wwPDB EM Validation Summary Report (i)

Nov 20, 2022 – 03:43 pm GMT

PDB ID	:	2XFC
EMDB ID	:	EMD-1015
Title	:	CHIKUNGUNYA E1 E2 ENVELOPE GLYCOPROTEINS FITTED IN SEMLIKI FOREST VIRUS cryo-EM MAP
Authors	:	Voss, J.E.; Vaney, M.C.; Duquerroy, S.; Rey, F.A.
Deposited on	:	2010-05-21
Resolution	:	9.00 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 (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:

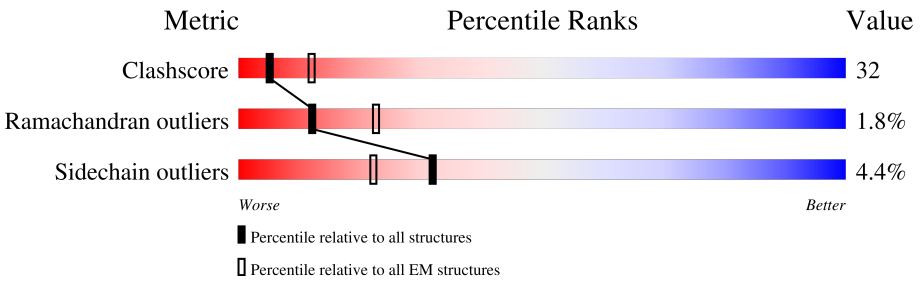
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 9.00 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)
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	439	38% 67% 20% • • 11%
1	D	439	40% 68% 18% • 11%
1	F	439	41% 66% 21% • 11%
1	H	439	46% 67% 20% • 11%
2	B	423	29% 63% 15% • 21%
2	E	423	32% 62% 16% • 21%
2	G	423	29% 59% 18% • 21%
2	I	423	34% 61% 17% • 21%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 22420 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 E1 ENVELOPE GLYCOPROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	391	2975	1881	499	571	24	0	1
1	D	391	2975	1881	499	571	24	0	1
1	F	391	2975	1881	499	571	24	0	1
1	H	391	2975	1881	499	571	24	0	1

- Molecule 2 is a protein called E2 ENVELOPE GLYCOPROTEIN.

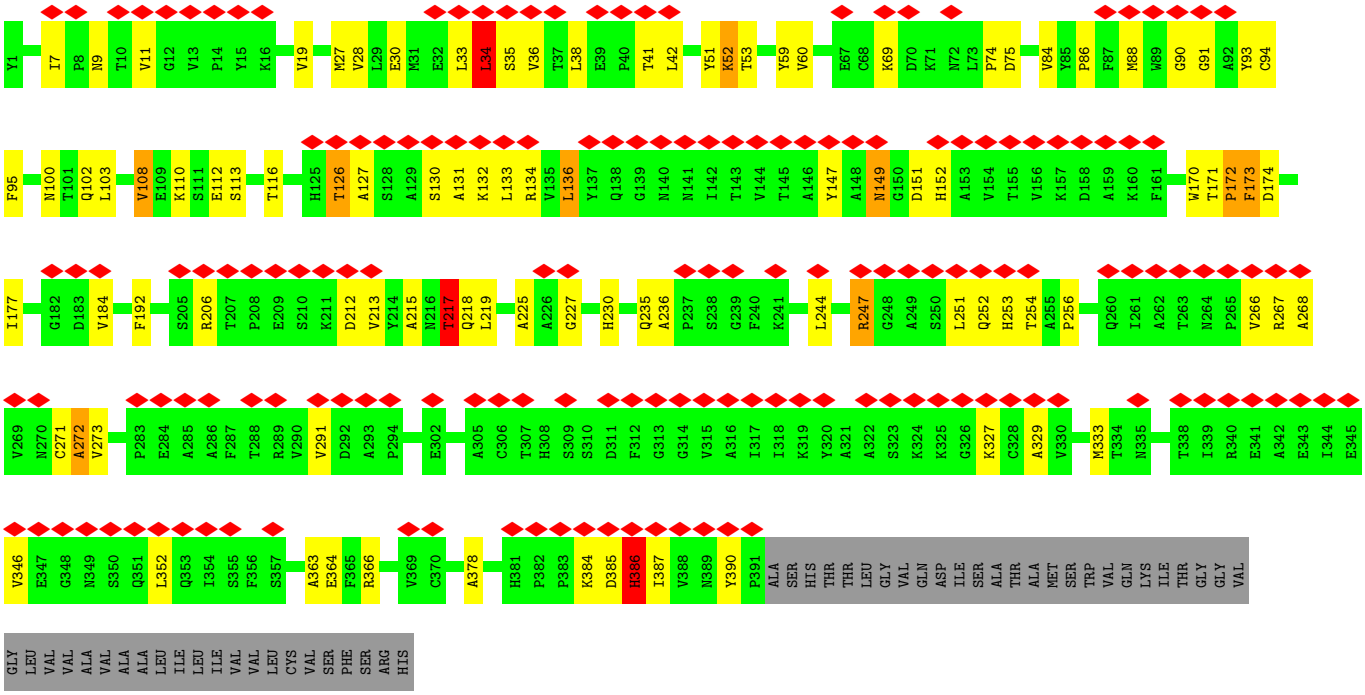
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	335	2630	1641	476	494	19	0	1
2	E	335	2630	1641	476	494	19	0	1
2	G	335	2630	1641	476	494	19	0	1
2	I	335	2630	1641	476	494	19	0	1

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.

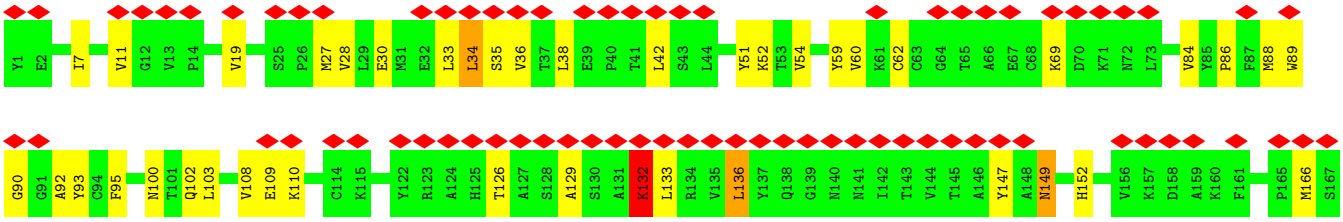
• Molecule 1: E1 ENVELOPE GLYCOPROTEIN

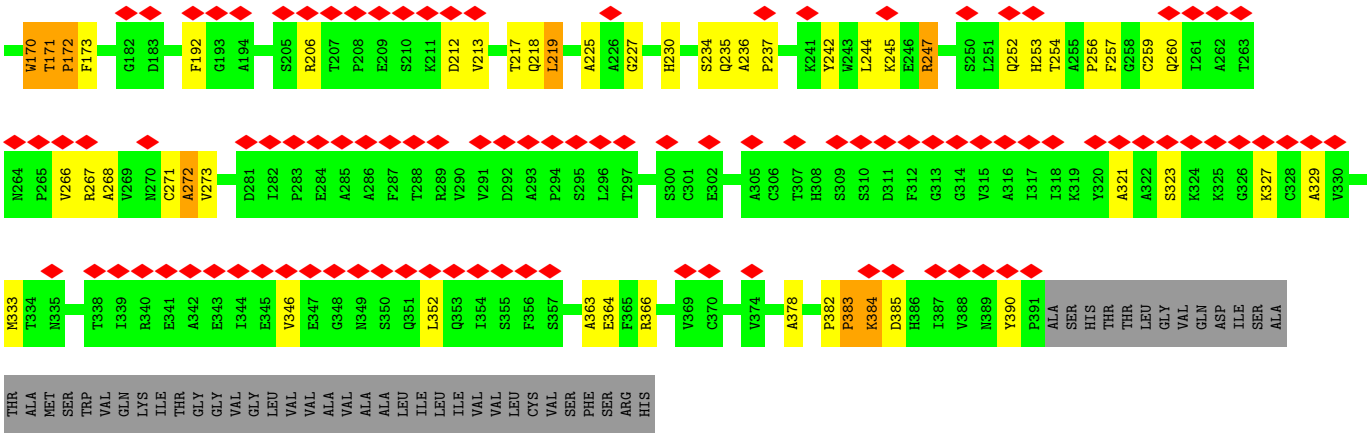
Chain A:



• Molecule 1: E1 ENVELOPE GLYCOPROTEIN

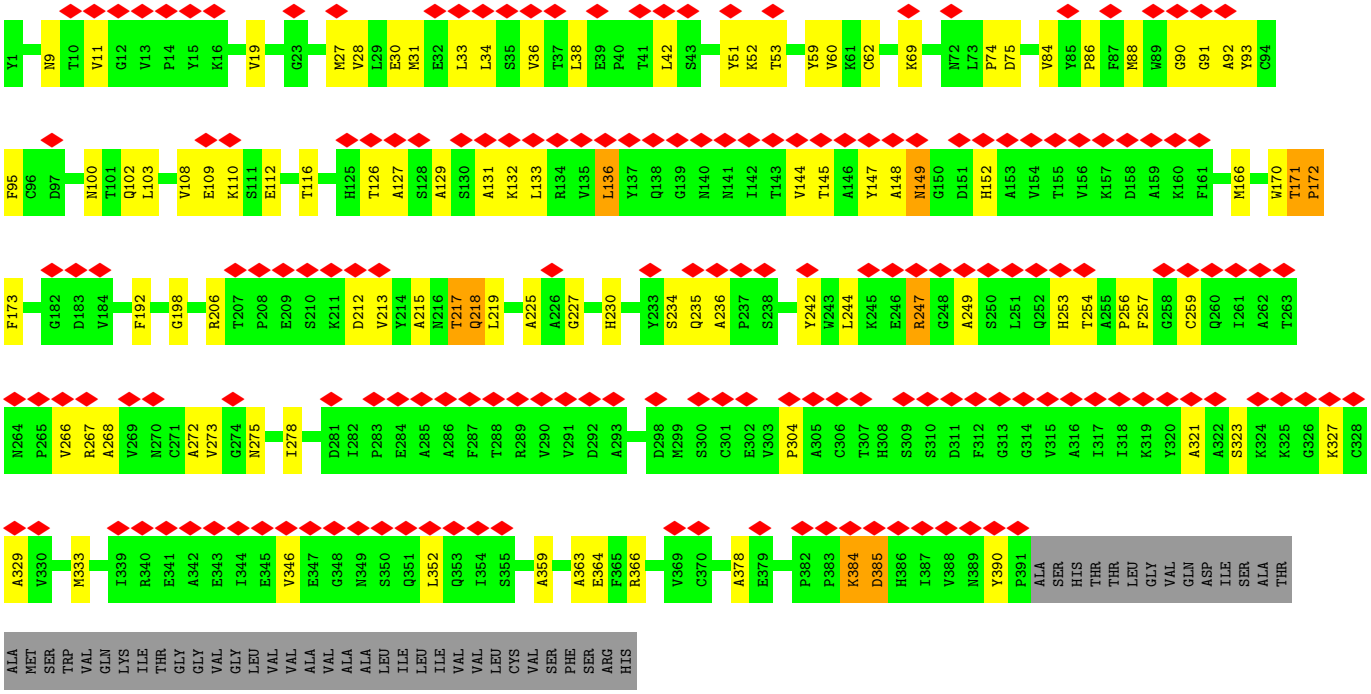
Chain D:





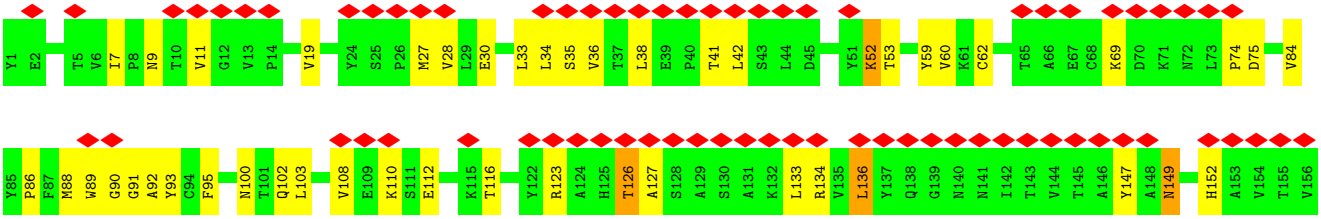
- Molecule 1: E1 ENVELOPE GLYCOPROTEIN

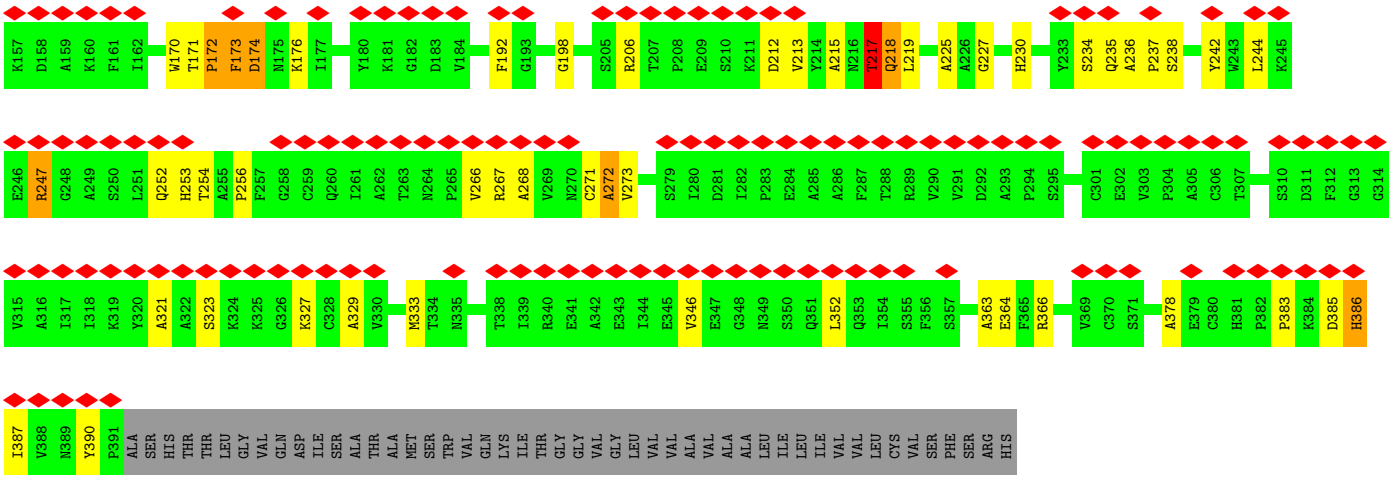
Chain F:



- Molecule 1: E1 ENVELOPE GLYCOPROTEIN

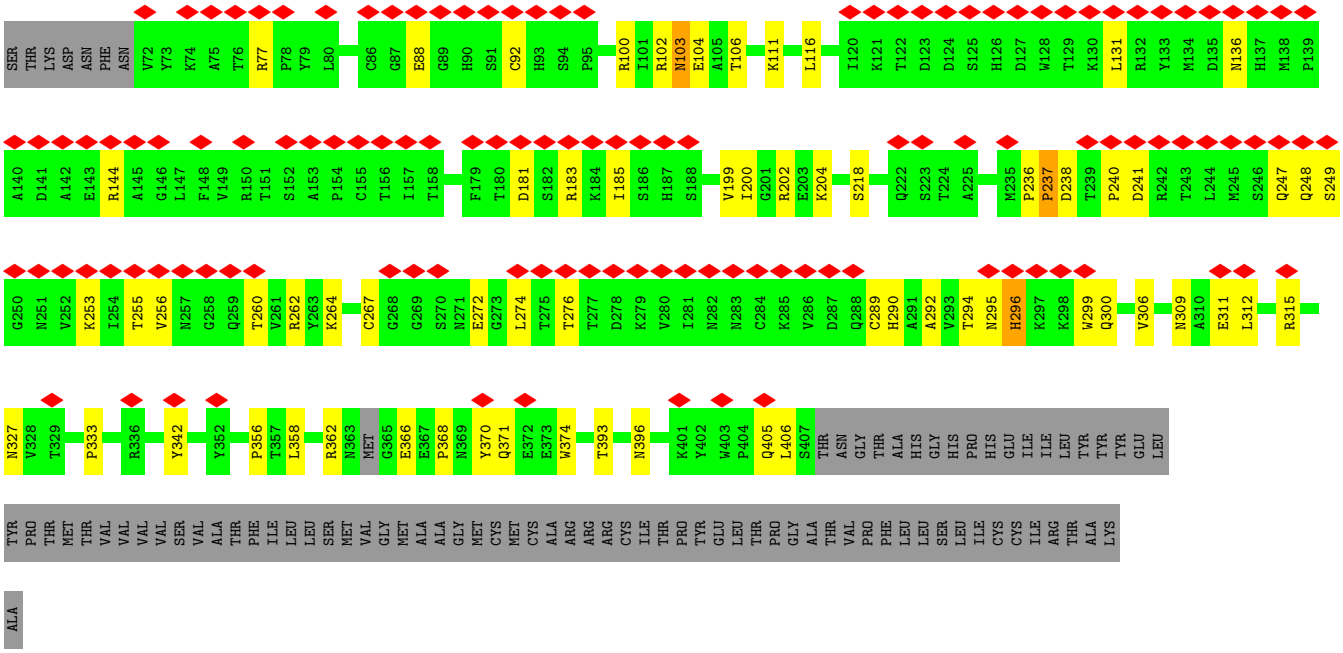
Chain H:





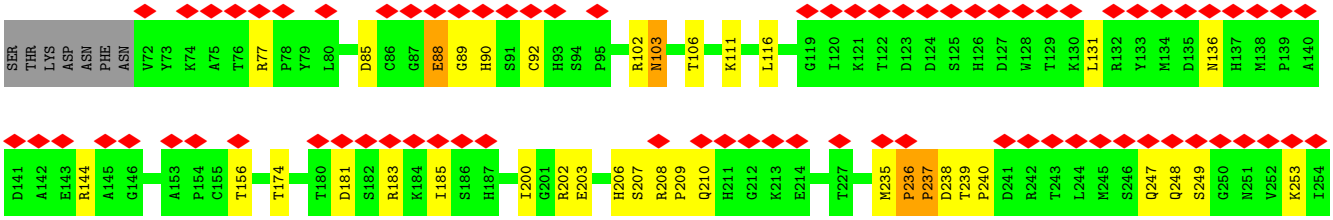
- Molecule 2: E2 ENVELOPE GLYCOPROTEIN

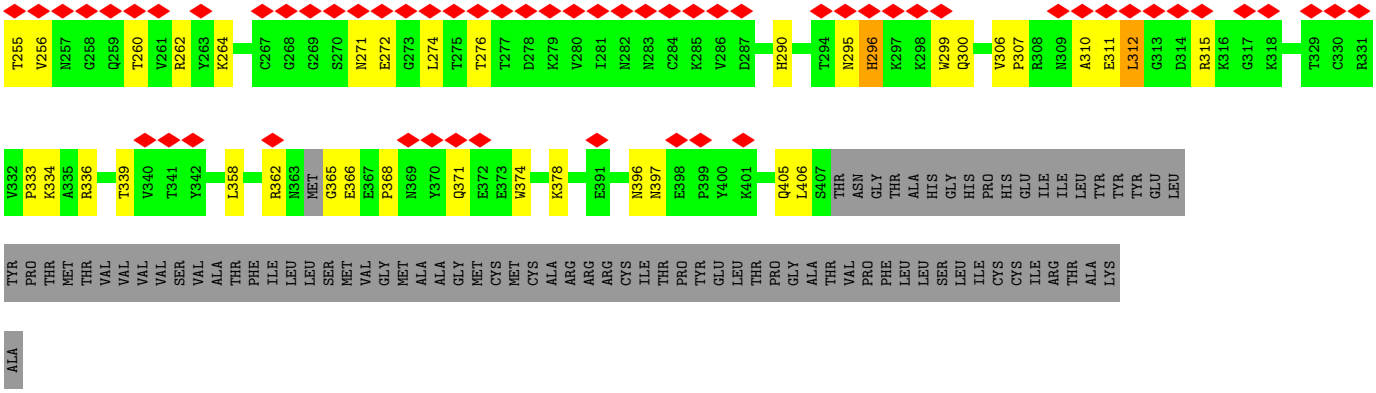
Chain B:



- Molecule 2: E2 ENVELOPE GLYCOPROTEIN

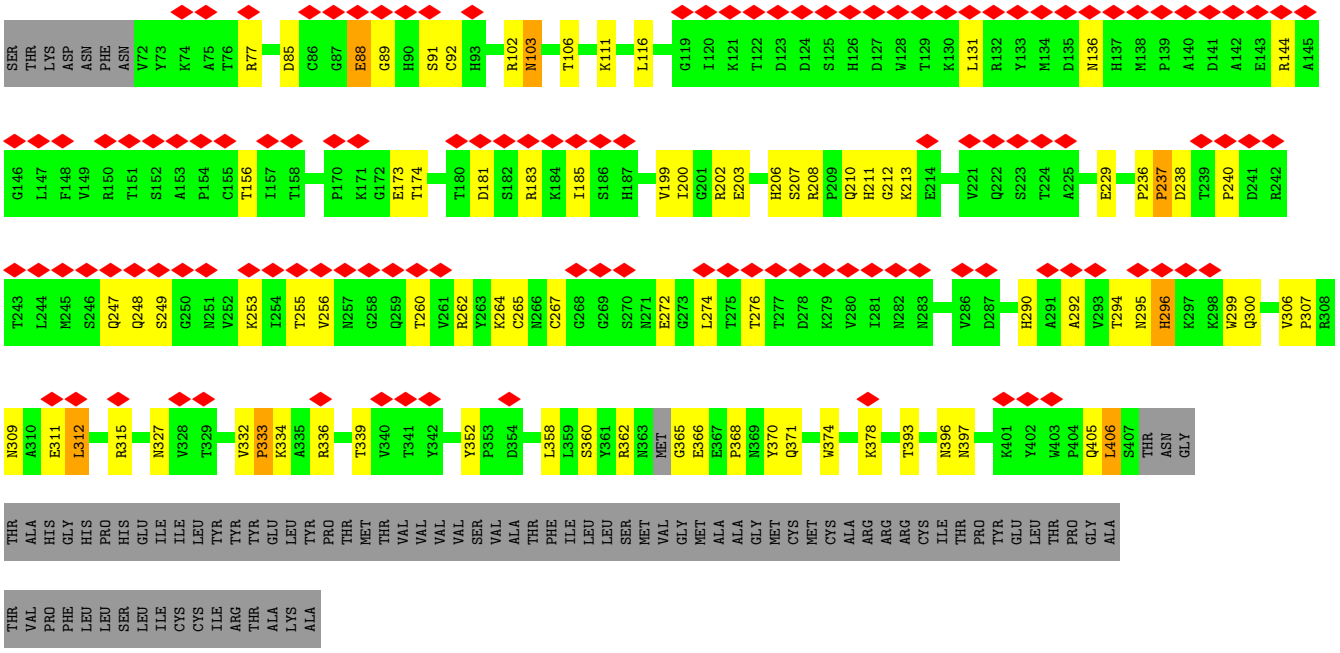
Chain E:





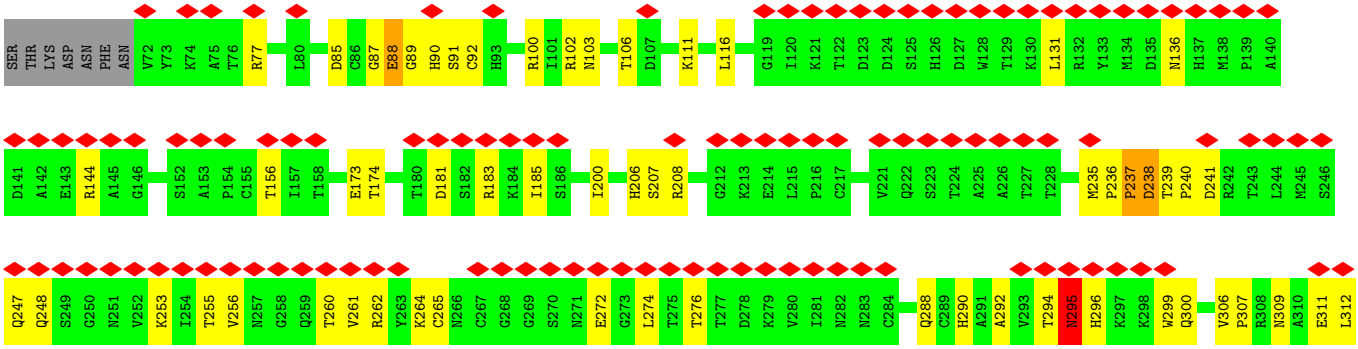
- Molecule 2: E2 ENVELOPE GLYCOPROTEIN

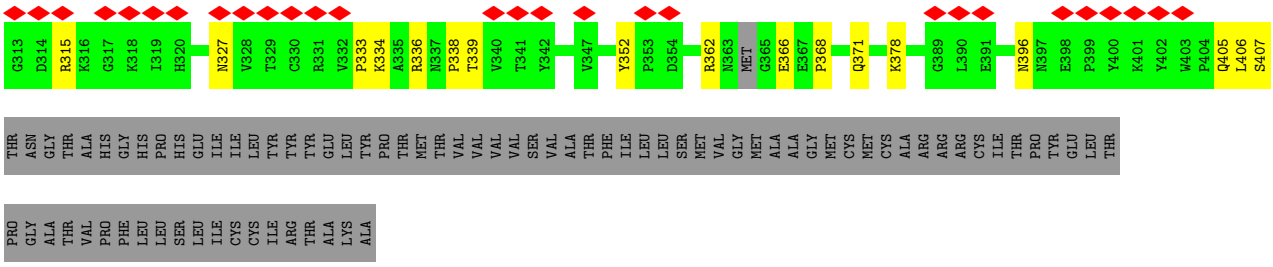
Chain G:



- Molecule 2: E2 ENVELOPE GLYCOPROTEIN

Chain I:




G313 THR
D314 ASN
R315 GLY
K316 THR
G317 ALA
K318 HIS
I319 GLY
H320 HIS
N327 GLU
V328 ILE
T329 ILE
C330 LEU
R331 TYR
V332 TYR
P333 GLU
K334 LEU
A335 TYR
R336 PRO
N337 THR
P338 MET
T339 THR
V340 VAL
T341 VAL
Y342 VAL
V347 VAL
Y352 VAL
P353 ALA
D354 THR
R362 PHE
N363 ILE
MET LEU
G365 LEU
E366 SER
E367 MET
P368 VAL
Q371 GLY
K378 MET
G389 ALA
L390 ALA
E391 GLY
N396 MET
N397 CYS
E398 MET
P399 CYS
Y400 ALA
K401 ARG
Y402 ARG
W403 ARG
P404 CYS
Q405 ILE
L406 THR
S407 PRO
TYR
GLU
LEU
THR
PRO
GLY
ALA
THR
VAL
PRO
PHE
LEU
LEU
SER
LEU
ILE
CYS
CYS
ILE
ARG
THR
ALA
LYS
ALA

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	6000	Depositor
Resolution determination method	Not provided	
CTF correction method	Not provided	
Microscope	FEI/PHILIPS CM200FEG/ST	Depositor
Voltage (kV)	100	Depositor
Electron dose ($e^-/Å^2$)	8	Depositor
Minimum defocus (nm)	975	Depositor
Maximum defocus (nm)	7628	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 FILM	Depositor
Maximum map value	25955.000	Depositor
Minimum map value	-25432.000	Depositor
Average map value	-39.429	Depositor
Map value standard deviation	4659.710	Depositor
Recommended contour level	9460.0	Depositor
Map size (Å)	758.52, 758.52, 758.52	wwPDB
Map dimensions	301, 301, 301	wwPDB
Map angles (°)	90, 90, 90	wwPDB
Pixel spacing (Å)	2.52, 2.52, 2.52	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	0.52	2/3046 (0.1%)	1.09	5/4147 (0.1%)
1	D	0.37	0/3047	0.59	2/4150 (0.0%)
1	F	0.58	1/3048 (0.0%)	0.77	5/4153 (0.1%)
1	H	0.61	3/3050 (0.1%)	0.83	7/4159 (0.2%)
2	B	0.37	0/2697	0.62	0/3668
2	E	0.40	0/2698	0.63	0/3671
2	G	0.37	0/2698	0.62	0/3671
2	I	0.37	0/2699	0.61	0/3674
All	All	0.46	6/22983 (0.0%)	0.74	19/31293 (0.1%)

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
1	D	0	1
1	F	0	1
1	H	0	2
All	All	0	5

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	217	THR	C-N	24.85	1.91	1.34
1	H	217	THR	C-N	-16.38	0.96	1.34
1	A	52	LYS	C-N	15.38	1.69	1.34
1	H	52	LYS	C-N	15.20	1.69	1.34
1	A	108	VAL	C-N	14.51	1.67	1.34

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	217	THR	O-C-N	-56.87	31.71	122.70
1	H	217	THR	O-C-N	-27.84	78.15	122.70
1	F	217	THR	O-C-N	-26.85	79.73	122.70
1	H	217	THR	CA-C-N	16.72	153.99	117.20
1	H	217	THR	C-N-CA	14.46	157.85	121.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	217	THR	Mainchain
1	D	321	ALA	Mainchain
1	F	321	ALA	Mainchain
1	H	217	THR	Mainchain
1	H	321	ALA	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	2975	0	2877	286	0
1	D	2975	0	2875	274	0
1	F	2975	0	2876	270	0
1	H	2975	0	2877	259	0
2	B	2630	0	2550	179	0
2	E	2630	0	2546	252	0
2	G	2630	0	2546	302	0
2	I	2630	0	2544	314	0
All	All	22420	0	21691	1424	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 32.

