wwPDB EM Validation Summary Report (i)

Nov 29, 2022 – 11:20 AM JST

PDB ID	:	7WB0
EMDB ID	:	EMD-32391
Title	:	PlmCasX-sgRNAv1-dsDNA ternary complex at nts loading state with flexible H2 domain
Authors	:	Zhang, S.; Liu, J.J.G.
Deposited on	:	2021-12-15
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 (i) 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 (i)) were used in the production of this report:

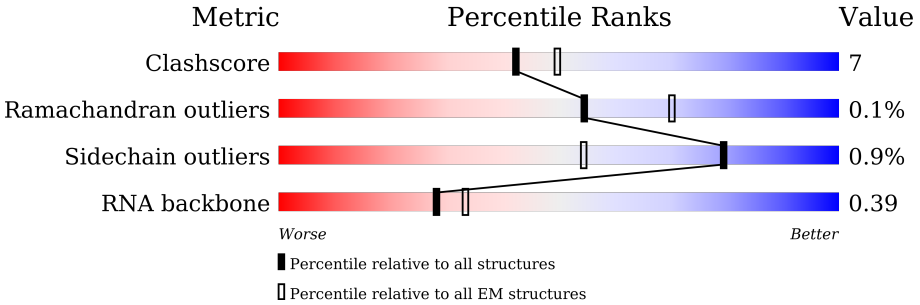
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

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
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	B	40	25%; 25%, 58%, 18%
2	C	40	35%; 50%, 18%, 32%
3	D	122	32%; 46%, 40%, 7%, • , 6%
4	A	978	26%; 67%, 14%, 19%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10096 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 DNA chain called TS-DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	B	33	674	323	115	203	33	0	0

- Molecule 2 is a DNA chain called NTS-DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	C	27	547	260	97	163	27	0	0

- Molecule 3 is a RNA chain called RNA (115-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	D	115	2446	1095	437	799	115	0	0

- Molecule 4 is a protein called dPlmCasX.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	797	6429	4087	1155	1163	24	0	0

There are 3 discrepancies between the modelled and reference sequences:

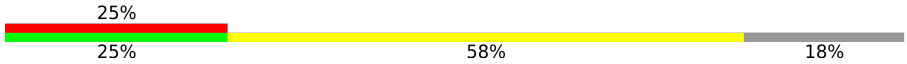
Chain	Residue	Modelled	Actual	Comment	Reference
A	659	ALA	ASP	engineered mutation	UNP A0A1G3BXR9
A	756	ALA	GLU	engineered mutation	UNP A0A1G3BXR9
A	922	ALA	ASP	engineered mutation	UNP A0A1G3BXR9

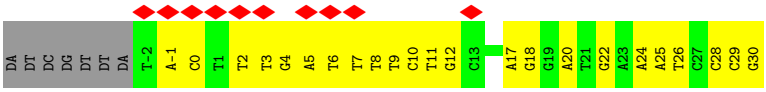
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: TS-DNA

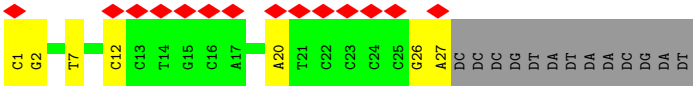
Chain B:





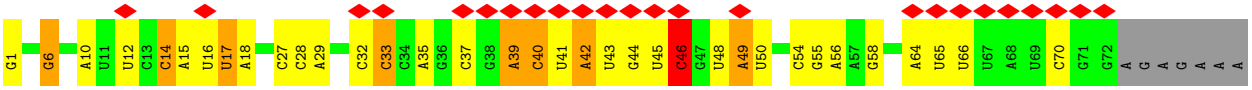
- Molecule 2: NTS-DNA

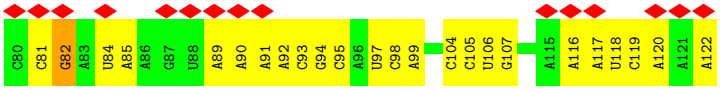
Chain C:



- Molecule 3: RNA (115-MER)

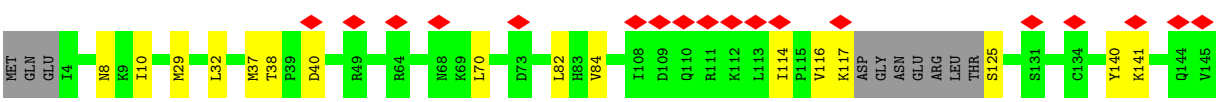
Chain D:

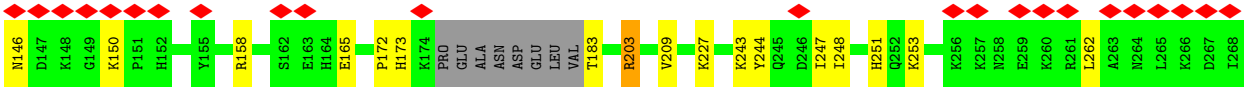


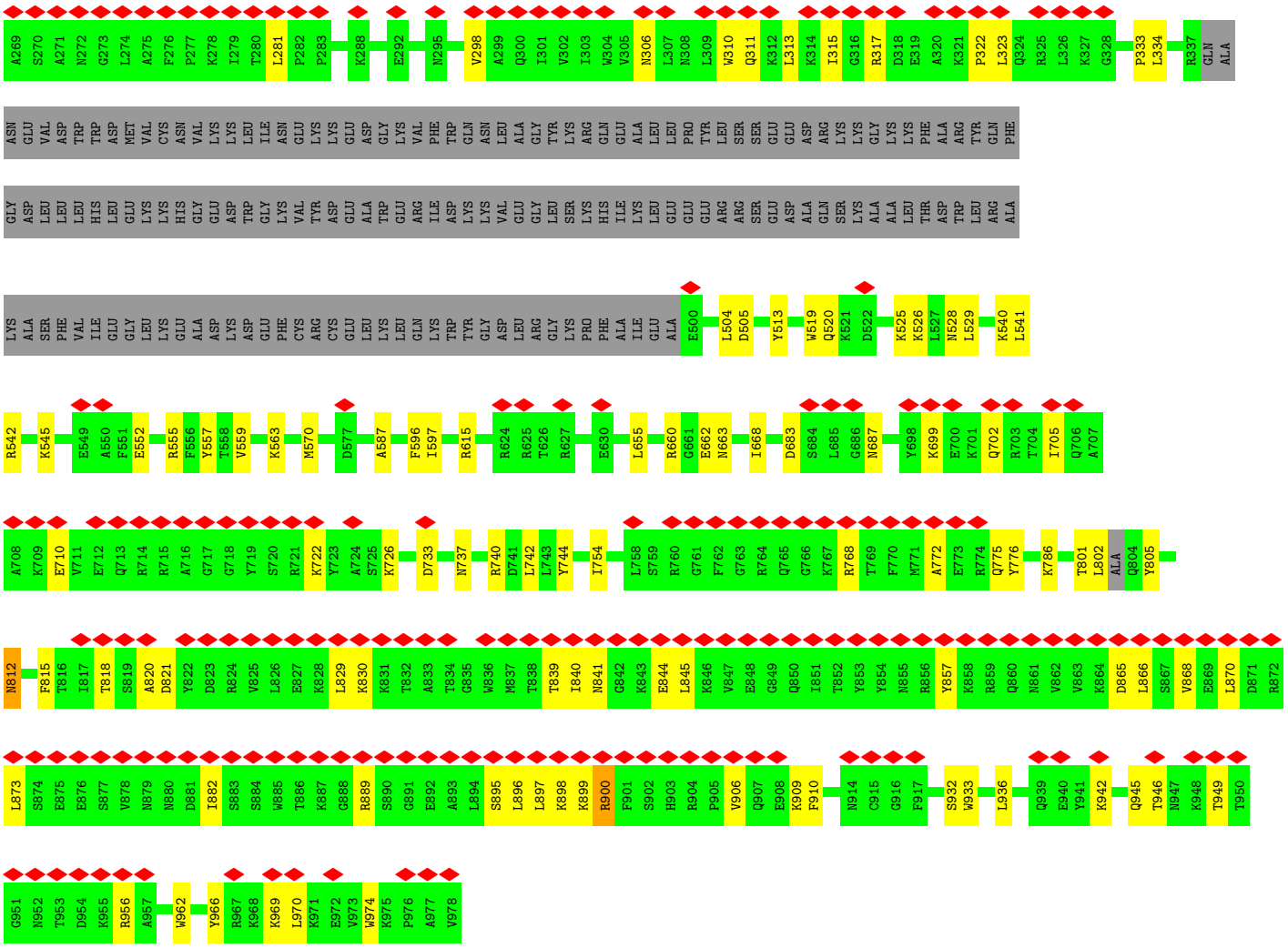


- Molecule 4: dPlmCasX

Chain A:






A269 S270 A271 N272 G273 L274 A275 F276 P277 K278 I279 T280 L281 P282 P283 K288 E292 N295 V298 A299 Q300 I301 V302 I303 W304 V305 N306 L307 N308 L309 W310 Q311 K312 L313 K314 I315 G316 R317 D318 E319 A320 K321 P322 L323 Q324 R325 L326 K327 G328 P333 L334 R337 GLN ALA
ASN GLU VAL ASP TRP TRP ASP MET VAL CYS ASN VAL LYS LYS LEU ILE ASN GLU LYS LYS LYS GLU ASP GLY LYS VAL PHE TRP GLN ASN LEU ALA GLY TYR LYS ARG GLN GLU ALA LEU LEU PRO TYR LEU SER SER GLU GLU ASP ARG LYS LYS GLY LYS LYS PHE ALA ARG TYR GLN PHE
GLY ASP LEU LEU HIS LEU GLU LYS LYS HIS GLY GLU ASP TRP GLY LYS VAL TYR ASP GLU ALA TRP GLU ARG ILE ASP LYS LYS VAL GLU GLY LEU SER LYS HIS ILE LYS LEU GLU GLU GLU ARG ARG SER GLU ASP ALA GLN SER LYS ALA ALA LEU THR ASP TRP LEU ARG ALA
LYS ALA SER PHE VAL ILE GLU GLY LEU LYS GLU ALA ASP LYS ASP GLU PHE CYS ARG CYS GLU LEU LYS LEU GLN LYS TRP TYR GLY ASP LEU ARG GLY LYS PRO PHE ALA ILE GLU ALA E500 L504 D505 Y513 W519 Q520 K521 D522 K525 K526 L527 N528 L529 K540 L541
R542 K545 E549 A550 F551 E552 R555 F556 Y557 T558 V559 K563 M570 D577 A587 F596 I597 R615 R624 R625 T626 R627 E630 L655 R660 G661 E662 N663 I668 D683 S684 L685 G686 N687 Y698 K699 E700 K701 Q702 R703 T704 I705 Q706 A707
A708 K709 E710 V711 E712 Q713 R714 R715 A716 G717 G718 Y719 S720 R721 K722 Y723 A724 S725 K726 D733 N737 R740 D741 L742 L743 Y744 I754 L758 S759 R760 G761 F762 G763 R764 Q765 G766 K767 R768 T769 F770 M771 A772 E773 R774 Q775 Y776 K786 T801 L802 ALA Q804 Y805
N812 F815 T816 I817 T818 S819 A820 D821 Y822 D823 R824 V825 L826 E827 K828 L829 K830 K831 T832 A833 T834 G835 W836 M837 T838 T839 I840 N841 G842 K843 E844 L845 K846 V847 E848 G849 Q850 I851 T852 Y853 Y854 N855 R856 Y857 K858 R859 Q860 N861 V862 V863 K864 D865 L866 S867 V868 E869 L870 D871 R872
L873 S874 E875 E876 S877 V878 N879 N880 D881 I882 S883 S884 W885 T886 K887 G888 R889 S890 G891 E892 A893 L894 S895 L896 L897 K898 K899 R900 F901 S902 H903 R904 P905 V906 Q907 E908 K909 F910 N914 C915 G916 F917 S932 W933 L936 Q939 E940 Y941 K942 Q945 T946 N947 K948 T949 T950
G951 N952 T953 D954 K955 R956 A957 W962 Y966 R967 K968 K969 L970 K971 E972 V973 W974 K975 P976 A977 V978

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	710824	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	1.923	Depositor
Minimum map value	-1.096	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.054	Depositor
Recommended contour level	0.325	Depositor
Map size (Å)	240.64, 240.64, 240.64	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.94, 0.94, 0.94	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	B	0.51	0/753	0.96	0/1160
2	C	0.51	0/611	0.88	0/939
3	D	0.21	0/2734	0.82	2/4254 (0.0%)
4	A	0.25	0/6554	0.51	0/8822
All	All	0.29	0/10652	0.68	2/15175 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	46	C	N1-C2-O2	6.13	122.58	118.90
3	D	17	U	C2-N1-C1'	5.37	124.15	117.70

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	B	674	0	376	22	0
2	C	547	0	304	5	0
3	D	2446	0	1239	24	0
4	A	6429	0	6584	96	0
All	All	10096	0	8503	133	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 133 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)
4:A:897:LEU:HD22	4:A:900:ARG:HD2	1.62	0.82
4:A:733:ASP:OD1	4:A:737:ASN:ND2	2.14	0.80
4:A:873:LEU:HB3	4:A:889:ARG:HE	1.45	0.78
3:D:90:A:H2'	3:D:91:A:H8	1.52	0.74
4:A:663:ASN:HA	4:A:699:LYS:HD2	1.72	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
4	A	787/978 (80%)	760 (97%)	26 (3%)	1 (0%)	51 83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	900	ARG

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
4	A	700/855 (82%)	694 (99%)	6 (1%)	78 91

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	A	812	ASN
4	A	830	LYS
4	A	956	ARG
4	A	203	ARG
4	A	117	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
4	A	306	ASN
4	A	512	GLN
4	A	775	GLN
4	A	804	GLN
4	A	924	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	D	113/122 (92%)	35 (30%)	0

5 of 35 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	D	6	G
3	D	12	U
3	D	14	C
3	D	15	A
3	D	16	U

There are no RNA pucker outliers to report.

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.

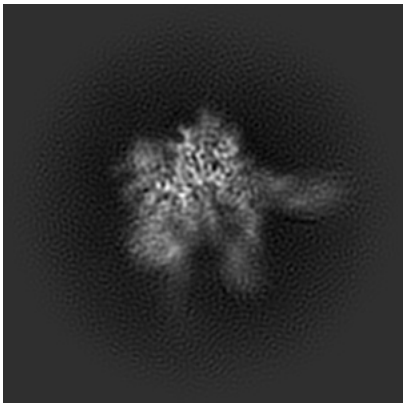
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-32391. These allow visual inspection of the internal detail of the map and identification of artifacts.

