Full wwPDB EM Validation Report (i)

Dec 12, 2022 – 06:47 am GMT

PDB ID	:	6Y5E
EMDB ID	:	EMD-10695
Title	:	Structure of human cGAS (K394E) bound to the nucleosome (focused refinement of cGAS-NCP subcomplex)
Authors	:	Pathare, G.R.; Cavadini, S.; Kempf, G.; Thoma, N.H.
Deposited on	:	2020-02-25
Resolution	:	3.15 Å(reported)

This is a Full wwPDB EM Validation 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:

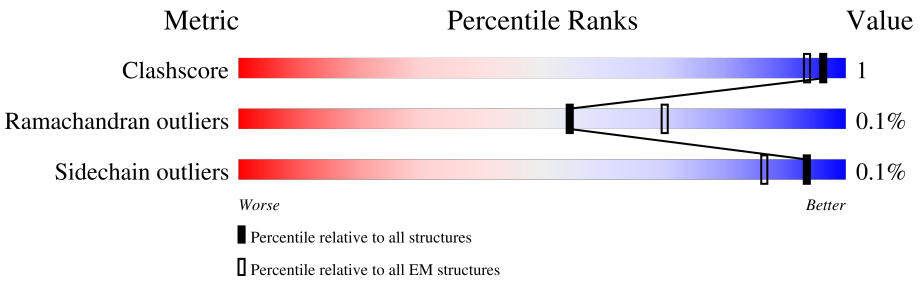
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 3.15 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



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

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	96	98% •
1	E	96	98% •
2	B	81	• 95% 5%
3	C	107	95% 5%
4	D	93	100%
5	F	83	• 93% 7%
6	G	108	• 94% 6%
7	H	94	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	I	153	98%
9	J	153	97%
10	K	362	98%

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 15292 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 Histone H3.2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	96	789	499	151	137	2	0	0
1	E	96	789	499	151	137	2	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	111	ALA	CYS	conflict	UNP Q71DI3
E	111	ALA	CYS	conflict	UNP Q71DI3

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	81	646	407	126	112	1	0	0

- Molecule 3 is a protein called Histone H2A type 2-C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	107	822	518	161	142	1	0	0

- Molecule 4 is a protein called Histone H2B type 1-K.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	93	724	455	130	137	2	0	0

- Molecule 5 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	83	662	418	129	114	1	0	0

- Molecule 6 is a protein called Histone H2A type 2-C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	G	108	831	524	163	143	1	0	0

- Molecule 7 is a protein called Histone H2B type 1-K.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	H	94	735	461	134	138	2	0	0

- Molecule 8 is a DNA chain called DNA (153-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	I	153	3119	1482	564	920	153	0	0

- Molecule 9 is a DNA chain called DNA (153-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
9	J	153	3154	1493	592	916	153	0	0

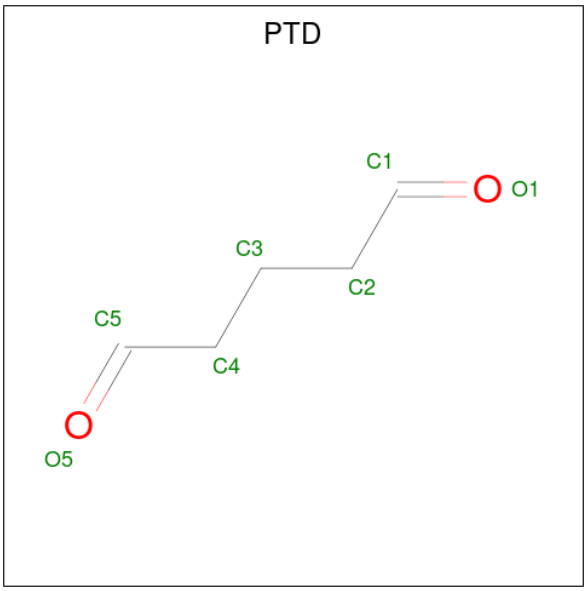
- Molecule 10 is a protein called Cyclic GMP-AMP synthase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	K	362	2980	1903	514	548	15	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	394	GLU	LYS	conflict	UNP Q8N884

- Molecule 11 is PENTANEDIAL (three-letter code: PTD) (formula: $C_5H_8O_2$).



Mol	Chain	Residues	Atoms		AltConf
			Total	C	
11	A	1	5	5	0
11	C	1	5	5	0
11	D	1	5	5	0
11	E	1	10	10	0
11	E	1	10	10	0
11	F	1	5	5	0
11	H	1	5	5	0
11	K	1	5	5	0

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

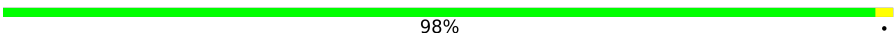
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
12	K	1	1	1	0

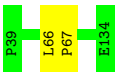
3 Residue-property plots (i)

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: Histone H3.2

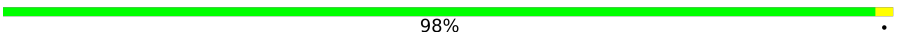
Chain A:

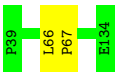




- Molecule 1: Histone H3.2

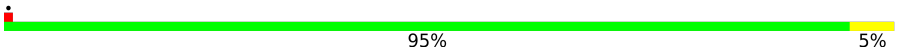
Chain E:

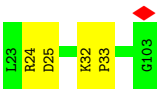




- Molecule 2: Histone H4

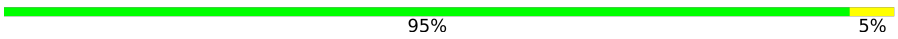
Chain B:

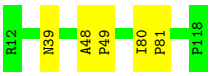




- Molecule 3: Histone H2A type 2-C

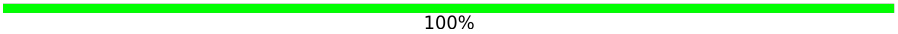
Chain C:





- Molecule 4: Histone H2B type 1-K

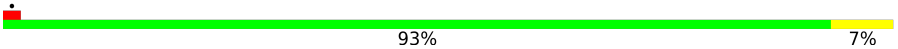
Chain D:

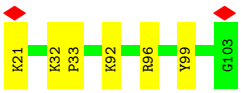


There are no outlier residues recorded for this chain.

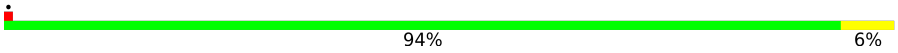
- Molecule 5: Histone H4

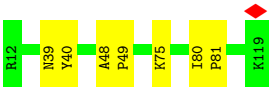
Chain F:



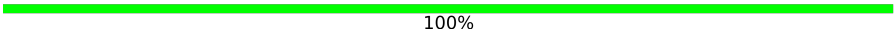


• Molecule 6: Histone H2A type 2-C

Chain G: 

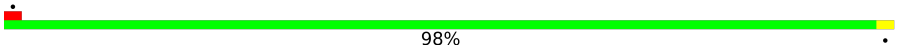


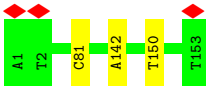
• Molecule 7: Histone H2B type 1-K

Chain H: 

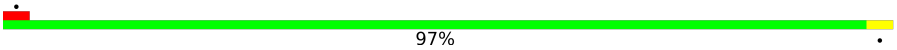
There are no outlier residues recorded for this chain.

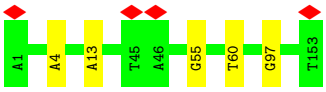
• Molecule 8: DNA (153-MER)

Chain I: 

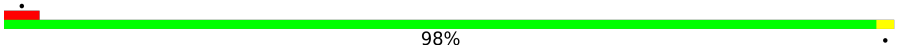


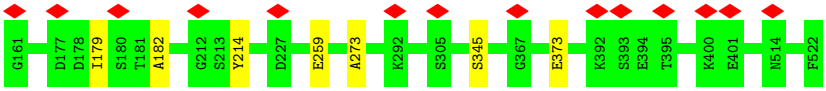
• Molecule 9: DNA (153-MER)

Chain J: 



• Molecule 10: Cyclic GMP-AMP synthase

Chain K: 



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	62036	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 ($e^-/Å^2$)	45	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 QUANTUM (4k x 4k)	Depositor
Maximum map value	1.177	Depositor
Minimum map value	-0.433	Depositor
Average map value	0.005	Depositor
Map value standard deviation	0.041	Depositor
Recommended contour level	0.139	Depositor
Map size (Å)	275.2, 275.2, 275.2	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.86, 0.86, 0.86	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: ZN, PTD

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	1.08	0/801	0.61	0/1075
1	E	1.08	0/801	0.63	0/1075
2	B	1.09	0/653	0.67	0/873
3	C	0.97	0/832	0.60	0/1121
4	D	0.94	0/735	0.62	0/989
5	F	1.02	0/669	0.71	1/894 (0.1%)
6	G	1.01	0/841	0.63	0/1132
7	H	0.94	0/746	0.63	0/1003
8	I	0.53	0/3494	0.94	2/5386 (0.0%)
9	J	0.55	0/3542	0.94	3/5469 (0.1%)
10	K	0.81	0/3040	0.64	1/4076 (0.0%)
All	All	0.80	0/16154	0.79	7/23093 (0.0%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	J	60	DT	P-O3'-C3'	6.86	127.93	119.70
8	I	150	DT	P-O3'-C3'	6.78	127.84	119.70
9	J	97	DG	P-O3'-C3'	6.50	127.50	119.70
9	J	4	DA	P-O3'-C3'	6.45	127.44	119.70
8	I	81	DC	P-O3'-C3'	5.93	126.82	119.70
10	K	214	TYR	CB-CG-CD2	-5.52	117.69	121.00
5	F	99	TYR	CB-CG-CD2	-5.21	117.87	121.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	789	0	818	1	0
1	E	789	0	820	1	0
2	B	646	0	685	2	0
3	C	822	0	875	4	0
4	D	724	0	740	0	0
5	F	662	0	707	3	0
6	G	831	0	885	5	0
7	H	735	0	753	0	0
8	I	3119	0	1719	1	0
9	J	3154	0	1717	2	0
10	K	2980	0	3009	3	0
11	A	5	0	6	0	0
11	C	5	0	6	0	0
11	D	5	0	6	0	0
11	E	10	0	12	0	0
11	F	5	0	6	0	0
11	H	5	0	6	0	0
11	K	5	0	6	0	0
12	K	1	0	0	0	0
All	All	15292	0	12776	19	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 1.

