



# wwPDB EM Validation Summary Report ⓘ

Mar 12, 2024 – 12:40 pm GMT

PDB ID : 8QEN  
EMDB ID : EMD-18373  
Title : cryo-EM structure of apo Clostridioides difficile toxin B  
Authors : Kinsolving, J.; Bous, J.; Structural Genomics Consortium (SGC)  
Deposited on : 2023-09-01  
Resolution : 3.40 Å (reported)  
Based on initial model : .

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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.dev70  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.13  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

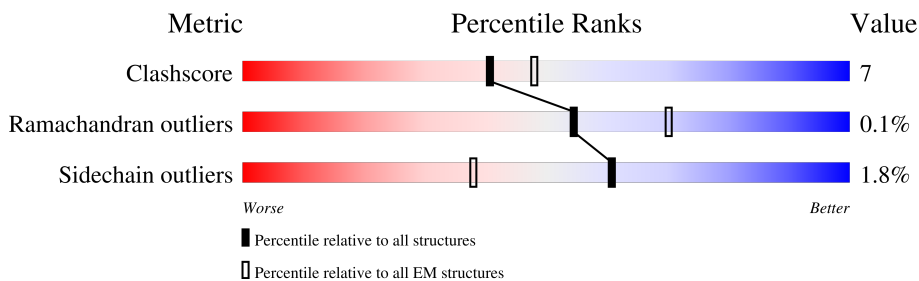
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.40 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  $\geq 3$ , 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	2397	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 19007 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 Toxin B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	2363	19006	12122	2981	3850	53	0	0

There are 31 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	initiating methionine	UNP P18177
A	-12	ASP	-	expression tag	UNP P18177
A	-11	LYS	-	expression tag	UNP P18177
A	-10	LEU	-	expression tag	UNP P18177
A	-9	VAL	-	expression tag	UNP P18177
A	-8	HIS	-	expression tag	UNP P18177
A	-7	LEU	-	expression tag	UNP P18177
A	-6	ASN	-	expression tag	UNP P18177
A	-5	GLN	-	expression tag	UNP P18177
A	-4	ARG	-	expression tag	UNP P18177
A	-3	GLY	-	expression tag	UNP P18177
A	-2	LYS	-	expression tag	UNP P18177
A	-1	CYS	-	expression tag	UNP P18177
A	0	THR	-	expression tag	UNP P18177
A	2367	GLY	-	expression tag	UNP P18177
A	2368	TYR	-	expression tag	UNP P18177
A	2369	ARG	-	expression tag	UNP P18177
A	2370	PRO	-	expression tag	UNP P18177
A	2371	HIS	-	expression tag	UNP P18177
A	2372	ALA	-	expression tag	UNP P18177
A	2373	GLY	-	expression tag	UNP P18177
A	2374	LEU	-	expression tag	UNP P18177
A	2375	ARG	-	expression tag	UNP P18177
A	2376	GLY	-	expression tag	UNP P18177
A	2377	SER	-	expression tag	UNP P18177
A	2378	HIS	-	expression tag	UNP P18177
A	2379	HIS	-	expression tag	UNP P18177
A	2380	HIS	-	expression tag	UNP P18177

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	2381	HIS	-	expression tag	UNP P18177
A	2382	HIS	-	expression tag	UNP P18177
A	2383	HIS	-	expression tag	UNP P18177

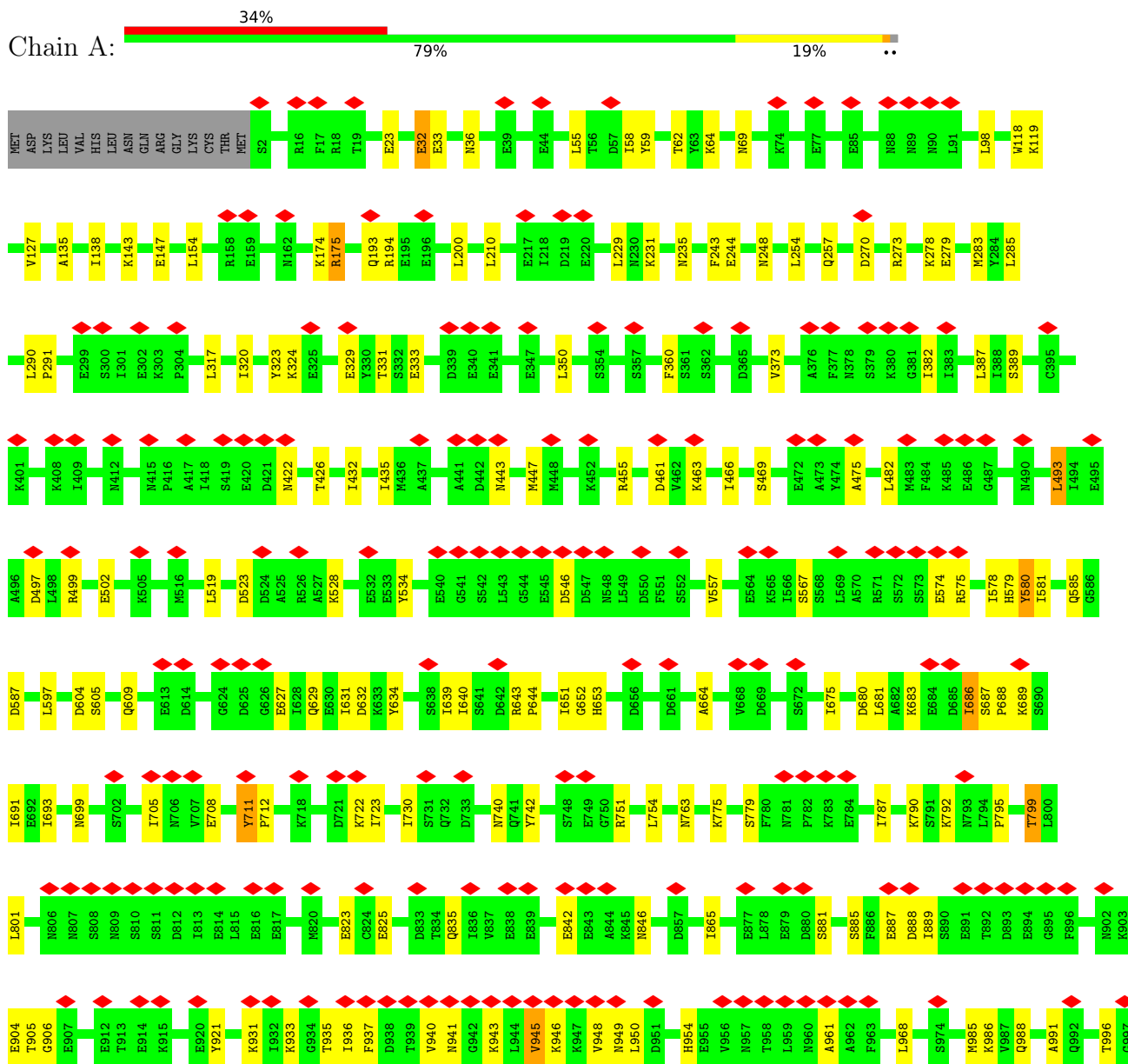
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

