Full wwPDB EM Validation Report ⓘ

Aug 23, 2022 – 07:11 pm BST

PDB ID	:	7PHR
EMDB ID	:	EMD-13427
Title	:	Structure of a fully assembled T-cell receptor engaging a tumor-associated peptide-MHC I
Authors	:	Susac, L.; Thomas, C.; Tampe, R.
Deposited on	:	2021-08-18
Resolution	:	3.08 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the ⓘ 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:

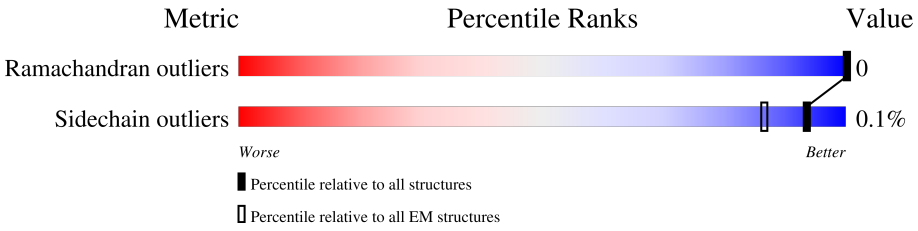
EMDB validation analysis	:	0.0.1.dev8
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.30

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 3.08 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.



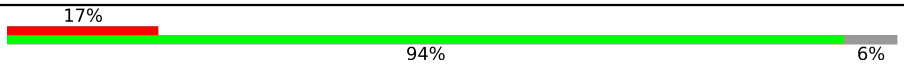
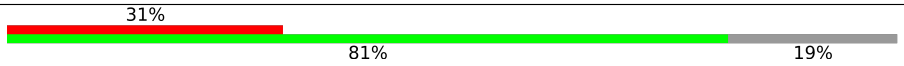
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	251	97%
2	B	290	98%
3	C	122	12% 93% 7%
4	D	356	30% 70%
5	E	136	13% 90% 10%
5	e	136	10% 91% 9%
6	H	304	10% 90% 10%
7	L	101	7% 98%
8	P	9	89% 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	Z	36	
9	z	36	

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 11511 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 T-cell receptor alpha chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	243	Total	C	N	O	S	0	0
			1898	1181	328	379	10		

- Molecule 2 is a protein called T-cell receptor beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	285	Total	C	N	O	S	0	0
			2252	1436	380	428	8		

- Molecule 3 is a protein called T-cell surface glycoprotein CD3 gamma chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	114	Total	C	N	O	S	0	0
			872	562	145	158	7		

- Molecule 4 is a protein called T-cell surface glycoprotein CD3 delta chain, green fluorescent protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	108	Total	C	N	O	S	0	0
			826	523	139	158	6		

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	112	ASP	-	linker	UNP P04234
D	113	PRO	-	linker	UNP P04234
D	114	PRO	-	linker	UNP P04234
D	115	VAL	-	linker	UNP P04234
D	116	ALA	-	linker	UNP P04234
D	117	THR	-	linker	UNP P04234
D	119	VAL	-	insertion	UNP P42212
D	182	LEU	PHE	engineered mutation	UNP P42212

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	349	LEU	HIS	conflict	UNP P42212

- Molecule 5 is a protein called T-cell surface glycoprotein CD3 epsilon chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	123	Total	C	N	O	S	0	0
			950	604	151	187	8		
5	e	124	Total	C	N	O	S	0	0
			979	622	158	191	8		

- Molecule 6 is a protein called HLA class I histocompatibility antigen, A alpha chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	275	Total	C	N	O	S	0	0
			2246	1403	409	425	9		

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	-11	MET	-	initiating methionine	UNP P04439
H	-10	GLY	-	expression tag	UNP P04439
H	-9	SER	-	expression tag	UNP P04439
H	-8	SER	-	expression tag	UNP P04439
H	-7	HIS	-	expression tag	UNP P04439
H	-6	HIS	-	expression tag	UNP P04439
H	-5	HIS	-	expression tag	UNP P04439
H	-4	HIS	-	expression tag	UNP P04439
H	-3	HIS	-	expression tag	UNP P04439
H	-2	HIS	-	expression tag	UNP P04439
H	-1	GLY	-	expression tag	UNP P04439
H	0	SER	-	expression tag	UNP P04439
H	62	GLY	GLN	variant	UNP P04439
H	66	LYS	ASN	variant	UNP P04439
H	70	HIS	GLN	variant	UNP P04439
H	74	HIS	ASP	variant	UNP P04439
H	95	VAL	ILE	variant	UNP P04439
H	97	ARG	ILE	variant	UNP P04439
H	107	TRP	GLY	variant	UNP P04439
H	114	HIS	ARG	variant	UNP P04439
H	116	TYR	ASP	variant	UNP P04439
H	127	LYS	ASN	variant	UNP P04439
H	142	THR	ILE	variant	UNP P04439

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	145	HIS	ARG	variant	UNP P04439
H	152	VAL	GLU	variant	UNP P04439
H	161	GLU	ASP	variant	UNP P04439
H	184	ALA	PRO	variant	UNP P04439
H	193	ALA	PRO	variant	UNP P04439
H	194	VAL	ILE	variant	UNP P04439
H	207	SER	GLY	variant	UNP P04439
H	253	GLN	GLU	variant	UNP P04439
H	276	PRO	LEU	variant	UNP P04439
H	281	GLU	-	expression tag	UNP P04439
H	282	ASP	-	expression tag	UNP P04439
H	283	GLN	-	expression tag	UNP P04439
H	284	VAL	-	expression tag	UNP P04439
H	285	ASP	-	expression tag	UNP P04439
H	286	PRO	-	expression tag	UNP P04439
H	287	ARG	-	expression tag	UNP P04439
H	288	LEU	-	expression tag	UNP P04439
H	289	ILE	-	expression tag	UNP P04439
H	290	ASP	-	expression tag	UNP P04439
H	291	GLY	-	expression tag	UNP P04439
H	292	LYS	-	expression tag	UNP P04439

