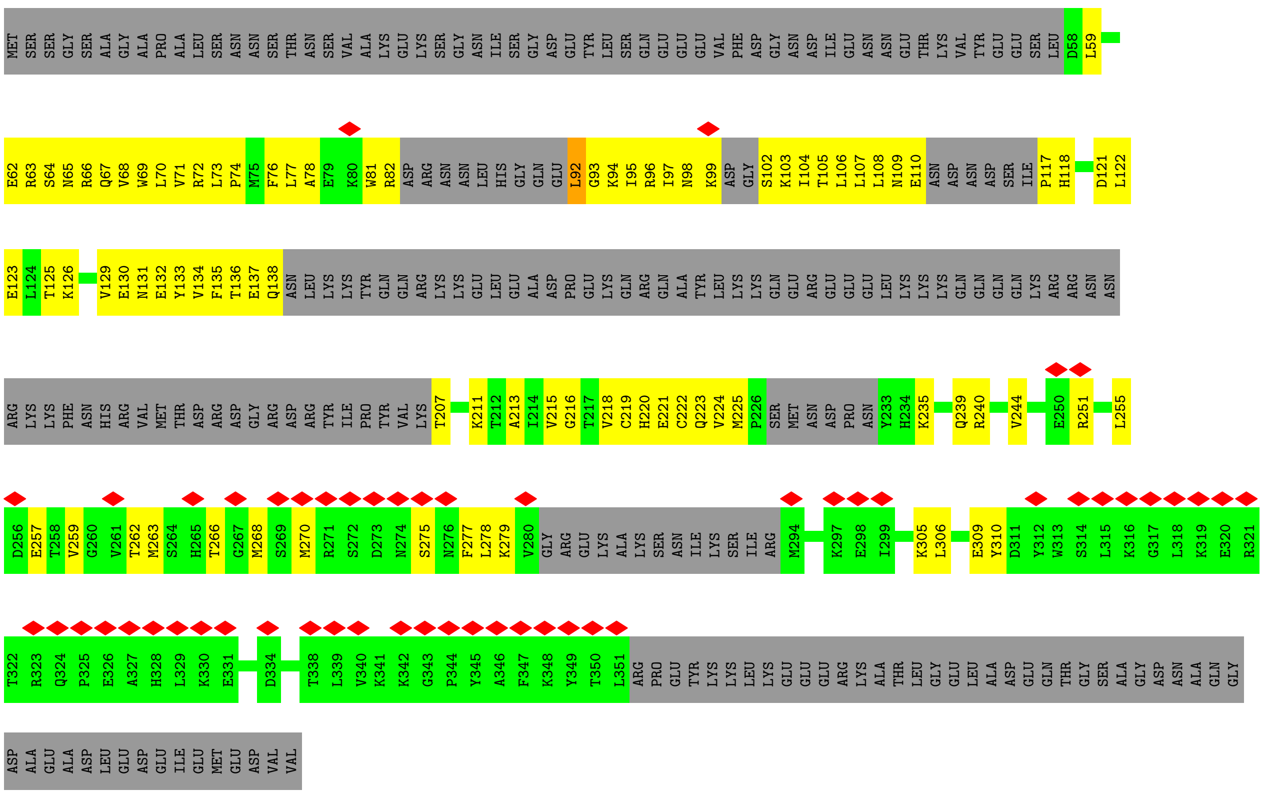
- Molecule 14: Transcription initiation factor IIF subunit alpha





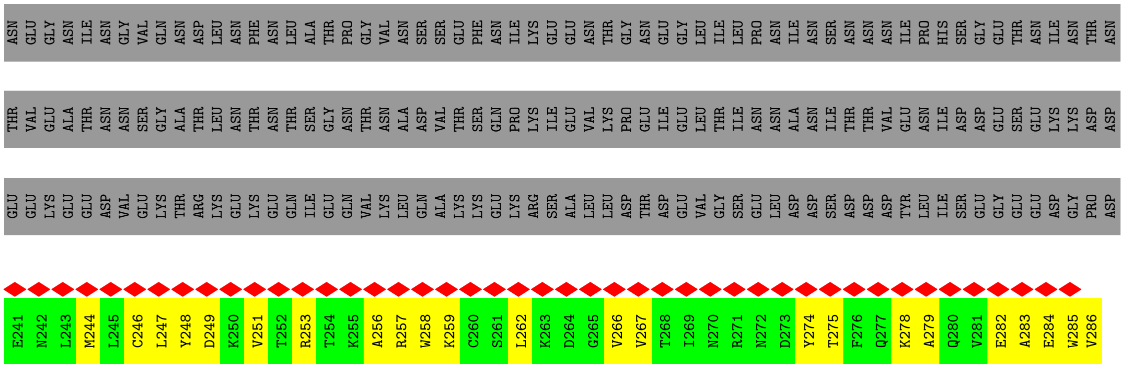
- Molecule 15: Transcription initiation factor IIF subunit beta

Chain R:

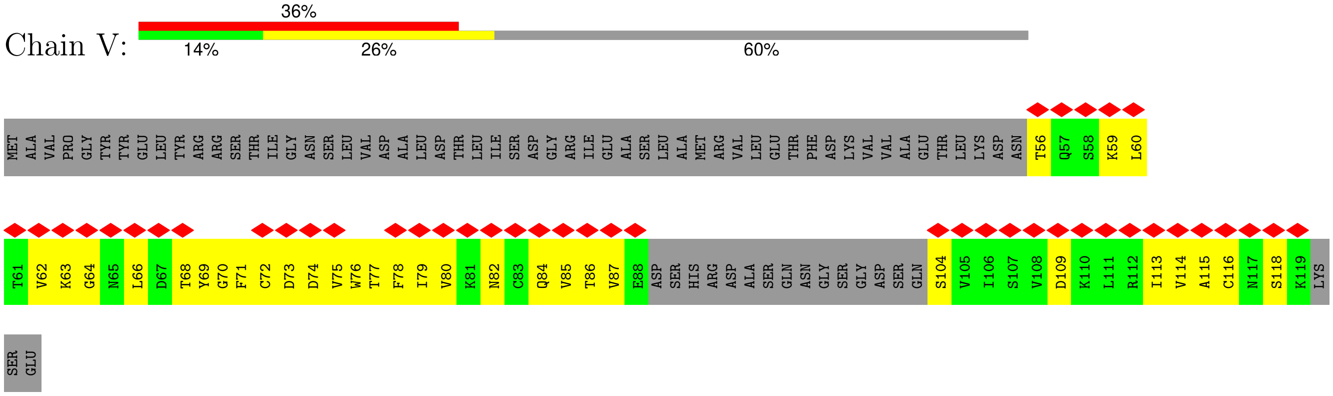


- Molecule 16: Transcription initiation factor IIA large subunit

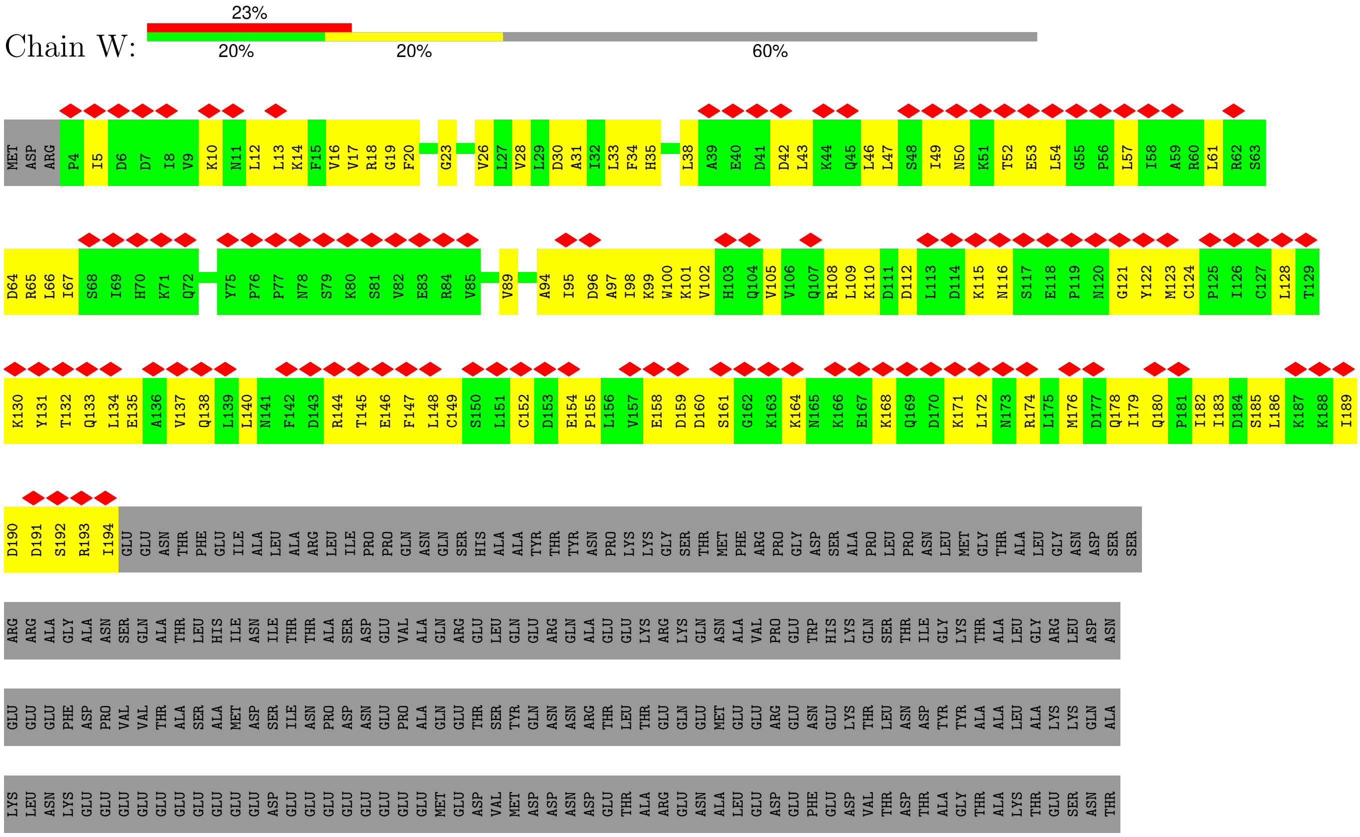
Chain U:



• Molecule 17: Transcription initiation factor IIA subunit 2

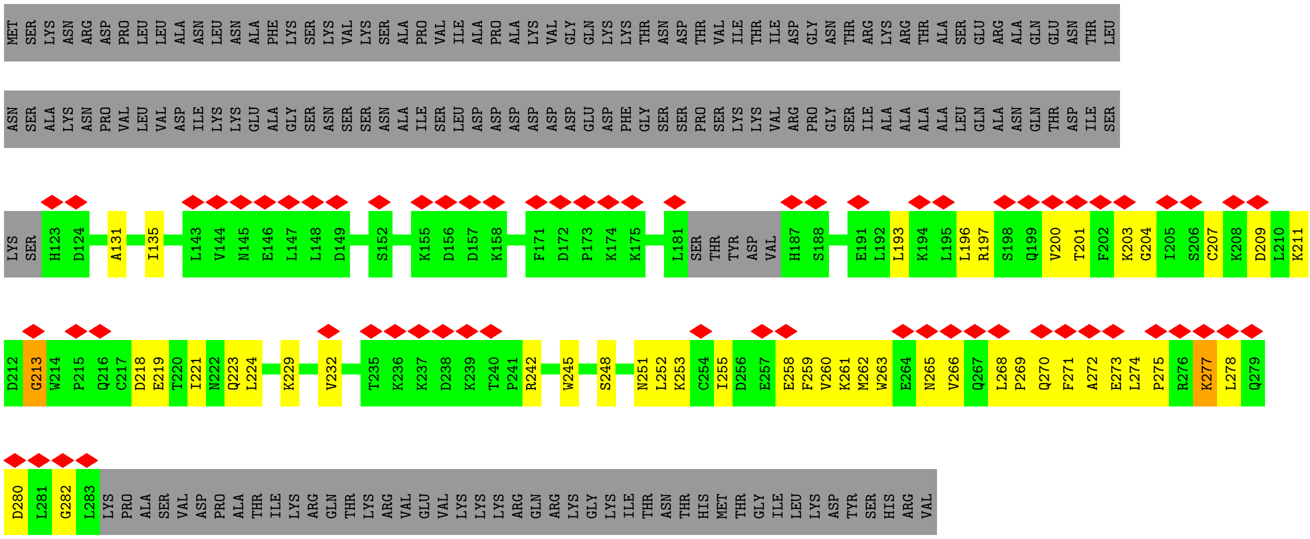


• Molecule 18: Transcription initiation factor IIE subunit alpha



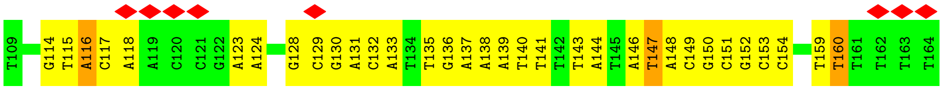
- Molecule 19: Transcription initiation factor IIE subunit beta

Chain X:



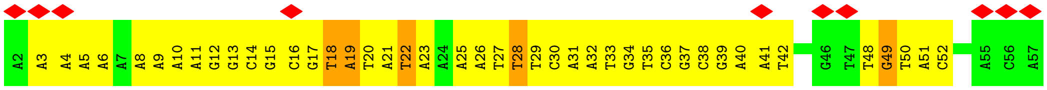
- Molecule 20: template strand DNA

Chain T:



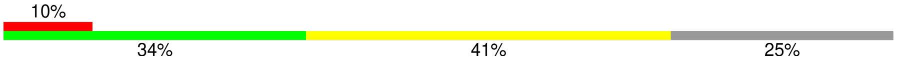
- Molecule 21: non-template strand DNA

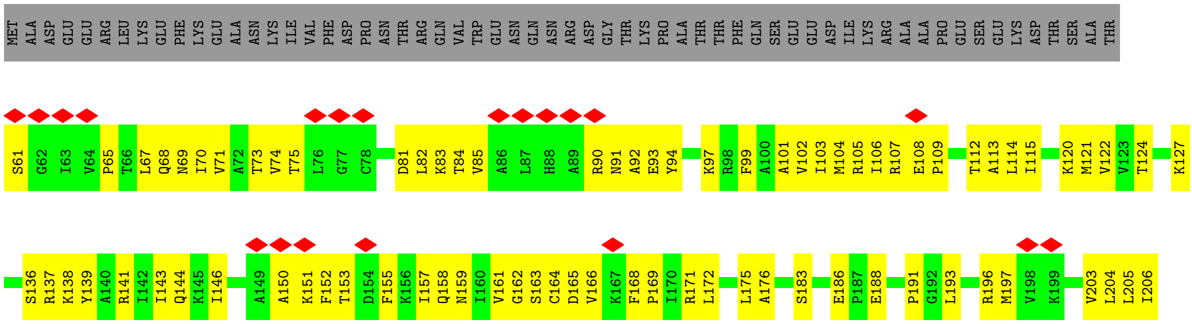
Chain N:

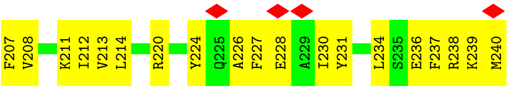


- Molecule 22: BJ4_G0004860.mRNA.1.CDS.1

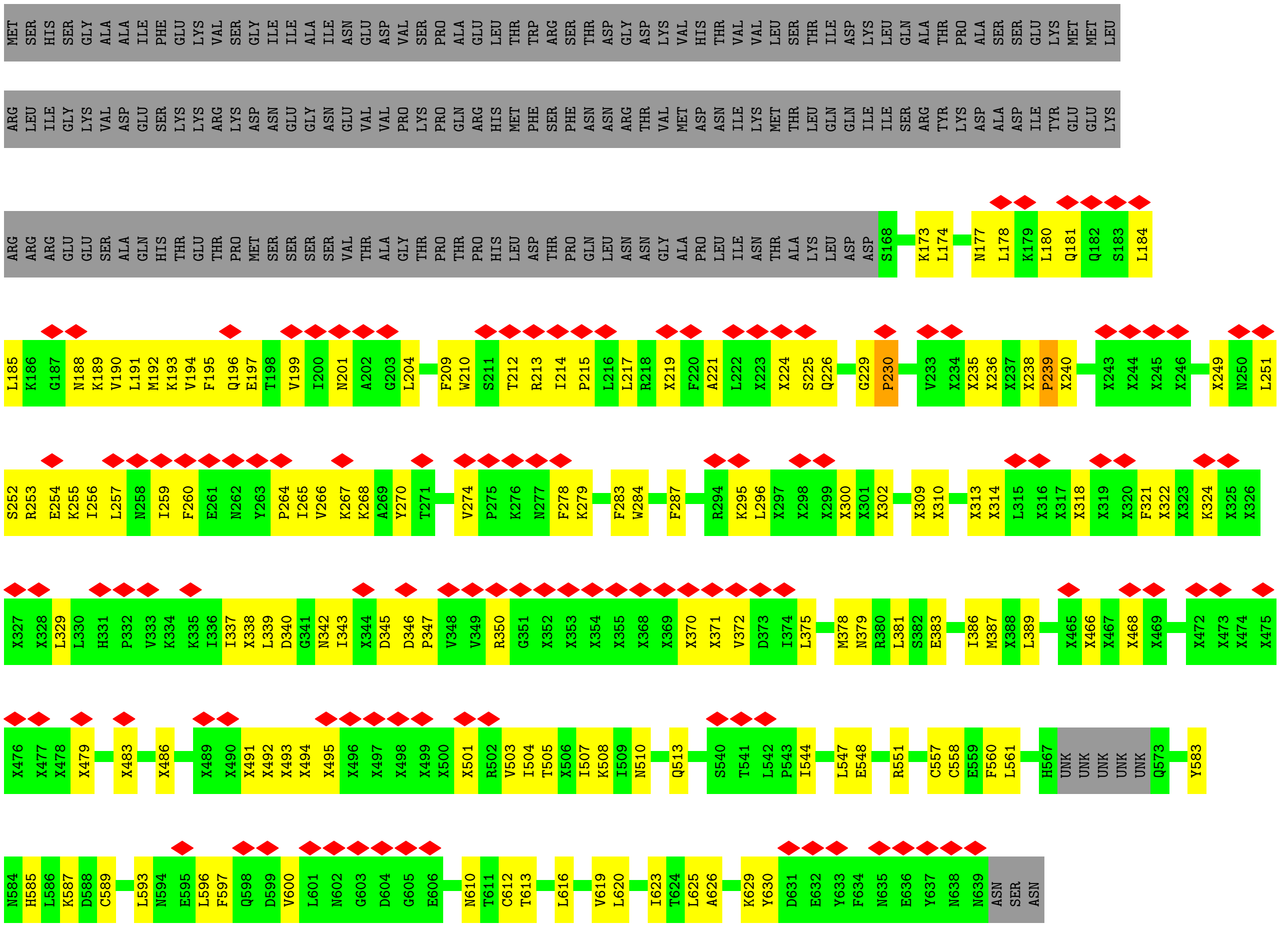
Chain O:



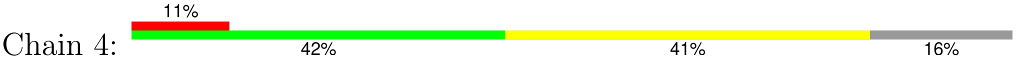


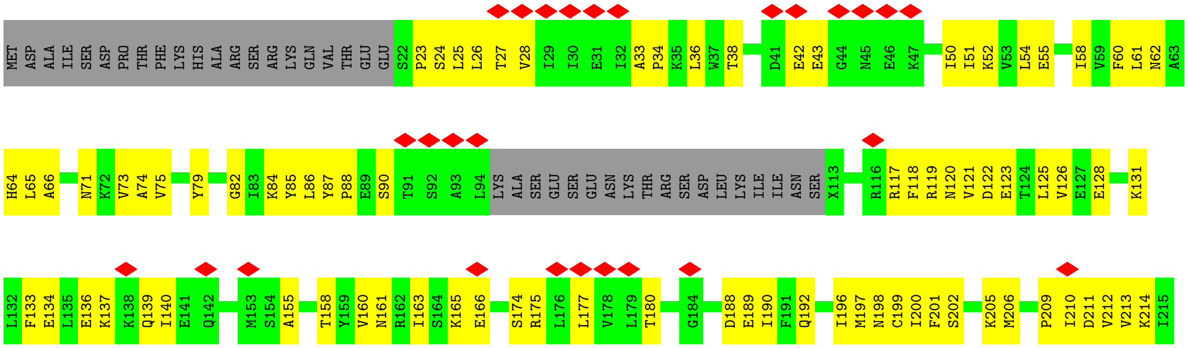


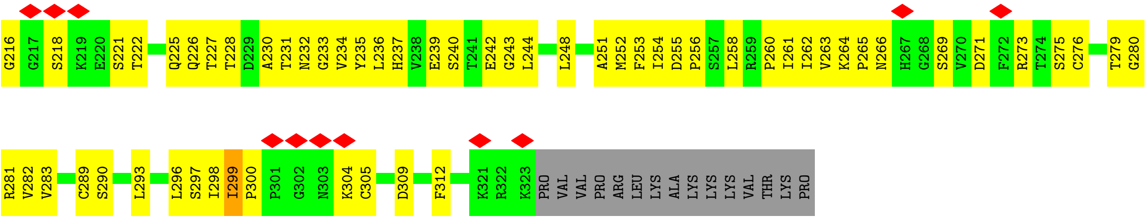
- Molecule 23: Tfb1



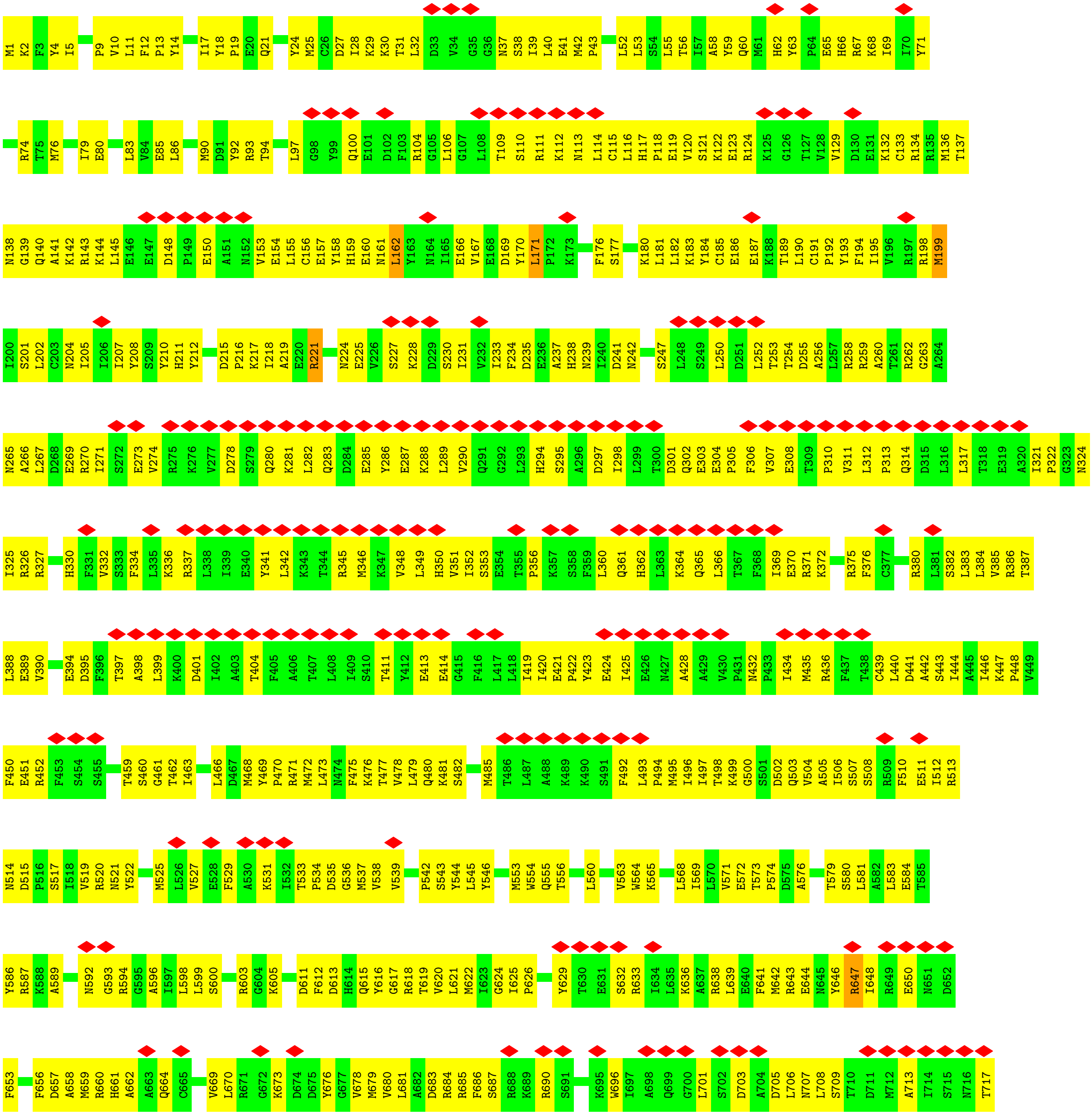
- Molecule 24: Tfb4

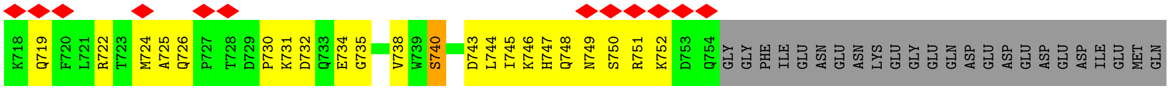






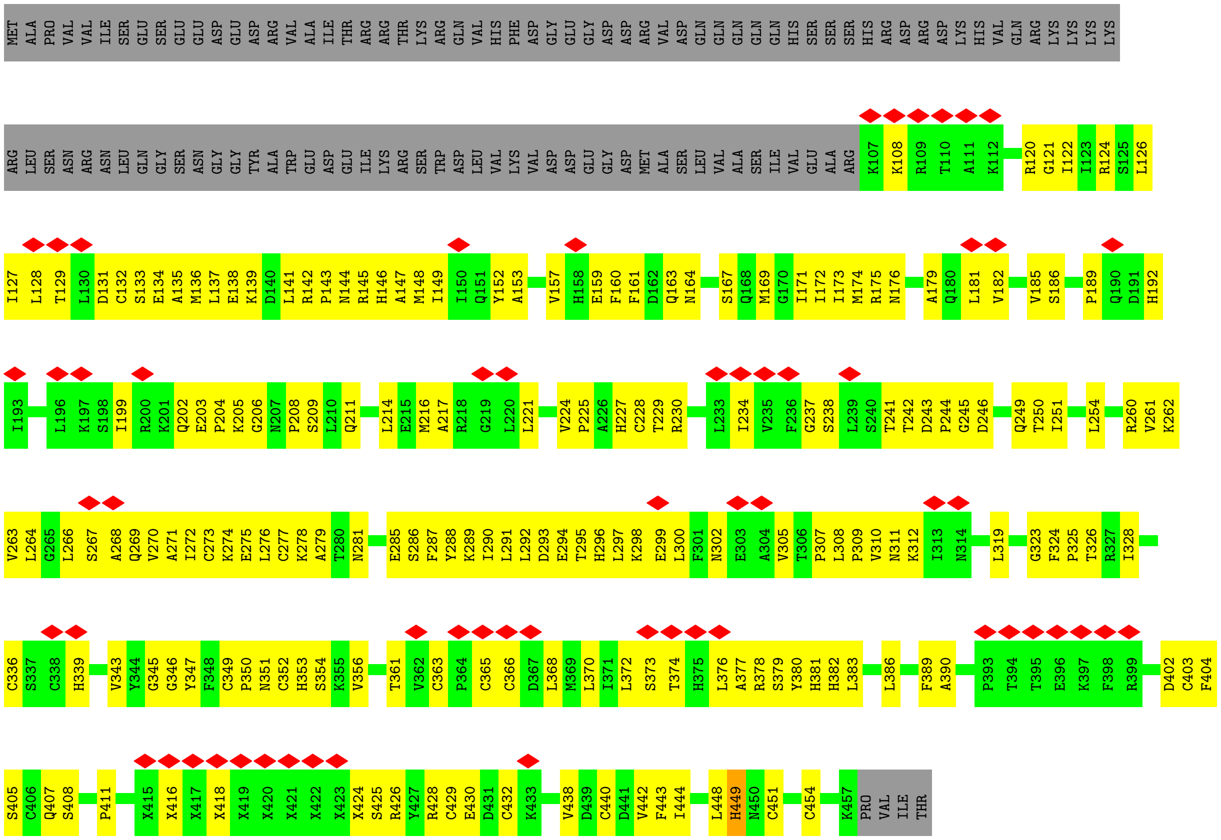
- Molecule 25: General transcription and DNA repair factor IIH helicase subunit XPD





