wwPDB EM Validation Summary Report (i)

Nov 20, 2022 – 12:17 PM EST

PDB ID	:	7ML0
EMDB ID	:	EMD-23904
Title	:	RNA polymerase II pre-initiation complex (PIC1)
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	:	3.00 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 (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:

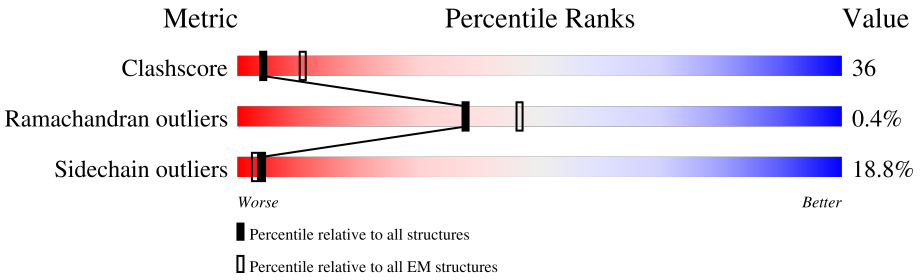
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.00 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	42% 32% 7% 19%
2	B	1224	54% 33% 6% 6%
3	C	318	51% 26% 5% 18%
4	D	221	23% 39% 9% 29%
5	E	215	42% 48% 9%
6	F	155	32% 18% 46%
7	G	171	28% 60% 12%
8	H	146	40% 43% 8% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	122	30% 52% 14% 5%
10	J	70	54% 31% 6% 7%
11	K	120	5% 59% 30% 7%
12	L	70	14% 36% 14% 36%
13	M	345	25% 32% 40% 9% 19%
14	Q	735	9% 9% 80%
15	R	400	26% 18% 52%
16	W	482	14% 20% 5% 60%
17	X	328	12% 30% 14% 52%
18	O	240	29% 28% 43% 25%
19	T	66	21% 45% 55%
20	N	66	17% 52% 48%
21	0	778	22% 20% 58% 18%
22	1	541	13% 28% 34% 5% 32%
23	4	338	5% 31% 41% 11% 16%
24	6	461	5% 28% 37% 11% 24%
25	2	513	11% 35% 43% 12% 10%
26	5	72	11% 6% 68% 18% 8%
27	3	321	19% 17% 24% 57%
28	7	843	28% 22% 42% 12% 24%

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
31	SF4	0	801	-	-	X	-

2 Entry composition (i)

There are 31 unique types of molecules in this entry. The entry contains 64214 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	10996	6931	1926	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	1147	9132	5775	1602	1700	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	C	N	O	S	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					AltConf	Trace
8	H	136	Total	C	N	O	S	0	0
			1089	686	184	215	4		

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- Molecule 16 is a protein called Transcription initiation factor IIE subunit alpha.

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

- Molecule 17 is a protein called Transcription initiation factor IIE subunit beta.

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

- Molecule 18 is a protein called TATA-box-binding protein.

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

- Molecule 19 is a DNA chain called template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	T	66	Total	C	N	O	P	0	0
			1345	649	230	400	66		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
20	N	66	Total	C	N	O	P	0	0
			1361	653	250	392	66		

- Molecule 21 is a protein called General transcription and DNA repair factor IIH helicase subunit XPD.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	0	753	Total	C	N	O	S	0	0
			6099	3886	1031	1144	38		

- Molecule 22 is a protein called DNA-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	1	367	Total	C	N	O	S	0	0
			2415	1538	439	431	7		

- Molecule 23 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	4	284	Total	C	N	O	S	0	0
			2041	1310	343	376	12		

- Molecule 24 is a protein called General transcription and DNA repair factor IIH subunit SSL1.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	6	351	Total	C	N	O	S	0	0
			2527	1590	454	456	27		

- Molecule 25 is a protein called RNA polymerase II transcription factor B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	2	461	Total	C	N	O	S	0	0
			3020	1862	564	585	9		

- Molecule 26 is a protein called General transcription and DNA repair factor IIH subunit TFB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	5	66	Total	C	N	O	S	0	0
			498	314	89	93	2		

- Molecule 27 is a protein called BJ4_G0050160.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	3	137	Total	C	N	O	S	0	0
			890	552	164	166	8		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
28	7	638	Total 4478	C 2739	N 832	O 883	S 24	0	0

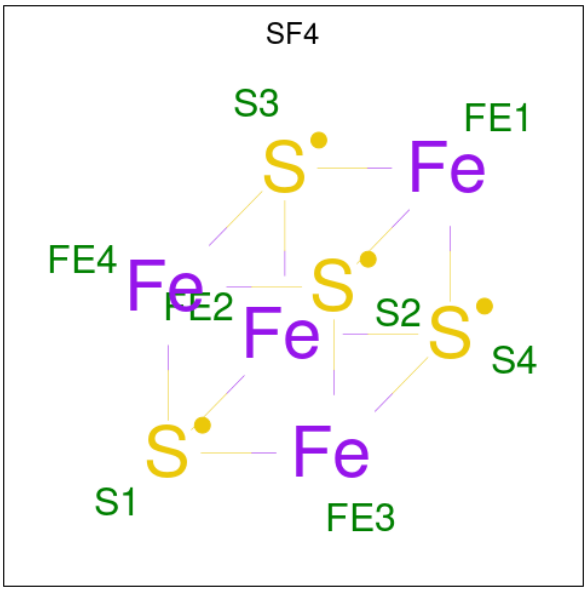
- Molecule 29 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
29	A	2	Total 2	Zn 2	0
29	B	1	Total 1	Zn 1	0
29	C	1	Total 1	Zn 1	0
29	I	2	Total 2	Zn 2	0
29	J	1	Total 1	Zn 1	0
29	L	1	Total 1	Zn 1	0
29	M	1	Total 1	Zn 1	0
29	W	1	Total 1	Zn 1	0
29	4	1	Total 1	Zn 1	0
29	6	4	Total 4	Zn 4	0
29	3	2	Total 2	Zn 2	0

- Molecule 30 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 31 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



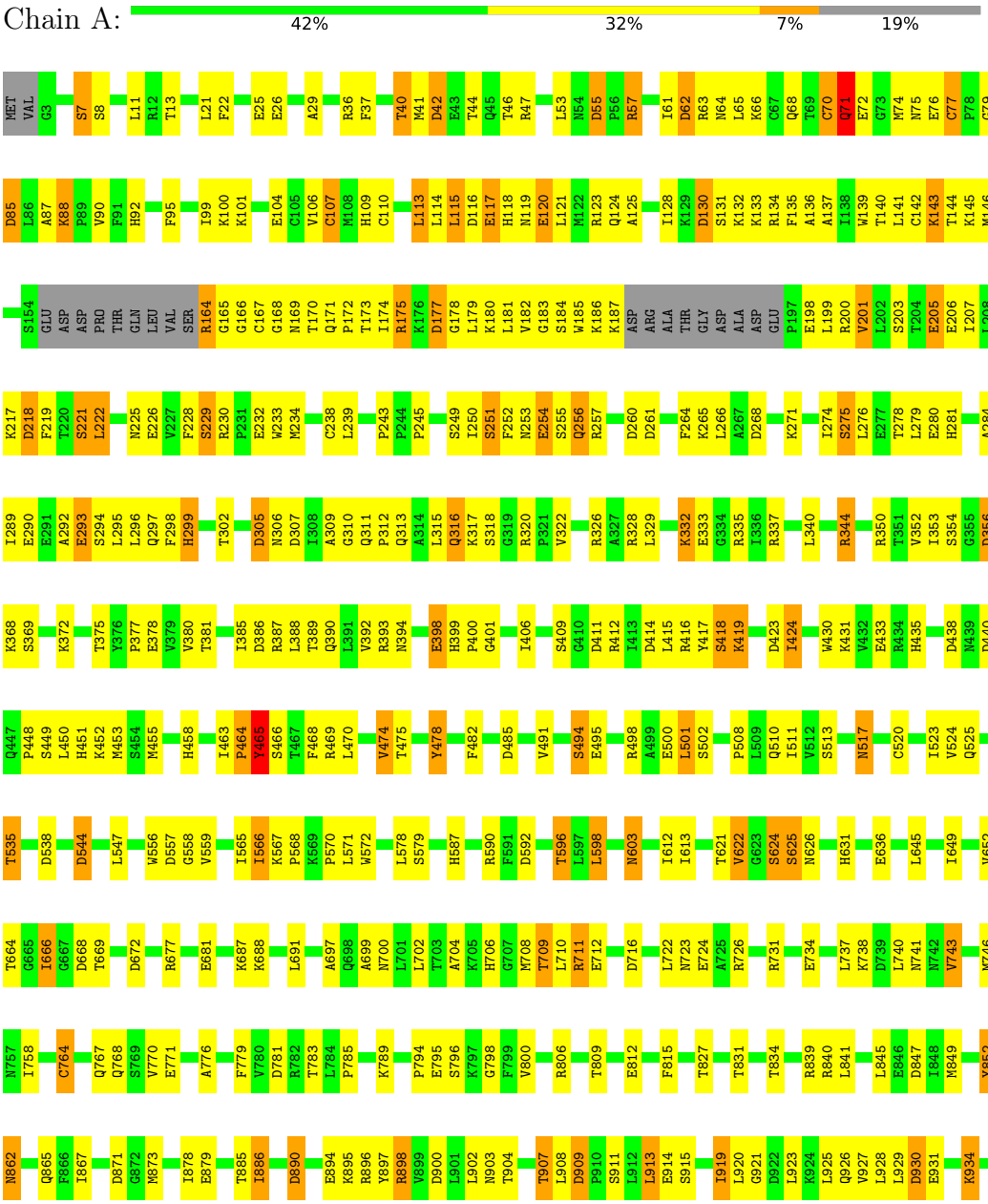
Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
31	0	1	8	4	4	0

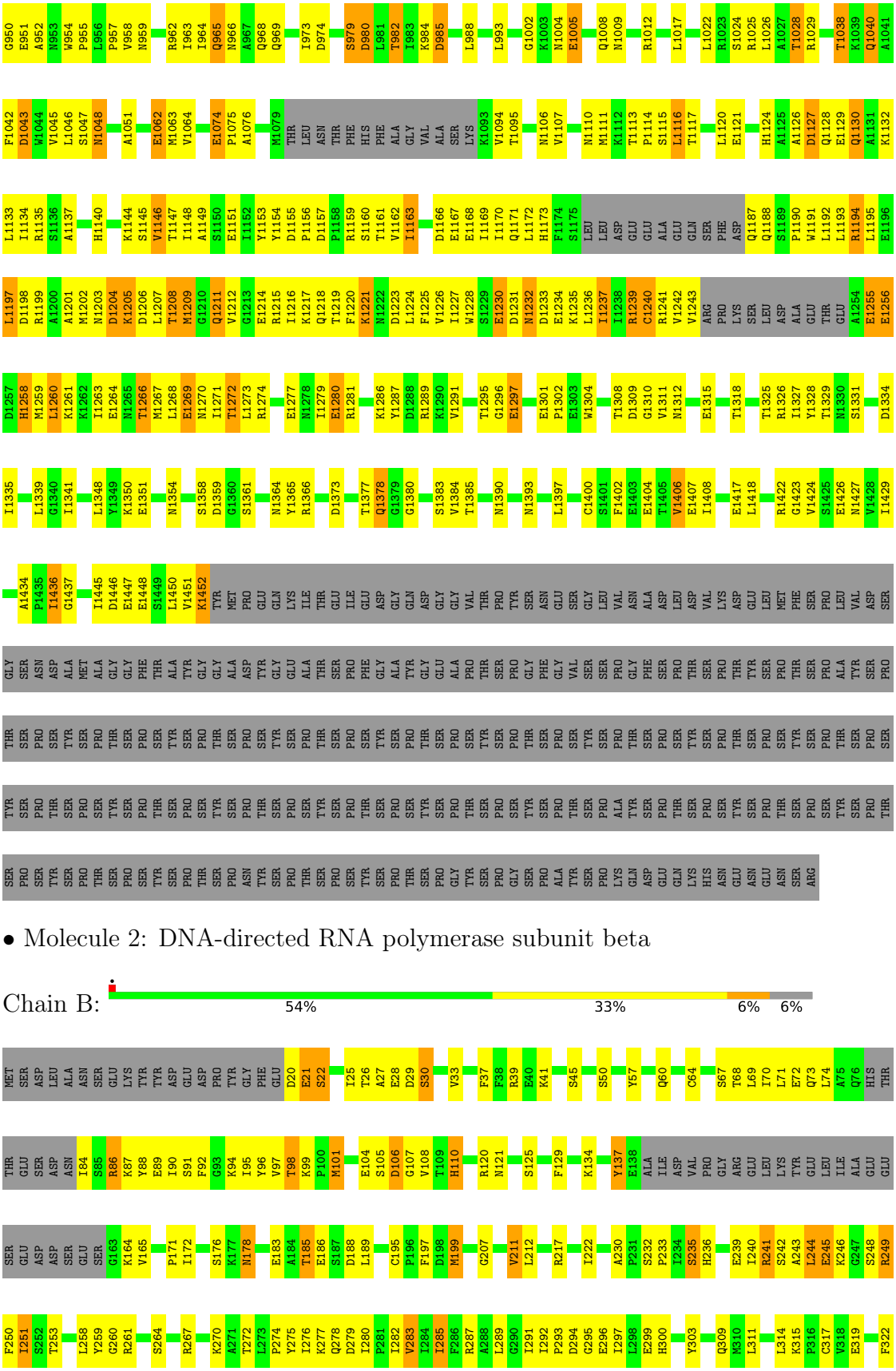
3 Residue-property plots (i)

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 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

