



## wwPDB EM Validation Summary Report ⓘ

Jul 26, 2021 – 12:23 pm BST

PDB ID : 7O72  
EMDB ID : EMD-12743  
Title : Yeast RNA polymerase II transcription pre-initiation complex with closed promoter DNA  
Authors : Schilbach, S.; Aibara, S.; Dienemann, C.; Grabbe, F.; Cramer, P.  
Deposited on : 2021-04-12  
Resolution : 3.40 Å(reported)  
Based on initial model : 5OQJ

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ symbol.

---

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

EMDB validation analysis : 0.0.0.dev84  
Mogul : 1.8.5 (274361), 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)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.22

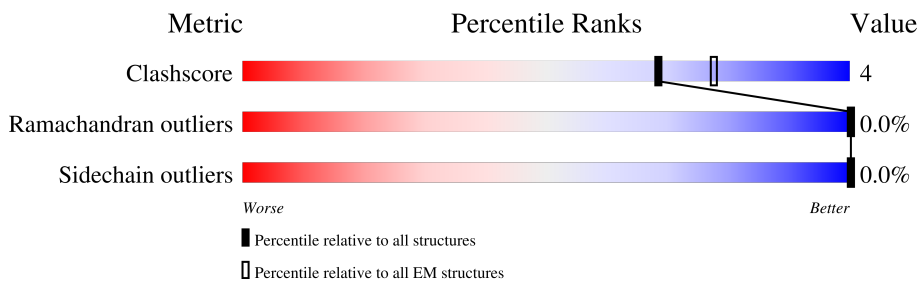
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.









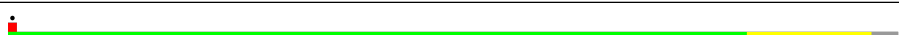
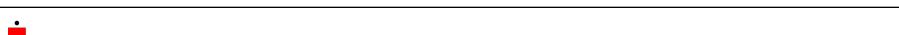
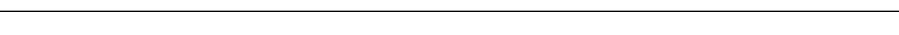
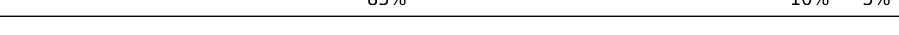
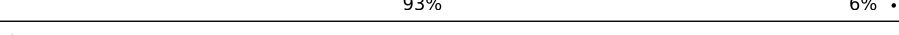
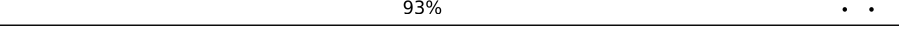










Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	0	778	
2	1	645	
3	2	517	
4	3	324	
5	4	341	
6	5	76	
7	6	464	
8	7	843	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	A	1733	
10	B	1224	
11	C	347	
12	D	221	
13	E	215	
14	F	155	
15	G	177	
16	H	146	
17	I	122	
18	J	70	
19	K	120	
20	L	70	
21	M	352	
22	N	106	
23	O	247	
24	Q	738	
25	R	400	
26	T	106	
27	U	286	
28	V	129	
29	W	492	
30	X	328	

## 2 Entry composition

There are 35 unique types of molecules in this entry. The entry contains 73664 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 General transcription and DNA repair factor IIIH helicase subunit XPD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	0	752	6091	3882	1029	1142	38	0	0

- Molecule 2 is a protein called General transcription and DNA repair factor IIIH subunit TFB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	1	522	4214	2660	734	798	22	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	-2	GLY	-	expression tag	UNP P32776
1	-1	GLY	-	expression tag	UNP P32776
1	0	SER	-	expression tag	UNP P32776

- Molecule 3 is a protein called General transcription and DNA repair factor IIIH subunit TFB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	445	3597	2327	591	663	16	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
2	-3	GLY	-	expression tag	UNP Q02939
2	-2	PRO	-	expression tag	UNP Q02939
2	-1	GLY	-	expression tag	UNP Q02939
2	0	SER	-	expression tag	UNP Q02939

- Molecule 4 is a protein called RNA polymerase II transcription factor B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	3	131	1089	692	180	209	8	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
3	-2	GLY	-	expression tag	UNP Q03290
3	-1	PRO	-	expression tag	UNP Q03290
3	0	HIS	-	expression tag	UNP Q03290

- Molecule 5 is a protein called General transcription and DNA repair factor IIH subunit TFB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	4	293	2274	1454	377	429	14	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
4	-2	SER	-	expression tag	UNP Q12004
4	-1	ASN	-	expression tag	UNP Q12004
4	0	ALA	-	expression tag	UNP Q12004

- Molecule 6 is a protein called General transcription and DNA repair factor IIH subunit TFB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	65	514	326	90	95	3	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
5	-3	GLY	-	expression tag	UNP Q3E7C1
5	-2	PRO	-	expression tag	UNP Q3E7C1
5	-1	GLY	-	expression tag	UNP Q3E7C1
5	0	SER	-	expression tag	UNP Q3E7C1

- Molecule 7 is a protein called General transcription and DNA repair factor IIH subunit SSL1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	6	356	2794	1771	482	513	28	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
6	-2	GLY	-	expression tag	UNP Q04673
6	-1	GLY	-	expression tag	UNP Q04673
6	0	SER	-	expression tag	UNP Q04673

- Molecule 8 is a protein called General transcription and DNA repair factor IIIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	7	615	4954	3153	860	914	27	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	A	1427	11230	7075	1961	2132	62	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	B	1180	9404	5946	1643	1760	55	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	C	266	2092	1315	348	416	13	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-28	MET	-	initiating methionine	UNP P16370
C	-27	GLY	-	expression tag	UNP P16370
C	-26	SER	-	expression tag	UNP P16370
C	-25	HIS	-	expression tag	UNP P16370
C	-24	HIS	-	expression tag	UNP P16370
C	-23	HIS	-	expression tag	UNP P16370
C	-22	HIS	-	expression tag	UNP P16370
C	-21	HIS	-	expression tag	UNP P16370
C	-20	HIS	-	expression tag	UNP P16370

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-19	SER	-	expression tag	UNP P16370
C	-18	ASN	-	expression tag	UNP P16370
C	-17	SER	-	expression tag	UNP P16370
C	-16	GLY	-	expression tag	UNP P16370
C	-15	LEU	-	expression tag	UNP P16370
C	-14	ASN	-	expression tag	UNP P16370
C	-13	ASP	-	expression tag	UNP P16370
C	-12	ILE	-	expression tag	UNP P16370
C	-11	PHE	-	expression tag	UNP P16370
C	-10	GLU	-	expression tag	UNP P16370
C	-9	ALA	-	expression tag	UNP P16370
C	-8	GLN	-	expression tag	UNP P16370
C	-7	LYS	-	expression tag	UNP P16370
C	-6	ILE	-	expression tag	UNP P16370
C	-5	GLU	-	expression tag	UNP P16370
C	-4	TRP	-	expression tag	UNP P16370
C	-3	HIS	-	expression tag	UNP P16370
C	-2	GLU	-	expression tag	UNP P16370
C	-1	ASP	-	expression tag	UNP P16370
C	0	THR	-	expression tag	UNP P16370
C	1	GLY	-	expression tag	UNP P16370
C	2	SER	-	expression tag	UNP P16370
C	3	SER	-	expression tag	UNP P16370

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	D	165	1321	817	234	268	2	0	0

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	E	214	1752	1111	309	321	11	0	0

- Molecule 14 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	F	86	697	445	118	131	3	0	0

- Molecule 15 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	G	171	1339	861	222	248	8	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	172	HIS	-	expression tag	UNP P34087
G	173	HIS	-	expression tag	UNP P34087
G	174	HIS	-	expression tag	UNP P34087
G	175	HIS	-	expression tag	UNP P34087
G	176	HIS	-	expression tag	UNP P34087
G	177	HIS	-	expression tag	UNP P34087

- Molecule 16 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	H	140	1120	704	188	224	4	0	0

- Molecule 17 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	I	116	944	581	172	181	10	0	0

- Molecule 18 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	J	69	569	362	101	100	6	0	0

- Molecule 19 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	K	115	924	593	157	172	2	0	0

- Molecule 20 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.



Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	L	45	359	221	71	63	4	0	0

- Molecule 21 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	M	309	2367	1493	407	449	18	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	346	LYS	-	expression tag	UNP P29055
M	347	HIS	-	expression tag	UNP P29055
M	348	HIS	-	expression tag	UNP P29055
M	349	HIS	-	expression tag	UNP P29055
M	350	HIS	-	expression tag	UNP P29055
M	351	HIS	-	expression tag	UNP P29055
M	352	HIS	-	expression tag	UNP P29055

- Molecule 22 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
22	N	65	1332	637	236	394	65	0	0

- Molecule 23 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	O	181	1422	925	243	248	6	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	241	LYS	-	expression tag	UNP P13393
O	242	HIS	-	expression tag	UNP P13393
O	243	HIS	-	expression tag	UNP P13393
O	244	HIS	-	expression tag	UNP P13393
O	245	HIS	-	expression tag	UNP P13393
O	246	HIS	-	expression tag	UNP P13393
O	247	HIS	-	expression tag	UNP P13393

- Molecule 24 is a protein called Transcription initiation factor IIF subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	Q	222	1878	1183	347	341	7	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	-2	GLY	-	expression tag	UNP P41895
Q	-1	PRO	-	expression tag	UNP P41895
Q	0	GLY	-	expression tag	UNP P41895

- Molecule 25 is a protein called Transcription initiation factor IIF subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	R	261	2173	1369	383	411	10	0	0

- Molecule 26 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
26	T	65	1333	636	246	386	65	0	0

- Molecule 27 is a protein called Transcription initiation factor IIA large subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	U	107	885	559	147	176	3	0	0

- Molecule 28 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	V	104	815	511	136	164	4	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	123	LYS	-	expression tag	UNP P32774
V	124	HIS	-	expression tag	UNP P32774
V	125	HIS	-	expression tag	UNP P32774

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
V	126	HIS	-	expression tag	UNP P32774
V	127	HIS	-	expression tag	UNP P32774
V	128	HIS	-	expression tag	UNP P32774
V	129	HIS	-	expression tag	UNP P32774

- Molecule 29 is a protein called Transcription initiation factor IIE subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	W	304	2470	1555	432	476	7	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	483	ALA	-	expression tag	UNP P36100
W	484	ALA	-	expression tag	UNP P36100
W	485	ALA	-	expression tag	UNP P36100
W	486	LEU	-	expression tag	UNP P36100
W	487	GLU	-	expression tag	UNP P36100
W	488	HIS	-	expression tag	UNP P36100
W	489	HIS	-	expression tag	UNP P36100
W	490	HIS	-	expression tag	UNP P36100
W	491	HIS	-	expression tag	UNP P36100
W	492	HIS	-	expression tag	UNP P36100

- Molecule 30 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	X	206	1653	1055	280	312	6	0	0

- Molecule 31 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).

