



wwPDB EM Validation Summary Report ⓘ

Jun 5, 2021 – 06:17 am BST

PDB ID : 7NVW
EMDB ID : EMD-12615
Title : TFIIH in a pre-translocated state (without ADP-BeF3)
Authors : Aibara, S.; Schilbach, S.; Cramer, P.
Deposited on : 2021-03-16
Resolution : 4.30 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 ⓘ symbol.

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

EMDB validation analysis : 0.0.0.dev75
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
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) : 2.19

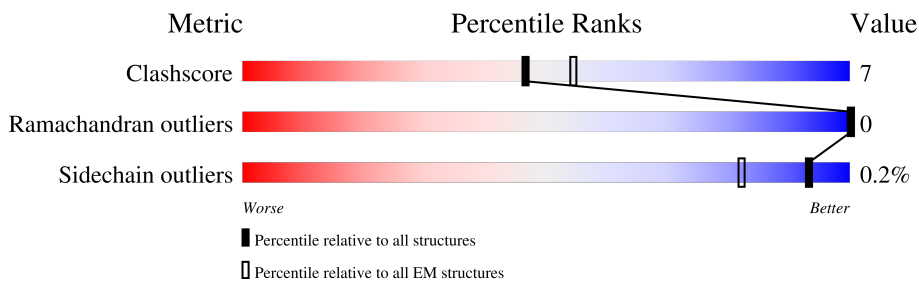
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.30 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.





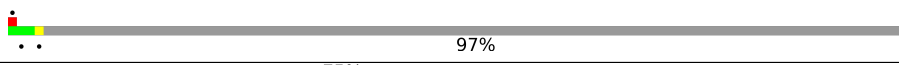

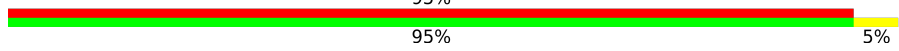
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 ≥ 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	760	
2	1	548	
3	2	462	
4	3	309	
5	4	308	
6	5	71	
7	6	395	
8	7	782	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	N	106	
10	T	106	
11	W	439	
12	Y	8	
13	Z	19	

2 Entry composition [i](#)

There are 15 unique types of molecules in this entry. The entry contains 23647 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 TFIID basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	0	714	5751	3683	999	1040	29	0	0

- Molecule 2 is a protein called General transcription factor IIF subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	1	265	2167	1382	378	395	12	0	0

- Molecule 3 is a protein called General transcription factor IIF subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	390	3158	2050	545	551	12	0	0

- Molecule 4 is a protein called CDK-activating kinase assembly factor MAT1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	3	149	1225	763	212	240	10	0	0

- Molecule 5 is a protein called General transcription factor IIF subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	4	263	2066	1323	344	380	19	0	0

- Molecule 6 is a protein called General transcription factor IIF subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	66	523	337	83	100	3	0	0

- Molecule 7 is a protein called General transcription factor IIIH subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	6	347	2732	1726	471	508	27	0	0

- 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	605	4890	3127	848	885	30	0	0

- Molecule 9 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
9	N	21	421	202	65	133	21	0	0

- Molecule 10 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
10	T	21	440	206	94	119	21	0	0

- Molecule 11 is a protein called General transcription factor IIE subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
11	W	15	124	78	24	22	0	0

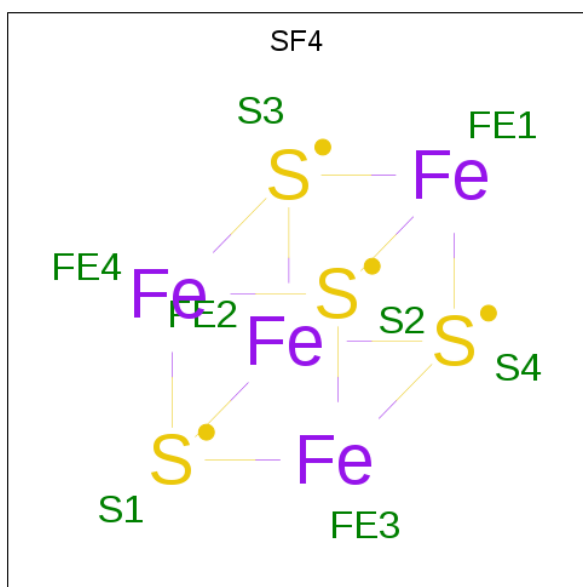
- Molecule 12 is a protein called Unassigned Peptide, likely XPB.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
12	Y	8	40	24	8	8	0	0

- Molecule 13 is a protein called Unassigned Peptide, likely TFIIIE-Beta.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
13	Z	19	95	57	19	19	0	0

- Molecule 14 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe₄S₄).



Mol	Chain	Residues	Atoms			AltConf
14	0	1	Total	Fe	S	0
			8	4	4	

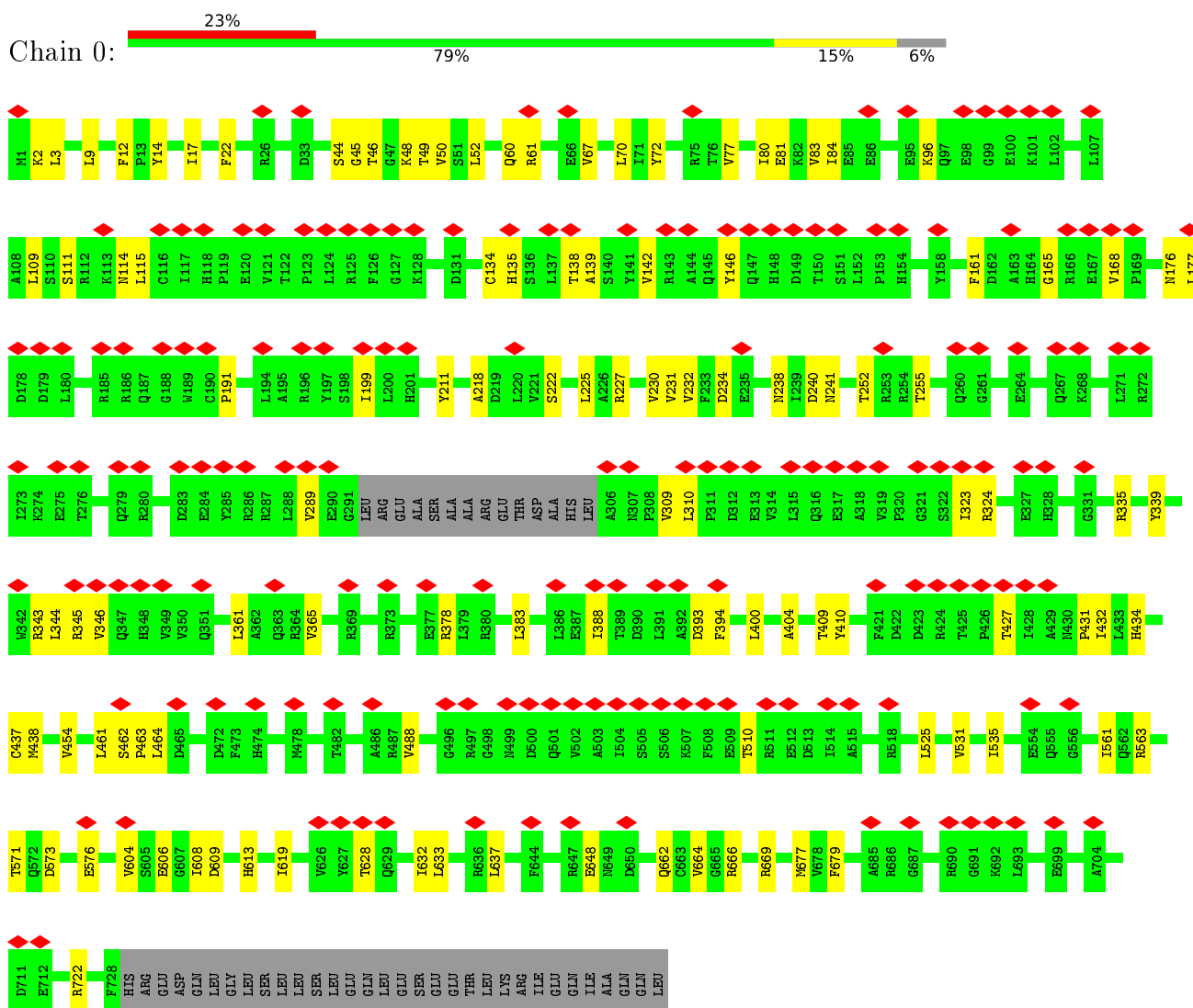
- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
15	3	2	Total	Zn	0
			2	2	
15	4	2	Total	Zn	0
			2	2	
15	6	3	Total	Zn	0
			3	3	

