# WORLDWIDE PDB

## PROTEIN DATA BANK

# wwPDB EM Validation Summary Report (i)

Apr 28, 2024 – 12:03 am BST

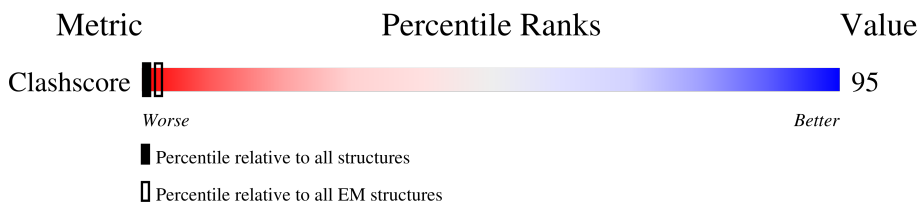| | | |
|---|---|---|
| PDB ID | : | 3ZFS |
| EMDB ID | : | EMD-2097 |
| Title | : | Cryo-EM structure of the F420-reducing NiFe-hydrogenase from a methanogenic archaeon with bound substrate |
| Authors | : | Mills, D.J.; Vitt, S.; Strauss, M.; Shima, S.; Vonck, J. |
| Deposited on | : | 2012-12-12 |
| Resolution | : | 4.00 Å(reported) |
| Based on initial model | : | 2WPN |

# 1 Overall quality at a glance ⓘ

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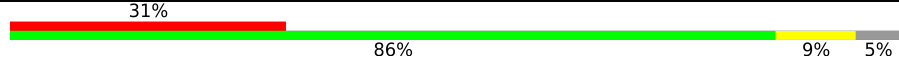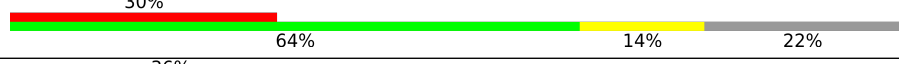*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.00 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 | | 95 |

*Worse*               *Better*

▌Percentile relative to all structures

▯Percentile relative to all EM structures

| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|--------|--------------------------|--------------------------|
| Clashscore | 158937 | 4297 |

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 | 405 | 31% ■■■ / 86% ... 9% 5% |
| 2 | B | 275 | 30% ■■■ / 64% ... 14% 22% |
| 3 | C | 281 | 36% ■■■ / 78% ... 21% • |

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

| Mol | Type | Chain | Res | Chirality | Geometry | Clashes | Electron density |
|-----|------|-------|-----|-----------|----------|---------|------------------|
| 7 | SF4 | B | 300 | - | - | X | - |
| 7 | SF4 | C | 303 | - | - | X | - |

WORLDWIDE wwPDB PROTEIN DATA BANK

# 2 Entry composition ⓘ

There are 9 unique types of molecules in this entry. The entry contains 985 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 F420-REDUCING HYDROGENASE, SUBUNIT ALPHA.

| Mol | Chain | Residues | Atoms | | AltConf | Trace |
|---|---|---|---|---|---|---|
| 1 | A | 385 | Total C<br>385 385 | | 0 | 385 |

- Molecule 2 is a protein called F420-REDUCING HYDROGENASE, SUBUNIT GAMMA.

| Mol | Chain | Residues | Atoms | | AltConf | Trace |
|---|---|---|---|---|---|---|
| 2 | B | 215 | Total C<br>215 215 | | 0 | 215 |

- Molecule 3 is a protein called F420-REDUCING HYDROGENASE, SUBUNIT BETA.

| Mol | Chain | Residues | Atoms | | AltConf | Trace |
|---|---|---|---|---|---|---|
| 3 | C | 278 | Total C<br>278 278 | | 0 | 278 |

- Molecule 4 is CARBONMONOXIDE-(DICYANO) IRON (three-letter code: FCO) (formula: $C_3FeN_2O$).
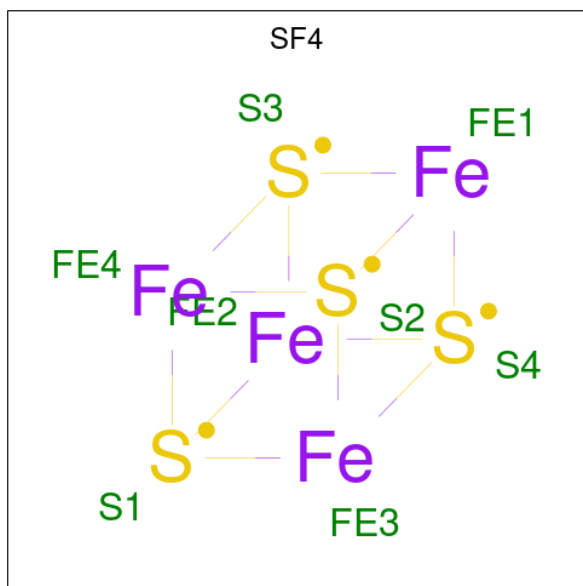


FCO
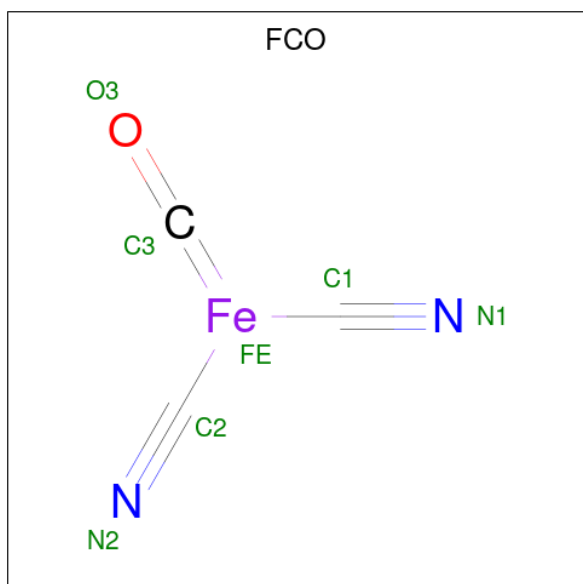
| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 4 | A | 1 | Total | Fe | 0 |
| | | | 1 | 1 | |

- Molecule 5 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 5 | A | 1 | Total | Ni | 0 |
| | | | 1 | 1 | |

- Molecule 6 is FE (II) ION (three-letter code: FE2) (formula: Fe).

| Mol | Chain | Residues | Atoms | | AltConf |
|---|---|---|---|---|---|
| 6 | A | 1 | Total | Fe | 0 |
| | | | 1 | 1 | |

- Molecule 7 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: $Fe_4S_4$).



| Mol | Chain | Residues | Atoms | | | AltConf |
|---|---|---|---|---|---|---|
| 7 | B | 1 | Total | Fe | S | 0 |
| | | | 8 | 4 | 4 | |
| 7 | B | 1 | Total | Fe | S | 0 |
| | | | 8 | 4 | 4 | |
| 7 | B | 1 | Total | Fe | S | 0 |
| | | | 8 | 4 | 4 | |
| 7 | C | 1 | Total | Fe | S | 0 |
| | | | 8 | 4 | 4 | |

