



wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 03:39 AM EST

PDB ID : 6XP5
EMDB ID : EMD-22284
Title : Head-Middle module of Mediator
Authors : Zhang, H.Q.; Chen, D.C.; Kornberg, R.D.
Deposited on : 2020-07-08
Resolution : 4.20 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 ⓘ 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
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

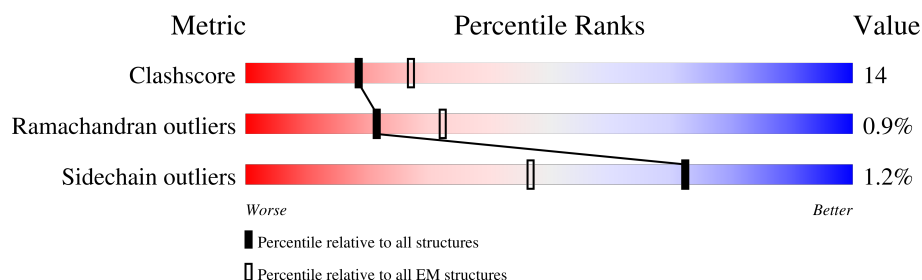
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.20 Å.

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





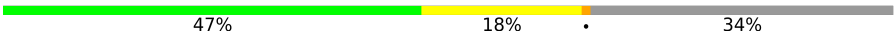
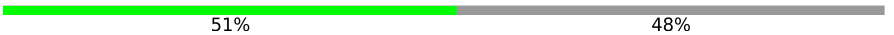

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	U	135	
2	D	341	
3	F	337	
4	G	246	
5	I	179	
6	N	1171	
7	Q	634	
8	Z	124	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	H	290	
10	K	238	
11	R	340	
12	V	169	
13	a	736	
14	b	21	
15	c	200	

2 Entry composition

There are 15 unique types of molecules in this entry. The entry contains 21065 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 Mediator of RNA polymerase II transcription subunit 21.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	U	129	Total	C	N	O	S	0	0
			1055	662	181	210	2		

- Molecule 2 is a protein called Mediator of RNA polymerase II transcription subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	D	134	Total	C	N	O	S	0	0
			994	616	178	198	2		

- Molecule 3 is a protein called Mediator of RNA polymerase II transcription subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	F	156	Total	C	N	O	S	0	0
			1202	763	202	231	6		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	136	ALA	-	insertion	UNP G0SGT8
F	206	PHE	THR	conflict	UNP G0SGT8

- Molecule 4 is a protein called Mediator of RNA polymerase II transcription subunit 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	122	Total	C	N	O	S	0	0
			836	524	150	159	3		

- Molecule 5 is a protein called Mediator of RNA polymerase II transcription subunit 9.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	I	96	Total	C	N	O	0	0
			470	278	96	96		

- Molecule 6 is a protein called Mediator of RNA polymerase II transcription subunit 14.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	N	544	Total	C	N	O	S	0	0
			4081	2590	719	762	10		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	741	ALA	-	insertion	UNP G0SCL5
N	742	ALA	-	insertion	UNP G0SCL5
N	743	ALA	-	insertion	UNP G0SCL5

- Molecule 7 is a protein called Mediator of RNA polymerase II transcription subunit 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	Q	476	Total	C	N	O	S	0	0
			3571	2227	649	680	15		

- Molecule 8 is a protein called Mediator of RNA polymerase II transcription subunit 31.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	Z	98	Total	C	N	O	S	0	0
			775	504	136	132	3		

- Molecule 9 is a protein called Mediator of RNA polymerase II transcription subunit 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	H	146	Total	C	N	O	S	0	0
			1148	722	201	222	3		

- Molecule 10 is a protein called Mediator of RNA polymerase II transcription subunit 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	107	Total	C	N	O	S	0	0
			777	482	137	158			

- Molecule 11 is a protein called Mediator of RNA polymerase II transcription subunit 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	R	207	Total	C	N	O	S	0	0
			1255	805	213	236	1		

- Molecule 12 is a protein called Med22.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	V	112	Total	C	N	O	S	0	0
			855	529	151	169	6		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	78	ALA	-	insertion	UNP G0S3J8

- Molecule 13 is a protein called Mediator of RNA polymerase II transcription subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	a	380	Total	C	N	O	S	0	0
			2941	1851	509	575	6		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
a	240	LEU	ASN	conflict	UNP G0SEC7
a	318	GLU	ASP	conflict	UNP G0SEC7

- Molecule 14 is a protein called Unknown peptide.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	b	21	Total	C	N	O	0	0
			105	63	21	21		

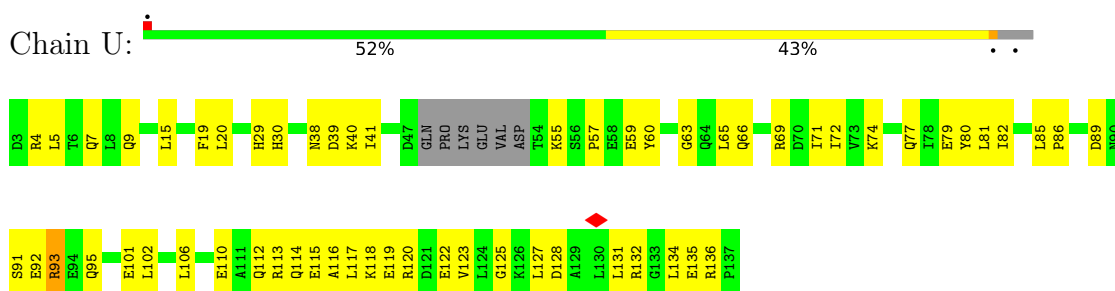
- Molecule 15 is a protein called HEAT.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	c	200	Total	C	N	O	0	0
			1000	600	200	200		

