



wwPDB EM Validation Summary Report ⓘ

Aug 5, 2026 – 05:00 AM EDT

PDB ID : 7ML4 / pdb_00007ml4
EMDB ID : EMD-23908
Title : RNA polymerase II initially transcribing complex (ITC)
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 : 3.10 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:

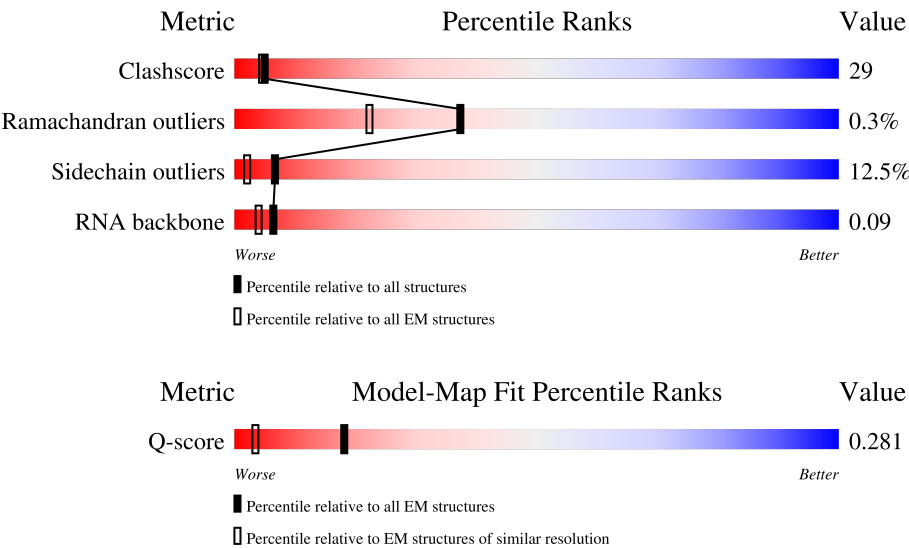
EMDB validation analysis : 0.0.1.dev133
Mogul : 2022.3.0, CSD as543be (2022)
MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : 202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

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 3.10 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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
RNA backbone	8273	3508	-
Q-score	-	25397	14724 (2.60 - 3.60)

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	Q	735	
2	R	398	
3	D	221	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	G	171	
5	M	345	
6	A	1733	
7	B	1224	
8	C	318	
9	E	215	
10	F	155	
11	H	146	
12	I	122	
13	J	70	
14	K	120	
15	L	70	
16	3	321	
17	0	778	
18	4	338	
19	6	461	
20	1	543	
21	7	843	
22	5	72	
23	2	513	
24	X	328	
25	U	286	
26	V	122	
27	N	38	
28	T	148	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
29	O	240	
30	W	482	
31	P	5	

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
34	SF4	0	801	-	-	X	-

2 Entry composition

There are 34 unique types of molecules in this entry. The entry contains 62733 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 Transcription initiation factor IIF subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	Q	148	Total	C	N	O	S	0	0
			1141	731	195	212	3		

- Molecule 2 is a protein called Transcription initiation factor IIF subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	R	154	Total	C	N	O	S	0	0
			1039	652	190	193	4		

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB4.

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

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB7.

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

- Molecule 5 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	M	234	Total	C	N	O	S	0	0
			1805	1152	304	333	16		

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	A	1405	Total	C	N	O	S	0	0
			11039	6962	1935	2081	61		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	B	1114	Total	C	N	O	S	0	0
			8861	5610	1549	1647	55		

- Molecule 8 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I,II,and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	F	85	Total	C	N	O	S	0	0
			688	439	116	130	3		

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	H	133	Total	C	N	O	S	0	0
			1068	673	180	211	4		

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases II subunit RPABC5.

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

- Molecule 14 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 15 is a protein called DNA-directed RNA polymerases II subunit RPABC4.

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

- Molecule 16 is a protein called BJ4_G0050160.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	3	72	Total	C	N	O	S	0	0
			361	215	72	74			

- Molecule 17 is a protein called General transcription and DNA repair factor IIIH helicase subunit XPD.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	0	754	Total	C	N	O	S	0	0
			6108	3891	1032	1147	38		

- Molecule 18 is a protein called General transcription and DNA repair factor IIIH subunit TFB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	4	284	Total	C	N	O	S	0	0
			2041	1310	343	376	12		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
4	113	UNK	ASP	conflict	UNP A0A7I9C5C2
4	114	UNK	MET	conflict	UNP A0A7I9C5C2

- Molecule 19 is a protein called General transcription and DNA repair factor IIIH.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	6	351	Total	C	N	O	S	0	0
			2527	1590	454	456	27		

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
6	412	UNK	ILE	conflict	UNP A0A7I9FQL5
6	413	UNK	LEU	conflict	UNP A0A7I9FQL5
6	414	UNK	LYS	conflict	UNP A0A7I9FQL5
6	415	UNK	ASN	conflict	UNP A0A7I9FQL5
6	416	UNK	HIS	conflict	UNP A0A7I9FQL5
6	417	UNK	LYS	conflict	UNP A0A7I9FQL5
6	418	UNK	ASN	conflict	UNP A0A7I9FQL5
6	419	UNK	ASP	conflict	UNP A0A7I9FQL5
6	420	UNK	LYS	conflict	UNP A0A7I9FQL5
6	421	UNK	LEU	conflict	UNP A0A7I9FQL5
6	422	UNK	LEU	conflict	UNP A0A7I9FQL5
6	423	UNK	THR	conflict	UNP A0A7I9FQL5
6	424	UNK	SER	conflict	UNP A0A7I9FQL5

- Molecule 20 is a protein called Tfb1.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	1	367	Total	C	N	O	S	0	0
			2411	1536	438	430	7		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
21	7	634	Total	C	N	O	S	0	0
			4447	2722	827	874	24		

- Molecule 22 is a protein called General transcription and DNA repair factor IIH subunit TFB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	5	66	Total	C	N	O	S	0	0
			498	314	89	93	2		

- Molecule 23 is a protein called RNA polymerase II transcription factor B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	2	460	Total	C	N	O	S	0	0
			3011	1856	562	584	9		

- Molecule 24 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	X	149	Total	C	N	O	S	0	0
			921	569	168	180	4		

- Molecule 25 is a protein called Transcription initiation factor IIA large subunit.

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

- Molecule 26 is a protein called Transcription initiation factor IIA subunit 2.

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

- Molecule 27 is a DNA chain called non-template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	N	38	Total	C	N	O	P	0	0
			791	376	161	216	38		

- Molecule 28 is a DNA chain called template strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	T	48	Total	C	N	O	P	0	0
			968	468	150	302	48		

- Molecule 29 is a protein called TATA-box-binding protein.

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

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

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

- Molecule 31 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	P	5	Total	C	N	O	P	0	0
			110	50	25	31	4		

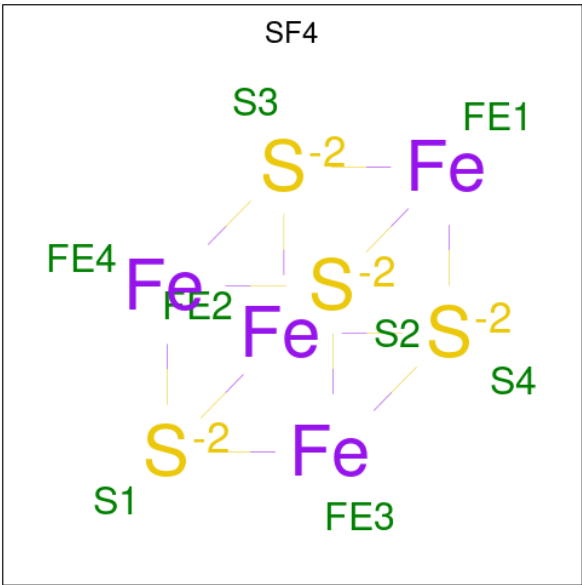
- Molecule 32 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
32	M	1	Total	Zn	0
			1	1	
32	A	2	Total	Zn	0
			2	2	
32	B	1	Total	Zn	0
			1	1	
32	C	1	Total	Zn	0
			1	1	
32	I	2	Total	Zn	0
			2	2	
32	J	1	Total	Zn	0
			1	1	
32	L	1	Total	Zn	0
			1	1	
32	4	1	Total	Zn	0
			1	1	
32	6	4	Total	Zn	0
			4	4	
32	W	1	Total	Zn	0
			1	1	

- Molecule 33 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 34 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe₄S₄).

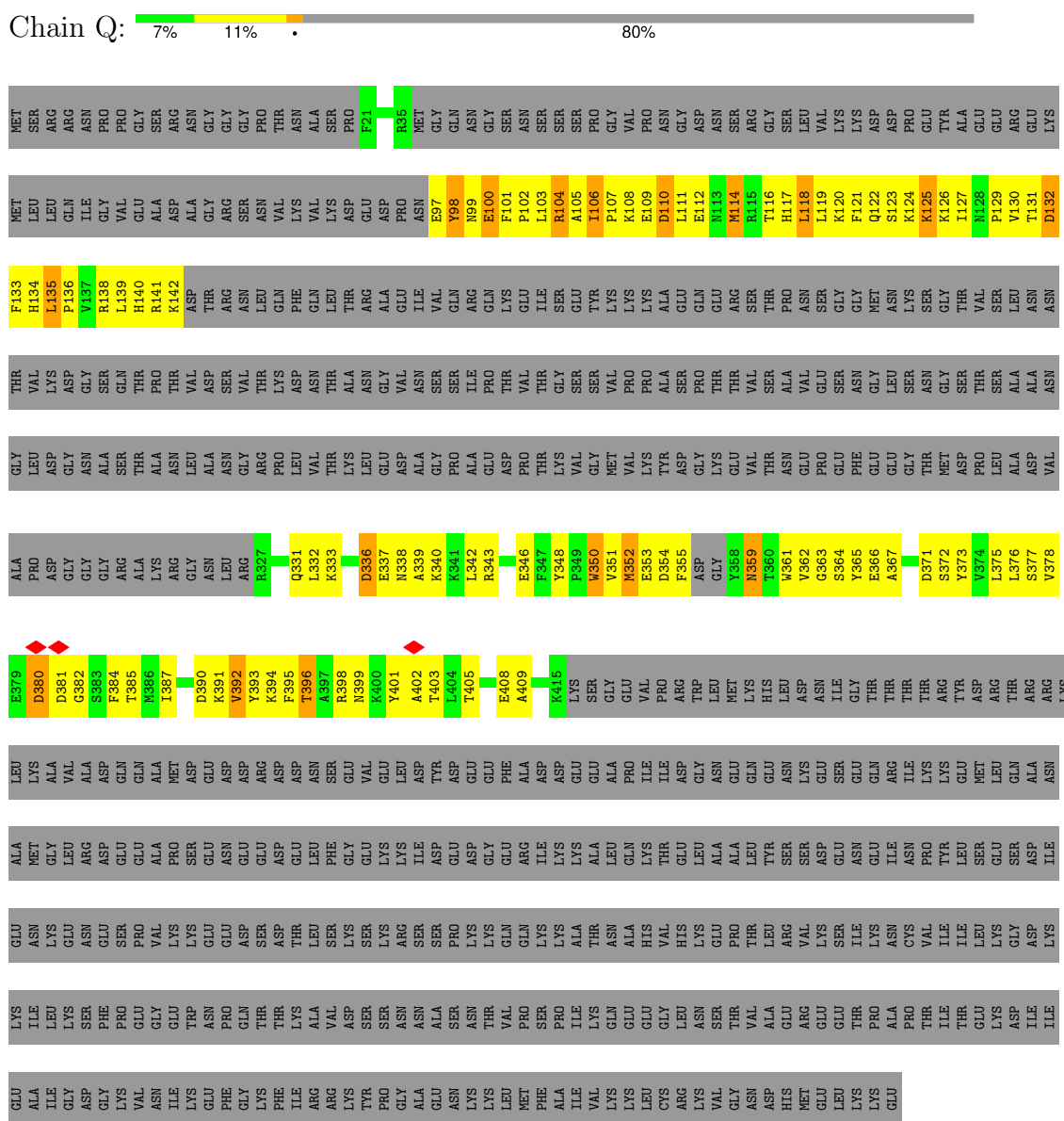


Mol	Chain	Residues	Atoms			AltConf
34	0	1	Total	Fe	S	0
			8	4	4	

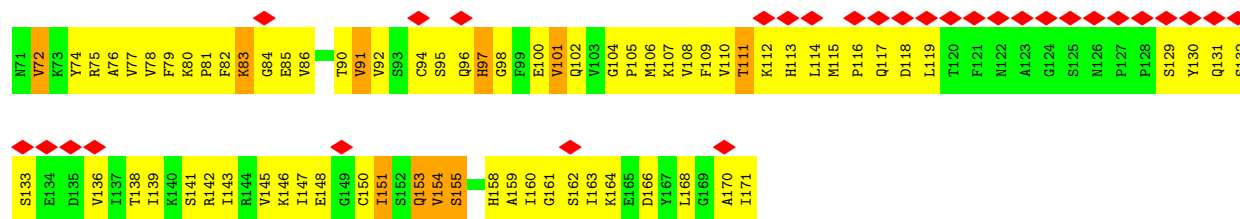
3 Residue-property plots

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.