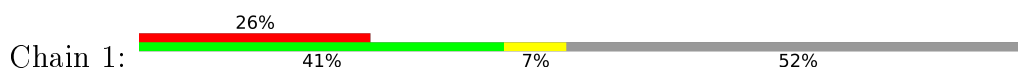
3 Residue-property plots

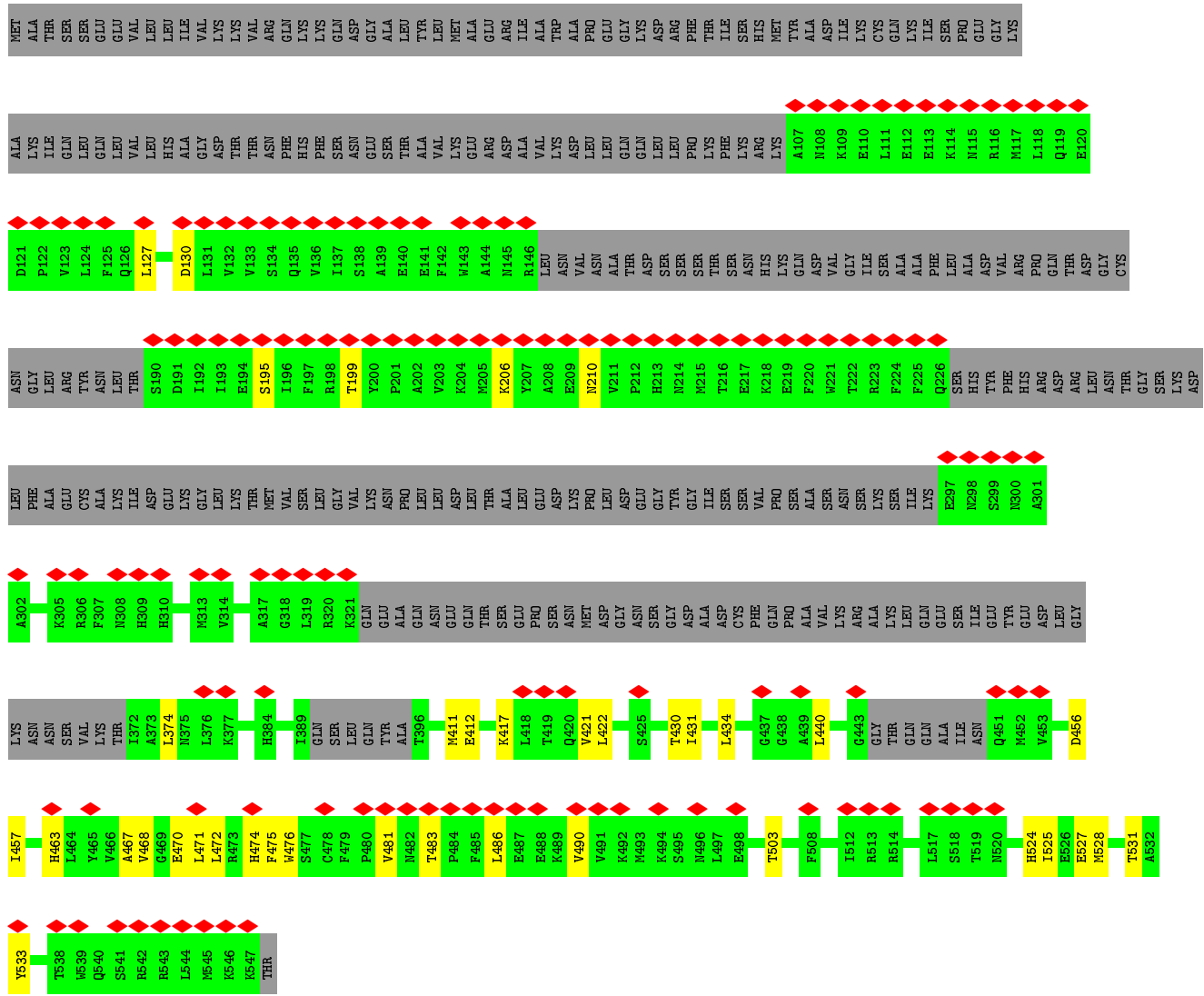
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

- Molecule 1: TFIIF basal transcription factor complex helicase XPD subunit

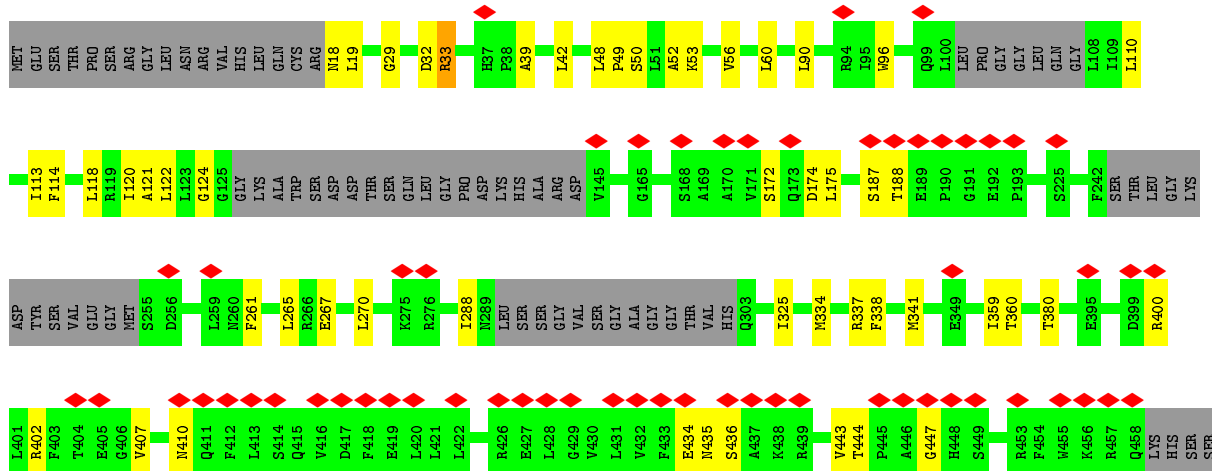
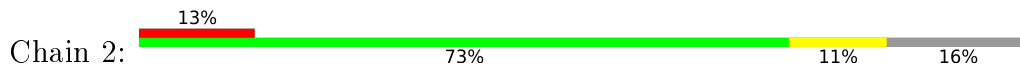


