



## wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 06:10 PM EST

PDB ID : 7KH0  
EMDB ID : EMD-22872  
Title : Cryo-EM structure of the human arginine vasopressin AVP-vasopressin receptor V2R-Gs signaling complex  
Authors : Wang, L.; Xu, J.; Gao, S.; Sun, D.; Liu, H.; Liu, Z.; Du, Y.; Zhang, C.  
Deposited on : 2020-10-19  
Resolution : 2.80 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](mailto: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) : 2.31.2

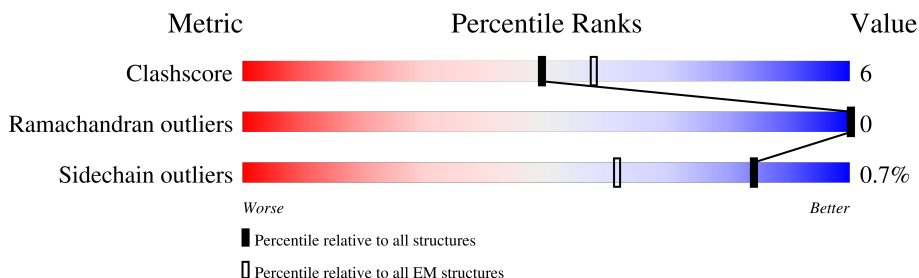
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.80 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  $\geq 3$ , 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	R	389	<div> <div>13%</div> <div>56%</div> <div>15%</div> <div>29%</div> </div>
2	B	350	<div> <div>84%</div> <div>13%</div> <div>.</div> </div>
3	A	372	<div> <div>7%</div> <div>52%</div> <div>9%</div> <div>39%</div> </div>
4	L	10	<div> <div>10%</div> <div>80%</div> <div>20%</div> </div>
5	G	67	<div> <div>9%</div> <div>76%</div> <div>13%</div> <div>10%</div> </div>
6	S	248	<div> <div>87%</div> <div>8%</div> <div>5%</div> </div>
7	N	134	<div> <div>81%</div> <div>13%</div> <div>6%</div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9898 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 Vasopressin V2 receptor.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	R	276	Total	C	N	O	S	0	0
			2117	1400	359	340	18		

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	-8	ASP	-	expression tag	UNP P30518
R	-7	TYR	-	expression tag	UNP P30518
R	-6	LYS	-	expression tag	UNP P30518
R	-5	ASP	-	expression tag	UNP P30518
R	-4	ASP	-	expression tag	UNP P30518
R	-3	ASP	-	expression tag	UNP P30518
R	-2	ASP	-	expression tag	UNP P30518
R	-1	VAL	-	expression tag	UNP P30518
R	0	ASP	-	expression tag	UNP P30518
R	372	ALA	-	expression tag	UNP P30518
R	373	ALA	-	expression tag	UNP P30518
R	374	ALA	-	expression tag	UNP P30518
R	375	LEU	-	expression tag	UNP P30518
R	376	GLU	-	expression tag	UNP P30518
R	377	VAL	-	expression tag	UNP P30518
R	378	LEU	-	expression tag	UNP P30518
R	379	PHE	-	expression tag	UNP P30518
R	380	GLN	-	expression tag	UNP P30518

- Molecule 2 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	339	Total	C	N	O	S	0	0
			2607	1607	468	511	21		

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	HIS	-	expression tag	UNP P62873
B	-8	HIS	-	expression tag	UNP P62873
B	-7	HIS	-	expression tag	UNP P62873
B	-6	HIS	-	expression tag	UNP P62873
B	-5	HIS	-	expression tag	UNP P62873
B	-4	HIS	-	expression tag	UNP P62873
B	-3	GLY	-	expression tag	UNP P62873
B	-2	SER	-	expression tag	UNP P62873
B	-1	LEU	-	expression tag	UNP P62873
B	0	LEU	-	expression tag	UNP P62873
B	1	GLN	-	expression tag	UNP P62873

- Molecule 3 is a protein called Guanine nucleotide-binding protein G(i) subunit alpha-3, Isoform Gnas-2 of Guanine nucleotide-binding protein G(s) subunit alpha isoforms short fusion.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	A	227	Total	C	N	O	S	0	0
			1869	1184	336	341	8		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	188	ASP	ALA	conflict	UNP P63092-2

- Molecule 4 is a protein called Arg-vasopressin.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	L	10	Total	C	N	O	S	0	1
			75	46	15	12	2		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	10	NH2	-	amidation	UNP P01185

- Molecule 5 is a protein called Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	G	60	Total	C	N	O	S	0	0
			462	288	82	89	3		

- Molecule 6 is a protein called Single Fab chain (scFv16).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	S	236	Total	C	N	O	S	0	0
			1807	1146	300	351	10		

