wwPDB EM Validation Summary Report

Apr 18, 2023 – 10:33 AM EDT

PDB ID	:	8FCK
EMDB ID	:	EMD-28981
Title	:	Structure of the vertebrate augmin complex
Authors	:	Travis, S.M.; Huang, W.; Zhang, R.; Petry, S.
Deposited on	:	2022-12-01
Resolution	:	6.88 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:

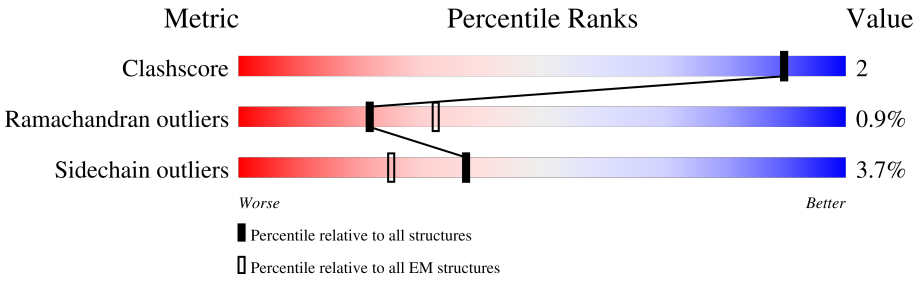
EMDB validation analysis	:	0.0.1.dev50
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.2

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 6.88 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	286	41% 91% 7%
2	B	597	26% 89% 10%
3	C	353	28% 88% 12%
4	D	666	20% 89% 9%
5	E	472	12% 44% 53%
6	F	442	49% 82% 6% 11%
7	G	348	39% 91% 5%
8	H	367	14% 53% 6% 41%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 24717 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 HAUS augmin-like complex subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	286	Total	C	N	O	S	0	0
			2281	1435	378	455	13		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	156	GLN	ARG	conflict	UNP A0A8J1L9M8

- Molecule 2 is a protein called HAUS augmin-like complex subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	597	Total	C	N	O	S	0	0
			4756	2980	811	942	23		

- Molecule 3 is a protein called HAUS augmin like complex subunit 4 L homeolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	353	Total	C	N	O	S	0	0
			2886	1807	508	555	16		

- Molecule 4 is a protein called HAUS augmin-like complex subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	666	Total	C	N	O	S	0	0
			5416	3362	1000	1023	31		

- Molecule 5 is a protein called HAUS augmin like complex subunit 2 L homeolog, Green fluorescent protein chimera.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	222	Total	C	N	O	S	0	0
			1755	1098	302	342	13		

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	223	VAL	-	linker	UNP Q6INL9
E	224	ASP	-	linker	UNP Q6INL9
E	225	MET	-	linker	UNP Q6INL9
E	226	VAL	-	linker	UNP Q6INL9
E	289	LEU	PHE	conflict	UNP P42212
E	290	THR	SER	conflict	UNP P42212
E	446	LYS	LEU	conflict	UNP P42212
E	456	LEU	HIS	conflict	UNP P42212
E	464	VAL	-	expression tag	UNP P42212
E	465	ASP	-	expression tag	UNP P42212
E	466	HIS	-	expression tag	UNP P42212
E	467	HIS	-	expression tag	UNP P42212
E	468	HIS	-	expression tag	UNP P42212
E	469	HIS	-	expression tag	UNP P42212
E	470	HIS	-	expression tag	UNP P42212
E	471	HIS	-	expression tag	UNP P42212
E	472	HIS	-	expression tag	UNP P42212

- Molecule 6 is a protein called HAUS augmin like complex subunit 6 L homeolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	393	Total	C	N	O	S	0	0
			3211	2039	581	571	20		

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-11	GLY	-	expression tag	UNP A0JPI0
F	-10	PRO	-	expression tag	UNP A0JPI0
F	-9	ALA	-	expression tag	UNP A0JPI0
F	-8	SER	-	expression tag	UNP A0JPI0
F	-7	GLY	-	expression tag	UNP A0JPI0
F	-6	SER	-	expression tag	UNP A0JPI0
F	-5	THR	-	expression tag	UNP A0JPI0
F	-4	ARG	-	expression tag	UNP A0JPI0
F	-3	GLY	-	expression tag	UNP A0JPI0
F	-2	ALA	-	expression tag	UNP A0JPI0
F	-1	GLU	-	expression tag	UNP A0JPI0
F	0	PHE	-	expression tag	UNP A0JPI0

- Molecule 7 is a protein called HAUS augmin like complex subunit 7 S homeolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	332	Total	C	N	O	S	0	0
			2627	1655	435	517	20		

- Molecule 8 is a protein called HAUS augmin-like complex subunit 8.

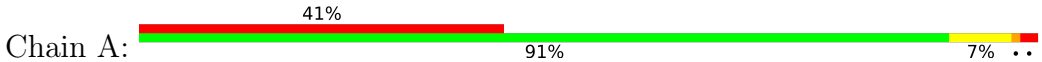
Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	218	Total	C	N	O	S	0	0
			1785	1122	305	351	7		

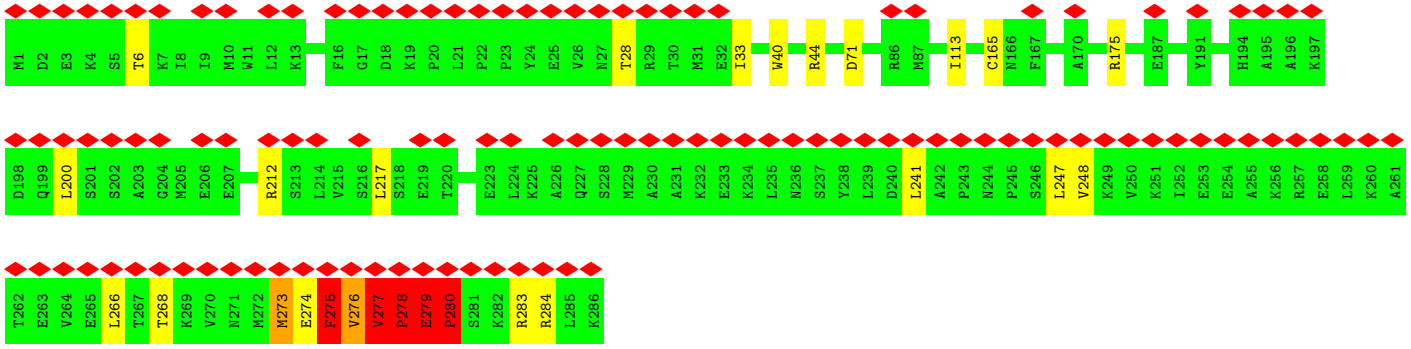
3 Residue-property plots (i)

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: HAUS augmin-like complex subunit 1

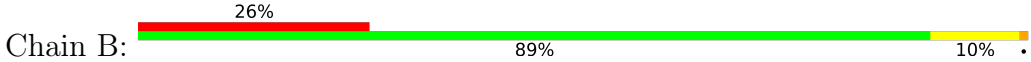
Chain A:

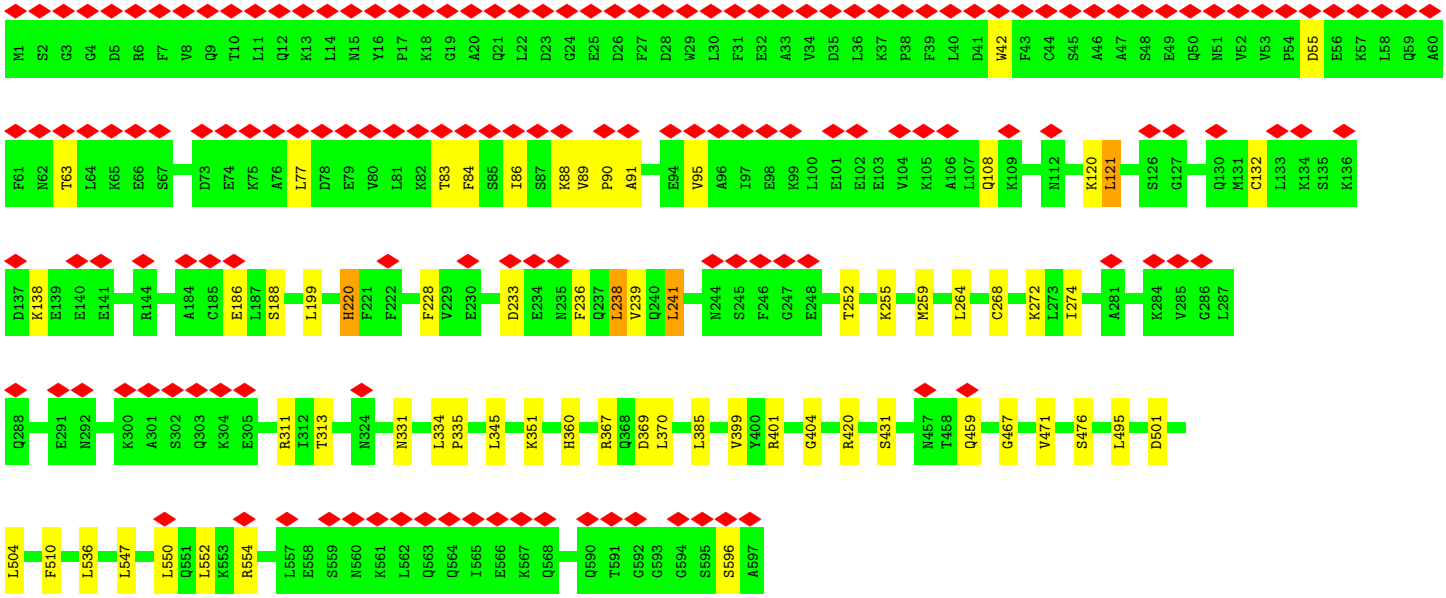




• Molecule 2: HAUS augmin-like complex subunit 3

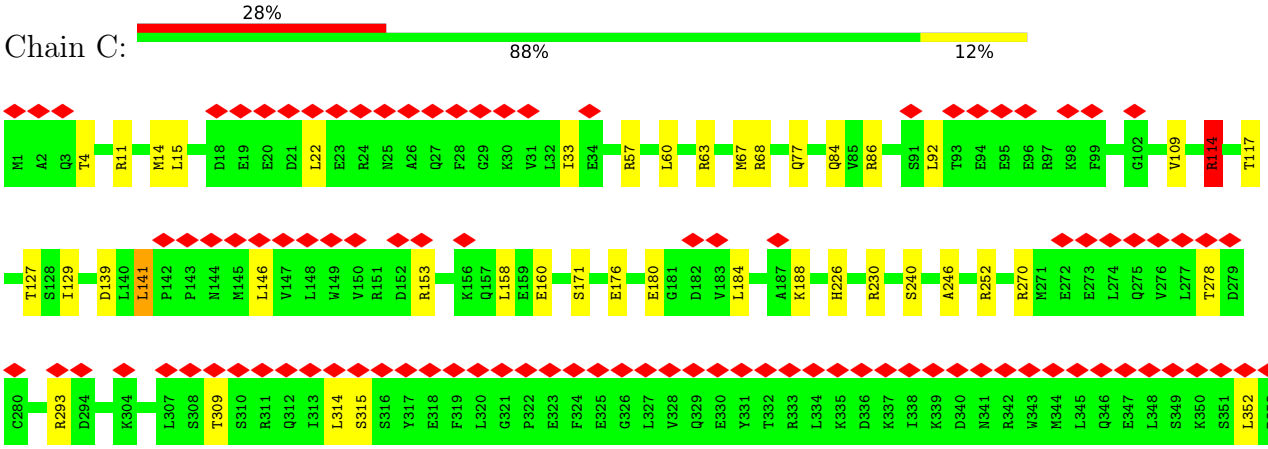
Chain B:





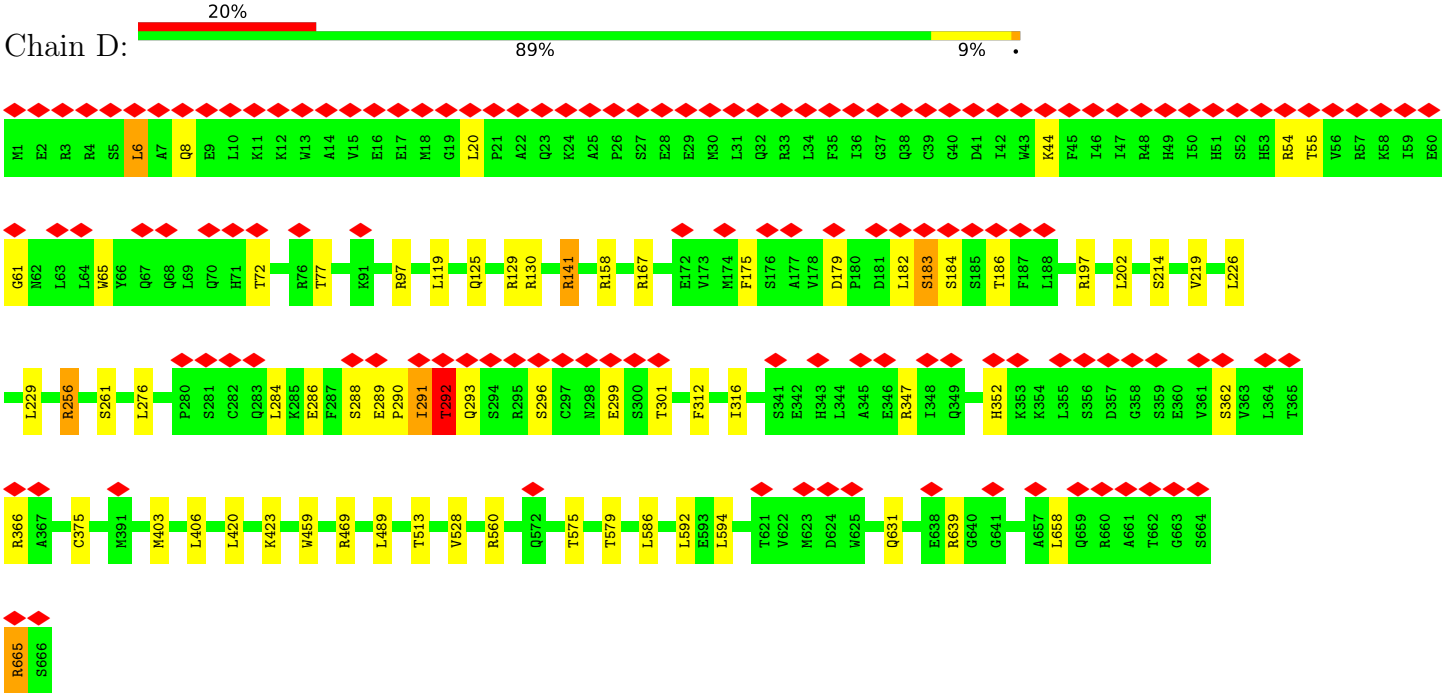
• Molecule 3: HAUS augmin like complex subunit 4 L homeolog

Chain C:



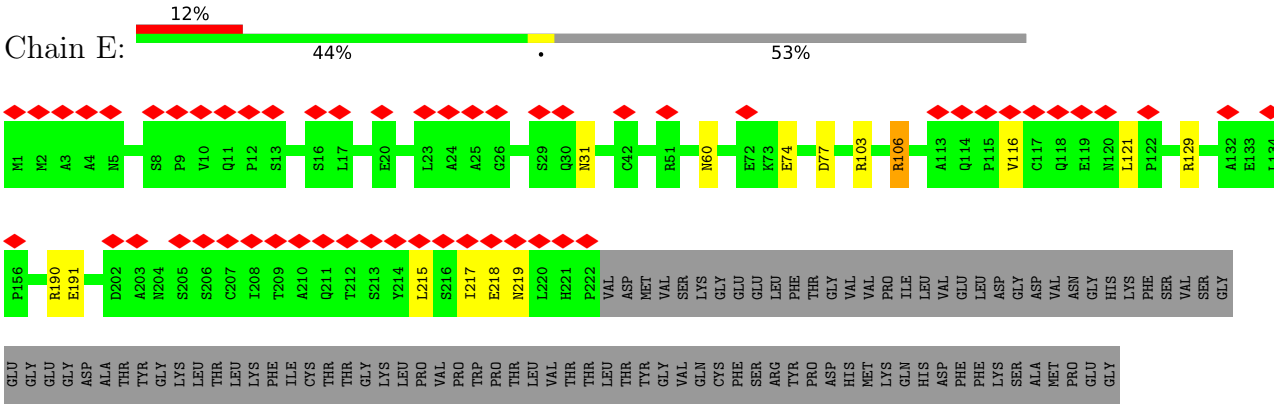
• Molecule 4: HAUS augmin-like complex subunit 5

