wwPDB EM Validation Summary Report (i)

Mar 3, 2024 – 09:59 AM EST

PDB ID	:	6C6T
EMDB ID	:	EMD-7350
Title	:	CryoEM structure of E.coli RNA polymerase elongation complex bound with RfaH
Authors	:	Kang, J.Y.; Artsimovitch, I.; Landick, R.; Darst, S.A.
Deposited on	:	2018-01-19
Resolution	:	3.50 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 (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:

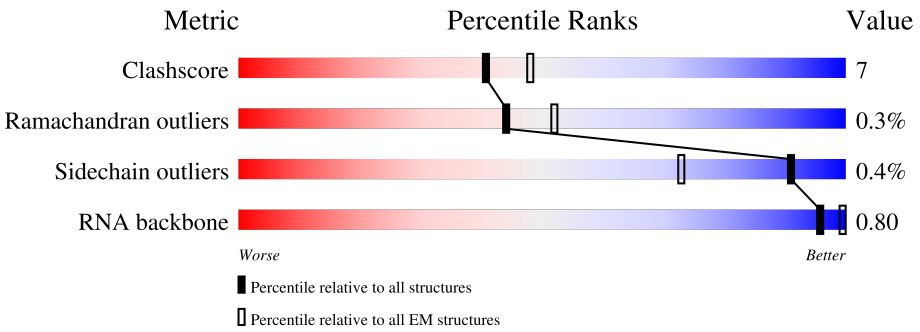
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.50 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 >=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	29	• 62% 34% •
2	B	29	83% 17%
3	R	20	30% 20% 50%
4	G	239	79% 13% 8%
4	H	239	66% 26% 8%
5	I	1342	6% 80% 18% •
6	J	1407	9% 74% 20% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	K	91	74% 18% 9%
8	D	162	41% 11% 48%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 26915 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 DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	A	29	598	282	117	171	28	0	0

- Molecule 2 is a DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	B	29	585	279	102	176	28	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	R	10	218	97	43	68	10	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	G	221	1694	1061	296	331	6	0	0
4	H	219	1687	1053	298	330	6	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	235	GLU	-	expression tag	UNP P0A7Z4
G	236	VAL	-	expression tag	UNP P0A7Z4
G	237	LEU	-	expression tag	UNP P0A7Z4
G	238	PHE	-	expression tag	UNP P0A7Z4
G	239	GLN	-	expression tag	UNP P0A7Z4
H	235	GLU	-	expression tag	UNP P0A7Z4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	236	VAL	-	expression tag	UNP P0A7Z4
H	237	LEU	-	expression tag	UNP P0A7Z4
H	238	PHE	-	expression tag	UNP P0A7Z4
H	239	GLN	-	expression tag	UNP P0A7Z4

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	I	1319	10407	6530	1814	2020	43	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	J	1335	10388	6526	1854	1958	50	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	K	83	655	399	123	132	1	0	0

- Molecule 8 is a protein called Transcription antitermination protein RfaH.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	D	85	680	438	119	120	3	0	0

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
9	R	1	1	1	0

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
10	J	2	2	2	0

3 Residue-property plots (i)

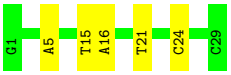
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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 (29-MER)

Chain A:

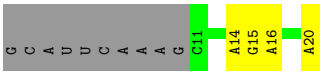
• Molecule 2: DNA (29-MER)

Chain B:



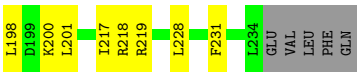
• Molecule 3: RNA (5'-R(*GP*CP*AP*UP*UP*CP*AP*AP*AP*GP*CP*CP*GP*AP*GP*AP*GP*GP*UP*A)-3')

Chain R:



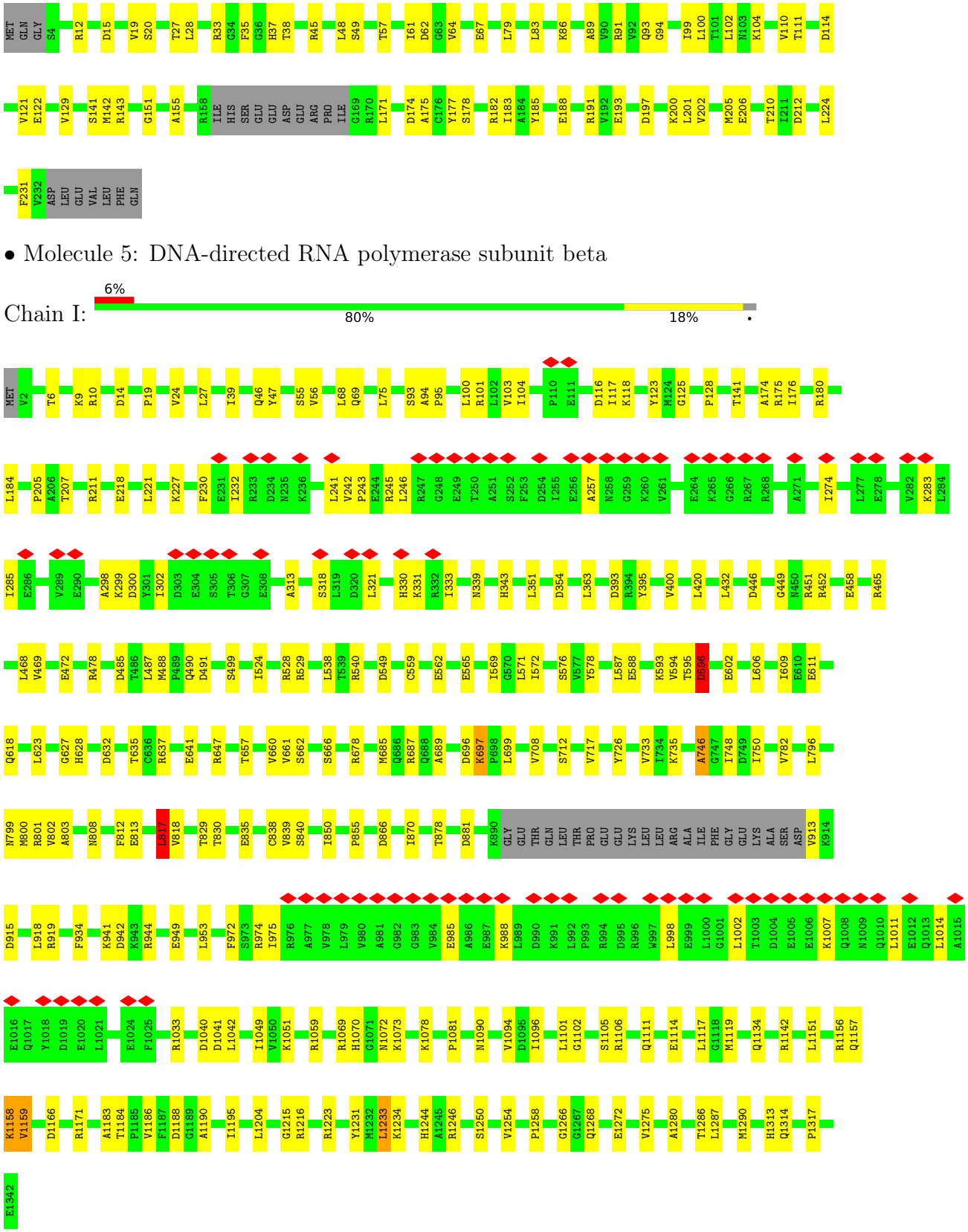
• Molecule 4: DNA-directed RNA polymerase subunit alpha