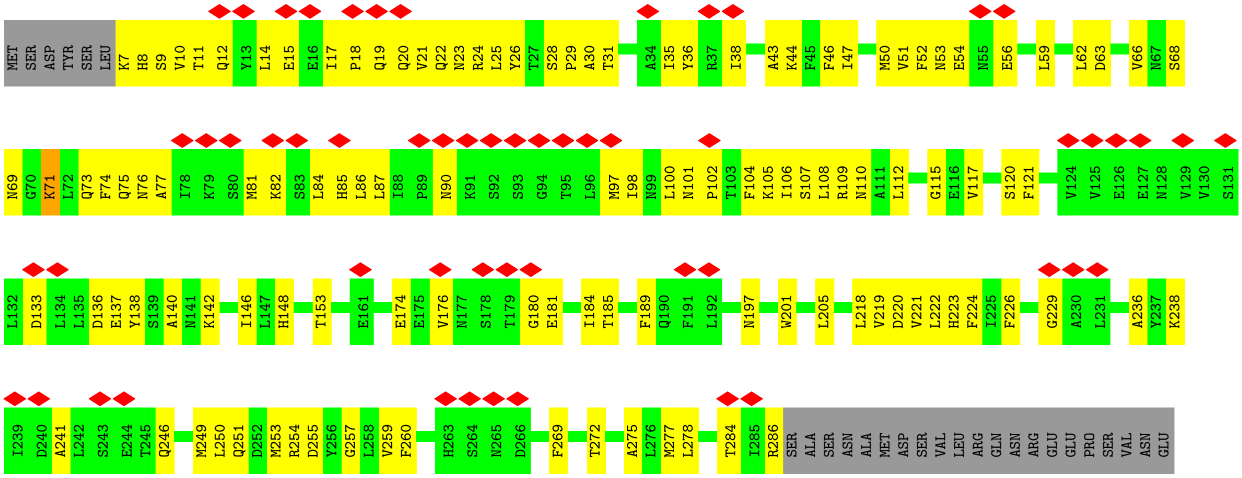
- Molecule 26: General transcription and DNA repair factor IIH

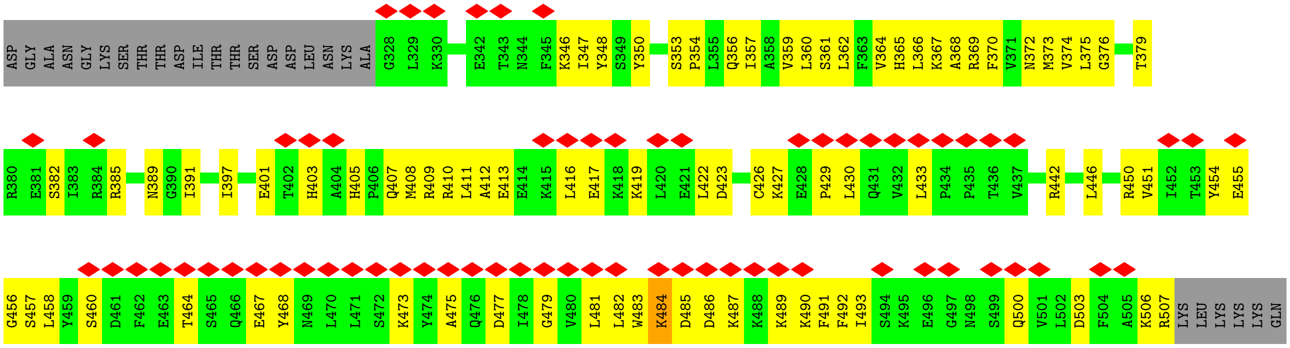
Chain 6:



- Molecule 27: RNA polymerase II transcription factor B subunit 2

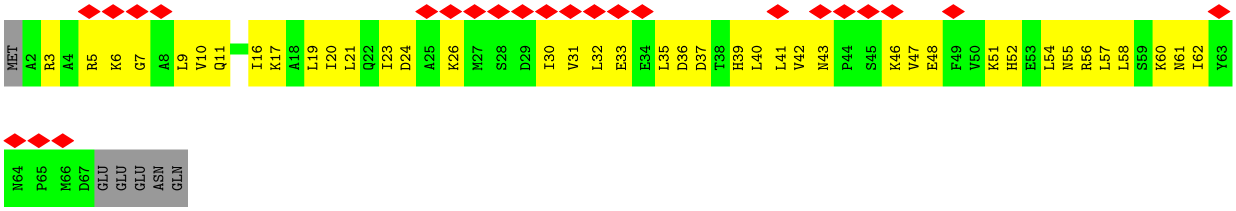
Chain 2:





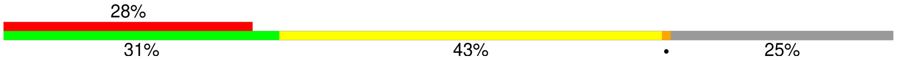
- Molecule 28: General transcription and DNA repair factor IIH subunit TFB5

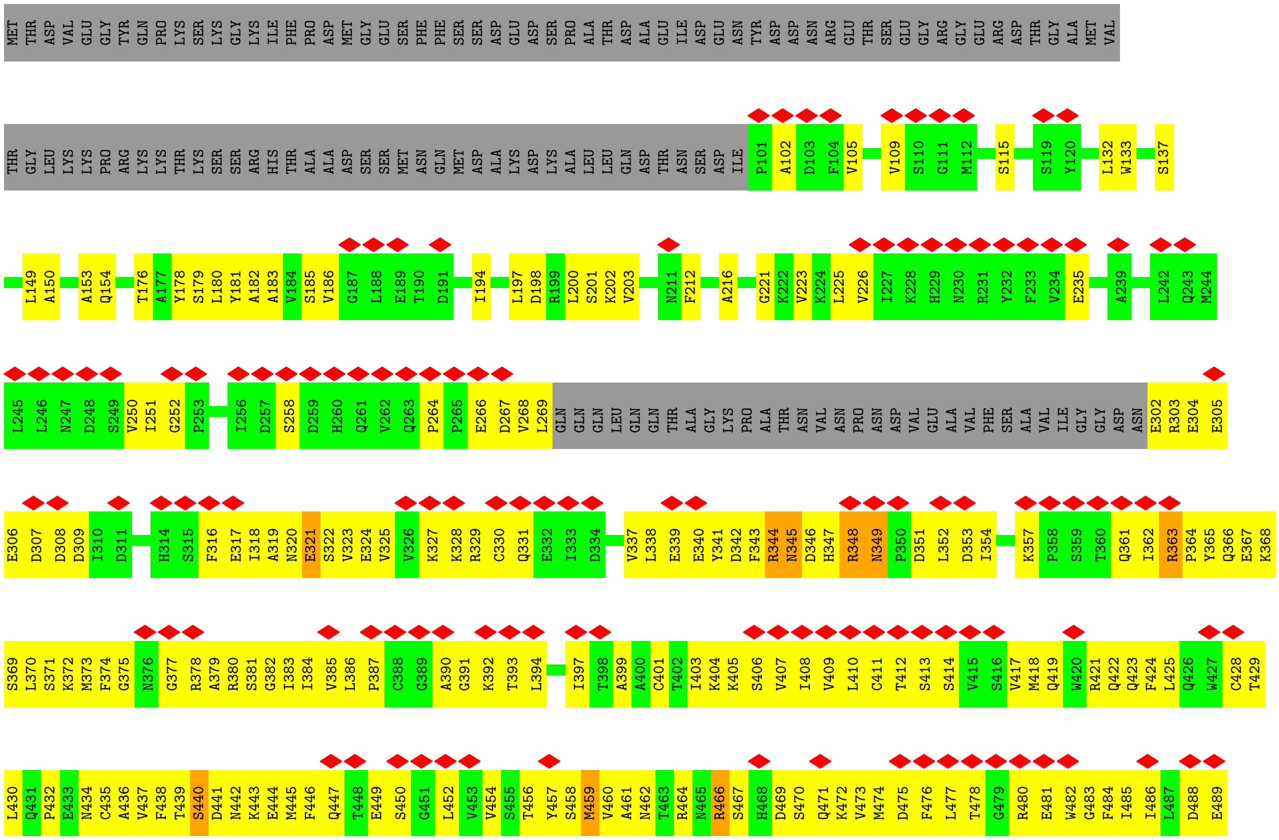
Chain 5:

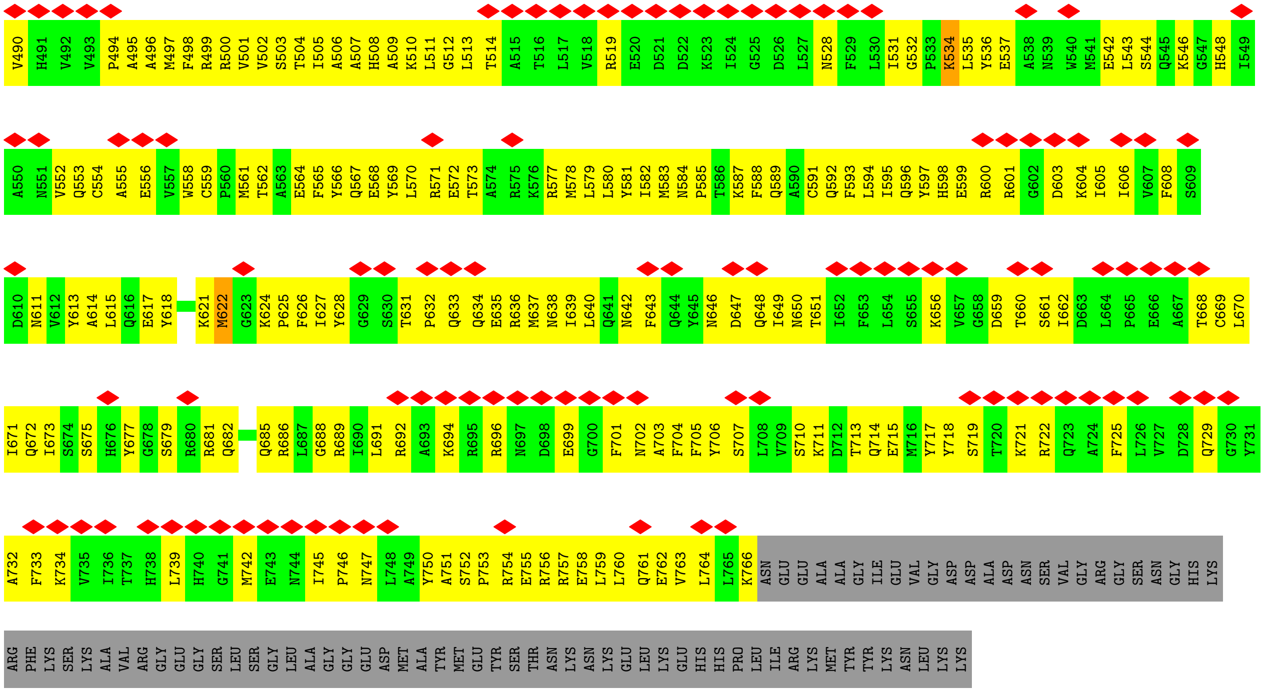


- Molecule 29: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 7:

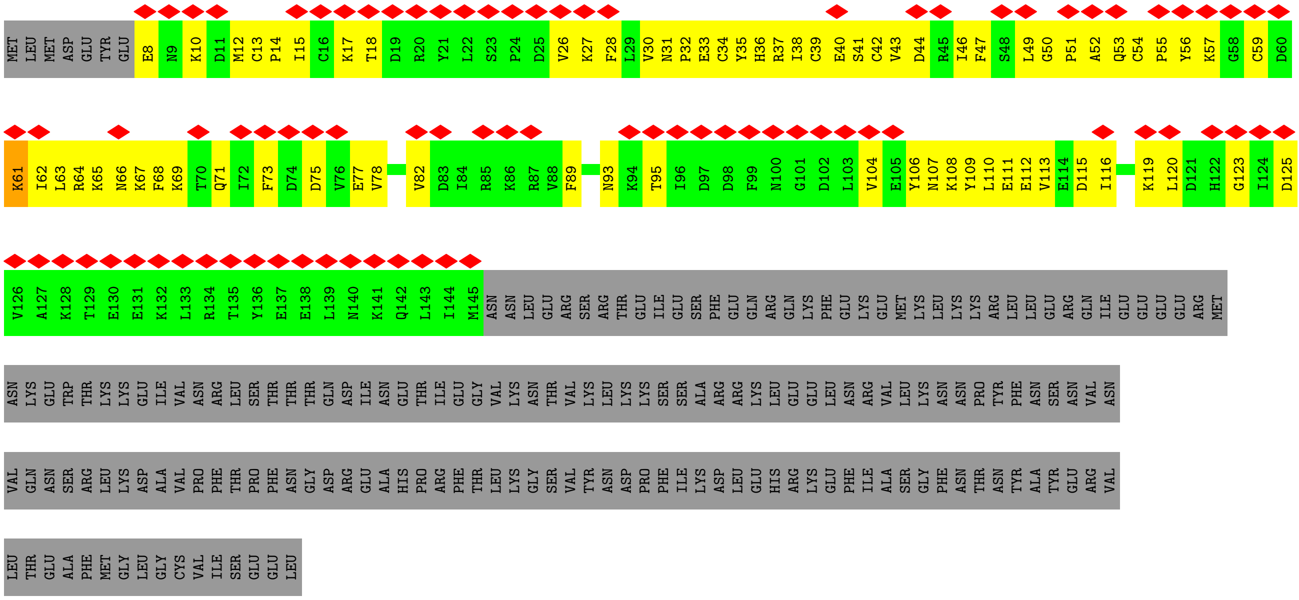






- Molecule 30: BJ4_G0050160.mRNA.1.CDS.1

Chain 3:



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	69513	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	45	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.056	Depositor
Minimum map value	0.000	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.018	Depositor
Map size (Å)	474.87997, 501.37997, 473.81998	wwPDB
Map dimensions	448, 473, 447	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MG, SF4, ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.49	3/11192 (0.0%)	0.66	6/15128 (0.0%)
2	B	0.50	0/9357	0.67	4/12618 (0.0%)
3	C	0.55	1/2099 (0.0%)	0.69	0/2845
4	D	0.34	0/1262	0.64	0/1693
5	E	0.45	0/1780	0.62	0/2395
6	F	0.65	1/682 (0.1%)	0.83	1/922 (0.1%)
7	G	0.38	0/1368	0.61	0/1844
8	H	0.55	0/1107	0.76	0/1499
9	I	0.49	0/962	0.77	1/1295 (0.1%)
10	J	0.62	0/541	0.83	1/727 (0.1%)
11	K	0.54	1/922 (0.1%)	0.90	4/1244 (0.3%)
12	L	0.37	0/360	0.71	0/478
13	M	0.37	0/2204	0.62	0/2963
14	Q	0.33	0/1168	0.53	0/1579
15	R	0.32	0/1312	0.57	1/1777 (0.1%)
16	U	0.25	0/389	0.57	0/523
17	V	0.28	0/384	0.48	0/518
18	W	0.29	0/1490	0.49	0/2014
19	X	0.25	0/993	0.50	0/1357
20	T	0.88	2/1273 (0.2%)	1.12	2/1962 (0.1%)
21	N	0.98	1/1301 (0.1%)	1.10	9/2006 (0.4%)
22	O	0.32	0/1443	0.53	0/1942
23	1	0.28	0/1896	0.50	0/2543
24	4	0.33	0/2062	0.57	0/2805
25	0	0.34	0/6226	0.57	2/8407 (0.0%)
26	6	0.33	0/2506	0.58	0/3402
27	2	0.30	0/3057	0.54	0/4071
28	5	0.27	0/502	0.58	0/677
29	7	0.38	0/4521	0.62	2/6036 (0.0%)
30	3	0.52	1/870 (0.1%)	0.62	1/1190 (0.1%)
All	All	0.46	10/65229 (0.0%)	0.66	34/88460 (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
1	A	0	7
2	B	0	4
8	H	0	3
9	I	0	1
10	J	0	1
13	M	0	3
18	W	0	1
19	X	0	1
24	4	0	2
25	0	0	2
26	6	0	1
29	7	0	8
30	3	0	2
All	All	0	36

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1074	GLU	CD-OE2	-10.39	1.14	1.25
30	3	61	LYS	CE-NZ	-7.98	1.29	1.49
3	C	166	GLU	CG-CD	7.34	1.62	1.51
20	T	116	DA	N9-C4	-7.07	1.33	1.37
6	F	135	ARG	CA-CB	-6.59	1.39	1.53

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
11	K	6	ARG	NE-CZ-NH1	-12.37	114.12	120.30
21	N	22	DT	O4’-C4’-C3’	-8.60	100.84	106.00
20	T	147	DT	O4’-C4’-C3’	-7.76	101.34	106.00
2	B	935	ARG	NE-CZ-NH1	-6.57	117.02	120.30
21	N	51	DA	O4’-C4’-C3’	-6.39	101.94	104.50

There are no chirality outliers.

5 of 36 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	22	PHE	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	250	ILE	Peptide
1	A	450	LEU	Peptide
1	A	465	TYR	Peptide
1	A	71	GLN	Peptide

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10997	0	11081	800	0
2	B	9178	0	9195	614	0
3	C	2061	0	2029	154	0
4	D	1253	0	1275	127	0
5	E	1744	0	1772	120	0
6	F	670	0	690	66	0
7	G	1340	0	1357	132	0
8	H	1089	0	1062	124	0
9	I	944	0	899	106	0
10	J	532	0	542	62	0
11	K	904	0	911	87	0
12	L	358	0	383	39	0
13	M	2175	0	2283	217	0
14	Q	1144	0	1034	113	0
15	R	1303	0	1110	122	0
16	U	383	0	384	24	0
17	V	381	0	388	35	0
18	W	1469	0	1432	110	0
19	X	984	0	722	58	0
20	T	1140	0	641	71	0
21	N	1156	0	631	87	0
22	O	1416	0	1493	99	0
23	1	2411	0	1881	121	0
24	4	2041	0	1954	140	0
25	0	6108	0	6168	537	0
26	6	2527	0	2321	221	0
27	2	3011	0	2600	225	0
28	5	498	0	506	58	0
29	7	4447	0	3905	483	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
30	3	860	0	623	103	0
31	3	2	0	0	0	0
31	4	1	0	0	0	0
31	6	4	0	0	0	0
31	A	2	0	0	0	0
31	B	1	0	0	0	0
31	C	1	0	0	0	0
31	I	2	0	0	0	0
31	J	1	0	0	0	0
31	L	1	0	0	0	0
31	M	1	0	0	0	0
31	W	1	0	0	0	0
32	A	1	0	0	0	0
33	0	8	0	0	3	0
All	All	64550	0	61272	4827	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 39.

The worst 5 of 4827 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)
12:L:48:CYS:SG	12:L:53:HIS:HB3	1.70	1.32
12:L:34:CYS:SG	12:L:36:SER:OG	1.98	1.18
2:B:649:LYS:NZ	2:B:736:THR:O	1.83	1.11
25:0:162:LEU:HD22	25:0:194:PHE:HB3	1.27	1.09
1:A:253:ASN:HA	2:B:935:ARG:HH12	1.17	1.07

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 PDB entries followed by that with respect to all EM entries.

