wwPDB EM Validation Summary Report (i)

Dec 9, 2023 – 12:09 PM EST

PDB ID	:	8SYO
EMDB ID	:	EMD-40886
Title	:	Human Retriever VPS35L/VPS29/VPS26C Complex (Composite Map)
Authors	:	Chen, Z.; Chen, B.; Burstein, E.; Han, Y.
Deposited on	:	2023-05-25
Resolution	:	2.94 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 (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:

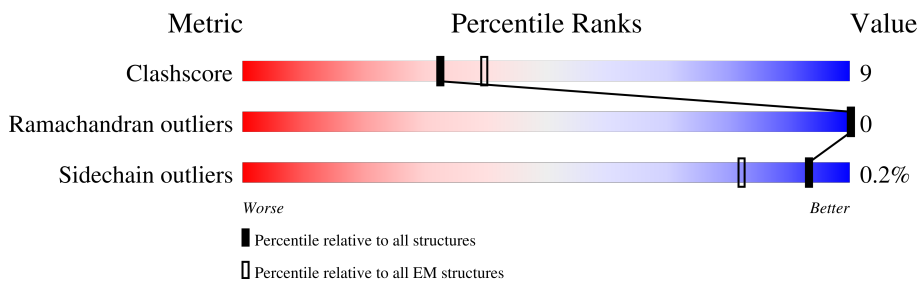|		
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.94 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	Percentile Ranks	Value
Clashscore		9
Ramachandran outliers		0
Sidechain outliers		0.2%

Worse   *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all EM structures

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	963	66% 17% 17%
2	B	205	61% 29% 10%
3	C	297	73% 19% 8%

2   Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10069 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 VPS35 endosomal protein-sorting factor-like.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	802	Total 6474	C 4172	N 1088	O 1169	S 45	0	0

- Molecule 2 is a protein called Vacuolar protein sorting-associated protein 29.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	184	Total 1463	C 945	N 249	O 264	S 5	0	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	187	GLU	-	expression tag	UNP Q9UBQ0
B	188	ASN	-	expression tag	UNP Q9UBQ0
B	189	LEU	-	expression tag	UNP Q9UBQ0
B	190	TYR	-	expression tag	UNP Q9UBQ0
B	191	PHE	-	expression tag	UNP Q9UBQ0
B	192	GLN	-	expression tag	UNP Q9UBQ0
B	193	GLY	-	expression tag	UNP Q9UBQ0
B	194	GLY	-	expression tag	UNP Q9UBQ0
B	195	GLY	-	expression tag	UNP Q9UBQ0
B	196	SER	-	expression tag	UNP Q9UBQ0
B	197	GLY	-	expression tag	UNP Q9UBQ0
B	198	GLY	-	expression tag	UNP Q9UBQ0
B	199	SER	-	expression tag	UNP Q9UBQ0
B	200	HIS	-	expression tag	UNP Q9UBQ0
B	201	HIS	-	expression tag	UNP Q9UBQ0
B	202	HIS	-	expression tag	UNP Q9UBQ0
B	203	HIS	-	expression tag	UNP Q9UBQ0
B	204	HIS	-	expression tag	UNP Q9UBQ0
B	205	HIS	-	expression tag	UNP Q9UBQ0

- Molecule 3 is a protein called Vacuolar protein sorting-associated protein 26C.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	273	Total 2132	C 1370	N 358	O 392	S 12	0	0

3 Residue-property plots ⓘ

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: VPS35 endosomal protein-sorting factor-like

• Molecule 2: Vacuolar protein sorting-associated protein 29

Chain B:

61% 29% 10%

MET ALA G3 L6 V7 L8 V9 D12 R18 P23 A24 K25 K34 I35 Q36 H37 I38 L39 C40 N43 A57 I62 F67 Y73 H90 G91 V94 I95 P96 W97 M100 A104 L105 L106 Q107 R108 Q109 I114 F123 E124 N130 K131 F132

Y133 I134 N135 S138 A139 T140 G141 A142 Y143 N144 A145 T148 S153 F154 V155 L156 M157 D158 I159 V164 V165 T166 Y167 V168 Y169 D174 D175 V176 E179 R180 I181 P186 GLU ASN LEU TYR PHE GLN GLY GLY GLY SER GLY GLY SER HIS HIS HIS HIS HIS HIS

• Molecule 3: Vacuolar protein sorting-associated protein 26C

Chain C:

73% 19% 8%

MET GLY T3 N13 K14 V15 L22 S29 SER LYS ASP SER VAL GLN HIS GLN G38 V39 T42 V47 K54 S55 V56 GLY VAL PHE GLU A61 V66 T75 V79 K80 P81 GLY LYS PHE PRO S86 E94 H98 L99 K100 Y106 E107 V112 I116