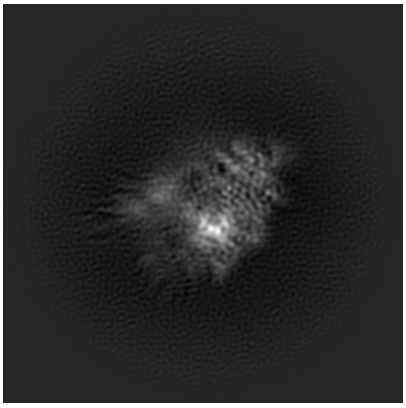
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

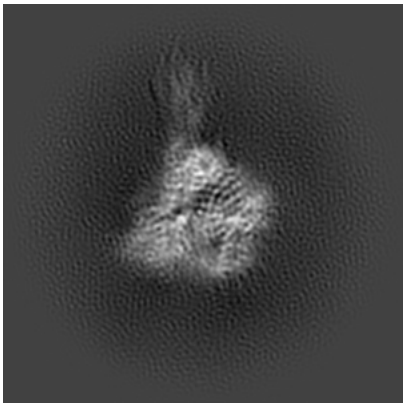
6.1.1 Primary map



X



Y

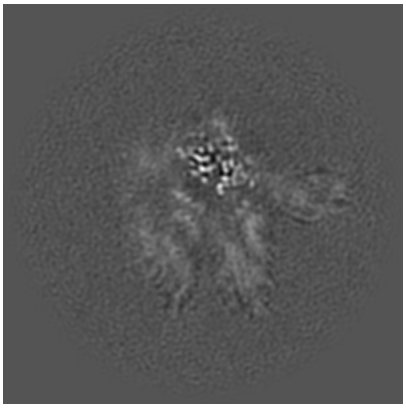


Z

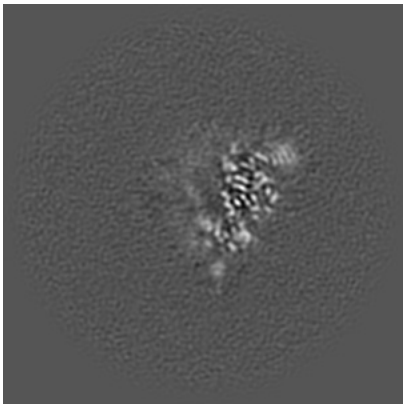
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

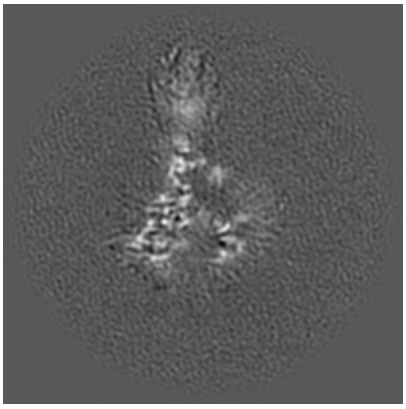
6.2.1 Primary map



X Index: 128



Y Index: 128

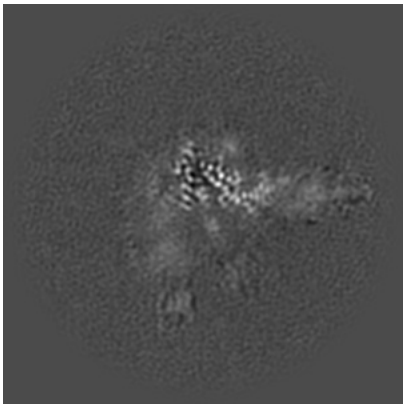


Z Index: 128

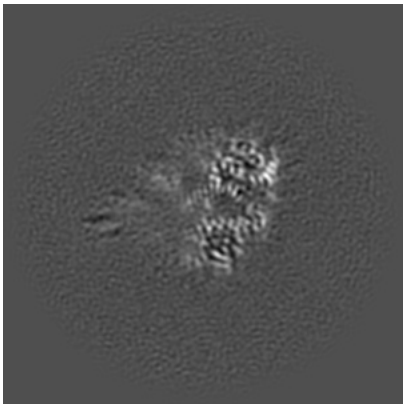
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

