

Full wwPDB EM Validation Report (i)

Jul 10, 2024 – 04:50 am BST

PDB ID : 7QOF

EMDB ID : EMD-14088

Title : Icosahedral capsid of the phicrAss001 virion

Authors: Bayfield, O.W.; Shkoporov, A.N.; Yutin, N.; Khokhlova, E.V.; Smith, J.L.R.;

Hawkins, D.E.D.P.; Koonin, E.V.; Hill, C.; Antson, A.A.

Deposited on : 2021-12-24

Resolution : 3.01 Å(reported)

This is a Full wwPDB EM Validation 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)

MapQ : 1.9.13

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

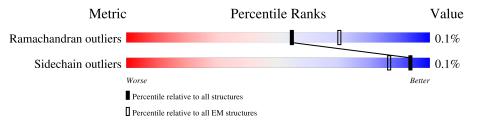
Validation Pipeline (wwPDB-VP) : 2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $ELECTRON\ MICROSCOPY$

The reported resolution of this entry is 3.01 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 $(\# \mathrm{Entries})$	${ m EM~structures} \ (\#{ m Entries})$	
Ramachandran outliers	154571	4023	
Sidechain outliers	154315	3826	

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

Mol	Chain	Length	Quality of chain
1	A	504	94% 6%
1	В	504	100%
1	С	504	100%
1	D	504	100%
1	Е	504	100%
1	F	504	100%
1	G	504	100%
1	Н	504	100%
1	I	504	100%



Continued from previous page...

Mol	Chain	Length		uality of chain
2	a	333	<u>-</u>	99%
2	b	333	5%	99%
2	c	333	•	99%
2	d	333	·	100%
2	e	333	<u>·</u>	99%
2	f	333	<u>·</u>	99%
2	g	333	•	99%
2	h	333	<u>•</u>	99%
2	i	333	<u>-</u>	99%
3	i	122	33%	52%
3	k	122	39%	52%
4	1	97	39%	61%



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 59983 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 Major capsid protein gp32.

Mol	Chain	Residues		Atoms					Trace
1	A	472	Total	С	N	О	S	0	0
1	Λ	412	3776	2412	650	695	19	0	
1	В	503	Total	С	N	О	S	0	0
1	D	505	4012	2552	695	745	20	0	
1	С	503	Total	С	N	О	S	0	0
1	C	505	4012	2552	695	745	20	0	U
1	D	503	Total	С	N	О	S	0	0
1	D	505	4012	2552	695	745	20	0	
1	E	503	Total	С	N	О	S	0	0
1	ינו	303	4012	2552	695	745	20		
1	F	503	Total	С	N	О	S	0	0
1	I'	303	4012	2552	695	745	20	0	
1	G	503	Total	С	N	О	S	0	0
1	G	303	4012	2552	695	745	20	0	
1	Н	503	Total	С	N	О	S	0	0
1	11	505	4012	2552	695	745	20		U
1	I	503	Total	С	N	О	S	0	0
1	1	909	4012	2552	695	745	20	U	

• Molecule 2 is a protein called Auxiliary capsid protein gp36.

Mol	Chain	Residues		At	oms			AltConf	Trace
2		333	Total	С	N	О	S	0	0
	a	333	2542	1624	410	498	10	U	0
2	b	999	Total	С	N	О	S	0	0
	D	333	2542	1624	410	498	10	U	0
2		333	Total	С	N	О	S	0	0
2	c		2542	1624	410	498	10		
2	d	333	Total	С	N	О	S	0	0
	u	333	2542	1624	410	498	10	0	0
2	0	333	Total	С	N	О	S	0	0
2	е	e 333	2542	1624	410	498	10	U	U
2	f	222	Total	С	N	О	S	0	0
	1	f 333	2542	1624	410	498	10		0



Continued from previous page...

Mol	Chain	Residues		Atoms				AltConf	Trace
2	ď	333	Total	С	N	О	S	0	0
	g	333	2542	1624	410	498	10	0	U
2	h	333	Total	С	N	О	S	0	0
	11	333	2542	1624	410	498	10	0	U
2	;	222	Total	С	N	О	S	0	0
	2 1	333	2542	1624	410	498	10	U	U

• Molecule 3 is a protein called Head fiber trimer protein gp21.

