



wwPDB EM Validation Summary Report ⓘ

Aug 27, 2026 – 09:15 am BST

PDB ID : 6RO4 / pdb_00006ro4
EMDB ID : EMD-4970
Title : Structure of the core TFIIH-XPA-DNA complex
Authors : Kokic, G.; Chernev, A.; Tegunov, D.; Dienemann, C.; Urlaub, H.; Cramer, P.
Deposited on : 2019-05-10
Resolution : 3.50 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

EMDB validation analysis : 0.0.1.dev133
Mogul : 1.8.4, CSD as541be (2020)
MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : 202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ : 1.9.13
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.50

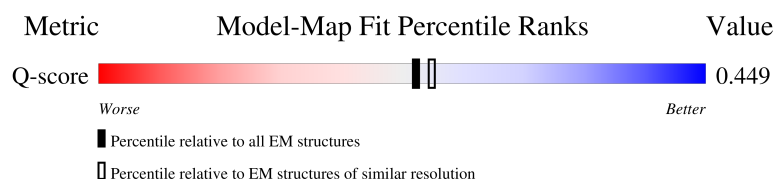
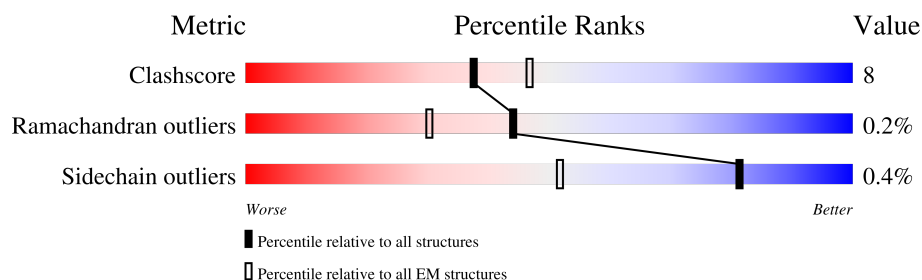
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 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.



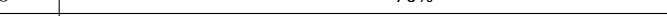
Metric	Whole archive (#Entries)	EM structures (#Entries)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
Q-score	-	25397	13950 (3.00 - 4.00)

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	J	49	 24% 14% 61%
2	K	49	 6% 33% 24% 43%
3	A	782	 15% 64% 9% 26%
4	F	71	 15% 72% 15% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	760	
6	E	308	
7	D	395	
8	C	462	
9	G	273	

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 19827 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 DNA chain called DNA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	J	19	Total	C	N	O	P	0	0
			387	184	71	113	19		

- Molecule 2 is a DNA chain called DNA2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	K	28	Total	C	N	O	P	0	0
			575	273	111	163	28		

- Molecule 3 is a protein called General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	A	575	Total	C	N	O	S	0	0
			4648	2977	806	837	28		

- Molecule 4 is a protein called General transcription factor IIH subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	F	62	Total	C	N	O	S	0	0
			490	315	78	94	3		

- Molecule 5 is a protein called TFIIH basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	B	675	Total	C	N	O	S	0	0
			5436	3490	941	976	29		

- Molecule 6 is a protein called General transcription factor IIH subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	E	237	Total	C	N	O	S	0	0
			1857	1196	309	333	19		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	D	280	Total	C	N	O	S	0	0
			2182	1382	376	400	24		

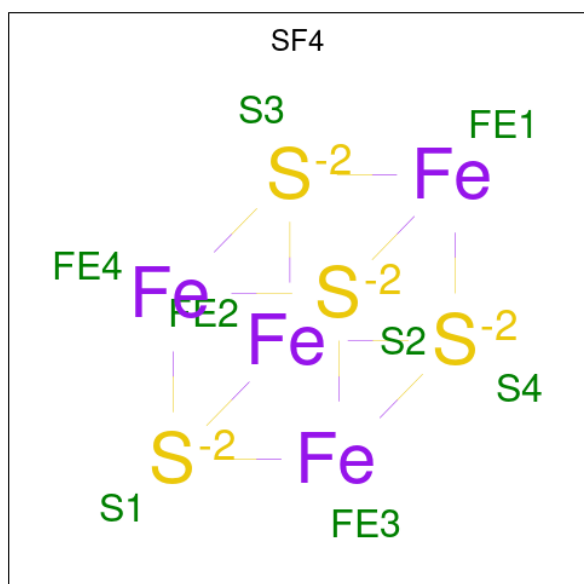
- Molecule 8 is a protein called General transcription factor IIH subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	C	386	Total	C	N	O	S	0	0
			3109	2016	538	543	12		

- Molecule 9 is a protein called DNA repair protein complementing XP-A cells.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	G	134	Total	C	N	O	S	0	0
			1129	710	204	206	9		

- Molecule 10 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe₄S₄).



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

- Molecule 11 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
11	E	2	Total 2	Zn 2	0
11	D	3	Total 3	Zn 3	0
11	G	1	Total 1	Zn 1	0

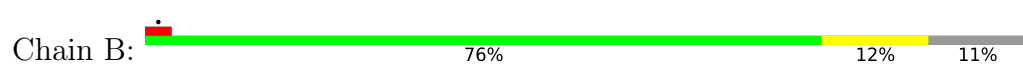
LEU
PHE
LYS
ARG
PHE
ARG
LYS

• Molecule 4: General transcription factor IIH subunit 5



MET VAL ASN V4 L5 K6 G7 T10 E11 C12 D13 T46 D24 E25 L29 G30 K31 K32 F33 I34 I35 Q36 D37 D40 E48 R56 V57 A65 PHE SER LEU THR GLN LYS

• Molecule 5: TFIIH basal transcription factor complex helicase XPD subunit



M1 K2 L3 I17 Y18 P19 E20 Q21 L28 G46 T46 G47 L52 L55 I56 Q60 L70 I71 T76 V77 P78 E85 G99 E100 K101 L109 S110 S111 K113 N114 L115 V121 R125 V130 D131 G132 K133 Q145 D149 L152 D162