- Molecule 7 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	L	99	Total	C	N	O	S	0	0
			829	528	140	158	3		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	-1	MET	-	initiating methionine	UNP P61769
L	0	GLY	ALA	conflict	UNP P61769

- Molecule 8 is a protein called Tumor-associated antigentic peptide gp100.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	P	9	Total	C	N	O	0	0
			69	46	9	14		

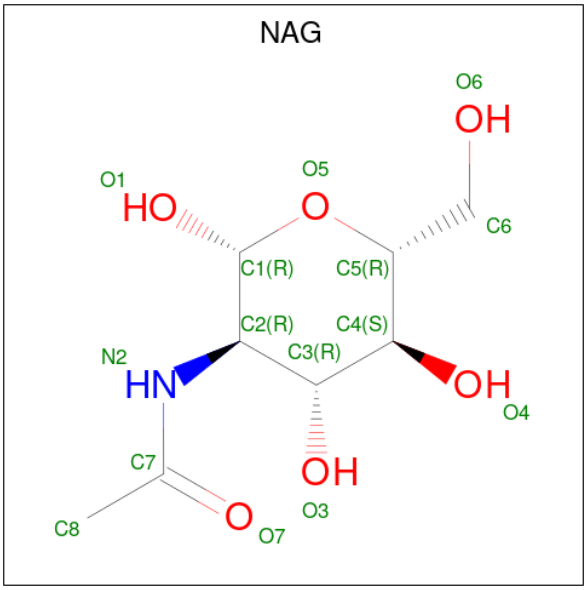
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	9	VAL	ALA	engineered mutation	UNP P40967

- Molecule 9 is a protein called T-cell surface glycoprotein CD3 zeta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	Z	34	Total	C	N	O	S	0	0
			275	191	40	43	1		
9	z	29	Total	C	N	O	S	0	0
			231	160	34	36	1		

- Molecule 10 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



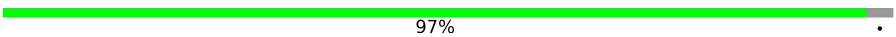
Mol	Chain	Residues	Atoms				AltConf
10	A	1	Total	C	N	O	0
			42	24	3	15	
10	A	1	Total	C	N	O	0
			42	24	3	15	
10	A	1	Total	C	N	O	0
			42	24	3	15	
10	B	1	Total	C	N	O	0
			14	8	1	5	
10	D	1	Total	C	N	O	0
			28	16	2	10	
10	D	1	Total	C	N	O	0
			28	16	2	10	

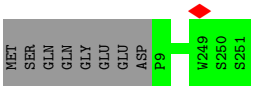
3 Residue-property plots (i)

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: T-cell receptor alpha chain

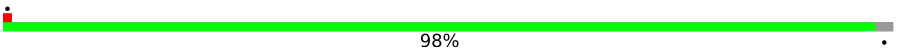
Chain A:

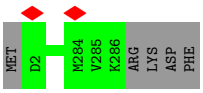




• Molecule 2: T-cell receptor beta chain

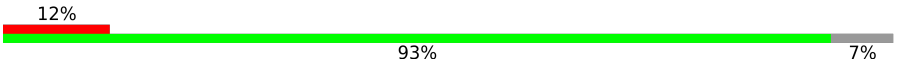
Chain B:

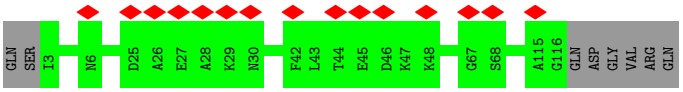




• Molecule 3: T-cell surface glycoprotein CD3 gamma chain

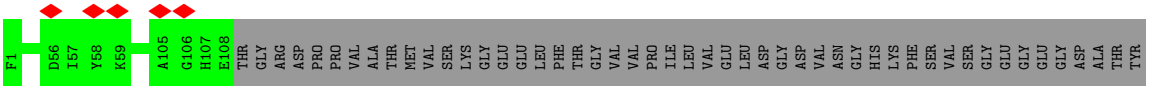
Chain C:





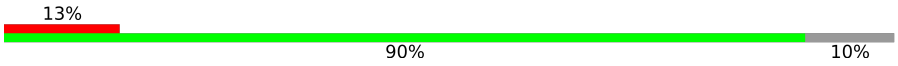
• Molecule 4: T-cell surface glycoprotein CD3 delta chain, green fluorescent protein

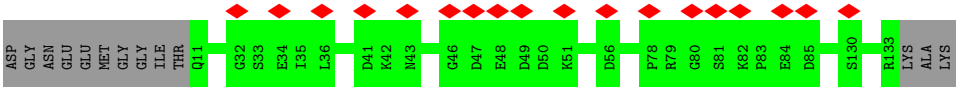
Chain D:



• Molecule 5: T-cell surface glycoprotein CD3 epsilon chain

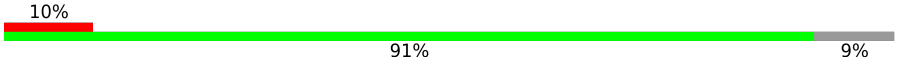
Chain E:

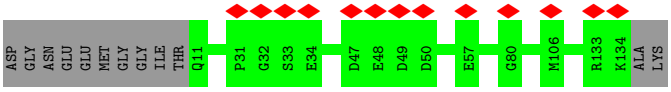




• Molecule 5: T-cell surface glycoprotein CD3 epsilon chain

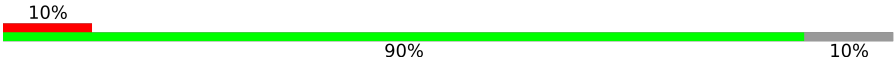
Chain e:

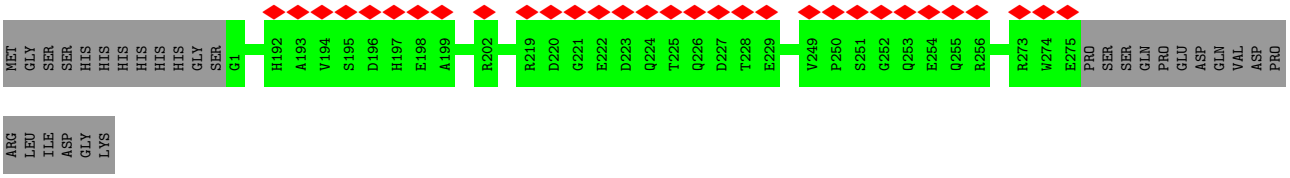




• Molecule 6: HLA class I histocompatibility antigen, A alpha chain

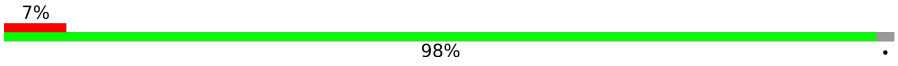
Chain H:

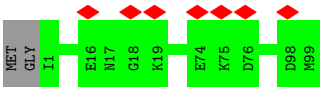




• Molecule 7: Beta-2-microglobulin

Chain L:



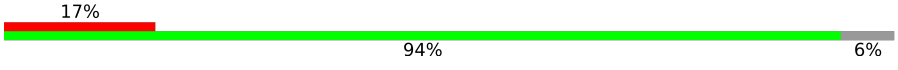


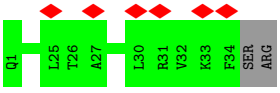
• Molecule 8: Tumor-associated antigentic peptide gp100

Chain P:

• Molecule 9: T-cell surface glycoprotein CD3 zeta chain

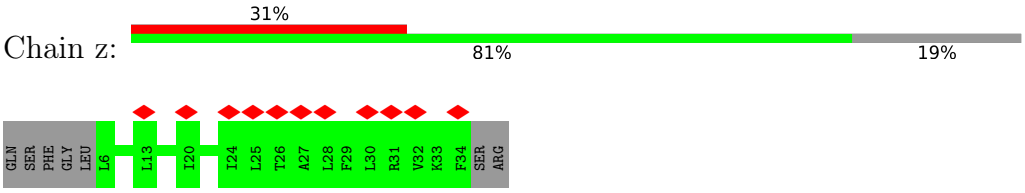
Chain Z:





• Molecule 9: T-cell surface glycoprotein CD3 zeta chain

Chain z:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	154408	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 ($e^-/Å^2$)	61	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	2.782	Depositor
Minimum map value	-1.896	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.035	Depositor
Recommended contour level	0.25	Depositor
Map size (Å)	403.19998, 403.19998, 403.19998	wwPDB
Map dimensions	384, 384, 384	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: NAG

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.35	0/1931	0.52	0/2610
2	B	0.38	0/2314	0.52	0/3153
3	C	0.33	0/890	0.50	0/1202
4	D	0.32	0/838	0.51	0/1141
5	E	0.29	0/971	0.47	0/1323
5	e	0.33	0/1000	0.48	0/1357
6	H	0.31	0/2311	0.53	0/3137
7	L	0.32	0/852	0.49	0/1152
8	P	0.39	0/71	0.57	0/97
9	Z	0.33	0/281	0.46	0/379
9	z	0.29	0/235	0.39	0/318
All	All	0.34	0/11694	0.51	0/15869

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

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

5.3 Torsion angles

5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	241/251 (96%)	239 (99%)	2 (1%)	0	100	100
2	B	283/290 (98%)	276 (98%)	7 (2%)	0	100	100
3	C	112/122 (92%)	104 (93%)	8 (7%)	0	100	100
4	D	106/356 (30%)	100 (94%)	6 (6%)	0	100	100
5	E	121/136 (89%)	117 (97%)	4 (3%)	0	100	100
5	e	122/136 (90%)	119 (98%)	3 (2%)	0	100	100
6	H	273/304 (90%)	269 (98%)	4 (2%)	0	100	100
7	L	97/101 (96%)	95 (98%)	2 (2%)	0	100	100
8	P	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
9	Z	32/36 (89%)	32 (100%)	0	0	100	100
9	z	27/36 (75%)	27 (100%)	0	0	100	100
All	All	1421/1777 (80%)	1384 (97%)	37 (3%)	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	A	220/227 (97%)	220 (100%)	0	100	100
2	B	241/249 (97%)	241 (100%)	0	100	100
3	C	89/103 (86%)	89 (100%)	0	100	100
4	D	92/311 (30%)	92 (100%)	0	100	100
5	E	104/120 (87%)	104 (100%)	0	100	100
5	e	109/120 (91%)	109 (100%)	0	100	100
6	H	231/257 (90%)	231 (100%)	0	100	100
7	L	94/95 (99%)	94 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	P	8/8 (100%)	7 (88%)	1 (12%)	4 17
9	Z	30/32 (94%)	30 (100%)	0	100 100
9	z	25/32 (78%)	25 (100%)	0	100 100
All	All	1243/1554 (80%)	1242 (100%)	1 (0%)	93 97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	P	1	TYR

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

Mol	Chain	Res	Type
2	B	210	GLN
5	e	39	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)

6 ligands are modelled in this entry.