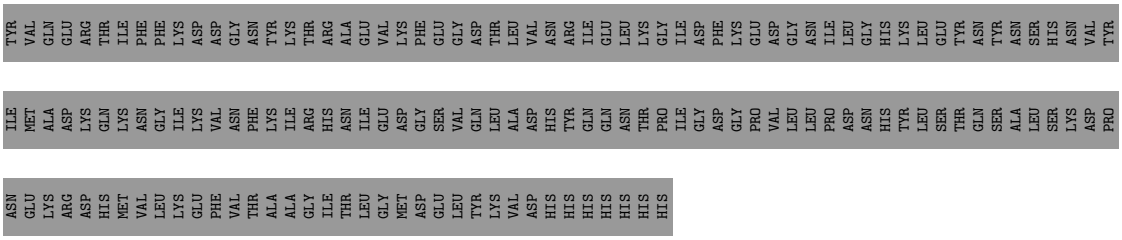
Chain D:



• Molecule 5: HAUS augmin like complex subunit 2 L homeolog, Green fluorescent protein chimera

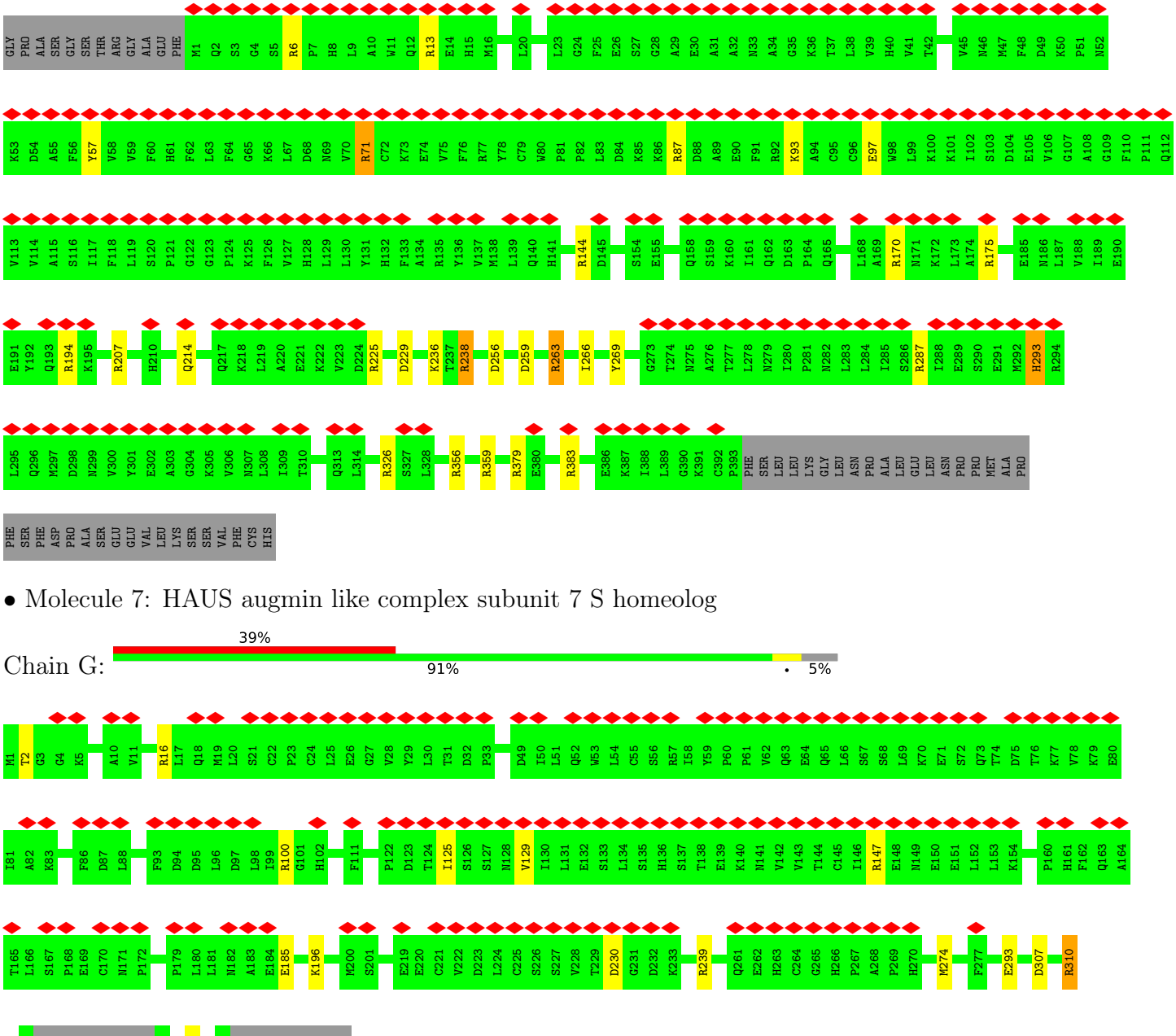
Chain E:





• Molecule 6: HAUS augmin like complex subunit 6 L homeolog

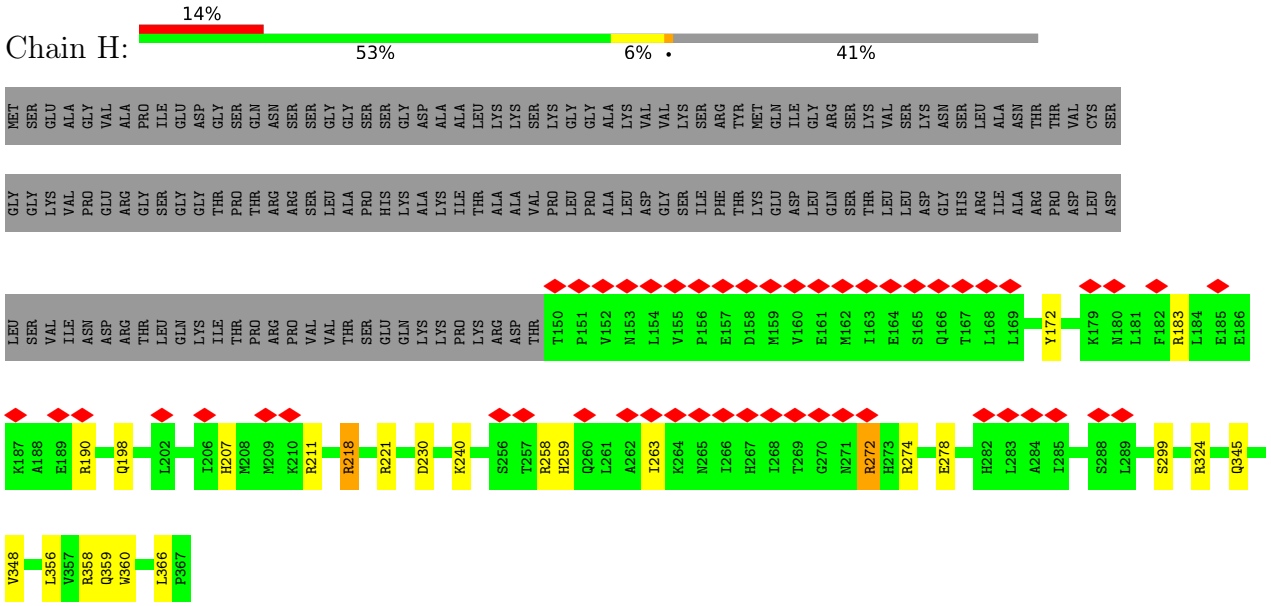
Chain F:



• Molecule 7: HAUS augmin like complex subunit 7 S homeolog

Chain G:

• Molecule 8: HAUS augmin-like complex subunit 8



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	114000	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 ($e^-/Å^2$)	66	Depositor
Minimum defocus (nm)	10105	Depositor
Maximum defocus (nm)	39903	Depositor
Magnification	64000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	1.224	Depositor
Minimum map value	-0.749	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.019	Depositor
Recommended contour level	0.161	Depositor
Map size (Å)	712.4, 712.4, 712.4	wwPDB
Map dimensions	520, 520, 520	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.37, 1.37, 1.37	Depositor

