Full wwPDB EM Validation Report ⓘ

Dec 30, 2024 – 12:21 AM EST

PDB ID	:	7ZSB
EMDB ID	:	EMD-14929
Title	:	Yeast RNA polymerase II transcription pre-initiation complex with the +1 nucleosome and NTP, complex C
Authors	:	Wang, H.; Cramer, P.
Deposited on	:	2022-05-06
Resolution	:	6.60 Å(reported)

This is a Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

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

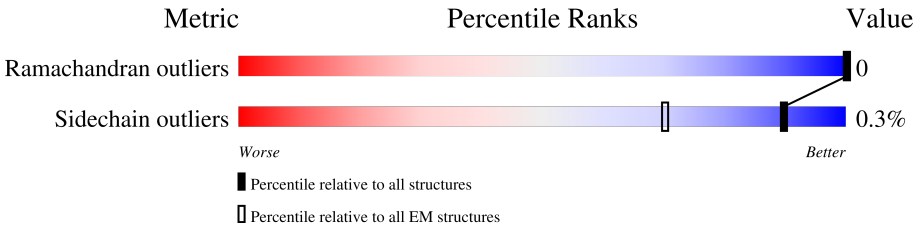
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 6.60 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)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1733	82% 18%
2	B	1224	96%
3	C	347	76% 23%
4	D	221	75% 24%
5	E	215	99%
6	F	155	75% 24%
7	G	177	97%
8	H	146	95%
9	I	122	95% 5%

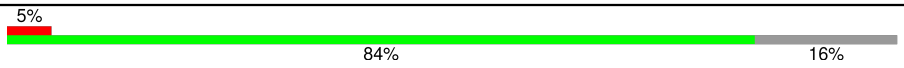
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	70	97%
11	K	120	95%
12	L	70	64% 36%
13	M	352	88% 12%
14	N	219	100%
15	O	247	73% 27%
16	Q	738	30% 70%
17	R	400	67% 33%
18	T	219	100%
19	U	171	62% 37%
20	V	129	79% 19%
21	W	492	5% 62% 38%
22	X	328	64% 36%
23	0	778	97%
24	1	645	80% 19%
25	2	517	87% 13%
26	3	324	40% 60%
27	4	341	89% 10%
28	5	76	83% 14%
29	6	464	82% 17%
30	7	843	73% 27%
31	a	135	72% 28%
31	e	135	73% 27%
32	b	102	79% 20%
32	f	102	78% 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
33	c	129	
33	g	129	
34	d	125	
34	h	125	

2 Entry composition (i)

There are 37 unique types of molecules in this entry. The entry contains 86808 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 II subunit RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1426	11221	7070	1960	2129	62	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1180	9404	5946	1643	1760	55	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	266	2092	1315	348	416	13	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-28	MET	-	initiating methionine	UNP P16370
C	-27	GLY	-	expression tag	UNP P16370
C	-26	SER	-	expression tag	UNP P16370
C	-25	HIS	-	expression tag	UNP P16370
C	-24	HIS	-	expression tag	UNP P16370
C	-23	HIS	-	expression tag	UNP P16370
C	-22	HIS	-	expression tag	UNP P16370
C	-21	HIS	-	expression tag	UNP P16370
C	-20	HIS	-	expression tag	UNP P16370
C	-19	SER	-	expression tag	UNP P16370
C	-18	ASN	-	expression tag	UNP P16370
C	-17	SER	-	expression tag	UNP P16370
C	-16	GLY	-	expression tag	UNP P16370
C	-15	LEU	-	expression tag	UNP P16370
C	-14	ASN	-	expression tag	UNP P16370

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-13	ASP	-	expression tag	UNP P16370
C	-12	ILE	-	expression tag	UNP P16370
C	-11	PHE	-	expression tag	UNP P16370
C	-10	GLU	-	expression tag	UNP P16370
C	-9	ALA	-	expression tag	UNP P16370
C	-8	GLN	-	expression tag	UNP P16370
C	-7	LYS	-	expression tag	UNP P16370
C	-6	ILE	-	expression tag	UNP P16370
C	-5	GLU	-	expression tag	UNP P16370
C	-4	TRP	-	expression tag	UNP P16370
C	-3	HIS	-	expression tag	UNP P16370
C	-2	GLU	-	expression tag	UNP P16370
C	-1	ASP	-	expression tag	UNP P16370
C	0	THR	-	expression tag	UNP P16370
C	1	GLY	-	expression tag	UNP P16370
C	2	SER	-	expression tag	UNP P16370
C	3	SER	-	expression tag	UNP P16370

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	167	Total	C	N	O	S	0	0
			1343	829	242	270	2		

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	118	Total	C	N	O	S	0	0
			977	620	161	193	3		

- 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
			1339	861	222	248	8		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	172	HIS	-	expression tag	UNP P34087
G	173	HIS	-	expression tag	UNP P34087
G	174	HIS	-	expression tag	UNP P34087
G	175	HIS	-	expression tag	UNP P34087
G	176	HIS	-	expression tag	UNP P34087
G	177	HIS	-	expression tag	UNP P34087

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

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	140	Total	C	N	O	S	0	0
			1120	704	188	224	4		

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	69	Total	C	N	O	S	0	0
			569	362	101	100	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	115	Total	C	N	O	S	0	0
			924	593	157	172	2		

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	45	Total	C	N	O	S	0	0
			359	221	71	63	4		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	310	Total	C	N	O	S	0	0
			2379	1504	408	449	18		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	346	LYS	-	expression tag	UNP P29055
M	347	HIS	-	expression tag	UNP P29055
M	348	HIS	-	expression tag	UNP P29055
M	349	HIS	-	expression tag	UNP P29055
M	350	HIS	-	expression tag	UNP P29055
M	351	HIS	-	expression tag	UNP P29055
M	352	HIS	-	expression tag	UNP P29055

- Molecule 14 is a DNA chain called Non-template DNA (219-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	219	Total	C	N	O	P	0	0
			4463	2131	794	1320	218		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	181	Total	C	N	O	S	0	0
			1422	925	243	248	6		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	241	LYS	-	expression tag	UNP P13393
O	242	HIS	-	expression tag	UNP P13393
O	243	HIS	-	expression tag	UNP P13393
O	244	HIS	-	expression tag	UNP P13393
O	245	HIS	-	expression tag	UNP P13393
O	246	HIS	-	expression tag	UNP P13393
O	247	HIS	-	expression tag	UNP P13393

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

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Q	221	Total	C	N	O	S	0	0
			1871	1179	346	339	7		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	-2	GLY	-	expression tag	UNP P41895
Q	-1	PRO	-	expression tag	UNP P41895
Q	0	GLY	-	expression tag	UNP P41895

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

Mol	Chain	Residues	Atoms					AltConf	Trace
17	R	268	Total	C	N	O	S	0	0
			2230	1409	392	419	10		

- Molecule 18 is a DNA chain called Template DNA (219-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	219	Total	C	N	O	P	0	0
			4510	2144	844	1304	218		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
19	U	107	Total	C	N	O	S	0	0
			885	559	147	176	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
20	V	104	Total	C	N	O	S	0	0
			815	511	136	164	4		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	123	LYS	-	expression tag	UNP P32774
V	124	HIS	-	expression tag	UNP P32774
V	125	HIS	-	expression tag	UNP P32774
V	126	HIS	-	expression tag	UNP P32774
V	127	HIS	-	expression tag	UNP P32774
V	128	HIS	-	expression tag	UNP P32774
V	129	HIS	-	expression tag	UNP P32774

- Molecule 21 is a protein called Transcription initiation factor IIE subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	W	304	Total	C	N	O	S	0	0
			2473	1558	431	477	7		

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	483	ALA	-	expression tag	UNP P36100
W	484	ALA	-	expression tag	UNP P36100
W	485	ALA	-	expression tag	UNP P36100
W	486	LEU	-	expression tag	UNP P36100
W	487	GLU	-	expression tag	UNP P36100
W	488	HIS	-	expression tag	UNP P36100
W	489	HIS	-	expression tag	UNP P36100
W	490	HIS	-	expression tag	UNP P36100
W	491	HIS	-	expression tag	UNP P36100
W	492	HIS	-	expression tag	UNP P36100

- Molecule 22 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	X	211	Total	C	N	O	S	0	0
			1708	1089	293	320	6		

- Molecule 23 is a protein called General transcription and DNA repair factor IIH helicase subunit XPD.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	0	752	Total	C	N	O	S	0	0
			6091	3882	1029	1142	38		

- Molecule 24 is a protein called General transcription and DNA repair factor IIH subunit TFB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	1	522	Total	C	N	O	S	0	0
			4214	2660	734	798	22		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	-2	GLY	-	expression tag	UNP P32776
1	-1	GLY	-	expression tag	UNP P32776
1	0	SER	-	expression tag	UNP P32776

- Molecule 25 is a protein called General transcription and DNA repair factor IIH subunit TFB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	2	452	Total	C	N	O	S	0	0
			3647	2354	600	677	16		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
2	-3	GLY	-	expression tag	UNP Q02939
2	-2	PRO	-	expression tag	UNP Q02939
2	-1	GLY	-	expression tag	UNP Q02939
2	0	SER	-	expression tag	UNP Q02939

- Molecule 26 is a protein called RNA polymerase II transcription factor B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	3	131	Total	C	N	O	S	0	0
			1089	692	180	209	8		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
3	-2	GLY	-	expression tag	UNP Q03290
3	-1	PRO	-	expression tag	UNP Q03290
3	0	HIS	-	expression tag	UNP Q03290

- Molecule 27 is a protein called General transcription and DNA repair factor IIH subunit TFB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	4	306	Total	C	N	O	S	0	0
			2372	1511	396	451	14		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
4	-2	SER	-	expression tag	UNP Q12004
4	-1	ASN	-	expression tag	UNP Q12004
4	0	ALA	-	expression tag	UNP Q12004

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

Mol	Chain	Residues	Atoms					AltConf	Trace
28	5	65	Total	C	N	O	S	0	0
			514	326	90	95	3		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
5	-3	GLY	-	expression tag	UNP Q3E7C1
5	-2	PRO	-	expression tag	UNP Q3E7C1
5	-1	GLY	-	expression tag	UNP Q3E7C1
5	0	SER	-	expression tag	UNP Q3E7C1

- Molecule 29 is a protein called General transcription and DNA repair factor IIH subunit SSL1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	6	383	Total	C	N	O	S	0	0
			3019	1915	523	552	29		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
6	-2	GLY	-	expression tag	UNP Q04673
6	-1	GLY	-	expression tag	UNP Q04673
6	0	SER	-	expression tag	UNP Q04673

- Molecule 30 is a protein called General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	7	616	Total	C	N	O	S	0	0
			4961	3158	861	915	27		

- Molecule 31 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	a	97	Total	C	N	O	S	0	0
			801	506	155	138	2		
31	e	98	Total	C	N	O	S	0	0
			810	512	157	139	2		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
a	102	ALA	GLY	conflict	UNP P84233
a	110	ALA	CYS	engineered mutation	UNP P84233
e	102	ALA	GLY	conflict	UNP P84233
e	110	ALA	CYS	engineered mutation	UNP P84233

- Molecule 32 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	b	82	Total	C	N	O	S	0	0
			653	412	127	113	1		
32	f	80	Total	C	N	O	S	0	0
			638	401	125	111	1		

- Molecule 33 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				AltConf	Trace
33	c	109	Total	C	N	O	0	0
			843	531	167	145		
33	g	106	Total	C	N	O	0	0
			818	516	160	142		

- Molecule 34 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	d	97	Total	C	N	O	S	0	0
			767	481	142	142	2		
34	h	95	Total	C	N	O	S	0	0
			745	469	134	140	2		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	29	THR	SER	conflict	UNP P02281
h	29	THR	SER	conflict	UNP P02281

- Molecule 35 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
35	A	2	Total	Zn	0
			2	2	
35	B	1	Total	Zn	0
			1	1	

Continued on next page...