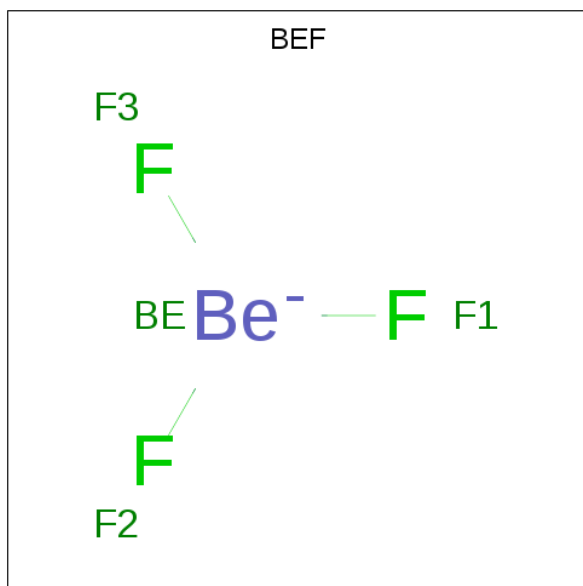


Mol	Chain	Residues	Atoms		AltConf
31	0	1	Total	Fe S	0
			8	4 4	

- Molecule 32 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
32	3	2	Total	Zn	0
			2	2	
32	4	1	Total	Zn	0
			1	1	
32	6	4	Total	Zn	0
			4	4	
32	A	2	Total	Zn	0
			2	2	
32	B	1	Total	Zn	0
			1	1	
32	C	1	Total	Zn	0
			1	1	
32	I	2	Total	Zn	0
			2	2	
32	J	1	Total	Zn	0
			1	1	
32	L	1	Total	Zn	0
			1	1	
32	M	1	Total	Zn	0
			1	1	
32	W	1	Total	Zn	0
			1	1	

- Molecule 33 is BERYLLIUM TRIFLUORIDE ION (three-letter code: BEF) (formula: BeF<sub>3</sub>).

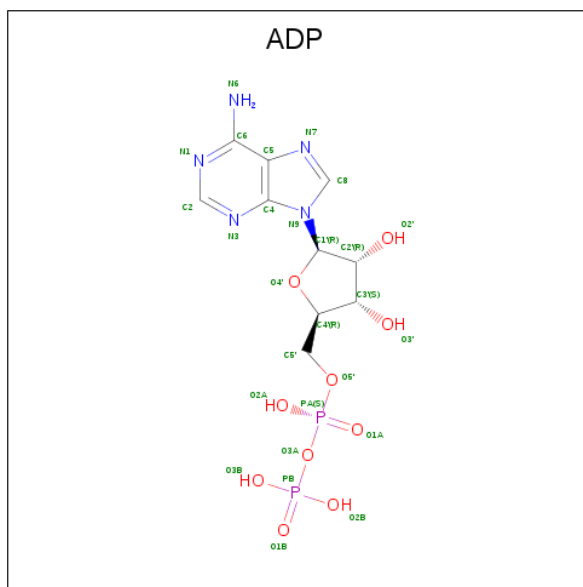


Mol	Chain	Residues	Atoms			AltConf
			Total	Be	F	
33	7	1	4	1	3	0

- Molecule 34 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

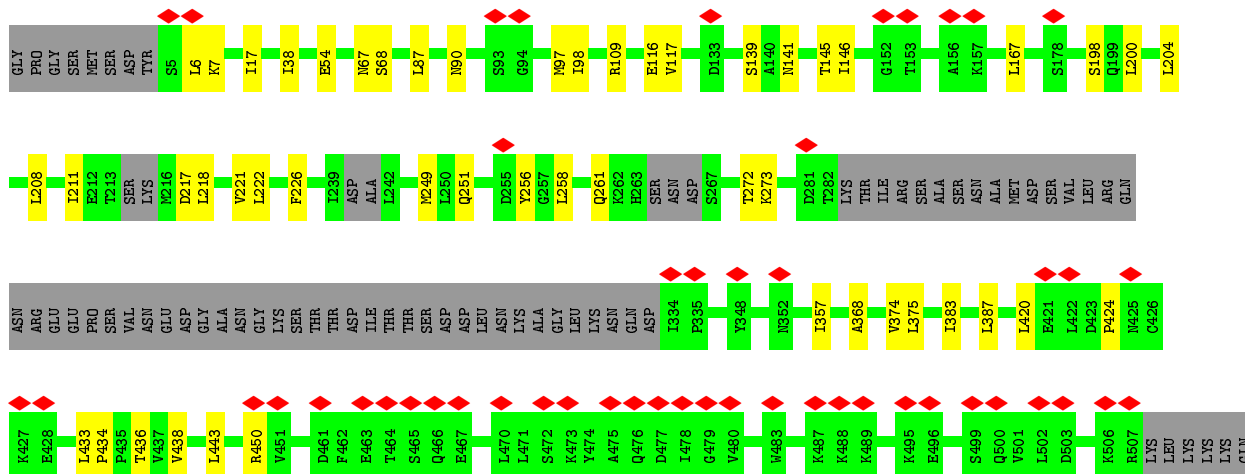
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
34	7	1	1	1	0
34	A	1	1	1	0

- Molecule 35 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>10</sub>P<sub>2</sub>).

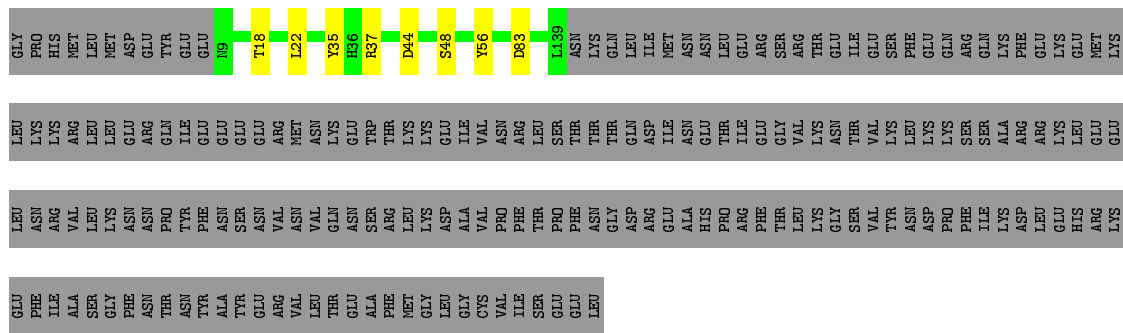


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
35	7	1	27	10	5	10	2	0

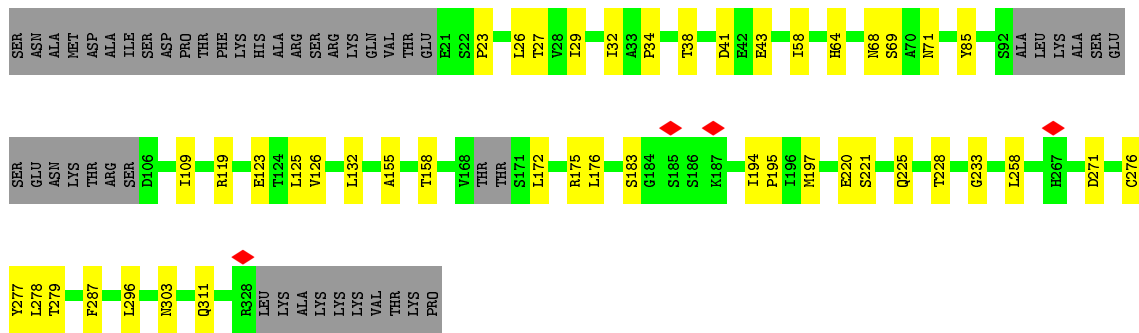




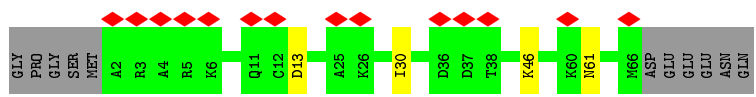
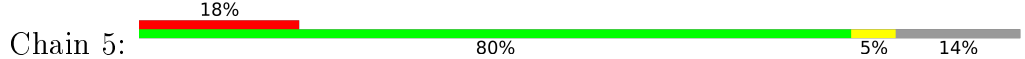
• Molecule 4: RNA polymerase II transcription factor B subunit 3



• Molecule 5: General transcription and DNA repair factor IIH subunit TFB4



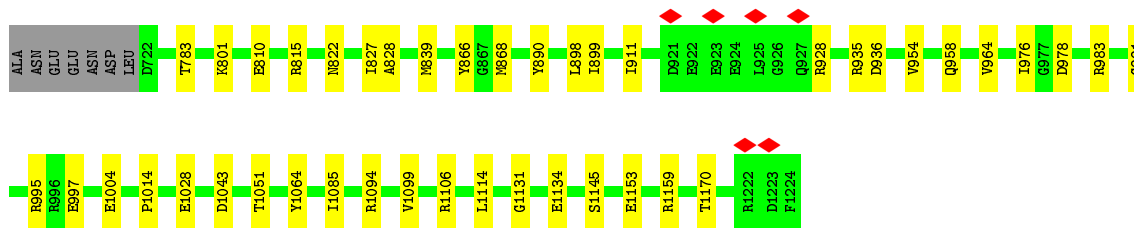
• Molecule 6: General transcription and DNA repair factor IIH subunit TFB5



