



wwPDB EM Validation Summary Report ⓘ

Mar 5, 2026 – 07:28 PM UTC

PDB ID : 6GH5 / pdb_00006gh5
EMDB ID : EMD-0001
Title : Cryo-EM structure of bacterial RNA polymerase-sigma54 holoenzyme transcription open complex
Authors : Glyde, R.; Ye, F.Z.; Zhang, X.D.
Deposited on : 2018-05-04
Resolution : 3.40 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ 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:

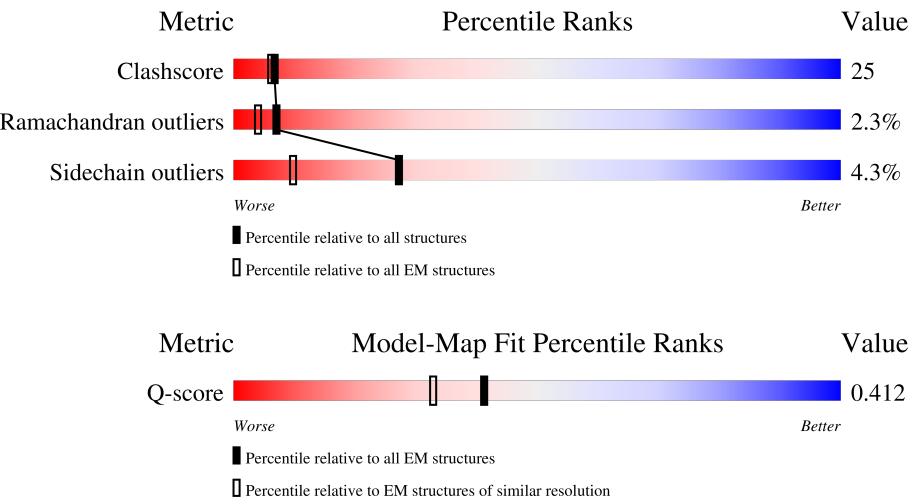
EMDB validation analysis : 0.0.1.dev132
MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : 202505.v01 (Using data in the EMDb archive up until May 2025)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
Q-score	-	25397	14717 (2.90 - 3.90)

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 $\leq 5\%$. 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	329	49%29%14%6%
1	B	329	38%28%5%29%
2	C	1342	46%34%17%
3	D	1407	41%33%18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	91	43% 26% 12% • 18%
5	M	497	• 53% 13% • 32%
6	F	63	6% 16% 43% 14% 27%
7	G	63	5% 30% 41% • 27%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 28316 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 DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	309	Total	C	N	O	S	0	0
			2302	1441	400	454	7		
1	B	235	Total	C	N	O	S	0	0
			1733	1085	301	341	6		

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1341	Total	C	N	O	S	0	0
			10034	6289	1746	1961	38		

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	1345	Total	C	N	O	S	0	0
			9790	6144	1746	1858	42		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	75	Total	C	N	O	S	0	0
			565	347	110	107	1		

- Molecule 5 is a protein called RNA polymerase sigma-54 factor,RNA polymerase sigma-54 factor,RNA polymerase sigma-54 factor,RNA polymerase sigma-54 factor,RNA polymerase sigma-54 factor.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	M	338	Total	C	N	O	S	0	0
			2002	1243	356	400	3		

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	-19	MET	-	initiating methionine	UNP A0A0J4U551
M	-18	GLY	-	expression tag	UNP A0A0J4U551
M	-17	SER	-	expression tag	UNP A0A0J4U551
M	-16	SER	-	expression tag	UNP A0A0J4U551
M	-15	HIS	-	expression tag	UNP A0A0J4U551
M	-14	HIS	-	expression tag	UNP A0A0J4U551
M	-13	HIS	-	expression tag	UNP A0A0J4U551
M	-12	HIS	-	expression tag	UNP A0A0J4U551
M	-11	HIS	-	expression tag	UNP A0A0J4U551
M	-10	HIS	-	expression tag	UNP A0A0J4U551
M	-9	SER	-	expression tag	UNP A0A0J4U551
M	-8	SER	-	expression tag	UNP A0A0J4U551
M	-7	GLY	-	expression tag	UNP A0A0J4U551
M	-6	LEU	-	expression tag	UNP A0A0J4U551
M	-5	VAL	-	expression tag	UNP A0A0J4U551
M	-4	PRO	-	expression tag	UNP A0A0J4U551
M	-3	ARG	-	expression tag	UNP A0A0J4U551
M	-2	GLY	-	expression tag	UNP A0A0J4U551
M	-1	SER	-	expression tag	UNP A0A0J4U551
M	0	HIS	-	expression tag	UNP A0A0J4U551
M	336	ALA	ARG	engineered mutation	UNP A0A0J4U551

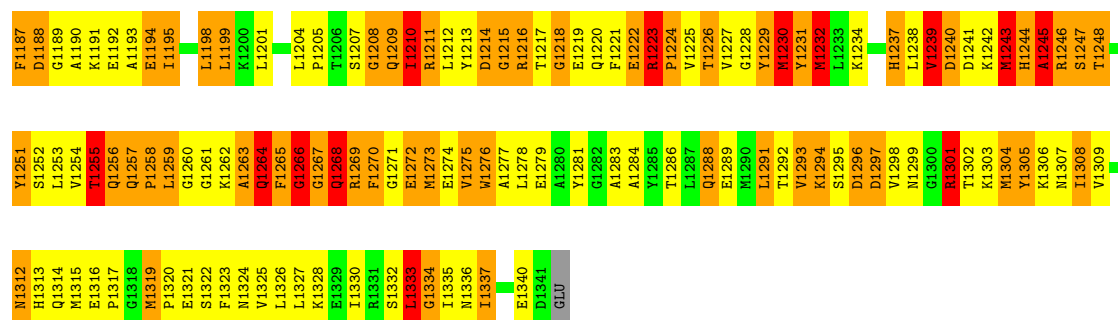
- Molecule 6 is a DNA chain called nifH promoter template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	46	Total	C	N	O	P	0	0
			944	445	182	271	46		

