



# wwPDB EM Validation Summary Report ⓘ

Nov 19, 2022 – 10:28 AM EST

PDB ID : 3J8V  
EMDB ID : EMD-6121  
Title : Cryo-EM reconstruction of quasi-HPV16 complex with H16.14J Fab  
Authors : Guan, J.; Hafenstein, S.  
Deposited on : 2014-11-19  
Resolution : 13.90 Å (reported)  
Based on initial model : 3OAE

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) : 2.31.3

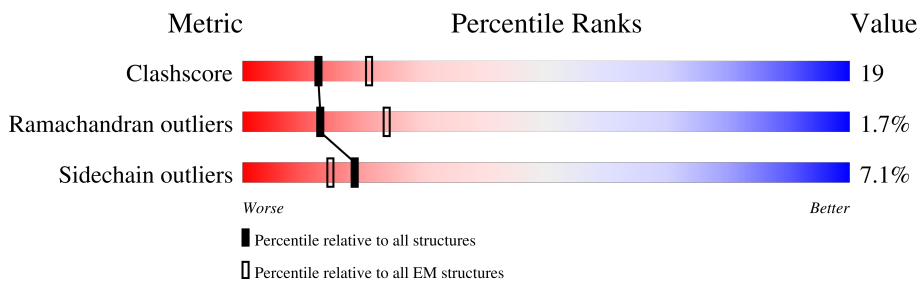
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 13.90 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	A	455	
1	B	455	
1	C	455	
1	D	455	
1	E	455	
2	J	109	
2	K	109	
2	L	109	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	M	109	 92% 7%
3	F	119	 90% 9%
3	G	119	 90% 9%
3	H	119	 88% 11%
3	I	119	 90% 9%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 30346 atoms, of which 6752 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 L1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	421	3322	2116	556	630	20	0	0
1	B	421	3322	2116	556	630	20	0	0
1	C	421	3322	2116	556	630	20	0	0
1	D	421	3322	2116	556	630	20	0	0
1	E	421	3322	2116	556	630	20	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
A	177	GLN	ASN	CONFLICT	UNP Q4VRM0
A	181	GLN	ASN	CONFLICT	UNP Q4VRM0
A	472	LEU	ALA	CONFLICT	UNP Q4VRM0
B	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
B	177	GLN	ASN	CONFLICT	UNP Q4VRM0
B	181	GLN	ASN	CONFLICT	UNP Q4VRM0
B	472	LEU	ALA	CONFLICT	UNP Q4VRM0
C	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
C	177	GLN	ASN	CONFLICT	UNP Q4VRM0
C	181	GLN	ASN	CONFLICT	UNP Q4VRM0
C	472	LEU	ALA	CONFLICT	UNP Q4VRM0
D	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
D	177	GLN	ASN	CONFLICT	UNP Q4VRM0
D	181	GLN	ASN	CONFLICT	UNP Q4VRM0
D	472	LEU	ALA	CONFLICT	UNP Q4VRM0
E	20	ALA	-	EXPRESSION TAG	UNP Q4VRM0
E	177	GLN	ASN	CONFLICT	UNP Q4VRM0
E	181	GLN	ASN	CONFLICT	UNP Q4VRM0
E	472	LEU	ALA	CONFLICT	UNP Q4VRM0

- Molecule 2 is a protein called H16.14J light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
2	L	109	Total	C	H	N	O	S	0	0
			1651	523	820	142	163	3		
2	J	109	Total	C	H	N	O	S	0	0
			1651	523	820	142	163	3		
2	K	109	Total	C	H	N	O	S	0	0
			1651	523	820	142	163	3		
2	M	109	Total	C	H	N	O	S	0	0
			1651	523	820	142	163	3		

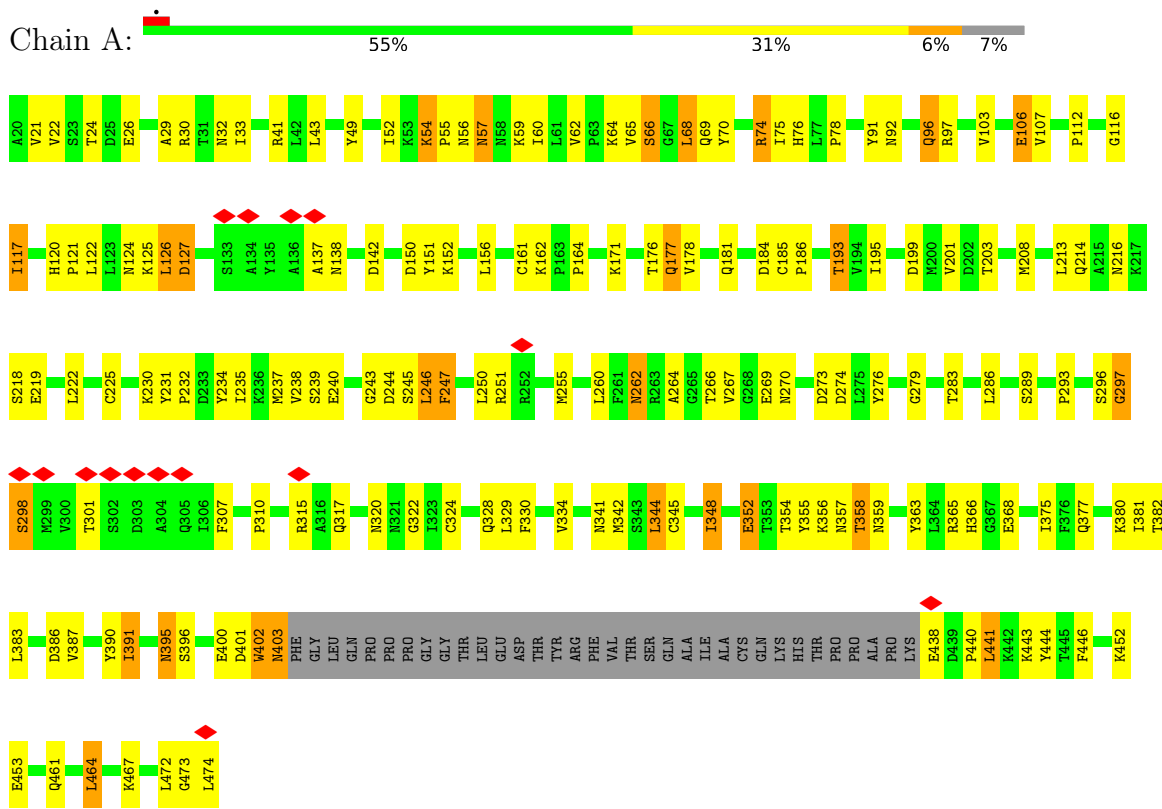
- Molecule 3 is a protein called H16.14J heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace	
			Total	C	H	N	O			S
3	H	119	Total	C	H	N	O	S	0	0
			1783	578	868	153	179	5		
3	F	119	Total	C	H	N	O	S	0	0
			1783	578	868	153	179	5		
3	G	119	Total	C	H	N	O	S	0	0
			1783	578	868	153	179	5		
3	I	119	Total	C	H	N	O	S	0	0
			1783	578	868	153	179	5		