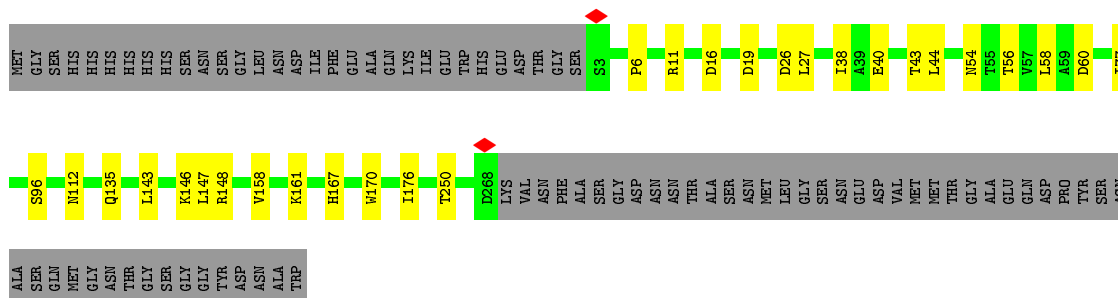




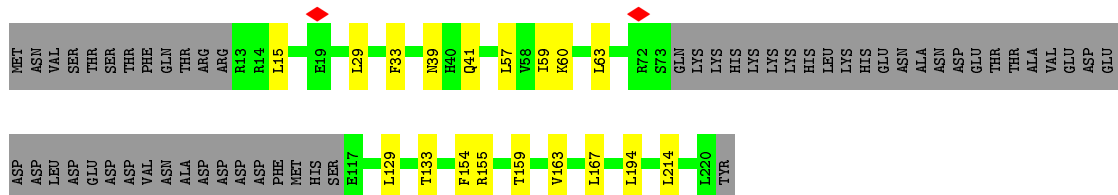




- Molecule 11: DNA-directed RNA polymerase II subunit RPB3



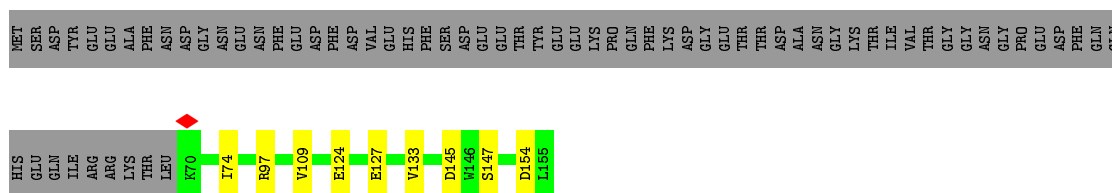
- Molecule 12: DNA-directed RNA polymerase II subunit RPB4




- Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC1

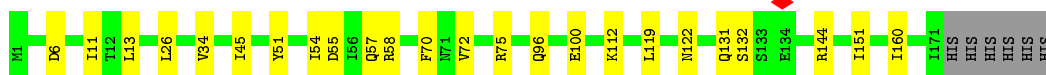


- Molecule 14: DNA-directed RNA polymerases I, II, and III subunit RPABC2




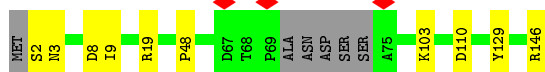
- Molecule 15: DNA-directed RNA polymerase II subunit RPB7

Chain G:  83% 14%




- Molecule 16: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:  89% 7%



- Molecule 17: DNA-directed RNA polymerase II subunit RPB9

Chain I:  85% 10% 5%



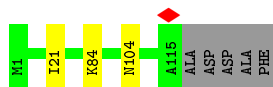
- Molecule 18: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:  93% 6%



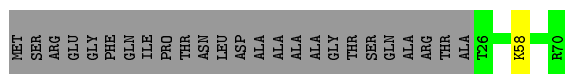
- Molecule 19: DNA-directed RNA polymerase II subunit RPB11

Chain K:  93%




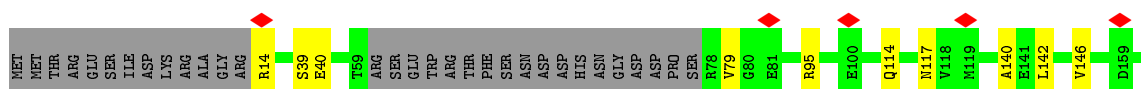
- Molecule 20: DNA-directed RNA polymerases I, II, and III subunit RPABC4

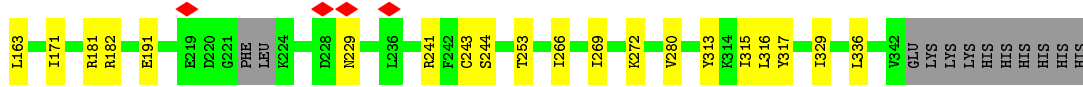
Chain L:  63% 36%



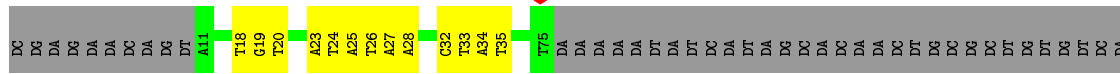
- Molecule 21: Transcription initiation factor IIB

Chain M:  79% 9% 12%

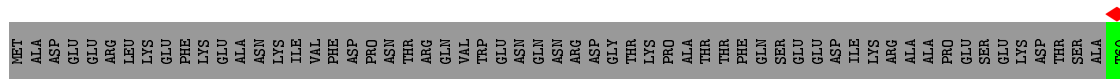




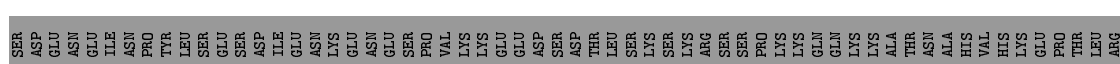
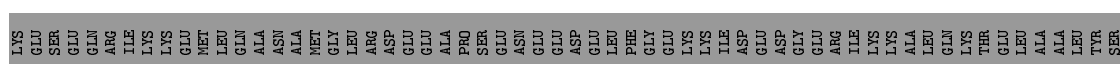
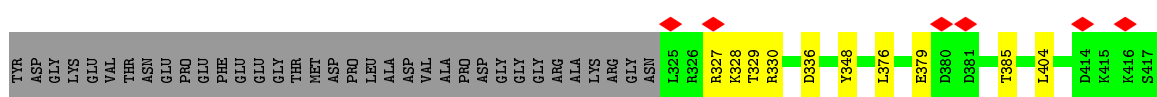
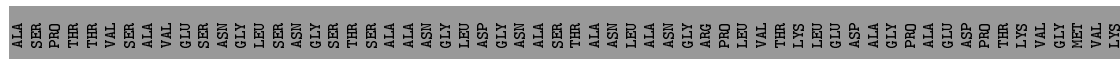
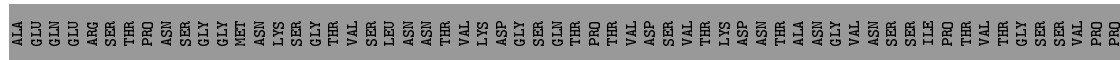
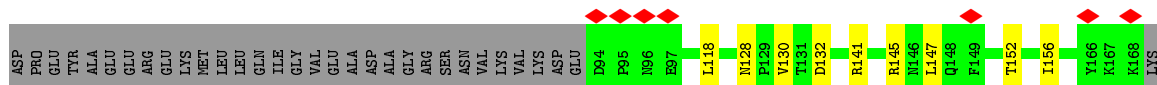
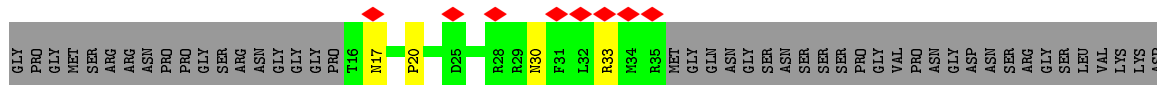
• Molecule 22: Non-template DNA



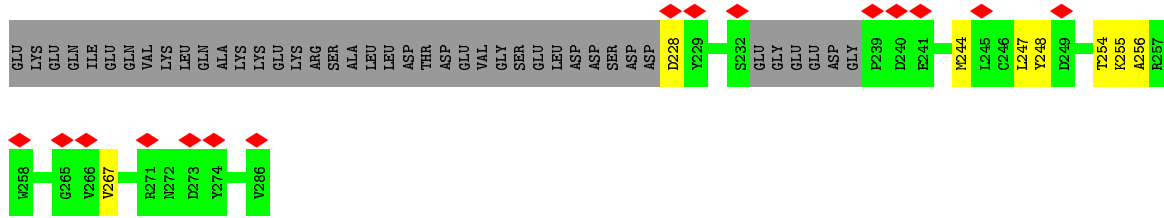
• Molecule 23: TATA-box-binding protein



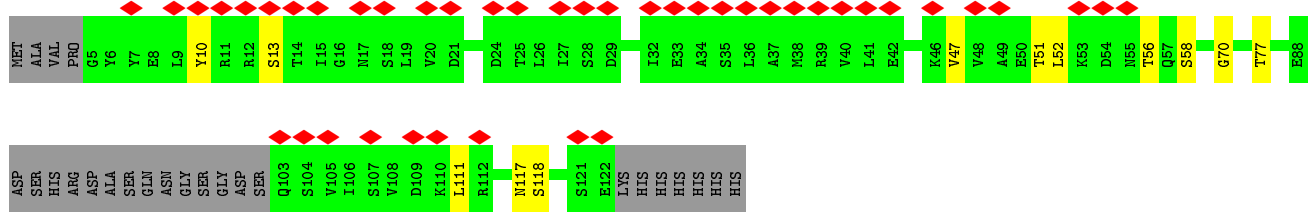
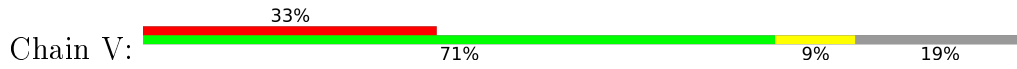
• Molecule 24: Transcription initiation factor IIF subunit alpha



