WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB EM Validation Summary Report (i)

Jun 20, 2024 – 06:31 AM JST

PDB ID	:	7WTX
EMDB ID	:	EMD-32804
Title	:	Cryo-EM structure of a human pre-40S ribosomal subunit - State RRP12-B1
Authors	:	Cheng, J.; Lau, B.; Thoms, M.; Ameismeier, M.; Berninghausen, O.; Hurt, E.; Beckmann, R.
Deposited on	:	2022-02-05
Resolution	:	3.10 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.

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 3.10 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
Ramachandran outliers		0.1%
Sidechain outliers		0.2%
RNA backbone		0.45

Worse   *Better*

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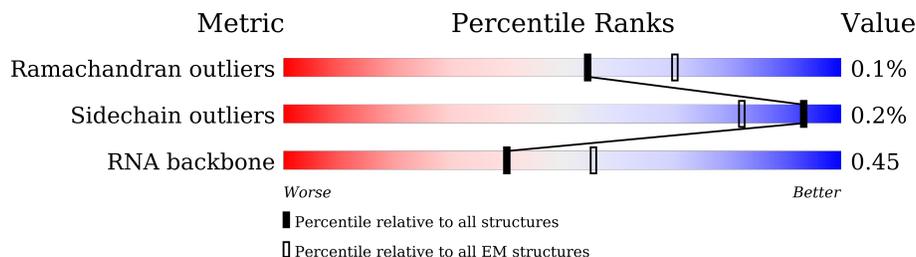
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	2	1873	59% 24% · 14%
2	F	204	92% 7%
3	M	132	14% 93% 7%
4	P	145	82% · 17%
5	Q	146	83% · 16%
6	R	135	83% 17%
7	S	152	86% 14%
8	T	145	99% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	Z	125	57% · 42%
10	c	69	87% · 12%
11	f	156	37% .. 61%
12	A	295	73% 27%
13	B	264	81% 19%
14	C	293	74% 26%
15	E	263	99%
16	G	249	91% · 8%
17	H	194	95% . .
18	I	208	98% ·
19	J	194	92% · 7%
20	L	158	95% . .
21	N	151	99% ·
22	O	151	89% 11%
23	V	83	98% ..
24	W	130	98% ..
25	X	143	95% . .
26	Y	133	93% 7%
27	b	84	98% ·
28	e	59	93% 7%
29	x	252	71% 29%
30	y	412	79% 21%
31	u	804	77% · 22%
32	w	437	63% 37%
33	t	475	28% 72%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	K	1297	5% · 72% · 28%

2 Entry composition ⓘ

There are 35 unique types of molecules in this entry. The entry contains 85074 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called 18S rRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	2	1605	Total 34277	C 15298	N 6153	O 11222	P 1604	0	0

- Molecule 2 is a protein called 40S ribosomal protein S5.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	F	189	Total 1495	C 934	N 284	O 270 S 7	0	0

- Molecule 3 is a protein called 40S ribosomal protein S12.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	M	123	Total 953	C 598	N 169	O 177 S 9	0	0

- Molecule 4 is a protein called 40S ribosomal protein S15.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	P	120	Total 984	C 625	N 184	O 168 S 7	0	0

- Molecule 5 is a protein called 40S ribosomal protein S16.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	Q	122	Total 969	C 616	N 180	O 170 S 3	0	0

- Molecule 6 is a protein called 40S ribosomal protein S17.

Mol	Chain	Residues	Atoms				AltConf	Trace
6	R	112	Total 911	C 571	N 169	O 168 S 3	0	0

- Molecule 7 is a protein called 40S ribosomal protein S18.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	S	130	Total	C	N	O	S	0	0
			1083	686	214	182	1		

- Molecule 8 is a protein called 40S ribosomal protein S19.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	T	144	Total	C	N	O	S	0	0
			1122	703	217	199	3		

- Molecule 9 is a protein called 40S ribosomal protein S25.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	Z	72	Total	C	N	O	S	0	0
			574	368	104	101	1		

- Molecule 10 is a protein called 40S ribosomal protein S28.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	c	61	Total	C	N	O	S	0	0
			471	288	95	86	2		

- Molecule 11 is a protein called Ubiquitin-40S ribosomal protein S27a.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	f	61	Total	C	N	O	S	0	0
			498	316	94	81	7		

- Molecule 12 is a protein called 40S ribosomal protein SA.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	A	216	Total	C	N	O	S	0	0
			1705	1083	299	315	8		

- Molecule 13 is a protein called 40S ribosomal protein S3a.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	B	213	Total	C	N	O	S	0	0
			1729	1098	309	308	14		

- Molecule 14 is a protein called 40S ribosomal protein S2.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	C	218	Total 1690	C 1094	N 289	O 297	S 10	0	0

- Molecule 15 is a protein called 40S ribosomal protein S4, X isoform.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	E	262	Total 2076	C 1324	N 386	O 358	S 8	0	0

- Molecule 16 is a protein called 40S ribosomal protein S6.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	G	230	Total 1862	C 1164	N 371	O 320	S 7	0	0

- Molecule 17 is a protein called 40S ribosomal protein S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	H	186	Total 1501	C 957	N 276	O 267	S 1	0	0

- Molecule 18 is a protein called 40S ribosomal protein S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	I	205	Total 1682	C 1056	N 331	O 290	S 5	0	0

- Molecule 19 is a protein called 40S ribosomal protein S9.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	J	180	Total 1499	C 955	N 300	O 242	S 2	0	0

- Molecule 20 is a protein called 40S ribosomal protein S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	L	151	Total 1229	C 782	N 230	O 211	S 6	0	0

- Molecule 21 is a protein called 40S ribosomal protein S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	N	149	Total 1202	C 770	N 228	O 203	S 1	0	0

- Molecule 22 is a protein called 40S ribosomal protein S14.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	O	135	Total 1009	C 618	N 198	O 187	S 6	0	0

- Molecule 23 is a protein called 40S ribosomal protein S21.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	V	82	Total 625	C 384	N 116	O 120	S 5	0	0

- Molecule 24 is a protein called 40S ribosomal protein S15a.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	W	129	Total 1034	C 659	N 193	O 176	S 6	0	0

- Molecule 25 is a protein called 40S ribosomal protein S23.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	X	141	Total 1098	C 693	N 219	O 183	S 3	0	0

- Molecule 26 is a protein called 40S ribosomal protein S24.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	Y	124	Total 1014	C 641	N 198	O 170	S 5	0	0

- Molecule 27 is a protein called 40S ribosomal protein S27.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	b	82	Total 640	C 402	N 118	O 113	S 7	0	0

- Molecule 28 is a protein called 40S ribosomal protein S30.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	e	55	Total	C	N	O	S	0	0
			438	271	95	71	1		

- Molecule 29 is a protein called RNA-binding protein PNO1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	x	178	Total	C	N	O	S	0	0
			1391	891	252	244	4		

- Molecule 30 is a protein called RNA-binding protein NOB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	y	325	Total	C	N	O	S	0	0
			2568	1622	473	463	10		

- Molecule 31 is a protein called Pre-rRNA-processing protein TSR1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	u	629	Total	C	N	O	S	0	0
			5062	3249	902	887	24		

- Molecule 32 is a protein called Bystin.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	w	275	Total	C	N	O	S	0	0
			2248	1455	397	387	9		

- Molecule 33 is a protein called Protein LTV1 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	t	132	Total	C	N	O	S	0	0
			1131	701	212	214	4		

- Molecule 34 is a protein called RRP12-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	K	940	Total	C	N	O	S	0	0
			7302	4680	1266	1311	45		

- Molecule 35 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
35	f	1	Total 1	Zn 1	0
35	y	1	Total 1	Zn 1	0

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: 18S rRNA

- Molecule 2: 40S ribosomal protein S5

Chain F:
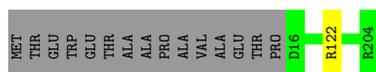
- Molecule 3: 40S ribosomal protein S12
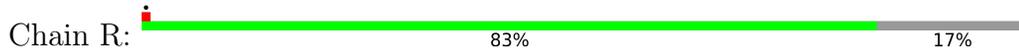
Chain M:

- Molecule 4: 40S ribosomal protein S15
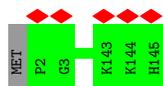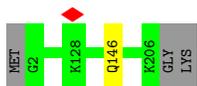
Chain P:

- Molecule 5: 40S ribosomal protein S16

Chain Q:

83%   16%

MET PRO SER LYS GLY PRO LEU Q8 K106 S129 LYS LYS PHE GLY GLY PRO GLY ALA ARG ALA ARG TYR GLN LYS SER TYR ARG

- Molecule 6: 40S ribosomal protein S17

Chain R:

83%   17%

MET G2 P65 VAL ARG GLY ILE SER ILE LYS LEU GLN GLU E76 T123 VAL GLY MET ASN PHE LYS THR PRO ARG GLY PRO VAL

- Molecule 7: 40S ribosomal protein S18

Chain S:

86%   14%

MET SER L3 E7 R132 GLY GLN HIS THR LYS THR THR GLY ARG ARG GLY ARG THR VAL GLY VAL SER LYS LYS LYS

- Molecule 8: 40S ribosomal protein S19

Chain T:

99%

MET P2 G3 K143 K144 H145

- Molecule 9: 40S ribosomal protein S25

Chain Z:

57%   42%

MET PRO PRO LYS ASP ASP LYS LYS LYS LYS ASP ALA GLY LYS SER ALA LYS LYS ASP LYS ASP PRO VAL ASN LYS SER GLY GLY LYS ALA LYS LYS LYS TRP SER LYS GLY LYS VAL ARG D42 L77 T113 LYS GLY GLY ASP ALA PRO ALA ALA PRO GLU ASP ALA

- Molecule 10: 40S ribosomal protein S28

Chain c:

87%   12%

MET ASP THR SER ARG VAL GLN P8 D37 L68 ARG

- Molecule 11: Ubiquitin-40S ribosomal protein S27a

Chain f:

37%   61%

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

ILE GLN LYS GLU SER THR LEU HIS LEU VAL LEU ARG LEU ARG GLY A77 Y106 LYS VAL ASP GLU ASN GLY LYS ILE SER ARG LEU ARG R119 F131 M132 R138 C149 PHE ASN LYS PRO GLU ASP LYS

• Molecule 12: 40S ribosomal protein SA

Chain A:

73%   27%

MET S2 A217 ALA GLU LYS VAL THR LYS GLU PHE GLN GLY GLU TRP THR ALA PRO ALA PRO GLU PHE THR ALA THR GLN PRO GLU VAL ALA ASP TRP SER GLU GLY VAL GLN VAL PRO SER VAL PRO ILE GLN GLN PHE PRO THR GLU ASP TRP SER ALA GLN PRO ALA THR

GLU ASP TRP SER ALA ALA PRO THR ALA GLN ALA THR GLU TRP VAL GLY ALA THR THR ASP TRP SER

• Molecule 13: 40S ribosomal protein S3a

Chain B:

81%   19%

MET ALA VAL GLY LYS ASN LYS ARG LEU THR LYS GLY GLY LYS LYS GLY ALA LYS LYS LYS V21 G233 GLU GLY SER SER GLY LYS ALA THR GLY ASP GLU THR GLY ALA LYS VAL GLU ALA ASP GLY TYR GLU PRO PRO VAL GLN GLU SER VAL

• Molecule 14: 40S ribosomal protein S2

Chain C:

74%   26%

MET ALA ASP ASP ALA GLY ALA ALA GLY GLY PRO GLY GLY PRO GLY GLY PRO GLY MET GLY ASN ARG GLY GLY PHE ARG GLY GLY PHE GLY SER GLY ILE ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ALA ARG GLY GLY LYS ALA GLU ASP LYS E59

G135 T276 HIS THR ARG VAL SER VAL GLN ARG THR GLN ALA PRO ALA VAL ALA THR THR

• Molecule 15: 40S ribosomal protein S4, X isoform

Chain E:

99%

MET A2 L38 G263

• Molecule 16: 40S ribosomal protein S6

Chain G:

91%   8%

M1 R98 L111 V180 K230 ARG ARG ARG LEU SER SER LEU ARG ALA SER THR SER LYS SER GLU SER SER GLN LYS

• Molecule 17: 40S ribosomal protein S7

Chain H:

95%

MET PHE SER SER ALA LYS I8 K15 D149 Q193 LEU

• Molecule 18: 40S ribosomal protein S8

Chain I: 98%

MET G2 K128 Q146 K206 GLY LYS

• Molecule 19: 40S ribosomal protein S9

Chain J: 92% 7%

MET P2 R138 L161 G181 GLN GLY GLY ALA GLY ALA GLY ASP ASP GLU GLU ASP

• Molecule 20: 40S ribosomal protein S11

Chain L: 95%

MET A2 E31 R69 K152 LYS GLN PHE GLN LYS PHE

• Molecule 21: 40S ribosomal protein S13

Chain N: 99%

MET G2 V150 ALA

• Molecule 22: 40S ribosomal protein S14

Chain O: 89% 11%

MET ALA PRO ARG LYS GLY LYS GLU LYS LYS GLU GLU GLN VAL ILE SER L17 L151

• Molecule 23: 40S ribosomal protein S21

Chain V: 98%

M1 N21 N82 PHE

• Molecule 24: 40S ribosomal protein S15a

Chain W: 98%

MET V2 D54 F130

• Molecule 25: 40S ribosomal protein S23

Chain X: 95%

MET G2 K60 P86 F105 I115 L130 R142 SER

• Molecule 26: 40S ribosomal protein S24

Chain Y: 93%   7%

MET ASN D3 G126 ALA GLY LYS LYS PRO LYS GLU

• Molecule 27: 40S ribosomal protein S27

Chain b: 98%

MET P2 Q83 HIS

• Molecule 28: 40S ribosomal protein S30

Chain e: 93%   7%

LYS VAL HIS GLY S5 S59

• Molecule 29: RNA-binding protein PNO1

Chain x: 71%   29%

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

PRO PRO LEU CYS GLY ASP GLY LEU LEU SER GLY LYS E73 D250 ARG PHE

• Molecule 30: RNA-binding protein NOB1

Chain y: 79%   21%

MET ALA PRO VAL GLU HIS V7 A47 V48 Y51 K71 P131 TYR LYS PRO LYS PRO PRO GLN GLU THR GLU LYS GLY HIS SER ALA CYS GLU PRO GLU ASN LEU GLU PHE SER SER PHE MET PHE TRP ARG ASN PRO LEU PRO ASN ILE ASP HIS GLU LEU GLN GLU LEU LEU

ILE ASP ARG GLY GLU ASP VAL PRO SER GLU GLU GLU GLU GLU GLU GLU ASN GLY PHE GLU ASP LYS ASP SER ASP ASP ASP GLY **G206** **L221 E222** GLN CYS ASP VAL PRO GLU ASP **V230** **R412**

• Molecule 31: Pre-rRNA-processing protein TSR1 homolog

Chain u:

77%　　　.　　　22%

MET ALA ALA HIS ARG PRO GLY PRO LEU LYS GLN GLN ASN LYS ALA HIS LYS GLY GLY ARG HIS ARG GLY ARG SER ALA GLN ARG ASP GLY LYS ARG ALA LEU LEU LYS THR LEU SER LYS LYS VAL ARG LYS **E47** **D105** ARG ASP THR **G109** **L222** SER **N266** **P313** ARG

GLY ILE LYS PRO GLN LYS ASP PRO ASP PRO MET ALA MET GLU ILE CYS ALA THR ASP ALA VAL ASP MET GLU GLU GLY LEU **K342** **R351** **P363** MET GLU GLY GLU GLN THR TRP PRO THR GLU GLU GLU LEU SER GLU ALA LYS ASP PHE LEU LYS GLU SER SER LYS TYR ASP LYS LYS

K392 **I405** LEU ASP GLY GLY SER GLN SER GLY GLY GLY ASP GLU TYR GLU THR ASP TYR ASP ASP MET GLU HIS GLU ASP ASP PHE MET GLU GLU SER GLN ASP GLU GLU SER GLU GLU GLU GLU TYR GLU THR MET THR ILE GLY GLU SER VAL HIS ASP LEU ASP LYS LYS

