



wwPDB EM Validation Summary Report ⓘ

Nov 20, 2022 – 12:35 PM EST

PDB ID : 7ML1
EMDB ID : EMD-23905
Title : RNA polymerase II pre-initiation complex (PIC2)
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 : 4.00 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 ⓘ 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.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

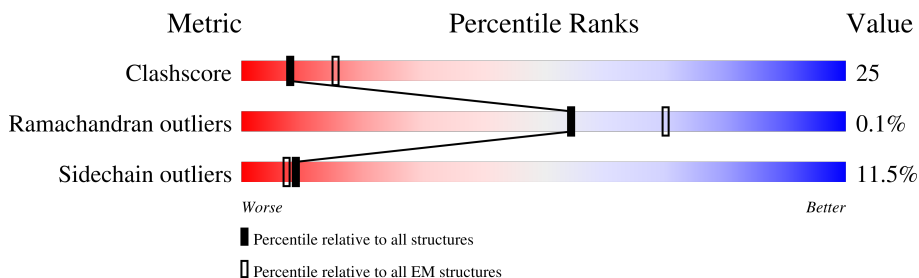
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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 $\leq 5\%$. 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	1	542	
2	4	338	
3	0	778	
4	6	461	
5	2	513	
6	5	72	
7	7	843	
8	3	321	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	O	240	
10	N	57	
11	T	57	
12	A	1733	
13	B	1224	
14	C	318	
15	D	221	
16	E	215	
17	F	155	
18	G	171	
19	H	146	
20	I	122	
21	J	70	
22	K	120	
23	L	70	
24	M	345	
25	Q	735	
26	R	400	
27	U	286	
28	V	122	
29	W	482	
30	X	328	

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
32	SF4	0	801	-	-	X	-

2 Entry composition i

There are 33 unique types of molecules in this entry. The entry contains 64538 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 Tfb1.

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

- Molecule 2 is a protein called General transcription and DNA repair factor IIIH subunit TFB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	4	284	Total	C	N	O	S	0	0
			2041	1310	343	376	12		

- Molecule 3 is a protein called General transcription and DNA repair factor IIIH helicase subunit XPD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	0	754	Total	C	N	O	S	0	0
			6108	3891	1032	1147	38		

- Molecule 4 is a protein called General transcription and DNA repair factor IIIH subunit SSL1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	6	351	Total	C	N	O	S	0	0
			2527	1590	454	456	27		

- Molecule 5 is a protein called RNA polymerase II transcription factor B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	2	460	Total	C	N	O	S	0	0
			3011	1856	562	584	9		

- Molecule 6 is a protein called General transcription and DNA repair factor IIIH subunit TFB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	66	Total	C	N	O	S	0	0
			498	314	89	93	2		

- Molecule 7 is a protein called General transcription and DNA repair factor IIIH helicase subunit XPB.

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

- Molecule 8 is a protein called BJ4_G0050160.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	3	138	860	533	160	160	7	0	0

- Molecule 9 is a protein called TATA-box-binding protein.

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

- Molecule 10 is a DNA chain called non-template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
10	N	57	1178	562	227	332	57	0	0

- Molecule 11 is a DNA chain called template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
11	T	57	1159	559	191	352	57	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerase subunit.

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

- Molecule 13 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	B	1147	9132	5775	1602	1700	55	0	0

- Molecule 14 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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

- Molecule 15 is a protein called DNA-directed RNA polymerase II subunit RPB4.

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

- Molecule 16 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		
16	E	213	1744	1107	308	318	11	0	0

- Molecule 17 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		
17	F	83	670	428	114	125	3	0	0

- Molecule 18 is a protein called DNA-directed RNA polymerase II subunit RPB7.

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

- Molecule 19 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 20 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 21 is a protein called DNA-directed RNA polymerases II subunit RPABC5.

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

- Molecule 22 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 23 is a protein called DNA-directed RNA polymerases II subunit RPABC4.

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

- Molecule 24 is a protein called Transcription initiation factor IIB.

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

- Molecule 25 is a protein called Transcription initiation factor IIF subunit alpha.

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

- Molecule 26 is a protein called Transcription initiation factor IIF subunit beta.

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

- Molecule 27 is a protein called Transcription initiation factor IIA large subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	U	44	Total	C	N	O	S	0	0
			366	233	64	66	3		

- Molecule 28 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	V	50	389	245	65	76	3	0	0

- Molecule 29 is a protein called Transcription initiation factor IIE subunit alpha.

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

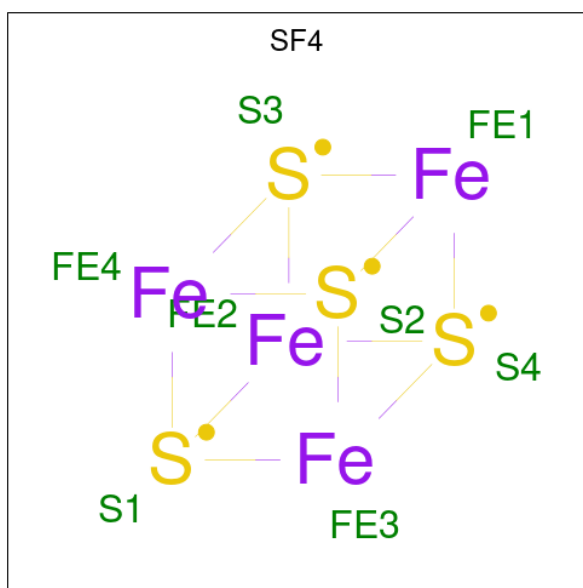
- Molecule 30 is a protein called Transcription initiation factor IIE subunit beta.

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

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

Mol	Chain	Residues	Atoms		AltConf
31	4	1	Total	Zn	0
			1	1	
31	6	4	Total	Zn	0
			4	4	
31	3	2	Total	Zn	0
			2	2	
31	A	3	Total	Zn	0
			3	3	
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	X	1	Total	Zn	0
			1	1	

- Molecule 32 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
32	0	1	8	4	4	0

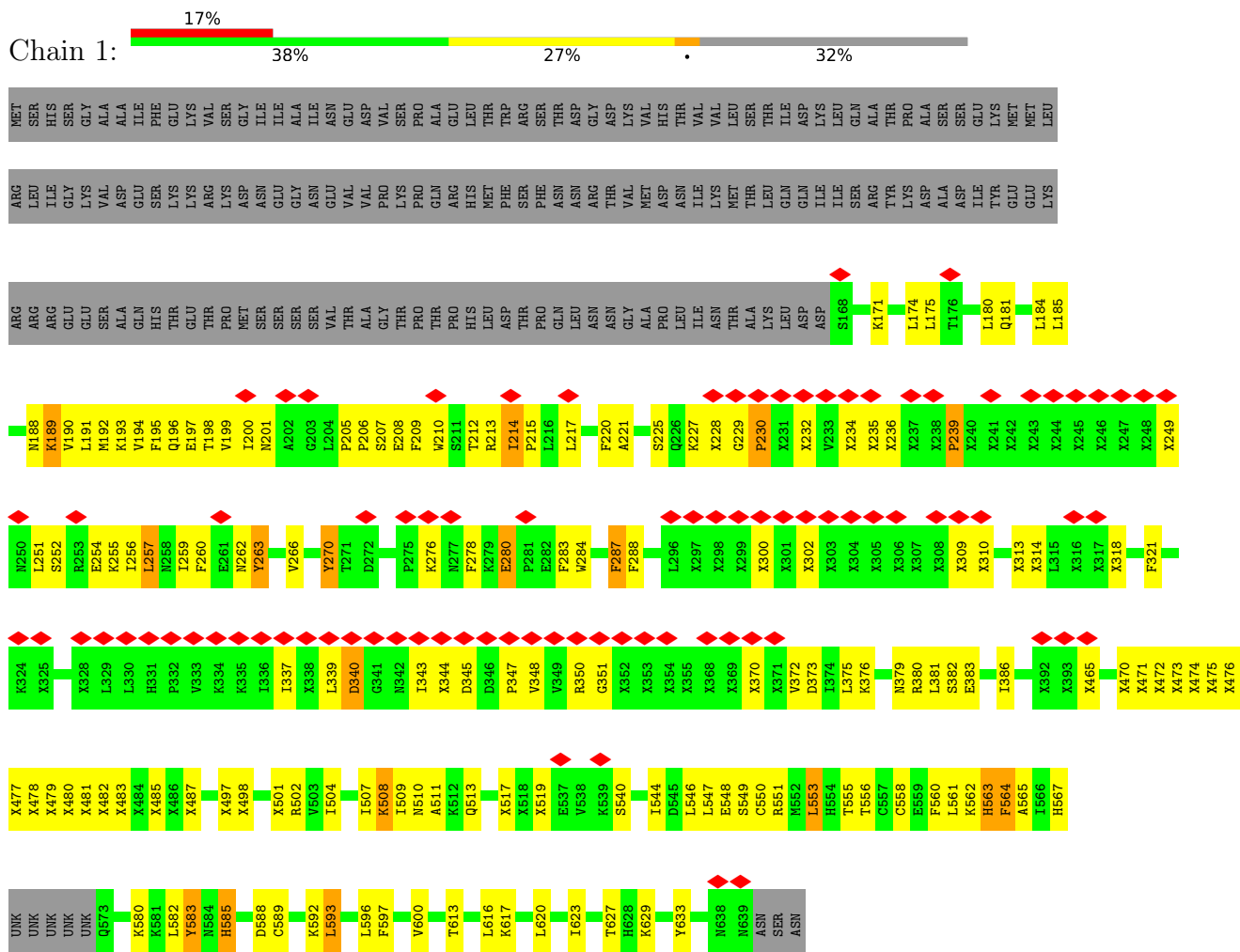
- Molecule 33 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

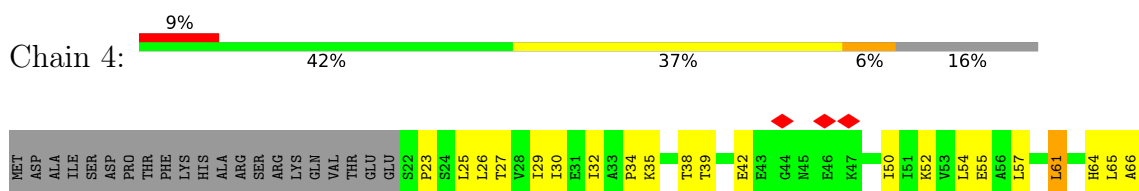
3 Residue-property plots i

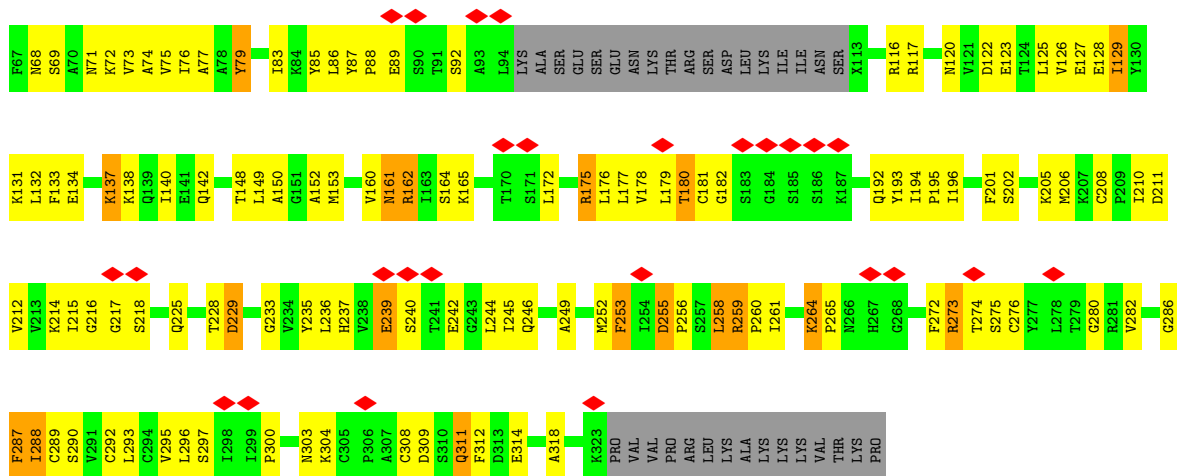
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Tfb1

