wwPDB EM Validation Summary Report

Nov 13, 2023 – 03:24 am GMT

PDB ID	:	8PBY
EMDB ID	:	EMD-17588
Title	:	Single particle cryo-EM of the P140-P110 heterodimer with an alternative conformation in the P140 stalk of Mycoplasma genitalium at a resolution of 3.7 Angstrom.
Authors	:	Sprankel, L.; Scheffer, M.P.; Frangakis, A.S.
Deposited on	:	2023-06-09
Resolution	:	3.70 Å(reported)
Based on initial models	:	6R3T, 6RUT

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 (i) 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 (i)) were used in the production of this report:

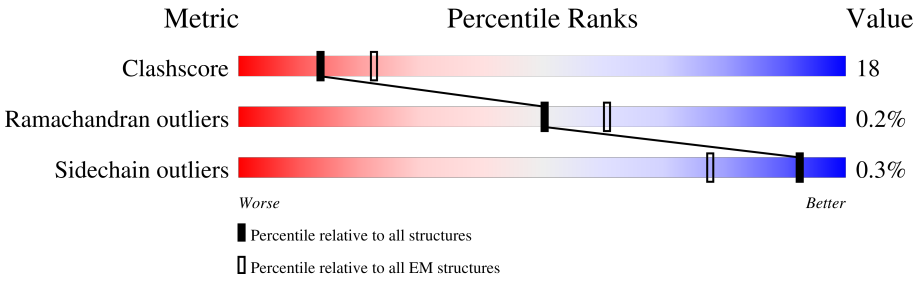
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.70 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	1059	14% 58% 27% 15%
2	B	1444	16% 54% 34% 11%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 33480 atoms, of which 16474 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 Mgp-operon protein 3.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
1	A	902	13658	4350	6714	1165	1423	6	2	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1053	HIS	-	expression tag	UNP P22747
A	1054	HIS	-	expression tag	UNP P22747
A	1055	HIS	-	expression tag	UNP P22747
A	1056	HIS	-	expression tag	UNP P22747
A	1057	HIS	-	expression tag	UNP P22747
A	1058	HIS	-	expression tag	UNP P22747

- Molecule 2 is a protein called Adhesin P1.

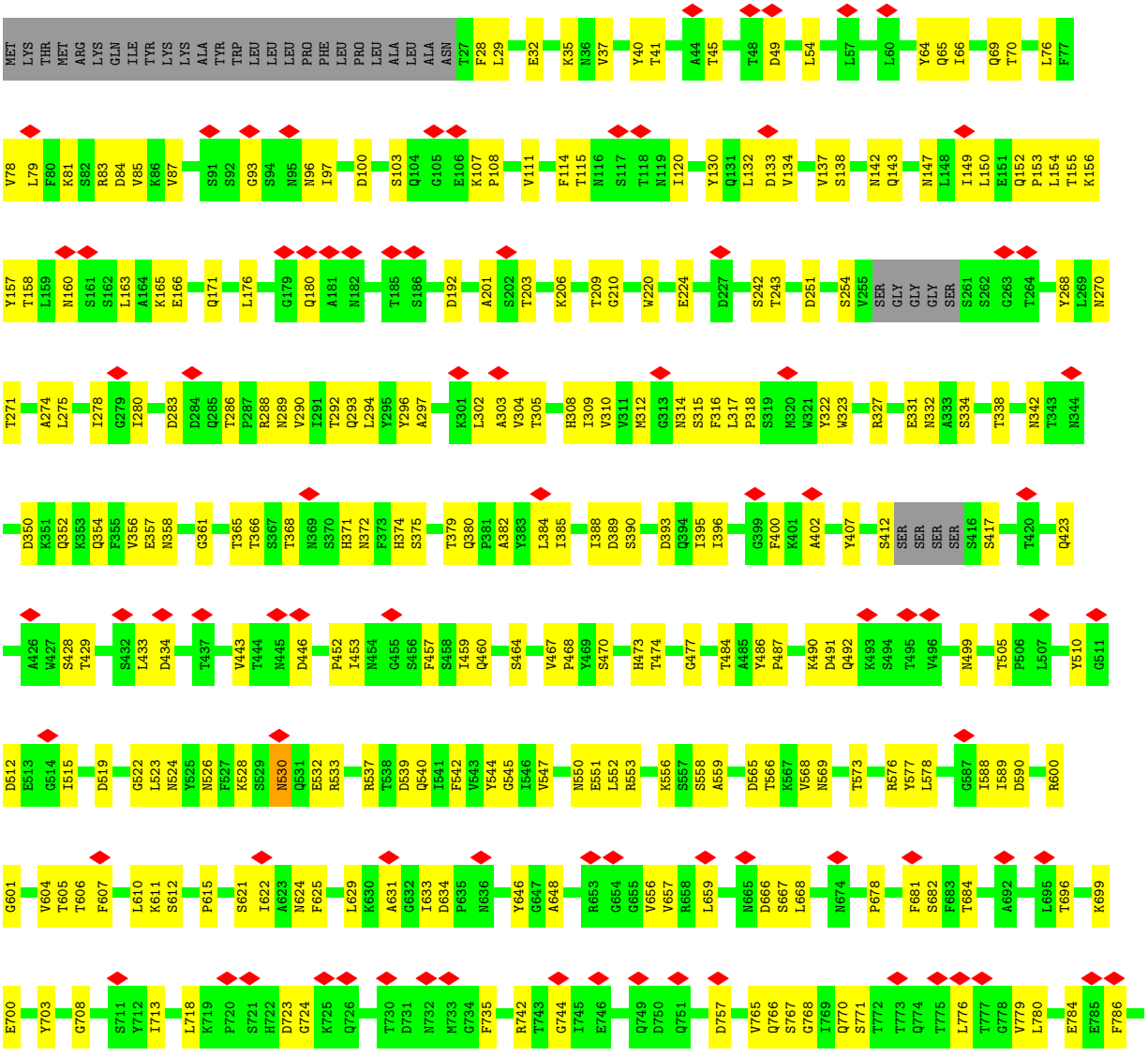
Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
2	B	1285	19822	6373	9760	1701	1974	14	0	0

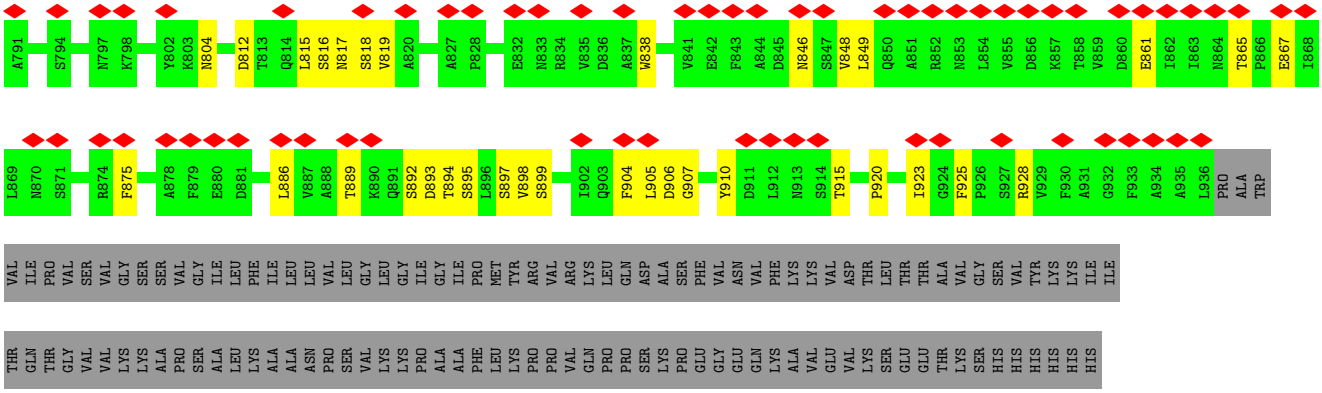
3 Residue-property plots (i)

