



# wwPDB EM Validation Summary Report ⓘ

Apr 28, 2024 – 05:16 am BST

PDB ID : 4BIL  
EMDB ID : EMD-2356  
Title : Threading model of the T7 large terminase within the gp8gp19 complex  
Authors : Dauden, M.I.; Martin-Benito, J.; Sanchez-Ferrero, J.C.; Pulido-Cid, M.;  
Valpuesta, J.M.; Carrascosa, J.L.  
Deposited on : 2013-04-10  
Resolution : 29.00 Å(reported)  
Based on initial model : 3CPE

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.dev92  
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) : 2.36.2

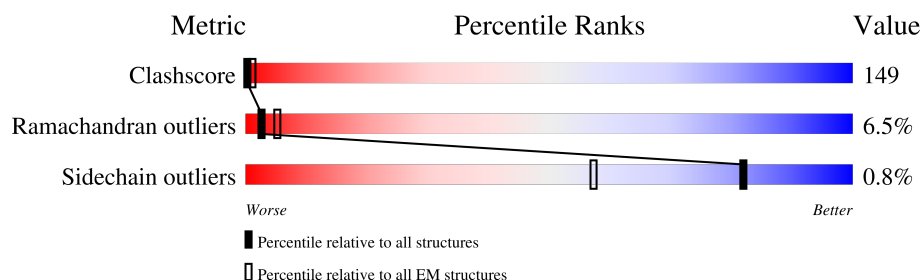
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 29.00 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	476	<div> <div>6%</div> <div>22%</div> <div>61%</div> <div>12%</div> <div>.</div> </div>
1	B	476	<div> <div>6%</div> <div>19%</div> <div>64%</div> <div>13%</div> <div>.</div> </div>
1	C	476	<div> <div>6%</div> <div>19%</div> <div>64%</div> <div>13%</div> <div>.</div> </div>
1	D	476	<div> <div>6%</div> <div>19%</div> <div>64%</div> <div>12%</div> <div>5%</div> </div>
1	E	476	<div> <div>6%</div> <div>19%</div> <div>64%</div> <div>13%</div> <div>.</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 18925 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.

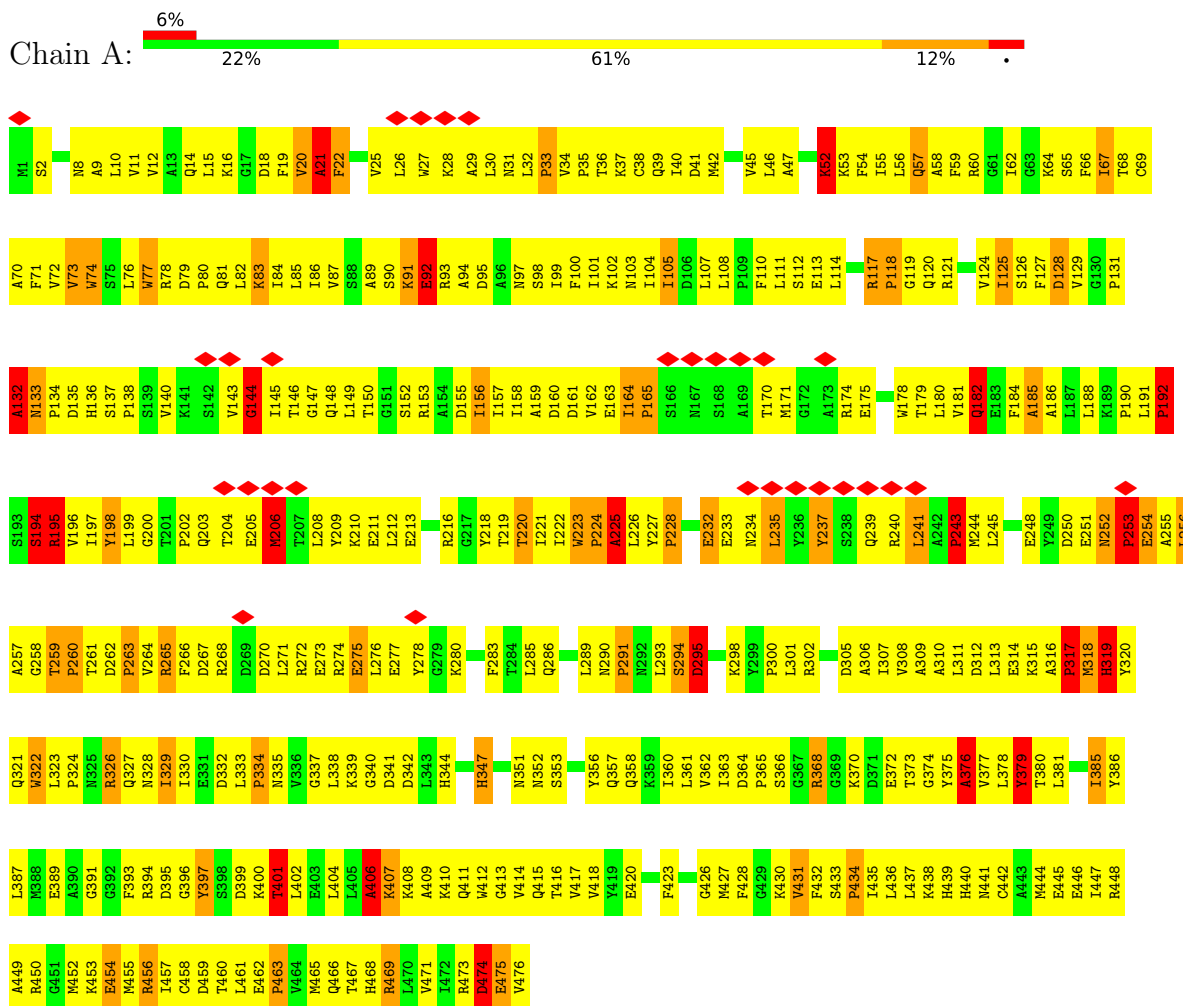
- Molecule 1 is a protein called DNA MATURASE B.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	476	Total	C	N	O	S	0	0
			3785	2412	650	705	18		
1	B	476	Total	C	N	O	S	0	0
			3785	2412	650	705	18		
1	C	476	Total	C	N	O	S	0	0
			3785	2412	650	705	18		
1	D	476	Total	C	N	O	S	0	0
			3785	2412	650	705	18		
1	E	476	Total	C	N	O	S	0	0
			3785	2412	650	705	18		

