



wwPDB EM Validation Summary Report ⓘ

Nov 22, 2021 – 10:16 am GMT

PDB ID : 7PW4
EMDB ID : EMD-13674
Title : Human SMG1-8-9 kinase complex bound to a SMG1 inhibitor
Authors : Langer, L.M.; Conti, E.
Deposited on : 2021-10-06
Resolution : 3.27 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : 0.0.0.dev97
Mogul : 1.8.4 (270009), CSD as541be (2020)
MolProbity : 4.02b-467
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.23.2

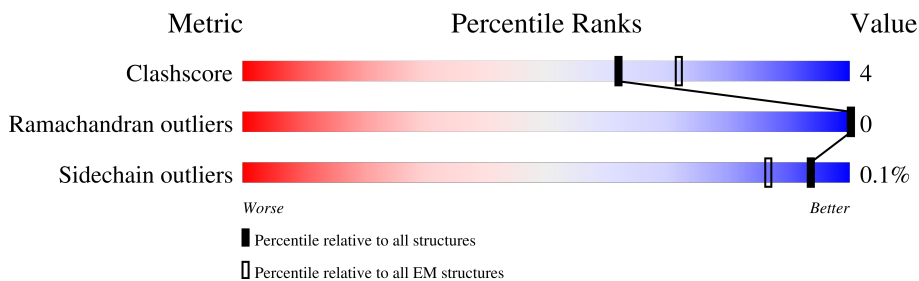
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.27 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	3657	
2	B	991	
3	C	520	

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 18965 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 Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1944	13719	8801	2444	2411	63	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	743	ARG	LYS	conflict	UNP Q96Q15
A	1209	SER	ALA	conflict	UNP Q96Q15

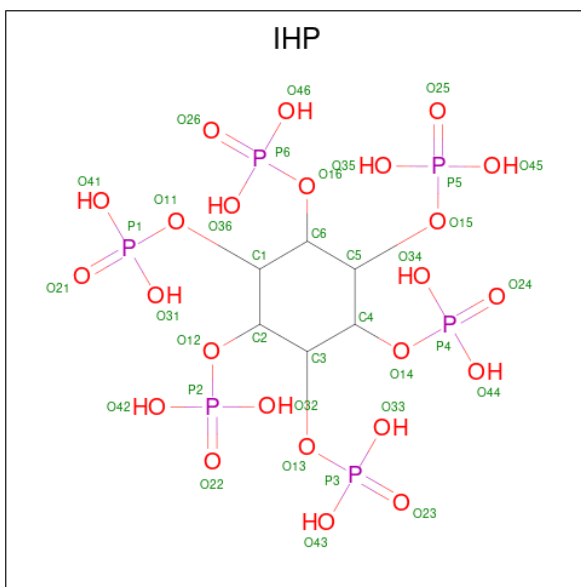
- Molecule 2 is a protein called Protein SMG8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	376	2718	1770	486	449	13	0	0

- Molecule 3 is a protein called Protein SMG9.

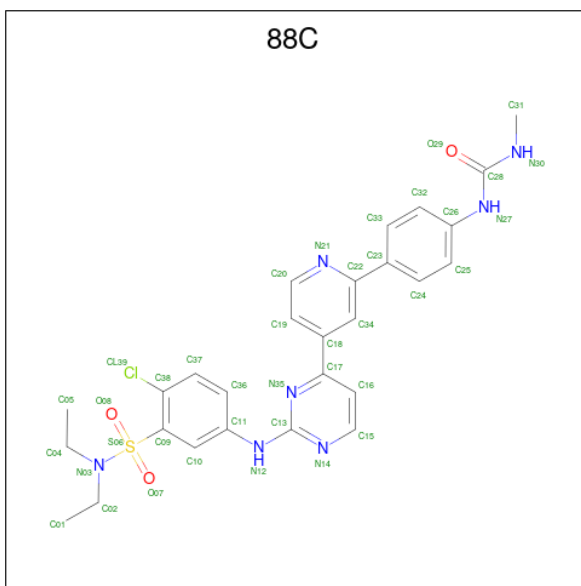
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	310	2421	1574	417	413	17	0	0

- Molecule 4 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: C₆H₁₈O₂₄P₆).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
4	A	1	36	6	24	6	0

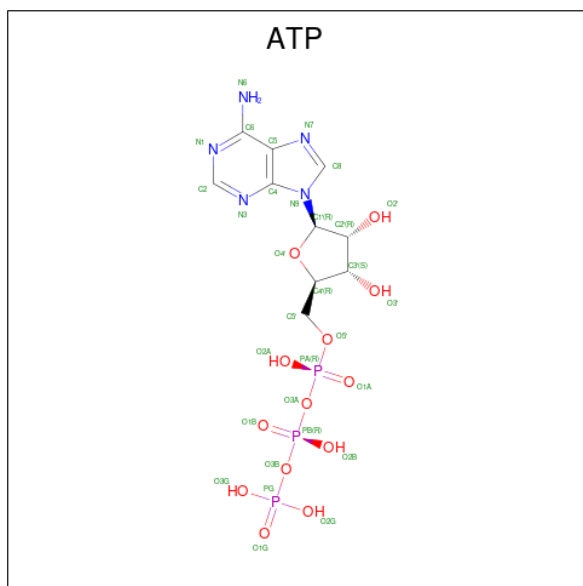
- Molecule 5 is 1-[4-[4-[2-[[4-chloranyl-3-(diethylsulfamoyl)phenyl]amino]pyrimidin-4-yl]pyridin-2-yl]phenyl]-3-methyl-urea (three-letter code: 88C) (formula: $C_{27}H_{28}ClN_7O_3S$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					AltConf	
			Total	C	Cl	N	O		S
5	A	1	39	27	1	7	3	1	0

- Molecule 6 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:

C₁₀H₁₆N₅O₁₃P₃).



