wwPDB EM Validation Summary Report (i)

Nov 6, 2022 – 12:10 PM EST

PDB ID	:	6BSP
EMDB ID	:	EMD-7136
Title	:	High-Resolution Structure Analysis of Antibody V5 and U4 Conformational Epitope on Human Papillomavirus 16
Authors	:	Guan, J.; Bywaters, S.M.; Brendle, S.A.; Ashley, R.E.; Makhov, A.M.; Conway, J.F.; Christenson, N.D.; Hafenstein, S.
Deposited on	:	2017-12-04
Resolution	:	4.70 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

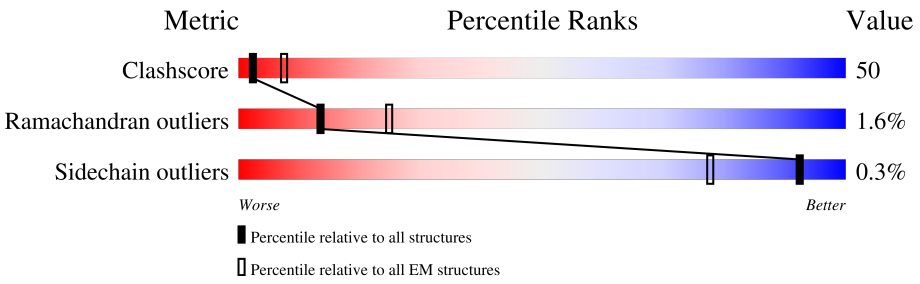
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 4.70 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 <=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	109	81%; 26%, 73%, •
2	B	110	85%; 26%, 74%
3	C	469	17%; 47%, 48%, • •
3	D	469	•; 15%, 77%, 6%, •
3	E	469	•; 47%, 50%, • •
3	F	469	6%; 46%, 50%, • •
3	G	469	•; 47%, 47%, 5%, •
3	H	469	7%; 46%, 50%, • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 23690 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 U4 Heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	109	851	534	145	167	5	0	0

- Molecule 2 is a protein called U4 Light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	110	850	541	135	170	4	0	0

- Molecule 3 is a protein called Major capsid protein L1.

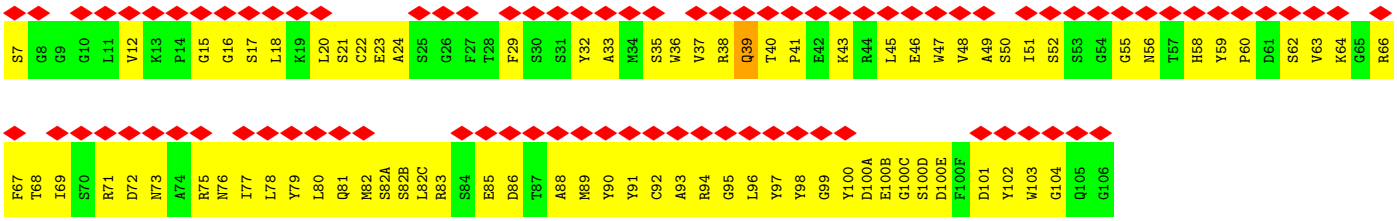
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	468	3675	2347	617	690	21	0	0
3	D	460	3614	2304	609	680	21	0	0
3	E	469	3686	2356	618	691	21	0	0
3	F	468	3675	2347	617	690	21	0	0
3	G	467	3666	2341	615	689	21	0	0
3	H	468	3673	2345	617	690	21	0	0

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.

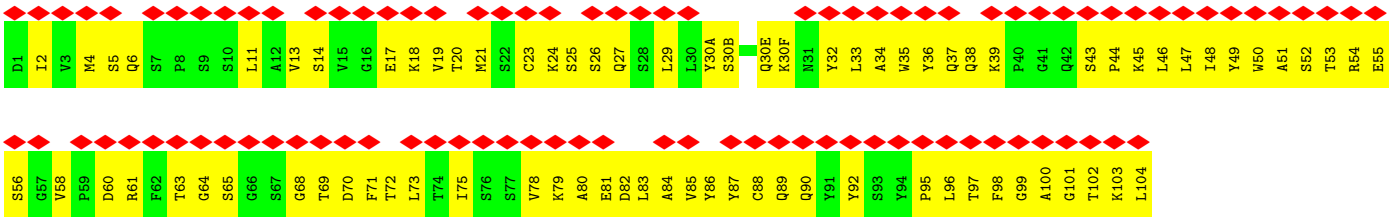
- Molecule 1: U4 Heavy chain

Chain A:



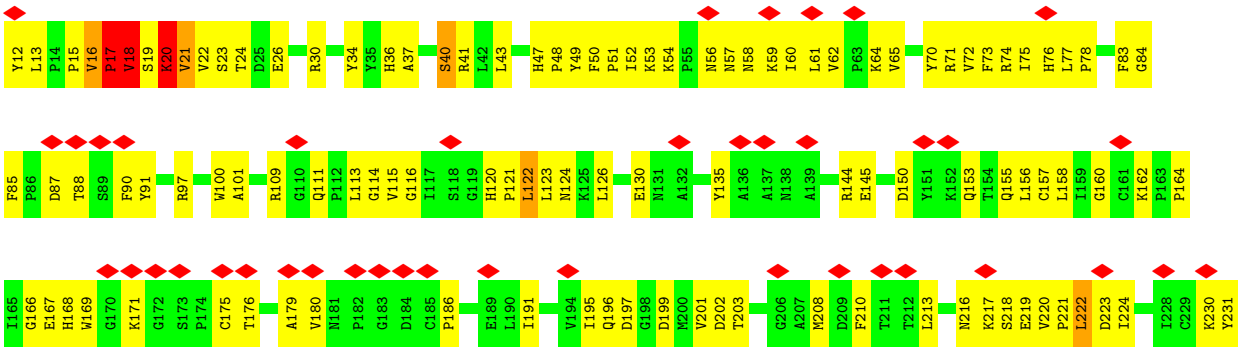
- Molecule 2: U4 Light chain

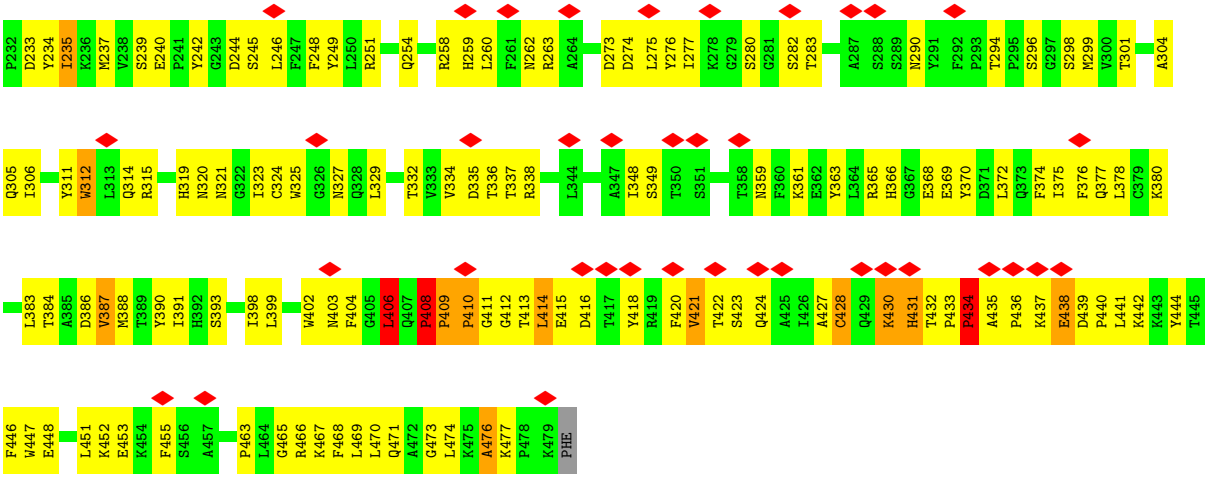
Chain B:



- Molecule 3: Major capsid protein L1

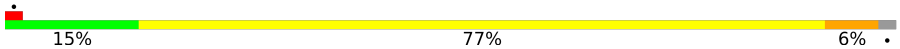
Chain C:

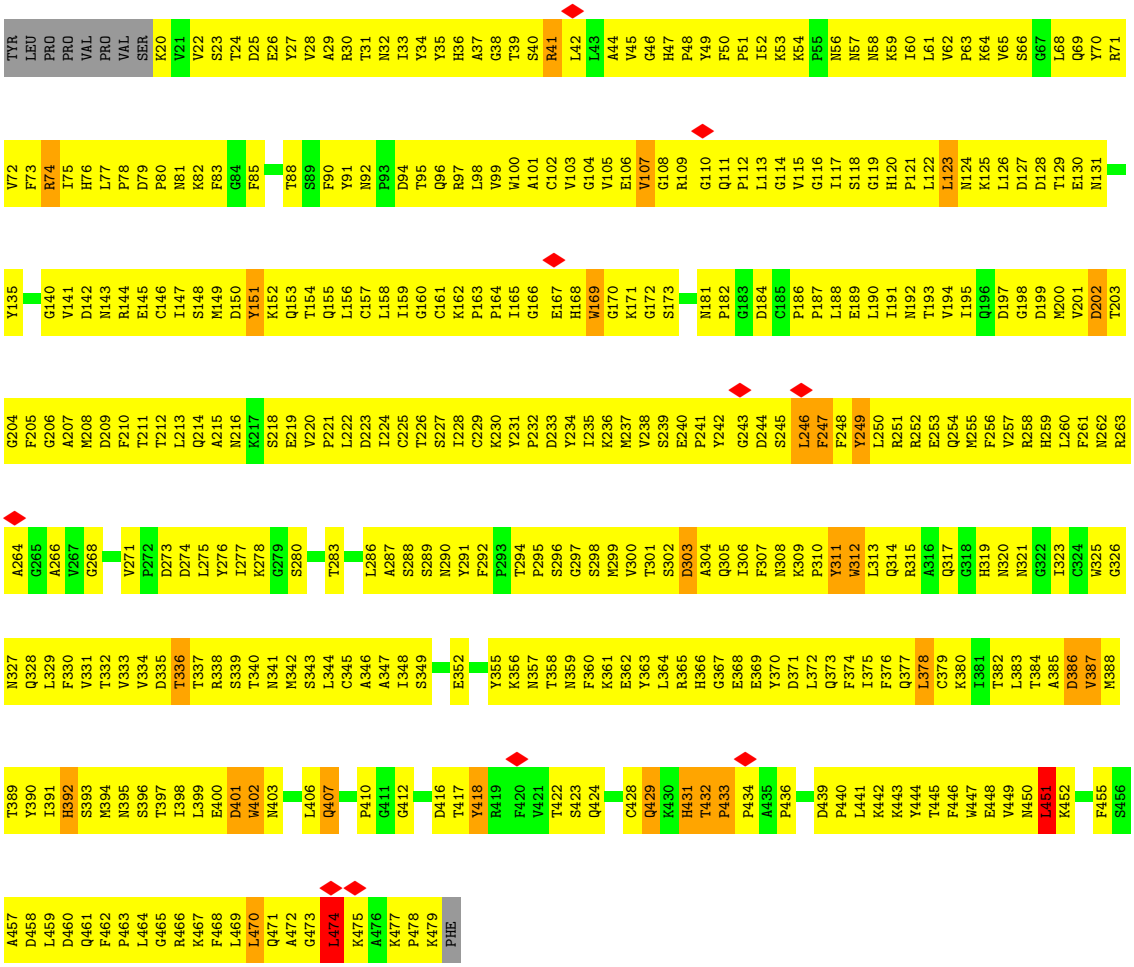




• Molecule 3: Major capsid protein L1

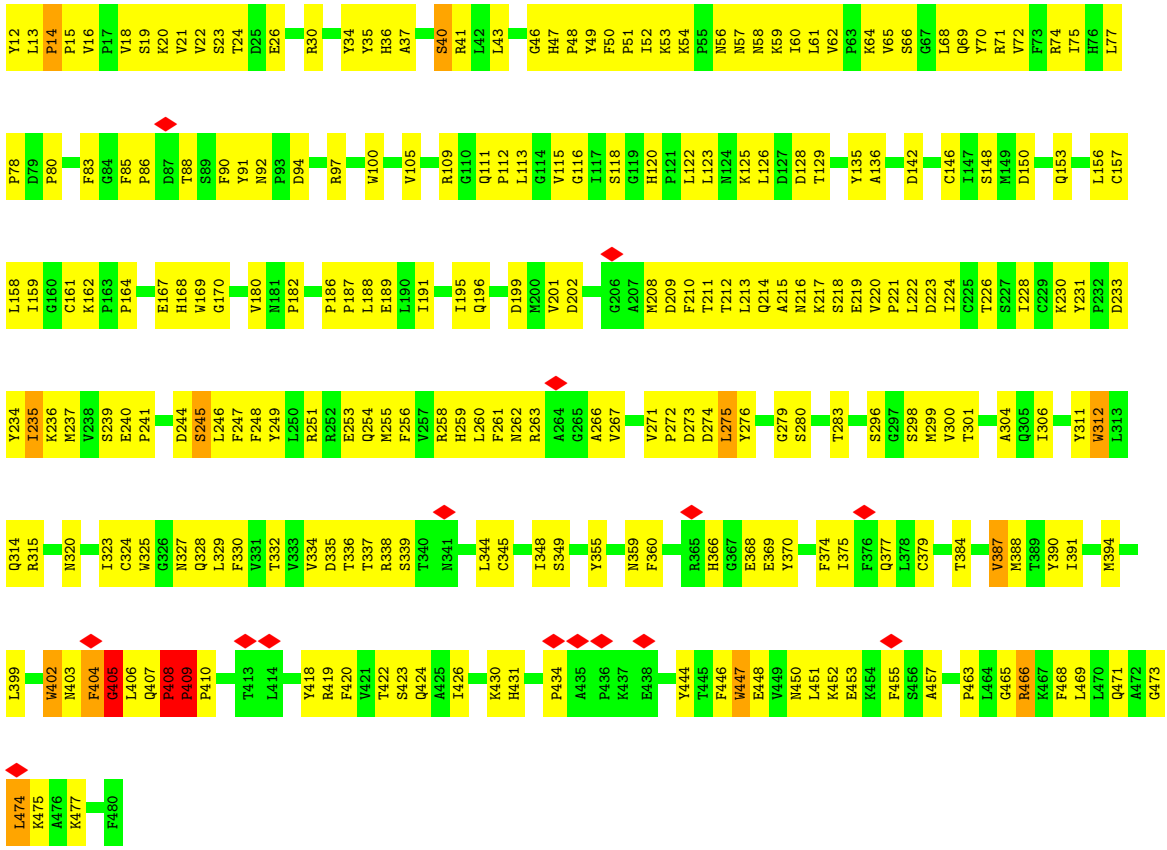
Chain D:





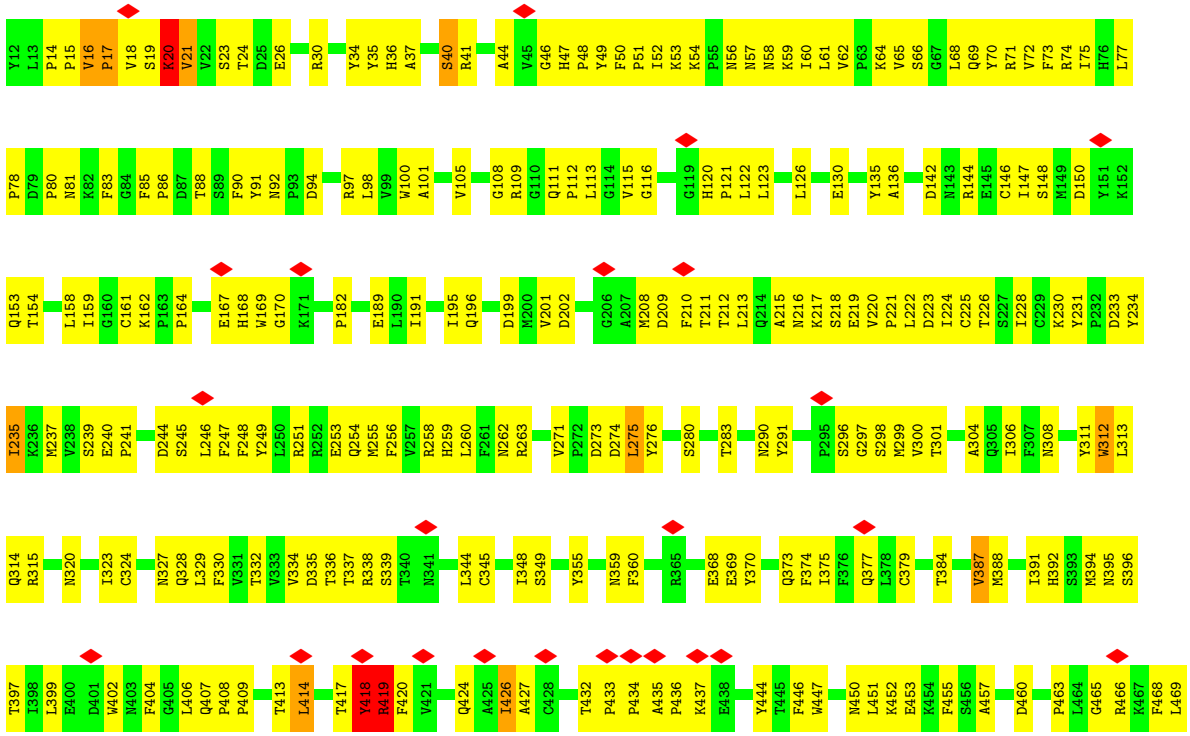
• Molecule 3: Major capsid protein L1

Chain E:



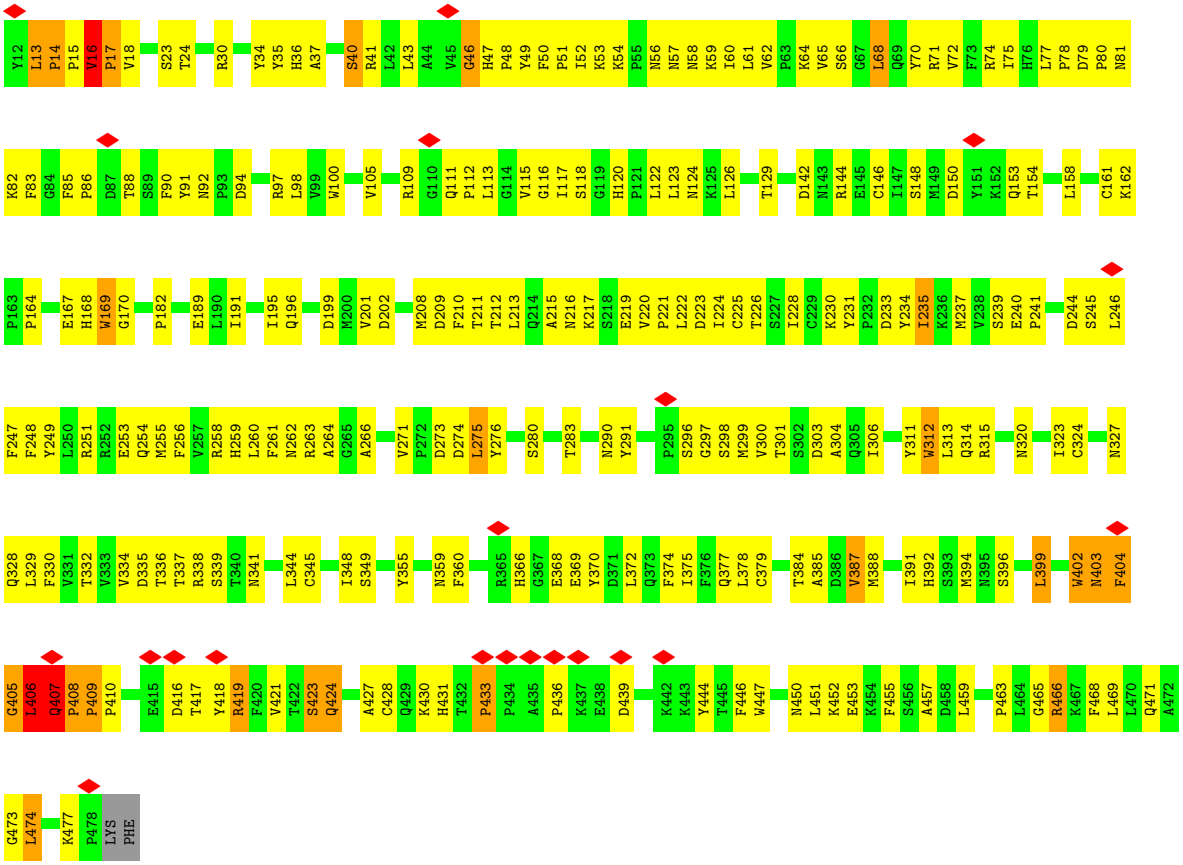
- Molecule 3: Major capsid protein L1

Chain F:



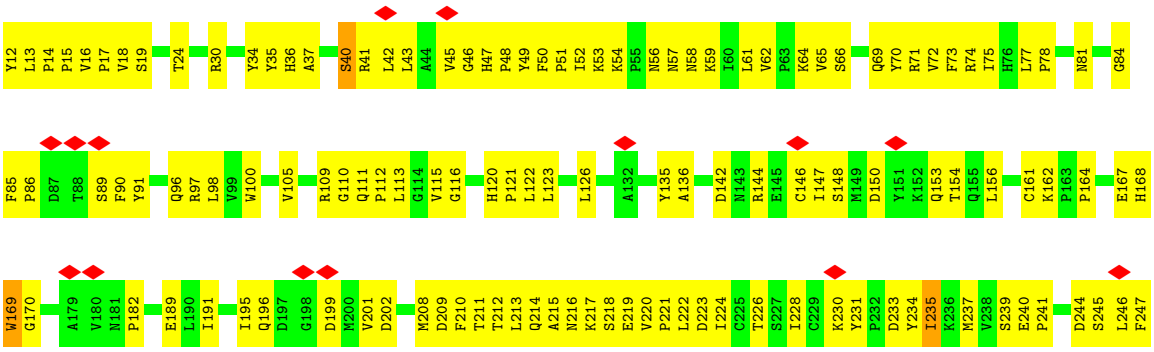
• Molecule 3: Major capsid protein L1

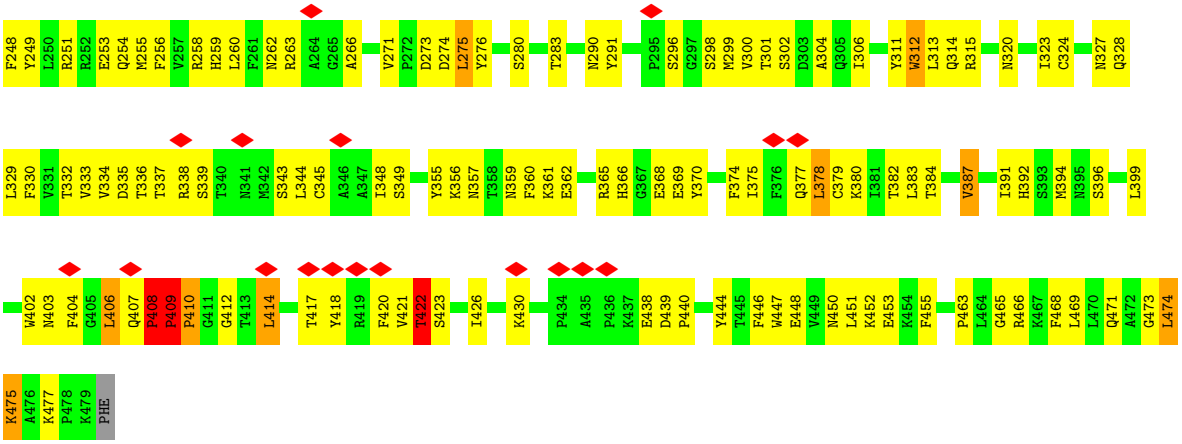
Chain G:



• Molecule 3: Major capsid protein L1

Chain H:




F248
Y249
L250
R251
R252
E253
Q254
M255
F256
V257
R258
H259
L260
F261
N262
R263
A264
G265
A266
V271
P272
D273
D274
L275
Y276
S280
T283
N290
Y291
P295
S296
G297
S298
M299
V300
T301
S302
D303
A304
Q305
I306
Y311
W312
L313
Q314
R315
N320
I323
C324
N327
Q328
L329
F330
V331
T332
V333
V334
D335
T336
T337
R338
S339
T340
N341
M342
S343
L344
C345
A346
A347
I348
S349
Y355
K356
N357
T358
N359
F360
K361
E362
R365
H366
G367
E368
E369
Y370
F374
I375
F376
Q377
L378
C379
K380
I381
T382
L383
T384
V387
I391
H392
S393
M394
N395
S396
L399
W402
N403
F404
G405
L406
Q407
P408
P409
P410
G411
G412
T413
L414
T417
Y418
R419
F420
V421
T422
S423
I426
K430
P434
A435
P436
K437
E438
D439
P440
Y444
T445
F446
W447
E448
V449
N450
L451
K452
E453
K454
F455
P463
L464
G465
R466
K467
F468
L469
L470
Q471
A472
G473
L474
K475
A476
K477
P478
K479
PHE

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	17612	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	7, 7	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON II (4k x 4k), FEI FALCON II (4k x 4k)	Depositor
Maximum map value	14.058	Depositor
Minimum map value	-8.247	Depositor
Average map value	0.000	Depositor
Map value standard deviation	1.000	Depositor
Recommended contour level	2.1	Depositor
Map size (Å)	779.95996, 779.95996, 779.95996	wwPDB
Map dimensions	680, 680, 680	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.147, 1.147, 1.147	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	0.50	1/873 (0.1%)	0.71	1/1177 (0.1%)
2	B	0.44	0/870	0.76	0/1181
3	C	0.96	10/3776 (0.3%)	0.93	22/5142 (0.4%)
3	D	0.82	4/3711 (0.1%)	0.95	6/5049 (0.1%)
3	E	0.76	4/3788 (0.1%)	0.79	7/5158 (0.1%)
3	F	0.73	1/3776 (0.0%)	0.79	11/5142 (0.2%)
3	G	0.75	6/3767 (0.2%)	0.82	11/5131 (0.2%)
3	H	0.75	5/3774 (0.1%)	0.85	13/5139 (0.3%)
All	All	0.78	31/24335 (0.1%)	0.85	71/33119 (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	1
3	C	0	12
3	D	0	21
3	E	0	2
3	F	0	7
3	G	0	9
3	H	0	5
All	All	0	57

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	18	VAL	CA-CB	28.75	2.15	1.54
3	C	18	VAL	CB-CG1	21.24	1.97	1.52
3	C	17	PRO	CA-C	14.74	1.82	1.52
3	E	402	TRP	CB-CG	-10.49	1.31	1.50
3	H	409	PRO	N-CA	9.26	1.62	1.47

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	H	409	PRO	C-N-CD	-20.45	75.60	120.60
3	C	17	PRO	C-N-CA	12.92	154.00	121.70
3	G	405	GLY	N-CA-C	-11.90	83.35	113.10
3	G	16	VAL	C-N-CD	-11.61	95.06	120.60
3	C	20	LYS	CD-CE-NZ	-11.38	85.52	111.70

There are no chirality outliers.

5 of 57 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	98	TYR	Peptide
3	C	17	PRO	Peptide
3	C	18	VAL	Peptide
3	C	20	LYS	Peptide
3	C	408	PRO	Mainchain

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	851	0	792	120	0
2	B	850	0	831	126	0
3	C	3675	0	3592	289	0
3	D	3614	0	3528	811	0
3	E	3686	0	3601	295	0
3	F	3675	0	3592	280	0
3	G	3666	0	3579	281	0
3	H	3673	0	3585	292	0
All	All	23690	0	23100	2321	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 50.

The worst 5 of 2321 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)
3:C:17:PRO:C	3:C:17:PRO:CA	1.82	1.43
3:C:18:VAL:CB	3:C:18:VAL:CG1	1.97	1.39
3:F:19:SER:O	3:F:20:LYS:HG2	1.15	1.33
3:C:18:VAL:CB	3:C:18:VAL:CA	2.15	1.25
3:G:16:VAL:HB	3:G:17:PRO:CD	1.67	1.25

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	107/109 (98%)	95 (89%)	12 (11%)	0	100	100
2	B	108/110 (98%)	101 (94%)	6 (6%)	1 (1%)	17	56
3	C	466/469 (99%)	425 (91%)	33 (7%)	8 (2%)	9	43
3	D	458/469 (98%)	401 (88%)	54 (12%)	3 (1%)	22	62
3	E	467/469 (100%)	425 (91%)	33 (7%)	9 (2%)	8	40
3	F	466/469 (99%)	421 (90%)	36 (8%)	9 (2%)	8	40
3	G	465/469 (99%)	425 (91%)	29 (6%)	11 (2%)	6	35
3	H	466/469 (99%)	431 (92%)	28 (6%)	7 (2%)	10	46
All	All	3003/3033 (99%)	2724 (91%)	231 (8%)	48 (2%)	13	45

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	409	PRO
3	F	20	LYS
3	G	16	VAL
3	G	17	PRO
3	H	410	PRO

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	89/89 (100%)	89 (100%)	0	100	100
2	B	96/96 (100%)	96 (100%)	0	100	100
3	C	407/408 (100%)	405 (100%)	2 (0%)	88	93
3	D	399/408 (98%)	393 (98%)	6 (2%)	65	80
3	E	408/408 (100%)	408 (100%)	0	100	100
3	F	407/408 (100%)	407 (100%)	0	100	100
3	G	406/408 (100%)	406 (100%)	0	100	100
3	H	406/408 (100%)	406 (100%)	0	100	100
All	All	2618/2633 (99%)	2610 (100%)	8 (0%)	92	95

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	424	GLN
3	D	401	ASP
3	D	386	ASP
3	D	303	ASP
3	D	392	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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.

