wwPDB EM Validation Summary Report (i)

Jun 8, 2024 – 07:43 PM EDT

PDB ID	:	8EH8
EMDB ID	:	EMD-28143
Title	:	Cryo-EM structure of his-elemental paused elongation complex with a folded TL and a rotated RH-FL (1)
Authors	:	Kang, J.Y.; Chen, J.; Llewellyn, E.; Landick, R.; Darst, S.A.
Deposited on	:	2022-09-13
Resolution	:	3.40 Å(reported)
Based on initial model	:	6C6T

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 (1)) were used in the production of this report:

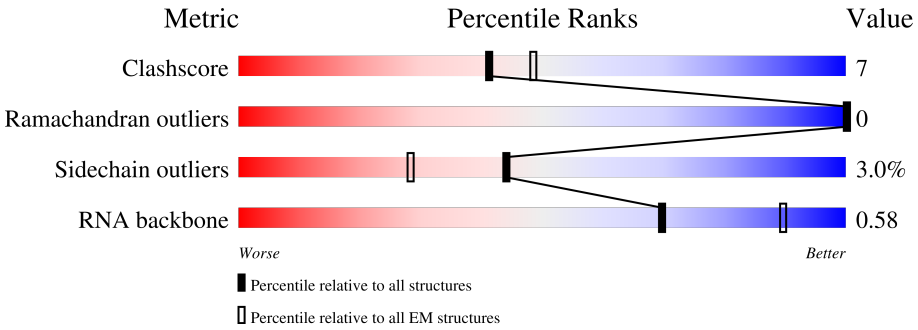
EMDB validation analysis	:	0.0.1.dev92
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 3.40 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.



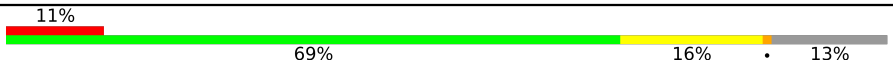
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	32	16% 31% 34% 6% 28%
2	B	32	19% 56% 41% .
3	R	19	37% 16% 47%
4	G	239	. 74% 17% 9%
4	H	239	7% 70% 21% 8%
5	I	1342	8% 79% 18% . .
6	J	1407	10% 77% 19% . .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	K	91	

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 26246 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 DNA chain called non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	A	23	472	225	90	135	22	0	0

- Molecule 2 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	B	31	627	299	112	185	31	0	0

- Molecule 3 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	R	10	214	95	36	73	10	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	G	218	1679	1051	297	325	6	0	0
4	H	219	Total	C	N	O	S	0	0
			1689	1056	298	329	6		

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	235	GLU	-	expression tag	UNP P0A7Z6
G	236	VAL	-	expression tag	UNP P0A7Z6
G	237	LEU	-	expression tag	UNP P0A7Z6
G	238	PHE	-	expression tag	UNP P0A7Z6
G	239	GLN	-	expression tag	UNP P0A7Z6
H	235	GLU	-	expression tag	UNP P0A7Z6
H	236	VAL	-	expression tag	UNP P0A7Z6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	237	LEU	-	expression tag	UNP P0A7Z6
H	238	PHE	-	expression tag	UNP P0A7Z6
H	239	GLN	-	expression tag	UNP P0A7Z6

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	1316	Total	C	N	O	S	0	0
			10375	6511	1807	2014	43		

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	1358	Total	C	N	O	S	0	0
			10533	6614	1882	1987	50		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	1	VAL	-	expression tag	UNP C3SIA2

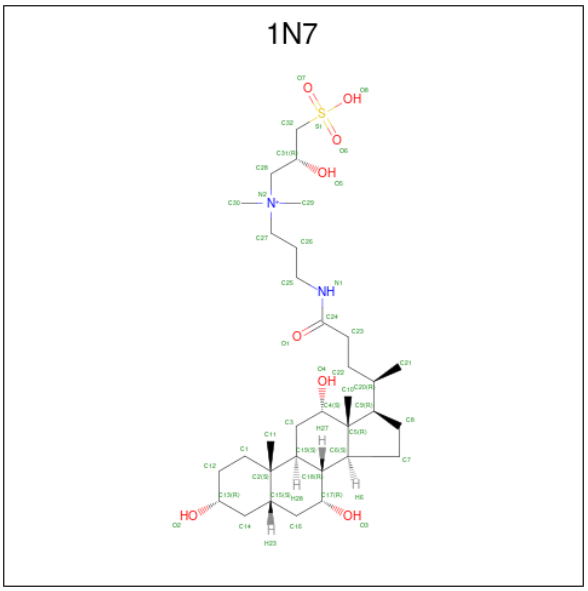
- Molecule 7 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	K	79	Total	C	N	O	S	0	0
			627	382	118	126	1		

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
8	R	1	Total	Mg	0
			1	1	

- Molecule 9 is CHAPSO (three-letter code: 1N7) (formula: $C_{32}H_{59}N_2O_8S$).



Mol	Chain	Residues	Atoms			AltConf
9	I	1	Total	C	O	0
			27	24	3	

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

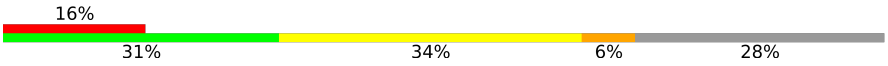
Mol	Chain	Residues	Atoms		AltConf
10	J	2	Total	Zn	0
			2	2	

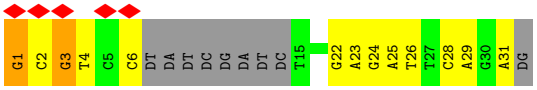
3 Residue-property plots (i)

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: non-template DNA

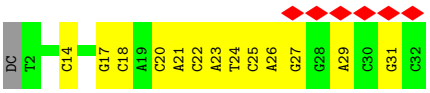
Chain A:





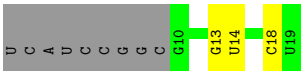
• Molecule 2: template DNA

Chain B:



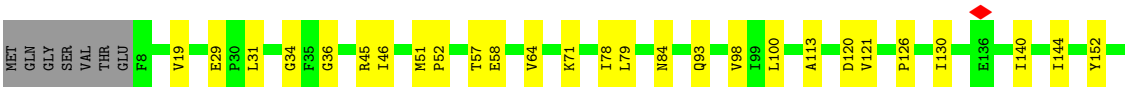
• Molecule 3: RNA

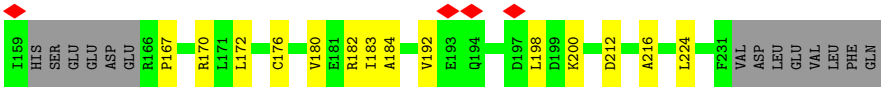
Chain R:



• Molecule 4: DNA-directed RNA polymerase subunit alpha

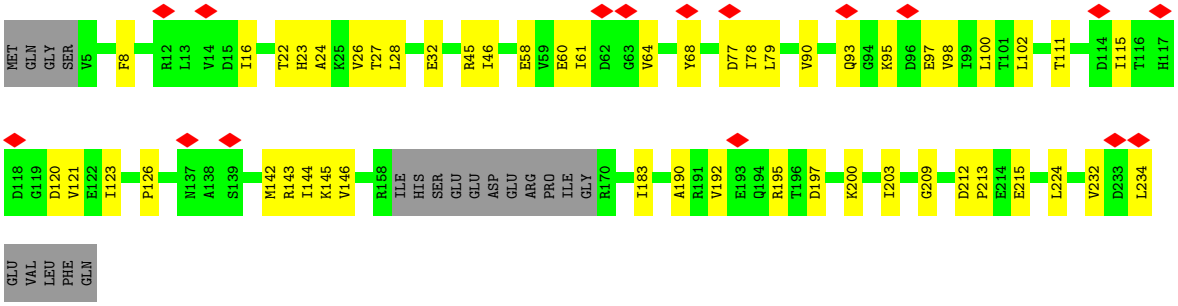
Chain G:





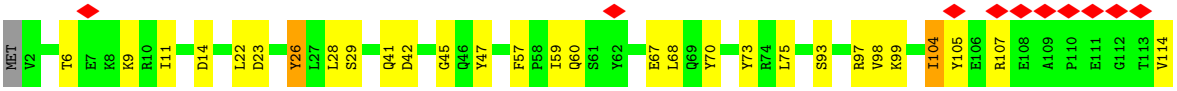
• Molecule 4: DNA-directed RNA polymerase subunit alpha

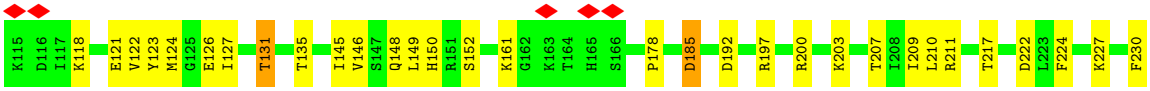
Chain H:

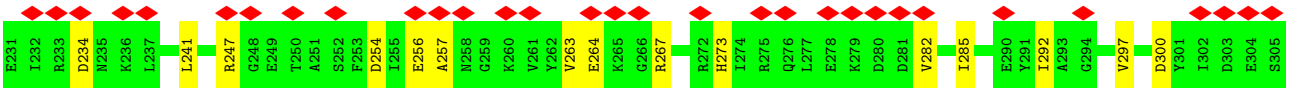


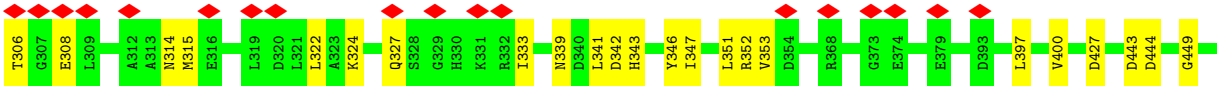
• Molecule 5: DNA-directed RNA polymerase subunit beta

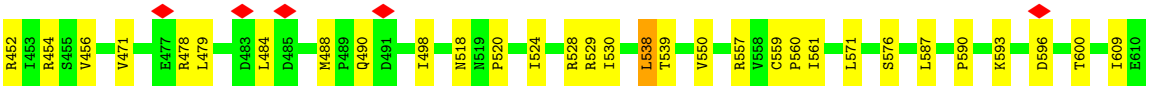
Chain I:

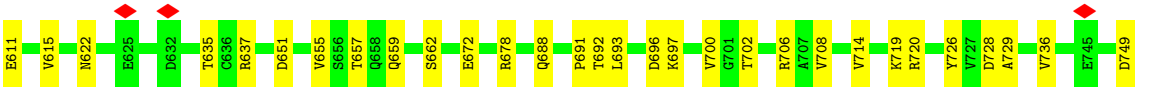


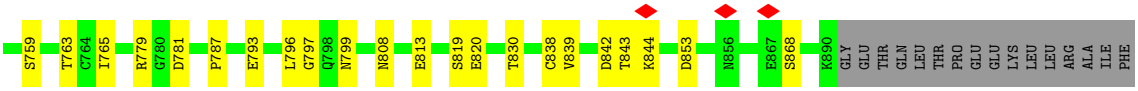


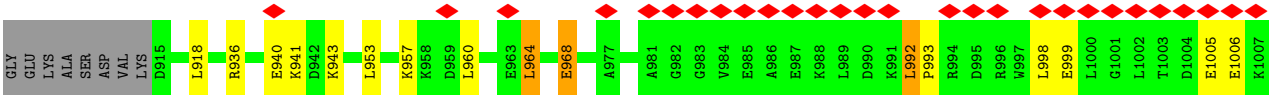






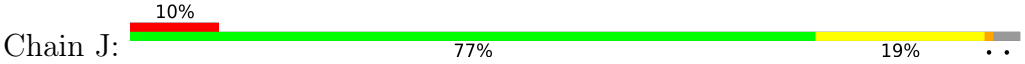


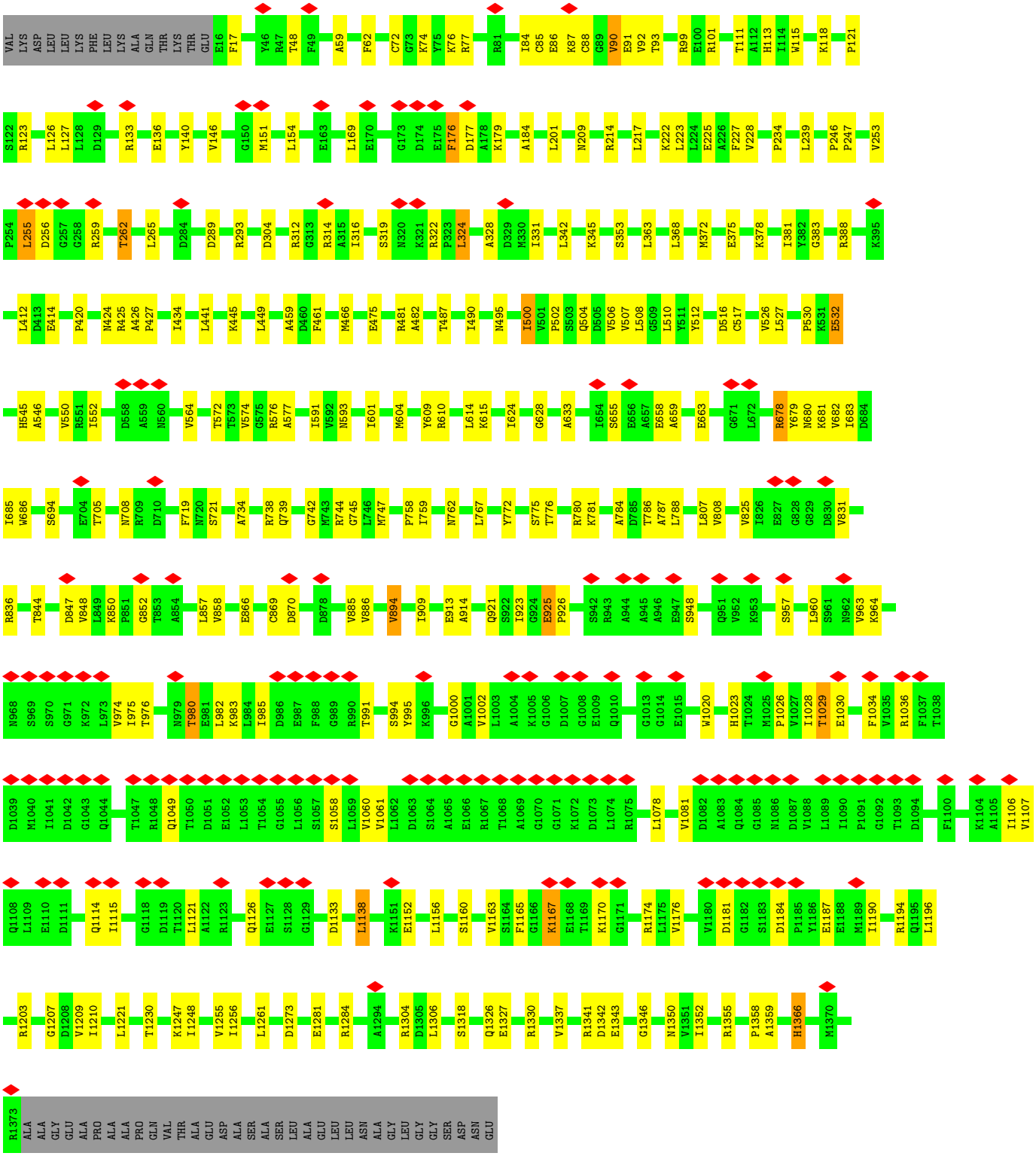




• Molecule 6: DNA-directed RNA polymerase subunit beta'

