



wwPDB EM Validation Summary Report ⓘ

Nov 22, 2022 – 06:58 PM JST

PDB ID : 7ETJ
EMDB ID : EMD-31298
Title : C5 portal vertex in the partially-enveloped virion capsid
Authors : Li, Z.; Yu, X.
Deposited on : 2021-05-13
Resolution : 4.00 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:

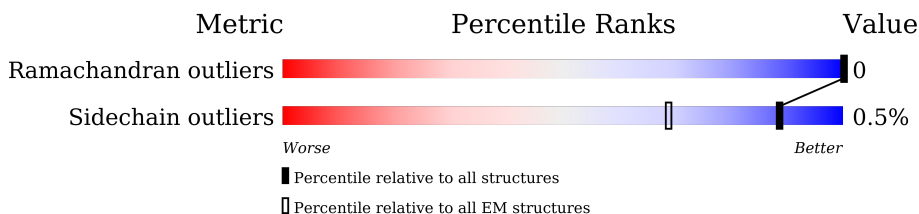
EMDB validation analysis : 0.0.1.dev43
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.3

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 4.00 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)
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	I	306	
1	h	306	
1	n	306	
1	o	306	
2	H	2241	
2	P	2241	
3	g	290	
3	m	290	
4	M	594	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	N	642	 12% 88%
5	O	642	 11% 89%
6	1	1048	 10% 27% 73%
7	R	75	 27% 84% 16%
7	S	75	 16% 84% 16%
7	T	75	 49% 81% 16%
7	i	75	 40% 81% 16%
7	j	75	 35% 83% 16%
8	B	1370	 97%
8	C	1370	 96%
8	D	1370	 94% 5%
8	Y	1370	 8% 98%
8	Z	1370	 97%
8	a	1370	 92% 7%

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 86613 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 Triplex capsid protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	h	300	Total 2384	C 1528	N 415	O 424	S 17	0	0
1	I	301	Total 2389	C 1531	N 417	O 424	S 17	0	0
1	n	294	Total 2329	C 1499	N 401	O 410	S 19	0	0
1	o	289	Total 2291	C 1473	N 393	O 407	S 18	0	0

- Molecule 2 is a protein called Large tegument protein deneddylase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	H	20	Total 172	C 110	N 32	O 29	S 1	0	0
2	P	20	Total 172	C 110	N 32	O 29	S 1	0	0

- Molecule 3 is a protein called Triplex capsid protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	g	227	Total 1822	C 1171	N 319	O 321	S 11	0	0
3	m	290	Total 2325	C 1485	N 411	O 417	S 12	0	0

- Molecule 4 is a protein called Capsid vertex component 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	M	468	Total 3848	C 2408	N 740	O 686	S 14	0	0

- Molecule 5 is a protein called Capsid vertex component 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	N	76	Total	C	N	O	S	0	0
			648	408	127	109	4		
5	O	69	Total	C	N	O	S	0	0
			589	371	113	102	3		

- Molecule 6 is a protein called ORFL92C_UL32.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	1	284	Total	C	N	O	S	0	0
			2320	1463	425	420	12		

- Molecule 7 is a protein called Small capsomere-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	R	63	Total	C	N	O	S	0	0
			513	321	97	91	4		
7	S	63	Total	C	N	O	S	0	0
			513	321	97	91	4		
7	T	63	Total	C	N	O	S	0	0
			513	321	97	91	4		
7	i	63	Total	C	N	O	S	0	0
			513	321	97	91	4		
7	j	63	Total	C	N	O	S	0	0
			513	321	97	91	4		

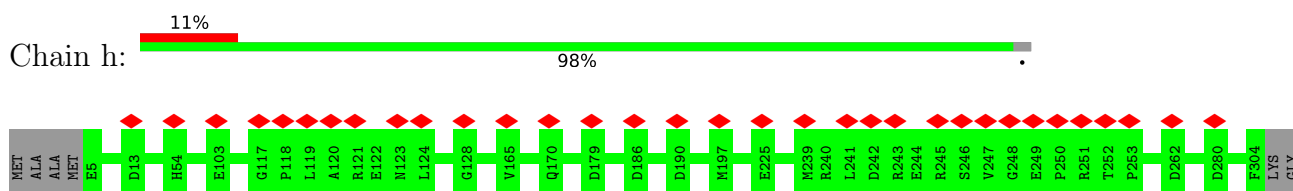
- Molecule 8 is a protein called Major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	a	1275	Total	C	N	O	S	0	0
			10079	6410	1756	1855	58		
8	B	1339	Total	C	N	O	S	0	0
			10617	6762	1838	1956	61		
8	C	1331	Total	C	N	O	S	0	0
			10536	6711	1827	1938	60		
8	D	1297	Total	C	N	O	S	0	0
			10269	6538	1785	1887	59		
8	Y	1347	Total	C	N	O	S	0	0
			10676	6799	1850	1966	61		
8	Z	1337	Total	C	N	O	S	0	0
			10582	6740	1831	1952	59		

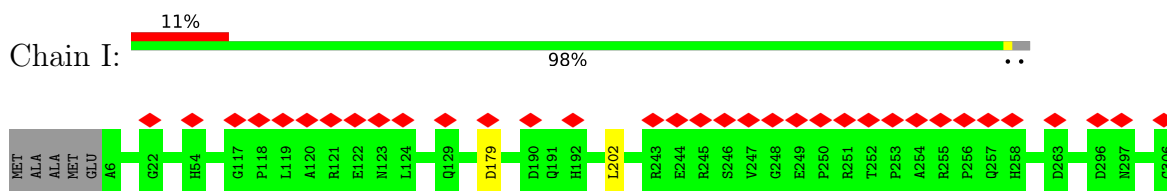
3 Residue-property plots [i](#)

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.

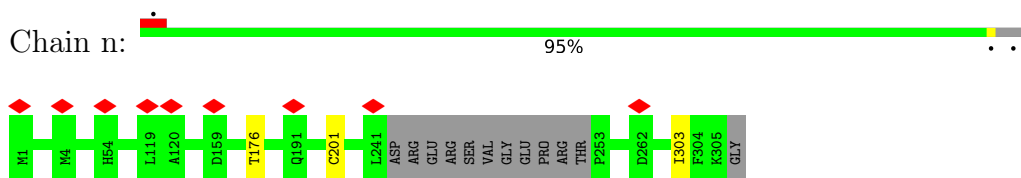
- Molecule 1: Triplex capsid protein 2



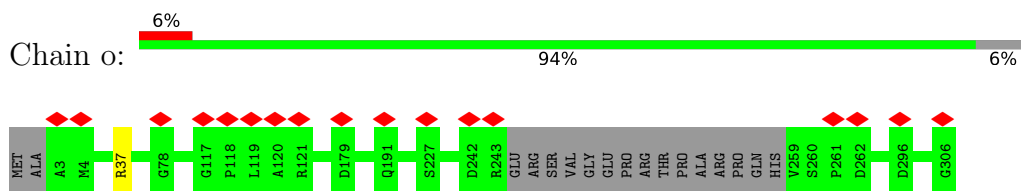
- Molecule 1: Triplex capsid protein 2



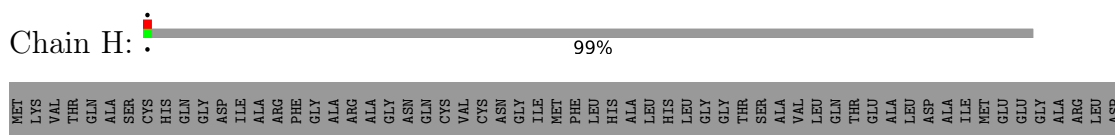
- Molecule 1: Triplex capsid protein 2



- Molecule 1: Triplex capsid protein 2



- Molecule 2: Large tegument protein deneddylase

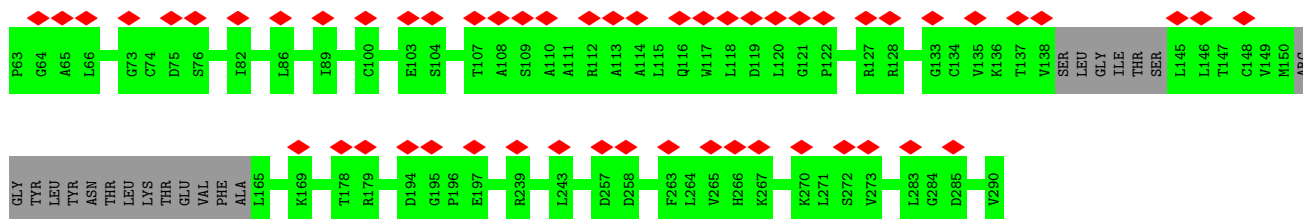


L2221	R2222	Q2223	L2224	A2225	Q2226	S2227	Y2228	D2230	Q2233	H2234	H2235	F2237	Y2239	L2238	L2240	LEU
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----

