wwPDB EM Validation Summary Report

Mar 17, 2025 – 08:31 AM EDT

PDB ID	:	9B2K
EMDB ID	:	EMD-44111
Title	:	SpCas9 with dual-guide RNA in open conformation
Authors	:	Korolev, S.; Gagnon, K.
Deposited on	:	2024-03-15
Resolution	:	3.67 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*
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:

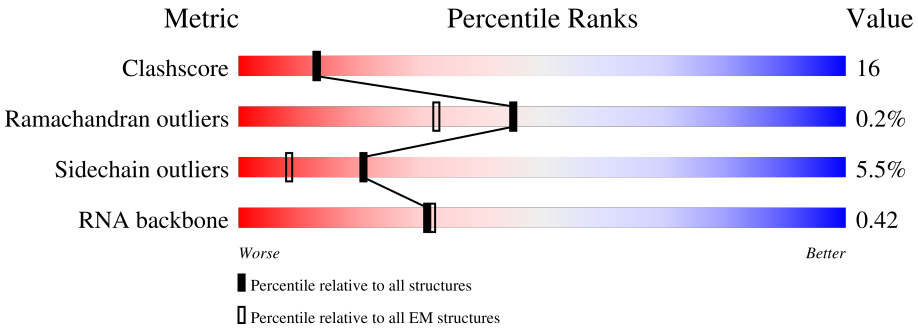
EMDB validation analysis	:	0.0.1.dev117
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

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 3.67 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	42	12% 17% 26% 57%
2	B	70	10% 27% 37% 14% • 20%
3	P	1384	10% 48% 26% • 23%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10242 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 RNA chain called RNA (5'-R(P*GP*UP*UP*UP*UP*AP*GP*AP*GP*CP *UP*AP*UP*GP*CP*UP*GP*U)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	18	Total	C	N	O	P	0	0
			381	170	62	131	18		

- Molecule 2 is a RNA chain called RNA (56-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	56	Total	C	N	O	P	0	0
			1206	540	229	381	56		

- Molecule 3 is a protein called CRISPR-associated endonuclease Cas9/Csn1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	P	1062	Total	C	N	O	S	0	0
			8655	5503	1511	1625	16		

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	1369	PRO	-	expression tag	UNP Q99ZW2
P	1370	LYS	-	expression tag	UNP Q99ZW2
P	1371	LYS	-	expression tag	UNP Q99ZW2
P	1372	LYS	-	expression tag	UNP Q99ZW2
P	1373	ARG	-	expression tag	UNP Q99ZW2
P	1374	LYS	-	expression tag	UNP Q99ZW2
P	1375	VAL	-	expression tag	UNP Q99ZW2
P	1376	MET	-	expression tag	UNP Q99ZW2
P	1377	ASP	-	expression tag	UNP Q99ZW2
P	1378	LYS	-	expression tag	UNP Q99ZW2
P	1379	HIS	-	expression tag	UNP Q99ZW2
P	1380	HIS	-	expression tag	UNP Q99ZW2
P	1381	HIS	-	expression tag	UNP Q99ZW2
P	1382	HIS	-	expression tag	UNP Q99ZW2

Continued on next page...

Continued from previous page...

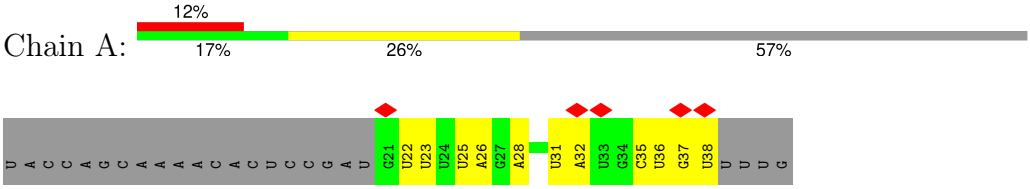
Chain	Residue	Modelled	Actual	Comment	Reference
P	1383	HIS	-	expression tag	UNP Q99ZW2
P	1384	HIS	-	expression tag	UNP Q99ZW2

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.

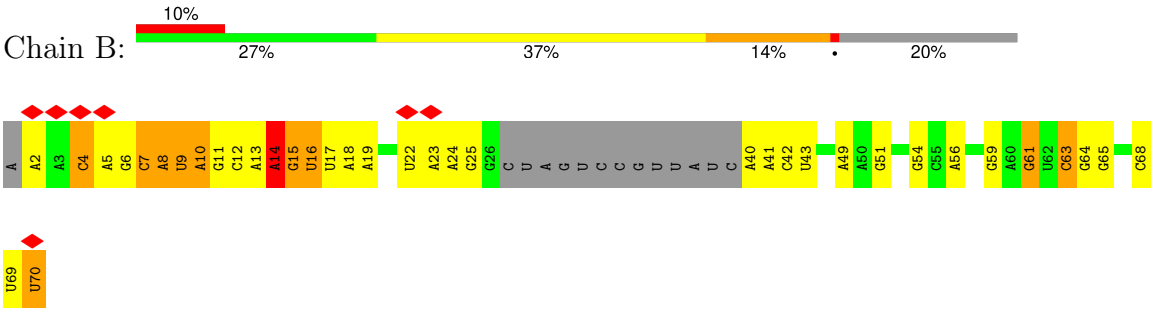
• Molecule 1: RNA (5'-R(P*GP*UP*UP*UP*UP*AP*GP*AP*GP*CP*UP*AP*UP*GP*CP*UP*GP*U)-3')

Chain A:



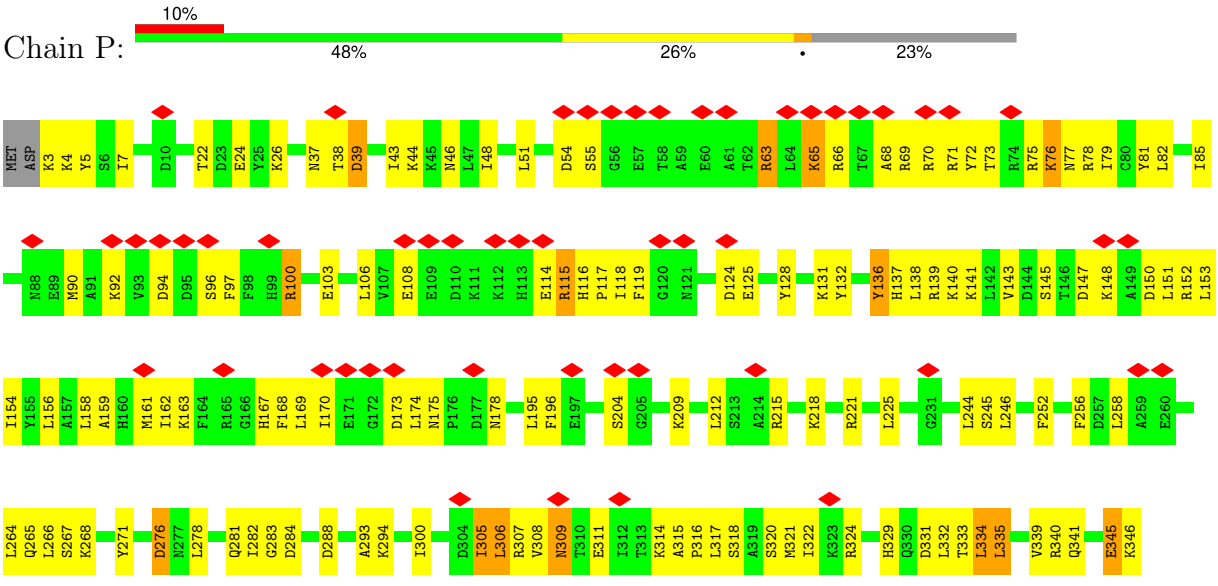
• Molecule 2: RNA (56-MER)

