



wwPDB EM Validation Summary Report ⓘ

Oct 26, 2024 – 05:09 PM EDT

PDB ID : 6DBV
EMDB ID : EMD-7851
Title : Cryo-EM structure of RAG in complex with 12-RSS and 23-RSS substrate DNAs
Authors : Wu, H.; Liao, M.; Ru, H.; Mi, W.
Deposited on : 2018-05-03
Resolution : 4.29 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.dev113
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

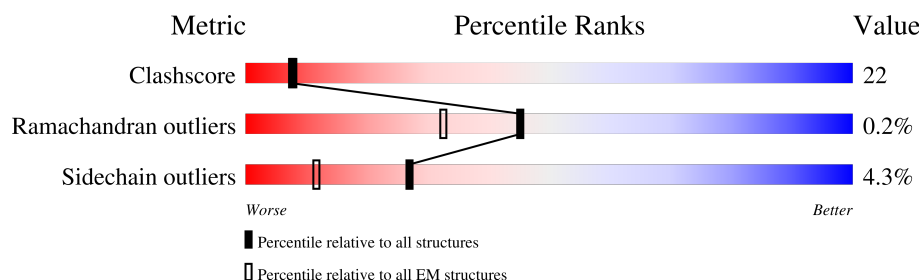
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.29 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	1159	
2	B	533	
2	D	533	
3	C	1159	
4	E	50	
5	F	50	
6	G	61	
7	H	61	

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 19958 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 Recombination activating gene 1 - MBP chimera.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	616	Total	C	N	O	S	0	0
			4974	3113	895	929	37		

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-127	MET	-	initiating methionine	UNP P0AEX9
A	-126	GLY	-	expression tag	UNP P0AEX9
A	-125	SER	-	expression tag	UNP P0AEX9
A	-124	SER	-	expression tag	UNP P0AEX9
A	-123	HIS	-	expression tag	UNP P0AEX9
A	-122	HIS	-	expression tag	UNP P0AEX9
A	-121	HIS	-	expression tag	UNP P0AEX9
A	-120	HIS	-	expression tag	UNP P0AEX9
A	-119	HIS	-	expression tag	UNP P0AEX9
A	-118	HIS	-	expression tag	UNP P0AEX9
A	-117	GLY	-	expression tag	UNP P0AEX9
A	-116	THR	-	expression tag	UNP P0AEX9
A	-115	LYS	-	expression tag	UNP P0AEX9
A	-114	THR	-	expression tag	UNP P0AEX9
A	251	GLY	-	linker	UNP P0AEX9
A	252	THR	-	linker	UNP P0AEX9
A	253	ASP	-	linker	UNP P0AEX9
A	254	TYR	-	linker	UNP P0AEX9
A	255	ASP	-	linker	UNP P0AEX9
A	256	ILE	-	linker	UNP P0AEX9
A	257	PRO	-	linker	UNP P0AEX9
A	258	THR	-	linker	UNP P0AEX9
A	259	THR	-	linker	UNP P0AEX9
A	260	LEU	-	linker	UNP P0AEX9
A	261	GLU	-	linker	UNP P0AEX9
A	262	VAL	-	linker	UNP P0AEX9
A	263	LEU	-	linker	UNP P0AEX9
A	264	PHE	-	linker	UNP P0AEX9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	265	GLN	-	linker	UNP P0AEX9
A	266	GLY	-	linker	UNP P0AEX9
A	267	PRO	-	linker	UNP P0AEX9
A	268	LEU	-	linker	UNP P0AEX9
A	269	GLY	-	linker	UNP P0AEX9
A	270	SER	-	linker	UNP P0AEX9
A	465	ARG	LYS	conflict	UNP O13033

- Molecule 2 is a protein called Recombination activating gene 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	351	Total	C	N	O	S	0	0
			2714	1716	470	509	19		
2	D	351	Total	C	N	O	S	0	0
			2714	1716	470	509	19		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	expression tag	UNP Q1RLW7
B	-1	GLY	-	expression tag	UNP Q1RLW7
B	0	SER	-	expression tag	UNP Q1RLW7
D	-2	GLY	-	expression tag	UNP Q1RLW7
D	-1	GLY	-	expression tag	UNP Q1RLW7
D	0	SER	-	expression tag	UNP Q1RLW7

- Molecule 3 is a protein called Recombination activating gene 1 - MBP chimera.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	621	Total	C	N	O	S	0	0
			4999	3128	899	935	37		

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-127	MET	-	initiating methionine	UNP P0AEX9
C	-126	GLY	-	expression tag	UNP P0AEX9
C	-125	SER	-	expression tag	UNP P0AEX9
C	-124	SER	-	expression tag	UNP P0AEX9
C	-123	HIS	-	expression tag	UNP P0AEX9
C	-122	HIS	-	expression tag	UNP P0AEX9
C	-121	HIS	-	expression tag	UNP P0AEX9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-120	HIS	-	expression tag	UNP P0AEX9
C	-119	HIS	-	expression tag	UNP P0AEX9
C	-118	HIS	-	expression tag	UNP P0AEX9
C	-117	GLY	-	expression tag	UNP P0AEX9
C	-116	THR	-	expression tag	UNP P0AEX9
C	-115	LYS	-	expression tag	UNP P0AEX9
C	-114	THR	-	expression tag	UNP P0AEX9
C	251	GLY	-	linker	UNP P0AEX9
C	252	THR	-	linker	UNP P0AEX9
C	253	ASP	-	linker	UNP P0AEX9
C	254	TYR	-	linker	UNP P0AEX9
C	255	ASP	-	linker	UNP P0AEX9
C	256	ILE	-	linker	UNP P0AEX9
C	257	PRO	-	linker	UNP P0AEX9
C	258	THR	-	linker	UNP P0AEX9
C	259	THR	-	linker	UNP P0AEX9
C	260	LEU	-	linker	UNP P0AEX9
C	261	GLU	-	linker	UNP P0AEX9
C	262	VAL	-	linker	UNP P0AEX9
C	263	LEU	-	linker	UNP P0AEX9
C	264	PHE	-	linker	UNP P0AEX9
C	265	GLN	-	linker	UNP P0AEX9
C	266	GLY	-	linker	UNP P0AEX9
C	267	PRO	-	linker	UNP P0AEX9
C	268	LEU	-	linker	UNP P0AEX9
C	269	GLY	-	linker	UNP P0AEX9
C	270	SER	-	linker	UNP P0AEX9

