wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 03:54 AM EST

PDB ID	:	7KZF
EMDB ID	:	EMD-23081
Title	:	High resolution cryo EM analysis of HPV16 identifies minor structural protein L2 and describes capsid flexibility
Authors	:	Hartmann, S.R.; Goetschius, D.J.; Hafenstein, S.
Deposited on	:	2020-12-10
Resolution	:	3.10 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:

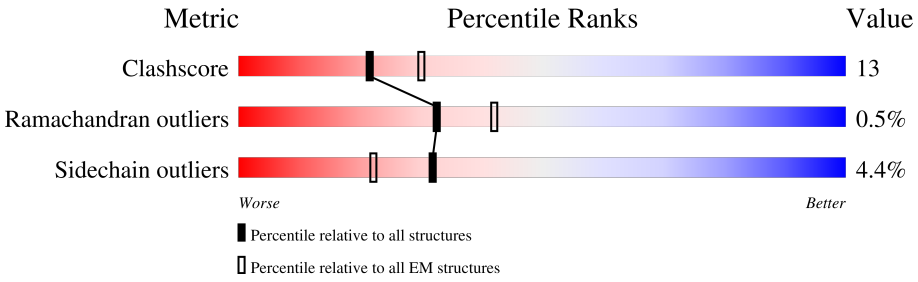
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.10 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 <=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	503	13% 70% 23% • 6%
1	B	503	14% 67% 25% • 7%
1	C	503	16% 70% 24% • 5%
1	D	503	15% 73% 21% • 5%
1	E	503	15% 71% 23% • •
1	F	503	14% 70% 24% • •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 22491 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.

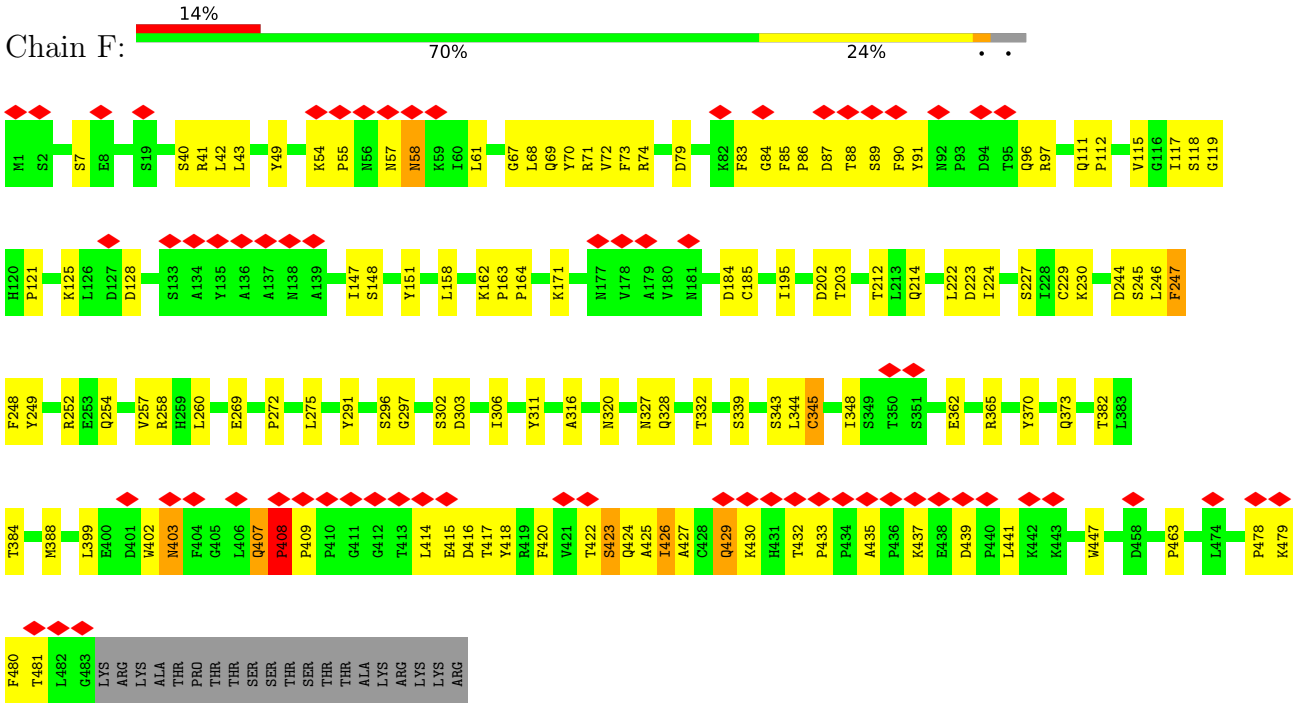
- Molecule 1 is a protein called Major capsid protein L1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	F	483	Total	C	N	O	S	0	0
			3790	2424	633	711	22		
1	B	466	Total	C	N	O	S	0	0
			3659	2335	615	688	21		
1	C	480	Total	C	N	O	S	0	0
			3771	2412	630	707	22		
1	D	480	Total	C	N	O	S	0	0
			3771	2412	630	707	22		
1	E	481	Total	C	N	O	S	0	0
			3776	2416	631	708	21		
1	A	474	Total	C	N	O	S	0	0
			3724	2380	626	697	21		

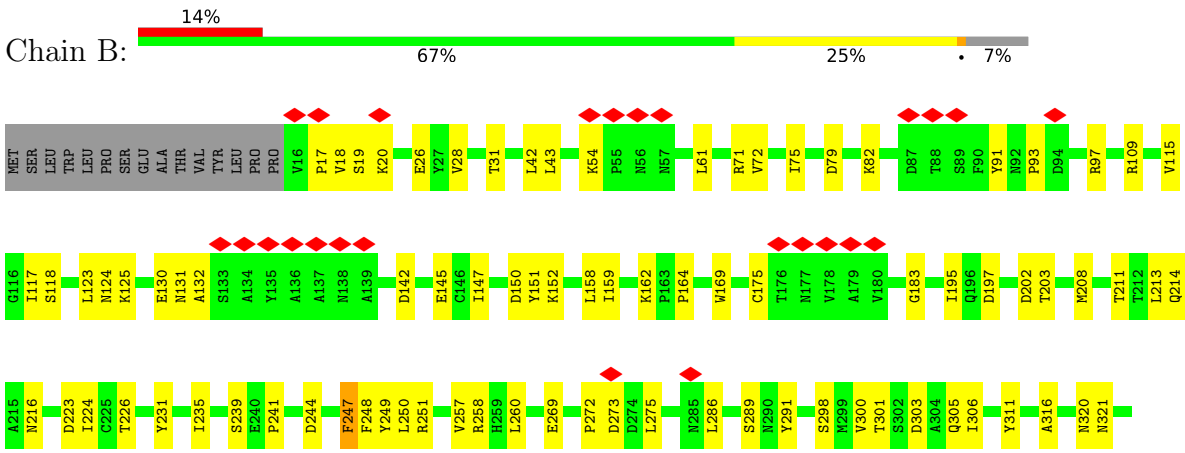
3 Residue-property plots (i)