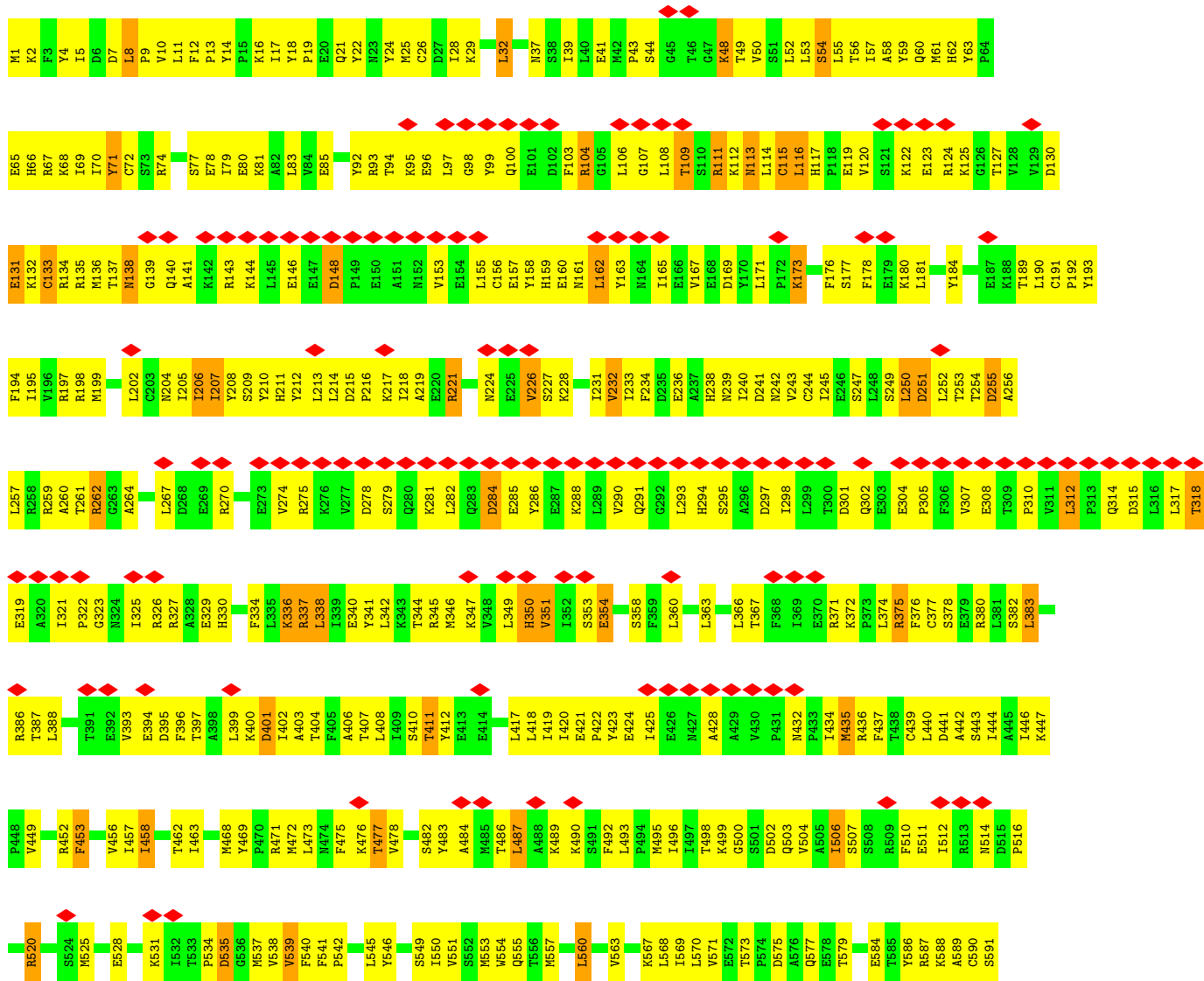


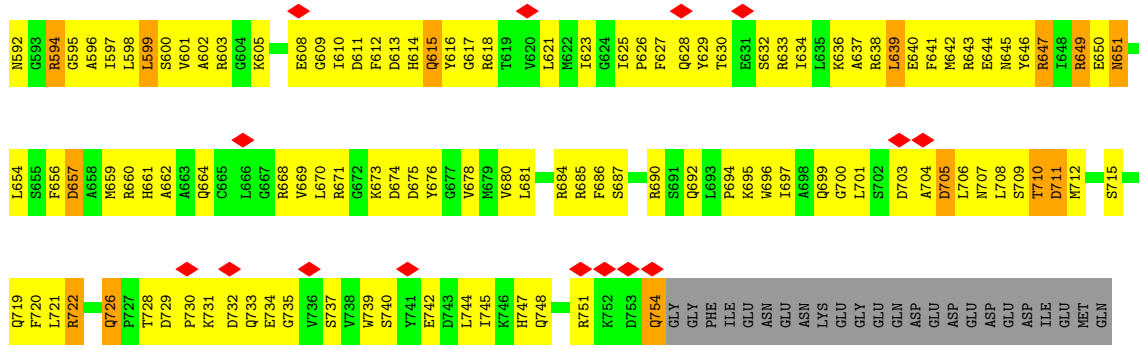
- Molecule 2: General transcription and DNA repair factor IIIH subunit TFB4



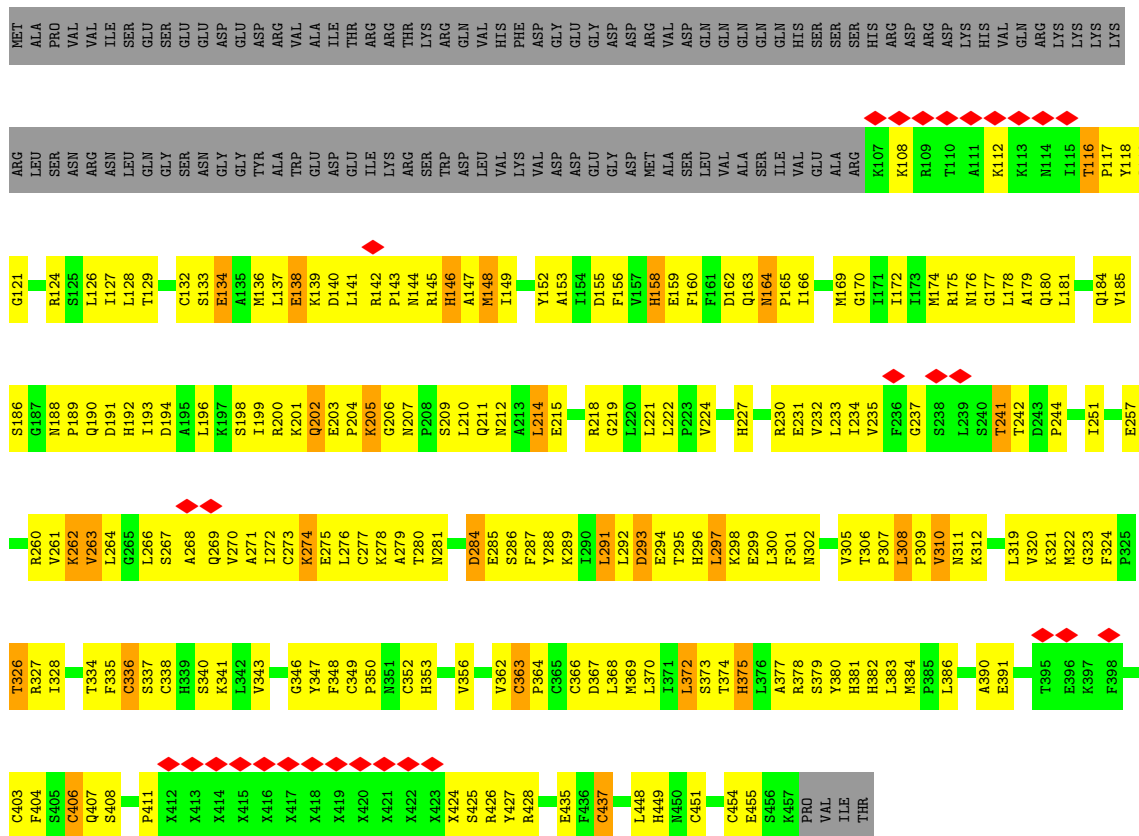


• Molecule 3: General transcription and DNA repair factor IIIH helicase subunit XPD

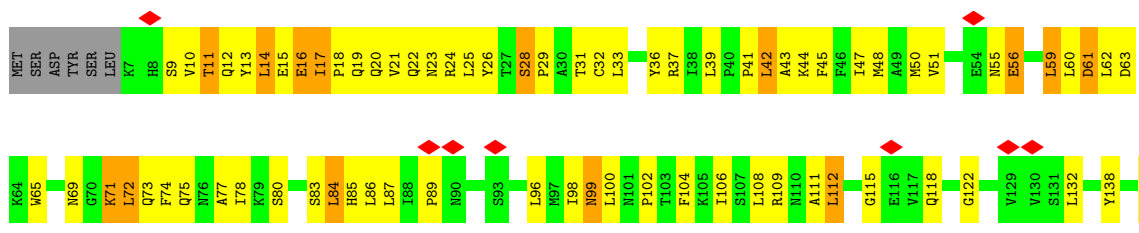


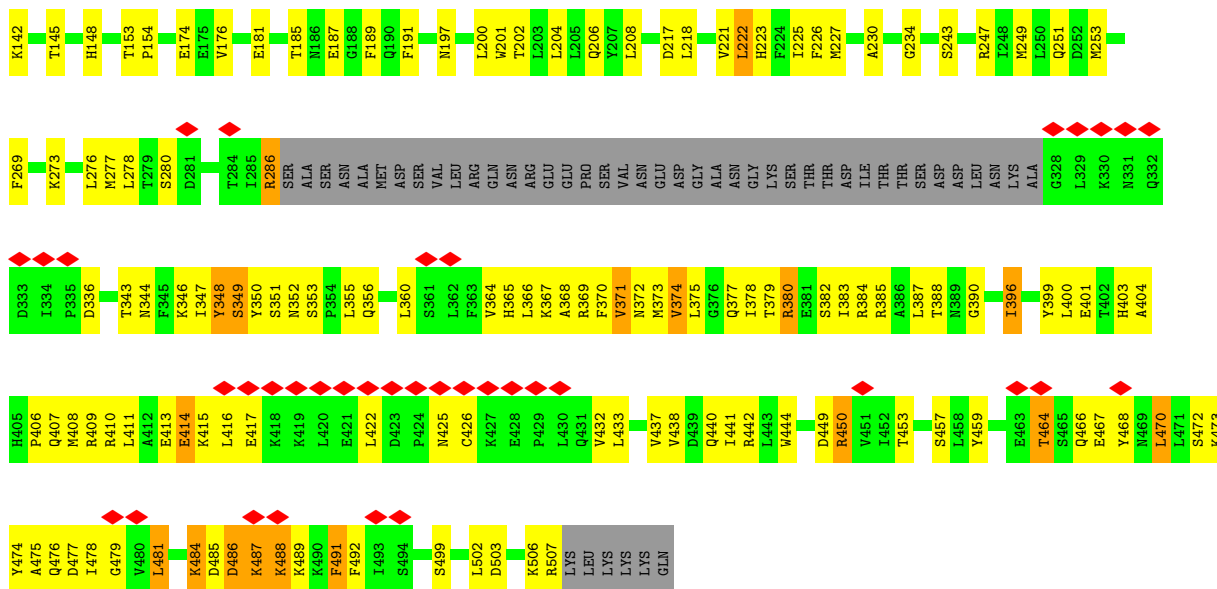


● Molecule 4: General transcription and DNA repair factor IIH subunit SSL1

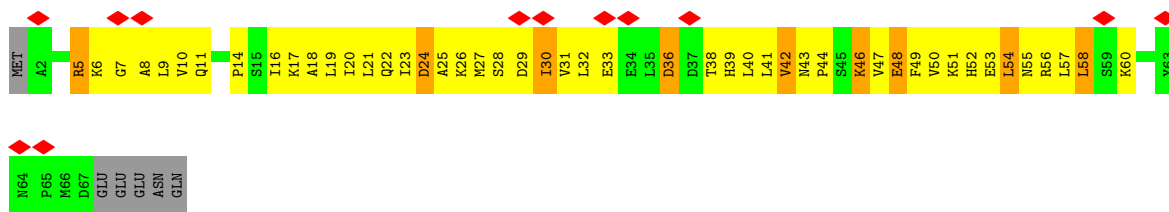


● Molecule 5: RNA polymerase II transcription factor B subunit 2

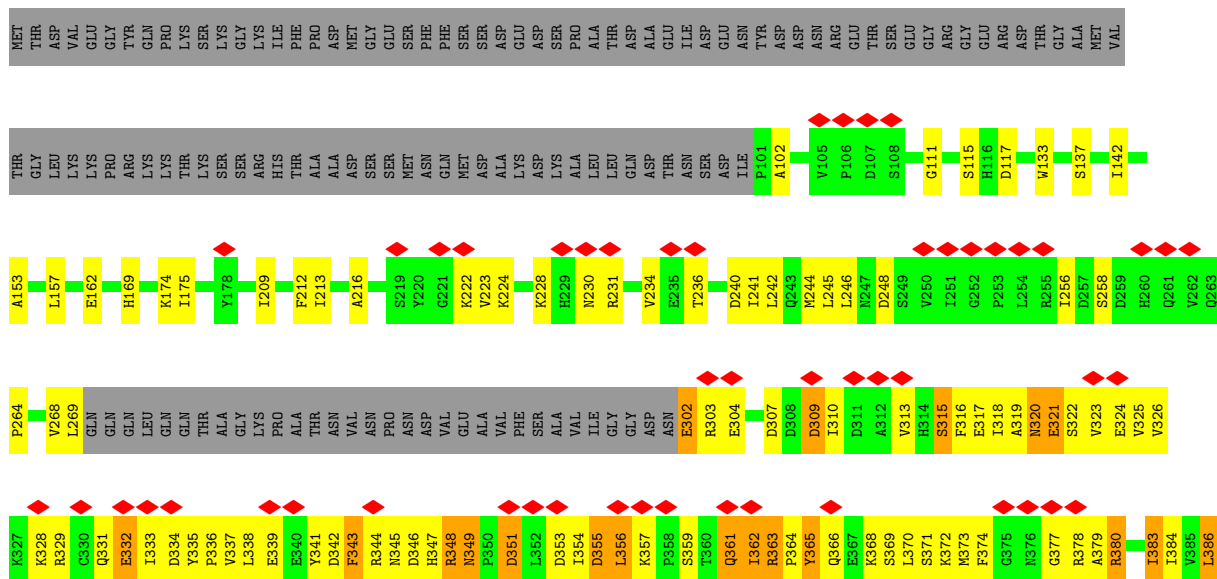
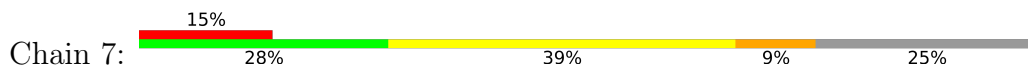