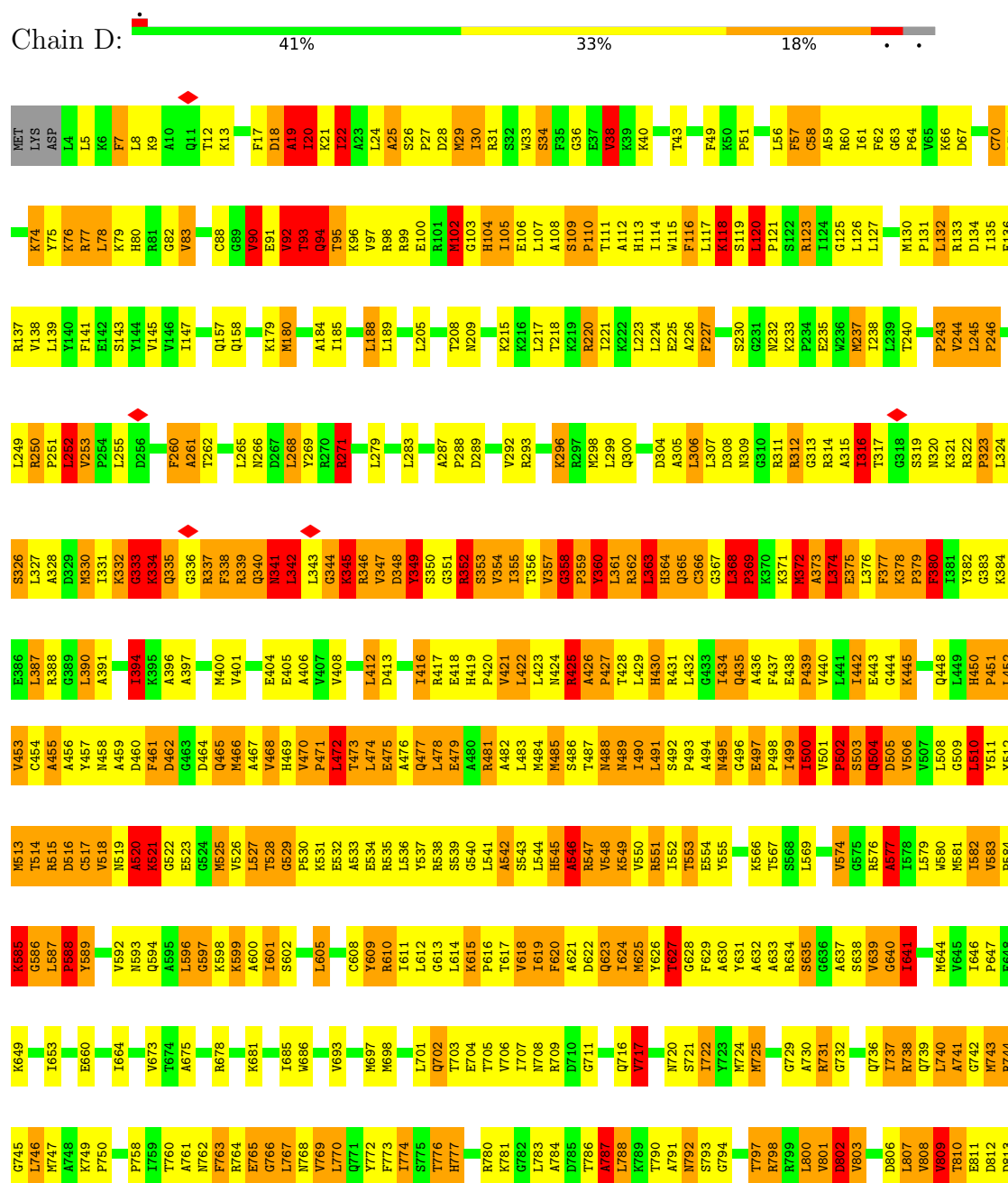
- Molecule 7 is a DNA chain called nifH promoter non-template DNA.

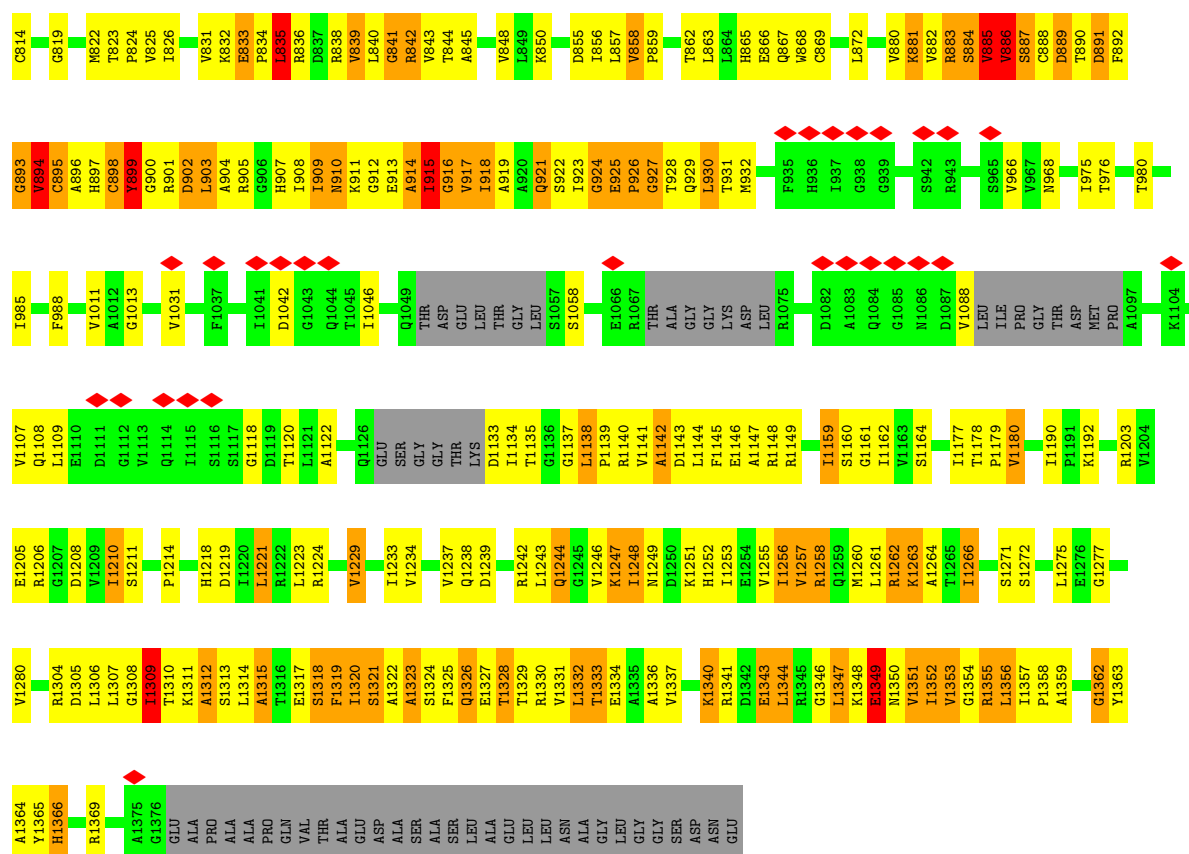
Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	46	Total	C	N	O	P	0	0
			946	448	173	279	46		