- Molecule 8 is COENZYME F420 (three-letter code: F42) (formula: $C_{29}H_{36}N_5O_{18}P$).
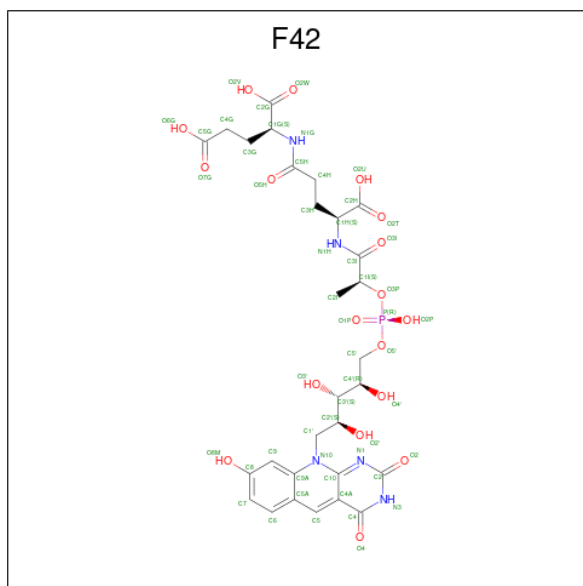


| Mol | Chain | Residues | Atoms | | | | AltConf |
|-----|-------|----------|-------|---|---|---|---------|
| | | | Total | C | N | O | |
| 8 | C | 1 | 19 | 13 | 3 | 3 | 0 |

- Molecule 9 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



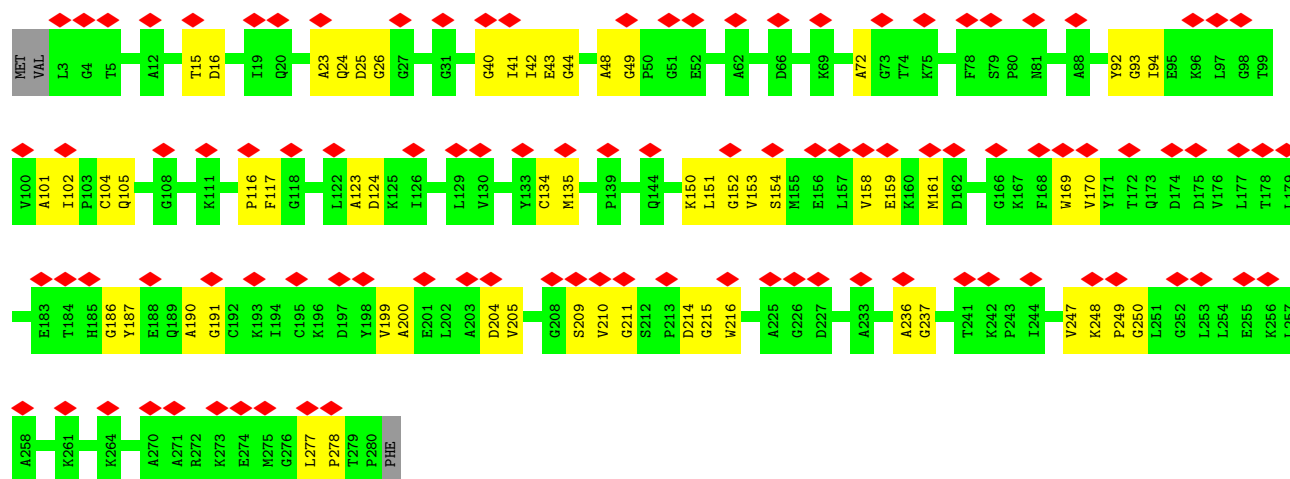| Mol | Chain | Residues | Atoms | | | | | AltConf |
|-----|-------|----------|-------|---|---|---|---|---------|
| | | | Total | C | N | O | P | |
| 9 | C | 1 | 53 | 27 | 9 | 15 | 2 | 0 |

# 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: F420-REDUCING HYDROGENASE, SUBUNIT ALPHA



Chain A:

• Molecule 2: F420-REDUCING HYDROGENASE, SUBUNIT GAMMA



Chain B:

L219 C220 I221 G224 A227 A234 P242 D247 R248 C249 I250 K251 C252 G253 S262 W263 W264 E267 K270 K271 E272 LEU GLY LEU

• Molecule 3: F420-REDUCING HYDROGENASE, SUBUNIT BETA

Chain C:

36%

78% 21%

MET VAL L3 G4 T5 A12 T15 D16 I19 Q20 A23 Q24 D25 G26 G27 G31 G40 I41 I42 E43 G44 A48 G49 P50 G51 E52 A62 D66 K69 A72 G73 T74 K75 F78 S79 P80 N81 A88 Y92 G93 I94 E95 K96 L97 G98 I99

V100 A101 I102 P103 C104 Q105 G108 K111 P116 F117 G118 L122 A123 D124 K125 I126 L129 V130 Y133 C134 M135 P139 Q144 K150 L151 G152 V153 S154 M155 E156 L157 V158 E159 K160 M161 D162 G166 K167 F168 W169 V170 Y171 T172 Q173 D174 D175 V176 L177 T178 L179

E183 T184 H185 G186 Y187 E188 Q189 A190 G191 C192 K193 I194 C195 K196 D197 V198 V199 A200 E201 L202 A203 D204 V205 G208 S209 V210 G211 S212 P213 D214 G215 W216 A225 G226 D227 A233 A236 G237 T241 K242 P243 I244 V247 K248 P249 G250 L251 G252 L253 L254 E255 K256 L257

A258 K261 K264 A270 A271 R272 K273 E274 M275 G276 L277 P278 T279 P280 PHE

# 4   Experimental information ⓘ

| Property | Value | Source |
|---|---|---|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, T | Depositor |
| Number of particles used | 97290 | Depositor |
| Resolution determination method | Not provided | |
| CTF correction method | PER MICROGRAPH | Depositor |
| Microscope | FEI POLARA 300 | Depositor |
| Voltage (kV) | 200 | Depositor |
| Electron dose $(e^-/\text{Å}^2)$ | 15 | Depositor |
| Minimum defocus (nm) | 1650 | Depositor |
| Maximum defocus (nm) | 3820 | Depositor |
| Magnification | 61400 | Depositor |
| Image detector | KODAK SO-163 FILM | Depositor |
| Maximum map value | 1390.650 | Depositor |
| Minimum map value | -464.710 | Depositor |
| Average map value | 1.698 | Depositor |
| Map value standard deviation | 46.708 | Depositor |
| Recommended contour level | 212.0 | Depositor |
| Map size (Å) | 328.32, 328.32, 328.32 | wwPDB |
| Map dimensions | 288, 288, 288 | wwPDB |
| Map angles (°) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (Å) | 1.14, 1.14, 1.14 | Depositor |

# 5  Model quality ⓘ

## 5.1  Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: FCO, SF4, FE2, F42, NI, FAD

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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

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 | 385 | 0 | 0 | 21 | 0 |
| 2 | B | 215 | 0 | 0 | 23 | 0 |
| 3 | C | 278 | 0 | 0 | 49 | 0 |
| 4 | A | 1 | 0 | 0 | 0 | 0 |
| 5 | A | 1 | 0 | 0 | 0 | 0 |
| 6 | A | 1 | 0 | 0 | 0 | 0 |
| 7 | B | 24 | 0 | 0 | 3 | 0 |
| 7 | C | 8 | 0 | 0 | 3 | 0 |
| 8 | C | 19 | 0 | 8 | 8 | 0 |
| 9 | C | 53 | 0 | 30 | 12 | 0 |
| All | All | 985 | 0 | 38 | 97 | 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 95.

