



wwPDB EM Validation Summary Report ⓘ

Nov 8, 2022 – 02:44 PM JST

PDB ID : 7W5Y
EMDB ID : EMD-32324
Title : Cryo-EM structure of SoxS-dependent transcription activation complex with
fpr promoter DNA
Authors : Lin, W.; Feng, Y.
Deposited on : 2021-11-30
Resolution : 4.20 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:

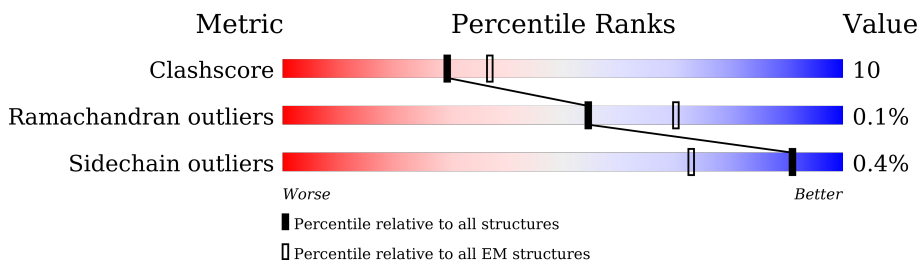
EMDB validation analysis : 0.0.1.dev43
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : 1.9.9
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.31.2

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 4.20 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	329	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">13%</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 13%, orange 13%, yellow 21%, green 70%, grey 9%);"></div> <div style="text-align: center;">70%</div> <div style="text-align: center;">21%</div> <div style="text-align: center;">9%</div> </div>
1	B	329	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">20%</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 20%, orange 20%, yellow 22%, green 72%, grey 6%);"></div> <div style="text-align: center;">72%</div> <div style="text-align: center;">22%</div> <div style="text-align: center;">6%</div> </div>
2	C	1342	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">•</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 0%, orange 0%, yellow 19%, green 80%, grey 0%);"></div> <div style="text-align: center;">80%</div> <div style="text-align: center;">19%</div> <div style="text-align: center;">•</div> </div>
3	D	1407	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">•</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 0%, orange 0%, yellow 21%, green 74%, grey 5%);"></div> <div style="text-align: center;">74%</div> <div style="text-align: center;">21%</div> <div style="text-align: center;">5%</div> </div>
4	E	91	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">34%</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 34%, orange 34%, yellow 23%, green 64%, grey 13%);"></div> <div style="text-align: center;">64%</div> <div style="text-align: center;">23%</div> <div style="text-align: center;">13%</div> </div>
5	F	613	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">•</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 0%, orange 0%, yellow 14%, green 67%, grey 19%);"></div> <div style="text-align: center;">67%</div> <div style="text-align: center;">14%</div> <div style="text-align: center;">19%</div> </div>
6	1	86	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">26%</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 26%, orange 26%, yellow 49%, green 48%, grey 0%);"></div> <div style="text-align: center;">48%</div> <div style="text-align: center;">49%</div> <div style="text-align: center;">•</div> </div>
7	2	86	<div style="display: flex; justify-content: space-between; align-items: center;"> <div style="text-align: center;">24%</div> <div style="width: 100%; height: 15px; background: linear-gradient(to right, red 24%, orange 24%, yellow 43%, green 53%, grey 0%);"></div> <div style="text-align: center;">53%</div> <div style="text-align: center;">43%</div> <div style="text-align: center;">•</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	K	107	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a green segment on the left, a red segment in the middle, and a yellow segment on the right. Above the bar, the text '93%' is centered. Below the bar, '79%' is positioned under the green segment and '20%' is positioned under the yellow segment. A small black dot is located at the far right end of the bar.</p>

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 34577 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-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	300	Total	C	N	O	S	0	0
			2334	1459	412	455	8		
1	B	308	Total	C	N	O	S	0	0
			2380	1485	418	469	8		

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	1338	Total	C	N	O	S	0	0
			10541	6616	1838	2044	43		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	516	VAL	ASP	variant	UNP P0A8V2

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	1335	Total	C	N	O	S	0	0
			10381	6520	1851	1960	50		

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	79	Total	C	N	O	S	0	0
			627	382	118	126	1		

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	496	Total	C	N	O	S	0	0
			3986	2488	713	762	23		

- Molecule 6 is a DNA chain called fpr promoter DNA forward strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	1	83	1709	815	304	507	83	0	0

- Molecule 7 is a DNA chain called fpr promoter DNA reverse strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	2	83	1707	811	326	487	83	0	0

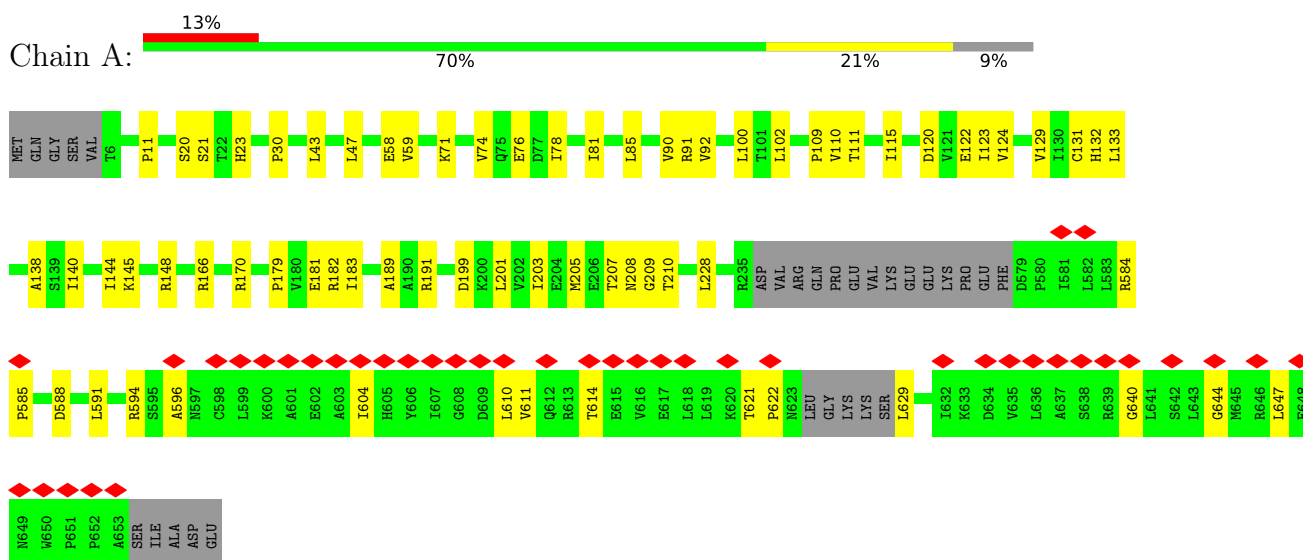
- Molecule 8 is a protein called Regulatory protein SoxS.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	K	107	912	579	170	160	3	0	0

