wwPDB EM Validation Summary Report (i)

Jul 15, 2024 – 02:22 pm BST

PDB ID	:	8BYQ
EMDB ID	:	EMD-16331
Title	:	RNA polymerase II pre-initiation complex with the proximal +1 nucleosome (PIC-Nuc10W)
Authors	:	Abril-Garrido, J.; Dienemann, C.; Grabbe, F.; Velychko, T.; Lidschreiber, M.; Wang, H.; Cramer, P.
Deposited on	:	2022-12-14
Resolution	:	4.10 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:

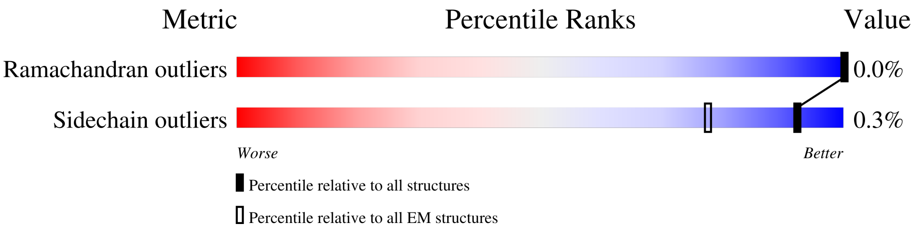
EMDB validation analysis	:	0.0.1.dev92
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 4.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
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	0	782	84% 16%
2	1	760	100%
3	2	548	74% • 26%
4	3	462	98% •
5	4	395	92% 7%
6	5	308	85% 15%
7	6	71	93% 7%
8	7	309	• 88% • 12%
9	8	346	• 84% • 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	9	323	14% 88% • 11%
11	A	1970	72% 28%
12	B	1174	97% •
13	C	275	93% 7%
14	D	142	90% 10%
15	E	210	99% •
16	F	127	62% 38%
17	G	172	98% ••
18	H	150	99% •
19	I	125	91% 9%
20	J	67	96% •
21	K	117	98% •
22	L	58	76% 24%
23	M	316	80% 20%
24	N	210	90% 5% 6%
25	O	339	53% 47%
26	Q	517	26% 73%
27	R	249	89% 11%
28	T	210	91% • 6%
29	U	376	30% 70%
30	V	109	89% • 9%
31	W	196	95% • 5%
32	X	291	59% 41%
33	a	136	71% 29%
33	e	136	72% 28%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	b	103	80% 20%
34	f	103	80% 20%
35	c	130	84% 16%
35	g	130	82% 18%
36	d	126	77% 23%
36	h	126	75% 25%

2 Entry composition

There are 39 unique types of molecules in this entry. The entry contains 85971 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 General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	0	653	5189	3303	899	957	30	0	0

- Molecule 2 is a protein called TFIIH basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	1	760	6034	3853	1053	1099	29	0	0

- Molecule 3 is a protein called General transcription factor IIH subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	408	2643	1646	486	503	8	0	0

- Molecule 4 is a protein called General transcription factor IIH subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	3	452	3589	2307	631	637	14	0	0

- Molecule 5 is a protein called General transcription factor IIH subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	4	366	2710	1702	476	506	26	0	0

- Molecule 6 is a protein called General transcription factor IIH subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	263	2064	1322	344	379	19	0	0

- Molecule 7 is a protein called General transcription factor IIH subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	6	66	Total 522	C 336	N 83	O 100	S 3	0	0

- Molecule 8 is a protein called CDK-activating kinase assembly factor MAT1.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	7	273	Total 2233	C 1397	N 391	O 432	S 13	0	0

- Molecule 9 is a protein called Cyclin-dependent kinase 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	8	298	Total 2370	C 1531	N 404	O 424	S 11	0	0

- Molecule 10 is a protein called Cyclin-H.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	9	287	Total 2334	C 1493	N 402	O 422	S 17	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
11	A	1423	Total 11274	C 7092	N 2016	O 2094	S 72	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	TYR	deletion	UNP A0A7M4DUC2
A	?	-	SER	deletion	UNP A0A7M4DUC2
A	?	-	PRO	deletion	UNP A0A7M4DUC2
A	?	-	THR	deletion	UNP A0A7M4DUC2
A	?	-	SER	deletion	UNP A0A7M4DUC2
A	?	-	PRO	deletion	UNP A0A7M4DUC2
A	?	-	SER	deletion	UNP A0A7M4DUC2
A	?	-	TYR	deletion	UNP A0A7M4DUC2
A	?	-	SER	deletion	UNP A0A7M4DUC2
A	?	-	PRO	deletion	UNP A0A7M4DUC2
A	?	-	THR	deletion	UNP A0A7M4DUC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP A0A7M4DUC2
A	?	-	PRO	deletion	UNP A0A7M4DUC2
A	?	-	SER	deletion	UNP A0A7M4DUC2

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

Mol	Chain	Residues	Atoms					AltConf	Trace
12	B	1136	Total 9076	C 5739	N 1597	O 1676	S 64	0	0

- Molecule 13 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	C	257	Total 2059	C 1294	N 351	O 408	S 6	0	0

- Molecule 14 is a protein called RNA polymerase II subunit D.

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

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

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

- Molecule 16 is a protein called DNA-directed RNA polymerase II subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	F	79	Total 636	C 406	N 108	O 117	S 5	0	0

- Molecule 17 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	G	171	Total 1351	C 875	N 219	O 249	S 8	0	0

- Molecule 18 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	H	148	1186	750	194	237	5	0	0

- Molecule 19 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 20 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 21 is a protein called DNA-directed RNA polymerase II subunit RPB11-a.

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

- Molecule 22 is a protein called RNA polymerase II subunit K.

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

- Molecule 23 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	M	252	1953	1224	346	366	17	0	0

- Molecule 24 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
24	N	198	4051	1921	746	1187	197	0	0

- Molecule 25 is a protein called TATA-box-binding protein.

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

- Molecule 26 is a protein called General transcription factor IIF subunit 1.

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

- Molecule 27 is a protein called General transcription factor IIF subunit 2.

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

- Molecule 28 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	T	198	Total 4061	C 1924	N 755	O 1185	P 197	0	0

- Molecule 29 is a protein called Transcription initiation factor IIA subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	U	113	Total 930	C 585	N 152	O 189	S 4	0	0

- Molecule 30 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	V	99	Total 806	C 510	N 142	O 151	S 3	0	0

- Molecule 31 is a protein called General transcription factor IIE subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	W	187	Total 1535	C 964	N 275	O 285	S 11	0	0

- Molecule 32 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	X	171	Total 1403	C 895	N 243	O 261	S 4	0	0

- Molecule 33 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	a	97	Total 802	C 506	N 155	O 138	S 3	0	0
33	e	98	Total 811	C 512	N 157	O 139	S 3	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
a	103	ALA	GLY	engineered mutation	UNP P84233
e	103	ALA	GLY	engineered mutation	UNP P84233

- Molecule 34 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	b	82	Total 653	C 412	N 127	O 113	S 1	0	0
34	f	82	Total 653	C 412	N 127	O 113	S 1	0	0

- Molecule 35 is a protein called Histone H2A type 1.

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

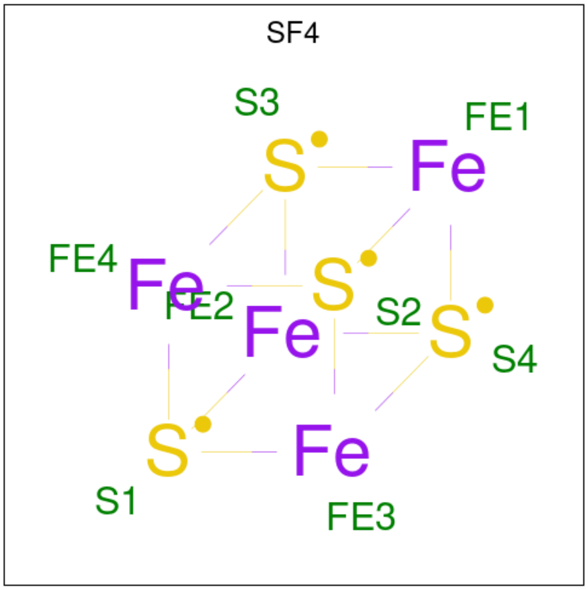
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
c	100	ARG	GLY	conflict	UNP P06897
g	100	ARG	GLY	conflict	UNP P06897

- Molecule 36 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	d	97	Total 766	C 480	N 142	O 142	S 2	0	0
36	h	95	Total 744	C 468	N 134	O 140	S 2	0	0

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



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

- Molecule 38 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
38	4	3	Total 3	Zn 3	0
38	5	1	Total 1	Zn 1	0
38	7	2	Total 2	Zn 2	0
38	A	2	Total 2	Zn 2	0
38	B	1	Total 1	Zn 1	0
38	C	1	Total 1	Zn 1	0
38	I	2	Total 2	Zn 2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		AltConf
38	J	1	Total 1	Zn 1	0
38	L	1	Total 1	Zn 1	0
38	M	1	Total 1	Zn 1	0
38	W	1	Total 1	Zn 1	0

- Molecule 39 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

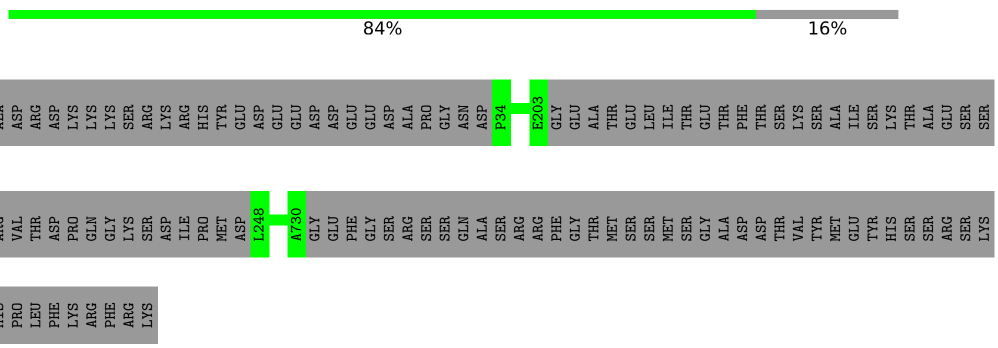
Mol	Chain	Residues	Atoms		AltConf
39	A	1	Total 1	Mg 1	0

