



## wwPDB EM Validation Summary Report ⓘ

Nov 19, 2022 – 07:23 pm GMT

PDB ID : 5OQJ  
EMDB ID : EMD-3846  
Title : STRUCTURE OF YEAST TRANSCRIPTION PRE-INITIATION COMPLEX WITH TFIIH  
Authors : Schilbach, S.; Hantsche, M.; Tegunov, D.; Dienemann, C.; Wigge, C.; Henning, U.; Cramer, P.  
Deposited on : 2017-08-11  
Resolution : 4.70 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

EMDB validation analysis : 0.0.1.dev43  
Mogul : 1.8.4, CSD as541be (2020)  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.2

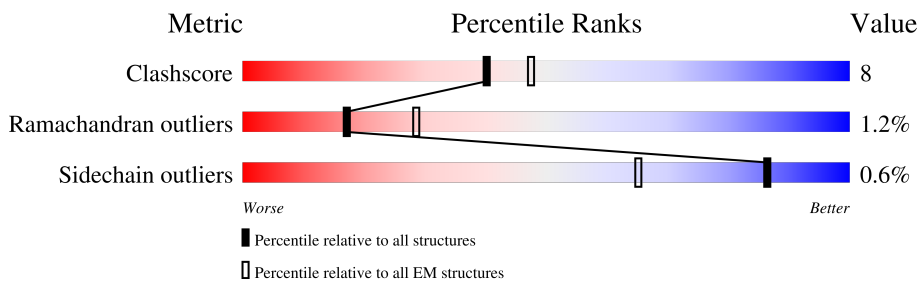
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.70 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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

Mol	Chain	Length	Quality of chain
1	A	1733	
2	B	1224	
3	C	318	
4	D	221	
5	E	215	
6	F	155	
7	G	171	
8	H	146	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	122	71% 23% 5%
10	J	70	80% 13% 7%
11	K	120	77% 17% 7%
12	L	70	50% 14% 36%
13	M	345	63% 17% 19%
14	N	106	9% 38% 34% 28%
15	O	240	58% 17% 25%
16	Q	735	15% 5% 80%
17	R	400	38% 10% 52%
18	T	106	11% 38% 30% 32%
19	U	286	24% 8% 68%
20	V	122	70% 12% 18%
21	W	586	5% 38% 6% 56%
22	X	328	44% 5% 51%
23	0	778	5% 75% 14% 10%
24	1	641	13% 70% 6% 23%
25	2	462	81% 5% 15%
26	3	321	38% 5% 57%
27	4	338	83% 12% 5%
28	5	72	18% 65% 26% 8%
29	6	461	67% 6% 27%
30	7	843	36% 12% 52%
31	Z	43	88% 12%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
34	SF4	0	801	-	-	X	-

## 2 Entry composition [i](#)

There are 34 unique types of molecules in this entry. The entry contains 62931 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	1398	10997	6931	1927	2078	61	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	1152	9178	5807	1608	1708	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	262	2061	1299	343	406	13	0	0

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	213	1744	1107	308	318	11	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	83	670	428	114	125	3	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

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

- Molecule 8 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		
8	H	136	1089	686	184	215	4	0	0

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

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

- Molecule 10 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		
10	J	65	532	339	93	94	6	0	0

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

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

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

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

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

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

- Molecule 14 is a DNA chain called NONTEMPLATE DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
14	N	76	1533	747	288	422	76	0	0

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

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

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

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

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

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

- Molecule 18 is a DNA chain called TEMPLATE DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
18	T	72	1440	706	254	408	72	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	U	92	757	474	130	150	3	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	V	100	782	492	130	156	4	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
21	W	258	Total	C	N	O	S	0	0
			1825	1147	321	351	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
22	X	160	Total	C	N	O	S	0	0
			1004	620	184	196	4		

- Molecule 23 is a protein called DNA repair helicase RAD3.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	0	698	Total	C	N	O	S	0	0
			4844	3051	856	907	30		

- Molecule 24 is a protein called RNA polymerase II transcription factor B subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	1	491	Total	C	N	O	S	0	0
			3060	1893	570	590	7		

- Molecule 25 is a protein called RNA polymerase II transcription factor B subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
25	2	395	Total	C	N	O	0	0
			2160	1317	413	430		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
26	3	138	Total	C	N	O	S	0	0
			860	533	160	160	7		

- Molecule 27 is a protein called RNA polymerase II transcription factor B subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	4	297	Total	C	N	O	S	0	0
			1475	877	297	297	4		

- Molecule 28 is a protein called RNA polymerase II transcription factor B subunit 5.



Mol	Chain	Residues	Atoms					AltConf	Trace
28	5	66	Total	C	N	O	S	0	0
			498	314	89	93	2		

- Molecule 29 is a protein called Suppressor of stem-loop protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	6	335	Total	C	N	O	S	0	0
			2197	1355	399	422	21		

- Molecule 30 is a protein called DNA repair helicase RAD25.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	7	408	Total	C	N	O	S	0	0
			3148	2000	557	572	19		

- Molecule 31 is a protein called Unknown protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
31	Z	43	Total	C	N	O	0	0
			215	129	43	43		

- Molecule 32 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
32	A	2	Total	Zn	0
			2	2	
32	B	1	Total	Zn	0
			1	1	
32	C	1	Total	Zn	0
			1	1	
32	I	2	Total	Zn	0
			2	2	
32	J	1	Total	Zn	0
			1	1	
32	L	1	Total	Zn	0
			1	1	
32	M	1	Total	Zn	0
			1	1	
32	W	1	Total	Zn	0
			1	1	
32	3	2	Total	Zn	0
			2	2	

*Continued on next page...*

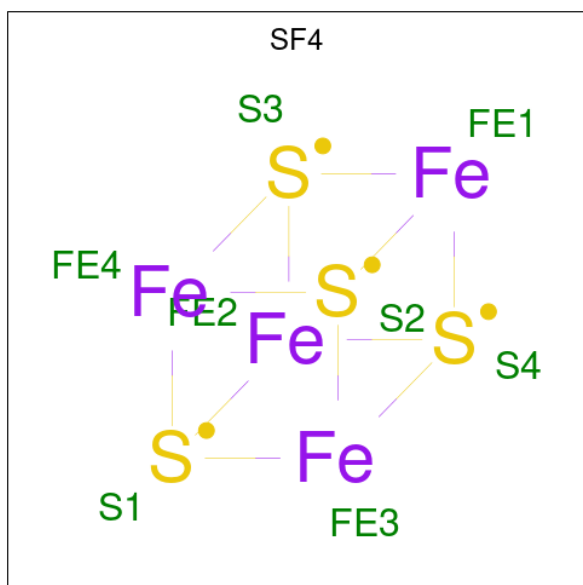
Continued from previous page...

Mol	Chain	Residues	Atoms		AltConf
32	4	1	Total	Zn	0
			1	1	
32	6	3	Total	Zn	0
			3	3	

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

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

- Molecule 34 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).

