wwPDB EM Validation Summary Report ⓘ

Feb 26, 2024 – 03:29 PM EST

PDB ID	:	8G9T
EMDB ID	:	EMD-29878
Title	:	Exploiting Activation and Inactivation Mechanisms in Type I-C CRISPR-Cas3 for Genome Editing Applications
Authors	:	Hu, C.; Nam, K.H.; Ke, A.
Deposited on	:	2023-02-22
Resolution	:	3.60 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

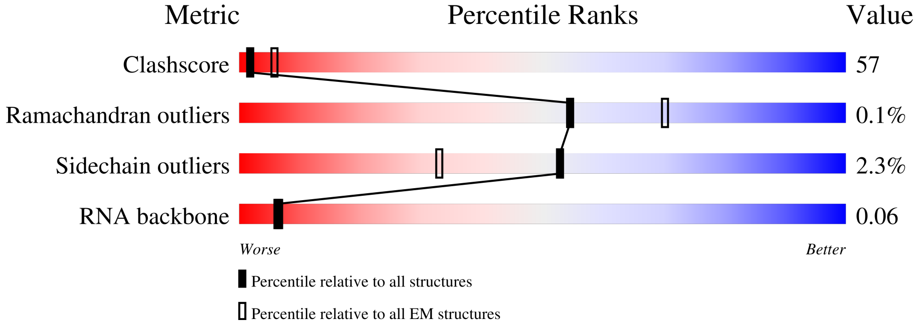
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

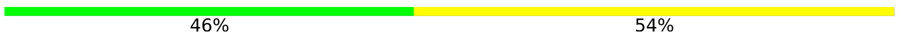
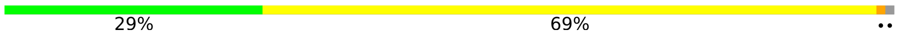
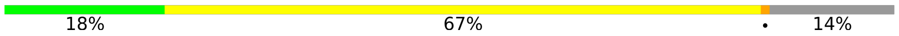
The reported resolution of this entry is 3.60 Å.

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.



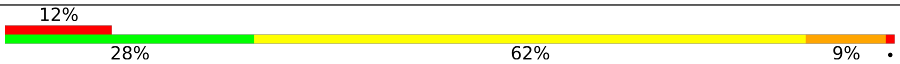
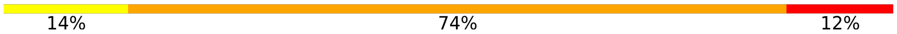
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	70	
2	B	283	
2	C	283	
2	D	283	
2	E	283	
2	F	283	
2	H	283	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	283	
3	G	124	
3	I	124	
3	J	124	
3	L	124	
4	K	582	
5	N	205	
6	O	43	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 26804 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 AcrIC9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	70	583	370	92	119	2	0	0

- Molecule 2 is a protein called Cas7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	275	2206	1379	397	419	11	0	0
2	C	280	2241	1397	403	430	11	0	0
2	D	276	2214	1383	399	421	11	0	0
2	E	278	2225	1389	400	425	11	0	0
2	F	261	2097	1314	375	399	9	0	0
2	H	242	1955	1229	348	370	8	0	0
2	M	263	2101	1312	377	401	11	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	284	GLY	-	expression tag	UNP A0A378VEU0
C	284	GLY	-	expression tag	UNP A0A378VEU0
D	284	GLY	-	expression tag	UNP A0A378VEU0
E	284	GLY	-	expression tag	UNP A0A378VEU0
F	284	GLY	-	expression tag	UNP A0A378VEU0
H	284	GLY	-	expression tag	UNP A0A378VEU0
M	284	GLY	-	expression tag	UNP A0A378VEU0

- Molecule 3 is a protein called Cas11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	G	124	1007	646	180	179	2	0	0
3	I	124	1007	646	180	179	2	0	0
3	J	124	1007	646	180	179	2	0	0
3	L	124	1007	646	180	179	2	0	0

- Molecule 4 is a protein called Cas8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	K	582	4559	2894	800	849	16	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	190	ALA	VAL	conflict	UNP A0A378VF47
K	239	ALA	ILE	conflict	UNP A0A378VF47
K	242	ILE	VAL	conflict	UNP A0A378VF47
K	260	GLY	SER	conflict	UNP A0A378VF47
K	271	THR	ALA	conflict	UNP A0A378VF47

- Molecule 5 is a protein called Cas5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	N	205	1674	1067	289	306	12	0	0

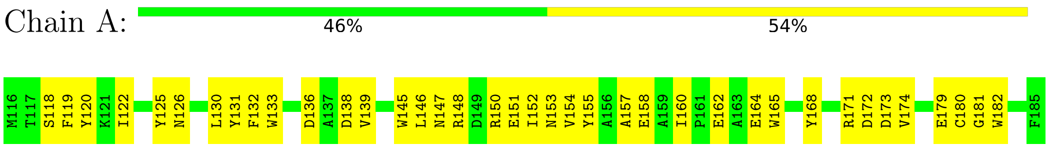
- Molecule 6 is a RNA chain called crRNA (43-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	O	43	921	411	169	298	43	0	0

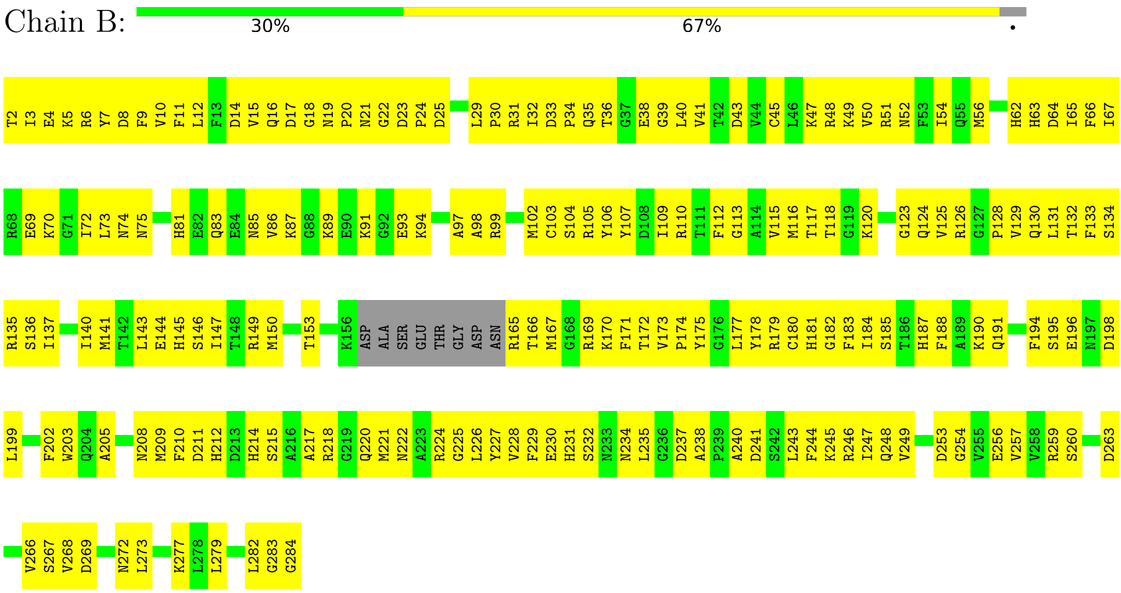
3 Residue-property plots ⓘ

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.