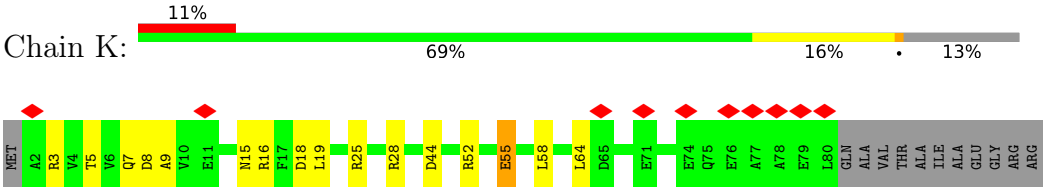
Chain J:





• Molecule 7: DNA-directed RNA polymerase subunit omega

Chain K:


11%
69%
16%
13%
MET A2 R3 V4 T5 V6 Q7 D8 A9 V10 E11 N15 R16 F17 D18 L19 R25 R28 D44 R52 E55 L58 L64 D65 E71 E74 Q75 E76 A77 A78 E79 L80 GLN ALA VAL THR ALA ILE ALA GLU GLY ARG ARG

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	150900	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 ($e^-/Å^2$)	47.34	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.215	Depositor
Minimum map value	-0.147	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.028	Depositor
Map size (Å)	332.8, 332.8, 332.8	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.3, 1.3, 1.3	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 1N7, MG, ZN

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	0.71	2/529 (0.4%)	0.98	0/813
2	B	0.66	1/701 (0.1%)	0.94	0/1077
3	R	0.32	0/238	0.86	0/369
4	G	0.25	0/1699	0.51	0/2302
4	H	0.24	0/1708	0.51	0/2315
5	I	0.26	0/10541	0.50	0/14225
6	J	0.25	0/10694	0.52	0/14441
7	K	0.23	0/629	0.52	0/847
All	All	0.29	3/26739 (0.0%)	0.55	0/36389

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
6	J	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	3	DG	C1'-N9	-5.61	1.39	1.47
1	A	1	DG	C1'-N9	-5.18	1.40	1.47
2	B	31	DG	C1'-N9	-5.12	1.40	1.47

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
6	J	1138	LEU	Peptide
6	J	925	GLU	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	472	0	261	18	0
2	B	627	0	349	18	0
3	R	214	0	107	2	0
4	G	1679	0	1717	26	0
4	H	1689	0	1725	28	0
5	I	10375	0	10381	149	0
6	J	10533	0	10743	164	0
7	K	627	0	634	11	0
8	R	1	0	0	0	0
9	I	27	0	37	9	0
10	J	2	0	0	0	0
All	All	26246	0	25954	384	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 7.

The worst 5 of 384 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)
9:I:1401:1N7:C3	9:I:1401:1N7:C19	1.82	1.52
1:A:3:DG:H2"	1:A:4:DT:H71	1.33	1.05
1:A:3:DG:H2"	1:A:4:DT:C7	1.86	1.03
1:A:3:DG:C6	2:B:29:DA:N6	2.32	0.97
6:J:72:CYS:HB3	6:J:88:CYS:SG	2.18	0.83

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	
4	G	214/239 (90%)	205 (96%)	9 (4%)	0	100	100
4	H	215/239 (90%)	202 (94%)	13 (6%)	0	100	100
5	I	1312/1342 (98%)	1234 (94%)	78 (6%)	0	100	100
6	J	1356/1407 (96%)	1277 (94%)	79 (6%)	0	100	100
7	K	77/91 (85%)	71 (92%)	6 (8%)	0	100	100
All	All	3174/3318 (96%)	2989 (94%)	185 (6%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	G	185/206 (90%)	183 (99%)	2 (1%)	73	86
4	H	187/206 (91%)	181 (97%)	6 (3%)	39	67
5	I	1134/1157 (98%)	1102 (97%)	32 (3%)	43	70
6	J	1131/1168 (97%)	1093 (97%)	38 (3%)	37	65
7	K	67/75 (89%)	64 (96%)	3 (4%)	27	58
All	All	2704/2812 (96%)	2623 (97%)	81 (3%)	44	68

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	J	564	VAL
6	J	1107	VAL
6	J	678	ARG
6	J	980	THR
6	J	1209	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/19 (47%)	0	0

There are no RNA backbone outliers to report.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
9	1N7	I	1401	-	30,30,46	5.06	16 (53%)	47,48,72	2.32	12 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	1N7	I	1401	-	-	6/7/72/92	0/4/4/4

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	1401	1N7	C3-C19	17.50	1.82	1.53
9	I	1401	1N7	C3-C4	11.91	1.73	1.53
9	I	1401	1N7	C5-C4	-9.28	1.40	1.54
9	I	1401	1N7	C2-C19	-8.35	1.40	1.56
9	I	1401	1N7	C8-C7	6.18	1.70	1.54

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	I	1401	1N7	C9-C5-C4	-8.47	109.93	117.67
9	I	1401	1N7	C5-C9-C20	-5.55	112.86	119.50
9	I	1401	1N7	C6-C5-C4	5.25	112.29	107.40
9	I	1401	1N7	C7-C6-C18	-4.92	111.45	118.33
9	I	1401	1N7	C19-C3-C4	-4.09	108.90	114.30

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	I	1401	1N7	C21-C20-C9-C5
9	I	1401	1N7	C22-C20-C9-C8
9	I	1401	1N7	C21-C20-C9-C8
9	I	1401	1N7	C22-C20-C9-C5
9	I	1401	1N7	C20-C22-C23-C24

There are no ring outliers.

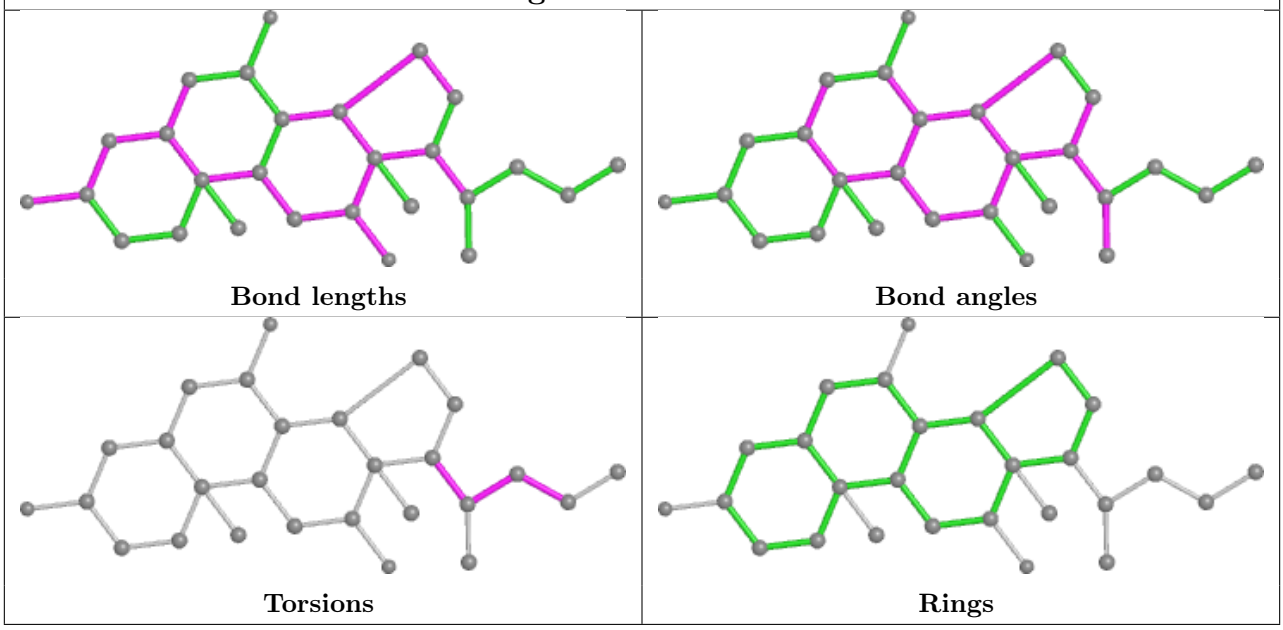
1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	I	1401	1N7	9	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 1N7 I 1401



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.

