



wwPDB EM Validation Summary Report

Feb 29, 2024 – 06:15 PM EST

PDB ID : 8GCR
EMDB ID : EMD-29941
Title : HPV16 E6-E6AP-p53 complex
Authors : Bratkowski, M.A.; Wang, J.C.K.; Hao, Q.; Nile, A.H.
Deposited on : 2023-03-02
Resolution : 3.38 Å (reported)
Based on initial model : ?

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.dev70
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) : 2.36

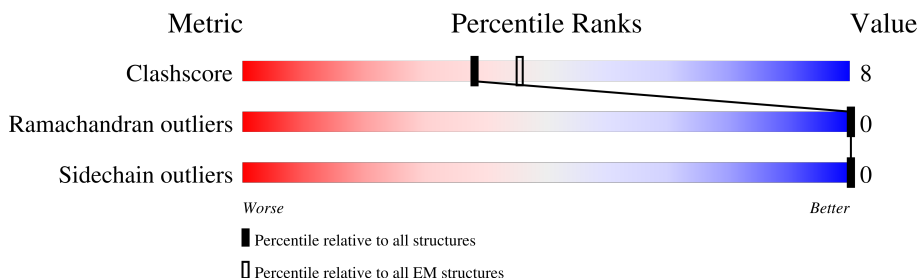
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.38 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	A	541	
2	B	220	
3	R	903	

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 7206 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 Maltose/maltodextrin-binding periplasmic protein, Protein E6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	140	1182	746	213	211	12	0	0

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-382	MET	-	initiating methionine	UNP P0AEY0
A	-11	GLY	-	linker	UNP P0AEY0
A	-10	GLY	-	linker	UNP P0AEY0
A	-9	GLY	-	linker	UNP P0AEY0
A	-8	GLY	-	linker	UNP P0AEY0
A	-7	SER	-	linker	UNP P0AEY0
A	-6	GLU	-	linker	UNP P0AEY0
A	-5	ASN	-	linker	UNP P0AEY0
A	-4	LEU	-	linker	UNP P0AEY0
A	-3	TYR	-	linker	UNP P0AEY0
A	-2	PHE	-	linker	UNP P0AEY0
A	-1	GLN	-	linker	UNP P0AEY0
A	0	GLY	-	linker	UNP P0AEY0
A	87	SER	CYS	engineered mutation	UNP P03126
A	104	SER	CYS	engineered mutation	UNP P03126
A	118	SER	CYS	engineered mutation	UNP P03126
A	147	SER	CYS	engineered mutation	UNP P03126

- Molecule 2 is a protein called Cellular tumor antigen p53.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	176	1410	877	262	256	15	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	93	GLY	-	expression tag	UNP P04637

- Molecule 3 is a protein called Ubiquitin-protein ligase E3A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	R	561	4611	2956	756	872	27	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	876	GLU	-	expression tag	UNP Q05086
R	877	ASN	-	expression tag	UNP Q05086
R	878	LEU	-	expression tag	UNP Q05086
R	879	TYR	-	expression tag	UNP Q05086
R	880	PHE	-	expression tag	UNP Q05086
R	881	GLN	-	expression tag	UNP Q05086
R	882	GLY	-	expression tag	UNP Q05086
R	883	HIS	-	expression tag	UNP Q05086
R	884	HIS	-	expression tag	UNP Q05086
R	885	HIS	-	expression tag	UNP Q05086
R	886	HIS	-	expression tag	UNP Q05086
R	887	HIS	-	expression tag	UNP Q05086
R	888	HIS	-	expression tag	UNP Q05086
R	889	GLY	-	expression tag	UNP Q05086
R	890	LEU	-	expression tag	UNP Q05086
R	891	ASN	-	expression tag	UNP Q05086
R	892	ASP	-	expression tag	UNP Q05086
R	893	ILE	-	expression tag	UNP Q05086
R	894	PHE	-	expression tag	UNP Q05086
R	895	GLU	-	expression tag	UNP Q05086
R	896	ALA	-	expression tag	UNP Q05086
R	897	GLN	-	expression tag	UNP Q05086
R	898	LYS	-	expression tag	UNP Q05086
R	899	ILE	-	expression tag	UNP Q05086
R	900	GLU	-	expression tag	UNP Q05086
R	901	TRP	-	expression tag	UNP Q05086
R	902	HIS	-	expression tag	UNP Q05086
R	903	GLU	-	expression tag	UNP Q05086

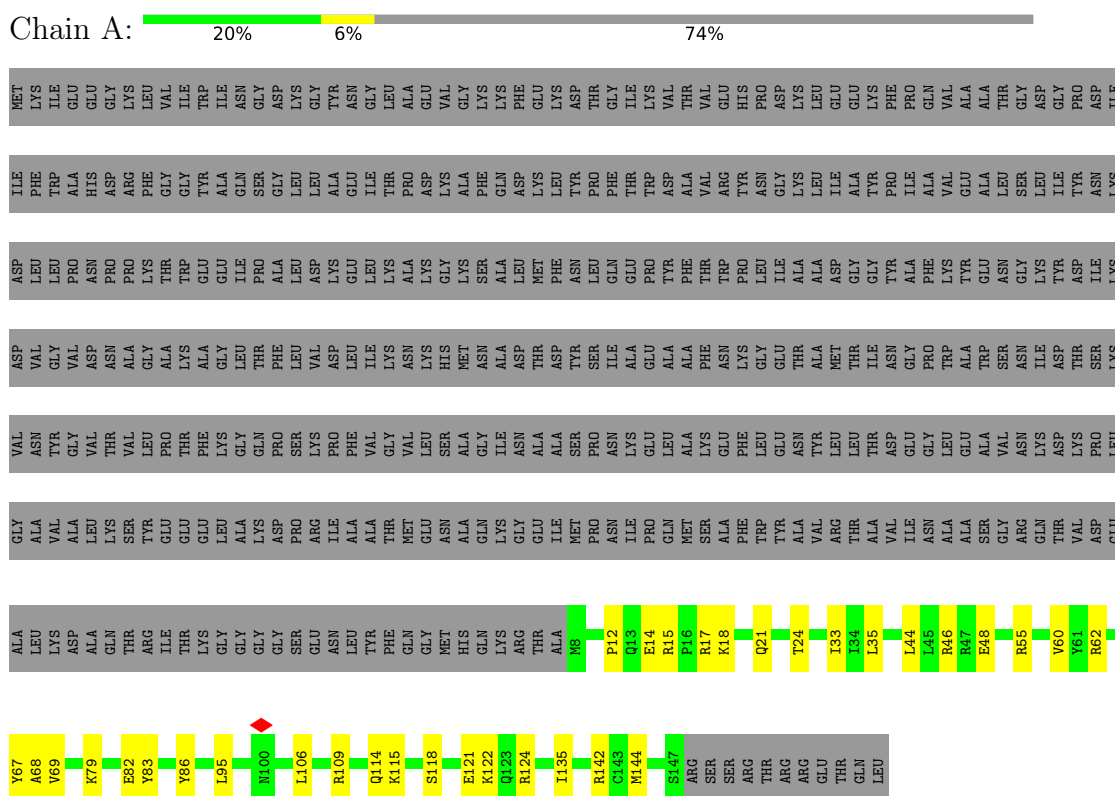
- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
4	A	2	Total 2	Zn 2	0
4	B	1	Total 1	Zn 1	0

