



wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 12:06 PM EST

PDB ID : 7UIM
EMDB ID : EMD-26549
Title : CryoEM Structure of an Group II Intron Retroelement (apo-complex)
Authors : Chung, K.; Xu, L.
Deposited on : 2022-03-29
Resolution : 3.10 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 ⓘ 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:

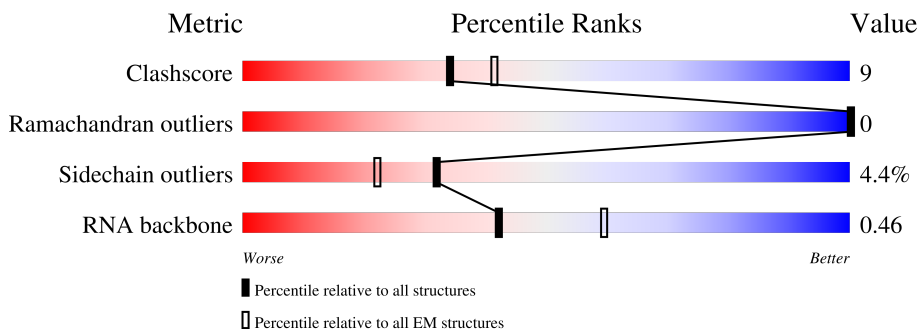
EMDB validation analysis : 0.0.1.dev43
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

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 3.10 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)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	B	638	
2	D	427	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 14772 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 RNA chain called E.r IIC Intron.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	B	541	11624	5197	2167	3724	536	0	0

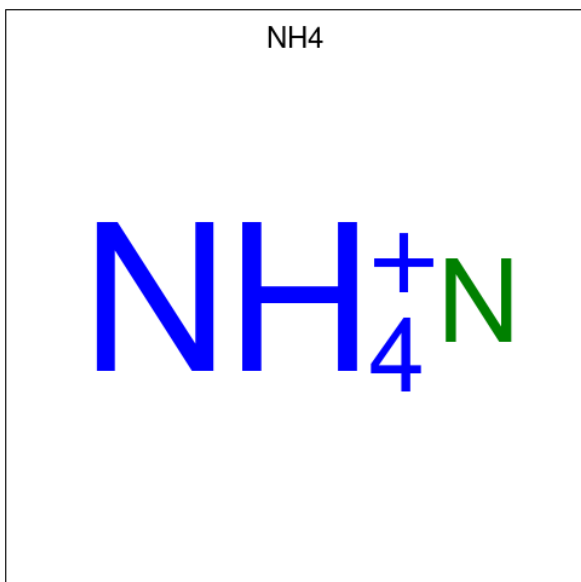
- Molecule 2 is a protein called Group II intron reverse transcriptase/maturase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	D	387	3133	1987	567	558	21	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
3	B	13	13	13	0

- Molecule 4 is AMMONIUM ION (three-letter code: NH4) (formula: H₄N) (labeled as "Ligand of Interest" by depositor).