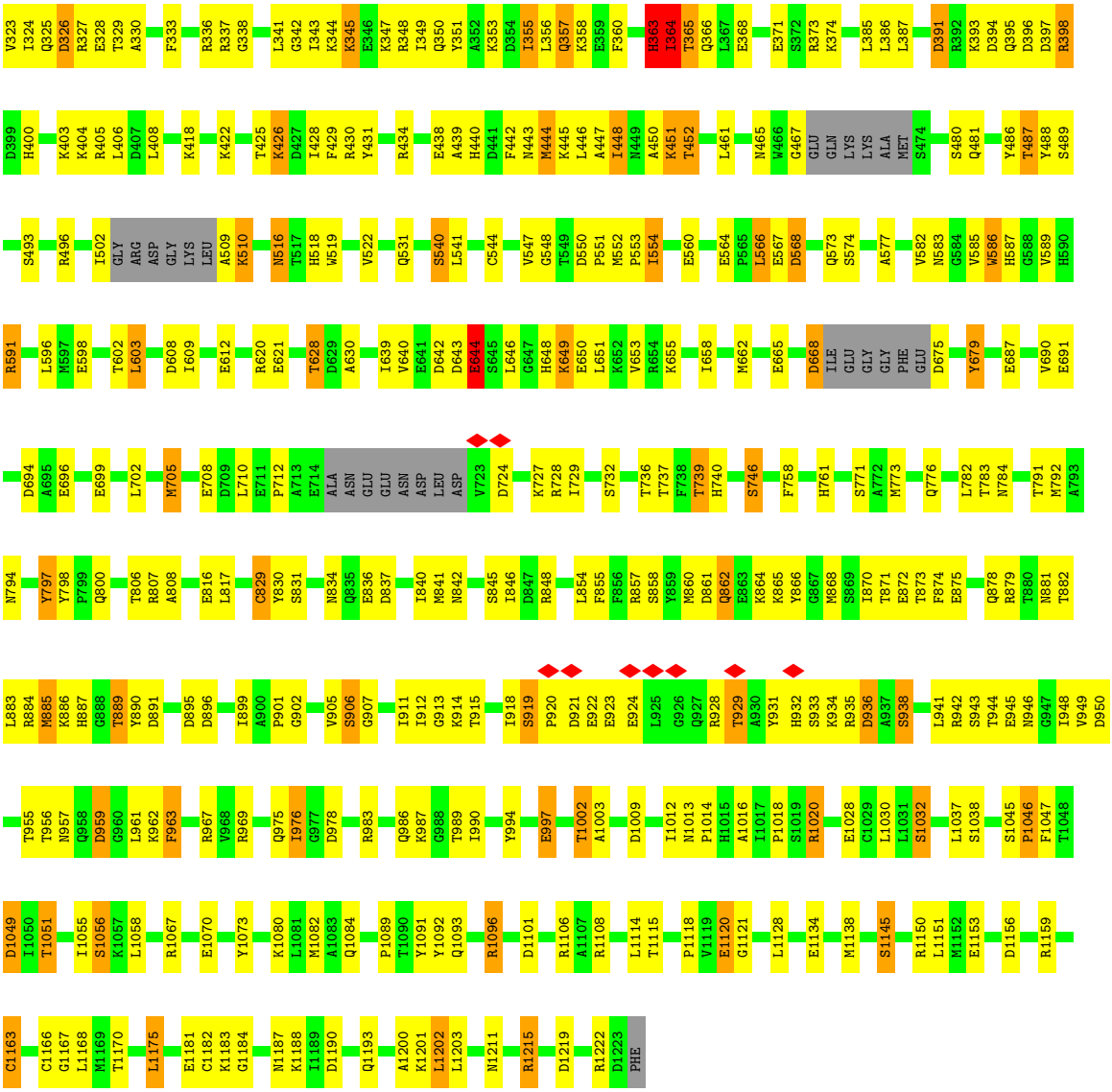
Chain A:





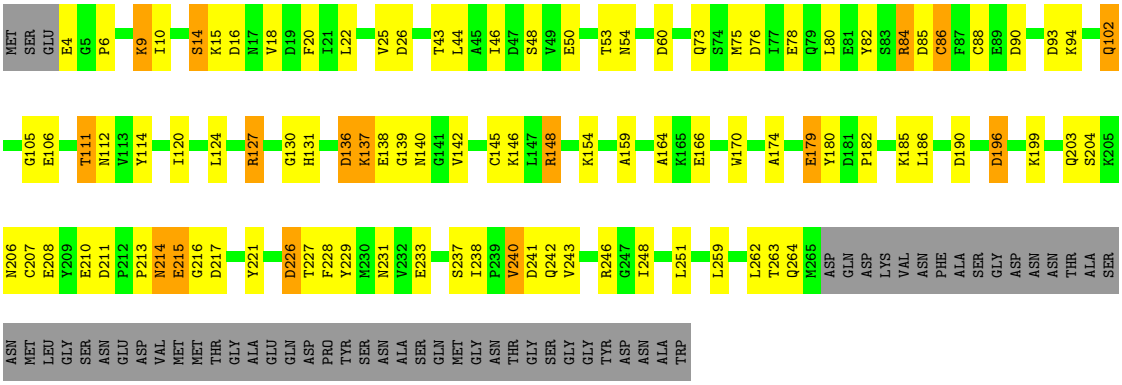
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain B:



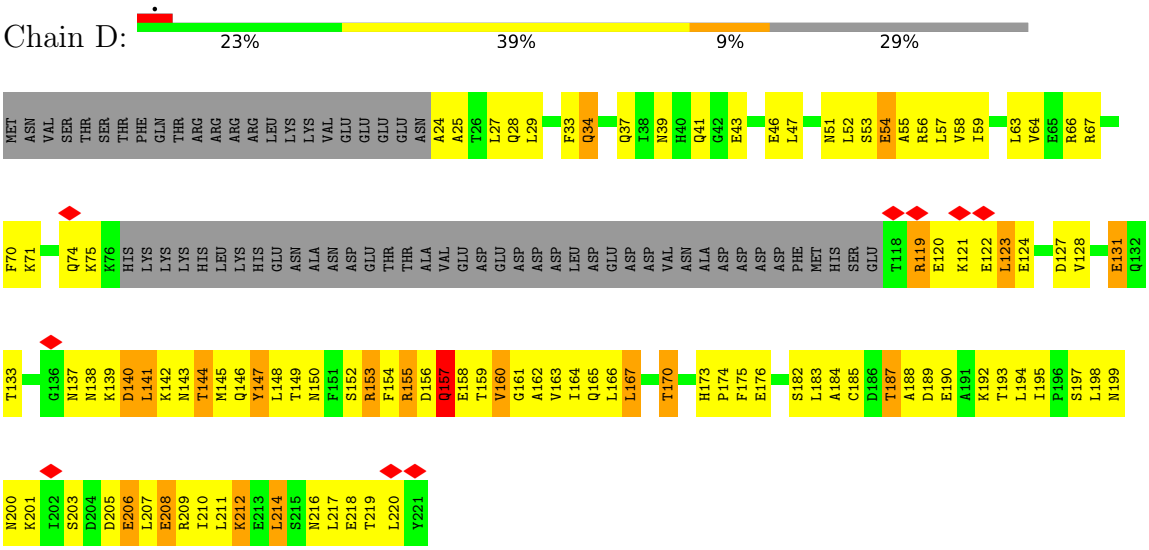
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C:



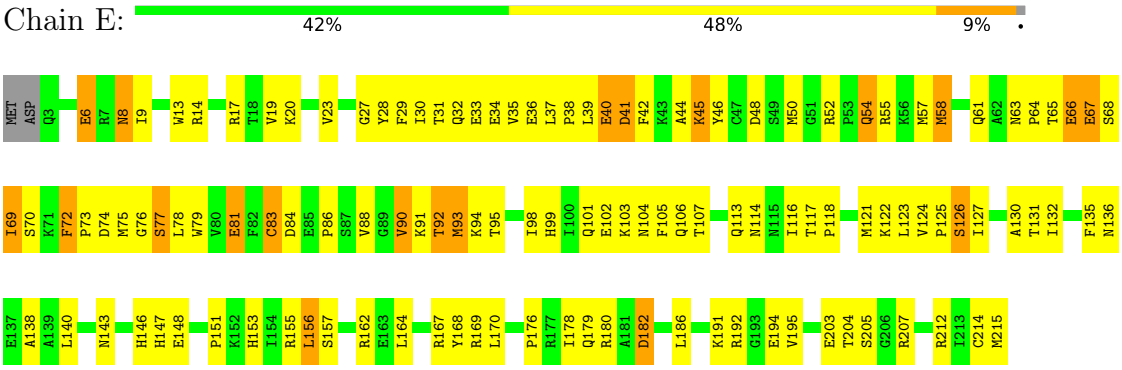
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D:



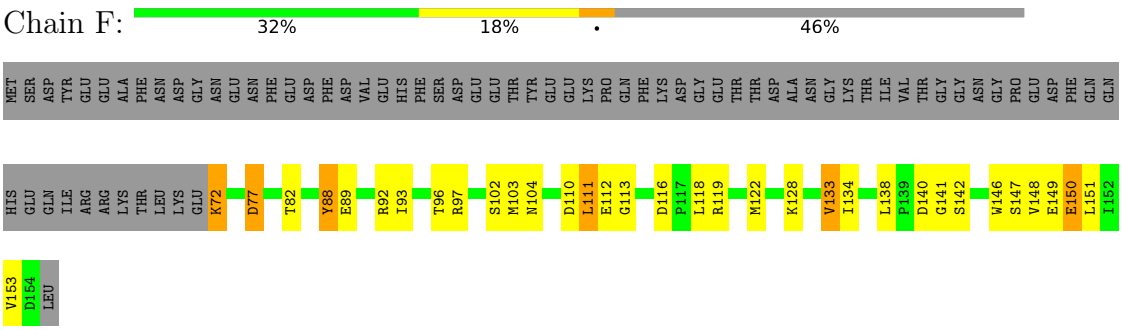
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



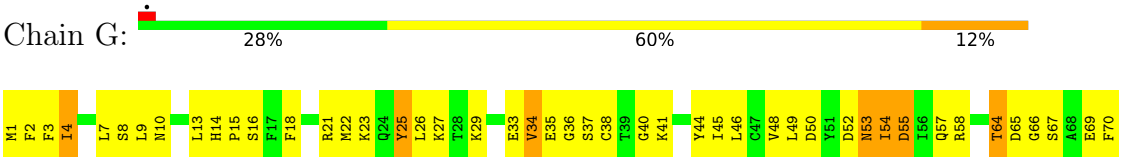
• Molecule 6: DNA-directed RNA polymerases I,II,and III subunit RPABC2

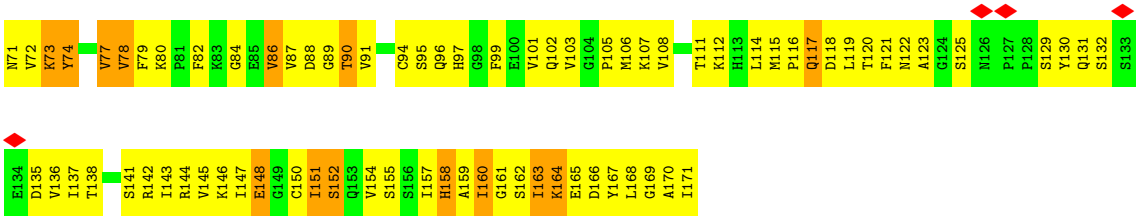
Chain F:



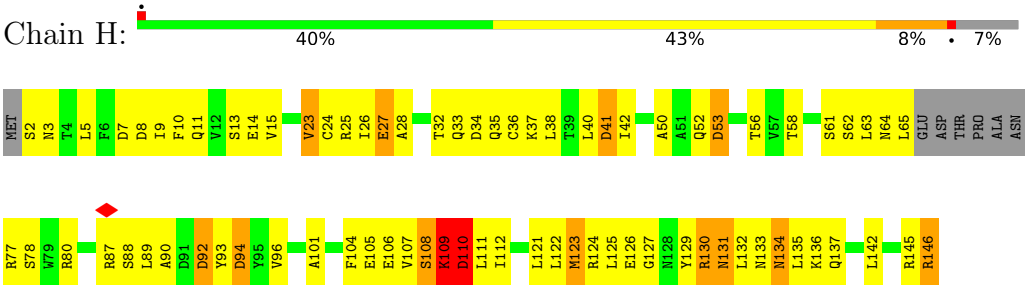
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G:

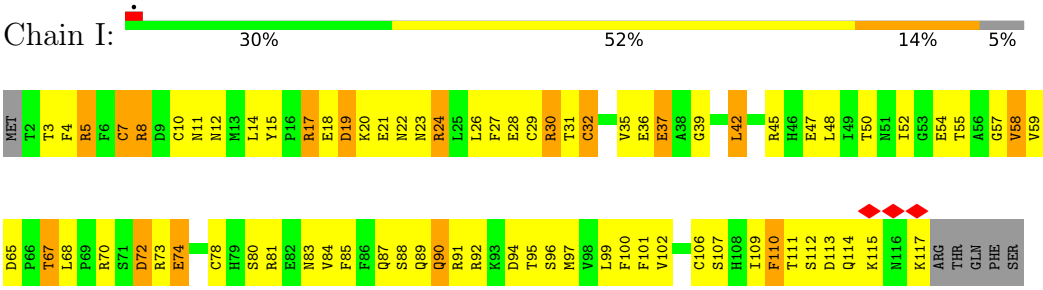




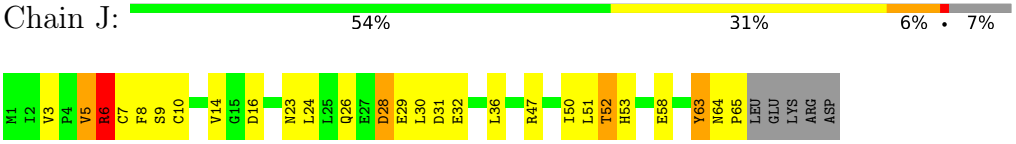
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



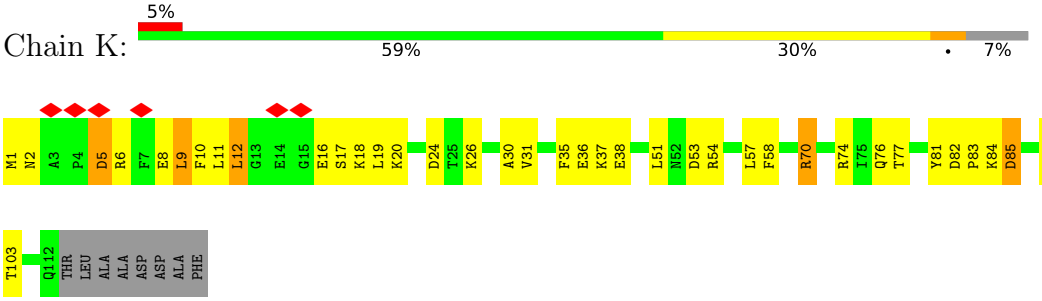
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9



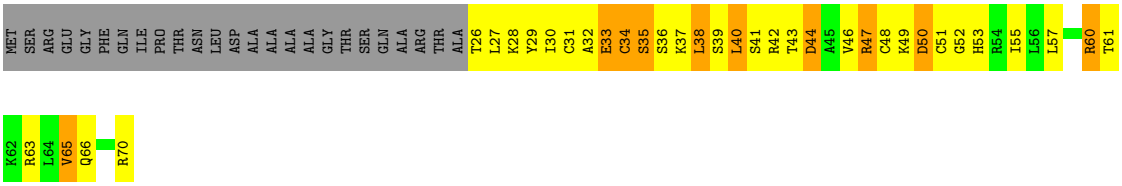
• Molecule 10: DNA-directed RNA polymerases II subunit RPABC5



