wwPDB EM Validation Summary Report (i)

Nov 23, 2022 – 05:43 AM JST

PDB ID	:	7FG7
EMDB ID	:	EMD-31578
Title	:	Cryo-EM structure of S protein trimer of SARS-CoV2
Authors	:	Song, C.; Murata, K.; Katayama, K.
Deposited on	:	2021-07-26
Resolution	:	6.90 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:

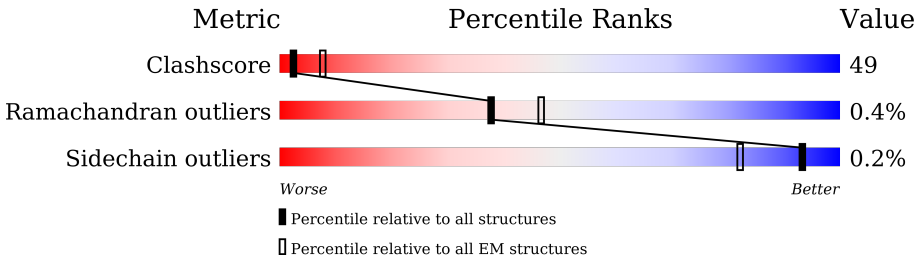
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 6.90 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 <=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	1273	33% 30% 57% · 12%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 8690 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 Spike glycoprotein.

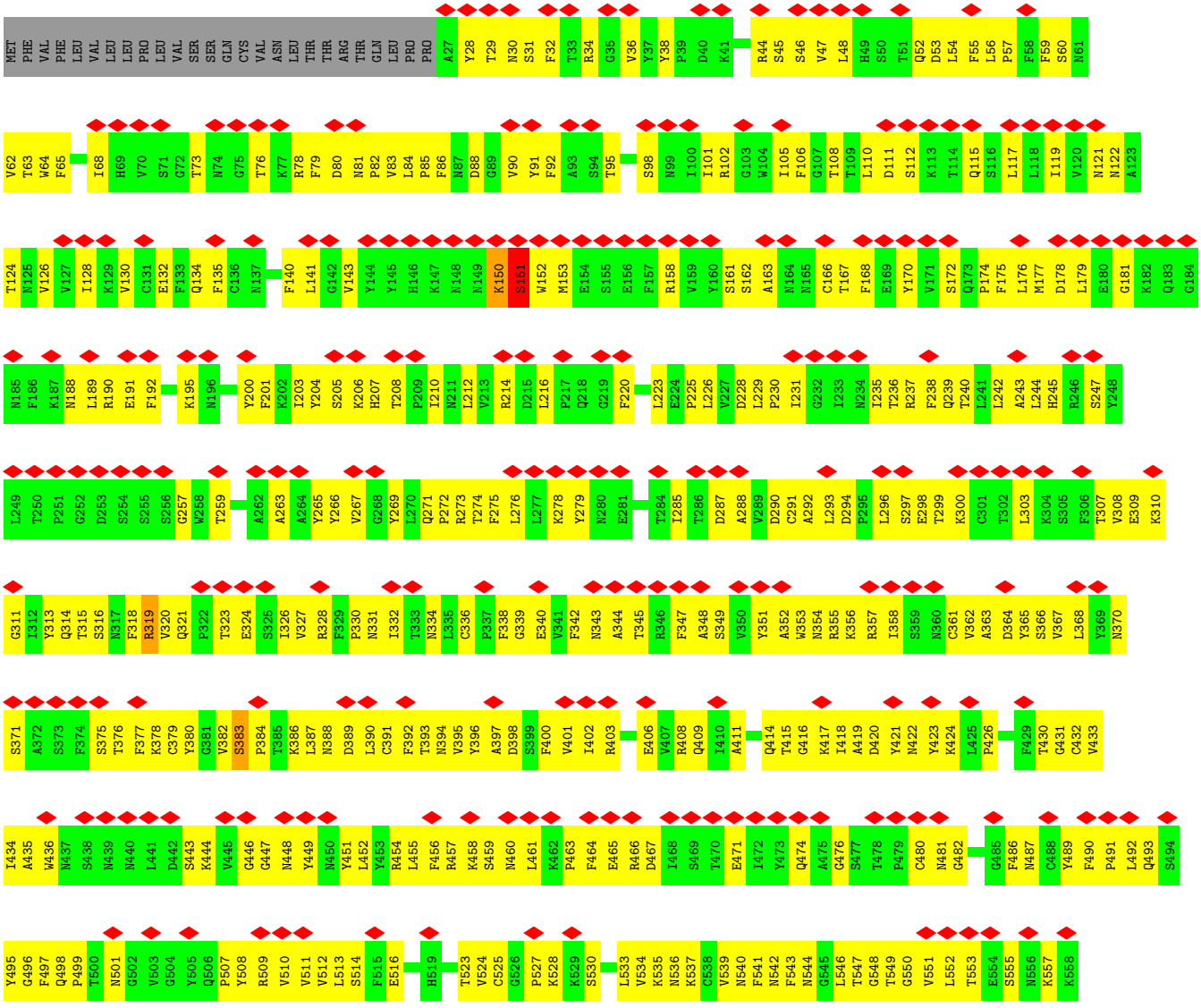
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1114	8690	5534	1460	1657	39	0	0