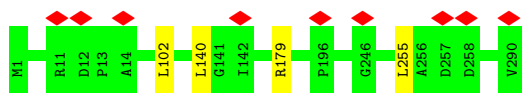
• Molecule 2: Large tegument protein deneddylase

Chain P:  99%

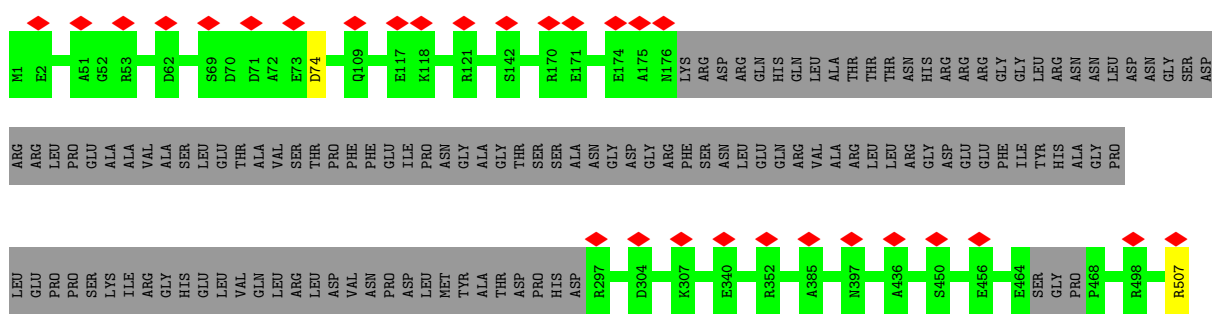
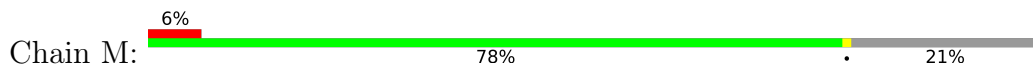
MET	LYS	VAL	THR	GLN	ALA	SER	CYS	HIS	GLY	ASP	ILE	ARG	GLY	PRO	THR	ASP
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



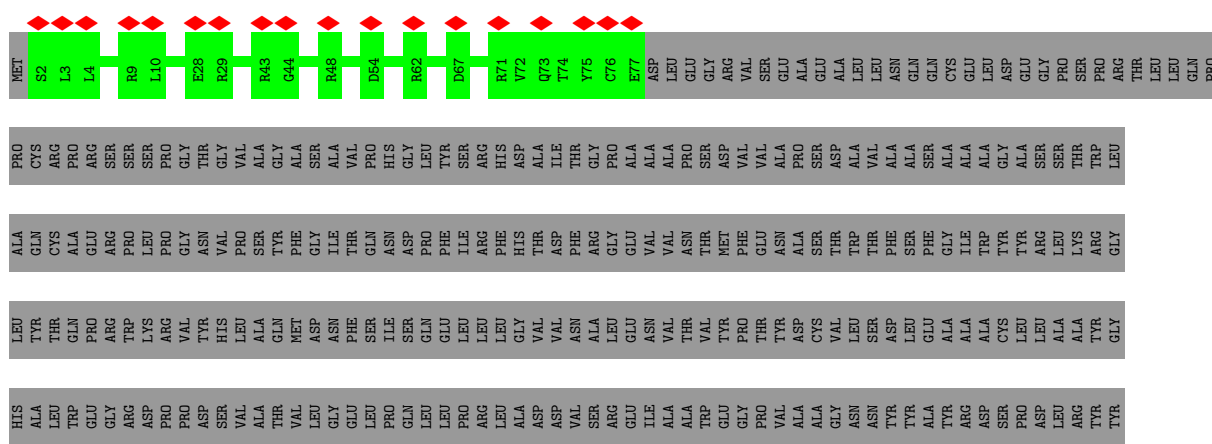
• Molecule 3: Triplex capsid protein 1



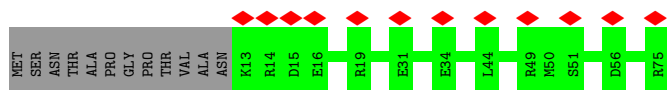
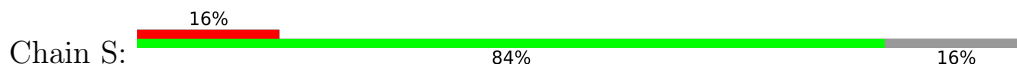
• Molecule 4: Capsid vertex component 1



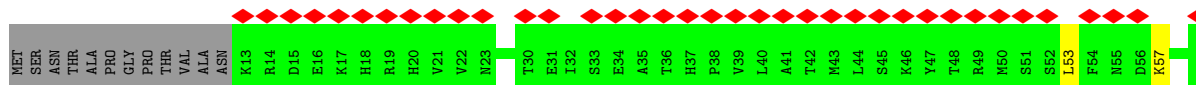
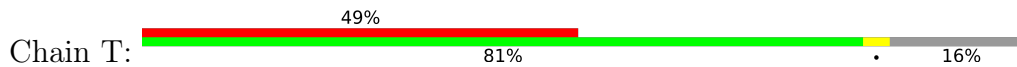
• Molecule 5: Capsid vertex component 2



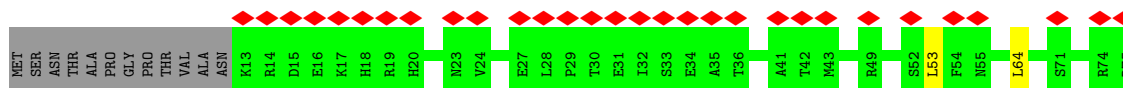
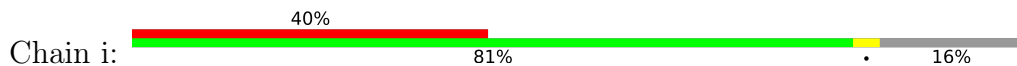
• Molecule 7: Small capsomere-interacting protein



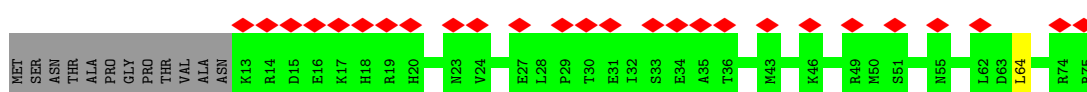
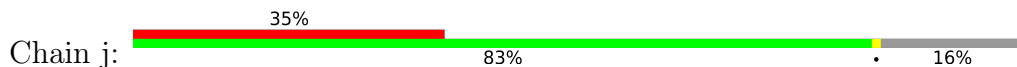
• Molecule 7: Small capsomere-interacting protein