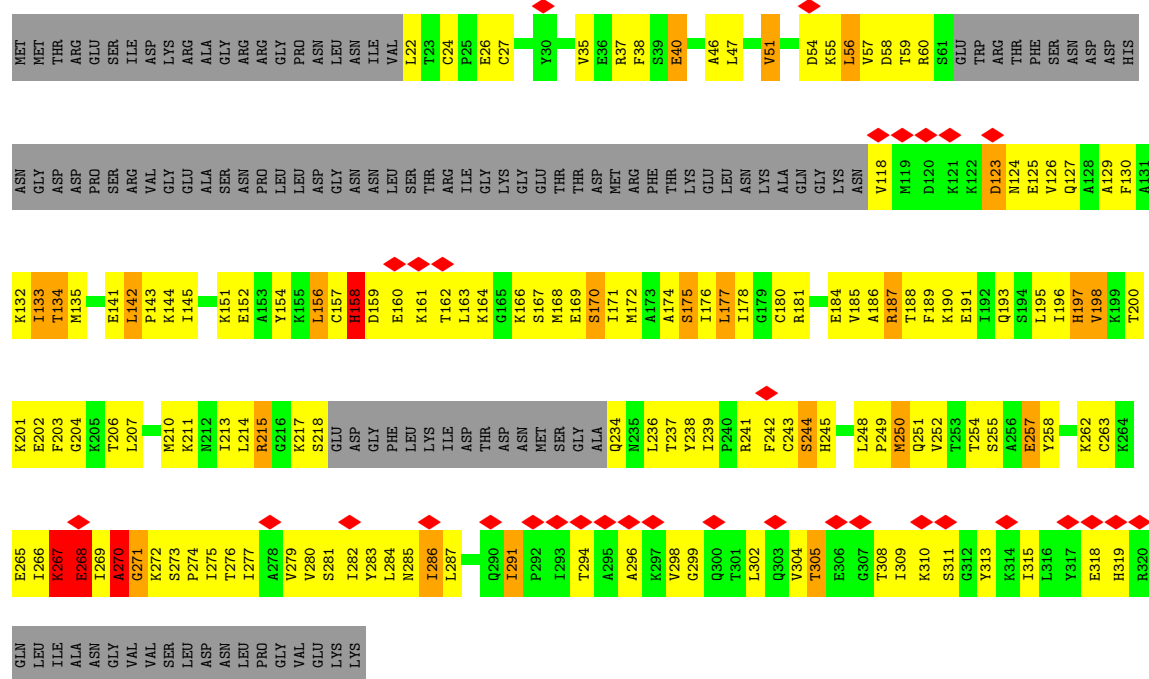
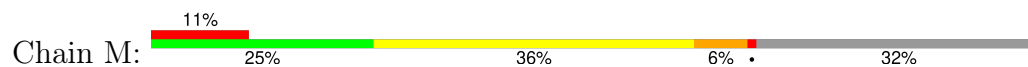
- Molecule 1: Transcription initiation factor IIF subunit alpha



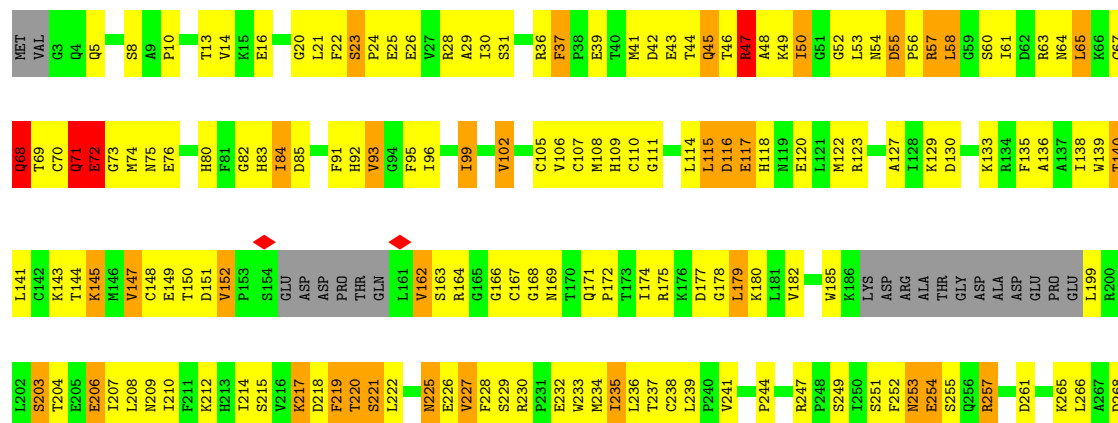
- Molecule 2: Transcription initiation factor IIF subunit beta



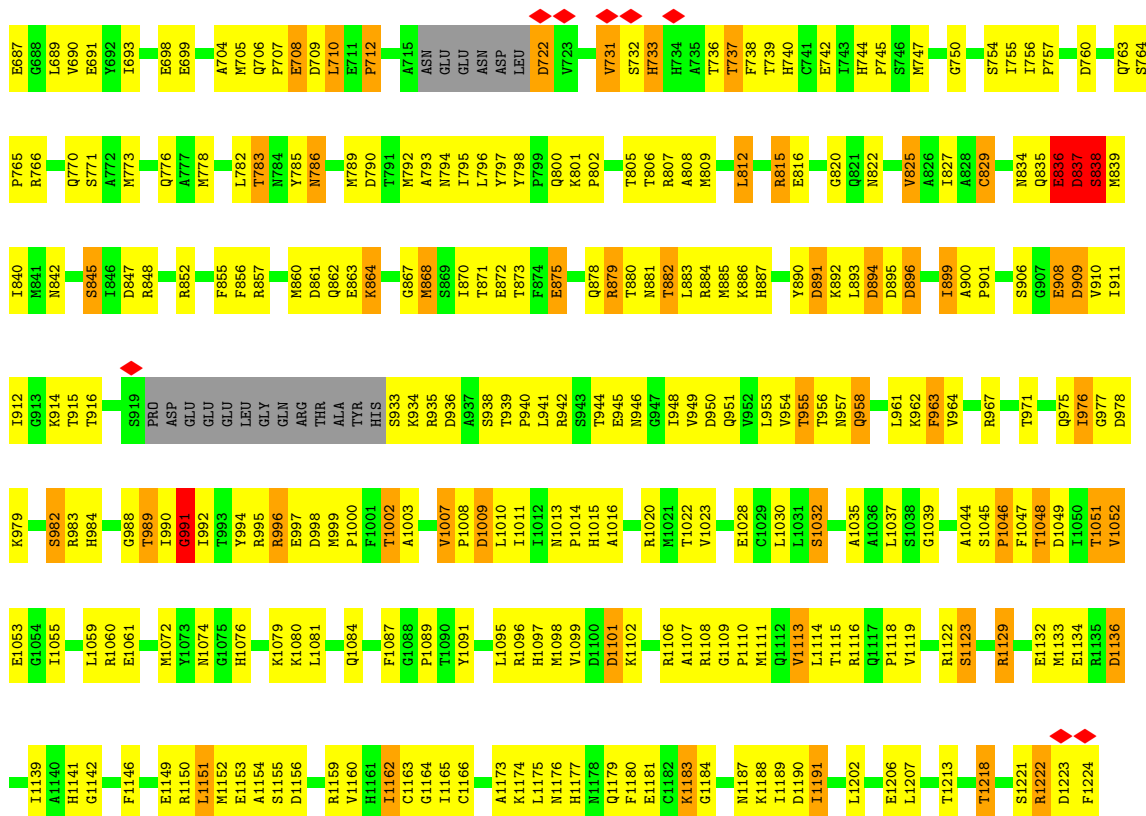
• Molecule 5: Transcription initiation factor IIB



• Molecule 6: DNA-directed RNA polymerase subunit

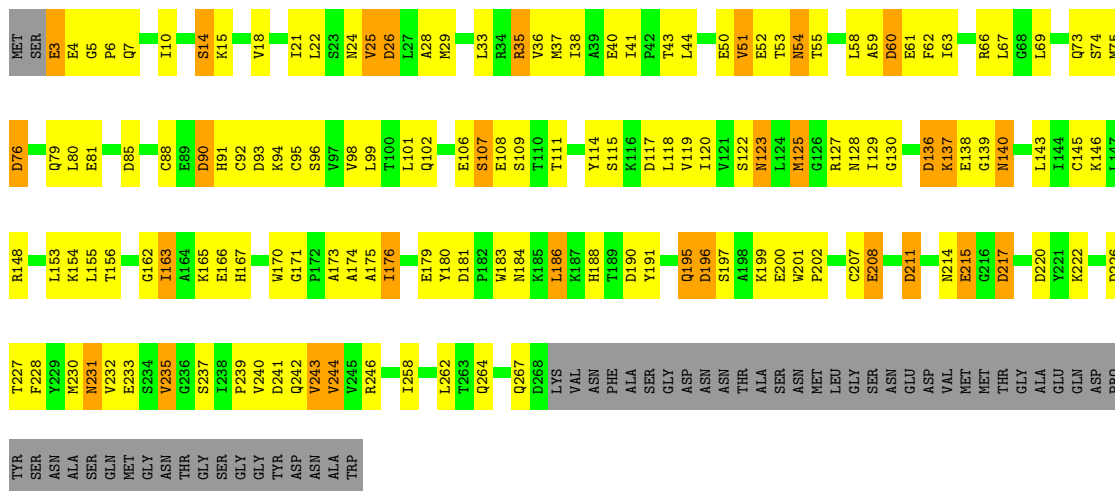


I1327	I1328	I1329	I1330	I1331	I1332	I1333	I1334	I1335	M1336	I1341	I1342	L1348	E1351	S1358	D1359	S1360	S1361	Y1362	V1363	M1364	Y1365	M1368	V1372	T1376	T1377	Q1378	G1379	G1380	G1381	T1382	S1383	R1386	F1389	M1390	R1391	N1392	T1393	T1394	G1395	A1396	L1397	M1399	N1399	C1400	E1404	T1405	V1406	E1407	I1408										
ALA	GLU	THR	GLU	A1254	E1255	E1256	D1257	H1258	K1261	K1262	I1263	E1264	N1265	T1266	M1267	N1270	I1271	T1272	L1273	L1274	G1275	V1276	E1277	N1278	M1284	M1285	K1286	Y1287	K1290	G1291	P1292	G1293	P1294	T1295	G1296	E1297	Y1298	V1299	K1300	E1301	S1302	E1303	Y1304	E1307	A1396	L1397	V1311	N1312	E1315	T1318	V1405	E1407	I1408						
F1190	W1191	L1192	L1133	L1134	L1135	E1196	L1197	D1198	R1199	A1200	A1201	M1202	M1203	D1204	K1205	D1206	T1207	T1208	M1209	G1210	Q1211	T1212	G1213	G1214	R1215	I1216	K1217	Q1218	T1219	F1220	K1221	N1222	D1223	E1224	K1225	L1226	V1227	M1228	T1229	K1230	D1231	M1232	D1233	E1234	K1235	L1236	I1237	I1238	R1239	C1240	R1241	V1242	V1243	ARG	PRO	LYS	SER	LEU	ASP
E1129	Q1130	L1133	L1134	R1135	S1136	L1137	I1138	R1139	H1140	T1141	T1142	L1143	K1144	S1145	T1146	T1147	I1148	A1149	M1150	G1151	E1152	Y1153	Y1154	D1155	P1156	D1157	P1158	R1159	S1160	T1161	V1162	L1163	P1164	E1165	L1166	E1167	T1168	I1169	T1170	Q1171	L1172	H1173	F1174	S1175	L1176	ASP	GLU	GLU	GLU	ALA	GLU	GLN	SER	PHE	ASP	Q1187	Q1188	S1189	
F1190	W1191	L1192	L1133	L1134	L1135	E1196	L1197	D1198	R1199	A1200	A1201	M1202	M1203	D1204	K1205	D1206	T1207	T1208	M1209	G1210	Q1211	T1212	G1213	G1214	R1215	I1216	K1217	Q1218	T1219	F1220	K1221	N1222	D1223	E1224	K1225	L1226	V1227	M1228	T1229	K1230	D1231	M1232	D1233	E1234	K1235	L1236	I1237	I1238	R1239	C1240	R1241	V1242	V1243	ARG	PRO	LYS	SER	LEU	ASP
ALA	GLU	THR	GLU	A1254	E1255	E1256	D1257	H1258	K1261	K1262	I1263	E1264	N1265	T1266	M1267	N1270	I1271	T1272	L1273	L1274	G1275	V1276	E1277	N1278	M1284	M1285	K1286	Y1287	K1290	G1291	P1292	G1293	P1294	T1295	G1296	E1297	Y1298	V1299	K1300	E1301	S1302	E1303	Y1304	E1307	A1396	L1397	V1311	N1312	E1315	T1318	V1405	E1407	I1408						
K419	R420	D423	I424	Q425	L426	Q427	Y428	Q429	W430	K431	W432	E433	R434	H435	I436	D438	M439	D440	P441	N445	R446	Q447	P448	S449	L450	H451	K452	M453	S454	M455	A457	H458	R459	V460	D544	Q545	V546	M549	L550	D555	W556	D557	G558	V559	I560	P561	T562	P563	L566	K567	P568	K569	P570						
L269	L270	N273	S274	L275	L276	E277	T278	L279	E280	H281	H286	H287	A288	I289	E290	E291	E292	F298	H299	V300	A301	T302	D305	G310	Q311	L315	Q316	K317	S318	P321	V322	I325	R326	R328	L329	K332	E333	R407	G334	R335	I336	R337	G338	N339	L340	R344	V345	D346											