3 Residue-property plots

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

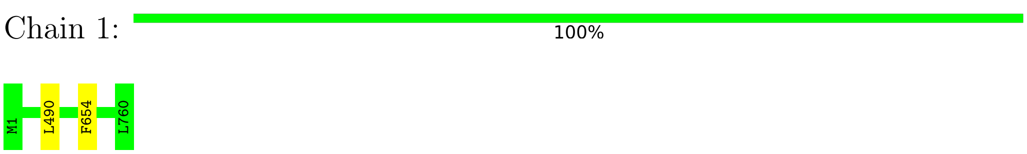
- Molecule 1: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 0:



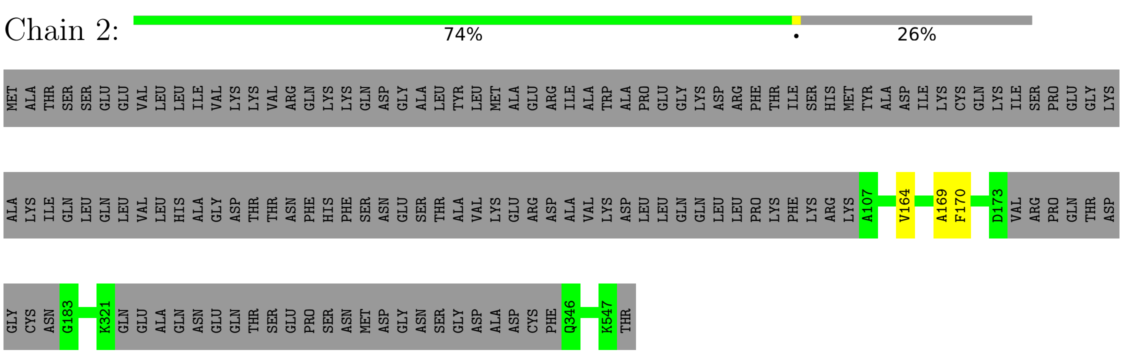
- Molecule 2: TFIIH basal transcription factor complex helicase XPD subunit

Chain 1:



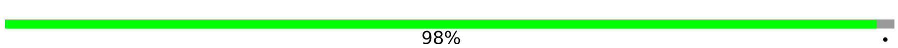
- Molecule 3: General transcription factor IIH subunit 1

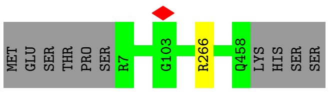
Chain 2:



- Molecule 4: General transcription factor IIH subunit 4

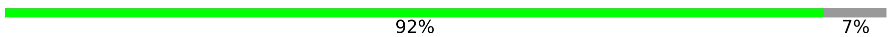
Chain 3:

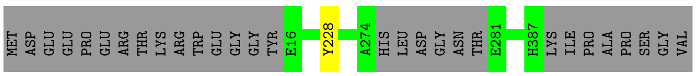




• Molecule 5: General transcription factor IIH subunit 2

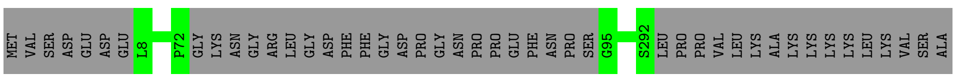
Chain 4:





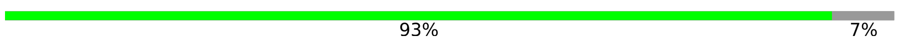
• Molecule 6: General transcription factor IIH subunit 3

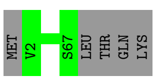
Chain 5:



• Molecule 7: General transcription factor IIH subunit 5

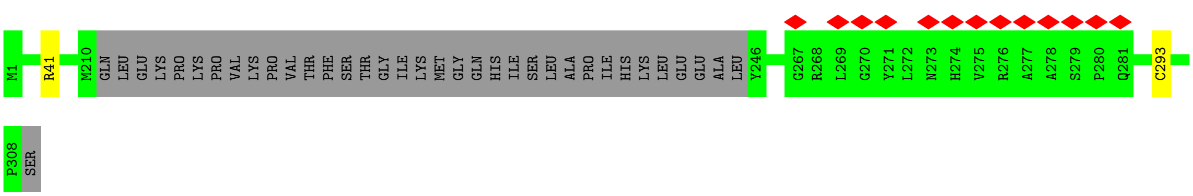
Chain 6:





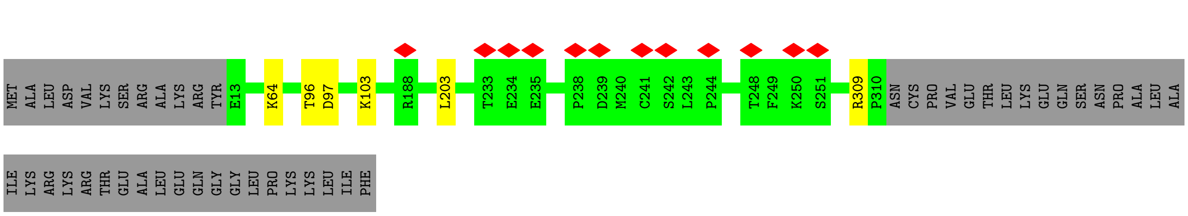
• Molecule 8: CDK-activating kinase assembly factor MAT1

Chain 7:



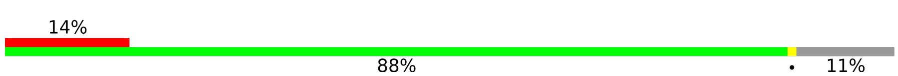
• Molecule 9: Cyclin-dependent kinase 7

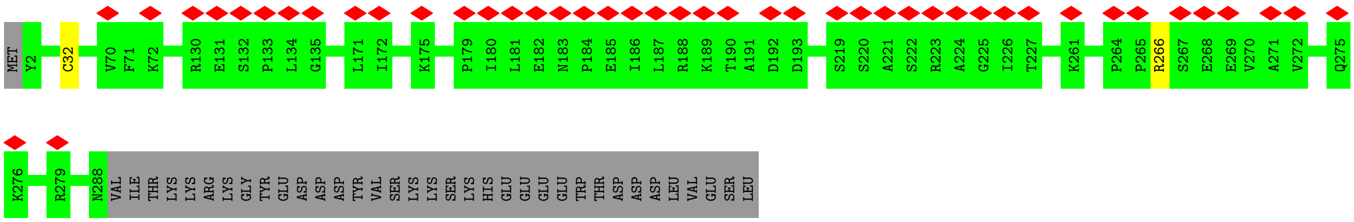
Chain 8:



• Molecule 10: Cyclin-H

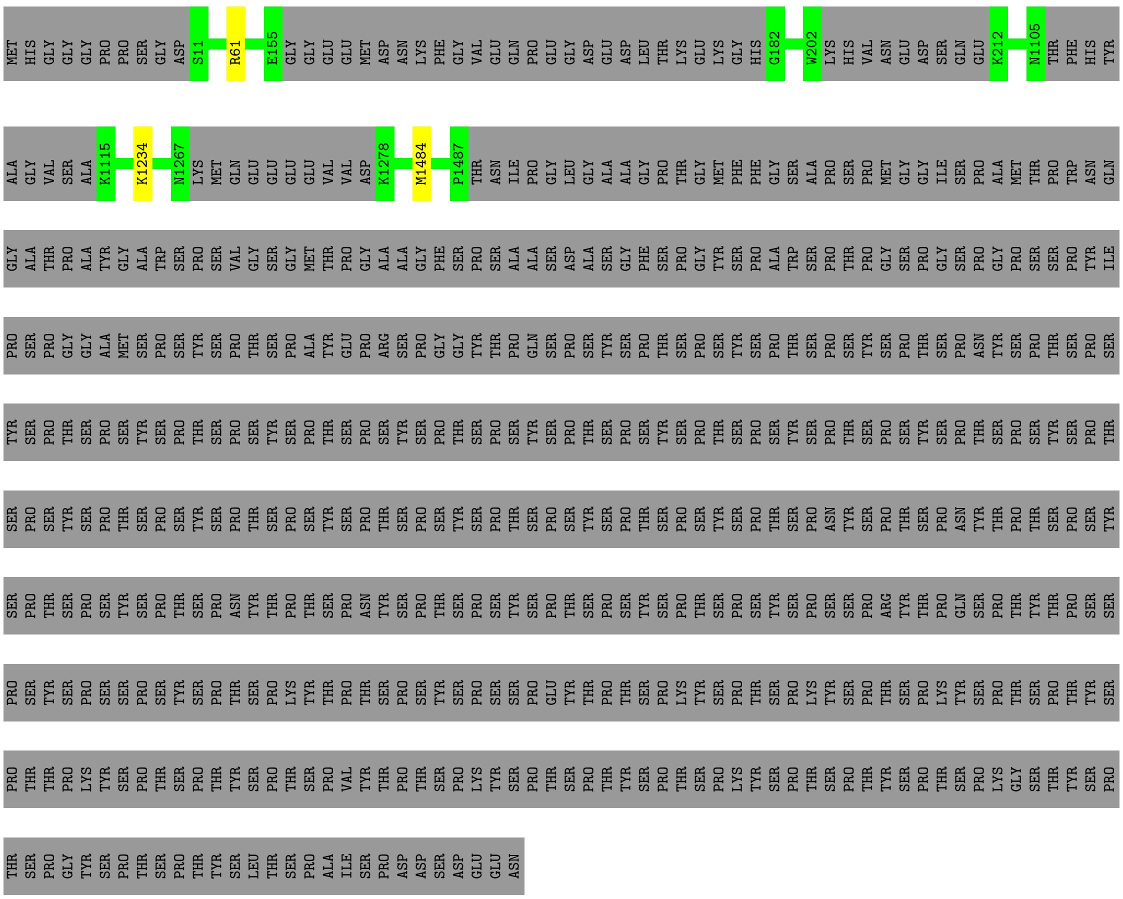
Chain 9:





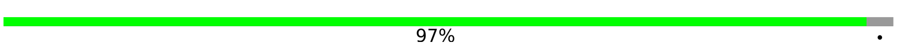
- Molecule 11: DNA-directed RNA polymerase subunit

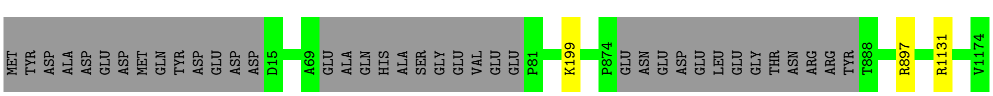
Chain A:



- Molecule 12: DNA-directed RNA polymerase subunit beta

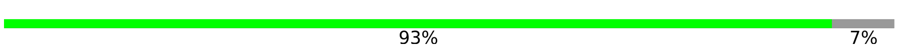
Chain B:





- Molecule 13: DNA-directed RNA polymerase II subunit RPB3

Chain C:



MET P2 S132 ARG ASN ARG ASP ASN ASP PRO SER ASP TYR VAL GLU GLN D146 V271 LEU THR ILE ASN

- Molecule 14: RNA polymerase II subunit D

Chain D: 90% 10%

MET ALA ALA GLY GLY SER ASP PRO ARG ALA GLY ASP VAL E14 Q141 TYR

- Molecule 15: DNA-directed RNA polymerase II subunit E

Chain E: 99%

MET D2 R52 R162 Q210

- Molecule 16: DNA-directed RNA polymerase II subunit F

Chain F: 62% 38%

MET SER ASP ASN GLU ASP ASN PHE ASP GLY ASP ASP PHE ASP ASP VAL GLU GLU ASP GLU GLY LEU ASP ASP LEU GLU ASN ALA GLU GLU GLU GLY GLN GLU ASN VAL GLU ILE LEU PRO SER GLY GLU ARG PRO GLN ALA ASN Q49 D127

- Molecule 17: DNA-directed RNA polymerase II subunit RPB7

Chain G: 98%

M1 D52 K81 V171 SER

- Molecule 18: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 99%

MET A2 A149 PHE

- Molecule 19: DNA-directed RNA polymerase II subunit RPB9

Chain I: 91% 9%

MET GLU PRO ASP GLY THR TYR GLU PRO GLY PHE V12 E125

- Molecule 20: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J: 96%

M1 P64 LEU GLU LYS

• Molecule 21: DNA-directed RNA polymerase II subunit RPB11-a

Chain K: 98% •

M1 G115 ILE GLU

• Molecule 22: RNA polymerase II subunit K

Chain L: 76% 24%

MET ASP THR GLN LYS ASP VAL GLN PRO PRO LYS GLN GLN PRO M15 R58

• Molecule 23: Transcription initiation factor IIB

Chain M: 80% 20%

MET ALA SER THR SER ARG LEU ASP ALA LEU PRO ARG VAL T14 T54 PHE SER ASN ASP LYS ALA THR LYS ASP PRO SER ARG VAL GLY ASP SER GLN ASN PRO LEU LEU SER ASP GLY ASP LEU SER THR MET ILE GLY LYS GLY THR GLY ALA ALA SER PHE ASP GLU PHE GLY ASN

SER LYS TYR GLN ASN ARG ARG T106 L316

• Molecule 24: Non-template DNA

Chain N: 90% 5% 6%

DC DG DG DG DT DG DT DT DC DC DT G-9 G11 T38 G48 C54 G58 G82 C89 C137 C165 C182 A188 DT

• Molecule 25: TATA-box-binding protein

Chain O: 53% 47%

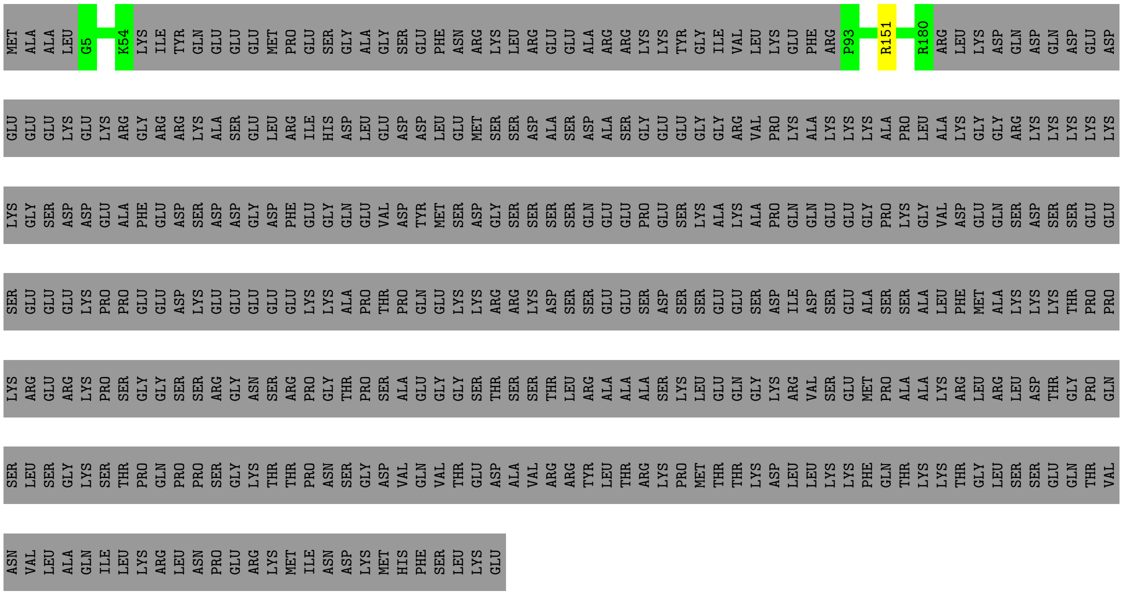
MET ASP GLN ASN ASN SER LEU PRO PRO TYR ALA GLN GLY LEU ALA SER PRO GLN GLY ALA MET THR PRO GLY ILE PRO ILE PHE SER PRO MET MET PRO TYR GLY THR GLY LEU THR PRO GLN PRO ILE GLN ASN THR ASN SER LEU SER ILE LEU GLU GLU GLN GLN ARG GLN GLN

GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN GLN ALA VAL ALA ALA ALA ALA VAL GLN GLN SER THR SER GLN GLN ALA THR GLN GLY THR SER GLY GLN ALA PRO GLN

LEU PHE HIS SER GLN THR LEU THR THR ALA PRO LEU PRO GLY THR THR PRO LEU TYR PRO SER PRO MET THR PRO MET THR PRO ILE THR PRO ALA THR PRO ALA SER GLU SER S159 K337 THR THR

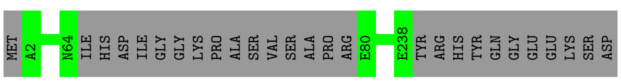
• Molecule 26: General transcription factor IIF subunit 1

Chain Q: 26% 73%



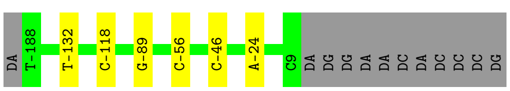
• Molecule 27: General transcription factor IIF subunit 2

Chain R:

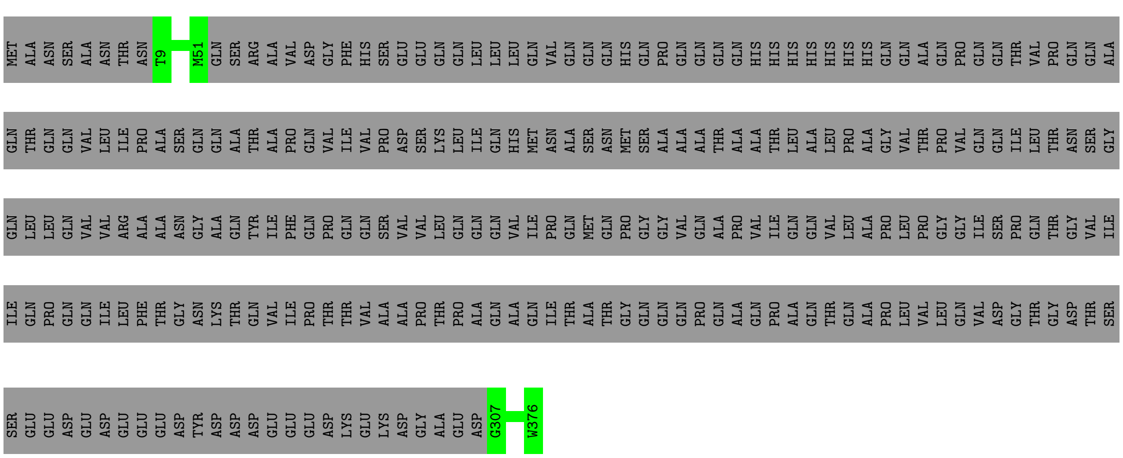
• Molecule 28: Template DNA

Chain T:

• Molecule 29: Transcription initiation factor IIA subunit 1

Chain U:

• Molecule 30: Transcription initiation factor IIA subunit 2

Chain V: 89% • 9%

M1 R51 R82 D99 GLY LYS ASN THR GLY SER ASN THR THR GLU

• Molecule 31: General transcription factor IIE subunit 1

Chain W: 95% • 5%

MET ALA ASP PRO ASP VAL LEU THR GLU V10 R56 D196

• Molecule 32: Transcription initiation factor IIE subunit beta

Chain X: 59% 41%

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

SER PHE ASN LEU LYS ALA LEU SER GLY SER SER G72 I242 SER SER MET GLN GLU SER GLY PRO LYS LYS VAL ALA PRO ILE GLN ARG ARG LYS LYS PRO ALA SER GLN LYS LYS ARG ARG PHE LYS THR HIS ASN GLU HIS LEU ALA GLY VAL LEU LYS ASP TYR SER ASP ILE THR

SER SER LYS

• Molecule 33: Histone H3.2

Chain a: 71% 29%

MET ALA ARG THR LYS GLN THR ALA ARG LYS SER THR GLY GLY LYS ALA PRO ARG LYS GLN LEU ALA THR LYS ALA ALA ARG LYS SER ALA PRO ALA THR GLY GLY VAL LYS LYS P39 R135 ALA

• Molecule 33: Histone H3.2

Chain e: 72% 28%

MET ALA ARG THR LYS GLN THR ALA ARG LYS SER THR GLY GLY LYS ALA PRO ARG LYS GLN LEU ALA THR LYS ALA ALA ARG LYS SER ALA PRO ALA THR GLY GLY VAL LYS K38 R135 ALA

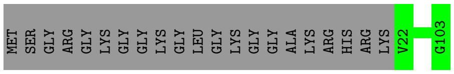
• Molecule 34: Histone H4

Chain b: 80% 20%

MET SER GLY ARG GLY LYS GLY GLY LYS GLY LEU GLY LYS GLY GLY ALA LYS ARG HIS ARG LYS V22 G103

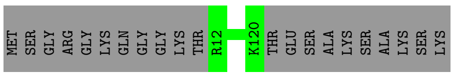
• Molecule 34: Histone H4

Chain f: 80% 20%



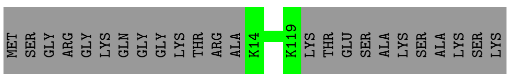
- Molecule 35: Histone H2A type 1

Chain c:

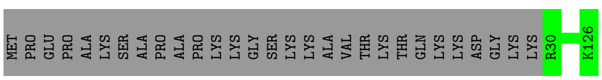
- Molecule 35: Histone H2A type 1

Chain g:

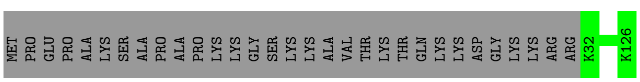
- Molecule 36: Histone H2B 1.1

Chain d:

- Molecule 36: Histone H2B 1.1

Chain h:

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	101994	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	50.45	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	1300	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	21.074	Depositor
Minimum map value	-5.711	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	0.4	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: SF4, ZN, MG

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	0	0.26	0/5294	0.53	0/7158
2	1	0.30	0/6158	0.59	1/8342 (0.0%)
3	2	0.30	0/2683	0.46	0/3672
4	3	0.26	0/3672	0.54	0/4979
5	4	0.29	0/2764	0.56	0/3754
6	5	0.28	0/2101	0.55	0/2843
7	6	0.26	0/528	0.49	0/713
8	7	0.27	0/2269	0.53	0/3053
9	8	0.26	0/2429	0.56	0/3295
10	9	0.26	0/2384	0.50	0/3220
11	A	0.38	0/11479	0.59	0/15496
12	B	0.45	0/9257	0.62	0/12493
13	C	0.43	0/2102	0.60	0/2857
14	D	0.26	0/1064	0.49	0/1428
15	E	0.33	0/1752	0.56	0/2366
16	F	0.39	0/646	0.61	0/871
17	G	0.31	0/1382	0.56	0/1874
18	H	0.39	0/1207	0.60	0/1628
19	I	0.32	0/949	0.56	0/1284
20	J	0.50	0/516	0.66	0/696
21	K	0.39	0/939	0.55	0/1271
22	L	0.44	0/378	0.74	0/500
23	M	0.32	0/1983	0.57	0/2679
24	N	0.74	0/4543	1.08	10/7009 (0.1%)
25	O	0.31	0/1448	0.56	0/1948
26	Q	0.27	0/1167	0.54	0/1576
27	R	0.28	0/1817	0.51	0/2445
28	T	0.71	0/4557	1.03	6/7033 (0.1%)
29	U	0.26	0/945	0.51	0/1274
30	V	0.28	0/816	0.54	0/1105
31	W	0.26	0/1560	0.56	0/2097
32	X	0.26	0/1427	0.49	0/1916

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
33	a	0.25	0/814	0.55	0/1092
33	e	0.26	0/823	0.54	0/1104
34	b	0.25	0/660	0.57	0/883
34	f	0.25	0/660	0.59	0/883
35	c	0.24	0/853	0.52	0/1149
35	g	0.24	0/828	0.55	0/1117
36	d	0.26	0/777	0.52	0/1041
36	h	0.27	0/755	0.46	0/1013
All	All	0.39	0/88386	0.64	17/121157 (0.0%)

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
24	N	82	DG	O4'-C1'-N9	7.46	113.22	108.00
2	1	490	LEU	CA-CB-CG	7.34	132.18	115.30
24	N	38	DT	O4'-C1'-N1	6.73	112.71	108.00
24	N	48	DG	O4'-C1'-N9	6.15	112.31	108.00
24	N	58	DG	O4'-C1'-N9	6.09	112.27	108.00

There are no chirality outliers.

There are no planarity outliers.

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	0	649/782 (83%)	629 (97%)	20 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	1	758/760 (100%)	736 (97%)	22 (3%)	0	100	100
3	2	402/548 (73%)	352 (88%)	47 (12%)	3 (1%)	22	60
4	3	450/462 (97%)	437 (97%)	13 (3%)	0	100	100
5	4	362/395 (92%)	340 (94%)	22 (6%)	0	100	100
6	5	259/308 (84%)	244 (94%)	15 (6%)	0	100	100
7	6	64/71 (90%)	64 (100%)	0	0	100	100
8	7	269/309 (87%)	264 (98%)	5 (2%)	0	100	100
9	8	296/346 (86%)	281 (95%)	15 (5%)	0	100	100
10	9	285/323 (88%)	279 (98%)	6 (2%)	0	100	100
11	A	1413/1970 (72%)	1376 (97%)	37 (3%)	0	100	100
12	B	1130/1174 (96%)	1099 (97%)	31 (3%)	0	100	100
13	C	253/275 (92%)	248 (98%)	5 (2%)	0	100	100
14	D	126/142 (89%)	125 (99%)	1 (1%)	0	100	100
15	E	207/210 (99%)	205 (99%)	2 (1%)	0	100	100
16	F	77/127 (61%)	76 (99%)	1 (1%)	0	100	100
17	G	169/172 (98%)	166 (98%)	3 (2%)	0	100	100
18	H	146/150 (97%)	142 (97%)	4 (3%)	0	100	100
19	I	112/125 (90%)	107 (96%)	5 (4%)	0	100	100
20	J	62/67 (92%)	61 (98%)	1 (2%)	0	100	100
21	K	113/117 (97%)	110 (97%)	3 (3%)	0	100	100
22	L	42/58 (72%)	41 (98%)	1 (2%)	0	100	100
23	M	248/316 (78%)	242 (98%)	6 (2%)	0	100	100
25	O	177/339 (52%)	176 (99%)	1 (1%)	0	100	100
26	Q	134/517 (26%)	129 (96%)	5 (4%)	0	100	100
27	R	218/249 (88%)	215 (99%)	3 (1%)	0	100	100
29	U	109/376 (29%)	108 (99%)	1 (1%)	0	100	100
30	V	97/109 (89%)	94 (97%)	3 (3%)	0	100	100
31	W	185/196 (94%)	183 (99%)	2 (1%)	0	100	100
32	X	169/291 (58%)	162 (96%)	7 (4%)	0	100	100
33	a	95/136 (70%)	95 (100%)	0	0	100	100
33	e	96/136 (71%)	96 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
34	b	80/103 (78%)	80 (100%)	0	0	100	100
34	f	80/103 (78%)	80 (100%)	0	0	100	100
35	c	107/130 (82%)	107 (100%)	0	0	100	100
35	g	104/130 (80%)	103 (99%)	1 (1%)	0	100	100
36	d	95/126 (75%)	94 (99%)	1 (1%)	0	100	100
36	h	93/126 (74%)	93 (100%)	0	0	100	100
All	All	9731/12274 (79%)	9439 (97%)	289 (3%)	3 (0%)	100	100

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	2	164	VAL
3	2	170	PHE
3	2	169	ALA

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	0	558/688 (81%)	558 (100%)	0	100	100
2	1	639/664 (96%)	638 (100%)	1 (0%)	93	96
3	2	176/484 (36%)	176 (100%)	0	100	100
4	3	384/399 (96%)	383 (100%)	1 (0%)	92	95
5	4	280/352 (80%)	279 (100%)	1 (0%)	91	94
6	5	234/272 (86%)	234 (100%)	0	100	100
7	6	59/64 (92%)	59 (100%)	0	100	100
8	7	250/283 (88%)	248 (99%)	2 (1%)	81	88
9	8	258/299 (86%)	252 (98%)	6 (2%)	50	70
10	9	259/296 (88%)	257 (99%)	2 (1%)	81	88
11	A	1254/1749 (72%)	1251 (100%)	3 (0%)	93	96

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
12	B	994/1027 (97%)	991 (100%)	3 (0%)	92	95
13	C	234/252 (93%)	234 (100%)	0	100	100
14	D	118/126 (94%)	118 (100%)	0	100	100
15	E	191/192 (100%)	189 (99%)	2 (1%)	76	85
16	F	69/111 (62%)	69 (100%)	0	100	100
17	G	152/153 (99%)	150 (99%)	2 (1%)	69	81
18	H	129/131 (98%)	129 (100%)	0	100	100
19	I	103/112 (92%)	103 (100%)	0	100	100
20	J	53/56 (95%)	53 (100%)	0	100	100
21	K	104/106 (98%)	104 (100%)	0	100	100
22	L	41/55 (74%)	41 (100%)	0	100	100
23	M	215/268 (80%)	215 (100%)	0	100	100
25	O	154/293 (53%)	154 (100%)	0	100	100
26	Q	121/448 (27%)	120 (99%)	1 (1%)	81	88
27	R	196/218 (90%)	196 (100%)	0	100	100
29	U	105/324 (32%)	105 (100%)	0	100	100
30	V	90/98 (92%)	88 (98%)	2 (2%)	52	71
31	W	169/177 (96%)	168 (99%)	1 (1%)	86	92
32	X	154/261 (59%)	154 (100%)	0	100	100
33	a	85/111 (77%)	85 (100%)	0	100	100
33	e	86/111 (78%)	86 (100%)	0	100	100
34	b	67/79 (85%)	67 (100%)	0	100	100
34	f	67/79 (85%)	67 (100%)	0	100	100
35	c	86/101 (85%)	86 (100%)	0	100	100
35	g	84/101 (83%)	84 (100%)	0	100	100
36	d	83/106 (78%)	83 (100%)	0	100	100
36	h	81/106 (76%)	81 (100%)	0	100	100
All	All	8382/10752 (78%)	8355 (100%)	27 (0%)	92	95

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	A	1234	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
12	B	897	ARG
30	V	51	ARG
12	B	199	LYS
12	B	1131	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
11	A	1394	ASN
18	H	133	HIS
26	Q	53	ASN
19	I	100	HIS
9	8	258	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 18 ligands modelled in this entry, 17 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	1	1000	2	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	1	1000	2	-	-	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.

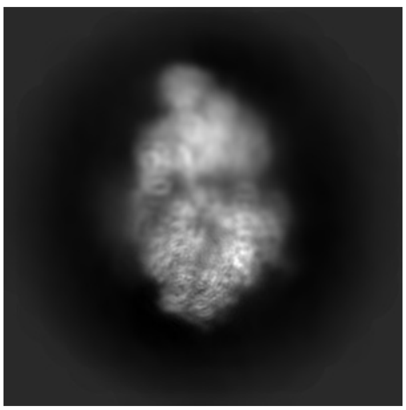
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-16331. These allow visual inspection of the internal detail of the map and identification of artifacts.