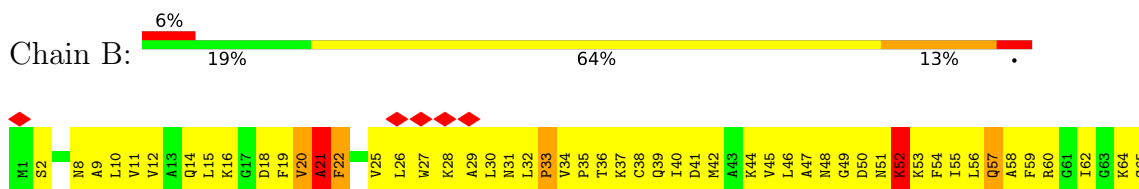
### 3 Residue-property plots

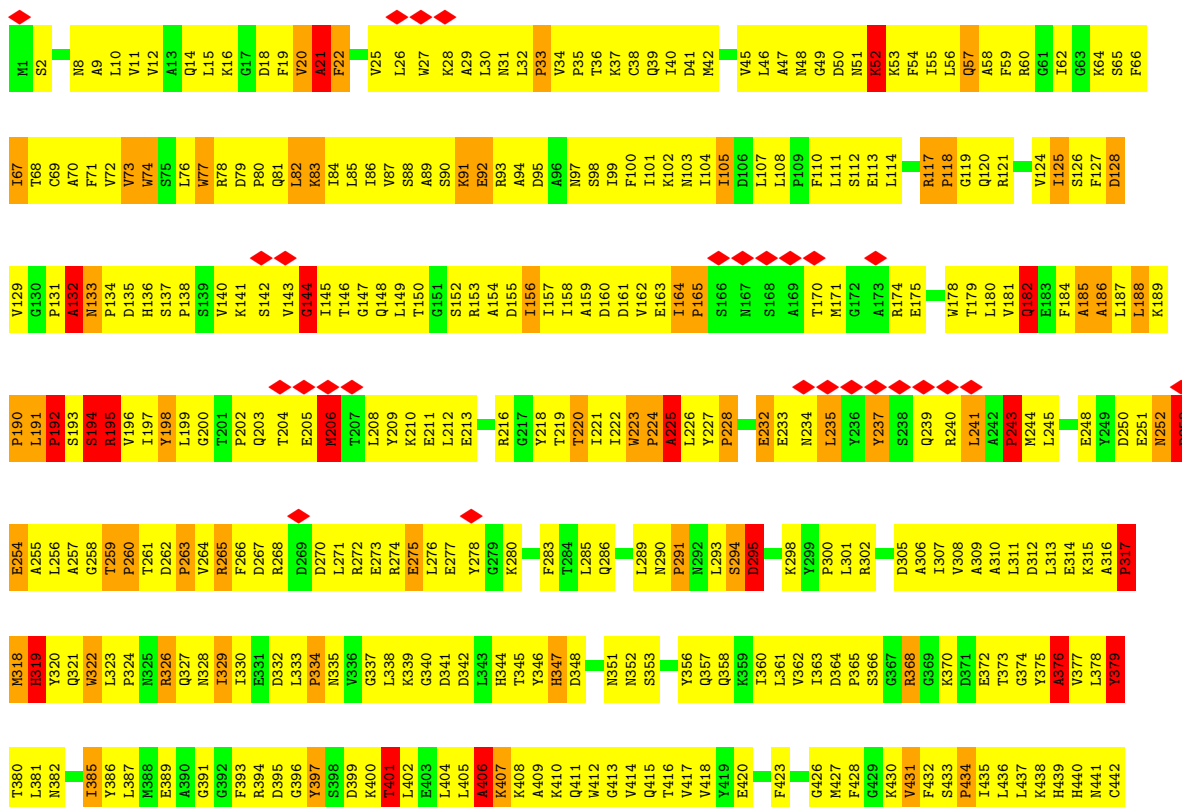
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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.

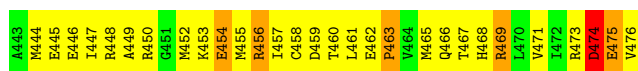
#### • Molecule 1: DNA MATURASE B



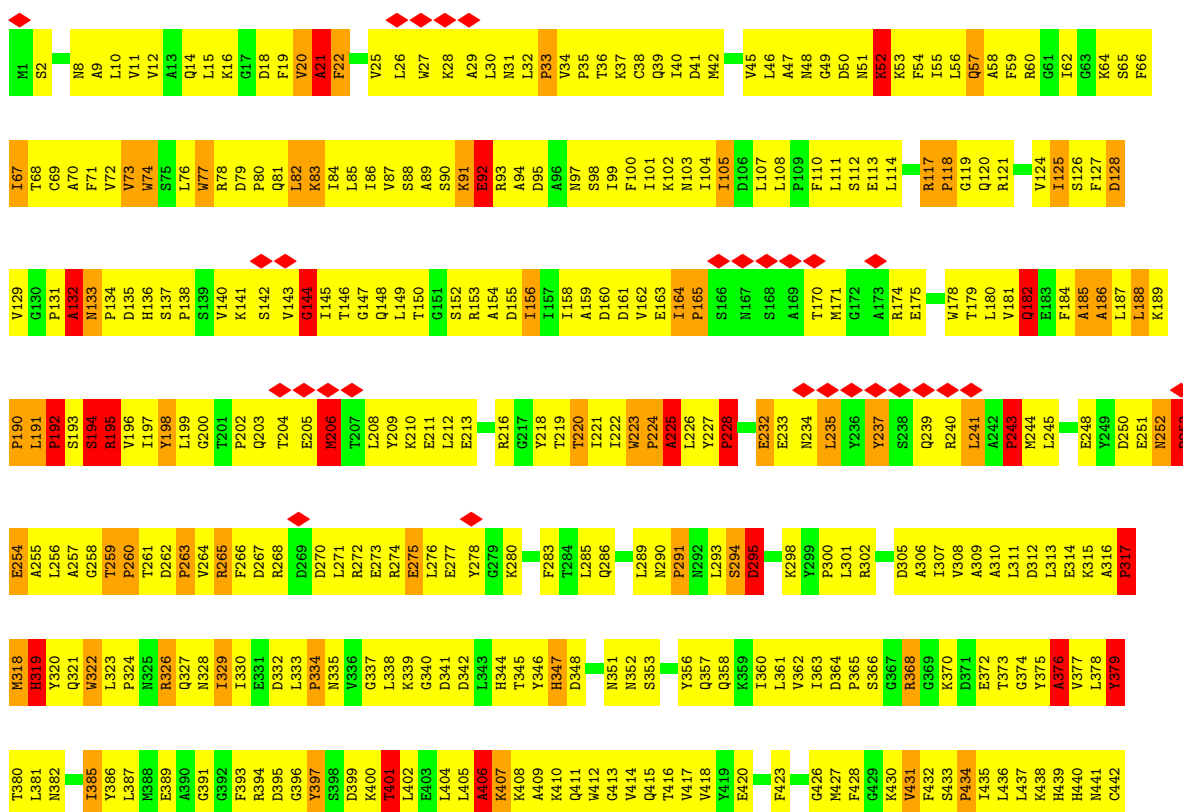
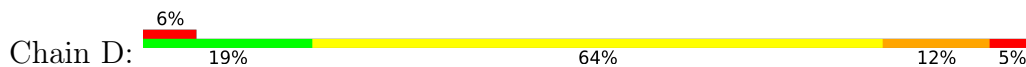
#### • Molecule 1: DNA MATURASE B