Mol	Chain	Residues	Atoms				AltConf	
			Total	C	N	O		P
6	C	1	31	10	5	13	3	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
7	C	1	1	1	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: Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1, Serine/threonine-protein kinase SMG1

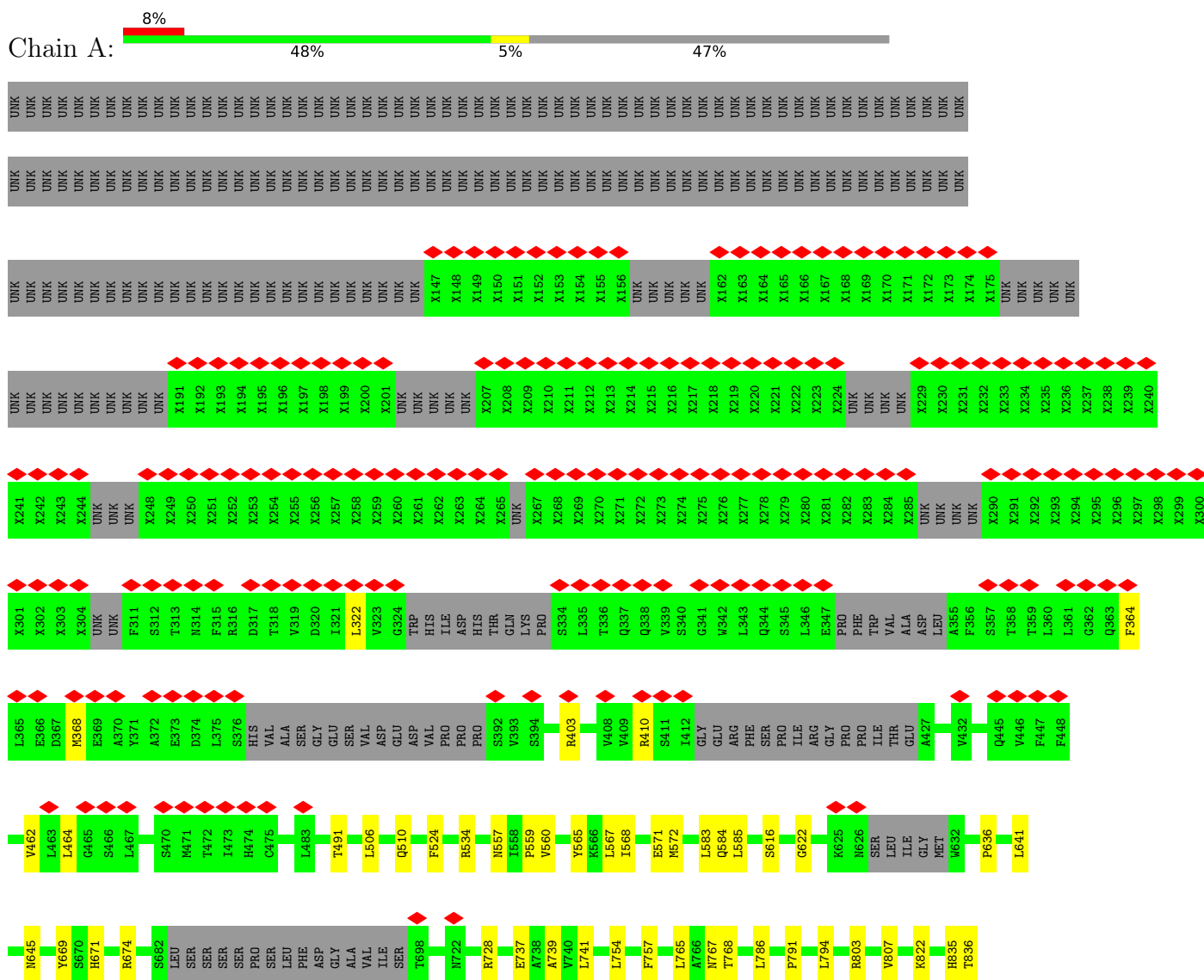
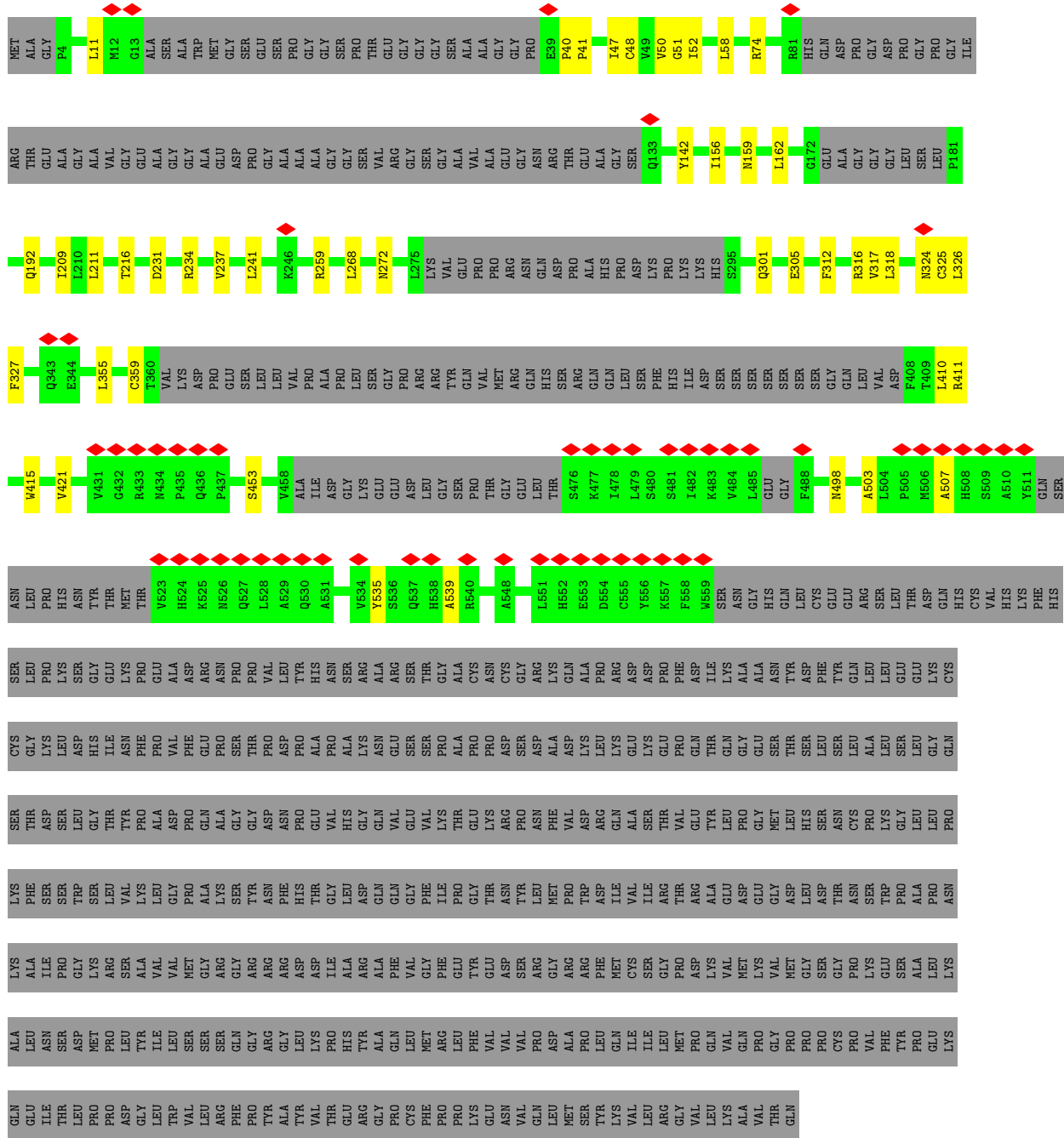