Mol	Chain	Residues		Ato	ms			AltConf	Trace
9	;	50	Total	С	N	О	S	0	0
3	J	58	461	292	78	90	1	U	0
9	1,	58	Total	С	N	О	S	0	0
3	K	30	461	292	78	90	1	0	U

• Molecule 4 is a protein called Head fiber dimer protein gp29.

Mol	Chain	Residues		Ato	$\mathbf{m}\mathbf{s}$			AltConf	Trace
4	1	38	Total	С	N	О	S	0	0
4	1	30	310	201	51	56	2	0	U

• Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
5	D	1	Total Mg 1 1	0



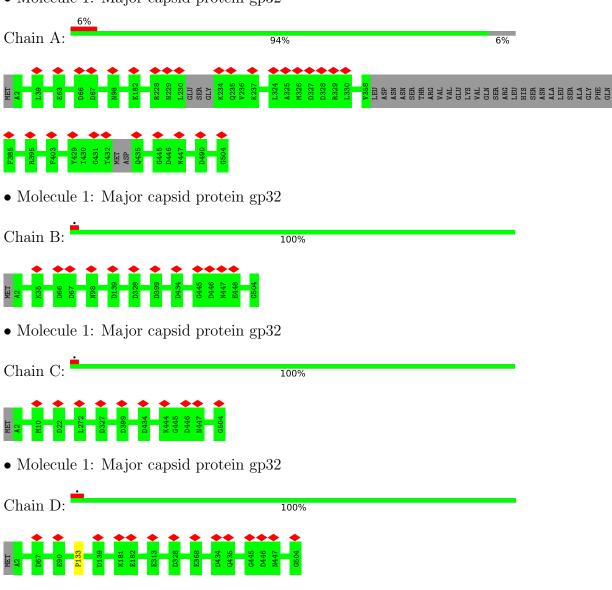
3 Residue-property plots (i)

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: Major capsid protein gp32

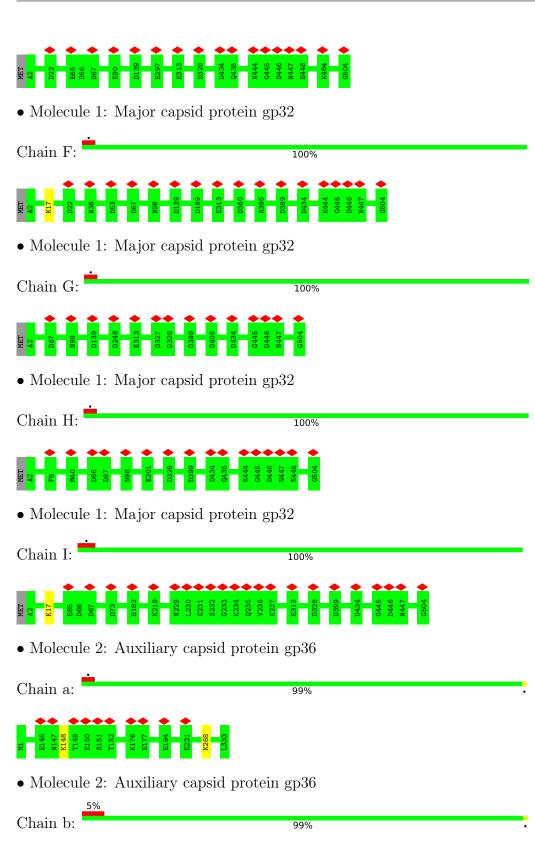
• Molecule 1: Major capsid protein gp32

Chain E:



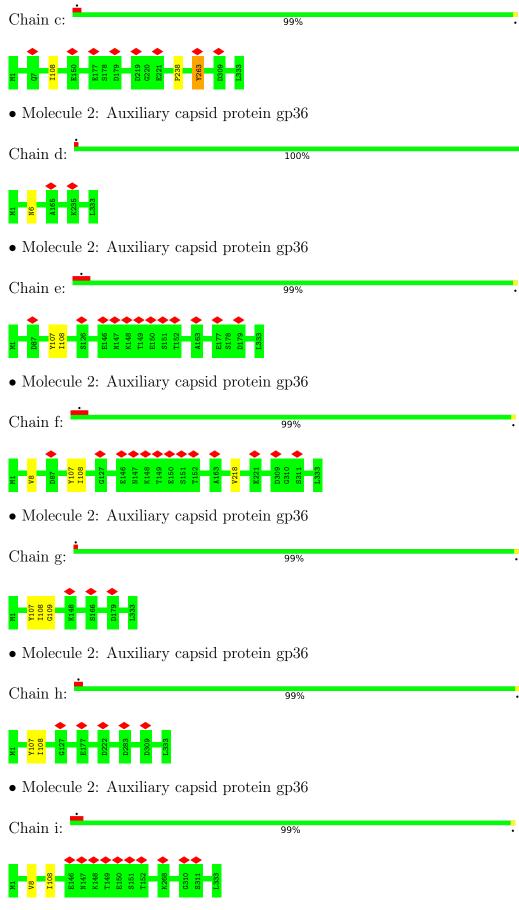


100%



• Molecule 2: Auxiliary capsid protein gp36









4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	55063	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 $(e^-/\text{Å}^2)$	51	Depositor
Minimum defocus (nm)	300	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	0.194	Depositor
Minimum map value	-0.102	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.035	Depositor
Map size (Å)	1250.7347, 1250.7347, 1250.7347	wwPDB
Map dimensions	916, 916, 916	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.365431, 1.365431, 1.365431	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	Во	ond angles
MIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	A	0.31	0/3865	0.45	0/5224
1	В	0.31	0/4106	0.47	0/5553
1	С	0.31	0/4106	0.46	0/5553
1	D	0.31	0/4106	0.47	0/5553
1	Е	0.31	0/4106	0.48	0/5553
1	F	0.30	0/4106	0.46	0/5553
1	G	0.31	0/4106	0.46	0/5553
1	Н	0.31	0/4106	0.46	0/5553
1	I	0.31	0/4106	0.47	0/5553
2	a	0.30	0/2591	0.48	0/3533
2	b	0.30	0/2591	0.49	0/3533
2	С	0.31	0/2591	0.50	1/3533~(0.0%)
2	d	0.31	0/2591	0.47	0/3533
2	е	0.31	0/2591	0.49	0/3533
2	f	0.31	0/2591	0.49	0/3533
2	g	0.31	0/2591	0.48	0/3533
2	h	0.31	0/2591	0.49	0/3533
2	i	0.31	0/2591	0.49	0/3533
3	j	0.25	0/462	0.47	0/619
3	k	0.25	0/462	0.44	0/619
4	1	0.29	0/318	0.47	0/430
All	All	0.31	0/61274	0.47	1/83113 (0.0%)

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 maintenain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	b	0	2
2	С	1	0



Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
2	d	0	1
2	е	0	1
2	f	0	2
2	g	0	1
2	h	0	1
All	All	1	8

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$Ideal(^{o})$
2	С	263	TYR	N-CA-CB	5.17	119.90	110.60

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	c	263	TYR	CA

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	b	107	TYR	Peptide
2	b	218	VAL	Peptide
2	d	6	ASN	Peptide
2	е	107	TYR	Peptide
2	f	107	TYR	Peptide
2	f	218	VAL	Peptide
2	g	107	TYR	Peptide
2	h	107	TYR	Peptide

5.2 Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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



entries.

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	ntiles
1	A	464/504~(92%)	445 (96%)	19 (4%)	0	100	100
1	В	501/504 (99%)	476 (95%)	25 (5%)	0	100	100
1	С	501/504 (99%)	477 (95%)	24 (5%)	0	100	100
1	D	501/504 (99%)	481 (96%)	20 (4%)	0	100	100
1	Е	501/504 (99%)	478 (95%)	23 (5%)	0	100	100
1	F	501/504 (99%)	484 (97%)	17 (3%)	0	100	100
1	G	501/504 (99%)	475 (95%)	26 (5%)	0	100	100
1	Н	501/504 (99%)	482 (96%)	19 (4%)	0	100	100
1	I	501/504 (99%)	474 (95%)	27 (5%)	0	100	100
2	a	331/333 (99%)	309 (93%)	22 (7%)	0	100	100
2	b	331/333 (99%)	309 (93%)	20 (6%)	2 (1%)	25	62
2	С	331/333 (99%)	310 (94%)	20 (6%)	1 (0%)	41	75
2	d	331/333 (99%)	309 (93%)	22 (7%)	0	100	100
2	e	331/333 (99%)	315 (95%)	15 (4%)	1 (0%)	41	75
2	f	331/333 (99%)	313 (95%)	16 (5%)	2 (1%)	25	62
2	g	331/333 (99%)	316 (96%)	13 (4%)	2 (1%)	25	62
2	h	331/333 (99%)	314 (95%)	16 (5%)	1 (0%)	41	75
2	i	331/333 (99%)	309 (93%)	20 (6%)	2 (1%)	25	62
3	j	56/122 (46%)	53 (95%)	3 (5%)	0	100	100
3	k	56/122 (46%)	56 (100%)	0	0	100	100
4	1	36/97 (37%)	34 (94%)	2 (6%)	0	100	100
All	All	7599/7874 (96%)	7219 (95%)	369 (5%)	11 (0%)	54	85

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	b	108	ILE
2	е	108	ILE
2	f	108	ILE
2	g	108	ILE
2	h	108	ILE



Continued from previous page...

