



wwPDB EM Validation Summary Report ⓘ

Mar 11, 2024 – 06:08 AM EDT

PDB ID : 6OKB
EMDB ID : EMD-20099
Title : Prohead 2 of the phage T5
Authors : Huet, A.; Duda, R.L.; Boulanger, P.; Conway, J.F.
Deposited on : 2019-04-12
Resolution : 6.70 Å (reported)
Based on initial model : 2FT1

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:

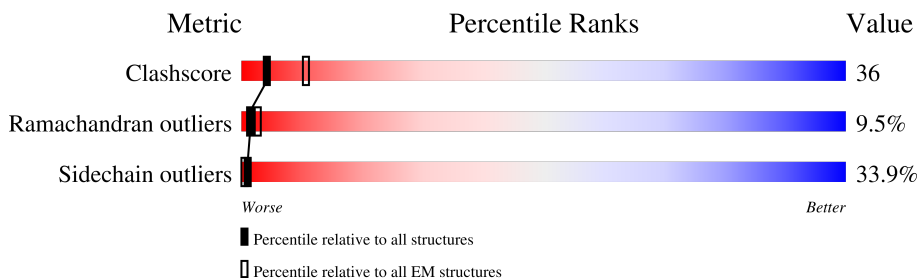
EMDB validation analysis : 0.0.1.dev70
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 6.70 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.



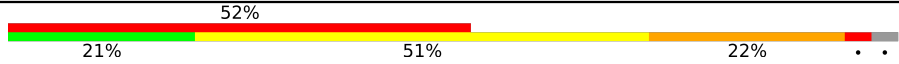
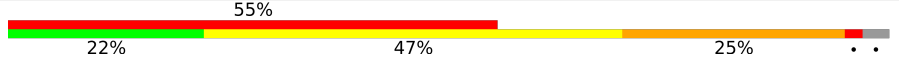
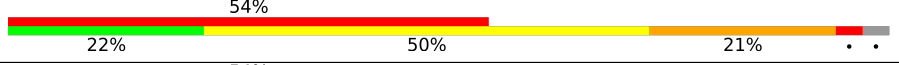
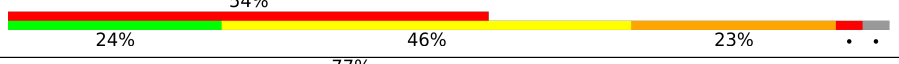

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	299	49% 25% 44% 24% . .
1	B	299	52% 20% 49% 24% . .
1	C	299	51% 27% 44% 23% . .
1	D	299	48% 24% 46% 24% . .
1	E	299	49% 25% 42% 26% . .
1	F	299	53% 24% 43% 24% 5% .
1	G	299	50% 21% 49% 21% 5% .
1	H	299	51% 19% 47% 27% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	299	
1	J	299	
1	K	299	
1	L	299	
1	M	299	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 29211 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.

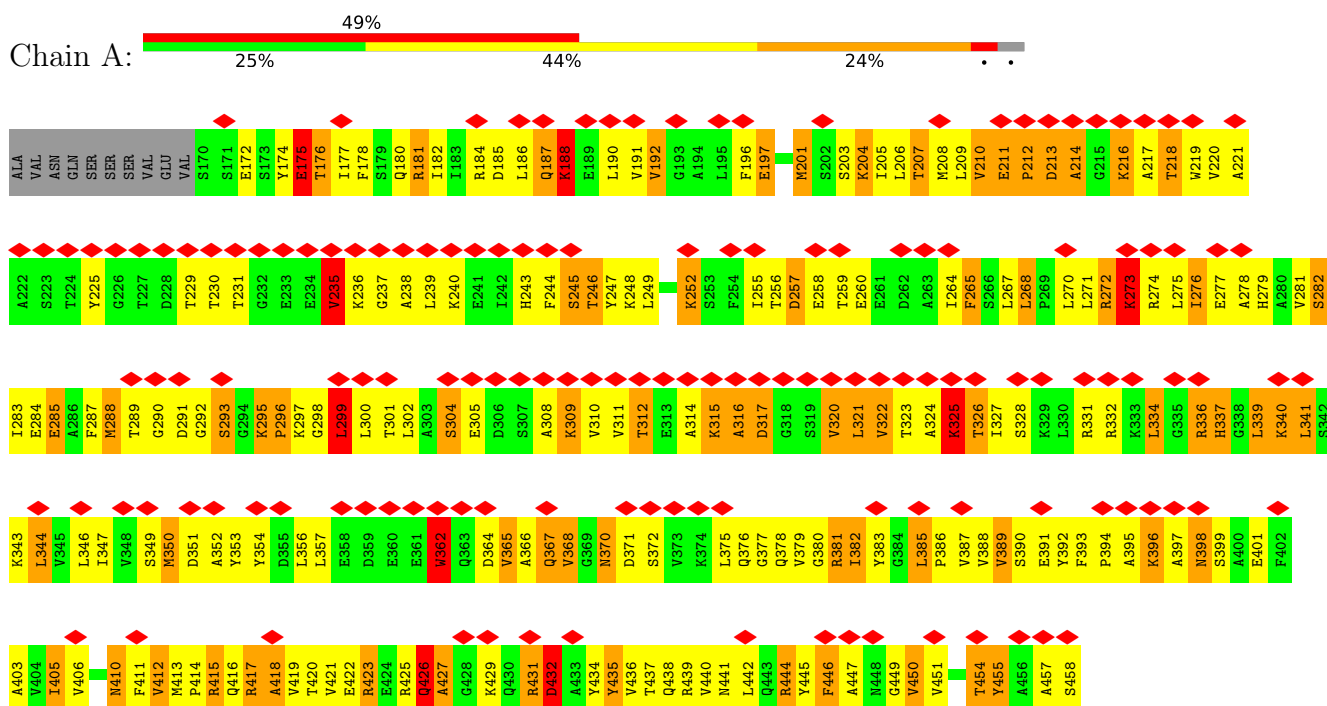
- Molecule 1 is a protein called Major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	289	2247	1425	379	438	5	0	0
1	B	289	2247	1425	379	438	5	0	0
1	C	289	2247	1425	379	438	5	0	0
1	D	289	2247	1425	379	438	5	0	0
1	E	289	2247	1425	379	438	5	0	0
1	F	289	2247	1425	379	438	5	0	0
1	G	289	2247	1425	379	438	5	0	0
1	H	289	2247	1425	379	438	5	0	0
1	I	289	2247	1425	379	438	5	0	0
1	J	289	2247	1425	379	438	5	0	0
1	K	289	2247	1425	379	438	5	0	0
1	L	289	2247	1425	379	438	5	0	0
1	M	289	2247	1425	379	438	5	0	0

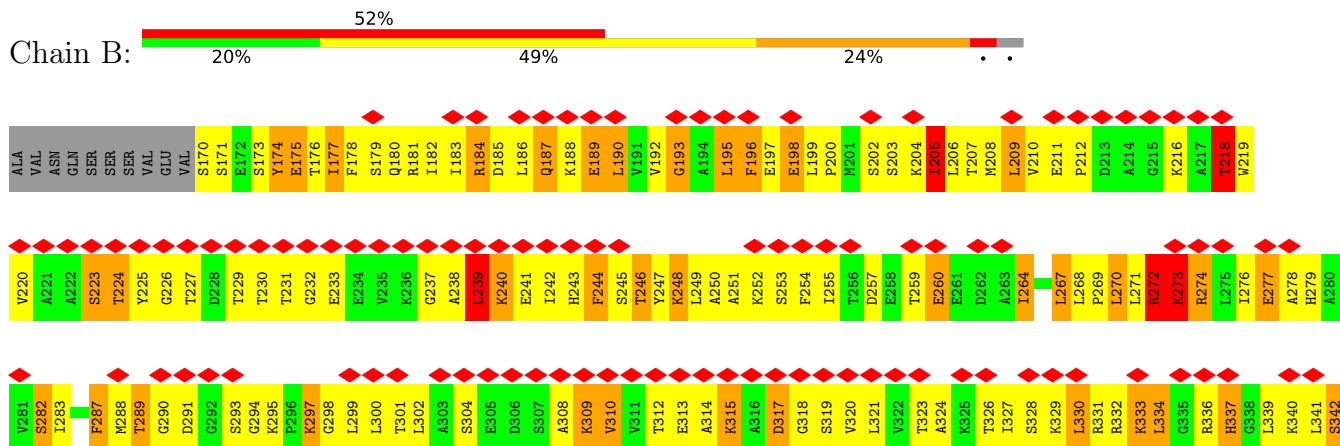
3 Residue-property plots

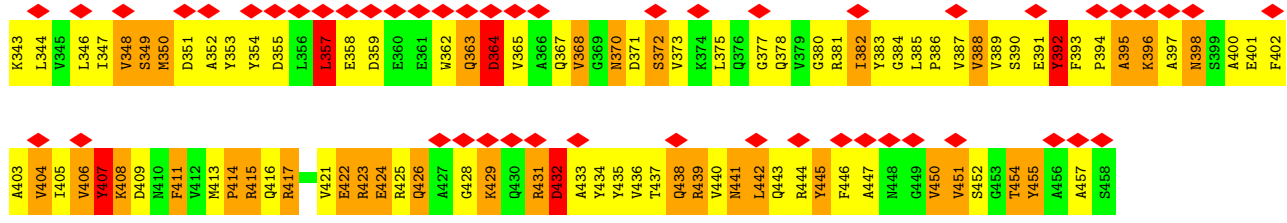
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

