wwPDB EM Validation Summary Report (i)

Dec 2, 2024 – 06:29 PM JST

PDB ID	:	9JXS
EMDB ID	:	EMD-61880
Title	:	Cryo-EM structure of Cas5-HNH Cascade bound with dsDNA
Authors	:	Liu, Y.N.; Zhang, H.; Zhu, H.
Deposited on	:	2024-10-11
Resolution	:	2.93 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:

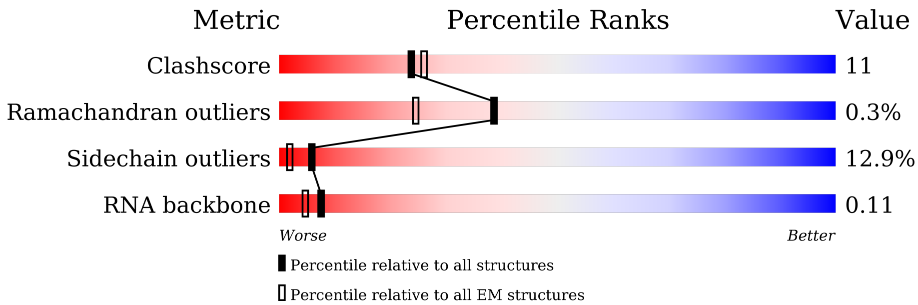
EMDB validation analysis	:	0.0.1.dev113
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 2.93 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	C	535	• 52% 26% • 18%
2	E	174	72% 22% • 5%
3	F	378	• 67% 27% • •
3	G	378	• 37% 23% 6% 34%
3	H	378	• 70% 27% • •
3	I	378	• 71% 24% • •
3	J	378	• 70% 27% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	K	378	66% 22% • 9%
4	B	388	55% 28% 6% 11%
5	M	54	72% 19% 9%
6	D	272	60% 32% • • •
7	A	61	15% 51% 30% • •
8	N	54	13% 87%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 27905 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 CRISPR-associated protein Cse1 (CRISPR_cse1).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	C	438	3460	2226	591	624	19	0	0

- Molecule 2 is a protein called CRISPR-associated protein Cse2 (CRISPR_cse2).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	E	165	1352	875	238	233	6	0	0

- Molecule 3 is a protein called CRISPR system Cascade subunit CasC.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	F	368	2819	1782	489	536	12	0	0
3	H	375	2870	1808	503	547	12	0	0
3	I	370	2809	1769	493	535	12	0	0
3	J	371	2819	1779	491	537	12	0	0
3	K	345	2650	1672	466	501	11	0	0
3	G	249	1881	1191	327	356	7	0	0

- Molecule 4 is a protein called CRISPR system Cascade subunit CasD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	B	346	2710	1709	500	484	17	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	333	ALA	HIS	conflict	UNP A0A1V6F8C5

- Molecule 5 is a DNA chain called DNA (54-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	M	49	Total 995	C 475	N 176	O 295	P 49	0	0

- Molecule 6 is a protein called CRISPR-associated endoribonuclease Cse3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	D	264	Total 2134	C 1368	N 383	O 380	S 3	0	0

- Molecule 7 is a RNA chain called RNA (61-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	A	59	Total 1258	C 562	N 225	O 413	P 58	0	0

- Molecule 8 is a DNA chain called non-target DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	N	7	Total 146	C 69	N 27	O 43	P 7	0	0

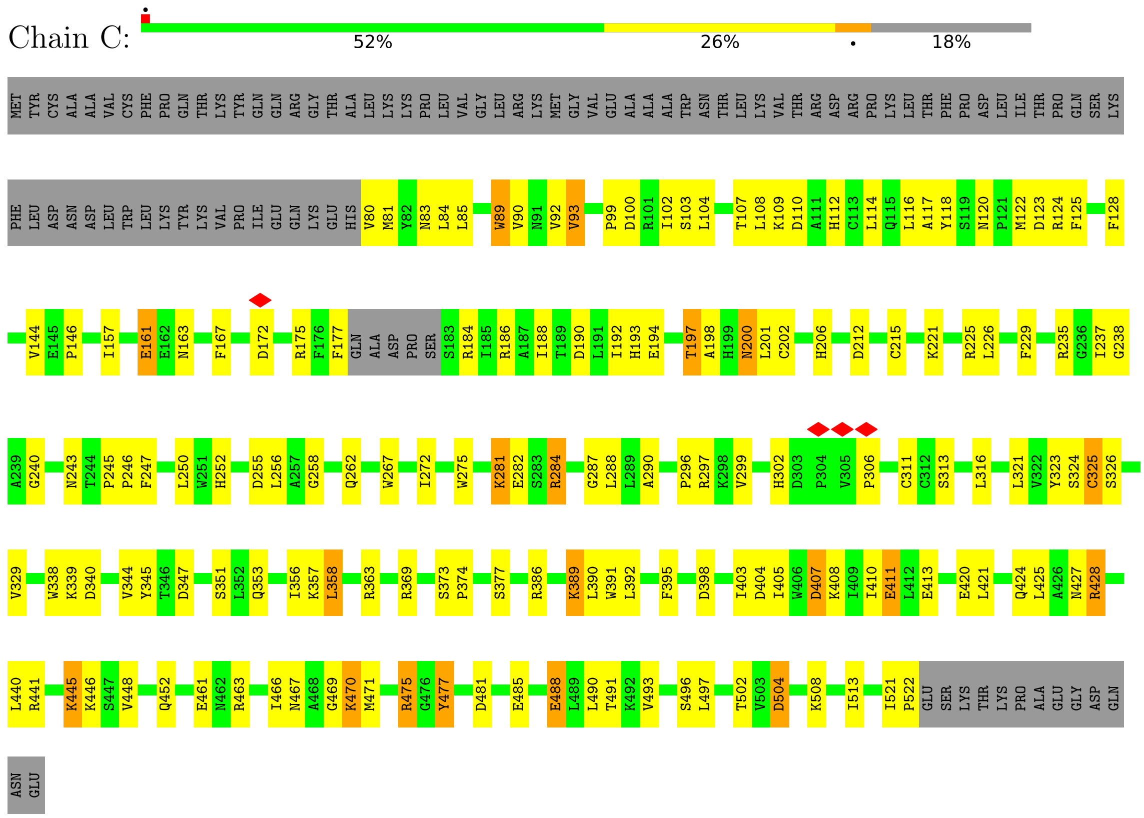
- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
9	C	1	Total 1	Mg 1	0
9	B	1	Total 1	Mg 1	0

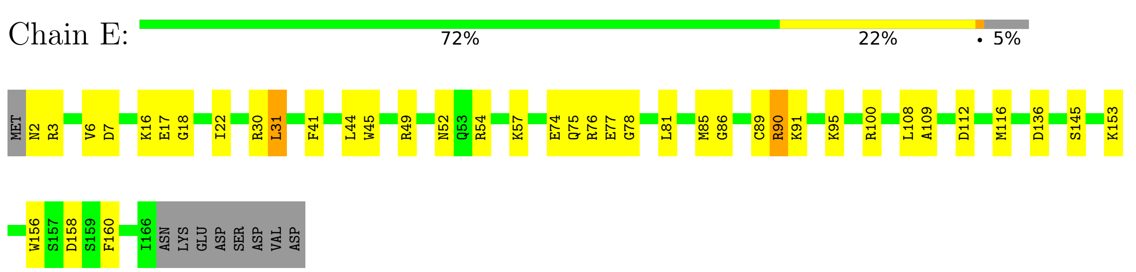
3 Residue-property plots

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: CRISPR-associated protein Cse1 (CRISPR_cse1)

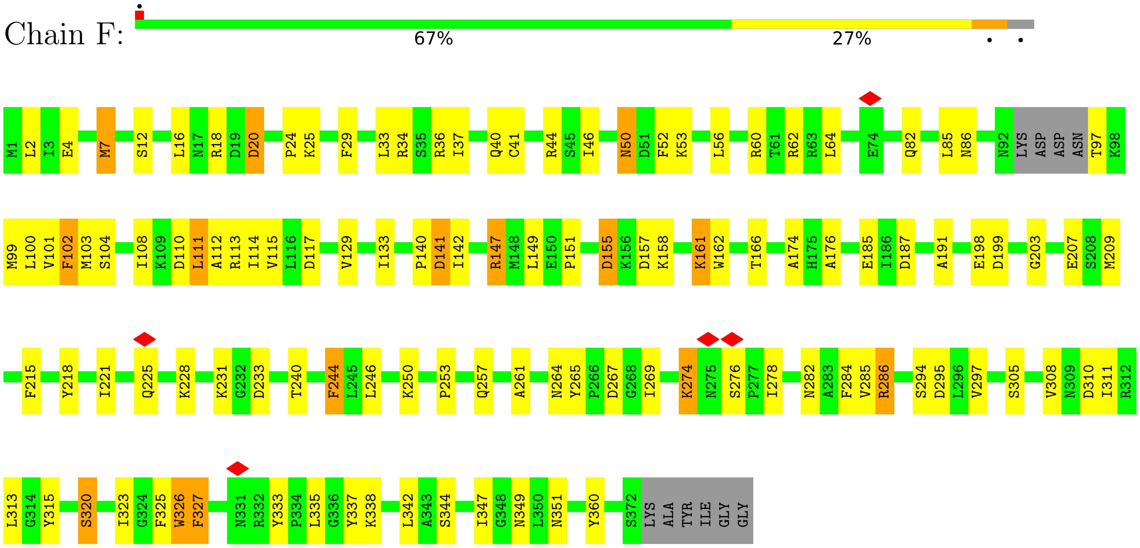


- Molecule 2: CRISPR-associated protein Cse2 (CRISPR_cse2)