E167 N176 L177 D178 P191 R196 N203 Y207 Y211 D219 L220 V221 S222 K223 E224 L225 A226 V232 H237 C243 L251 T255 L256 Q267 R272 ILE LYS GLU THR ASP GLN ARG LEU ARG ASP GLU TYR ARG ARG LEU VAL GLU GLY LEU

ARG GLU ALA SER ALA ALA ARG GLU THR ASP ALA HIS LEU ALA ASN PRO VAL LEU PRO ASP VAL GLN GLU ALA VAL PRO GLY SER ILE ARG THR A326 R334 Y339 R345 S353 A356 V365 R380 L383 L386 E387 I388 L391 A392 D393 L399

F421 D422 D423 T427 I428 A429 A443 I444 T460 V476 T477 T484 L485 V488 C489 R497 D500 S505 T510 R511 E512 D513 V516 I517 P532 S541 T548 L581 E606 G607 I608 R616 V623 E648 T653 M657 R666

L676 D681 K682 R683 F684 A685 K689 P694 N707 F728 HIS ARG GLU ASP GLN LEU SER LEU SER LEU GLN LEU SER GLU THR LYS ARG ILE GLN ILE ALA GLN LEU

• Molecule 6: General transcription factor IIH subunit 3



MET VAL SER ASP GLU L8 L11 V12 I13 Q25 K28 E29 S30 Q31 F32 L34 A40 V41 L44 G45 N46 F50 Q65 G73 LYS ASN GLY ARG LEU GLY ASP PHE GLY ASP PRO GLY ASN PRO PRO PHE GLU ASN PRO PRO PHE SER GLY LYS ASP

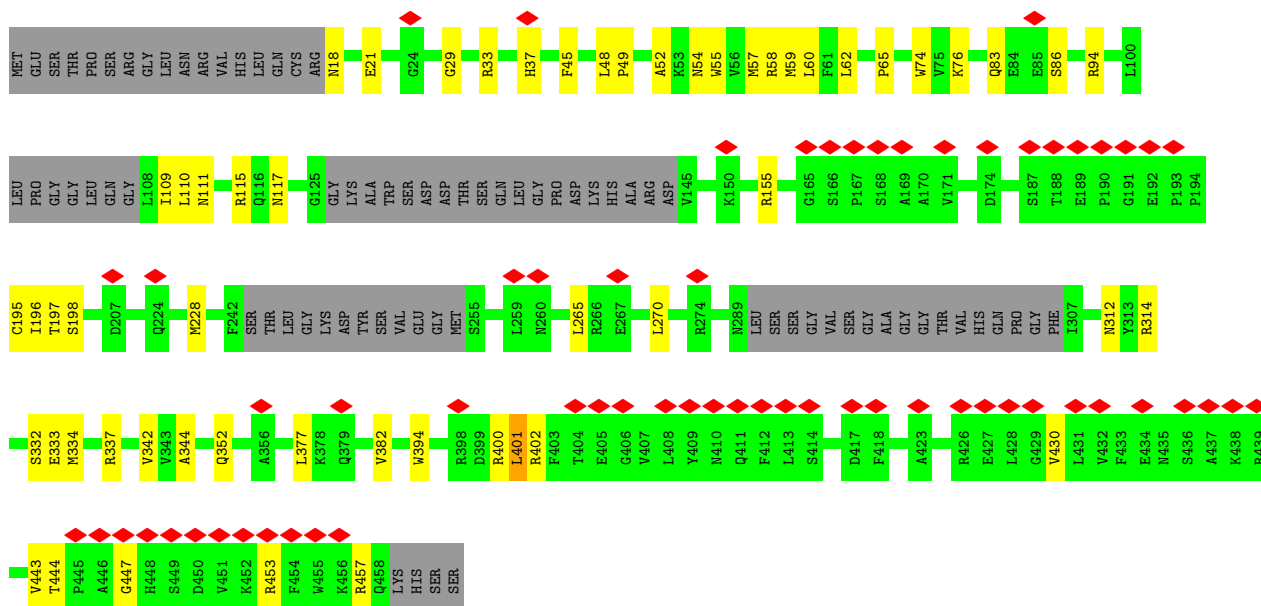
GLY LYS TYR GLU L103 L104 E113 M119 T120 L120 LYS SER ASP ILE LYS GLN H128 K149 E150 V151 LYS ASP ASN GLU M157 M175 N176 L213 Y214 L215 S222 L223 L224 Q225 Y226 L227 L228 W229 V230 F231 L232 R238 L241 R252 C255 F256 L261

C268 S269 V270 C271 L272 S273 I274 N277 F278 S279 P280 I281 C282 T283 T284 C285 G285 THR ALA PHE LYS ILE SER LEU PRO GLN VAL LEU LYS ALA LYS LYS LYS LYS LYS LYS VAL SER ALA

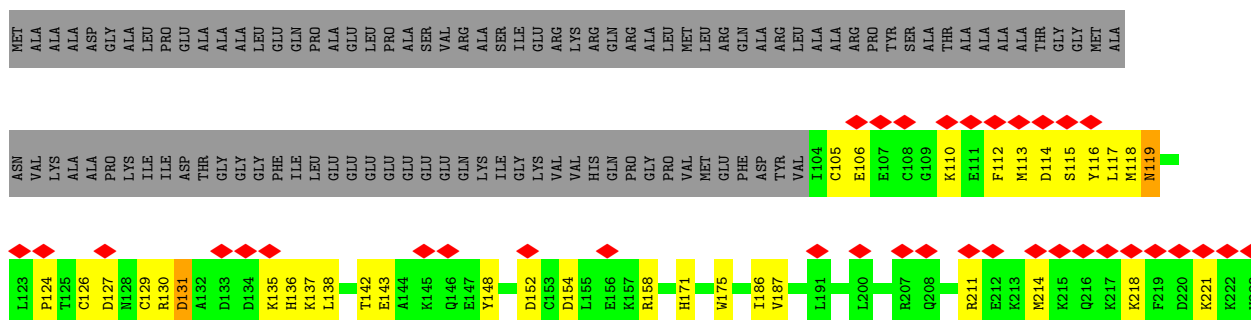
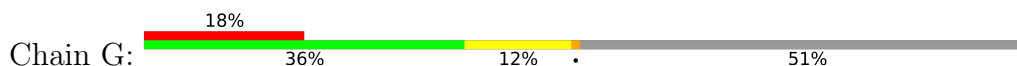
• Molecule 7: General transcription factor IIH subunit 2



