wwPDB EM Validation Summary Report

Nov 20, 2022 – 08:20 am GMT

PDB ID	:	6HLS
EMDB ID	:	EMD-0241
Title	:	Yeast apo RNA polymerase I*
Authors	:	Tafur, L.; Sadian, Y.; Weis, F.; Muller, C.W.
Deposited on	:	2018-09-11
Resolution	:	3.21 Å(reported)
Based on initial models	:	4C3I, 4C2M

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) were used in the production of this report:

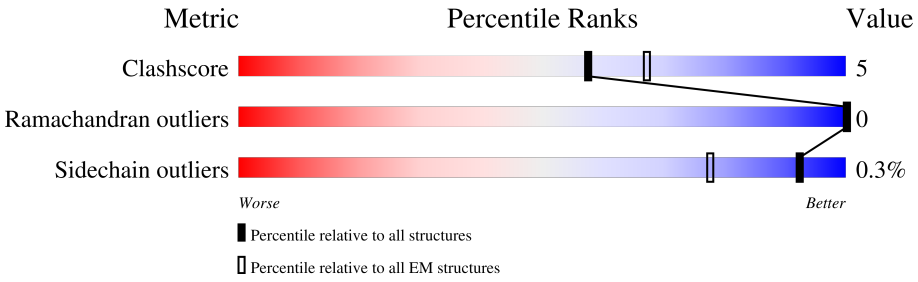
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.21 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.



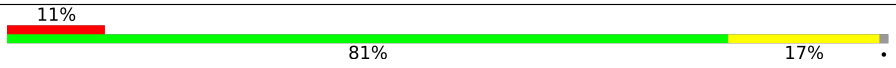
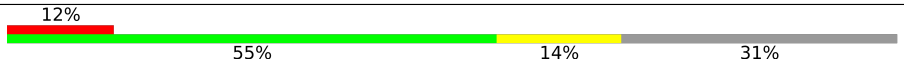
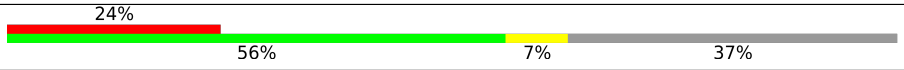
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 >=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 <=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	1664	18% 74% 12% 14%
2	B	1203	15% 80% 16% ·
3	C	335	23% 78% 13% 9%
4	D	137	32% 37% 6% 57%
5	E	215	38% 89% 11%
6	F	155	10% 56% 8% 35%
7	G	326	33% 47% 7% 46%
8	H	146	21% 82% 8% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	
10	J	70	
11	K	142	
12	L	70	

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 30954 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 I subunit RPA190.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1431	11308	7147	1964	2136	61	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase I subunit RPA135.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1163	9248	5849	1625	1724	50	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerases I and III subunit RPAC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	304	2415	1535	414	458	8	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase I subunit RPA14.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	D	59	467	293	80	94	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	214	1751	1111	309	320	11	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	F	100	823	522	144	154	3	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase I subunit RPA43.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	176	Total	C	N	O	S	0	0
			1386	897	232	252	5		

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	131	Total	C	N	O	S	0	0
			1052	664	176	208	4		

- Molecule 9 is a protein called DNA-directed RNA polymerase I subunit RPA12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	107	Total	C	N	O	S	0	0
			810	502	137	162	9		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	69	Total	C	N	O	S	0	0
			569	362	101	100	6		

- Molecule 11 is a protein called DNA-directed RNA polymerases I and III subunit RPAC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	98	Total	C	N	O	S	0	0
			766	481	124	156	5		

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	44	Total	C	N	O	S	0	0
			352	217	70	61	4		

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
13	A	2	Total	Zn	0
			2	2	
13	B	1	Total	Zn	0
			1	1	

Continued on next page...

Continued from previous page...

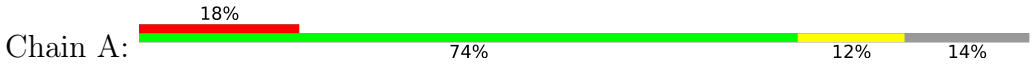
Mol	Chain	Residues	Atoms	AltConf
13	I	2	Total Zn 2 2	0
13	J	1	Total Zn 1 1	0
13	L	1	Total Zn 1 1	0

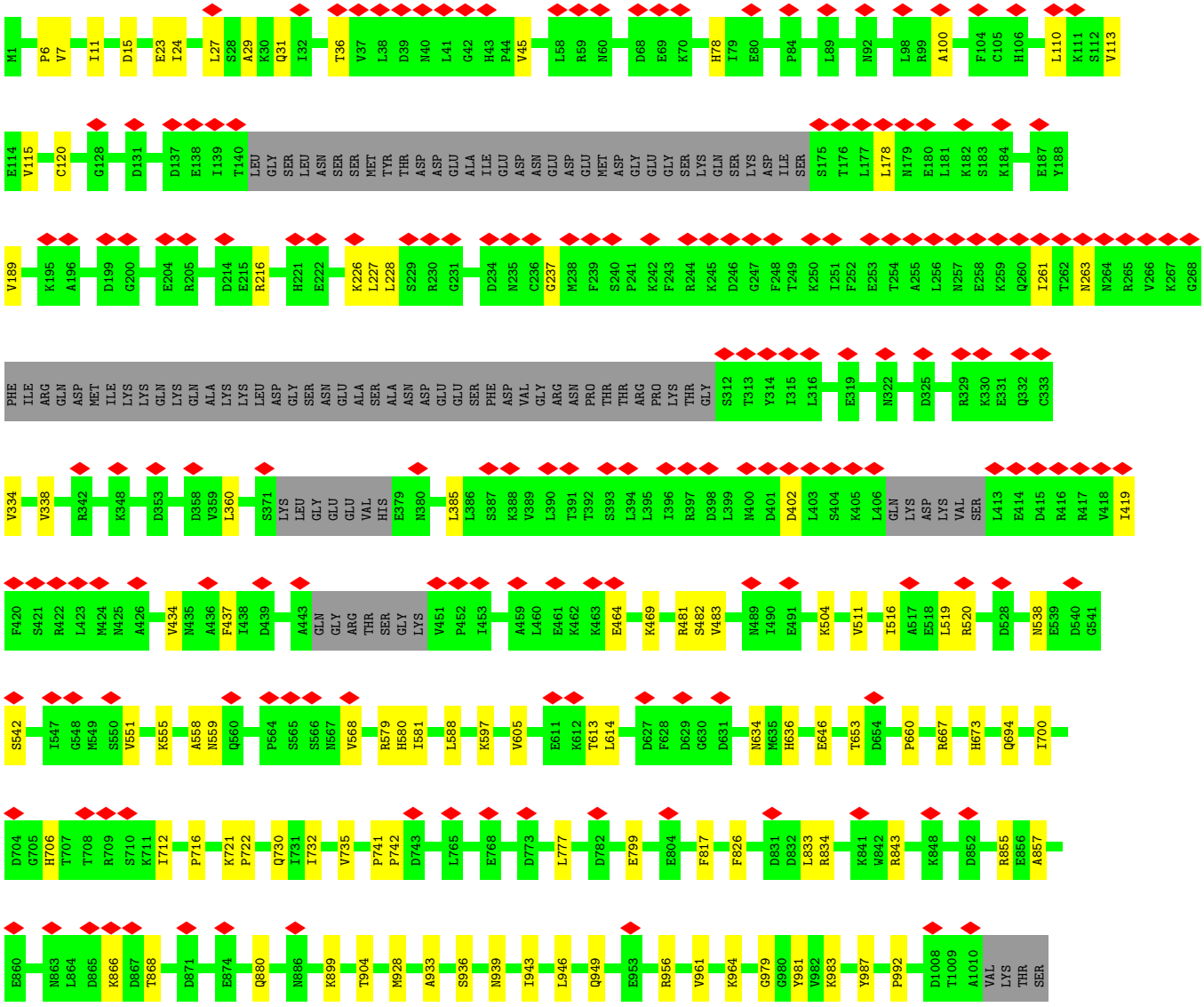
3 Residue-property plots (i)

