wwPDB EM Validation Summary Report (i)

Nov 20, 2022 – 08:21 am GMT

PDB ID	:	6HKO
EMDB ID	:	EMD-0238
Title	:	Yeast RNA polymerase I elongation complex bound to nucleotide analog GM-PCPP
Authors	:	Tafur, L.; Sadian, Y.; Weis, F.; Muller, C.W.
Deposited on	:	2018-09-07
Resolution	:	3.42 Å(reported)
Based on initial models	:	2E2J, 4C3I, 4C2M, 5M5X, 4A3J

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

EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
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   ⓘ

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

The reported resolution of this entry is 3.42 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.
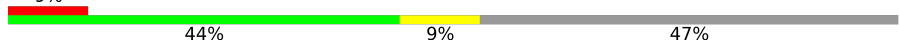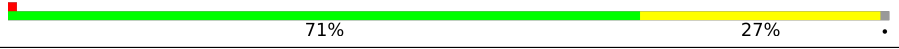
Metric	Percentile Ranks	Value
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		0.2%
RNA backbone		0.49

Worse   *Better*

■ Percentile relative to all structures

▯ Percentile relative to all EM structures
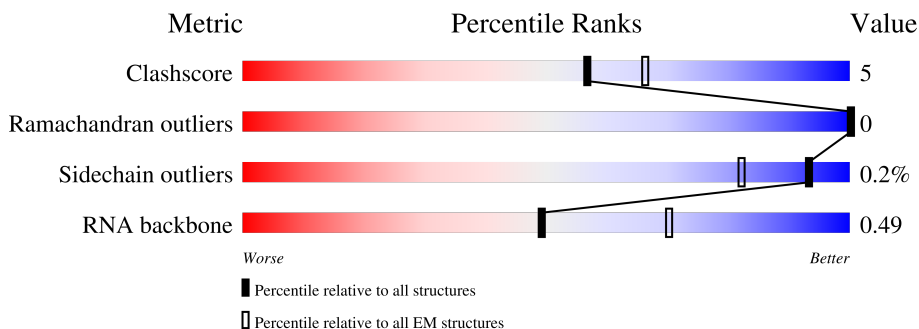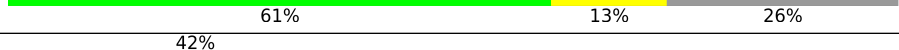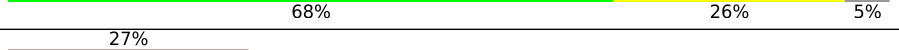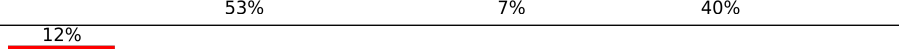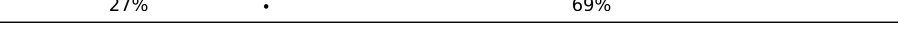
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	1664	5% 76% 12% 12%
2	B	1203	· 81% 16% ·
3	C	335	· 78% 13% 9%
4	D	137	7% 36% 7% 57%
5	E	215	6% 90% 10%
6	F	155	· 54% 11% 35%
7	G	326	21% 48% 6% 46%

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	H	146	82% 8% 10%
9	I	125	9% 44% 9% 47%
10	J	70	71% 27%
11	K	142	54% 15% 31%
12	L	70	56% 7% 37%
13	R	20	40% 5% 55%
14	S	38	50% 61% 13% 26%
15	T	38	42% 68% 26% 5%
16	N	233	27% 53% 7% 40%
17	M	415	12% 27% 69%

2 Entity composition (i)

There are 20 unique types of molecules in this entry. The entry contains 34518 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
1	A	1458	Total	C	N	O	S	0	0
			11513	7269	2002	2180	62		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1166	Total	C	N	O	S	0	0
			9270	5862	1628	1729	51		

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

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

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

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

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

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

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

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

- 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	66	Total	C	N	O	S	0	0
			486	305	80	97	4		

- 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 a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	R	9	Total	C	N	O	P	0	0
			195	87	39	60	9		

- Molecule 14 is a DNA chain called Non-template strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	S	28	Total 581	C 276	N 114	O 163	P 28	0	0

- Molecule 15 is a DNA chain called Template strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	T	36	Total 732	C 350	N 127	O 219	P 36	0	0

- Molecule 16 is a protein called DNA-directed RNA polymerase I subunit RPA34.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	N	140	Total 1110	C 710	N 183	O 213	S 4	0	0

- Molecule 17 is a protein called DNA-directed RNA polymerase I subunit RPA49.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	M	129	Total 1011	C 642	N 171	O 198	0	0

- Molecule 18 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
18	A	1	Total 1	Mg 1	0

- Molecule 19 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
19	A	2	Total 2	Zn 2	0
19	B	1	Total 1	Zn 1	0
19	I	1	Total 1	Zn 1	0
19	J	1	Total 1	Zn 1	0
19	L	1	Total 1	Zn 1	0

- Molecule 20 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-

letter code: G2P) (formula: $C_{11}H_{18}N_5O_{13}P_3$).
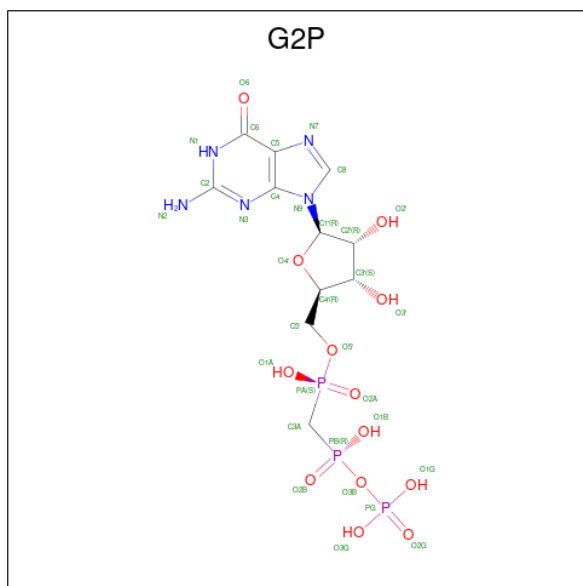
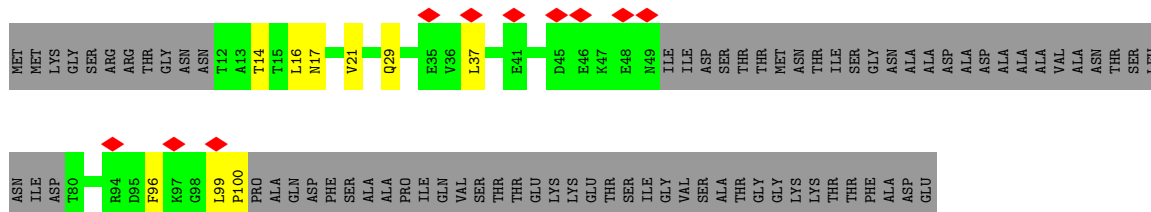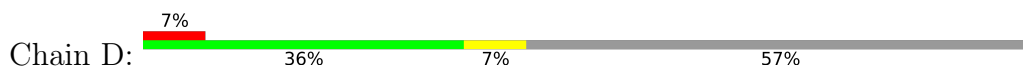| Mol	Chain	Residues	Atoms					AltConf
20	T	1	Total	C	N	O	P	0
			32	11	5	13	3	