- Molecule 4 is a DNA chain called Forward strand of 12-RSS substrate DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	50	Total	C	N	O	P	0	0
			1023	486	192	295	50		

- Molecule 5 is a DNA chain called Reverse strand of 12-RSS substrate DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	50	Total	C	N	O	P	0	0
			1027	489	183	305	50		

- Molecule 6 is a DNA chain called Forward strand of 23-RSS substrate DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	G	61	Total	C	N	O	P	0	0
			1245	593	223	368	61		

- Molecule 7 is a DNA chain called Reverse strand of 23-RSS substrate DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	H	61	Total	C	N	O	P	0	0
			1256	596	235	364	61		

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

Mol	Chain	Residues	Atoms		AltConf
8	A	1	Total	Zn	0
			1	1	
8	C	1	Total	Zn	0
			1	1	

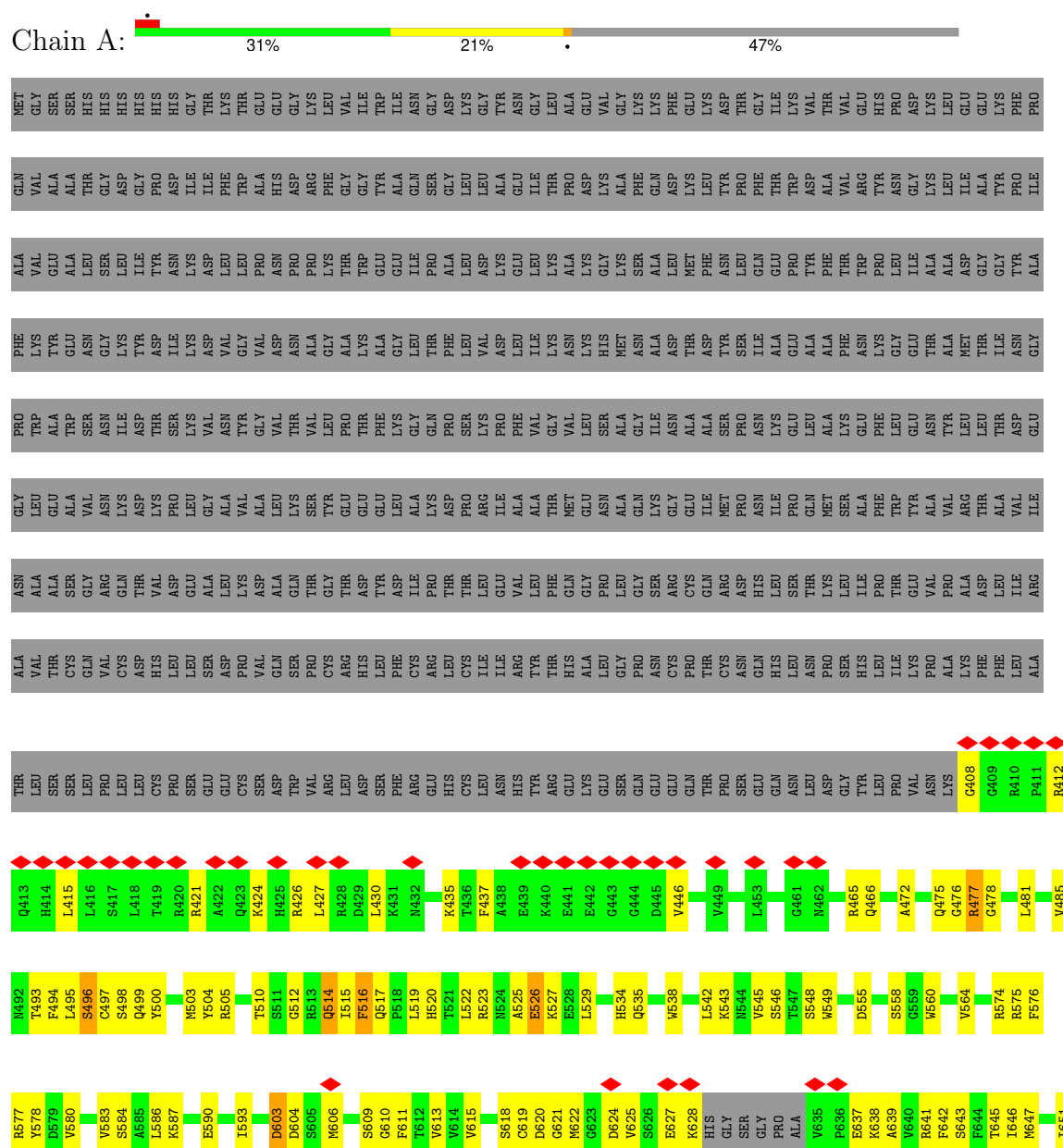
- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

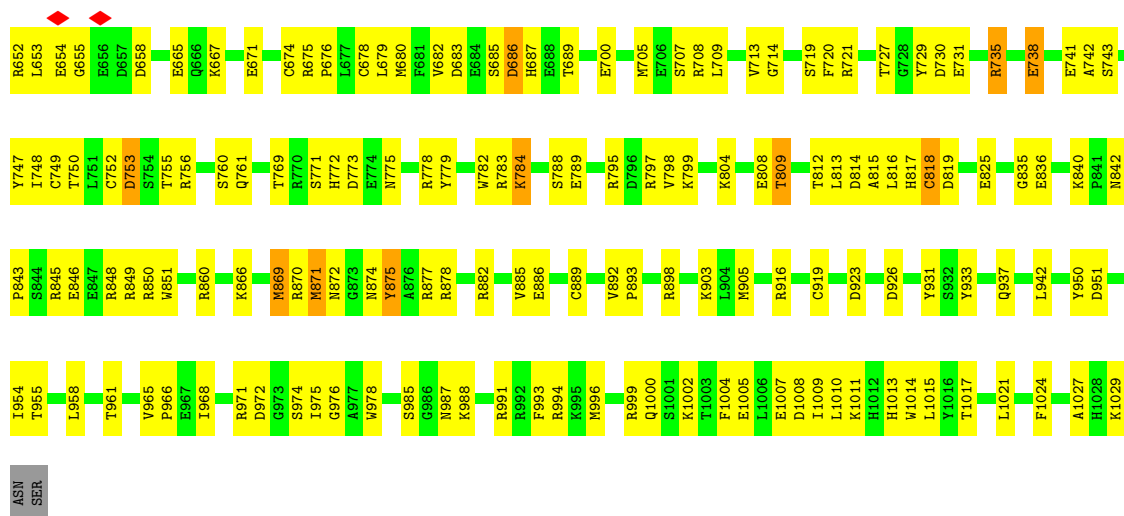
Mol	Chain	Residues	Atoms		AltConf
9	A	2	Total	Ca	0
			2	2	
9	C	2	Total	Ca	0
			2	2	