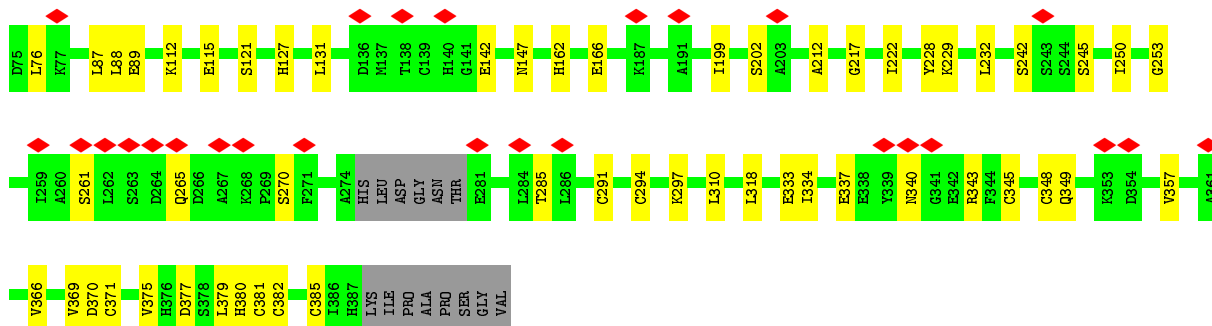
- Molecule 2: General transcription factor IIH subunit 1



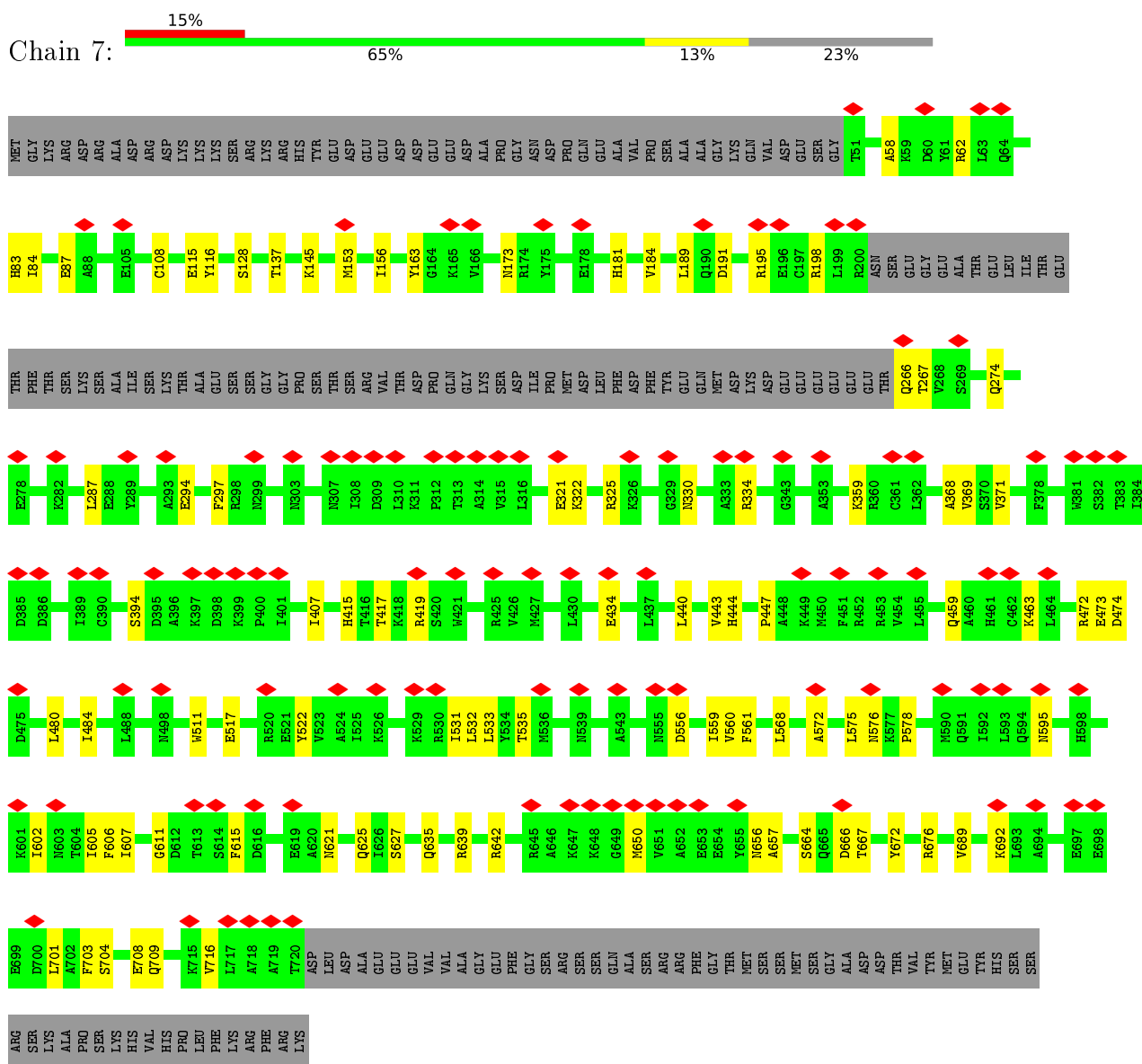


- Molecule 3: General transcription factor IIH subunit 4



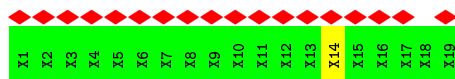
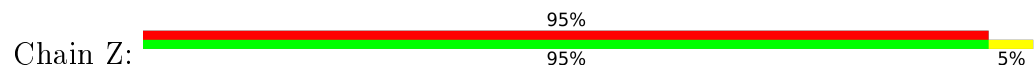


• Molecule 8: General transcription and DNA repair factor IIIH helicase subunit XPB



• Molecule 9: Non-template DNA





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	129156	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 ($e^-/\text{\AA}^2$)	41.1	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	86.632	Depositor
Minimum map value	-56.218	Depositor
Average map value	-0.005	Depositor
Map value standard deviation	0.935	Depositor
Recommended contour level	11.0	Depositor
Map size (Å)	472.49997, 472.49997, 472.49997	wwPDB
Map dimensions	450, 450, 450	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, 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	0/5875	0.39	0/7955
2	1	0.24	0/2210	0.37	0/2975
3	2	0.24	0/3230	0.38	0/4376
4	3	0.24	0/1241	0.37	0/1665
5	4	0.25	0/2103	0.41	0/2846
6	5	0.24	0/529	0.38	0/714
7	6	0.25	0/2793	0.40	0/3780
8	7	0.24	0/4994	0.38	0/6745
9	N	0.55	0/467	1.01	0/717
10	T	0.47	0/497	0.75	0/766
11	W	0.21	0/126	0.46	0/169
All	All	0.26	0/24065	0.42	0/32708

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 [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	0	5751	0	5794	78	0
2	1	2167	0	2175	33	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	2	3158	0	3213	43	0
4	3	1225	0	1224	13	0
5	4	2066	0	2097	51	0
6	5	523	0	530	7	0
7	6	2732	0	2698	44	0
8	7	4890	0	4949	65	0
9	N	421	0	240	4	0
10	T	440	0	233	2	0
11	W	124	0	126	2	0
12	Y	40	0	10	1	0
13	Z	95	0	24	1	0
14	0	8	0	0	0	0
15	3	2	0	0	0	0
15	4	2	0	0	0	0
15	6	3	0	0	0	0
All	All	23647	0	23313	306	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

