



## wwPDB EM Validation Summary Report ⓘ

Feb 12, 2024 – 08:47 PM EST

PDB ID : 3J1N  
EMDB ID : EMD-5407  
Title : Cryo-EM map of a yeast minimal preinitiation complex interacting with the Mediator Head module  
Authors : Asturias, F.J.; Imasaki, T.  
Deposited on : 2012-03-29  
Resolution : 16.00 Å(reported)  
Based on initial model : 1WCM

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:

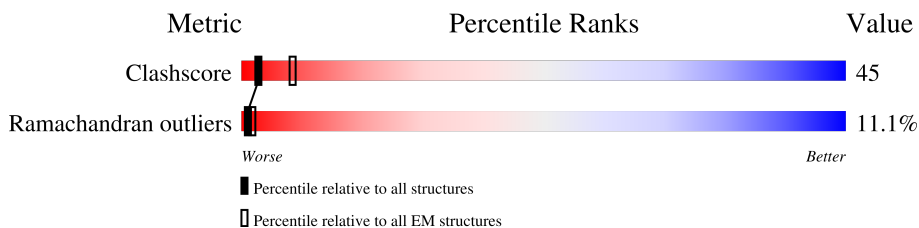
EMDB validation analysis : 0.0.1.dev70  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:  
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 16.00 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.



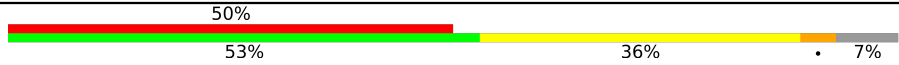
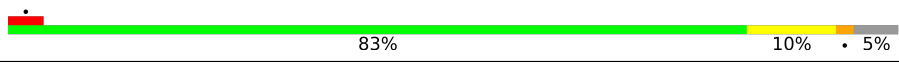

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023

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	1455	
2	B	1224	
3	C	268	
4	D	218	
5	E	215	
6	F	84	
7	G	171	
8	H	146	
9	I	122	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
10	J	70	
11	K	120	
12	L	70	

## 2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 19237 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 II subunit RPB1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	A	1412	6964	4140	1412	1412	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	B	1094	5397	3209	1094	1094	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	C	266	1317	785	266	266	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	D	177	878	524	177	177	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPABC1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	E	214	1062	634	214	214	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPABC2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	F	84	417	249	84	84	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
7	G	171	841	499	171	171	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerase II subunit RPABC3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
8	H	133	659	393	133	133	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
9	I	118	587	351	118	118	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPABC5.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
10	J	65	321	191	65	65	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
11	K	114	565	337	114	114	0	0

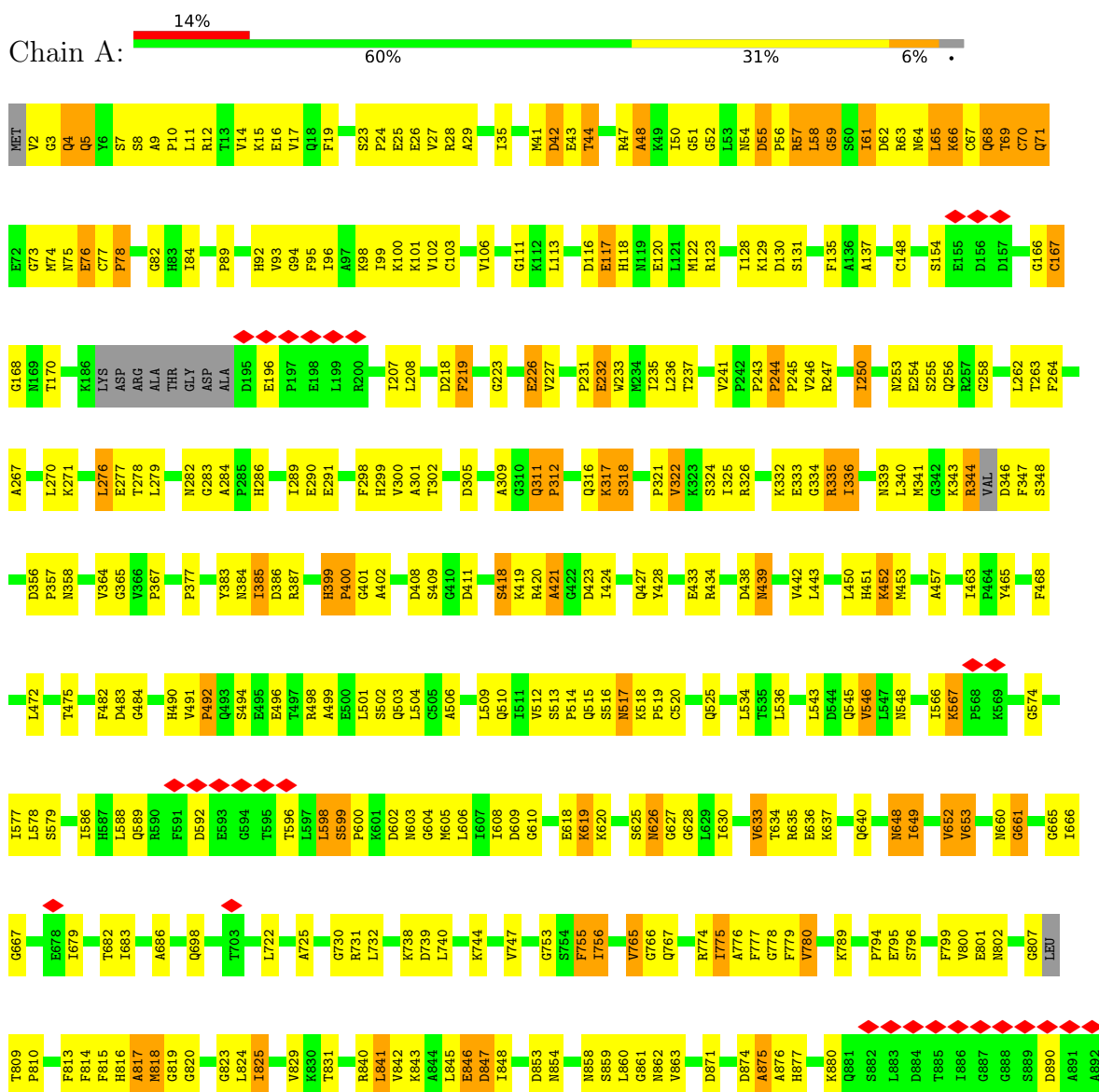
- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPABC4.