The worst 5 of 1424 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)
2:I:274:LEU:HD21	2:I:296:HIS:CE1	1.20	1.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:34:LEU:HD12	1:A:132:LYS:CE	1.18	1.58
1:H:88:MET:HB3	2:I:240:PRO:CD	1.29	1.57
1:A:34:LEU:CD1	1:A:132:LYS:HE2	1.27	1.57
2:G:206:HIS:CG	2:I:174:THR:HG23	1.34	1.57

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	380/439 (87%)	356 (94%)	16 (4%)	8 (2%)	7	36
1	D	381/439 (87%)	358 (94%)	15 (4%)	8 (2%)	7	36
1	F	383/439 (87%)	365 (95%)	14 (4%)	4 (1%)	15	55
1	H	387/439 (88%)	364 (94%)	14 (4%)	9 (2%)	6	34
2	B	327/423 (77%)	303 (93%)	19 (6%)	5 (2%)	10	46
2	E	329/423 (78%)	301 (92%)	22 (7%)	6 (2%)	8	40
2	G	329/423 (78%)	303 (92%)	19 (6%)	7 (2%)	7	36
2	I	331/423 (78%)	300 (91%)	26 (8%)	5 (2%)	10	46
All	All	2847/3448 (83%)	2650 (93%)	145 (5%)	52 (2%)	12	40

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	34	LEU
1	A	173	PHE
1	A	272	ALA
1	A	386	HIS
1	A	390	TYR

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	327/366 (89%)	314 (96%)	13 (4%)	31	55
1	D	327/366 (89%)	313 (96%)	14 (4%)	29	53
1	F	327/366 (89%)	312 (95%)	15 (5%)	27	52
1	H	327/366 (89%)	316 (97%)	11 (3%)	37	60
2	B	295/372 (79%)	282 (96%)	13 (4%)	28	53
2	E	295/372 (79%)	281 (95%)	14 (5%)	26	51
2	G	295/372 (79%)	281 (95%)	14 (5%)	26	51
2	I	295/372 (79%)	280 (95%)	15 (5%)	24	48
All	All	2488/2952 (84%)	2379 (96%)	109 (4%)	32	53

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	171	THR
2	G	103	ASN
2	I	238	ASP
1	F	219	LEU
1	F	352	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
2	G	210	GLN
1	H	230	HIS
2	G	296	HIS
2	G	396	ASN
1	H	373	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	7
1	D	4
1	F	4
1	H	4
2	B	2
2	G	1
2	E	1

The worst 5 of 23 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	321:ALA	C	322:ALA	N	20.90
1	A	323:SER	C	324:LYS	N	19.06
1	A	283:PRO	C	284:GLU	N	13.12
1	A	284:GLU	C	285:ALA	N	12.29
1	D	235:GLN	C	236:ALA	N	4.87

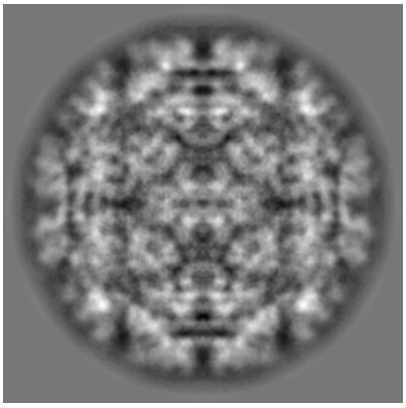
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-1015. These allow visual inspection of the internal detail of the map and identification of artifacts.

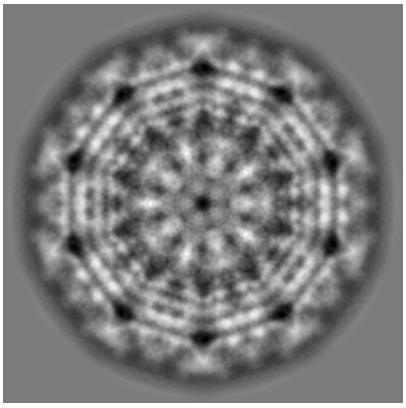
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

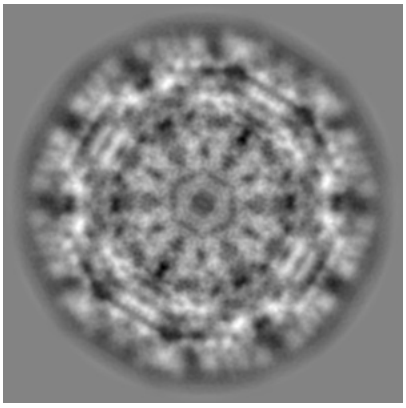
6.1.1 Primary map



X



Y

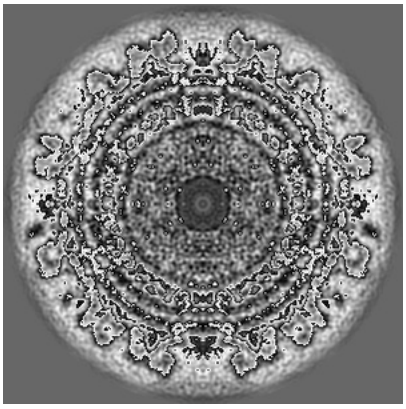


Z

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

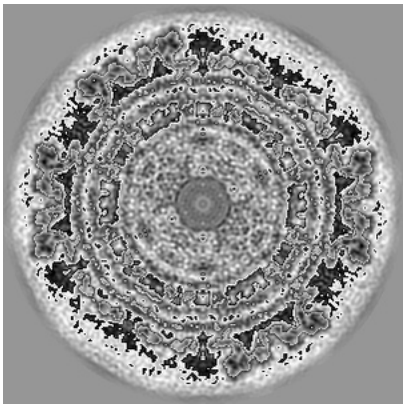
6.2 Central slices (i)

6.2.1 Primary map



X Index: 150

Y Index: 150

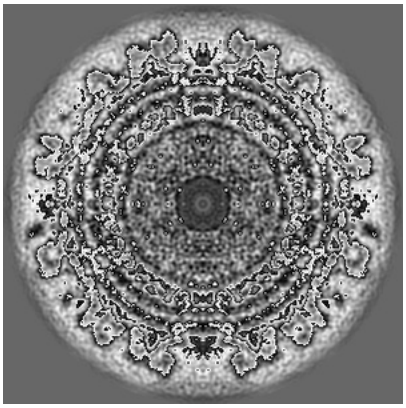


Z Index: 150

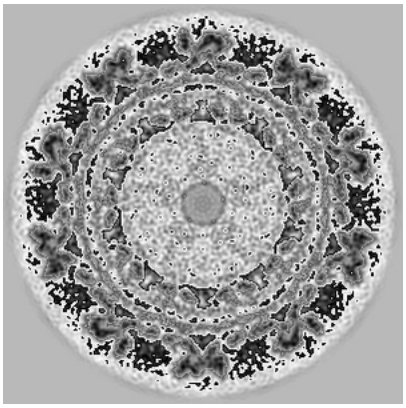
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

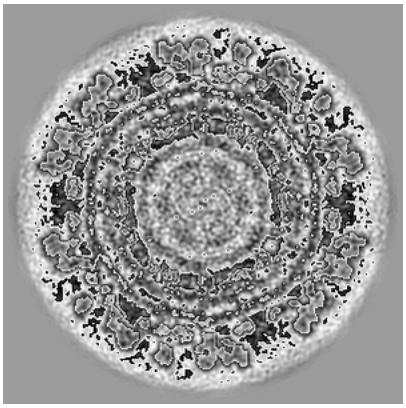
6.3.1 Primary map



X Index: 150



Y Index: 137

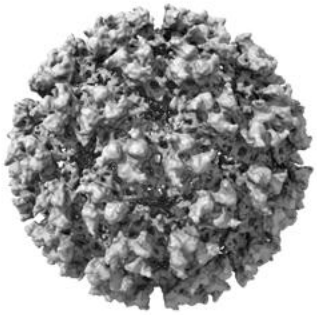


Z Index: 112

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

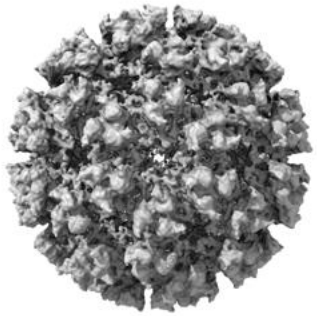
6.4 Orthogonal surface views (i)

