Full wwPDB EM Validation Report (i)

Dec 11, 2022 – 04:43 am GMT

PDB ID	:	4D1K
EMDB ID	:	EMD-2638
Title	:	Cryo-electron microscopy of tubular arrays of HIV-1 Gag resolves structures essential for immature virus assembly.
Authors	:	Bharat, T.A.M.; Castillo-Menendez, L.R.; Hagen, W.J.H.; Lux, V.; Igonet, S.; Schorb, M.; Schur, F.K.M.; Krauesslich, H.G.; Briggs, J.A.G.
Deposited on	:	2014-05-02
Resolution	:	9.40 Å(reported)
Based on initial model	:	4COP

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 (i) 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 (i)) were used in the production of this report:

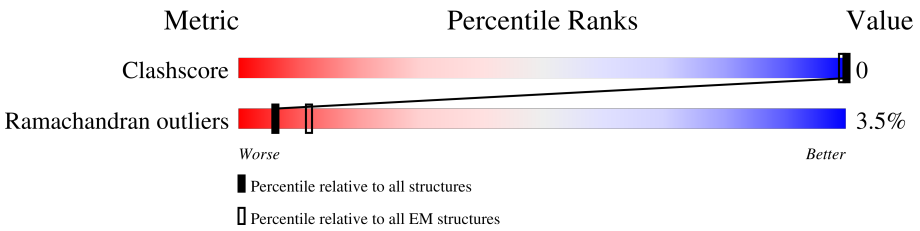
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 9.40 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

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	219	33% 94% 5%
1	B	219	40% 97% .
1	C	219	44% 97% .
1	D	219	37% 95% 5%
1	E	219	41% 95% 5%
1	F	219	42% 95% 5%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3942 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 GAG PROTEIN.

Mol	Chain	Residues	Atoms			AltConf	Trace
			Total	C	N		
1	A	219	657	438	219	0	0
1	B	219	657	438	219	0	0
1	C	219	657	438	219	0	0
1	D	219	657	438	219	0	0
1	E	219	657	438	219	0	0
1	F	219	657	438	219	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	ILE	LEU	variant	UNP Q5D0H3
A	87	HIS	GLN	variant	UNP Q5D0H3
A	120	HIS	SER	variant	UNP Q5D0H3
A	169	SER	TYR	engineered mutation	UNP Q5D0H3
B	15	ILE	LEU	variant	UNP Q5D0H3
B	87	HIS	GLN	variant	UNP Q5D0H3
B	120	HIS	SER	variant	UNP Q5D0H3
B	169	SER	TYR	engineered mutation	UNP Q5D0H3
C	15	ILE	LEU	variant	UNP Q5D0H3
C	87	HIS	GLN	variant	UNP Q5D0H3
C	120	HIS	SER	variant	UNP Q5D0H3
C	169	SER	TYR	engineered mutation	UNP Q5D0H3
D	15	ILE	LEU	variant	UNP Q5D0H3
D	87	HIS	GLN	variant	UNP Q5D0H3
D	120	HIS	SER	variant	UNP Q5D0H3
D	169	SER	TYR	engineered mutation	UNP Q5D0H3
E	15	ILE	LEU	variant	UNP Q5D0H3
E	87	HIS	GLN	variant	UNP Q5D0H3

Continued on next page...

Continued from previous page...

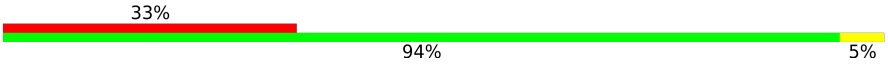
Chain	Residue	Modelled	Actual	Comment	Reference
E	120	HIS	SER	variant	UNP Q5D0H3
E	169	SER	TYR	engineered mutation	UNP Q5D0H3
F	15	ILE	LEU	variant	UNP Q5D0H3
F	87	HIS	GLN	variant	UNP Q5D0H3
F	120	HIS	SER	variant	UNP Q5D0H3
F	169	SER	TYR	engineered mutation	UNP Q5D0H3

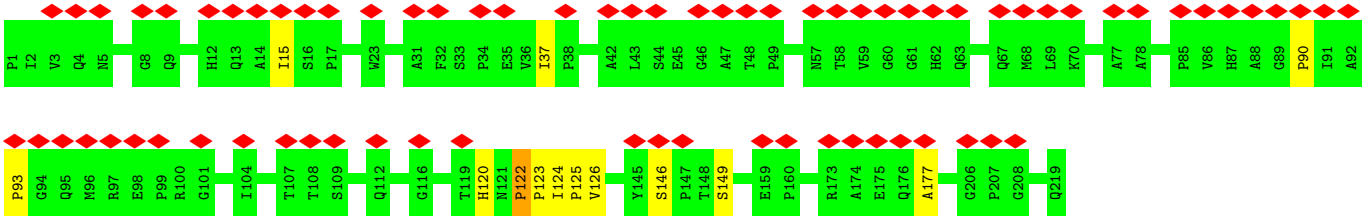
3 Residue-property plots (i)

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: GAG PROTEIN

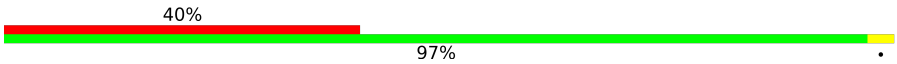
Chain A:

