wwPDB EM Validation Summary Report (i)

Jul 24, 2023 – 08:25 PM JST

PDB ID	:	7YJL
EMDB ID	:	EMD-31529
Title	:	Melbournevirus major capsid protein built from 4.4A reconstruction
Authors	:	Burton-Smith, R.N.; Okamoto, K.; Murata, K.
Deposited on	:	2022-07-20
Resolution	:	4.42 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 (i)) were used in the production of this report:

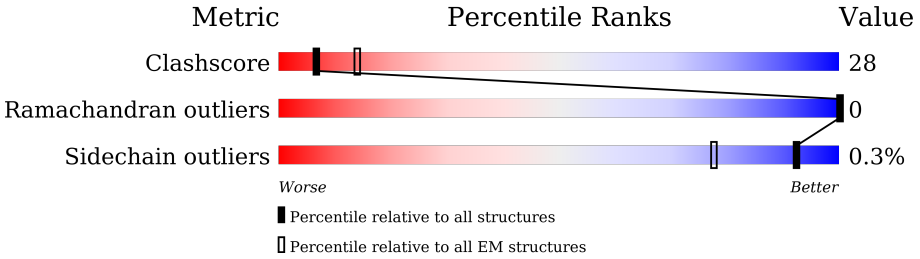
EMDB validation analysis	:	0.0.1.dev50
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

1 Overall quality at a glance (i)

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

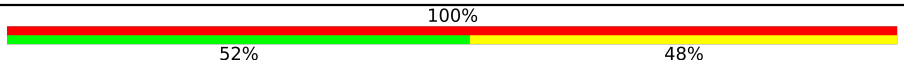
The reported resolution of this entry is 4.42 Å.

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 <=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	459	

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 7083 atoms, of which 3484 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.

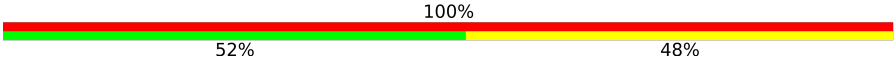
Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
1	A	459	7083	2302	3484	605	683	9	0	0

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

Chain A: 100%



V10 T11 A12 A13 S14 G15 F16 V17 D18 L19 A20 T21 F22 S23 D24 L25 E26 A27 Y28 L29 Y30 G31 G32 C33 S34 A35 V36 T37 Y38 F39 V40 R41 A42 I43 K44 K45 A46 N47 W48 F49 S50 F51 L52 P53 V54 V55 L56 R57 N58 I59 S60 G61 L62 P63 G64 F65 G66 S67 E68 F69

S70 A71 S72 V73 N74 R75 S76 G77 D78 Y79 V80 L81 N82 T83 W84 L85 R86 V87 R88 L89 P90 L91 V92 A93 I94 R95 P96 T97 N98 T99 G100 G101 A102 I103 N104 A105 N106 A107 T108 I109 R110 W111 T112 R113 N114 F115 M116 H117 N118 L119 V120 E121 K122 V123 N124 I125 T126 F127 N128 D129

L130 I131 V132 H133 E134 F135 D136 S137 Y138 W139 F140 D141 F142 N143 S144 Q145 F146 N147 I148 D149 A150 S151 K152 R153 V154 G155 Y156 R157 N158 M159 I160 G161 D162 I163 P164 A165 M166 I167 N168 P169 V170 T171 T172 G173 N174 P175 L176 G177 T178 G179 E180 F181 F182 N183 L184 P185 I186 P187 L188 F189

Y190 T191 E192 D193 S194 G195 L196 A197 L198 A199 V200 S201 A202 L203 P204 F205 N206 D207 I208 K209 I210 N211 Y212 C213 L214 R215 R216 W217 Q218 D219 L220 I221 V222 L223 N224 V225 G226 V227 G228 G229 N230 P231 P232 T233 Y234 D235 D236 I237 V238 Q239 V240 S241 Y242 D243 S244 T245 F246 T247 L248 I249

Y250 S251 S252 N253 A254 P255 A256 I257 T258 N259 V260 E261 T262 W263 C264 H265 Y266 A267 V268 V269 H270 N271 D272 E273 R274 V275 K276 M277 G278 K279 N280 P281 R282 D283 M284 V285 I286 K287 Q288 V289 Q290 K291 V292 N293 E294 T295 T296 I297 N298 L299 S300 Q301 L302 N303 A304 L305 V306 P307 I308 D309

I310 R311 V312 S313 H314 A315 V316 V317 G318 Y319 F320 Y321 A322 I323 R324 N325 S326 S327 T328 T329 G330 E331 W332 S333 N334 Y335 T336 T337 E338 P339 A340 Y341 A342 G343 L344 D345 P346 L347 E348 A349 A350 Q351 L352 V353 Y354 E355 S356 T357 A358 R359 V360 S361 N362 G363 S364 D365 Y366 Y367 S368 L369

