WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB EM Validation Report (i)

Oct 18, 2021 – 11:07 am BST

PDB ID	:	7OOP
EMDB ID	:	EMD-13010
Title	:	Pol II-CSB-CSA-DDB1-UVSSA-PAF-SPT6 (Structure 3)
Authors	:	Kokic, G.; Cramer, P.
Deposited on	:	2021-05-28
Resolution	:	2.90 Å(reported)

This is a Full wwPDB EM Validation 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.

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 2.90 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		0.1%
RNA backbone		0.33

Worse   *Better*

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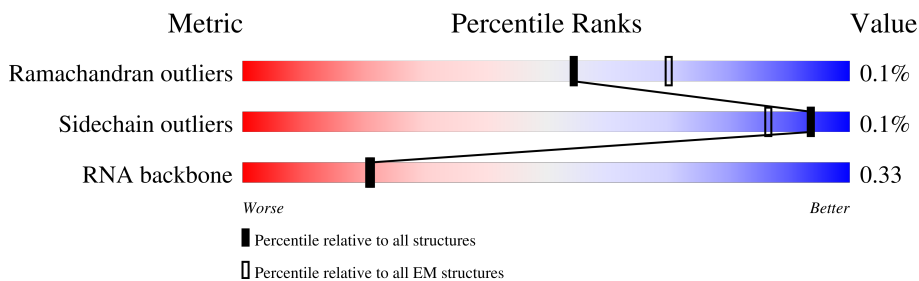
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	A	1970	5% / 71% / 28%
2	B	1174	6% / 96% / •
3	C	275	• / 94% / 5%
4	D	142	87% / 90% / 10%
5	E	210	• / 99%
6	F	127	6% / 65% / 35%
7	G	172	42% / 97% / ••
8	H	150	• / 99% / •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	14% · 94% · 6%
10	J	67	· 100%
11	K	117	· 98% ·
12	L	58	· 79% · 21%
13	M	1726	5% · 46% · 53%
14	N	47	11% · 74% · 26%
15	P	45	24% · 18% · 16% · 13% · 53%
16	R	40	68% · 100%
17	S	1173	57% · 76% · 24%
18	T	47	17% · 96% ·
19	U	666	16% · 15% · 84%
20	V	531	38% · 40% · 59%
21	Y	305	58% · 98% ·
22	Z	531	6% · 8% · 92%
23	a	396	· 92% · 8%
24	b	1493	11% · 35% · 64%
25	c	709	20% · 20% · 80%
26	d	1140	23% · 96% ·

2 Entity composition ⓘ

There are 28 unique types of molecules in this entry. The entry contains 62441 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 II subunit RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1412	Total 11179	C 7033	N 2002	O 2074	S 70	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1131	Total 9052	C 5727	N 1592	O 1669	S 64	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	260	Total 2089	C 1309	N 359	O 415	S 6	0	0

- Molecule 4 is a protein called RPOL4c domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	128	Total 1013	C 636	N 172	O 201	S 4	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	209	Total 1720	C 1089	N 300	O 323	S 8	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	82	Total 657	C 418	N 113	O 121	S 5	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	171	Total	C	N	O	S	0	0
			1334	867	216	243	8		

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	148	Total	C	N	O	S	0	0
			1186	750	194	237	5		

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	117	Total	C	N	O	S	0	0
			949	587	169	182	11		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	67	Total	C	N	O	S	0	0
			533	345	90	92	6		

- Molecule 11 is a protein called RNA_pol_L_2 domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	115	Total	C	N	O	S	0	0
			920	593	152	173	2		

- Molecule 12 is a protein called RNA polymerase II subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	46	Total	C	N	O	S	0	0
			388	241	75	66	6		

- Molecule 13 is a protein called Transcription elongation factor SPT6.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	810	Total	C	N	O	S	0	0
			6648	4226	1155	1234	33		

- Molecule 14 is a DNA chain called NTS.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	35	Total 727	C 344	N 142	O 206	P 35	0	0

- Molecule 15 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	P	21	Total 454	C 204	N 89	O 140	P 21	0	0

- Molecule 16 is a protein called LEO1 helix.

Mol	Chain	Residues	Atoms			AltConf	Trace
16	R	40	Total 160	C 80	N 40	O 40	0	0

- Molecule 17 is a protein called RNA polymerase-associated protein CTR9 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	S	890	Total 3560	C 1780	N 890	O 890	0	0

- Molecule 18 is a DNA chain called TS.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	47	Total 947	C 453	N 159	O 288	P 47	0	0

- Molecule 19 is a protein called RNA polymerase-associated protein LEO1.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	U	104	Total 416	C 208	N 104	O 104	0	0

- Molecule 20 is a protein called RNA polymerase II-associated factor 1 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	V	217	Total 868	C 434	N 217	O 217	0	0

- Molecule 21 is a protein called WD repeat-containing protein 61.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	Y	300	Total 1200	C 600	N 300	O 300	0	0

- Molecule 22 is a protein called Parafibromin.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	Z	43	Total 172	C 86	N 43	O 43	0	0

- Molecule 23 is a protein called DNA excision repair protein ERCC-8.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	a	365	Total 2849	C 1775	N 507	O 548	S 19	0	0

- Molecule 24 is a protein called DNA excision repair protein ERCC-6.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	b	534	Total 4356	C 2803	N 763	O 769	S 21	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
b	538	ARG	LYS	conflict	UNP Q03468

- Molecule 25 is a protein called UV-stimulated scaffold protein A.

Mol	Chain	Residues	Atoms				AltConf	Trace
25	c	141	Total 564	C 282	N 141	O 141	0	0

- Molecule 26 is a protein called DNA damage-binding protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	d	1096	Total 8491	C 5397	N 1423	O 1625	S 46	0	0

- Molecule 27 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 28 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
28	A	1	Total 1	Mg 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: DNA-directed RNA polymerase II subunit RPB1
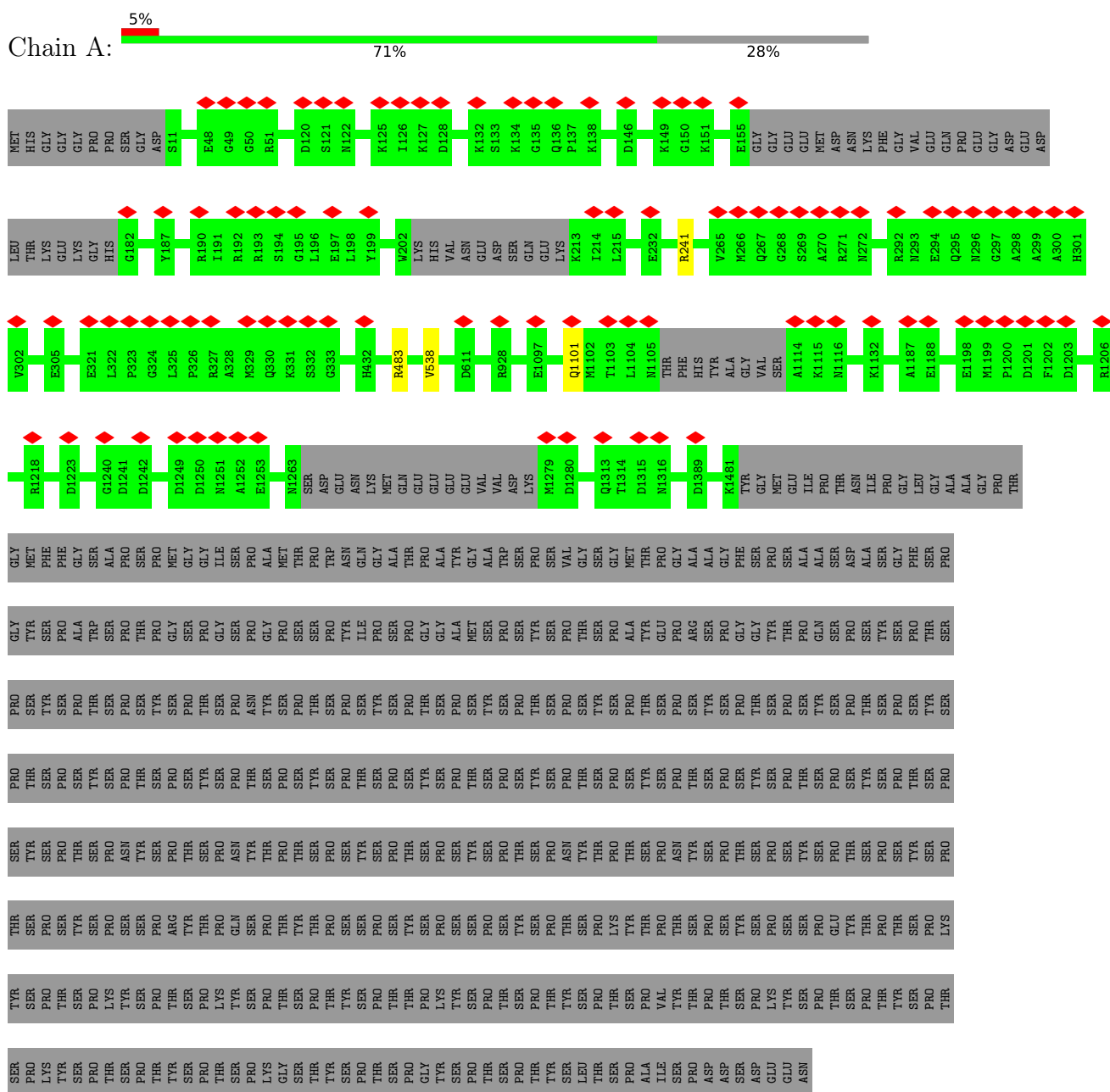
• Molecule 2: DNA-directed RNA polymerase subunit beta
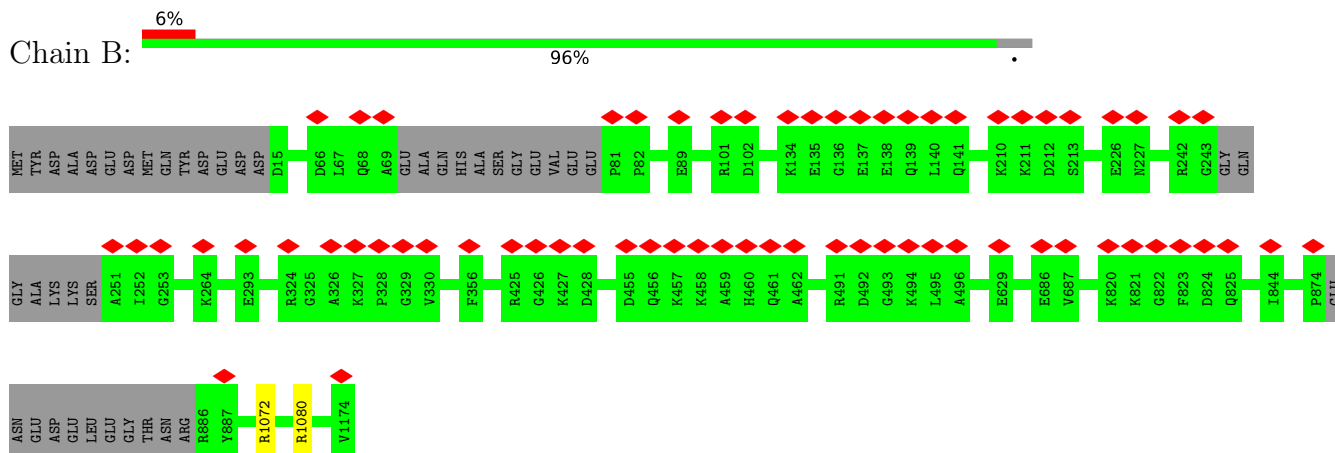
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3
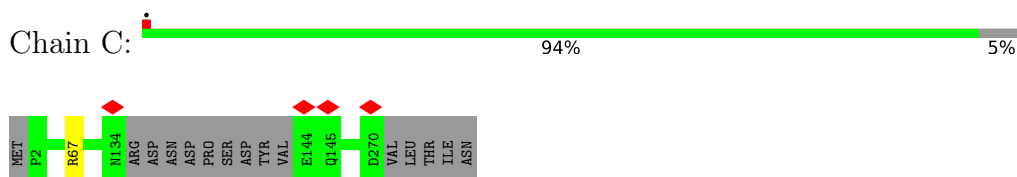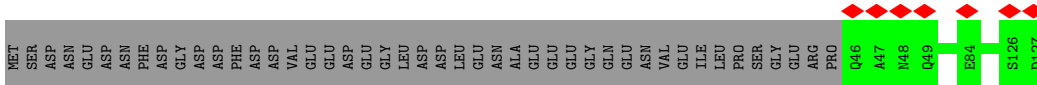