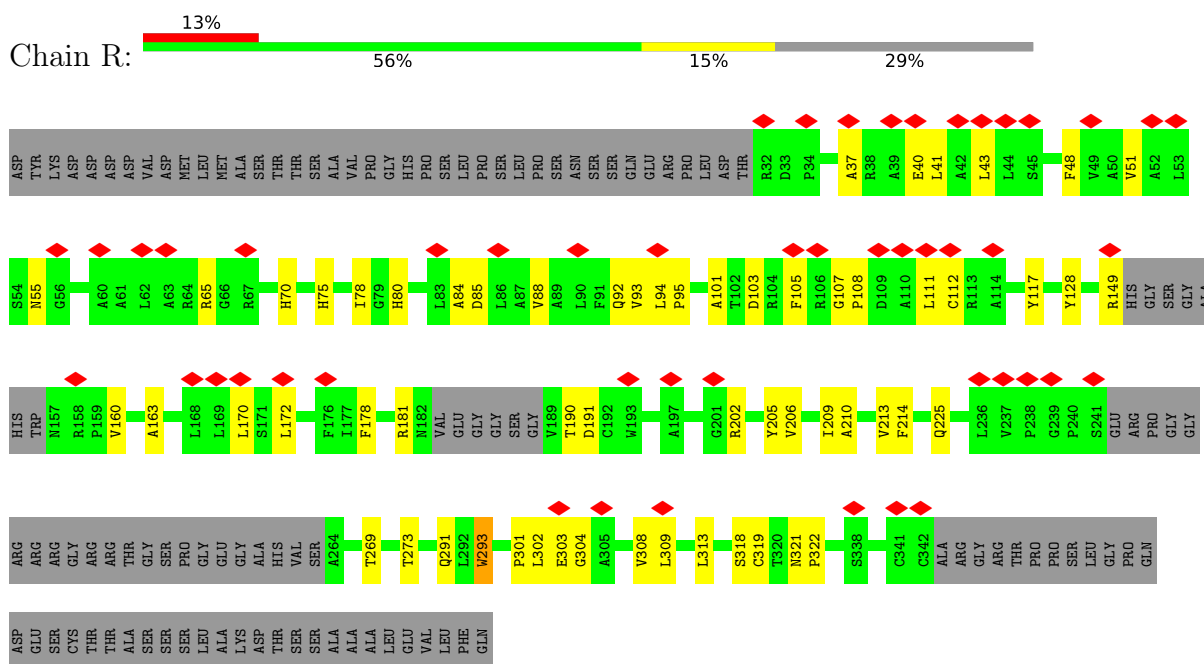
- Molecule 7 is a protein called Nanobody 35.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	N	126	Total	C	N	O	S	0	0
			961	599	168	188	6		

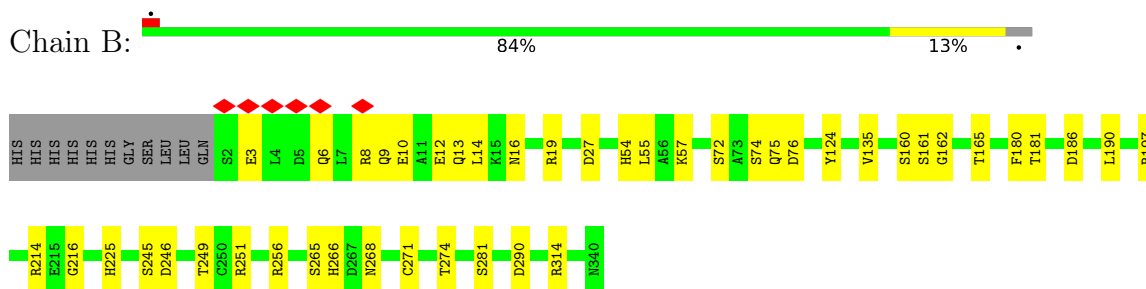
### 3 Residue-property plots

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: Vasopressin V2 receptor

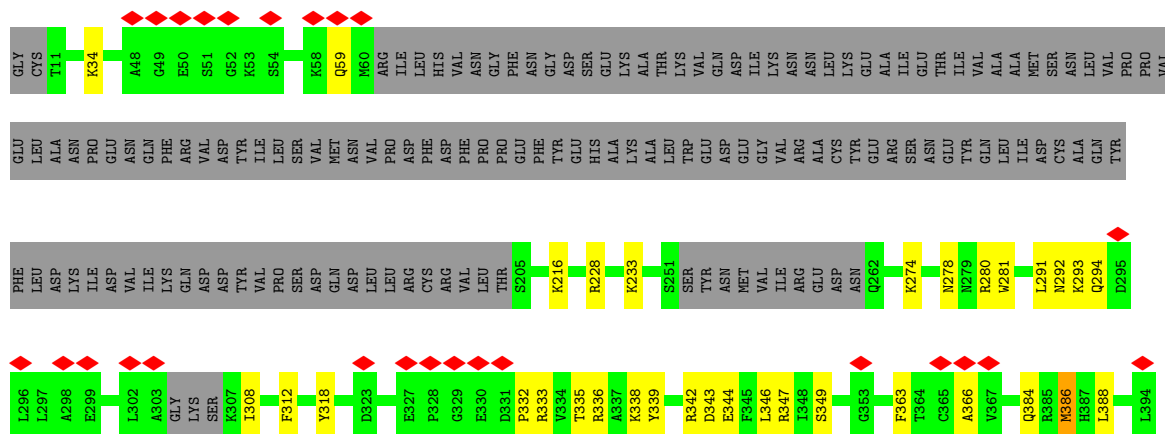


- Molecule 2: Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

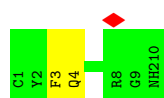
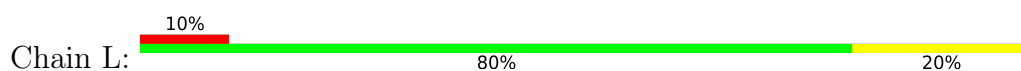


- Molecule 3: Guanine nucleotide-binding protein G(i) subunit alpha-3, Isoform Gnas-2 of Guanine nucleotide-binding protein G(s) subunit alpha isoforms short fusion