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: DNA-directed RNA polymerase I subunit RPA190

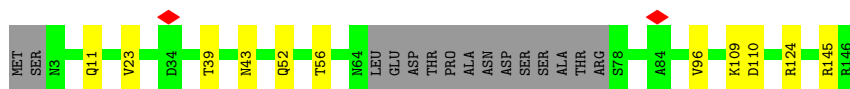THR LYS GLN ALA VAL SER TYR ASP GLU PRO ASP GLU ASP GLU ILE GLU THR MET ARG GLU ALA ALA GLU LYS SER SER ASP GLU GLU GLY ILE ASP SER ASP LYS GLU SER ASP SER SER GLU ASP GLU GLU ASP GLU VAL ASP MET ASN GLU GLN ILE ASN LYS SER ILE **V1433 E1434 A1435 N1436 N1437**

N1438 M1439 N1440 K1441 R1446 **H1454** D1464 E1465 S1466 A1478 D1479 T1480 E1481 K1482 I1498 I1504 E1513 N1514 G1515 V1518 L1519 G1523 V1524 I1536 I1541 R1559 N1560 T1561 I1562 R1580 D1583 M1588 R1591 Y1595 L1622 Q1633 L1634 R1640 V1643

G1644 K1645 T1651 L1657 V1660 P1661 N1662 A1663 ALA

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

Chain B: 81% 16% •

MET SER LYS VAL ILE LYS PRO PRO GLY GLN **A11 R12 T13** A14 R17 T18 L19 E20 P29 S33 **D57** G58 N62 E70 K71 V72 I73 F74 D75 P78 LEU ASN SER GLU ASP GLU ILE SER ASN SER G89 G92 S96 V97 S98 D111 GLY VAL SER SER

A116 R129 L130 G135 K136 L137 W143 L152 L161 Q166 S167 N168 S176 P177 E186 S187 **D188** E189 I190 I199 I203 V208 Q209 R210 R211 N212 A216 **A227** I234 R237 R240 S245 L250 H251 Y252 L253 Q257 F260 V271

P272 V273 L280 S284 D285 R286 E287 D296 R305 L320 Q321 N322 R323 V326 R335 S341 L354 V360 H361 L362 **D365** R372 L381 L384 P394 D395 Q400 E401 L404 L408 **I417** D418 I424 V428 R429 M430 V449

V453 I457 L465 S466 T467 **Q473** T485 E489 K490 I491 N492 R495 R505 **K513** S524 W525 L528 H532 I551 S574 **F577** C584 G590 K591 I626 G639 L642 L656 I674 A675 Q679 N683 H688 V689 E690

T704 S707 N710 Q711 S712 P713 C719 T725 V731 R743 P750 K753 D758 V773 I776 G780 Y781 D782 A786 N790 K791 S792 R796 Y804 K808 **V809 D810** L813 **N816 R817 G818** **E831** D843 P857 I858 I871

L887 **E891 S892 N893 K894 F895** L898 I903 R908 I913 G914 R920 V926 C927 W931 I934 D935 M936 S939 D946 I947 R957 M958 T959 I960 F963 H975 D980 S981 N987 T991 P992 A993 A1003 N1006 N1010 M1013 G1016

R1023 I1026 Y1027 Y1032 N1041 D1042 K1043 F1044 Q1045 V1046 R1047 L1055 T1056 R1063 G1067 R1070 H1082 G1083 T1084 D1090 N1094 R1105 L1111 **I1120** R1129 R1130 C1131 S1132 F1135 E1136 K1139 LYS LEU LEU THR LYS SER GLU ASP GLY GLU LYS ILE

PHE **I1153 D1154** F1167 V1168 V1176 L1189 I1194 R1195 **E1201** P1202 K1203

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

Chain C: 78% 13% 9%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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

Chain D:

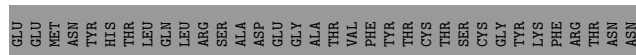• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:

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

Chain F:

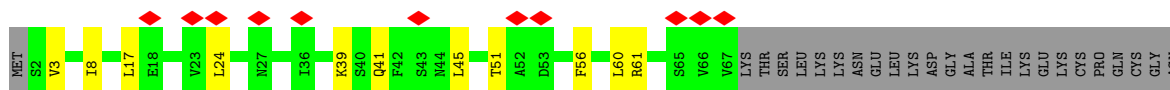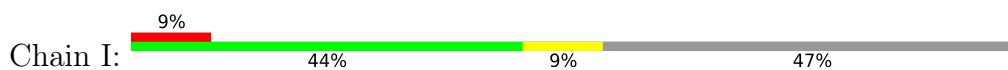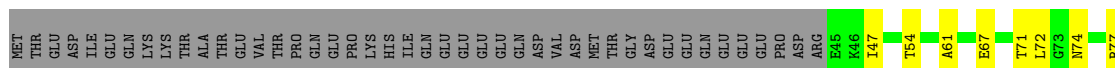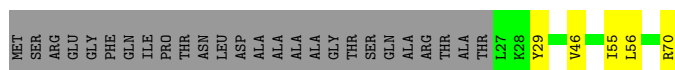• Molecule 7: DNA-directed RNA polymerase I subunit RPA43

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:

• Molecule 13: RNA

Chain R:

40%   5%   55%

U A U A U G C A U A A12 G18 G19 C20

• Molecule 14: Non-template strand

50%

Chain S:

61%   13%   26%

G1 G2 C3 A4 G5 T6 A7 C8 T9 A10 G11 T12 A13 A14 A15 DC DT DA DG DT DA DT DT G24 G28 T32 T33 G34 A35 C36 DT DT

• Molecule 15: Template strand

42%

Chain T:

68%   26%   5%

DA DA G3 T4 C5 A6 C11 C17 C18 T22 C23 A24 T25 T26 A27 C28 T29 A30 G31 T32 A33 C34 T35 G36 C37 C38

• Molecule 16: DNA-directed RNA polymerase I subunit RPA34

27%

Chain N:

53%   7%   40%

MET SER LYS LEU SER LYS ASP TYR VAL SER ASP SER ASP SER ASP GLU VAL ILE SER ASN GLU PHE S24 I25 P26 F29 K33 H34 N37 F38 P39 L40 N41 GLY ASP ASN LYS LYS ALA K49 Q50 Q51 Q52 K57 F58 P59 S60 N61 V62 D63 I64 S65 K66 L67

K68 S69 L70 P71 V72 ASP PHE GLU SER SER T78 T81 I82 D83 K84 I89 M90 D91 D92 T93 D94 I95 E96 S97 S98 L99 T100 Q101 D102 N103 L104 S105 N106 L109 S113 E114 S115 K116 E117 A122 S123 T124 A125 K126 D127 N128 A129 P130 L131 Q132 F133 D134 K135 V136 F137

S138 V139 I148 R154 V155 P156 R157 E169 G174 Y175 ASP ALA GLU ASP PHE HIS VAL ALA GLU VAL LYS GLU ASN LYS GLU PRO LYS ARG SER HIS HIS ASP ASP GLU GLU SER SER GLU LYS LYS LYS LYS LYS LYS GLU LYS ARG GLU LYS ARG GLU LYS