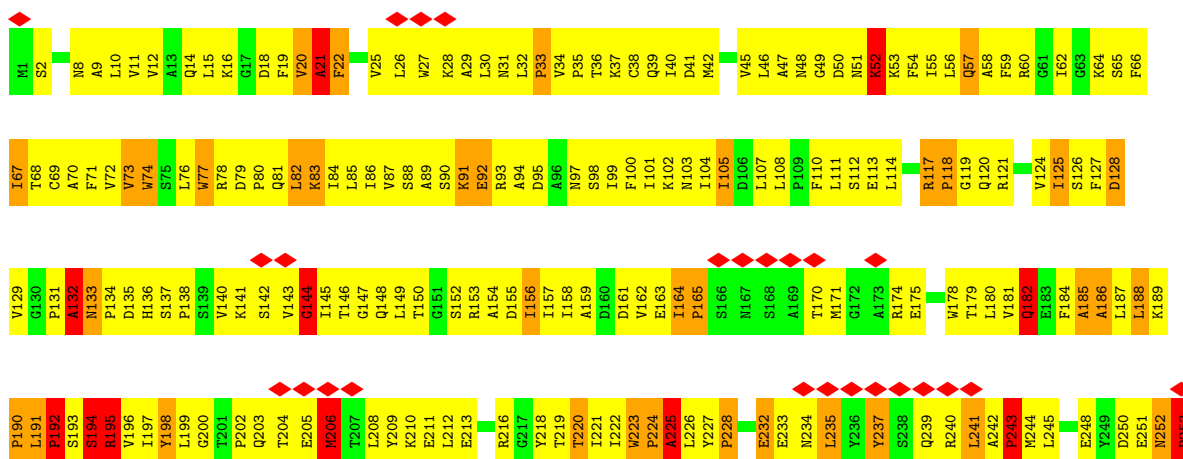
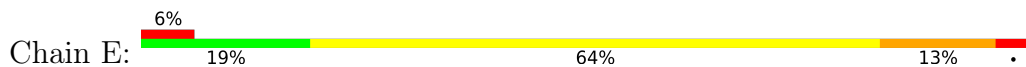




• Molecule 1: DNA MATURASE B



• Molecule 1: DNA MATURASE B



T380	A443	M444	E445	E446	I447	R448	A449	R450	G451	M452	K453	E454	M455	R456	I457	C458	D459	T460	L461	E462	P463	V464	M465	Q466	T467	H468	R469	L470	V471	I472	R473	D474	E475	V476																							
L381	N382	I385	Y386	L387	R388	E389	G390	G391	G392	F393	R394	D395	G396	Y397	S398	D399	K400	T401	L402	E403	L404	L405	A406	K407	K408	A409	K410	Q411	V412	G413	V414	Q415	T416	V417	V418	Y419	E420	F423	G426	M427	F428	G429	K430	V431	F432	S433	P434	I435	L436	L437	K438	H439	H440	N441	C442		
M318	R319	Y320	Q321	W322	L323	P324	N325	R326	Q327	N328	I329	I330	E331	D332	L333	P334	N335	V336	G337	L338	K339	G340	D341	D342	L343	H344	T345	Y346	H347	D348	N351	N352	S353	Y356	Q357	Q358	K359	I360	L361	V362	I363	D364	P365	S366	G367	R368	G369	K370	D371	E372	T373	G374	Y375	A376	V377	L378	Y379
E254	A255	L256	A257	G258	T259	P260	T261	D262	P263	V264	R265	F266	D267	R268	D269	D270	L271	R272	E273	R274	E275	L276	E277	Y278	G279	K280	F283	T284	L285	Q286	L289	N290	P291	N292	L293	S294	D295	K298	Y299	P300	L301	R302	D305	A306	I307	V308	A309	A310	L311	D312	L313	E314	K315	A316	P317		

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C5	Depositor
Number of particles used	837	Depositor
Resolution determination method	Not provided	
CTF correction method	EACH PLATE	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3200	Depositor
Magnification	67000	Depositor
Image detector	FEI EAGLE (4k x 4k)	Depositor
Maximum map value	0.068	Depositor
Minimum map value	-0.027	Depositor
Average map value	-0.001	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.0248	Depositor
Map size ( $\text{\AA}$ )	470.39996, 470.39996, 470.39996	wwPDB
Map dimensions	112, 112, 112	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	4.2, 4.2, 4.2	Depositor



## 5 Model quality i

### 5.1 Standard geometry i

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	0.86	13/3868 (0.3%)	1.52	86/5242 (1.6%)
1	B	0.86	13/3868 (0.3%)	1.52	85/5242 (1.6%)
1	C	0.86	13/3868 (0.3%)	1.52	85/5242 (1.6%)
1	D	0.86	13/3868 (0.3%)	1.52	86/5242 (1.6%)
1	E	0.86	12/3868 (0.3%)	1.52	84/5242 (1.6%)
All	All	0.86	64/19340 (0.3%)	1.52	426/26210 (1.6%)

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	14
1	B	0	13
1	C	0	14
1	D	0	14
1	E	0	14
All	All	0	69

The worst 5 of 64 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	223	TRP	NE1-CE2	-7.16	1.28	1.37
1	A	223	TRP	NE1-CE2	-7.11	1.28	1.37
1	E	223	TRP	NE1-CE2	-7.10	1.28	1.37
1	C	223	TRP	NE1-CE2	-7.10	1.28	1.37
1	B	223	TRP	NE1-CE2	-7.07	1.28	1.37

The worst 5 of 426 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	77	TRP	CH2-CZ2-CE2	14.72	132.12	117.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	77	TRP	CH2-CZ2-CE2	14.68	132.08	117.40
1	E	77	TRP	CH2-CZ2-CE2	14.66	132.06	117.40
1	B	77	TRP	CH2-CZ2-CE2	14.65	132.05	117.40
1	C	77	TRP	CH2-CZ2-CE2	14.63	132.03	117.40

There are no chirality outliers.

