



## wwPDB EM Validation Summary Report ⓘ

Feb 25, 2024 – 08:11 AM EST

PDB ID : 7JOA  
EMDB ID : EMD-22409  
Title : 2:1 cGAS-nucleosome complex  
Authors : Boyer, J.A.; Spangler, C.J.; Strauss, J.D.; Cesmat, A.P.; Liu, P.; McGinty, R.K.; Zhang, Q.  
Deposited on : 2020-08-06  
Resolution : 3.30 Å (reported)  
Based on initial models : 6FQ5, 4K8V

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

EMDB validation analysis : 0.0.1.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

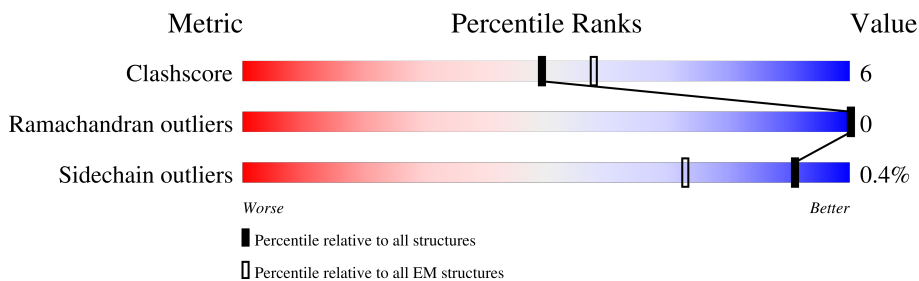
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.30 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  $\geq 3$ , 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  $\leq 5\%$ . 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	136	
1	E	136	
2	B	103	
2	F	103	
3	C	129	
3	G	129	
4	D	125	
4	H	125	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	I	147	 77% 22%
6	J	147	 64% 35%
7	K	366	 78% 79% 19%

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 14994 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	98	810	511	157	139	3	0	0
1	E	97	801	505	155	138	3	0	0

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	81	648	410	126	111	1	0	0
2	F	85	683	430	136	116	1	0	0

- Molecule 3 is a protein called Histone H2A type 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	C	109	837	529	164	144	0	0
3	G	109	837	529	164	144	0	0

- Molecule 4 is a protein called Histone H2B type 1-C/E/F/G/I.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	94	736	461	134	139	2	0	0
4	H	94	736	461	134	139	2	0	0

- Molecule 5 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	I	145	2990	1415	559	871	145	0	0

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	J	145	2955	1403	538	869	145	0	0

- Molecule 7 is a protein called Cyclic GMP-AMP synthase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	K	358	2960	1902	504	541	13	0	0

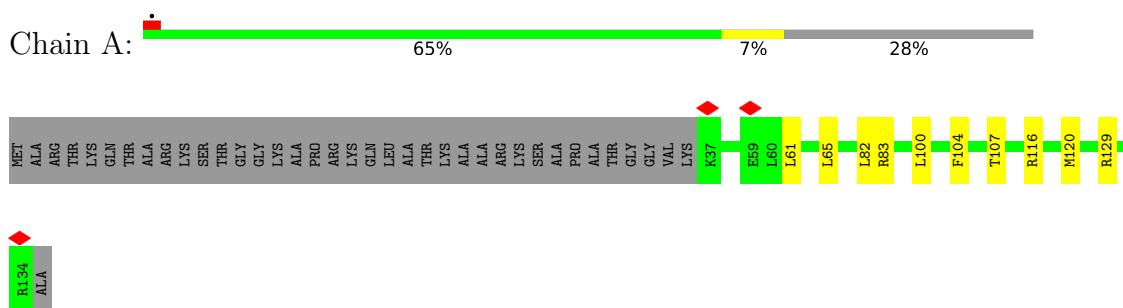
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

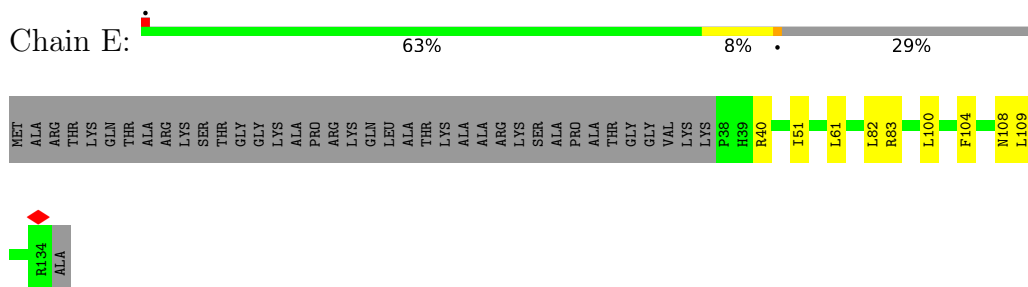
### 3 Residue-property plots [i](#)

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.

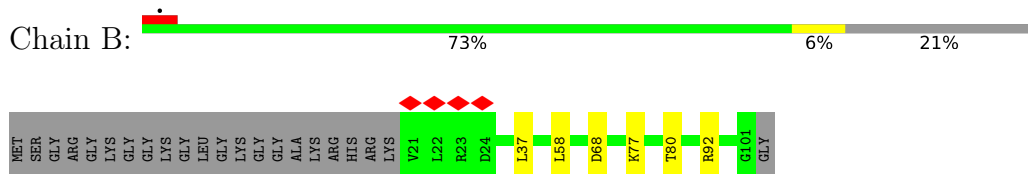
- Molecule 1: Histone H3.2



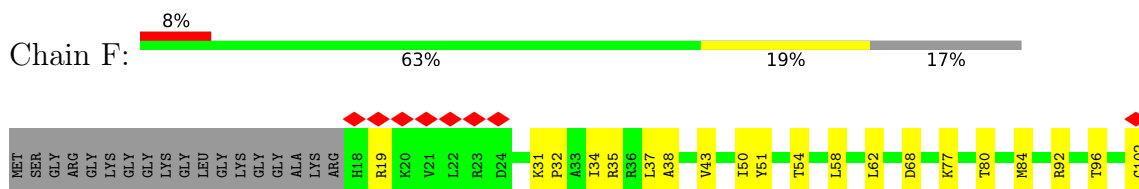
- Molecule 1: Histone H3.2



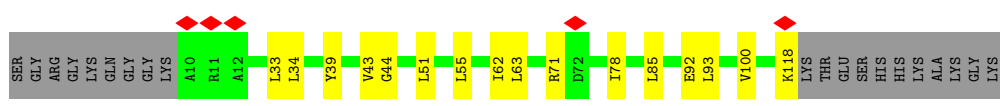
- Molecule 2: Histone H4



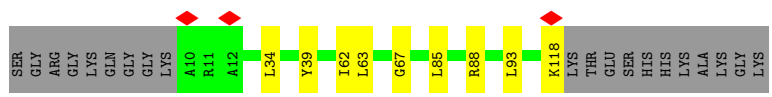
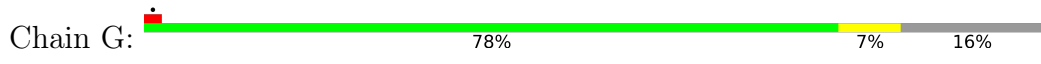
- Molecule 2: Histone H4