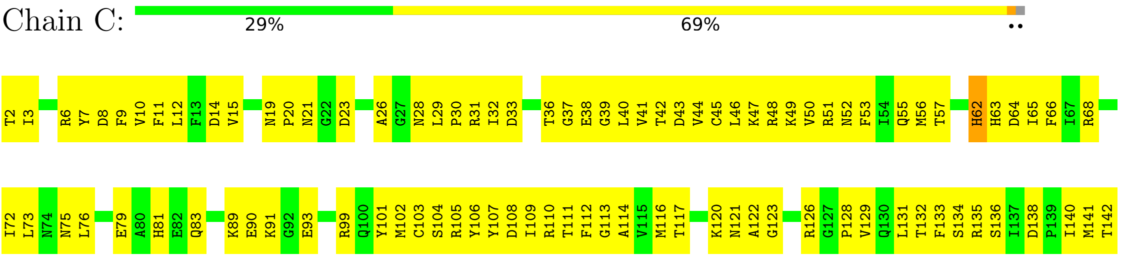
- Molecule 1: AcrIC9

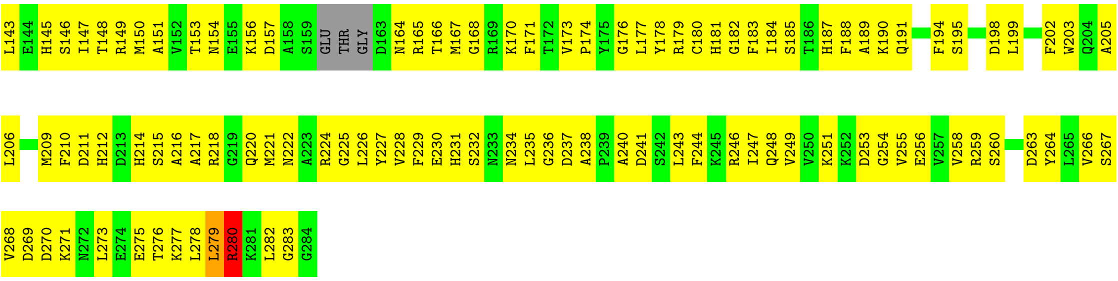


- Molecule 2: Cas7



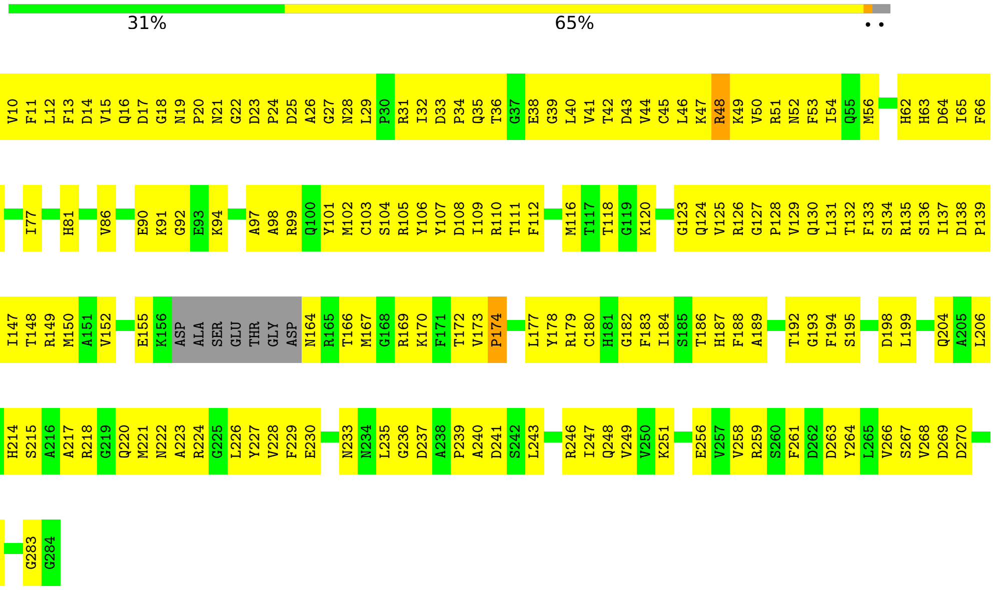
- Molecule 2: Cas7





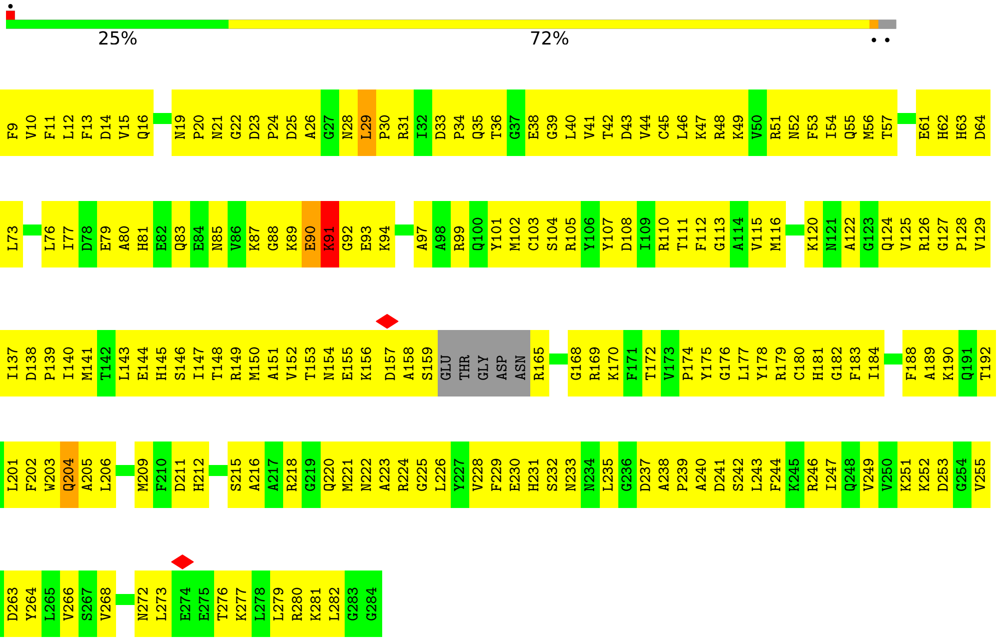
- Molecule 2: Cas7

Chain D:

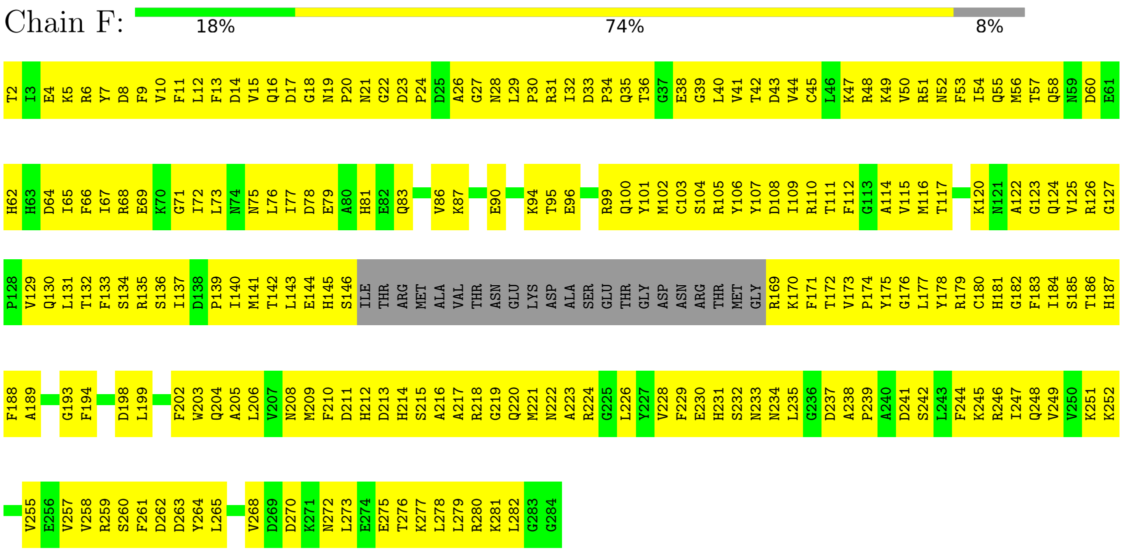


- Molecule 2: Cas7

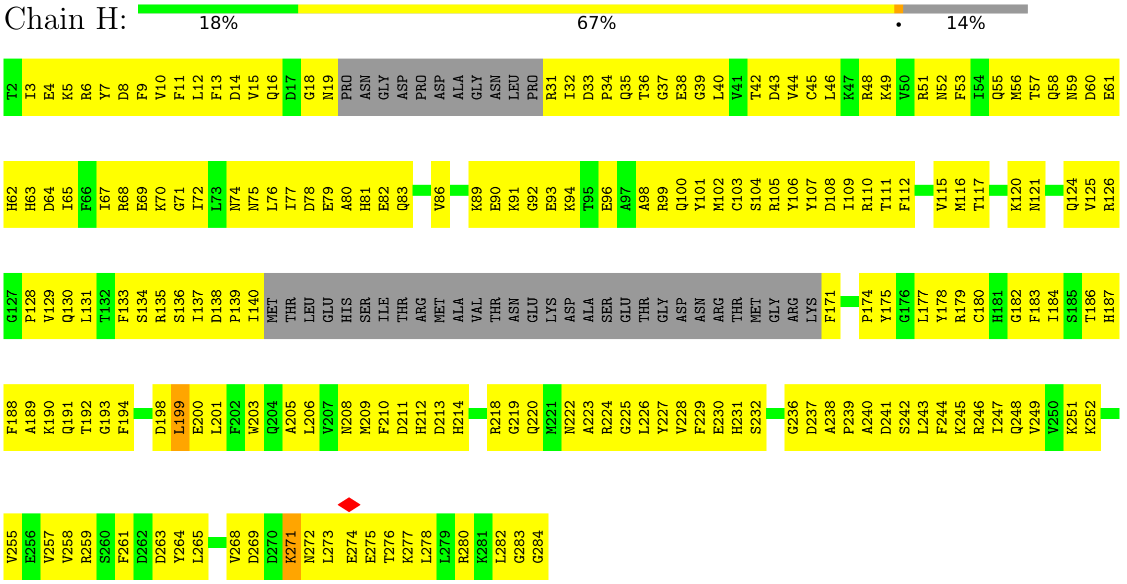
Chain E:



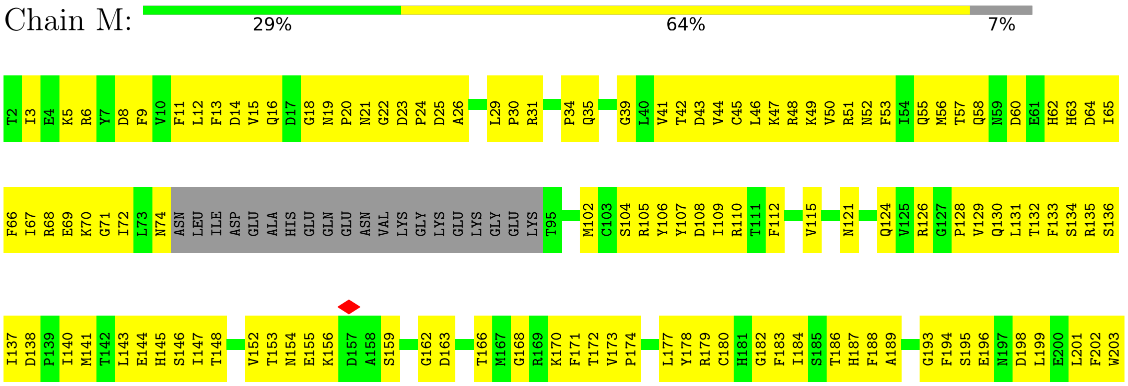
- Molecule 2: Cas7

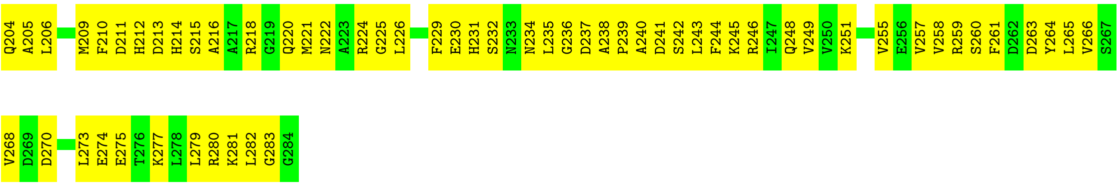


- Molecule 2: Cas7



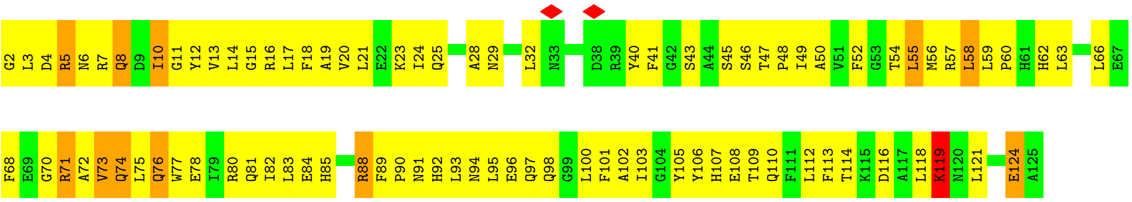
- Molecule 2: Cas7





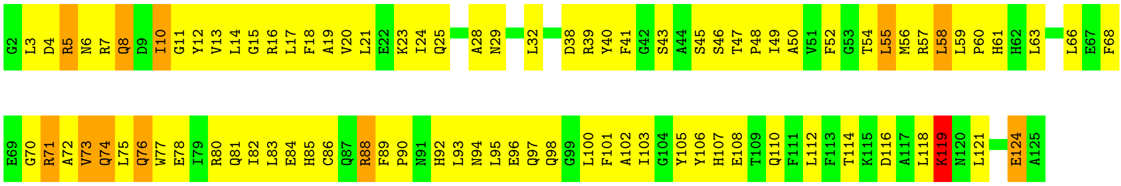
- Molecule 3: Cas11

Chain G:



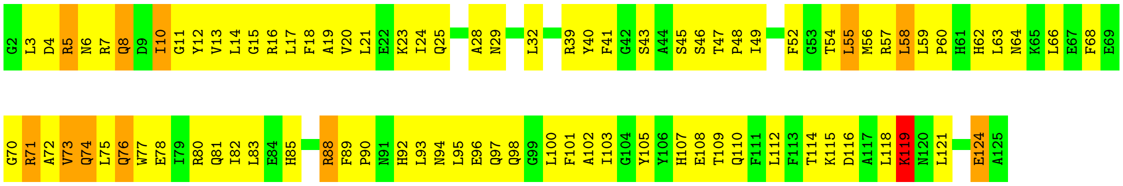
- Molecule 3: Cas11

Chain I:



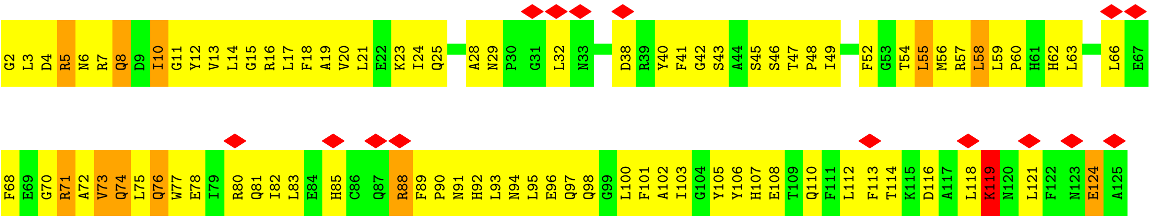
- Molecule 3: Cas11

Chain J:

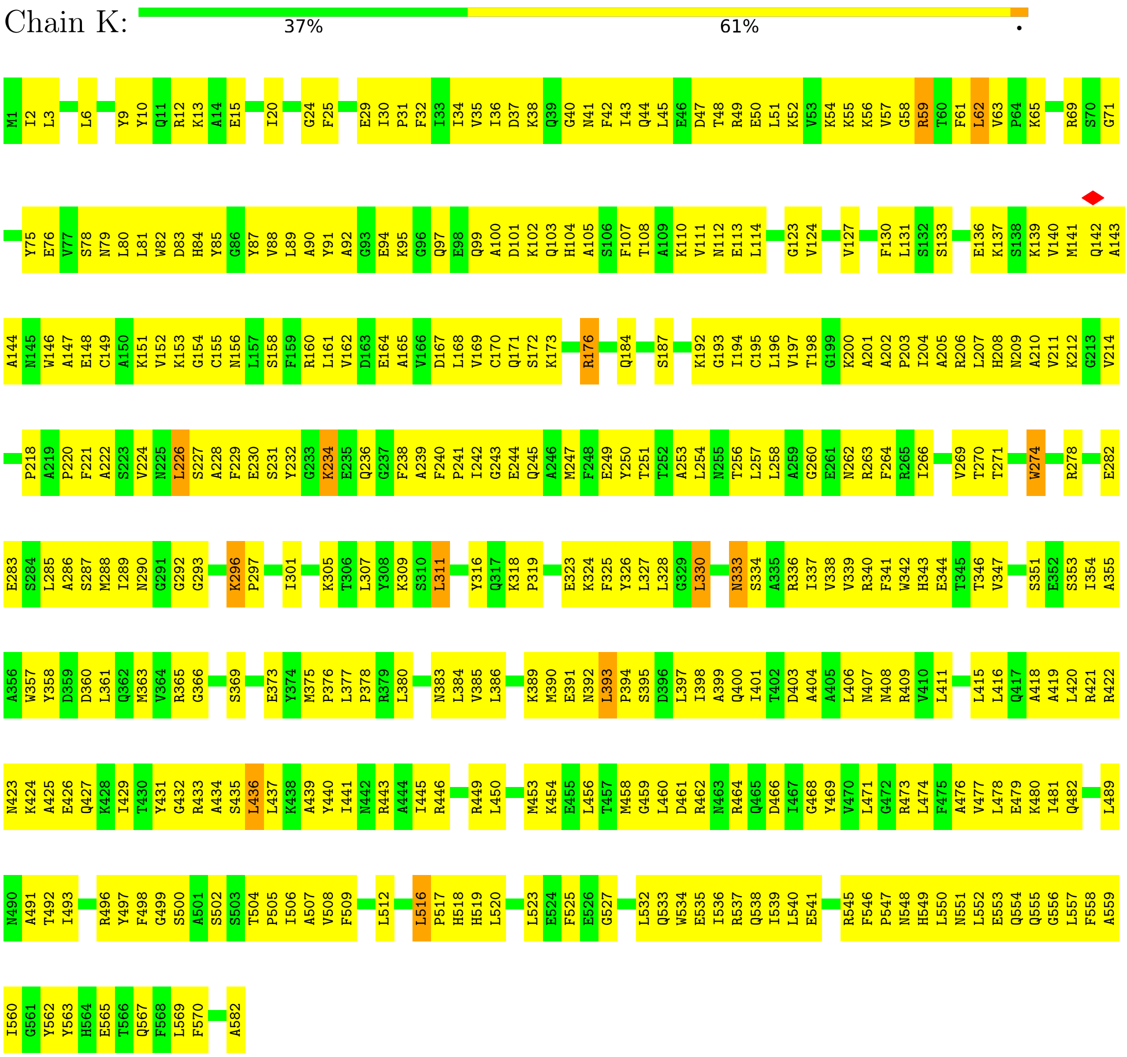


- Molecule 3: Cas11

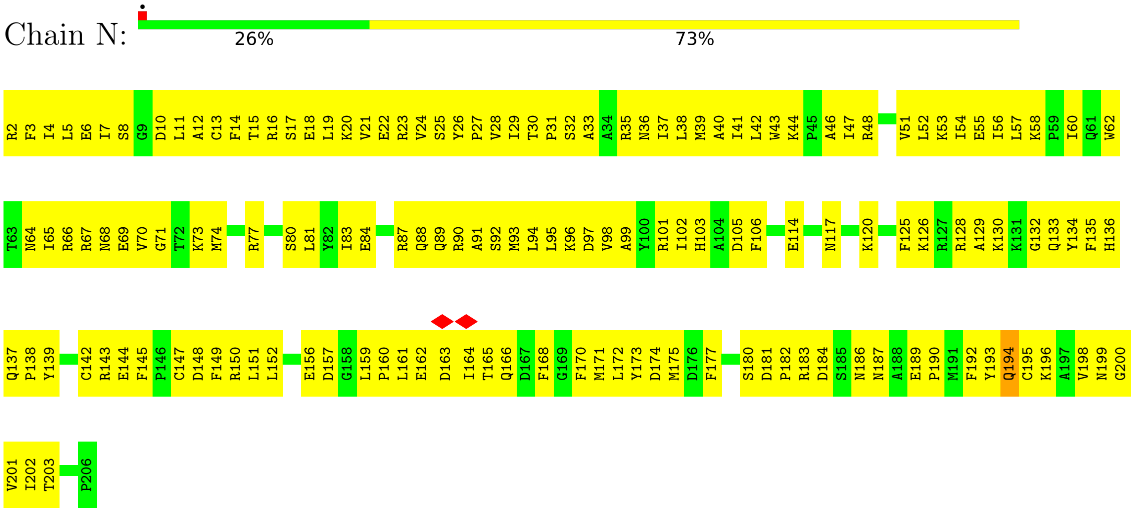
Chain L:



- Molecule 4: Cas8

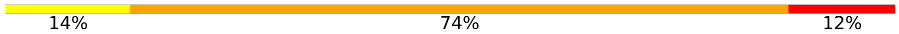


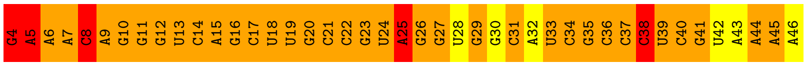
- Molecule 5: Cas5



- Molecule 6: crRNA (43-MER)

Chain O:


14%
74%
12%


G4
A5
A6
A7
C8
A9
G10
G11
G12
U13
C14
A15
G16
C17
U18
U19
G20
C21
C22
G23
U24
A25
G26
G27
U28
G29
G30
C31
A32
U33
C34
G35
C36
C37
C38
U39
C40
G41
U42
A43
A44
A45
A46

4 Experimental information ⓘ

Property	**Value**	**Source**
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	200000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose $(e^-/\text{Å}^2)$	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	67000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	2.613	Depositor
Minimum map value	-1.561	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.060	Depositor
Recommended contour level	0.165	Depositor
Map size (Å)	381.184, 381.184, 381.184	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.489, 1.489, 1.489	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	0.66	0/600	0.59	0/818
2	B	0.75	0/2247	0.68	0/3023
2	C	0.74	1/2282 (0.0%)	0.70	1/3071 (0.0%)
2	D	0.75	1/2255 (0.0%)	0.75	4/3034 (0.1%)
2	E	0.64	0/2266	0.68	2/3049 (0.1%)
2	F	0.54	0/2138	0.60	0/2878
2	H	0.41	0/1991	0.62	1/2676 (0.0%)
2	M	0.73	0/2141	0.66	0/2885
3	G	0.55	0/1030	0.83	2/1392 (0.1%)
3	I	0.55	0/1030	0.83	2/1392 (0.1%)
3	J	0.55	0/1030	0.83	2/1392 (0.1%)
3	L	0.55	0/1030	0.83	2/1392 (0.1%)
4	K	0.60	1/4648 (0.0%)	0.70	8/6281 (0.1%)
5	N	0.68	0/1712	0.69	0/2307
6	O	1.40	4/1030 (0.4%)	1.69	33/1604 (2.1%)
All	All	0.68	7/27430 (0.0%)	0.77	57/37194 (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
4	K	0	2
5	N	0	1
All	All	0	3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	48	ARG	CB-CG	-8.94	1.28	1.52
6	O	4	G	N7-C5	-7.39	1.34	1.39
6	O	4	G	C6-N1	-7.27	1.34	1.39

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	K	274	TRP	CB-CG	-7.09	1.37	1.50
6	O	4	G	C8-N7	-5.78	1.27	1.30

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	48	ARG	CG-CD-NE	-11.72	87.19	111.80
6	O	4	G	C8-N9-C1'	-11.53	112.01	127.00
6	O	4	G	C4-N9-C1'	11.38	141.30	126.50
6	O	4	G	C6-C5-N7	-11.29	123.63	130.40
6	O	4	G	N1-C2-N2	-10.43	106.81	116.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	K	234	LYS	Peptide
4	K	333	ASN	Peptide
5	N	194	GLN	Peptide

5.2 Too-close contacts ⓘ

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	583	0	508	57	0
2	B	2206	0	2152	269	0
2	C	2241	0	2176	306	0
2	D	2214	0	2158	296	0
2	E	2225	0	2166	303	0
2	F	2097	0	2034	311	0
2	H	1955	0	1899	243	0
2	M	2101	0	2037	264	0
3	G	1007	0	992	150	0
3	I	1007	0	992	159	0
3	J	1007	0	992	132	0
3	L	1007	0	992	128	0
4	K	4559	0	4565	456	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	N	1674	0	1663	240	0
6	O	921	0	466	256	0
All	All	26804	0	25792	3011	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 57.

The worst 5 of 3011 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:48:ARG:NH2	6:O:26:G:OP2	1.67	1.27
4:K:473:ARG:HA	4:K:565:GLU:HG3	1.33	1.10
3:G:109:THR:HG22	3:I:80:ARG:HE	1.03	1.07
2:C:205:ALA:O	2:C:209:MET:HB3	1.55	1.04
2:D:206:LEU:HA	2:D:209:MET:HE3	1.42	1.00

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	A	68/70 (97%)	63 (93%)	5 (7%)	0	100	100
2	B	271/283 (96%)	234 (86%)	37 (14%)	0	100	100
2	C	276/283 (98%)	230 (83%)	46 (17%)	0	100	100
2	D	272/283 (96%)	237 (87%)	35 (13%)	0	100	100
2	E	274/283 (97%)	229 (84%)	43 (16%)	2 (1%)	22	61
2	F	257/283 (91%)	223 (87%)	34 (13%)	0	100	100
2	H	236/283 (83%)	205 (87%)	31 (13%)	0	100	100
2	M	259/283 (92%)	211 (82%)	48 (18%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	G	122/124 (98%)	113 (93%)	9 (7%)	0	100	100
3	I	122/124 (98%)	113 (93%)	9 (7%)	0	100	100
3	J	122/124 (98%)	113 (93%)	9 (7%)	0	100	100
3	L	122/124 (98%)	113 (93%)	9 (7%)	0	100	100
4	K	580/582 (100%)	487 (84%)	92 (16%)	1 (0%)	47	79
5	N	203/205 (99%)	163 (80%)	40 (20%)	0	100	100
All	All	3184/3334 (96%)	2734 (86%)	447 (14%)	3 (0%)	54	83

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	91	LYS
2	E	90	GLU
4	K	52	LYS

5.3.2 Protein sidechains ⓘ

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	58/58 (100%)	58 (100%)	0	100	100
2	B	239/245 (98%)	239 (100%)	0	100	100
2	C	243/245 (99%)	240 (99%)	3 (1%)	71	87
2	D	240/245 (98%)	240 (100%)	0	100	100
2	E	241/245 (98%)	238 (99%)	3 (1%)	71	87
2	F	227/245 (93%)	227 (100%)	0	100	100
2	H	211/245 (86%)	209 (99%)	2 (1%)	78	90
2	M	228/245 (93%)	228 (100%)	0	100	100
3	G	104/104 (100%)	91 (88%)	13 (12%)	4	25
3	I	104/104 (100%)	91 (88%)	13 (12%)	4	25
3	J	104/104 (100%)	91 (88%)	13 (12%)	4	25

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	L	104/104 (100%)	91 (88%)	13 (12%)	4	25
4	K	475/477 (100%)	473 (100%)	2 (0%)	91	97
5	N	182/182 (100%)	181 (100%)	1 (0%)	88	95
All	All	2760/2848 (97%)	2697 (98%)	63 (2%)	53	76

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	73	VAL
3	L	58	LEU
3	J	8	GLN
3	L	54	THR
3	L	76	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	52	ASN
3	L	81	GLN
2	M	187	HIS
5	N	36	ASN
2	D	181	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	O	43/43 (100%)	34 (79%)	5 (11%)

5 of 34 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	O	5	A
6	O	6	A
6	O	8	C
6	O	9	A
6	O	10	G

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	O	4	G
6	O	5	A
6	O	9	A
6	O	22	C
6	O	34	C

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.

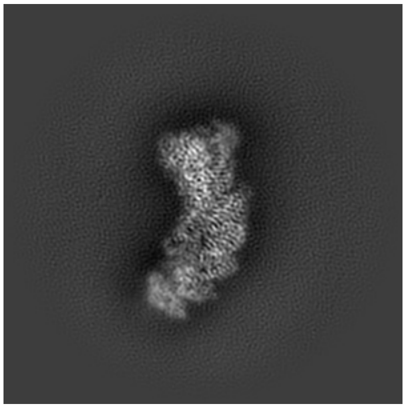
6 Map visualisation ⓘ

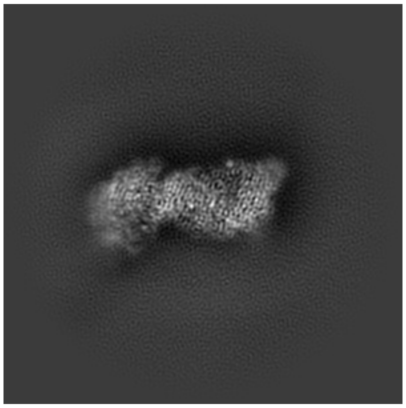
This section contains visualisations of the EMDB entry EMD-29878. These allow visual inspection of the internal detail of the map and identification of artifacts.