The worst 5 of 97 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---|---|---|---|
| 3:C:25:ASP:CA | 9:C:304:FAD:C5' | 1.78 | 1.58 |
| 3:C:104:CYS:CA | 7:C:303:SF4:S3 | 2.04 | 1.45 |
| 3:C:25:ASP:CA | 9:C:304:FAD:O4' | 1.71 | 1.38 |
| 3:C:25:ASP:CA | 9:C:304:FAD:C4' | 2.11 | 1.28 |
| 3:C:26:GLY:CA | 9:C:304:FAD:H51A | 1.85 | 1.06 |

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

There are no protein backbone outliers to report in this entry.

### 5.3.2 Protein sidechains ⓘ

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

### 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 ⓘ

Of 9 ligands modelled in this entry, 1 is modelled with single atom and 2 are monoatomic - leaving 6 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|-----|------|-------|-----|------|--------|------|---------|--------|------|---------|
| | | | | | Counts | RMSZ | #\|Z\| > 2 | Counts | RMSZ | #\|Z\| > 2 |
| 7 | SF4 | B | 300 | - | 0,12,12 | - | - | - | | |
| 9 | FAD | C | 304 | - | 53,58,58 | 1.28 | 5 (9%) | 68,89,89 | 1.35 | 10 (14%) |
| 8 | F42 | C | 285 | - | 18,21,55 | 1.20 | 1 (5%) | 20,31,79 | 2.91 | 6 (30%) |
| 7 | SF4 | B | 301 | - | 0,12,12 | - | - | - | | |
| 7 | SF4 | B | 302 | - | 0,12,12 | - | - | - | | |
| 7 | SF4 | C | 303 | - | 0,12,12 | - | - | - | | |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 7 | SF4 | B | 300 | - | - | - | 0/6/5/5 |
| 9 | FAD | C | 304 | - | - | 20/30/50/50 | 0/6/6/6 |
| 8 | F42 | C | 285 | - | - | 0/2/2/53 | 0/3/3/3 |
| 7 | SF4 | B | 301 | - | - | - | 0/6/5/5 |
| 7 | SF4 | B | 302 | - | - | - | 0/6/5/5 |
| 7 | SF4 | C | 303 | - | - | - | 0/6/5/5 |

The worst 5 of 6 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|--------|------|-------------|----------|
| 9 | C | 304 | FAD | C9A-C5X | 5.05 | 1.49 | 1.41 |
| 9 | C | 304 | FAD | C8-C7 | 3.29 | 1.49 | 1.40 |
| 9 | C | 304 | FAD | C4X-N5 | 2.50 | 1.35 | 1.30 |
| 9 | C | 304 | FAD | C5A-C4A | 2.48 | 1.47 | 1.40 |
| 9 | C | 304 | FAD | C4-N3 | -2.45 | 1.34 | 1.38 |

The worst 5 of 16 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|--------|------|----------------|-------------|
| 8 | C | 285 | F42 | C4A-C4-N3 | -9.73 | 117.62 | 124.40 |
| 8 | C | 285 | F42 | C2-N3-C4 | 5.45 | 119.74 | 115.14 |
| 8 | C | 285 | F42 | C4A-C5-C5A | -4.01 | 116.31 | 122.27 |
| 9 | C | 304 | FAD | P-O3P-PA | -3.60 | 120.46 | 132.83 |
| 9 | C | 304 | FAD | C3B-C2B-C1B | 3.51 | 106.26 | 100.98 |

There are no chirality outliers.
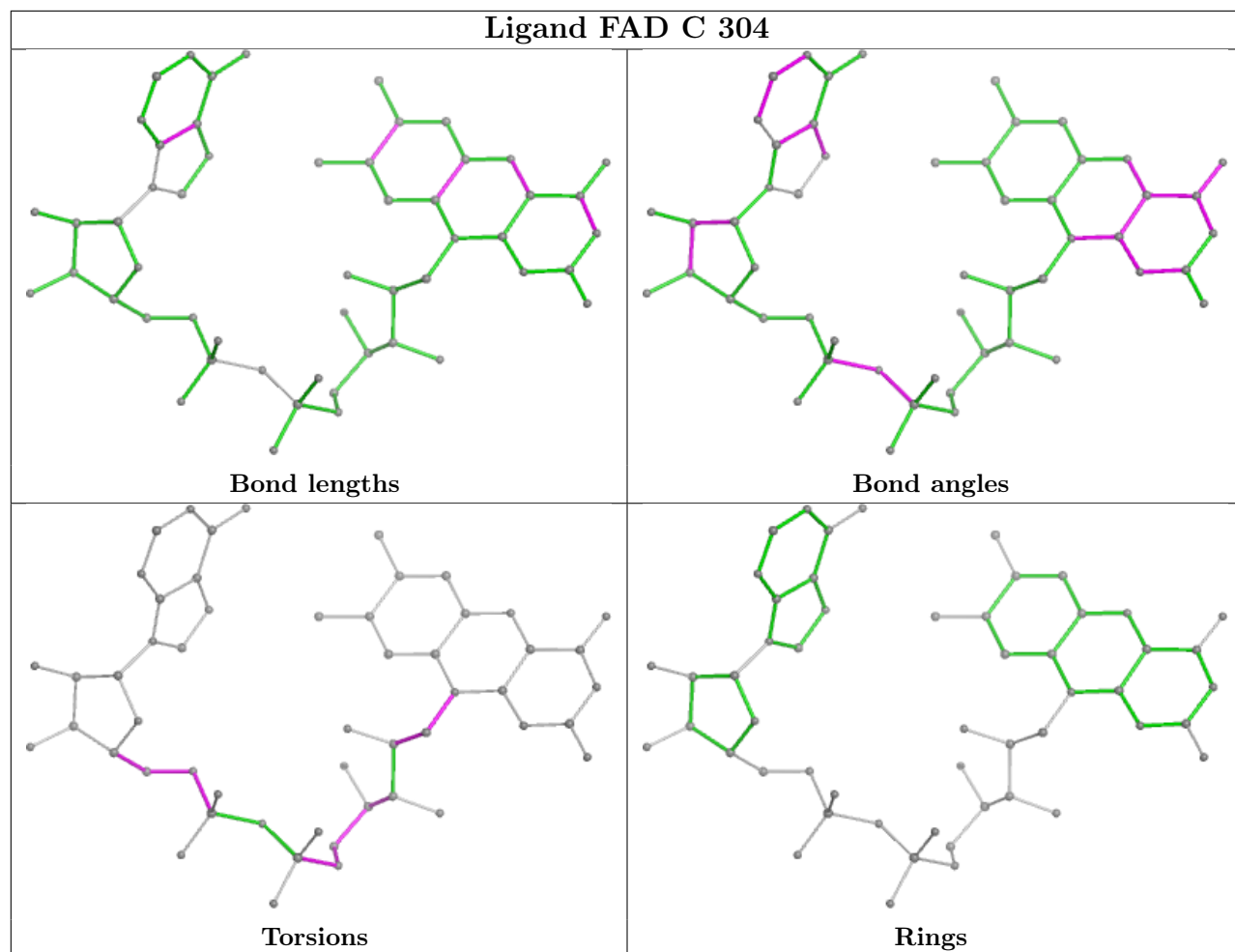
5 of 20 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 9 | C | 304 | FAD | C5B-O5B-PA-O3P |
| 9 | C | 304 | FAD | C3B-C4B-C5B-O5B |
| 9 | C | 304 | FAD | C2'-C1'-N10-C9A |
| 9 | C | 304 | FAD | C2'-C1'-N10-C10 |
| 9 | C | 304 | FAD | N10-C1'-C2'-O2' |