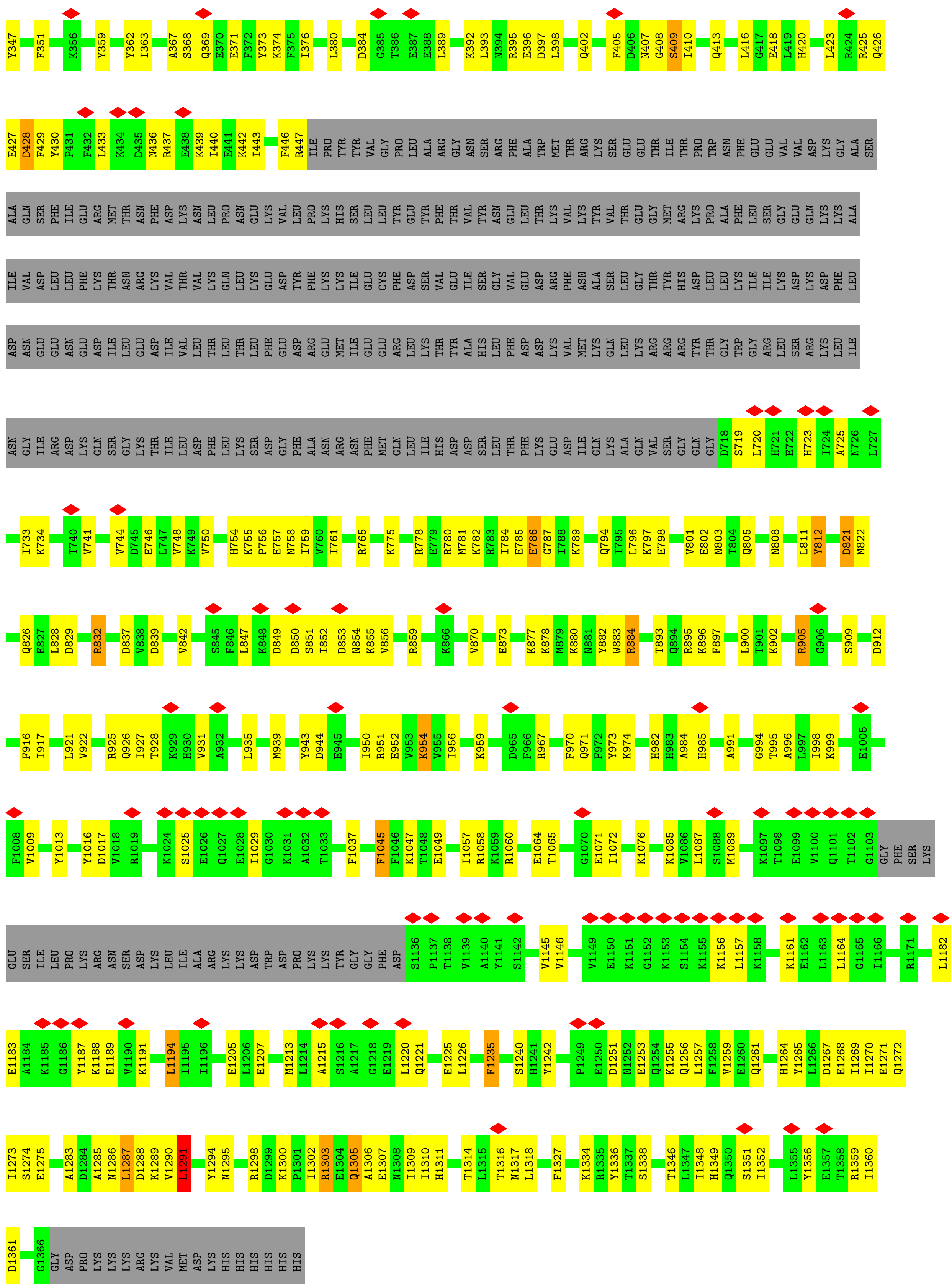
Chain B:



• Molecule 3: CRISPR-associated endonuclease Cas9/Csn1

Chain P:




Y347 F351 K356 Y359 Y362 I363 A367 S368 Q369 E370 E371 F372 Y373 K374 F375 I376 L380 D384 G385 T386 E387 E388 L389 K392 L393 N394 R395 E396 D397 L398 Q402 F405 D406 N407 G408 S409 I410 Q413 L416 G417 E418 L419 H420 L423 R424 R425 Q426
E427 D428 F429 Y430 P431 F432 L433 K434 D435 N436 R437 E438 K439 I440 E441 K442 I443 F446 R447 ILE PRO TYR TYR VAL GLY PRO LEU ALA ARG GLY ASN SER ARG PHE ALA TRP MET THR ARG LYS SER GLU GLU THR ILE THR PRO TRP ASN PHE GLU GLU VAL VAL ASP LYS GLY ALA SER
ALA GLN SER PHE ILE GLU ARG MET THR ASN PHE ASP LYS ASN LEU PRO ASN GLU LYS VAL LEU PRO LYS HIS SER LEU LEU TYR GLU TYR PHE THR VAL TYR ASN GLU LEU THR LYS VAL LYS TYR VAL THR GLU GLY MET ARG LYS PRO ALA PHE LEU SER GLY GLU GLN LYS LYS ALA
ILE VAL ASP LEU LEU PHE LYS THR ASN ARG LYS VAL THR VAL LYS GLN LEU LYS GLU ASP TYR PHE LYS LYS ILE GLU CYS PHE ASP SER VAL GLU ILE SER GLY VAL GLU ASP ARG PHE ASN ALA SER LEU GLY THR TYR HIS ASP LEU LEU LYS ILE ILE LYS ASP LYS ASP PHE LEU
ASP ASN GLU GLU ASN GLU ASP ILE LEU GLU ASP ILE VAL LEU THR LEU THR LEU PHE GLU ASP ARG GLU MET ILE GLU GLU ARG LEU LYS THR TYR ALA HIS LEU PHE ASP ASP LYS VAL MET LYS GLN LEU LYS ARG ARG ARG TYR THR GLY TRP GLY ARG LEU SER ARG LYS LEU ILE
ASN GLY ILE ARG ASP LYS GLN SER GLY LYS THR ILE LEU ASP PHE LEU LYS SER ASP GLY PHE ALA ASN ARG ASN PHE MET GLN LEU ILE HIS ASP ASP SER LEU THR PHE LYS GLU ASP ILE GLN LYS ALA GLN VAL SER GLY GLN GLY D718 S719 L720 H721 E722 H723 I724 A725 N726 L727
I733 K734 T740 V741 V744 D745 E746 L747 V748 K749 V750 H754 K755 P756 E757 N758 I759 V760 I761 R765 K775 R778 E779 R780 M781 K782 R783 I784 E785 E786 G787 I788 K789 Q794 I795 L796 K797 E798 V801 E802 N803 T804 Q805 N808 L811 Y812 D821 M822
Q826 E827 L828 D829 R832 D837 V838 D839 V842 S845 F846 L847 K848 D849 D850 S851 I852 D853 N854 K855 V856 R859 K866 V870 E873 K877 K878 M879 K880 N881 Y882 W883 R884 T893 Q894 R895 K896 F897 L900 T901 K902 R905 G906 S909 D912
F916 I917 L921 V922 R925 Q926 I927 T928 K929 H930 V931 A932 L935 M939 Y943 D944 E945 I950 R951 E952 V953 K954 V955 I956 K959 D965 F966 R967 F970 Q971 F972 Y973 K974 H982 H983 A984 H985 A991 G994 T995 A996 L997 I998 K999 E1005
F1008 V1009 Y1013 Y1016 D1017 V1018 R1019 K1024 S1025 E1026 Q1027 E1028 I1029 G1030 K1031 A1032 T1033 F1037 F1045 F1046 K1047 T1048 E1049 I1057 R1058 K1059 R1060 E1064 T1065 G1070 E1071 I1072 K1076 K1085 V1086 L1087 S1088 M1089 K1097 T1098 E1099 V1100 Q1101 T1102 G1103 GLY PHE SER LYS
GLU SER ILE LEU PRO LYS ARG ASN SER ASP LYS LEU ILE ALA ARG LYS LYS ASP TRP ASP PRO LYS LYS TYR GLY GLY PHE ASP S1136 P1137 T1138 V1139 A1140 Y1141 S1142 V1145 V1146 V1149 E1150 K1151 G1152 K1153 S1154 K1155 K1156 L1157 K1158 K1161 E1162 L1163 L1164 G1165 I1166 R1171 L1182
E1183 A1184 K1185 G1186 Y1187 K1188 E1189 V1190 K1191 L1194 I1195 I1196 E1205 L1206 E1207 M1213 L1214 A1215 S1216 A1217 G1218 E1219 L1220 Q1221 E1225 L1226 F1235 S1240 H1241 Y1242 P1249 E1250 D1251 N1252 E1253 Q1254 K1255 Q1256 L1257 F1258 V1259 E1260 Q1261 H1264 Y1265 L1266 D1267 E1268 I1269 I1270 E1271 Q1272
I1273 S1274 E1275 A1283 D1284 A1285 N1286 L1287 D1288 K1289 V1290 L1291 Y1294 N1295 R1298 D1299 K1300 P1301 I1302 R1303 E1304 Q1305 A1306 E1307 N1308 I1309 I1310 H1311 T1314 L1315 T1316 N1317 L1318 F1327 K1334 R1335 Y1336 T1337 S1338 T1346 L1347 I1348 H1349 Q1350 S1351 I1352 L1355 Y1356 E1357 T1358 R1359 I1360
D1361 G1366 GLY ASP PRO LYS LYS LYS ARG LYS VAL MET ASP LYS HIS HIS HIS HIS HIS HIS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	164524	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 ($e^-/Å^2$)	51.37	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	0.557	Depositor
Minimum map value	-0.267	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.014	Depositor
Recommended contour level	0.1	Depositor
Map size (Å)	280.0, 280.0, 280.0	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.7, 0.7, 0.7	Depositor

