



wwPDB EM Validation Summary Report ⓘ

Mar 14, 2023 – 04:05 PM EDT

PDB ID : 8CZI
EMDB ID : EMD-27098
Title : Cryo-EM structure of the SARS-CoV-2 HR1HR2 fusion core complex with extended HR2
Authors : Yang, K.; Brunger, A.T.
Deposited on : 2022-05-24
Resolution : 2.22 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.dev43
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) : 2.32.1

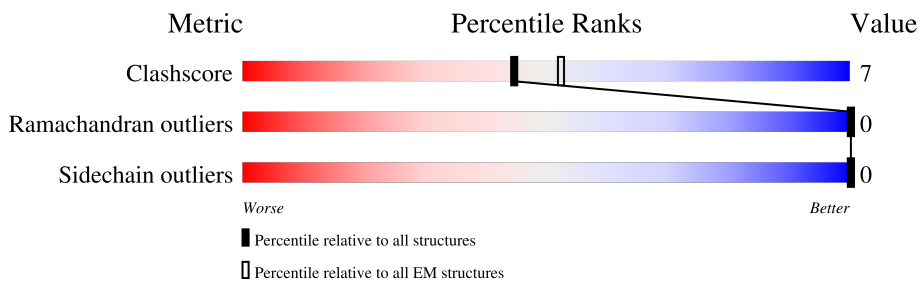
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.22 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



| Metric | Whole archive (#Entries) | EM structures (#Entries) |
|-----------------------|--------------------------|--------------------------|
| Clashscore | 158937 | 4297 |
| Ramachandran outliers | 154571 | 4023 |
| Sidechain outliers | 154315 | 3826 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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 | 257 | |
| 1 | B | 257 | |
| 1 | C | 257 | |
| 2 | D | 45 | |
| 2 | E | 45 | |
| 2 | F | 45 | |

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2550 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 Scaffolded Spike protein S2' HR1.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|----|-----|---------|-------|
| | | | Total | C | N | O | | |
| 1 | A | 71 | 531 | 327 | 94 | 110 | 0 | 0 |
| 1 | B | 71 | 531 | 327 | 94 | 110 | 0 | 0 |
| 1 | C | 71 | 531 | 327 | 94 | 110 | 0 | 0 |

There are 33 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|-----------------------|------------|
| A | 732 | MET | - | initiating methionine | UNP B2J981 |
| A | 733 | SER | - | expression tag | UNP B2J981 |
| A | 734 | HIS | - | expression tag | UNP B2J981 |
| A | 735 | HIS | - | expression tag | UNP B2J981 |
| A | 736 | HIS | - | expression tag | UNP B2J981 |
| A | 737 | HIS | - | expression tag | UNP B2J981 |
| A | 738 | HIS | - | expression tag | UNP B2J981 |
| A | 739 | HIS | - | expression tag | UNP B2J981 |
| A | 740 | GLY | - | expression tag | UNP B2J981 |
| A | 741 | SER | - | expression tag | UNP B2J981 |
| A | 916 | ALA | - | linker | UNP B2J981 |
| B | 732 | MET | - | initiating methionine | UNP B2J981 |
| B | 733 | SER | - | expression tag | UNP B2J981 |
| B | 734 | HIS | - | expression tag | UNP B2J981 |
| B | 735 | HIS | - | expression tag | UNP B2J981 |
| B | 736 | HIS | - | expression tag | UNP B2J981 |
| B | 737 | HIS | - | expression tag | UNP B2J981 |
| B | 738 | HIS | - | expression tag | UNP B2J981 |
| B | 739 | HIS | - | expression tag | UNP B2J981 |
| B | 740 | GLY | - | expression tag | UNP B2J981 |
| B | 741 | SER | - | expression tag | UNP B2J981 |
| B | 916 | ALA | - | linker | UNP B2J981 |
| C | 732 | MET | - | initiating methionine | UNP B2J981 |
| C | 733 | SER | - | expression tag | UNP B2J981 |

Continued on next page...

Continued from previous page...

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| C | 734 | HIS | - | expression tag | UNP B2J981 |
| C | 735 | HIS | - | expression tag | UNP B2J981 |
| C | 736 | HIS | - | expression tag | UNP B2J981 |
| C | 737 | HIS | - | expression tag | UNP B2J981 |
| C | 738 | HIS | - | expression tag | UNP B2J981 |
| C | 739 | HIS | - | expression tag | UNP B2J981 |
| C | 740 | GLY | - | expression tag | UNP B2J981 |
| C | 741 | SER | - | expression tag | UNP B2J981 |
| C | 916 | ALA | - | linker | UNP B2J981 |

- Molecule 2 is a protein called Spike protein S2' HR2.

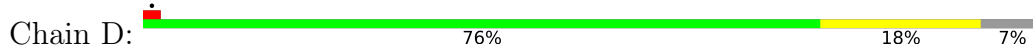
| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|----|----|---------|-------|
| | | | Total | C | N | O | | |
| 2 | D | 42 | 319 | 195 | 55 | 69 | 0 | 0 |
| 2 | E | 42 | 319 | 195 | 55 | 69 | 0 | 0 |
| 2 | F | 42 | 319 | 195 | 55 | 69 | 0 | 0 |

GLU GLY
GLN GLY
SER MET
GLU VAL
PHE GLU
TYR TYR
SER SER
LEU LEU
HIS HIS
GLU GLU
PHE PHE
PHE PHE
ASN ASN
GLU GLU
TYR TYR
ASN ASN
GLN GLN
VAL VAL
ILE ILE
LEU LEU
GLY GLY
ASP ASP
GLY GLY
THR THR
ARG ARG
TYR TYR
VAL VAL
LEU LEU
PRO PRO
VAL VAL
ALA ALA
THR THR
LEU LEU
PHE PHE
SER SER
LYS LYS
LEU LEU
ALA ALA
GLU GLU
THR THR
CYS CYS
HIS HIS
PHE PHE
GLN GLN
SER SER
GLU GLU
VAL VAL
TYR TYR
SER SER
LEU LEU