6.4.1 Primary map



X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 9460.0. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

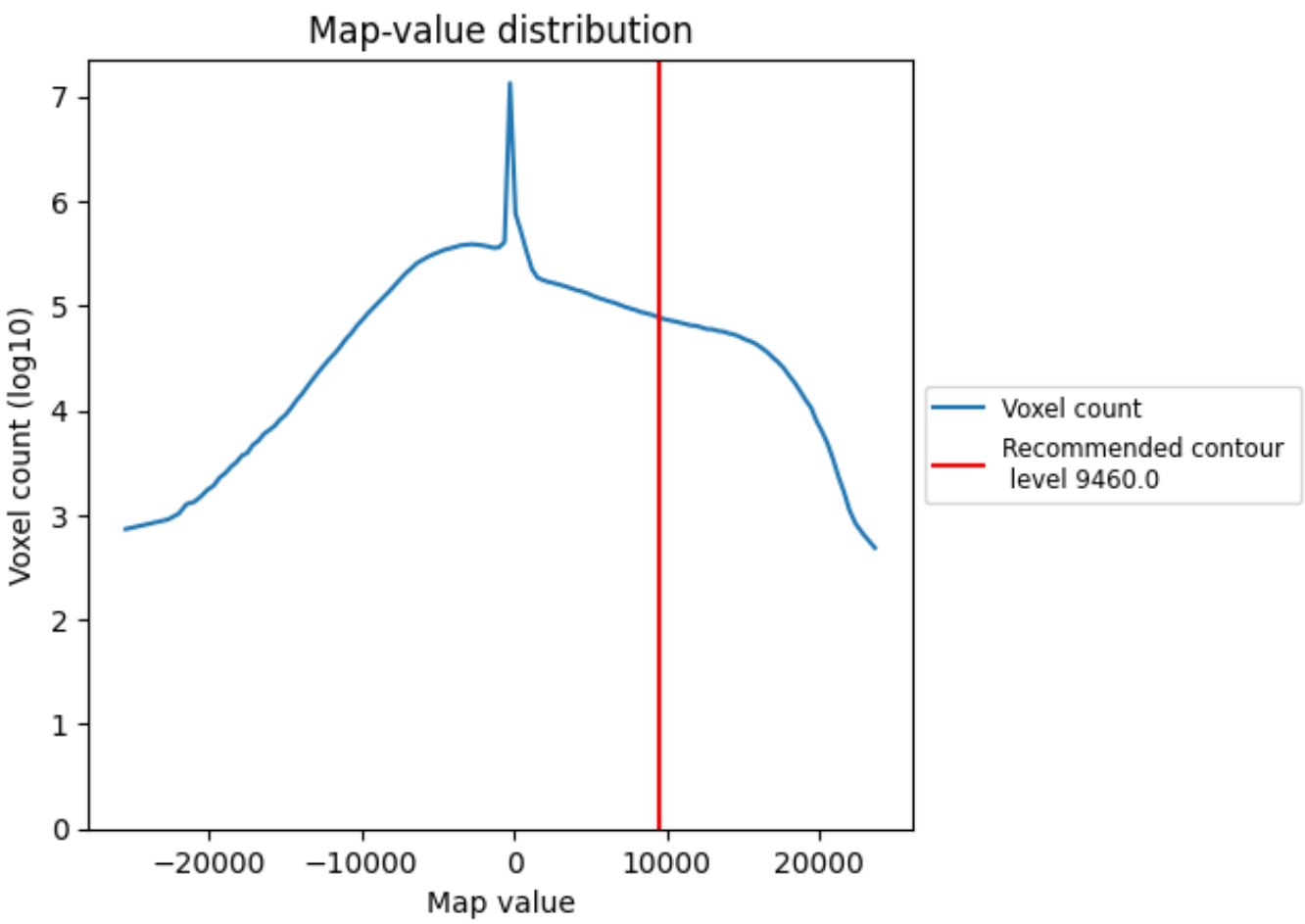
6.5 Mask visualisation (i)

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

7 Map analysis (i)

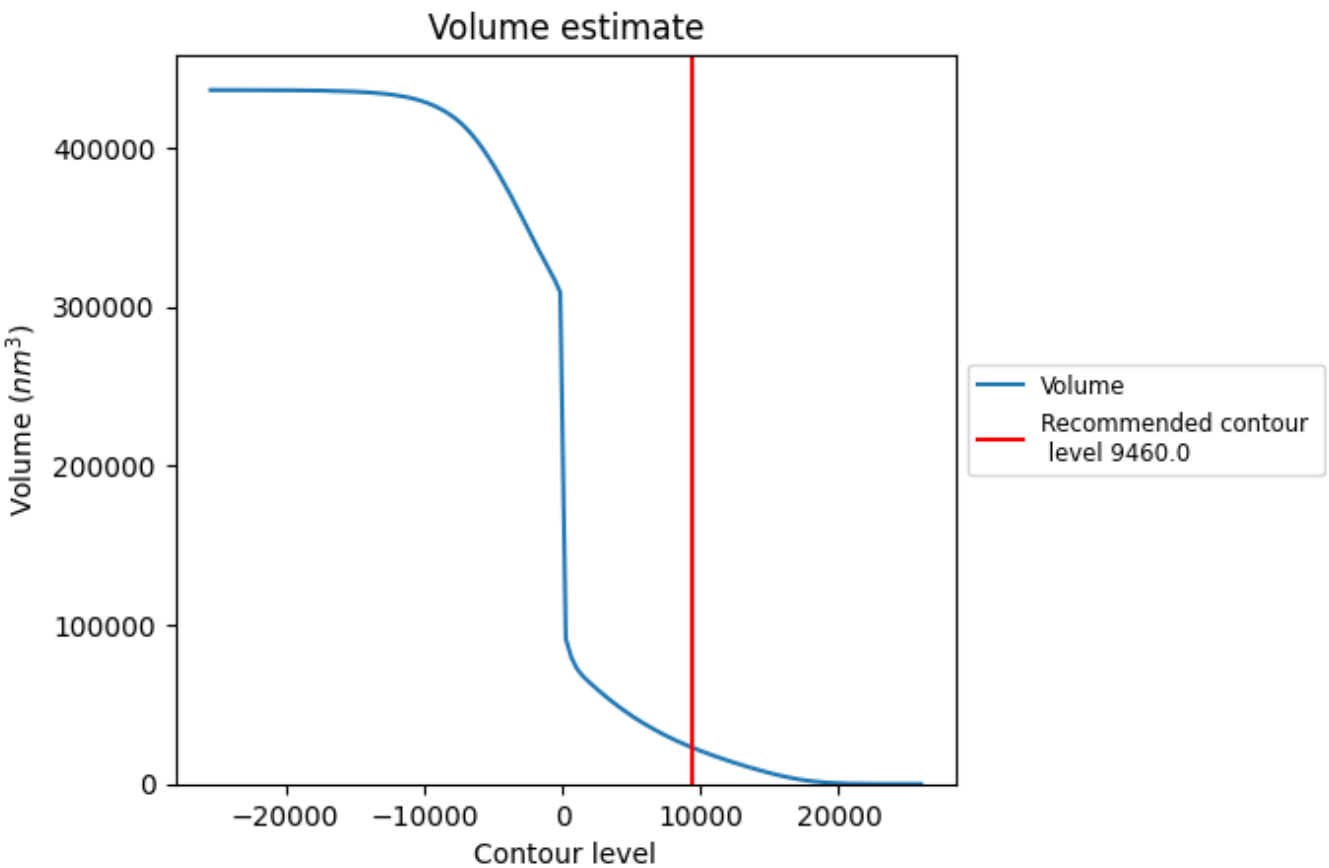
This section contains the results of statistical analysis of the map.

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.