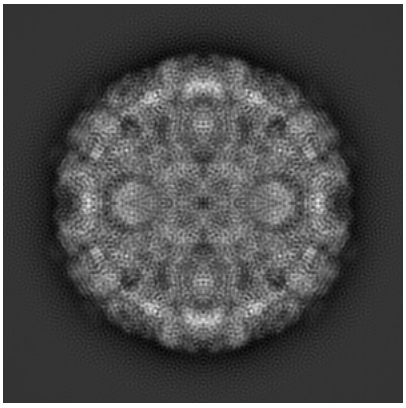
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-7136. These allow visual inspection of the internal detail of the map and identification of artifacts.

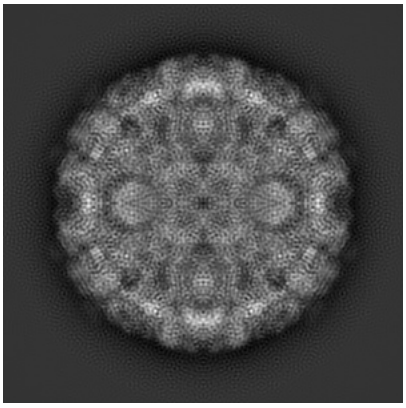
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

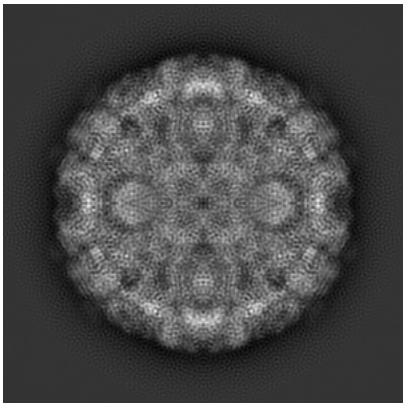
6.1.1 Primary map



X



Y

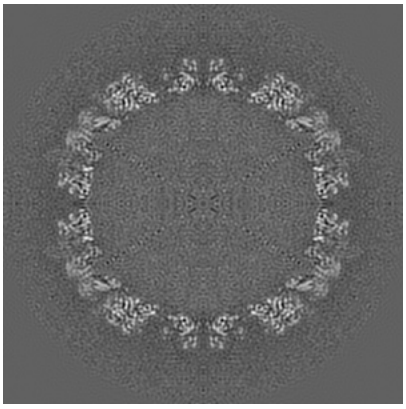


Z

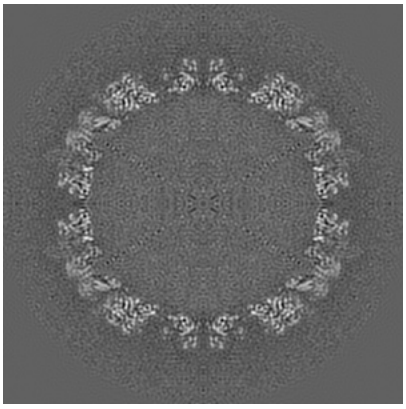
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

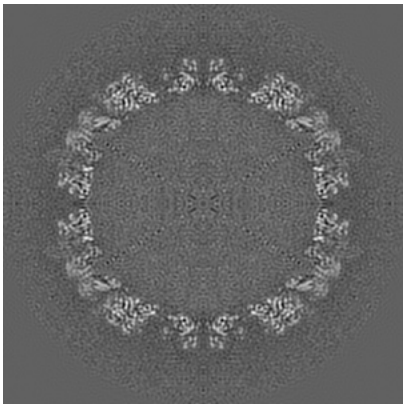
6.2.1 Primary map



X Index: 340



Y Index: 340

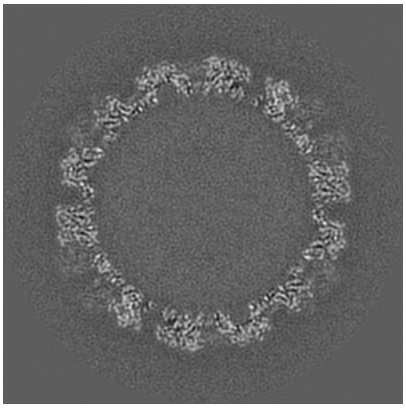


Z Index: 340

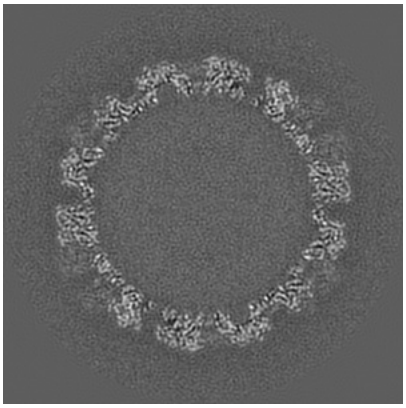
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

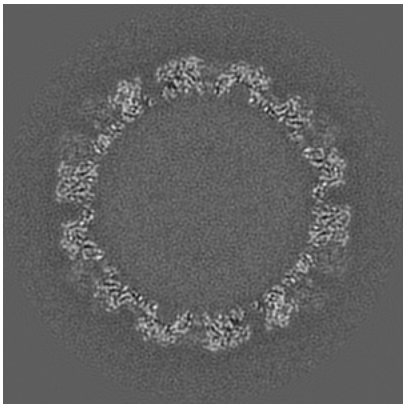
6.3.1 Primary map



X Index: 288



Y Index: 288

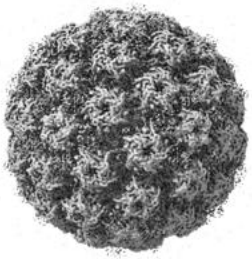


Z Index: 392

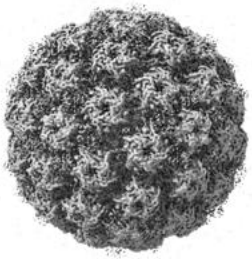
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

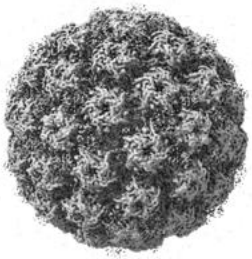
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 2.1. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