C122 D123 M124 K125 R126 S127 LEU LEU ALA LYS D132 L133 T134 H142 K149 S153 I159 E162 Q165 K168 E169 R170 L173 F176 T185 A204 R207 S208 V209 E210 L211 Q212 L213 V214 R215 V216 E217 C221 A222 GLU GLY Y225 D228 I235

A238 V242 I250 Y251 M252 P255 T259 C260 L263 I276 V277 V278 L279 L280 H281 I286 T287 E288 N289 F290 I297

4   Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	426624	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 $(e^-/\text{Å}^2)$	60	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	105000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	3.053	Depositor
Minimum map value	-0.002	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.025	Depositor
Recommended contour level	0.13	Depositor
Map size (Å)	318.72, 318.72, 318.72	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0624, 1.0624, 1.0624	Depositor

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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	0.28	0/6617	0.47	0/8949
2	B	0.27	0/1498	0.50	0/2032
3	C	0.29	0/2168	0.51	0/2939
All	All	0.28	0/10283	0.49	0/13920

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	6474	0	6519	121	0
2	B	1463	0	1470	42	0
3	C	2132	0	2188	33	0
All	All	10069	0	10177	189	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

The worst 5 of 189 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:547:GLN:HE21	1:A:581:VAL:HG13	1.46	0.81

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:7:VAL:HB	2:B:157:MET:HB2	1.64	0.79
1:A:708:CYS:O	1:A:712:ILE:HD12	1.83	0.78
2:B:25:LYS:NZ	2:B:174:ASP:O	2.18	0.76
1:A:705:VAL:HG21	1:A:740:GLN:HE21	1.51	0.75

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	794/963 (82%)	777 (98%)	17 (2%)	0	100	100
2	B	182/205 (89%)	177 (97%)	5 (3%)	0	100	100
3	C	261/297 (88%)	250 (96%)	11 (4%)	0	100	100
All	All	1237/1465 (84%)	1204 (97%)	33 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	725/867 (84%)	724 (100%)	1 (0%)	93	98
2	B	161/176 (92%)	161 (100%)	0	100	100
3	C	244/266 (92%)	243 (100%)	1 (0%)	91	97

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1130/1309 (86%)	1128 (100%)	2 (0%)	93 98

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	612	HIS
3	C	125	LYS

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

Mol	Chain	Res	Type
1	A	547	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-40886. 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 ⓘ

6.1.1 Primary map

| X | Y | Z |
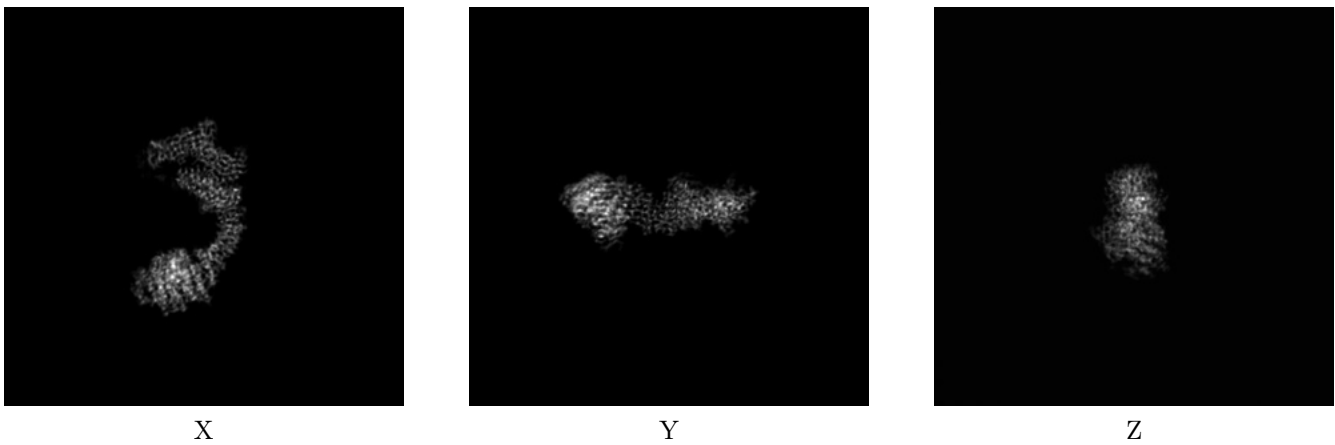
The images above show the map projected in three orthogonal directions.

