wwPDB EM Validation Summary Report (i)

May 19, 2024 – 10:46 AM EDT

PDB ID	:	7MI8
EMDB ID	:	EMD-23842
Title	:	Signal subtracted reconstruction of AAA5 and AAA6 domains of dynein in the presence of a pyrazolo-pyrimidinone-based compound, Model 5
Authors	:	Santarossa, C.C.; Coudray, N.; Urnavicius, L.; Ekiert, D.C.; Bhabha, G.; Kapoor, T.M.
Deposited on	:	2021-04-16
Resolution	:	3.70 Å(reported)
Based on initial model	:	7MI1

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:

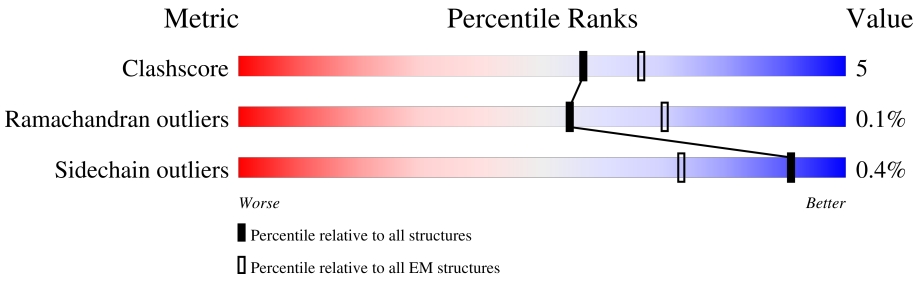
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.70 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	2661	

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 7034 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 Fusion protein of Dynein and Endolysin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	868	7034	4542	1146	1316	30	0	0

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1363	GLY	-	expression tag	UNP P36022
A	1849	GLN	GLU	engineered mutation	UNP P36022
A	3120	GLY	-	linker	UNP P36022
A	3121	SER	-	linker	UNP P36022
A	3122	GLY	-	linker	UNP P36022
A	3123	SER	-	linker	UNP P36022
A	3124	GLY	-	linker	UNP P36022
A	3125	SER	-	linker	UNP P36022
A	3136	GLY	ARG	conflict	UNP P00720
A	3178	THR	CYS	conflict	UNP P00720
A	3221	ALA	CYS	conflict	UNP P00720
A	3261	ARG	ILE	conflict	UNP P00720
A	3286	GLY	-	linker	UNP P00720
A	3287	SER	-	linker	UNP P00720
A	3288	GLY	-	linker	UNP P00720
A	3289	SER	-	linker	UNP P00720
A	3290	GLY	-	linker	UNP P00720
A	3291	SER	-	linker	UNP P00720
A	3742	ASP	ASN	conflict	UNP P36022
A	3895	VAL	PHE	conflict	UNP P36022
A	4072	ASP	ASN	conflict	UNP P36022
A	4093	GLY	-	expression tag	UNP P36022
A	4094	SER	-	expression tag	UNP P36022
A	4095	GLY	-	expression tag	UNP P36022
A	4096	SER	-	expression tag	UNP P36022
A	4097	GLY	-	expression tag	UNP P36022
A	4098	SER	-	expression tag	UNP P36022
A	4099	HIS	-	expression tag	UNP P36022

Continued on next page...

Continued from previous page...

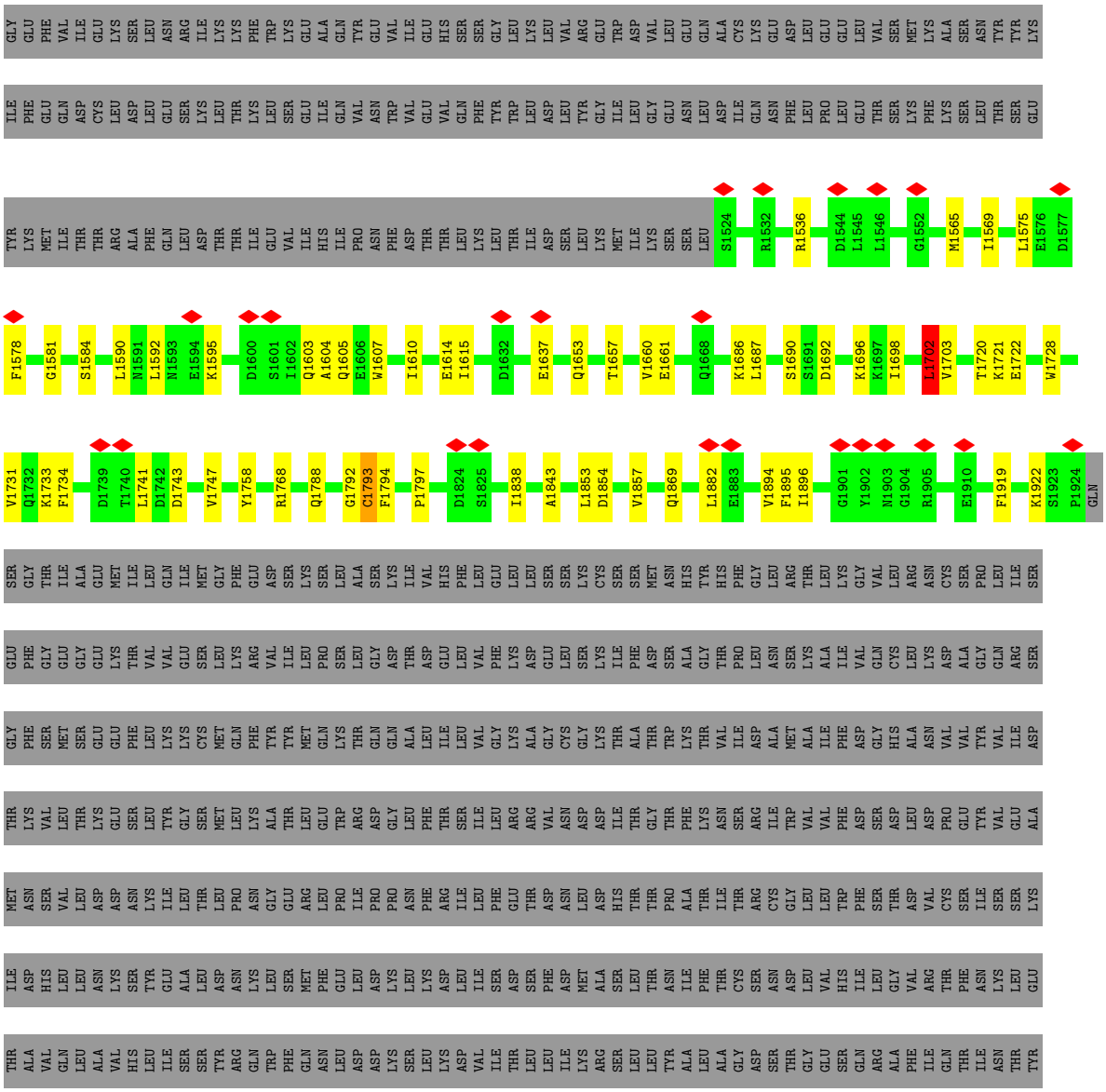
Chain	Residue	Modelled	Actual	Comment	Reference
A	4100	HIS	-	expression tag	UNP P36022
A	4101	HIS	-	expression tag	UNP P36022
A	4102	HIS	-	expression tag	UNP P36022
A	4103	HIS	-	expression tag	UNP P36022
A	4104	HIS	-	expression tag	UNP P36022