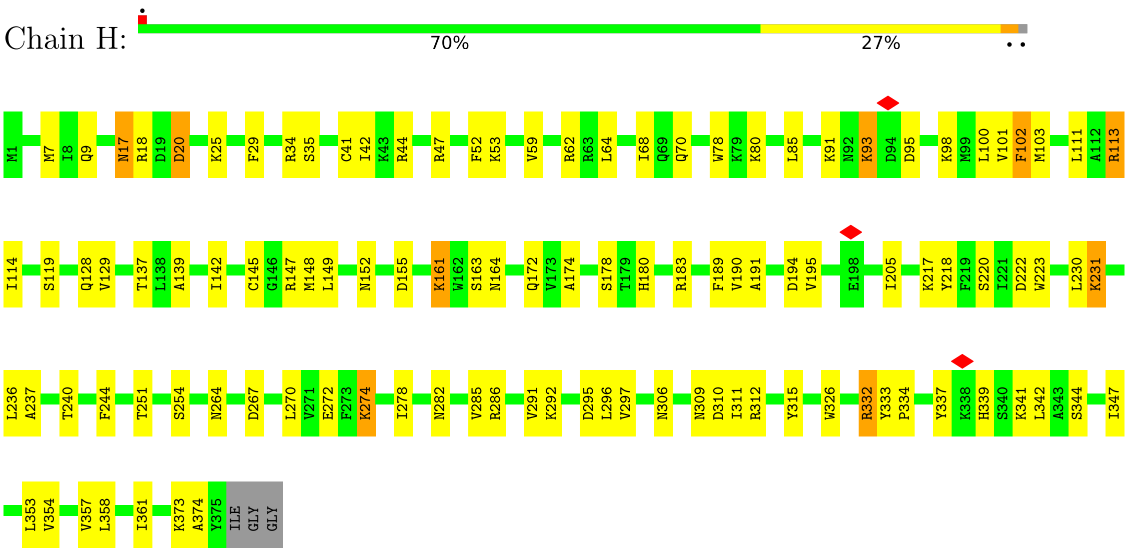
- Molecule 3: CRISPR system Cascade subunit CasC

Chain F:



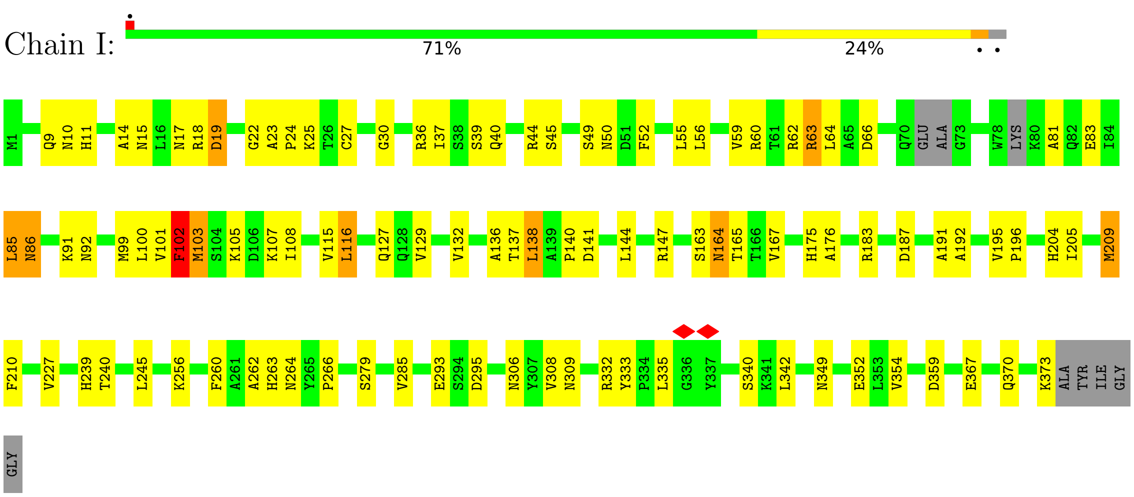
• Molecule 3: CRISPR system Cascade subunit CasC

Chain H:

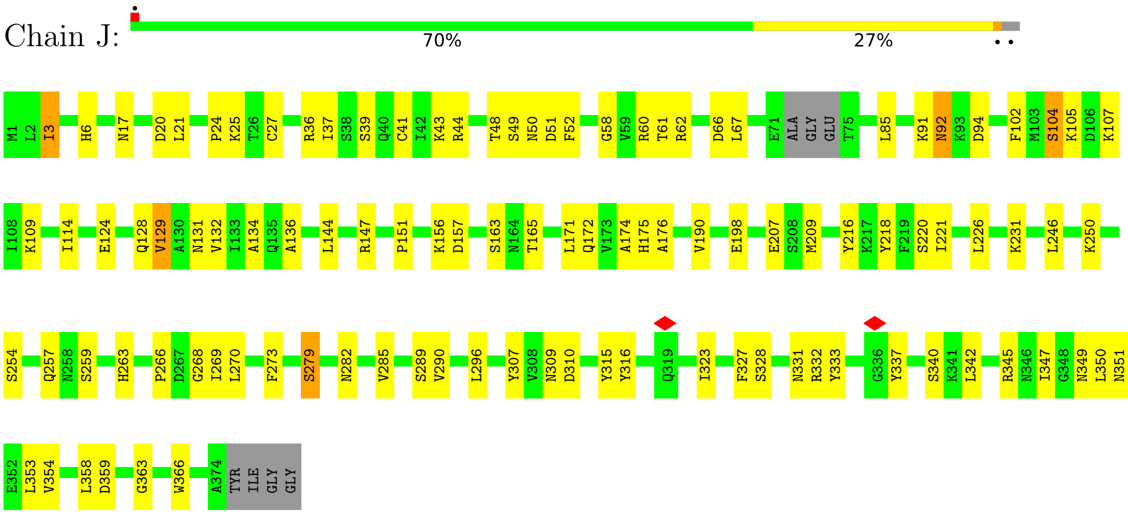


• Molecule 3: CRISPR system Cascade subunit CasC

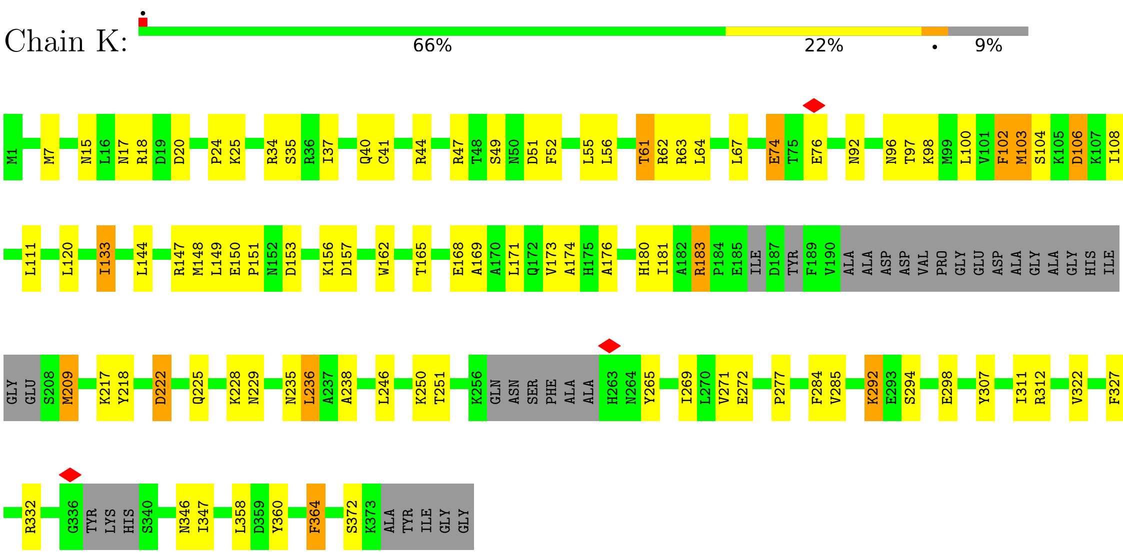
Chain I:



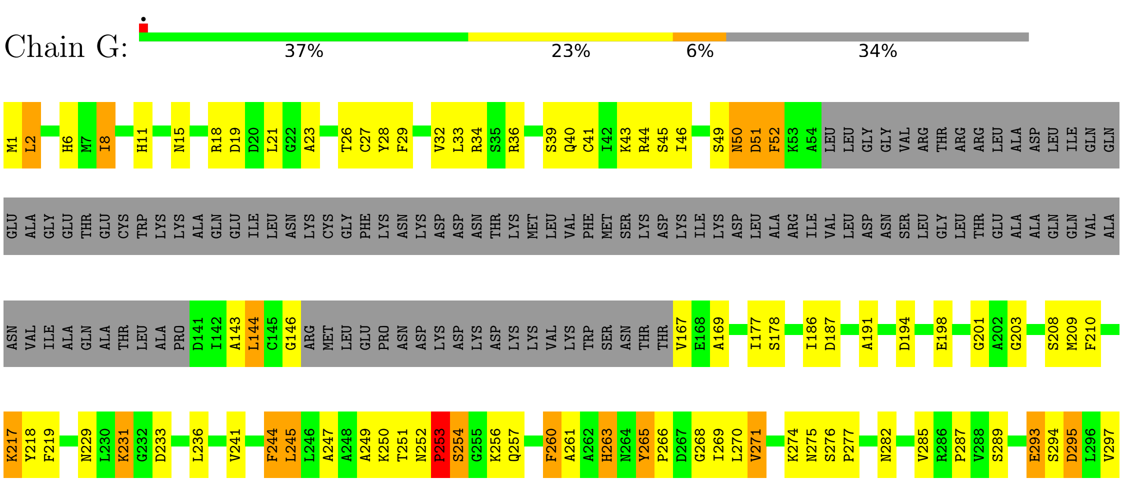
- Molecule 3: CRISPR system Cascade subunit CasC

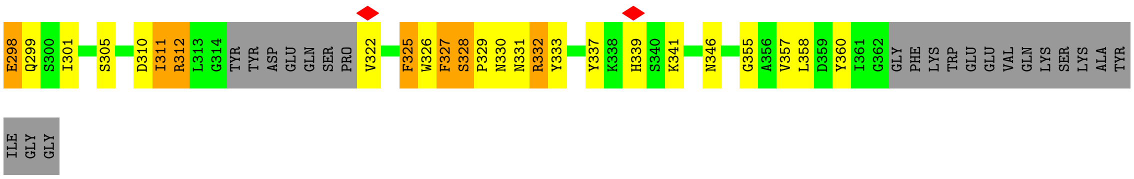


- Molecule 3: CRISPR system Cascade subunit CasC



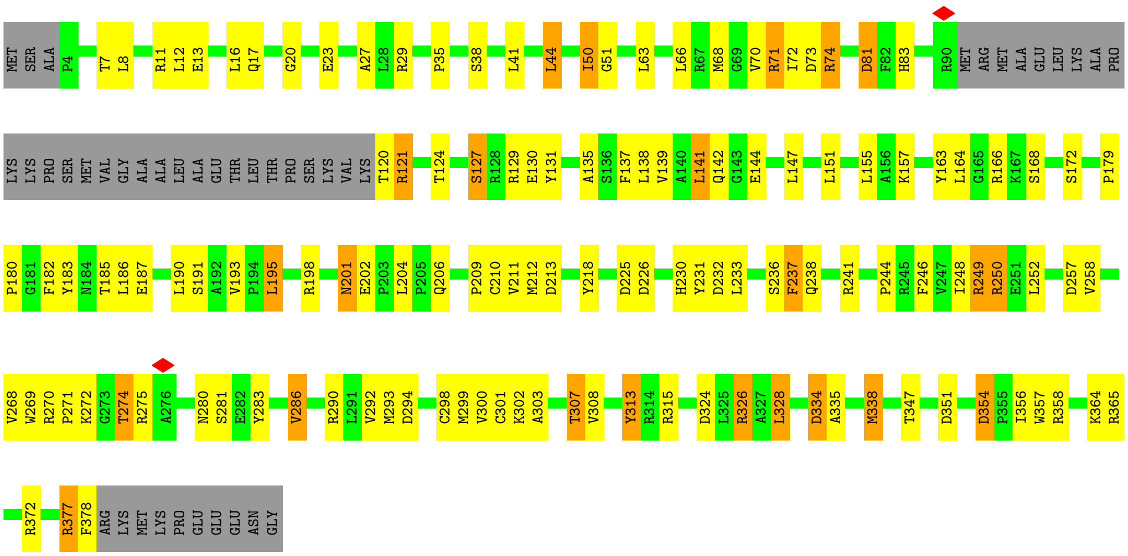
- Molecule 3: CRISPR system Cascade subunit CasC





• Molecule 4: CRISPR system Cascade subunit CasD

Chain B: 55% 28% 6% 11%

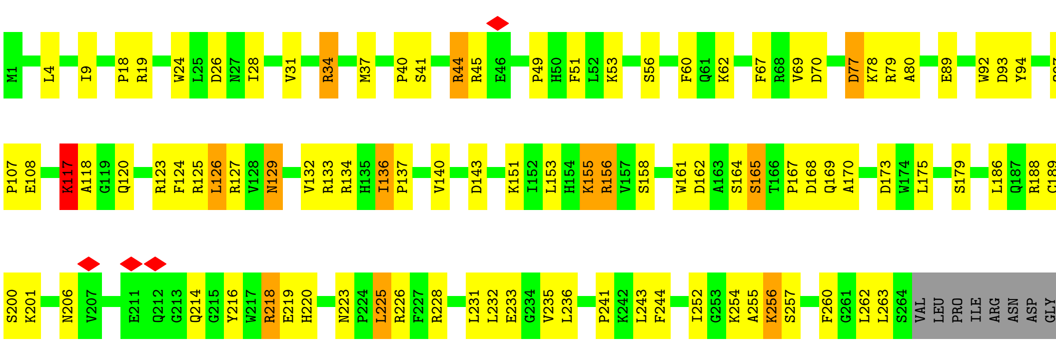


• Molecule 5: DNA (54-MER)

Chain M: 72% 19% 9%

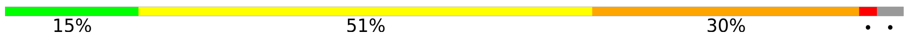
• Molecule 6: CRISPR-associated endoribonuclease Cse3

Chain D: 60% 32%



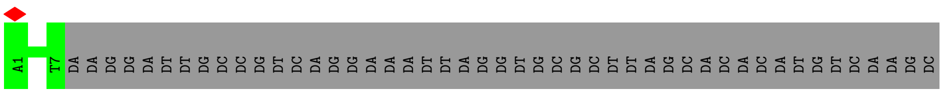
• Molecule 7: RNA (61-MER)

Chain A:



• Molecule 8: non-target DNA

Chain N:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	73951	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{Å}^2$)	48.62	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1200	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	16.150	Depositor
Minimum map value	-0.144	Depositor
Average map value	0.012	Depositor
Map value standard deviation	0.662	Depositor
Recommended contour level	1.04	Depositor
Map size (Å)	291.0, 291.0, 291.0	wwPDB
Map dimensions	358, 358, 358	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.81284916, 0.81284916, 0.81284916	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	C	0.31	0/3559	0.69	5/4848 (0.1%)
2	E	0.31	0/1382	0.69	0/1862
3	F	0.33	0/2875	0.66	2/3900 (0.1%)
3	G	0.41	1/1920 (0.1%)	0.70	4/2606 (0.2%)
3	H	0.30	0/2926	0.59	1/3967 (0.0%)
3	I	0.30	0/2861	0.61	3/3878 (0.1%)
3	J	0.35	0/2874	0.62	2/3901 (0.1%)
3	K	0.29	0/2695	0.58	2/3644 (0.1%)
4	B	0.30	0/2779	0.75	7/3777 (0.2%)
5	M	0.66	0/1113	1.08	2/1713 (0.1%)
6	D	0.31	0/2191	0.79	10/2968 (0.3%)
7	A	0.28	0/1405	1.01	3/2187 (0.1%)
8	N	0.47	0/163	0.93	0/250
All	All	0.34	1/28743 (0.0%)	0.71	41/39501 (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
2	E	0	1
3	I	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	253	PRO	N-CD	10.22	1.62	1.47

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	108	LEU	CA-CB-CG	8.81	135.56	115.30
6	D	173	ASP	CB-CG-OD1	8.51	125.96	118.30
6	D	175	LEU	CA-CB-CG	8.39	134.61	115.30
3	G	245	LEU	CA-CB-CG	8.26	134.29	115.30
6	D	93	ASP	CB-CG-OD1	7.41	124.97	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	90	ARG	Sidechain
3	I	102	PHE	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	C	3460	0	3373	92	0
2	E	1352	0	1387	21	0
3	F	2819	0	2745	76	0
3	G	1881	0	1802	74	0
3	H	2870	0	2797	57	0
3	I	2809	0	2716	62	0
3	J	2819	0	2714	56	0
3	K	2650	0	2599	43	0
4	B	2710	0	2673	83	0
5	M	995	0	553	8	0
6	D	2134	0	2126	52	0
7	A	1258	0	638	36	0
8	N	146	0	80	0	0
9	B	1	0	0	0	0
9	C	1	0	0	0	0
All	All	27905	0	26203	575	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 11.