There are no ring outliers.

5 monomers are involved in 26 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 7 | B | 300 | SF4 | 2 | 0 |
| 9 | C | 304 | FAD | 12 | 0 |
| 8 | C | 285 | F42 | 8 | 0 |
| 7 | B | 302 | SF4 | 1 | 0 |
| 7 | C | 303 | SF4 | 3 | 0 |

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

### Ligand FAD C 304



Bond lengths

Bond angles

Torsions

Rings

**Ligand F42 C 285**



Bond lengths



Bond angles



Torsions



Rings
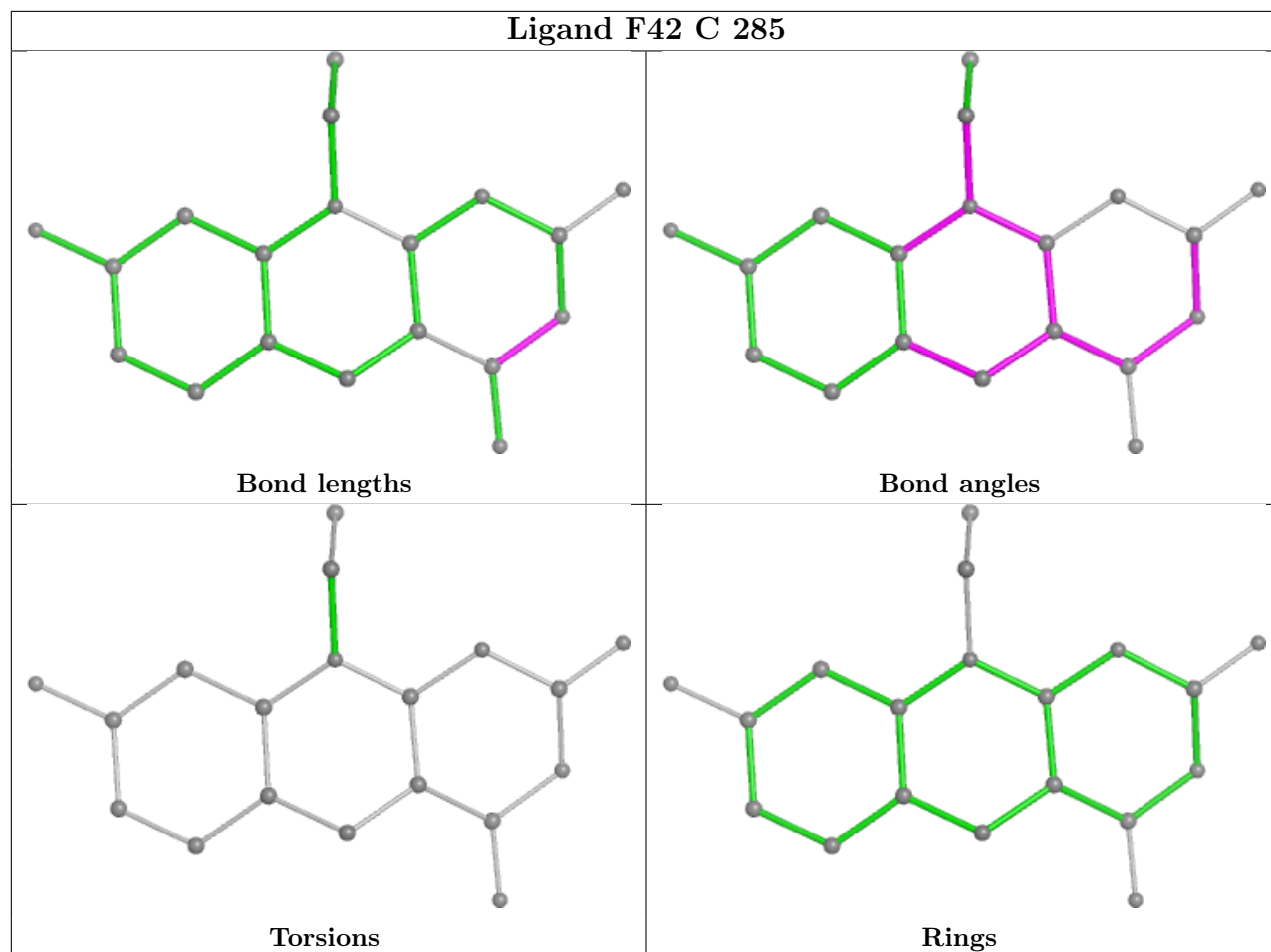
## 5.7 Other polymers ⓘ

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-2097. 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 |
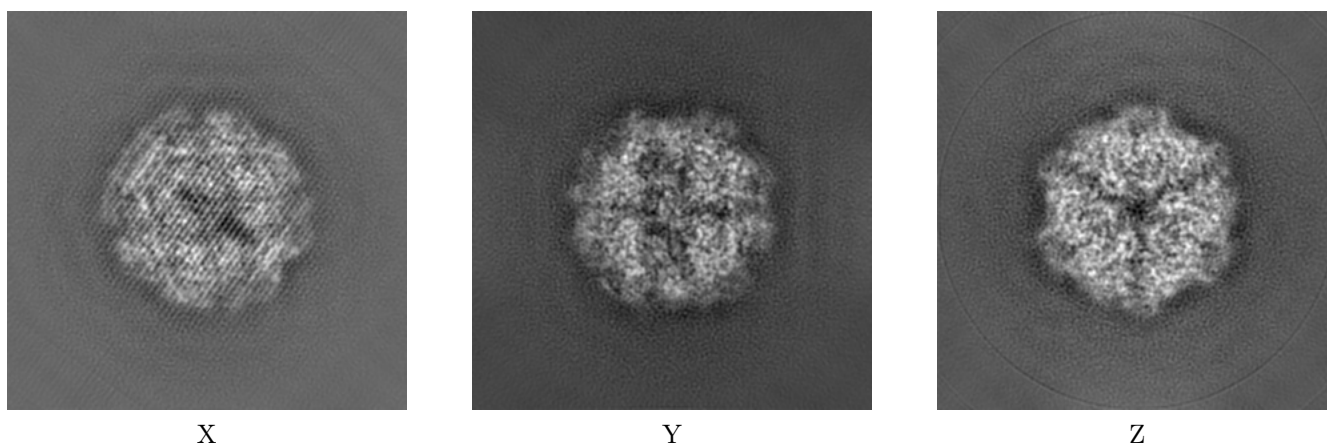
The images above show the map projected in three orthogonal directions.

