



wwPDB EM Validation Summary Report ⓘ

Nov 19, 2022 – 10:36 AM EST

PDB ID : 7R76
EMDB ID : EMD-24292
Title : Cryo-EM structure of DNMT5 in apo state
Authors : Wang, J.; Patel, D.J.
Deposited on : 2021-06-24
Resolution : 3.20 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) : 2.31.3

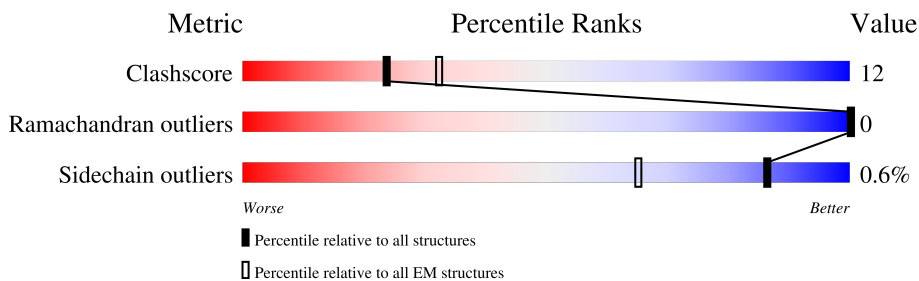
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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	2348	 57% 21% 22%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14500 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 DNA repair protein Rad8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1834	14495	9110	2618	2697	70	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	expression tag	UNP J9VI03
A	31	SER	-	expression tag	UNP J9VI03
A	32	TYR	-	expression tag	UNP J9VI03
A	33	TYR	-	expression tag	UNP J9VI03
A	34	HIS	-	expression tag	UNP J9VI03
A	35	HIS	-	expression tag	UNP J9VI03
A	36	HIS	-	expression tag	UNP J9VI03
A	37	HIS	-	expression tag	UNP J9VI03
A	38	HIS	-	expression tag	UNP J9VI03
A	39	HIS	-	expression tag	UNP J9VI03
A	40	ASP	-	expression tag	UNP J9VI03
A	41	TYR	-	expression tag	UNP J9VI03
A	42	ASP	-	expression tag	UNP J9VI03
A	43	ILE	-	expression tag	UNP J9VI03
A	44	PRO	-	expression tag	UNP J9VI03
A	45	THR	-	expression tag	UNP J9VI03
A	46	THR	-	expression tag	UNP J9VI03
A	47	GLU	-	expression tag	UNP J9VI03
A	48	ASN	-	expression tag	UNP J9VI03
A	49	LEU	-	expression tag	UNP J9VI03
A	50	TYR	-	expression tag	UNP J9VI03
A	51	PHE	-	expression tag	UNP J9VI03
A	52	GLN	-	expression tag	UNP J9VI03
A	53	GLY	-	expression tag	UNP J9VI03
A	54	ALA	-	expression tag	UNP J9VI03
A	55	MET	-	expression tag	UNP J9VI03
A	56	GLY	-	expression tag	UNP J9VI03
A	57	SER	-	expression tag	UNP J9VI03

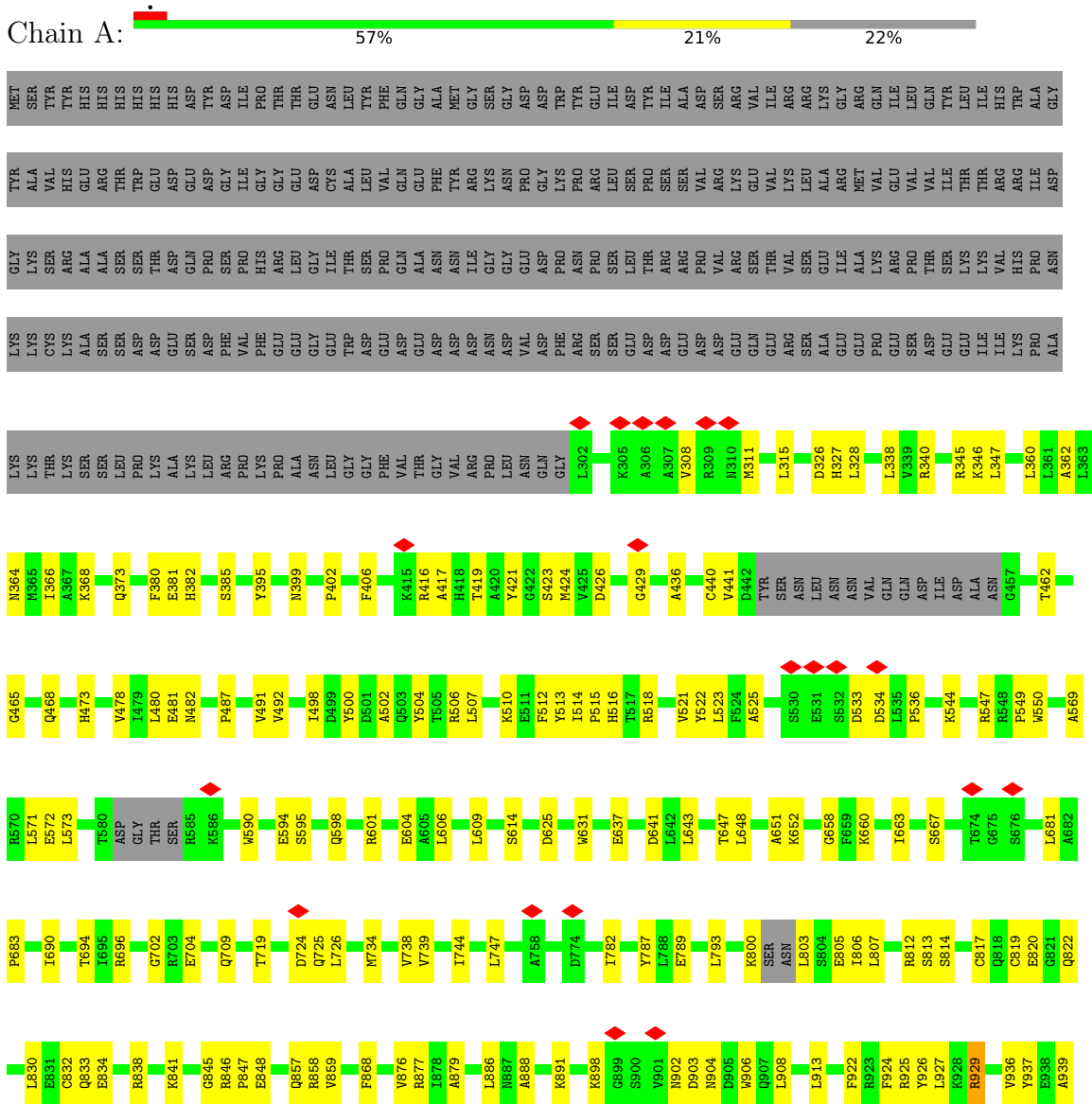
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
2	A	5	Total 5	Zn 5	0