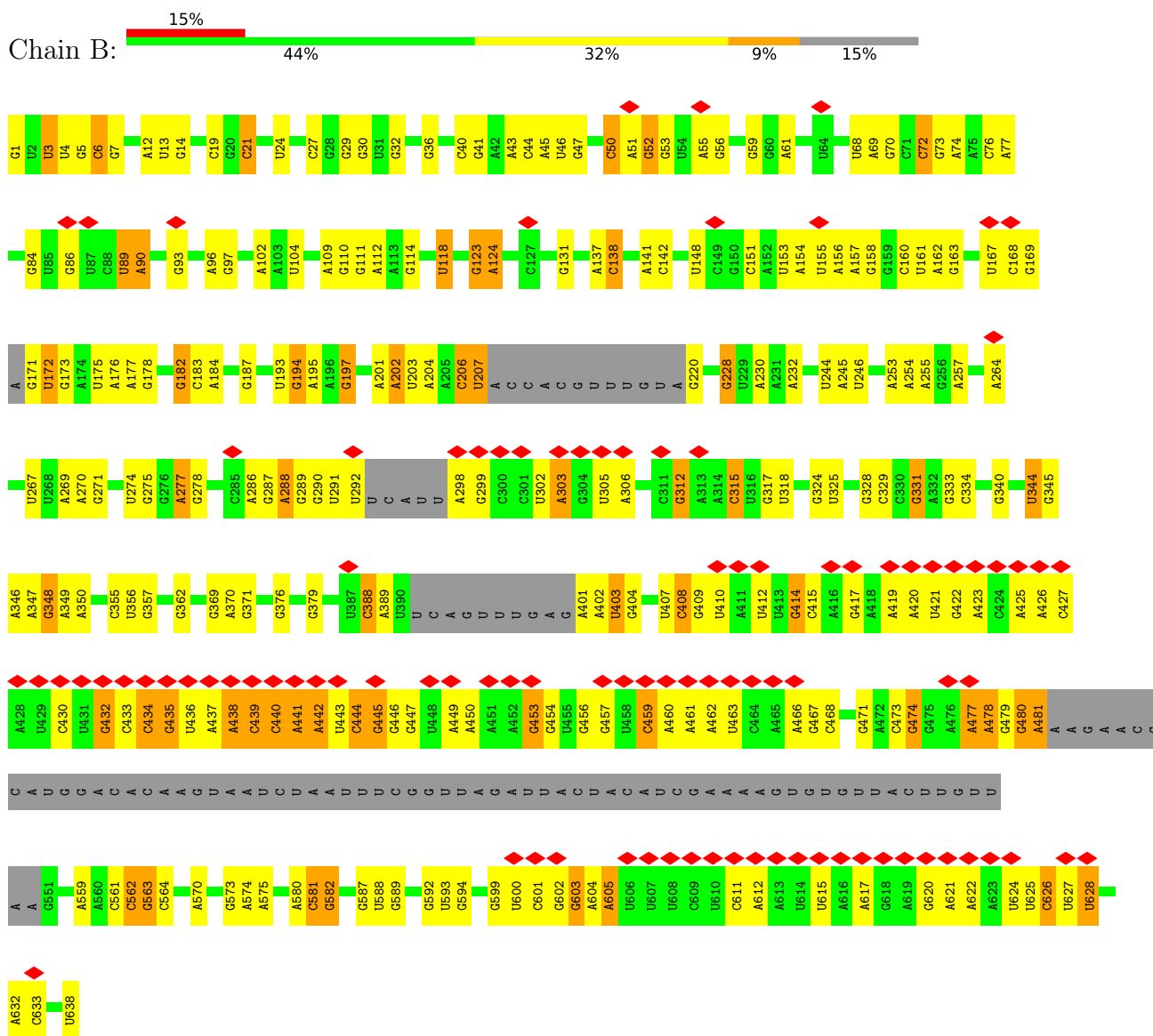


Mol	Chain	Residues	Atoms	AltConf
4	B	1	Total N 2 2	0
4	B	1	Total N 2 2	0

3 Residue-property plots [i](#)

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

- Molecule 1: E.r IIC Intron



- Molecule 2: Group II intron reverse transcriptase/maturase



MET	ASP	THR	SER	H5	ARG	ARG	L6	M7	E8	Q9	I10	L11	S12	S13	D14	H15	L16	H17	R18	A19	Y20	L21	Q22	V23	V24	R25	M26	K27	G28	A29	E30	G31	V32	D33	G34	M35	K36	Y37	T38	E39	L40	K41	E42	H43	L44	A45	K46	N47	G48	E49	T50	I51	K52	L55	R56	T57	R58	H59	Y60	K61
P62	Q63	P64	ALA	ARG	ARG	VAL	VAL	GLU	ILE	PRO	PRO	PRO	ASP	GLY	GLY	VAL	ARG	ASN	LEU	G81	V82	P83	T84	V85	T86	D87	R88	F89	I90	Q91	A95	Q96	V97	L98	T99	P100	T101	Y102	E103	E104	Q105	F106	H107	D108	H109	S110	Y111	G114	P115	M116	R117	C118	A119	Q120	Q121	A122	I123	L124		
T125	A126	L127	N128	I129	M130	N131	D132	D135	M136	I137	V138	D139	I140	D141	L142	E143	K144	F145	F146	D147	T148	V149	M150	H151	D152	K153	L154	M155	G159	R160	T161	I162	K163	S169	I170	V171	R172	K173	Y174	L175	V176	S177	GLY	ILE	ILE	ILE	ASP	ASP	GLU	TYR	GLU	ASP	SER	ILE	VAL	GLY				
THR	PRO	GLN	G195	G196	M197	L198	S199	L200	L201	L202	A203	N204	I205	M206	L207	N208	E209	L210	D211	K212	E213	M214	E215	K216	R217	G218	L219	N220	F221	V222	R223	Y224	A225	D226	D227	C228	I229	L230	M231	V232	G233	S234	E235	M236	S237	A238	M239	R240	V241	M242	R243	N244	I245	S246	R247	F248	I249	E250	E251	
K252	L253	G254	L255	K256	V257	N258	M259	T260	K261	S262	K263	V264	D265	R266	G269	L270	K271	Y272	L273	G274	F277	Y278	F279	D280	P281	R282	A283	H284	Q285	F286	K287	A288	V295	A296	K297	F298	K299	K300	R301	M302	K303	E304	L305	T306	S309	W310	G311	N314	S315	Y316	E319	K320	L321							
N322	Q323	L324	I325	R326	G327	W328	I329	N330	Y331	F332	K333	I334	G335	S336	M337	K338	T339	L340	C341	K342	E343	L344	D345	I348	R349	Y350	R351	L352	R353	W357	K358	Q359	K368	V371	A296	L373	G374	I375	D376	R377	R381	R382	V383	A384	Y385	T386	G387	K388	R389	I390	A391	Y392	V393	C394						
N395	K396	G397	A398	V399	M400	V401	A402	I403	S404	N405	K406	R407	L408	A409	S410	F411	G412	L413	I414	S415	M416	L417	D418	Y419	Y420	I421	A422	K423	C424	VAL	THR	CYS																												

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	434467	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 ($e^-/\text{\AA}^2$)	50.3	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.425	Depositor
Minimum map value	-0.145	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.015	Depositor
Recommended contour level	0.12	Depositor
Map size (Å)	319.488, 319.488, 319.488	wwPDB
Map dimensions	384, 384, 384	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.832, 0.832, 0.832	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: NH4, 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	B	0.50	0/13028	0.98	27/20316 (0.1%)
2	D	0.34	0/3186	0.69	4/4273 (0.1%)
All	All	0.47	0/16214	0.94	31/24589 (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
2	D	0	1

There are no bond length outliers.

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	138	C	N1-C2-O2	7.84	123.60	118.90
1	B	27	C	C5-C6-N1	7.14	124.57	121.00
1	B	110	G	C4-N9-C1'	7.09	135.71	126.50
1	B	3	U	N3-C2-O2	-6.88	117.38	122.20
1	B	3	U	N1-C2-O2	6.66	127.47	122.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	281	PRO	Peptide

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	11624	0	5848	120	0
2	D	3133	0	3219	67	0
3	B	13	0	0	0	0
4	B	2	0	0	0	0
All	All	14772	0	9067	184	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 9.

The worst 5 of 184 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)
1:B:401:A:H3'	1:B:480:G:H1	1.21	1.03
1:B:303:A:H62	1:B:312:G:H21	1.21	0.88
1:B:444:C:H2'	1:B:445:G:C8	2.08	0.87
1:B:434:C:H42	1:B:445:G:H1	1.21	0.84
1:B:477:A:H2'	1:B:478:A:O4'	1.90	0.71

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
2	D	381/427 (89%)	367 (96%)	14 (4%)	0	100 100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	D	341/376 (91%)	326 (96%)	15 (4%)	28 61

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	247	ARG
2	D	419	TYR
2	D	259	MET
2	D	422	GLU
2	D	350	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	535/638 (83%)	162 (30%)	7 (1%)

5 of 162 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	3	U
1	B	4	U
1	B	5	G
1	B	6	C
1	B	12	A

5 of 7 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	B	206	C
1	B	333	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	480	G
1	B	438	A
1	B	123	G

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 15 ligands modelled in this entry, 13 are monoatomic and 2 are modelled with single atom - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

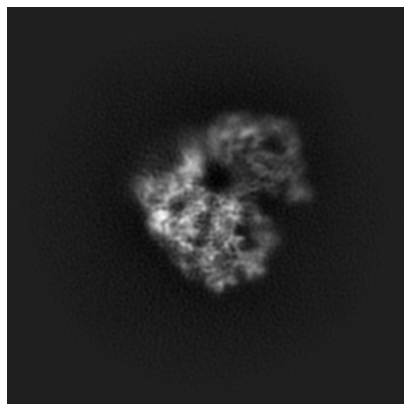
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-26549. These allow visual inspection of the internal detail of the map and identification of artifacts.