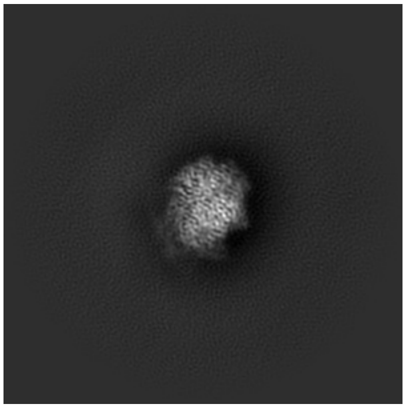
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections ⓘ

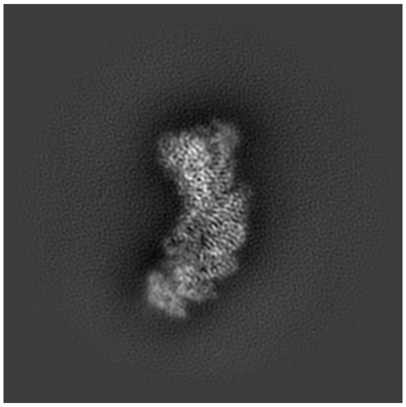
6.1.1 Primary map

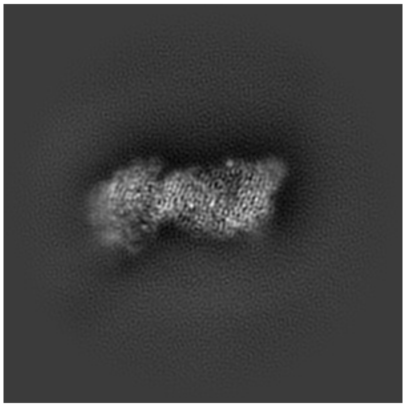

X

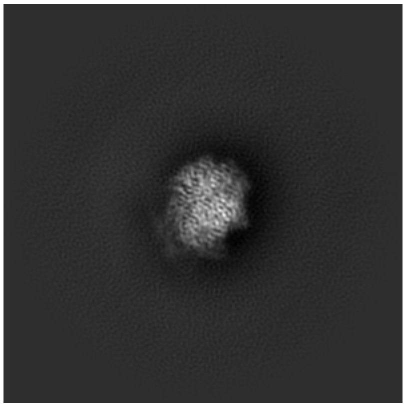

Y


Z

6.1.2 Raw map


X

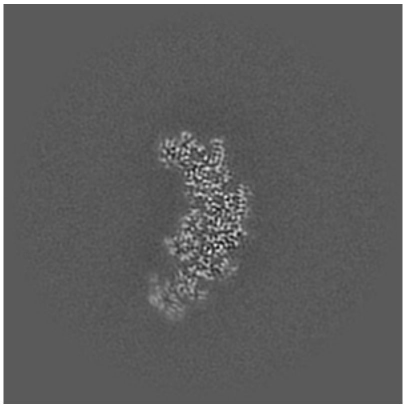

Y

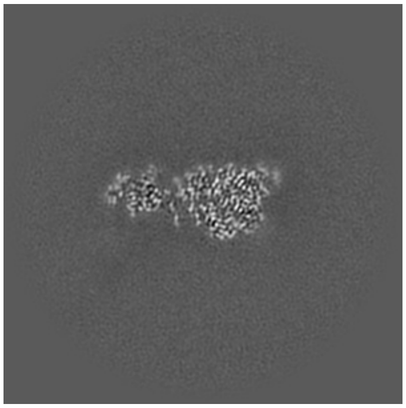

Z

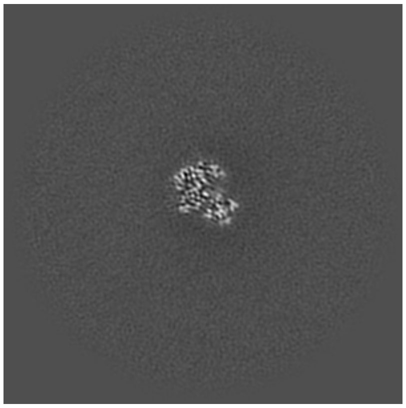
The images above show the map projected in three orthogonal directions.

6.2 Central slices

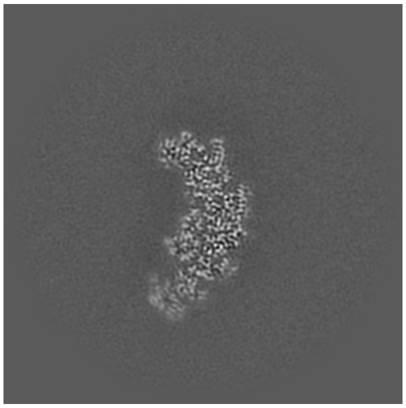
6.2.1 Primary map

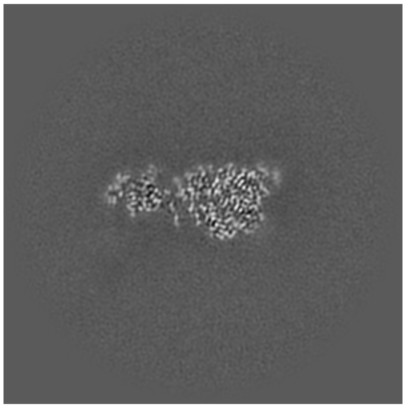

X Index: 128

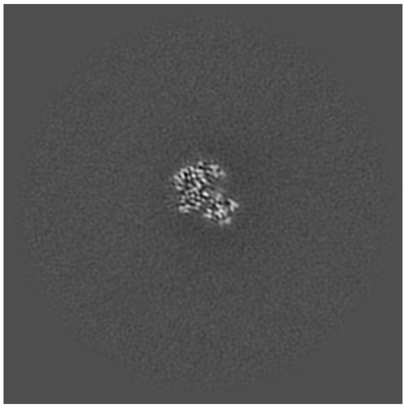

Y Index: 128


Z Index: 128

6.2.2 Raw map


X Index: 128

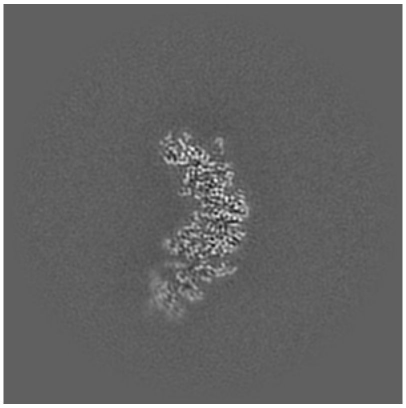

Y Index: 128


Z Index: 128

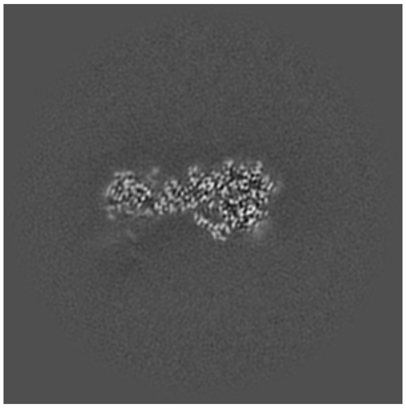
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ

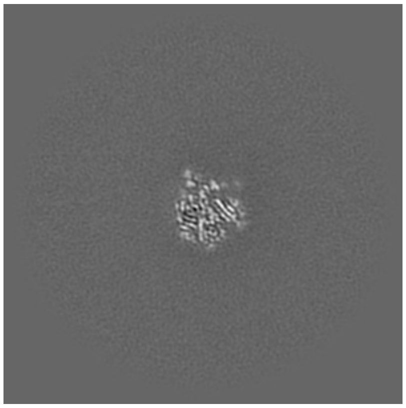
6.3.1 Primary map



X Index: 126

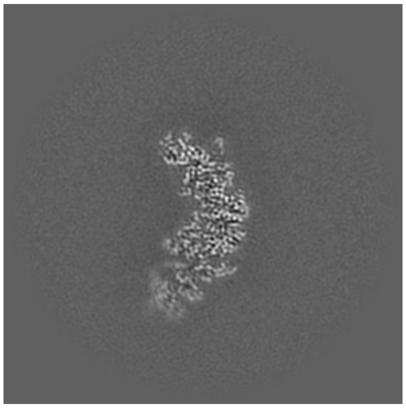


Y Index: 123

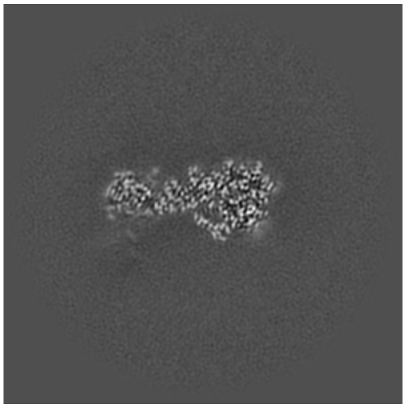


Z Index: 154

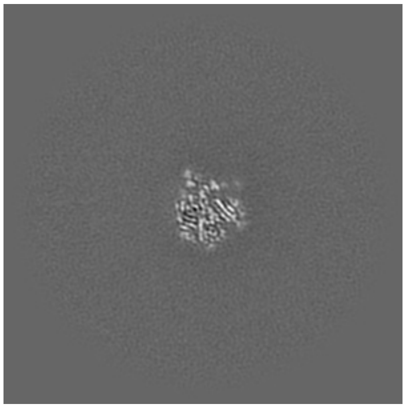
6.3.2 Raw map



X Index: 126



Y Index: 123

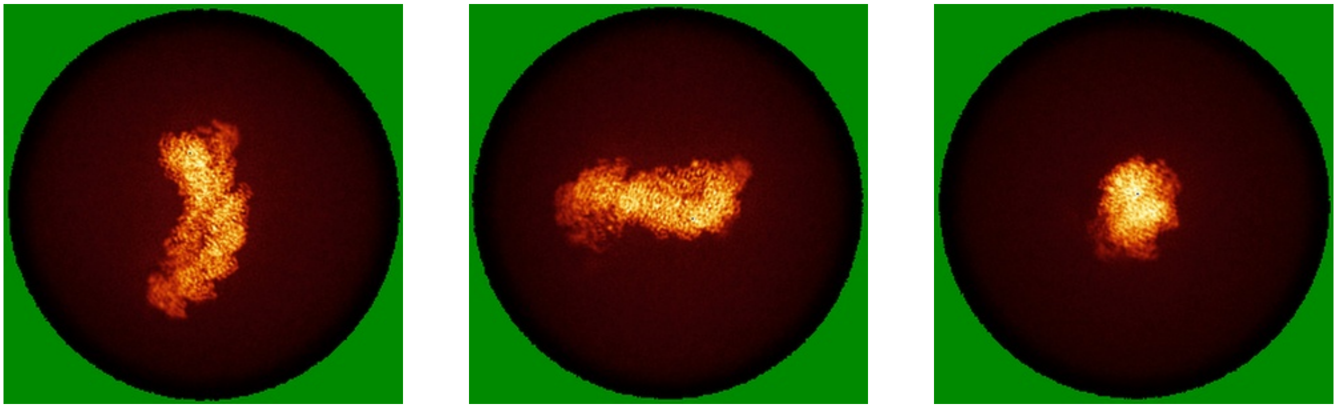


Z Index: 154

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

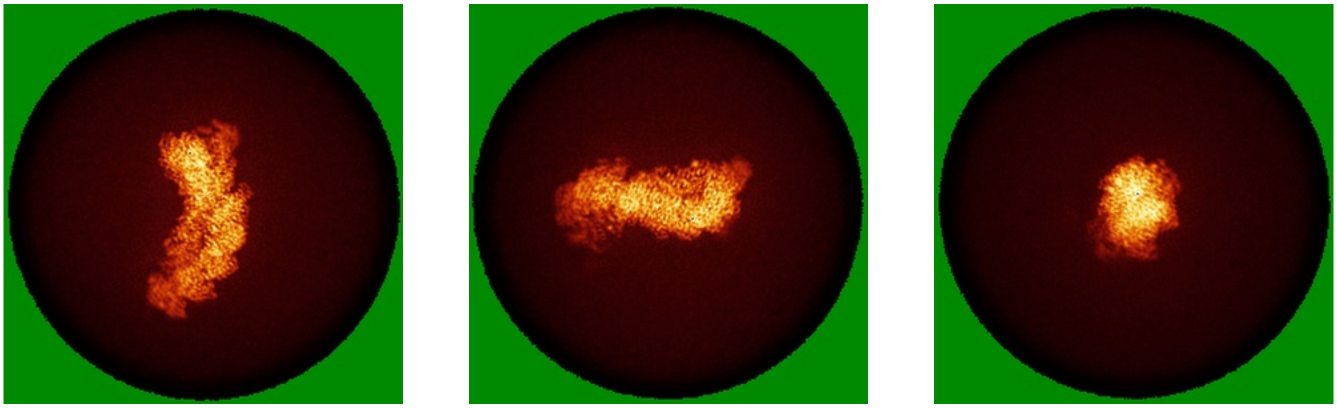
6.4 Orthogonal standard-deviation projections (False-color) ⓘ

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.

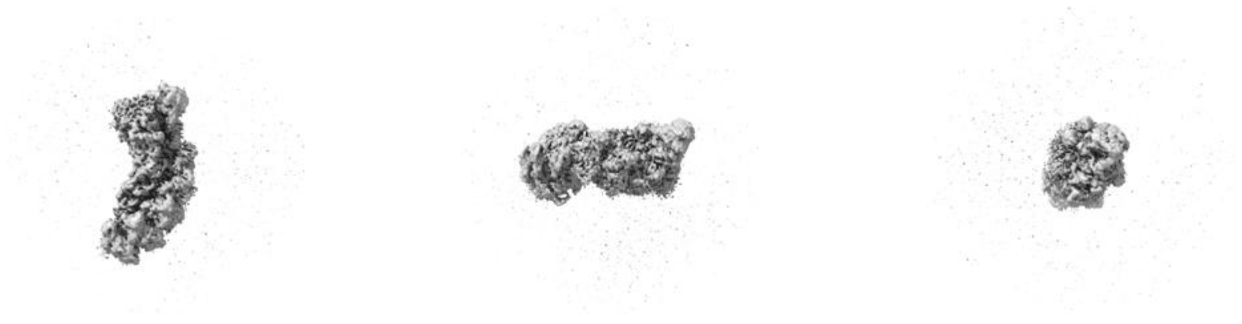
6.5 Orthogonal surface views ⓘ

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.165. 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.

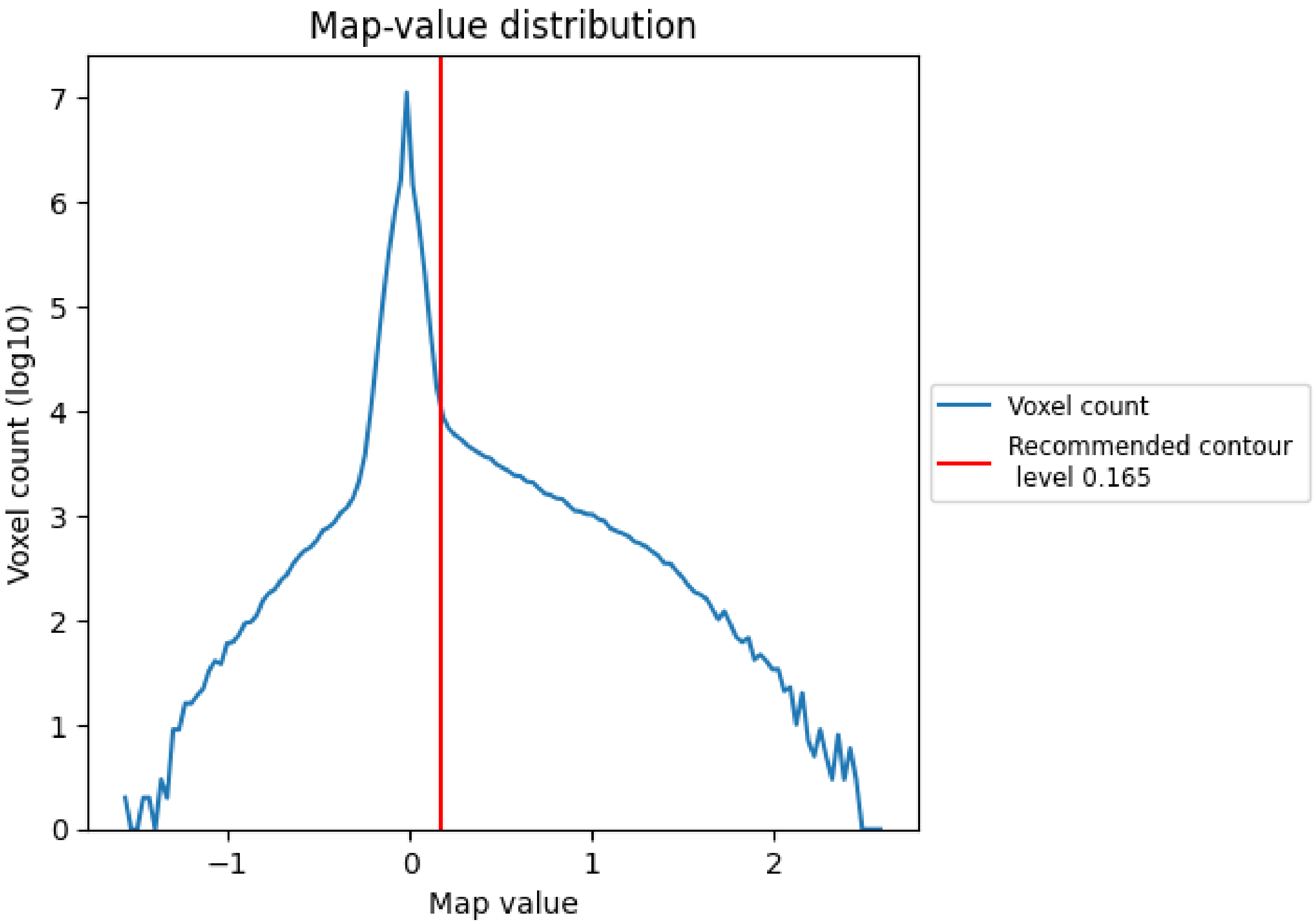
6.6 Mask visualisation ⓘ

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

7 Map analysis ⓘ

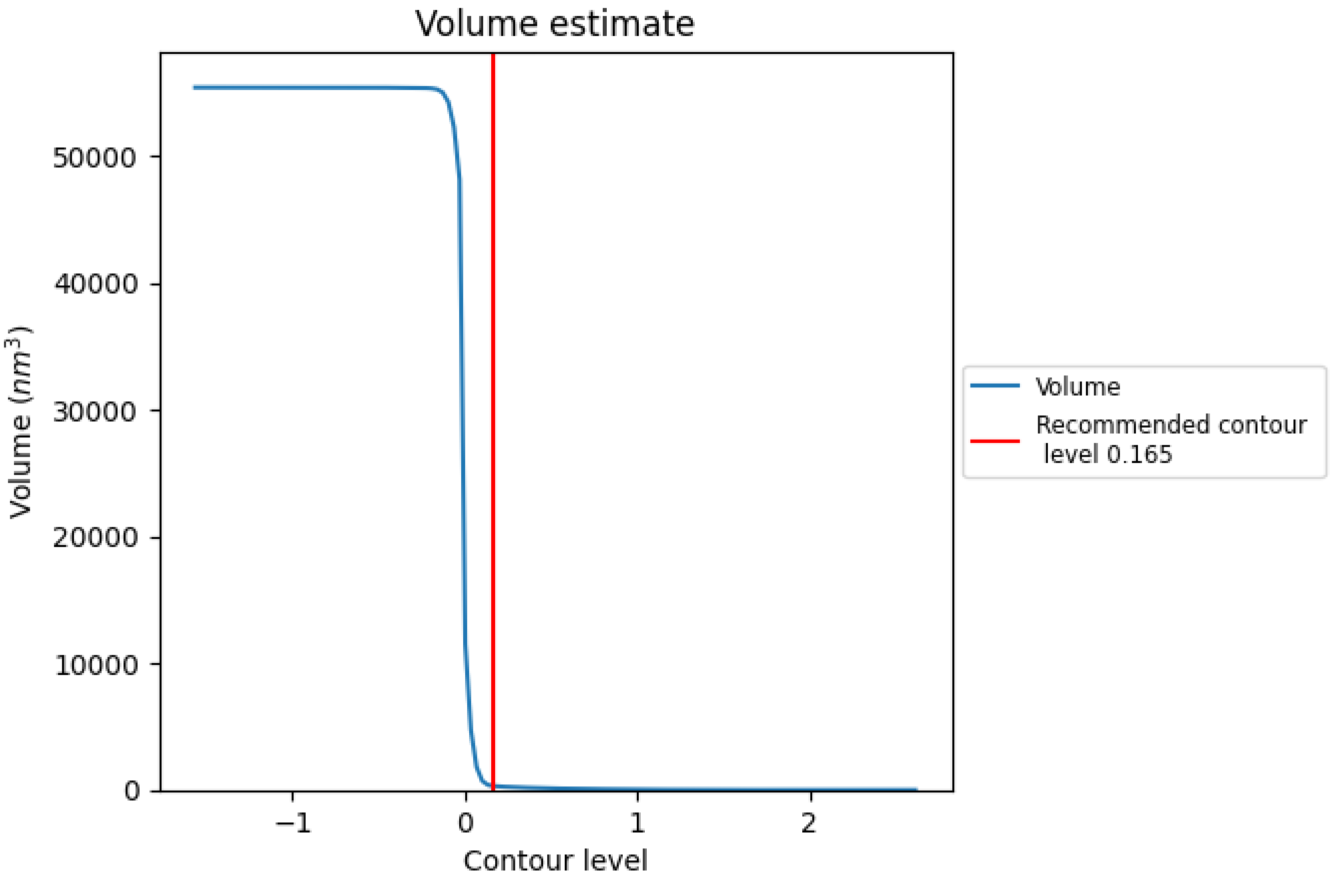
This section contains the results of statistical analysis of the map.