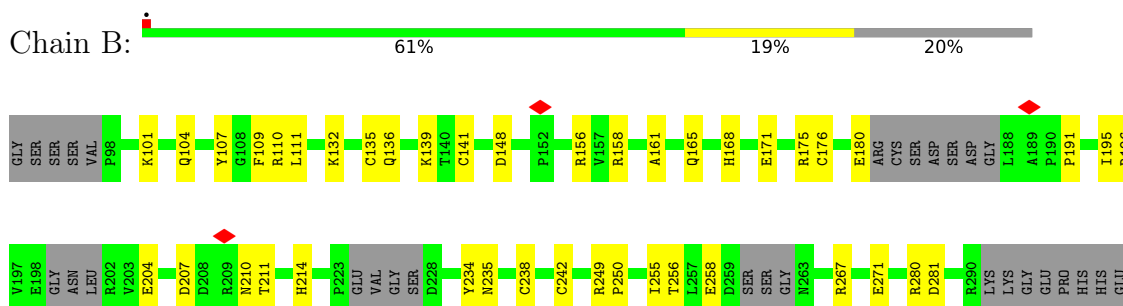
3 Residue-property plots [i](#)

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: Maltose/maltodextrin-binding periplasmic protein, Protein E6



- Molecule 2: Cellular tumor antigen p53



LEU
PRO
PRO
GLY
SER
THR
LYS
ARG
ALA
LEU
PRO
ASN
ASN
THR

• Molecule 3: Ubiquitin-protein ligase E3A



MET
GLU
LYS
LEU
HIS
GLN
CYS
TYR
ALA
TRP
LYS
SER
GLY
LEU
PRO
GLN
SER
ASP
ASP
ASP
ILE
GLU
ALA
SER
ARG
MET
PRO
LYS
ARG
ALA
ALA
SER
LEU
LYS
ALA
ALA
LYS
LYS
HIS
LEU
LEU
LEU

LEU
ARG
MET
ASP
VAL
ASN
ASN
ALA
ALA
ILE
LEU
LYS
ALA
LEU
GLU
LEU
TYR
LYS
SER
ASP
ASP
ASN
ASN
ILE
ALA
GLY
LEU
LEU
CYS
ARG
HIS
PRO
ALA
SER
LYS
LYS
LYS
GLY
ALA
ASP
LEU
ILE

ASP
PHE
LYS
VAL
T126
Y127
L128
I136
L137
E138
I139
C140
R141
R143
Y146
S147
P148
L149
V152
V156
F157
L163
V170
LYS
GLN
HIS
THR
LYS
TYR
LEU
GLU
LEU
ASN
SER
LYS
GLY
ALA
PRO
ASN
ASN
ASP
GLU
ASP
ASP
LYS
ASP
ASP
ASP
GLU
ILE
ASP
GLU
MET
PHE
LYS
ASN
LYS
LYS
LYS
ALA
ALA
GLY
CYS
PRO
THR
PHE

ALA
ALA
MET
GLU
ASP
SER
GLU
ALA
SER
SER
SER
ARG
ILE
GLY
SER
SER
SER
GLN
GLY
ASP
ASN
ASN
LEU
GLN
LYS
LYS
HIS
GLY
ASN
PRO
ASP
VAL
V233
V234
D237
L246
E250
K251
I252
E253
L260
L263
V267
D270
L286
F287
I288
L296
L306

P307
L308
F309
P317
Q321
R326
L346
I347
T348
D364
D365
D366
A367
I368
K376
G386
GLU
VAL
ASP
THR
ASN
HIS
ASN
ASN
E394
E399
E403
L407
T408
L409
Q410
E411
L412
R417
R418
M419
K420
K421
V425
D426
P427
L428
E429
T430
E431
L432
GLY
VAL
LYS

THR
LEU
ASP
C439
P442
F446
M451
E452
P453
L454
M455
M460
D463
M478
F482
K489
D496
M497
I499
R500
M501
Q515
G516
Q517
M520
P521
Y522
L523
R524
L525
K526
R529
L540
K553
Q554
E558
F559
E560
G561
E562
Q563
G564
V565
D566

E567
K572
E573
Q576
L577
V578
D593
K597
F601
N602
P603
E609
F612
V618
I623
I628
K640
L641
M642
L662
E669
G670
M671
S683
Q684
T685
D686
L687
F688
G689
N690
P691
Y694
D701
E712
F713
V714
N715
L716
Y717
D719
Y720

M723
K724
S725
A732
F733
R734
L749
P752
E753
E754
I759
C760
G761
SER
ARG
ASN
LYS
MET
ASP
ASP
PHE
GLN
ALA
LYS
ASN
GLY
GLU
ASP
PRO
THR
THR
THR
GLU
TYR
ASP
GLY
PHE
THR
SER
SER
HIS
LYS
PHE
LYS
THR
THR
ASP
VAL
PHE
THR
ASP
GLU
GLN
LYS
ARG

LEU
PHE
LEU
GLN
THR
PHE
THR
GLY
THR
THR
ASP
ARG
ALA
PRO
VAL
GLY
GLY
LEU
PHE
GLN
LYS
LEU
HIS
LYS
MET
ILE
ILE
HIS
HIS
HIS
GLY
ALA
ASN
GLY
ASP
ILE
PHE
GLU
ALA
ARG
LEU
PRO
GLY
TRP
ILE
VAL
HIS
SER
SER
HIS
LYS
PHE
LYS
THR
ASP
GLU
ARG
GLN
LYS
LEU
ARG

LYS
ALA
THR
TYR
ALA
LYS
PHE
GLY
MET
LEU
LEU
ASN
LEU
TYR
PHE
GLY
HIS
HIS
HIS
HIS
HIS
GLY
LEU
ASN
ASN
ASP
ILE
PHE
GLU
ALA
GLN
LEU
GLN
LYS
PRO
THR
THR
THR
GLU
ALA
ARG
LEU
PRO
GLY
THR
SER
HIS
THR
CYS
PHE
VAL
VAL
LEU
LEU
LEU
PRO
GLU
TYR
SER
SER
HIS
LYS
LYS
LYS
LYS
GLU
ARG
LEU
LEU
ARG

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	105794	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 ($e^-/\text{\AA}^2$)	80.5	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	4.580	Depositor
Minimum map value	-2.710	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.056	Depositor
Recommended contour level	0.457	Depositor
Map size (Å)	332.0, 332.0, 332.0	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.83, 0.83, 0.83	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section:
ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.32	0/1209	0.45	0/1627
2	B	0.30	0/1441	0.52	0/1948
3	R	0.30	0/4702	0.49	0/6348
All	All	0.30	0/7352	0.49	0/9923

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1182	0	1165	21	0
2	B	1410	0	1381	30	0
3	R	4611	0	4576	68	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
All	All	7206	0	7122	114	0

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

The worst 5 of 114 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)
3:R:403:GLU:HA	3:R:403:GLU:OE1	1.81	0.80
3:R:138:GLU:HA	3:R:141:ARG:HG2	1.73	0.68
3:R:489:LYS:NZ	3:R:754:GLU:OE2	2.28	0.66
3:R:719:ASP:HA	3:R:723:ASN:HD22	1.61	0.65
3:R:157:PHE:HA	3:R:163:LEU:HD11	1.79	0.64

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	138/541 (26%)	134 (97%)	4 (3%)	0	100	100
2	B	166/220 (76%)	162 (98%)	4 (2%)	0	100	100
3	R	553/903 (61%)	523 (95%)	30 (5%)	0	100	100
All	All	857/1664 (52%)	819 (96%)	38 (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	
1	A	134/456 (29%)	134 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	160/196 (82%)	160 (100%)	0	100	100
3	R	517/810 (64%)	517 (100%)	0	100	100
All	All	811/1462 (56%)	811 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	179	HIS

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 3 ligands modelled in this entry, 3 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](#)

There are no chain breaks in this entry.