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: DNA repair protein Rad8



P2246	P2250	S2251	L2252	Y2255	R2256	E2259	A2262	L2263	G2264	R2265	V2266	R2267	R2268	Y2269	H2276	L2280	L2281	A2282	L2283	D2284	T2285	L2286	D2287	M2288	T2289	L2290	F2291	R2294	R2295	L2298	K2301	E2310	E2311	Y2312	R2315	GLY	SER	SER	ILE	SER	SER	THR	ASN	GLU	ARG	THR									
S2146	G2147	A2151	K2152	L2153	E2154	H2155	L2156	T2160	P2163	L2164	E2167	R2168	F2172	L2173	V2182	V2194	T2195	L2196	S2197	G2198	S2199	A2200	K2201	F2210	Q2211	S2212	T2213	ASN	ALA	ASP	S2217	A2218	L2221	L2222	L2223	K2224	D2227	ALA	SER	ALA	ALA	ALA	G2232	S2233	M2234	H2240	L2244	G2245							
S2070	D2071	S2079	E2080	C2081	C2082	H2083	K2084	PRO	SER	THR	ASN	PRO	ASP	ILE	E2092	M2093	A2094	L2096	C2099	G2100	H2101	V2102	K2106	E2109	K2110	A2111	A2112	A2113	R2116	C2117	V2118	K2119	S2120	GLY	GLU	C2123	Q2124	A2125	A2126	V2127	R2128	P2129	K2134	V2135	S2136	S2137	E2141	G2142	E2143	L2144	S2145				
H1986	D1989	I1990	E1991	R1994	I1995	K1998	V1999	I2000	G2004	V2005	K2006	D2007	G2008	M2009	T2010	P2011	F1907	K1914	A1931	R1932	A1933	R1934	Q1935	L1936	L1937	E1941	M1948	Q1949	A1950	I1951	A1952	K1959	F1963	S1964	K1965	M1966	D1967	D1968	Q1971	F1972	F1973	I1977	K1985												
SER	VAL	LYS	LYS	ASP	TVR	K1707	P1710	V1713	L1714	H1715	M1716	F1717	R1718	F1719	R1720	R1721	V1722	I1723	K1732	A1737	R1740	L1741	S1742	S1743	S1744	H1745	R1746	M1747	I1748	L1749	P1753	D1757	F1758	A1759	A1760	S1763	I1764	H1771	G1780	D1781	Y1784	V1804	R1812	R1813	I1829	K1985									
ALA	GLU	I1832	H1840	F1844	K1845	L1846	E1850	L1861	A1863	Q1867	R1882	L1890	A1896	F1907	K1914	A1931	R1932	A1933	R1934	Q1935	L1936	L1937	E1941	M1948	Q1949	A1950	I1951	A1952	K1959	F1963	S1964	K1965	M1966	D1967	D1968	Q1971	F1972	F1973	I1977	K1985															
MET	LYS	GLY	GLN	ALA	TYR	ARG	ASP	HIS	ASP	SER	THR	LYS	GLU	LEU	GLY	ARG	TRP	SER	GLU	ALA	SER	GLU	ASP	GLU	ALA	ASP	GLU	ASP	GLU	ASN	SER	LYS	SER	LEU	VAL	ASP	ASN	ILE	VAL	PRO	GLY	GLY	THR	HIS	SER	ALA	VAL	ASN	PHE	ILE	SER	PHE	GLY	LYS	ARG
LYS	ASP	ALA	ALA	ALA	ASN	GLU	ILE	HIS	ARG	ASP	THR	LYS	GLU	LEU	ARG	TRP	GLU	ALA	SER	GLU	ASP	GLU	ALA	ASP	GLU	ASP	GLU	ASN	SER	LYS	SER	LEU	VAL	ASP	ASN	ILE	VAL	PRO	GLY	GLY	THR	HIS	SER	ALA	VAL	ASN	PHE	ILE	SER	PHE	GLY	LYS	ARG		
ASN	ASP	ALA	ALA	ALA	ASN	GLU	ILE	HIS	ARG	ASP	THR	LYS	GLU	LEU	ARG	TRP	GLU	ALA	SER	GLU	ASP	GLU	ALA	ASP	GLU	ASP	GLU	ASN	SER	LYS	SER	LEU	VAL	ASP	ASN	ILE	VAL	PRO	GLY	GLY	THR	HIS	SER	ALA	VAL	ASN	PHE	ILE	SER	PHE	GLY	LYS	ARG		
G1057	T1058	Y1059	L1062	M1069	R1076	S1080	D1081	D1082	G1083	L1084	P1085	F1090	D1102	R1103	Y1104	Y1105	F1106	S1107	V1121	L1125	D1126	S1127	H1128	W1129	R1130	E1134	R1137	K1138	V1139	K1140	S1144	W1147	C1150	H1154	I1158	GLY	GLY	ASP	ASP	ILE	ALA	ALA	VAL	VAL	ALA	ALA									
L944	S945	L946	P953	E954	W955	E965	S969	O970	L971	R972	L982	O983	I984	D985	I986	A987	G988	Q989	L992	W996	F997	L998	C999	M1003	I1008	E1009	I1010	T1011	G1012	K1013	S1019	W1020	R1035	E1040	I1041	S1042	L1043	Q1044	D1047	L1051	D1052	R1053	K1054	L1055	S1056										