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	NAG	A	301	1	14,14,15	0.20	0	17,19,21	0.54	0
10	NAG	B	301	2	14,14,15	0.21	0	17,19,21	0.54	0
10	NAG	A	303	1	14,14,15	0.19	0	17,19,21	0.43	0
10	NAG	D	402	4	14,14,15	0.25	0	17,19,21	0.46	0
10	NAG	A	302	1	14,14,15	0.21	0	17,19,21	0.48	0
10	NAG	D	401	4	14,14,15	0.23	0	17,19,21	0.41	0

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	NAG	A	301	1	-	3/6/23/26	0/1/1/1
10	NAG	B	301	2	-	4/6/23/26	0/1/1/1
10	NAG	A	303	1	-	2/6/23/26	0/1/1/1
10	NAG	D	402	4	-	1/6/23/26	0/1/1/1
10	NAG	A	302	1	-	2/6/23/26	0/1/1/1
10	NAG	D	401	4	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	D	401	NAG	C4-C5-C6-O6
10	D	401	NAG	O5-C5-C6-O6
10	A	302	NAG	O5-C5-C6-O6
10	A	303	NAG	O5-C5-C6-O6
10	A	302	NAG	C4-C5-C6-O6
10	A	303	NAG	C4-C5-C6-O6
10	A	301	NAG	O5-C5-C6-O6
10	A	301	NAG	C4-C5-C6-O6
10	B	301	NAG	C4-C5-C6-O6
10	B	301	NAG	O5-C5-C6-O6
10	D	402	NAG	O5-C5-C6-O6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
10	A	301	NAG	C3-C2-N2-C7
10	B	301	NAG	C3-C2-N2-C7
10	B	301	NAG	C1-C2-N2-C7

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.

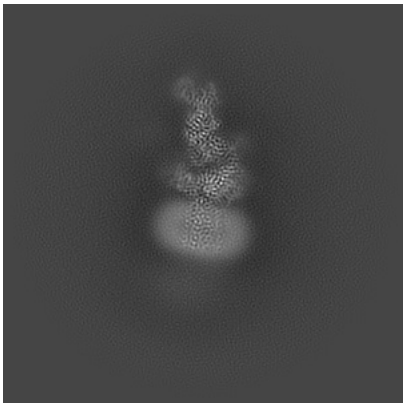
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-13427. These allow visual inspection of the internal detail of the map and identification of artifacts.

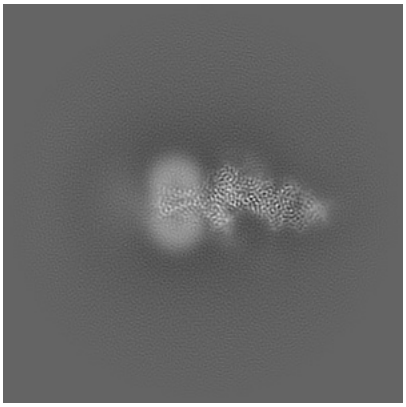
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

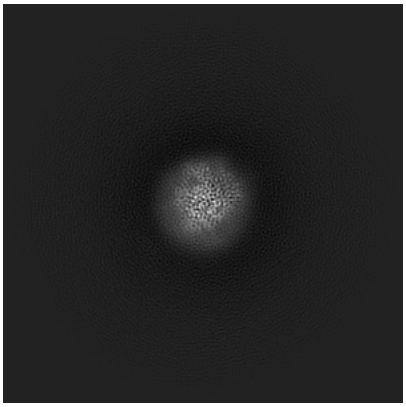
6.1.1 Primary map



X



Y

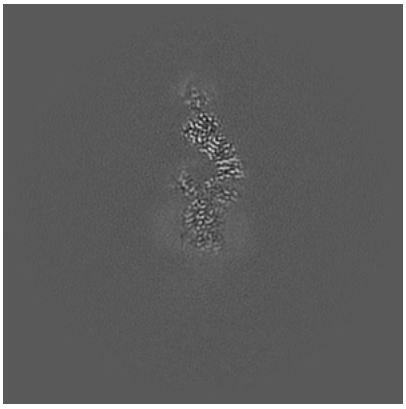


Z

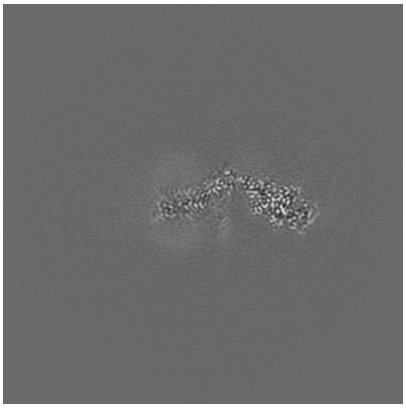
The images above show the map projected in three orthogonal directions.

6.2 Central slices

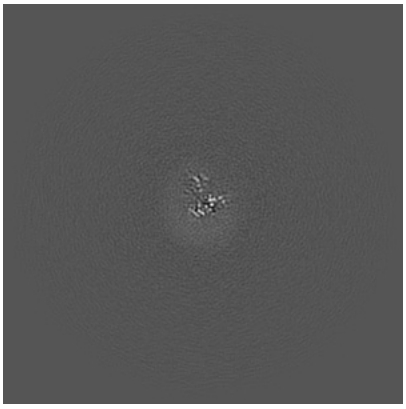
6.2.1 Primary map



X Index: 192



Y Index: 192

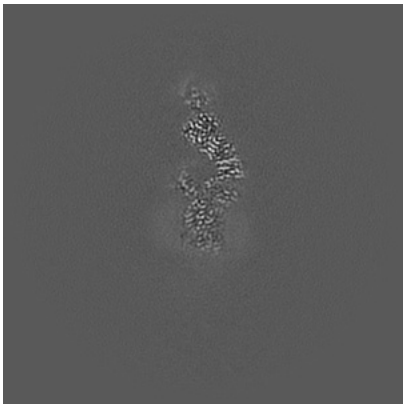


Z Index: 192

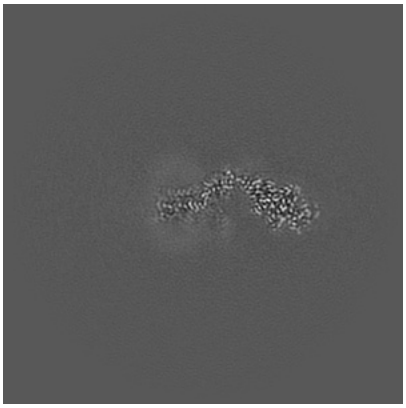
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