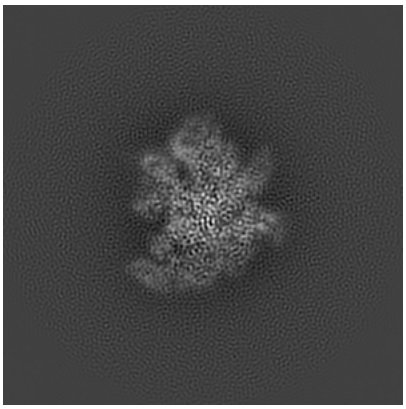
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-28143. These allow visual inspection of the internal detail of the map and identification of artifacts.

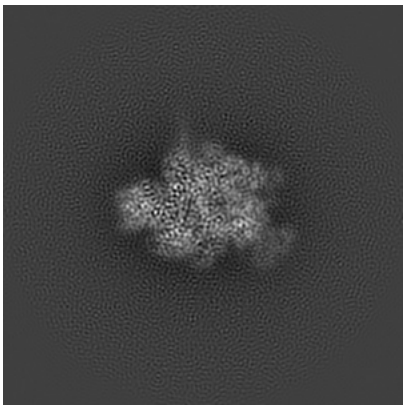
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

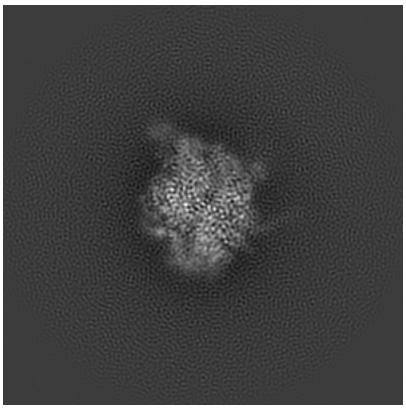
6.1.1 Primary map



X

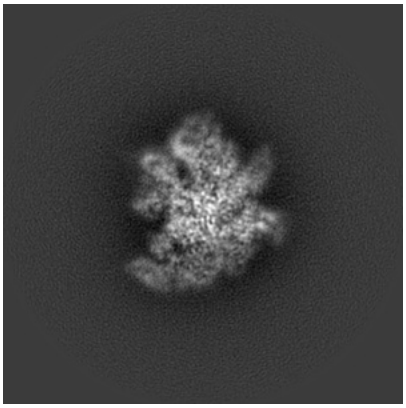


Y

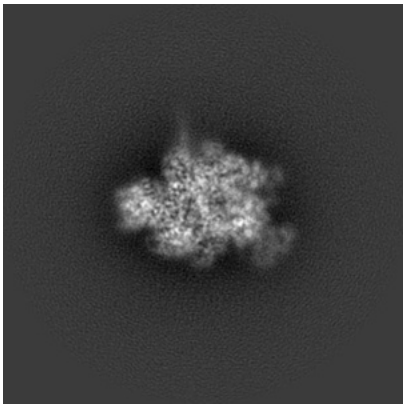


Z

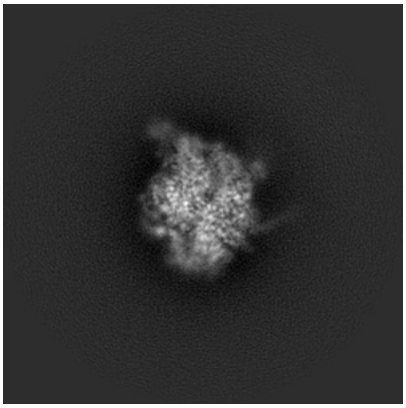
6.1.2 Raw map



X



Y

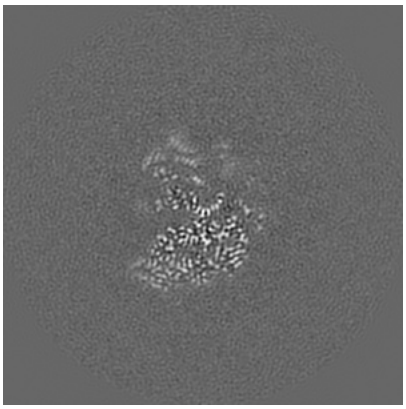


Z

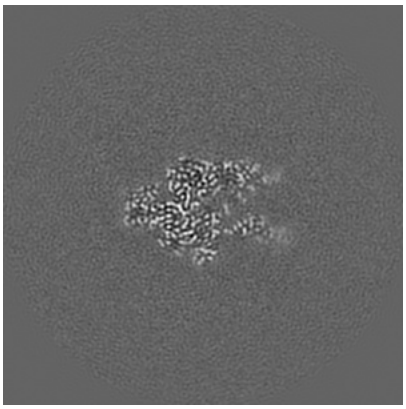
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

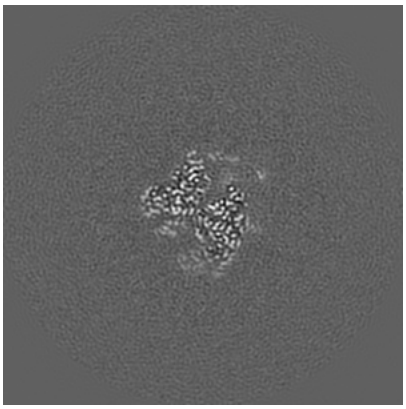
6.2.1 Primary map



X Index: 128

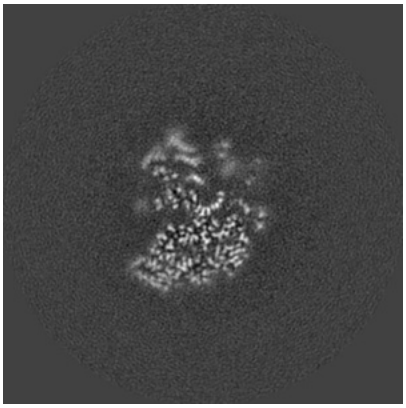


Y Index: 128

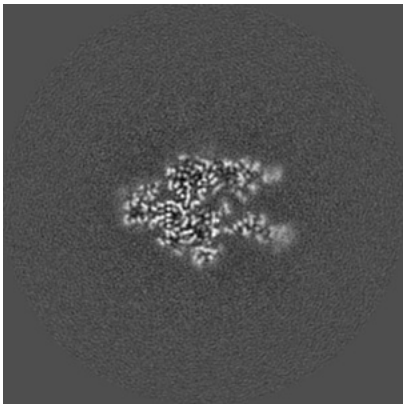


Z Index: 128

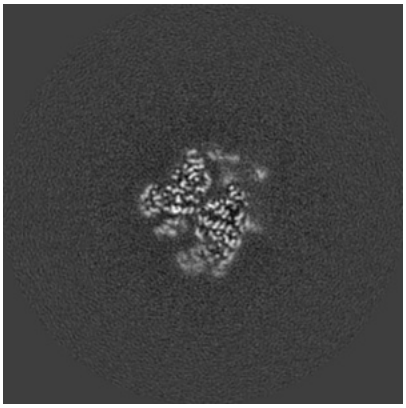
6.2.2 Raw map



X Index: 128



Y Index: 128

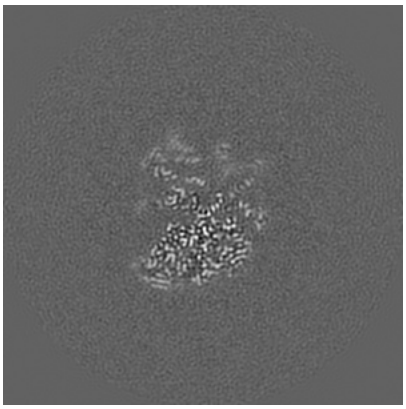


Z Index: 128

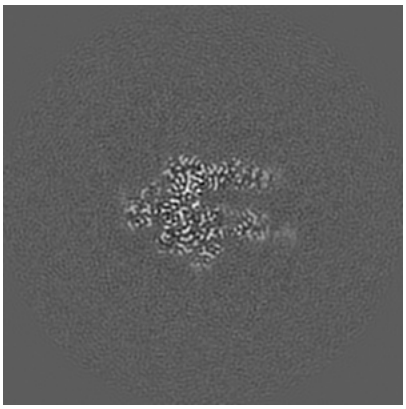
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