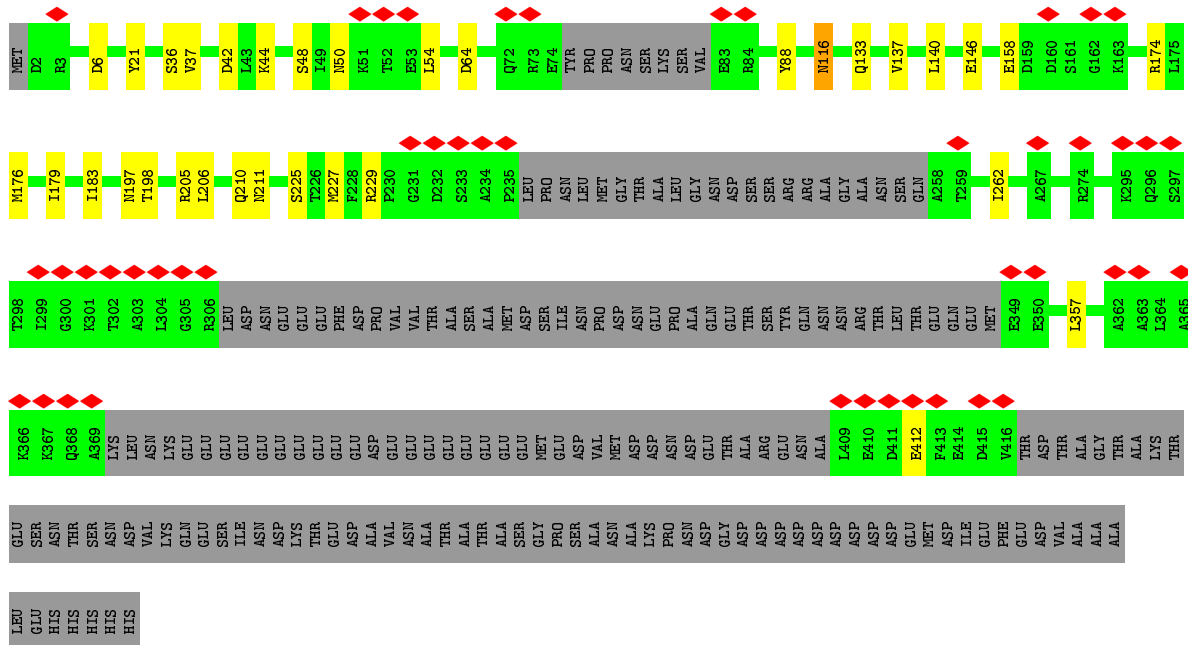




• Molecule 28: Transcription initiation factor IIA subunit 2

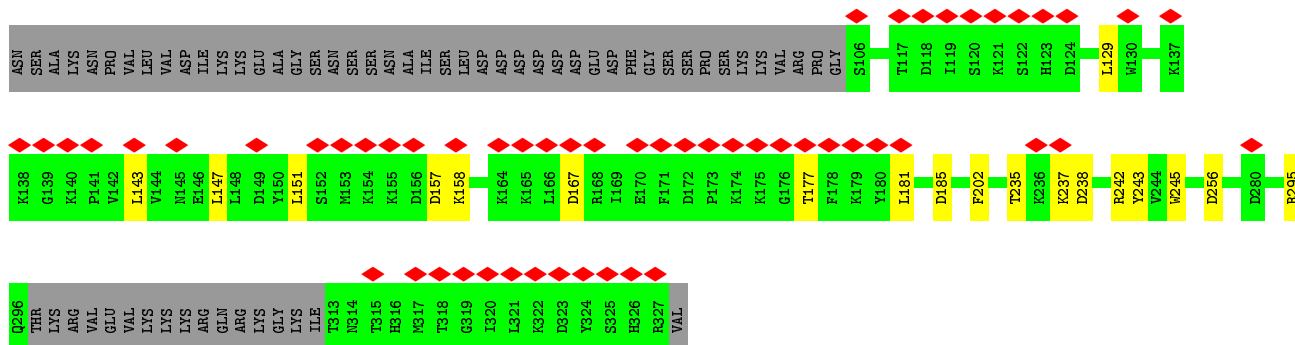


• Molecule 29: Transcription initiation factor IIE subunit alpha



• Molecule 30: Transcription initiation factor IIE subunit beta







## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	11014	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 ( $e^-/\text{\AA}^2$ )	43.6	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.206	Depositor
Minimum map value	-0.088	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.02	Depositor
Map size ( $\text{\AA}$ )	398.99997, 398.99997, 398.99997	wwPDB
Map dimensions	380, 380, 380	wwPDB
Map angles ( $^\circ$ )	90.0, 90.0, 90.0	wwPDB
Pixel spacing ( $\text{\AA}$ )	1.05, 1.05, 1.05	Depositor

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BEF, MG, ZN, ADP, SF4

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	0	0.24	0/6209	0.39	0/8384
2	1	0.23	0/4277	0.37	0/5755
3	2	0.23	0/3664	0.37	0/4952
4	3	0.23	0/1109	0.38	0/1492
5	4	0.24	0/2312	0.40	0/3127
6	5	0.23	0/520	0.37	0/701
7	6	0.23	0/2851	0.39	0/3856
8	7	0.24	0/5059	0.39	0/6841
9	A	0.24	0/11431	0.41	0/15458
10	B	0.24	0/9589	0.41	0/12934
11	C	0.24	0/2130	0.40	0/2887
12	D	0.23	0/1329	0.39	0/1783
13	E	0.24	0/1788	0.39	0/2406
14	F	0.24	0/709	0.40	0/956
15	G	0.25	0/1367	0.42	0/1844
16	H	0.24	0/1139	0.42	0/1544
17	I	0.24	0/962	0.43	0/1295
18	J	0.24	0/578	0.39	0/775
19	K	0.24	0/942	0.39	0/1272
20	L	0.23	0/361	0.41	0/478
21	M	0.23	0/2394	0.39	0/3221
22	N	0.56	0/1492	0.97	0/2301
23	O	0.25	0/1449	0.44	0/1952
24	Q	0.24	0/1914	0.40	0/2566
25	R	0.23	0/2209	0.41	0/2969
26	T	0.54	0/1496	0.91	0/2306
27	U	0.23	0/898	0.40	0/1212
28	V	0.22	0/822	0.40	0/1109
29	W	0.23	0/2509	0.37	0/3380
30	X	0.23	0/1684	0.38	0/2269
All	All	0.26	0/75193	0.44	0/102025

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	0	6091	0	6155	38	0
2	1	4214	0	4288	22	0
3	2	3597	0	3687	38	0
4	3	1089	0	1069	6	0
5	4	2274	0	2332	30	0
6	5	514	0	541	3	0
7	6	2794	0	2815	18	0
8	7	4954	0	4946	53	0
9	A	11230	0	11297	100	0
10	B	9404	0	9398	77	0
11	C	2092	0	2050	17	0
12	D	1321	0	1340	12	0
13	E	1752	0	1776	11	0
14	F	697	0	720	8	0
15	G	1339	0	1357	18	0
16	H	1120	0	1086	6	0
17	I	944	0	899	9	0
18	J	569	0	585	4	0
19	K	924	0	934	2	0
20	L	359	0	381	1	0
21	M	2367	0	2474	19	0
22	N	1332	0	737	13	0
23	O	1422	0	1500	10	0
24	Q	1878	0	1890	25	0
25	R	2173	0	2194	34	0
26	T	1333	0	733	10	0
27	U	885	0	866	14	0
28	V	815	0	822	10	0
29	W	2470	0	2467	27	0
30	X	1653	0	1683	14	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	0	8	0	0	0	0
32	3	2	0	0	0	0
32	4	1	0	0	0	0
32	6	4	0	0	0	0
32	A	2	0	0	0	0
32	B	1	0	0	0	0
32	C	1	0	0	0	0
32	I	2	0	0	0	0
32	J	1	0	0	0	0
32	L	1	0	0	0	0
32	M	1	0	0	0	0
32	W	1	0	0	0	0
33	7	4	0	0	0	0
34	7	1	0	0	0	0
34	A	1	0	0	0	0
35	7	27	0	12	1	0
All	All	73664	0	73034	558	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 4.

The worst 5 of 558 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)
9:A:30:ILE:HD11	10:B:1170:THR:HG21	1.38	0.99
13:E:127:ILE:HD11	13:E:132:ILE:HD11	1.57	0.86
10:B:810:GLU:OE1	10:B:815:ARG:NH1	2.09	0.85
10:B:868:MET:SD	21:M:182:ARG:NH1	2.50	0.84
21:M:181:ARG:O	21:M:229:ASN:ND2	2.11	0.83

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	0	750/778 (96%)	735 (98%)	15 (2%)	0	100	100
2	1	508/645 (79%)	506 (100%)	2 (0%)	0	100	100
3	2	435/517 (84%)	429 (99%)	6 (1%)	0	100	100
4	3	129/324 (40%)	127 (98%)	2 (2%)	0	100	100
5	4	287/341 (84%)	286 (100%)	1 (0%)	0	100	100
6	5	63/76 (83%)	61 (97%)	2 (3%)	0	100	100
7	6	352/464 (76%)	350 (99%)	2 (1%)	0	100	100
8	7	609/843 (72%)	595 (98%)	14 (2%)	0	100	100
9	A	1419/1733 (82%)	1384 (98%)	34 (2%)	1 (0%)	51	82
10	B	1168/1224 (95%)	1135 (97%)	33 (3%)	0	100	100
11	C	264/347 (76%)	259 (98%)	5 (2%)	0	100	100
12	D	161/221 (73%)	160 (99%)	1 (1%)	0	100	100
13	E	212/215 (99%)	207 (98%)	5 (2%)	0	100	100
14	F	84/155 (54%)	84 (100%)	0	0	100	100
15	G	169/177 (96%)	167 (99%)	2 (1%)	0	100	100
16	H	136/146 (93%)	134 (98%)	2 (2%)	0	100	100
17	I	114/122 (93%)	106 (93%)	8 (7%)	0	100	100
18	J	67/70 (96%)	66 (98%)	1 (2%)	0	100	100
19	K	113/120 (94%)	112 (99%)	1 (1%)	0	100	100
20	L	43/70 (61%)	42 (98%)	1 (2%)	0	100	100
21	M	303/352 (86%)	296 (98%)	7 (2%)	0	100	100
23	O	179/247 (72%)	174 (97%)	5 (3%)	0	100	100
24	Q	216/738 (29%)	210 (97%)	6 (3%)	0	100	100
25	R	255/400 (64%)	246 (96%)	9 (4%)	0	100	100
27	U	101/286 (35%)	98 (97%)	3 (3%)	0	100	100
28	V	100/129 (78%)	100 (100%)	0	0	100	100
29	W	294/492 (60%)	283 (96%)	11 (4%)	0	100	100
30	X	202/328 (62%)	191 (95%)	11 (5%)	0	100	100
All	All	8733/11560 (76%)	8543 (98%)	189 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
9	A	465	TYR