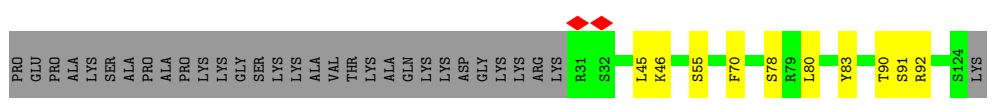
- Molecule 3: Histone H2A type 1



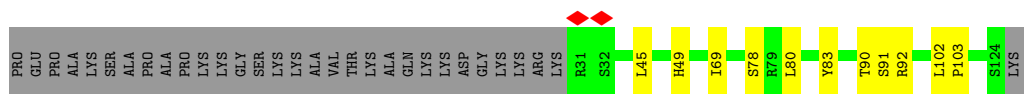
• Molecule 3: Histone H2A type 1



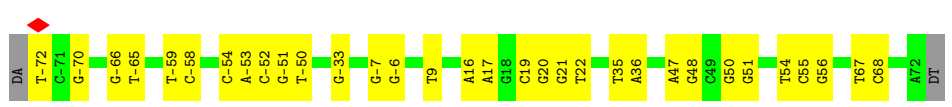
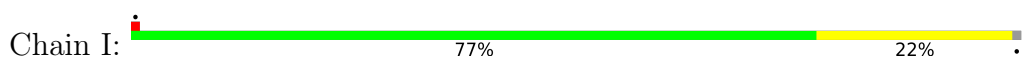
• Molecule 4: Histone H2B type 1-C/E/F/G/I



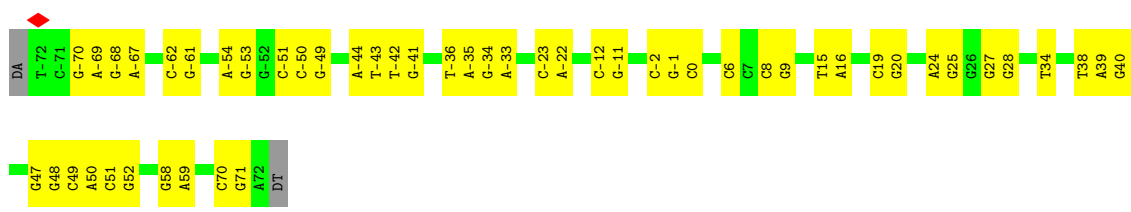
• Molecule 4: Histone H2B type 1-C/E/F/G/I



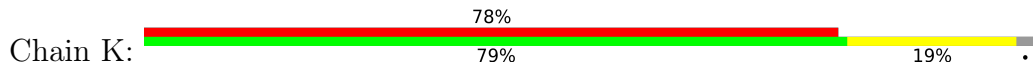
• Molecule 5: DNA (145-MER)



• Molecule 6: DNA (145-MER)



• Molecule 7: Cyclic GMP-AMP synthase



GLY	D148	K149	L150	K151	K152	V153	L154	D155	K156	L157	R158	L159	K160	R161	K162	D163	I164	S165	E166	A167	A168	E169	T170	V171	M172	K173	V174	V175	E176	R177	L178	L179	R180	R181	M182	Q183	K184	R185	E186	S187	E188	F189	K190	G191	V192	E193	Q194	L195	N196	T197	G198	S199	Y200	Y201				
E202	H203	V204	K205	I206	S207	A208	P209	N210	E211	F212	D213	V214	M215	F216	K217	L218	E219	I223	Q226	E227	Y228	Y229	E230	T231	G232	A233	F239	K240	R241	R244	G245	N246	P247	L248	S249	H250	F251	L252	E253	G254	E255	S258	A259	T260	L263	S264	K265	F266	R267	K268	I269	I270	K271					
E273	E274	V274	K275	E276	I277	K278	D279	L280	D281	V282	S283	V284	E285	K286	E287	K288	P289	G290	S291	P292	A293	V294	T295	L296	L297	I298	R299	N300	P301	E302	E303	I304	S305	V306	D307	I308	I309	L310	A311	L312	E313	S317	T322	K323	E324	G325	L326	P327	I328	Q329	G330	W331	L332	G333	T334	K350		
N351	A352	K353	D354	G355	N356	S357	F358	Q359	G360	E361	R364	L365	S366	F367	S368	H369	T370	E371	K372	Y373	I374	L375	N376	N377	H378	G379	I380	E381	K382	T383	C384	C385	E386	S387	S388	G389	A390	K391	C392	C393	R394	K395	E396	C397	L398	K399	L400	M401	K402	Y403	L404	L405	E406	Q407	L408	K409	K410	E411
F412	Q413	E414	L415	D416	A417	F418	C419	S420	Y421	H422	V423	K424	T425	A426	I427	F428	H429	M430	W431	T432	Q433	D434	F435	Q436	D437	S438	Q439	W440	D441	P442	R443	M444	L445	S446	S447	C448	F449	D450	K451	L452	L453	A454	F455	F456	L457	E458	C459	L460	R461	T462	E463	K464	L465	D466	H467	Y468	P471	K472
F473	M474	L475	F476	S477	Q478	E479	L480	L481	D482	R483	K484	S485	K486	E487	F488	L489	S490	K491	K492	L493	E494	Y495	E496	R497	M498	M499	G500	F501	P502	I503	F504	D505	LYS	LEU																								



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	45587	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 ( $e^-/\text{\AA}^2$ )	53	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.111	Depositor
Minimum map value	-0.051	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	300.30002, 300.30002, 300.30002	wwPDB
Map dimensions	330, 330, 330	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.91, 0.91, 0.91	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

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	0.51	0/822	0.55	0/1102
1	E	0.52	0/813	0.54	0/1090
2	B	0.54	0/655	0.53	0/878
2	F	0.53	0/691	0.54	0/923
3	C	0.48	0/847	0.54	0/1142
3	G	0.48	0/847	0.54	0/1142
4	D	0.55	0/747	0.53	0/1004
4	H	0.55	0/747	0.53	0/1004
5	I	0.87	0/3357	0.97	0/5184
6	J	0.89	0/3311	0.95	0/5103
7	K	0.30	0/3024	0.48	0/4060
All	All	0.67	0/15861	0.75	0/22632

There are no bond length outliers.

