wwPDB EM Validation Summary Report

Apr 3, 2024 – 01:54 pm BST

PDB ID	:	8C8M
EMDB ID	:	EMD-16484
Title	:	In vitro structure of the Nitrosopumilus maritimus S-layer - Composite map between two and six-fold symmetrised
Authors	:	von Kuegelgen, A.; Bharat, T.
Deposited on	:	2023-01-20
Resolution	:	2.87 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:

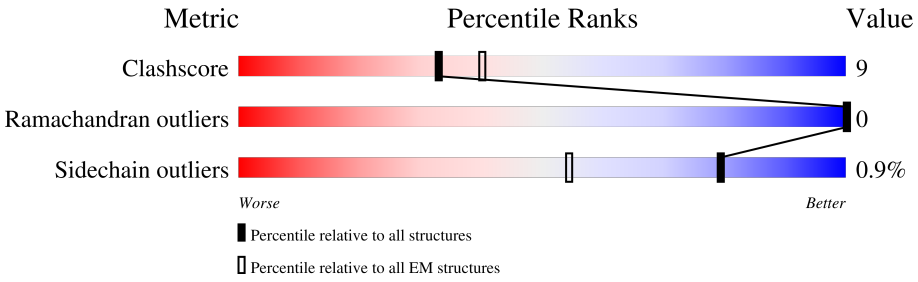
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.87 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	1734	72% 18% • 9%
1	B	1734	73% 18% 9%
1	C	1734	73% 18% 9%
1	D	1734	73% 18% 9%
1	E	1734	73% 18% 9%
1	F	1734	72% 18% 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 70506 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 Cell surface protein.

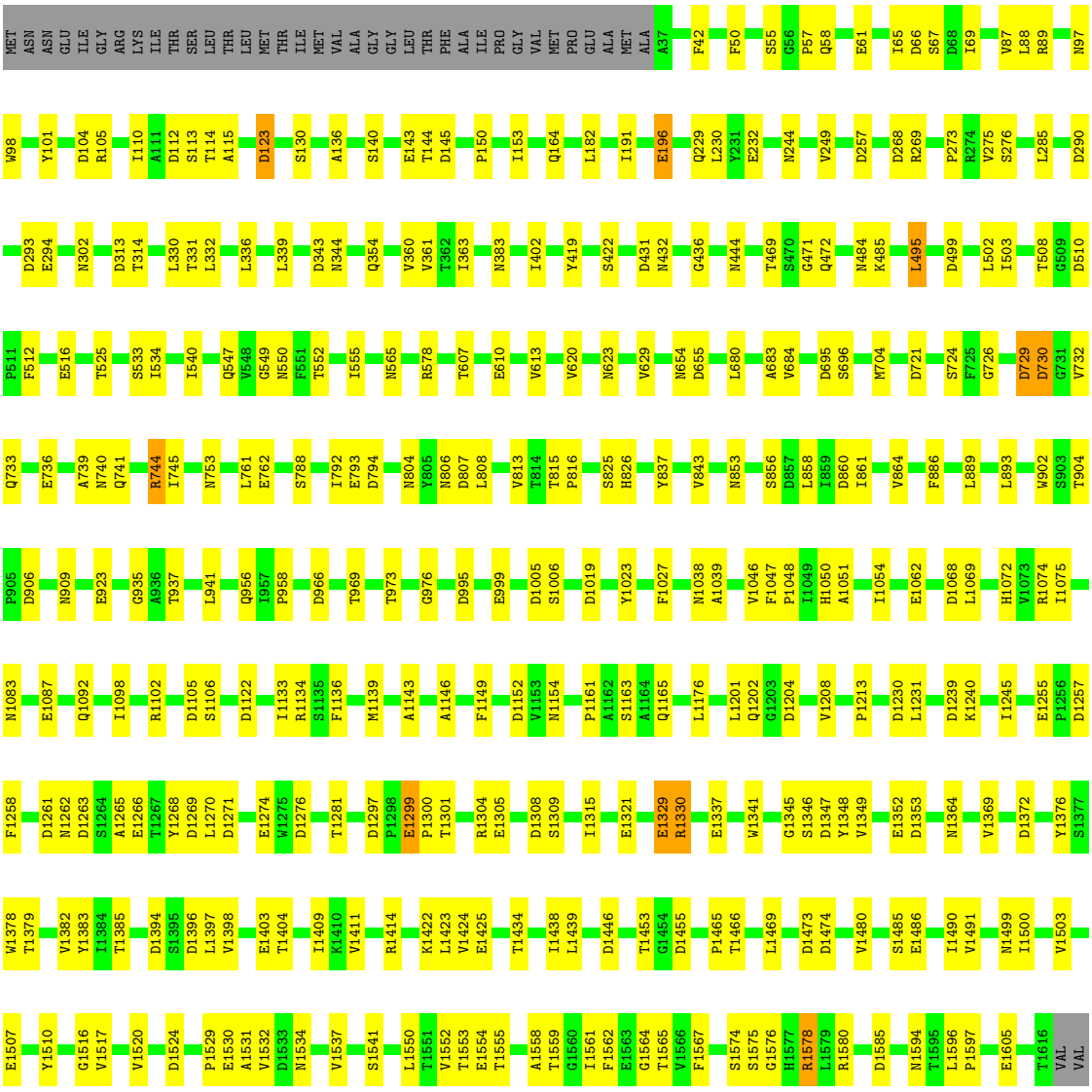
Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1580	Total	C	N	O	S	0	0
			11751	7234	1888	2610	19		
1	B	1580	Total	C	N	O	S	0	0
			11751	7234	1888	2610	19		
1	C	1580	Total	C	N	O	S	0	0
			11751	7234	1888	2610	19		
1	D	1580	Total	C	N	O	S	0	0
			11751	7234	1888	2610	19		
1	E	1580	Total	C	N	O	S	0	0
			11751	7234	1888	2610	19		
1	F	1580	Total	C	N	O	S	0	0
			11751	7234	1888	2610	19		

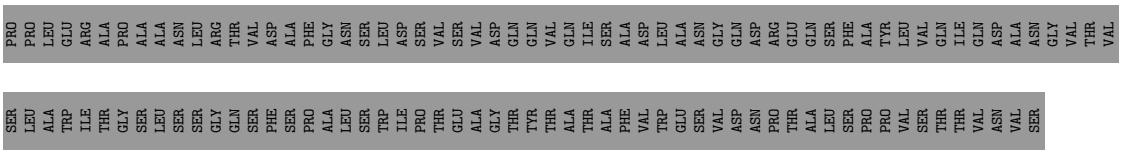
3 Residue-property plots (i)

