



# wwPDB EM Validation Summary Report ⓘ

May 18, 2024 – 08:34 PM EDT

PDB ID : 7M8E  
EMDB ID : EMD-23716  
Title : E.coli RNAP-RapA elongation complex  
Authors : Shi, W.; Liu, B.  
Deposited on : 2021-03-29  
Resolution : 3.40 Å (reported)  
Based on initial model : 4S20

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:

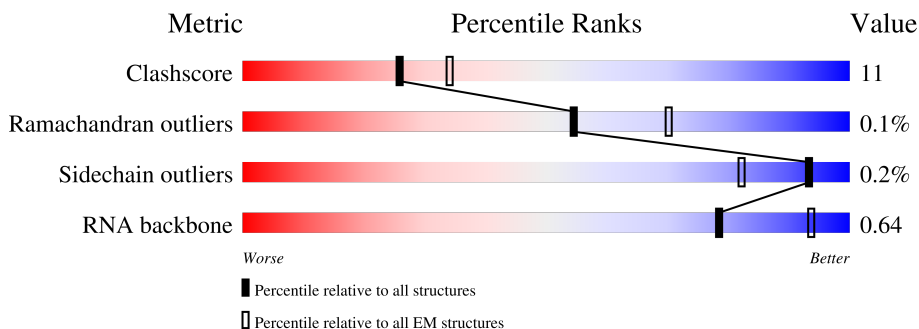
EMDB validation analysis : 0.0.1.dev92  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 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 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.



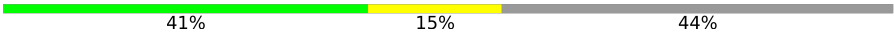

Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	329	
1	B	329	
2	C	1342	
3	D	1416	
4	E	91	
5	F	974	
6	1	39	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	2	39	 41% 15% 44%
8	3	9	 67% 33%

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 33880 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
			Total	C	N	O	S		
1	A	230	Total	C	N	O	S	0	0
			1786	1112	317	351	6		
1	B	228	Total	C	N	O	S	0	0
			1767	1100	312	349	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	1340	Total	C	N	O	S	0	0
			10570	6631	1841	2055	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	1337	Total	C	N	O	S	0	0
			10396	6531	1853	1962	50		

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1408	LEU	-	expression tag	UNP D8ED86
D	1409	GLU	-	expression tag	UNP D8ED86
D	1410	VAL	-	expression tag	UNP D8ED86
D	1411	HIS	-	expression tag	UNP D8ED86
D	1412	HIS	-	expression tag	UNP D8ED86
D	1413	HIS	-	expression tag	UNP D8ED86
D	1414	HIS	-	expression tag	UNP D8ED86
D	1415	HIS	-	expression tag	UNP D8ED86
D	1416	HIS	-	expression tag	UNP D8ED86

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	79	627	382	118	126	1	0	0

- Molecule 5 is a protein called RNA polymerase-associated protein RapA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	967	7719	4831	1382	1477	29	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-5	HIS	-	expression tag	UNP C3TR27
F	-4	HIS	-	expression tag	UNP C3TR27
F	-3	HIS	-	expression tag	UNP C3TR27
F	-2	HIS	-	expression tag	UNP C3TR27
F	-1	HIS	-	expression tag	UNP C3TR27
F	0	HIS	-	expression tag	UNP C3TR27

- Molecule 6 is a DNA chain called Nontemplate DNA (39-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	1	18	370	174	72	106	18	0	0

- Molecule 7 is a DNA chain called Template DNA (39-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	2	22	451	213	81	135	22	0	0

- Molecule 8 is a RNA chain called RNA (5'-R(P\*AP\*UP\*CP\*GP\*GP\*CP\*UP\*CP\*A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	3	9	191	85	33	64	9	0	0

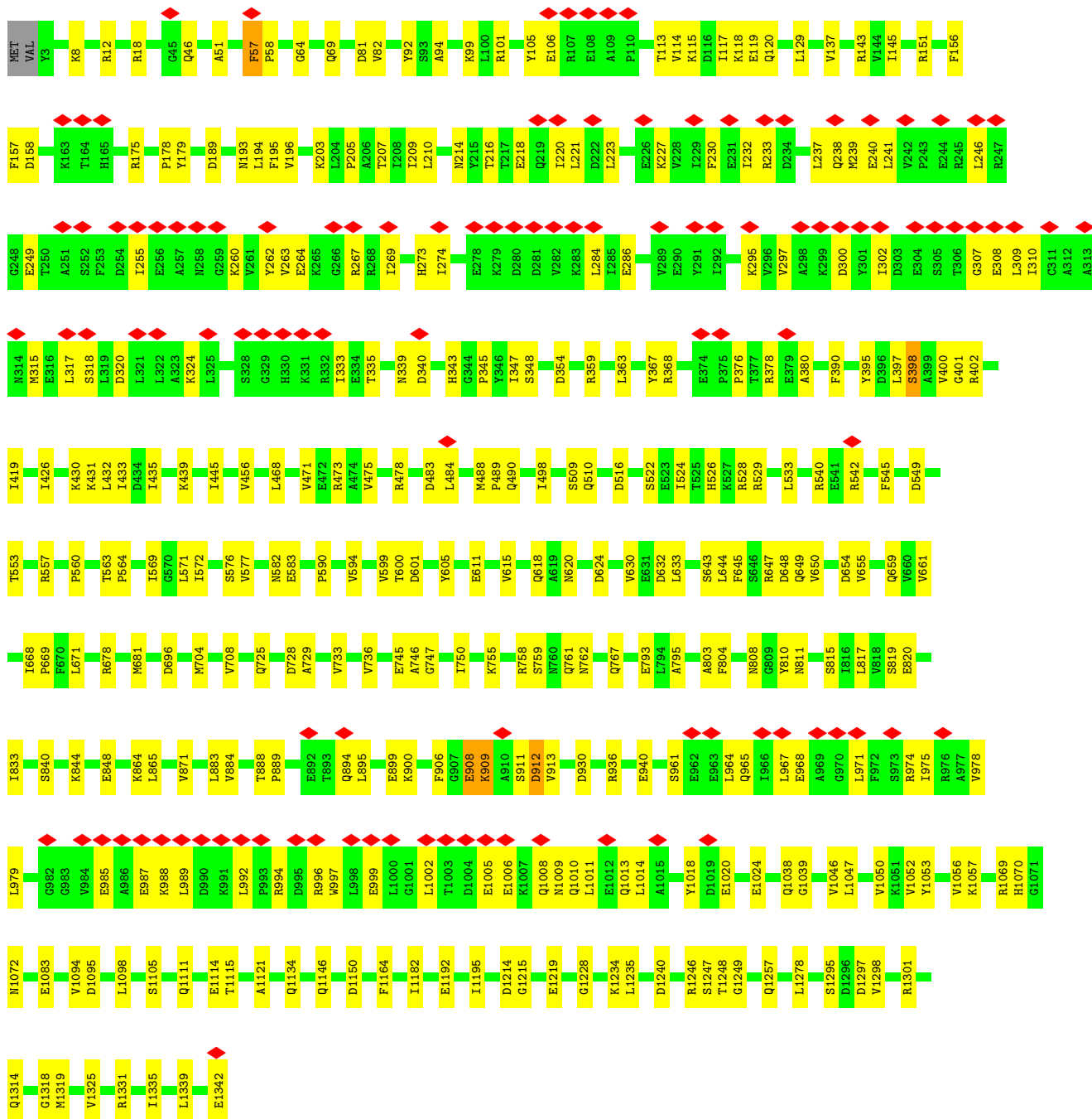
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
9	D	2	Total	Zn	0
			2	2	