5 Model quality

5.1 Standard geometry

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	0.25	0/424	0.92	0/658
2	B	0.30	0/1352	0.95	7/2104 (0.3%)
3	P	0.29	0/8801	0.59	3/11828 (0.0%)
All	All	0.29	0/10577	0.68	10/14590 (0.1%)

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
3	P	0	2

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	9	U	C2-N1-C1'	7.23	126.38	117.70
2	B	7	C	N1-C2-O2	6.47	122.78	118.90
2	B	7	C	N3-C2-O2	-6.42	117.41	121.90
3	P	1291	LEU	CA-CB-CG	5.95	128.99	115.30
2	B	9	U	C6-N1-C1'	-5.88	112.96	121.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	P	305	ILE	Mainchain
3	P	63	ARG	Mainchain

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	381	0	191	6	0
2	B	1206	0	607	31	0
3	P	8655	0	8817	297	0
All	All	10242	0	9615	324	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 16.

The worst 5 of 324 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:P:305:ILE:HG13	3:P:306:LEU:HD13	1.26	1.17
3:P:306:LEU:HG	3:P:316:PRO:HB2	1.58	0.85
3:P:215:ARG:HG3	3:P:307:ARG:HH22	1.42	0.84
3:P:1311:HIS:O	3:P:1314:THR:HG23	1.80	0.81
3:P:79:ILE:HA	3:P:82:LEU:HD12	1.65	0.79

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
3	P	1056/1384 (76%)	982 (93%)	72 (7%)	2 (0%)	44 72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	P	309	ASN
3	P	409	SER

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
3	P	950/1243 (76%)	898 (94%)	52 (6%)	18 45

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	P	409	SER
3	P	832	ARG
3	P	1303	ARG
3	P	426	GLN
3	P	803	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
3	P	1241	HIS
3	P	990	ASN
3	P	369	GLN
3	P	771	GLN
3	P	354	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	17/42 (40%)	2 (11%)	0
2	B	54/70 (77%)	18 (33%)	0
All	All	71/112 (63%)	20 (28%)	0

5 of 20 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	28	A
1	A	37	G
2	B	8	A
2	B	10	A
2	B	11	G

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

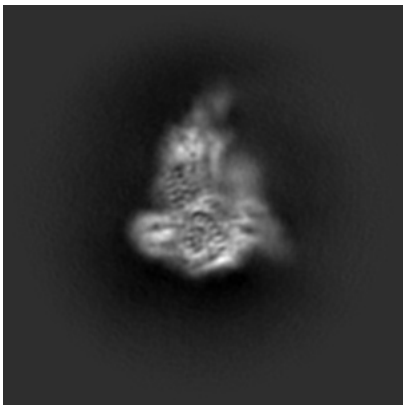
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-44111. These allow visual inspection of the internal detail of the map and identification of artifacts.

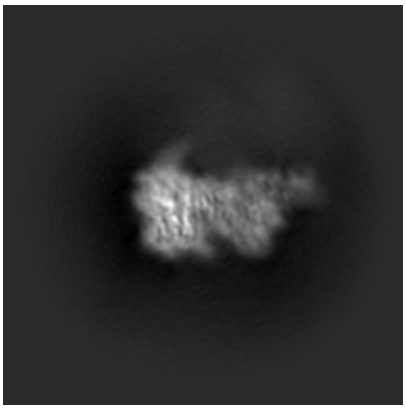
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