The worst 5 of 306 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)
8:7:371:VAL:HG23	8:7:407:ILE:HG22	1.57	0.85
1:0:722:ARG:NH2	7:6:202:SER:OG	2.12	0.82
1:0:252:THR:HG22	1:0:432:ILE:HG22	1.61	0.82
8:7:394:SER:OG	10:T:-41:DG:OP2	1.97	0.81
6:5:24:ASP:OD1	6:5:25:GLU:N	2.15	0.79

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	710/760 (93%)	680 (96%)	30 (4%)	0	100	100
2	1	253/548 (46%)	240 (95%)	13 (5%)	0	100	100
3	2	380/462 (82%)	365 (96%)	15 (4%)	0	100	100
4	3	147/309 (48%)	137 (93%)	10 (7%)	0	100	100
5	4	259/308 (84%)	252 (97%)	7 (3%)	0	100	100
6	5	64/71 (90%)	63 (98%)	1 (2%)	0	100	100
7	6	341/395 (86%)	325 (95%)	16 (5%)	0	100	100
8	7	601/782 (77%)	572 (95%)	29 (5%)	0	100	100
11	W	13/439 (3%)	12 (92%)	1 (8%)	0	100	100
All	All	2768/4074 (68%)	2646 (96%)	122 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	624/664 (94%)	623 (100%)	1 (0%)	93	96
2	1	241/484 (50%)	241 (100%)	0	100	100
3	2	342/399 (86%)	341 (100%)	1 (0%)	92	95
4	3	143/283 (50%)	142 (99%)	1 (1%)	84	90
5	4	234/272 (86%)	234 (100%)	0	100	100
6	5	59/64 (92%)	59 (100%)	0	100	100
7	6	311/352 (88%)	310 (100%)	1 (0%)	92	95
8	7	536/688 (78%)	534 (100%)	2 (0%)	91	94
11	W	13/373 (4%)	13 (100%)	0	100	100
All	All	2503/3579 (70%)	2497 (100%)	6 (0%)	93	96

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	6	343	ARG
8	7	198	ARG
8	7	692	LYS
3	2	33	ARG
1	0	96	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	0	114	ASN
5	4	258	HIS

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

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
14	SF4	0	1000	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
14	SF4	0	1000	1	-	-	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

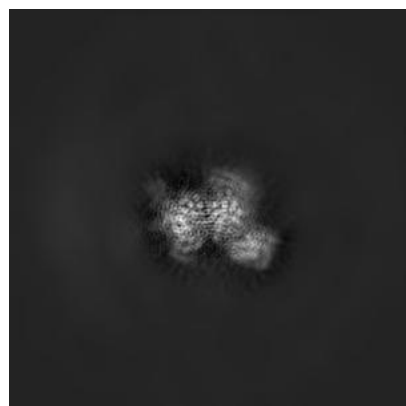
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-12615. These allow visual inspection of the internal detail of the map and identification of artifacts.

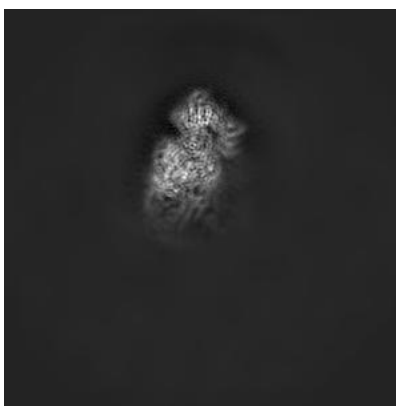
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