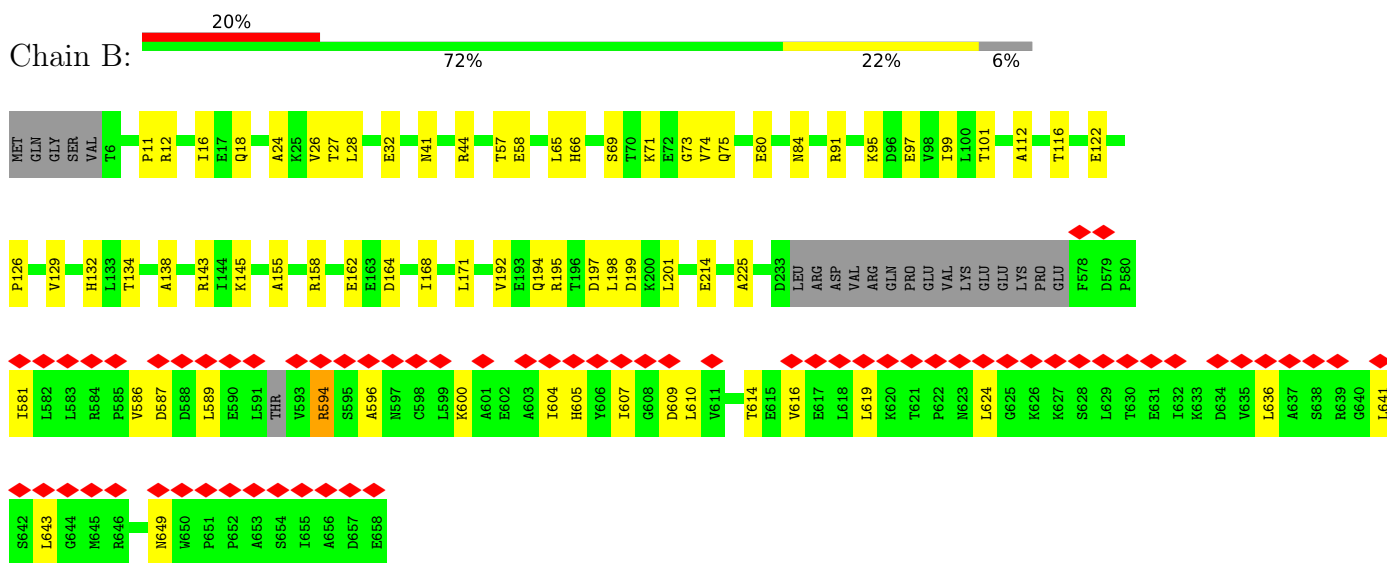
3 Residue-property plots

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-directed RNA polymerase subunit alpha