- Molecule 8: General transcription factor IIH subunit 4



- Molecule 9: DNA repair protein complementing XP-A cells



K224	E225	L226	R227	R228	A229	V230	R231	S232	V233	V234	W235	K236	R237	GLU	THR	ILE	VAL	HIS	GLN	HIS	HIS	GLU	TYR	GLY	PRO	GLU	GLU	ASN	LEU	GLU	ASP	ASP	MET	TYR	ARG	LYS	THR	CYS	THR	MET	CYS	GLY	HIS	GLU	LEU	THR	TYR	GLU	LYS	MET
------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	227776	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	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.150	Depositor
Minimum map value	-0.050	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	336.0, 336.0, 336.0	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	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: SF4, ZN

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z > 5$	RMSZ	# $ Z > 5$
1	J	0.44	0/433	0.51	0/665
2	K	0.47	0/646	0.53	0/994
3	A	0.20	0/4748	0.50	0/6416
4	F	0.17	0/495	0.48	0/668
5	B	0.24	0/5556	0.50	0/7522
6	E	0.33	0/1890	0.56	0/2561
7	D	0.42	0/2228	0.69	2/3017 (0.1%)
8	C	0.24	0/3178	0.56	0/4307
9	G	0.49	0/1149	0.81	1/1531 (0.1%)
All	All	0.30	0/20323	0.56	3/27681 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
9	G	119	ASN	N-CA-C	-5.24	105.97	114.09
7	D	201	LEU	CA-C-N	5.14	131.36	121.54
7	D	201	LEU	C-N-CA	5.14	131.36	121.54

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	J	387	0	214	13	0
2	K	575	0	314	21	0
3	A	4648	0	4704	44	0
4	F	490	0	499	10	0
5	B	5436	0	5483	62	0
6	E	1857	0	1892	45	0
7	D	2182	0	2173	51	0
8	C	3109	0	3155	53	0
9	G	1129	0	1145	46	0
10	B	8	0	0	0	0
11	D	3	0	0	0	0
11	E	2	0	0	0	0
11	G	1	0	0	0	0
All	All	19827	0	19579	297	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 8.

The worst 5 of 297 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)
7:D:80:ARG:HH21	7:D:201:LEU:HB3	1.03	1.17
7:D:80:ARG:NH2	7:D:201:LEU:HB3	1.61	1.15
7:D:80:ARG:HH21	7:D:201:LEU:CB	1.69	1.06
9:G:148:TYR:HE2	9:G:187:VAL:HG12	1.29	0.98
9:G:115:SER:HB2	9:G:127:ASP:HB3	1.48	0.96

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	
3	A	569/782 (73%)	518 (91%)	51 (9%)	0	100	100
4	F	60/71 (84%)	59 (98%)	1 (2%)	0	100	100
5	B	671/760 (88%)	619 (92%)	52 (8%)	0	100	100
6	E	229/308 (74%)	217 (95%)	11 (5%)	1 (0%)	30	62
7	D	272/395 (69%)	255 (94%)	16 (6%)	1 (0%)	30	62
8	C	376/462 (81%)	337 (90%)	37 (10%)	2 (0%)	24	57
9	G	132/273 (48%)	120 (91%)	11 (8%)	1 (1%)	16	49
All	All	2309/3051 (76%)	2125 (92%)	179 (8%)	5 (0%)	44	74

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
7	D	78	PRO
8	C	401	LEU
9	G	131	ASP
6	E	279	SER
8	C	49	PRO

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	
3	A	511/688 (74%)	509 (100%)	2 (0%)	84	81
4	F	55/64 (86%)	54 (98%)	1 (2%)	51	69
5	B	589/664 (89%)	587 (100%)	2 (0%)	86	83
6	E	211/272 (78%)	208 (99%)	3 (1%)	59	71
7	D	251/352 (71%)	250 (100%)	1 (0%)	84	81
8	C	334/399 (84%)	334 (100%)	0	100	100
9	G	126/233 (54%)	126 (100%)	0	100	100
All	All	2077/2672 (78%)	2068 (100%)	9 (0%)	81	81

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	E	284	THR
7	D	78	PRO
5	B	444	ILE
5	B	488	VAL
6	E	11	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
8	C	111	ASN
8	C	410	ASN
8	C	116	GLN
8	C	312	ASN
8	C	458	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 6 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
10	SF4	B	1000	-	0,12,12	-	-	-		

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
10	SF4	B	1000	-	-	-	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.

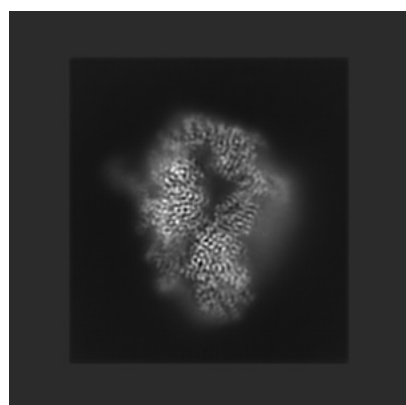
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-4970. These allow visual inspection of the internal detail of the map and identification of artifacts.