• Molecule 4: RPOL4c domain-containing protein
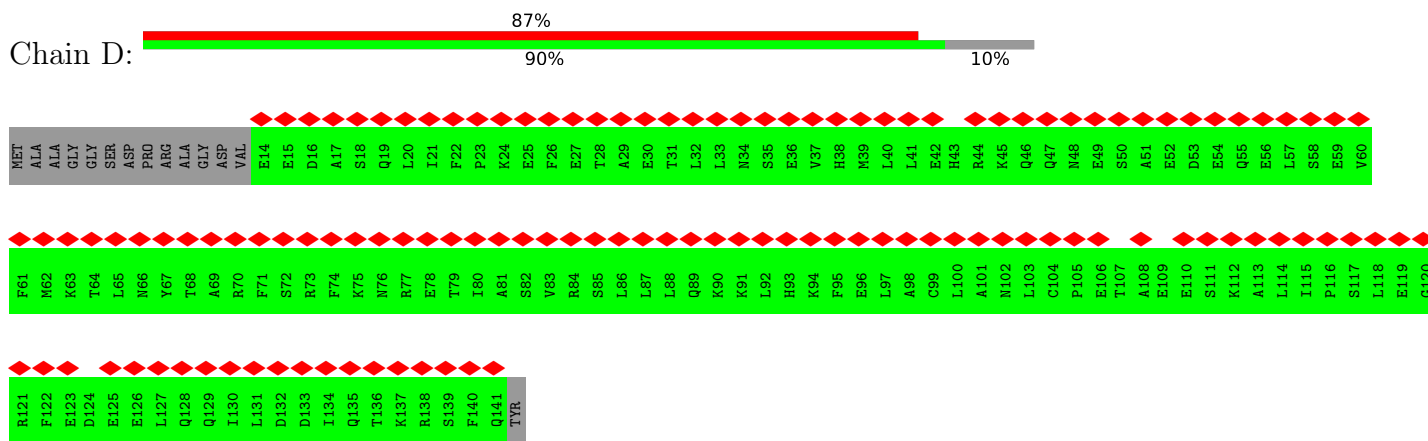
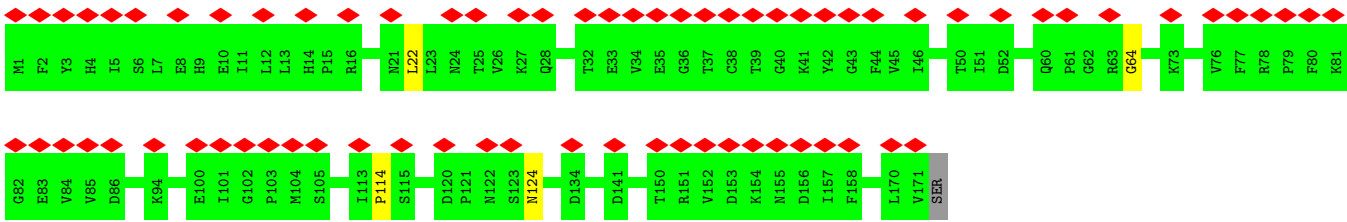• Molecule 5: DNA-directed RNA polymerase II subunit E
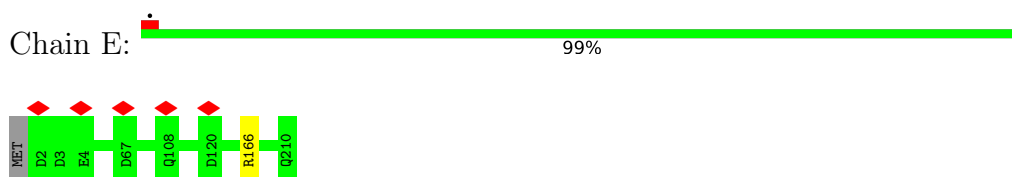
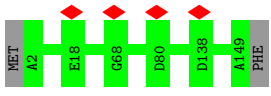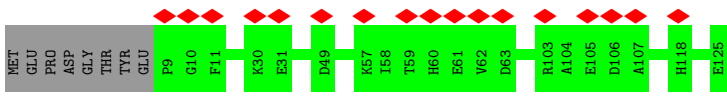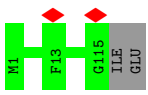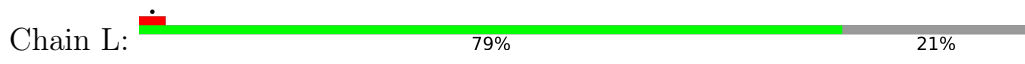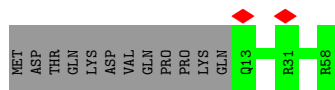• Molecule 6: DNA-directed RNA polymerase II subunit F

MET SER ASP ASN GLU ASP ASN PHE ASP GLY ASP ASP PHE ASP ASP VAL GLU GLU GLU ASP GLY GLY LEU ASP ASP LEU GLU ASN ALA GLU GLU GLU GLY GLN GLU GLU VAL GLU ILE LEU PRO SER GLY GLU ARG PRO Q46 A47 N48 Q49 E84 S126 D127

• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G: 42% 97%

M1 F2 Y3 H4 I5 S6 L7 E8 H9 E10 I11 L12 L13 H14 P15 R16 N21 L22 L23 N24 T25 V26 K27 Q28 T32 E33 V34 E35 G36 T37 C38 T39 G40 K41 Y42 G43 F44 V45 I46 T50 I51 D52 Q60 P61 G62 R63 G64 K73 V76 F77 R78 P79 F80 K81

G82 E83 V84 V85 D86 K94 E100 I101 G102 P103 M104 S105 I113 P114 S115 D120 P121 N122 S123 N124 D134 D141 T150 R151 V152 D153 K154 N155 D156 I157 F158 L170 V171 SER

• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 99%

MET A2 E18 G68 D80 D138 A149 PHE

• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I: 14% 94% 6%

MET GLU PRO ASP GLY THR TYR GLU P9 G10 F11 K30 E31 D49 K57 I58 T59 H60 E61 V62 D63 R103 A104 E105 D106 A107 H118 E125

• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J: 100%

M1 K67

• Molecule 11: RNA_pol_L_2 domain-containing protein

Chain K: 98%

M1 F13 G115 ILE GLU

• Molecule 12: RNA polymerase II subunit K

Chain L: 79% 21%

MET ASP THR GLN LYS ASP VAL GLN PRO PRO LYS GLN Q13 R31 R58

• Molecule 13: Transcription elongation factor SPT6

Chain M:

5%

46%   •   53%

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

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

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

ALA GLN TRP ALA SER SER GLN GLY TYR GLY TYR GLY GLY SER GLY SER GLY SER ALA TYR HIS VAL PHE PRO THR PRO ALA GLN GLN PRO VAL ALA THR PRO LEU MET THR PRO SER TYR SER THR THR PRO SER GLN PRO ILE THR THR PRO GLN TYR HIS GLN LEU GLN ALA SER

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

PRO SER PRO MET ILE GLU SER THR PRO MET SER ILE ALA GLY ASP ALA THR PRO LEU LEU ASP GLU GLU MET ASP ARG

● Molecule 14: NTS

Chain N:

● Molecule 15: RNA

Chain P:

● Molecule 16: LEO1 helix

Chain R:

● Molecule 17: RNA polymerase-associated protein CTR9 homolog

Chain S:

GLY SER ASP SER ASP

• Molecule 18: TS

Chain T:

17%

96%

C1 G30 G31 C32 T33 A34 T35 G36 A46 G47

• Molecule 19: RNA polymerase-associated protein LEO1

Chain U:

16%

15%

84%

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

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

HIS ARG SER ASP GLY GLY SER HIS HIS SER GLU ALA GLU GLY SER GLU LYS ALA HIS SER ASP SER GLU LYS TRP GLY ARG GLU GLU ASP LYS ALA ASN SER ASP GLN GLN LEU SER GLU GLU GLU ASP HIS LYS SER GLU SER ALA ARG GLY SER SER GLU GLU GLN GLU LYS GLN GLY SER ASP GLU LYS LEU GLN ASN SER ASP

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

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

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

GLU ILE PRO LYS VAL ASN THR ASP LEU **G370 N371 D372 L373 Y374 F375 V376 K377 L378 P379 N380 F381 L382 S383 V384 E385 P386 R387 P388 F389 D390 P391 Q392 Y393 Y394 E395 D396 E397 F398 E399** ASP GLU MET LEU ASP **E406 E407 G408 R409 T410 R411 L412 K413 L414 K415 V416 E417 N418** THR ILE

ARG TRP ARG ILE ARG ARG ASP GLU GLU GLY ASN GLU ILE LYS GLU SER **N437 A438 R439 I440 V441 K442 W443 S444 D445 G446 S447 N448 S449 L450 H451 L452 G453 N454 E455 V456 F457 D458 V459 Y460 K461 A462 P463 L464 Q465 G466 D467 H468 N469 H470 L471 F472 I473 R474 Q475 C476 T477 G478 L479 Q480**

G481 Q482 A483 V484 F485 K486 T487 K488 L489 T490 F491 R492 P493 H494 S495 T496 D497 SER ALA THR HIS ARG LYS MET THR THR LEU SER LEU ALA ASP ARG CYS SER LYS THR GLN LYS ILE ARG LEU LEU PRO MET ALA GLY ARG ASP PRO GLU CYS GLN ARG THR GLU MET ILE LYS LYS GLU GLU

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

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

GLU GLU GLU ASP ASP

• Molecule 20: RNA polymerase II-associated factor 1 homolog

Chain V:

38%

40%

59%

• Molecule 21: WD repeat-containing protein 61

Chain Y:

58%

98%

• Molecule 22: Parafibromin

6%

Chain Z:
8%　　　　　　　　　　　　　92%

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

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

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

LYS ILE ALA ALA ILE LYS ALA LYS ILE MET ALA LYS LYS ARG SER THR ILE LYS THR ASP LEU ASP ASP ASP ILE THR ALA LEU LYS GLN ARG SER PHE VAL ASP ALA GLU **V218 D219 V220 T221 R222 D223 I224 V225 S226 R227 E228** **W231** **T235 T236** **Q239 S240 T241 G242 K243 N244**

F245 S246 K247 N248 I249 F250 A251 I252 L253 Q254 S255 V256 K257 A258 R259 E260 GLU GLY ARG ALA PRO GLU GLN ARG ALA PRO ALA PRO ASN ALA ALA GLN ALA PRO VAL ASP PRO THR LEU ARG THR LYS GLN PRO ILE PRO ALA ALA TYR ASN ARG TYR ASP GLU ILE ARG PHE LYS GLY LYS GLU GLU THR

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

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

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

LEU ASP PRO ASN VAL GLN TRP ASP LYS THR VAL LEU GLU LEU SER TYR HIS LYS ARG HIS LEU ASP ARG PRO VAL PHE LEU ARG PHE GLU THR LEU ASP ARG MET VAL LYS HIS SER HIS LEU ARG PHE

• Molecule 23: DNA excision repair protein ERCC-8

•

Chain a:
　　　　　　　　92%　　　　　　　　•　　8%

M1 **K34** **K174 S175** **K232 K233 S234 Q235 A236 V237 E238 S239 A240** **K292** **L365** TYR GLU PRO VAL PRO ASP ASP ASP GLU THR THR THR LYS SER GLN LEU ASN PRO ALA PHE GLU ASP ALA TRP SER SER SER ASP GLU GLU GLY

• Molecule 24: DNA excision repair protein ERCC-6

11%

Chain b:
35%　　　　　　　　　　　64%

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

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

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

LYS GLU GLN GLN LEU LEU LYS LYS ILE THR ALA LYS GLN GLU LYS HIS LEU GLN ALA ILE LEU GLY TYR ALA GLU GLN PRO GLN ASP ALA GLU GLN PRO GLY PRO SER SER LEU GLU SER MET LEU LEU MET PRO VAL GLN GLU GLY THR ALA TRP GLU GLN GLU LEU ILE

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

• Molecule 25: UV-stimulated scaffold protein A

Chain c:

20%

20%   80%

• Molecule 26: DNA damage-binding protein 1

Chain d:

Q634 P635 R639 S643 L644 S645 T646 T647 N648 S653 D654 R655 V658 I659 Y660 S661 S662 N663 H664 K665 L666 V667 F668 N672 L673 K674 E675 Y678 L682 N683 S684 D685 G686 Y687 P688 D689 L693 A694 N695 N696 S697 T698 G702 T703 I704 D705 E706 I707 Q708 K709

Q743 ASP THR SER GLY G748 K769 S772 SER SER THR ALA PRO HIS GLU THR SER PHE GLY E784 E839 E840 D962 S981 ALA ALA THR THR D986 M1014 GLN ASN LEU GLY GLU THR SER THR P1023 R1080 D1099 N1111 LEU GLN TYR ASP ASP GLY SER GLY MET

LYS ARG GLU A1124 D1127 H1140

4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	100000	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 $(e^-/\text{Å}^2)$	40.4	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	81000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.130	Depositor
Minimum map value	-0.062	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.01	Depositor
Map size (Å)	419.99997, 419.99997, 419.99997	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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	0.30	0/11382	0.54	1/15368 (0.0%)
2	B	0.31	0/9233	0.53	0/12463
3	C	0.34	0/2132	0.56	1/2896 (0.0%)
4	D	0.37	0/1027	0.64	0/1384
5	E	0.29	0/1751	0.53	0/2366
6	F	0.31	0/667	0.50	0/901
7	G	0.50	0/1365	0.71	2/1853 (0.1%)
8	H	0.33	0/1207	0.53	0/1628
9	I	0.30	0/972	0.54	0/1316
10	J	0.32	0/542	0.50	0/730
11	K	0.30	0/939	0.49	0/1271
12	L	0.32	0/394	0.59	0/524
13	M	0.43	0/6770	0.71	23/9119 (0.3%)
14	N	0.89	0/817	0.95	0/1258
15	P	1.23	5/510 (1.0%)	2.12	8/793 (1.0%)
17	S	0.28	0/3559	0.52	0/4447
18	T	1.14	2/1056 (0.2%)	1.05	0/1624
19	U	0.28	0/413	0.47	0/511
20	V	0.28	0/864	0.54	0/1073
21	Y	0.32	0/1199	0.62	0/1497
22	Z	0.29	0/171	0.52	0/212
23	a	0.61	0/2908	0.61	0/3939
24	b	0.44	0/4460	0.63	2/6024 (0.0%)
25	c	0.22	0/563	0.42	0/702
26	d	0.47	1/8646 (0.0%)	0.62	0/11725
All	All	0.43	8/63547 (0.0%)	0.64	37/85624 (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
1	A	0	1
7	G	0	1
23	a	0	2
24	b	0	2
All	All	0	6

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	P	29	A	C1'-N9	-5.60	1.39	1.46
15	P	32	C	C1'-N1	5.49	1.56	1.48
18	T	36	DG	C3'-O3'	-5.25	1.37	1.44
15	P	26	U	C1'-N1	5.23	1.56	1.48
15	P	25	A	C1'-N9	-5.12	1.39	1.46
18	T	35	DT	C3'-O3'	-5.12	1.37	1.44
15	P	30	U	C1'-N1	5.03	1.56	1.48
26	d	840	GLU	CG-CD	-5.00	1.44	1.51

