

wwPDB EM Validation Summary Report (i)

Jul 3, 2024 – 10:53 am BST

PDB ID : 7O5H

EMDB ID : EMD-12736

Title : Ribosomal methyltransferase KsgA bound to small ribosomal subunit

Authors: Stephan, N.C.; Ries, A.B.; Boehringer, D.; Ban, N.

Deposited on : 2021-04-08

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 (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 (1)) were used in the production of this report:

EMDB validation analysis : 0.0.1.dev92

MolProbity : 4.02b-467

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $\mathrm{MapQ} \quad : \quad 1.9.13$

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

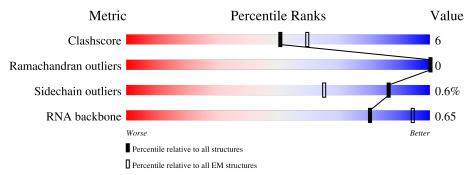
Validation Pipeline (wwPDB-VP) : 2.37.1

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 3.10 Å.

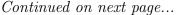
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive $(\# \mathrm{Entries})$	${ m EM\ structures} \ (\#{ m Entries})$
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 <=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	- •								
			90%									
1	V	252	84%	15%								
			12%									
2	A	964	69%	27% •								
			97%									
3	В	225	79%	20% •								
			65%									
4	D	205	82%	18%								
			65%									
5	E	158	79%	21%								
			81%									
6	F	106	75%	22%								
			46%									
7	Н	129	80%	20%								





Continued from previous page...

Mol	Chain	Length	Quality of chain	
			89%	
8	K	117	79%	21%
			47%	
9	L	123	79%	21%
			56%	
10	О	88	83%	16% •
			38%	
11	Р	82	82%	18%
			56%	
12	Q	80	88%	11% •
			60%	
13	R	55	84%	16%
			56%	
14	Т	86	93%	7%
			54%	
15	U	56	48% 5% 46%	



2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 34436 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 Ribosomal RNA small subunit methyltransferase A.

Mol	Chain	Residues		At	oms			AltConf	Trace
1	V	251	Total 1955	C 1247	N 336	O 359	S 13	0	0

• Molecule 2 is a RNA chain called 16S rRNA.

Mol	Chain	Residues	Atoms Total C N O P 20726 9239 3830 6693 964				AltConf	Trace	
2	Λ	964	Total	С	N	О	Р	0	0
	A	904	20726	9239	3830	6693	964	U	U

• Molecule 3 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues		Ato	oms			AltConf	Trace
3	В	225	Total 1760	C 1113	N 316	O 323	S 8	0	0

• Molecule 4 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues		Ato	oms			AltConf	Trace
1	D	205	Total	С	N	О	S	0	0
4	ט	200	1643	1026	315	298	4		U

• Molecule 5 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues		At	oms			AltConf	Trace
5	Е	158	Total 1165	C 725	N 220	O 214	S 6	0	0

• Molecule 6 is a protein called 30S ribosomal protein S6, fully modified isoform.

Mol	Chain	Residues		At	oms			AltConf	Trace
6	F	103	Total		N	О	S	0	0
	_		839	530	151	151	7		Ů



• Molecule 7 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues		At	oms			AltConf	Trace
7	П	129	Total	С	N	О	S	0	0
'	11	129	979	616	173	184	6	0	U

• Molecule 8 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues		At	oms			AltConf	Trace
8	К	117	Total 877	C 540	N 174	O 160	S 3	0	0

• Molecule 9 is a protein called 30S ribosomal protein S12.

\mathbf{Mol}	Chain	Residues		At	oms			AltConf	Trace
9	L	123	Total 955	C 590	N 196	O 165	S 4	0	0

• Molecule 10 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	0	87	Total	С	N	О	S	0	0
10		01	702	433	140	128	1	0	0

• Molecule 11 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	Р	82	Total 649	C 406	N 128	O 114	S 1	0	0

• Molecule 12 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	${f Atoms}$					AltConf	Trace
12	Q	79	Total 641	C 406	N 120	O 112	S 3	0	0

• Molecule 13 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	R	55	Total 456	C 288	N 86	O 82	0	0

• Molecule 14 is a protein called 30S ribosomal protein S20.



