



wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 12:22 PM EST

PDB ID : 7UVK
EMDB ID : EMD-26812
Title : G. haemolysans IgA1 protease
Authors : Eisenmesser, E.Z.; Zheng, H.
Deposited on : 2022-05-02
Resolution : 3.28 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.dev43
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

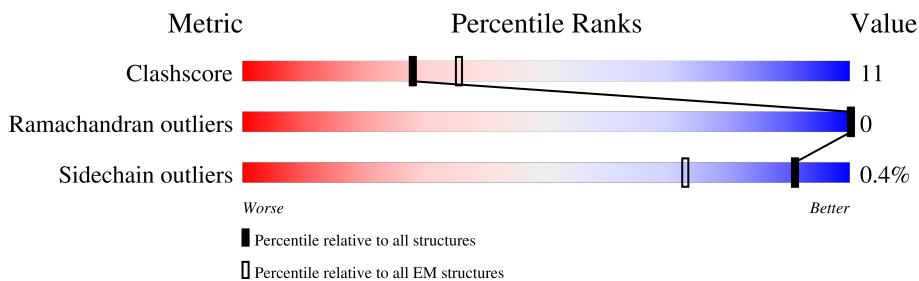
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.28 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	A	2201	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 10350 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.

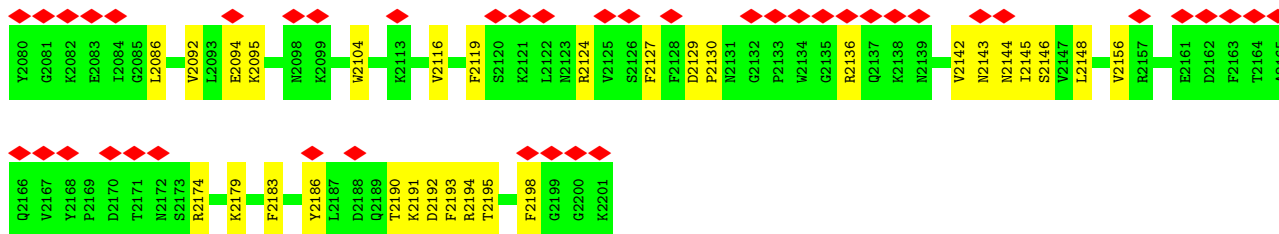
- Molecule 1 is a protein called IgA1 Protease.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1295	10350	6536	1750	2045	19	0	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP C5NYF3
A	2	ARG	-	expression tag	UNP C5NYF3
A	3	LYS	-	expression tag	UNP C5NYF3
A	4	TYR	-	expression tag	UNP C5NYF3
A	5	LEU	-	expression tag	UNP C5NYF3
A	6	GLU	-	expression tag	UNP C5NYF3
A	7	GLU	-	expression tag	UNP C5NYF3
A	8	LYS	-	expression tag	UNP C5NYF3
A	9	TYR	-	expression tag	UNP C5NYF3
A	10	ASN	-	expression tag	UNP C5NYF3
A	11	LYS	-	expression tag	UNP C5NYF3
A	12	PHE	-	expression tag	UNP C5NYF3
A	13	SER	-	expression tag	UNP C5NYF3
A	14	LEU	-	expression tag	UNP C5NYF3
A	15	ARG	-	expression tag	UNP C5NYF3
A	16	LYS	-	expression tag	UNP C5NYF3
A	17	LEU	-	expression tag	UNP C5NYF3
A	18	THR	-	expression tag	UNP C5NYF3
A	19	VAL	-	expression tag	UNP C5NYF3
A	20	GLY	-	expression tag	UNP C5NYF3
A	21	VAL	-	expression tag	UNP C5NYF3
A	22	CYS	-	expression tag	UNP C5NYF3
A	23	SER	-	expression tag	UNP C5NYF3