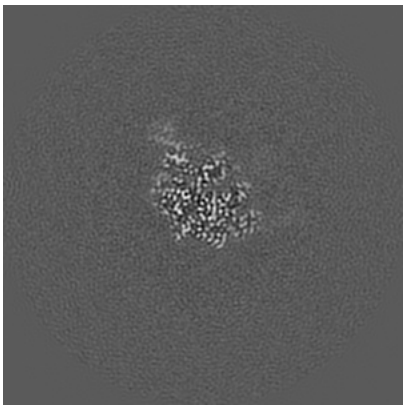
6.3.1 Primary map



X Index: 127

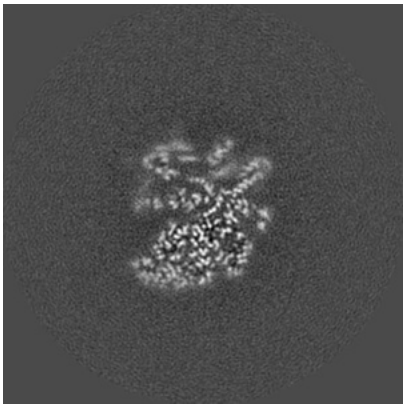


Y Index: 131

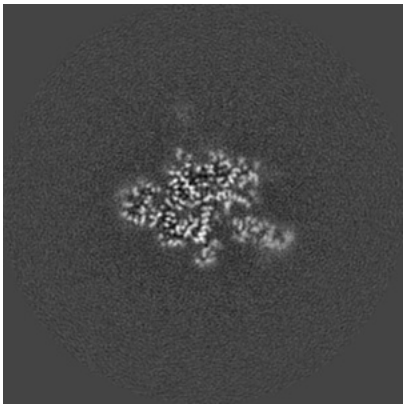


Z Index: 111

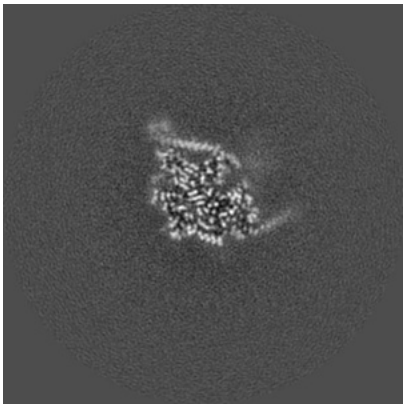
6.3.2 Raw map



X Index: 125



Y Index: 124

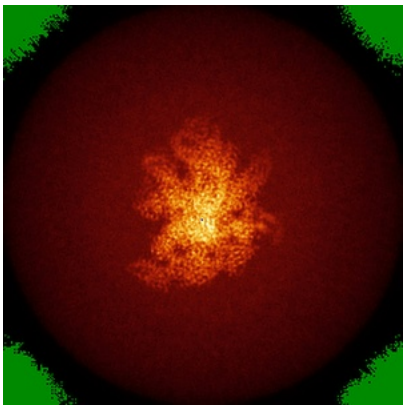


Z Index: 113

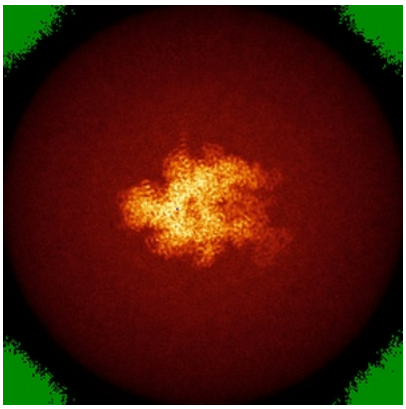
The images above show the largest variance slices of the map in three orthogonal directions.

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

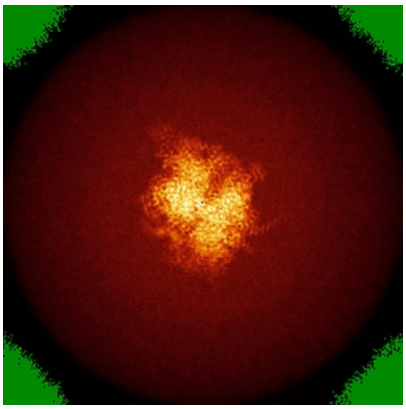
6.4.1 Primary map



X

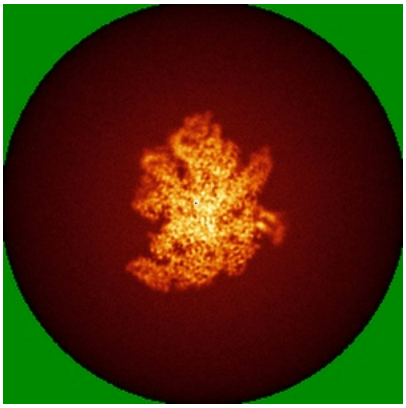


Y

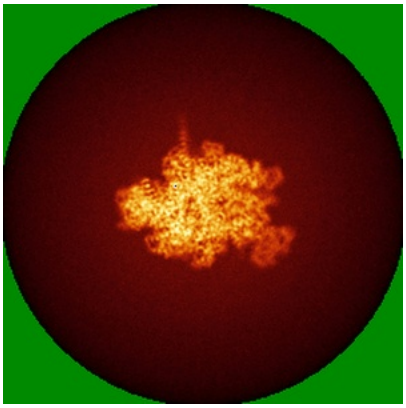


Z

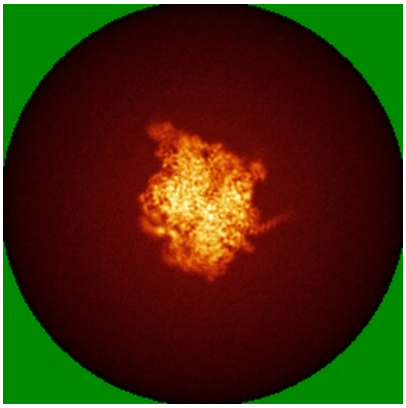
6.4.2 Raw map



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.

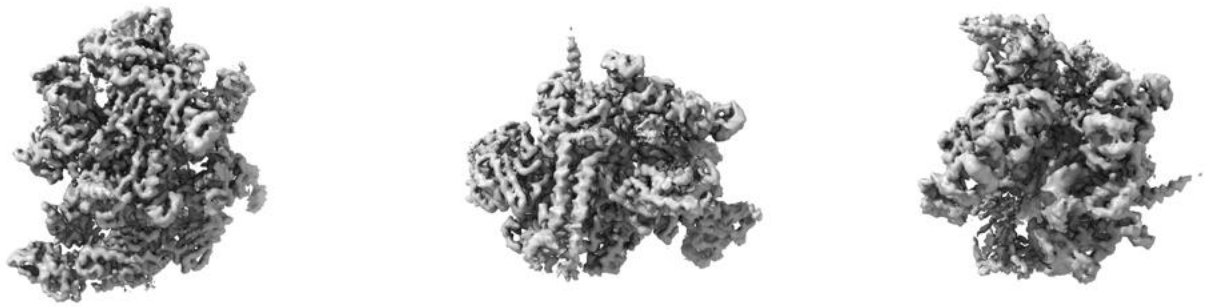
6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.028. 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.