The worst 5 of 575 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:H:311:ILE:O	3:H:315:TYR:HB2	1.55	1.07
6:D:117:LYS:HD2	6:D:118:ALA:H	1.40	0.84
3:J:6:HIS:HB2	3:J:270:LEU:HB3	1.61	0.82
3:K:102:PHE:HB3	3:K:147:ARG:HD2	1.62	0.80
3:F:274:LYS:HG2	3:F:276:SER:H	1.46	0.78

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	C	434/535 (81%)	389 (90%)	44 (10%)	1 (0%)	44	66
2	E	163/174 (94%)	154 (94%)	9 (6%)	0	100	100
3	F	364/378 (96%)	339 (93%)	25 (7%)	0	100	100
3	G	241/378 (64%)	209 (87%)	29 (12%)	3 (1%)	11	28
3	H	373/378 (99%)	348 (93%)	25 (7%)	0	100	100
3	I	364/378 (96%)	328 (90%)	36 (10%)	0	100	100
3	J	367/378 (97%)	336 (92%)	29 (8%)	2 (0%)	25	50
3	K	334/378 (88%)	309 (92%)	25 (8%)	0	100	100
4	B	342/388 (88%)	284 (83%)	58 (17%)	0	100	100
6	D	262/272 (96%)	227 (87%)	32 (12%)	3 (1%)	12	30
All	All	3244/3637 (89%)	2923 (90%)	312 (10%)	9 (0%)	38	60

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	306	PRO
3	J	151	PRO
3	G	253	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	G	311	ILE
6	D	117	LYS

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	C	366/459 (80%)	320 (87%)	46 (13%)	3	10
2	E	143/153 (94%)	134 (94%)	9 (6%)	15	33
3	F	295/313 (94%)	256 (87%)	39 (13%)	3	9
3	G	194/313 (62%)	161 (83%)	33 (17%)	1	4
3	H	300/313 (96%)	258 (86%)	42 (14%)	3	7
3	I	291/313 (93%)	263 (90%)	28 (10%)	7	18
3	J	290/313 (93%)	261 (90%)	29 (10%)	6	17
3	K	280/313 (90%)	242 (86%)	38 (14%)	3	8
4	B	281/321 (88%)	240 (85%)	41 (15%)	2	7
6	D	228/238 (96%)	188 (82%)	40 (18%)	1	3
All	All	2668/3049 (88%)	2323 (87%)	345 (13%)	6	9

5 of 345 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	1	MET
4	B	232	ASP
3	G	52	PHE
3	G	332	ARG
4	B	326	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
6	D	97	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	D	120	GLN
3	I	239	HIS
3	J	128	GLN
4	B	142	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
7	A	57/61 (93%)	36 (63%)	0

5 of 36 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
7	A	-6	U
7	A	-1	C
7	A	1	G
7	A	5	G
7	A	6	C

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.

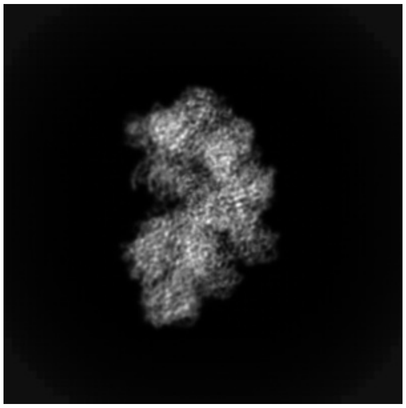
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-61880. These allow visual inspection of the internal detail of the map and identification of artifacts.

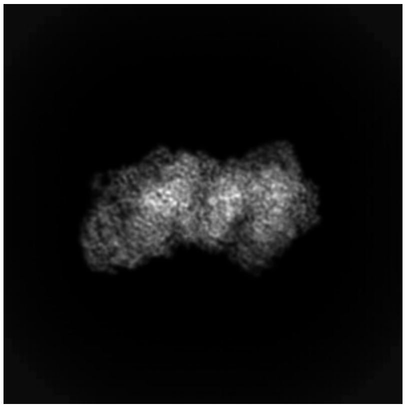
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