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: Cell surface protein

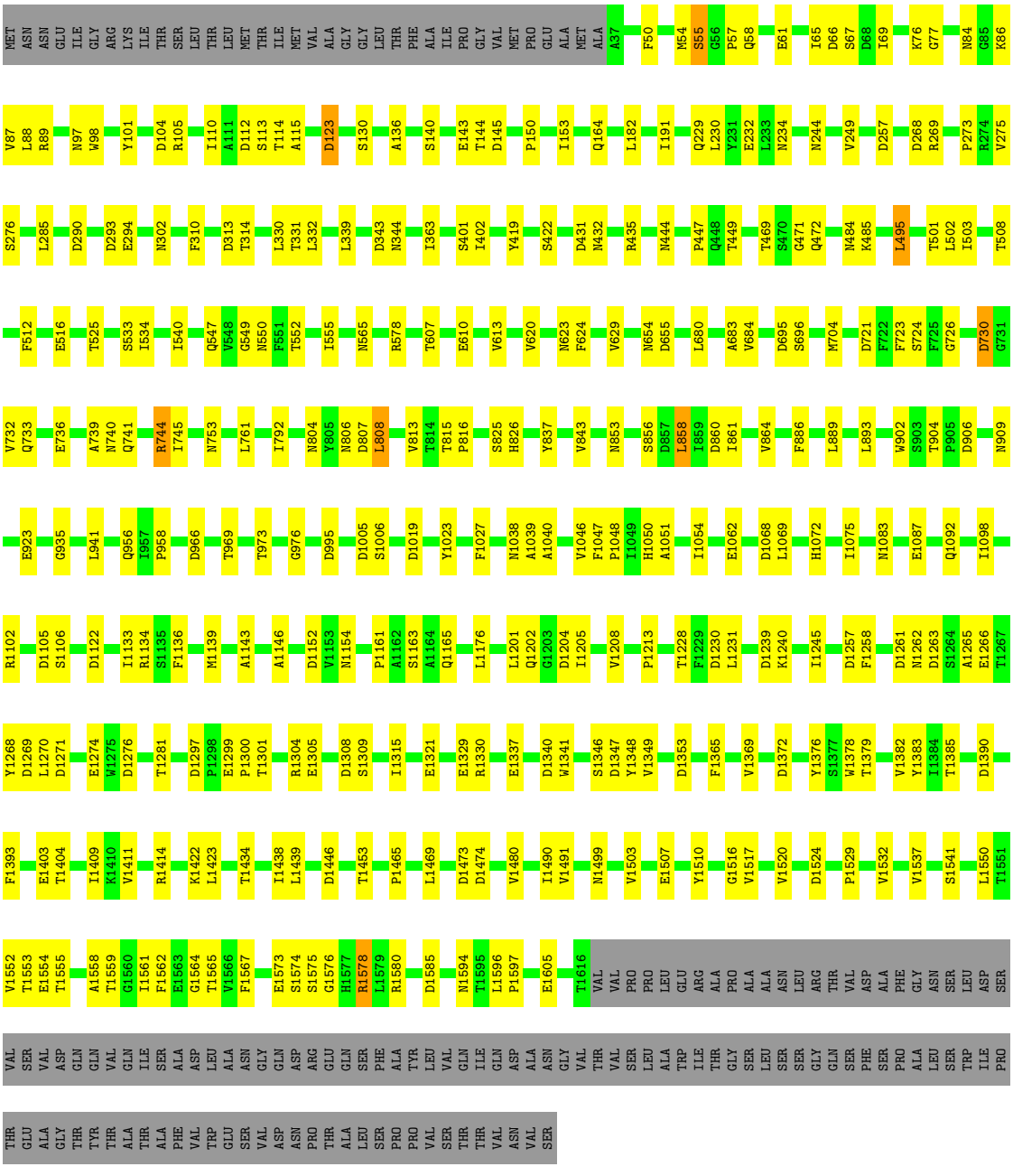
Chain A:





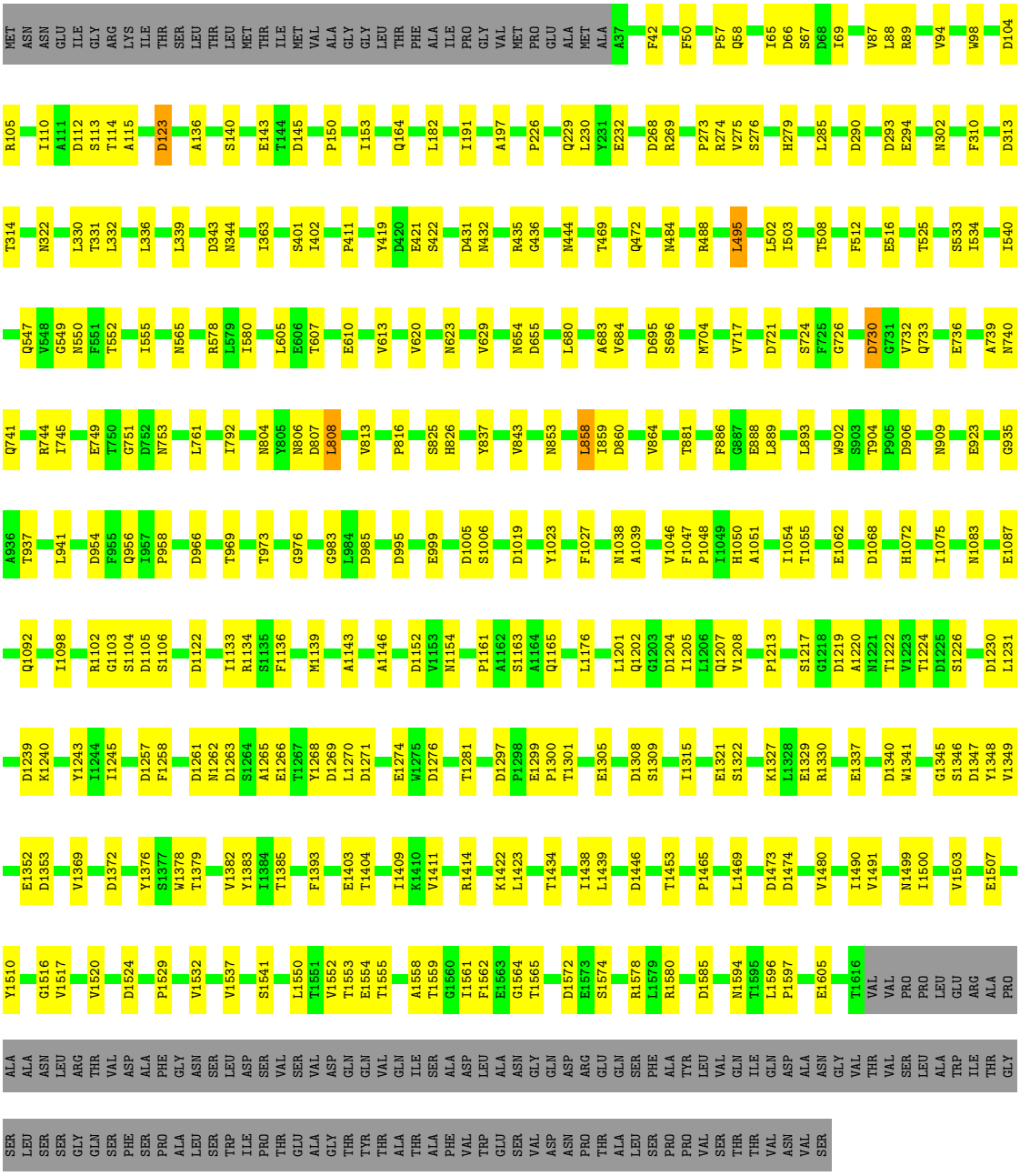
• Molecule 1: Cell surface protein

Chain B:



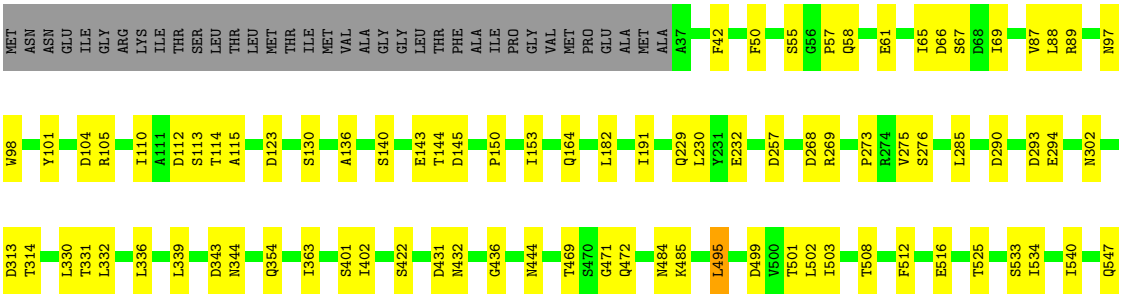
• Molecule 1: Cell surface protein

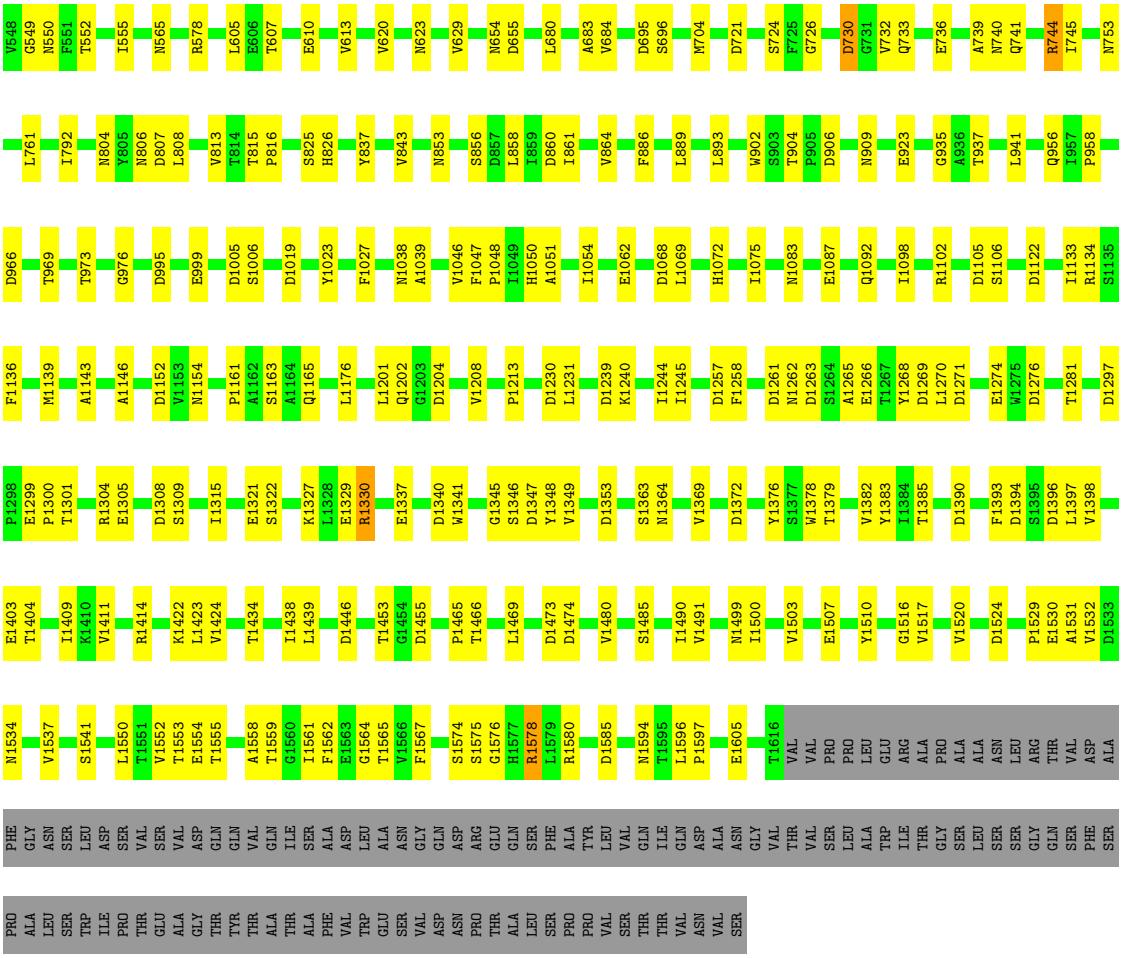
Chain C:



• Molecule 1: Cell surface protein

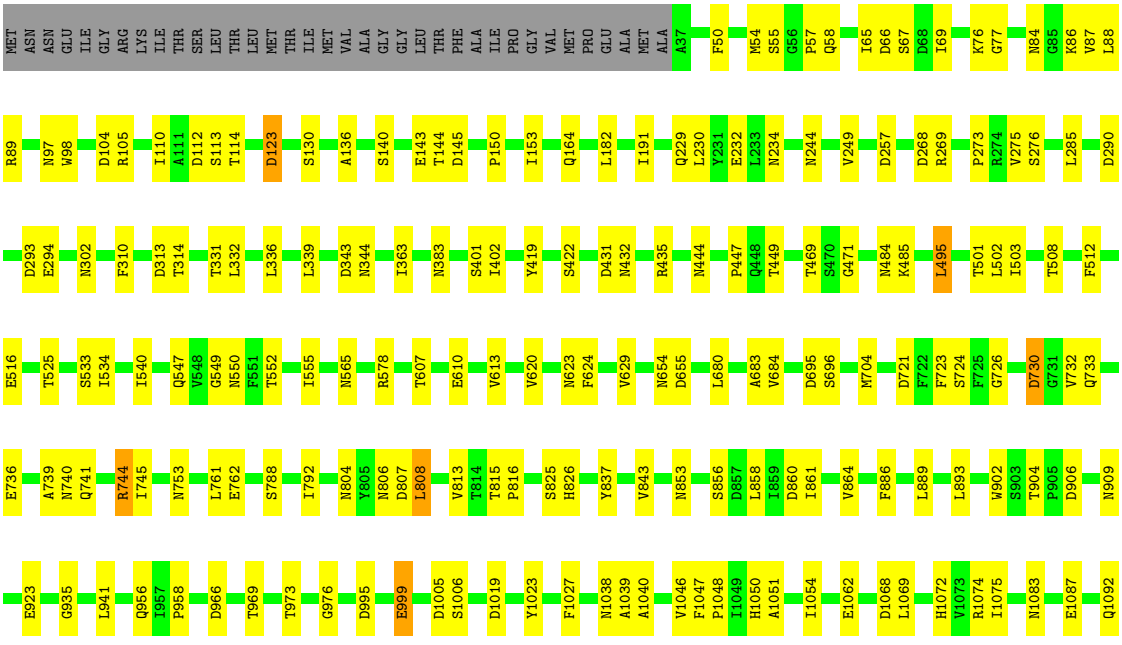
Chain D:

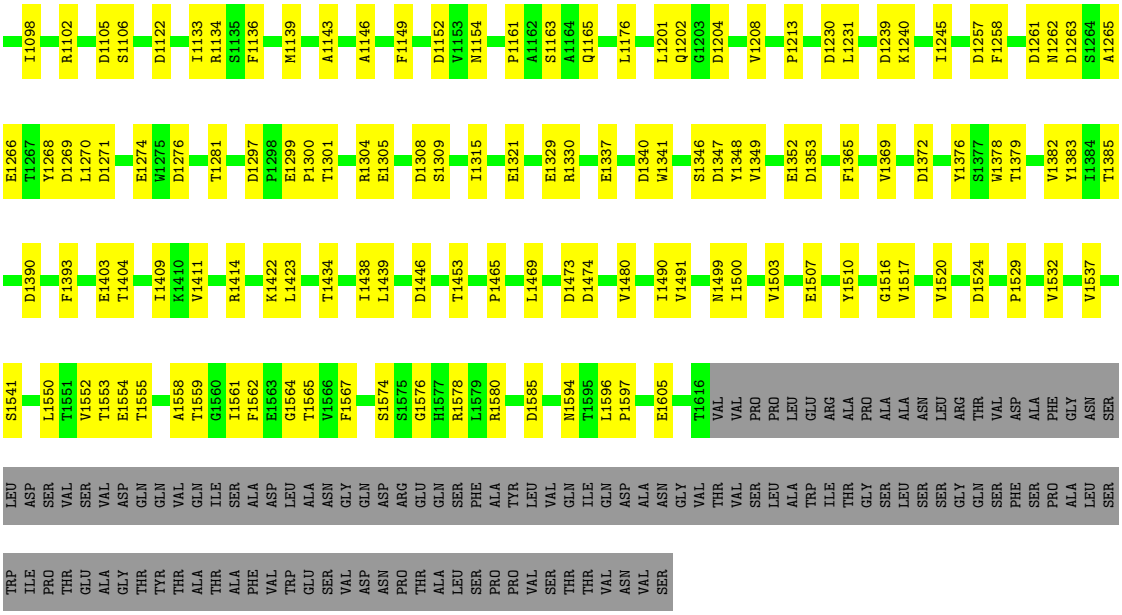




• Molecule 1: Cell surface protein

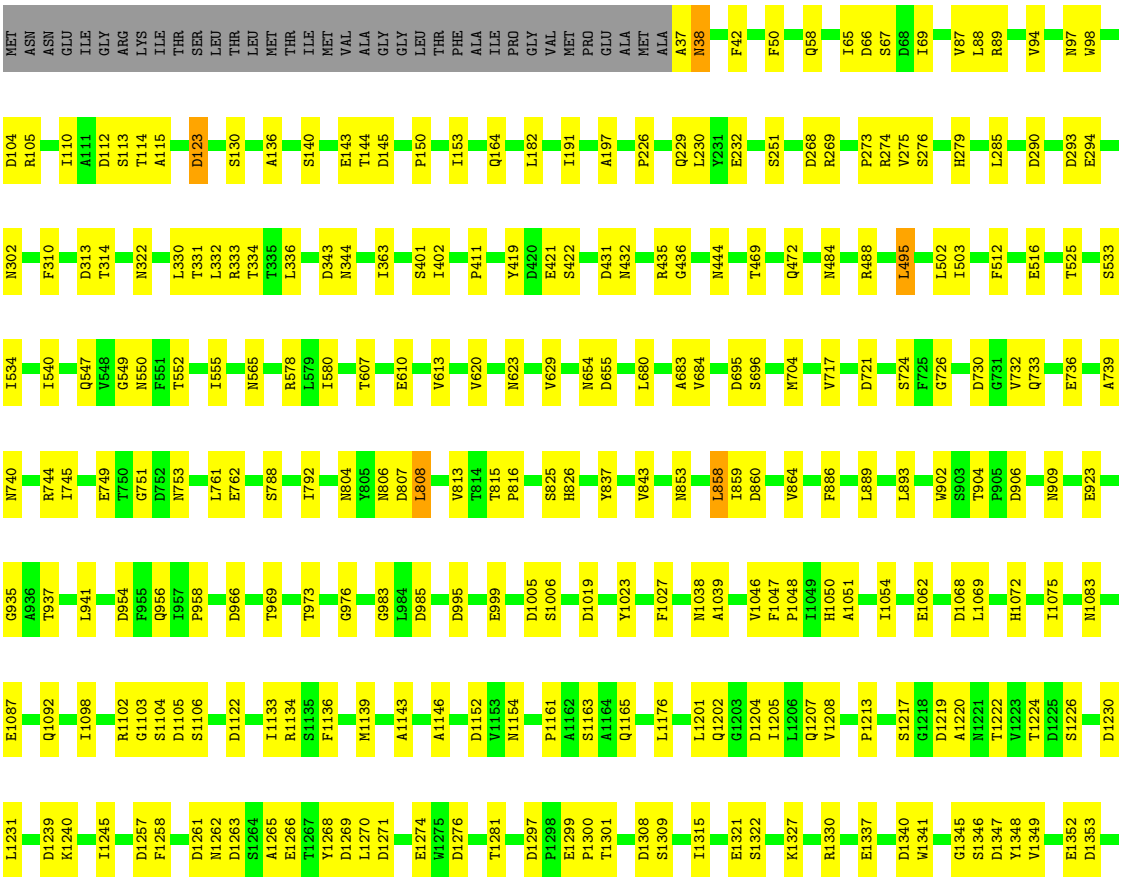
Chain E:

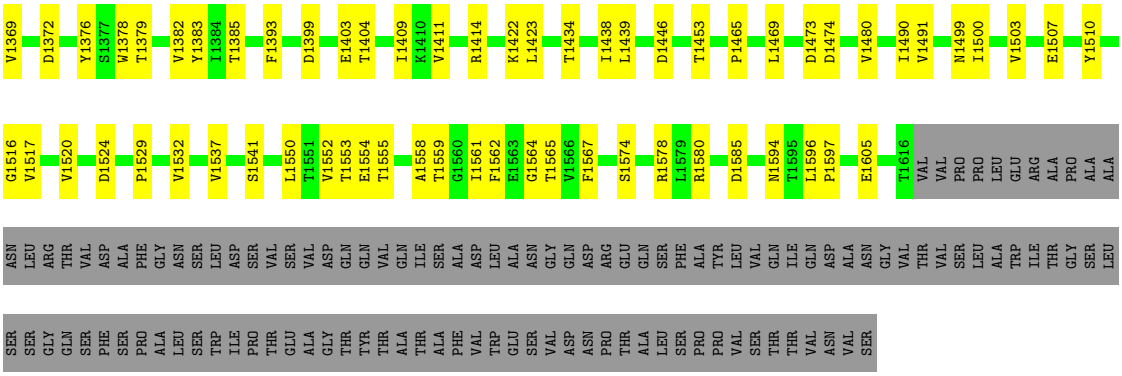




• Molecule 1: Cell surface protein

Chain F:




V1369 D1372 Y1376 S1377 W1378 T1379 V1382 Y1383 I1384 T1385 F1393 D1399 E1403 T1404 I1409 K1410 V1411 R1414 K1422 L1423 T1434 I1438 L1439 D1446 T1453 P1465 L1469 D1473 D1474 V1480 I1490 V1491 N1499 I1500 V1503 E1507 Y1510
G1516 V1517 V1520 D1524 P1529 V1532 V1537 S1541 L1550 T1551 V1552 T1553 E1554 T1555 A1558 T1559 G1560 I1561 F1562 E1563 G1564 T1565 V1566 F1567 S1574 R1578 L1579 R1580 D1585 N1594 T1595 L1596 P1597 E1605 T1616 VAL VAL PRO PRO LEU GLU ARG ALA PRO ALA ALA
ASN LEU ARG THR VAL ASP ALA PHE GLY ASN SER LEU ASP SER VAL SER VAL ASP GLN GLN VAL GLN ILE SER ALA ASP LEU ALA ASN GLY GLN ASP ARG GLU GLN SER PHE ALA TYR LEU VAL VAL GLN ILE GLN ASP ALA ASN GLY VAL THR VAL SER LEU ALA TRP ILE THR GLY SER LEU
SER SER GLY GLN SER PHE SER PRO ALA LEU SER TRP ILE PRO THR GLU ALA GLY THR TYR THR ALA THR ALA PHE VAL TRP GLU SER VAL ASP ASN PRO THR ALA LEU SER PRO PRO VAL SER THR THR VAL ASN VAL SER