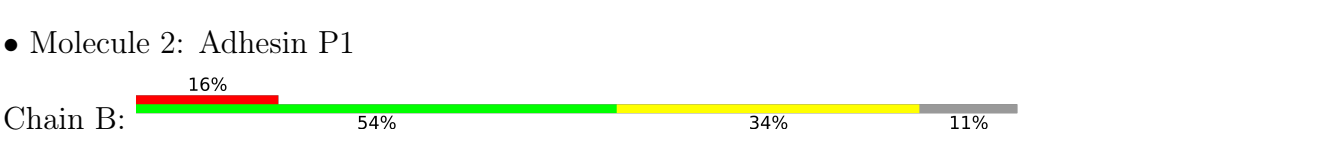
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

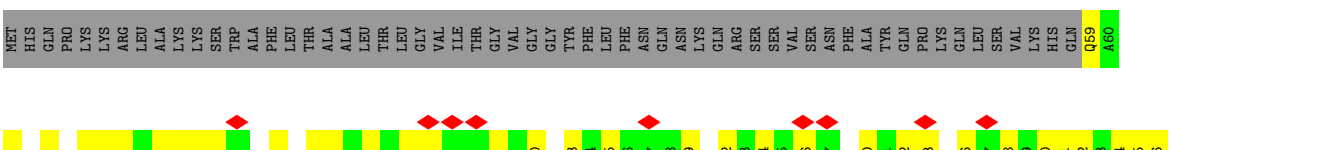
- Molecule 1: Mgp-operon protein 3

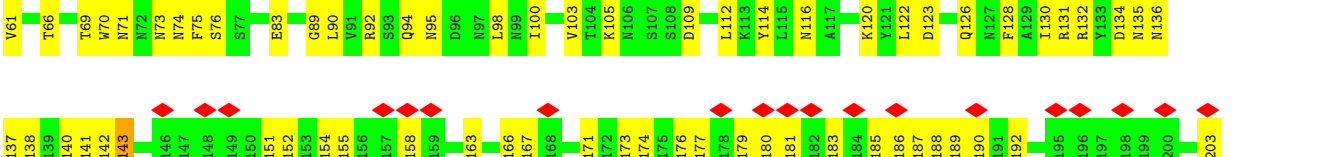
Chain A:

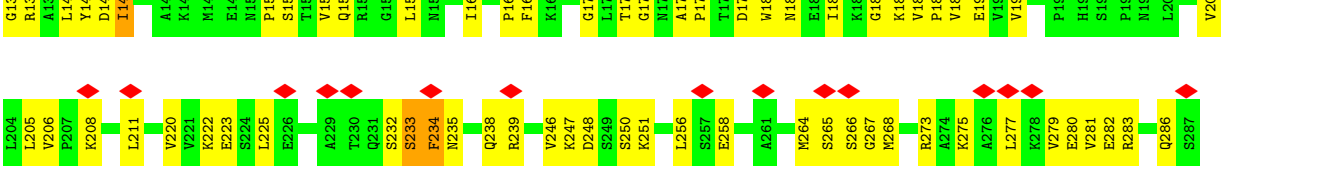






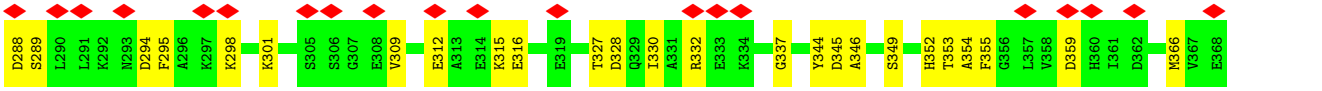


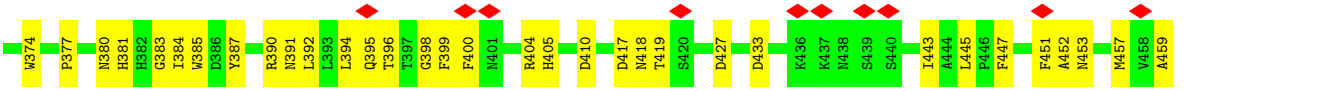


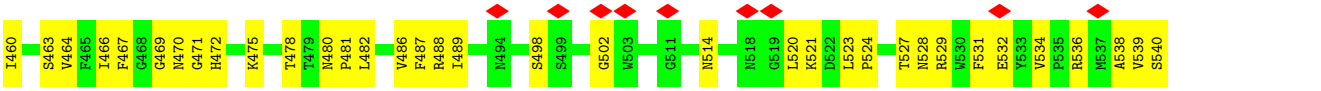


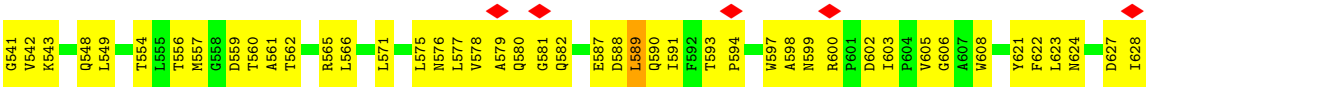
• Molecule 2: Adhesin P1

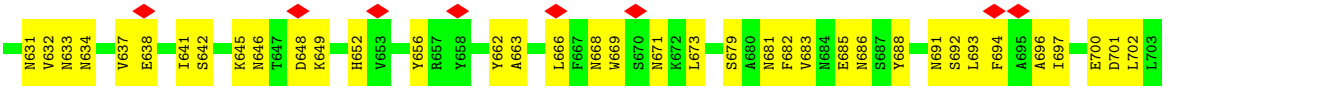
Chain B:

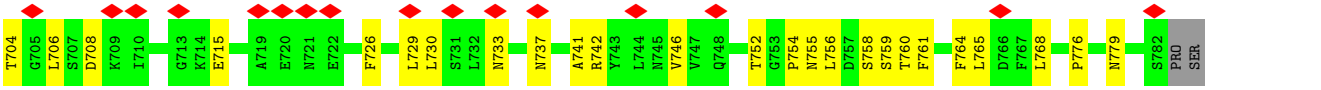


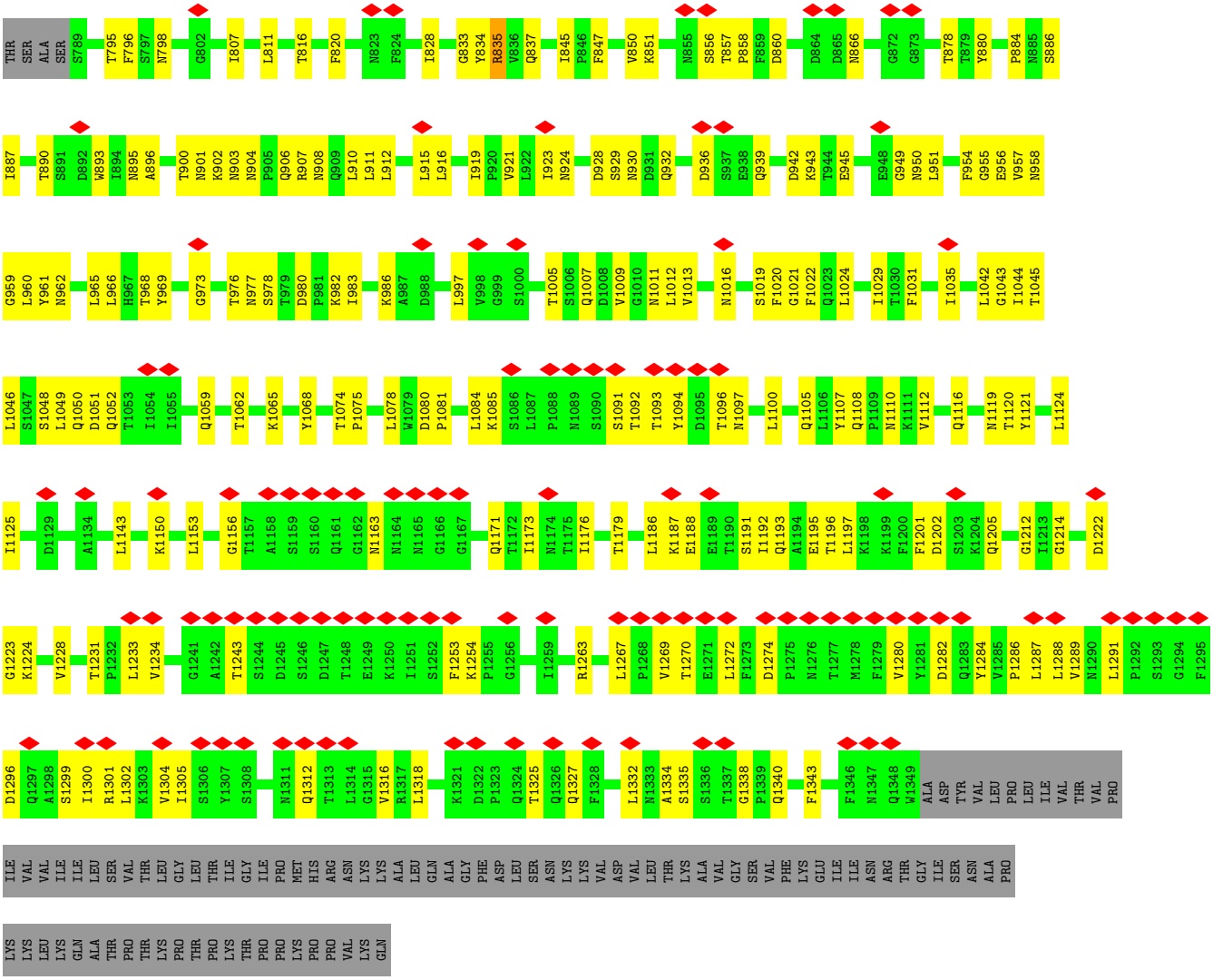











THR SER ALA SER S789 T795 F796 S797 N798 G802 I807 L811 T816 F820 N823 F824 I828 G833 Y834 R835 V836 Q837 I845 P846 F847 V850 K851 N855 S856 T857 P858 F859 D860 D864 D865 N866 G872 G873 T878 T879 Y880 P884 N885 S886
I887 T890 S891 D892 W893 I894 N895 A896 T900 N901 K902 N903 N904 P905 Q906 R907 N908 Q909 L910 L911 L912 L915 L916 I919 P920 V921 L922 I923 N924 D928 S929 N930 D931 Q932 D936 S937 E938 Q939 D942 K943 T944 E945 E948 G949 N950 L951 F954 G955 E956 V957 N958
G959 L960 Y961 N962 L965 L966 H967 T968 Y969 G973 T976 N977 S978 T979 D980 P981 K982 I983 K986 A987 D988 L997 V998 G999 S1000 T1005 S1006 Q1007 D1008 V1009 G1010 N1011 L1012 V1013 N1016 S1019 F1020 G1021 F1022 Q1023 L1024 I1029 T1030 F1031 I1035 L1042 G1043 I1044 T1045
L1046 S1047 S1048 L1049 Q1050 D1051 Q1052 T1053 I1054 I1055 Q1059 T1062 K1065 Y1068 T1074 P1075 L1078 W1079 D1080 P1081 L1084 K1085 S1086 L1087 P1088 N1089 S1090 S1091 T1092 T1093 Y1094 D1095 T1096 N1097 L1100 Q1105 L1106 Y1107 Q1108 P1109 N1110 K1111 V1112 Q1116 N1119 T1120 Y1121 L1124
I1125 D1129 A1134 L1143 K1150 L1153 G1156 T1157 A1158 S1159 S1160 Q1161 G1162 N1163 N1164 N1165 G1166 G1167 Q1171 T1172 I1173 N1174 T1175 I1176 T1179 L1186 K1187 E1188 E1189 T1190 S1191 I1192 Q1193 A1194 E1195 T1196 L1197 K1198 K1199 F1200 F1201 D1202 S1203 K1204 Q1205 G1212 I1213 G1214 D1222
G1223 K1224 V1228 T1231 P1232 L1233 V1234 G1241 A1242 T1243 S1244 D1245 S1246 D1247 T1248 E1249 K1250 I1251 S1252 F1253 K1254 P1255 G1256 I1259 R1263 L1267 P1268 V1269 T1270 E1271 L1272 F1273 D1274 P1275 N1276 T1277 M1278 F1279 V1280 Y1281 D1282 Q1283 Y1284 V1285 P1286 L1287 L1288 V1289 N1290 L1291 P1292 S1293 G1294 F1295
D1296 Q1297 A1298 S1299 I1300 R1301 L1302 K1303 V1304 I1305 S1306 Y1307 S1308 N1311 Q1312 T1313 L1314 G1315 V1316 R1317 L1318 K1321 D1322 P1323 Q1324 T1325 Q1326 Q1327 F1328 L1332 N1333 A1334 S1335 S1336 T1337 G1338 P1339 Q1340 F1343 F1346 N1347 Q1348 W1349 ALA ASP TYR VAL LEU PRO LEU ILE VAL THR VAL PRO
ILE VAL VAL ILE ILE LEU SER VAL THR LEU GLY LEU THR ILE GLY ILE PRO MET HIS ARG ASN LYS LYS ALA LEU GLN ALA GLY PHE ASP LEU SER ASN LYS LYS VAL ASP VAL LEU THR LYS ALA VAL GLY SER VAL PHE LYS GLU ILE ILE ASN ARG THR GLY ILE SER ASN ALA PRO
LYS LYS LEU LYS GLN ALA THR PRO THR LYS PRO THR PRO LYS THR PRO PRO PRO LYS PRO PRO VAL LYS GLN

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	67091	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 ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	1.945	Depositor
Minimum map value	-1.485	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.062	Depositor
Recommended contour level	0.34	Depositor
Map size (Å)	315.0, 315.0, 315.0	wwPDB
Map dimensions	300, 300, 300	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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	0/7076	0.47	0/9624
2	B	0.26	0/10316	0.48	0/14045
All	All	0.26	0/17392	0.47	0/23669

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
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	589	LEU	Peptide

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	6944	6714	6705	227	0
2	B	10062	9760	9759	383	0
All	All	17006	16474	16464	603	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 18.

The worst 5 of 603 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:530:ASN:O	1:A:530:ASN:ND2	2.03	0.91
1:A:165:LYS:NZ	1:A:206:LYS:O	2.05	0.90
2:B:387:TYR:OH	2:B:418:ASN:OD1	1.89	0.89
2:B:273:ARG:NH2	2:B:878:THR:OG1	2.07	0.88
2:B:391:ASN:OD1	2:B:418:ASN:ND2	2.08	0.87

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	896/1059 (85%)	839 (94%)	56 (6%)	1 (0%)	51 83
2	B	1281/1444 (89%)	1151 (90%)	126 (10%)	4 (0%)	41 74
All	All	2177/2503 (87%)	1990 (91%)	182 (8%)	5 (0%)	50 78

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	233	SER
2	B	234	PHE
2	B	1300	ILE
1	A	120	ILE
2	B	143	ILE

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	786/920 (85%)	784 (100%)	2 (0%)	92 96
2	B	1127/1267 (89%)	1124 (100%)	3 (0%)	92 96
All	All	1913/2187 (88%)	1908 (100%)	5 (0%)	92 96

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	530	ASN
1	A	556	LYS
2	B	120	LYS
2	B	835	ARG
2	B	1263	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	631	ASN
2	B	668	ASN
2	B	737	ASN
2	B	735	ASN
1	A	358	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)

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.

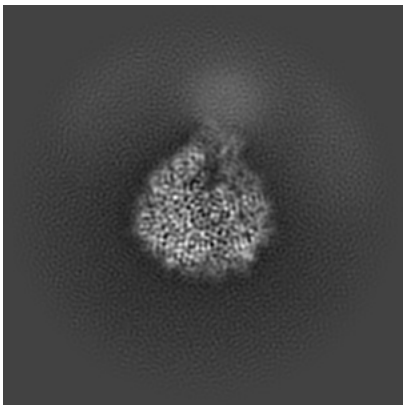
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-17588. These allow visual inspection of the internal detail of the map and identification of artifacts.