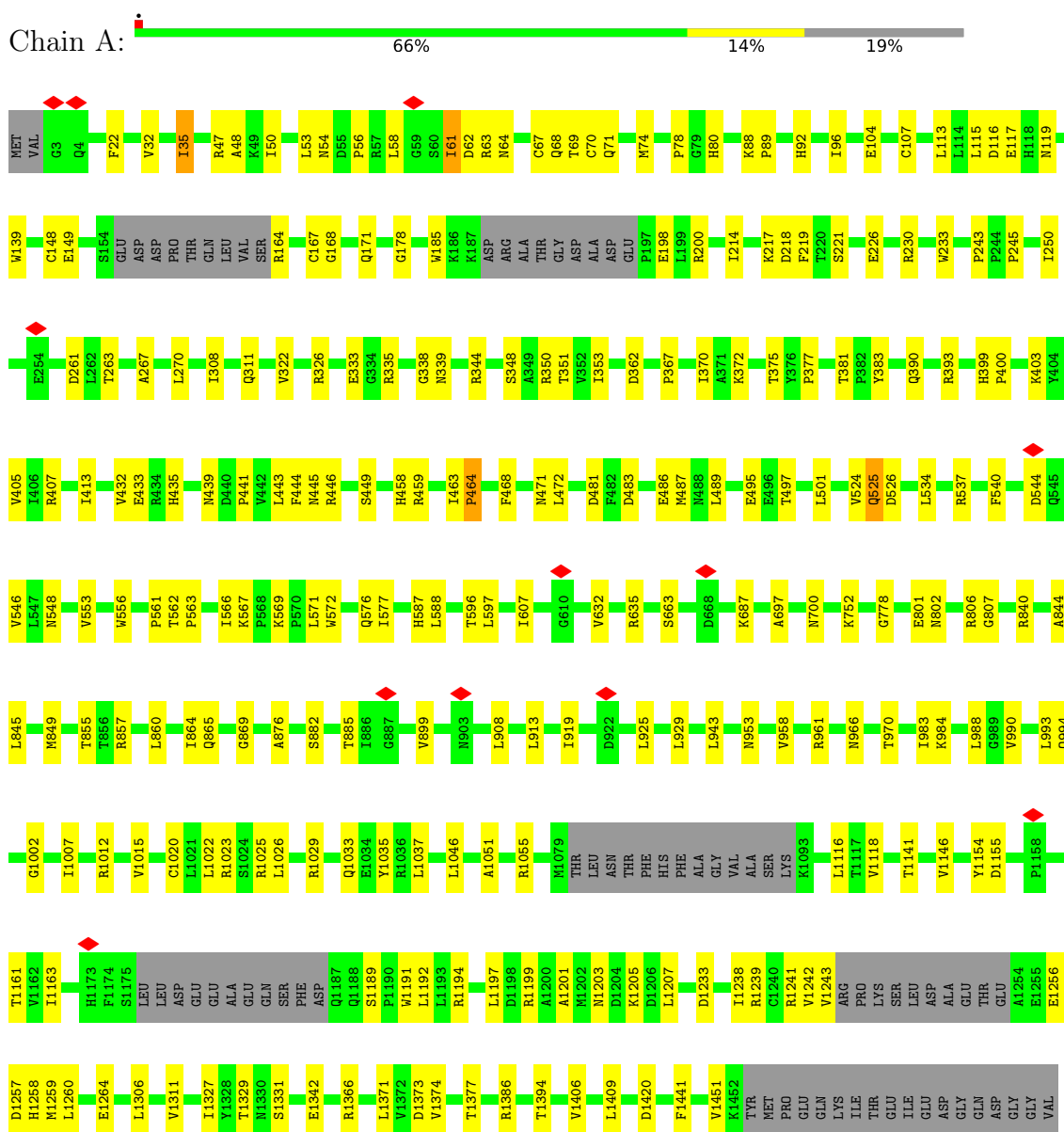


Mol	Chain	Residues	Atoms			AltConf
34	0	1	Total	Fe	S	0
			8	4	4	

### 3 Residue-property plots

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

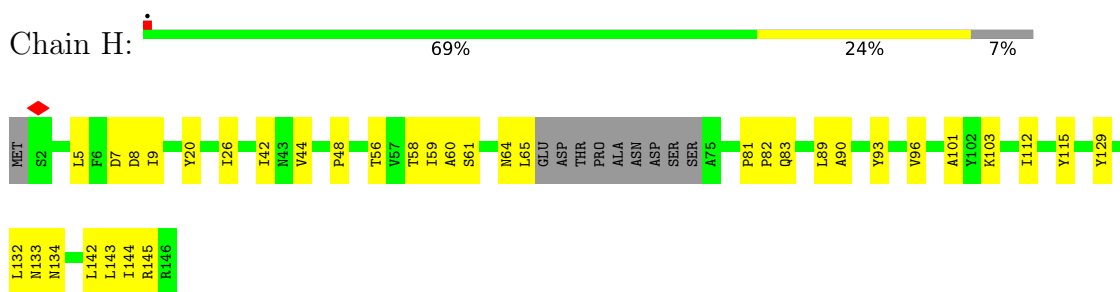
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1



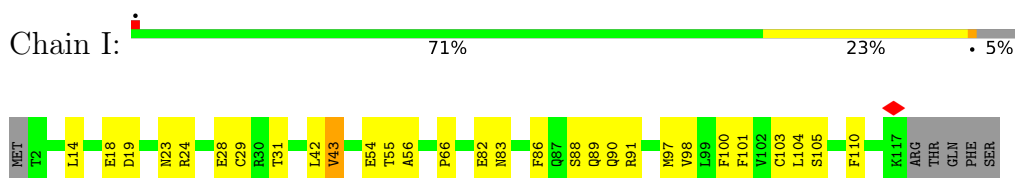




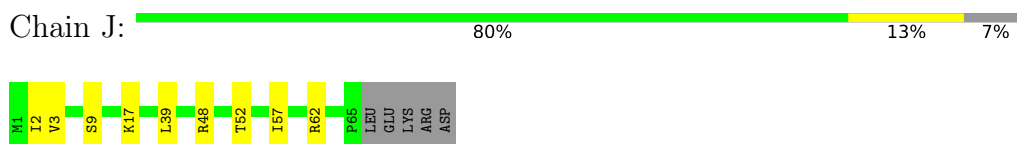
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



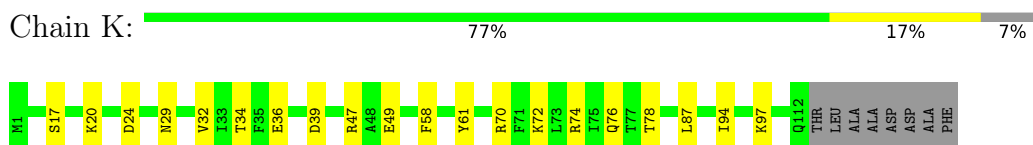
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