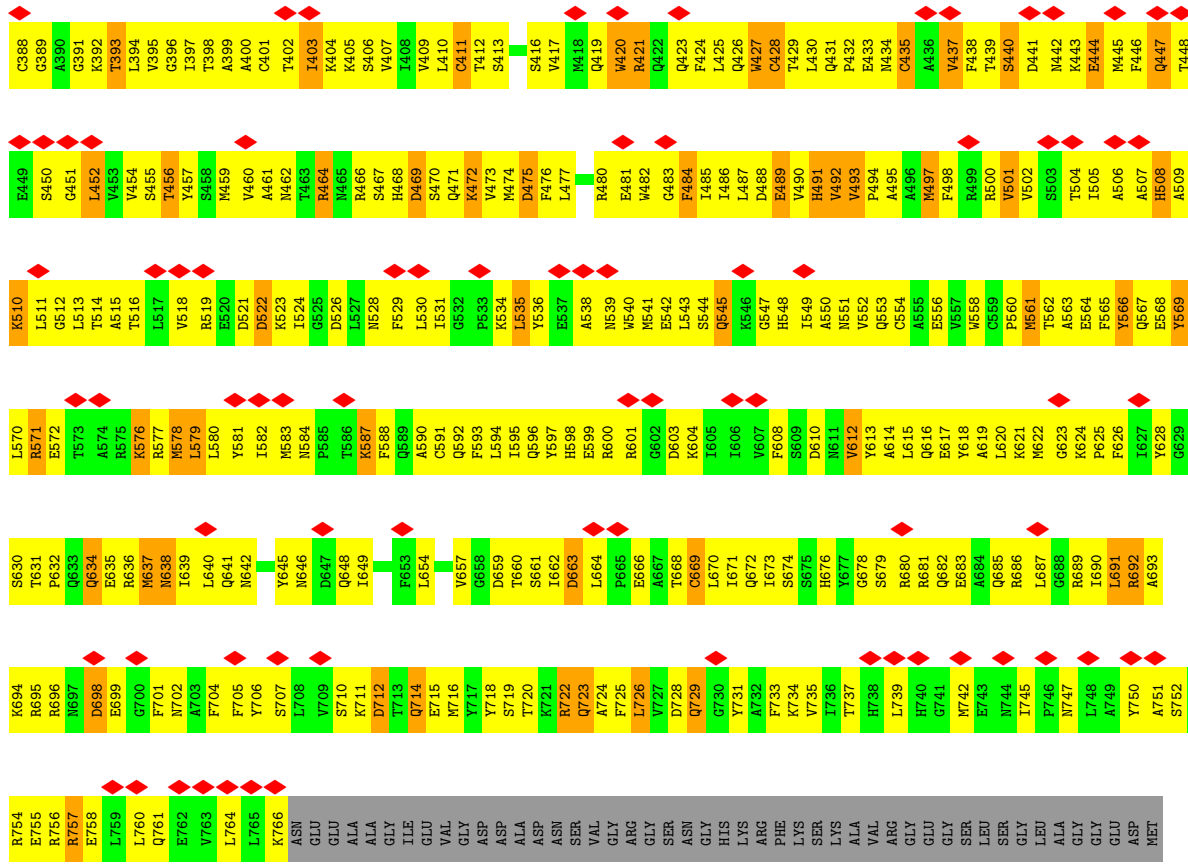


• Molecule 6: General transcription and DNA repair factor IIH subunit TFB5

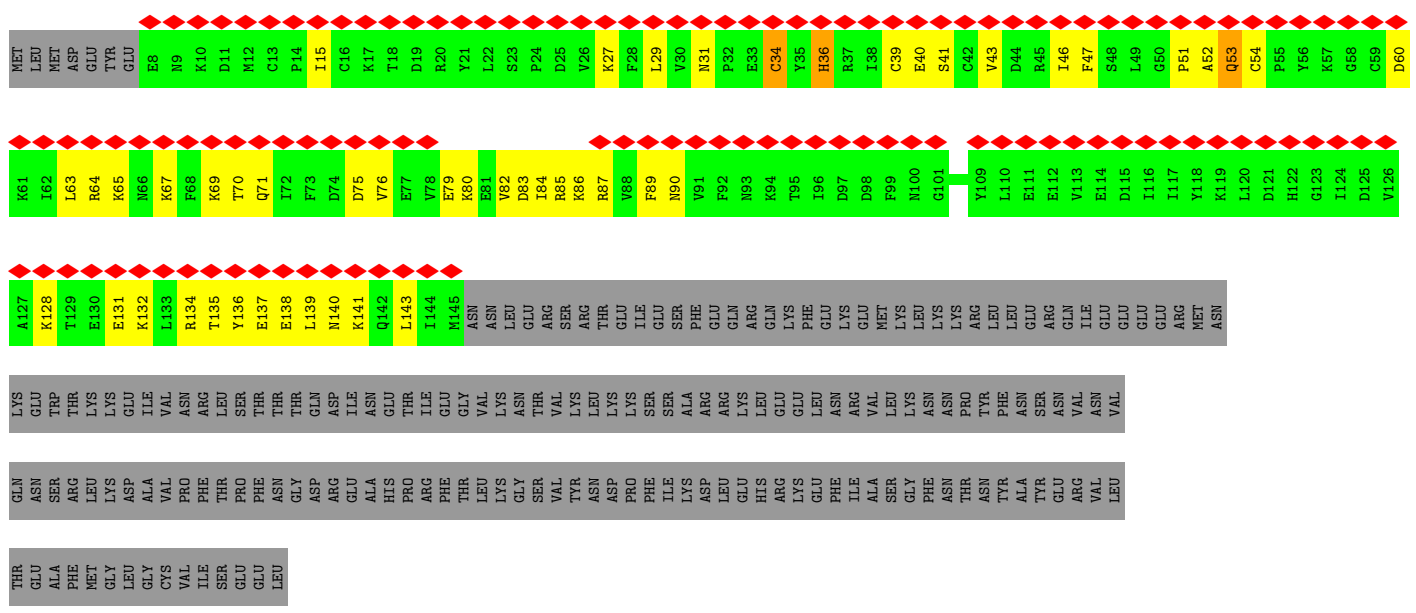


• Molecule 7: General transcription and DNA repair factor IIH helicase subunit XPB