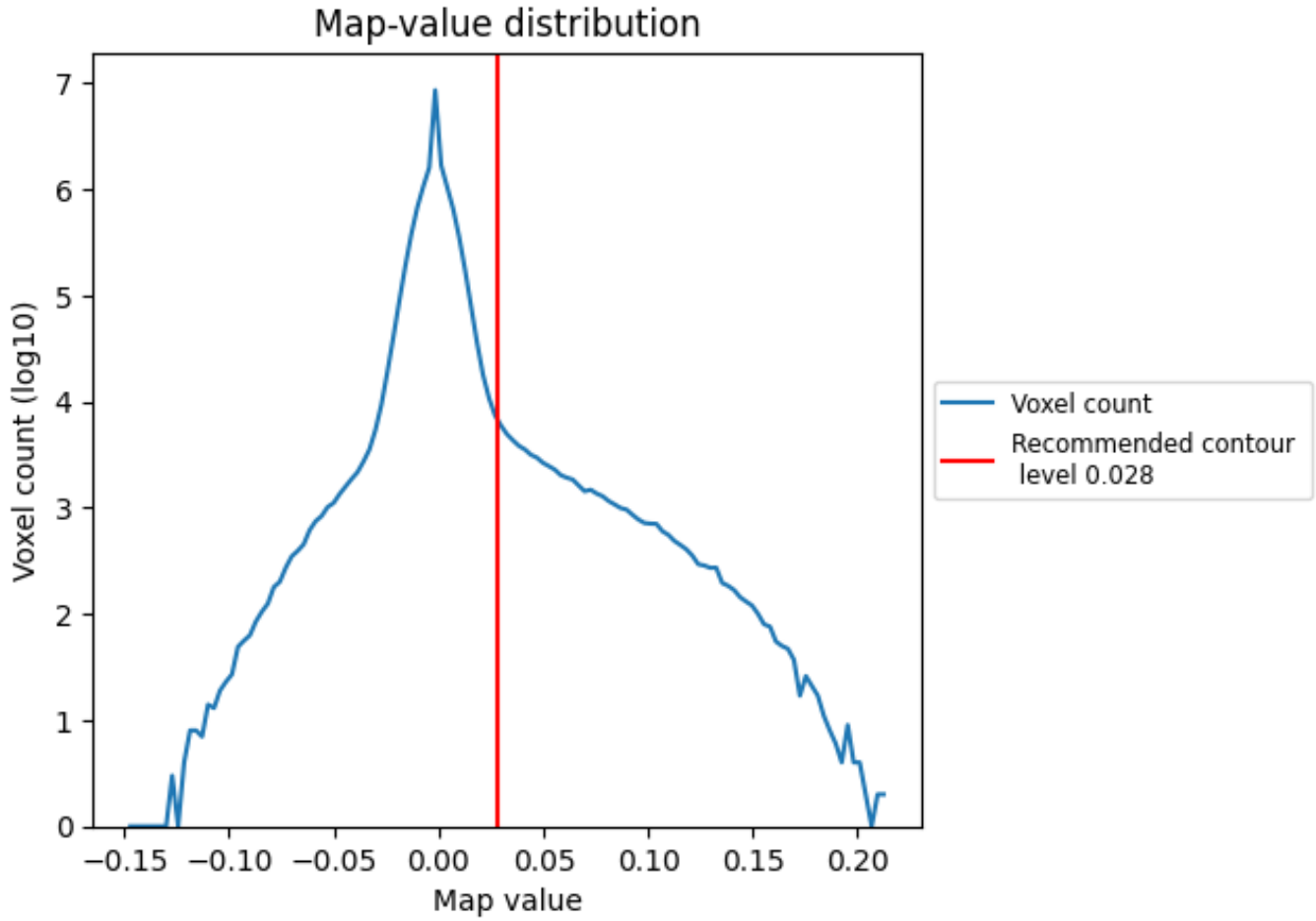
6.6 Mask visualisation (i)

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

7 Map analysis

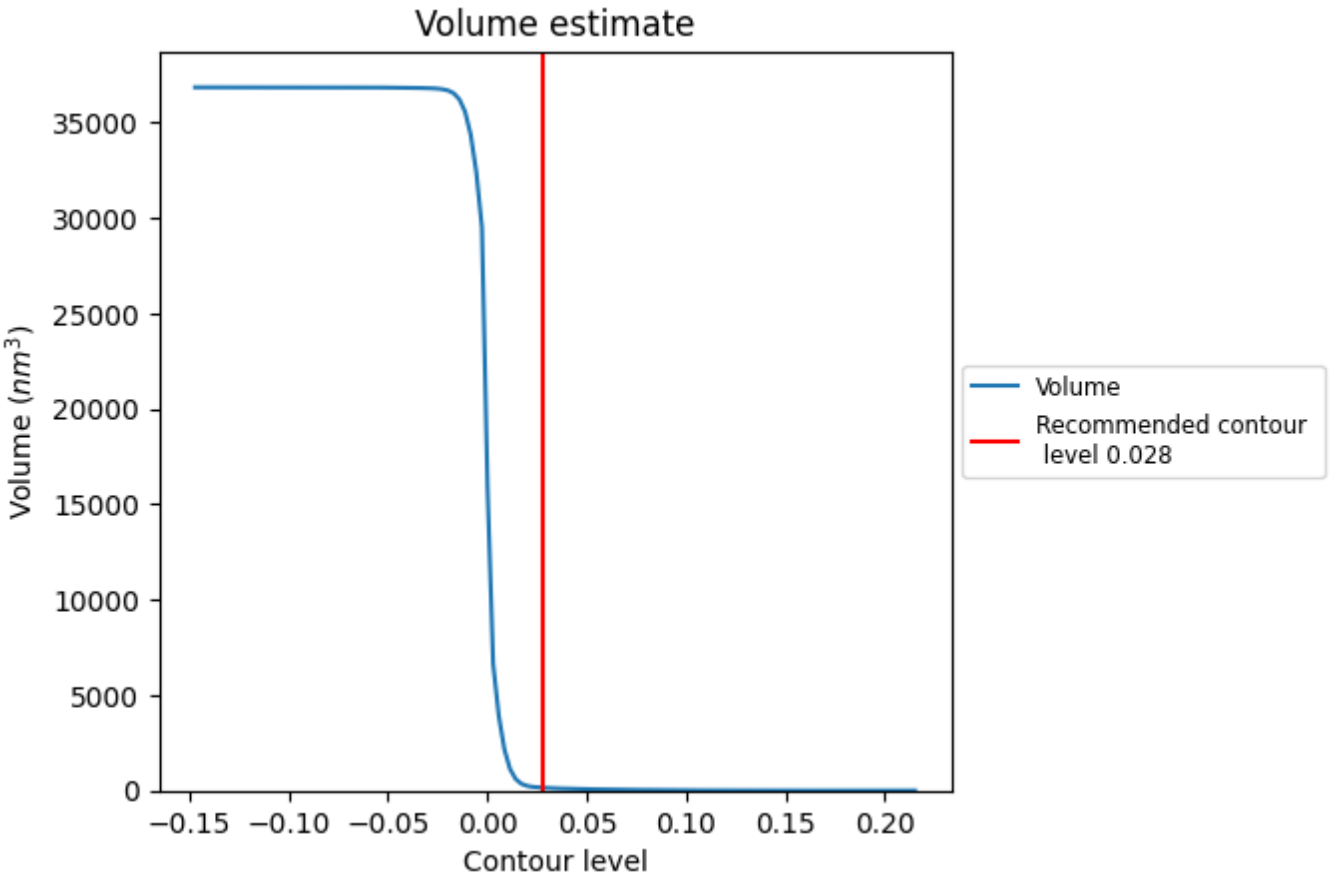
This section contains the results of statistical analysis of the map.

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.