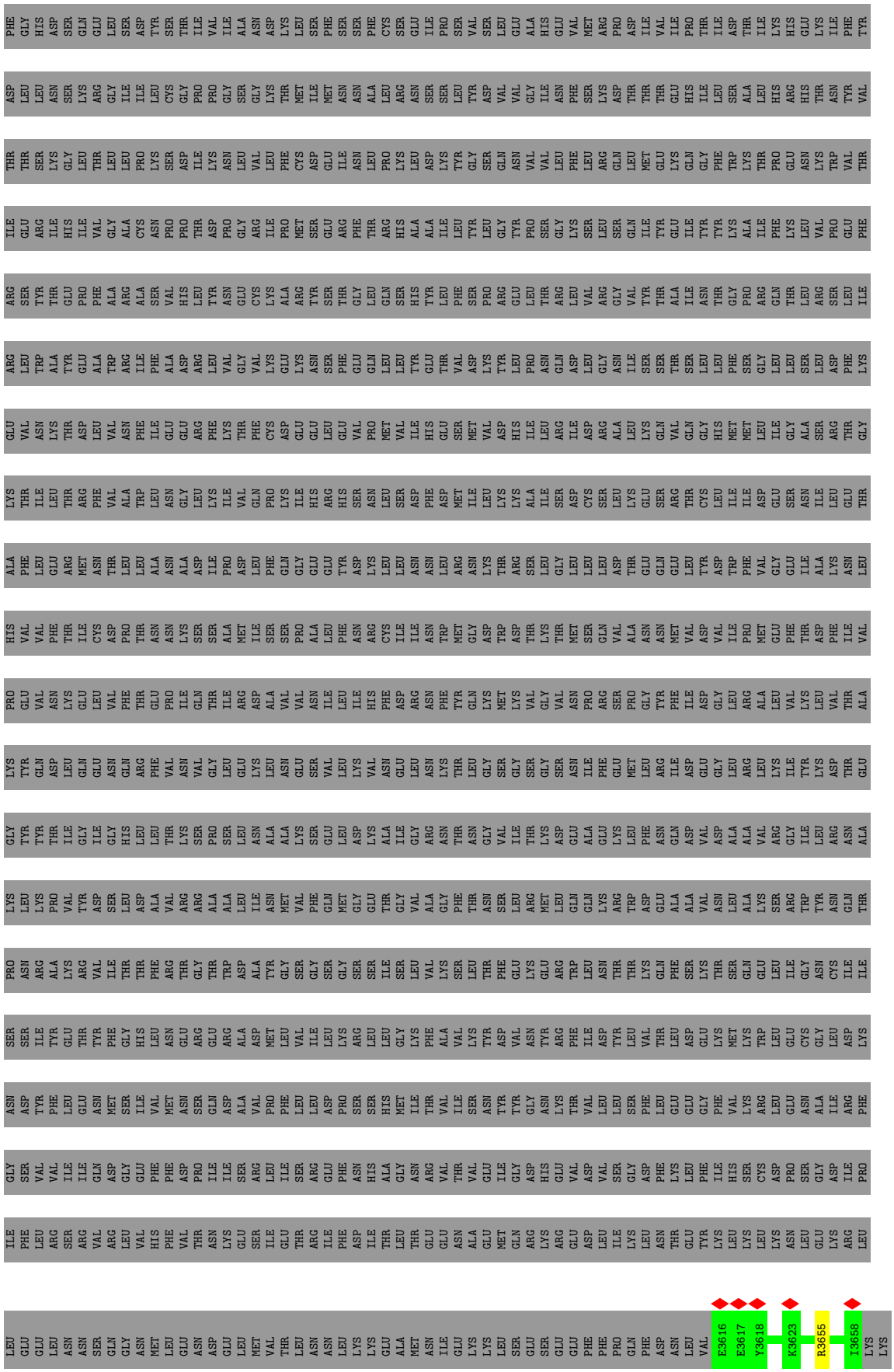
3 Residue-property plots (i)

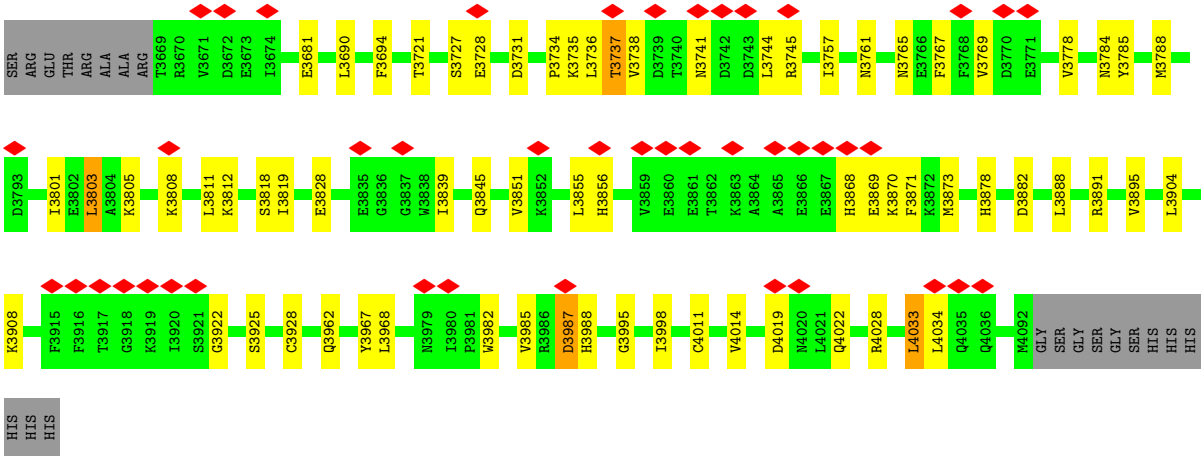
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

Chain A:




E3616
E3617
Y3618
K3623
R3655
I3658


SER
ARG
GLU
THR
ARG
ALA
ALA
ARG
T3669
R3670
V3671
D3672
E3673
I3674
E3681
L3690
F3694
T3721
S3727
E3728
D3731
P3734
K3735
L3736
T3737
V3738
D3739
T3740
N3741
D3742
D3743
L3744
R3745
I3757
N3761
N3765
E3766
F3767
F3768
V3769
D3770
E3771
V3778
N3784
Y3785
M3788
D3793
I3801
E3802
L3803
A3804
K3805
K3808
L3811
K3812
S3818
I3819
E3828
E3835
G3836
G3837
W3838
I3839
Q3845
V3851
K3852
L3855
H3856
V3859
E3860
E3861
T3862
K3863
A3864
A3865
E3866
E3867
H3868
E3869
K3870
F3871
K3872
M3873
H3878
D3882
L3888
R3891
V3895
L3904
K3908
F3915
F3916
T3917
G3918
K3919
I3920
S3921
G3922
S3925
C3928
Q3962
Y3967
L3968
N3979
I3980
P3981
W3982
V3985
R3986
D3987
H3988
G3995
I3998
C4011
V4014
D4019
N4020
L4021
Q4022
R4028
L4033
L4034
Q4035
Q4036
M4092
GLY
SER
GLY
SER
GLY
SER
HIS
HIS
HIS
HIS
HIS
HIS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	128004	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 ($e^-/Å^2$)	44	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	3700	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.025	Depositor
Minimum map value	-0.015	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.006	Depositor
Map size (Å)	190.62401, 190.62401, 190.62401	wwPDB
Map dimensions	368, 368, 368	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.518, 0.518, 0.518	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	0.34	0/7187	0.58	6/9712 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	4019	ASP	CB-CG-OD2	8.23	125.71	118.30
1	A	1882	LEU	CA-CB-CG	7.54	132.64	115.30
1	A	1741	LEU	CA-CB-CG	7.05	131.52	115.30
1	A	4033	LEU	CA-CB-CG	6.67	130.64	115.30
1	A	3987	ASP	CB-CG-OD1	5.76	123.48	118.30

There are no chirality outliers.