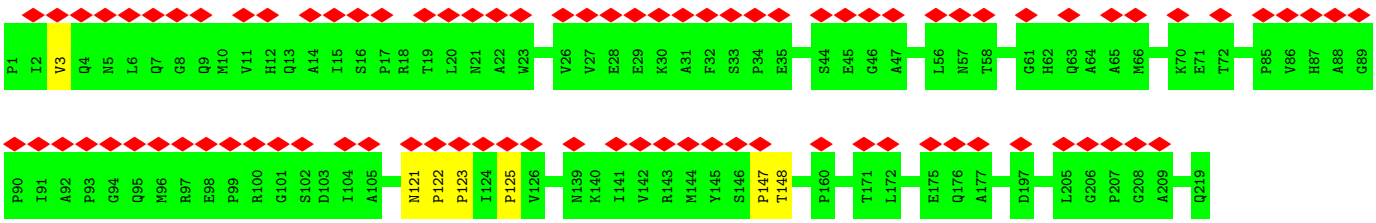




- Molecule 1: GAG PROTEIN

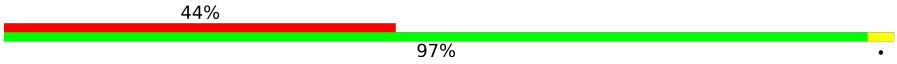
Chain B:

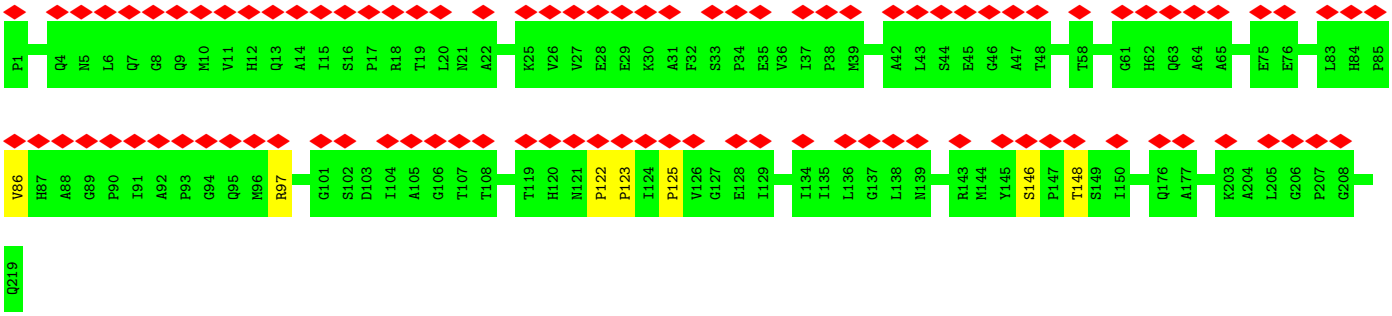




- Molecule 1: GAG PROTEIN

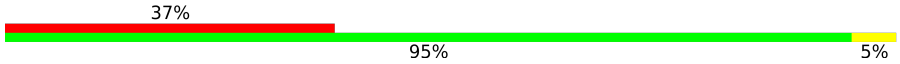
Chain C:

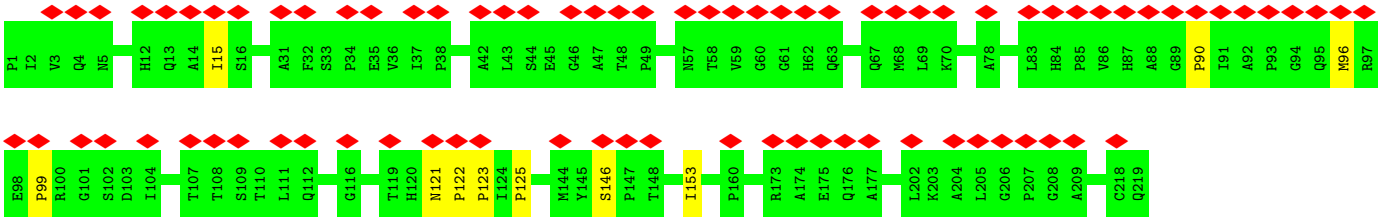




- Molecule 1: GAG PROTEIN

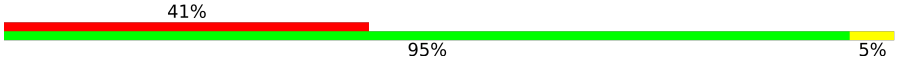
Chain D:

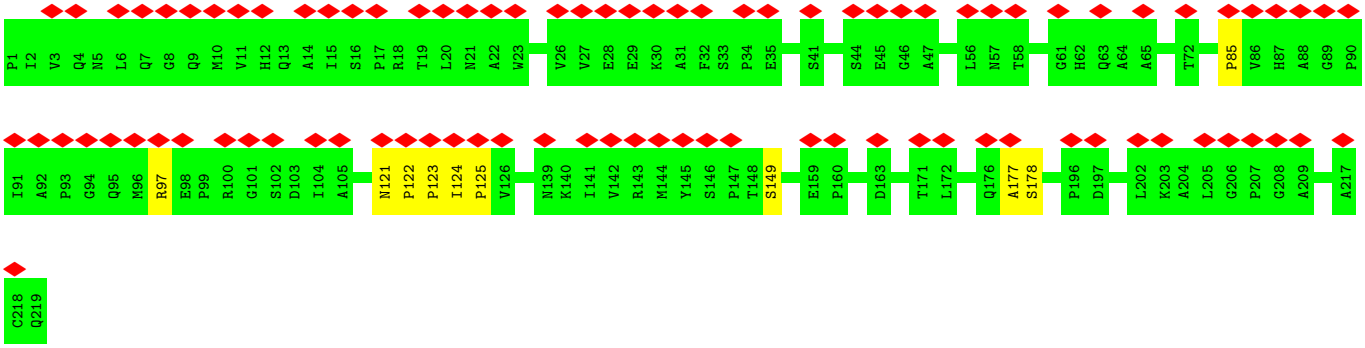




- Molecule 1: GAG PROTEIN

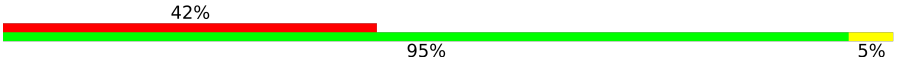
Chain E:

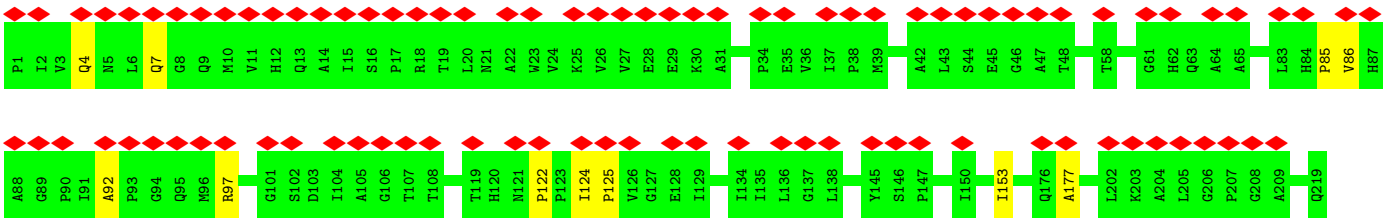




- Molecule 1: GAG PROTEIN

Chain F:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=Not provided°, rise=Not provided Å, axial sym=Not provided	Depositor
Number of segments used	Not provided	
Resolution determination method	Not provided	
CTF correction method	Not provided	
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	23	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4500	Depositor
Magnification	Not provided	
Image detector	KODAK SO-163 FILM	Depositor
Maximum map value	86.177	Depositor
Minimum map value	-69.385	Depositor
Average map value	2.531	Depositor
Map value standard deviation	19.957	Depositor
Recommended contour level	27.0	Depositor
Map size (Å)	153.0, 153.0, 153.0	wwPDB
Map dimensions	100, 100, 100	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.53, 1.53, 1.53	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	1.51	3/656 (0.5%)	1.46	1/655 (0.2%)
1	B	1.44	0/656	1.44	0/655
1	C	1.46	0/656	1.46	0/655
1	D	1.47	0/656	1.39	0/655
1	E	1.49	0/656	1.43	0/655
1	F	1.46	1/656 (0.2%)	1.45	0/655
All	All	1.47	4/3936 (0.1%)	1.44	1/3930 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
1	D	0	1
1	E	0	2
1	F	0	2
All	All	0	8

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	93	PRO	CA-C	5.87	1.64	1.52
1	F	85	PRO	CA-C	-5.52	1.41	1.52
1	A	37	ILE	C-N	5.52	1.44	1.34
1	A	126	VAL	C-N	5.13	1.42	1.33

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	122	PRO	N-CA-C	-5.15	98.71	112.10

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	124	ILE	Peptide
1	B	121	ASN	Peptide
1	B	3	VAL	Peptide
1	D	121	ASN	Peptide
1	E	121	ASN	Peptide
1	E	124	ILE	Peptide
1	F	124	ILE	Peptide
1	F	4	GLN	Peptide

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	657	0	230	0	0
1	B	657	0	230	0	0
1	C	657	0	230	0	0
1	D	657	0	230	0	0
1	E	657	0	230	0	0
1	F	657	0	230	0	0
All	All	3942	0	1380	0	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 0.

There are no clashes within the asymmetric unit.

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	217/219 (99%)	200 (92%)	8 (4%)	9 (4%)	3	23
1	B	217/219 (99%)	205 (94%)	7 (3%)	5 (2%)	6	34
1	C	217/219 (99%)	202 (93%)	8 (4%)	7 (3%)	4	26
1	D	217/219 (99%)	202 (93%)	6 (3%)	9 (4%)	3	23
1	E	217/219 (99%)	205 (94%)	4 (2%)	8 (4%)	3	24
1	F	217/219 (99%)	201 (93%)	8 (4%)	8 (4%)	3	24
All	All	1302/1314 (99%)	1215 (93%)	41 (3%)	46 (4%)	6	25

All (46) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	PRO
1	A	120	HIS
1	A	122	PRO
1	A	123	PRO
1	A	146	SER
1	C	97	ARG
1	C	122	PRO
1	C	123	PRO
1	C	146	SER
1	D	15	ILE
1	D	122	PRO
1	D	123	PRO
1	D	125	PRO
1	E	123	PRO
1	F	97	ARG
1	F	122	PRO
1	A	15	ILE
1	A	177	ALA
1	C	148	THR
1	D	96	MET
1	D	146	SER
1	E	97	ARG
1	E	122	PRO
1	E	178	SER
1	F	125	PRO
1	F	177	ALA
1	A	125	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	122	PRO
1	B	147	PRO
1	E	125	PRO
1	E	149	SER
1	E	177	ALA
1	F	7	GLN
1	A	149	SER
1	B	148	THR
1	C	125	PRO
1	E	85	PRO
1	D	99	PRO
1	F	92	ALA
1	B	123	PRO
1	D	90	PRO
1	C	86	VAL
1	F	86	VAL
1	B	125	PRO
1	F	153	ILE
1	D	153	ILE

5.3.2 Protein sidechains (i)

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

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.