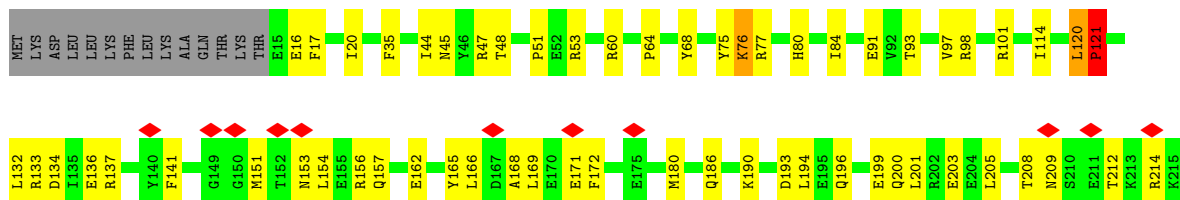
- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
10	D	1	Total	Mg	0
			1	1	

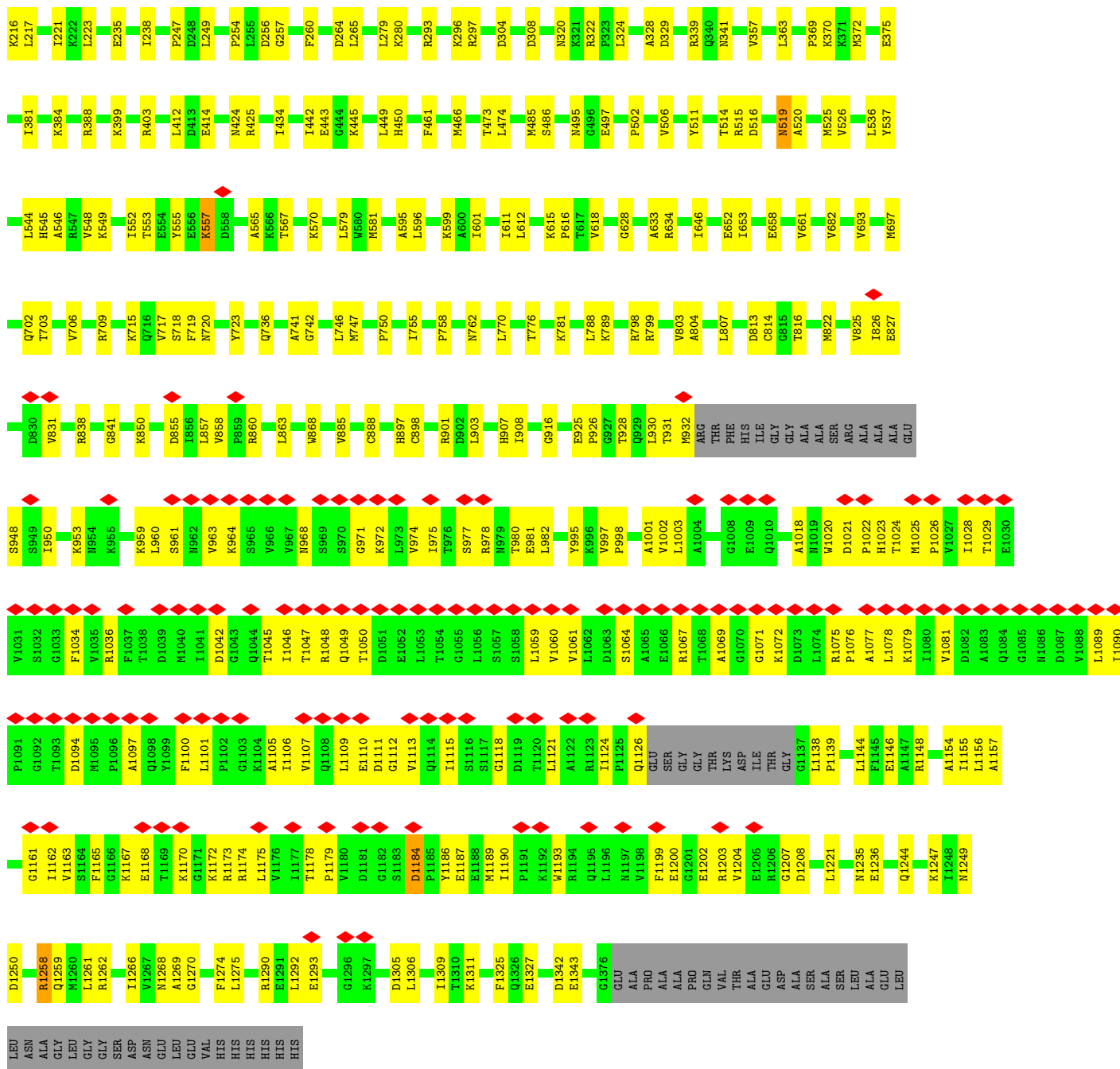




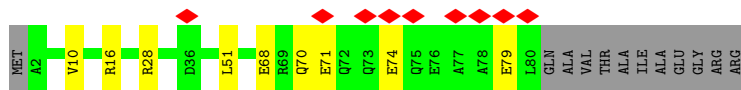
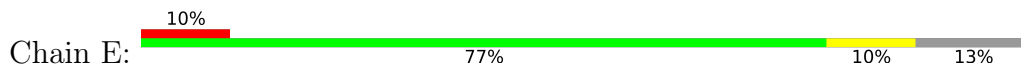
• Molecule 3: DNA-directed RNA polymerase subunit beta'