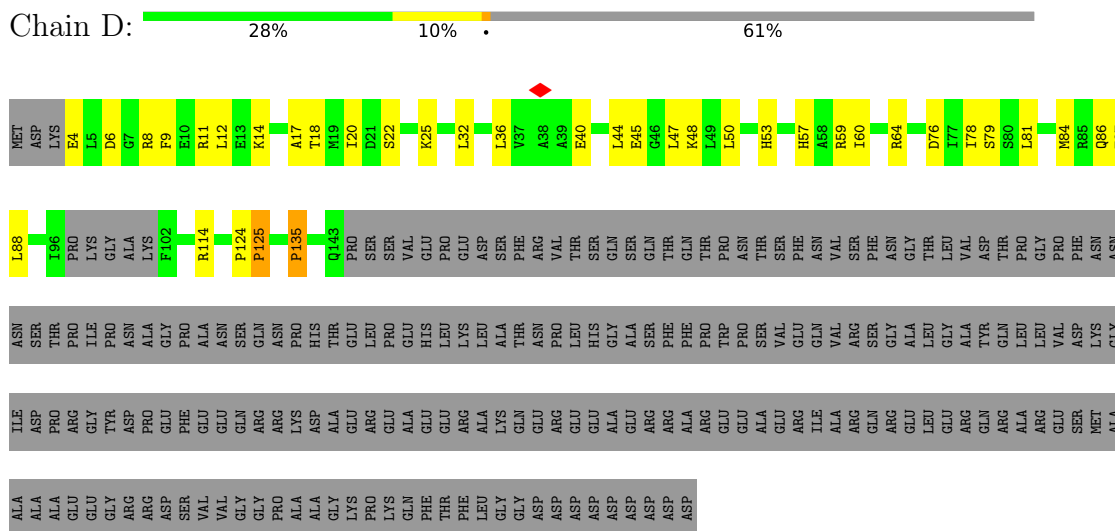
3 Residue-property plots

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.

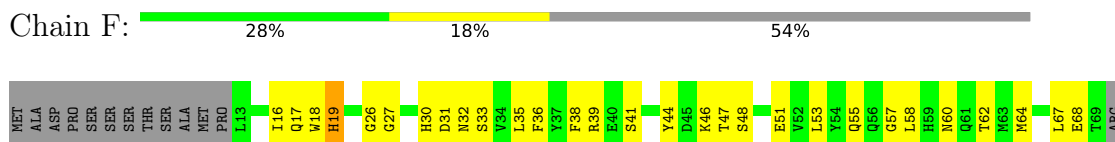
- Molecule 1: Mediator of RNA polymerase II transcription subunit 21

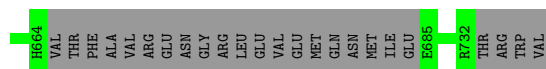


- Molecule 2: Mediator of RNA polymerase II transcription subunit 4



- Molecule 3: Mediator of RNA polymerase II transcription subunit 6





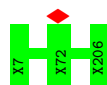
- Molecule 14: Unknown peptide

Chain b:  100%

There are no outlier residues recorded for this chain.

- Molecule 15: HEAT

Chain c:  100%



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	108348	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 ($e^-/\text{\AA}^2$)	60	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	1.272	Depositor
Minimum map value	-0.409	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.032	Depositor
Recommended contour level	0.17	Depositor
Map size (Å)	526.07996, 526.07996, 526.07996	wwPDB
Map dimensions	480, 480, 480	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.096, 1.096, 1.096	Depositor

5 Model quality

5.1 Standard geometry

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	U	0.25	0/1068	0.50	0/1437
2	D	0.25	0/1002	0.55	3/1353 (0.2%)
3	F	0.25	0/1228	0.49	0/1664
4	G	0.26	0/848	0.47	0/1142
5	I	0.27	0/469	0.85	10/649 (1.5%)
6	N	0.28	0/4169	0.61	14/5665 (0.2%)
7	Q	0.26	0/3628	0.53	3/4926 (0.1%)
8	Z	0.25	0/794	0.49	2/1079 (0.2%)
9	H	0.27	0/1168	0.49	0/1585
10	K	0.26	0/785	0.53	2/1066 (0.2%)
11	R	0.25	0/1259	0.50	0/1691
12	V	0.27	0/868	0.50	0/1176
13	a	0.25	0/2992	0.51	0/4048
All	All	0.26	0/20278	0.54	34/27481 (0.1%)

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
3	F	0	1
6	N	0	3
7	Q	0	2
10	K	0	1
11	R	0	1
13	a	0	3
All	All	0	11

There are no bond length outliers.

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	N	678	PRO	N-CA-CB	8.49	113.50	103.30
6	N	231	PRO	N-CA-CB	6.70	111.34	103.30
6	N	789	PRO	N-CA-CB	6.65	111.28	103.30
6	N	779	PRO	N-CA-CB	6.51	111.12	103.30
6	N	808	PRO	N-CA-CB	6.40	110.98	103.30

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	F	19	HIS	Peptide
6	N	298	TRP	Peptide
6	N	435	SER	Peptide
6	N	497	LEU	Peptide
7	Q	46	GLN	Peptide

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	U	1055	0	1051	57	0
2	D	994	0	946	27	0
3	F	1202	0	1094	54	0
4	G	836	0	730	21	0
5	I	470	0	248	5	0
6	N	4081	0	3631	135	0
7	Q	3571	0	3475	106	0
8	Z	775	0	736	18	0
9	H	1148	0	1097	35	0
10	K	777	0	749	27	0
11	R	1255	0	921	19	0
12	V	855	0	852	34	0
13	a	2941	0	2905	0	0
14	b	105	0	23	0	0
15	c	1000	0	213	0	0
All	All	21065	0	18671	489	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 14.

