# wwPDB EM Validation Summary Report

Nov 3, 2024 – 06:54 am GMT

| | | |
|---|---|---|
| PDB ID | : | 7QFQ |
| EMDB ID | : | EMD-13947 |
| Title | : | Cryo-EM structure of Botulinum neurotoxin serotype B |
| Authors | : | Kosenina, S.; Martinez-Carranza, M.; Davies, J.R.; Masuyer, G.; Stenmark, P. |
| Deposited on | : | 2021-12-06 |
| Resolution | : | 3.60 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 ⓘ 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:

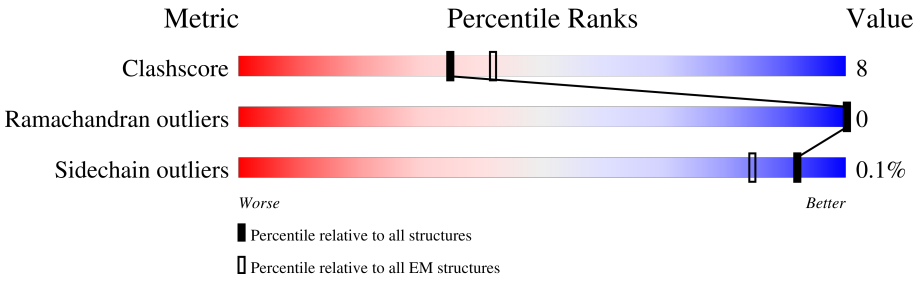
| | | |
|---|---|---|
| EMDB validation analysis | : | 0.0.1.dev113 |
| MolProbity | : | 4.02b-467 |
| Percentile statistics | : | 20231227.v01 (using entries in the PDB archive December 27th 2023) |
| MapQ | : | 1.9.13 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.39 |

# 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 3.60 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 | 210492 | 15764 |
| Ramachandran outliers | 207382 | 16835 |
| Sidechain outliers | 206894 | 16415 |

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 | 1307 | 30% 79% 20% |

# 2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 10659 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 Botulinum neurotoxin type B.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | |
| 1 | A | 1291 | 10659 | 6878 | 1714 | 2032 | 35 | 0 | 0 |

There are 20 discrepancies between the modelled and reference sequences:

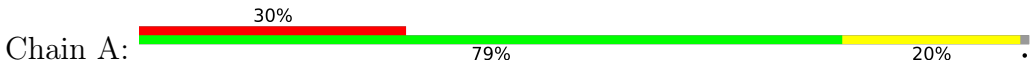
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 231 | GLN | GLU | engineered mutation | UNP P10844 |
| A | 234 | TYR | HIS | engineered mutation | UNP P10844 |
| A | 1191 | MET | GLU | conflict | UNP P10844 |
| A | 1199 | TYR | SER | conflict | UNP P10844 |
| A | 1292 | LEU | - | expression tag | UNP P10844 |
| A | 1293 | GLU | - | expression tag | UNP P10844 |
| A | 1294 | VAL | - | expression tag | UNP P10844 |
| A | 1295 | LEU | - | expression tag | UNP P10844 |
| A | 1296 | PHE | - | expression tag | UNP P10844 |
| A | 1297 | GLN | - | expression tag | UNP P10844 |
| A | 1298 | GLY | - | expression tag | UNP P10844 |
| A | 1299 | PRO | - | expression tag | UNP P10844 |
| A | 1300 | LEU | - | expression tag | UNP P10844 |
| A | 1301 | GLU | - | expression tag | UNP P10844 |
| A | 1302 | HIS | - | expression tag | UNP P10844 |
| A | 1303 | HIS | - | expression tag | UNP P10844 |
| A | 1304 | HIS | - | expression tag | UNP P10844 |
| A | 1305 | HIS | - | expression tag | UNP P10844 |
| A | 1306 | HIS | - | expression tag | UNP P10844 |
| A | 1307 | HIS | - | expression tag | UNP P10844 |

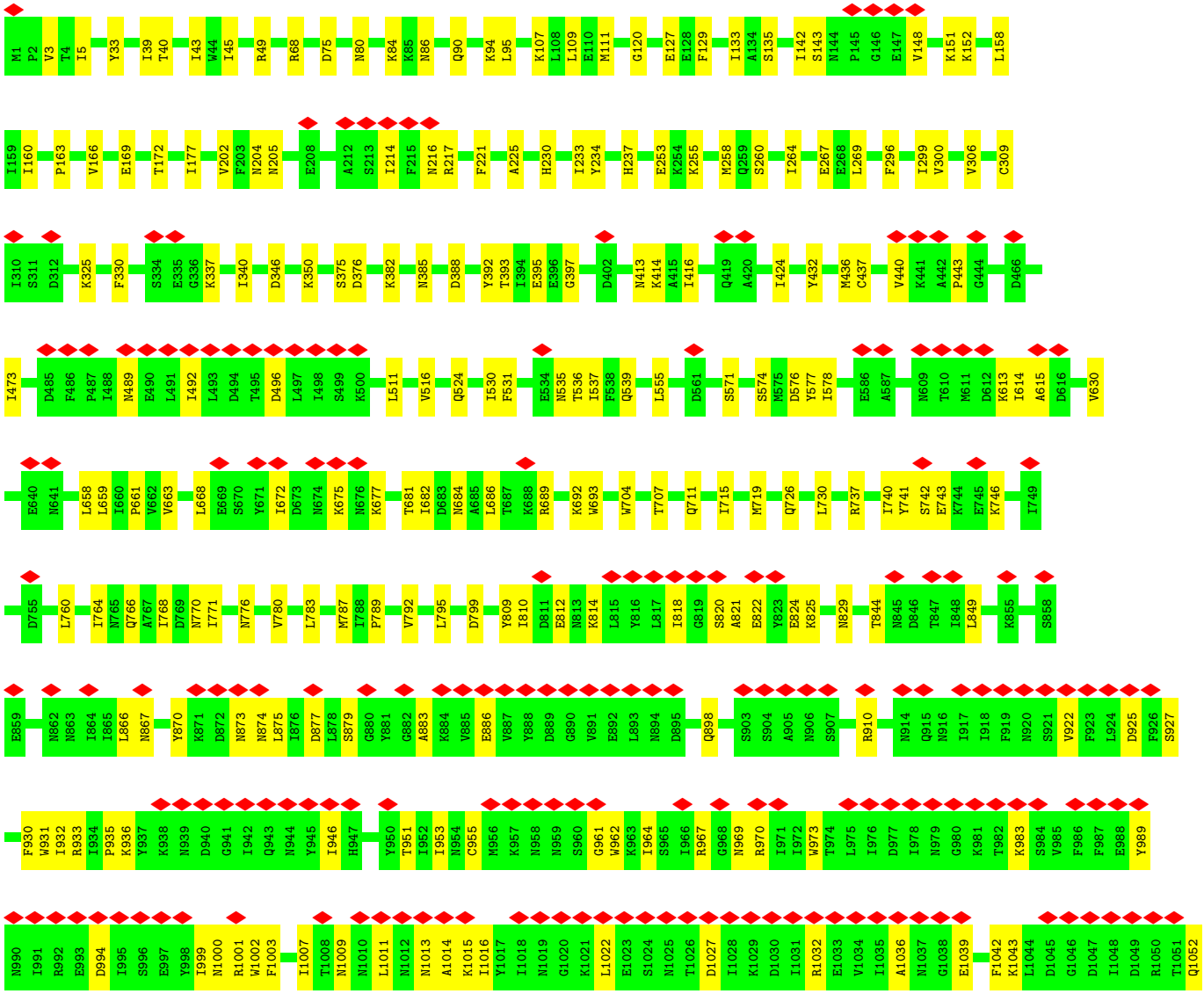
# 3 Residue-property plots (i)

