wwPDB EM Validation Summary Report (i)

Jul 15, 2024 – 08:50 am BST

PDB ID	:	7ZX8
EMDB ID	:	EMD-15007
Title	:	Structure of SNAPc containing Pol II pre-initiation complex bound to U1 snRNA promoter (OC)
Authors	:	Rengachari, S.; Schilbach, S.; Kaliyappan, T.; Gouge, J.; Zumer, K.; Schwarz, J.; Urlaub, H.; Dienemann, C.; Vannini, A.; Cramer, P.
Deposited on	:	2022-05-20
Resolution	:	3.00 Å(reported)

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:

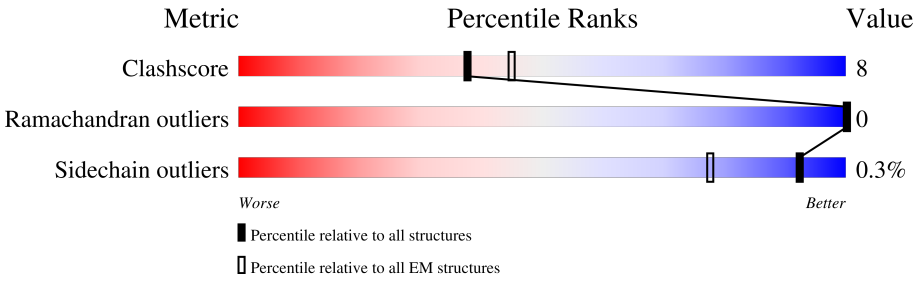
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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	1970	64% 8% 28%
2	B	1174	89% 8%
3	C	275	84% 9% 7%
4	D	142	23% 86% 10%
5	E	210	87% 13%
6	F	127	56% 6% 38%
7	G	172	41% 87% 12%
8	H	150	96%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	78% 14% 9%
10	J	67	87% 9%
11	K	117	85% 13%
12	L	58	67% 9% 24%
13	M	316	84% 10% 6%
14	N	96	36% 27% 34%
15	O	339	40% 13% 47%
16	Q	517	24% 73%
17	R	249	79% 10% 11%
18	T	96	9% 39% 25% 34%
19	U	376	14% 10% 77%
20	V	109	48% 39% 12%
21	a	368	18% 58% 42%
22	b	411	9% 88% 11%
23	c	1469	20% 79%
24	d	98	43% 51% 49%

2 Entry composition (i)

There are 26 unique types of molecules in this entry. The entry contains 49534 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	1423	11274	7092	2016	2094	72	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	1136	9076	5739	1597	1676	64	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	257	2059	1294	351	408	6	0	0

- Molecule 4 is a protein called RNA polymerase II subunit D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	128	1050	656	178	212	4	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	209	1721	1089	300	324	8	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	79	636	406	108	117	5	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
			1351	875	219	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	148	Total	C	N	O	S	0	0
			1186	750	194	237	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	114	Total	C	N	O	S	0	0
			928	571	166	180	11		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	64	Total	C	N	O	S	0	0
			507	328	86	87	6		

- Molecule 11 is a protein called RNA polymerase II subunit J.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	115	Total	C	N	O	S	0	0
			920	593	152	173	2		

- Molecule 12 is a protein called RNA polymerase II subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	44	Total	C	N	O	S	0	0
			373	231	72	64	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	298	Total	C	N	O	S	0	0
			2301	1435	409	439	18		

- Molecule 14 is a DNA chain called Non-template strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	63	Total	C	N	O	P	0	0
			1316	620	253	380	63		

- Molecule 15 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	179	Total	C	N	O	S	0	0
			1422	923	251	241	7		

- Molecule 16 is a protein called General transcription factor IIF subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Q	138	Total	C	N	O	S	0	0
			1138	719	208	208	3		

- Molecule 17 is a protein called General transcription factor IIF subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	R	222	Total	C	N	O	S	0	0
			1788	1127	320	338	3		

- Molecule 18 is a DNA chain called Template strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	63	Total	C	N	O	P	0	0
			1253	595	221	374	63		

- Molecule 19 is a protein called Transcription initiation factor IIA subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	U	88	Total	C	N	O	S	0	0
			734	470	124	136	4		

- Molecule 20 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	V	96	Total	C	N	O	S	0	0
			785	498	139	146	2		

- Molecule 21 is a protein called snRNA-activating protein complex subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	a	215	Total	C	N	O	S	0	0
			1807	1165	313	319	10		

- Molecule 22 is a protein called snRNA-activating protein complex subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	b	364	Total	C	N	O	S	0	0
			2972	1887	509	556	20		

- Molecule 23 is a protein called snRNA-activating protein complex subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	c	303	Total	C	N	O	S	0	0
			2517	1575	453	479	10		

- Molecule 24 is a protein called snRNA-activating protein complex subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	d	50	Total	C	N	O	S	0	0
			408	248	79	79	2		

- Molecule 25 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
25	A	2	Total	Zn	0
			2	2	
25	B	1	Total	Zn	0
			1	1	
25	C	1	Total	Zn	0
			1	1	
25	I	2	Total	Zn	0
			2	2	
25	J	1	Total	Zn	0
			1	1	
25	L	1	Total	Zn	0
			1	1	
25	M	1	Total	Zn	0
			1	1	
25	b	2	Total	Zn	0
			2	2	

- Molecule 26 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Lig-

and of Interest" by depositor).

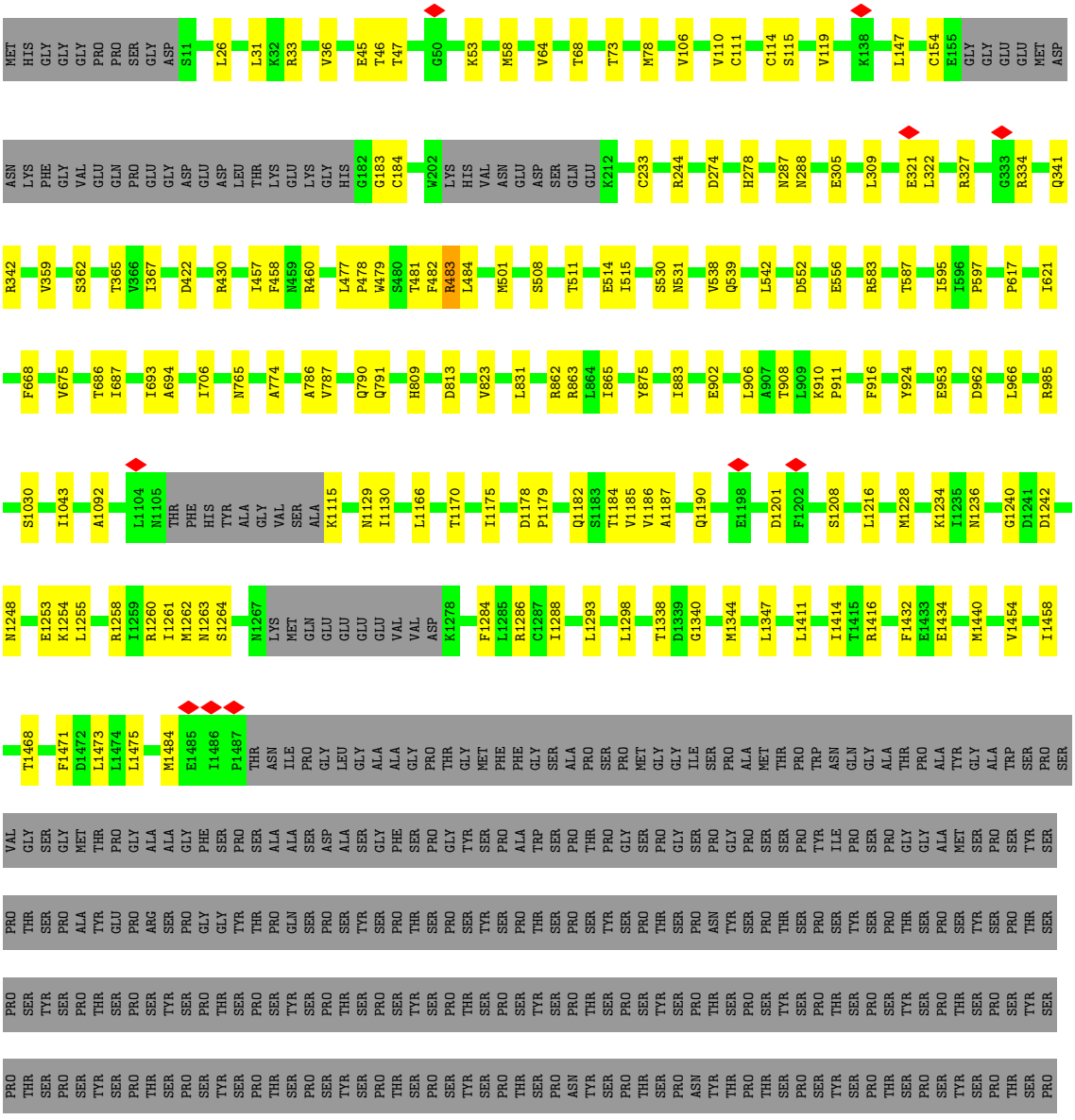
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
26	A	1	1	1	0

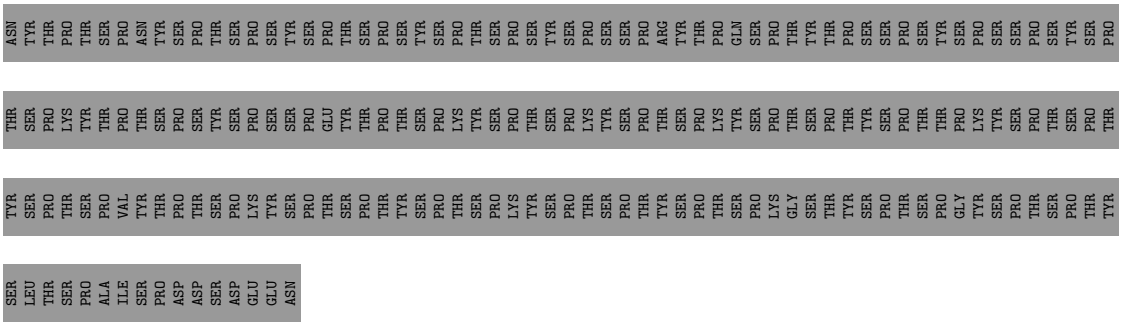
3 Residue-property plots (i)

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: DNA-directed RNA polymerase subunit

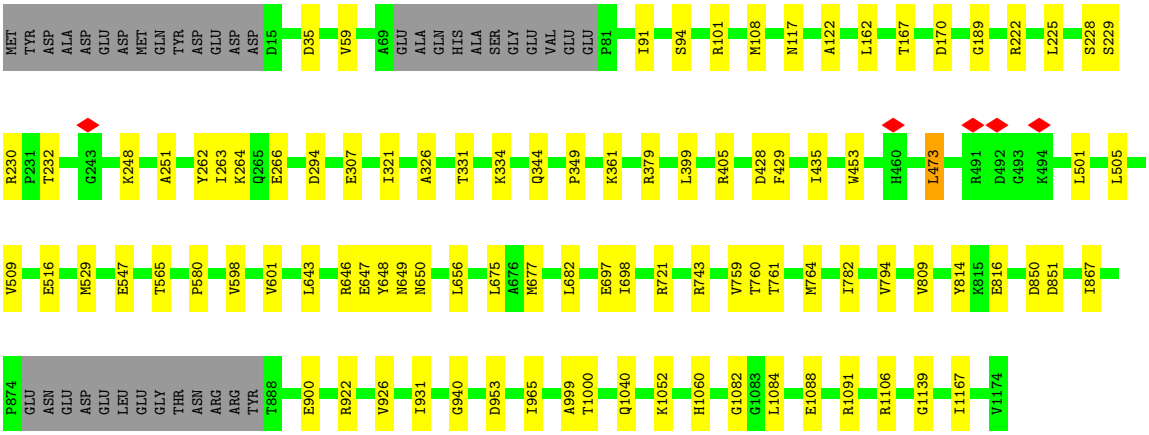
Chain A:





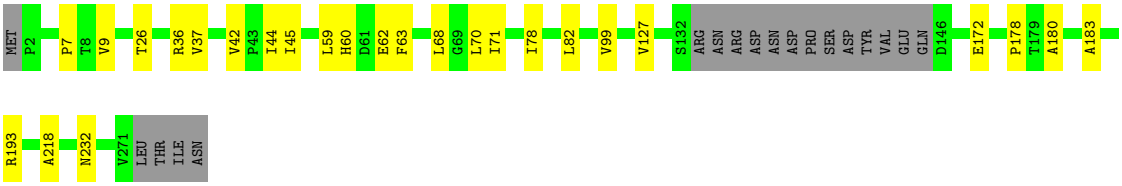
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain B:



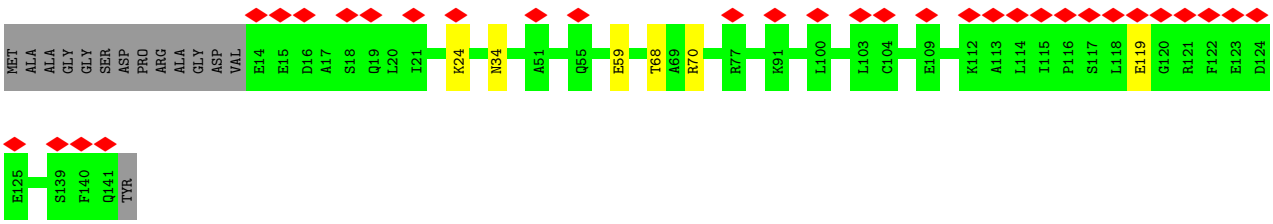
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C:



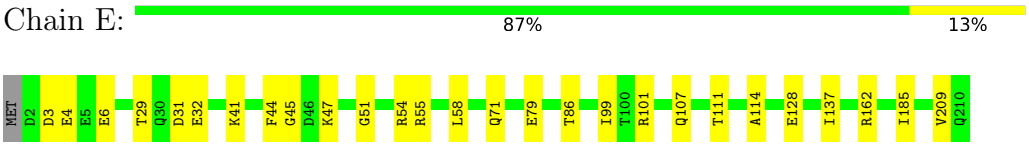
• Molecule 4: RNA polymerase II subunit D

Chain D:



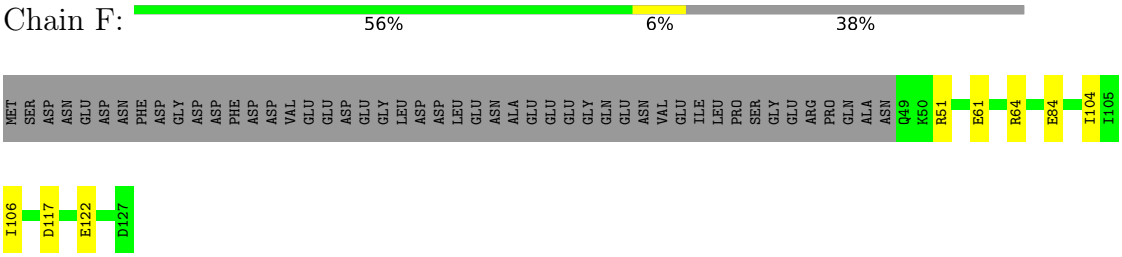
• Molecule 5: DNA-directed RNA polymerase II subunit E

Chain E:



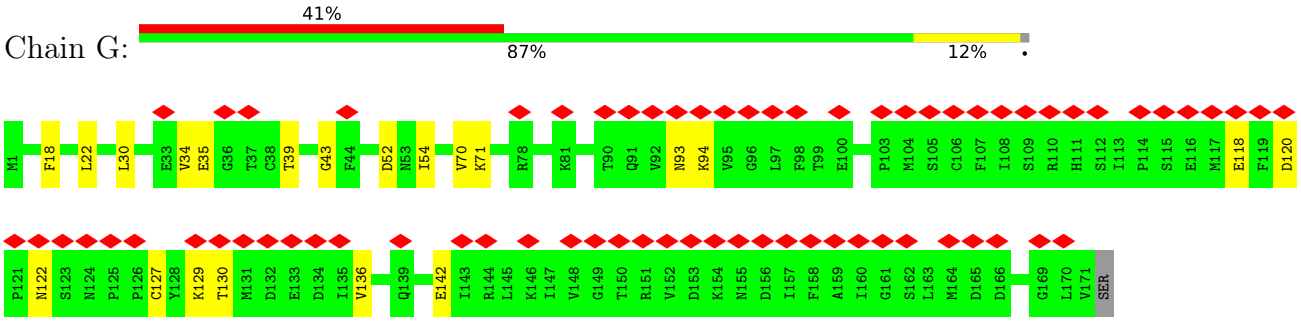
• Molecule 6: DNA-directed RNA polymerase II subunit F

Chain F:



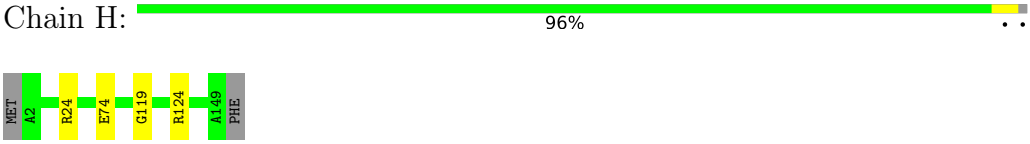
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G:



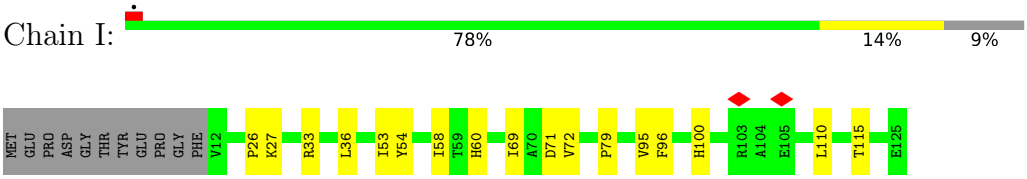
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



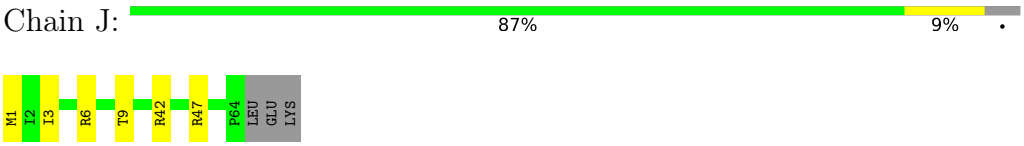
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:



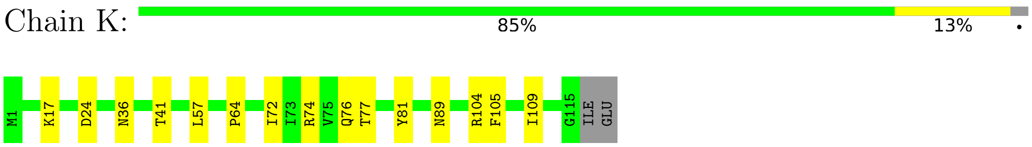
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



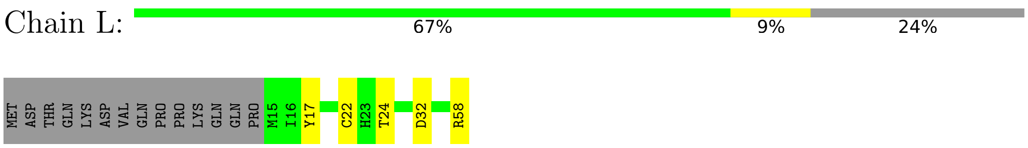
• Molecule 11: RNA polymerase II subunit J

Chain K:



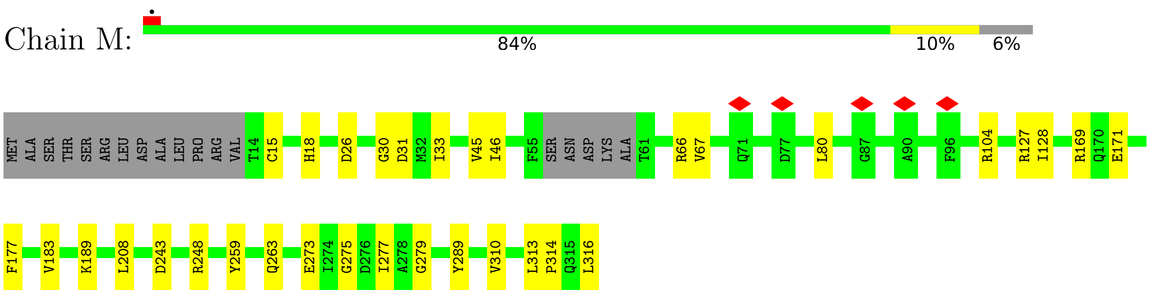
• Molecule 12: RNA polymerase II subunit K

Chain L:



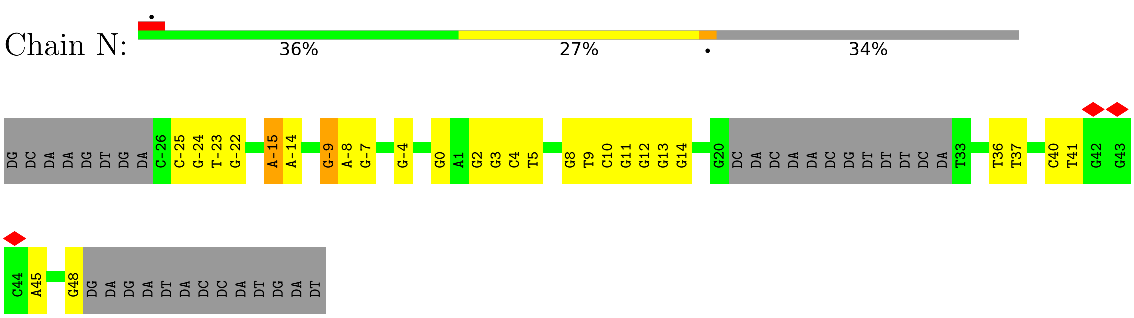
• Molecule 13: Transcription initiation factor IIB

Chain M:



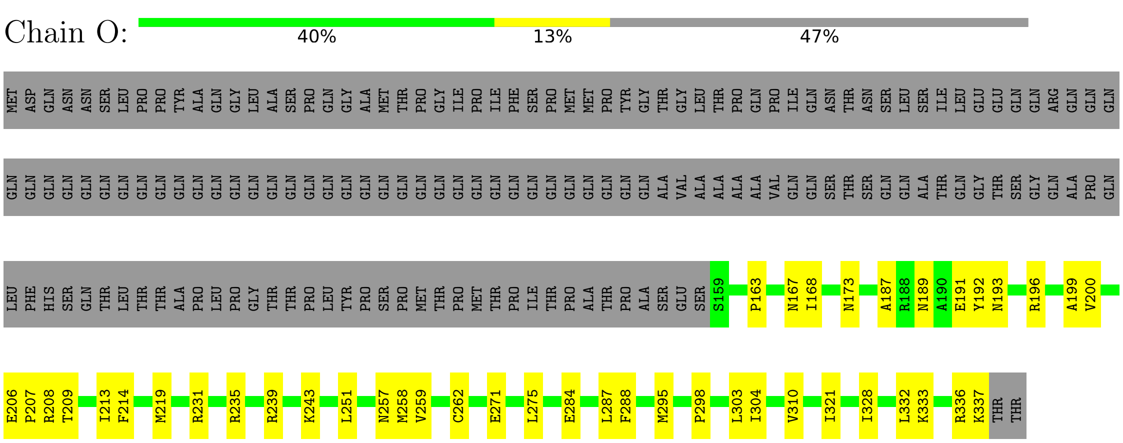
• Molecule 14: Non-template strand

Chain N:



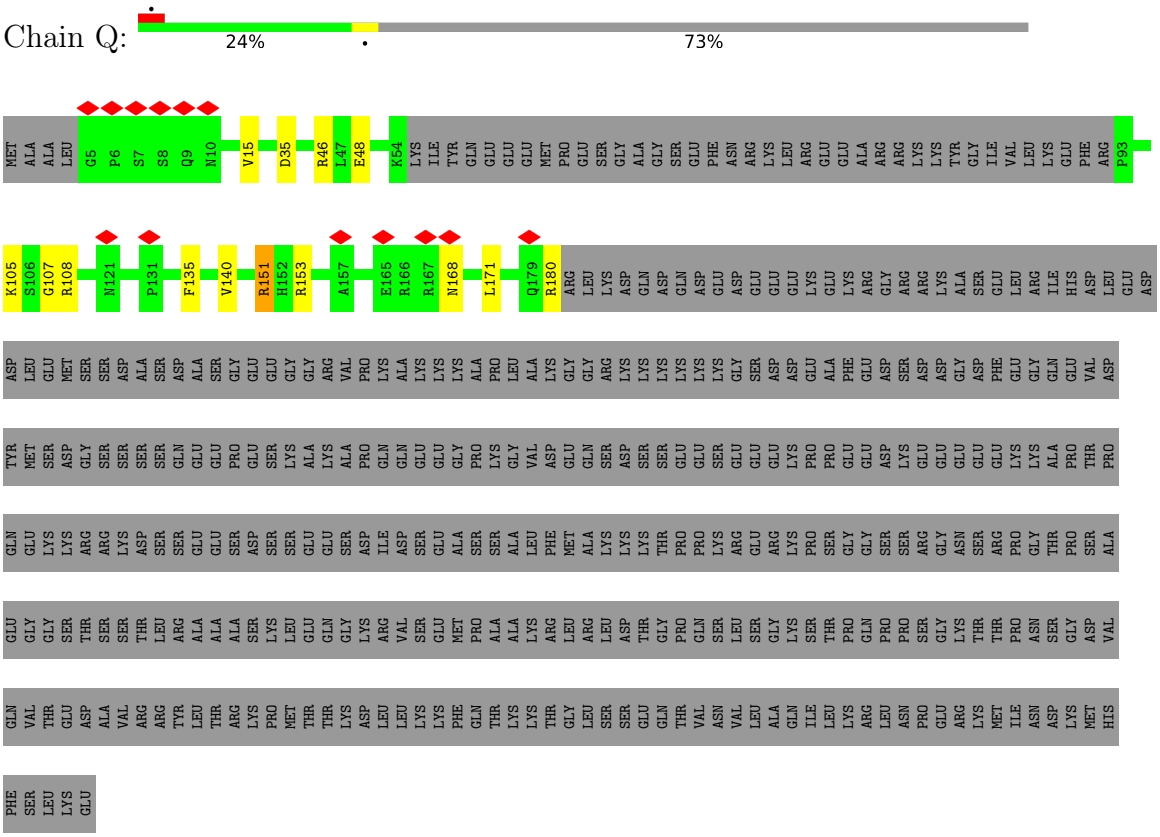
• Molecule 15: TATA-box-binding protein

Chain O:



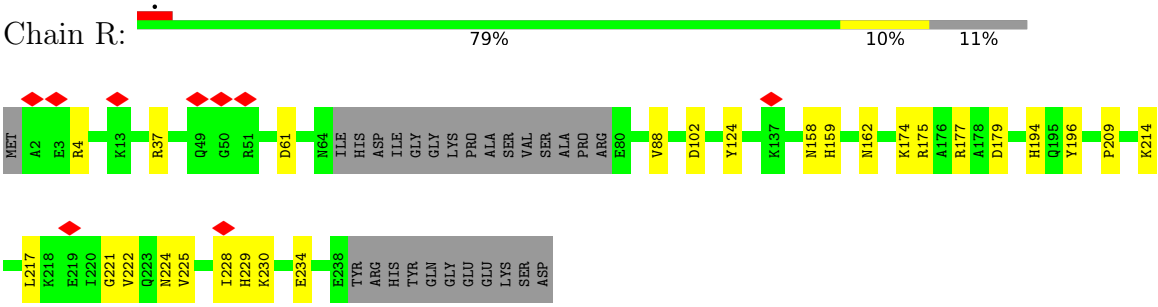
• Molecule 16: General transcription factor IIF subunit 1

Chain Q:



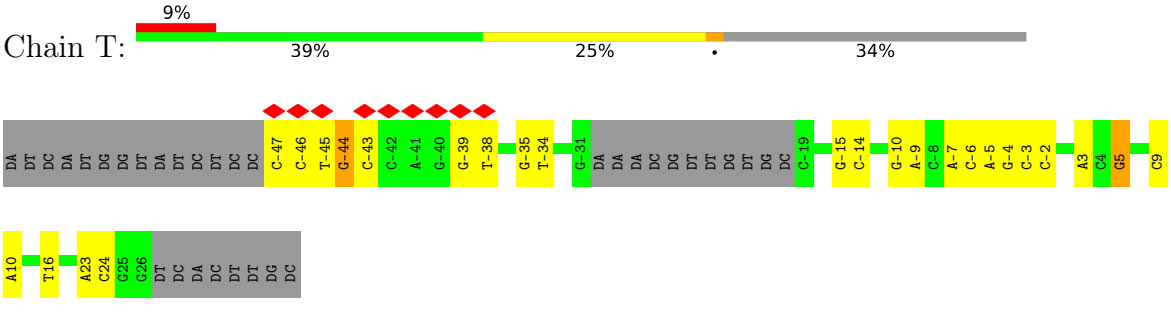
• Molecule 17: General transcription factor IIF subunit 2

Chain R:



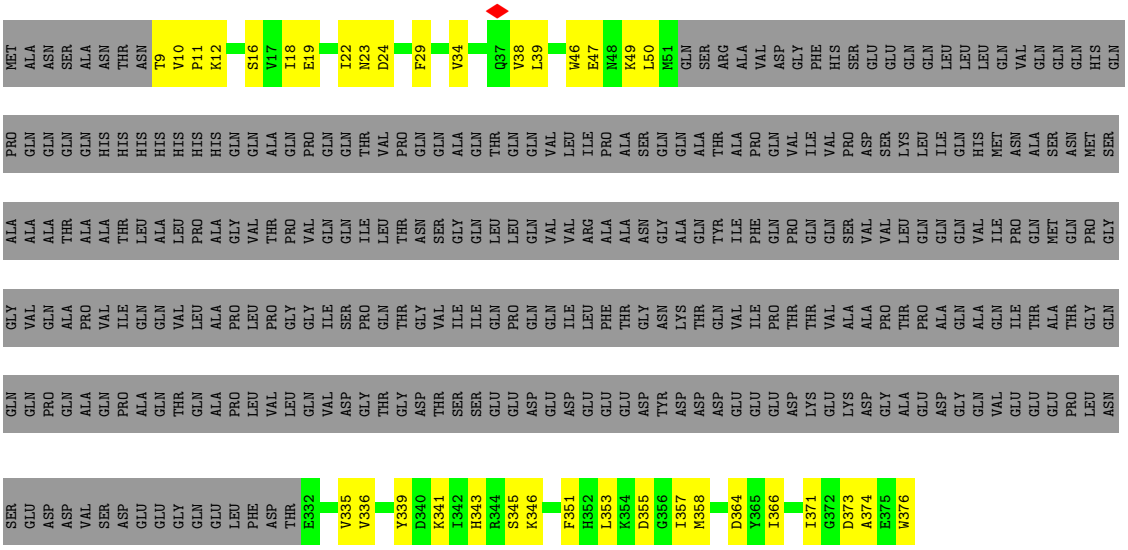
• Molecule 18: Template strand

Chain T:



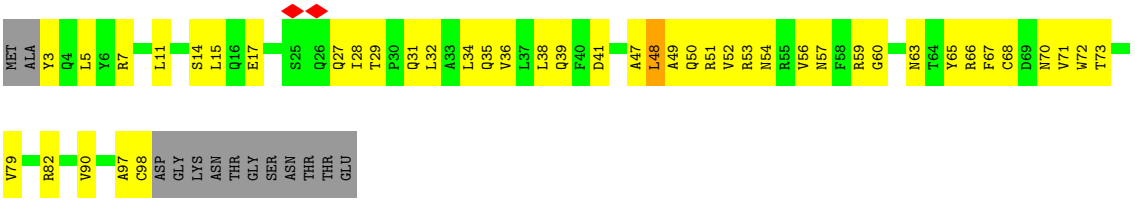
• Molecule 19: Transcription initiation factor IIA subunit 1

Chain U:



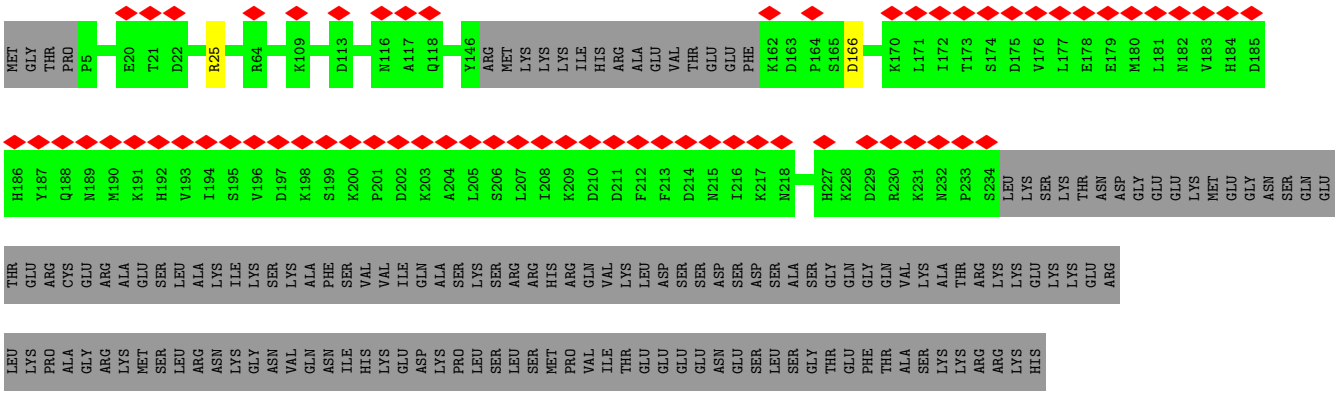
• Molecule 20: Transcription initiation factor IIA subunit 2