VAL **D464** **R499** **R624** **M716** **K742** **L745** **S792** GLU ILE SER SER THR VAL PRO GLN GLY GLY MET GLU

• Molecule 32: Bystin

Chain w:

63%　　　37%

MET PRO LYS PHE LYS ALA ALA ARG GLY VAL GLY GLY GLN GLU LYS ALA PRO ALA ASP LEU ALA GLY ASN ALA VAL ARG ALA GLY VAL GLU LYS ARG GLY ARG GLY THR GLY GLU ALA GLU **E47** **E72** HIS GLY THR GLY ASP LYS PRO ALA ALA PRO ARG

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

ARG ARG THR LEU ALA ASP ILE ILE MET GLU LYS LEU THR GLU LYS GLN THR GLU VAL GLU VAL MET SER GLU VAL SER GLY PHE PRO MET PRO GLN LEU ASP **P179** **N240** **R427** ASP VAL GLU ASP VAL PRO ILE THR VAL GLU

• Molecule 33: Protein LTV1 homolog

Chain t:

28%　　　72%

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

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

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

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

THR GLU TYR SER MET THR **S247** **E261** **E268 Q269 Y270** ASP ASP ASP GLU ILE GLY ALA LEU ASP LEU ASN ALA GLU GLY LEU GLU GLU SER ILE GLN VAL ASP **S291** **T314** LEU GLU PRO LEU GLU GLU ASP LEU MET ASN GLU LEU LEU ASP GLU GLU SER GLU GLU GLY GLU MET ILE THR

VAL VAL LEU GLU GLU ALA LYS GLU K347 Y369 GLN PRO LYS PRO LYS GLN ILE ARG ILE SER SER LYS THR GLY ILE PRO LEU ASN VAL LEU PRO LYS LYS GLY LEU THR ALA LYS GLN THR GLU ARG ILE GLN MET ILE ASN GLY SER SER ASP LEU PRO LYS VAL S414 K474 LEU

• Molecule 34: RRP12-like protein

Chain K:

5%

72%   28%

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

GLU LEU GLN SER GLY SER LEU ARG LEU GLY LYS SER GLU ALA PRO GLU THR PRO MET GLU GLU GLU ALA GLU GLU LEU VAL LEU THR GLU LYS SER SER GLY THR PHE LEU SER GLY LEU SER ASP CYS THR ASN VAL T106 Q111 W114 E115 S116 N117 S118 A119 A120 H121 K122 E123 I124

C125 A126 V127 L128 A129 A130 V131 T132 E133 V134 I135 R136 S137 Q138 G139 G140 K141 E142 T143 A150 T153 T154 M155 E156 A157 V158 E159 S160 P161 K176 R177 V178 K186 M198 A202 S203 S204 Q224 ASP LEU GLU ALA TRP GLY TYR P232 K265 G266 SER GLU PHE MET

PHE GLU LYS ALA PRO ALA HIS H278 S295 GLY GLY SER LYS GLU A301 L330 V378 S460 E499 H524 E560 Q605 A606 G607 I638 E653 A677 R684 V701 ALA ALA GLY ASP THR PRO A708 D725 P743 A744 S893 V911 G912 A913

K1043 ARG HIS ARG ALA LEU SER GLN ALA ALA VAL GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU PRO ALA GLN GLY LYS GLY ASP SER ILE GLU GLU ILE LEU LEU ALA ASP SER GLU ASP GLU GLU ASP ASN GLU GLU GLU GLU ARG SER ARG GLY LYS GLU GLN GLN ARG LYS LEU ALA

R1102 Q1103 R1104 E1111 GLY GLY GLY D1115 G1135 PRO GLY ARG GLY ARG LYS ASP HIS GLY PHE LYS VAL SER ALA ASP GLY TYR GLN ALA GLY LYS GLY SER GLY ILE LEU ARG PRO VAL ALA LYS LYS ALA MET PRO GLY ALA GLU GLU LYS GLY GLY TYR LYS GLU GLU MET ALA ASP PRO MET GLU ASP

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

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

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	27437	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; Relion	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	44	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.498	Depositor
Minimum map value	-0.260	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.011	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	381.24, 381.24, 381.24	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.059, 1.059, 1.059	Depositor

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z| > 5$	RMSZ	$\#|Z| > 5$
1	2	0.85	18/38323 (0.0%)	1.17	252/59711 (0.4%)
2	F	0.30	0/1516	0.61	0/2037
3	M	0.28	0/963	0.55	0/1291
4	P	0.37	0/1003	0.67	0/1341
5	Q	0.33	0/982	0.65	0/1318
6	R	0.29	0/922	0.61	0/1237
7	S	0.31	0/1100	0.65	0/1475
8	T	0.31	0/1142	0.60	0/1530
9	Z	0.29	0/580	0.67	1/780 (0.1%)
10	c	0.33	0/473	0.77	1/633 (0.2%)
11	f	0.35	0/508	0.74	2/670 (0.3%)
12	A	0.45	0/1742	0.66	0/2367
13	B	0.40	0/1756	0.63	0/2350
14	C	0.45	0/1726	0.66	0/2332
15	E	0.48	0/2118	0.70	1/2849 (0.0%)
16	G	0.40	0/1885	0.70	2/2510 (0.1%)
17	H	0.36	0/1524	0.63	1/2042 (0.0%)
18	I	0.48	0/1711	0.70	1/2282 (0.0%)
19	J	0.47	0/1524	0.72	0/2035
20	L	0.48	0/1250	0.65	0/1673
21	N	0.41	0/1226	0.63	0/1649
22	O	0.42	0/1022	0.73	0/1372
23	V	0.44	0/631	0.65	0/844
24	W	0.52	0/1051	0.73	0/1406
25	X	0.52	0/1116	0.71	2/1490 (0.1%)
26	Y	0.43	0/1031	0.68	0/1370
27	b	0.42	0/653	0.63	0/876
28	e	0.42	0/443	0.68	0/582
29	x	0.40	0/1413	0.68	0/1906
30	y	0.37	0/2618	0.67	0/3536
31	u	0.42	0/5187	0.65	4/7011 (0.1%)
32	w	0.30	0/2296	0.60	0/3102

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
33	t	0.35	0/1143	0.67	0/1515
34	K	0.28	0/7441	0.59	2/10080 (0.0%)
All	All	0.63	18/90019 (0.0%)	0.93	269/129202 (0.2%)

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
11	f	0	2
24	W	0	1
25	X	0	1
34	K	0	1
All	All	0	5

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	2	1196	A	C6-N1	-9.93	1.28	1.35
1	2	1196	A	N7-C5	-7.05	1.35	1.39
1	2	1221	G	C6-O6	-6.93	1.18	1.24
1	2	685	A	C6-N6	-6.83	1.28	1.33
1	2	492	C	C2-O2	-6.28	1.18	1.24

The worst 5 of 269 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	2	844	U	C5-C4-O4	24.83	140.80	125.90
1	2	1215	C	N3-C4-N4	-24.79	100.65	118.00
1	2	1221	G	N1-C6-O6	-24.26	105.34	119.90
1	2	844	U	N3-C4-O4	-22.20	103.86	119.40
1	2	914	U	C5-C4-O4	20.40	138.14	125.90

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
34	K	378	VAL	Peptide
24	W	54	ASP	Peptide
25	X	60	LYS	Peptide

Continued on next page...

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

Continued from previous page...