The worst 5 of 489 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:N:609:SER:HG	6:N:636:THR:N	1.65	0.95
3:F:98:THR:N	9:H:87:THR:HG1	1.66	0.92
6:N:284:VAL:HG12	6:N:299:PHE:HB2	1.52	0.89
7:Q:256:ARG:HG3	7:Q:328:VAL:HA	1.54	0.89
7:Q:491:LEU:HD23	7:Q:492:HIS:H	1.40	0.87

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	U	125/135 (93%)	108 (86%)	17 (14%)	0	100	100
2	D	128/341 (38%)	113 (88%)	13 (10%)	2 (2%)	9	45
3	F	146/337 (43%)	130 (89%)	16 (11%)	0	100	100
4	G	114/246 (46%)	103 (90%)	11 (10%)	0	100	100
5	I	94/179 (52%)	81 (86%)	9 (10%)	4 (4%)	2	25
6	N	528/1171 (45%)	415 (79%)	98 (19%)	15 (3%)	5	33
7	Q	460/634 (73%)	375 (82%)	84 (18%)	1 (0%)	47	80
8	Z	94/124 (76%)	83 (88%)	11 (12%)	0	100	100
9	H	142/290 (49%)	128 (90%)	14 (10%)	0	100	100
10	K	103/238 (43%)	90 (87%)	12 (12%)	1 (1%)	15	54
11	R	201/340 (59%)	163 (81%)	38 (19%)	0	100	100
12	V	110/169 (65%)	93 (84%)	17 (16%)	0	100	100
13	a	362/736 (49%)	287 (79%)	74 (20%)	1 (0%)	41	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2607/4940 (53%)	2169 (83%)	414 (16%)	24 (1%)	21	56

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	135	PRO
5	I	27	ASN
5	I	74	PRO
6	N	110	PRO
6	N	230	PRO

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	U	115/121 (95%)	114 (99%)	1 (1%)	78	87
2	D	98/286 (34%)	98 (100%)	0	100	100
3	F	122/287 (42%)	122 (100%)	0	100	100
4	G	71/208 (34%)	69 (97%)	2 (3%)	43	65
6	N	377/1026 (37%)	370 (98%)	7 (2%)	57	74
7	Q	370/539 (69%)	366 (99%)	4 (1%)	73	84
8	Z	74/110 (67%)	74 (100%)	0	100	100
9	H	121/237 (51%)	120 (99%)	1 (1%)	81	89
10	K	76/192 (40%)	76 (100%)	0	100	100
11	R	65/302 (22%)	59 (91%)	6 (9%)	9	31
12	V	90/133 (68%)	89 (99%)	1 (1%)	73	84
13	a	319/635 (50%)	319 (100%)	0	100	100
All	All	1898/4076 (47%)	1876 (99%)	22 (1%)	72	83

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	H	74	ARG
11	R	52	PRO
11	R	27	THR
11	R	95	PRO
6	N	462	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
7	Q	90	GLN
13	a	622	GLN
7	Q	254	GLN
10	K	158	GLN
7	Q	251	GLN

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](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
7	Q	1
2	D	1
6	N	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	Q	66:ALA	C	75:LYS	N	44.21
1	D	122:PRO	C	124:PRO	N	22.98
1	N	715:ASP	C	741:ALA	N	11.22

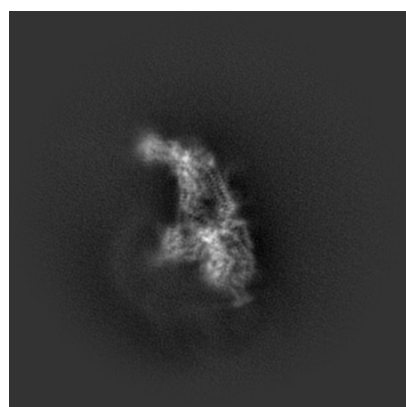
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22284. These allow visual inspection of the internal detail of the map and identification of artifacts.

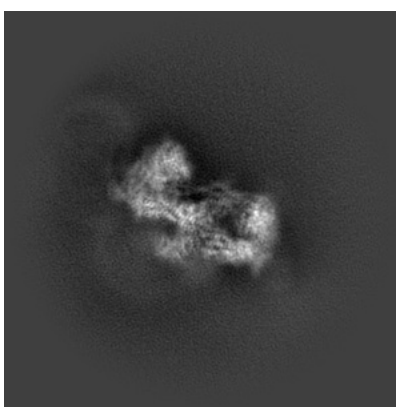
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

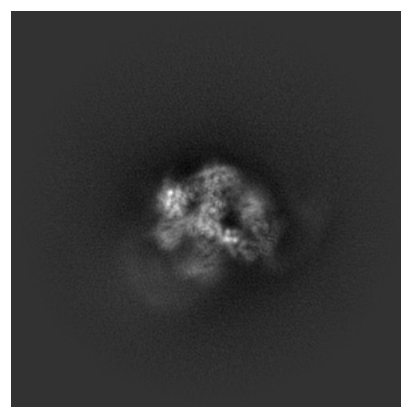
6.1.1 Primary map



X



Y

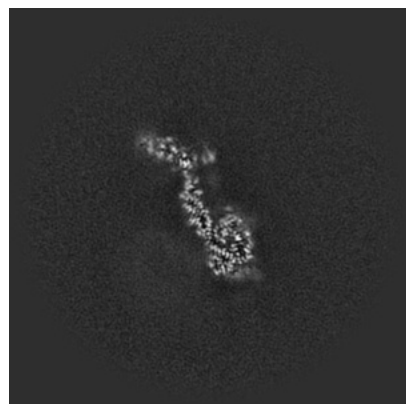


Z

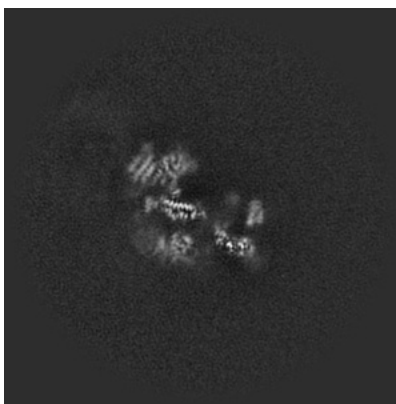
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