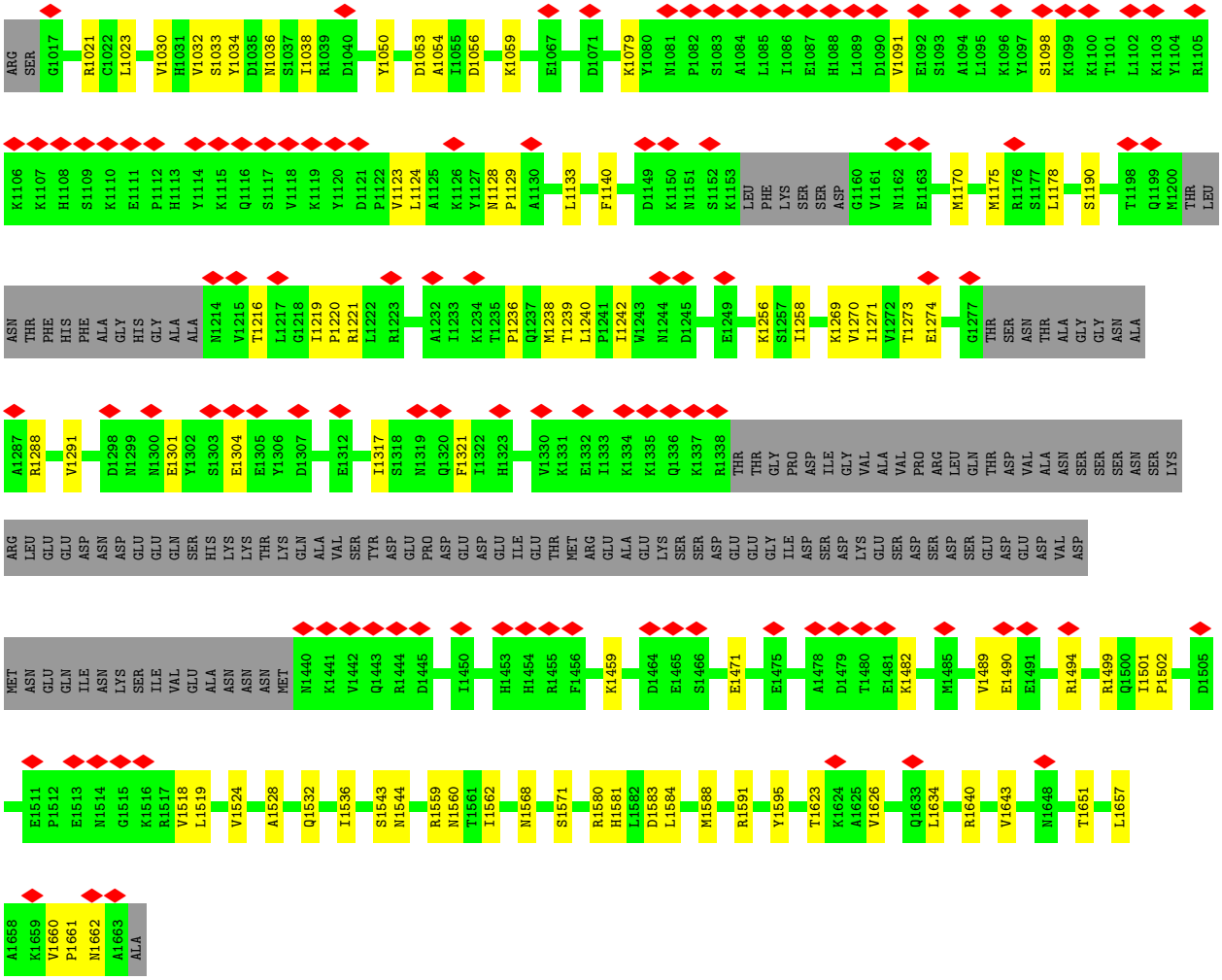
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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 I subunit RPA190

Chain A:

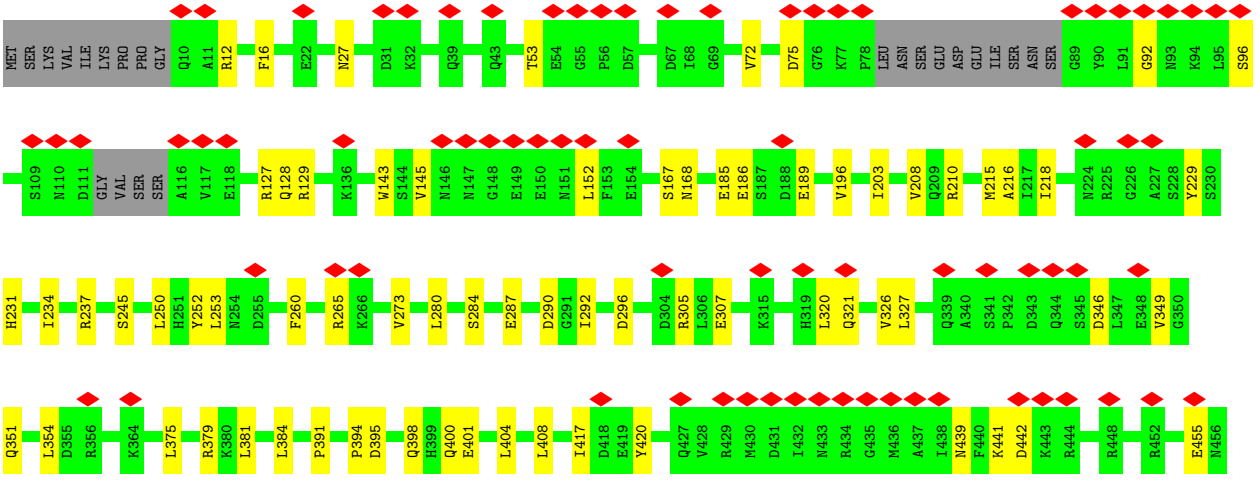


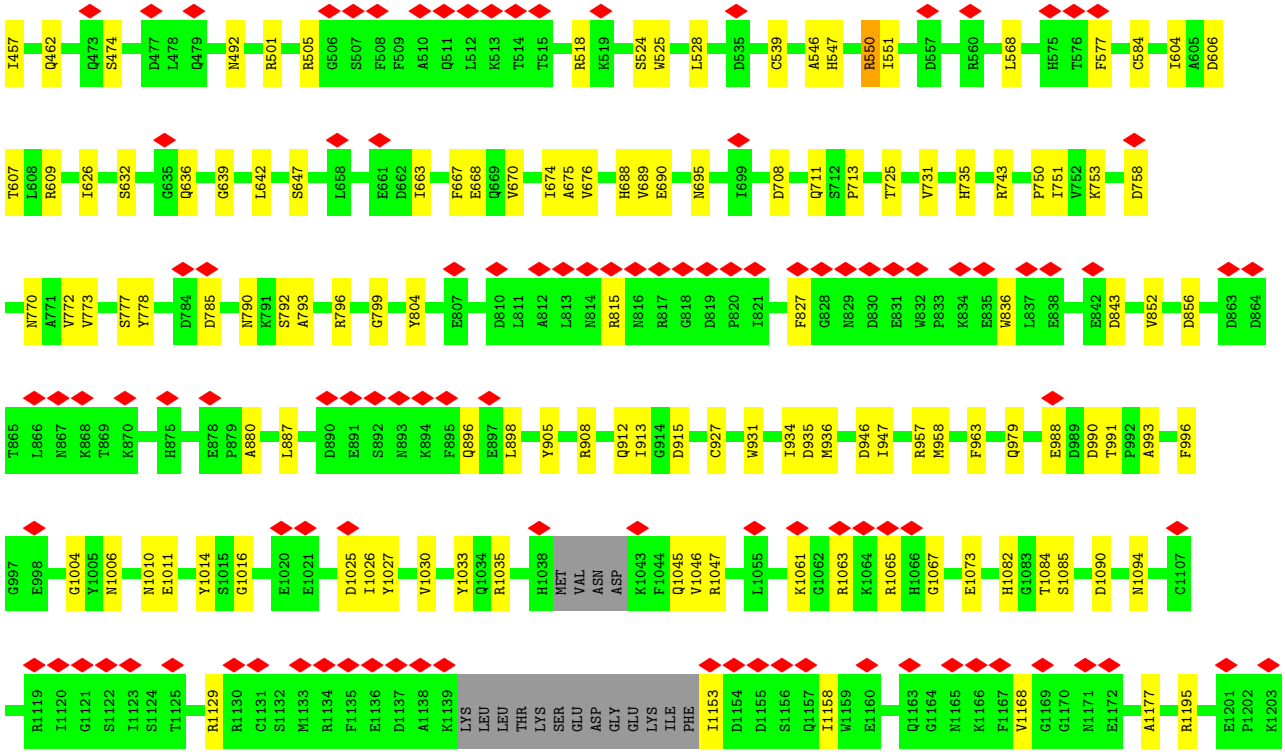




• Molecule 2: DNA-directed RNA polymerase I subunit RPA135

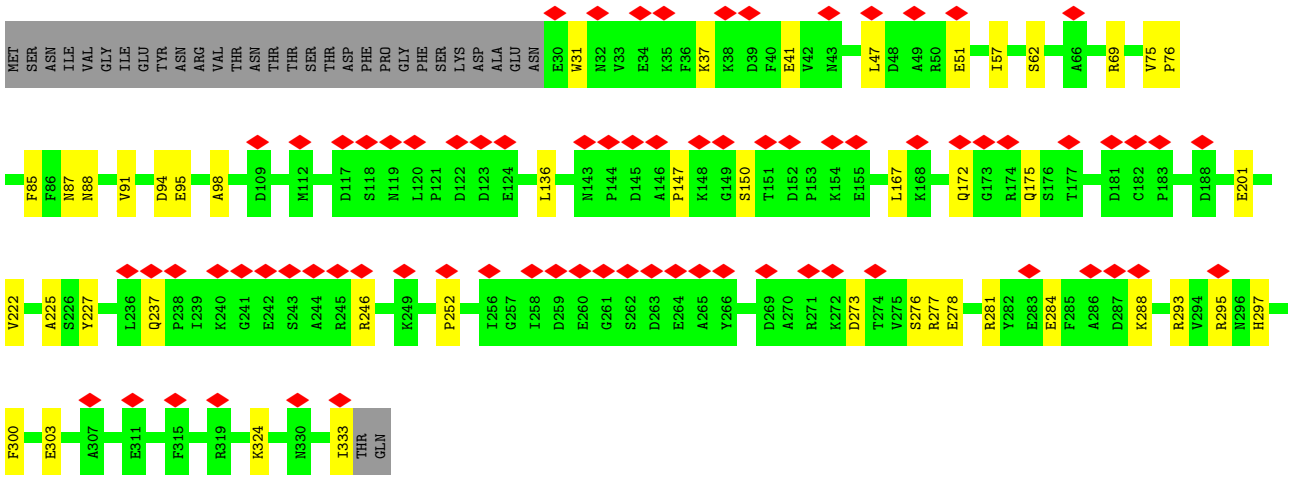
Chain B:





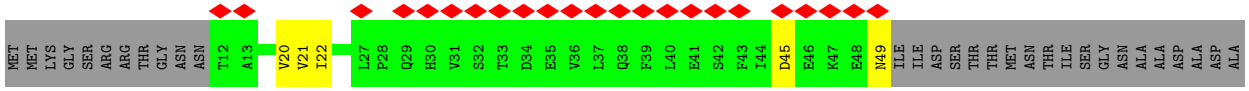
• Molecule 3: DNA-directed RNA polymerases I and III subunit RPAC1

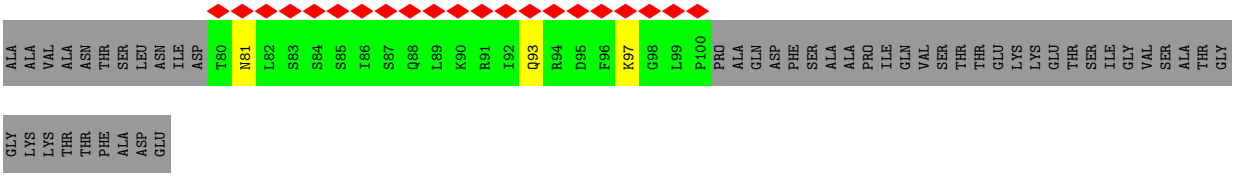
Chain C:



• Molecule 4: DNA-directed RNA polymerase I subunit RPA14

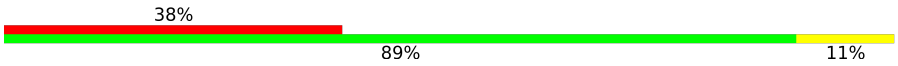
Chain D:

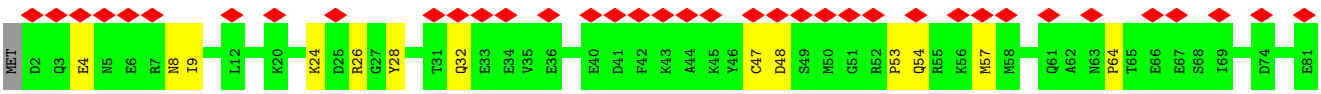


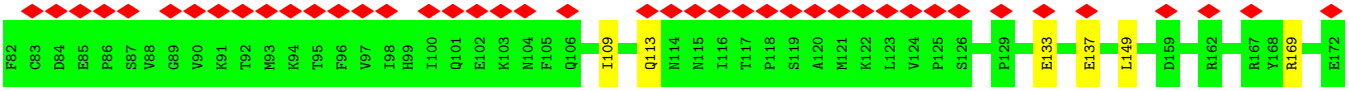


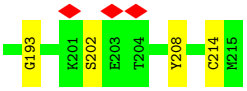
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



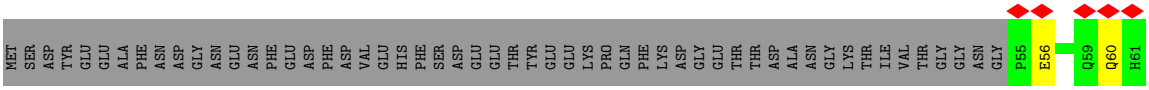


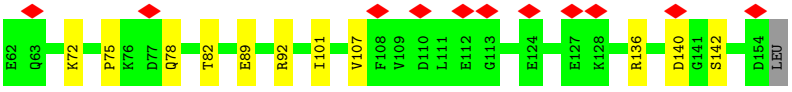




• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

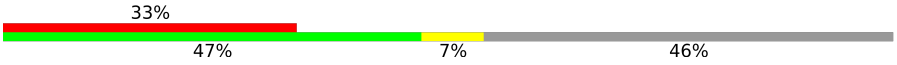
Chain F:

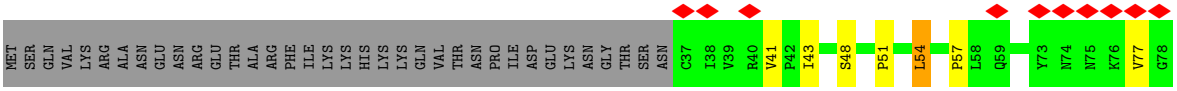


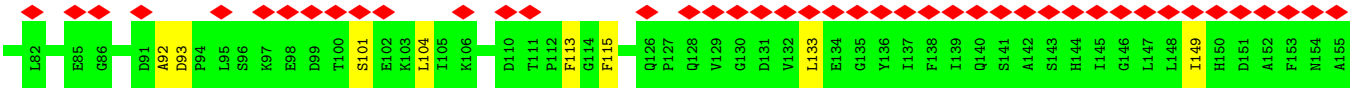


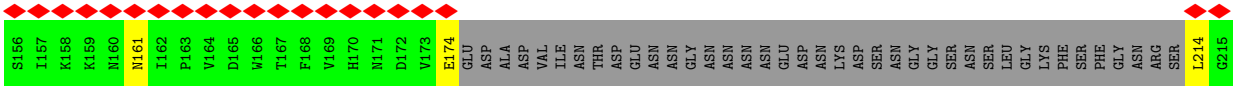
• Molecule 7: DNA-directed RNA polymerase I subunit RPA43

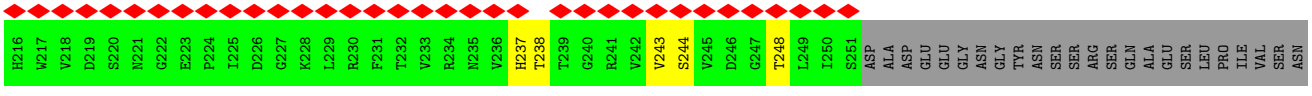
Chain G:





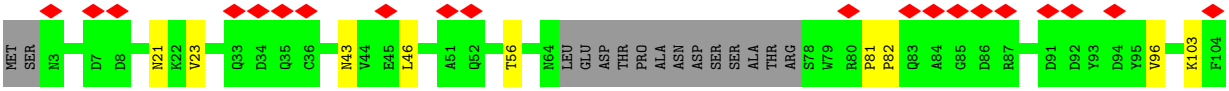


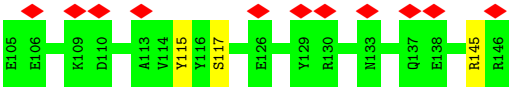




• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

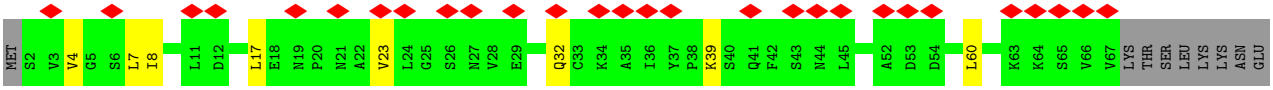
Chain H:

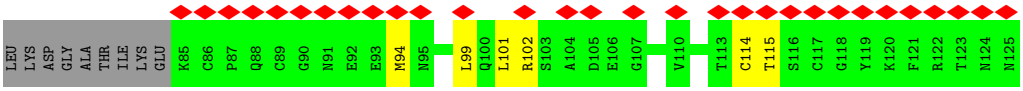




• Molecule 9: DNA-directed RNA polymerase I subunit RPA12

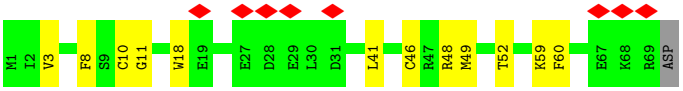
Chain I:





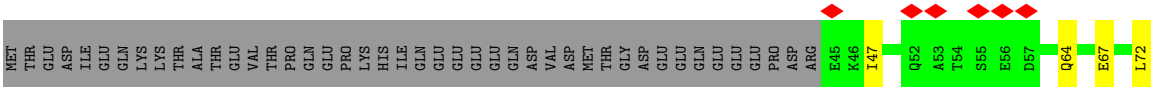
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



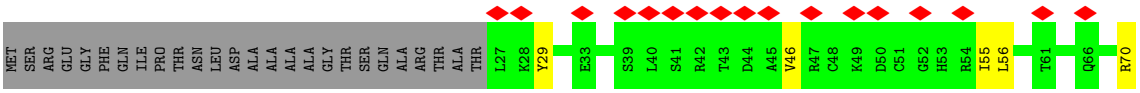
• Molecule 11: DNA-directed RNA polymerases I and III subunit RPAC2

Chain K:



• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

Chain L:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	73660	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 ($e^-/Å^2$)	39	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.316	Depositor
Minimum map value	-0.195	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.017	Depositor
Recommended contour level	0.0777	Depositor
Map size (Å)	270.4, 270.4, 270.4	wwPDB
Map dimensions	260, 260, 260	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.04, 1.04, 1.04	Depositor

5 Model quality

5.1 Standard geometry

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

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	0.36	0/11513	0.58	0/15547
2	B	0.38	0/9452	0.59	0/12774
3	C	0.34	0/2467	0.54	0/3344
4	D	0.27	0/473	0.52	0/641
5	E	0.34	0/1787	0.52	0/2406
6	F	0.34	0/838	0.51	0/1129
7	G	0.29	0/1421	0.52	0/1938
8	H	0.34	0/1070	0.58	0/1449
9	I	0.30	0/822	0.57	0/1111
10	J	0.39	0/578	0.62	0/775
11	K	0.34	0/776	0.61	0/1047
12	L	0.34	0/354	0.58	0/468
All	All	0.36	0/31551	0.57	0/42629

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	11308	0	11384	122	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	9248	0	9141	120	0
3	C	2415	0	2403	30	0
4	D	467	0	468	5	0
5	E	1751	0	1776	13	0
6	F	823	0	841	8	0
7	G	1386	0	1386	16	0
8	H	1052	0	1021	8	0
9	I	810	0	785	8	0
10	J	569	0	586	8	0
11	K	766	0	765	17	0
12	L	352	0	374	5	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
All	All	30954	0	30930	308	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 5.

The worst 5 of 308 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:712:ILE:H	11:K:106:GLN:HE22	1.33	0.73
1:A:1482:LYS:HD3	2:B:307:GLU:HG2	1.75	0.67
3:C:172:GLN:H	3:C:175:GLN:HE21	1.42	0.67
1:A:464:GLU:HA	1:A:469:LYS:HD2	1.79	0.65
11:K:104:ARG:HE	11:K:106:GLN:HE21	1.43	0.65

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	1409/1664 (85%)	1341 (95%)	68 (5%)	0	100	100
2	B	1153/1203 (96%)	1073 (93%)	80 (7%)	0	100	100
3	C	302/335 (90%)	291 (96%)	11 (4%)	0	100	100
4	D	55/137 (40%)	53 (96%)	2 (4%)	0	100	100
5	E	212/215 (99%)	204 (96%)	8 (4%)	0	100	100
6	F	98/155 (63%)	97 (99%)	1 (1%)	0	100	100
7	G	172/326 (53%)	163 (95%)	9 (5%)	0	100	100
8	H	127/146 (87%)	122 (96%)	5 (4%)	0	100	100
9	I	103/125 (82%)	94 (91%)	9 (9%)	0	100	100
10	J	67/70 (96%)	62 (92%)	5 (8%)	0	100	100
11	K	96/142 (68%)	92 (96%)	4 (4%)	0	100	100
12	L	42/70 (60%)	37 (88%)	5 (12%)	0	100	100
All	All	3836/4588 (84%)	3629 (95%)	207 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	1263/1465 (86%)	1261 (100%)	2 (0%)	93	97
2	B	1016/1053 (96%)	1013 (100%)	3 (0%)	92	96
3	C	268/296 (90%)	268 (100%)	0	100	100
4	D	56/116 (48%)	56 (100%)	0	100	100
5	E	196/197 (100%)	196 (100%)	0	100	100
6	F	90/137 (66%)	90 (100%)	0	100	100
7	G	156/291 (54%)	155 (99%)	1 (1%)	86	93

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	H	115/128 (90%)	114 (99%)	1 (1%)	78 90
9	I	94/110 (86%)	92 (98%)	2 (2%)	53 79
10	J	64/65 (98%)	64 (100%)	0	100 100
11	K	88/130 (68%)	88 (100%)	0	100 100
12	L	39/57 (68%)	39 (100%)	0	100 100
All	All	3445/4045 (85%)	3436 (100%)	9 (0%)	92 96

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	4	VAL
9	I	7	LEU
2	B	607	THR
2	B	1195	ARG
7	G	54	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	121	ASN
8	H	35	GLN
11	K	106	GLN
8	H	21	ASN
2	B	684	ASN

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)

Of 7 ligands modelled in this entry, 7 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.