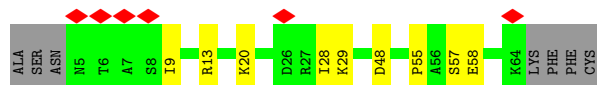
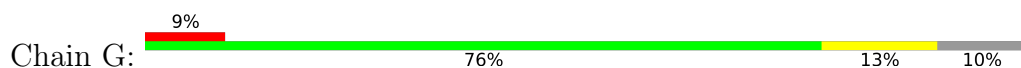




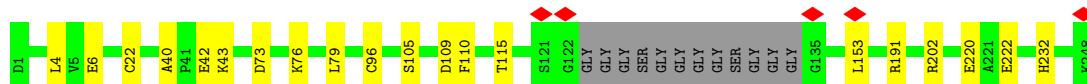
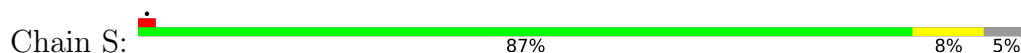
• Molecule 4: Arg-vasopressin



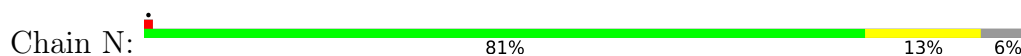
• Molecule 5: Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2



• Molecule 6: Single Fab chain (scFv16)



• Molecule 7: Nanobody 35



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	398038	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 ( $e^-/\text{\AA}^2$ )	55	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	1.344	Depositor
Minimum map value	-0.536	Depositor
Average map value	0.003	Depositor
Map value standard deviation	0.051	Depositor
Recommended contour level	0.3	Depositor
Map size ( $\text{\AA}$ )	217.6, 217.6, 217.6	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	0.85, 0.85, 0.85	Depositor



## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NH2

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	R	0.37	0/2174	0.57	1/2975 (0.0%)
2	B	0.29	0/2654	0.54	0/3597
3	A	0.30	0/1903	0.53	0/2558
4	L	0.52	0/76	1.03	0/101
5	G	0.27	0/468	0.52	0/631
6	S	0.30	0/1851	0.51	0/2508
7	N	0.31	0/981	0.53	0/1329
All	All	0.32	0/10107	0.54	1/13699 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	178	PHE	CB-CA-C	5.02	120.44	110.40

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	R	2117	0	2164	40	0
2	B	2607	0	2510	33	0
3	A	1869	0	1860	21	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	75	0	65	2	0
5	G	462	0	473	10	0
6	S	1807	0	1748	13	0
7	N	961	0	928	16	0
All	All	9898	0	9748	120	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 6.

The worst 5 of 120 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:A:280:ARG:NH1	3:A:281:TRP:CH2	2.34	0.95
2:B:54:HIS:HE2	2:B:72:SER:HG	1.22	0.88
7:N:100:PRO:HD2	7:N:107:CYS:HB2	1.59	0.85
1:R:43:LEU:HD21	1:R:313:LEU:HD11	1.66	0.77
2:B:8:ARG:NH2	7:N:120:GLN:OE1	2.18	0.77

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	R	268/389 (69%)	260 (97%)	8 (3%)	0	100	100
2	B	337/350 (96%)	328 (97%)	9 (3%)	0	100	100
3	A	219/372 (59%)	214 (98%)	5 (2%)	0	100	100
4	L	8/10 (80%)	7 (88%)	1 (12%)	0	100	100
5	G	58/67 (87%)	56 (97%)	2 (3%)	0	100	100
6	S	232/248 (94%)	224 (97%)	8 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	N	124/134 (92%)	119 (96%)	5 (4%)	0	100	100
All	All	1246/1570 (79%)	1208 (97%)	38 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	R	217/315 (69%)	213 (98%)	4 (2%)	59	86
2	B	282/292 (97%)	282 (100%)	0	100	100
3	A	202/334 (60%)	199 (98%)	3 (2%)	65	89
4	L	8/8 (100%)	8 (100%)	0	100	100
5	G	49/55 (89%)	49 (100%)	0	100	100
6	S	199/201 (99%)	199 (100%)	0	100	100
7	N	104/112 (93%)	104 (100%)	0	100	100
All	All	1061/1317 (81%)	1054 (99%)	7 (1%)	84	95

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	R	303	GLU
3	A	384	GLN
3	A	388	LEU
3	A	386	MET
1	R	302	LEU

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

Mol	Chain	Res	Type
1	R	174	GLN
3	A	294	GLN

### 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.

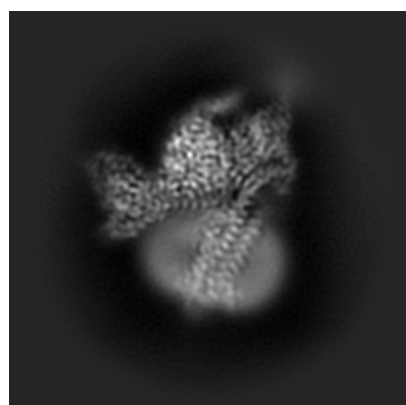
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22872. These allow visual inspection of the internal detail of the map and identification of artifacts.