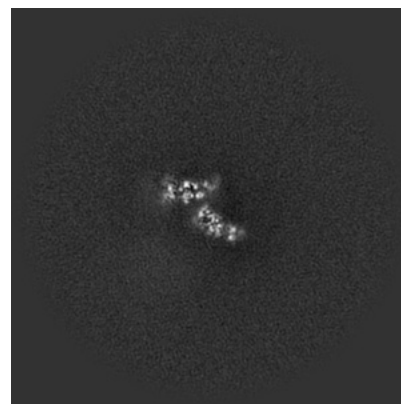
6.2.1 Primary map



X Index: 240



Y Index: 240

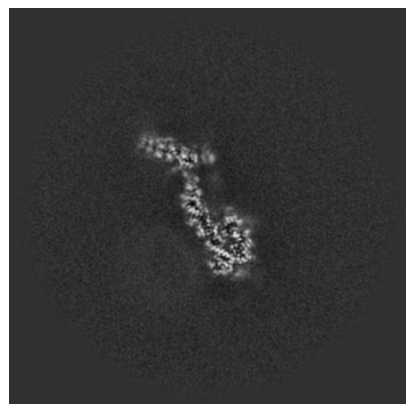


Z Index: 240

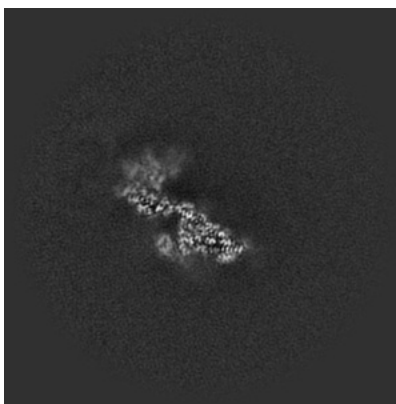
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

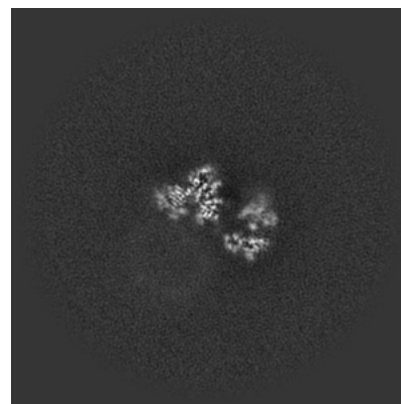
6.3.1 Primary map



X Index: 238



Y Index: 257



Z Index: 213

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

6.4 Orthogonal surface views [i](#)

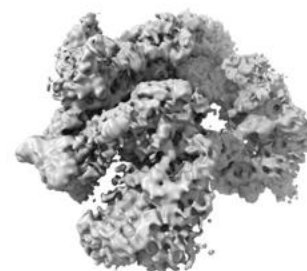
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.17. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