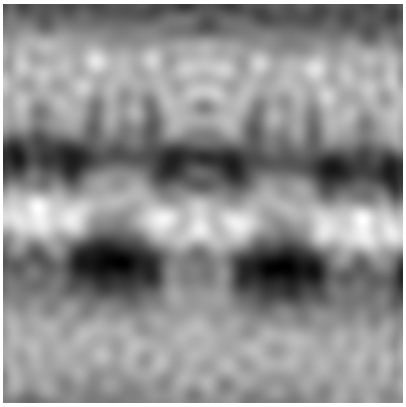
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-2638. These allow visual inspection of the internal detail of the map and identification of artifacts.

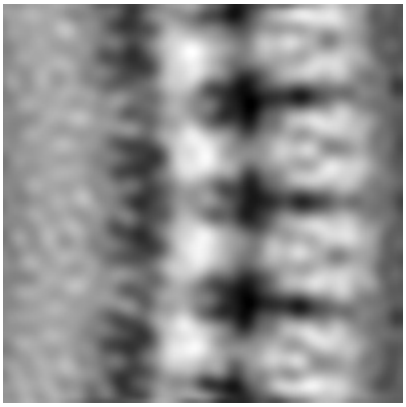
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

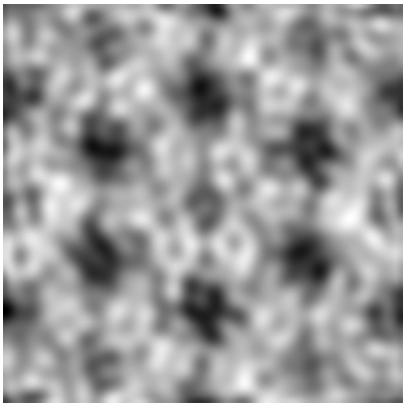
6.1.1 Primary map



X



Y



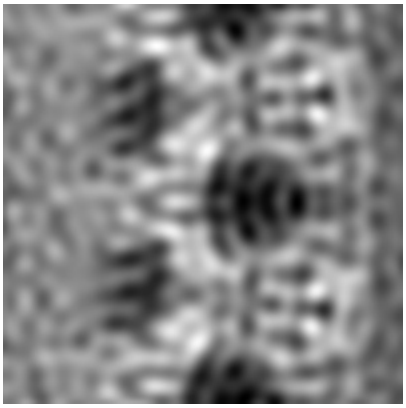
Z

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

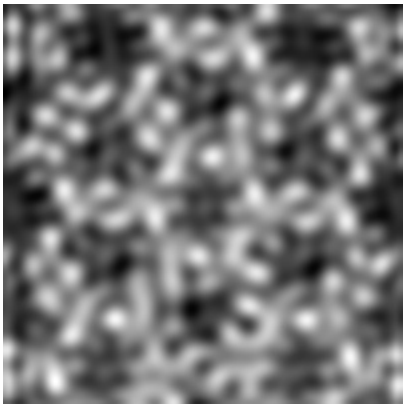
6.2 Central slices (i)

6.2.1 Primary map

X Index: 50



Y Index: 50

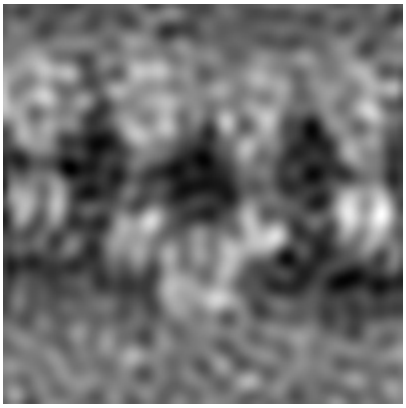


Z Index: 50

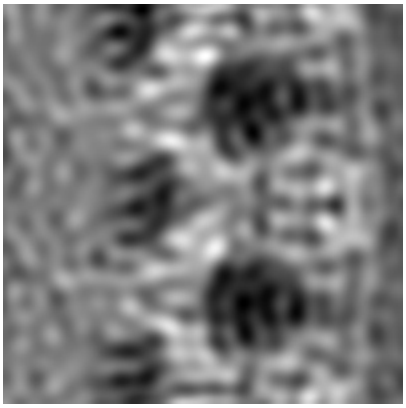
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

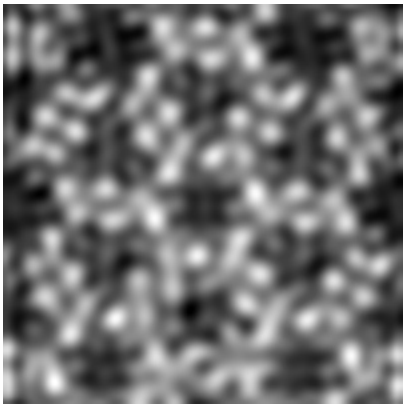
6.3.1 Primary map



X Index: 2



Y Index: 10

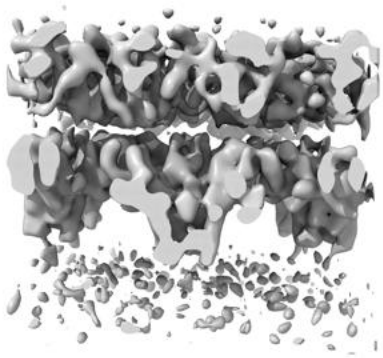


Z Index: 51

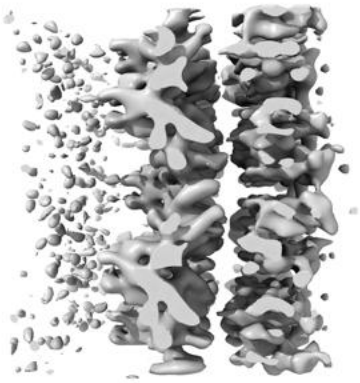
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