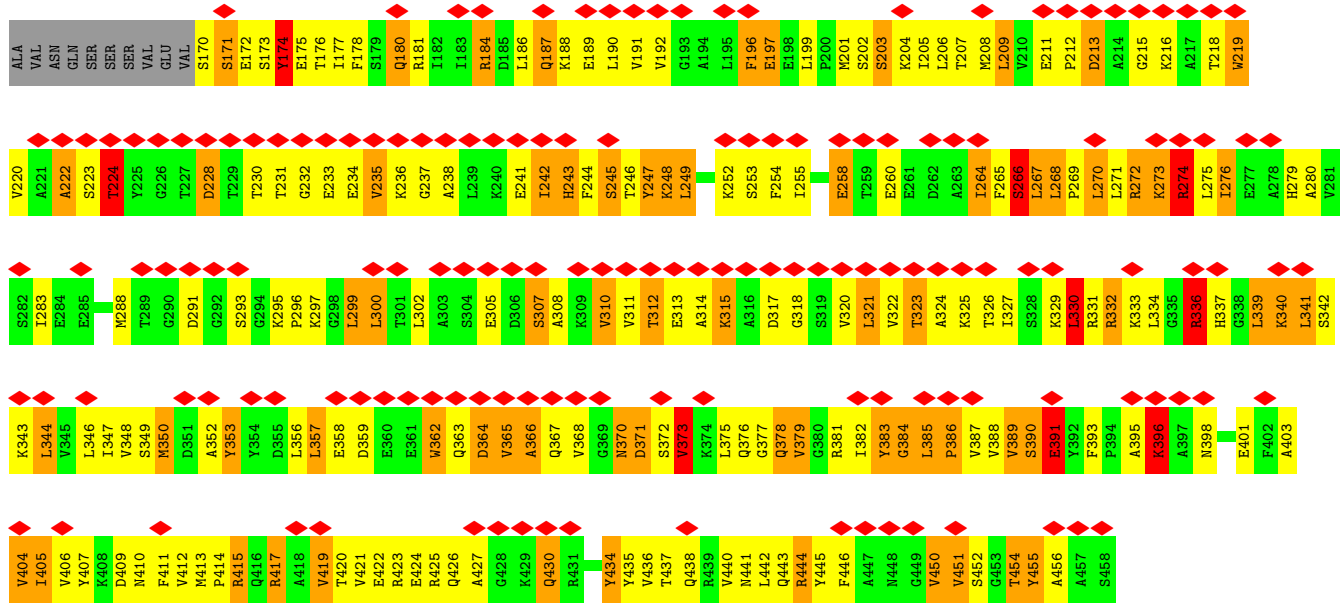


- Molecule 1: Major capsid protein

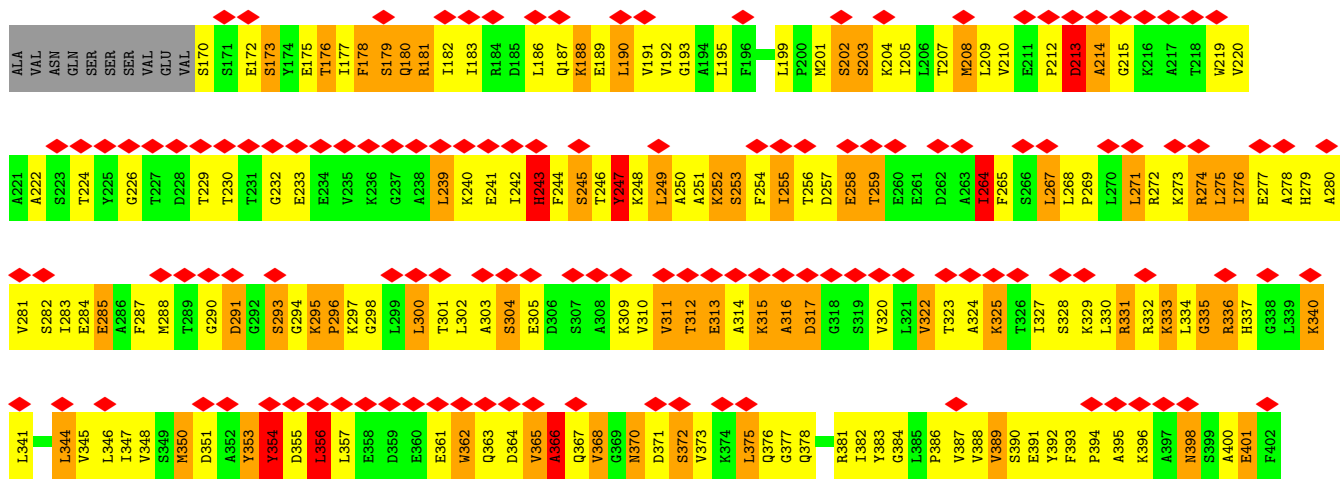


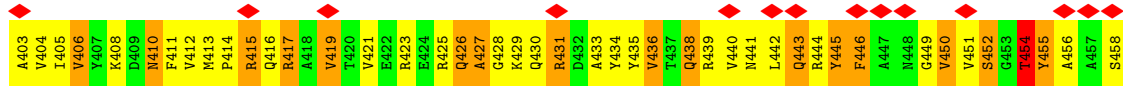


• Molecule 1: Major capsid protein

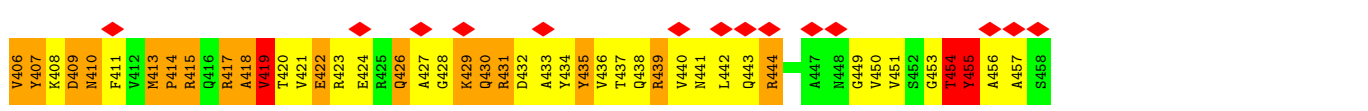
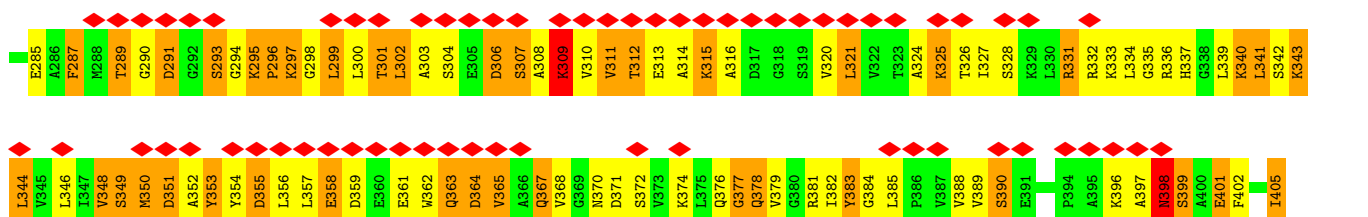
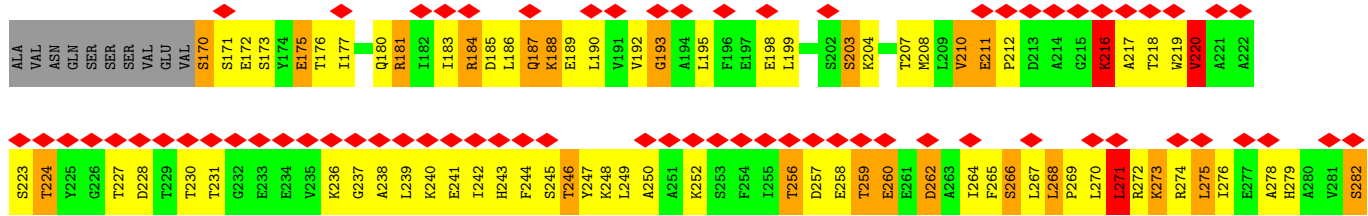
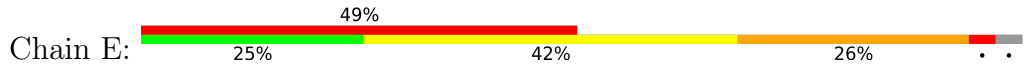


• Molecule 1: Major capsid protein

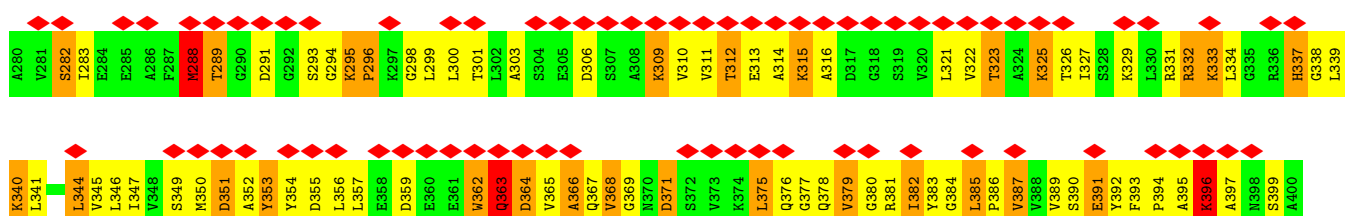
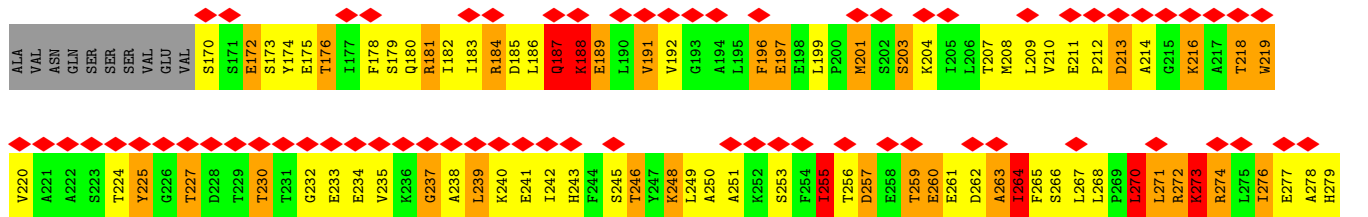
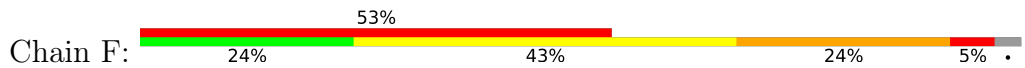