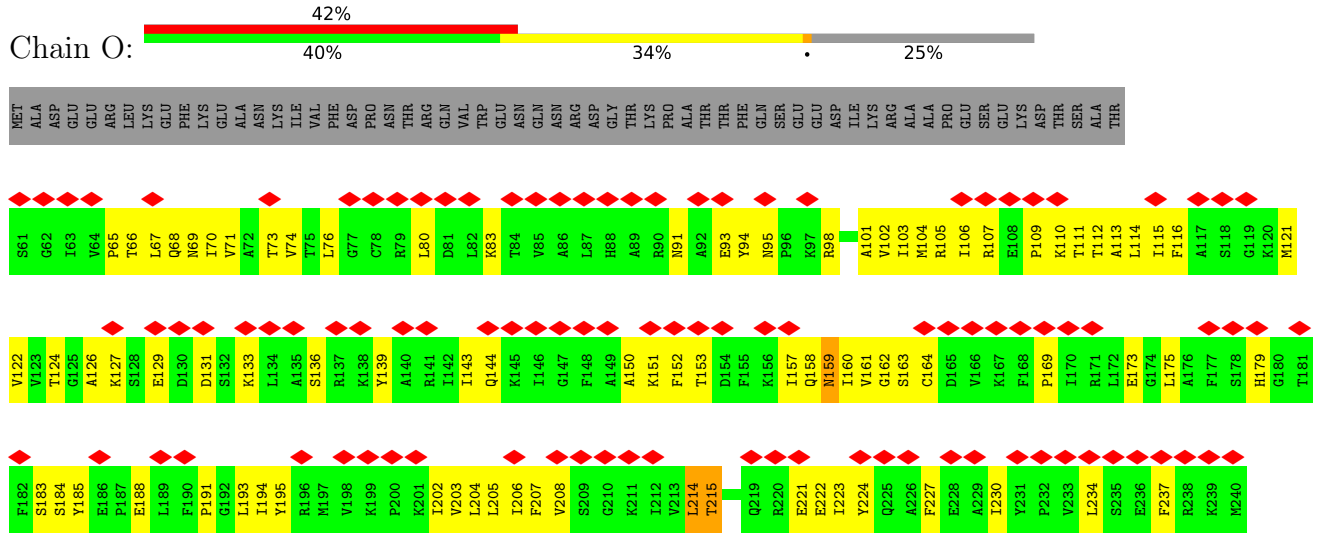




● Molecule 8: BJ4_G0050160.mRNA.1.CDS.1



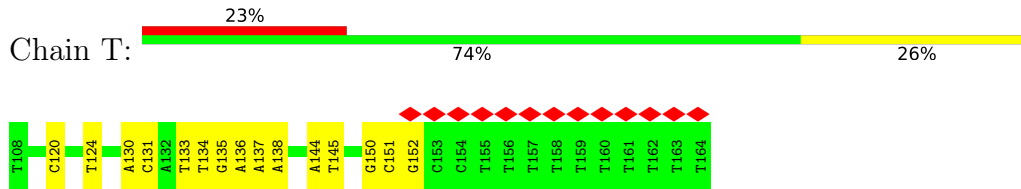
• Molecule 9: TATA-box-binding protein



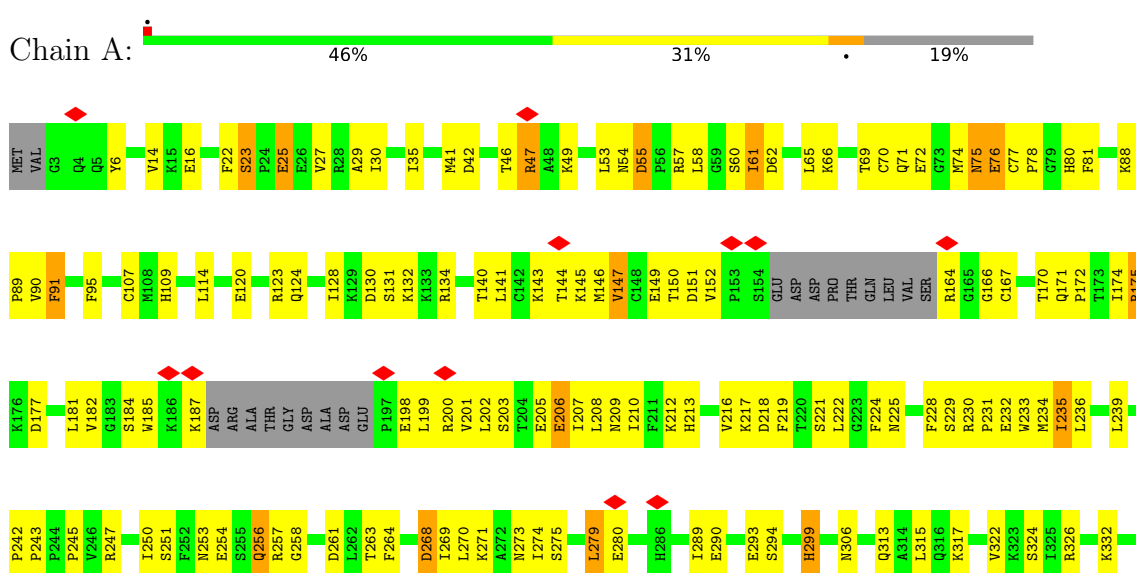
• Molecule 10: non-template strand DNA



• Molecule 11: template strand DNA



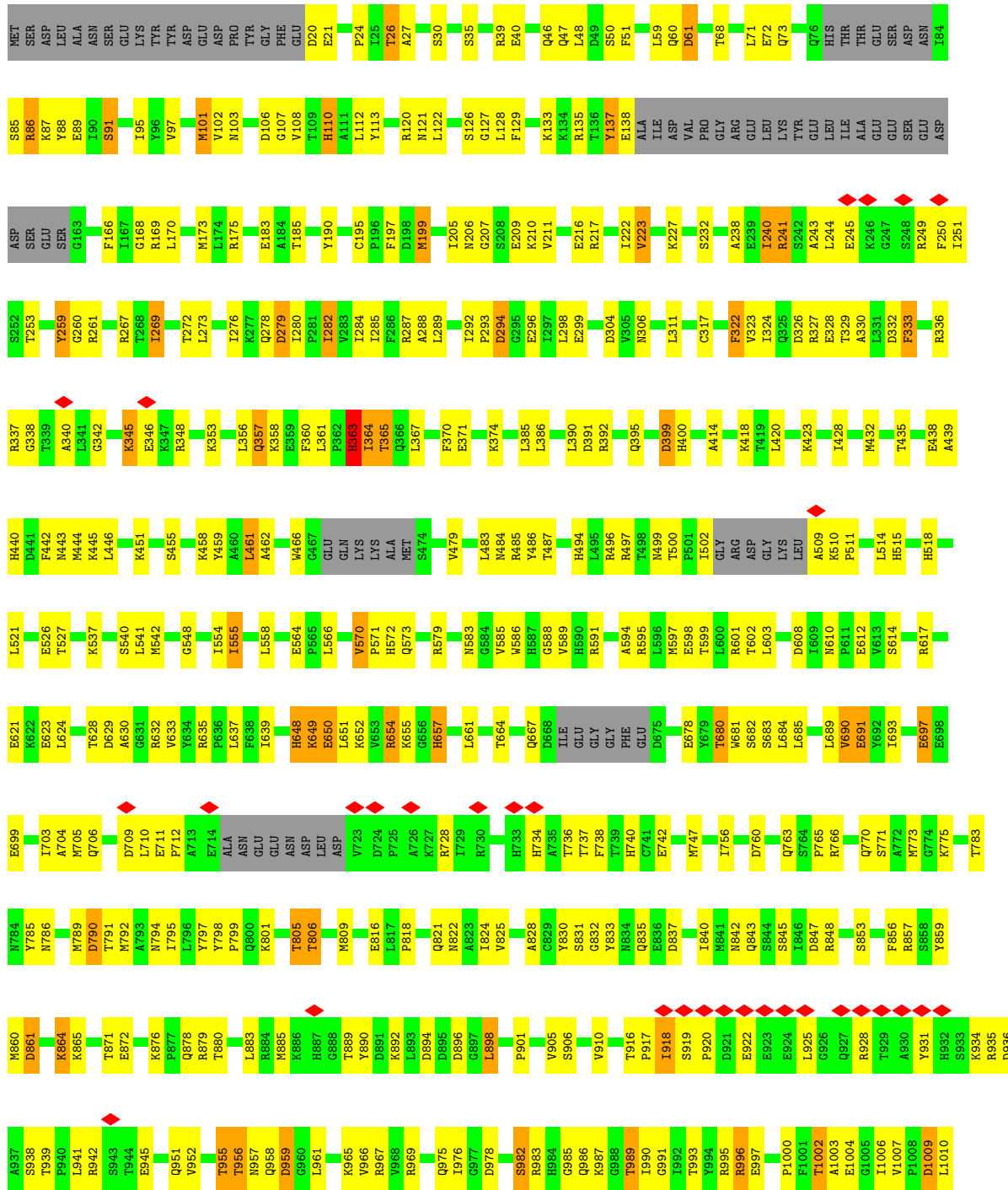
• Molecule 12: DNA-directed RNA polymerase subunit

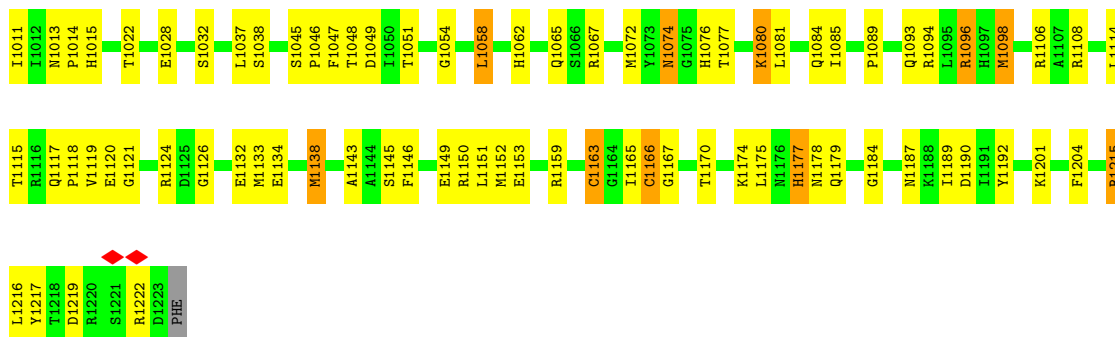


SER	PRO	THR	THR	LEU	SER	PRO	PRO	ASN	SER	TYR	SER	PRO	PRO	THR	THR	PRO	PRO	GLY	GLY	SER	SER	ALA	ALA	THR	THR	TYR	PRO	GLN
ASP	GLU	GLN	LYS	SER	HIS	ASN	HIS	GLU	GLU	ASN	GLU	ASN	ARG															

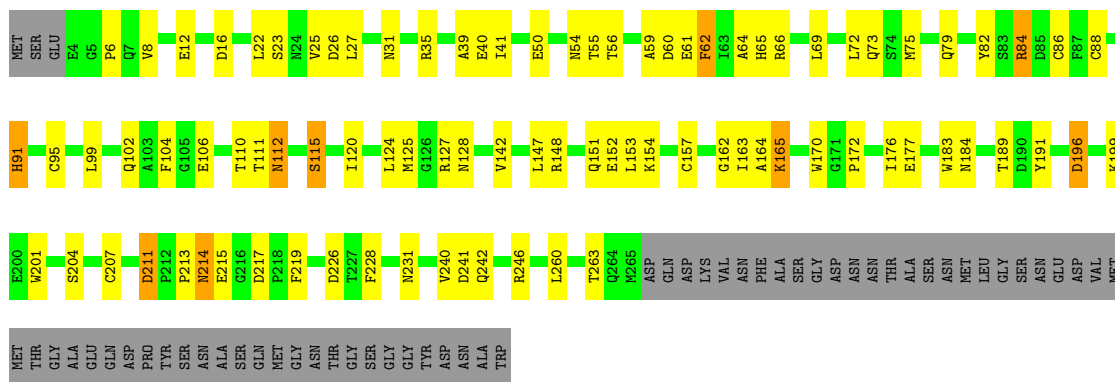
● Molecule 13: DNA-directed RNA polymerase subunit beta

Chain B: 55% 34% 5% 6%

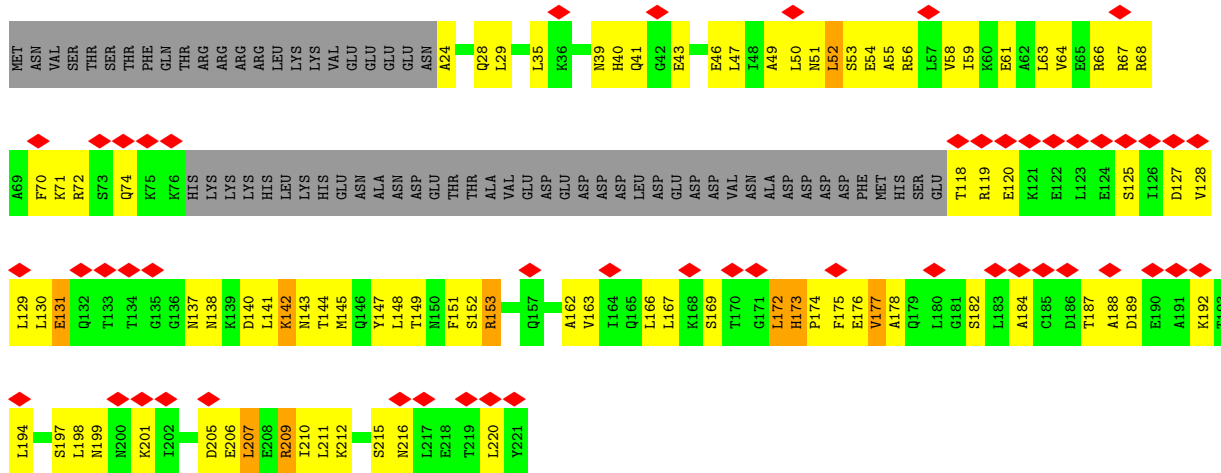




• Molecule 14: DNA-directed RNA polymerase II subunit RPB3

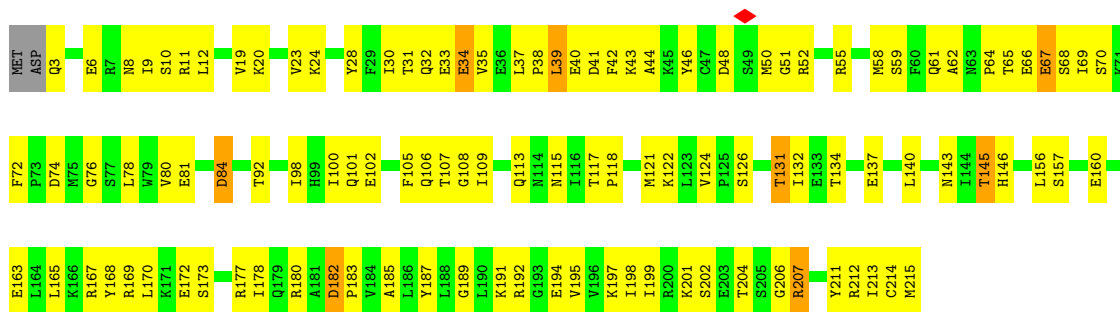


• Molecule 15: DNA-directed RNA polymerase II subunit RPB4

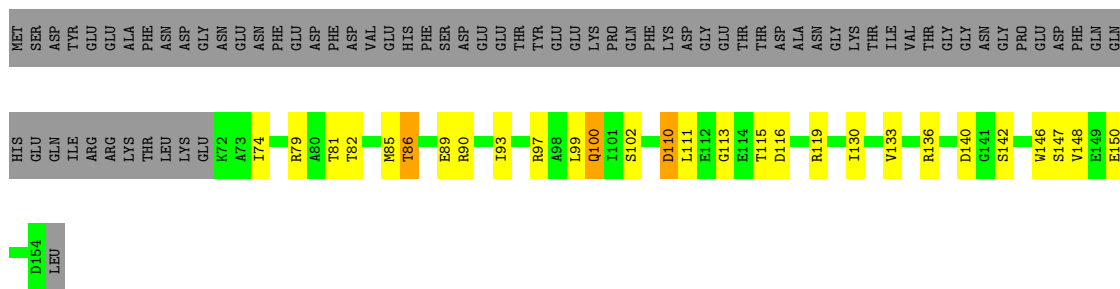
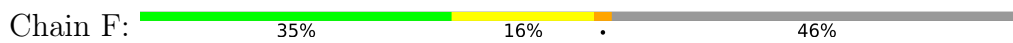


• Molecule 16: DNA-directed RNA polymerases I, II, and III subunit RPABC1

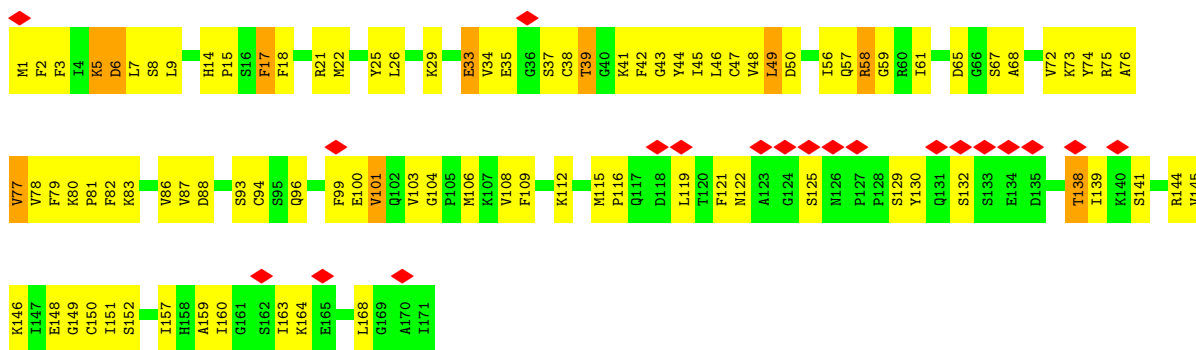




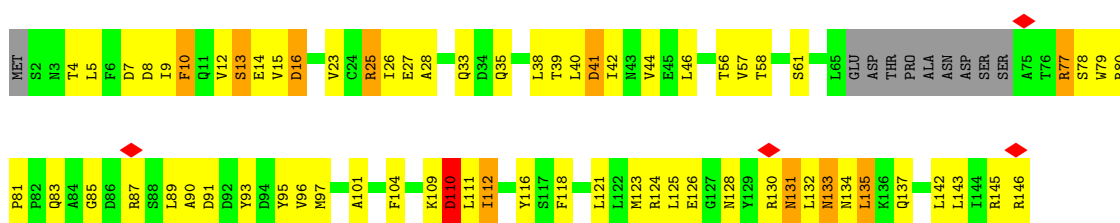
• Molecule 17: DNA-directed RNA polymerases I,II,and III subunit RPABC2