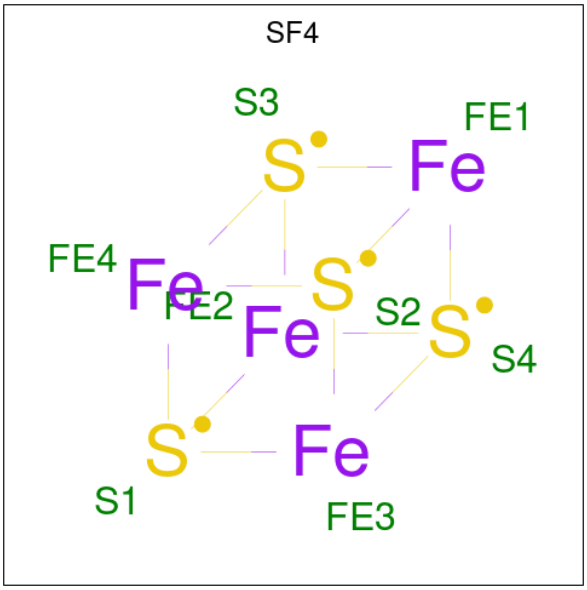
Continued from previous page...

Mol	Chain	Residues	Atoms	AltConf
35	C	1	Total Zn 1 1	0
35	I	2	Total Zn 2 2	0
35	J	1	Total Zn 1 1	0
35	L	1	Total Zn 1 1	0
35	M	1	Total Zn 1 1	0
35	W	1	Total Zn 1 1	0
35	3	2	Total Zn 2 2	0
35	4	1	Total Zn 1 1	0
35	6	4	Total Zn 4 4	0

- Molecule 36 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 37 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



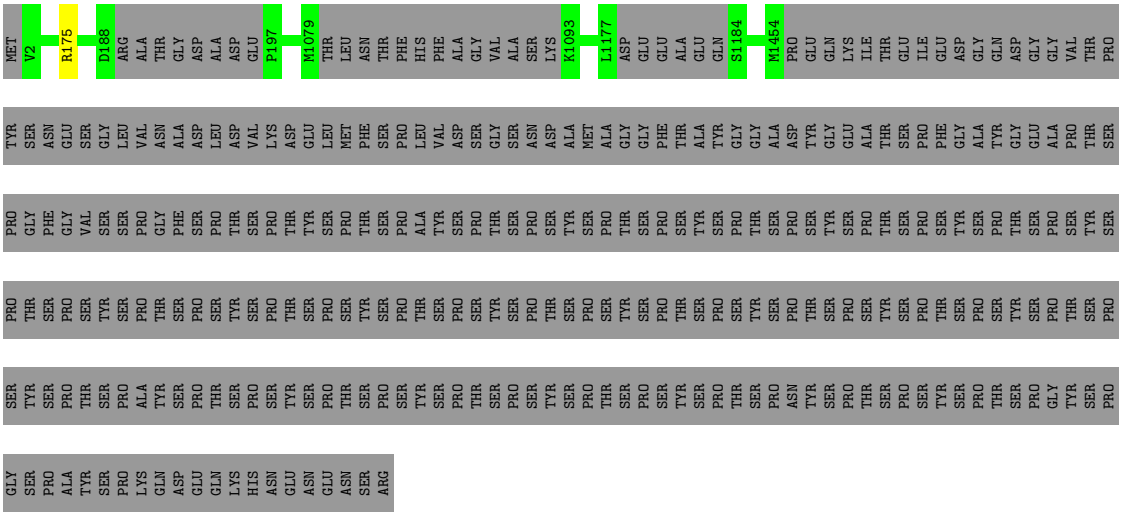
Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
37	0	1	8	4	4	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

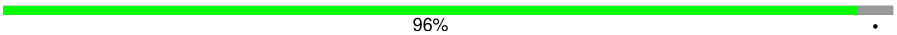
• Molecule 1: DNA-directed RNA polymerase II subunit RPB1

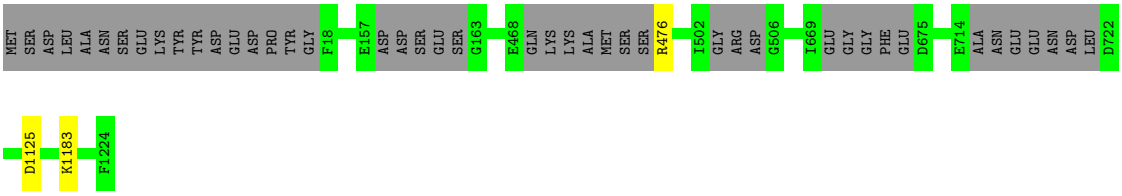
Chain A:



• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

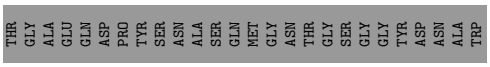
Chain B:





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

Chain C:

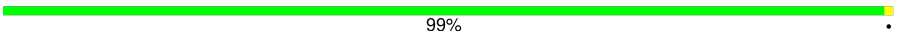


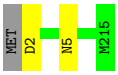
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D:

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

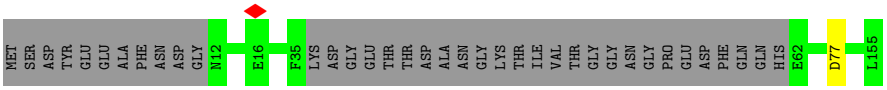
Chain E:





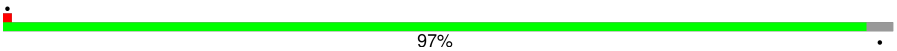
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

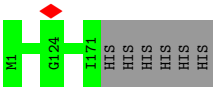
Chain F:



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

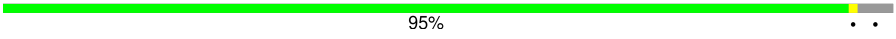
Chain G:

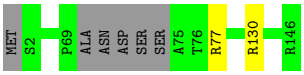




• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

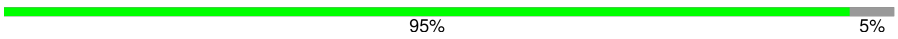
Chain H:

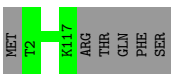




• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

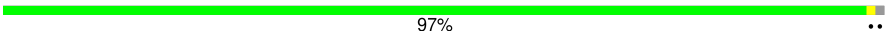
Chain I:

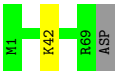




• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

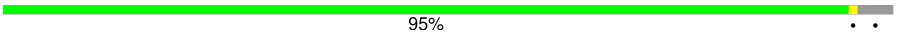
Chain J:

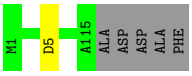




• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain K:





• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

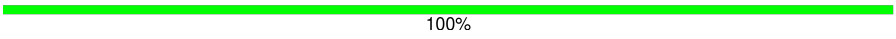
Chain L:

• Molecule 13: Transcription initiation factor IIB

Chain M:

• Molecule 14: Non-template DNA (219-MER)

Chain N:



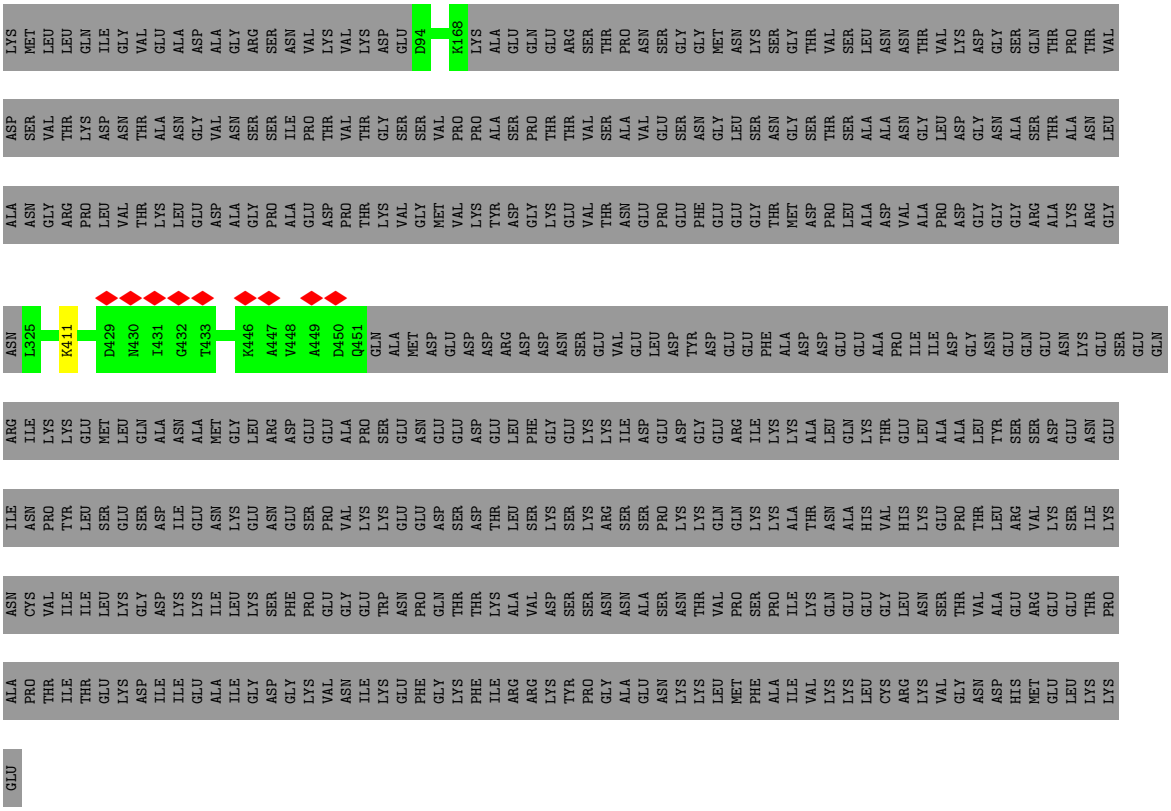
There are no outlier residues recorded for this chain.

• Molecule 15: TATA-box-binding protein

Chain O:

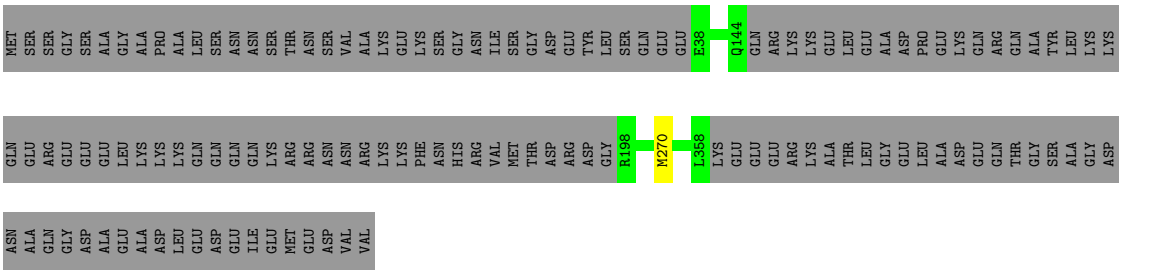
• Molecule 16: Transcription initiation factor IIF subunit alpha

Chain Q:



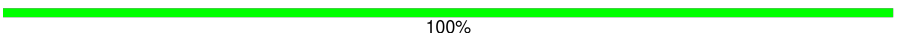
• Molecule 17: Transcription initiation factor IIF subunit beta

Chain R:



• Molecule 18: Template DNA (219-MER)

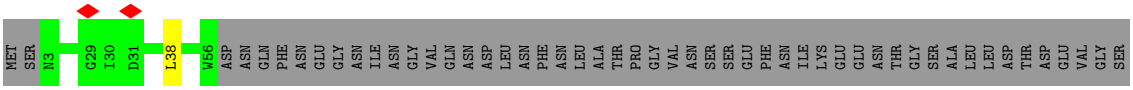
Chain T:

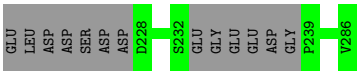


There are no outlier residues recorded for this chain.