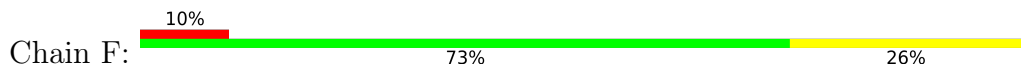


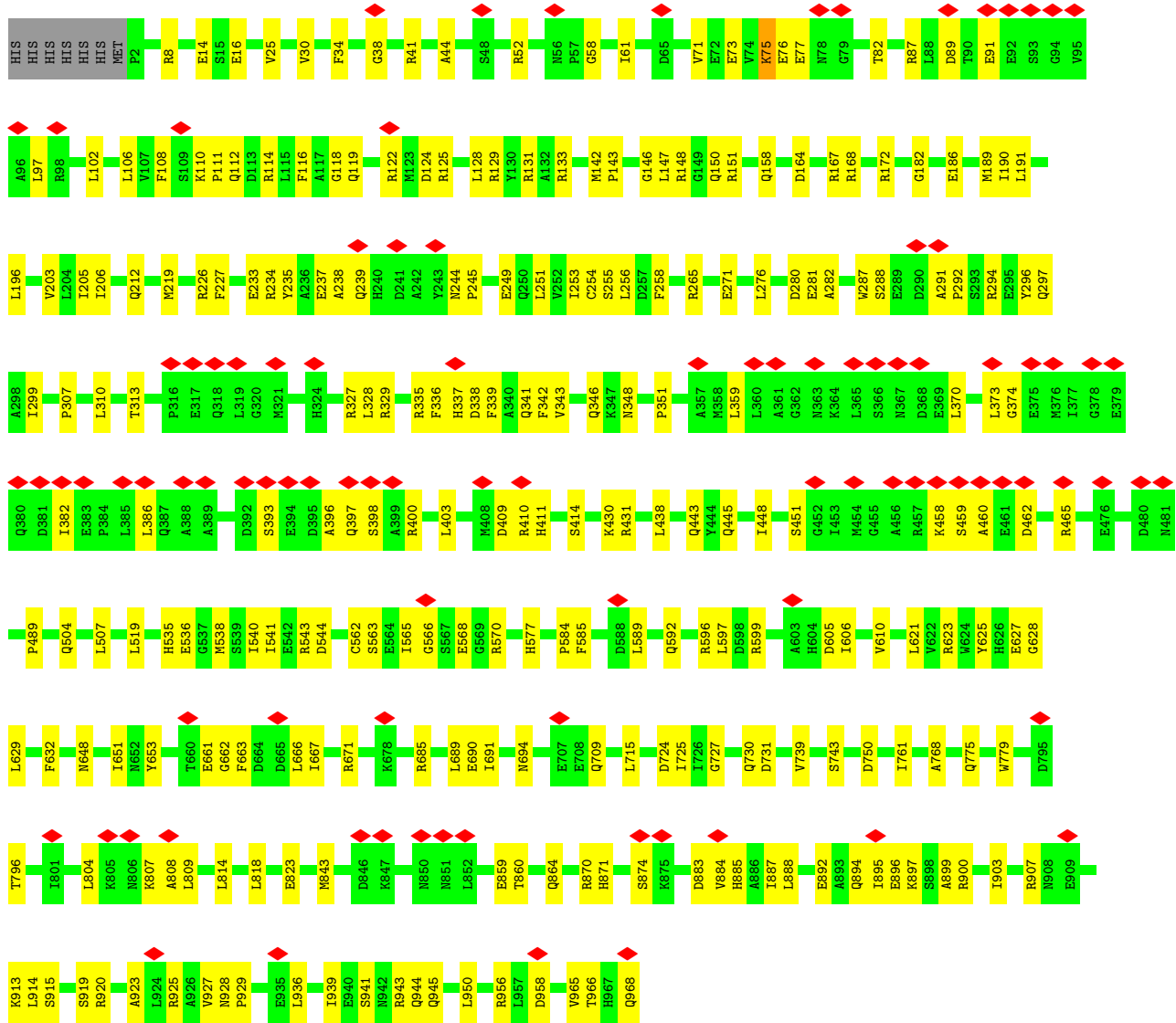


• Molecule 4: DNA-directed RNA polymerase subunit omega

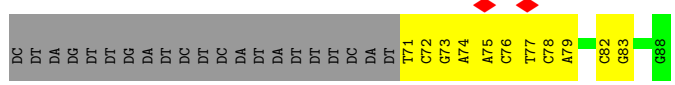


• Molecule 5: RNA polymerase-associated protein RapA

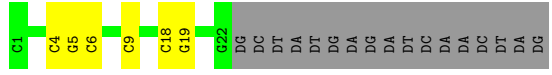




• Molecule 6: Nontemplate DNA (39-MER)



• Molecule 7: Template DNA (39-MER)



• Molecule 8: RNA (5'-R(P\*AP\*UP\*CP\*GP\*GP\*CP\*UP\*CP\*A)-3')



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	58761	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	40	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	96000	Depositor
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	2.160	Depositor
Minimum map value	-1.536	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.056	Depositor
Recommended contour level	0.22	Depositor
Map size ( $\text{\AA}$ )	341.76, 341.76, 341.76	wwPDB
Map dimensions	384, 384, 384	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	0.89000005, 0.89000005, 0.89000005	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.52	0/1808	0.64	2/2450 (0.1%)
1	B	0.44	0/1789	0.60	0/2425
2	C	0.54	0/10739	0.60	0/14489
3	D	0.49	0/10553	0.61	0/14248
4	E	0.36	0/629	0.53	0/847
5	F	0.32	0/7863	0.49	0/10650
6	1	0.68	0/415	0.85	0/638
7	2	0.85	0/504	0.89	0/776
8	3	1.08	1/212 (0.5%)	0.79	0/326
All	All	0.49	1/34512 (0.0%)	0.59	2/46849 (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	B	0	1
2	C	0	4
3	D	0	6
All	All	0	11

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	3	1	A	OP3-P	-10.79	1.48	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	28	LEU	CA-CB-CG	5.23	127.32	115.30
1	A	88	LEU	CA-CB-CG	5.08	126.98	115.30

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	192	VAL	Peptide
2	C	57	PHE	Peptide
2	C	894	GLN	Peptide
2	C	908	GLU	Peptide
2	C	912	ASP	Peptide

## 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	1786	0	1813	27	0
1	B	1767	0	1789	29	0
2	C	10570	0	10582	264	0
3	D	10396	0	10613	278	0
4	E	627	0	634	8	0
5	F	7719	0	7583	162	0
6	1	370	0	201	10	0
7	2	451	0	248	6	0
8	3	191	0	97	2	0
9	D	2	0	0	0	0
10	D	1	0	0	0	0
All	All	33880	0	33560	755	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 11.