Chain V:



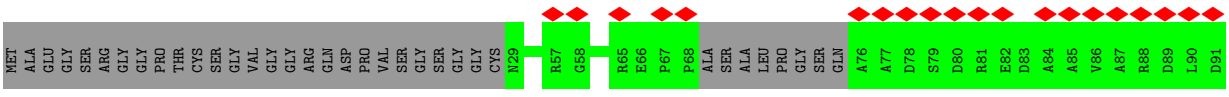
• Molecule 21: snRNA-activating protein complex subunit 1

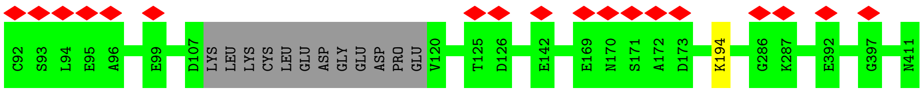
Chain a:



• Molecule 22: snRNA-activating protein complex subunit 3

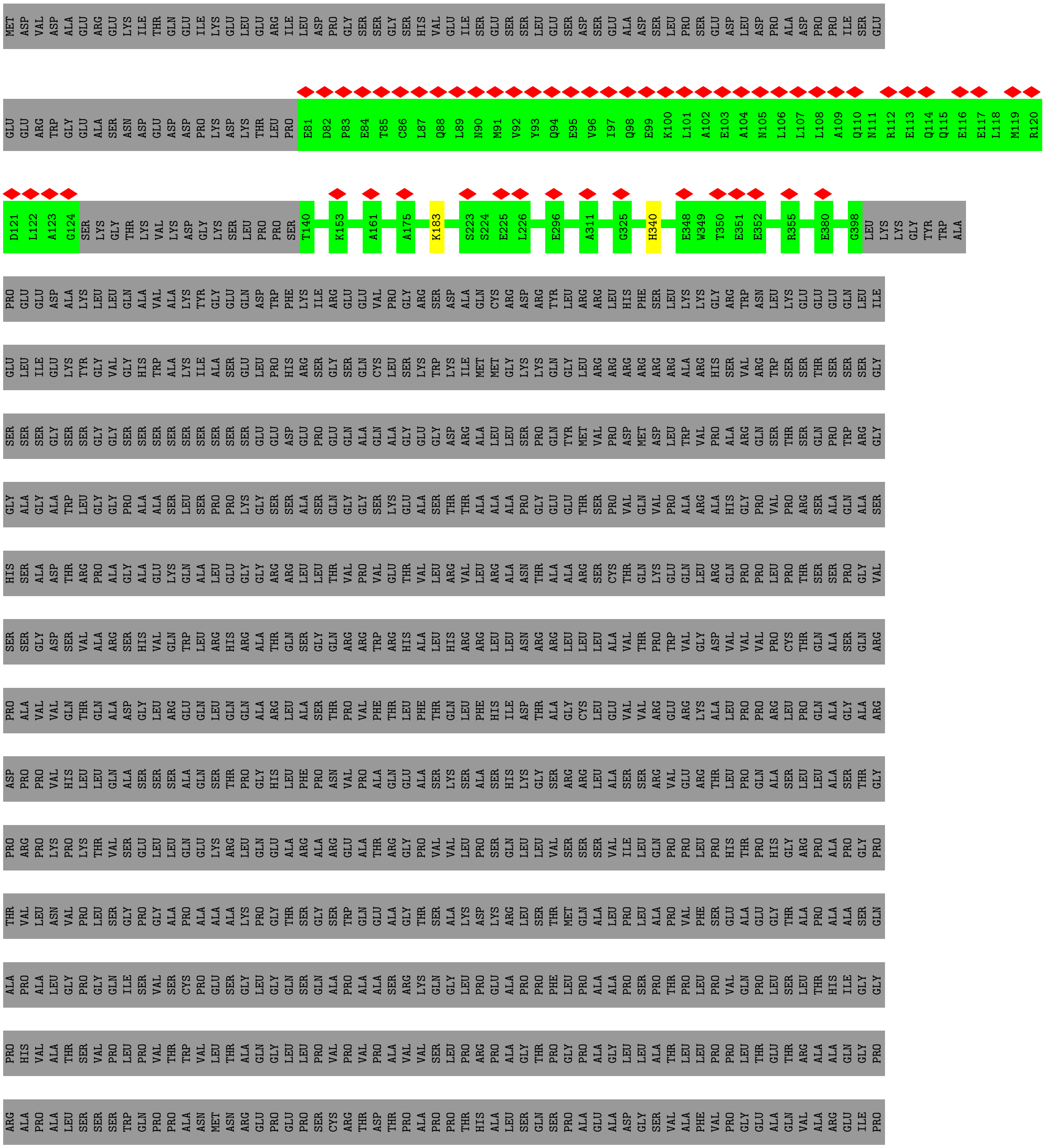
Chain b:





- Molecule 23: snRNA-activating protein complex subunit 4

Chain c:



GLU PRO ARG THR SER SER HIS ALA ASP PRO PRO GLU ALA GLU PRO PRO TRP SER GLY ARG LEU PRO ALA PHE GLY GLY VAL ILE PRO ALA THR GLU PRO ARG GLY THR PRO GLY SER PRO SER GLY THR GLN GLU PRO ARG GLY PRO LEU GLY LEU GLU LYS LEU PRO LEU ARG GLN PRO

GLY PRO GLU LYS GLY ALA LEU ASP LEU GLU LYS PRO PRO LEU PRO PRO GLN PRO GLY PRO GLU LYS GLY ALA LEU ASP LEU GLY LEU LEU SER GLN GLU GLY GLU ALA ALA THR GLN GLN TRP LEU GLY GLY GLN ARG GLY VAL ARG VAL PRO LEU LEU GLY SER ARG LEU PRO TYR GLN PRO

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

ALA ARG PHE LEU ALA ALA PHE THR LEU PRO ALA LEU LEU ALA THR LEU ALA PRO GLN GLY VAL ARG THR THR LEU SER VAL PRO SER ARG VAL GLY SER GLU SER GLU ASP GLU ASP LEU LEU SER GLU LEU GLU LEU ALA ASP ARG ASP GLY GLN PRO GLY CYS THR THR ALA THR CYS

PRO ILE GLN GLY ALA PRO ASP SER GLY LYS CYS SER ALA SER SER CYS LEU ASP THR SER ASN ASP PRO ASP ASP LEU ASP VAL LEU ARG THR ARG HIS ALA ARG HIS THR ARG LYS ARG ARG ARG LEU VAL

- Molecule 24: snRNA-activating protein complex subunit 5

Chain d:

MET LEU S3 R4 L5 L8 R9 K10 E11 E12 L15 L18 K19 A20 A21 L22 H23 D24 Q25 L26 N27 R28 L29 K30 V31 E32 E33 L34 A35 L36 Q37 S38 M39 I40 S41 S42 R43 R44 G45 D46 E47 M48 L49 S50 S51 H52 THR VAL PRO GLU GLN SER HIS ASP MET LEU VAL

HIS VAL ASP ASN GLU ALA SER ILE ASN GLN THR THR LEU GLU LEU SER THR LYS SER HIS VAL THR GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU SER ASP SER

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	137246	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	54.45	Depositor
Minimum defocus (nm)	300	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.177	Depositor
Minimum map value	-0.069	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.0132	Depositor
Map size (Å)	419.99997, 419.99997, 419.99997	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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: ZN, MG

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.26	0/11479	0.42	0/15496
2	B	0.28	0/9257	0.42	0/12493
3	C	0.29	0/2102	0.43	0/2857
4	D	0.24	0/1064	0.35	0/1428
5	E	0.26	0/1752	0.41	0/2366
6	F	0.27	0/646	0.41	0/871
7	G	0.27	0/1382	0.42	0/1874
8	H	0.29	0/1207	0.44	0/1628
9	I	0.26	0/949	0.43	0/1284
10	J	0.31	0/516	0.44	0/696
11	K	0.27	0/939	0.40	0/1271
12	L	0.29	0/378	0.44	0/500
13	M	0.26	0/2337	0.41	0/3154
14	N	0.83	8/1479 (0.5%)	0.88	0/2285
15	O	0.40	0/1448	0.49	0/1948
16	Q	0.24	0/1167	0.39	0/1576
17	R	0.24	0/1817	0.39	0/2445
18	T	0.81	6/1399 (0.4%)	0.85	0/2149
19	U	0.38	0/747	0.47	0/1005
20	V	0.43	0/795	0.53	0/1077
21	a	0.32	0/1851	0.45	0/2493
22	b	0.36	0/3048	0.49	0/4124
23	c	0.31	0/2563	0.44	0/3442
24	d	0.23	0/409	0.45	0/543
All	All	0.34	14/50731 (0.0%)	0.47	0/69005

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
20	V	0	1

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	N	-7	DG	C1'-N9	-7.94	1.36	1.47
14	N	-9	DG	C1'-N9	-7.84	1.36	1.47
18	T	-9	DA	C1'-N9	-7.51	1.36	1.47
18	T	5	DG	C1'-N9	-7.18	1.37	1.47
18	T	3	DA	C1'-N9	-6.79	1.37	1.47

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
20	V	48	LEU	Mainchain

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	11274	0	11406	128	0
2	B	9076	0	9116	64	0
3	C	2059	0	2007	23	0
4	D	1050	0	1033	4	0
5	E	1721	0	1737	18	0
6	F	636	0	665	7	0
7	G	1351	0	1358	16	0
8	H	1186	0	1147	4	0
9	I	928	0	859	13	0
10	J	507	0	523	4	0
11	K	920	0	942	12	0
12	L	373	0	378	4	0
13	M	2301	0	2316	45	0
14	N	1316	0	709	43	0
15	O	1422	0	1514	43	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	Q	1138	0	1103	8	0
17	R	1788	0	1819	32	0
18	T	1253	0	696	38	0
19	U	734	0	729	75	0
20	V	785	0	797	85	0
21	a	1807	0	1800	0	0
22	b	2972	0	2836	0	0
23	c	2517	0	2485	0	0
24	d	408	0	424	0	0
25	A	2	0	0	0	0
25	B	1	0	0	0	0
25	C	1	0	0	0	0
25	I	2	0	0	0	0
25	J	1	0	0	0	0
25	L	1	0	0	0	0
25	M	1	0	0	0	0
25	b	2	0	0	0	0
26	A	1	0	0	0	0
All	All	49534	0	48399	500	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 8.