LYS ASP LYS LYS ASP LYS LYS LYS HIS ARG ASP

• Molecule 17: DNA-directed RNA polymerase I subunit RPA49

12%

Chain M:

27%   •   69%

MET SER VAL LYS ARG SER VAL S8 E9 I10 E11 I12 V15 Q16 D17 Q18 V21 G24 K28 R31 S34 D35 Y41 K44 K45 S46 E47 K48 D49 E50 G55 E56 N57 E58 E61 Y62 S69 S70 Q71 A72 Q75 L80 F81 E84 K85 K86

S87 I88 Q89 L90 Y91 K92 A93 V100 V101 S105 S115 LYS SER ASP THR ARG PRO SER ALA L124 R125 N126 A127 L128 G129 E130 A131 F132 G133 T134 K135 K136 A137 K138 K139 A140 I141 A142 D143 L144 GLU ARG ASN ARG ILE ASP SER ASP LYS LEU THR ASP SER ALA ILE ASP ILE

VAL ASP SER VAL ARG THR ALA SER LYS ASP LEU PRO THR ARG ALA GLN LEU ASP GLU ILE THR SER ASN ASP ARG PRO THR PRO LEU ALA ASN ILE ASP ALA THR ASP VAL GLU GLN ILE ILE TYR PRO ILE GLU SER ILE ILE PRO LYS LYS GLU GLN LEU GLN PHE ILE ARG VAL SER SER ILE

LEU LYS GLU ALA ASP LYS GLU LYS LYS LEU GLU GLU LEU PHE PRO TYR GLN ASN ASN SER LYS TYR VAL ALA LYS LEU LEU ASP SER LEU THR GLN PRO SER GLN MET THR LYS LEU GLN LEU LEU TYR TYR LEU SER LEU LEU LEU GLY VAL TYR GLU ASN ARG ARG VAL ASN ASN LYS THR

LYS LEU LEU GLU ARG LEU ASN SER PRO GLU ILE VAL ASP GLY ILE LEU SER ARG PHE THR VAL ILE LYS PRO GLY GLN PHE GLY ARG SER LYS ASP ARG SER TYR PHE ILE ASP PRO GLN ASN GLU ASP ASP LYS ILE LEU CYS TYR ILE LEU ALA ILE MET HIS LEU ASP ASN

PHE ILE VAL GLU ILE THR PRO LEU ALA HIS GLU LEU ASN LEU LYS PRO SER LYS VAL VAL SER LEU PHE ARG LEU GLY VAL ILE ALA ILE VAL LYS GLY ALA THR VAL ALA GLN ALA GLU PHE GLY ILE ASP GLN ASN GLU ASP LYS ILE LEU SER THR ALA ALA SER TYR LYS ILE ALA THR MET LYS VAL PRO PHE

LYS LEU PRO GLU MET THR ARG ARG GLY ARG GLY PRO ARG ARG

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	34232	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	0.210	Depositor
Minimum map value	-0.107	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.011	Depositor
Recommended contour level	0.0388	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, MG, G2P

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	0.33	0/11721	0.56	0/15830
2	B	0.36	0/9475	0.57	0/12807
3	C	0.32	0/2467	0.52	0/3344
4	D	0.26	0/473	0.52	0/641
5	E	0.30	0/1787	0.48	0/2406
6	F	0.32	0/838	0.48	0/1129
7	G	0.28	0/1421	0.54	0/1938
8	H	0.32	0/1070	0.56	0/1449
9	I	0.29	0/492	0.60	0/667
10	J	0.36	0/578	0.60	0/775
11	K	0.32	0/776	0.59	0/1047
12	L	0.32	0/354	0.55	0/468
13	R	0.39	0/218	0.89	0/338
14	S	0.56	0/653	0.92	0/1004
15	T	0.62	0/818	0.96	0/1259
16	N	0.28	0/1130	0.50	0/1523
17	M	0.28	0/1027	0.52	0/1374
All	All	0.35	0/35298	0.58	0/47999

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	11513	0	11588	122	0
2	B	9270	0	9162	122	0
3	C	2415	0	2403	28	0
4	D	467	0	468	7	0
5	E	1751	0	1776	14	0
6	F	823	0	841	10	0
7	G	1386	0	1386	13	0
8	H	1052	0	1021	7	0
9	I	486	0	489	7	0
10	J	569	0	585	11	0
11	K	766	0	765	15	0
12	L	352	0	374	5	0
13	R	195	0	99	0	0
14	S	581	0	316	5	0
15	T	732	0	408	11	0
16	N	1110	0	1141	12	0
17	M	1011	0	1026	11	0
18	A	1	0	0	0	0
19	A	2	0	0	0	0
19	B	1	0	0	0	0
19	I	1	0	0	0	0
19	J	1	0	0	0	0
19	L	1	0	0	0	0
20	T	32	0	14	3	0
All	All	34518	0	33862	336	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 336 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)
15:T:24:DA:H2"	15:T:25:DT:OP1	1.83	0.76
1:A:712:ILE:H	11:K:106:GLN:HE22	1.37	0.70
2:B:129:ARG:HH22	2:B:891:GLU:HG3	1.56	0.69
2:B:675:ALA:HB3	2:B:689:VAL:HG12	1.75	0.67
1:A:464:GLU:HA	1:A:469:LYS:HD2	1.77	0.66

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	1440/1664 (86%)	1366 (95%)	74 (5%)	0	100	100
2	B	1158/1203 (96%)	1081 (93%)	77 (7%)	0	100	100
3	C	302/335 (90%)	292 (97%)	10 (3%)	0	100	100
4	D	55/137 (40%)	55 (100%)	0	0	100	100
5	E	212/215 (99%)	203 (96%)	9 (4%)	0	100	100
6	F	98/155 (63%)	96 (98%)	2 (2%)	0	100	100
7	G	172/326 (53%)	166 (96%)	6 (4%)	0	100	100
8	H	127/146 (87%)	124 (98%)	3 (2%)	0	100	100
9	I	64/125 (51%)	57 (89%)	7 (11%)	0	100	100
10	J	67/70 (96%)	61 (91%)	6 (9%)	0	100	100
11	K	96/142 (68%)	90 (94%)	6 (6%)	0	100	100
12	L	42/70 (60%)	38 (90%)	4 (10%)	0	100	100
16	N	134/233 (58%)	121 (90%)	13 (10%)	0	100	100
17	M	125/415 (30%)	111 (89%)	14 (11%)	0	100	100
All	All	4092/5236 (78%)	3861 (94%)	231 (6%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	1287/1465 (88%)	1285 (100%)	2 (0%)	93	98
2	B	1019/1053 (97%)	1014 (100%)	5 (0%)	88	95
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%)	156 (100%)	0	100	100
8	H	115/128 (90%)	115 (100%)	0	100	100
9	I	58/110 (53%)	57 (98%)	1 (2%)	60	82
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
16	N	131/220 (60%)	131 (100%)	0	100	100
17	M	112/371 (30%)	112 (100%)	0	100	100
All	All	3679/4636 (79%)	3671 (100%)	8 (0%)	93	98