All (19) 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:C:48:ALA:N	3:C:49:PRO:CD	2.69	0.55
6:G:48:ALA:N	6:G:49:PRO:CD	2.70	0.55
2:B:32:LYS:HB3	2:B:33:PRO:HD3	1.90	0.52
6:G:39:ASN:O	6:G:40:TYR:C	2.49	0.51
3:C:39:ASN:HB3	6:G:39:ASN:O	2.15	0.46
10:K:273:ALA:N	10:K:373:GLU:OE1	2.49	0.45
6:G:80:ILE:HB	6:G:81:PRO:HD2	2.00	0.43
5:F:32:LYS:HB3	5:F:33:PRO:HD3	2.01	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:F:21:LYS:NZ	9:J:55:DG:OP1	2.52	0.43
10:K:259:GLU:N	10:K:259:GLU:OE1	2.52	0.43
2:B:24:ARG:O	2:B:25:ASP:C	2.58	0.42
1:E:66:LEU:N	1:E:67:PRO:CD	2.82	0.42
6:G:48:ALA:HB3	6:G:49:PRO:HD3	2.02	0.42
10:K:179:ILE:O	10:K:182:ALA:N	2.52	0.41
1:A:66:LEU:N	1:A:67:PRO:CD	2.82	0.41
3:C:80:ILE:HB	3:C:81:PRO:HD2	2.01	0.41
8:I:142:DA:C2	9:J:13:DA:C2	3.09	0.41
5:F:92:LYS:HA	5:F:96:ARG:O	2.21	0.41
3:C:48:ALA:HB3	3:C:49:PRO:HD3	2.02	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	94/96 (98%)	94 (100%)	0	0	100	100
1	E	94/96 (98%)	93 (99%)	1 (1%)	0	100	100
2	B	79/81 (98%)	76 (96%)	3 (4%)	0	100	100
3	C	105/107 (98%)	104 (99%)	1 (1%)	0	100	100
4	D	91/93 (98%)	91 (100%)	0	0	100	100
5	F	81/83 (98%)	80 (99%)	1 (1%)	0	100	100
6	G	106/108 (98%)	106 (100%)	0	0	100	100
7	H	92/94 (98%)	92 (100%)	0	0	100	100
10	K	360/362 (99%)	344 (96%)	15 (4%)	1 (0%)	41	73
All	All	1102/1120 (98%)	1080 (98%)	21 (2%)	1 (0%)	54	83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
10	K	345	SER

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	83/83 (100%)	83 (100%)	0	100	100
1	E	83/83 (100%)	83 (100%)	0	100	100
2	B	66/66 (100%)	66 (100%)	0	100	100
3	C	83/83 (100%)	83 (100%)	0	100	100
4	D	79/79 (100%)	79 (100%)	0	100	100
5	F	68/68 (100%)	68 (100%)	0	100	100
6	G	84/84 (100%)	83 (99%)	1 (1%)	71	87
7	H	80/80 (100%)	80 (100%)	0	100	100
10	K	334/334 (100%)	334 (100%)	0	100	100
All	All	960/960 (100%)	959 (100%)	1 (0%)	93	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	G	75	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
10	K	260	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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	$\#\|Z\| > 2$
11	PTD	A	201	1,2	4,4,6	0.49	0	3,3,5	0.37	0
11	PTD	E	202	1	4,4,6	0.38	0	3,3,5	0.48	0
11	PTD	C	201	3,6	4,4,6	0.87	0	3,3,5	0.32	0
11	PTD	F	201	1,5	4,4,6	0.41	0	3,3,5	0.46	0
11	PTD	H	301	3,7	4,4,6	0.46	0	3,3,5	0.41	0
11	PTD	E	201	1	4,4,6	0.40	0	3,3,5	0.44	0
11	PTD	D	301	4,6	4,4,6	0.38	0	3,3,5	0.47	0
11	PTD	K	602	10,6	4,4,6	0.44	0	3,3,5	0.41	0

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
11	PTD	A	201	1,2	-	1/2/2/4	-
11	PTD	E	202	1	-	0/2/2/4	-
11	PTD	C	201	3,6	-	2/2/2/4	-
11	PTD	F	201	1,5	-	1/2/2/4	-
11	PTD	H	301	3,7	-	1/2/2/4	-
11	PTD	E	201	1	-	0/2/2/4	-
11	PTD	D	301	4,6	-	0/2/2/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
11	PTD	K	602	10,6	-	1/2/2/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
11	C	201	PTD	C1-C2-C3-C4
11	C	201	PTD	C2-C3-C4-C5
11	A	201	PTD	C1-C2-C3-C4
11	F	201	PTD	C1-C2-C3-C4
11	K	602	PTD	C1-C2-C3-C4
11	H	301	PTD	C1-C2-C3-C4

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.

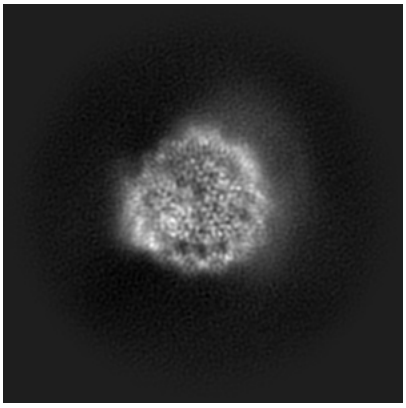
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-10695. These allow visual inspection of the internal detail of the map and identification of artifacts.