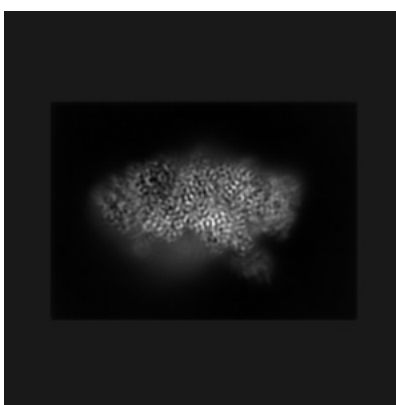
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

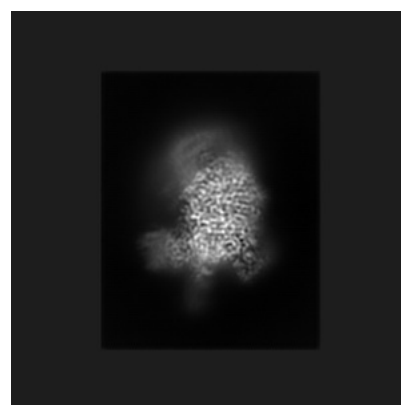
6.1.1 Primary map



X



Y

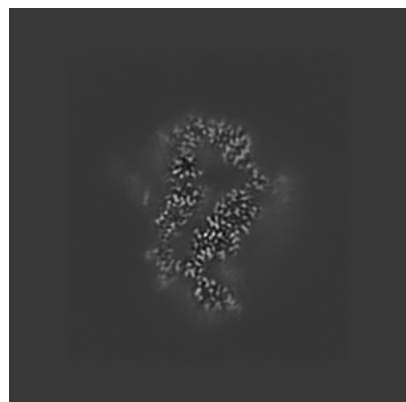


Z

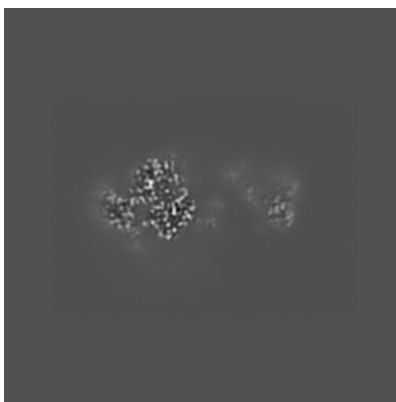
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

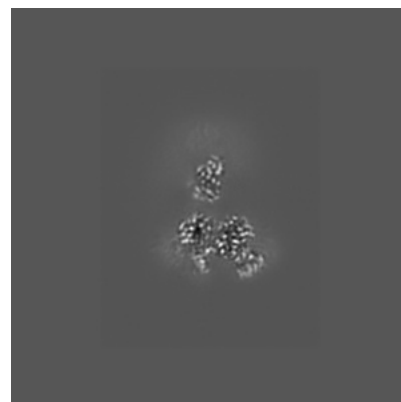
6.2.1 Primary map



X Index: 160



Y Index: 160

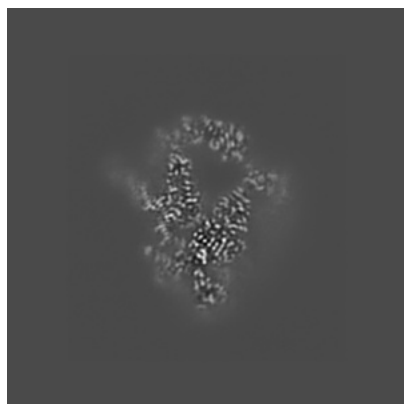


Z Index: 160

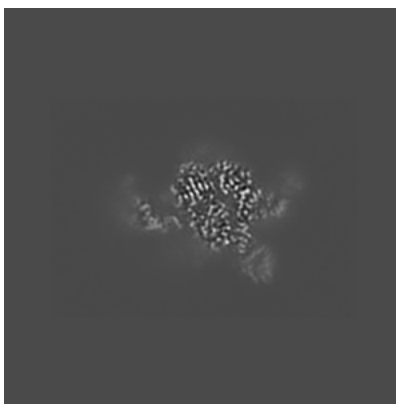
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

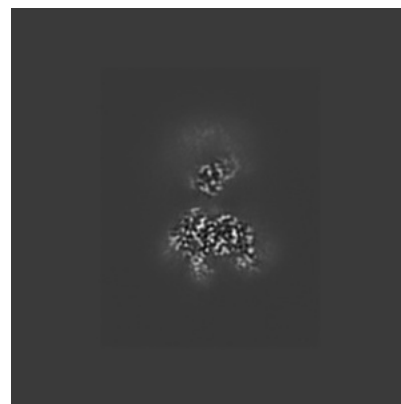
6.3.1 Primary map



X Index: 153



Y Index: 134

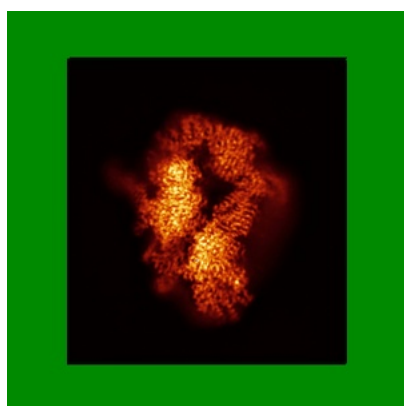


Z Index: 166

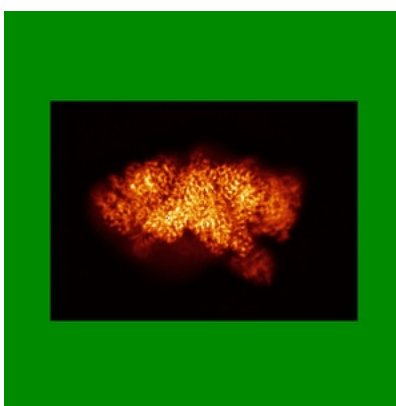
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [i](#)