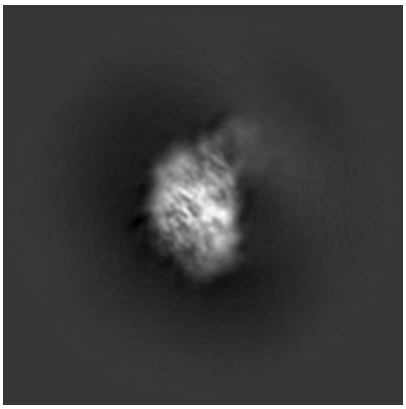
6.1.1 Primary map



X

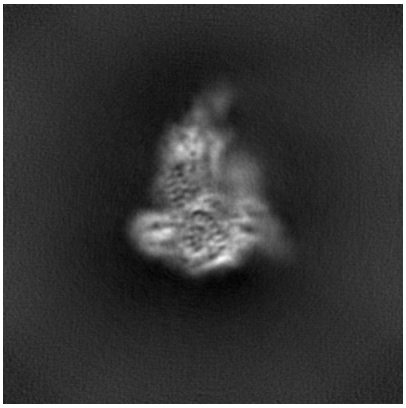


Y

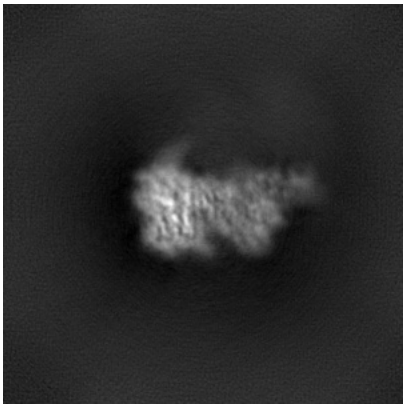


Z

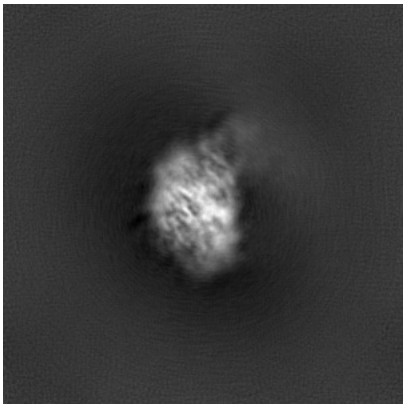
6.1.2 Raw map



X



Y

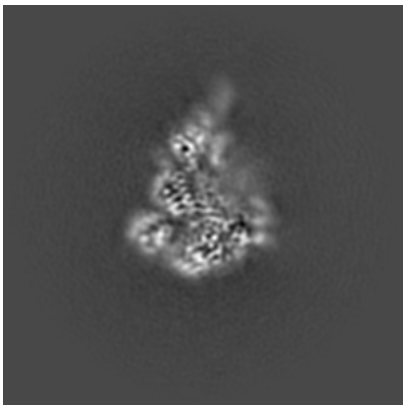


Z

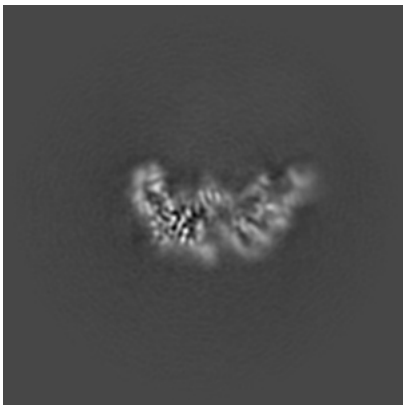
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

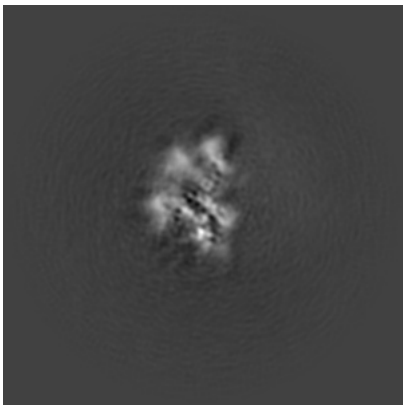
6.2.1 Primary map



X Index: 200

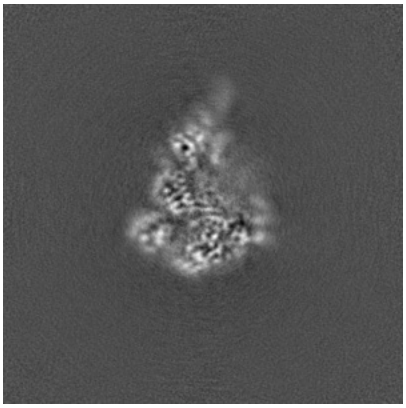


Y Index: 200

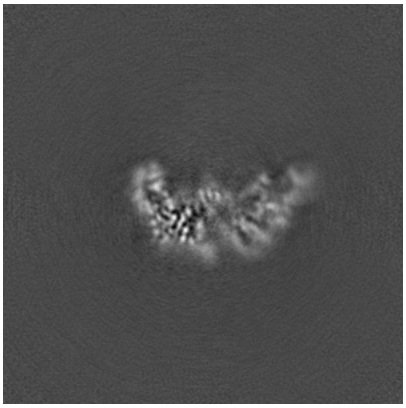


Z Index: 200

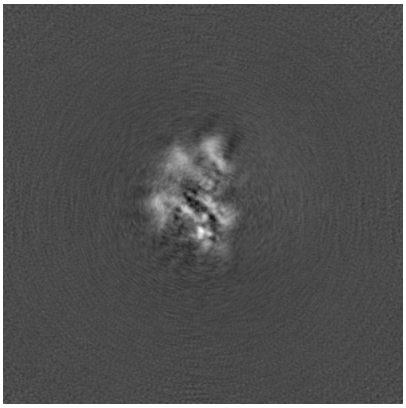
6.2.2 Raw map



X Index: 200



Y Index: 200

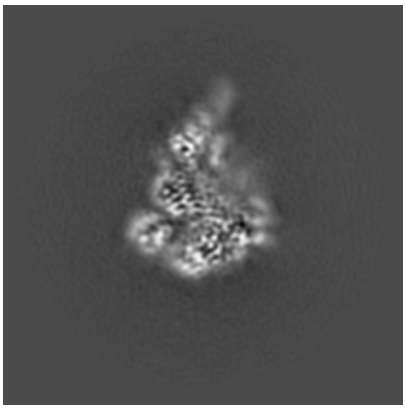


Z Index: 200

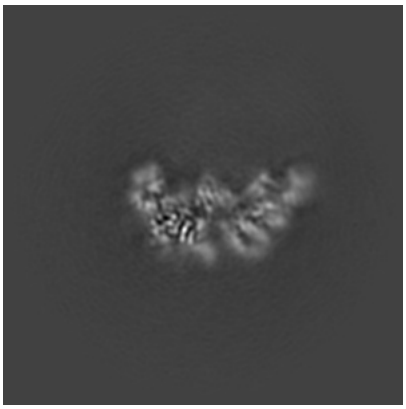
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