3 Residue-property plots

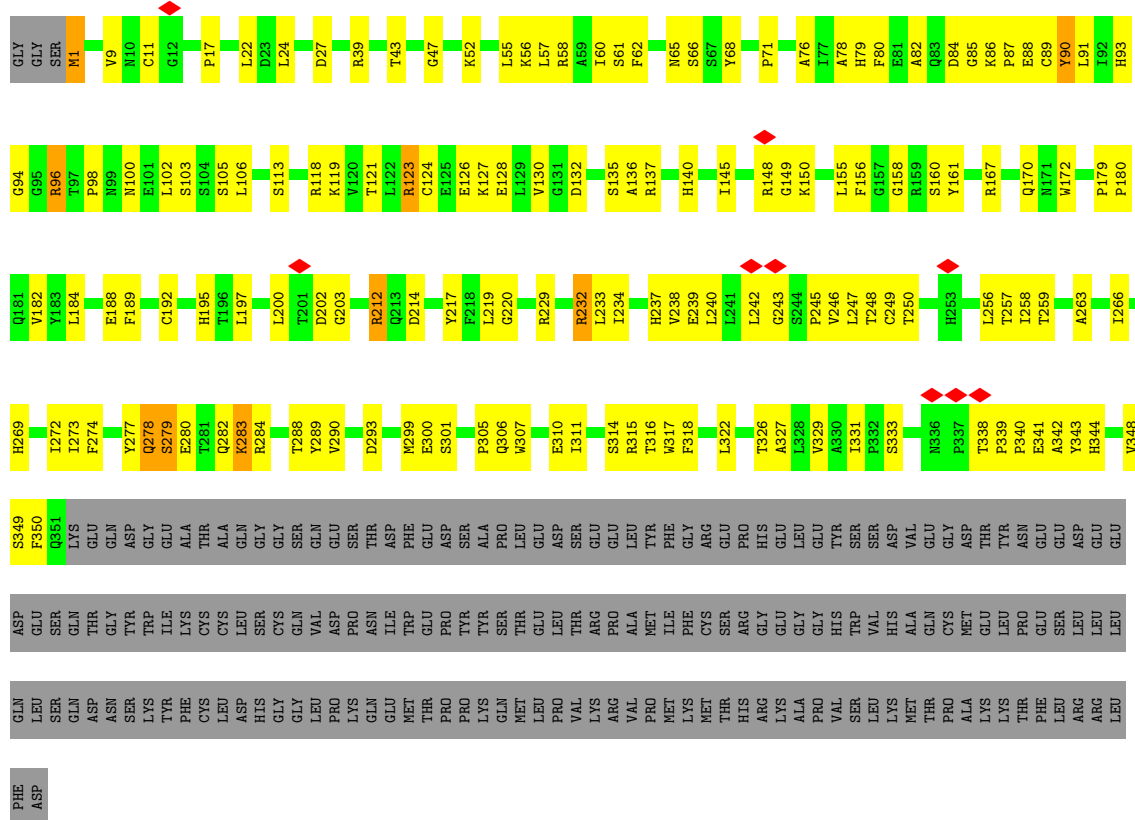
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Recombination activating gene 1 - MBP chimera



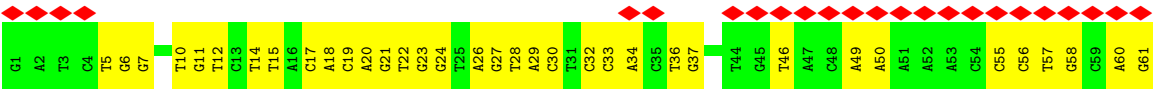


• Molecule 2: Recombination activating gene 2

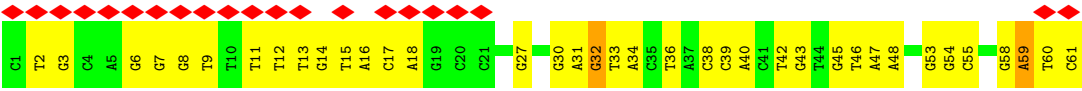


• Molecule 2: Recombination activating gene 2





• Molecule 7: Reverse strand of 23-RSS substrate DNA



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	45159	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	47	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	0.166	Depositor
Minimum map value	-0.078	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.009	Depositor
Recommended contour level	0.03	Depositor
Map size (Å)	237.69601, 237.69601, 237.69601	wwPDB
Map dimensions	192, 192, 192	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.238, 1.238, 1.238	Depositor

5 Model quality

5.1 Standard geometry

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

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	0.34	0/5069	0.51	2/6815 (0.0%)
2	B	0.31	0/2784	0.54	1/3784 (0.0%)
2	D	0.30	0/2784	0.54	1/3784 (0.0%)
3	C	0.35	0/5097	0.52	1/6856 (0.0%)
4	E	0.73	0/1148	0.93	1/1768 (0.1%)
5	F	0.74	0/1150	1.05	2/1774 (0.1%)
6	G	0.70	0/1394	0.94	0/2148
7	H	0.70	0/1410	0.98	2/2175 (0.1%)
All	All	0.46	0/20836	0.68	10/29104 (0.0%)

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	1

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	34	DG	O4'-C1'-N9	11.35	115.95	108.00
2	B	263	ALA	C-N-CA	-6.96	104.30	121.70
3	C	674	CYS	C-N-CA	-6.92	104.40	121.70
1	A	809	THR	C-N-CA	-6.47	105.51	121.70
5	F	34	DG	C1'-O4'-C4'	-5.38	104.72	110.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	512	GLY	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	4974	0	4937	214	0
2	B	2714	0	2665	139	0
2	D	2714	0	2665	135	0
3	C	4999	0	4954	257	0
4	E	1023	0	561	49	0
5	F	1027	0	566	41	0
6	G	1245	0	688	41	0
7	H	1256	0	686	47	0
8	A	1	0	0	0	0
8	C	1	0	0	0	0
9	A	2	0	0	0	0
9	C	2	0	0	0	0
All	All	19958	0	17722	836	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 22.