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	0	684/707 (97%)	684 (100%)	0	100	100
2	1	483/590 (82%)	482 (100%)	1 (0%)	93	98
3	2	408/470 (87%)	408 (100%)	0	100	100
4	3	125/305 (41%)	125 (100%)	0	100	100
5	4	260/302 (86%)	260 (100%)	0	100	100
6	5	59/68 (87%)	59 (100%)	0	100	100
7	6	323/419 (77%)	323 (100%)	0	100	100
8	7	547/737 (74%)	547 (100%)	0	100	100
9	A	1249/1520 (82%)	1249 (100%)	0	100	100
10	B	1024/1061 (96%)	1023 (100%)	1 (0%)	93	98
11	C	234/299 (78%)	234 (100%)	0	100	100
12	D	147/200 (74%)	147 (100%)	0	100	100
13	E	196/197 (100%)	196 (100%)	0	100	100
14	F	76/137 (56%)	76 (100%)	0	100	100
15	G	152/158 (96%)	152 (100%)	0	100	100
16	H	123/128 (96%)	123 (100%)	0	100	100
17	I	110/116 (95%)	110 (100%)	0	100	100
18	J	64/65 (98%)	64 (100%)	0	100	100
19	K	99/102 (97%)	99 (100%)	0	100	100
20	L	40/57 (70%)	40 (100%)	0	100	100
21	M	266/306 (87%)	266 (100%)	0	100	100
23	O	153/212 (72%)	153 (100%)	0	100	100
24	Q	205/642 (32%)	205 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
25	R	247/363 (68%)	247 (100%)	0	100	100
27	U	99/260 (38%)	99 (100%)	0	100	100
28	V	94/115 (82%)	94 (100%)	0	100	100
29	W	273/436 (63%)	272 (100%)	1 (0%)	91	95
30	X	185/295 (63%)	185 (100%)	0	100	100
All	All	7925/10267 (77%)	7922 (100%)	3 (0%)	100	100

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	1	440	ASN
10	B	928	ARG
29	W	116	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
9	A	253	ASN
9	A	1130	GLN
12	D	37	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 19 are monoatomic - leaving 3 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	$\# Z  > 2$
33	BEF	7	901	35	0,3,3	0.00	-	-		
35	ADP	7	903	34,33	24,29,29	0.69	0	29,45,45	0.70	0
31	SF4	0	801	1	0,12,12	0.00	-	-		

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
35	ADP	7	903	34,33	-	1/12/32/32	0/3/3/3
31	SF4	0	801	1	-	-	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
35	7	903	ADP	PA-O3A-PB-O2B

There are no ring outliers.

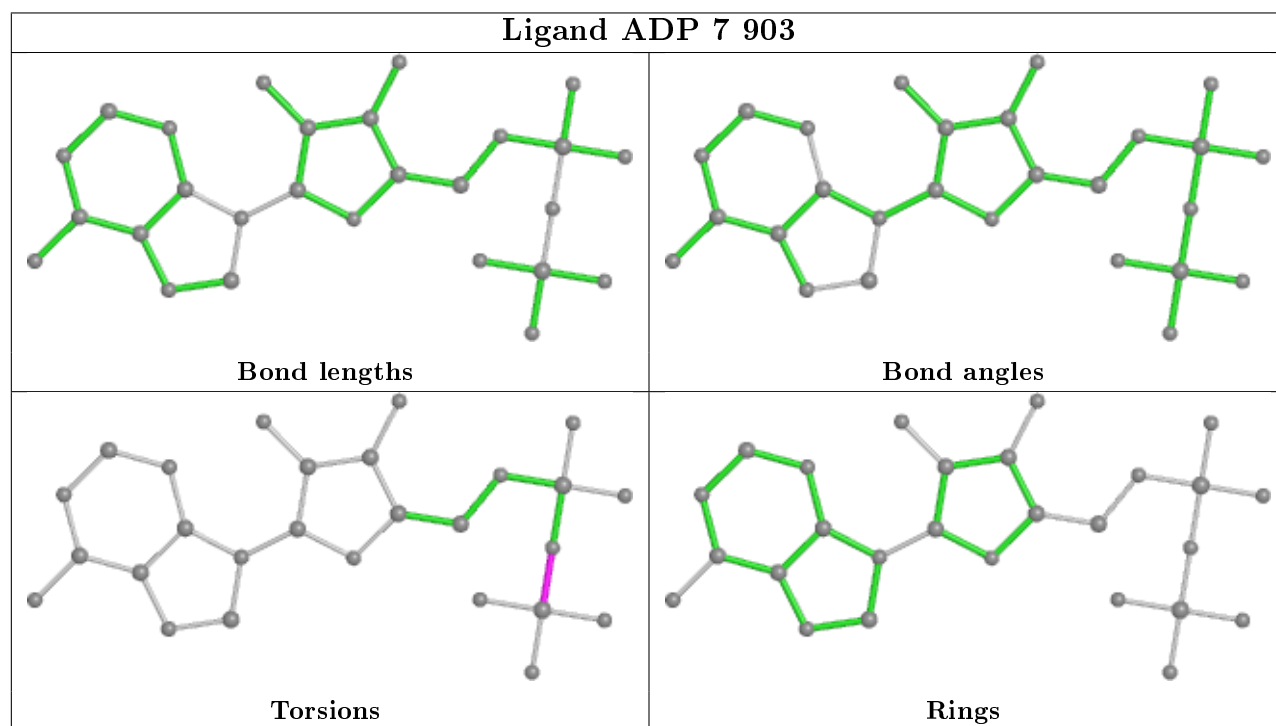
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
35	7	903	ADP	1	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 than 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.



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

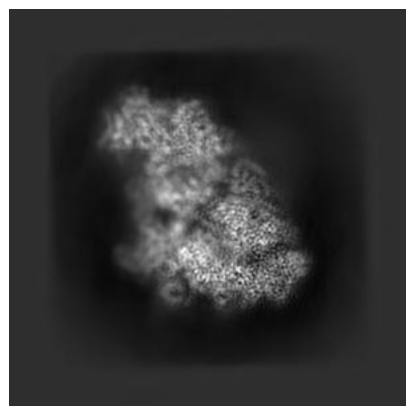
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-12743. These allow visual inspection of the internal detail of the map and identification of artifacts.

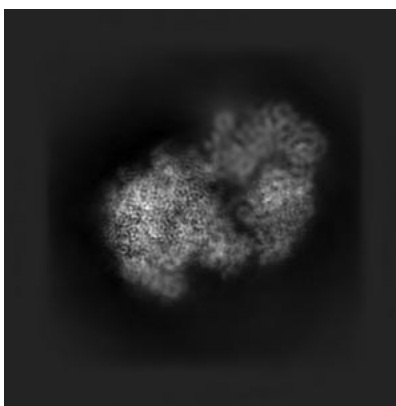
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

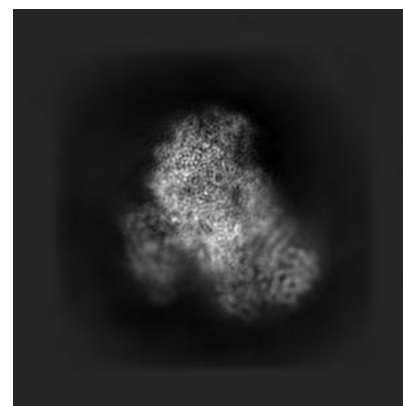
#### 6.1.1 Primary map



X

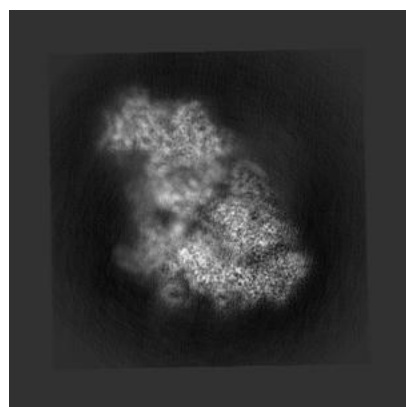


Y

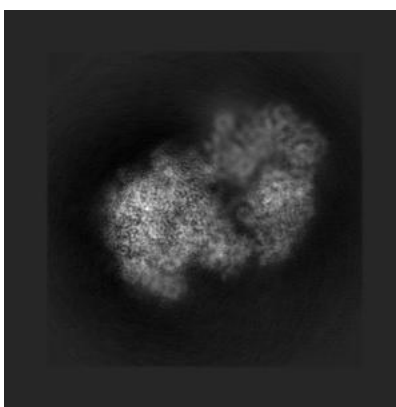


Z

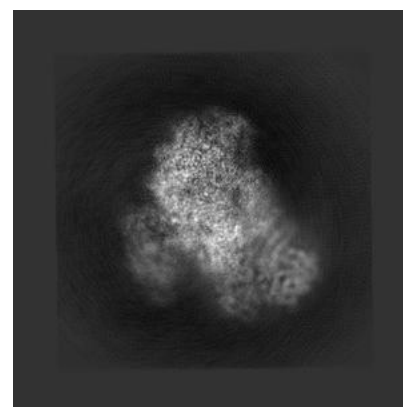
#### 6.1.2 Raw map



X



Y

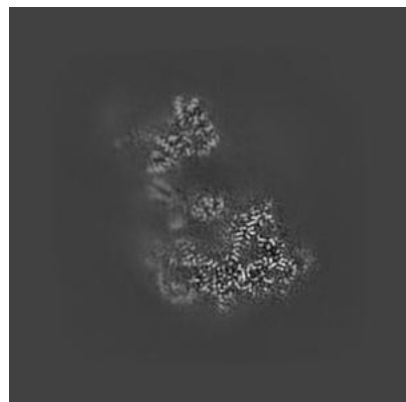


Z

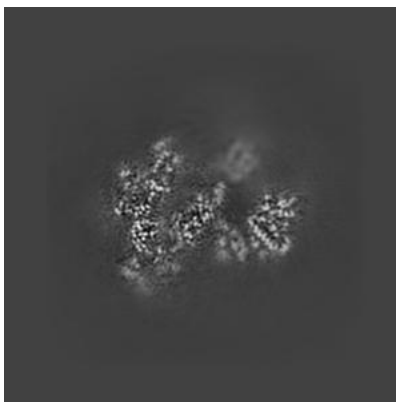
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