ARG GLN
GLY MET
VAL VAL
GLU PHE
ASN ASN
ASP ASP
LEU LEU
ALA ALA
ALA ALA
GLN GLN
ALA ALA
ILE ILE
SER SER
GLY TYR
VAL VAL
ASN ASN
GLN GLN
VAL VAL
ARG ARG
GLN GLN
ALA ALA
HIS HIS
ILE ILE
ALA ALA
GLN GLN
HIS HIS
GLU GLU
ILE ILE
SER SER
LEU LEU
GLY GLY
ASP ASP
ARG ARG
GLY GLY
THR THR
ARG ARG
TYR TYR
LEU LEU
VAL VAL
TYR TYR
PRO PRO
VAL VAL
LYS LYS
ILE ILE
THR THR
LEU LEU
PHE PHE
LEU LEU
SER SER
LYS LYS
THR THR
GLU GLU
ALA ALA
ARG ARG
LEU LEU
ALA ALA
TYR TYR
CYS CYS
HIS HIS
SER SER
LEU LEU
HIS HIS
PHE PHE
SER SER
PHE PHE
ALA ALA
LYS LYS
VAL VAL
ASP ASP
SER SER
SER SER
LEU LEU

THR LEU
GLY PHE
ALA ALA
TYR TYR
E918
I923
K933
I934
Q935
L938
Q949
L962
L966
N978
L984
E988

- Molecule 2: Spike protein S2' HR2



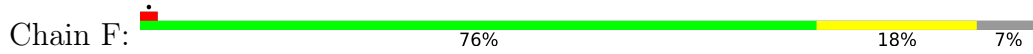
LYS ASN
H1159
D1163
M1173
A1174
V1177
L1186
V1189
A1190
I1198
D1199
L1200
GLN

- Molecule 2: Spike protein S2' HR2



LYS ASN
H1159
D1163
M1173
A1174
V1177
L1186
V1189
A1190
I1198
D1199
L1200
GLN

- Molecule 2: Spike protein S2' HR2



LYS ASN
H1159
D1163
M1173
A1174
V1177
L1186
V1189
A1190
I1198
D1199
L1200
GLN

4 Experimental information

| Property | Value | Source |
|--------------------------------------|---------------------------|-----------|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 751443 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | NONE | Depositor |
| Microscope | FEI TITAN KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose ($e^-/\text{\AA}^2$) | 47 | Depositor |
| Minimum defocus (nm) | 300 | Depositor |
| Maximum defocus (nm) | 2000 | Depositor |
| Magnification | Not provided | |
| Image detector | GATAN K3 (6k x 4k) | Depositor |
| Maximum map value | 1.583 | Depositor |
| Minimum map value | -0.002 | Depositor |
| Average map value | 0.001 | Depositor |
| Map value standard deviation | 0.019 | Depositor |
| Recommended contour level | 0.05 | Depositor |
| Map size (\AA) | 208.96, 208.96, 208.96 | wwPDB |
| Map dimensions | 256, 256, 256 | wwPDB |
| Map angles ($^\circ$) | 90.0, 90.0, 90.0 | wwPDB |
| Pixel spacing (\AA) | 0.81625, 0.81625, 0.81625 | Depositor |

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|-----|-------|--------------|---------|-------------|---------|
| | | RMSZ | # Z >5 | RMSZ | # Z >5 |
| 1 | A | 0.26 | 0/532 | 0.36 | 0/717 |
| 1 | B | 0.26 | 0/532 | 0.36 | 0/717 |
| 1 | C | 0.26 | 0/532 | 0.36 | 0/717 |
| 2 | D | 0.26 | 0/320 | 0.43 | 0/434 |
| 2 | E | 0.26 | 0/320 | 0.43 | 0/434 |
| 2 | F | 0.26 | 0/320 | 0.43 | 0/434 |
| All | All | 0.26 | 0/2556 | 0.39 | 0/3453 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | A | 531 | 0 | 545 | 14 | 0 |
| 1 | B | 531 | 0 | 545 | 14 | 0 |
| 1 | C | 531 | 0 | 545 | 13 | 0 |
| 2 | D | 319 | 0 | 317 | 7 | 0 |
| 2 | E | 319 | 0 | 317 | 7 | 0 |
| 2 | F | 319 | 0 | 317 | 7 | 0 |
| All | All | 2550 | 0 | 2586 | 35 | 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 7.

The worst 5 of 35 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:923:ILE:HG12 | 2:F:1198:ILE:HD13 | 1.85 | 0.58 |
| 1:C:923:ILE:HG12 | 2:E:1198:ILE:HD13 | 1.86 | 0.58 |
| 1:B:923:ILE:HG12 | 2:D:1198:ILE:HD13 | 1.84 | 0.57 |
| 1:C:978:ASN:ND2 | 2:F:1163:ASP:OD1 | 2.38 | 0.56 |
| 1:A:978:ASN:ND2 | 2:D:1163:ASP:OD1 | 2.38 | 0.56 |

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 | 69/257 (27%) | 69 (100%) | 0 | 0 | 100 | 100 |
| 1 | B | 69/257 (27%) | 69 (100%) | 0 | 0 | 100 | 100 |
| 1 | C | 69/257 (27%) | 69 (100%) | 0 | 0 | 100 | 100 |
| 2 | D | 40/45 (89%) | 40 (100%) | 0 | 0 | 100 | 100 |
| 2 | E | 40/45 (89%) | 40 (100%) | 0 | 0 | 100 | 100 |
| 2 | F | 40/45 (89%) | 40 (100%) | 0 | 0 | 100 | 100 |
| All | All | 327/906 (36%) | 327 (100%) | 0 | 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 | 61/220 (28%) | 61 (100%) | 0 | 100 | 100 |
| 1 | B | 61/220 (28%) | 61 (100%) | 0 | 100 | 100 |
| 1 | C | 61/220 (28%) | 61 (100%) | 0 | 100 | 100 |
| 2 | D | 38/41 (93%) | 38 (100%) | 0 | 100 | 100 |
| 2 | E | 38/41 (93%) | 38 (100%) | 0 | 100 | 100 |
| 2 | F | 38/41 (93%) | 38 (100%) | 0 | 100 | 100 |
| All | All | 297/783 (38%) | 297 (100%) | 0 | 100 | 100 |

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

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-27098. These allow visual inspection of the internal detail of the map and identification of artifacts.