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

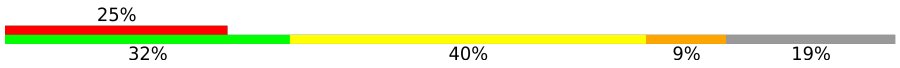


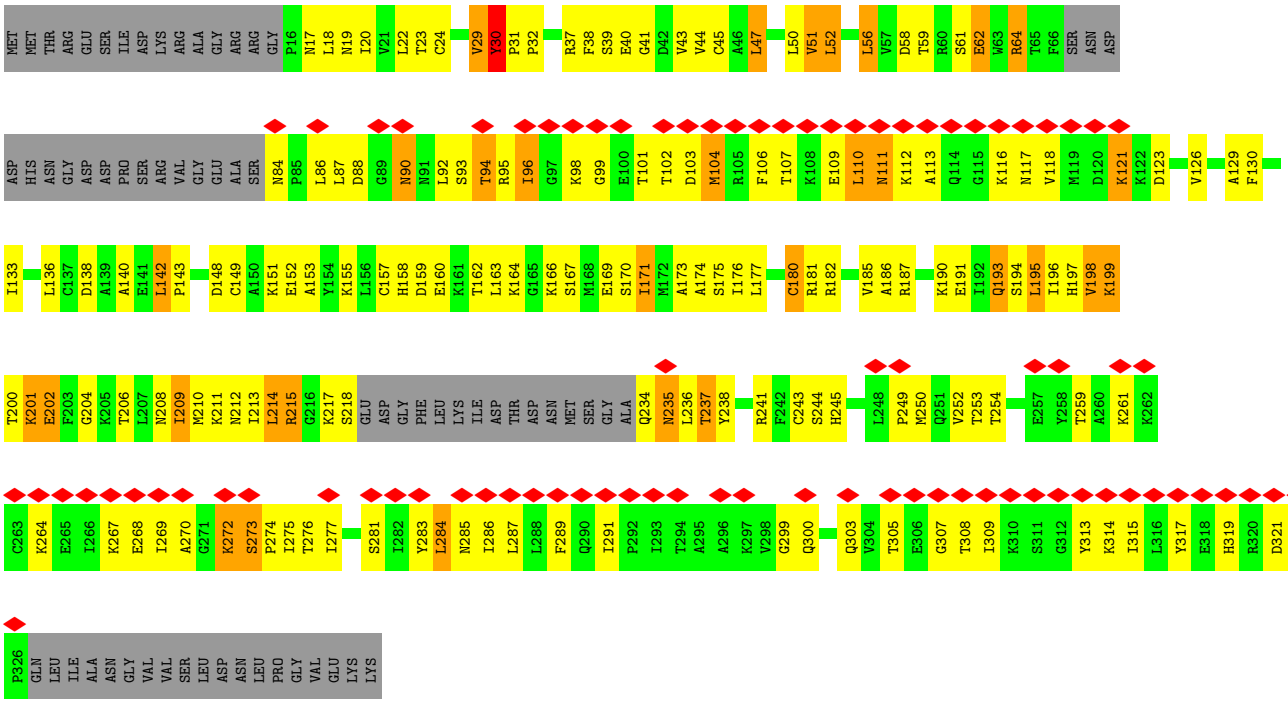
• Molecule 12: DNA-directed RNA polymerases II subunit RPABC4



• Molecule 13: Transcription initiation factor IIB

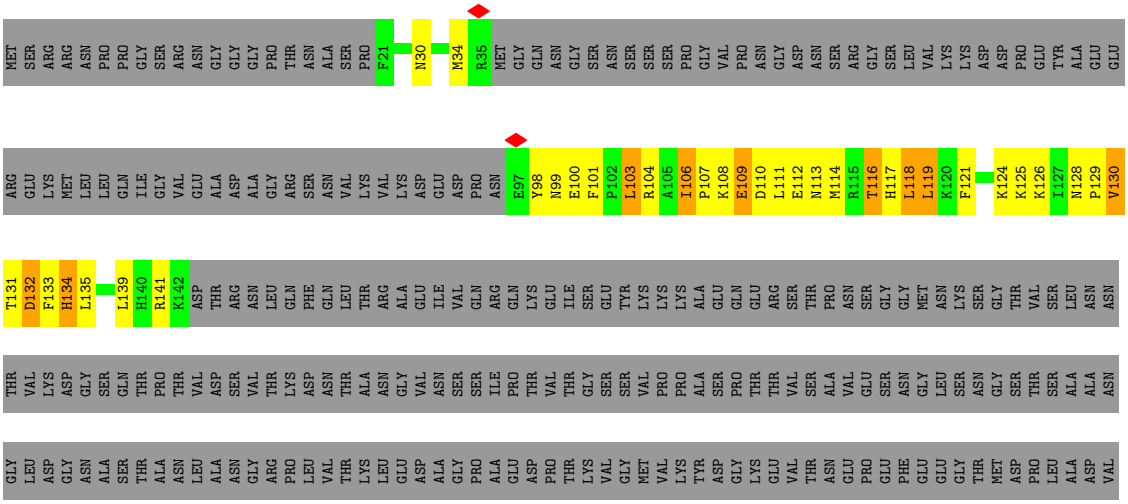
Chain M:

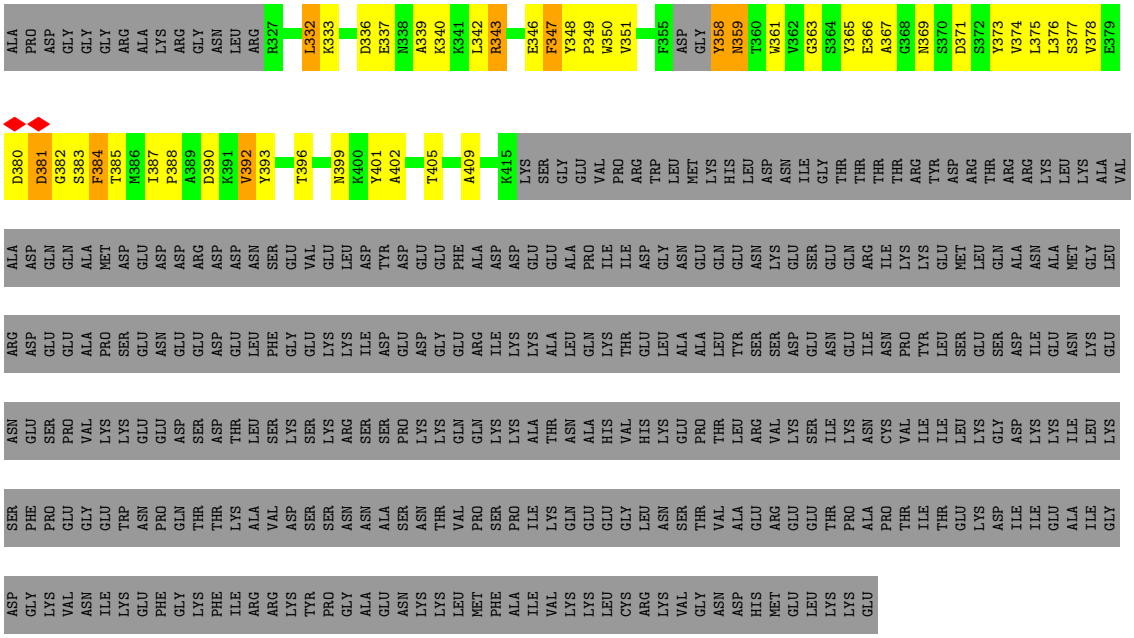




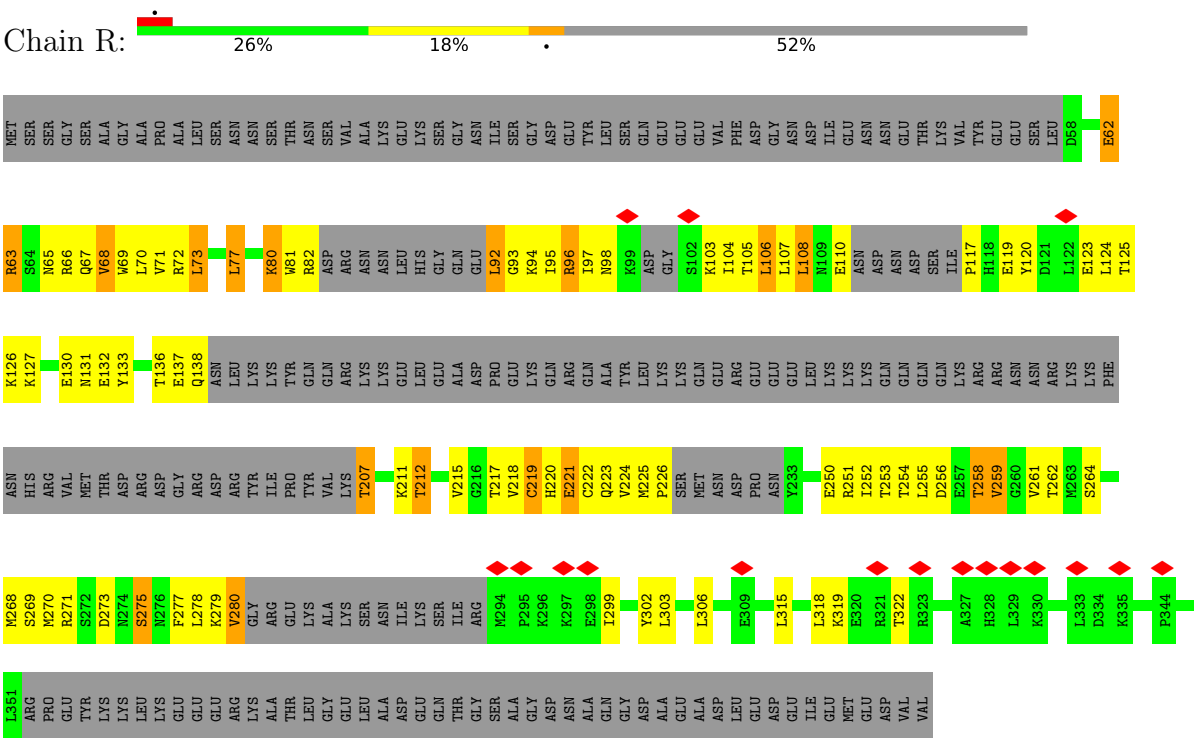
• Molecule 14: Transcription initiation factor IIF subunit alpha

Chain Q:

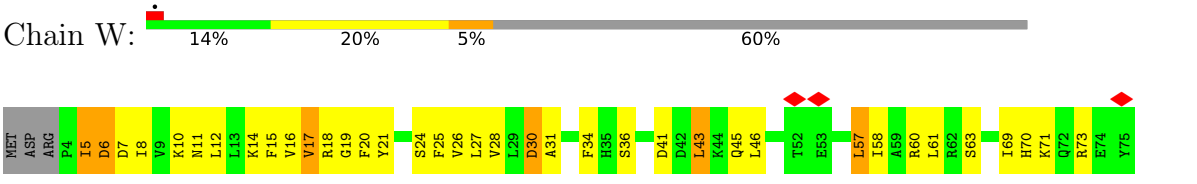


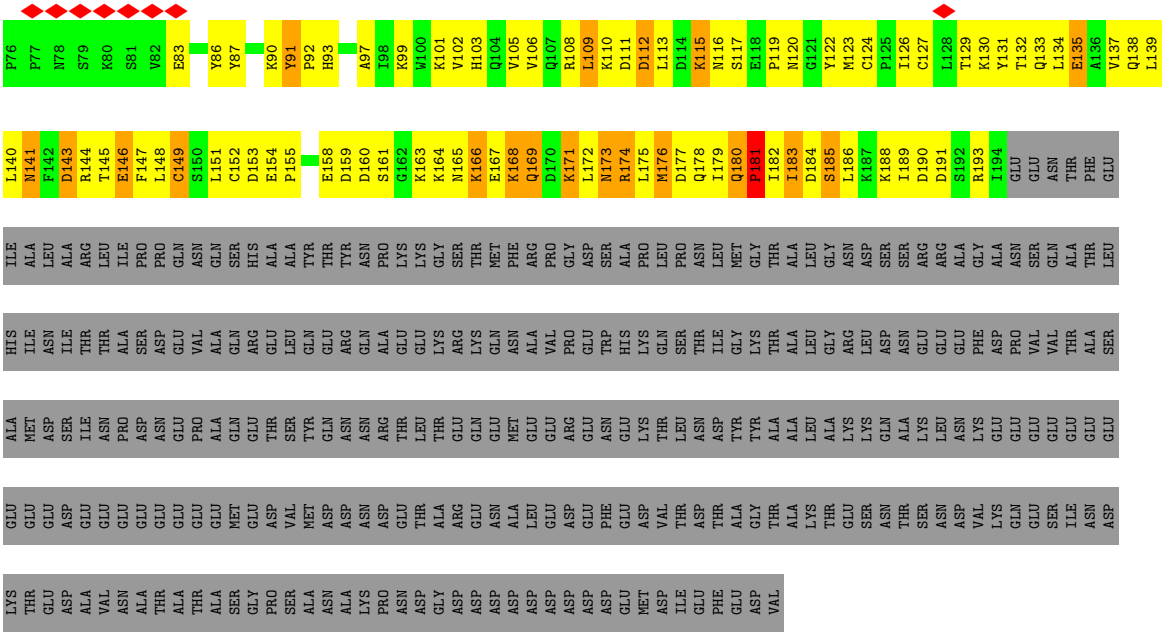


• Molecule 15: Transcription initiation factor IIF subunit beta

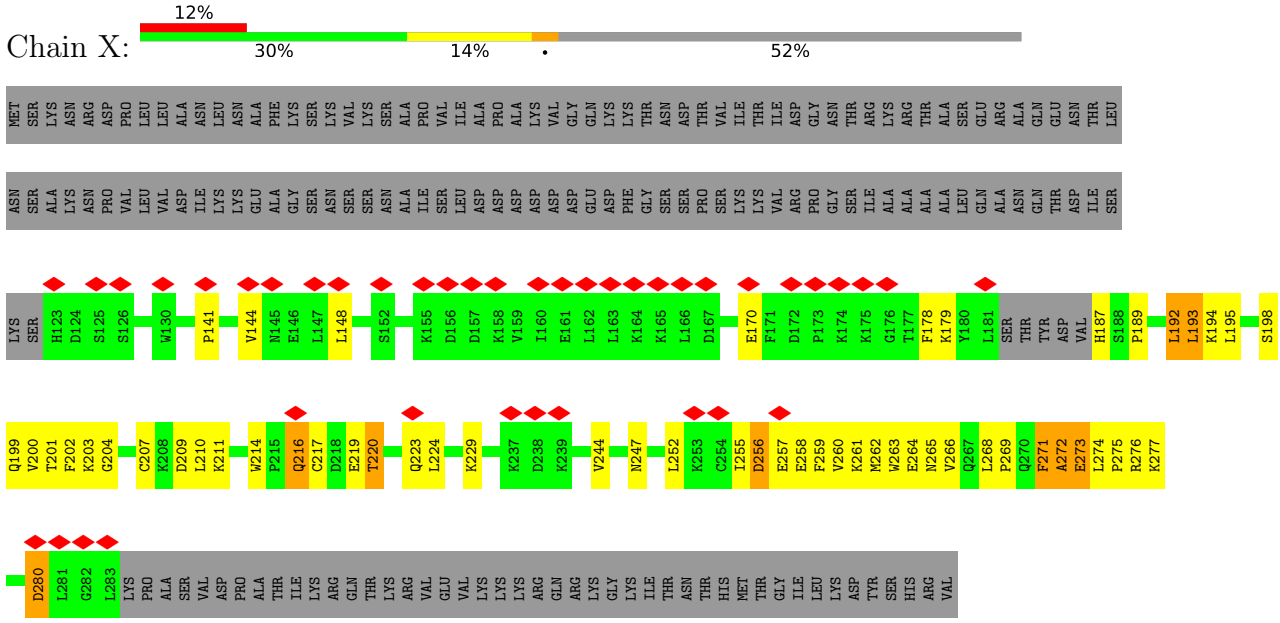


• Molecule 16: Transcription initiation factor IIE subunit alpha

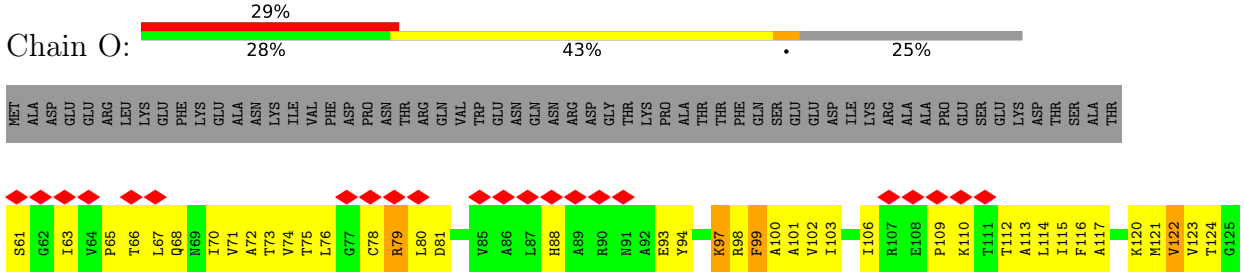


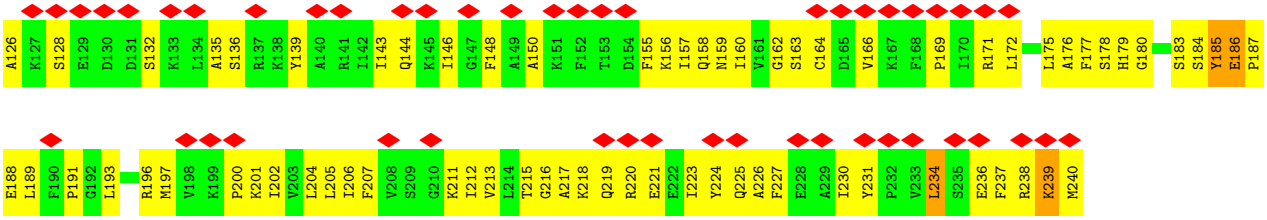


• Molecule 17: Transcription initiation factor IIE subunit beta

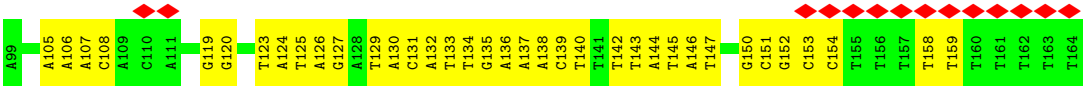


• Molecule 18: TATA-box-binding protein

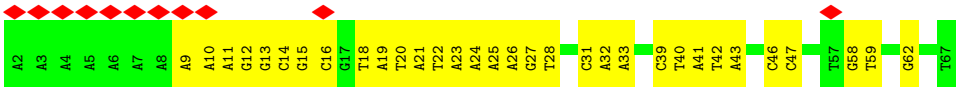




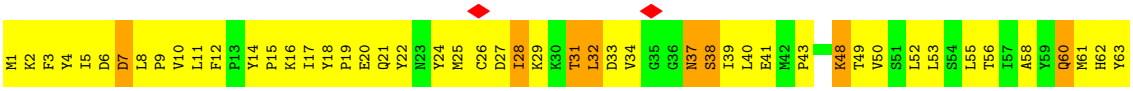
• Molecule 19: template strand DNA

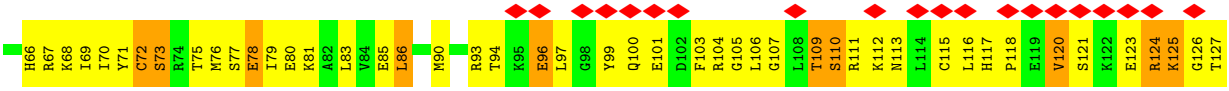


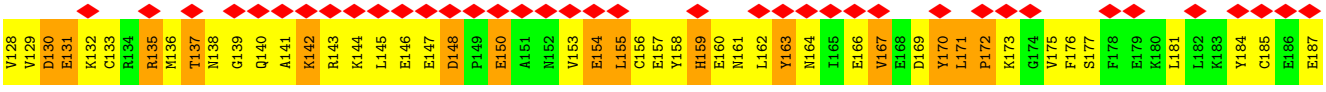
• Molecule 20: non-template strand DNA

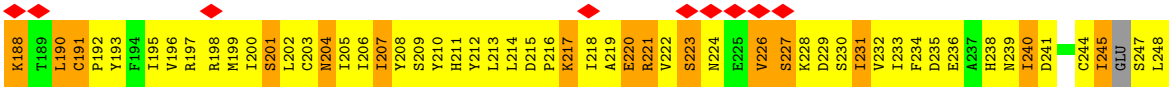


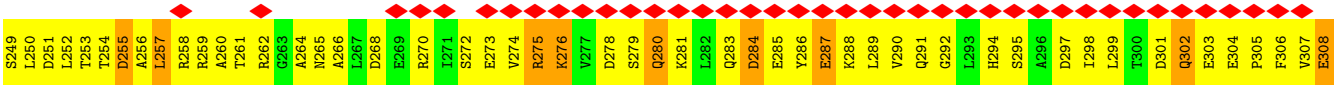
• Molecule 21: General transcription and DNA repair factor IIH helicase subunit XPD

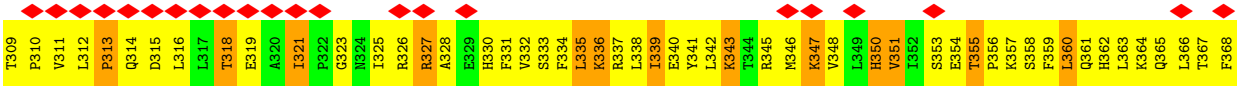


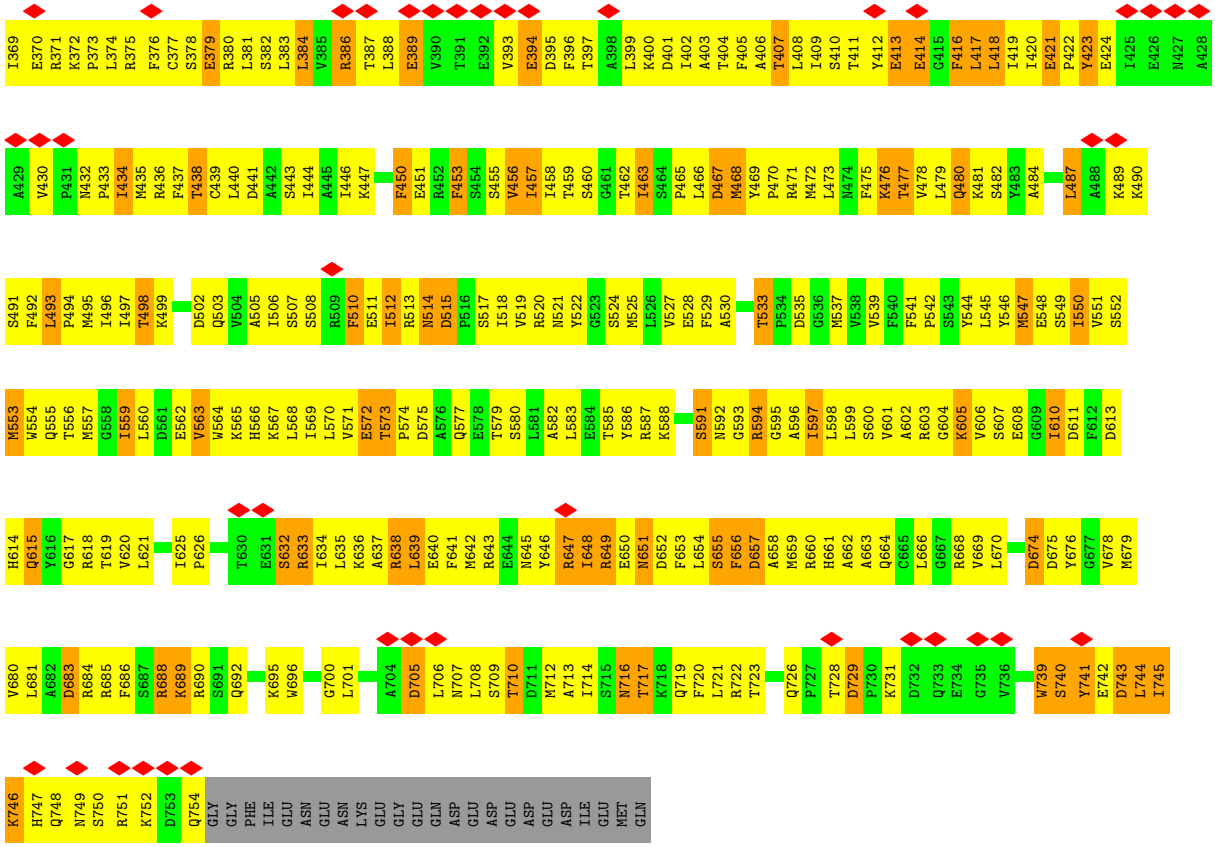




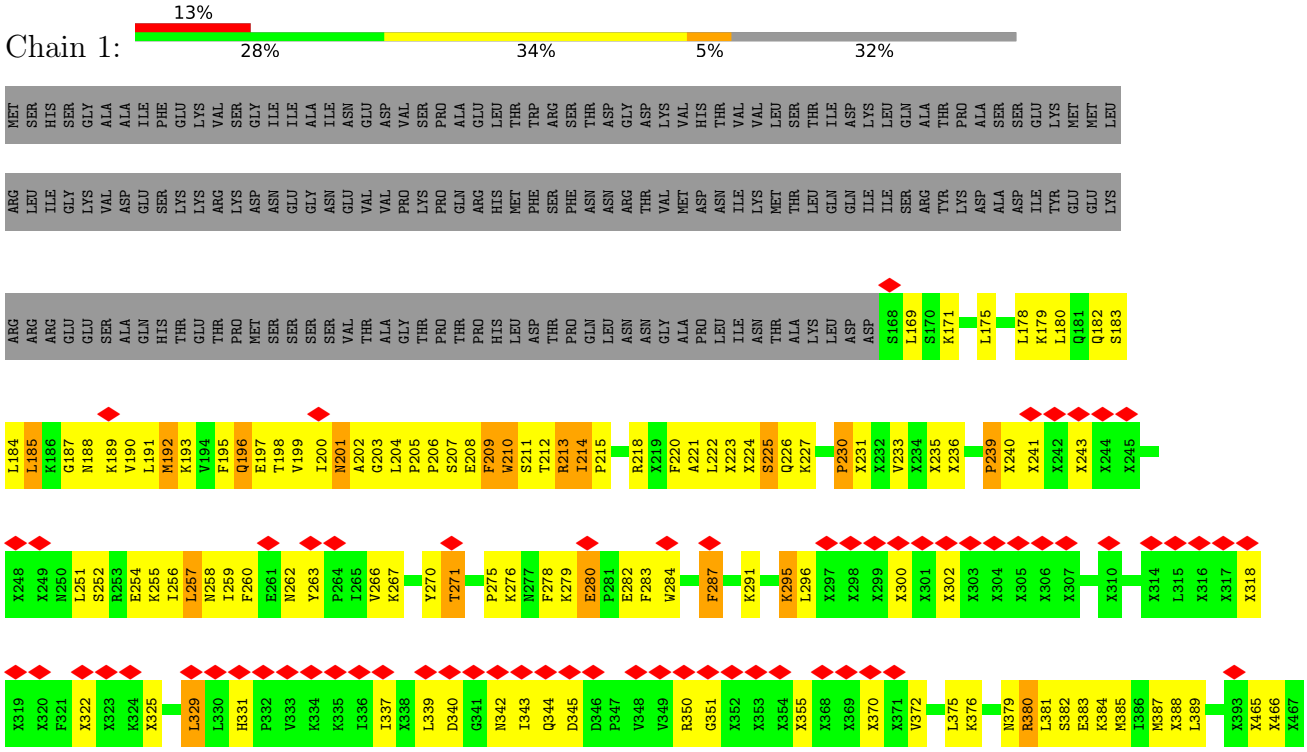


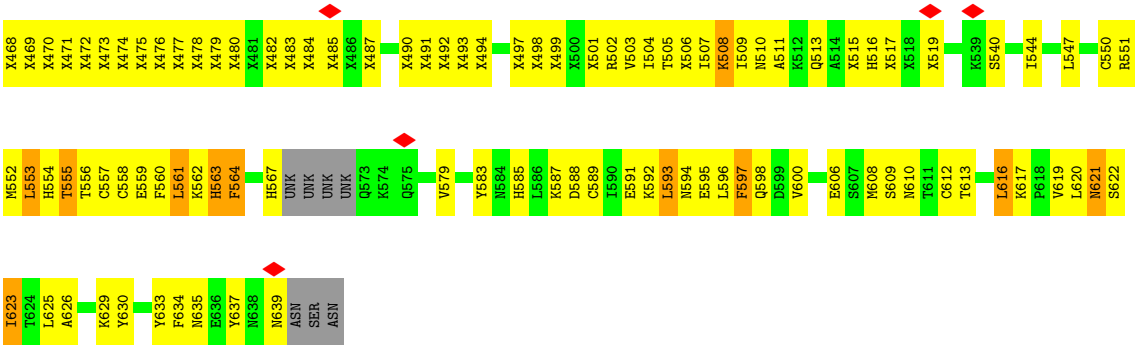






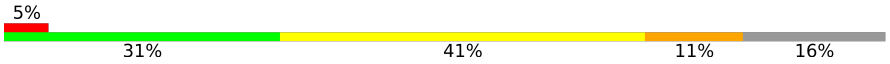
• Molecule 22: DNA-directed RNA polymerase subunit