Table of amino acid sequences for Chain B, containing 15 columns of residues.

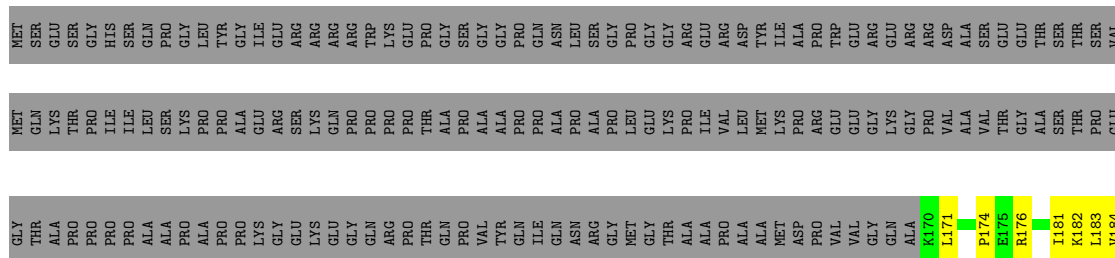
Vertical legend for residue types: R3623, D3624, V3625, N3628, R3629, R3630, A3644, E3655, V3661.

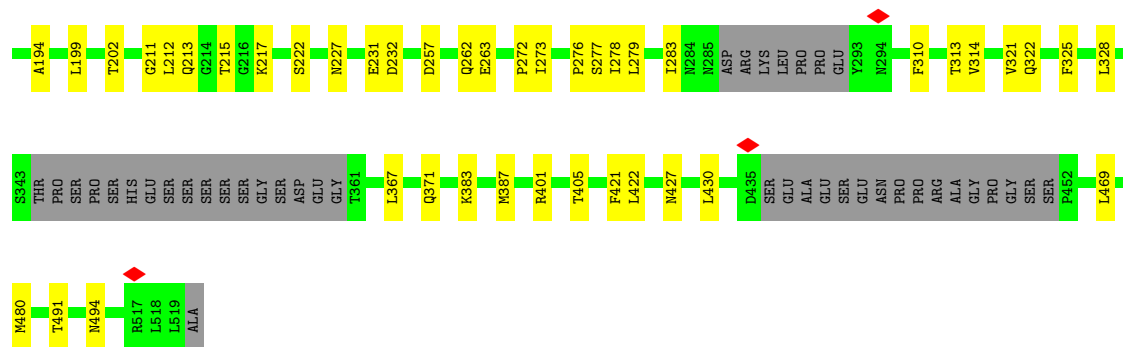
• Molecule 2: Protein SMG8





• Molecule 3: Protein SMG9





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	228291	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 ($e^-/\text{\AA}^2$)	89.32	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	3.406	Depositor
Minimum map value	-1.579	Depositor
Average map value	0.005	Depositor
Map value standard deviation	0.076	Depositor
Recommended contour level	0.599	Depositor
Map size (Å)	326.86078, 326.86078, 326.86078	wwPDB
Map dimensions	384, 384, 384	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.8511999, 0.8511999, 0.8511999	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: IHP, ATP, 88C, MG

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	0.25	0/12988	0.46	1/17732 (0.0%)
2	B	0.24	0/2776	0.45	0/3781
3	C	0.24	0/2479	0.50	0/3365
All	All	0.25	0/18243	0.46	1/24878 (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	A	1392	PRO	CA-N-CD	-5.49	103.82	111.50

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	13719	0	12364	106	0
2	B	2718	0	2505	29	0
3	C	2421	0	2373	34	0
4	A	36	0	6	0	0
5	A	39	0	0	0	0
6	C	31	0	12	1	0
7	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	18965	0	17260	162	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 4.