PRO
THR
LEU
THR
VAL
LYS
SER
ASN
PRO
PHE
LYS
ARG
SER
SER
SER
TRP
ALA
LEU
ALA
SER
SER
PHE
ARG
SER
LYS
LYS
ARG
SER
MET
GLU
ALA
ARG
ASP
ALA
GLU
GLY
VAL
SER
ASP
ASP
ASP
GLU
ASN
SER
SER
GLU
LEU
SER
SER
ASP
ASP
ILE
ILE

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	72216	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 ($e^-/\text{\AA}^2$)	53	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	47262	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	0.103	Depositor
Minimum map value	-0.052	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	272.384, 272.384, 272.384	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.064, 1.064, 1.064	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

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	0.40	0/14788	0.58	0/20010

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	A	14495	0	14487	350	0
2	A	5	0	0	0	0
All	All	14500	0	14487	350	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 12.

The worst 5 of 350 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:1010:ILE:HG21	1:A:1129:TRP:HZ2	1.03	1.14
1:A:2259:GLU:O	1:A:2263:ILE:HG13	1.50	1.10

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1010:ILE:HG21	1:A:1129:TRP:CZ2	1.88	1.07
1:A:1602:LEU:HD12	1:A:1707:LYS:NZ	1.70	1.07
1:A:2117:CYS:HG	1:A:2123:CYS:N	1.59	0.99

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	1808/2348 (77%)	1645 (91%)	163 (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	1572/2015 (78%)	1562 (99%)	10 (1%)	86 94

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1994	ARG
1	A	2234	ASN
1	A	2267	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1130	ARG
1	A	1342	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2234	ASN
1	A	2240	HIS
1	A	2278	HIS
1	A	2276	HIS
1	A	2101	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 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

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.

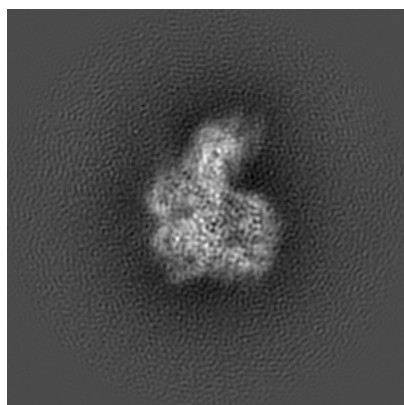
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-24292. These allow visual inspection of the internal detail of the map and identification of artifacts.