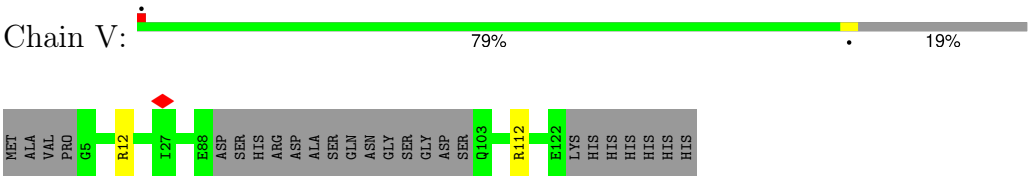
• Molecule 19: Transcription initiation factor IIA large subunit

Chain U:

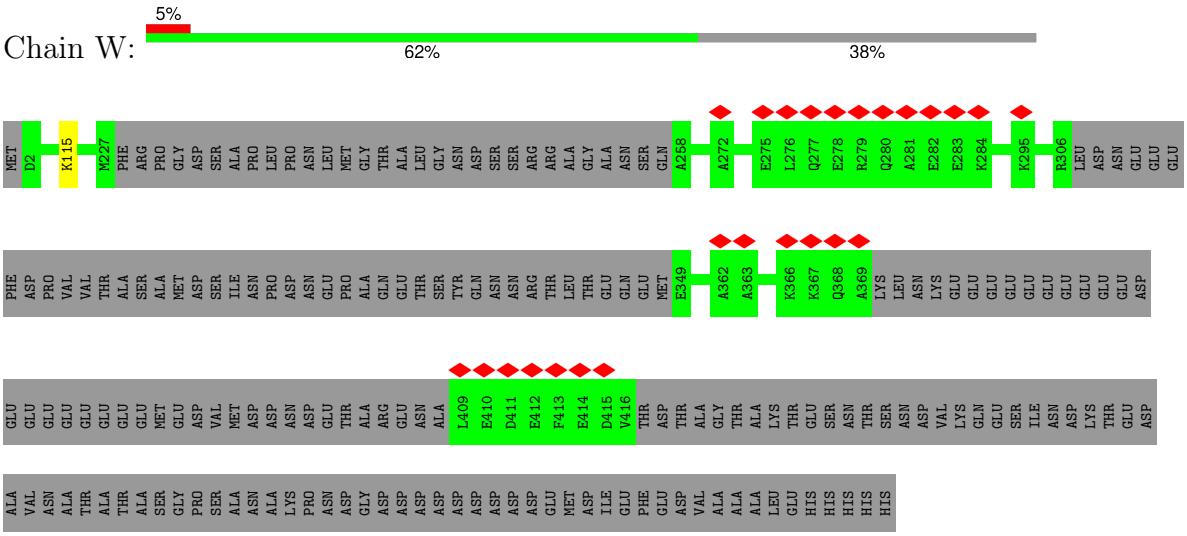




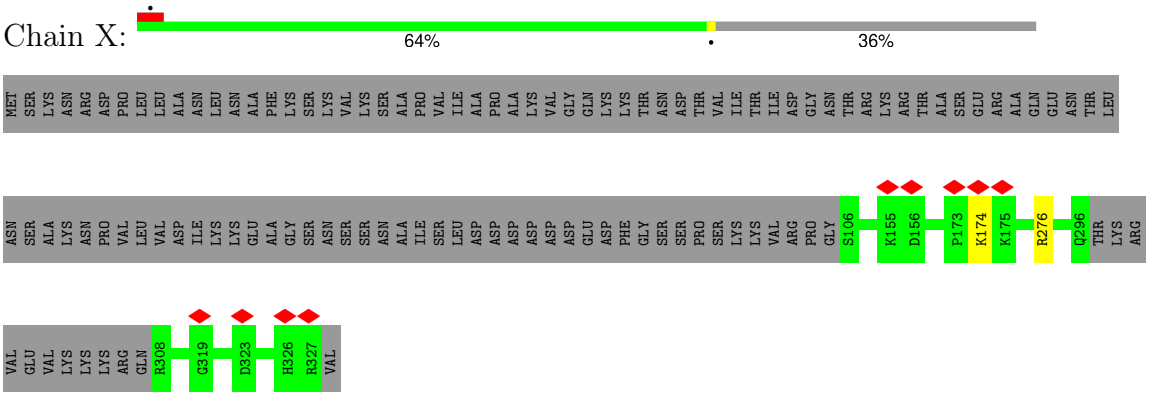
• Molecule 20: Transcription initiation factor IIA subunit 2



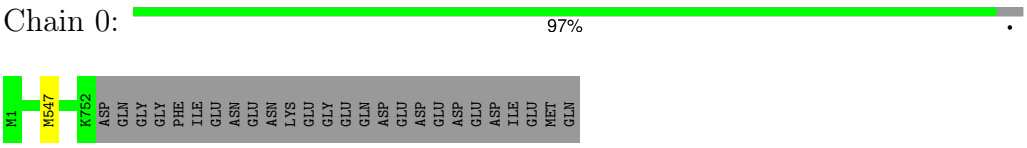
• Molecule 21: Transcription initiation factor IIE subunit alpha



• Molecule 22: Transcription initiation factor IIE subunit beta

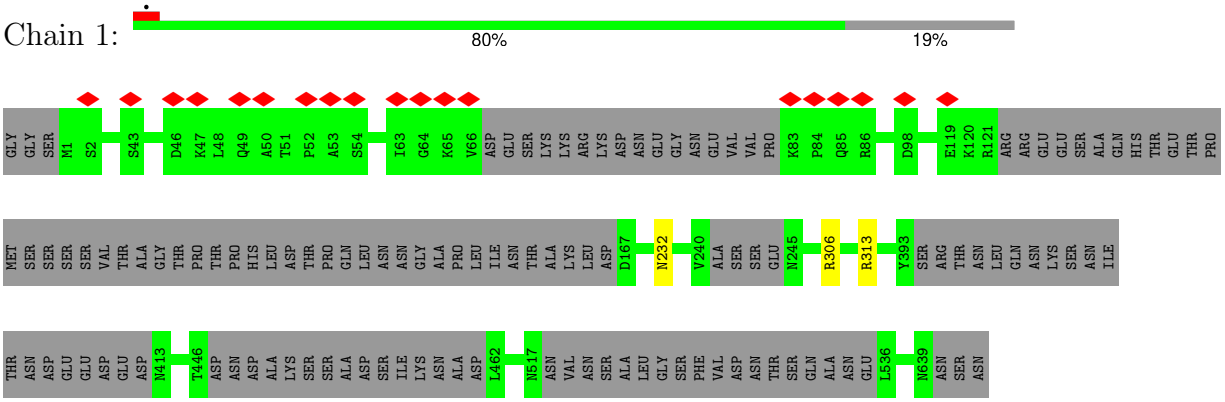


• Molecule 23: General transcription and DNA repair factor IIH helicase subunit XPD



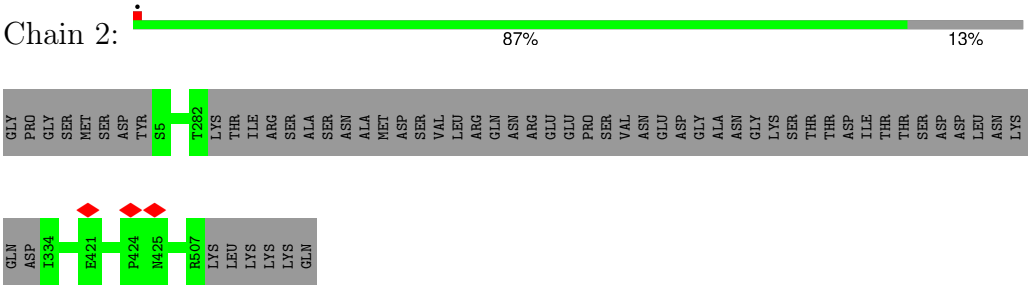
• Molecule 24: General transcription and DNA repair factor IIH subunit TFB1

Chain 1:



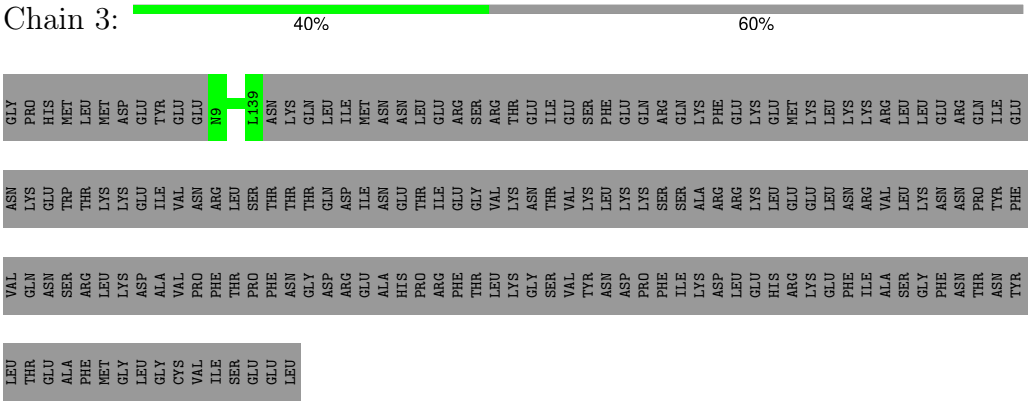
• Molecule 25: General transcription and DNA repair factor IIH subunit TFB2

Chain 2:



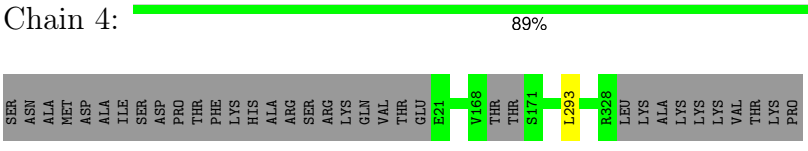
• Molecule 26: RNA polymerase II transcription factor B subunit 3

Chain 3:



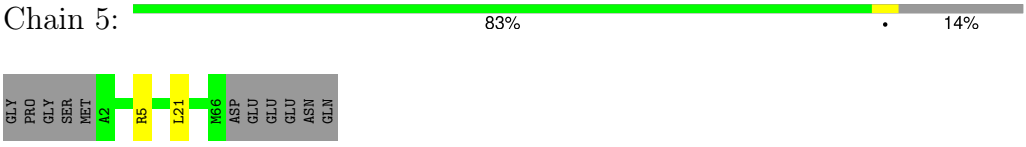
• Molecule 27: General transcription and DNA repair factor IIH subunit TFB4

Chain 4:



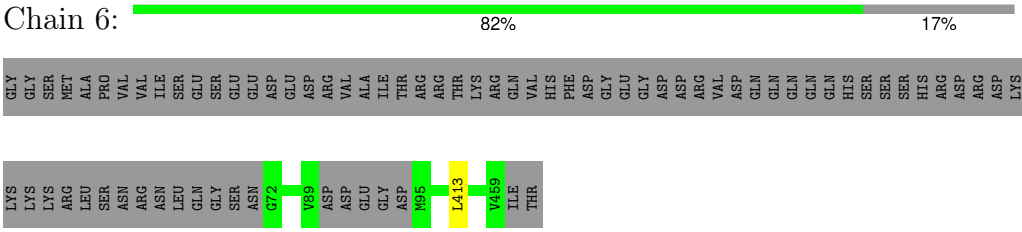
• Molecule 28: General transcription and DNA repair factor IIH subunit TFB5

Chain 5:



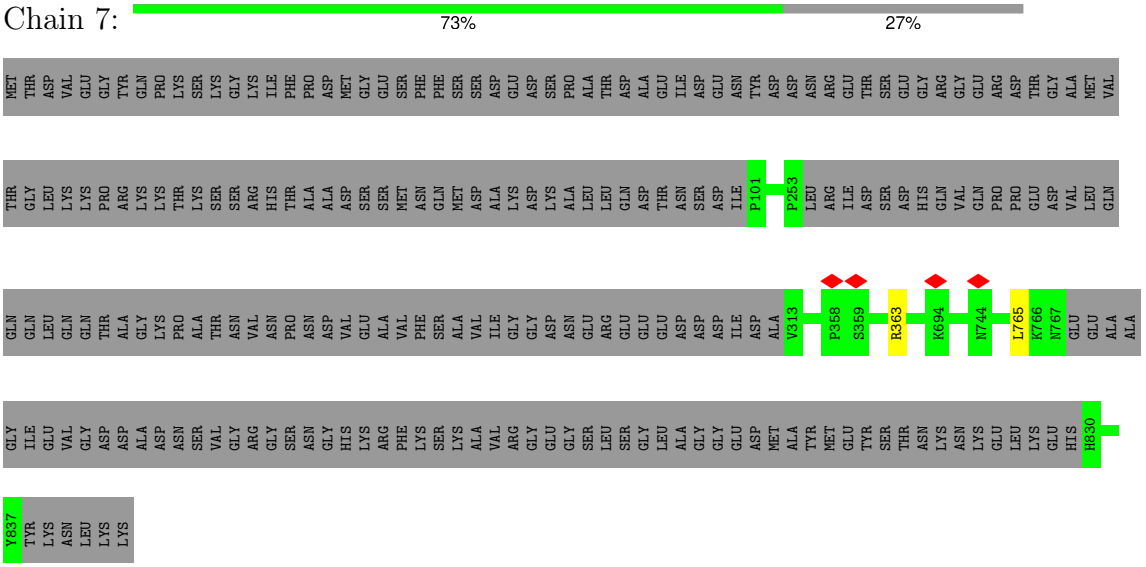
• Molecule 29: General transcription and DNA repair factor IIH subunit SSL1

Chain 6:



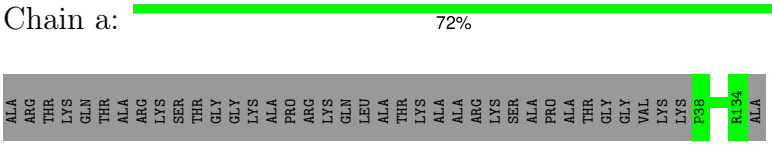
• Molecule 30: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 7:



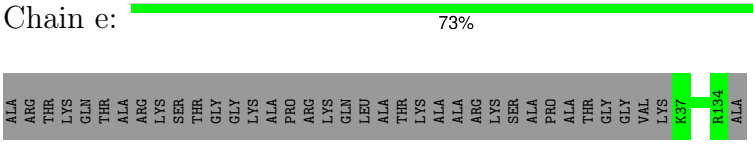
• Molecule 31: Histone H3.2

Chain a:



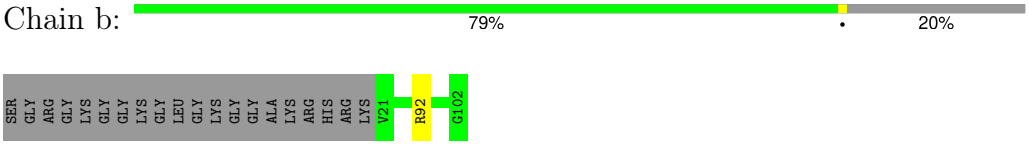
• Molecule 31: Histone H3.2

Chain e:



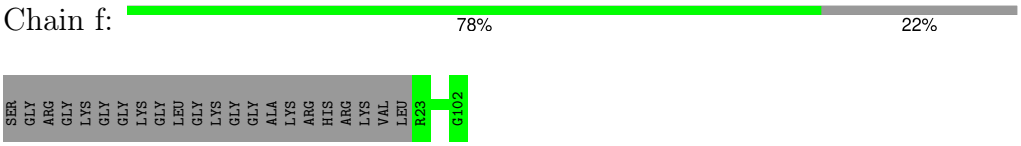
• Molecule 32: Histone H4

Chain b:



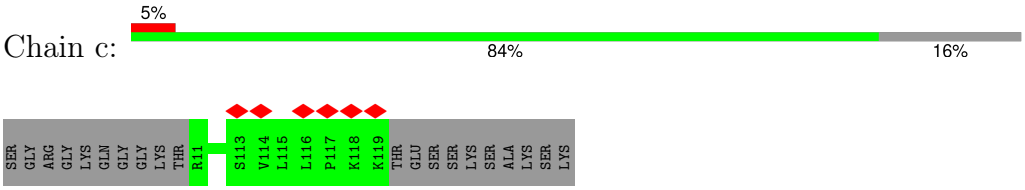
• Molecule 32: Histone H4

Chain f:



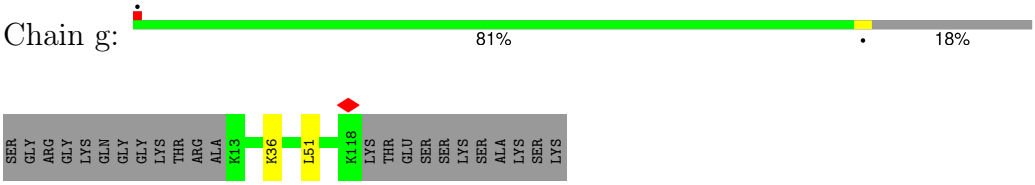
• Molecule 33: Histone H2A

Chain c:



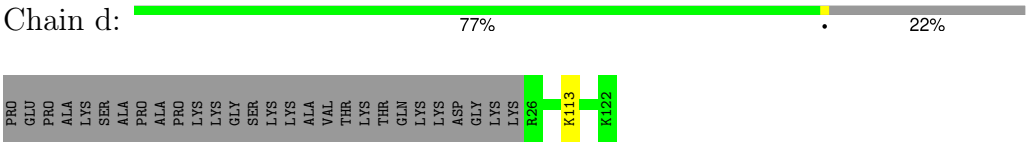
• Molecule 33: Histone H2A

Chain g:



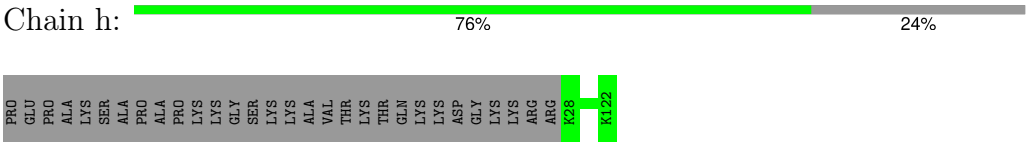
• Molecule 34: Histone H2B 1.1

Chain d:



• Molecule 34: Histone H2B 1.1

Chain h:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	82942	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	41	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.351	Depositor
Minimum map value	-0.171	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.007	Depositor
Map size (Å)	377.99997, 377.99997, 377.99997	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	Depositor

5 Model quality

5.1 Standard geometry

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

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	0.31	0/11422	0.54	0/15445
2	B	0.34	0/9589	0.56	1/12934 (0.0%)
3	C	0.34	0/2130	0.54	0/2887
4	D	0.25	0/1351	0.55	0/1811
5	E	0.30	0/1788	0.57	1/2406 (0.0%)
6	F	0.33	0/995	0.63	1/1340 (0.1%)
7	G	0.29	0/1367	0.56	0/1844
8	H	0.34	0/1139	0.58	0/1544
9	I	0.31	0/962	0.57	0/1295
10	J	0.37	0/578	0.65	0/775
11	K	0.34	0/942	0.62	0/1272
12	L	0.37	0/361	0.72	0/478
13	M	0.28	0/2408	0.54	1/3241 (0.0%)
14	N	0.53	0/4999	0.96	0/7709
15	O	0.29	0/1449	0.58	0/1952
16	Q	0.26	0/1907	0.58	0/2556
17	R	0.25	0/2270	0.56	1/3052 (0.0%)
18	T	0.53	0/5067	0.94	0/7824
19	U	0.25	0/898	0.56	1/1212 (0.1%)
20	V	0.25	0/822	0.55	0/1109
21	W	0.25	0/2513	0.53	0/3388
22	X	0.24	0/1739	0.53	0/2339
23	0	0.28	0/6209	0.57	1/8384 (0.0%)
24	1	0.25	0/4277	0.51	0/5755
25	2	0.27	0/3717	0.57	0/5028
26	3	0.30	0/1109	0.55	0/1492
27	4	0.30	0/2411	0.59	1/3260 (0.0%)
28	5	0.27	0/520	0.66	1/701 (0.1%)
29	6	0.29	0/3082	0.57	1/4165 (0.0%)
30	7	0.27	0/5067	0.55	1/6853 (0.0%)
31	a	0.26	0/813	0.58	0/1091
31	e	0.28	0/822	0.60	0/1103

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
32	b	0.29	0/660	0.63	0/883
32	f	0.26	0/645	0.58	0/862
33	c	0.27	0/853	0.55	0/1149
33	g	0.27	0/828	0.62	1/1117 (0.1%)
34	d	0.26	0/778	0.56	0/1043
34	h	0.31	0/756	0.60	0/1015
All	All	0.33	0/89243	0.62	12/122314 (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
11	K	0	1

There are no bond length outliers.