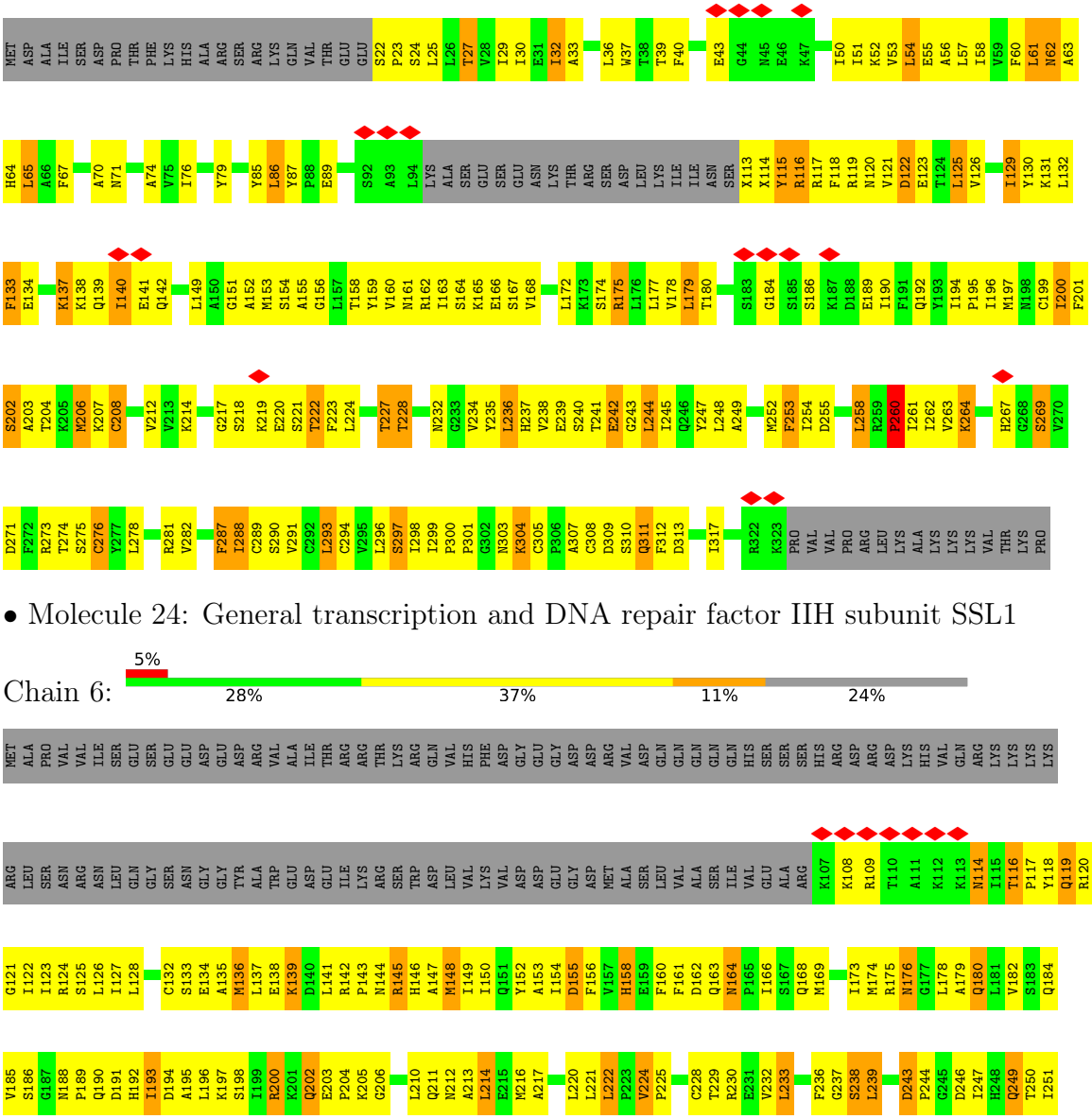




• Molecule 23: DNA-directed RNA polymerase II subunit RPB4

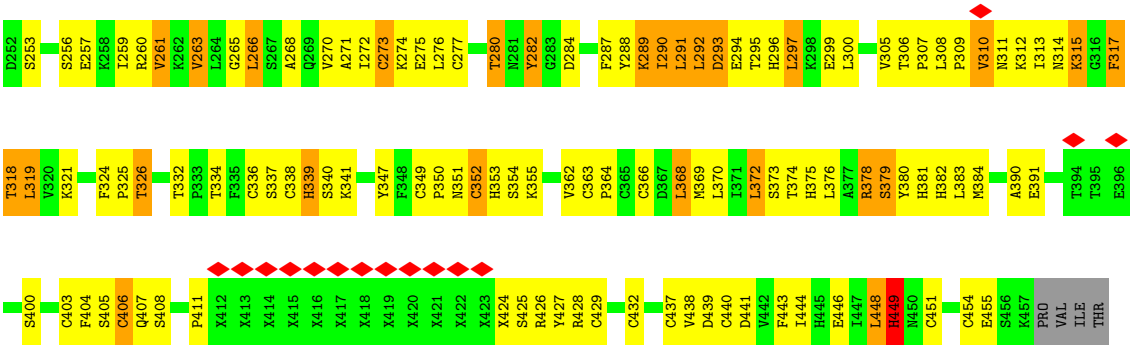
Chain 4:





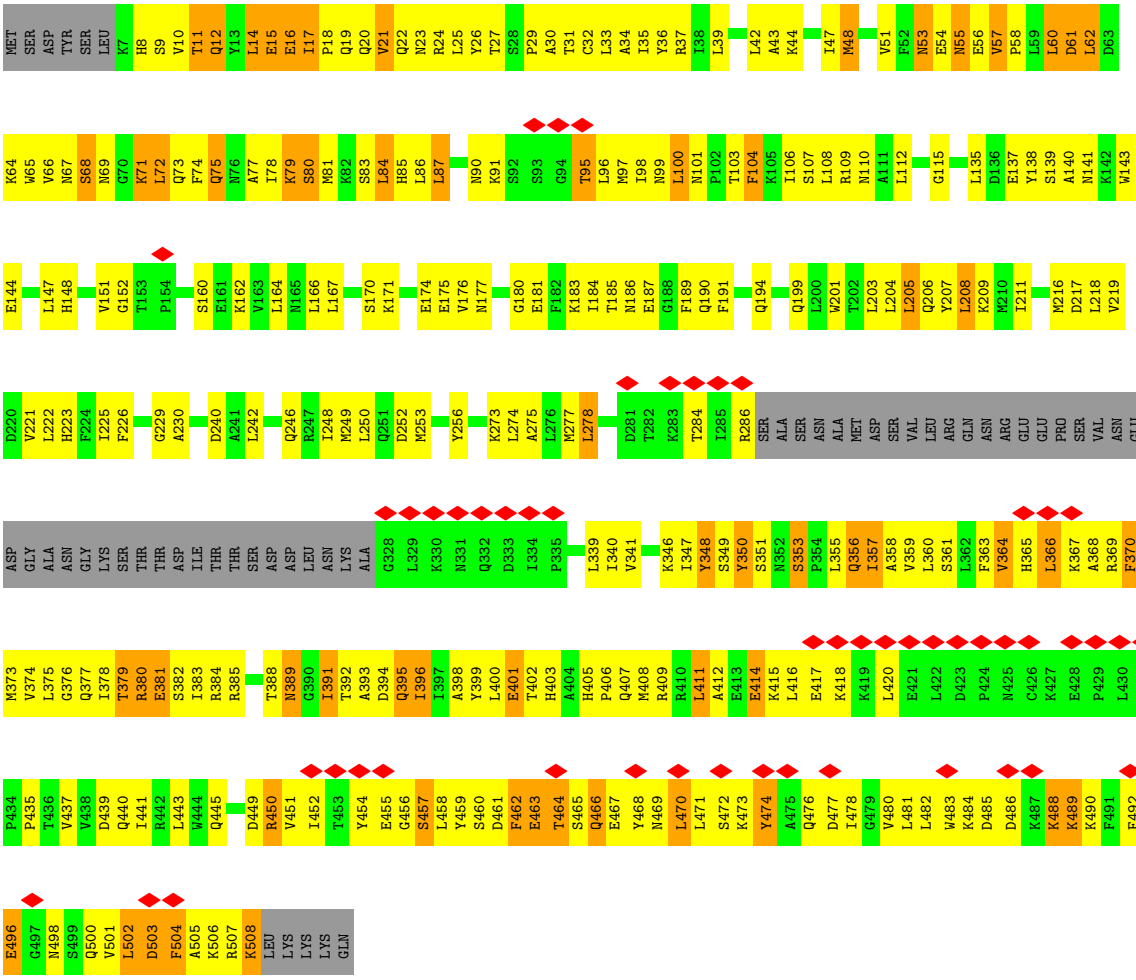
• Molecule 24: General transcription and DNA repair factor IIH subunit SSL1

Chain 6:



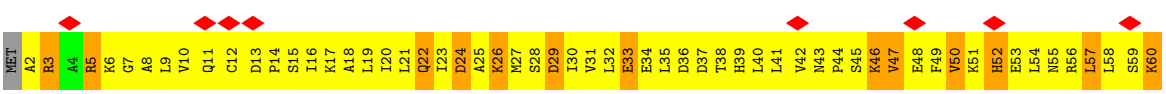
• Molecule 25: RNA polymerase II transcription factor B subunit 2

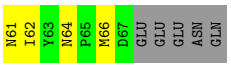
Chain 2:



• Molecule 26: General transcription and DNA repair factor IIH subunit TFB5

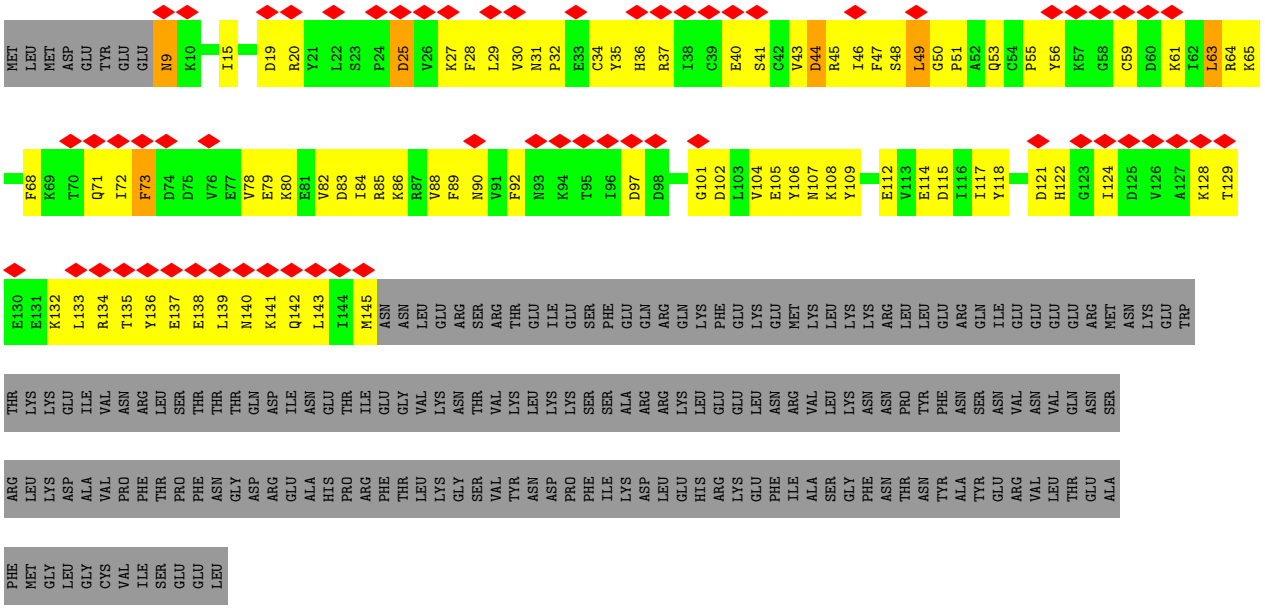
Chain 5:





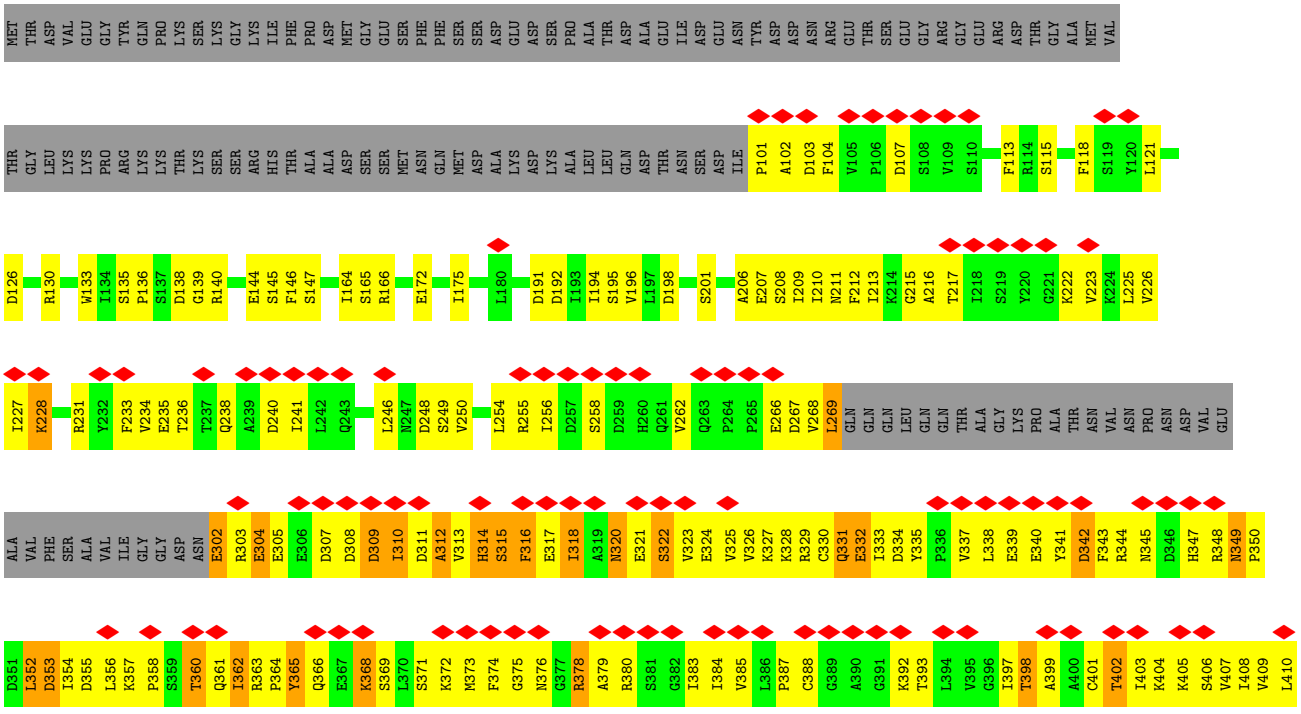
• Molecule 27: BJ4_G0050160.mRNA.1.CDS.1