• Molecule 1: Major capsid protein

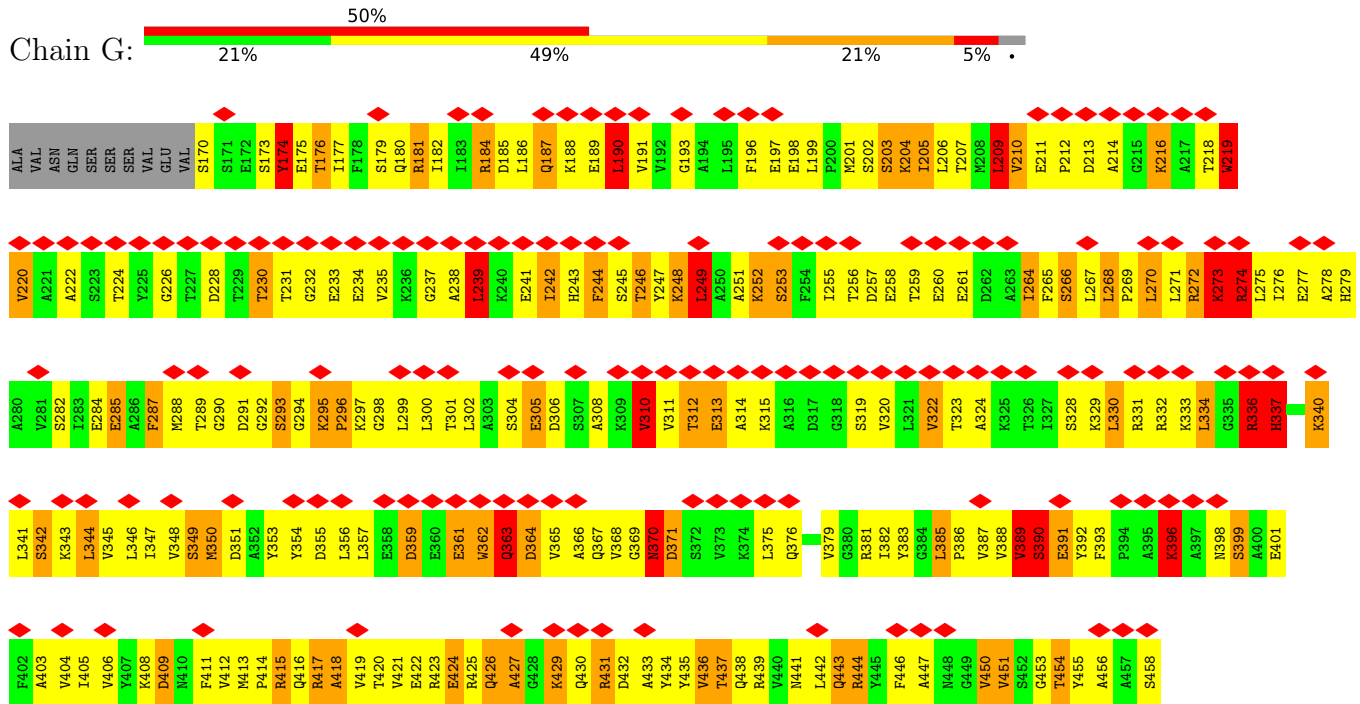


• Molecule 1: Major capsid protein

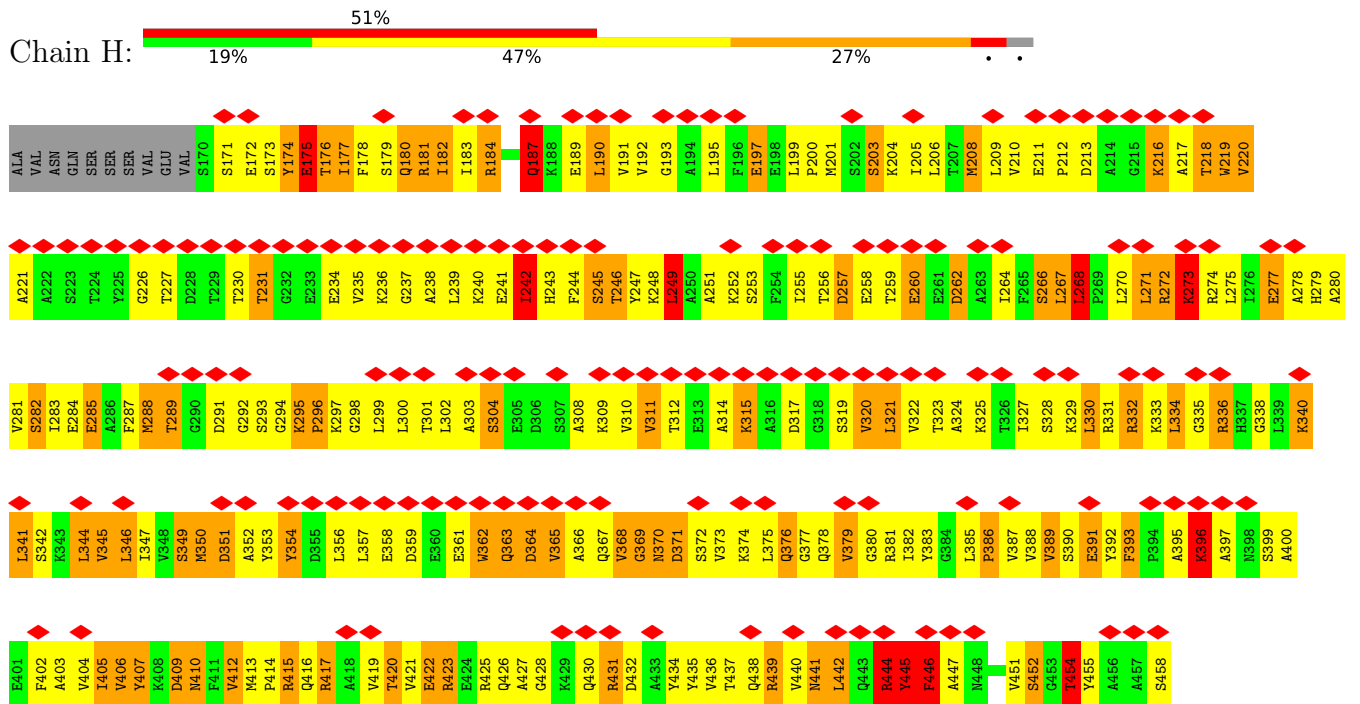


• Molecule 1: Major capsid protein

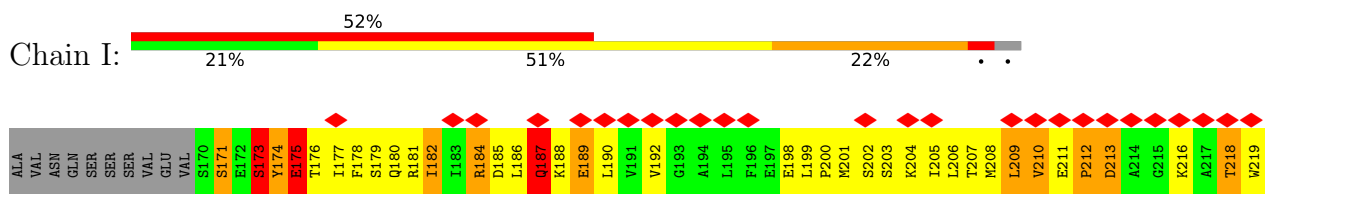


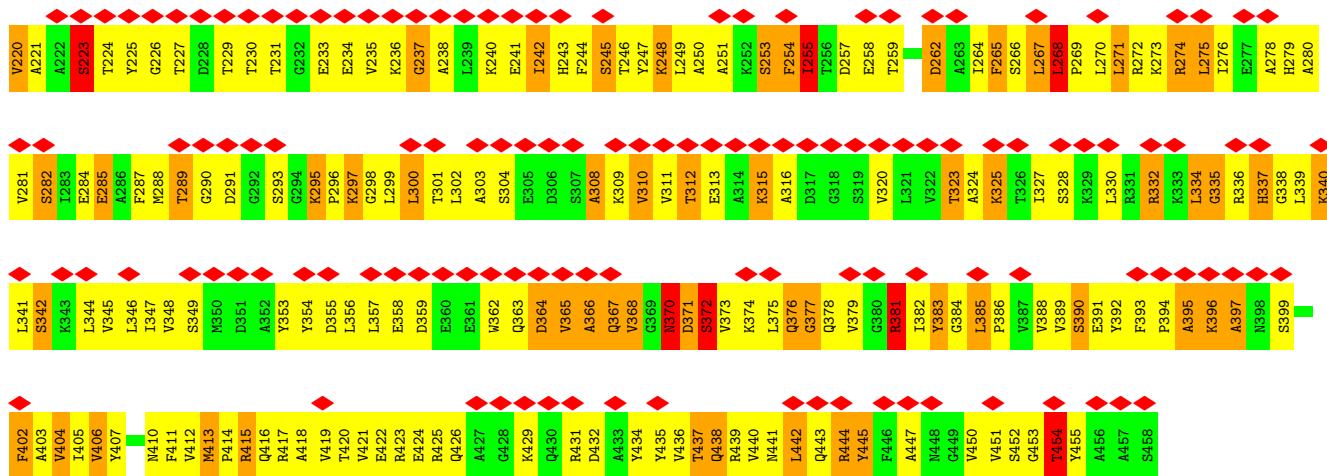


• Molecule 1: Major capsid protein

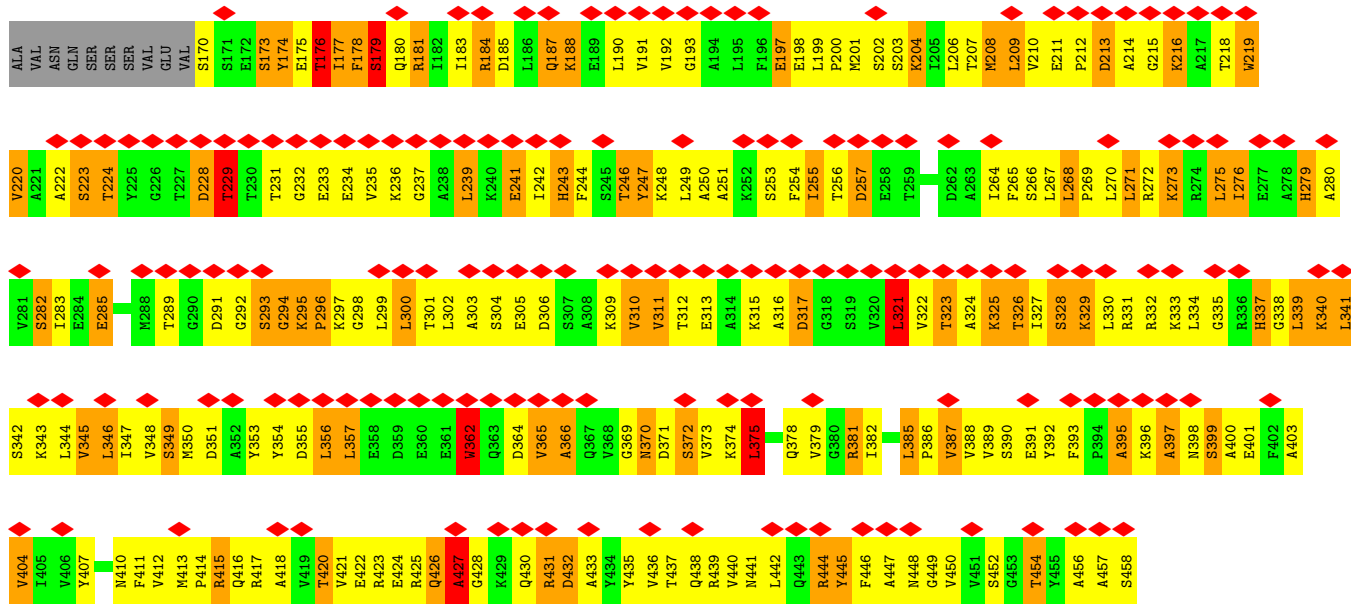
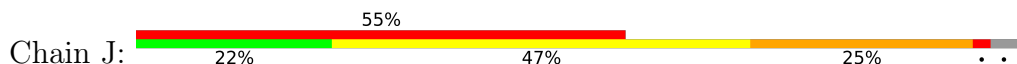