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.

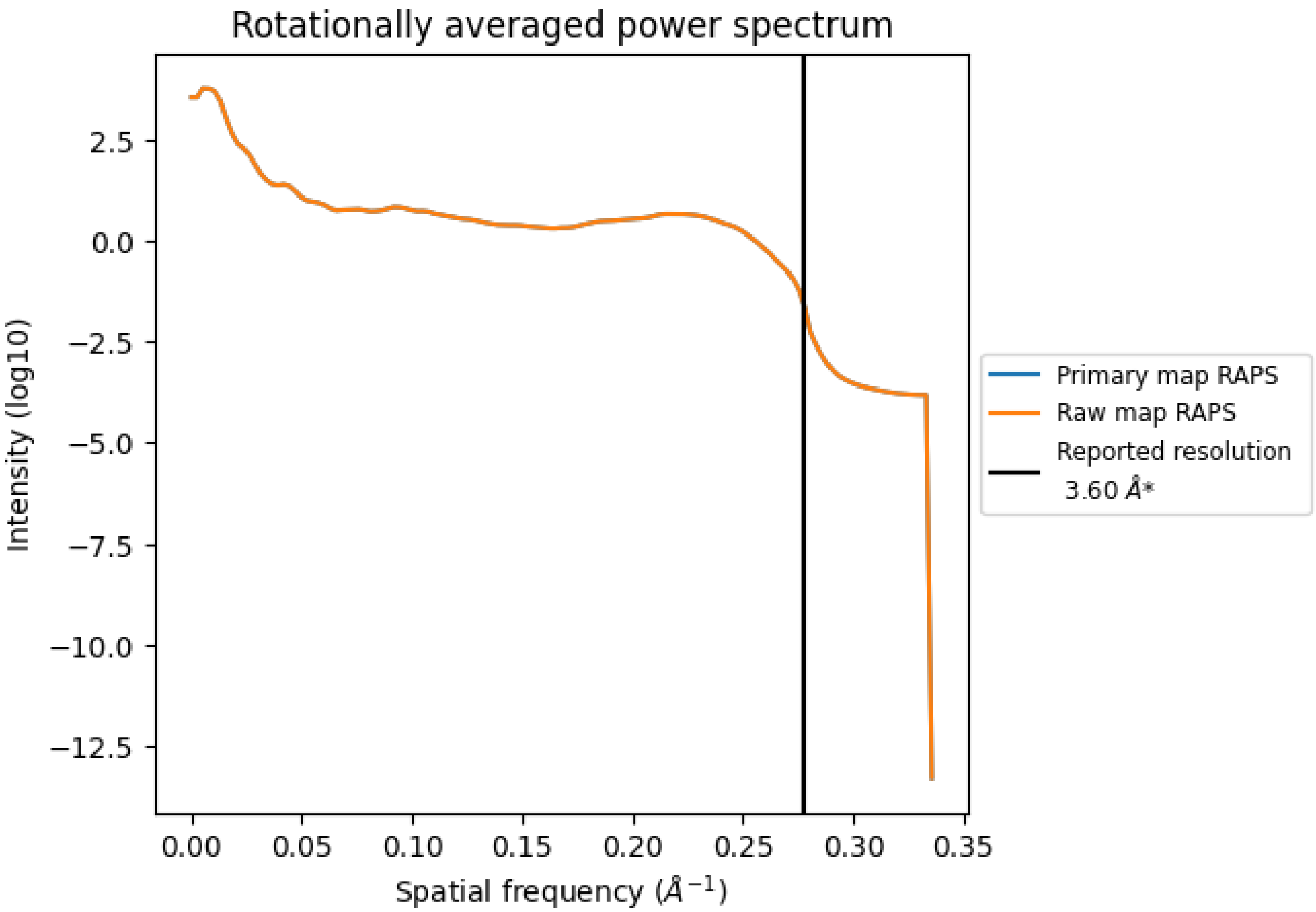
7.2 Volume estimate ⓘ



The volume at the recommended contour level is 316 nm^3; this corresponds to an approximate mass of 286 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 ⓘ