There are no planarity outliers.

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	7034	0	7008	71	0
All	All	7034	0	7008	71	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 5.

The worst 5 of 71 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:3788:MET:HA	1:A:3895:VAL:O	1.89	0.72
1:A:3761:ASN:O	1:A:3765:ASN:ND2	2.28	0.66
1:A:3928:CYS:HB3	1:A:3968:LEU:HD13	1.79	0.63
1:A:1794:PHE:HB3	1:A:1919:PHE:HB3	1.82	0.61
1:A:1692:ASP:O	1:A:1696:LYS:NZ	2.34	0.61

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	862/2661 (32%)	840 (97%)	21 (2%)	1 (0%)	51 83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3737	THR

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	782/2406 (32%)	779 (100%)	3 (0%)	91 95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1702	LEU
1	A	1793	CYS
1	A	3803	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3962	GLN
1	A	3948	HIS
1	A	3780	ASN
1	A	1899	ASN
1	A	3784	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.

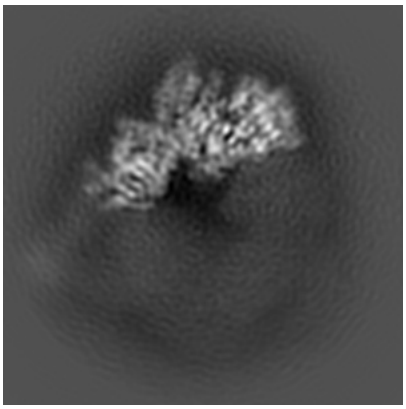
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-23842. These allow visual inspection of the internal detail of the map and identification of artifacts.

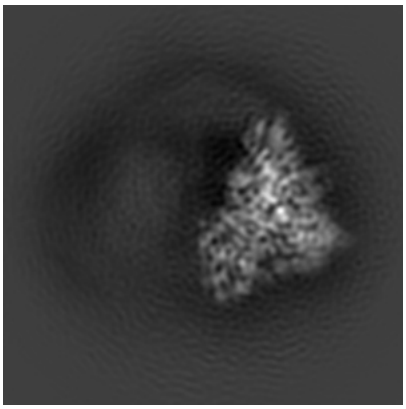
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

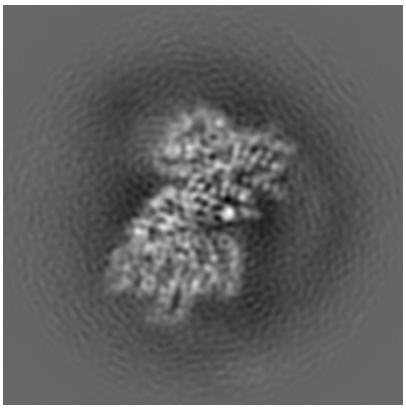
6.1.1 Primary map



X

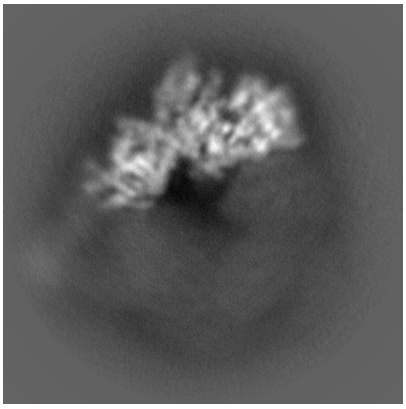


Y

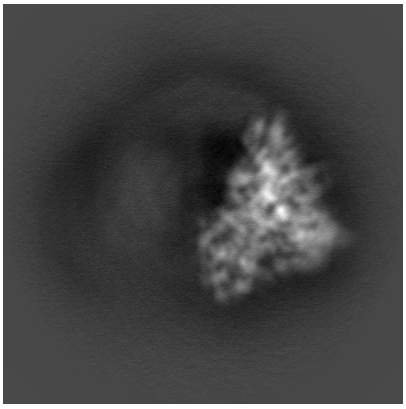


Z

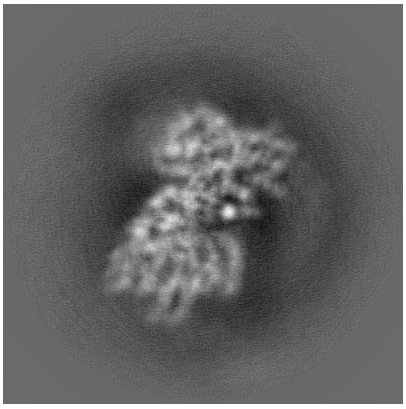
6.1.2 Raw map



X



Y

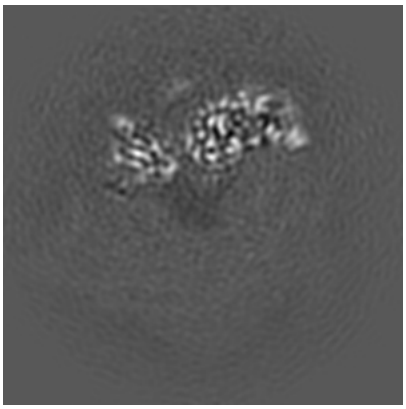


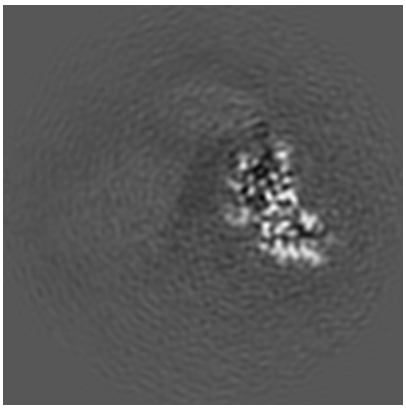
Z

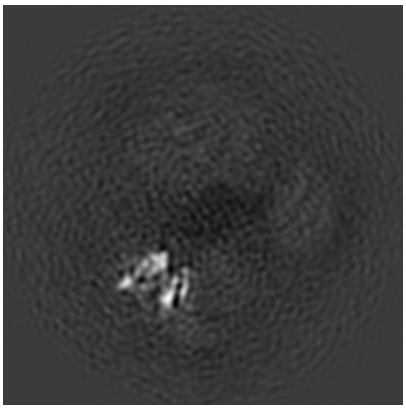
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