## 6.2    Central slices  ⓘ

### 6.2.1    Primary map



| X Index: 144 | Y Index: 144 | Z Index: 144 |
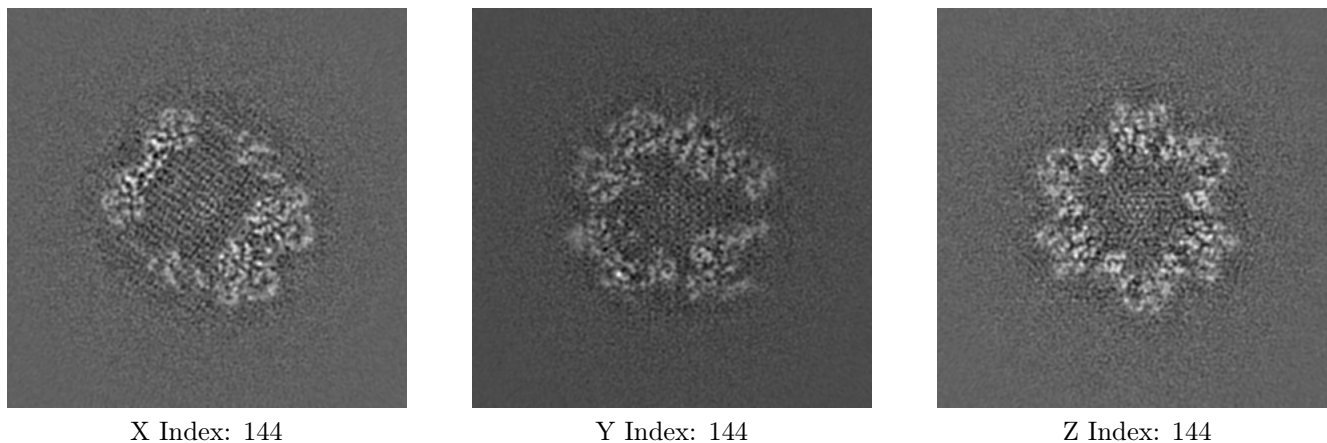
The images above show central slices of the map in three orthogonal directions.

## 6.3　Largest variance slices　ⓘ
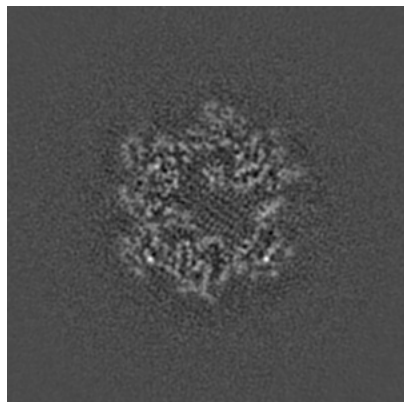
### 6.3.1　Primary map
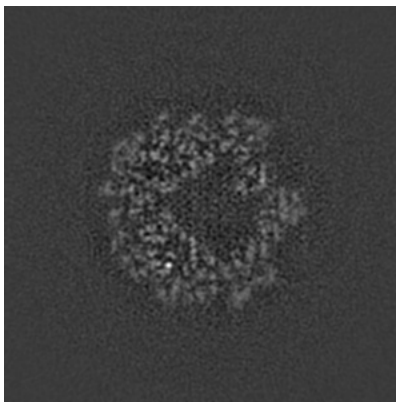


| X Index: 168 | Y Index: 128 | Z Index: 168 |
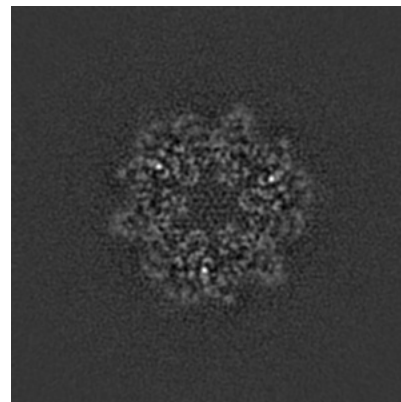
The images above show the largest variance slices of the map in three orthogonal directions.

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

### 6.4.1　Primary map



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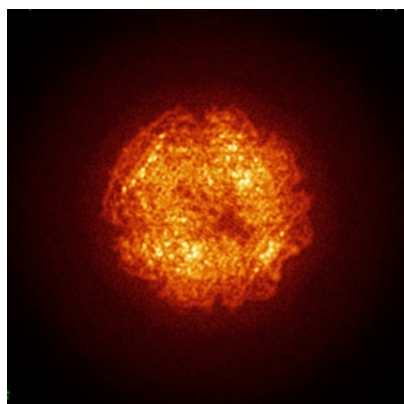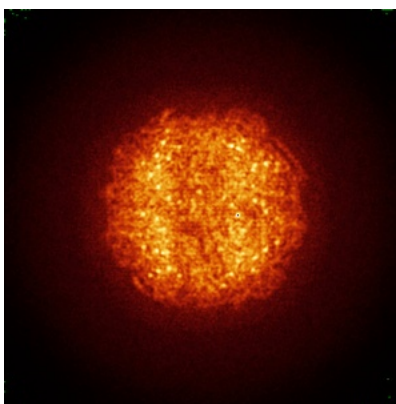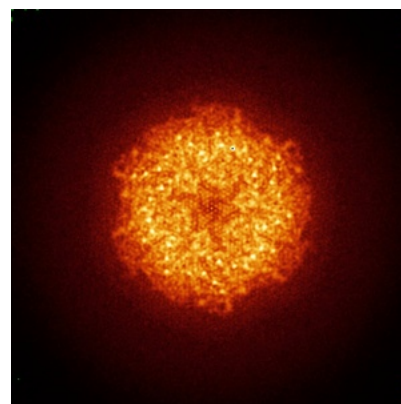
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