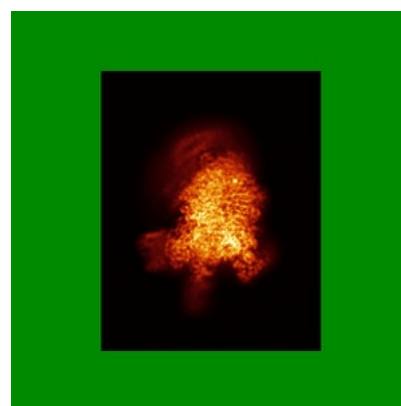
6.4.1 Primary map



X



Y

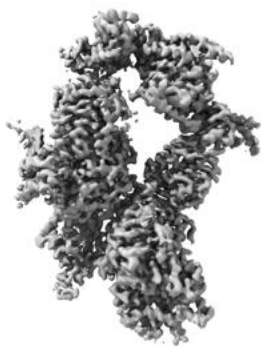


Z

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

6.5 Orthogonal surface views [i](#)

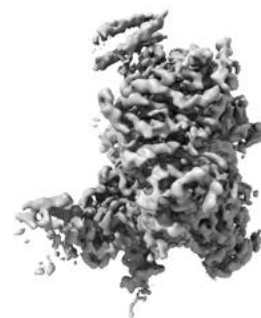
6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

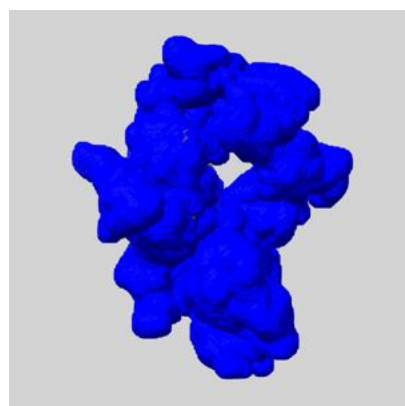
6.6 Mask visualisation [i](#)

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

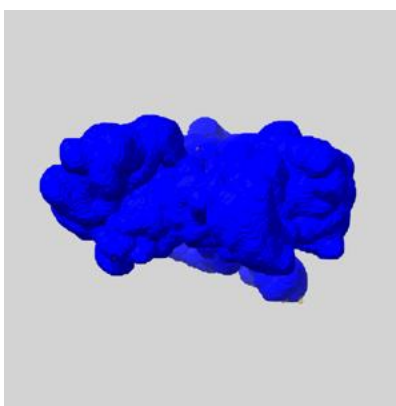
A mask typically either:

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

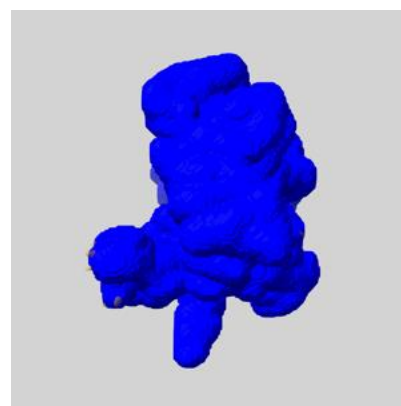
6.6.1 emd_4970_msk_1.map [i](#)