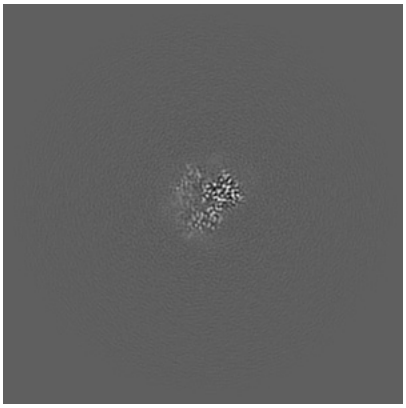
6.3.1 Primary map



X Index: 192



Y Index: 191



Z Index: 212

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

X

Y

Z

The images above show the 3D surface view of the map at the recommended contour level 0.25. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

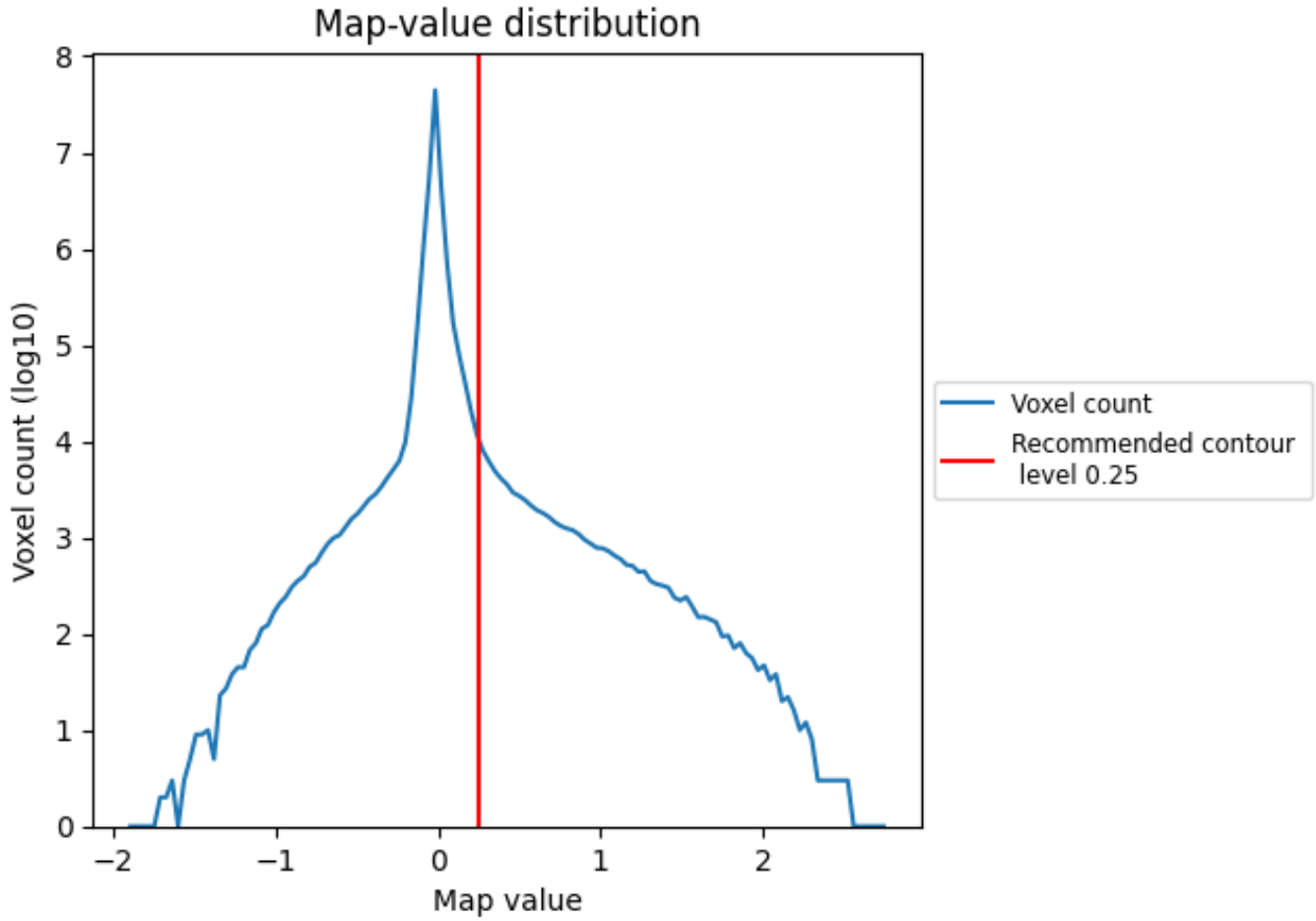
6.5 Mask visualisation (i)