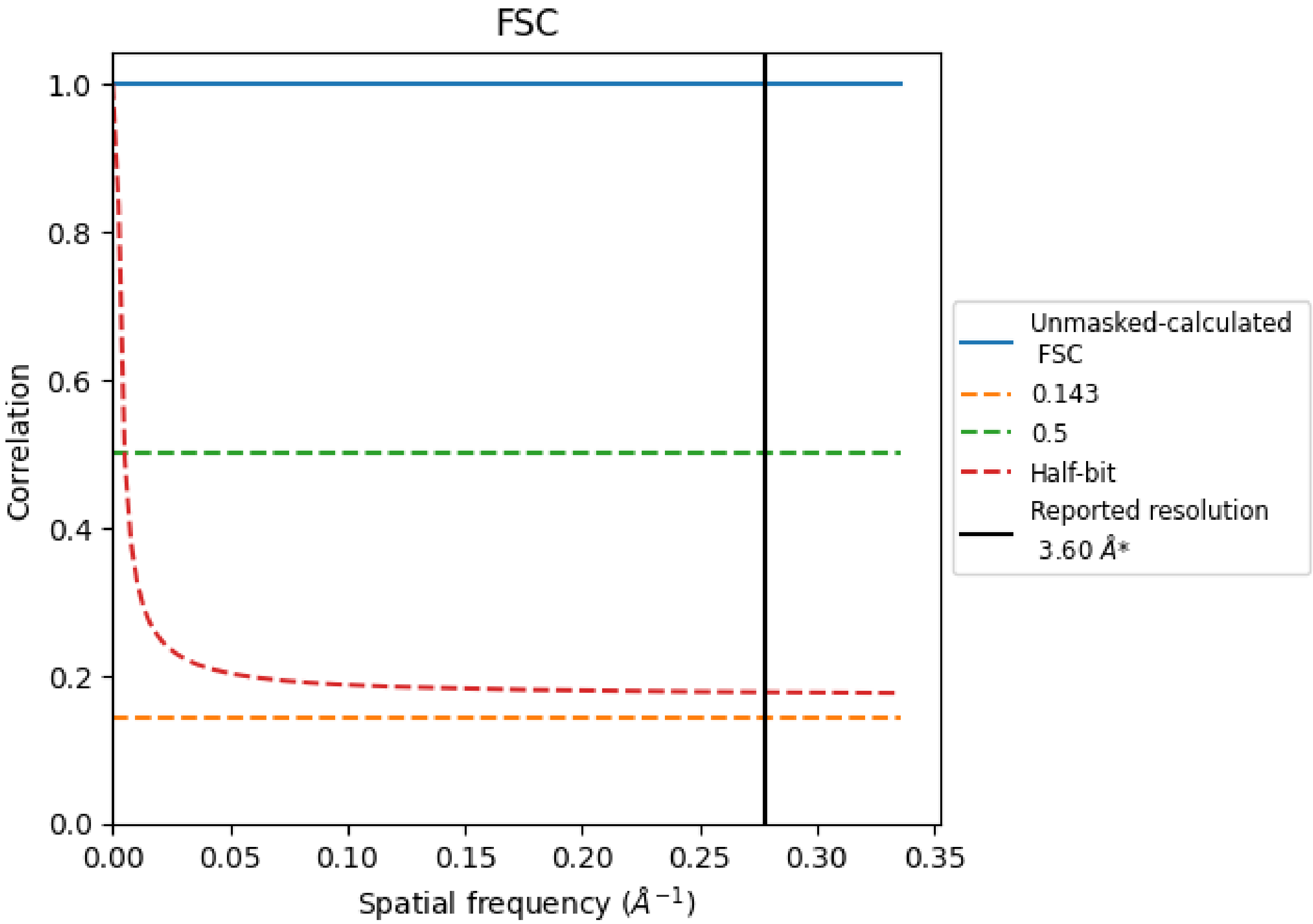


*Reported resolution corresponds to spatial frequency of 0.278 Å^{-1}

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 0.278 $Å^{-1}$

8.2 Resolution estimates ⓘ

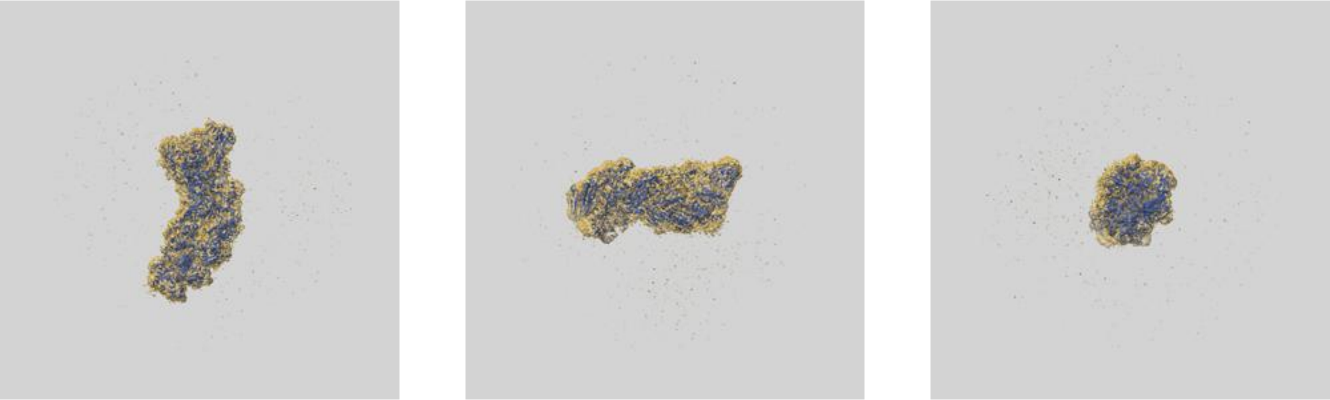
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.60	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	-	-	-

*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-29878 and PDB model 8G9T. Per-residue inclusion information can be found in section 3 on page 6.

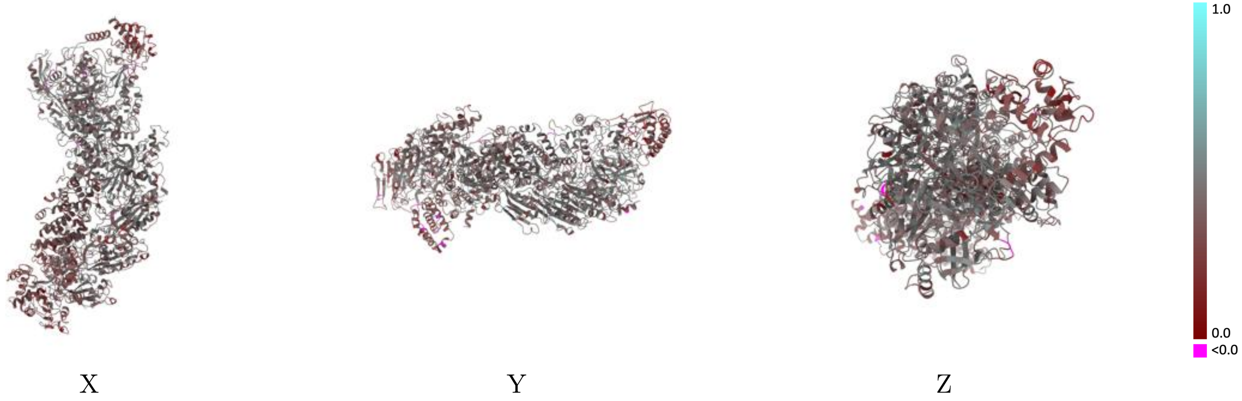
9.1 Map-model overlay (i)



X Y Z

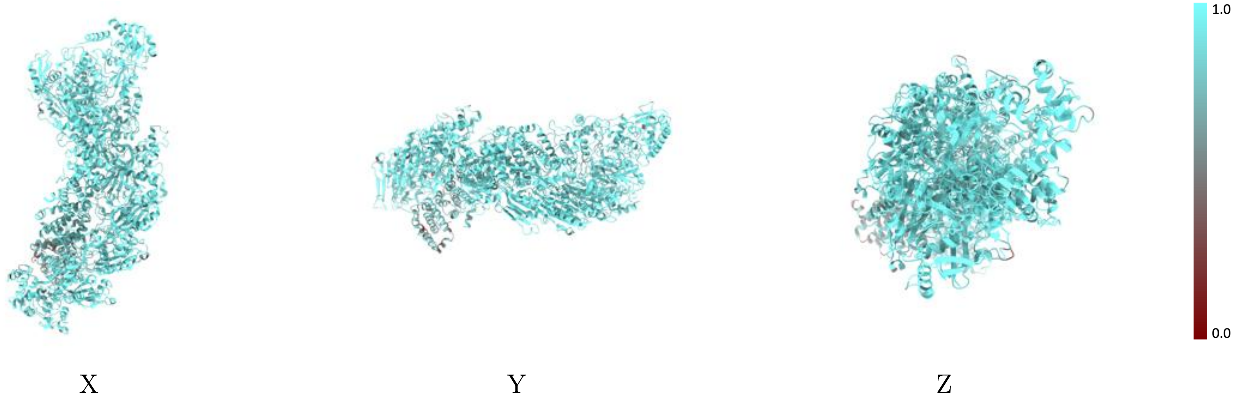
The images above show the 3D surface view of the map at the recommended contour level 0.165 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model ⓘ



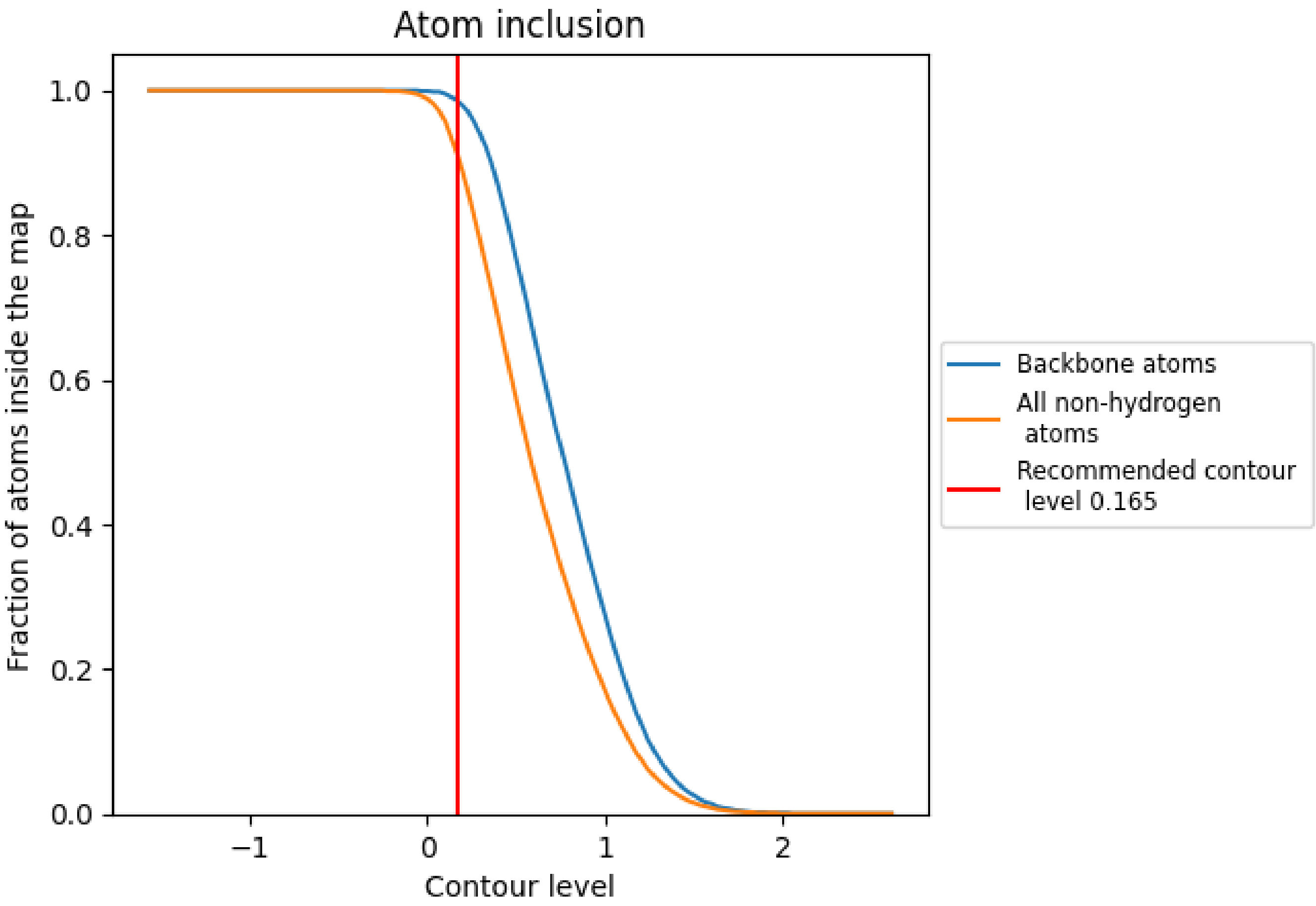
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions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



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 (0.165).

9.4 Atom inclusion ⓘ



At the recommended contour level, 99% of all backbone atoms, 92% 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 (0.165) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9160	0.3910
A	0.9400	0.4340
B	0.9440	0.4390
C	0.9360	0.4300
D	0.9410	0.4310
E	0.9200	0.4050
F	0.9200	0.3780
G	0.7920	0.3050
H	0.9090	0.2970
I	0.8920	0.3630
J	0.9240	0.3580
K	0.9360	0.3900
L	0.6650	0.2250
M	0.9360	0.4370
N	0.9380	0.4260
O	0.9640	0.4240