The worst 5 of 836 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:999:ARG:HG3	3:C:1008:ASP:OD2	1.17	1.27
3:C:505:ARG:NH1	5:F:25:DT:OP2	1.71	1.22
1:A:523:ARG:NH1	6:G:24:DG:OP2	1.81	1.13
3:C:760:SER:HA	3:C:954:ILE:HD11	1.20	1.07
1:A:590:GLU:HG2	1:A:713:VAL:HG23	1.34	1.04

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	612/1159 (53%)	558 (91%)	54 (9%)	0	100	100
2	B	349/533 (66%)	308 (88%)	41 (12%)	0	100	100
2	D	349/533 (66%)	306 (88%)	43 (12%)	0	100	100
3	C	619/1159 (53%)	546 (88%)	70 (11%)	3 (0%)	25	63
All	All	1929/3384 (57%)	1718 (89%)	208 (11%)	3 (0%)	45	78

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	512	GLY
3	C	818	CYS
3	C	657	ASP

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	545/1000 (54%)	519 (95%)	26 (5%)	21	44
2	B	303/465 (65%)	291 (96%)	12 (4%)	27	49
2	D	303/465 (65%)	292 (96%)	11 (4%)	30	52
3	C	547/1000 (55%)	523 (96%)	24 (4%)	24	47
All	All	1698/2930 (58%)	1625 (96%)	73 (4%)	27	48

5 of 73 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	859	LEU
2	D	292	LEU
3	C	942	LEU
2	D	79	HIS
1	A	996	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	775	ASN
3	C	874	ASN
3	C	983	ASN
3	C	936	GLN
2	B	269	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 are 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.

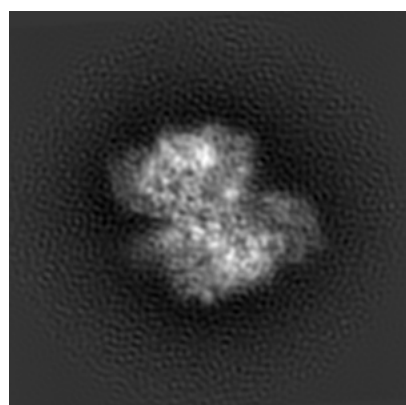
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-7851. These allow visual inspection of the internal detail of the map and identification of artifacts.

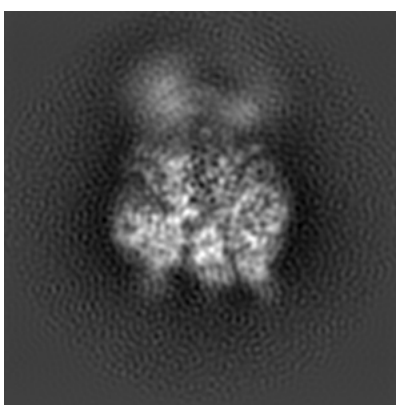
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

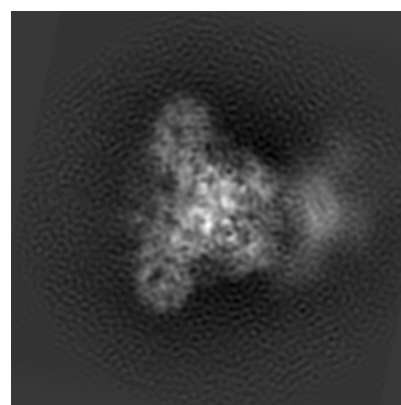
6.1.1 Primary map



X



Y

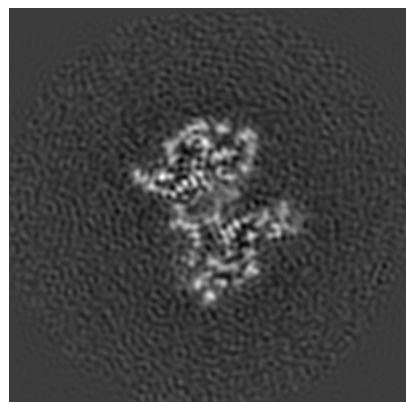


Z

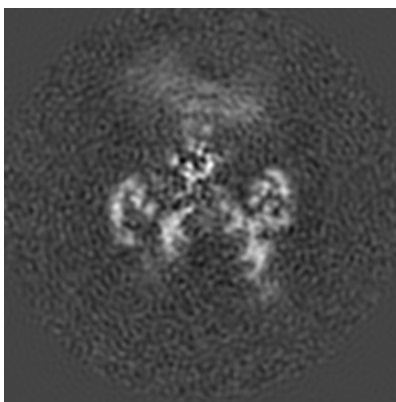
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

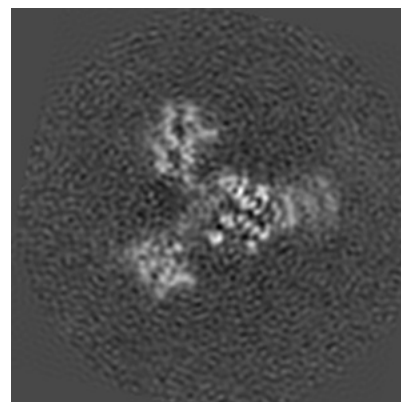
6.2.1 Primary map



X Index: 96



Y Index: 96

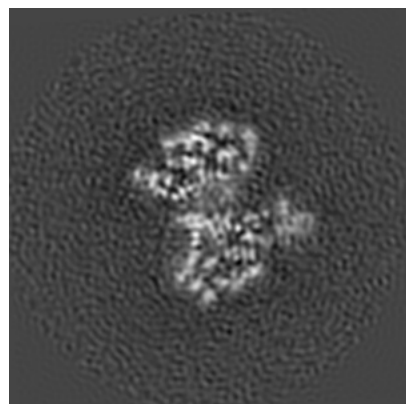


Z Index: 96

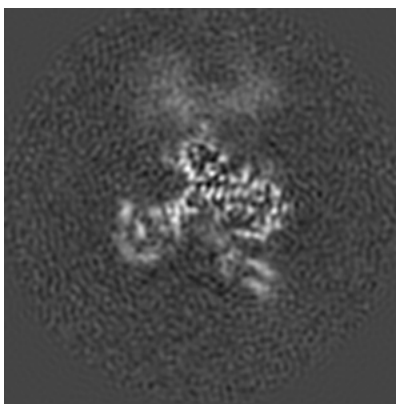
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