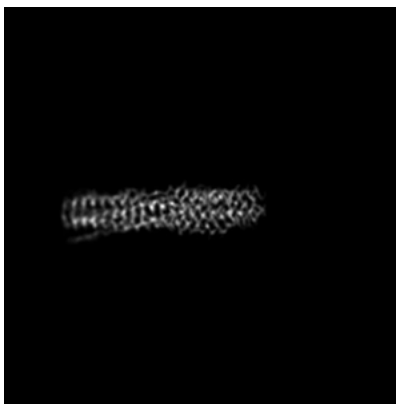
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

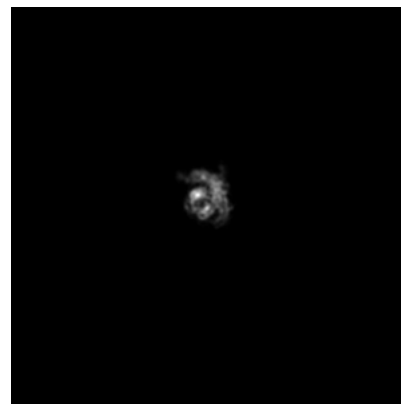
6.1.1 Primary map



X

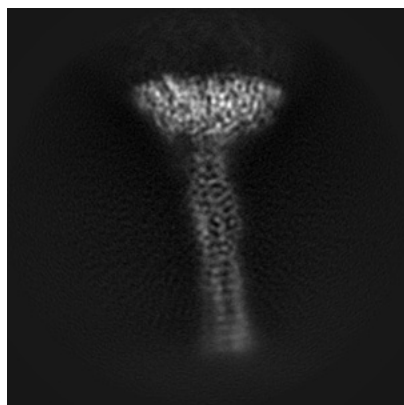


Y

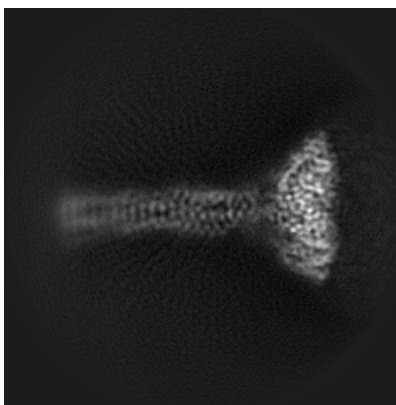


Z

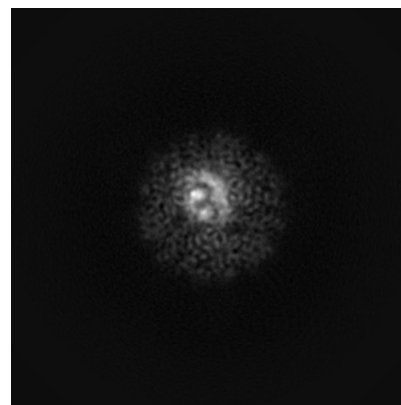
6.1.2 Raw map



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: 128

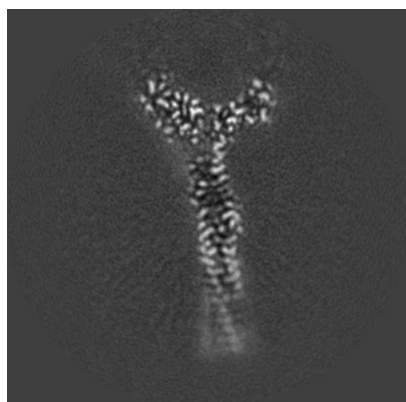


Y Index: 128

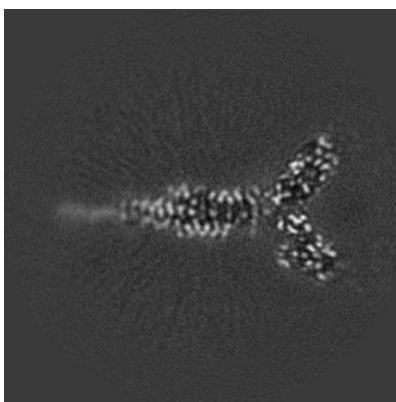


Z Index: 128

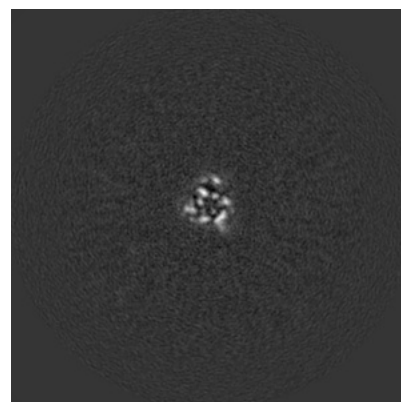
6.2.2 Raw map



X Index: 128



Y Index: 128

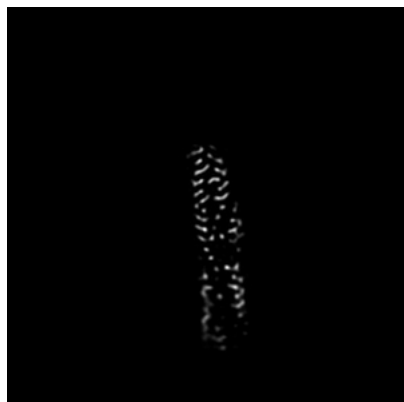


