



## wwPDB EM Validation Summary Report ⓘ

Feb 25, 2024 – 07:40 AM EST

PDB ID : 6X43  
EMDB ID : EMD-22039  
Title : Mfd-bound E.coli RNA polymerase elongation complex - II state  
Authors : Llewellyn, E.; Chen, J.; Kang, J.Y.; Darst, S.A.  
Deposited on : 2020-05-22  
Resolution : 3.60 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](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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
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

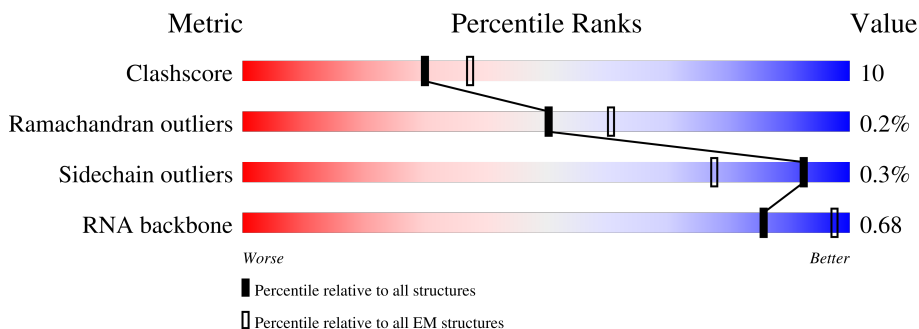
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.60 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	1148	
2	G	329	
2	H	329	
3	I	1342	
4	J	1407	
5	K	91	
6	R	21	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	P	64	 48% 33% 19%
8	Q	64	 5% 44% 30% 23%

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 36055 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 Transcription-repair-coupling factor.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1144	Total	C	N	O	S	0	0
			8982	5683	1600	1663	36		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	G	224	Total	C	N	O	S	0	0
			1725	1080	303	336	6		
2	H	217	Total	C	N	O	S	0	0
			1667	1043	293	325	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	I	1316	Total	C	N	O	S	0	0
			10354	6495	1806	2010	43		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	J	1336	Total	C	N	O	S	0	0
			10383	6524	1852	1957	50		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	1384	VAL	MET	conflict	UNP A0A4S1NBU2

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

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

- Molecule 6 is a RNA chain called RNA (20-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	R	10	221	98	45	68	10	0	0

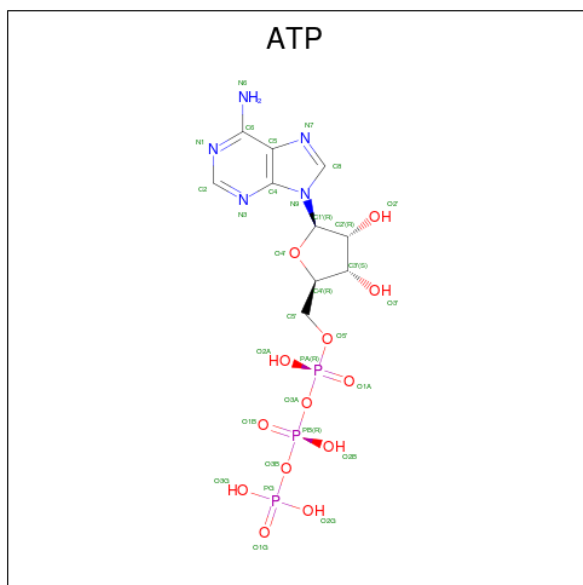
- Molecule 7 is a DNA chain called DNA (64-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	P	52	1050	502	179	318	51	0	0

- Molecule 8 is a DNA chain called DNA (64-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	Q	49	1011	478	194	290	49	0	0

- Molecule 9 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
9	A	1	31	10	5	13	3	0

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
10	A	1	Total 1	Mg 1	0
10	J	1	Total 1	Mg 1	0

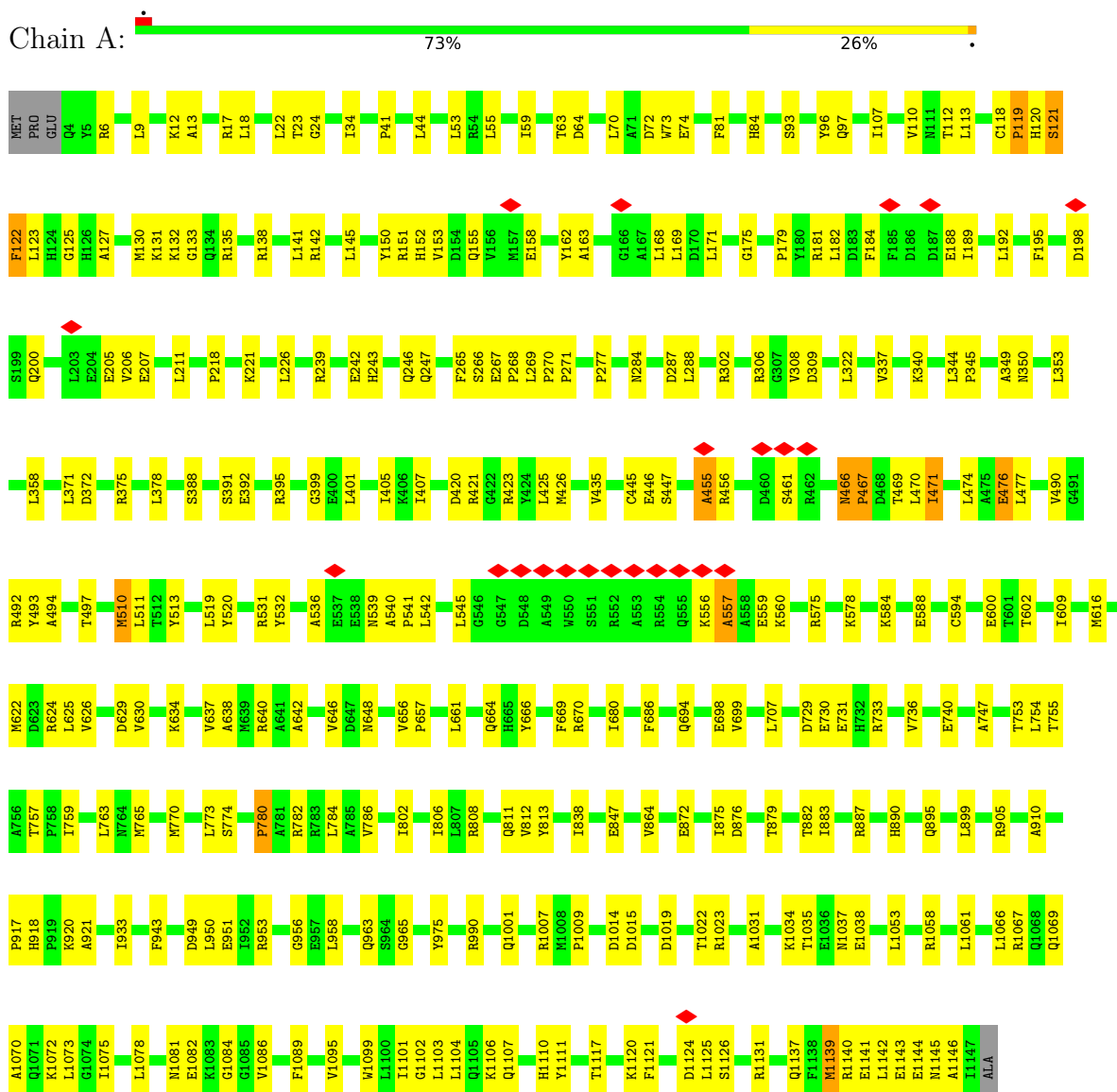
- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
11	J	2	Total 2	Zn 2	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.