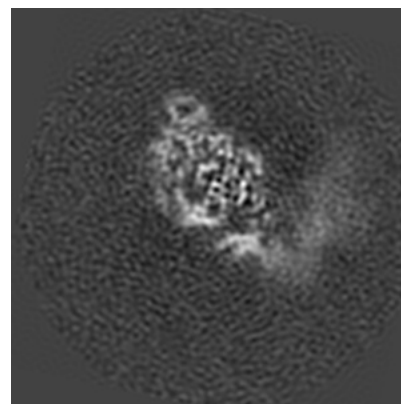
6.3.1 Primary map



X Index: 94



Y Index: 89

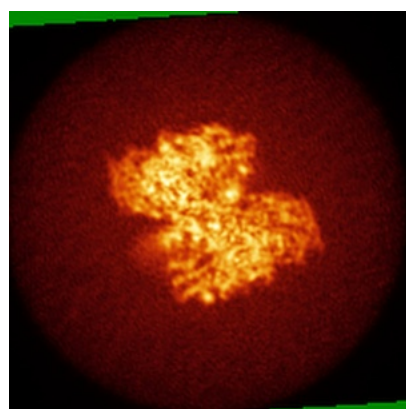


Z Index: 83

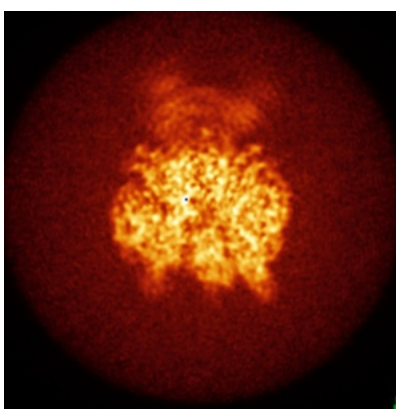
The images above show the largest variance slices of the map in three orthogonal directions.

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

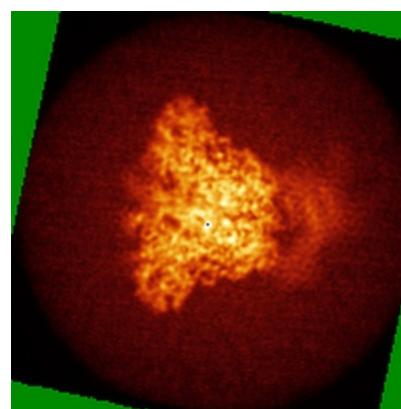
6.4.1 Primary map



X



Y

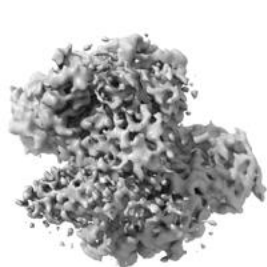


Z

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](#)

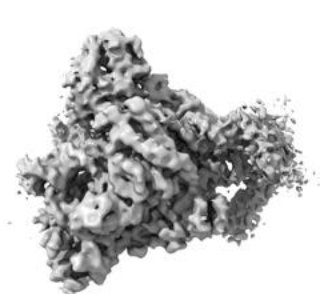
6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.03. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

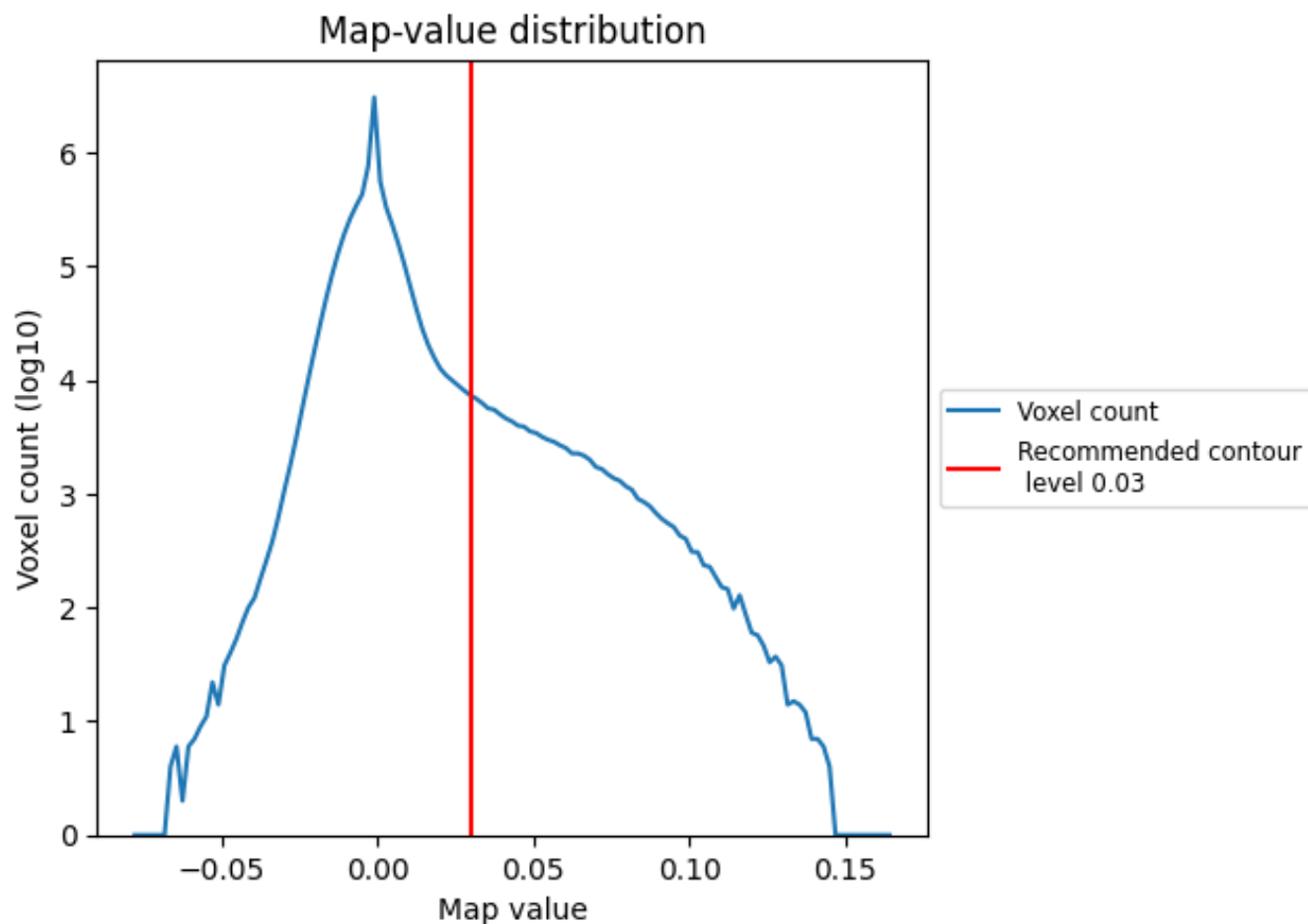
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