The worst 5 of 500 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)
19:U:29:PHE:CE1	20:V:32:LEU:HD11	1.26	1.67
20:V:47:ALA:HB1	20:V:51:ARG:CD	1.63	1.29
15:O:189:ASN:ND2	19:U:376:TRP:HH2	1.29	1.25
19:U:16:SER:OG	20:V:51:ARG:NH1	1.70	1.24
19:U:29:PHE:CE1	20:V:32:LEU:CD1	2.22	1.22

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	1413/1970 (72%)	1375 (97%)	38 (3%)	0	100	100
2	B	1130/1174 (96%)	1091 (96%)	39 (4%)	0	100	100
3	C	253/275 (92%)	248 (98%)	5 (2%)	0	100	100
4	D	126/142 (89%)	124 (98%)	2 (2%)	0	100	100
5	E	207/210 (99%)	203 (98%)	4 (2%)	0	100	100
6	F	77/127 (61%)	77 (100%)	0	0	100	100
7	G	169/172 (98%)	167 (99%)	2 (1%)	0	100	100
8	H	146/150 (97%)	142 (97%)	4 (3%)	0	100	100
9	I	112/125 (90%)	106 (95%)	6 (5%)	0	100	100
10	J	62/67 (92%)	61 (98%)	1 (2%)	0	100	100
11	K	113/117 (97%)	110 (97%)	3 (3%)	0	100	100
12	L	42/58 (72%)	38 (90%)	4 (10%)	0	100	100
13	M	294/316 (93%)	289 (98%)	5 (2%)	0	100	100
15	O	177/339 (52%)	167 (94%)	10 (6%)	0	100	100
16	Q	134/517 (26%)	130 (97%)	4 (3%)	0	100	100
17	R	218/249 (88%)	215 (99%)	3 (1%)	0	100	100
19	U	84/376 (22%)	76 (90%)	8 (10%)	0	100	100
20	V	94/109 (86%)	86 (92%)	8 (8%)	0	100	100
21	a	211/368 (57%)	202 (96%)	9 (4%)	0	100	100
22	b	358/411 (87%)	303 (85%)	55 (15%)	0	100	100
23	c	299/1469 (20%)	277 (93%)	22 (7%)	0	100	100
24	d	48/98 (49%)	48 (100%)	0	0	100	100
All	All	5767/8839 (65%)	5535 (96%)	232 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	1254/1749 (72%)	1253 (100%)	1 (0%)	93 98
2	B	994/1027 (97%)	993 (100%)	1 (0%)	93 98
3	C	234/252 (93%)	234 (100%)	0	100 100
4	D	118/126 (94%)	117 (99%)	1 (1%)	81 93
5	E	191/192 (100%)	191 (100%)	0	100 100
6	F	69/111 (62%)	69 (100%)	0	100 100
7	G	152/153 (99%)	151 (99%)	1 (1%)	84 94
8	H	129/131 (98%)	129 (100%)	0	100 100
9	I	103/112 (92%)	103 (100%)	0	100 100
10	J	53/56 (95%)	52 (98%)	1 (2%)	57 84
11	K	104/106 (98%)	104 (100%)	0	100 100
12	L	41/55 (74%)	41 (100%)	0	100 100
13	M	253/268 (94%)	252 (100%)	1 (0%)	91 97
15	O	154/293 (53%)	154 (100%)	0	100 100
16	Q	121/448 (27%)	119 (98%)	2 (2%)	60 85
17	R	196/218 (90%)	195 (100%)	1 (0%)	88 96
19	U	82/324 (25%)	82 (100%)	0	100 100
20	V	88/98 (90%)	86 (98%)	2 (2%)	50 80
21	a	198/334 (59%)	196 (99%)	2 (1%)	76 91
22	b	323/356 (91%)	322 (100%)	1 (0%)	92 97
23	c	270/1213 (22%)	268 (99%)	2 (1%)	84 94
24	d	46/93 (50%)	46 (100%)	0	100 100
All	All	5173/7715 (67%)	5157 (100%)	16 (0%)	92 97

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
23	c	183	LYS
22	b	194	LYS
17	R	230	LYS
21	a	166	ASP
16	Q	151	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
2	B	980	HIS
22	b	319	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

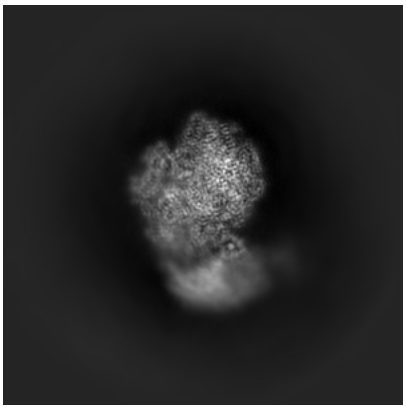
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-15007. These allow visual inspection of the internal detail of the map and identification of artifacts.

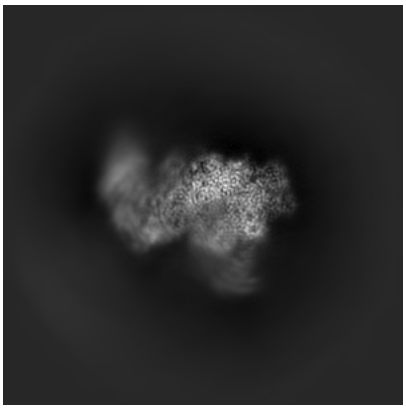
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

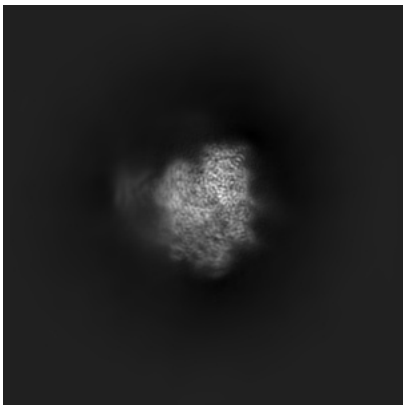
6.1.1 Primary map



X

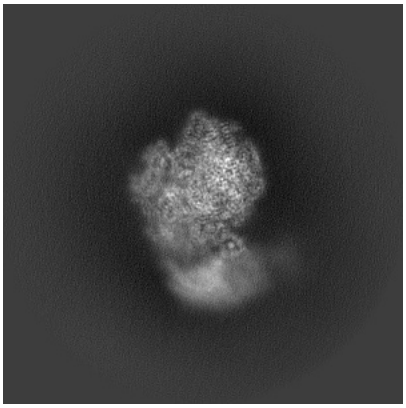


Y

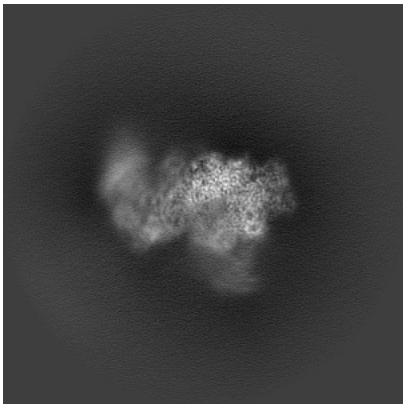


Z

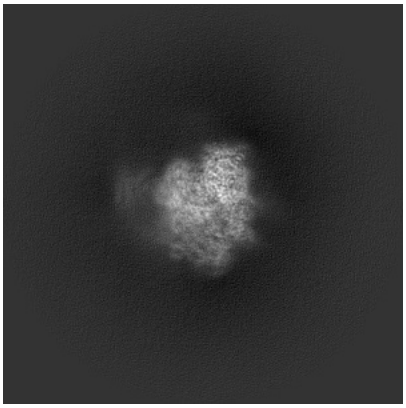
6.1.2 Raw map



X



Y

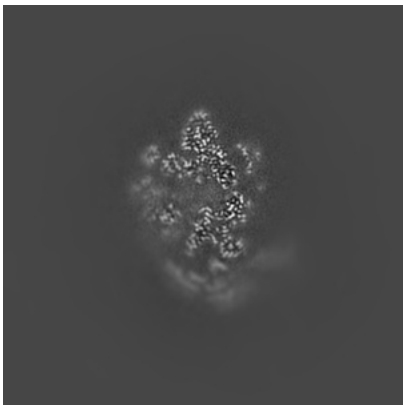


Z

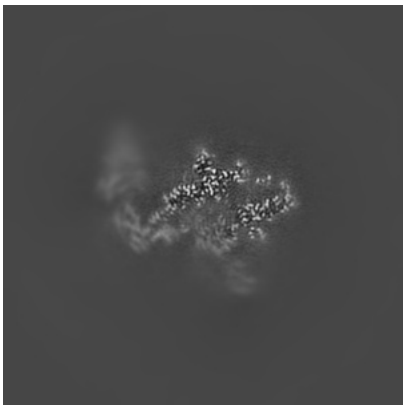
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

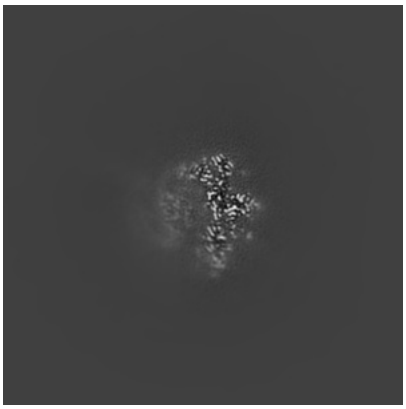
6.2.1 Primary map



X Index: 200

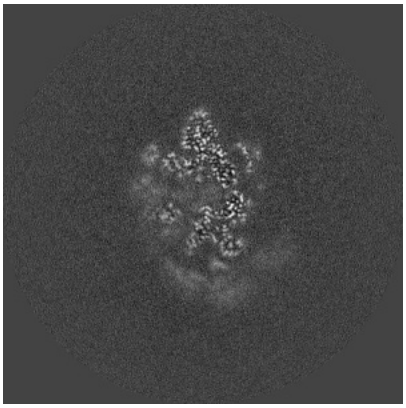


Y Index: 200

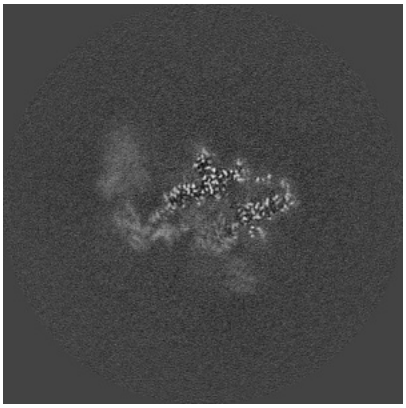


Z Index: 200

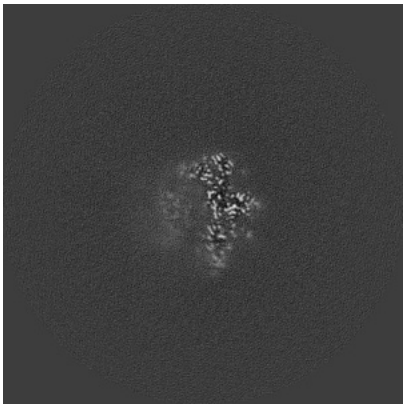
6.2.2 Raw map



X Index: 200



Y Index: 200

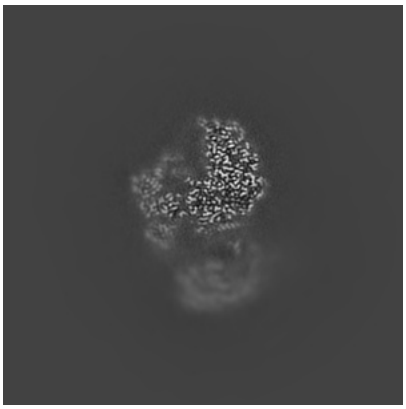


Z Index: 200

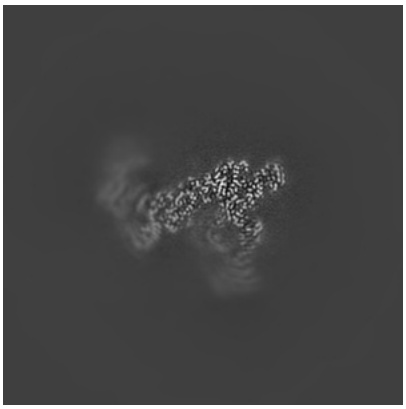
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