## 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 212.0. 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 ⓘ
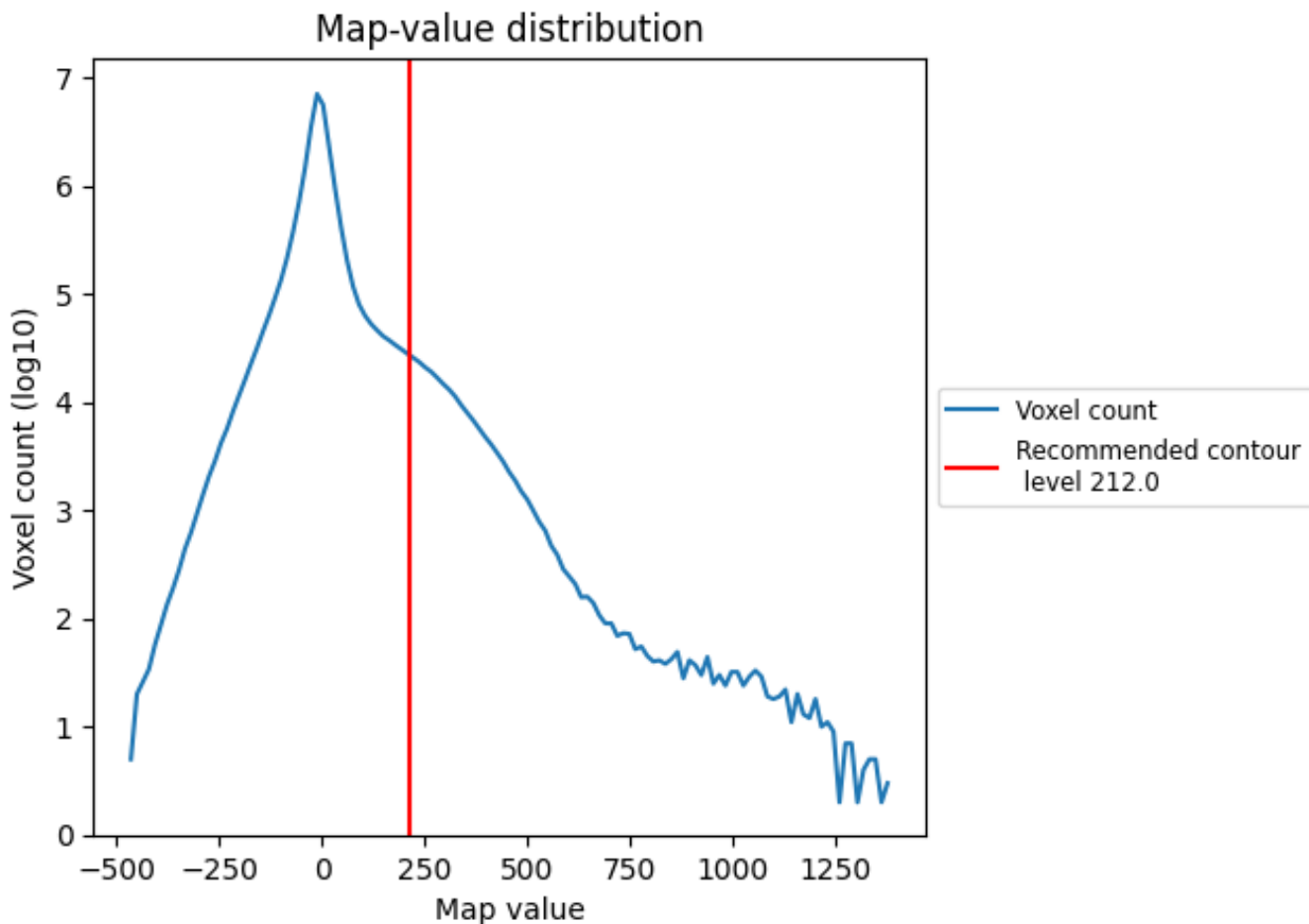


The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

## 7.2 Volume estimate ⓘ
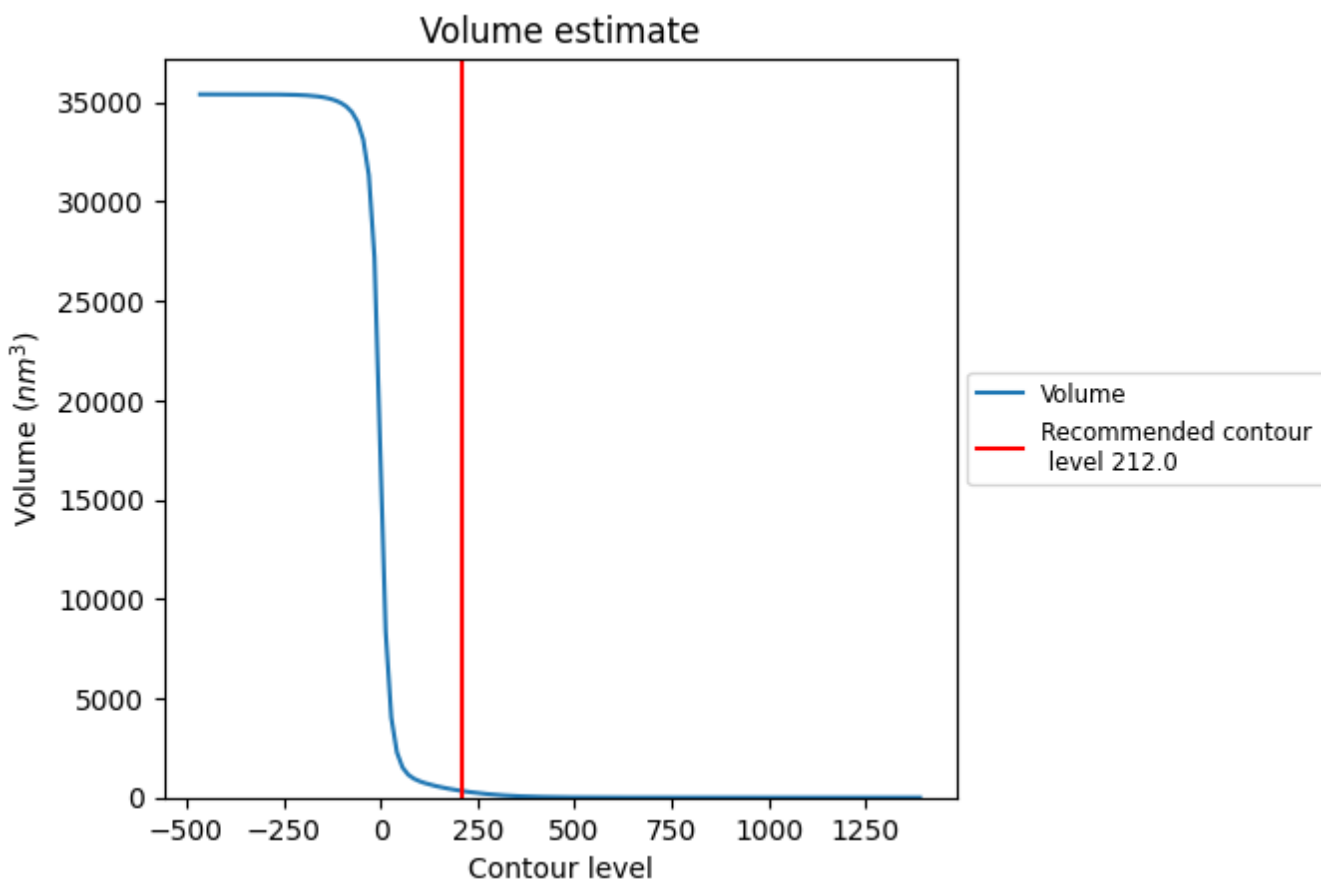


The volume at the recommended contour level is 328 nm$^3$; this corresponds to an approximate mass of 297 kDa.