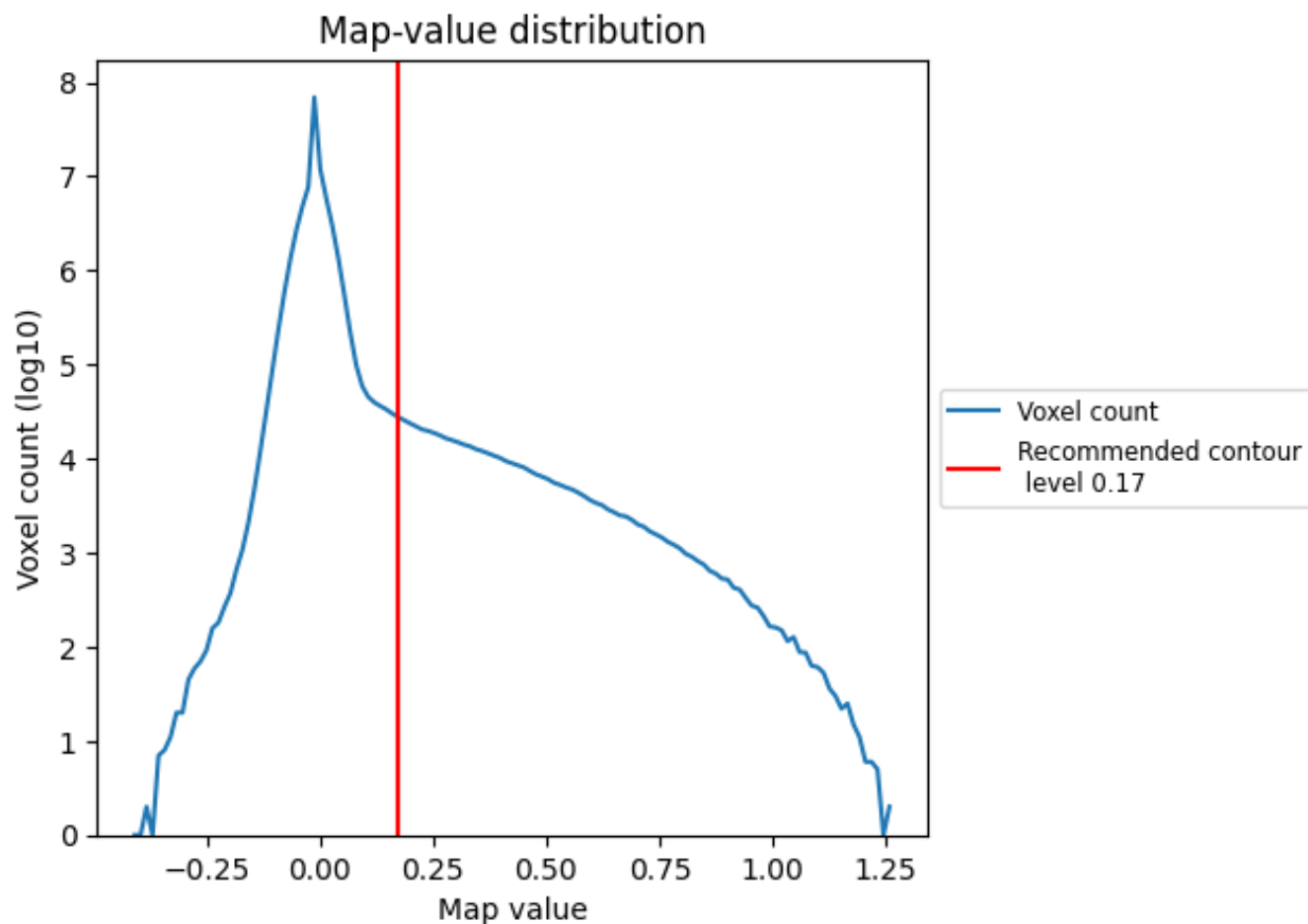
6.5 Mask visualisation

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

7 Map analysis [i](#)

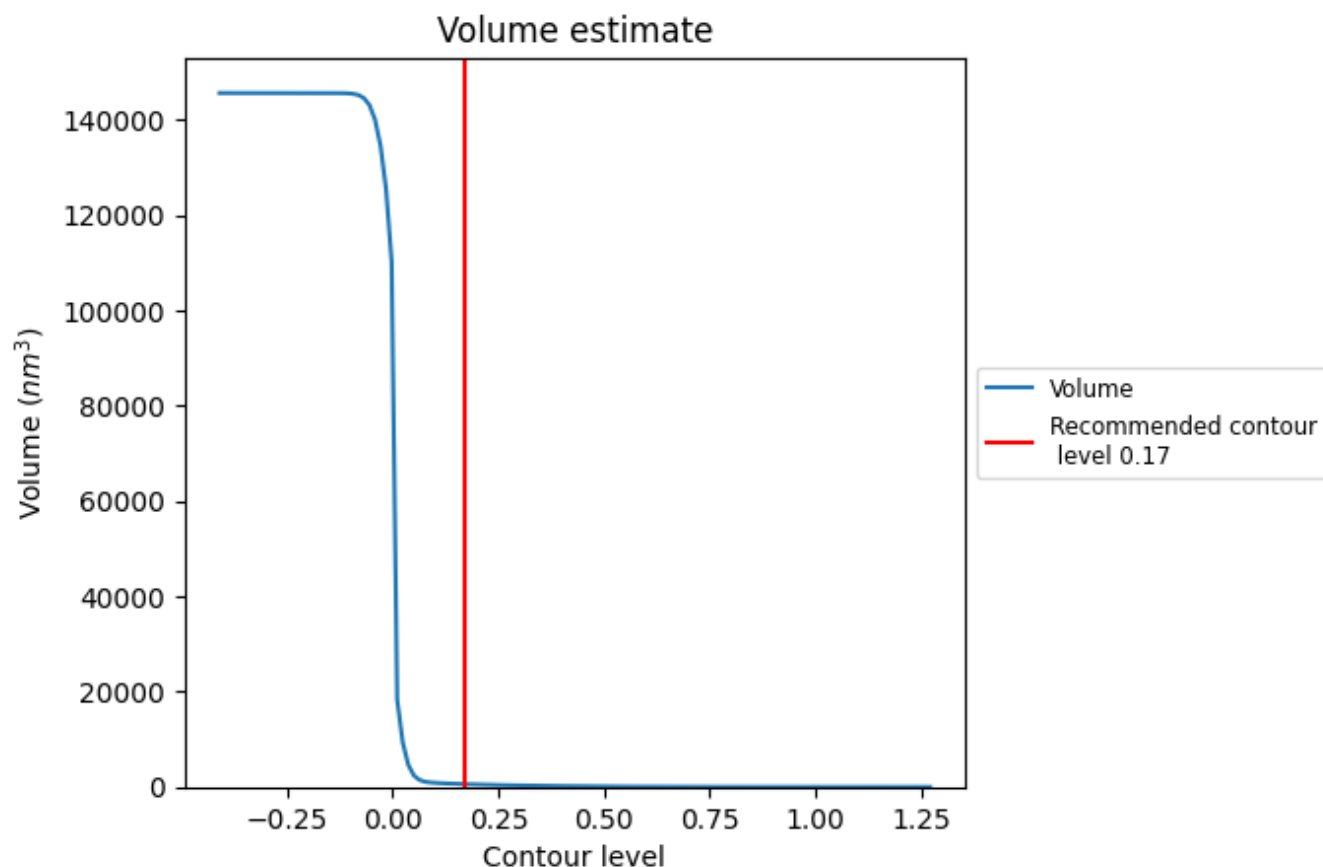
This section contains the results of statistical analysis of the map.

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.