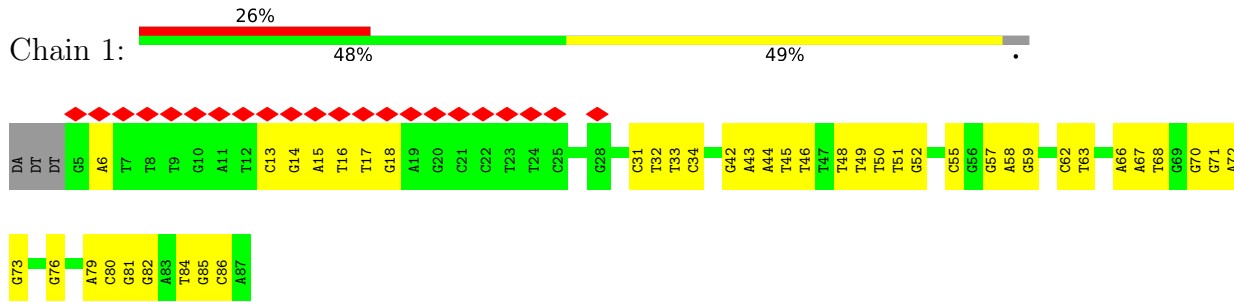


- Molecule 1: DNA-directed RNA polymerase subunit alpha

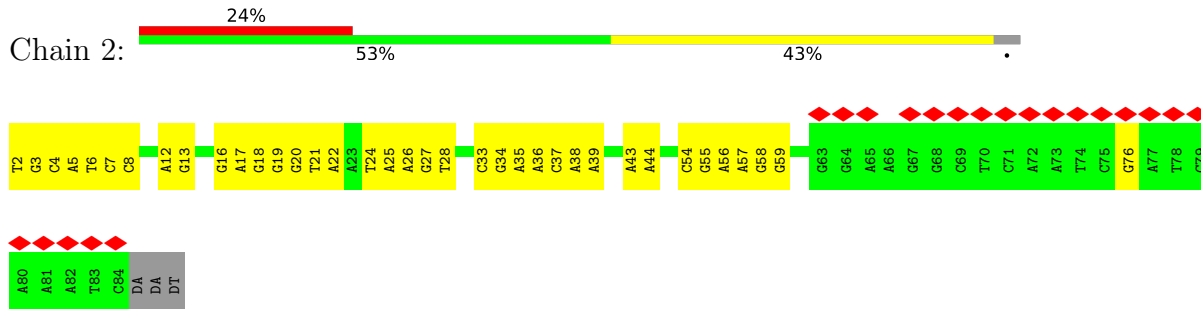


- Molecule 2: DNA-directed RNA polymerase subunit beta

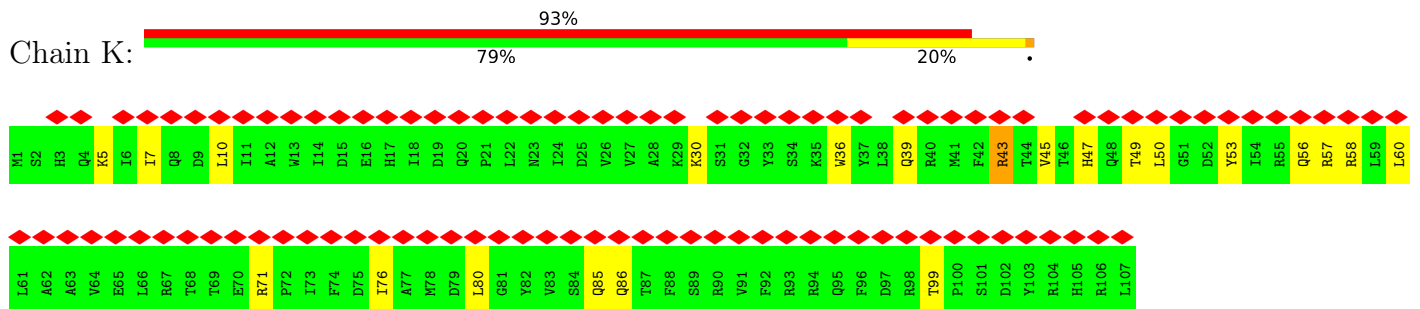
• Molecule 6: fpr promoter DNA forward strand



• Molecule 7: fpr promoter DNA reverse strand



• Molecule 8: Regulatory protein SoxS



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	74601	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	52	Depositor
Minimum defocus (nm)	1400	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.088	Depositor
Minimum map value	-0.032	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.012	Depositor
Map size (\AA)	287.54, 287.54, 287.54	wwPDB
Map dimensions	220, 220, 220	wwPDB
Map angles ($^\circ$)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (\AA)	1.307, 1.307, 1.307	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	0.32	0/2363	0.58	0/3204
1	B	0.30	0/2408	0.58	0/3260
2	C	0.36	0/10707	0.55	0/14443
3	D	0.35	0/10536	0.58	0/14222
4	E	0.30	0/629	0.56	0/847
5	F	0.31	0/4037	0.56	0/5428
6	1	0.68	0/1915	0.97	0/2957
7	2	0.63	0/1919	0.91	0/2959
8	K	0.27	0/932	0.54	0/1256
All	All	0.39	0/35446	0.63	0/48576

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	2334	0	2381	44	0
1	B	2380	0	2418	46	0
2	C	10541	0	10561	173	0
3	D	10381	0	10602	177	0
4	E	627	0	634	15	0
5	F	3986	0	4020	57	0
6	1	1709	0	941	103	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	2	1707	0	931	111	0
8	K	912	0	917	17	0
All	All	34577	0	33405	677	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 10.

The worst 5 of 677 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)
6:1:80:DC:C5	6:1:81:DG:C5	2.31	1.19
7:2:55:DG:H2''	7:2:56:DA:H5'	1.12	1.12
6:1:43:DA:C6	6:1:44:DA:N6	2.25	1.05
7:2:38:DA:N7	7:2:39:DA:C5	2.25	1.04
6:1:80:DC:C5	6:1:81:DG:C6	2.47	1.03

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	294/329 (89%)	271 (92%)	23 (8%)	0	100	100
1	B	302/329 (92%)	275 (91%)	27 (9%)	0	100	100
2	C	1332/1342 (99%)	1212 (91%)	120 (9%)	0	100	100
3	D	1327/1407 (94%)	1196 (90%)	131 (10%)	0	100	100
4	E	77/91 (85%)	72 (94%)	5 (6%)	0	100	100
5	F	490/613 (80%)	443 (90%)	45 (9%)	2 (0%)	34	72
8	K	105/107 (98%)	97 (92%)	8 (8%)	0	100	100
All	All	3927/4218 (93%)	3566 (91%)	359 (9%)	2 (0%)	54	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	F	507	MET
5	F	325	PRO

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	260/286 (91%)	260 (100%)	0	100	100
1	B	264/286 (92%)	262 (99%)	2 (1%)	81	89
2	C	1150/1157 (99%)	1148 (100%)	2 (0%)	93	96
3	D	1118/1168 (96%)	1110 (99%)	8 (1%)	84	90
4	E	67/75 (89%)	67 (100%)	0	100	100
5	F	430/540 (80%)	429 (100%)	1 (0%)	93	96
8	K	99/99 (100%)	98 (99%)	1 (1%)	76	86
All	All	3388/3611 (94%)	3374 (100%)	14 (0%)	91	94

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	639	VAL
3	D	849	LEU
8	K	43	ARG
3	D	1093	THR
5	F	562	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	128	ASN
8	K	85	GLN
5	F	242	HIS
5	F	461	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	606	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.