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	41	GLN
2	B	1195	ARG
2	B	1105	ARG
2	B	1043	LYS
2	B	1111	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	175	GLN
3	C	232	GLN
17	M	75	GLN
11	K	106	GLN
16	N	132	GLN

5.3.3   RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	R	8/20 (40%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	R	18	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 7 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
20	G2P	T	3000	-	27,34,34	1.96	8 (29%)	33,54,54	1.89	8 (24%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
20	G2P	T	3000	-	-	2/15/38/38	0/3/3/3

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
20	T	3000	G2P	PA-O2A	4.09	1.61	1.51
20	T	3000	G2P	PB-O2B	4.09	1.61	1.51
20	T	3000	G2P	C5-C6	3.93	1.48	1.41
20	T	3000	G2P	PB-O1B	-3.39	1.48	1.56
20	T	3000	G2P	PA-O1A	-3.39	1.48	1.56

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
20	T	3000	G2P	C2-N3-C4	4.80	120.84	115.36
20	T	3000	G2P	C5-C6-N1	-3.93	118.06	123.43
20	T	3000	G2P	C2-N1-C6	3.85	122.04	115.93
20	T	3000	G2P	C3'-C2'-C1'	3.55	106.32	100.98
20	T	3000	G2P	PB-O3B-PG	-3.33	120.88	132.62

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
20	T	3000	G2P	PB-C3A-PA-O2A
20	T	3000	G2P	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
20	T	3000	G2P	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand G2P T 3000

Bond lengths

Bond angles

Torsions

Rings
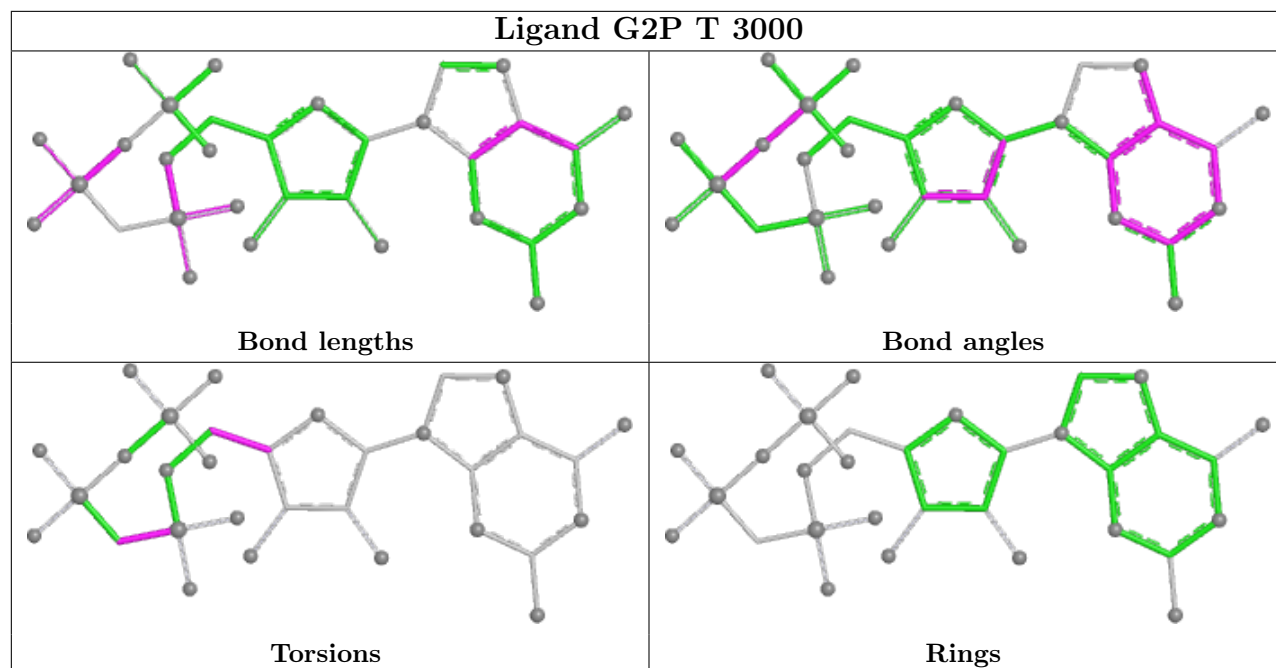
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-0238. 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 |
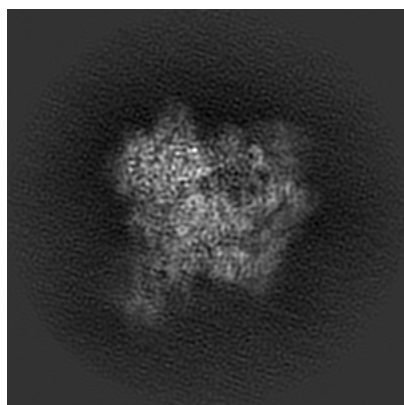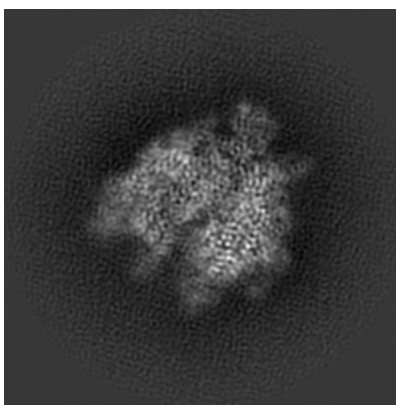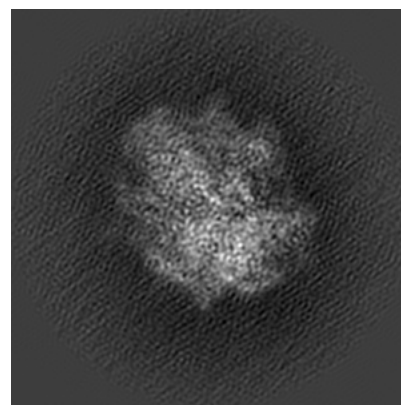
The images above show the map projected in three orthogonal directions.