5 Model quality

5.1 Standard geometry

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	0.62	0/2308	0.91	5/3100 (0.2%)
2	B	0.61	0/4821	0.86	6/6477 (0.1%)
3	C	0.65	0/2921	0.94	12/3925 (0.3%)
4	D	0.63	0/5504	0.95	16/7400 (0.2%)
5	E	0.63	0/1783	0.98	4/2416 (0.2%)
6	F	0.94	5/3270 (0.2%)	1.11	20/4392 (0.5%)
7	G	0.65	0/2678	0.96	6/3622 (0.2%)
8	H	0.66	0/1809	1.04	10/2440 (0.4%)
All	All	0.68	5/25094 (0.0%)	0.96	79/33772 (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	12
2	B	0	2
3	C	0	4
4	D	0	8
5	E	0	3
6	F	0	5
7	G	0	1
8	H	0	3
All	All	0	38

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	F	293	HIS	CD2-NE2	19.01	1.81	1.42
6	F	293	HIS	CG-CD2	18.79	1.67	1.35
6	F	293	HIS	CG-ND1	16.37	1.74	1.38
6	F	293	HIS	CE1-NE2	14.40	1.65	1.32

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	F	293	HIS	ND1-CE1	12.77	1.66	1.34

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	F	326	ARG	NE-CZ-NH1	10.89	125.74	120.30
3	C	114	ARG	NE-CZ-NH2	-10.68	114.96	120.30
6	F	238	ARG	NE-CZ-NH1	9.97	125.29	120.30
6	F	287	ARG	NE-CZ-NH1	9.32	124.96	120.30
4	D	97	ARG	NE-CZ-NH1	9.10	124.85	120.30

There are no chirality outliers.

5 of 38 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	212	ARG	Sidechain
1	A	273	MET	Peptide
1	A	274	GLU	Peptide
1	A	275	PHE	Peptide
1	A	276	VAL	Peptide

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	2281	0	2360	21	0
2	B	4756	0	4769	36	0
3	C	2886	0	2957	15	0
4	D	5416	0	5455	21	0
5	E	1755	0	1763	6	0
6	F	3211	0	3280	7	0
7	G	2627	0	2592	0	0
8	H	1785	0	1807	22	0
All	All	24717	0	24983	82	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 2.

The worst 5 of 82 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)
6:F:293:HIS:ND1	6:F:293:HIS:CG	1.74	1.52
6:F:293:HIS:NE2	6:F:293:HIS:CD2	1.81	1.46
2:B:268:CYS:SG	8:H:345:GLN:HA	1.91	1.11
5:E:60:ASN:HB3	8:H:207:HIS:CD2	1.85	1.10
5:E:60:ASN:HB3	8:H:207:HIS:NE2	1.83	0.92

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	284/286 (99%)	274 (96%)	7 (2%)	3 (1%)	14	52
2	B	595/597 (100%)	578 (97%)	9 (2%)	8 (1%)	12	48
3	C	351/353 (99%)	344 (98%)	7 (2%)	0	100	100
4	D	664/666 (100%)	634 (96%)	21 (3%)	9 (1%)	11	46
5	E	220/472 (47%)	210 (96%)	9 (4%)	1 (0%)	29	69
6	F	391/442 (88%)	364 (93%)	27 (7%)	0	100	100
7	G	328/348 (94%)	305 (93%)	19 (6%)	4 (1%)	13	50
8	H	216/367 (59%)	207 (96%)	7 (3%)	2 (1%)	17	57
All	All	3049/3531 (86%)	2916 (96%)	106 (4%)	27 (1%)	21	57

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	279	GLU
1	A	280	PRO
2	B	186	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	184	SER
7	G	2	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	260/260 (100%)	249 (96%)	11 (4%)	30	54
2	B	538/541 (99%)	511 (95%)	27 (5%)	24	49
3	C	326/326 (100%)	308 (94%)	18 (6%)	21	47
4	D	605/605 (100%)	579 (96%)	26 (4%)	29	53
5	E	195/415 (47%)	190 (97%)	5 (3%)	46	66
6	F	347/387 (90%)	342 (99%)	5 (1%)	67	80
7	G	306/320 (96%)	301 (98%)	5 (2%)	62	79
8	H	205/328 (62%)	199 (97%)	6 (3%)	42	64
All	All	2782/3182 (87%)	2679 (96%)	103 (4%)	37	58

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	8	GLN
4	D	316	ILE
8	H	258	ARG
4	D	72	THR
4	D	214	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
4	D	161	GLN
4	D	211	HIS
4	D	426	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	D	272	GLN
3	C	84	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)

There are no chain breaks in this entry.

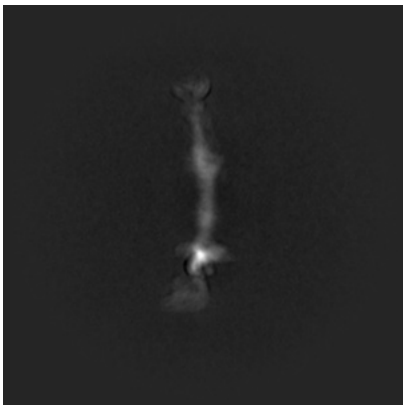
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-28981. These allow visual inspection of the internal detail of the map and identification of artifacts.

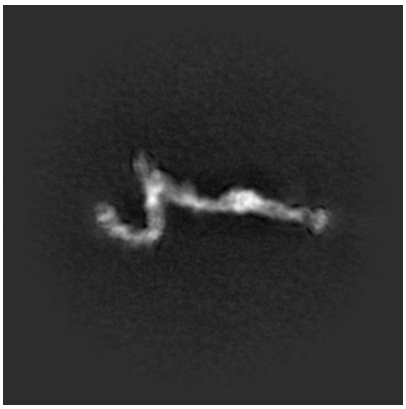
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

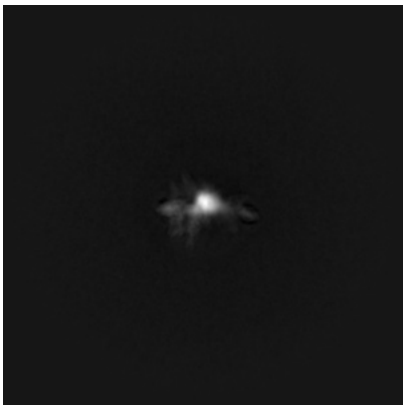
6.1.1 Primary map



X

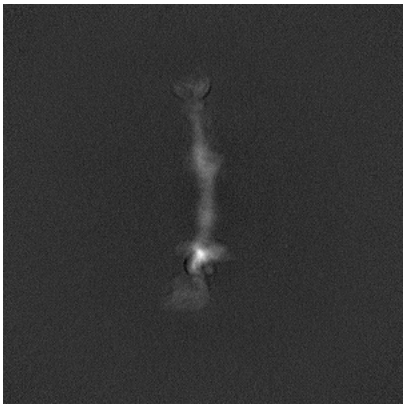


Y

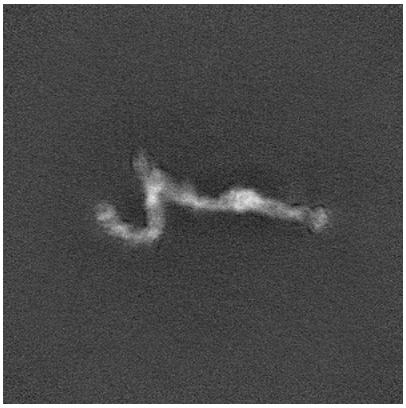


Z

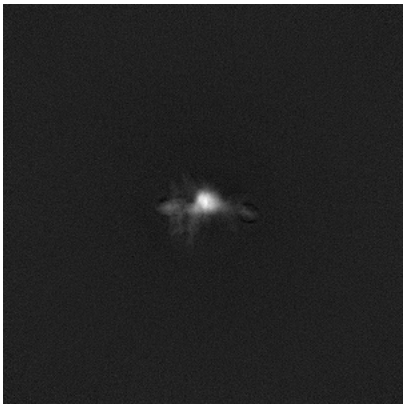
6.1.2 Raw map



X



Y

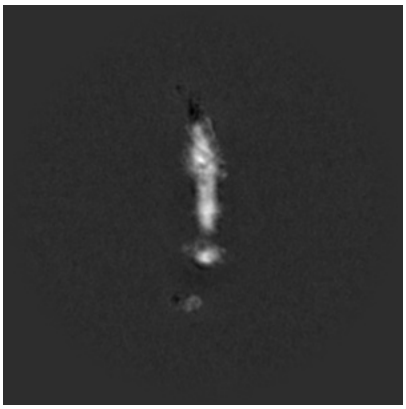


Z

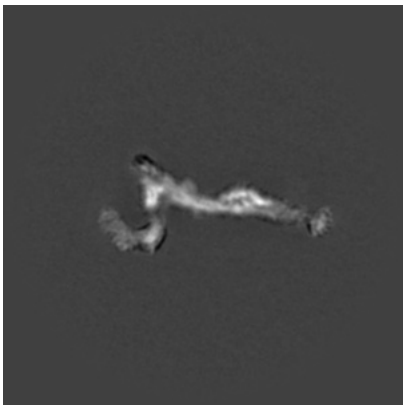
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