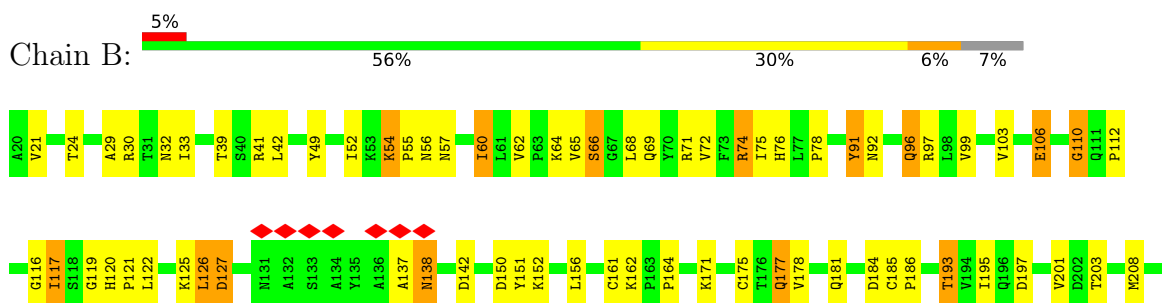
### 3 Residue-property plots [i](#)

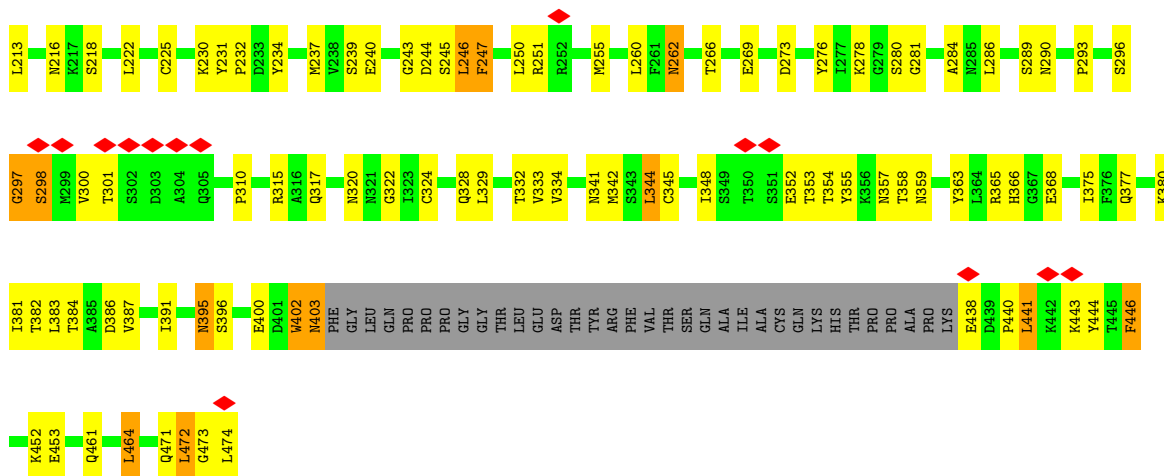
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