Z Index: 128

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: 124



Y Index: 135

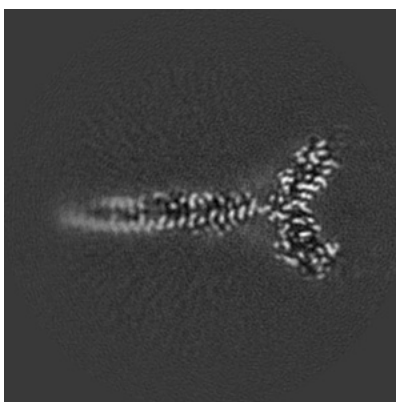


Z Index: 88

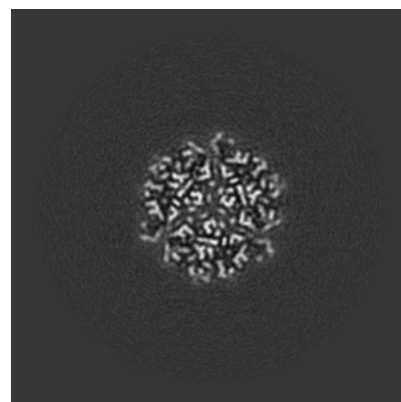
6.3.2 Raw map



X Index: 124



Y Index: 134

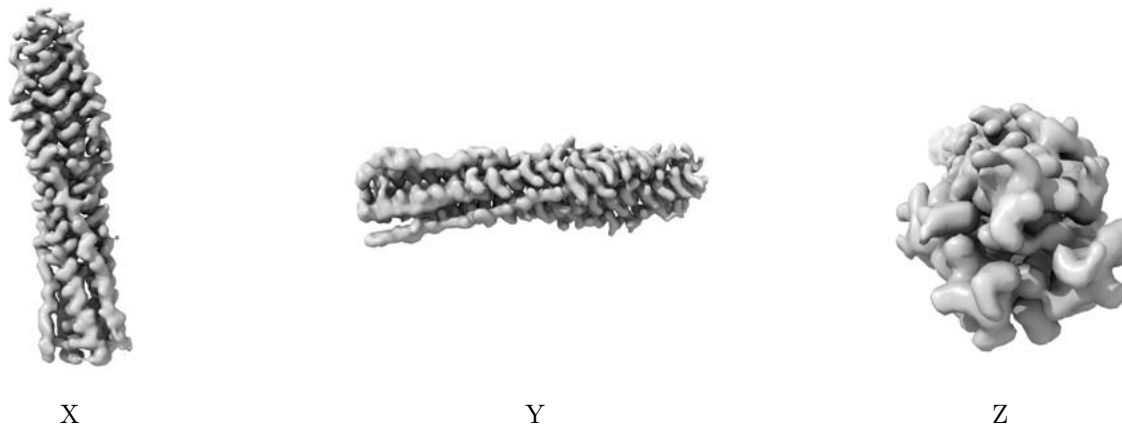


Z Index: 195

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

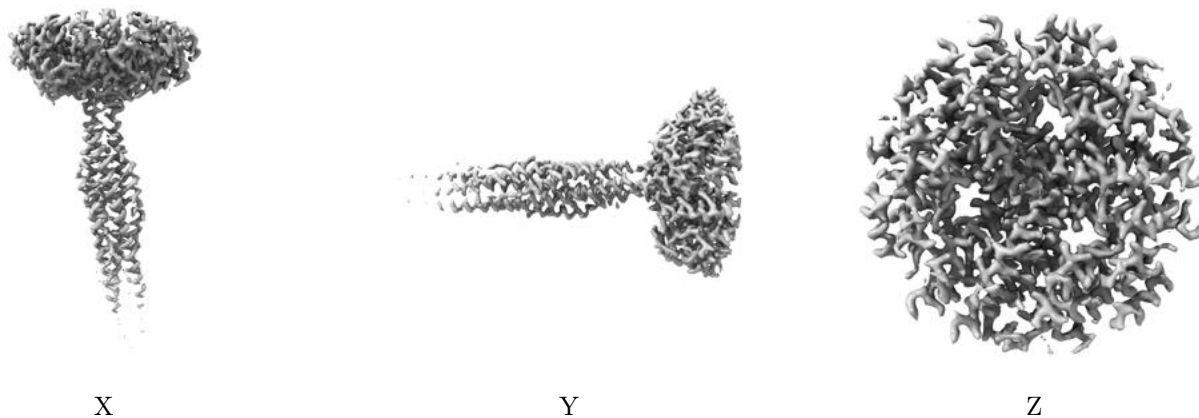
6.4 Orthogonal surface views [i](#)

6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.05. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.4.2 Raw map



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.