6.2   Central slices ⓘ
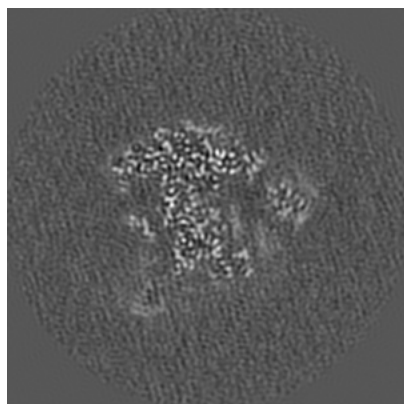
6.2.1   Primary map
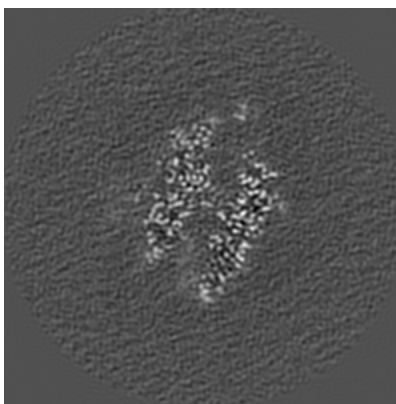
| X Index: 130 | Y Index: 130 | Z Index: 130 |
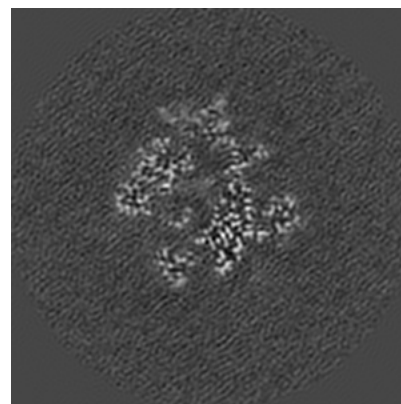
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map
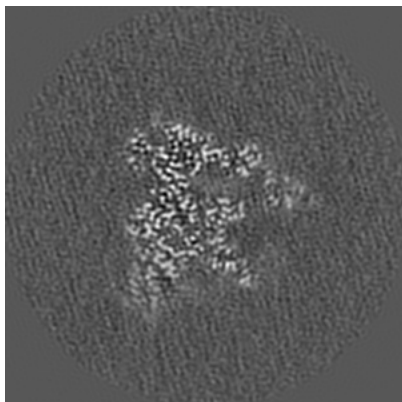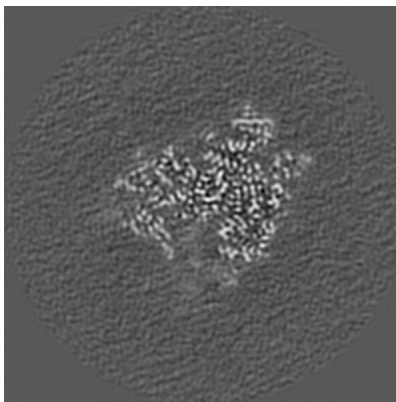
| X Index: 140 | Y Index: 109 | Z Index: 155 |
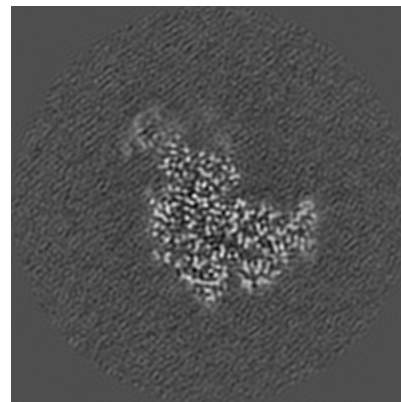
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

6.4.1 Primary map

| X | Y | Z |
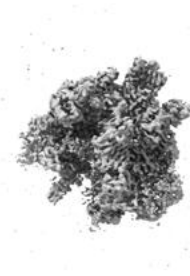
The images above show the 3D surface view of the map at the recommended contour level 0.0388. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

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 ⓘ
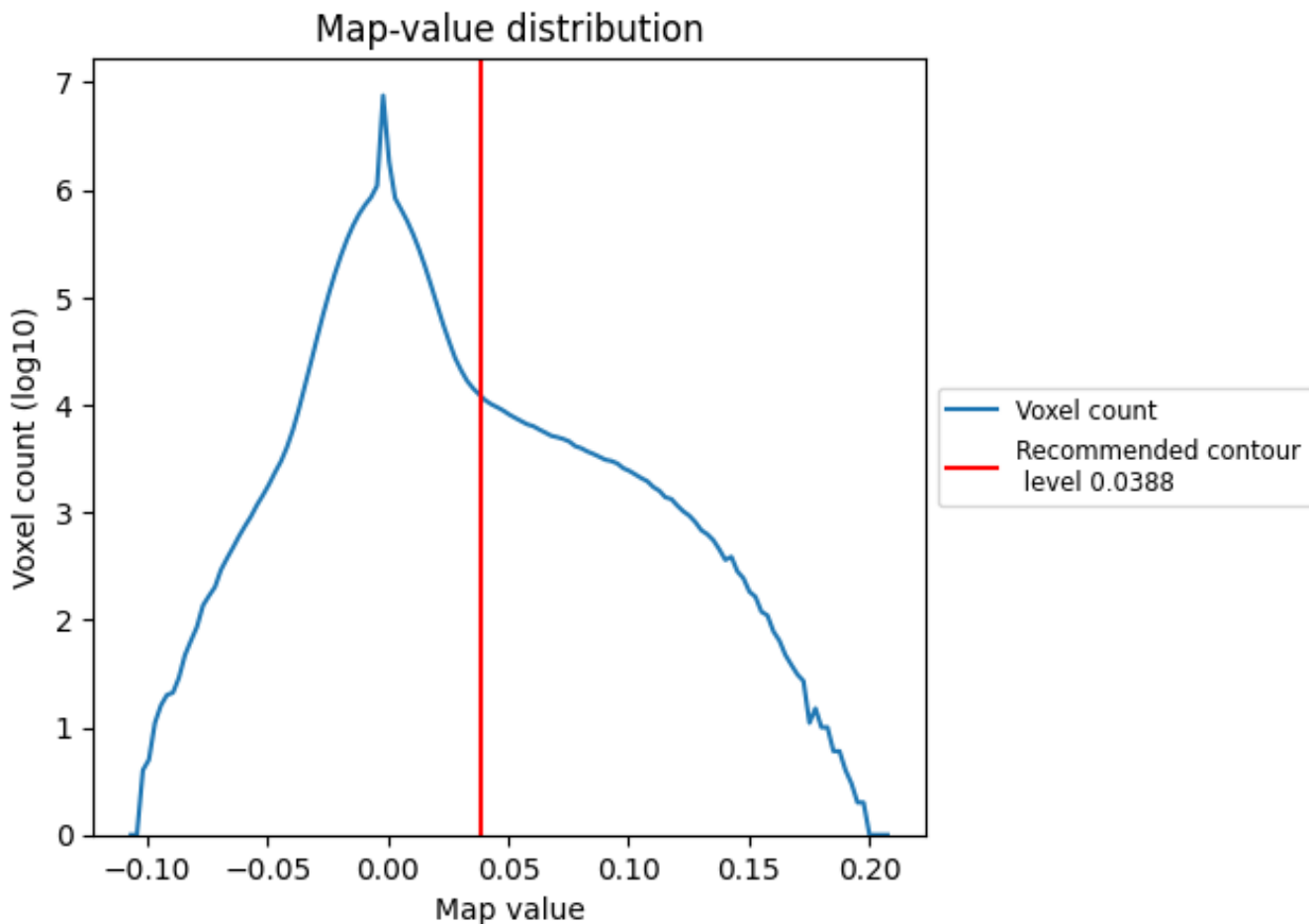
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2   Volume estimate ⓘ
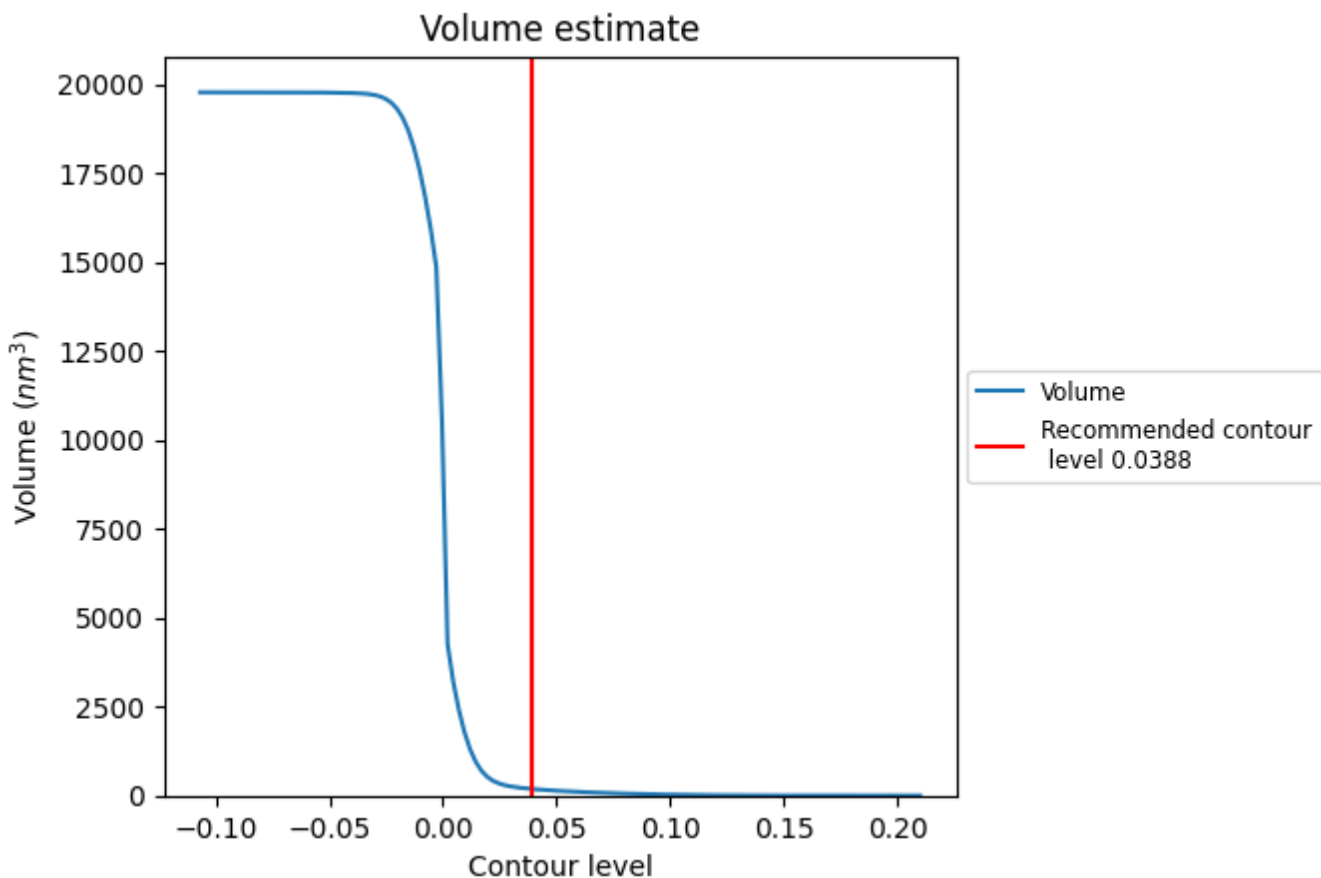
Volume estimate