• Molecule 8: DNA-directed RNA polymerase II subunit RPB3

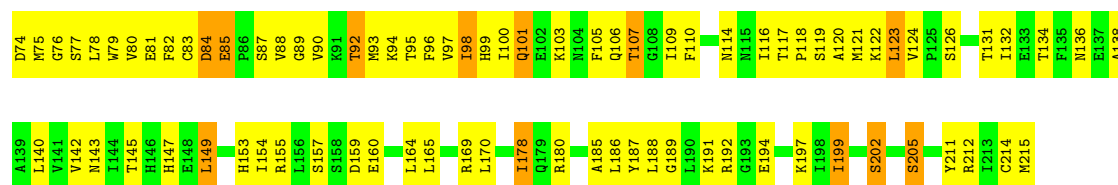
Chain C: 38% 37% 9% 16%



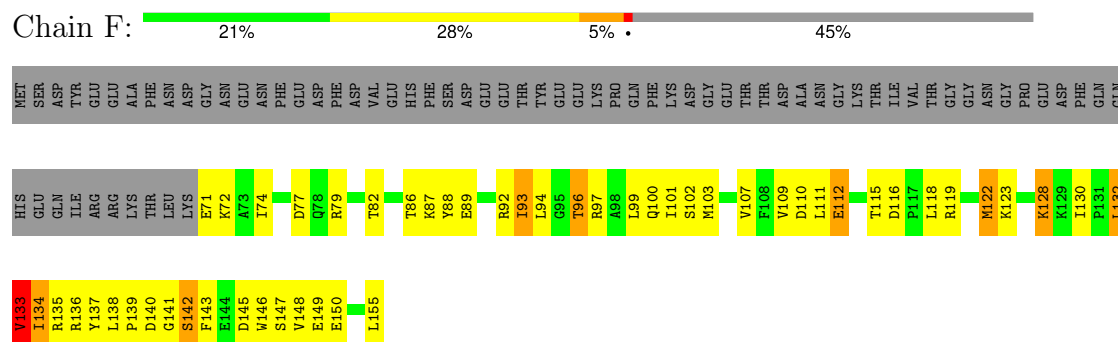
• Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 39% 51% 10%

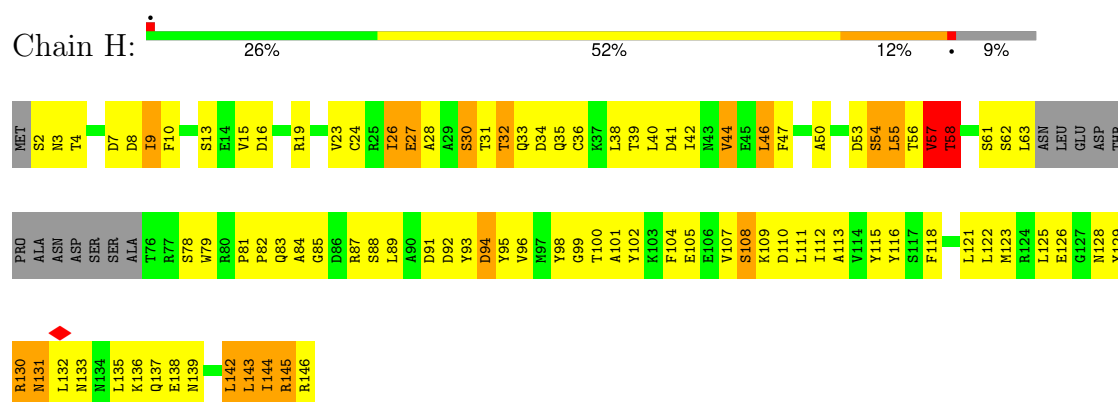




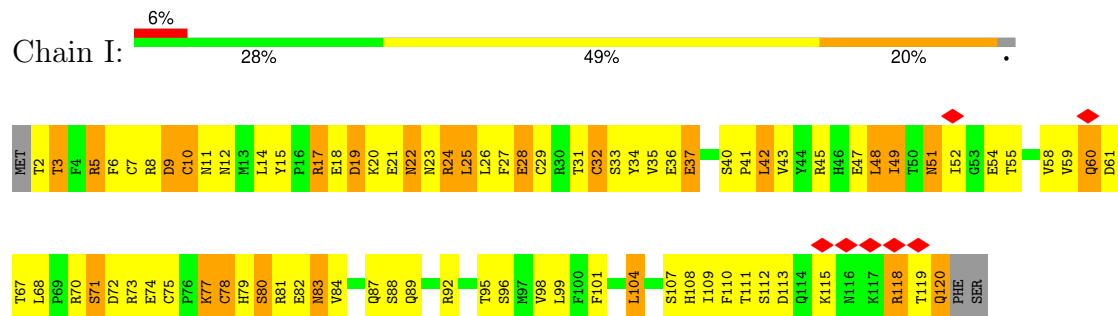
• Molecule 10: DNA-directed RNA polymerases I,II,and III subunit RPABC2



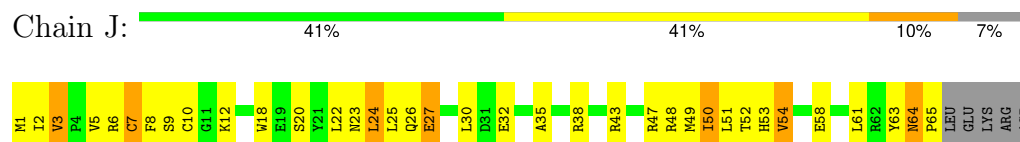
• Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC3



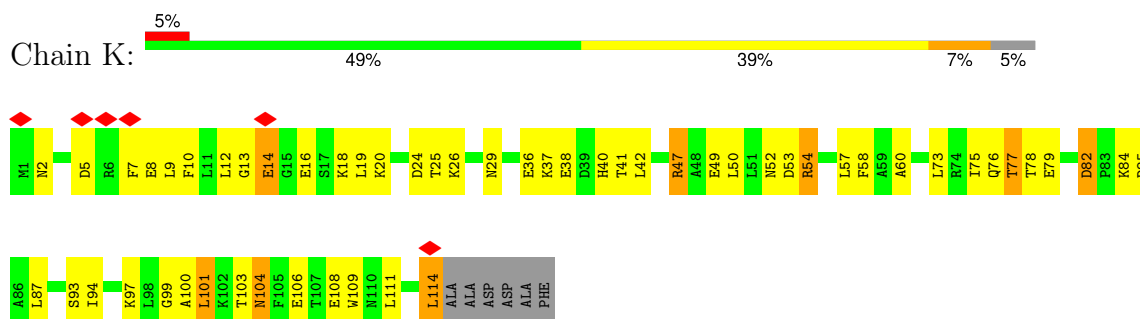
• Molecule 12: DNA-directed RNA polymerase II subunit RPB9



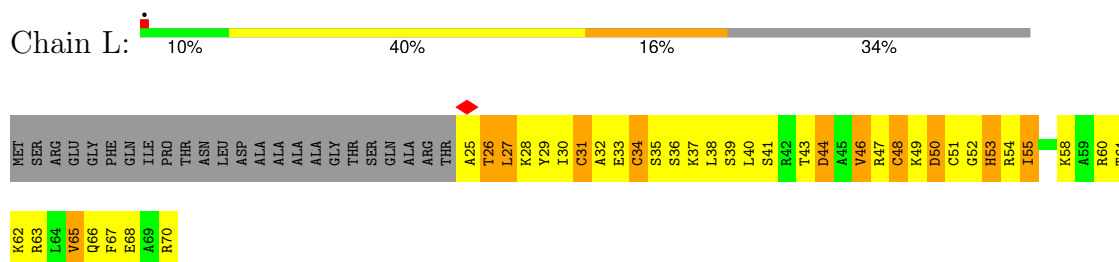
• Molecule 13: DNA-directed RNA polymerases II subunit RPABC5



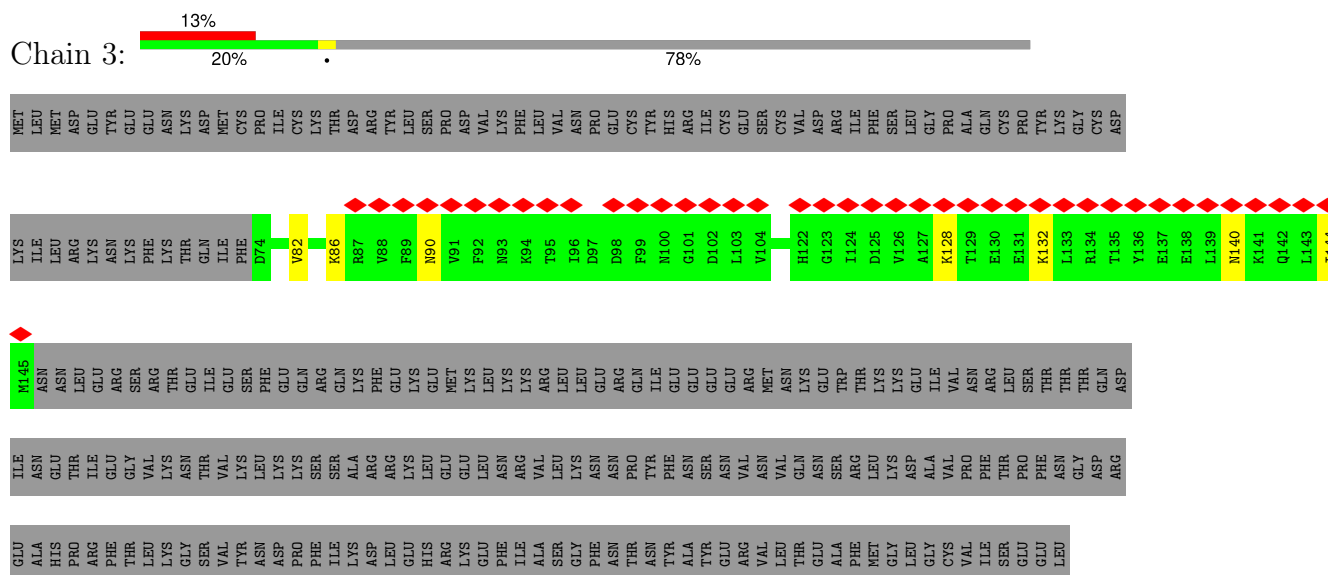
- Molecule 14: DNA-directed RNA polymerase II subunit RPB11