### 6.2.1 Primary map



X Index: 190

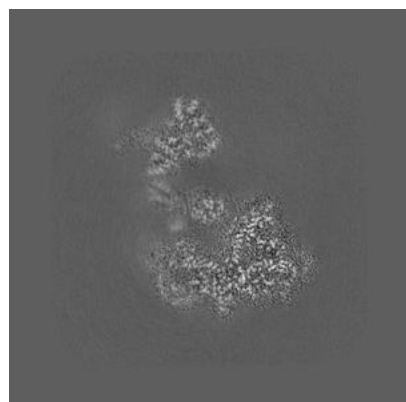


Y Index: 190

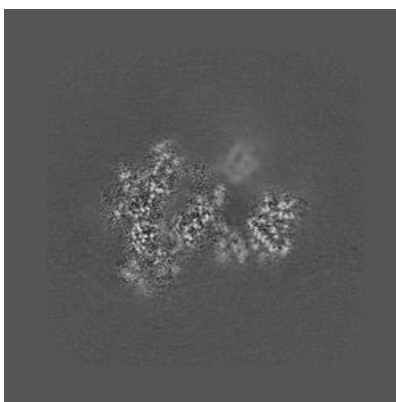


Z Index: 190

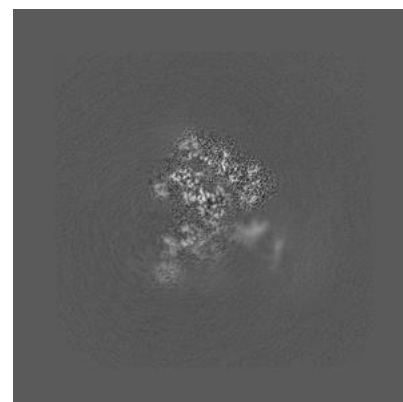
### 6.2.2 Raw map



X Index: 190



Y Index: 190

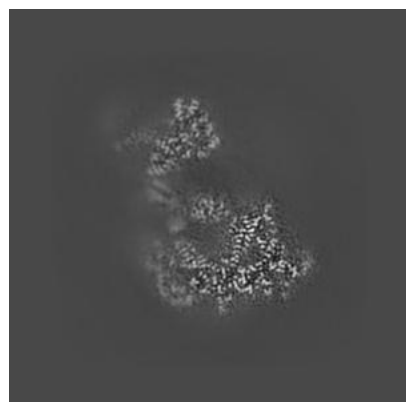


Z Index: 190

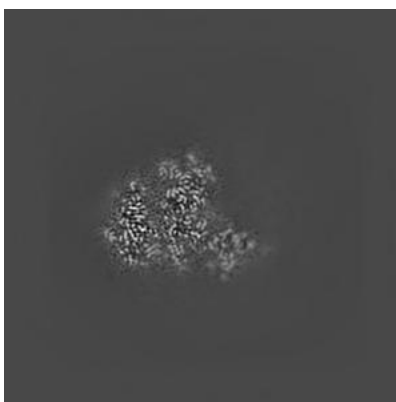
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

### 6.3.1 Primary map



X Index: 191

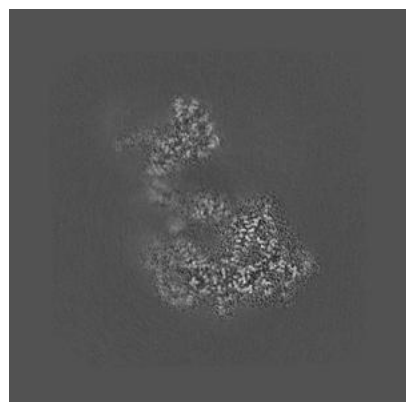


Y Index: 223

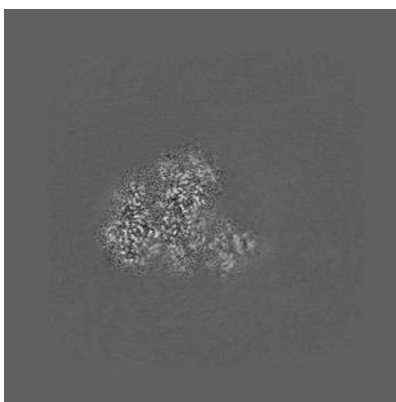


Z Index: 132

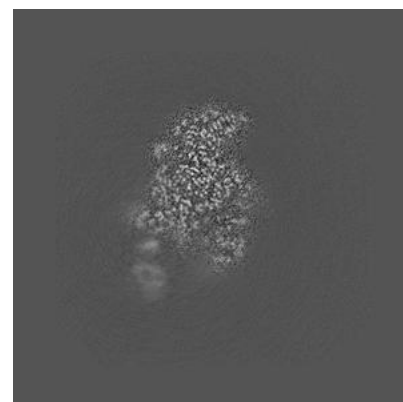
### 6.3.2 Raw map



X Index: 191



Y Index: 224

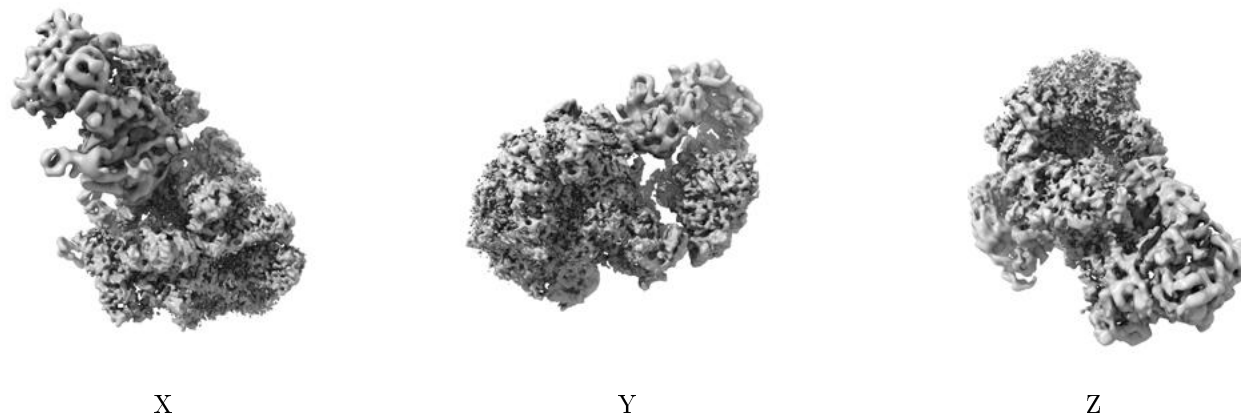


Z Index: 132

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

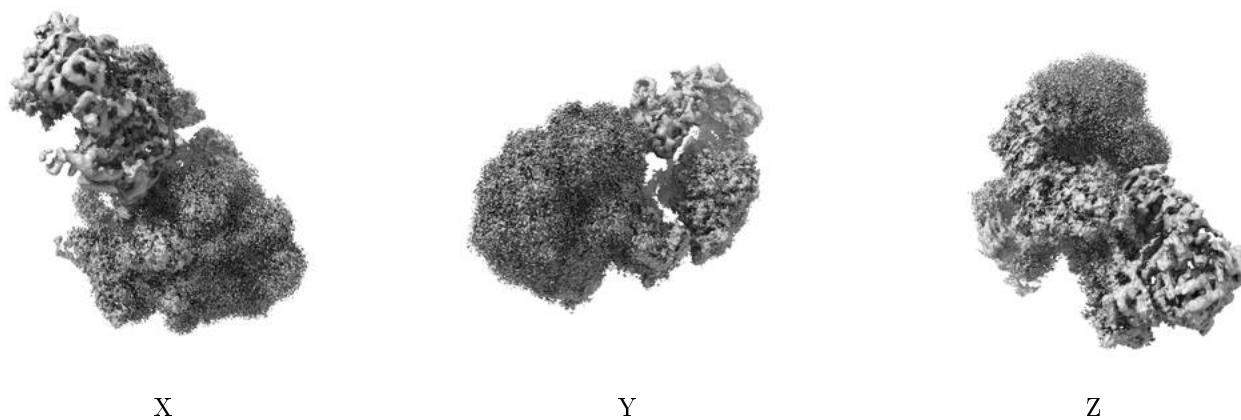
## 6.4 Orthogonal surface views [i](#)

### 6.4.1 Primary map



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.

### 6.4.2 Raw map



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.

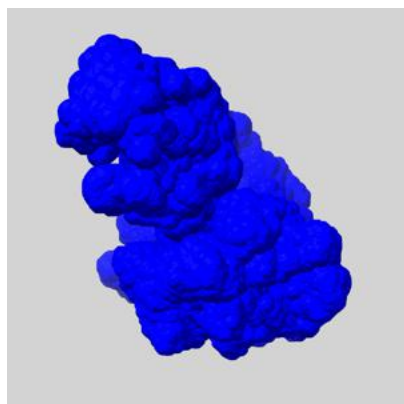
## 6.5 Mask visualisation [i](#)

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

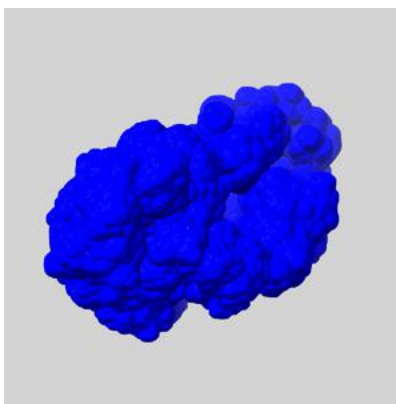
A mask typically either:

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

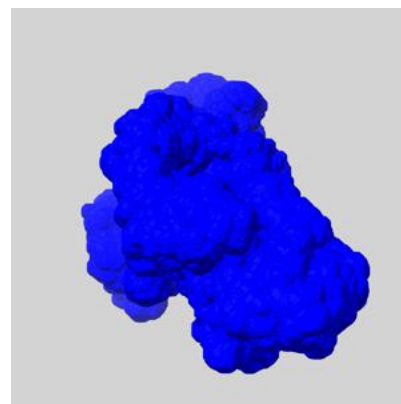
### 6.5.1 emd\_12743\_msk\_1.map [i](#)