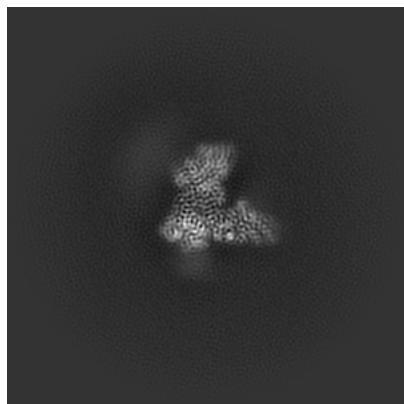
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-29941. These allow visual inspection of the internal detail of the map and identification of artifacts.

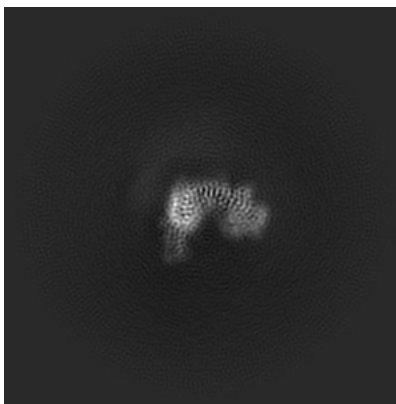
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

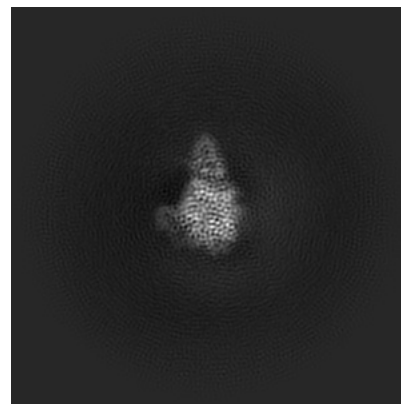
6.1.1 Primary map



X

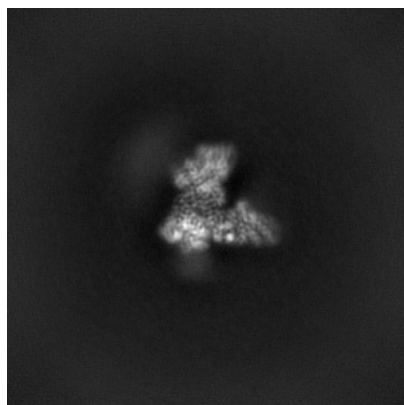


Y

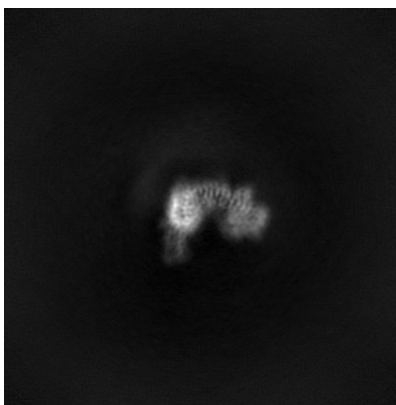


Z

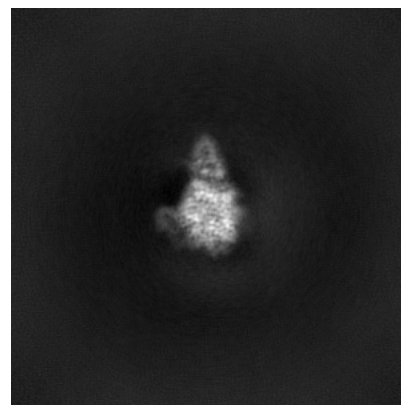
6.1.2 Raw map



X



Y

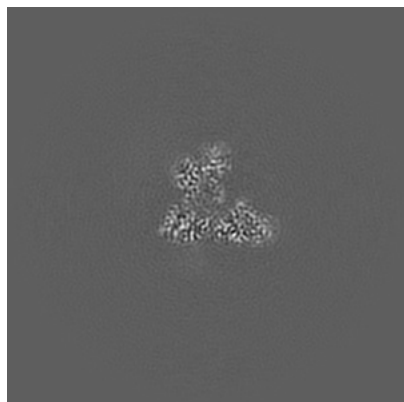


Z

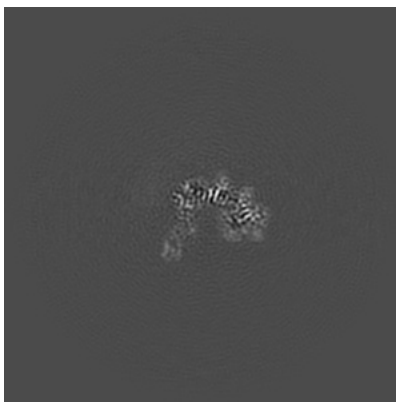
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