6.2 Central slices ⓘ

6.2.1 Primary map

| X Index: 150 | Y Index: 150 | Z Index: 150 |
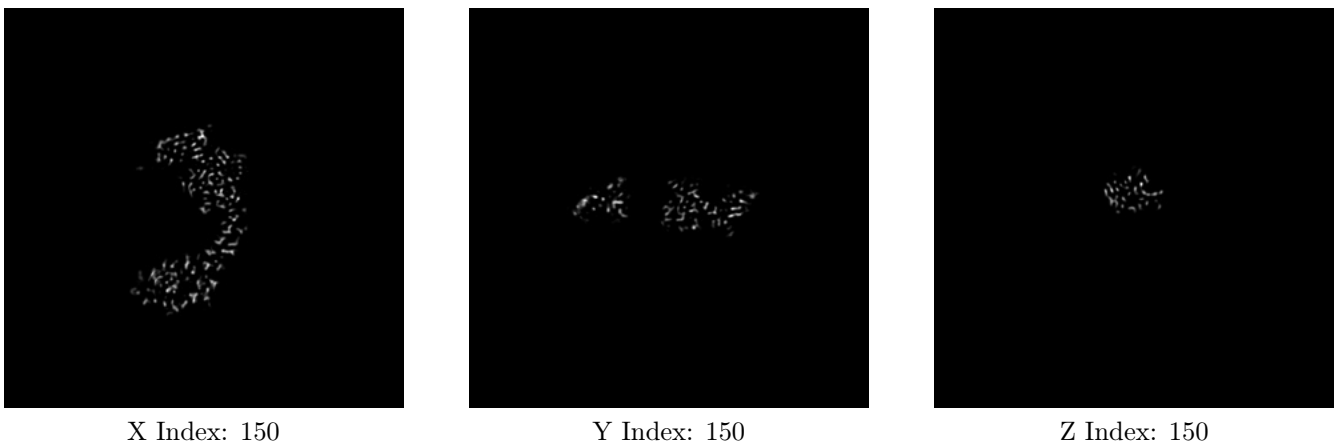
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ

6.3.1 Primary map

| X Index: 152 | Y Index: 155 | Z Index: 100 |
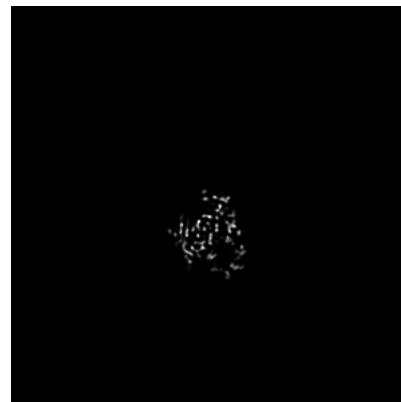
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map

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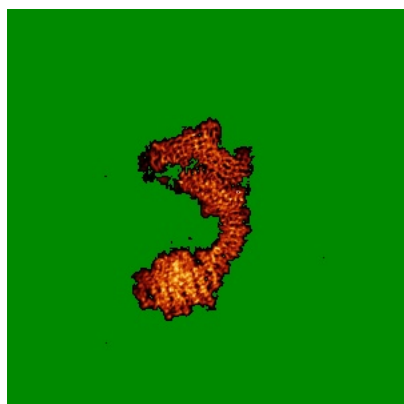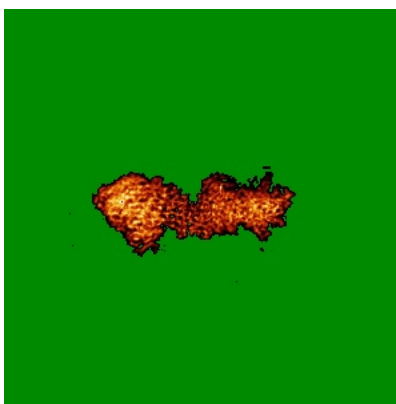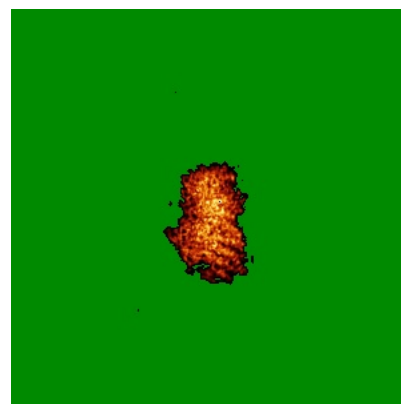
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5　Orthogonal surface views ⓘ