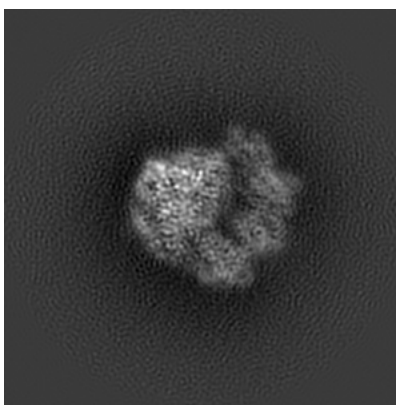
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

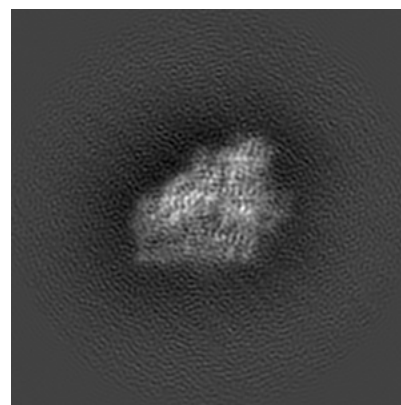
6.1.1 Primary map



X



Y

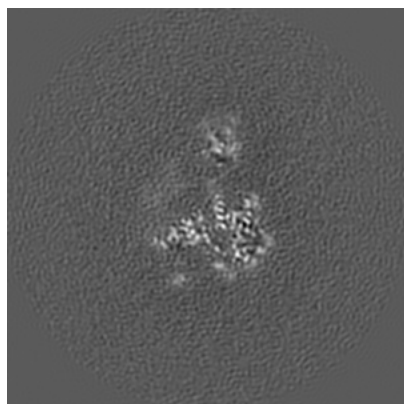


Z

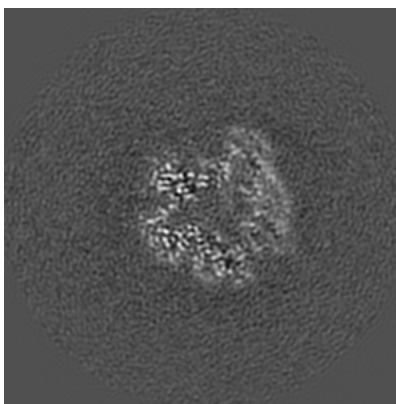
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

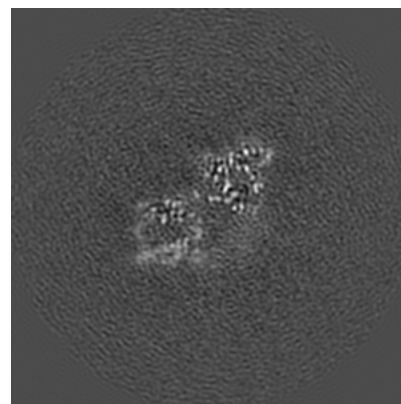
6.2.1 Primary map



X Index: 128



Y Index: 128

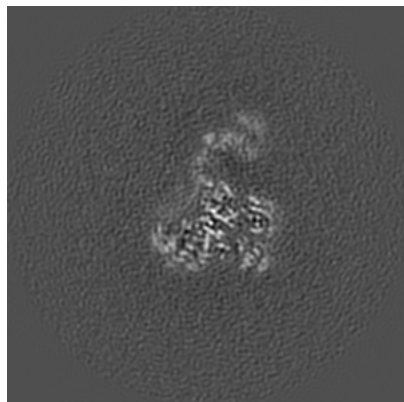


Z Index: 128

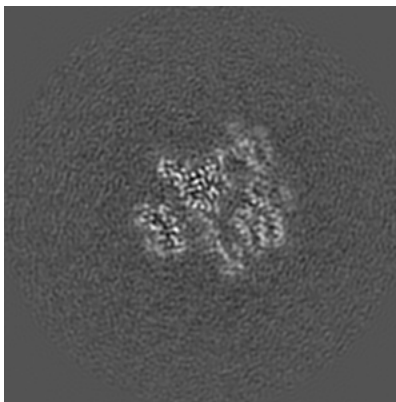
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

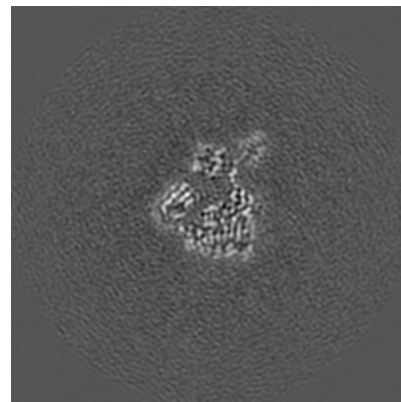
6.3.1 Primary map



X Index: 148



Y Index: 135



Z Index: 110

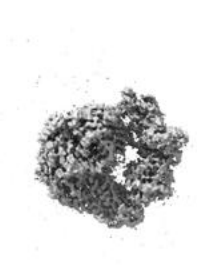
The images above show the largest variance slices of the map in three orthogonal directions.

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 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