Mol	Chain	Res	Type
2	i	8	VAL
2	f	8	VAL
2	b	8	VAL
2	С	108	ILE
2	i	108	ILE
2	g	109	GLY

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	Perce	ntiles
1	A	399/427~(93%)	399 (100%)	0	100	100
1	В	426/427 (100%)	426 (100%)	0	100	100
1	С	426/427 (100%)	426 (100%)	0	100	100
1	D	426/427 (100%)	425 (100%)	1 (0%)	93	98
1	Е	426/427 (100%)	426 (100%)	0	100	100
1	F	426/427 (100%)	425 (100%)	1 (0%)	93	98
1	G	$426/427 \; (100\%)$	426 (100%)	0	100	100
1	Н	426/427 (100%)	426 (100%)	0	100	100
1	I	426/427 (100%)	425 (100%)	1 (0%)	93	98
2	a	275/275 (100%)	273 (99%)	2 (1%)	84	94
2	b	275/275~(100%)	275 (100%)	0	100	100
2	С	275/275 (100%)	273 (99%)	2 (1%)	84	94
2	d	275/275 (100%)	275 (100%)	0	100	100
2	е	275/275 (100%)	275 (100%)	0	100	100
2	f	275/275~(100%)	275 (100%)	0	100	100
2	g	275/275 (100%)	275 (100%)	0	100	100
2	h	275/275 (100%)	275 (100%)	0	100	100
2	i	275/275 (100%)	275 (100%)	0	100	100
3	j	54/107~(50%)	54 (100%)	0	100	100



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Perce	$_{ m ntiles}$
3	k	54/107 (50%)	54 (100%)	0	100	100
4	1	34/78 (44%)	34 (100%)	0	100	100
All	All	6424/6610 (97%)	6417 (100%)	7 (0%)	93	98

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	133	PRO
1	F	17	LYS
1	I	17	LYS
2	a	148	LYS
2	a	268	LYS
2	С	238	PRO
2	c	263	TYR

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. All (51) such side chains are listed below:

Mol	Chain	Res	Type
1	A	91	ASN
1	A	209	HIS
1	A	261	ASN
1	A	305	ASN
1	B C C C	414	HIS
1	С	435	GLN
1	С	479	HIS
1	С	503	GLN
1		68	ASN
1	D	273	ASN
1	D	278	ASN
1	D	447	ASN
1	Е	229	ASN
1	F	229	ASN
1	F	261	ASN
1	F	503	GLN
1	G	9	GLN
1	G	57	ASN
1	G	91	ASN
1	G	200	ASN
1	G	241	ASN
1	G	305	ASN



Continued from previous page...

Mol	Chain	Res	Type
1	G	392	ASN
1	Н	200	ASN ASN
1	Н	229	ASN
1	Н	392	ASN
1	I	81	ASN
1	I	254	GLN
1	I	392	ASN
2	a	58	ASN
2 2 2 2 2 2 2 2 2 2 2	a	61	HIS
2	a	115	GLN
2	b	96	GLN
2	b	147	ASN
2	b	241	HIS
2	С	265	ASN
2	c	292	ASN
2	c	316	ASN
2	d	115	GLN
2 2 2	d	296	GLN
2	е	61	HIS
2	е	312	HIS
2 2 2 2	f	20	ASN
2	f	195	GLN
2	h	20	ASN
2	h	97	ASN
2	h	292	ASN
2	i	7	GLN
2	i	61	HIS
2 2 2 2 4	i	292	ASN
4	1	31	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 1 ligands modelled in this entry, 1 is 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)

There are no chain breaks in this entry.