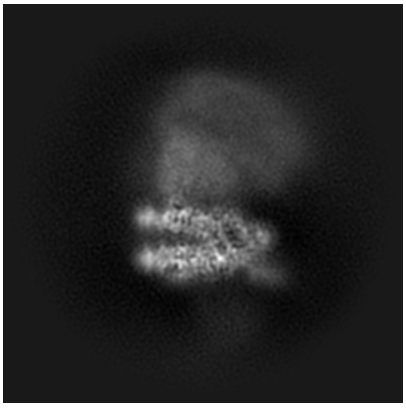
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

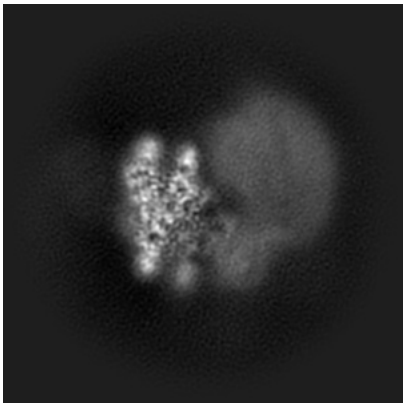
6.1.1 Primary map



X



Y

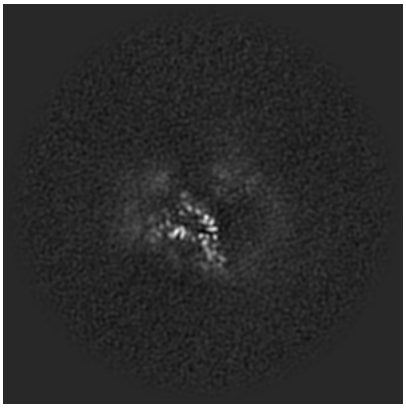


Z

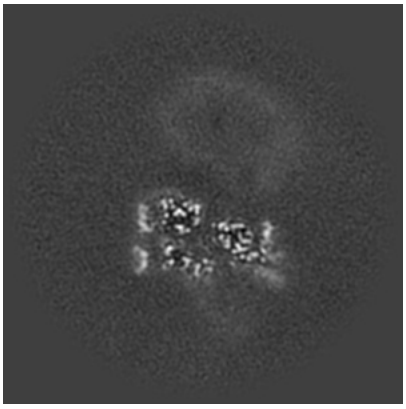
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

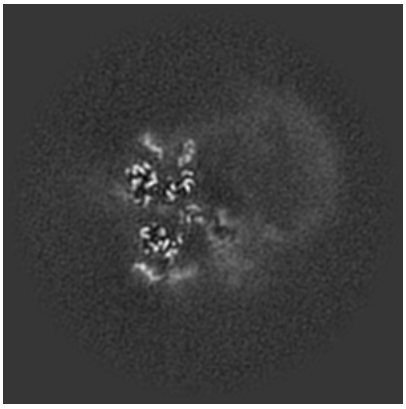
6.2.1 Primary map



X Index: 160



Y Index: 160

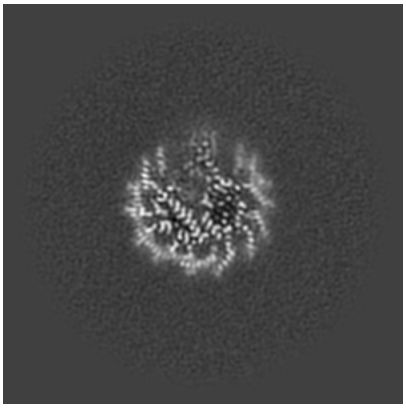


Z Index: 160

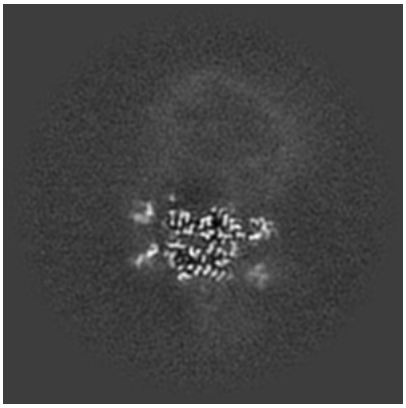
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

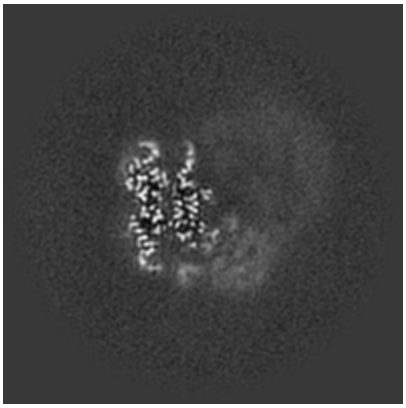
6.3.1 Primary map



X Index: 114



Y Index: 174



Z Index: 141

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

6.4 Orthogonal surface views (i)

6.4.1 Primary map

X

Y

Z

The images above show the 3D surface view of the map at the recommended contour level 0.139. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