Chain G:



• Molecule 4: DNA-directed RNA polymerase subunit alpha

Chain H:



• Molecule 6: DNA-directed RNA polymerase subunit beta'

Chain J: 9% 74% 20% 5%

MET LYS ASP LEU LEU LYS PHE LEU LYS ALA GLN THR LYS THR GLU E16 K40 P41 E42 T43 I44 N45 Y46 K50 R53 C58 A59 R60 C70 L71 C72 G73 K76 R77 L78 K79 G82 C85 E86 V92 V97 R98 R99 E100 R101 I114 L120 P121

S122 R123 L126 L127 L132 R137 V138 L139 Y140 F141 E142 S143 Y144 R156 D174 G181 S191 L201 N209 R214 T218 K219 R220 I221 K222 L223 L224 S230 G231 P246 L249 R275 R278 L279 I291 V292 R293 L299 D304 R312

G313 S319 N320 L342 L343 G351 R352 L363 H364 K370 R388 T393 I394 K395 A396 L412 D413 E414 V415 I416 R417 P420 V421 R425 A426 A436 F437 L441 V453 N458 F461 A467 M485 L491 S503 V506 V507 L508 G509 L510

M513 T514 R515 D516 V526 K531 E532 A533 E534 L544 K549 V550 D558 A559 E562 S568 L569 K570 R576 L579 G586 N593 L596 I601 L605 Y609 R610 I611 L612 P616 I619 G628 A633 S638 V639 P647 Q665

R678 I683 M697 R709 V717 S718 S721 M724 R731 A741 L746 K749 P750 I754 I755 P758 I759 N762 L770 Q771 I774 H777 D785 K789 Y795 L796 T797 R798 R799 L800 V801 D802 V803 A804 Q805 D806 V809 C814

I820 P824 E827 D830 V831 K832 E833 P834 L835 R836 V839 R842 D847 V848 L849 K850 D855 I856 L857 V858 L864 R883 S884 D891 V894 D902 L903 A904 R905 G906 H907 I908 I909 I918 A919 A920 G924 E925 P926 L930 R933 THR PHE

HIS ILE GLY GLY ALA ALA SER ARG ALA ALA ALA GLU S948 S949 I950 Q951 V952 K953 N954 K955 K959 L960 S961 N962 V963 K964 S965 V966 V967 N968 S969 S970 G971 K972 L973 V974 I975 T976 S977 R978 M979 L984 I985 D986 E987 F988 E993 Y999 L1003 A1004 K1005 G1006 D1007 Q1010

G1014 E1015 A1018 T1024 M1025 P1026 V1027 I1028 T1029 E1030 V1031 S1032 G1033 F1034 V1035 R1036 F1037 T1038 D1039 M1040 I1041 D1042 G1043 Q1044 T1045 I1046 T1047 R1048 Q1049 T1050 D1051 E1052 L1053 T1054 G1055 L1056 S1057 S1058 L1059 V1060 V1061 L1062 D1063 S1064 A1065 E1066 R1067 T1068 A1069 G1070 G1071 K1072 D1073 L1074 R1075 P1076 A1077

L1078 K1079 I1080 V1081 D1082 A1083 Q1084 G1085 N1086 D1087 V1088 L1089 I1090 P1091 G1092 T1093 D1094 M1095 P1096 A1097 Q1098 Y1099 F1100 L1101 P1102 G1103 K1104 A1105 I1106 V1107 Q1108 L1109 E1110 D1111 G1112 V1113 G1114 I1115 S1116 S1117 G1118 L1121 A1122 R1123 I1124 P1125 Q1126 GLU SER GLY GLY THR LYS ASP ILE THR G1136 D1143

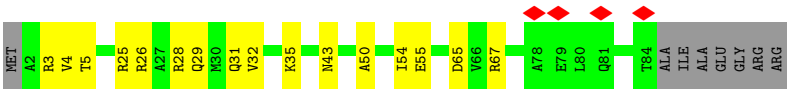
E1146 A1147 R1148 R1149 P1150 I1155 E1158 G1161 G1166 K1167 E1168 T1169 K1170 G1171 K1172 R1173 R1174 L1175 P1179 V1180 D1181 G1182 S1183 D1184 P1185 I1190 P1191 K1192 N1197 R1203 V1204 E1205 I1210 G1213 P1214 D1219 I1220 L1221 R1222 L1223 H1227 T1230 Y1241 R1242 L1243