Mol	Chain	Res	Type	Group
11	f	131	PHE	Peptide
11	f	138	ARG	Sidechain

5.2　Too-close contacts　ⓘ

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	
2	F	187/204 (92%)	179 (96%)	8 (4%)	0	100	100
3	M	121/132 (92%)	113 (93%)	8 (7%)	0	100	100
4	P	118/145 (81%)	114 (97%)	4 (3%)	0	100	100
5	Q	120/146 (82%)	114 (95%)	6 (5%)	0	100	100
6	R	108/135 (80%)	104 (96%)	4 (4%)	0	100	100
7	S	128/152 (84%)	119 (93%)	9 (7%)	0	100	100
8	T	142/145 (98%)	142 (100%)	0	0	100	100
9	Z	70/125 (56%)	66 (94%)	4 (6%)	0	100	100
10	c	59/69 (86%)	55 (93%)	4 (7%)	0	100	100
11	f	57/156 (36%)	48 (84%)	9 (16%)	0	100	100
12	A	214/295 (72%)	206 (96%)	8 (4%)	0	100	100
13	B	211/264 (80%)	201 (95%)	10 (5%)	0	100	100
14	C	216/293 (74%)	207 (96%)	8 (4%)	1 (0%)	29	64
15	E	260/263 (99%)	250 (96%)	10 (4%)	0	100	100
16	G	228/249 (92%)	223 (98%)	5 (2%)	0	100	100
17	H	184/194 (95%)	177 (96%)	7 (4%)	0	100	100
18	I	203/208 (98%)	195 (96%)	8 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
19	J	178/194 (92%)	173 (97%)	3 (2%)	2 (1%)	14	46
20	L	149/158 (94%)	145 (97%)	4 (3%)	0	100	100
21	N	147/151 (97%)	147 (100%)	0	0	100	100
22	O	133/151 (88%)	126 (95%)	7 (5%)	0	100	100
23	V	80/83 (96%)	80 (100%)	0	0	100	100
24	W	127/130 (98%)	124 (98%)	3 (2%)	0	100	100
25	X	139/143 (97%)	136 (98%)	2 (1%)	1 (1%)	22	57
26	Y	122/133 (92%)	117 (96%)	5 (4%)	0	100	100
27	b	80/84 (95%)	78 (98%)	2 (2%)	0	100	100
28	e	53/59 (90%)	49 (92%)	4 (8%)	0	100	100
29	x	176/252 (70%)	172 (98%)	4 (2%)	0	100	100
30	y	319/412 (77%)	301 (94%)	18 (6%)	0	100	100
31	u	619/804 (77%)	588 (95%)	31 (5%)	0	100	100
32	w	271/437 (62%)	266 (98%)	5 (2%)	0	100	100
33	t	124/475 (26%)	108 (87%)	15 (12%)	1 (1%)	19	54
34	K	926/1297 (71%)	885 (96%)	40 (4%)	1 (0%)	51	83
All	All	6269/8138 (77%)	6008 (96%)	255 (4%)	6 (0%)	54	83

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
34	K	725	ASP
14	C	135	GLY
19	J	138	ARG
25	X	86	PRO
33	t	261	GLU

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	
2	F	159/170 (94%)	158 (99%)	1 (1%)	86	94
3	M	104/108 (96%)	104 (100%)	0	100	100
4	P	107/130 (82%)	106 (99%)	1 (1%)	78	91
5	Q	103/121 (85%)	102 (99%)	1 (1%)	76	90
6	R	101/122 (83%)	101 (100%)	0	100	100
7	S	114/132 (86%)	114 (100%)	0	100	100
8	T	114/115 (99%)	114 (100%)	0	100	100
9	Z	64/103 (62%)	64 (100%)	0	100	100
10	c	52/62 (84%)	52 (100%)	0	100	100
11	f	54/140 (39%)	54 (100%)	0	100	100
12	A	180/243 (74%)	180 (100%)	0	100	100
13	B	194/231 (84%)	194 (100%)	0	100	100
14	C	184/225 (82%)	184 (100%)	0	100	100
15	E	224/225 (100%)	224 (100%)	0	100	100
16	G	200/218 (92%)	199 (100%)	1 (0%)	88	94
17	H	167/174 (96%)	167 (100%)	0	100	100
18	I	178/180 (99%)	178 (100%)	0	100	100
19	J	160/168 (95%)	160 (100%)	0	100	100
20	L	135/142 (95%)	134 (99%)	1 (1%)	84	93
21	N	130/131 (99%)	130 (100%)	0	100	100
22	O	105/119 (88%)	105 (100%)	0	100	100
23	V	66/67 (98%)	65 (98%)	1 (2%)	65	85
24	W	112/113 (99%)	112 (100%)	0	100	100
25	X	113/115 (98%)	112 (99%)	1 (1%)	78	91
26	Y	108/115 (94%)	108 (100%)	0	100	100
27	b	74/76 (97%)	74 (100%)	0	100	100
28	e	45/48 (94%)	45 (100%)	0	100	100
29	x	150/208 (72%)	150 (100%)	0	100	100
30	y	285/367 (78%)	284 (100%)	1 (0%)	91	96
31	u	550/705 (78%)	548 (100%)	2 (0%)	91	96
32	w	240/370 (65%)	239 (100%)	1 (0%)	91	96
33	t	126/434 (29%)	126 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
34	K	805/1094 (74%)	803 (100%)	2 (0%)	93	97
All	All	5503/6971 (79%)	5490 (100%)	13 (0%)	93	97

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
30	y	71	LYS
31	u	499	ARG
34	K	1104	ARG
32	w	240	ASN
34	K	186	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
32	w	240	ASN
32	w	251	ASN
33	t	308	ASN
9	Z	106	GLN
2	F	95	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	1589/1873 (84%)	441 (27%)	29 (1%)

5 of 441 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	3	C
1	2	4	C
1	2	5	U
1	2	20	G
1	2	23	G

5 of 29 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	2	958	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	2	1726	G
1	2	1316	C
1	2	1558	C
1	2	1231	C

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

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

5.5　Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6　Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 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 ⓘ

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-32804. 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
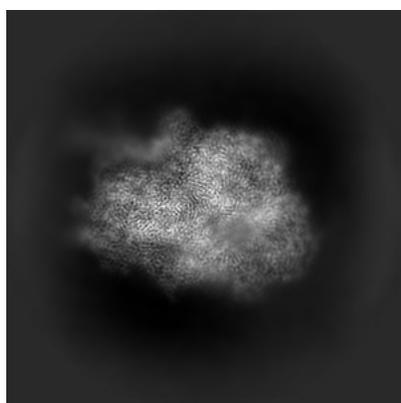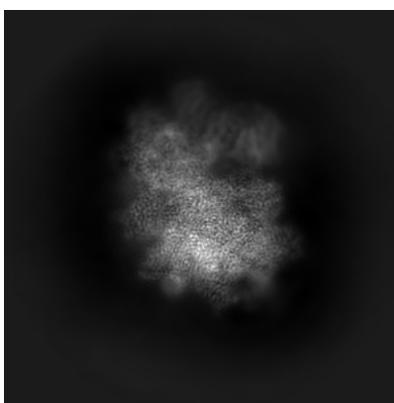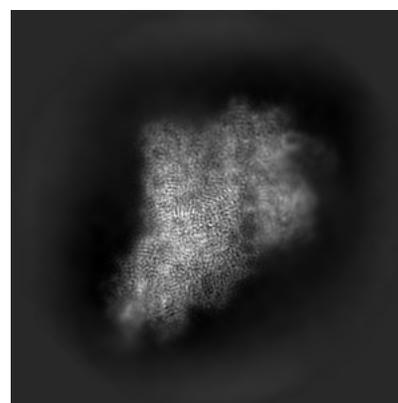
The images above show the map projected in three orthogonal directions.