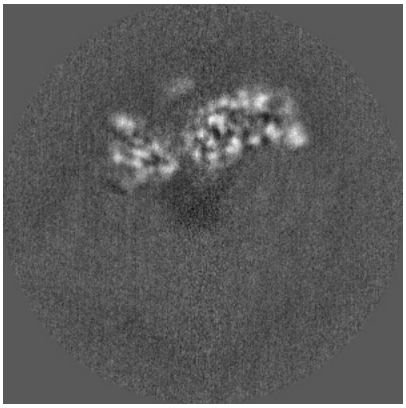
6.2.1 Primary map

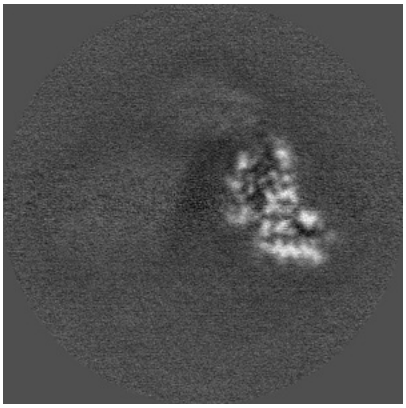

X Index: 184

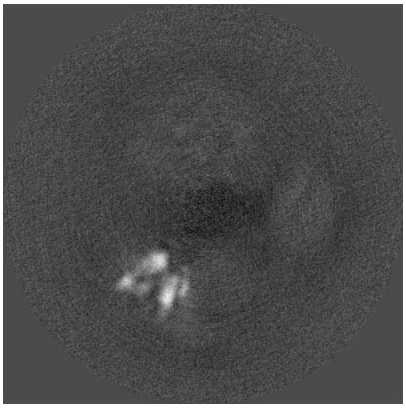

Y Index: 184


Z Index: 184

6.2.2 Raw map


X Index: 184

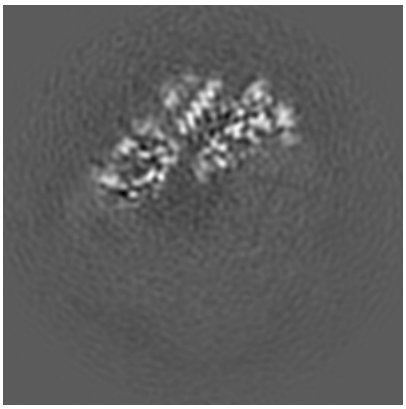

Y Index: 184

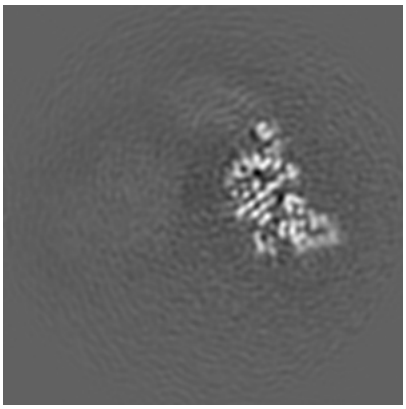

Z Index: 184

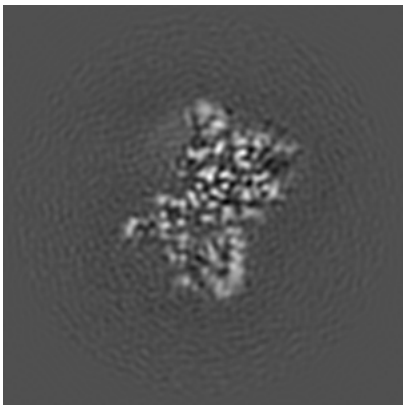
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

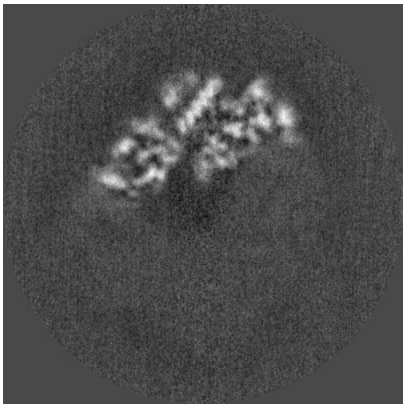
6.3.1 Primary map

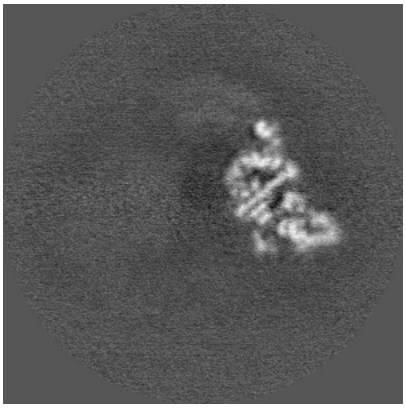

X Index: 170

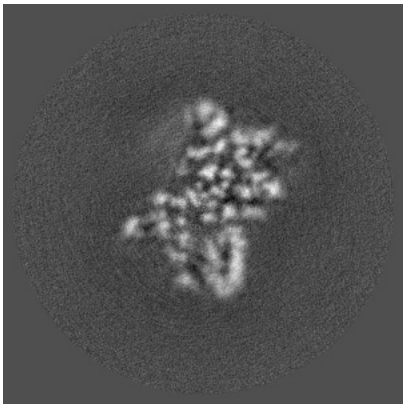

Y Index: 193


Z Index: 243

6.3.2 Raw map


X Index: 171

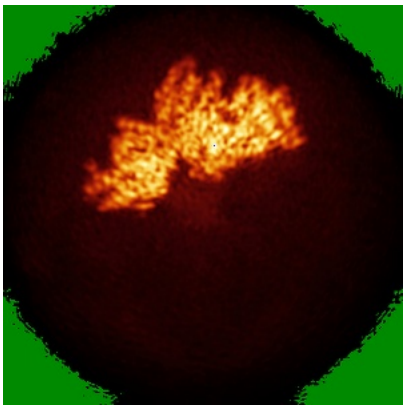

Y Index: 194

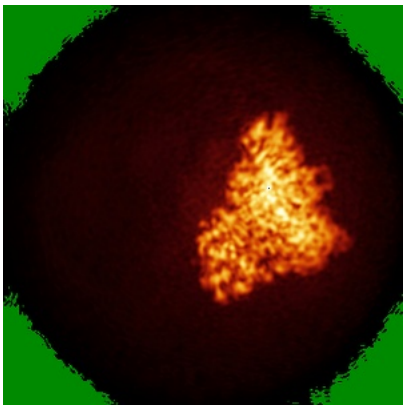

Z Index: 243

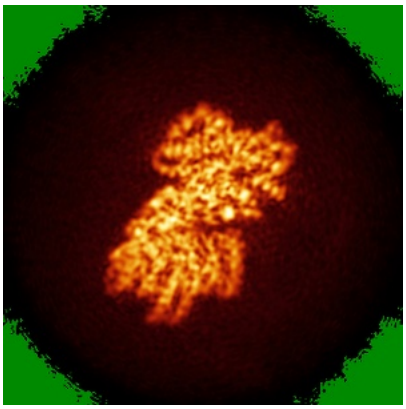
The images above show the largest variance slices of the map in three orthogonal directions.