Q1244 I1248 N1249 D1250 R1258 L1261 N1268 S1272 D1273 N1289 R1290 E1293 R1304 D1305 L1306 T1310 S1321 S1324 F1325 Q1326 R1330 L1347 K1348 E1349 N1350 V1351 I1352 V1353 G1354 R1355 L1356 I1357 P1358 A1359 R1369 M1370 R1371 R1372 R1373 ALA ALA GLY GLU ALA PRO ALA

ALA PRO GLN VAL THR ALA GLU ASP ALA SER ALA SER LEU ALA GLU LEU LEU ASN ALA GLY LEU GLY GLY SER ASP ASN GLU

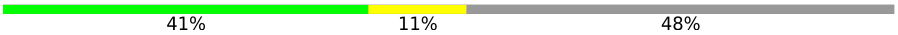
- Molecule 7: DNA-directed RNA polymerase subunit omega

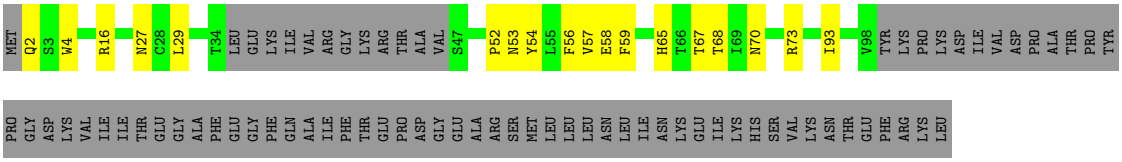
Chain K: 74% 18% 9%



- Molecule 8: Transcription antitermination protein RfaH

Chain D:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	174600	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 ($e^-/Å^2$)	71.0	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	38462	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.185	Depositor
Minimum map value	-0.106	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.016	Depositor
Map size (Å)	332.8, 332.8, 332.8	wwPDB
Map dimensions	256, 256, 256	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: ZN, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >5$	RMSZ	$\#\|Z\| >5$
1	A	0.92	0/672	1.00	2/1037 (0.2%)
2	B	1.03	0/653	1.03	0/1005
3	R	0.90	0/244	0.98	0/379
4	G	0.48	0/1714	0.67	1/2323 (0.0%)
4	H	0.41	0/1706	0.68	0/2311
5	I	0.50	0/10573	0.64	2/14265 (0.0%)
6	J	0.48	0/10545	0.65	2/14236 (0.0%)
7	K	0.41	0/657	0.66	0/886
8	D	0.49	0/698	0.61	0/951
All	All	0.52	0/27462	0.68	7/37393 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
5	I	0	9
6	J	0	5
7	K	0	1
All	All	0	15

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	J	903	LEU	C-N-CA	5.86	136.36	121.70
4	G	61	ILE	CG1-CB-CG2	-5.65	98.96	111.40
6	J	363	LEU	CA-CB-CG	5.59	128.16	115.30
5	I	1233	LEU	CA-CB-CG	5.46	127.85	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	2	DG	O4'-C4'-C3'	-5.38	102.35	104.50

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	I	205	PRO	Peptide
5	I	594	VAL	Peptide
5	I	595	THR	Peptide
5	I	596	ASP	Mainchain
5	I	696	ASP	Peptide

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	598	0	325	8	0
2	B	585	0	328	5	0
3	R	218	0	109	3	0
4	G	1694	0	1722	21	0
4	H	1687	0	1722	40	0
5	I	10407	0	10420	157	0
6	J	10388	0	10611	182	0
7	K	655	0	663	15	0
8	D	680	0	671	10	0
9	R	1	0	0	0	0
10	J	2	0	0	0	0
All	All	26915	0	26571	396	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 7.

The worst 5 of 396 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)
6:J:425:ARG:NH1	6:J:458:ASN:O	2.08	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:K:3:ARG:NH1	7:K:55:GLU:OE2	2.09	0.84
6:J:665:GLN:HE22	6:J:678:ARG:HH11	1.25	0.84
5:I:1314:GLN:HB2	7:K:28:ARG:HH12	1.43	0.83
6:J:40:LYS:NZ	6:J:42:GLU:OE1	2.15	0.80

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	
4	G	217/239 (91%)	197 (91%)	19 (9%)	1 (0%)	29	68
4	H	215/239 (90%)	202 (94%)	13 (6%)	0	100	100
5	I	1315/1342 (98%)	1204 (92%)	106 (8%)	5 (0%)	34	72
6	J	1329/1407 (94%)	1216 (92%)	111 (8%)	2 (0%)	47	81
7	K	81/91 (89%)	74 (91%)	7 (9%)	0	100	100
8	D	81/162 (50%)	74 (91%)	5 (6%)	2 (2%)	5	34
All	All	3238/3480 (93%)	2967 (92%)	261 (8%)	10 (0%)	44	75

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	I	596	ASP
5	I	1158	LYS
6	J	805	GLN
8	D	53	ASN
6	J	904	ALA

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	
4	G	186/206 (90%)	186 (100%)	0	100	100
4	H	187/206 (91%)	186 (100%)	1 (0%)	88	94
5	I	1138/1157 (98%)	1135 (100%)	3 (0%)	92	97
6	J	1120/1168 (96%)	1113 (99%)	7 (1%)	86	94
7	K	70/75 (93%)	70 (100%)	0	100	100
8	D	75/142 (53%)	75 (100%)	0	100	100
All	All	2776/2954 (94%)	2765 (100%)	11 (0%)	91	96

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

Mol	Chain	Res	Type
6	J	709	ARG
6	J	836	ARG
6	J	1373	ARG
6	J	1268	ASN
6	J	53	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
6	J	1108	GLN
6	J	1279	GLN
6	J	1268	ASN
6	J	1326	GLN
5	I	1157	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/20 (45%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

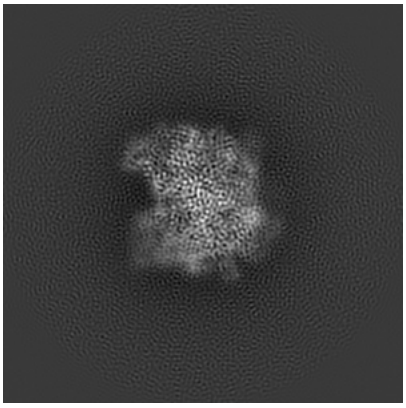
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-7350. These allow visual inspection of the internal detail of the map and identification of artifacts.