Mol	Chain	Residues	${f Atoms}$					AltConf	Trace
1.4	Т	86	Total	С	N	О	S	0	0
14	14 1		670	414	138	115	3		

• Molecule 15 is a protein called 30S ribosomal protein S21.

N	/Iol	Chain	Residues	Atoms				AltConf	Trace
	15	U	30	Total 252	C 156	N 54	O 42	0	0

• Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

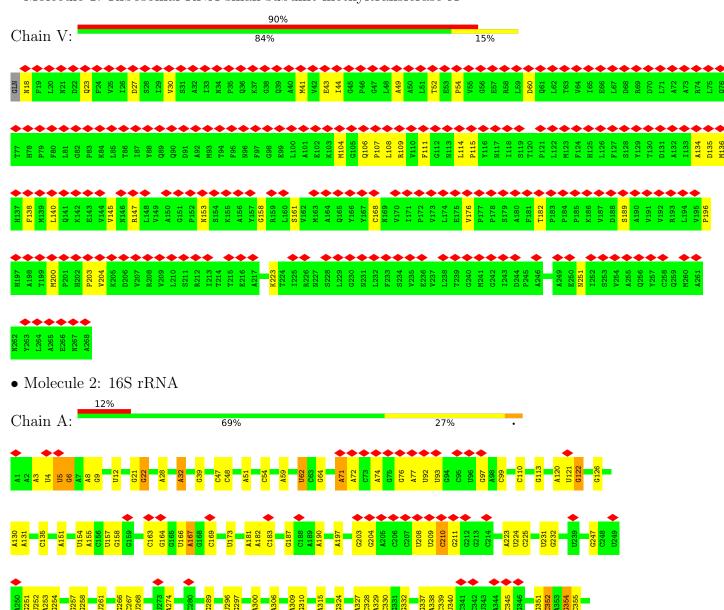
Mol	Chain	Residues	Atoms	AltConf
16	A	161	Total Mg 161 161	0
16	D	1	Total Mg 1 1	0
16	E	1	Total Mg 1 1	0
16	F	1	Total Mg 1 1	0
16	Н	1	Total Mg 1 1	0
16	К	1	Total Mg 1 1	0
16	U	1	Total Mg 1 1	0



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: Ribosomal RNA small subunit methyltransferase A









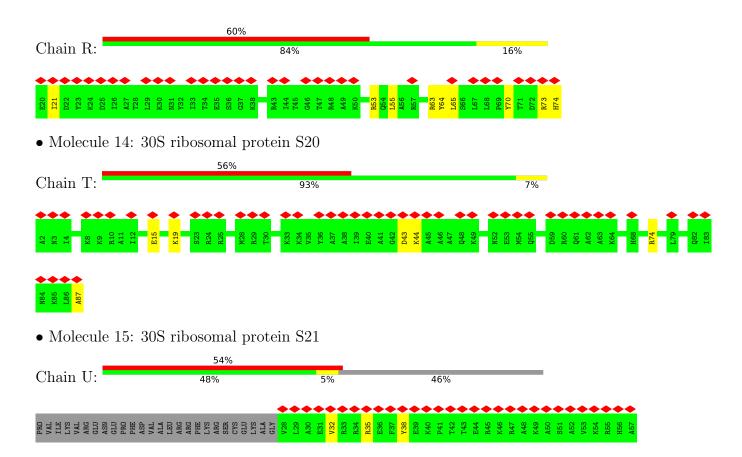






• Molecule 13: 30S ribosomal protein S18







4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	165073	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE	Depositor
	CORRECTION	
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	35	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	0.245	Depositor
Minimum map value	-0.079	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.081	Depositor
Map size (Å)	444.8, 444.8, 444.8	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.39, 1.39, 1.39	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: 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
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	V	0.25	0/1998	0.45	0/2718
2	A	0.21	0/23209	0.67	0/36199
3	В	0.24	0/1791	0.45	0/2413
4	D	0.24	0/1665	0.49	0/2227
5	Е	0.25	0/1178	0.49	0/1584
6	F	0.24	0/858	0.47	0/1160
7	Н	0.25	0/989	0.47	0/1326
8	K	0.24	0/893	0.52	0/1205
9	L	0.25	0/969	0.56	0/1300
10	О	0.22	0/710	0.48	0/950
11	Р	0.24	0/659	0.54	0/884
12	Q	0.24	0/650	0.51	0/871
13	R	0.24	0/463	0.51	0/621
14	Т	0.23	0/676	0.43	0/895
15	U	0.23	0/255	0.55	0/338
All	All	0.22	0/36963	0.61	0/54691