6.5.1　Primary map

X　　　　　　　　　　Y　　　　　　　　　　Z
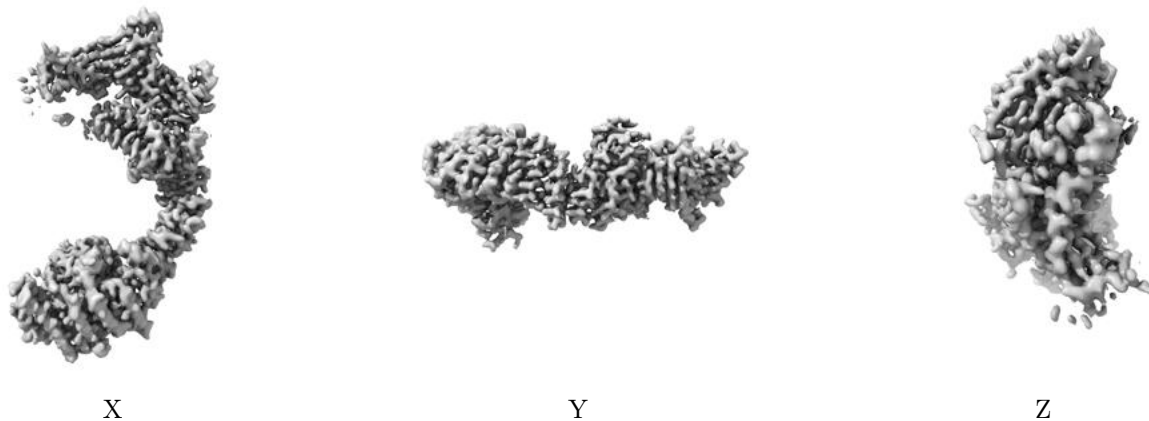
The images above show the 3D surface view of the map at the recommended contour level 0.13. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.6　Mask visualisation ⓘ