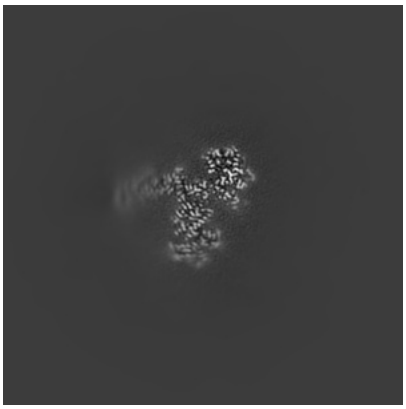
6.3.1 Primary map



X Index: 219

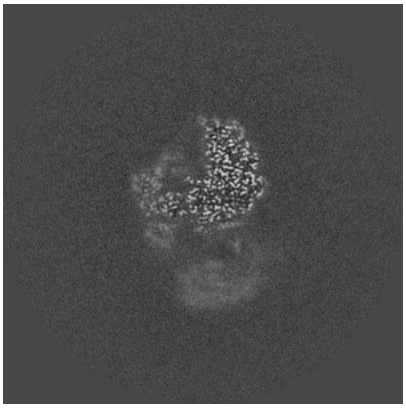


Y Index: 220

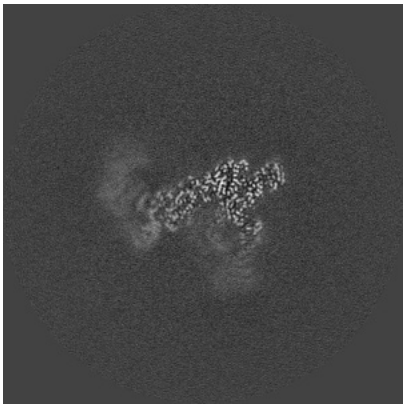


Z Index: 242

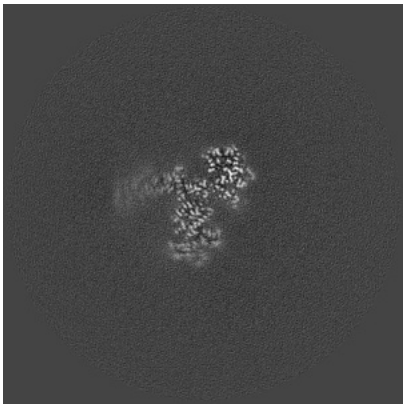
6.3.2 Raw map



X Index: 219



Y Index: 220

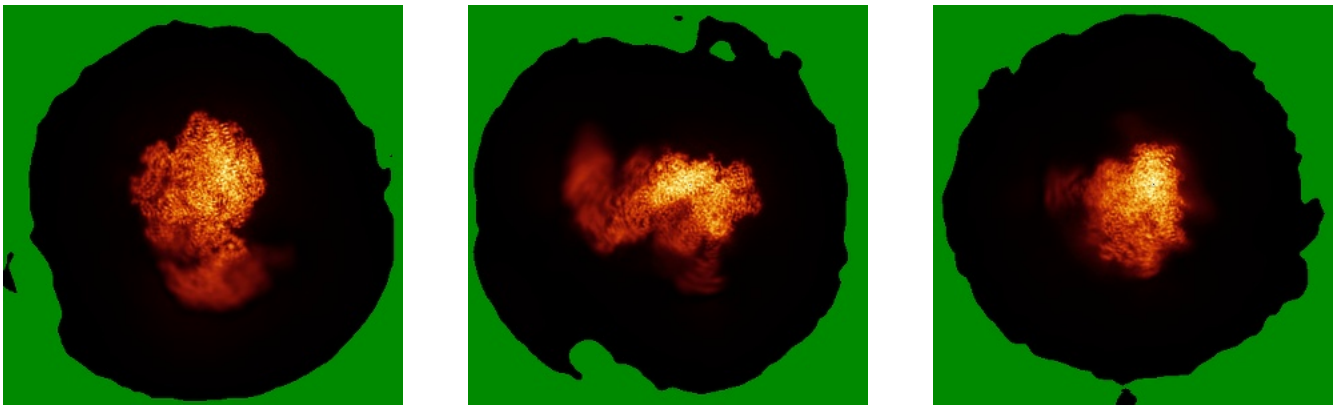


Z Index: 242

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

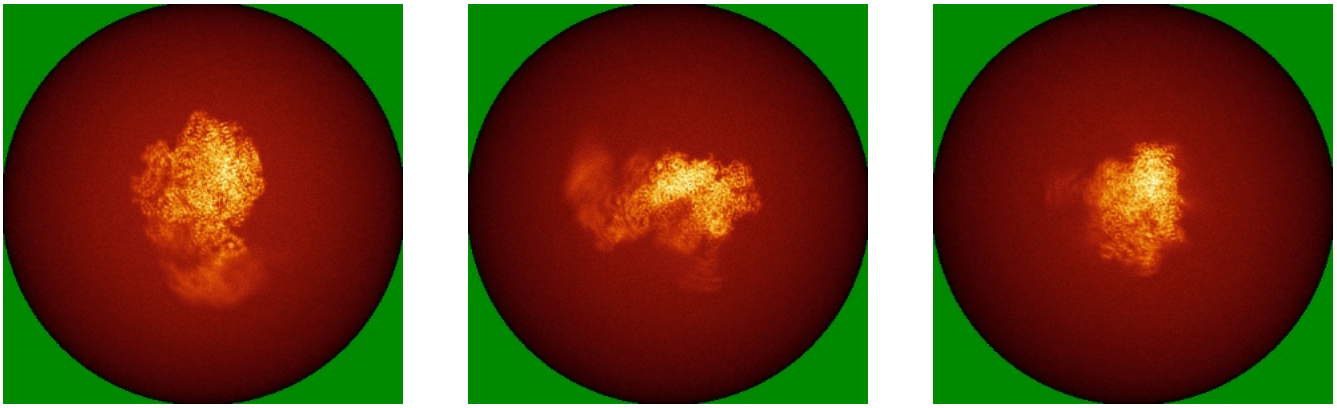
6.4 Orthogonal standard-deviation projections (False-color) (i)

6.4.1 Primary map



X Y Z

6.4.2 Raw 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.

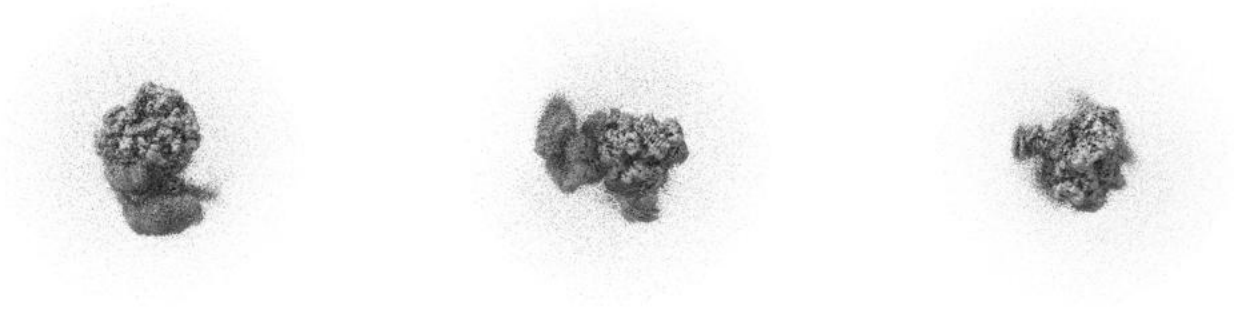
6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.0132. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map



X Y Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

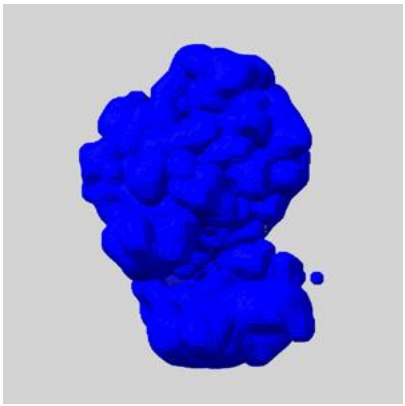
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

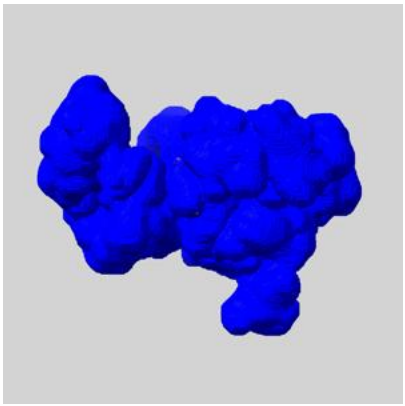
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

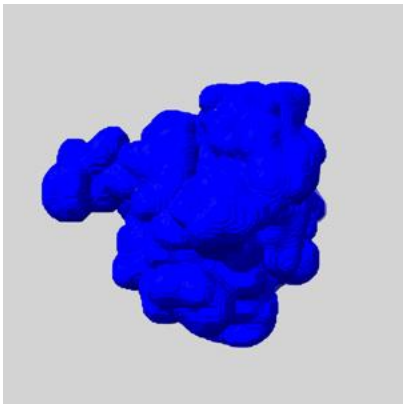
6.6.1 emd_15007_msk_1.map (i)