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

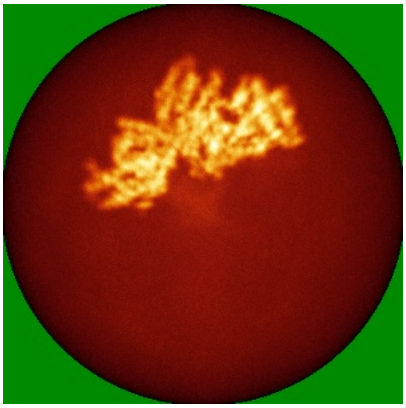
6.4.1 Primary map

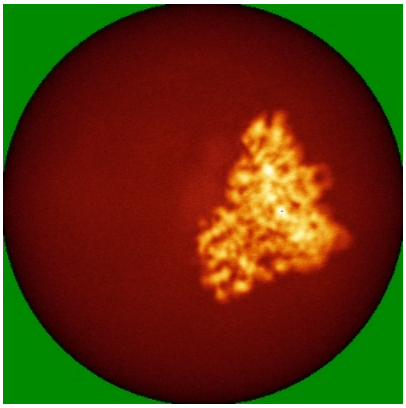

X

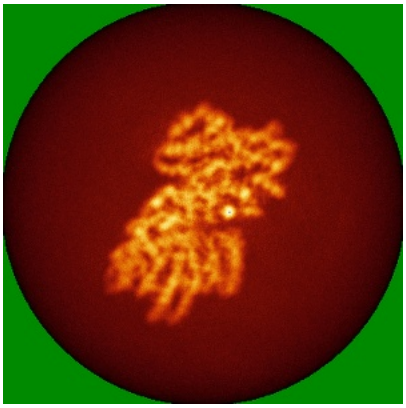

Y


Z

6.4.2 Raw map


X

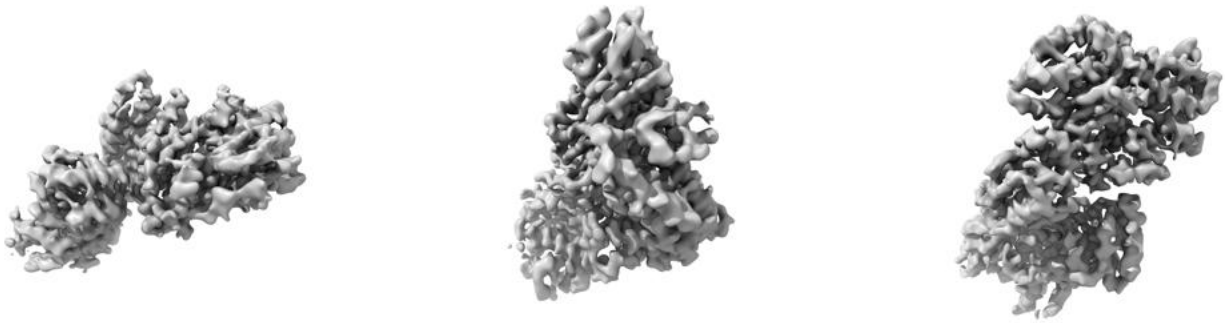

Y


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)

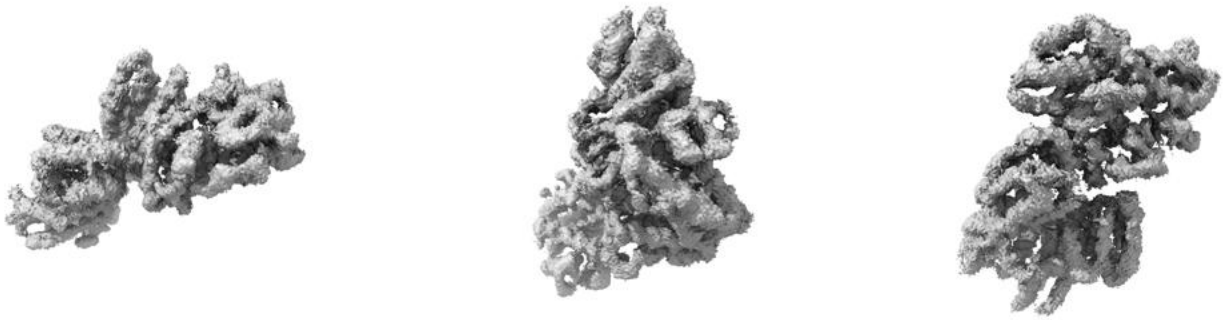
6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.006. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map



X Y Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

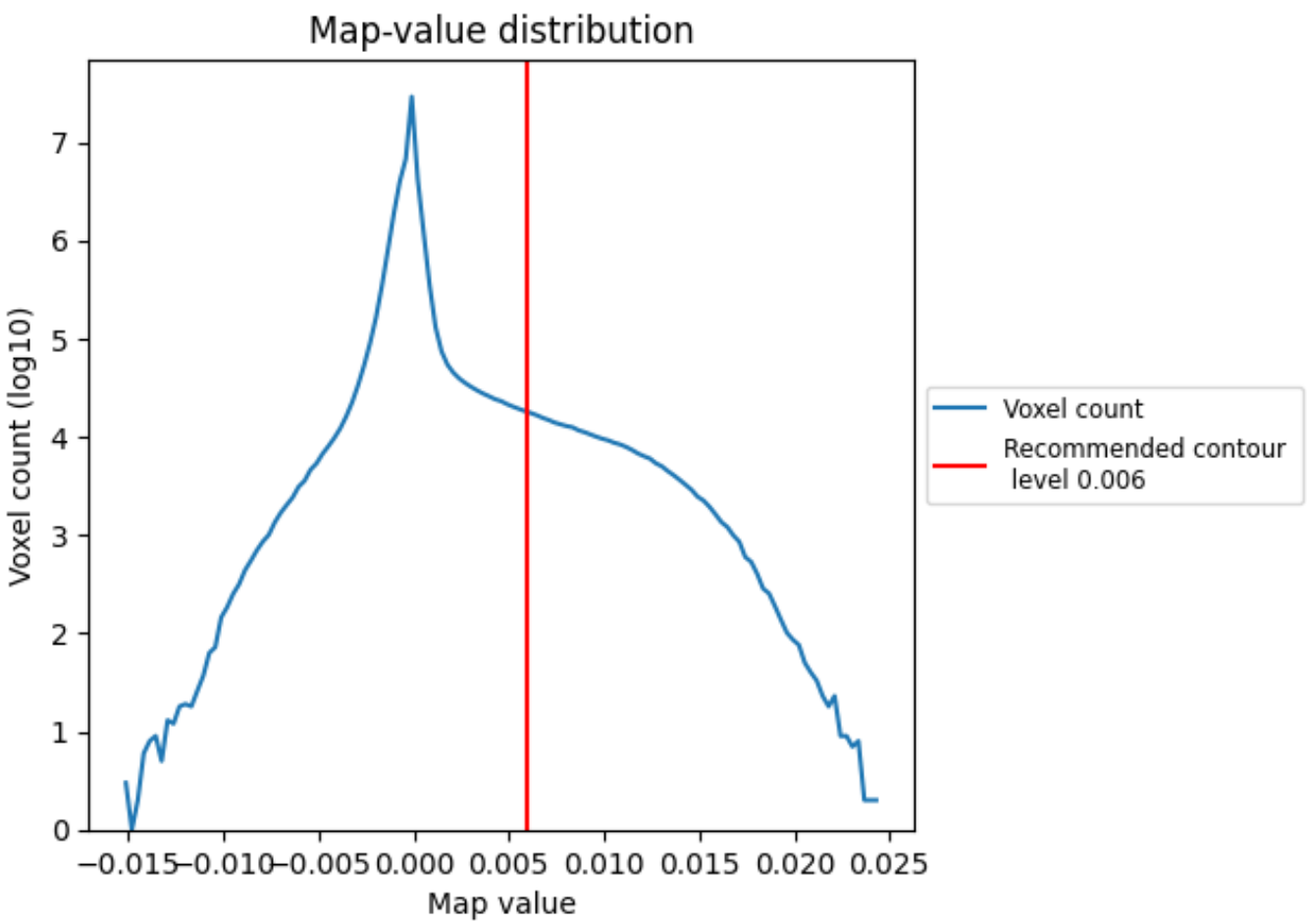
6.6 Mask visualisation (i)