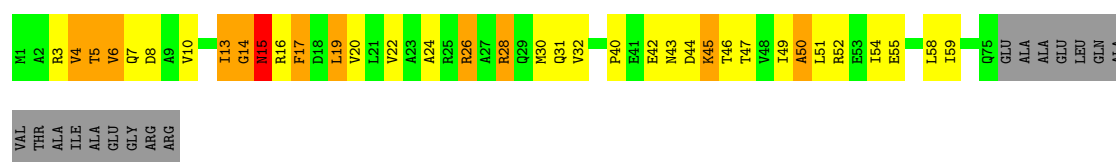
• Molecule 3: DNA-directed RNA polymerase subunit beta'





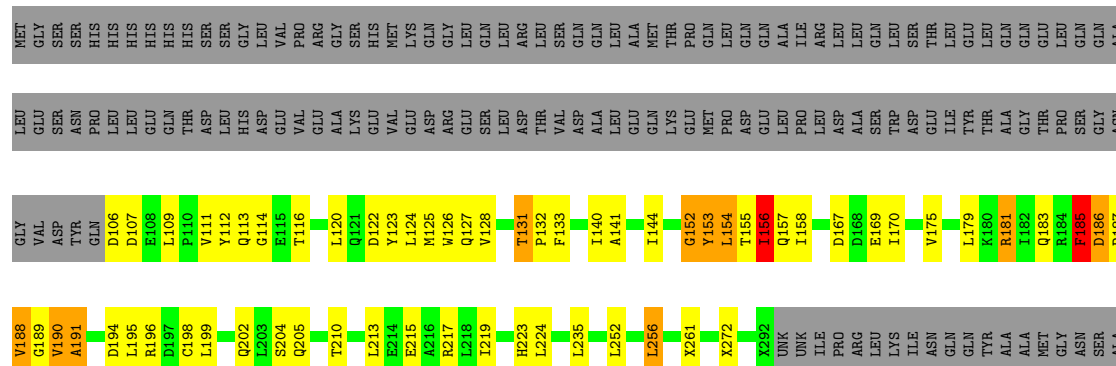
• Molecule 4: DNA-directed RNA polymerase subunit omega

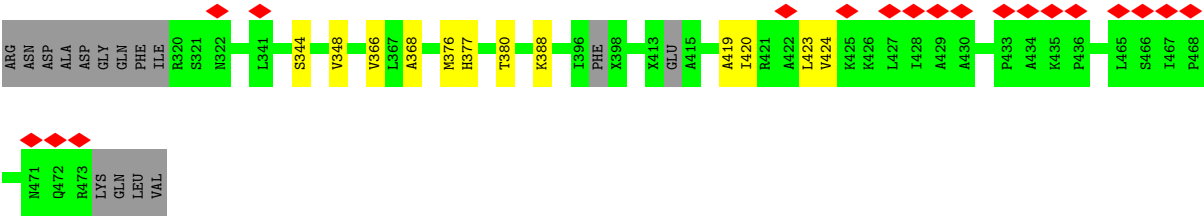
Chain E: 43% 26% 12% 18%



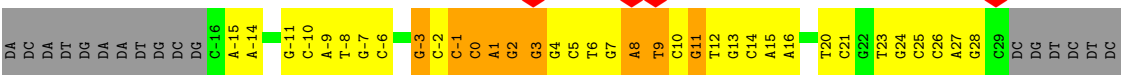
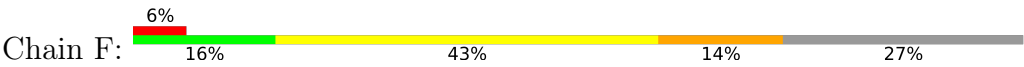
• Molecule 5: RNA polymerase sigma-54 factor, RNA polymerase sigma-54 factor, RNA polymerase sigma-54 factor, RNA polymerase sigma-54 factor, RNA polymerase sigma-54 factor

Chain M: 53% 13% 32%

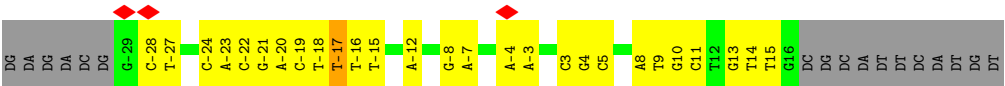




• Molecule 6: nifH promoter template DNA



• Molecule 7: nifH promoter non-template DNA



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	79678	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 ($e^-/\text{\AA}^2$)	45	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.227	Depositor
Minimum map value	-0.108	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.018	Depositor
Map size (Å)	271.36, 271.36, 271.36	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	Depositor

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	2.02	85/2331 (3.6%)	1.82	59/3177 (1.9%)
1	B	1.58	19/1752 (1.1%)	1.50	21/2385 (0.9%)
2	C	2.45	641/10187 (6.3%)	2.16	558/13822 (4.0%)
3	D	2.41	562/9923 (5.7%)	2.17	546/13482 (4.0%)
4	E	1.69	11/567 (1.9%)	1.85	18/767 (2.3%)
5	M	1.27	14/1764 (0.8%)	1.35	22/2445 (0.9%)
6	F	0.51	1/1060 (0.1%)	1.31	19/1633 (1.2%)
7	G	0.39	0/1060	0.75	1/1635 (0.1%)
All	All	2.19	1333/28644 (4.7%)	1.98	1244/39346 (3.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	6
2	C	0	28
3	D	0	28
4	E	0	1
5	M	0	1
All	All	0	64

The worst 5 of 1333 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	526	HIS	C-O	-17.73	1.03	1.24
2	C	555	TYR	C-O	17.35	1.44	1.24
2	C	1225	VAL	C-O	17.28	1.41	1.24
3	D	333	GLY	N-CA	17.13	1.70	1.45
3	D	99	ARG	CZ-NH1	16.91	1.56	1.32

The worst 5 of 1244 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	462	ASP	N-CA-C	-18.13	86.00	113.02
3	D	765	GLU	CA-C-N	-16.20	100.22	120.64
3	D	765	GLU	C-N-CA	-16.20	100.22	120.64
1	A	61	ILE	CB-CG1-CD1	-16.10	79.98	113.80
2	C	805	MET	CG-SD-CE	-16.07	65.55	100.90

There are no chirality outliers.