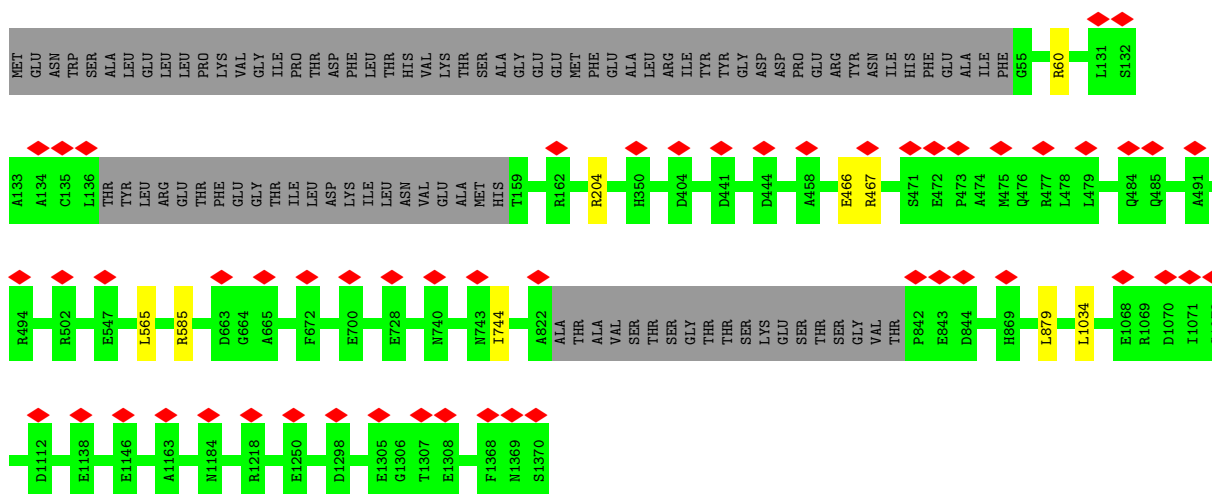
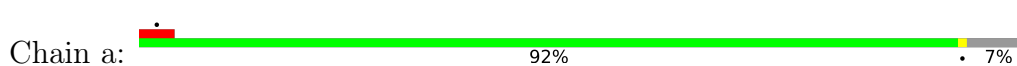
• Molecule 7: Small capsomere-interacting protein



• Molecule 7: Small capsomere-interacting protein

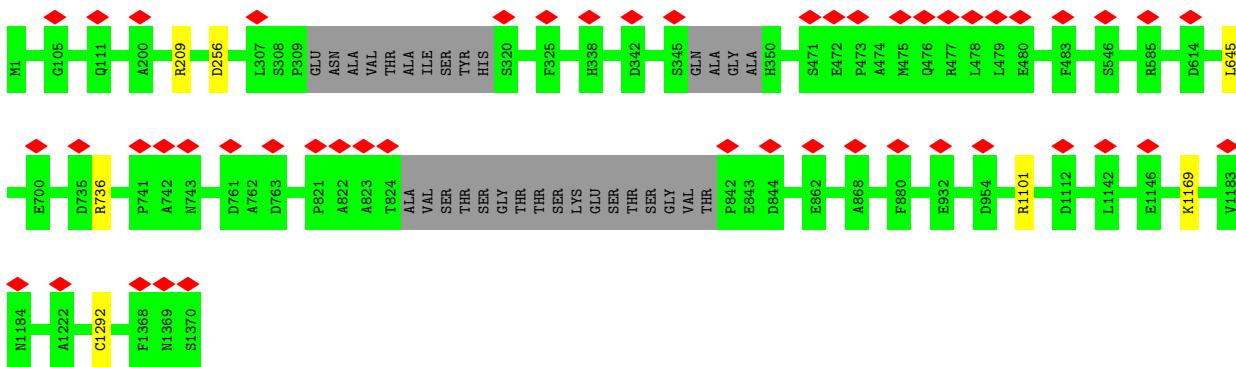
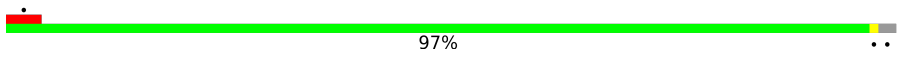


• Molecule 8: Major capsid protein



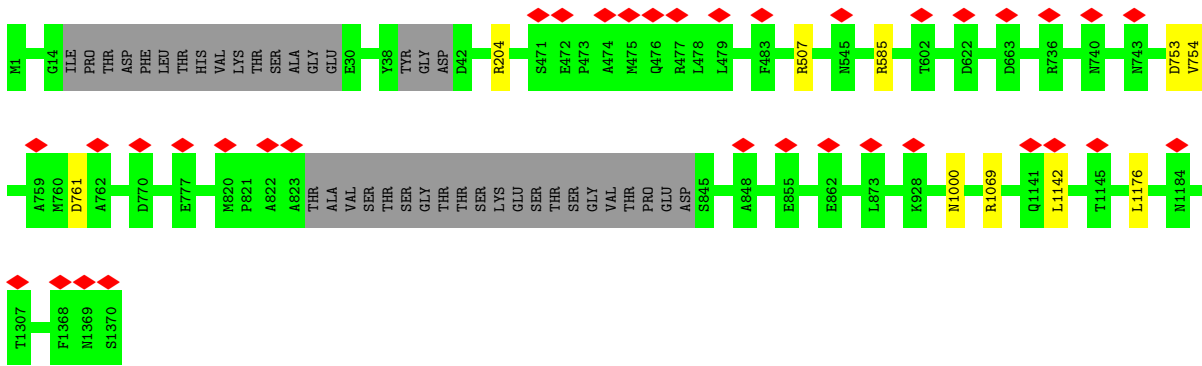
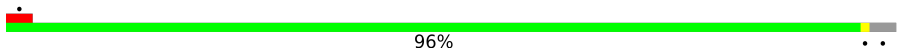
• Molecule 8: Major capsid protein

Chain B:



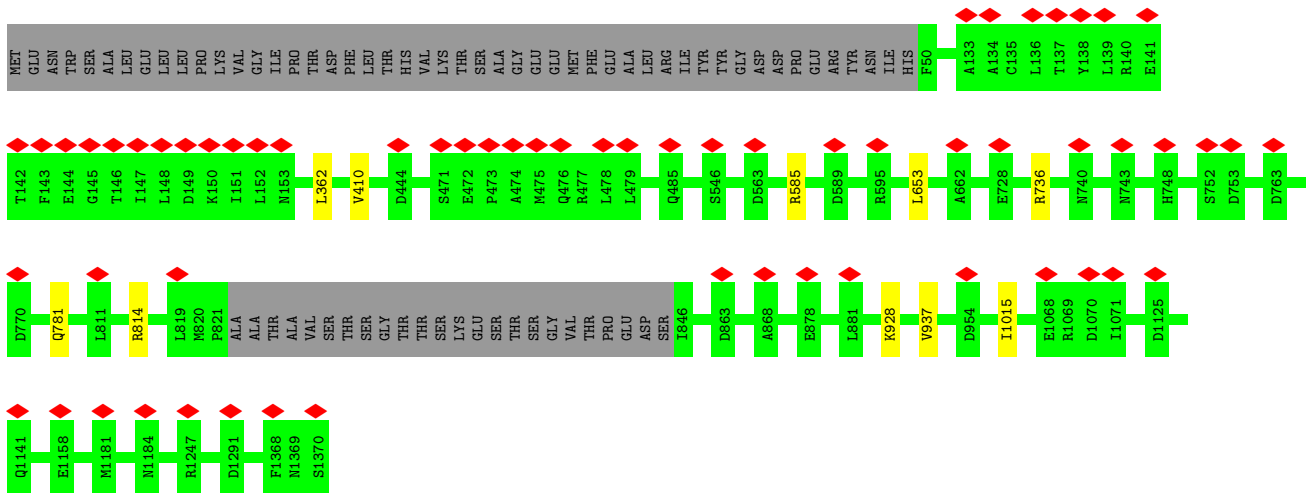
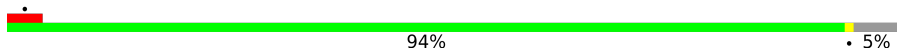
- Molecule 8: Major capsid protein

Chain C:

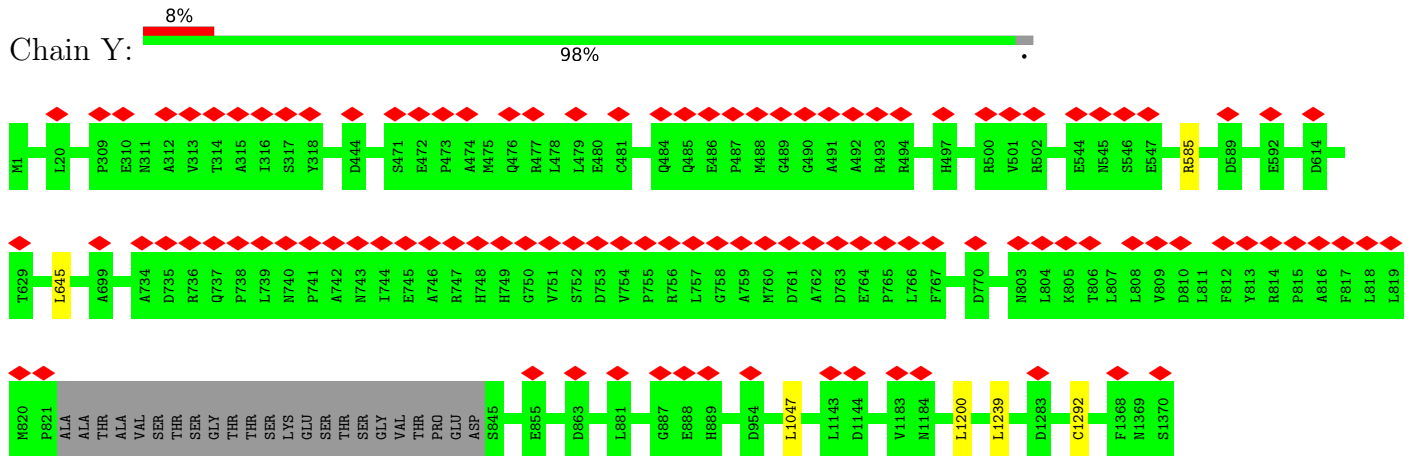


- Molecule 8: Major capsid protein

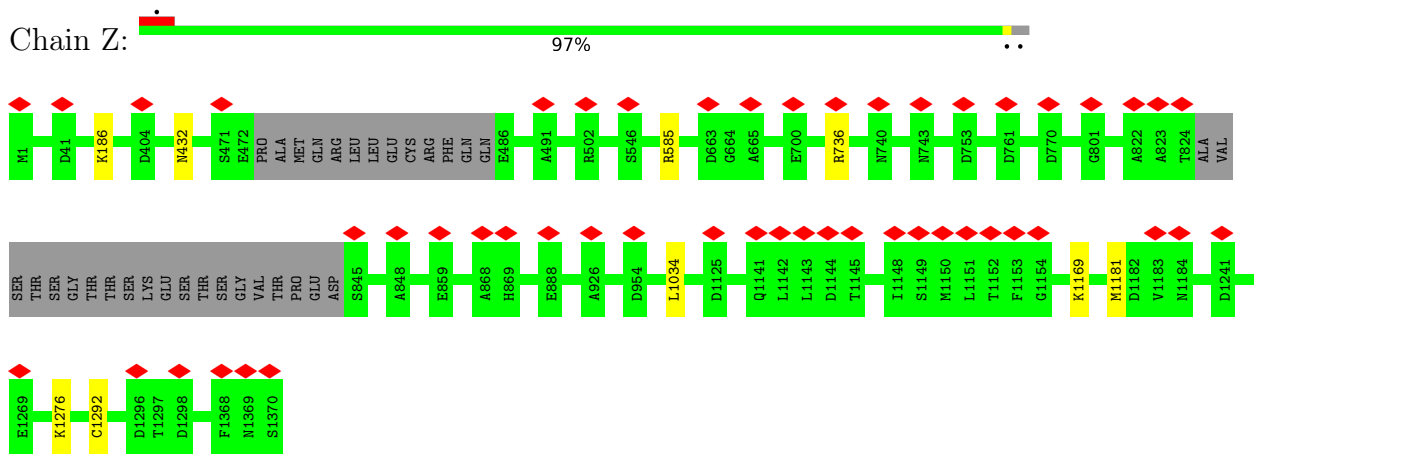
Chain D:



- Molecule 8: Major capsid protein



• Molecule 8: Major capsid protein



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	42849	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	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.093	Depositor
Minimum map value	-0.068	Depositor
Average map value	0.005	Depositor
Map value standard deviation	0.009	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	416.0, 416.0, 416.0	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.625, 1.625, 1.625	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	I	0.31	0/2436	0.65	2/3310 (0.1%)
1	h	0.31	0/2431	0.58	0/3306
1	n	0.30	0/2374	0.52	0/3225
1	o	0.30	0/2333	0.51	0/3167
2	H	0.27	0/174	0.60	0/233
2	P	0.28	0/174	0.51	0/233
3	g	0.31	0/1860	0.57	0/2521
3	m	0.34	0/2374	0.58	2/3221 (0.1%)
4	M	0.34	0/3935	0.57	0/5331
5	N	0.30	0/662	0.62	0/892
5	O	0.27	0/600	0.58	0/808
6	l	0.28	0/2358	0.55	0/3182
7	R	0.28	0/520	0.59	0/697
7	S	0.31	0/520	0.59	0/697
7	T	0.27	0/520	0.62	1/697 (0.1%)
7	i	0.29	0/520	0.74	2/697 (0.3%)
7	j	0.30	0/520	0.64	0/697
8	B	0.34	0/10870	0.58	2/14804 (0.0%)
8	C	0.34	0/10786	0.57	3/14692 (0.0%)
8	D	0.33	0/10513	0.54	1/14322 (0.0%)
8	Y	0.32	0/10932	0.54	3/14892 (0.0%)
8	Z	0.33	0/10835	0.54	1/14762 (0.0%)
8	a	0.34	0/10318	0.57	3/14057 (0.0%)
All	All	0.33	0/88565	0.56	20/120443 (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
8	C	0	1
8	D	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
8	Y	0	1
8	Z	0	1
8	a	0	1
All	All	0	5

There are no bond length outliers.

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	C	1142	LEU	CA-CB-CG	8.64	135.16	115.30
1	I	179	ASP	CB-CG-OD2	8.31	125.78	118.30
3	m	140	LEU	CA-CB-CG	6.81	130.96	115.30
7	i	64	LEU	CA-CB-CG	6.66	130.63	115.30
8	C	761	ASP	CB-CG-OD2	6.65	124.28	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
8	C	585	ARG	Peptide
8	D	585	ARG	Peptide
8	Y	585	ARG	Peptide
8	Z	585	ARG	Peptide
8	a	585	ARG	Peptide

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	I	299/306 (98%)	287 (96%)	12 (4%)	0	100	100
1	h	298/306 (97%)	283 (95%)	15 (5%)	0	100	100
1	n	290/306 (95%)	273 (94%)	17 (6%)	0	100	100
1	o	285/306 (93%)	274 (96%)	11 (4%)	0	100	100
2	H	18/2241 (1%)	18 (100%)	0	0	100	100
2	P	18/2241 (1%)	18 (100%)	0	0	100	100
3	g	221/290 (76%)	204 (92%)	17 (8%)	0	100	100
3	m	288/290 (99%)	274 (95%)	14 (5%)	0	100	100
4	M	462/594 (78%)	442 (96%)	20 (4%)	0	100	100
5	N	74/642 (12%)	70 (95%)	4 (5%)	0	100	100
5	O	65/642 (10%)	64 (98%)	1 (2%)	0	100	100
6	l	282/1048 (27%)	269 (95%)	13 (5%)	0	100	100
7	R	61/75 (81%)	60 (98%)	1 (2%)	0	100	100
7	S	61/75 (81%)	60 (98%)	1 (2%)	0	100	100
7	T	61/75 (81%)	59 (97%)	2 (3%)	0	100	100
7	i	61/75 (81%)	58 (95%)	3 (5%)	0	100	100
7	j	61/75 (81%)	58 (95%)	3 (5%)	0	100	100
8	B	1331/1370 (97%)	1260 (95%)	71 (5%)	0	100	100
8	C	1323/1370 (97%)	1268 (96%)	55 (4%)	0	100	100
8	D	1293/1370 (94%)	1210 (94%)	83 (6%)	0	100	100
8	Y	1343/1370 (98%)	1267 (94%)	76 (6%)	0	100	100
8	Z	1331/1370 (97%)	1264 (95%)	67 (5%)	0	100	100
8	a	1269/1370 (93%)	1216 (96%)	53 (4%)	0	100	100
All	All	10795/17807 (61%)	10256 (95%)	539 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	270/273 (99%)	270 (100%)	0	100	100
1	h	270/273 (99%)	270 (100%)	0	100	100
1	n	263/273 (96%)	260 (99%)	3 (1%)	73	85
1	o	259/273 (95%)	258 (100%)	1 (0%)	91	94
2	H	19/1941 (1%)	19 (100%)	0	100	100
2	P	19/1941 (1%)	19 (100%)	0	100	100
3	g	199/252 (79%)	198 (100%)	1 (0%)	88	93
3	m	252/252 (100%)	250 (99%)	2 (1%)	81	89
4	M	395/500 (79%)	392 (99%)	3 (1%)	81	89
5	N	69/526 (13%)	69 (100%)	0	100	100
5	O	64/526 (12%)	64 (100%)	0	100	100
6	l	255/883 (29%)	254 (100%)	1 (0%)	91	94
7	R	59/68 (87%)	59 (100%)	0	100	100
7	S	59/68 (87%)	59 (100%)	0	100	100
7	T	59/68 (87%)	58 (98%)	1 (2%)	60	78
7	i	59/68 (87%)	59 (100%)	0	100	100
7	j	59/68 (87%)	58 (98%)	1 (2%)	60	78
8	B	1169/1192 (98%)	1164 (100%)	5 (0%)	91	94
8	C	1156/1192 (97%)	1150 (100%)	6 (0%)	88	93
8	D	1130/1192 (95%)	1122 (99%)	8 (1%)	84	90
8	Y	1174/1192 (98%)	1172 (100%)	2 (0%)	93	96
8	Z	1163/1192 (98%)	1156 (99%)	7 (1%)	86	92
8	a	1110/1192 (93%)	1105 (100%)	5 (0%)	88	93
All	All	9531/15405 (62%)	9485 (100%)	46 (0%)	89	93