There are no bond angle outliers.

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	810	0	851	8	0
1	E	801	0	839	11	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	648	0	693	6	0
2	F	683	0	729	18	0
3	C	837	0	902	14	0
3	G	837	0	902	8	0
4	D	736	0	756	9	0
4	H	736	0	756	8	0
5	I	2990	0	1628	22	0
6	J	2955	0	1627	40	0
7	K	2960	0	3003	49	0
8	K	1	0	0	0	0
All	All	14994	0	12686	168	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 6.

The worst 5 of 168 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:71:ARG:HH22	7:K:317:SER:HA	1.41	0.83
3:C:92:GLU:OE1	7:K:241:ARG:NH1	2.21	0.73
1:E:129:ARG:O	1:E:129:ARG:NH1	2.23	0.71
7:K:274:VAL:HG11	7:K:284:VAL:HG23	1.72	0.71
7:K:427:ILE:HG23	7:K:431:TRP:CZ3	2.27	0.70

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	96/136 (71%)	91 (95%)	5 (5%)	0	100	100
1	E	95/136 (70%)	89 (94%)	6 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	79/103 (77%)	69 (87%)	10 (13%)	0	100	100
2	F	83/103 (81%)	73 (88%)	10 (12%)	0	100	100
3	C	107/129 (83%)	102 (95%)	5 (5%)	0	100	100
3	G	107/129 (83%)	102 (95%)	5 (5%)	0	100	100
4	D	92/125 (74%)	83 (90%)	9 (10%)	0	100	100
4	H	92/125 (74%)	83 (90%)	9 (10%)	0	100	100
7	K	356/366 (97%)	343 (96%)	13 (4%)	0	100	100
All	All	1107/1352 (82%)	1035 (94%)	72 (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	
1	A	86/111 (78%)	85 (99%)	1 (1%)	71	83
1	E	85/111 (77%)	84 (99%)	1 (1%)	71	83
2	B	67/79 (85%)	67 (100%)	0	100	100
2	F	70/79 (89%)	70 (100%)	0	100	100
3	C	84/98 (86%)	83 (99%)	1 (1%)	71	83
3	G	84/98 (86%)	83 (99%)	1 (1%)	71	83
4	D	81/105 (77%)	81 (100%)	0	100	100
4	H	81/105 (77%)	81 (100%)	0	100	100
7	K	331/338 (98%)	331 (100%)	0	100	100
All	All	969/1124 (86%)	965 (100%)	4 (0%)	91	95

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	129	ARG

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	118	LYS
1	E	129	ARG
3	G	118	LYS

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

### 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 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 5.7 Other polymers [i](#)

There are no such residues in this entry.

### 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

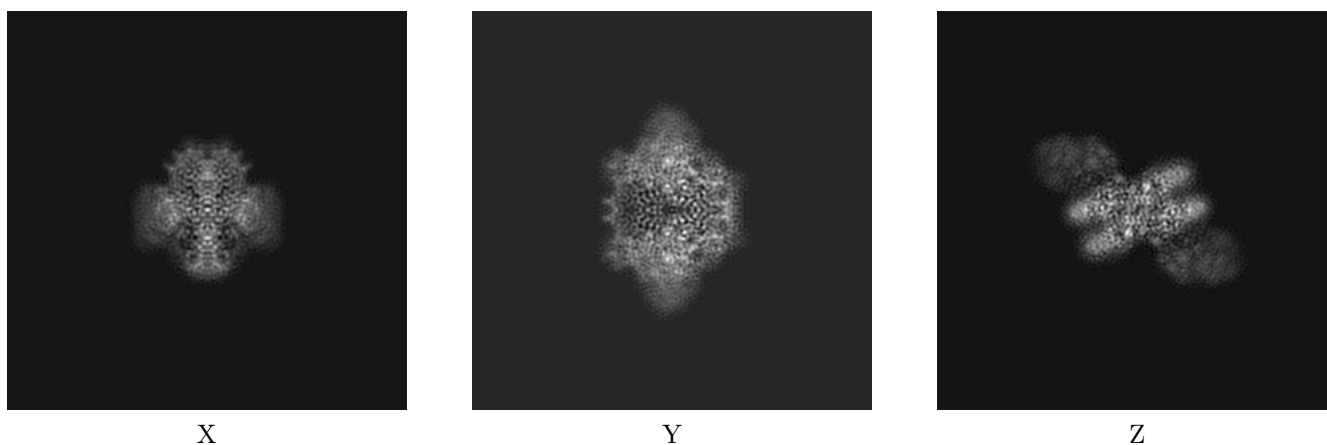
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22409. 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](#)