The worst 5 of 162 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:1741:PHE:HE1	1:A:1803:LEU:HD23	1.56	0.71
1:A:1126:GLN:OE1	1:A:1194:TYR:OH	2.10	0.69
3:C:202:THR:HG22	3:C:491:THR:HG22	1.74	0.69
1:A:1849:ALA:HB2	1:A:1856:ILE:HD11	1.75	0.68
1:A:2131:LEU:HB2	1:A:2137:PRO:HG2	1.74	0.68

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	1695/3657 (46%)	1657 (98%)	38 (2%)	0	100	100
2	B	358/991 (36%)	350 (98%)	8 (2%)	0	100	100
3	C	302/520 (58%)	283 (94%)	19 (6%)	0	100	100
All	All	2355/5168 (46%)	2290 (97%)	65 (3%)	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	1205/2797 (43%)	1205 (100%)	0	100	100
2	B	240/847 (28%)	239 (100%)	1 (0%)	91	95
3	C	253/450 (56%)	252 (100%)	1 (0%)	91	95
All	All	1698/4094 (42%)	1696 (100%)	2 (0%)	93	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	498	ASN
3	C	401	ARG

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

Mol	Chain	Res	Type
1	A	2337	HIS
2	B	159	ASN

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](#)

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

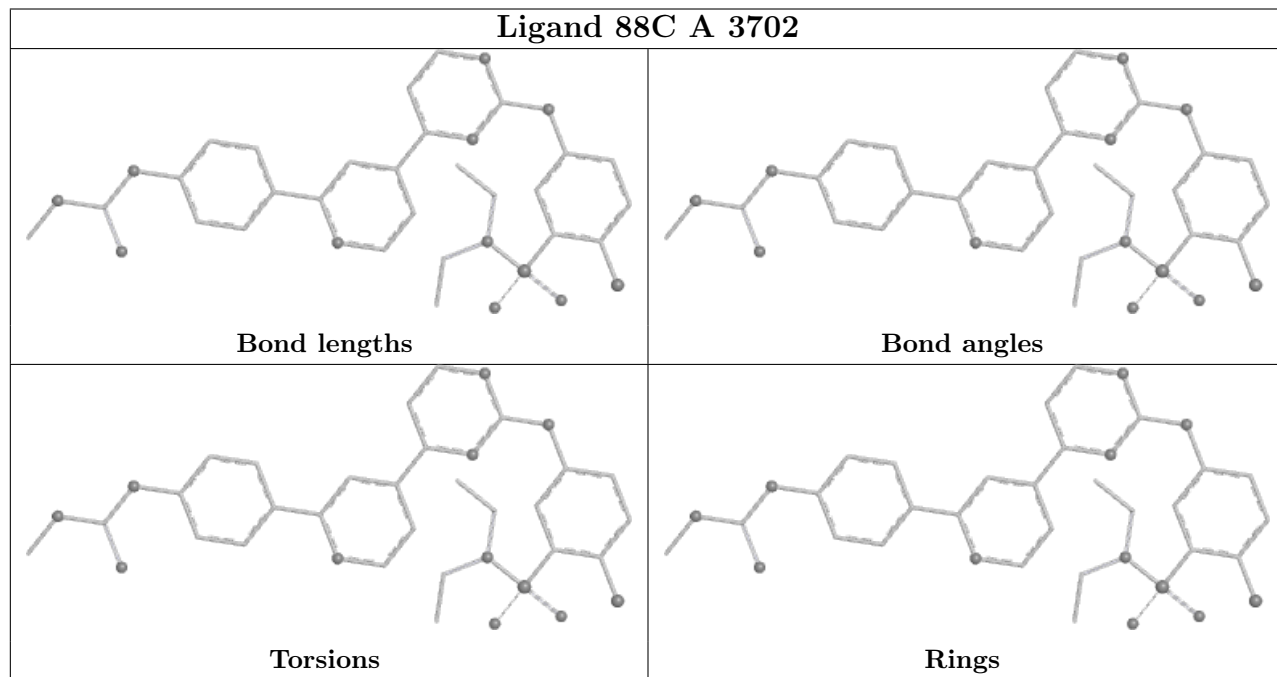
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 than 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.



5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

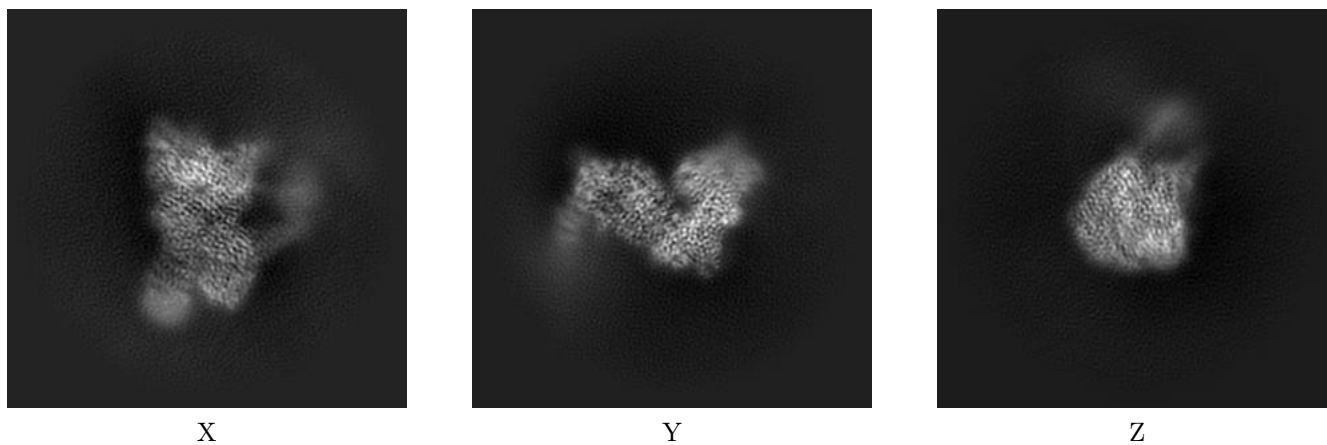
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-13674. 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](#)