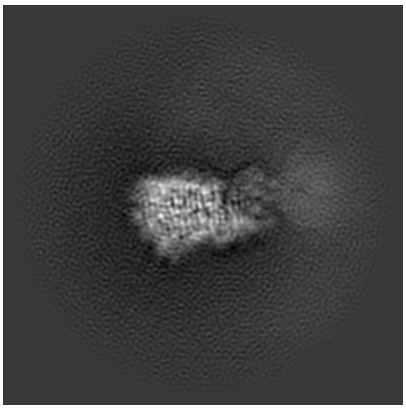
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

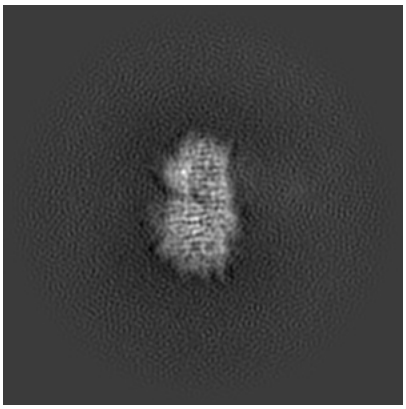
6.1.1 Primary map



X

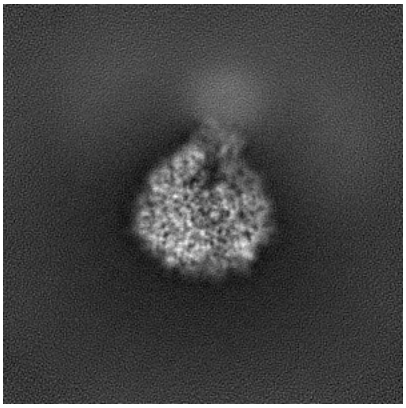


Y

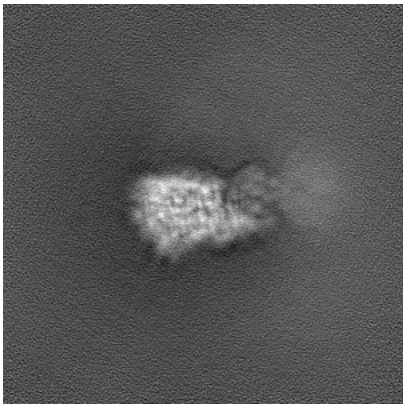


Z

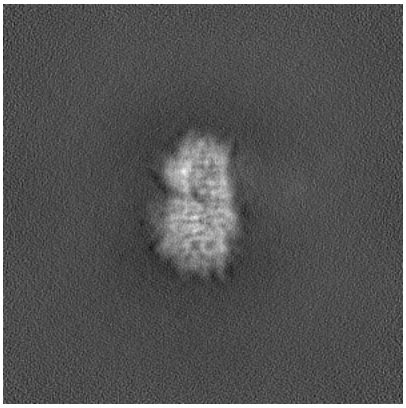
6.1.2 Raw map



X



Y

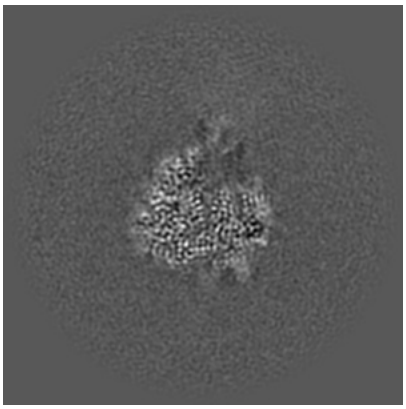


Z

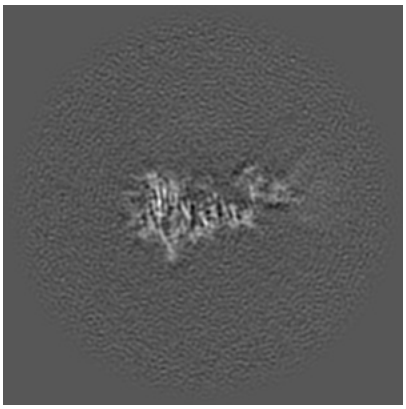
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

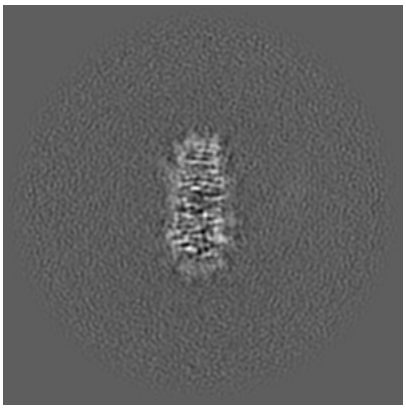
6.2.1 Primary map



X Index: 150

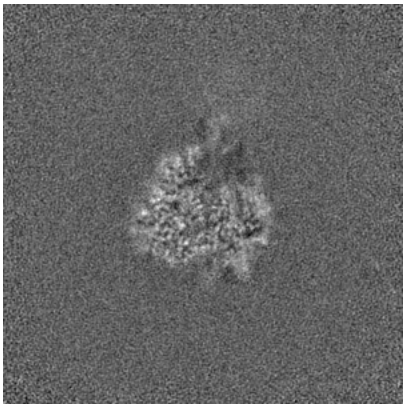


Y Index: 150

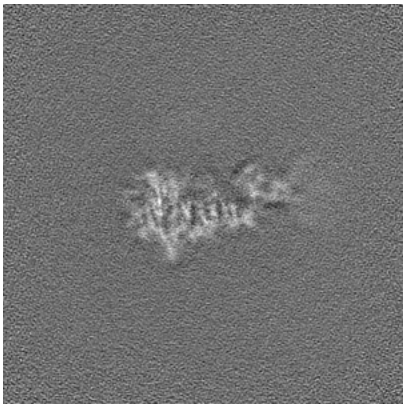


Z Index: 150

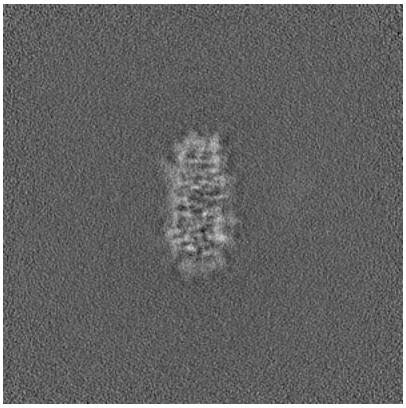
6.2.2 Raw map



X Index: 150



Y Index: 150

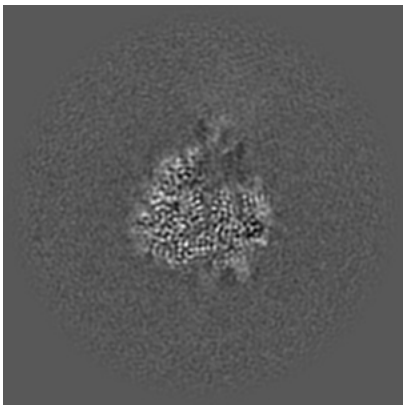


Z Index: 150

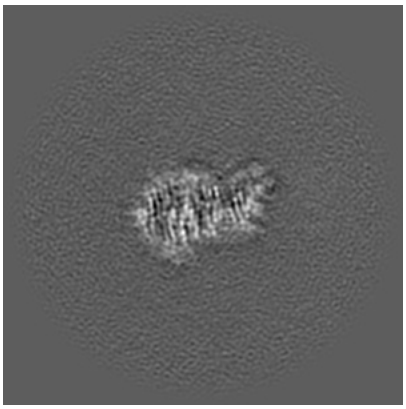
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