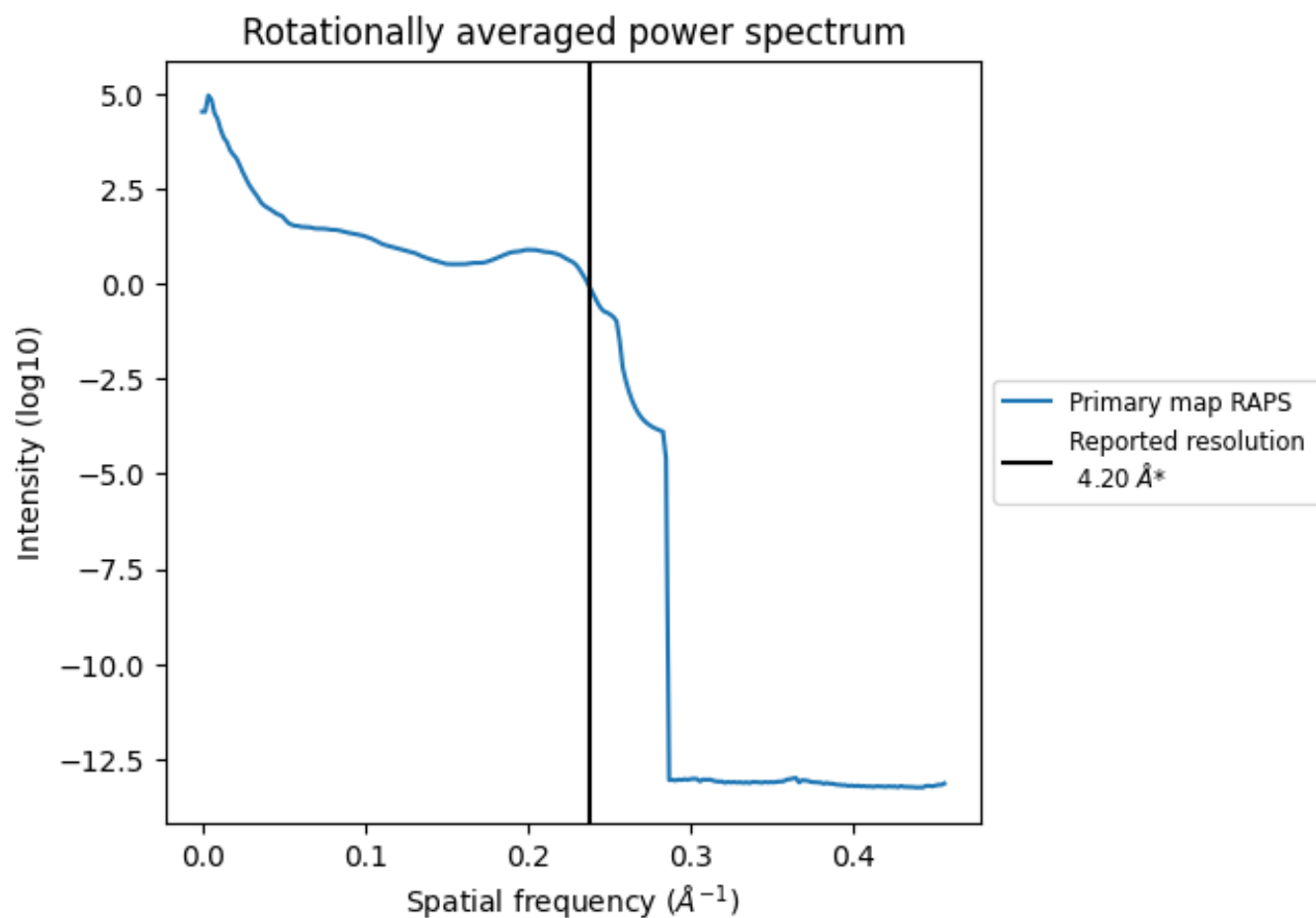
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 591 nm³; this corresponds to an approximate mass of 534 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 ⓘ