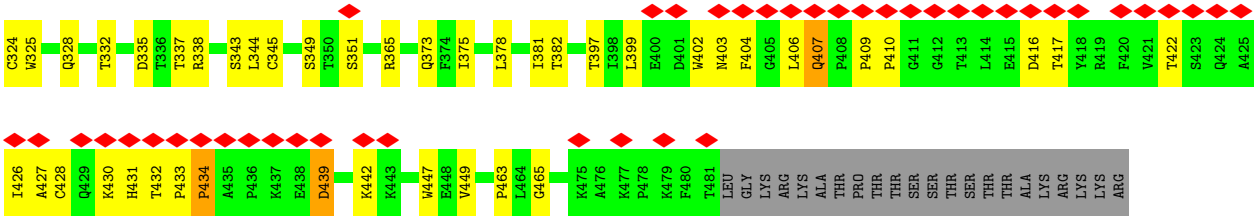
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Major capsid protein L1



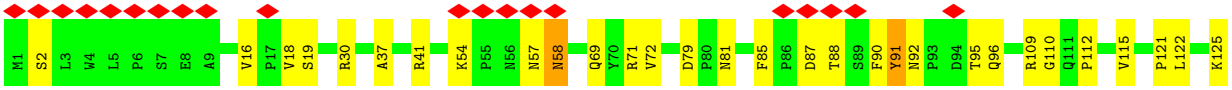
• Molecule 1: Major capsid protein L1

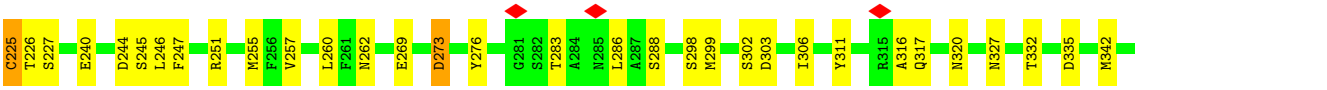


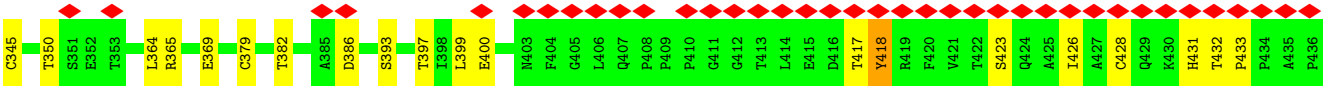


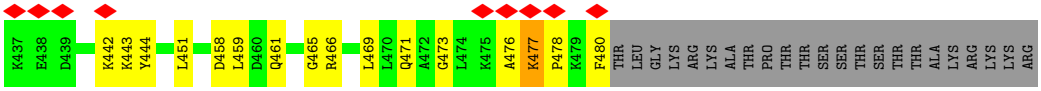
• Molecule 1: Major capsid protein L1

Chain C:



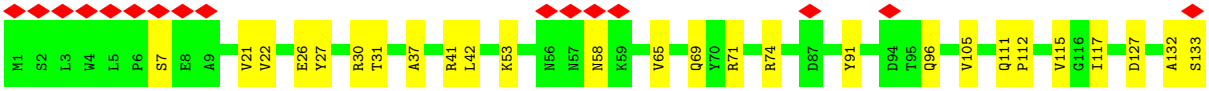


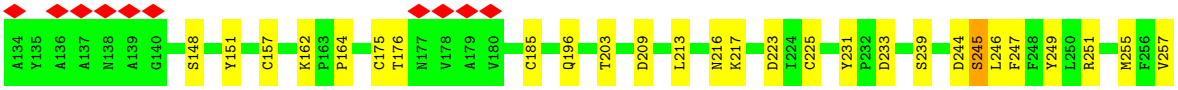


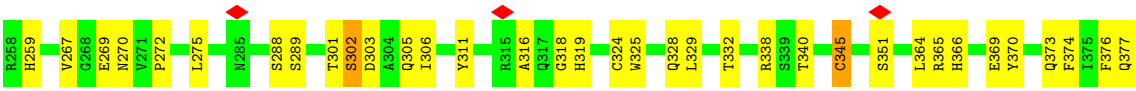


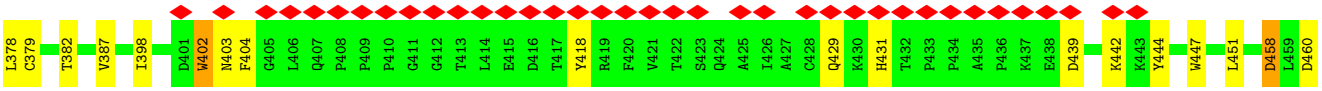
• Molecule 1: Major capsid protein L1

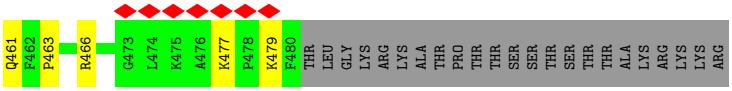
Chain D:





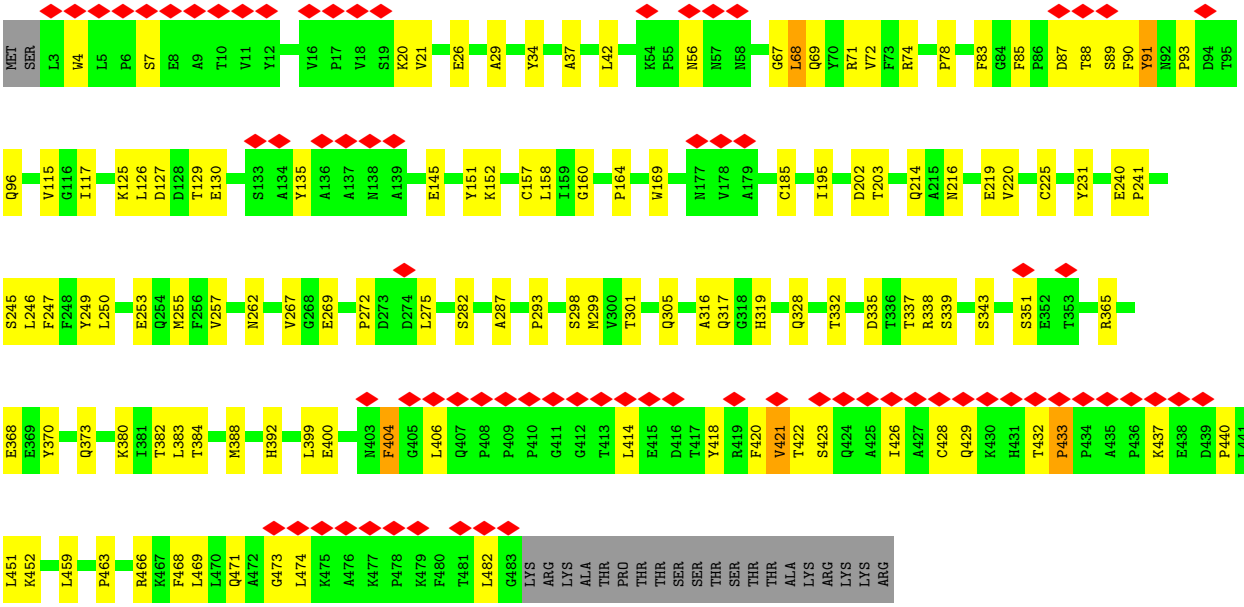






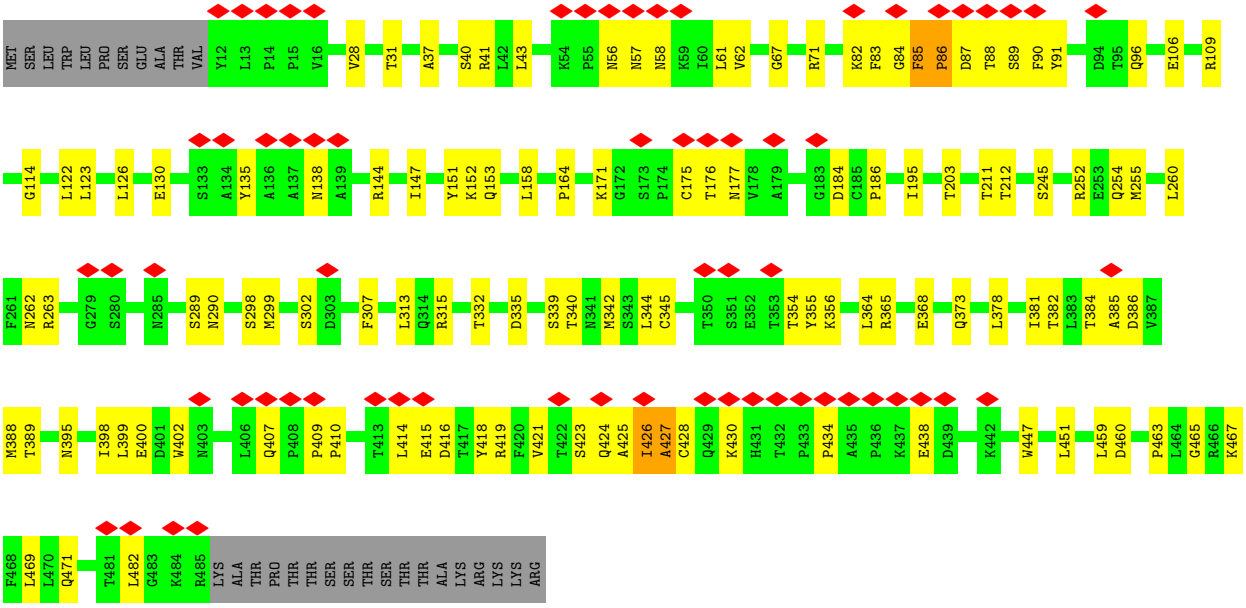
• Molecule 1: Major capsid protein L1

Chain E:



• Molecule 1: Major capsid protein L1

Chain A:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	181299	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 ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	17.655	Depositor
Minimum map value	-10.471	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	2.5	Depositor
Map size (Å)	660.0, 660.0, 660.0	wwPDB
Map dimensions	600, 600, 600	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.1, 1.1, 1.1	Depositor

5 Model quality

5.1 Standard geometry

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	0.35	0/3826	0.47	0/5207
1	B	0.32	0/3758	0.46	0/5115
1	C	0.32	0/3876	0.46	0/5280
1	D	0.32	0/3876	0.45	0/5280
1	E	0.32	0/3881	0.45	1/5288 (0.0%)
1	F	0.32	0/3895	0.46	0/5306
All	All	0.33	0/23112	0.46	1/31476 (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	2
1	B	0	1
1	C	0	4
1	E	0	1
1	F	0	2
All	All	0	10

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	474	LEU	CA-CB-CG	5.49	127.92	115.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	430	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	417	THR	Peptide
1	C	418	TYR	Peptide
1	F	407	GLN	Peptide
1	F	408	PRO	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	3724	0	3644	161	0
1	B	3659	0	3574	112	0
1	C	3771	0	3689	98	0
1	D	3771	0	3687	125	0
1	E	3776	0	3693	89	0
1	F	3790	0	3704	179	0
All	All	22491	0	21991	578	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 13.