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

7 Map analysis

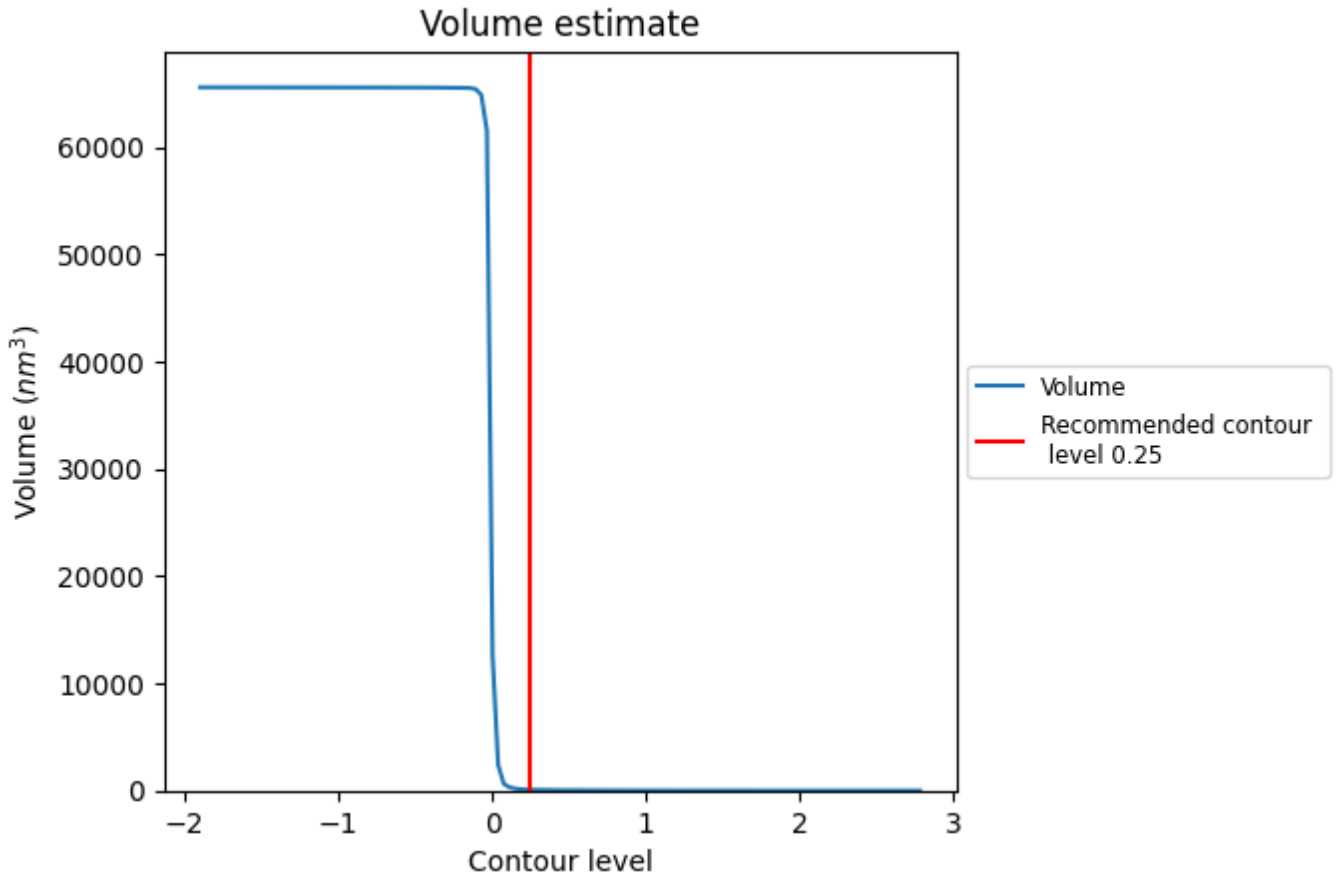
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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