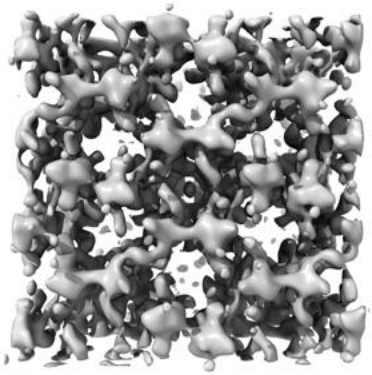
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 27.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

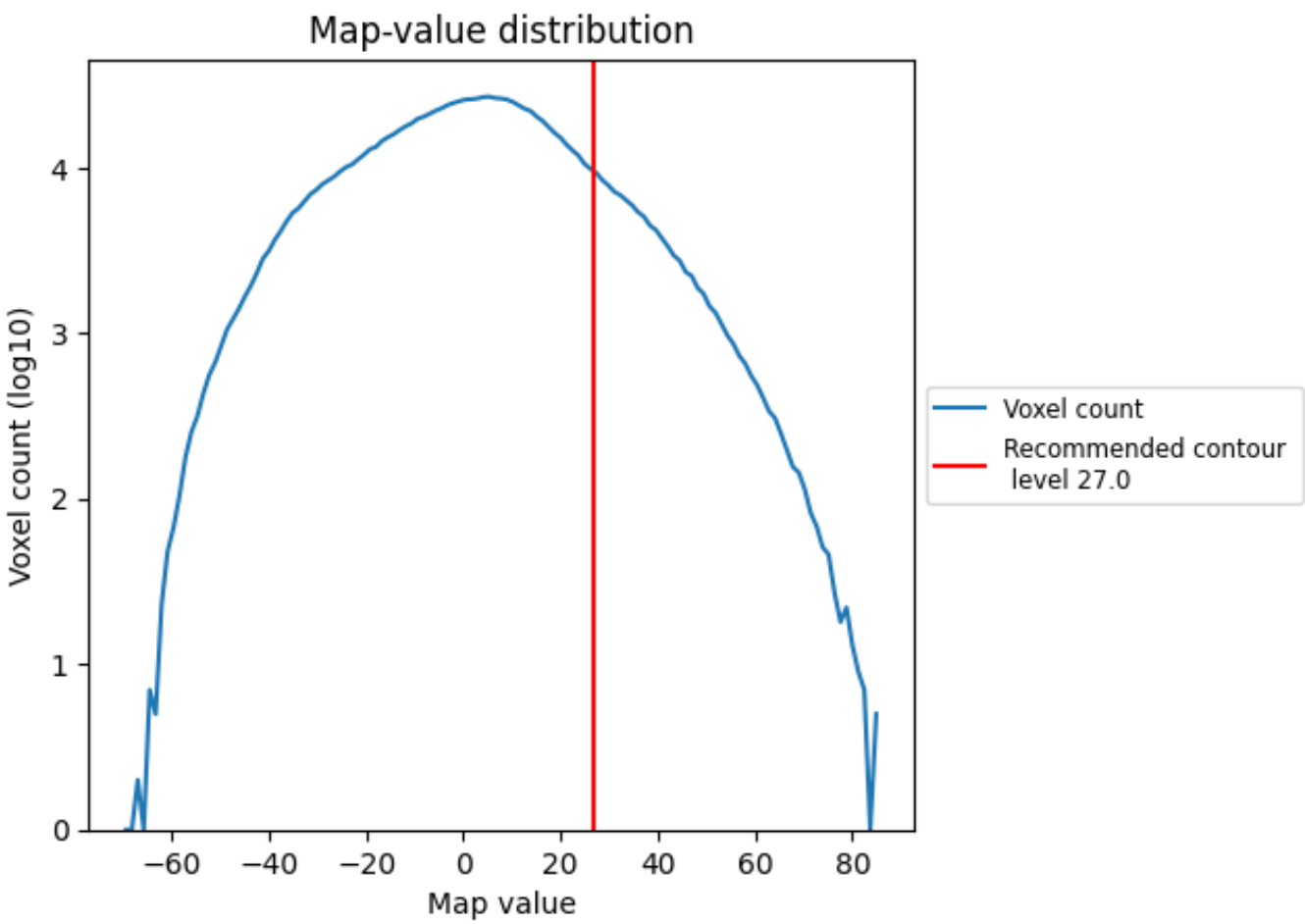
6.5 Mask visualisation (i)