The volume at the recommended contour level is 188 nm^3; this corresponds to an approximate mass of 170 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 ⓘ
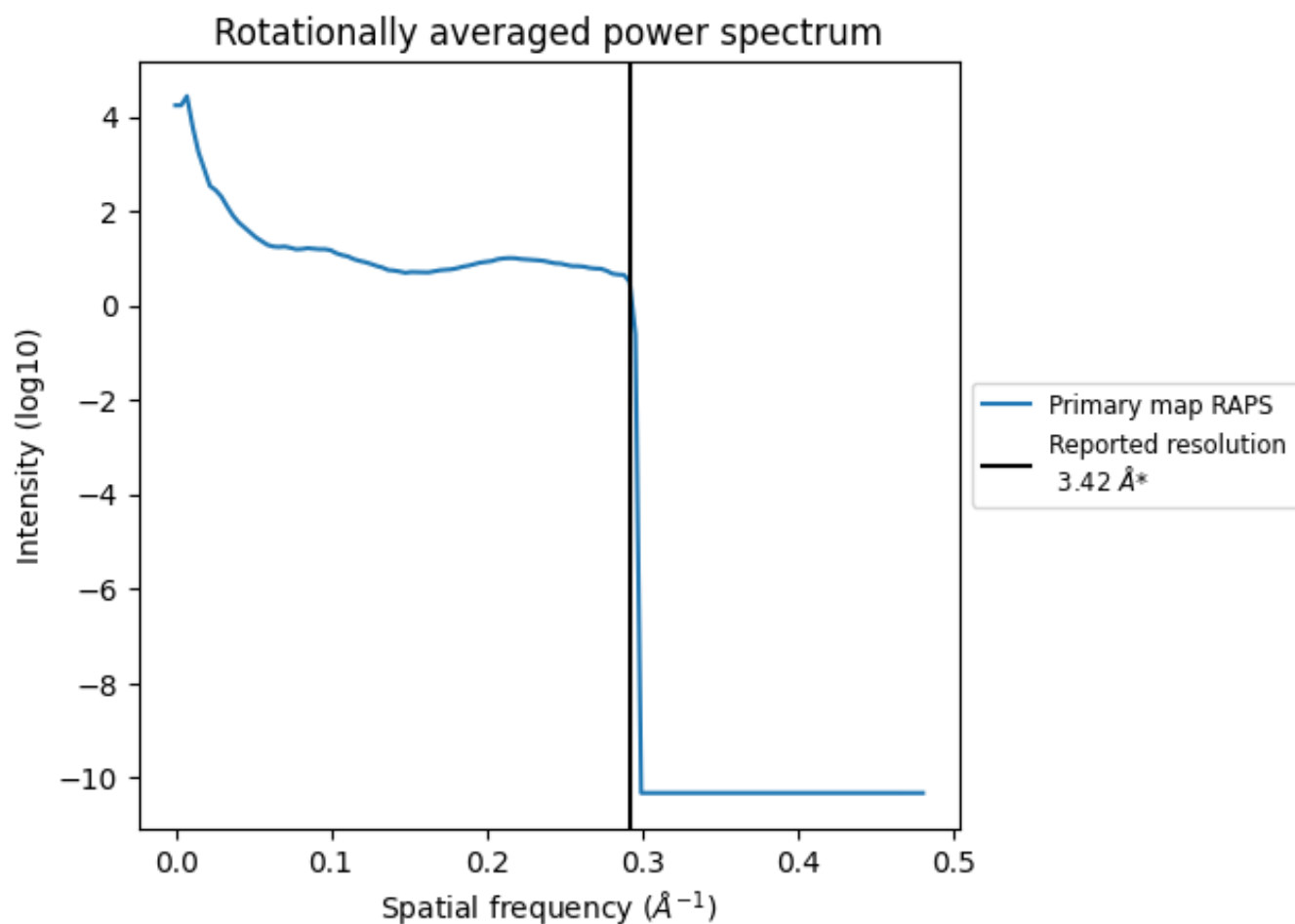
*Reported resolution corresponds to spatial frequency of 0.292 Å^{-1}