X



Y

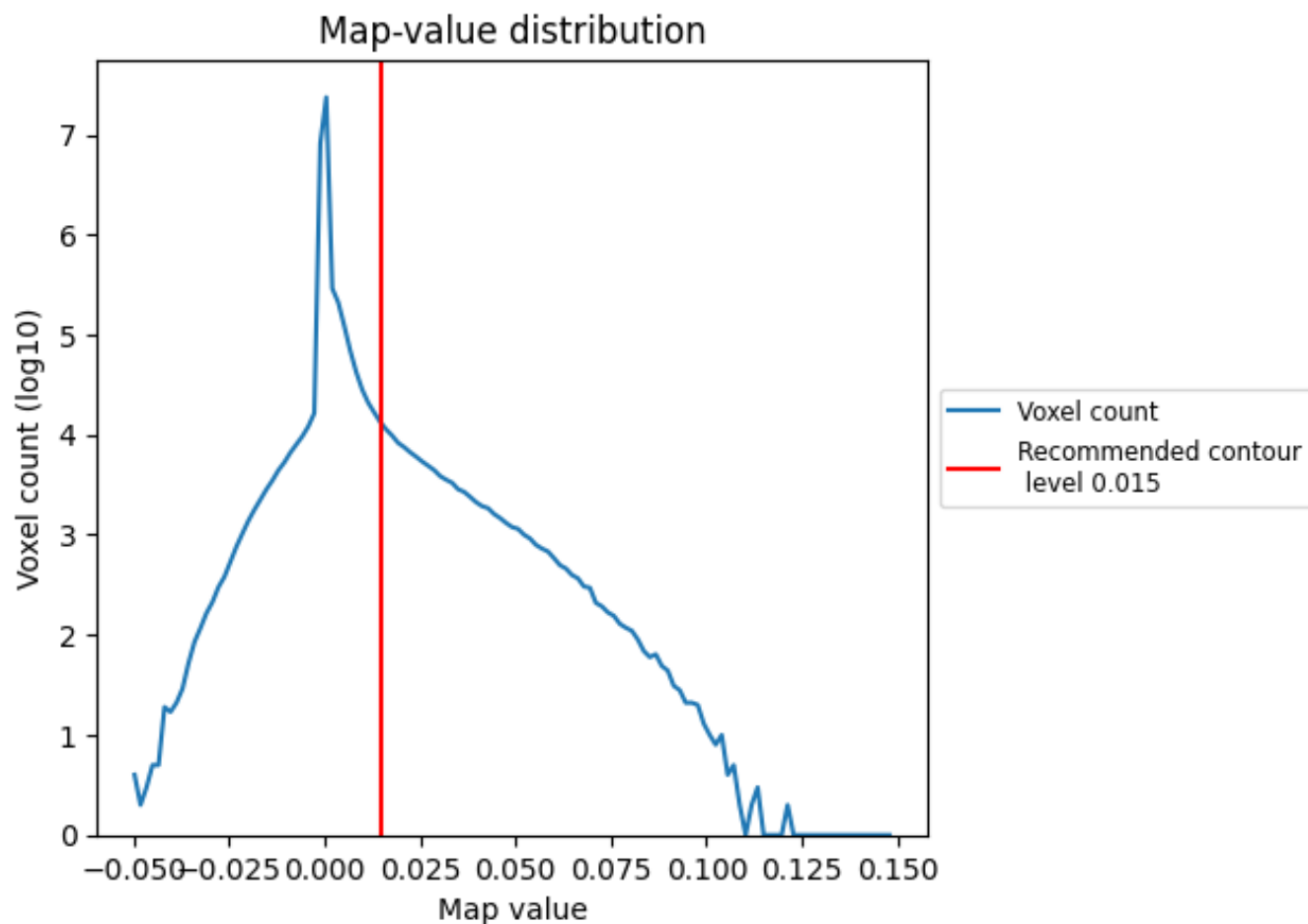


Z

7 Map analysis [i](#)

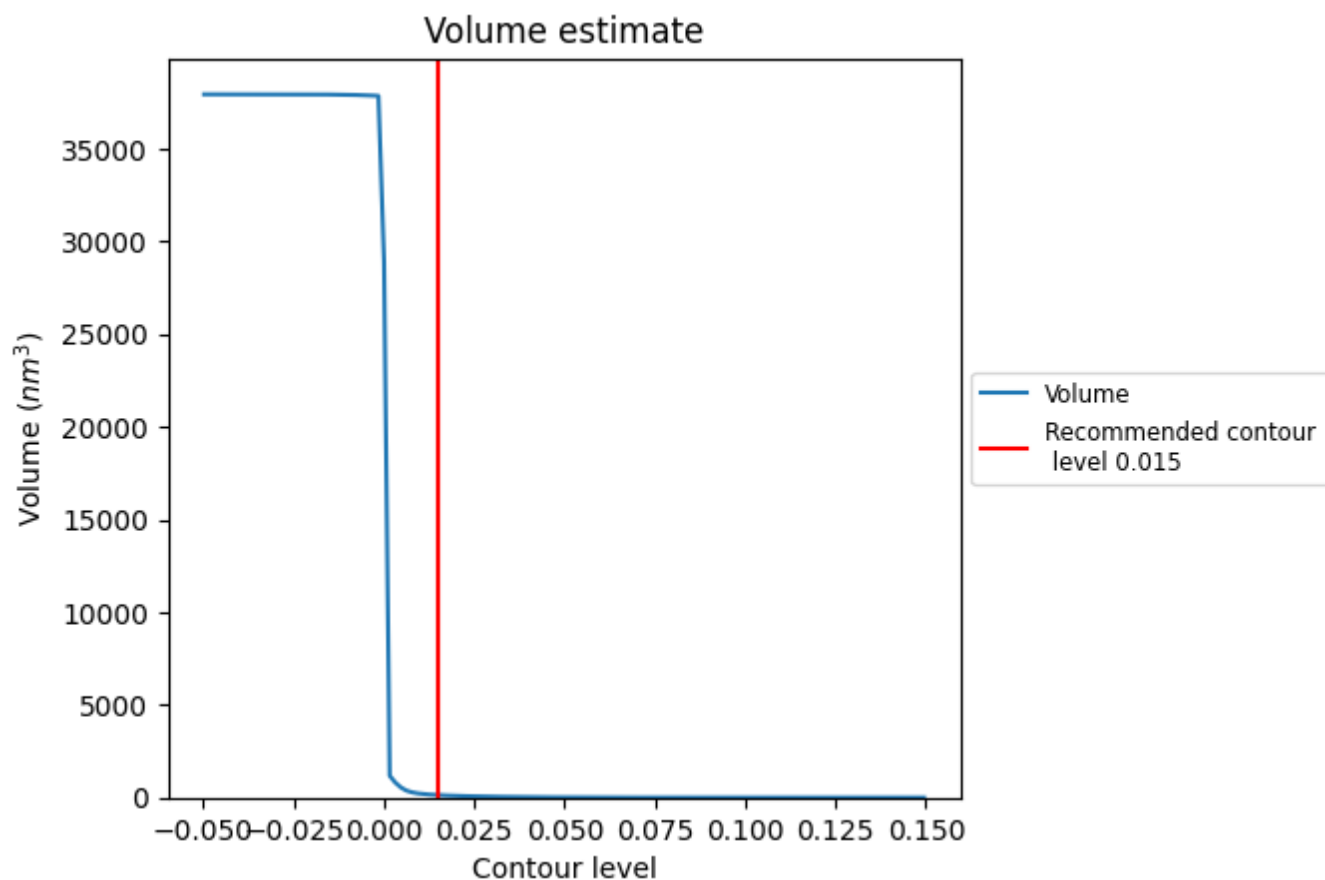
This section contains the results of statistical analysis of the map.

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.