ALA	S978	G1057	T1159	V1239	N1321	S1401	D1487	M1646	V1778	V1894	K1993	GLN
ARG	R980	E1059	I1162	V1243	N1322	V1402	R1490	K1655	Y1779	L1895	L1994	TYR
HIS	K981	S1060	D1167	D1248	L1325	K1406	D1494	T1666	A1781	D1898	D1899	GLY
LYS	D982	N1063	T1168	R1249	K1326	D1407	D1498	T1664	F1782	D1899	S1900	GLY
ALA	N983	K1064	T1169	E1250	K1327	I1411	N1498	I1665	R1783	E1901	E1905	ALA
GLY	N984	E1065	A1170	T1251	K1328	I1417	I1499	E1667	T1788	R1905	L1906	GLY
MET	Q984	F1066	T1171	S1252	K1329	K1418	I1509	V1688	N1789	Q1907	Q1908	MET
GLY	D985	S1067	Q1172	E1253	L1329	T1419	T1543	S1689	M1792	A1909	D1910	GLY
GLY	V990	G1068	G1173	L1260	L1331	K1420	L1525	Q1672	Q1793	I1916	E1919	GLY
ILE	D912	K1069	I1174	L1269	L1332	I1421	D1528	E1678	Q1799	L1931	L1932	ILE
PHE	D912	L1070	G1175	M1268	V1333	N1422	M1534	S1691	F1807	F1807	A1934	PHE
ASN	N994	I1071	G1176	M1268	P1333	M1431	S1535	K1683	F1808	G1809	E1936	ASN
LYS	K914	G1072	L1177	L1264	K1334	I1439	T1536	Y1694	G1809	K1819	G1940	LYS
SER	E915	E1073	L1181	S1265	K1335	L1441	L1554	T1703	H1820	H1820	V1943	SER
LYS	E916	M1074	I1188	D1266	L1339	Q1442	I1544	L1704	F1829	F1829	A1944	LYS
GLY	K917	K1077	E1189	V1267	E1340	Q1443	I1544	P1705	F1829	V1830	L1946	GLY
GLY	Q918	M1001	K1189	M1268	E1340	D1443	I1547	E1706	F1829	V1830	D1947	GLY
ILE	N919	N1002	K1190	V1269	D1341	N1431	L1554	E1707	F1830	V1831	E1950	ILE
GLY	V919	M1001	S1191	V1269	F1342	N1431	L1554	E1707	F1830	V1831	K1951	GLY
LEU	D920	I1003	I1192	M1273	N1343	I1439	L1554	T1717	F1837	L1837	K1951	LEU
HIS	D920	I1003	I1192	M1273	N1343	I1439	L1554	T1717	F1837	L1837	K1951	HIS
GLN	L923	L1004	S1193	M1273	F1344	L1441	L1554	T1718	S1840	S1840	K1951	GLN
ASN	R926	L1005	I1195	M1273	S1345	L1441	L1554	T1718	T1844	Y1845	K1957	ASN
GLY	N927	E1006	I1195	M1273	S1346	Q1442	L1554	T1718	Y1845	Y1845	K1957	GLY
ALA	N930	T1009	P1087	F1280	R1347	Q1443	I1547	T1718	H1846	H1846	K1957	ALA
GLY	R933	E1010	L1088	M1281	Y1348	F1444	L1547	T1718	H1847	E1848	K1966	GLY
GLY	L936	M1016	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	GLY
LEU	E937	M1017	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	LEU
LEU	N938	N1017	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	LEU
LEU	N939	E1019	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	LEU
SER	K940	E1019	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	SER
PRO	Y941	M1020	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	PRO
THR	Y941	M1020	F1089	M1281	E1349	F1444	L1547	T1718	H1847	E1848	P1966	THR
GLN	Y944	G1022	S1108	V1297	Y1352	L1460	L1588	R1740	S1853	D1854	M1969	GLN
LEU	S945	Y1023	E1111	V1300	Y1352	L1460	L1588	R1740	S1853	D1854	M1969	LEU
ASP	S946	T1024	A1114	Q1301	D1359	L1460	L1588	R1740	S1853	D1854	M1969	ASP
TRP	L947	L1029	A1117	V1301	D1359	L1460	L1588	R1740	S1853	D1854	M1969	TRP
ARG	S948	L1030	K1118	V1301	D1359	L1460	L1588	R1740	S1853	D1854	M1969	ARG
SER	S949	D1031	K1118	V1301	D1359	L1460	L1588	R1740	S1853	D1854	M1969	SER
ILE	V950	I1035	I1135	V1301	D1359	L1460	L1588	R1740	S1853	D1854	M1969	ILE
LYS	P951	Y1038	E1141	V1301	D1359	L1460	L1588	R1740	S1853	D1854	M1969	LYS
GLY	N953	Y1038	R1139	S1308	F1375	L1471	Y1616	K1762	P1870	P1870	L1980	GLY
LYS	P954	L1043	S1140	S1308	F1375	L1471	Y1616	K1762	P1870	P1870	L1980	LYS
GLY	Y957	C1044	I1141	S1308	F1375	L1471	Y1616	K1762	P1870	P1870	L1980	GLY
GLY	Y957	C1044	I1141	S1308	F1375	L1471	Y1616	K1762	P1870	P1870	L1980	GLY
ASN	F958	L1048	I1145	E1312	Y1382	D1476	M1623	K1768	M1992	M1992	S2073	ASN
THR	R959	D1048	V1148	Q1314	Y1382	D1476	M1623	K1768	M1992	M1992	S2073	THR
PRO	N965	A1049	V1148	Q1314	Y1382	D1476	M1623	K1768	M1992	M1992	S2073	PRO
THR	R966	R1050	E1051	R1315	D1389	S1481	E1634	L1770	M1992	M1992	S2073	THR
TYR	D968	E1053	V1052	R1315	D1389	S1481	E1634	L1770	M1992	M1992	S2073	TYR
ASP	L971	L1054	S1055	G1316	M1391	N1392	L1641	L1772	M1992	M1992	S2073	ASP
ASP	E977	D1056	D1056	E1320	T1396	K1397	L1641	L1772	M1992	M1992	S2073	ASP



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	443908	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	30	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	Not provided	
Image detector	GATAN K2 BASE (4k x 4k)	Depositor
Maximum map value	1.738	Depositor
Minimum map value	-1.169	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.045	Depositor
Recommended contour level	0.43	Depositor
Map size (\AA)	298.8, 298.8, 298.8	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	0.83, 0.83, 0.83	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.29	2/10533 (0.0%)	0.50	4/14226 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	954	PRO	CG-CD	-8.92	1.21	1.50
1	A	954	PRO	N-CD	5.05	1.54	1.47

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	954	PRO	CA-N-CD	-11.03	96.05	111.50
1	A	954	PRO	N-CD-CG	-10.91	86.83	103.20
1	A	954	PRO	CA-CB-CG	-6.55	91.55	104.00
1	A	953	ASN	C-N-CD	6.05	141.11	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1055	SER	Peptide

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	A	10350	0	10260	218	0
All	All	10350	0	10260	218	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 11.

The worst 5 of 218 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)
1:A:1969:ASN:HD21	1:A:2077:ILE:HB	1.48	0.78
1:A:1204:GLY:O	1:A:1238:LYS:NZ	2.17	0.76
1:A:1655:LYS:HB3	1:A:1707:GLU:HA	1.68	0.75
1:A:1370:GLU:OE1	1:A:1378:ARG:NH1	2.20	0.74
1:A:2191:LYS:HG3	1:A:2192:ASP:H	1.54	0.72

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	A	1293/2201 (59%)	1176 (91%)	117 (9%)	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	1148/1915 (60%)	1143 (100%)	5 (0%)	91 95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	926	ARG
1	A	1605	ASN
1	A	1622	LYS
1	A	2002	ARG
1	A	2174	ARG

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

Mol	Chain	Res	Type
1	A	1322	ASN
1	A	1715	ASN
1	A	1969	ASN

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](#)

There are no ligands in this entry.

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.

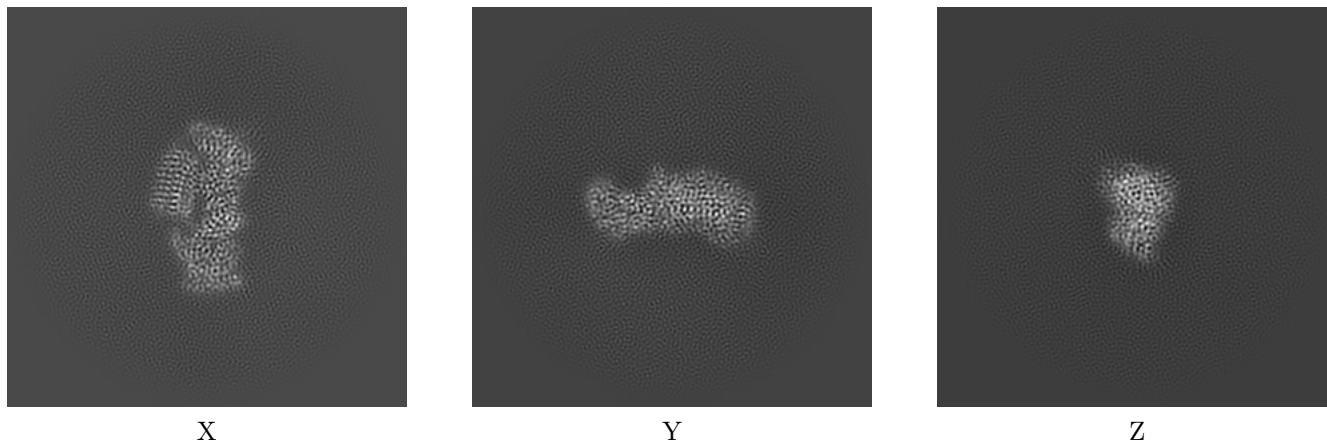
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-26812. These allow visual inspection of the internal detail of the map and identification of artifacts.