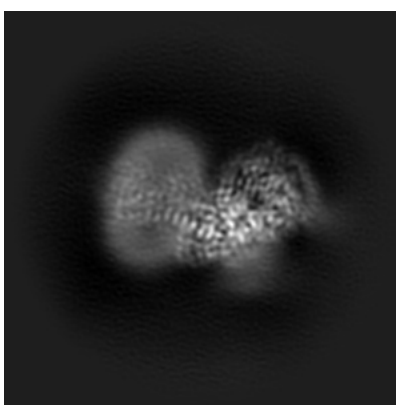
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

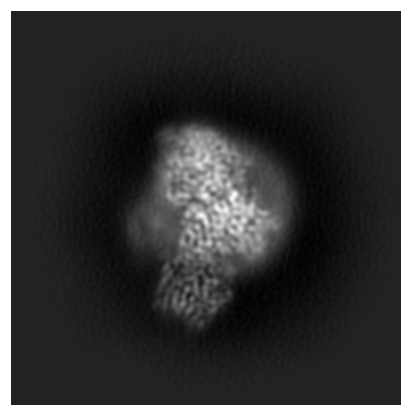
#### 6.1.1 Primary map



X



Y

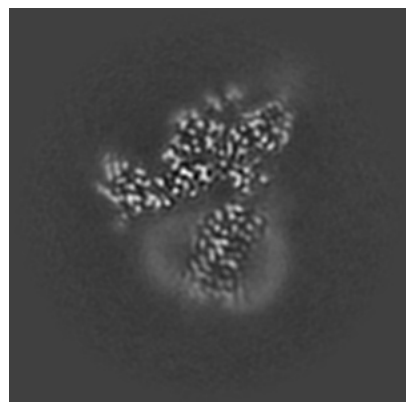


Z

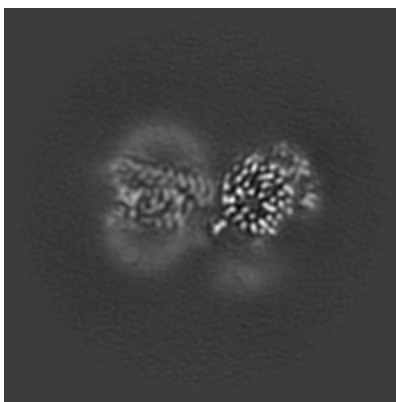
The images above show the map projected in three orthogonal directions.

### 6.2 Central slices [i](#)

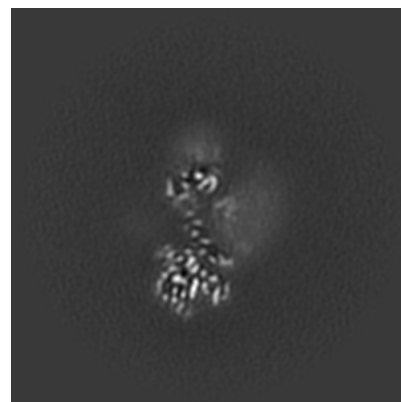
#### 6.2.1 Primary map



X Index: 128



Y Index: 128

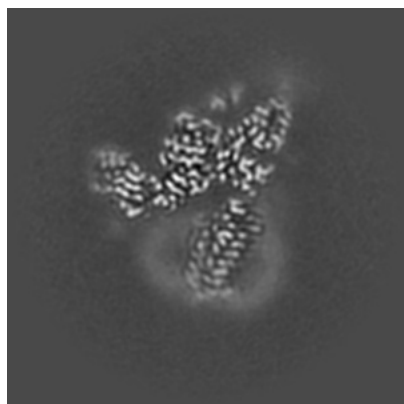


Z Index: 128

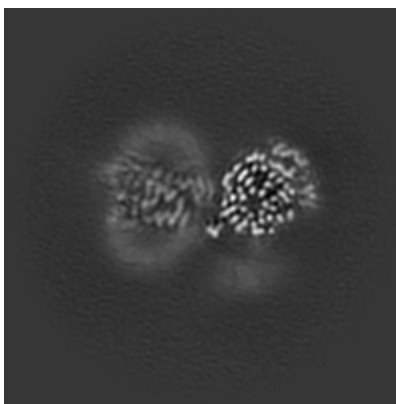
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

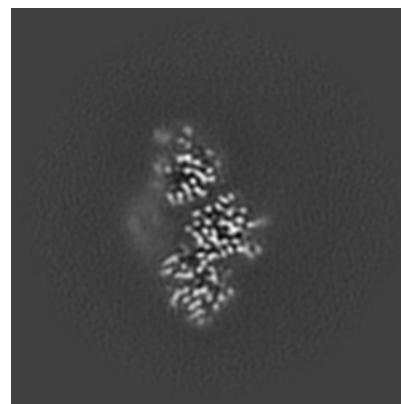
### 6.3.1 Primary map



X Index: 125



Y Index: 126

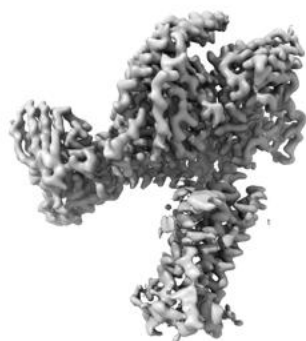


Z Index: 144

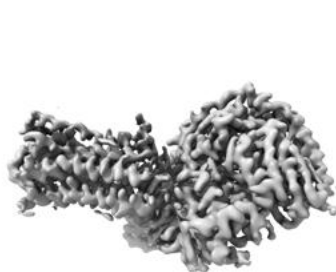
The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal surface views [i](#)

### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.3. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