The worst 5 of 755 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:D:186:GLN:HG3	3:D:238:ILE:HD11	1.43	1.00

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:519:ASN:HB2	3:D:709:ARG:HD3	1.50	0.93
2:C:81:ASP:OD1	2:C:82:VAL:N	2.06	0.89
3:D:857:LEU:HG	3:D:858:VAL:HG13	1.54	0.89
4:E:79:GLU:HB3	5:F:808:ALA:HB2	1.56	0.88

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	228/329 (69%)	206 (90%)	22 (10%)	0	100	100
1	B	226/329 (69%)	197 (87%)	29 (13%)	0	100	100
2	C	1338/1342 (100%)	1172 (88%)	164 (12%)	2 (0%)	51	82
3	D	1331/1416 (94%)	1176 (88%)	154 (12%)	1 (0%)	51	82
4	E	77/91 (85%)	70 (91%)	7 (9%)	0	100	100
5	F	965/974 (99%)	898 (93%)	67 (7%)	0	100	100
All	All	4165/4481 (93%)	3719 (89%)	443 (11%)	3 (0%)	54	82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	909	LYS
2	C	398	SER
3	D	121	PRO

### 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	198/286 (69%)	197 (100%)	1 (0%)	88	94
1	B	196/286 (68%)	195 (100%)	1 (0%)	88	94
2	C	1155/1157 (100%)	1155 (100%)	0	100	100
3	D	1120/1177 (95%)	1118 (100%)	2 (0%)	93	98
4	E	67/75 (89%)	67 (100%)	0	100	100
5	F	828/835 (99%)	824 (100%)	4 (0%)	88	94
All	All	3564/3816 (93%)	3556 (100%)	8 (0%)	93	98

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	925	ARG
5	F	327	ARG
5	F	75	LYS
3	D	1258	ARG
5	F	131	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	120	GLN
2	C	214	ASN
3	D	495	ASN
5	F	244	ASN
5	F	864	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	3	8/9 (88%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.



## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 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.

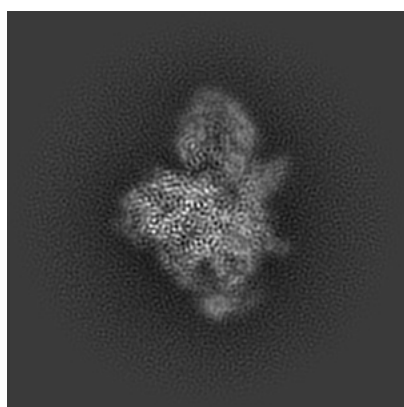
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-23716. These allow visual inspection of the internal detail of the map and identification of artifacts.

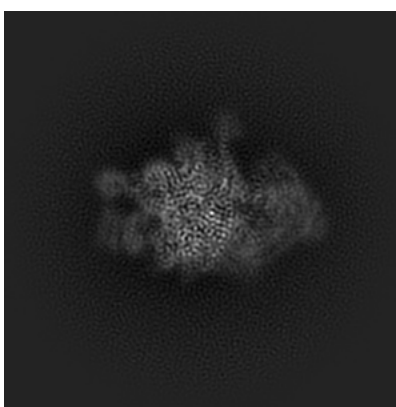
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

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

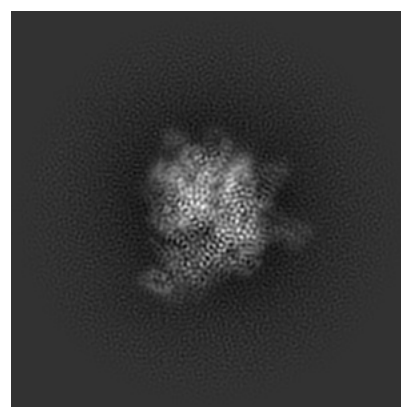
#### 6.1.1 Primary map



X



Y

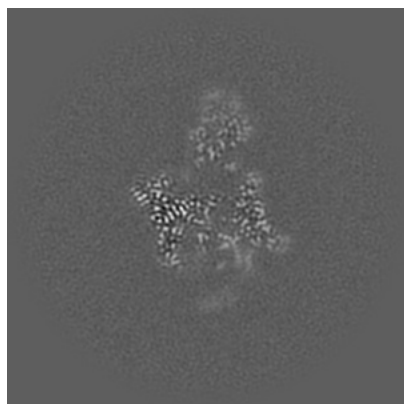


Z

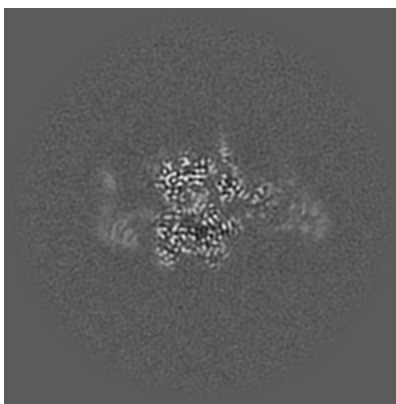
The images above show the map projected in three orthogonal directions.