• Molecule 1: Major capsid protein

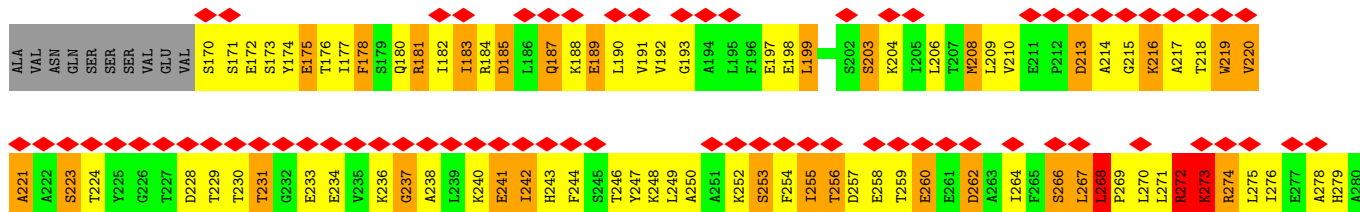


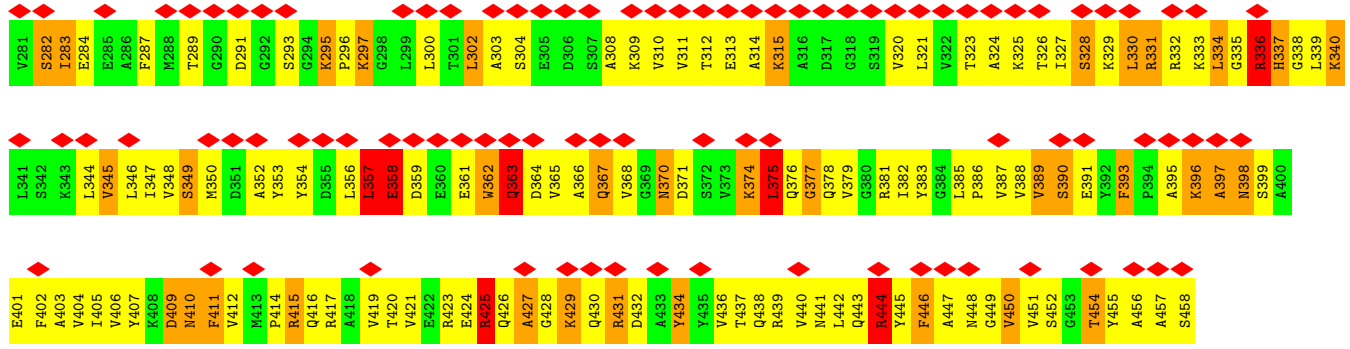


• Molecule 1: Major capsid protein

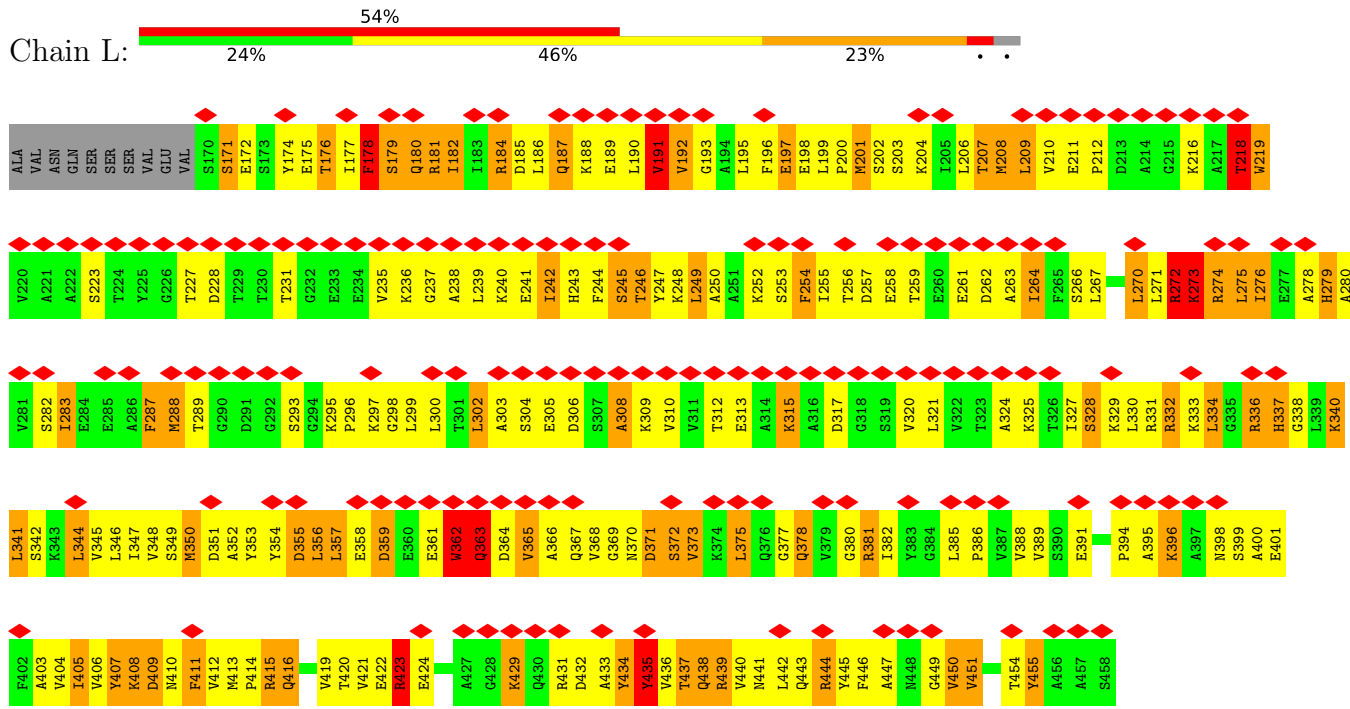


• Molecule 1: Major capsid protein

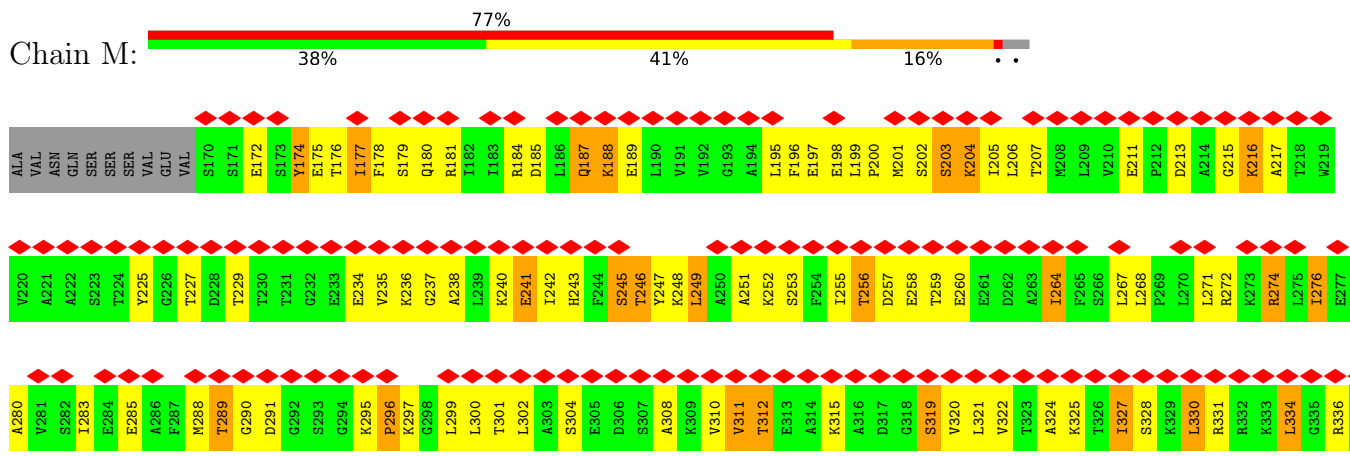




• Molecule 1: Major capsid protein



• Molecule 1: Major capsid protein



K340	L341	S342	K343	L344	V345	L346	I347	V348	S349	M350	D351	Y352	Y353	Y354	D355	L356	L357	E358	D359	E360	E361	W362	Q363	D364	V365	A366	Q367	V368	G369	N370	D371	S372	V373	K374	L375	Q376	G377	Q378	V379	G380	R381	I382	Y383	G384	L385	P386	V387	V388	V389	S390	E391	Y392	F393	P394	A395	K396	A397	N398	S399
A400	E401	F402	A403	V404	I405	V406	Y407	K408	D409	R410	F411	V412	M413	P414	R415	Q416	R417	A418	V419	T420	V421	E422	R423	E424	R425	Q426	A427	G428	K429	Q430	R431	D432	A433	Y434	Y435	V436	T437	Q438	R439	V440	N441	L442	Q443	R444	Y445	F446	A447	N448	C449	V450	V451	S452	G453	T454	Y455	A456	A457	S458	