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

7 Map analysis (i)

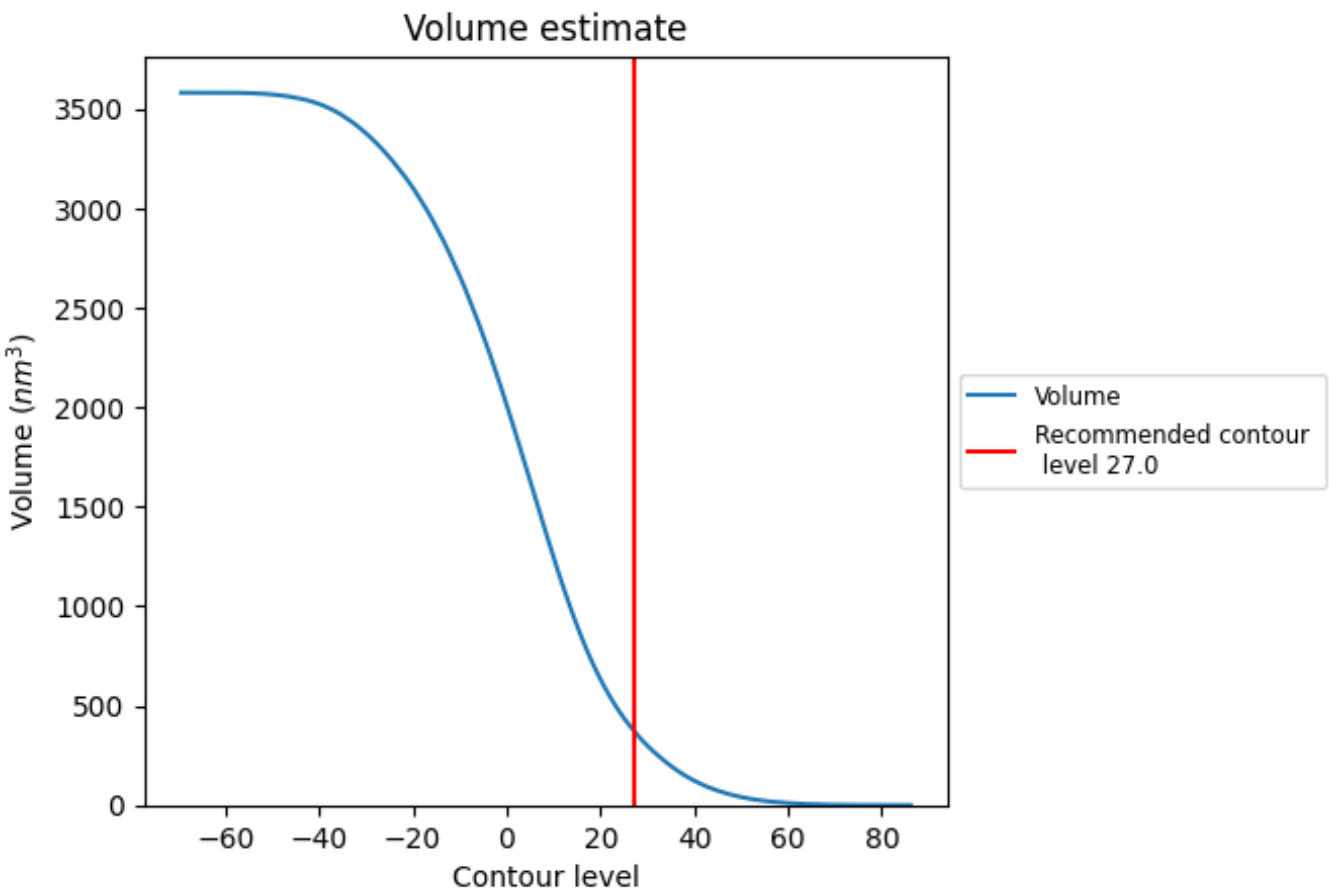
This section contains the results of statistical analysis of the map.

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.

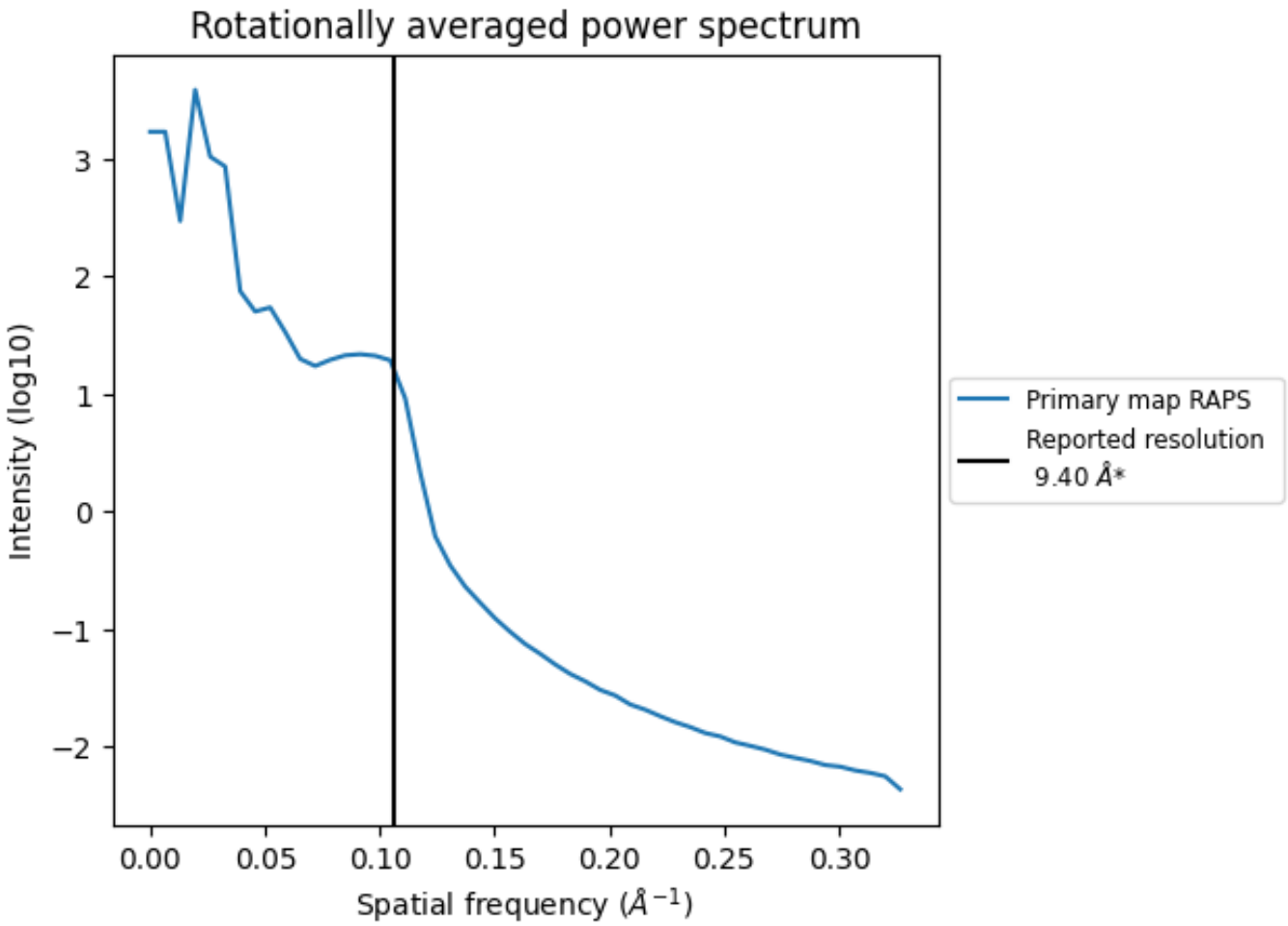
7.2 Volume estimate (i)



