# wwPDB EM Validation Summary Report

Sep 28, 2024 – 08:16 PM EDT

| | | |
|---|---|---|
| PDB ID | : | 8D1B |
| EMDB ID | : | EMD-27121 |
| Title | : | CryoEM structure of human orphan GPCR GPR179 in complex with extracellular matrix protein pikachurin |
| Authors | : | Patil, D.N.; Martemyanov, K.A. |
| Deposited on | : | 2022-05-27 |
| Resolution | : | 3.57 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:

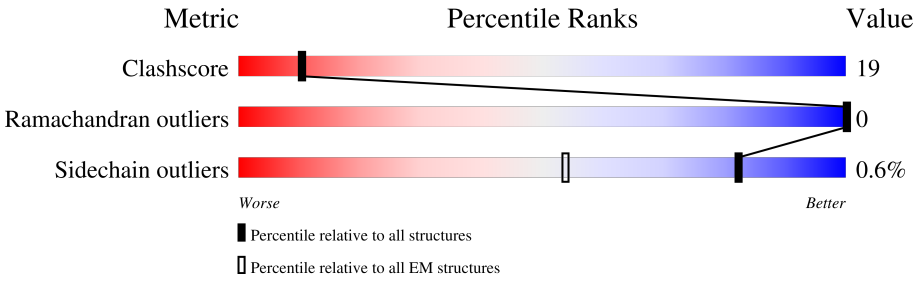
| | | |
|---|---|---|
| EMDB validation analysis | : | 0.0.1.dev113 |
| MolProbity | : | 4.02b-467 |
| Percentile statistics | : | 20231227.v01 (using entries in the PDB archive December 27th 2023) |
| MapQ | : | 1.9.13 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, 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 3.57 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 | C | 234 | 22% 46% 30% 24% |
| 1 | D | 234 | 23% 46% 30% 24% |
| 2 | A | 740 | 23% 12% 65% |
| 2 | B | 740 | 22% 13% 65% |

# 2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6724 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 Pikachurin.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 1 | D | 178 | Total | C | N | O | S | 0 | 0 |
| | | | 1379 | 862 | 253 | 255 | 9 | | |
| 1 | C | 178 | Total | C | N | O | S | 0 | 0 |
| | | | 1379 | 862 | 253 | 255 | 9 | | |

There are 2 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| D | 1017 | CYS | LYS | conflict | UNP Q63HQ2 |
| C | 1017 | CYS | LYS | conflict | UNP Q63HQ2 |

- Molecule 2 is a protein called Probable G-protein coupled receptor 179.

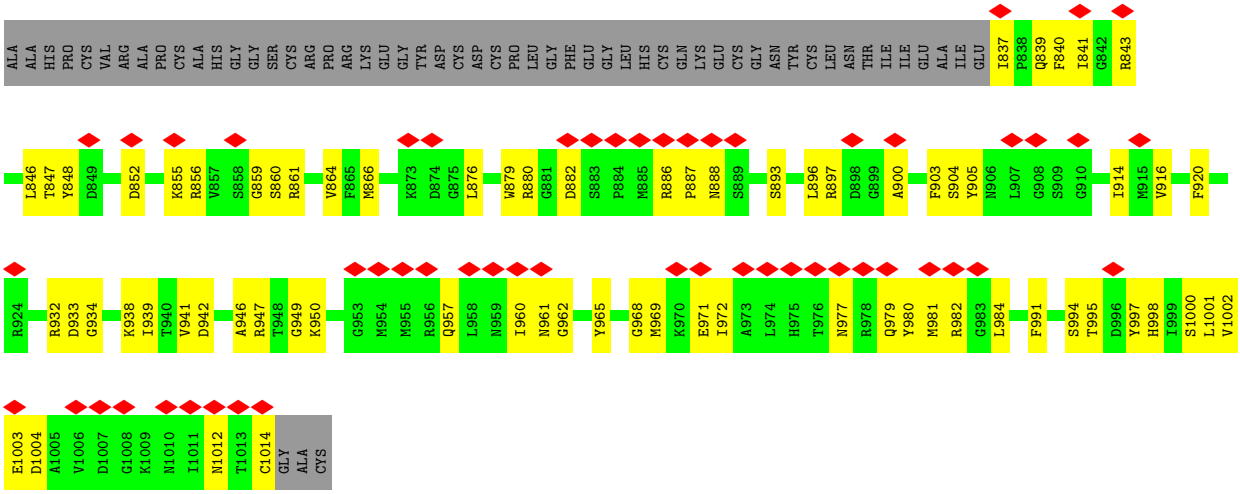
| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|
| 2 | B | 259 | Total | C | N | O | S | 0 | 0 |
| | | | 1983 | 1236 | 349 | 388 | 10 | | |
| 2 | A | 259 | Total | C | N | O | S | 0 | 0 |
| | | | 1983 | 1236 | 349 | 388 | 10 | | |

# 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.

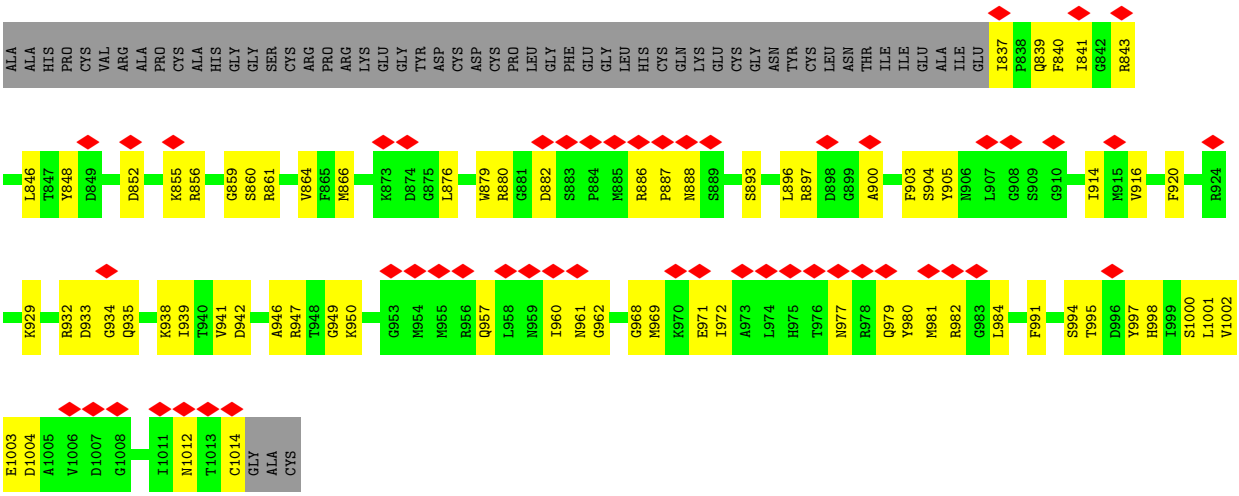
• Molecule 1: Pikachurin

Chain D:



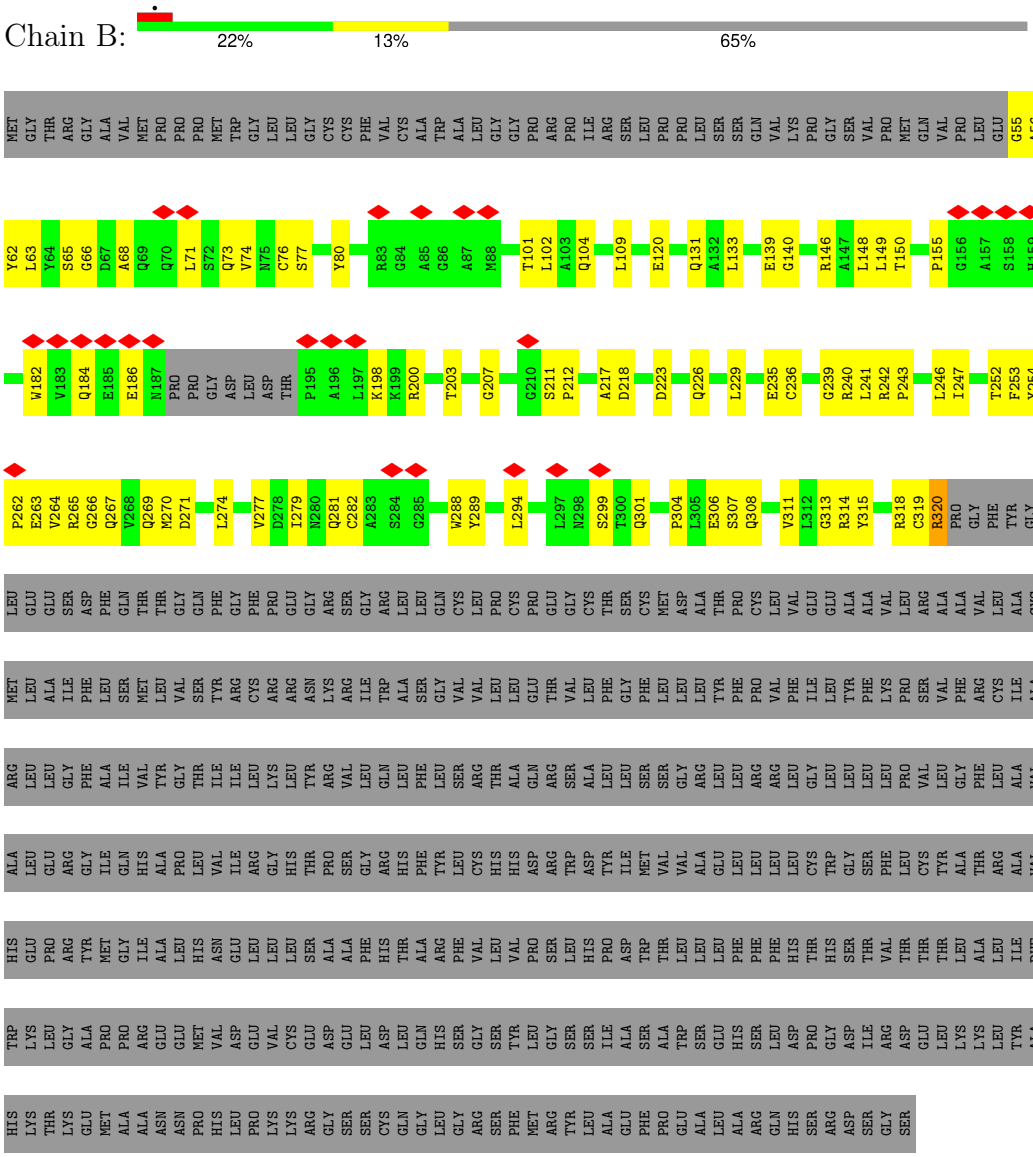

• Molecule 1: Pikachurin

Chain C:



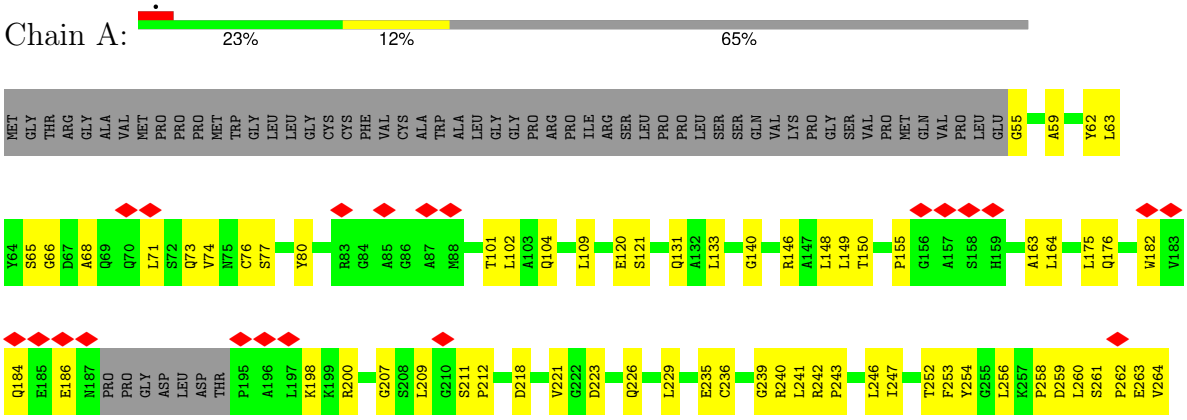
• Molecule 2: Probable G-protein coupled receptor 179

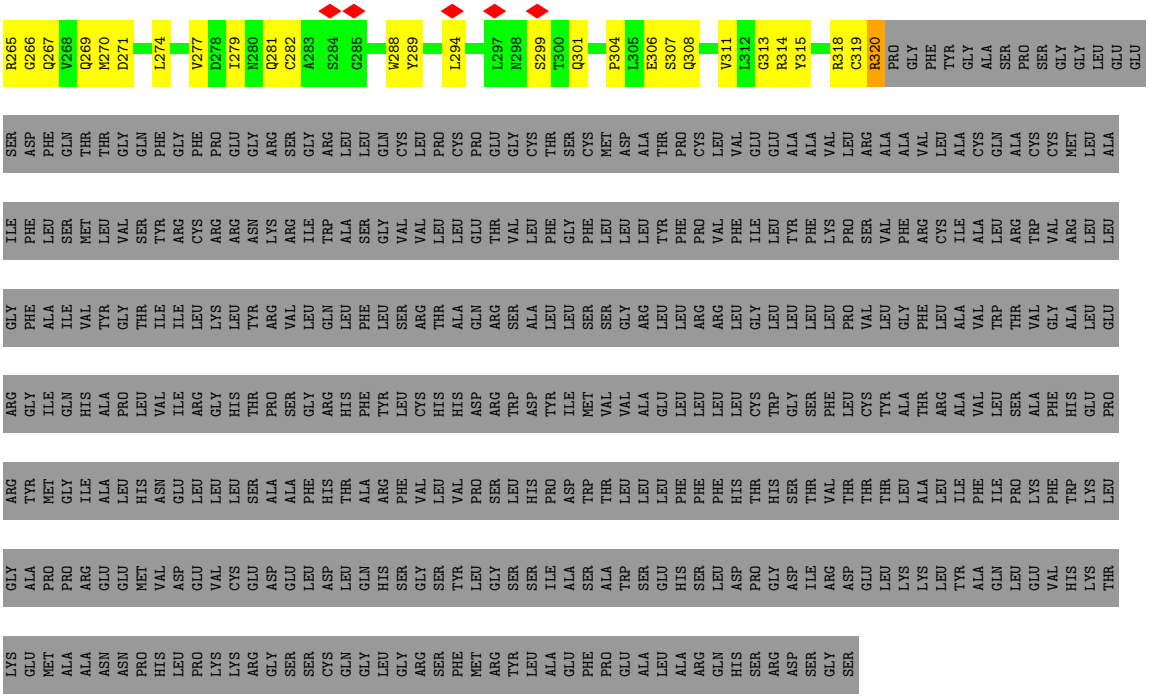
Chain B:




• Molecule 2: Probable G-protein coupled receptor 179

Chain A:




R265 G266 Q267 V268 Q269 M270 D271 L274 V277 D278 I279 N280 Q281 C282 A283 S284 G285 W288 Y289 L294 L297 N298 S299 T300 Q301 P304 L305 E306 S307 Q308 V311 L312 G313 R314 Y315 R318 C319 R320 PRO GLY PHE TYR GLY ALA SER PRO SER GLY GLY LEU GLU GLU
SER ASP PHE GLN THR THR GLY GLN PHE GLY PHE PRO GLU GLY ARG SER GLY ARG LEU LEU GLN CYS LEU PRO CYS PRO GLU GLY CYS THR SER CYS MET ASP ALA THR PRO CYS LEU VAL GLU GLU ALA ALA VAL LEU ARG ALA ALA VAL LEU ALA CYS GLN ALA CYS CYS MET LEU ALA
ILE PHE LEU SER MET LEU VAL SER TYR ARG CYS ARG ARG ASN LYS ARG ILE TRP ALA SER GLY VAL VAL LEU LEU GLU THR VAL LEU PHE GLY PHE LEU LEU LEU TYR PHE PRO VAL PHE ILE LEU TYR PHE LYS PRO SER VAL PHE ARG CYS ILE ALA LEU ARG TRP VAL ARG LEU LEU
GLY PHE ALA ILE VAL TYR GLY THR ILE ILE LEU LYS LEU TYR ARG VAL LEU GLN LEU PHE LEU SER ARG THR ALA GLN ARG SER ALA LEU LEU SER SER GLY ARG LEU LEU ARG ARG LEU GLY LEU LEU LEU LEU PRO VAL LEU GLY PHE LEU ALA VAL TRP THR VAL GLY ALA LEU GLU
ARG GLY ILE GLN HIS ALA PRO LEU VAL ILE ARG GLY HIS THR PRO SER GLY ARG HIS PHE TYR LEU CYS HIS HIS ASP ARG TRP ASP TYR ILE MET VAL VAL ALA GLU LEU LEU LEU LEU CYS TRP GLY SER PHE LEU CYS TYR ALA THR ARG ALA VAL LEU SER ALA PHE HIS GLU PRO
ARG TYR MET GLY ILE ALA LEU HIS ASN GLU LEU LEU LEU SER ALA ALA PHE HIS THR ALA ARG PHE VAL LEU VAL PRO SER LEU HIS PRO ASP TRP THR LEU LEU LEU PHE PHE PHE HIS THR HIS SER THR VAL THR THR THR LEU ALA LEU ILE PHE ILE PRO LYS PHE TRP LYS LEU
GLY ALA PRO PRO ARG GLU GLU MET VAL ASP GLU VAL CYS GLU ASP GLU LEU ASP LEU GLN HIS SER GLY SER TYR LEU GLY SER SER ILE ALA SER ALA TRP SER GLU HIS SER LEU ASP PRO GLY ASP ILE ARG ASP GLU LEU LYS LYS LEU TYR ALA GLN LEU GLU VAL HIS LYS THR
LYS GLU MET ALA ALA ASN ASN PRO HIS LEU PRO LYS LYS ARG GLY SER SER CYS GLN GLY LEU GLY ARG SER PHE MET ARG TYR LEU ALA GLU PHE PRO GLU ALA LEU ALA ARG GLN HIS SER ARG ASP SER GLY SER

# 4 Experimental information (i)

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 217570 | 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 ($e^-/Å^2$) | 40 | 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 | 3.352 | Depositor |
| Minimum map value | -1.821 | Depositor |
| Average map value | -0.001 | Depositor |
| Map value standard deviation | 0.041 | Depositor |
| Recommended contour level | 0.54 | Depositor |
| Map size (Å) | 410.40002, 410.40002, 410.40002 | wwPDB |
| Map dimensions | 480, 480, 480 | wwPDB |
| Map angles (°) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (Å) | 0.8550001, 0.8550001, 0.8550001 | 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 | C | 0.30 | 0/1405 | 0.57 | 0/1892 |
| 1 | D | 0.30 | 0/1405 | 0.57 | 0/1892 |
| 2 | A | 0.36 | 0/2024 | 0.58 | 0/2755 |
| 2 | B | 0.36 | 0/2024 | 0.58 | 0/2755 |
| All | All | 0.34 | 0/6858 | 0.57 | 0/9294 |

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 | C | 1379 | 0 | 1367 | 61 | 0 |
| 1 | D | 1379 | 0 | 1367 | 61 | 0 |
| 2 | A | 1983 | 0 | 1916 | 72 | 0 |
| 2 | B | 1983 | 0 | 1916 | 73 | 0 |
| All | All | 6724 | 0 | 6566 | 247 | 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 19.

The worst 5 of 247 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) |
|---|---|---|---|
| 2:B:73:GLN:OE1 | 2:B:240:ARG:NH2 | 2.08 | 0.86 |
| 2:A:73:GLN:OE1 | 2:A:240:ARG:NH2 | 2.08 | 0.86 |
| 2:A:101:THR:HG21 | 2:A:140:GLY:HA3 | 1.59 | 0.85 |
| 2:B:101:THR:HG21 | 2:B:140:GLY:HA3 | 1.59 | 0.83 |
| 1:C:950:LYS:NZ | 2:A:259:ASP:OD1 | 2.10 | 0.83 |

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 | C | 176/234 (75%) | 130 (74%) | 46 (26%) | 0 | 100 | 100 |
| 1 | D | 176/234 (75%) | 130 (74%) | 46 (26%) | 0 | 100 | 100 |
| 2 | A | 255/740 (34%) | 197 (77%) | 58 (23%) | 0 | 100 | 100 |
| 2 | B | 255/740 (34%) | 197 (77%) | 58 (23%) | 0 | 100 | 100 |
| All | All | 862/1948 (44%) | 654 (76%) | 208 (24%) | 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 | C | 149/192 (78%) | 148 (99%) | 1 (1%) | 81 | 90 |
| 1 | D | 149/192 (78%) | 148 (99%) | 1 (1%) | 81 | 90 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 2 | A | 212/617 (34%) | 211 (100%) | 1 (0%) | 86 93 |
| 2 | B | 212/617 (34%) | 211 (100%) | 1 (0%) | 86 93 |
| All | All | 722/1618 (45%) | 718 (99%) | 4 (1%) | 82 92 |

All (4) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | D | 998 | HIS |
| 1 | C | 998 | HIS |
| 2 | B | 320 | ARG |
| 2 | A | 320 | ARG |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | B | 165 | GLN |
| 2 | A | 165 | 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 oligosaccharides 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 (i)

There are no chain breaks in this entry.

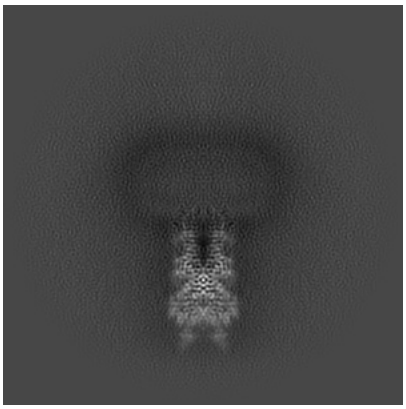
# 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-27121. These allow visual inspection of the internal detail of the map and identification of artifacts.