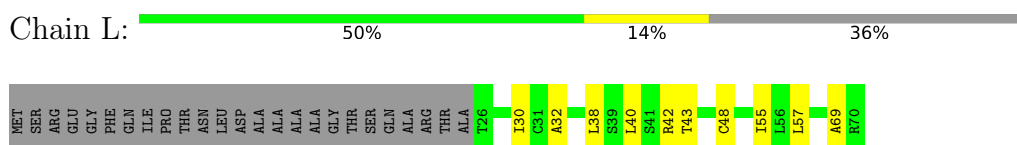
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



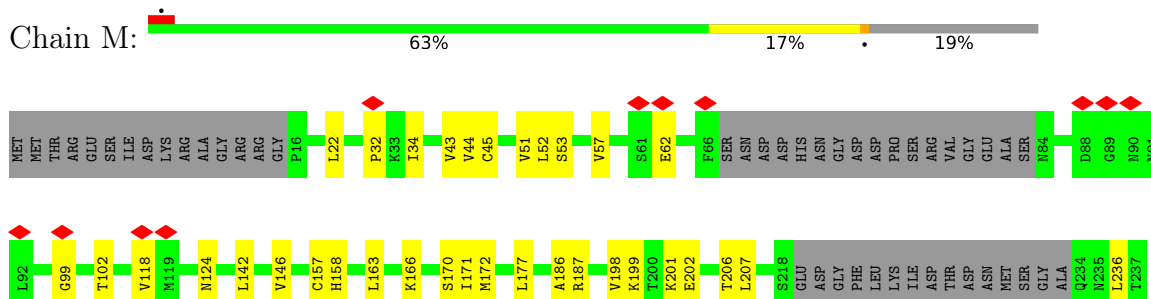
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11



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



- Molecule 13: Transcription initiation factor IIB







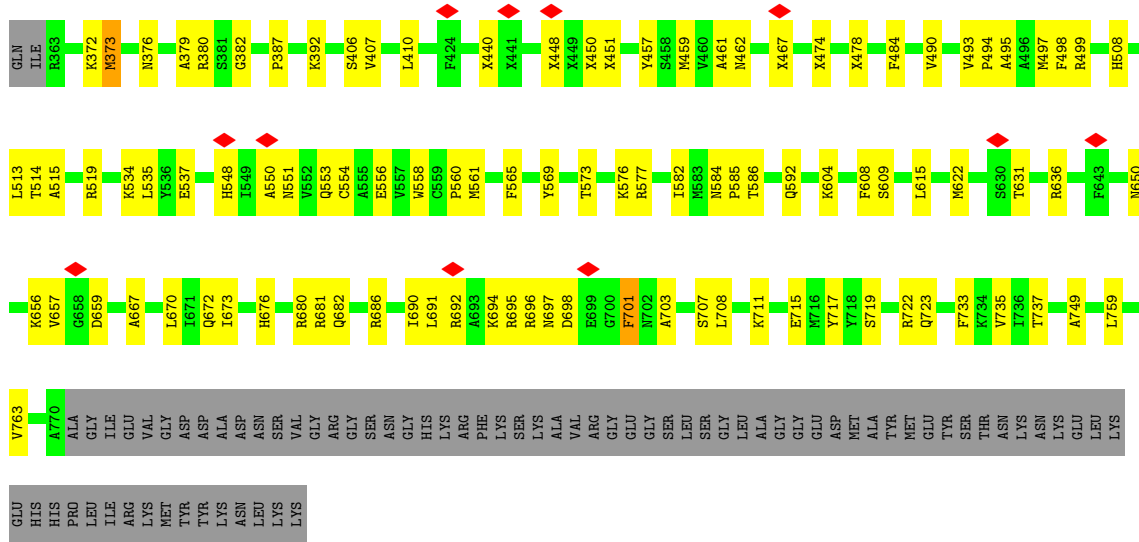




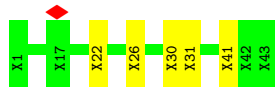
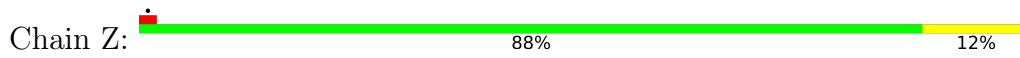








• Molecule 31: Unknown protein



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	58000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	42	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	105000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.592	Depositor
Minimum map value	-0.205	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.011	Depositor
Recommended contour level	0.066	Depositor
Map size (Å)	479.5, 479.5, 479.5	wwPDB
Map dimensions	350, 350, 350	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.37, 1.37, 1.37	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG, 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	0.25	0/11192	0.40	0/15128
2	B	0.24	0/9357	0.40	0/12618
3	C	0.24	0/2099	0.40	0/2845
4	D	0.23	0/1262	0.37	0/1693
5	E	0.24	0/1780	0.38	0/2395
6	F	0.23	0/682	0.38	0/922
7	G	0.25	0/1368	0.42	0/1844
8	H	0.25	0/1107	0.42	0/1499
9	I	0.24	0/962	0.43	0/1295
10	J	0.28	0/541	0.41	0/727
11	K	0.24	0/922	0.38	0/1244
12	L	0.23	0/360	0.45	0/478
13	M	0.24	0/2204	0.40	0/2963
14	N	0.54	0/1724	0.93	0/2614
15	O	0.25	0/1443	0.43	0/1942
16	Q	0.26	0/1168	0.44	0/1579
17	R	0.24	0/1312	0.42	0/1777
18	T	0.53	0/1614	0.96	0/2442
19	U	0.22	0/766	0.38	0/1032
20	V	0.23	0/789	0.39	0/1066
21	W	0.24	0/1551	0.38	0/2096
22	X	0.24	0/1013	0.40	0/1385
23	0	0.26	0/4739	0.41	0/6431
24	1	0.25	0/2238	0.39	0/3039
25	2	0.23	0/1631	0.38	0/2243
26	3	0.24	0/870	0.40	0/1190
27	4	0.27	0/1282	0.52	0/1780
28	5	0.23	0/502	0.41	0/677
29	6	0.27	0/1996	0.42	0/2713
30	7	0.24	0/2980	0.38	0/4019
All	All	0.27	0/61454	0.46	0/83676