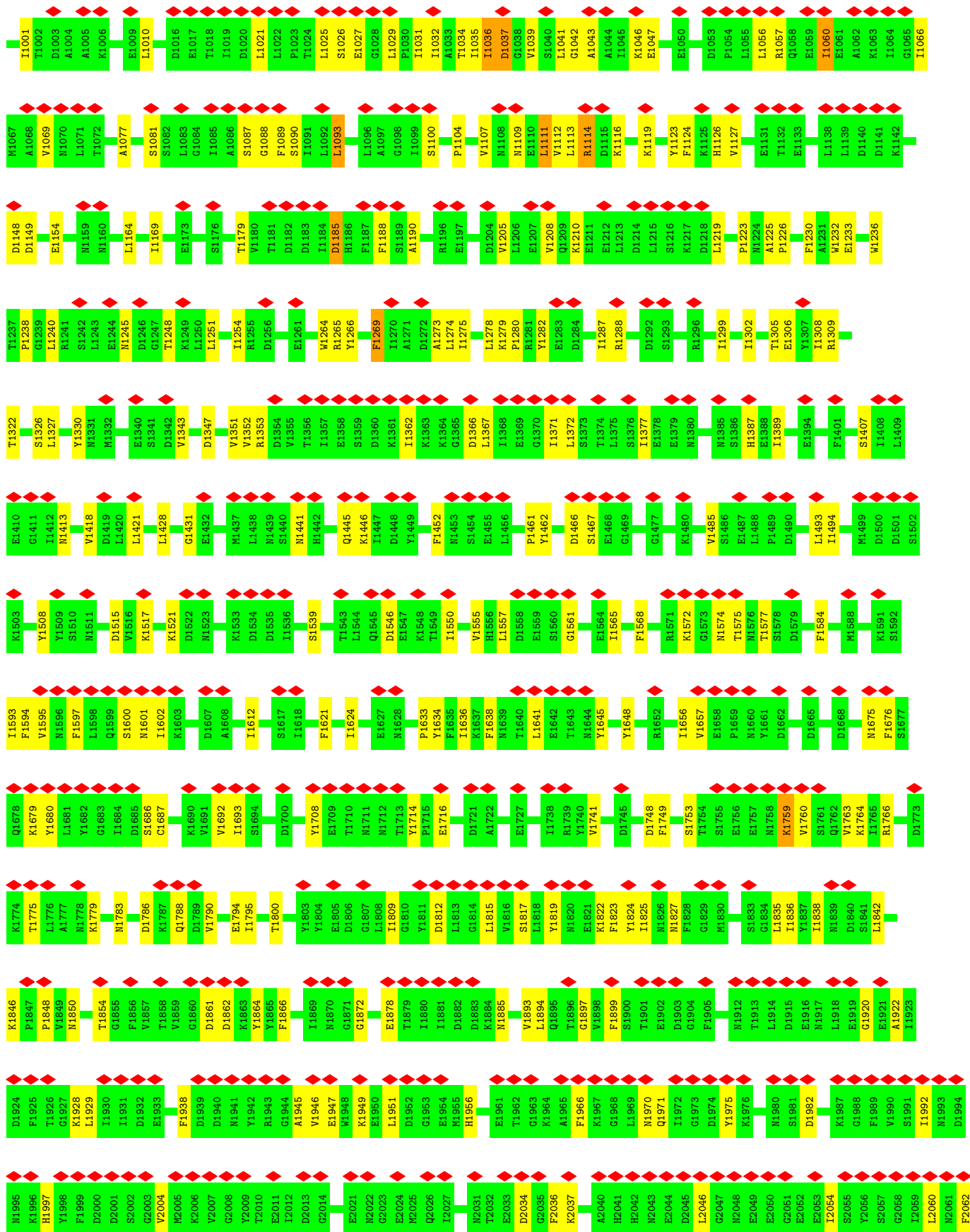
Mol	Chain	Residues	Atoms		AltConf
2	A	1	Total	Zn	0
			1	1	

### 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: Toxin B





N2063	N2064	R2065	I2066	Y2067	F2068	D2070	D2071	A2075	V2076	G2078	W2079	K2080	D2081	L2082	E2083	G2085	S2086	K2087	Y2088	Y2089	F2090	D2091	E2092	D2093	T2094	A2095	E2096	A2097	Y2098	I2099	G2100	L2101	L2103	L2104	N2105	D2106	G2107	Q2108	Y2109	Y2110	F2111	N2112	D2113	D2114	G2115	I2116	N2117	Q2118	W2119	G2120	F2121	V2122	T2123	I2124			
N2125	D2126	K2127	V2128	F2129	Y2130	F2131	S2132	D2133	S2134	G2135	I2136	I2137	E2138	S2139	G2140	Y2141	Q2142	N2143	I2144	D2145	D2146	N2147	Y2148	F2149	Y2150	I2151	D2152	D2153	N2154	G2155	I2156	Y2157	Q2158	I2159	F2162	D2163	T2164	S2165	D2166	K2169	Y2170	F2171	A2172	P2173	A2174	N2175	T2176	Y2177	N2178	D2179	N2180	I2181	Y2182	G2183	Q2184	A2185	V2186
E2187	Y2188	S2189	R2193	V2194	G2195	E2196	D2197	V2198	Y2199	Y2200	F2201	G2202	E2203	T2204	Y2205	T2206	I2207	E2208	T2209	G2210	W2211	D2214	M2215	E2216	N2217	E2218	S2219	D2220	K2221	F2224	N2225	P2226	E2227	T2228	K2229	K2230	A2231	C2232	N2236	D2239	D2240	Y2244	E2247	K2248	G2254	L2255	I2256	E2259	N2260	N2261							
F2265	N2266	E2267	E2270	F2273	G2274	N2277	I2278	E2279	D2280	K2281	N2282	E2287	Q2282	V2295	F2296	N2297	T2298	P2299	D2300	G2301	F2302	K2303	Y2304	F2305	A2306	H2307	Q2308	L2311	D2312	E2313	N2314	S2319	I2320	N2321	Y2322	T2323	G2324	W2325	L2326	D2327	L2328	D2329	E2330	K2331	R2332	Y2333	Y2334	F2335	N2336	D2337							
E2338	Y2339	I2340	A2341	A2342	T2343	G2344	S2345	V2346	I2347	I2348	D2349	G2350	E2351	E2352	Y2353	Y2354	F2355	D2356	P2357	D2358	T2359	A2360	Q2361	L2362	V2363	I2364	SER	GLU	GLY	TYR	ARG	PRO	HIS	ALA	GLY	LEU	ARG	GLY	SER	HIS	HIS	HIS	HIS	HIS	HIS	HIS	HIS	HIS									

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	41523	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 ( $e^-/\text{\AA}^2$ )	50.453	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	105000	Depositor
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	2.188	Depositor
Minimum map value	-0.441	Depositor
Average map value	0.015	Depositor
Map value standard deviation	0.042	Depositor
Recommended contour level	0.504	Depositor
Map size (Å)	579.6, 579.6, 579.6	wwPDB
Map dimensions	500, 500, 500	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.1592, 1.1592, 1.1592	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	0.28	0/19378	0.46	0/26233

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1037	ASP	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	19006	0	18363	280	0
2	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	19007	0	18363	280	0

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

The worst 5 of 280 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:1574:ASN:HD22	1:A:1686:SER:HB2	1.40	0.86
1:A:1035:ILE:HG23	1:A:1039:VAL:HG12	1.67	0.76
1:A:2034:ASP:O	1:A:2037:LYS:NZ	2.21	0.73
1:A:968:LEU:HB3	1:A:985:MET:HG3	1.71	0.72
1:A:686:ILE:HG23	1:A:688:PRO:HD3	1.71	0.71

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	2361/2397 (98%)	2240 (95%)	118 (5%)	3 (0%)	51 82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1036	ILE
1	A	1269	PHE
1	A	1069	VAL

### 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	2134/2163 (99%)	2096 (98%)	38 (2%)	59 79

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1185	ASP
1	A	1966	PHE
1	A	1305	THR
1	A	1648	TYR
1	A	2346	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1380	ASN
1	A	2236	ASN
1	A	1387	HIS
1	A	2261	ASN
1	A	2016	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 1 ligands modelled in this entry, 1 is 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.

