



Full wwPDB EM Validation Report ⓘ

Feb 25, 2024 – 12:11 PM EST

PDB ID : 6X0M
EMDB ID : EMD-21979
Title : Bridging of double-strand DNA break activates PARP2/HPF1 to modify chromatin
Authors : Halic, M.; Bilokapic, S.
Deposited on : 2020-05-16
Resolution : 6.30 Å(reported)

This is a Full wwPDB EM Validation 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:

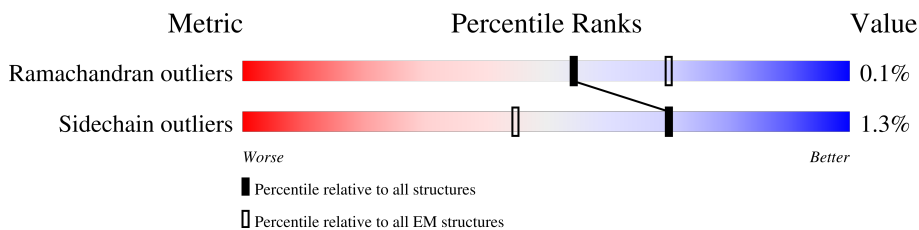
EMDB validation analysis : 0.0.1.dev70
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

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 6.30 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)
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	O	356	21% (red), 69% (green), 31% (grey), 0% (yellow), 0% (orange)
1	o	356	21% (red), 68% (green), 32% (grey), 0% (yellow), 0% (orange)
2	P	590	14% (red), 78% (green), 20% (grey), 0% (yellow), 0% (orange)
2	p	590	12% (red), 78% (green), 21% (grey), 0% (yellow), 0% (orange)
3	I	167	6% (green), 93% (grey), 0% (red), 0% (yellow), 0% (orange)
3	i	167	7% (green), 93% (grey), 0% (red), 0% (yellow), 0% (orange)
4	J	167	6% (green), 93% (grey), 0% (red), 0% (yellow), 0% (orange)
4	j	167	5% (green), 93% (grey), 0% (red), 0% (yellow), 0% (orange)

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 12421 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 Histone PARylation factor 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	O	246	1995	1280	338	367	10	0	0
1	o	242	1967	1264	332	361	10	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	-9	MET	-	initiating methionine	UNP Q9NWX4
O	-8	GLY	-	expression tag	UNP Q9NWX4
O	-7	HIS	-	expression tag	UNP Q9NWX4
O	-6	HIS	-	expression tag	UNP Q9NWX4
O	-5	HIS	-	expression tag	UNP Q9NWX4
O	-4	HIS	-	expression tag	UNP Q9NWX4
O	-3	HIS	-	expression tag	UNP Q9NWX4
O	-2	HIS	-	expression tag	UNP Q9NWX4
O	-1	GLY	-	expression tag	UNP Q9NWX4
O	0	GLY	-	expression tag	UNP Q9NWX4
O	174	LYS	ARG	variant	UNP Q9NWX4
o	-9	MET	-	initiating methionine	UNP Q9NWX4
o	-8	GLY	-	expression tag	UNP Q9NWX4
o	-7	HIS	-	expression tag	UNP Q9NWX4
o	-6	HIS	-	expression tag	UNP Q9NWX4
o	-5	HIS	-	expression tag	UNP Q9NWX4
o	-4	HIS	-	expression tag	UNP Q9NWX4
o	-3	HIS	-	expression tag	UNP Q9NWX4
o	-2	HIS	-	expression tag	UNP Q9NWX4
o	-1	GLY	-	expression tag	UNP Q9NWX4
o	0	GLY	-	expression tag	UNP Q9NWX4
o	174	LYS	ARG	variant	UNP Q9NWX4

- Molecule 2 is a protein called Poly [ADP-ribose] polymerase 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	P	470	Total	C	N	O	S	0	0
			3761	2388	649	697	27		
2	p	468	Total	C	N	O	S	0	0
			3754	2386	646	695	27		

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	-19	MET	-	initiating methionine	UNP Q9UGN5
P	-18	GLY	-	expression tag	UNP Q9UGN5
P	-17	SER	-	expression tag	UNP Q9UGN5
P	-16	SER	-	expression tag	UNP Q9UGN5
P	-15	HIS	-	expression tag	UNP Q9UGN5
P	-14	HIS	-	expression tag	UNP Q9UGN5
P	-13	HIS	-	expression tag	UNP Q9UGN5
P	-12	HIS	-	expression tag	UNP Q9UGN5
P	-11	HIS	-	expression tag	UNP Q9UGN5
P	-10	HIS	-	expression tag	UNP Q9UGN5
P	-9	SER	-	expression tag	UNP Q9UGN5
P	-8	SER	-	expression tag	UNP Q9UGN5
P	-7	GLY	-	expression tag	UNP Q9UGN5
P	-6	LEU	-	expression tag	UNP Q9UGN5
P	-5	VAL	-	expression tag	UNP Q9UGN5
P	-4	PRO	-	expression tag	UNP Q9UGN5
P	-3	ARG	-	expression tag	UNP Q9UGN5
P	-2	GLY	-	expression tag	UNP Q9UGN5
P	-1	SER	-	expression tag	UNP Q9UGN5
P	0	HIS	-	expression tag	UNP Q9UGN5
p	-19	MET	-	initiating methionine	UNP Q9UGN5
p	-18	GLY	-	expression tag	UNP Q9UGN5
p	-17	SER	-	expression tag	UNP Q9UGN5
p	-16	SER	-	expression tag	UNP Q9UGN5
p	-15	HIS	-	expression tag	UNP Q9UGN5
p	-14	HIS	-	expression tag	UNP Q9UGN5
p	-13	HIS	-	expression tag	UNP Q9UGN5
p	-12	HIS	-	expression tag	UNP Q9UGN5
p	-11	HIS	-	expression tag	UNP Q9UGN5
p	-10	HIS	-	expression tag	UNP Q9UGN5
p	-9	SER	-	expression tag	UNP Q9UGN5
p	-8	SER	-	expression tag	UNP Q9UGN5
p	-7	GLY	-	expression tag	UNP Q9UGN5
p	-6	LEU	-	expression tag	UNP Q9UGN5
p	-5	VAL	-	expression tag	UNP Q9UGN5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
p	-4	PRO	-	expression tag	UNP Q9UGN5
p	-3	ARG	-	expression tag	UNP Q9UGN5
p	-2	GLY	-	expression tag	UNP Q9UGN5
p	-1	SER	-	expression tag	UNP Q9UGN5
p	0	HIS	-	expression tag	UNP Q9UGN5