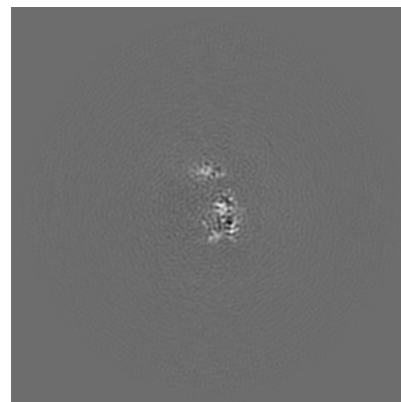
6.2.1 Primary map



X Index: 200

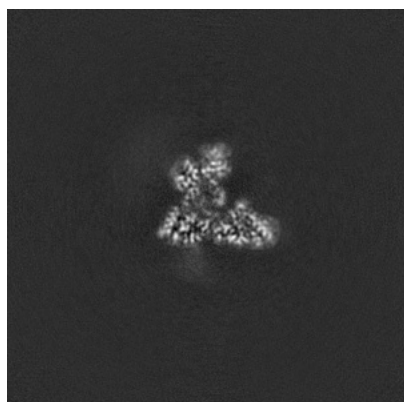


Y Index: 200

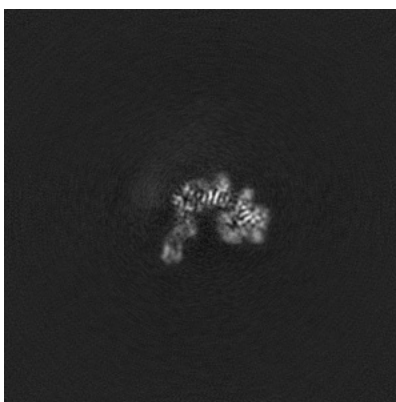


Z Index: 200

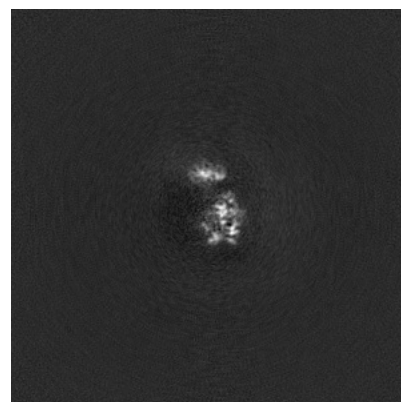
6.2.2 Raw map



X Index: 200



Y Index: 200

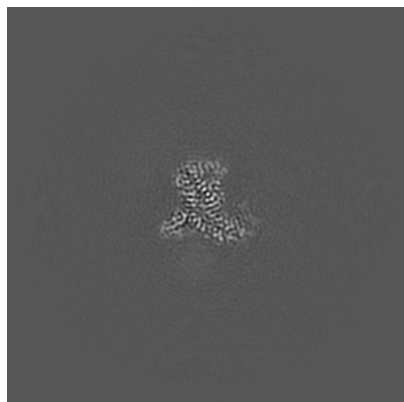


Z Index: 200

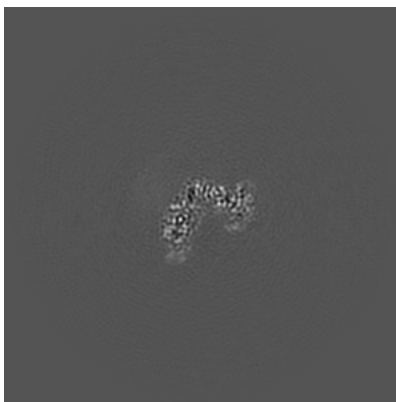
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

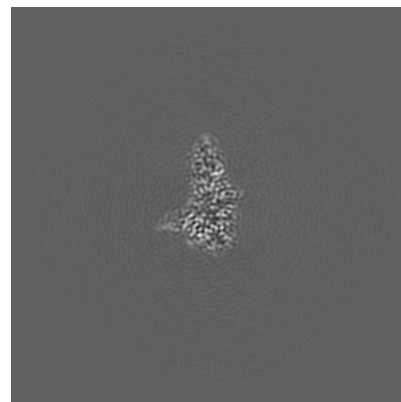
6.3.1 Primary map



X Index: 213

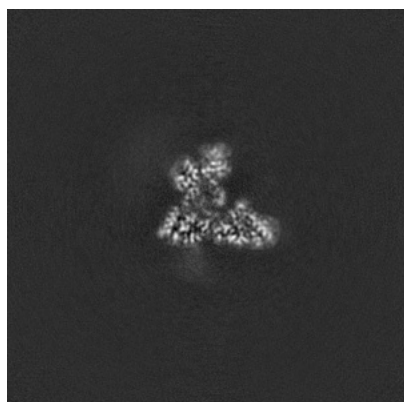


Y Index: 185

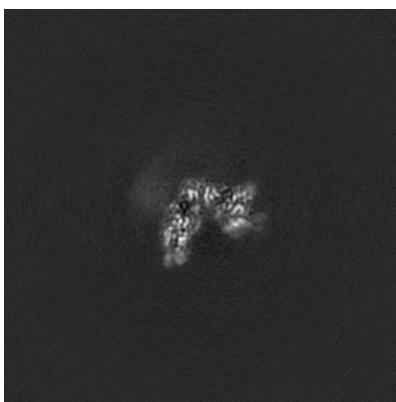


Z Index: 182

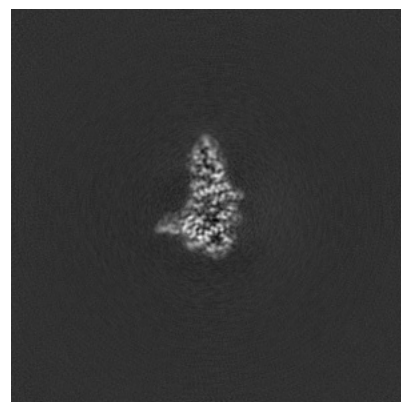
6.3.2 Raw map



X Index: 200



Y Index: 187

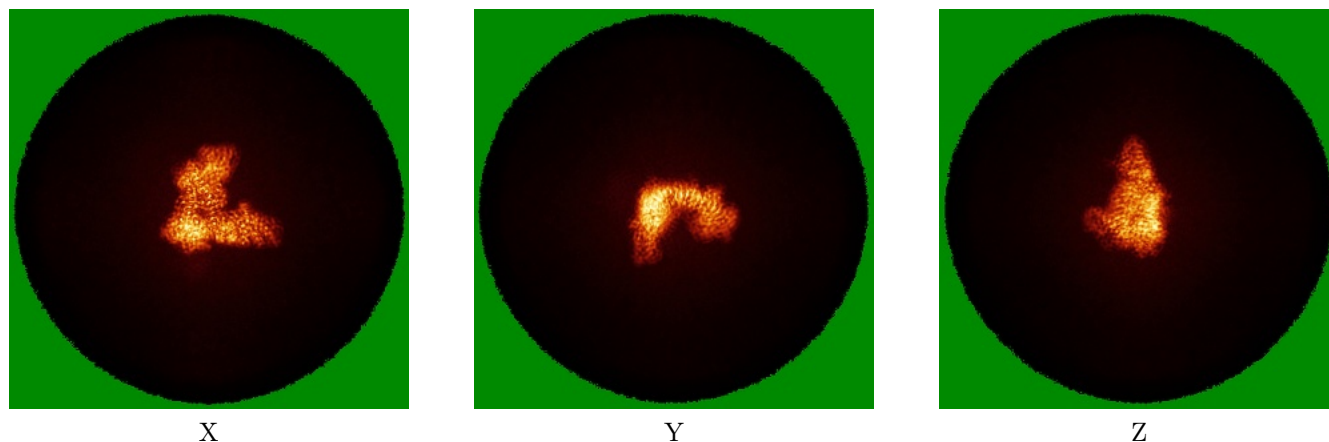


Z Index: 182

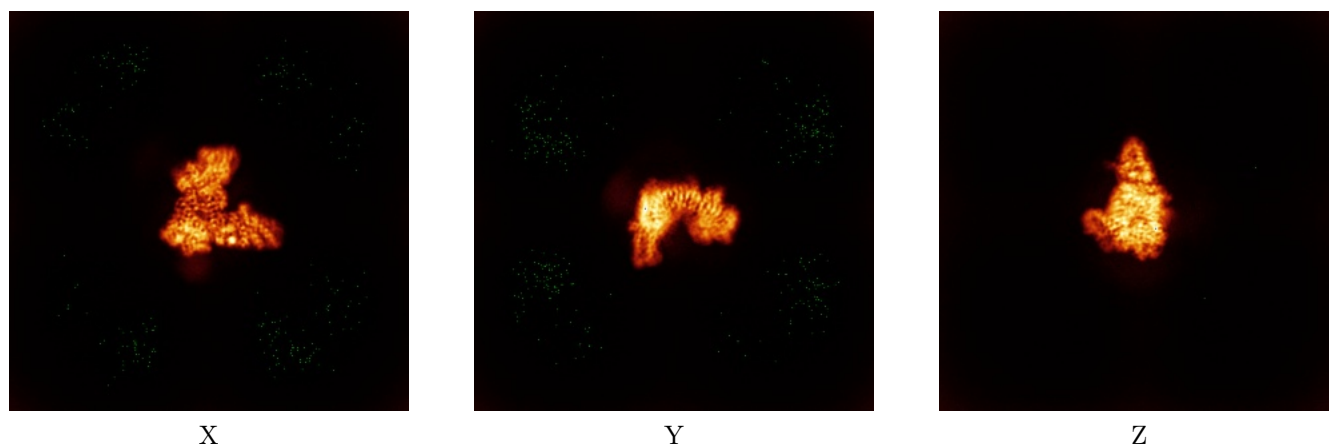
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [i](#)