M370 V371 P372 W373 Y374 W375 H376 K377 S378 I379 P380 E381 E382 T383 G384 Y385 H386 A387 Y388 S389 Y390 S391 L392 D393 T394 F395 A396 S397 D398 P399 K400 G401 S402 T403 N404 Y405 S406 K407 L408 T409 N410 V411 S412 M413 Q414 Y415 V416 P417 S418 T419 A420 A421 V422 N423 A424 S425 A426 G427 V428 T429

N430 T431 G432 I433 P434 I435 P436 S437 A438 T439 N440 P441 A442 V443 T444 Q445 Q446 N447 Q448 T449 F450 Q451 H452 I453 F454 R455 V456 L457 N458 F459 N460 V461 L462 R463 L464 S465 G466 G467 S468

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	187440	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 ($e^-/Å^2$)	26.4	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.093	Depositor
Minimum map value	-0.060	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.017	Depositor
Map size (Å)	972.4, 972.4, 972.4	wwPDB
Map dimensions	440, 440, 440	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.21, 2.21, 2.21	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

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	0.45	0/3696	0.53	0/5057

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	A	3599	3484	3483	196	0
All	All	3599	3484	3483	196	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 28.

The worst 5 of 196 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:A:57:ARG:NH2	1:A:58:ASN:O	1.94	0.99
1:A:196:LEU:HD22	1:A:198:LEU:HD22	1.45	0.94
1:A:75:ARG:NH1	1:A:202:ALA:O	2.03	0.90
1:A:243:ASP:OD1	1:A:247:THR:OG1	1.89	0.90
1:A:200:VAL:HG21	1:A:277:MET:HB2	1.54	0.89

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	457/459 (100%)	356 (78%)	101 (22%)	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	394/394 (100%)	393 (100%)	1 (0%)	92 95

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	359	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	446	GLN
1	A	414	GLN
1	A	280	ASN
1	A	218	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	325	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)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

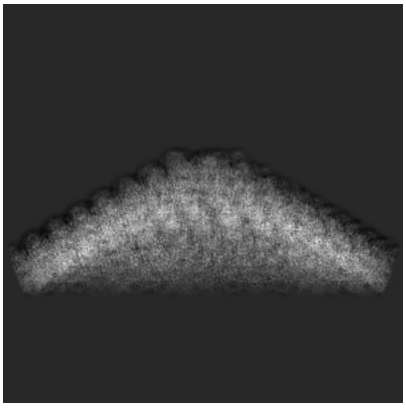
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-31529. These allow visual inspection of the internal detail of the map and identification of artifacts.

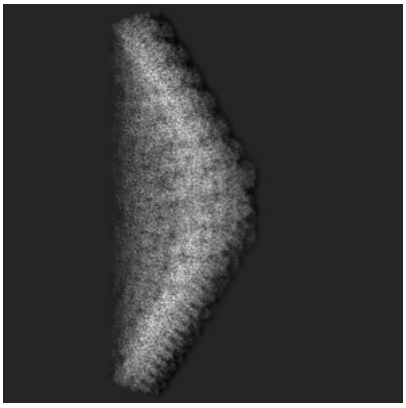
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

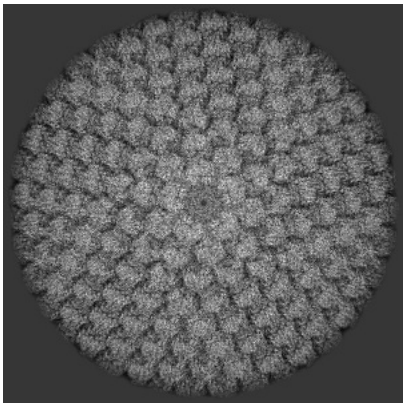
6.1.1 Primary map



X



Y

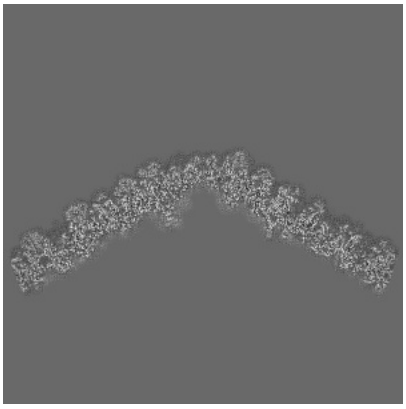


Z

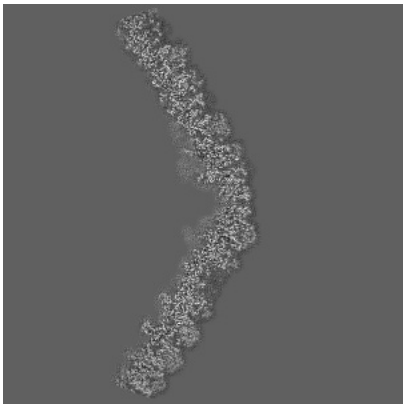
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