5 of 69 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	144	GLY	Peptide
1	A	194	SER	Mainchain
1	A	21	ALA	Mainchain
1	A	82	LEU	Peptide
1	A	91	LYS	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	3785	0	3794	1055	0
1	B	3785	0	3794	1141	0
1	C	3785	0	3794	1131	0
1	D	3785	0	3794	1139	0
1	E	3785	0	3794	1143	0
All	All	18925	0	18970	5609	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 149.

The worst 5 of 5609 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)
1:E:83:LYS:CE	1:E:192:PRO:HA	1.71	1.20
1:B:83:LYS:CE	1:B:192:PRO:HA	1.71	1.19
1:D:83:LYS:CE	1:D:192:PRO:HA	1.71	1.18

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:83:LYS:CE	1:C:192:PRO:HA	1.71	1.18
1:B:124:VAL:HB	1:B:125:ILE:HB	1.18	1.17

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	474/476 (100%)	396 (84%)	47 (10%)	31 (6%)	1	16
1	B	474/476 (100%)	396 (84%)	47 (10%)	31 (6%)	1	16
1	C	474/476 (100%)	396 (84%)	47 (10%)	31 (6%)	1	16
1	D	474/476 (100%)	396 (84%)	47 (10%)	31 (6%)	1	16
1	E	474/476 (100%)	396 (84%)	47 (10%)	31 (6%)	1	16
All	All	2370/2380 (100%)	1980 (84%)	235 (10%)	155 (6%)	2	16

5 of 155 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	83	LYS
1	A	92	GLU
1	A	125	ILE
1	A	132	ALA
1	A	182	GLN

### 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	411/411 (100%)	408 (99%)	3 (1%)	84	90
1	B	411/411 (100%)	407 (99%)	4 (1%)	76	86
1	C	411/411 (100%)	408 (99%)	3 (1%)	84	90
1	D	411/411 (100%)	407 (99%)	4 (1%)	76	86
1	E	411/411 (100%)	408 (99%)	3 (1%)	84	90
All	All	2055/2055 (100%)	2038 (99%)	17 (1%)	82	89

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	192	PRO
1	E	395	ASP
1	C	192	PRO
1	C	317	PRO
1	C	395	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	234	ASN
1	E	344	HIS
1	E	439	HIS
1	B	411	GLN
1	B	355	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides 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.

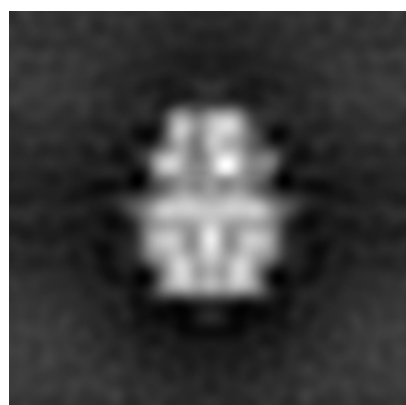
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-2356. These allow visual inspection of the internal detail of the map and identification of artifacts.

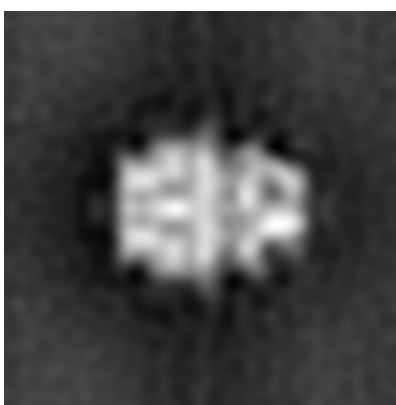
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

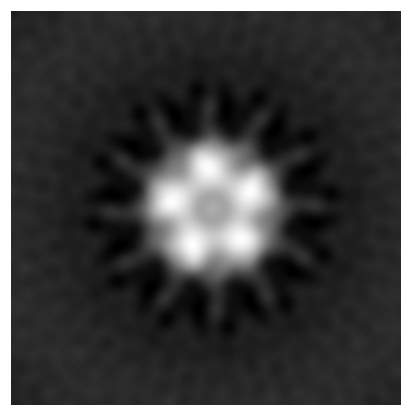
#### 6.1.1 Primary map



X



Y

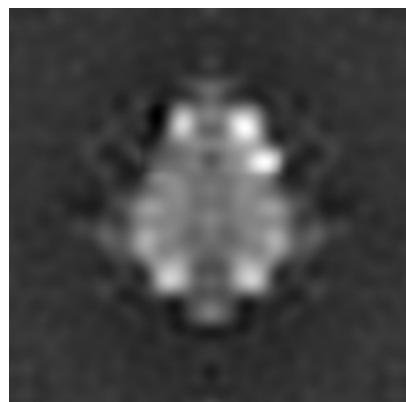


Z

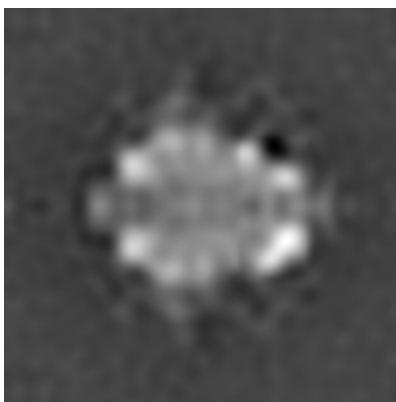
The images above show the map projected in three orthogonal directions.

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

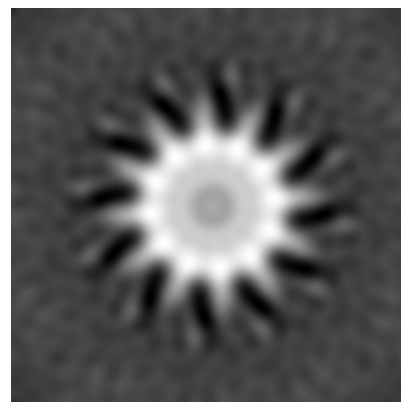
#### 6.2.1 Primary map



X Index: 56



Y Index: 56

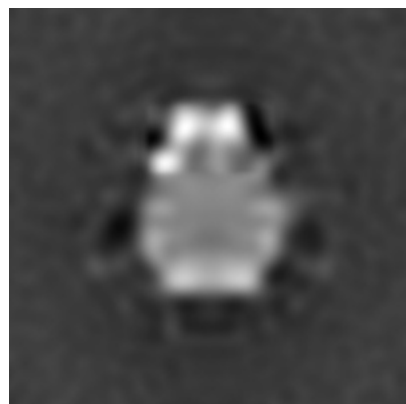


Z Index: 56

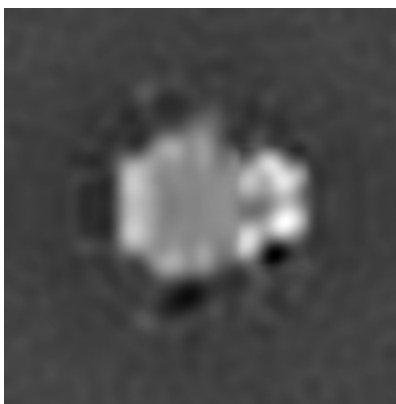
The images above show central slices of the map in three orthogonal directions.