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	8083	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 ($e^-/\text{\AA}^2$)	30	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON II (4k x 4k)	Depositor
Maximum map value	32.443	Depositor
Minimum map value	-15.158	Depositor
Average map value	0.119	Depositor
Map value standard deviation	4.035	Depositor
Recommended contour level	14.0	Depositor
Map size (Å)	845.05493, 845.05493, 845.05493	wwPDB
Map dimensions	801, 801, 801	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.055, 1.055, 1.055	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	0.71	1/2282 (0.0%)	0.90	4/3082 (0.1%)
1	B	0.71	1/2282 (0.0%)	0.93	4/3082 (0.1%)
1	C	0.68	1/2282 (0.0%)	0.94	5/3082 (0.2%)
1	D	0.72	2/2282 (0.1%)	0.95	6/3082 (0.2%)
1	E	0.72	0/2282	0.91	3/3082 (0.1%)
1	F	0.67	0/2282	0.92	5/3082 (0.2%)
1	G	0.77	1/2282 (0.0%)	0.97	7/3082 (0.2%)
1	H	0.77	1/2282 (0.0%)	0.95	8/3082 (0.3%)
1	I	0.69	1/2282 (0.0%)	0.86	2/3082 (0.1%)
1	J	0.70	2/2282 (0.1%)	0.94	9/3082 (0.3%)
1	K	0.71	1/2282 (0.0%)	0.92	4/3082 (0.1%)
1	L	0.68	1/2282 (0.0%)	0.94	4/3082 (0.1%)
1	M	0.53	0/2282	0.79	3/3082 (0.1%)
All	All	0.70	12/29666 (0.0%)	0.92	64/40066 (0.2%)

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	10
1	B	0	15
1	C	0	14
1	D	0	11
1	E	0	15
1	F	0	11
1	G	0	13
1	H	0	11
1	I	0	14
1	J	0	8
1	K	0	14
1	L	0	15

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	M	0	9
All	All	0	160

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	J	294	GLY	C-N	-11.25	1.08	1.34
1	G	210	VAL	C-N	-9.71	1.11	1.34
1	H	267	LEU	CA-CB	-9.11	1.32	1.53
1	I	267	LEU	CA-CB	-7.72	1.35	1.53
1	L	423	ARG	CA-CB	-7.22	1.38	1.53

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	188	LYS	C-N-CA	7.91	141.47	121.70
1	D	249	LEU	CB-CG-CD1	-7.43	98.37	111.00
1	G	190	LEU	CA-CB-CG	-7.00	99.20	115.30
1	H	393	PHE	CB-CG-CD1	6.76	125.53	120.80
1	L	435	TYR	CA-CB-CG	6.72	126.16	113.40

There are no chirality outliers.

5 of 160 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	187	GLN	Peptide
1	A	188	LYS	Peptide
1	A	197	GLU	Peptide
1	A	210	VAL	Peptide
1	A	220	VAL	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	2247	0	2275	178	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2247	0	2275	171	0
1	C	2247	0	2275	156	0
1	D	2247	0	2275	176	0
1	E	2247	0	2275	154	0
1	F	2247	0	2274	168	0
1	G	2247	0	2274	182	0
1	H	2247	0	2275	209	0
1	I	2247	0	2275	175	0
1	J	2247	0	2274	172	0
1	K	2247	0	2275	179	0
1	L	2247	0	2275	146	0
1	M	2247	0	2275	116	0
All	All	29211	0	29572	2121	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 36.

The worst 5 of 2121 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:352:ALA:O	1:A:356:LEU:HB2	1.54	1.07
1:K:327:ILE:O	1:K:331:ARG:HB2	1.60	1.01
1:L:351:ASP:O	1:L:355:ASP:HB2	1.64	0.95
1:B:350:MET:O	1:B:354:TYR:HB2	1.73	0.89
1:E:377:GLY:H	1:E:378:GLN:HG2	1.42	0.82

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	287/299 (96%)	188 (66%)	71 (25%)	28 (10%)	0	10
1	B	287/299 (96%)	190 (66%)	68 (24%)	29 (10%)	0	9
1	C	287/299 (96%)	176 (61%)	83 (29%)	28 (10%)	0	10
1	D	287/299 (96%)	181 (63%)	71 (25%)	35 (12%)	0	6
1	E	287/299 (96%)	191 (67%)	71 (25%)	25 (9%)	1	11
1	F	287/299 (96%)	190 (66%)	63 (22%)	34 (12%)	0	6
1	G	287/299 (96%)	193 (67%)	65 (23%)	29 (10%)	0	9
1	H	287/299 (96%)	185 (64%)	71 (25%)	31 (11%)	0	8
1	I	287/299 (96%)	180 (63%)	83 (29%)	24 (8%)	1	12
1	J	287/299 (96%)	188 (66%)	77 (27%)	22 (8%)	1	13
1	K	287/299 (96%)	200 (70%)	58 (20%)	29 (10%)	0	9
1	L	287/299 (96%)	190 (66%)	72 (25%)	25 (9%)	1	11
1	M	287/299 (96%)	190 (66%)	80 (28%)	17 (6%)	1	17
All	All	3731/3887 (96%)	2442 (66%)	933 (25%)	356 (10%)	1	10

5 of 356 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	239	LEU
1	A	446	PHE
1	B	233	GLU
1	B	273	LYS
1	B	390	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	241/250 (96%)	159 (66%)	82 (34%)	0	1
1	B	241/250 (96%)	155 (64%)	86 (36%)	0	1
1	C	241/250 (96%)	165 (68%)	76 (32%)	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	241/250 (96%)	168 (70%)	73 (30%)	0	2
1	E	241/250 (96%)	156 (65%)	85 (35%)	0	1
1	F	241/250 (96%)	153 (64%)	88 (36%)	0	1
1	G	241/250 (96%)	156 (65%)	85 (35%)	0	1
1	H	241/250 (96%)	161 (67%)	80 (33%)	0	2
1	I	241/250 (96%)	164 (68%)	77 (32%)	0	2
1	J	241/250 (96%)	158 (66%)	83 (34%)	0	1
1	K	241/250 (96%)	159 (66%)	82 (34%)	0	1
1	L	241/250 (96%)	143 (59%)	98 (41%)	0	0
1	M	241/250 (96%)	174 (72%)	67 (28%)	0	3
All	All	3133/3250 (96%)	2071 (66%)	1062 (34%)	1	1

5 of 1062 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	255	ILE
1	L	332	ARG
1	L	249	LEU
1	M	362	TRP
1	E	413	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	J	279	HIS
1	L	410	ASN
1	K	180	GLN
1	K	410	ASN
1	E	363	GLN

5.3.3 RNA

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	G	1
1	J	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	210:VAL	C	211:GLU	N	1.11
1	J	294:GLY	C	295:LYS	N	1.08

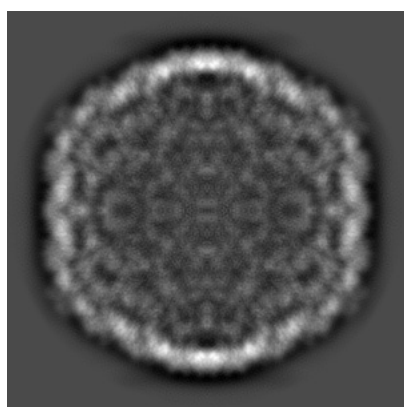
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-20099. These allow visual inspection of the internal detail of the map and identification of artifacts.

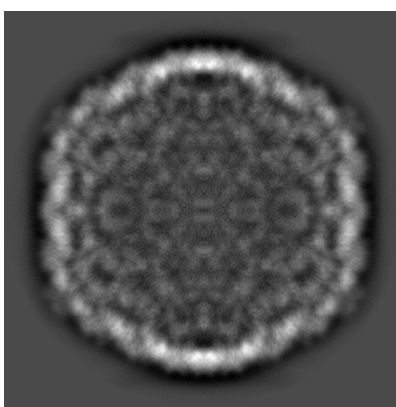
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

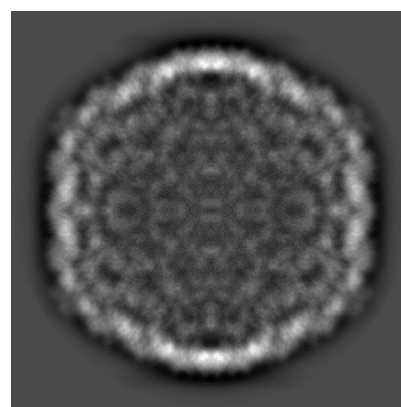
6.1.1 Primary map



X



Y

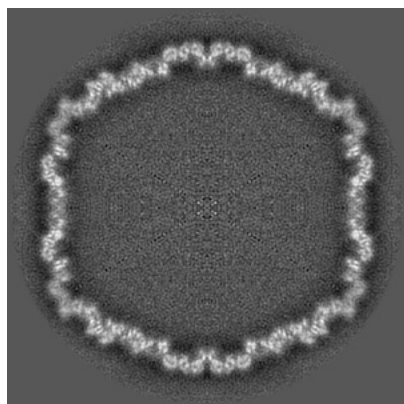


Z

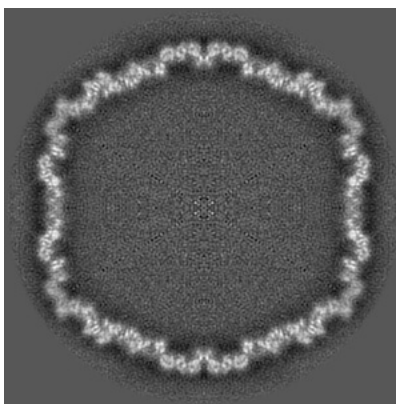
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

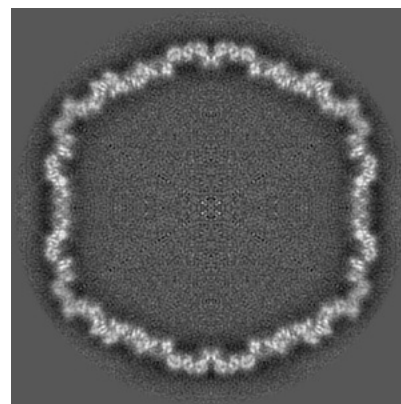
6.2.1 Primary map



X Index: 400



Y Index: 400

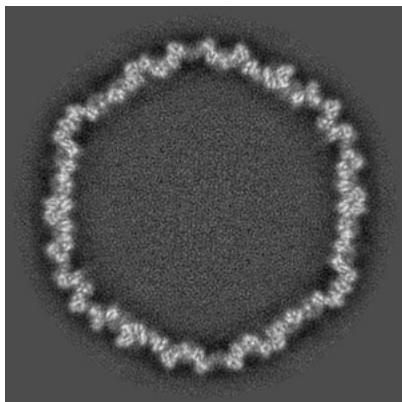


Z Index: 400

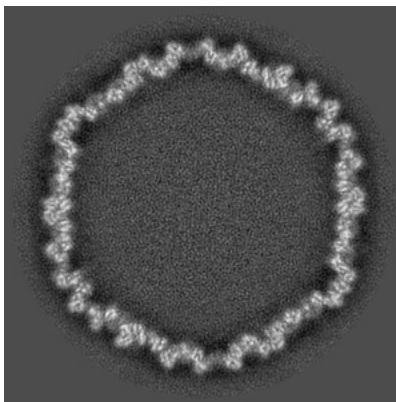
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