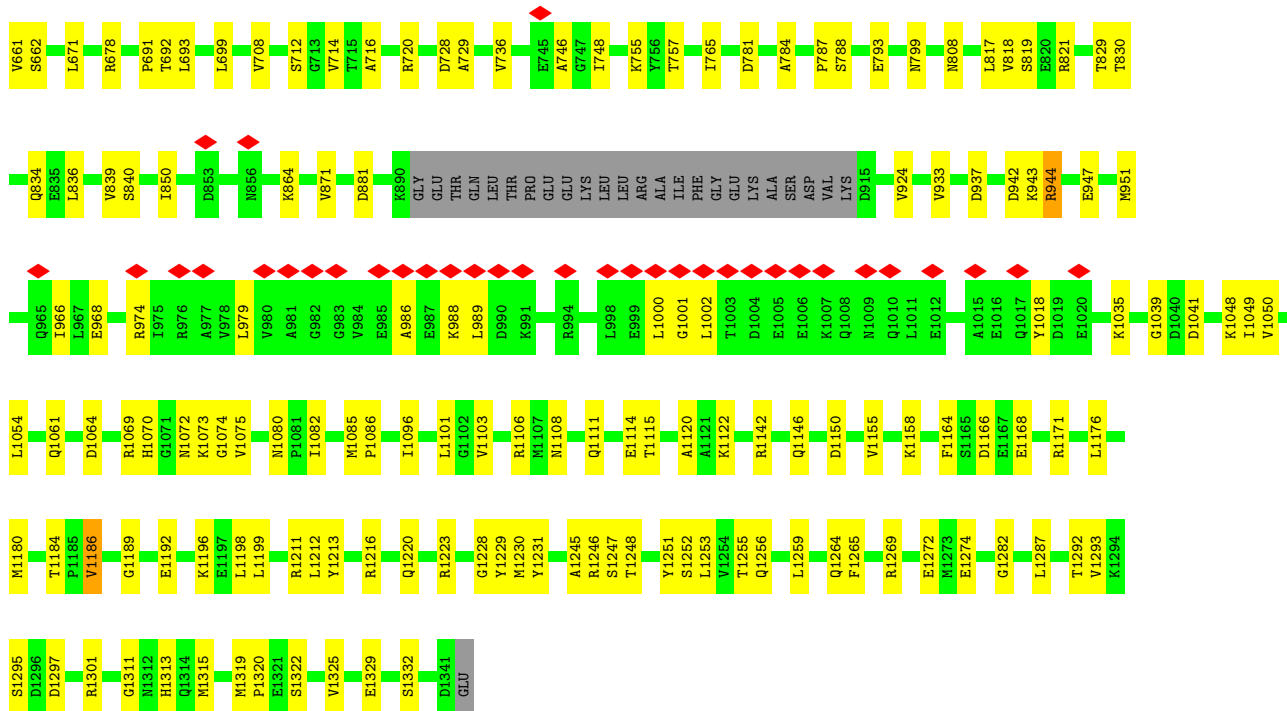
- Molecule 1: Transcription-repair-coupling factor



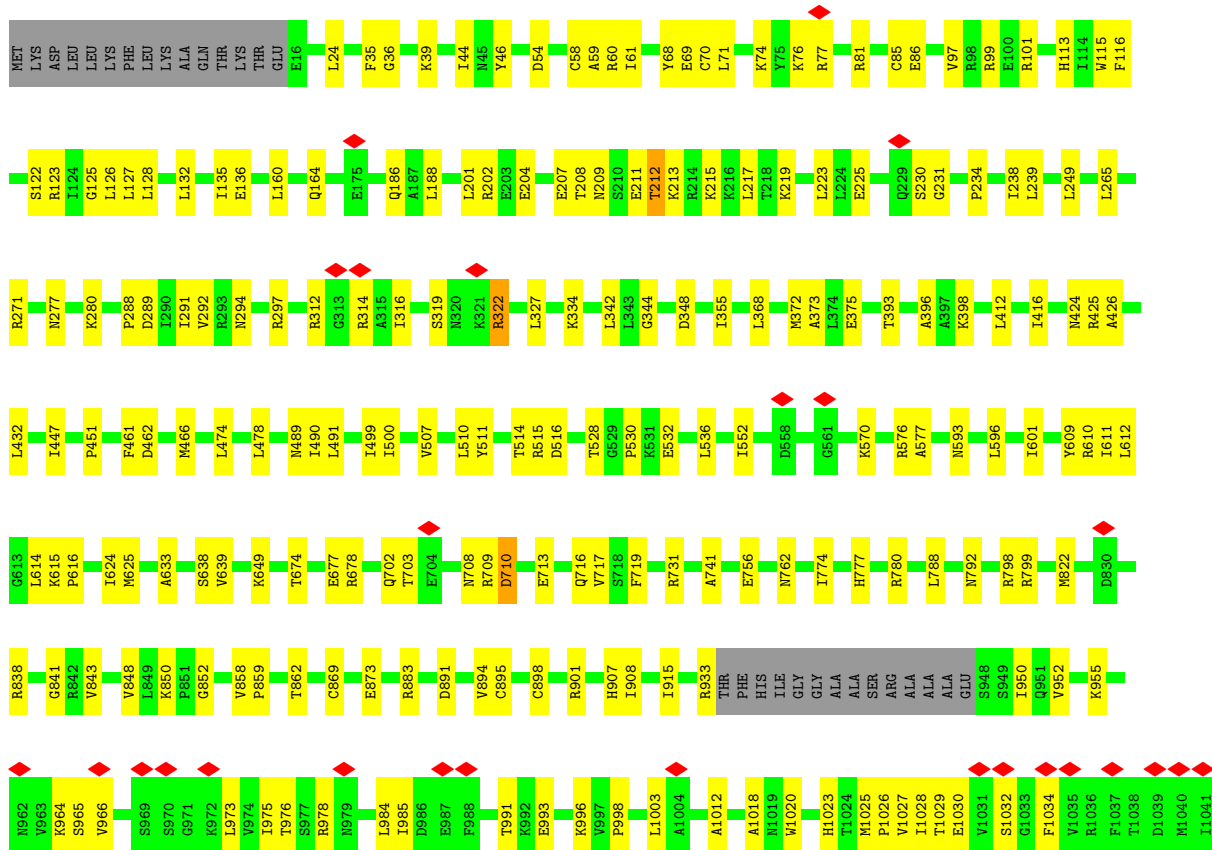
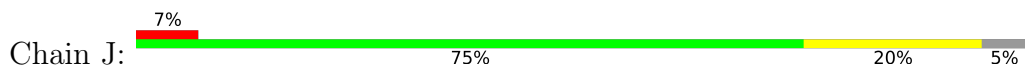
- Molecule 2: DNA-directed RNA polymerase subunit alpha