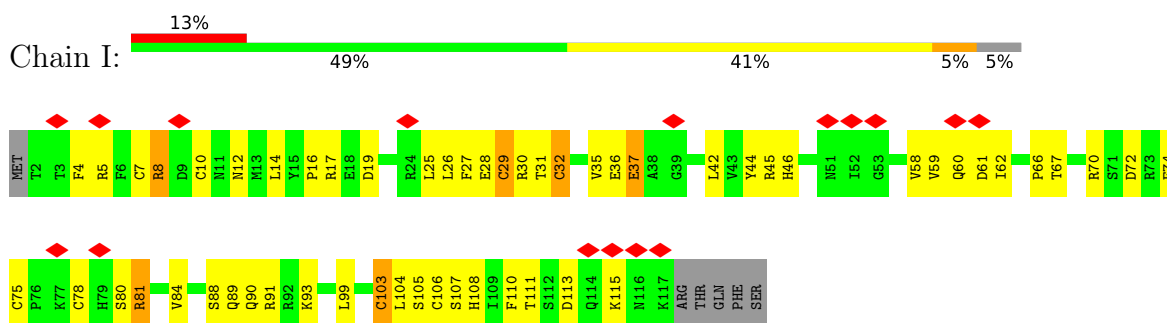
• Molecule 18: DNA-directed RNA polymerase II subunit RPB7



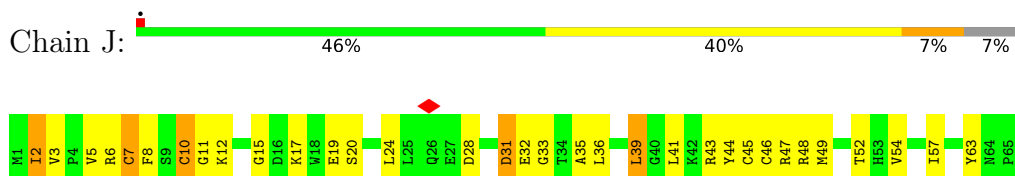
• Molecule 19: DNA-directed RNA polymerases I, II, and III subunit RPABC3



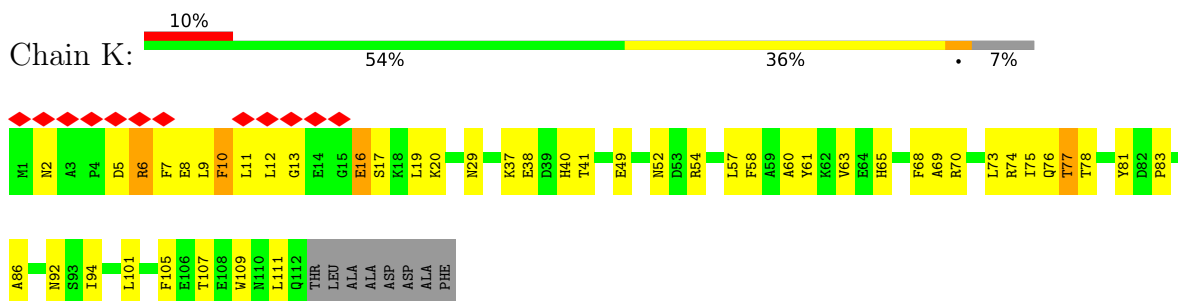
• Molecule 20: DNA-directed RNA polymerase II subunit RPB9



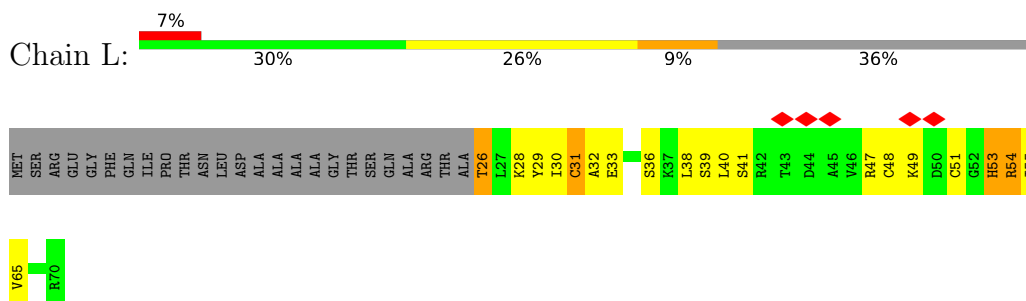
- Molecule 21: DNA-directed RNA polymerases II subunit RPABC5



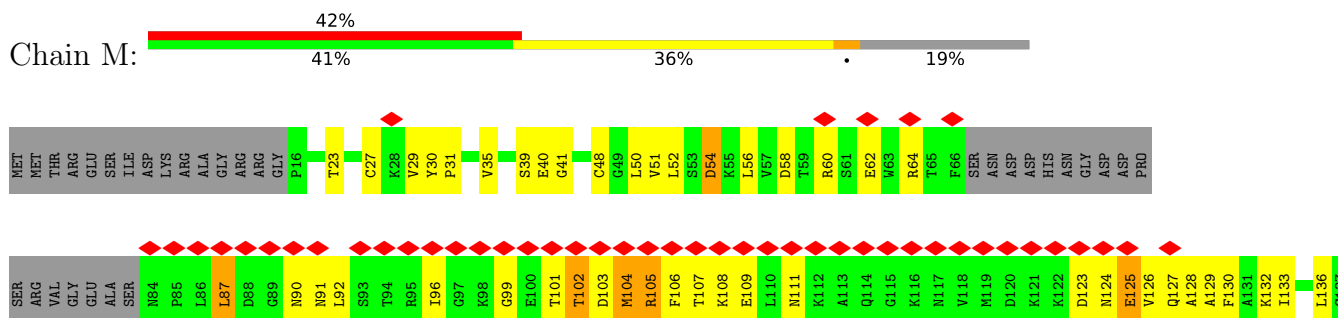
- Molecule 22: DNA-directed RNA polymerase II subunit RPB11

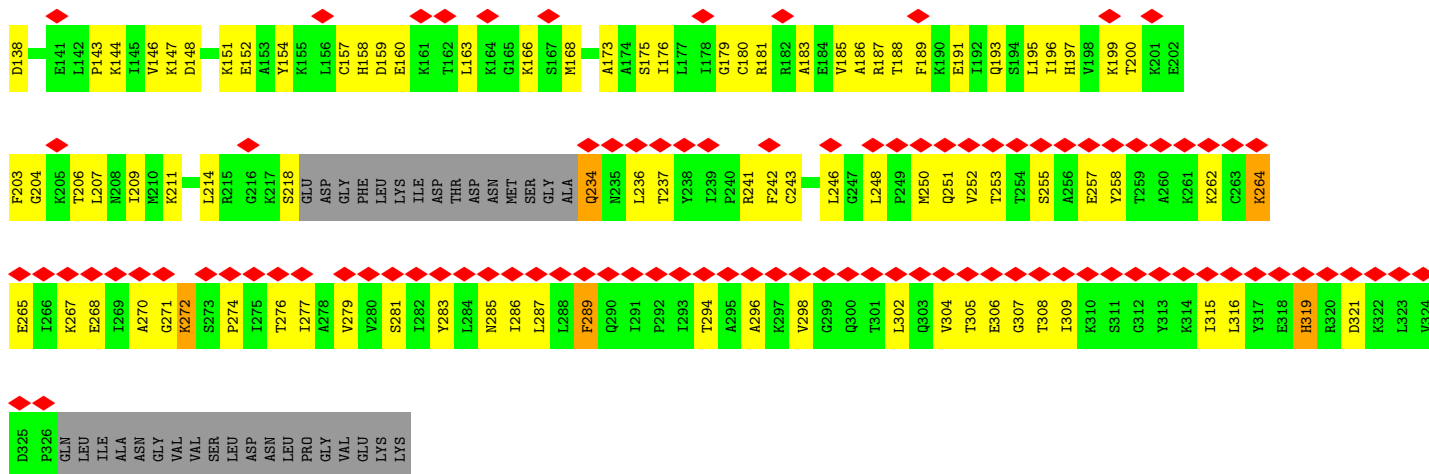


- Molecule 23: DNA-directed RNA polymerases II subunit RPABC4

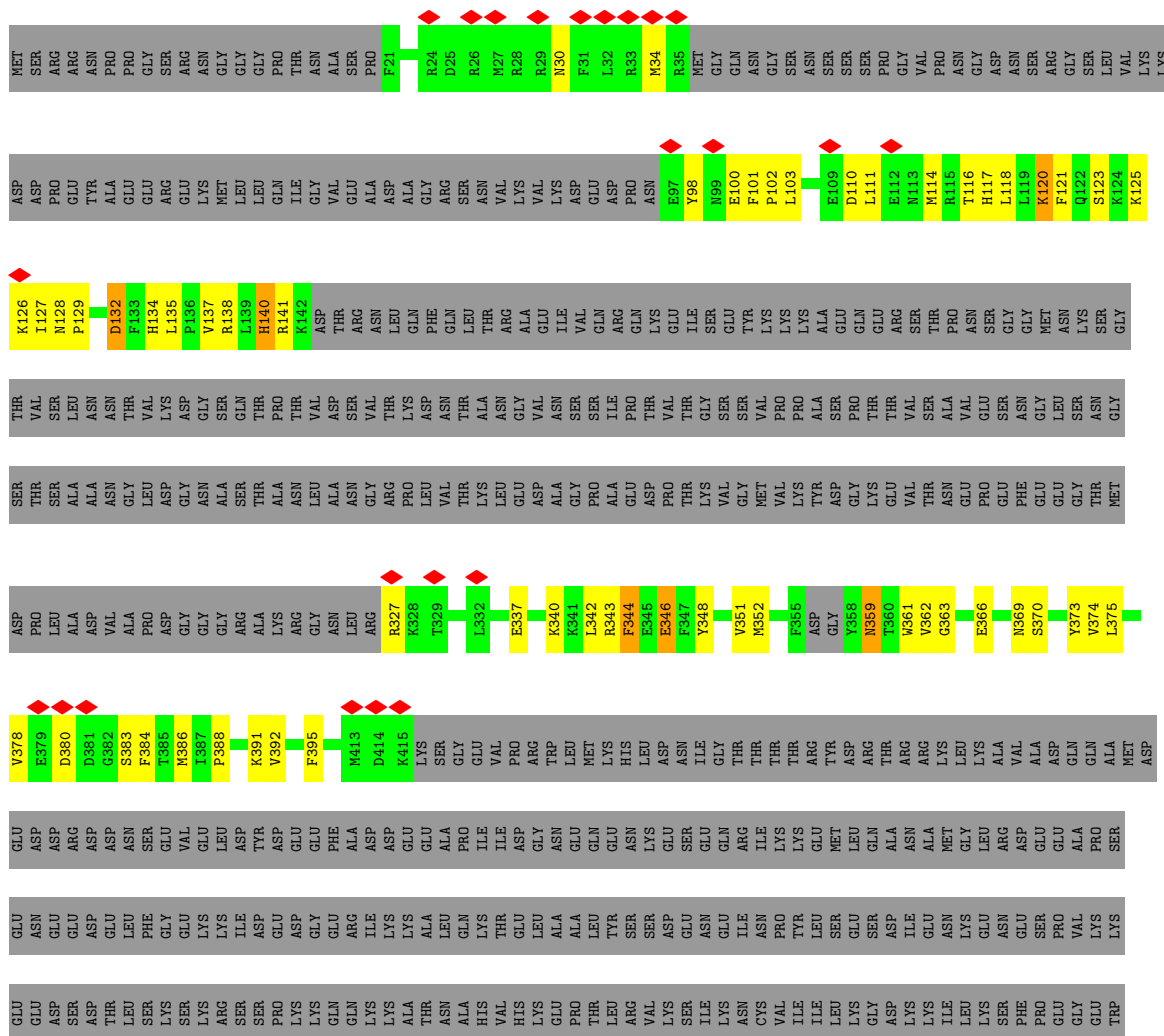


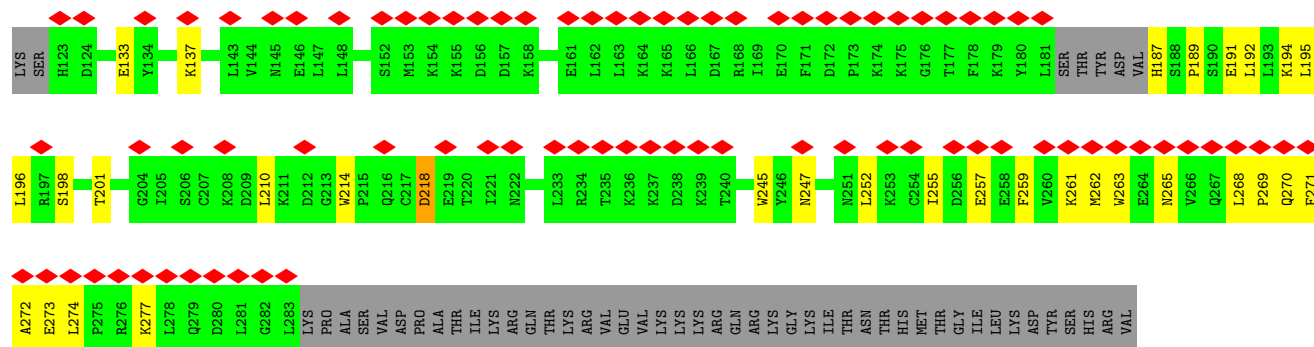
- Molecule 24: Transcription initiation factor IIB