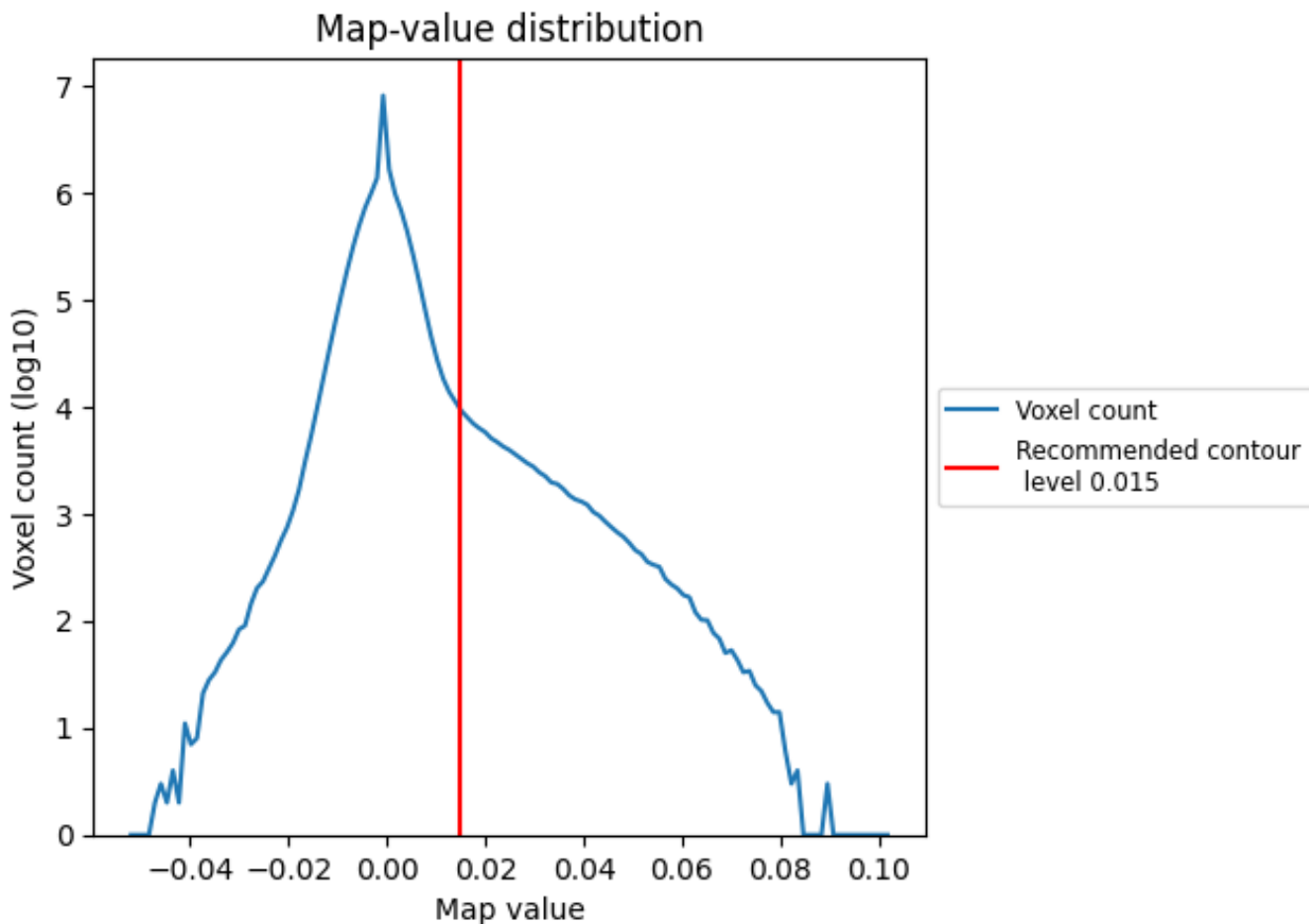
6.5 Mask visualisation

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

7 Map analysis [i](#)

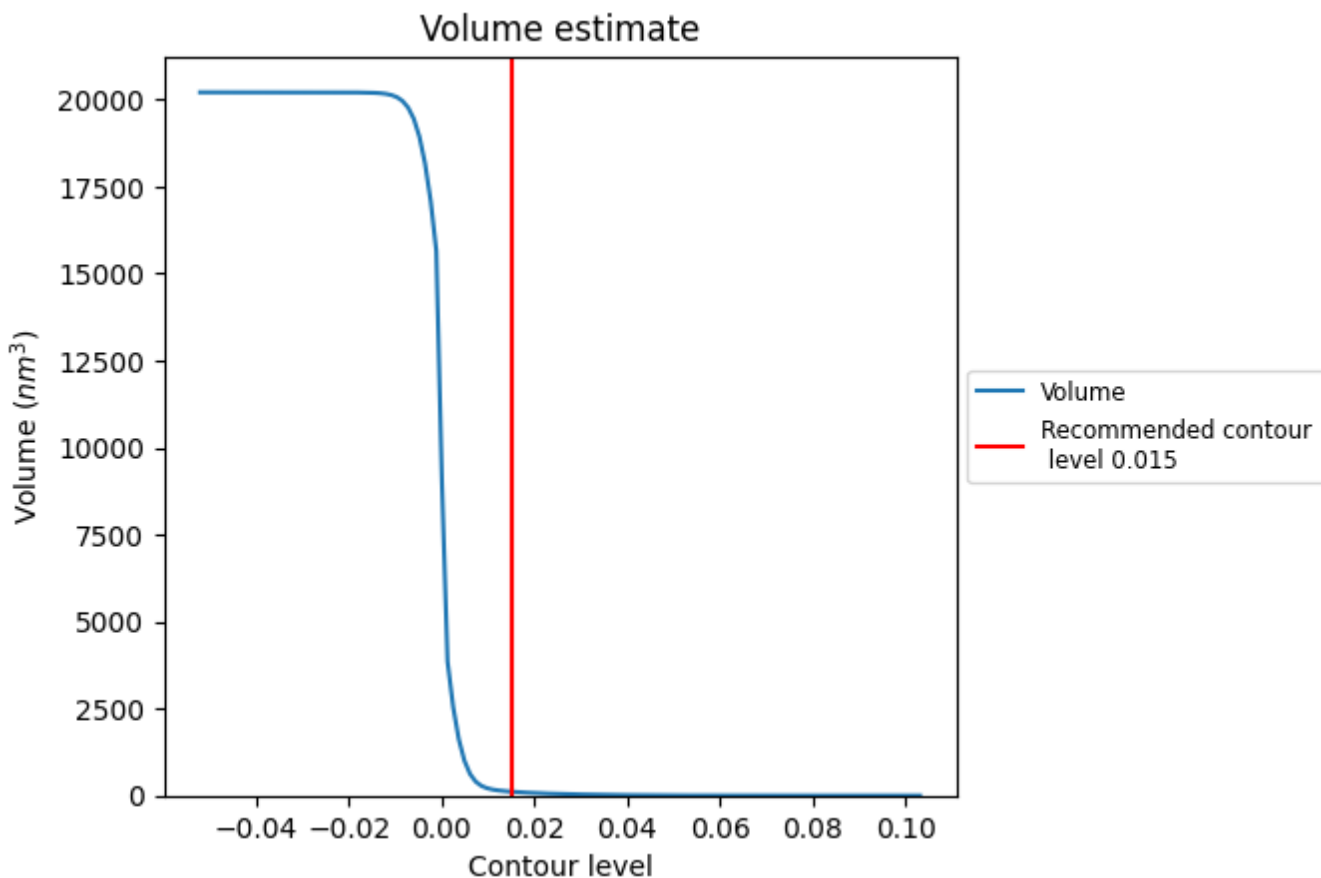
This section contains the results of statistical analysis of the map.