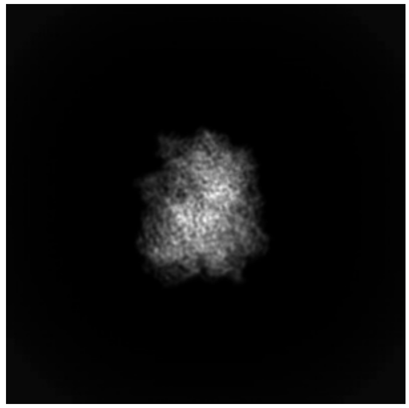
6.1.1 Primary map



X

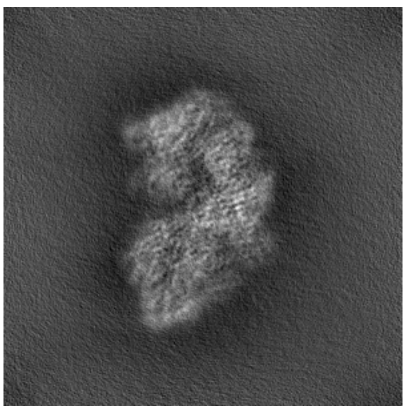


Y

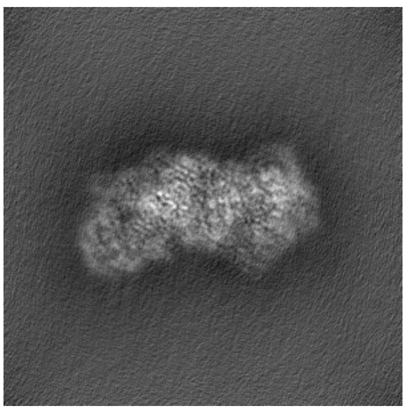


Z

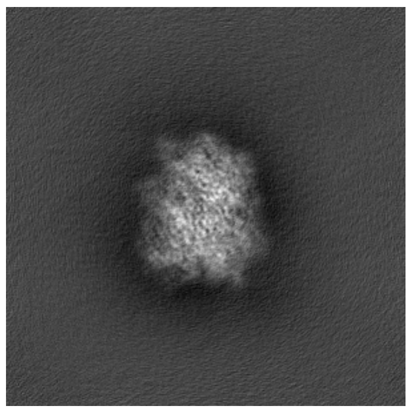
6.1.2 Raw map



X



Y



Z

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

6.2 Central slices (i)

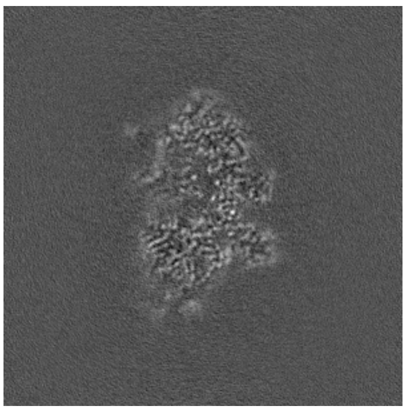
6.2.1 Primary map

X Index: 179

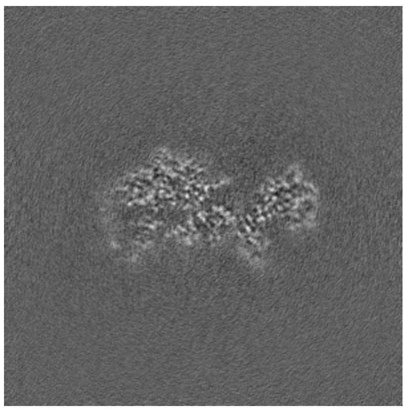
Y Index: 179

Z Index: 179

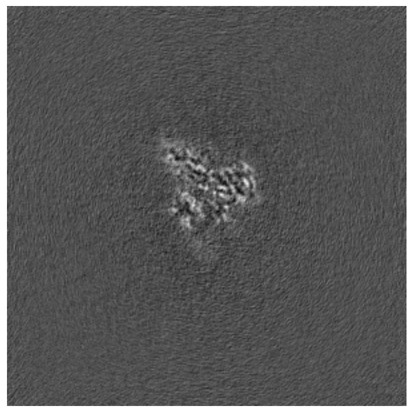
6.2.2 Raw map



X Index: 180



Y Index: 180

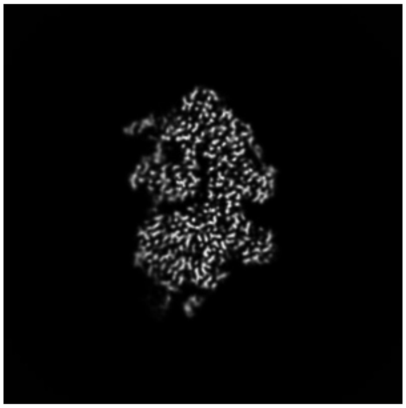


Z Index: 180

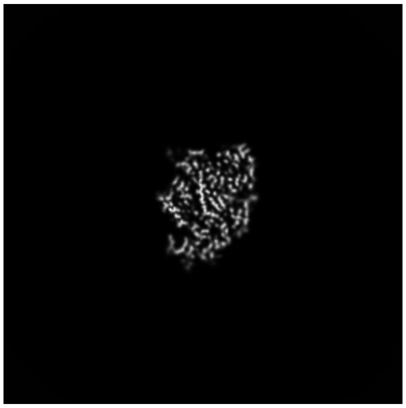
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

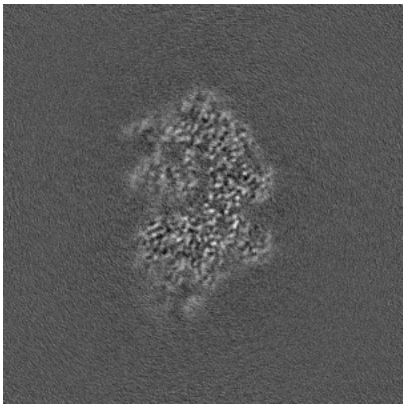
6.3.1 Primary map

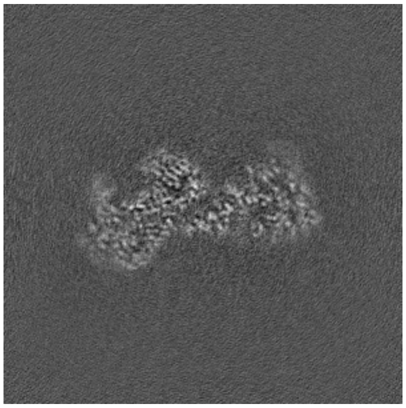

X Index: 183

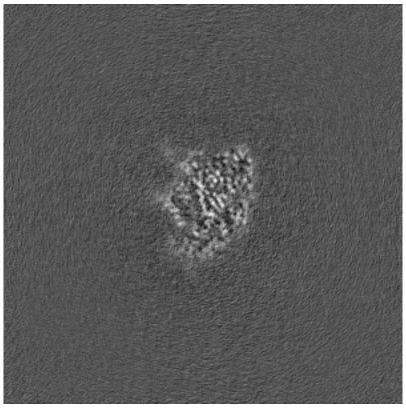

Y Index: 166


Z Index: 164

6.3.2 Raw map


X Index: 184

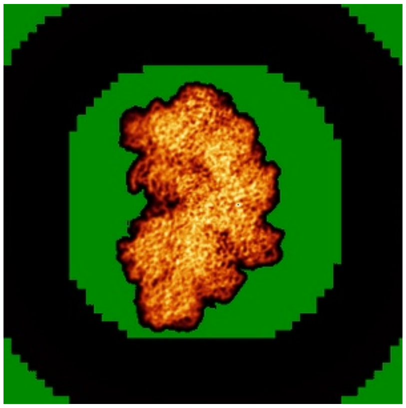

Y Index: 166

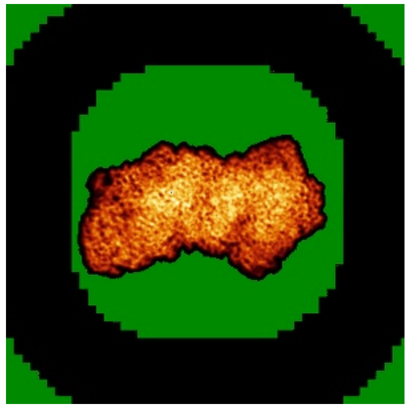

Z Index: 165

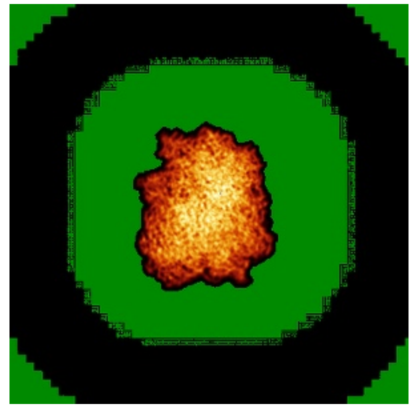
The images above show the largest variance slices of the map in three orthogonal directions.

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

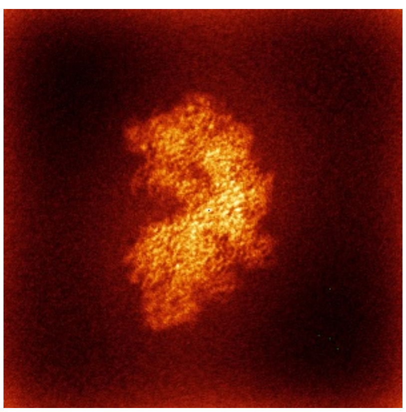
6.4.1 Primary map

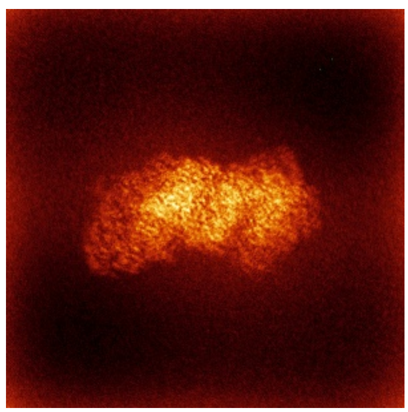

X

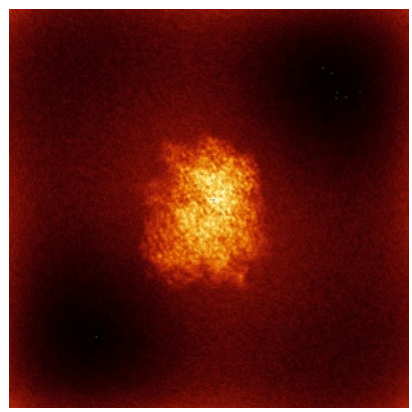

Y


Z

6.4.2 Raw map


X


Y

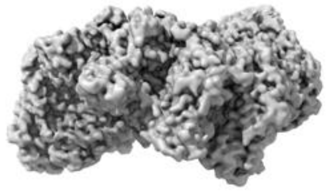

Z

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.

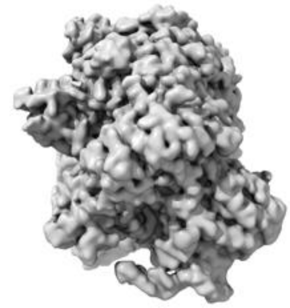
6.5 Orthogonal surface views (i)

6.5.1 Primary map

X



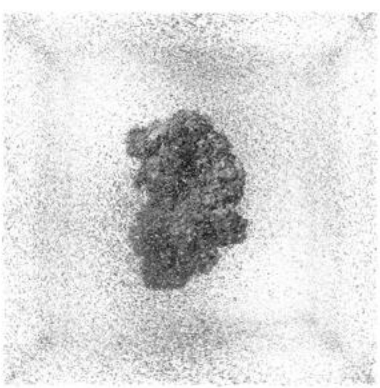
Y



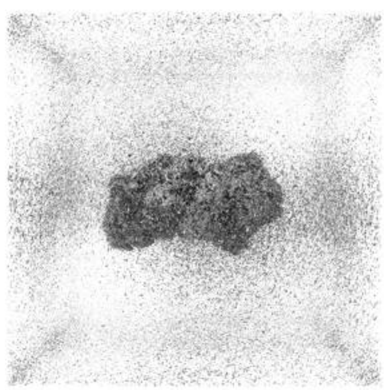
Z

The images above show the 3D surface view of the map at the recommended contour level 1.04. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

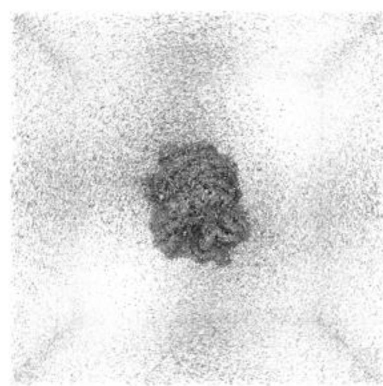
6.5.2 Raw map



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.