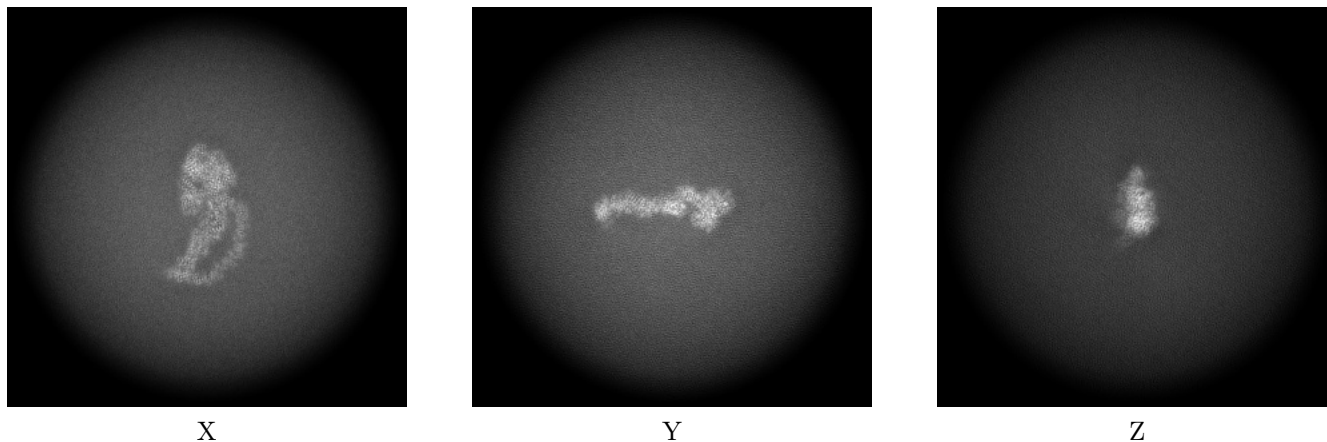
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-18373. These allow visual inspection of the internal detail of the map and identification of artifacts.

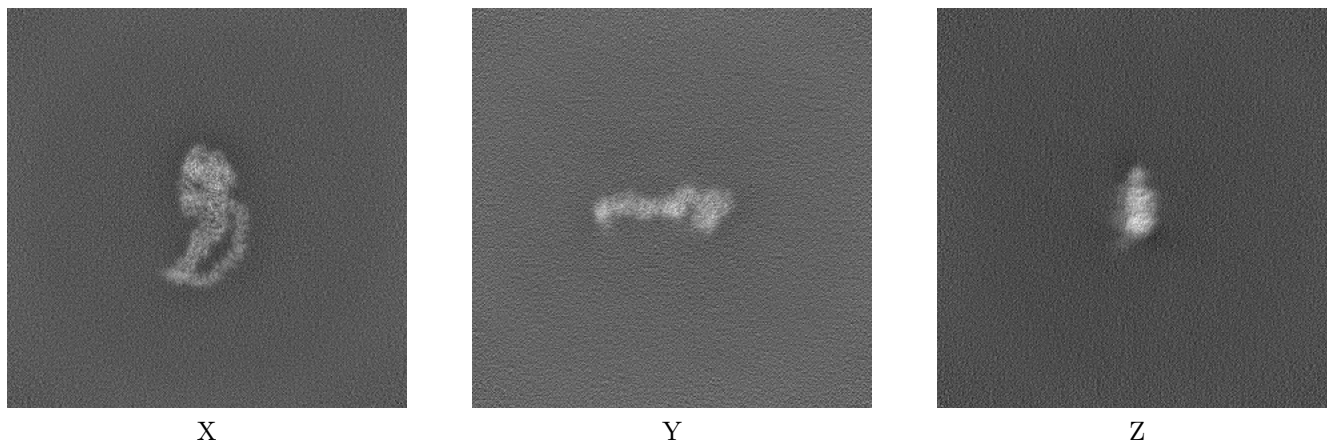
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections [i](#)

#### 6.1.1 Primary map



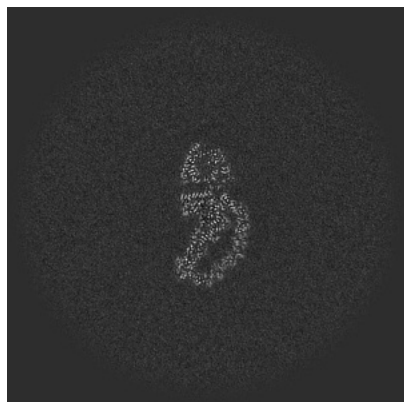
#### 6.1.2 Raw map



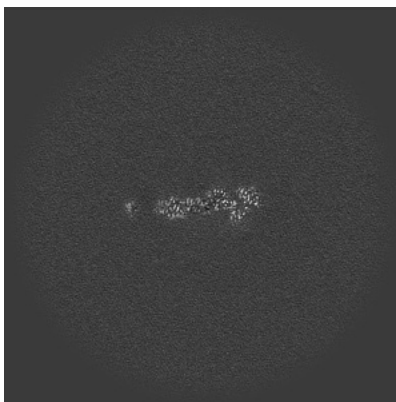
The images above show the map projected in three orthogonal directions.

## 6.2 Central slices [i](#)

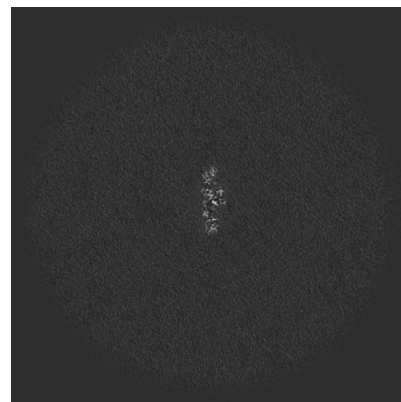
### 6.2.1 Primary map



X Index: 250

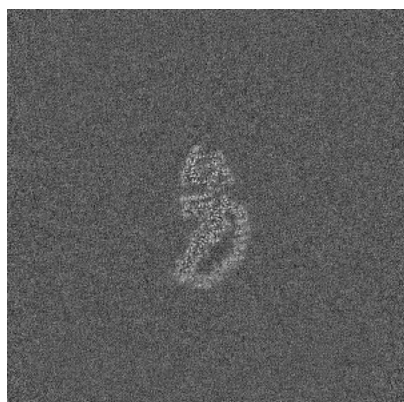


Y Index: 250

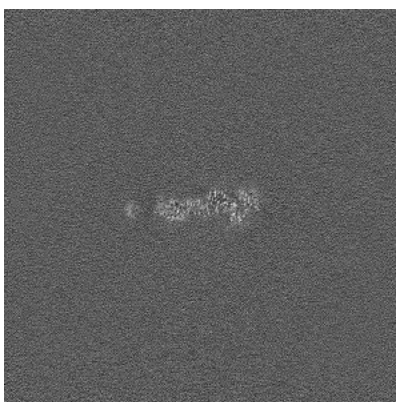


Z Index: 250

### 6.2.2 Raw map



X Index: 250



Y Index: 250

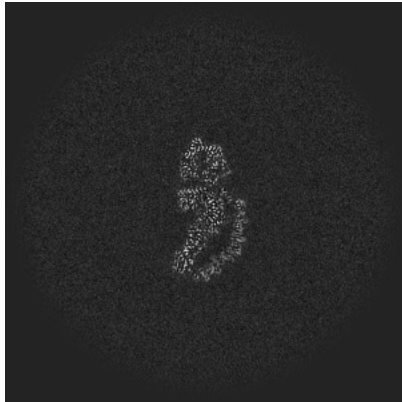


Z Index: 250

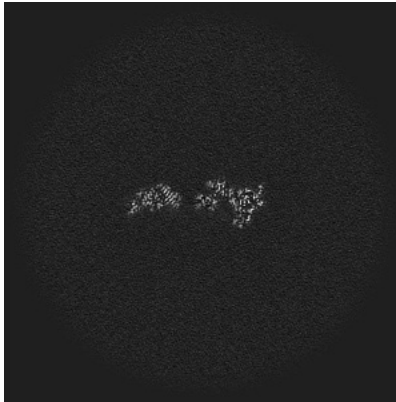
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