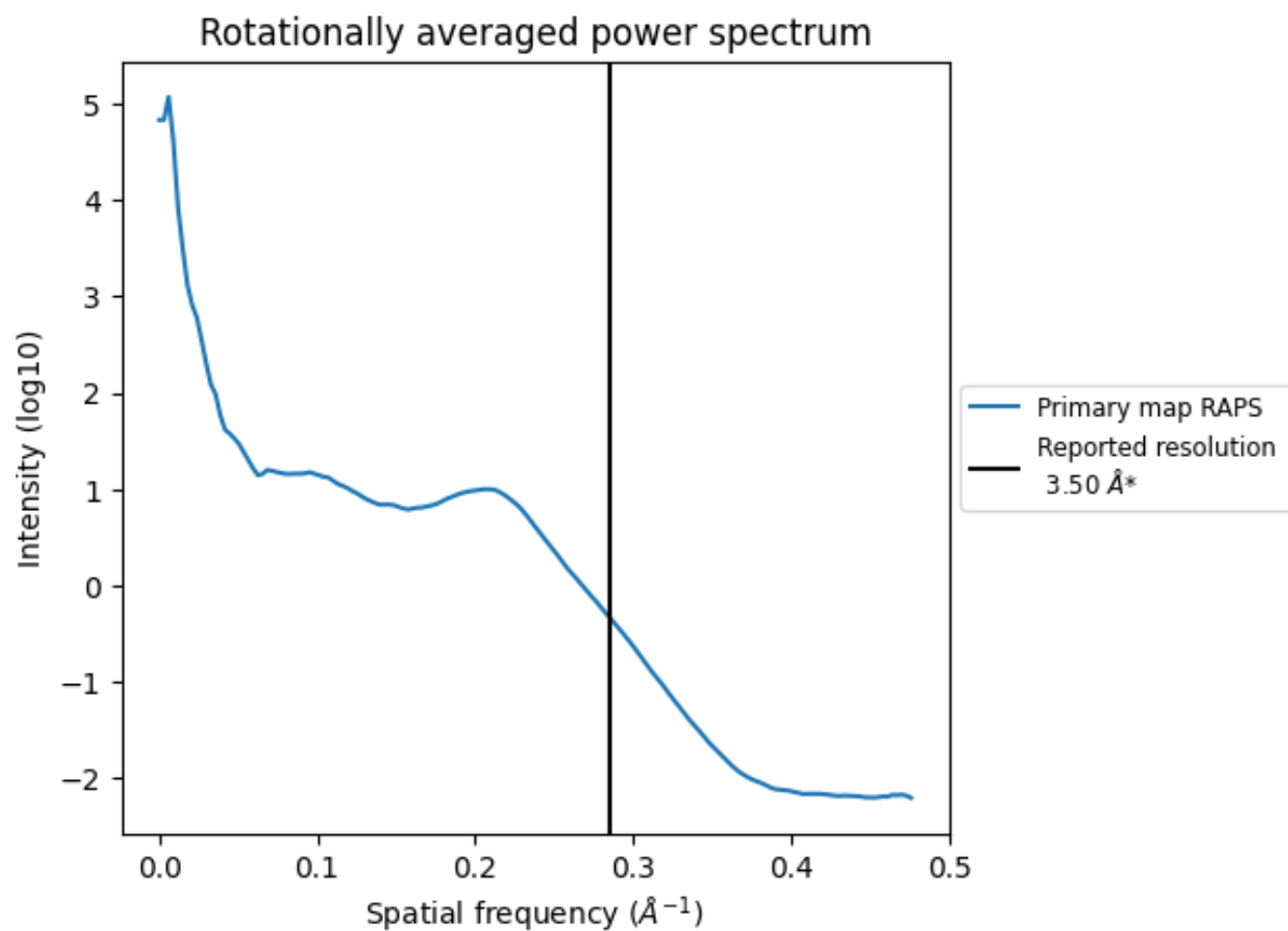
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 132 nm^3 ; this corresponds to an approximate mass of 119 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 0.286 Å⁻¹

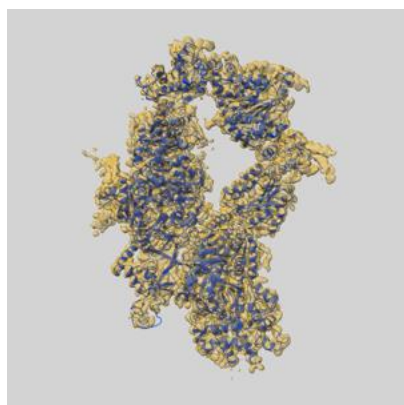
8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

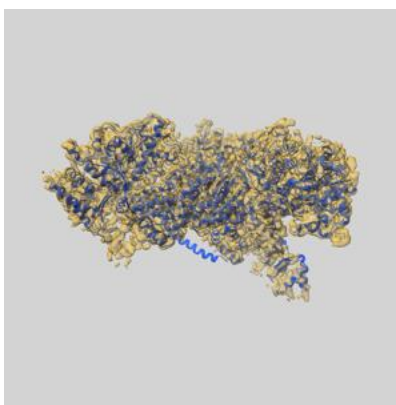
9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-4970 and PDB model 6RO4. Per-residue inclusion information can be found in section [3](#) on page [7](#).

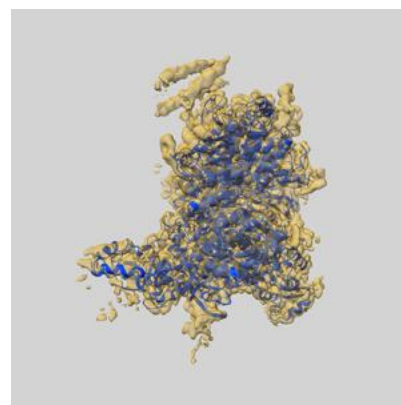
9.1 Map-model overlay [i](#)