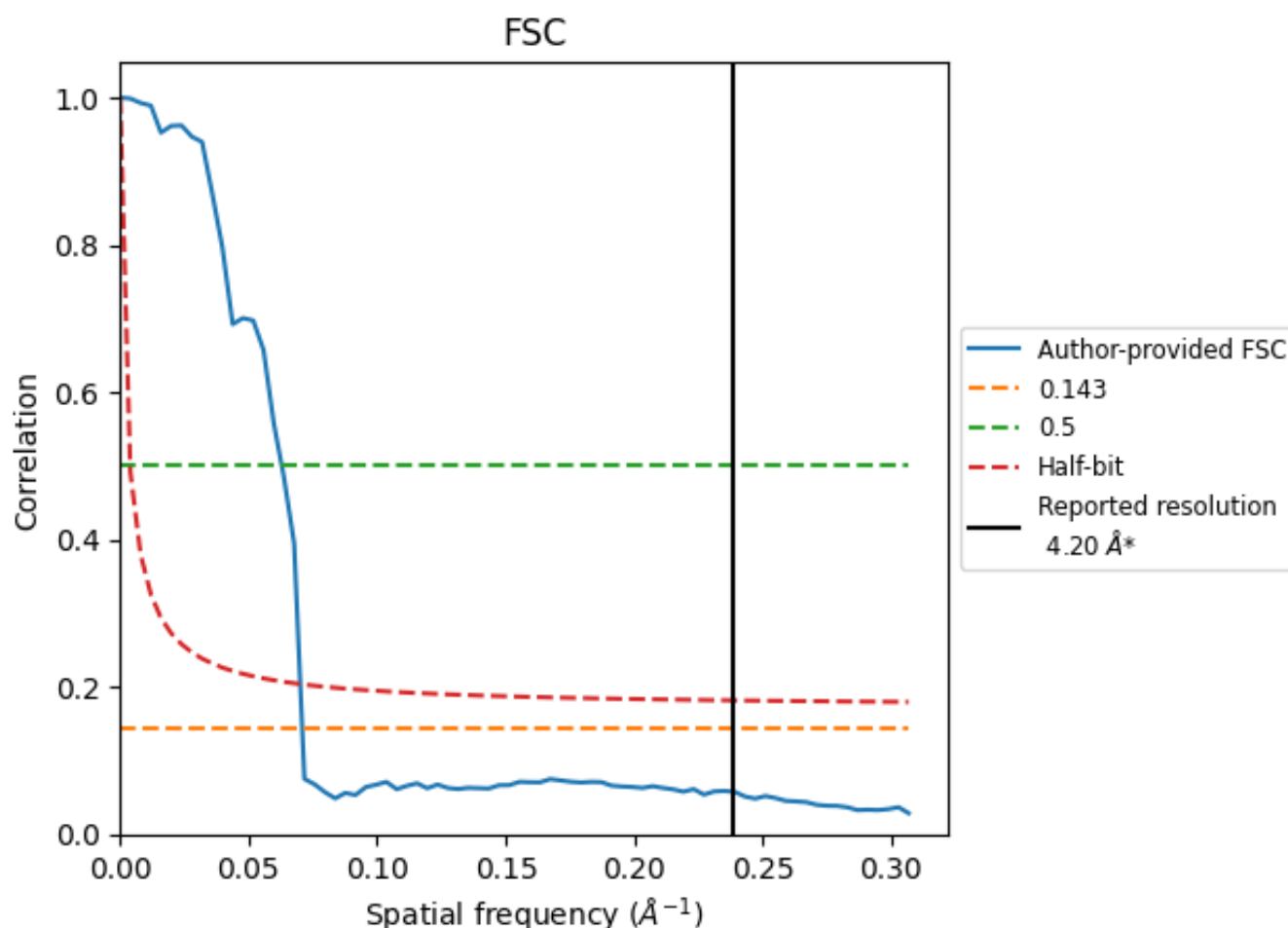


*Reported resolution corresponds to spatial frequency of 0.238 Å⁻¹

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 0.238 Å⁻¹

8.2 Resolution estimates [i](#)

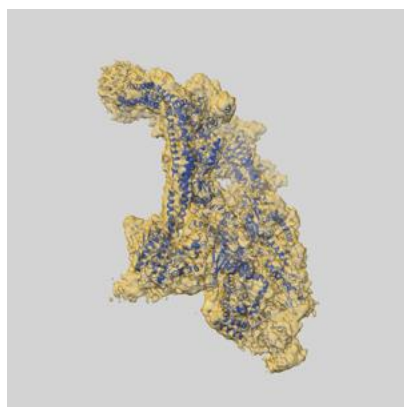
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.20	-	-
Author-provided FSC curve	14.10	15.90	14.27
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from author-provided FSC intersecting FSC 0.143 CUT-OFF 14.10 differs from the reported value 4.2 by more than 10 %

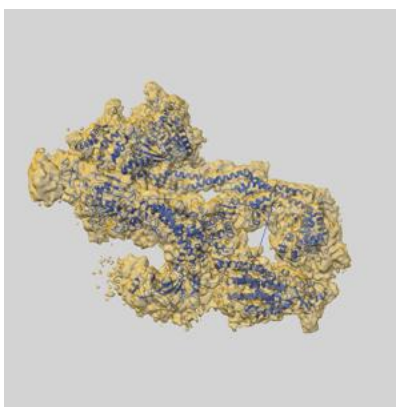
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-22284 and PDB model 6XP5. Per-residue inclusion information can be found in section [3](#) on page [7](#).

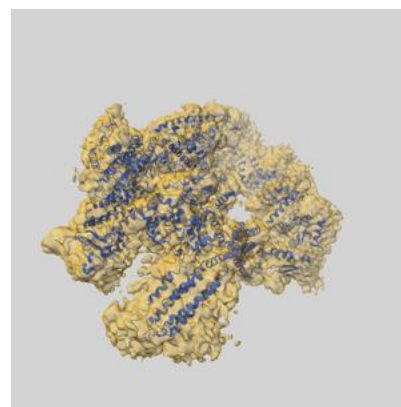
9.1 Map-model overlay [i](#)