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 ⓘ
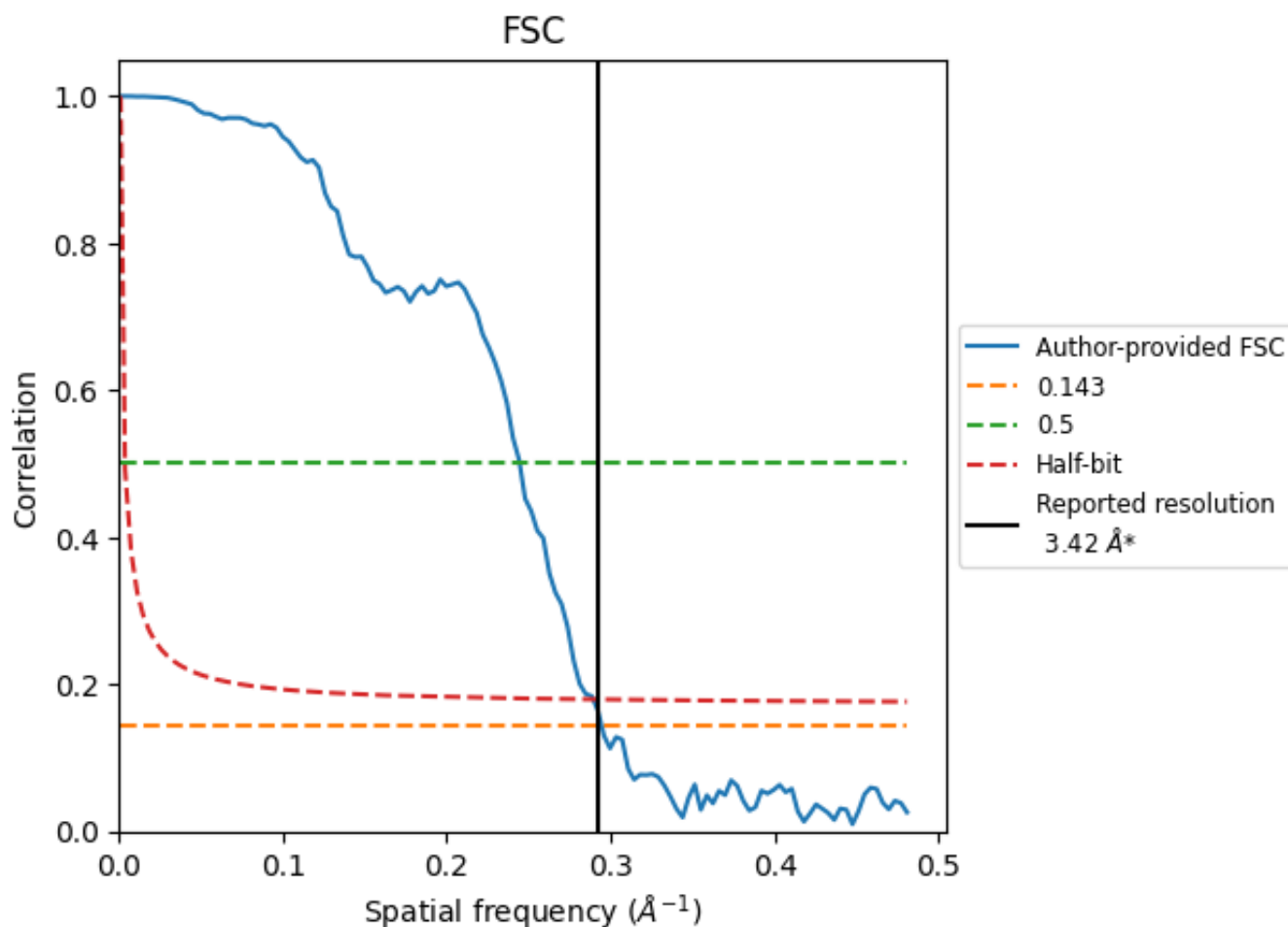
*Reported resolution corresponds to spatial frequency of 0.292 Å^{-1}

8.2　Resolution estimates　ⓘ

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.42	-	-
Author-provided FSC curve	3.40	4.09	3.46
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-0238 and PDB model 6HKO. Per-residue inclusion information can be found in section 3 on page 8.