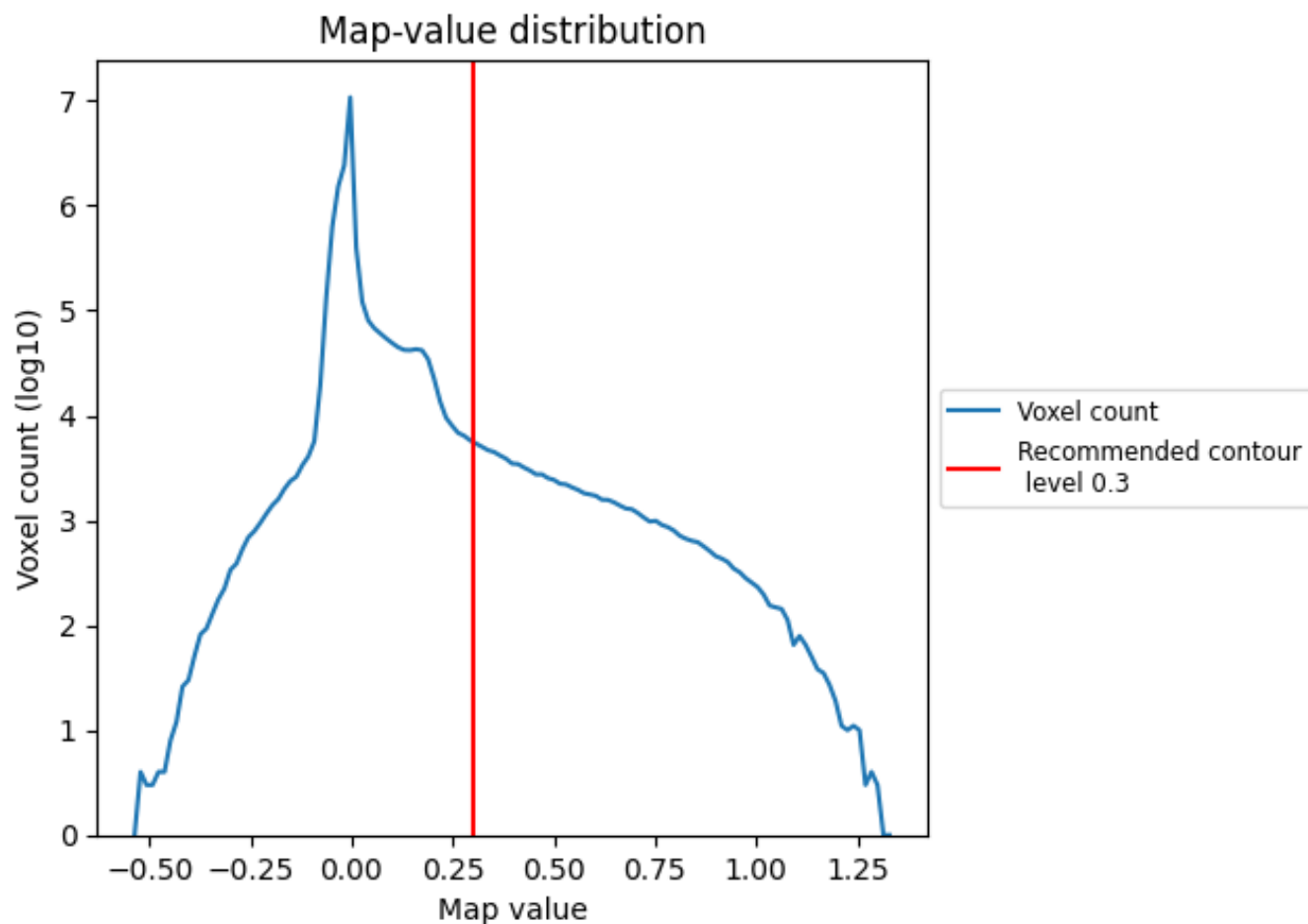
## 6.5 Mask visualisation

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.