All (37) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	P	36	A	O5'-P-OP2	-33.38	70.65	110.70
15	P	36	A	OP1-P-OP2	-25.02	82.08	119.60
15	P	36	A	O5'-P-OP1	20.52	135.32	110.70
13	M	717	LEU	CB-CG-CD1	14.25	135.22	111.00
15	P	34	A	O5'-P-OP1	-13.04	93.97	105.70
15	P	35	A	OP2-P-O3'	13.01	133.83	105.20
15	P	35	A	OP1-P-O3'	-10.84	81.36	105.20
24	b	985	ARG	NE-CZ-NH1	-10.61	114.99	120.30
13	M	616	LEU	CB-CG-CD1	7.79	124.25	111.00
13	M	836	LEU	CB-CG-CD2	7.30	123.41	111.00
13	M	702	VAL	CG1-CB-CG2	7.08	122.23	110.90
15	P	33	A	O4'-C1'-N9	-7.06	102.55	108.20
13	M	841	LEU	CB-CG-CD2	6.94	122.80	111.00
13	M	451	LEU	CB-CG-CD1	6.81	122.58	111.00
13	M	469	LEU	CB-CG-CD1	6.68	122.36	111.00
13	M	541	LEU	CB-CG-CD2	6.58	122.19	111.00
13	M	811	LEU	CB-CG-CD2	6.50	122.05	111.00
13	M	924	LEU	CB-CG-CD1	6.45	121.97	111.00
24	b	985	ARG	NE-CZ-NH2	6.26	123.43	120.30
13	M	585	LEU	CB-CG-CD1	6.14	121.45	111.00
15	P	35	A	O3'-P-O5'	-6.04	92.51	104.00
13	M	889	LEU	CB-CG-CD2	6.00	121.20	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	M	523	MET	CA-CB-CG	5.99	123.47	113.30
13	M	416	LEU	CB-CG-CD1	5.89	121.01	111.00
7	G	114	PRO	CA-N-CD	-5.84	103.33	111.50
13	M	884	LEU	CB-CG-CD1	5.82	120.90	111.00
13	M	443	LEU	CB-CG-CD1	5.71	120.72	111.00
3	C	67	ARG	NE-CZ-NH2	-5.63	117.49	120.30
13	M	443	LEU	CB-CG-CD2	5.61	120.53	111.00
13	M	1001	LEU	CB-CG-CD1	-5.59	101.49	111.00
13	M	654	LEU	CB-CG-CD1	5.53	120.40	111.00
13	M	469	LEU	CB-CG-CD2	5.37	120.14	111.00
1	A	483	ARG	CG-CD-NE	5.25	122.83	111.80
13	M	884	LEU	CB-CG-CD2	5.23	119.89	111.00
7	G	22	LEU	CA-CB-CG	5.22	127.30	115.30
13	M	924	LEU	CB-CG-CD2	5.14	119.74	111.00
13	M	836	LEU	CB-CG-CD1	5.04	119.57	111.00

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	538	VAL	Peptide
7	G	124	ASN	Peptide
23	a	174	LYS	Peptide
23	a	175	SER	Peptide
24	b	912	LEU	Peptide
24	b	995	PHE	Peptide

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	
1	A	1402/1970 (71%)	1368 (98%)	34 (2%)	0	100	100
2	B	1123/1174 (96%)	1075 (96%)	48 (4%)	0	100	100
3	C	256/275 (93%)	249 (97%)	7 (3%)	0	100	100
4	D	126/142 (89%)	121 (96%)	5 (4%)	0	100	100
5	E	207/210 (99%)	204 (99%)	3 (1%)	0	100	100
6	F	80/127 (63%)	75 (94%)	5 (6%)	0	100	100
7	G	169/172 (98%)	163 (96%)	5 (3%)	1 (1%)	25	58
8	H	146/150 (97%)	142 (97%)	4 (3%)	0	100	100
9	I	115/125 (92%)	111 (96%)	4 (4%)	0	100	100
10	J	65/67 (97%)	65 (100%)	0	0	100	100
11	K	113/117 (97%)	111 (98%)	2 (2%)	0	100	100
12	L	44/58 (76%)	40 (91%)	4 (9%)	0	100	100
13	M	788/1726 (46%)	743 (94%)	45 (6%)	0	100	100
17	S	888/1173 (76%)	842 (95%)	46 (5%)	0	100	100
19	U	98/666 (15%)	82 (84%)	14 (14%)	2 (2%)	7	27
20	V	209/531 (39%)	174 (83%)	31 (15%)	4 (2%)	8	28
21	Y	298/305 (98%)	278 (93%)	20 (7%)	0	100	100
22	Z	41/531 (8%)	40 (98%)	1 (2%)	0	100	100
23	a	363/396 (92%)	344 (95%)	19 (5%)	0	100	100
24	b	526/1493 (35%)	504 (96%)	22 (4%)	0	100	100
25	c	139/709 (20%)	136 (98%)	3 (2%)	0	100	100
26	d	1082/1140 (95%)	1012 (94%)	69 (6%)	1 (0%)	51	82
All	All	8278/13257 (62%)	7879 (95%)	391 (5%)	8 (0%)	54	82

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
20	V	248	PRO
26	d	707	ILE
7	G	64	GLY
19	U	481	GLY
20	V	301	ASN
19	U	463	PRO
20	V	238	GLU
20	V	259	ASP

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	1242/1749 (71%)	1240 (100%)	2 (0%)	93	98
2	B	992/1027 (97%)	990 (100%)	2 (0%)	93	98
3	C	237/252 (94%)	237 (100%)	0	100	100
4	D	108/126 (86%)	108 (100%)	0	100	100
5	E	191/192 (100%)	190 (100%)	1 (0%)	88	96
6	F	71/111 (64%)	71 (100%)	0	100	100
7	G	147/153 (96%)	147 (100%)	0	100	100
8	H	129/131 (98%)	129 (100%)	0	100	100
9	I	105/112 (94%)	105 (100%)	0	100	100
10	J	56/56 (100%)	56 (100%)	0	100	100
11	K	104/106 (98%)	104 (100%)	0	100	100
12	L	43/55 (78%)	43 (100%)	0	100	100
13	M	722/1522 (47%)	720 (100%)	2 (0%)	92	98
23	a	320/348 (92%)	320 (100%)	0	100	100
24	b	476/1297 (37%)	475 (100%)	1 (0%)	93	98
26	d	938/999 (94%)	938 (100%)	0	100	100
All	All	5881/8236 (71%)	5873 (100%)	8 (0%)	93	98

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	241	ARG
1	A	1101	GLN
2	B	1072	ARG
2	B	1080	ARG
5	E	166	ARG
13	M	625	ARG
13	M	1251	LYS
24	b	745	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	98	HIS
4	D	43	HIS
13	M	943	HIS
13	M	983	GLN
24	b	773	HIS
26	d	467	GLN

5.3.3   RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	P	20/45 (44%)	10 (50%)	1 (5%)

All (10) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
15	P	27	A
15	P	28	U
15	P	29	A
15	P	30	U
15	P	33	A
15	P	34	A
15	P	35	A
15	P	36	A
15	P	39	G
15	P	44	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
15	P	27	A

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 9 ligands modelled in this entry, 9 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-13010. These allow visual inspection of the internal detail of the map and identification of artifacts.

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

6.1   Orthogonal projections ⓘ
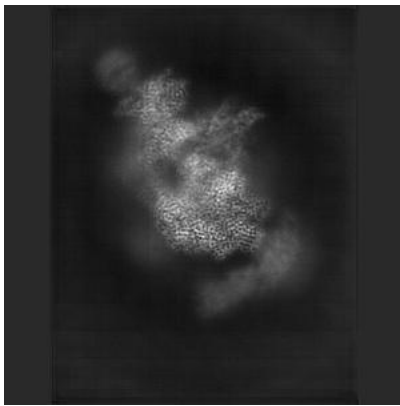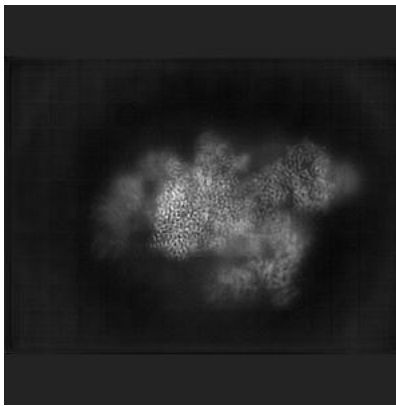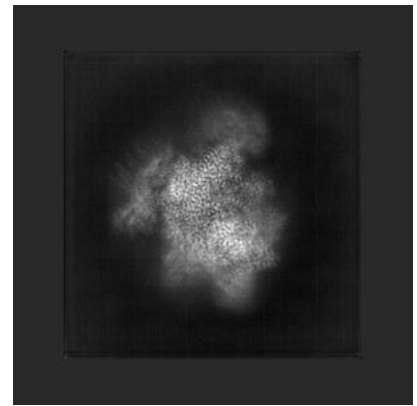
6.1.1   Primary map