6.2　Central slices ⓘ
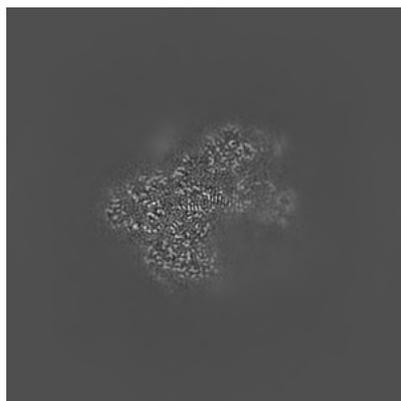
6.2.1　Primary map
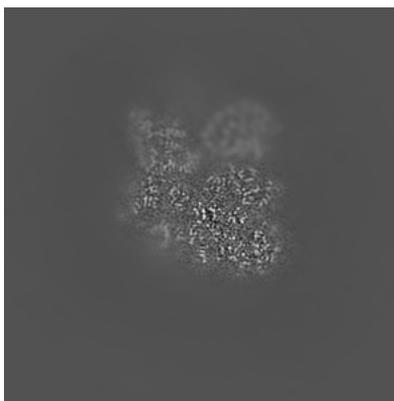
X Index: 180　　　　　　　　　　　Y Index: 180　　　　　　　　　　　Z Index: 180
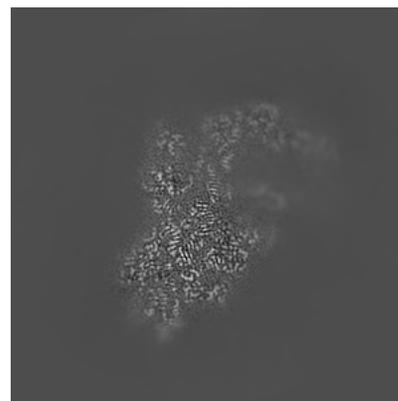
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ
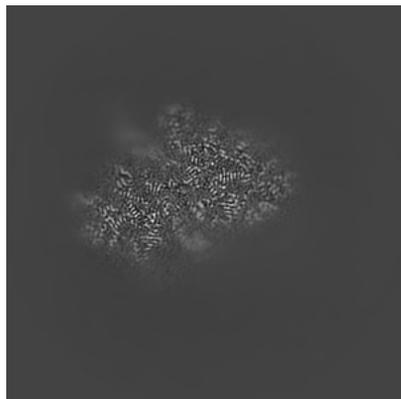
6.3.1 Primary map
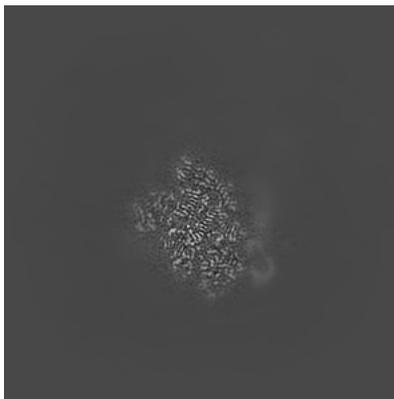
| X Index: 144 | Y Index: 133 | Z Index: 181 |
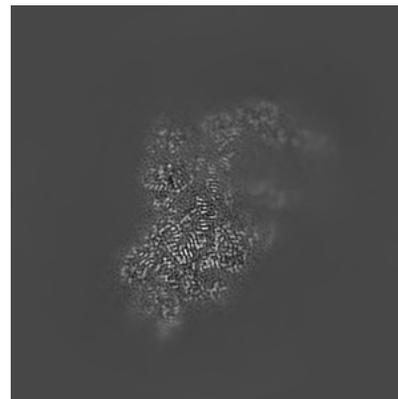
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map

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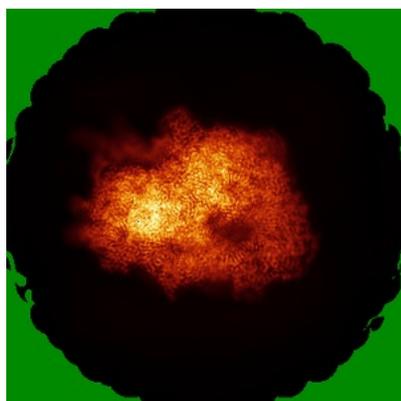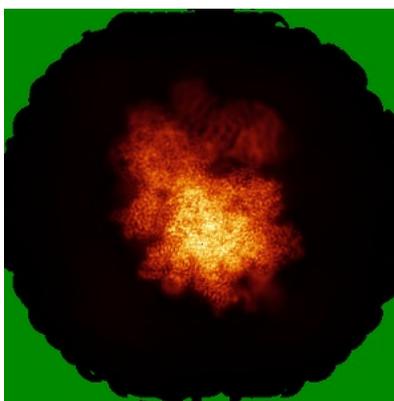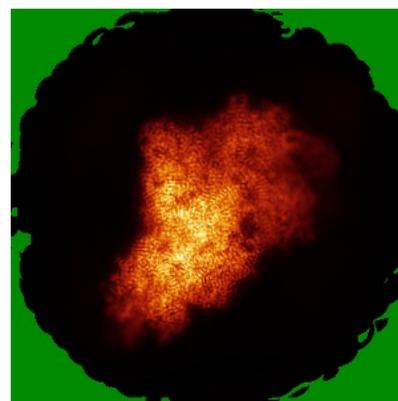
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X   Y   Z
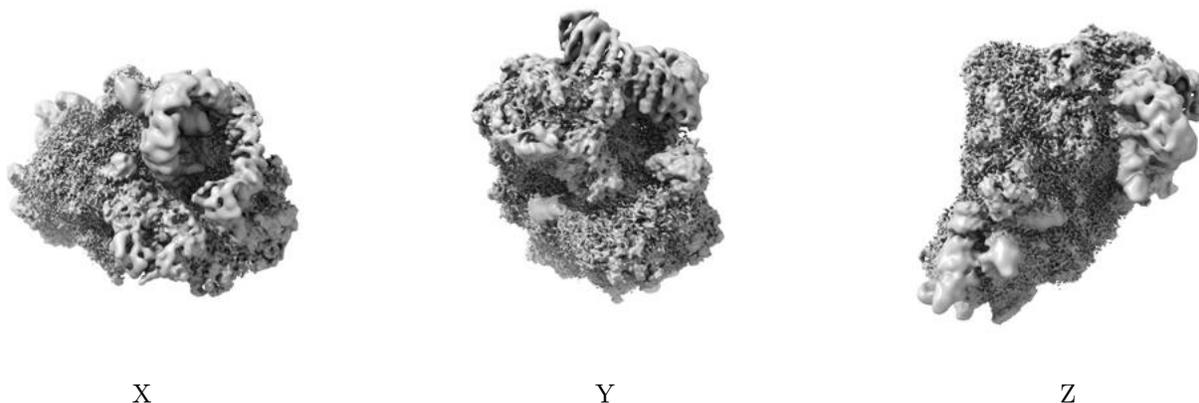
The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.6 Mask visualisation (i)