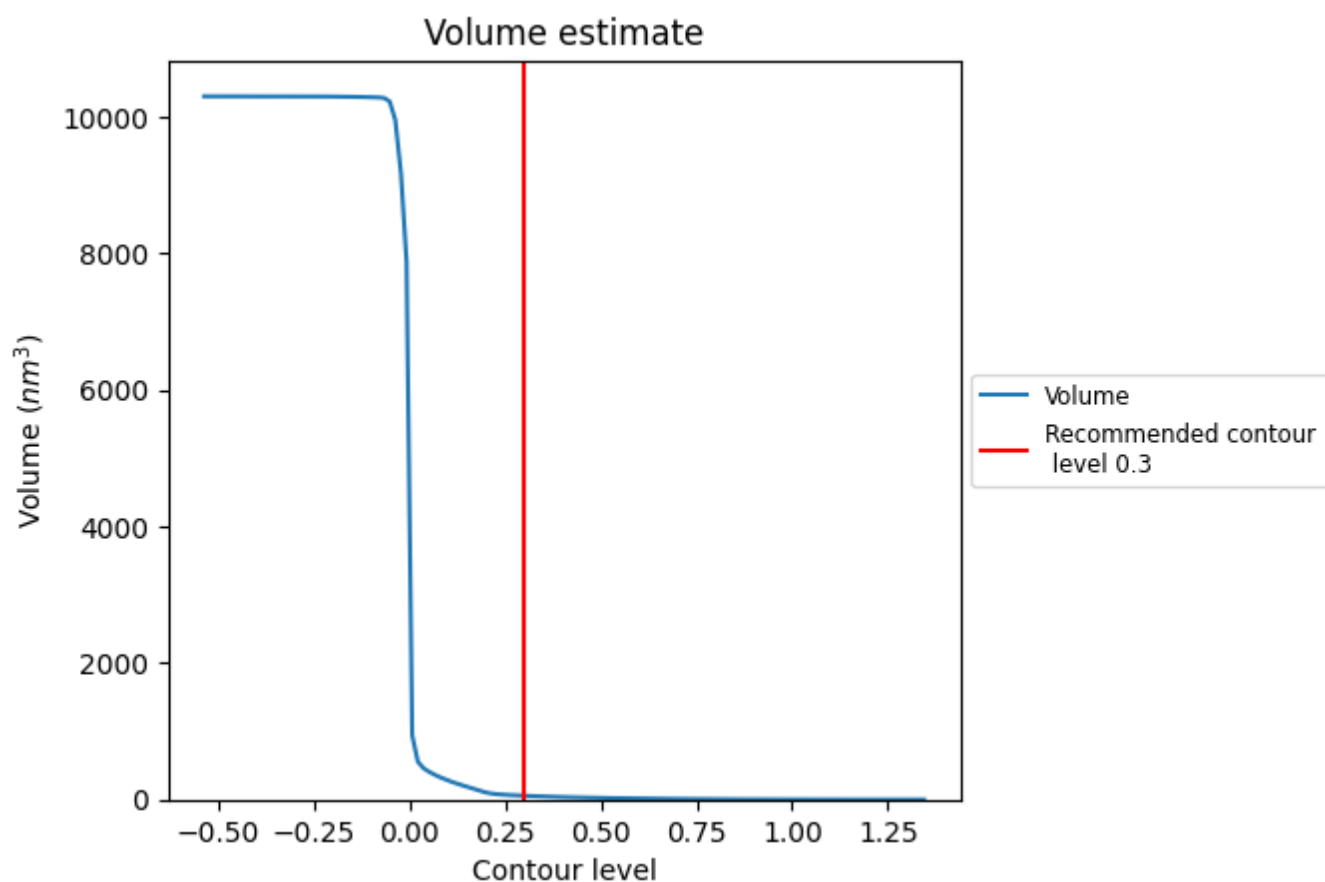
### 7.1 Map-value distribution [i](#)



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.



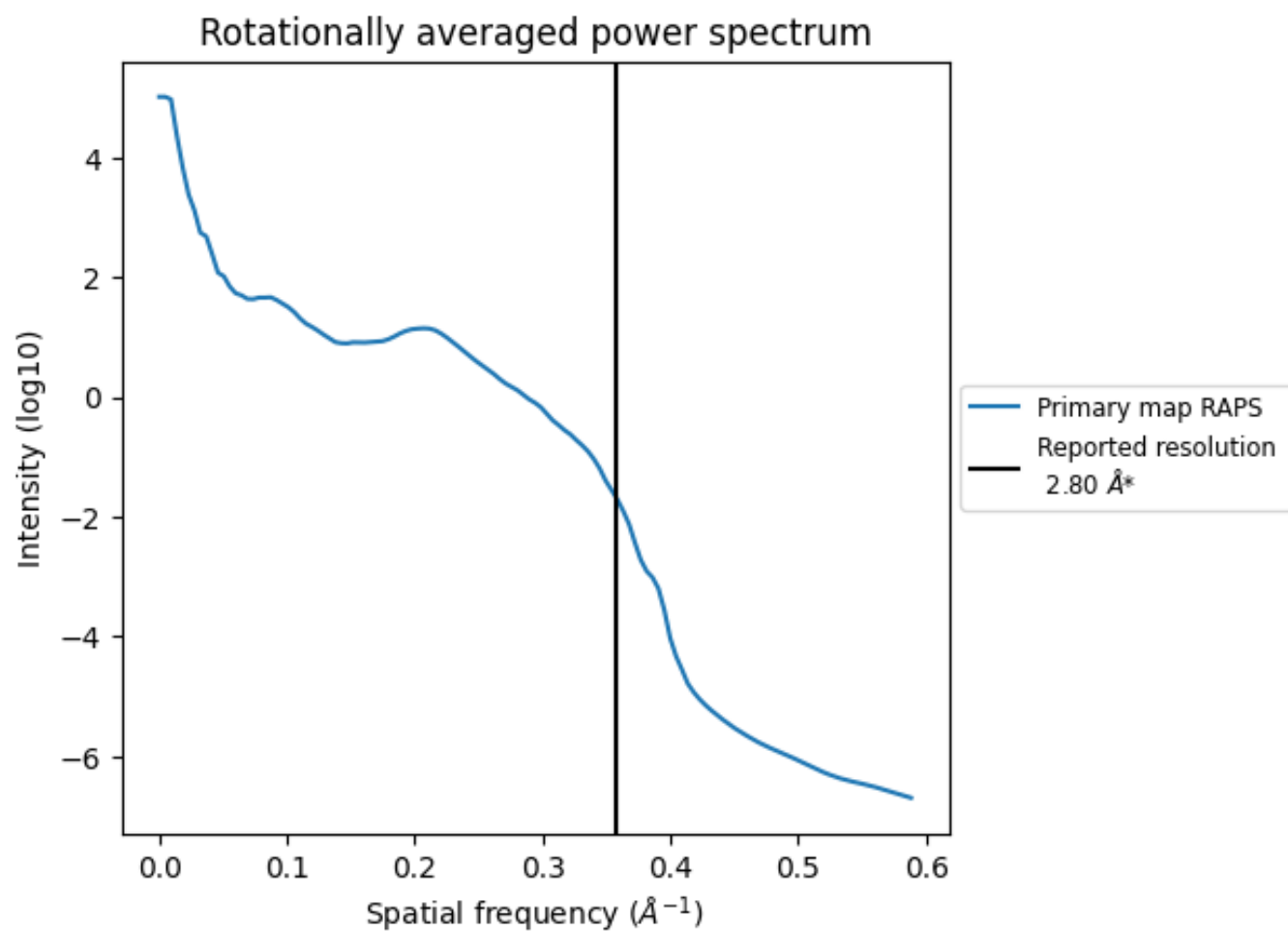
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 56 nm<sup>3</sup>; this corresponds to an approximate mass of 50 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 0.357 Å<sup>-1</sup>