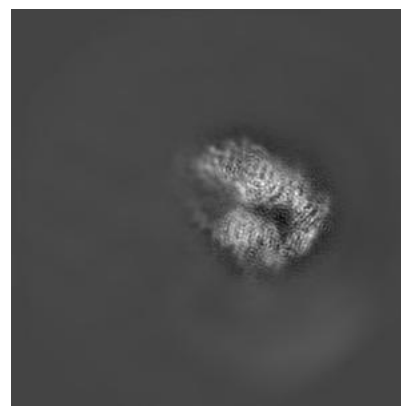
6.1.1 Primary map



X

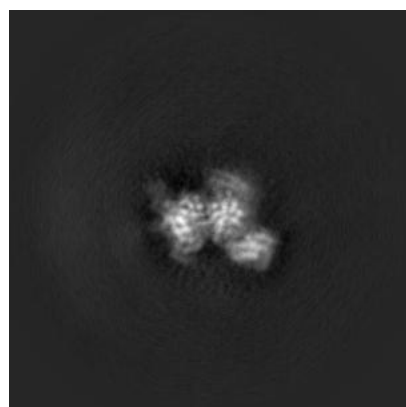


Y

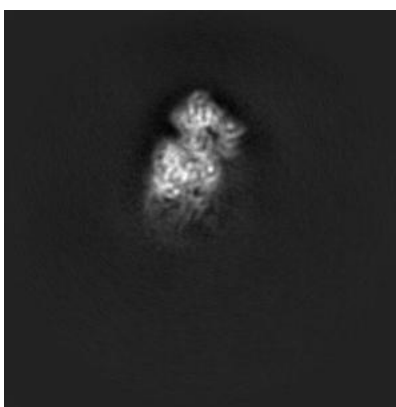


Z

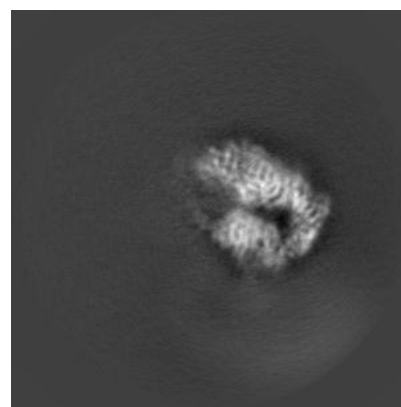
6.1.2 Raw map



X



Y

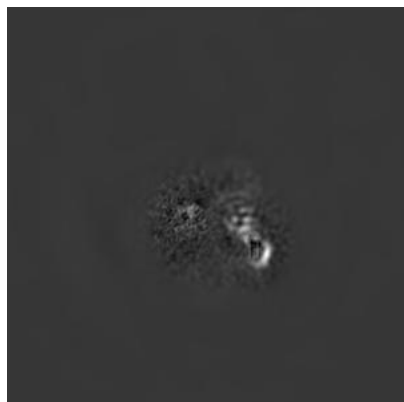


Z

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

6.2 Central slices [i](#)

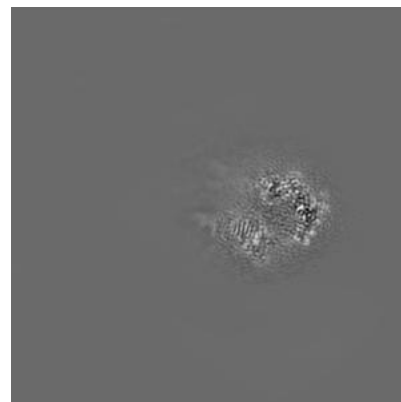
6.2.1 Primary map



X Index: 225

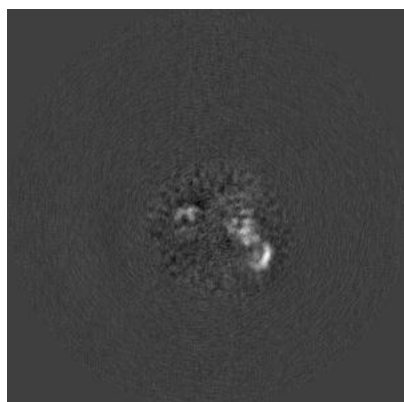


Y Index: 225

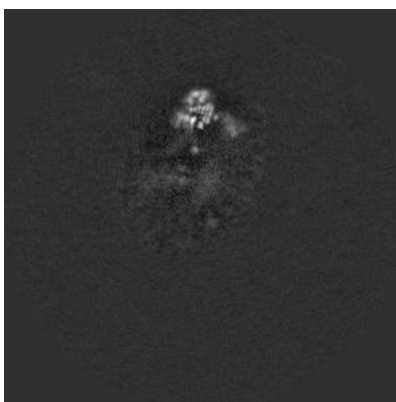


Z Index: 225

6.2.2 Raw map



X Index: 225



Y Index: 225

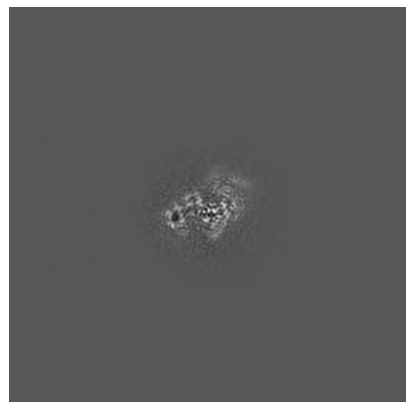


Z Index: 225

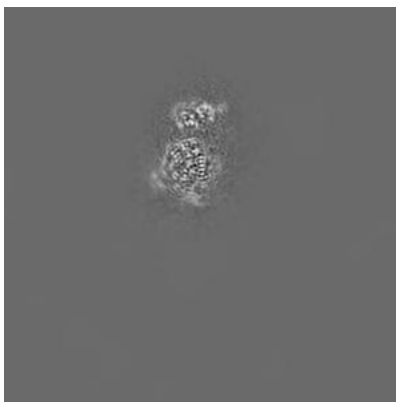
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

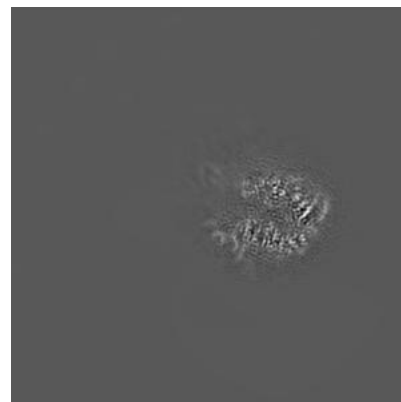
6.3.1 Primary map



X Index: 321

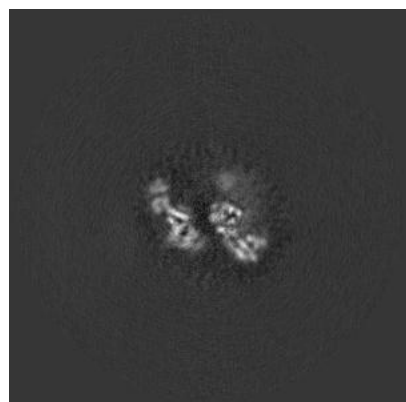


Y Index: 195

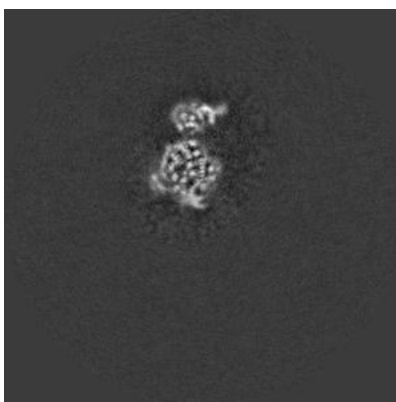


Z Index: 214

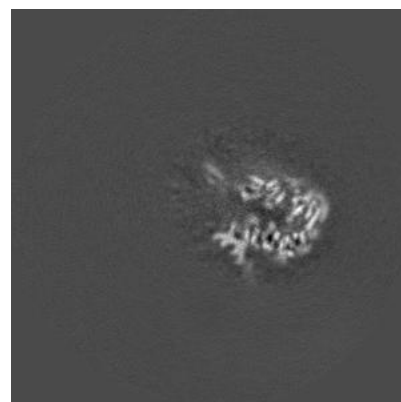
6.3.2 Raw map



X Index: 288



Y Index: 195

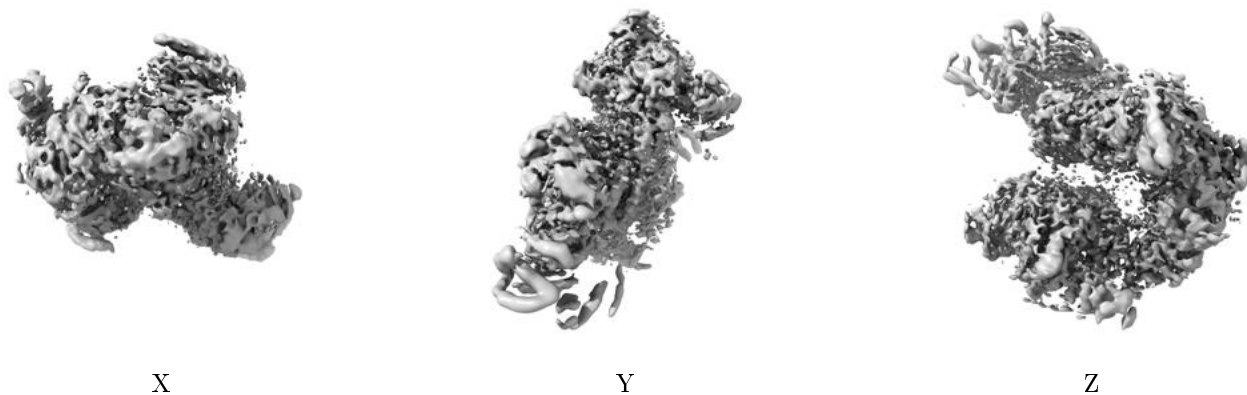


Z Index: 211

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 11.0. 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.

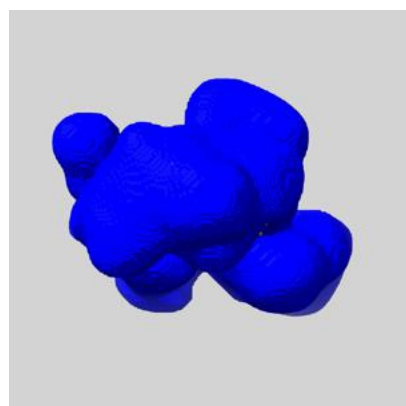
6.5 Mask visualisation [i](#)

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

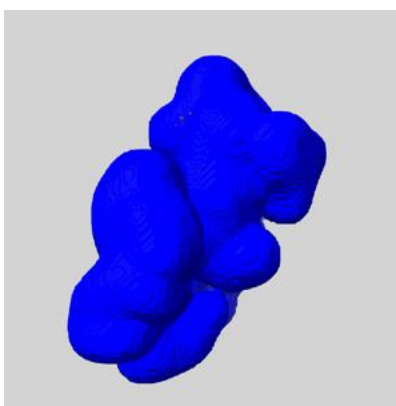
A mask typically either:

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

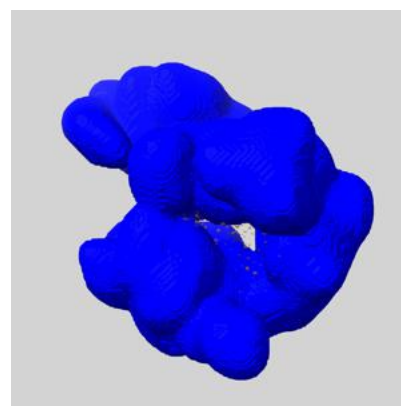
6.5.1 emd_12615_msk_1.map [i](#)



X



Y

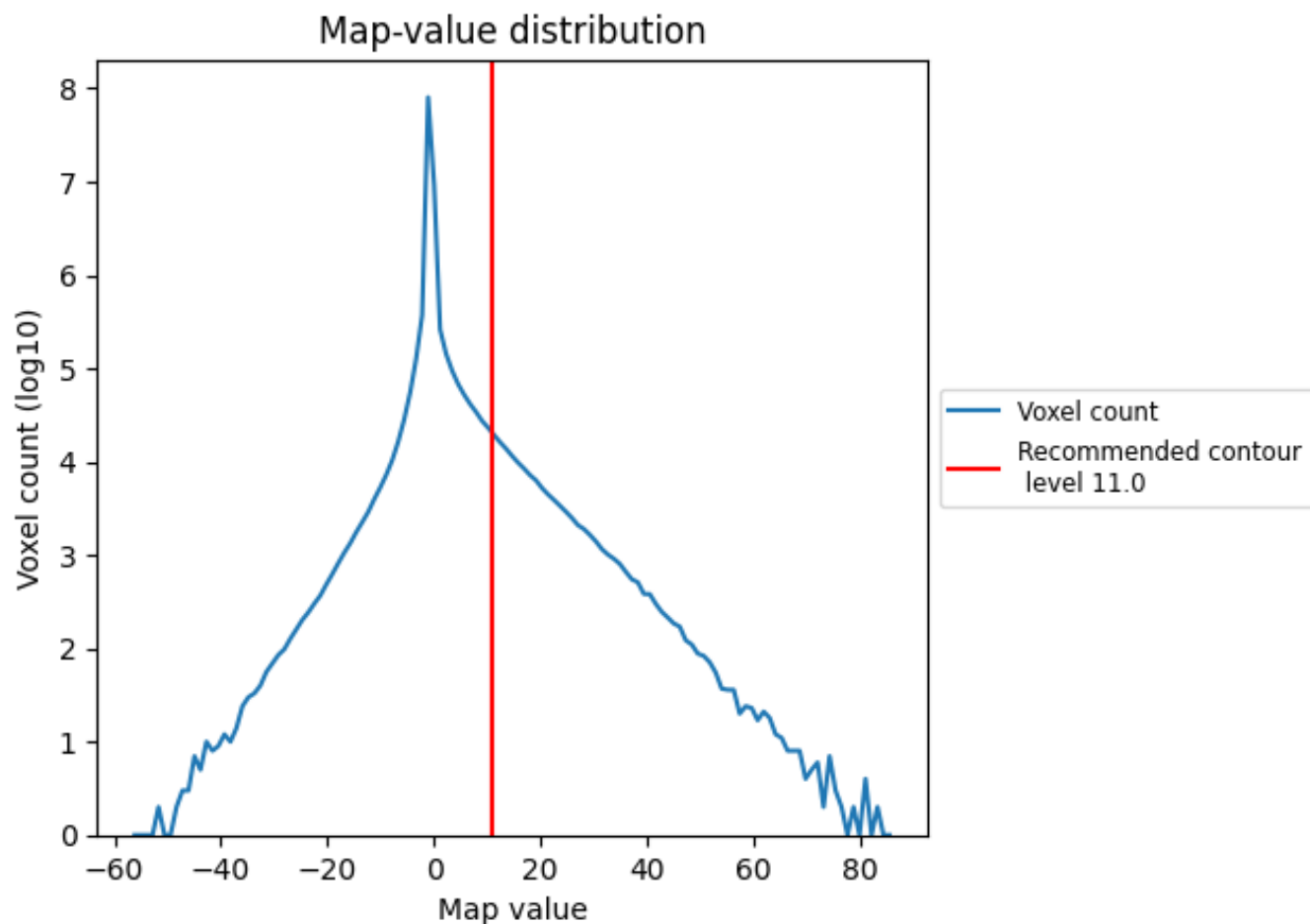


Z

7 Map analysis [i](#)

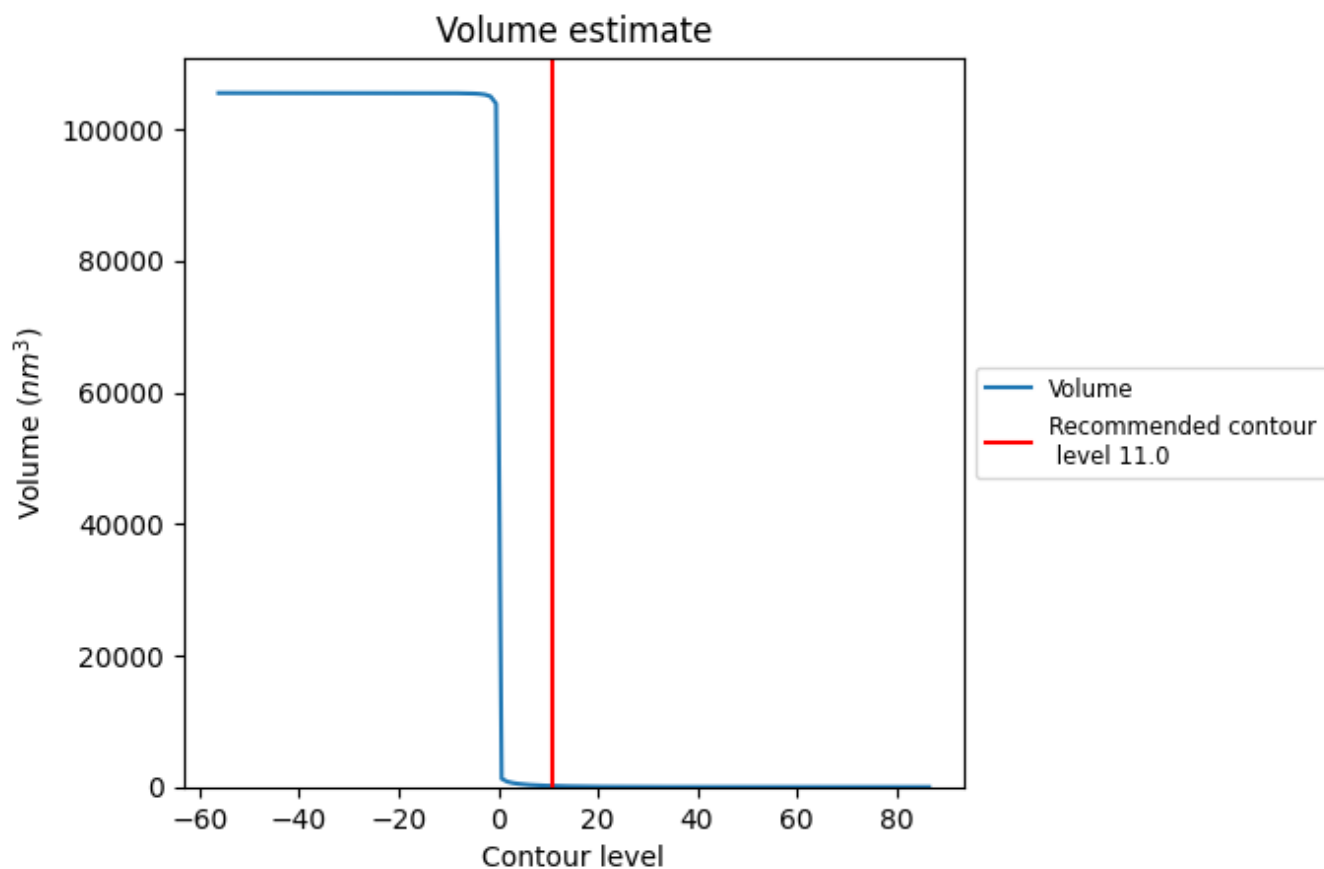
This section contains the results of statistical analysis of the map.