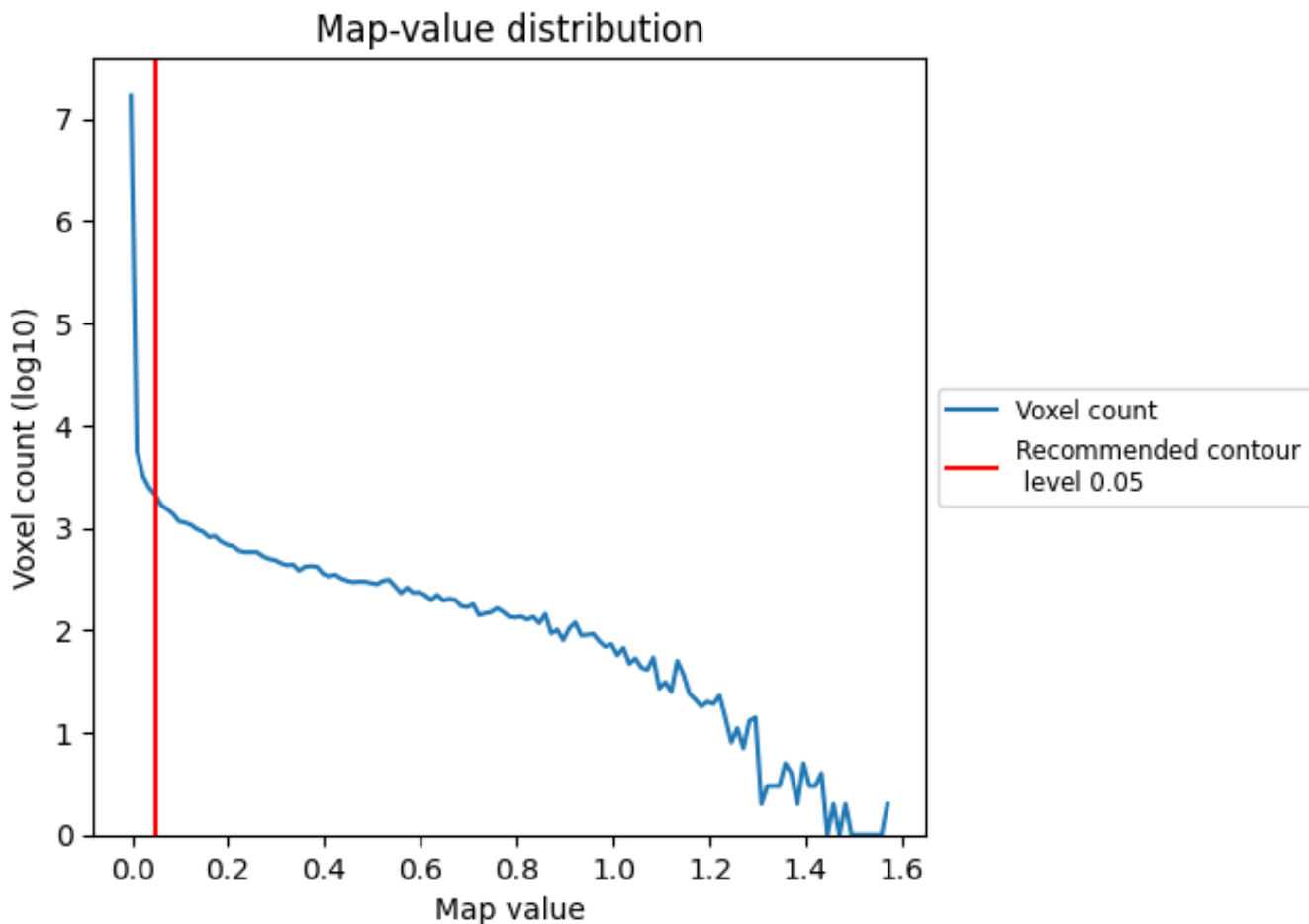
6.5 Mask visualisation [i](#)

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

7 Map analysis [i](#)

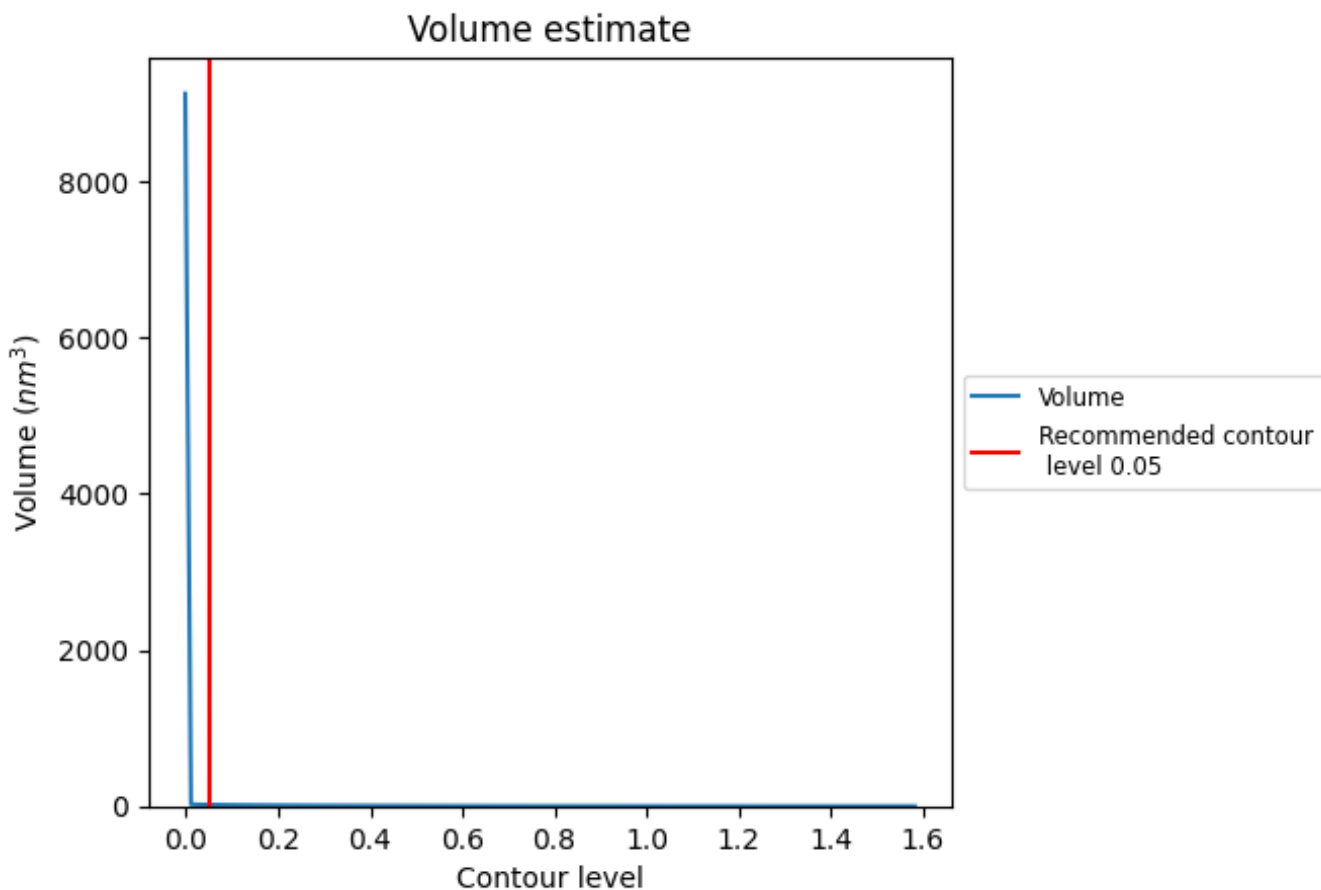
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