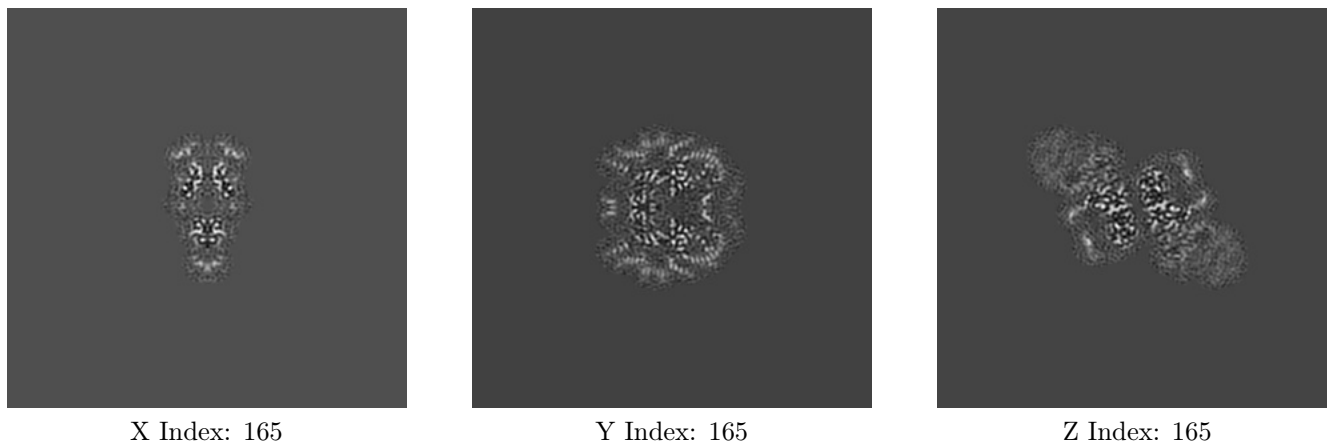
#### 6.1.1 Primary map



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

### 6.2 Central slices [i](#)

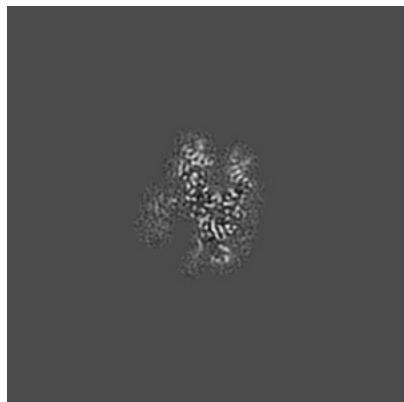
#### 6.2.1 Primary map



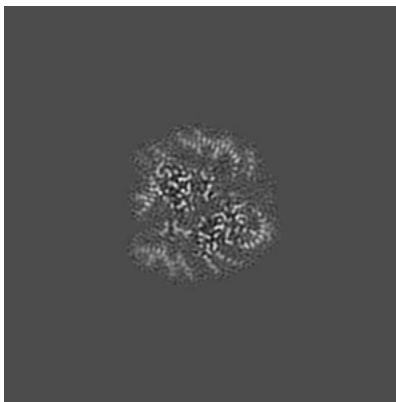
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

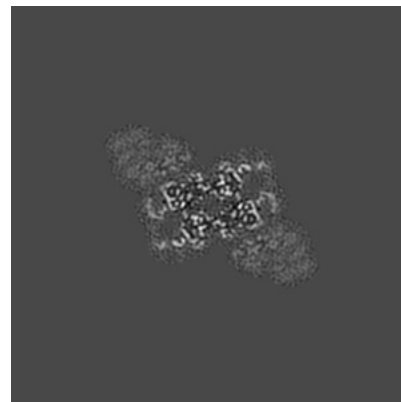
### 6.3.1 Primary map



X Index: 187



Y Index: 170

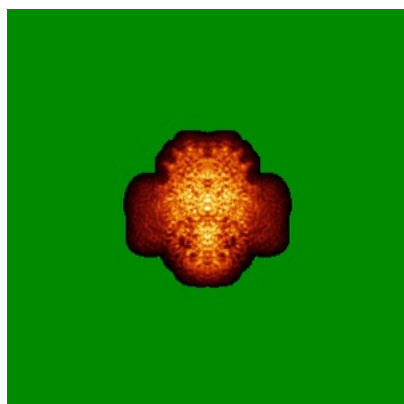


Z Index: 173

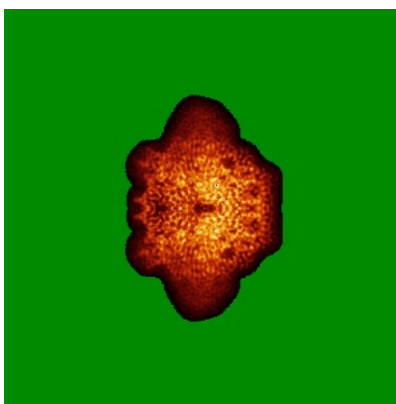
The images above show the largest variance slices of the map in three orthogonal directions.

## 6.4 Orthogonal standard-deviation projections (False-color) [i](#)

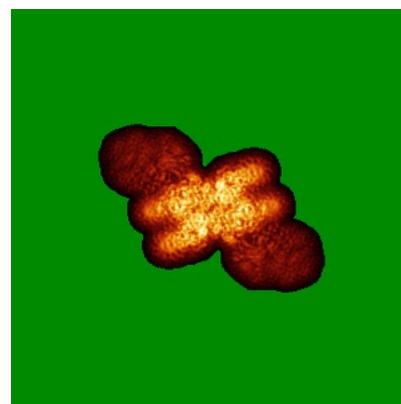
### 6.4.1 Primary map



X



Y



Z

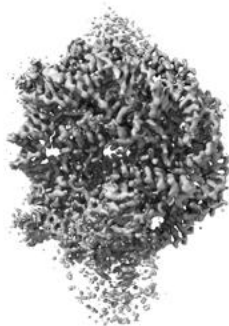
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

## 6.5 Orthogonal surface views [i](#)

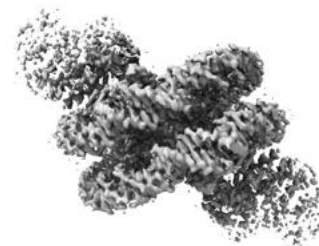
### 6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