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.

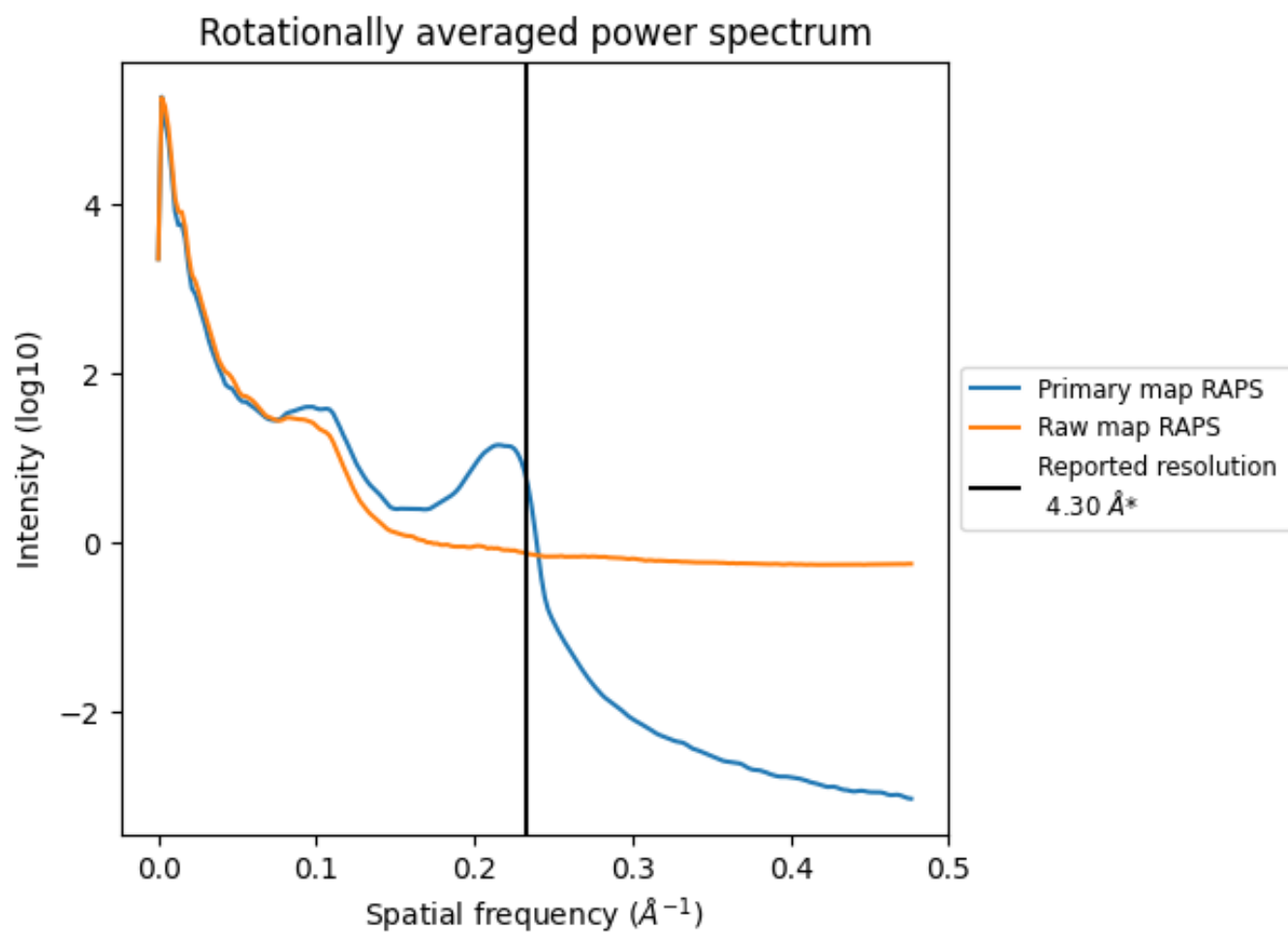
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 157 nm^3 ; this corresponds to an approximate mass of 142 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 [\(i\)](#)