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.

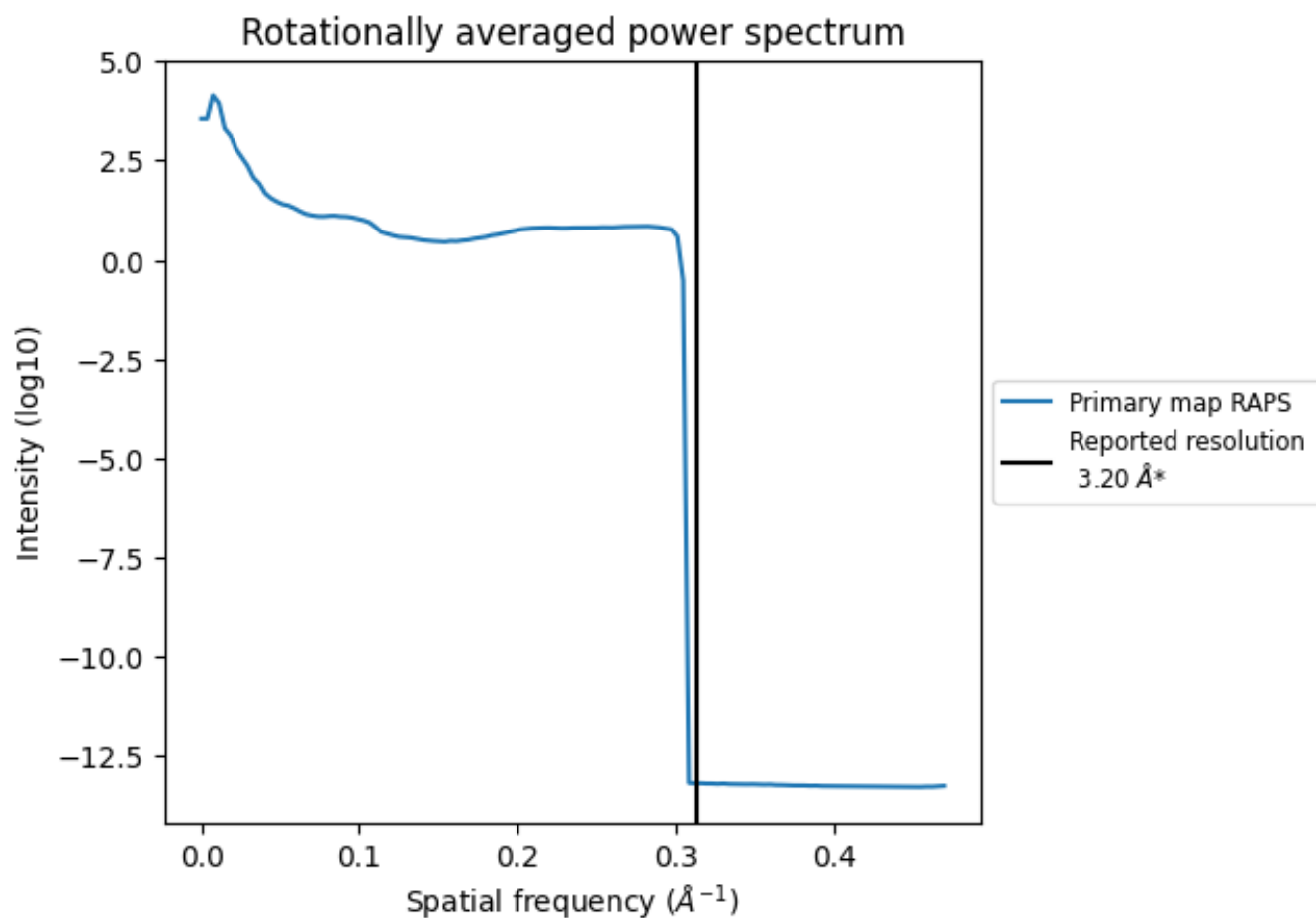
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 112 nm³; this corresponds to an approximate mass of 101 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 0.312 Å⁻¹