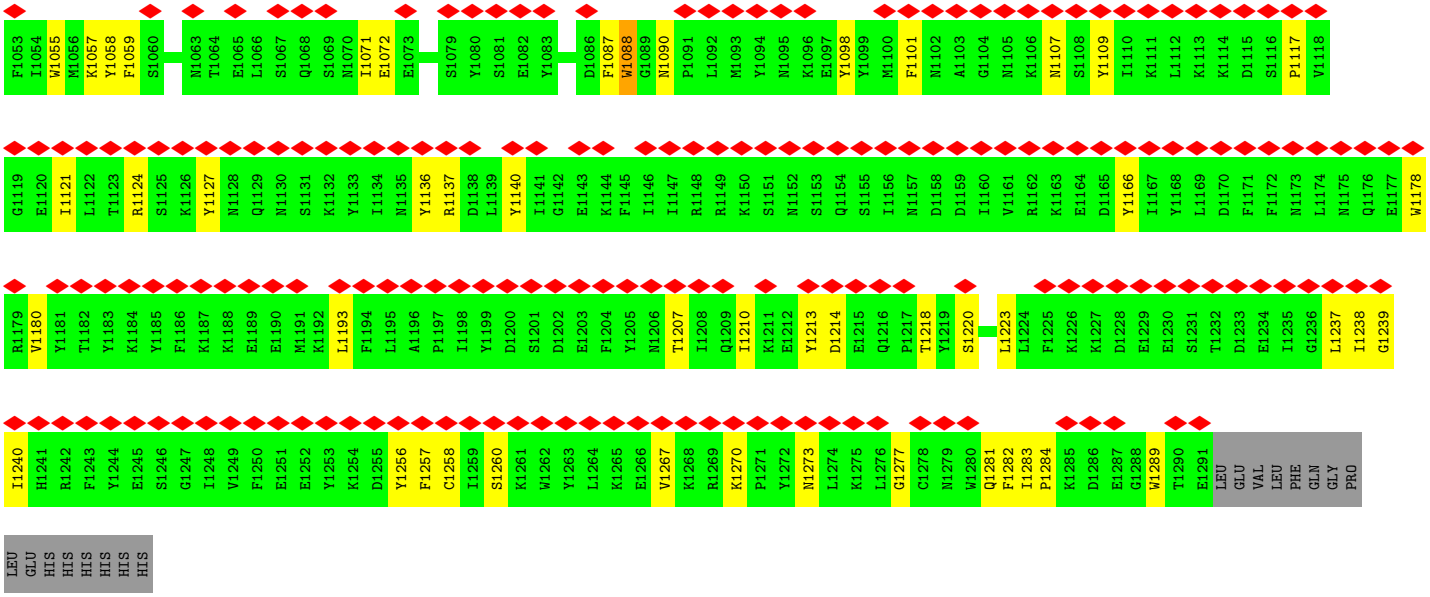
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Botulinum neurotoxin type B

Chain A:




F1053 I1054 W1055 M1056 K1057 Y1058 F1059 S1060 N1063 T1064 E1065 L1066 S1067 Q1068 S1069 N1070 I1071 E1072 E1073 S1079 Y1080 S1081 E1082 Y1083 D1086 F1087 W1088 G1089 N1090 P1091 L1092 M1093 Y1094 N1095 K1096 E1097 Y1098 Y1099 M1100 F1101 N1102 A1103 G1104 N1105 K1106 N1107 S1108 Y1109 I1110 K1111 L1112 K1113 K1114 D1115 S1116 P1117 V1118
G1119 E1120 I1121 L1122 T1123 R1124 S1125 K1126 Y1127 N1128 Q1129 N1130 S1131 K1132 Y1133 I1134 N1135 Y1136 R1137 D1138 L1139 Y1140 I1141 G1142 E1143 K1144 F1145 I1146 I1147 R1148 R1149 K1150 S1151 N1152 S1153 Q1154 S1155 I1156 N1157 D1158 D1159 I1160 V1161 R1162 K1163 E1164 D1165 Y1166 I1167 Y1168 L1169 D1170 F1171 F1172 N1173 L1174 N1175 Q1176 E1177 W1178
R1179 Y1180 Y1181 T1182 Y1183 K1184 Y1185 F1186 K1187 K1188 E1189 E1190 M1191 K1192 L1193 F1194 L1195 A1196 P1197 I1198 Y1199 D1200 S1201 D1202 E1203 F1204 Y1205 N1206 T1207 I1208 Q1209 I1210 K1211 E1212 Y1213 D1214 E1215 Q1216 P1217 T1218 Y1219 S1220 L1223 L1224 F1225 K1226 K1227 D1228 E1229 E1230 S1231 T1232 D1233 E1234 I1235 G1236 L1237 I1238 G1239
I1240 H1241 R1242 F1243 Y1244 E1245 S1246 G1247 I1248 V1249 F1250 E1251 E1252 Y1253 K1254 D1255 Y1256 F1257 C1258 I1259 S1260 K1261 W1262 Y1263 L1264 K1265 E1266 V1267 K1268 R1269 K1270 P1271 Y1272 N1273 L1274 K1275 L1276 G1277 C1278 N1279 W1280 Q1281 F1282 I1283 P1284 K1285 D1286 E1287 G1288 W1289 T1290 E1291 LEU GLU VAL LEU PHE GLN GLY PRO
LEU GLU HIS HIS HIS HIS HIS

# 4 Experimental information (i)

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 286802 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | PHASE FLIPPING AND AMPLITUDE CORRECTION | Depositor |
| Microscope | TFS KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose ($e^-/Å^2$) | 50.0 | 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 | 2.181 | Depositor |
| Minimum map value | -1.697 | Depositor |
| Average map value | 0.000 | Depositor |
| Map value standard deviation | 0.034 | Depositor |
| Recommended contour level | 0.252 | Depositor |
| Map size (Å) | 327.0, 327.0, 327.0 | wwPDB |
| Map dimensions | 300, 300, 300 | wwPDB |
| Map angles (°) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (Å) | 1.09, 1.09, 1.09 | 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 | 0.25 | 0/10893 | 0.46 | 0/14716 |

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 | 10659 | 0 | 10529 | 165 | 0 |
| All | All | 10659 | 0 | 10529 | 165 | 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 165 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:844:THR:HG21 | 1:A:849:LEU:HD12 | 1.66 | 0.77 |
| 1:A:668:LEU:HD11 | 1:A:681:THR:HB | 1.69 | 0.75 |
| 1:A:536:THR:HG22 | 1:A:537:ILE:H | 1.50 | 0.75 |
| 1:A:1237:LEU:HD12 | 1:A:1267:VAL:HG23 | 1.69 | 0.75 |
| 1:A:95:LEU:HD22 | 1:A:225:ALA:HB2 | 1.69 | 0.73 |

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 | 1289/1307 (99%) | 1219 (95%) | 70 (5%) | 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 | 1190/1205 (99%) | 1189 (100%) | 1 (0%) | 92 97 |

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

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 1088 | TRP |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 oligosaccharides 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.