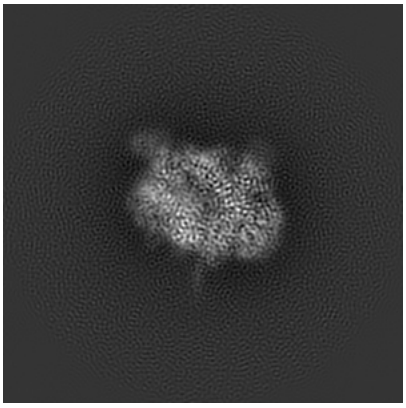
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

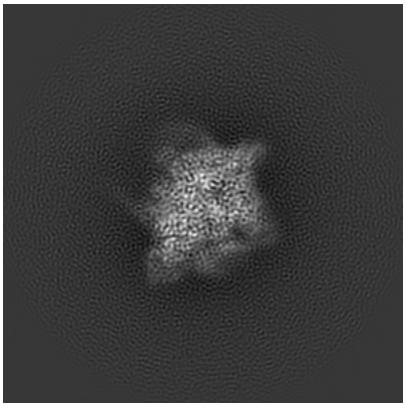
6.1.1 Primary map



X



Y

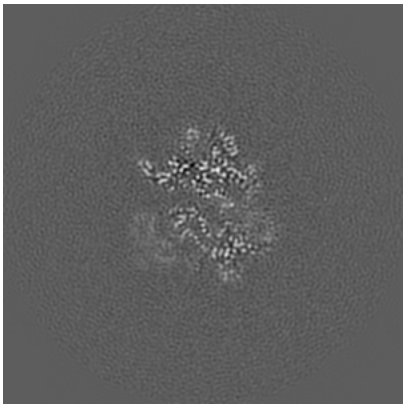


Z

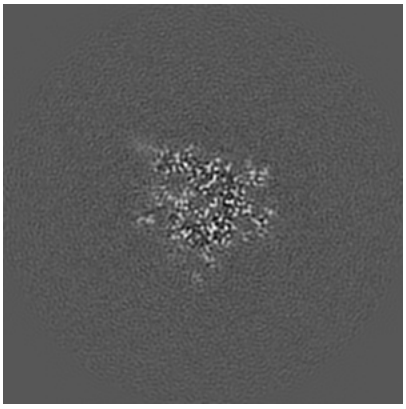
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

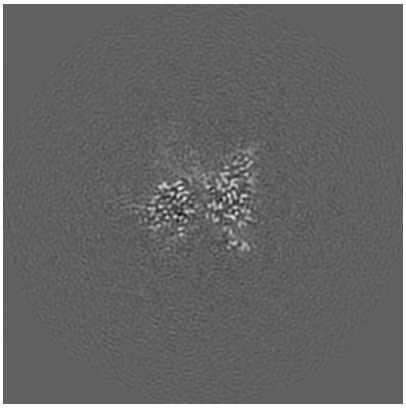
6.2.1 Primary map



X Index: 128



Y Index: 128

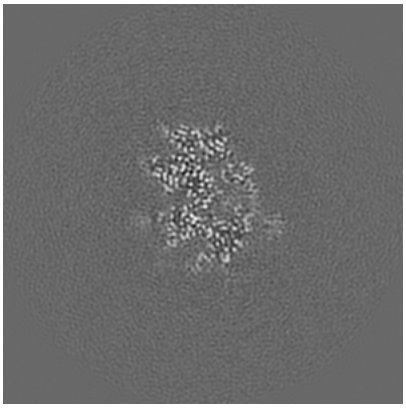


Z Index: 128

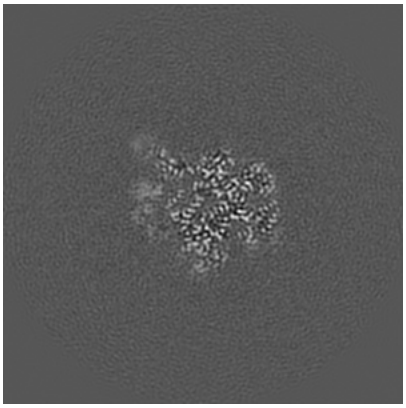
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

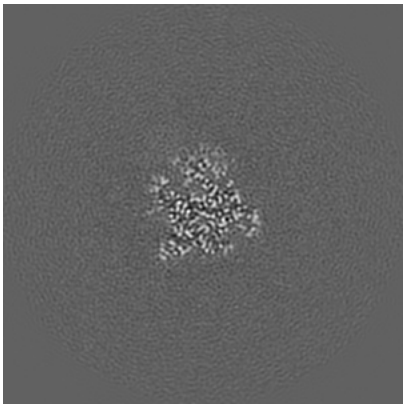
6.3.1 Primary map



X Index: 120



Y Index: 122

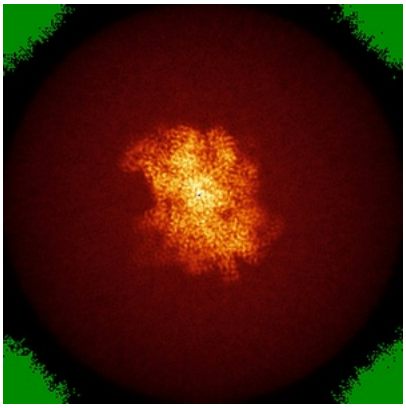


Z Index: 142

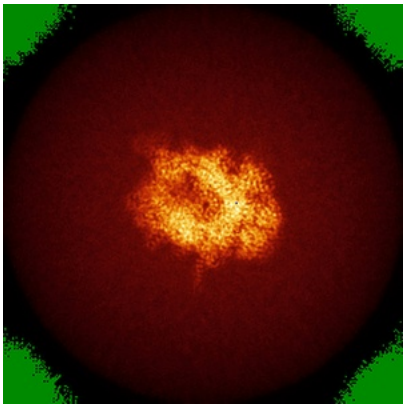
The images above show the largest variance slices of the map in three orthogonal directions.