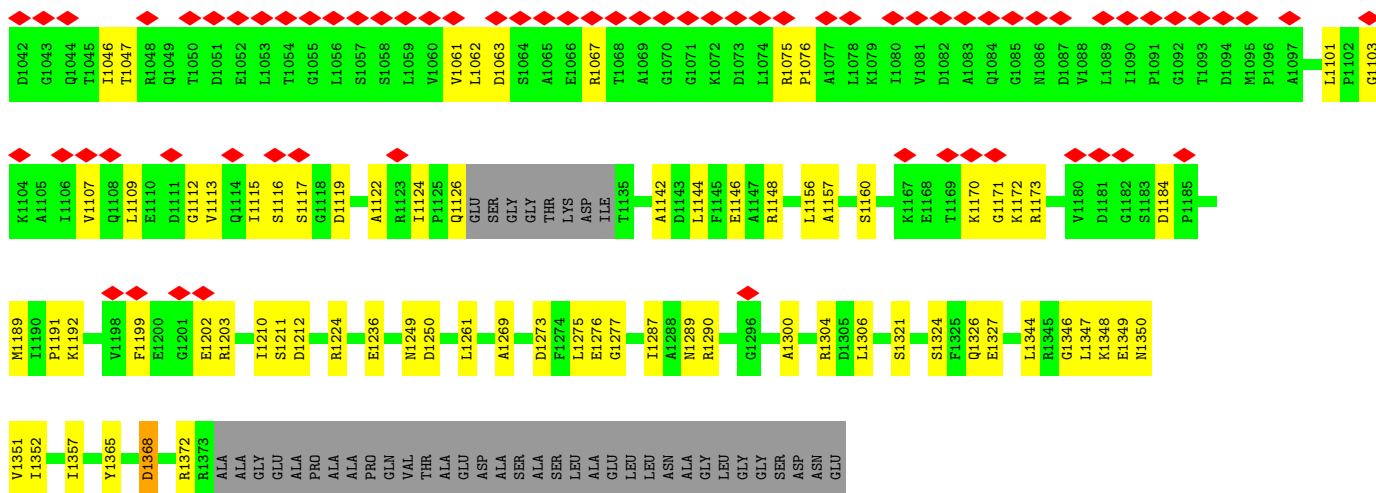






• Molecule 4: DNA-directed RNA polymerase subunit beta'

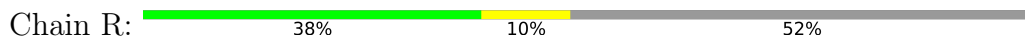




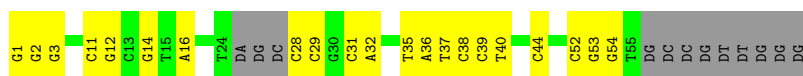
● Molecule 5: DNA-directed RNA polymerase subunit omega



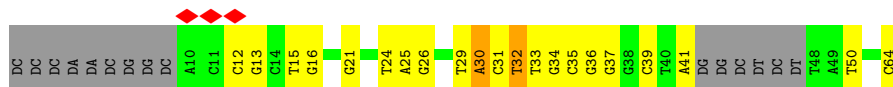
● Molecule 6: RNA (20-mer)



● Molecule 7: DNA (64-MER)



● Molecule 8: DNA (64-MER)



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	86000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	60	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	50.437	Depositor
Minimum map value	-22.539	Depositor
Average map value	0.006	Depositor
Map value standard deviation	1.083	Depositor
Recommended contour level	4.0	Depositor
Map size (Å)	390.0, 390.0, 390.0	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.3, 1.3, 1.3	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: MG, ZN, ATP

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	1/9161 (0.0%)	0.64	3/12421 (0.0%)
2	G	0.45	0/1745	0.60	1/2366 (0.0%)
2	H	0.39	0/1686	0.58	0/2286
3	I	0.51	0/10517	0.57	0/14193
4	J	0.47	0/10540	0.58	1/14231 (0.0%)
5	K	0.41	0/629	0.58	0/847
6	R	1.11	0/248	1.00	0/386
7	P	1.10	0/1171	1.03	0/1801
8	Q	0.92	1/1135 (0.1%)	0.92	1/1748 (0.1%)
All	All	0.54	2/36832 (0.0%)	0.63	6/50279 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	Q	30	DA	N9-C4	-5.19	1.34	1.37
1	A	81	PHE	C-N	-5.17	1.22	1.34

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	1368	ASP	CB-CG-OD1	6.02	123.72	118.30
1	A	474	LEU	N-CA-C	-5.70	95.62	111.00
8	Q	32	DT	N3-C4-O4	5.20	123.02	119.90
2	G	233	ASP	CB-CG-OD1	5.17	122.96	118.30
1	A	1139	MET	CG-SD-CE	-5.12	92.01	100.20

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	8982	0	8899	238	0
2	G	1725	0	1763	32	0
2	H	1667	0	1699	52	0
3	I	10354	0	10358	191	0
4	J	10383	0	10598	198	0
5	K	627	0	634	13	0
6	R	221	0	109	2	0
7	P	1050	0	590	30	0
8	Q	1011	0	550	21	0
9	A	31	0	12	2	0
10	A	1	0	0	0	0
10	J	1	0	0	0	0
11	J	2	0	0	0	0
All	All	36055	0	35212	712	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 10.

The worst 5 of 712 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:A:119:PRO:HB3	1:A:270:PRO:CD	1.43	1.49
1:A:119:PRO:CB	1:A:270:PRO:HD3	1.65	1.26
1:A:476:GLU:OE1	1:A:477:LEU:HD12	1.39	1.18
2:H:14:VAL:CG1	2:H:28:LEU:HG	1.76	1.16
4:J:211:GLU:CB	7:P:3:DG:OP2	2.01	1.06

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	1142/1148 (100%)	1006 (88%)	127 (11%)	9 (1%)	19	59
2	G	220/329 (67%)	198 (90%)	22 (10%)	0	100	100
2	H	213/329 (65%)	194 (91%)	19 (9%)	0	100	100
3	I	1312/1342 (98%)	1197 (91%)	115 (9%)	0	100	100
4	J	1330/1407 (94%)	1231 (93%)	98 (7%)	1 (0%)	51	83
5	K	77/91 (85%)	68 (88%)	9 (12%)	0	100	100
All	All	4294/4646 (92%)	3894 (91%)	390 (9%)	10 (0%)	50	79

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	121	SER
1	A	467	PRO
1	A	557	ALA
4	J	212	THR
1	A	345	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	937/973 (96%)	934 (100%)	3 (0%)	92	97
2	G	191/286 (67%)	190 (100%)	1 (0%)	88	95
2	H	184/286 (64%)	184 (100%)	0	100	100
3	I	1129/1157 (98%)	1126 (100%)	3 (0%)	92	97
4	J	1118/1168 (96%)	1114 (100%)	4 (0%)	91	97
5	K	67/75 (89%)	67 (100%)	0	100	100
All	All	3626/3945 (92%)	3615 (100%)	11 (0%)	92	97

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	J	77	ARG
4	J	322	ARG
4	J	792	ASN
4	J	710	ASP
3	I	352	ARG

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

Mol	Chain	Res	Type
1	A	466	ASN
1	A	664	GLN
1	A	963	GLN
4	J	209	ASN
4	J	424	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	R	9/21 (42%)	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 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	ATP	A	2000	10	26,33,33	0.96	0	31,52,52	1.74	5 (16%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	ATP	A	2000	10	-	1/18/38/38	0/3/3/3

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	2000	ATP	PA-O3A-PB	-4.89	116.04	132.83
9	A	2000	ATP	PB-O3B-PG	-4.25	118.23	132.83
9	A	2000	ATP	N3-C2-N1	-2.90	124.14	128.68
9	A	2000	ATP	C4-C5-N7	-2.68	106.60	109.40
9	A	2000	ATP	C3'-C2'-C1'	2.22	104.32	100.98

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	2000	ATP	O4'-C4'-C5'-O5'

There are no ring outliers.