- Molecule 3 is a DNA chain called DNA (167-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	I	12	243	116	46	69	12	0	0
3	i	11	229	108	45	65	11	0	0

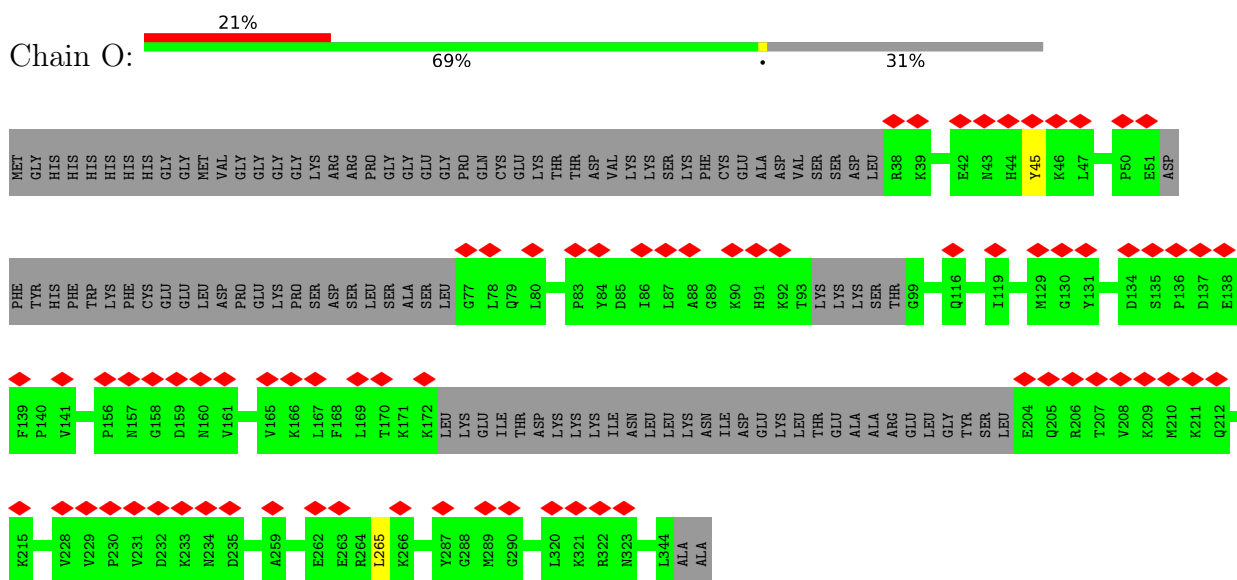
- Molecule 4 is a DNA chain called DNA (167-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	J	12	249	119	43	75	12	0	0
4	j	11	223	106	38	68	11	0	0

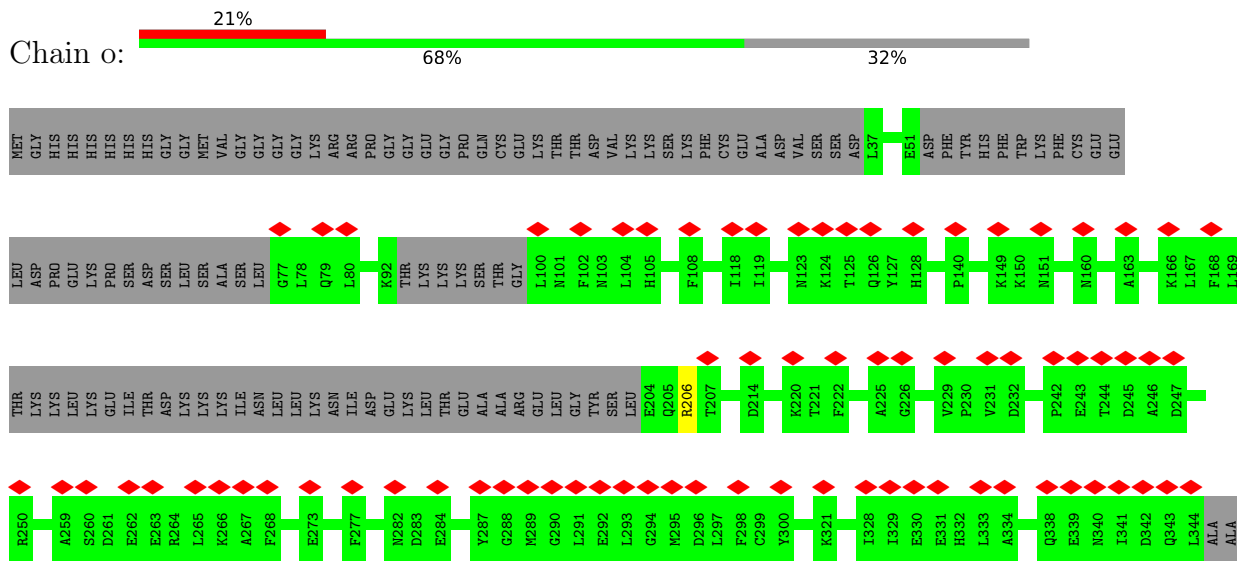
3 Residue-property plots

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: Histone PARylation factor 1