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

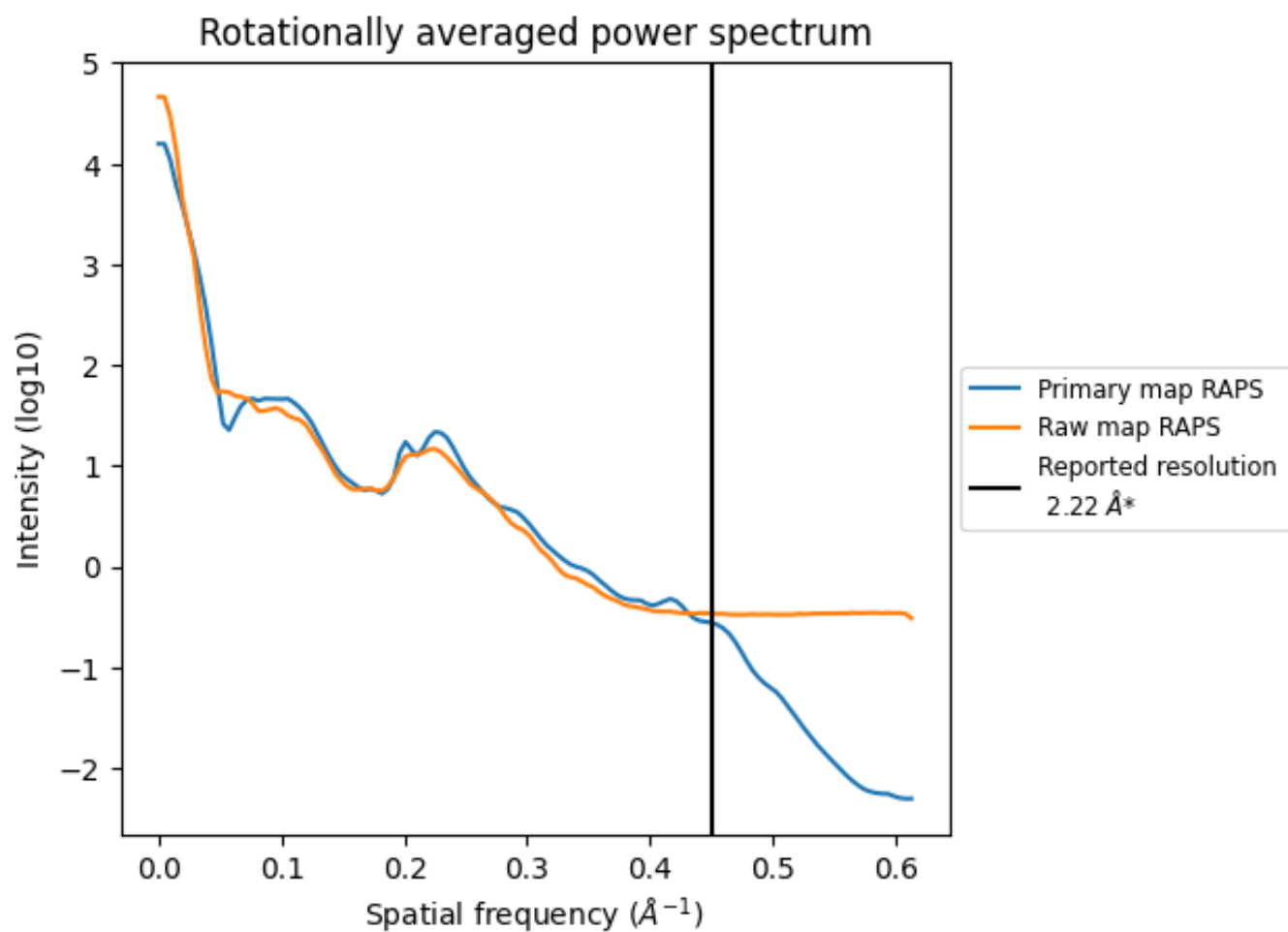
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 18 nm³; this corresponds to an approximate mass of 16 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