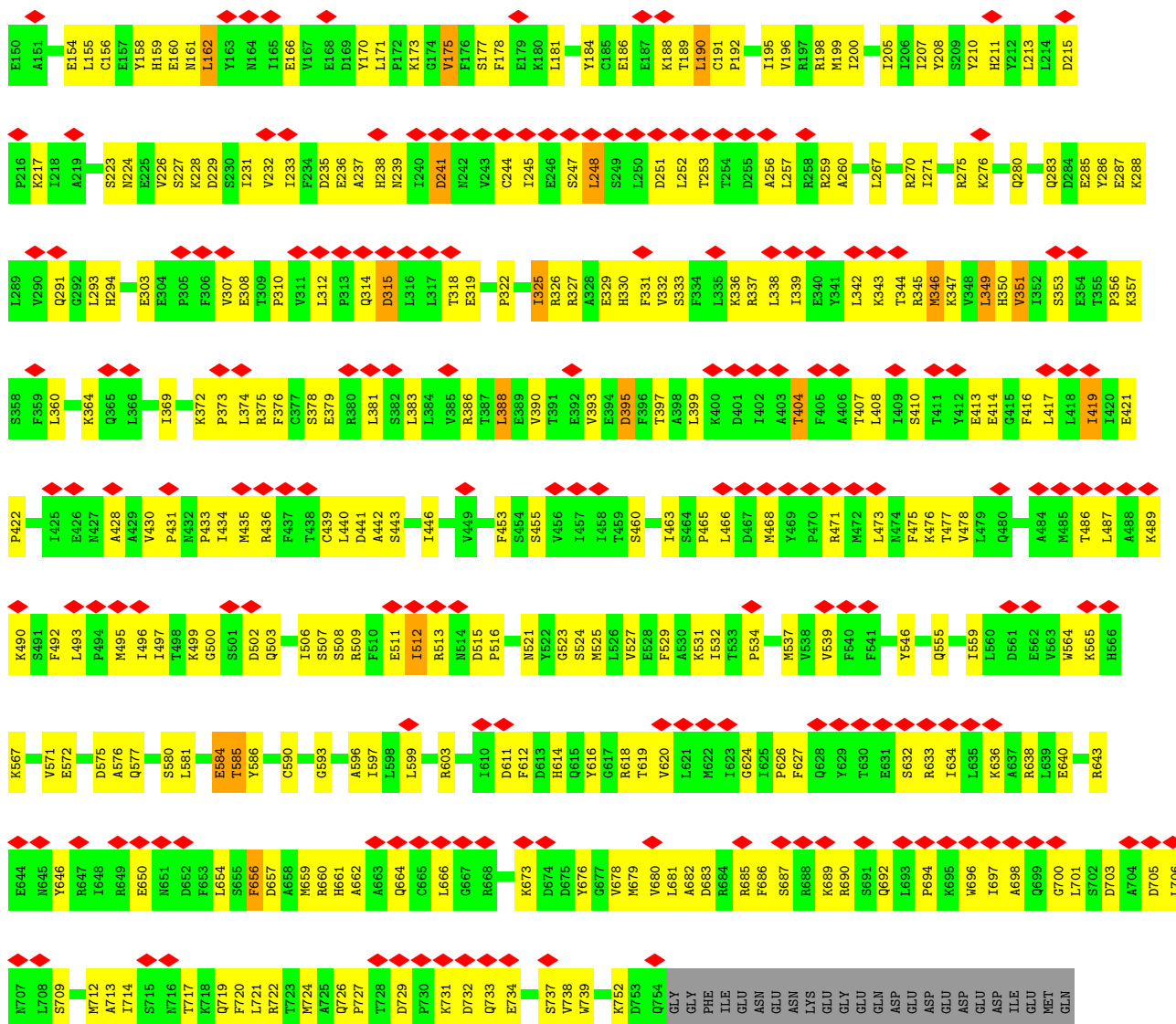


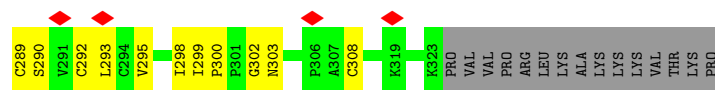
- Molecule 15: DNA-directed RNA polymerases II subunit RPABC4



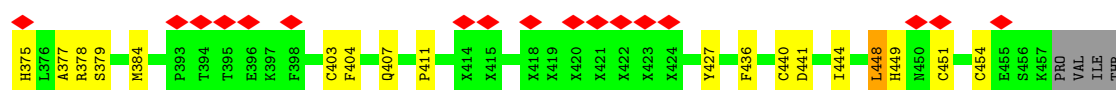
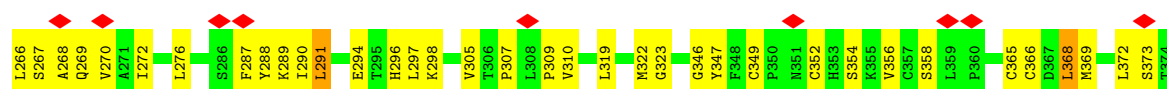
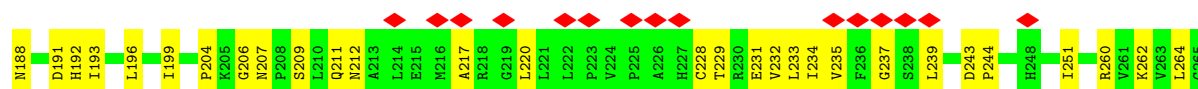
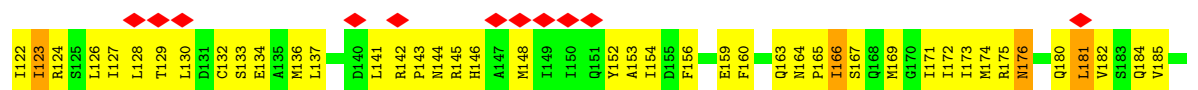
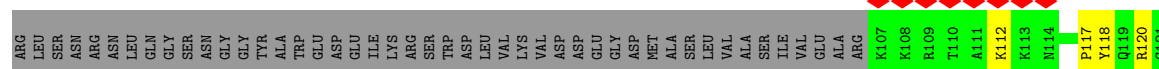
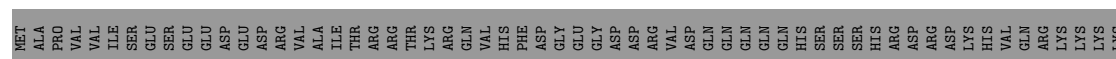
- Molecule 16: BJ4_G0050160.mRNA.1.CDS.1



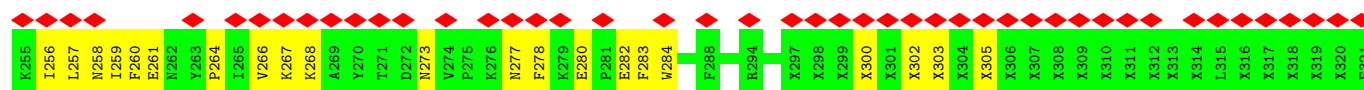
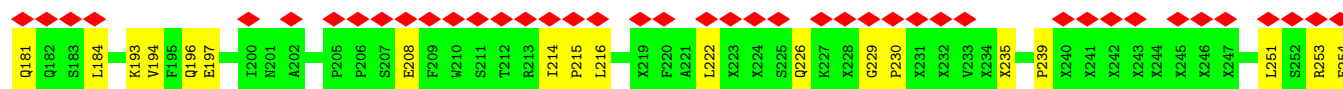
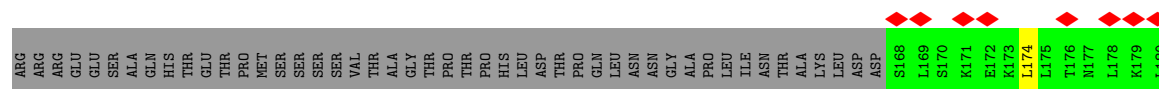
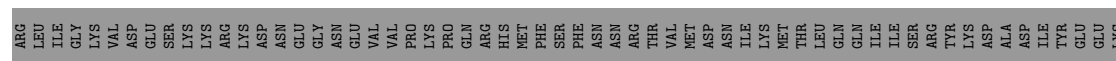




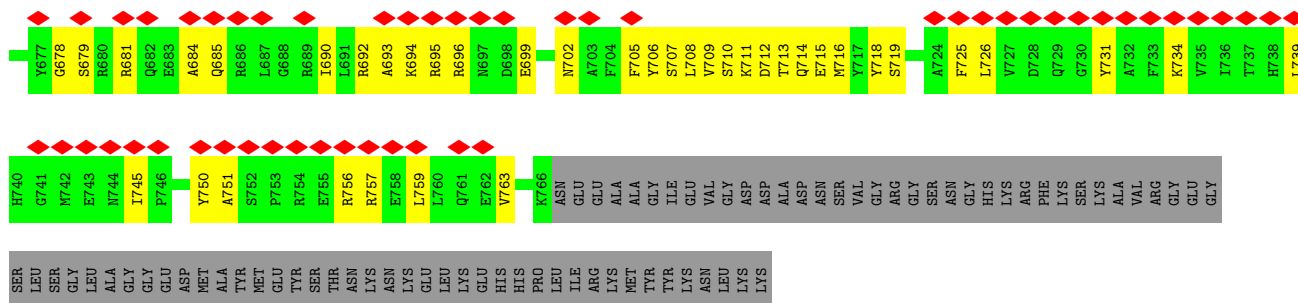
• Molecule 19: General transcription and DNA repair factor IHH



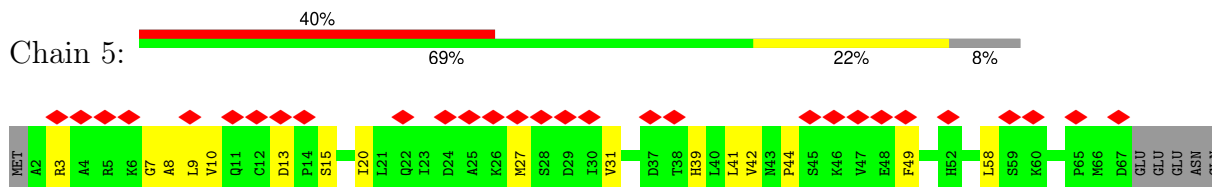
• Molecule 20: Tfb1



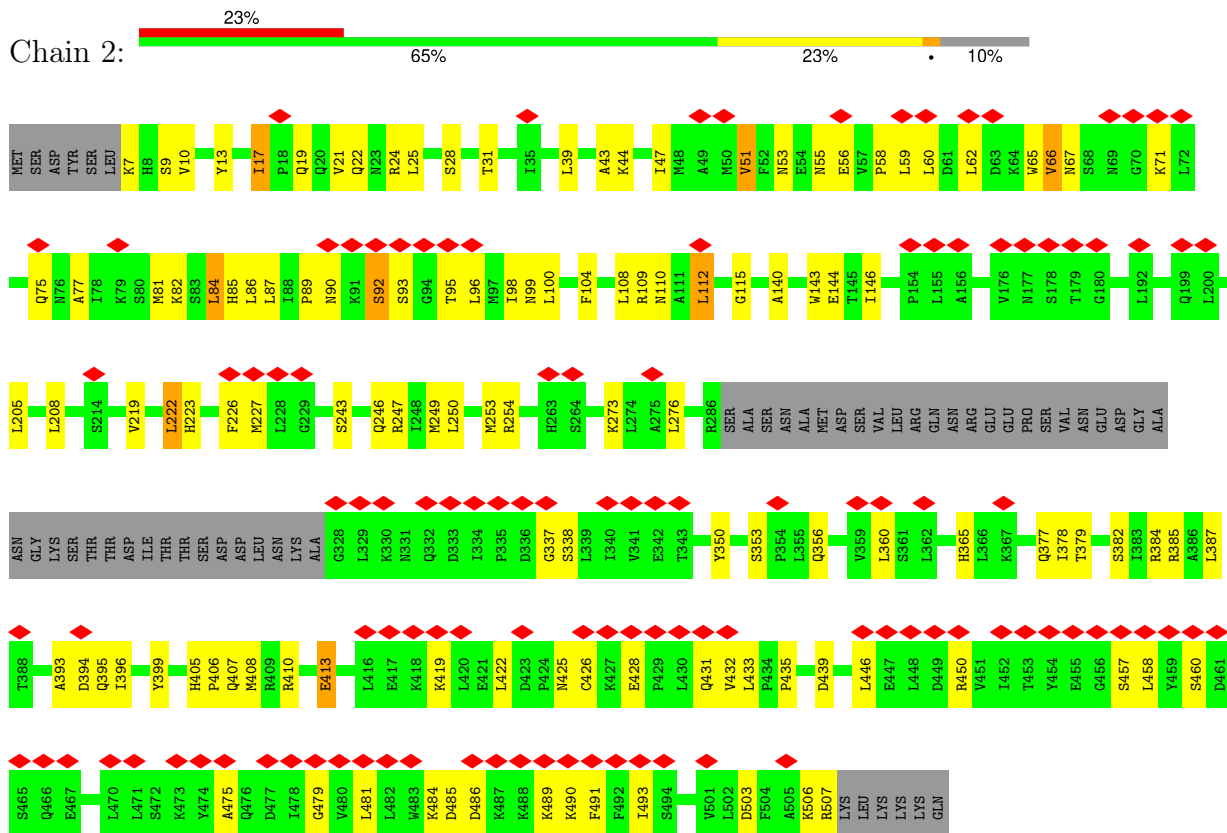




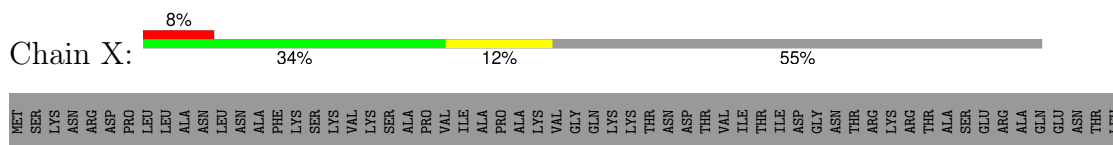
- Molecule 22: General transcription and DNA repair factor IIH subunit TFB5