The volume at the recommended contour level is 378 nm^3; this corresponds to an approximate mass of 341 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 0.106 Å^{-1}

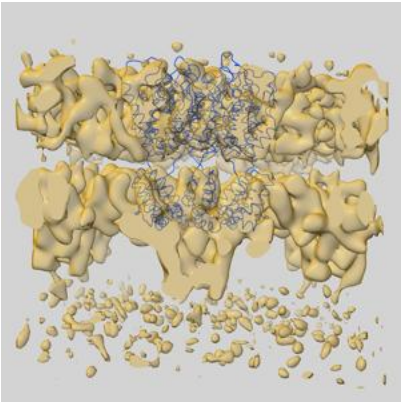
8 Fourier-Shell correlation (i)

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

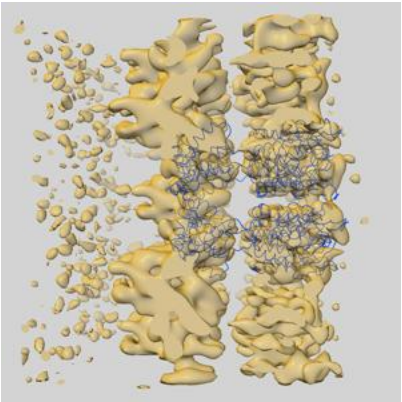
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-2638 and PDB model 4D1K. Per-residue inclusion information can be found in section 3 on page 5.

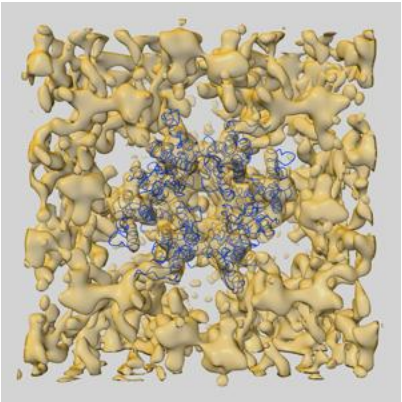
9.1 Map-model overlay (i)



X



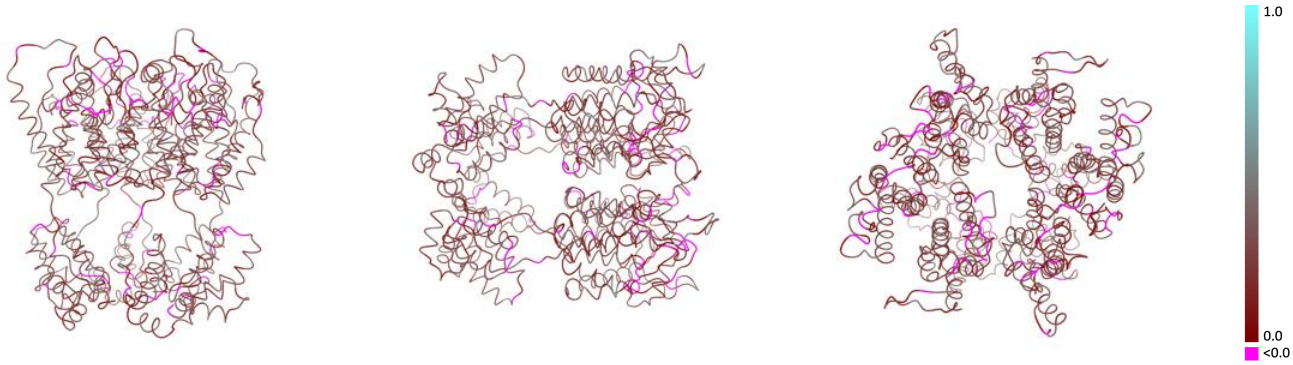
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 27.0 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.