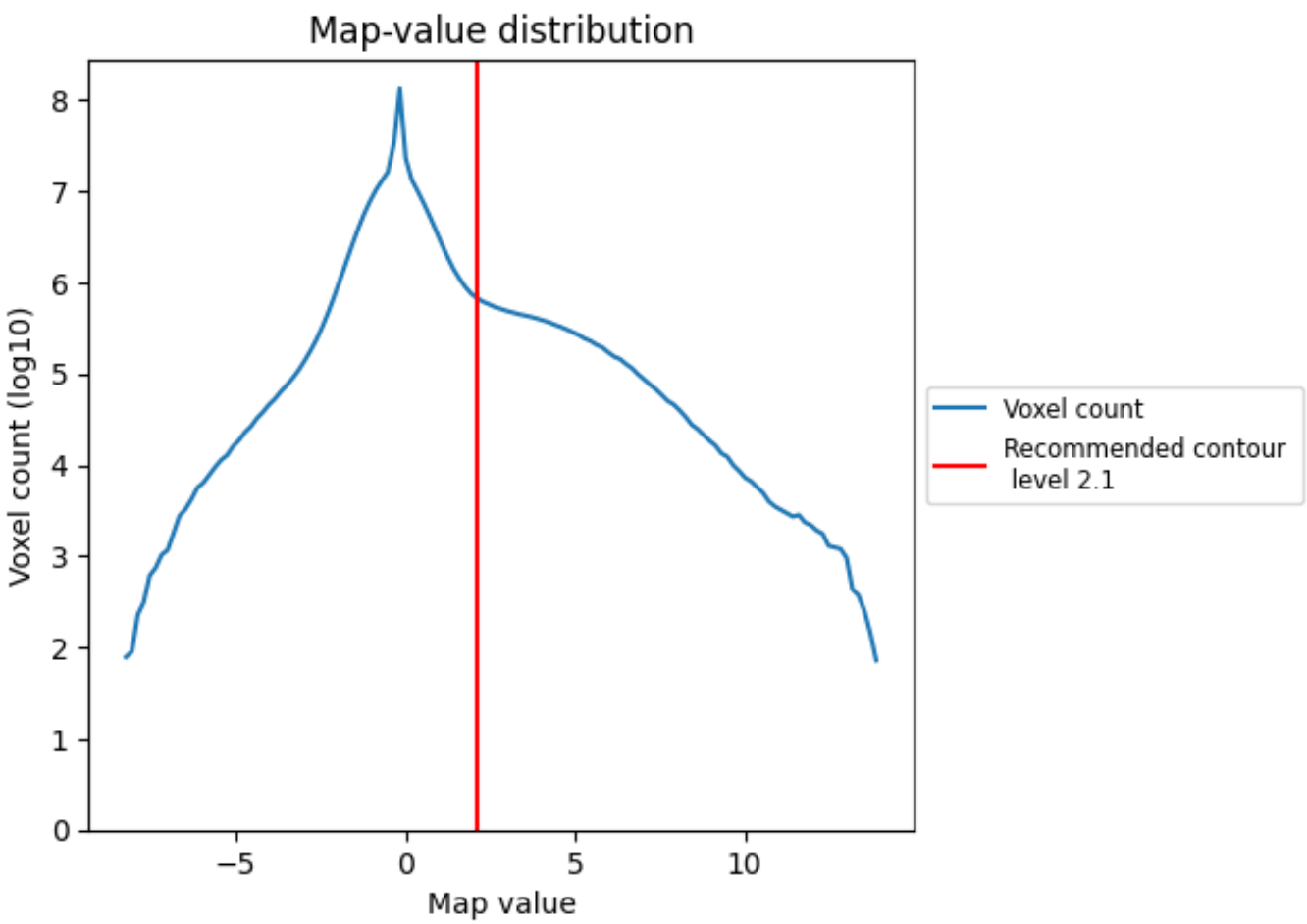
6.5 Mask visualisation (i)

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

7 Map analysis (i)

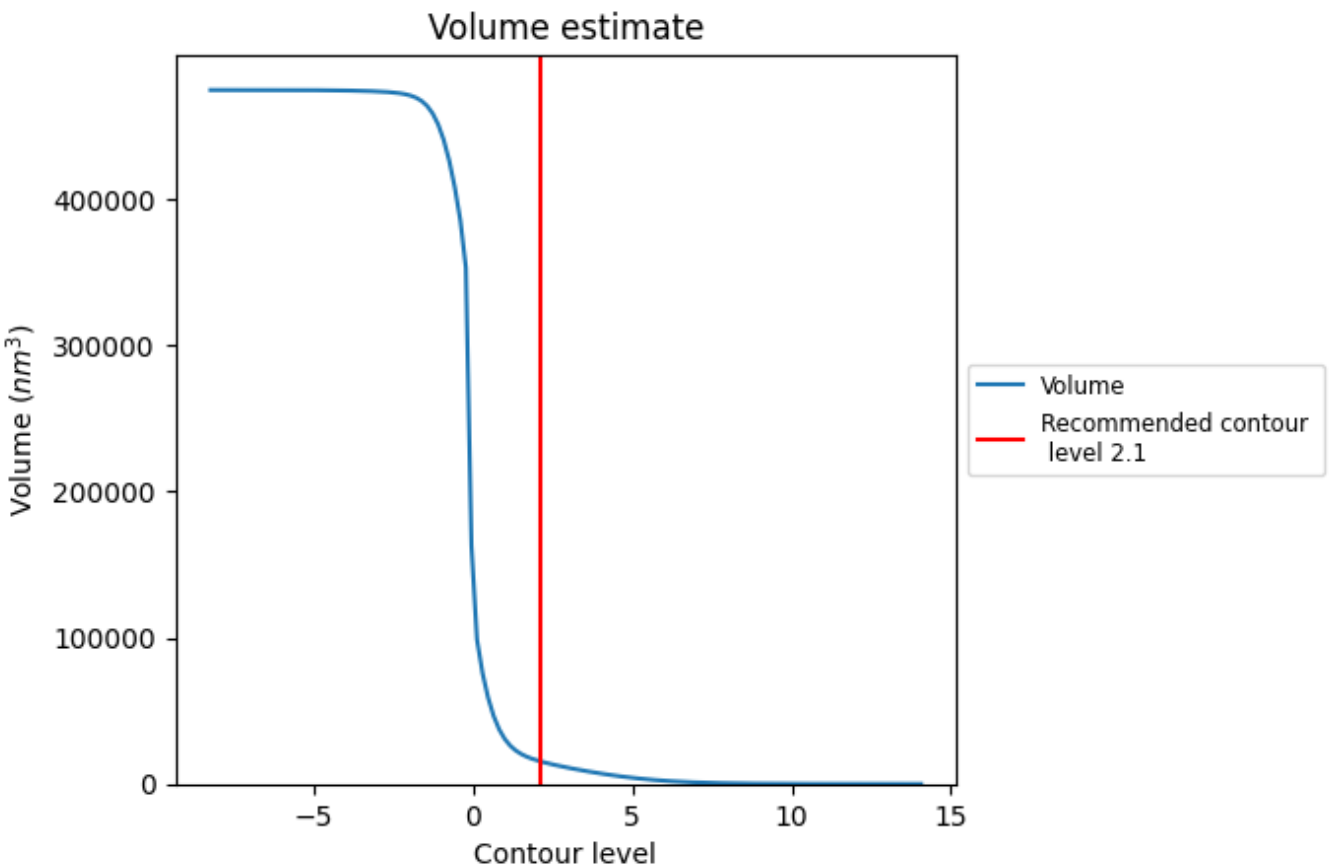
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