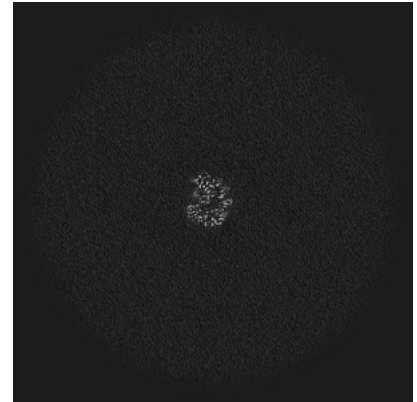
### 6.3.1 Primary map



X Index: 252

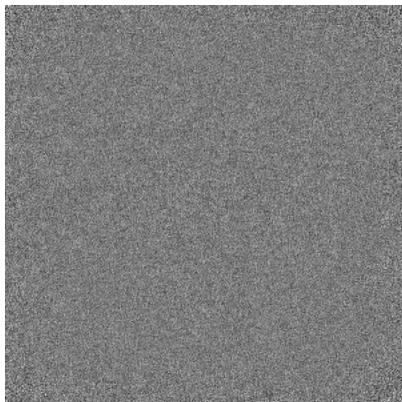


Y Index: 232

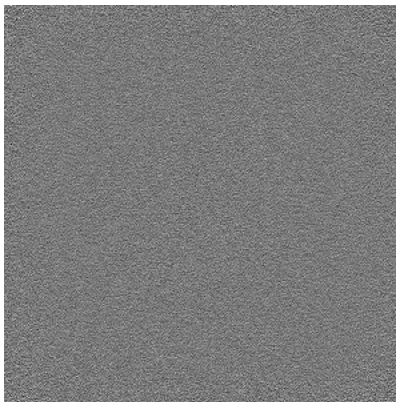


Z Index: 294

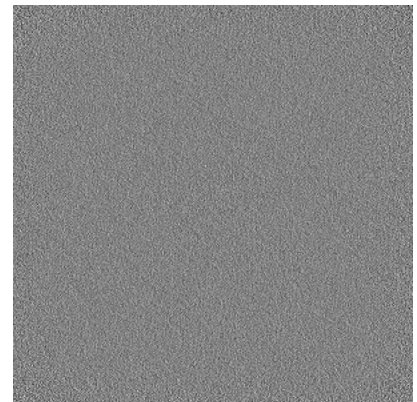
### 6.3.2 Raw map



X Index: 0



Y Index: 0

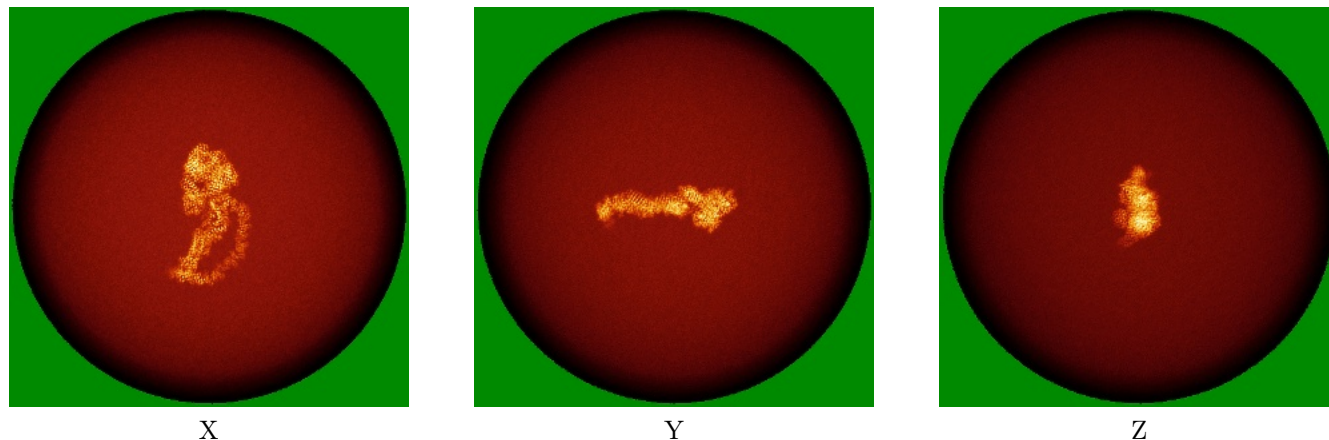


Z Index: 0

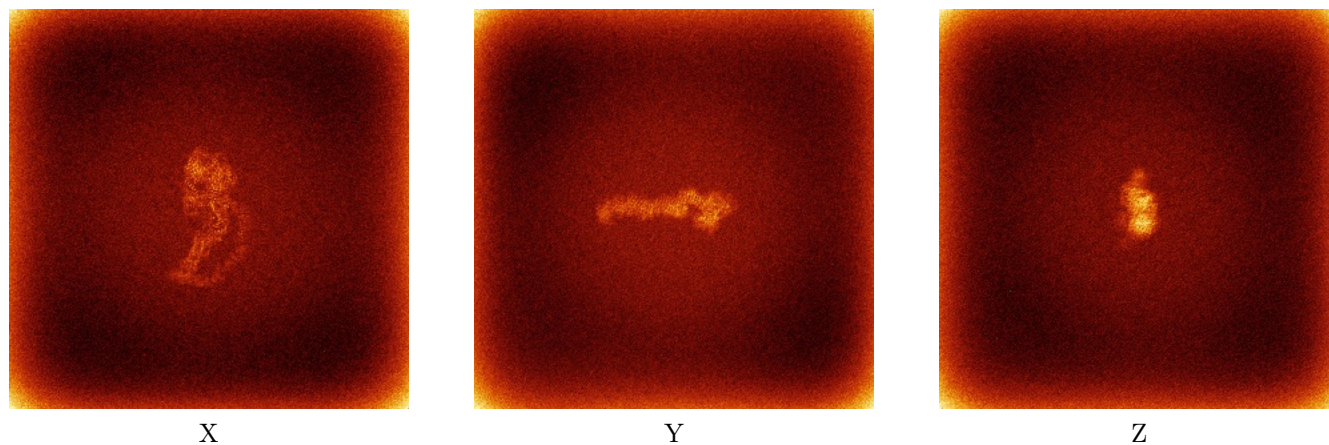
The images above show the largest variance slices of the map in three orthogonal directions.