The worst 5 of 578 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:F:425:ALA:CB	1:A:62:VAL:HG13	1.30	1.59
1:F:447:TRP:CZ3	1:A:419:ARG:HB3	1.39	1.54
1:F:447:TRP:CZ2	1:A:419:ARG:CA	1.97	1.44
1:D:477:LYS:CE	1:D:479:LYS:HD3	1.51	1.37
1:D:477:LYS:HE2	1:D:479:LYS:CB	1.52	1.36

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	472/503 (94%)	432 (92%)	38 (8%)	2 (0%)	34	69
1	B	464/503 (92%)	424 (91%)	37 (8%)	3 (1%)	25	59
1	C	478/503 (95%)	439 (92%)	36 (8%)	3 (1%)	25	59
1	D	478/503 (95%)	446 (93%)	31 (6%)	1 (0%)	47	79
1	E	479/503 (95%)	447 (93%)	31 (6%)	1 (0%)	47	79
1	F	481/503 (96%)	441 (92%)	37 (8%)	3 (1%)	25	59
All	All	2852/3018 (94%)	2629 (92%)	210 (7%)	13 (0%)	32	64

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	481	THR
1	C	418	TYR
1	D	403	ASN
1	A	427	ALA
1	F	403	ASN

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	412/438 (94%)	395 (96%)	17 (4%)	30	64
1	B	405/438 (92%)	391 (96%)	14 (4%)	36	68
1	C	418/438 (95%)	391 (94%)	27 (6%)	17	47

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	418/438 (95%)	396 (95%)	22 (5%)	22 54
1	E	418/438 (95%)	404 (97%)	14 (3%)	38 69
1	F	420/438 (96%)	404 (96%)	16 (4%)	33 66
All	All	2491/2628 (95%)	2381 (96%)	110 (4%)	32 61

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	7	SER
1	D	302	SER
1	A	482	LEU
1	A	211	THR
1	D	91	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	192	ASN
1	E	429	GLN
1	C	461	GLN
1	E	373	GLN
1	A	373	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.

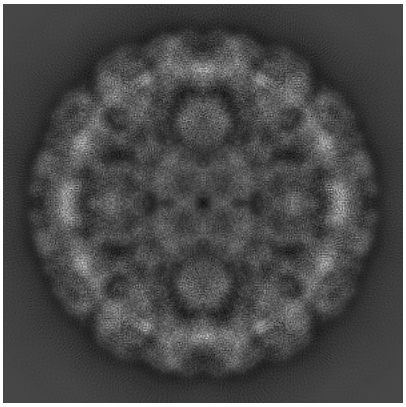
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-23081. These allow visual inspection of the internal detail of the map and identification of artifacts.

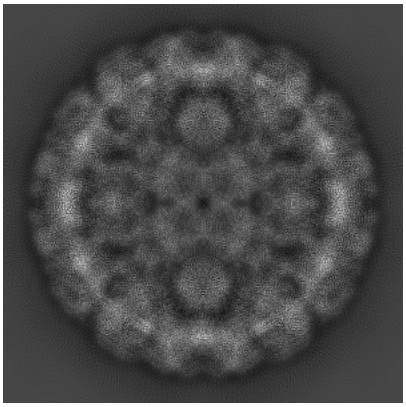
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

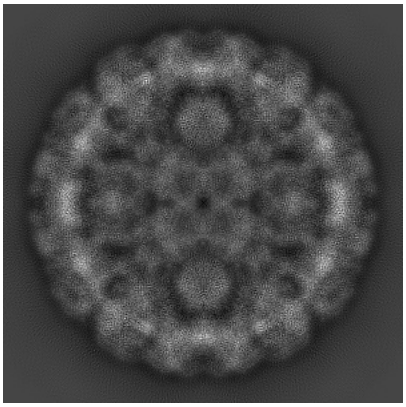
6.1.1 Primary map



X



Y

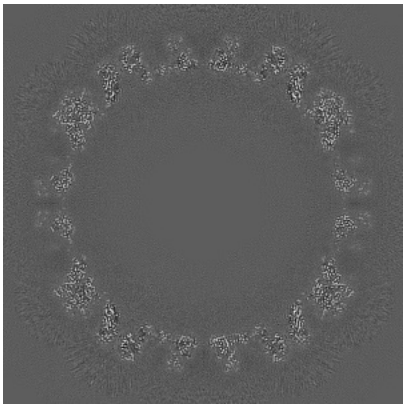


Z

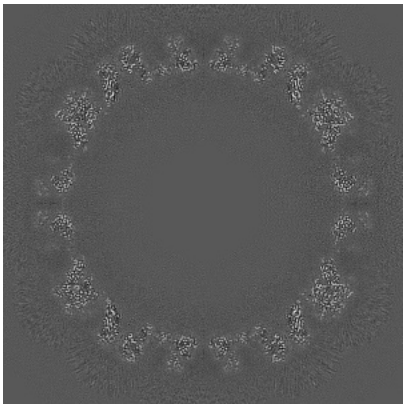
The images above show the map projected in three orthogonal directions.

6.2 Central slices

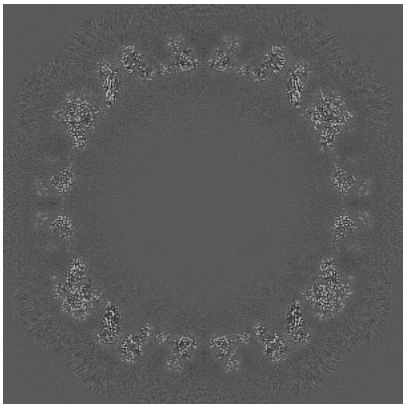
6.2.1 Primary map



X Index: 300



Y Index: 300

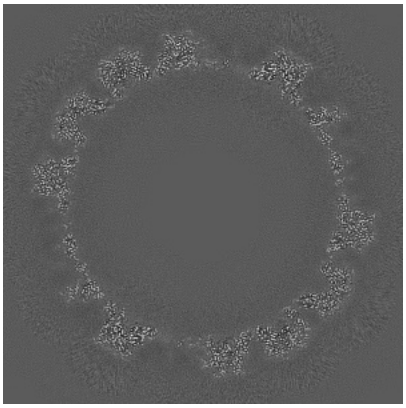


Z Index: 300

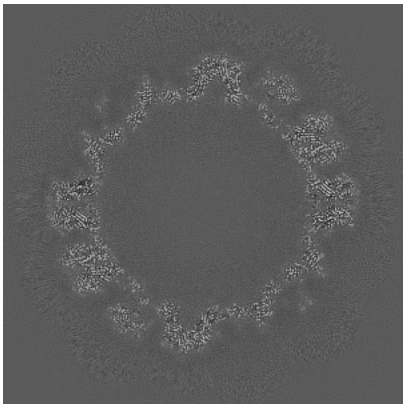
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