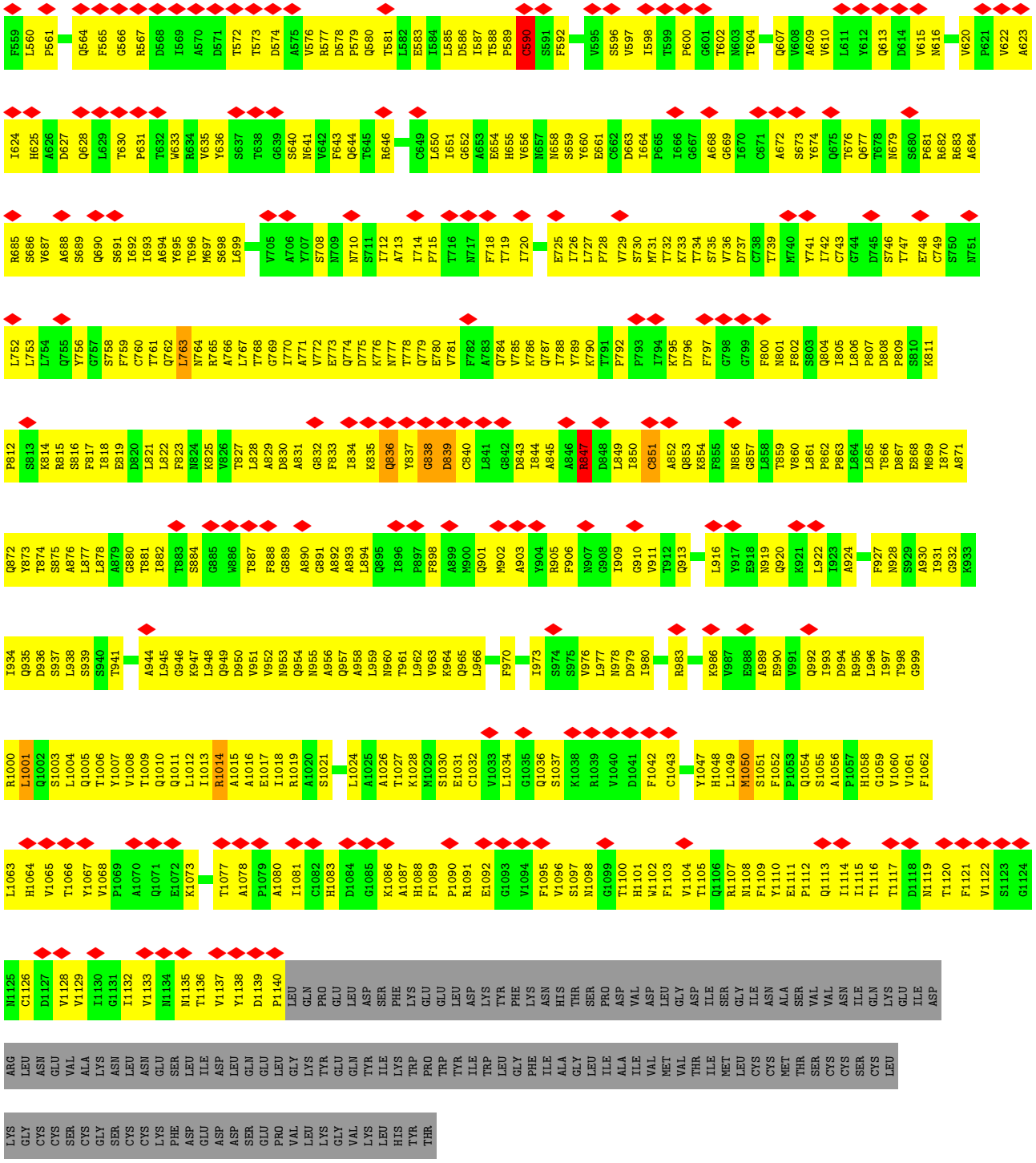
3 Residue-property plots (i)

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: Spike glycoprotein

Chain A:




F559 L560 P561 Q564 F565 G566 R567 D568 I569 A570 D571 T572 T573 D574 A575 V576 R577 D578 P579 Q580 T581 L582 E583 I584 L585 D586 I587 T588 P589 C590 S591 F592 V595 S596 V597 I598 T599 P600 G601 T602 N603 T604 Q607 V608 A609 V610 L611 Y612 Q613 D614 V615 N616 V620 P621 V622 A623
I624 H625 A626 D627 Q628 L629 T630 P631 T632 W633 R634 V635 Y636 S637 T638 G639 S640 N641 V642 F643 Q644 T645 R646 C649 L650 I651 G652 A653 E654 H655 V656 N657 N658 S659 Y660 E661 C662 D663 I664 F665 I666 G667 A668 G669 I670 C671 A672 S673 Y674 Q675 T676 Q677 T678 W679 S680 F681 R682 R683 A684
R685 S686 V687 A688 S689 Q690 S691 I692 I693 A694 Y695 T696 M697 S698 L699 V705 A706 Y707 S708 N709 N710 S711 I712 A713 I714 P715 T716 N717 F718 T719 I720 E725 I726 L727 P728 V729 S730 M731 T732 K733 T734 S735 V736 D737 G738 T739 M740 Y741 I742 C743 G744 D745 S746 T747 E748 C749 S750 N751
L752 L753 L754 Q755 Y756 G757 S758 F759 C760 T761 Q762 L763 N764 R765 A766 L767 T768 G769 I770 A771 V772 E773 Q774 D775 K776 N777 T778 Q779 E780 V781 F782 A783 Q784 V785 K786 Q787 I788 Y789 K790 T791 P792 P793 I794 K795 D796 F797 G798 G799 F800 N801 F802 S803 Q804 I805 L806 P807 D808 P809 S810 K811
P812 S813 K814 R815 S816 F817 I818 E819 D820 L821 L822 F823 N824 K825 V826 T827 L828 A829 D830 A831 G832 F833 I834 K835 Q836 Y837 G838 D839 C840 L841 G842 D843 I844 A845 A846 R847 D848 L849 I850 C851 A852 Q853 K854 F855 N856 G857 L858 T859 V860 L861 P862 P863 L864 L865 T866 D867 E868 M869 I870 A871
Q872 Y873 T874 S875 A876 L877 L878 A879 G880 T881 I882 T883 S884 G885 W886 T887 F888 G889 A890 G891 A892 A893 L894 Q895 I896 P897 F898 A899 N900 Q901 N902 A903 Y904 R905 F906 N907 G908 I909 G910 V911 T912 Q913 L916 Y917 E918 N919 Q920 K921 L922 I923 A924 F927 N928 S929 A930 I931 G932 K933
I934 Q935 D936 S937 L938 S939 S940 T941 A944 L945 G946 K947 L948 Q949 D950 V951 V952 N953 Q954 N955 A956 Q957 A958 L959 N960 T961 L962 V963 K964 Q965 L966 F970 I973 S974 S975 V976 L977 N978 D979 I980 R983 K986 V987 E988 A989 E990 Y991 Q992 I993 D994 R995 L996 I997 T998 G999
R1000 L1001 Q1002 S1003 L1004 Q1005 T1006 Y1007 V1008 T1009 Q1010 Q1011 L1012 I1013 R1014 A1015 A1016 E1017 I1018 R1019 A1020 S1021 L1024 A1025 A1026 T1027 K1028 M1029 S1030 E1031 C1032 V1033 L1034 G1035 Q1036 S1037 K1038 R1039 V1040 D1041 F1042 C1043 Y1047 H1048 L1049 M1050 S1051 F1052 P1053 Q1054 S1055 A1056 P1057 H1058 G1059 V1060 V1061 F1062
L1063 H1064 V1065 T1066 Y1067 V1068 P1069 A1070 Q1071 E1072 K1073 T1077 A1078 P1079 A1080 I1081 C1082 H1083 D1084 G1085 K1086 A1087 H1088 F1089 P1090 R1091 E1092 G1093 V1094 F1095 V1096 S1097 N1098 G1099 T1100 H1101 W1102 F1103 V1104 T1105 Q1106 R1107 N1108 F1109 Y1110 E1111 P1112 Q1113 I1114 I1115 T1116 T1117 D1118 N1119 T1120 F1121 V1122 S1123 G1124
N1125 C1126 D1127 V1128 V1129 I1130 G1131 I1132 V1133 N1134 N1135 T1136 V1137 Y1138 D1139 P1140 LEU GLN PRO GLU LEU ASP SER PHE LYS GLU GLU LEU ASP LYS TYR PHE LYS ASN HIS THR SER PRO ASP VAL ASP LEU GLY ASP ILE SER GLY ILE ASN ALA SER VAL VAL ASN ILE ILE GLN LYS GLU ILE ASP
ARG LEU ASN GLU VAL ALA LYS ASN LEU ASN GLU SER LEU ILE ASP LEU GLN GLU LEU GLY LYS TYR GLU GLN TYR ILE LYS TRP PRO TRP TYR ILE TRP LEU GLY PHE ILE ALA GLY LEU ILE ALA ILE VAL MET VAL THR ILE MET LEU CYS CYS MET THR SER CYS CYS SER CYS LEU
LYS GLY CYS CYS SER CYS GLY SER CYS CYS LYS PHE ASP GLU ASP ASP SER GLU PRO VAL LEU LYS GLY VAL LYS LEU HIS TYR THR