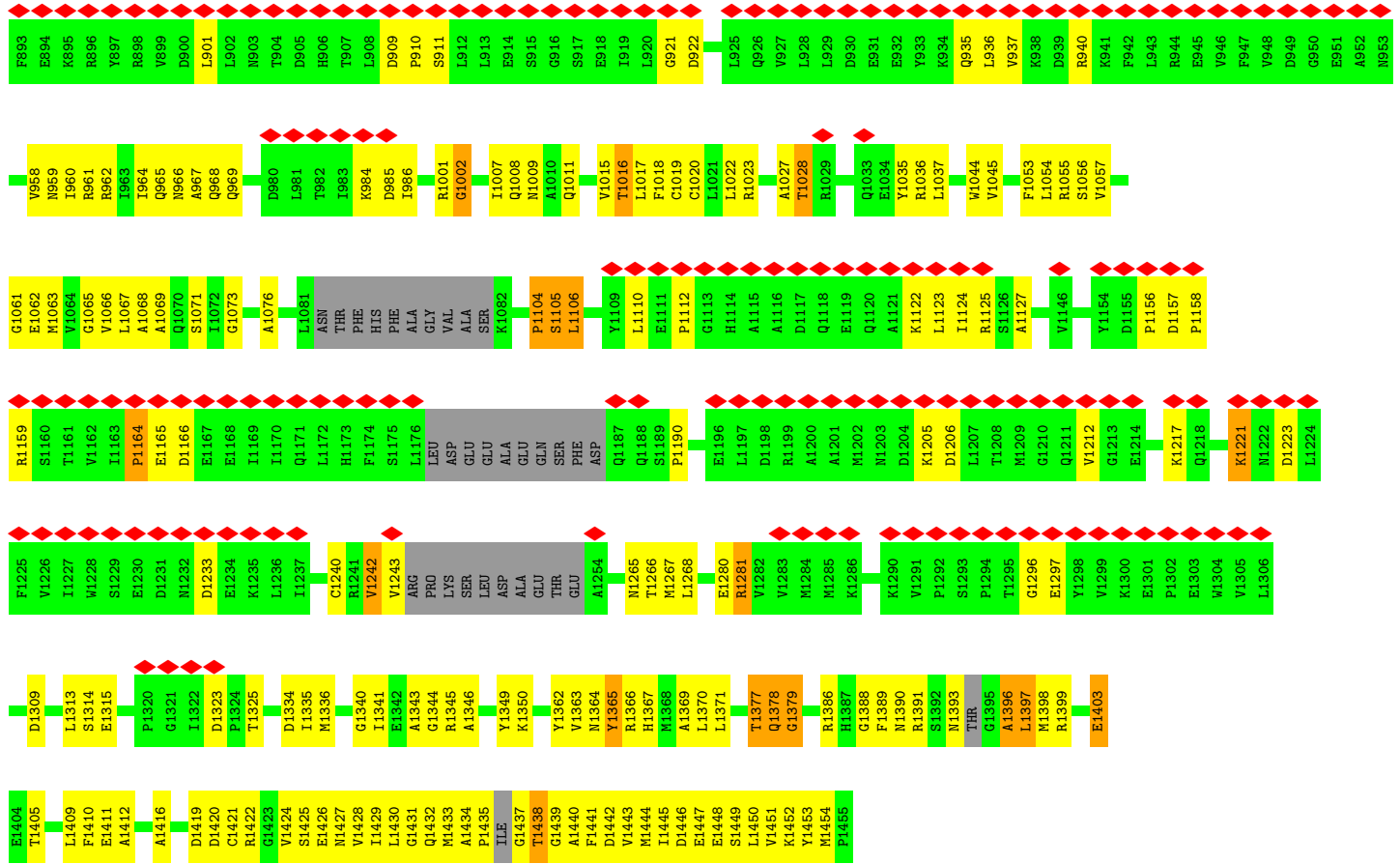
Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
12	L	46	229	137	46	46	0	0

### 3 Residue-property plots

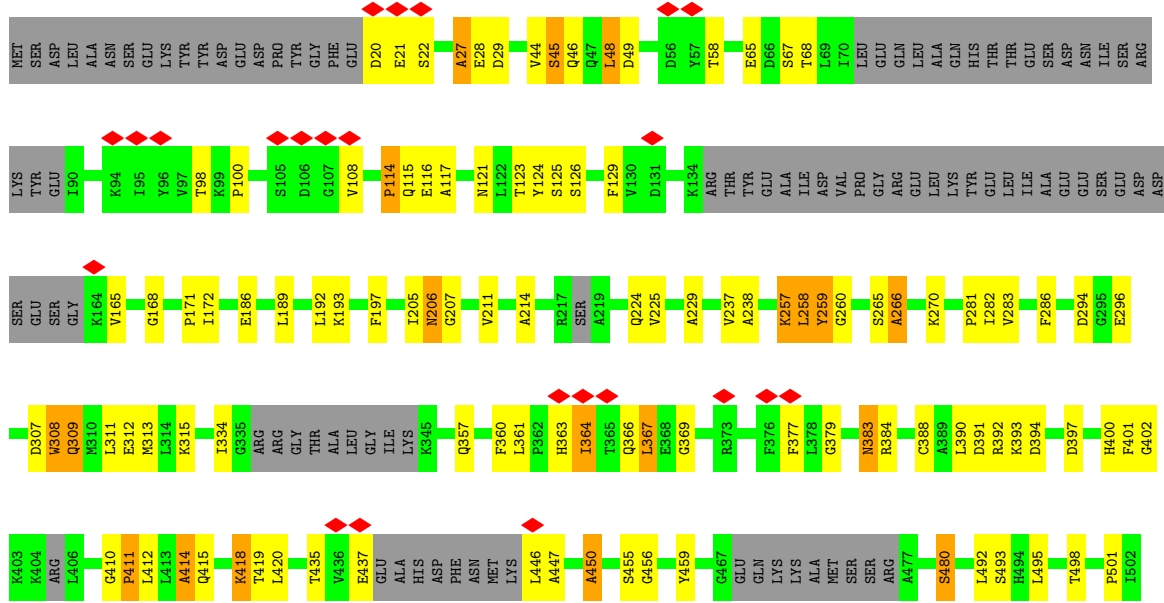
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

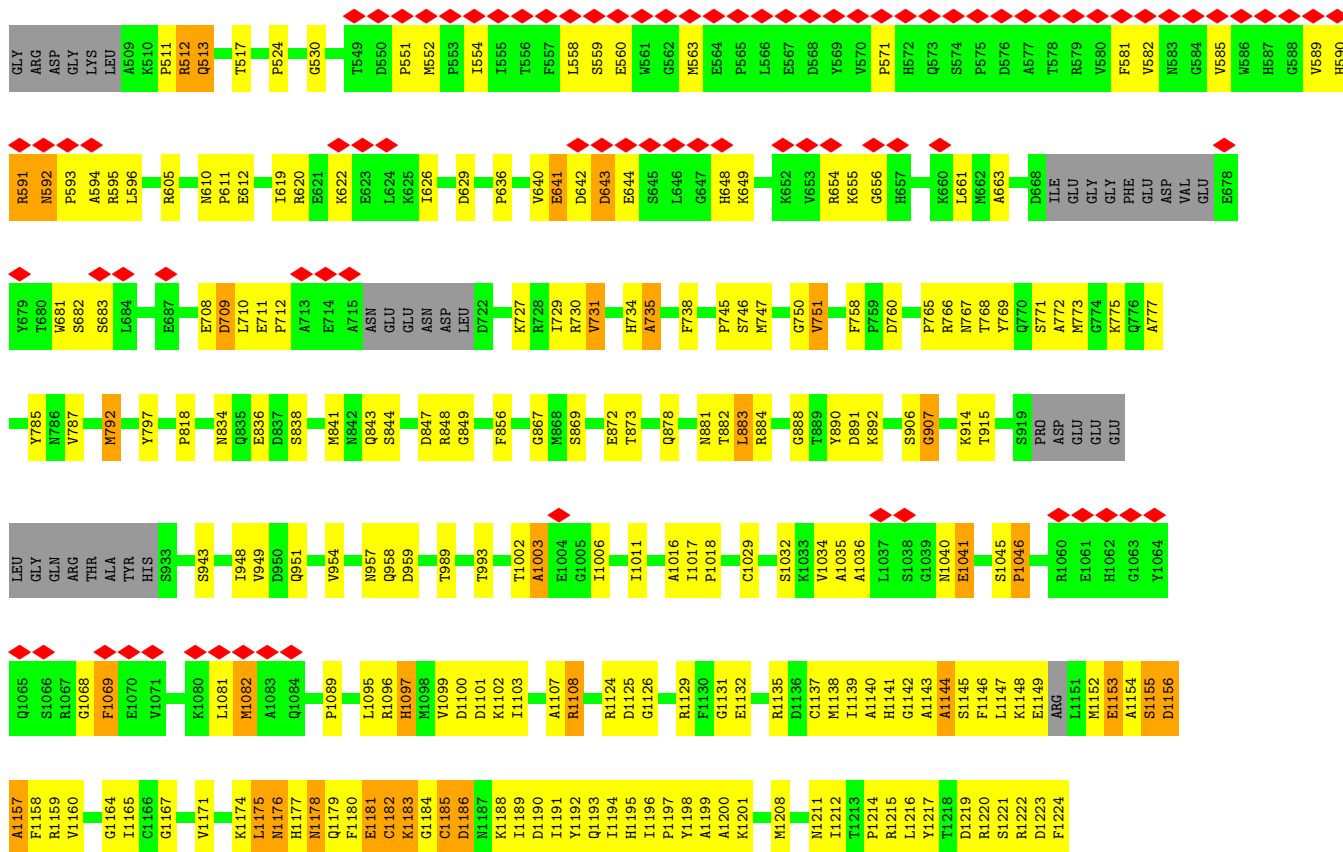
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1