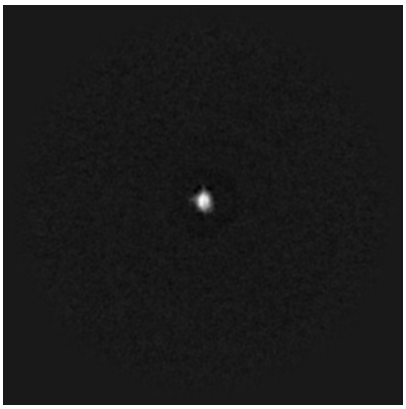
6.2.1 Primary map



X Index: 260

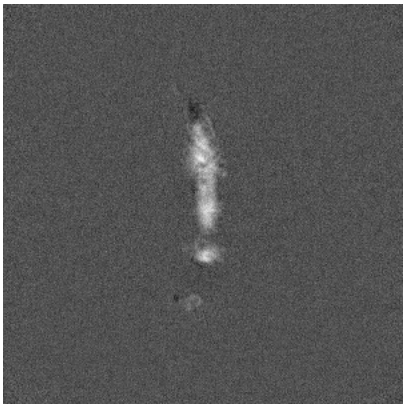


Y Index: 260



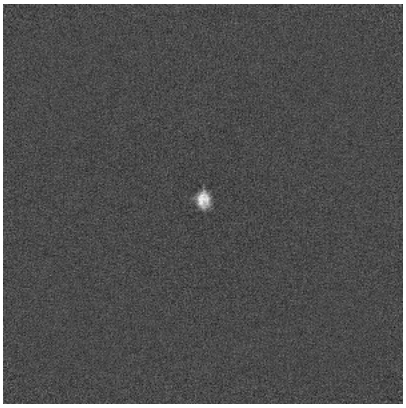
Z Index: 260

6.2.2 Raw map



X Index: 260

Y Index: 260

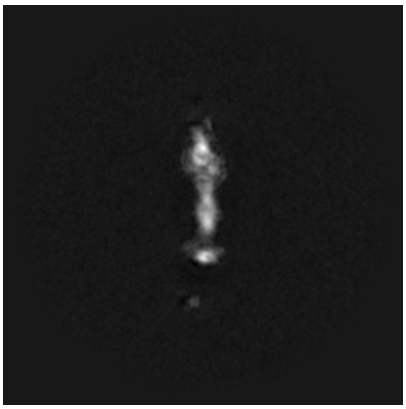


Z Index: 260

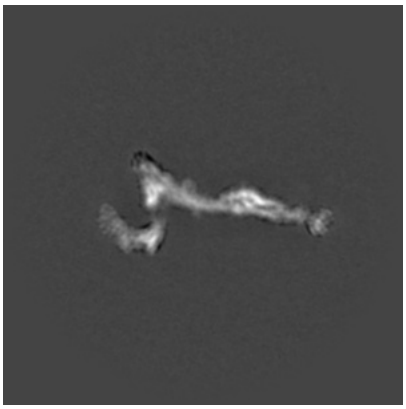
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

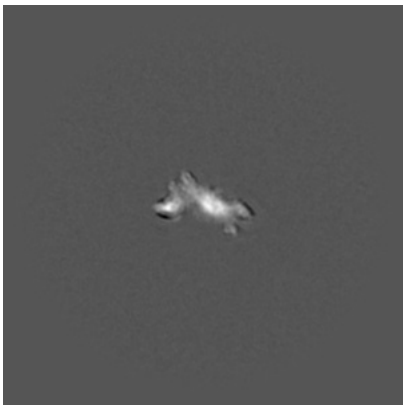
6.3.1 Primary map



X Index: 263

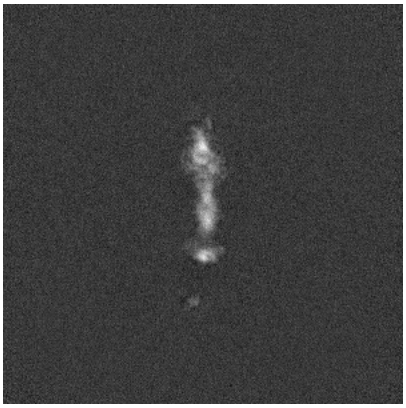


Y Index: 257



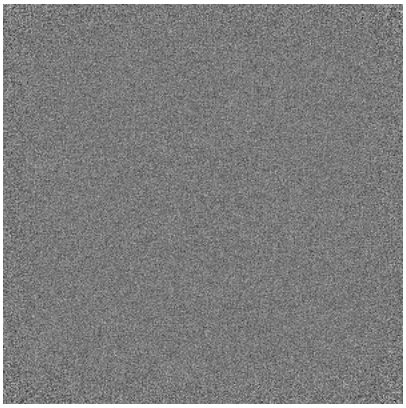
Z Index: 194

6.3.2 Raw map



X Index: 263

Y Index: 257

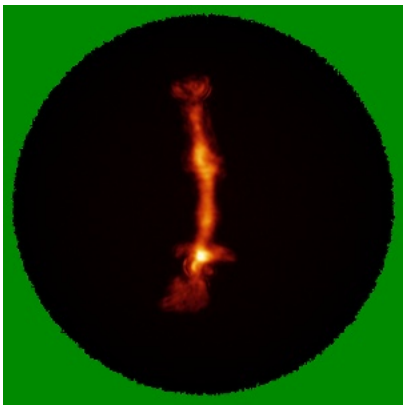


Z Index: 0

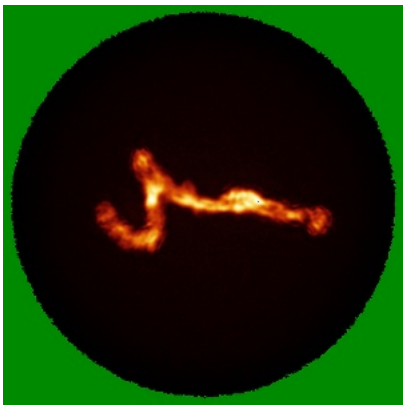
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map



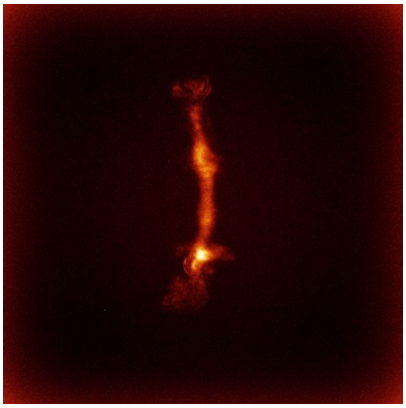
X



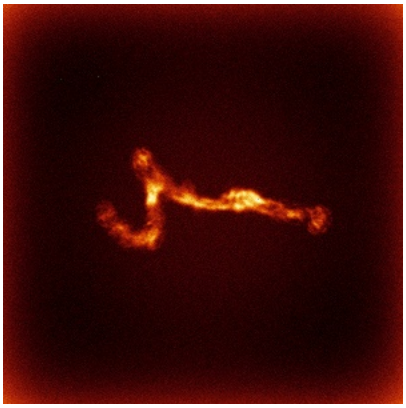
Y

Z

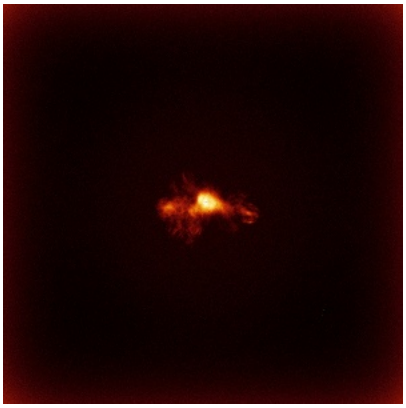
6.4.2 Raw map



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 0.161. 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.