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

7   Map analysis ⓘ

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

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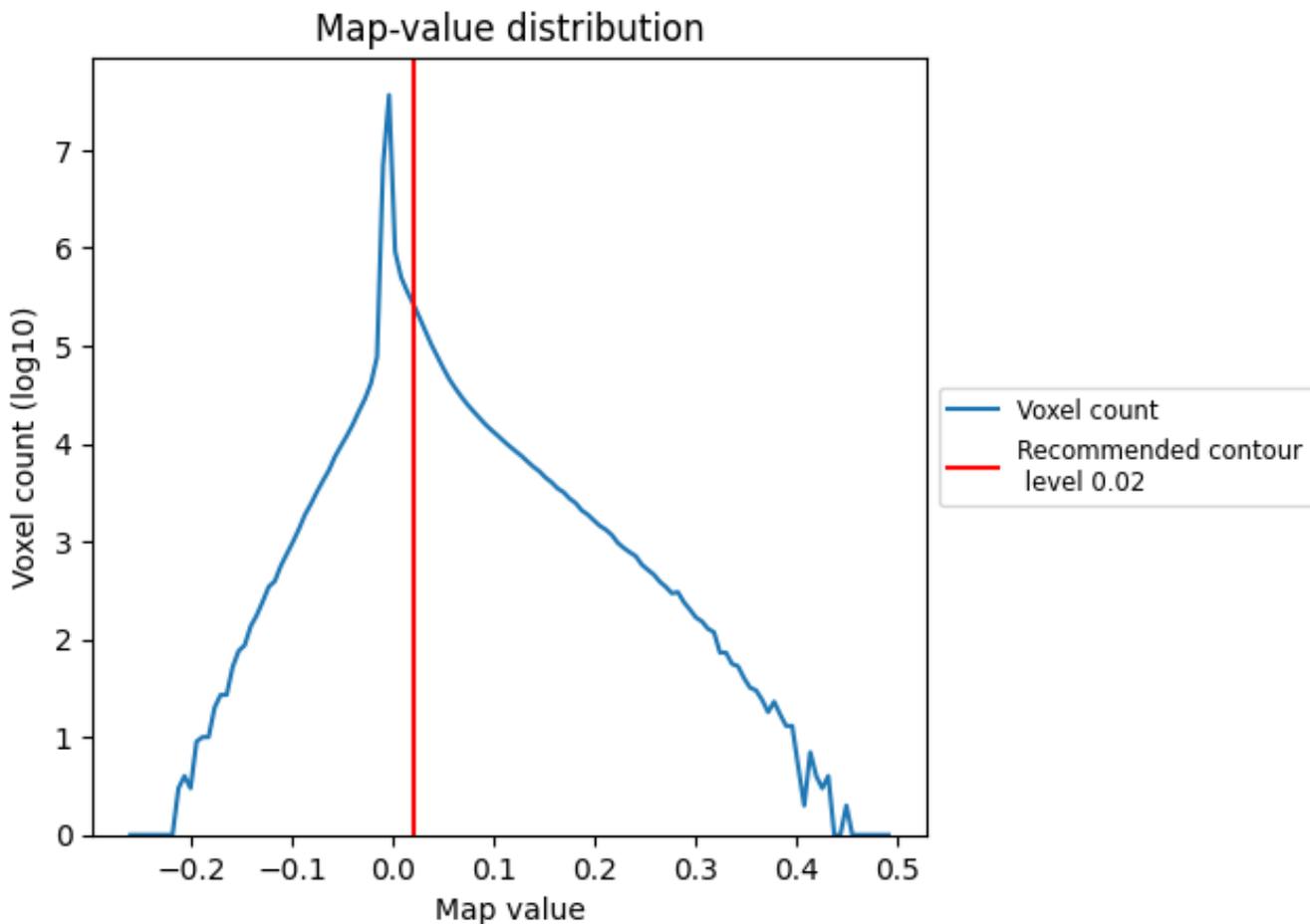
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2   Volume estimate ⓘ
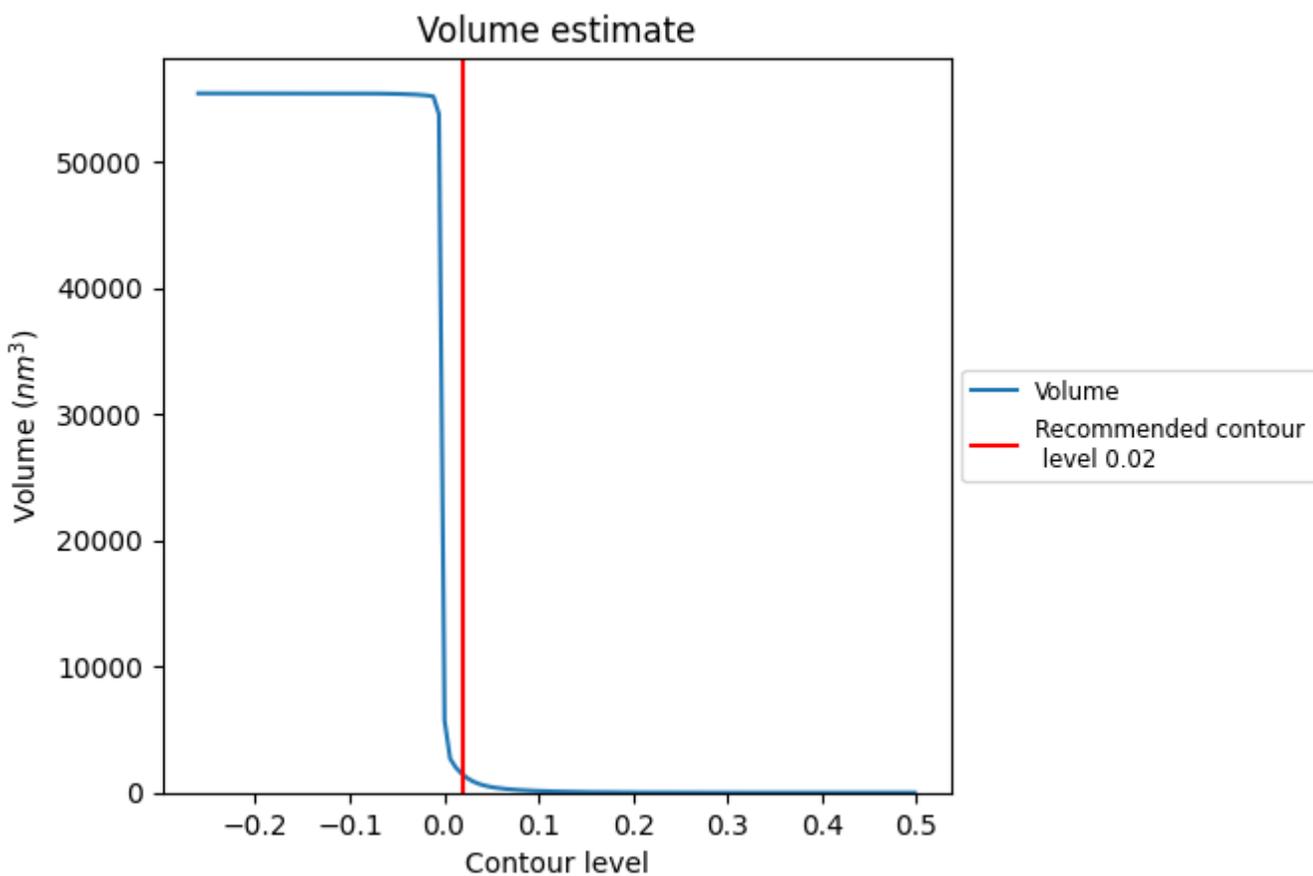
The volume at the recommended contour level is 1418 nm^3; this corresponds to an approximate mass of 1281 kDa.

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 ⓘ
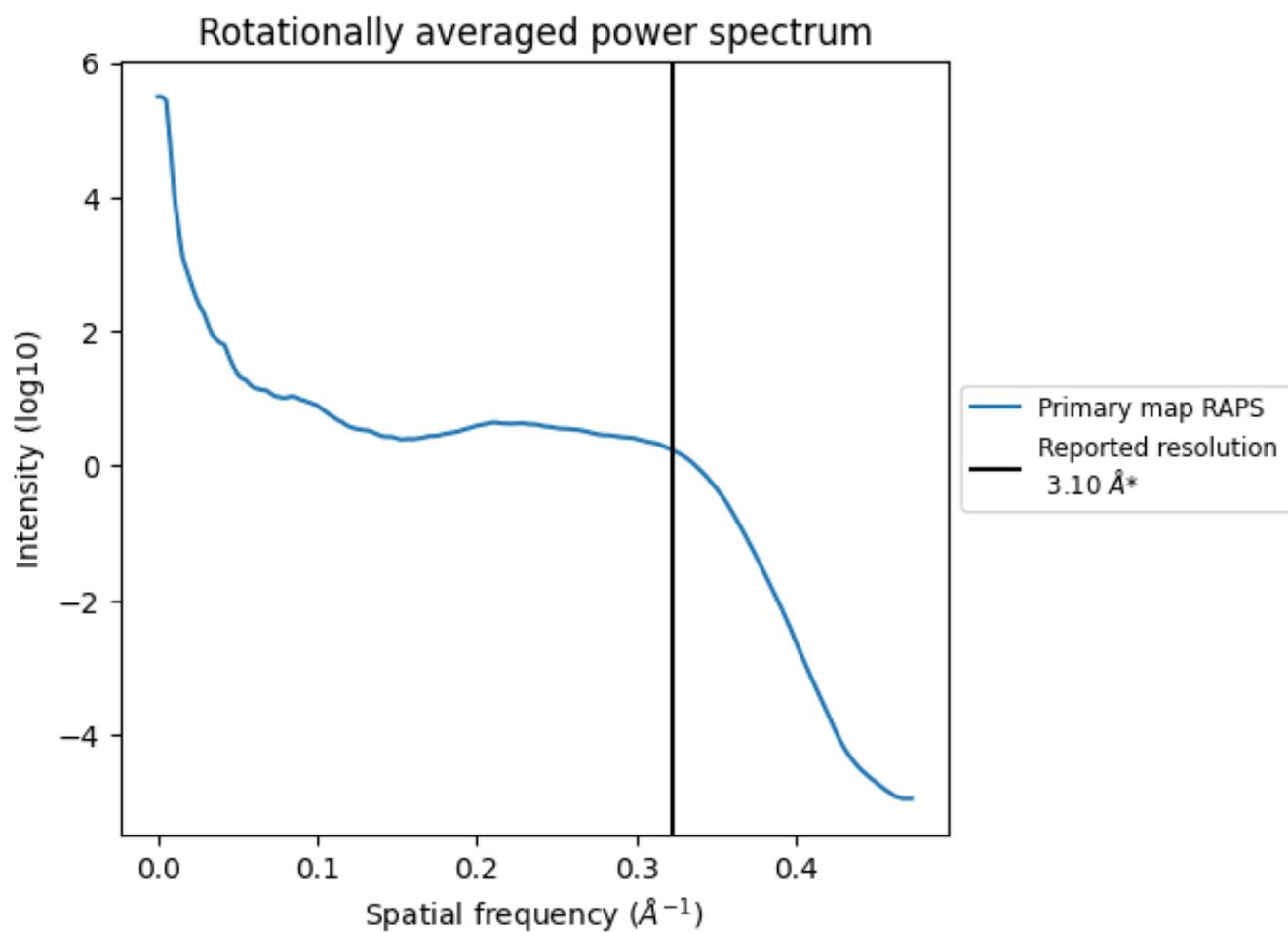
*Reported resolution corresponds to spatial frequency of 0.323 Å^{-1}