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

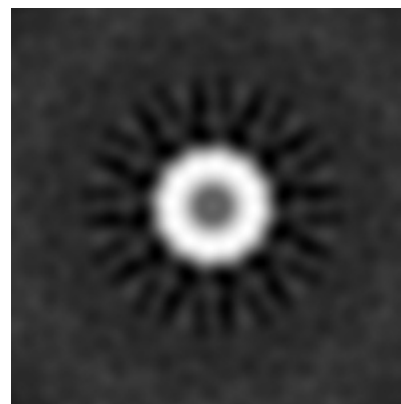
### 6.3.1 Primary map



X Index: 64



Y Index: 47

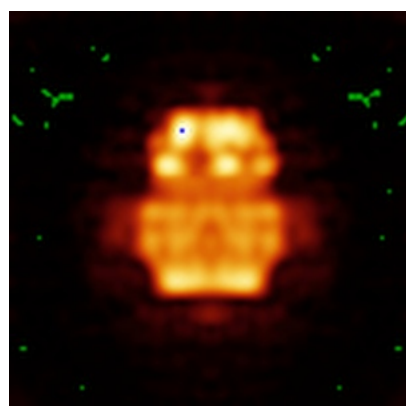


Z Index: 36

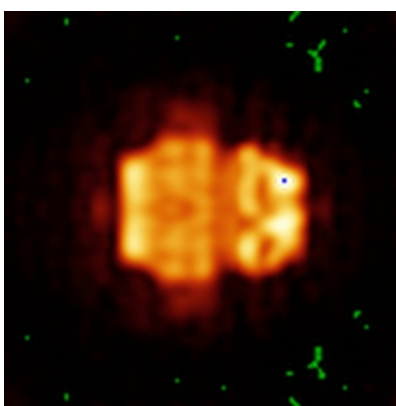
The images above show the largest variance slices of the map in three orthogonal directions.

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

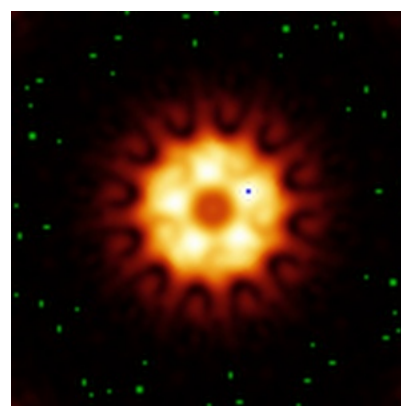
### 6.4.1 Primary map



X



Y

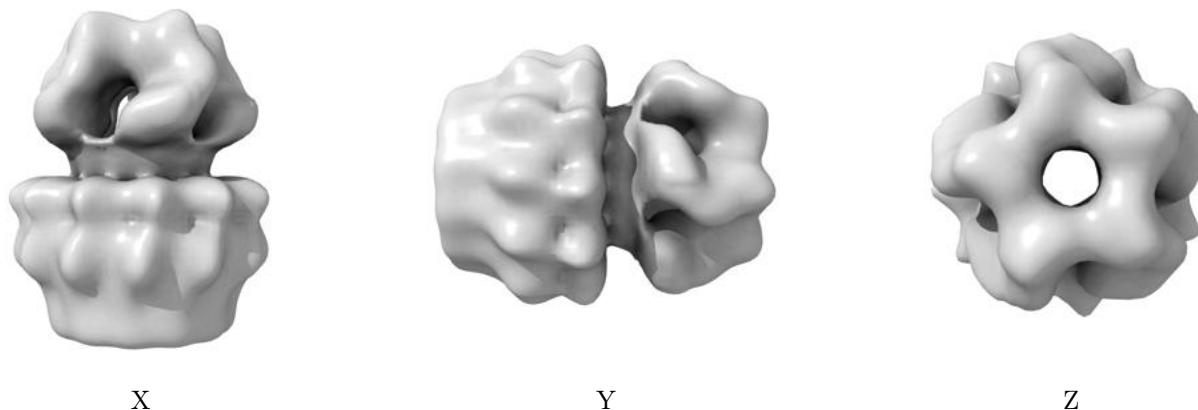


Z

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 0.0248. 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](#)