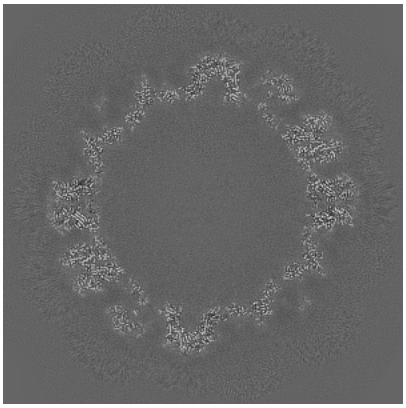
6.3.1 Primary map



X Index: 286



Y Index: 425

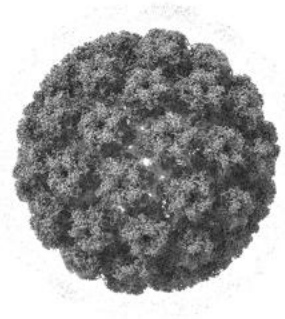


Z Index: 425

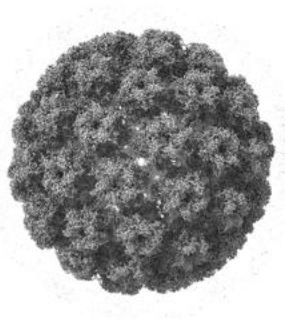
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

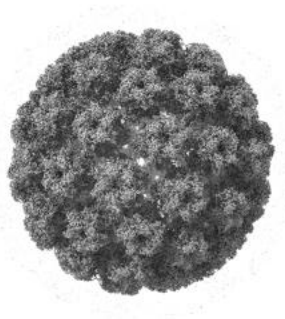
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 2.5. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

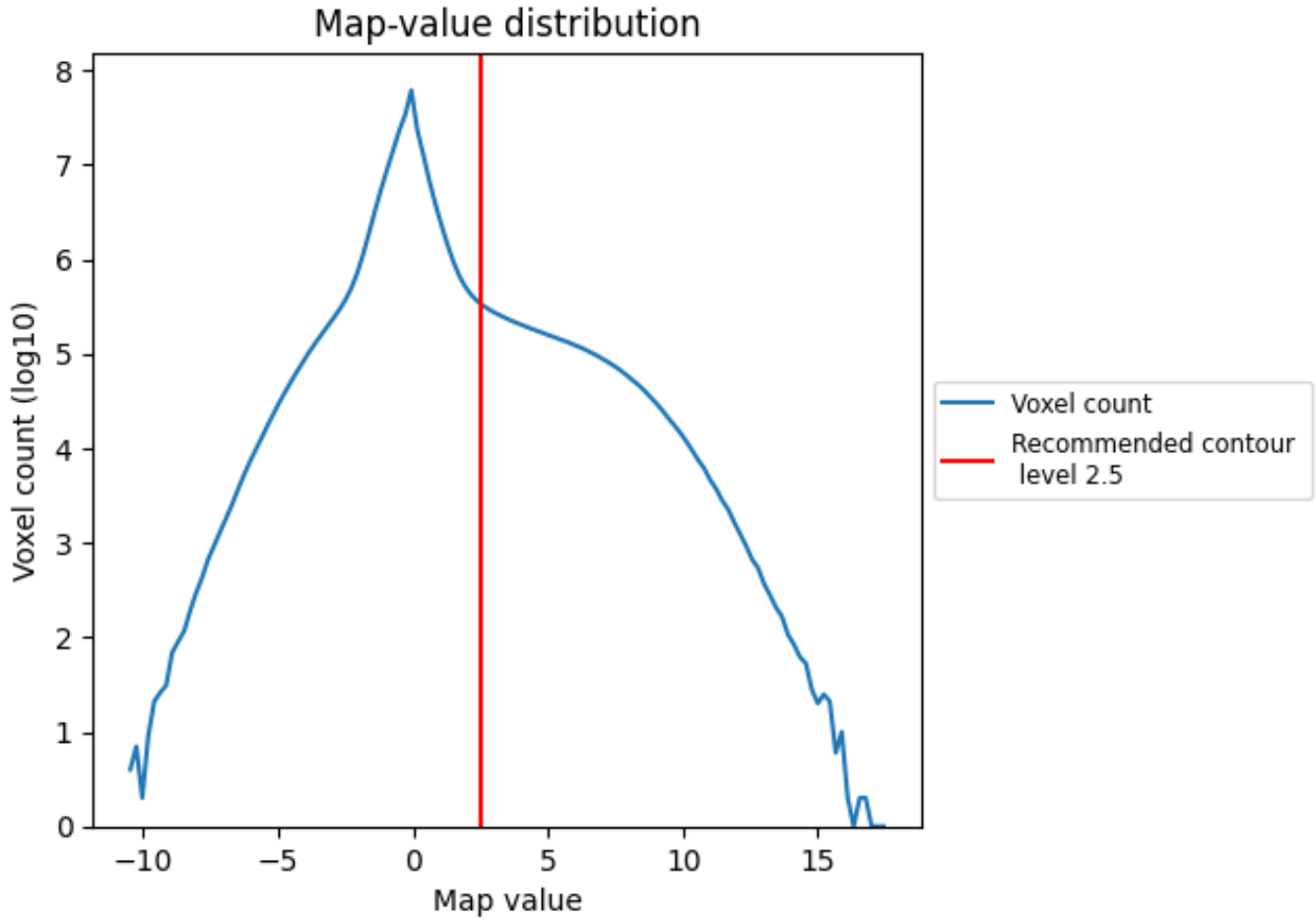
6.5 Mask visualisation (i)