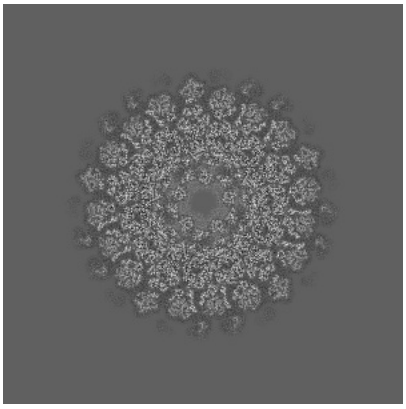
6.2.1 Primary map



X Index: 220



Y Index: 220

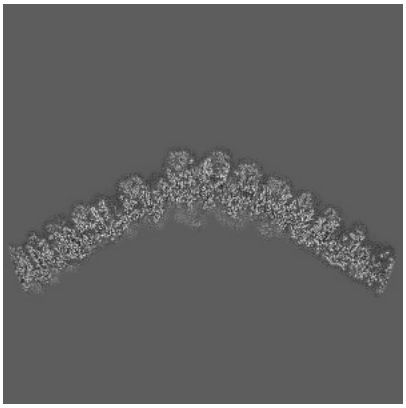


Z Index: 220

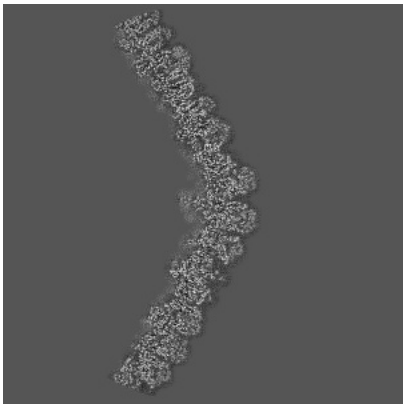
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

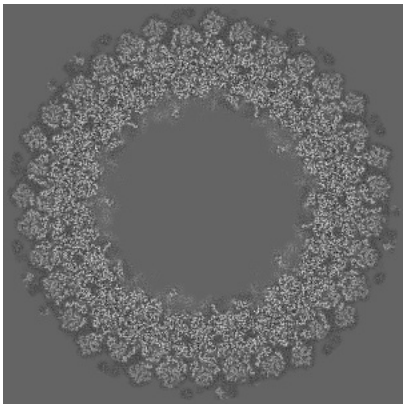
6.3.1 Primary map



X Index: 247



Y Index: 187

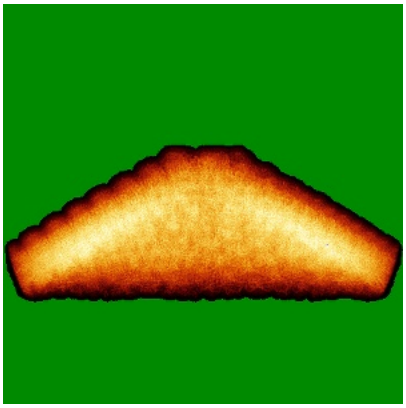


Z Index: 173

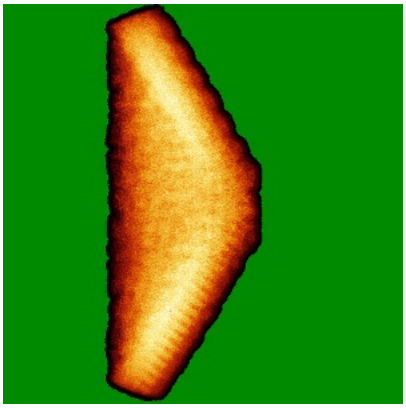
The images above show the largest variance slices of the map in three orthogonal directions.