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10997	0	11081	166	0
2	B	9178	0	9195	166	0
3	C	2061	0	2029	30	0
4	D	1253	0	1275	12	0
5	E	1744	0	1772	15	0
6	F	670	0	690	11	0
7	G	1340	0	1357	38	0
8	H	1089	0	1062	22	0
9	I	944	0	899	31	0
10	J	532	0	542	6	0
11	K	904	0	911	13	0
12	L	358	0	381	7	0
13	M	2175	0	2283	41	0
14	N	1533	0	860	33	0
15	O	1416	0	1493	25	0
16	Q	1144	0	1034	30	0
17	R	1303	0	1110	29	0
18	T	1440	0	820	32	0
19	U	757	0	747	13	0
20	V	782	0	790	10	0
21	W	1825	0	1543	27	0
22	X	1004	0	730	10	0
23	0	4844	0	4241	83	0
24	1	3060	0	2055	29	0
25	2	2160	0	1075	13	0
26	3	860	0	620	16	0
27	4	1475	0	624	9	0
28	5	498	0	506	12	0
29	6	2197	0	1738	15	0
30	7	3148	0	3001	67	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	Z	215	0	49	5	0
32	3	2	0	0	0	0
32	4	1	0	0	0	0
32	6	3	0	0	0	0
32	A	2	0	0	0	0
32	B	1	0	0	0	0
32	C	1	0	0	0	0
32	I	2	0	0	0	0
32	J	1	0	0	0	0
32	L	1	0	0	0	0
32	M	1	0	0	0	0
32	W	1	0	0	0	0
33	A	1	0	0	0	0
34	0	8	0	0	3	0
All	All	62931	0	56513	909	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

The worst 5 of 909 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:I:54:GLU:O	9:I:89:GLN:N	1.93	1.02
19:U:242:ASN:HA	19:U:268:THR:O	1.62	0.99
31:Z:31:UNK:HA	31:Z:41:UNK:CB	1.92	0.97
1:A:70:CYS:SG	1:A:80:HIS:NE2	2.40	0.95
31:Z:30:UNK:O	31:Z:41:UNK:CB	2.18	0.92

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1386/1733 (80%)	1297 (94%)	78 (6%)	11 (1%)	19	60
2	B	1136/1224 (93%)	1064 (94%)	62 (6%)	10 (1%)	17	56
3	C	260/318 (82%)	236 (91%)	20 (8%)	4 (2%)	10	46
4	D	153/221 (69%)	145 (95%)	7 (5%)	1 (1%)	22	62
5	E	211/215 (98%)	202 (96%)	8 (4%)	1 (0%)	29	68
6	F	81/155 (52%)	79 (98%)	2 (2%)	0	100	100
7	G	169/171 (99%)	156 (92%)	12 (7%)	1 (1%)	25	65
8	H	132/146 (90%)	117 (89%)	12 (9%)	3 (2%)	6	36
9	I	114/122 (93%)	99 (87%)	15 (13%)	0	100	100
10	J	63/70 (90%)	58 (92%)	3 (5%)	2 (3%)	4	30
11	K	110/120 (92%)	109 (99%)	1 (1%)	0	100	100
12	L	43/70 (61%)	37 (86%)	6 (14%)	0	100	100
13	M	273/345 (79%)	252 (92%)	16 (6%)	5 (2%)	8	41
15	O	178/240 (74%)	164 (92%)	13 (7%)	1 (1%)	25	65
16	Q	140/735 (19%)	119 (85%)	16 (11%)	5 (4%)	3	28
17	R	176/400 (44%)	160 (91%)	15 (8%)	1 (1%)	25	65
19	U	88/286 (31%)	82 (93%)	4 (4%)	2 (2%)	6	36
20	V	96/122 (79%)	92 (96%)	3 (3%)	1 (1%)	15	54
21	W	196/586 (33%)	187 (95%)	7 (4%)	2 (1%)	15	54
22	X	156/328 (48%)	129 (83%)	23 (15%)	4 (3%)	5	34
23	0	659/778 (85%)	590 (90%)	54 (8%)	15 (2%)	6	36
24	1	316/641 (49%)	303 (96%)	9 (3%)	4 (1%)	12	48
25	2	284/462 (62%)	266 (94%)	17 (6%)	1 (0%)	34	72
26	3	136/321 (42%)	115 (85%)	18 (13%)	3 (2%)	6	37
27	4	257/338 (76%)	231 (90%)	17 (7%)	9 (4%)	3	28
28	5	64/72 (89%)	59 (92%)	5 (8%)	0	100	100
29	6	289/461 (63%)	260 (90%)	26 (9%)	3 (1%)	15	54
30	7	361/843 (43%)	333 (92%)	23 (6%)	5 (1%)	11	47
All	All	7527/11523 (65%)	6941 (92%)	492 (6%)	94 (1%)	17	50

5 of 94 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
16	Q	127	ILE
21	W	77	PRO
23	0	156	CYS
23	0	254	THR
27	4	88	PRO

### 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	1221/1520 (80%)	1218 (100%)	3 (0%)	93	96
2	B	1000/1061 (94%)	997 (100%)	3 (0%)	92	95
3	C	230/274 (84%)	228 (99%)	2 (1%)	78	88
4	D	139/200 (70%)	137 (99%)	2 (1%)	67	81
5	E	195/197 (99%)	194 (100%)	1 (0%)	88	93
6	F	73/137 (53%)	73 (100%)	0	100	100
7	G	152/152 (100%)	152 (100%)	0	100	100
8	H	119/128 (93%)	119 (100%)	0	100	100
9	I	110/116 (95%)	108 (98%)	2 (2%)	59	77
10	J	60/65 (92%)	60 (100%)	0	100	100
11	K	97/102 (95%)	97 (100%)	0	100	100
12	L	40/57 (70%)	40 (100%)	0	100	100
13	M	245/299 (82%)	245 (100%)	0	100	100
15	O	152/205 (74%)	152 (100%)	0	100	100
16	Q	109/641 (17%)	108 (99%)	1 (1%)	78	88
17	R	107/363 (30%)	106 (99%)	1 (1%)	78	88
19	U	84/260 (32%)	84 (100%)	0	100	100
20	V	90/108 (83%)	90 (100%)	0	100	100
21	W	161/244 (66%)	161 (100%)	0	100	100
22	X	62/295 (21%)	62 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	0	413/677 (61%)	406 (98%)	7 (2%)	60	78
24	1	179/341 (52%)	177 (99%)	2 (1%)	73	85
25	2	53/273 (19%)	53 (100%)	0	100	100
26	3	53/303 (18%)	53 (100%)	0	100	100
27	4	4/260 (2%)	3 (75%)	1 (25%)	0	4
28	5	53/66 (80%)	53 (100%)	0	100	100
29	6	173/378 (46%)	173 (100%)	0	100	100
30	7	315/695 (45%)	308 (98%)	7 (2%)	52	71
All	All	5689/9417 (60%)	5657 (99%)	32 (1%)	86	92

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
30	7	695	ARG
30	7	701	PHE
9	I	43	VAL
9	I	42	LEU
30	7	715	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 92 such sidechains are listed below:

Mol	Chain	Res	Type
15	O	91	ASN
23	0	651	ASN
16	Q	117	HIS
21	W	178	GLN
24	1	92	ASN

### 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
34	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
34	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.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
34	0	801	SF4	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
21	W	2
24	1	2
25	2	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	W	373:UNK	C	407:UNK	N	85.52
1	1	394:UNK	C	465:UNK	N	84.00
1	W	289:UNK	C	349:UNK	N	45.11
1	2	419:UNK	C	433:LEU	N	13.48
1	1	519:UNK	C	537:GLU	N	11.51

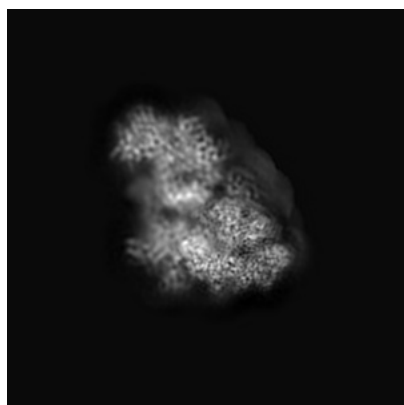
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-3846. These allow visual inspection of the internal detail of the map and identification of artifacts.