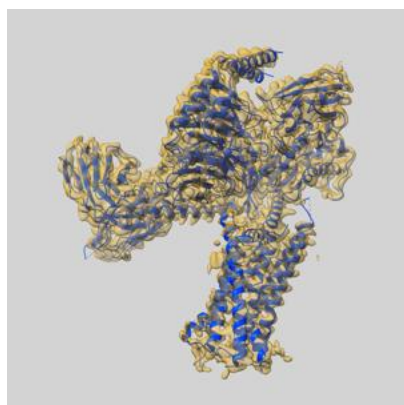
## 8 Fourier-Shell correlation ⓘ

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

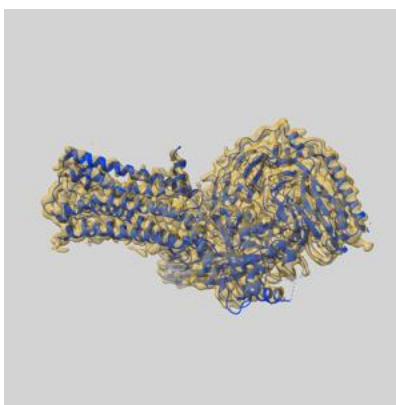
## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-22872 and PDB model 7KH0. Per-residue inclusion information can be found in section 3 on page 6.

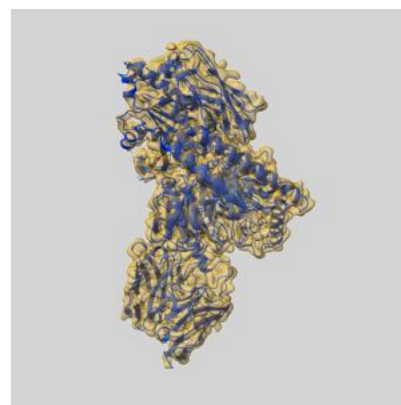
### 9.1 Map-model overlay [i](#)



X



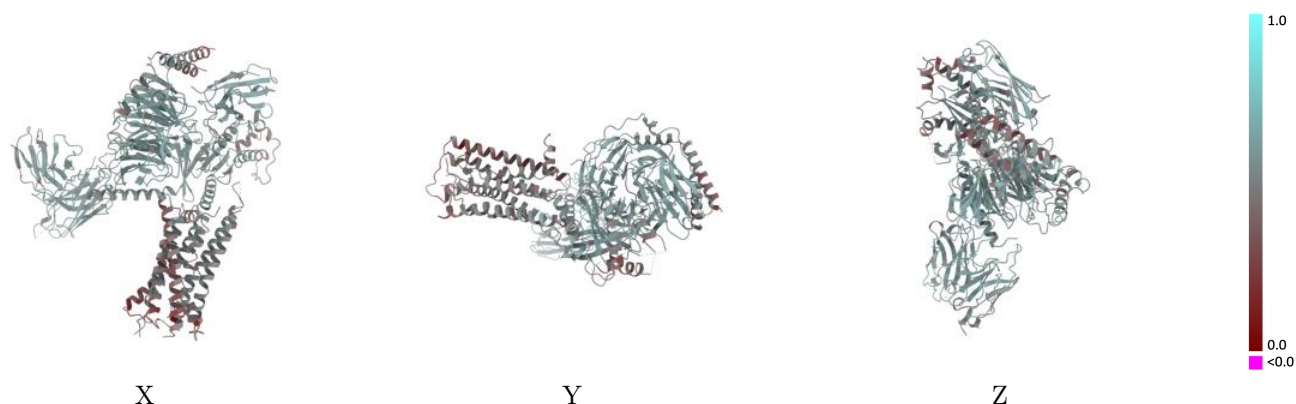
Y



Z

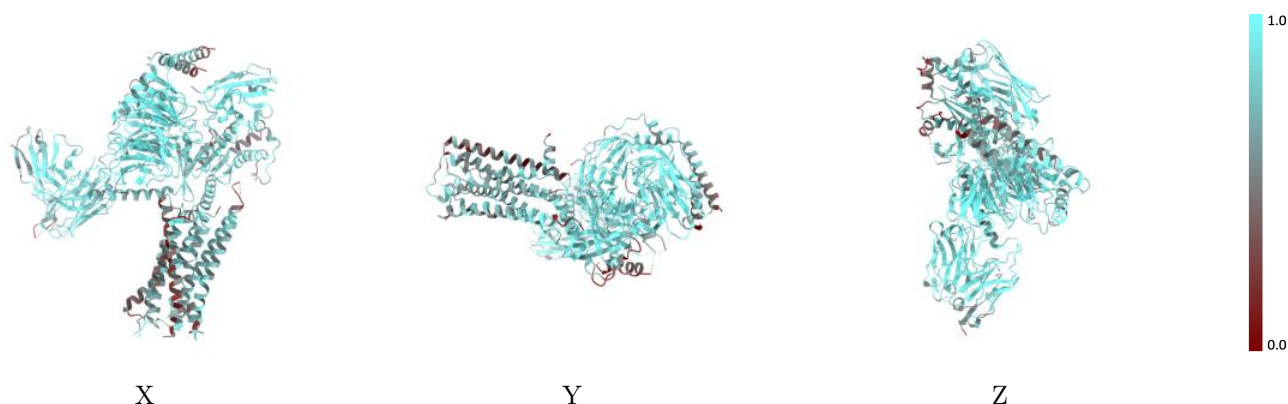
The images above show the 3D surface view of the map at the recommended contour level 0.3 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