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 (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	V	1955	0	1969	23	0
2	A	20726	0	10422	148	0
3	В	1760	0	1787	28	0
4	D	1643	0	1707	25	0
5	Е	1165	0	1212	18	0
6	F	839	0	833	16	0
7	Н	979	0	1031	18	0
8	K	877	0	886	19	0
9	L	955	0	1016	19	0
10	O	702	0	721	9	0
11	Р	649	0	666	12	0
12	Q	641	0	682	5	0
13	R	456	0	478	7	0
14	Т	670	0	719	4	0
15	U	252	0	267	3	0
16	A	161	0	0	0	0
16	D	1	0	0	0	0
16	Е	1	0	0	0	0
16	F	1	0	0	0	0
16	Н	1	0	0	0	0
16	K	1	0	0	0	0
16	U	1	0	0	0	0
All	All	34436	0	24396	307	0

The all-atom clash score is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clash score for this structure is 6.

The worst 5 of 307 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$egin{aligned} \operatorname{Clash} \ \operatorname{overlap}\ (\mathring{\mathbf{A}}) \end{aligned}$
2:A:437:U:HO2'	4:D:120:HIS:HD1	1.17	0.92
2:A:823:C:HO2'	7:H:2:SER:N	1.73	0.87
2:A:71:A:N6	2:A:99:C:O2	2.20	0.75
2:A:483:C:O2	11:P:13:LYS:NZ	2.22	0.72
8:K:23:ILE:HG22	8:K:32:VAL:HG12	1.70	0.72

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	Perce	ntiles
1	V	$249/252\ (99\%)$	246 (99%)	3 (1%)	0	100	100
3	В	$223/225\ (99\%)$	212 (95%)	11 (5%)	0	100	100
4	D	$203/205\ (99\%)$	200 (98%)	3 (2%)	0	100	100
5	E	156/158~(99%)	155 (99%)	1 (1%)	0	100	100
6	F	101/106~(95%)	101 (100%)	0	0	100	100
7	Н	127/129~(98%)	125 (98%)	2 (2%)	0	100	100
8	K	115/117 (98%)	110 (96%)	5 (4%)	0	100	100
9	L	121/123~(98%)	115 (95%)	6 (5%)	0	100	100
10	О	85/88~(97%)	84 (99%)	1 (1%)	0	100	100
11	Р	80/82 (98%)	77 (96%)	3 (4%)	0	100	100
12	Q	77/80~(96%)	76 (99%)	1 (1%)	0	100	100
13	R	53/55~(96%)	53 (100%)	0	0	100	100
14	Т	84/86 (98%)	84 (100%)	0	0	100	100
15	U	28/56 (50%)	28 (100%)	0	0	100	100
All	All	$1702/1762\ (97\%)$	1666 (98%)	36 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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

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



Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	V	215/216 (100%)	212 (99%)	3 (1%)	67 86
3	В	187/187 (100%)	185 (99%)	2 (1%)	73 89
4	D	172/172 (100%)	171 (99%)	1 (1%)	86 94
5	E	120/120 (100%)	119 (99%)	1 (1%)	81 92
6	F	90/92 (98%)	90 (100%)	0	100 100
7	Н	104/104 (100%)	104 (100%)	0	100 100
8	K	90/90 (100%)	90 (100%)	0	100 100
9	L	103/103 (100%)	102 (99%)	1 (1%)	76 90
10	О	75/76 (99%)	75 (100%)	0	100 100
11	Р	65/65 (100%)	65 (100%)	0	100 100
12	Q	73/74 (99%)	72 (99%)	1 (1%)	67 86
13	R	48/48 (100%)	48 (100%)	0	100 100
14	Т	65/65 (100%)	65 (100%)	0	100 100
15	U	25/48 (52%)	25 (100%)	0	100 100
All	All	1432/1460 (98%)	1423 (99%)	9 (1%)	86 94

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	L	18	LYS
12	Q	28	PHE
3	В	117	LEU
3	В	120	GLN
4	D	160	GLU

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such side chains are listed below:

Mol	Chain	Res	Type
9	L	5	ASN
13	R	74	HIS
14	Т	48	GLN
6	F	14	GLN
1	V	106	GLN

5.3.3 RNA (i)



Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	A	956/964 (99%)	126 (13%)	2 (0%)

5 of 126 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	A	5	U
2	A	6	G
2	A	9	G
2	A	22	G
2	A	32	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	A	411	A
2	A	1107	С

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 167 ligands modelled in this entry, 167 are monoatomic - 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	A	7

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1108:G	O3'	1392:G	P	42.05
1	A	1396:A	O3'	1507:A	P	29.66
1	A	924:C	O3'	1068:G	P	23.32
1	A	840:C	O3'	846:G	P	16.77
1	A	77:A	O3'	92:U	P	16.40



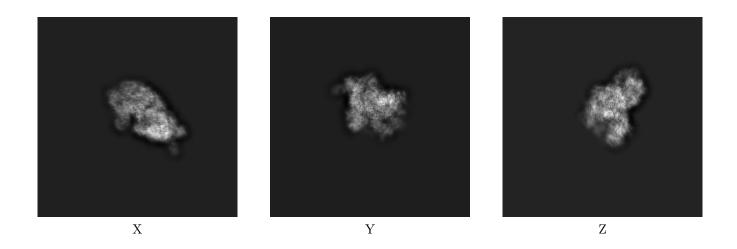
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-12736. These allow visual inspection of the internal detail of the map and identification of artifacts.

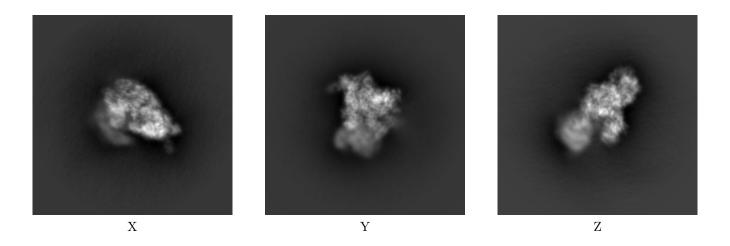
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

6.1.1 Primary map



6.1.2 Raw map

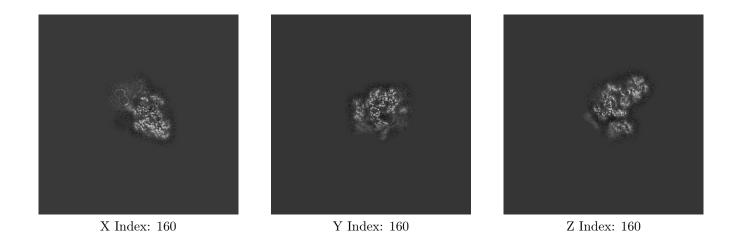


The images above show the map projected in three orthogonal directions.

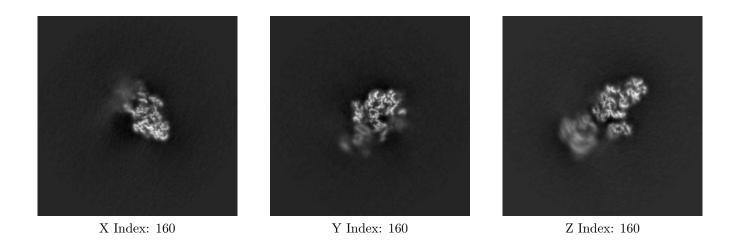


6.2 Central slices (i)

6.2.1 Primary map



6.2.2 Raw map

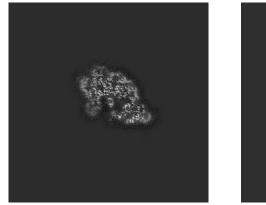


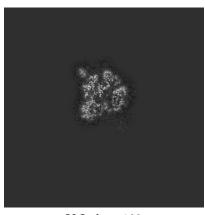
The images above show central slices of the map in three orthogonal directions.