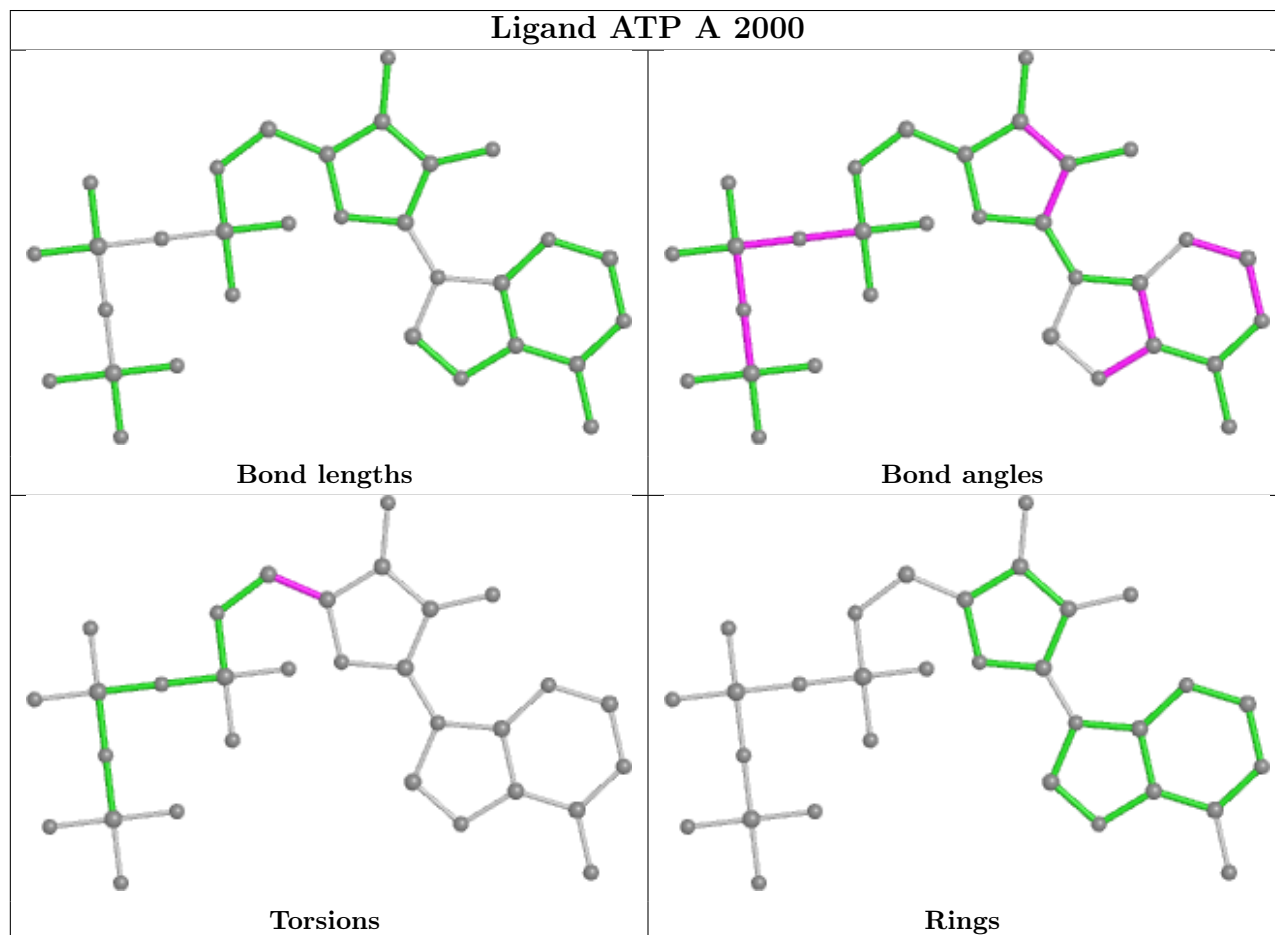
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	2000	ATP	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will



also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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.

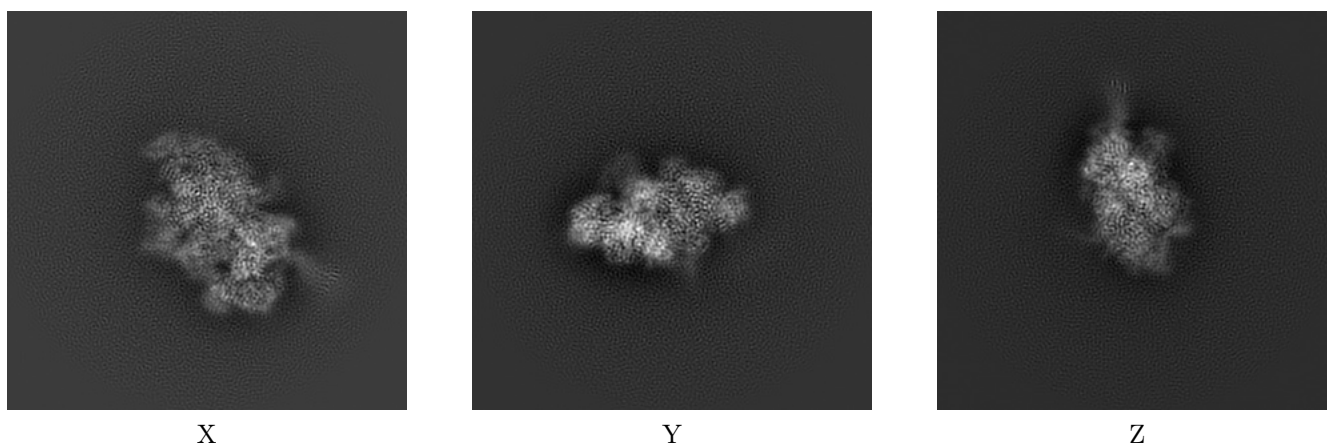
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22039. 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](#)