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

7 Map analysis ⓘ

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

7.1 Map-value distribution ⓘ
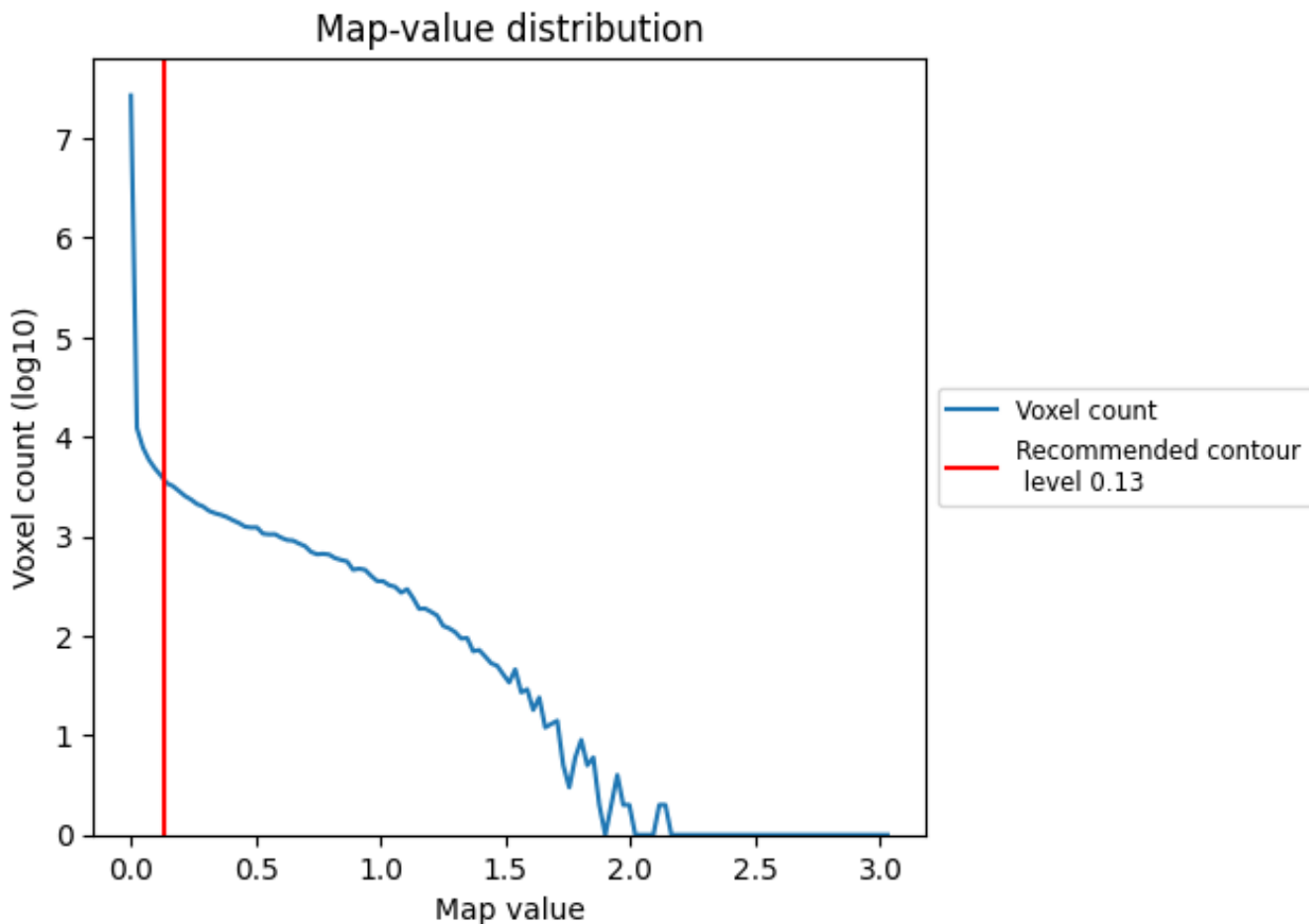
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2 Volume estimate ⓘ
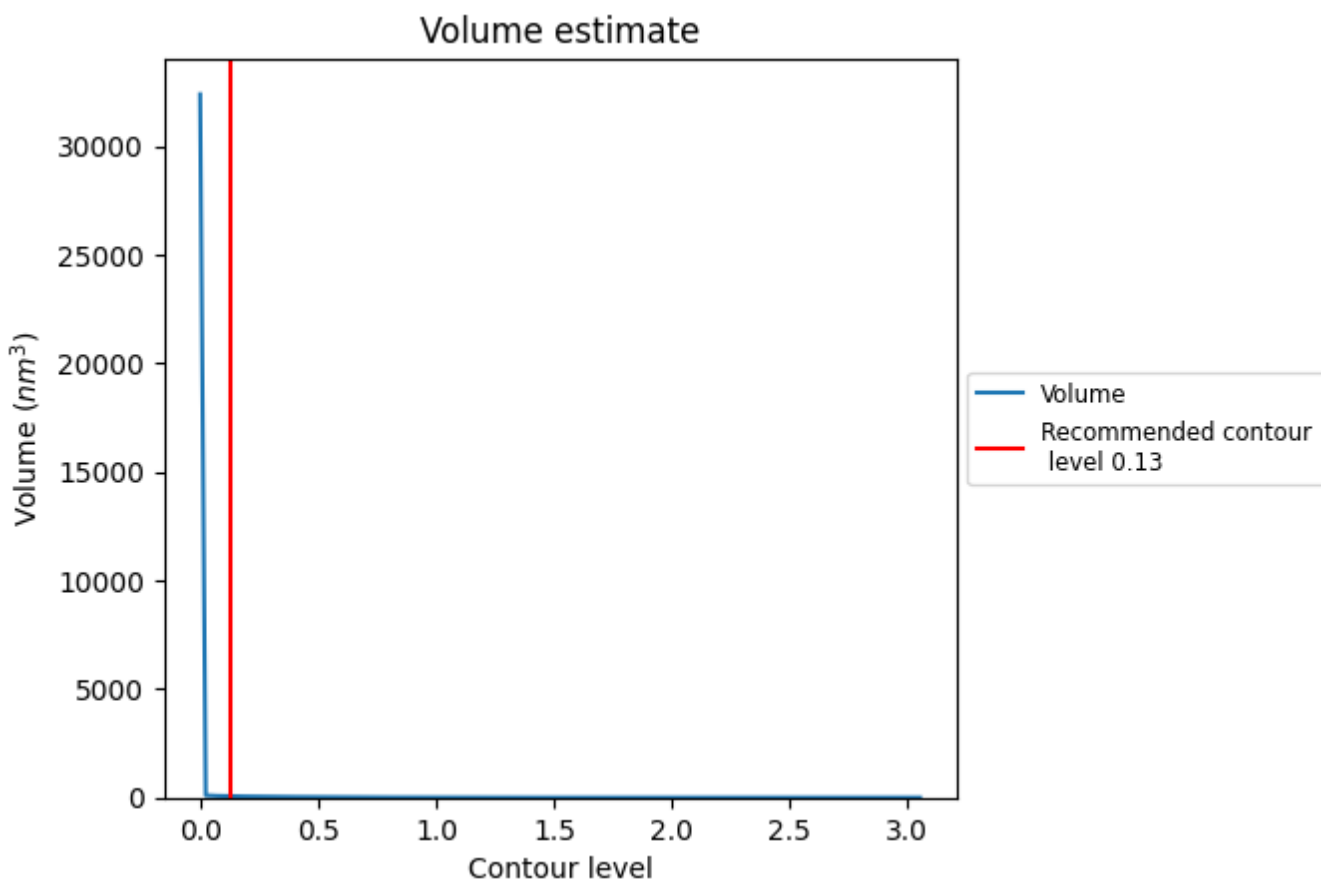
The volume at the recommended contour level is 62 nm^3; this corresponds to an approximate mass of 56 kDa.