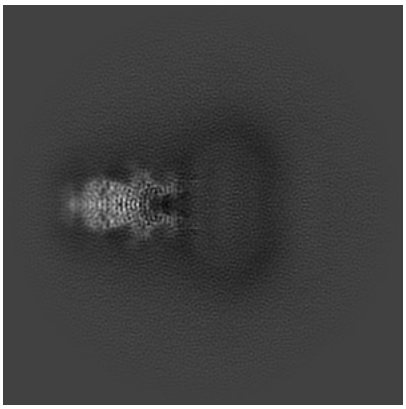
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

## 6.1 Orthogonal projections

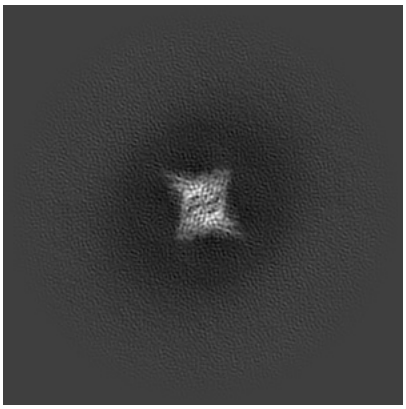
### 6.1.1 Primary map



X

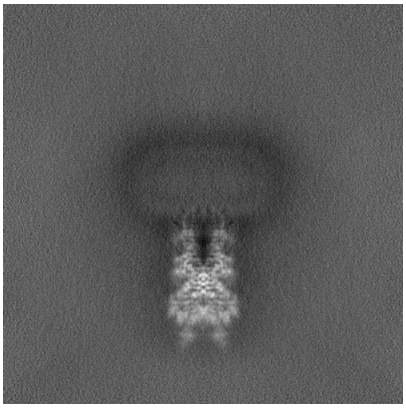


Y

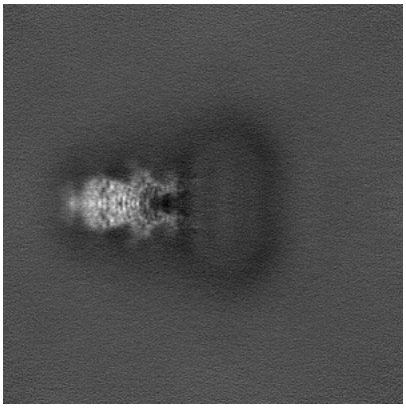


Z

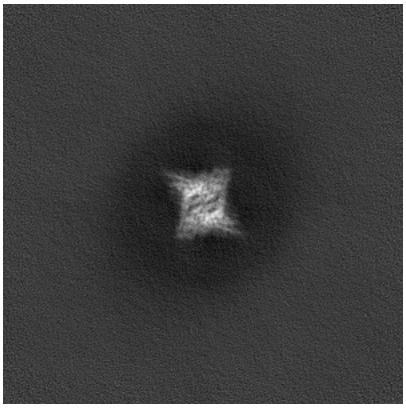
### 6.1.2 Raw map



X



Y

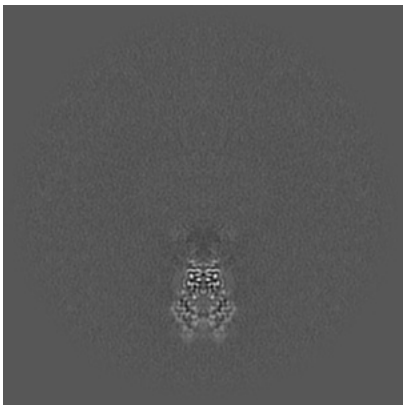


Z

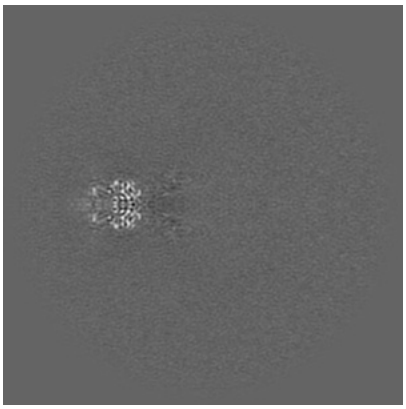
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices (i)

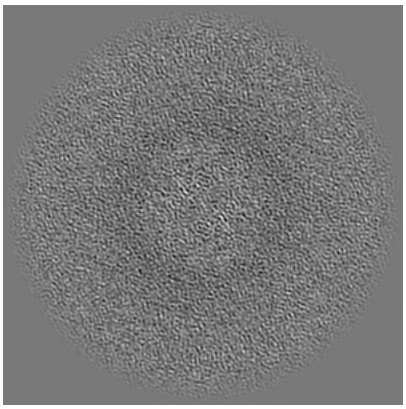
### 6.2.1 Primary map



X Index: 240

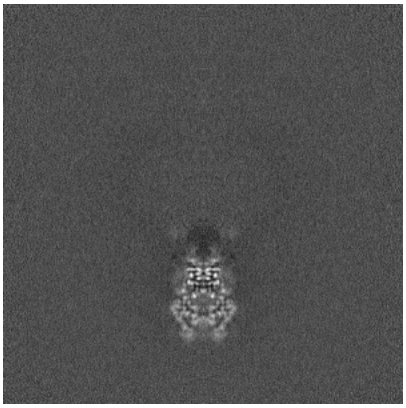


Y Index: 240

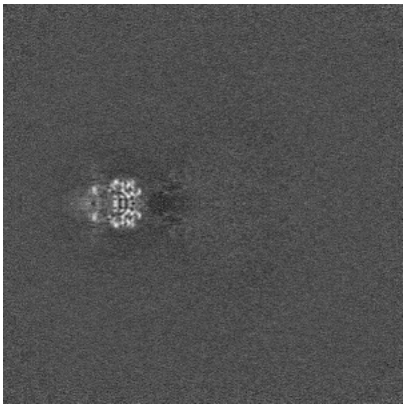


Z Index: 240

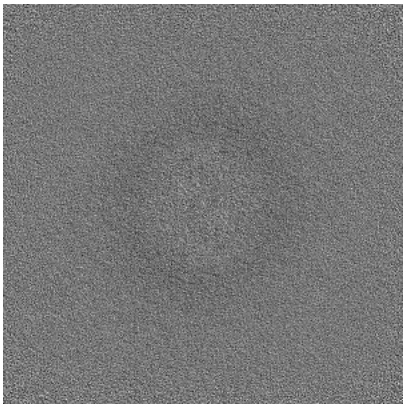
### 6.2.2 Raw map



X Index: 240



Y Index: 240

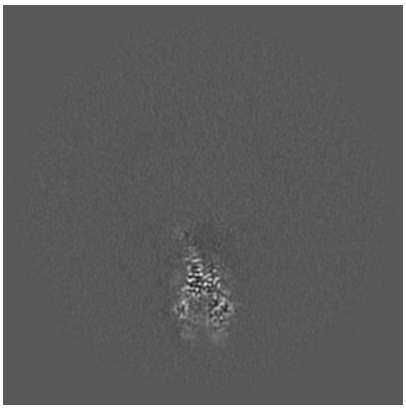


Z Index: 240

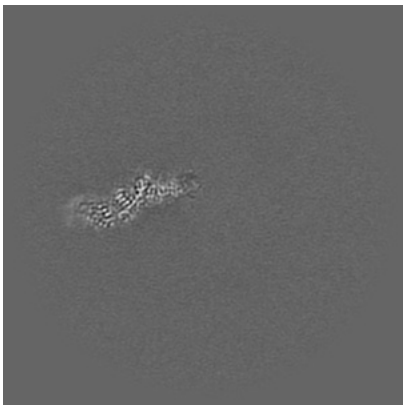
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices (i)