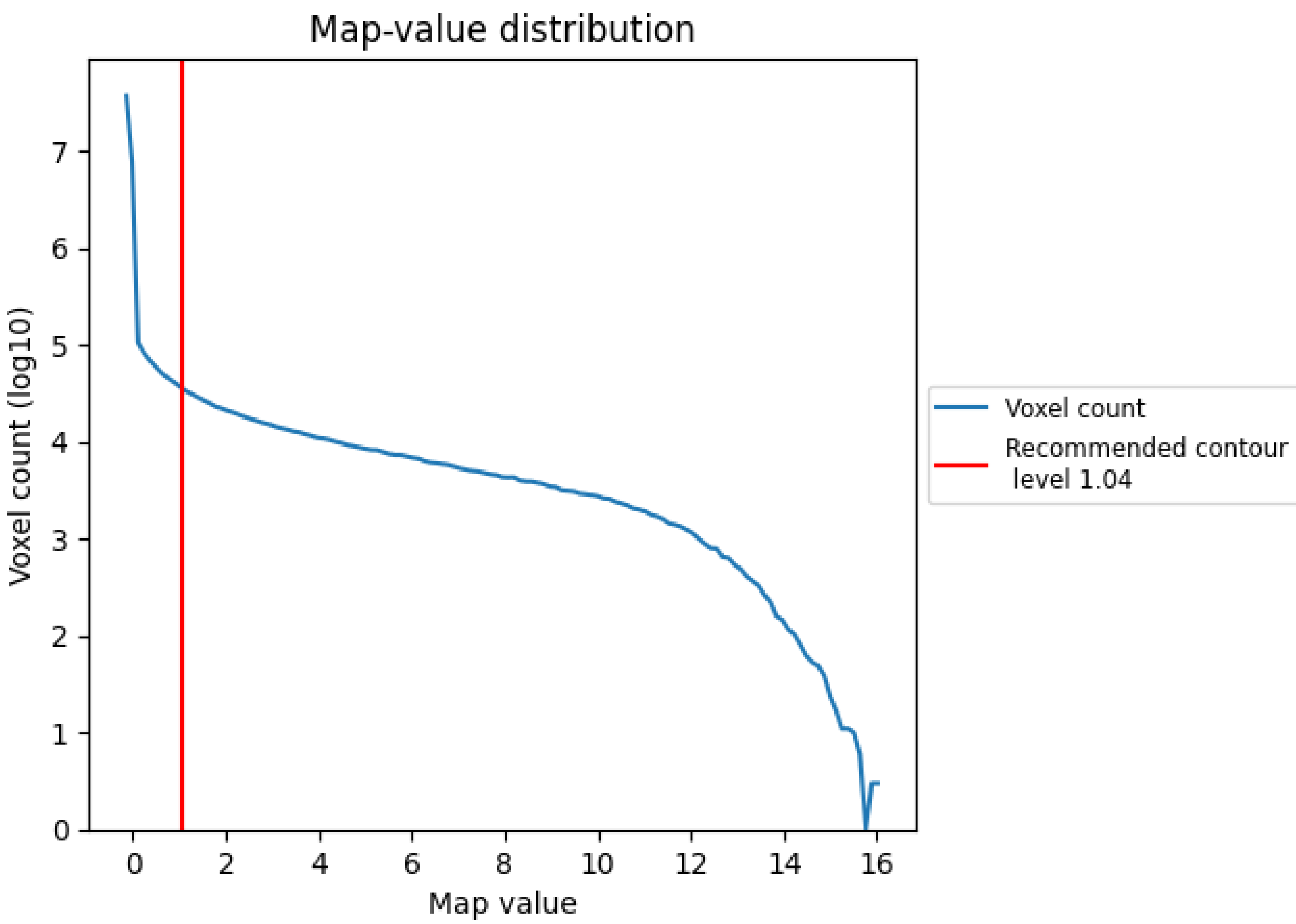
6.6 Mask visualisation (i)

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

7 Map analysis (i)

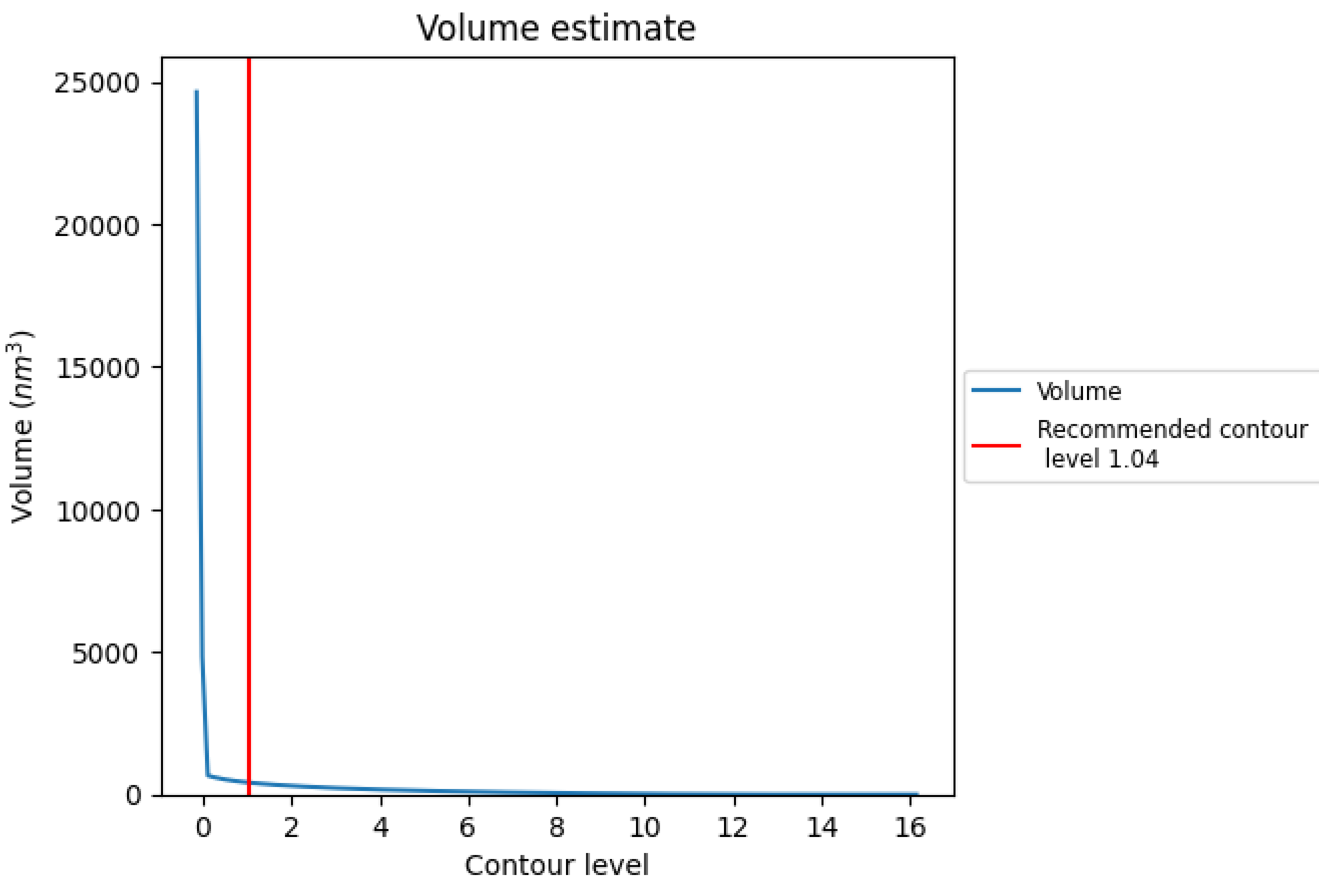
This section contains the results of statistical analysis of the map.

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.