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	C	1069	ARG
8	D	937	VAL
8	D	410	VAL
8	D	781	GLN
8	Y	1200	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 157

such sidechains are listed below:

Mol	Chain	Res	Type
8	D	1000	ASN
8	Z	543	GLN
8	D	1166	HIS
8	Y	940	HIS
8	Z	1082	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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.

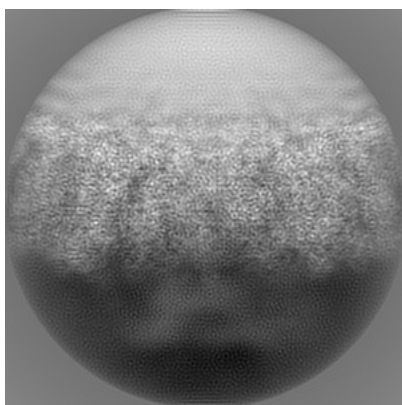
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-31298. These allow visual inspection of the internal detail of the map and identification of artifacts.

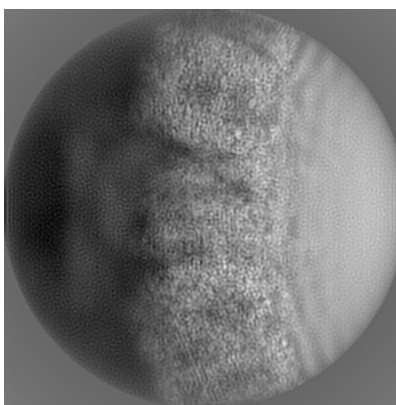
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

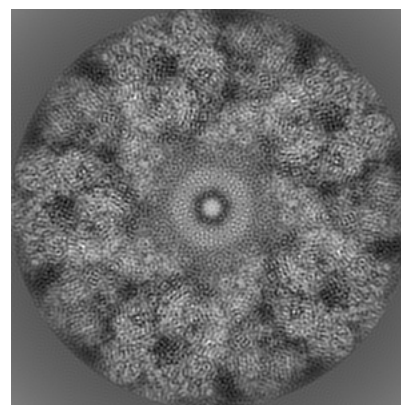
6.1.1 Primary map



X



Y

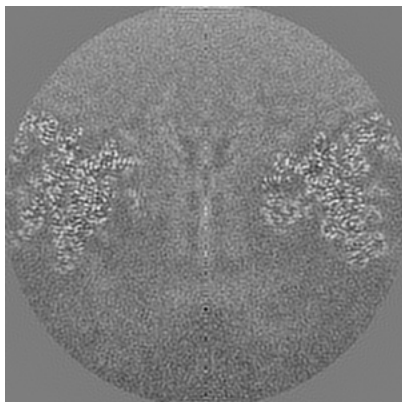


Z

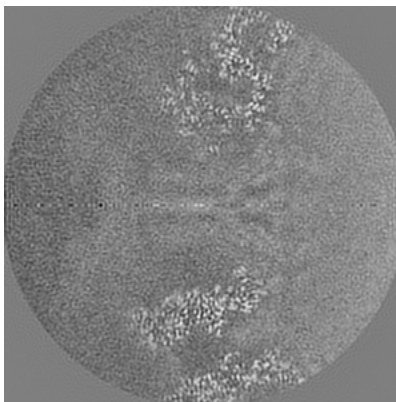
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

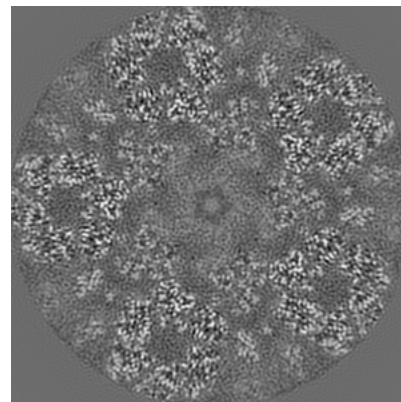
6.2.1 Primary map



X Index: 128



Y Index: 128

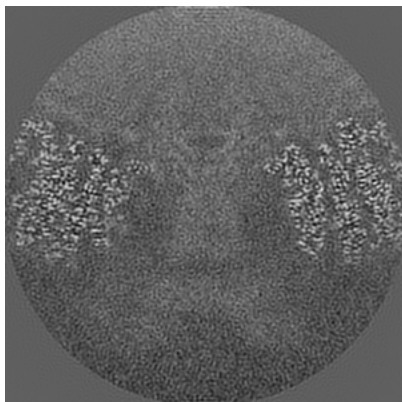


Z Index: 128

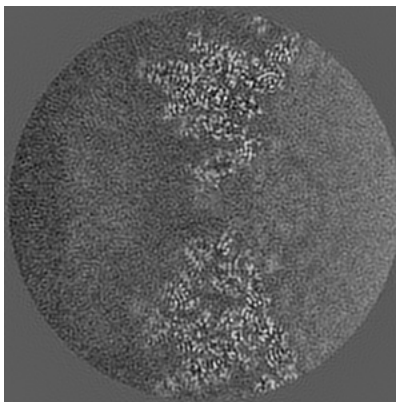
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

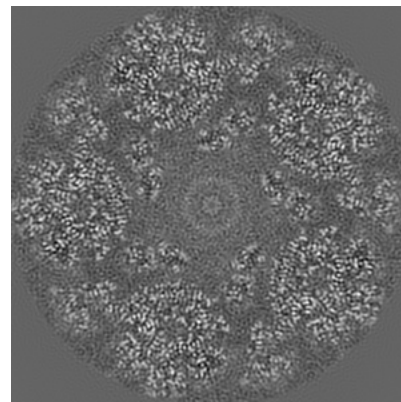
6.3.1 Primary map



X Index: 116



Y Index: 97

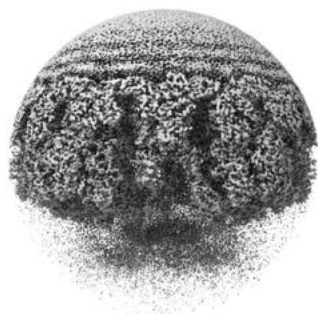


Z Index: 141

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

6.4 Orthogonal surface views [i](#)

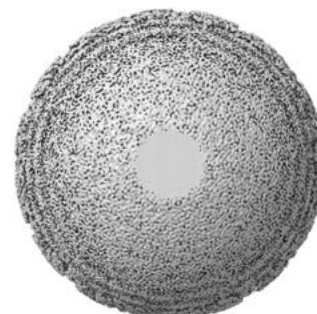
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