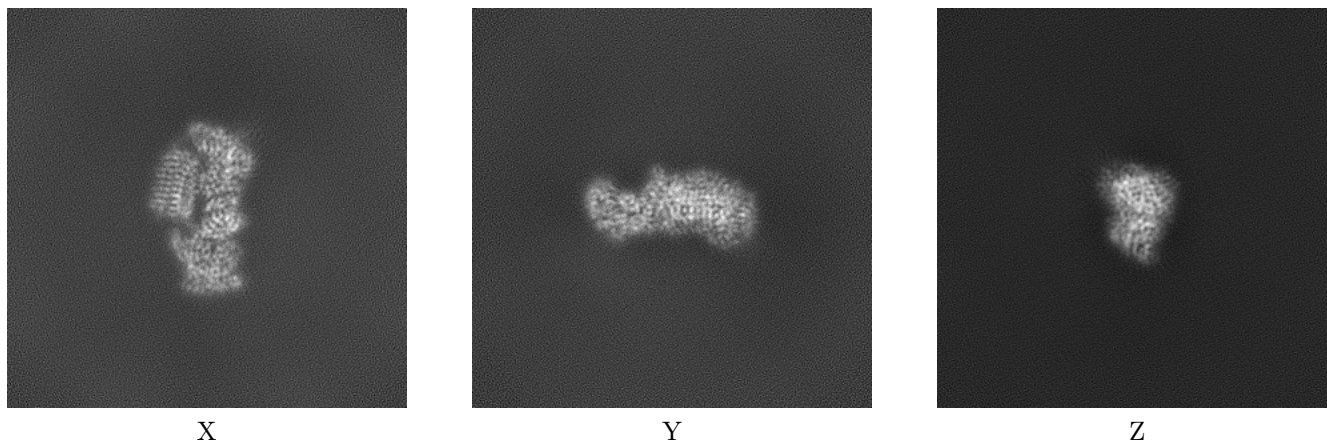
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

6.1.1 Primary map



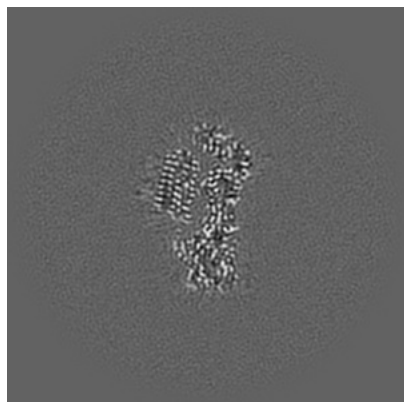
6.1.2 Raw map



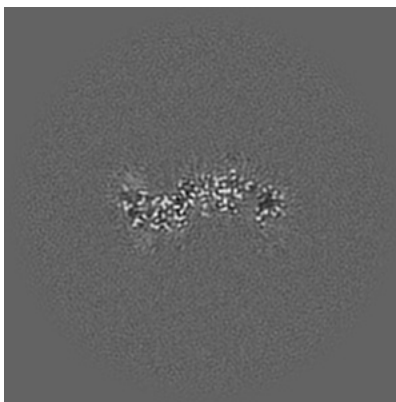
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