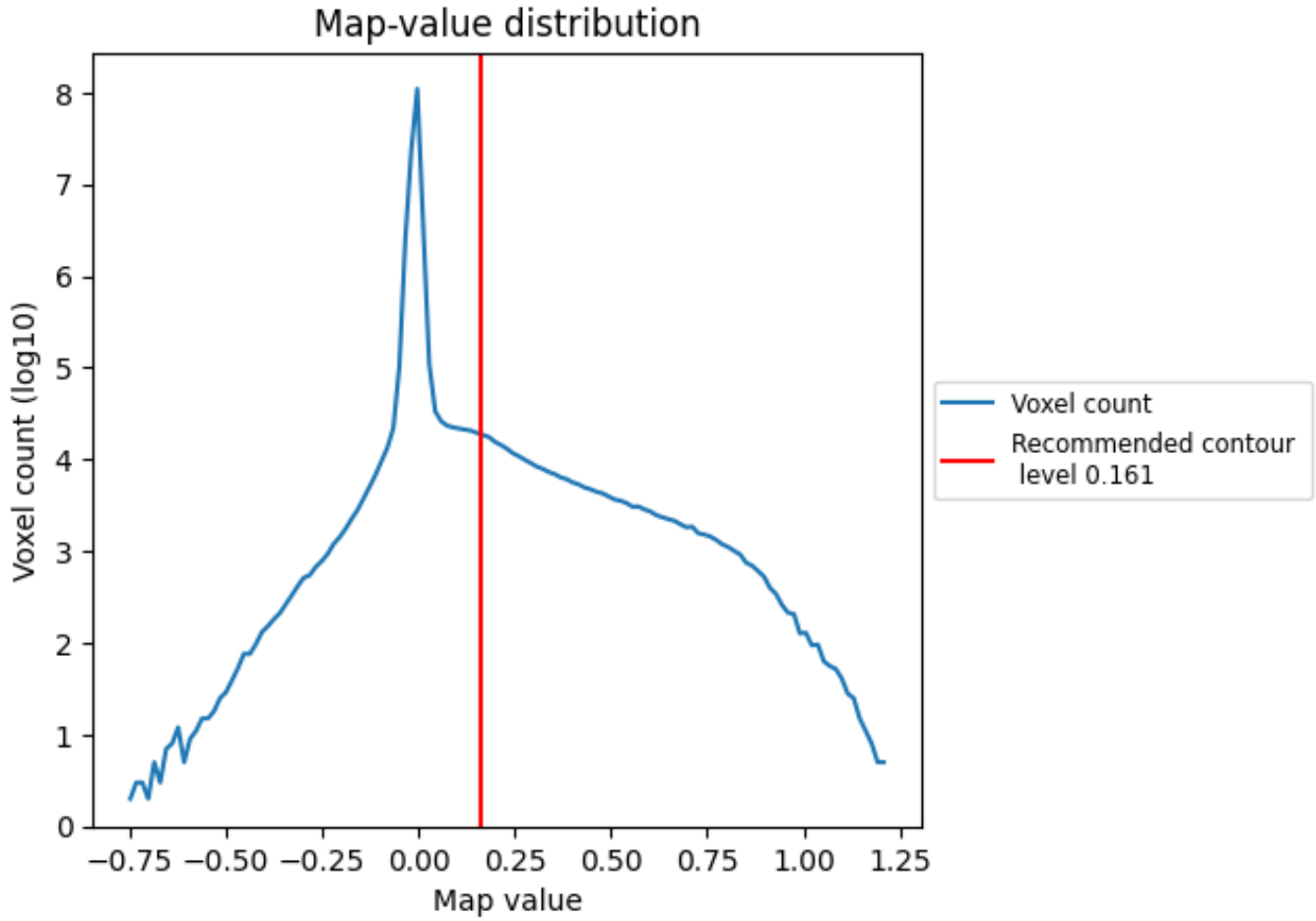
6.6 Mask visualisation (i)

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

7 Map analysis

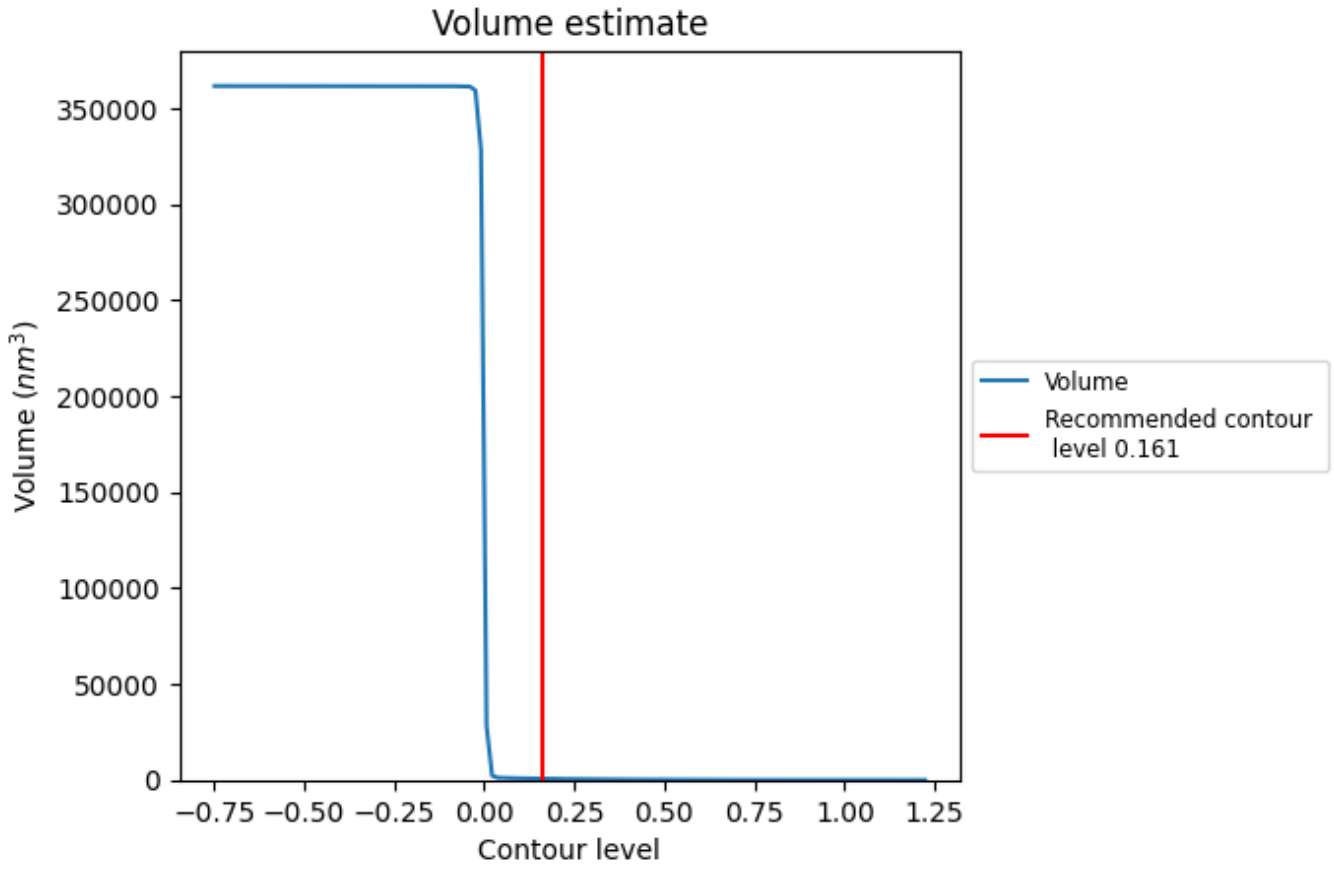
This section contains the results of statistical analysis of the map.

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.