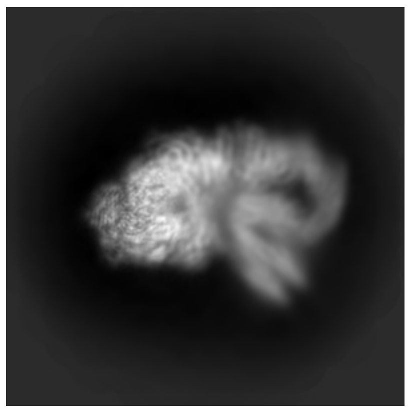
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

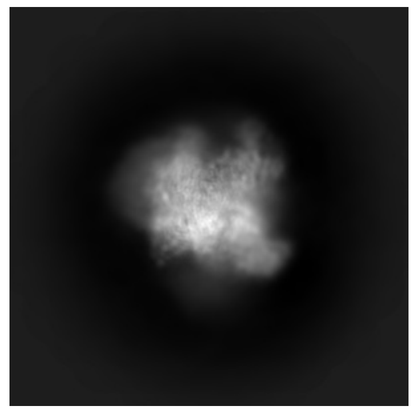
6.1.1 Primary map



X

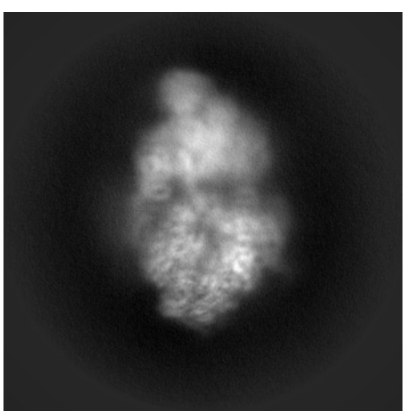


Y

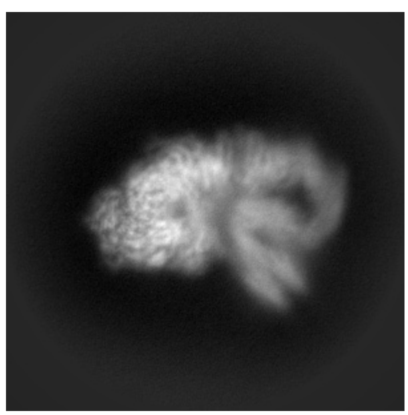


Z

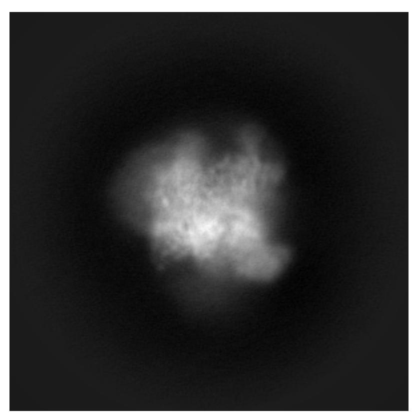
6.1.2 Raw map



X



Y

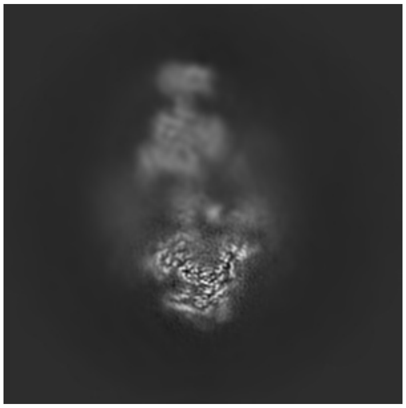


Z

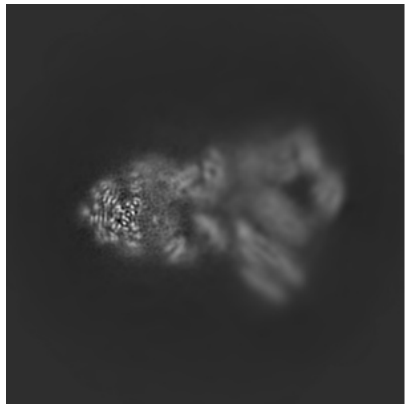
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

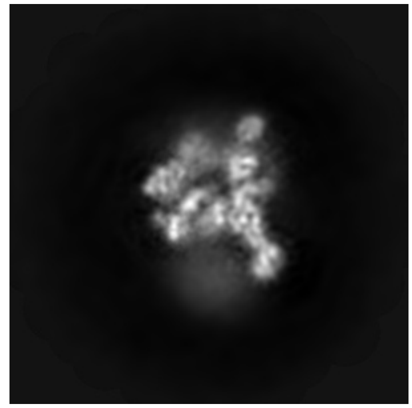
6.2.1 Primary map



X Index: 200

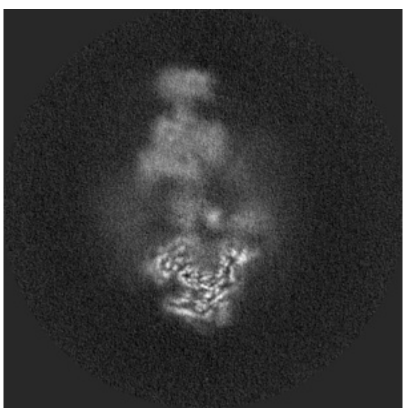


Y Index: 200

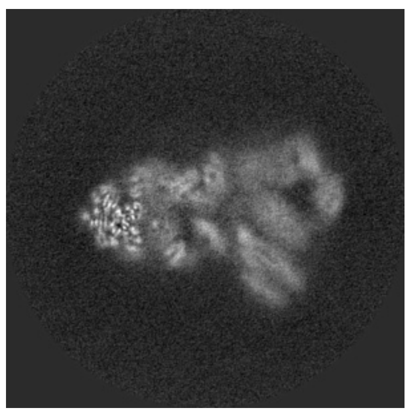


Z Index: 200

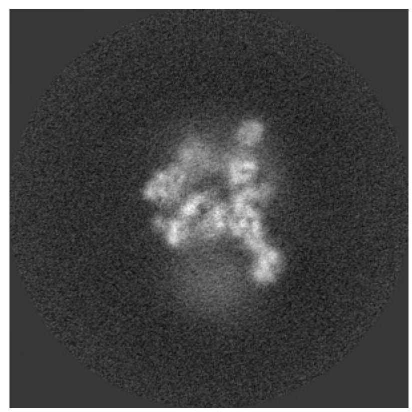
6.2.2 Raw map



X Index: 200



Y Index: 200

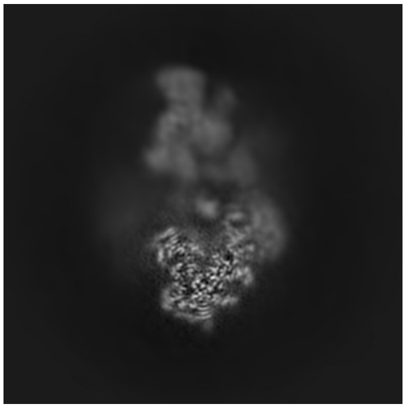


Z Index: 200

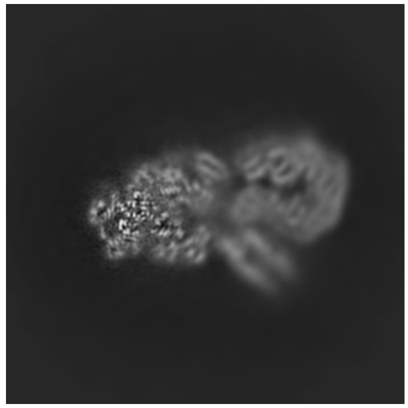
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ

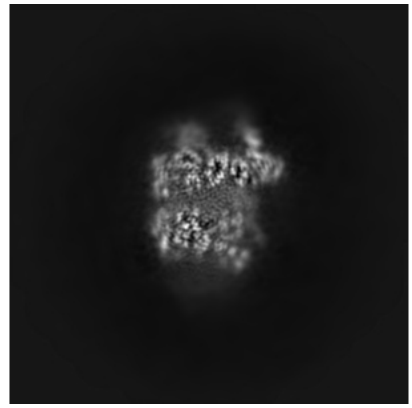
6.3.1 Primary map



X Index: 186

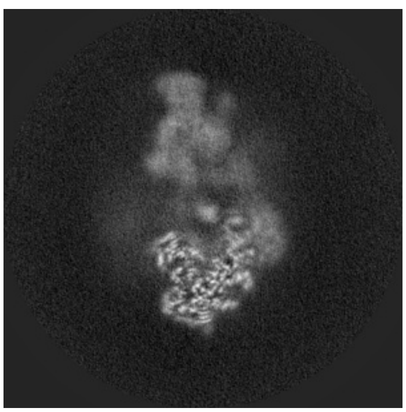


Y Index: 185

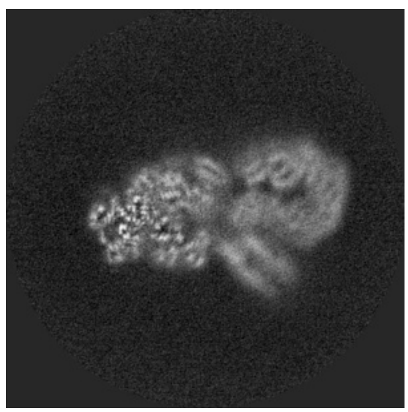


Z Index: 154

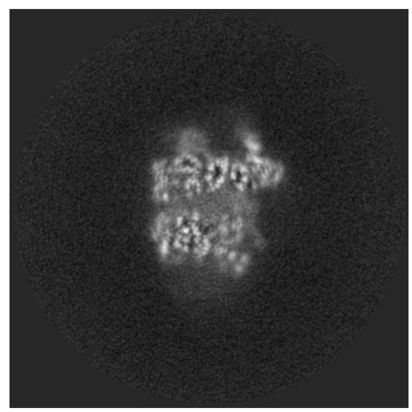
6.3.2 Raw map



X Index: 186



Y Index: 185

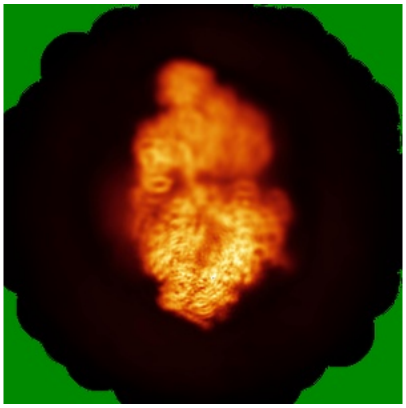


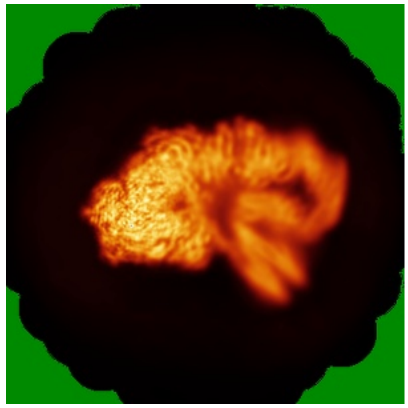
Z Index: 154

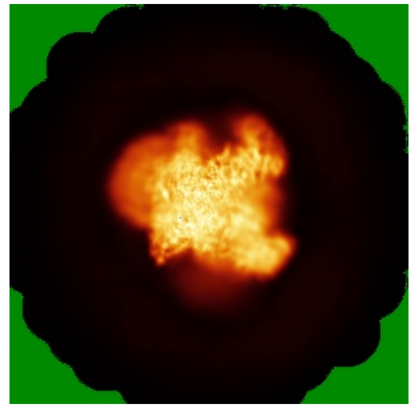
The images above show the largest variance slices of the map in three orthogonal directions.