- Molecule 1: Histone PARylation factor 1



- Molecule 2: Poly [ADP-ribose] polymerase 2

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	16000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	80, 80	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k), GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.299	Depositor
Minimum map value	-0.102	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.011	Depositor
Recommended contour level	0.06	Depositor
Map size (\AA)	508.8, 508.8, 508.8	wwPDB
Map dimensions	240, 240, 240	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	2.12, 2.12, 2.12	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	O	0.44	0/2038	0.70	1/2746 (0.0%)
1	o	0.40	0/2010	0.65	0/2710
2	P	0.40	0/3842	0.70	2/5182 (0.0%)
2	p	0.39	0/3836	0.70	2/5176 (0.0%)
3	I	0.82	0/272	1.21	2/416 (0.5%)
3	i	0.91	0/257	1.04	0/395
4	J	0.91	0/278	1.29	2/428 (0.5%)
4	j	1.10	1/248 (0.4%)	1.15	1/378 (0.3%)
All	All	0.47	1/12781 (0.0%)	0.75	10/17431 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	j	-83	DC	OP3-P	-10.68	1.48	1.61

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	82	DT	O4'-C1'-N1	6.78	112.74	108.00
3	I	-73	DA	O4'-C1'-N9	-6.14	103.70	108.00
2	P	221	LEU	CA-CB-CG	5.99	129.07	115.30
2	p	333	LEU	CA-CB-CG	5.74	128.50	115.30
3	I	-81	DA	O4'-C1'-N9	5.34	111.74	108.00
2	p	194	ASP	CB-CG-OD2	5.25	123.02	118.30
2	P	194	ASP	CB-CG-OD2	5.19	122.97	118.30
4	j	-82	DT	N3-C4-O4	5.16	123.00	119.90
1	O	265	LEU	CA-CB-CG	5.15	127.16	115.30
4	J	77	DT	O4'-C1'-N1	5.07	111.55	108.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	O	238/356 (67%)	233 (98%)	5 (2%)	0	100	100
1	o	234/356 (66%)	230 (98%)	4 (2%)	0	100	100
2	P	466/590 (79%)	443 (95%)	22 (5%)	1 (0%)	47	81
2	p	464/590 (79%)	448 (97%)	16 (3%)	0	100	100
All	All	1402/1892 (74%)	1354 (97%)	47 (3%)	1 (0%)	54	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	P	246	LYS

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	O	215/308 (70%)	214 (100%)	1 (0%)	88	93
1	o	212/308 (69%)	211 (100%)	1 (0%)	88	93
2	P	409/511 (80%)	402 (98%)	7 (2%)	60	78
2	p	408/511 (80%)	401 (98%)	7 (2%)	60	78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1244/1638 (76%)	1228 (99%)	16 (1%)	70 82

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	O	45	TYR
2	P	137	ARG
2	P	220	GLN
2	P	338	LEU
2	P	410	ARG
2	P	432	ILE
2	P	495	LEU
2	P	565	ASN
1	o	206	ARG
2	p	137	ARG
2	p	280	GLN
2	p	339	GLN
2	p	387	THR
2	p	432	ILE
2	p	556	MET
2	p	557	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	O	272	GLN
1	O	323	ASN
2	P	174	ASN
2	P	226	GLN
2	P	311	GLN
2	P	319	GLN
2	P	343	HIS
2	P	409	ASN
2	P	514	HIS
2	P	563	GLN
2	P	565	ASN
1	o	272	GLN
2	p	174	ASN
2	p	226	GLN
2	p	267	GLN
2	p	281	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	p	290	ASN
2	p	553	GLN
2	p	563	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.

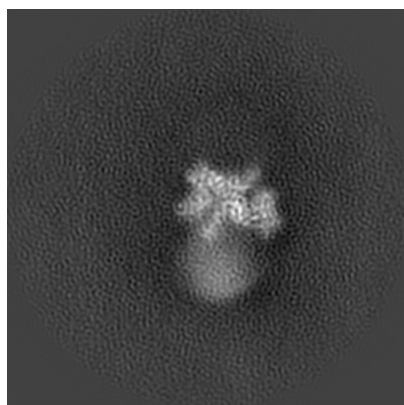
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-21979. These allow visual inspection of the internal detail of the map and identification of artifacts.

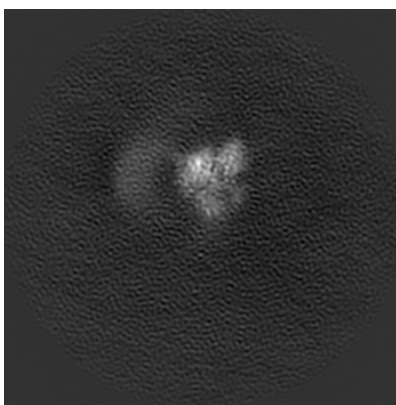
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