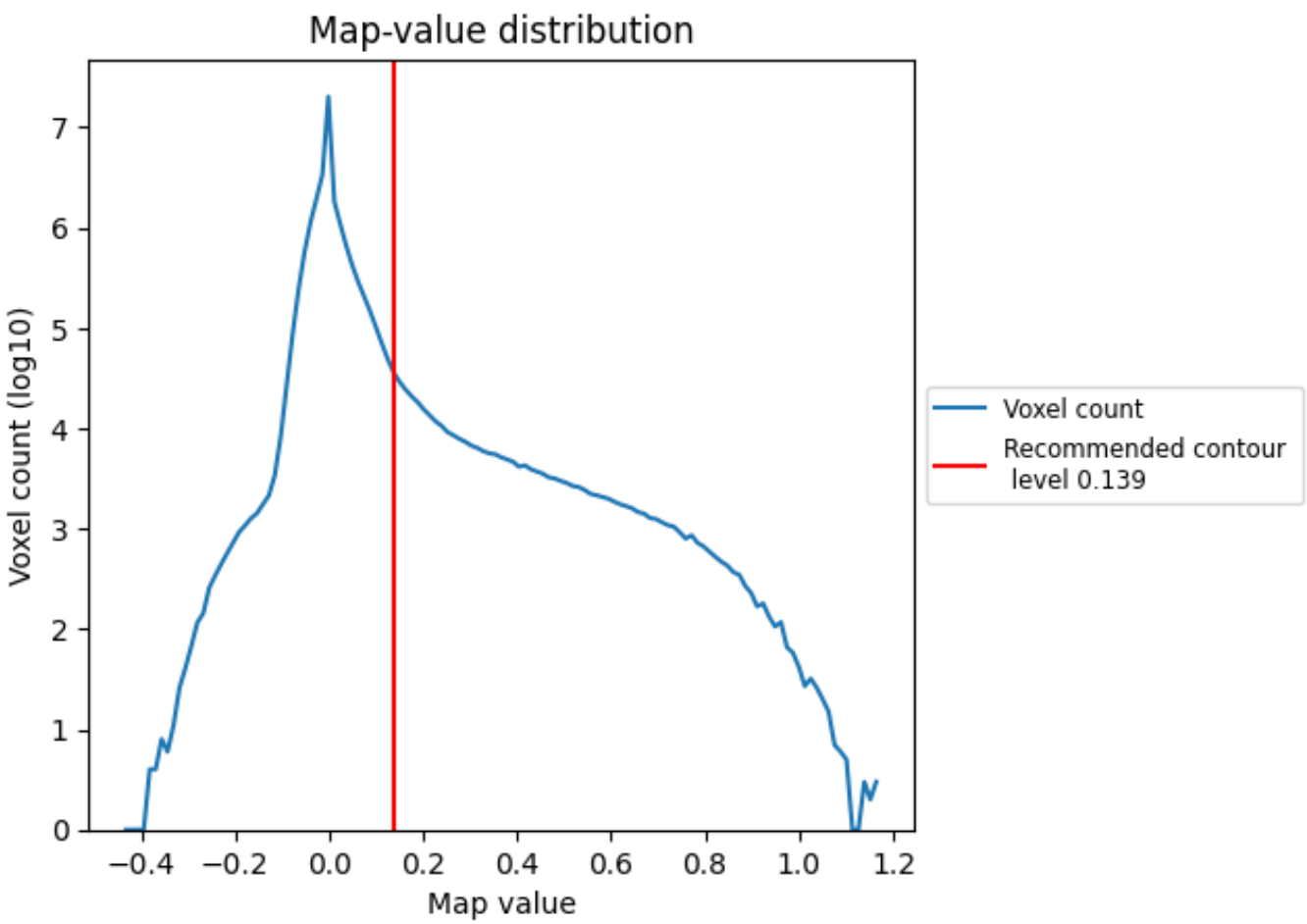
6.5 Mask visualisation (i)