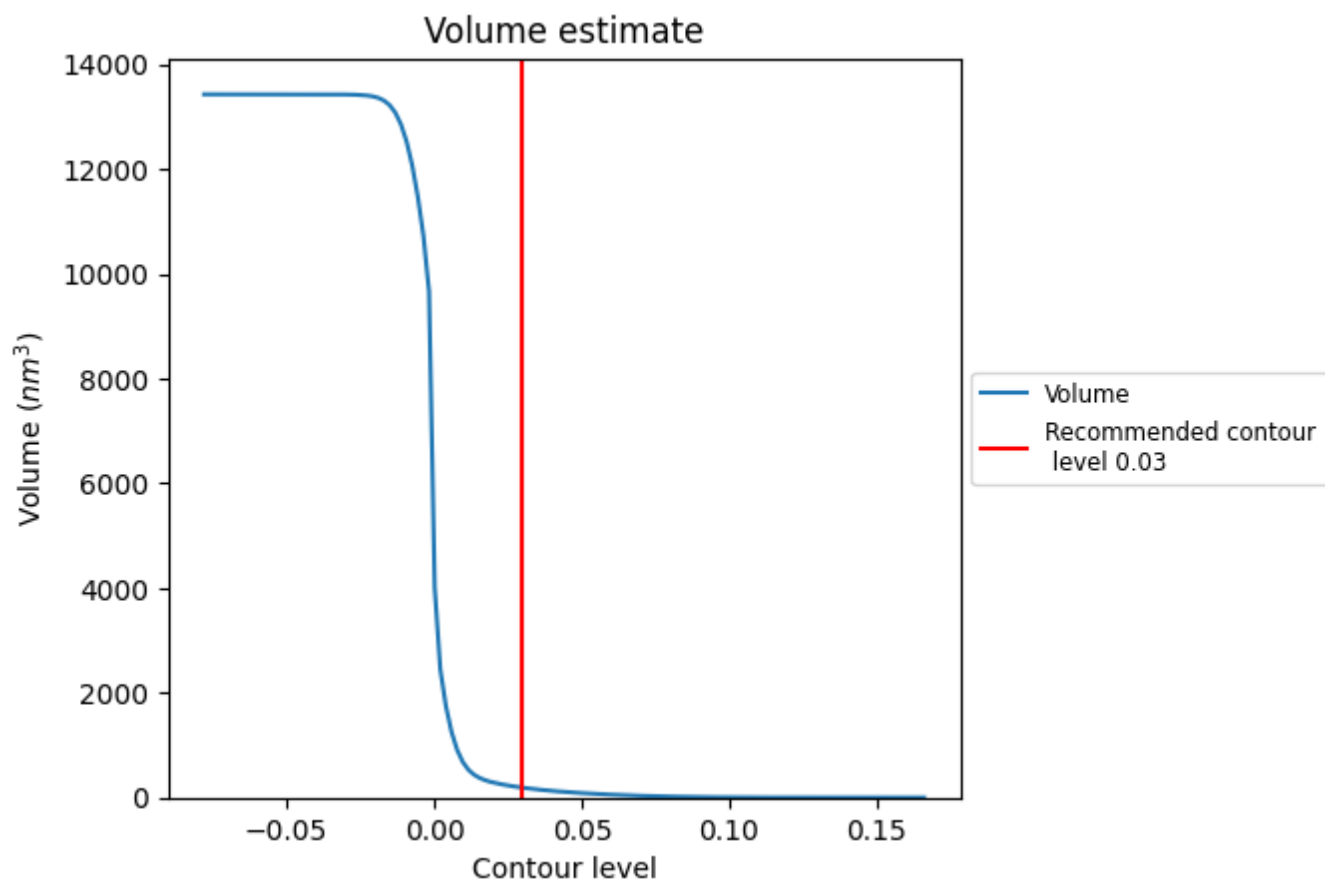
This section contains the results of statistical analysis of the map.

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.