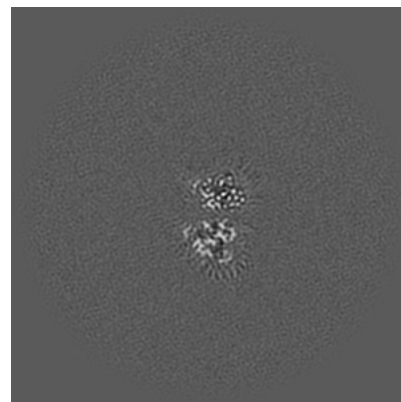
6.2.1 Primary map



X Index: 180

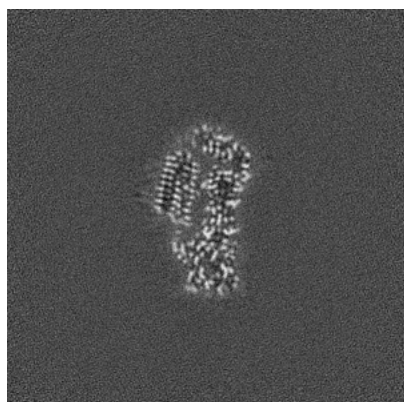


Y Index: 180

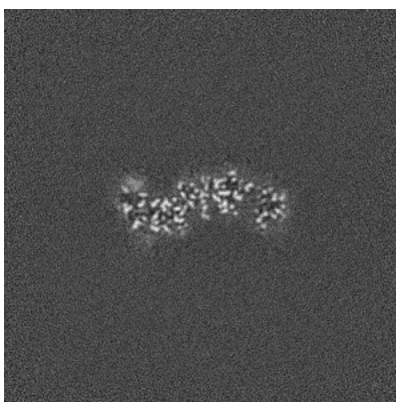


Z Index: 180

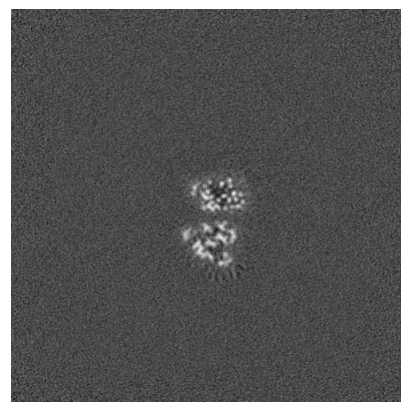
6.2.2 Raw map



X Index: 180



Y Index: 180

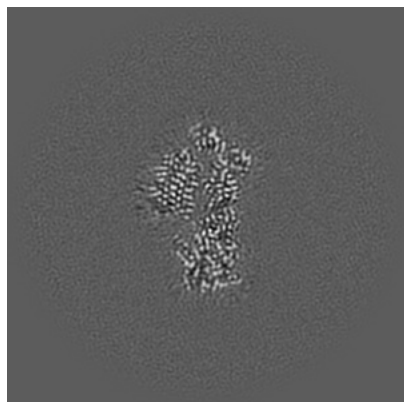


Z Index: 180

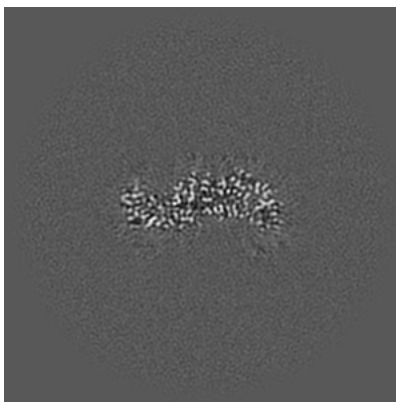
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

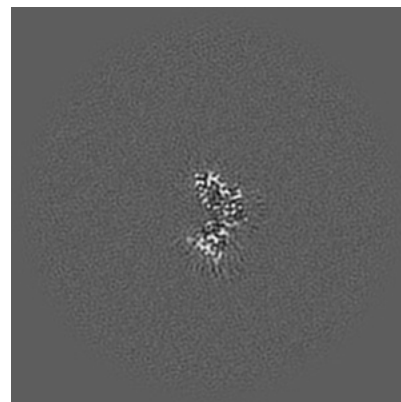
6.3.1 Primary map



X Index: 184

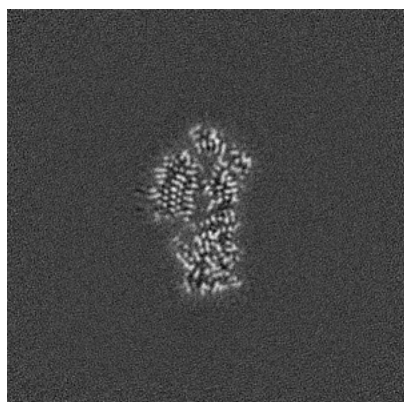


Y Index: 188

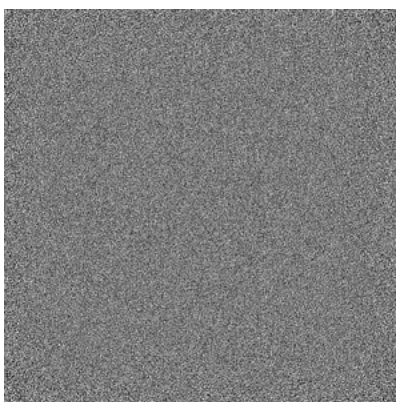


Z Index: 197

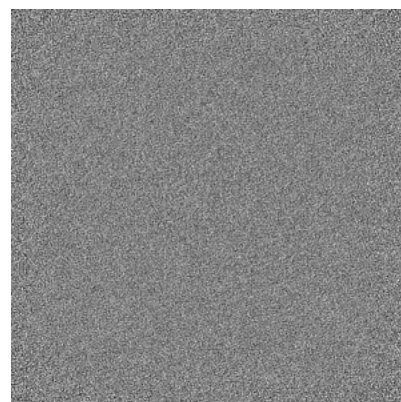
6.3.2 Raw map



X Index: 184



Y Index: 0

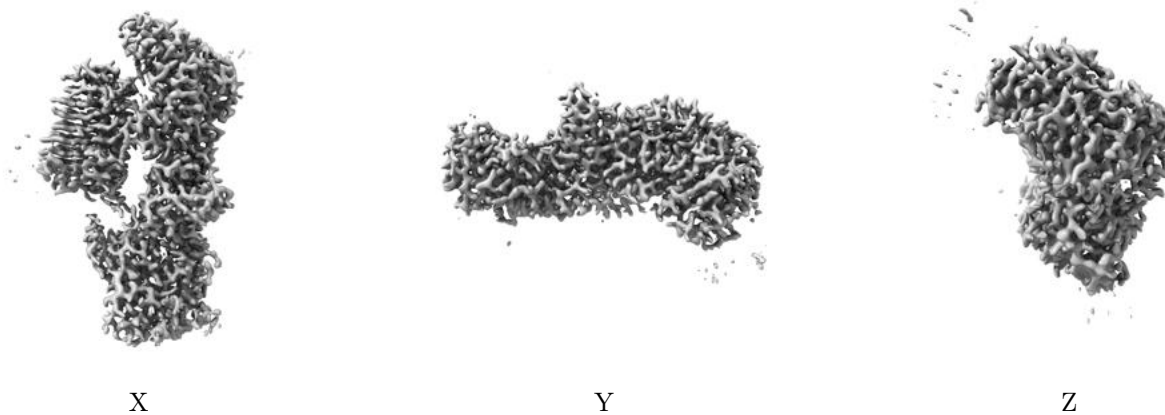


Z Index: 0

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

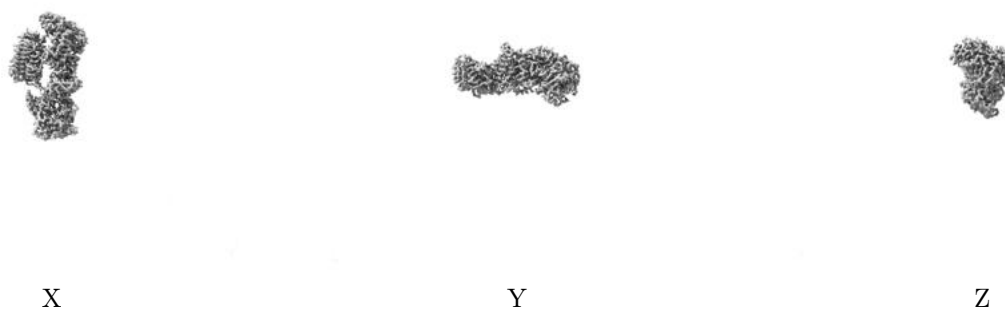
6.4 Orthogonal surface views [i](#)

6.4.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.43. 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.