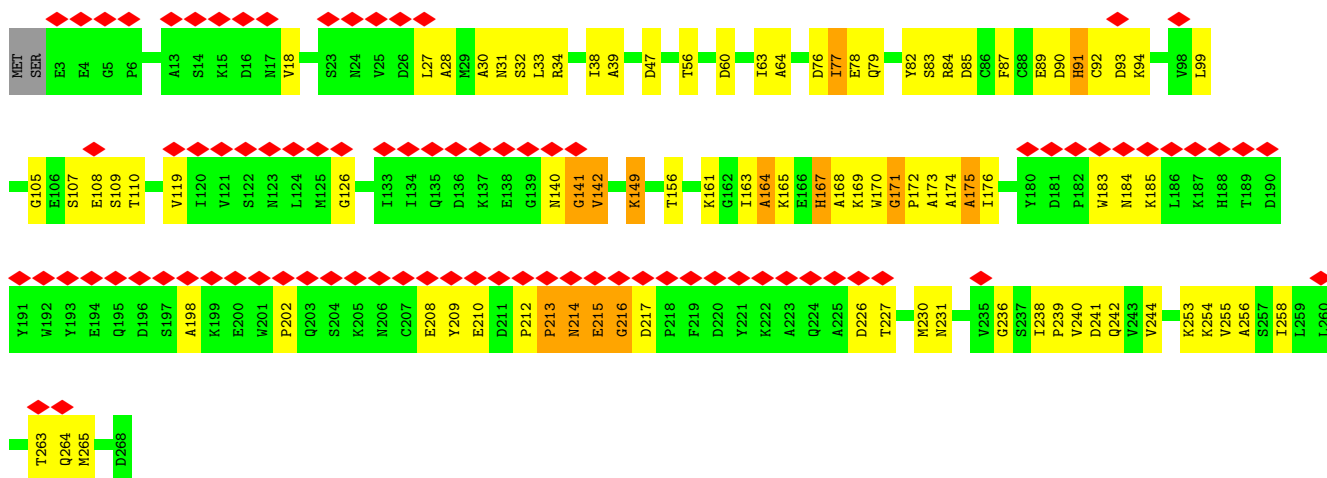


• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

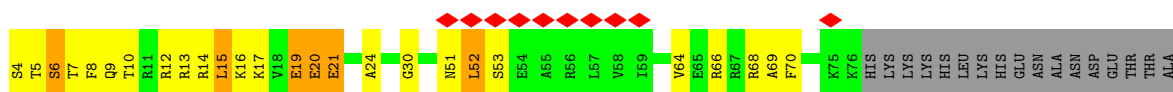




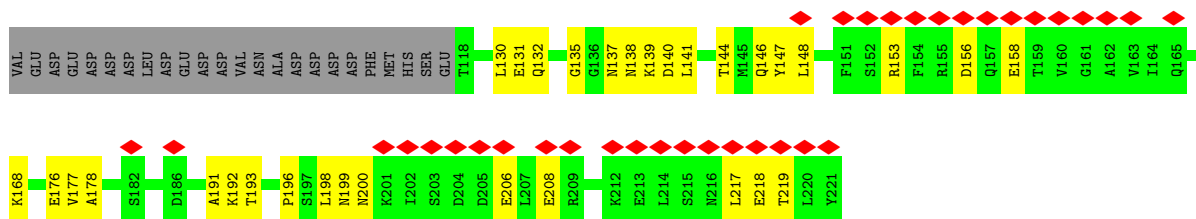
● Molecule 3: DNA-directed RNA polymerase II subunit RPB3



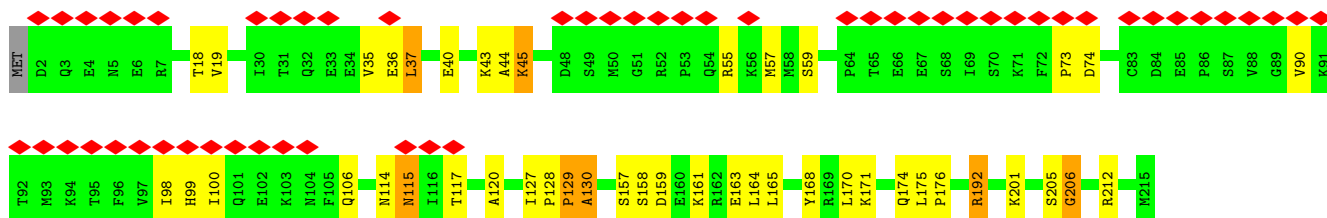
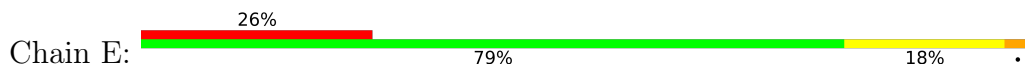
● Molecule 4: DNA-directed RNA polymerase II subunit RPB4







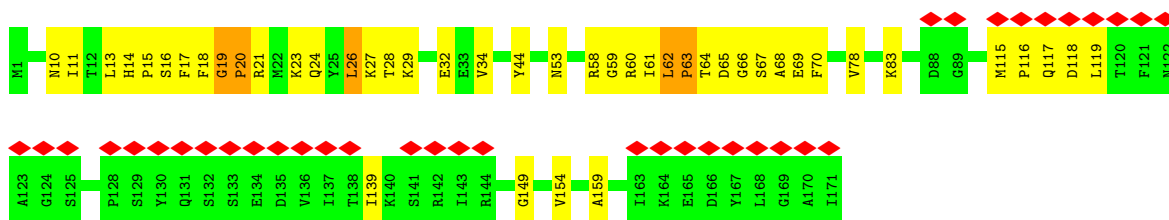
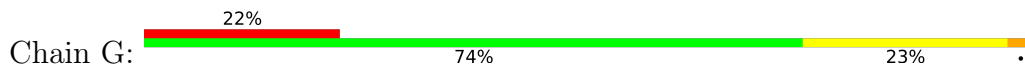
• Molecule 5: DNA-directed RNA polymerase II subunit RPABC1