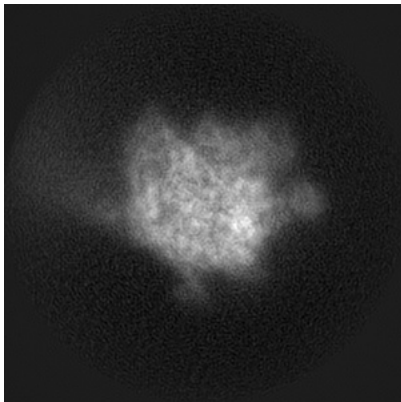
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-32324. These allow visual inspection of the internal detail of the map and identification of artifacts.

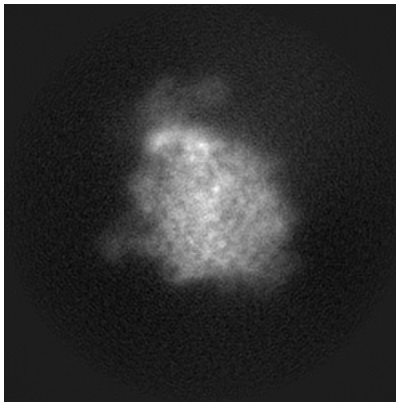
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

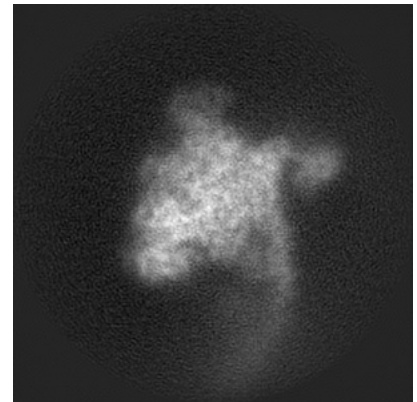
6.1.1 Primary map



X



Y

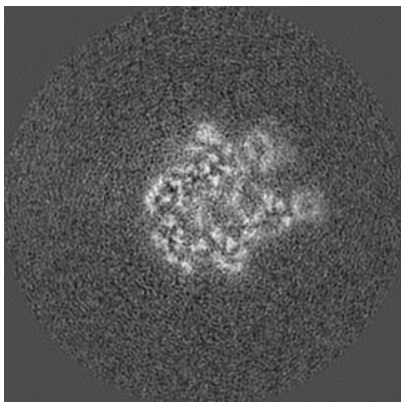


Z

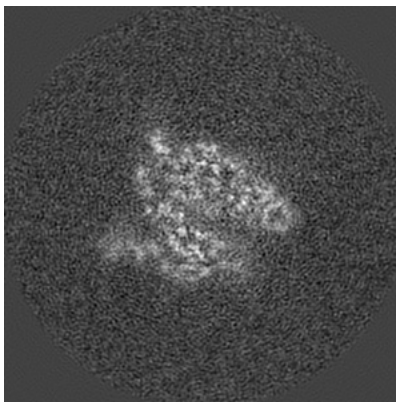
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

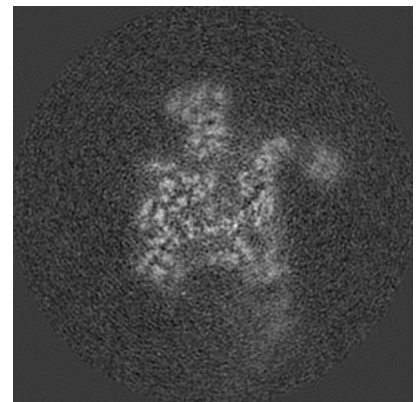
6.2.1 Primary map



X Index: 110



Y Index: 110

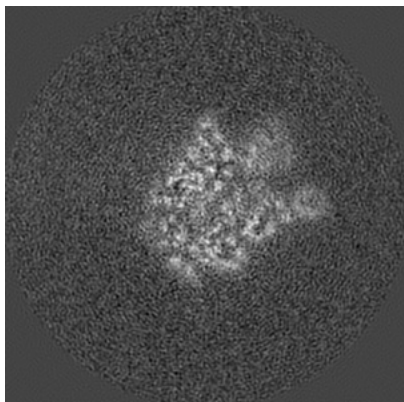


Z Index: 110

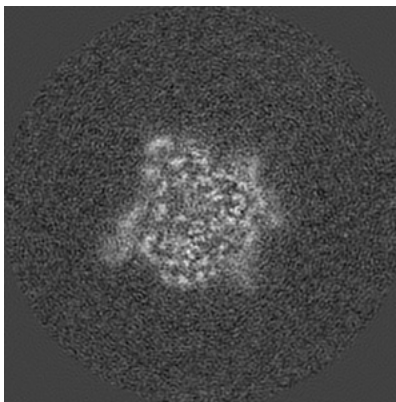
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

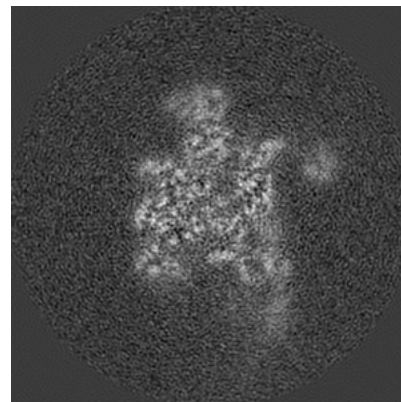
6.3.1 Primary map



X Index: 107



Y Index: 103

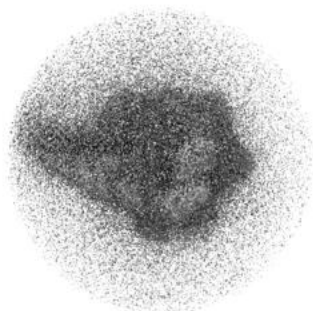


Z Index: 108

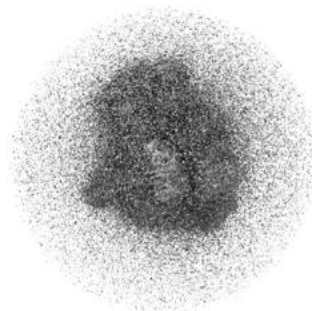
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views [i](#)