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

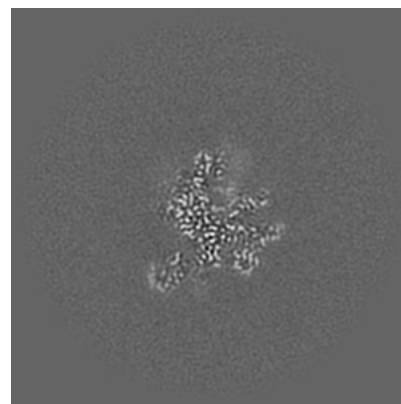
#### 6.2.1 Primary map



X Index: 192



Y Index: 192

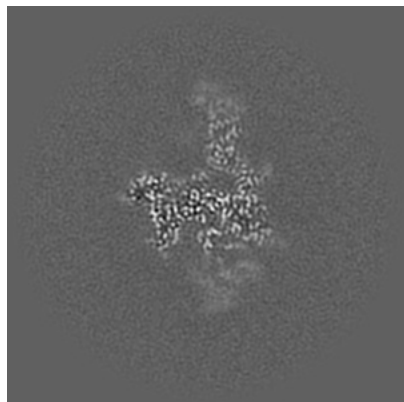


Z Index: 192

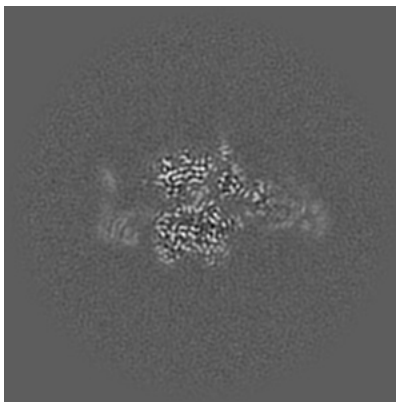
The images above show central slices of the map in three orthogonal directions.

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

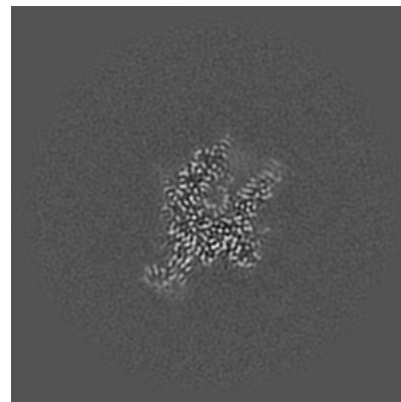
### 6.3.1 Primary map



X Index: 183



Y Index: 193

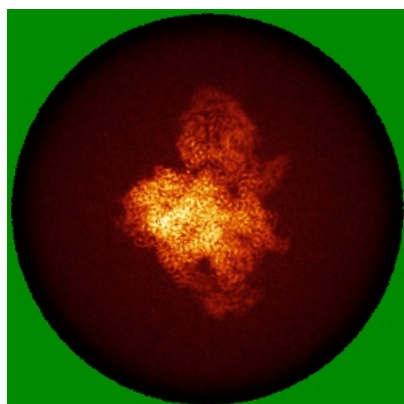


Z Index: 183

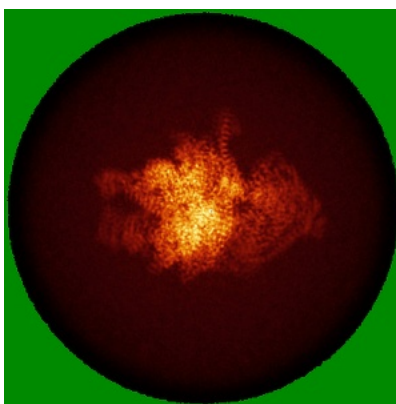
The images above show the largest variance slices of the map in three orthogonal directions.

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

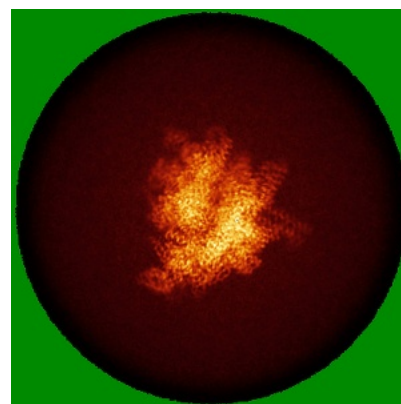
### 6.4.1 Primary map



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

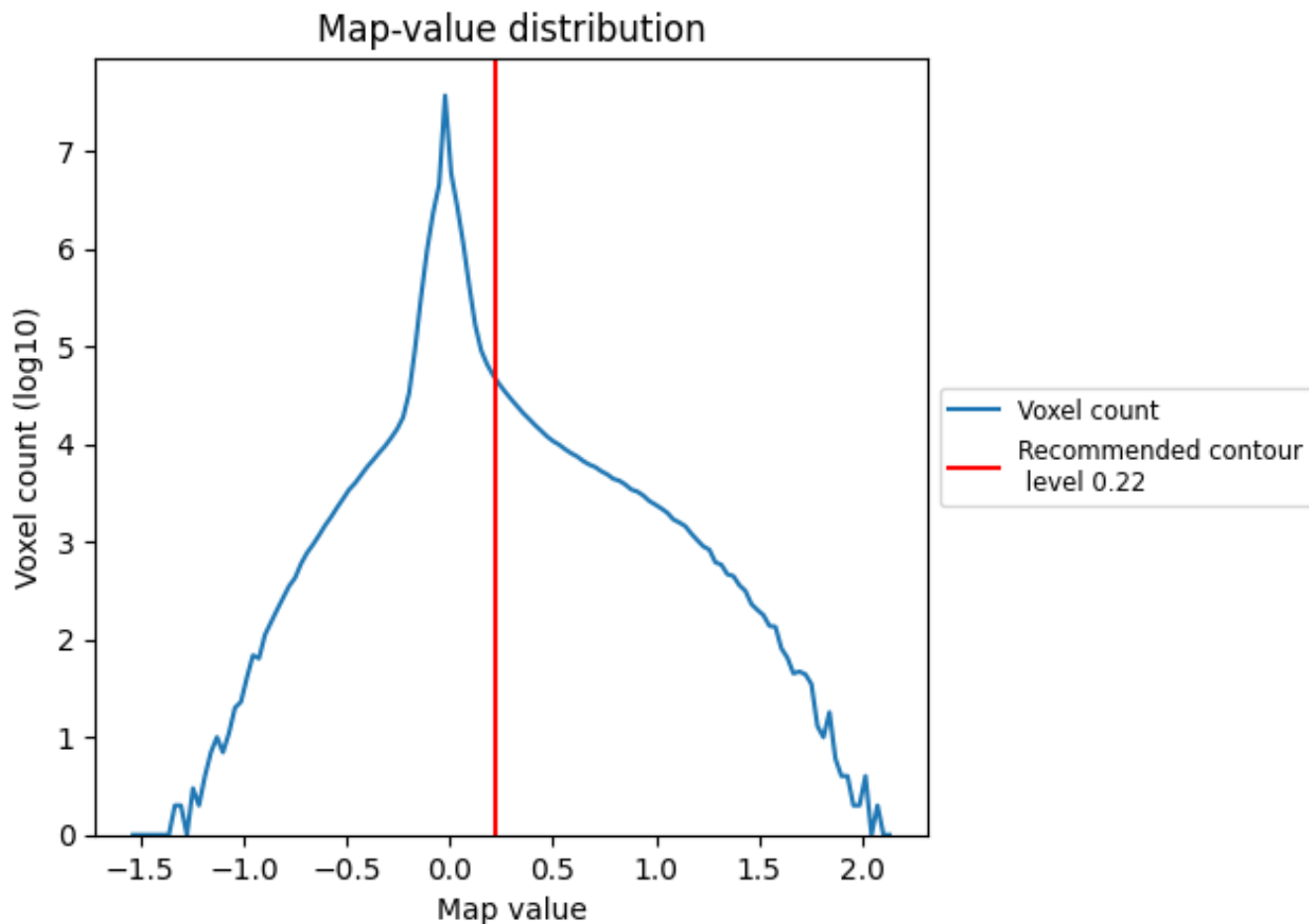
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