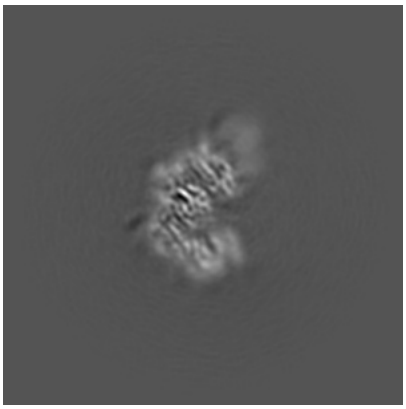
6.3.1 Primary map



X Index: 201

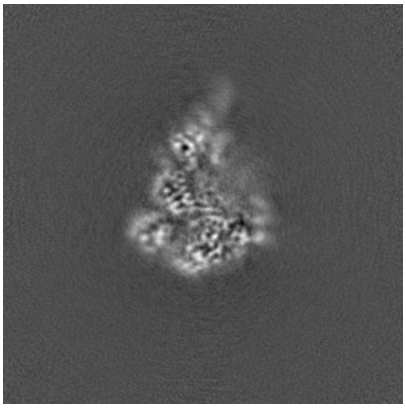


Y Index: 197

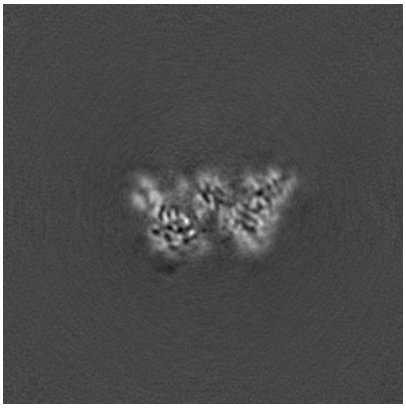


Z Index: 166

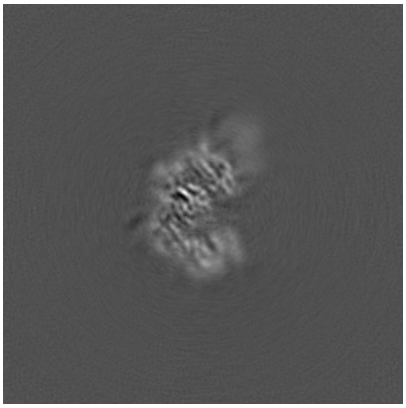
6.3.2 Raw map



X Index: 200



Y Index: 185

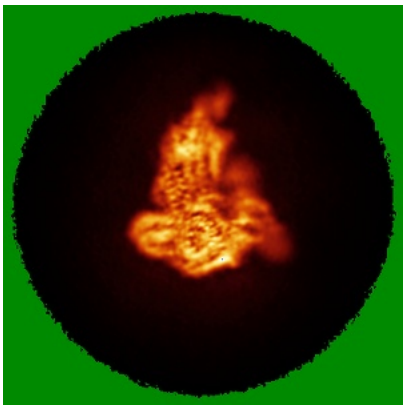


Z Index: 166

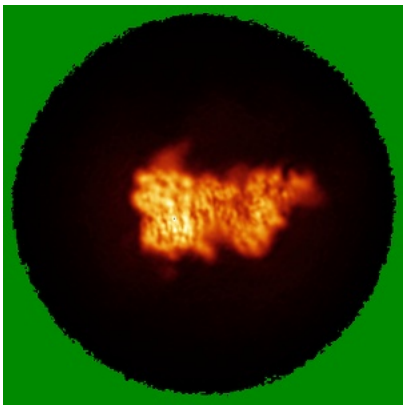
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

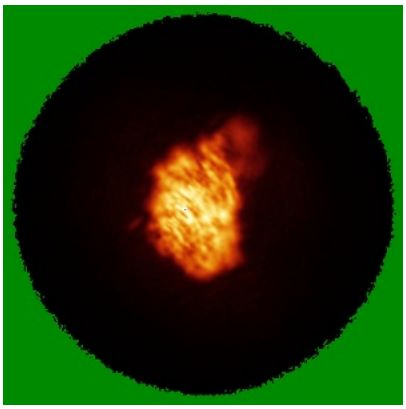
6.4.1 Primary map



X

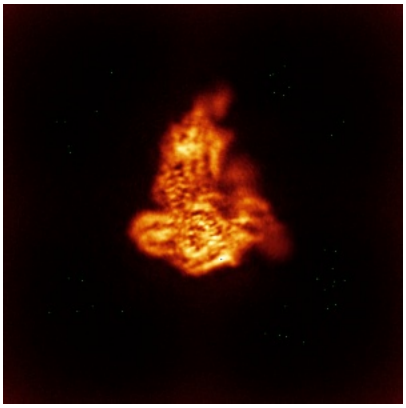


Y

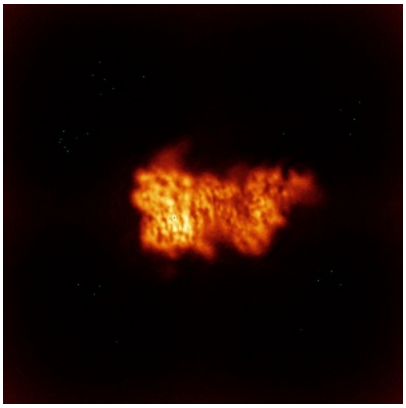


Z

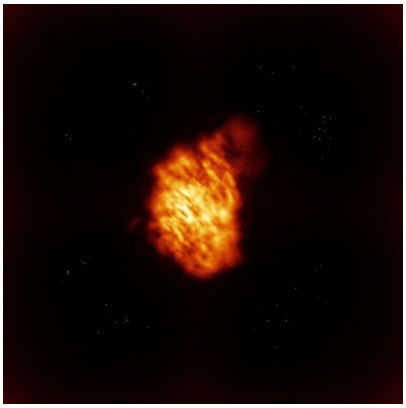
6.4.2 Raw map



X



Y

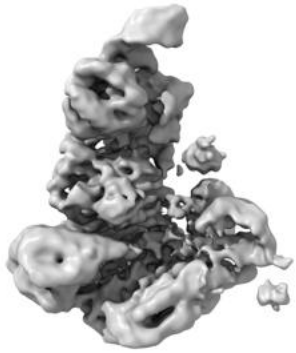


Z

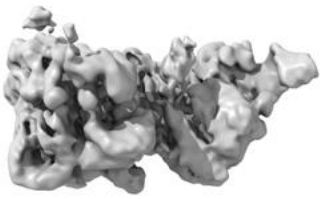
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

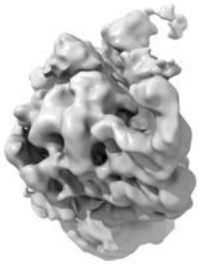
6.5.1 Primary map



X



Y

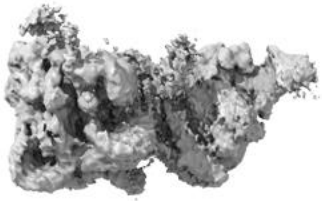


Z

The images above show the 3D surface view of the map at the recommended contour level 0.1. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

X



Y

Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

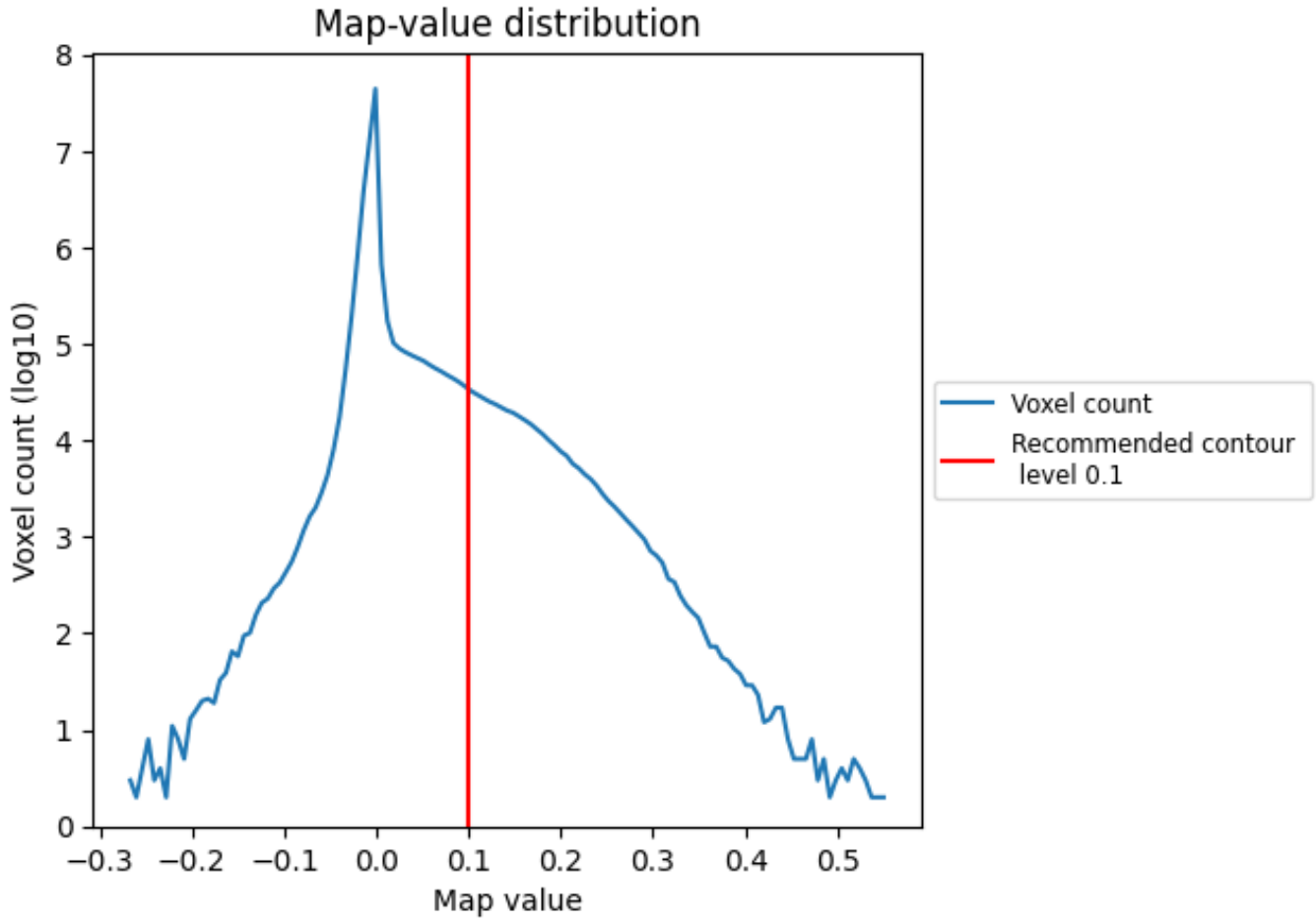
6.6 Mask visualisation (i)