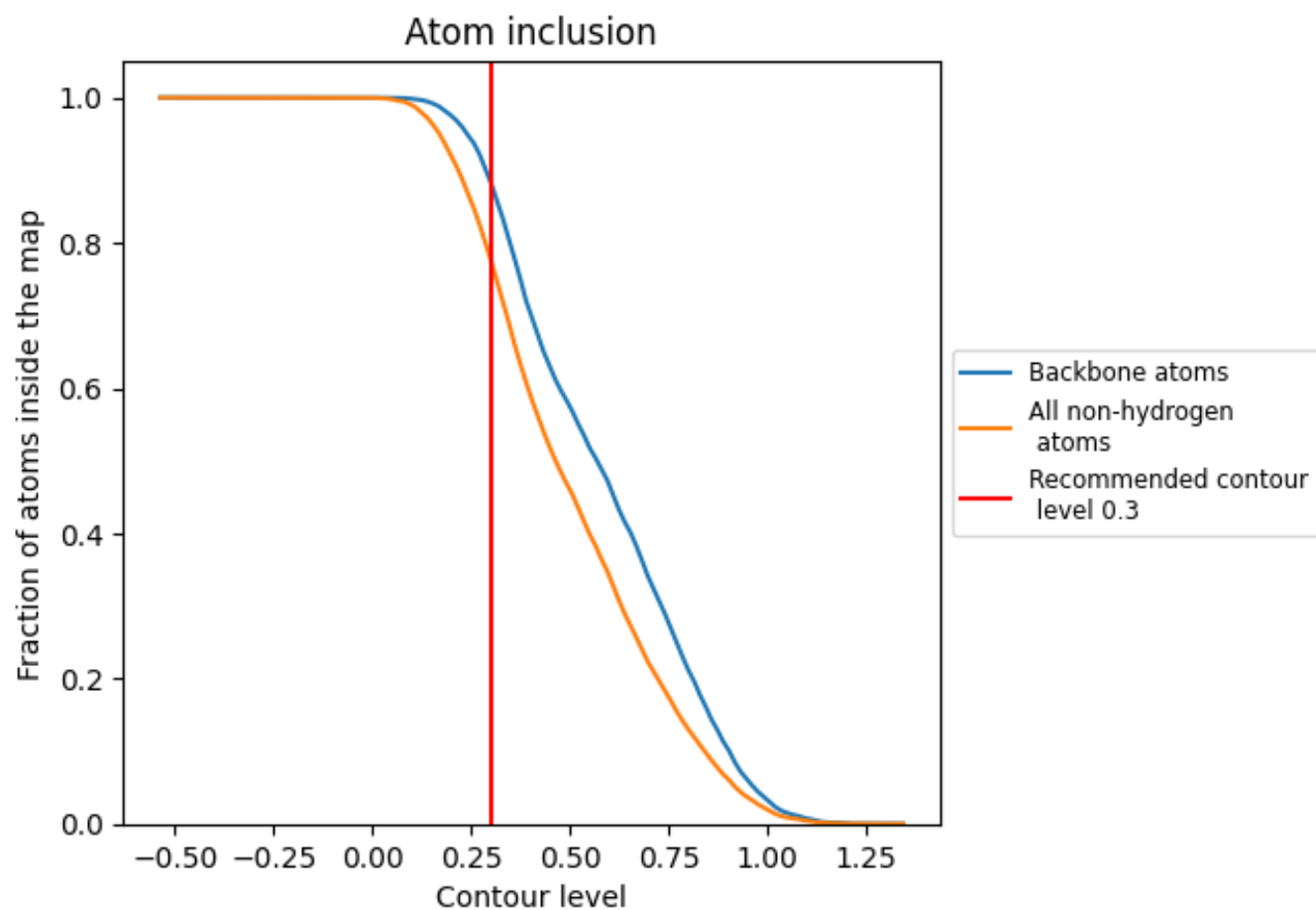
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.3).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 88% of all backbone atoms, 78% 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 (0.3) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div><div></div>0.7782</div>	<div><div></div>0.5230</div>
A	<div><div></div>0.7471</div>	<div><div></div>0.5220</div>
B	<div><div></div>0.8787</div>	<div><div></div>0.5630</div>
G	<div><div></div>0.6725</div>	<div><div></div>0.4820</div>
L	<div><div></div>0.7222</div>	<div><div></div>0.5000</div>
N	<div><div></div>0.8422</div>	<div><div></div>0.5580</div>
R	<div><div></div>0.6215</div>	<div><div></div>0.4420</div>
S	<div><div></div>0.8444</div>	<div><div></div>0.5570</div>

1.0

0.0

<0.0