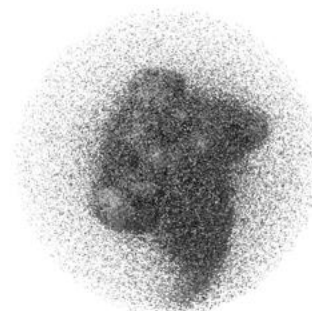
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.012. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

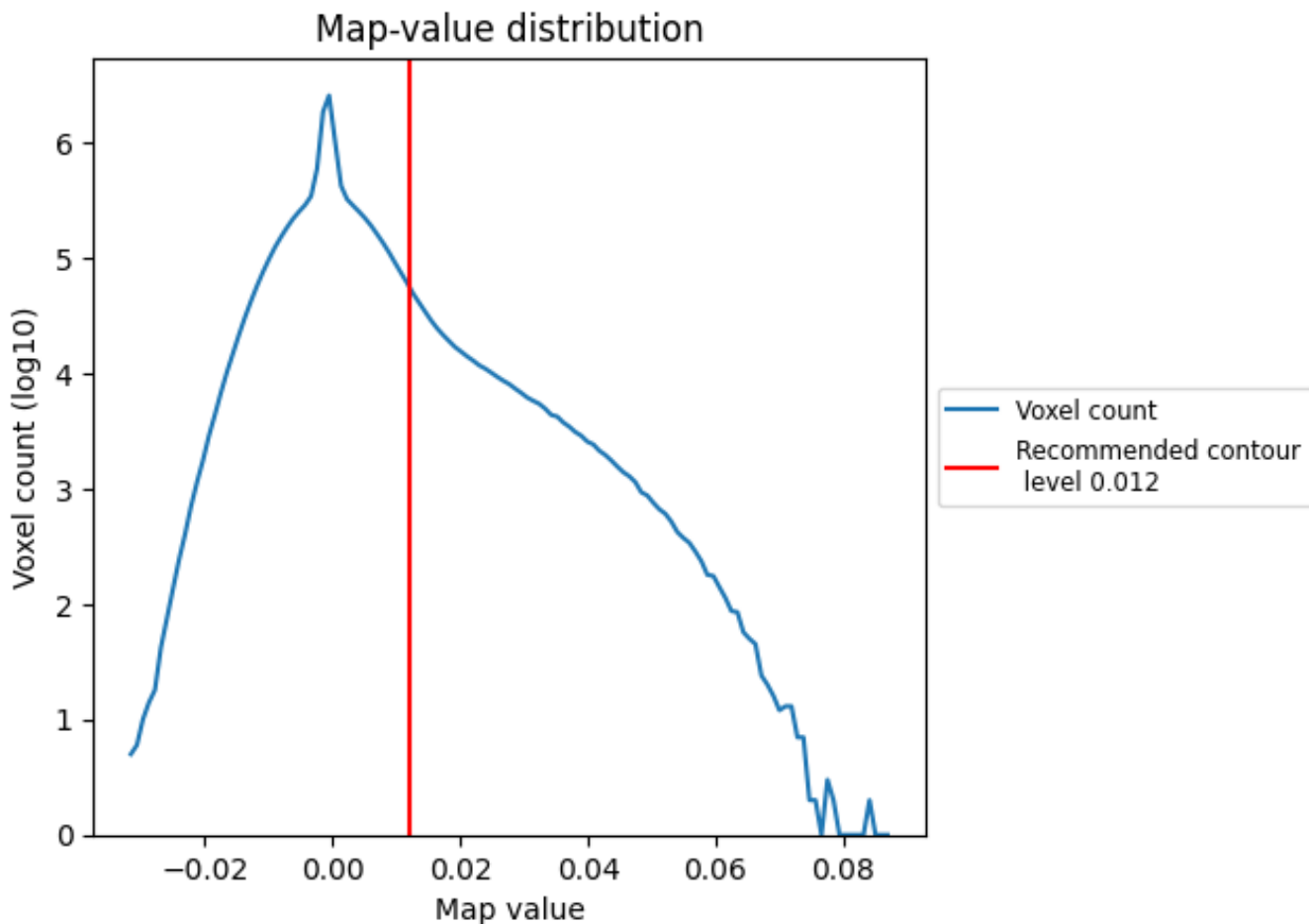
6.5 Mask visualisation

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

7 Map analysis [i](#)

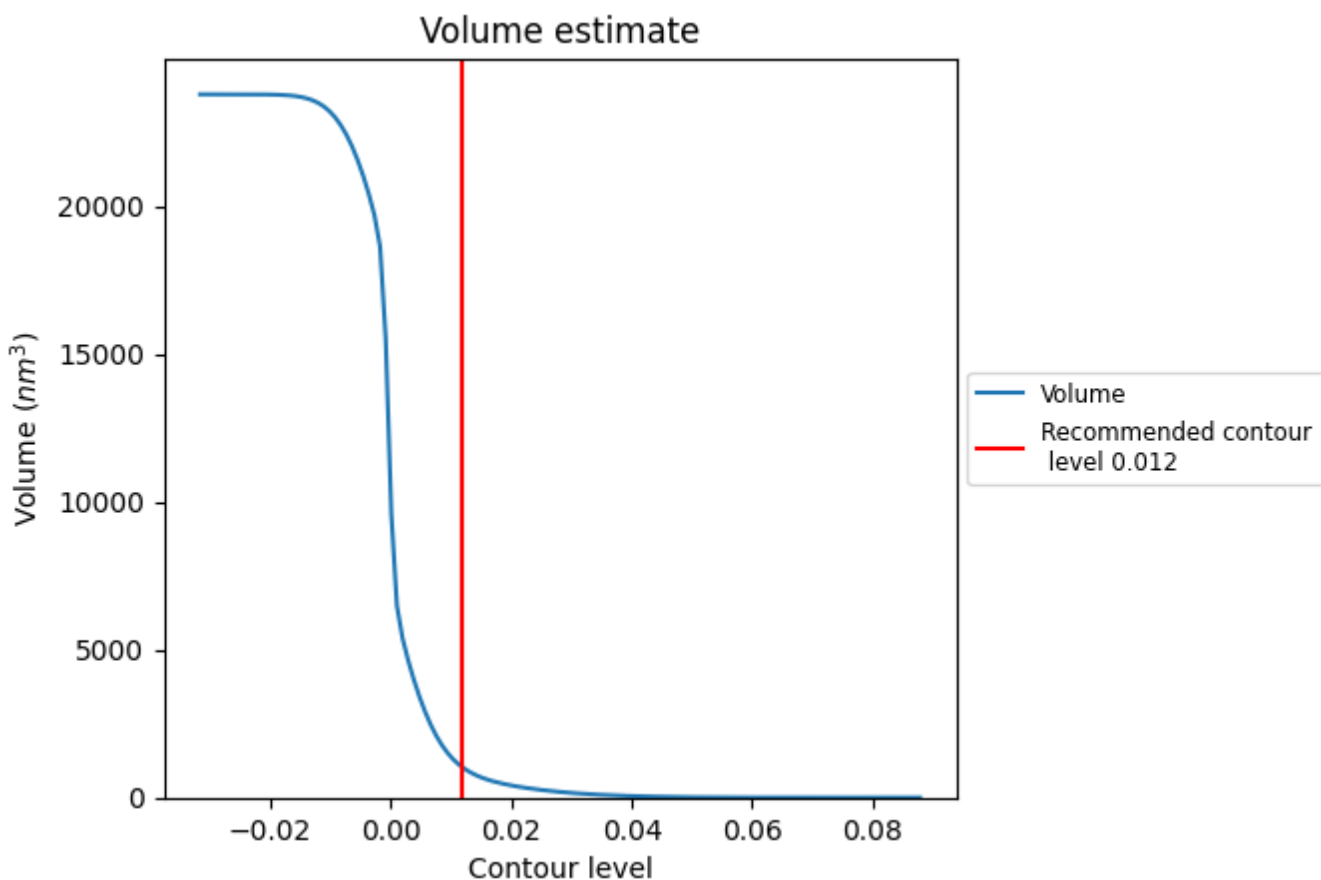