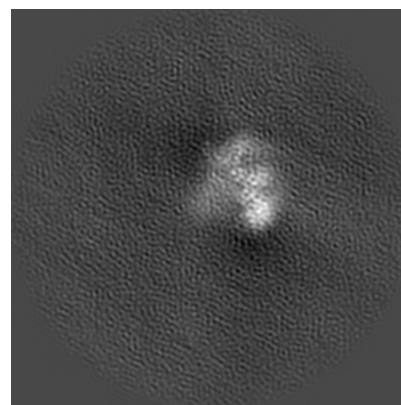
6.1.1 Primary map



X



Y

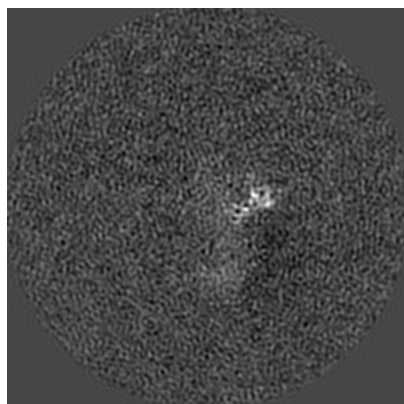


Z

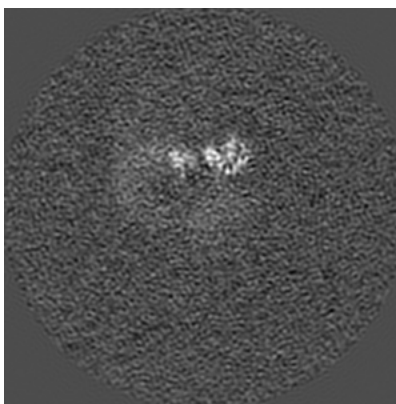
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

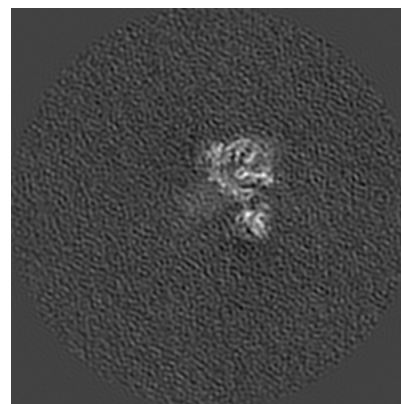
6.2.1 Primary map



X Index: 120



Y Index: 120

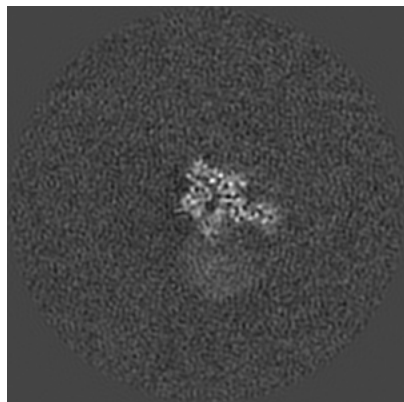


Z Index: 120

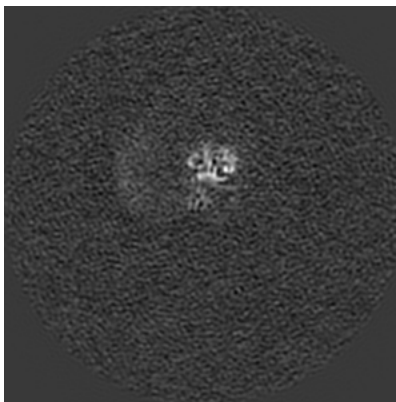
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

6.3.1 Primary map



X Index: 145



Y Index: 138

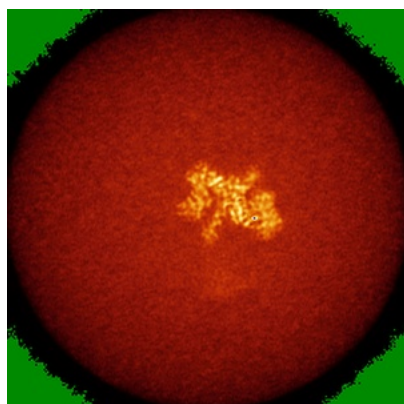


Z Index: 123

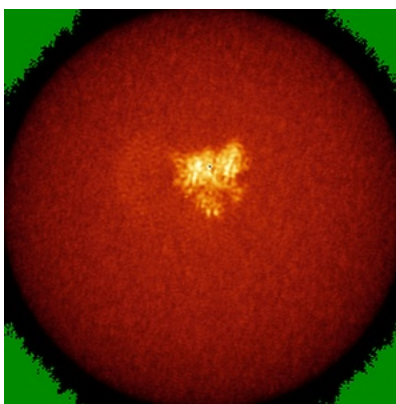
The images above show the largest variance slices of the map in three orthogonal directions.