X   Y   Z

6.1.2   Raw map

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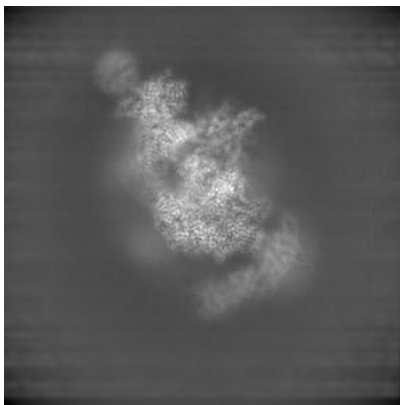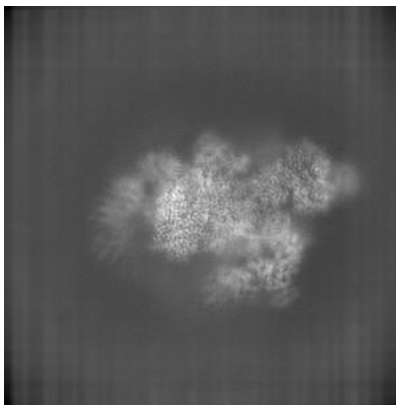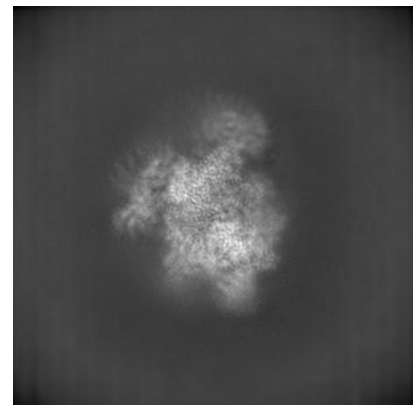
The images above show the map projected in three orthogonal directions.

6.2   Central slices ⓘ
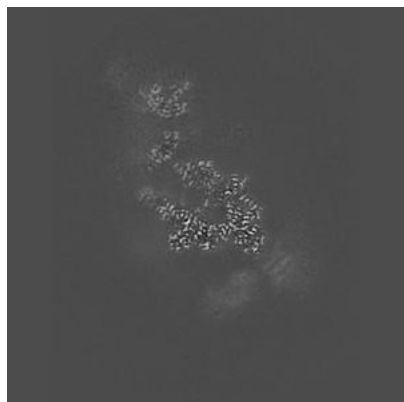
6.2.1   Primary map
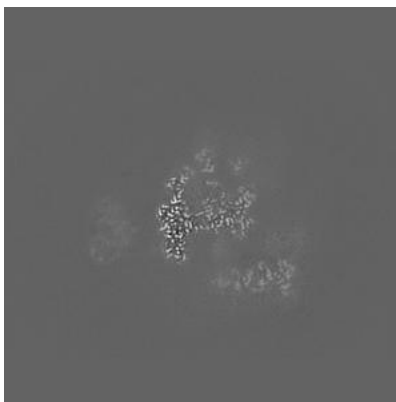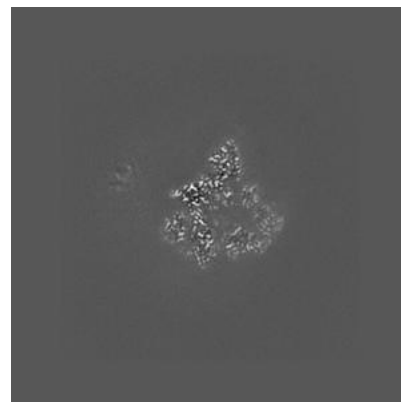
X Index: 200   Y Index: 200   Z Index: 200

6.2.2   Raw map
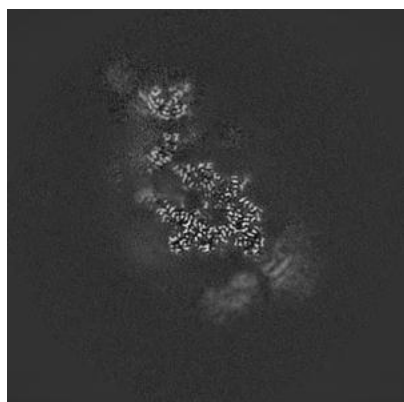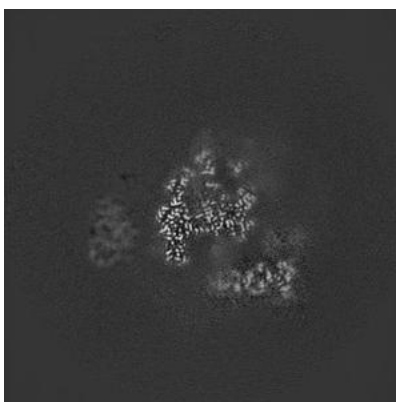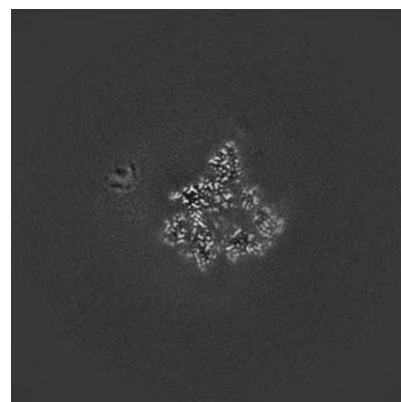
X Index: 200   Y Index: 200   Z Index: 200

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

6.3 Largest variance slices (i)

6.3.1 Primary map
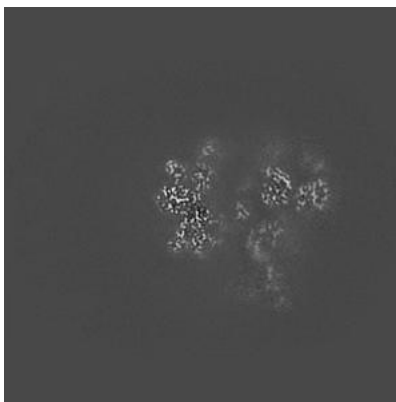
X Index: 195	Y Index: 173	Z Index: 173
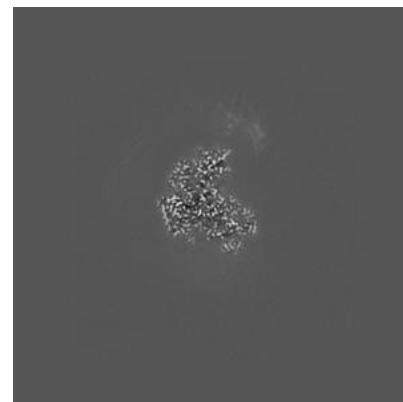
6.3.2 Raw map
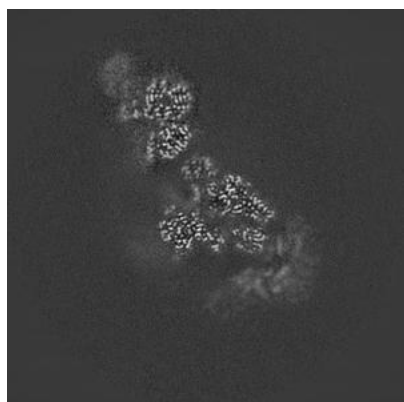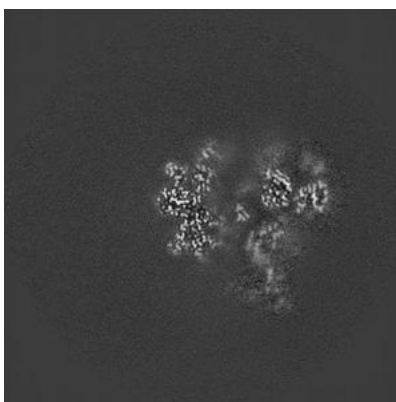
X Index: 211	Y Index: 173	Z Index: 173
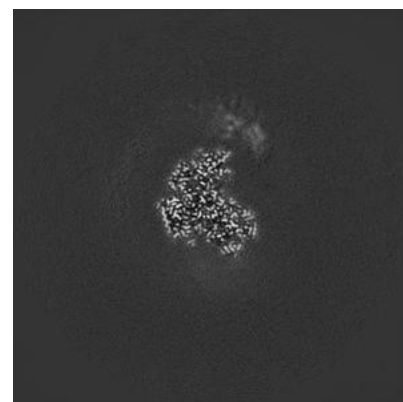
The images above show the largest variance slices of the map in three orthogonal directions.