5 of 64 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	131	CYS	Mainchain
1	A	151	GLY	Mainchain
1	A	47	LEU	Mainchain
1	A	63	GLY	Peptide
1	A	93	GLN	Peptide

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2302	0	2266	95	0
1	B	1733	0	1720	82	0
2	C	10034	0	9668	455	0
3	D	9790	0	9539	604	0
4	E	565	0	563	27	0
5	M	2002	0	1412	64	0
6	F	944	0	513	100	0
7	G	946	0	518	42	0
All	All	28316	0	26199	1340	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 25.

The worst 5 of 1340 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)
7:G:14:DT:H2''	7:G:15:DT:C7	1.31	1.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:1260:GLY:CA	3:D:346:ARG:HH12	1.20	1.52
2:C:1066:MET:CE	2:C:1232:MET:HE2	1.38	1.52
3:D:333:GLY:N	3:D:333:GLY:CA	1.70	1.51
3:D:316:ILE:CD1	3:D:317:THR:H	1.23	1.50

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	305/329 (93%)	269 (88%)	32 (10%)	4 (1%)	9	33
1	B	233/329 (71%)	209 (90%)	22 (9%)	2 (1%)	14	42
2	C	1339/1342 (100%)	1183 (88%)	128 (10%)	28 (2%)	5	24
3	D	1335/1407 (95%)	1127 (84%)	169 (13%)	39 (3%)	3	19
4	E	73/91 (80%)	68 (93%)	3 (4%)	2 (3%)	4	20
5	M	283/497 (57%)	236 (83%)	40 (14%)	7 (2%)	4	21
All	All	3568/3995 (89%)	3092 (87%)	394 (11%)	82 (2%)	7	23

5 of 82 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	ILE
2	C	29	SER
2	C	347	ILE
2	C	541	GLU
2	C	1155	VAL

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	244/286 (85%)	235 (96%)	9 (4%)	30	55
1	B	183/286 (64%)	176 (96%)	7 (4%)	29	54
2	C	1029/1157 (89%)	987 (96%)	42 (4%)	27	52
3	D	961/1168 (82%)	911 (95%)	50 (5%)	21	48
4	E	56/75 (75%)	55 (98%)	1 (2%)	51	67
5	M	113/393 (29%)	111 (98%)	2 (2%)	51	67
All	All	2586/3365 (77%)	2475 (96%)	111 (4%)	27	51

5 of 111 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1319	MET
5	M	156	ILE
3	D	249	LEU
5	M	140	ILE
3	D	886	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	335	GLN
3	D	792	ASN
3	D	364	HIS
3	D	477	GLN
3	D	907	HIS

5.3.3 RNA ⓘ

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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.

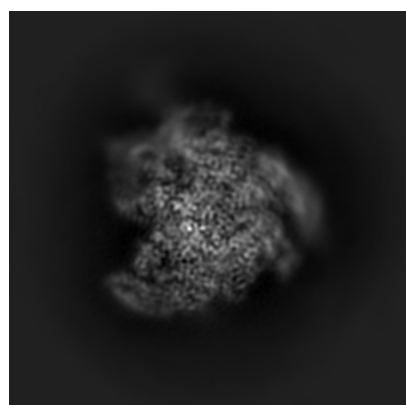
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-0001. These allow visual inspection of the internal detail of the map and identification of artifacts.

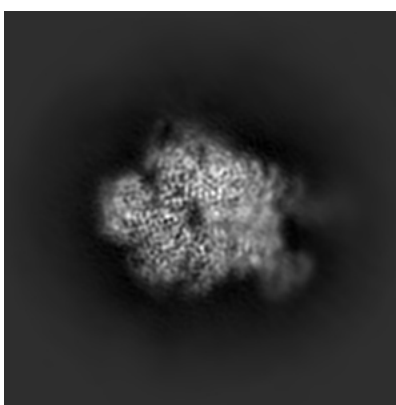
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

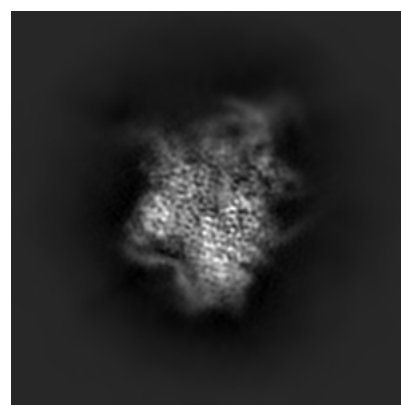
6.1.1 Primary map



X



Y

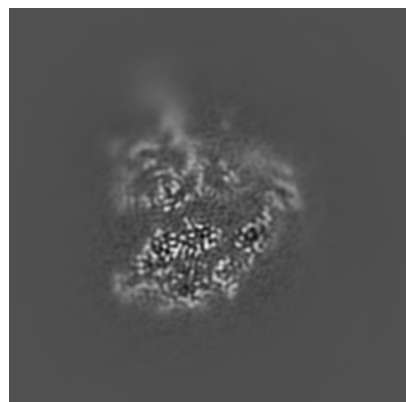


Z

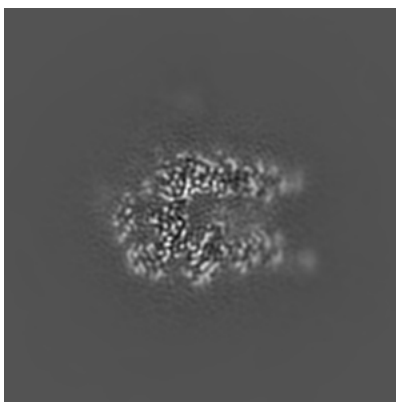
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