X



Y

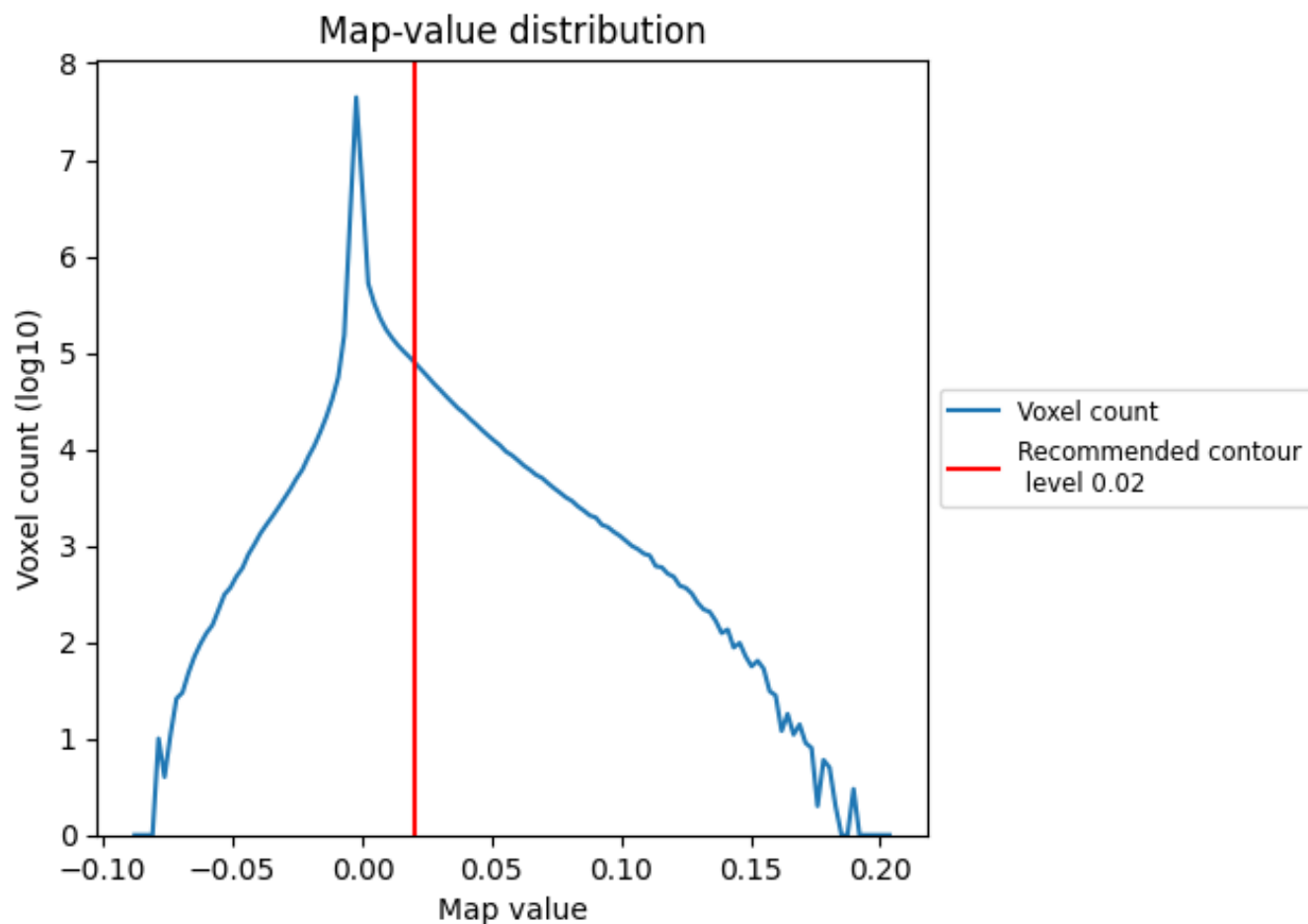


Z

## 7 Map analysis [i](#)

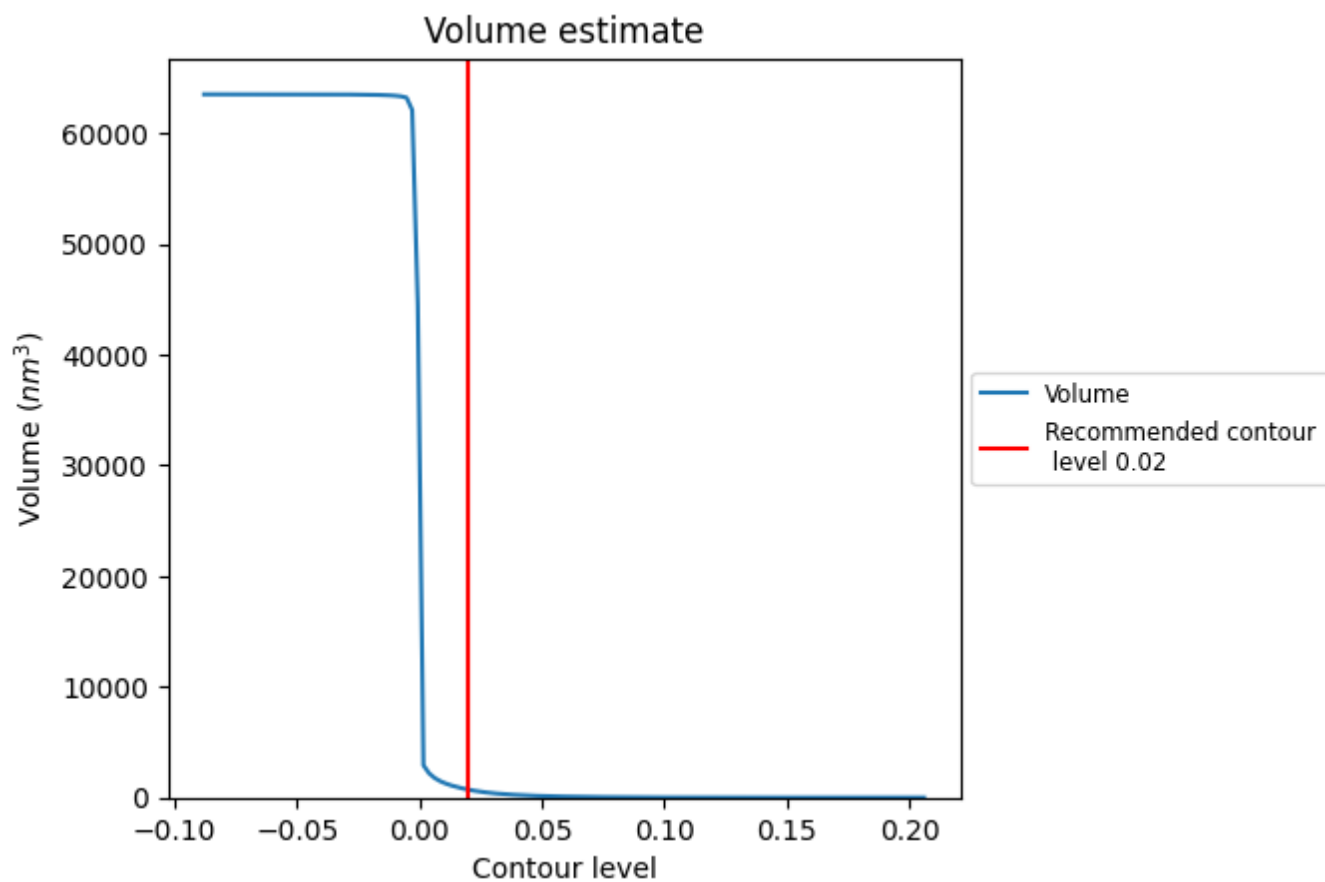
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



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

## 7.2 Volume estimate [\(i\)](#)

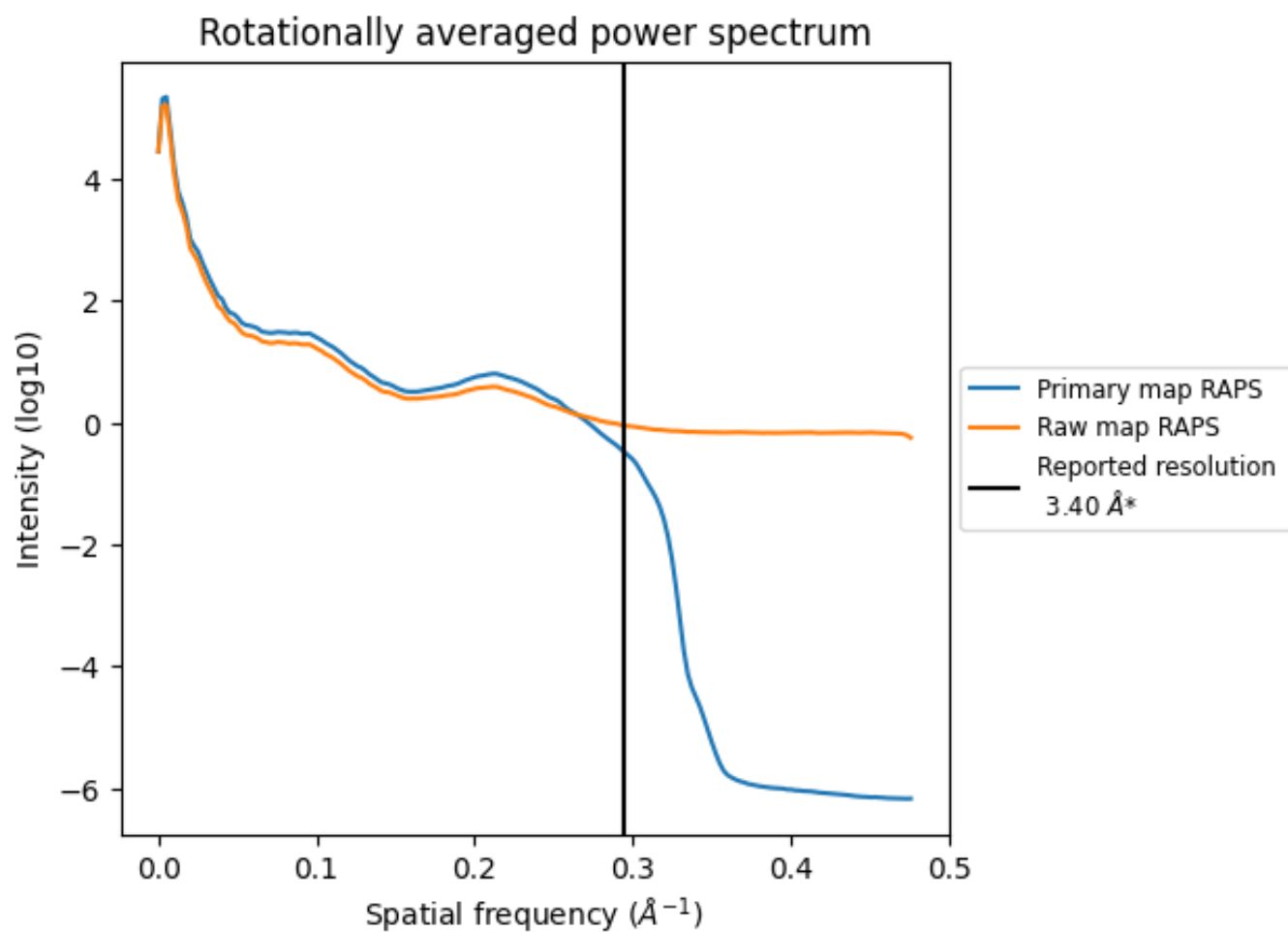


The volume at the recommended contour level is 712  $\text{nm}^3$ ; this corresponds to an approximate mass of 643 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.