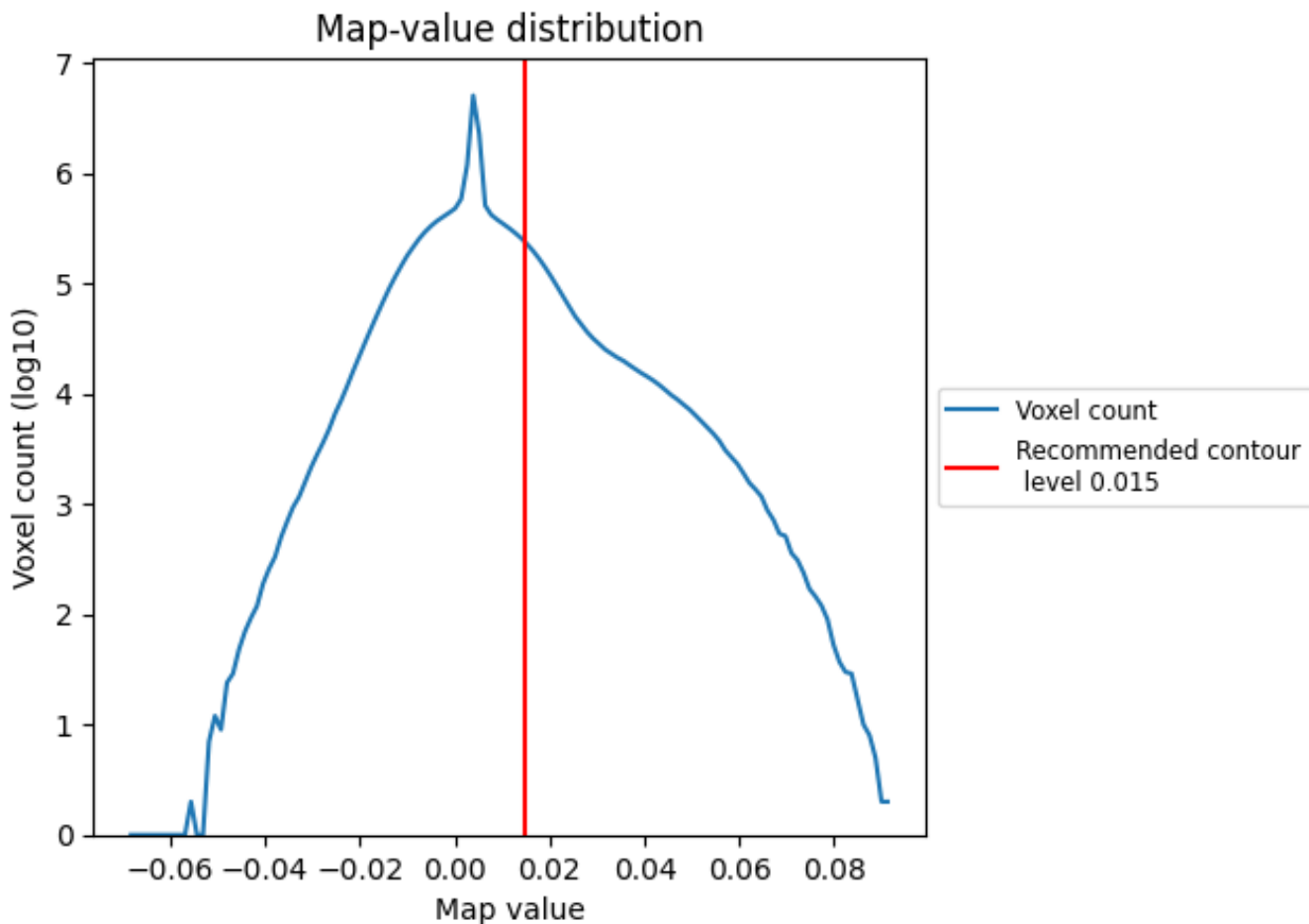
6.5 Mask visualisation

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

7 Map analysis [i](#)

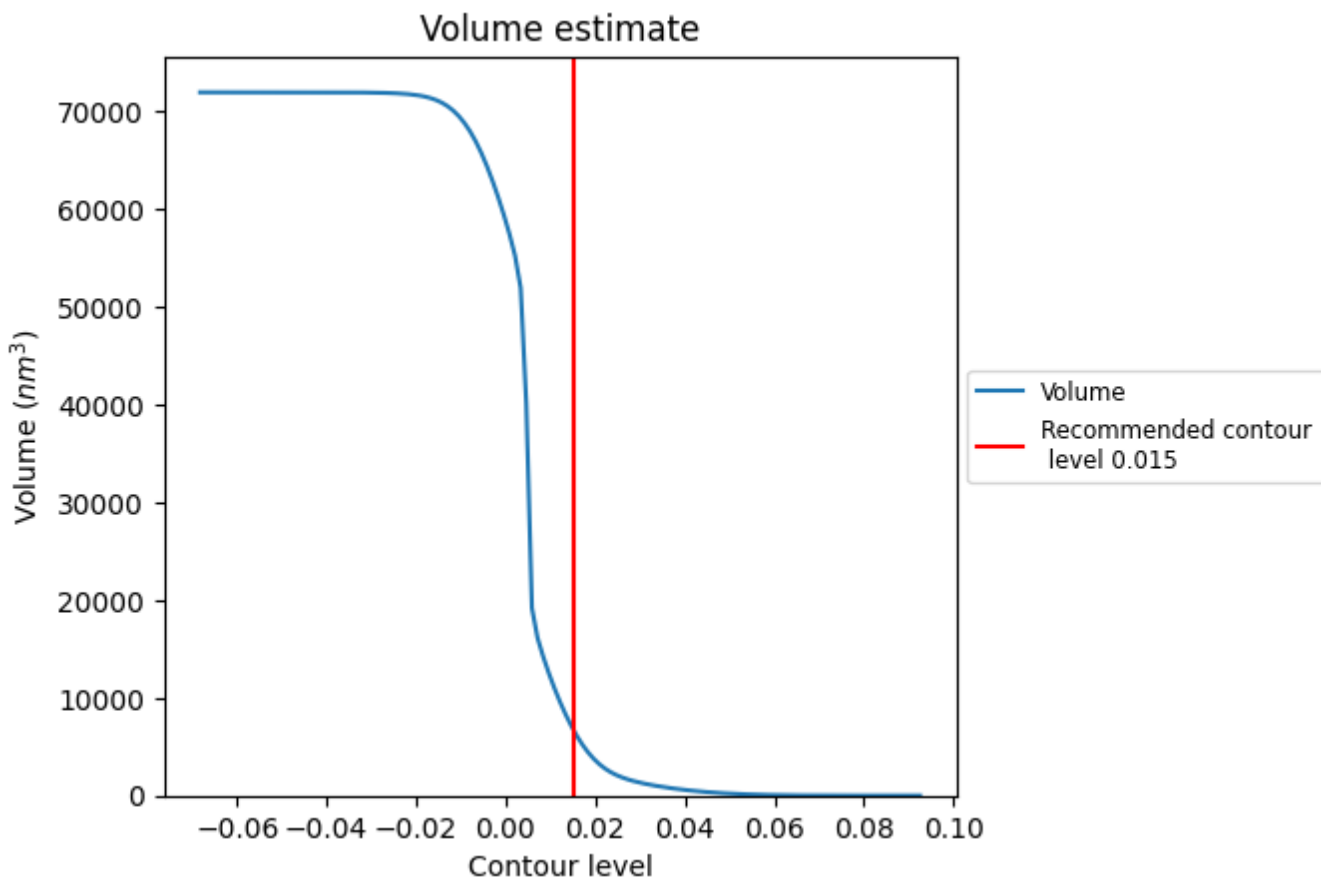
This section contains the results of statistical analysis of the map.

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.