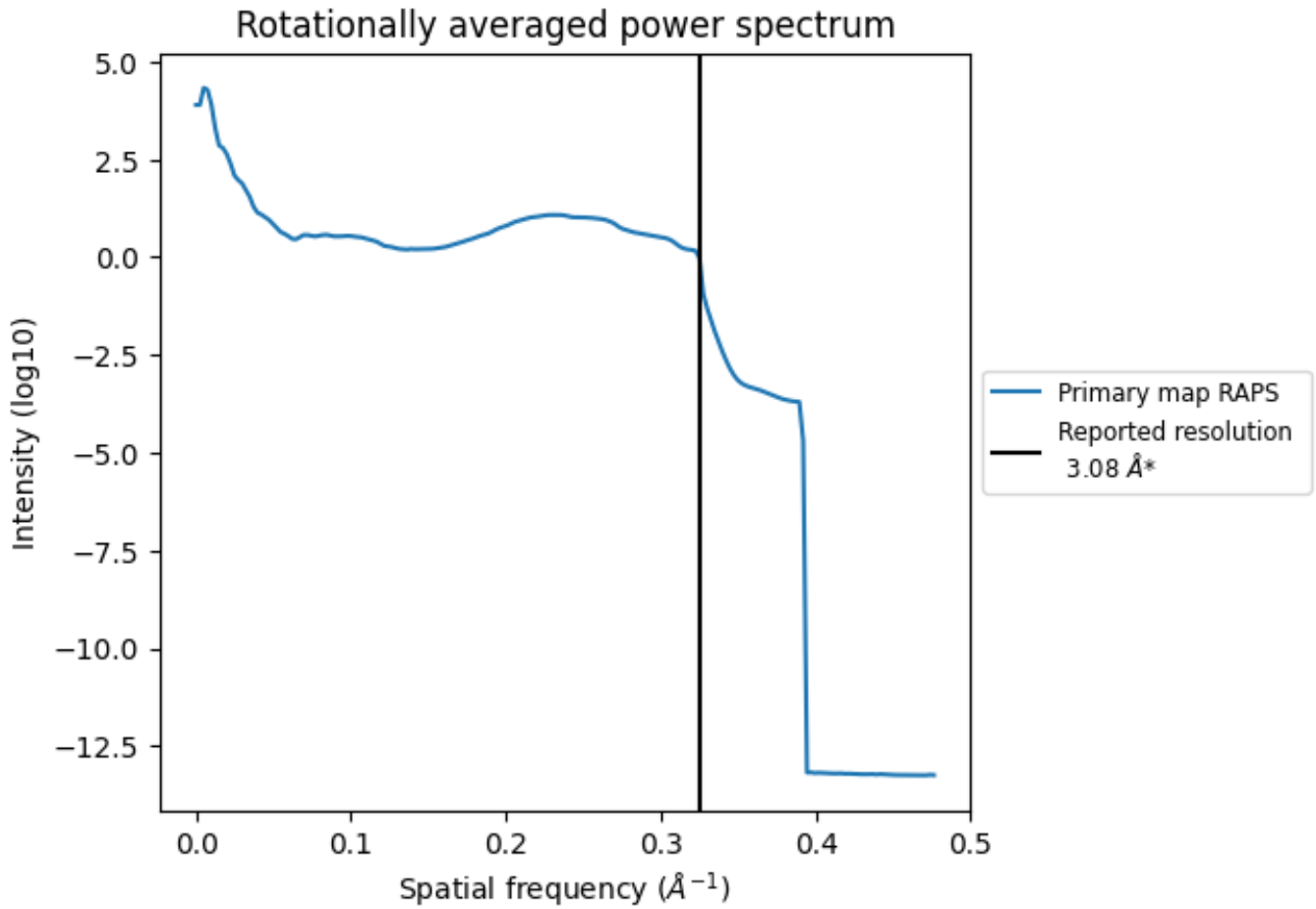
7.2 Volume estimate



The volume at the recommended contour level is 79 nm^3; this corresponds to an approximate mass of 71 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)

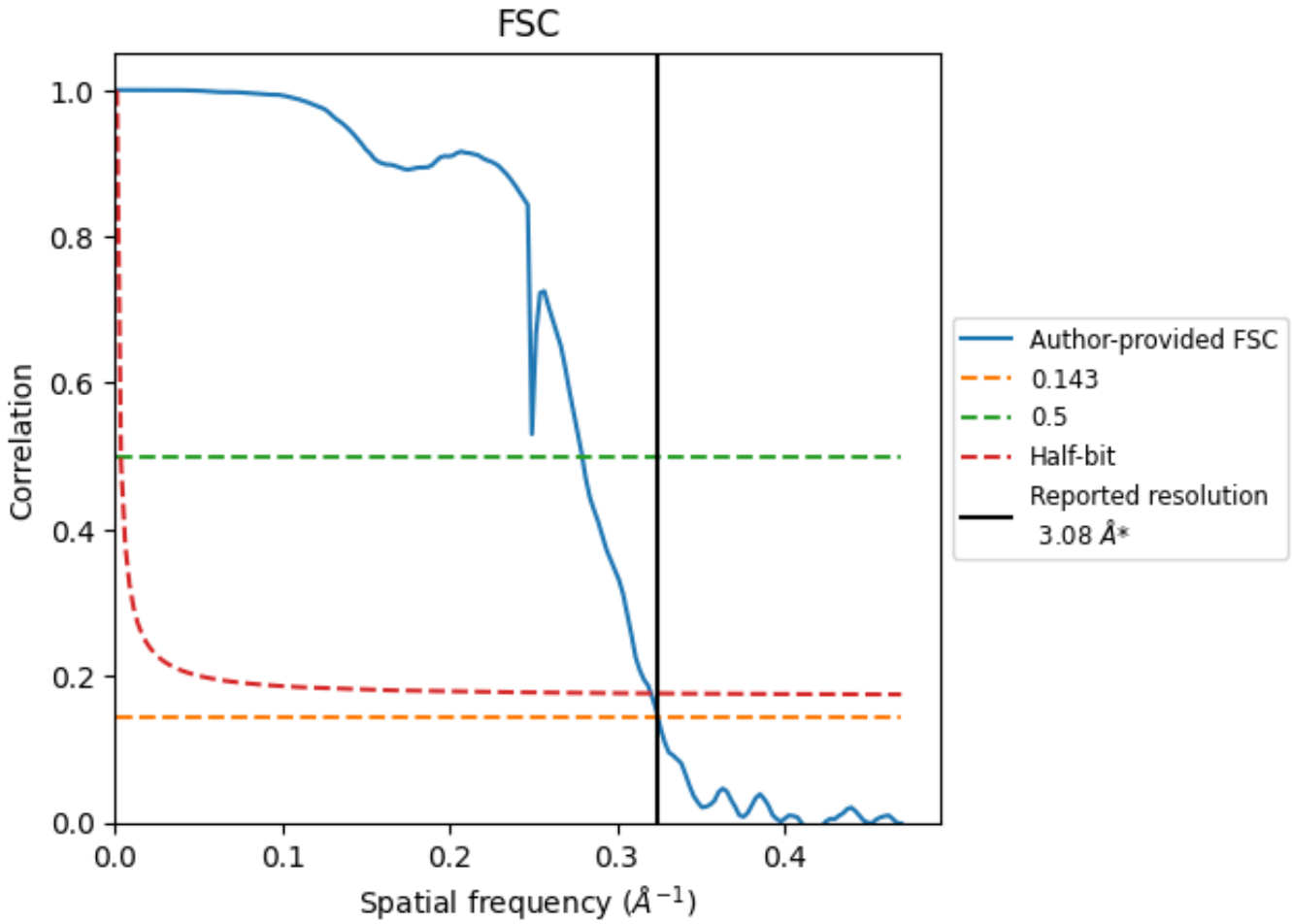


*Reported resolution corresponds to spatial frequency of 0.325 $Å^{-1}$

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.325 Å^{-1}

8.2 Resolution estimates (i)

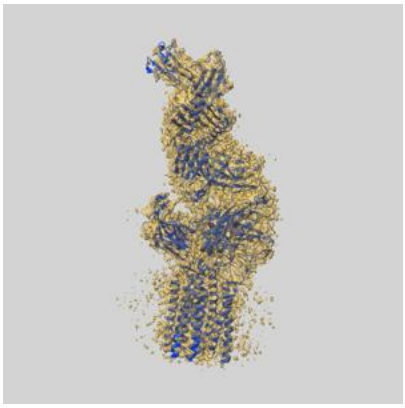
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.08	-	-
Author-provided FSC curve	3.08	3.58	3.12
Unmasked-calculated*	-	-	-