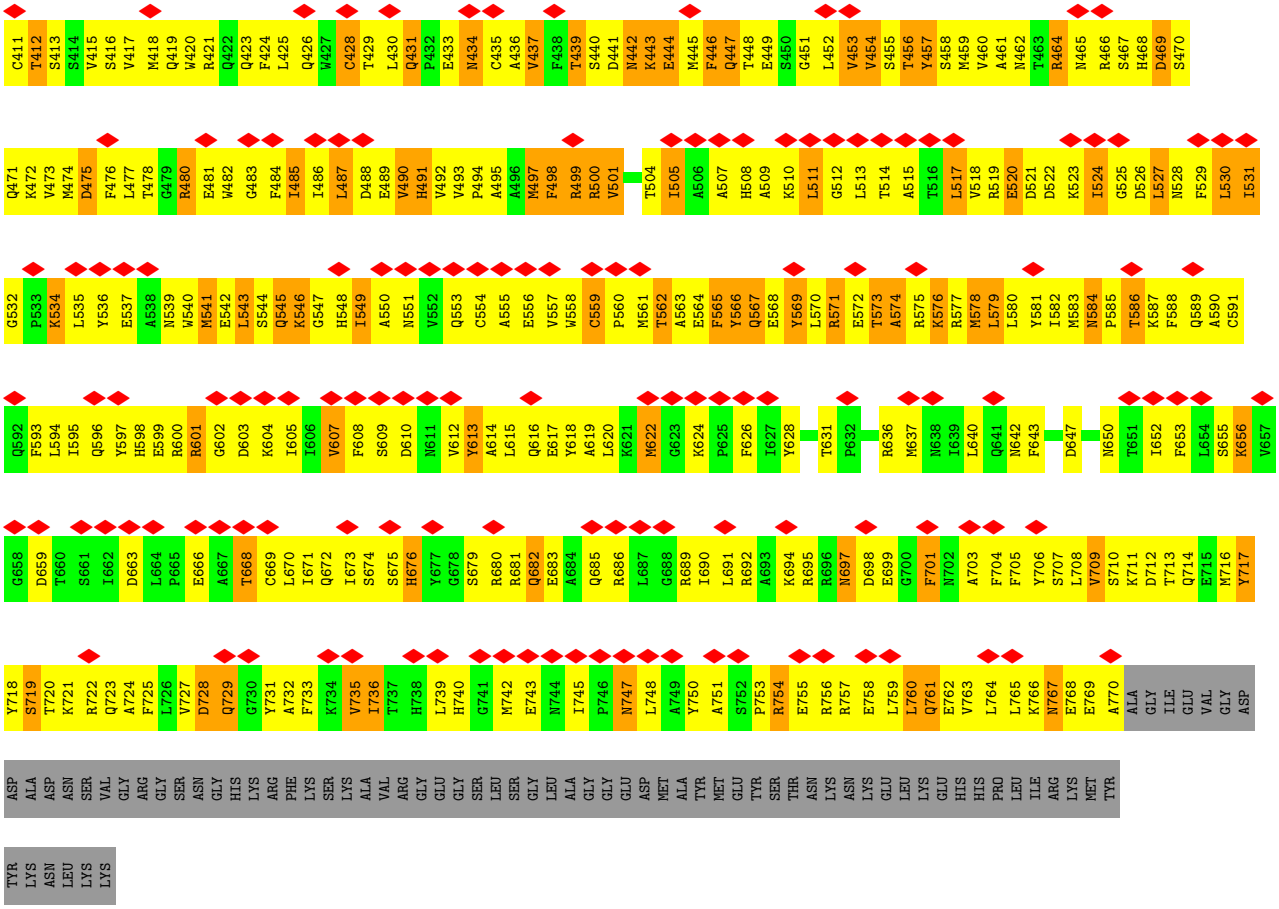
Chain 3:



• Molecule 28: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 7:




C411 T412 S413 S414 V415 S416 V417 M418 Q419 W420 R421 Q422 Q423 F424 L425 Q426 W427 C428 T429 L430 Q431 P432 E433 N434 C435 A436 V437 F438 T439 S440 D441 N442 K443 E444 M445 F446 Q447 T448 E449 S450 G451 L452 V453 V454 S455 T456 Y457 S458 M459 V460 A461 N462 T463 R464 N465 R466 S467 H468 D469 S470
Q471 K472 V473 M474 D475 F476 L477 T478 G479 R480 E481 W482 G483 F484 I485 I486 L487 D488 E489 V490 H491 V492 V493 P494 A495 A496 M497 F498 R499 R500 V501 T504 I505 A506 A507 H508 A509 K510 L511 G512 L513 T514 A515 T516 L517 V518 R519 E520 D521 D522 K523 I524 G525 D526 L527 N528 F529 L530 I531
G532 P533 K534 L535 Y536 E537 A538 N539 W540 M541 E542 L543 S544 Q545 K546 G547 H548 I549 A550 N551 V552 Q553 C554 A555 E556 V557 W558 C559 P560 M561 T562 A563 E564 F565 Y566 Q567 E568 Y569 L570 R571 E572 T573 A574 R575 K576 R577 M578 L579 L580 Y581 I582 M583 N584 P585 T586 K587 F588 Q589 A590 C591
Q592 F593 L594 I595 Q596 Y597 H598 E599 R600 R601 G602 D603 K604 I605 I606 V607 F608 S609 D610 N611 V612 Y613 A614 L615 Q616 E617 Y618 A619 L620 K621 M622 G623 K624 P625 F626 I627 Y628 T631 P632 R636 M637 N638 I639 L640 Q641 N642 F643 D647 N650 T651 I652 F653 L654 S655 K656 V657
G658 D659 T660 S661 I662 D663 L664 P665 E666 A667 T668 C669 L670 I671 Q672 I673 S674 S675 H676 Y677 G678 S679 R680 R681 Q682 E683 A684 Q685 R686 L687 G688 R689 I690 L691 R692 A693 K694 R695 R696 N697 D698 E699 G700 F701 N702 A703 F704 F705 Y706 S707 L708 V709 S710 K711 D712 T713 Q714 E715 M716 Y717
Y718 S719 T720 K721 R722 Q723 A724 F725 L726 V727 D728 Q729 G730 Y731 A732 F733 K734 V735 I736 T737 H738 L739 H740 G741 M742 E743 N744 I745 P746 N747 L748 A749 Y750 A751 S752 P753 R754 E755 R756 R757 E758 L759 L760 Q761 E762 V763 L764 L765 K766 N767 E768 E769 A770 ALA GLY ILE GLU VAL GLY ASP
ASP ALA ASP ASN SER VAL GLY ARG GLY SER ASN GLY HIS LYS ARG PHE LYS SER LYS ALA VAL ARG GLY GLU GLY SER LEU SER GLY LEU ALA GLY GLY GLU ASP MET ALA TYR MET GLU TYR SER THR ASN LYS ASN LYS GLU LEU LYS GLU HIS HIS PRO LEU ILE ARG LYS MET TYR
TYR LYS ASN LEU LYS LYS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	137466	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 ($e^-/Å^2$)	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	0.087	Depositor
Minimum map value	0.000	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.0164	Depositor
Map size (Å)	545.89996, 469.58, 507.73996	wwPDB
Map dimensions	515, 443, 479	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SF4, 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	1.13	12/11191 (0.1%)	0.72	0/15127
2	B	1.25	19/9311 (0.2%)	0.71	0/12558
3	C	1.26	3/2099 (0.1%)	0.74	1/2845 (0.0%)
4	D	0.40	0/1262	0.57	0/1693
5	E	1.01	1/1780 (0.1%)	0.64	0/2395
6	F	1.31	1/682 (0.1%)	0.72	0/922
7	G	0.66	1/1368 (0.1%)	0.59	0/1844
8	H	1.10	0/1107	0.70	0/1499
9	I	0.81	0/962	0.64	0/1295
10	J	1.50	2/541 (0.4%)	0.85	0/727
11	K	1.17	0/922	0.72	0/1244
12	L	0.96	0/360	0.68	0/478
13	M	0.46	0/2204	0.55	0/2963
14	Q	0.52	0/1168	0.54	0/1579
15	R	0.36	0/1312	0.49	0/1777
16	W	0.30	0/1490	0.47	0/2014
17	X	0.28	0/993	0.45	0/1357
18	O	0.31	0/1443	0.47	0/1942
19	T	0.75	0/1505	1.06	0/2319
20	N	0.82	0/1529	0.99	0/2359
21	0	0.43	0/6216	0.53	0/8392
22	1	0.38	0/1906	0.51	0/2558
23	4	0.55	0/2062	0.61	0/2805
24	6	0.52	0/2506	0.60	0/3402
25	2	0.43	0/3066	0.58	0/4082
26	5	0.43	0/502	0.65	0/677
27	3	0.33	0/902	0.44	0/1230
28	7	0.39	0/4552	0.60	0/6078
All	All	0.86	39/64941 (0.1%)	0.66	1/88161 (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	3
2	B	0	2
4	D	0	1
8	H	0	1
13	M	0	2
16	W	0	1
21	0	0	5
23	4	0	2
24	6	0	3
28	7	0	7
All	All	0	27

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1404	GLU	CA-CB	-7.34	1.37	1.53
3	C	170	TRP	CB-CG	-7.18	1.37	1.50
2	B	1091	TYR	CD1-CE1	-6.99	1.28	1.39
2	B	1073	TYR	CD2-CE2	-6.80	1.29	1.39
2	B	1091	TYR	CD2-CE2	-6.59	1.29	1.39

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	130	GLY	C-N-CA	-7.60	102.69	121.70

There are no chirality outliers.

5 of 27 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	465	TYR	Peptide
1	A	70	CYS	Peptide
1	A	71	GLN	Peptide
2	B	363	HIS	Peptide
2	B	644	GLU	Peptide

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	10996	0	11079	508	0
2	B	9132	0	9146	370	0
3	C	2061	0	2029	66	0
4	D	1253	0	1275	122	0
5	E	1744	0	1772	96	0
6	F	670	0	690	23	0
7	G	1340	0	1357	120	0
8	H	1089	0	1062	65	0
9	I	944	0	899	74	0
10	J	532	0	542	19	0
11	K	904	0	911	35	0
12	L	358	0	381	38	0
13	M	2175	0	2283	150	0
14	Q	1144	0	1034	79	0
15	R	1303	0	1110	95	0
16	W	1469	0	1432	123	0
17	X	984	0	722	54	0
18	O	1416	0	1493	105	0
19	T	1345	0	753	55	0
20	N	1361	0	749	48	0
21	0	6099	0	6160	668	0
22	1	2415	0	1887	218	0
23	4	2041	0	1954	180	0
24	6	2527	0	2321	223	0
25	2	3020	0	2613	352	0
26	5	498	0	506	123	0
27	3	890	0	680	58	0
28	7	4478	0	3928	641	0
29	3	2	0	0	0	0
29	4	1	0	0	0	0
29	6	4	0	0	0	0
29	A	2	0	0	0	0
29	B	1	0	0	0	0
29	C	1	0	0	0	0
29	I	2	0	0	0	0
29	J	1	0	0	0	0
29	L	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
29	M	1	0	0	0	0
29	W	1	0	0	0	0
30	A	1	0	0	0	0
31	0	8	0	0	3	0
All	All	64214	0	60768	4422	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 36.

The worst 5 of 4422 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:A:70:CYS:SG	1:A:80:HIS:CE1	2.32	1.23
18:O:71:VAL:HA	18:O:123:VAL:O	1.40	1.17
18:O:106:ILE:O	18:O:110:LYS:HA	1.50	1.11
25:2:457:SER:HA	26:5:6:LYS:HA	1.33	1.07
24:6:403:CYS:SG	24:6:437:CYS:HB2	1.91	1.04

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%)	1206 (87%)	174 (13%)	6 (0%)	34 72
2	B	1133/1224 (93%)	1030 (91%)	100 (9%)	3 (0%)	41 76
3	C	260/318 (82%)	229 (88%)	31 (12%)	0	100 100
4	D	153/221 (69%)	131 (86%)	21 (14%)	1 (1%)	22 60
5	E	211/215 (98%)	190 (90%)	21 (10%)	0	100 100
6	F	81/155 (52%)	75 (93%)	6 (7%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	G	169/171 (99%)	145 (86%)	24 (14%)	0	100	100
8	H	132/146 (90%)	106 (80%)	25 (19%)	1 (1%)	19	57
9	I	114/122 (93%)	85 (75%)	29 (25%)	0	100	100
10	J	63/70 (90%)	56 (89%)	5 (8%)	2 (3%)	4	22
11	K	110/120 (92%)	98 (89%)	12 (11%)	0	100	100
12	L	43/70 (61%)	32 (74%)	11 (26%)	0	100	100
13	M	273/345 (79%)	225 (82%)	48 (18%)	0	100	100
14	Q	140/735 (19%)	117 (84%)	23 (16%)	0	100	100
15	R	176/400 (44%)	150 (85%)	26 (15%)	0	100	100
16	W	189/482 (39%)	170 (90%)	18 (10%)	1 (0%)	29	68
17	X	152/328 (46%)	131 (86%)	19 (12%)	2 (1%)	12	45
18	O	178/240 (74%)	169 (95%)	9 (5%)	0	100	100
21	0	749/778 (96%)	631 (84%)	116 (16%)	2 (0%)	41	76
22	1	257/541 (48%)	223 (87%)	30 (12%)	4 (2%)	9	40
23	4	279/338 (82%)	202 (72%)	76 (27%)	1 (0%)	34	72
24	6	336/461 (73%)	263 (78%)	71 (21%)	2 (1%)	25	64
25	2	457/513 (89%)	345 (76%)	111 (24%)	1 (0%)	47	82
26	5	64/72 (89%)	46 (72%)	18 (28%)	0	100	100
27	3	135/321 (42%)	121 (90%)	14 (10%)	0	100	100
28	7	634/843 (75%)	448 (71%)	184 (29%)	2 (0%)	41	76
All	All	7874/10962 (72%)	6624 (84%)	1222 (16%)	28 (0%)	38	72

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	465	TYR
1	A	525	GLN
2	B	364	ILE
8	H	110	ASP
10	J	6	ARG