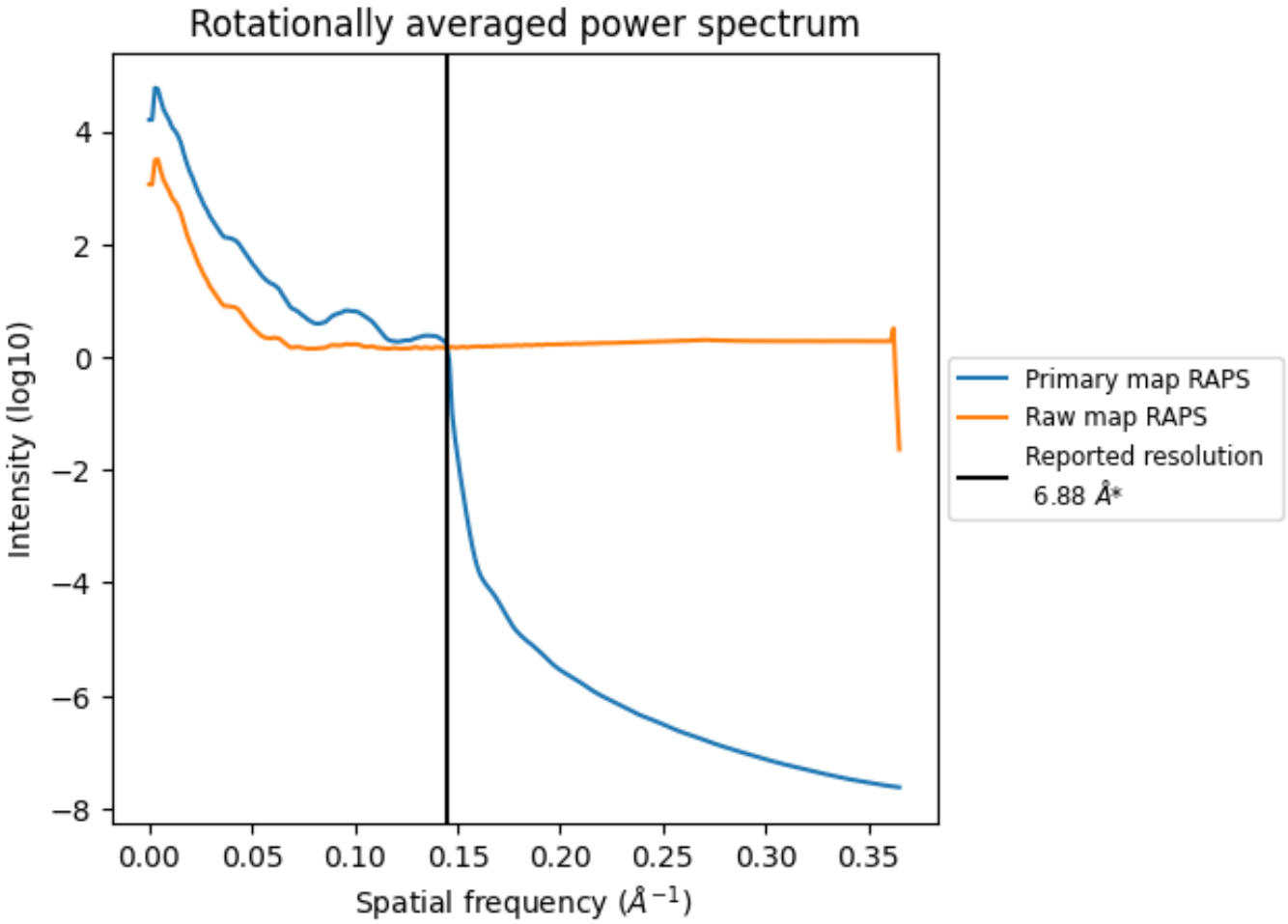
7.2 Volume estimate (i)



The volume at the recommended contour level is 661 nm^3; this corresponds to an approximate mass of 597 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)

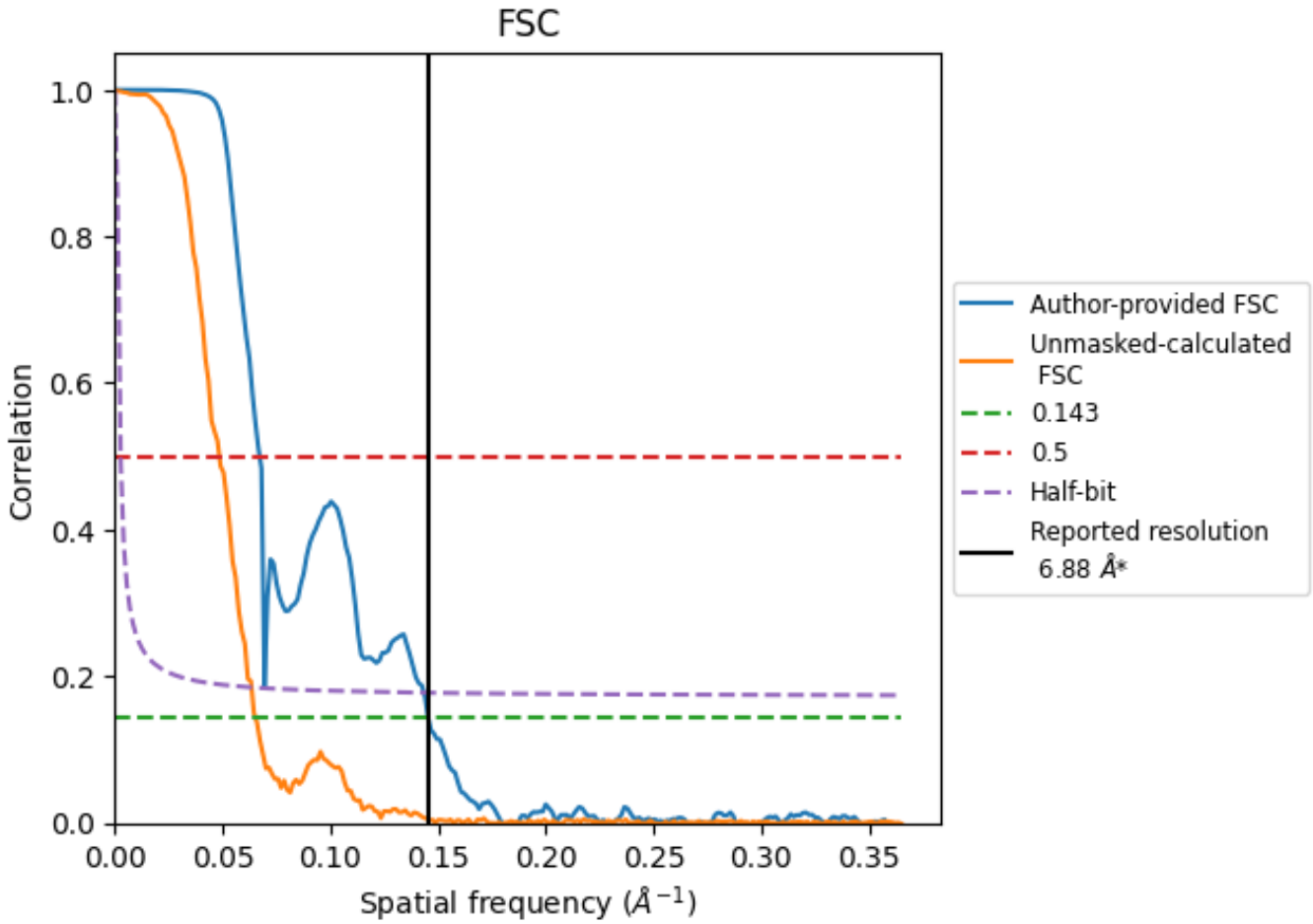


*Reported resolution corresponds to spatial frequency of 0.145 Å^{-1}

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.145 Å^{-1}

8.2 Resolution estimates (i)

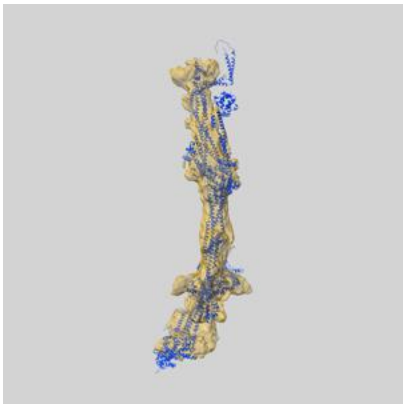
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	6.88	-	-
Author-provided FSC curve	6.88	14.88	6.97
Unmasked-calculated*	15.31	20.58	15.77

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 15.31 differs from the reported value 6.88 by more than 10 %

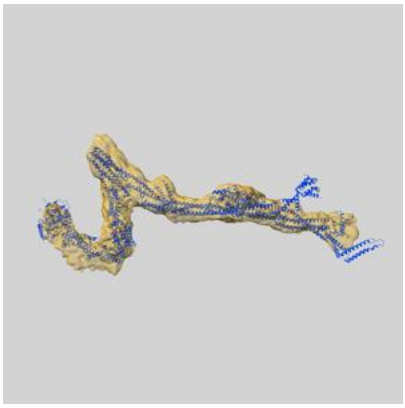
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-28981 and PDB model 8FCK. Per-residue inclusion information can be found in section 3 on page 6.