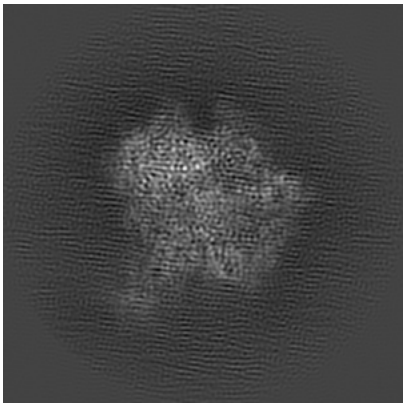
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-0241. These allow visual inspection of the internal detail of the map and identification of artifacts.

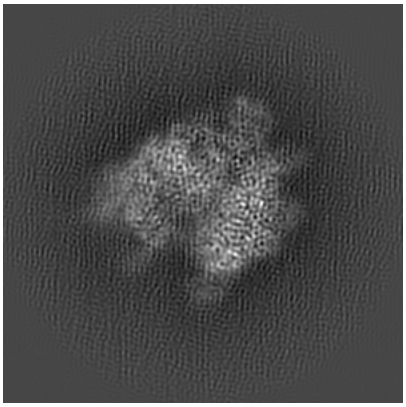
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

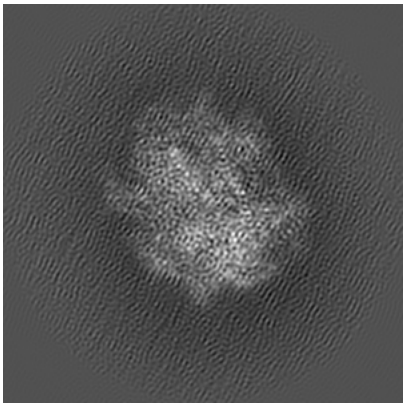
6.1.1 Primary map



X



Y

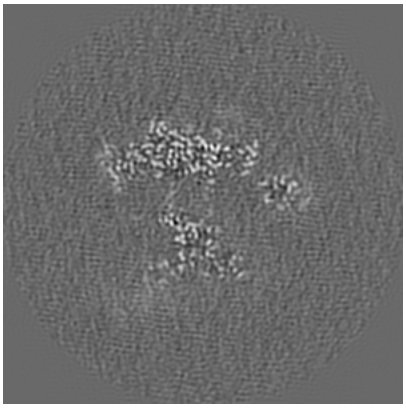


Z

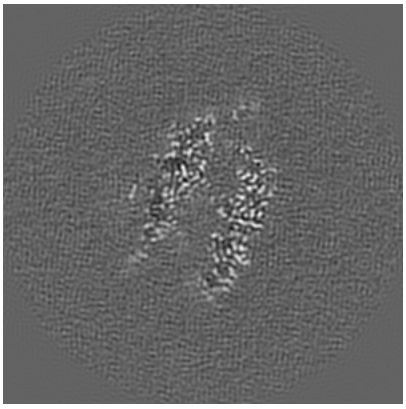
The images above show the map projected in three orthogonal directions.

6.2 Central slices

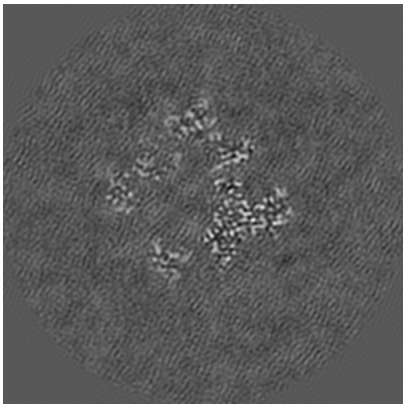
6.2.1 Primary map



X Index: 130



Y Index: 130

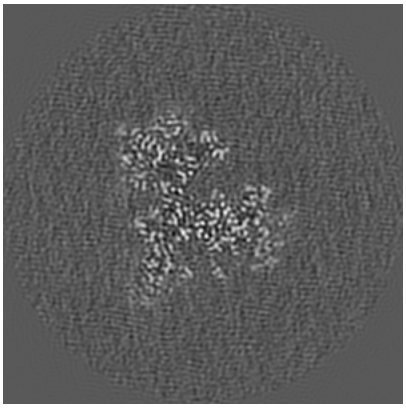


Z Index: 130

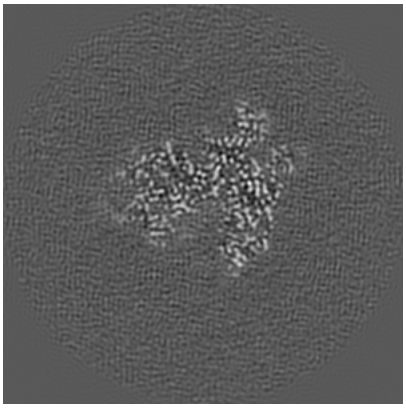
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

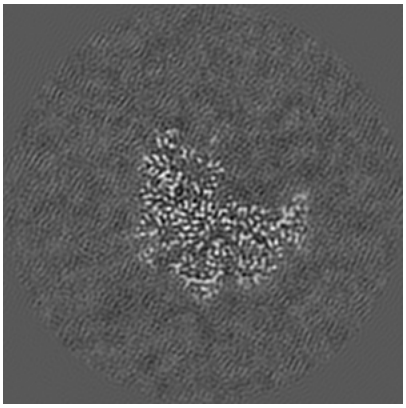
6.3.1 Primary map



X Index: 152



Y Index: 110



Z Index: 153

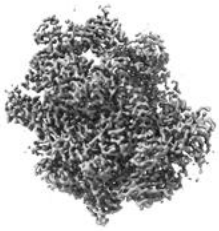
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

6.4.1 Primary map

X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0777. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