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

7 Map analysis

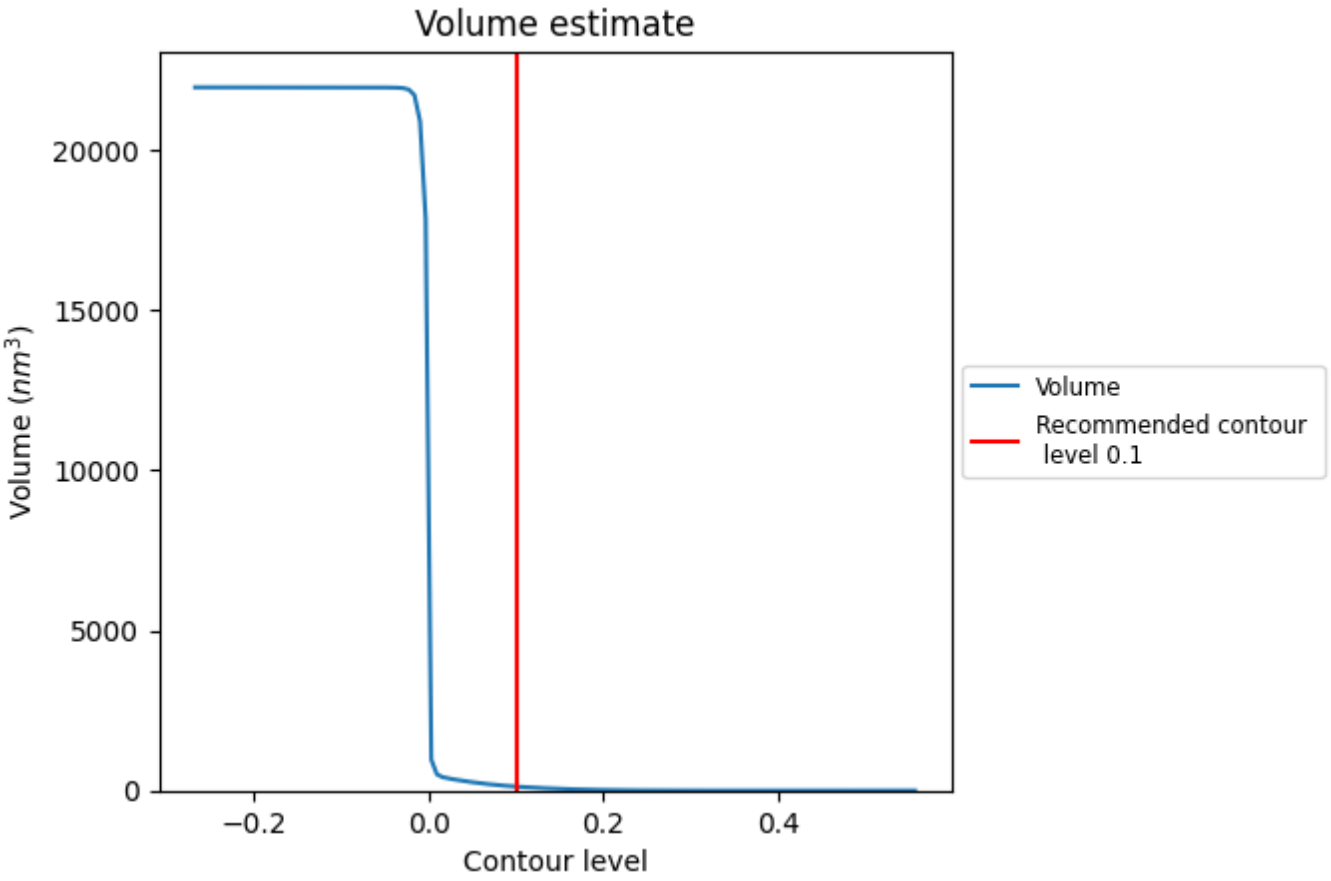
This section contains the results of statistical analysis of the map.

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.

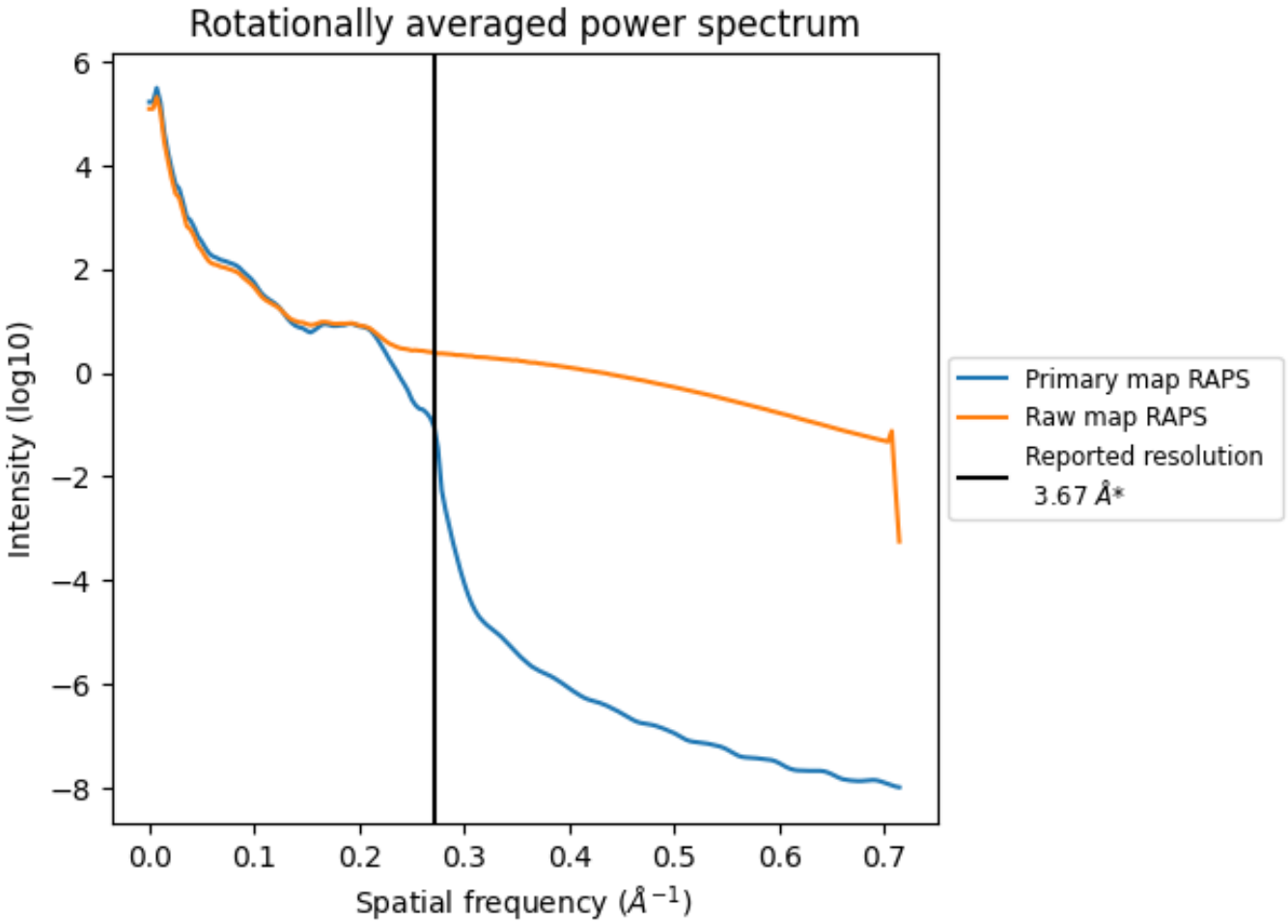
7.2 Volume estimate (i)