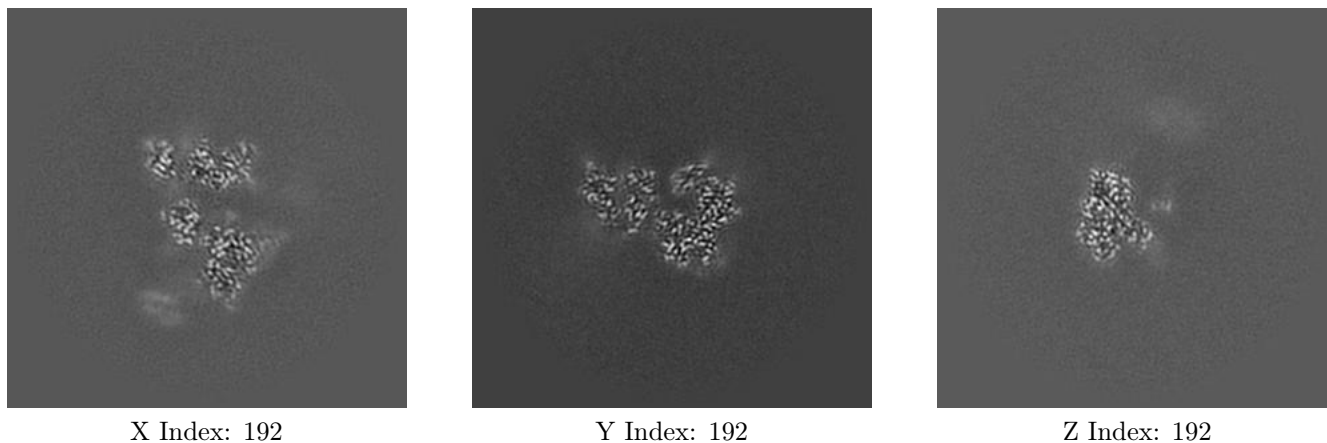
6.1.1 Primary map



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

6.2 Central slices [i](#)

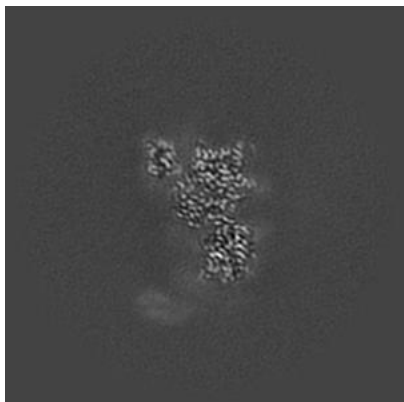
6.2.1 Primary map



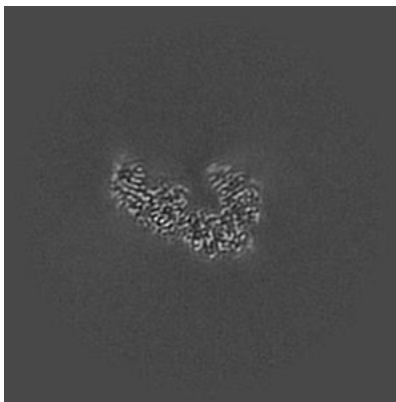
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

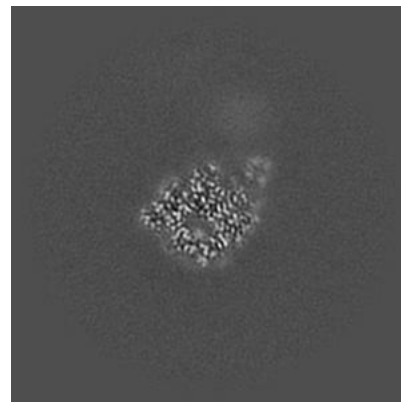
6.3.1 Primary map



X Index: 180



Y Index: 203



Z Index: 225

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

6.4 Orthogonal surface views [i](#)

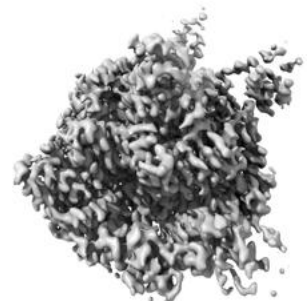
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.599. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