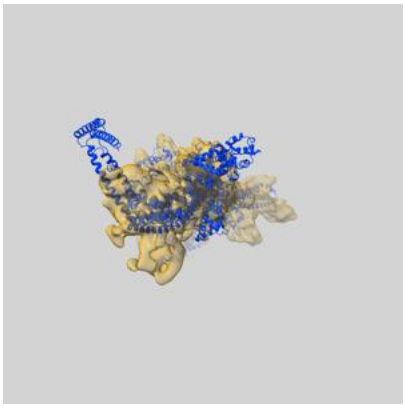
9.1 Map-model overlay



X



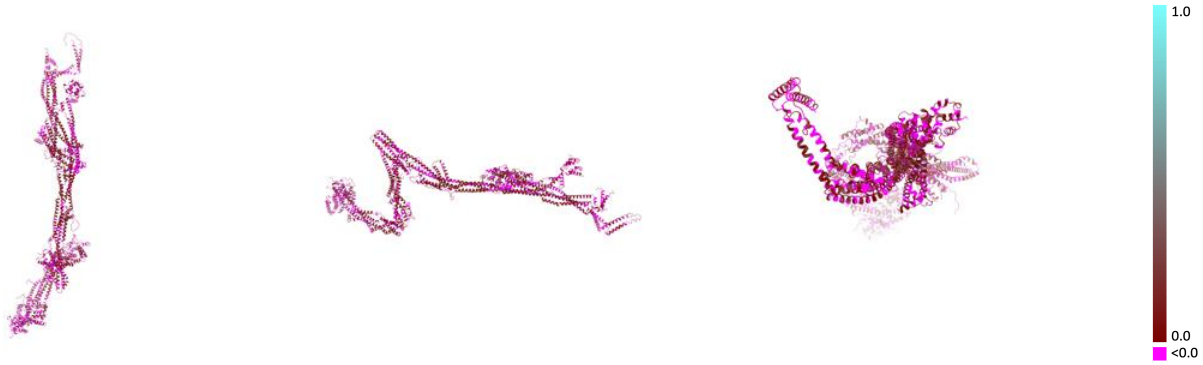
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.161 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.

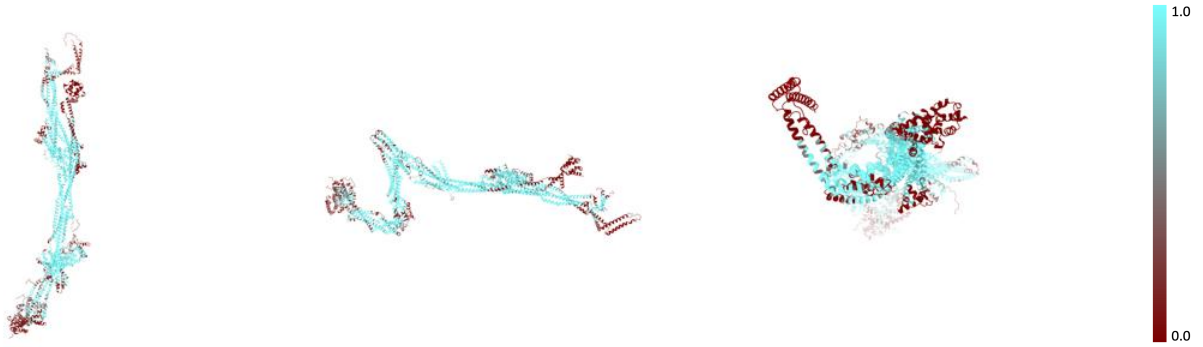
9.2 Q-score mapped to coordinate model (i)



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.

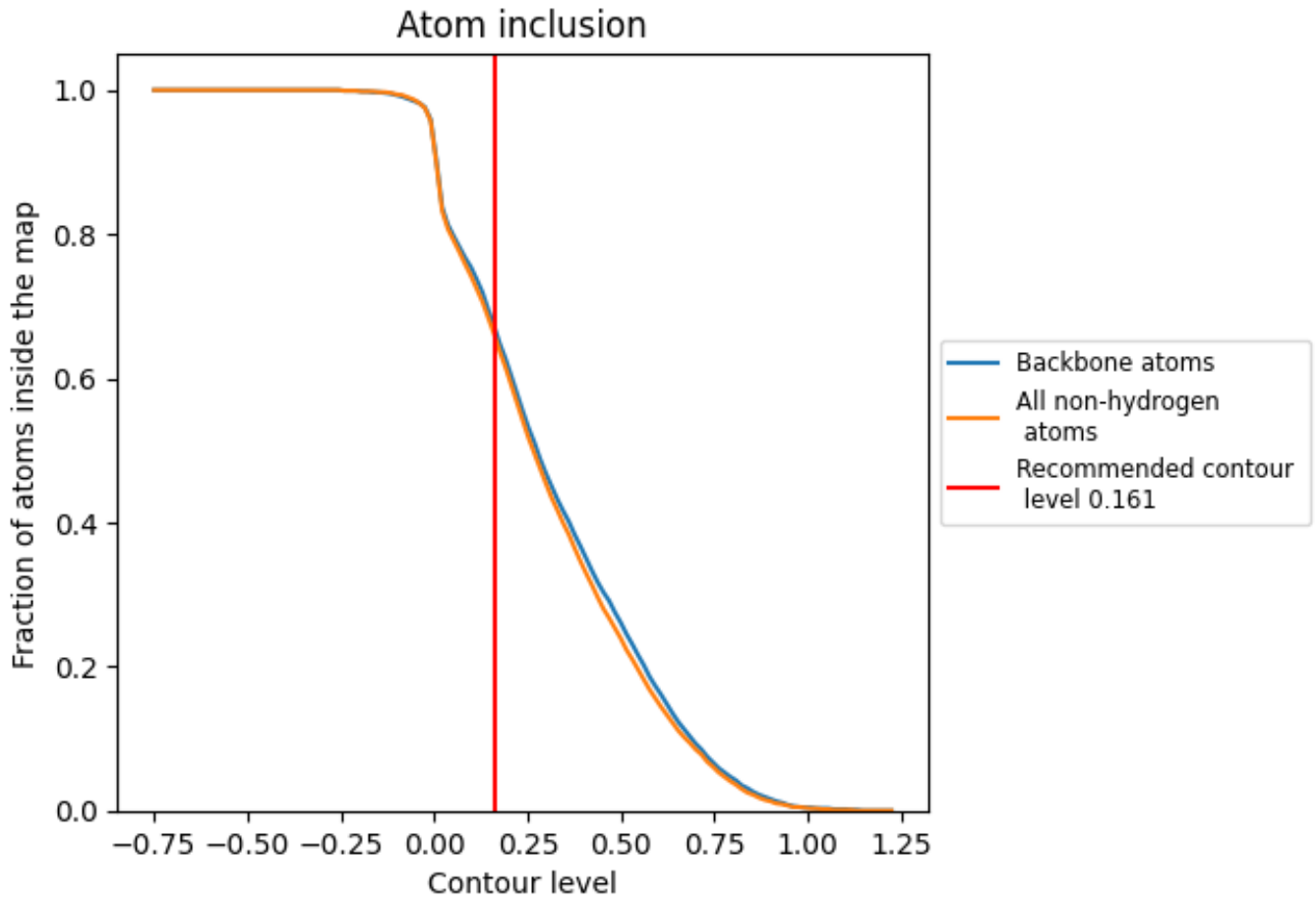
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.161).

9.4 Atom inclusion (i)



At the recommended contour level, 67% of all backbone atoms, 66% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.161) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6610	0.0350
A	0.5820	0.0470
B	0.7120	0.0240
C	0.6800	0.0350
D	0.7680	0.0530
E	0.7380	0.0430
F	0.4320	0.0210
G	0.5600	0.0230
H	0.7500	0.0280