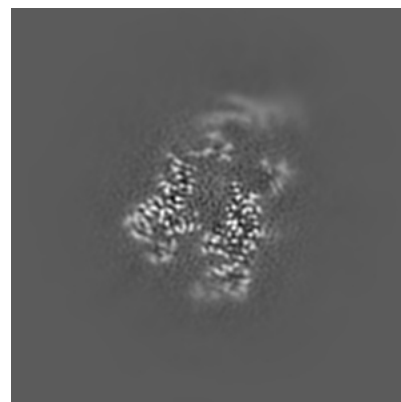
6.2.1 Primary map



X Index: 128



Y Index: 128

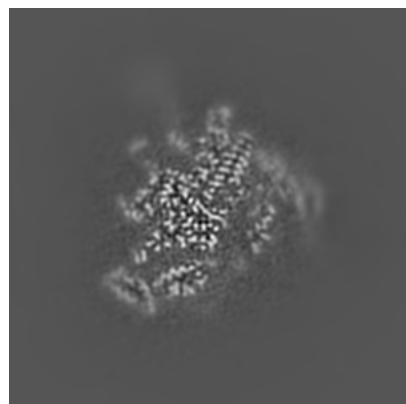


Z Index: 128

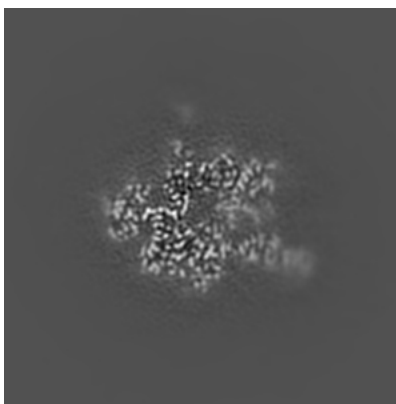
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

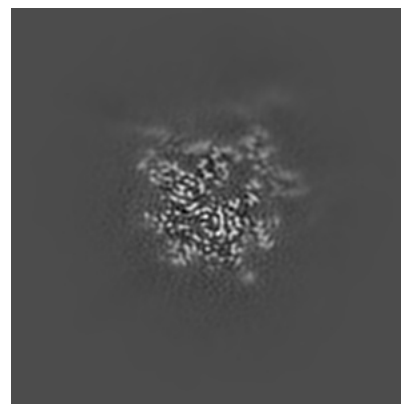
6.3.1 Primary map



X Index: 140



Y Index: 122

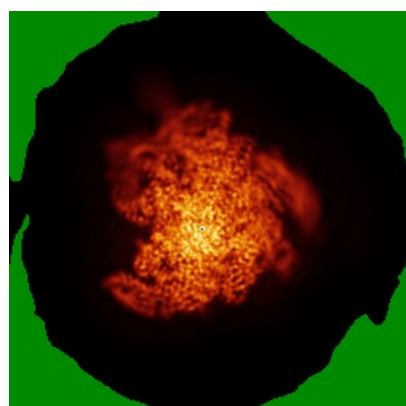


Z Index: 108

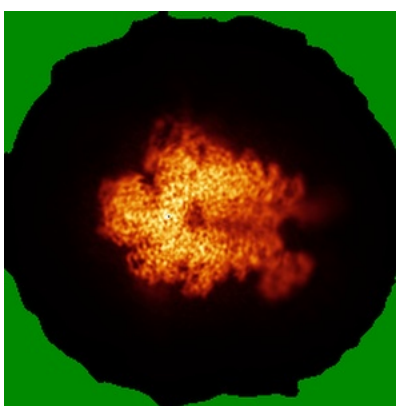
The images above show the largest variance slices of the map in three orthogonal directions.

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

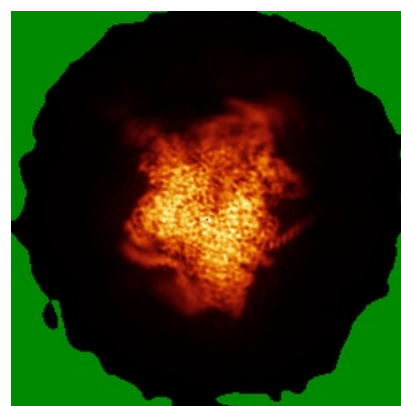
6.4.1 Primary map



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



The images above show the 3D surface view of the map at the recommended contour level 0.018. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

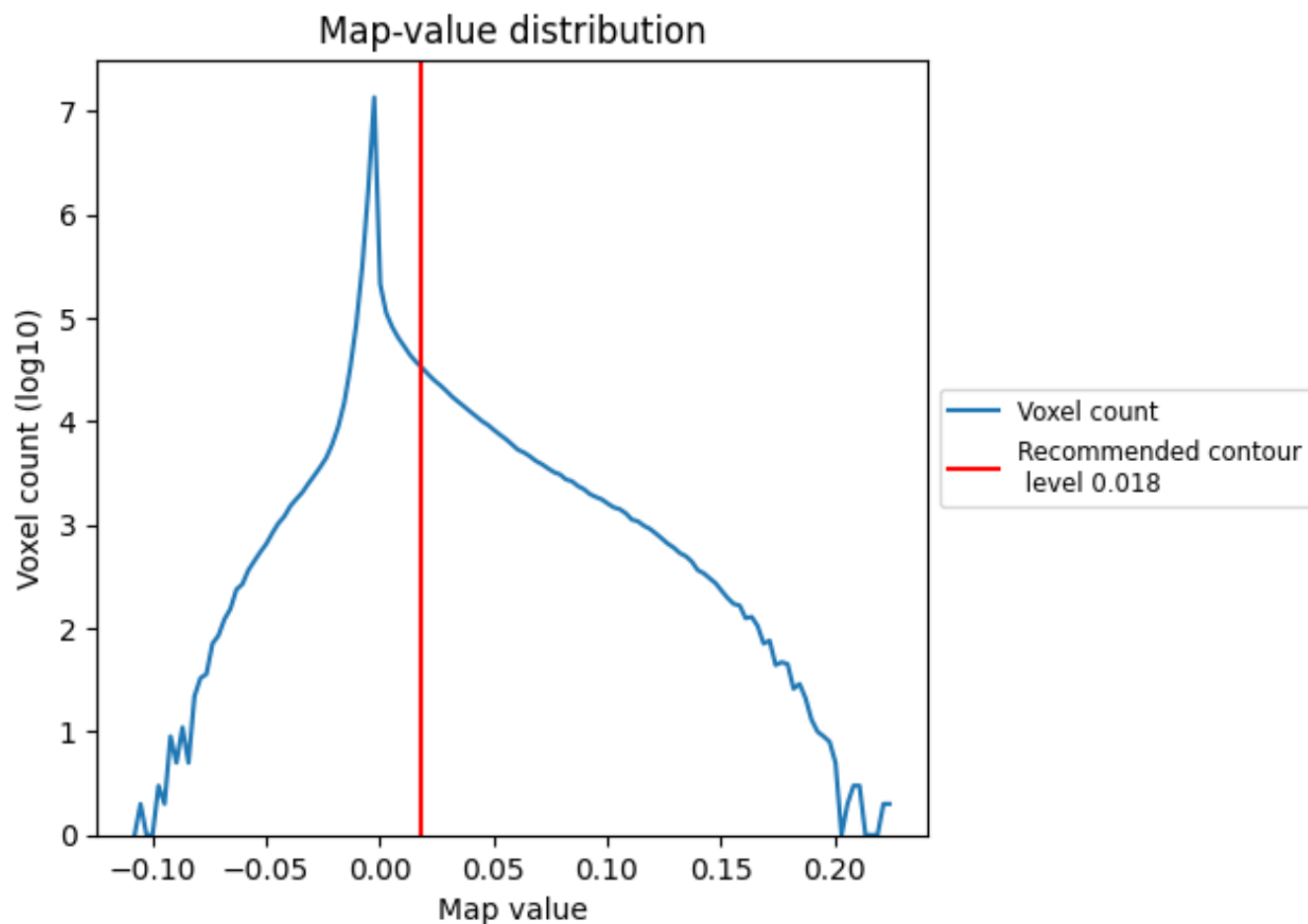
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