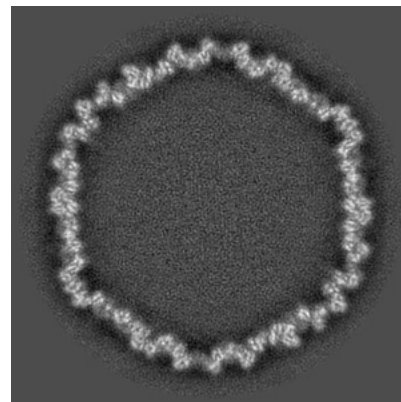
6.3.1 Primary map



X Index: 318



Y Index: 318

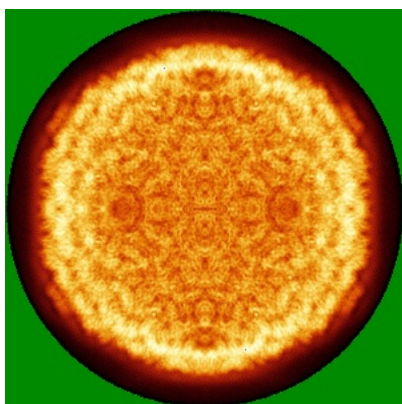


Z Index: 482

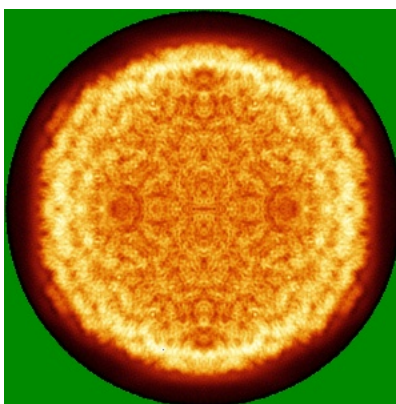
The images above show the largest variance slices of the map in three orthogonal directions.

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

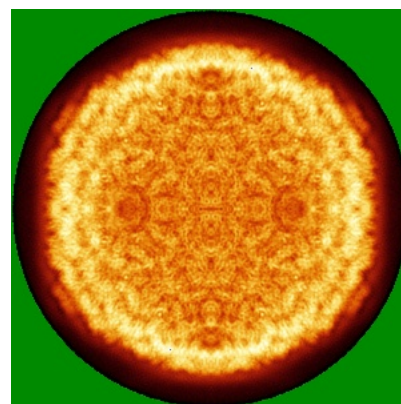
6.4.1 Primary map



X



Y

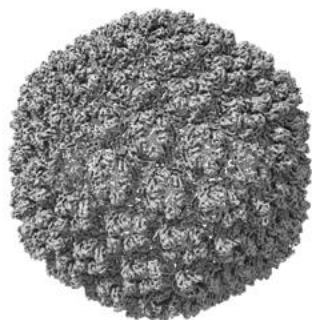


Z

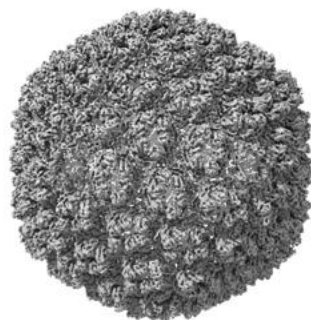
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 14.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

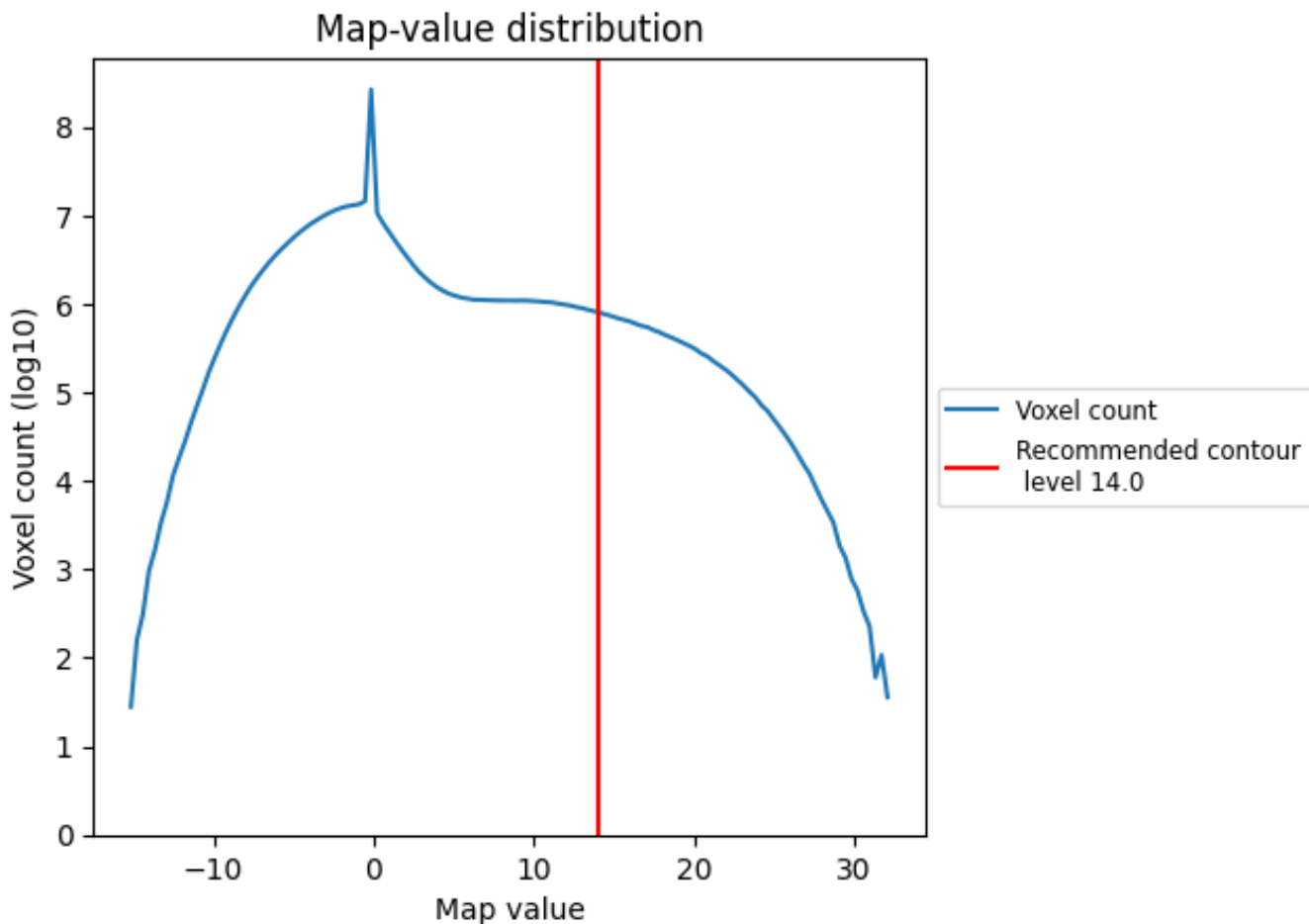
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

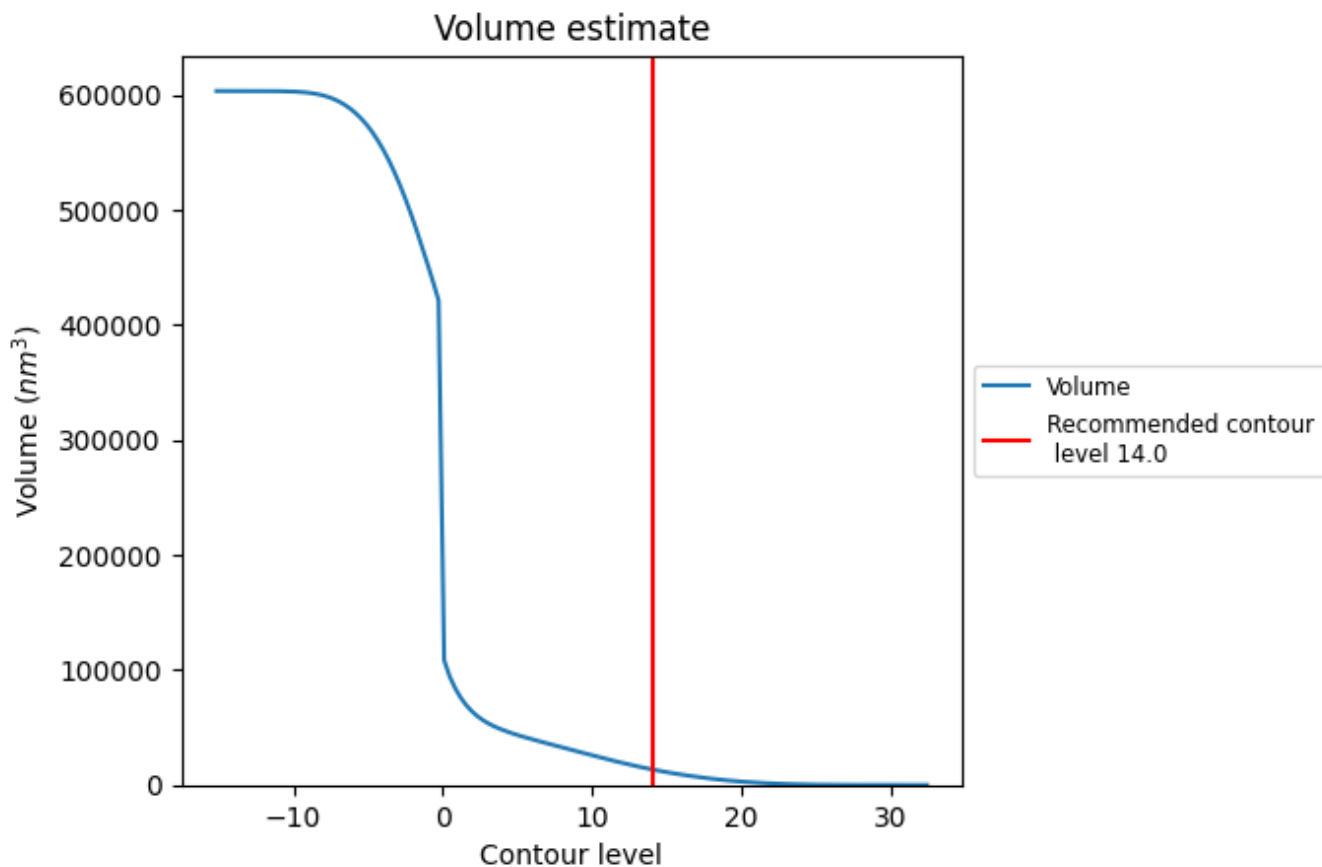
This section contains the results of statistical analysis of the map.