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	1386/1733 (80%)	1192 (86%)	189 (14%)	5 (0%)	30	60
2	B	1136/1224 (93%)	987 (87%)	146 (13%)	3 (0%)	37	66
3	C	260/318 (82%)	214 (82%)	46 (18%)	0	100	100
4	D	153/221 (69%)	135 (88%)	18 (12%)	0	100	100
5	E	211/215 (98%)	193 (92%)	18 (8%)	0	100	100
6	F	81/155 (52%)	69 (85%)	12 (15%)	0	100	100
7	G	169/171 (99%)	152 (90%)	17 (10%)	0	100	100
8	H	132/146 (90%)	106 (80%)	24 (18%)	2 (2%)	8	30
9	I	114/122 (93%)	94 (82%)	20 (18%)	0	100	100
10	J	63/70 (90%)	44 (70%)	19 (30%)	0	100	100
11	K	110/120 (92%)	95 (86%)	15 (14%)	0	100	100
12	L	43/70 (61%)	33 (77%)	7 (16%)	3 (7%)	1	6
13	M	273/345 (79%)	221 (81%)	52 (19%)	0	100	100
14	Q	140/735 (19%)	120 (86%)	20 (14%)	0	100	100
15	R	176/400 (44%)	163 (93%)	13 (7%)	0	100	100
16	U	44/286 (15%)	38 (86%)	6 (14%)	0	100	100
17	V	45/122 (37%)	44 (98%)	1 (2%)	0	100	100
18	W	189/482 (39%)	182 (96%)	7 (4%)	0	100	100
19	X	152/328 (46%)	136 (90%)	16 (10%)	0	100	100
22	O	178/240 (74%)	167 (94%)	11 (6%)	0	100	100
23	1	256/542 (47%)	234 (91%)	19 (7%)	3 (1%)	11	35
24	4	279/338 (82%)	222 (80%)	57 (20%)	0	100	100
25	0	752/778 (97%)	670 (89%)	82 (11%)	0	100	100
26	6	336/461 (73%)	298 (89%)	36 (11%)	2 (1%)	22	50
27	2	456/513 (89%)	390 (86%)	66 (14%)	0	100	100
28	5	64/72 (89%)	58 (91%)	6 (9%)	0	100	100
29	7	630/843 (75%)	539 (86%)	90 (14%)	1 (0%)	44	72
30	3	136/321 (42%)	105 (77%)	31 (23%)	0	100	100
All	All	7964/11371 (70%)	6901 (87%)	1044 (13%)	19 (0%)	45	72

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	465	TYR
2	B	364	ILE
8	H	110	ASP
26	6	411	PRO
29	7	349	ASN

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 PDB entries followed by that with respect to all EM entries.

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	1221/1520 (80%)	1218 (100%)	3 (0%)	92	96
2	B	1000/1061 (94%)	1000 (100%)	0	100	100
3	C	230/274 (84%)	230 (100%)	0	100	100
4	D	139/200 (70%)	139 (100%)	0	100	100
5	E	195/197 (99%)	195 (100%)	0	100	100
6	F	73/137 (53%)	73 (100%)	0	100	100
7	G	152/152 (100%)	152 (100%)	0	100	100
8	H	119/128 (93%)	118 (99%)	1 (1%)	79	87
9	I	110/116 (95%)	110 (100%)	0	100	100
10	J	60/65 (92%)	60 (100%)	0	100	100
11	K	97/102 (95%)	97 (100%)	0	100	100
12	L	40/57 (70%)	34 (85%)	6 (15%)	2	9
13	M	245/299 (82%)	245 (100%)	0	100	100
14	Q	109/641 (17%)	109 (100%)	0	100	100
15	R	107/363 (30%)	107 (100%)	0	100	100
16	U	42/260 (16%)	42 (100%)	0	100	100
17	V	46/108 (43%)	46 (100%)	0	100	100
18	W	155/429 (36%)	155 (100%)	0	100	100
19	X	62/295 (21%)	61 (98%)	1 (2%)	58	75
22	O	152/205 (74%)	151 (99%)	1 (1%)	81	88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	1	169/395 (43%)	169 (100%)	0	100	100
24	4	198/298 (66%)	198 (100%)	0	100	100
25	0	686/707 (97%)	684 (100%)	2 (0%)	91	95
26	6	247/406 (61%)	246 (100%)	1 (0%)	89	93
27	2	258/468 (55%)	255 (99%)	3 (1%)	67	80
28	5	53/66 (80%)	53 (100%)	0	100	100
29	7	414/737 (56%)	413 (100%)	1 (0%)	92	96
30	3	53/303 (18%)	53 (100%)	0	100	100
All	All	6432/9989 (64%)	6413 (100%)	19 (0%)	90	95

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
26	6	108	LYS
27	2	484	LYS
29	7	534	LYS
27	2	410	ARG
12	L	38	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
23	1	196	GLN
26	6	163	GLN
25	0	60	GLN
25	0	294	HIS
26	6	302	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 19 ligands modelled in this entry, 18 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
33	SF4	0	801	25	0,12,12	-	-	-		

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
33	SF4	0	801	25	-	-	0/6/5/5

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.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
33	0	801	SF4	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
23	1	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	1	393:UNK	C	465:UNK	N	84.96
1	1	519:UNK	C	537:GLU	N	11.53
1	1	355:UNK	C	368:UNK	N	10.44

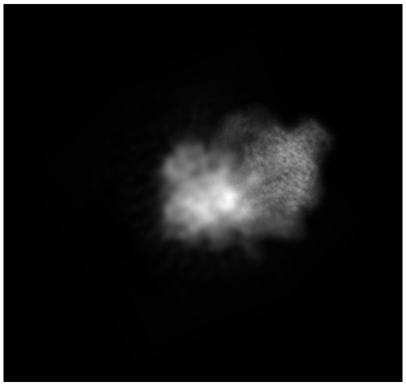
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-23906. These allow visual inspection of the internal detail of the map and identification of artifacts.

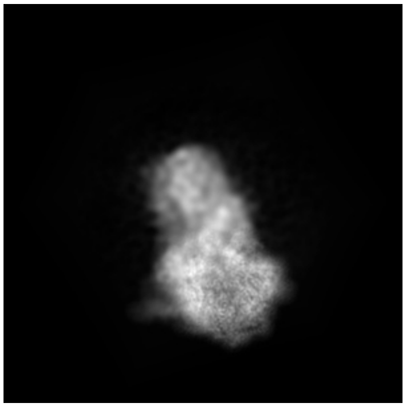
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

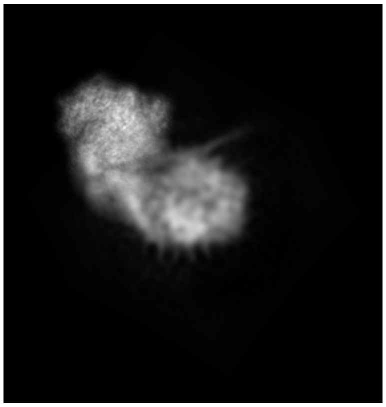
6.1.1 Primary map



X



Y

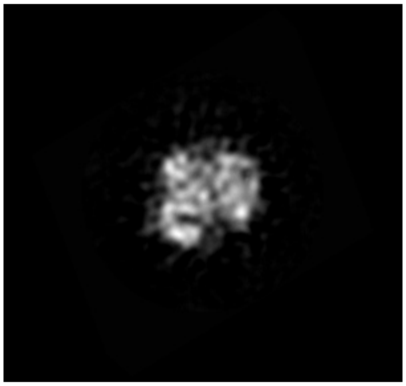


Z

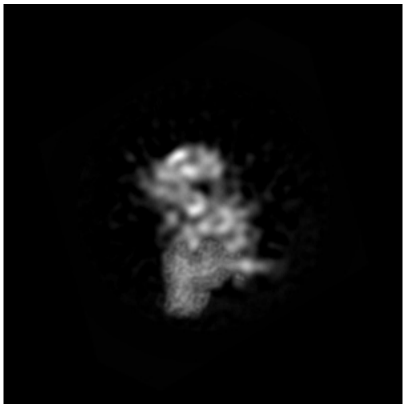
The images above show the map projected in three orthogonal directions.

6.2 Central slices

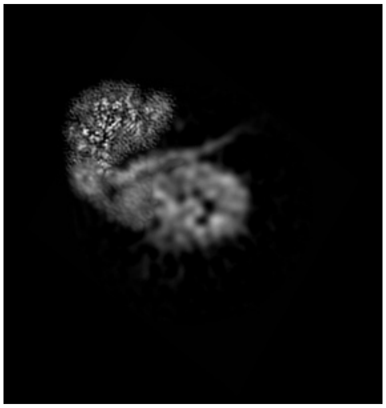
6.2.1 Primary map



X Index: 224



Y Index: 236

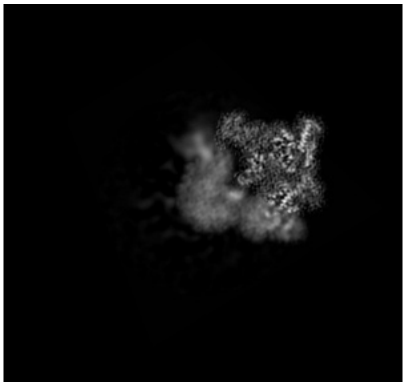


Z Index: 223

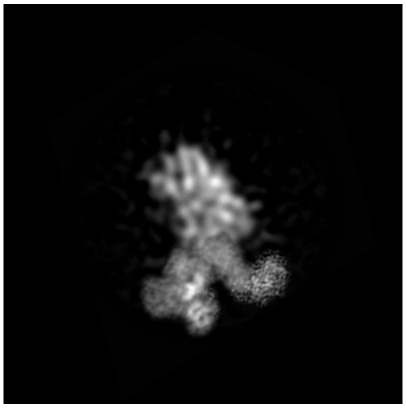
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices

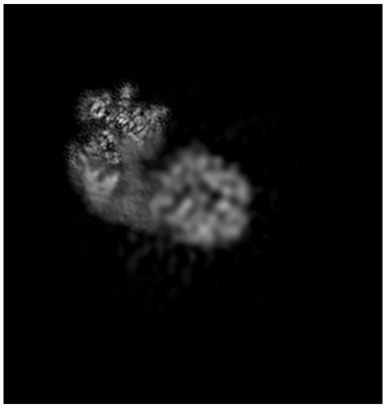
6.3.1 Primary map



X Index: 152



Y Index: 263

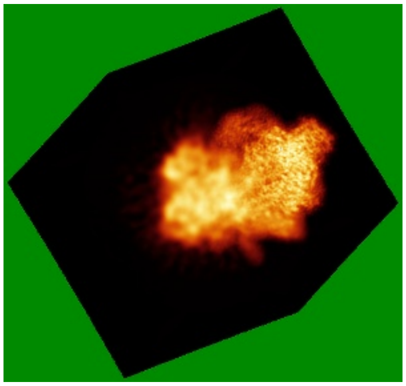


Z Index: 210

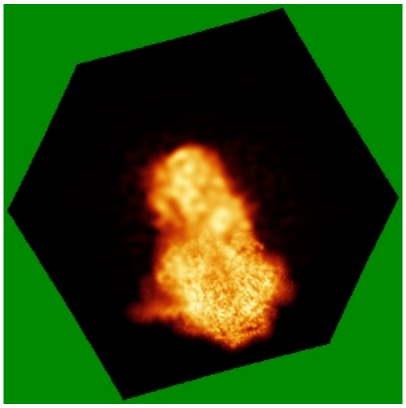
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color)

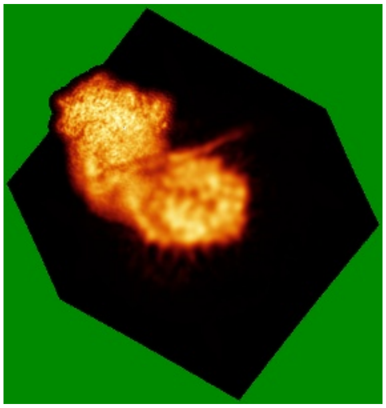
6.4.1 Primary map



X



Y

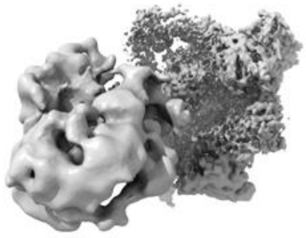


Z

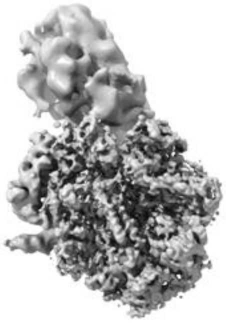
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views

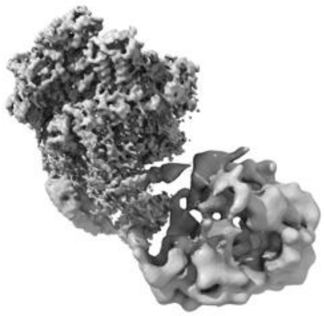
6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.018. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

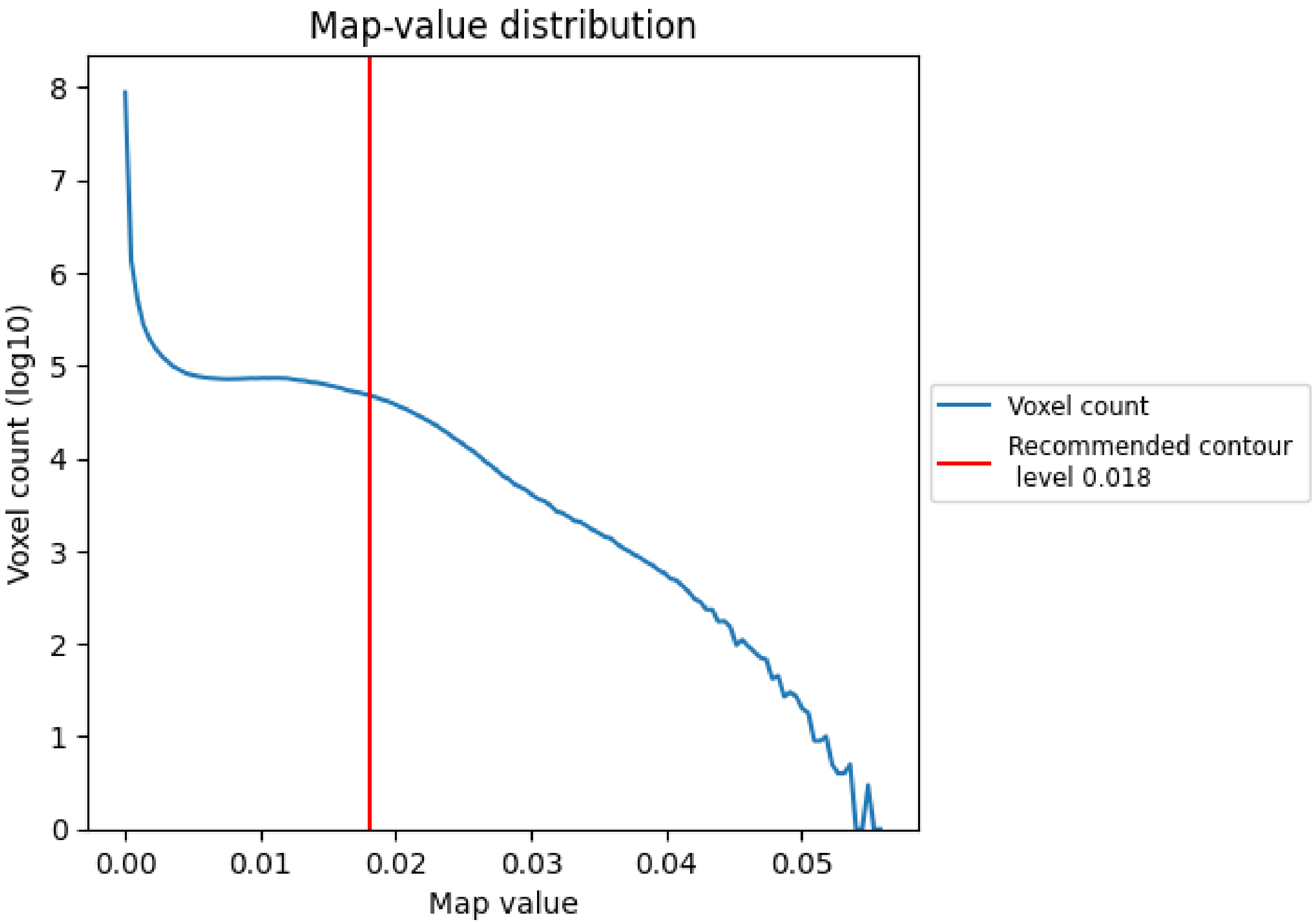
6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis ⓘ

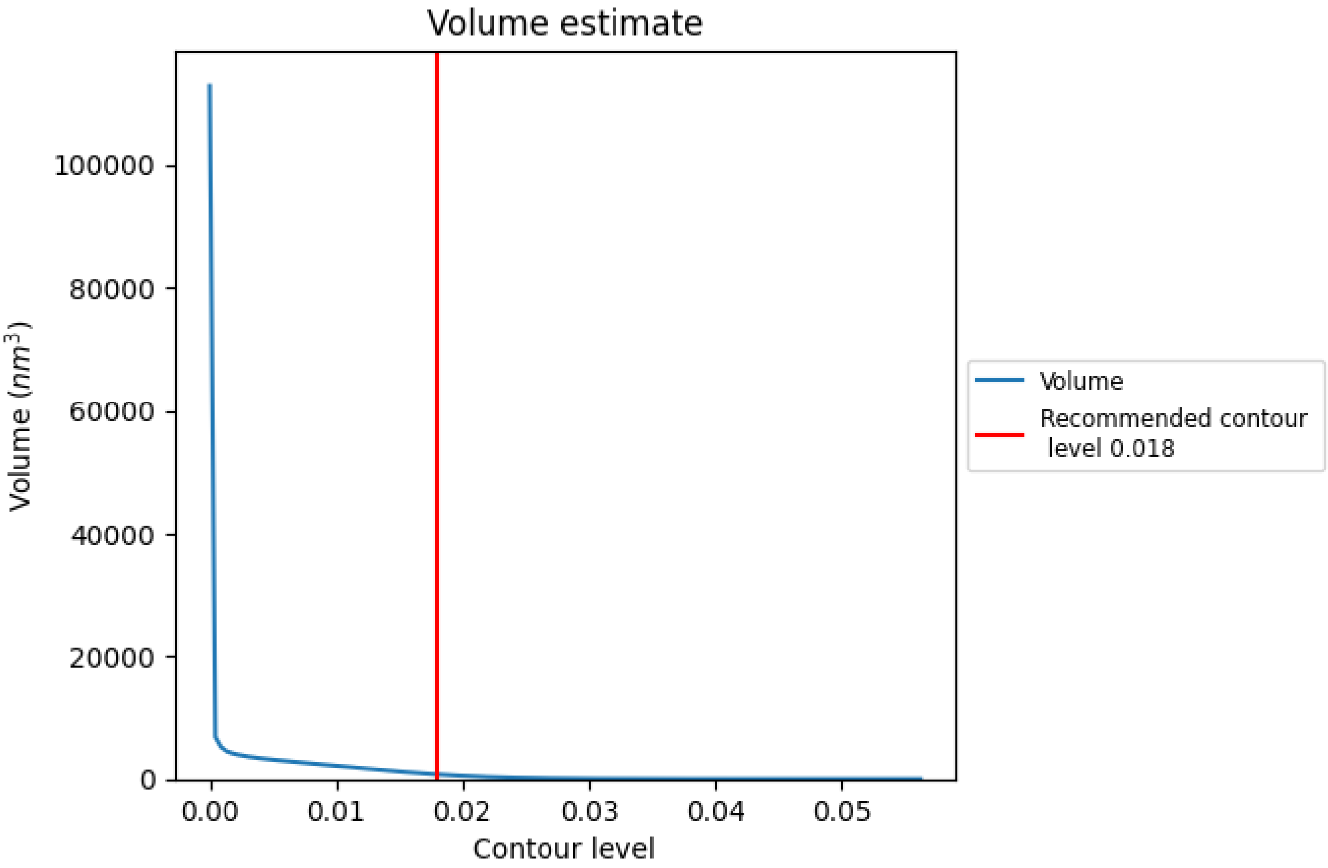
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2 Volume estimate ⓘ



The volume at the recommended contour level is 762 nm^3; this corresponds to an approximate mass of 688 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum is only generated for cubic maps.