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

### 6.4.1 Primary map



### 6.4.2 Raw map

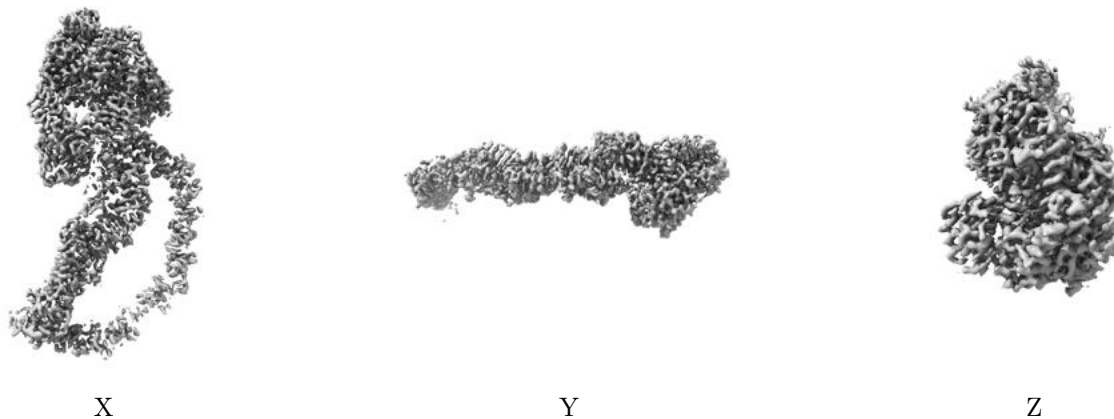


The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.



## 6.5 Orthogonal surface views [i](#)

### 6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.504. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

### 6.5.2 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.

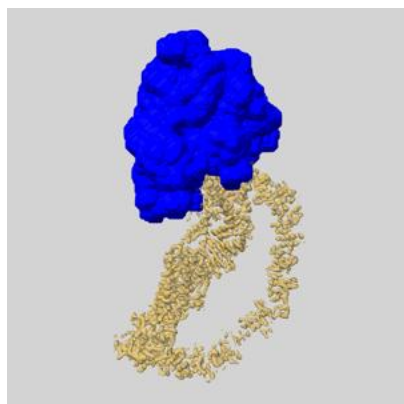
## 6.6 Mask visualisation [i](#)

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

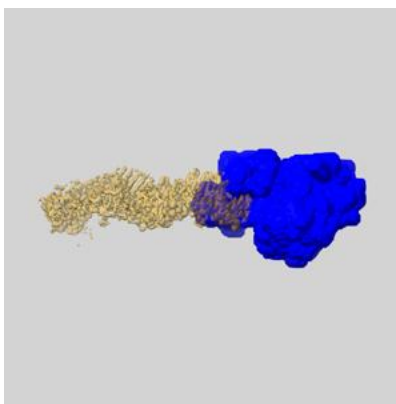
A mask typically either:

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

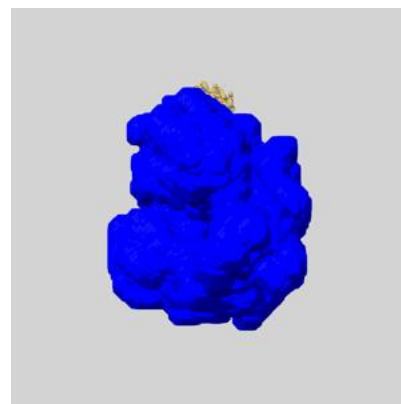
### 6.6.1 emd\_18373\_msk\_1.map [i](#)



X

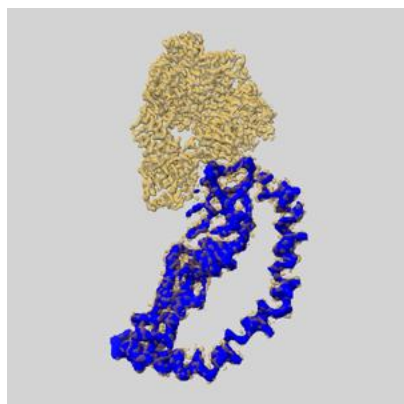


Y

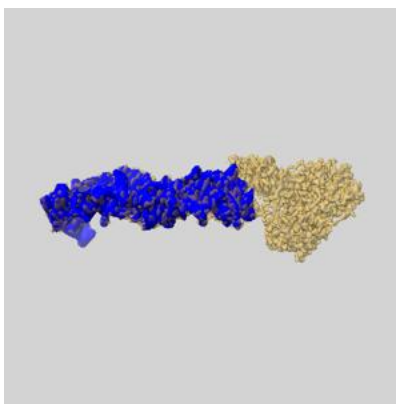


Z

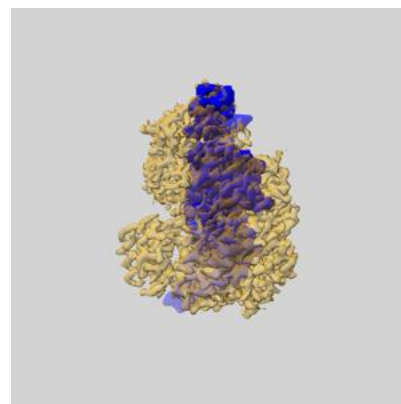
### 6.6.2 emd\_18373\_msk\_2.map [i](#)