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

7 Map analysis

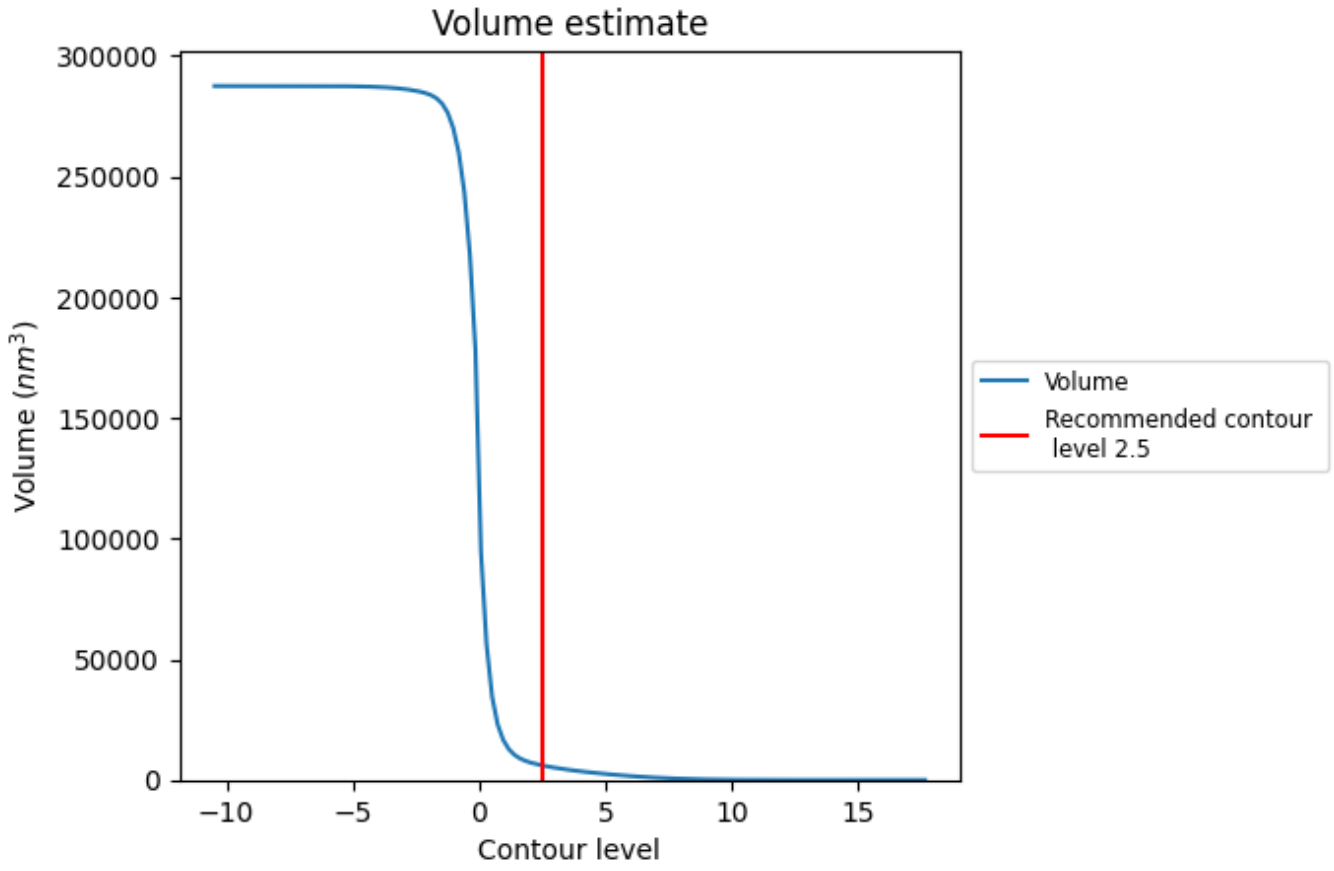
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

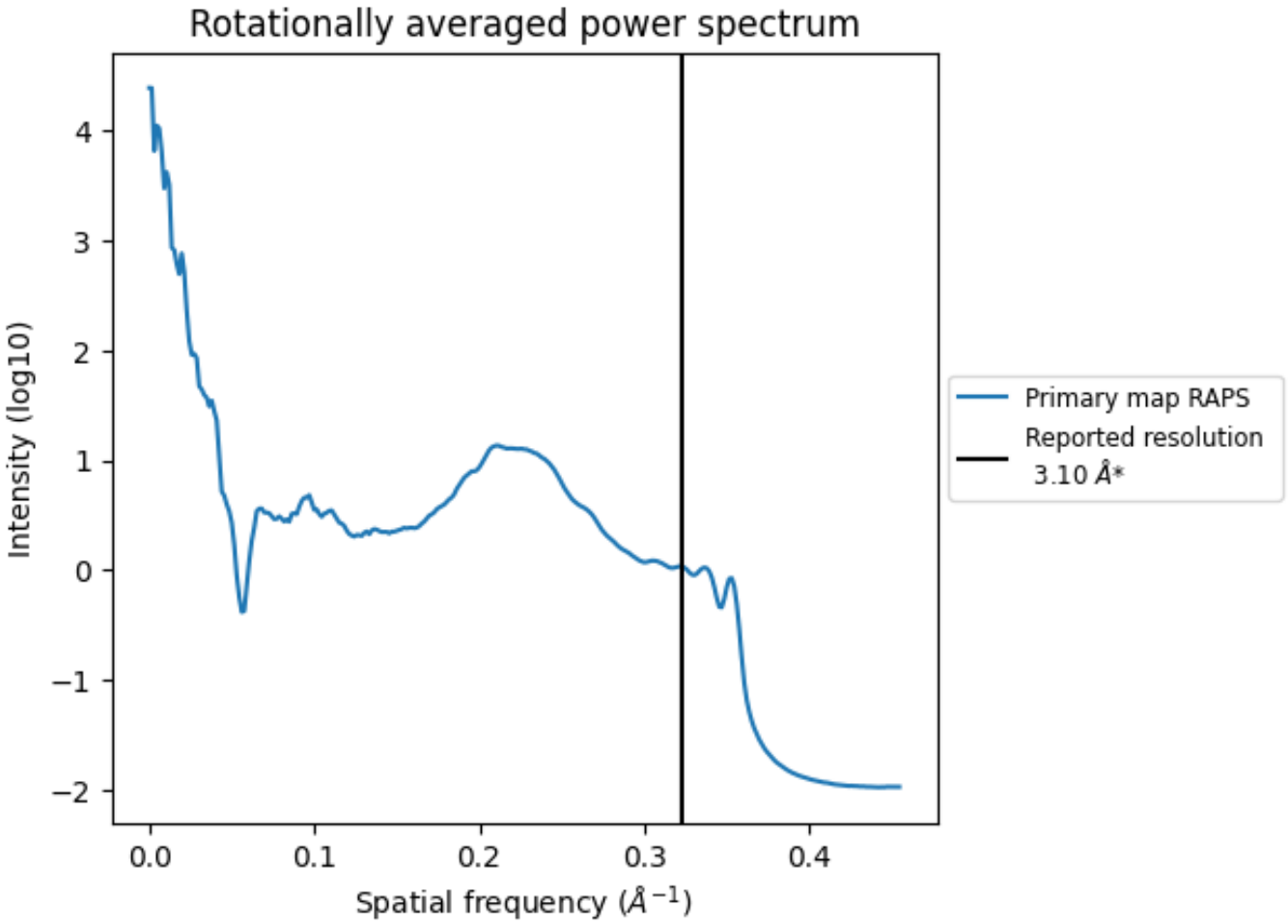
7.2 Volume estimate (i)