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
33	g	51	LEU	CA-CB-CG	6.54	130.34	115.30
2	B	1125	ASP	CB-CG-OD1	5.97	123.67	118.30
27	4	293	LEU	CA-CB-CG	5.94	128.97	115.30
17	R	270	MET	CA-CB-CG	5.85	123.24	113.30
13	M	250	MET	CB-CG-SD	-5.80	95.00	112.40
23	0	547	MET	CA-CB-CG	5.65	122.91	113.30
29	6	413	LEU	CA-CB-CG	5.63	128.25	115.30
28	5	21	LEU	CA-CB-CG	5.38	127.68	115.30
19	U	38	LEU	CA-CB-CG	5.24	127.35	115.30
6	F	77	ASP	CB-CG-OD2	5.22	123.00	118.30
5	E	2	ASP	CB-CG-OD2	5.09	122.88	118.30
30	7	765	LEU	CA-CB-CG	5.02	126.84	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
11	K	5	ASP	Peptide

5.2 Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	1418/1733 (82%)	1379 (97%)	39 (3%)	0	100	100
2	B	1168/1224 (95%)	1121 (96%)	47 (4%)	0	100	100
3	C	264/347 (76%)	254 (96%)	10 (4%)	0	100	100
4	D	163/221 (74%)	163 (100%)	0	0	100	100
5	E	212/215 (99%)	204 (96%)	8 (4%)	0	100	100
6	F	114/155 (74%)	109 (96%)	5 (4%)	0	100	100
7	G	169/177 (96%)	165 (98%)	4 (2%)	0	100	100
8	H	136/146 (93%)	134 (98%)	2 (2%)	0	100	100
9	I	114/122 (93%)	109 (96%)	5 (4%)	0	100	100
10	J	67/70 (96%)	65 (97%)	2 (3%)	0	100	100
11	K	113/120 (94%)	112 (99%)	1 (1%)	0	100	100
12	L	43/70 (61%)	40 (93%)	3 (7%)	0	100	100
13	M	306/352 (87%)	298 (97%)	8 (3%)	0	100	100
15	O	179/247 (72%)	175 (98%)	4 (2%)	0	100	100
16	Q	215/738 (29%)	208 (97%)	7 (3%)	0	100	100
17	R	264/400 (66%)	258 (98%)	6 (2%)	0	100	100
19	U	101/171 (59%)	98 (97%)	3 (3%)	0	100	100
20	V	100/129 (78%)	99 (99%)	1 (1%)	0	100	100
21	W	296/492 (60%)	291 (98%)	5 (2%)	0	100	100
22	X	207/328 (63%)	204 (99%)	3 (1%)	0	100	100
23	0	750/778 (96%)	730 (97%)	20 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
24	1	508/645 (79%)	501 (99%)	7 (1%)	0	100	100
25	2	448/517 (87%)	441 (98%)	7 (2%)	0	100	100
26	3	129/324 (40%)	122 (95%)	7 (5%)	0	100	100
27	4	302/341 (89%)	290 (96%)	12 (4%)	0	100	100
28	5	63/76 (83%)	62 (98%)	1 (2%)	0	100	100
29	6	379/464 (82%)	363 (96%)	16 (4%)	0	100	100
30	7	610/843 (72%)	590 (97%)	20 (3%)	0	100	100
31	a	95/135 (70%)	93 (98%)	2 (2%)	0	100	100
31	e	96/135 (71%)	96 (100%)	0	0	100	100
32	b	80/102 (78%)	79 (99%)	1 (1%)	0	100	100
32	f	78/102 (76%)	77 (99%)	1 (1%)	0	100	100
33	c	107/129 (83%)	105 (98%)	2 (2%)	0	100	100
33	g	104/129 (81%)	100 (96%)	4 (4%)	0	100	100
34	d	95/125 (76%)	94 (99%)	1 (1%)	0	100	100
34	h	93/125 (74%)	89 (96%)	4 (4%)	0	100	100
All	All	9586/12427 (77%)	9318 (97%)	268 (3%)	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	1248/1520 (82%)	1247 (100%)	1 (0%)	92	95
2	B	1024/1061 (96%)	1022 (100%)	2 (0%)	92	94
3	C	234/299 (78%)	233 (100%)	1 (0%)	89	91
4	D	149/200 (74%)	148 (99%)	1 (1%)	81	87
5	E	196/197 (100%)	195 (100%)	1 (0%)	86	89
6	F	107/137 (78%)	107 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	G	152/158 (96%)	152 (100%)	0	100 100
8	H	123/128 (96%)	121 (98%)	2 (2%)	58 74
9	I	110/116 (95%)	110 (100%)	0	100 100
10	J	64/65 (98%)	63 (98%)	1 (2%)	58 74
11	K	99/102 (97%)	99 (100%)	0	100 100
12	L	40/57 (70%)	40 (100%)	0	100 100
13	M	267/306 (87%)	267 (100%)	0	100 100
15	O	153/212 (72%)	153 (100%)	0	100 100
16	Q	204/642 (32%)	202 (99%)	2 (1%)	73 82
17	R	252/363 (69%)	252 (100%)	0	100 100
19	U	99/154 (64%)	99 (100%)	0	100 100
20	V	94/115 (82%)	92 (98%)	2 (2%)	48 66
21	W	275/436 (63%)	274 (100%)	1 (0%)	89 91
22	X	193/295 (65%)	191 (99%)	2 (1%)	73 82
23	0	684/707 (97%)	684 (100%)	0	100 100
24	1	483/590 (82%)	480 (99%)	3 (1%)	84 88
25	2	414/470 (88%)	414 (100%)	0	100 100
26	3	125/305 (41%)	125 (100%)	0	100 100
27	4	271/302 (90%)	271 (100%)	0	100 100
28	5	59/68 (87%)	58 (98%)	1 (2%)	56 72
29	6	346/419 (83%)	346 (100%)	0	100 100
30	7	548/737 (74%)	547 (100%)	1 (0%)	92 94
31	a	84/109 (77%)	84 (100%)	0	100 100
31	e	85/109 (78%)	85 (100%)	0	100 100
32	b	67/78 (86%)	66 (98%)	1 (2%)	60 75
32	f	65/78 (83%)	65 (100%)	0	100 100
33	c	86/101 (85%)	86 (100%)	0	100 100
33	g	84/101 (83%)	83 (99%)	1 (1%)	67 79
34	d	83/105 (79%)	82 (99%)	1 (1%)	67 79
34	h	81/105 (77%)	81 (100%)	0	100 100
All	All	8648/10947 (79%)	8624 (100%)	24 (0%)	90 92

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	175	ARG
2	B	476	ARG
2	B	1183	LYS
3	C	84	ARG
4	D	14	ARG
5	E	5	ASN
8	H	77	ARG
8	H	130	ARG
10	J	42	LYS
16	Q	29	ARG
16	Q	411	LYS
20	V	12	ARG
20	V	112	ARG
21	W	115	LYS
22	X	174	LYS
22	X	276	ARG
24	1	232	ASN
24	1	306	ARG
24	1	313	ARG
28	5	5	ARG
30	7	363	ARG
32	b	92	ARG
34	d	113	LYS
33	g	36	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	171	GLN
1	A	510	GLN
1	A	736	ASN
2	B	776	GLN
2	B	1097	HIS
2	B	1141	HIS
4	D	132	GLN
6	F	63	GLN
13	M	127	GLN
13	M	303	GLN
17	R	328	HIS
20	V	17	ASN
22	X	270	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
25	2	498	ASN
27	4	246	GLN
30	7	634	GLN
30	7	761	GLN
34	h	60	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

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

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
37	SF4	0	801	23	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
37	SF4	0	801	23	-	-	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.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

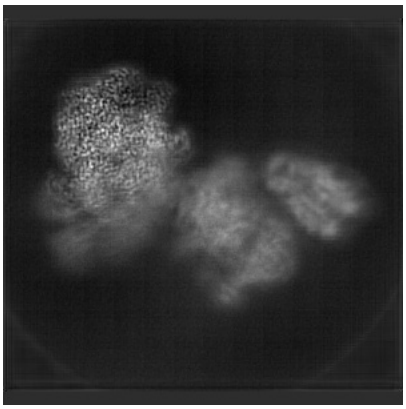
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-14929. These allow visual inspection of the internal detail of the map and identification of artifacts.

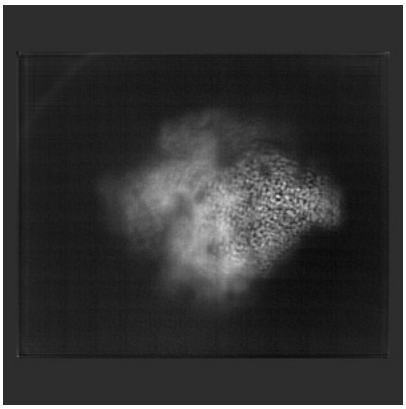
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

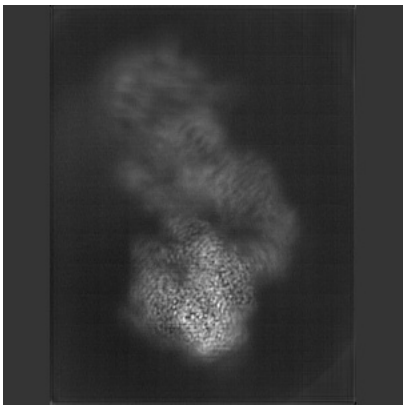
6.1.1 Primary map



X

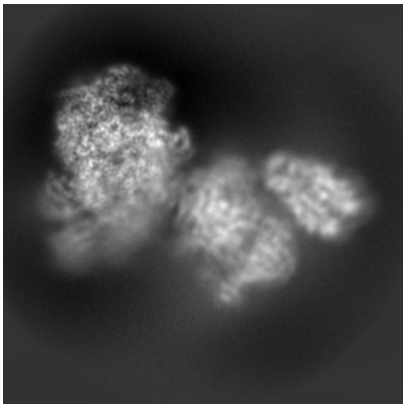


Y

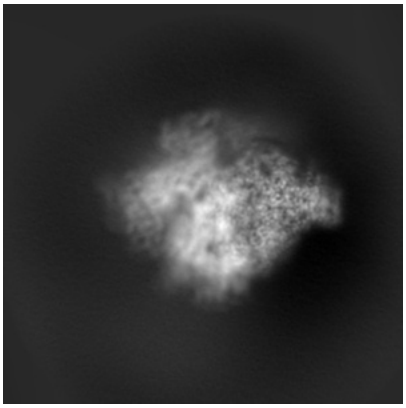


Z

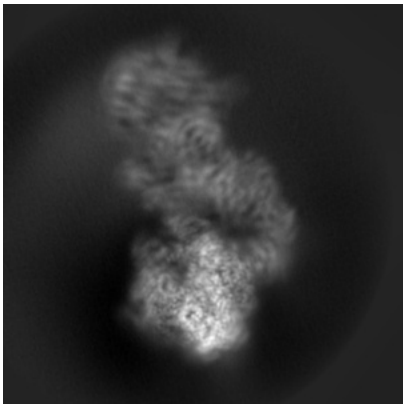
6.1.2 Raw map



X



Y

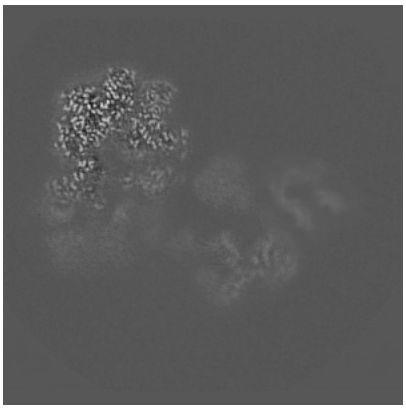


Z

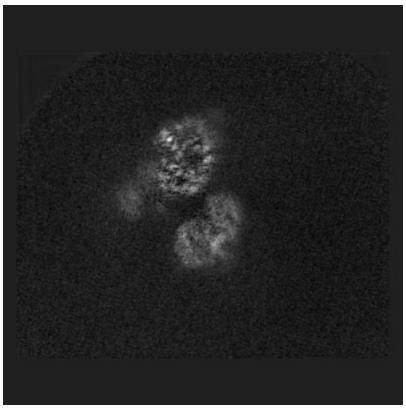
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