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.

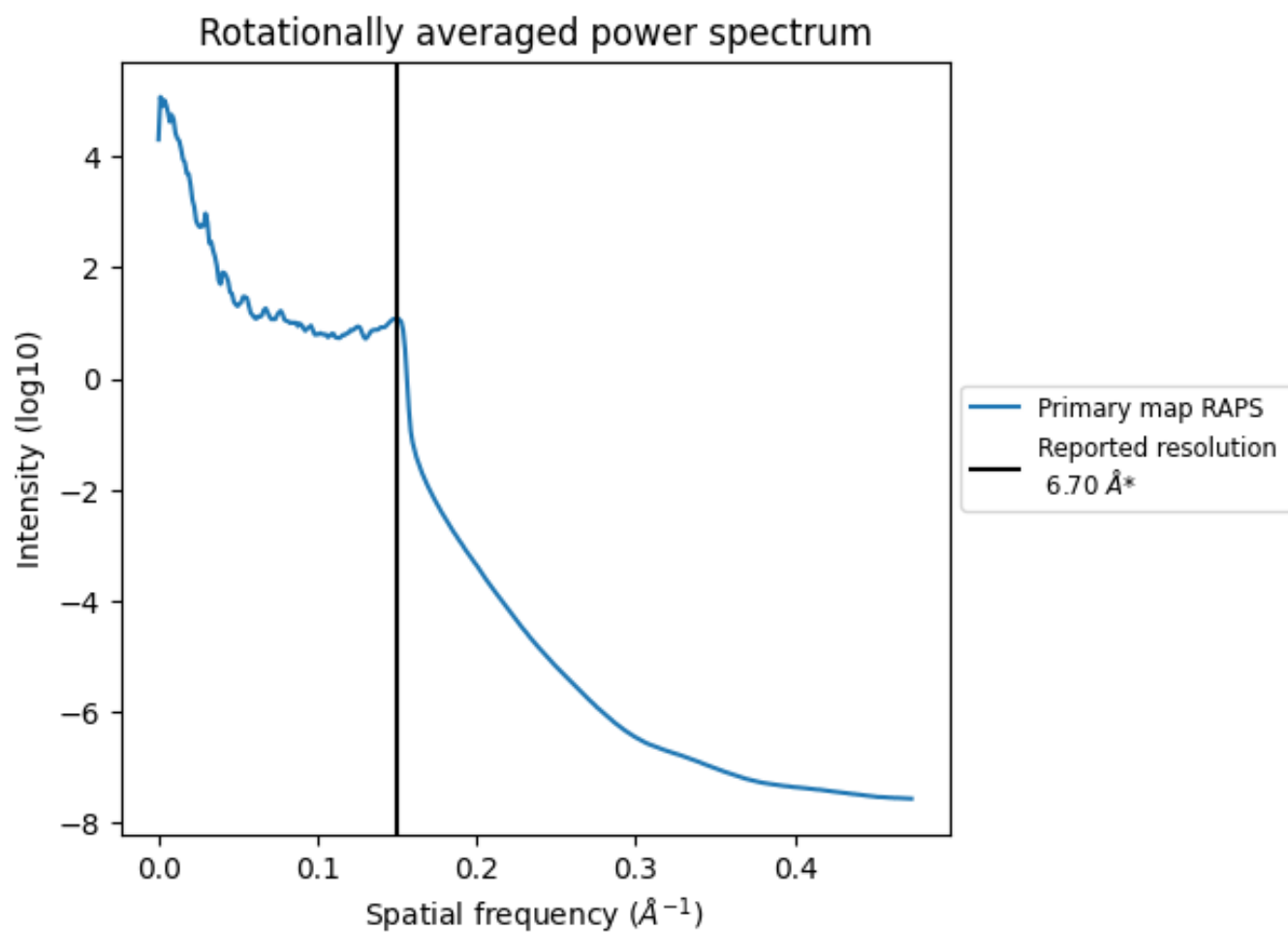
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 13553 nm³; this corresponds to an approximate mass of 12243 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](#)

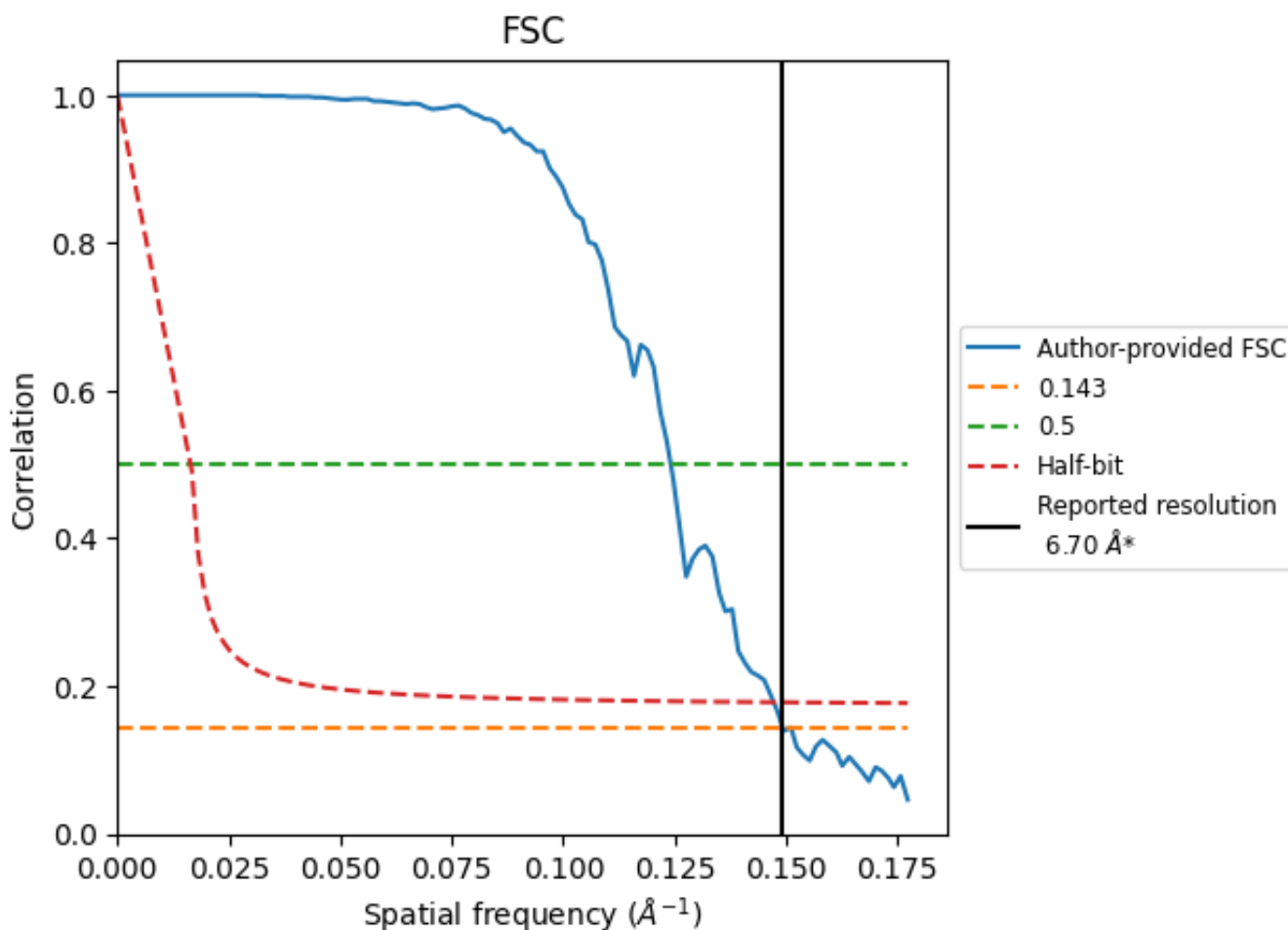


*Reported resolution corresponds to spatial frequency of 0.149\AA^{-1}

8 Fourier-Shell correlation [i](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC [i](#)



*Reported resolution corresponds to spatial frequency of 0.149 Å⁻¹

8.2 Resolution estimates [i](#)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	6.70	-	-
Author-provided FSC curve	6.70	8.06	6.79
Unmasked-calculated*	-	-	-

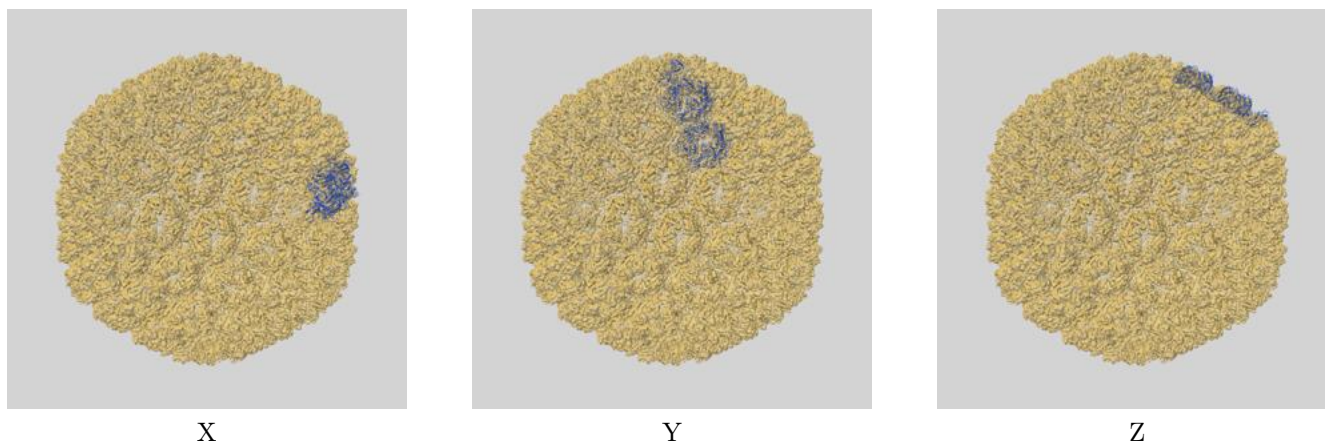
*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