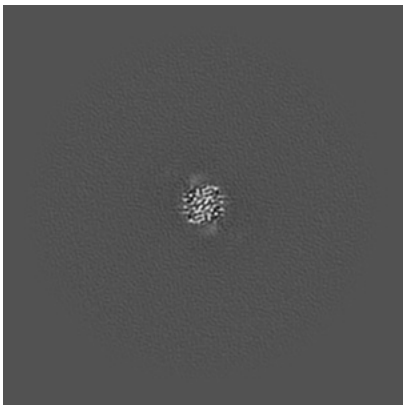
### 6.3.1 Primary map



X Index: 243

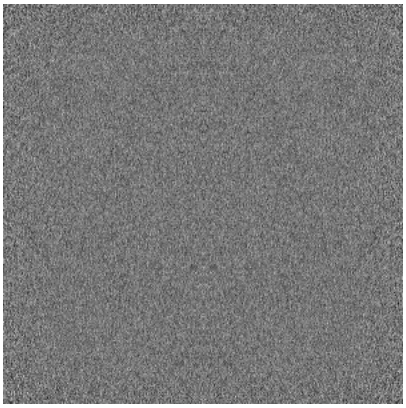


Y Index: 224

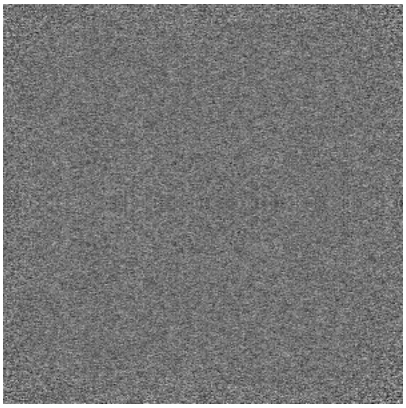


Z Index: 145

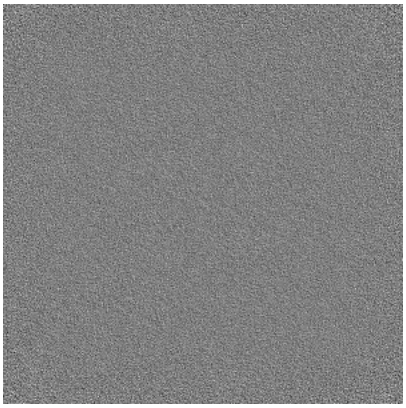
### 6.3.2 Raw map



X Index: 0



Y Index: 0

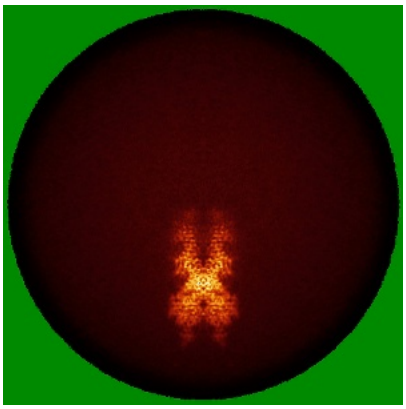


Z Index: 0

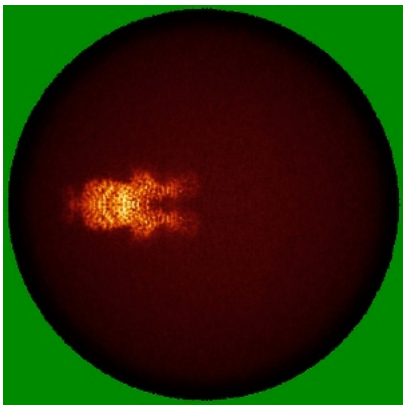
The images above show the largest variance slices of the map in three orthogonal directions.

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

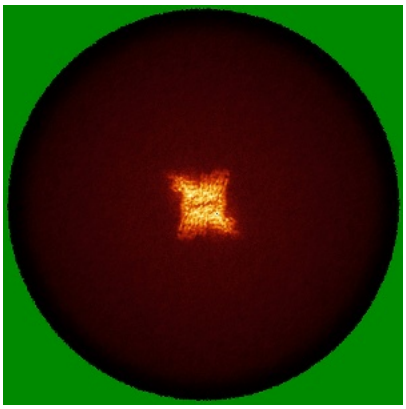
### 6.4.1 Primary map



X

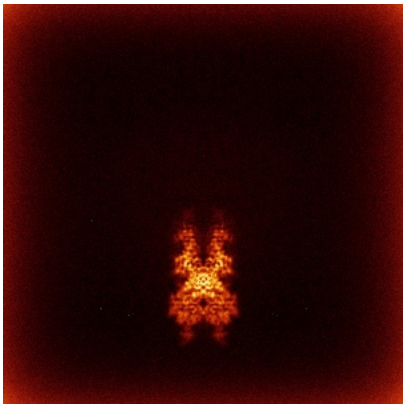


Y

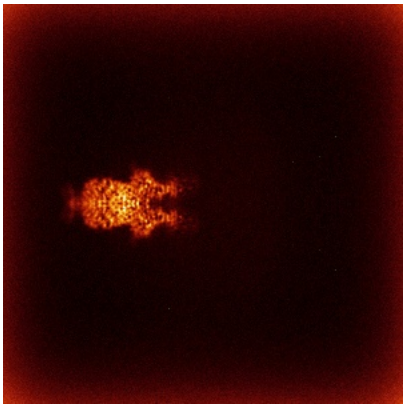


Z

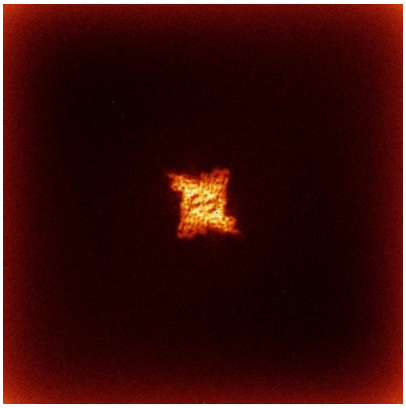
### 6.4.2 Raw map



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 0.54. 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.

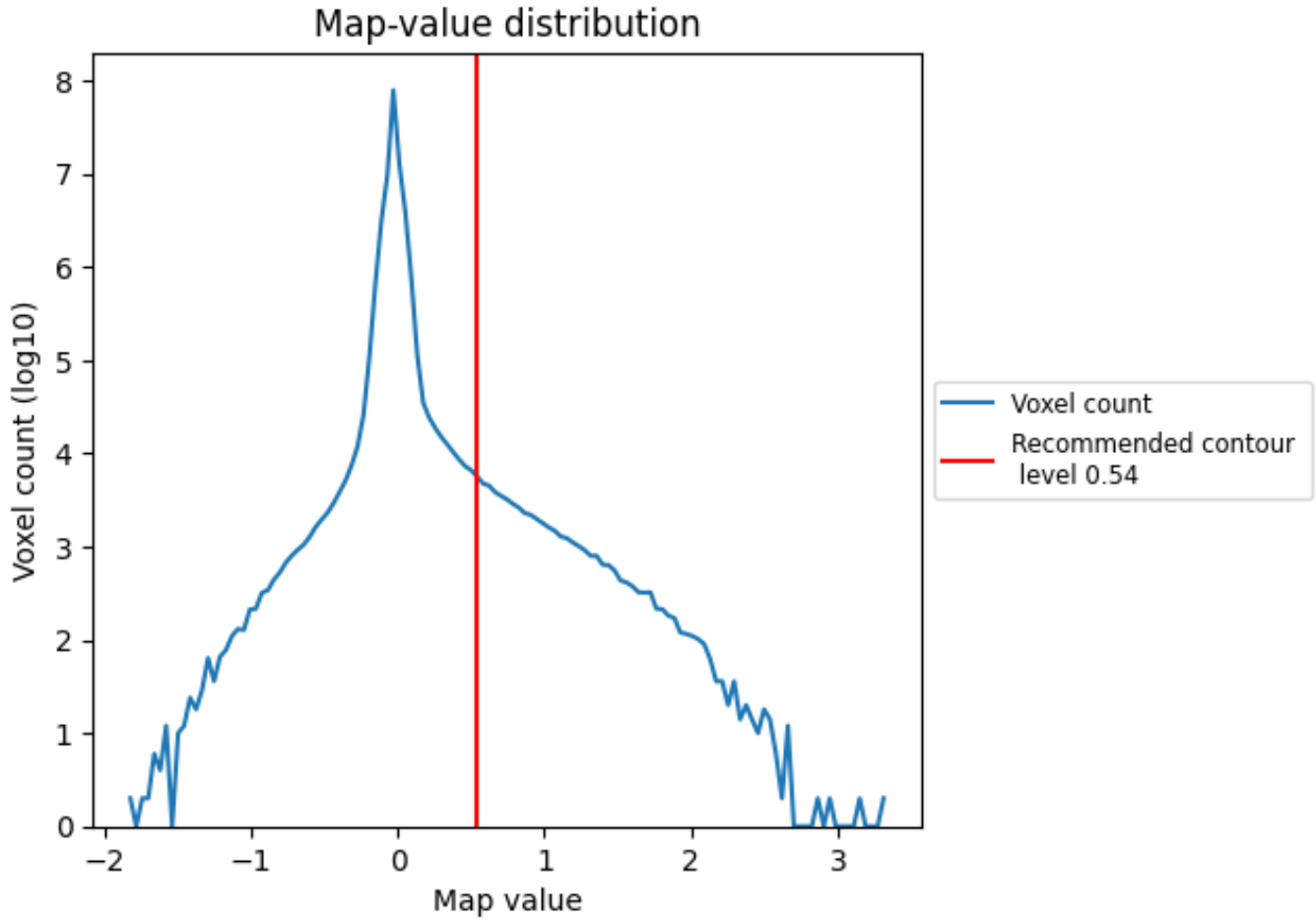
## 6.6 Mask visualisation (i)