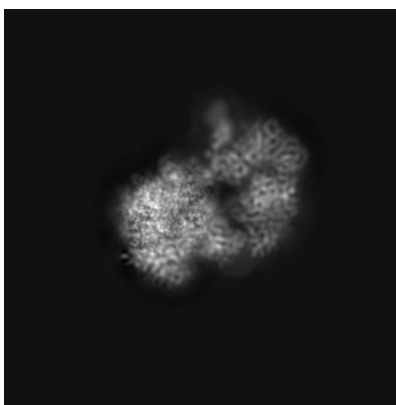
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

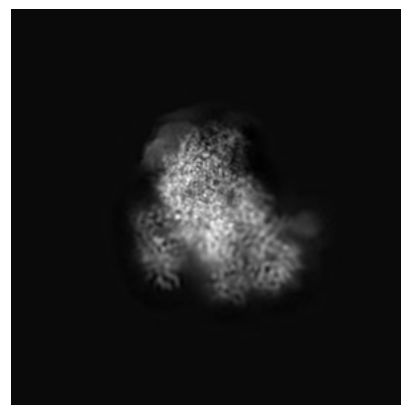
#### 6.1.1 Primary map



X



Y



Z

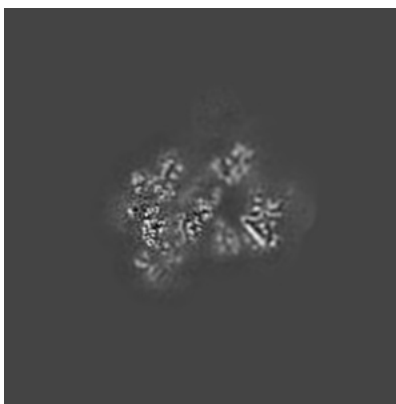
The images above show the map projected in three orthogonal directions.

### 6.2 Central slices [i](#)

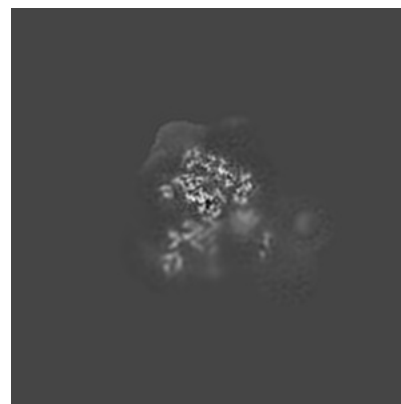
#### 6.2.1 Primary map



X Index: 175



Y Index: 175

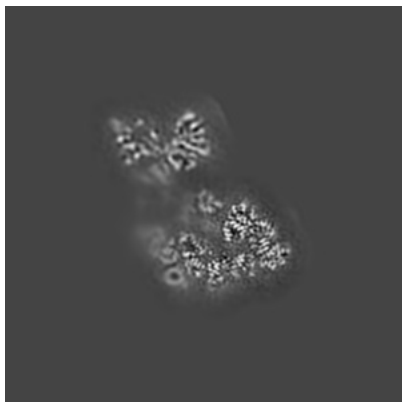


Z Index: 175

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

## 6.3 Largest variance slices [i](#)

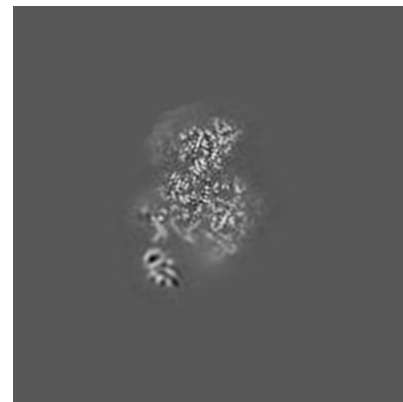
### 6.3.1 Primary map



X Index: 182



Y Index: 167

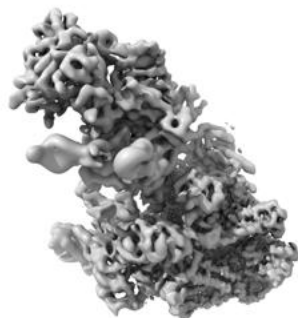


Z Index: 133

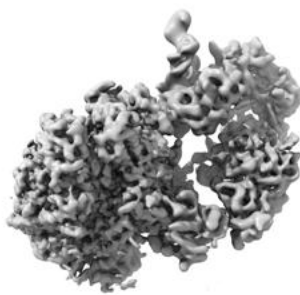
The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal surface views [i](#)

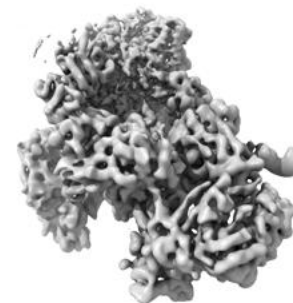
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.066. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.