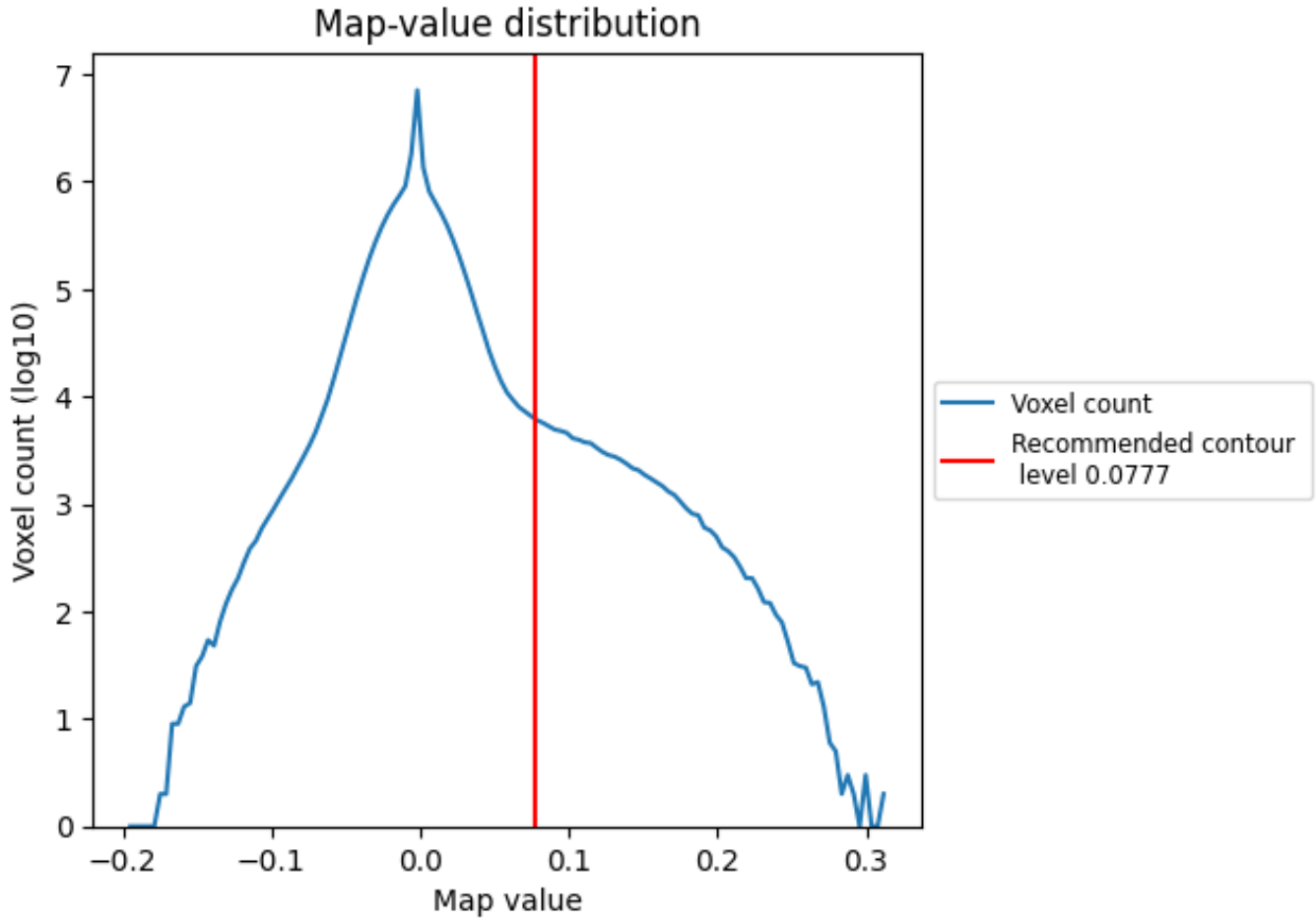
6.5 Mask visualisation (i)

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

7 Map analysis

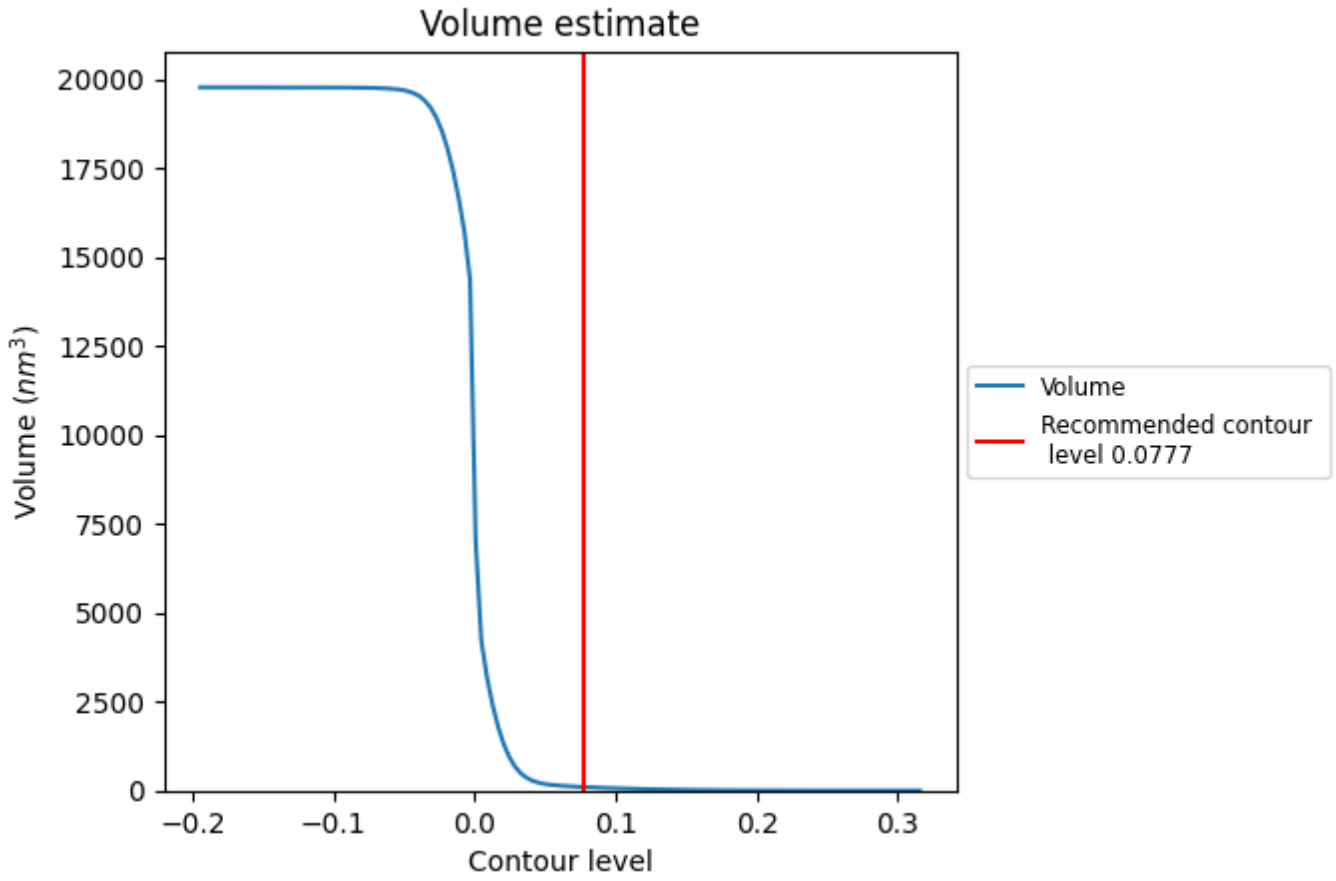
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