The volume at the recommended contour level is 6022 nm^3; this corresponds to an approximate mass of 5440 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 0.323 $Å^{-1}$

8 Fourier-Shell correlation (i)

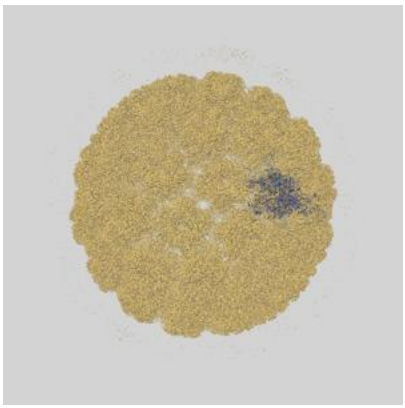
This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

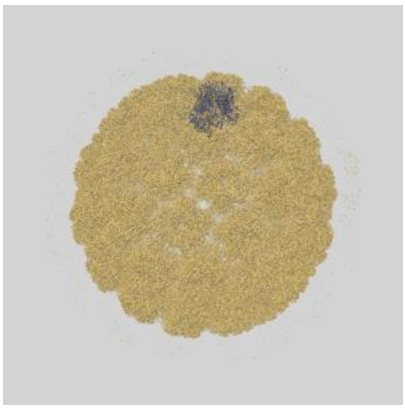
This section contains information regarding the fit between EMDB map EMD-23081 and PDB model 7KZF. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlays

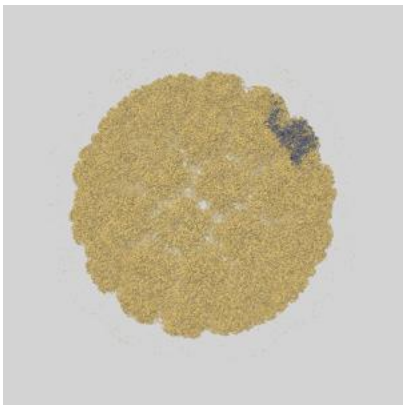
9.1.1 Map-model overlay



X

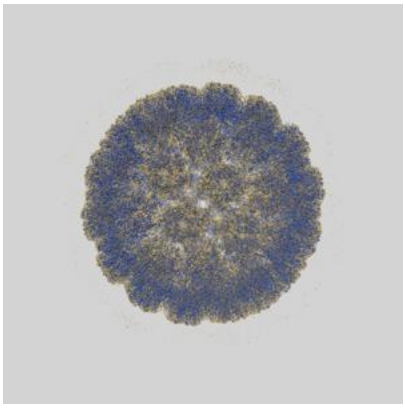


Y

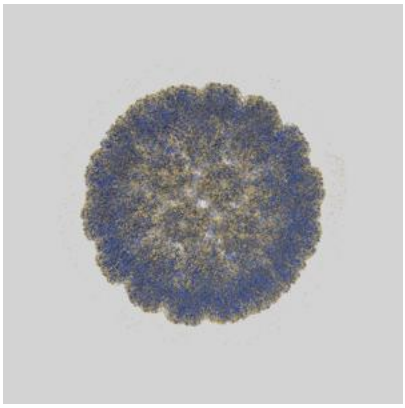


Z

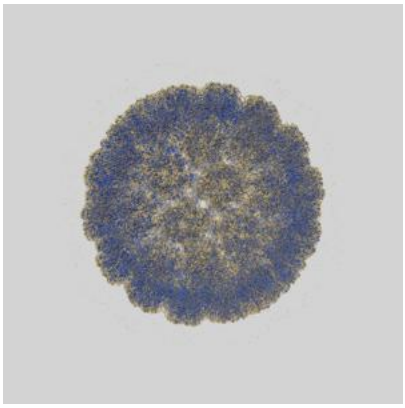
9.1.2 Map-model assembly overlay



X



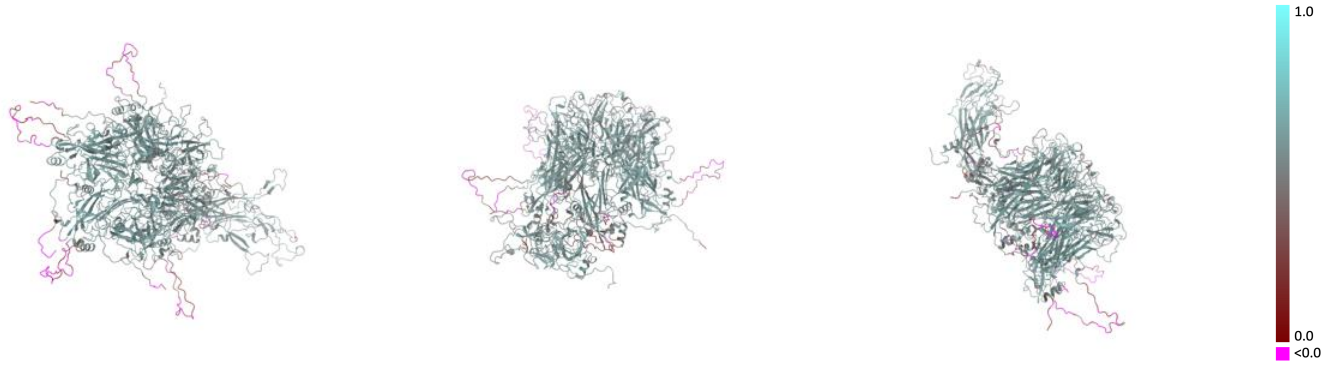
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 2.5 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.