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   ⓘ
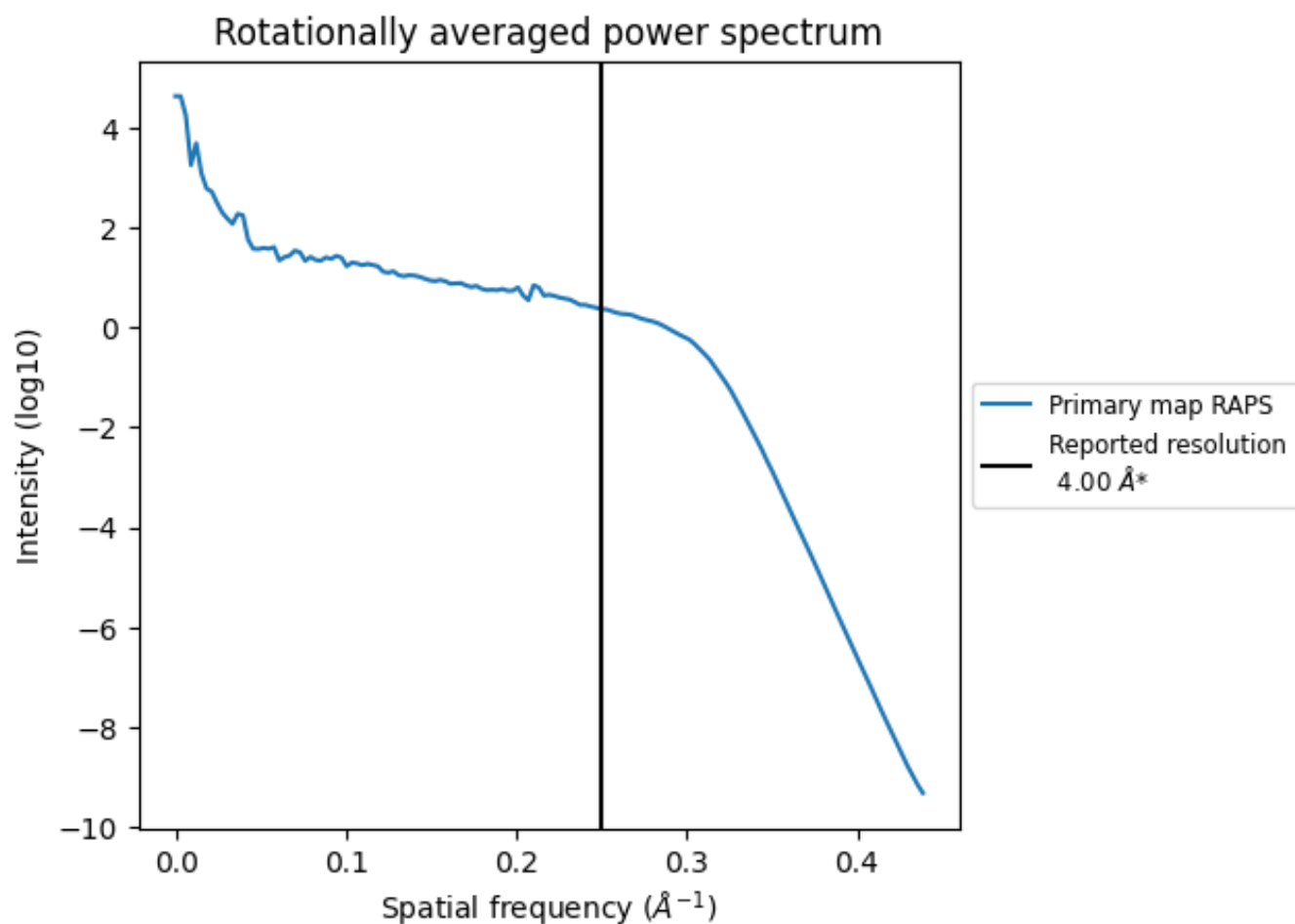


*Reported resolution corresponds to spatial frequency of 0.250 $\text{Å}^{-1}$

# 8 Fourier-Shell correlation ⓘ

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

# 9   Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-2097 and PDB model 3ZFS. Per-residue inclusion information can be found in section 3 on page 6.
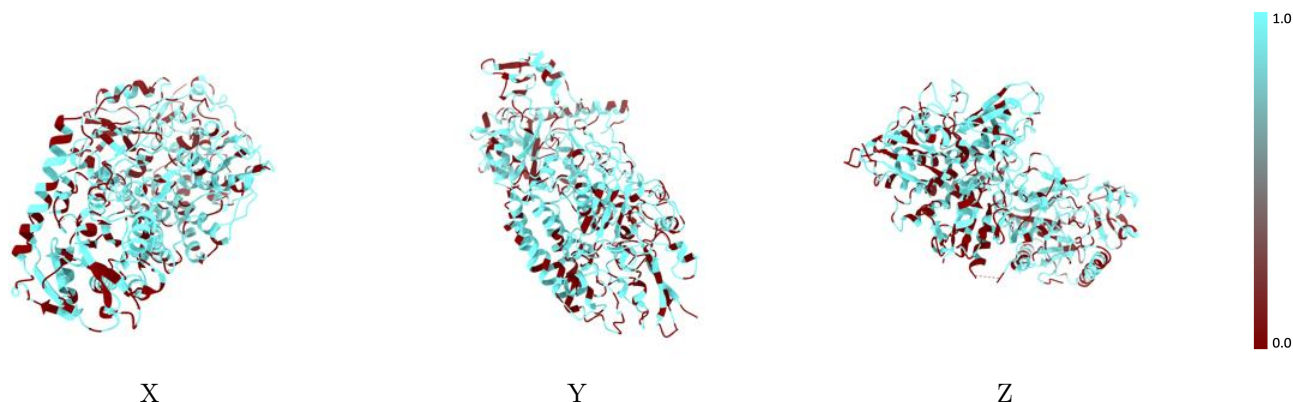
## 9.1   Map-model overlay ⓘ



<center>X                        Y                        Z</center>
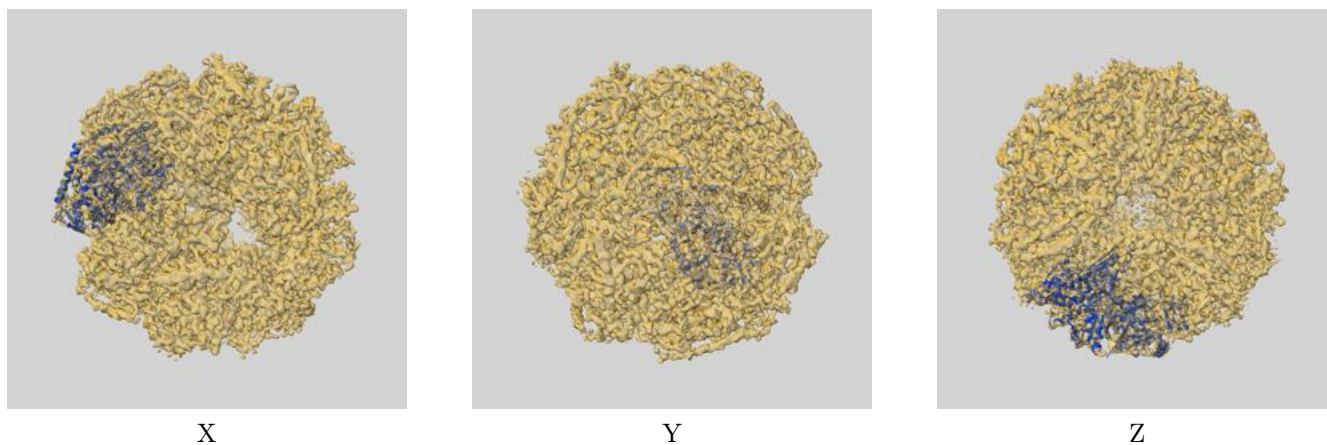
The images above show the 3D surface view of the map at the recommended contour level 212.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2    Q-score mapped to coordinate model ⓘ