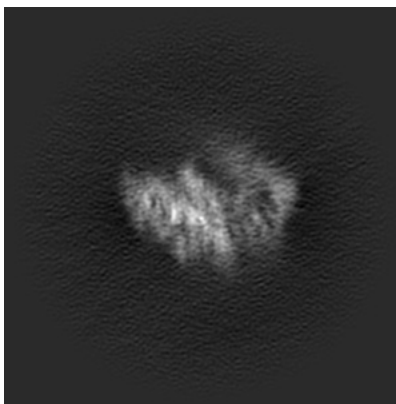
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

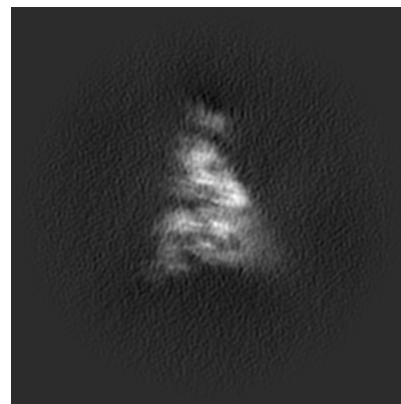
6.1.1 Primary map



X

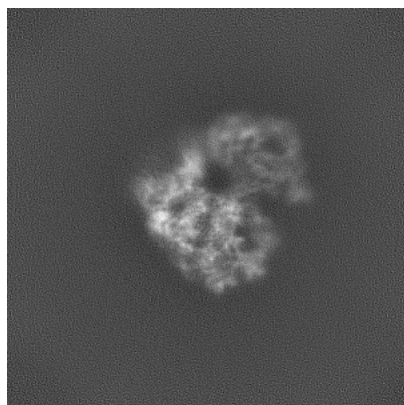


Y

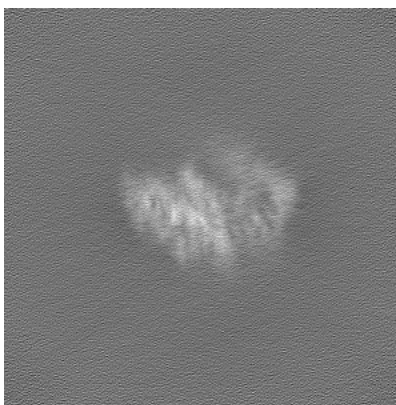


Z

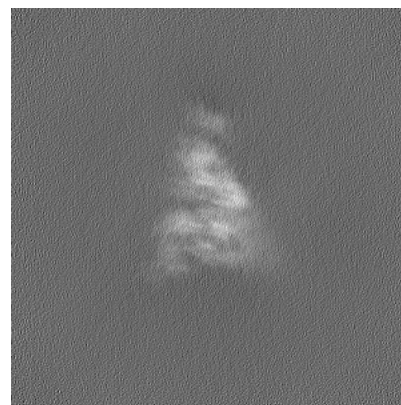
6.1.2 Raw map



X



Y

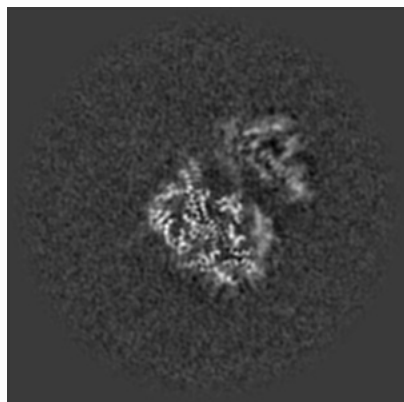


Z

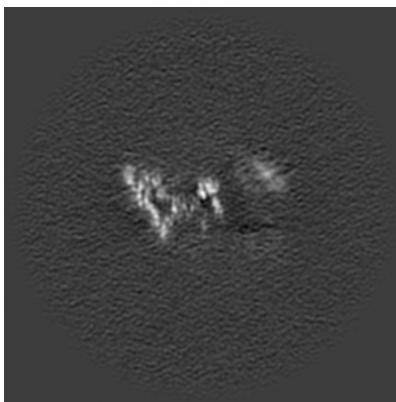
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

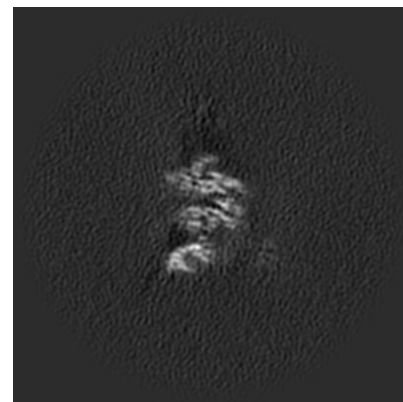
6.2.1 Primary map



X Index: 192

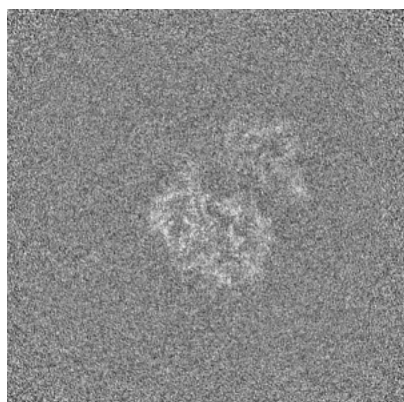


Y Index: 192

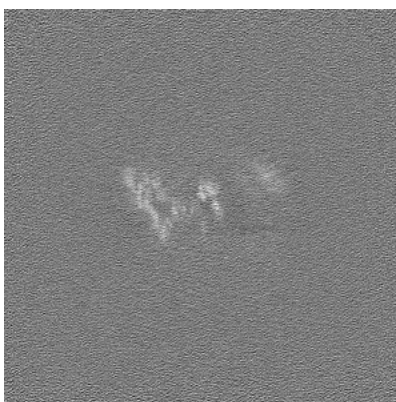


Z Index: 192

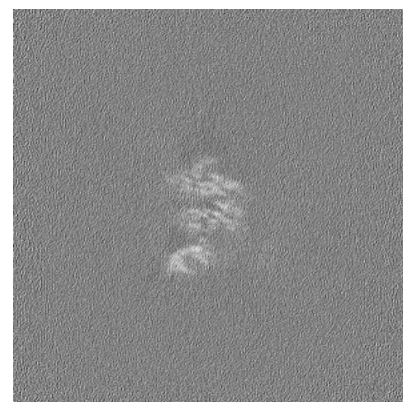
6.2.2 Raw map



X Index: 192



Y Index: 192

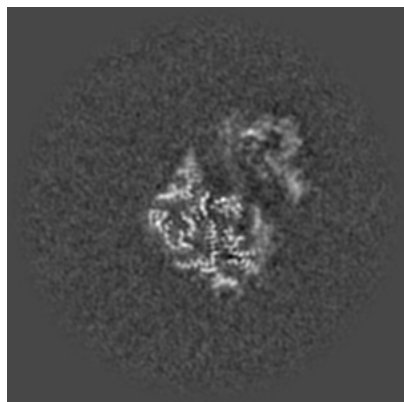


Z Index: 192

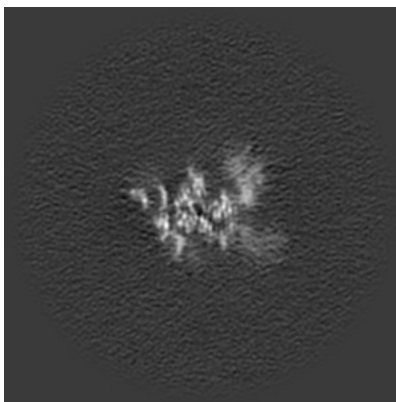
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