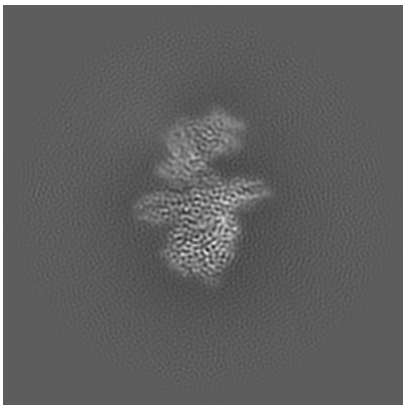
# 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-13947. These allow visual inspection of the internal detail of the map and identification of artifacts.

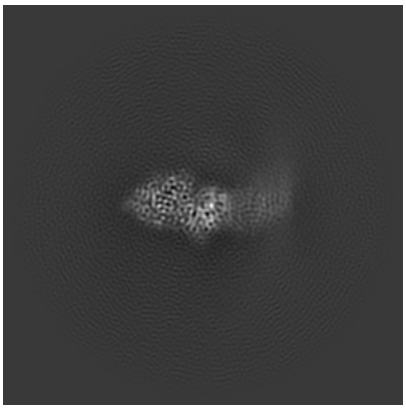
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

## 6.1 Orthogonal projections

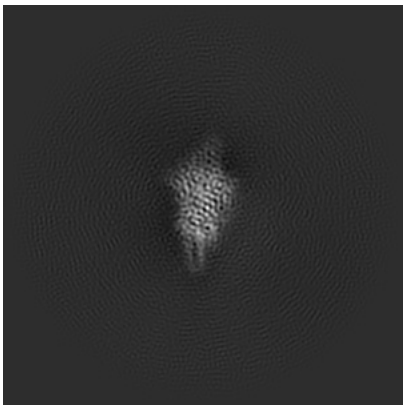
### 6.1.1 Primary map



X

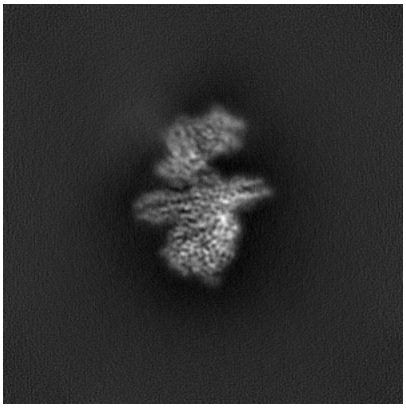


Y

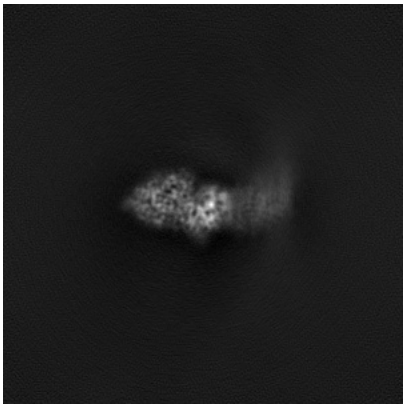


Z

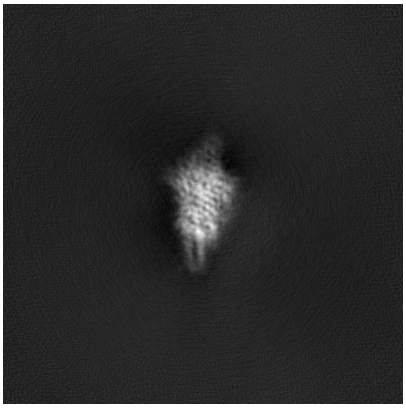
### 6.1.2 Raw map



X



Y

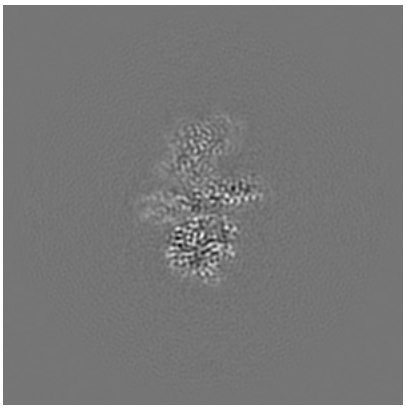


Z

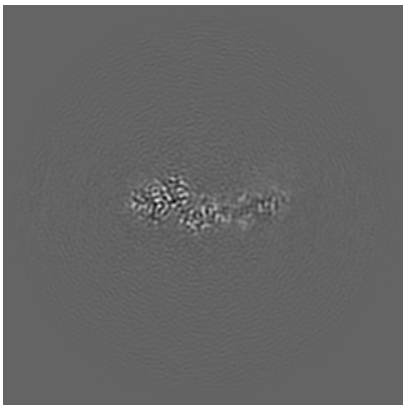
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices (i)

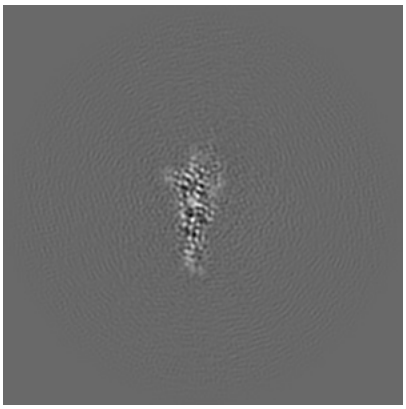
### 6.2.1 Primary map



X Index: 150

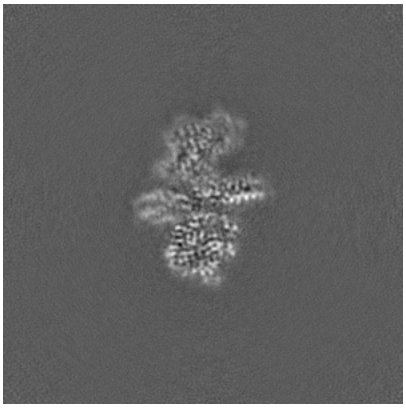


Y Index: 150

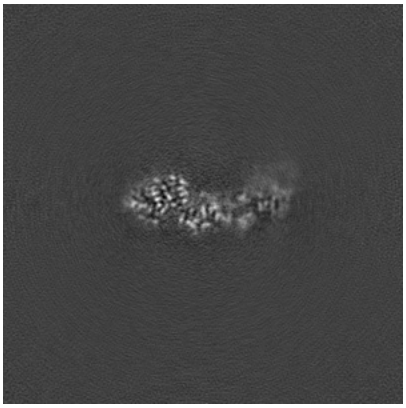


Z Index: 150

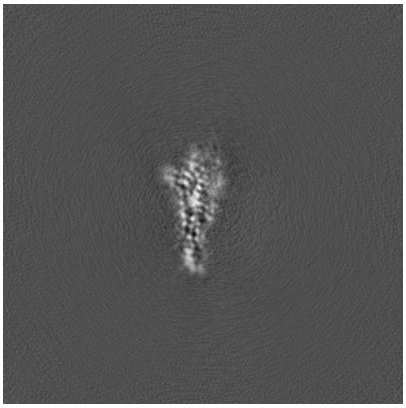
### 6.2.2 Raw map



X Index: 150



Y Index: 150

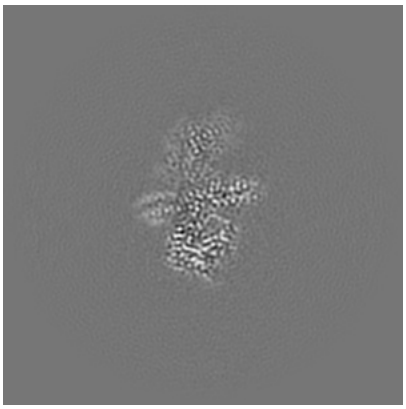


Z Index: 150

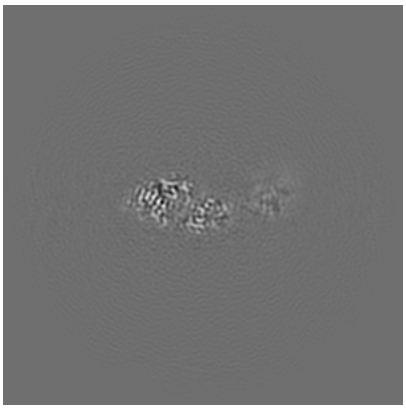
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices (i)