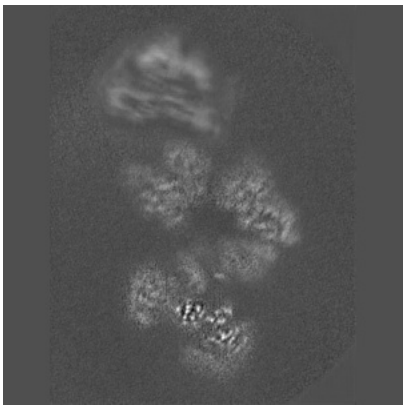
6.2.1 Primary map



X Index: 180

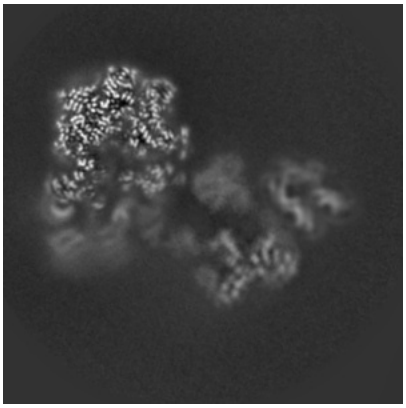


Y Index: 180

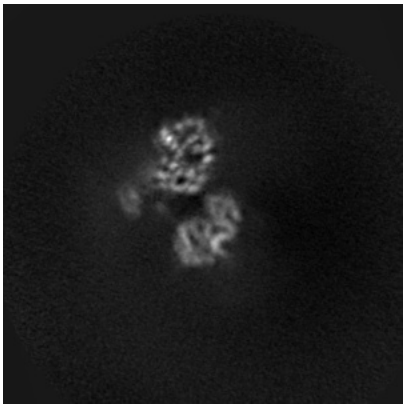


Z Index: 180

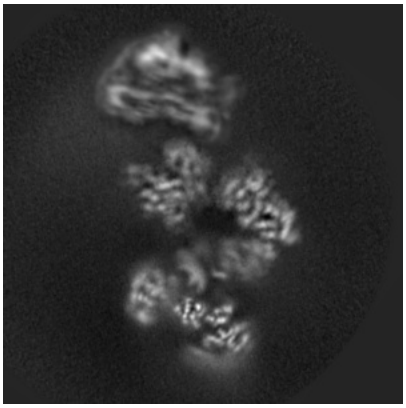
6.2.2 Raw map



X Index: 180



Y Index: 180

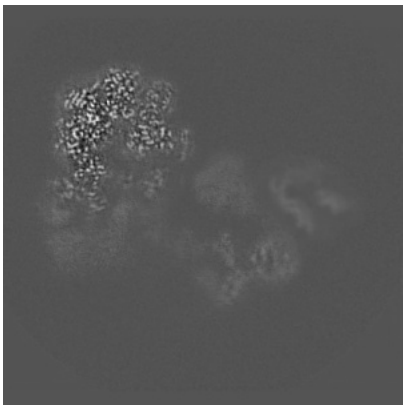


Z Index: 180

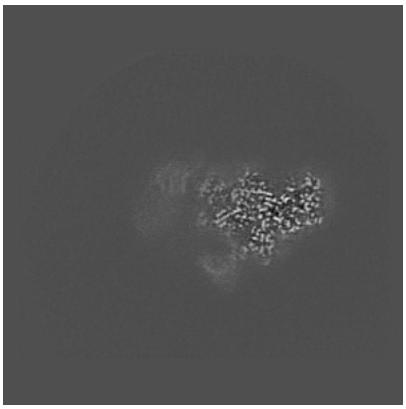
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

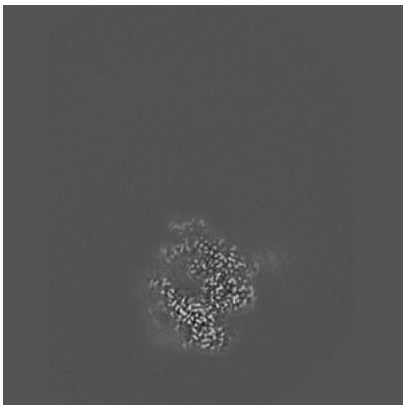
6.3.1 Primary map



X Index: 178

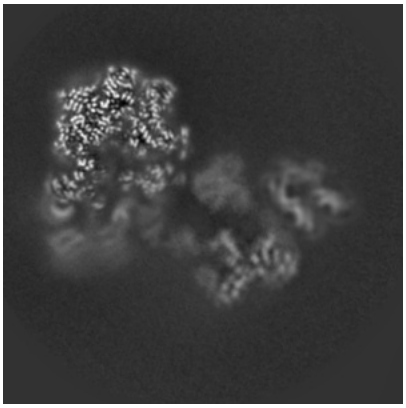


Y Index: 74

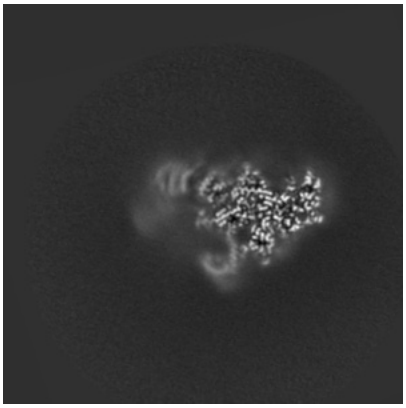


Z Index: 246

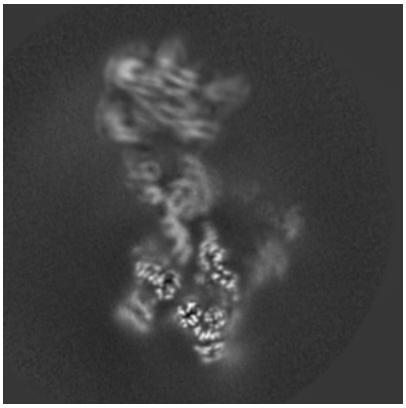
6.3.2 Raw map



X Index: 180



Y Index: 74

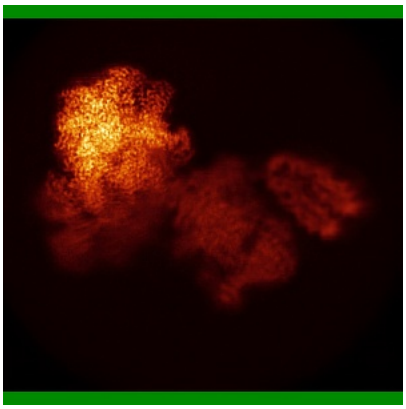


Z Index: 197

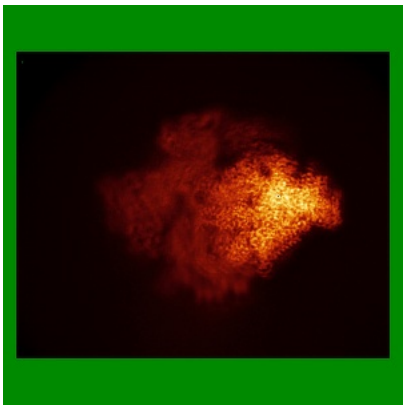
The images above show the largest variance slices of the map in three orthogonal directions.

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

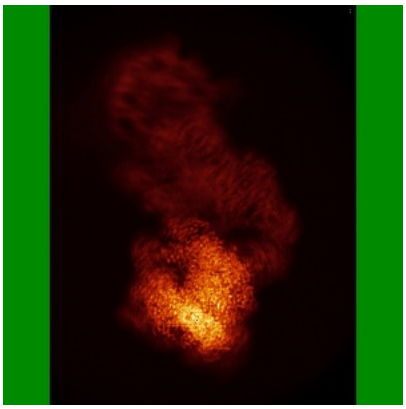
6.4.1 Primary map



X

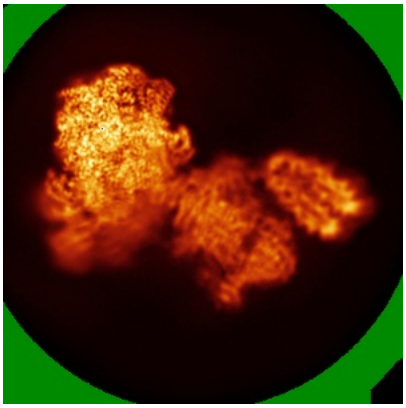


Y

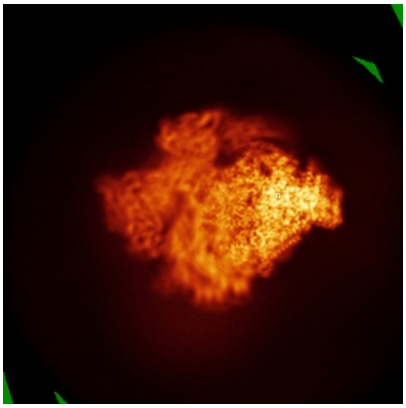


Z

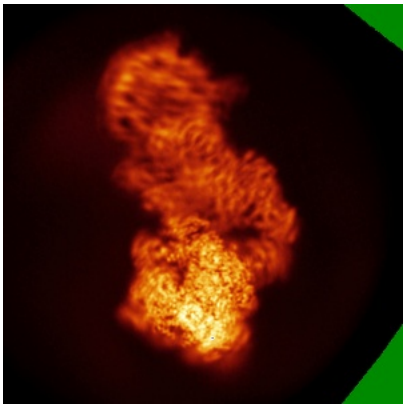
6.4.2 Raw map



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 0.007. 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.

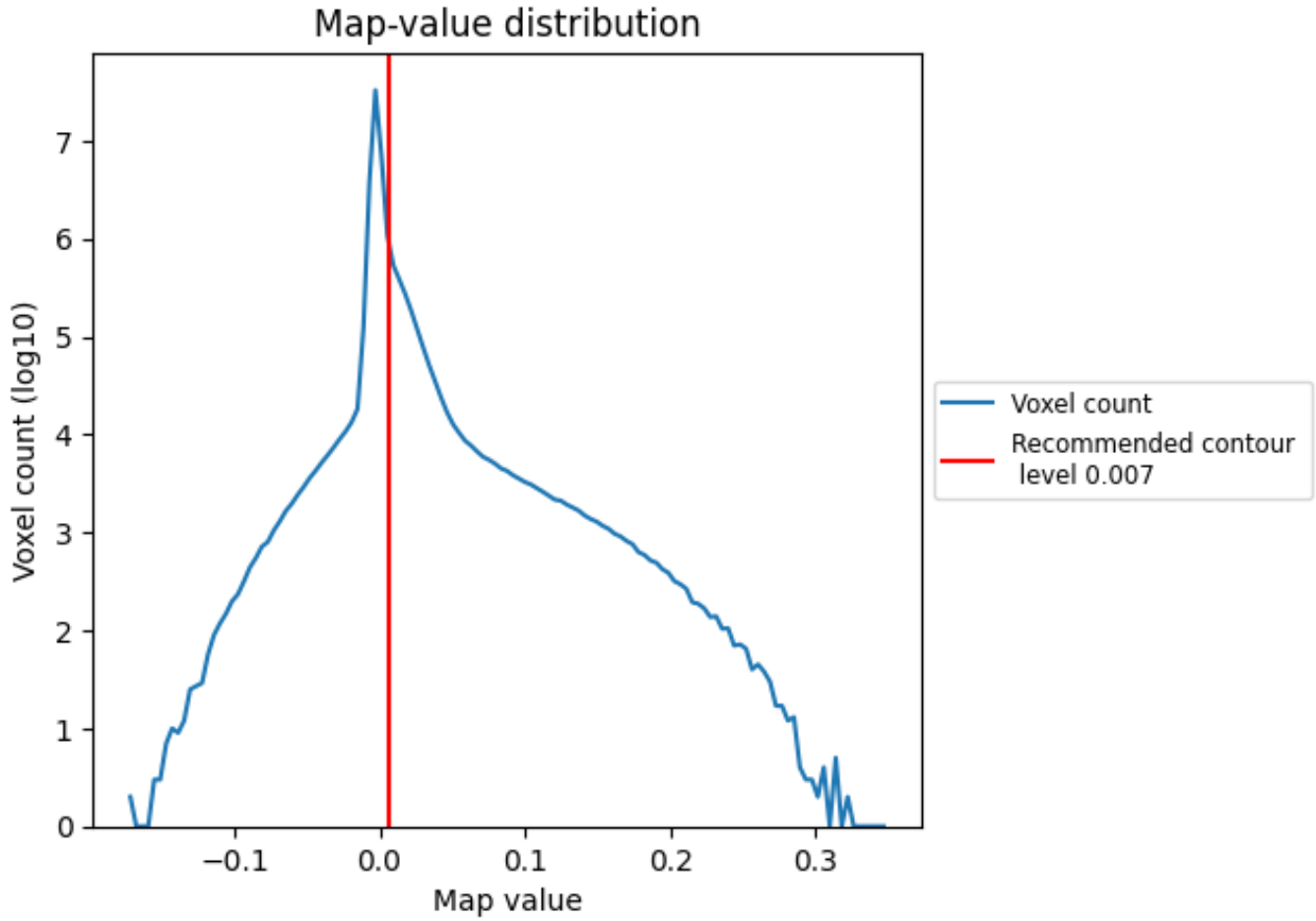
6.6 Mask visualisation (i)