## 6.6 Mask visualisation [i](#)

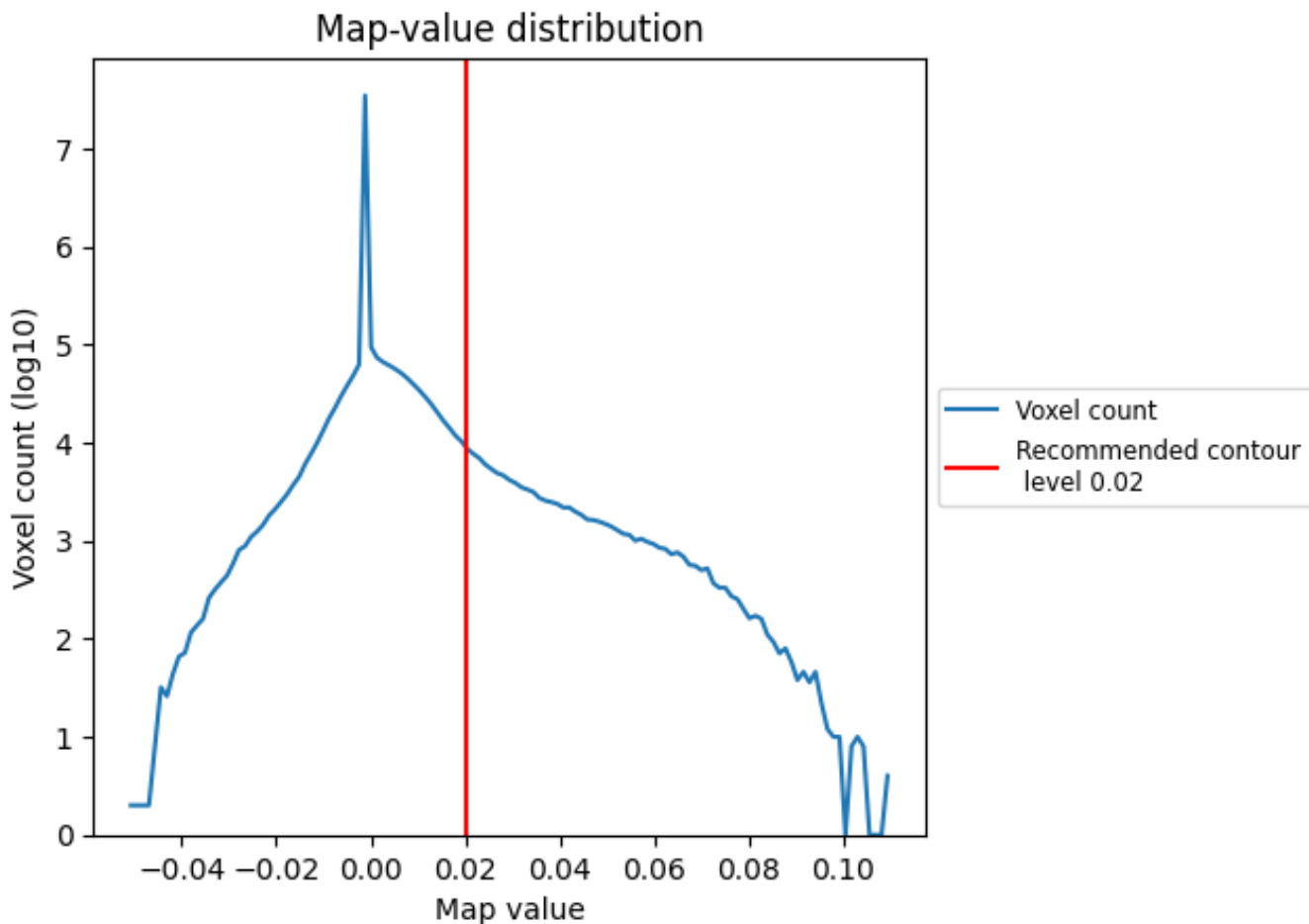
This section was not generated. No masks/segmentation were deposited.



## 7 Map analysis [i](#)

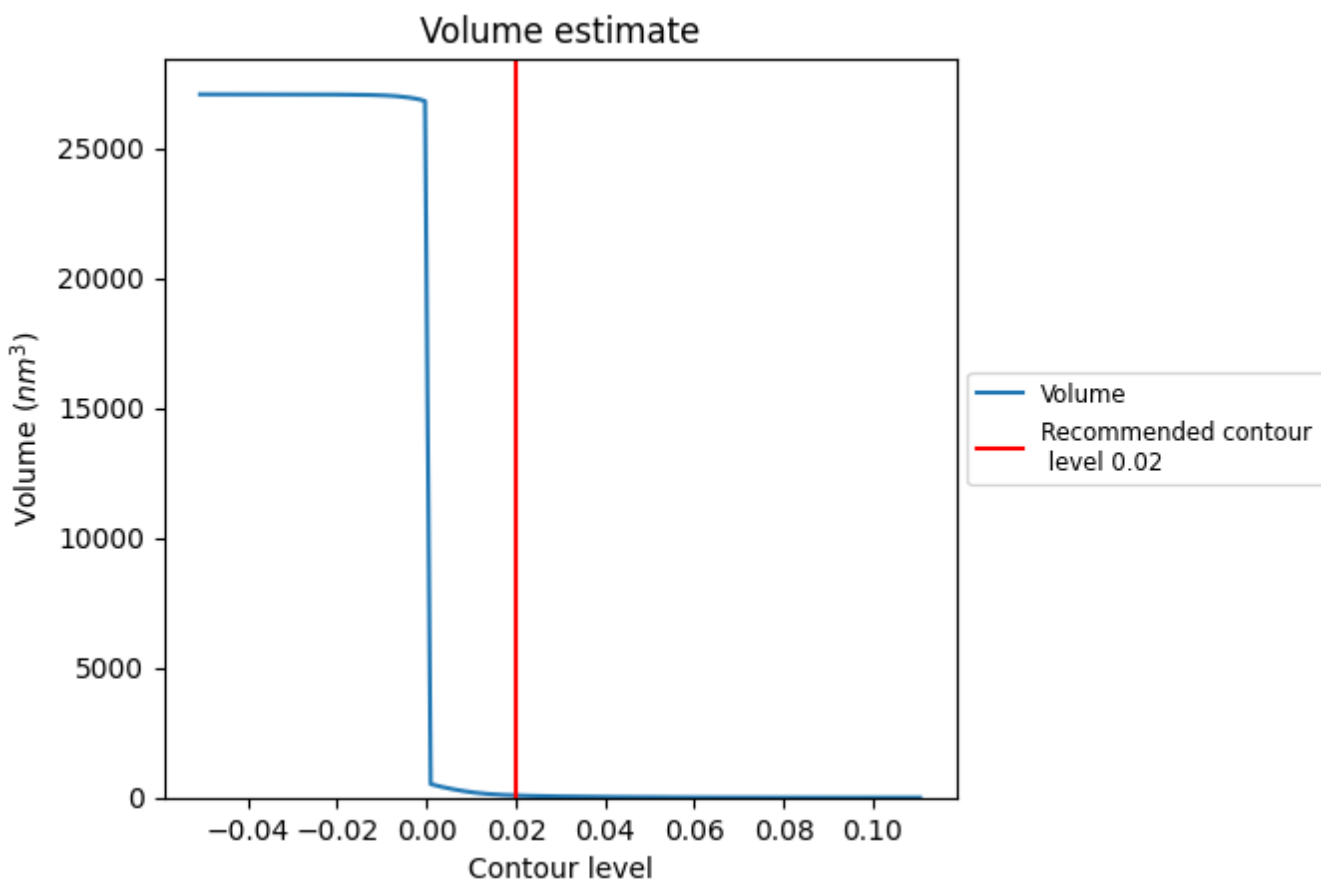
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