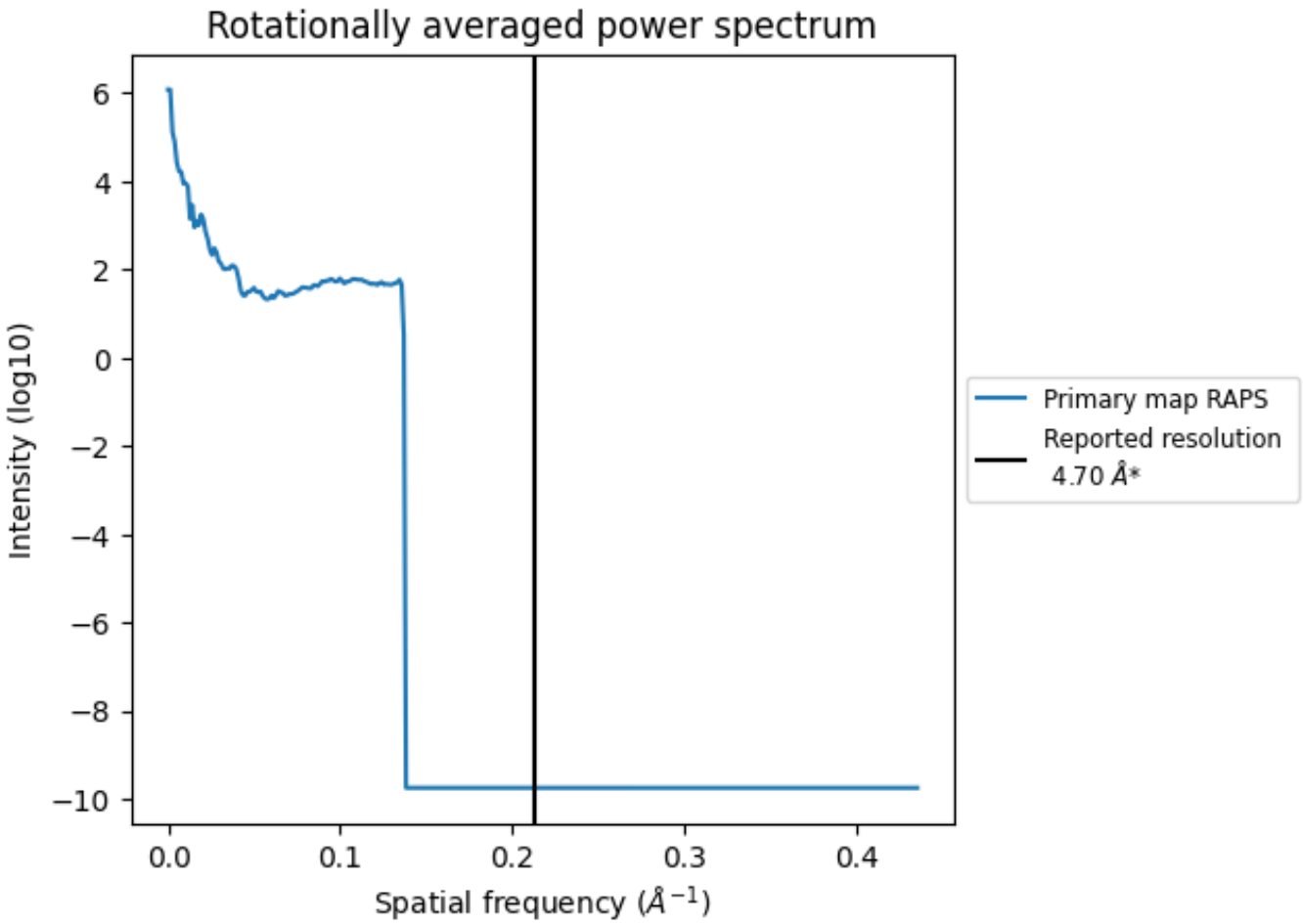
7.2 Volume estimate (i)



The volume at the recommended contour level is 15512 nm^3; this corresponds to an approximate mass of 14013 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 0.213 Å^{-1}

8 Fourier-Shell correlation (i)

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

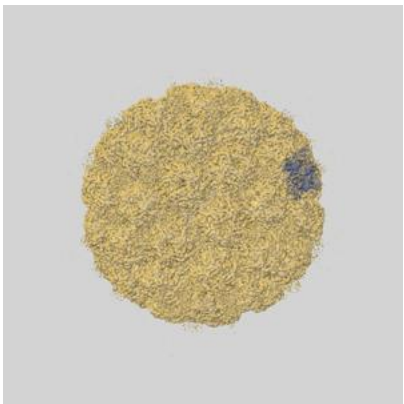
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-7136 and PDB model 6BSP. Per-residue inclusion information can be found in section 3 on page 4.

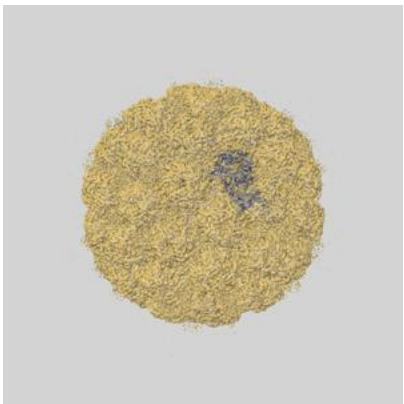
9.1 Map-model overlays

9.1.1 Map-model overlay (i)

X

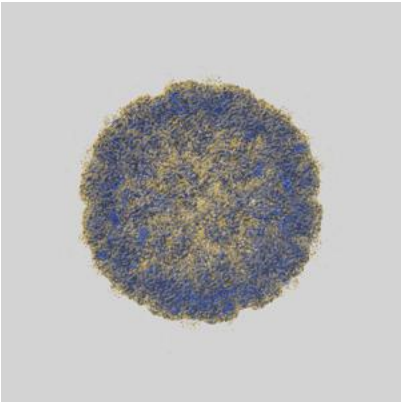


Y

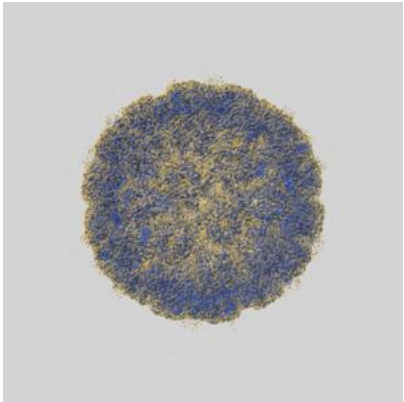


Z

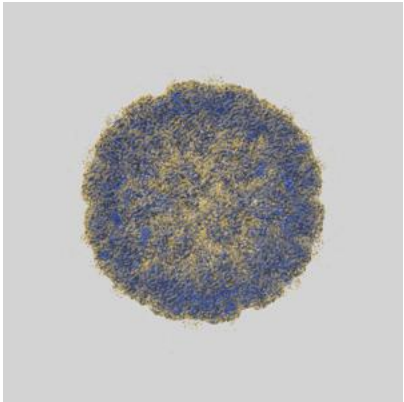
9.1.2 Map-model assembly overlay (i)