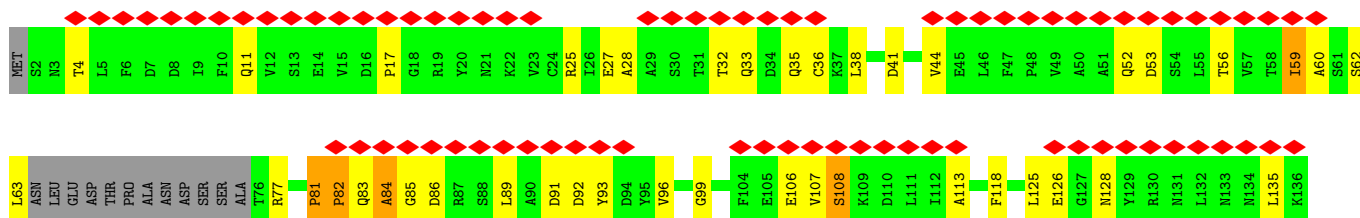
• Molecule 6: DNA-directed RNA polymerase II subunit RPABC2

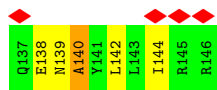


• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

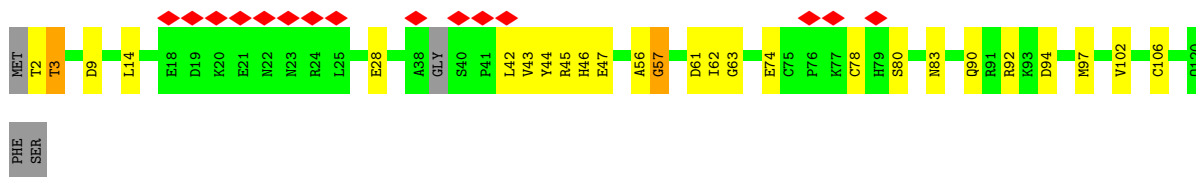
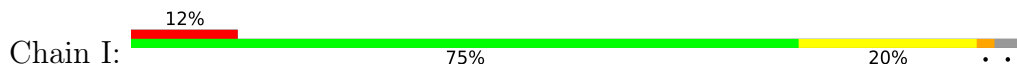


• Molecule 8: DNA-directed RNA polymerase II subunit RPABC3

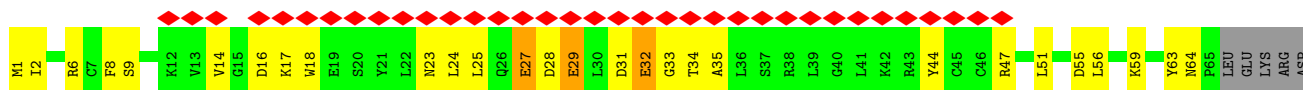




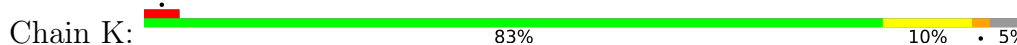
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9



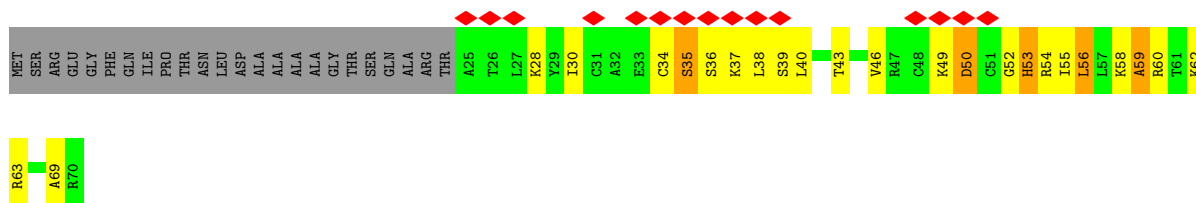
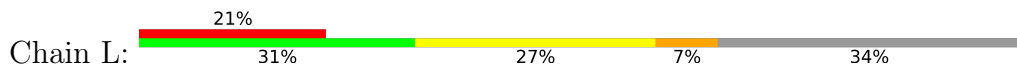
• Molecule 10: DNA-directed RNA polymerase II subunit RPABC5



• Molecule 11: DNA-directed RNA polymerase II subunit RPB11



• Molecule 12: DNA-directed RNA polymerase II subunit RPABC4



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	51000	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	Each CCD frame	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	120	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	10	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	3800	Depositor
Magnification	50000	Depositor
Image detector	TVIPS TEMCAM-F415 (4k x 4k)	Depositor
Maximum map value	32.476	Depositor
Minimum map value	-4.340	Depositor
Average map value	0.557	Depositor
Map value standard deviation	2.374	Depositor
Recommended contour level	3.25	Depositor
Map size ( $\text{\AA}$ )	268.1, 268.1, 268.1	wwPDB
Map dimensions	70, 70, 70	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	3.83, 3.83, 3.83	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	0.41	0/6953	0.77	3/9659 (0.0%)
2	B	0.39	0/5384	0.74	1/7473 (0.0%)
3	C	0.44	0/1316	0.81	0/1833
4	D	0.36	0/876	0.75	1/1219 (0.1%)
5	E	0.34	0/1061	0.66	0/1479
6	F	0.48	0/416	0.80	0/579
7	G	0.45	0/840	0.80	0/1166
8	H	0.34	0/657	0.69	0/913
9	I	0.43	0/585	0.82	0/814
10	J	0.45	0/320	0.88	0/444
11	K	0.40	0/564	0.70	0/785
12	L	0.52	0/228	0.82	0/317
All	All	0.41	0/19200	0.76	5/26681 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1403	GLU	N-CA-C	5.38	125.52	111.00
2	B	1185	CYS	N-CA-C	-5.30	96.70	111.00
1	A	452	LYS	N-CA-C	-5.21	96.93	111.00
4	D	7	THR	N-CA-C	5.14	124.88	111.00
1	A	344	ARG	N-CA-C	-5.04	97.38	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6964	0	3039	671	0
2	B	5397	0	2399	433	0
3	C	1317	0	590	53	0
4	D	878	0	387	63	0
5	E	1062	0	453	43	0
6	F	417	0	180	123	0
7	G	841	0	354	181	0
8	H	659	0	296	30	0
9	I	587	0	235	31	0
10	J	321	0	137	15	0
11	K	565	0	260	6	0
12	L	229	0	107	12	0
All	All	19237	0	8437	1250	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 45.

The worst 5 of 1250 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)
6:F:143:PHE:CB	7:G:68:ALA:CA	1.76	1.61
1:A:344:ARG:CA	2:B:1129:ARG:HA	1.27	1.60
6:F:143:PHE:CA	7:G:68:ALA:HB1	1.11	1.58
2:B:1148:LYS:HA	2:B:1200:ALA:CB	1.22	1.57
1:A:1444:MET:CB	7:G:10:ASN:CB	1.82	1.57

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	1390/1455 (96%)	938 (68%)	290 (21%)	162 (12%)	0	6
2	B	1068/1224 (87%)	727 (68%)	220 (21%)	121 (11%)	0	7
3	C	264/268 (98%)	159 (60%)	66 (25%)	39 (15%)	0	3
4	D	173/218 (79%)	122 (70%)	34 (20%)	17 (10%)	0	10
5	E	212/215 (99%)	148 (70%)	49 (23%)	15 (7%)	1	14
6	F	82/84 (98%)	64 (78%)	14 (17%)	4 (5%)	2	20
7	G	169/171 (99%)	131 (78%)	26 (15%)	12 (7%)	1	14
8	H	129/146 (88%)	84 (65%)	30 (23%)	15 (12%)	0	6
9	I	114/122 (93%)	77 (68%)	30 (26%)	7 (6%)	1	17
10	J	63/70 (90%)	37 (59%)	10 (16%)	16 (25%)	0	1
11	K	112/120 (93%)	89 (80%)	18 (16%)	5 (4%)	2	22
12	L	44/70 (63%)	19 (43%)	14 (32%)	11 (25%)	0	1
All	All	3820/4163 (92%)	2595 (68%)	801 (21%)	424 (11%)	1	7