X



Y

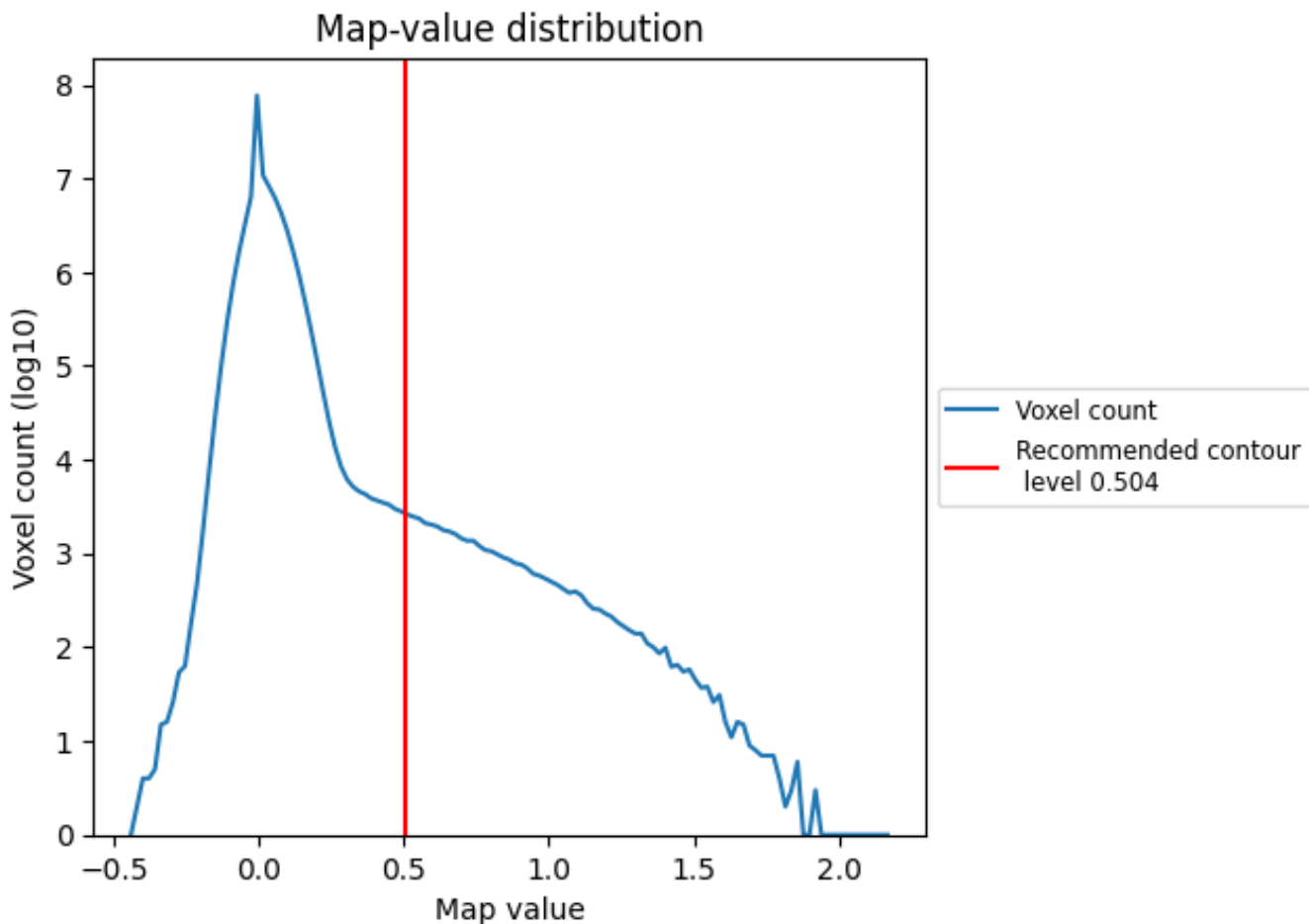


Z

## 7 Map analysis [i](#)

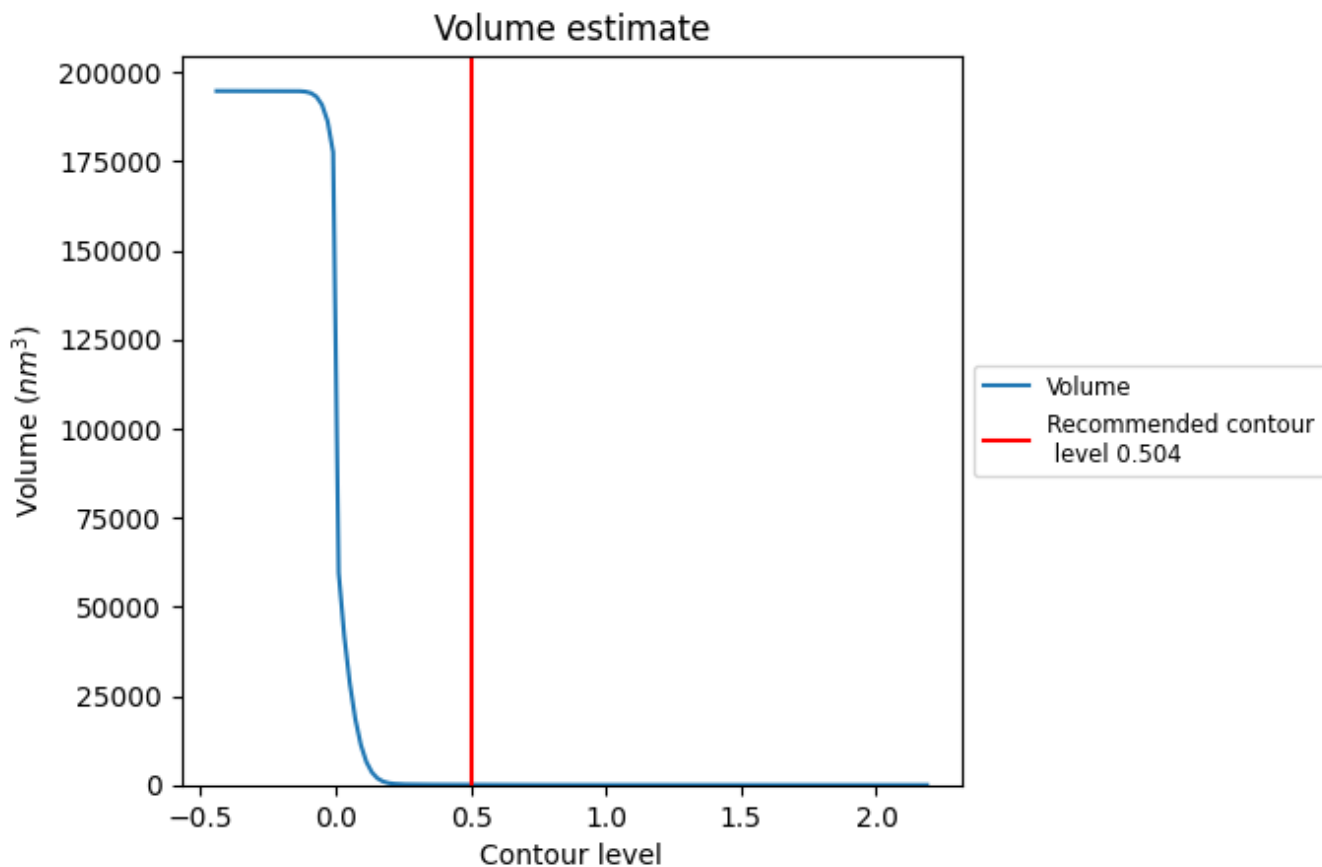
This section contains the results of statistical analysis of the map.

### 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.

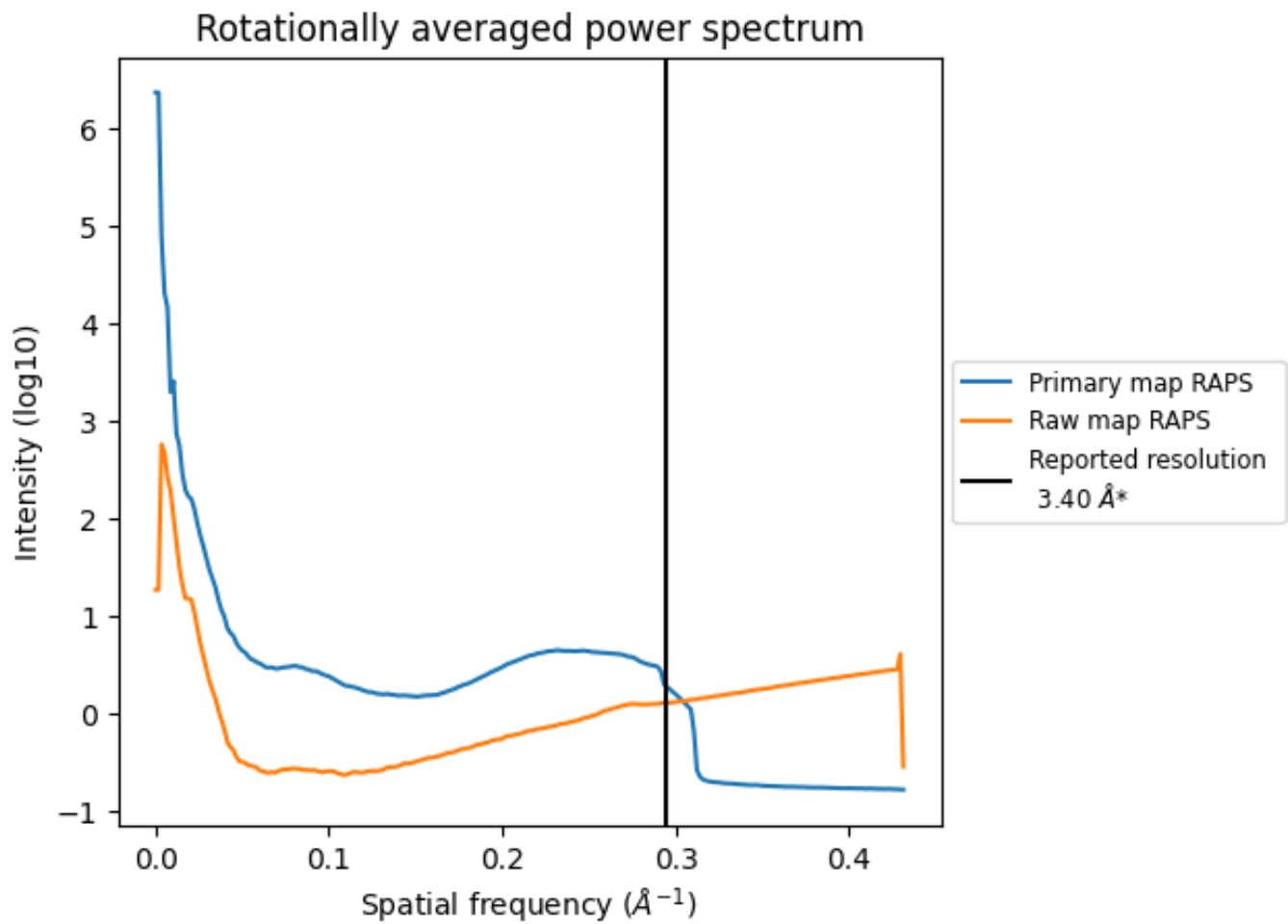
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 61  $\text{nm}^3$ ; this corresponds to an approximate mass of 55 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