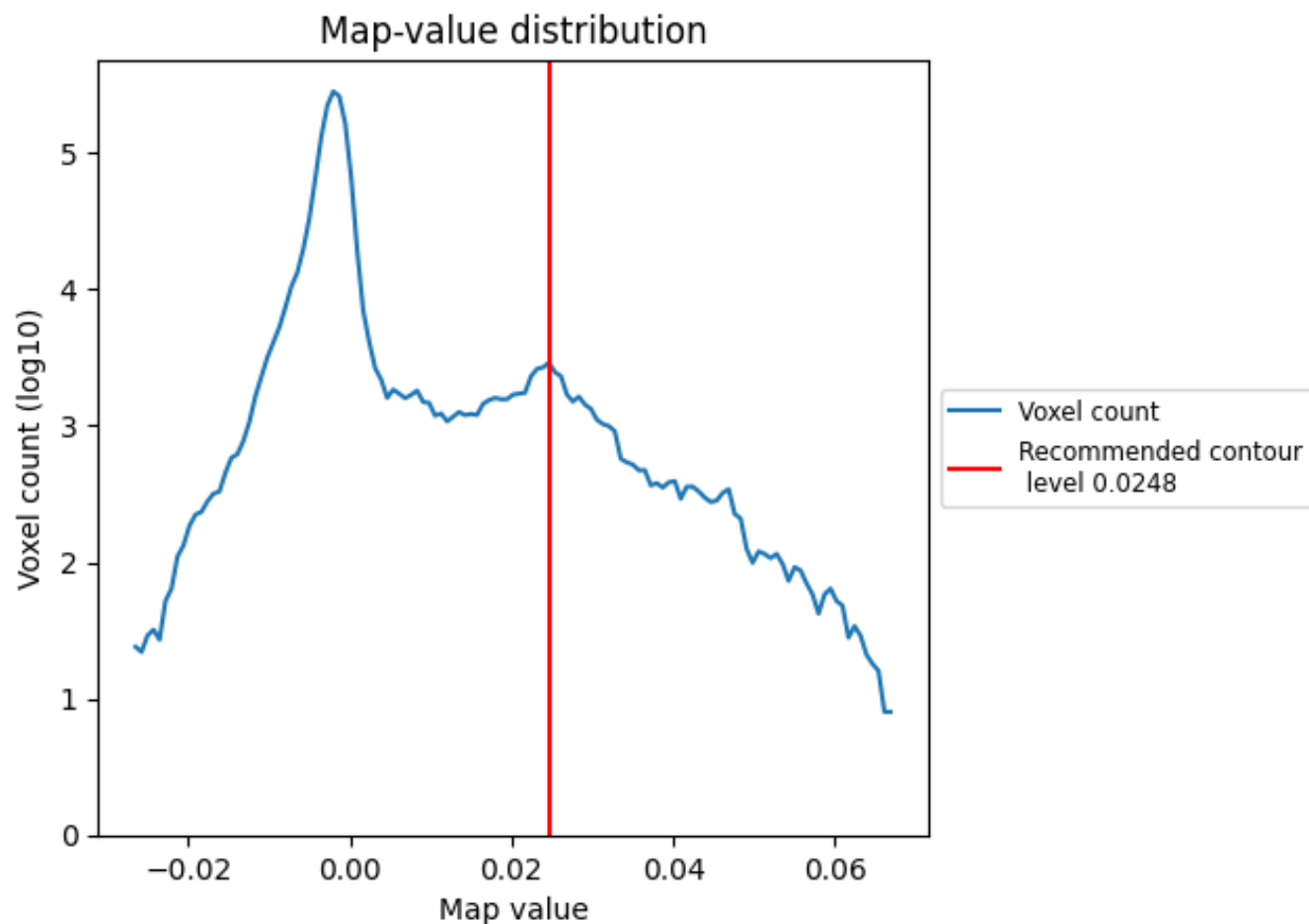
This section was not generated. No masks/segmentation were deposited.



## 7 Map analysis [i](#)

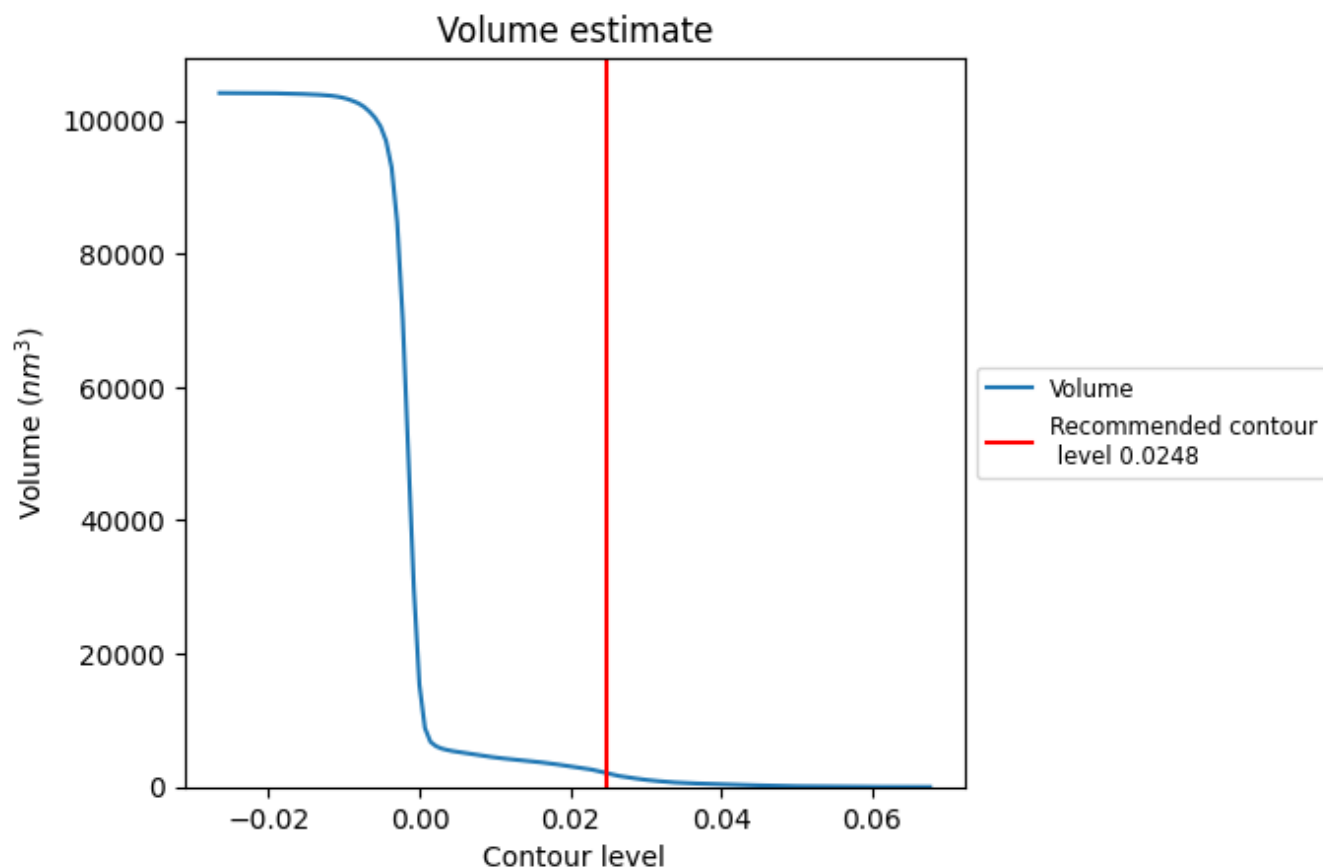
This section contains the results of statistical analysis of the map.