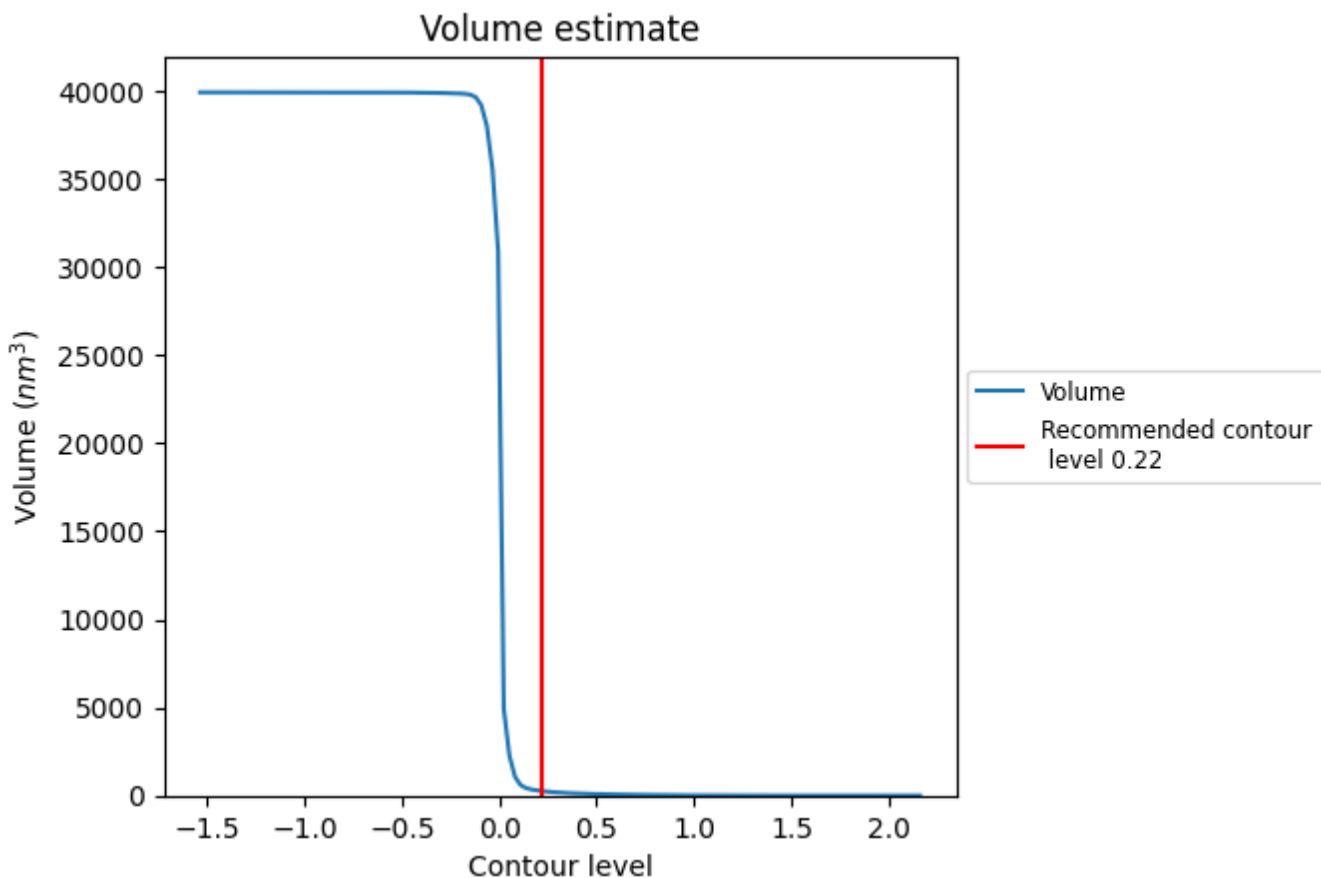
This section contains the results of statistical analysis of the map.