9 Map-model fit [i](#)

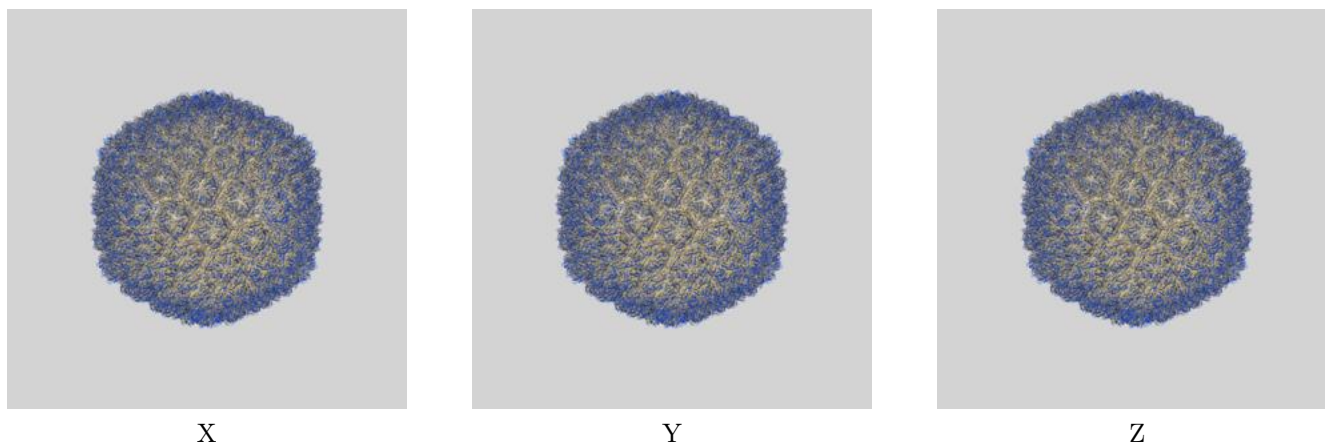
This section contains information regarding the fit between EMDB map EMD-20099 and PDB model 6OKB. Per-residue inclusion information can be found in section 3 on page 5.

9.1 Map-model overlays

9.1.1 Map-model overlay [i](#)

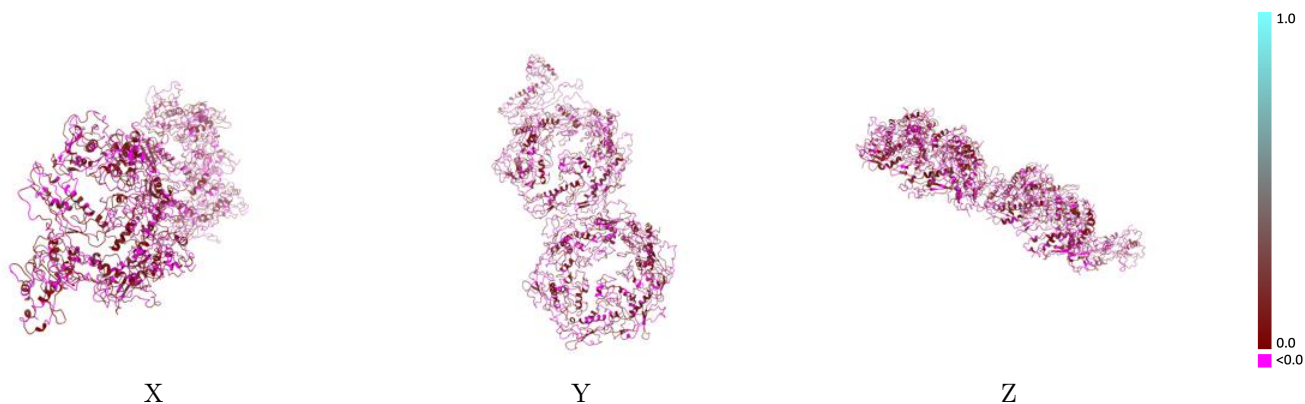


9.1.2 Map-model assembly overlay [i](#)



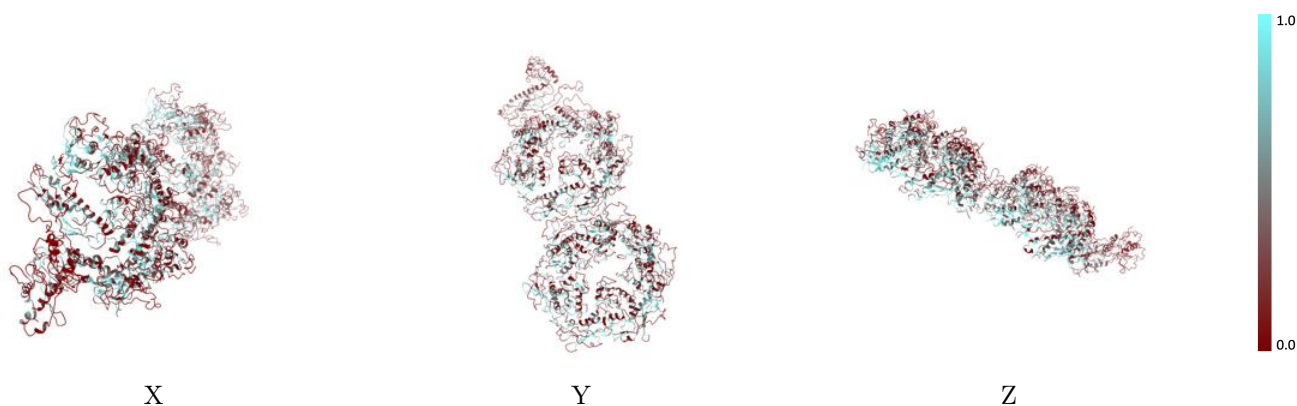
The images above show the 3D surface view of the map at the recommended contour level 14.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



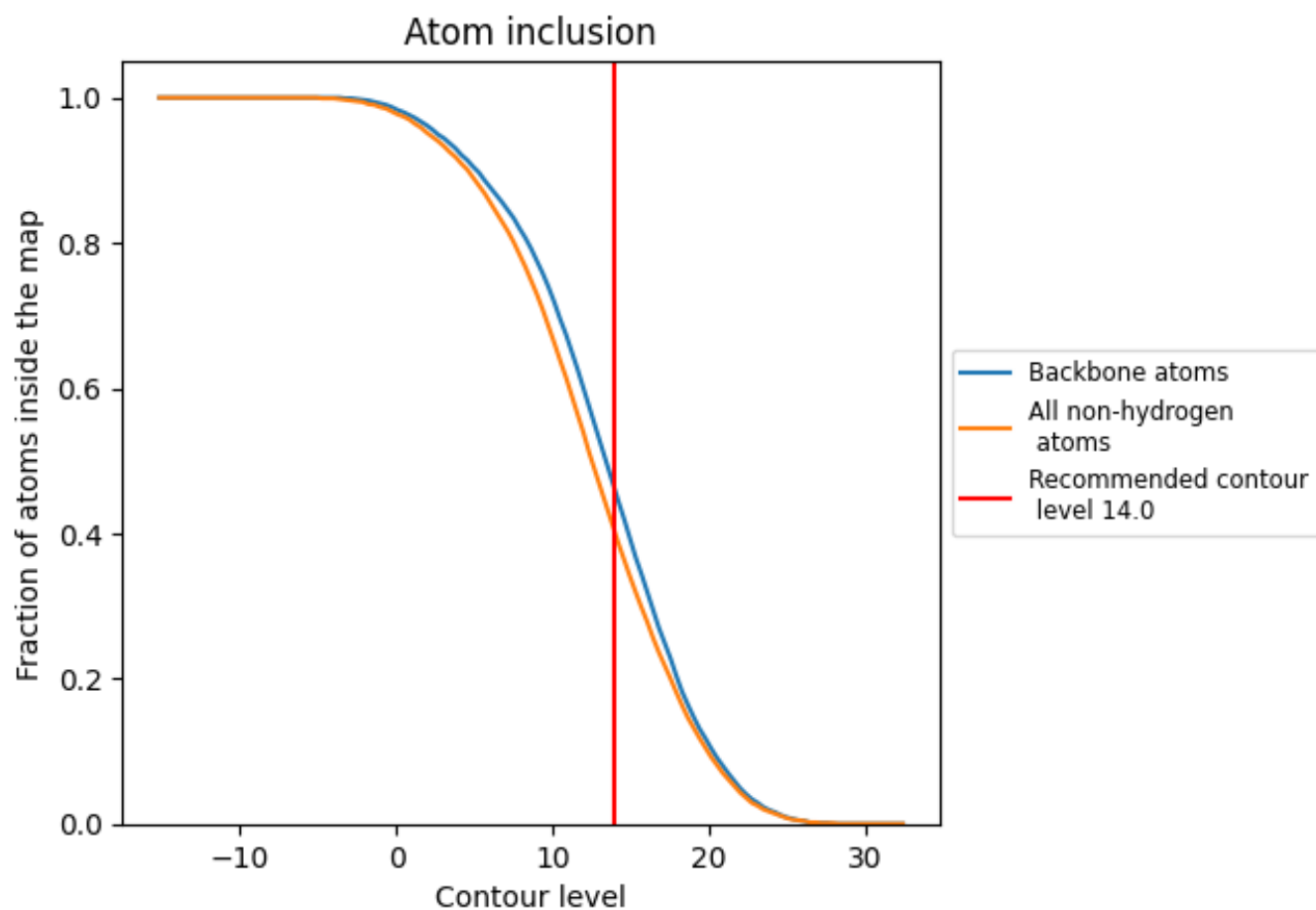
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (14.0).





























9.4 Atom inclusion [i](#)



At the recommended contour level, 46% of all backbone atoms, 40% 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 (14.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.4000	 0.0760
A	 0.4330	 0.0880
B	 0.4220	 0.0790
C	 0.4180	 0.0670
D	 0.4510	 0.0790
E	 0.4350	 0.0710
F	 0.4110	 0.0720
G	 0.4090	 0.0660
H	 0.4080	 0.0760
I	 0.4130	 0.0740
J	 0.4050	 0.0660
K	 0.4020	 0.0800
L	 0.3980	 0.0830
M	 0.1970	 0.0830