X



Y

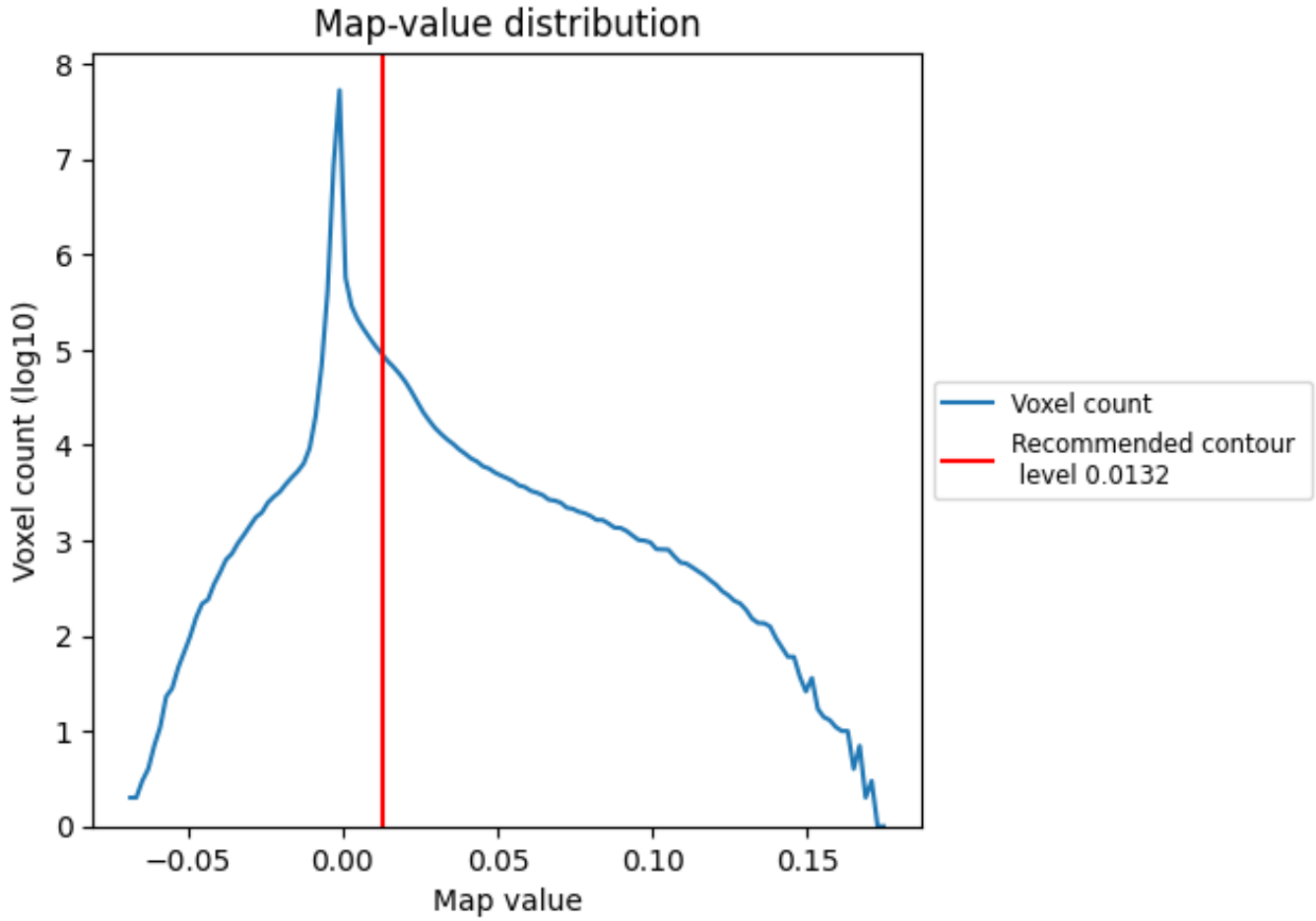


Z

7 Map analysis

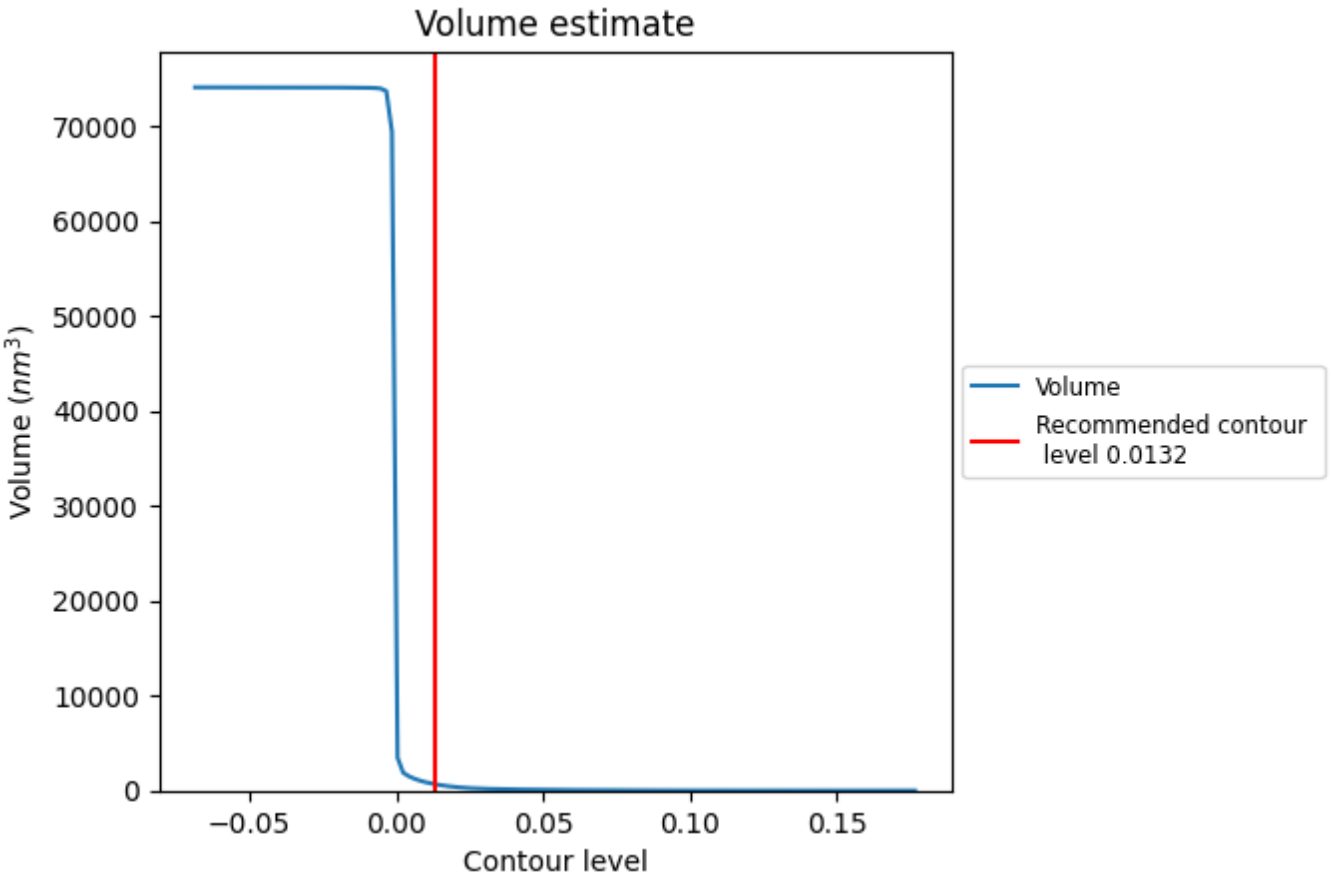
This section contains the results of statistical analysis of the map.

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.

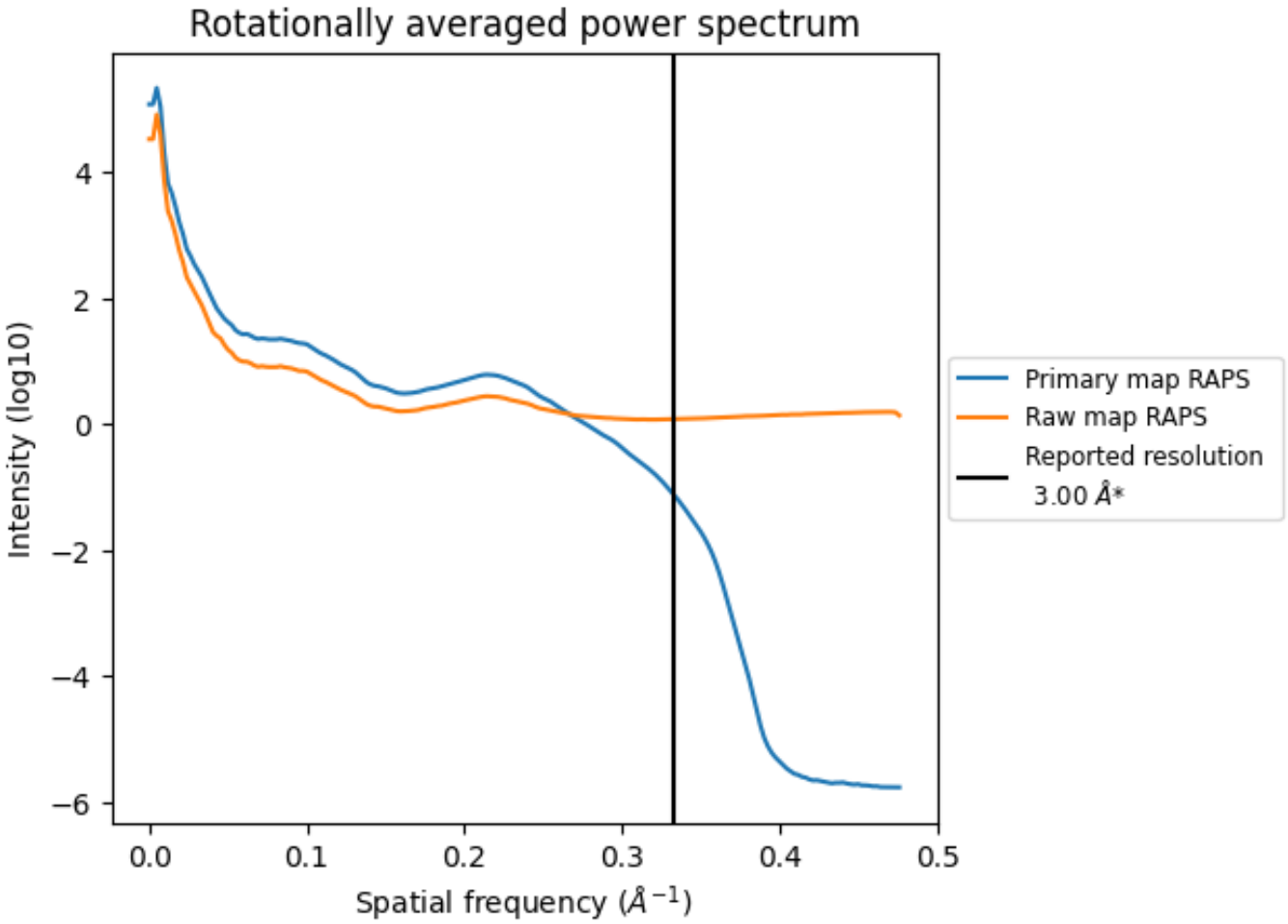
7.2 Volume estimate (i)



The volume at the recommended contour level is 667 nm^3; this corresponds to an approximate mass of 603 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)

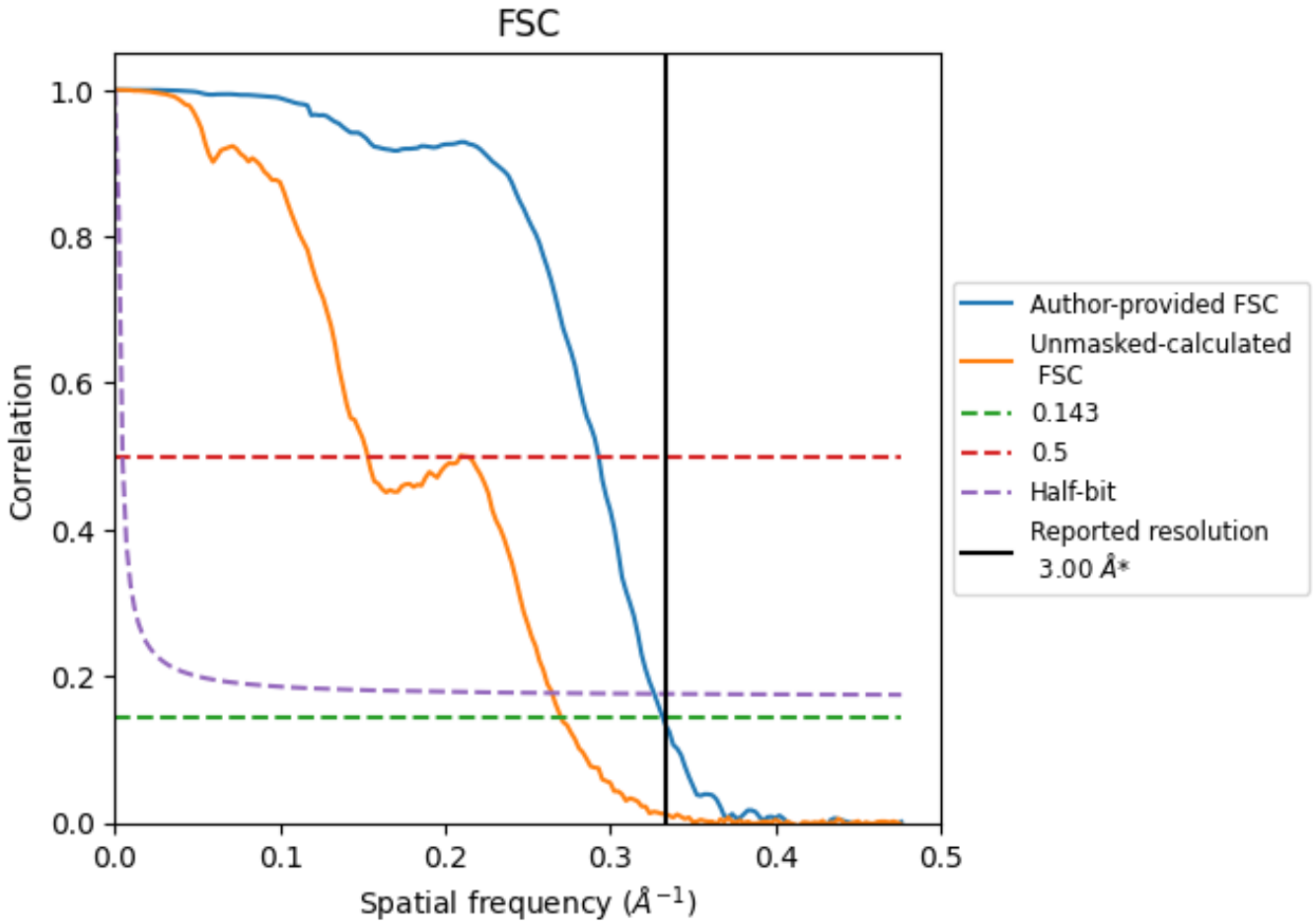


*Reported resolution corresponds to spatial frequency of 0.333 Å^{-1}

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)