The volume at the recommended contour level is 125 nm^3; this corresponds to an approximate mass of 113 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)

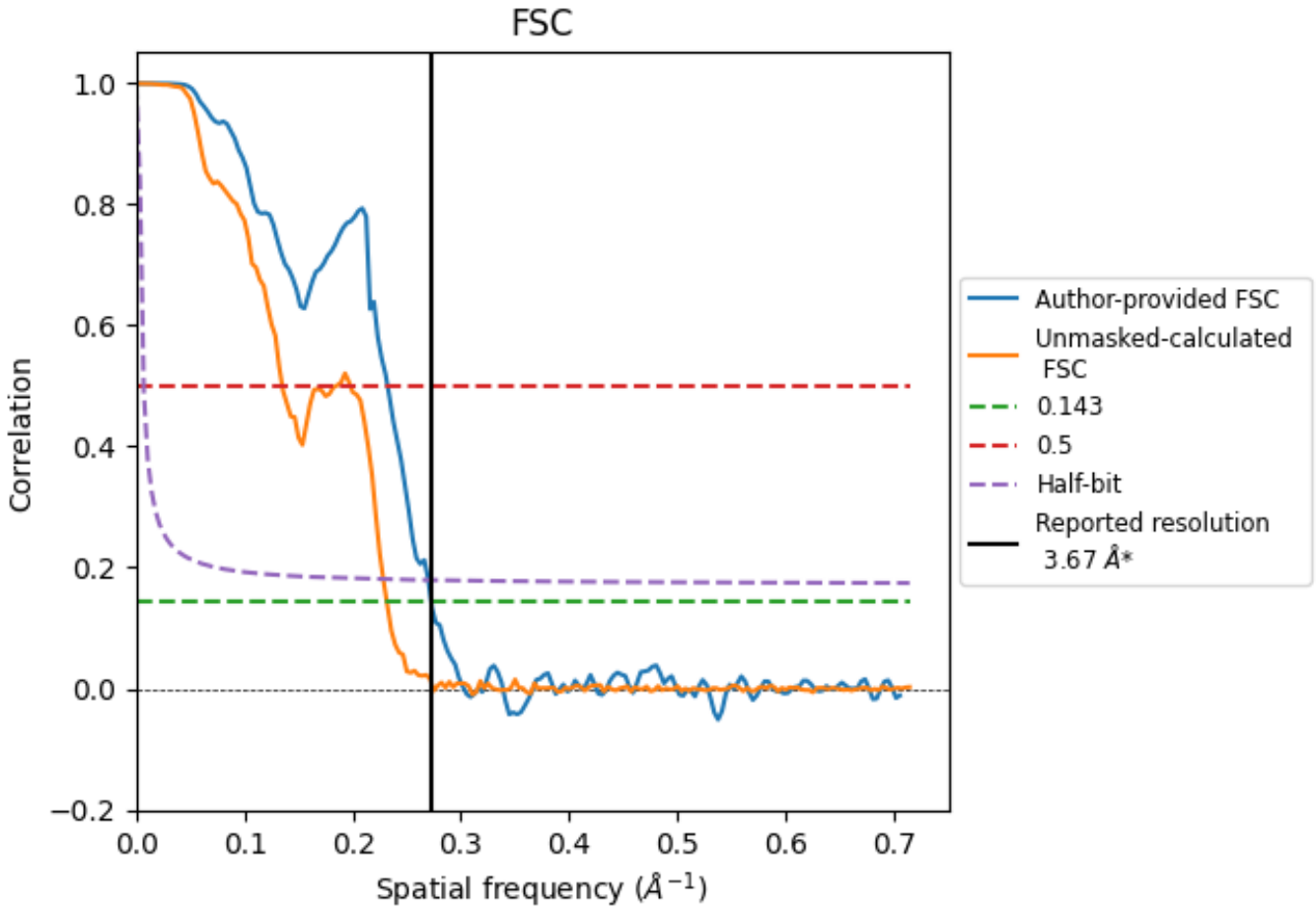


*Reported resolution corresponds to spatial frequency of 0.272 Å^{-1}

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 0.272 Å^{-1}

8.2 Resolution estimates (i)

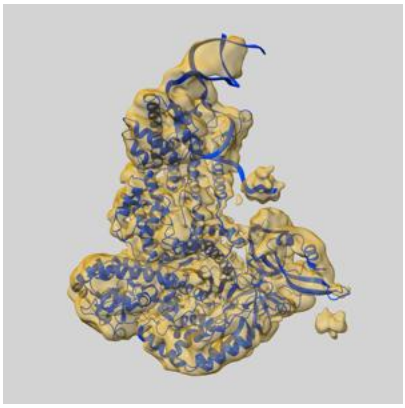
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.67	-	-
Author-provided FSC curve	3.67	4.31	3.71
Unmasked-calculated*	4.32	7.41	4.38

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 4.32 differs from the reported value 3.67 by more than 10 %

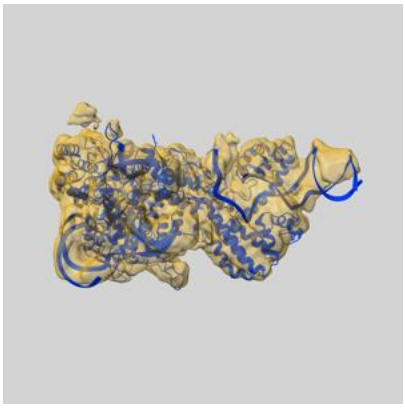
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-44111 and PDB model 9B2K. Per-residue inclusion information can be found in section 3 on page 5.