5 of 424 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	48	ALA
1	A	54	ASN
1	A	55	ASP
1	A	57	ARG

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 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](#)

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1274:ARG	C	1275:GLY	N	10.90
1	A	1131:THR	C	1142:THR	N	8.08

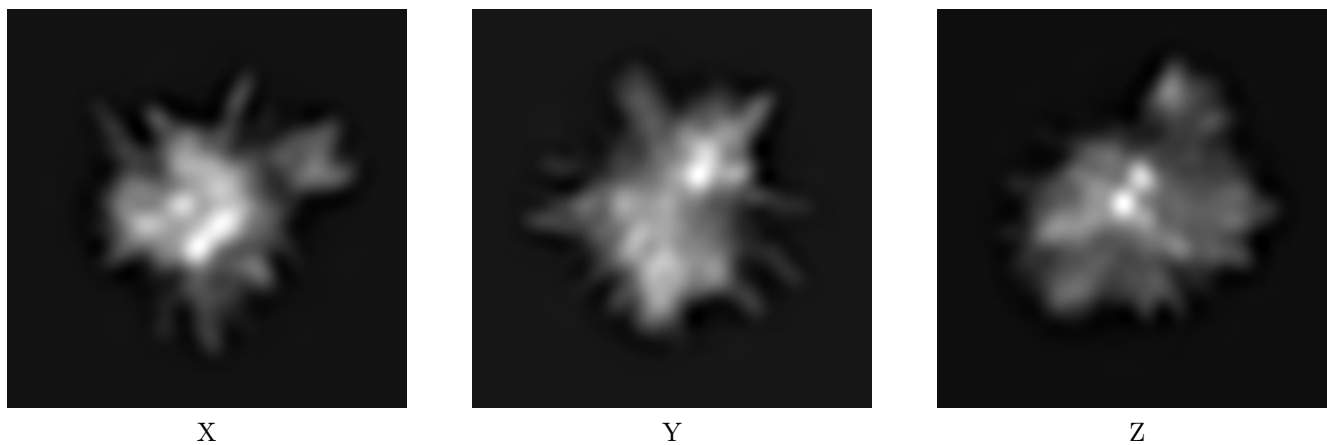
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-5407. 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](#)

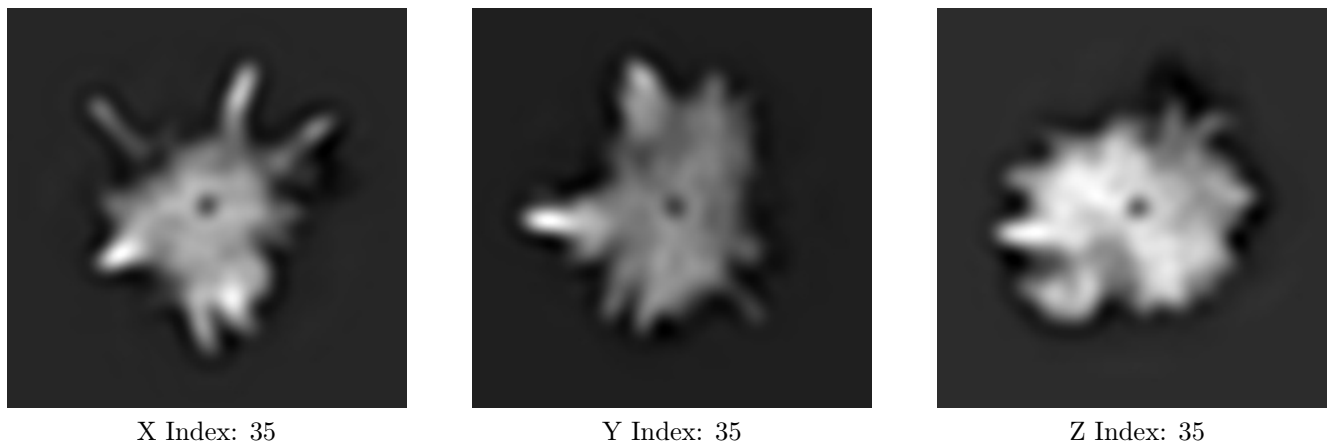
#### 6.1.1 Primary map



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: 41



Y Index: 35

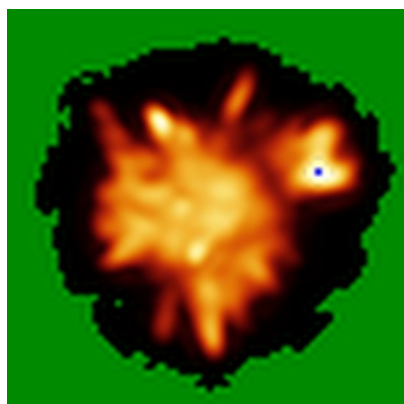


Z Index: 40

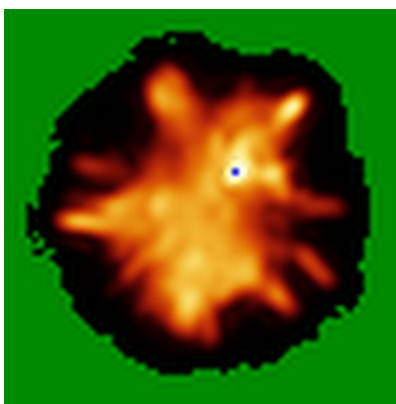
The images above show the largest variance slices of the map in three orthogonal directions.

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

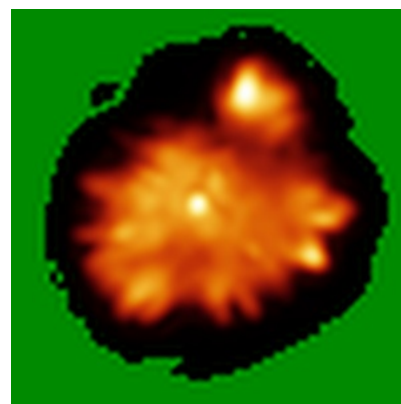
### 6.4.1 Primary map



X



Y

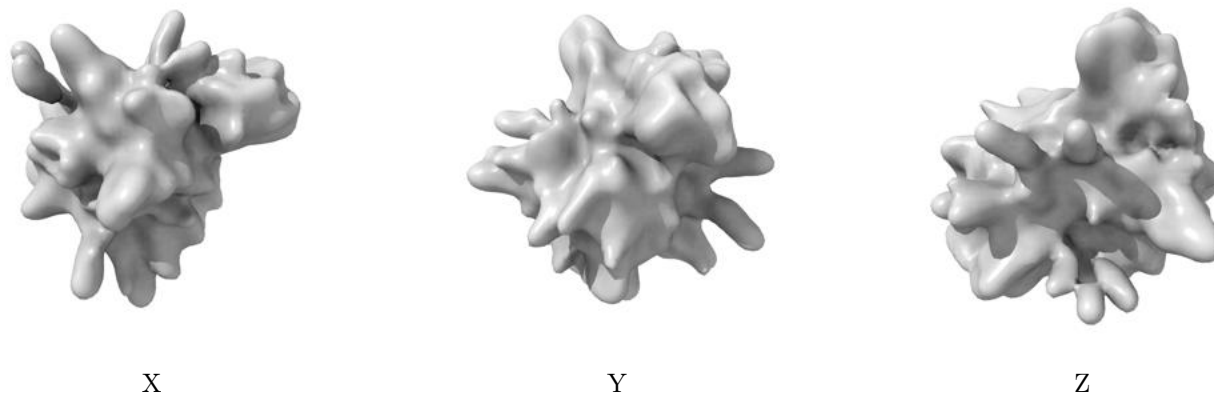


Z

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 3.25. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