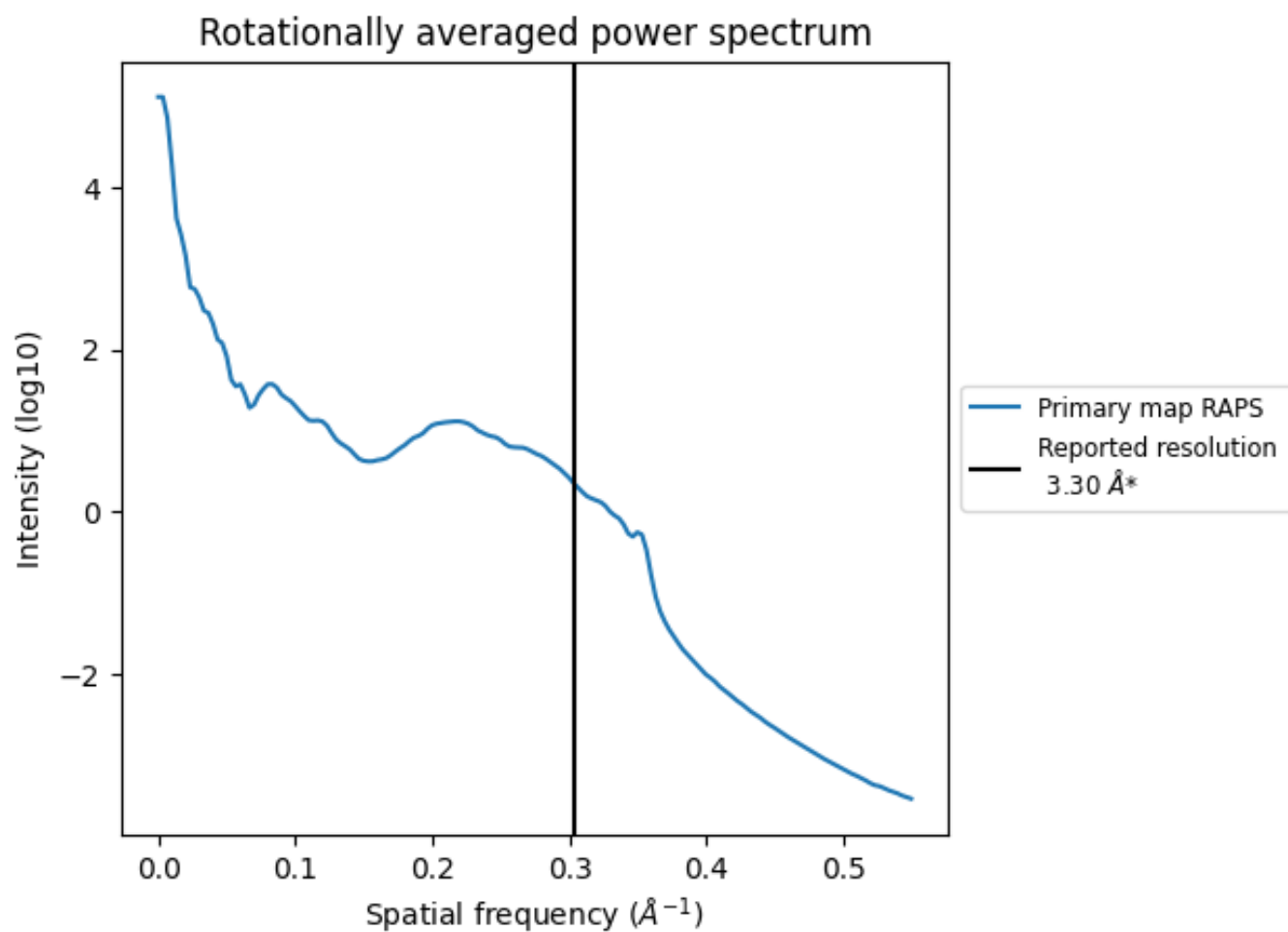
## 7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 82 nm<sup>3</sup>; this corresponds to an approximate mass of 74 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](#)