● Molecule 25: Transcription initiation factor IIF subunit alpha





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	33150	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 ($e^-/\text{\AA}^2$)	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	0.048	Depositor
Minimum map value	0.000	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.0135	Depositor
Map size (Å)	453.67996, 405.97998, 419.75998	wwPDB
Map dimensions	428, 383, 396	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: 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	1	0.25	0/1896	0.39	0/2543
2	4	0.27	0/2062	0.46	0/2805
3	0	0.26	0/6226	0.43	0/8407
4	6	0.27	0/2506	0.46	0/3402
5	2	0.25	0/3057	0.45	0/4071
6	5	0.24	0/502	0.50	0/677
7	7	0.27	0/4521	0.48	0/6036
8	3	0.24	0/870	0.39	0/1190
9	O	0.26	0/1443	0.43	0/1942
10	N	0.64	0/1326	1.00	2/2045 (0.1%)
11	T	0.53	0/1294	1.00	0/1994
12	A	0.34	0/11192	0.45	0/15128
13	B	0.36	0/9311	0.45	0/12558
14	C	0.38	0/2099	0.45	0/2845
15	D	0.25	0/1262	0.42	0/1693
16	E	0.35	0/1780	0.44	0/2395
17	F	0.37	0/682	0.44	0/922
18	G	0.30	0/1368	0.46	0/1844
19	H	0.36	0/1107	0.48	0/1499
20	I	0.31	0/962	0.47	0/1295
21	J	0.41	0/541	0.49	0/727
22	K	0.35	0/922	0.46	0/1244
23	L	0.36	0/360	0.55	0/478
24	M	0.26	0/2204	0.45	0/2963
25	Q	0.28	0/1168	0.43	0/1579
26	R	0.27	0/1312	0.47	0/1777
27	U	0.23	0/372	0.46	0/500
28	V	0.25	0/392	0.43	0/529
29	W	0.24	0/1490	0.39	0/2014
30	X	0.23	0/993	0.41	0/1357
All	All	0.32	0/65220	0.49	2/88459 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	4	0	1
4	6	0	1
7	7	0	1
12	A	0	2
13	B	0	2
24	M	0	1
All	All	0	8

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
10	N	20	DT	O4'-C4'-C3'	-6.98	101.71	104.50
10	N	20	DT	O4'-C1'-N1	5.50	111.85	108.00

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	4	255	ASP	Peptide
4	6	116	THR	Peptide
7	7	321	GLU	Peptide
12	A	465	TYR	Peptide
12	A	71	GLN	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	1	2411	0	1880	129	0
2	4	2041	0	1954	112	0
3	0	6108	0	6168	424	0
4	6	2527	0	2321	181	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	2	3011	0	2600	177	0
6	5	498	0	506	49	0
7	7	4447	0	3905	424	0
8	3	860	0	622	35	0
9	O	1416	0	1493	79	0
10	N	1178	0	642	52	0
11	T	1159	0	652	17	0
12	A	10997	0	11081	418	0
13	B	9132	0	9146	330	0
14	C	2061	0	2029	60	0
15	D	1253	0	1275	66	0
16	E	1744	0	1772	78	0
17	F	670	0	690	18	0
18	G	1340	0	1357	78	0
19	H	1089	0	1062	57	0
20	I	944	0	899	44	0
21	J	532	0	542	34	0
22	K	904	0	911	34	0
23	L	358	0	381	27	0
24	M	2175	0	2283	95	0
25	Q	1144	0	1034	64	0
26	R	1303	0	1110	59	0
27	U	366	0	372	41	0
28	V	389	0	394	31	0
29	W	1469	0	1432	78	0
30	X	984	0	722	26	0
31	3	2	0	0	0	0
31	4	1	0	0	0	0
31	6	4	0	0	0	0
31	A	3	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
31	X	1	0	0	0	0
32	0	8	0	0	3	0
33	A	1	0	0	0	0
All	All	64538	0	61235	3081	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 25.

The worst 5 of 3081 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:0:133:CYS:HB2	32:0:801:SF4:S4	1.97	1.03
7:7:303:ARG:H	7:7:323:VAL:HG13	1.29	0.97
7:7:477:LEU:HA	7:7:482:TRP:HE1	1.36	0.91
7:7:592:GLN:HE22	7:7:747:ASN:HB3	1.38	0.89
7:7:234:VAL:H	7:7:316:PHE:H	1.20	0.88

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	1	256/542 (47%)	236 (92%)	18 (7%)	2 (1%)	19	58
2	4	279/338 (82%)	237 (85%)	42 (15%)	0	100	100
3	0	752/778 (97%)	677 (90%)	75 (10%)	0	100	100
4	6	336/461 (73%)	297 (88%)	37 (11%)	2 (1%)	25	63
5	2	456/513 (89%)	406 (89%)	50 (11%)	0	100	100
6	5	64/72 (89%)	54 (84%)	10 (16%)	0	100	100
7	7	630/843 (75%)	541 (86%)	89 (14%)	0	100	100
8	3	136/321 (42%)	122 (90%)	14 (10%)	0	100	100
9	O	178/240 (74%)	173 (97%)	5 (3%)	0	100	100
12	A	1386/1733 (80%)	1269 (92%)	115 (8%)	2 (0%)	51	84
13	B	1133/1224 (93%)	1057 (93%)	74 (6%)	2 (0%)	47	79
14	C	260/318 (82%)	243 (94%)	17 (6%)	0	100	100
15	D	153/221 (69%)	142 (93%)	11 (7%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
16	E	211/215 (98%)	205 (97%)	6 (3%)	0	100	100
17	F	81/155 (52%)	78 (96%)	3 (4%)	0	100	100
18	G	169/171 (99%)	156 (92%)	13 (8%)	0	100	100
19	H	132/146 (90%)	118 (89%)	13 (10%)	1 (1%)	19	58
20	I	114/122 (93%)	101 (89%)	13 (11%)	0	100	100
21	J	63/70 (90%)	54 (86%)	9 (14%)	0	100	100
22	K	110/120 (92%)	105 (96%)	5 (4%)	0	100	100
23	L	43/70 (61%)	33 (77%)	10 (23%)	0	100	100
24	M	273/345 (79%)	243 (89%)	30 (11%)	0	100	100
25	Q	140/735 (19%)	130 (93%)	10 (7%)	0	100	100
26	R	176/400 (44%)	167 (95%)	9 (5%)	0	100	100
27	U	42/286 (15%)	38 (90%)	4 (10%)	0	100	100
28	V	46/122 (38%)	44 (96%)	2 (4%)	0	100	100
29	W	189/482 (39%)	183 (97%)	6 (3%)	0	100	100
30	X	152/328 (46%)	139 (91%)	12 (8%)	1 (1%)	22	61
All	All	7960/11371 (70%)	7248 (91%)	702 (9%)	10 (0%)	54	84

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	6	411	PRO
13	B	364	ILE
1	1	230	PRO
4	6	425	SER
13	B	363	HIS

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	1	169/395 (43%)	150 (89%)	19 (11%)	6	25

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	4	198/298 (66%)	171 (86%)	27 (14%)	3	21
3	0	686/707 (97%)	604 (88%)	82 (12%)	5	23
4	6	247/406 (61%)	209 (85%)	38 (15%)	2	16
5	2	258/468 (55%)	210 (81%)	48 (19%)	1	10
6	5	53/66 (80%)	42 (79%)	11 (21%)	1	6
7	7	414/737 (56%)	323 (78%)	91 (22%)	1	6
8	3	53/303 (18%)	50 (94%)	3 (6%)	20	49
9	O	152/205 (74%)	144 (95%)	8 (5%)	22	51
12	A	1221/1520 (80%)	1101 (90%)	120 (10%)	8	29
13	B	995/1061 (94%)	885 (89%)	110 (11%)	6	26
14	C	230/274 (84%)	207 (90%)	23 (10%)	7	29
15	D	139/200 (70%)	124 (89%)	15 (11%)	6	27
16	E	195/197 (99%)	176 (90%)	19 (10%)	8	30
17	F	73/137 (53%)	68 (93%)	5 (7%)	16	44
18	G	152/152 (100%)	139 (91%)	13 (9%)	10	37
19	H	119/128 (93%)	103 (87%)	16 (13%)	4	21
20	I	110/116 (95%)	97 (88%)	13 (12%)	5	24
21	J	60/65 (92%)	54 (90%)	6 (10%)	7	29
22	K	97/102 (95%)	87 (90%)	10 (10%)	7	28
23	L	40/57 (70%)	33 (82%)	7 (18%)	2	12
24	M	245/299 (82%)	220 (90%)	25 (10%)	7	28
25	Q	109/641 (17%)	102 (94%)	7 (6%)	17	45
26	R	107/363 (30%)	99 (92%)	8 (8%)	13	41
27	U	40/260 (15%)	38 (95%)	2 (5%)	24	52
28	V	47/108 (44%)	45 (96%)	2 (4%)	29	56
29	W	155/429 (36%)	148 (96%)	7 (4%)	27	55
30	X	62/295 (21%)	61 (98%)	1 (2%)	62	79
All	All	6426/9989 (64%)	5690 (88%)	736 (12%)	9	25

5 of 736 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	B	244	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
14	C	176	ILE
13	B	346	GLU
13	B	241	ARG
13	B	916	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
23	L	53	HIS
24	M	114	GLN
7	7	697	ASN
7	7	682	GLN
24	M	127	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 21 ligands modelled in this entry, 20 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
32	SF4	0	801	3	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
32	SF4	0	801	3	-	-	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
32	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
1	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	84.64
1	1	355:UNK	C	368:UNK	N	13.08
1	1	519:UNK	C	537:GLU	N	12.00

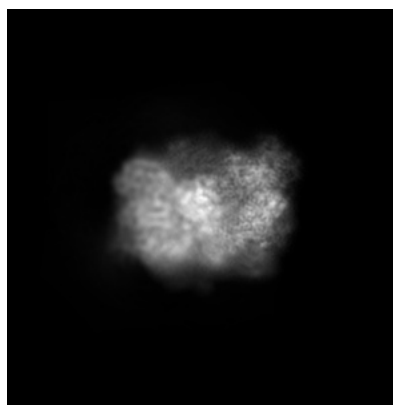
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-23905. These allow visual inspection of the internal detail of the map and identification of artifacts.