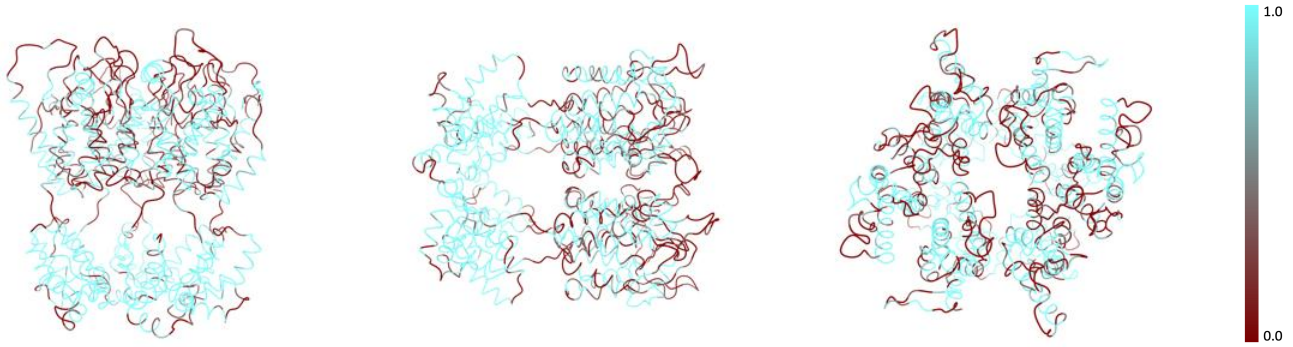
9.2 Q-score mapped to coordinate model (i)



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.

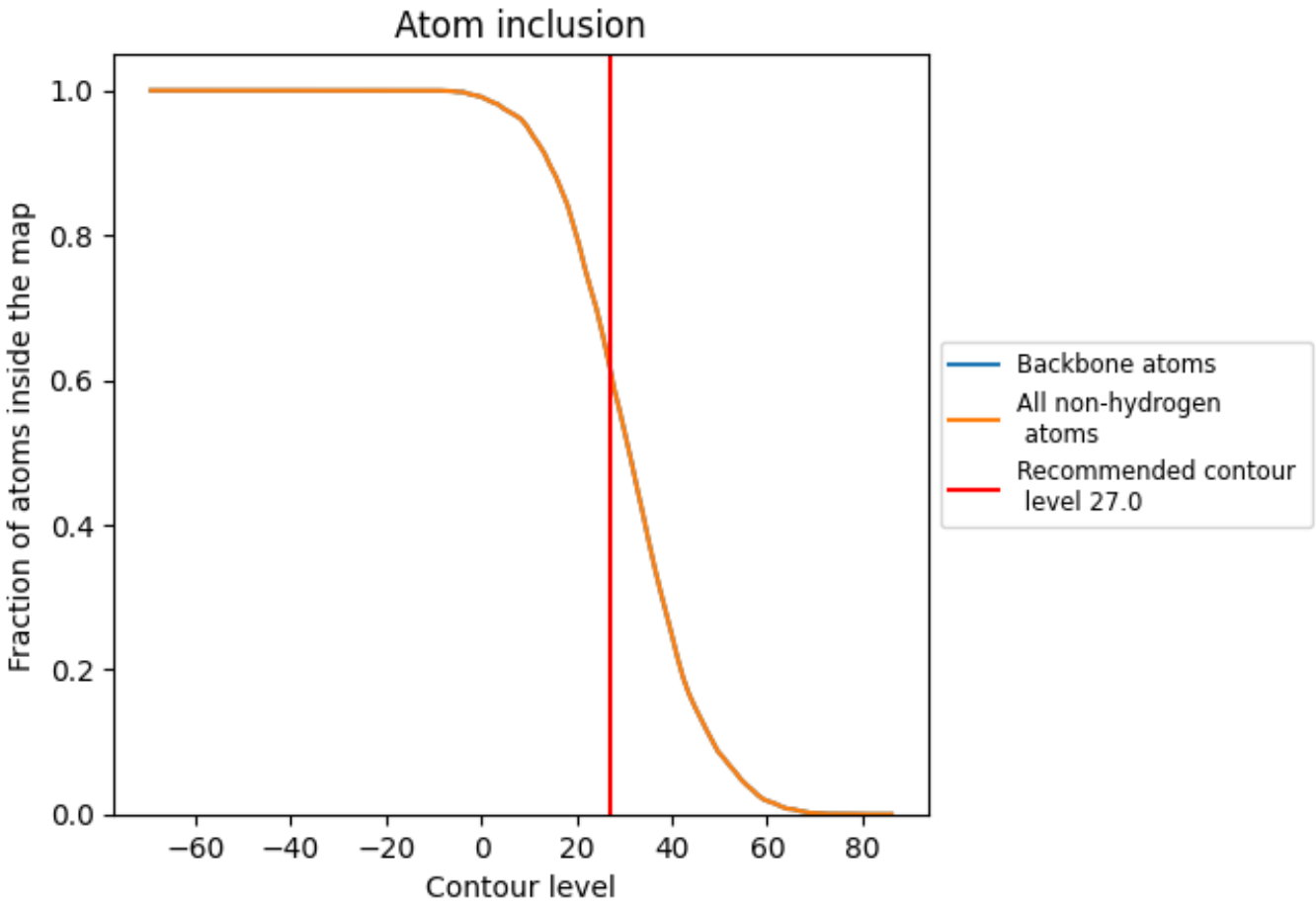
9.3 Atom inclusion mapped to coordinate model (i)



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 (27.0).

9.4 Atom inclusion (i)



At the recommended contour level, 62% of all backbone atoms, 62% 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 (27.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6180	0.2120
A	0.6575	0.2200
B	0.6149	0.1990
C	0.5830	0.2160
D	0.6514	0.2160
E	0.6043	0.2080
F	0.5967	0.2130