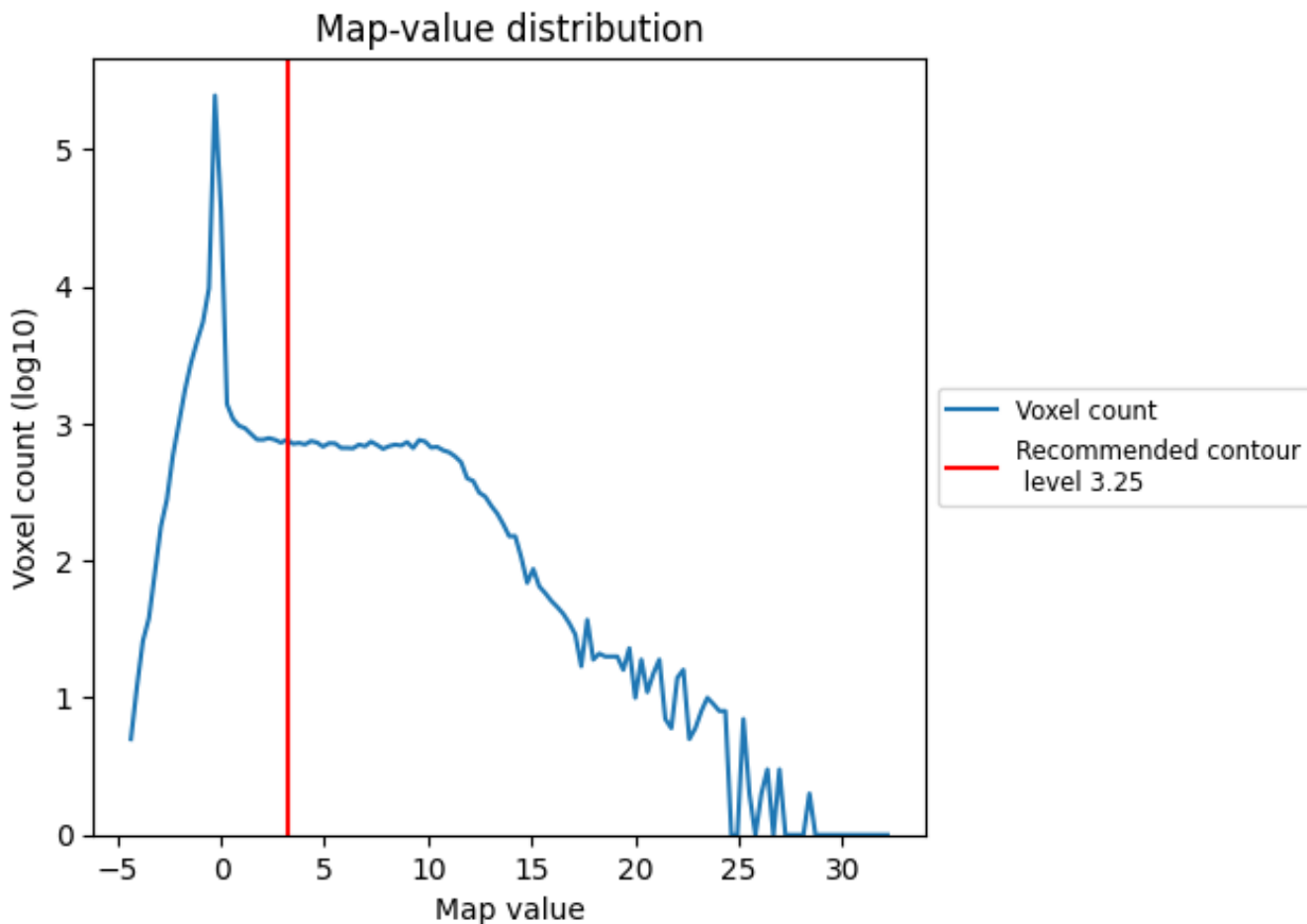
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

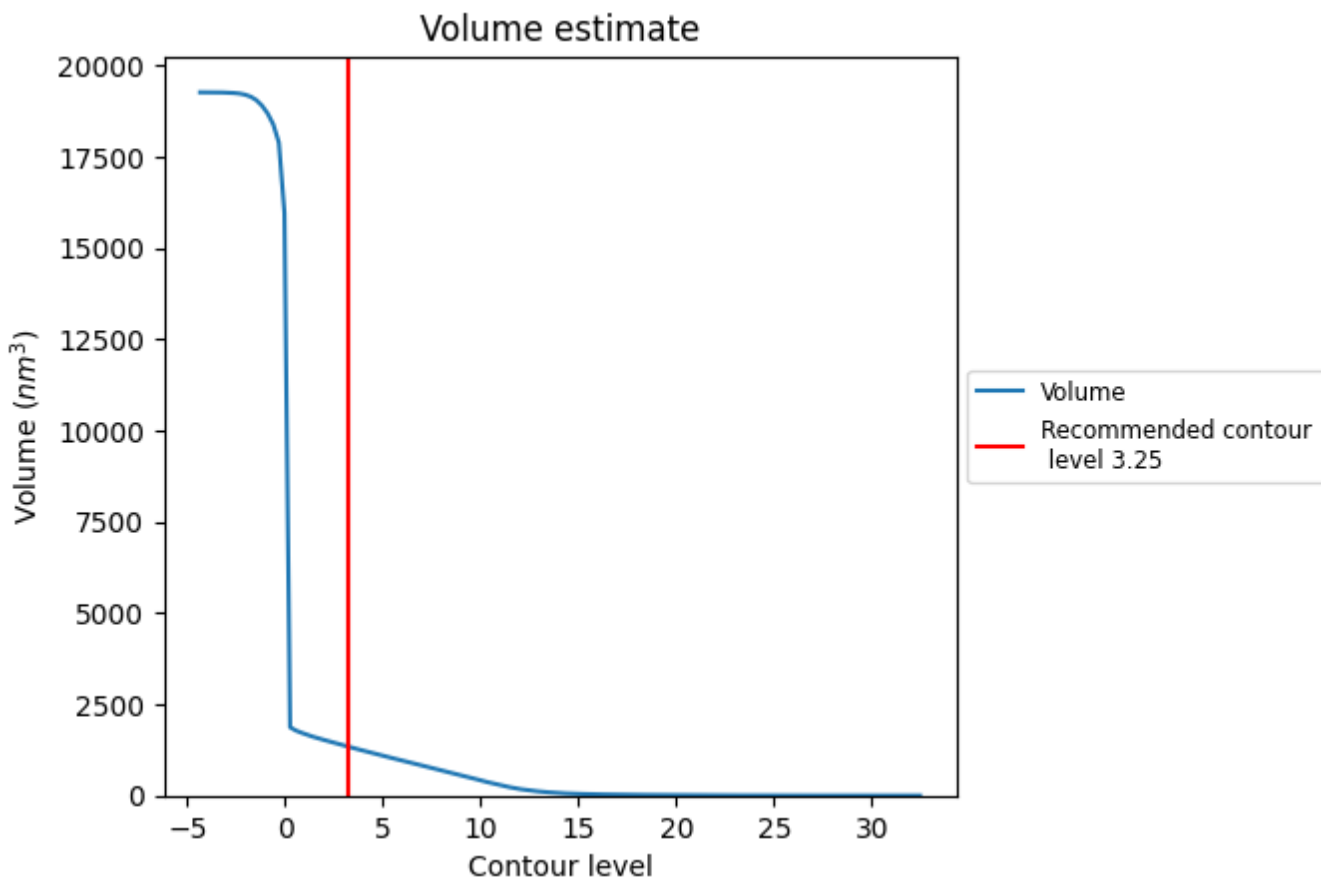
This section contains the results of statistical analysis of the map.