6.4　Orthogonal surface views ⓘ

6.4.1　Primary map

X　　　　　　　　Y　　　　　　　　Z

The images above show the 3D surface view of the map at the recommended contour level 0.01. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.4.2　Raw map

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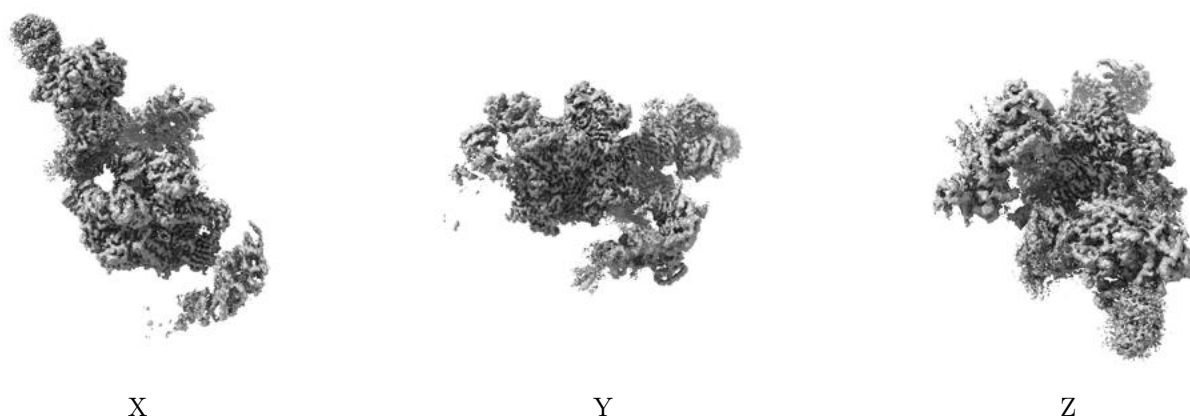
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

6.5 Mask visualisation ⓘ

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

A mask typically either:

- Encompasses the whole structure

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

6.5.1 emd_13010_msk_1.map ⓘ
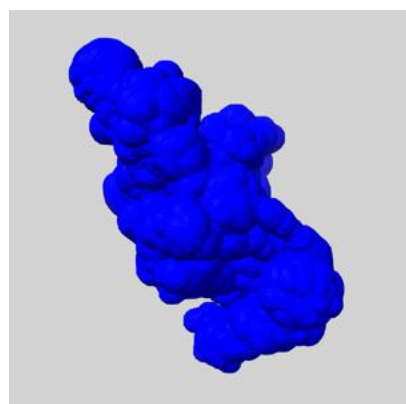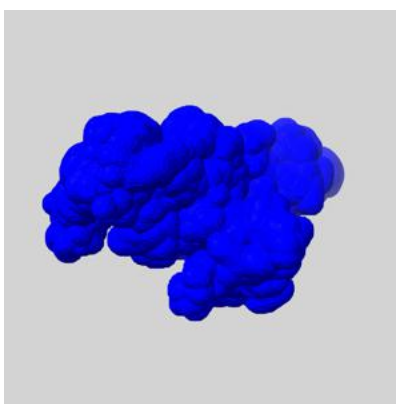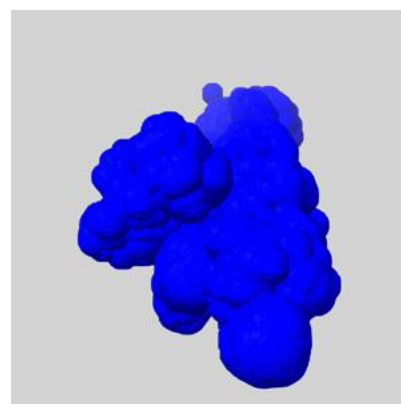
X　　　　　　　　　　　Y　　　　　　　　　　　Z

6.5.2 emd_13010_msk_2.map ⓘ
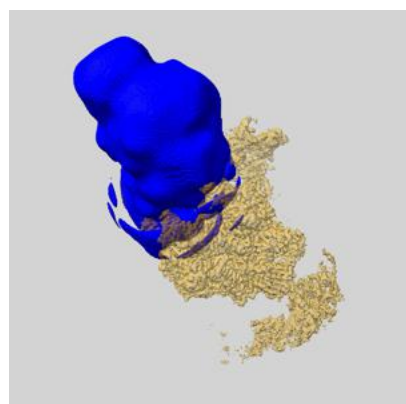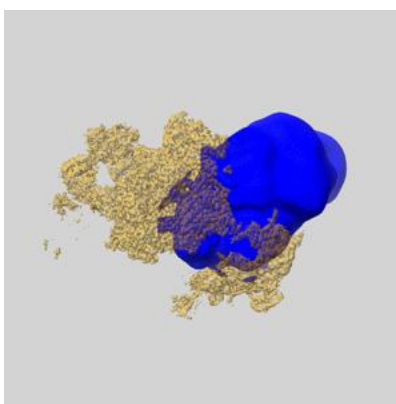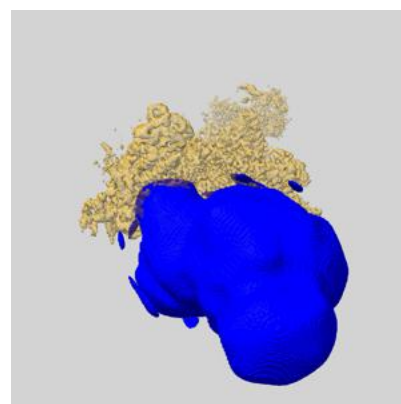
X　　　　　　　　　　　Y　　　　　　　　　　　Z

6.5.3 emd_13010_msk_3.map ⓘ
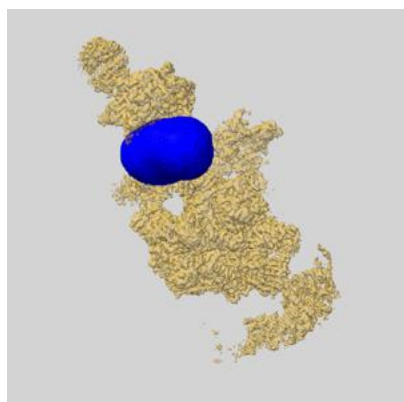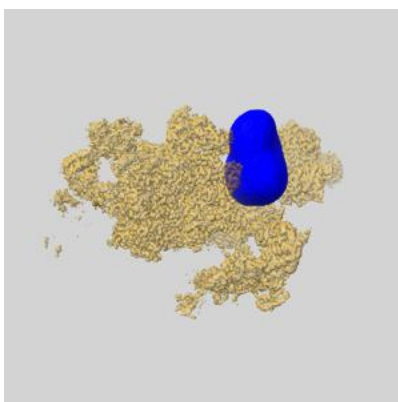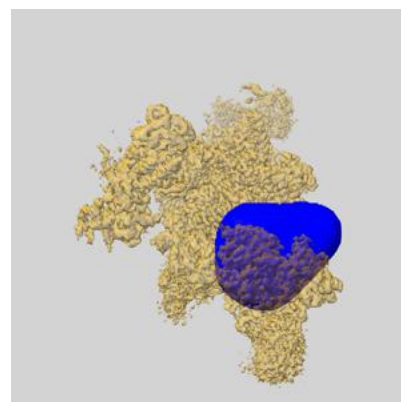
| X | Y | Z |

6.5.4 emd_13010_msk_4.map ⓘ
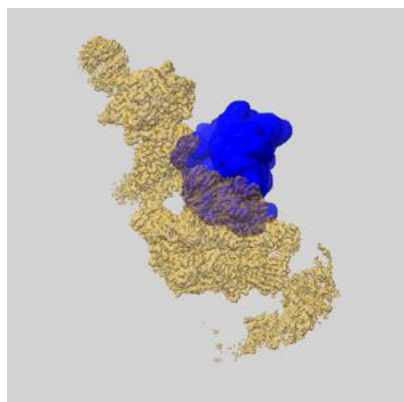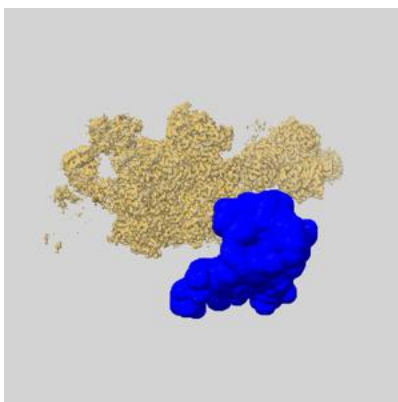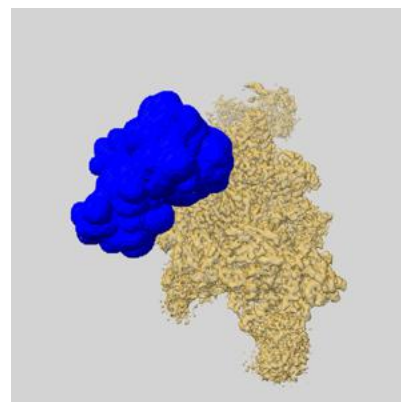
| X | Y | Z |

6.5.5 emd_13010_msk_5.map ⓘ

| X | Y | Z |
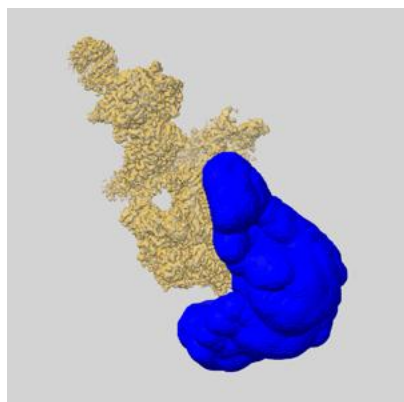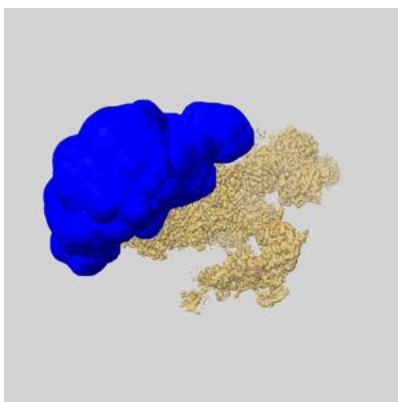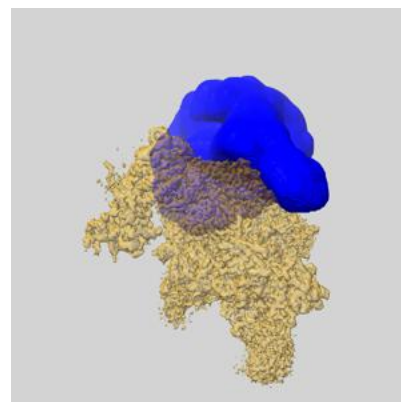
7 Map analysis ⓘ