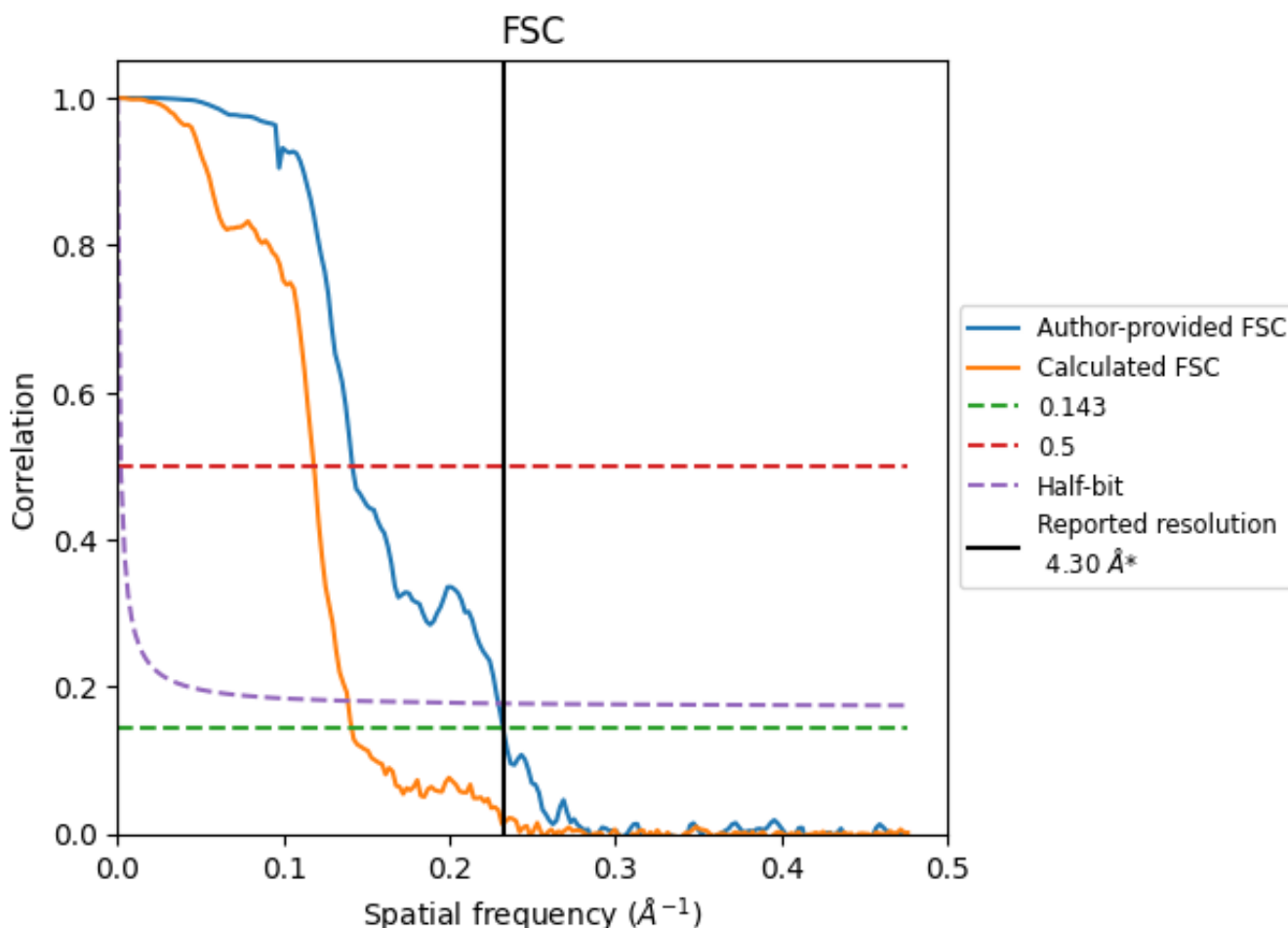


*Reported resolution corresponds to spatial frequency of 0.233 Å⁻¹

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 0.233 Å⁻¹

8.2 Resolution estimates [i](#)

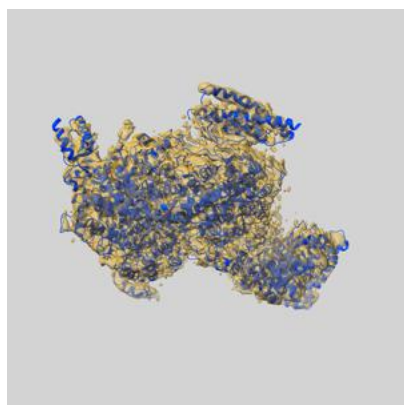
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.30	-	-
Author-provided FSC curve	4.30	7.06	4.36
Calculated*	7.08	8.45	7.20

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 7.08 differs from the reported value 4.3 by more than 10 %

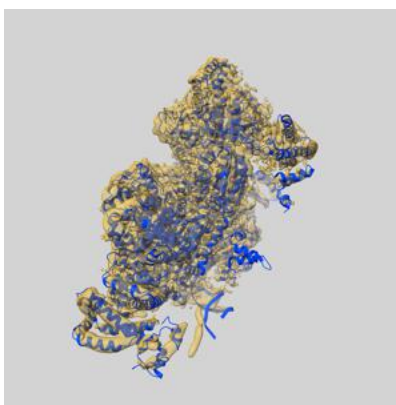
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-12615 and PDB model 7NVW. Per-residue inclusion information can be found in section 3 on page 7.

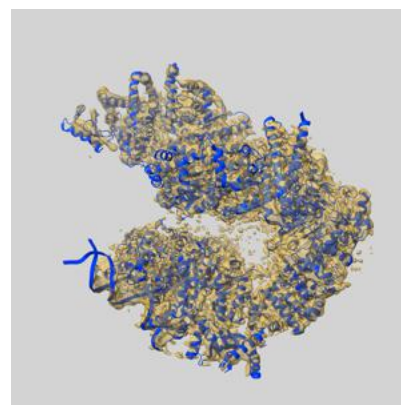
9.1 Map-model overlay [i](#)



X



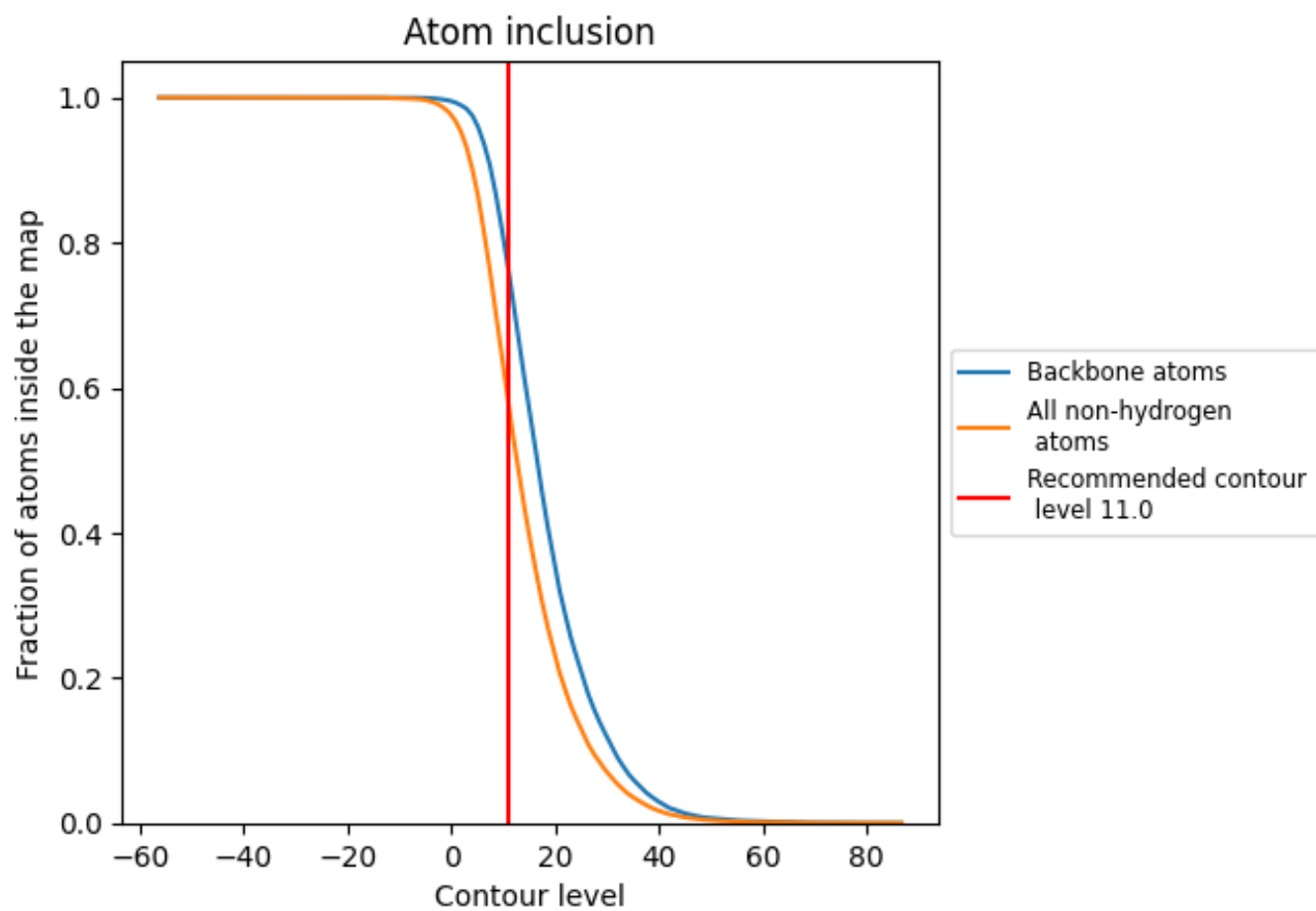
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 11.0 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, 76% of all backbone atoms, 58% of all non-hydrogen atoms, are inside the map.