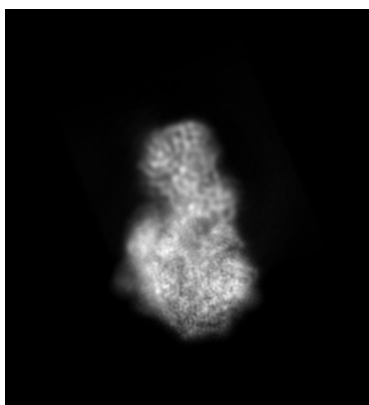
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

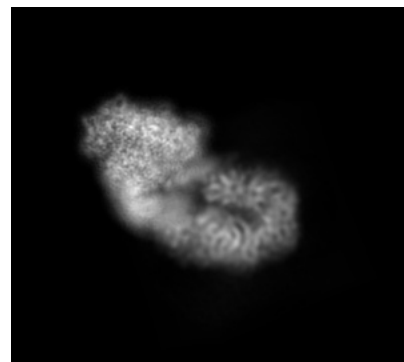
6.1.1 Primary map



X



Y

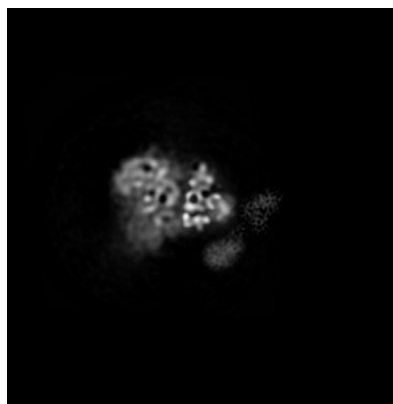


Z

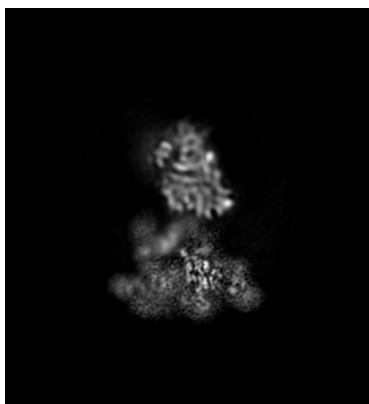
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

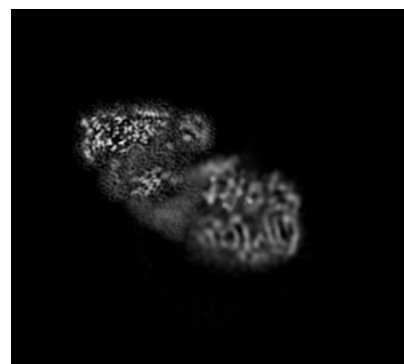
6.2.1 Primary map



X Index: 214



Y Index: 191

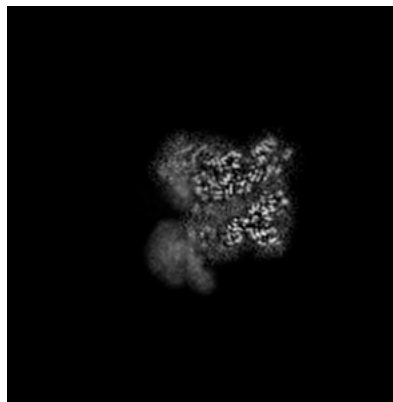


Z Index: 198

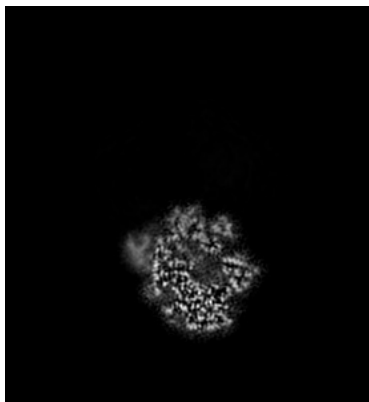
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

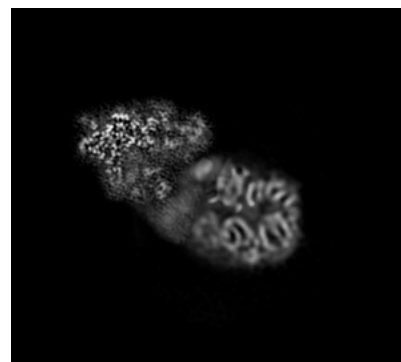
6.3.1 Primary map



X Index: 131



Y Index: 252

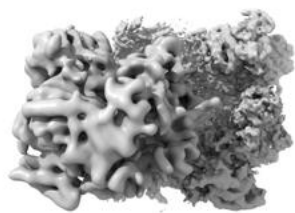


Z Index: 192

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

6.4 Orthogonal surface views [i](#)

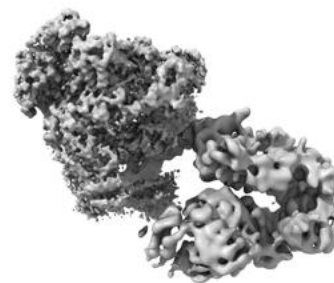
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0135. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

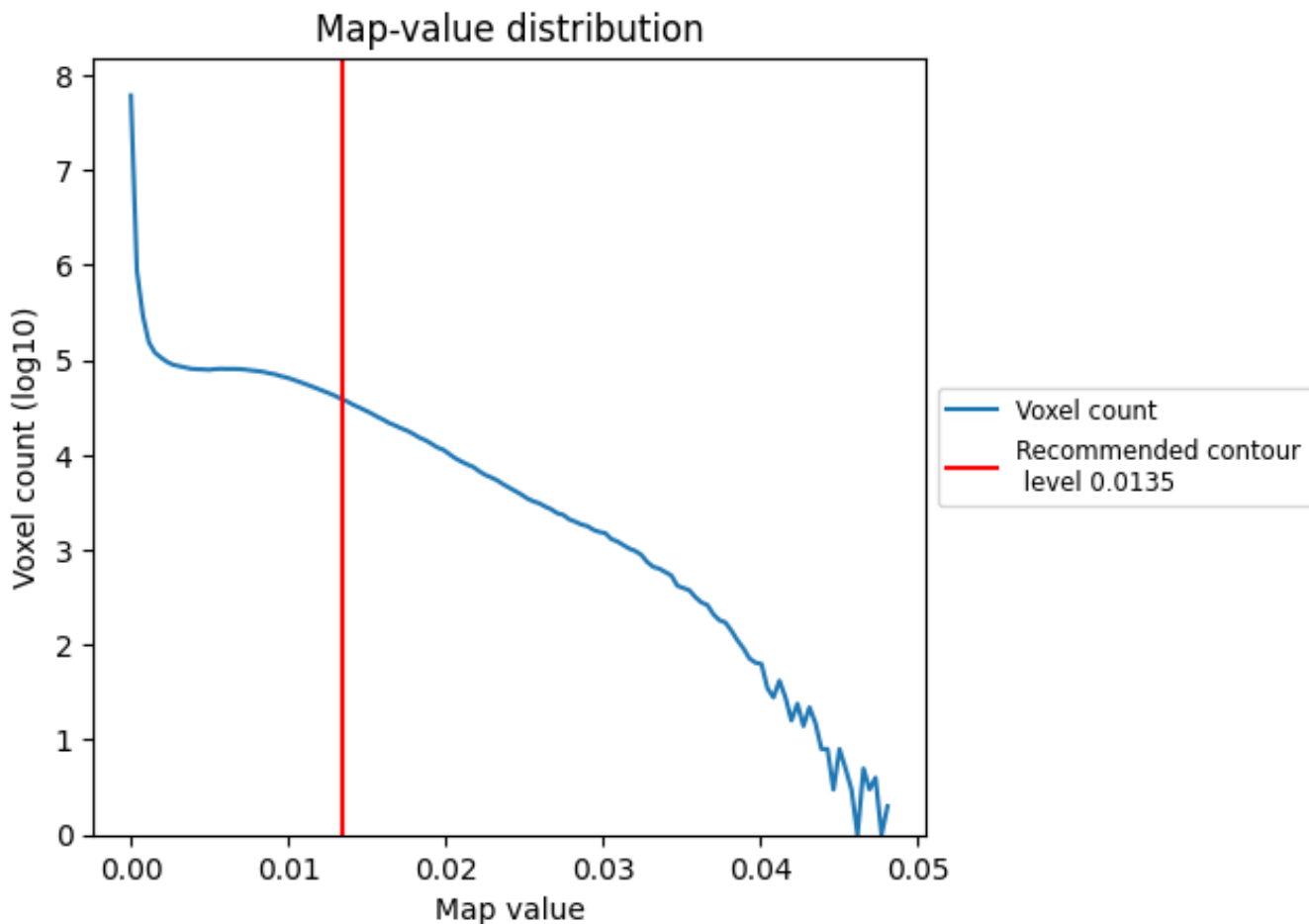
6.5 Mask visualisation

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

7 Map analysis [i](#)

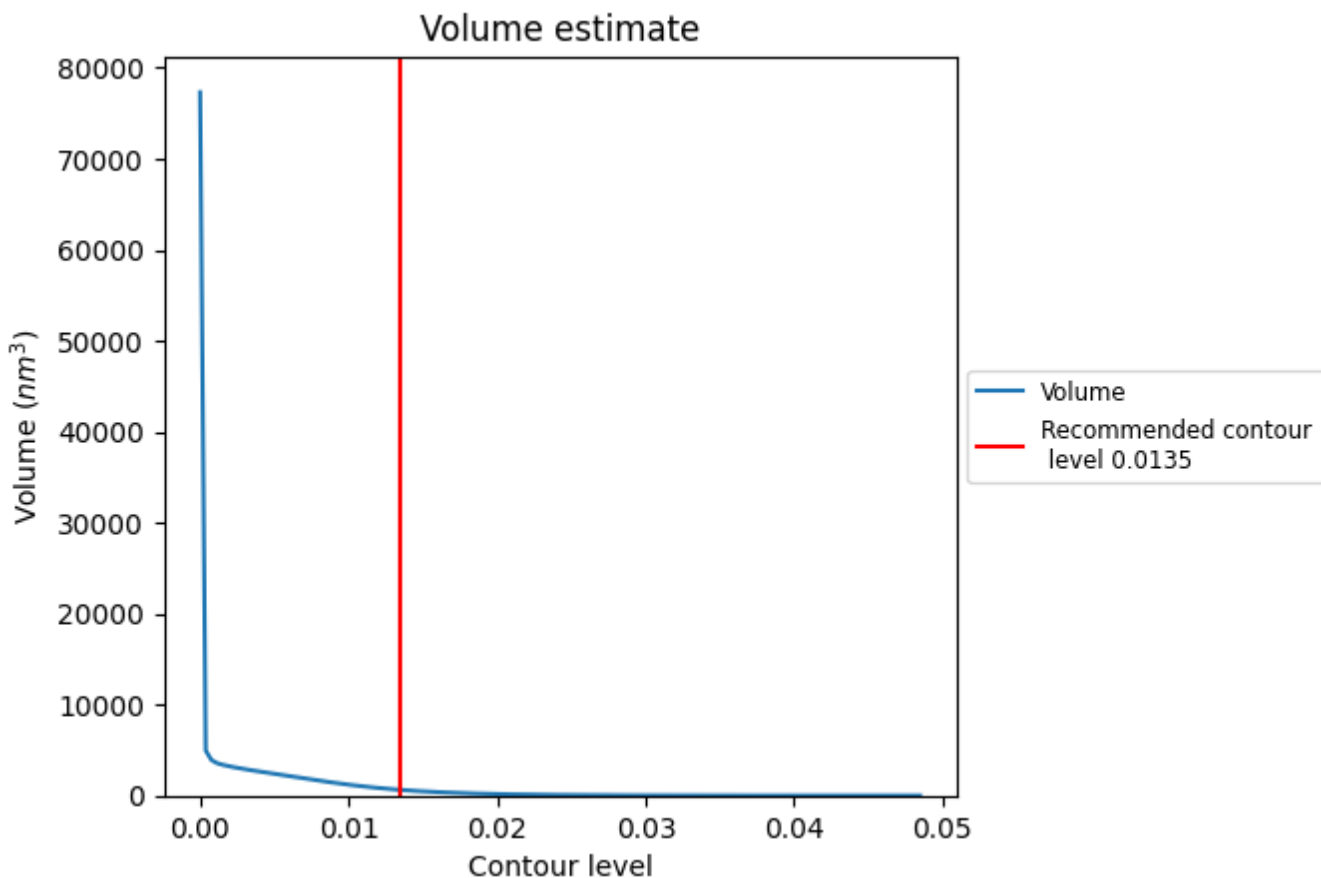
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



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 626 nm³; this corresponds to an approximate mass of 565 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.