6.4.1 Primary map



6.4.2 Raw map



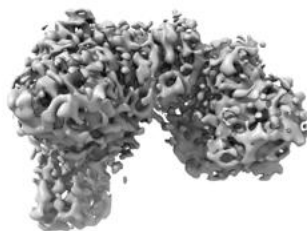
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



X



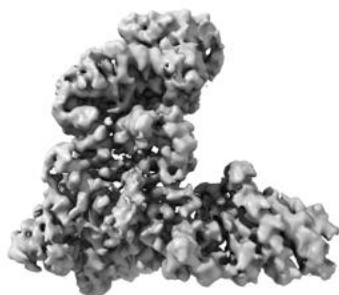
Y



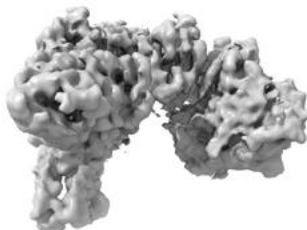
Z

The images above show the 3D surface view of the map at the recommended contour level 0.457. 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



X



Y



Z

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.

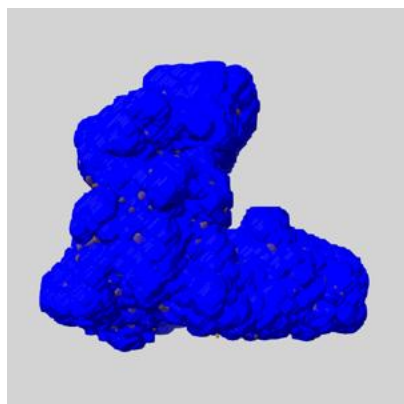
6.6 Mask visualisation [i](#)

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

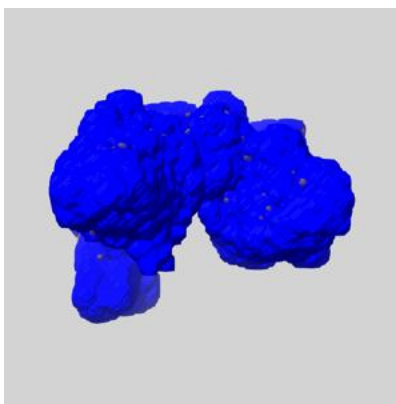
A mask typically either:

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

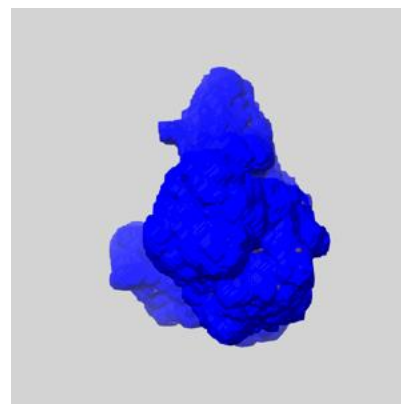
6.6.1 emd_29941_msk_1.map [i](#)