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

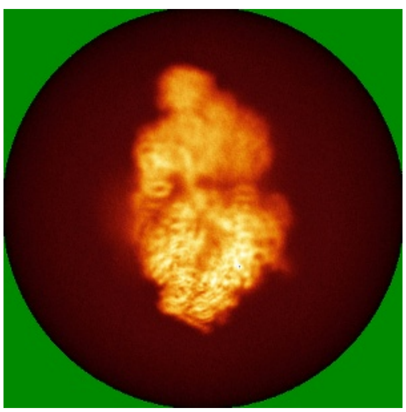
6.4.1 Primary map

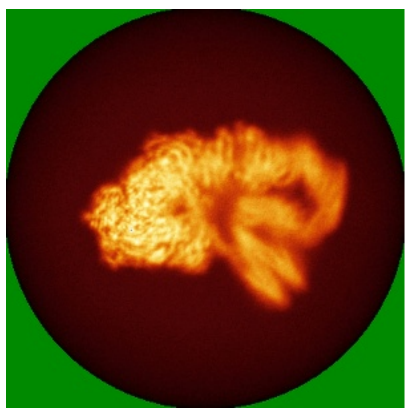

X

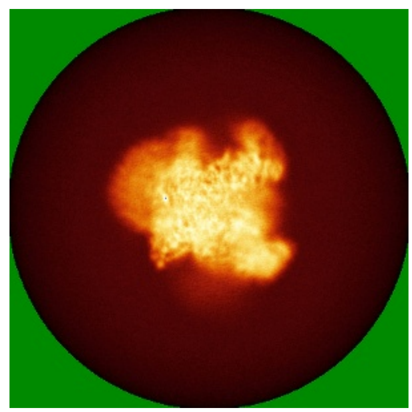

Y


Z

6.4.2 Raw map


X

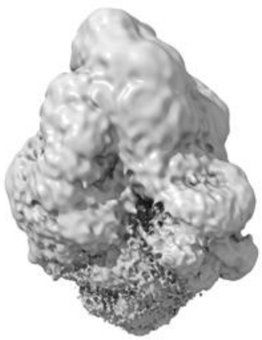

Y


Z

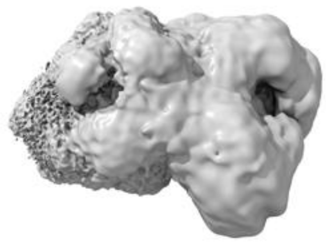
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

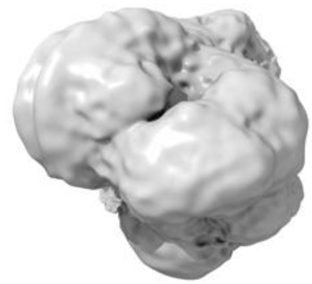
6.5.1 Primary map



X



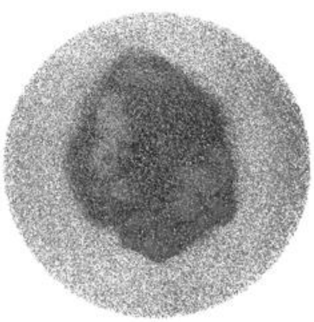
Y



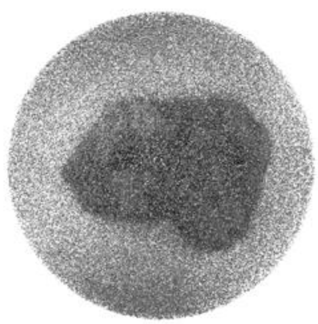
Z

The images above show the 3D surface view of the map at the recommended contour level 0.4. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

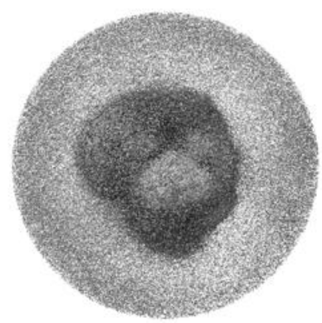
6.5.2 Raw map



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.

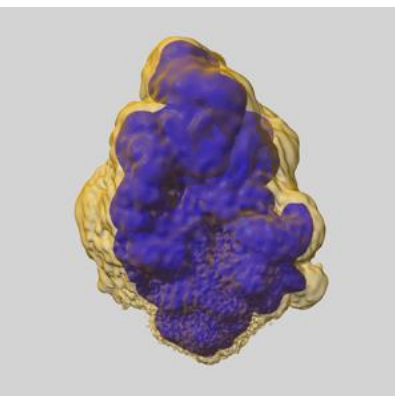
6.6 Mask visualisation (i)

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

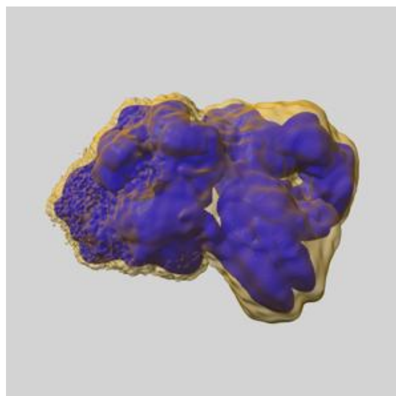
A mask typically either:

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

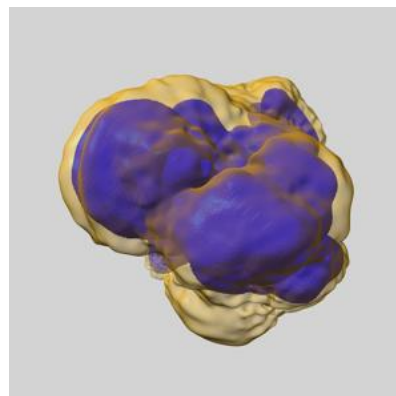
6.6.1 emd_16331_msk_1.map (i)