8   Fourier-Shell correlation (i)

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 (i)
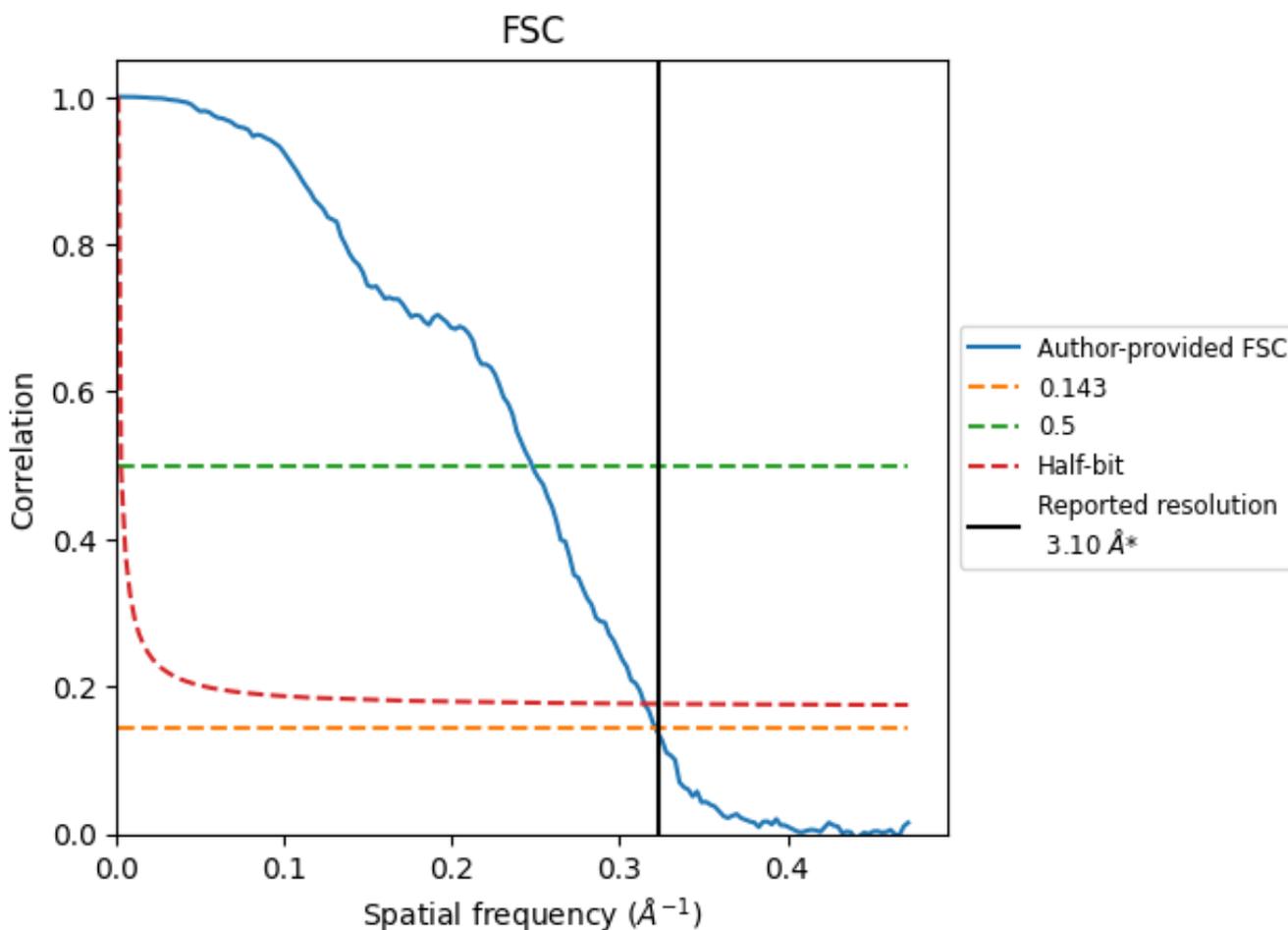
*Reported resolution corresponds to spatial frequency of 0.323 $Å^{-1}$

8.2   Resolution estimates ⓘ

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.10	-	-
Author-provided FSC curve	3.11	4.04	3.18
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-32804 and PDB model 7WTX. Per-residue inclusion information can be found in section 3 on page 11.