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

7.1 Map-value distribution ⓘ

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

7.2 Volume estimate ⓘ
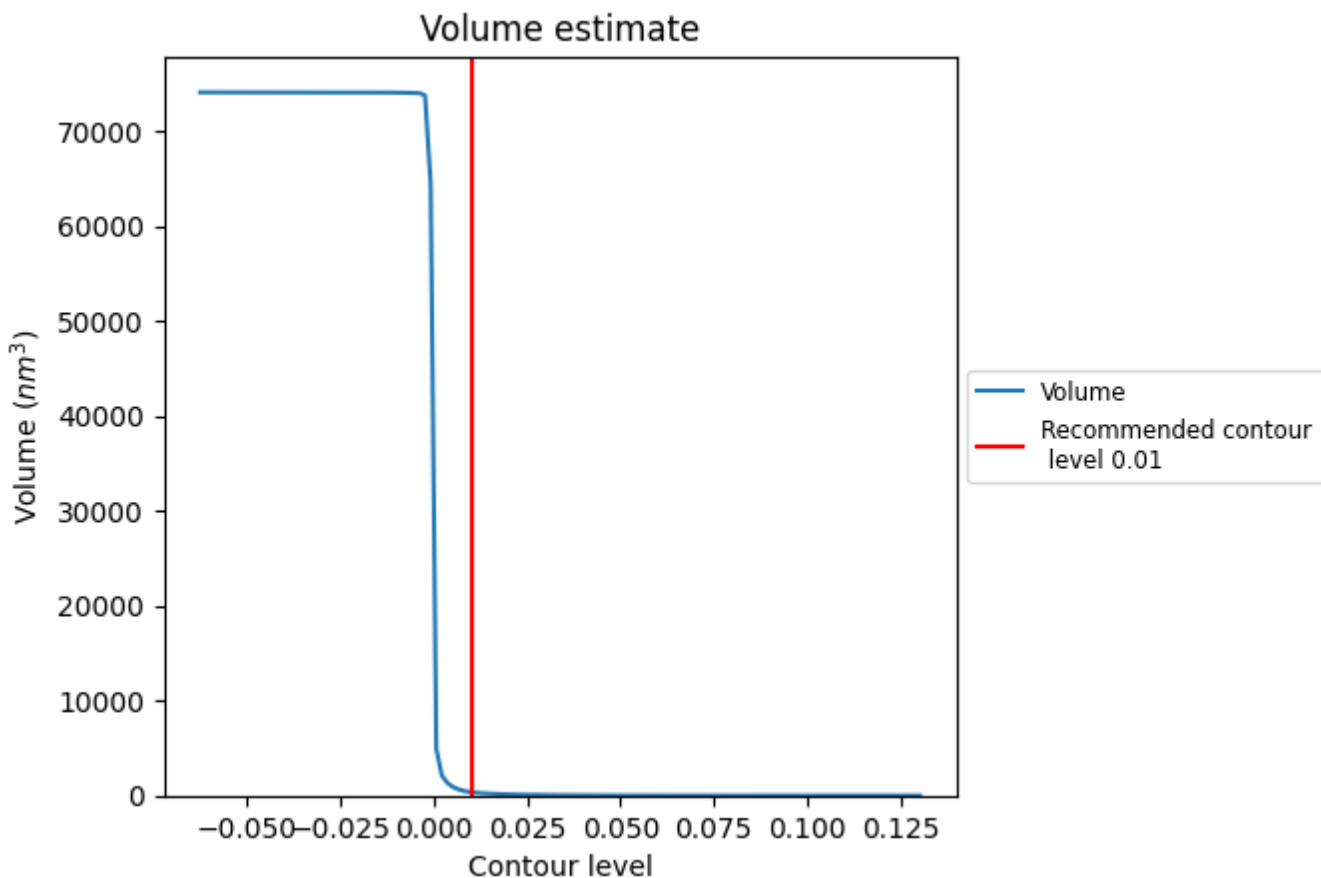
The volume at the recommended contour level is 363 nm^3; this corresponds to an approximate mass of 328 kDa.

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 ⓘ
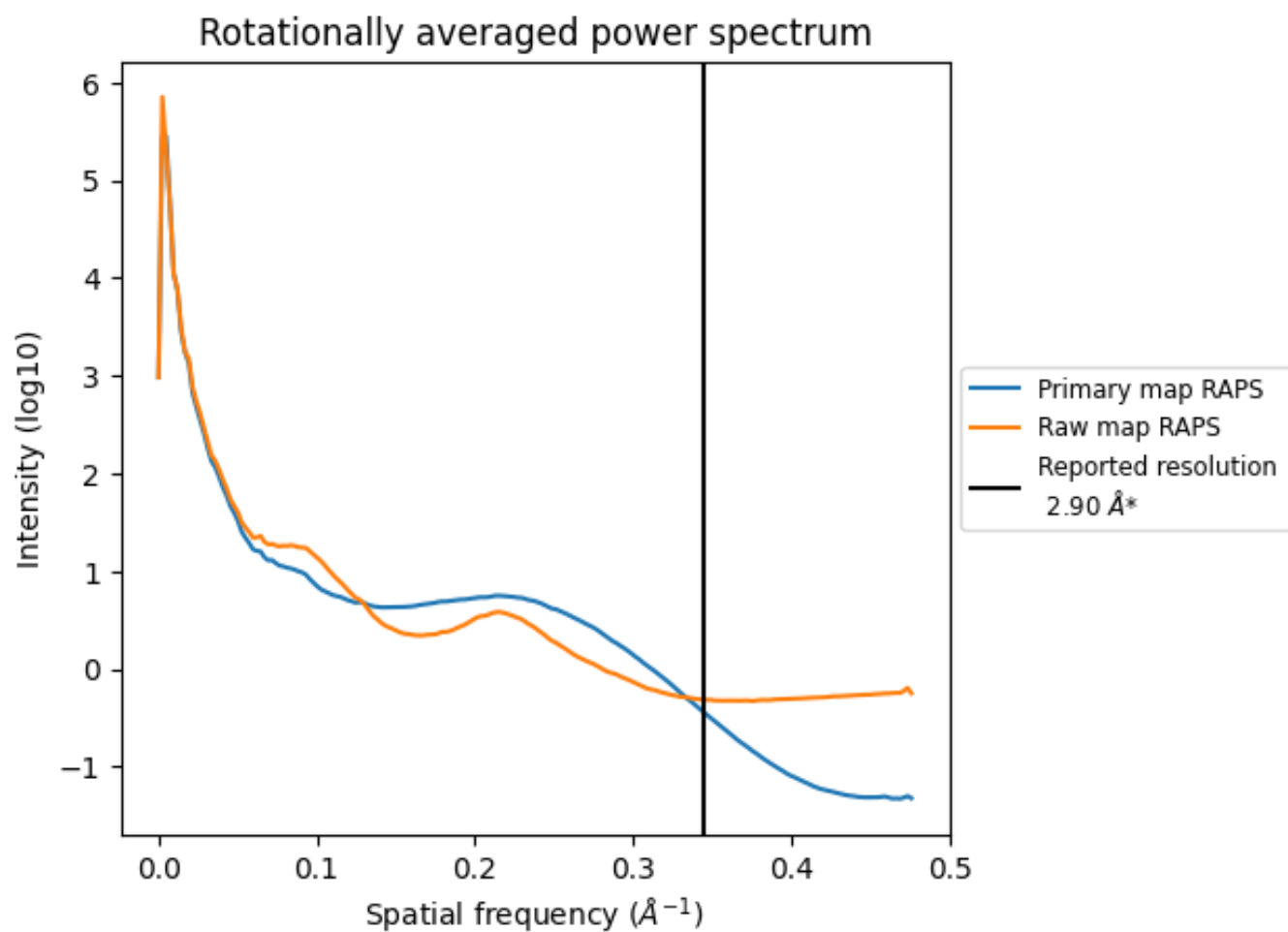
*Reported resolution corresponds to spatial frequency of 0.345 Å^{-1}