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

7 Map analysis (i)

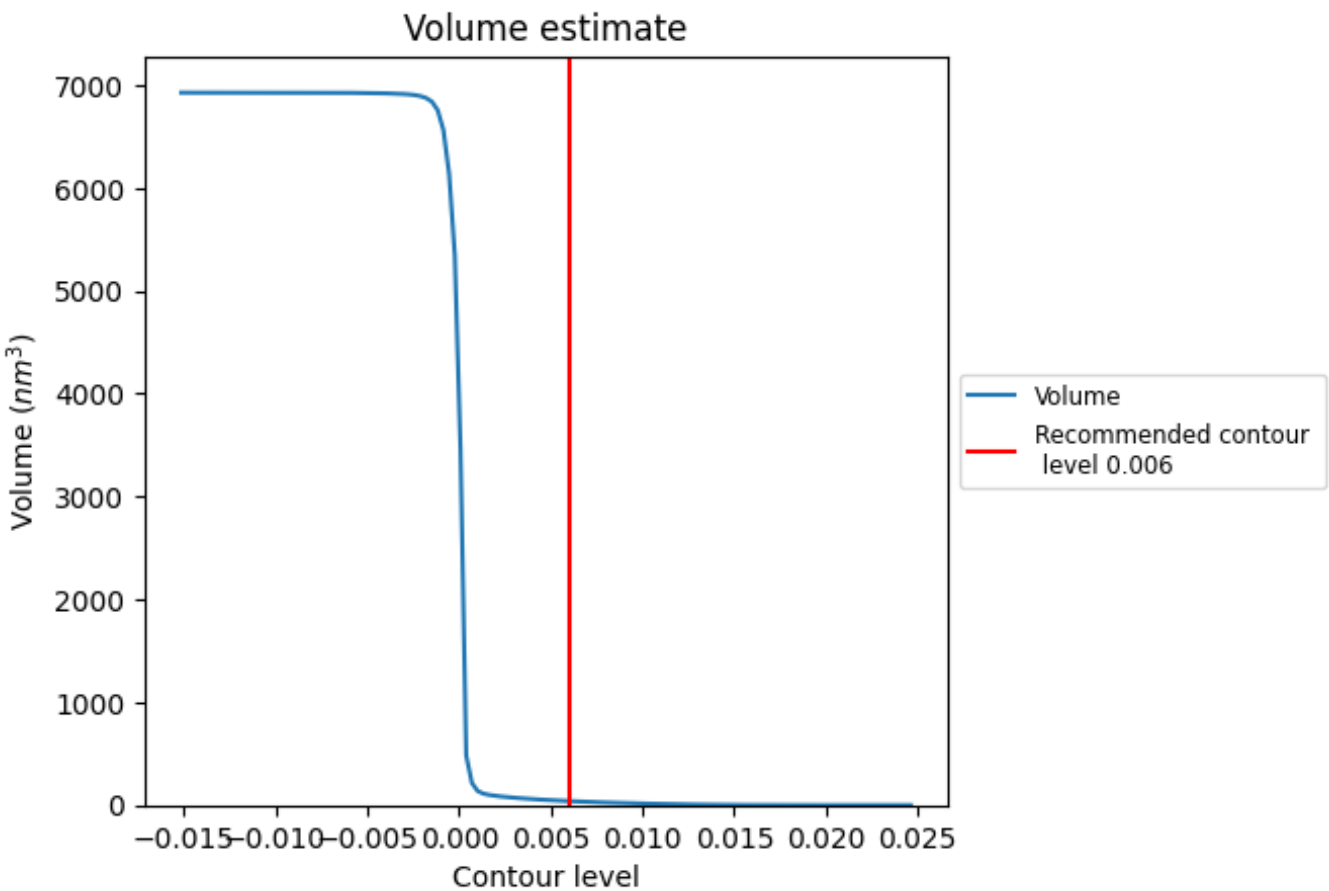
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



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

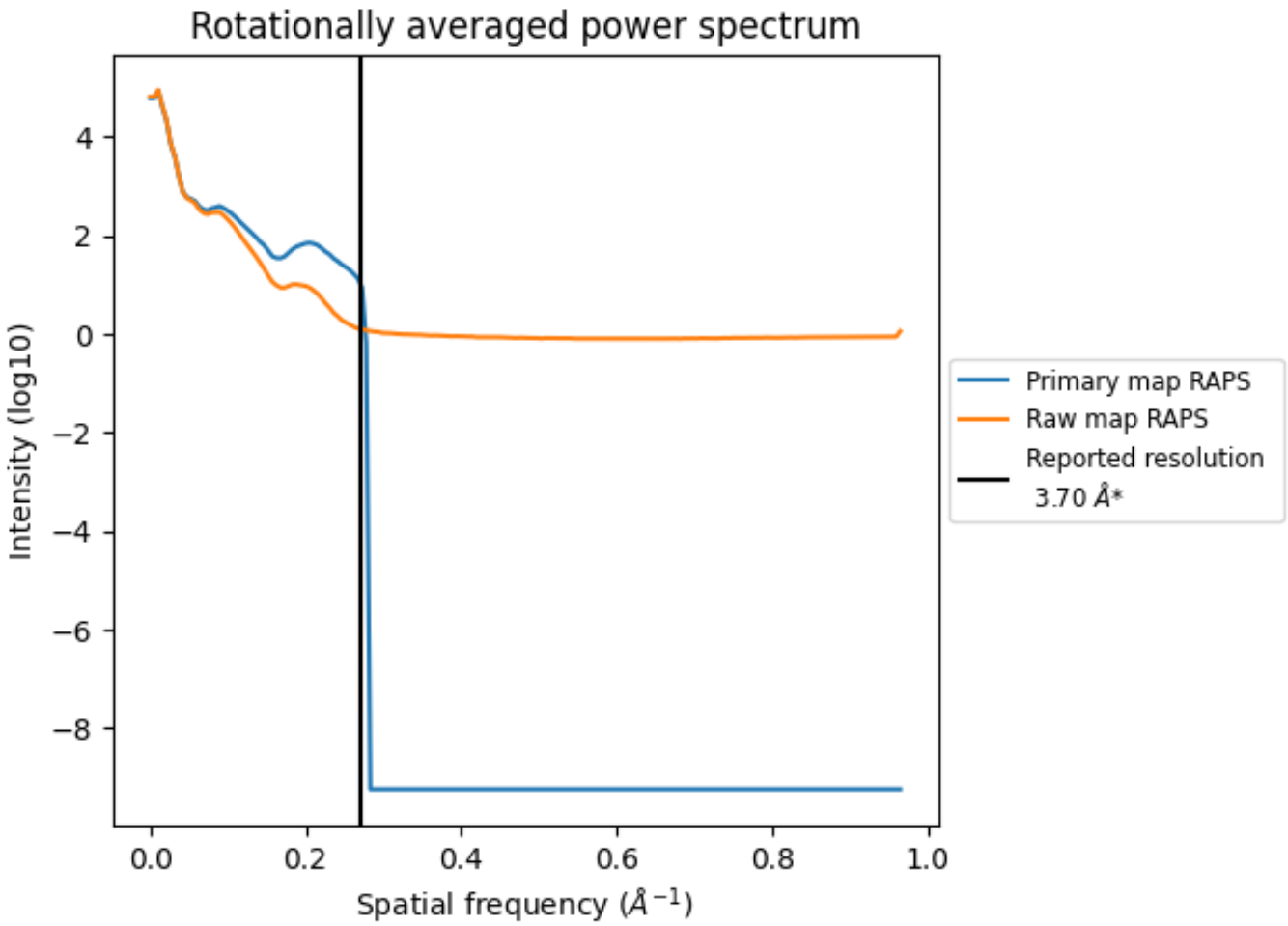
7.2 Volume estimate (i)