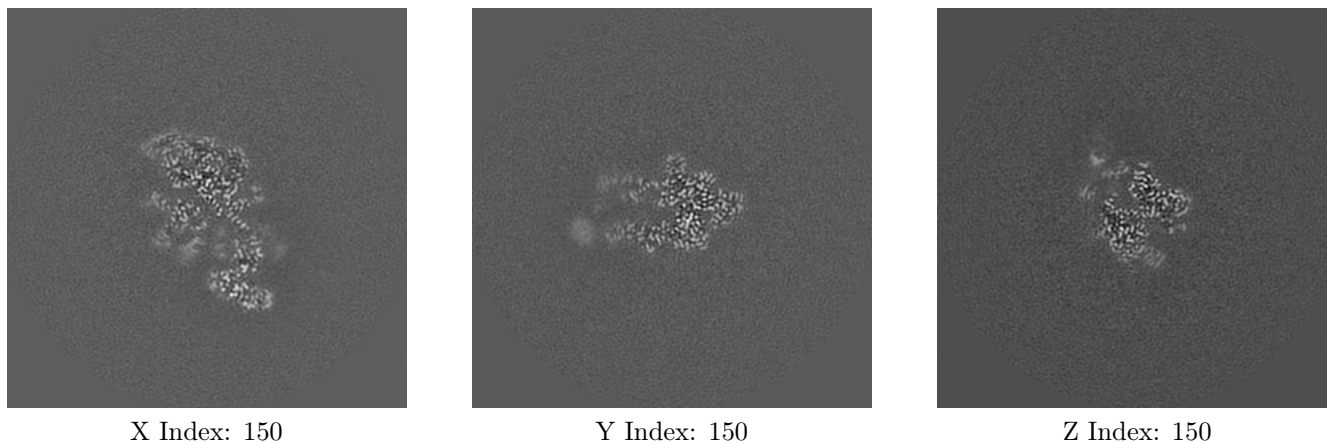
#### 6.1.1 Primary map



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

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

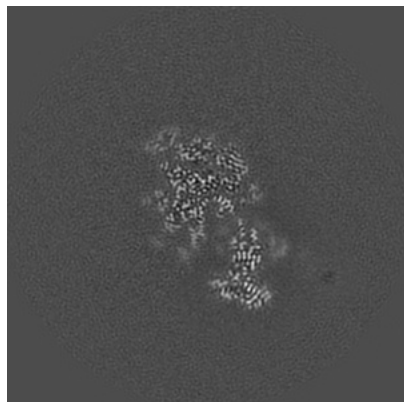
#### 6.2.1 Primary map



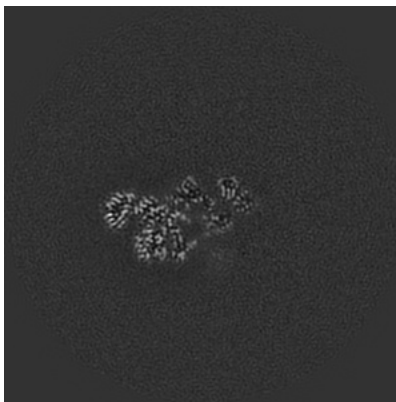
The images above show central slices of the map in three orthogonal directions.

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

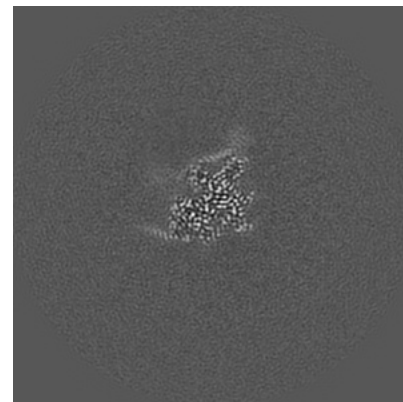
### 6.3.1 Primary map



X Index: 146



Y Index: 178

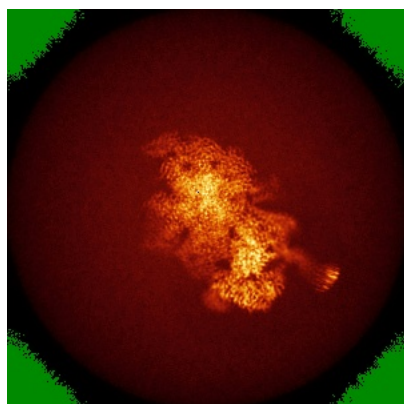


Z Index: 165

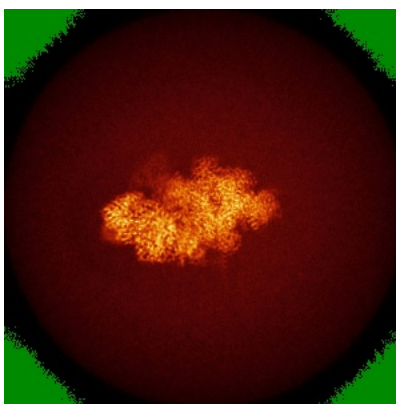
The images above show the largest variance slices of the map in three orthogonal directions.