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

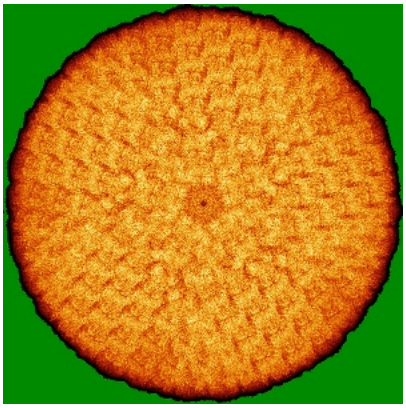
6.4.1 Primary map



X



Y

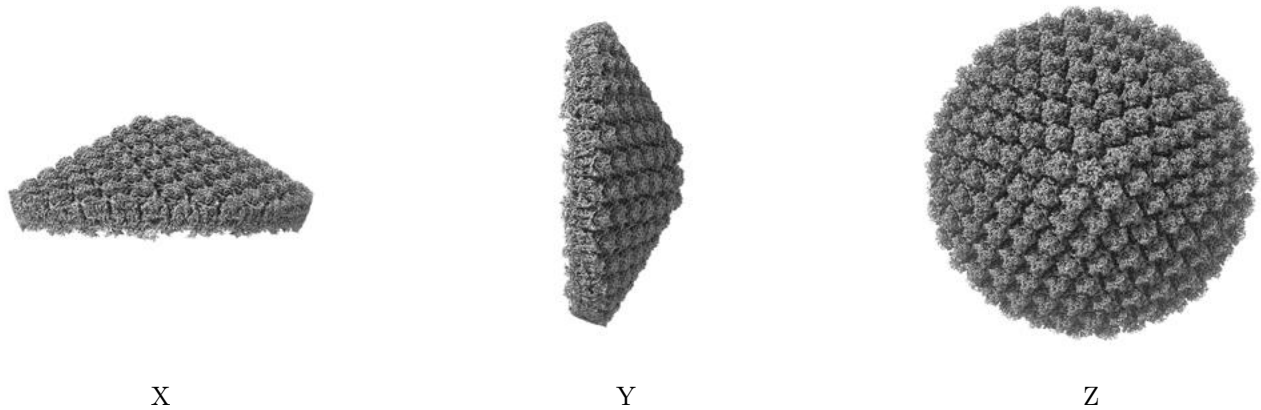


Z

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

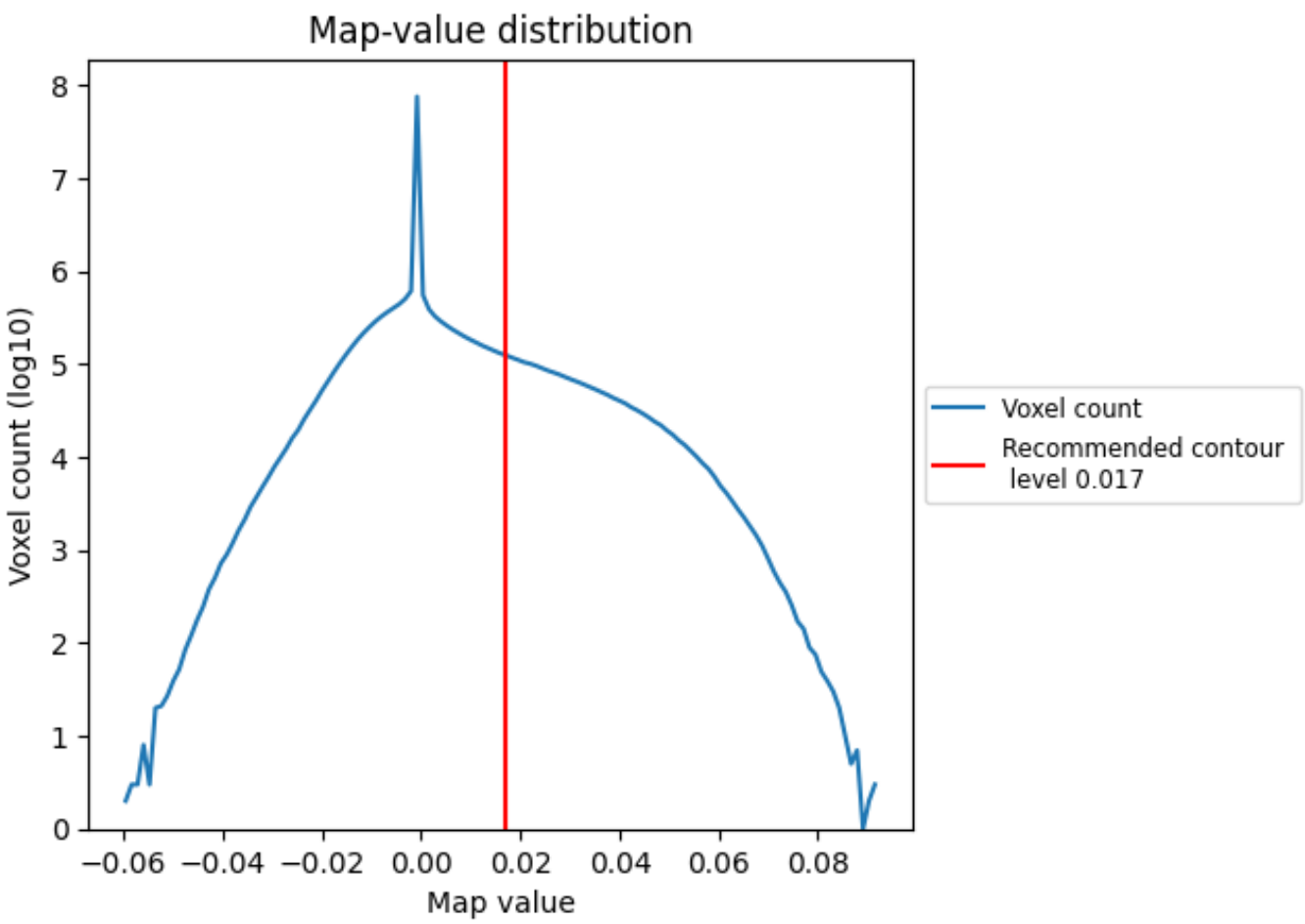
6.6 Mask visualisation (i)