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

7 Map analysis (i)

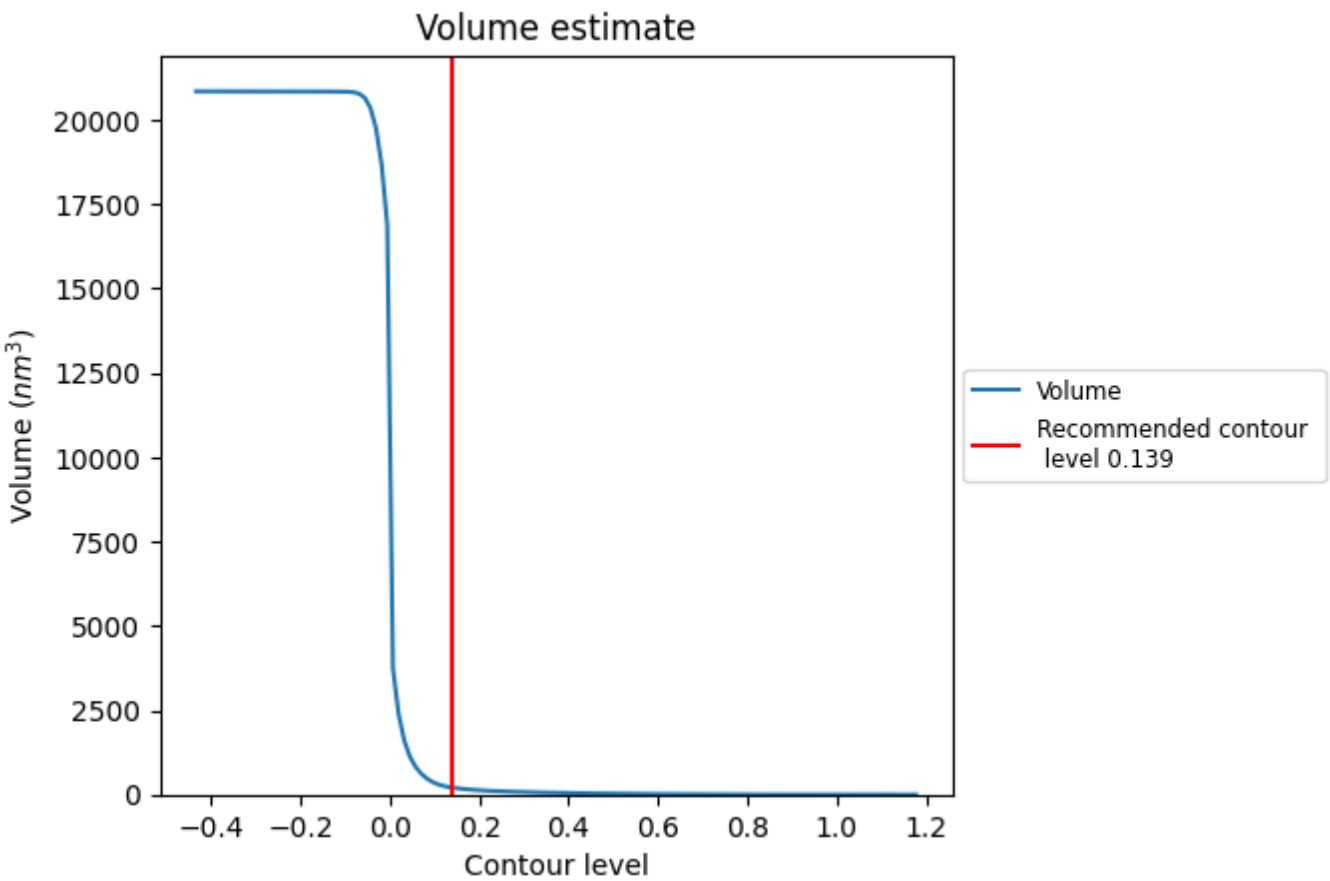
This section contains the results of statistical analysis of the map.

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.

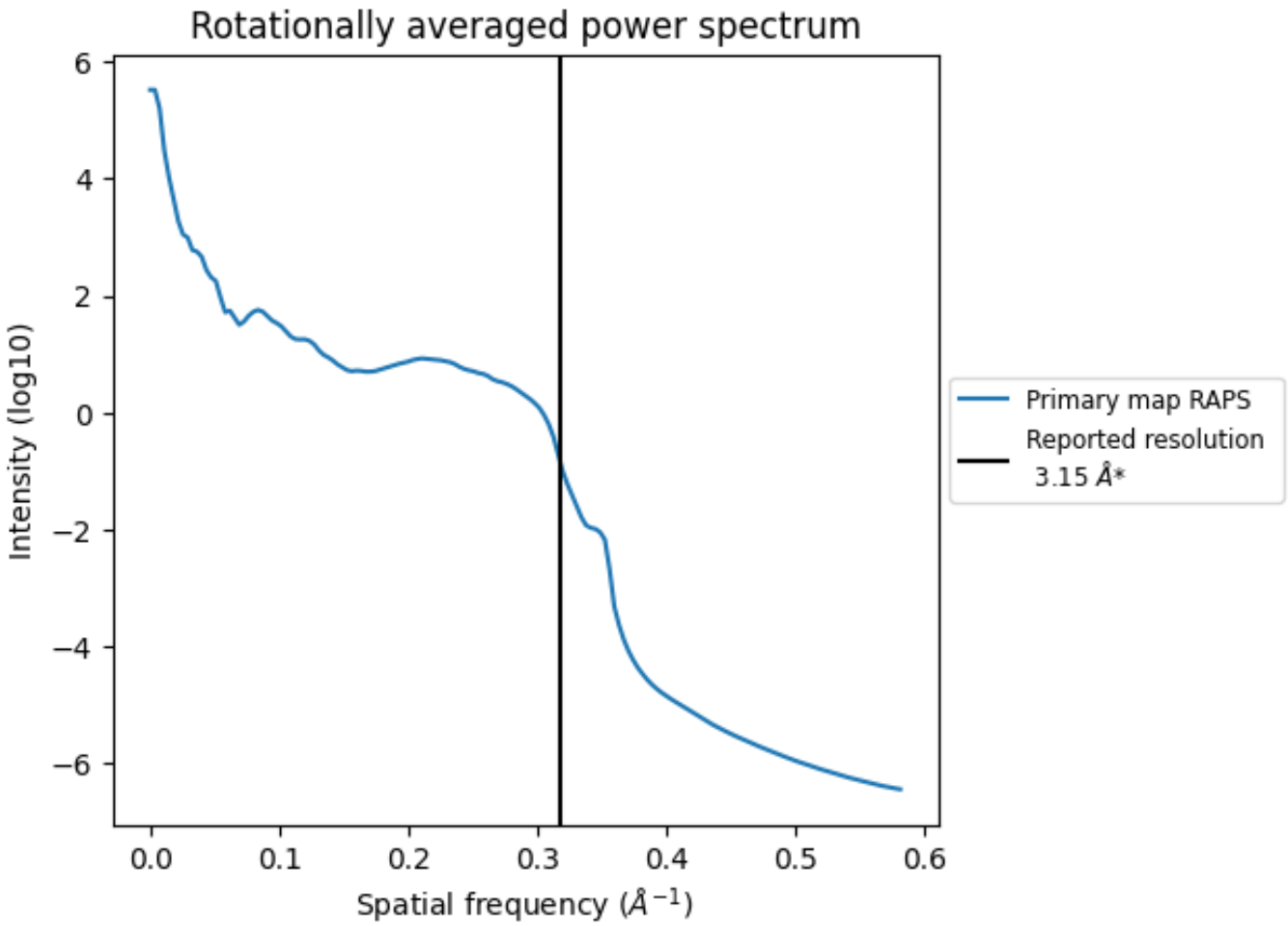
7.2 Volume estimate (i)