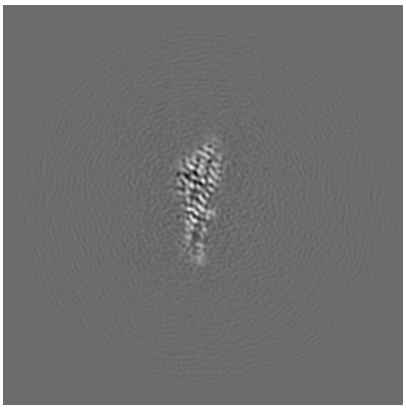
### 6.3.1 Primary map



X Index: 148

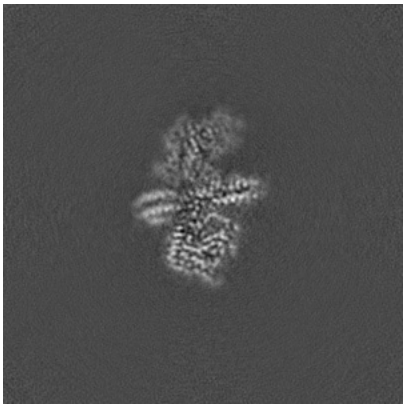


Y Index: 157

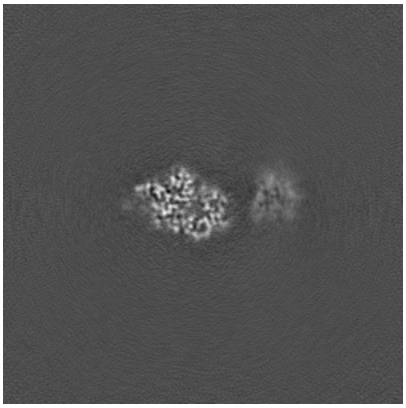


Z Index: 156

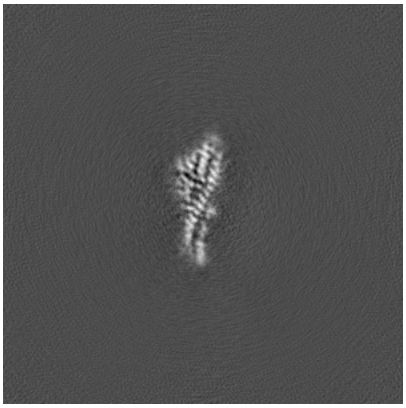
### 6.3.2 Raw map



X Index: 148



Y Index: 164

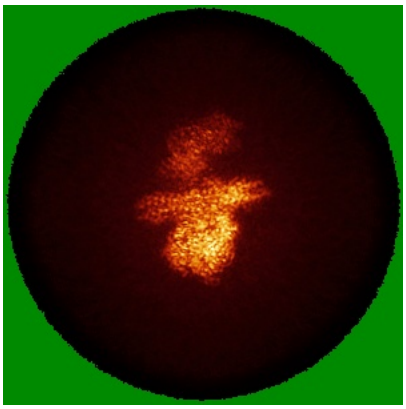


Z Index: 156

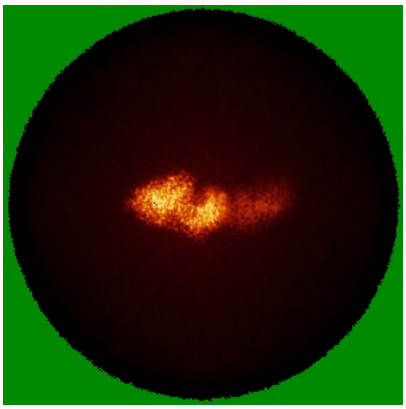
The images above show the largest variance slices of the map in three orthogonal directions.

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

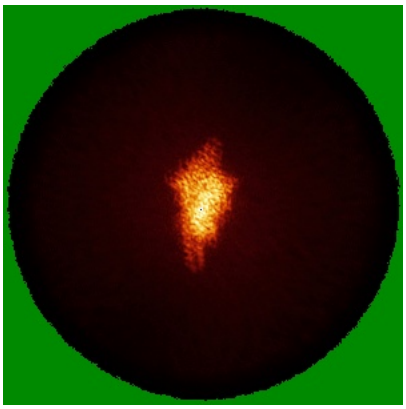
### 6.4.1 Primary map



X

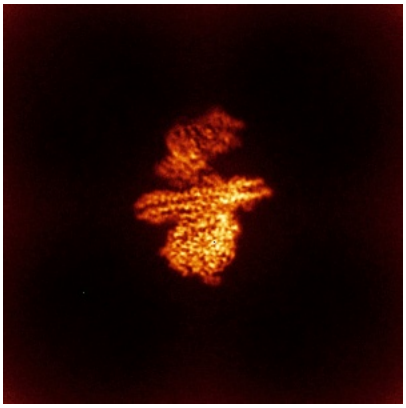


Y

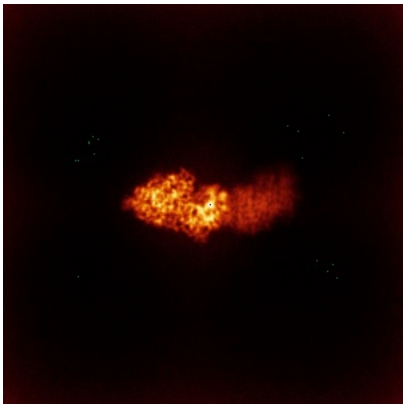


Z

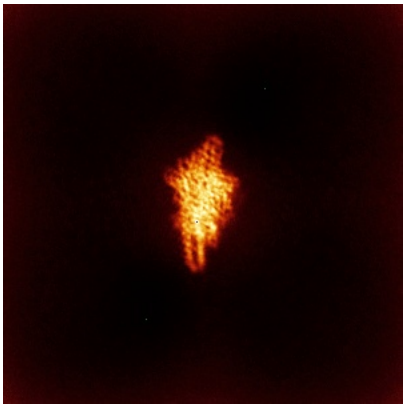
### 6.4.2 Raw map



X



Y

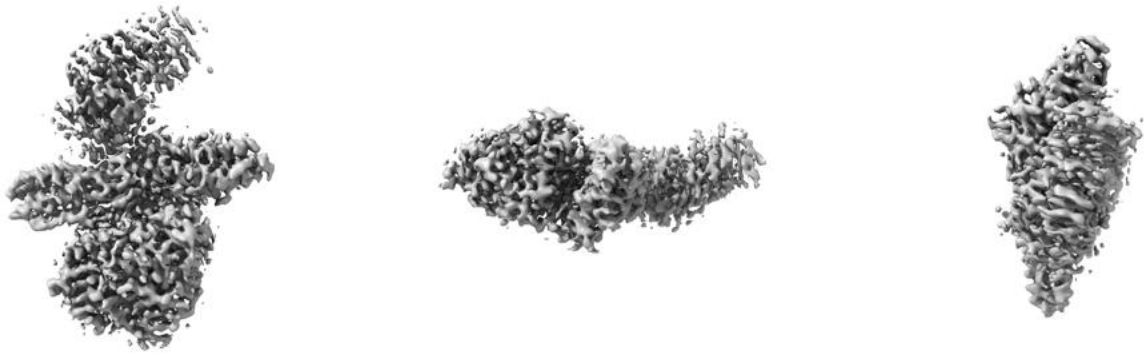


Z

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)