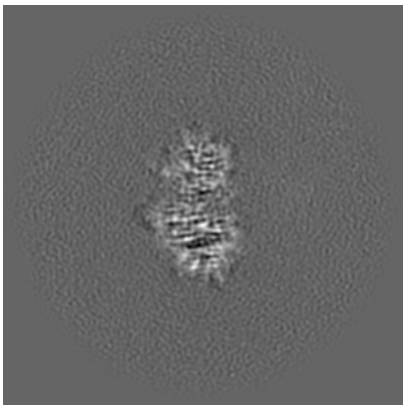
6.3.1 Primary map



X Index: 150

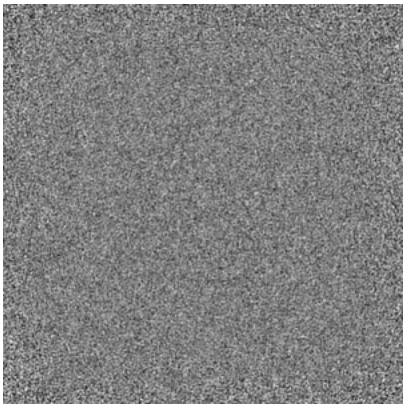


Y Index: 139

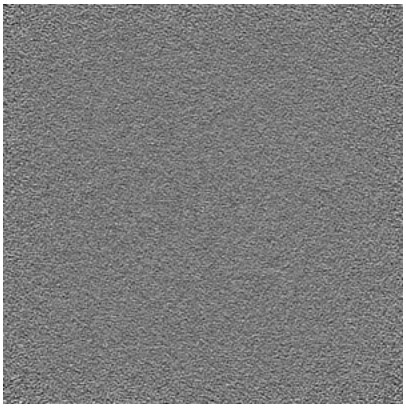


Z Index: 134

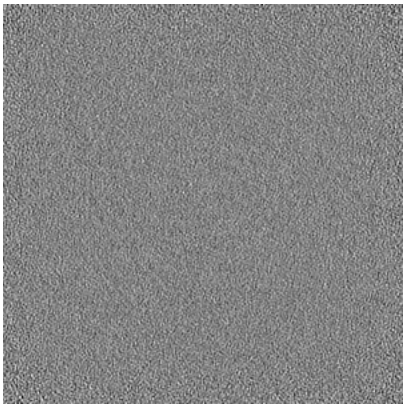
6.3.2 Raw map



X Index: 0



Y Index: 0

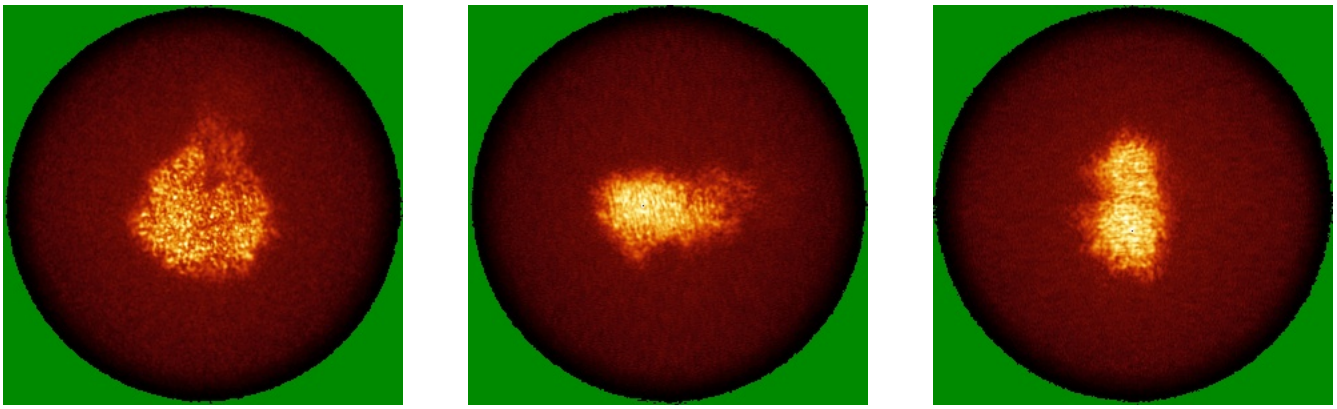


Z Index: 0

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

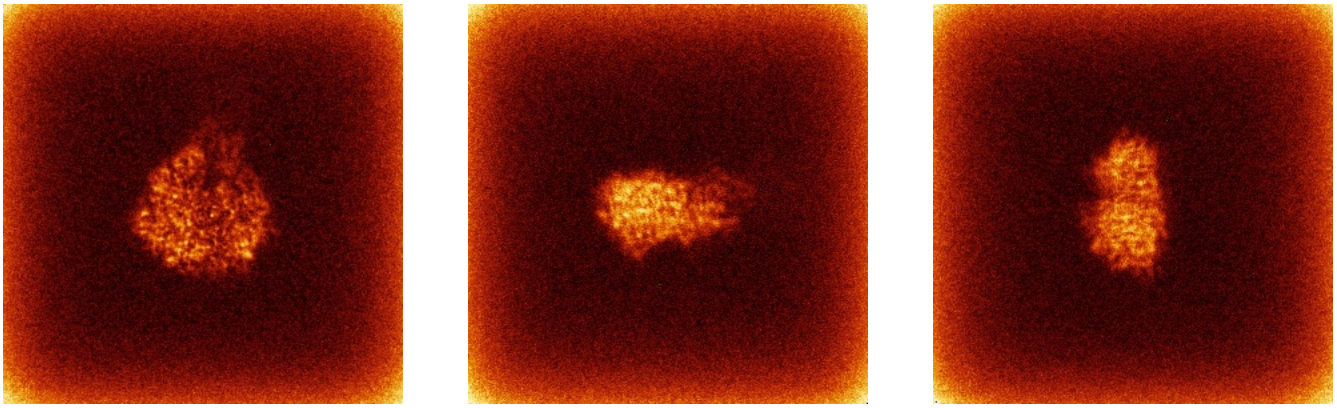
6.4 Orthogonal standard-deviation projections (False-color) (i)

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 0.34. 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.

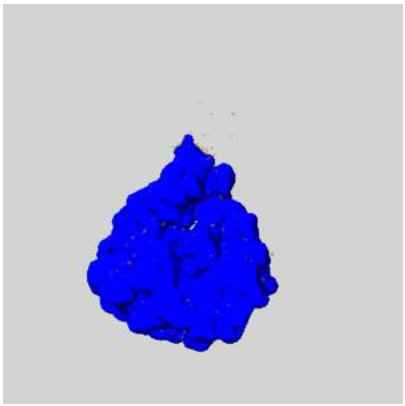
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

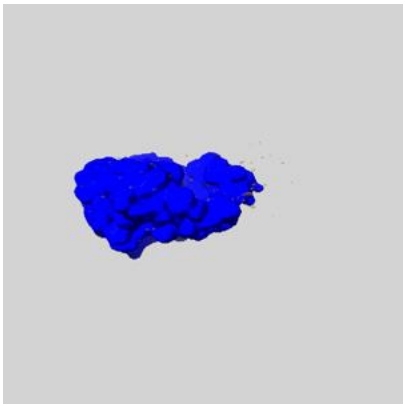
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

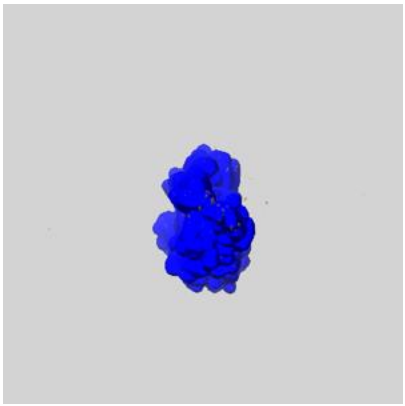
6.6.1 emd_17588_msk_1.map (i)