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

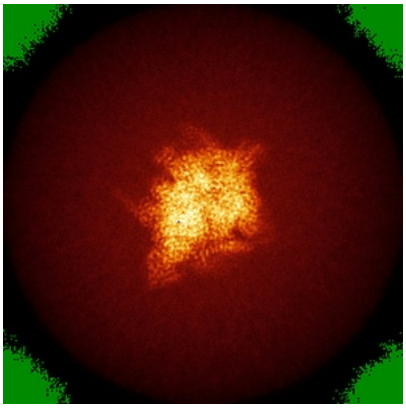
6.4.1 Primary map



X



Y

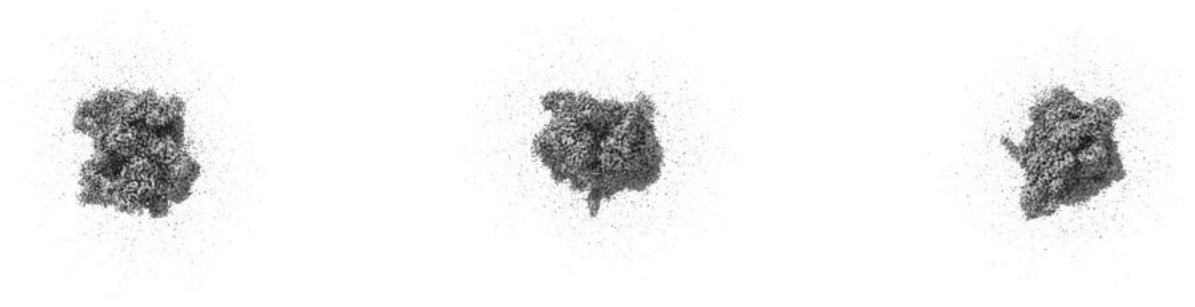


Z

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

6.5 Orthogonal surface views (i)

6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.016. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

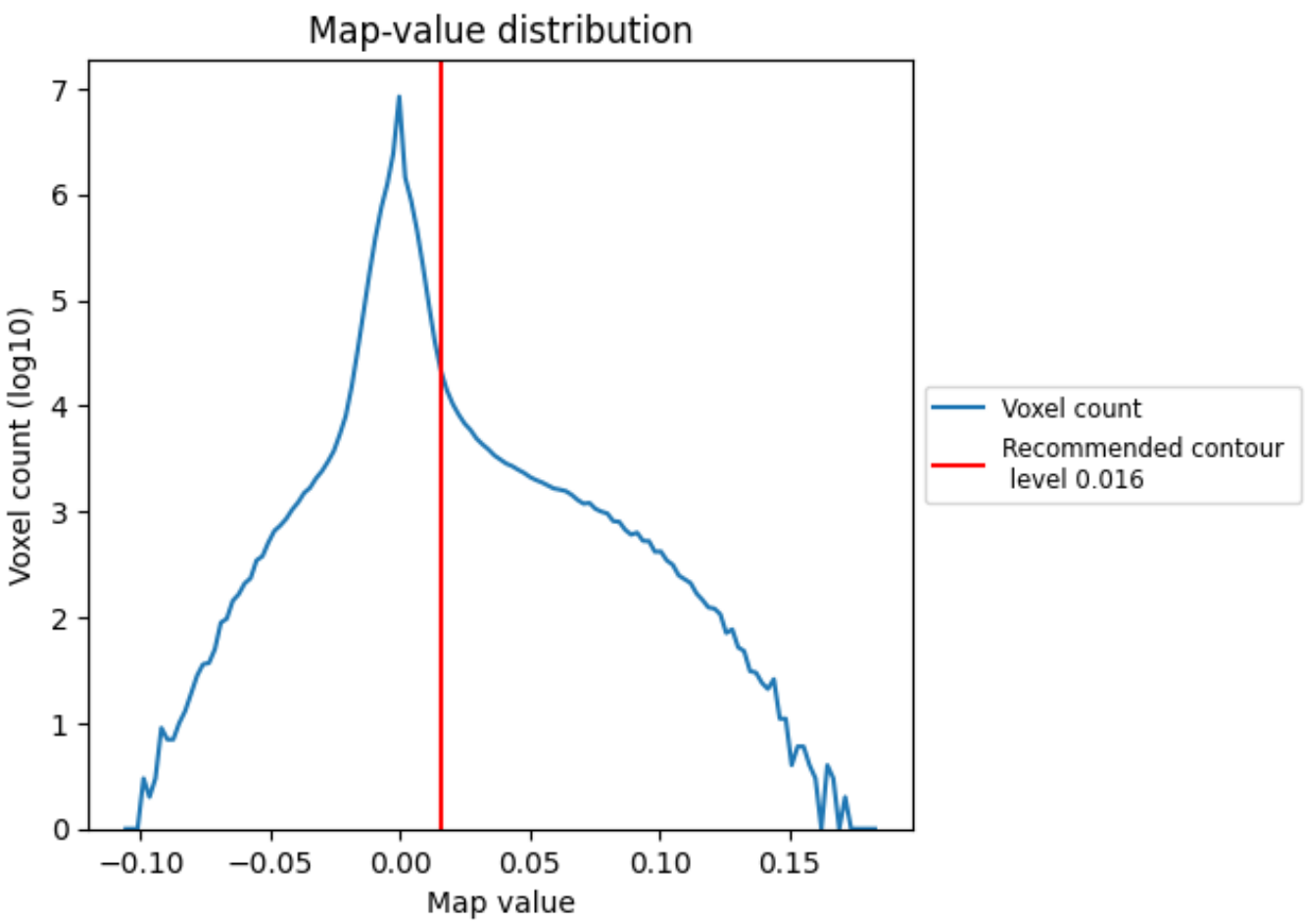
6.6 Mask visualisation (i)