X



Y

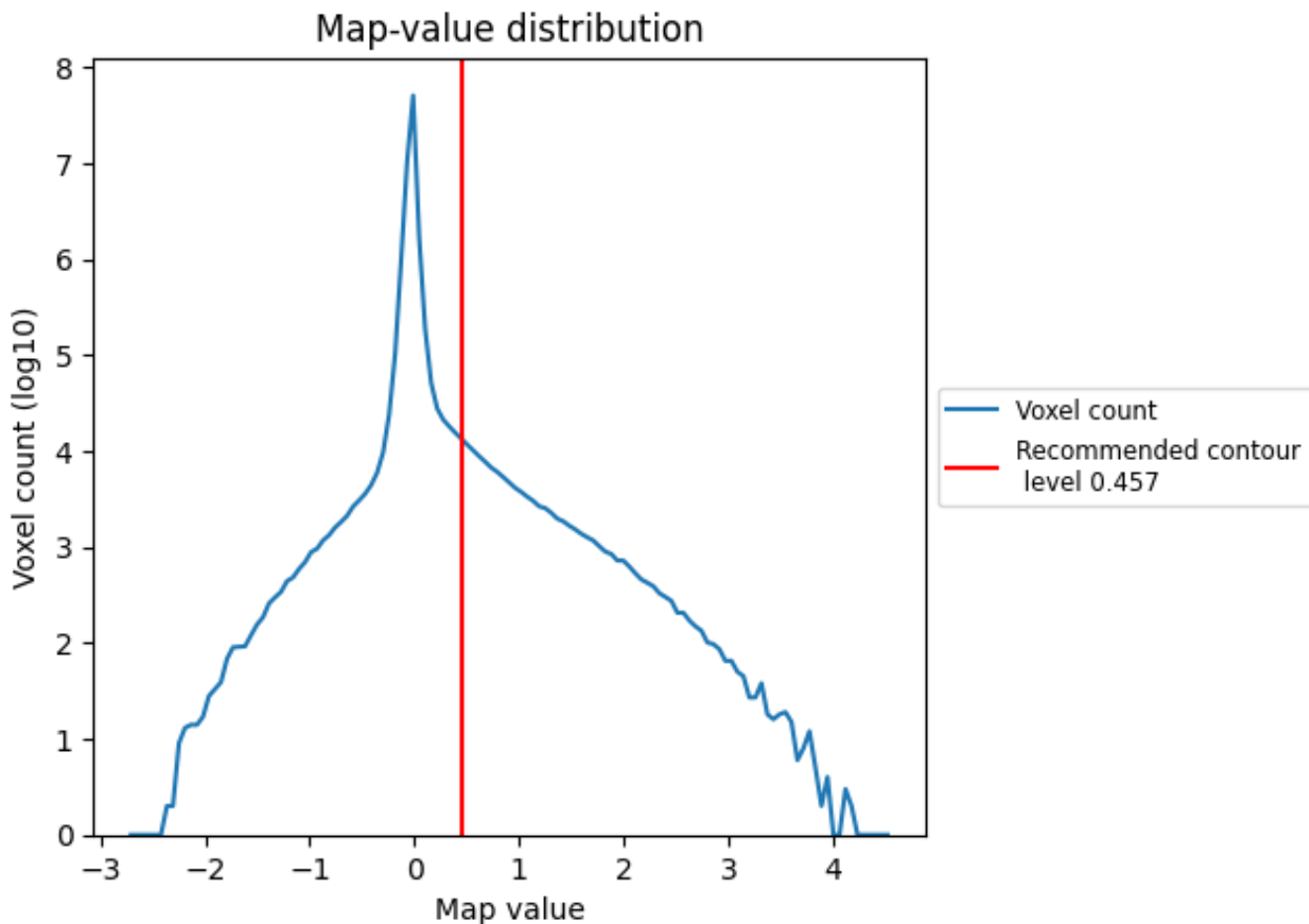


Z

7 Map analysis [i](#)

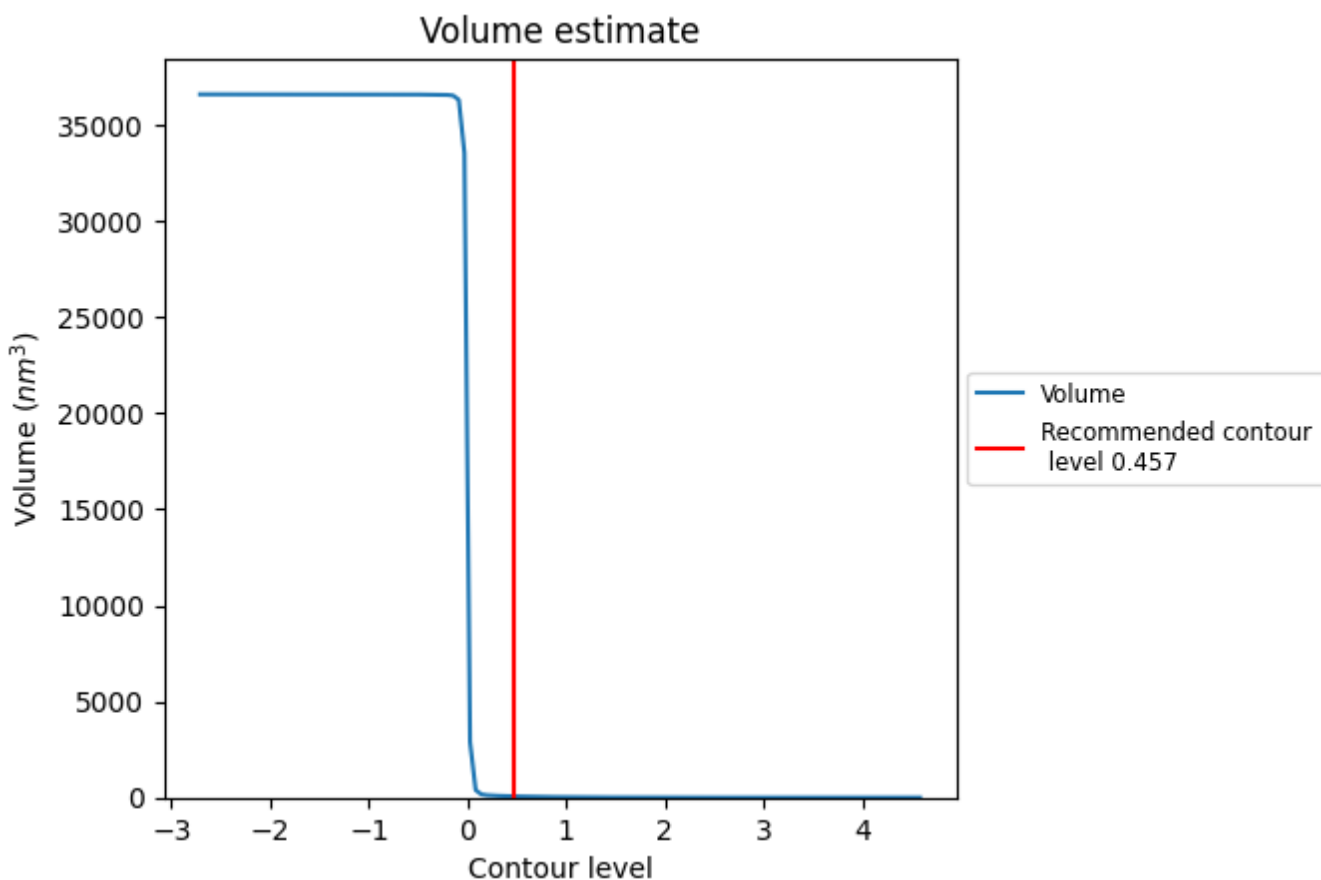
This section contains the results of statistical analysis of the map.

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.

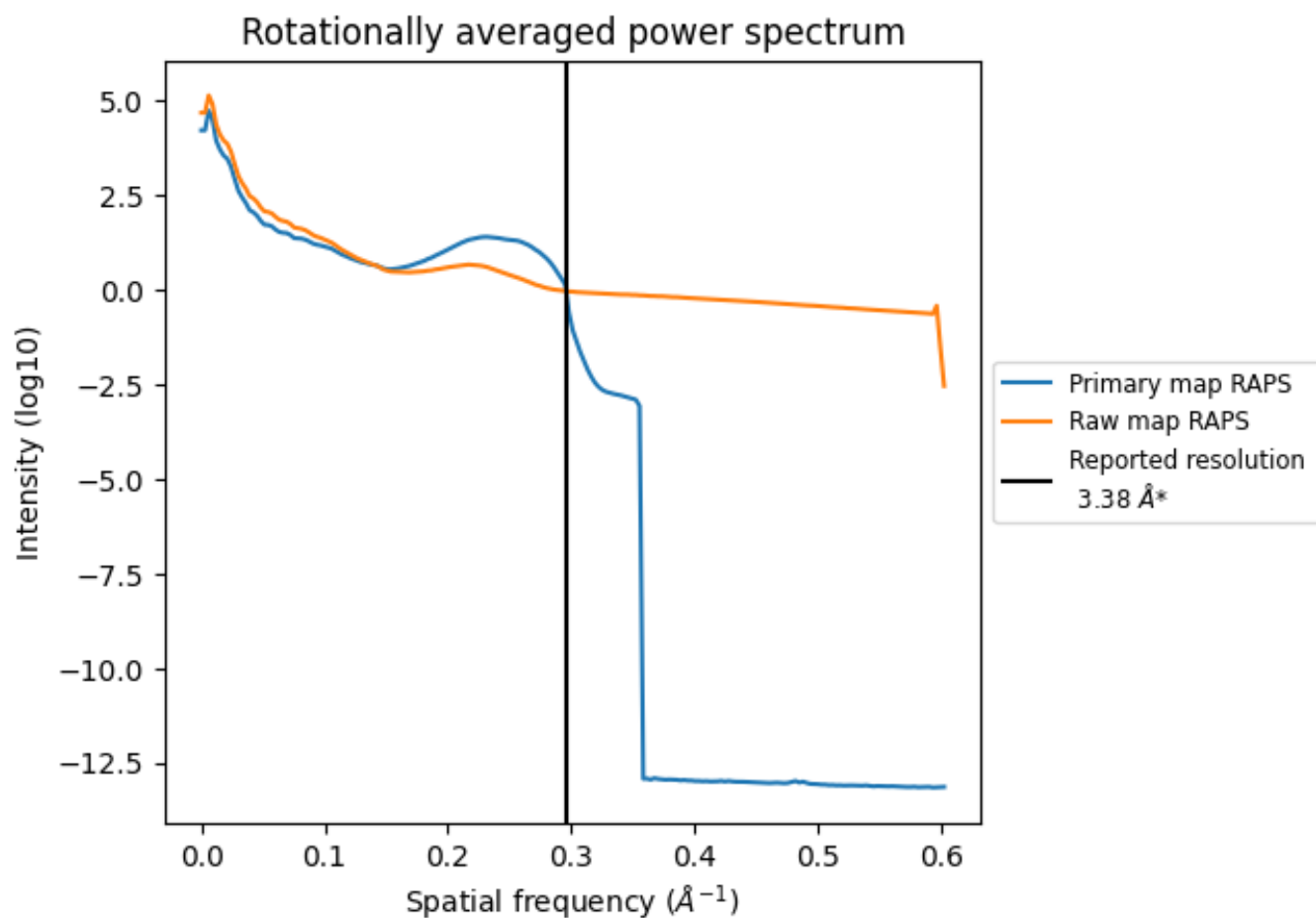
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 65 nm³; this corresponds to an approximate mass of 59 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](#)