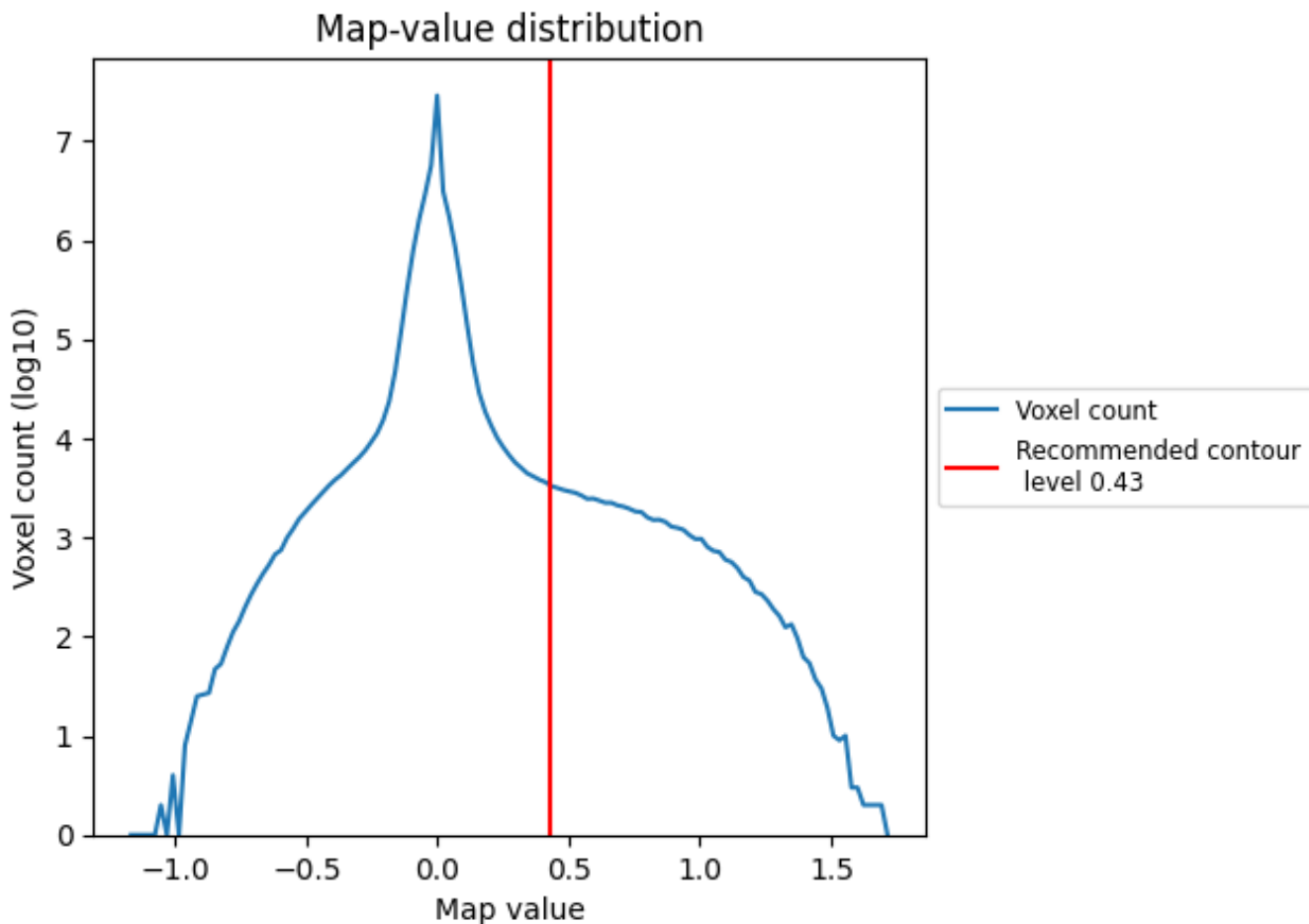
6.5 Mask visualisation [i](#)

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

7 Map analysis [i](#)

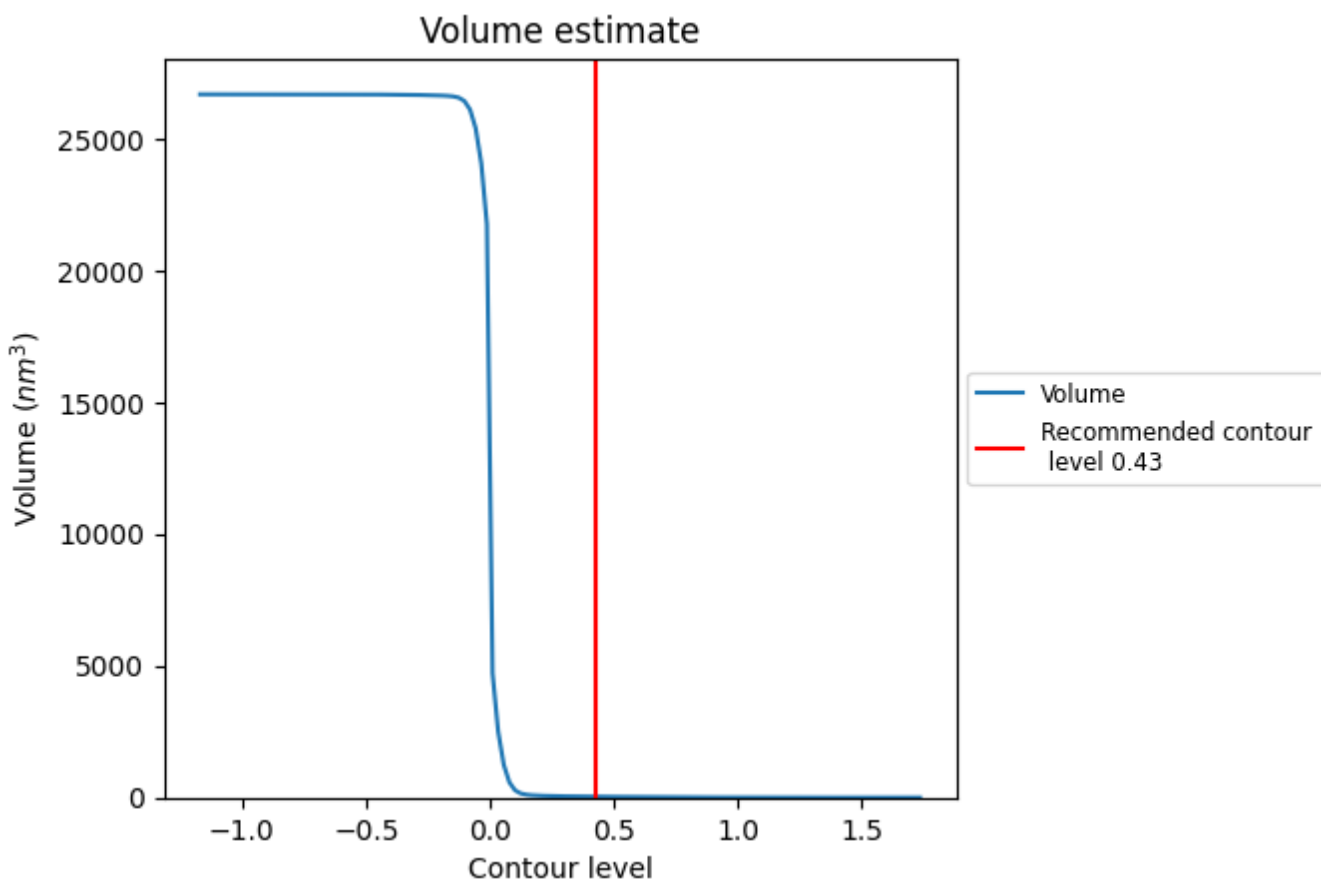
This section contains the results of statistical analysis of the map.