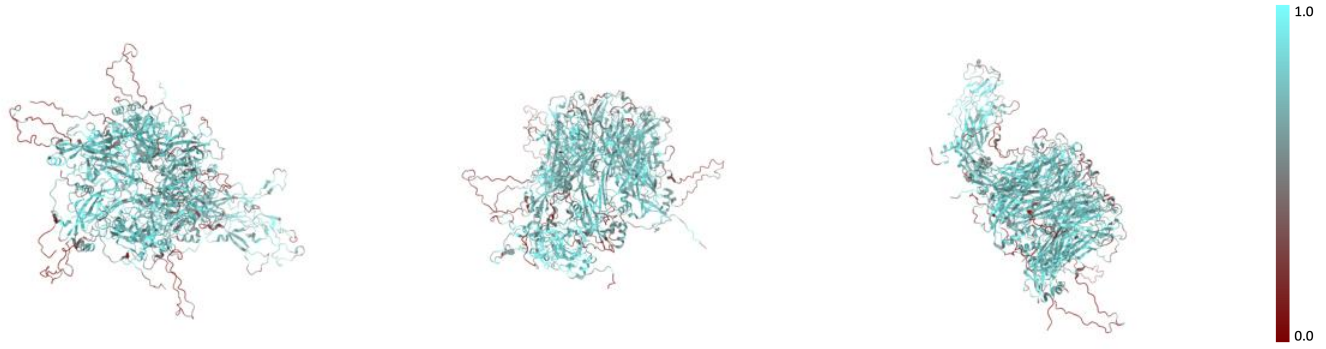
9.2 Q-score mapped to coordinate model



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.

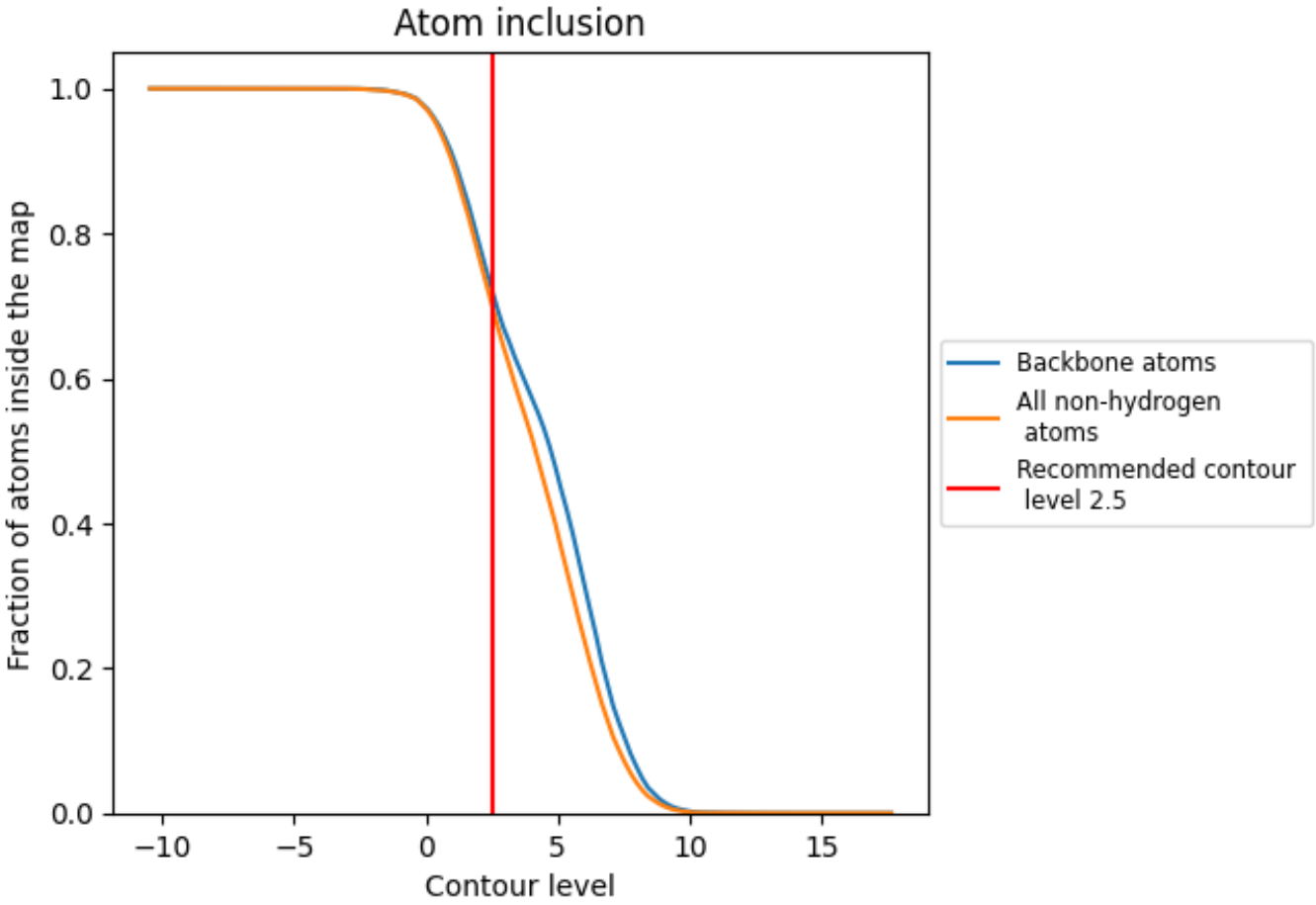
9.3 Atom inclusion mapped to coordinate model



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 (2.5).

9.4 Atom inclusion (i)



At the recommended contour level, 72% of all backbone atoms, 70% 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 (2.5) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7008	0.5280
A	0.6934	0.5230
B	0.7094	0.5390
C	0.7043	0.5170
D	0.7032	0.5310
E	0.6869	0.5270
F	0.7077	0.5320