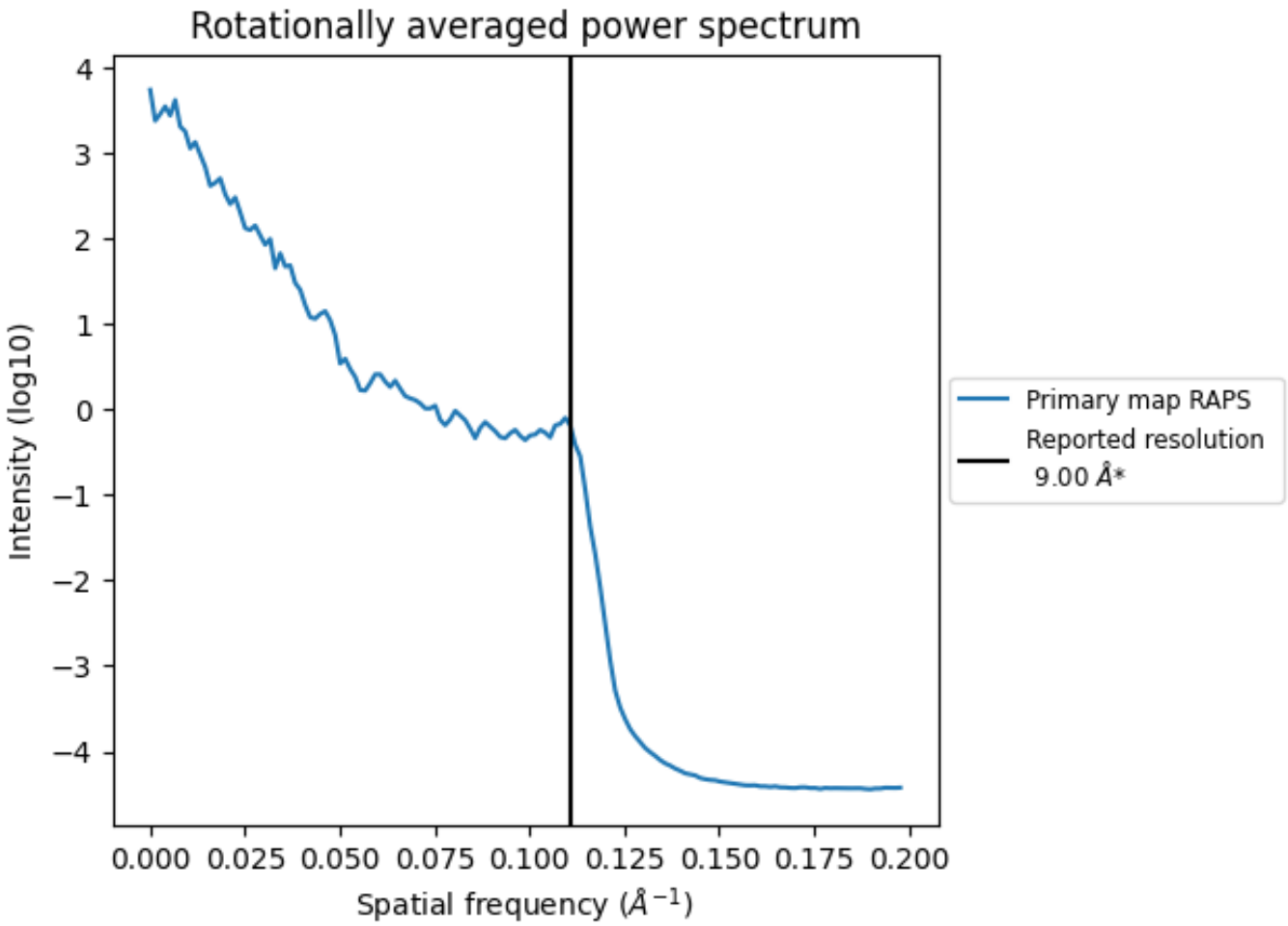
7.2 Volume estimate (i)



The volume at the recommended contour level is 22687 nm^3; this corresponds to an approximate mass of 20494 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 0.111 Å^{-1}

8 Fourier-Shell correlation (i)

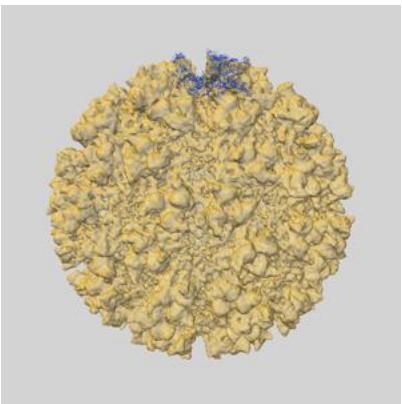
This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit (i)

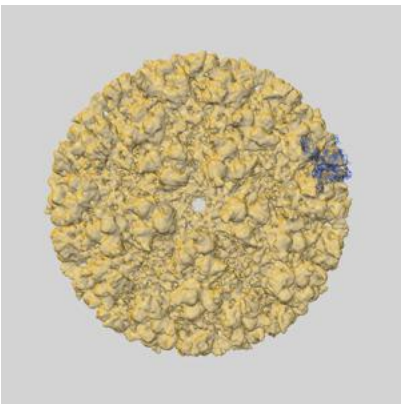
This section contains information regarding the fit between EMDB map EMD-1015 and PDB model 2XFC. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlays

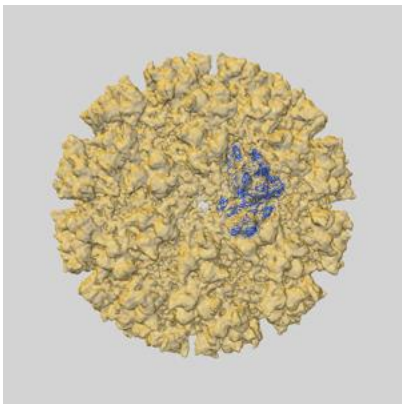
9.1.1 Map-model overlay (i)



X

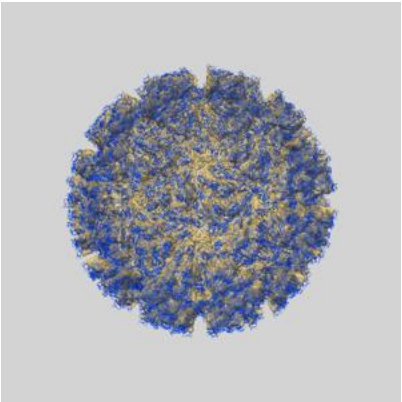


Y

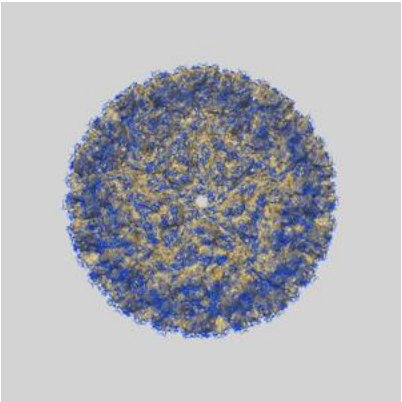


Z

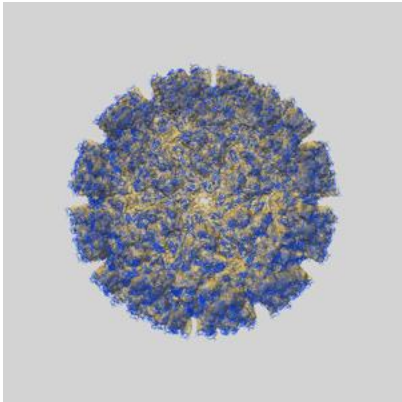
9.1.2 Map-model assembly overlay (i)