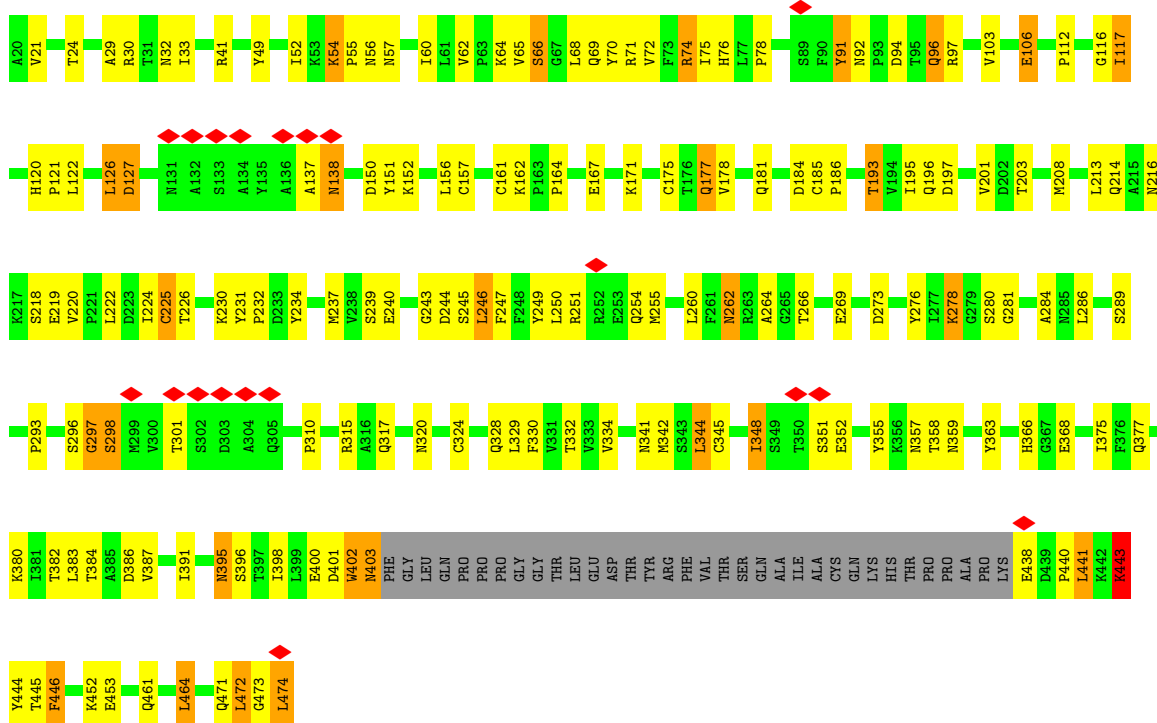


#### • Molecule 1: L1

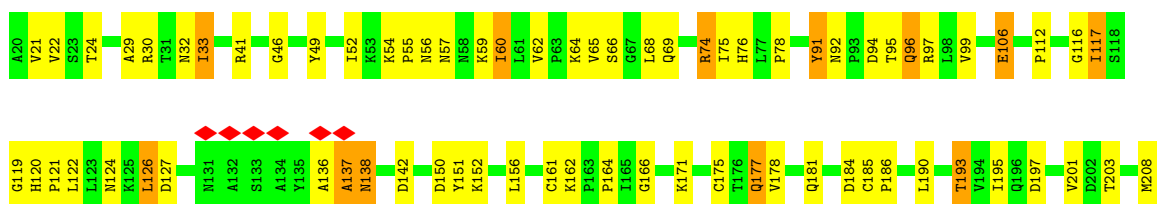


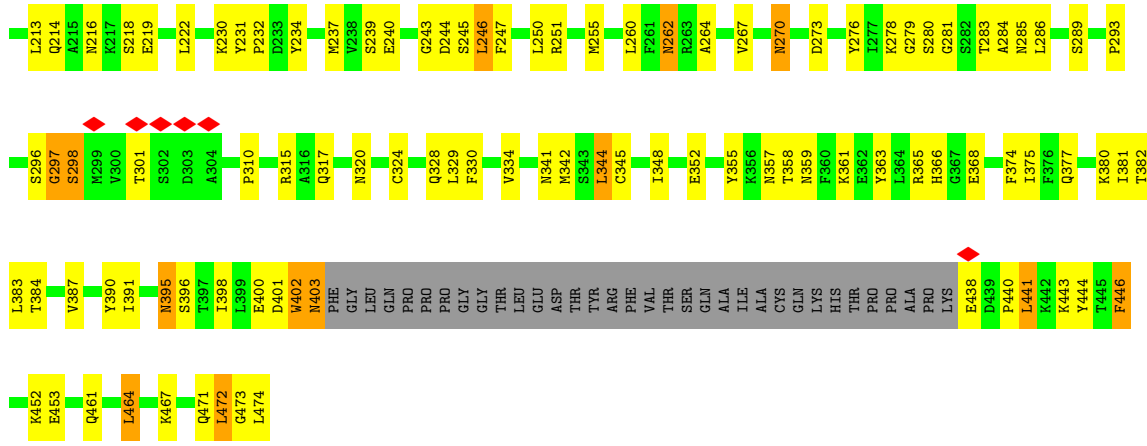


• Molecule 1: L1

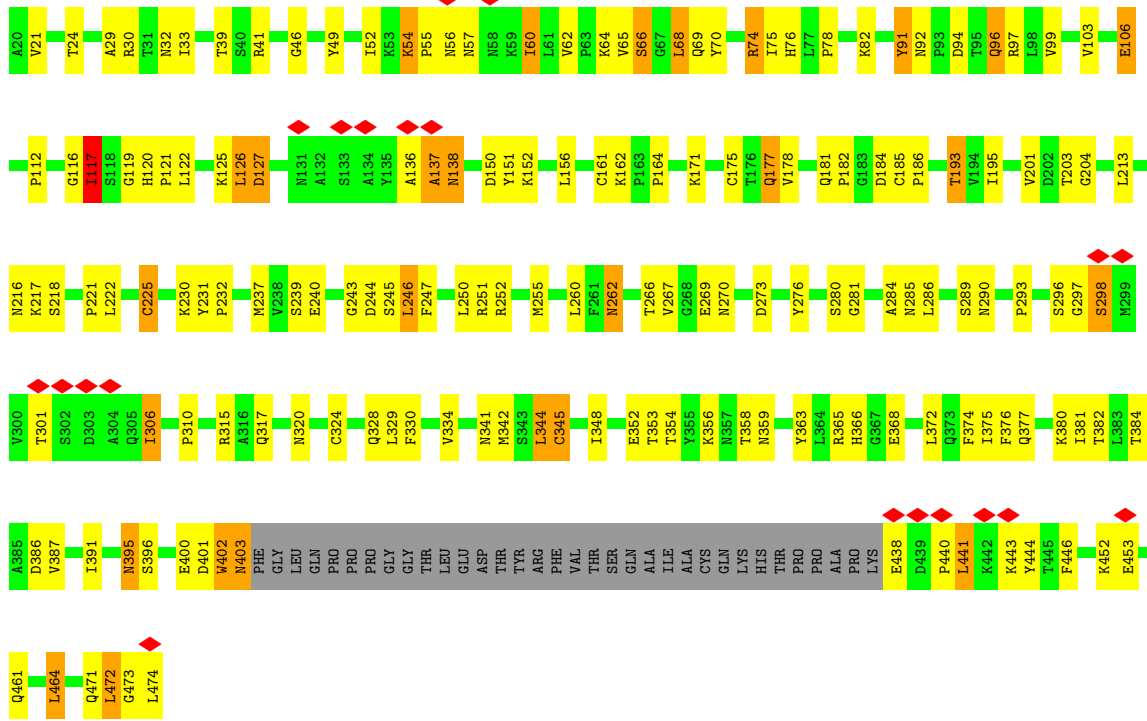


• Molecule 1: L1

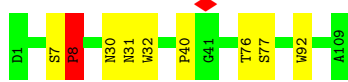
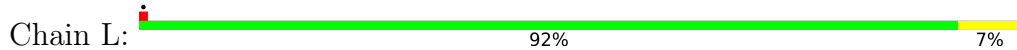




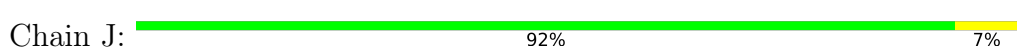
• Molecule 1: L1



• Molecule 2: H16.14J light chain



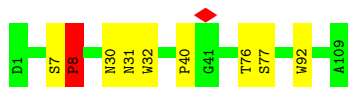
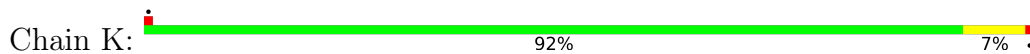
• Molecule 2: H16.14J light chain







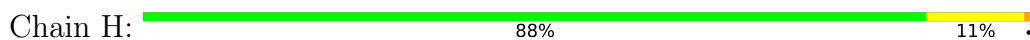
- Molecule 2: H16.14J light chain



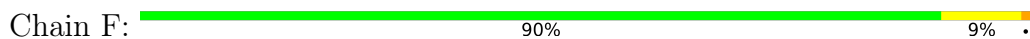
- Molecule 2: H16.14J light chain



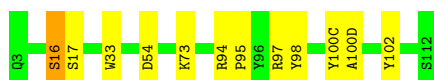
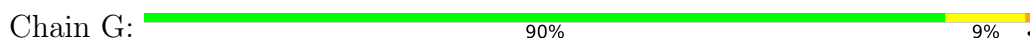
- Molecule 3: H16.14J heavy chain