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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%)	1025 (84%)	196 (16%)	2 12
2	B	995/1061 (94%)	860 (86%)	135 (14%)	3 17
3	C	230/274 (84%)	201 (87%)	29 (13%)	4 20
4	D	139/200 (70%)	114 (82%)	25 (18%)	1 9
5	E	195/197 (99%)	163 (84%)	32 (16%)	2 11
6	F	73/137 (53%)	62 (85%)	11 (15%)	3 14
7	G	152/152 (100%)	124 (82%)	28 (18%)	1 9
8	H	119/128 (93%)	97 (82%)	22 (18%)	1 8
9	I	110/116 (95%)	84 (76%)	26 (24%)	1 3
10	J	60/65 (92%)	52 (87%)	8 (13%)	4 17
11	K	97/102 (95%)	88 (91%)	9 (9%)	9 33
12	L	40/57 (70%)	27 (68%)	13 (32%)	0 1
13	M	245/299 (82%)	194 (79%)	51 (21%)	1 5
14	Q	109/641 (17%)	85 (78%)	24 (22%)	1 4
15	R	107/363 (30%)	81 (76%)	26 (24%)	0 3
16	W	155/429 (36%)	123 (79%)	32 (21%)	1 6
17	X	62/295 (21%)	53 (86%)	9 (14%)	3 15
18	O	152/205 (74%)	136 (90%)	16 (10%)	7 27
21	0	685/707 (97%)	534 (78%)	151 (22%)	1 4
22	1	170/396 (43%)	135 (79%)	35 (21%)	1 6
23	4	198/298 (66%)	148 (75%)	50 (25%)	0 3
24	6	247/406 (61%)	187 (76%)	60 (24%)	0 3
25	2	259/468 (55%)	188 (73%)	71 (27%)	0 2
26	5	53/66 (80%)	39 (74%)	14 (26%)	0 2
27	3	63/303 (21%)	53 (84%)	10 (16%)	2 12
28	7	417/737 (57%)	303 (73%)	114 (27%)	0 2
All	All	6353/9622 (66%)	5156 (81%)	1197 (19%)	4 8

5 of 1197 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
23	4	304	LYS
28	7	531	ILE
24	6	224	VAL
23	4	297	SER
25	2	417	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
24	6	382	HIS
28	7	471	GLN
25	2	75	GLN
28	7	331	GLN
28	7	545	GLN

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 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	$\#\|Z\|>2$
31	SF4	0	801	21	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
31	SF4	0	801	21	-	-	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
31	0	801	SF4	3	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
22	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	86.50
1	1	519:UNK	C	537:GLU	N	13.64
1	1	355:UNK	C	368:UNK	N	9.80

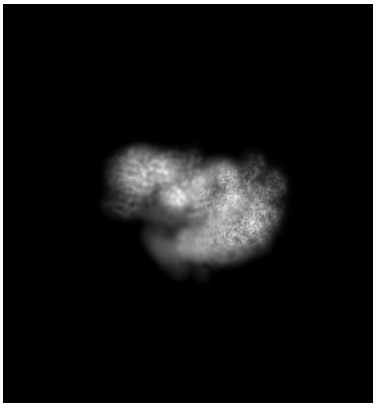
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-23904. These allow visual inspection of the internal detail of the map and identification of artifacts.

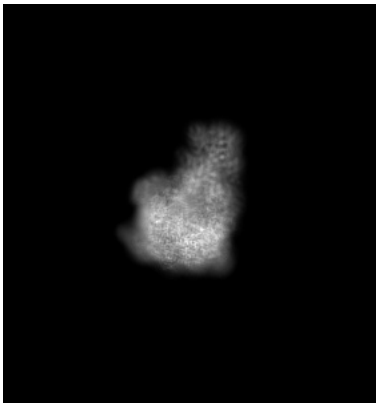
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

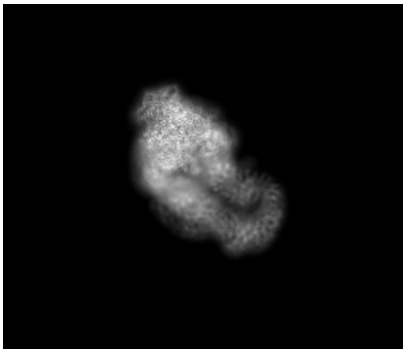
6.1.1 Primary map



X



Y

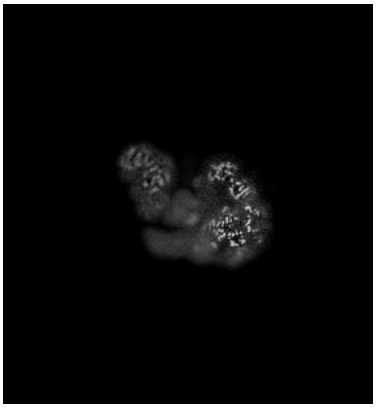


Z

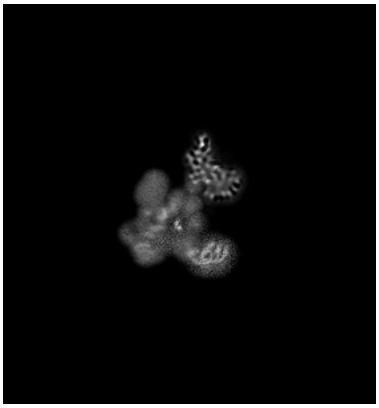
The images above show the map projected in three orthogonal directions.

6.2 Central slices

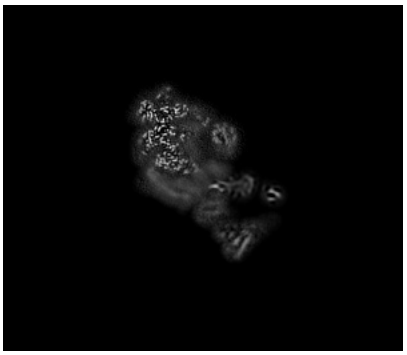
6.2.1 Primary map



X Index: 257



Y Index: 221

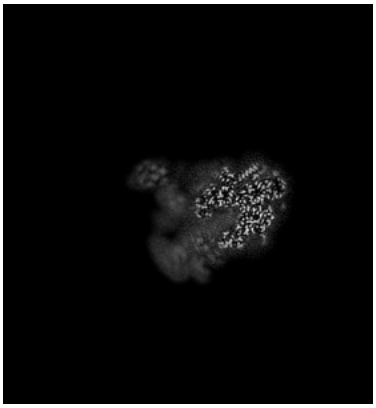


Z Index: 239

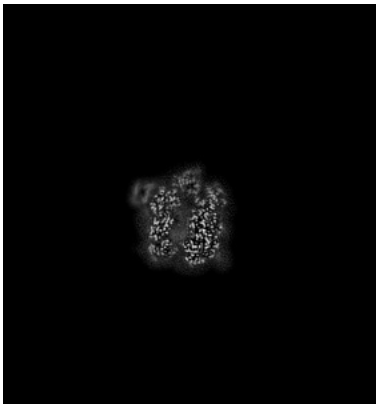
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

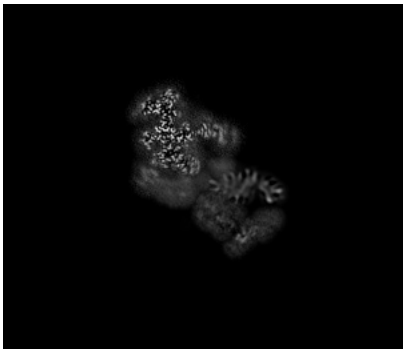
6.3.1 Primary map



X Index: 215



Y Index: 273

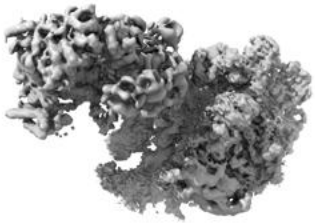


Z Index: 248

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

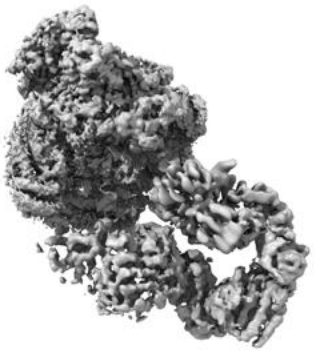
6.4 Orthogonal surface views (i)

6.4.1 Primary map



X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0164. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

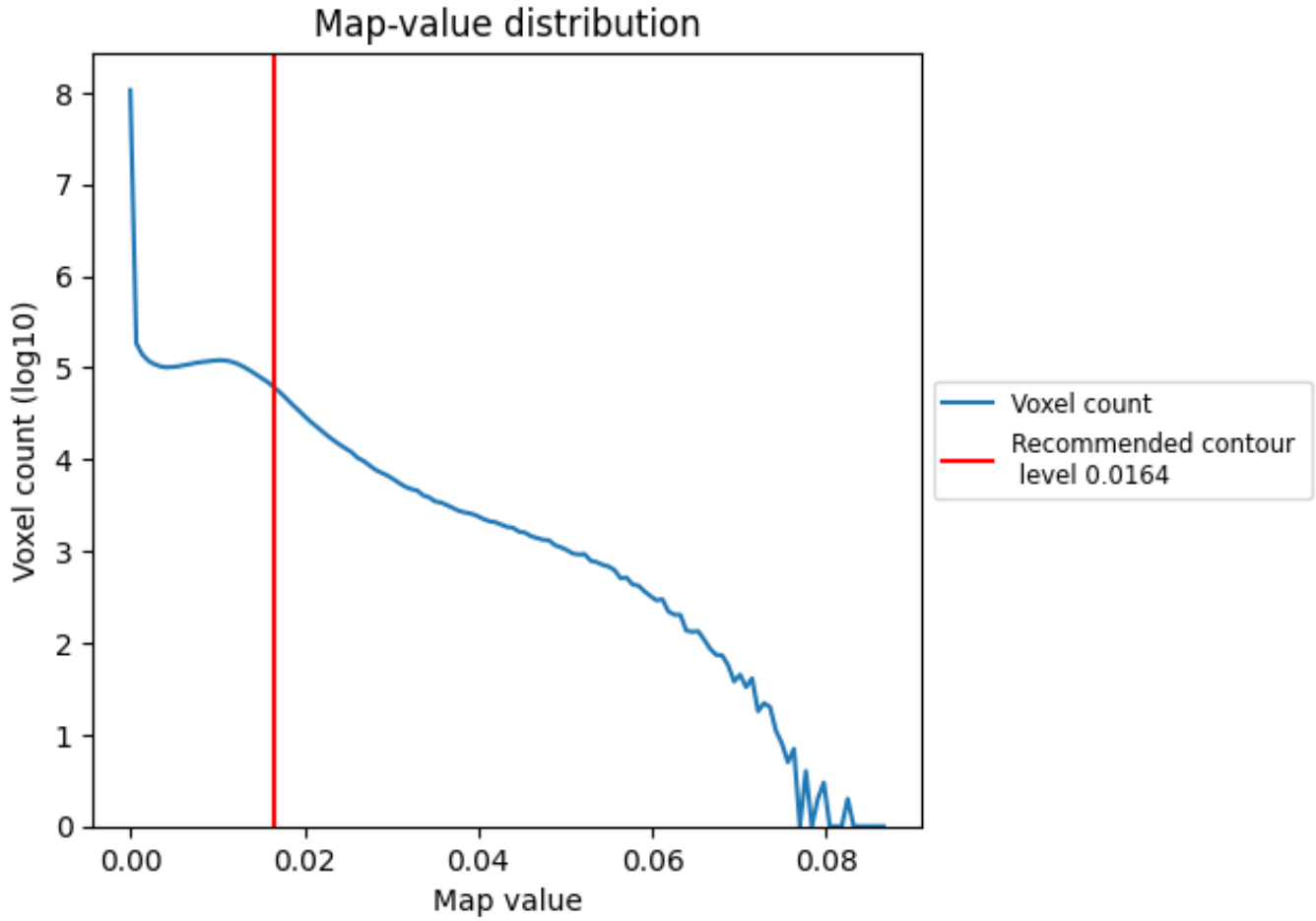
6.5 Mask visualisation (i)

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

7 Map analysis

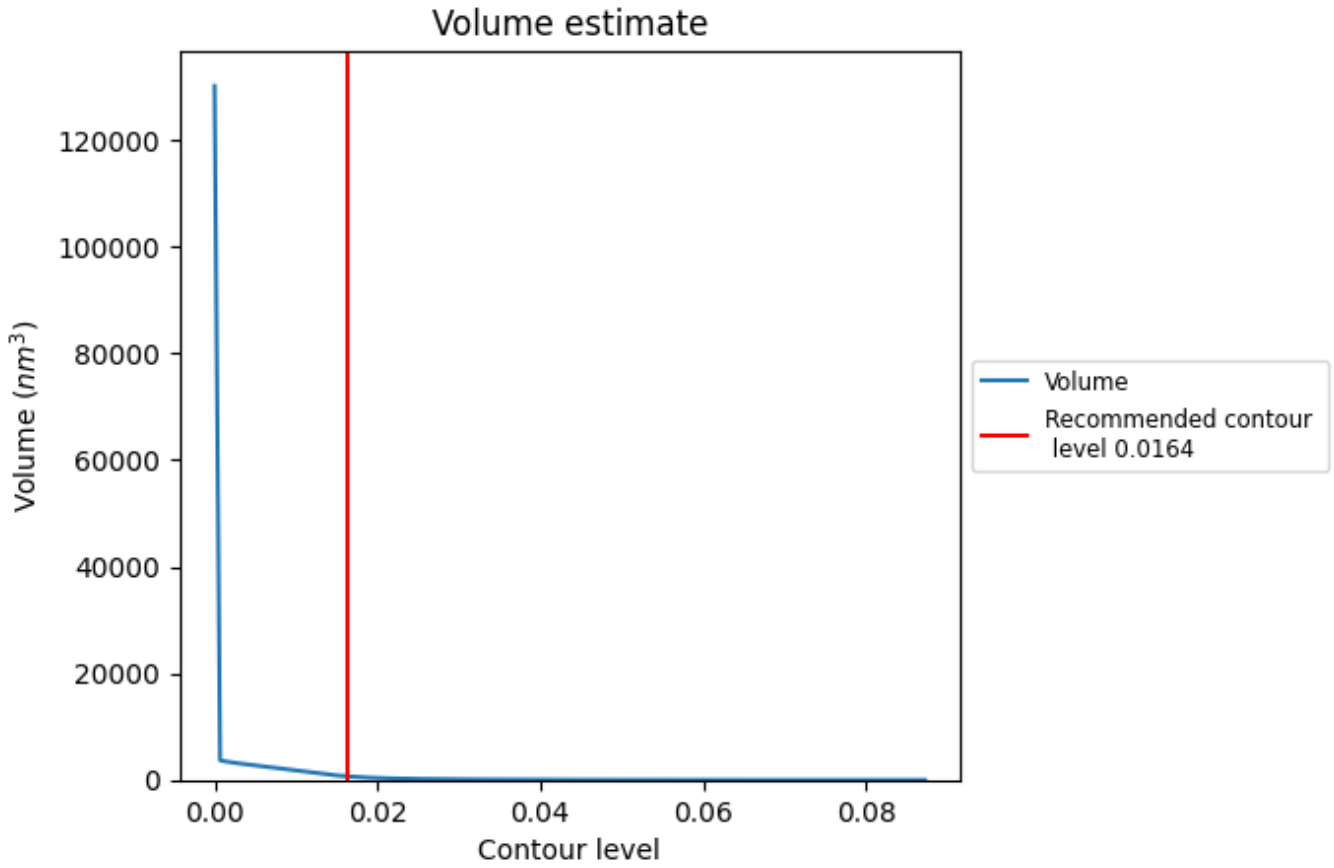
This section contains the results of statistical analysis of the map.