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

7 Map analysis

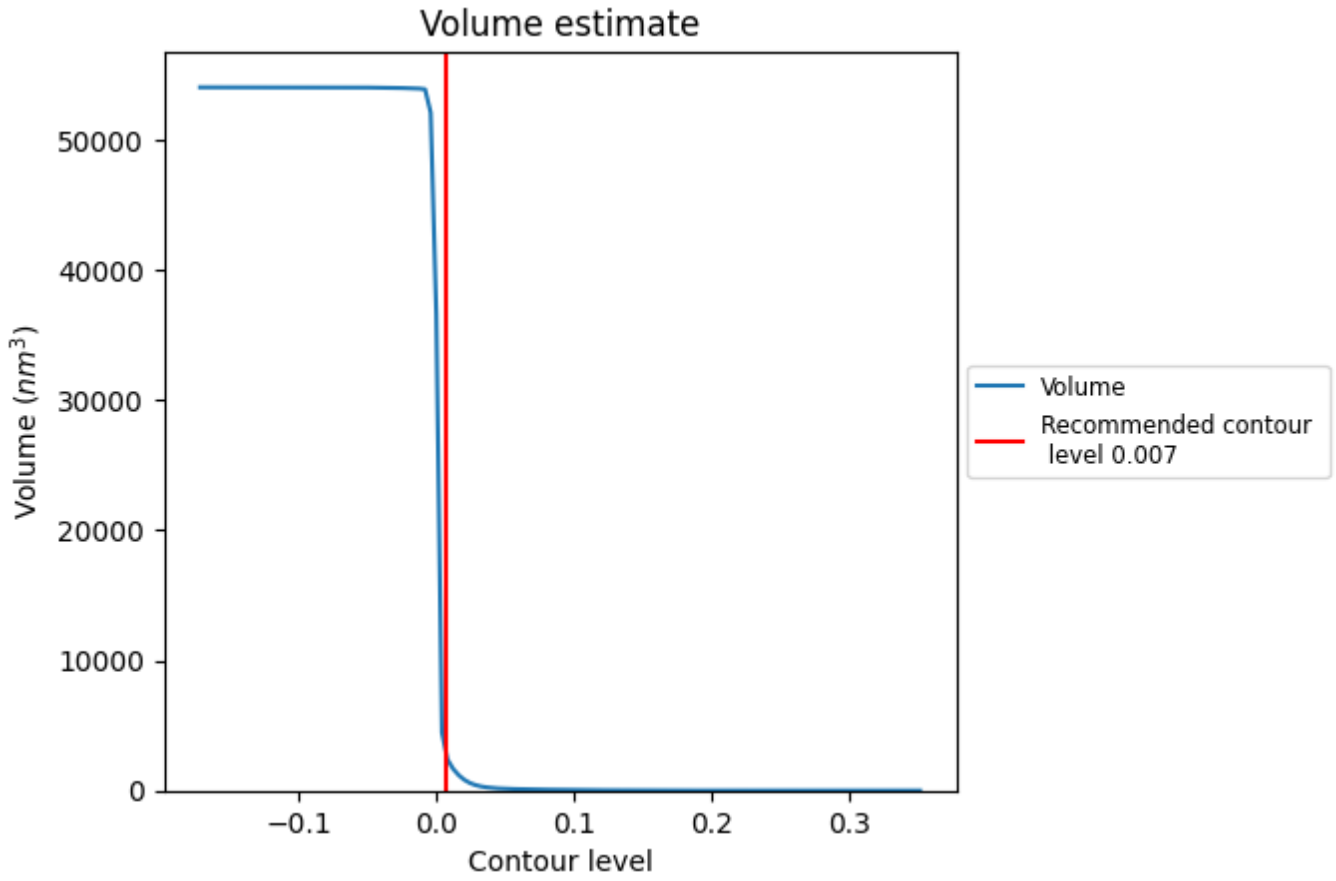
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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

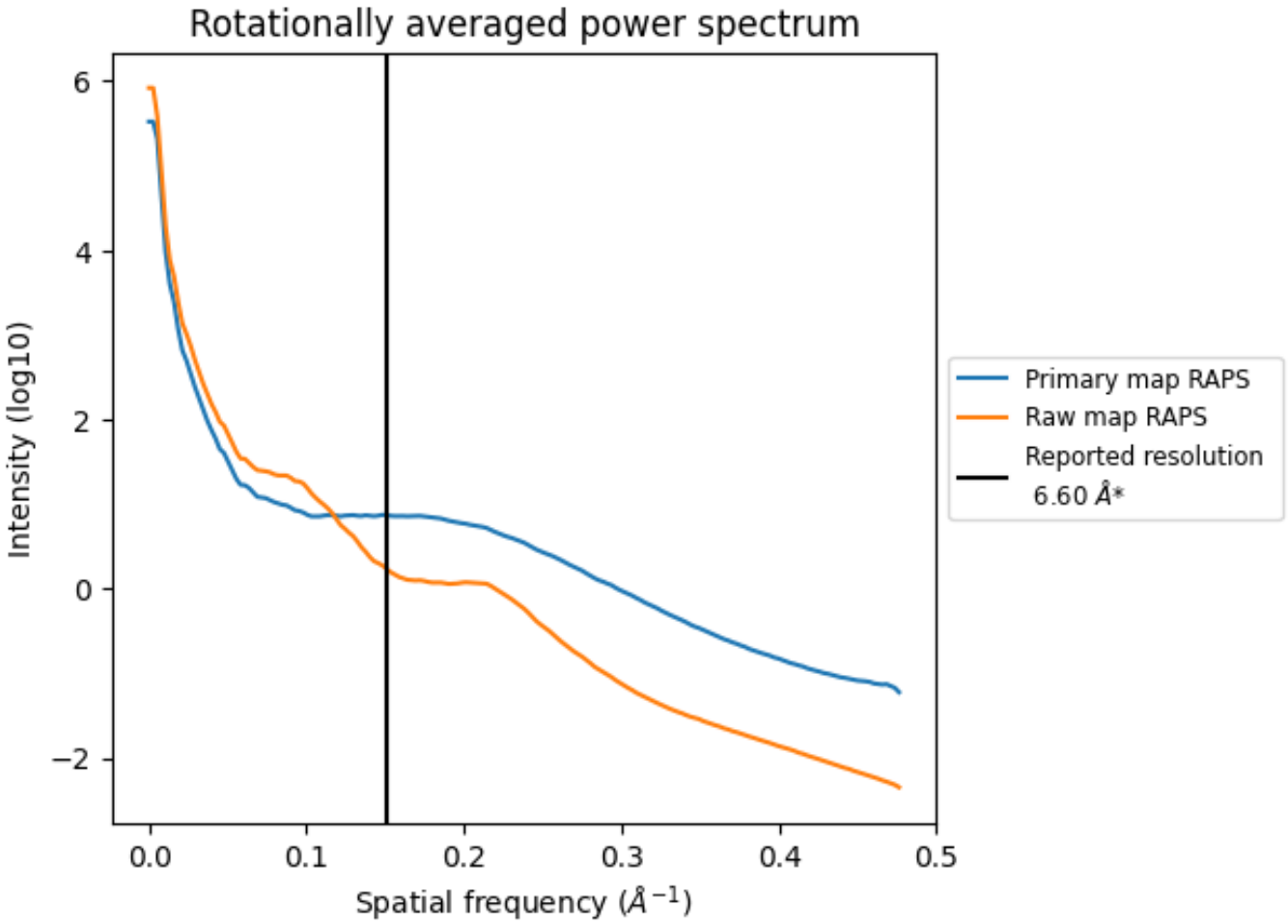
7.2 Volume estimate



The volume at the recommended contour level is 2948 nm^3; this corresponds to an approximate mass of 2663 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)