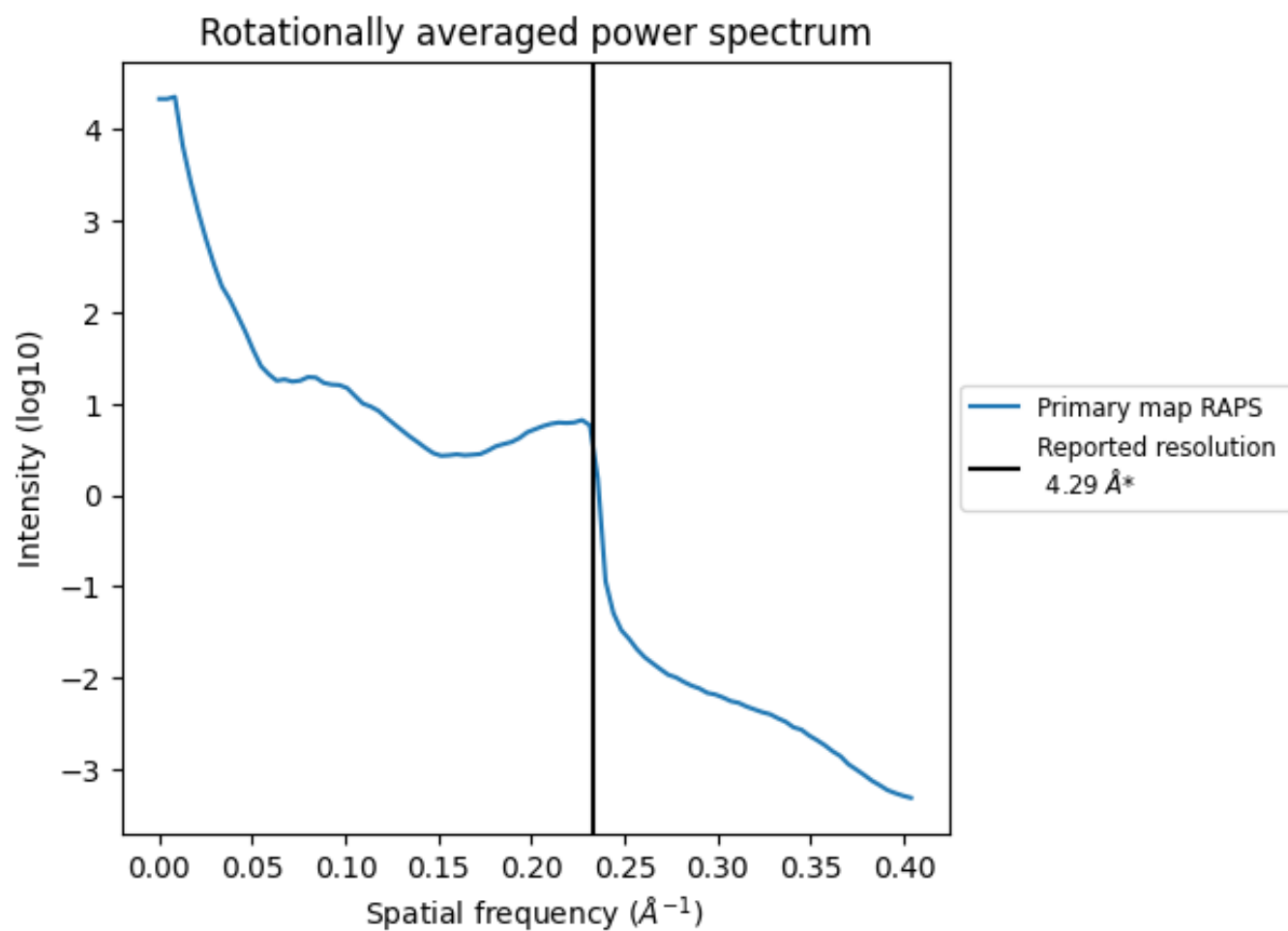
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 188 nm³; this corresponds to an approximate mass of 170 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 0.233 Å⁻¹

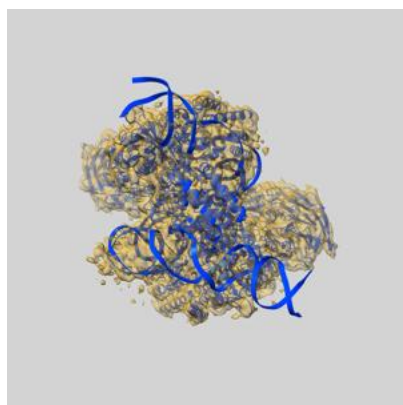
8 Fourier-Shell correlation

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

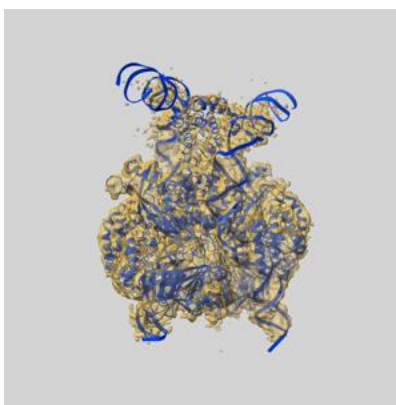
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-7851 and PDB model 6DBV. Per-residue inclusion information can be found in [section 3](#) on [page 7](#).

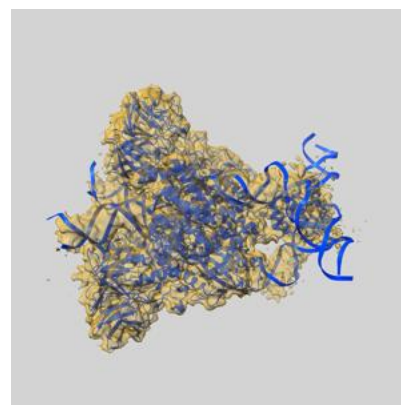
9.1 Map-model overlay [i](#)