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 (i)



The volume at the recommended contour level is 672 nm^3; this corresponds to an approximate mass of 607 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.

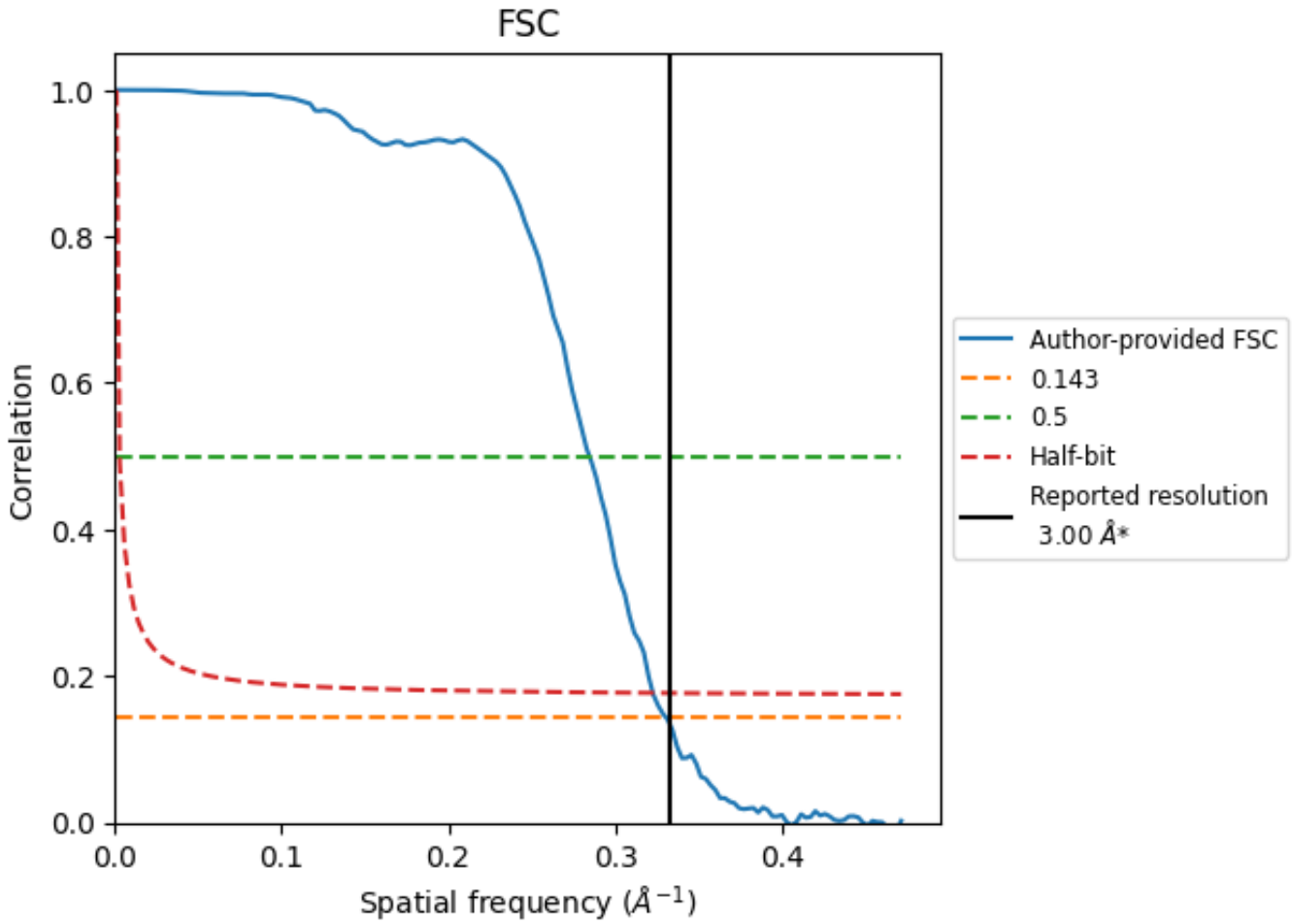
7.3 Rotationally averaged power spectrum (i)

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

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

8.2 Resolution estimates (i)

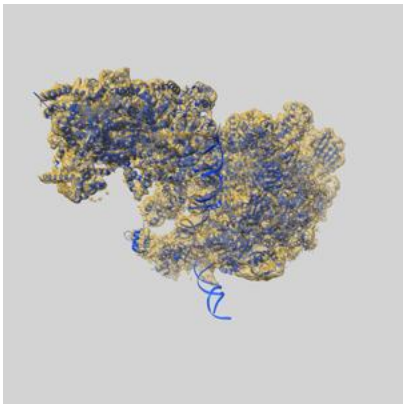
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.00	-	-
Author-provided FSC curve	3.02	3.51	3.10
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

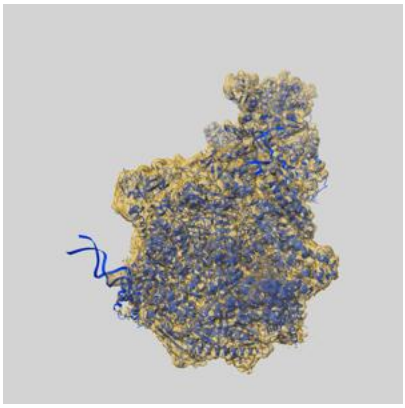
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-23904 and PDB model 7ML0. Per-residue inclusion information can be found in section 3 on page 10.

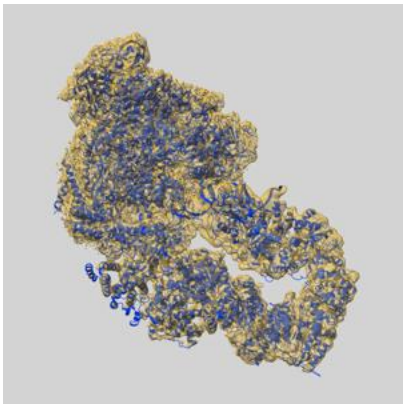
9.1 Map-model overlay



X



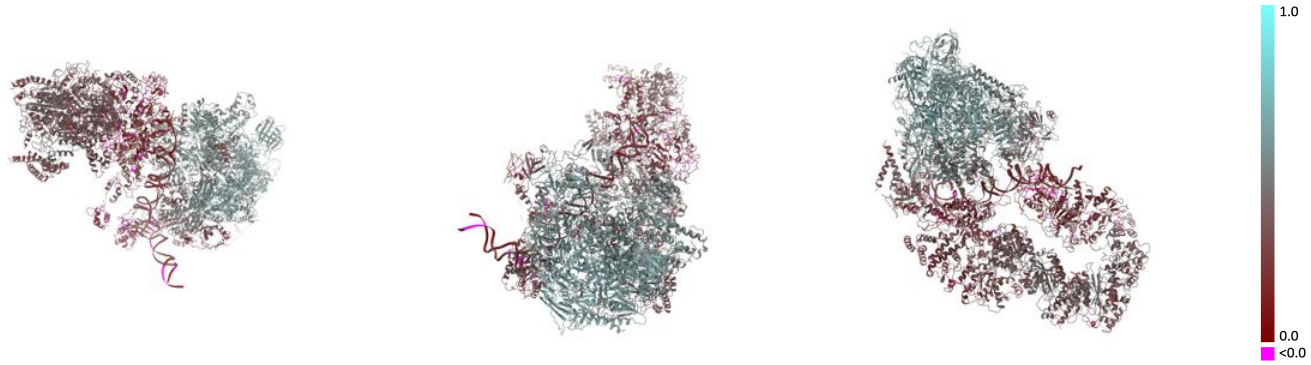
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0164 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.

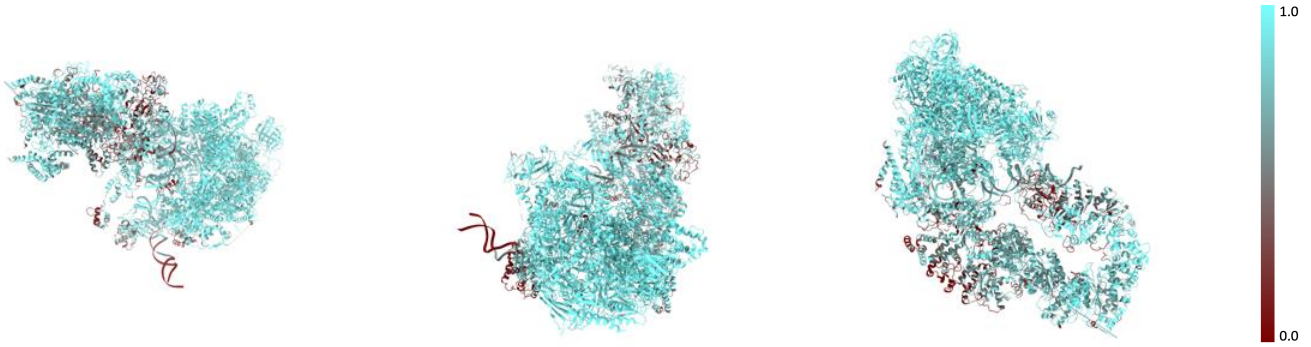
9.2 Q-score mapped to coordinate model



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.

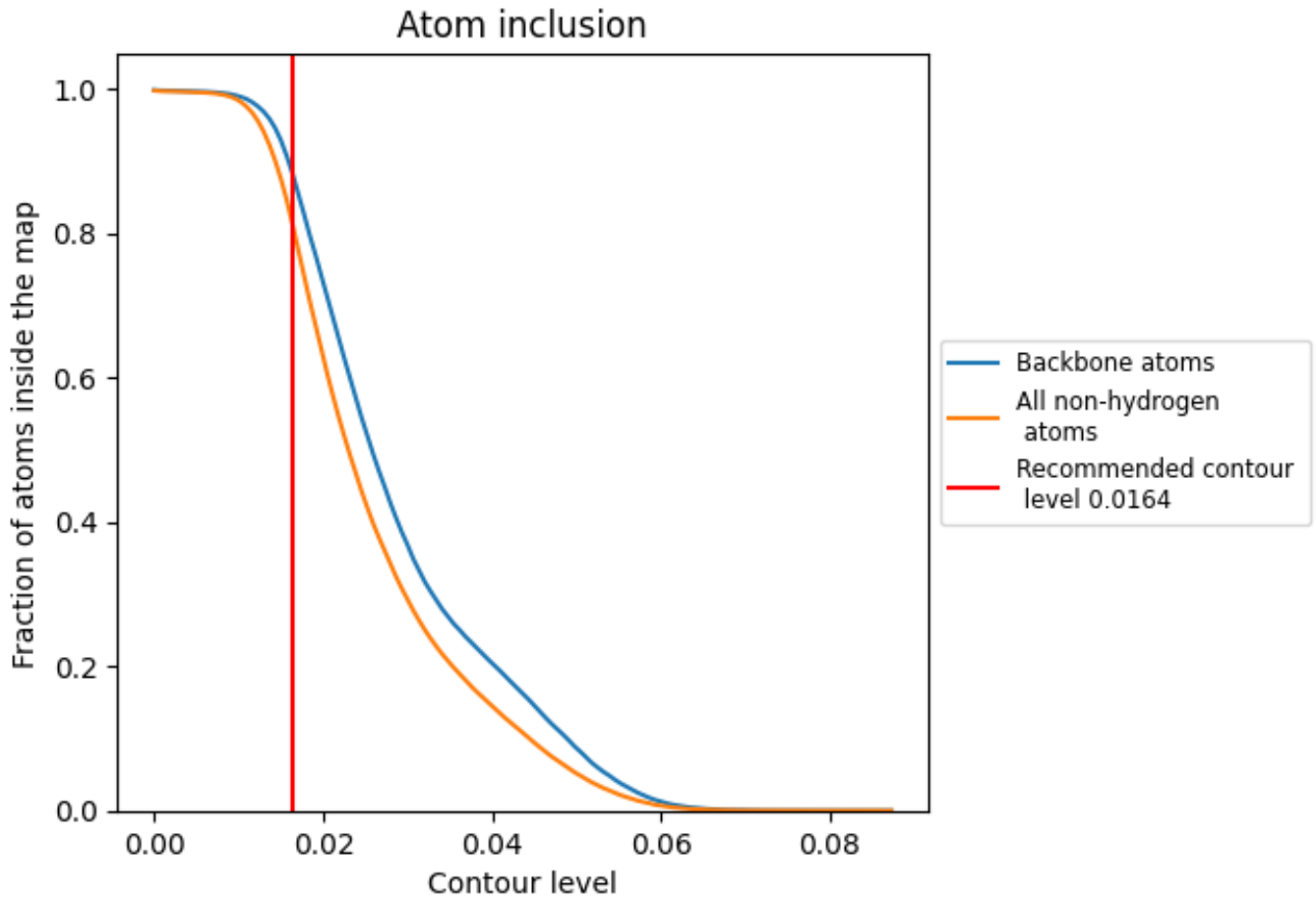
9.3 Atom inclusion mapped to coordinate model



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 (0.0164).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 81% 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 (0.0164) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8114	0.4170
0	0.5869	0.3220
1	0.6461	0.3280
2	0.7523	0.3140
3	0.4586	0.2560
4	0.7883	0.3870
5	0.6423	0.2530
6	0.8185	0.4040
7	0.5225	0.2200
A	0.9763	0.5500
B	0.9756	0.5650
C	0.9798	0.5840
D	0.8240	0.3720
E	0.9653	0.5270
F	0.9877	0.5850
G	0.9274	0.4510
H	0.9566	0.5370
I	0.9294	0.4960
J	0.9884	0.5880
K	0.9270	0.5450
L	0.9827	0.5410
M	0.6342	0.3970
N	0.6525	0.1870
O	0.5259	0.1790
Q	0.9136	0.3960
R	0.8381	0.3200
T	0.6513	0.1850
W	0.8847	0.2360
X	0.7127	0.2050