X



Y

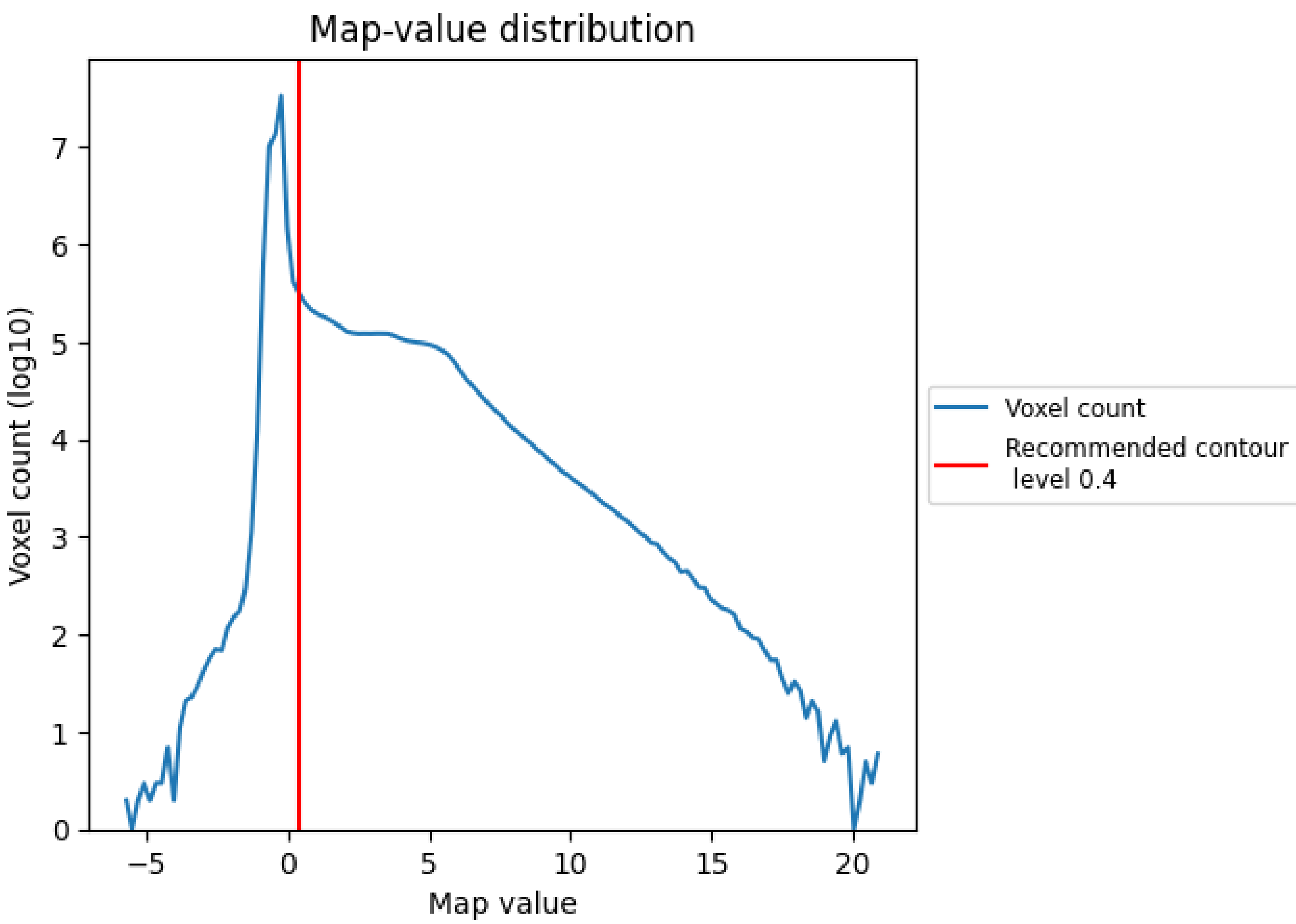


Z

7 Map analysis (i)

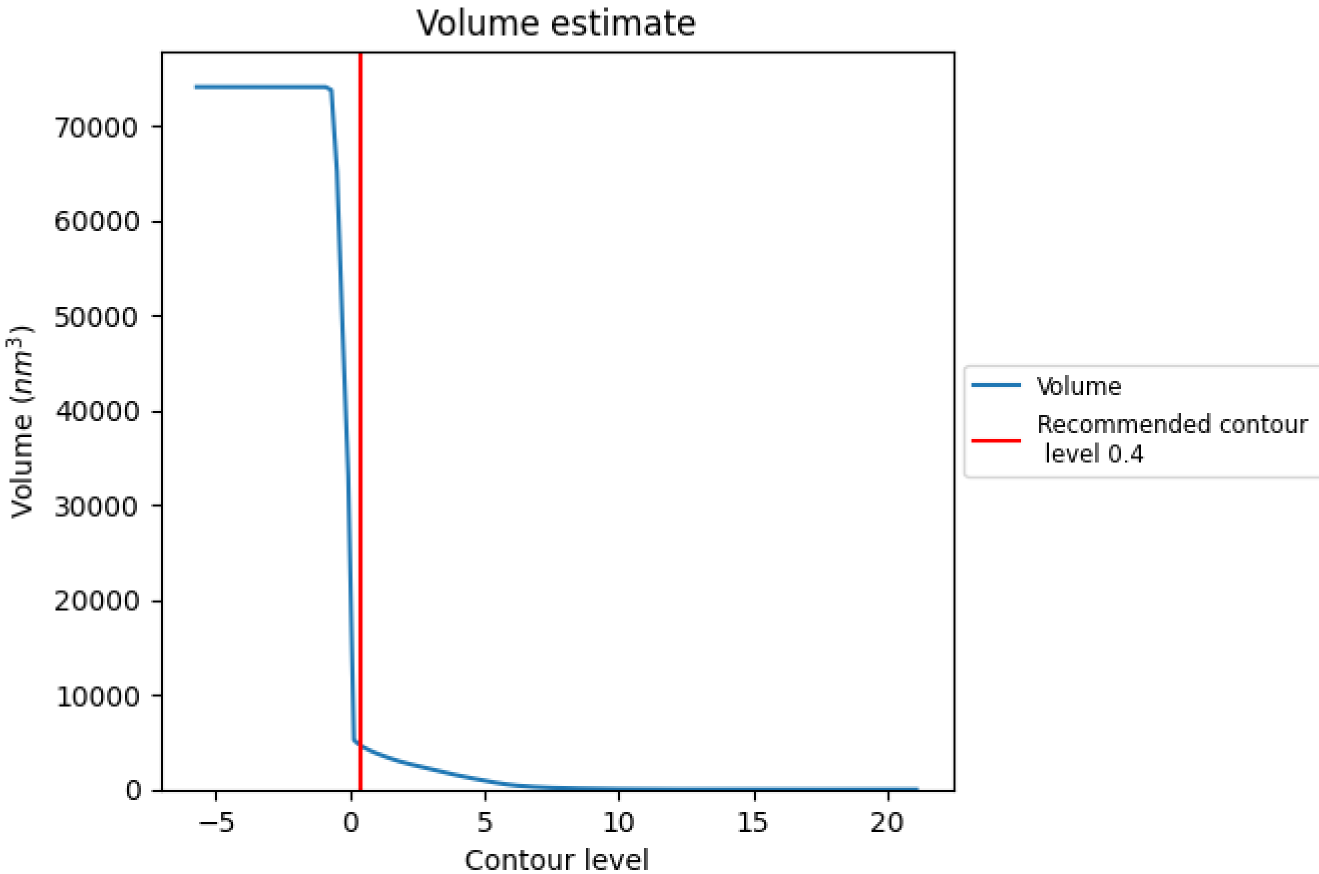
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



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

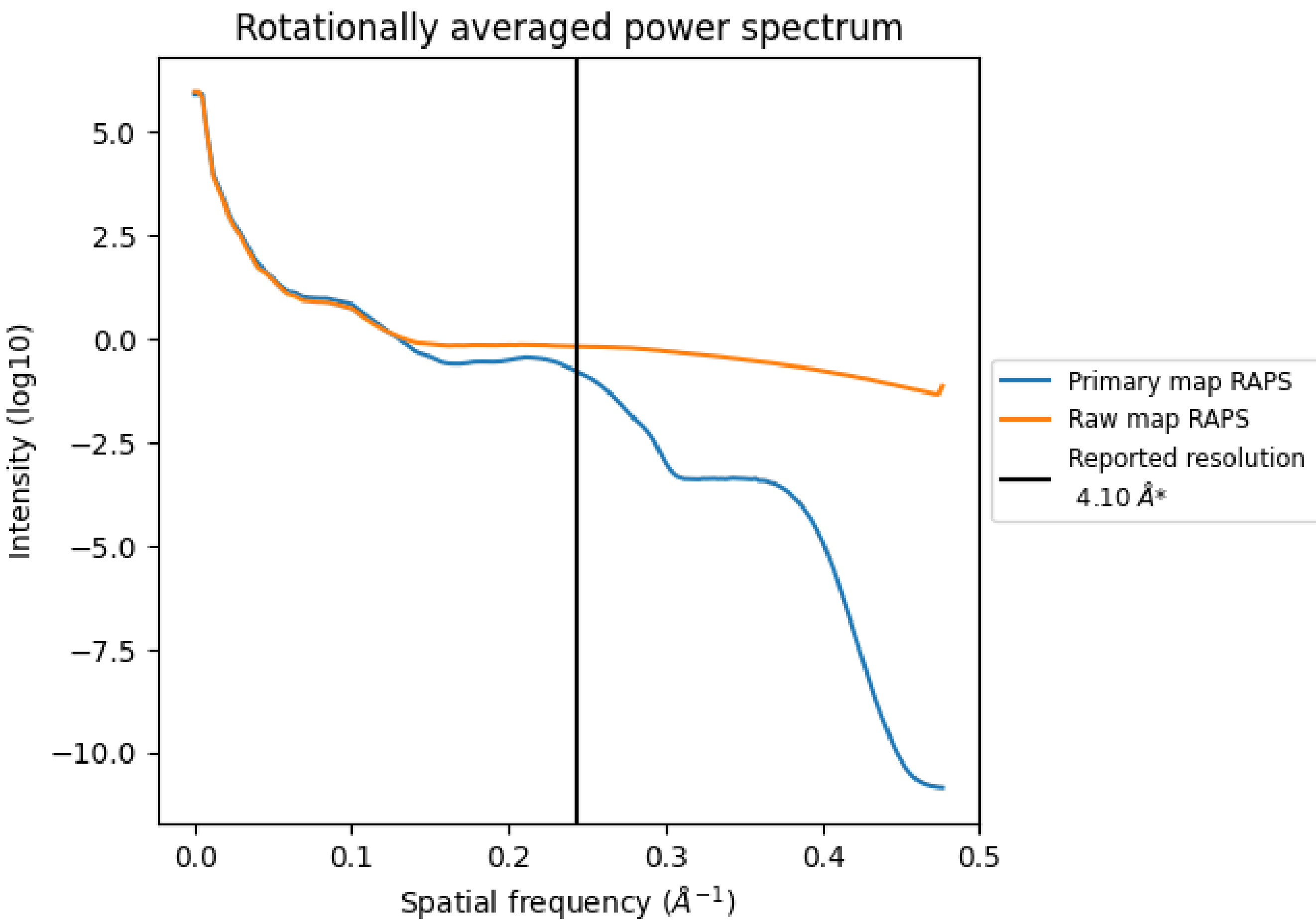
7.2 Volume estimate (i)



The volume at the recommended contour level is 4663 nm^3; this corresponds to an approximate mass of 4212 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)

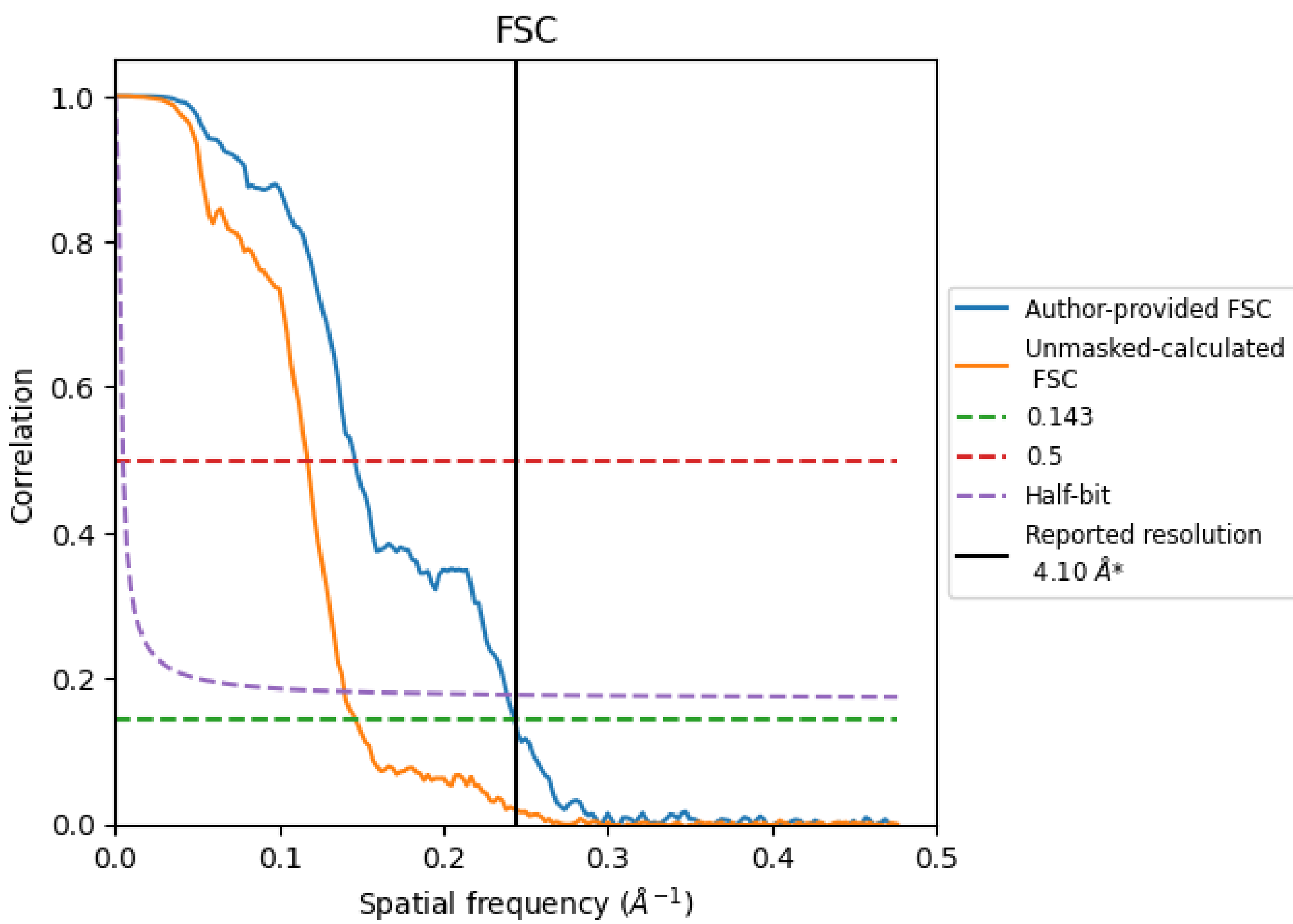


*Reported resolution corresponds to spatial frequency of 0.244 Å^{-1}

8 Fourier-Shell correlation (i)

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

8.1 FSC (i)



*Reported resolution corresponds to spatial frequency of 0.244 Å^{-1}

8.2 Resolution estimates (i)

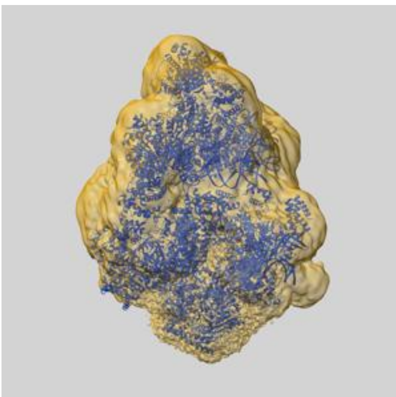
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.10	-	-
Author-provided FSC curve	4.11	6.85	4.19
Unmasked-calculated*	6.83	8.54	7.15