### 7.3 Rotationally averaged power spectrum [\(i\)](#)

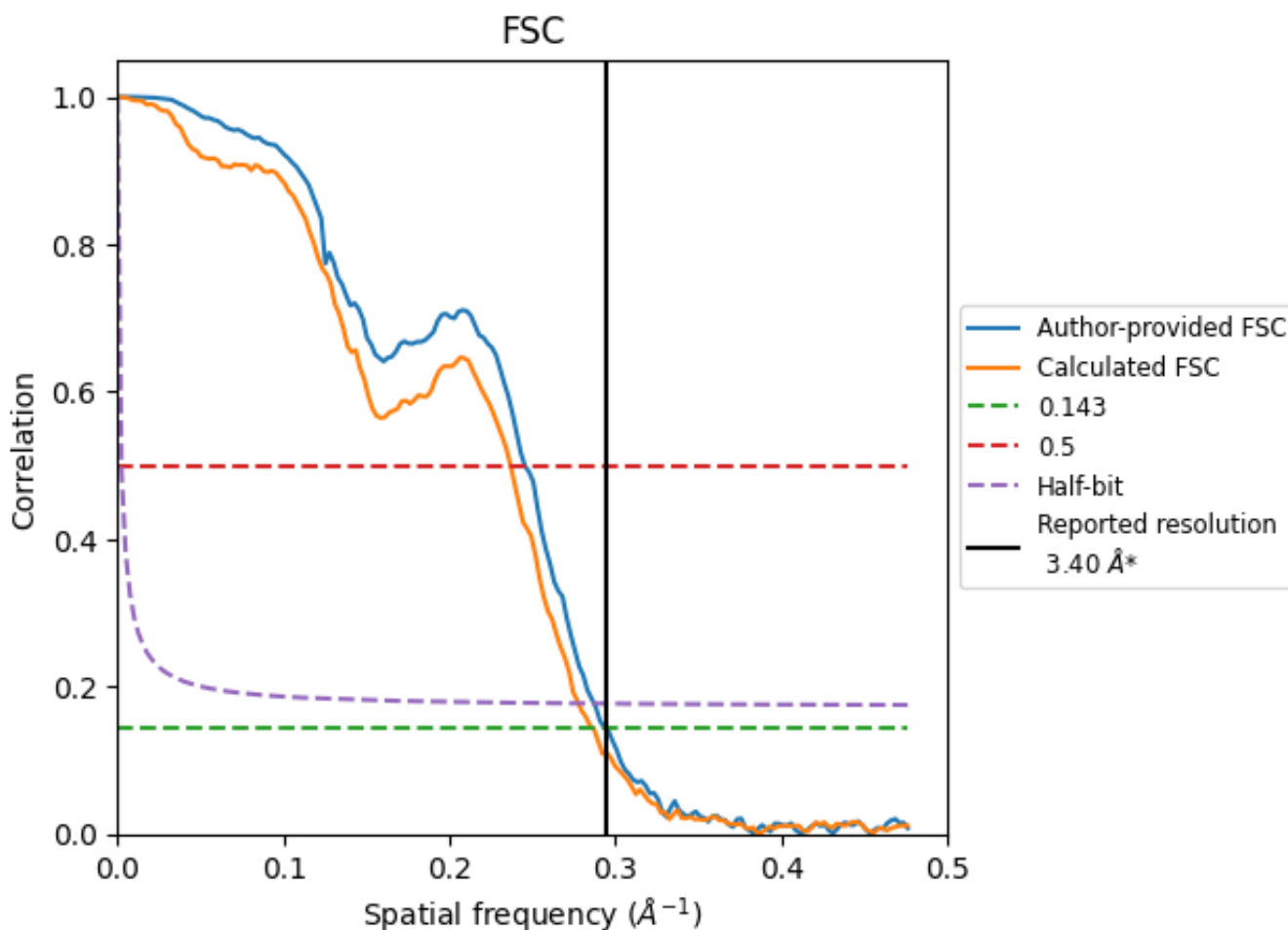


\*Reported resolution corresponds to spatial frequency of 0.294 Å<sup>-1</sup>

## 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 0.294 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

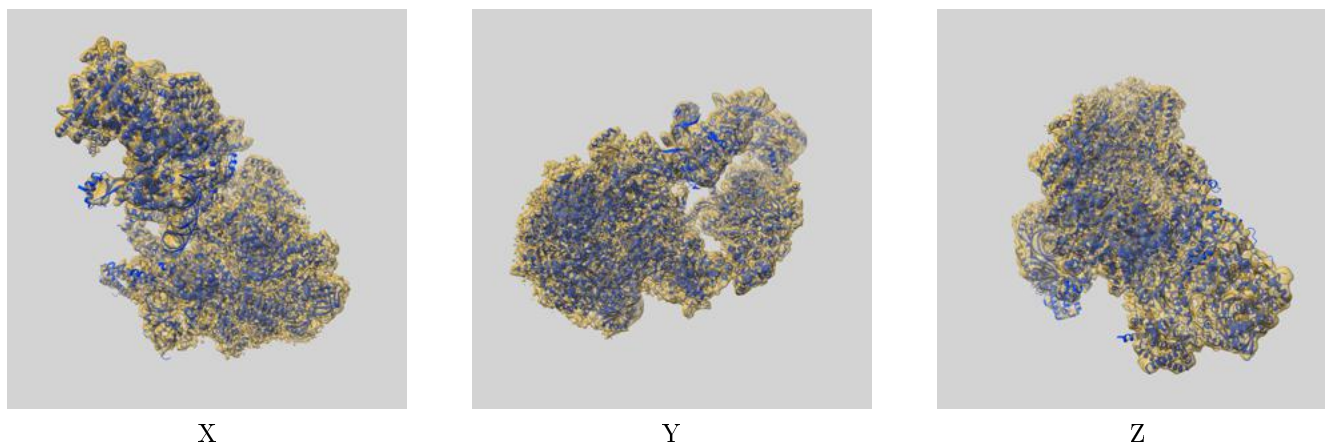
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.40	-	-
Author-provided FSC curve	3.40	4.07	3.48
Calculated*	3.49	4.22	3.60

\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

## 9 Map-model fit [i](#)

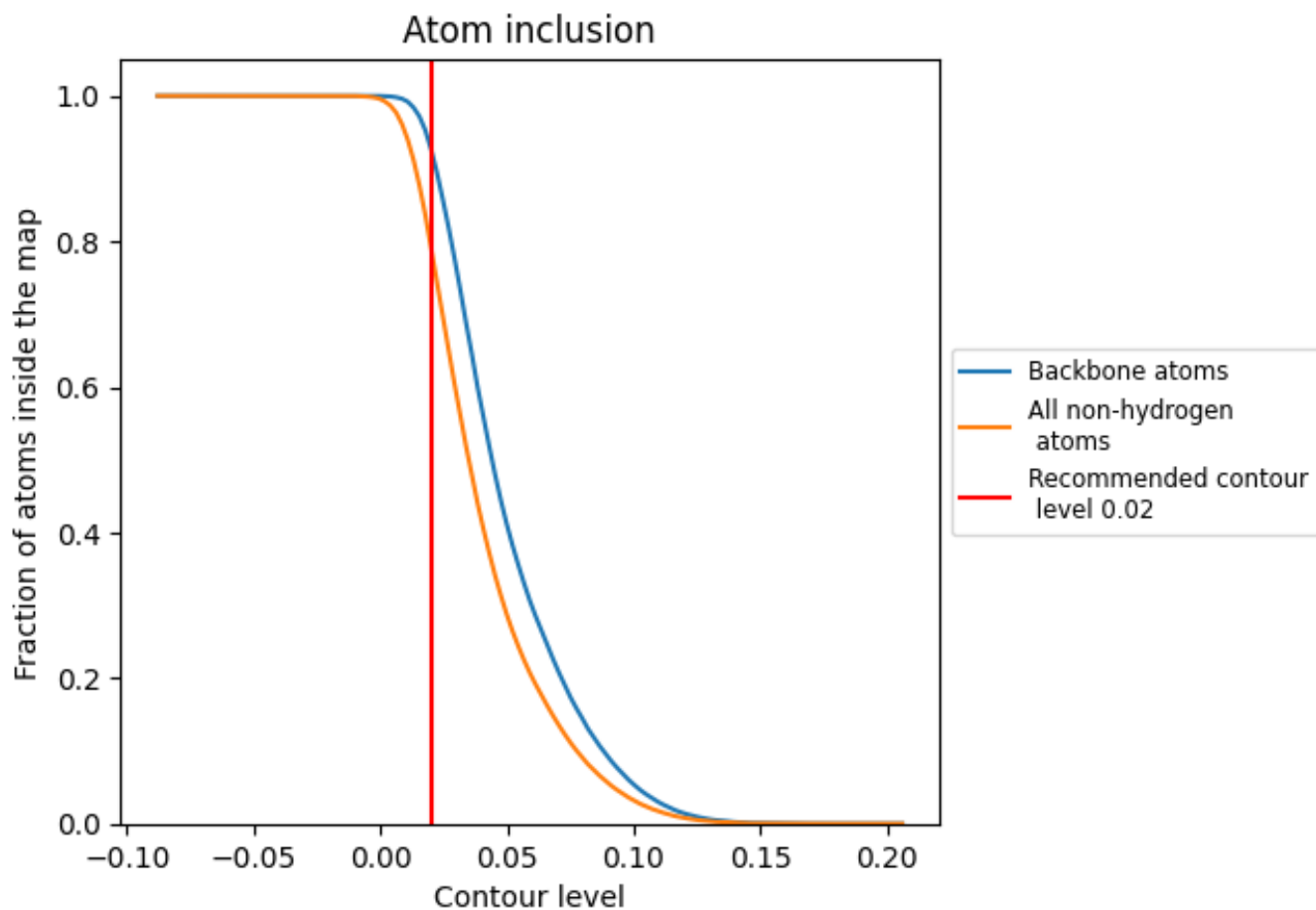
This section contains information regarding the fit between EMDB map EMD-12743 and PDB model 7O72. Per-residue inclusion information can be found in section 3 on page 15.

### 9.1 Map-model overlay [i](#)



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 Atom inclusion [i](#)



At the recommended contour level, 93% of all backbone atoms, 79% of all non-hydrogen atoms, are inside the map.