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

7 Map analysis (i)

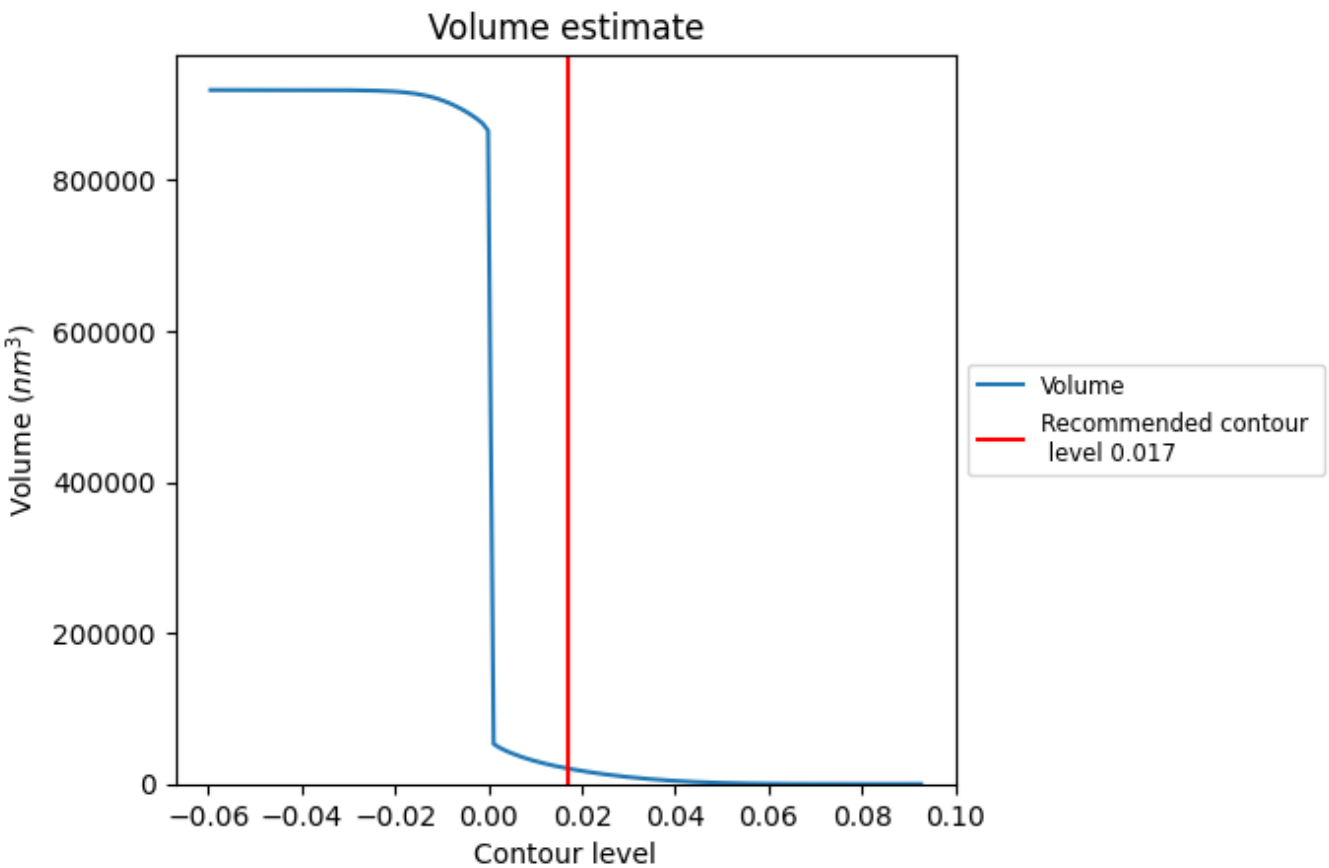
This section contains the results of statistical analysis of the map.

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.