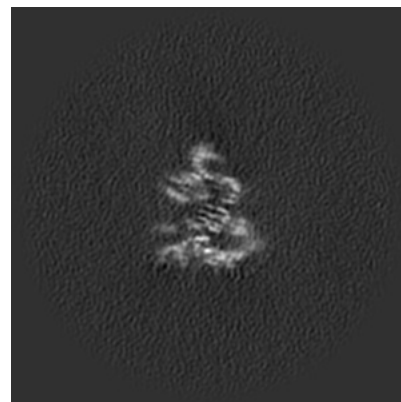
6.3.1 Primary map



X Index: 195

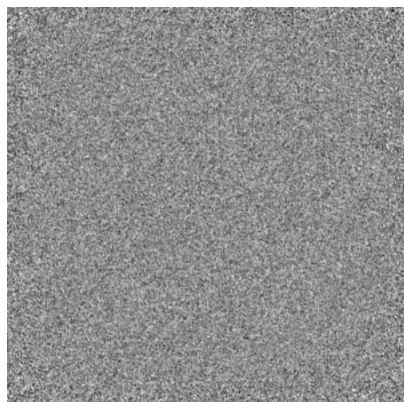


Y Index: 175

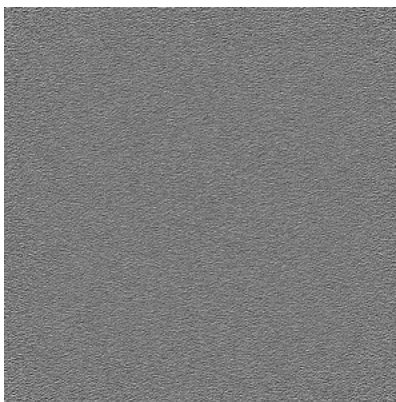


Z Index: 181

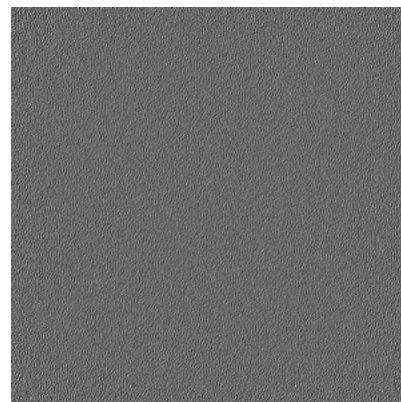
6.3.2 Raw map



X Index: 0



Y Index: 0

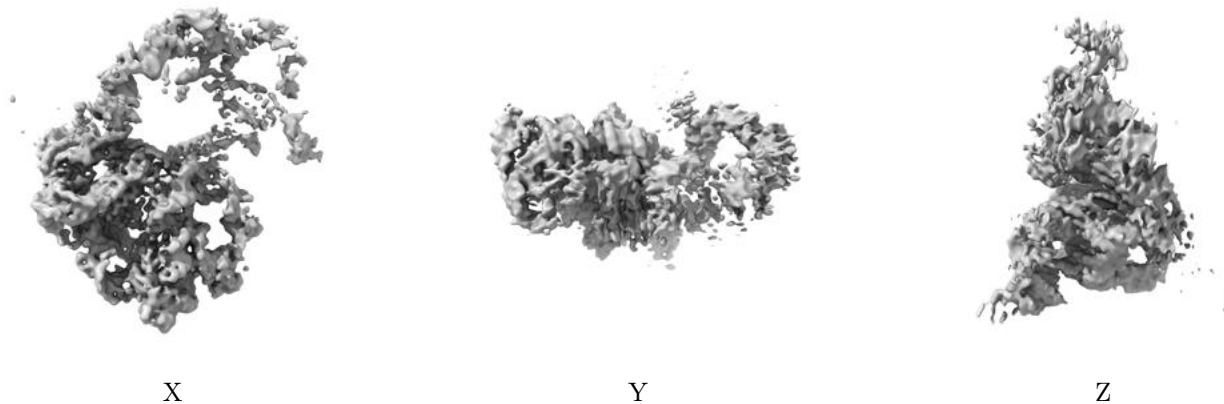


Z Index: 0

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

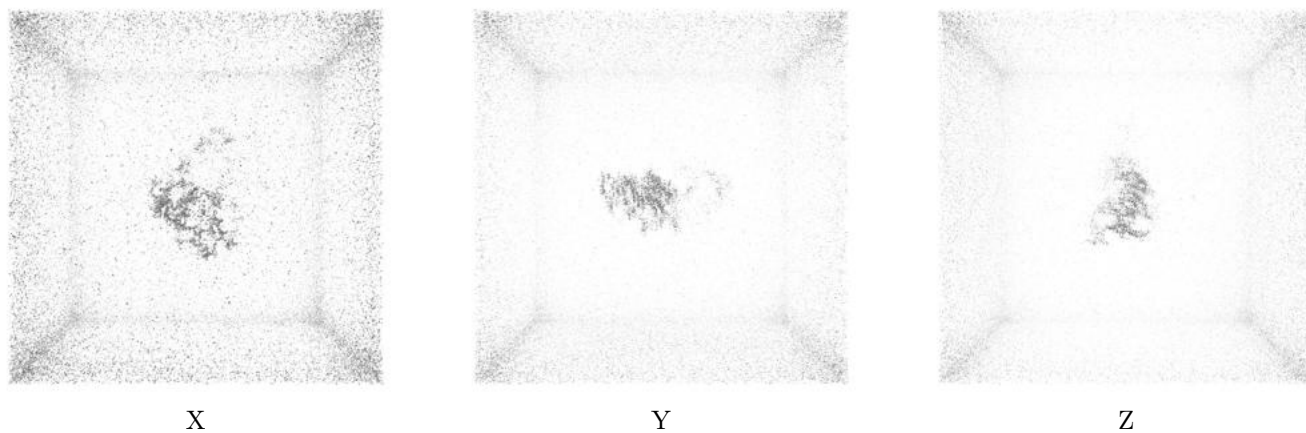
6.4 Orthogonal surface views [i](#)

6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.12. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.4.2 Raw map



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.