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	354860	Depositor
Resolution determination method	OTHER	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; RELION refinement with in-built CTF correction. The function is similar to a Wiener filter, so amplitude correction included.	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	48.5	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	10.250	Depositor
Minimum map value	-4.662	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.422	Depositor
Recommended contour level	1.05514	Depositor
Map size (Å)	349.44, 349.44, 349.44	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.092, 1.092, 1.092	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	0.75	9/11941 (0.1%)	0.81	18/16339 (0.1%)
1	B	0.69	2/11941 (0.0%)	0.77	8/16339 (0.0%)
1	C	0.69	2/11941 (0.0%)	0.77	6/16339 (0.0%)
1	D	0.68	1/11941 (0.0%)	0.76	7/16339 (0.0%)
1	E	0.69	2/11941 (0.0%)	0.77	7/16339 (0.0%)
1	F	0.70	3/11941 (0.0%)	0.77	5/16339 (0.0%)
All	All	0.70	19/71646 (0.0%)	0.77	51/98034 (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	2
1	B	0	1
All	All	0	3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	793	GLU	CD-OE2	16.03	1.43	1.25
1	A	1255	GLU	CD-OE1	-15.38	1.08	1.25
1	A	793	GLU	CD-OE1	-13.35	1.10	1.25
1	A	196	GLU	CD-OE1	-10.86	1.13	1.25
1	A	1299	GLU	CD-OE1	-9.95	1.14	1.25

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	729	ASP	CB-CG-OD1	15.40	132.16	118.30
1	A	123	ASP	CB-CG-OD1	11.60	128.74	118.30
1	A	794	ASP	CB-CG-OD2	10.58	127.82	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	729	ASP	CB-CG-OD2	-9.00	110.20	118.30
1	E	1257	ASP	CB-CG-OD1	-8.15	110.96	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	196	GLU	Sidechain
1	A	729	ASP	Sidechain
1	B	55	SER	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	11751	0	10951	245	0
1	B	11751	0	10951	221	0
1	C	11751	0	10951	248	0
1	D	11751	0	10951	252	0
1	E	11751	0	10949	221	0
1	F	11751	0	10951	248	0
All	All	70506	0	65704	1252	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 9.

The worst 5 of 1252 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:C:1217:SER:CB	1:D:1424:VAL:HG21	1.65	1.25
1:A:1424:VAL:HG21	1:F:1217:SER:CB	1.65	1.24
1:C:816:PRO:HD2	1:D:1083:ASN:HD21	1.01	1.14
1:A:1424:VAL:CG2	1:F:1217:SER:HB2	1.82	1.09
1:C:1217:SER:HB2	1:D:1424:VAL:CG2	1.83	1.08

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	1578/1734 (91%)	1517 (96%)	61 (4%)	0	100	100
1	B	1578/1734 (91%)	1518 (96%)	60 (4%)	0	100	100
1	C	1578/1734 (91%)	1520 (96%)	58 (4%)	0	100	100
1	D	1578/1734 (91%)	1518 (96%)	60 (4%)	0	100	100
1	E	1578/1734 (91%)	1518 (96%)	60 (4%)	0	100	100
1	F	1578/1734 (91%)	1520 (96%)	58 (4%)	0	100	100
All	All	9468/10404 (91%)	9111 (96%)	357 (4%)	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	A	1312/1438 (91%)	1300 (99%)	12 (1%)	78	92
1	B	1312/1438 (91%)	1301 (99%)	11 (1%)	81	93
1	C	1312/1438 (91%)	1302 (99%)	10 (1%)	81	93
1	D	1312/1438 (91%)	1300 (99%)	12 (1%)	78	92
1	E	1312/1438 (91%)	1300 (99%)	12 (1%)	78	92
1	F	1312/1438 (91%)	1300 (99%)	12 (1%)	78	92
All	All	7872/8628 (91%)	7803 (99%)	69 (1%)	79	92

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	38	ASN
1	F	269	ARG
1	F	1047	PHE
1	C	274	ARG
1	C	269	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 75 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	740	ASN
1	F	956	GLN
1	E	853	ASN
1	F	97	ASN
1	B	1154	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

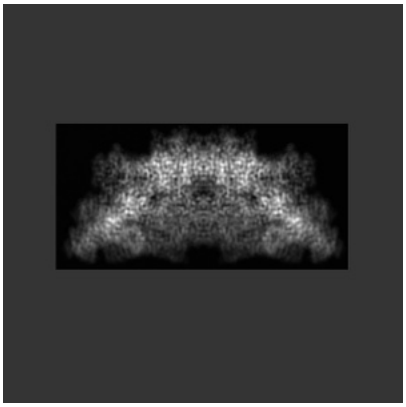
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-16484. These allow visual inspection of the internal detail of the map and identification of artifacts.

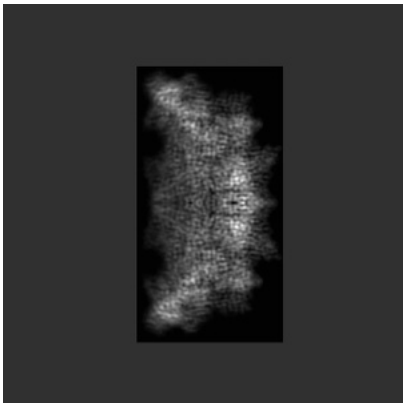
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