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.

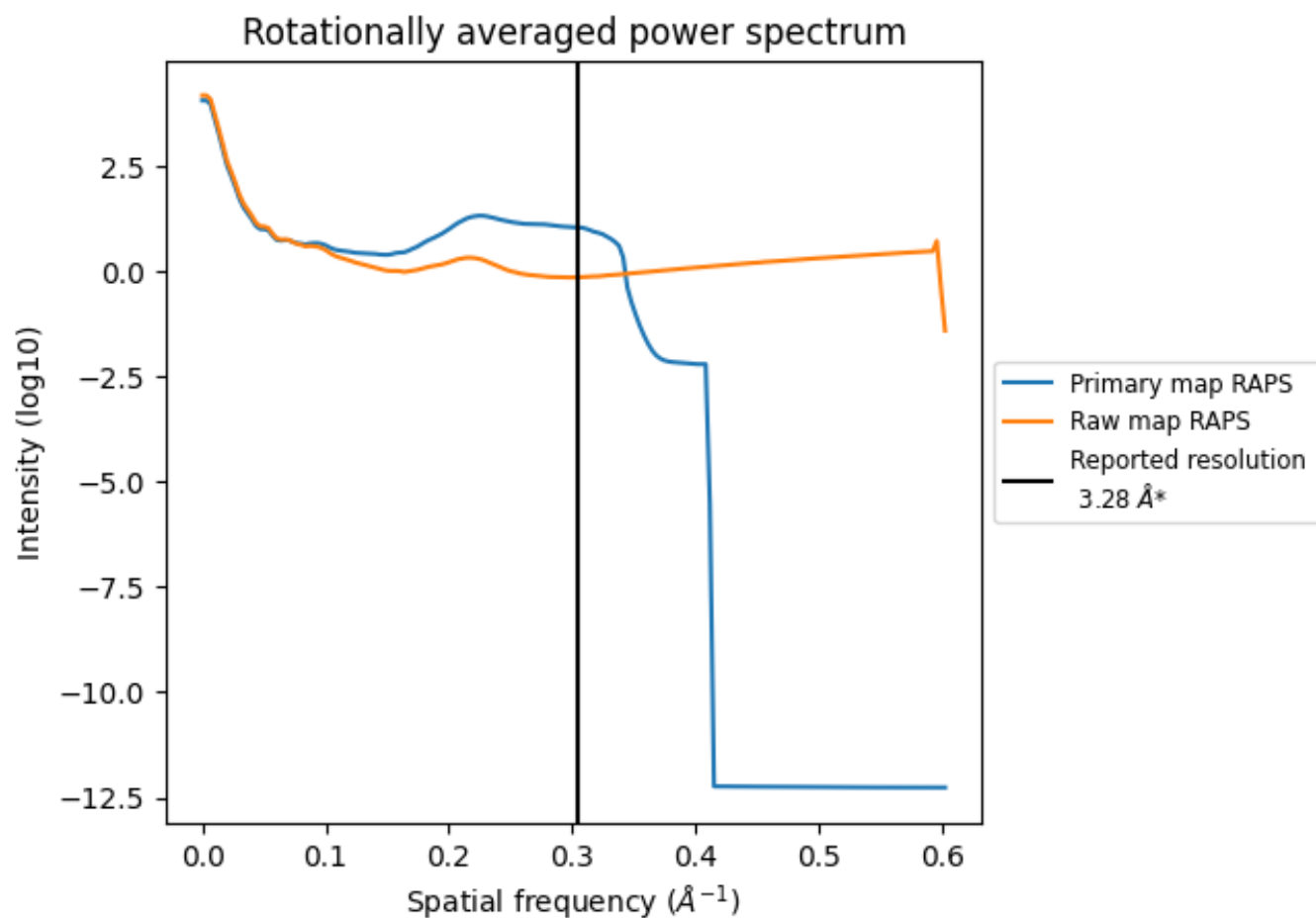
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 34 nm^3 ; this corresponds to an approximate mass of 31 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](#)