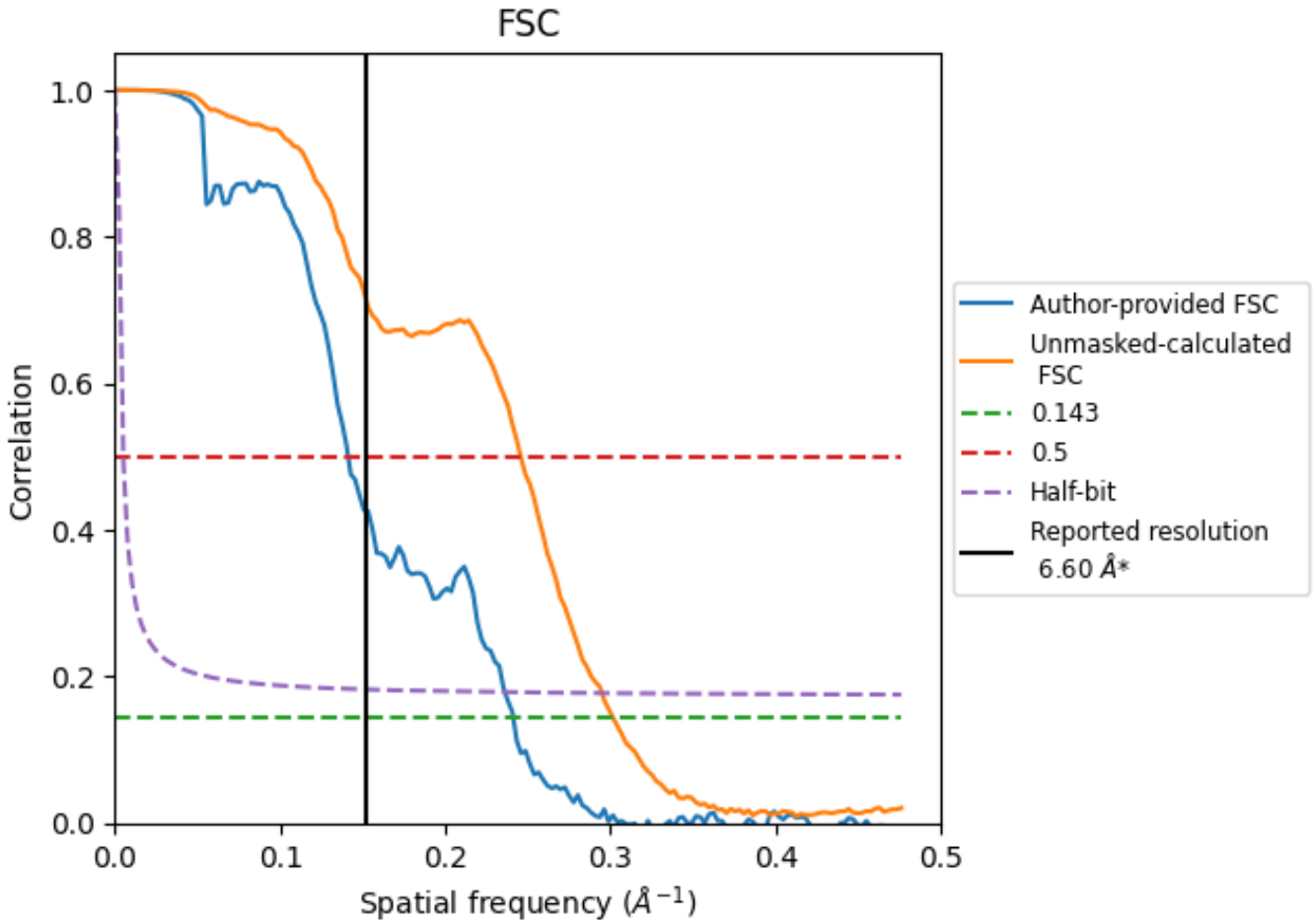


*Reported resolution corresponds to spatial frequency of 0.152 $Å^{-1}$

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 0.152 Å^{-1}

8.2 Resolution estimates (i)

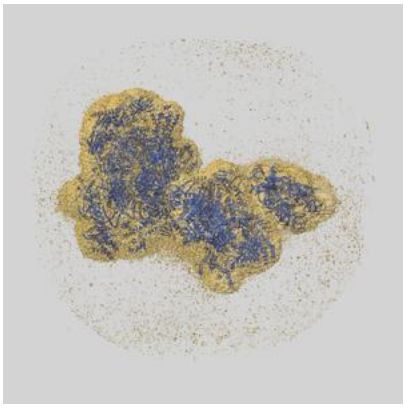
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	-	6.60	-
Author-provided FSC curve	4.14	7.08	4.23
Unmasked-calculated*	3.31	4.07	3.39

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.5 CUT-OFF 4.07 differs from the reported value 6.6 by more than 10 %

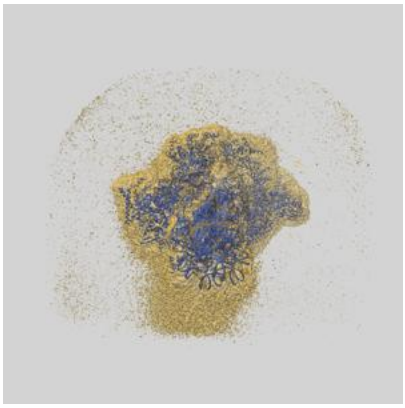
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-14929 and PDB model 7ZSB. Per-residue inclusion information can be found in section 3 on page 16.

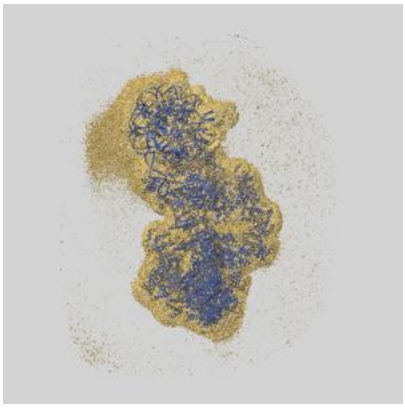
9.1 Map-model overlay



X



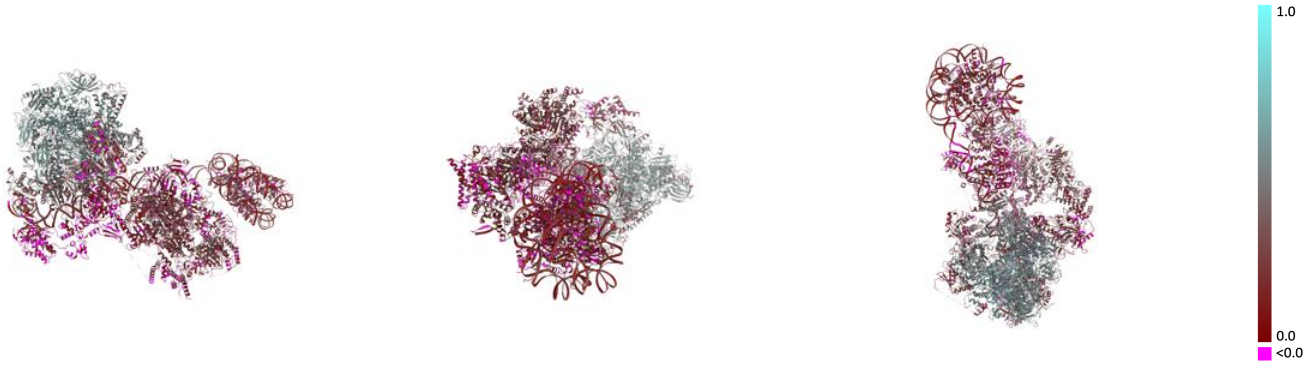
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.007 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.

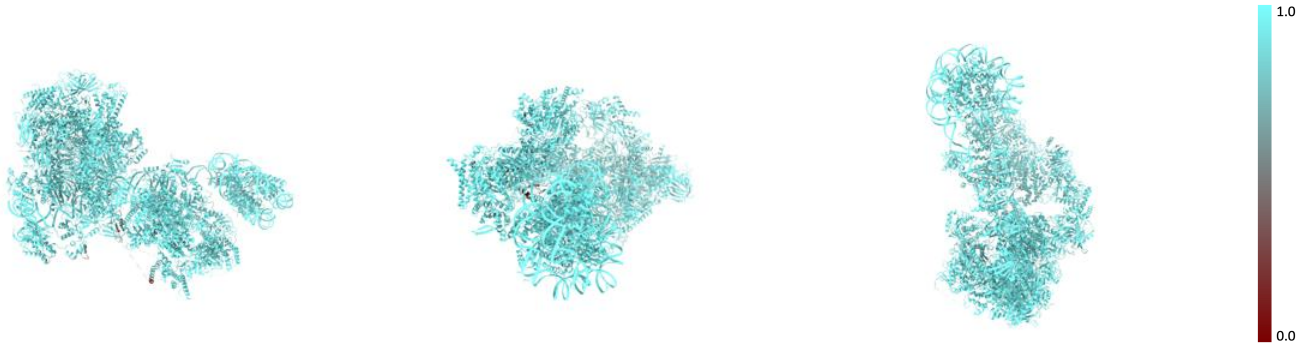
9.2 Q-score mapped to coordinate model (i)



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.

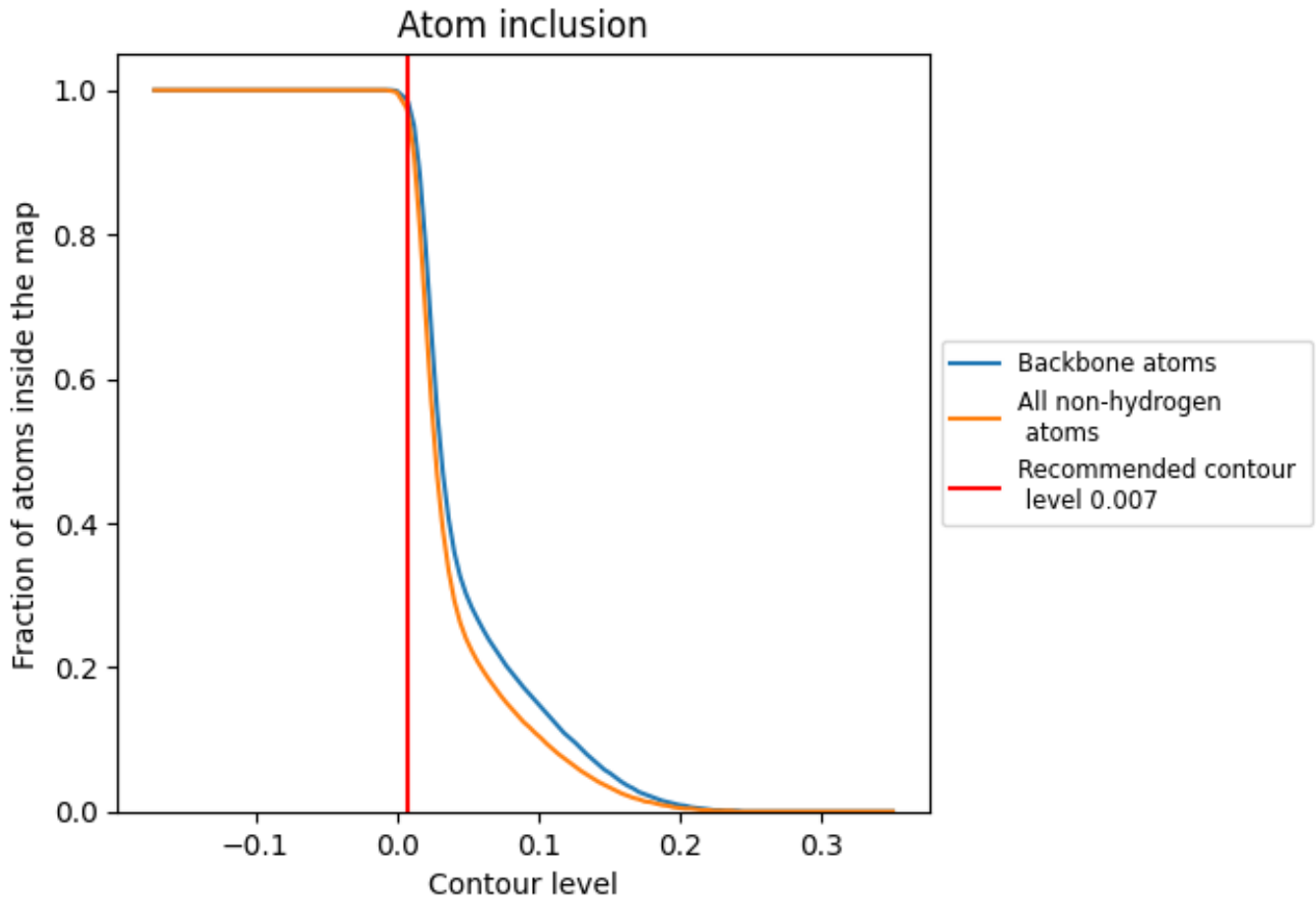
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.007).

9.4 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 98% 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 (0.007) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9750	0.2770
0	0.9900	0.2590
1	0.9490	0.1230
2	0.9770	0.1580
3	0.9970	0.1650
4	0.9890	0.2200
5	0.9940	0.1310
6	0.9760	0.1950
7	0.9760	0.1060
A	0.9840	0.4720
B	0.9820	0.5070
C	0.9890	0.5300
D	0.9560	0.1680
E	0.9870	0.4170
F	0.9560	0.4040
G	0.9650	0.2790
H	0.9910	0.4870
I	0.9860	0.3880
J	0.9910	0.5400
K	0.9790	0.5130
L	0.9830	0.4640
M	0.9750	0.3620
N	0.9920	0.1640
O	0.9910	0.1850
Q	0.9280	0.2070
R	0.9780	0.1830
T	0.9940	0.1580
U	0.9440	0.0420
V	0.9500	0.0520
W	0.8900	0.0740
X	0.9080	0.0550
a	0.9820	0.1370
b	0.9780	0.1540
c	0.9260	0.1880
d	0.9770	0.1650

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
e	0.9770	0.1800
f	0.9940	0.1840
g	0.9800	0.1730
h	0.9920	0.1670