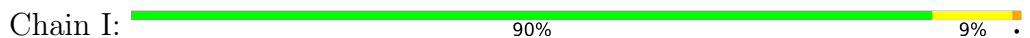
- Molecule 3: H16.14J heavy chain



- Molecule 3: H16.14J heavy chain



- Molecule 3: H16.14J heavy chain



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	5642	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	Each particle	Depositor
Microscope	JEOL 2100	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	15	Depositor
Minimum defocus (nm)	620	Depositor
Maximum defocus (nm)	4690	Depositor
Magnification	50000	Depositor
Image detector	GATAN ULTRASCAN 4000 (4k x 4k)	Depositor
Maximum map value	8.402	Depositor
Minimum map value	-8.250	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	1.0	Depositor
Map size ( $\text{\AA}$ )	1088.11, 1088.11, 1088.11	wwPDB
Map dimensions	467, 467, 467	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	2.33, 2.33, 2.33	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	0.58	6/3407 (0.2%)	0.72	3/4632 (0.1%)
1	B	0.49	0/3407	0.70	2/4632 (0.0%)
1	C	0.53	2/3407 (0.1%)	0.71	1/4632 (0.0%)
1	D	0.55	3/3407 (0.1%)	0.71	3/4632 (0.1%)
1	E	0.52	1/3407 (0.0%)	0.72	0/4632
2	J	0.69	1/850 (0.1%)	0.70	1/1155 (0.1%)
2	K	0.69	1/850 (0.1%)	0.70	1/1155 (0.1%)
2	L	0.70	1/850 (0.1%)	0.70	1/1155 (0.1%)
2	M	0.69	1/850 (0.1%)	0.70	1/1155 (0.1%)
3	F	0.74	2/939 (0.2%)	0.68	0/1269
3	G	0.74	2/939 (0.2%)	0.68	0/1269
3	H	0.74	2/939 (0.2%)	0.68	0/1269
3	I	0.74	2/939 (0.2%)	0.68	0/1269
All	All	0.60	24/24191 (0.1%)	0.71	13/32856 (0.0%)

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	8	PRO	N-CD	9.84	1.61	1.47
2	L	8	PRO	N-CD	9.77	1.61	1.47
2	M	8	PRO	N-CD	9.77	1.61	1.47
2	J	8	PRO	N-CD	9.77	1.61	1.47
3	H	102	TYR	C-N	9.13	1.55	1.34

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	8	PRO	CA-N-CD	-6.87	101.89	111.50
2	M	8	PRO	CA-N-CD	-6.86	101.90	111.50
2	J	8	PRO	CA-N-CD	-6.86	101.90	111.50
2	K	8	PRO	CA-N-CD	-6.84	101.92	111.50
1	A	443	LYS	CD-CE-NZ	-6.41	96.96	111.70

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	3322	0	3227	141	0
1	B	3322	0	3227	190	0
1	C	3322	0	3226	190	0
1	D	3322	0	3227	190	0
1	E	3322	0	3227	196	0
2	J	831	820	820	26	0
2	K	831	820	820	30	0
2	L	831	820	820	29	0
2	M	831	820	820	22	0
3	F	915	868	865	45	0
3	G	915	868	864	40	0
3	H	915	868	865	53	0
3	I	915	868	865	30	0
All	All	23594	6752	22873	880	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 19.

The worst 5 of 880 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:B:280:SER:HB2	2:M:92:TRP:CZ2	1.26	1.59
1:C:358:THR:HG21	3:F:98:TYR:CZ	1.32	1.59
1:C:284:ALA:CB	2:K:32:TRP:HH2	1.24	1.50
1:E:181:GLN:CB	3:H:73:LYS:HZ3	1.25	1.48
1:B:181:GLN:CB	3:I:73:LYS:NZ	1.74	1.47

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	417/455 (92%)	378 (91%)	29 (7%)	10 (2%)	6	33
1	B	417/455 (92%)	378 (91%)	29 (7%)	10 (2%)	6	33
1	C	417/455 (92%)	375 (90%)	34 (8%)	8 (2%)	8	38
1	D	417/455 (92%)	377 (90%)	30 (7%)	10 (2%)	6	33
1	E	417/455 (92%)	378 (91%)	31 (7%)	8 (2%)	8	38
2	J	107/109 (98%)	100 (94%)	7 (6%)	0	100	100
2	K	107/109 (98%)	100 (94%)	7 (6%)	0	100	100
2	L	107/109 (98%)	100 (94%)	7 (6%)	0	100	100
2	M	107/109 (98%)	100 (94%)	7 (6%)	0	100	100
3	F	117/119 (98%)	112 (96%)	4 (3%)	1 (1%)	17	57
3	G	117/119 (98%)	112 (96%)	4 (3%)	1 (1%)	17	57
3	H	117/119 (98%)	112 (96%)	4 (3%)	1 (1%)	17	57
3	I	117/119 (98%)	112 (96%)	4 (3%)	1 (1%)	17	57
All	All	2981/3187 (94%)	2734 (92%)	197 (7%)	50 (2%)	13	42

5 of 50 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	ALA
1	A	298	SER
1	A	402	TRP
1	B	137	ALA
1	B	298	SER

### 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	368/396 (93%)	334 (91%)	34 (9%)	9	29
1	B	368/396 (93%)	335 (91%)	33 (9%)	9	30
1	C	368/396 (93%)	329 (89%)	39 (11%)	6	24
1	D	368/396 (93%)	335 (91%)	33 (9%)	9	30
1	E	368/396 (93%)	330 (90%)	38 (10%)	7	25
2	J	92/92 (100%)	90 (98%)	2 (2%)	52	71
2	K	92/92 (100%)	90 (98%)	2 (2%)	52	71
2	L	92/92 (100%)	90 (98%)	2 (2%)	52	71
2	M	92/92 (100%)	90 (98%)	2 (2%)	52	71
3	F	95/95 (100%)	95 (100%)	0	100	100
3	G	95/95 (100%)	95 (100%)	0	100	100
3	H	95/95 (100%)	95 (100%)	0	100	100
3	I	95/95 (100%)	95 (100%)	0	100	100
All	All	2588/2728 (95%)	2403 (93%)	185 (7%)	18	39

5 of 185 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	138	ASN
1	E	66	SER
1	D	247	PHE
1	D	395	ASN
1	E	117	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	262	ASN
1	E	285	ASN
1	C	461	GLN
1	E	262	ASN
1	E	377	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-6121. These allow visual inspection of the internal detail of the map and identification of artifacts.

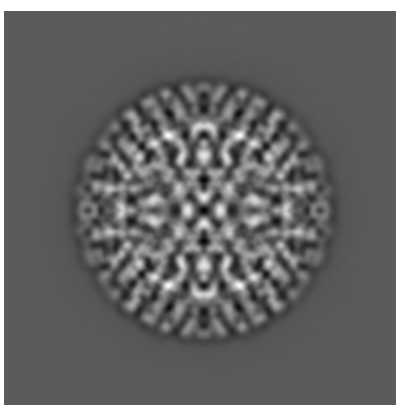
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

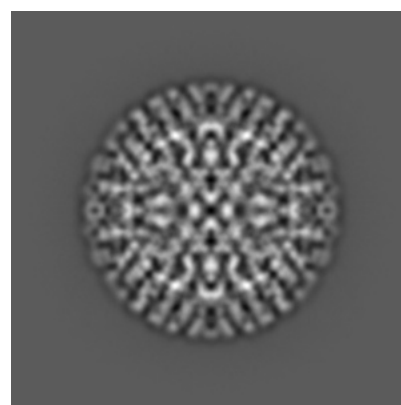
#### 6.1.1 Primary map



X



Y

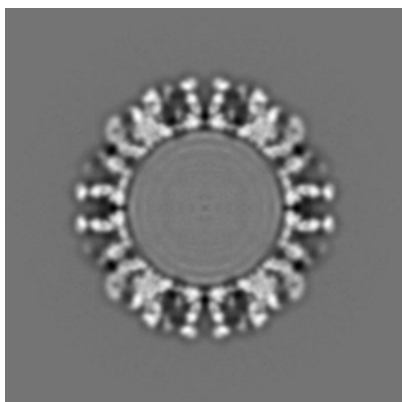


Z

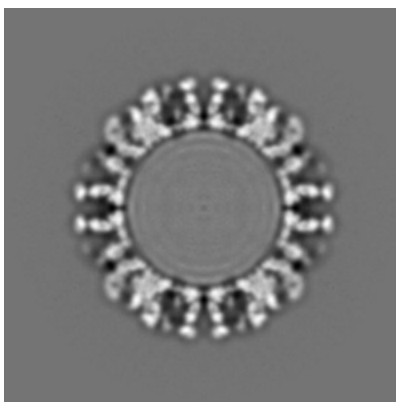
The images above show the map projected in three orthogonal directions.