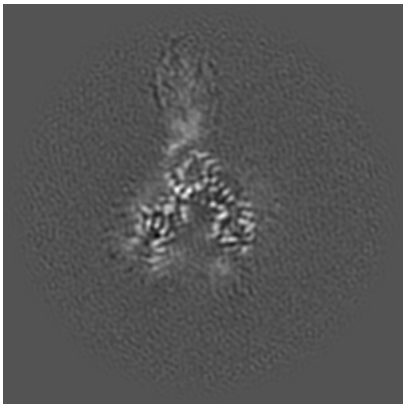
6.3.1 Primary map



X Index: 112



Y Index: 117



Z Index: 141

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

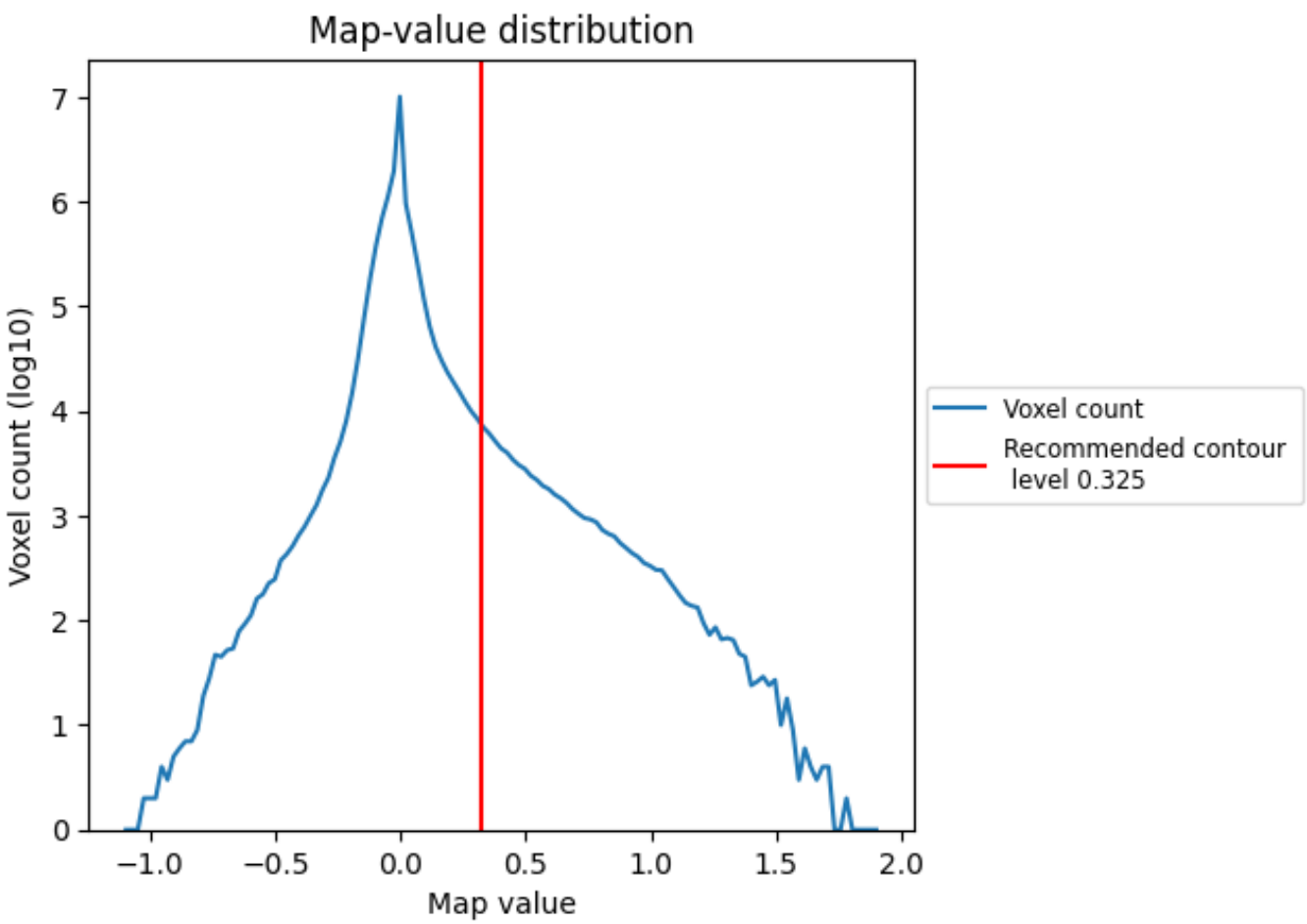
6.5 Mask visualisation (i)