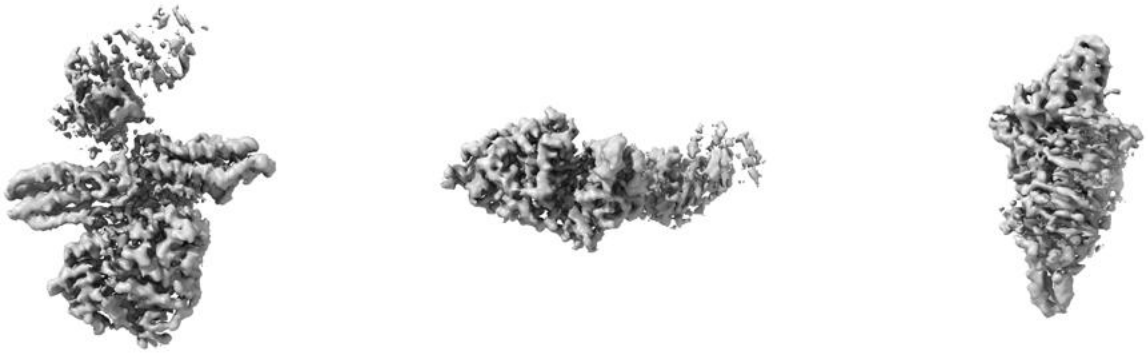
### 6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.252. 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.

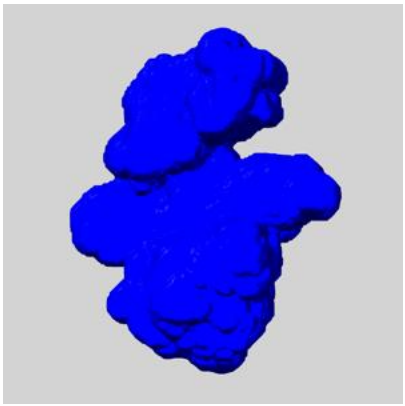
## 6.6 Mask visualisation (i)

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

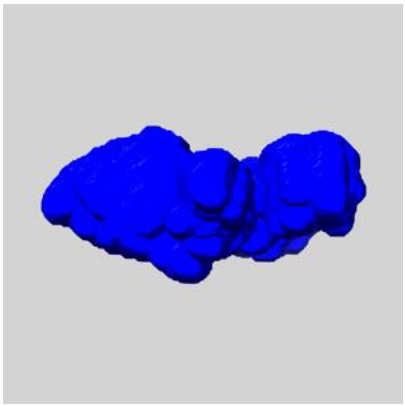
A mask typically either:

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

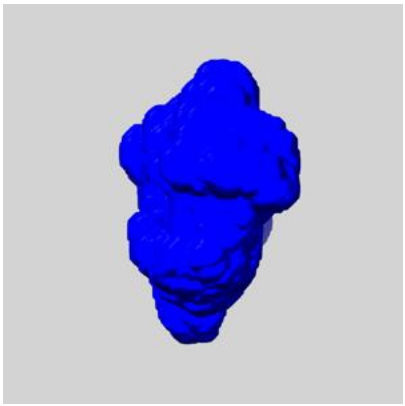
### 6.6.1 emd_13947_msk_1.map (i)