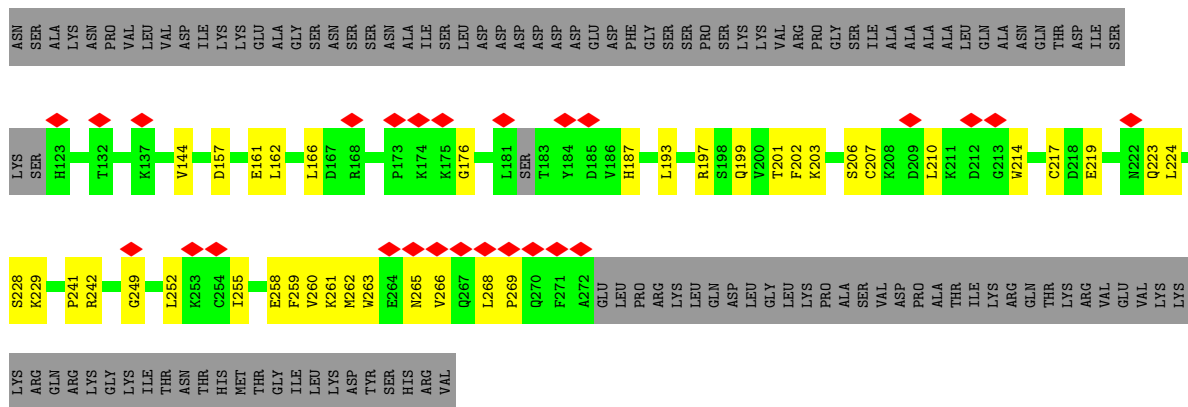


- Molecule 23: RNA polymerase II transcription factor B subunit 2

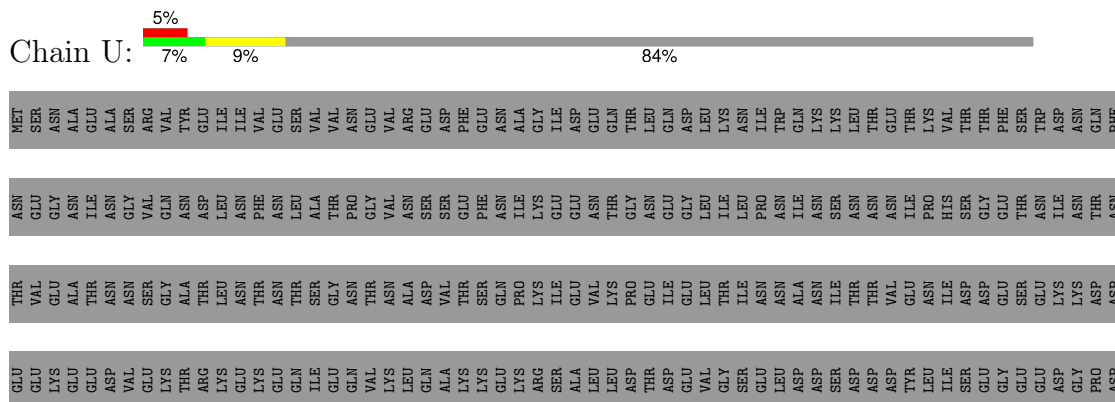


- Molecule 24: Transcription initiation factor IIE subunit beta

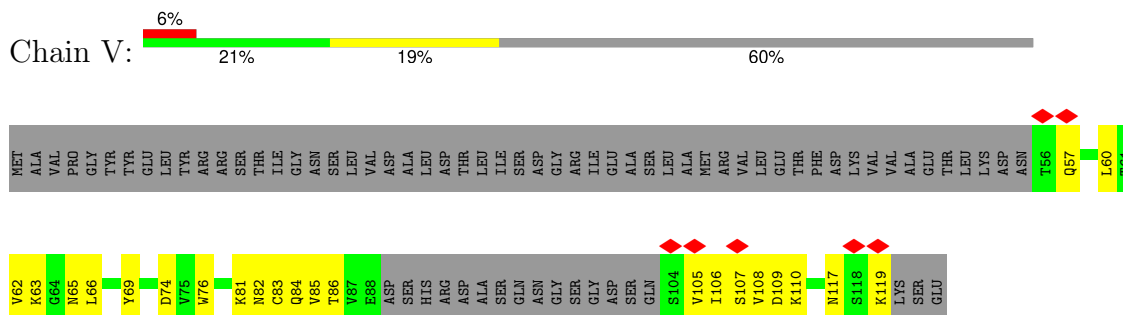




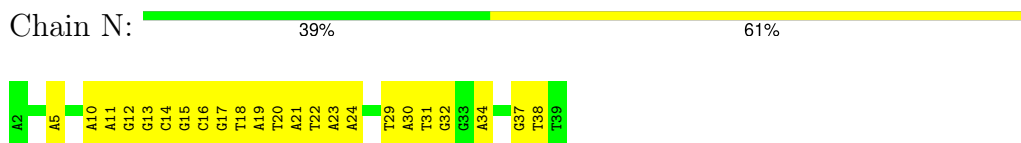
- Molecule 25: Transcription initiation factor IIA large subunit



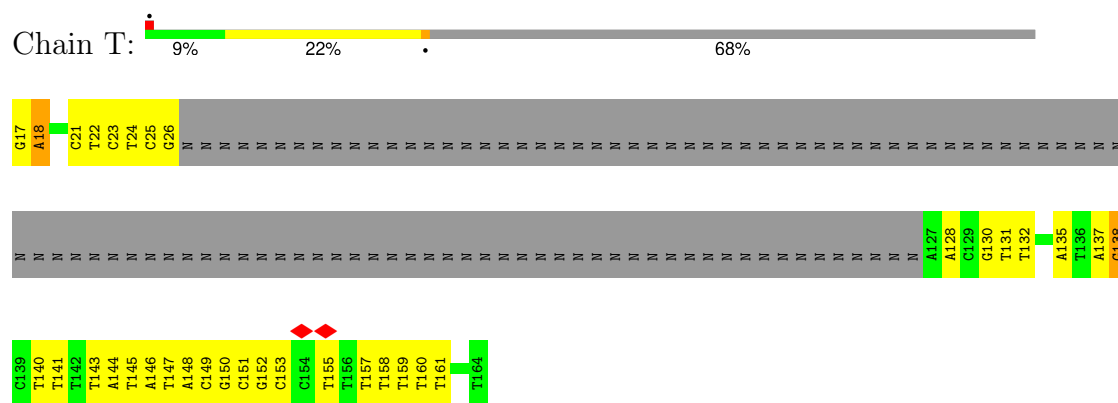
- Molecule 26: Transcription initiation factor IIA subunit 2



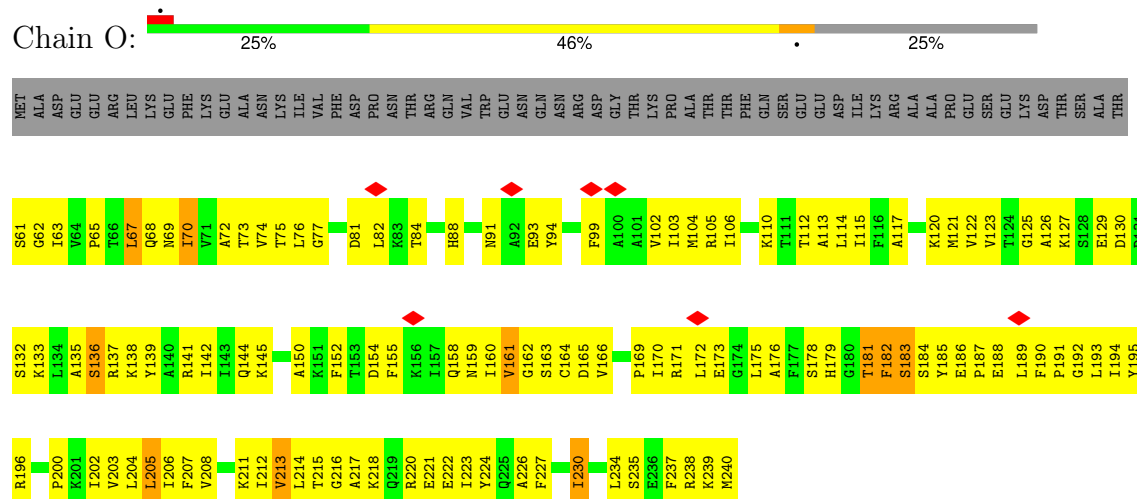
- Molecule 27: non-template strand DNA



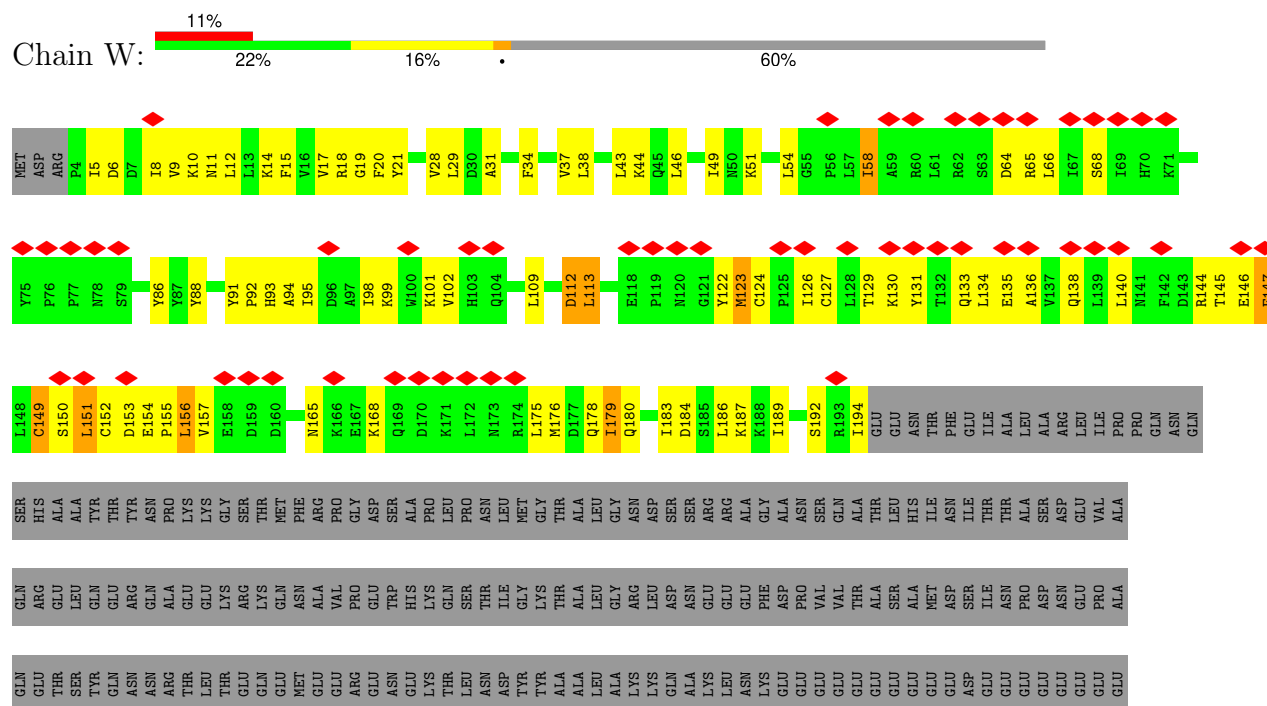
- Molecule 28: template strand DNA



• Molecule 29: TATA-box-binding protein



• Molecule 30: Transcription initiation factor IIE subunit alpha



MET
GLY
ASP
VAL
SER
ALA
MET
ASP
ASP
ASN
ASN
GLU
THR
ALA
ARG
GLU
ASN
ALA
LEU
GLU
ASP
GLU
PHE
GLU
MET
ASP
VAL
THR
ASP
GLY
THR
ALA
LYS
THR
GLU
SER
ASN
THR
SER
ASN
ASP
VAL
LYS
GLN
GLU
SER
ILE
ASN
ASP
LYS
THR
GLU
ASP
VAL
ASN
ALA
THR
ALA
THR
ALA

SER
GLY
PRO
SER
ALA
ASN
ALA
LYS
PRO
ASN
ASP
GLY
ASP
ASP
ASP
ASP
ASP
PHE
GLU
MET
ASP
ILE
GLU
PHE
GLU
ASP
VAL

- Molecule 31: RNA

Chain P:

60%

40%

G6
A7
G8
G9
A10

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	254448	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)	500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.057	Depositor
Minimum map value	-0.001	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	444.13998, 503.49997, 508.8	wwPDB
Map dimensions	419, 475, 480	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, MG, 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	Q	0.42	0/1165	0.57	0/1576
2	R	0.30	0/1047	0.51	0/1422
3	D	0.28	0/1262	0.48	0/1693
4	G	0.50	0/1368	0.56	0/1844
5	M	0.37	0/1828	0.57	5/2459 (0.2%)
6	A	1.02	2/11237 (0.0%)	0.82	11/15195 (0.1%)
7	B	1.10	1/9033 (0.0%)	0.80	4/12181 (0.0%)
8	C	0.98	0/2133	0.74	0/2891
9	E	0.91	0/1788	0.69	0/2406
10	F	1.14	0/700	0.91	0/945
11	H	0.90	0/1086	0.84	3/1470 (0.2%)
12	I	0.72	0/989	0.67	0/1331
13	J	1.12	0/541	0.88	0/727
14	K	0.97	0/937	0.66	0/1265
15	L	0.78	0/365	0.84	0/485
16	3	0.13	0/360	0.30	0/501
17	0	0.14	0/6226	0.37	1/8407 (0.0%)
18	4	0.16	0/2062	0.41	2/2805 (0.1%)
19	6	0.14	0/2506	0.36	0/3402
20	1	0.13	0/1896	0.32	0/2543
21	7	0.13	0/4521	0.34	0/6036
22	5	0.12	0/502	0.30	0/677
23	2	0.12	0/3057	0.34	0/4071
24	X	0.15	0/929	0.41	0/1272
25	U	0.15	0/389	0.40	0/523
26	V	0.16	0/384	0.30	0/518
27	N	0.34	0/893	1.02	0/1377
28	T	0.37	0/1076	1.15	2/1654 (0.1%)
29	O	0.19	0/1443	0.38	0/1942
30	W	0.18	0/1490	0.41	0/2014
31	P	0.54	0/124	1.37	1/193 (0.5%)
All	All	0.71	3/63337 (0.0%)	0.65	29/85825 (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	Q	0	1
4	G	0	1
5	M	0	3
6	A	0	9
7	B	0	5
10	F	0	1
11	H	0	3
All	All	0	23

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	687	GLU	CA-C	-8.99	1.40	1.52
6	A	786	HIS	CA-C	-8.31	1.37	1.52
6	A	512	VAL	C-N	-5.70	1.21	1.33

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
31	P	9	G	N9-C1'-C2'	6.36	121.54	112.00
6	A	1123	GLY	CA-C-N	6.28	133.01	121.70
6	A	1123	GLY	C-N-CA	6.28	133.01	121.70
5	M	268	GLU	CA-C-N	6.28	133.27	121.97
5	M	268	GLU	C-N-CA	6.28	133.27	121.97

There are no chirality outliers.