9.1 Map-model overlay ⓘ

X   Y   Z
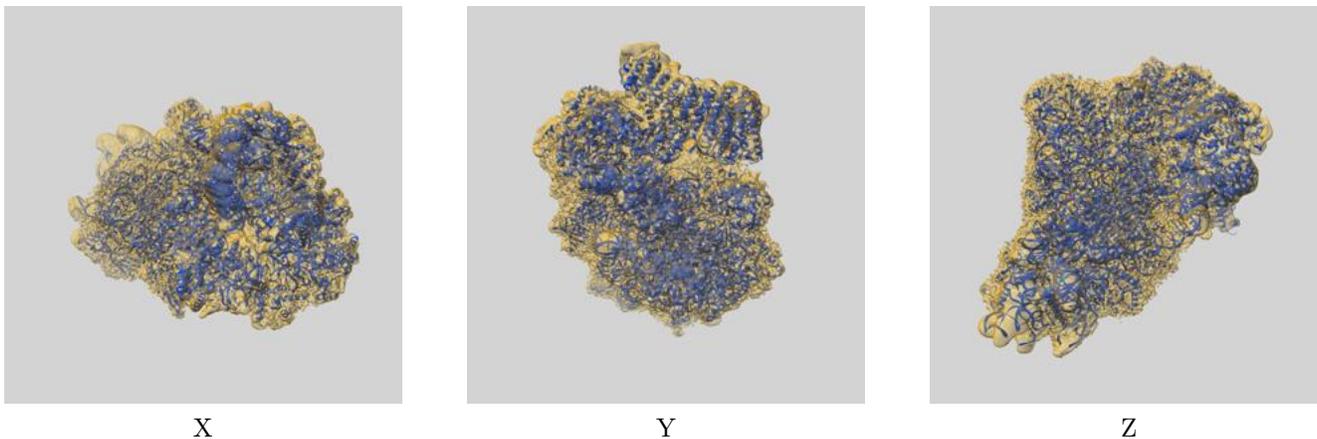
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2   Q-score mapped to coordinate model ⓘ
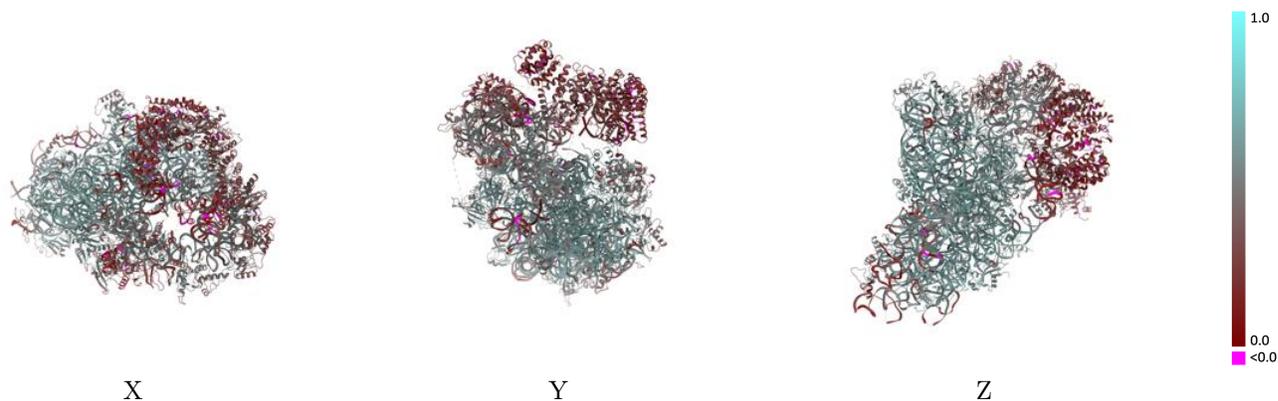
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 ⓘ
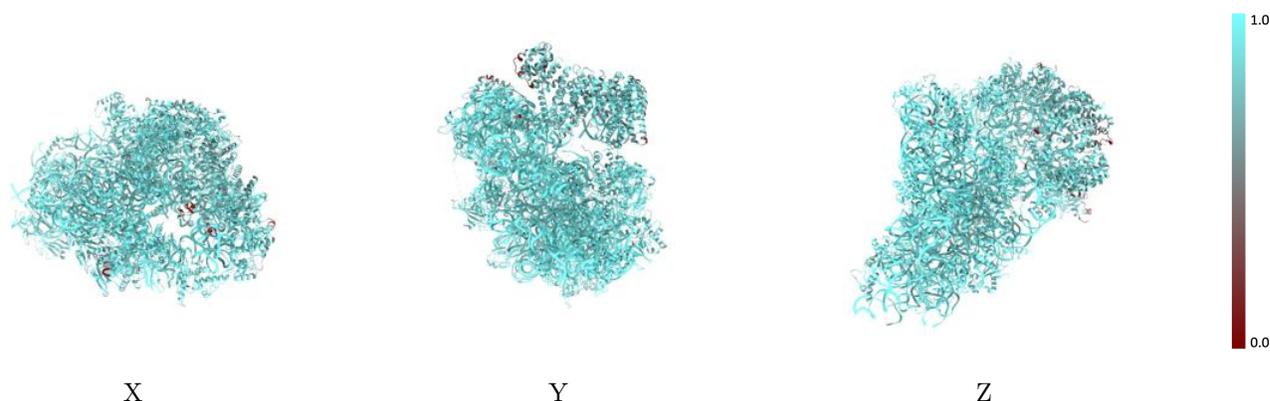
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.02).

9.4   Atom inclusion ⓘ
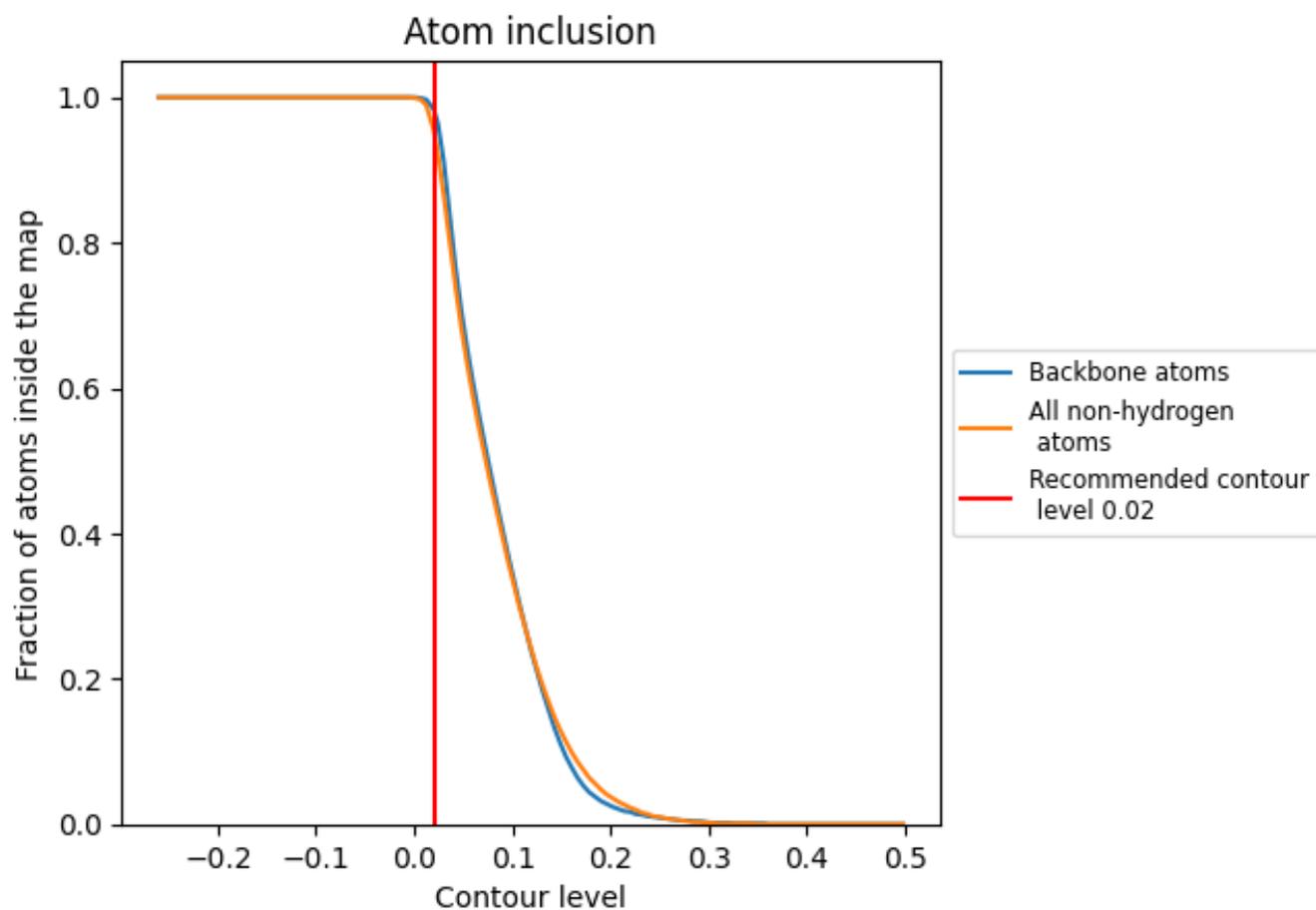
At the recommended contour level, 98% of all backbone atoms, 95% of all non-hydrogen atoms, are inside the map.

9.5　Map-model fit summary　ⓘ

The table lists the average atom inclusion at the recommended contour level (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9540	0.4780
2	0.9880	0.5110
A	0.9830	0.5700
B	0.9730	0.5390
C	0.9880	0.5800
E	0.9840	0.5910
F	0.9440	0.4290
G	0.9620	0.5040
H	0.9390	0.4720
I	0.9600	0.5260
J	0.9940	0.6000
K	0.7860	0.1970
L	0.9610	0.5710
M	0.7220	0.2240
N	0.9930	0.5760
O	0.9690	0.5220
P	0.9390	0.4610
Q	0.8940	0.4160
R	0.9120	0.3240
S	0.9110	0.4010
T	0.8860	0.3800
V	0.9800	0.5730
W	0.9910	0.6160
X	0.9980	0.6220
Y	0.9930	0.5730
Z	0.8410	0.3270
b	0.9860	0.5550
c	0.9490	0.4070
e	0.9480	0.5300
f	0.8990	0.2570
t	0.9520	0.4070
u	0.9780	0.5400
w	0.9410	0.4050
x	0.9880	0.5620
y	0.9430	0.4510

1.0

0.0
<0.0