### 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.

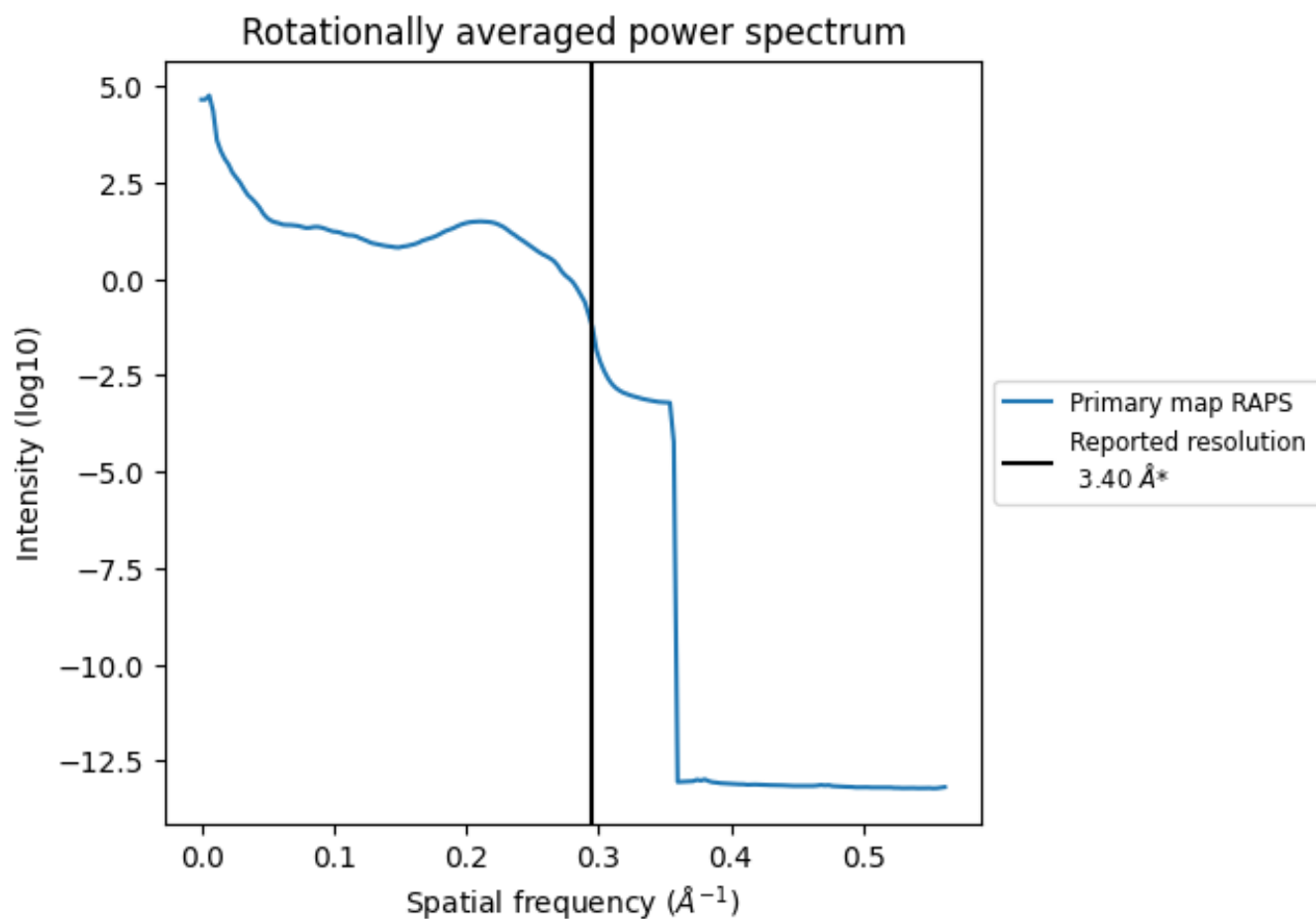
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 255 nm<sup>3</sup>; this corresponds to an approximate mass of 230 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  $0.294 \text{\AA}^{-1}$

## 8 Fourier-Shell correlation

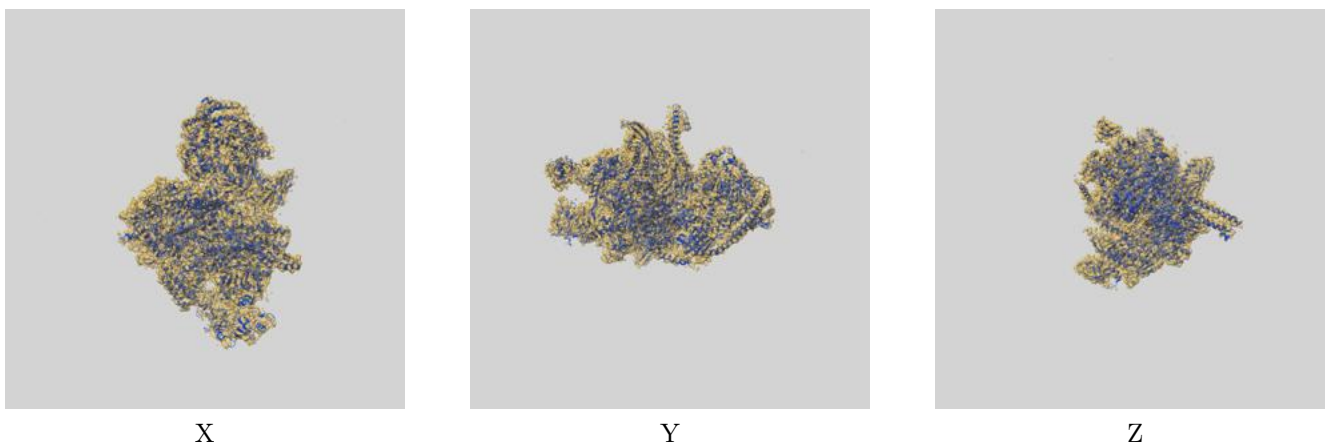
This section was not generated. No FSC curve or half-maps provided.



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

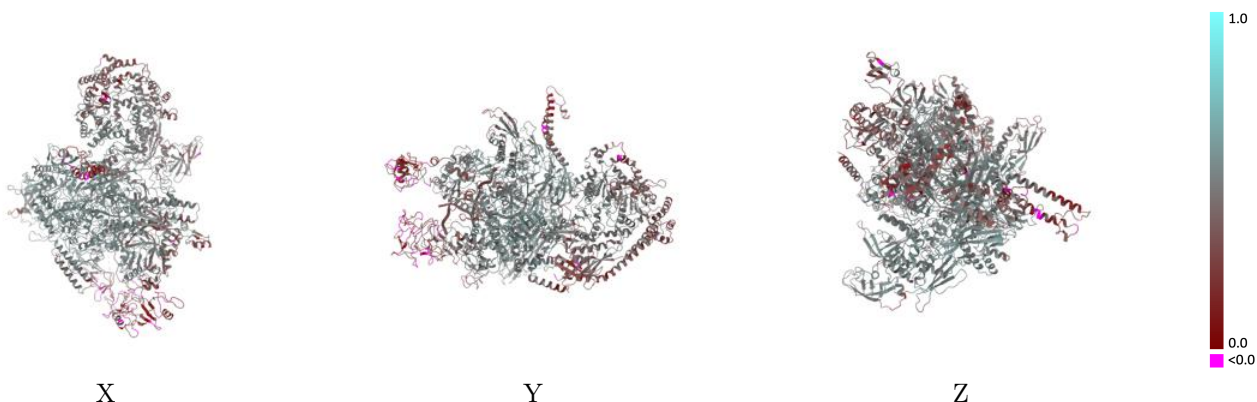
This section contains information regarding the fit between EMDB map EMD-23716 and PDB model 7M8E. Per-residue inclusion information can be found in section 3 on page 7.