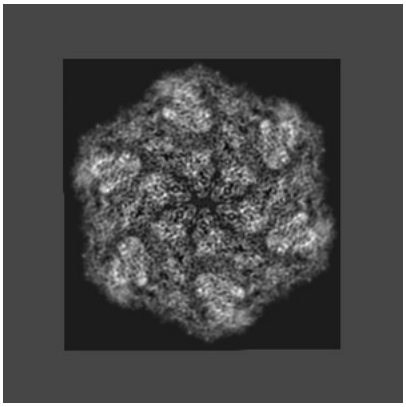
6.1.1 Primary map



X



Y



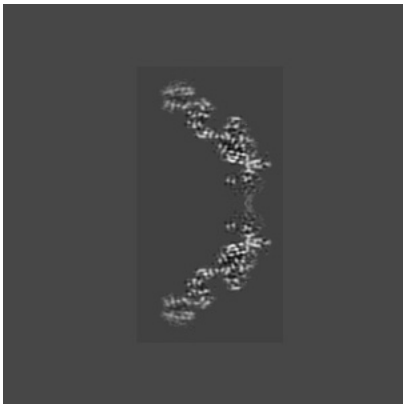
Z

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

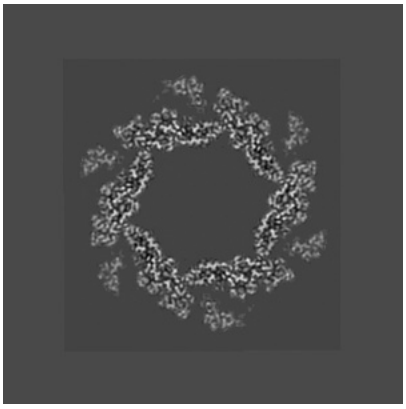
6.2 Central slices

6.2.1 Primary map

X Index: 160



Y Index: 160



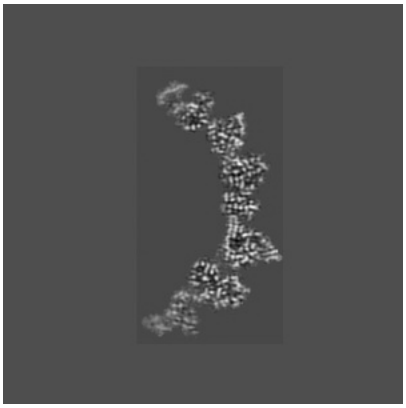
Z Index: 160

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: 156



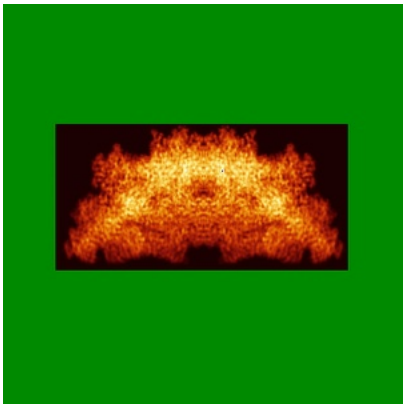
Y Index: 175

Z Index: 186

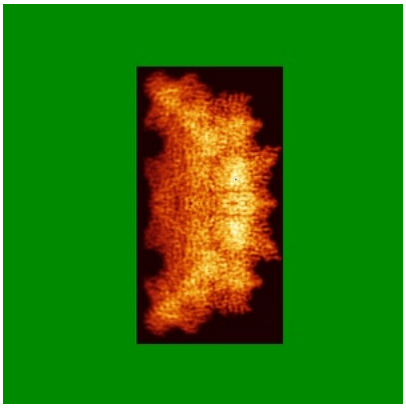
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

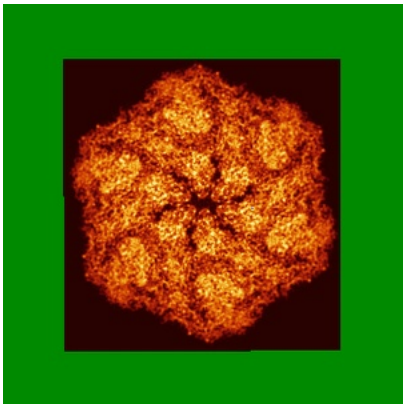
6.4.1 Primary map



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 1.05514. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

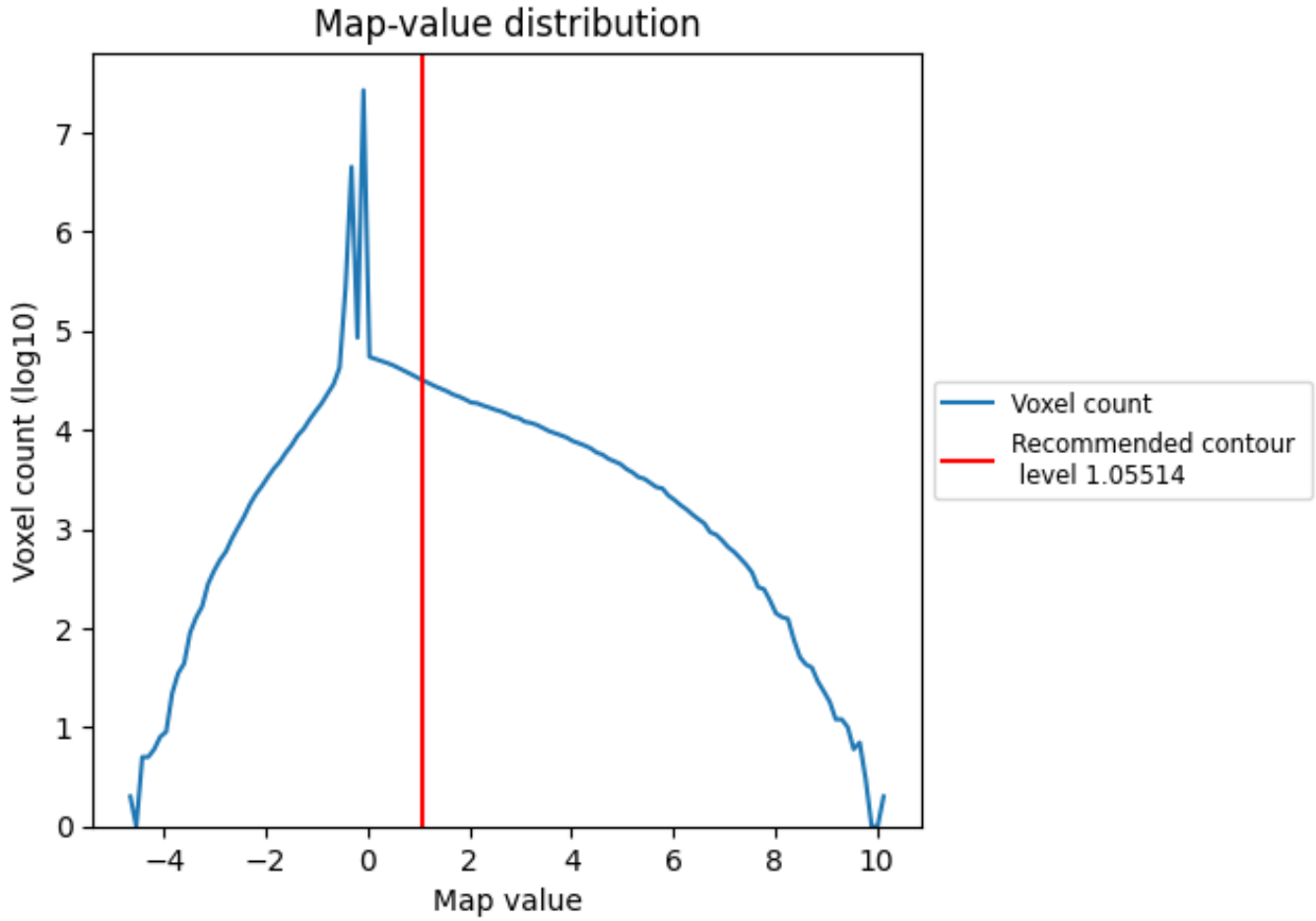
6.6 Mask visualisation (i)