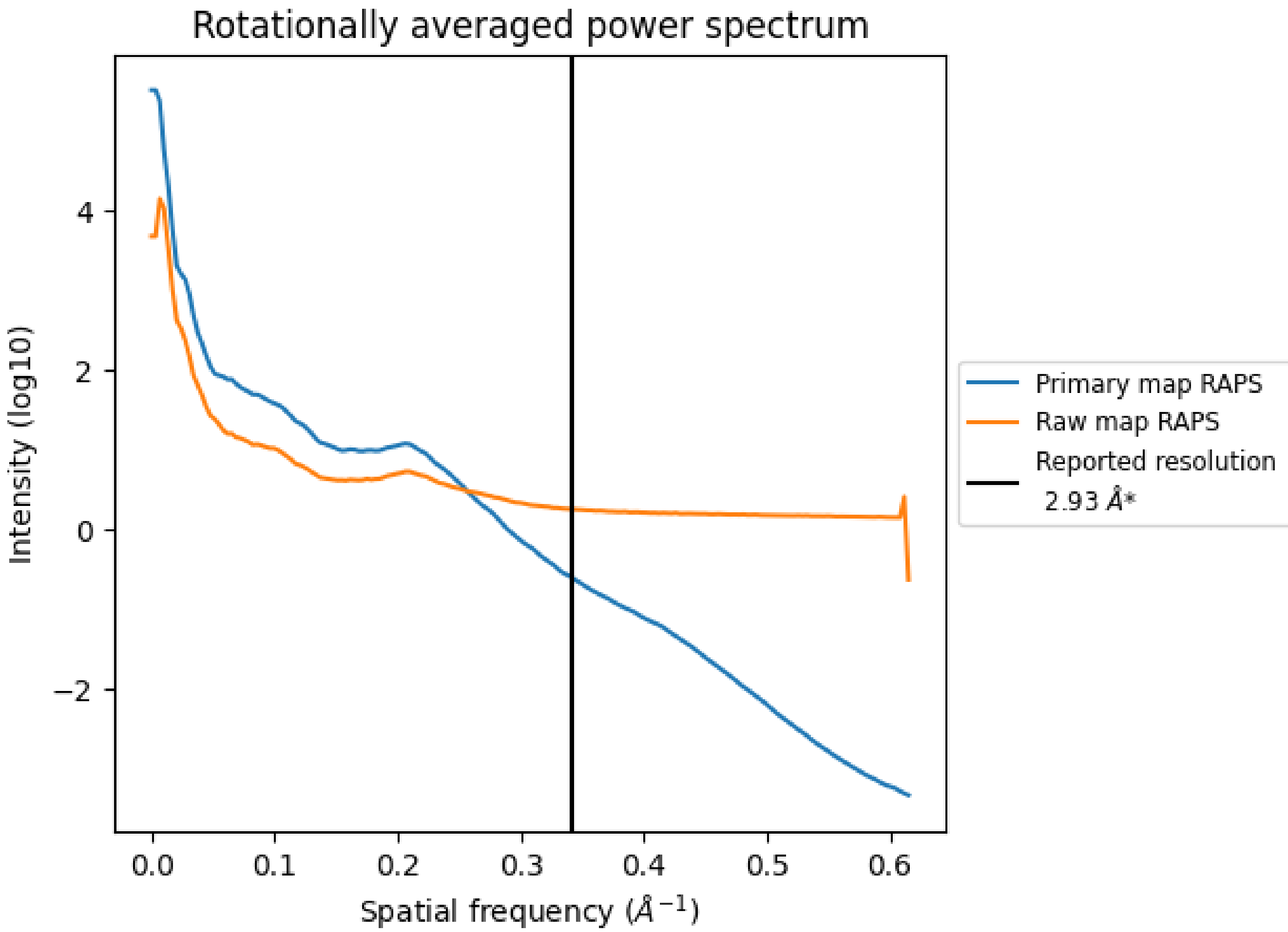
7.2 Volume estimate (i)



The volume at the recommended contour level is 415 nm^3; this corresponds to an approximate mass of 375 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 ⓘ

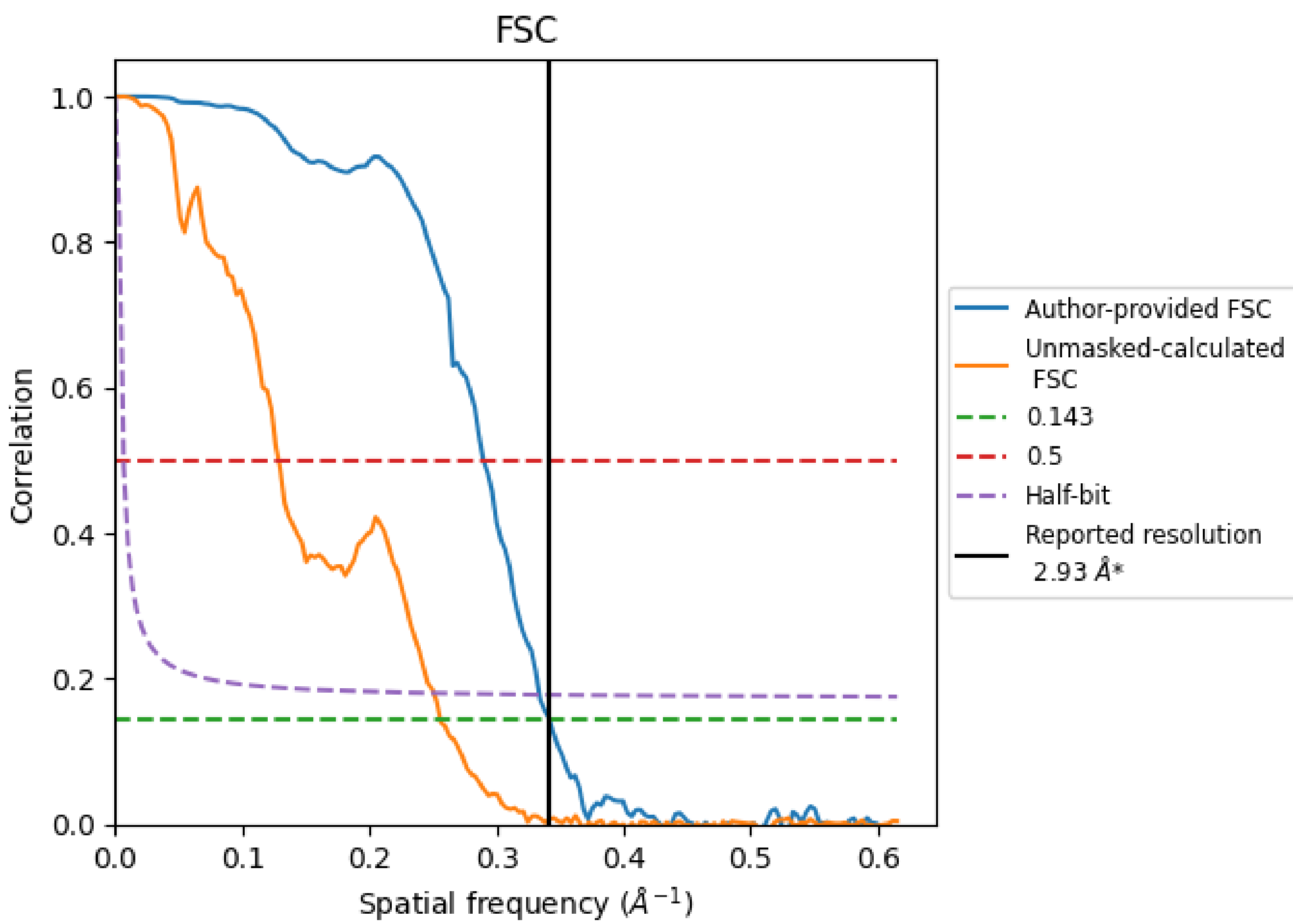


*Reported resolution corresponds to spatial frequency of 0.341 $Å^{-1}$

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 0.341 $Å^{-1}$

8.2 Resolution estimates (i)

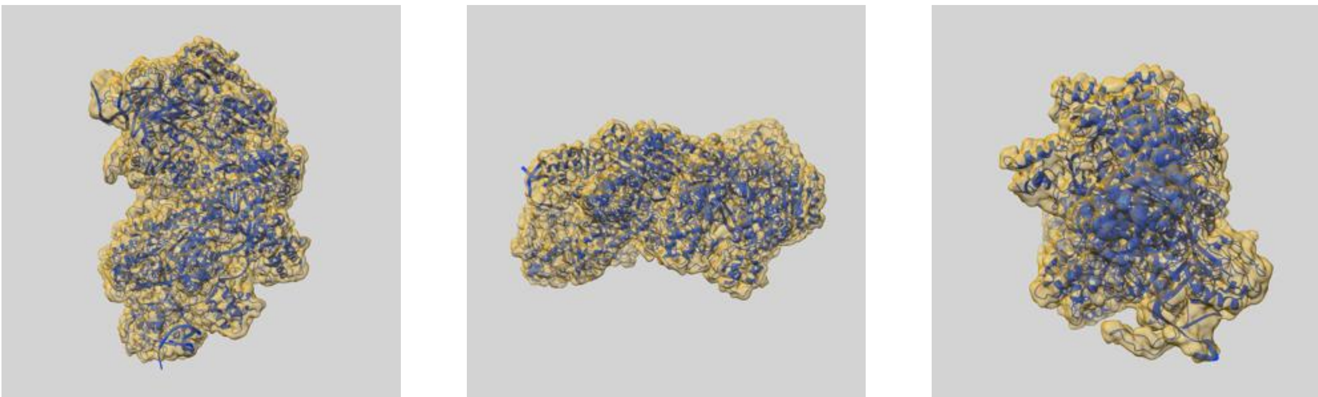
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.93	-	-
Author-provided FSC curve	2.93	3.45	3.00
Unmasked-calculated*	3.91	7.78	3.99

*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 3.91 differs from the reported value 2.93 by more than 10 %

9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-61880 and PDB model 9JXS. Per-residue inclusion information can be found in section 3 on page 6.