X



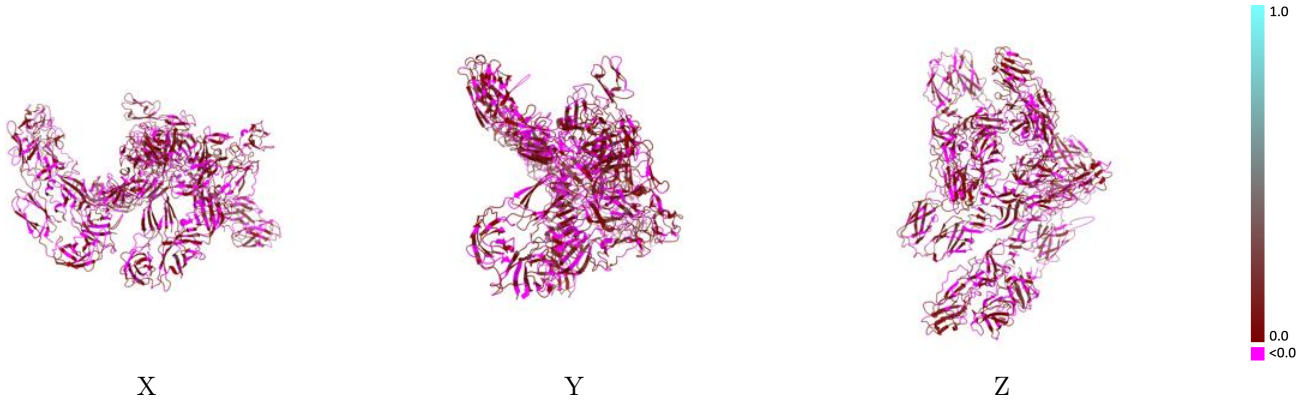
Y



Z

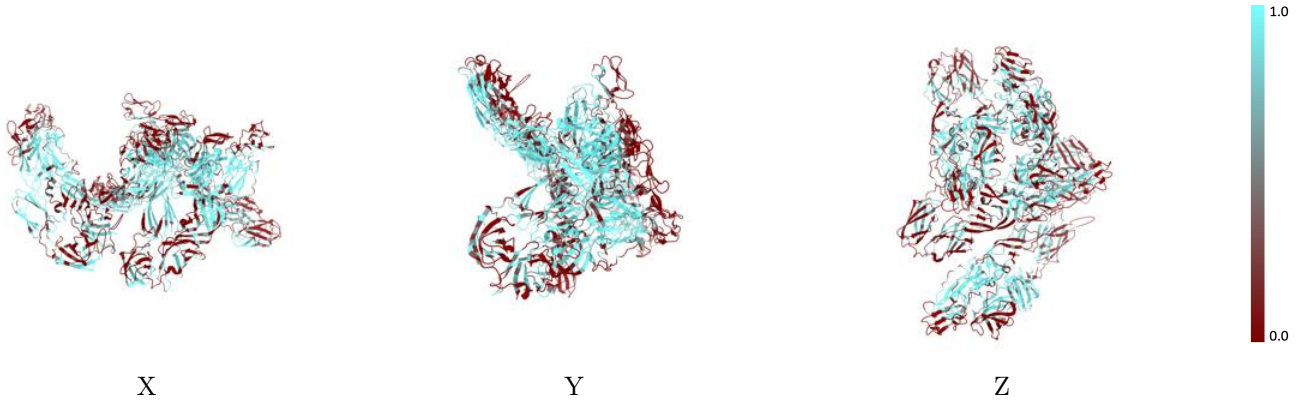
The images above show the 3D surface view of the map at the recommended contour level 9460.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



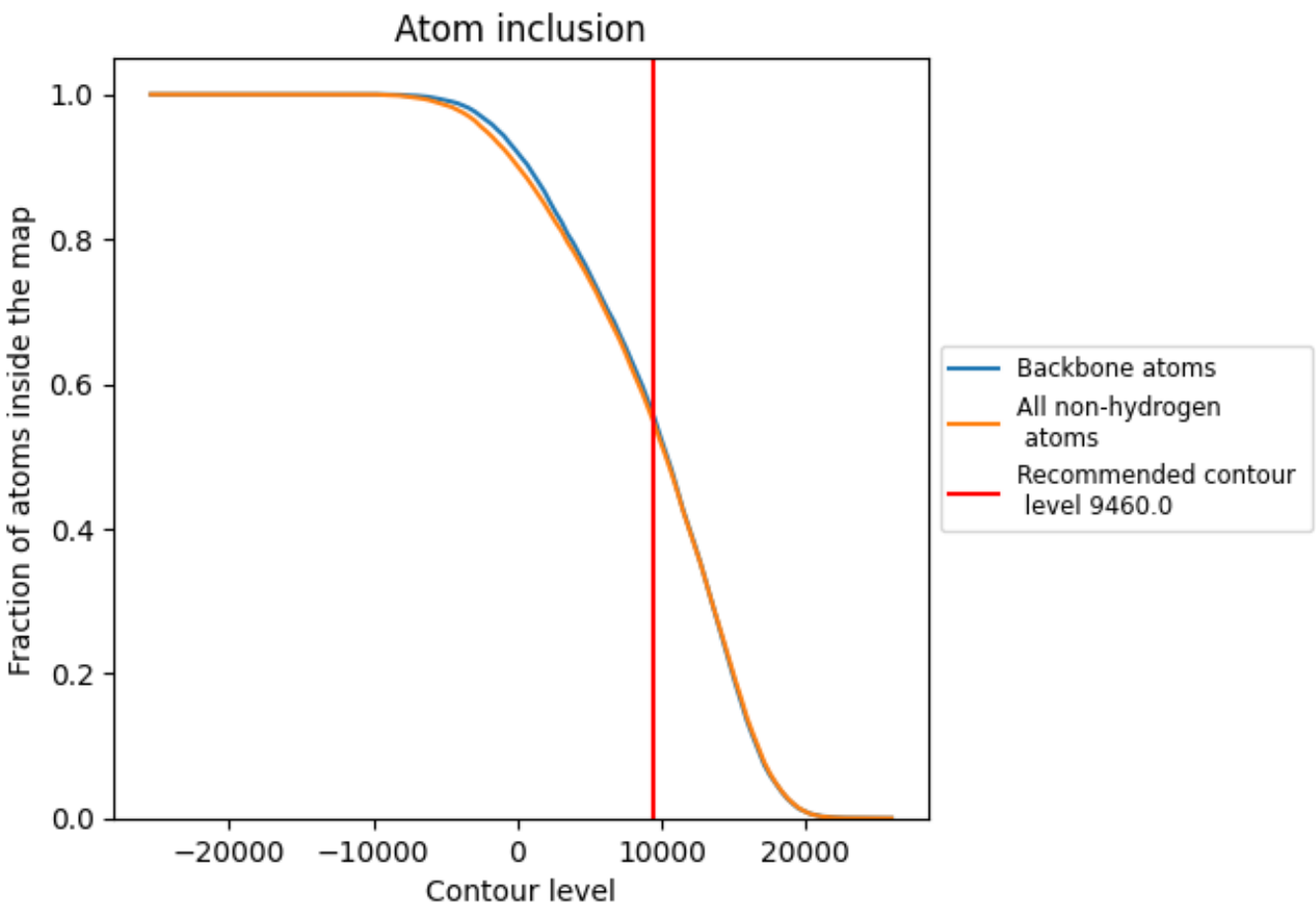
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (9460.0).

9.4 Atom inclusion (i)



At the recommended contour level, 56% of all backbone atoms, 55% 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 (9460.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5468	0.0400
A	0.5526	0.0290
B	0.5983	0.0460
D	0.5239	0.0400
E	0.5680	0.0440
F	0.5324	0.0260
G	0.6169	0.0530
H	0.4515	0.0360
I	0.5483	0.0510