### 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.

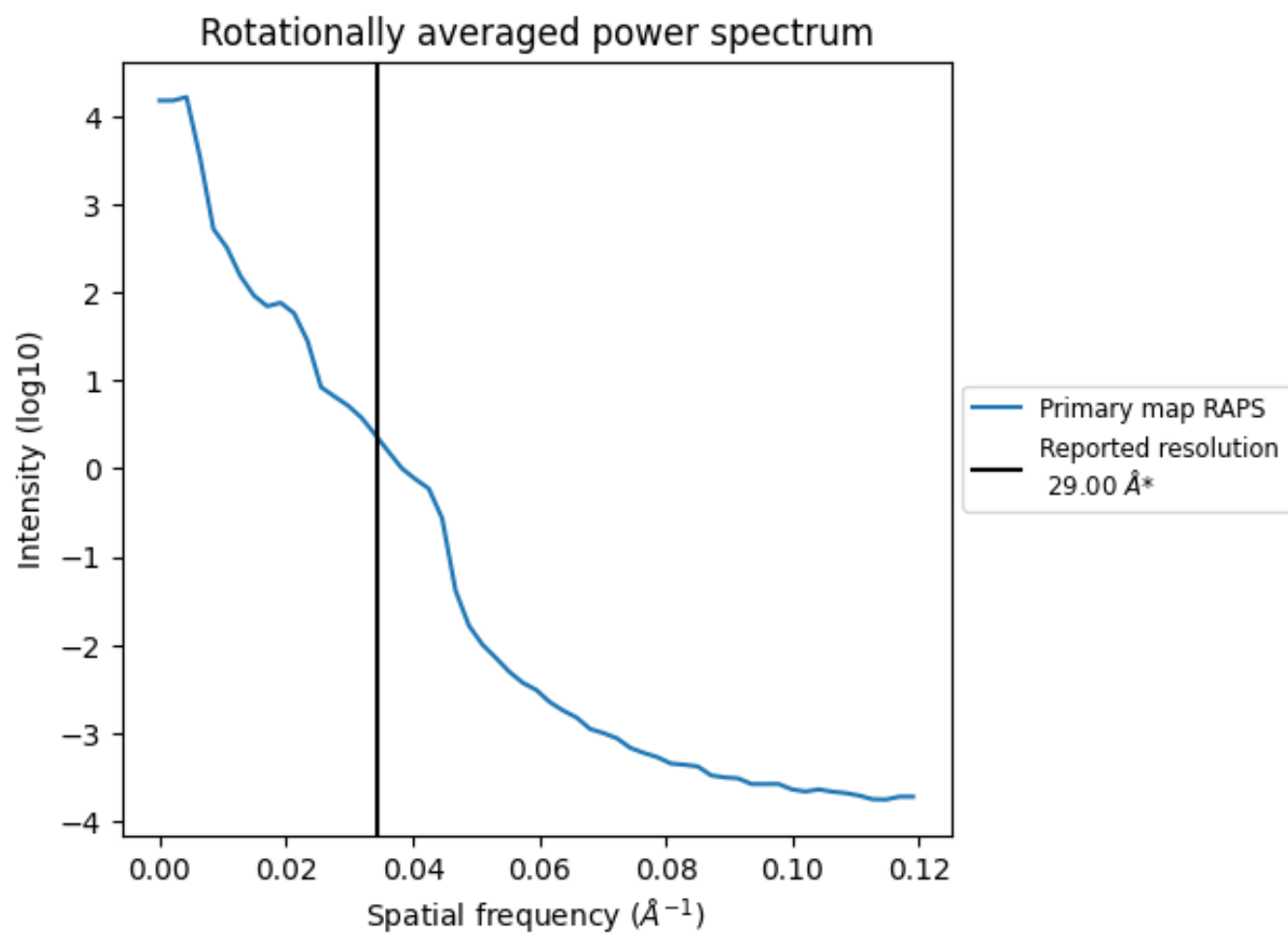
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 2076 nm<sup>3</sup>; this corresponds to an approximate mass of 1875 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 0.034 Å<sup>-1</sup>

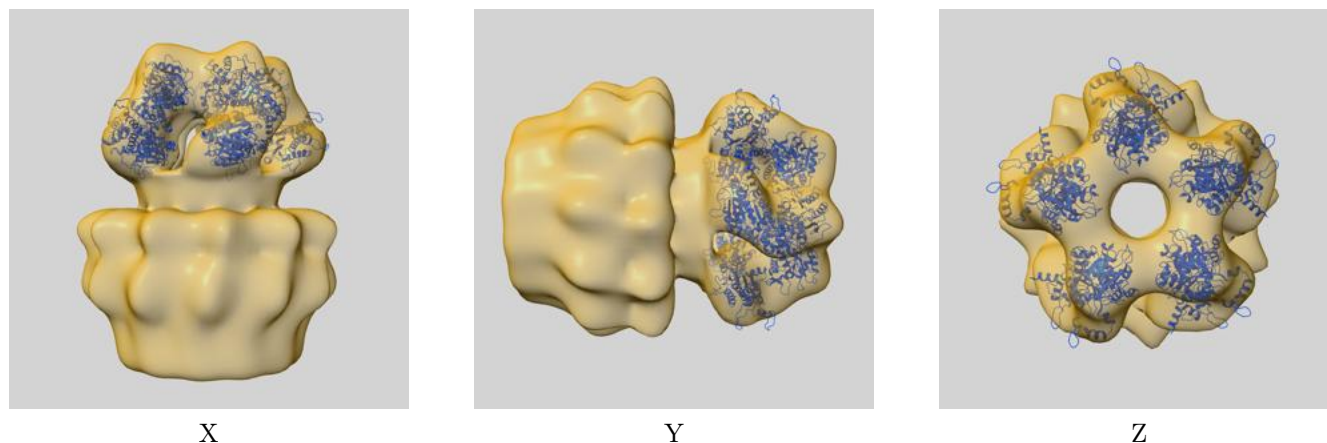
## 8 Fourier-Shell correlation

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

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

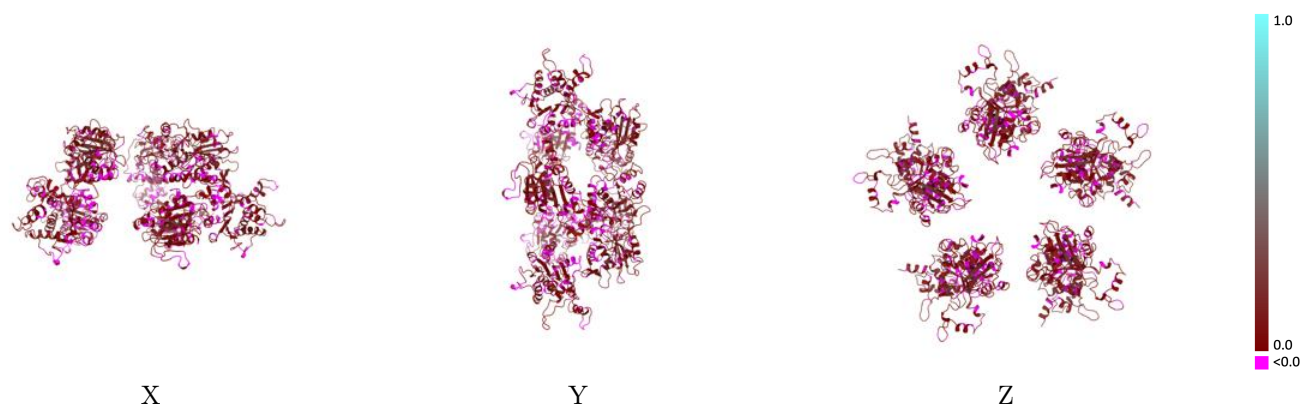
This section contains information regarding the fit between EMDB map EMD-2356 and PDB model 4BIL. Per-residue inclusion information can be found in section [3](#) on page [4](#).