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

7 Map analysis (i)

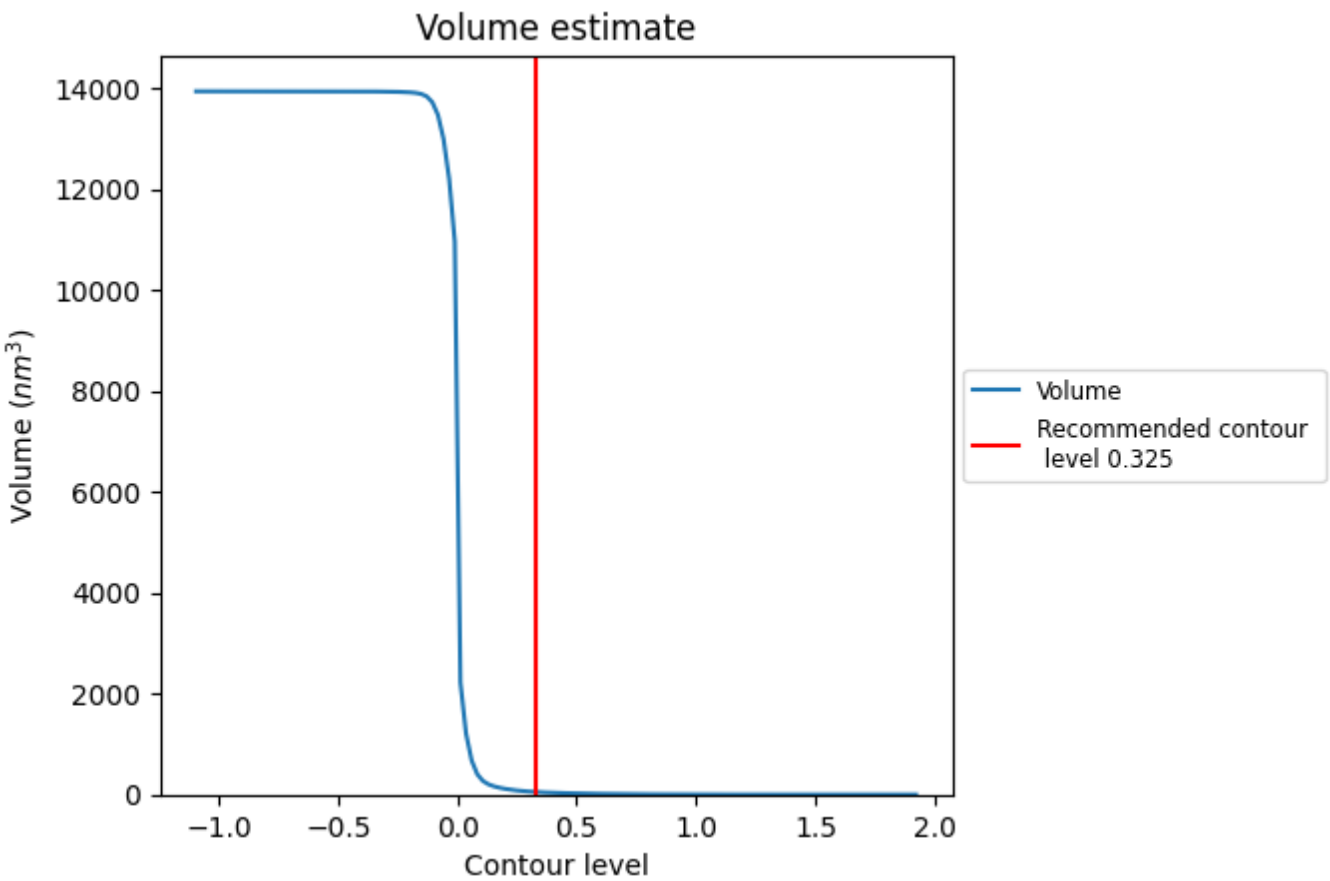
This section contains the results of statistical analysis of the map.

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.

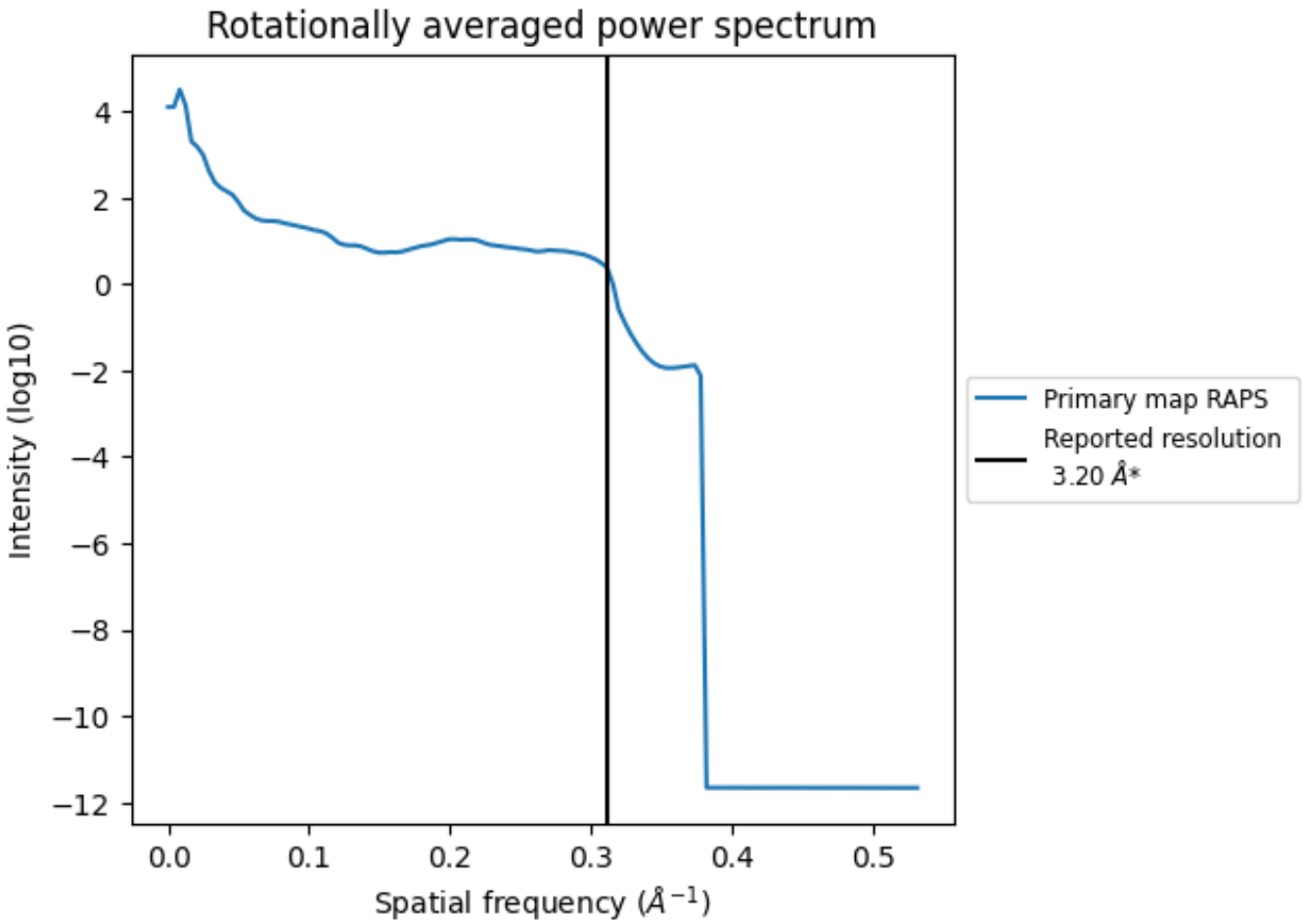
7.2 Volume estimate (i)