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

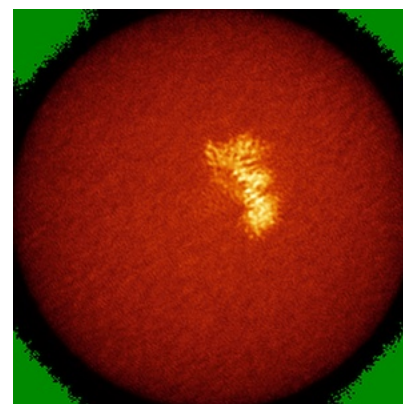
6.4.1 Primary map



X



Y

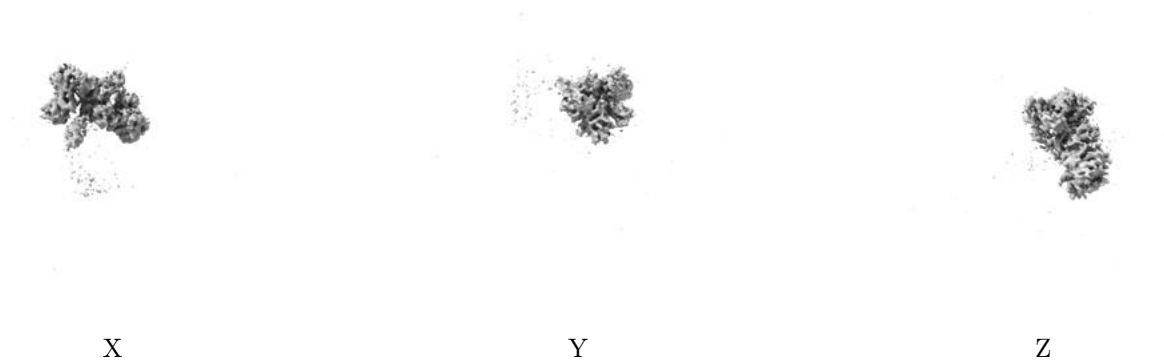


Z

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



The images above show the 3D surface view of the map at the recommended contour level 0.06. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

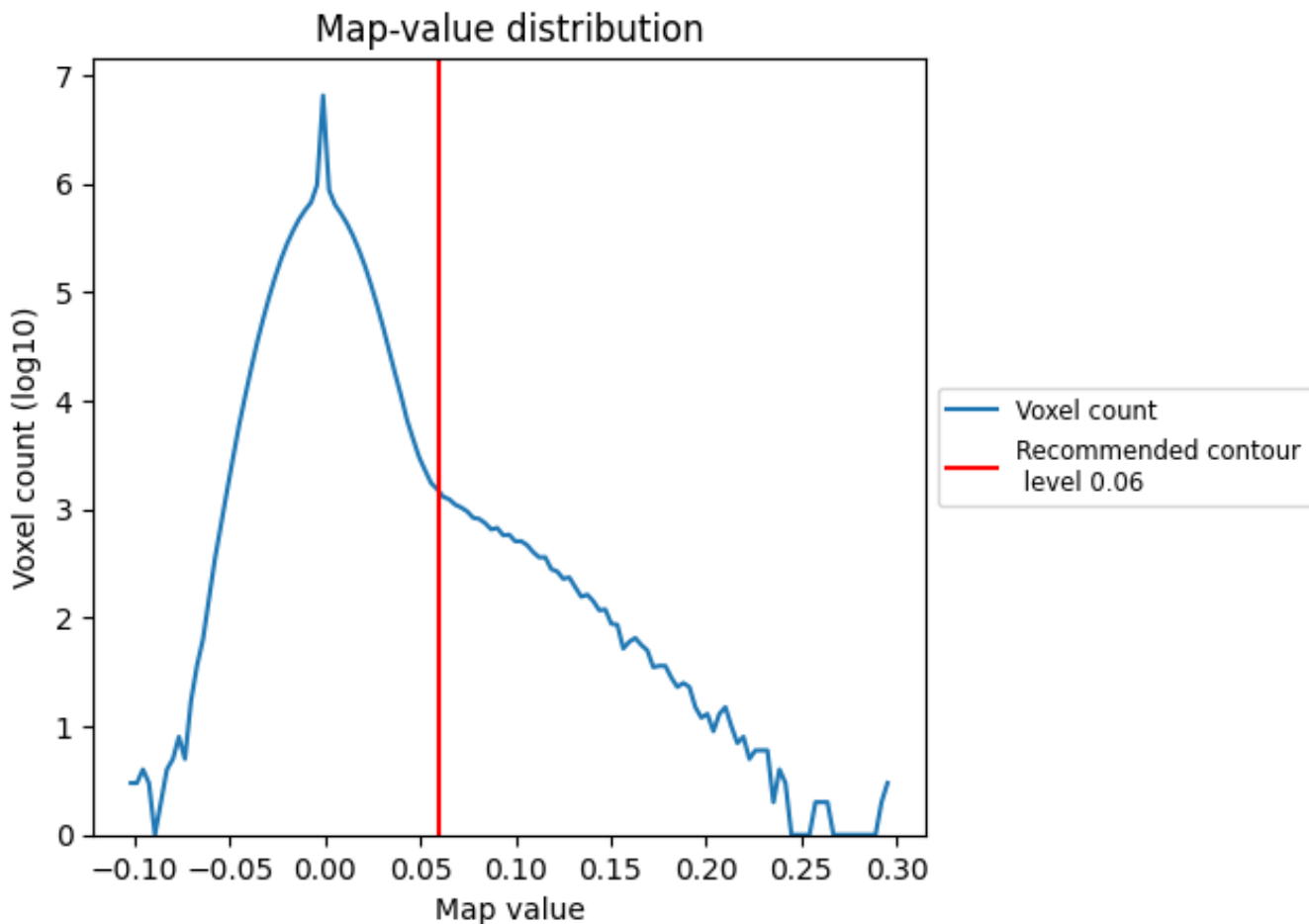
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