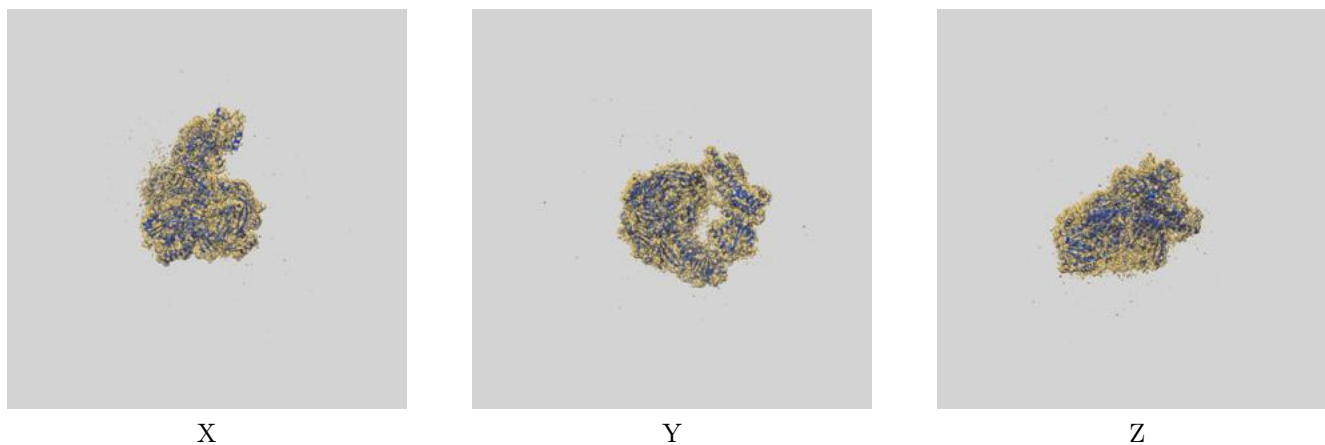
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

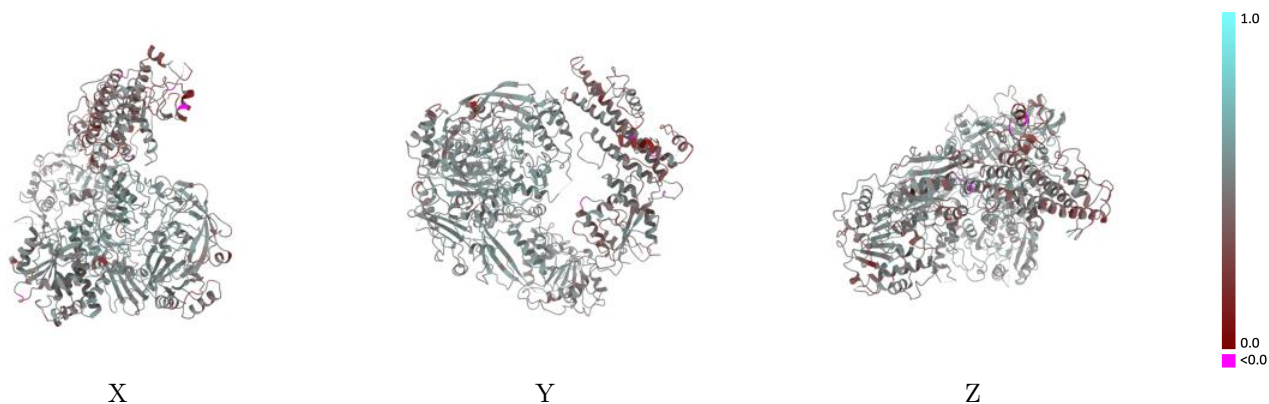
This section contains information regarding the fit between EMDB map EMD-24292 and PDB model 7R76. Per-residue inclusion information can be found in section [3](#) on page [5](#).