X



Y

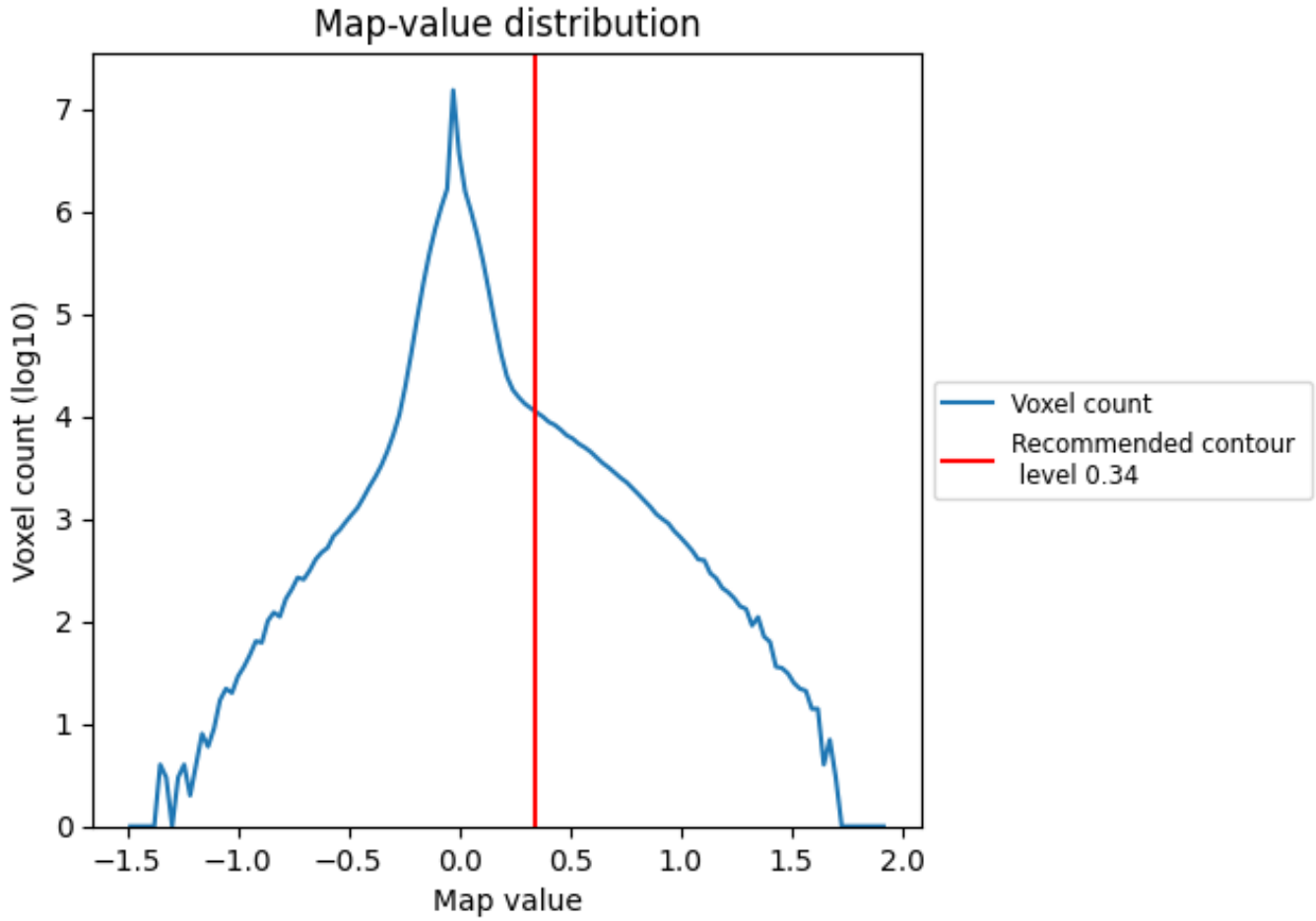


Z

7 Map analysis

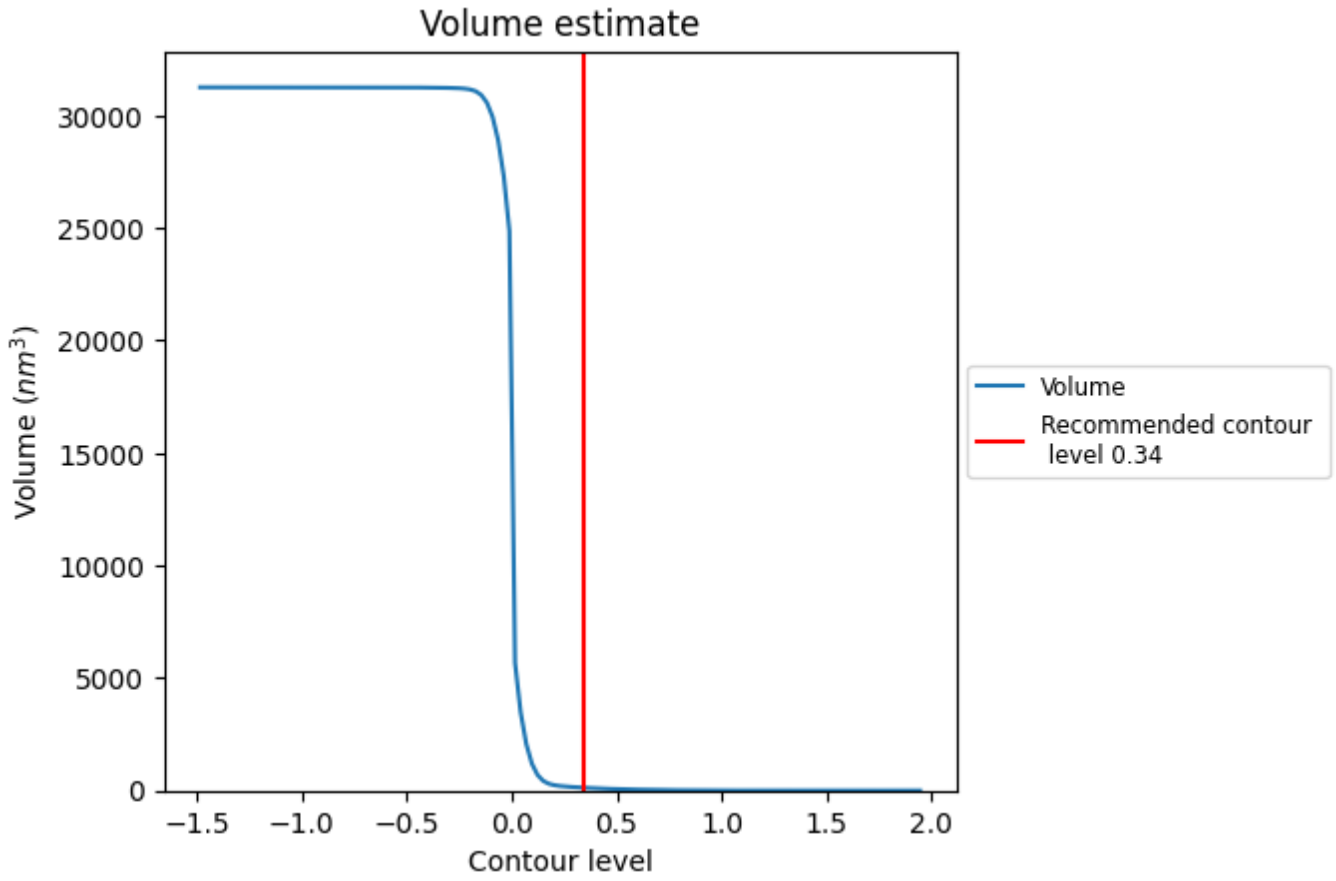
This section contains the results of statistical analysis of the map.

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.