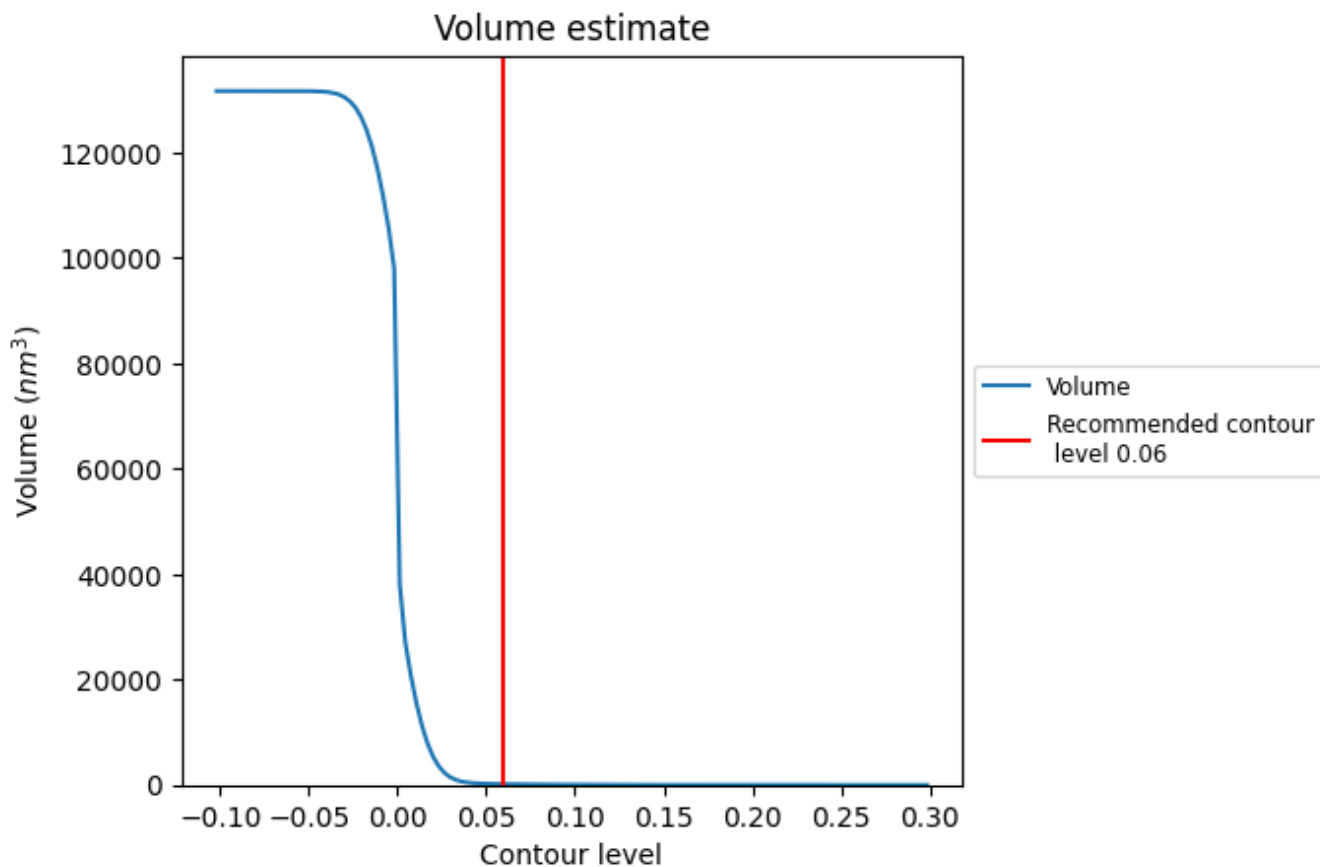
This section contains the results of statistical analysis of the map.

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.

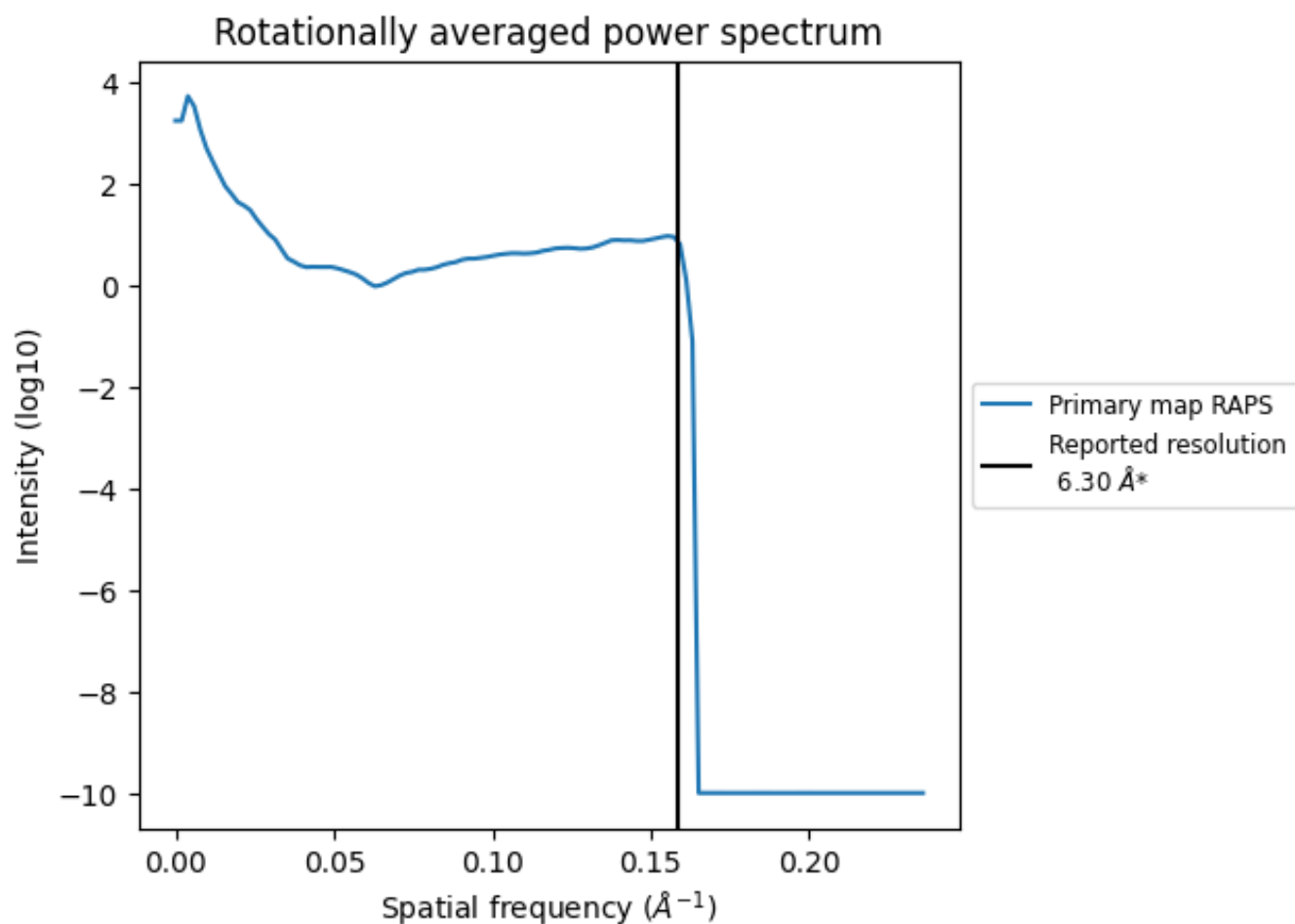
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 159 nm^3 ; this corresponds to an approximate mass of 144 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 0.159\AA^{-1}