### 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.

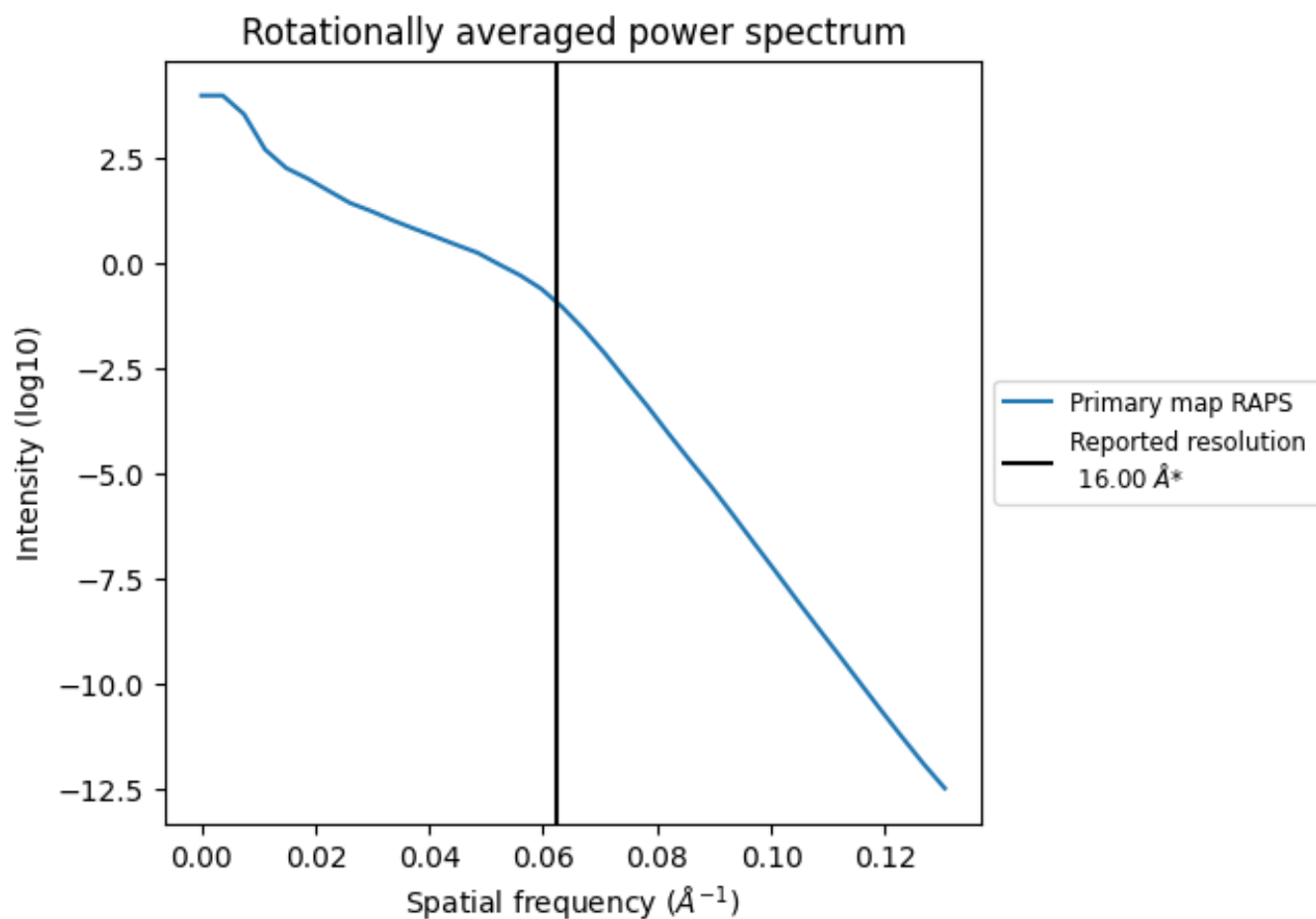
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 1339 nm<sup>3</sup>; this corresponds to an approximate mass of 1210 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  $0.062 \text{\AA}^{-1}$