5 of 23 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	G	56	ILE	Peptide
5	M	267	LYS	Peptide
5	M	268	GLU	Peptide
5	M	270	ALA	Peptide
1	Q	125	LYS	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	Q	1141	0	1027	133	0
2	R	1039	0	857	128	0
3	D	1253	0	1275	103	0
4	G	1340	0	1357	119	0
5	M	1805	0	1895	122	0
6	A	11039	0	11122	785	0
7	B	8861	0	8884	525	0
8	C	2095	0	2051	118	0
9	E	1752	0	1776	122	0
10	F	688	0	707	78	0
11	H	1068	0	1040	129	0
12	I	971	0	927	101	0
13	J	532	0	542	32	0
14	K	919	0	929	49	0
15	L	363	0	389	53	0
16	3	361	0	150	4	0
17	0	6108	0	6167	278	0
18	4	2041	0	1954	111	0
19	6	2527	0	2321	99	0
20	1	2411	0	1881	58	0
21	7	4447	0	3905	200	0
22	5	498	0	506	12	0
23	2	3011	0	2600	84	0
24	X	921	0	650	34	0
25	U	383	0	384	25	0
26	V	381	0	388	22	0
27	N	791	0	426	40	0
28	T	968	0	552	55	0
29	O	1416	0	1493	153	0
30	W	1469	0	1433	91	0
31	P	110	0	56	13	0
32	4	1	0	0	0	0
32	6	4	0	0	0	0
32	A	2	0	0	0	0
32	B	1	0	0	0	0
32	C	1	0	0	0	0
32	I	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
32	J	1	0	0	0	0
32	L	1	0	0	0	0
32	M	1	0	0	0	0
32	W	1	0	0	0	0
33	A	1	0	0	0	0
34	0	8	0	0	2	0
All	All	62733	0	59644	3536	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 29.

The worst 5 of 3536 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)
6:A:68:GLN:NE2	6:A:80:HIS:ND1	1.92	1.16
11:H:55:LEU:H	11:H:146:ARG:HA	1.12	1.10
6:A:1444:MET:HA	10:F:133:VAL:HA	1.19	1.09
6:A:68:GLN:NE2	6:A:80:HIS:CE1	2.23	1.07
6:A:68:GLN:HE22	6:A:80:HIS:CG	1.74	1.05

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	Q	140/735 (19%)	119 (85%)	21 (15%)	0	100	100
2	R	142/398 (36%)	126 (89%)	16 (11%)	0	100	100
3	D	153/221 (69%)	140 (92%)	13 (8%)	0	100	100
4	G	169/171 (99%)	156 (92%)	13 (8%)	0	100	100
5	M	228/345 (66%)	196 (86%)	29 (13%)	3 (1%)	9	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	A	1395/1733 (80%)	1227 (88%)	161 (12%)	7 (0%)	24	57
7	B	1096/1224 (90%)	942 (86%)	149 (14%)	5 (0%)	24	57
8	C	264/318 (83%)	241 (91%)	22 (8%)	1 (0%)	30	61
9	E	212/215 (99%)	199 (94%)	13 (6%)	0	100	100
10	F	83/155 (54%)	73 (88%)	10 (12%)	0	100	100
11	H	129/146 (88%)	97 (75%)	31 (24%)	1 (1%)	16	47
12	I	117/122 (96%)	95 (81%)	22 (19%)	0	100	100
13	J	63/70 (90%)	52 (82%)	11 (18%)	0	100	100
14	K	112/120 (93%)	101 (90%)	11 (10%)	0	100	100
15	L	44/70 (63%)	30 (68%)	14 (32%)	0	100	100
16	3	70/321 (22%)	65 (93%)	5 (7%)	0	100	100
17	0	752/778 (97%)	701 (93%)	51 (7%)	0	100	100
18	4	279/338 (82%)	246 (88%)	33 (12%)	0	100	100
19	6	336/461 (73%)	295 (88%)	39 (12%)	2 (1%)	21	52
20	1	256/543 (47%)	237 (93%)	17 (7%)	2 (1%)	16	47
21	7	630/843 (75%)	578 (92%)	52 (8%)	0	100	100
22	5	64/72 (89%)	59 (92%)	5 (8%)	0	100	100
23	2	456/513 (89%)	412 (90%)	44 (10%)	0	100	100
24	X	145/328 (44%)	126 (87%)	19 (13%)	0	100	100
25	U	44/286 (15%)	39 (89%)	5 (11%)	0	100	100
26	V	45/122 (37%)	45 (100%)	0	0	100	100
29	O	178/240 (74%)	165 (93%)	13 (7%)	0	100	100
30	W	189/482 (39%)	179 (95%)	10 (5%)	0	100	100
All	All	7791/11370 (68%)	6941 (89%)	829 (11%)	21 (0%)	37	67

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	M	269	ILE
5	M	270	ALA
6	A	68	GLN
6	A	72	GLU
6	A	609	ASP

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	Q	108/641 (17%)	86 (80%)	22 (20%)	1	5
2	R	78/362 (22%)	60 (77%)	18 (23%)	1	4
3	D	139/200 (70%)	125 (90%)	14 (10%)	7	28
4	G	152/152 (100%)	131 (86%)	21 (14%)	3	15
5	M	202/299 (68%)	166 (82%)	36 (18%)	2	9
6	A	1224/1520 (80%)	1004 (82%)	220 (18%)	2	8
7	B	967/1061 (91%)	817 (84%)	150 (16%)	2	12
8	C	234/274 (85%)	191 (82%)	43 (18%)	1	8
9	E	196/197 (100%)	166 (85%)	30 (15%)	3	12
10	F	75/137 (55%)	63 (84%)	12 (16%)	2	12
11	H	117/128 (91%)	95 (81%)	22 (19%)	1	7
12	I	113/116 (97%)	82 (73%)	31 (27%)	0	1
13	J	60/65 (92%)	49 (82%)	11 (18%)	2	8
14	K	99/102 (97%)	88 (89%)	11 (11%)	6	24
15	L	40/57 (70%)	27 (68%)	13 (32%)	0	0
16	3	1/303 (0%)	1 (100%)	0	100	100
17	0	686/707 (97%)	656 (96%)	30 (4%)	25	56
18	4	198/298 (66%)	180 (91%)	18 (9%)	9	32
19	6	247/406 (61%)	234 (95%)	13 (5%)	20	51
20	1	169/396 (43%)	163 (96%)	6 (4%)	31	62
21	7	414/737 (56%)	393 (95%)	21 (5%)	21	52
22	5	53/66 (80%)	53 (100%)	0	100	100
23	2	258/468 (55%)	247 (96%)	11 (4%)	26	57
24	X	54/295 (18%)	54 (100%)	0	100	100
25	U	42/260 (16%)	40 (95%)	2 (5%)	23	54
26	V	46/108 (43%)	45 (98%)	1 (2%)	45	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
29	O	152/205 (74%)	139 (91%)	13 (9%)	10	34
30	W	155/429 (36%)	142 (92%)	13 (8%)	10	35
All	All	6279/9989 (63%)	5497 (88%)	782 (12%)	7	19

5 of 782 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	B	1051	THR
11	H	44	VAL
7	B	1183	LYS
7	B	1048	THR
8	C	240	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
20	1	621	ASN
24	X	265	ASN
21	7	431	GLN
23	2	99	ASN
29	O	68	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
31	P	4/5 (80%)	1 (25%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
31	P	8	G

There are no RNA pucker outliers to report.

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 [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 17 ligands modelled in this entry, 16 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	$\# Z > 2$
34	SF4	0	801	17	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
34	SF4	0	801	17	-	-	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 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
34	0	801	SF4	2	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
20	1	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	1	355:UNK	C	368:UNK	N	14.81
1	1	519:UNK	C	537:GLU	N	11.86

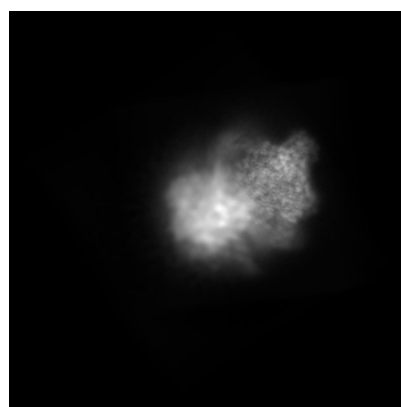
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-23908. These allow visual inspection of the internal detail of the map and identification of artifacts.

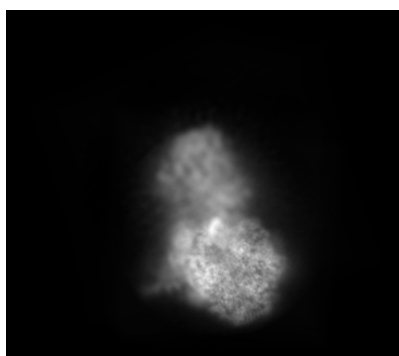
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

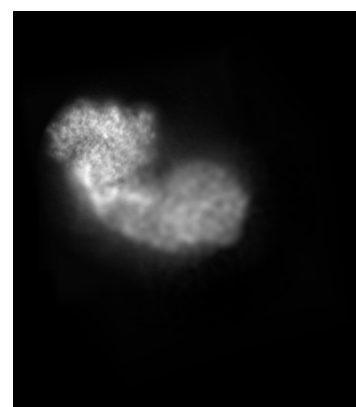
6.1.1 Primary map



X



Y

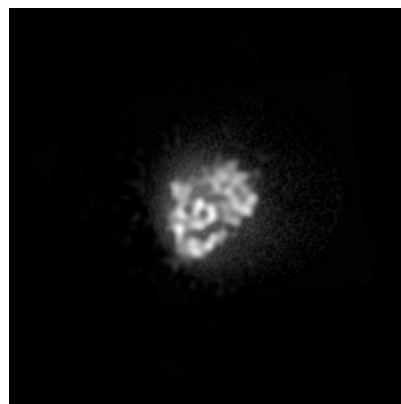


Z

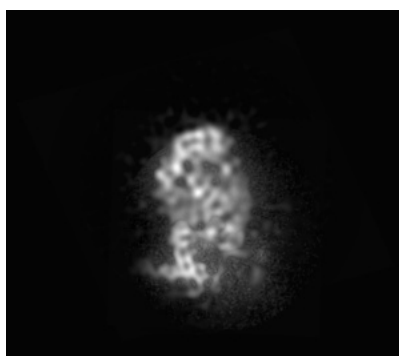
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

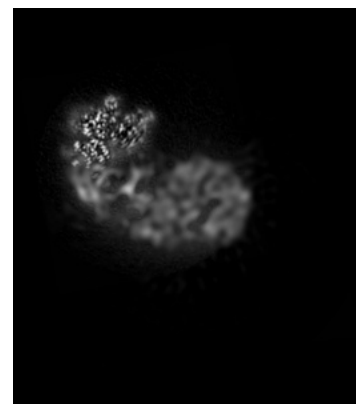
6.2.1 Primary map



X Index: 209



Y Index: 237

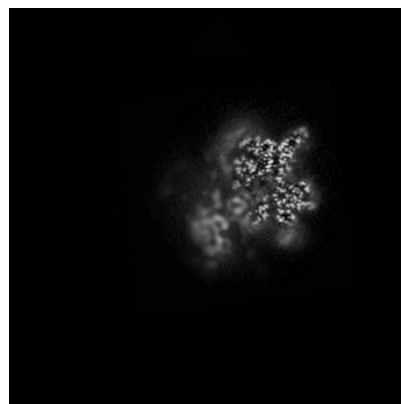


Z Index: 240

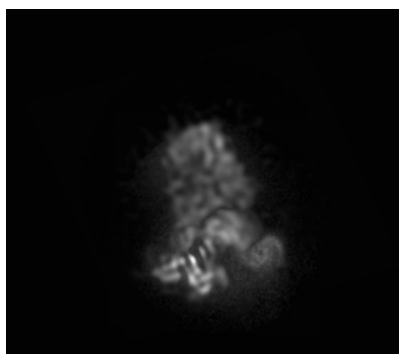
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