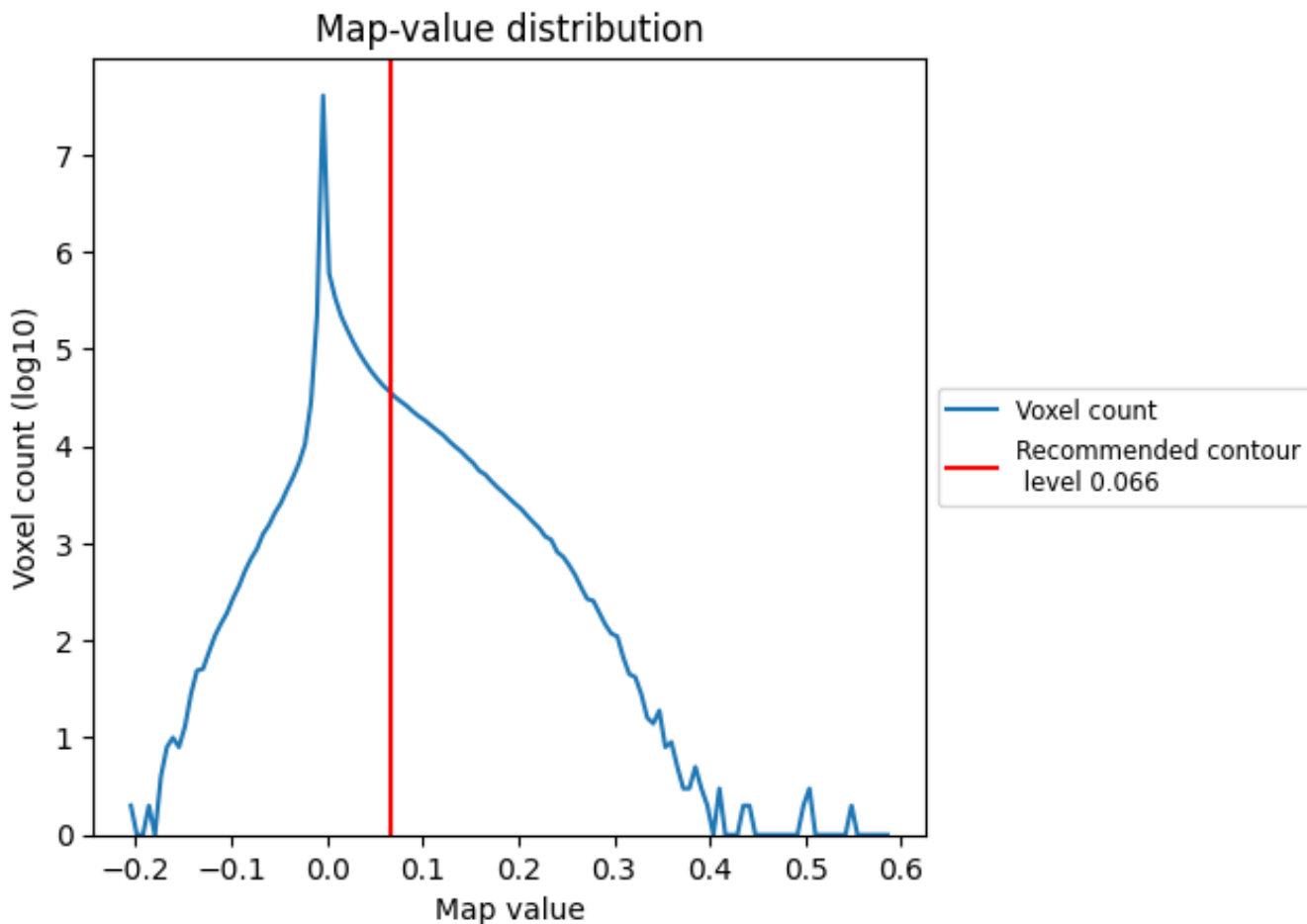
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

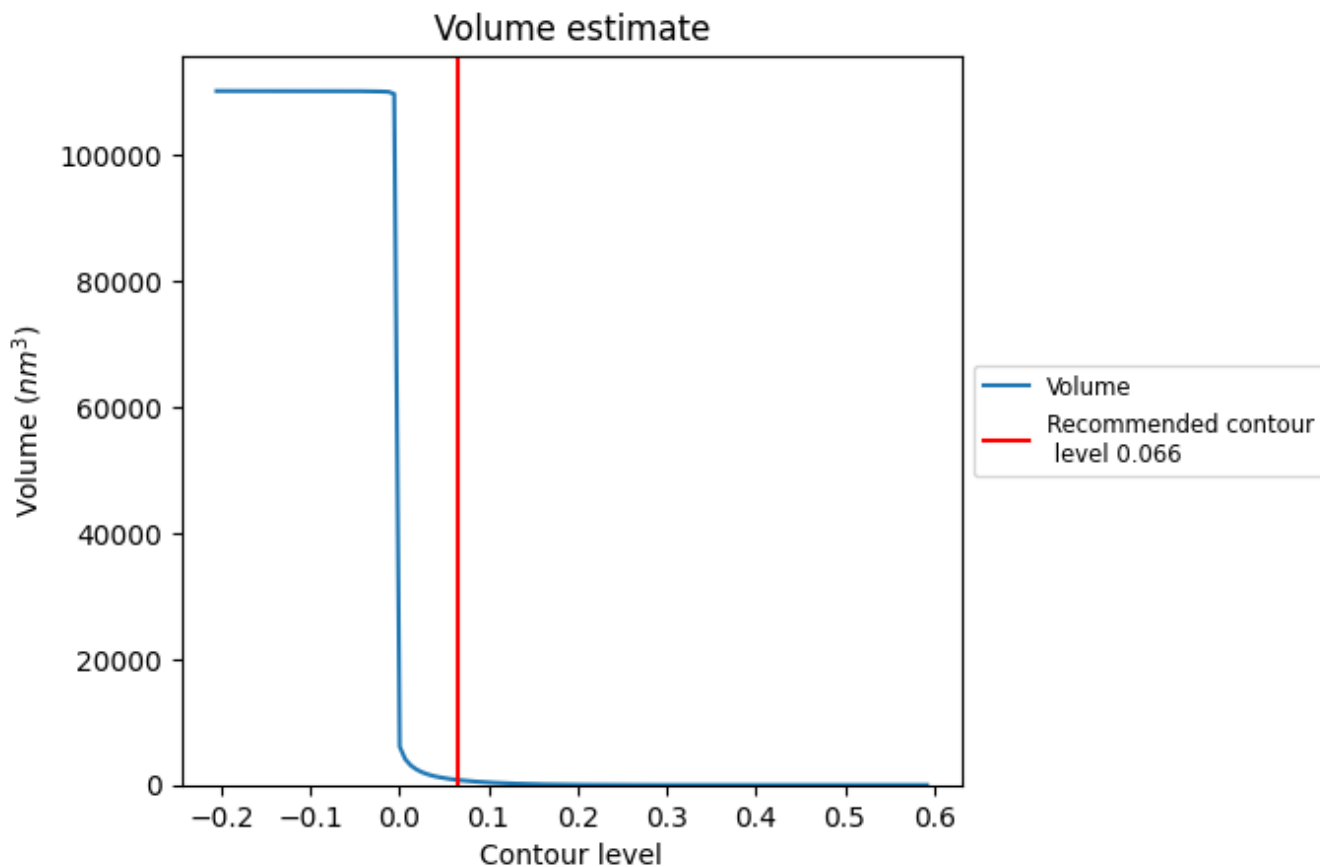
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