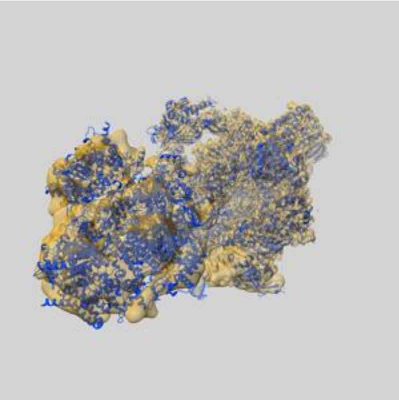
8 Fourier-Shell correlation ⓘ

This section was not generated. No FSC curve or half-maps provided.

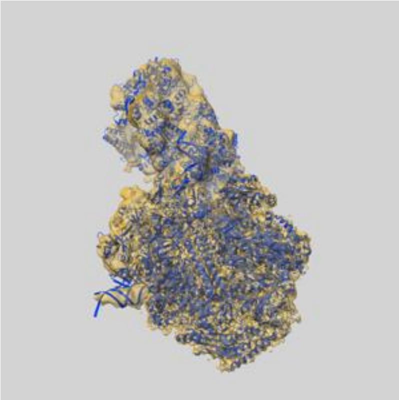
9 Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-23906 and PDB model 7ML2. Per-residue inclusion information can be found in section 3 on page 11.

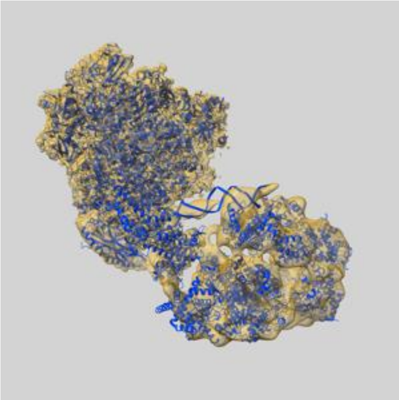
9.1 Map-model overlay ⓘ



X



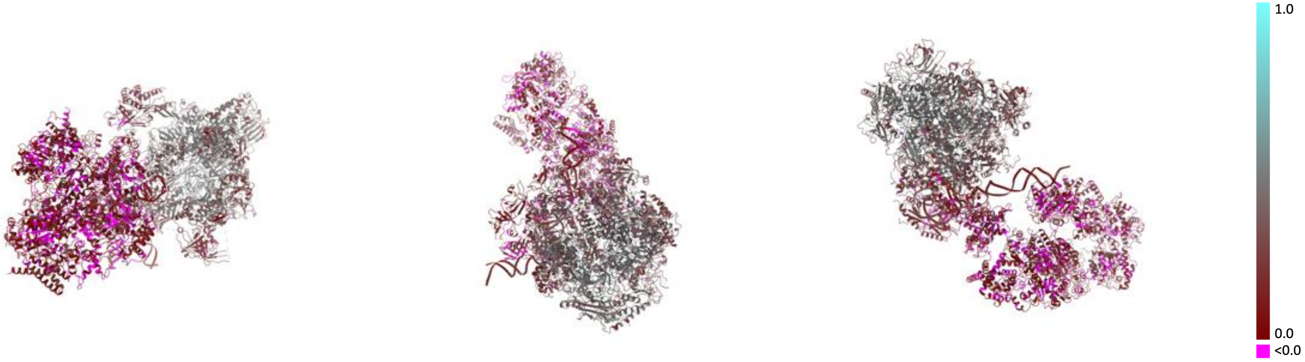
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.018 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

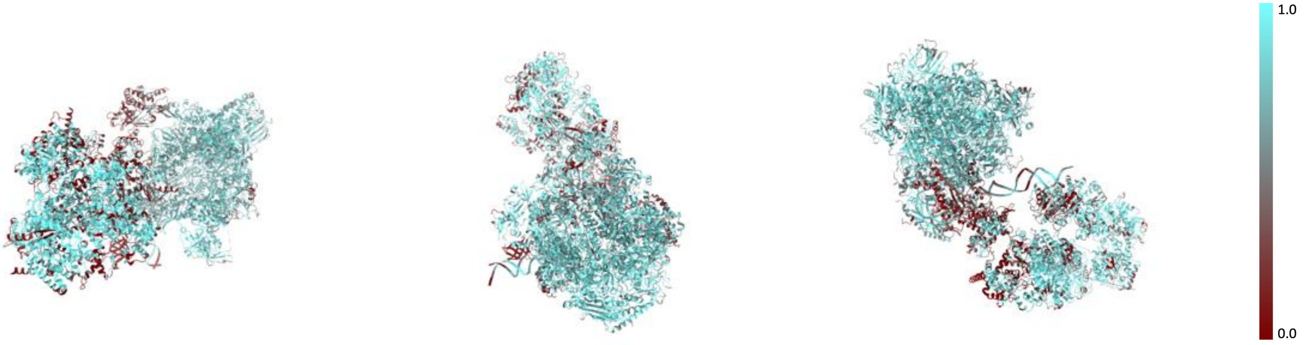
9.2 Q-score mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

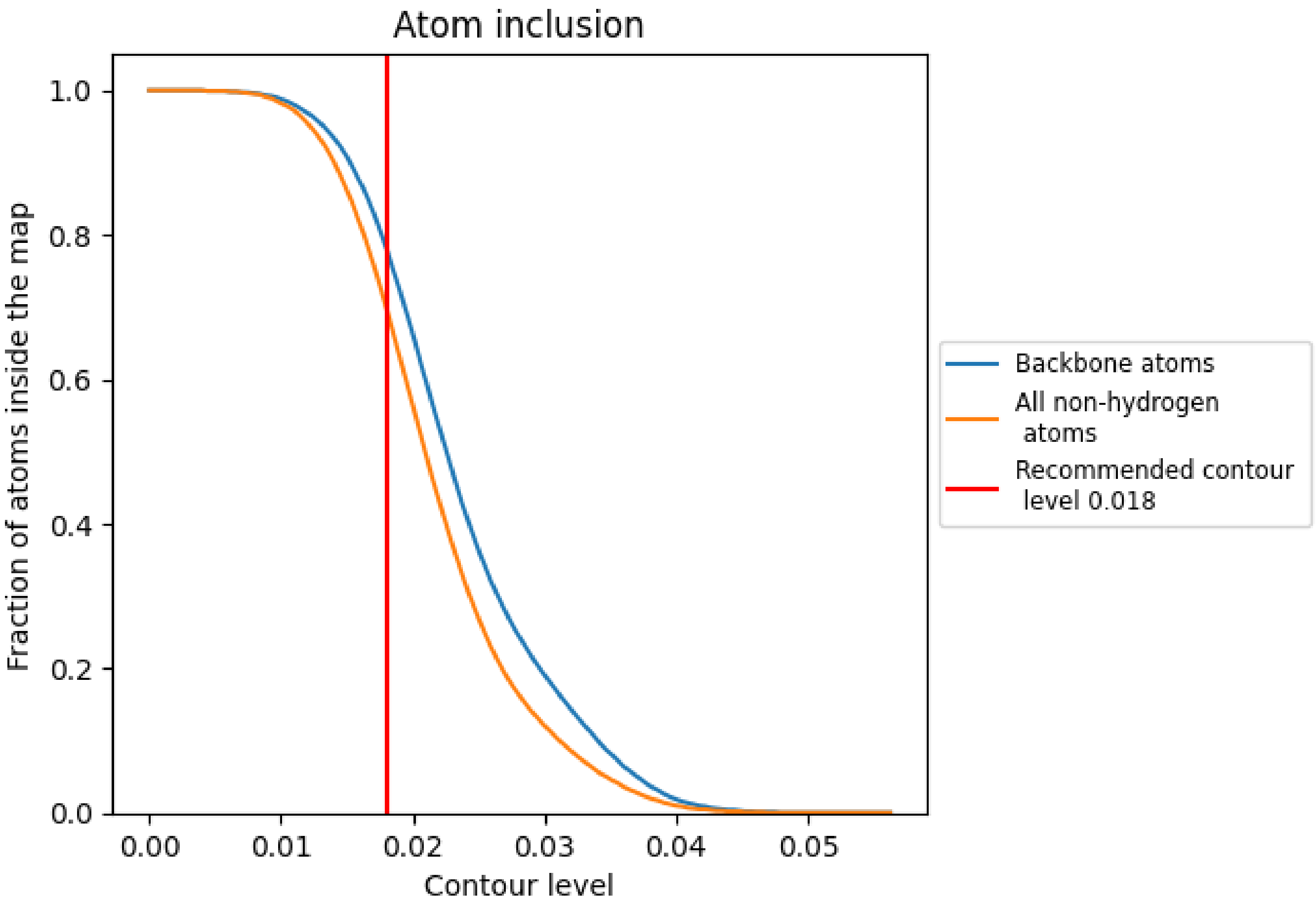
9.3 Atom inclusion mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.018).

9.4 Atom inclusion ⓘ



At the recommended contour level, 78% of all backbone atoms, 70% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.018) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6980	0.2410
0	0.6760	0.0760
1	0.6300	0.0700
2	0.6610	0.0730
3	0.3780	0.0870
4	0.8350	0.0840
5	0.6360	0.0880
6	0.8020	0.0860
7	0.5850	0.0670
A	0.7980	0.4070
B	0.8270	0.4220
C	0.8410	0.4250
D	0.2470	0.2040
E	0.7770	0.3820
F	0.7930	0.3810
G	0.4360	0.2550
H	0.7370	0.3440
I	0.6100	0.2690
J	0.8380	0.4100
K	0.7150	0.3710
L	0.6970	0.2880
M	0.5480	0.2340
N	0.6980	0.1450
O	0.7610	0.1580
Q	0.8080	0.2110
R	0.6770	0.1820
T	0.7070	0.1650
U	0.0290	0.0510
V	0.1140	0.0470
W	0.4040	0.0900
X	0.5360	0.0920