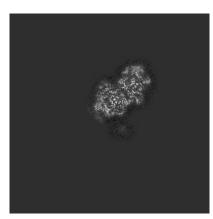


6.3 Largest variance slices (i)

6.3.1 Primary map





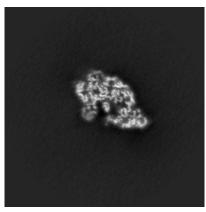


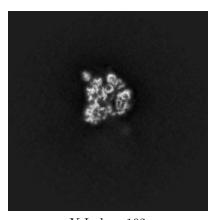
X Index: 184

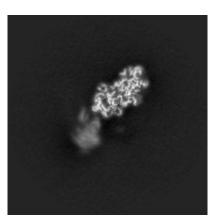
Y Index: 192

Z Index: 137

6.3.2 Raw map







X Index: 184

Y Index: 192

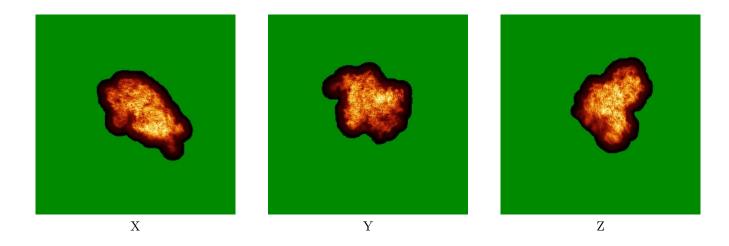
Z Index: 135

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

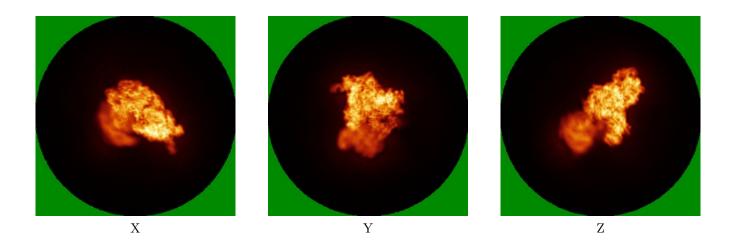


6.4 Orthogonal standard-deviation projections (False-color) (i)

6.4.1 Primary map



6.4.2 Raw map

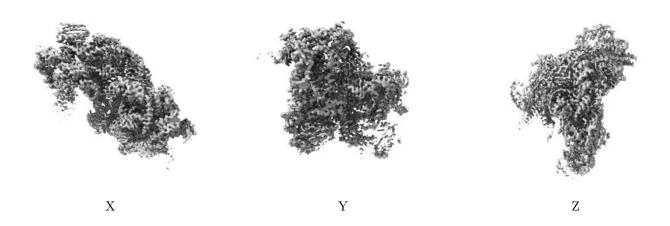


The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.



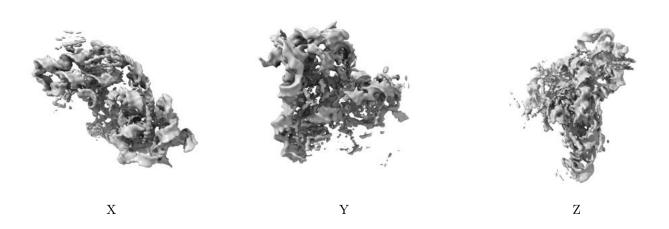
6.5 Orthogonal surface views (i)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.081. 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



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.



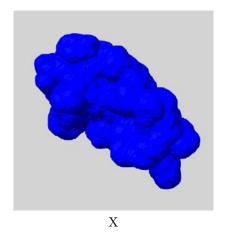
6.6 Mask visualisation (i)

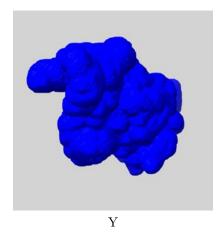
This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

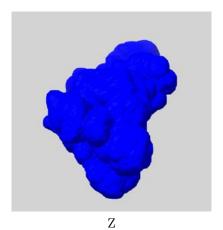
A mask typically either:

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

$6.6.1 \quad \mathrm{emd}_12736_\mathrm{msk}_1.\mathrm{map} \ \ \mathbf{\mathring{1}}$