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

7 Map analysis

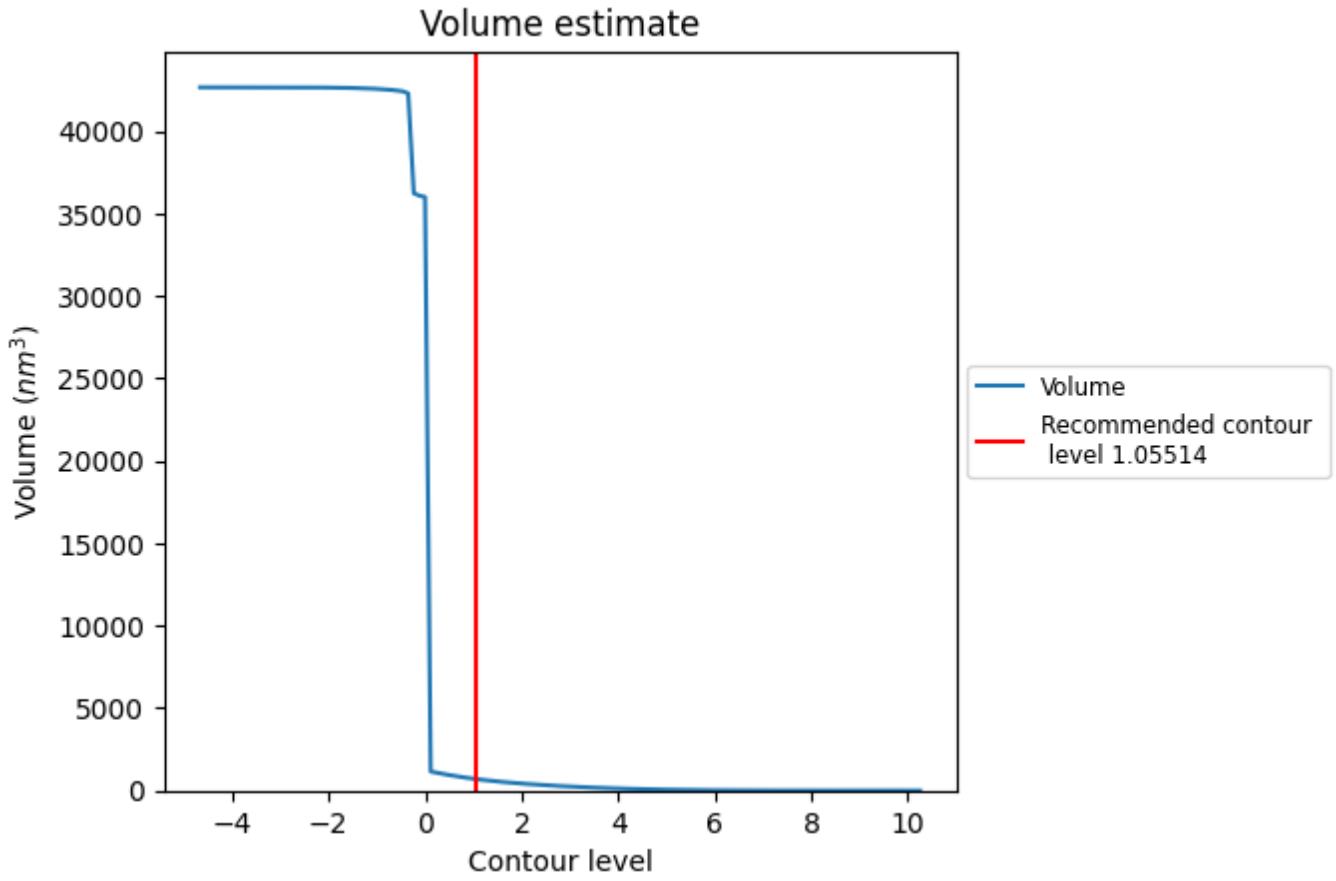
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

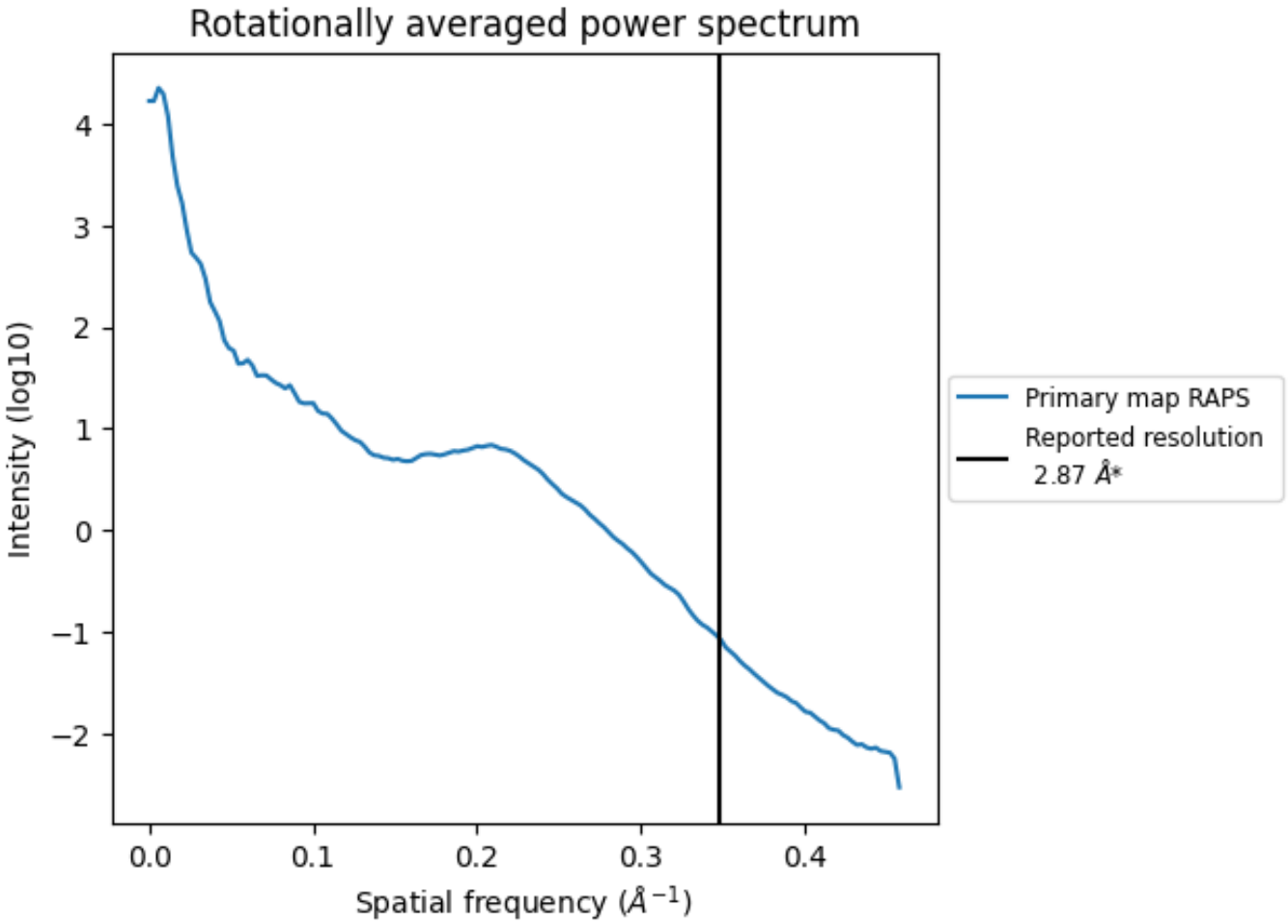
7.2 Volume estimate (i)



The volume at the recommended contour level is 697 nm^3; this corresponds to an approximate mass of 630 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 0.348 Å^{-1}

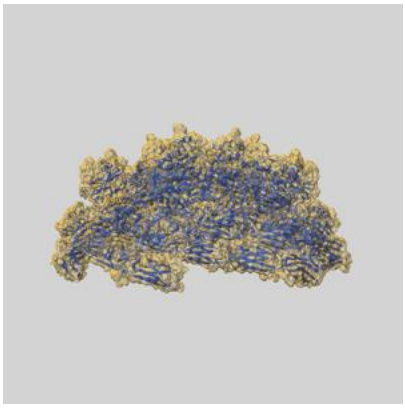
8 Fourier-Shell correlation

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

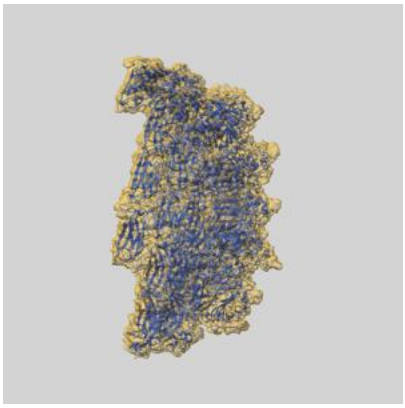
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-16484 and PDB model 8C8M. Per-residue inclusion information can be found in section 3 on page 4.