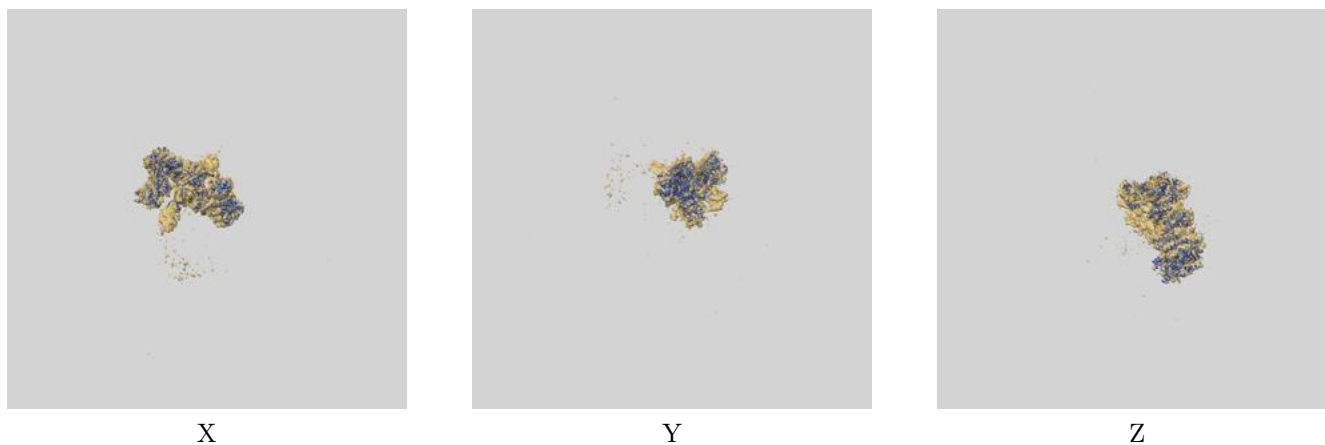
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

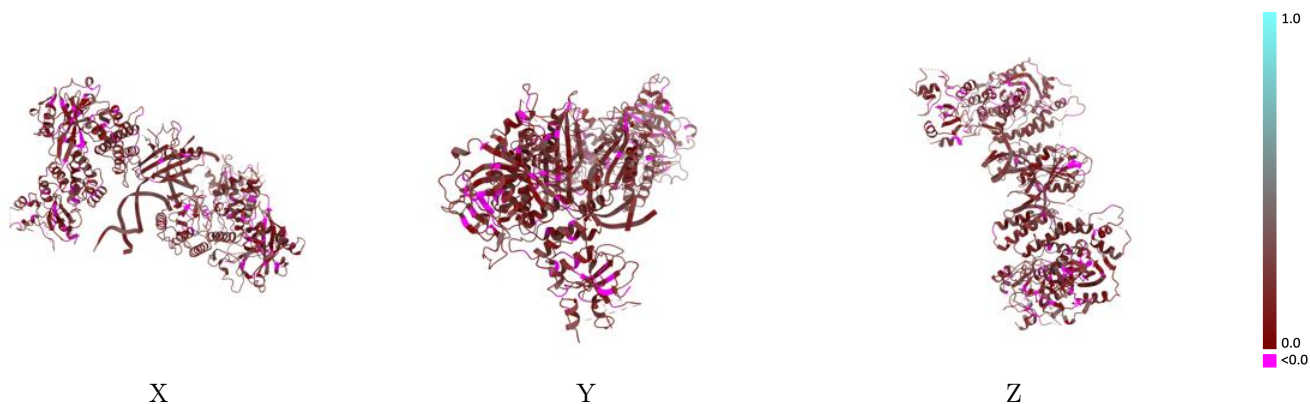
This section contains information regarding the fit between EMDB map EMD-21979 and PDB model 6X0M. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay [i](#)