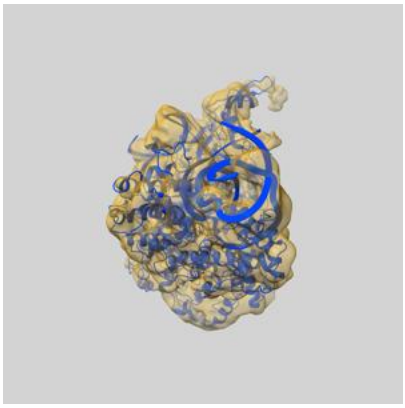
9.1 Map-model overlay



X



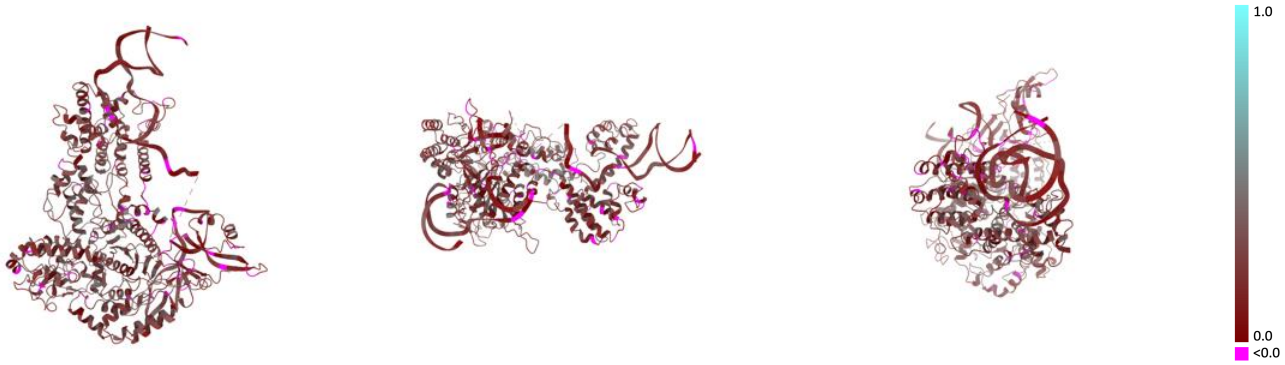
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.1 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.

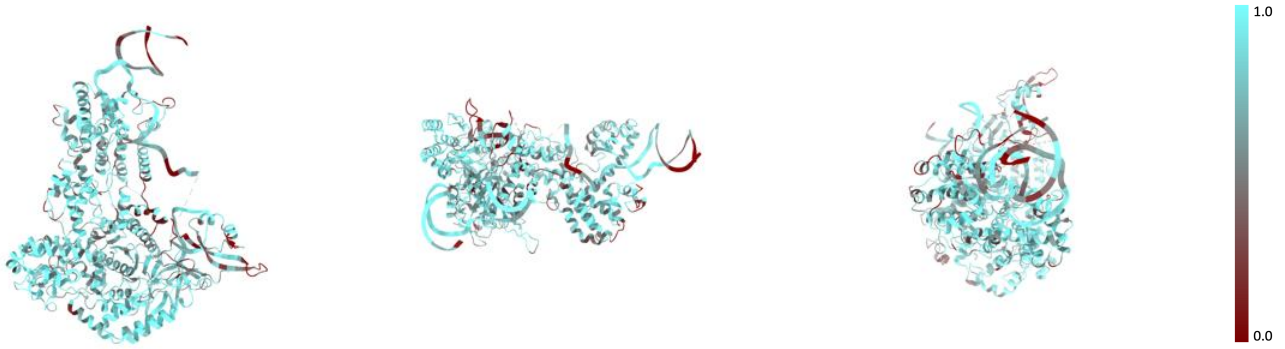
9.2 Q-score mapped to coordinate model



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.

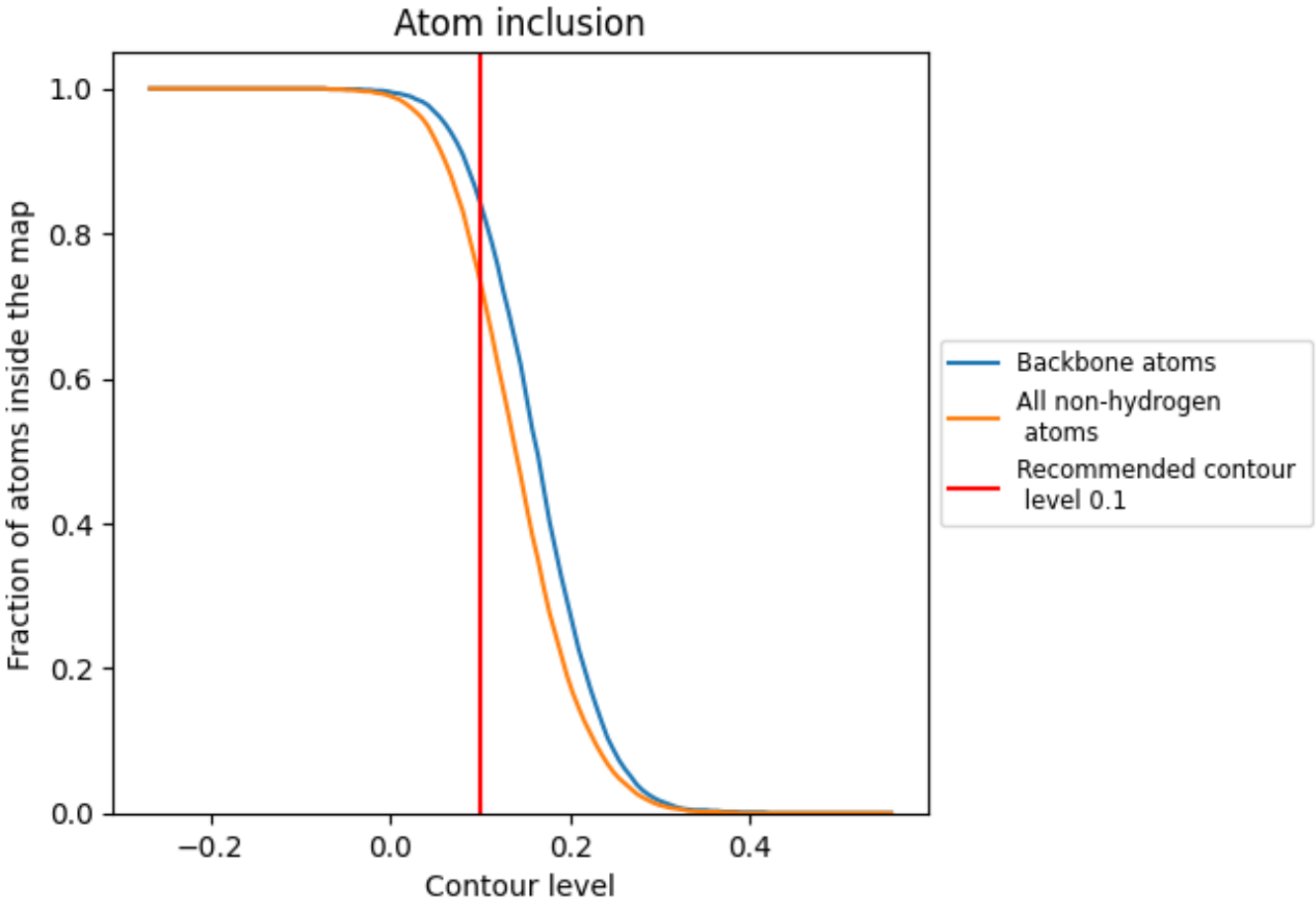
9.3 Atom inclusion mapped to coordinate model



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 (0.1).

9.4 Atom inclusion (i)



At the recommended contour level, 84% of all backbone atoms, 73% 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 (0.1) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7340	0.2160
A	0.6400	0.1900
B	0.7110	0.1480
P	0.7420	0.2270