The volume at the recommended contour level is 207 nm^3; this corresponds to an approximate mass of 187 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 0.317 Å^{-1}

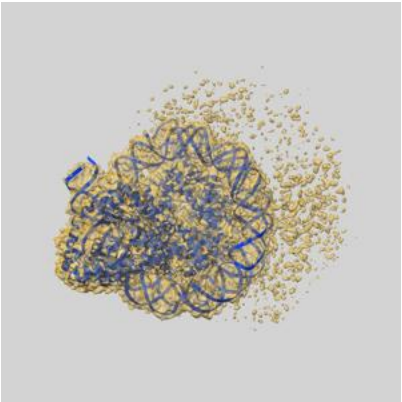
8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit (i)

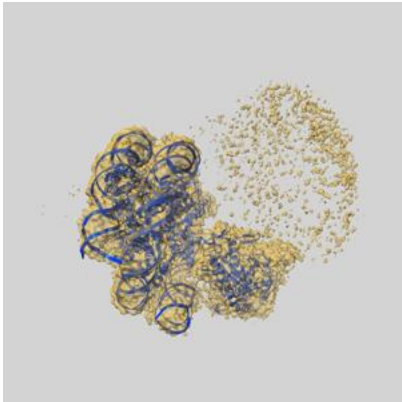
This section contains information regarding the fit between EMDB map EMD-10695 and PDB model 6Y5E. Per-residue inclusion information can be found in section 3 on page 7.

9.1 Map-model overlay (i)



X

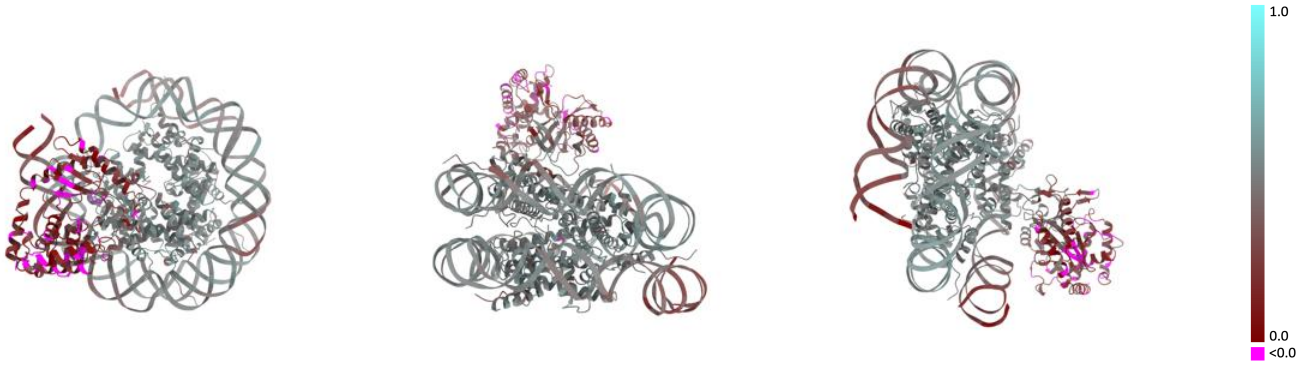
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.139 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