*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 6.83 differs from the reported value 4.1 by more than 10 %

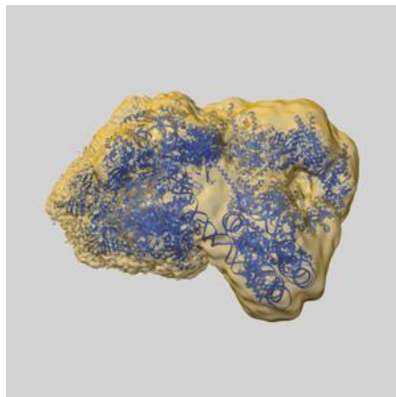
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-16331 and PDB model 8BYQ. Per-residue inclusion information can be found in section 3 on page 13.

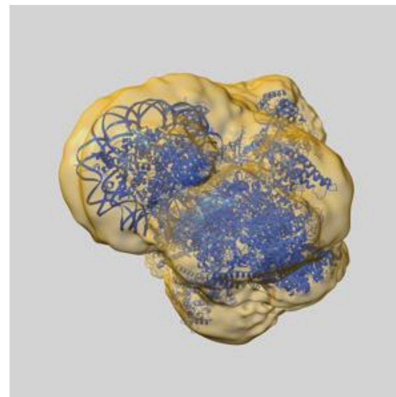
9.1 Map-model overlay (i)



X



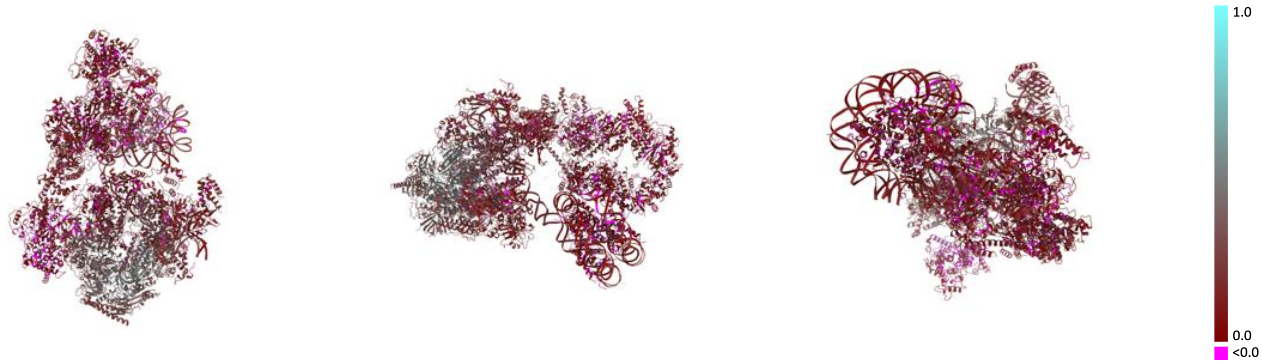
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.4 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.

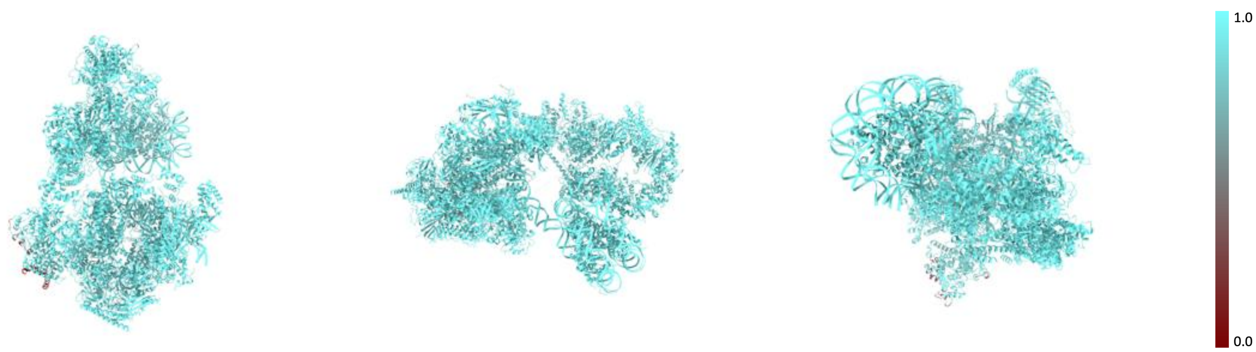
9.2 Q-score mapped to coordinate model (i)



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.

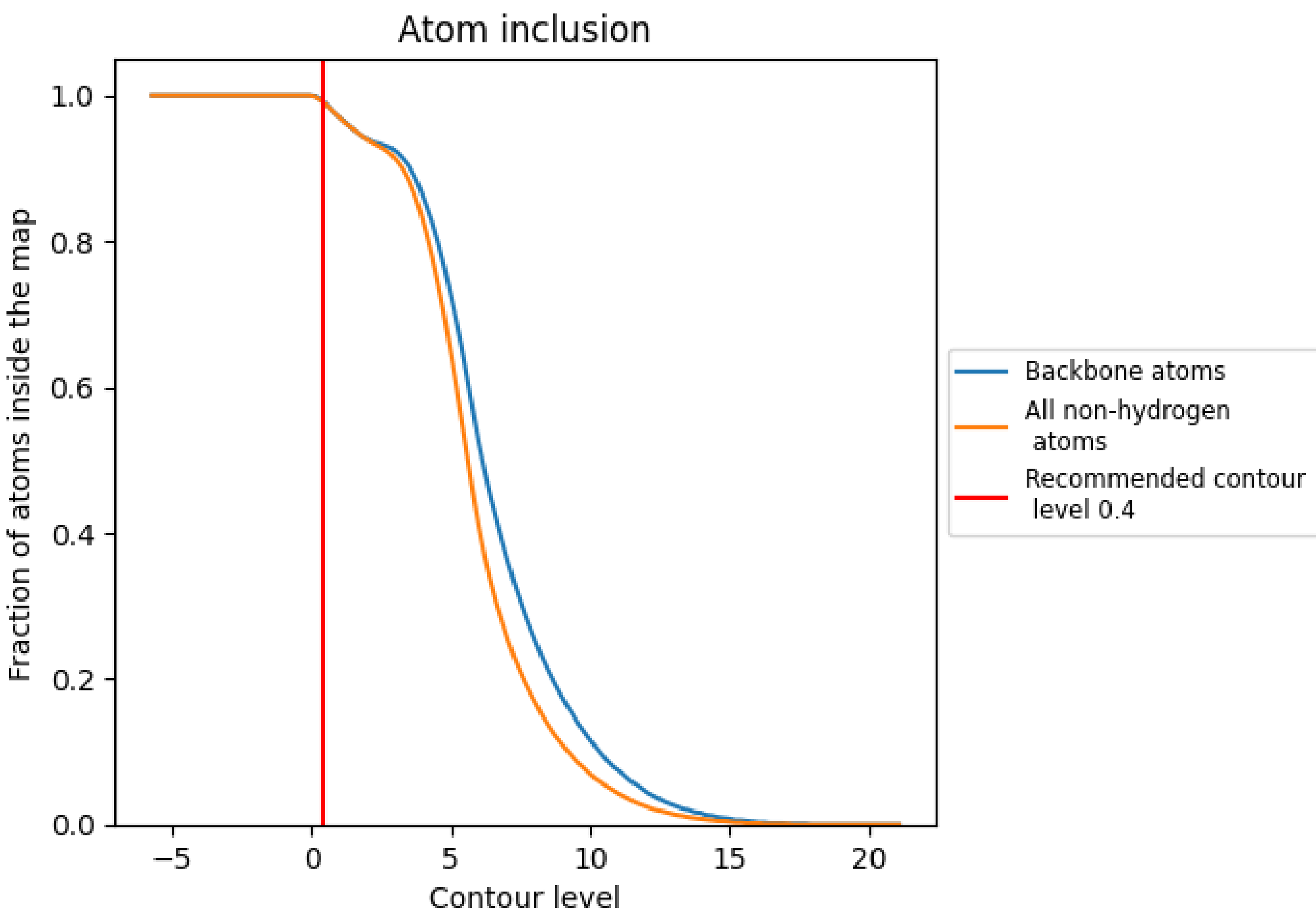
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.4).

9.4 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 99% 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 (0.4) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9920	0.1870
0	1.0000	0.1050
1	1.0000	0.1460
2	1.0000	0.0750
3	0.9990	0.1110
4	1.0000	0.1230
5	1.0000	0.1010
6	1.0000	0.0890
7	0.9570	0.1170
8	0.9540	0.0540
9	0.8120	0.0410
A	1.0000	0.2890
B	1.0000	0.3560
C	1.0000	0.3560
D	1.0000	0.1380
E	1.0000	0.2410
F	1.0000	0.3430
G	1.0000	0.1690
H	1.0000	0.3070
I	1.0000	0.2160
J	1.0000	0.3860
K	1.0000	0.3430
L	1.0000	0.3410
M	1.0000	0.2650
N	1.0000	0.1450
O	1.0000	0.1900
Q	1.0000	0.1400
R	0.9990	0.1590
T	1.0000	0.1470
U	0.9950	0.1500
V	1.0000	0.1430
W	1.0000	0.1230
X	1.0000	0.1340
a	1.0000	0.0880
b	1.0000	0.0830

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
c	1.0000	0.0740
d	1.0000	0.0770
e	1.0000	0.1130
f	1.0000	0.1010
g	1.0000	0.0760
h	1.0000	0.0900