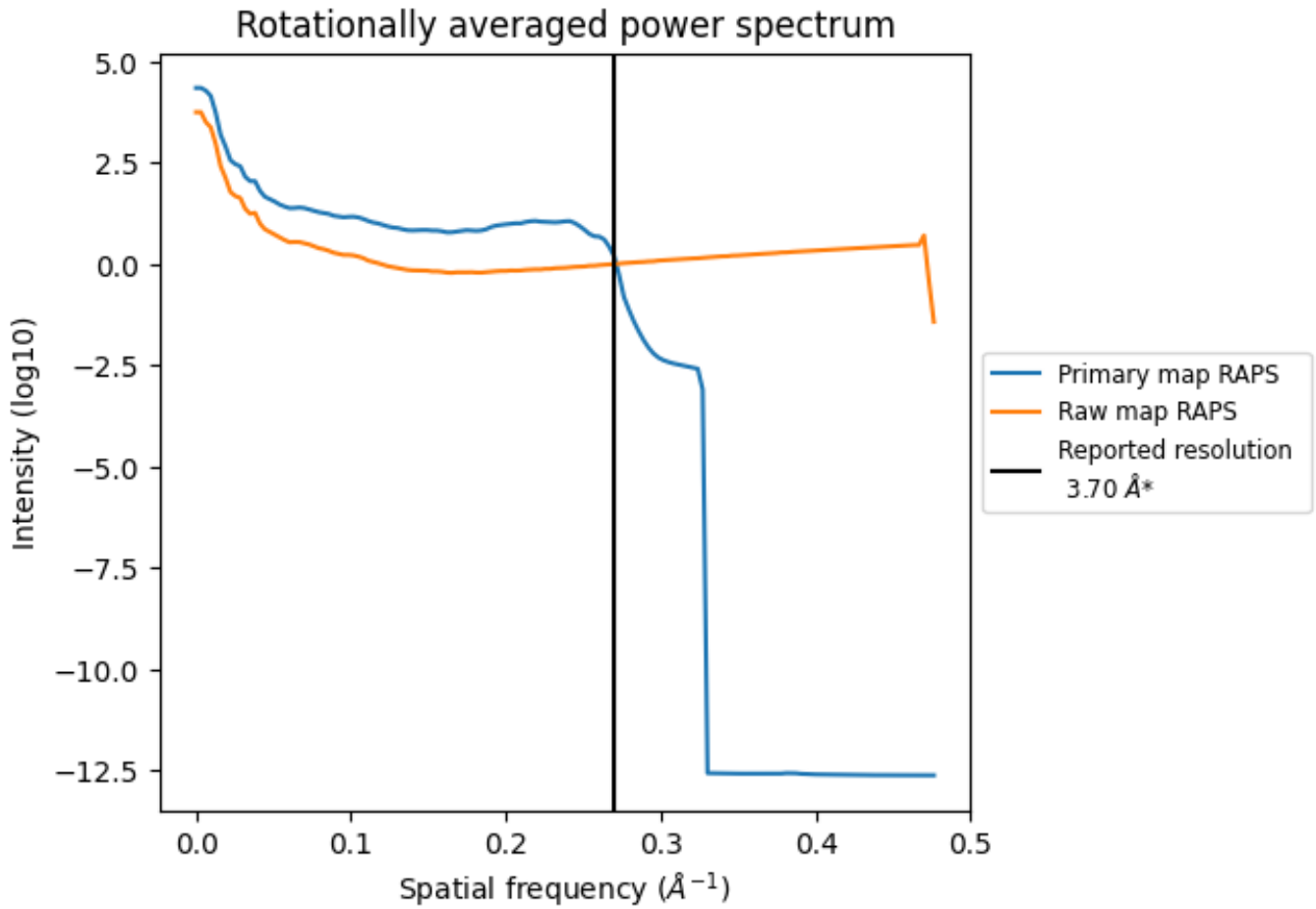
7.2 Volume estimate (i)



The volume at the recommended contour level is 128 nm^3; this corresponds to an approximate mass of 116 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)

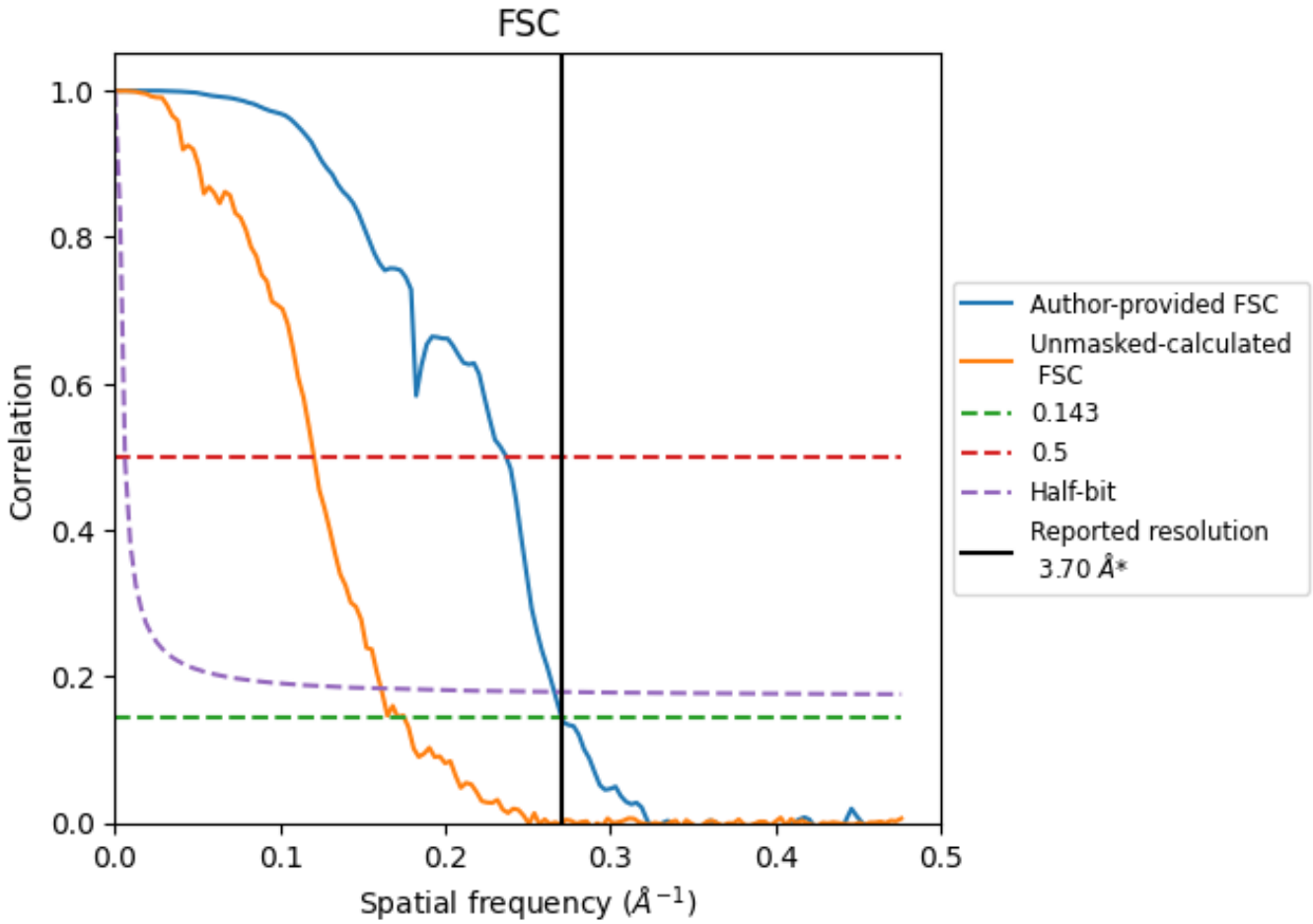


*Reported resolution corresponds to spatial frequency of 0.270 $Å^{-1}$

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 0.270 Å^{-1}

8.2 Resolution estimates (i)

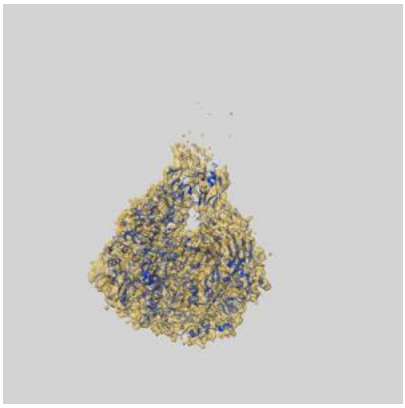
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.70	-	-
Author-provided FSC curve	3.69	4.22	3.76
Unmasked-calculated*	5.70	8.29	6.20

*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 5.70 differs from the reported value 3.7 by more than 10 %

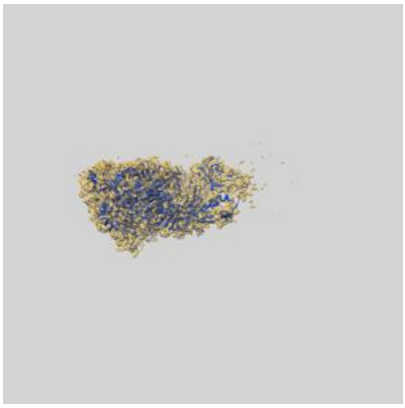
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-17588 and PDB model 8PBY. Per-residue inclusion information can be found in section 3 on page 4.