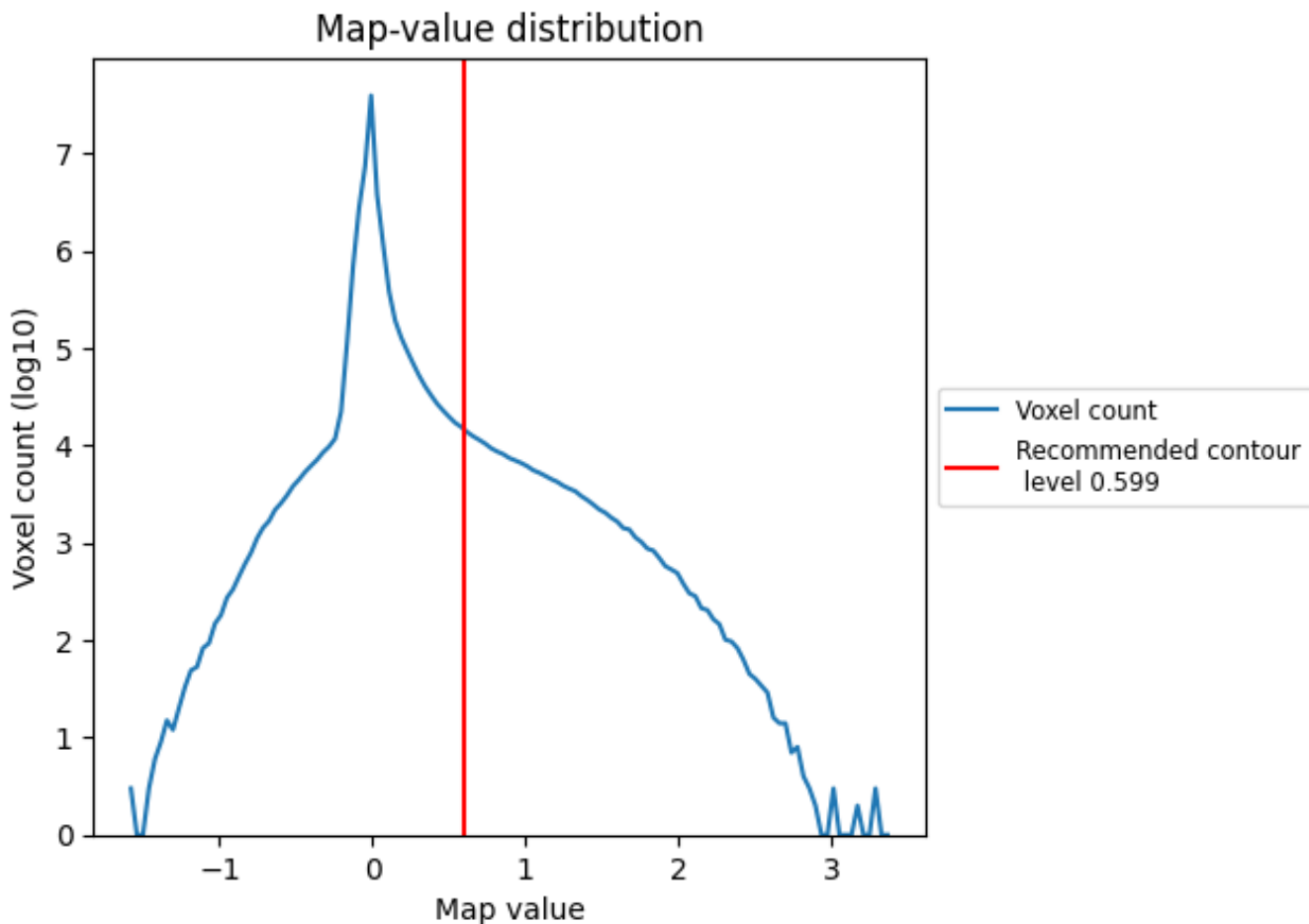
6.5 Mask visualisation

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

7 Map analysis [i](#)

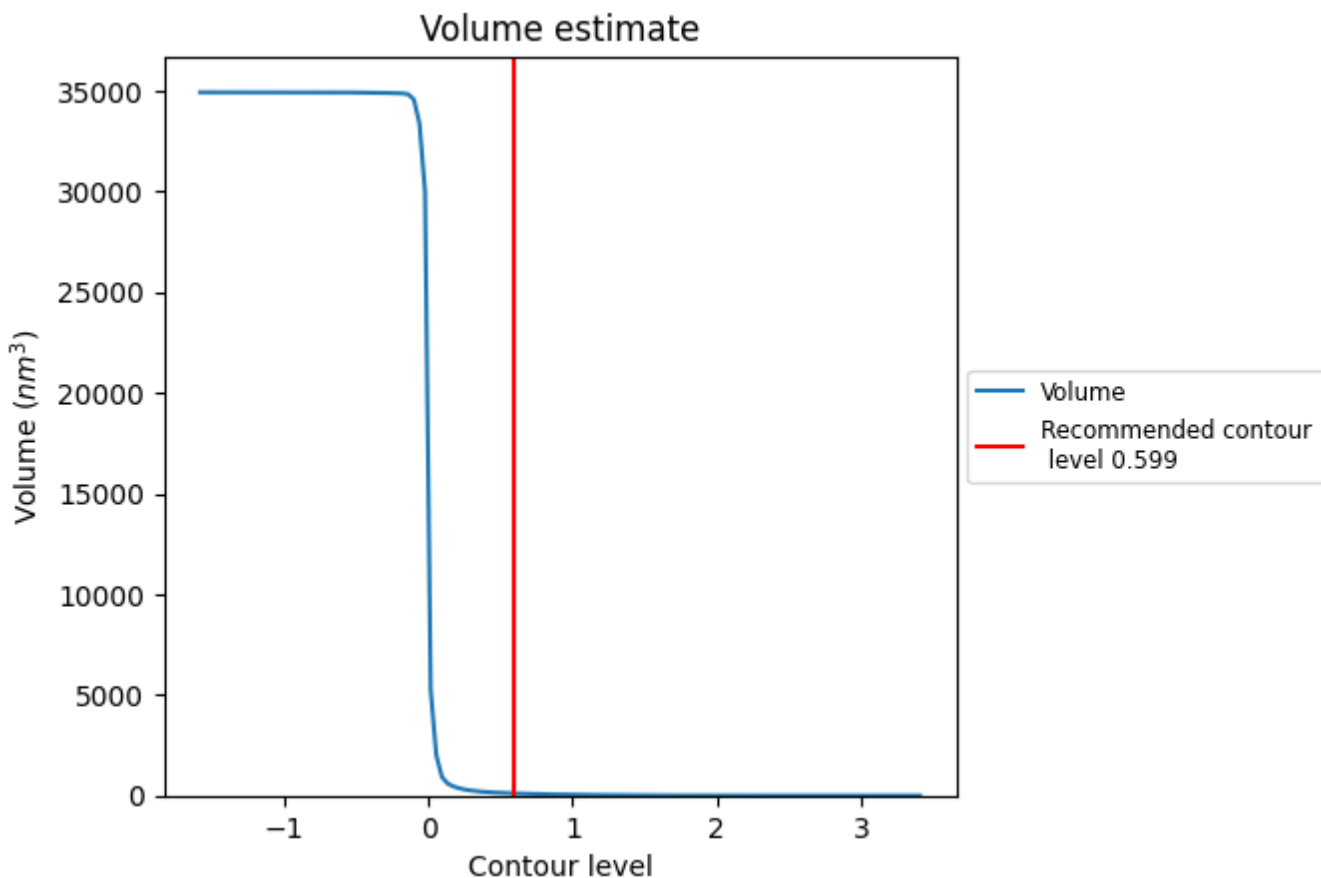
This section contains the results of statistical analysis of the map.