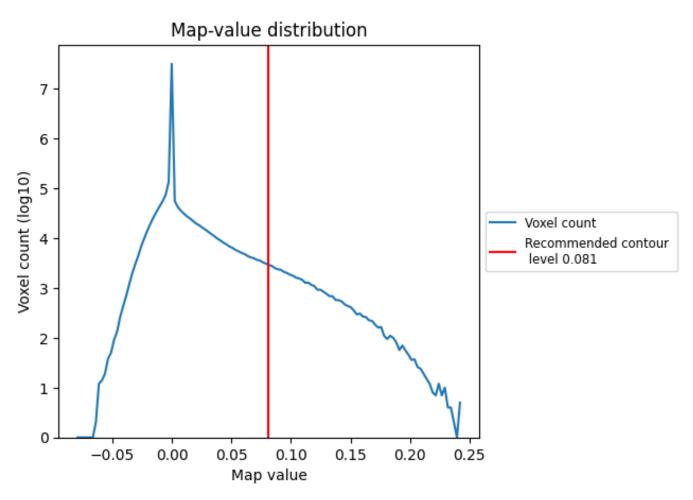




7 Map analysis (i)

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

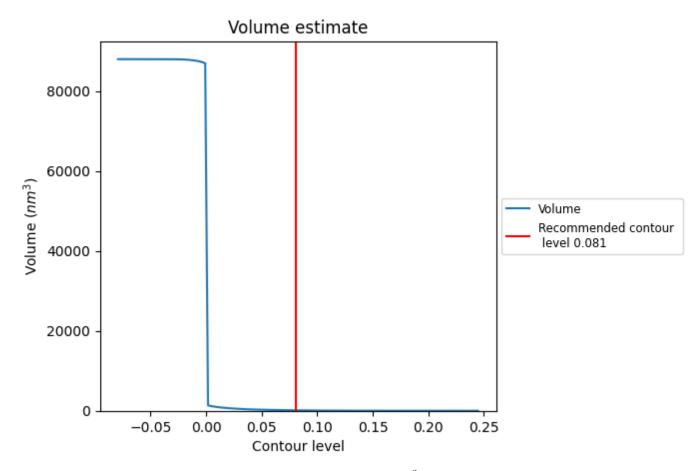
7.1 Map-value distribution (i)



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)

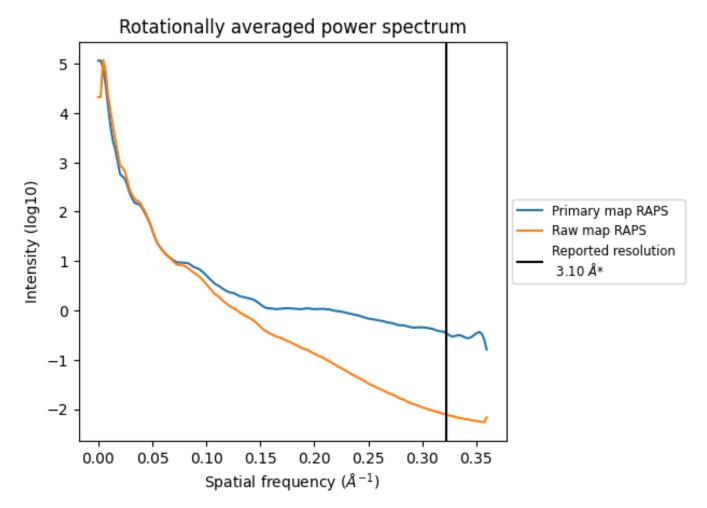


The volume at the recommended contour level is $112~\mathrm{nm}^3$; this corresponds to an approximate mass of $101~\mathrm{kDa}$.

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)



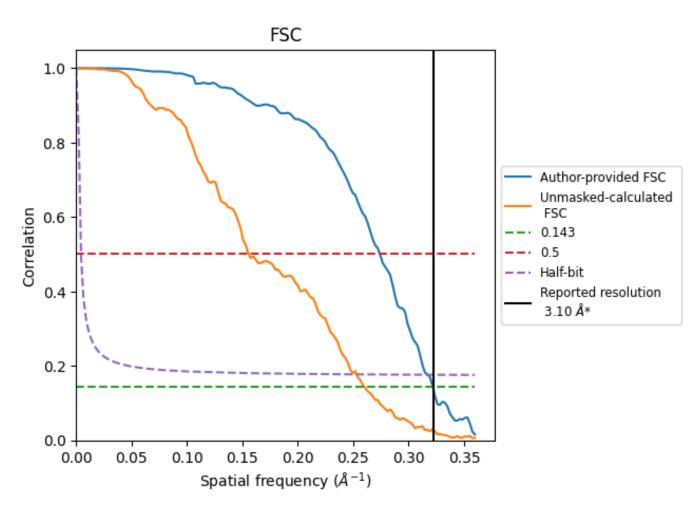
^{*}Reported resolution corresponds to spatial frequency of 0.323 $\rm \mathring{A}^{-1}$