X



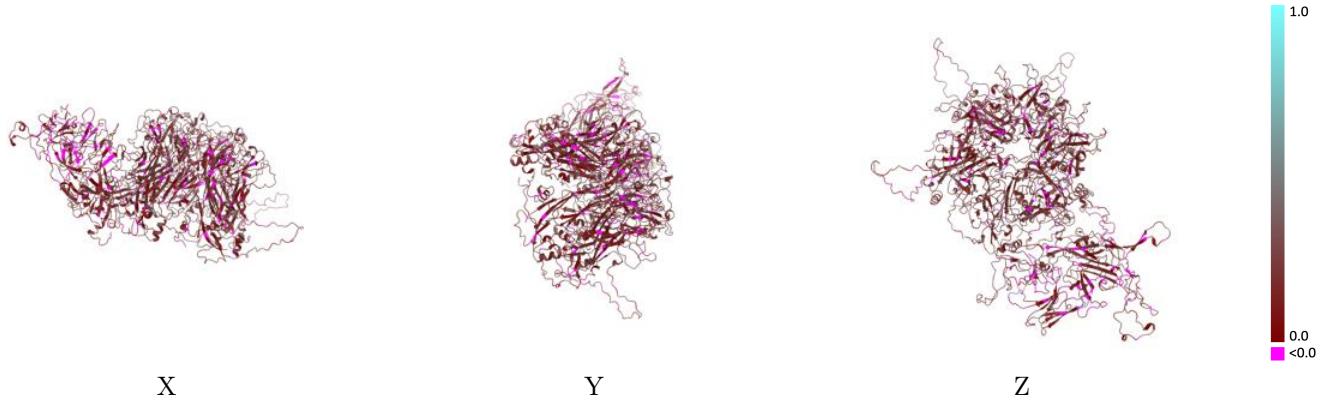
Y



Z

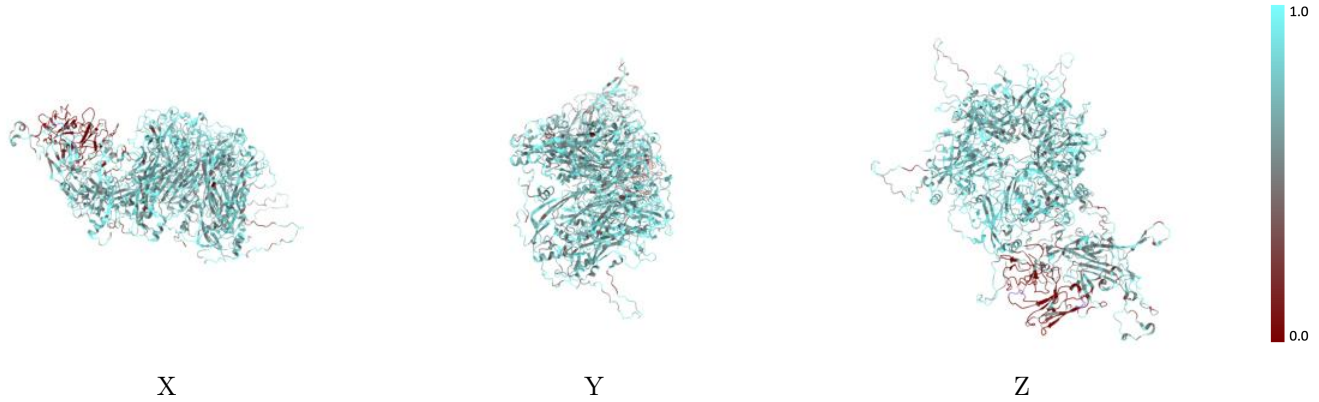
The images above show the 3D surface view of the map at the recommended contour level 2.1 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



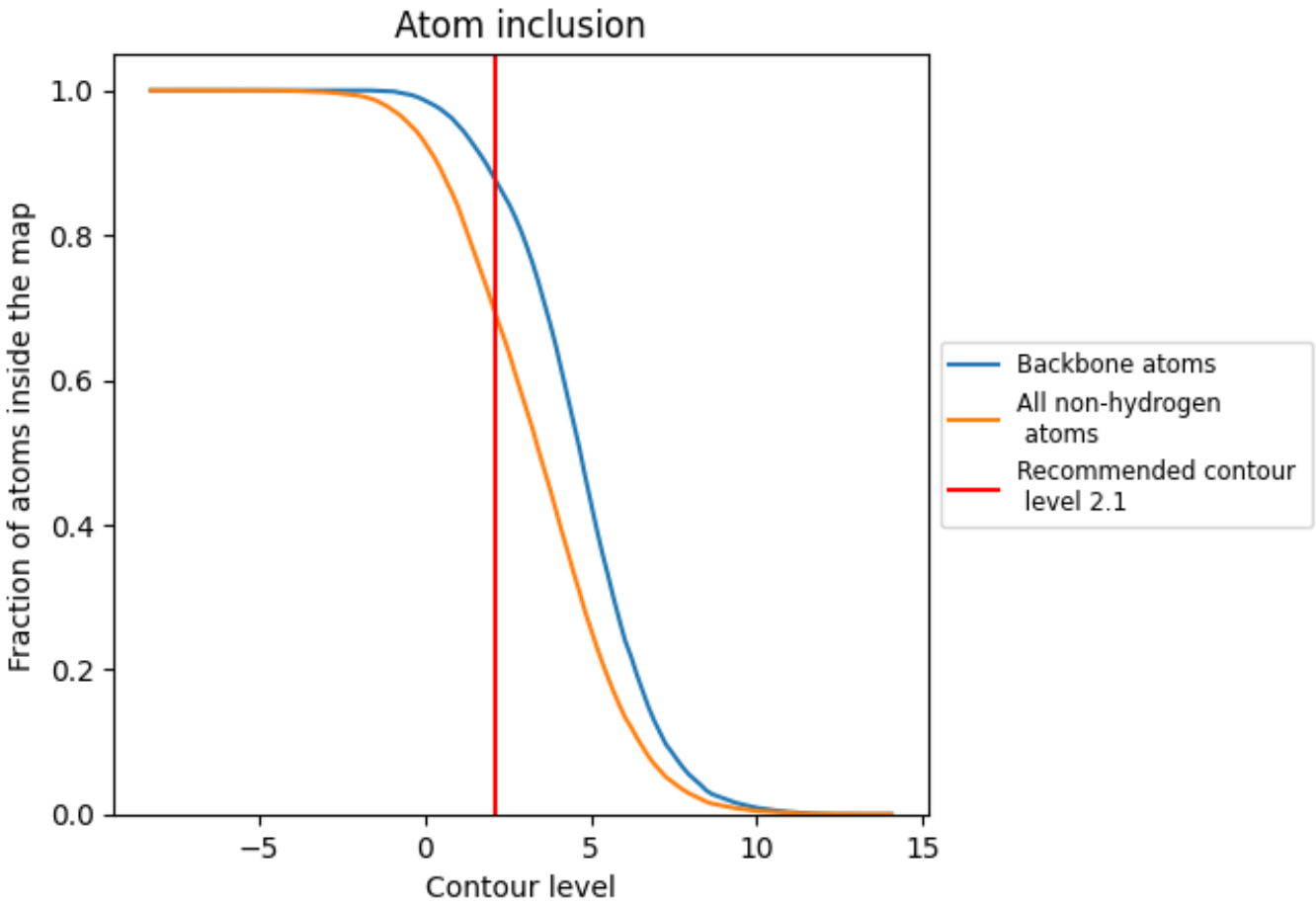
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (2.1).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 70% 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 (2.1) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6951	0.1930
A	0.1212	0.1620
B	0.1210	0.1410
C	0.6492	0.1550
D	0.8416	0.2590
E	0.7405	0.1890
F	0.7419	0.1960
G	0.7352	0.1910
H	0.7279	0.1870