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

# 7 Map analysis

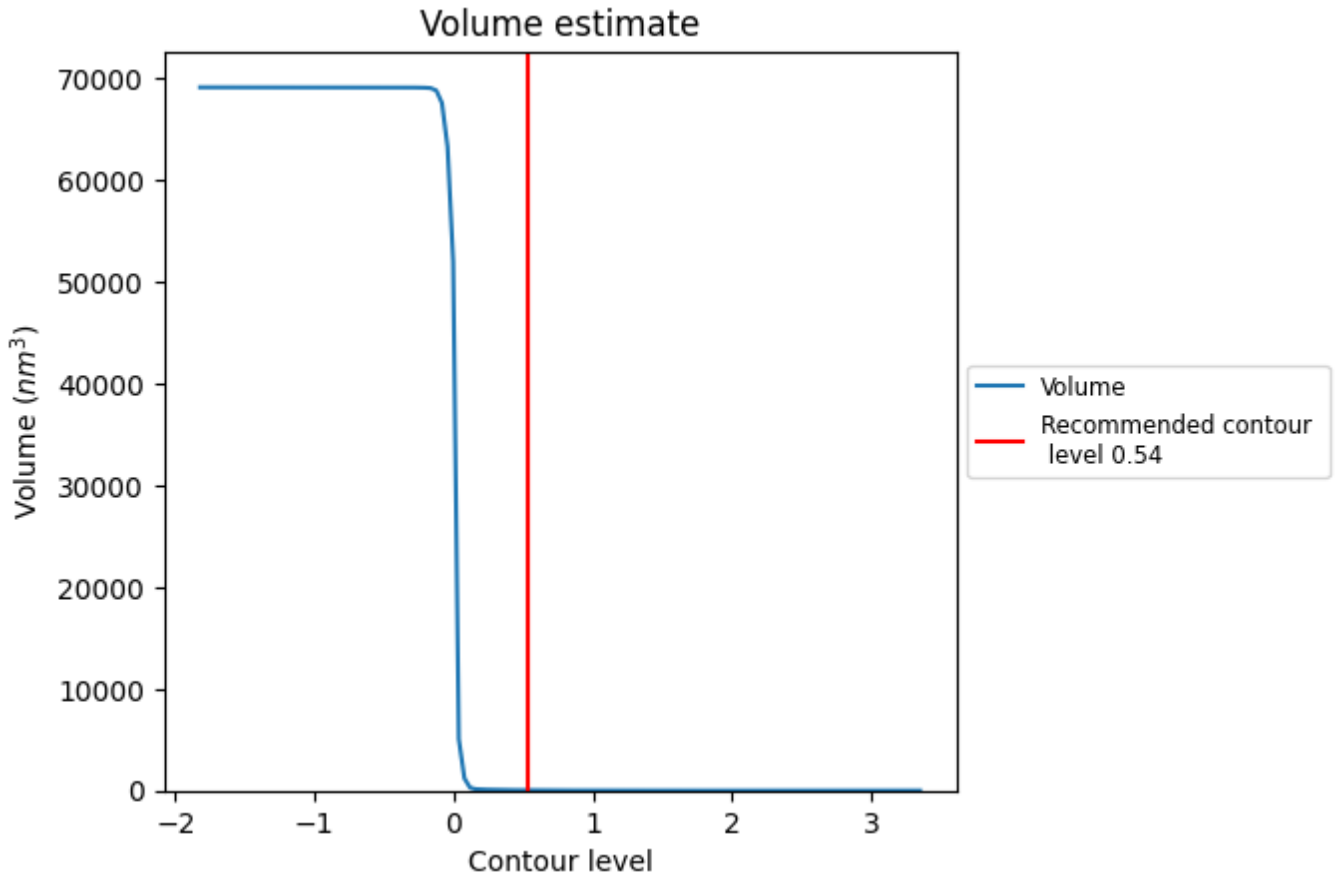
This section contains the results of statistical analysis of the map.

## 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.

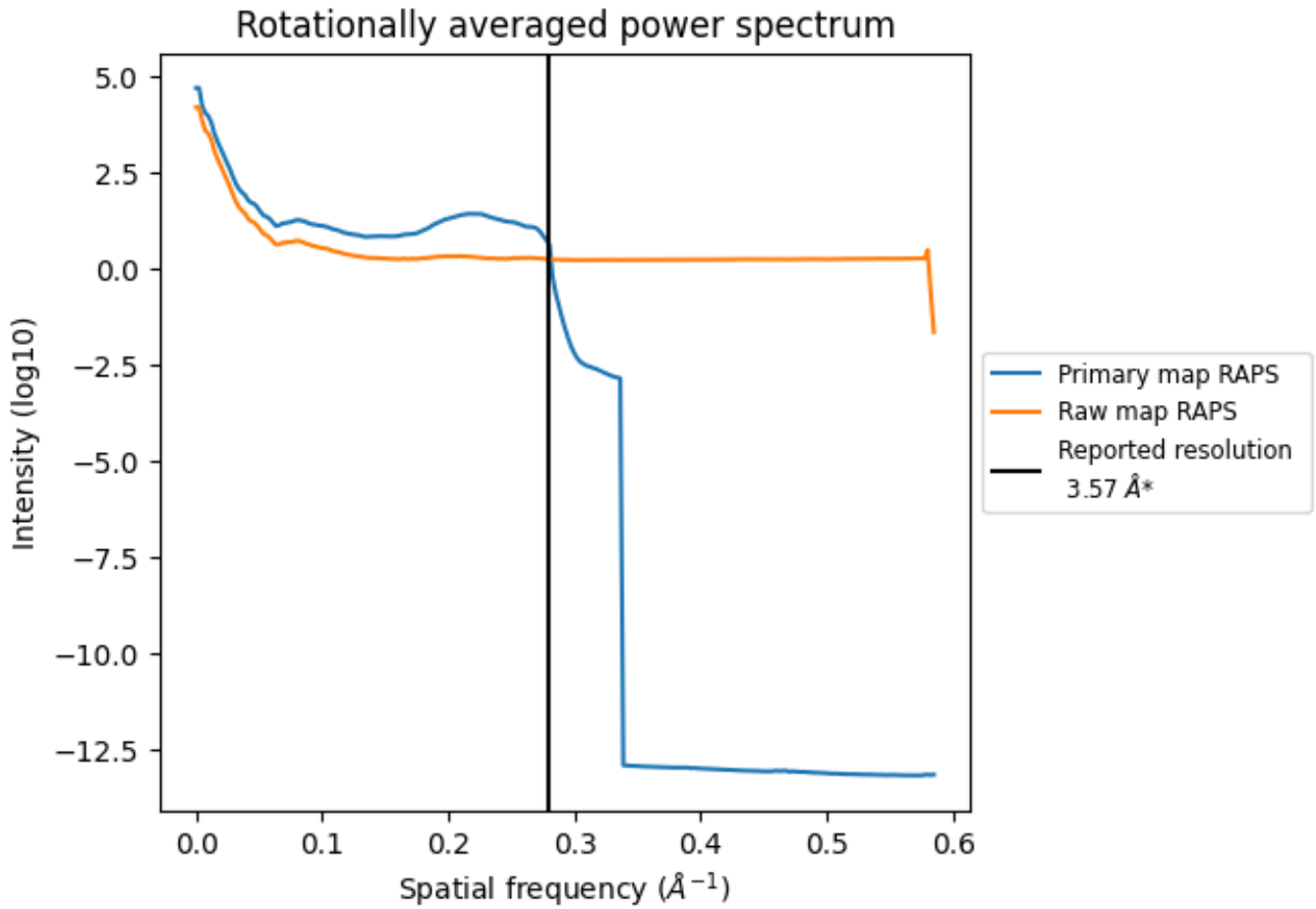
## 7.2 Volume estimate (i)