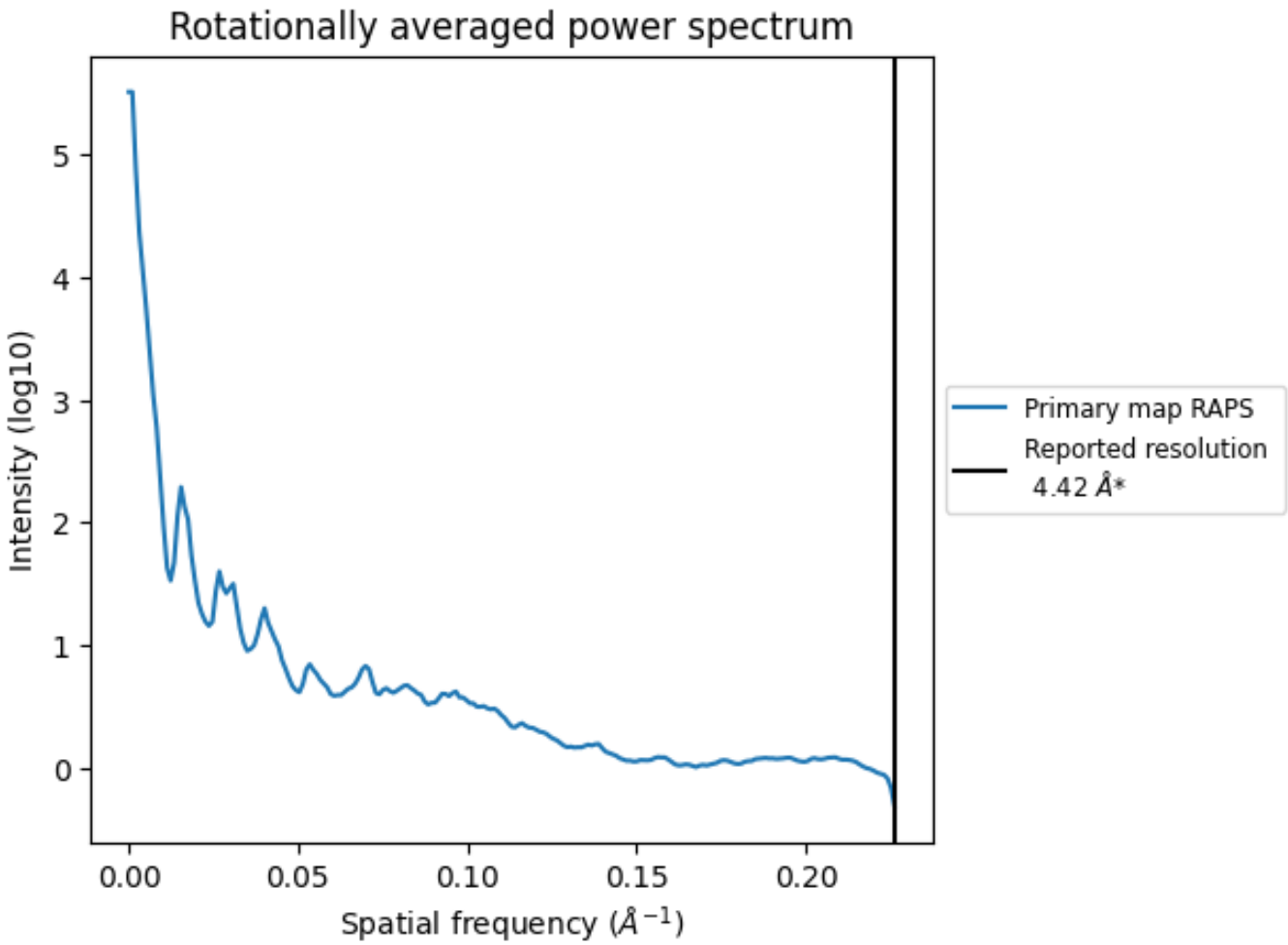
7.2 Volume estimate (i)



The volume at the recommended contour level is 20522 nm^3; this corresponds to an approximate mass of 18539 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)