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

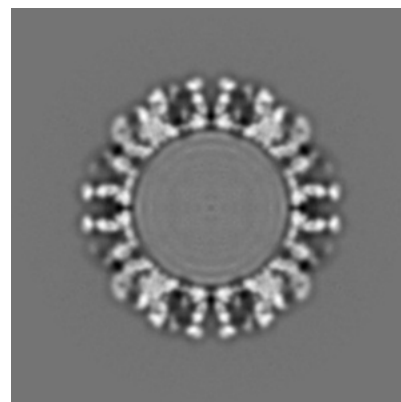
#### 6.2.1 Primary map



X Index: 233



Y Index: 233



Z Index: 233



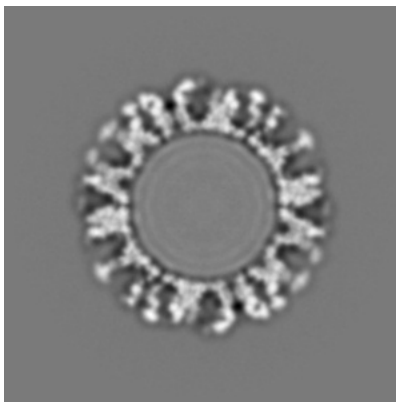
The images above show central slices of the map in three orthogonal directions.

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

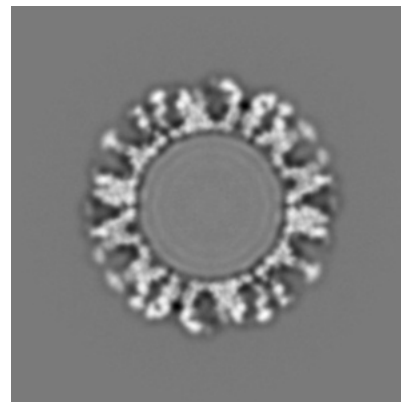
### 6.3.1 Primary map



X Index: 203



Y Index: 203



Z Index: 263

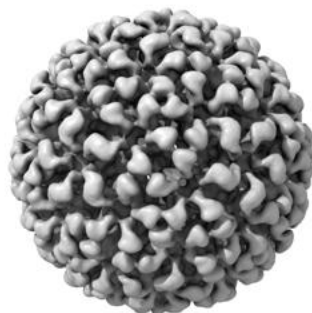
The images above show the largest variance slices of the map in three orthogonal directions.

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

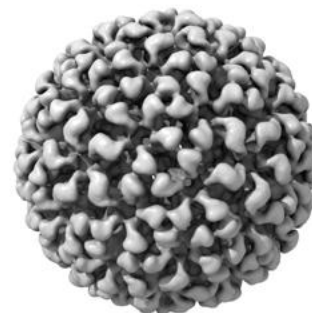
### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 1.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

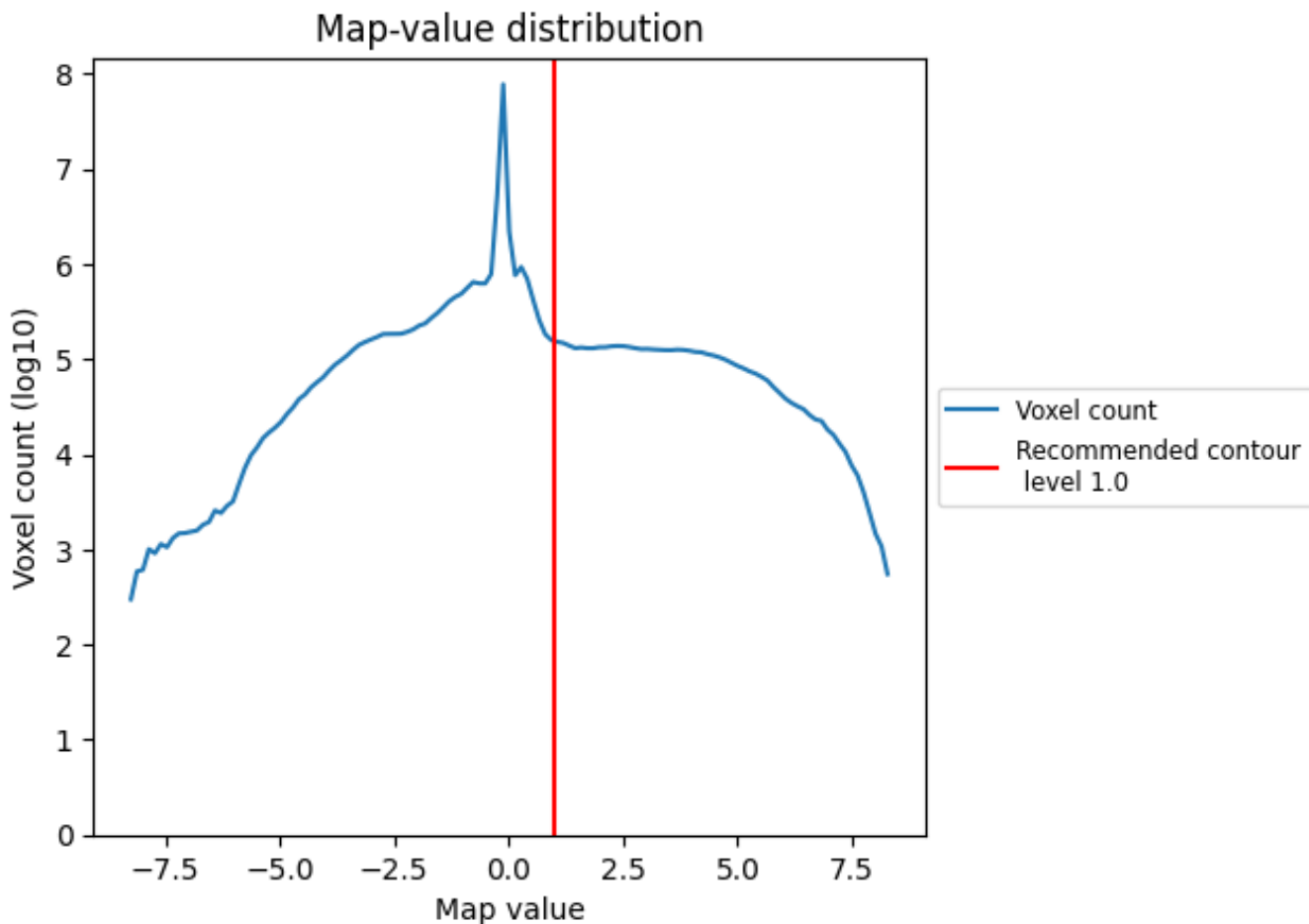
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