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 ⓘ
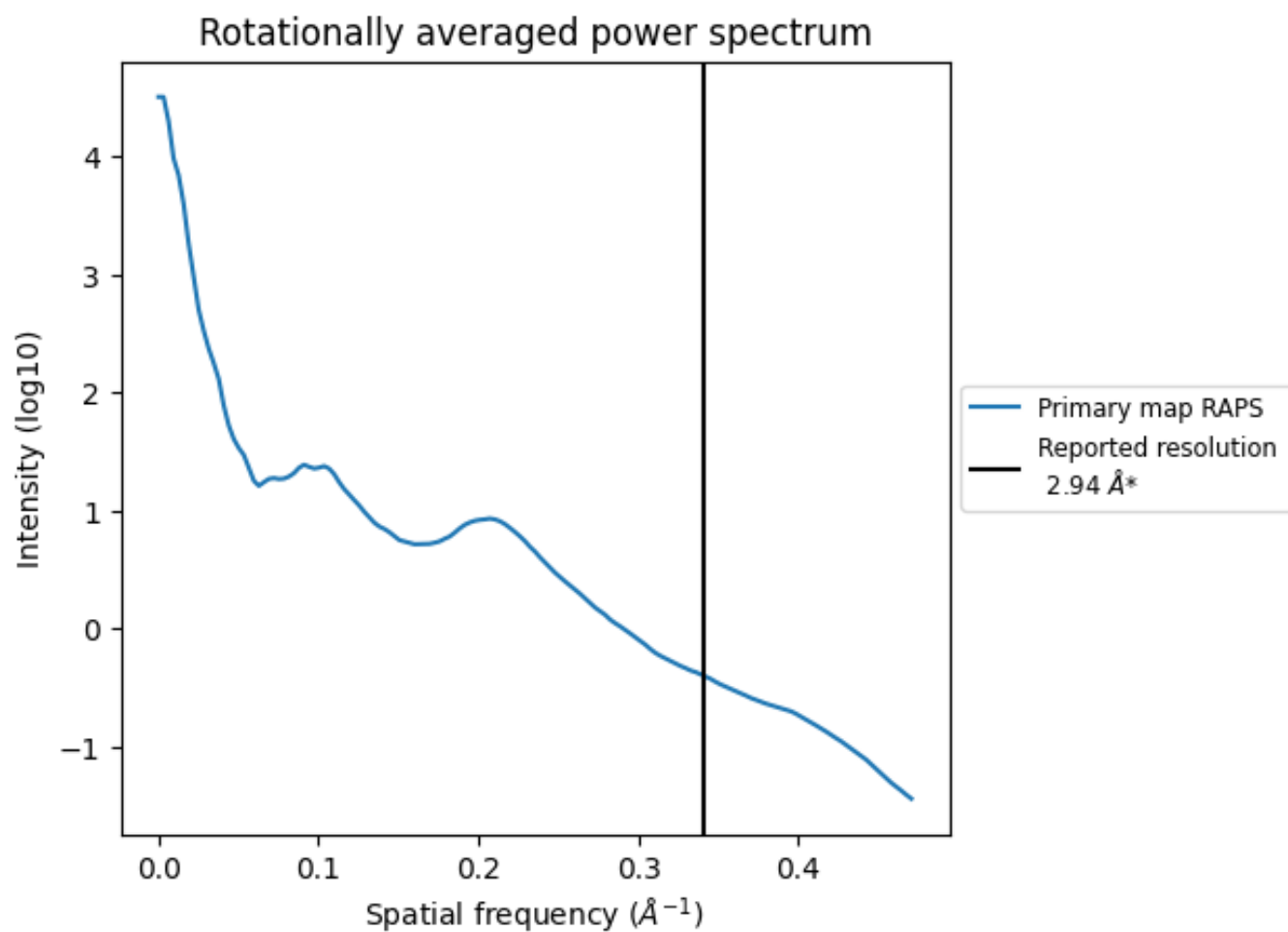
*Reported resolution corresponds to spatial frequency of 0.340 Å^{-1}

8 Fourier-Shell correlation ⓘ

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-40886 and PDB model 8SYO. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay (i)