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

7 Map analysis (i)

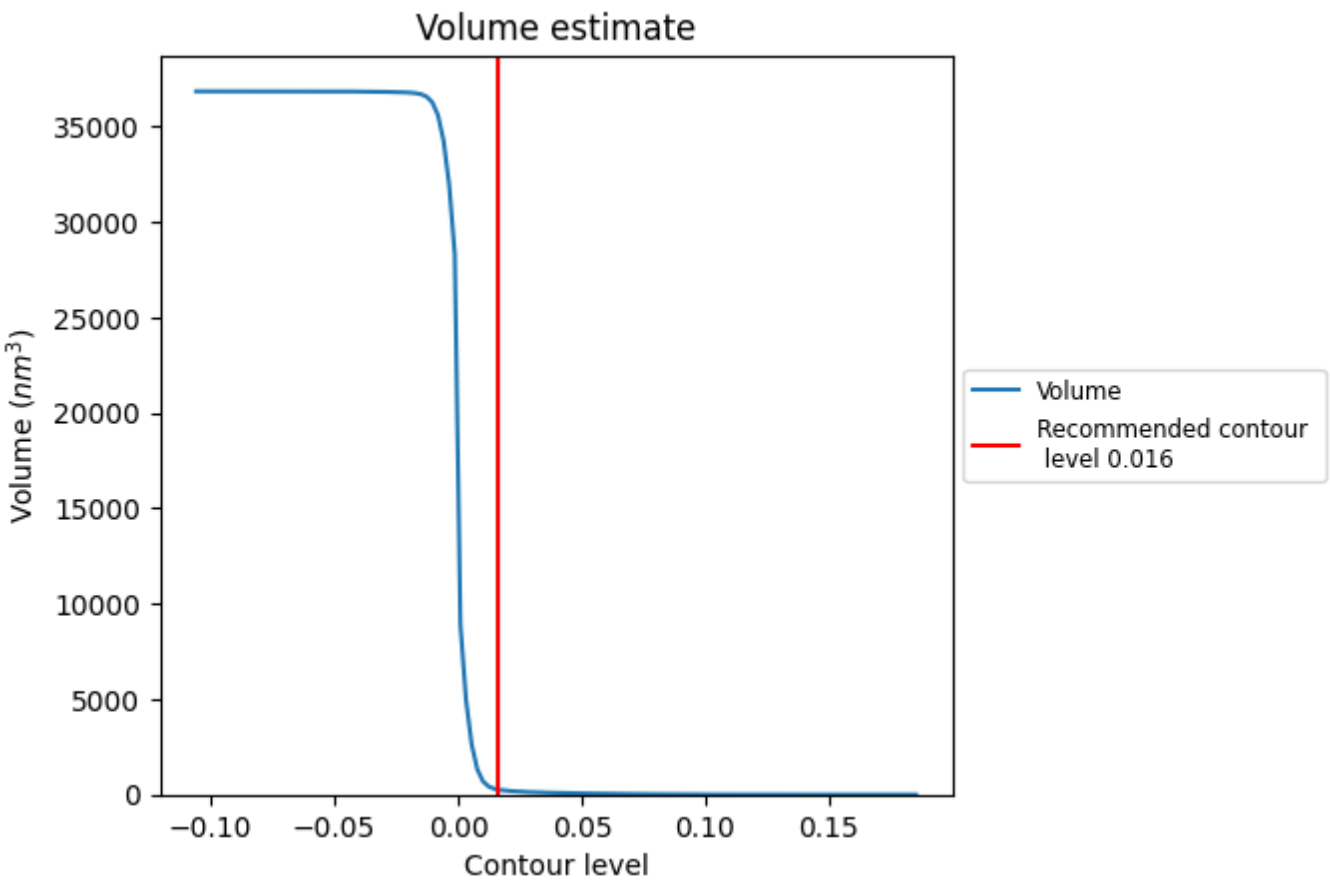
This section contains the results of statistical analysis of the map.

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.