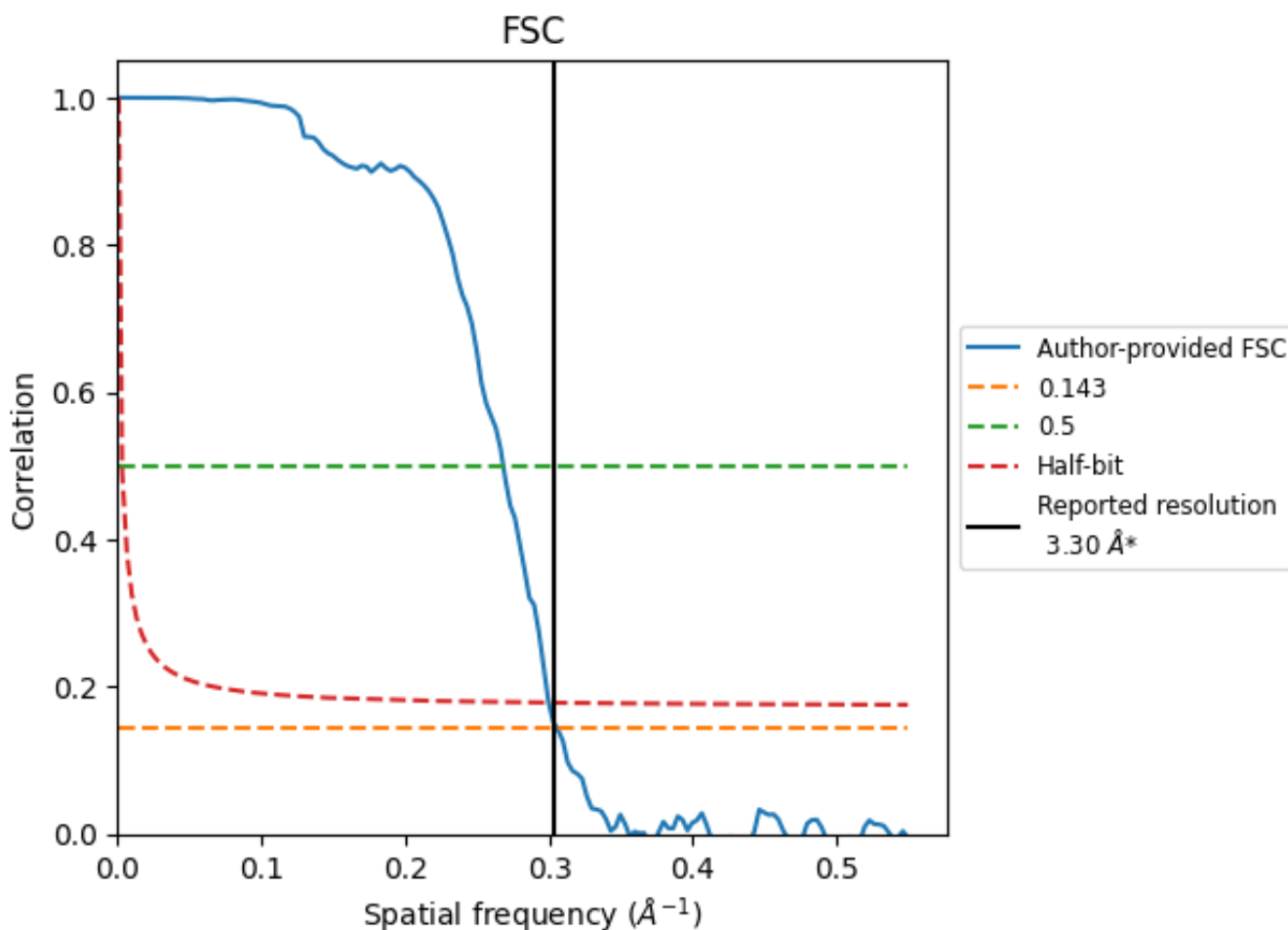


\*Reported resolution corresponds to spatial frequency of  $0.303 \text{ \AA}^{-1}$

## 8 Fourier-Shell correlation [i](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

### 8.1 FSC [i](#)



\*Reported resolution corresponds to spatial frequency of 0.303 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

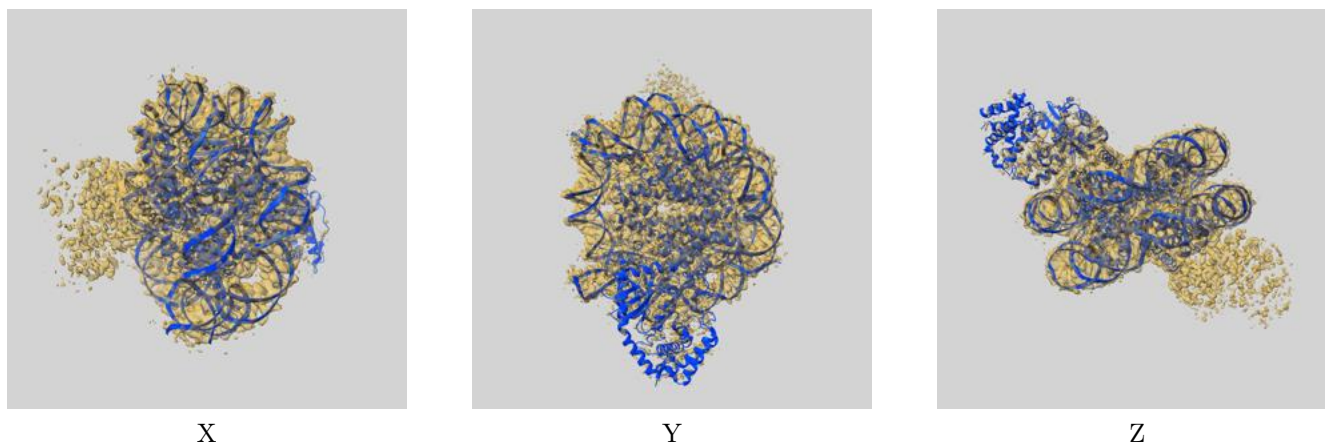
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.30	-	-
Author-provided FSC curve	3.27	3.73	3.34
Unmasked-calculated*	-	-	-

\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

## 9 Map-model fit [i](#)

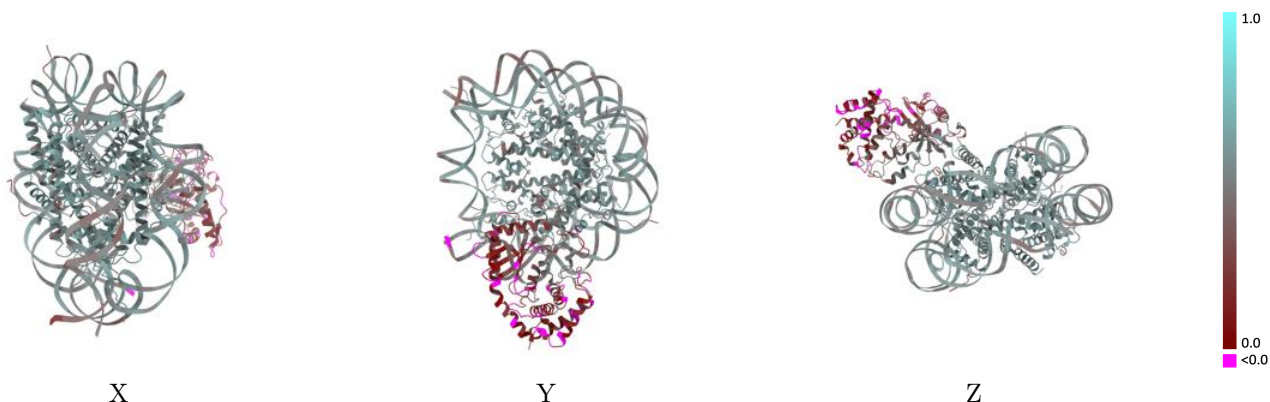
This section contains information regarding the fit between EMDB map EMD-22409 and PDB model 7JOA. Per-residue inclusion information can be found in section 3 on page 6.