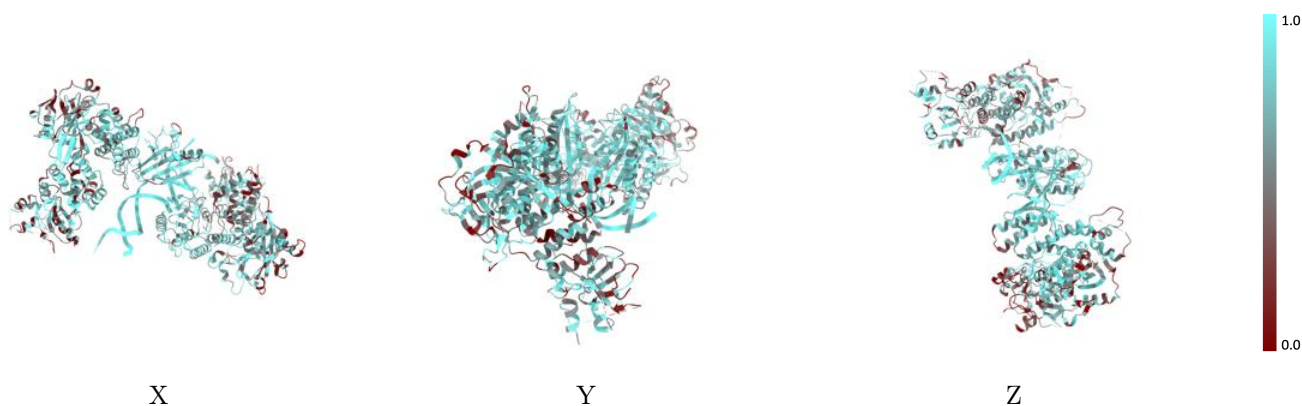
The images above show the 3D surface view of the map at the recommended contour level 0.06 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



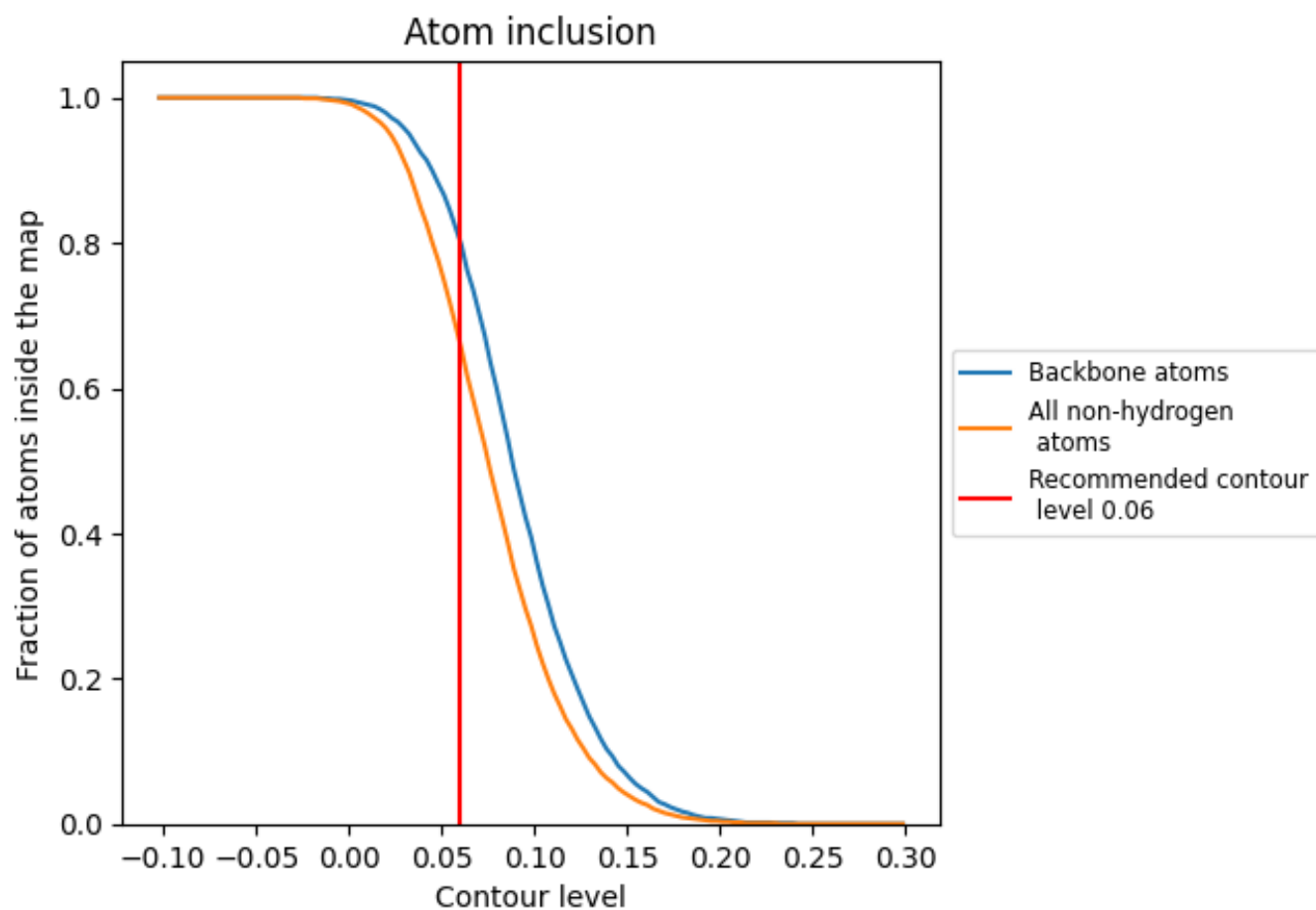
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.06).



















9.4 Atom inclusion [i](#)



At the recommended contour level, 80% of all backbone atoms, 66% 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 (0.06) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6610	 0.1670
I	 0.9420	 0.2600
J	 0.9120	 0.2290
O	 0.5970	 0.1620
P	 0.6800	 0.1660
i	 0.9260	 0.2570
j	 0.9150	 0.2210
o	 0.5650	 0.1510
p	 0.6600	 0.1610