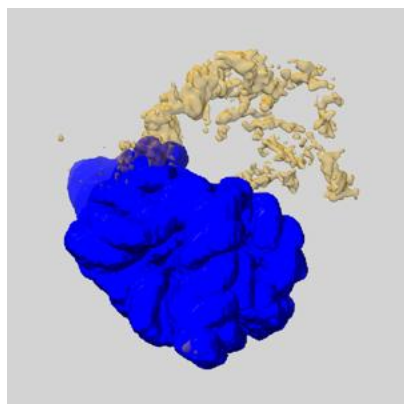
6.5 Mask visualisation [i](#)

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

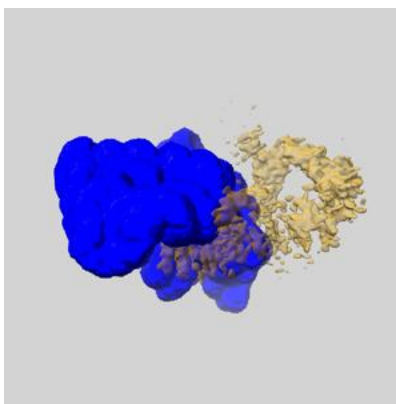
A mask typically either:

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

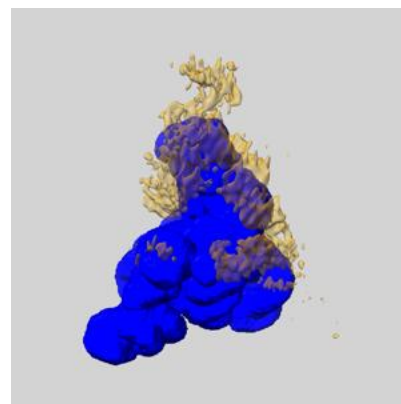
6.5.1 emd_26549_msk_1.map [i](#)



X

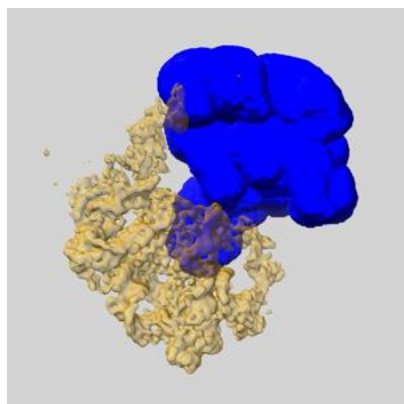


Y

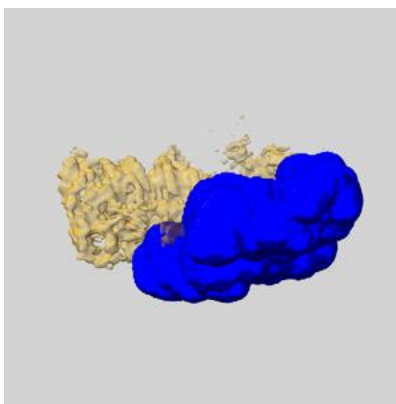


Z

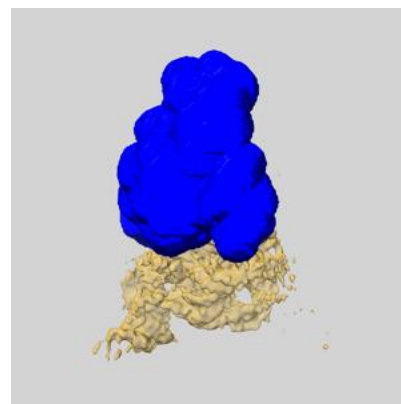
6.5.2 emd_26549_msk_2.map [i](#)