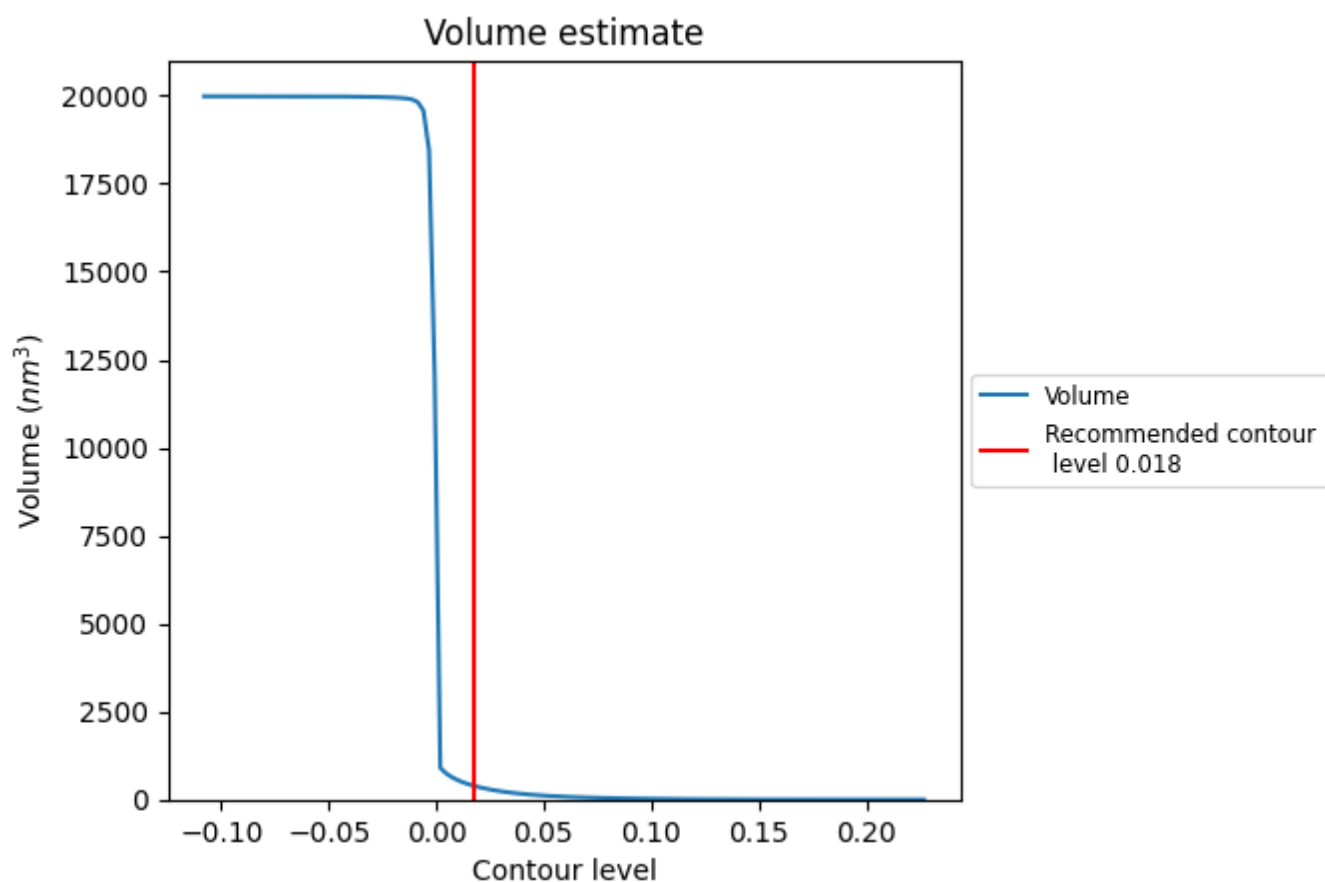
This section contains the results of statistical analysis of the map.

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.