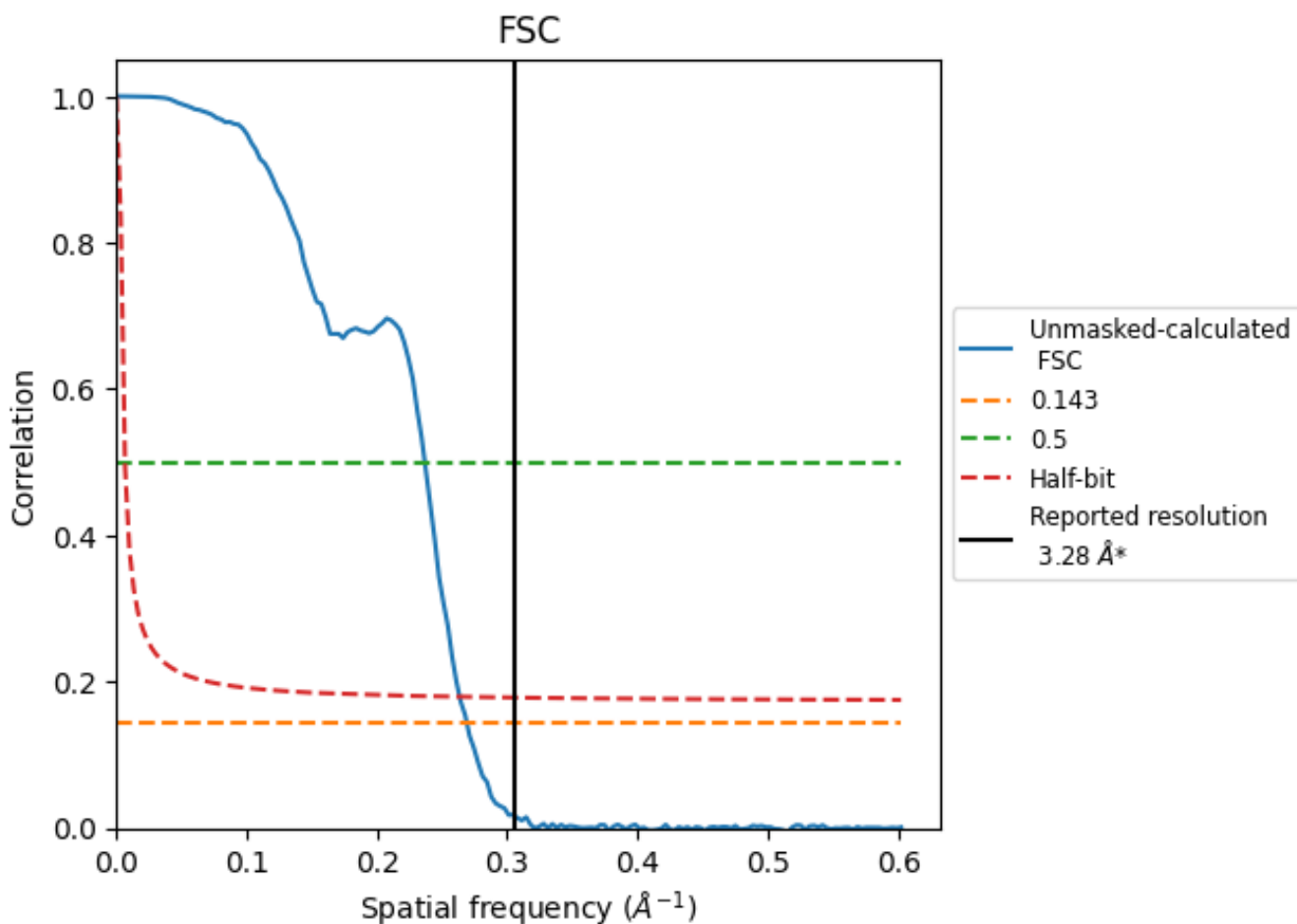


*Reported resolution corresponds to spatial frequency of 0.305 Å⁻¹

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 0.305 Å⁻¹

8.2 Resolution estimates [i](#)

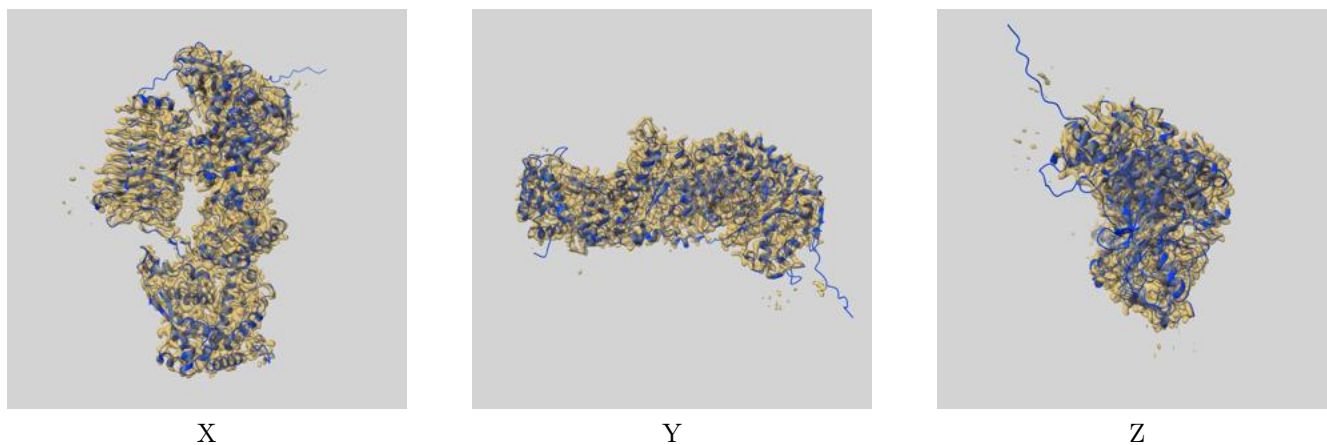
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.28	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	3.72	4.22	3.80

*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 3.72 differs from the reported value 3.28 by more than 10 %

9 Map-model fit [i](#)

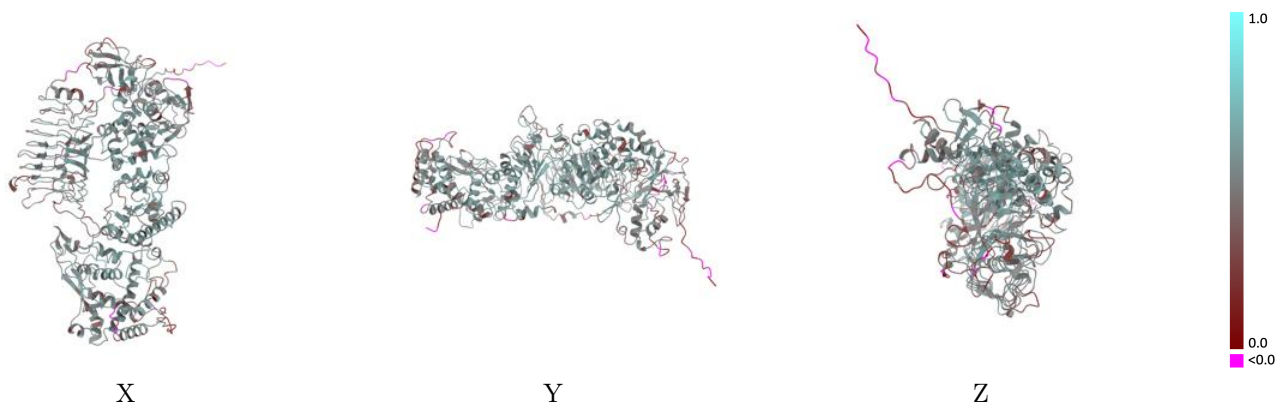
This section contains information regarding the fit between EMDB map EMD-26812 and PDB model 7UVK. Per-residue inclusion information can be found in section [3](#) on page [4](#).