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

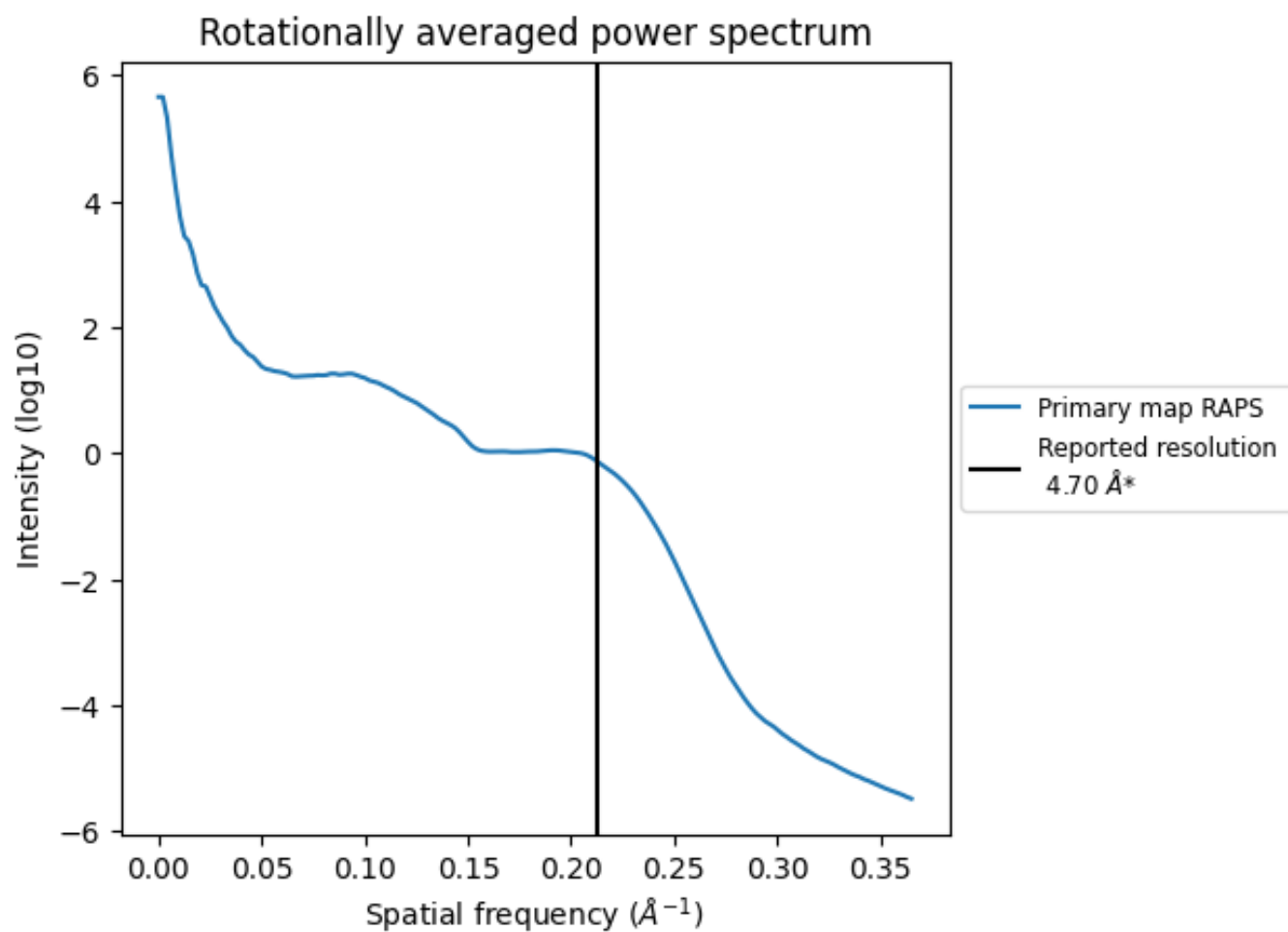
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 790  $\text{nm}^3$ ; this corresponds to an approximate mass of 714 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 0.213 Å<sup>-1</sup>

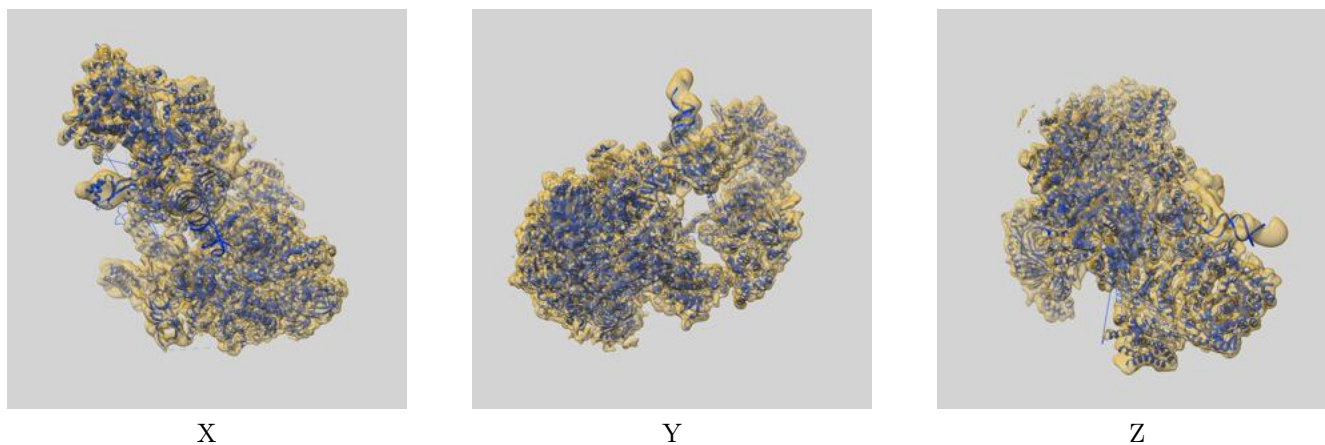
## 8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

## 9 Map-model fit [i](#)

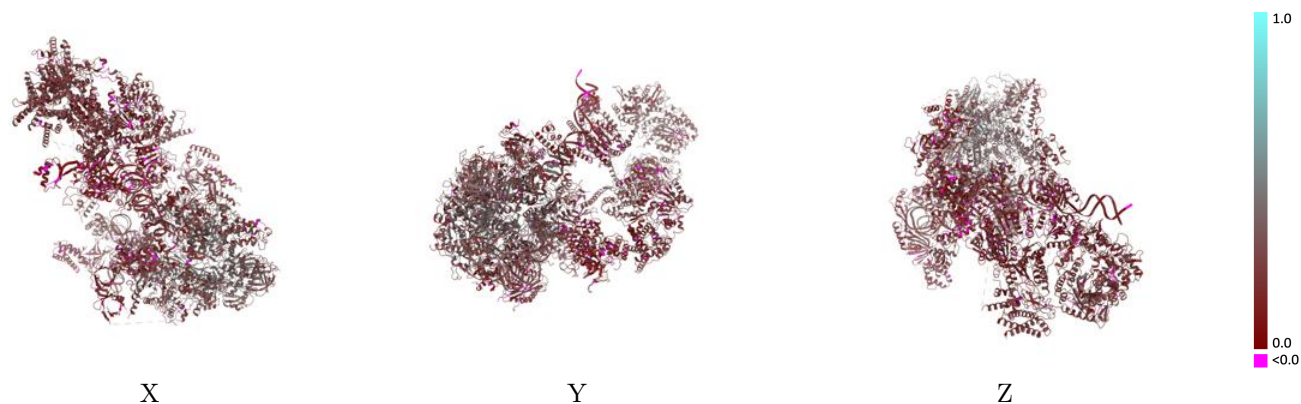
This section contains information regarding the fit between EMDB map EMD-3846 and PDB model 5OQJ. Per-residue inclusion information can be found in section 3 on page 11.