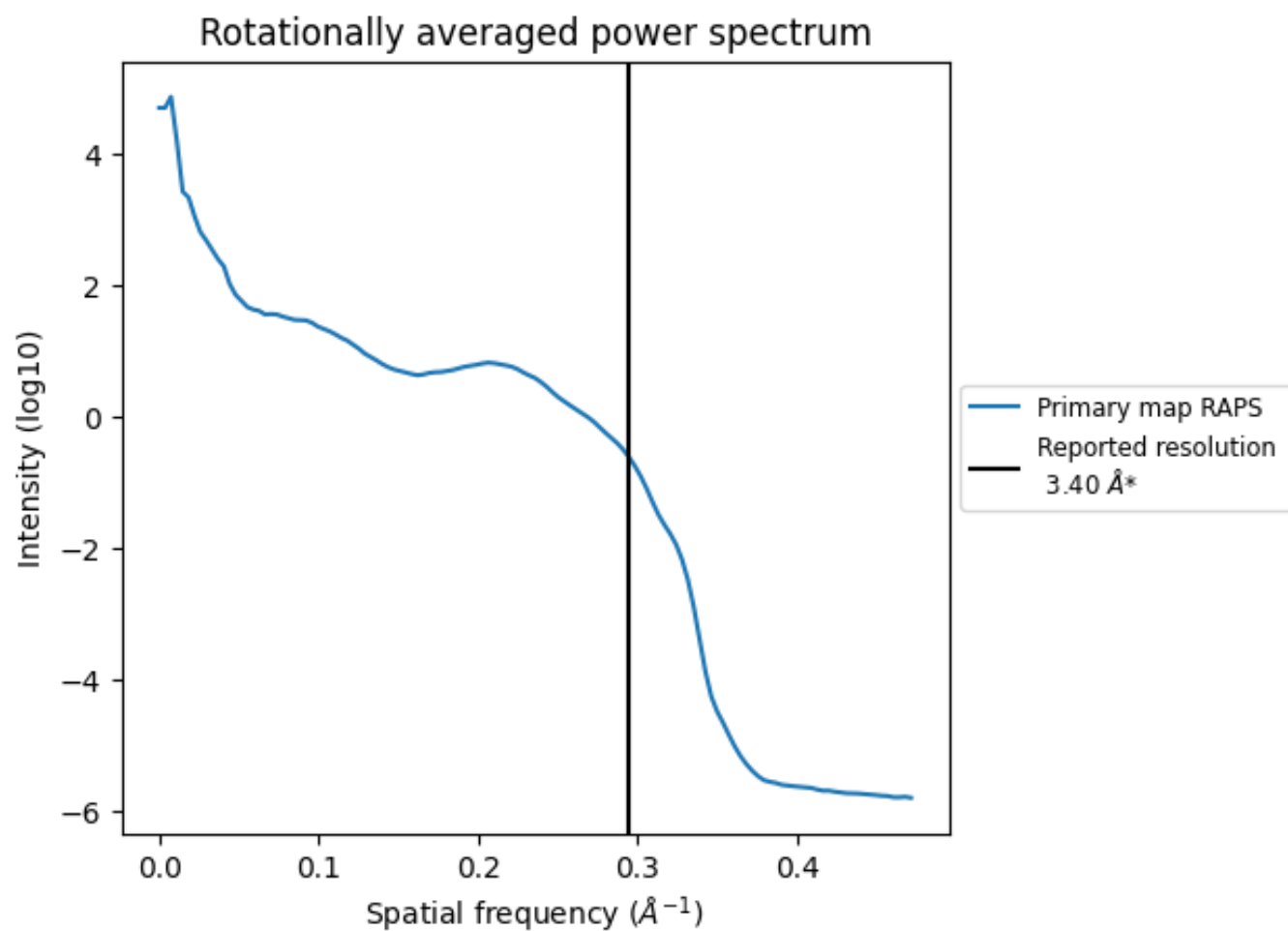
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 383 nm^3 ; this corresponds to an approximate mass of 346 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum ⓘ



*Reported resolution corresponds to spatial frequency of 0.294 Å⁻¹

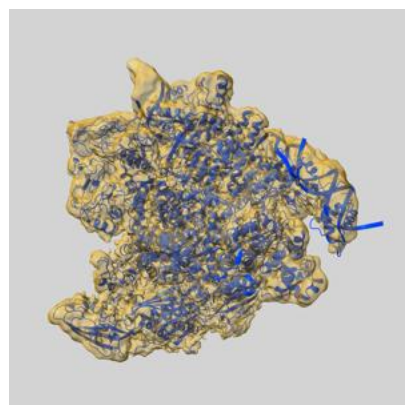
8 Fourier-Shell correlation ⓘ

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

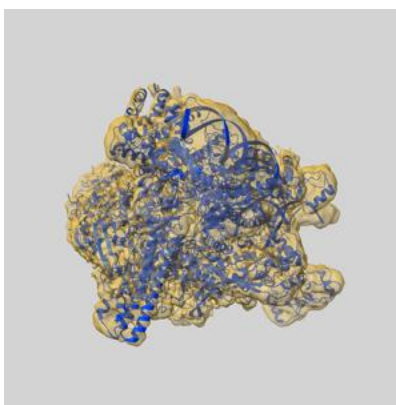
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-0001 and PDB model 6GH5. Per-residue inclusion information can be found in [section 3](#) on [page 6](#).

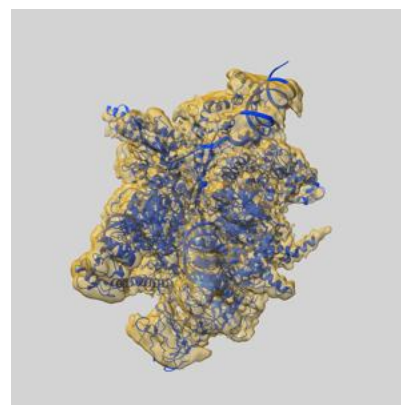
9.1 Map-model overlay [i](#)