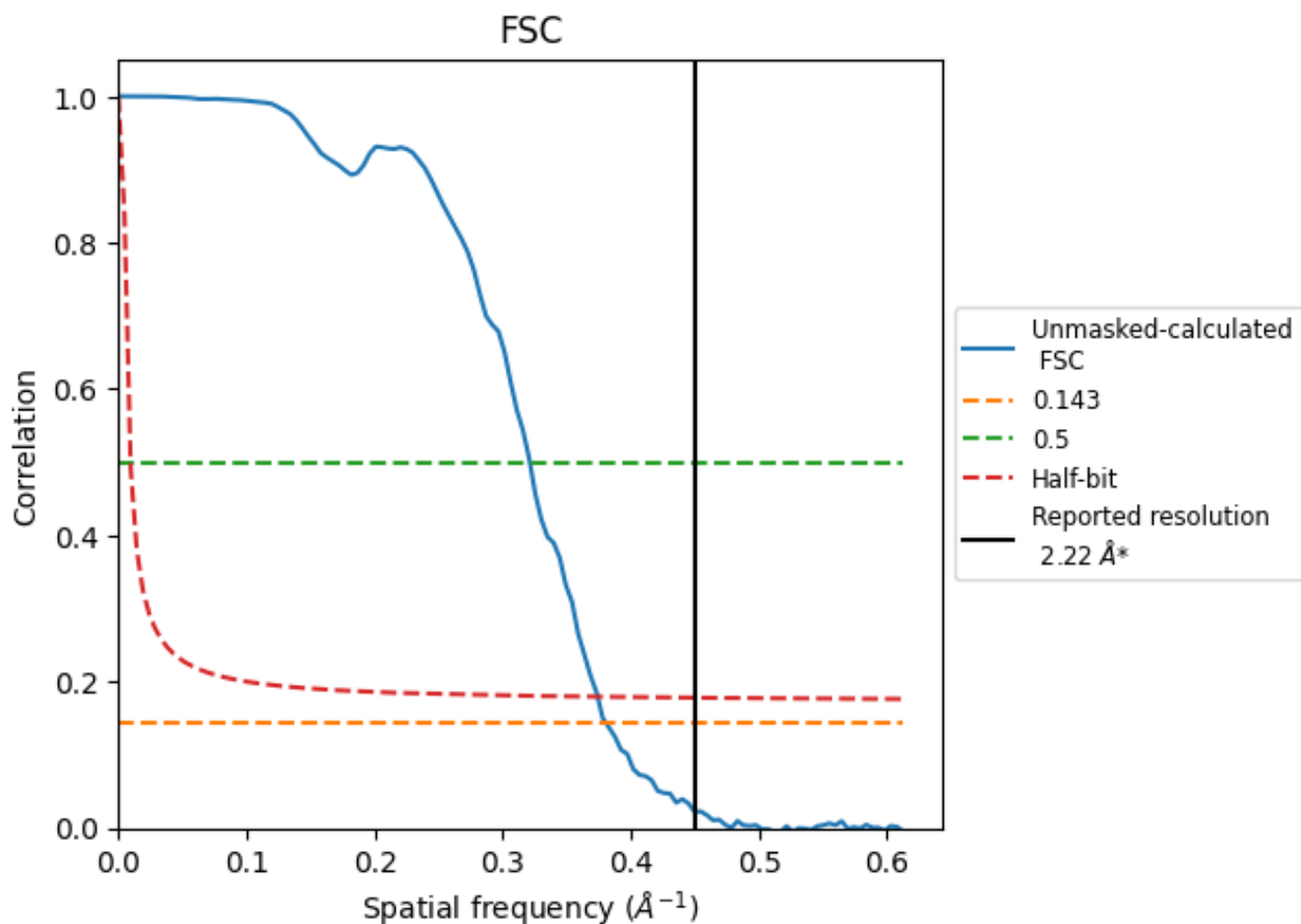


*Reported resolution corresponds to spatial frequency of 0.450 Å⁻¹

8 Fourier-Shell correlation [i](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC [i](#)



*Reported resolution corresponds to spatial frequency of 0.450 Å⁻¹

8.2 Resolution estimates [i](#)

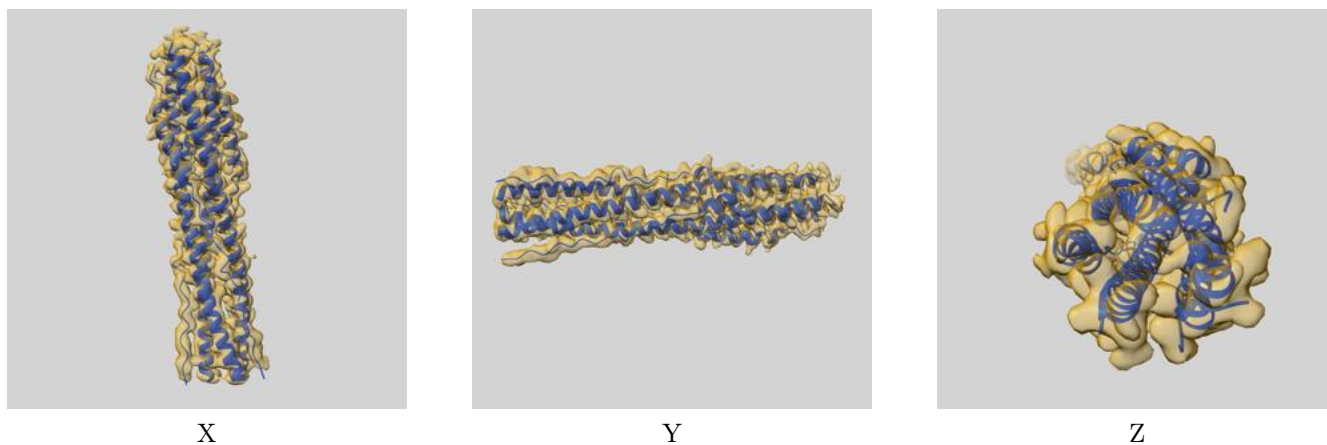
| Resolution estimate (Å) | Estimation criterion (FSC cut-off) | | |
|---------------------------|------------------------------------|------|----------|
| | 0.143 | 0.5 | Half-bit |
| Reported by author | 2.22 | - | - |
| Author-provided FSC curve | - | - | - |
| Unmasked-calculated* | 2.63 | 3.11 | 2.67 |

*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 2.63 differs from the reported value 2.22 by more than 10 %

9 Map-model fit [i](#)

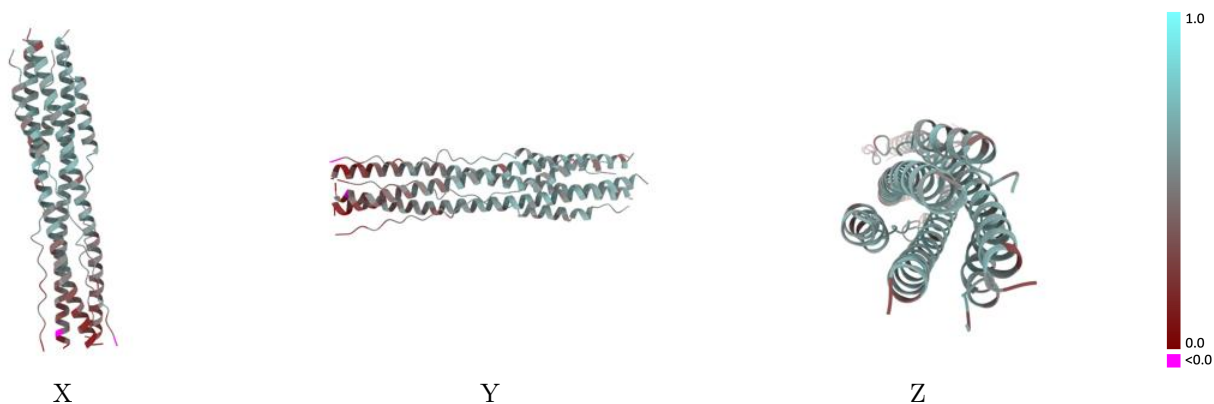
This section contains information regarding the fit between EMDB map EMD-27098 and PDB model 8CZI. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlay [i](#)