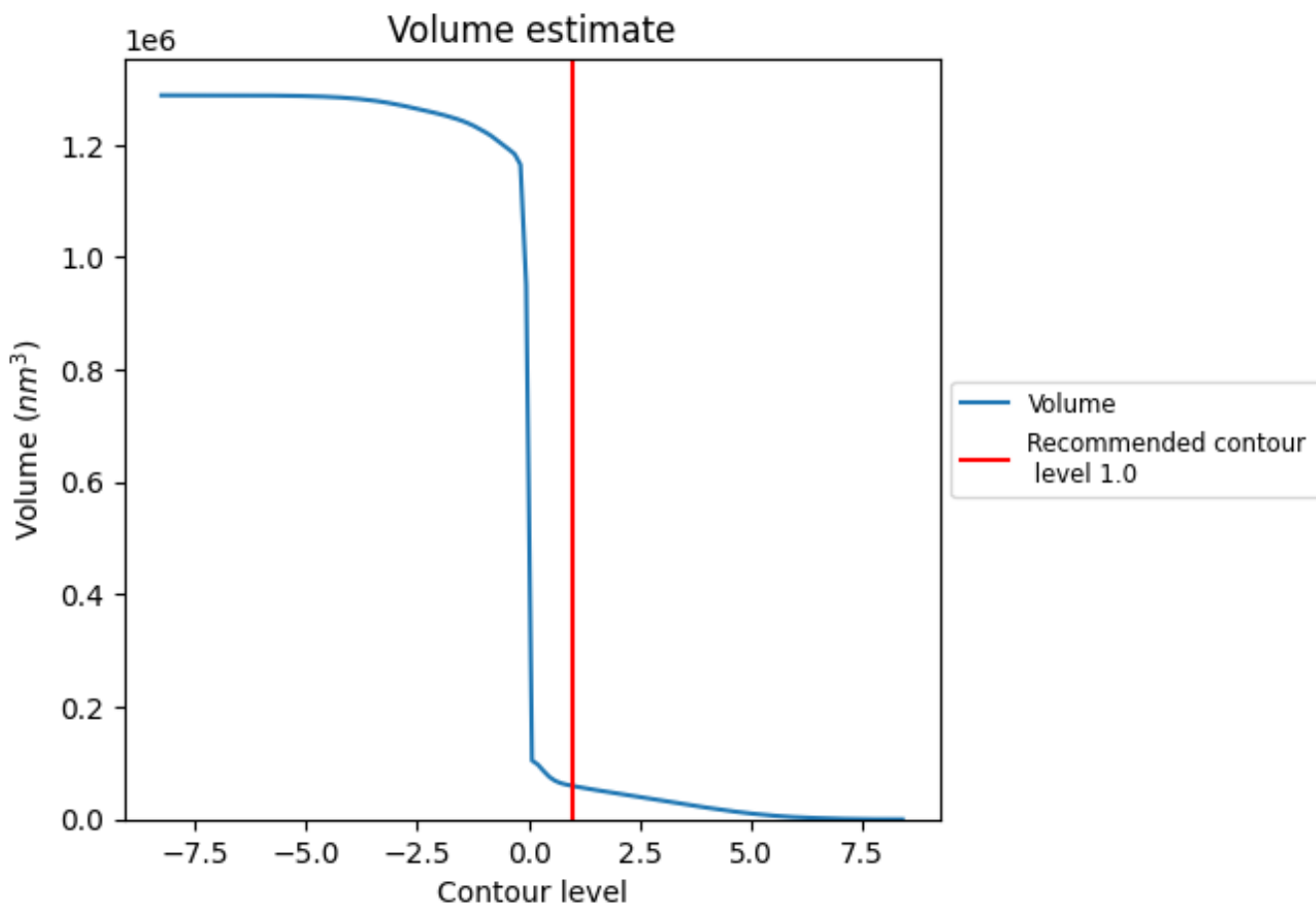
This section contains the results of statistical analysis of the map.

### 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.

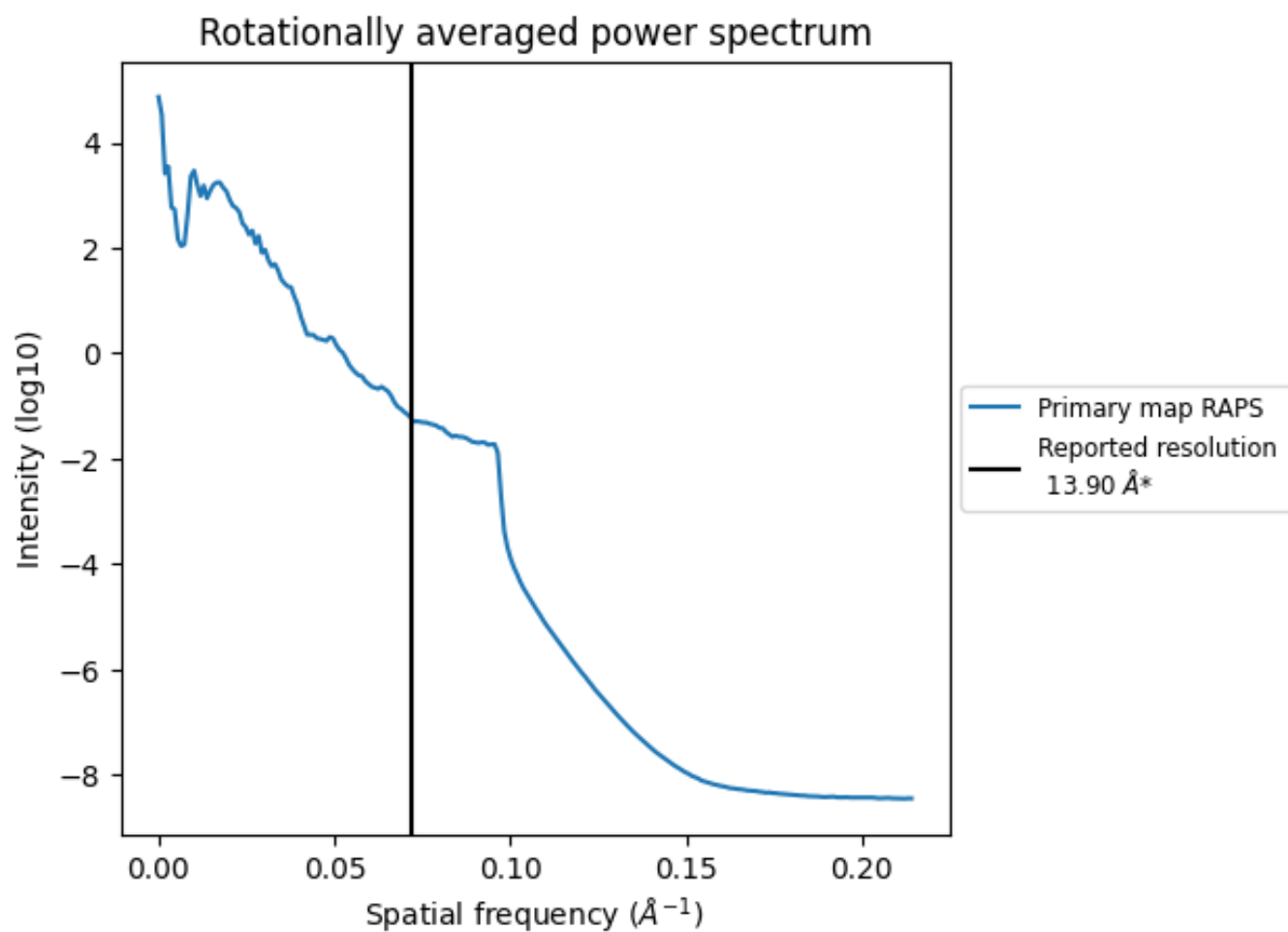
## 7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 59294 nm<sup>3</sup>; this corresponds to an approximate mass of 53562 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  $0.072 \text{\AA}^{-1}$