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 $\rm \mathring{A}^{-1}$



8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
rtesolution estimate (A)	0.143	0.5	Half-bit
Reported by author	3.10	-	-
Author-provided FSC curve	3.11	3.65	3.15
Unmasked-calculated*	3.84	6.42	3.95

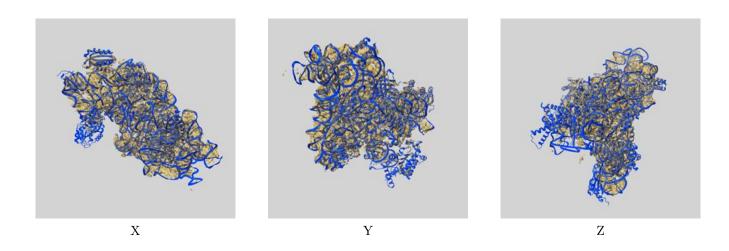
^{*}Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.84 differs from the reported value 3.1 by more than 10 %



9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-12736 and PDB model 705H. Per-residue inclusion information can be found in section 3 on page 7.

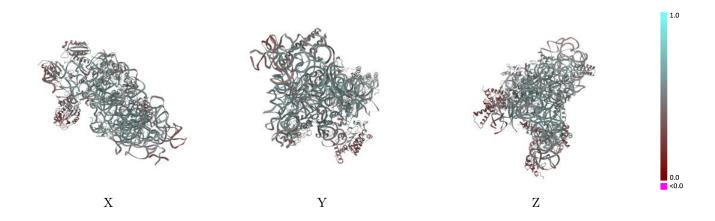
9.1 Map-model overlay (i)



The images above show the 3D surface view of the map at the recommended contour level 0.081 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.

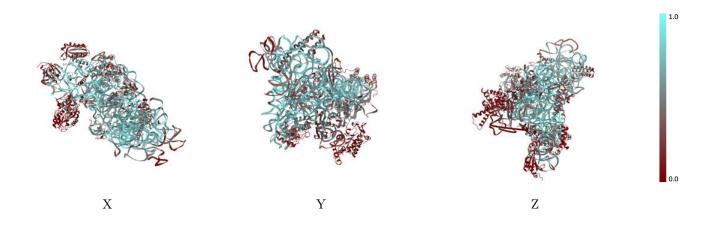


9.2 Q-score mapped to coordinate model (i)



The images above show the model with each residue coloured according 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.

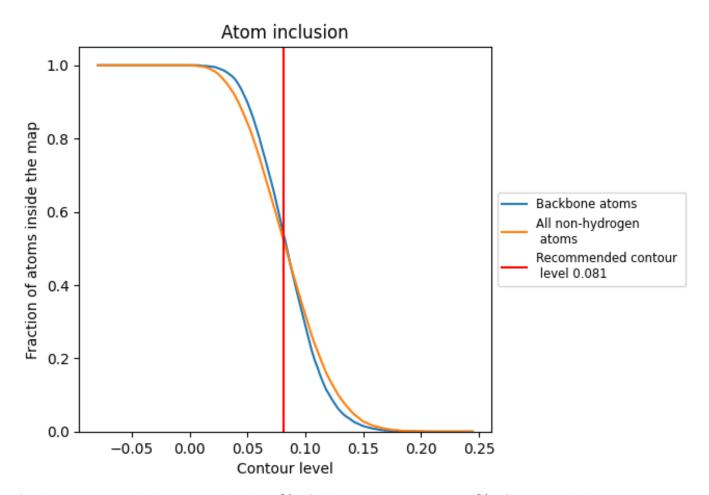
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.081).



9.4 Atom inclusion (i)



At the recommended contour level, 54% of all backbone atoms, 52% 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 (0.081) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5230	0.5030
A	0.6820	0.5280
В	0.0860	0.3850
D	0.3380	0.4890
Е	0.3410	0.5090
F	0.2440	0.4540
Н	0.4100	0.5280
K	0.1160	0.4040
L	0.3990	0.5300
О	0.4160	0.5170
Р	0.4740	0.5240
Q	0.3500	0.5110
R	0.3910	0.4940
Т	0.3920	0.4810
U	0.0000	0.4100
V	0.1170	0.4000