X



Y

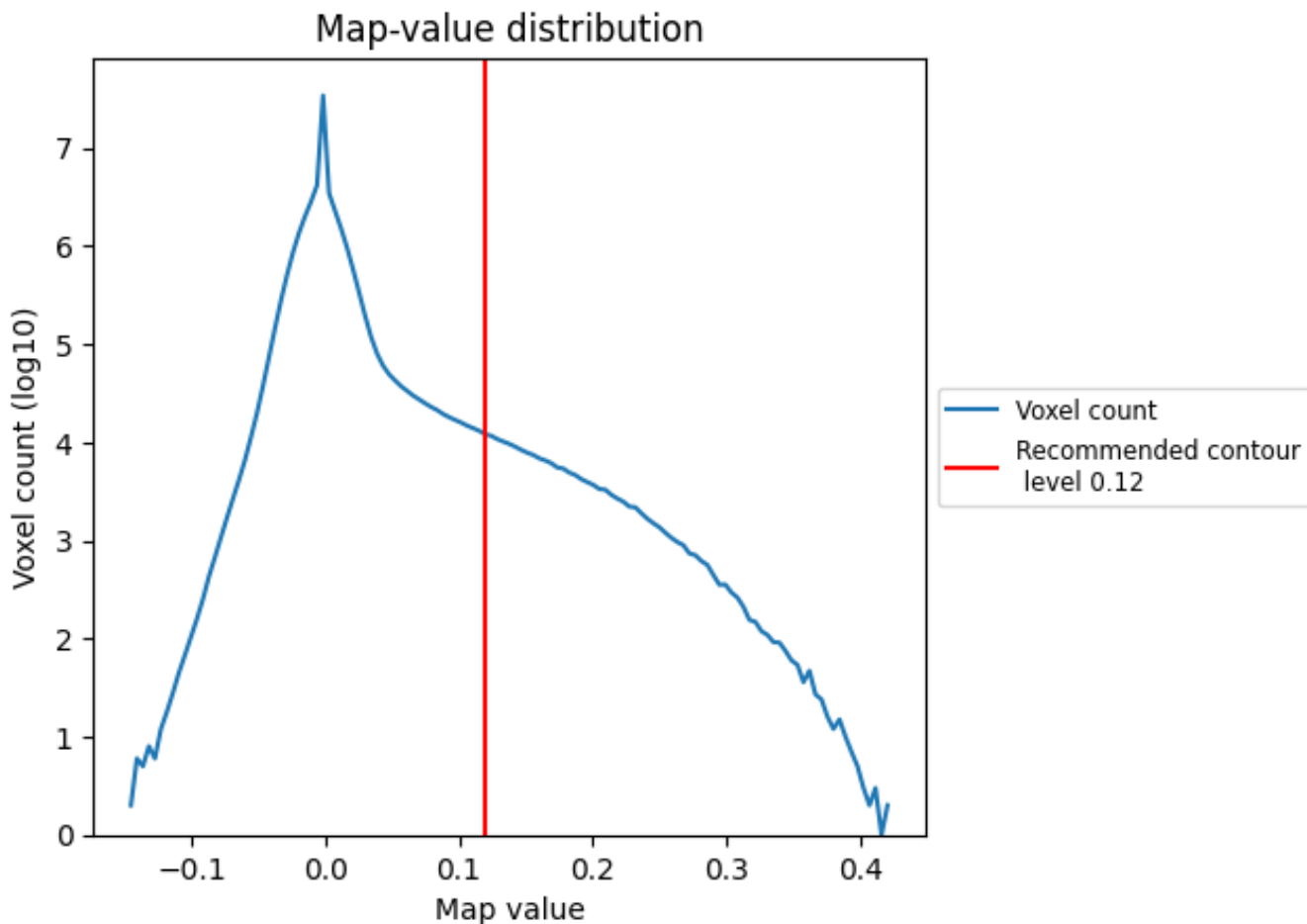


Z

7 Map analysis [i](#)

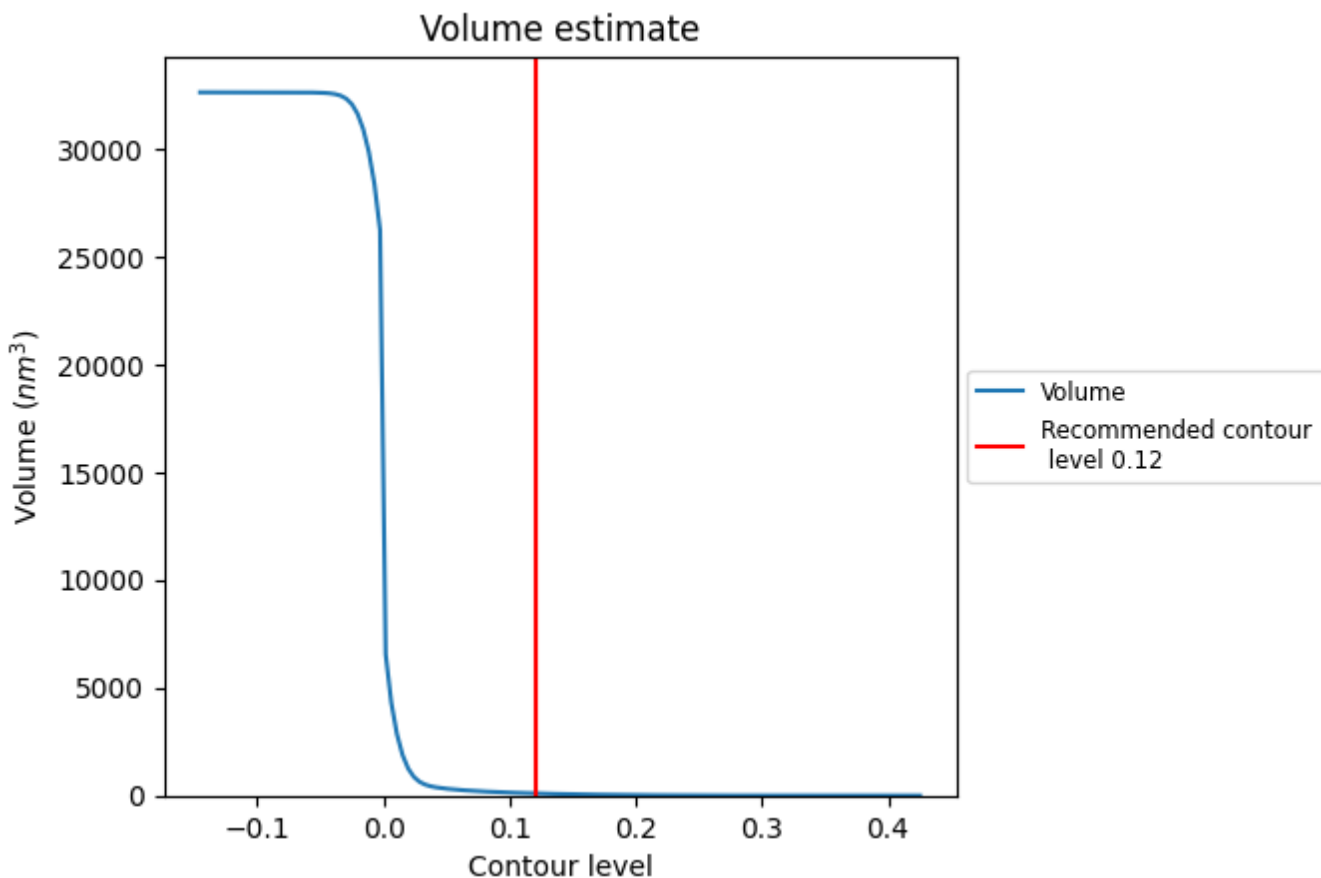
This section contains the results of statistical analysis of the map.

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.

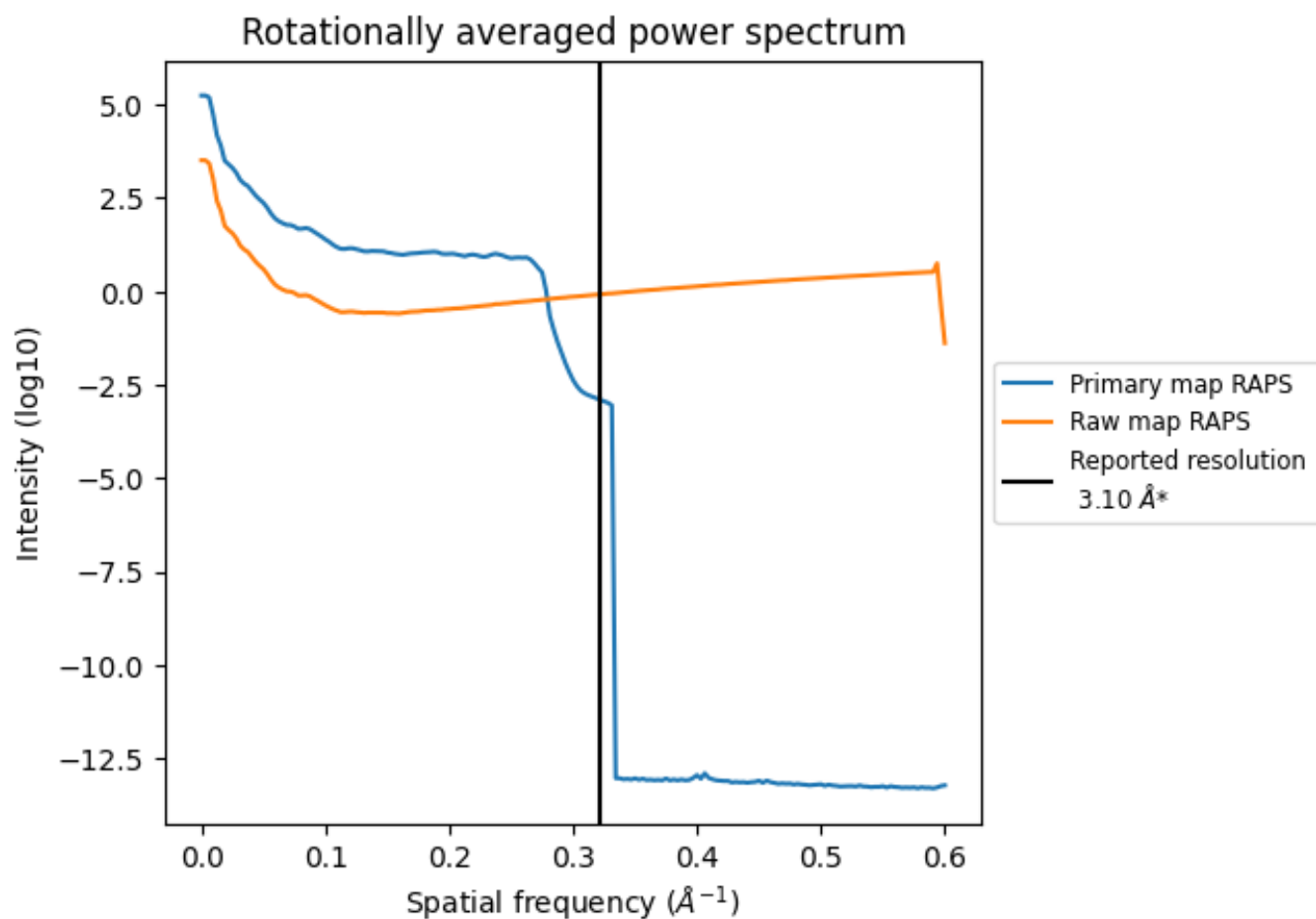
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 100 nm³; this corresponds to an approximate mass of 90 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](#)