9.1 Map-model overlay [i](#)



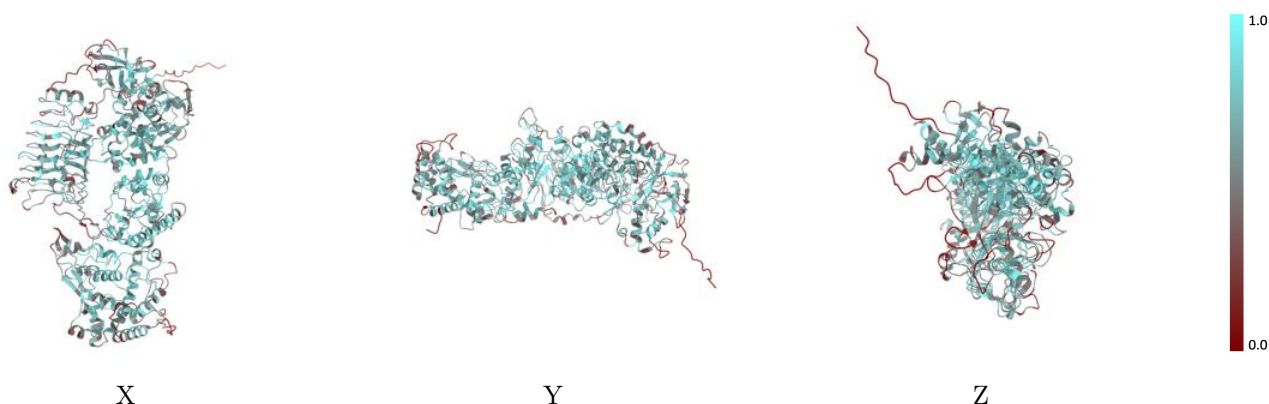
The images above show the 3D surface view of the map at the recommended contour level 0.43 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



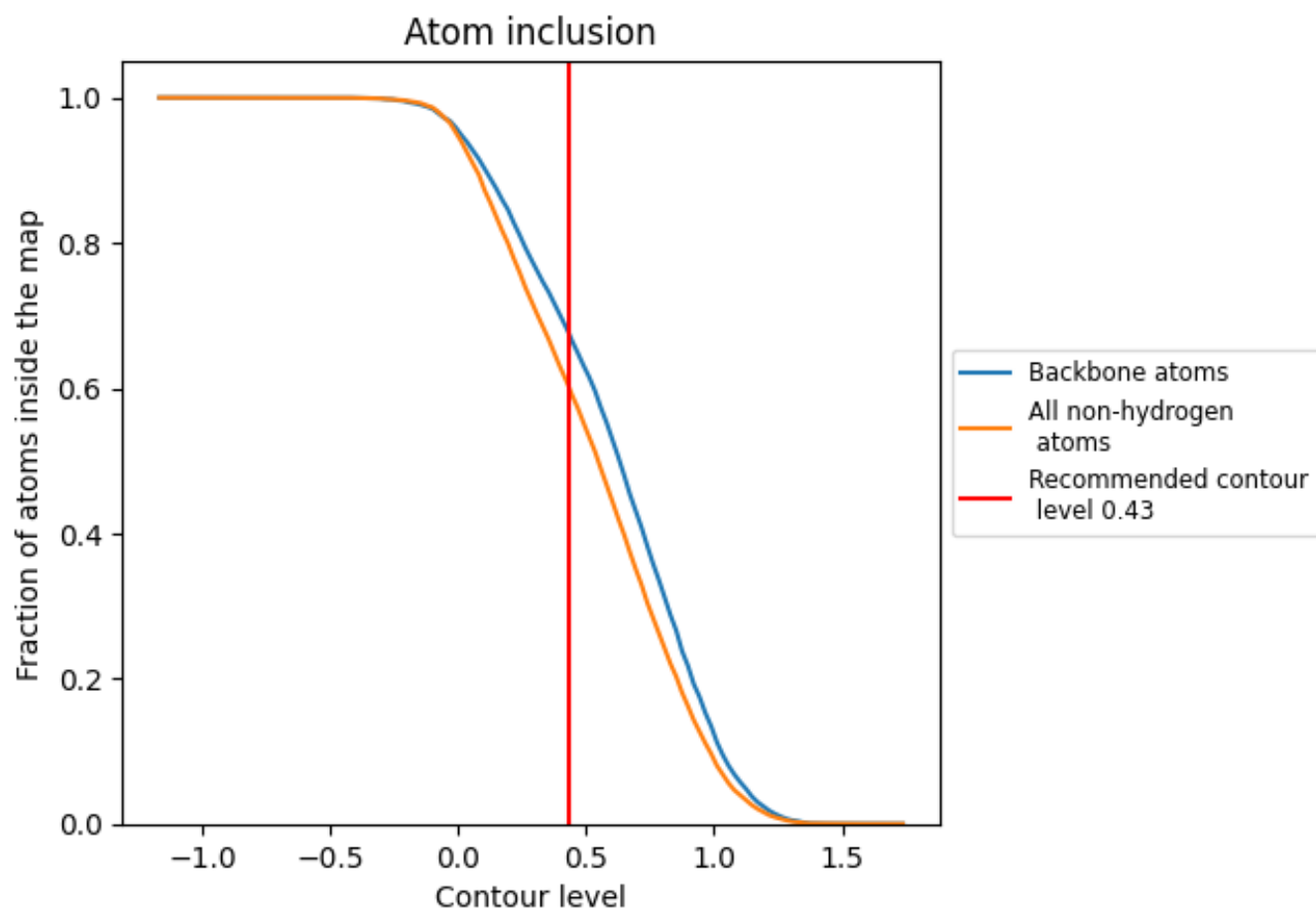
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.43).





9.4 Atom inclusion [i](#)



At the recommended contour level, 68% of all backbone atoms, 60% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary [i](#)

The table lists the average atom inclusion at the recommended contour level (0.43) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6037	 0.4820
A	 0.6037	 0.4820