### 9.1 Map-model overlay [i](#)



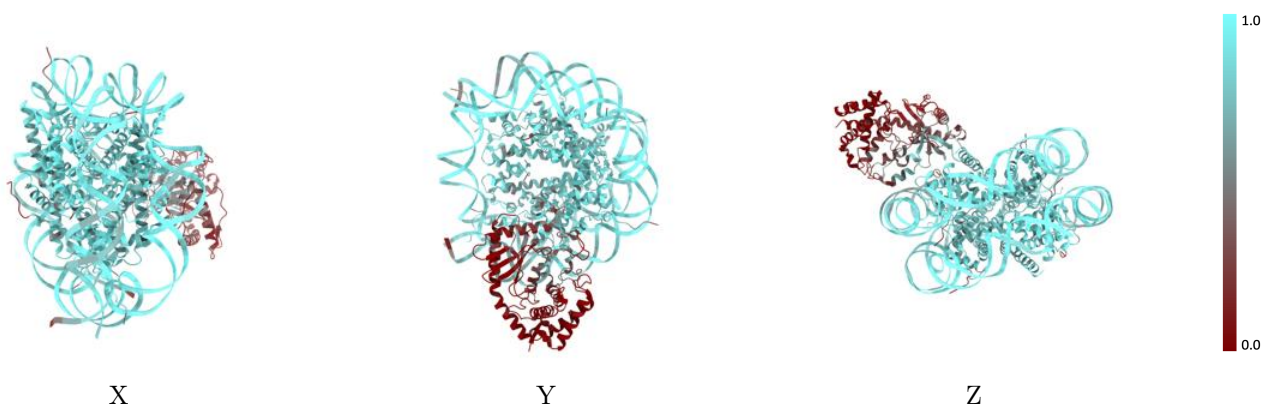
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2 Q-score mapped to coordinate model [i](#)



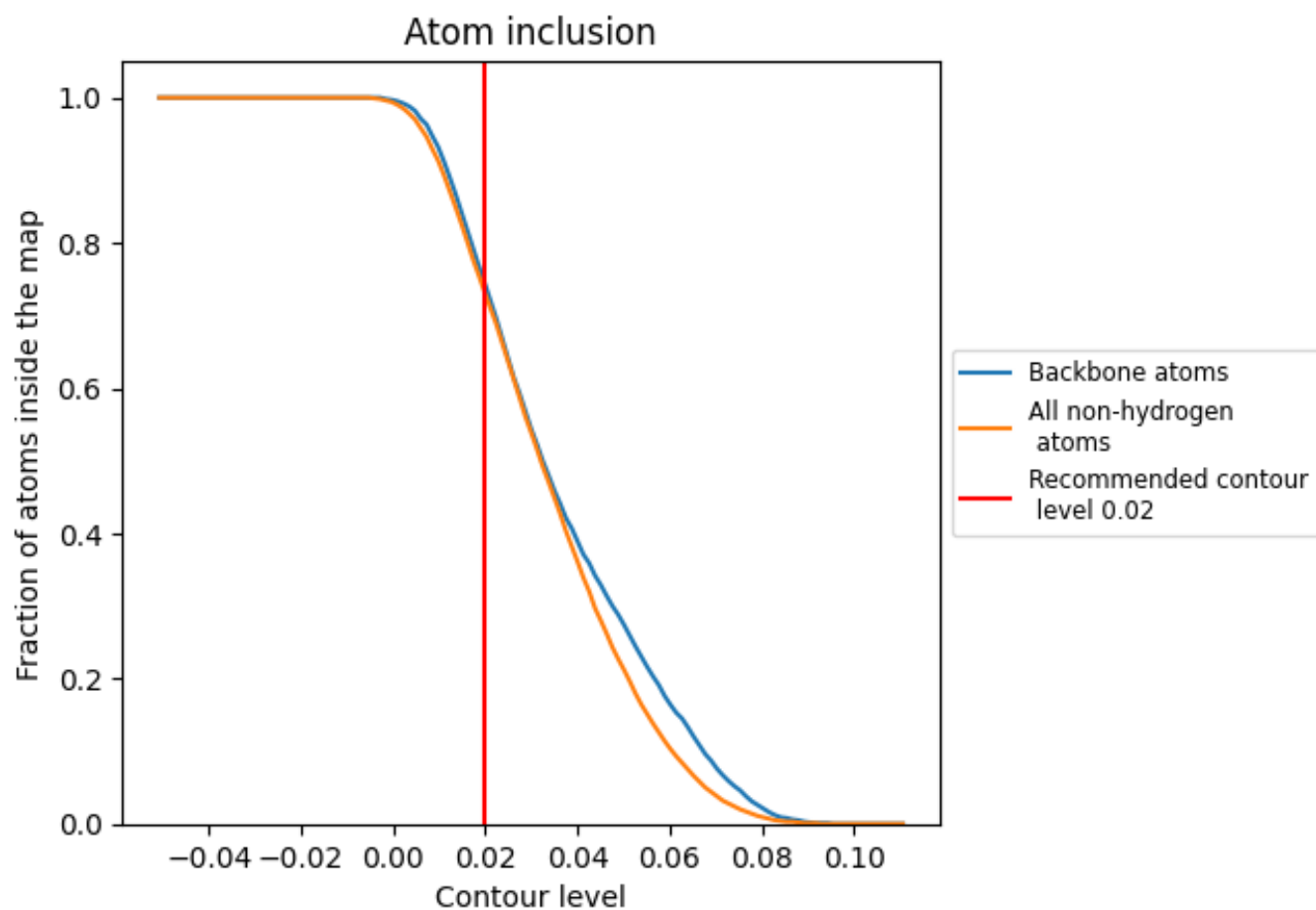
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions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



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 (0.02).

## 9.4 Atom inclusion [i](#)



























At the recommended contour level, 74% of all backbone atoms, 73% 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 (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7290	 0.4740
A	 0.8560	 0.5620
B	 0.8490	 0.5610
C	 0.8560	 0.5460
D	 0.8580	 0.5420
E	 0.8640	 0.5670
F	 0.8090	 0.5350
G	 0.8530	 0.5540
H	 0.8590	 0.5380
I	 0.8670	 0.5230
J	 0.8590	 0.5050
K	 0.2060	 0.2360