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

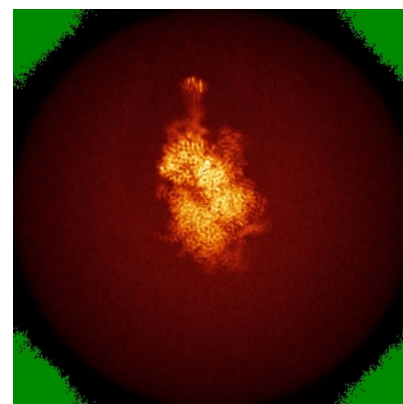
### 6.4.1 Primary map



X



Y

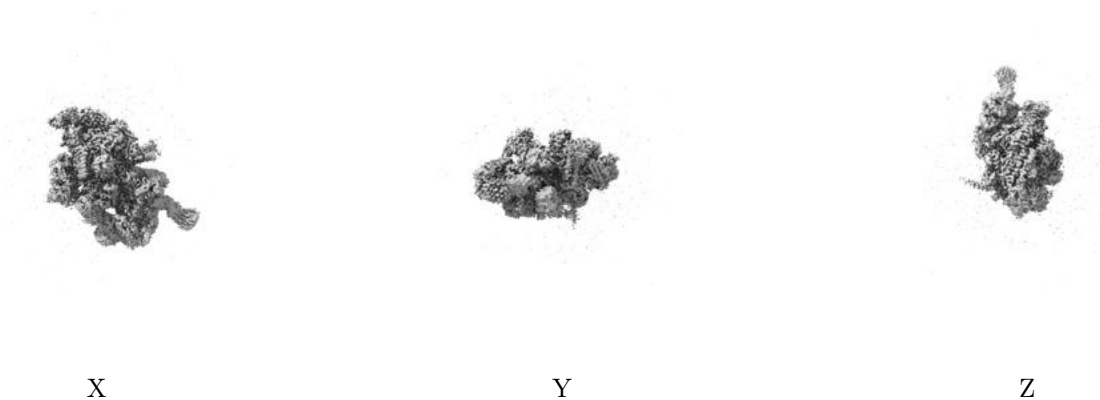


Z

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

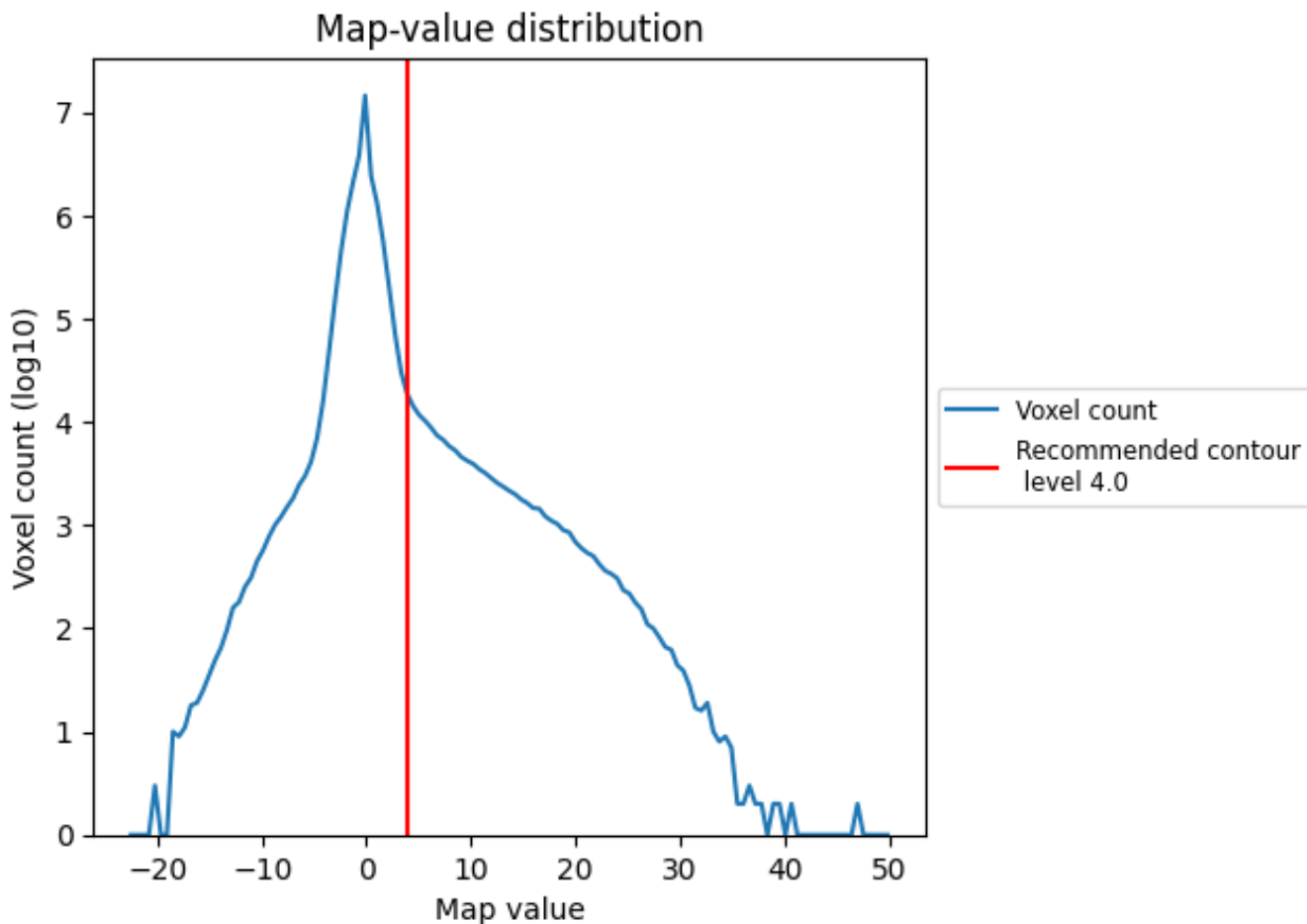
## 6.6 Mask visualisation [i](#)

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

## 7 Map analysis [i](#)

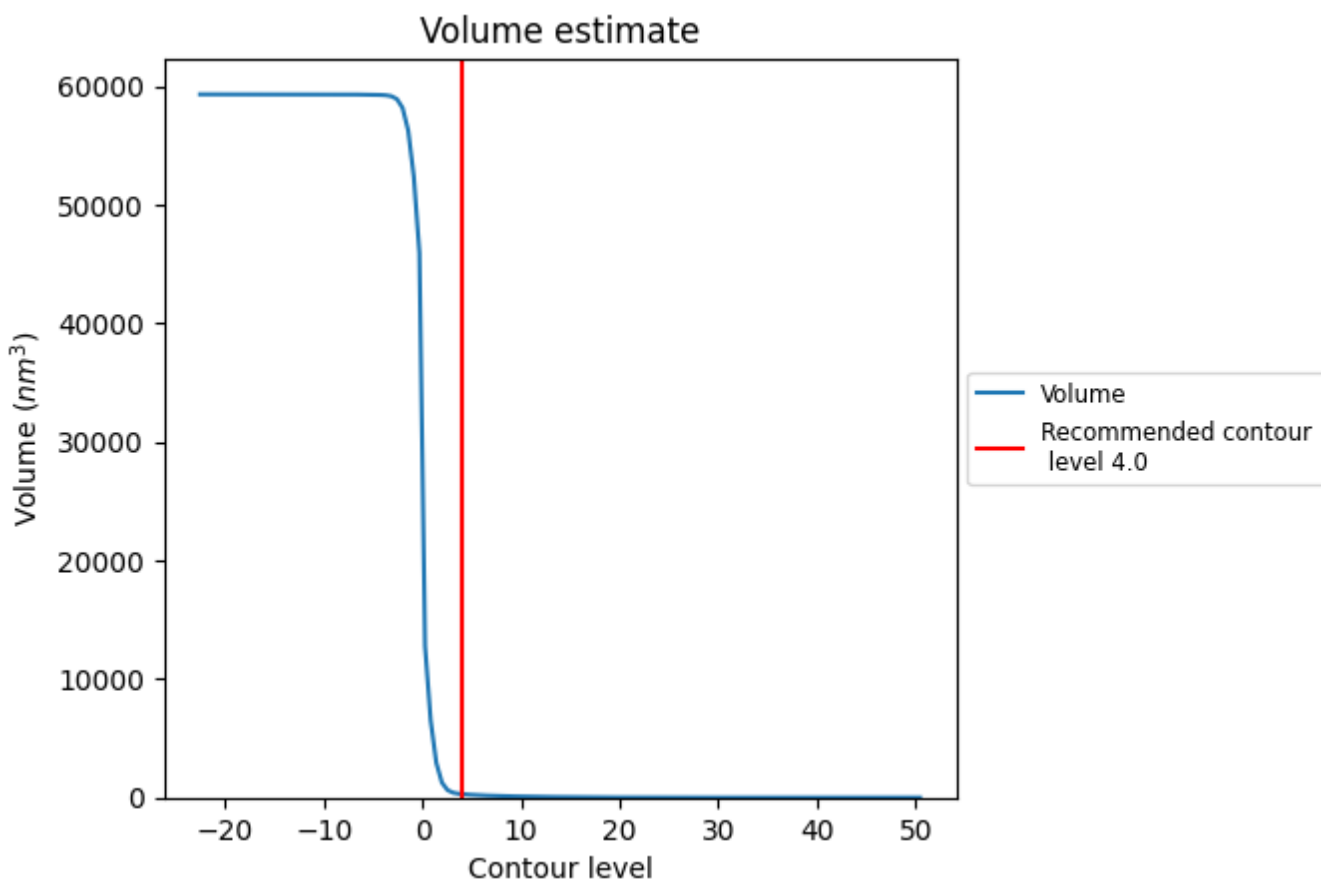
This section contains the results of statistical analysis of the map.