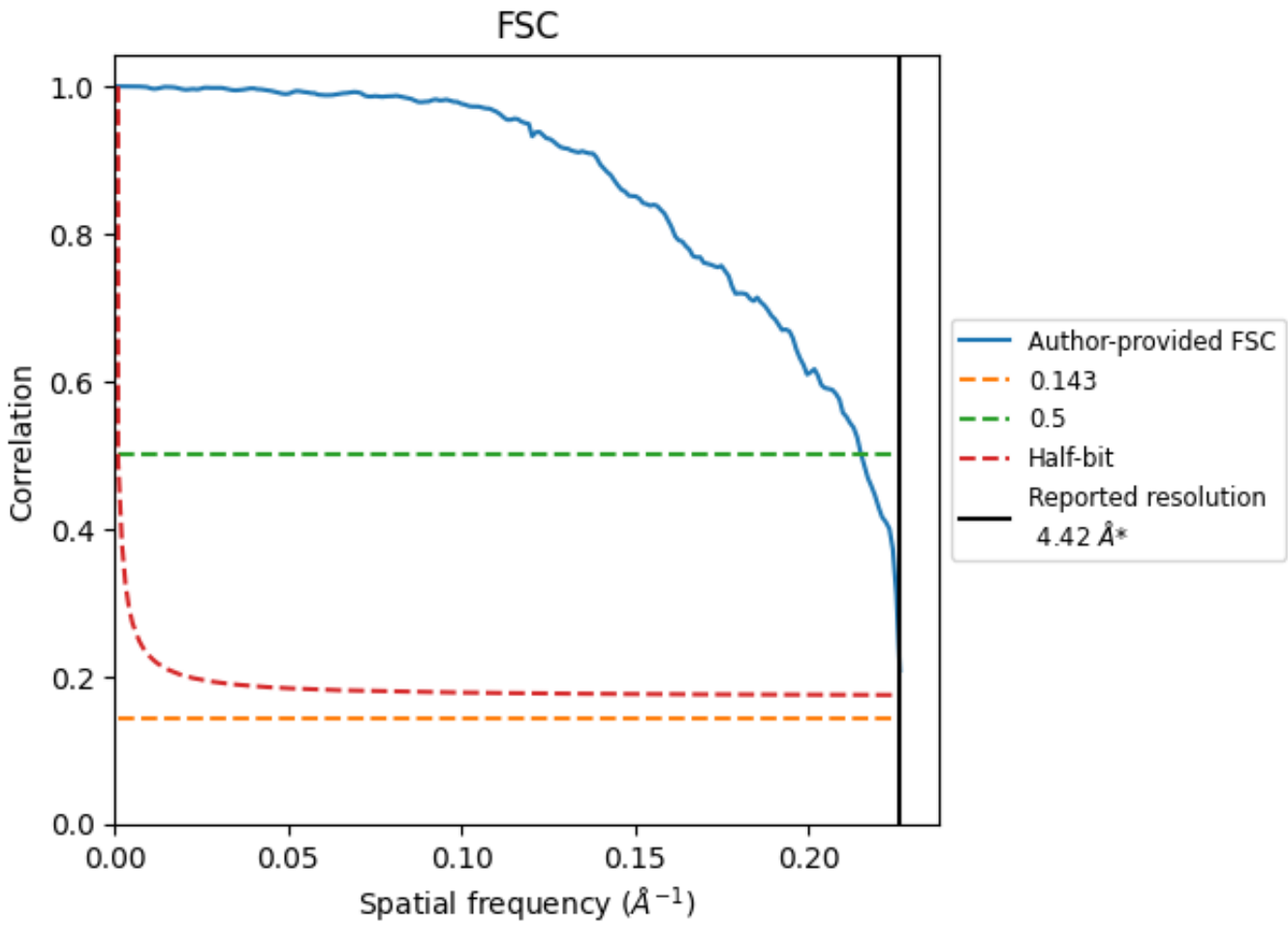


*Reported resolution corresponds to spatial frequency of 0.226 Å^{-1}

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 0.226 Å^{-1}

8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.42	-	-
Author-provided FSC curve	-	4.65	-
Unmasked-calculated*	-	-	-

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

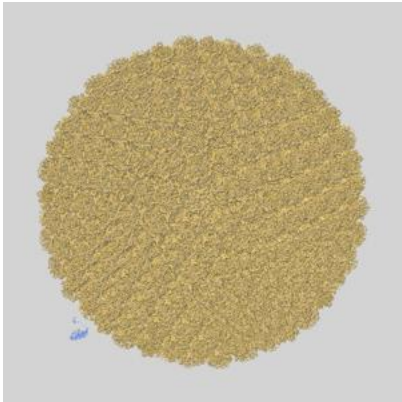
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-31529 and PDB model 7YJL. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlay (i)