| X | Y | Z |
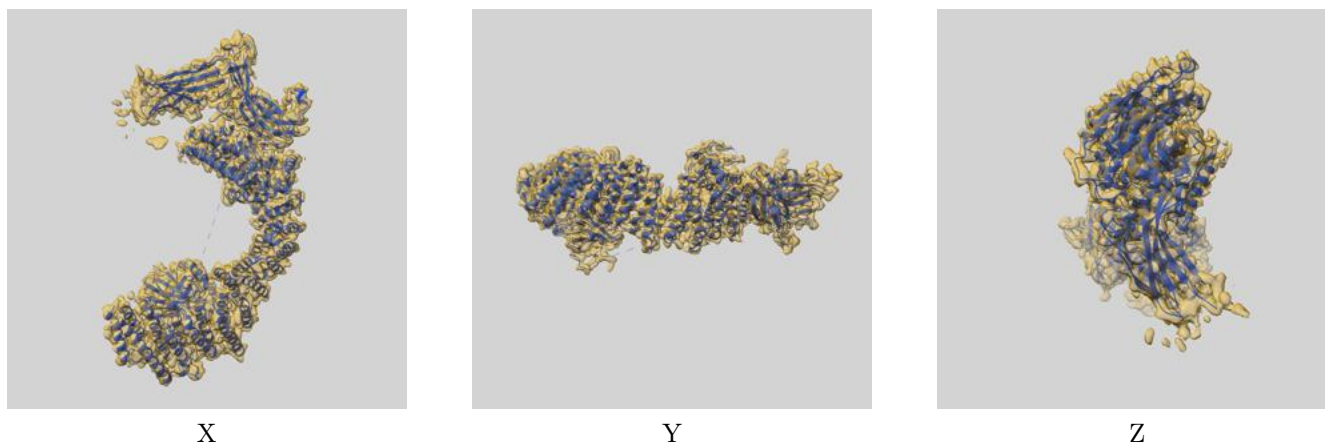
The images above show the 3D surface view of the map at the recommended contour level 0.13 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2　Q-score mapped to coordinate model ⓘ

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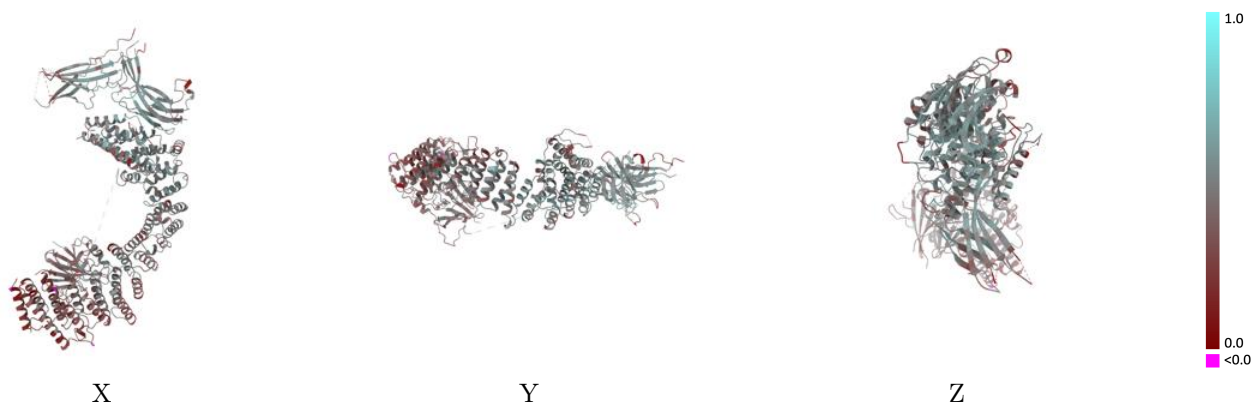
The images above show the model with each residue coloured according its Q-score. This shows
their resolvability in the map with higher Q-score values reflecting better resolvability. Please
note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at
resolutions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

X　　　　　　　　　　　Y　　　　　　　　　　　Z
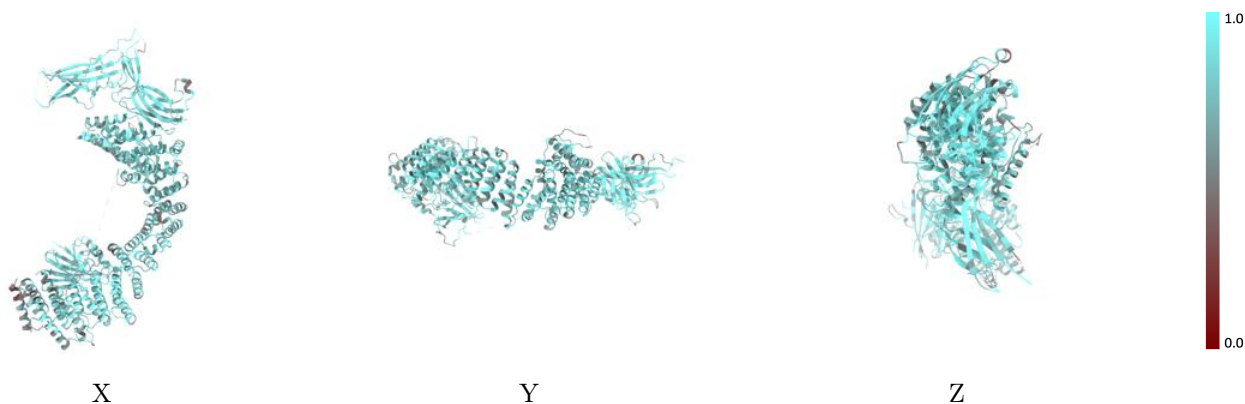
The images above show the model with each residue coloured according to its atom inclusion. This
shows to what extent they are inside the map at the recommended contour level (0.13).