The volume at the recommended contour level is 39 nm^3; this corresponds to an approximate mass of 36 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)

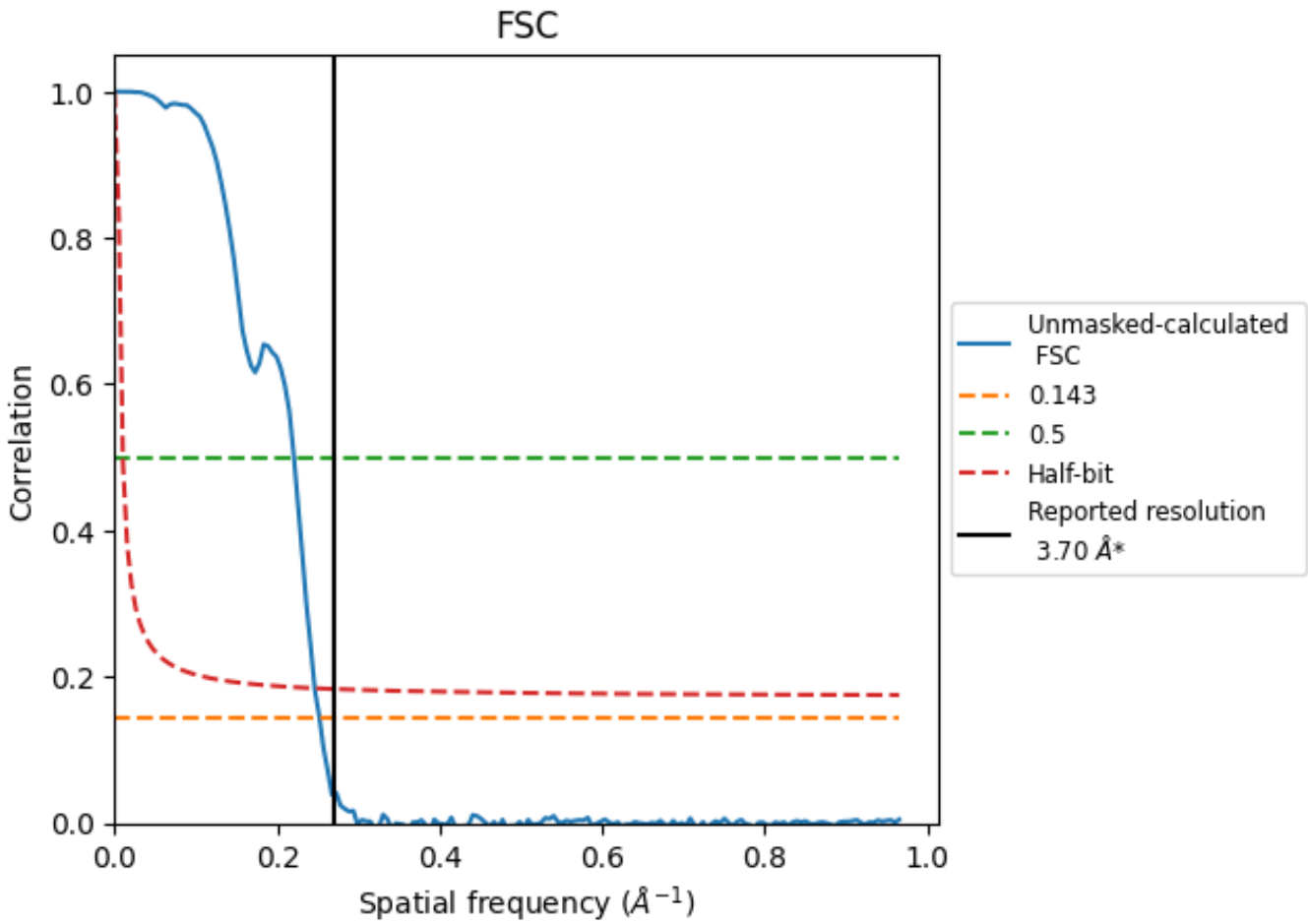


*Reported resolution corresponds to spatial frequency of 0.270 Å^{-1}

8 Fourier-Shell correlation (i)

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

8.1 FSC (i)



*Reported resolution corresponds to spatial frequency of 0.270 Å^{-1}

8.2 Resolution estimates (i)

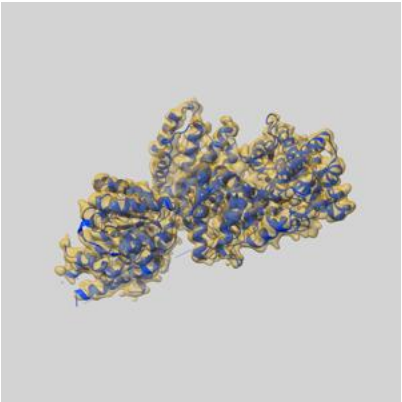
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.70	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.97	4.53	4.06

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

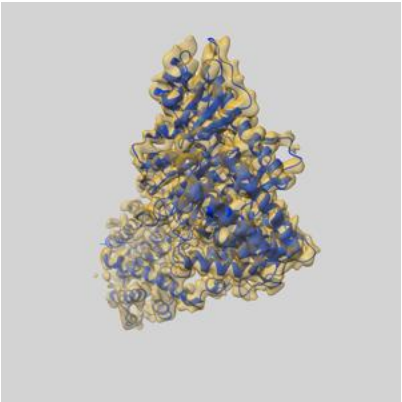
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-23842 and PDB model 7MI8. Per-residue inclusion information can be found in section 3 on page 5.