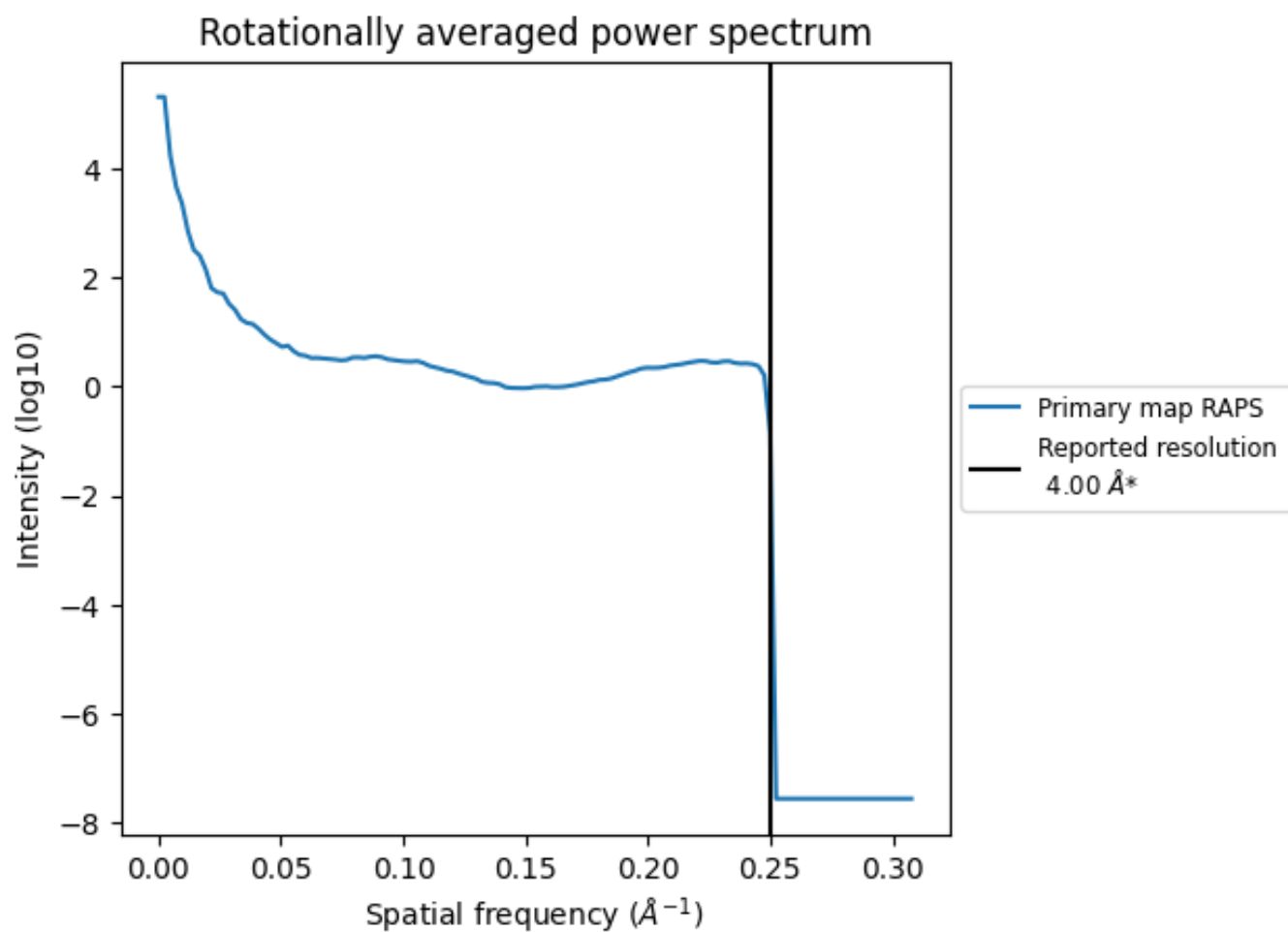
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 6849 nm³; this corresponds to an approximate mass of 6187 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 0.250 Å⁻¹

8 Fourier-Shell correlation

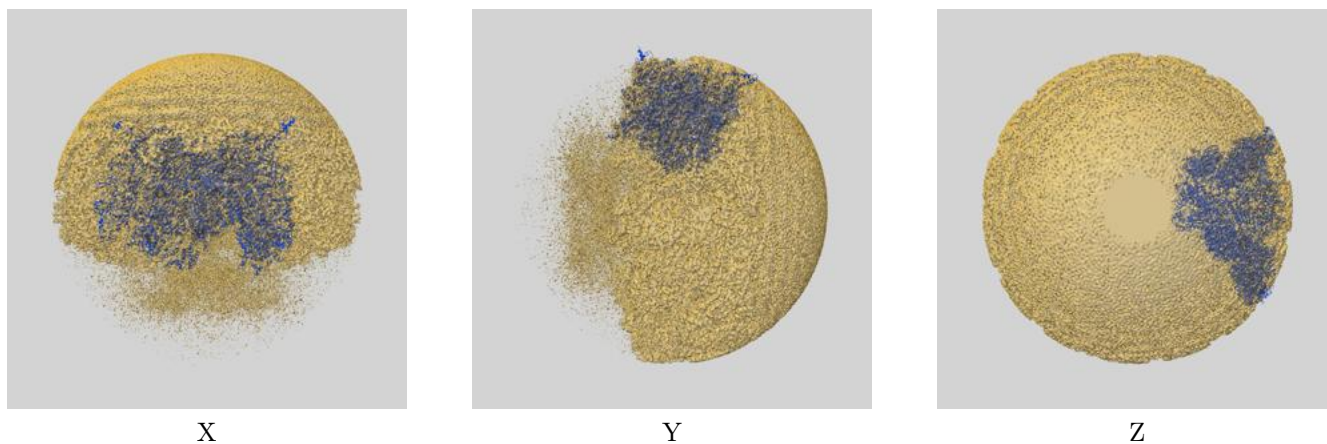
This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit [i](#)

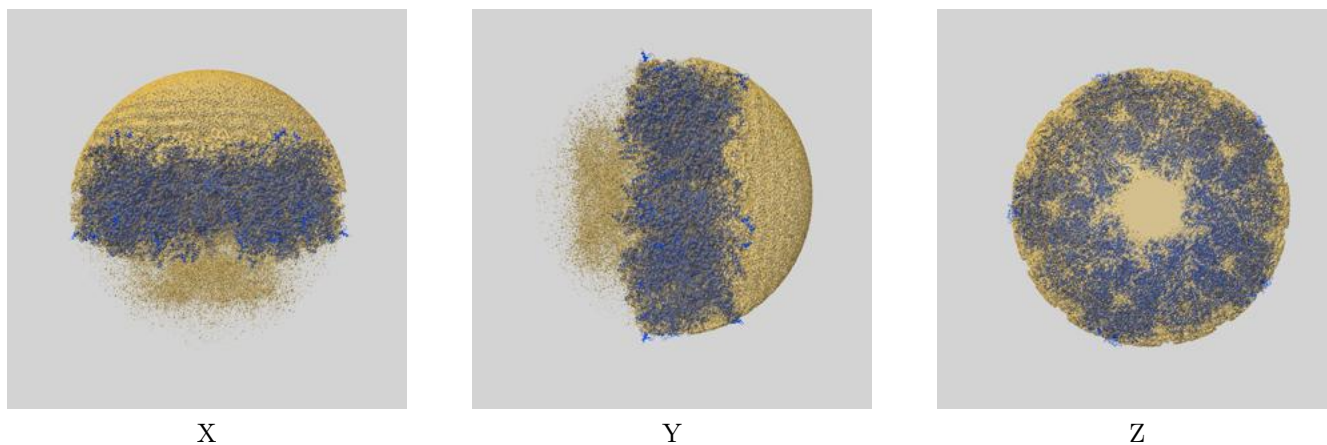
This section contains information regarding the fit between EMDB map EMD-31298 and PDB model 7ETJ. Per-residue inclusion information can be found in section 3 on page 6.