9.1 Map-model overlay ⓘ
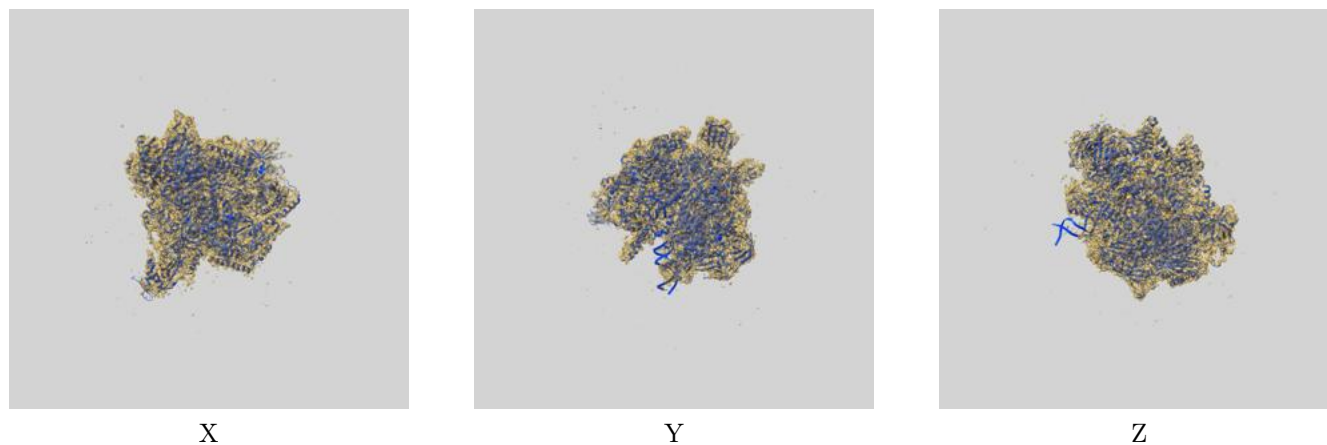
| X | Y | Z |

The images above show the 3D surface view of the map at the recommended contour level 0.0388 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
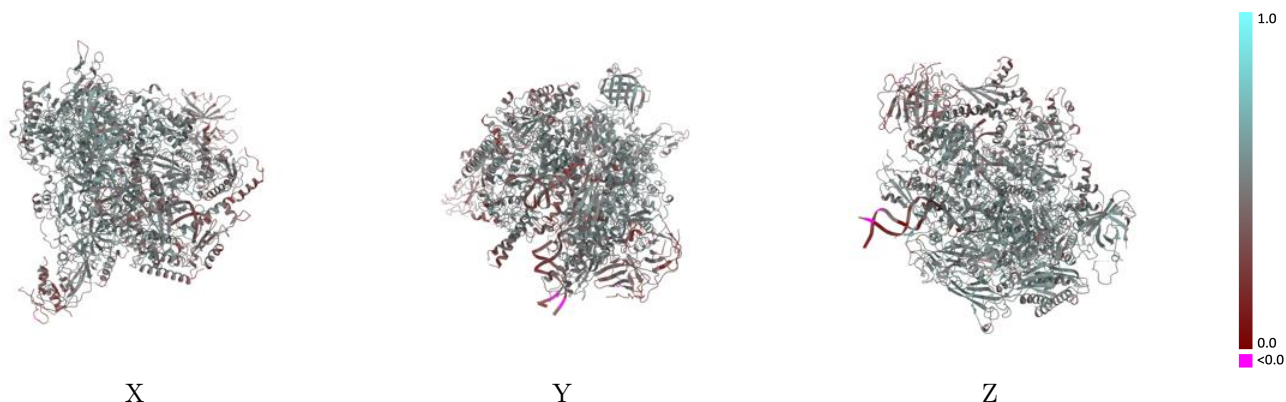
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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
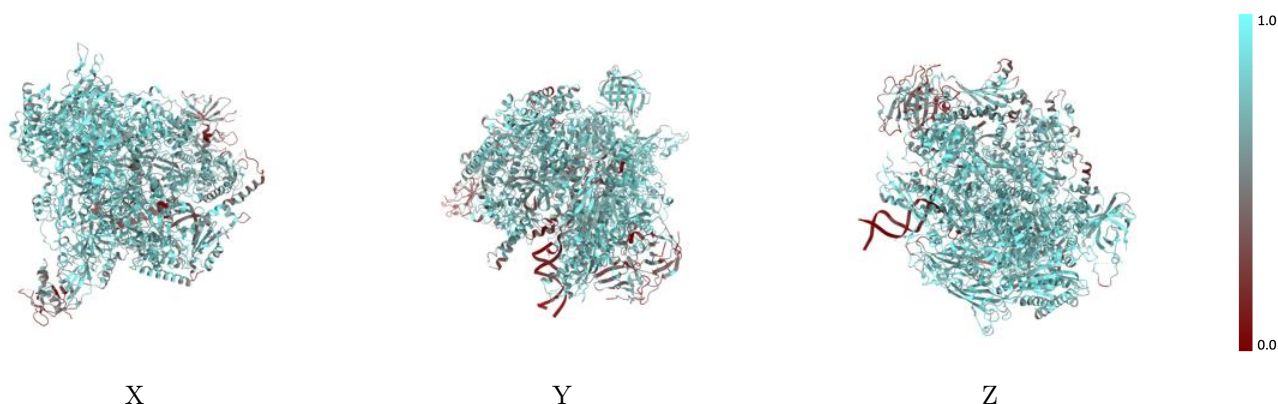
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.0388).

9.4 Atom inclusion ⓘ
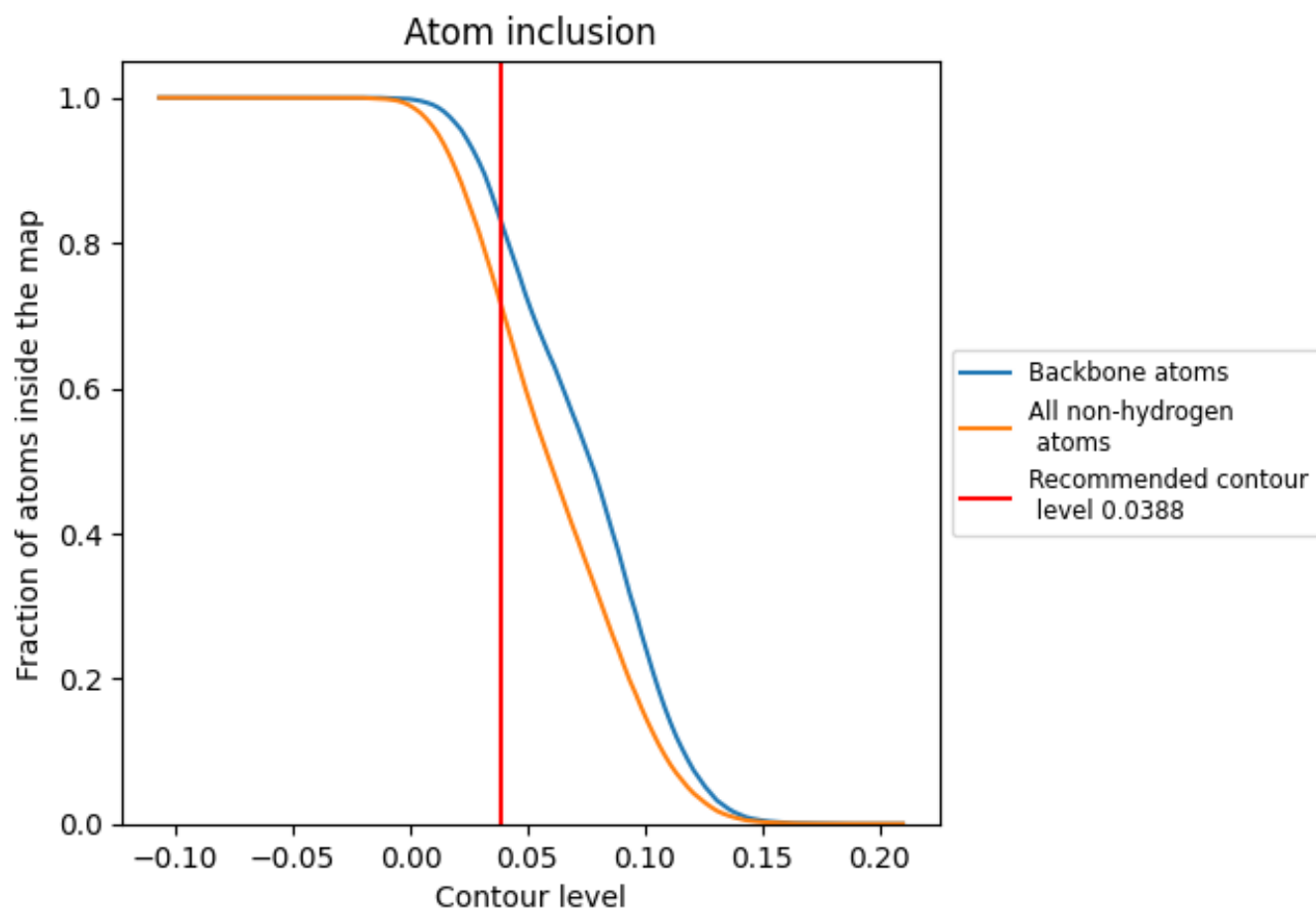
At the recommended contour level, 83% of all backbone atoms, 71% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.0388) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7143	0.4780
A	0.7414	0.4910
B	0.7908	0.5090
C	0.7811	0.5090
D	0.5896	0.4240
E	0.7137	0.4450
F	0.7665	0.5080
G	0.4963	0.4080
H	0.7554	0.5000
I	0.6260	0.4480
J	0.8152	0.5270
K	0.7879	0.5010
L	0.8029	0.5000
M	0.4539	0.3830
N	0.4338	0.3750
R	0.8615	0.5160
S	0.2788	0.2780
T	0.4620	0.3530

1.0

0.0
<0.0