This section contains the results of statistical analysis of the map.

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.

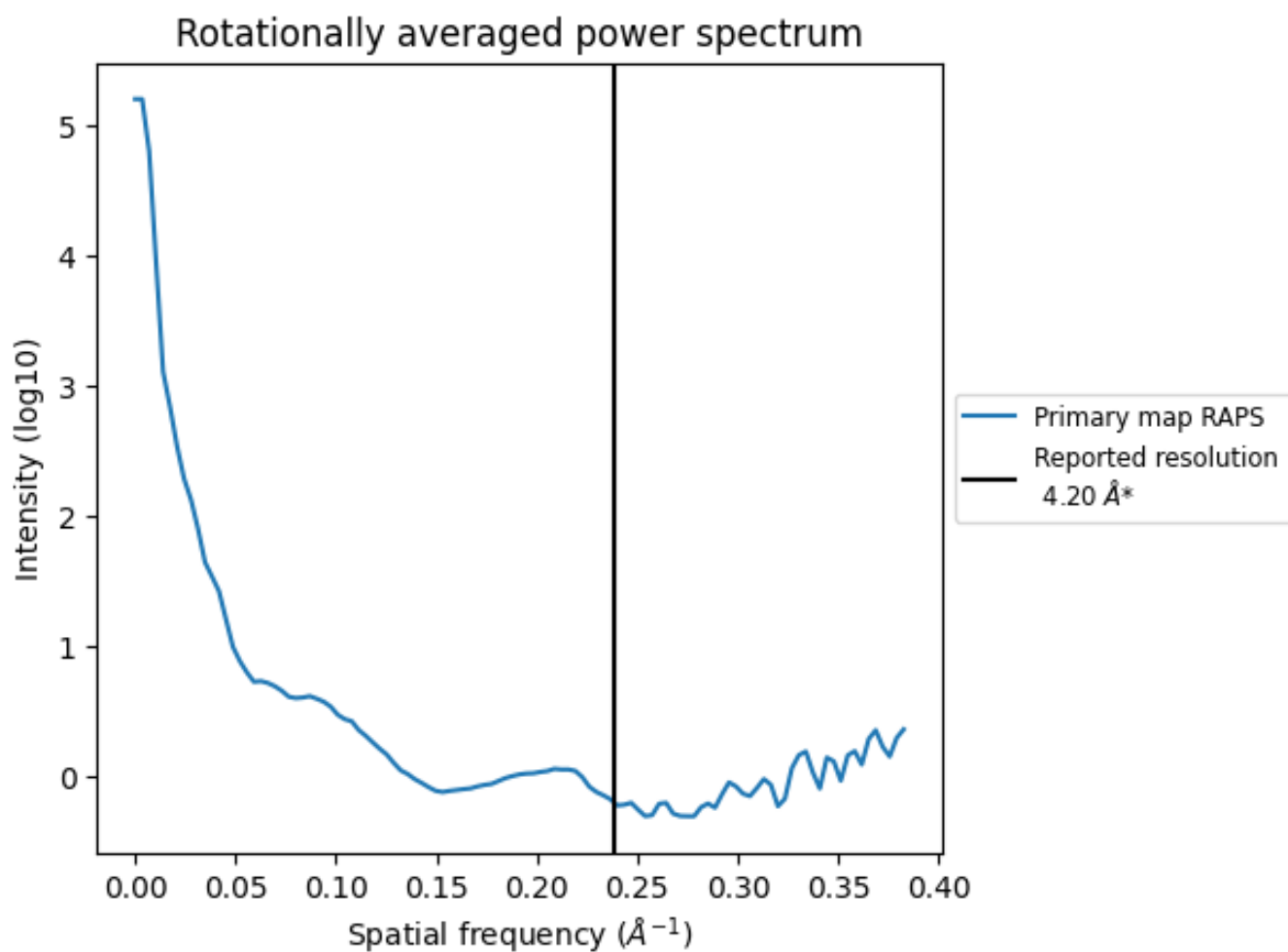
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 1020 nm^3 ; this corresponds to an approximate mass of 922 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 0.238\AA^{-1}