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.

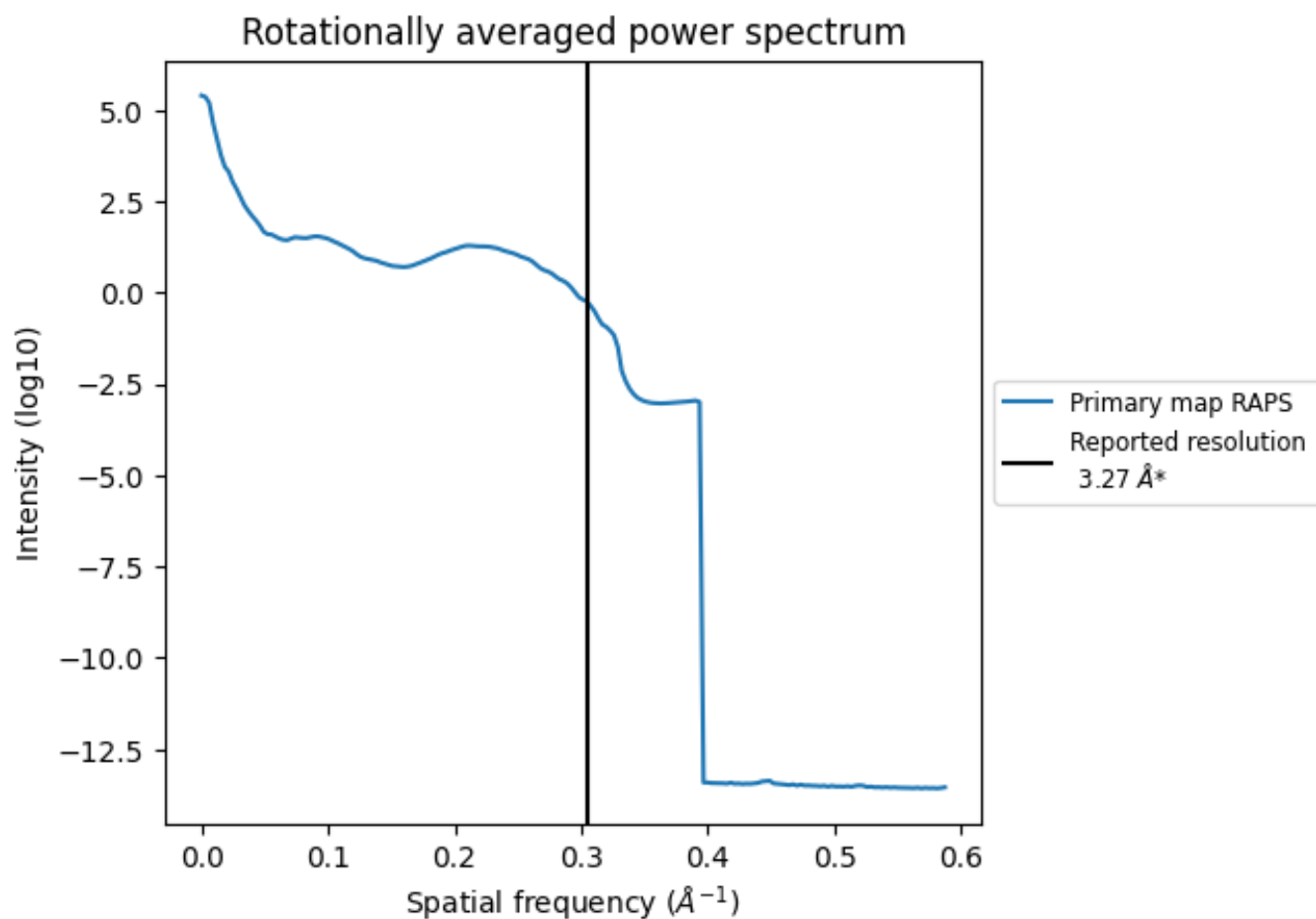
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 106 nm³; this corresponds to an approximate mass of 96 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 0.306 Å⁻¹