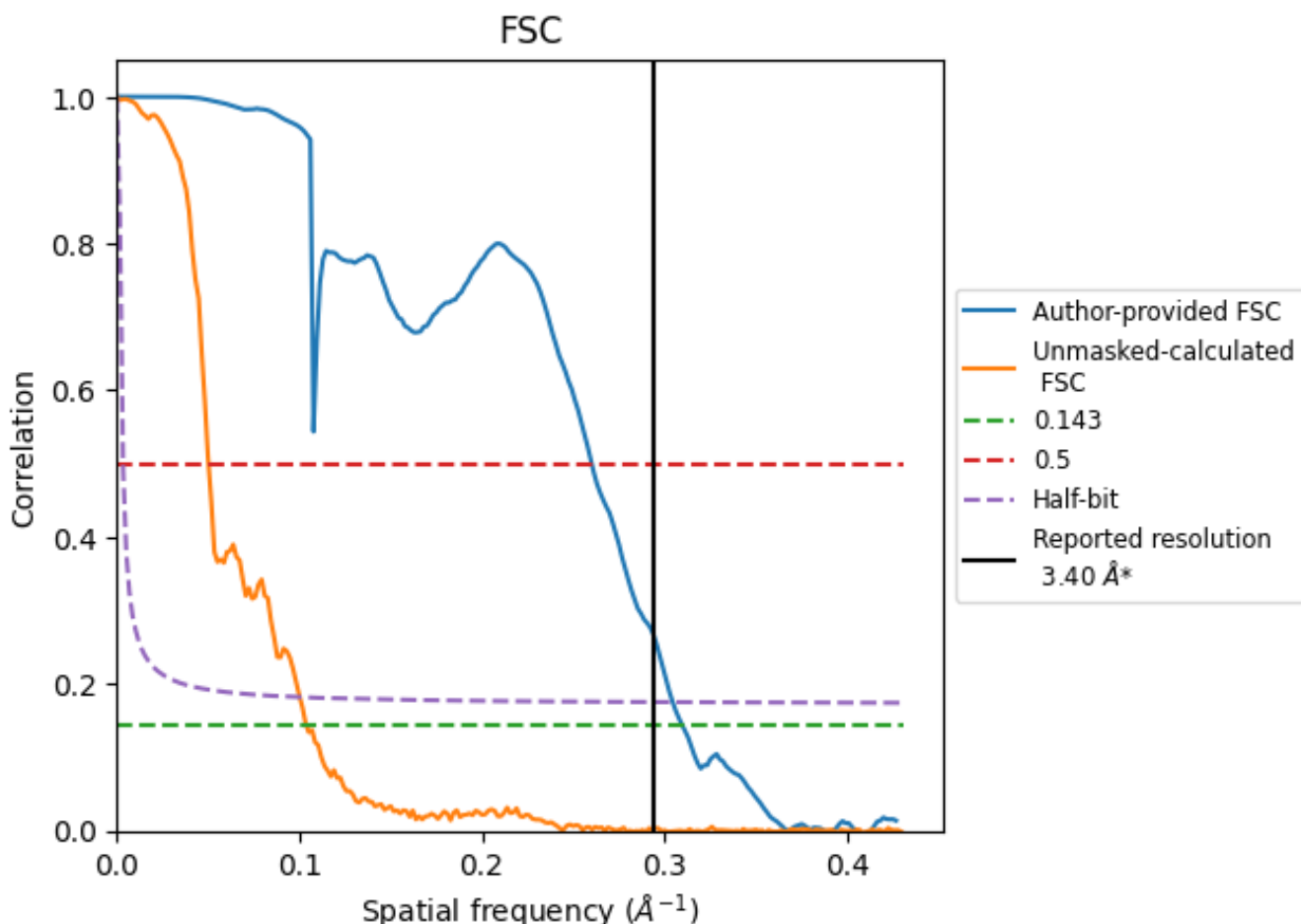


\*Reported resolution corresponds to spatial frequency of 0.294 Å<sup>-1</sup>

## 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 0.294 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

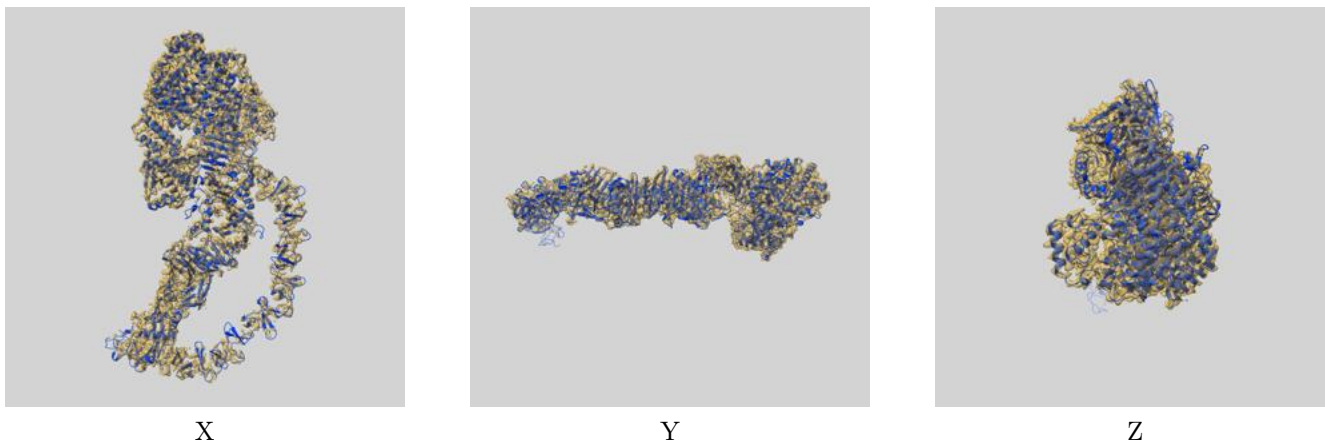
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.40	-	-
Author-provided FSC curve	3.23	3.84	3.28
Unmasked-calculated*	9.62	19.88	9.98

\*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 9.62 differs from the reported value 3.4 by more than 10 %

## 9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-18373 and PDB model 8QEN. Per-residue inclusion information can be found in section 3 on page 5.