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



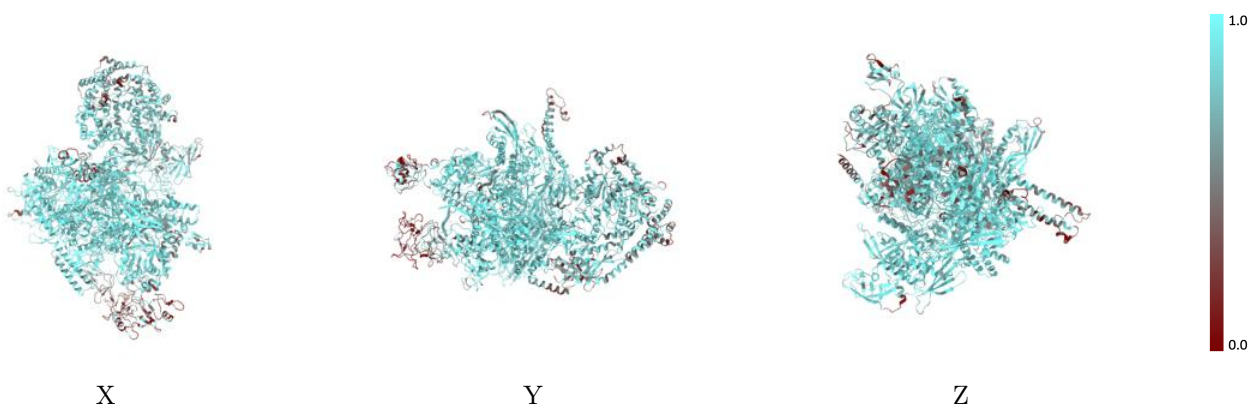
The images above show the 3D surface view of the map at the recommended contour level 0.22 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



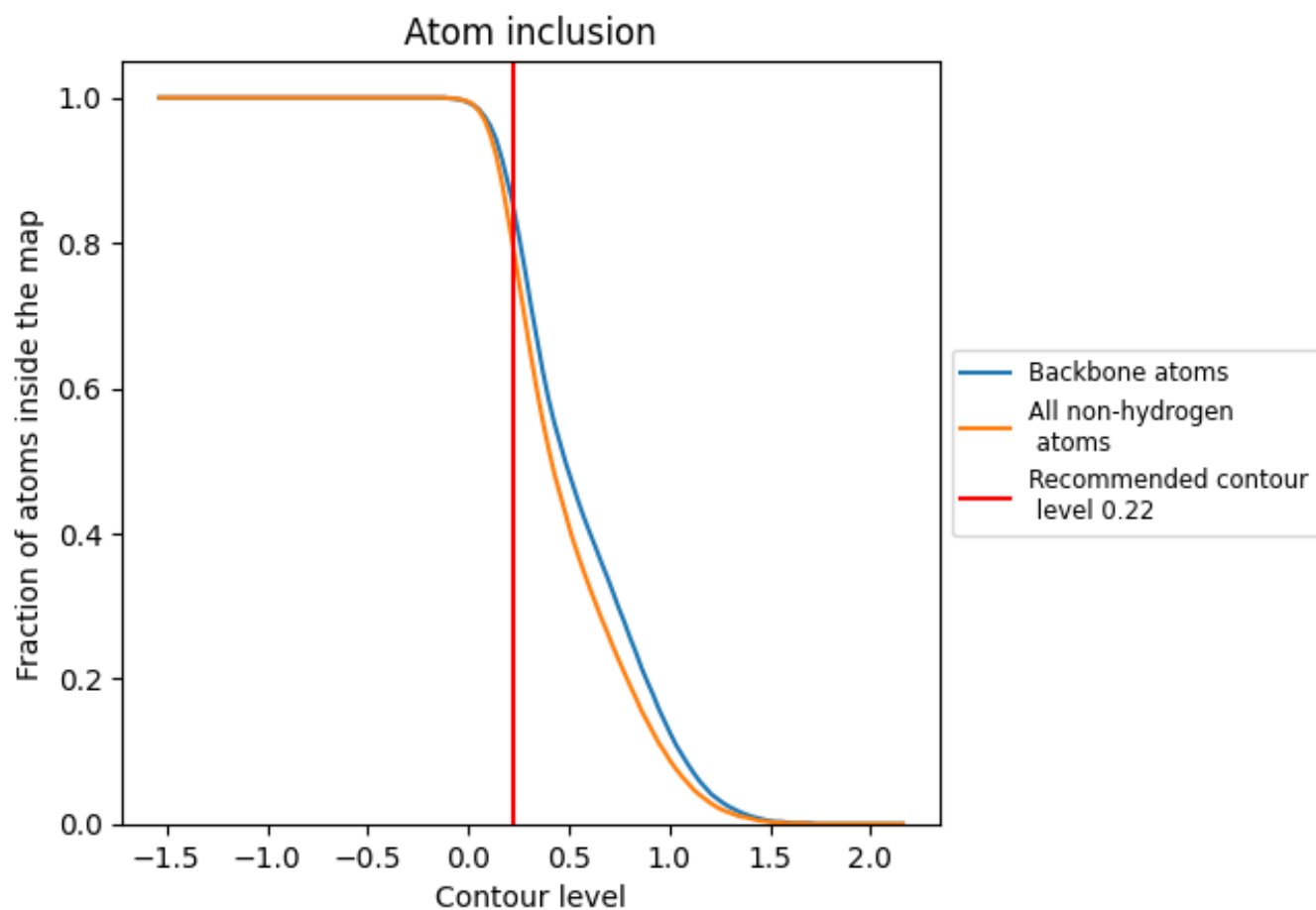
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.22).





















## 9.4 Atom inclusion [i](#)



At the recommended contour level, 86% of all backbone atoms, 80% 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 (0.22) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7990	 0.4550
1	 0.7460	 0.3610
2	 0.9180	 0.4600
3	 0.9580	 0.5090
A	 0.8910	 0.5290
B	 0.8500	 0.4900
C	 0.8280	 0.4700
D	 0.7990	 0.4530
E	 0.7220	 0.4720
F	 0.7260	 0.4150