X



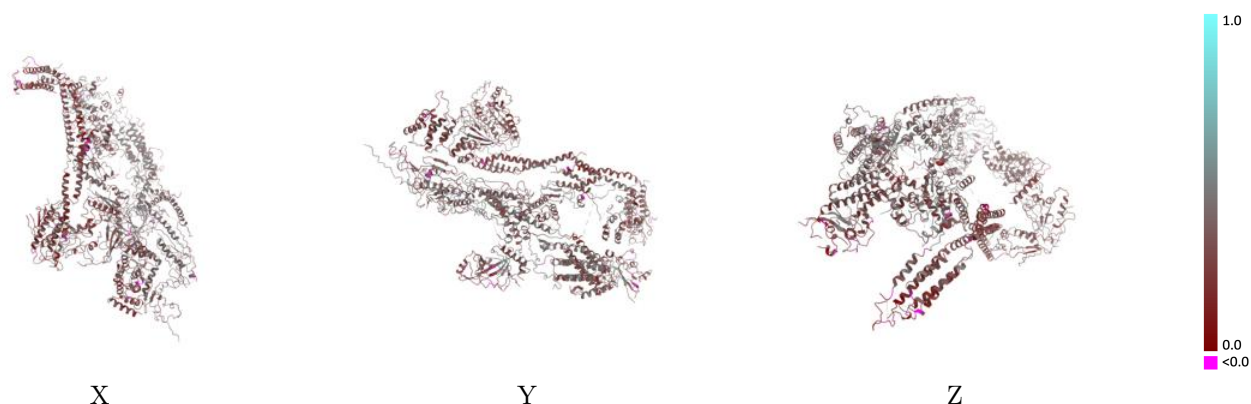
Y



Z

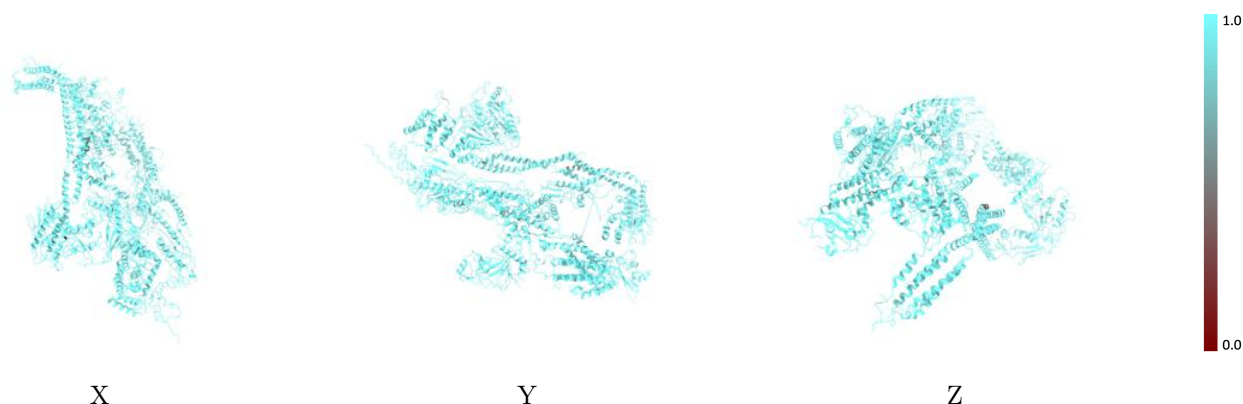
The images above show the 3D surface view of the map at the recommended contour level 0.17 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



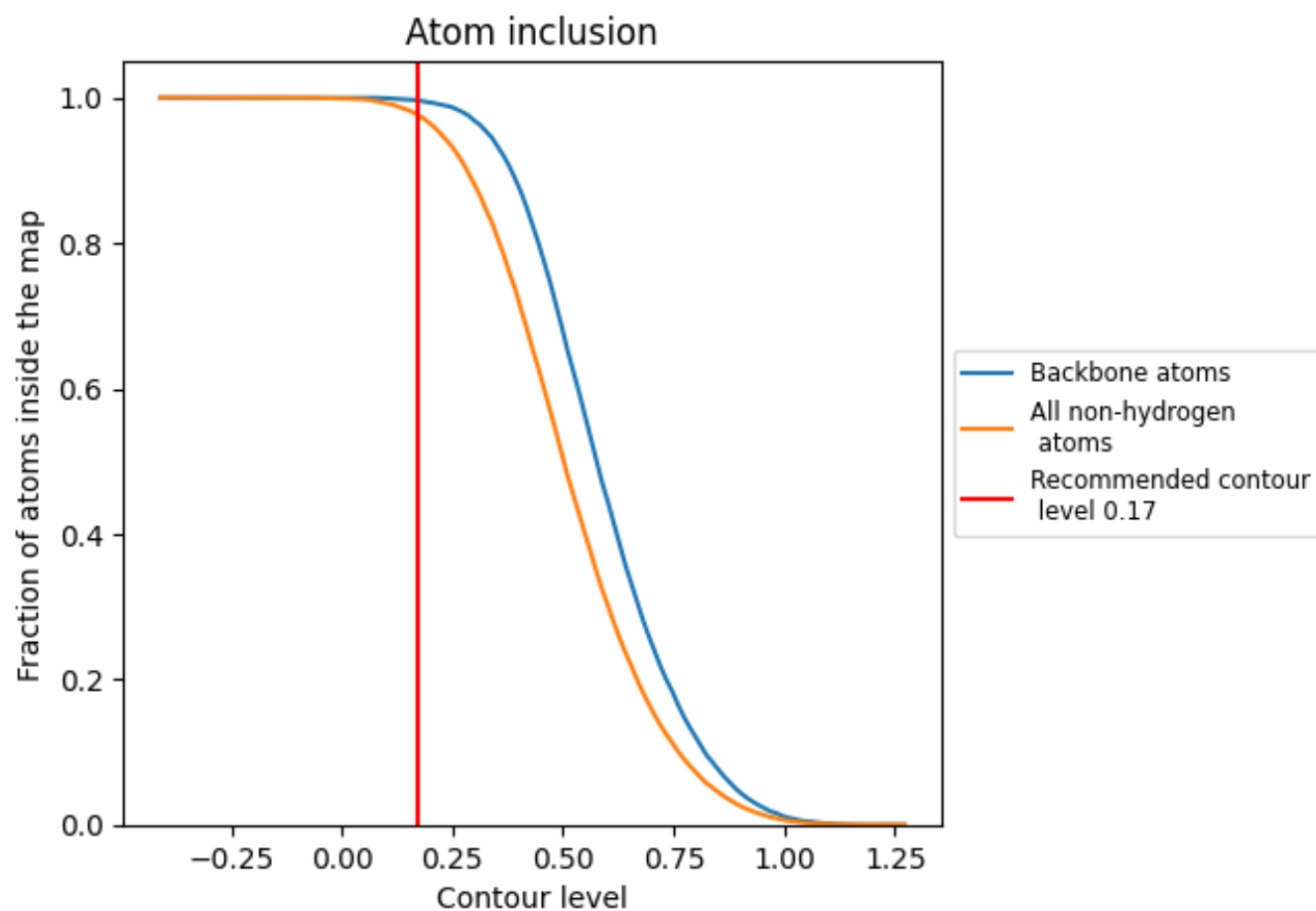
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.17).

































9.4 Atom inclusion ⓘ



At the recommended contour level, 100% 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 (0.17) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9771	 0.3370
D	 0.9725	 0.3070
F	 0.9907	 0.3010
G	 0.9795	 0.3030
H	 0.9867	 0.3260
I	 0.9957	 0.2910
K	 0.9922	 0.3910
N	 0.9700	 0.3630
Q	 0.9788	 0.3830
R	 0.9888	 0.3200
U	 0.9381	 0.2020
V	 0.9544	 0.3750
Z	 0.9828	 0.3680
a	 0.9765	 0.3250
b	 1.0000	 0.2840
c	 0.9940	 0.3180