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



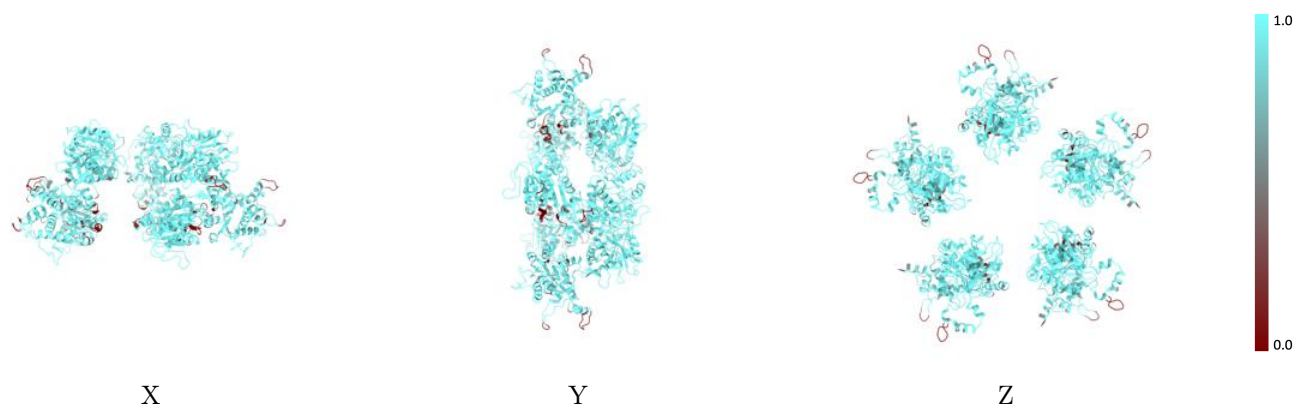
The images above show the 3D surface view of the map at the recommended contour level 0.0248 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



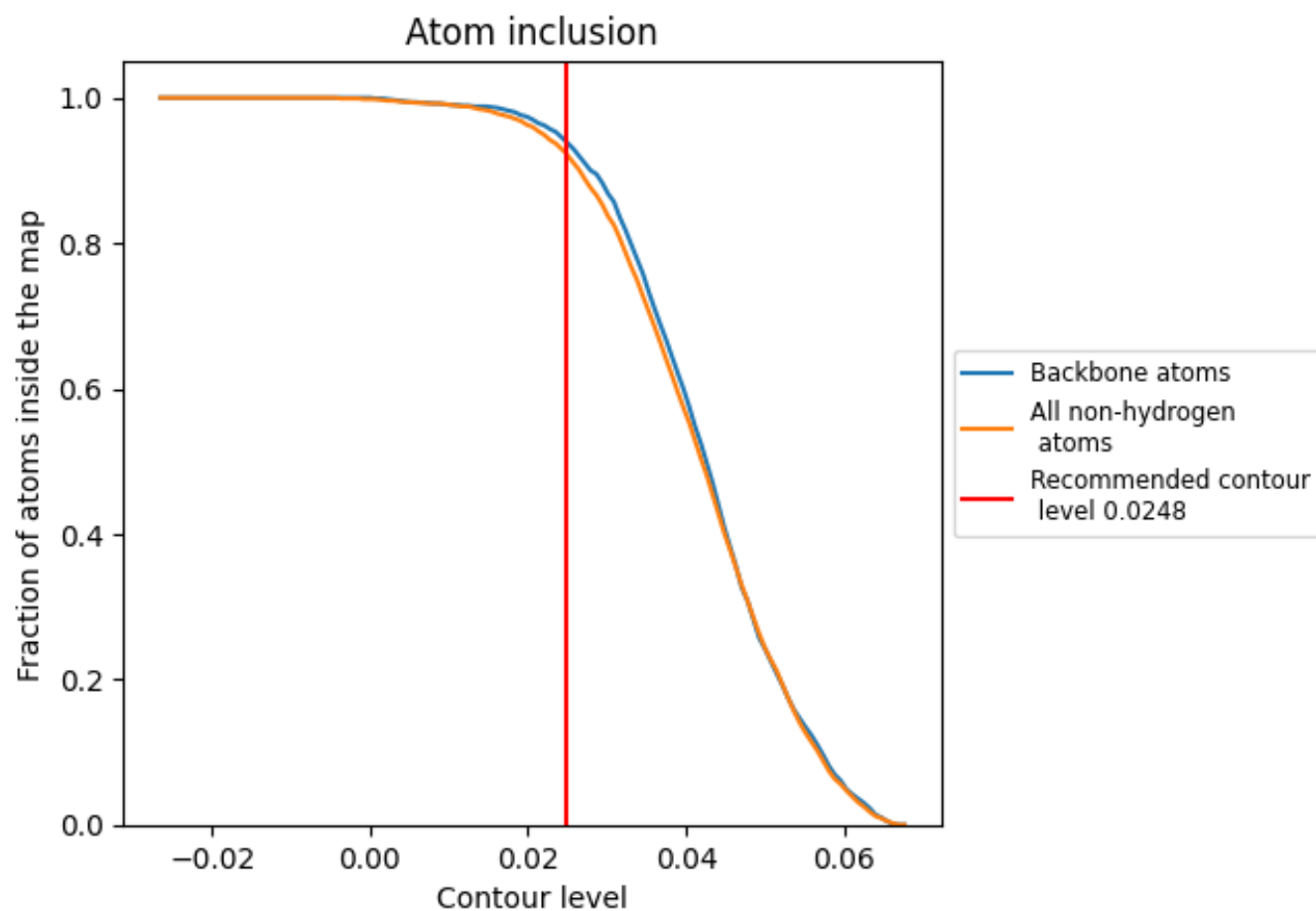
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.0248).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 94% of all backbone atoms, 92% 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 (0.0248) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div></div> 0.9240	<div></div> 0.0510
A	<div></div> 0.9240	<div></div> 0.0500
B	<div></div> 0.9240	<div></div> 0.0520
C	<div></div> 0.9240	<div></div> 0.0520
D	<div></div> 0.9240	<div></div> 0.0520
E	<div></div> 0.9230	<div></div> 0.0490