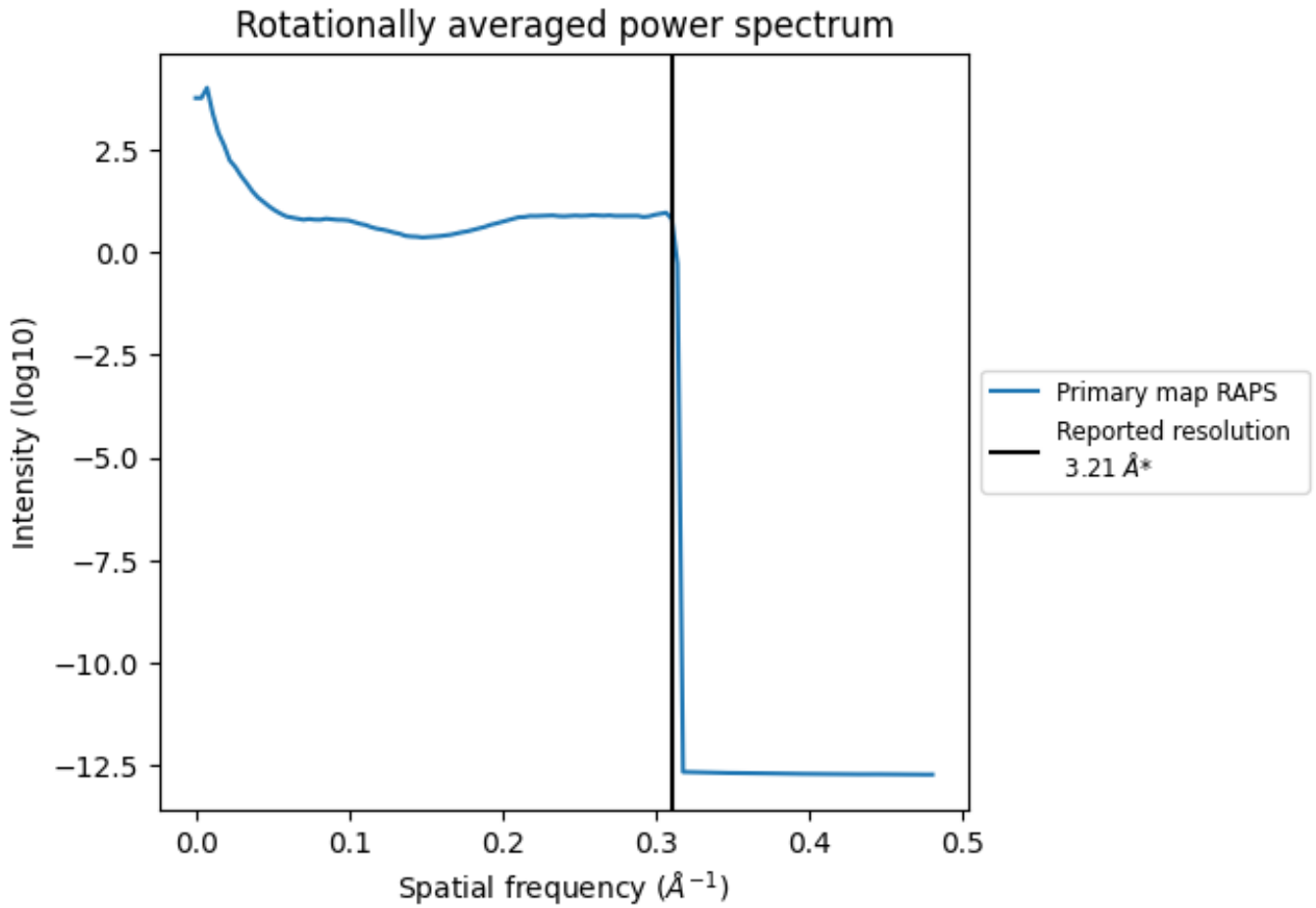
7.2 Volume estimate (i)



The volume at the recommended contour level is 97 nm^3; this corresponds to an approximate mass of 88 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)

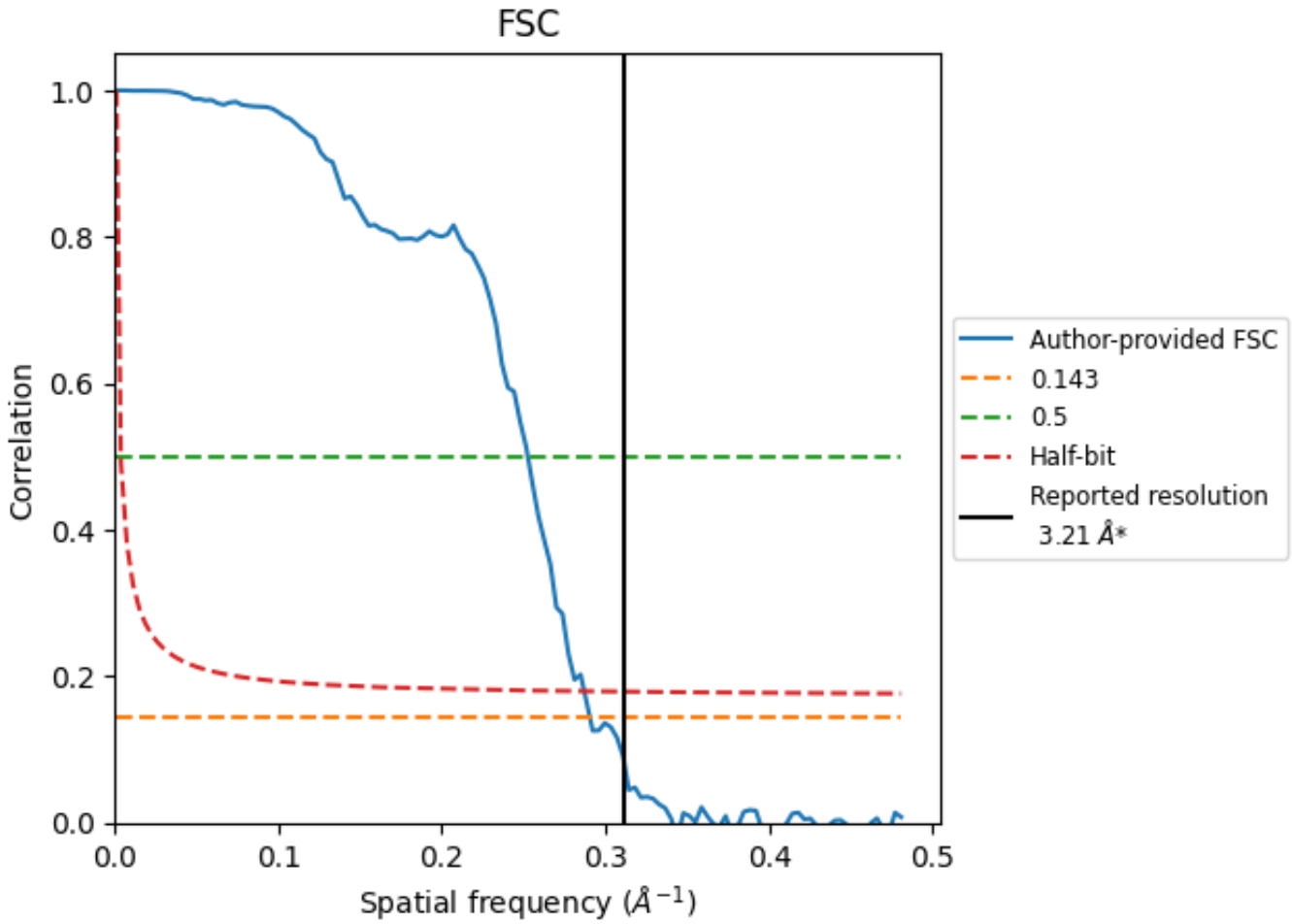


*Reported resolution corresponds to spatial frequency of 0.312 $Å^{-1}$

8 Fourier-Shell correlation

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.312 Å^{-1}

8.2 Resolution estimates (i)

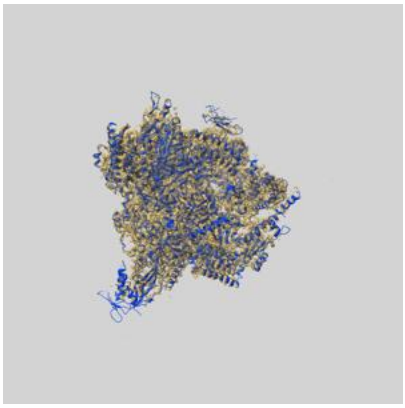
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.21	-	-
Author-provided FSC curve	3.44	3.96	3.48
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

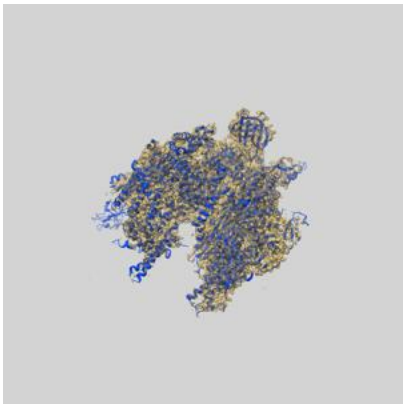
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-0241 and PDB model 6HLS. Per-residue inclusion information can be found in section 3 on page 7.