## 8 Fourier-Shell correlation

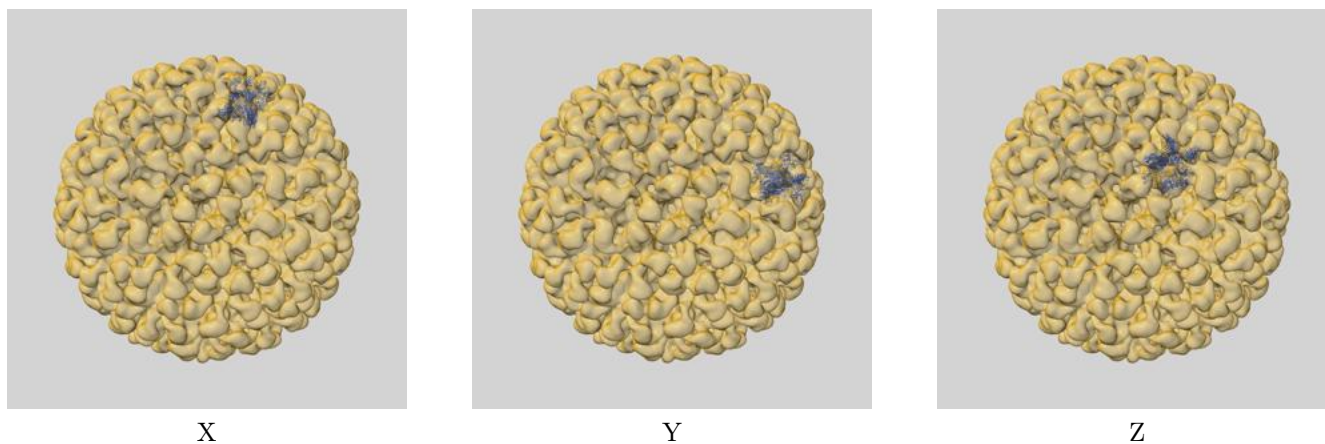
This section was not generated. No FSC curve or half-maps provided.

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

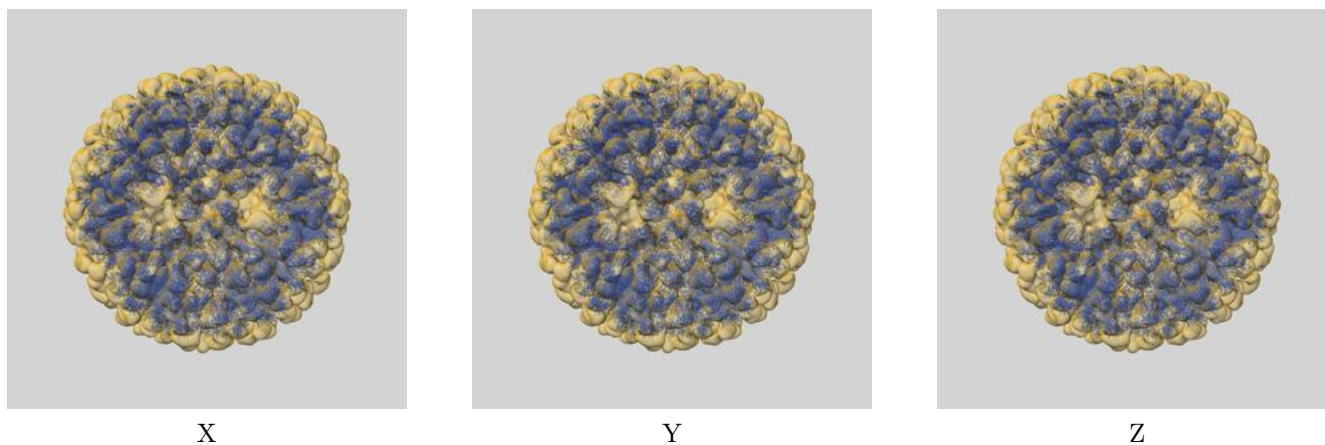
This section contains information regarding the fit between EMDB map EMD-6121 and PDB model 3J8V. Per-residue inclusion information can be found in section 3 on page 6.