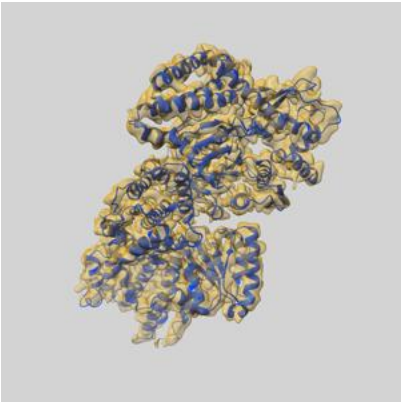
9.1 Map-model overlay (i)



X



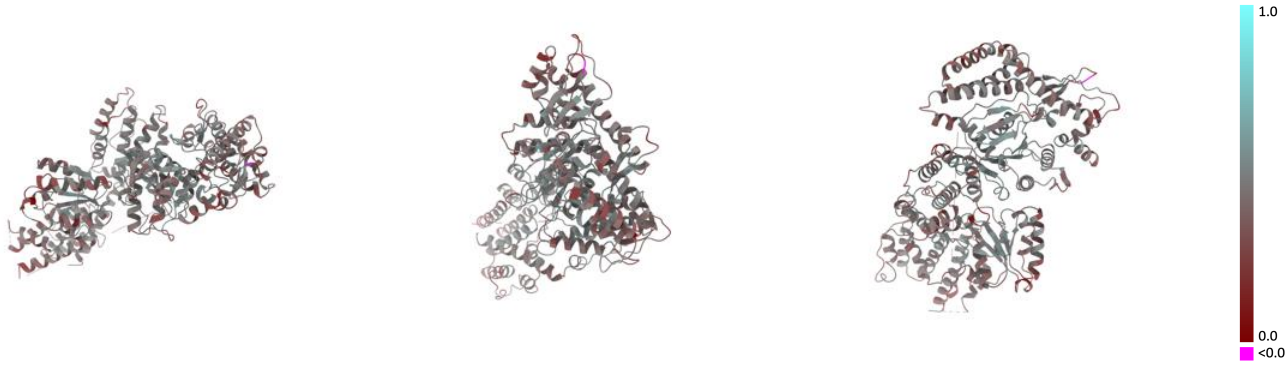
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.006 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.

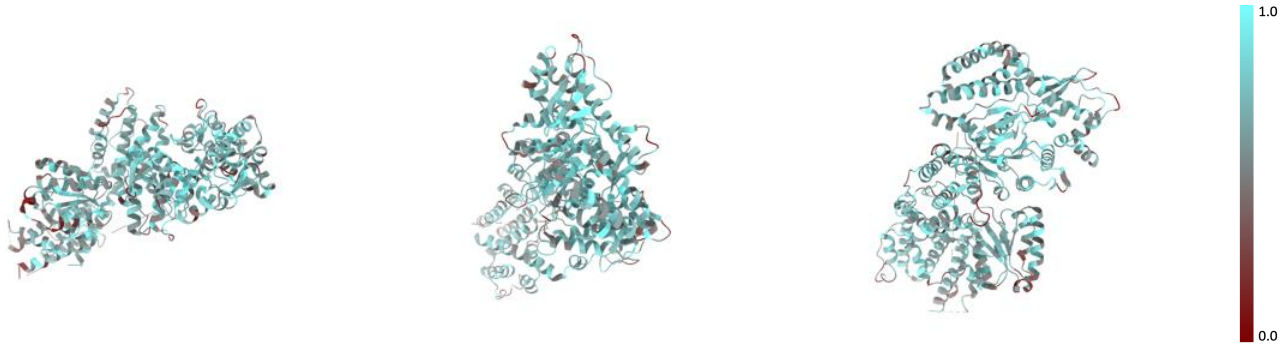
9.2 Q-score mapped to coordinate model (i)



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.

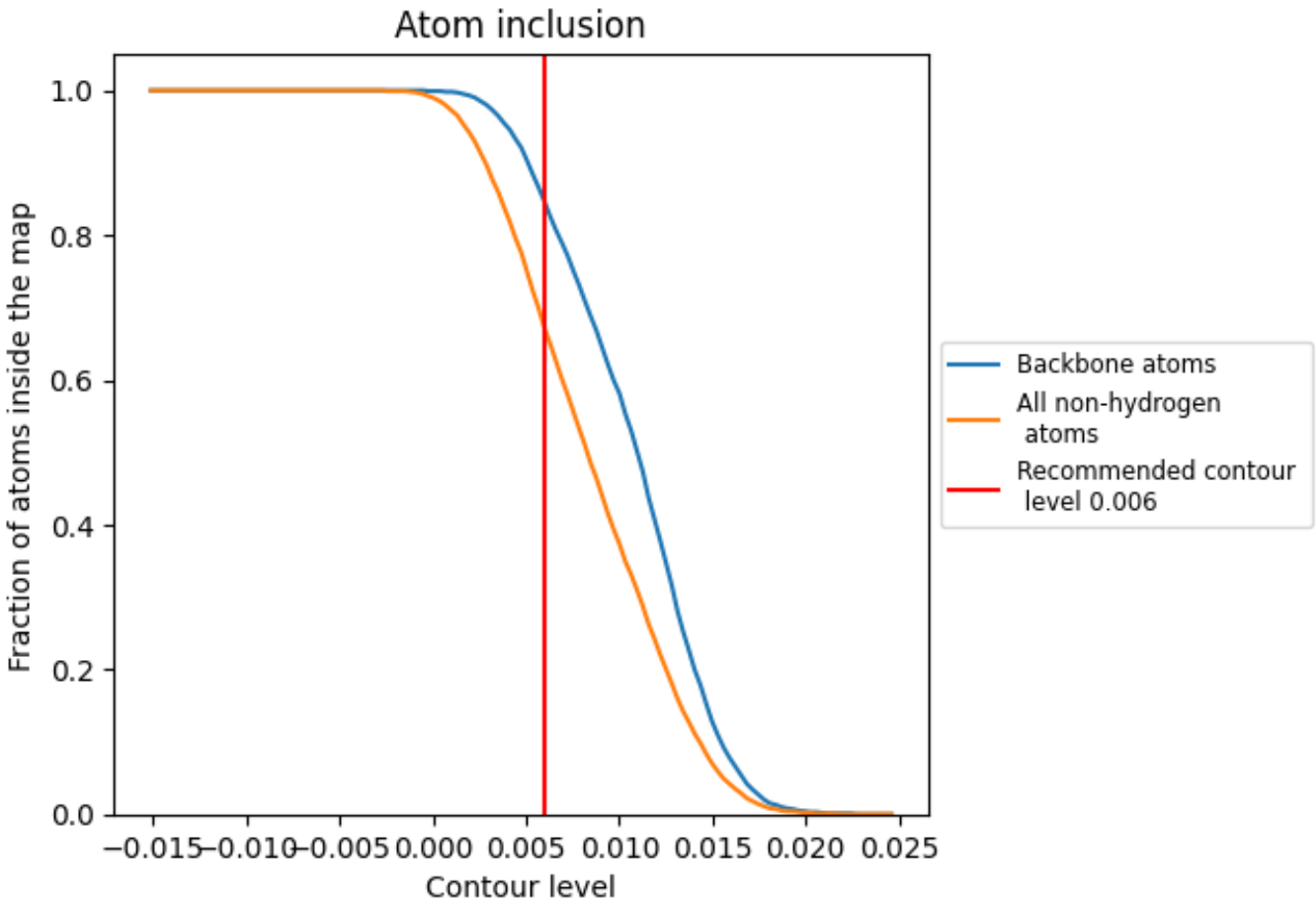
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.006).

9.4 Atom inclusion (i)



At the recommended contour level, 85% of all backbone atoms, 67% 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 (0.006) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6710	0.4270
A	0.6710	0.4270