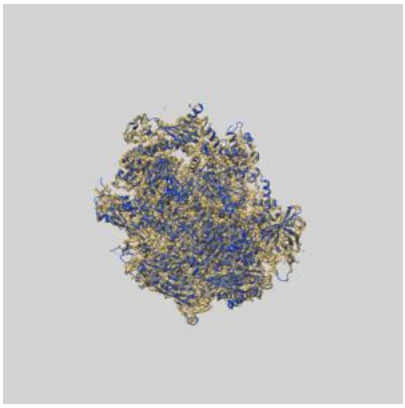
9.1 Map-model overlay



X



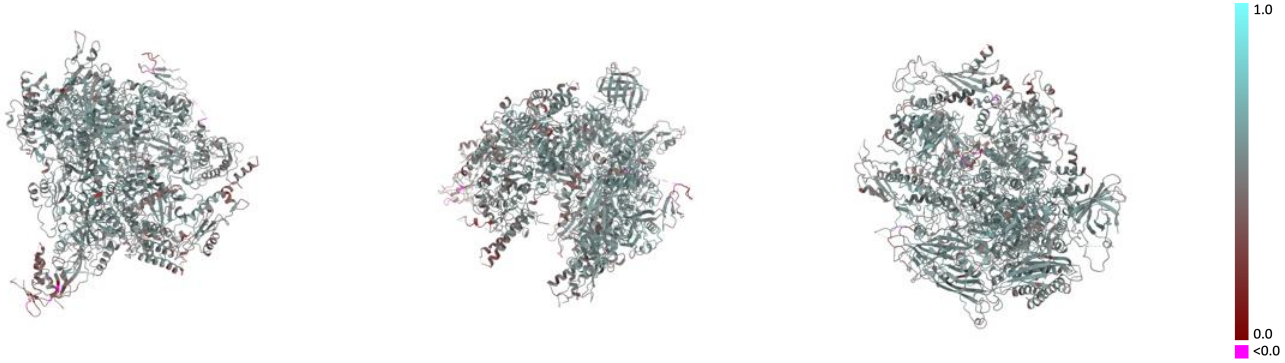
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0777 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.

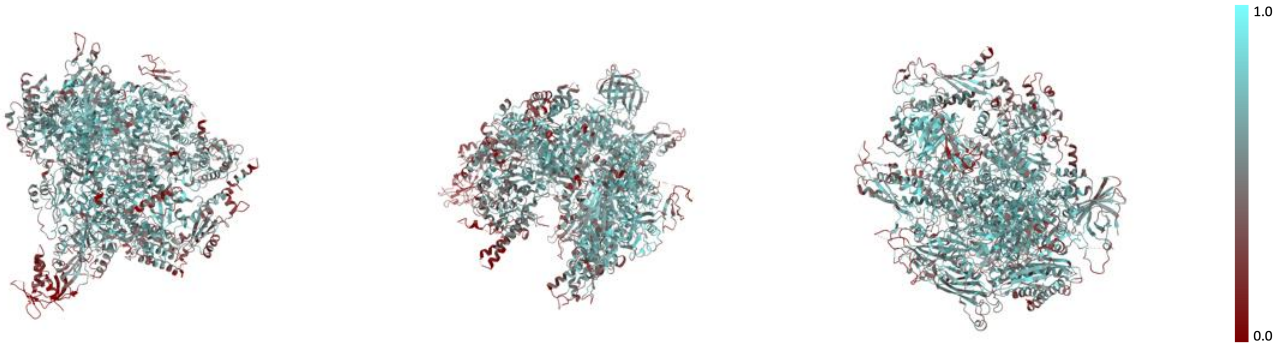
9.2 Q-score mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according 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.

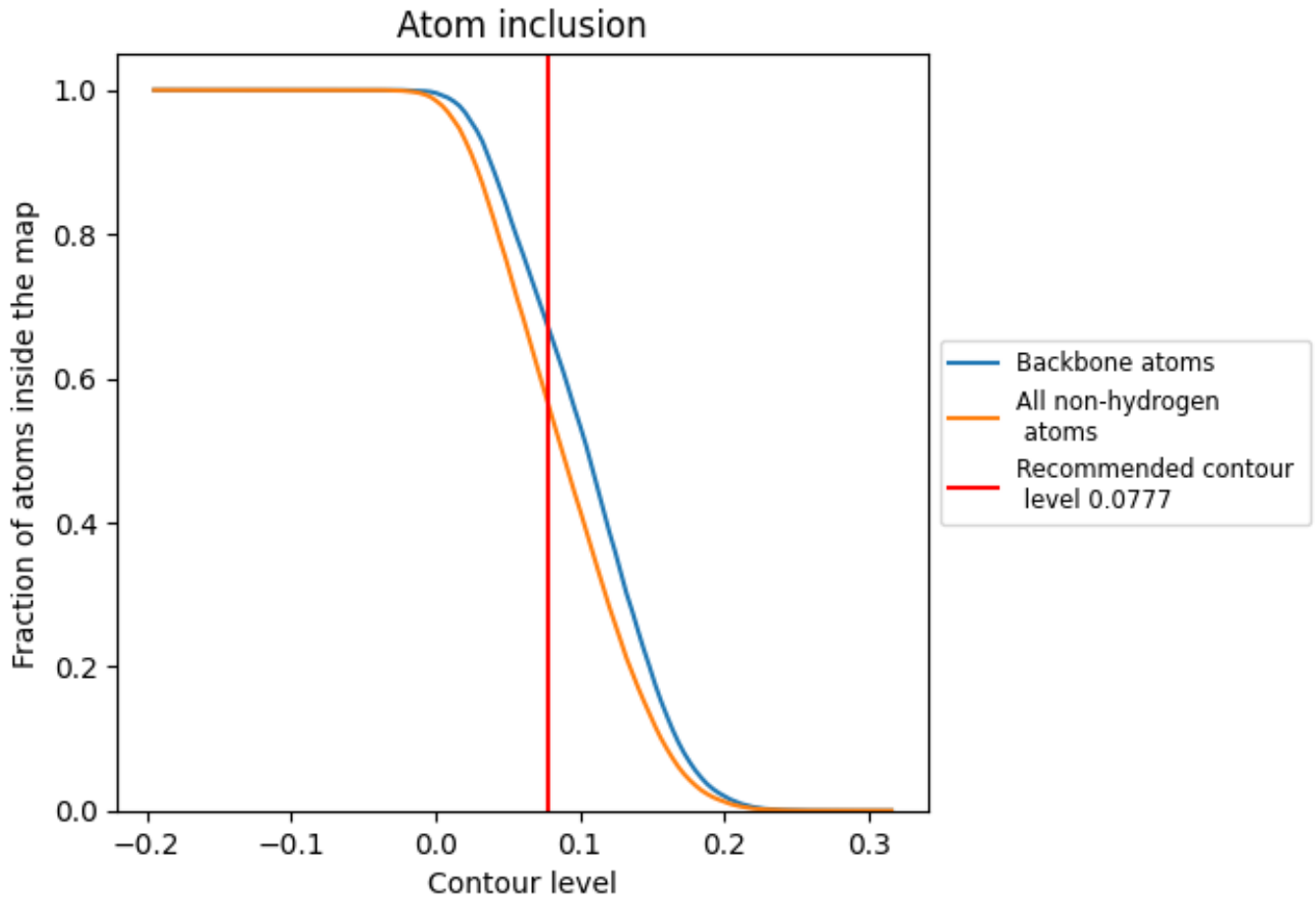
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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 (0.0777).

9.4 Atom inclusion (i)



At the recommended contour level, 67% of all backbone atoms, 57% 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 (0.0777) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5667	0.5150
A	0.5854	0.5200
B	0.6298	0.5330
C	0.5504	0.5260
D	0.2549	0.4150
E	0.5094	0.4900
F	0.6130	0.5390
G	0.2887	0.4170
H	0.5341	0.5190
I	0.3591	0.4360
J	0.6721	0.5550
K	0.5744	0.5290
L	0.5088	0.5090