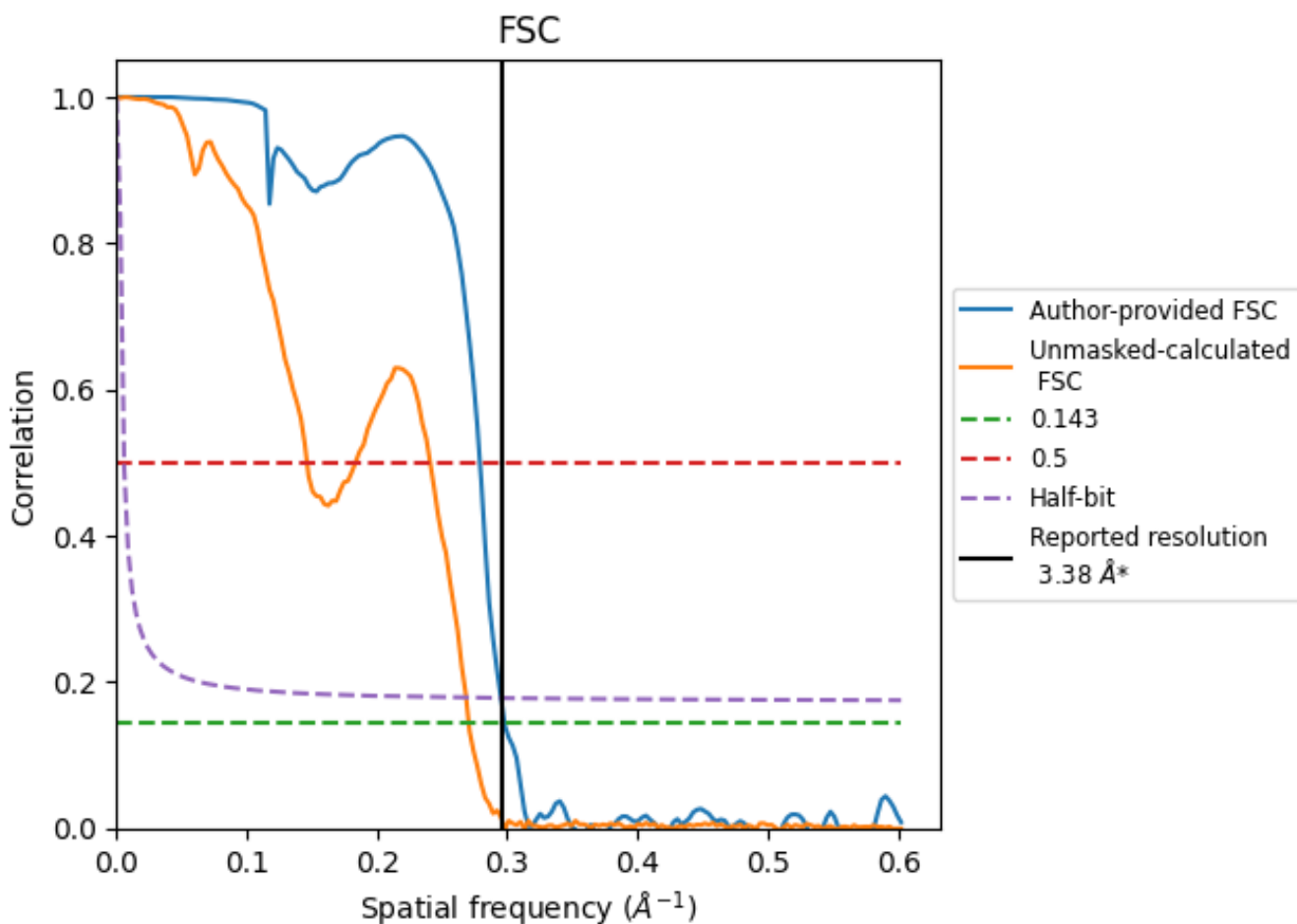


*Reported resolution corresponds to spatial frequency of 0.296 Å⁻¹

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 0.296 Å⁻¹

8.2 Resolution estimates [i](#)

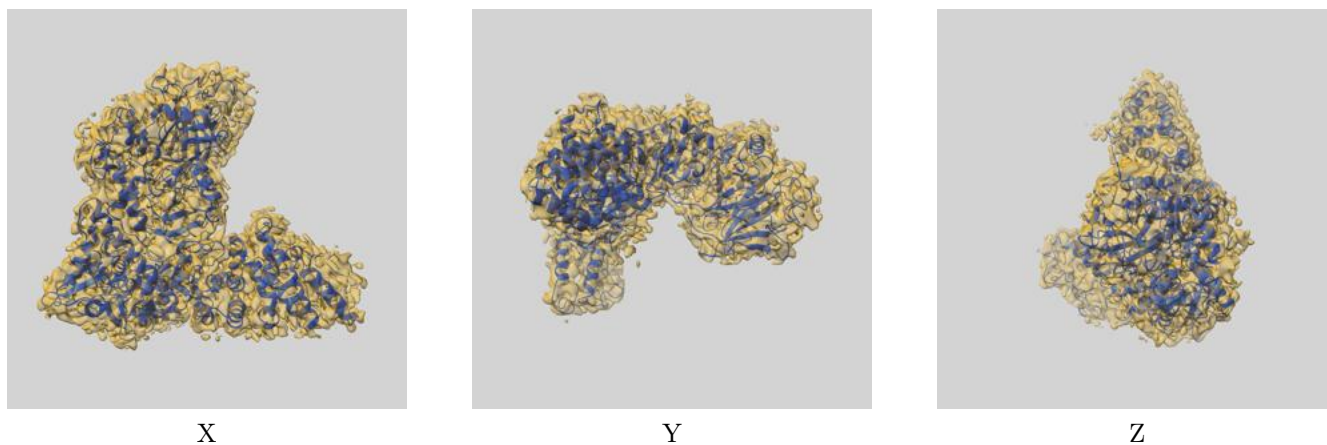
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.38	-	-
Author-provided FSC curve	3.36	3.58	3.39
Unmasked-calculated*	3.70	6.84	3.73