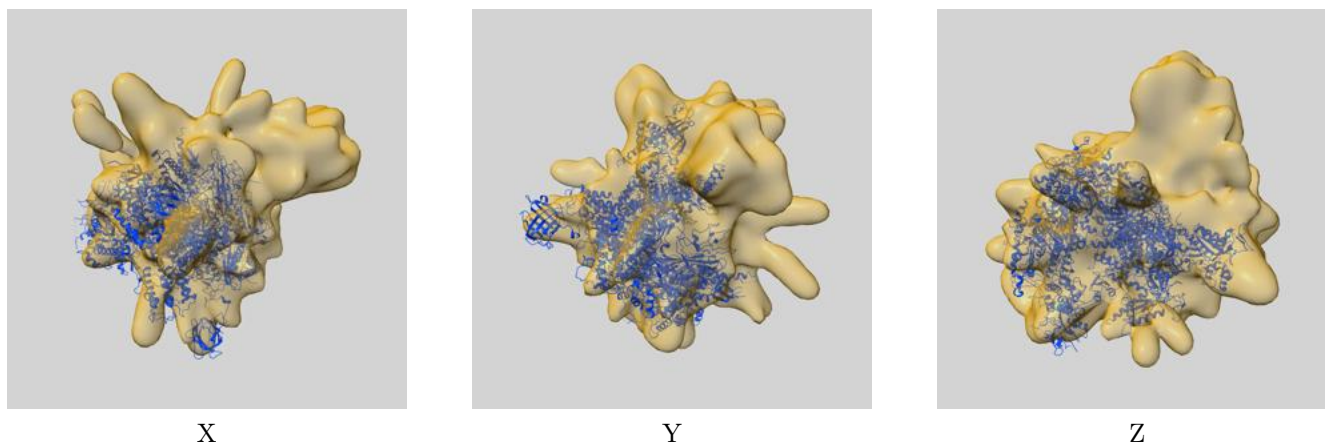
## 8 Fourier-Shell correlation

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

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

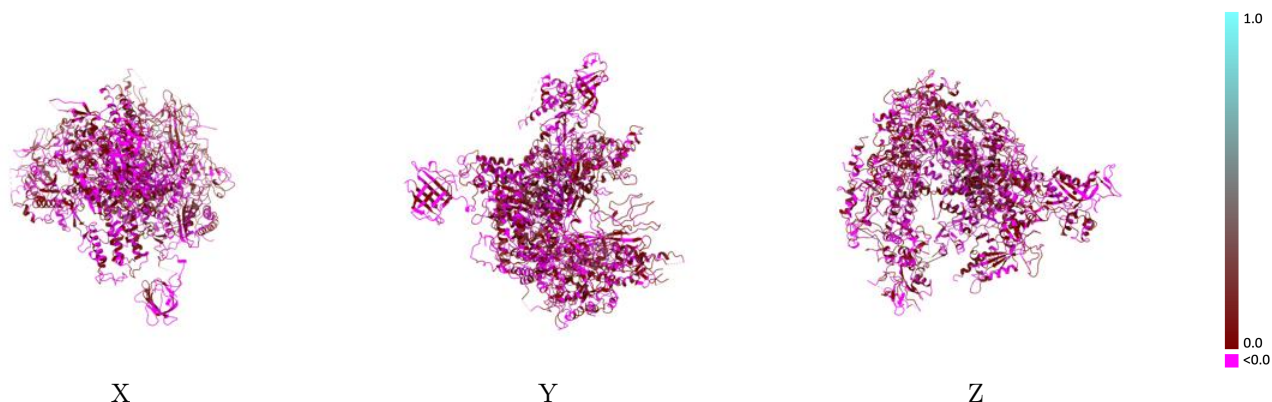
This section contains information regarding the fit between EMDB map EMD-5407 and PDB model 3J1N. Per-residue inclusion information can be found in section 3 on page 6.

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



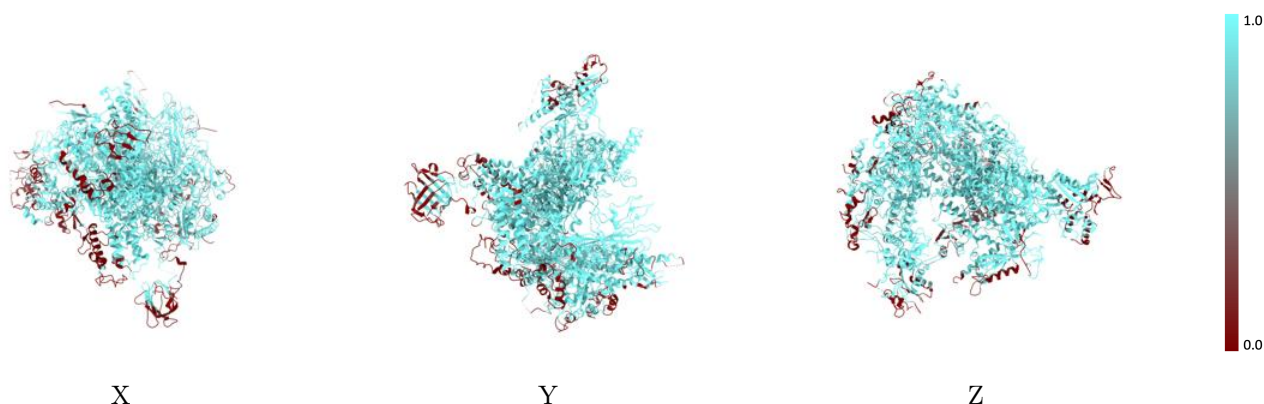
The images above show the 3D surface view of the map at the recommended contour level 3.25 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 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.

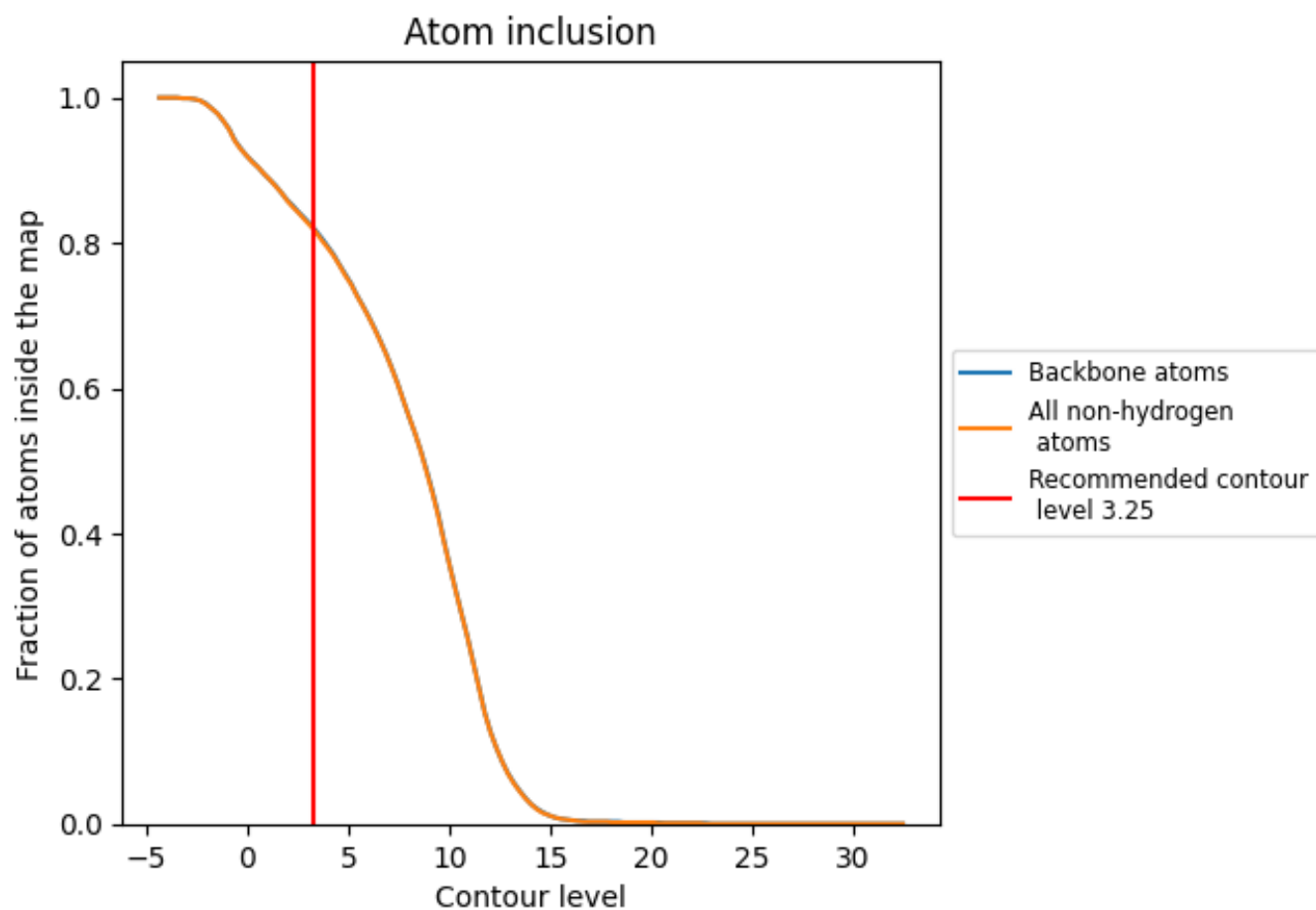
## 9.3 Atom inclusion mapped to coordinate model [i](#)



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 (3.25).





























## 9.4 Atom inclusion [i](#)



At the recommended contour level, 82% of all backbone atoms, 82% 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 (3.25) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8200	 0.0150
A	 0.8530	 0.0170
B	 0.8990	 0.0240
C	 0.6700	 0.0100
D	 0.7480	 0.0150
E	 0.7410	 0.0060
F	 1.0000	 0.0350
G	 0.7780	 -0.0200
H	 0.3760	 -0.0470
I	 0.8620	 0.0560
J	 0.4580	 -0.0390
K	 0.9500	 0.0370
L	 0.6510	 -0.0100