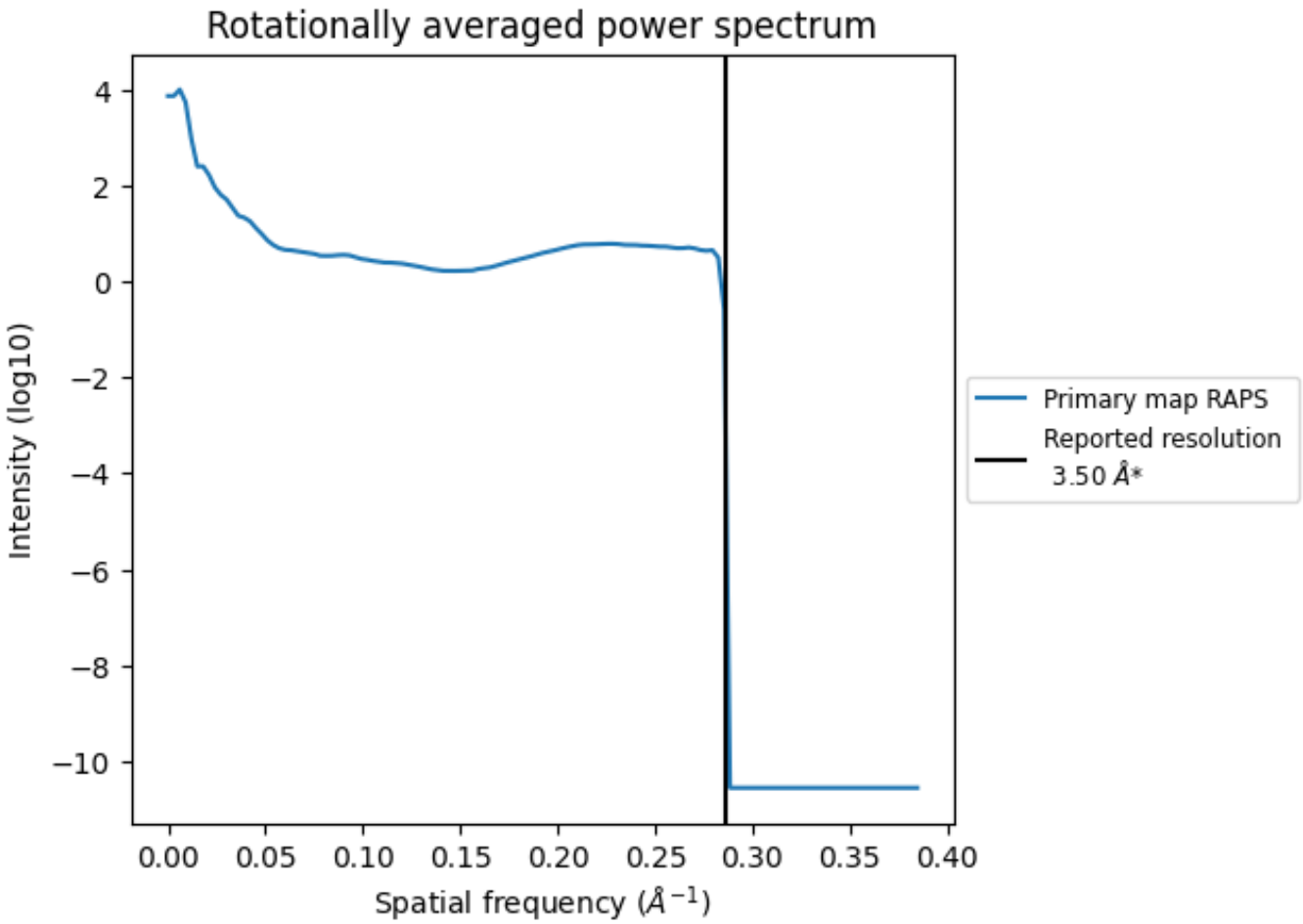
7.2 Volume estimate (i)



The volume at the recommended contour level is 269 nm^3; this corresponds to an approximate mass of 243 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 0.286 Å^{-1}

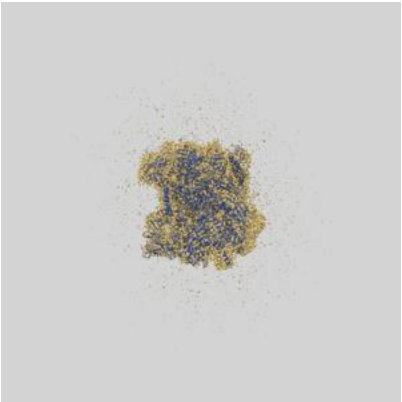
8 Fourier-Shell correlation (i)

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

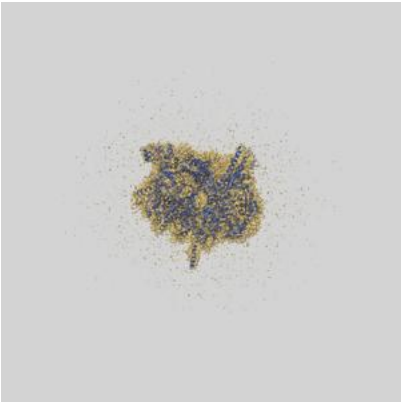
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-7350 and PDB model 6C6T. Per-residue inclusion information can be found in section 3 on page 6.

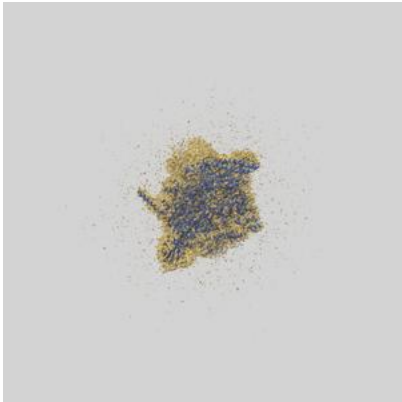
9.1 Map-model overlay (i)



X



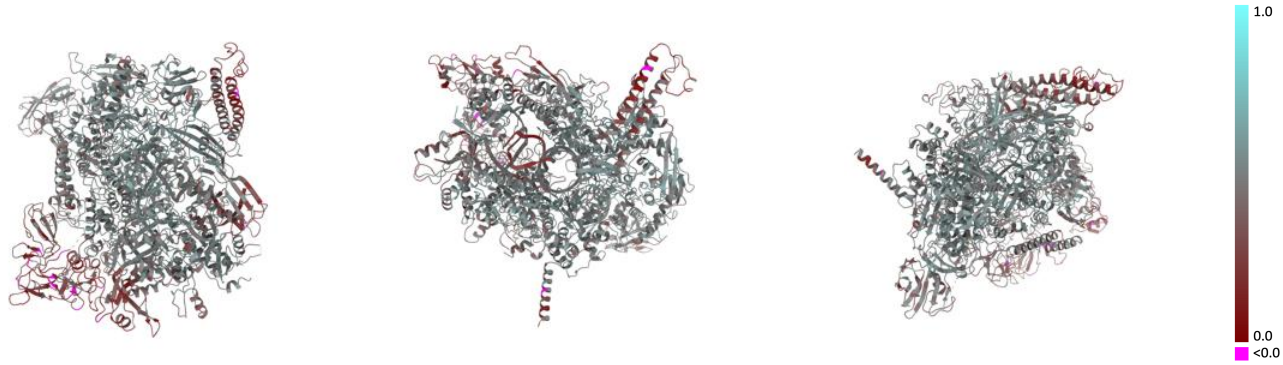
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.016 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.