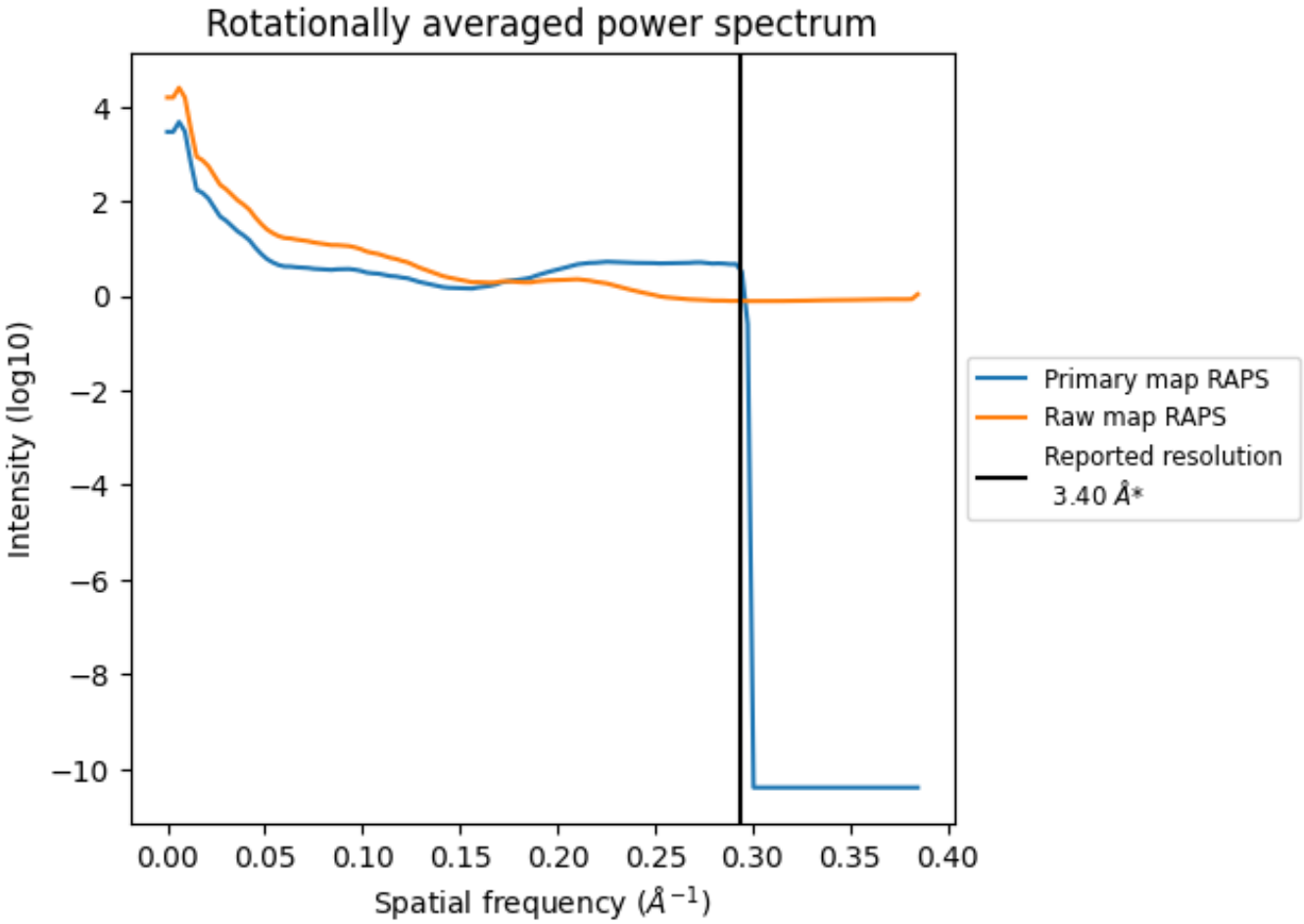
7.2 Volume estimate (i)



The volume at the recommended contour level is 148 nm^3; this corresponds to an approximate mass of 133 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)

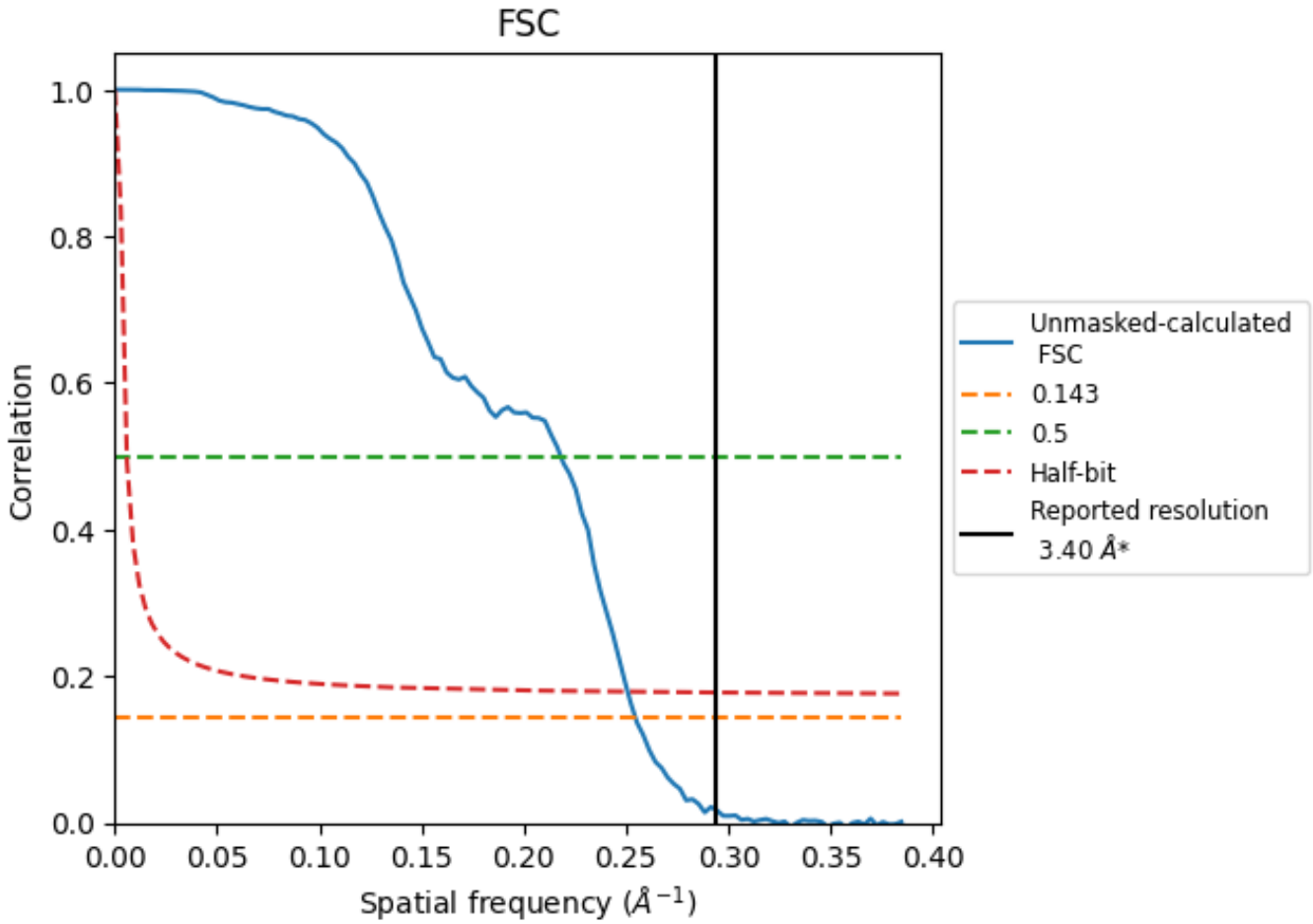


*Reported resolution corresponds to spatial frequency of 0.294 Å^{-1}

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 0.294 Å^{-1}

8.2 Resolution estimates (i)

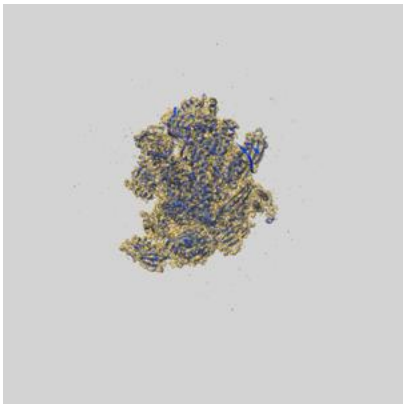
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.40	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.93	4.59	3.99

*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 3.93 differs from the reported value 3.4 by more than 10 %

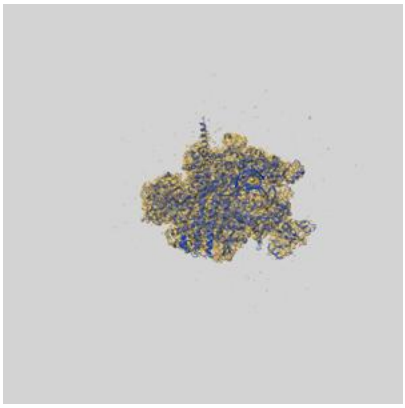
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-28143 and PDB model 8EH8. Per-residue inclusion information can be found in section 3 on page 7.