X



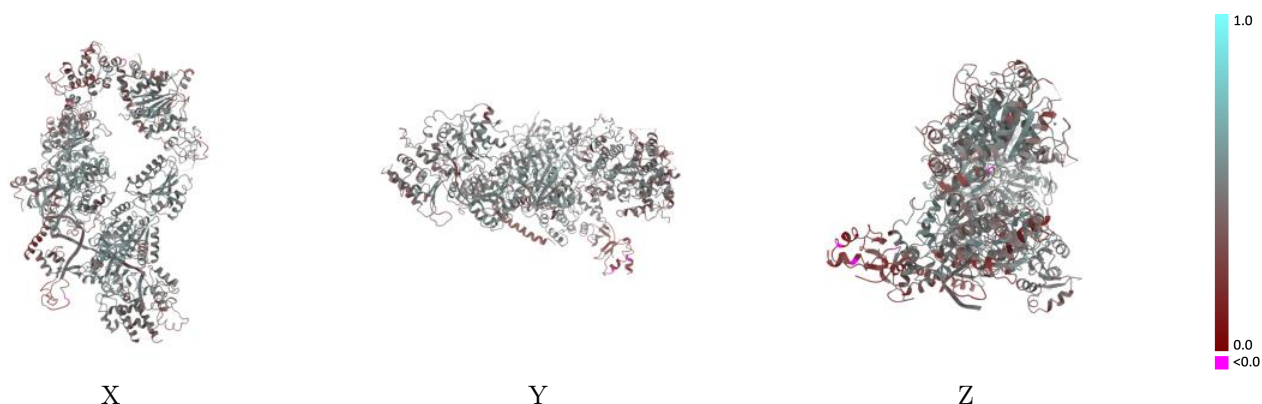
Y



Z

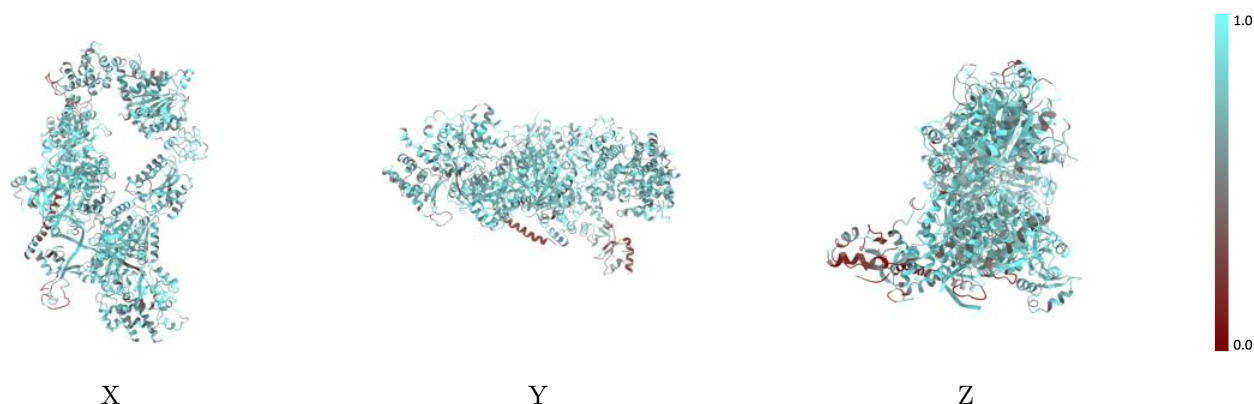
The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [i](#)



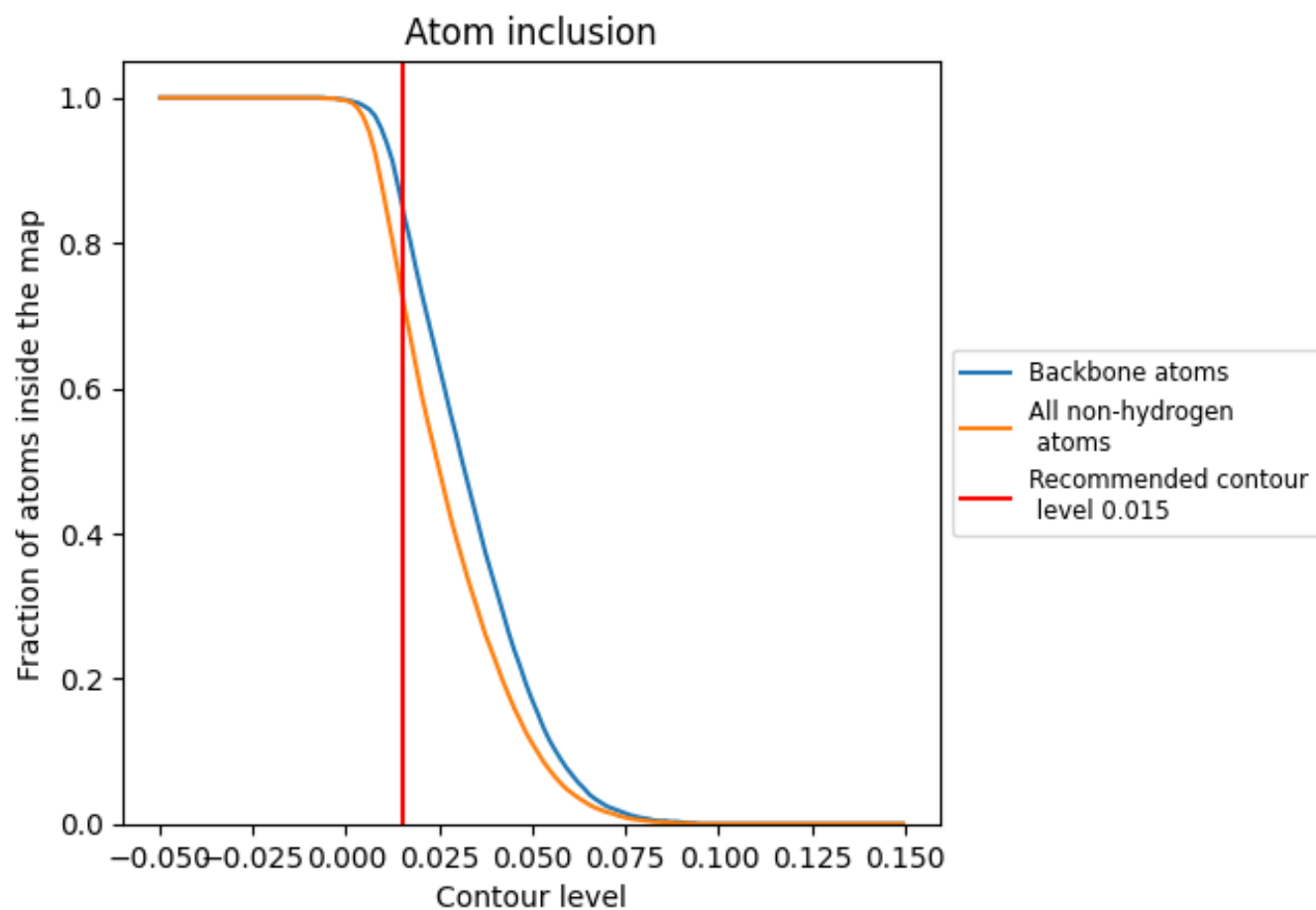
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.015).

9.4 Atom inclusion [i](#)



At the recommended contour level, 85% of all backbone atoms, 73% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7310	0.4490
A	0.7860	0.4860
B	0.7770	0.4820
C	0.6510	0.3850
D	0.7290	0.4630
E	0.7340	0.4620
F	0.6450	0.3640
G	0.4570	0.3040
J	0.8730	0.4430
K	0.8000	0.4590

1.0

0.0

<0.0