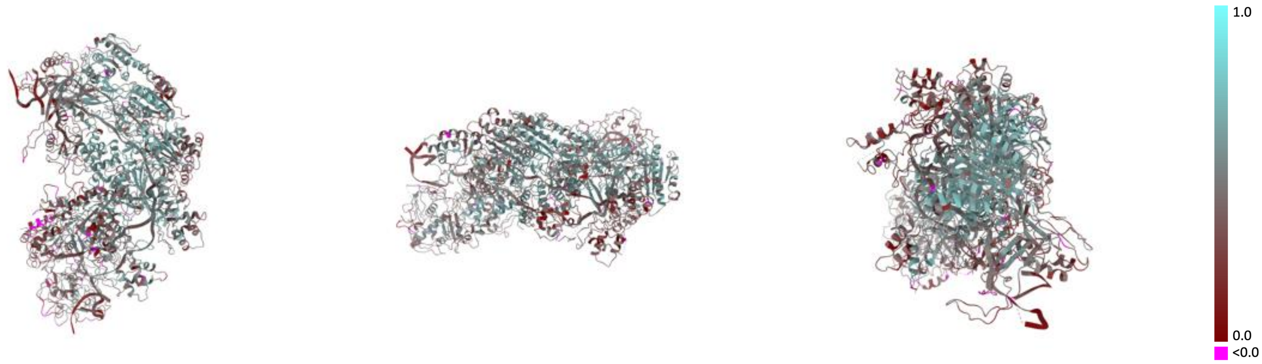
9.1 Map-model overlay (i)



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 1.04 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.

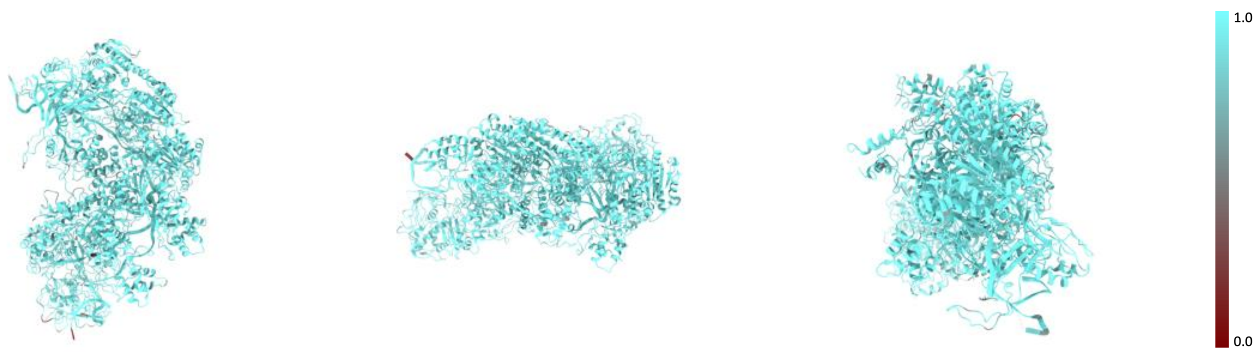
9.2 Q-score mapped to coordinate model (i)



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.

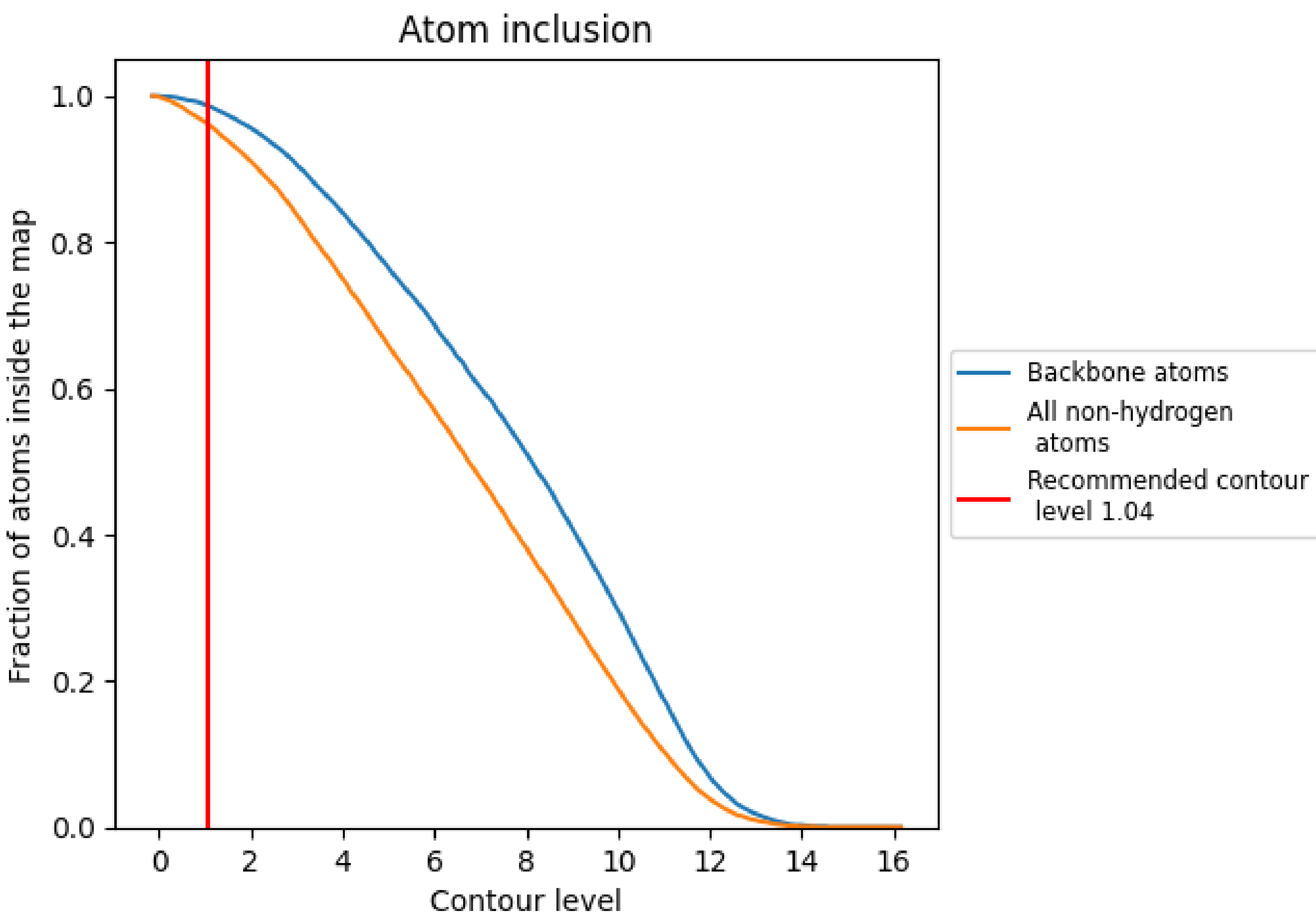
9.3 Atom inclusion mapped to coordinate model (i)



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 (1.04).

9.4 Atom inclusion ⓘ



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 (i)

The table lists the average atom inclusion at the recommended contour level (1.04) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9630	0.4460
A	0.9650	0.3660
B	0.9640	0.4070
C	0.9620	0.4280
D	0.9470	0.3810
E	0.9840	0.5960
F	0.9640	0.4470
G	0.9500	0.3950
H	0.9690	0.5060
I	0.9640	0.4650
J	0.9660	0.4440
K	0.9720	0.5050
M	0.9590	0.3980
N	0.8150	0.2420