The volume at the recommended contour level is 35 $nm^3$; this corresponds to an approximate mass of 31 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)

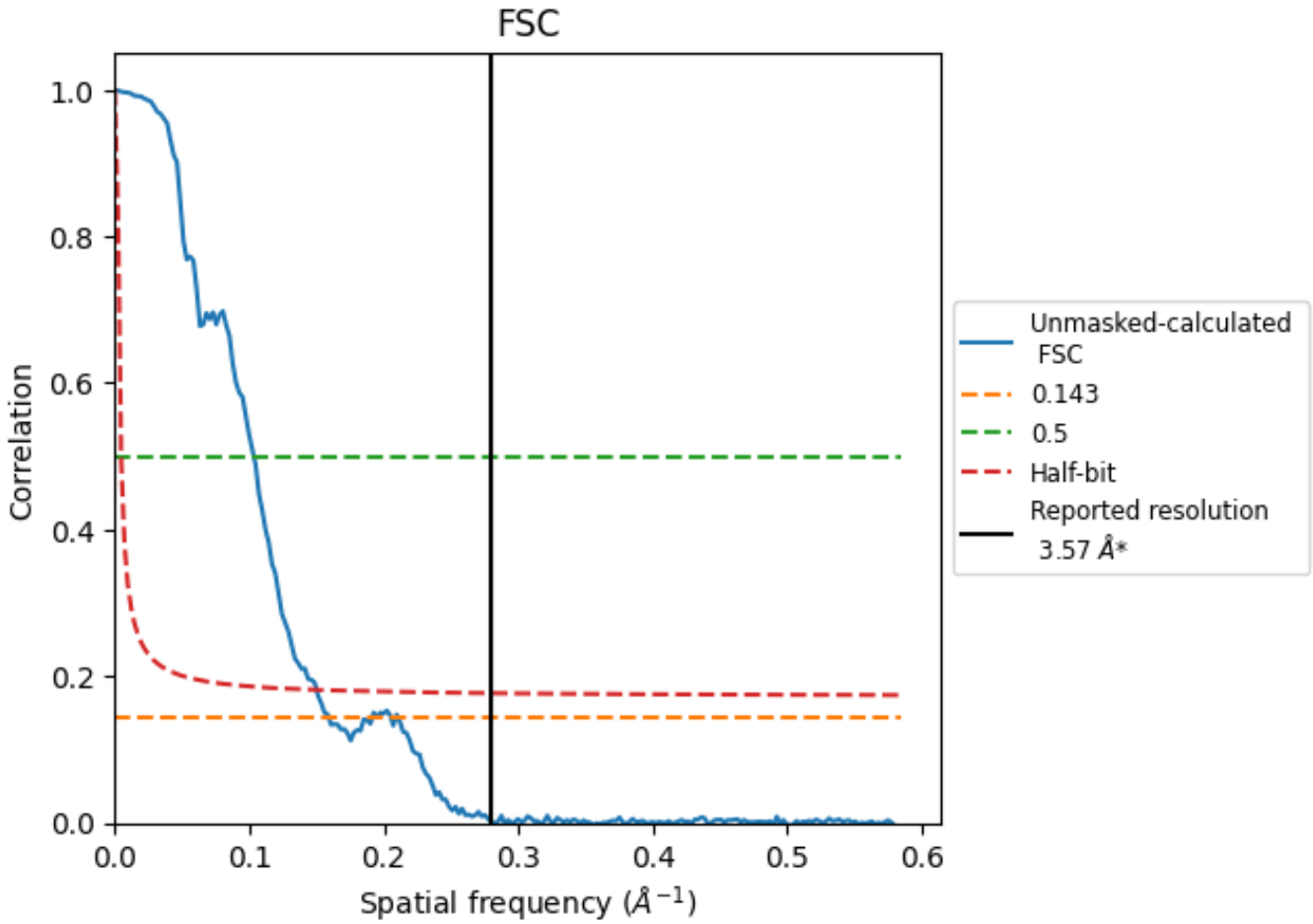



*Reported resolution corresponds to spatial frequency of 0.280 $Å^{-1}$

# 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 0.280 Å$^{-1}$

## 8.2 Resolution estimates (i)

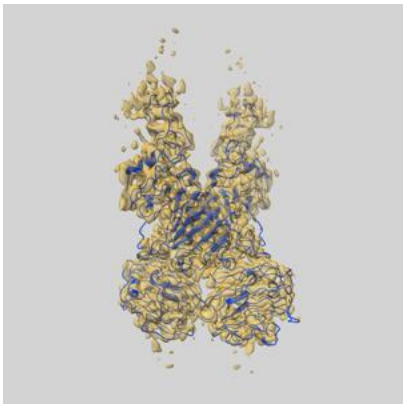
| Resolution estimate (Å) | Estimation criterion (FSC cut-off) | | |
|---|---|---|---|
| | 0.143 | 0.5 | Half-bit |
| Reported by author | 3.57 | - | - |
| Author-provided FSC curve | - | - | - |
| Unmasked-calculated* | 6.28 | 9.68 | 6.67 |

*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 6.28 differs from the reported value 3.57 by more than 10 %

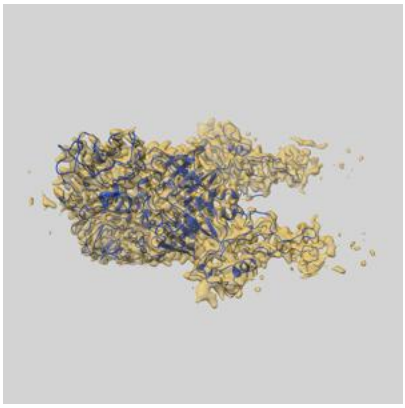
# 9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-27121 and PDB model 8D1B. Per-residue inclusion information can be found in section 3 on page 4.