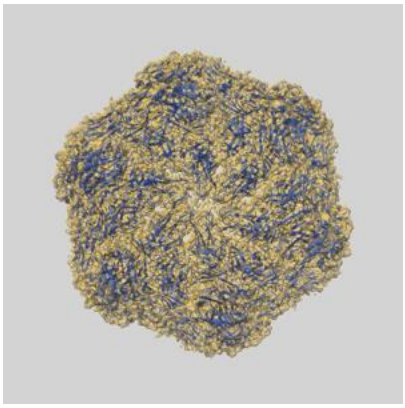
9.1 Map-model overlay



X



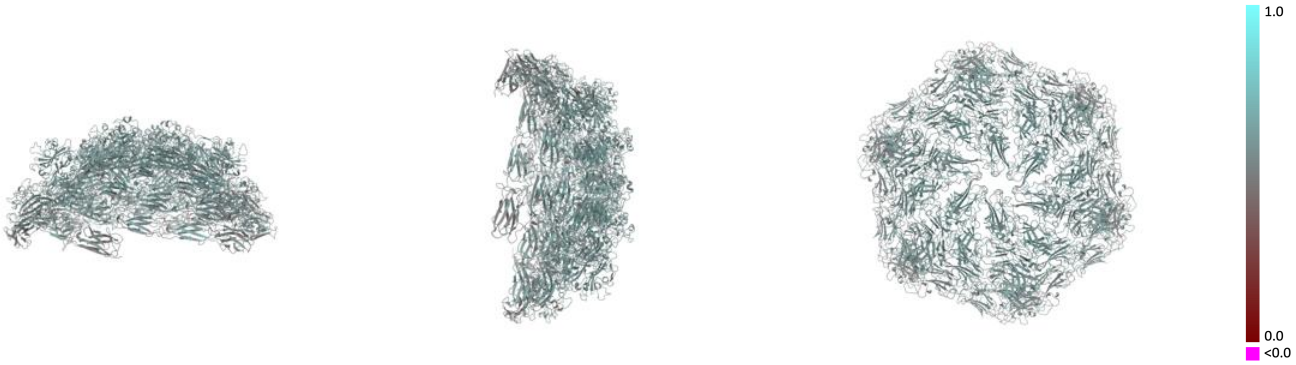
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 1.05514 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.

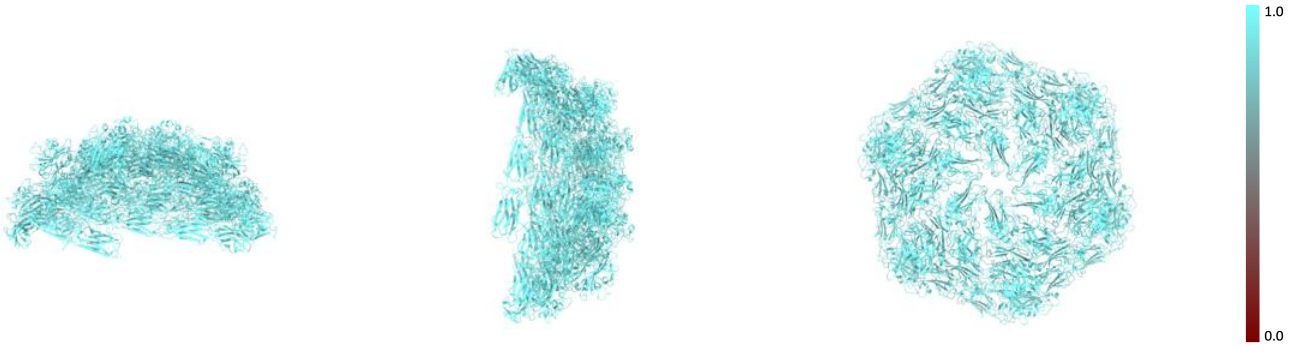
9.2 Q-score mapped to coordinate model



X Y Z

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.

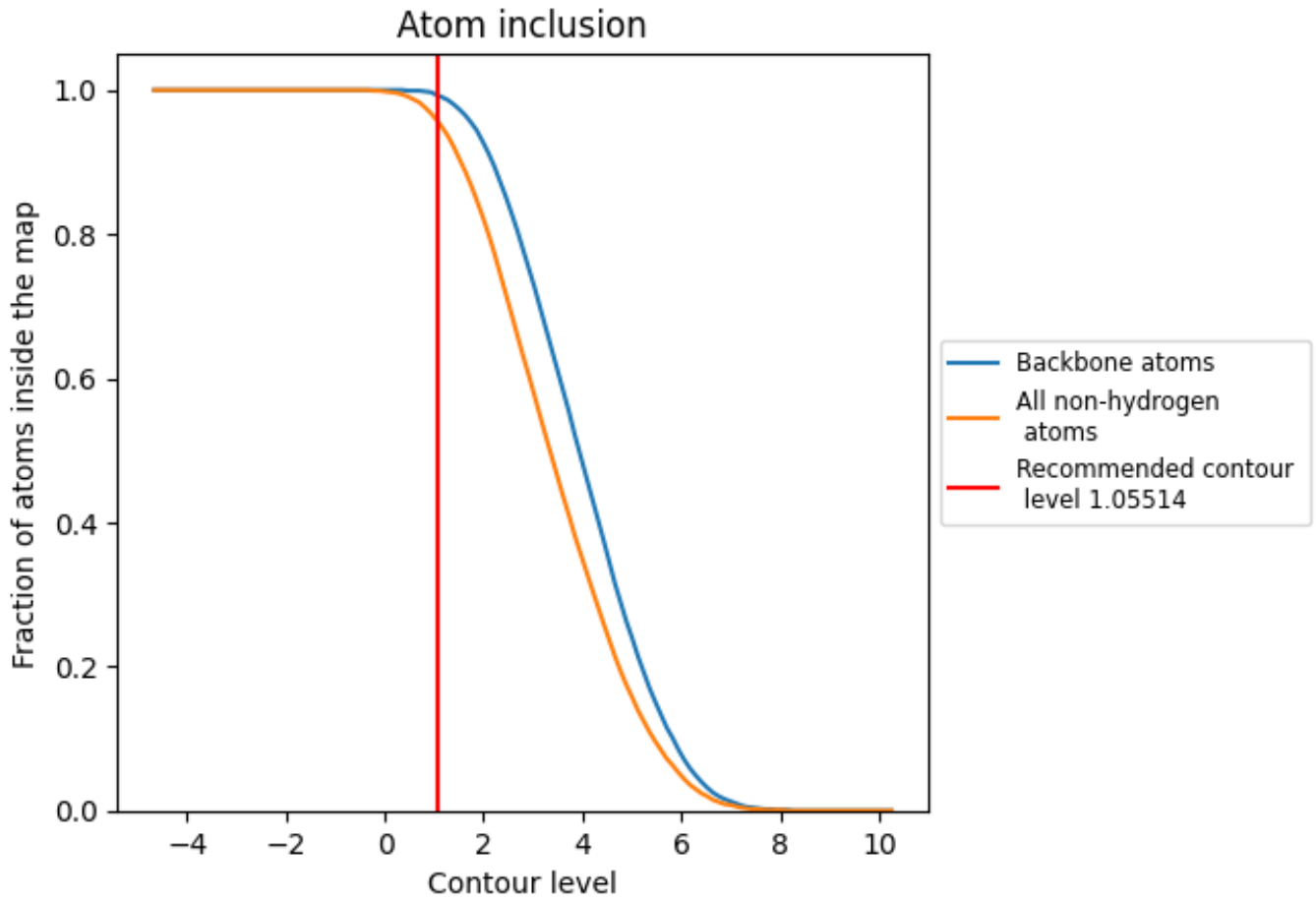
9.3 Atom inclusion mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (1.05514).

9.4 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 96% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (1.05514) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9580	0.5560
A	0.9620	0.5630
B	0.9580	0.5530
C	0.9530	0.5520
D	0.9630	0.5620
E	0.9580	0.5530
F	0.9540	0.5520