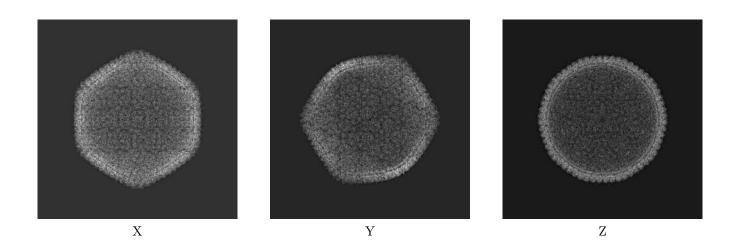
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-14088. 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)

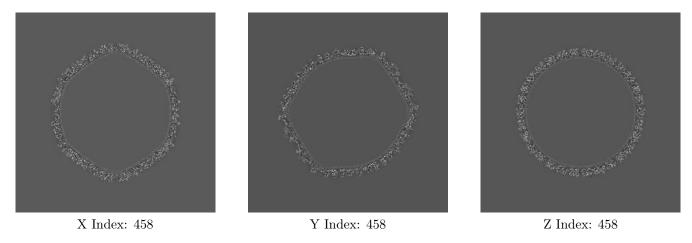
6.1.1 Primary map



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

6.2 Central slices (i)

6.2.1 Primary map

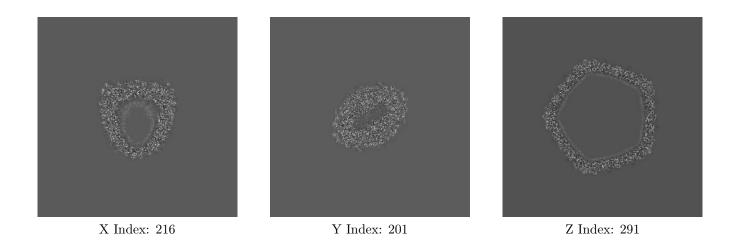




The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

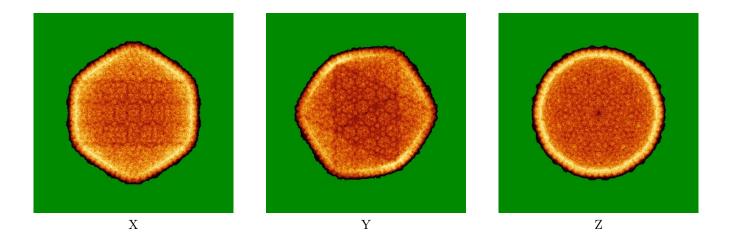
6.3.1 Primary map



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

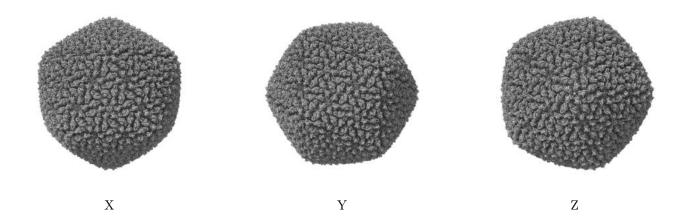


The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.



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 0.035. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.6 Mask visualisation (i)

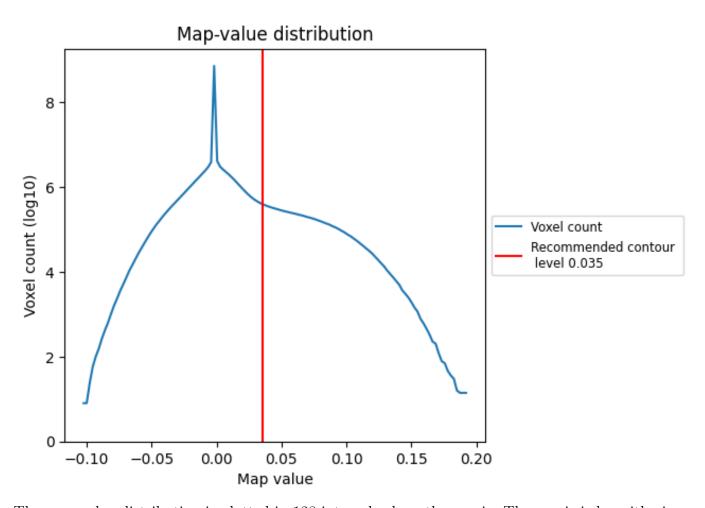
This section was not generated. No masks/segmentation were deposited.