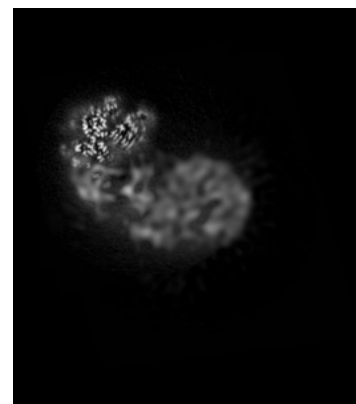
6.3.1 Primary map



X Index: 109



Y Index: 253

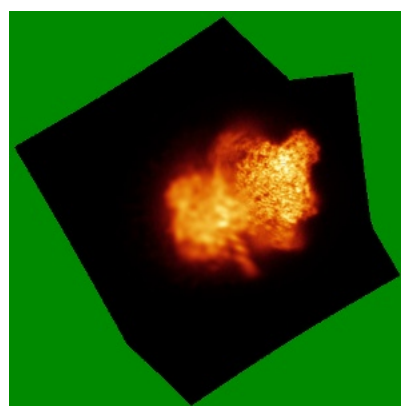


Z Index: 242

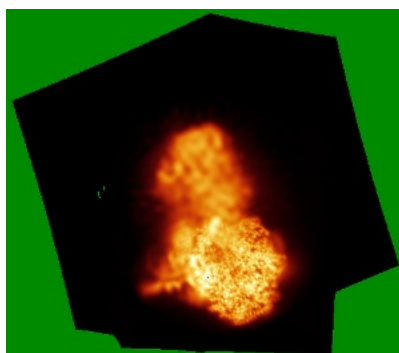
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [i](#)

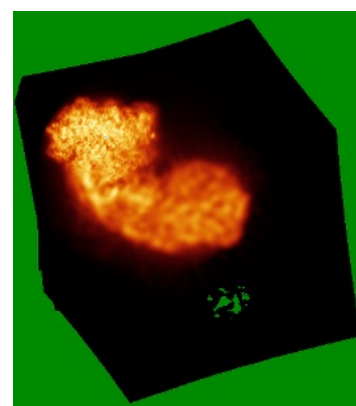
6.4.1 Primary map



X



Y

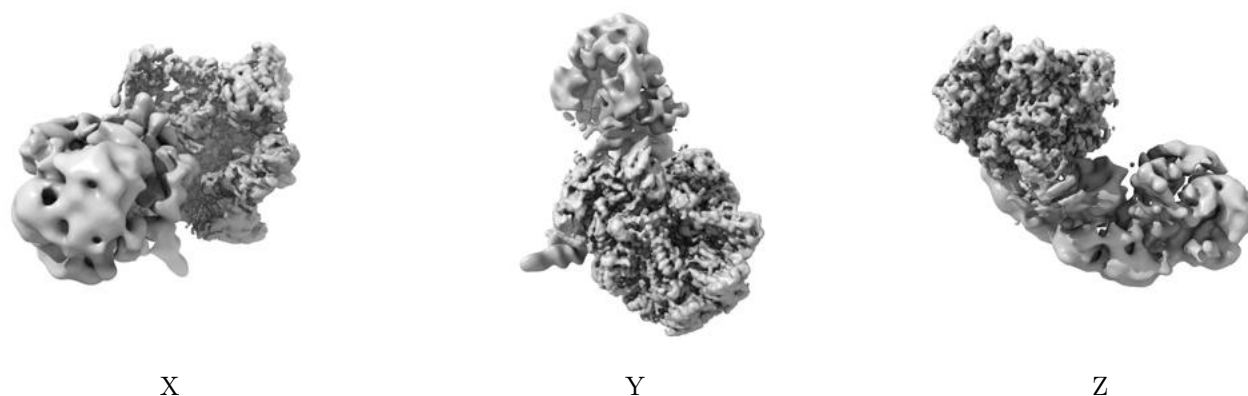


Z

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 [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

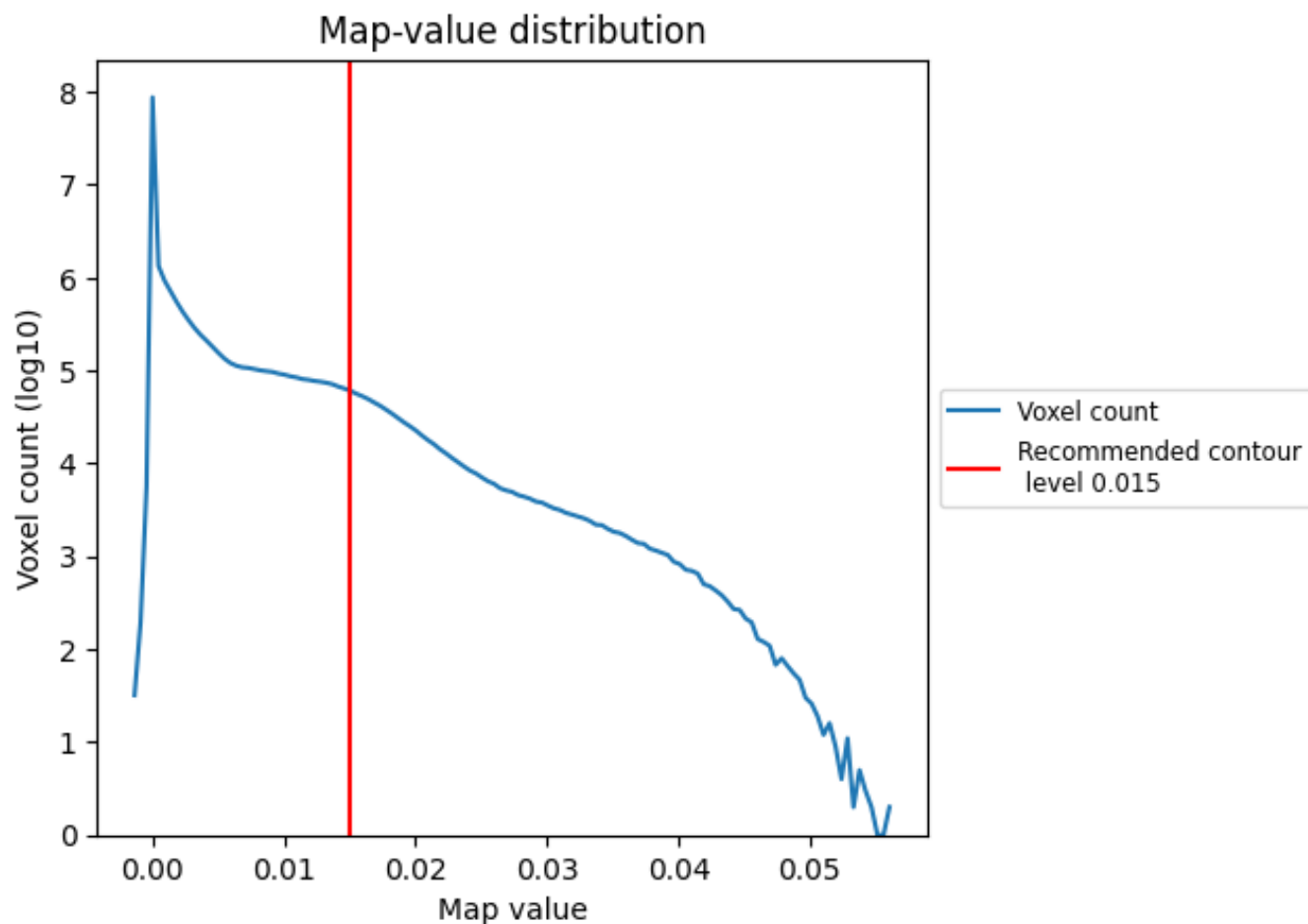
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

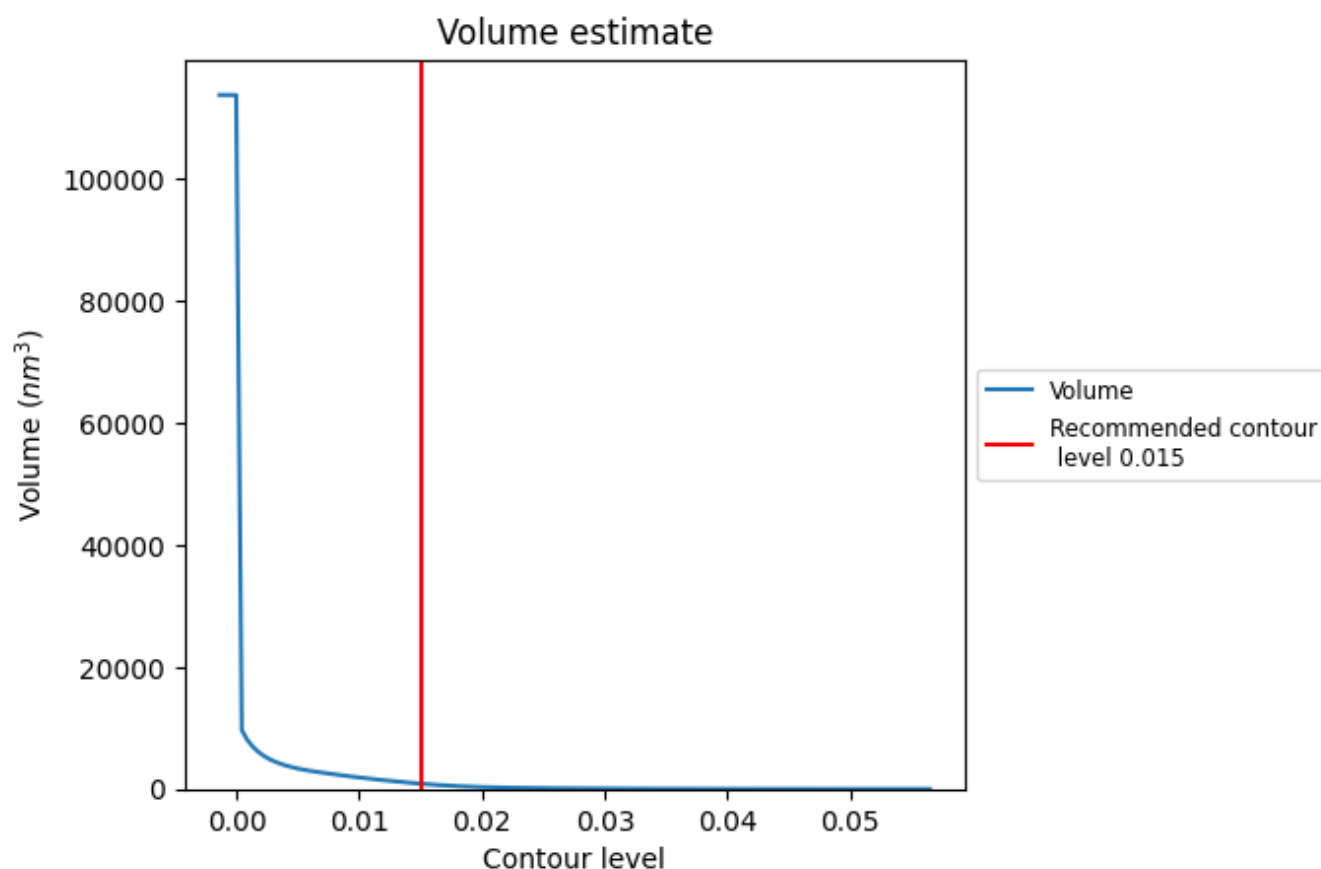
This section contains the results of statistical analysis of the map.

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 863 nm³; this corresponds to an approximate mass of 779 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.