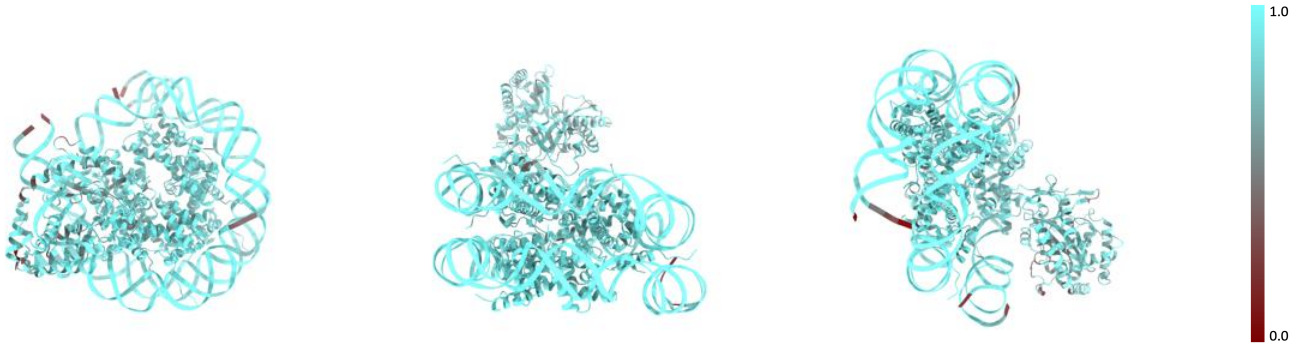
9.2 Q-score mapped to coordinate model (i)



X Y Z

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

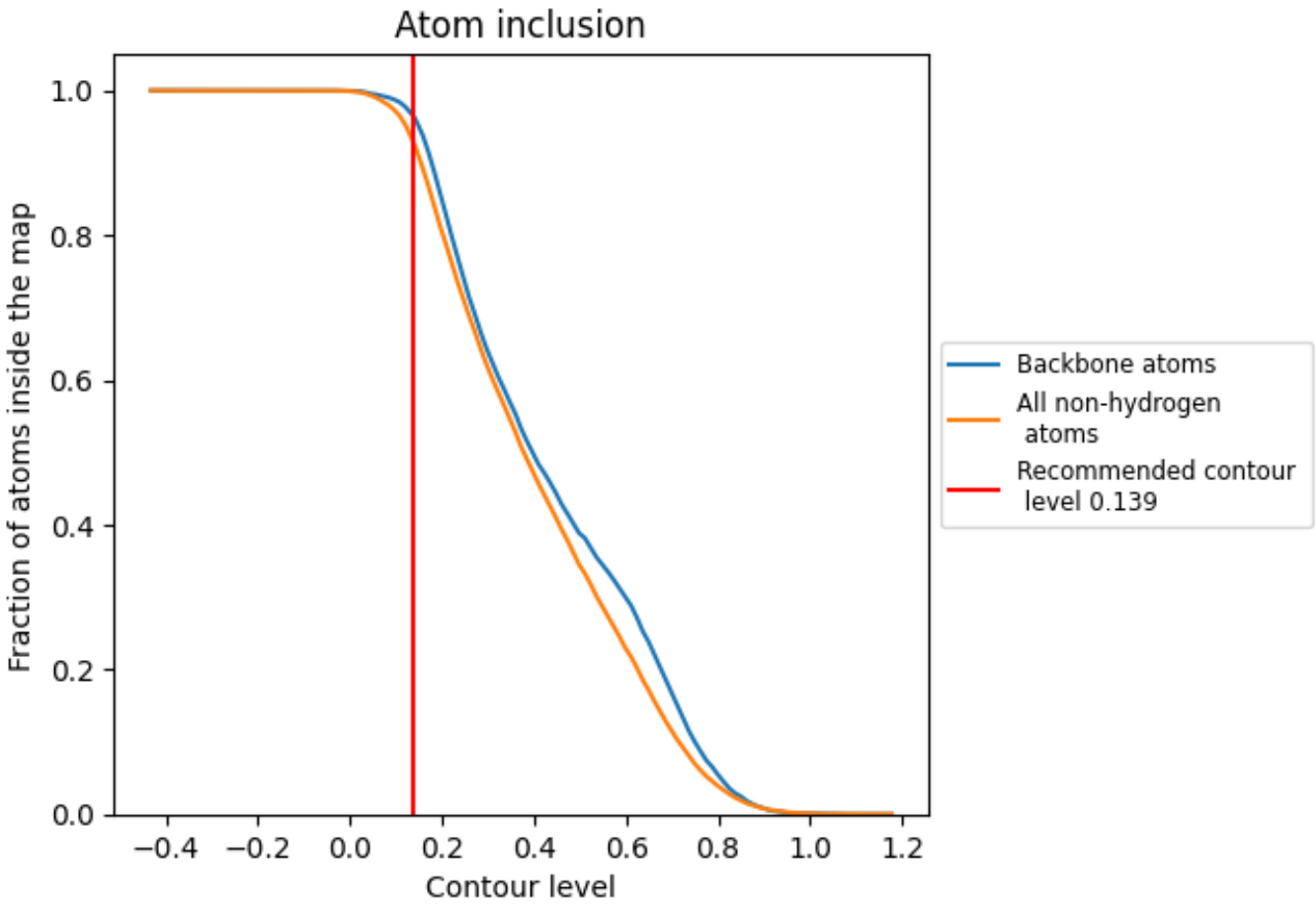
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.139).

9.4 Atom inclusion (i)



At the recommended contour level, 96% of all backbone atoms, 93% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.139) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9270	0.4300
A	0.9490	0.5230
B	0.9597	0.5130
C	0.9526	0.5080
D	0.9719	0.5100
E	0.9325	0.5140
F	0.9282	0.5190
G	0.9318	0.5130
H	0.9362	0.5030
I	0.9651	0.4570
J	0.9667	0.4540
K	0.8079	0.2100