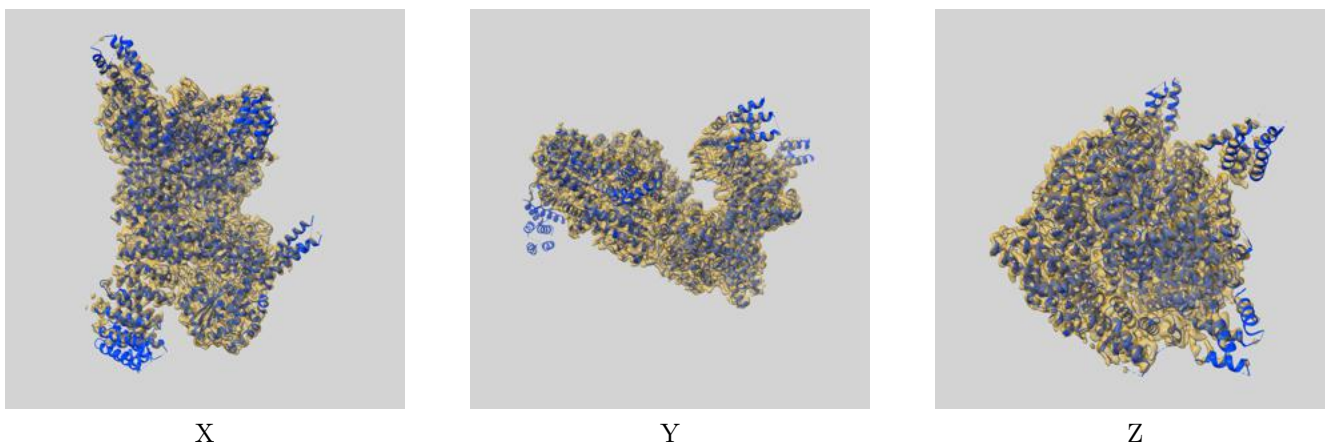
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

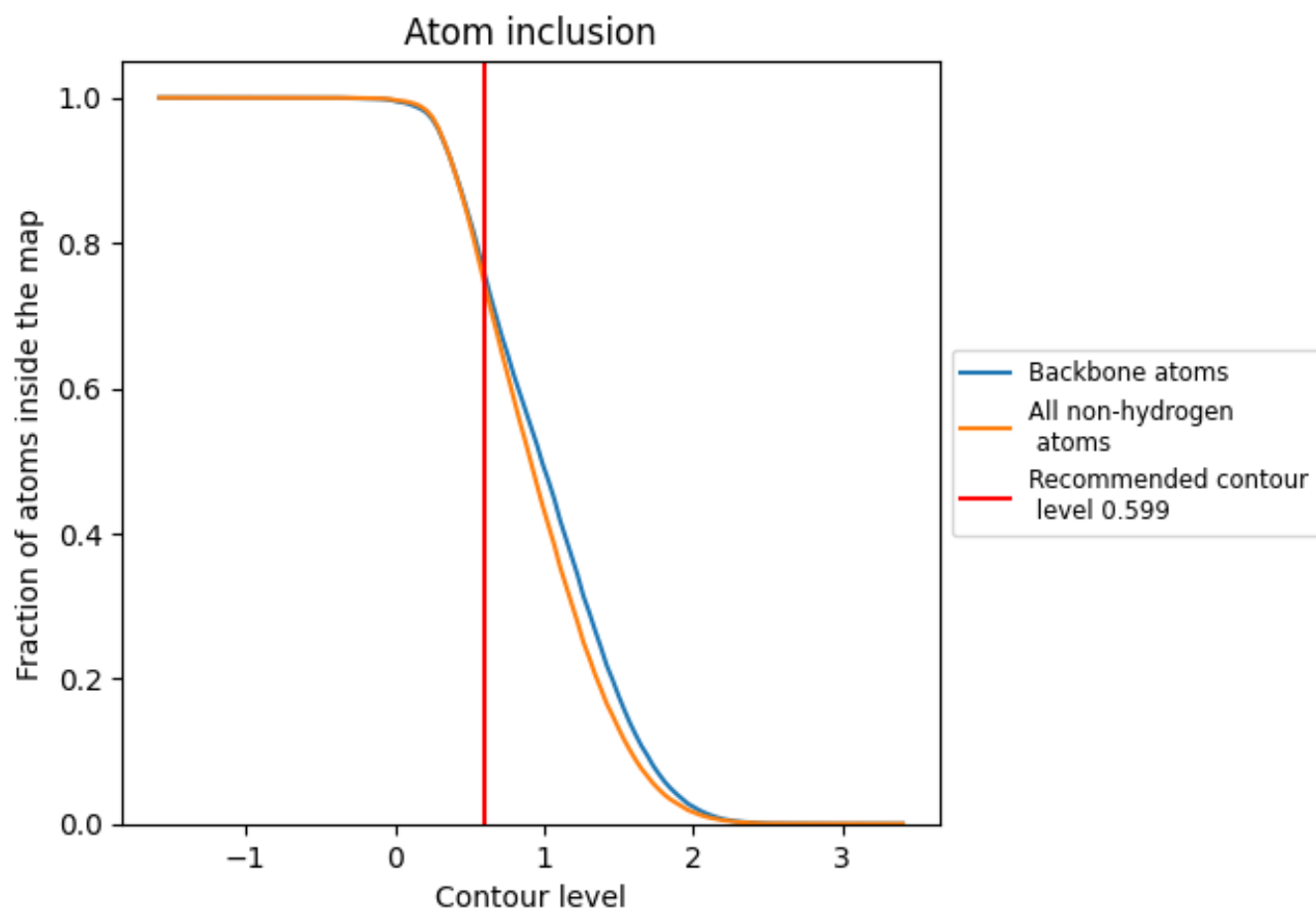
This section contains information regarding the fit between EMDB map EMD-13674 and PDB model 7PW4. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 0.599 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 Atom inclusion [i](#)



At the recommended contour level, 76% of all backbone atoms, 74% of all non-hydrogen atoms, are inside the map.