X



Y

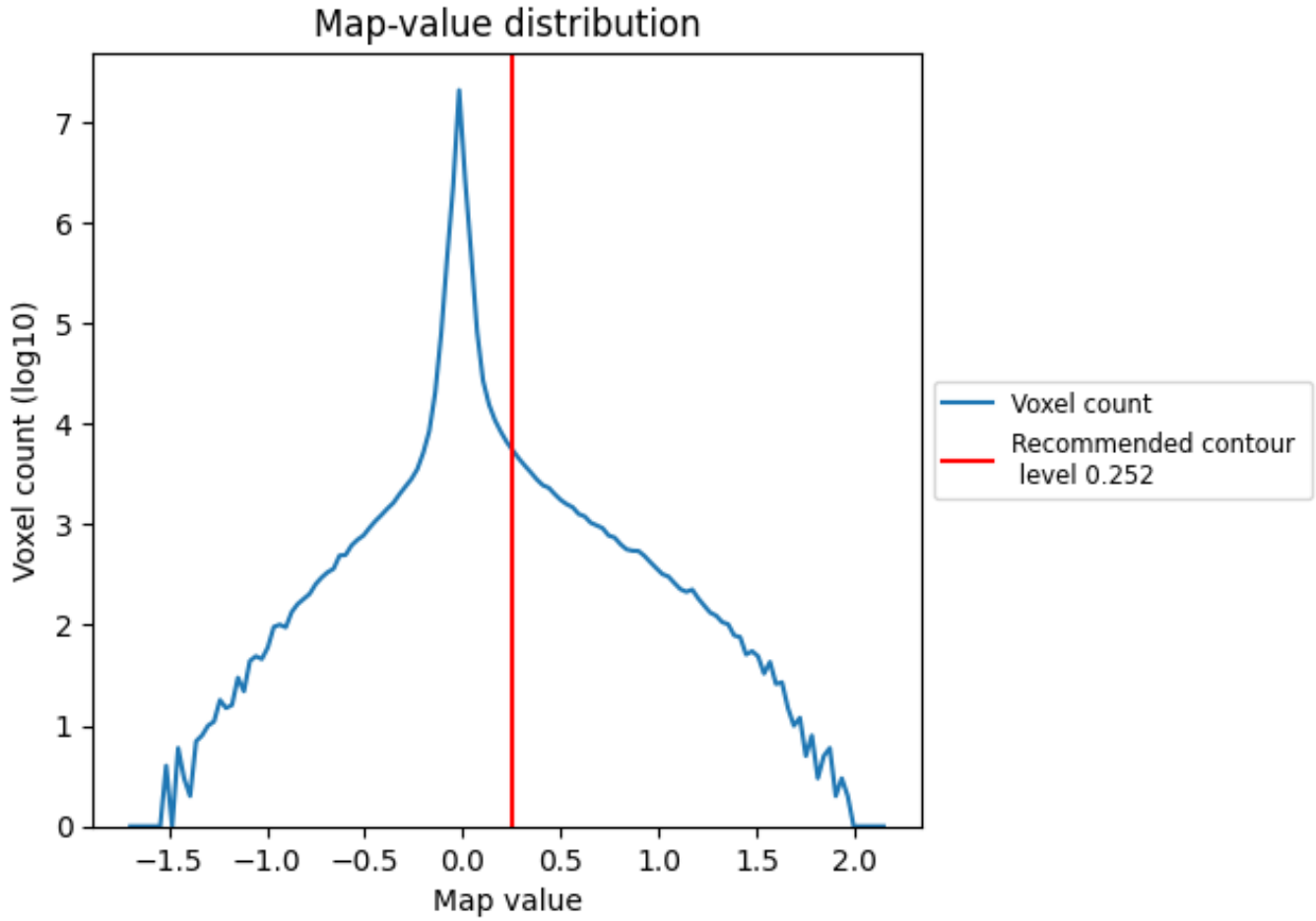


Z

# 7 Map analysis

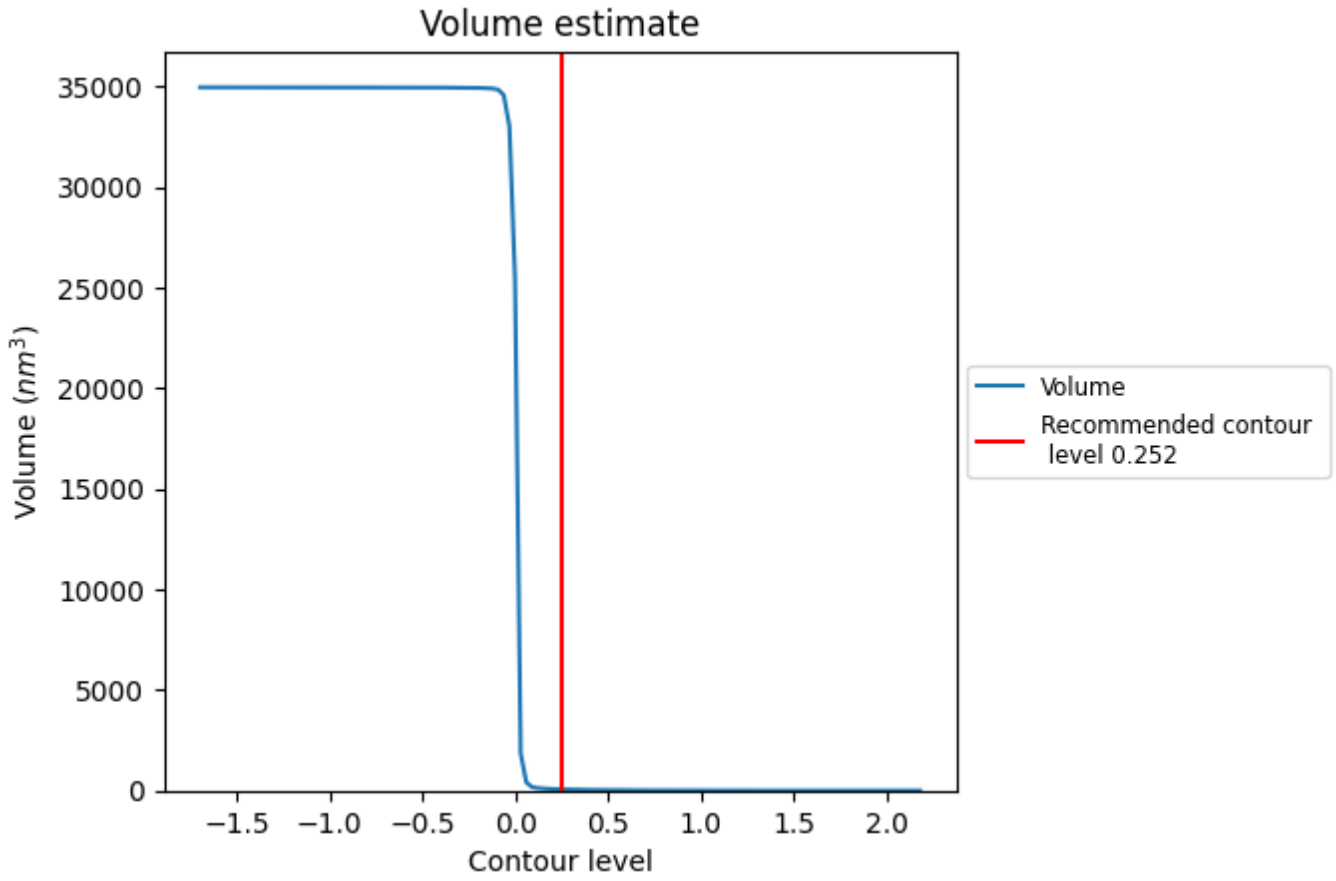
This section contains the results of statistical analysis of the map.

## 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.

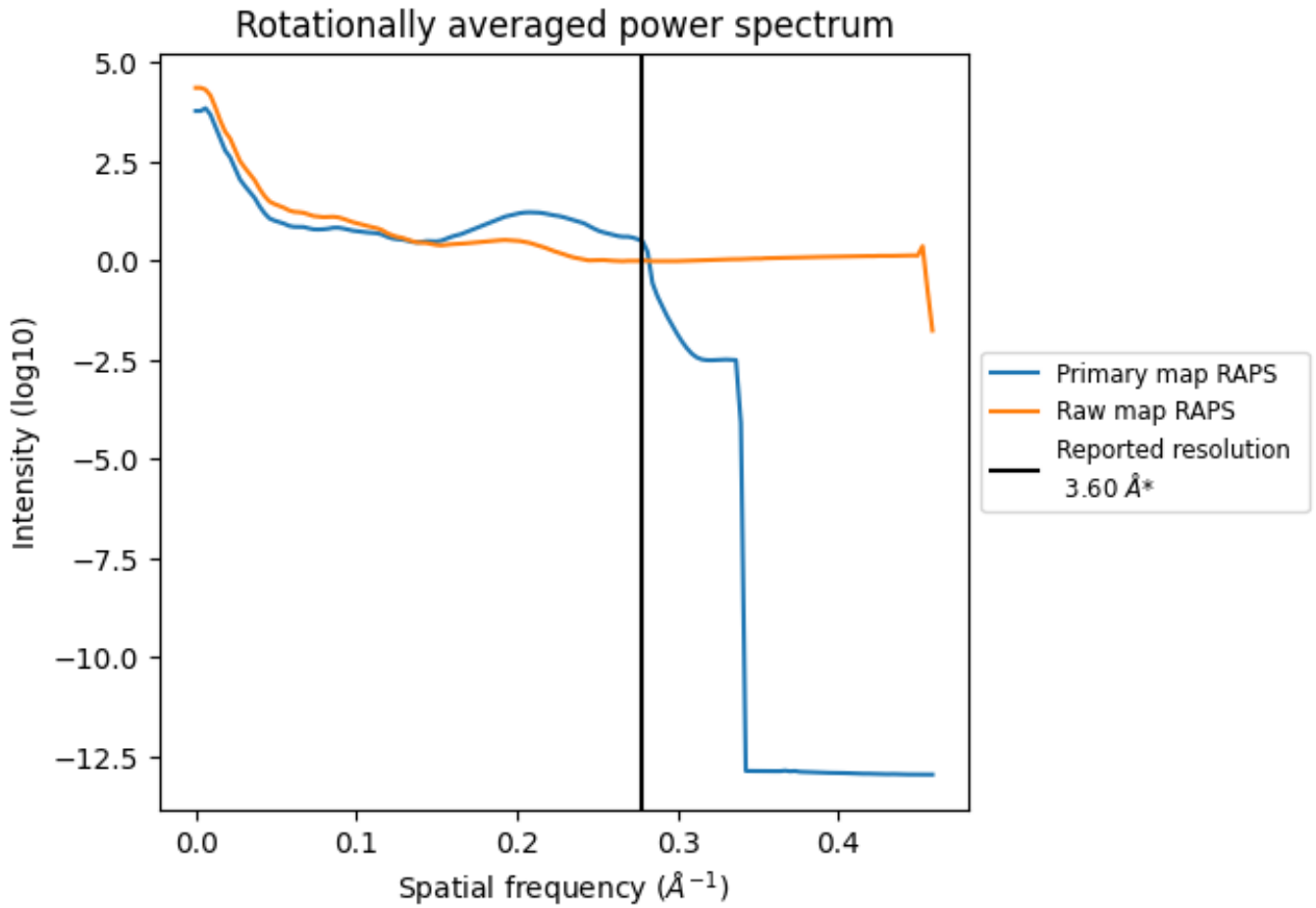
## 7.2 Volume estimate (i)