X                  Y                  Z
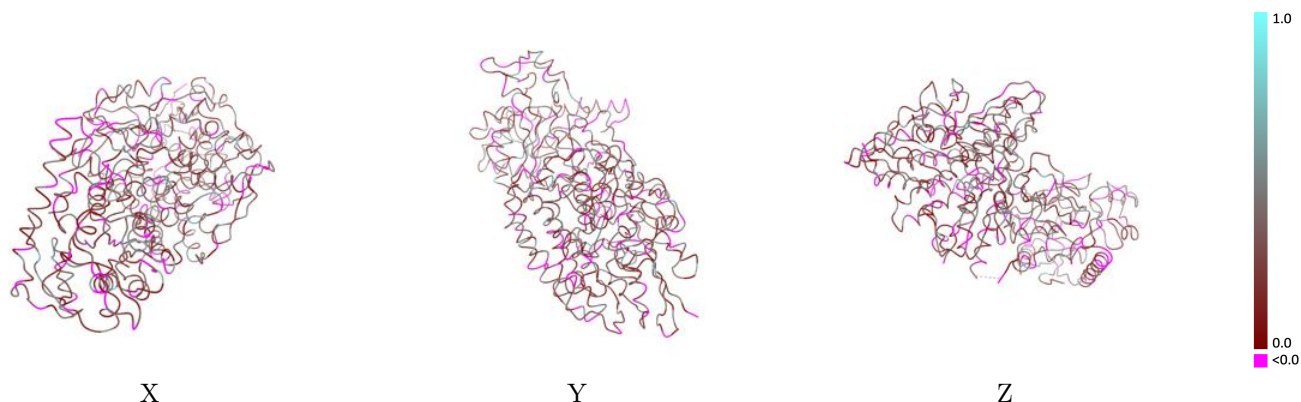
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (212.0).

## 9.4    Atom inclusion  ⓘ
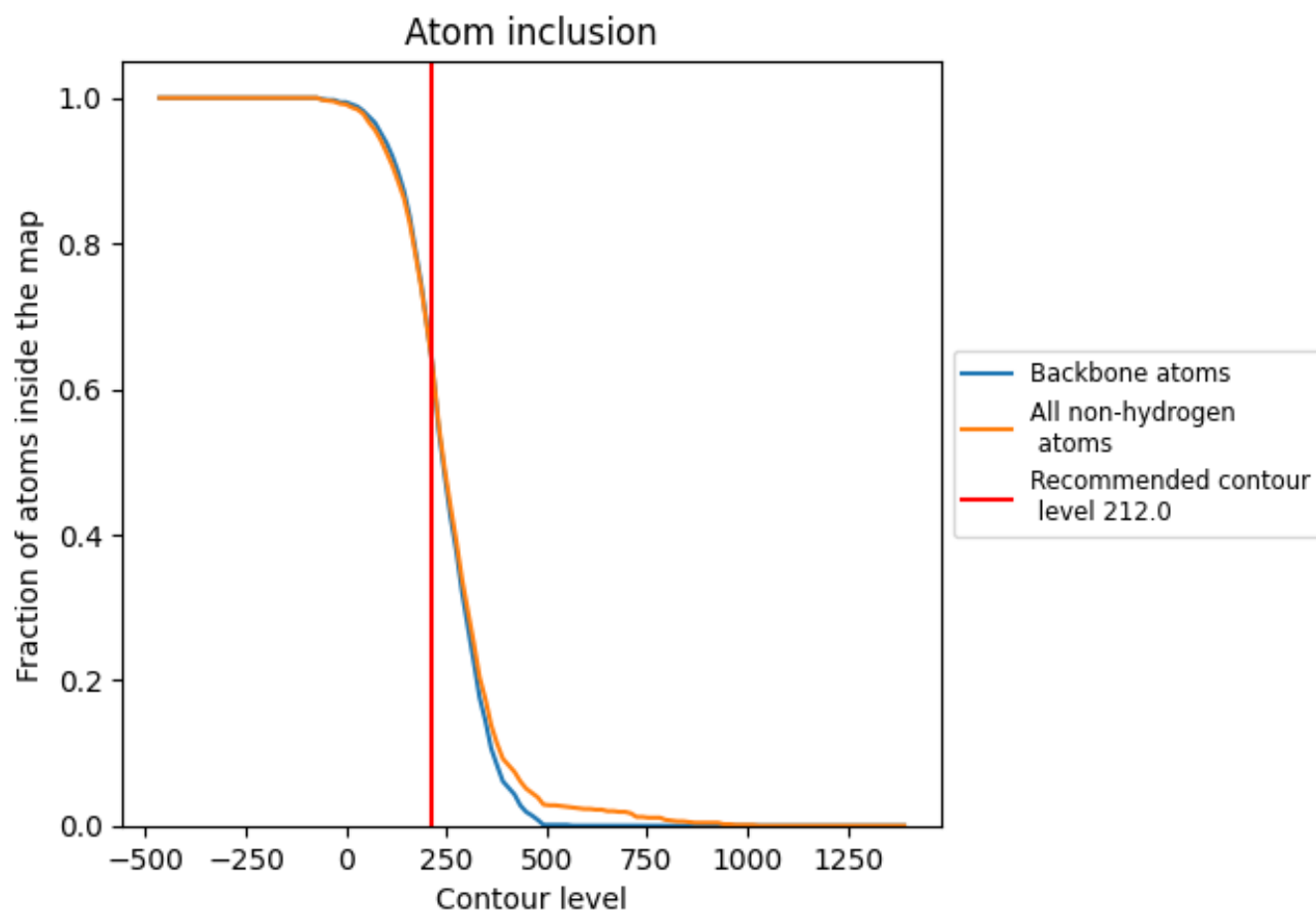


At the recommended contour level, 65% of all backbone atoms, 64% of all non-hydrogen atoms, are inside the map.

## 9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (212.0) and Q-score for the entire model and for each chain.

| Chain | Atom inclusion | Q-score |
|:-----:|:--------------:|:-------:|
| All | 0.6450 | 0.2490 |
| A | 0.6780 | 0.2440 |
| B | 0.6530 | 0.2530 |
| C | 0.6030 | 0.2520 |

1.0

0.0

<0.0

W O R L D W I D E

PDB

PROTEIN DATA BANK