7 Map analysis (i)

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

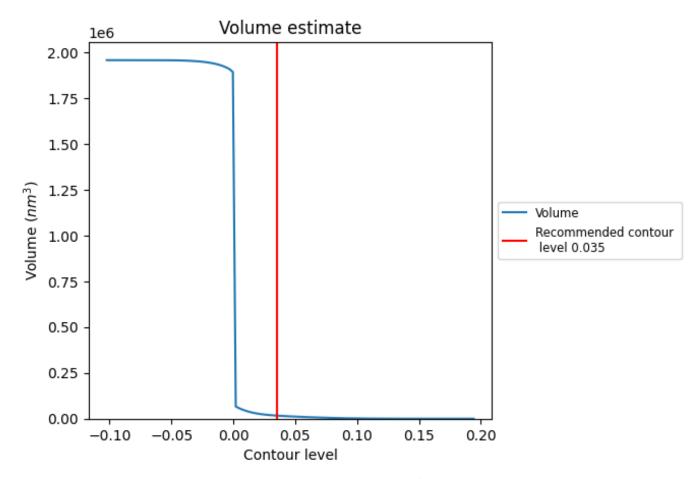
7.1 Map-value distribution (i)



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)

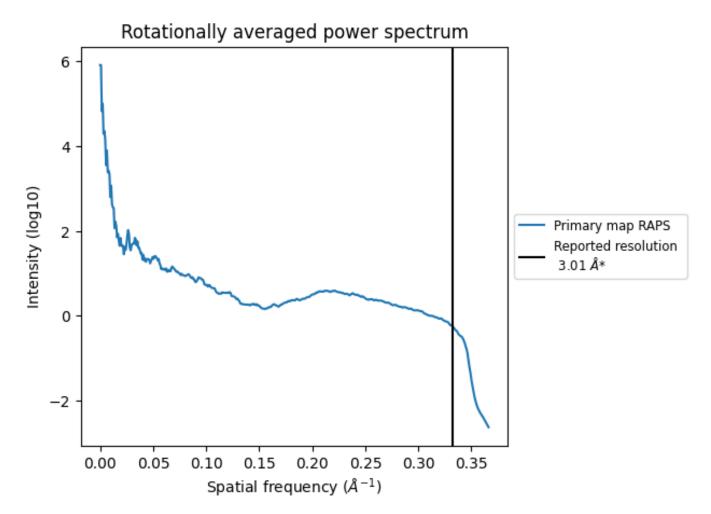


The volume at the recommended contour level is $17018~\mathrm{nm}^3$; this corresponds to an approximate mass of $15373~\mathrm{kDa}$.

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)



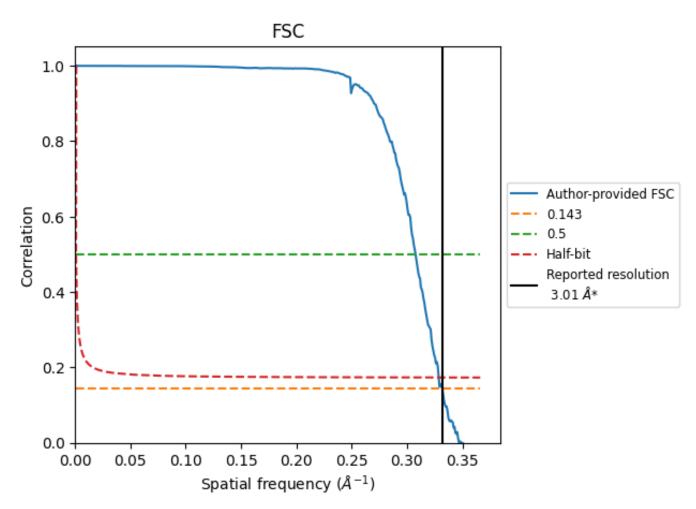
^{*}Reported resolution corresponds to spatial frequency of 0.332 $\rm \AA^{-1}$



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 0.332 $\rm \AA^{-1}$



8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
resolution estimate (A)	0.143	0.5	Half-bit
Reported by author	3.01	-	-
Author-provided FSC curve	3.01	3.25	3.04
Unmasked-calculated*	-	-	-

^{*}Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.