X

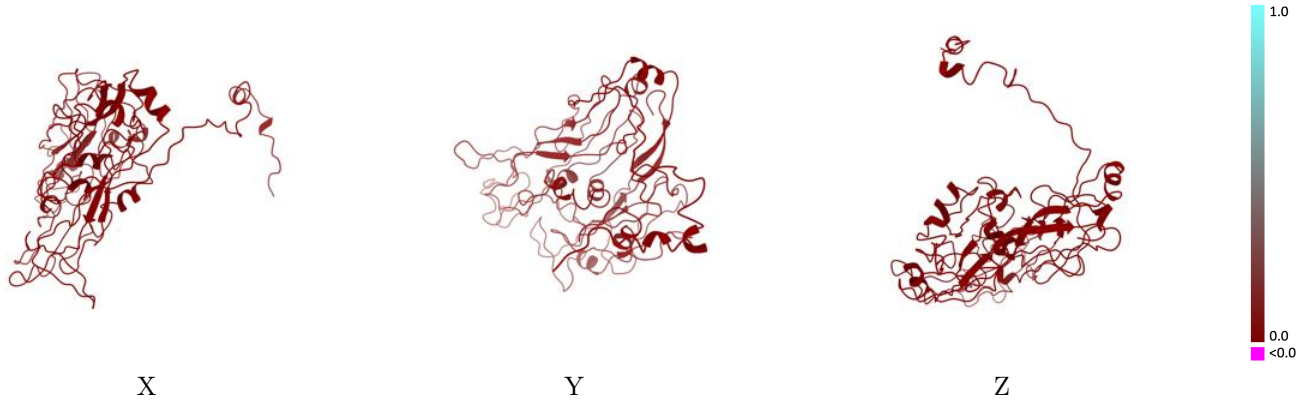
Y



Z

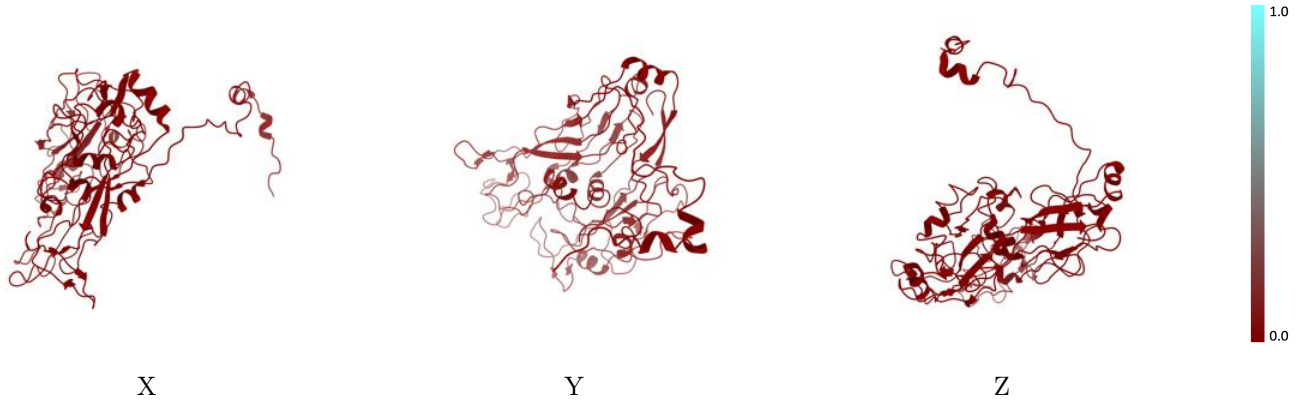
The images above show the 3D surface view of the map at the recommended contour level 0.017 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



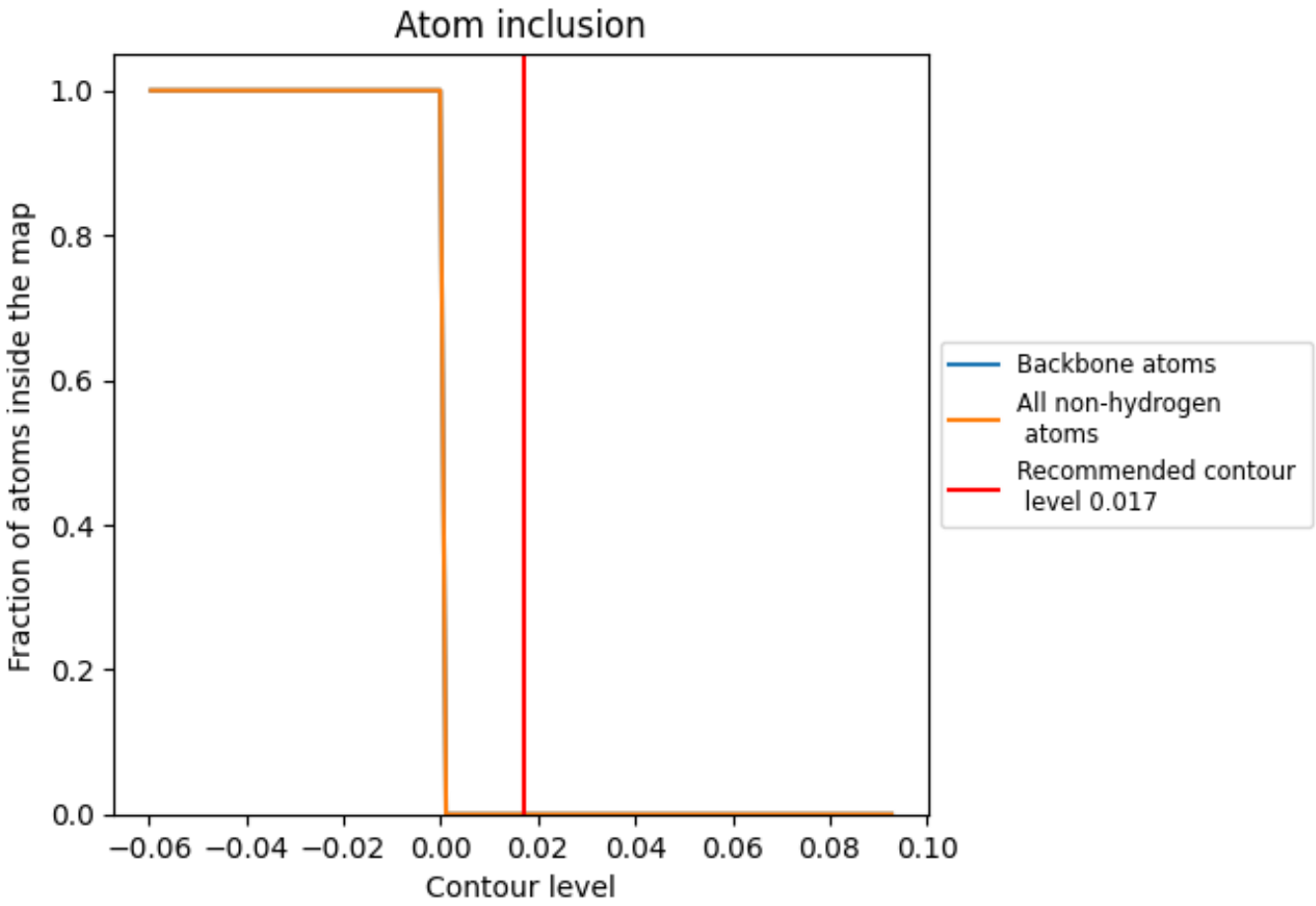
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.017).

9.4 Atom inclusion (i)



At the recommended contour level, 0% of all backbone atoms, 0% 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 (0.017) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.0000	0.0000
A	0.0000	0.0000