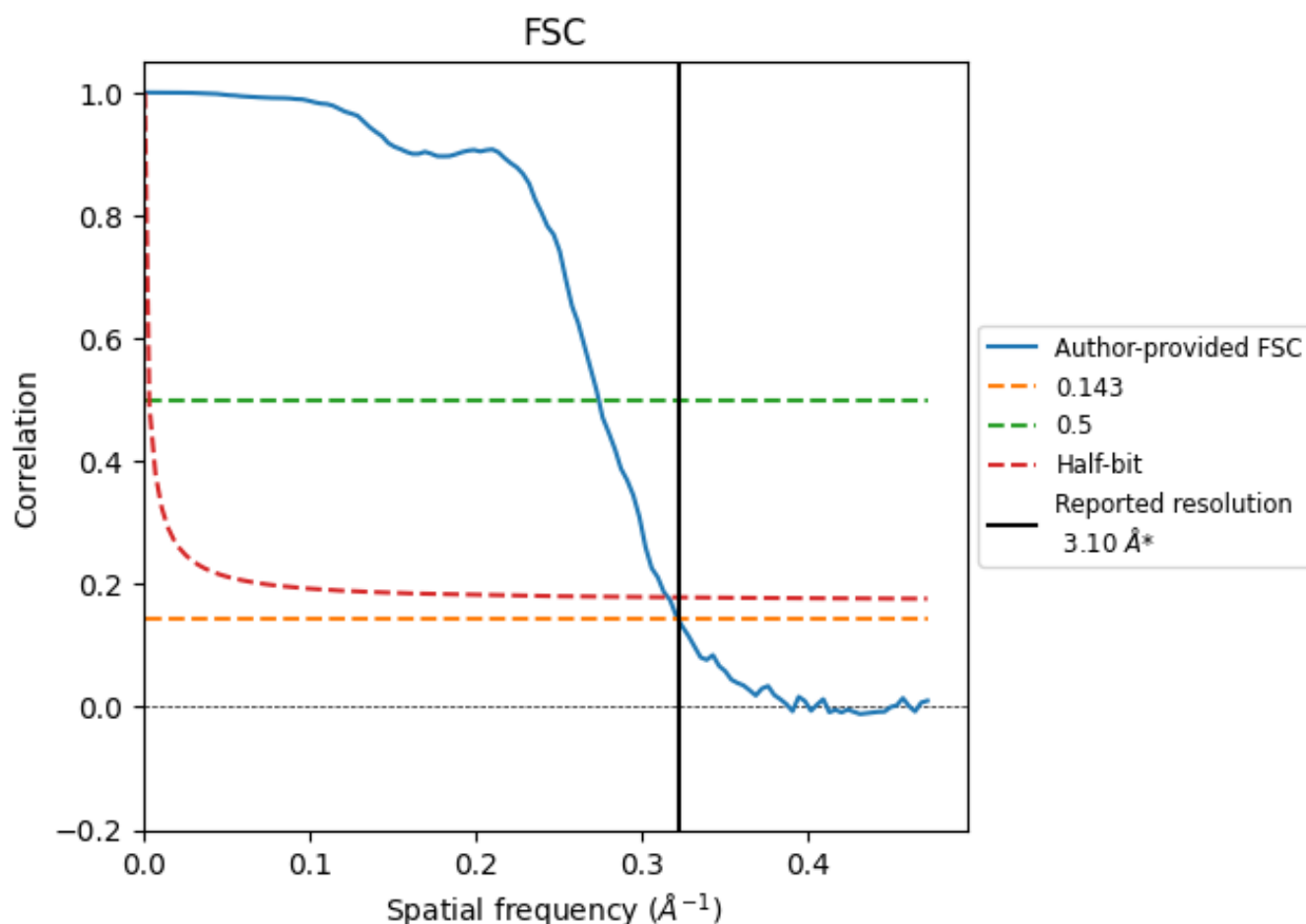
7.3 Rotationally averaged power spectrum [i](#)

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 0.323 Å⁻¹

8.2 Resolution estimates [i](#)

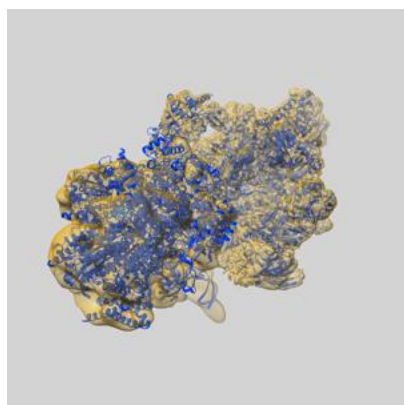
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.10	-	-
Author-provided FSC curve	3.11	3.65	3.17
Unmasked-calculated*	-	-	-

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

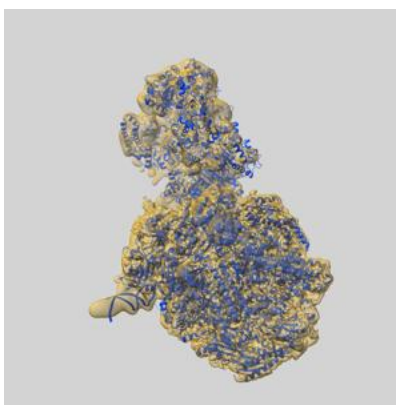
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-23908 and PDB model 7ML4. Per-residue inclusion information can be found in section 3 on page 12.

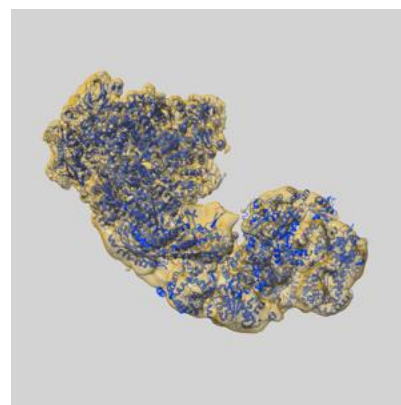
9.1 Map-model overlay [i](#)



X



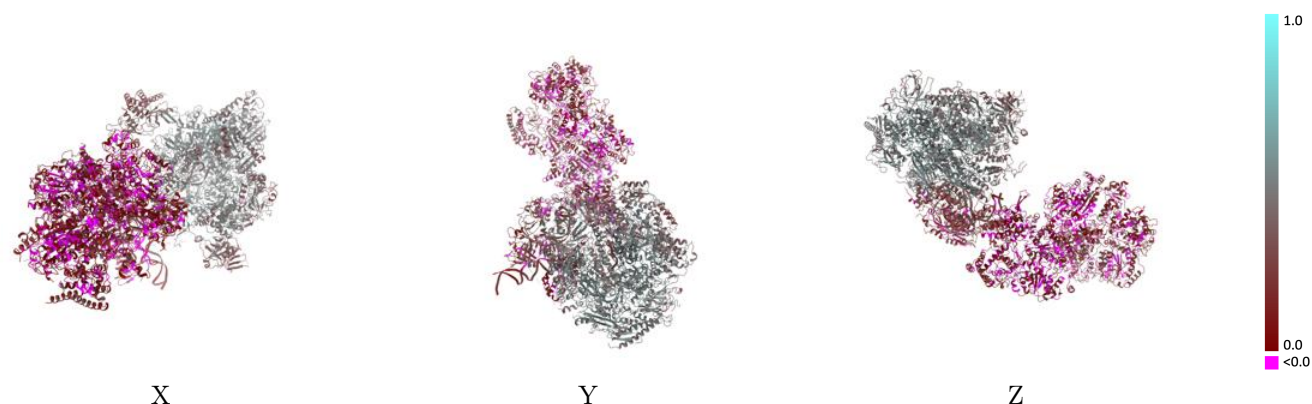
Y



Z

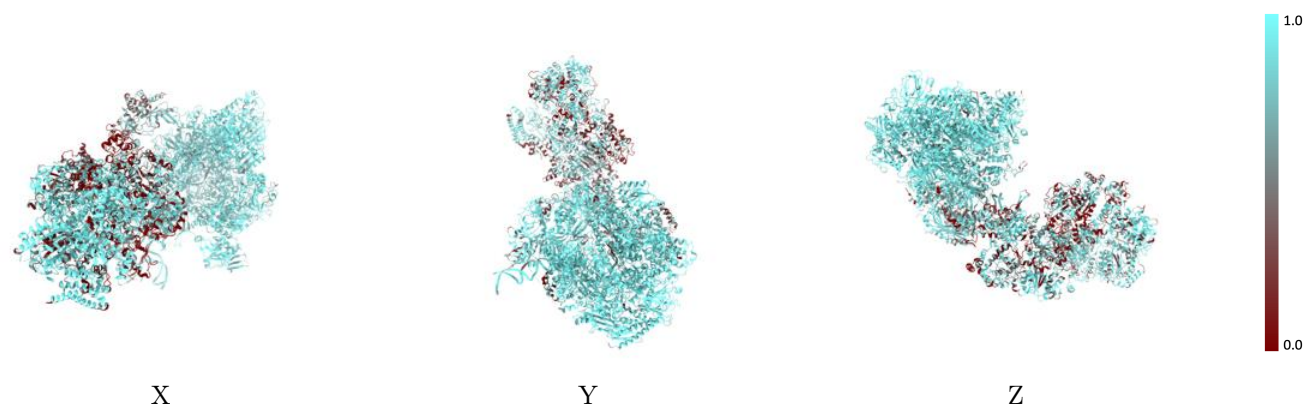
The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



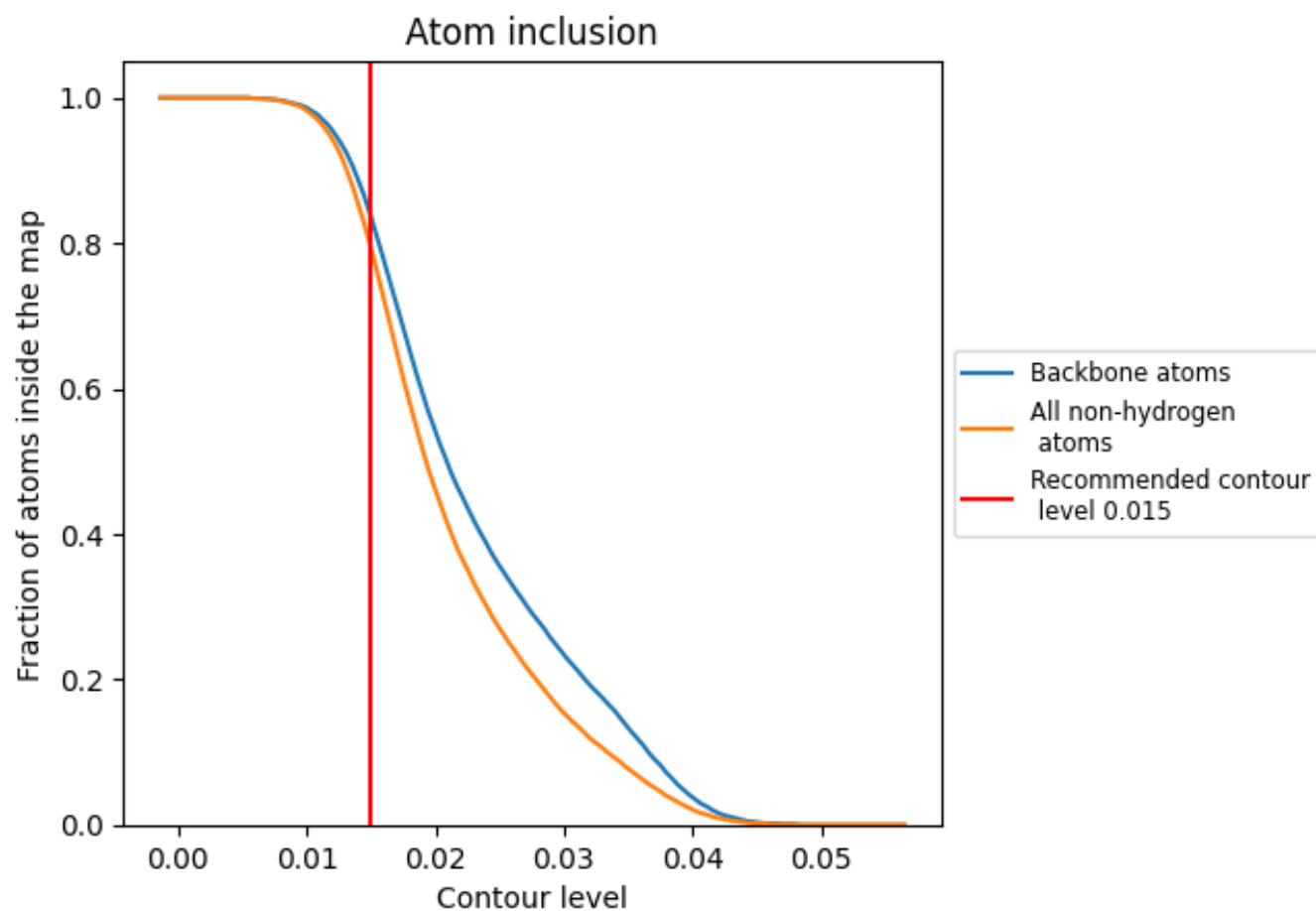
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.015).

9.4 Atom inclusion [i](#)



At the recommended contour level, 84% of all backbone atoms, 79% 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 (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7930	 0.2810
0	 0.6660	 0.0610
1	 0.5790	 0.0820
2	 0.6950	 0.0880
3	 0.4380	 0.1210
4	 0.7980	 0.0710
5	 0.4680	 0.0950
6	 0.7830	 0.0790
7	 0.4380	 0.0470
A	 0.9260	 0.4730
B	 0.9410	 0.4960
C	 0.9570	 0.4930
D	 0.4630	 0.2750
E	 0.9250	 0.4610
F	 0.9540	 0.5000
G	 0.6800	 0.3600
H	 0.9410	 0.4440
I	 0.8650	 0.4200
J	 0.9790	 0.5080
K	 0.8770	 0.4750
L	 0.9600	 0.4430
M	 0.7540	 0.1890
N	 0.9170	 0.1710
O	 0.8960	 0.0690
P	 0.6540	 0.2610
Q	 0.8950	 0.3570
R	 0.8490	 0.2590
T	 0.8460	 0.1830
U	 0.6490	 0.0900
V	 0.7880	 0.1090
W	 0.6690	 0.1000
X	 0.7940	 0.0640