### 9.1 Map-model overlays

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

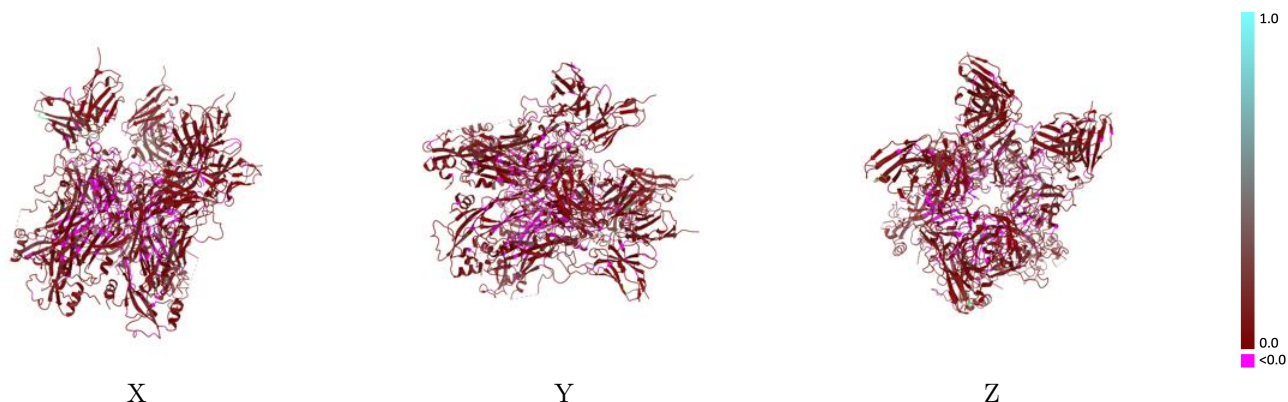


#### 9.1.2 Map-model assembly overlay [i](#)



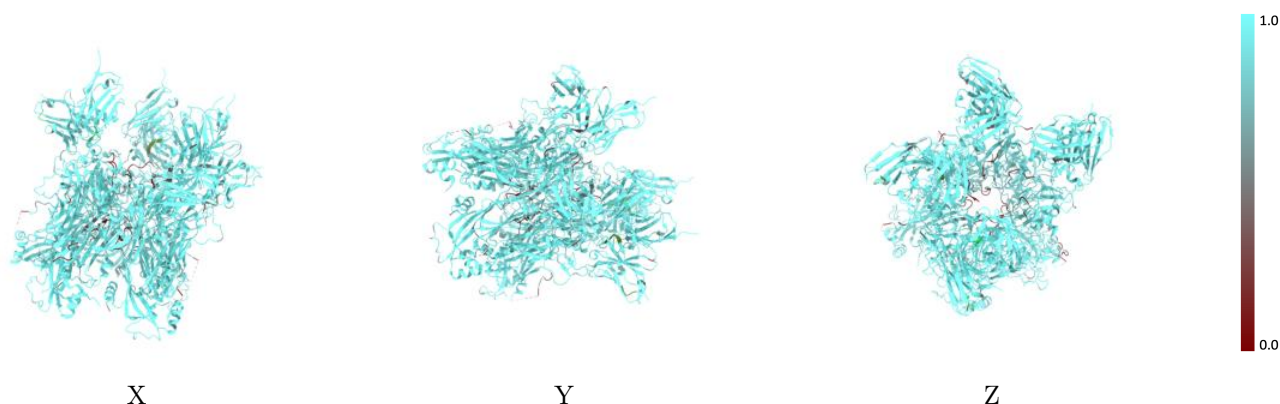
The images above show the 3D surface view of the map at the recommended contour level 1.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



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.

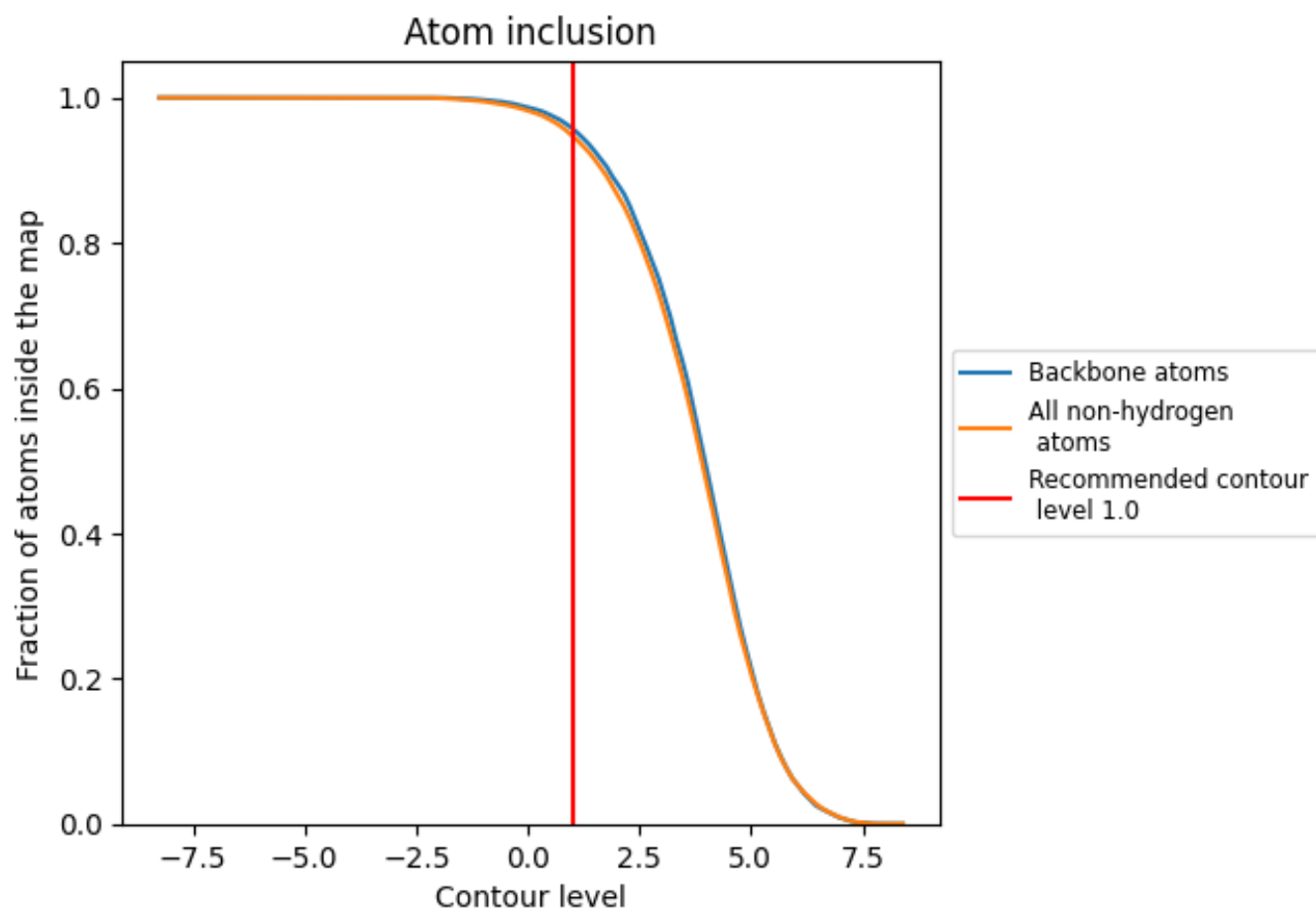
## 9.3 Atom inclusion mapped to coordinate model [\(i\)](#)



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 (1.0).































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 96% of all backbone atoms, 95% 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 (1.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9475	 0.0680
A	 0.9452	 0.0630
B	 0.9351	 0.0580
C	 0.9385	 0.0650
D	 0.9501	 0.0690
E	 0.9375	 0.0670
F	 0.9732	 0.0720
G	 0.9620	 0.0760
H	 0.9720	 0.0770
I	 0.9754	 0.0760
J	 0.9902	 0.0800
K	 0.9779	 0.0750
L	 0.9902	 0.0810
M	 0.9939	 0.0750