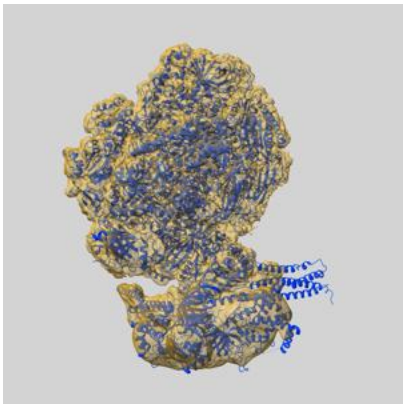
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.00	-	-
Author-provided FSC curve	3.01	3.41	3.06
Unmasked-calculated*	3.70	6.53	3.78

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.70 differs from the reported value 3.0 by more than 10 %

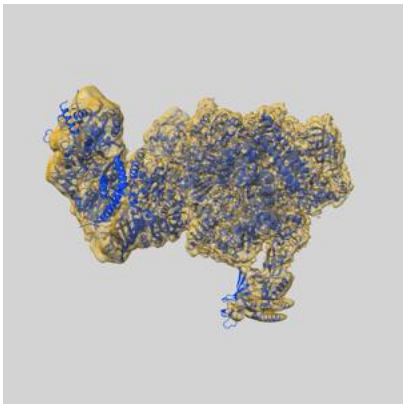
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-15007 and PDB model 7ZX8. Per-residue inclusion information can be found in section 3 on page 9.

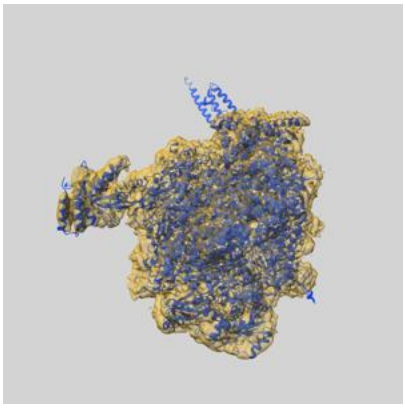
9.1 Map-model overlay



X



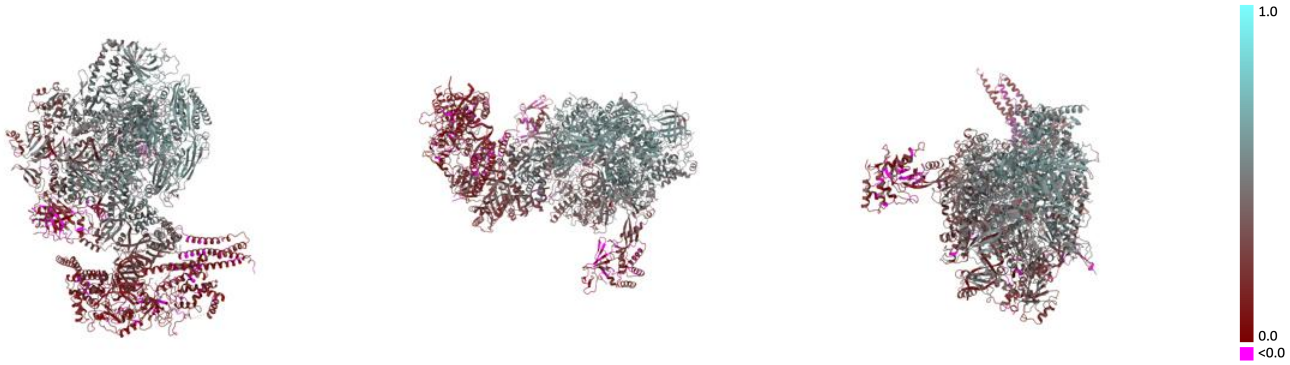
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0132 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.

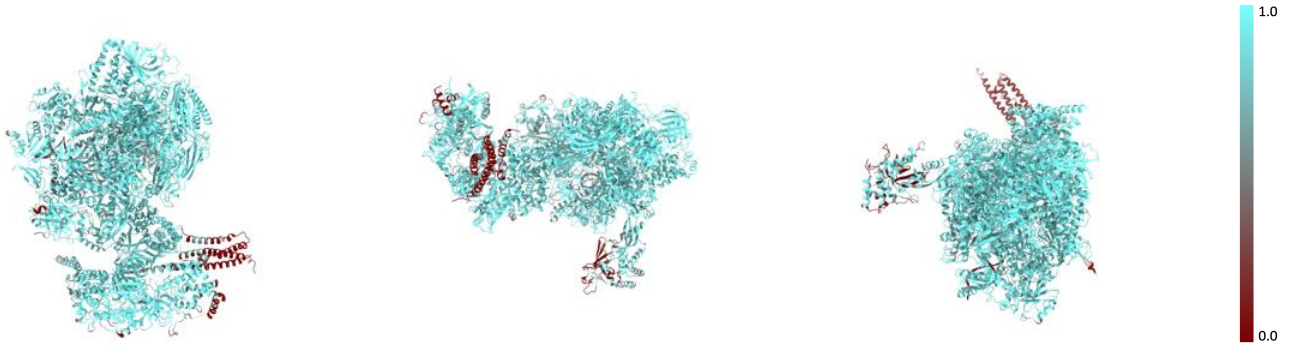
9.2 Q-score mapped to coordinate model (i)



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.

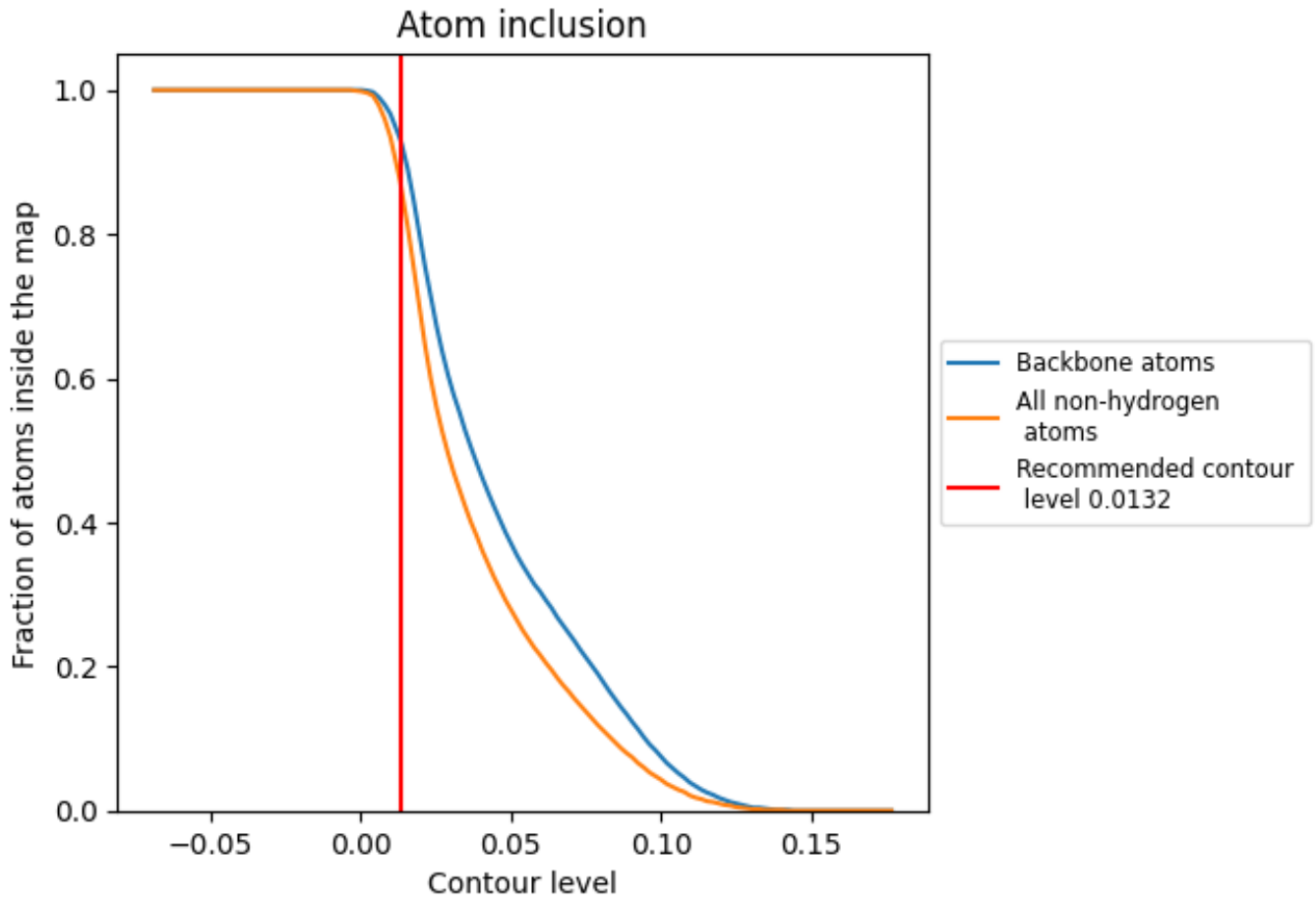
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.0132).

9.4 Atom inclusion (i)



At the recommended contour level, 93% of all backbone atoms, 87% 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 (0.0132) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8740	0.3560
A	0.9320	0.4520
B	0.9460	0.4910
C	0.9730	0.5450
D	0.6090	0.1440
E	0.9360	0.4120
F	0.9420	0.4680
G	0.5160	0.1670
H	0.9670	0.5110
I	0.9360	0.4030
J	0.9880	0.5700
K	0.9680	0.5530
L	0.9410	0.4500
M	0.8800	0.3990
N	0.8670	0.1920
O	0.9350	0.2850
Q	0.8120	0.1940
R	0.8230	0.1840
T	0.8140	0.2150
U	0.9140	0.1980
V	0.9030	0.1980
a	0.6470	0.1290
b	0.8450	0.1210
c	0.7340	0.1230
d	0.1440	0.0680