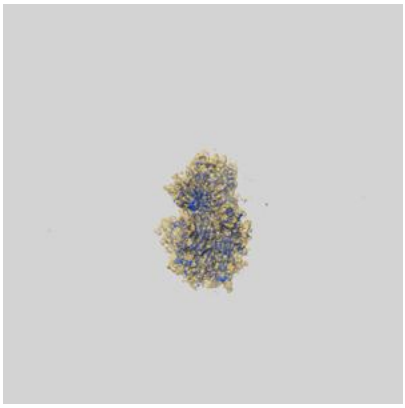
9.1 Map-model overlay



X



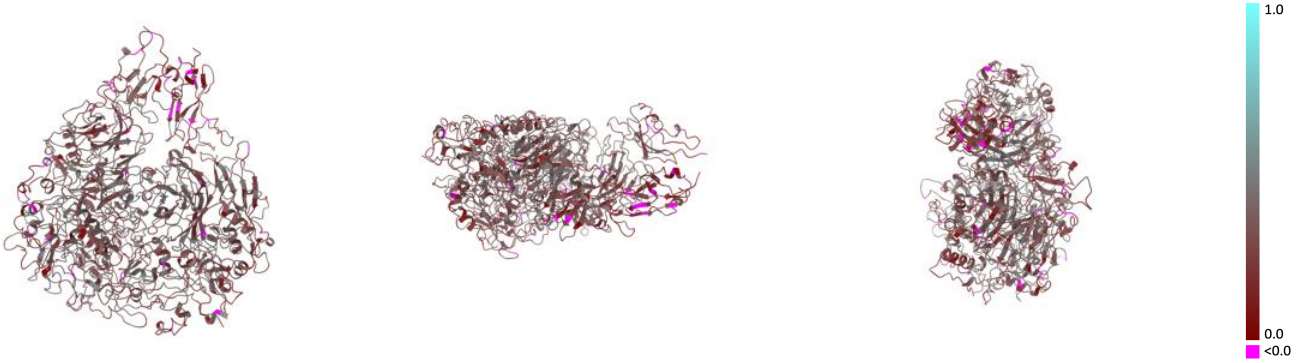
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.34 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.

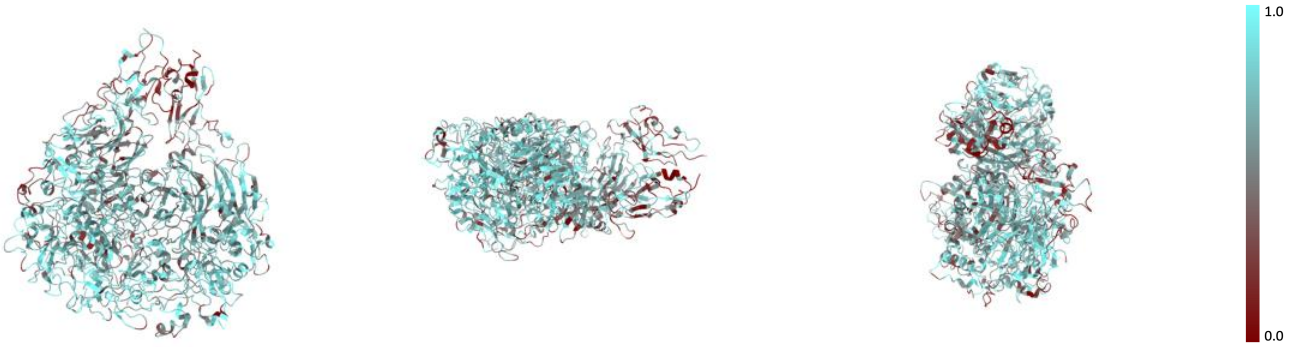
9.2 Q-score mapped to coordinate model



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.

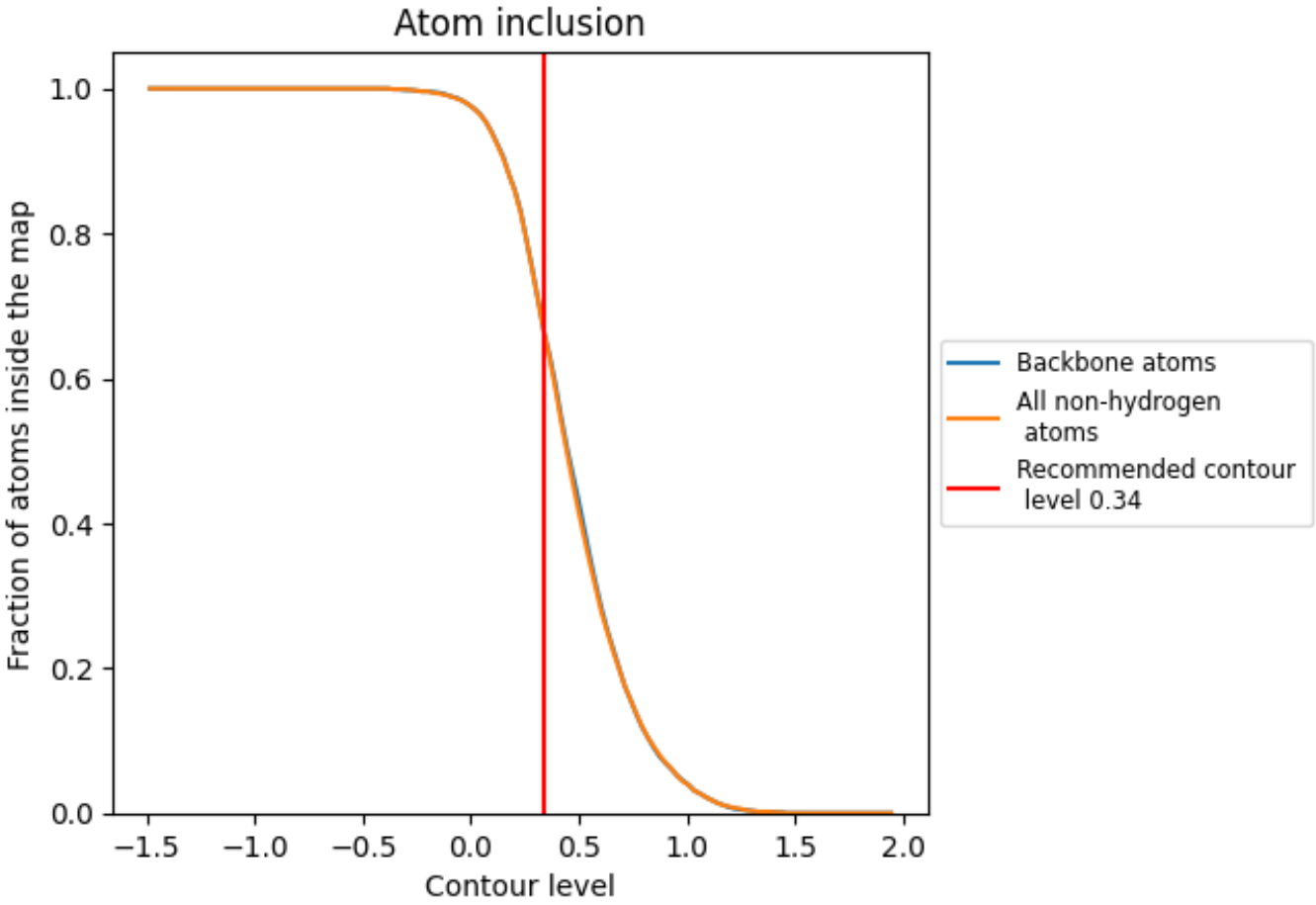
9.3 Atom inclusion mapped to coordinate model



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 (0.34).

9.4 Atom inclusion (i)



At the recommended contour level, 66% of all backbone atoms, 66% 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 (0.34) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6610	0.3200
A	0.6610	0.3130
B	0.6650	0.3240