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)
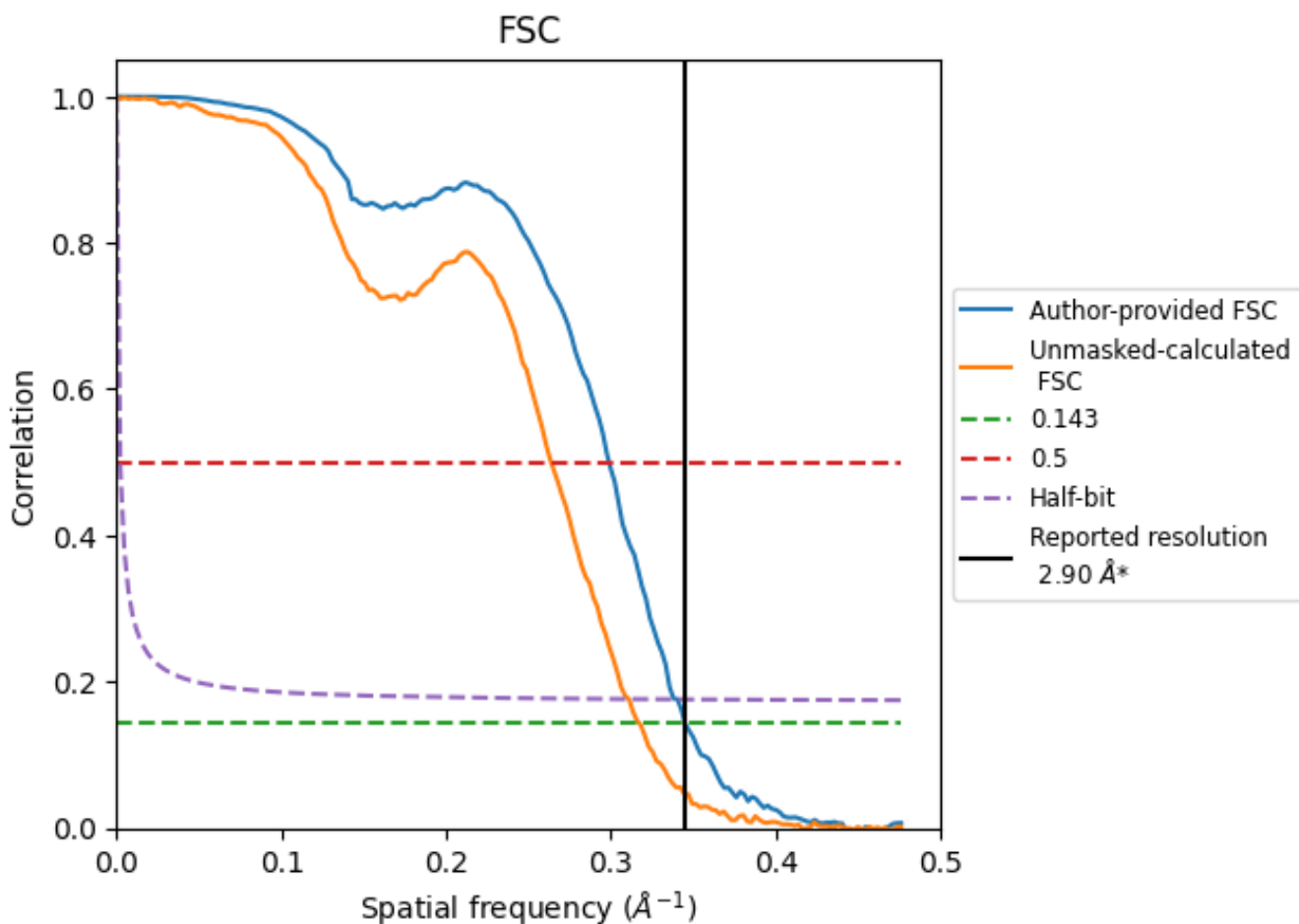
*Reported resolution corresponds to spatial frequency of 0.345 Å^{-1}

8.2　Resolution estimates ⓘ

| Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.90	-	-
Author-provided FSC curve	2.90	3.34	2.94
Unmasked-calculated*	3.15	3.79	3.22

*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-13010 and PDB model 7OOP. Per-residue inclusion information can be found in section 3 on page 9.
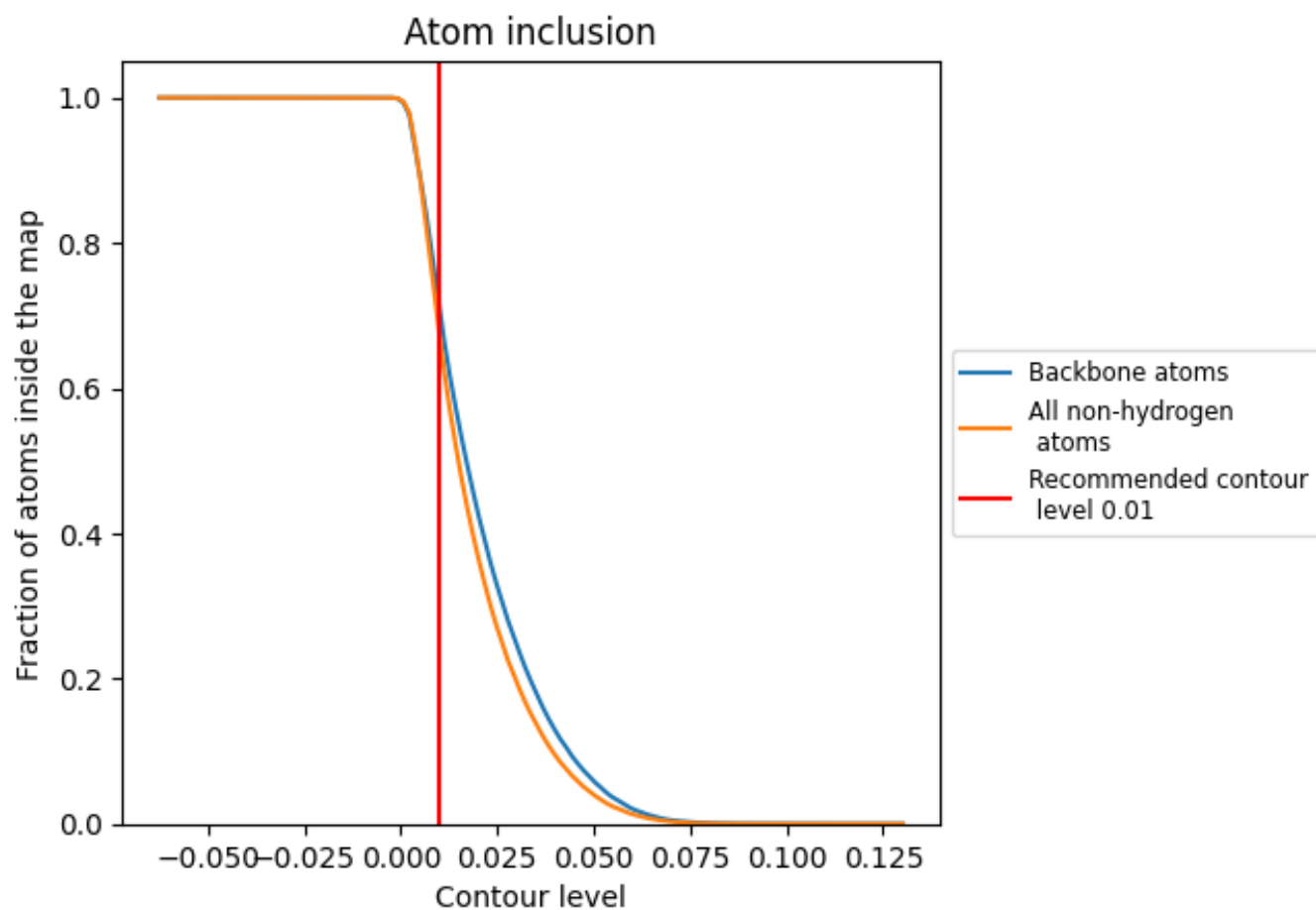
9.1 Map-model overlay ⓘ

X　　　　　　　　　　　Y　　　　　　　　　　　Z
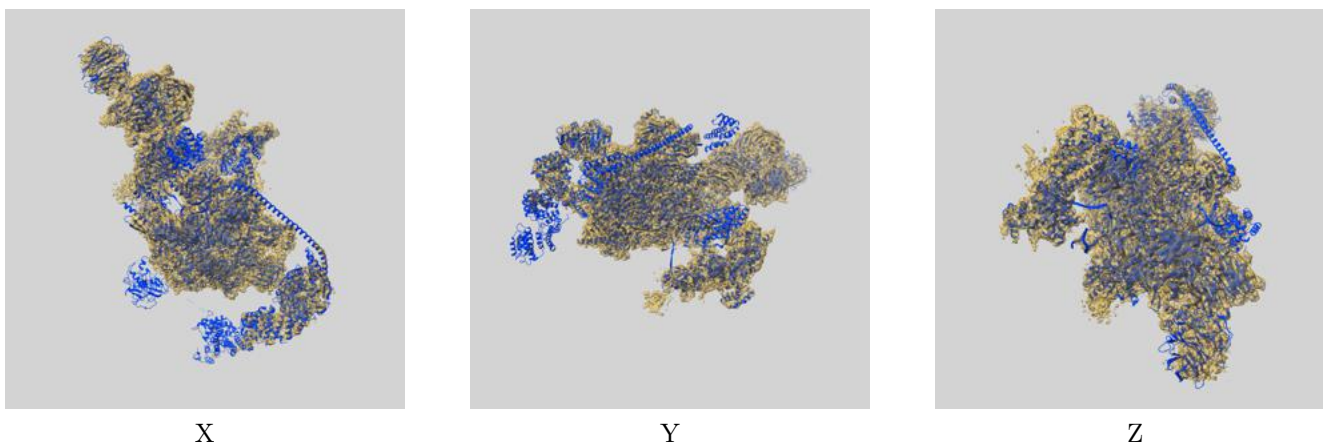
The images above show the 3D surface view of the map at the recommended contour level 0.01 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Atom inclusion ⓘ

At the recommended contour level, 71% of all backbone atoms, 68% of all non-hydrogen atoms, are inside the map.