### 9.1 Map-model overlay [i](#)



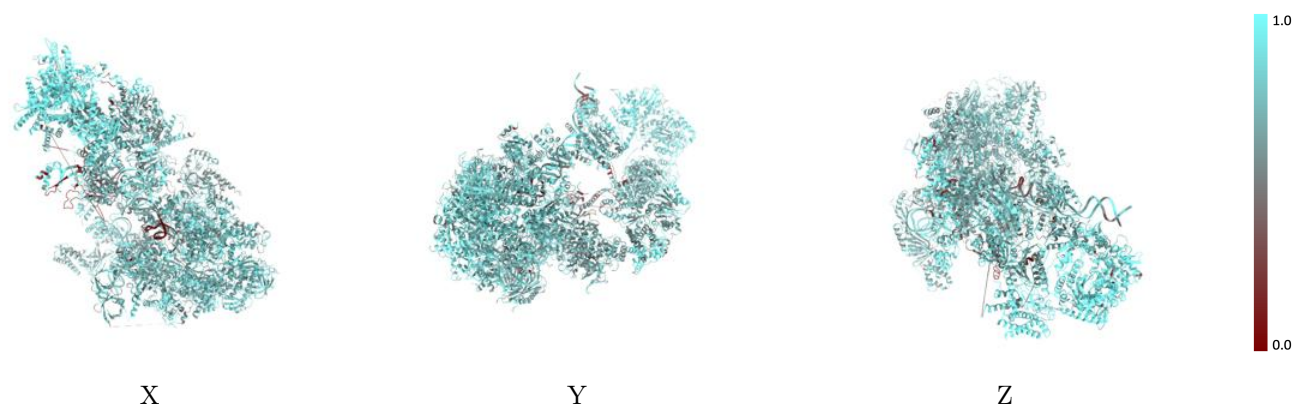
The images above show the 3D surface view of the map at the recommended contour level 0.066 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



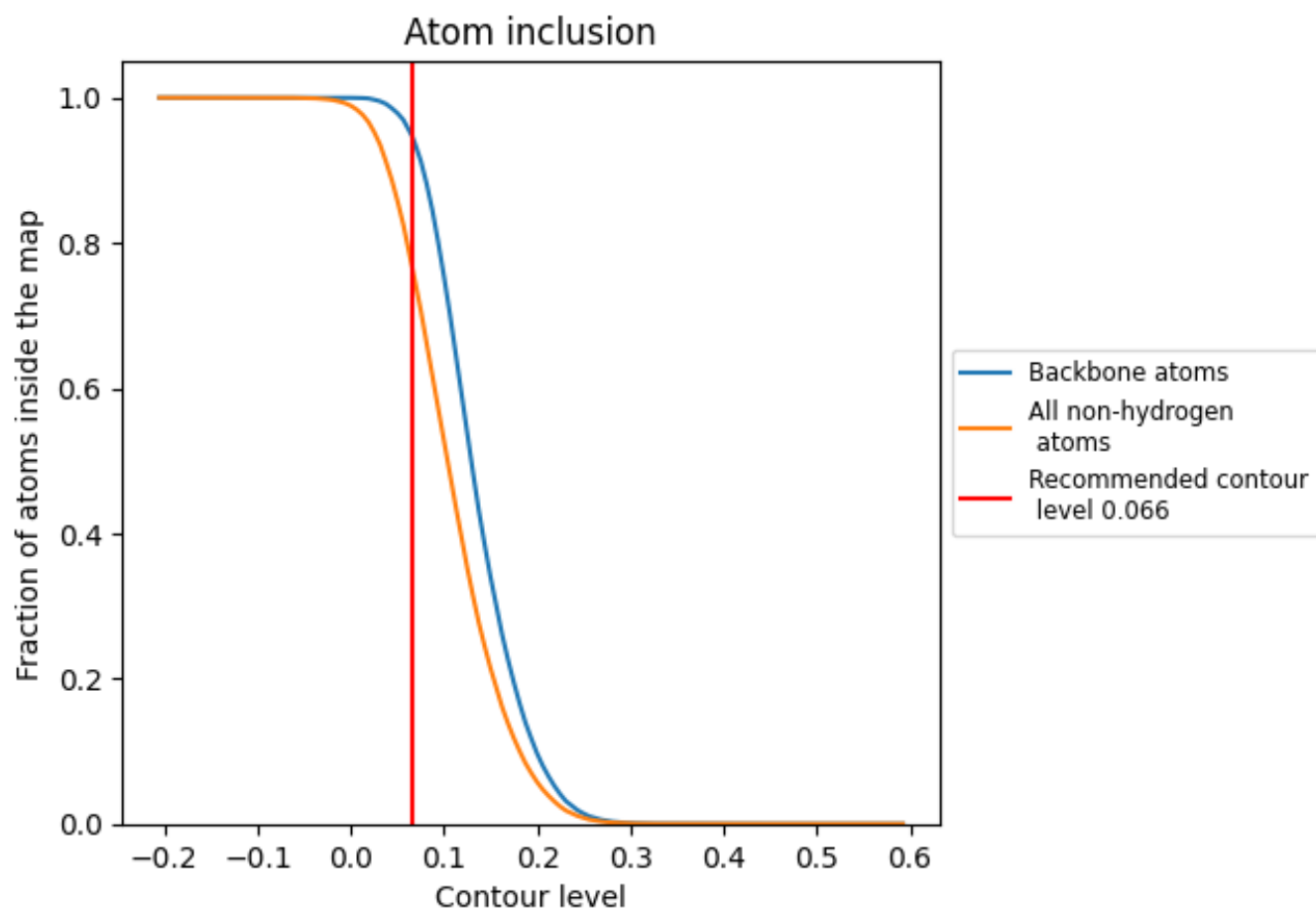
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

## 9.3 Atom inclusion mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.066).

## 9.4 Atom inclusion [i](#)



































































At the recommended contour level, 95% of all backbone atoms, 77% 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 (0.066) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7669	 0.2490
0	 0.7360	 0.1840
1	 0.6904	 0.2060
2	 0.8871	 0.2210
3	 0.7576	 0.1950
4	 0.9099	 0.2680
5	 0.6402	 0.1280
6	 0.7987	 0.2070
7	 0.7480	 0.1620
A	 0.7919	 0.3100
B	 0.7911	 0.3170
C	 0.8309	 0.3480
D	 0.6342	 0.1680
E	 0.8054	 0.2850
F	 0.8037	 0.3440
G	 0.7065	 0.2100
H	 0.7964	 0.3310
I	 0.7970	 0.2700
J	 0.8472	 0.3430
K	 0.8238	 0.3430
L	 0.8006	 0.3200
M	 0.6696	 0.2080
N	 0.7365	 0.2070
O	 0.7089	 0.1580
Q	 0.7337	 0.2000
R	 0.7798	 0.2240
T	 0.7417	 0.2150
U	 0.7085	 0.1350
V	 0.7119	 0.1490
W	 0.6565	 0.1770
X	 0.7495	 0.1850
Z	 0.9302	 0.2610