X



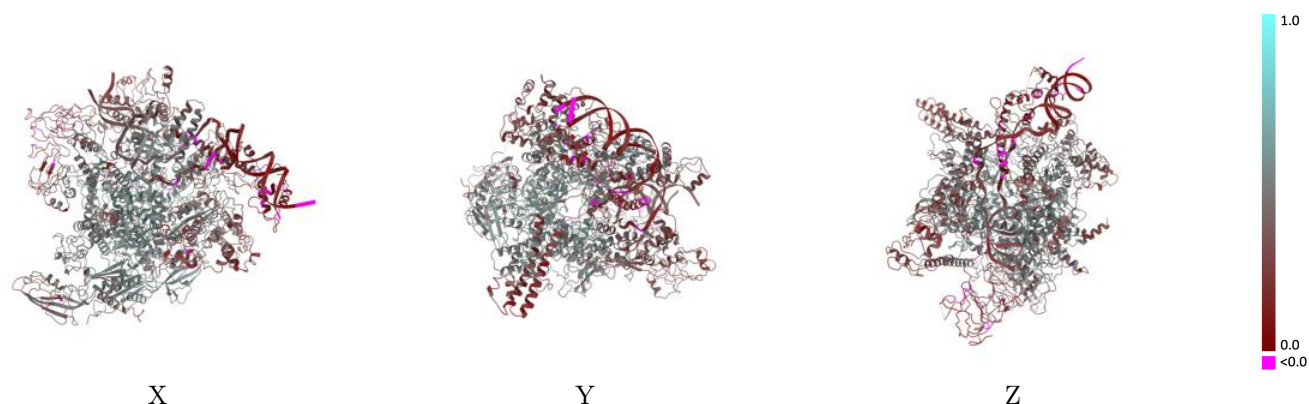
Y



Z

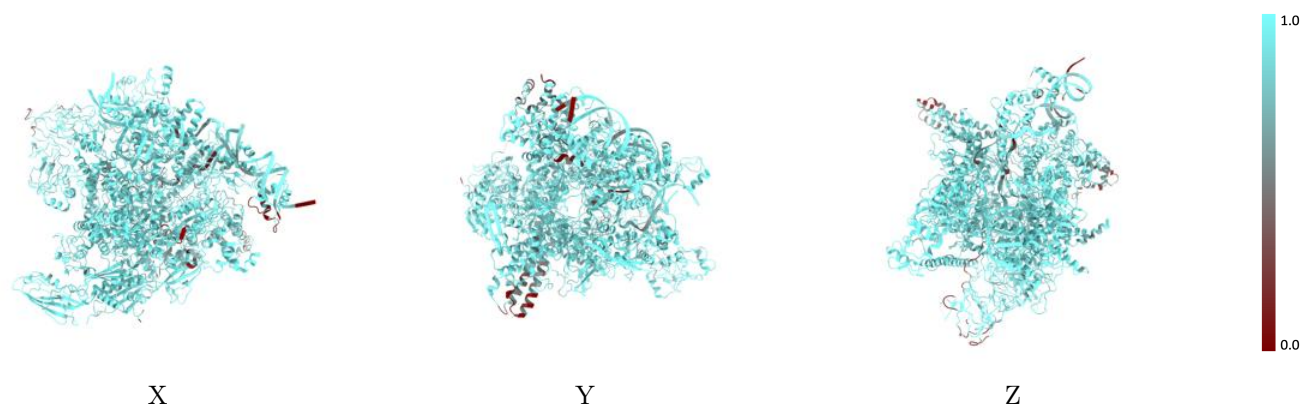
The images above show the 3D surface view of the map at the recommended contour level 0.018 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



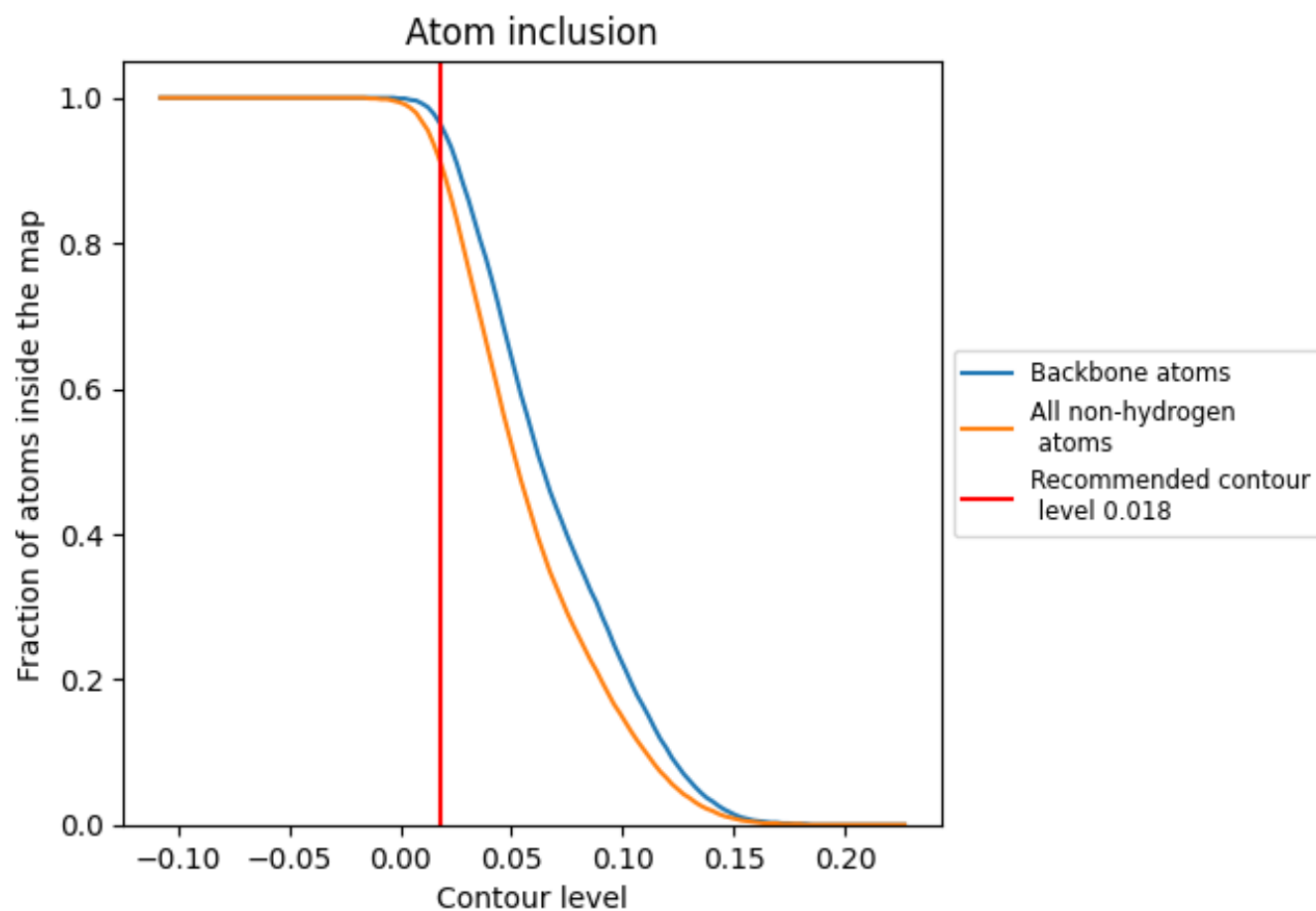
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.018).

9.4 Atom inclusion [i](#)



At the recommended contour level, 96% of all backbone atoms, 91% 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 (0.018) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9120	0.4120
A	0.8810	0.4420
B	0.9370	0.4200
C	0.9280	0.4460
D	0.9210	0.4290
E	0.9400	0.4620
F	0.7720	0.1950
G	0.8450	0.2120
M	0.9010	0.3010