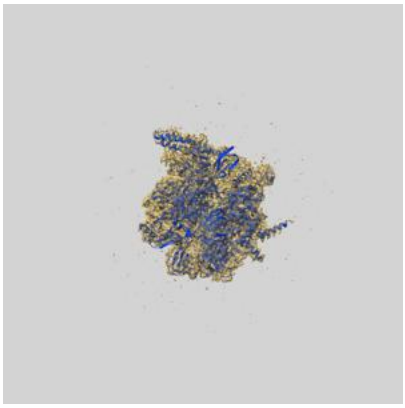
9.1 Map-model overlay



X



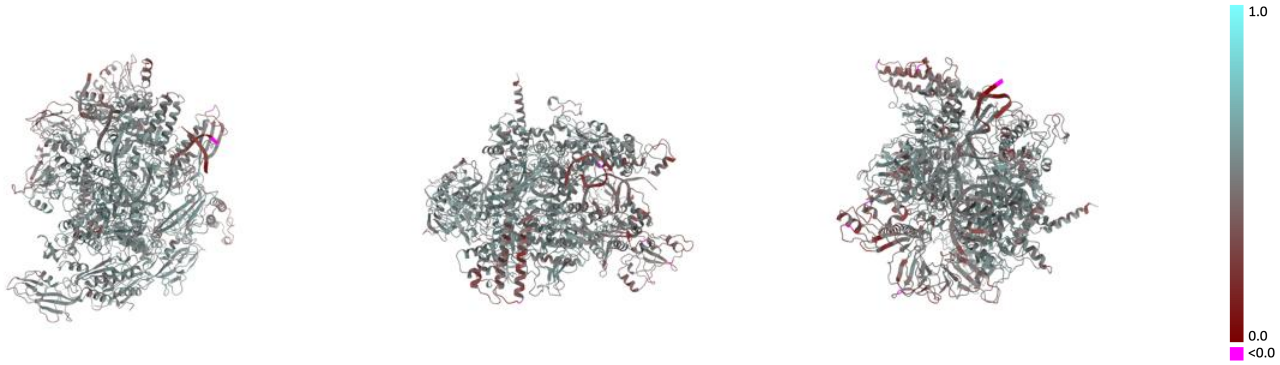
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.028 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

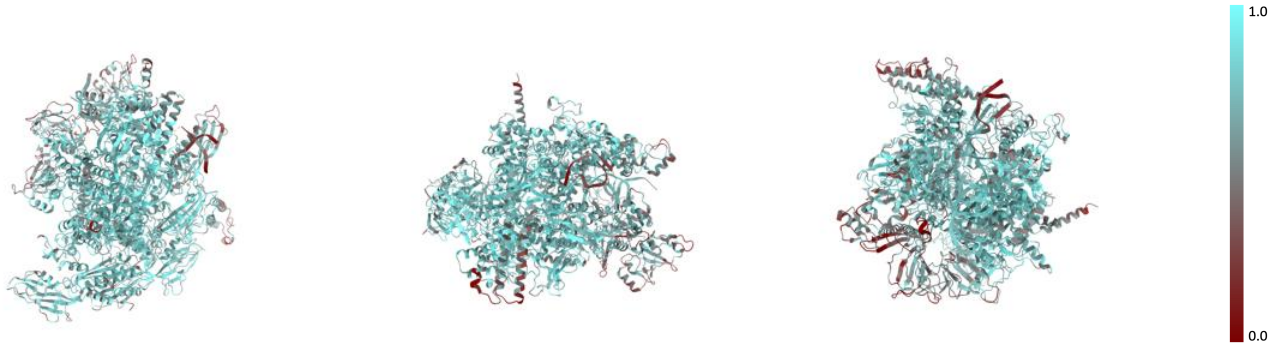
9.2 Q-score mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

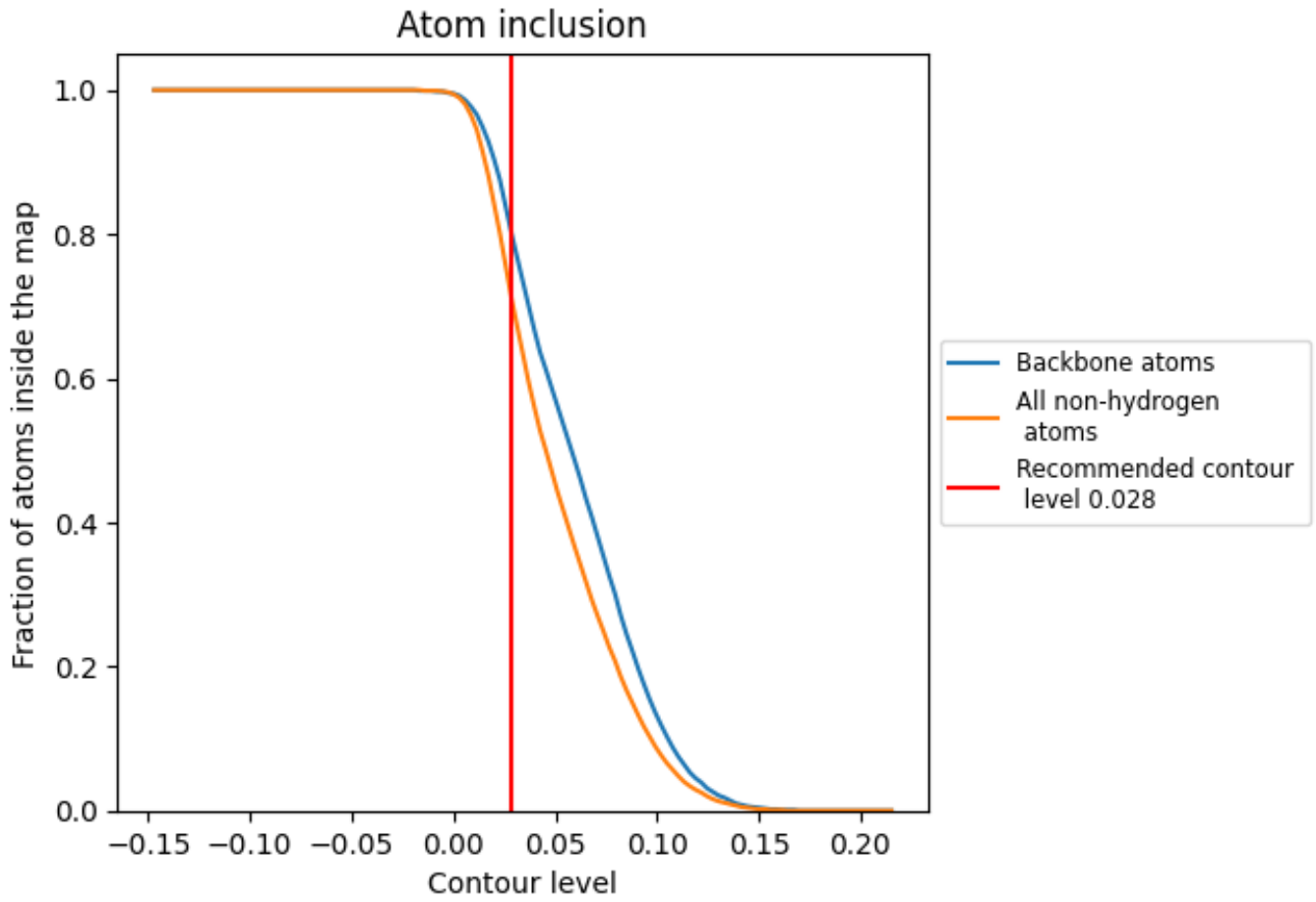
9.3 Atom inclusion mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.028).

9.4 Atom inclusion (i)



At the recommended contour level, 81% of all backbone atoms, 72% 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 (0.028) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7190	0.4960
A	0.6360	0.3710
B	0.7020	0.4170
G	0.7680	0.5300
H	0.7110	0.4920
I	0.7300	0.4990
J	0.7060	0.4980
K	0.6560	0.4980
R	0.8880	0.5110