X



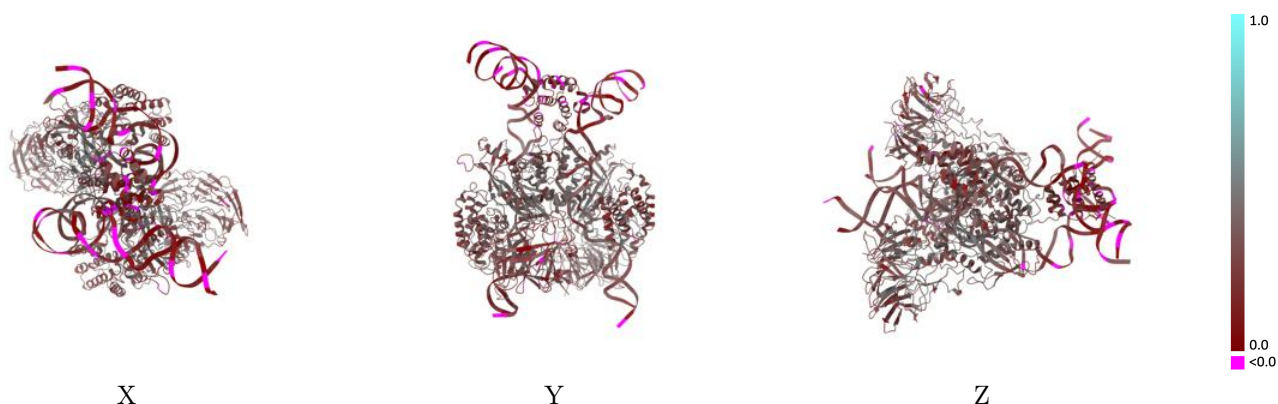
Y



Z

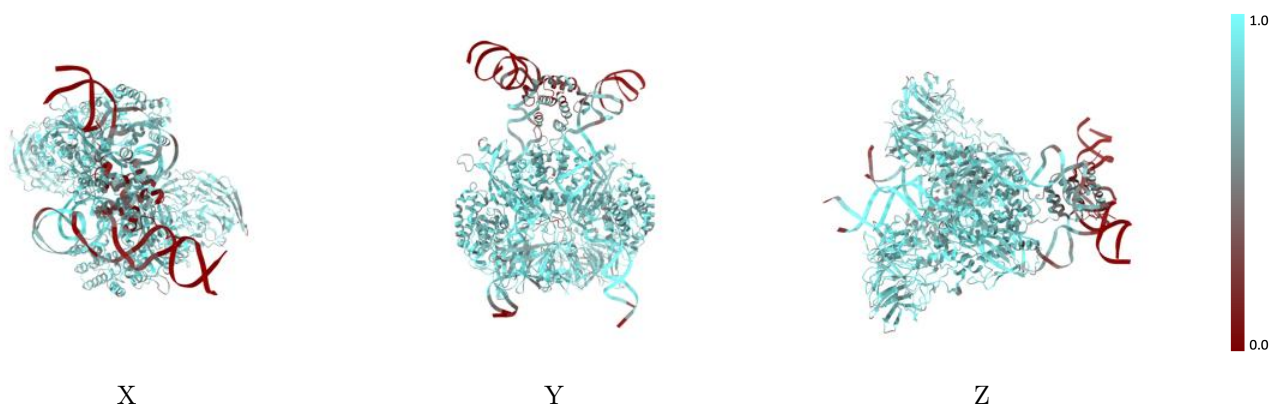
The images above show the 3D surface view of the map at the recommended contour level 0.03 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



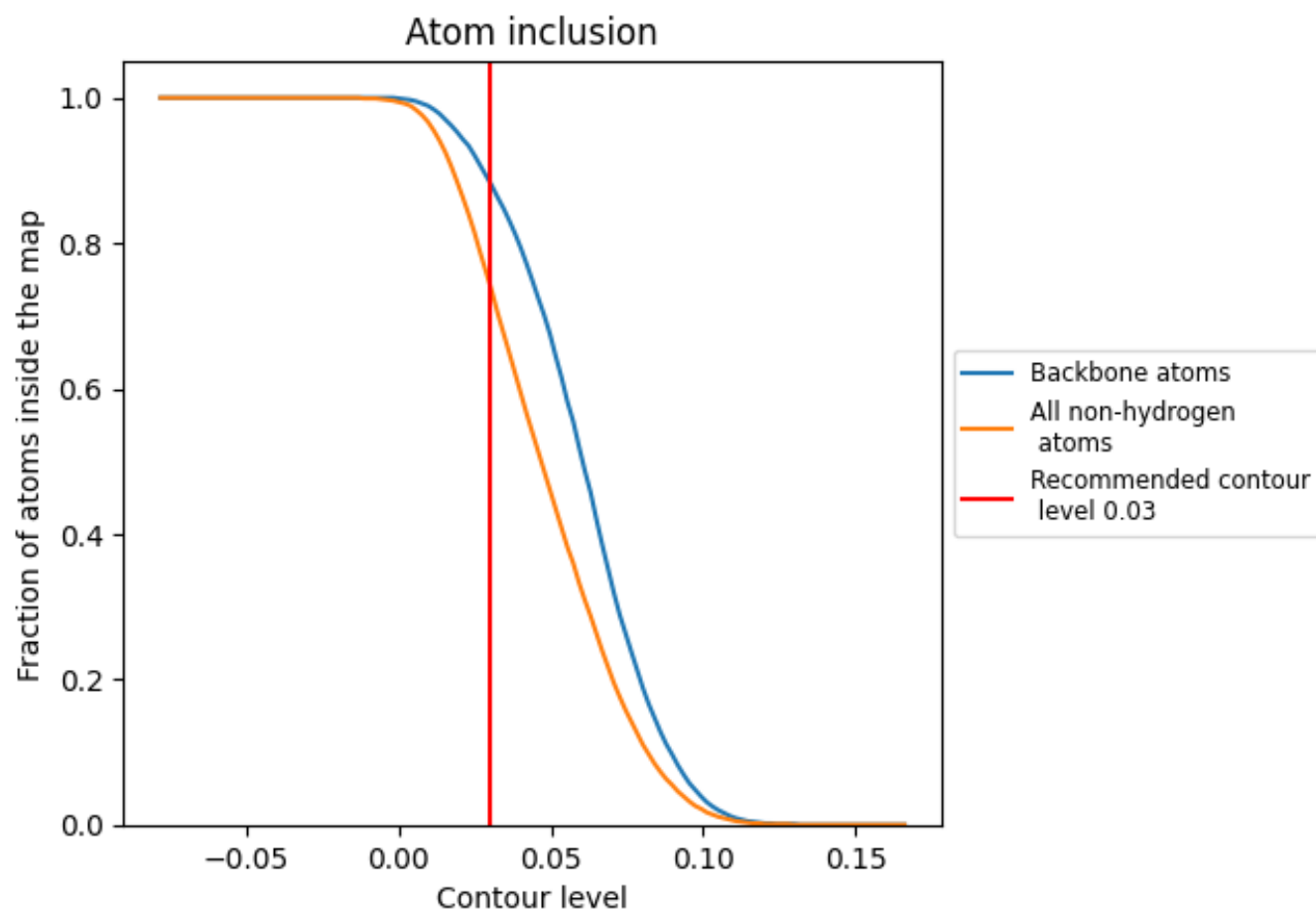
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.03).

9.4 Atom inclusion [i](#)



At the recommended contour level, 88% of all backbone atoms, 74% 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 (0.03) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div><div></div></div> 0.7440	<div><div></div></div> 0.2960
A	<div><div></div></div> 0.7730	<div><div></div></div> 0.3290
B	<div><div></div></div> 0.8170	<div><div></div></div> 0.3400
C	<div><div></div></div> 0.7620	<div><div></div></div> 0.3160
D	<div><div></div></div> 0.7860	<div><div></div></div> 0.3110
E	<div><div></div></div> 0.6720	<div><div></div></div> 0.2060
F	<div><div></div></div> 0.6810	<div><div></div></div> 0.2370
G	<div><div></div></div> 0.5820	<div><div></div></div> 0.1880
H	<div><div></div></div> 0.5880	<div><div></div></div> 0.1990

1.0

0.0

<0.0