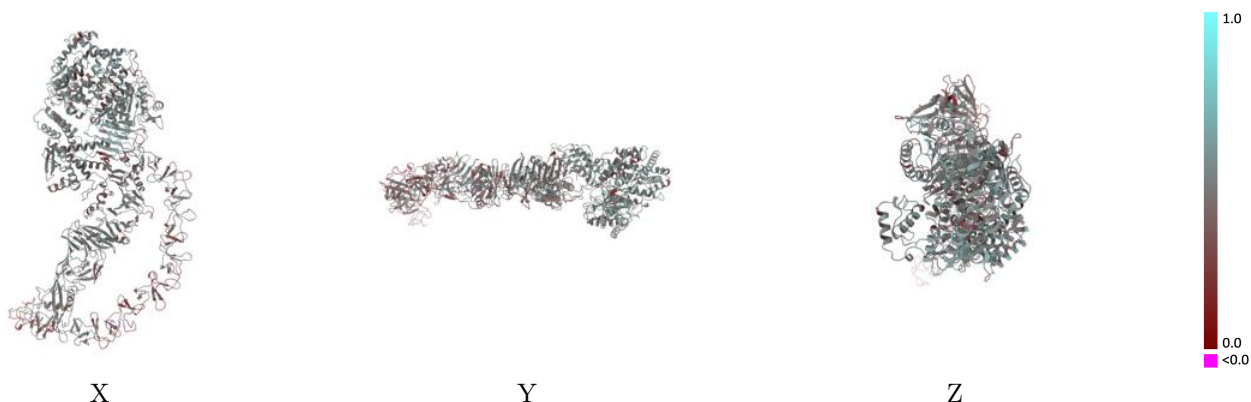
### 9.1 Map-model overlay [i](#)



The images above show the 3D surface view of the map at the recommended contour level 0.504 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.

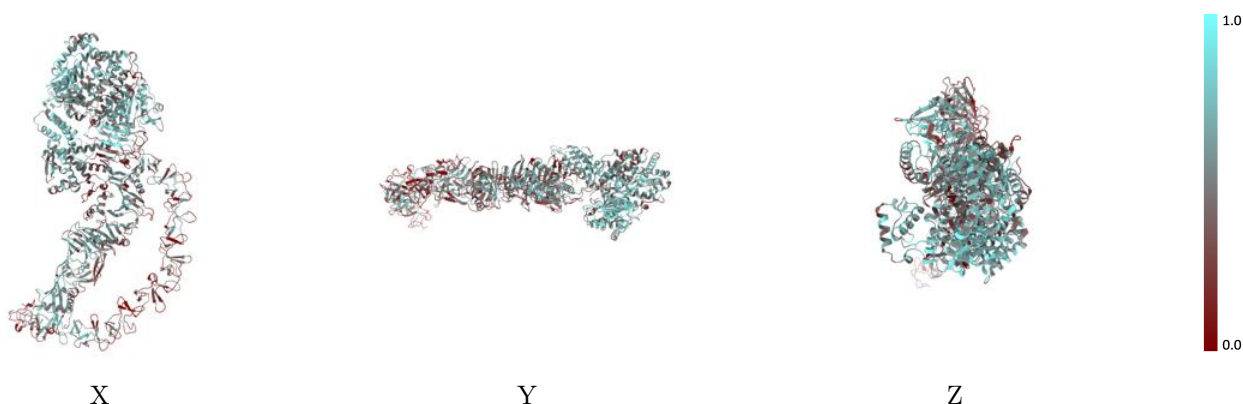


## 9.2 Q-score mapped to coordinate model [\(i\)](#)



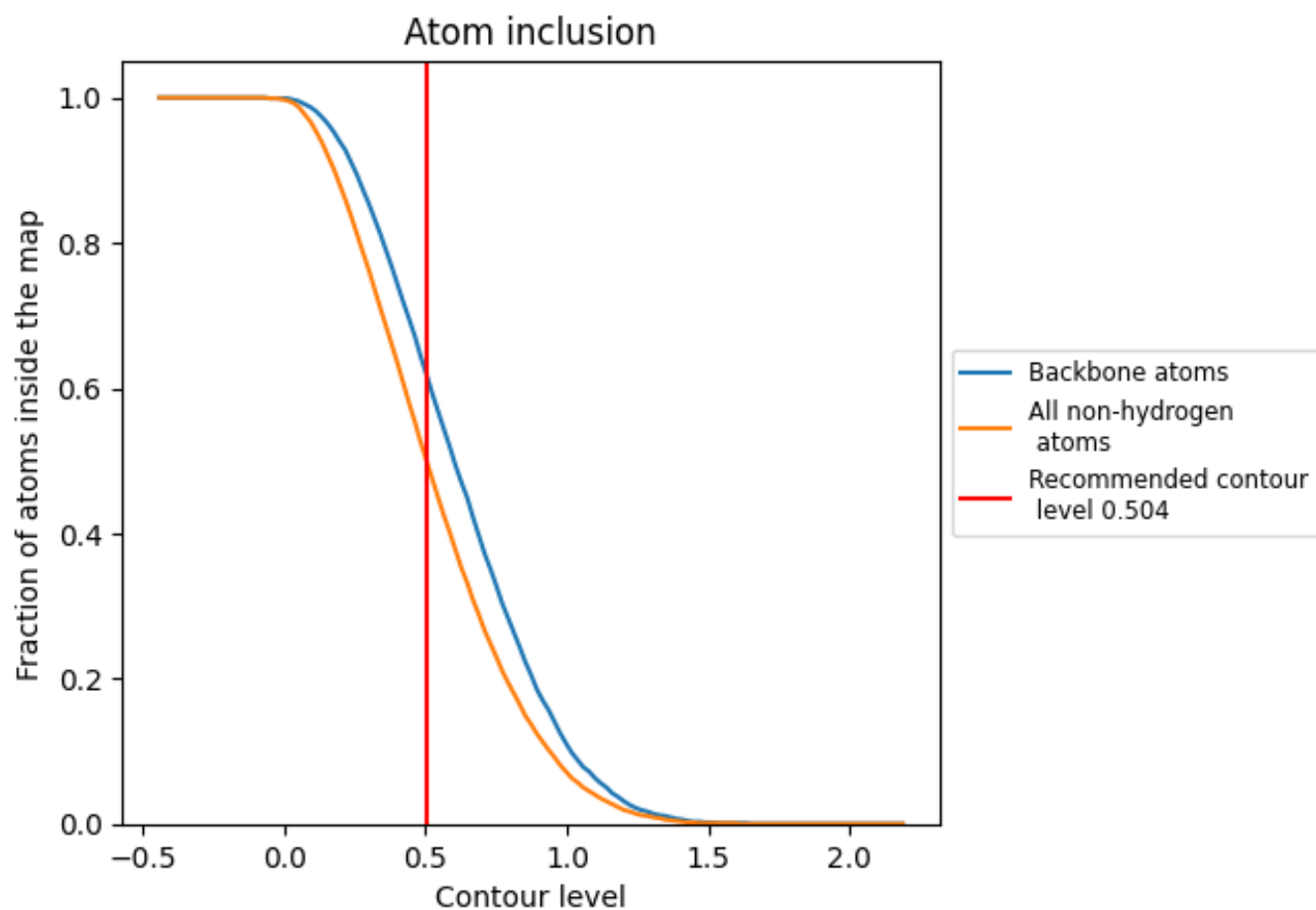
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.504).





## 9.4 Atom inclusion [i](#)



At the recommended contour level, 62% of all backbone atoms, 50% 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 (0.504) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.4990	 0.4520
A	 0.4990	 0.4520