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

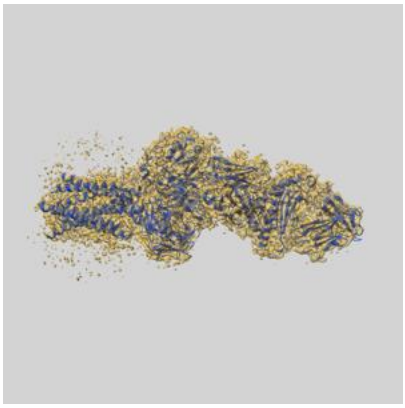
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-13427 and PDB model 7PHR. Per-residue inclusion information can be found in section 3 on page 8.

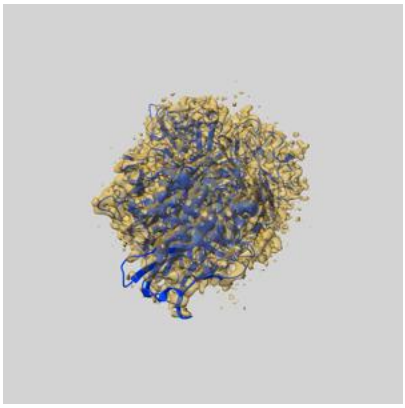
9.1 Map-model overlay



X



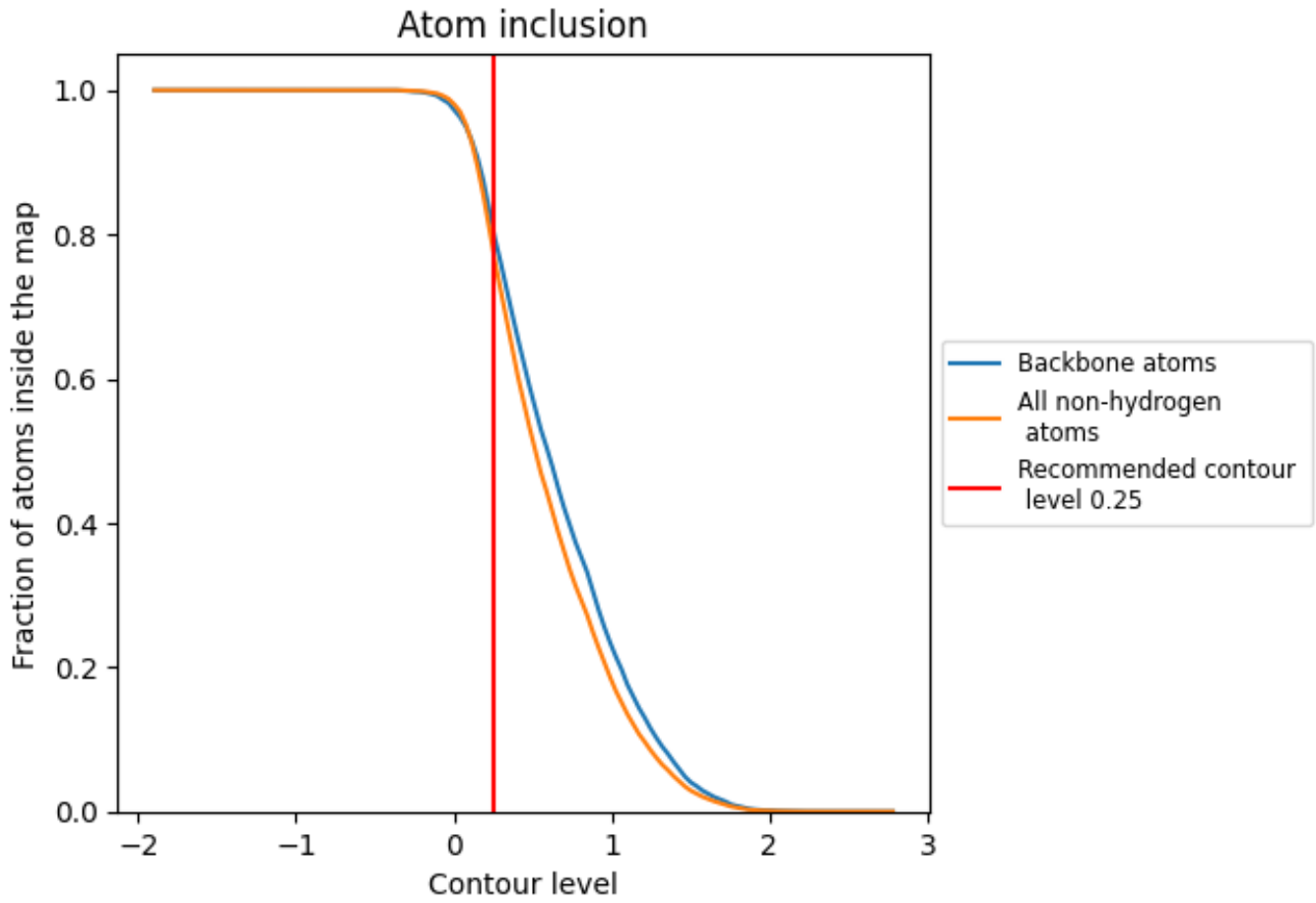
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.25 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, 81% of all backbone atoms, 78% of all non-hydrogen atoms, are inside the map.