9.1 Map-model overlay [i](#)



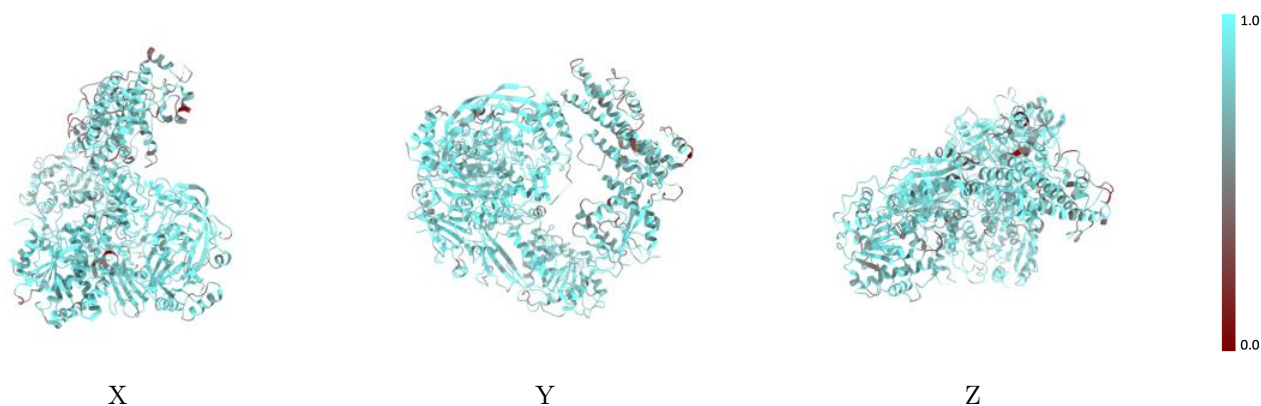
The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



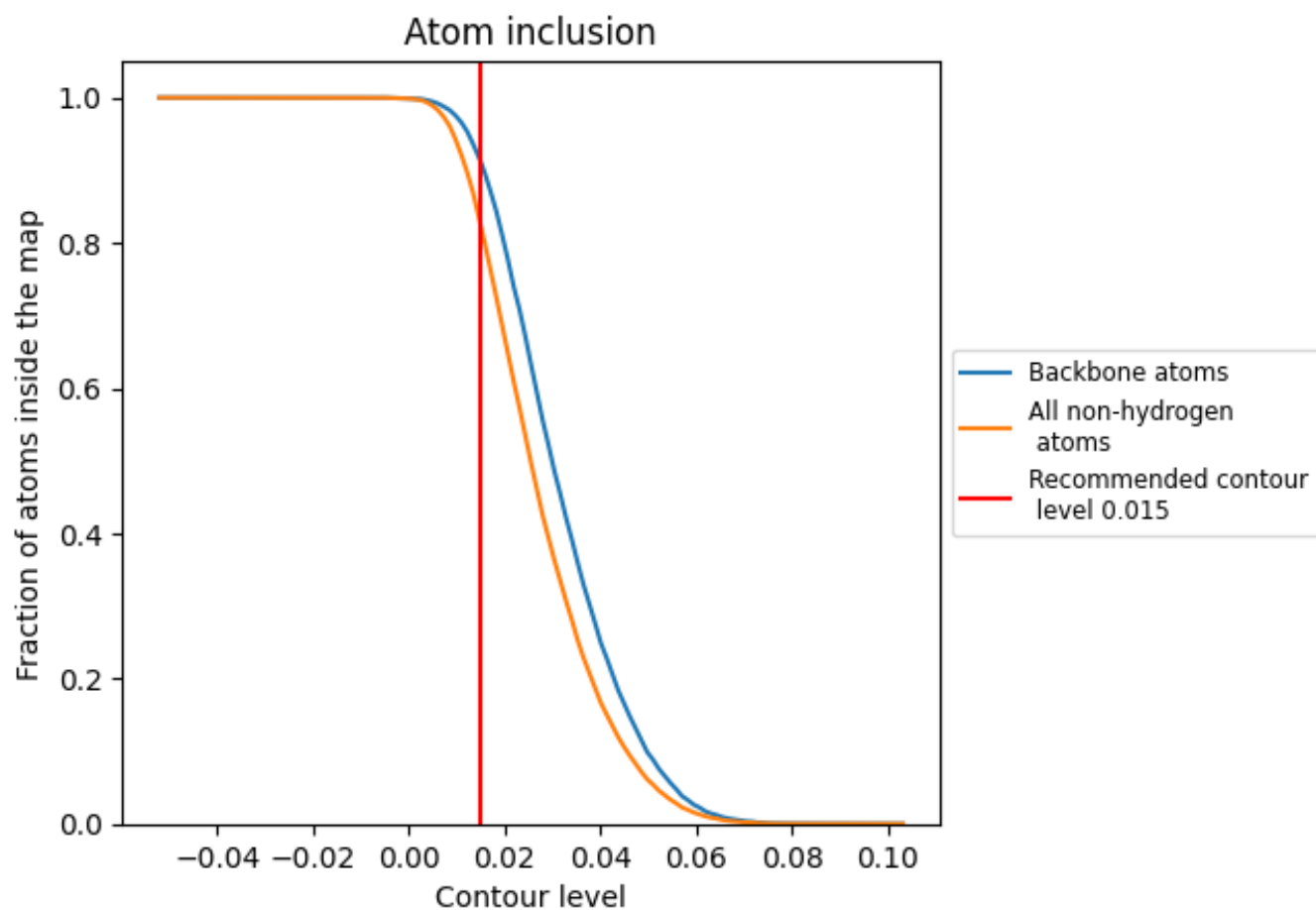
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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, 91% of all backbone atoms, 82% 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 (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.8242	 0.4790
A	 0.8242	 0.4790