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

9 Map-model fit [i](#)

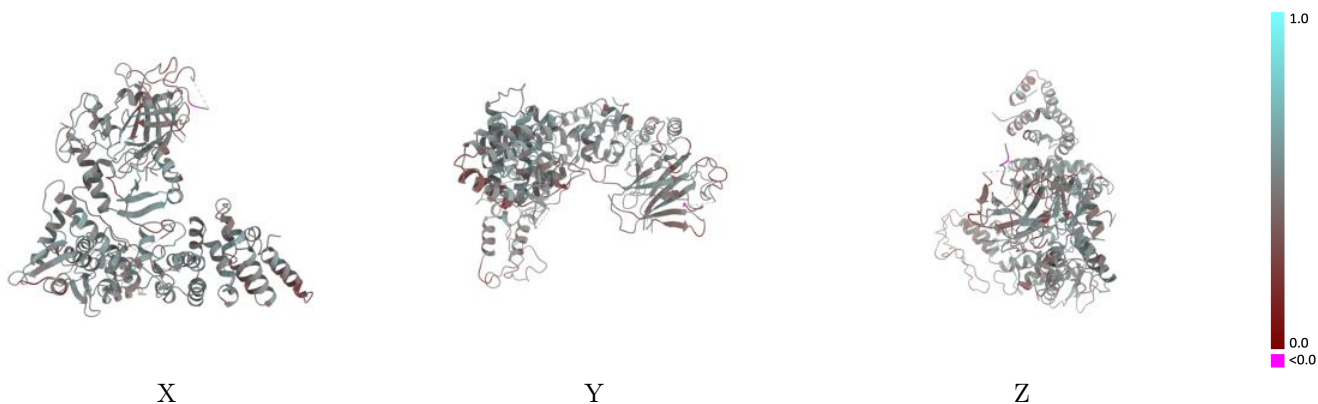
This section contains information regarding the fit between EMDB map EMD-29941 and PDB model 8GCR. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay [i](#)



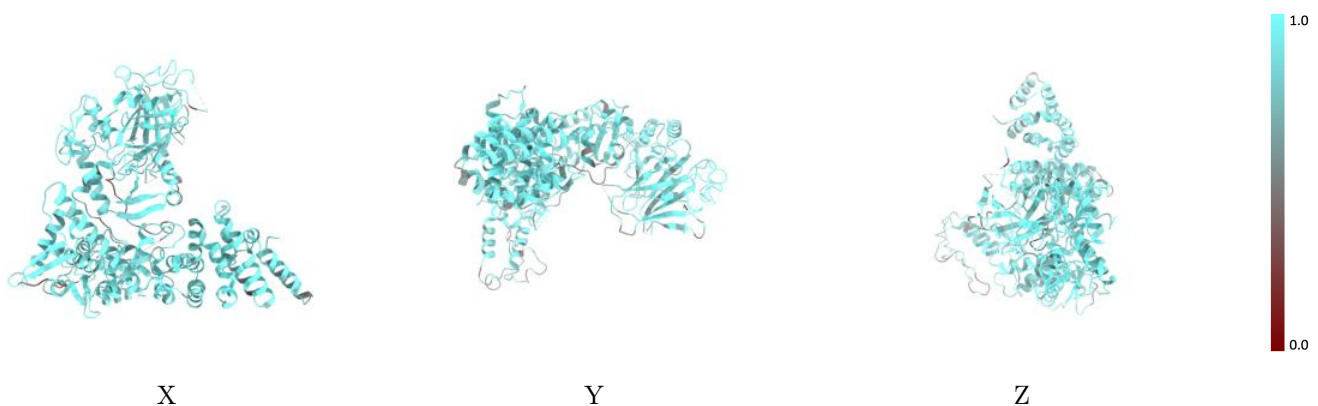
The images above show the 3D surface view of the map at the recommended contour level 0.457 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



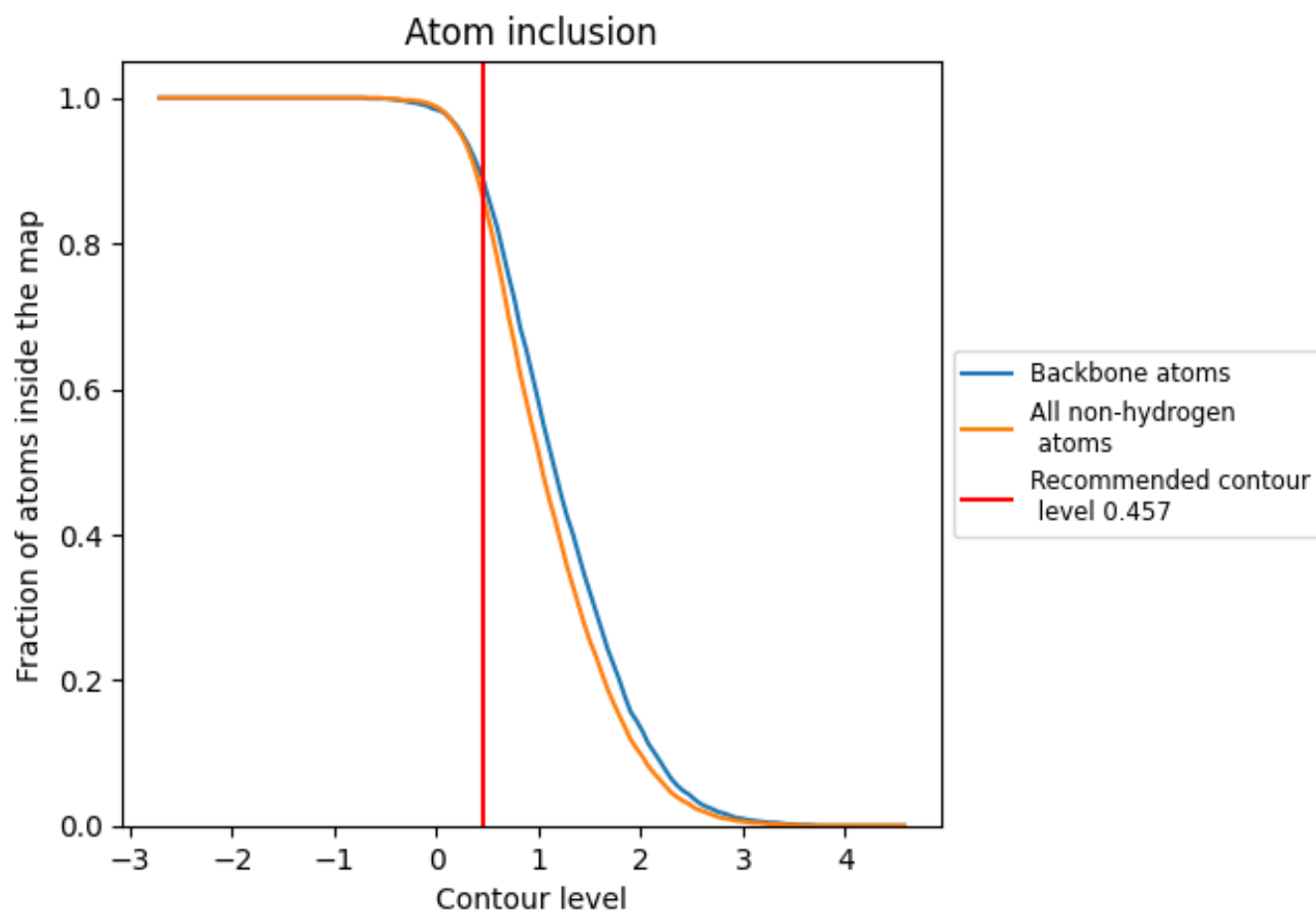
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.457).




9.4 Atom inclusion [i](#)



At the recommended contour level, 89% of all backbone atoms, 86% 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 (0.457) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8630	 0.4680
A	 0.8920	 0.5020
B	 0.8710	 0.4420
R	 0.8530	 0.4680