The volume at the recommended contour level is 52 nm^3; this corresponds to an approximate mass of 47 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}

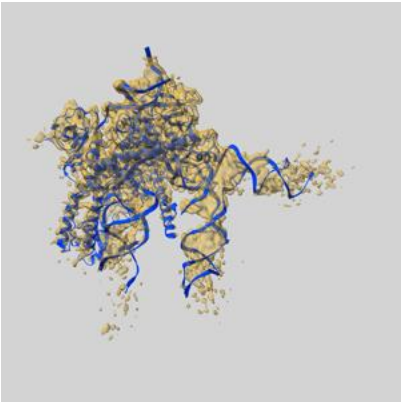
8 Fourier-Shell correlation (i)

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

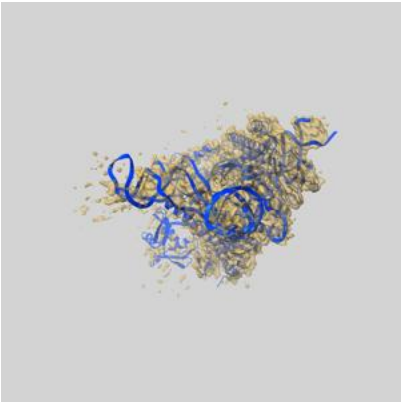
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-32391 and PDB model 7WB0. Per-residue inclusion information can be found in section 3 on page 4.

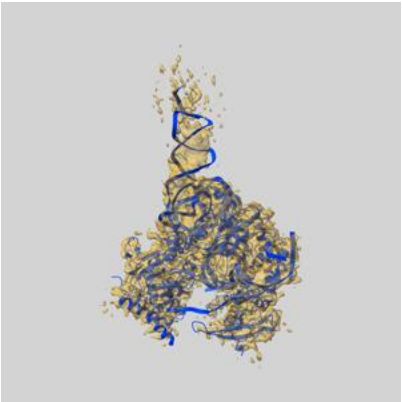
9.1 Map-model overlay (i)



X



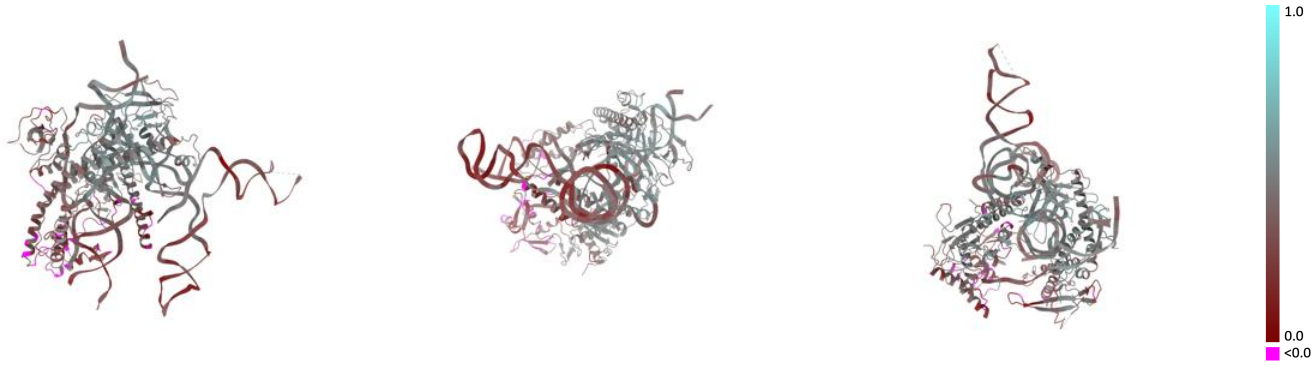
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.325 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.