9.4   Atom inclusion ⓘ
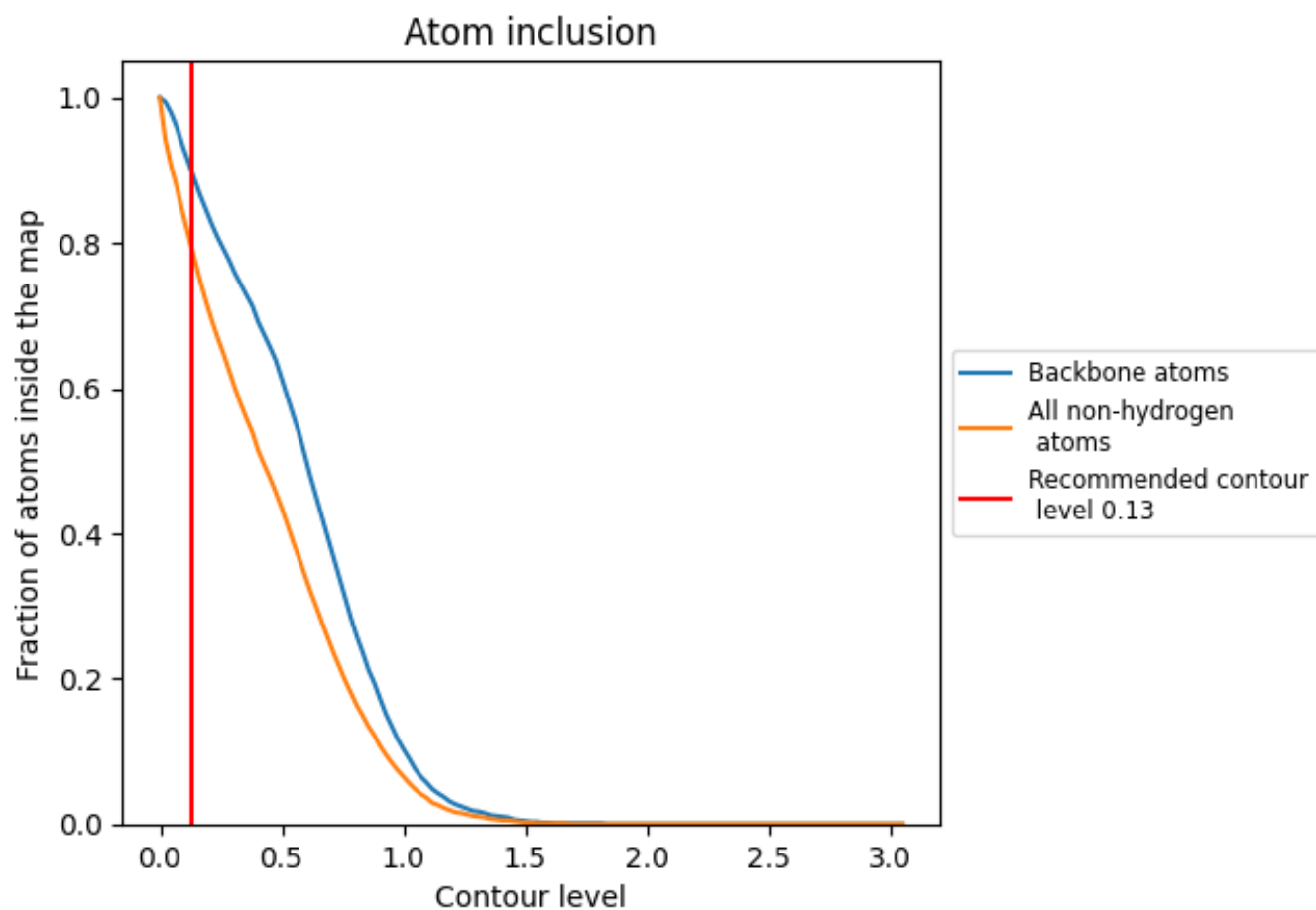
At the recommended contour level, 90% of all backbone atoms, 79% of all non-hydrogen atoms, are inside the map.

9.5   Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.13) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7930	0.4290
A	0.7740	0.4150
B	0.8300	0.4100
C	0.8250	0.4870

1.0

0.0

<0.0