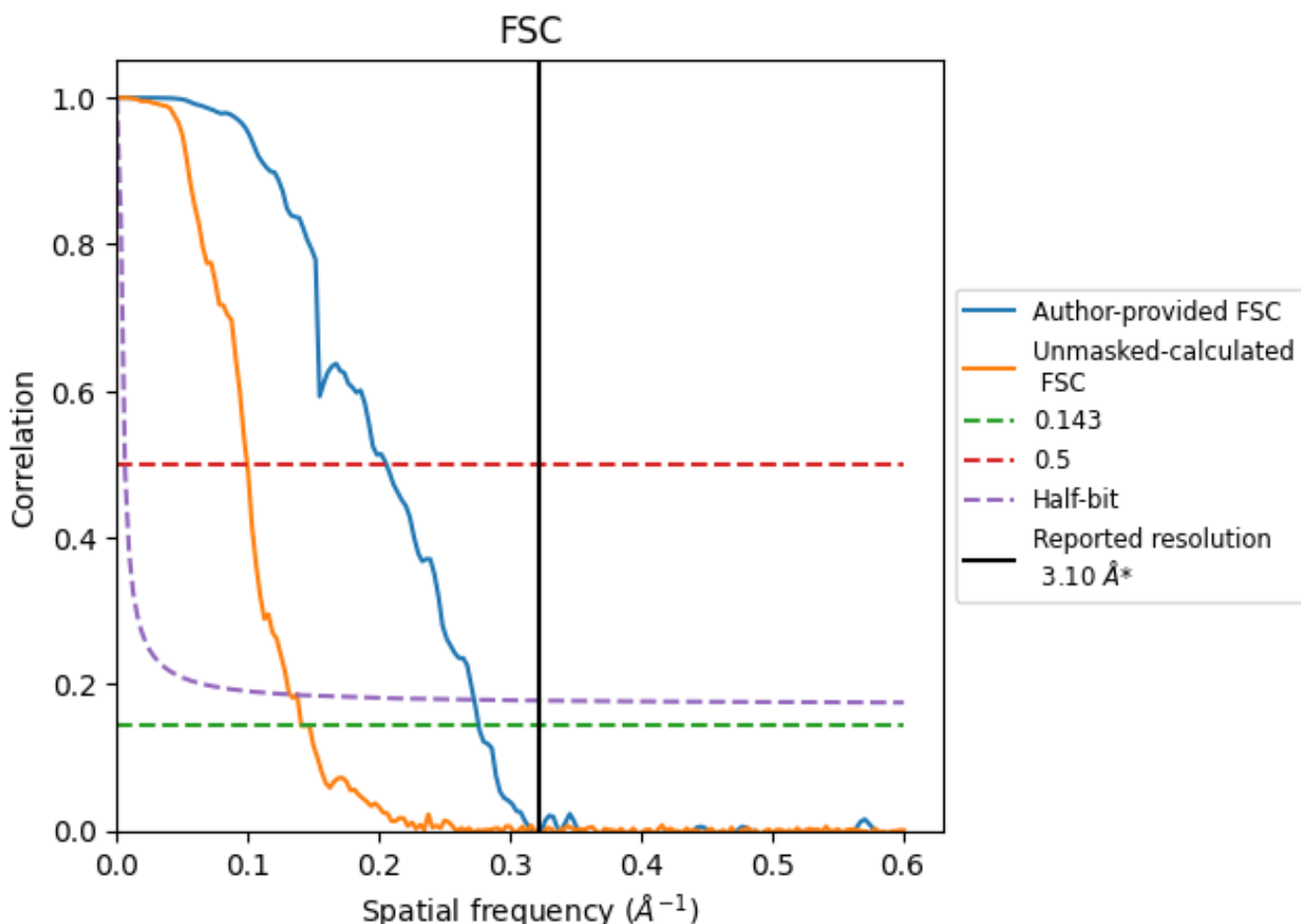


*Reported resolution corresponds to spatial frequency of 0.323 Å⁻¹

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 0.323 \AA^{-1}

8.2 Resolution estimates [i](#)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.10	-	-
Author-provided FSC curve	3.62	4.86	3.66
Unmasked-calculated*	7.10	10.05	7.55

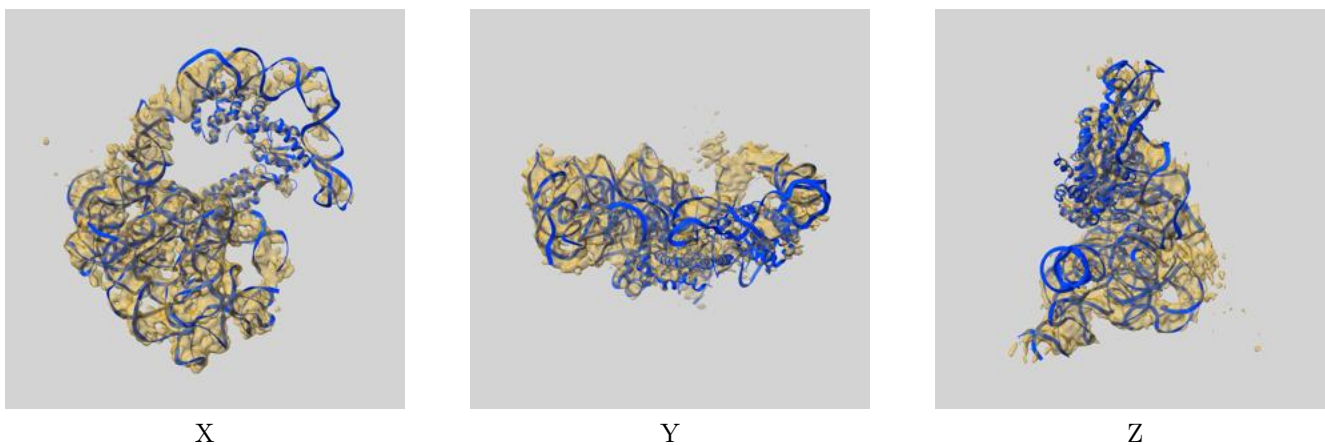
*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from author-provided FSC intersecting FSC 0.143 CUT-OFF 3.62 differs from the reported value 3.1 by more than 10 %