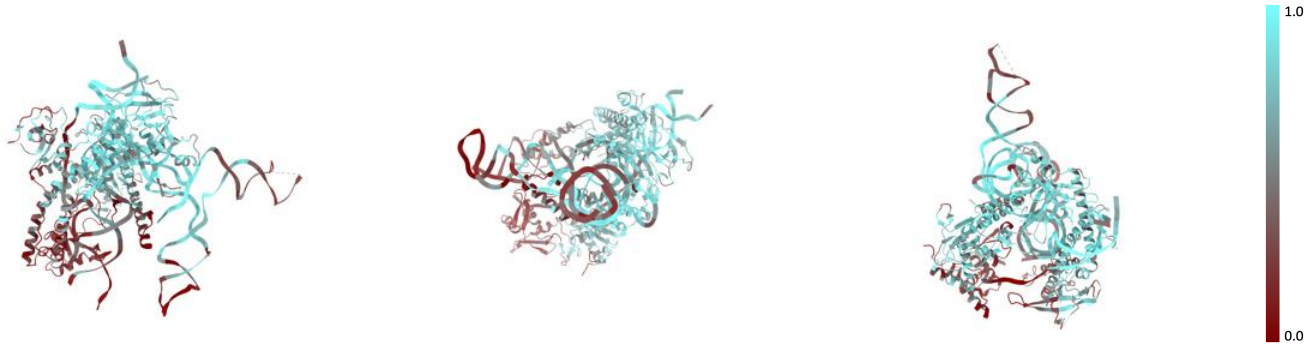
9.2 Q-score mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according 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.

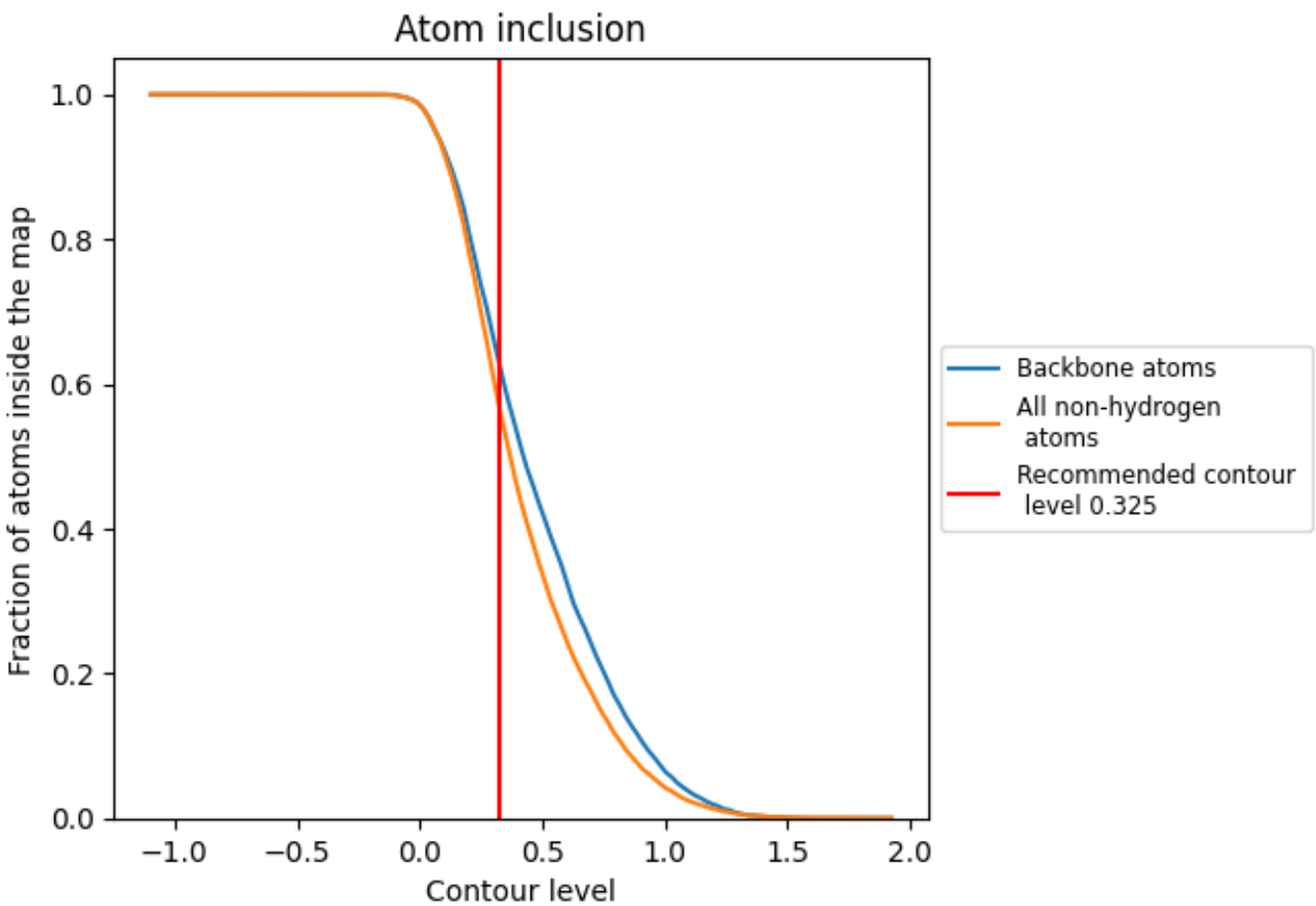
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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 (0.325).

9.4 Atom inclusion (i)



At the recommended contour level, 63% of all backbone atoms, 57% 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 (0.325) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5654	0.3740
A	0.5608	0.3770
B	0.5742	0.3980
C	0.4424	0.3880
D	0.6022	0.3550