The volume at the recommended contour level is 60 $nm^3$; this corresponds to an approximate mass of 54 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)

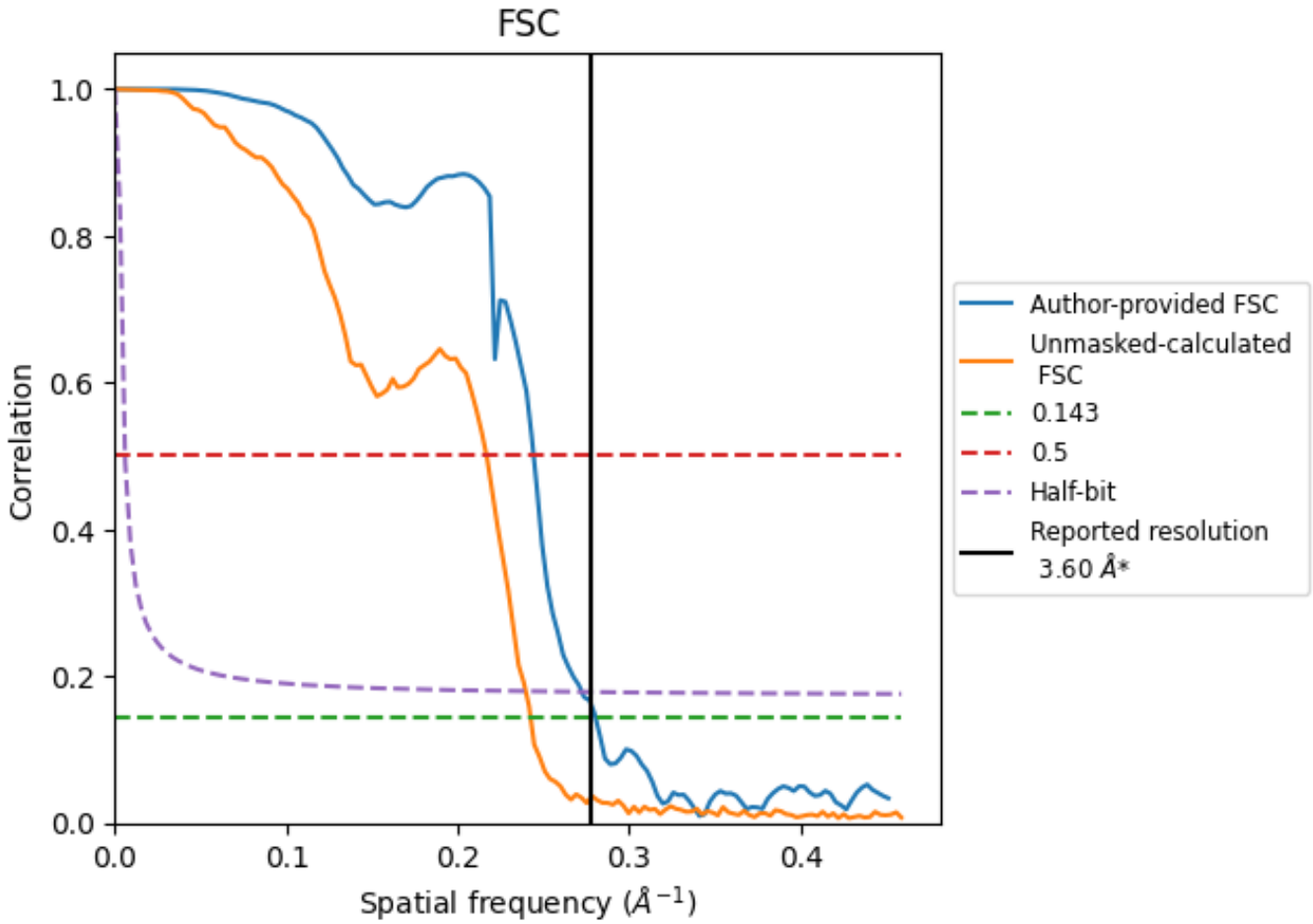



*Reported resolution corresponds to spatial frequency of 0.278 $Å^{-1}$

# 8 Fourier-Shell correlation

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




*Reported resolution corresponds to spatial frequency of 0.278 Å$^{-1}$

## 8.2 Resolution estimates (i)

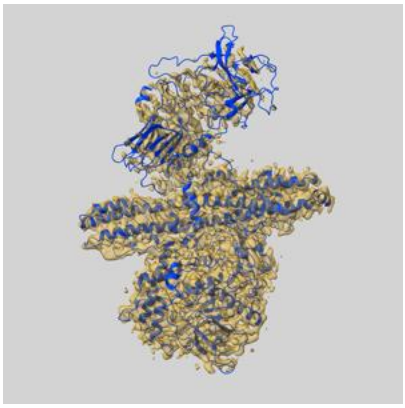
| Resolution estimate (Å) | Estimation criterion (FSC cut-off) | | |
|---|---|---|---|
| | 0.143 | 0.5 | Half-bit |
| Reported by author | 3.60 | - | - |
| Author-provided FSC curve | 3.57 | 4.09 | 3.67 |
| Unmasked-calculated* | 4.12 | 4.61 | 4.17 |

*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 4.12 differs from the reported value 3.6 by more than 10 %

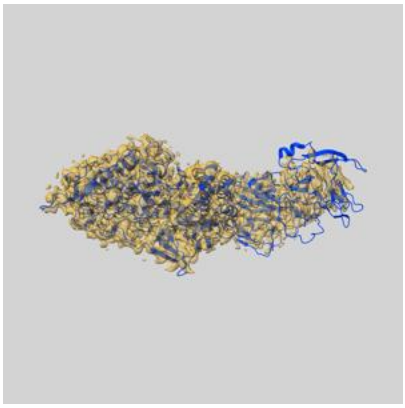
# 9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-13947 and PDB model 7QFQ. Per-residue inclusion information can be found in section 3 on page 4.