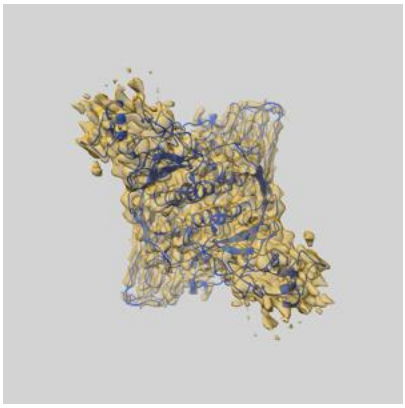
## 9.1 Map-model overlay



X



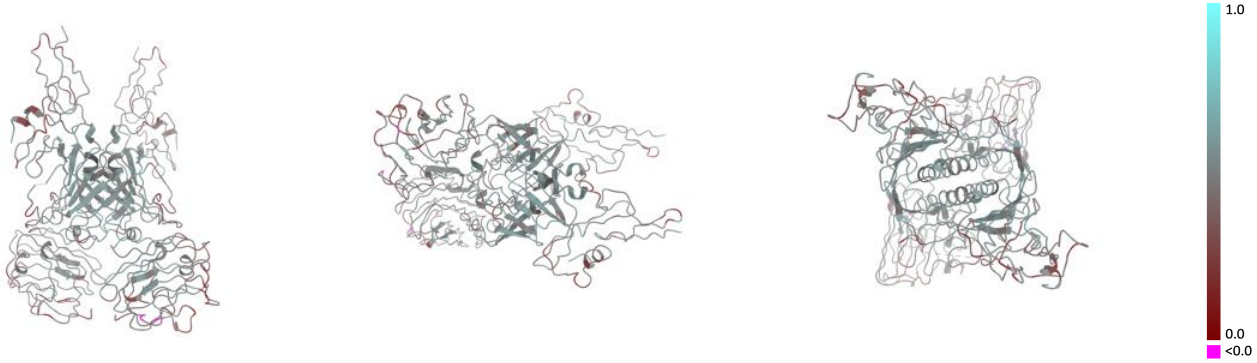
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.54 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.

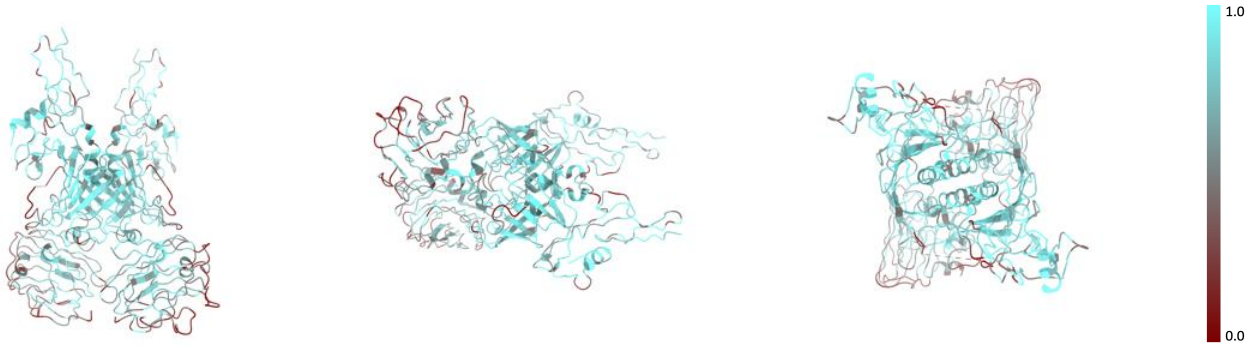
## 9.2 Q-score mapped to coordinate model (i)




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.

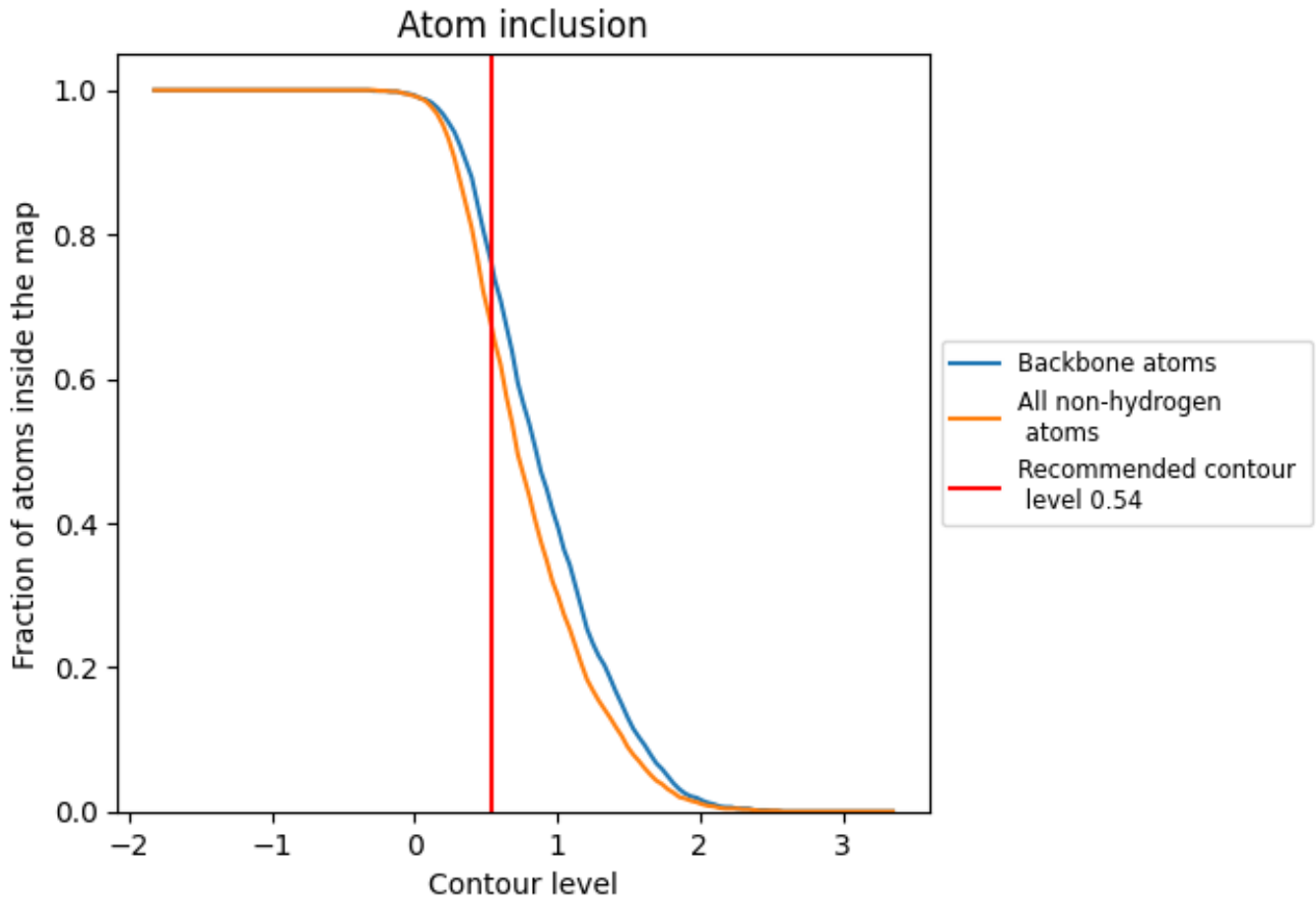
## 9.3 Atom inclusion mapped to coordinate model (i)




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 (0.54).

## 9.4 Atom inclusion (i)




At the recommended contour level, 76% of all backbone atoms, 67% 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 (0.54) and Q-score for the entire model and for each chain.

| Chain | Atom inclusion | Q-score |
|---|---|---|
| All | 0.6730 | 0.4760 |
| A | 0.7450 | 0.4920 |
| B | 0.7430 | 0.4920 |
| C | 0.5720 | 0.4500 |
| D | 0.5670 | 0.4550 |