9.1 Map-model overlays

9.1.1 Map-model overlay [i](#)

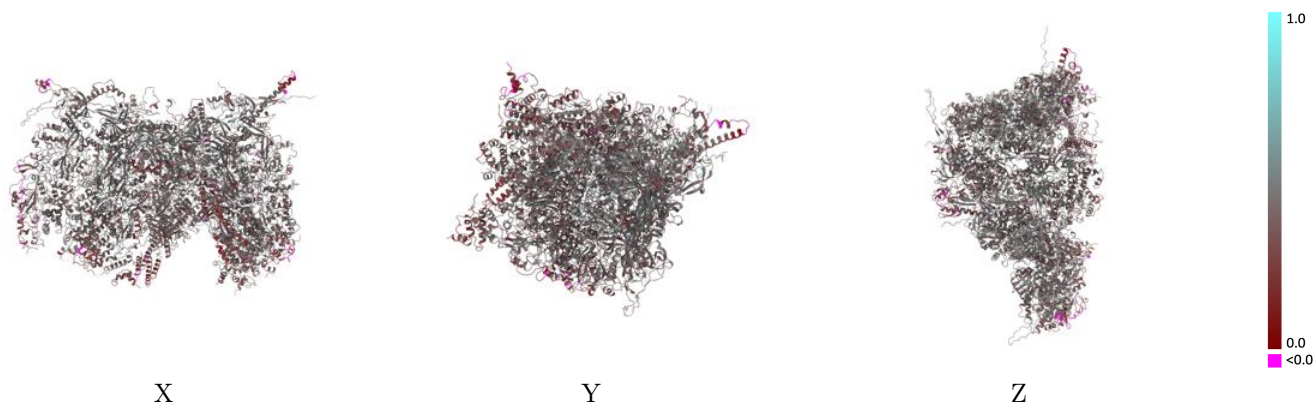


9.1.2 Map-model assembly overlay [i](#)



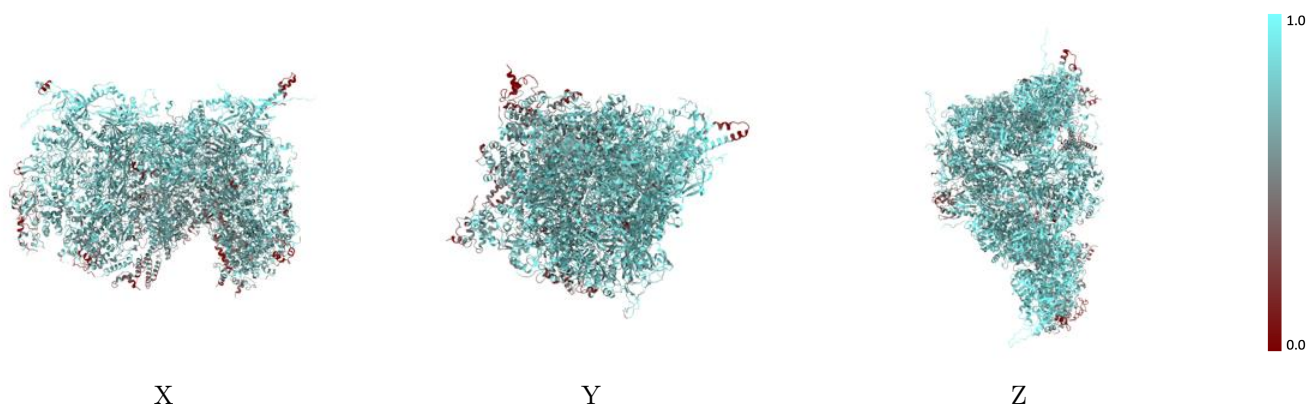
The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



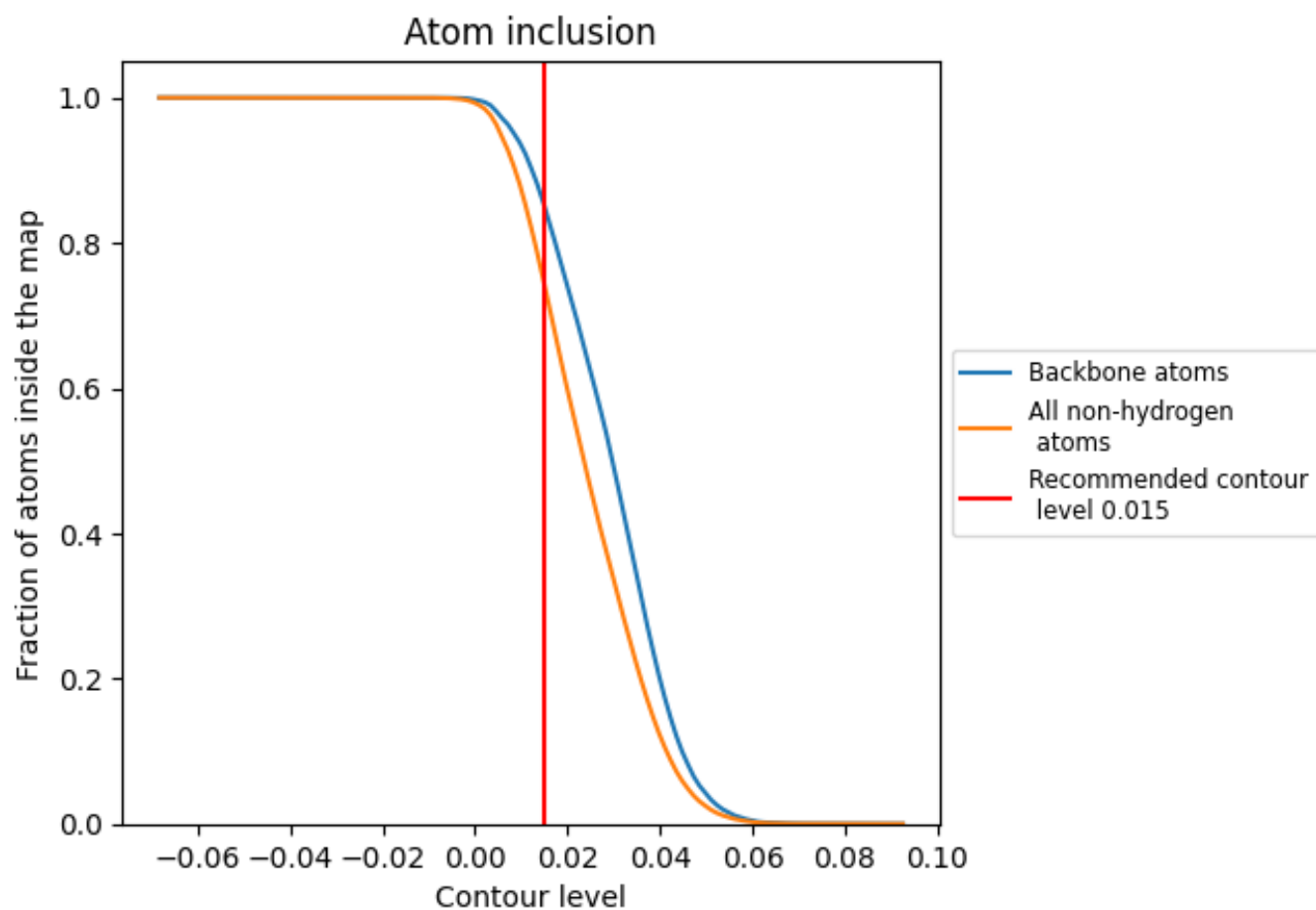
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.015).

















































9.4 Atom inclusion [i](#)



At the recommended contour level, 85% of all backbone atoms, 74% 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 (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7435	 0.4060
1	 0.5073	 0.3130
B	 0.7709	 0.4180
C	 0.7895	 0.4290
D	 0.7805	 0.4160
H	 0.4132	 0.2030
I	 0.6933	 0.3800
M	 0.7212	 0.4020
N	 0.6100	 0.3640
O	 0.5018	 0.2500
P	 0.5329	 0.2810
R	 0.5060	 0.3210
S	 0.5964	 0.3500
T	 0.3032	 0.2460
Y	 0.7567	 0.4060
Z	 0.7865	 0.4280
a	 0.7768	 0.4250
g	 0.5517	 0.3090
h	 0.6854	 0.3930
i	 0.3855	 0.2820
j	 0.4478	 0.3160
m	 0.7788	 0.4300
n	 0.7739	 0.4140
o	 0.7692	 0.4100