### 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.

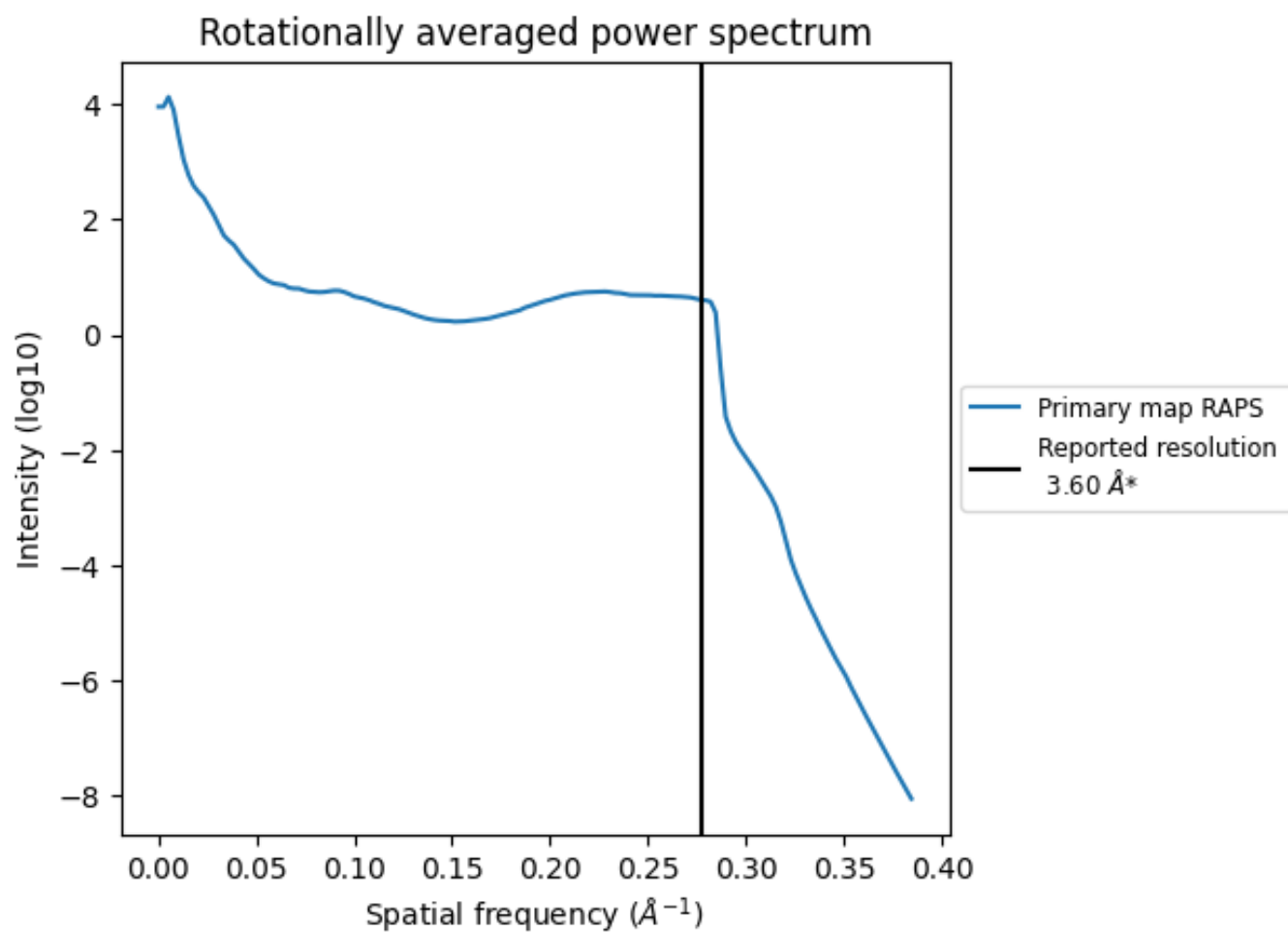
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 295 nm<sup>3</sup>; this corresponds to an approximate mass of 266 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  $0.278 \text{\AA}^{-1}$

## 8 Fourier-Shell correlation

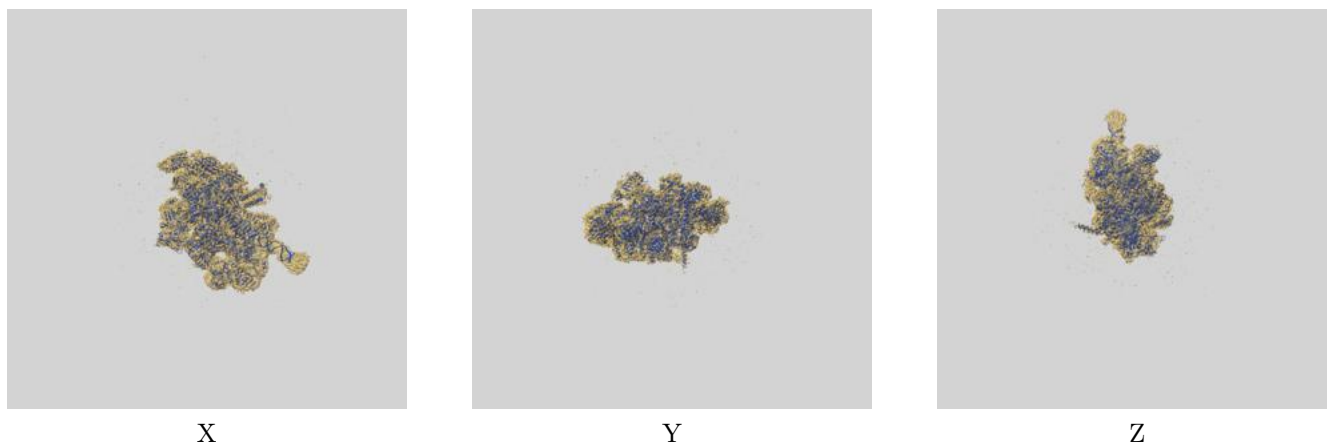
This section was not generated. No FSC curve or half-maps provided.