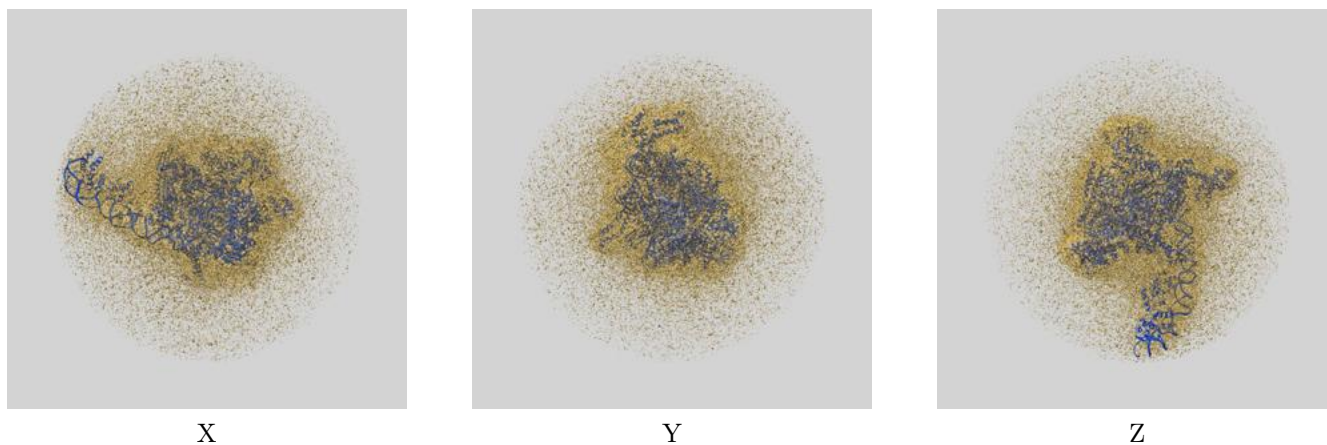
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

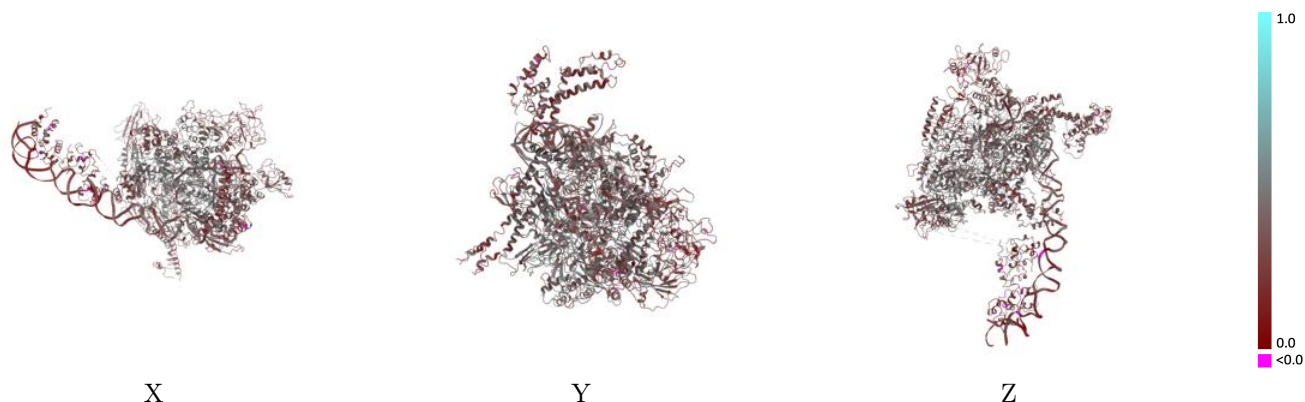
This section contains information regarding the fit between EMDB map EMD-32324 and PDB model 7W5Y. Per-residue inclusion information can be found in section [3](#) on page [6](#).