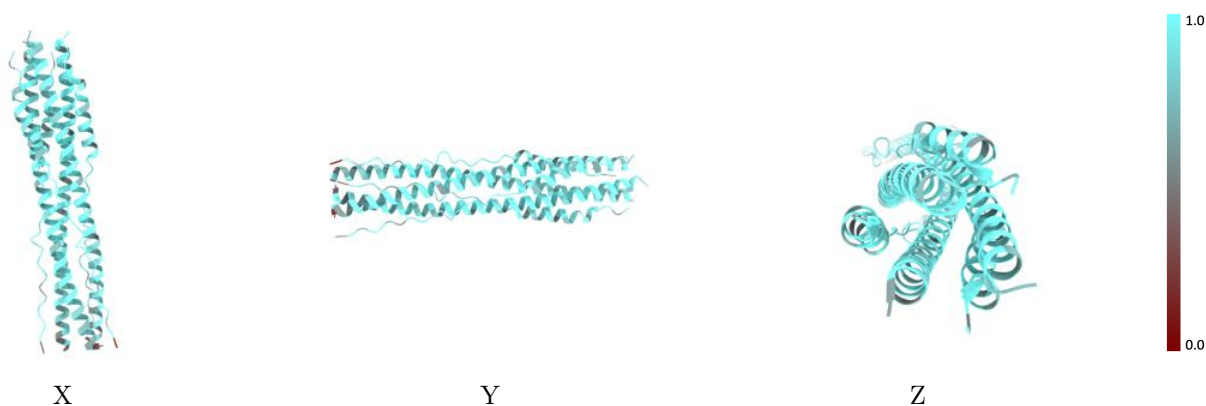
The images above show the 3D surface view of the map at the recommended contour level 0.05 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



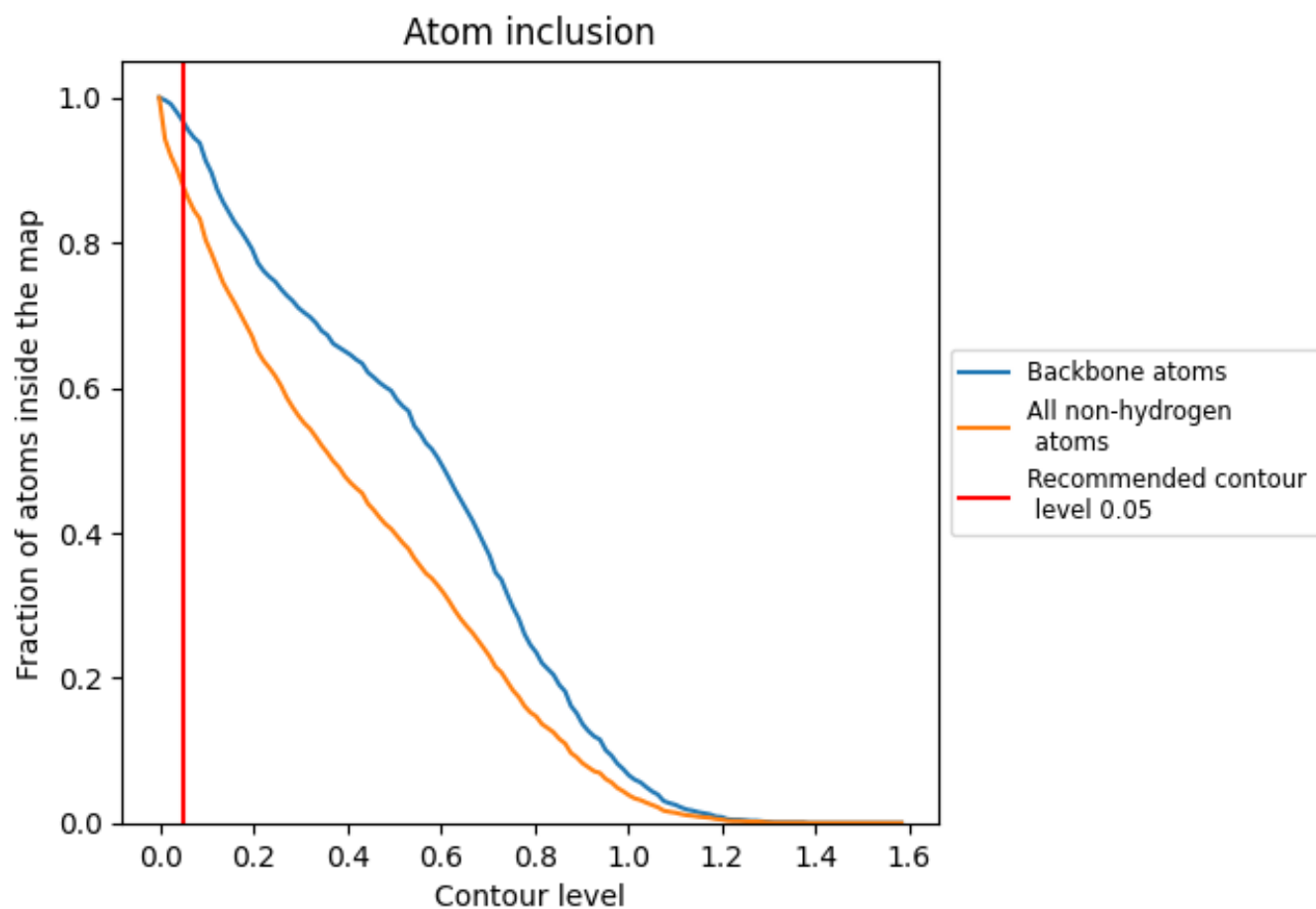
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model [\(i\)](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.05).















9.4 Atom inclusion [i](#)



At the recommended contour level, 97% of all backbone atoms, 88% 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 (0.05) and Q-score for the entire model and for each chain.

| Chain | Atom inclusion | Q-score |
|-------|--|--|
| All |  0.8767 |  0.4890 |
| A |  0.8790 |  0.4830 |
| B |  0.8809 |  0.4900 |
| C |  0.8979 |  0.4920 |
| D |  0.8580 |  0.4830 |
| E |  0.8707 |  0.4970 |
| F |  0.8549 |  0.4910 |