The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 7.10 differs from the reported value 3.1 by more than 10 %

9 Map-model fit [i](#)

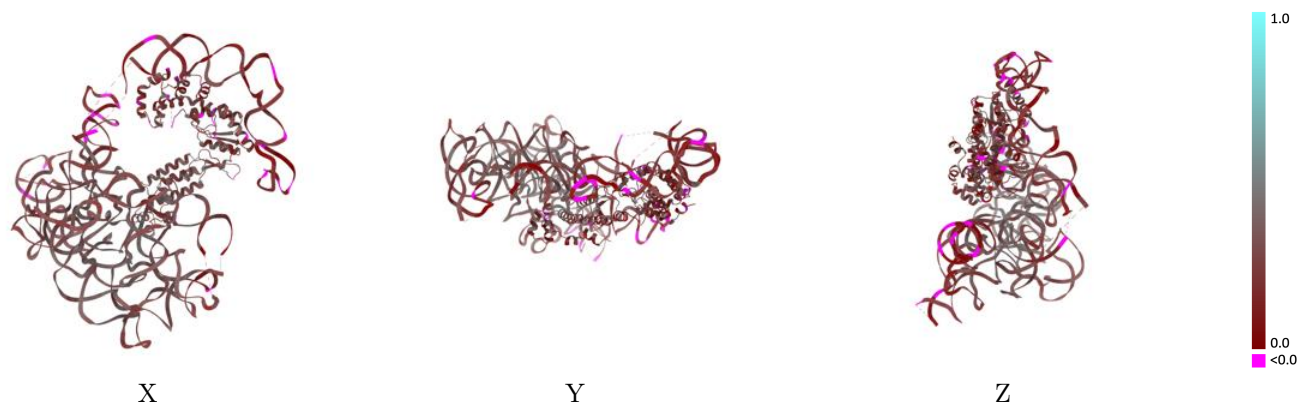
This section contains information regarding the fit between EMDB map EMD-26549 and PDB model 7UIM. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 0.12 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](#)



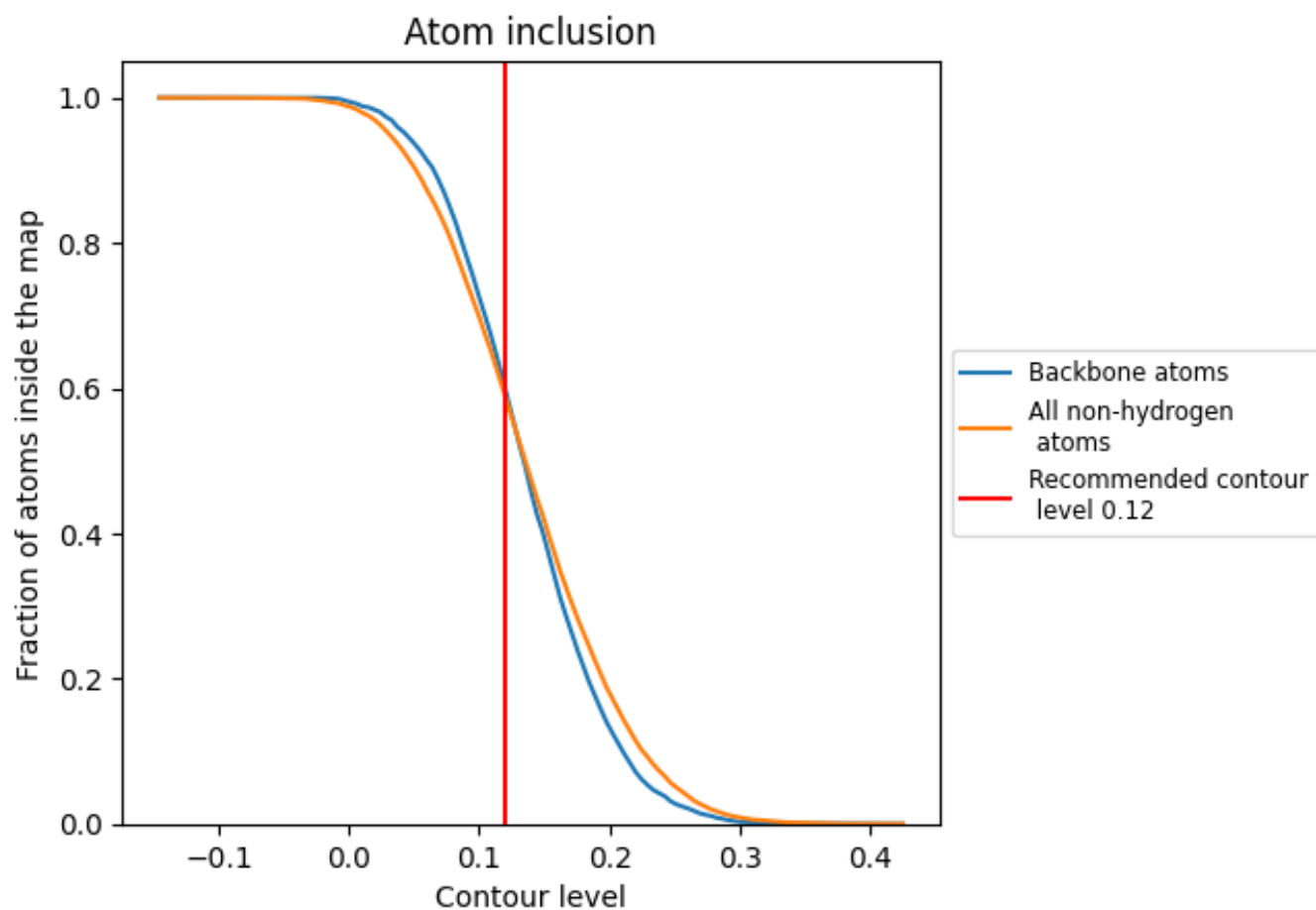
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.12).







9.4 Atom inclusion [i](#)



At the recommended contour level, 60% of all backbone atoms, 59% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary [i](#)

The table lists the average atom inclusion at the recommended contour level (0.12) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.5891	 0.2620
B	 0.6702	 0.2700
D	 0.2801	 0.2310