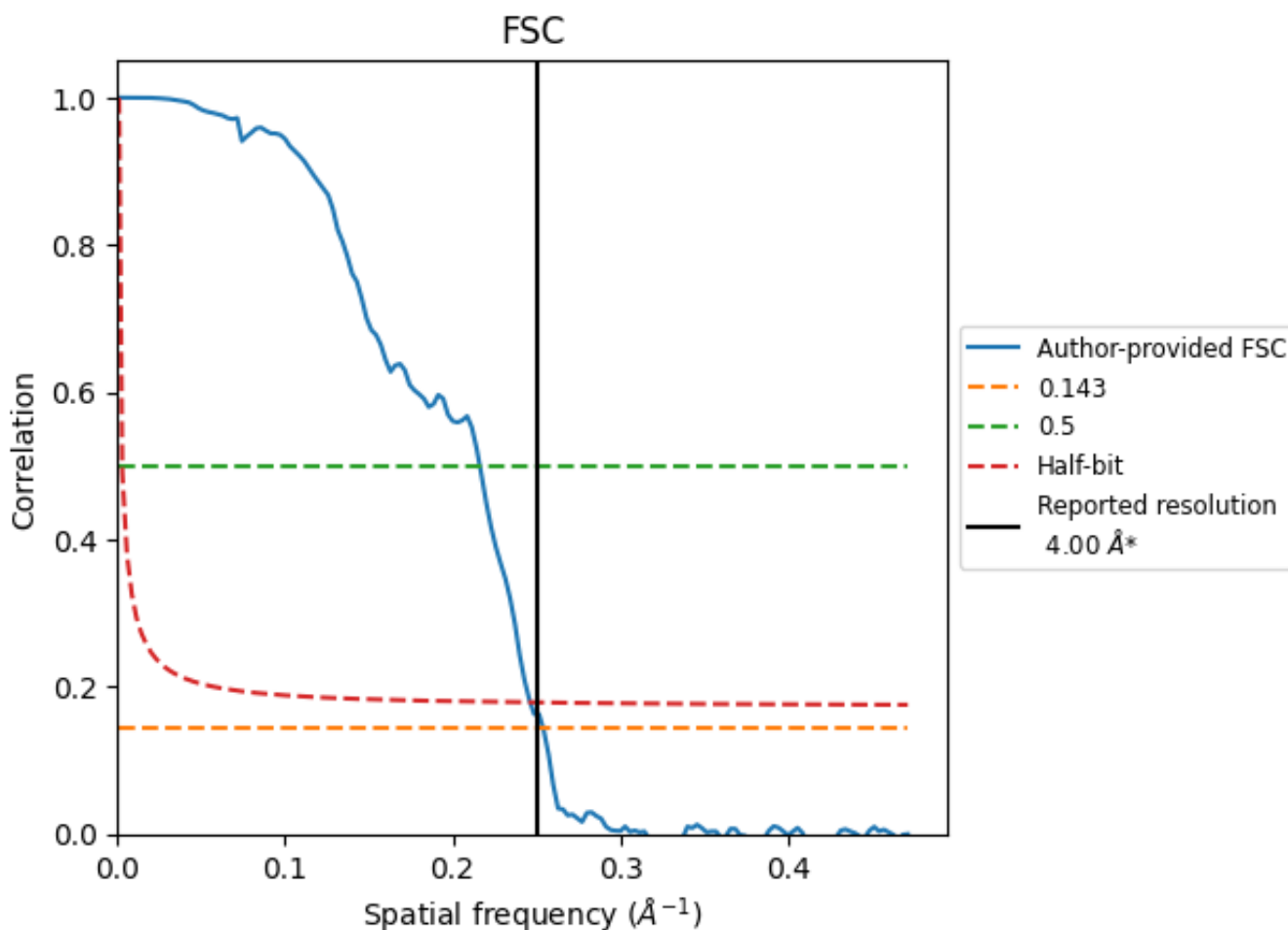
7.3 Rotationally averaged power spectrum [i](#)

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

8 Fourier-Shell correlation [i](#)

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 [i](#)



*Reported resolution corresponds to spatial frequency of 0.250 Å⁻¹

8.2 Resolution estimates [i](#)

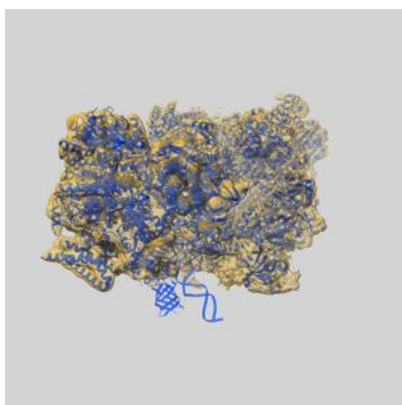
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.00	-	-
Author-provided FSC curve	3.94	4.63	4.06
Unmasked-calculated*	-	-	-

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

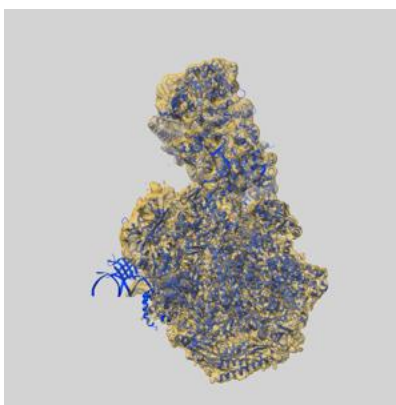
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-23905 and PDB model 7ML1. Per-residue inclusion information can be found in section 3 on page 11.

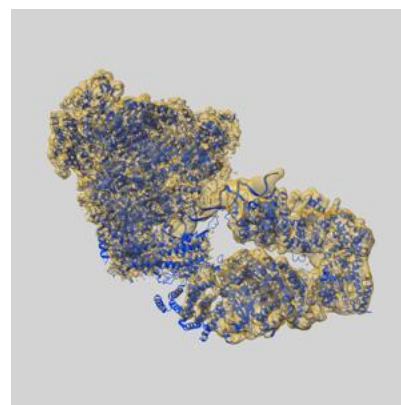
9.1 Map-model overlay [i](#)



X



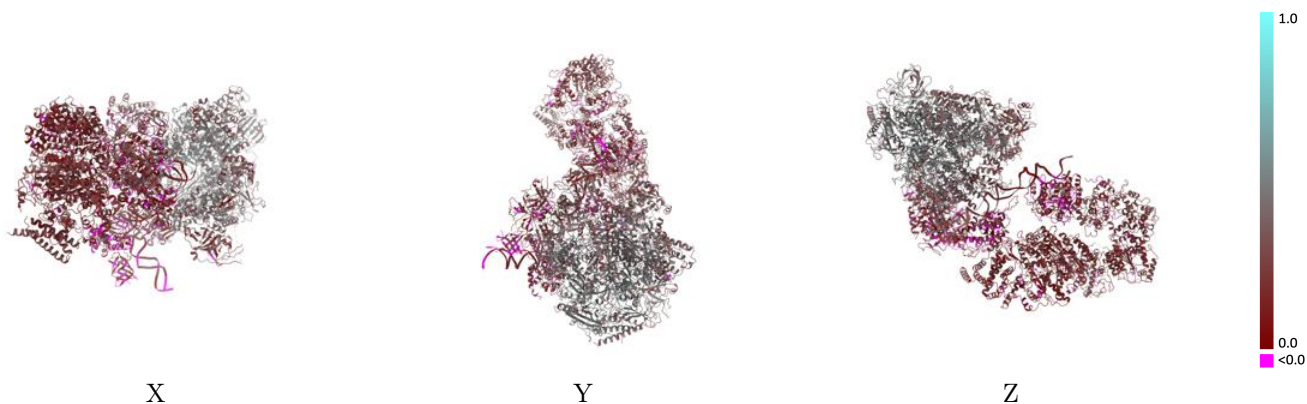
Y



Z

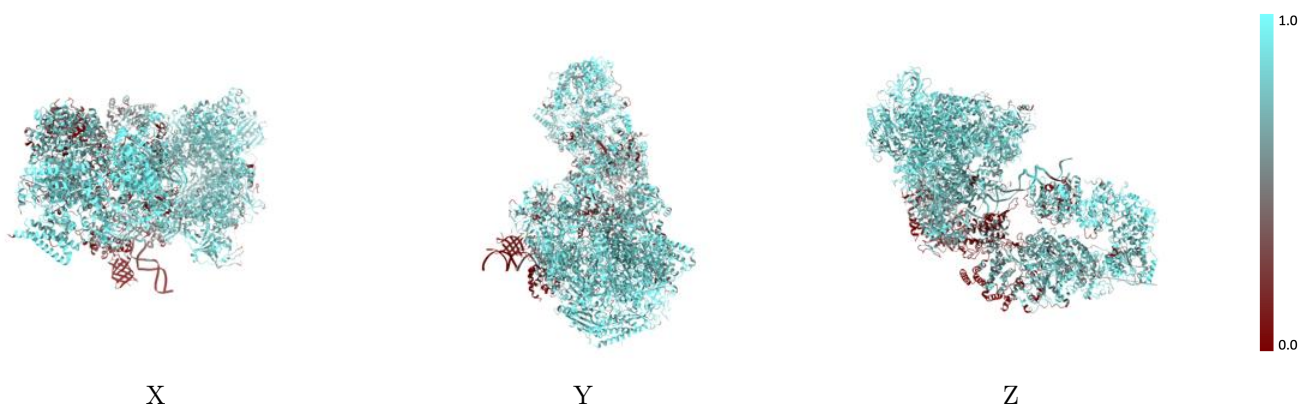
The images above show the 3D surface view of the map at the recommended contour level 0.0135 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [i](#)



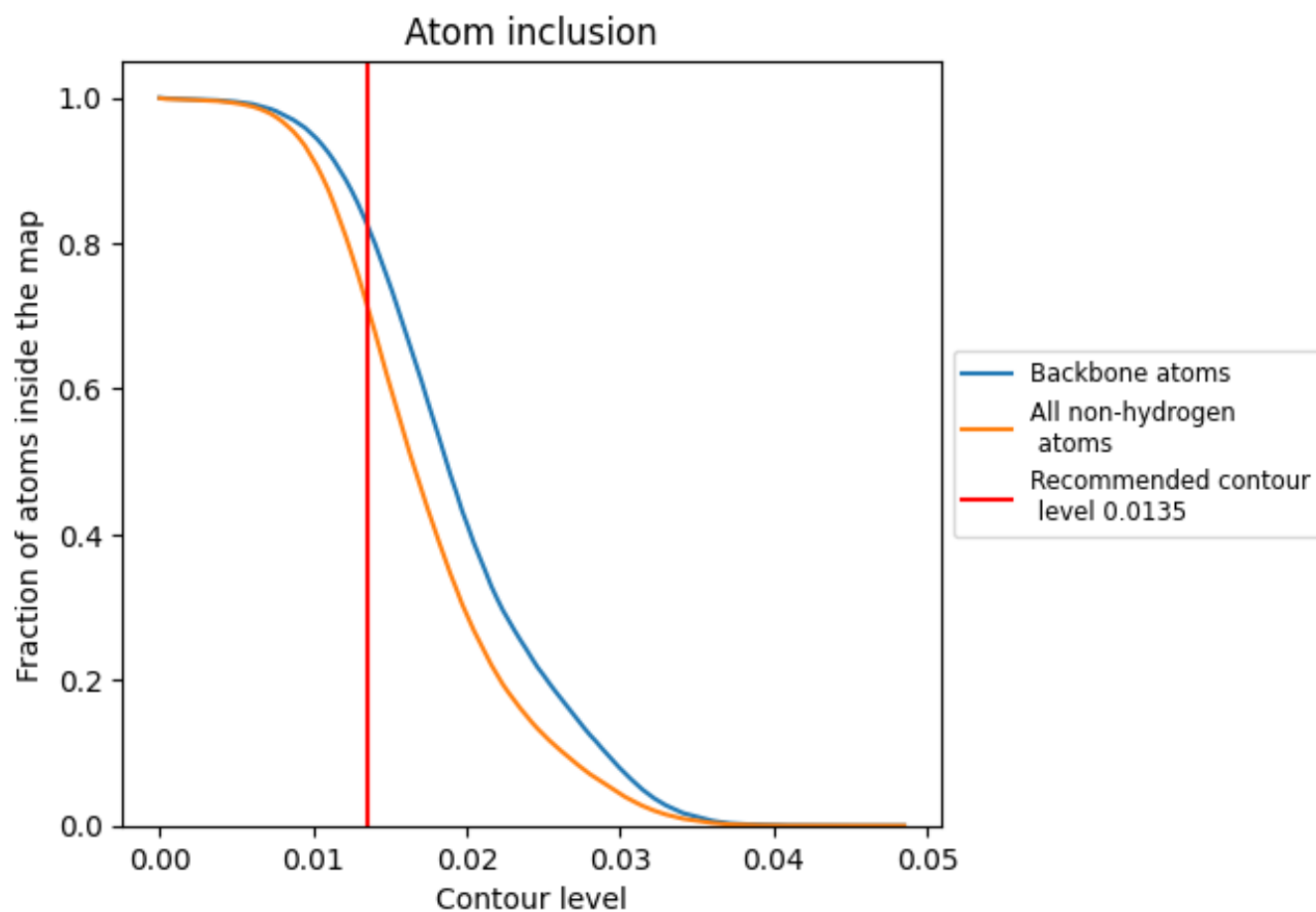
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 [i](#)



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 (0.0135).































































9.4 Atom inclusion [i](#)



At the recommended contour level, 83% of all backbone atoms, 72% 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 (0.0135) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7155	 0.2910
0	 0.6338	 0.1760
1	 0.6446	 0.2160
2	 0.8023	 0.1990
3	 0.0820	 0.1450
4	 0.7705	 0.1940
5	 0.7154	 0.1830
6	 0.7741	 0.2030
7	 0.7138	 0.1470
A	 0.8464	 0.4220
B	 0.8510	 0.4330
C	 0.8899	 0.4570
D	 0.5215	 0.2800
E	 0.8377	 0.4000
F	 0.8804	 0.4500
G	 0.7239	 0.3180
H	 0.8464	 0.4060
I	 0.7492	 0.3410
J	 0.8665	 0.4430
K	 0.7823	 0.3930
L	 0.8006	 0.3570
M	 0.3863	 0.2450
N	 0.6129	 0.1410
O	 0.3703	 0.0900
Q	 0.7418	 0.2460
R	 0.6117	 0.2190
T	 0.6549	 0.1440
U	 0.0056	 0.0640
V	 0.0182	 0.0760
W	 0.4293	 0.1480
X	 0.4157	 0.1520