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

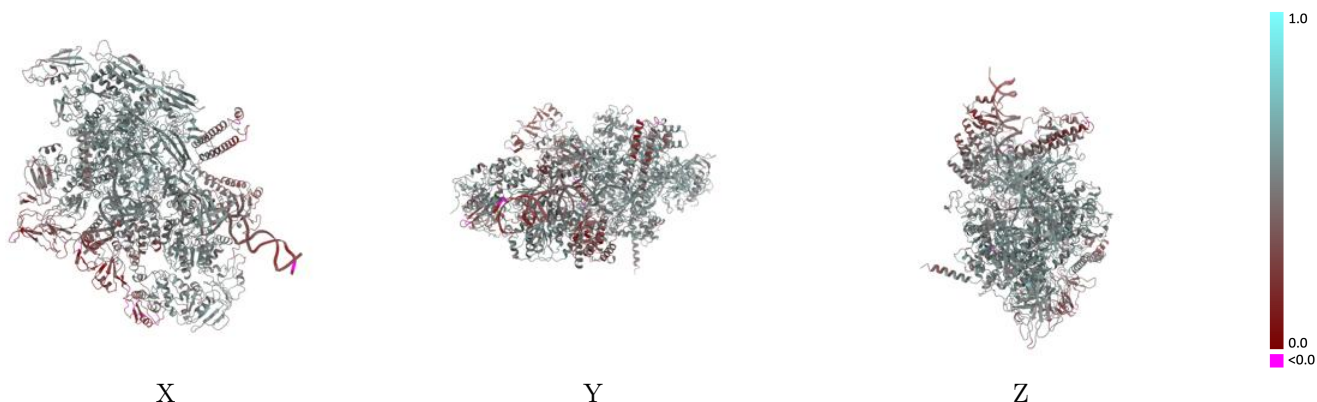
This section contains information regarding the fit between EMDB map EMD-22039 and PDB model 6X43. Per-residue inclusion information can be found in section 3 on page 7.

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



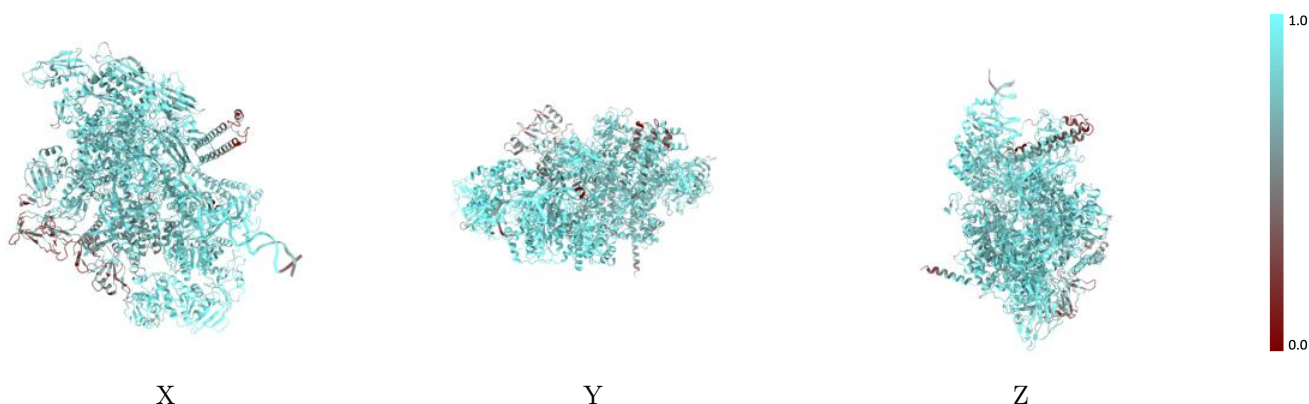
The images above show the 3D surface view of the map at the recommended contour level 4.0 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



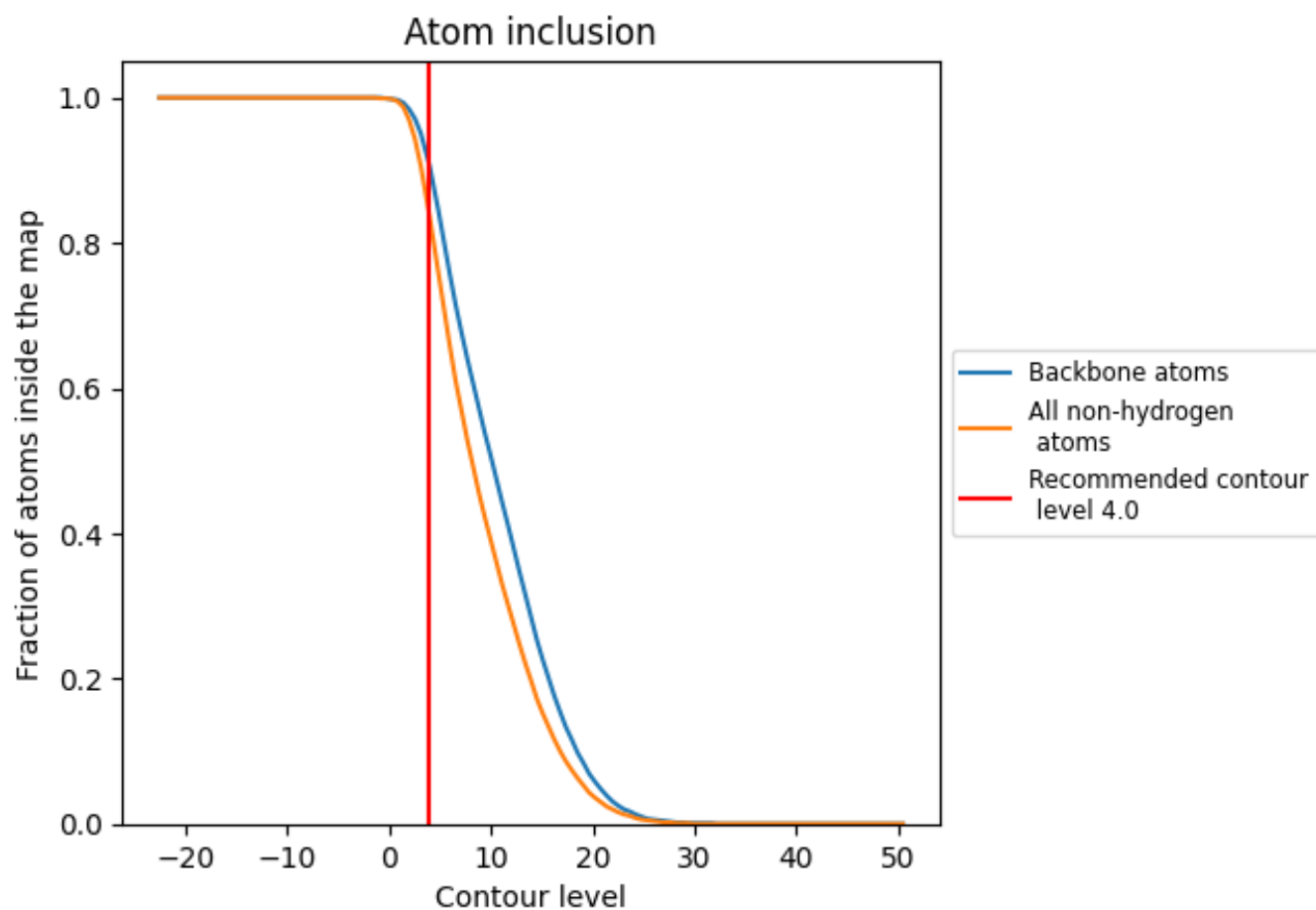
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (4.0).





















## 9.4 Atom inclusion [i](#)



At the recommended contour level, 90% of all backbone atoms, 84% 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 (4.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8350	 0.4760
A	 0.9130	 0.4480
G	 0.8460	 0.5240
H	 0.8140	 0.4930
I	 0.8080	 0.4920
J	 0.7850	 0.4870
K	 0.7660	 0.4950
P	 0.8980	 0.3980
Q	 0.8920	 0.3880
R	 0.9410	 0.5030