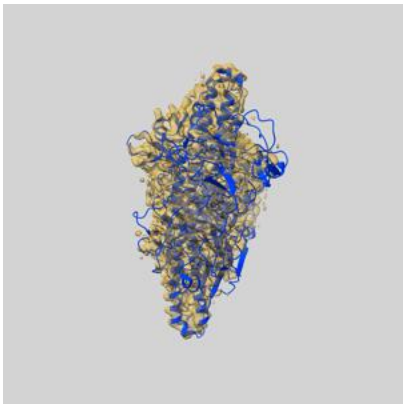
## 9.1 Map-model overlay



X



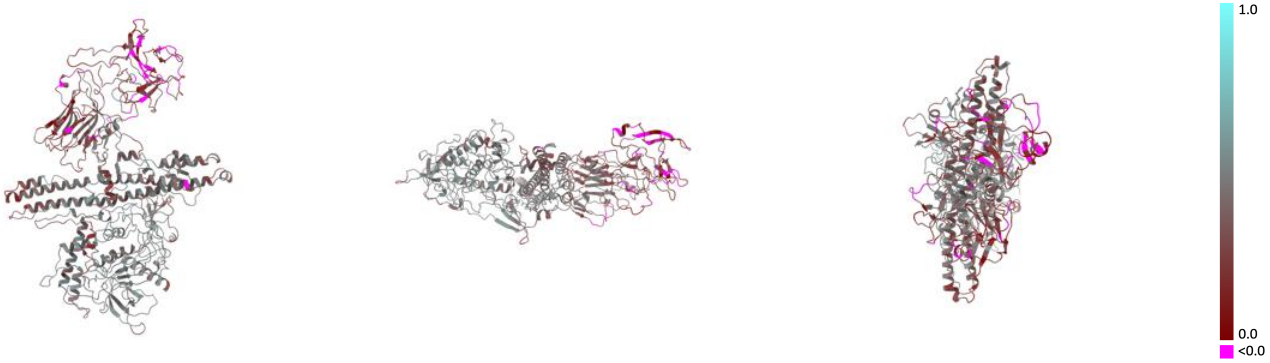
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.252 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.

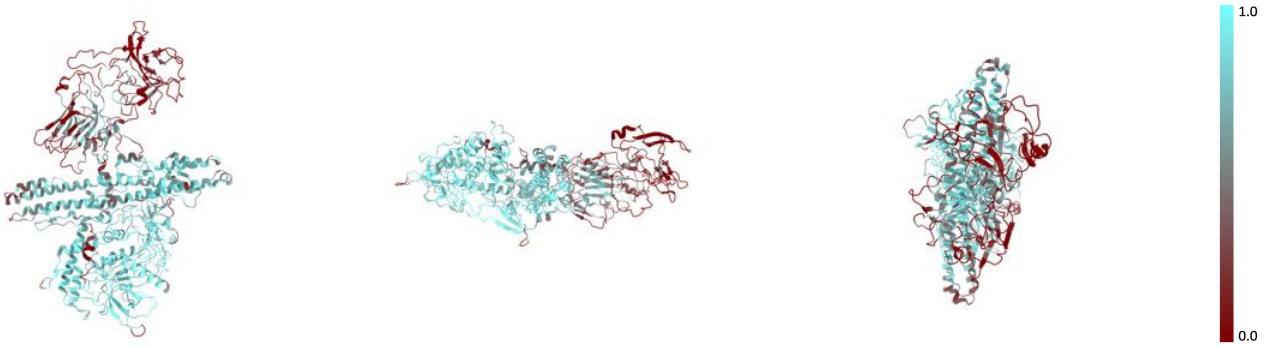
## 9.2 Q-score mapped to coordinate model




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.

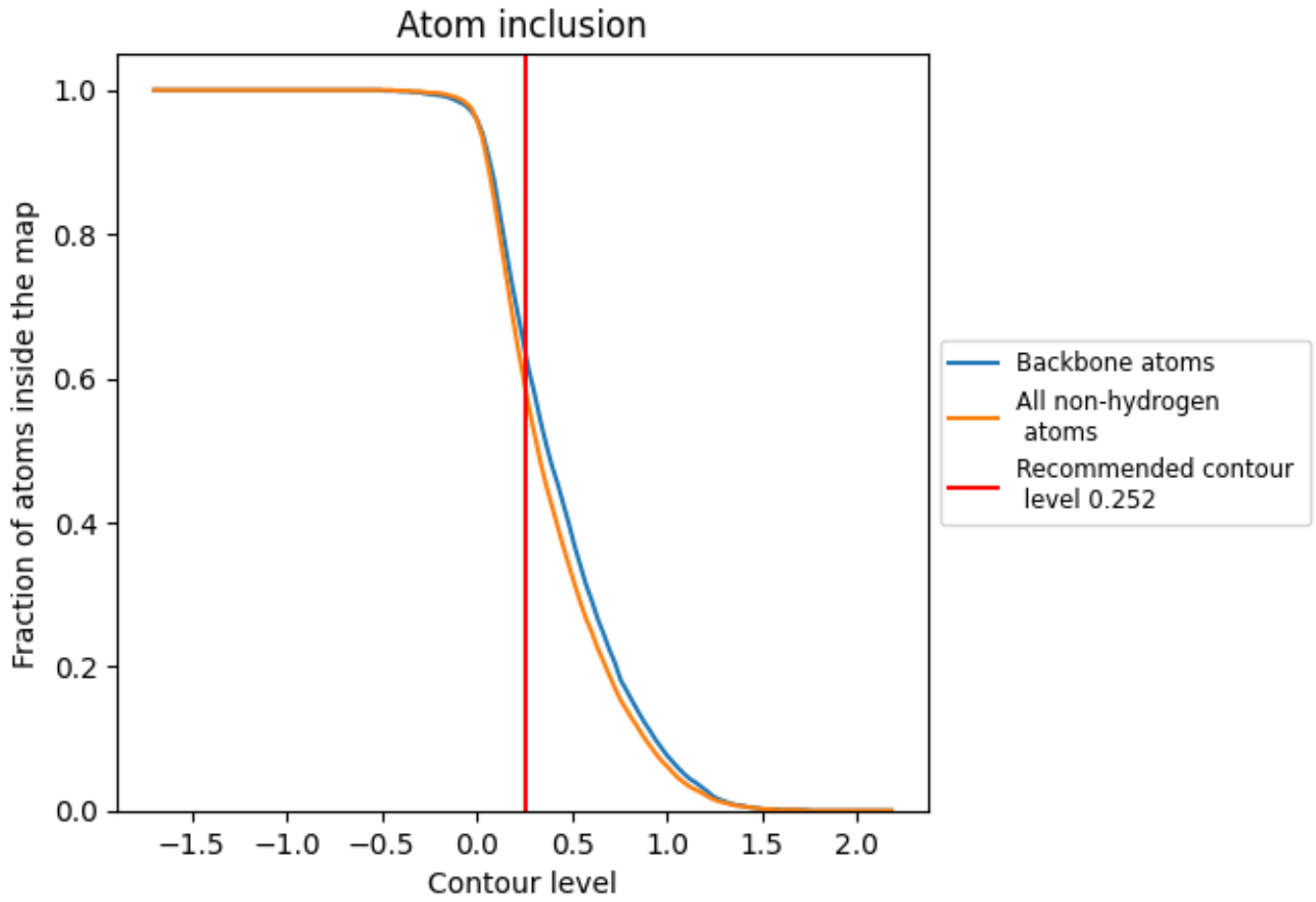
## 9.3 Atom inclusion mapped to coordinate model




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 (0.252).

## 9.4 Atom inclusion (i)




At the recommended contour level, 64% of all backbone atoms, 59% of all non-hydrogen atoms, are inside the map.

## 9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.252) and Q-score for the entire model and for each chain.


| Chain | Atom inclusion | Q-score |
|---|---|---|
| All | 0.5900 | 0.3800 |
| A | 0.5900 | 0.3800 |