9.1 Map-model overlay [i](#)



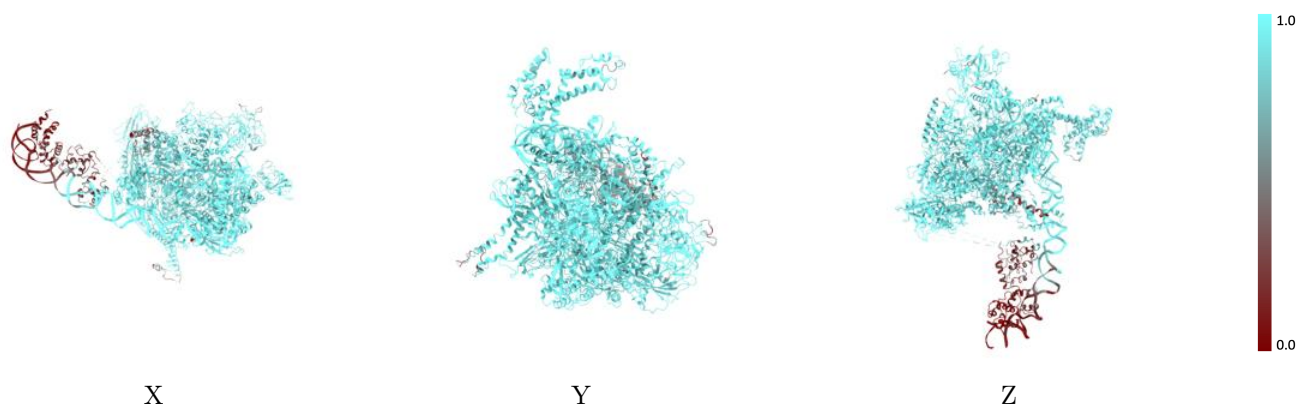
The images above show the 3D surface view of the map at the recommended contour level 0.012 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



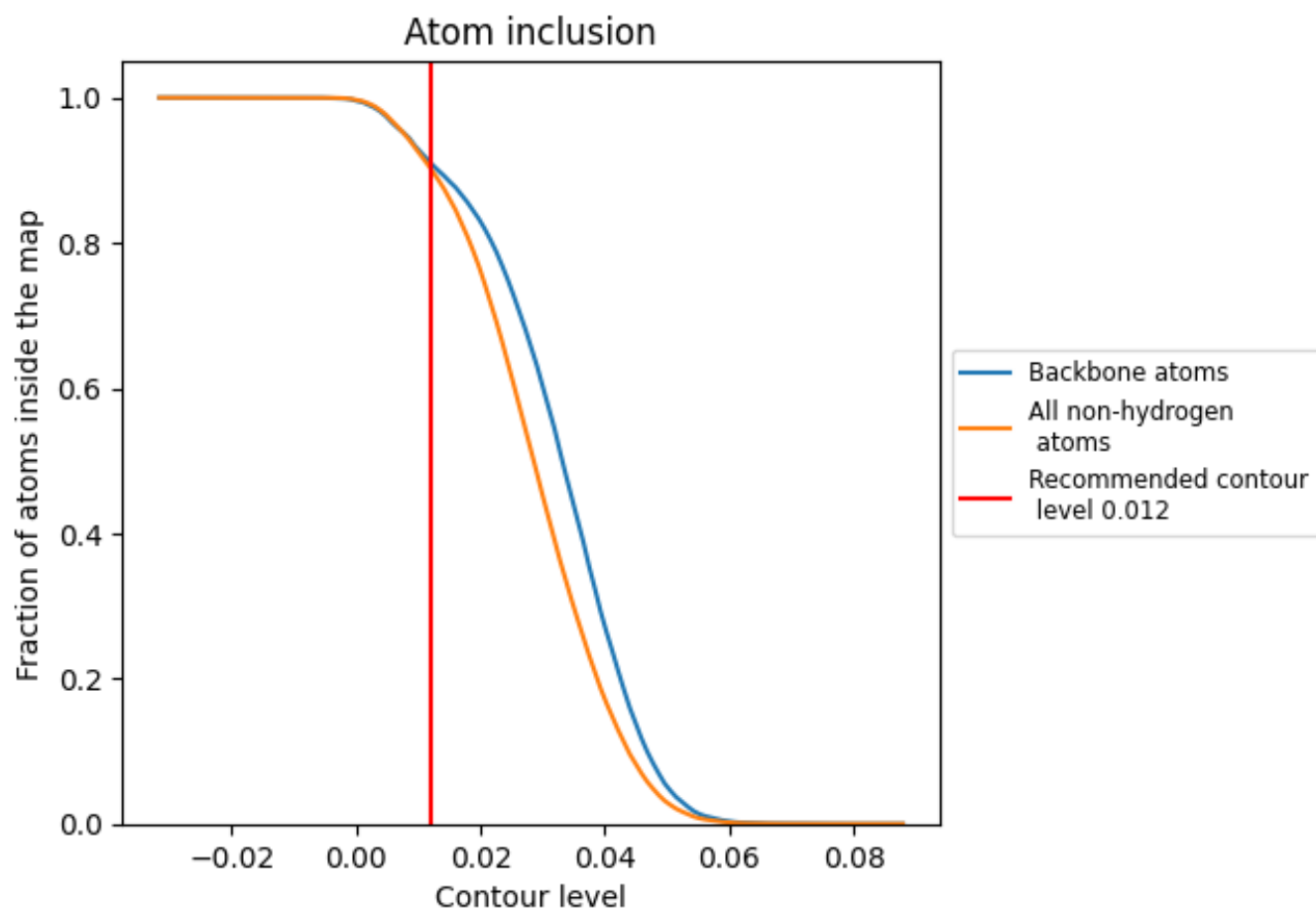
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.012).





















9.4 Atom inclusion [i](#)



At the recommended contour level, 91% of all backbone atoms, 90% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.012) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.9026	 0.3690
1	 0.7408	 0.2890
2	 0.7387	 0.2770
A	 0.8516	 0.3870
B	 0.7998	 0.3470
C	 0.9817	 0.4100
D	 0.9790	 0.3810
E	 0.5188	 0.2790
F	 0.9647	 0.3450
K	 0.1251	 0.2660