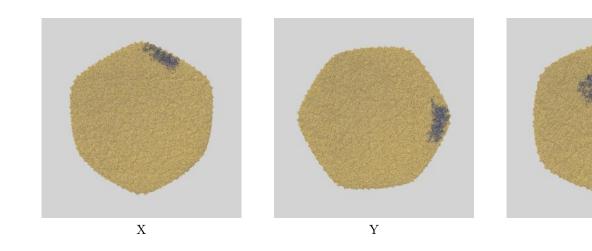
 \mathbf{Z}

9 Map-model fit (i)

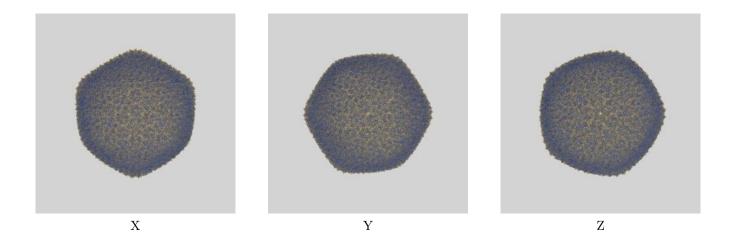
This section contains information regarding the fit between EMDB map EMD-14088 and PDB model 7QOF. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlays

9.1.1 Map-model overlay (i)



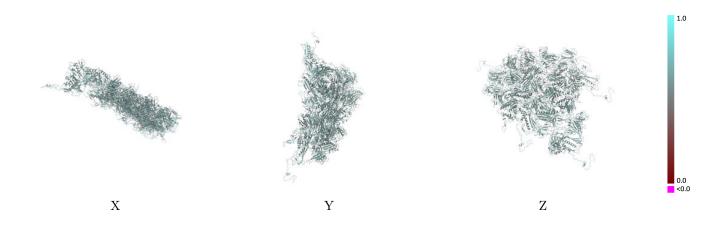
9.1.2 Map-model assembly overlay (i)



The images above show the 3D surface view of the map at the recommended contour level 0.035 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.

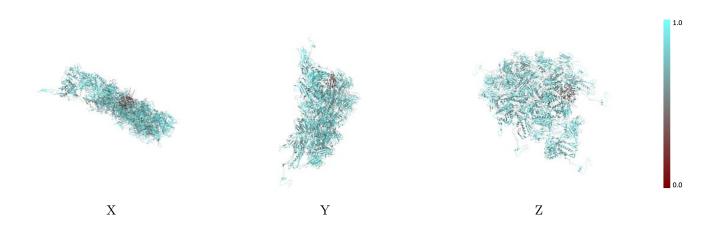


9.2 Q-score mapped to coordinate model (i)



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.

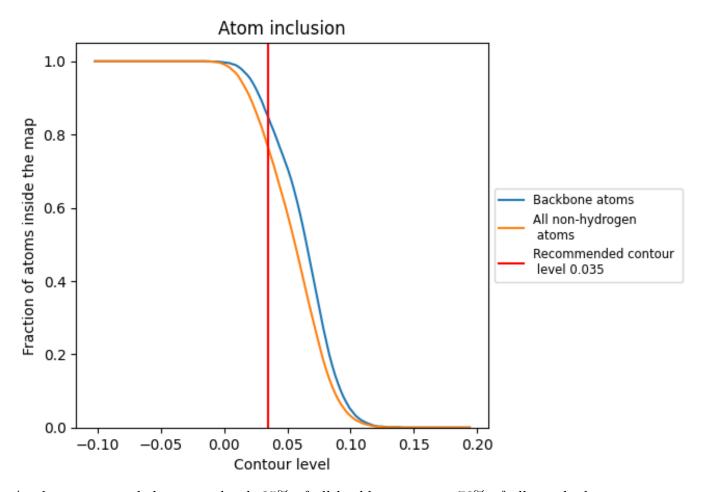
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.035).



9.4 Atom inclusion (i)



At the recommended contour level, 85% of all backbone atoms, 76% 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 (0.035) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7630	0.5750
A	0.7440	0.5800
В	0.7690	0.5800
С	0.7720	0.5850
D	0.7700	0.5760
Е	0.7640	0.5780
F	0.7600	0.5720
G	0.7610	0.5720
Н	0.7670	0.5810
I	0.7600	0.5800
a	0.7770	0.5690
b	0.7630	0.5690
С	0.7760	0.5740
d	0.7860	0.5720
e	0.7740	0.5640
f	0.7830	0.5690
g	0.7980	0.5760
h	0.7940	0.5790
i	0.7800	0.5750
j	0.2950	0.5180
k	0.2530	0.5180
1	0.7860	0.5770