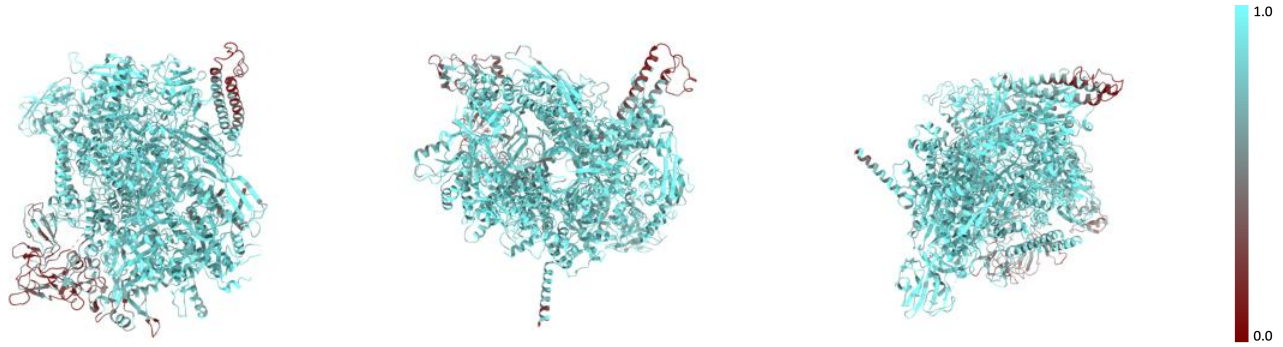
9.2 Q-score mapped to coordinate model (i)



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.

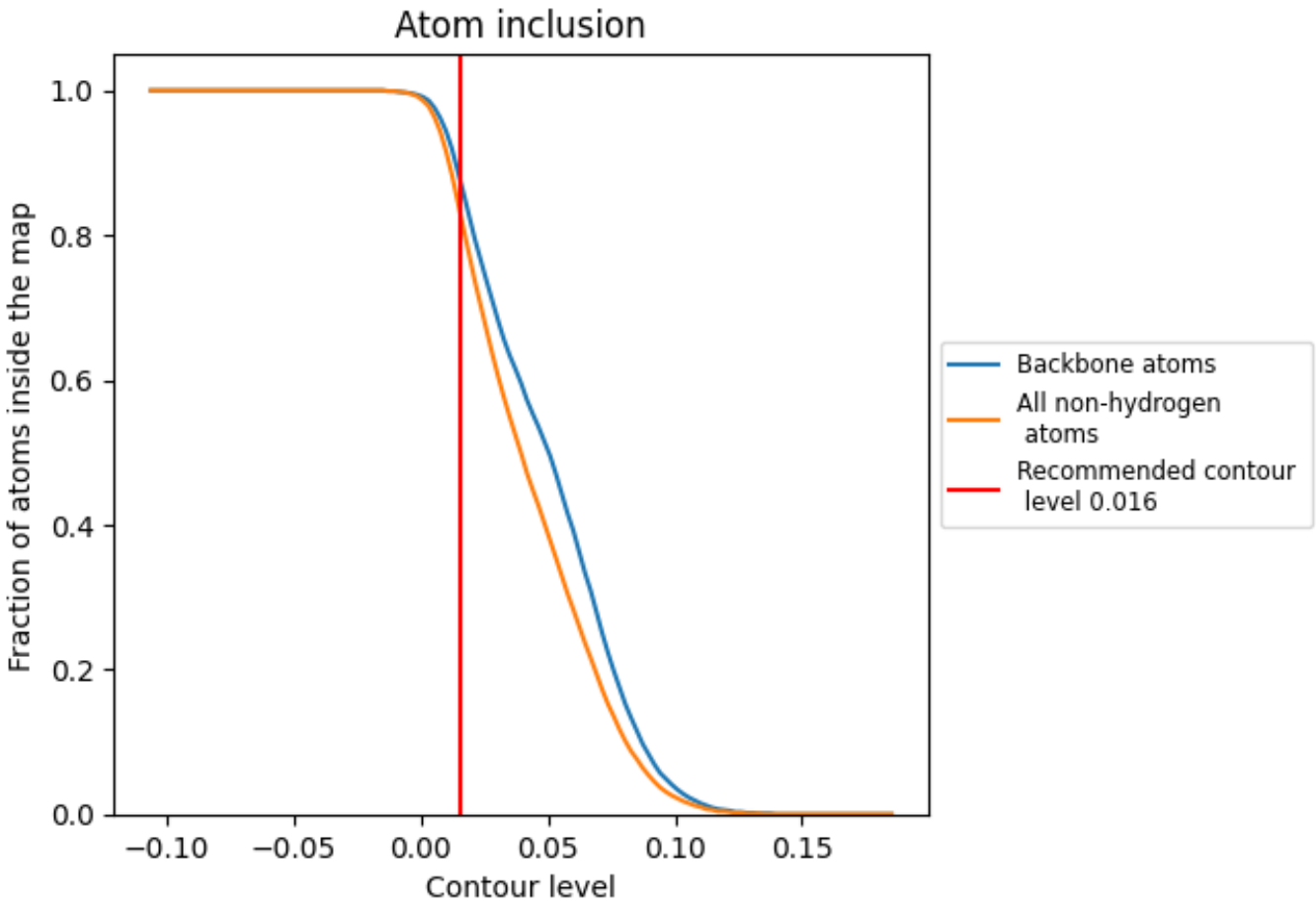
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.016).

9.4 Atom inclusion (i)



At the recommended contour level, 87% of all backbone atoms, 83% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.016) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8260	0.4620
A	0.9000	0.3980
B	0.8840	0.4110
D	0.9200	0.4950
G	0.8940	0.5060
H	0.8480	0.4630
I	0.8250	0.4660
J	0.7960	0.4560
K	0.7930	0.4470
R	0.9680	0.4870