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	14235	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	JEOL 2200FS	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	1578.3	Depositor
Maximum defocus (nm)	4967.2	Depositor
Magnification	45065	Depositor
Image detector	DIRECT ELECTRON DE-20 (5k x 3k)	Depositor
Maximum map value	0.213	Depositor
Minimum map value	-0.145	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	363.52, 363.52, 363.52	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.42, 1.42, 1.42	Depositor

5 Model quality

5.1 Standard geometry

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	0.43	2/8893 (0.0%)	0.63	4/12105 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	8

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	590	CYS	C-N	6.64	1.49	1.34
1	A	274	THR	C-N	6.34	1.48	1.34

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1050	MET	CG-SD-CE	6.62	110.78	100.20
1	A	763	LEU	CA-CB-CG	-6.05	101.39	115.30
1	A	1001	LEU	CB-CG-CD2	-5.58	101.51	111.00
1	A	1014	ARG	NE-CZ-NH1	-5.21	117.69	120.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	151	SER	Peptide
1	A	383	SER	Peptide
1	A	590	CYS	Peptide
1	A	837	TYR	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	845	ALA	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	8690	0	8483	833	0
All	All	8690	0	8483	833	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 49.

The worst 5 of 833 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:620:VAL:HG11	1:A:623:ALA:HB2	1.44	1.00
1:A:781:VAL:HG12	1:A:1026:ALA:HB2	1.48	0.96
1:A:1010:GLN:HB3	1:A:1014:ARG:HH12	1.30	0.94
1:A:448:ASN:HB3	1:A:497:PHE:HB3	1.51	0.93
1:A:739:THR:O	1:A:743:CYS:N	2.01	0.92

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	1112/1273 (87%)	927 (83%)	180 (16%)	5 (0%)	34 72

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	838	GLY
1	A	839	ASP
1	A	150	LYS
1	A	151	SER
1	A	836	GLN

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	965/1112 (87%)	963 (100%)	2 (0%)	93 96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	319	ARG
1	A	847	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	655	HIS
1	A	955	ASN
1	A	1010	GLN
1	A	409	GLN
1	A	121	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

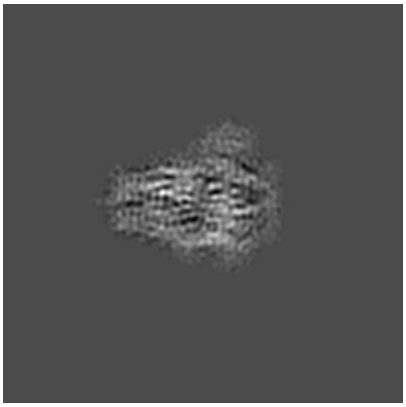
This section contains visualisations of the EMDB entry EMD-31578. These allow visual inspection of the internal detail of the map and identification of artifacts.

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.

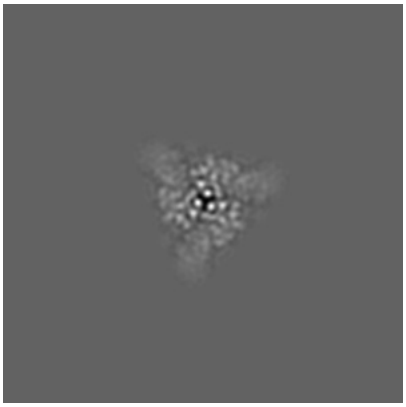
6.1 Orthogonal projections

6.1.1 Primary map

X



Y



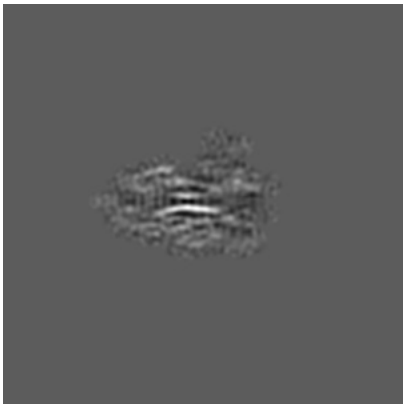
Z

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

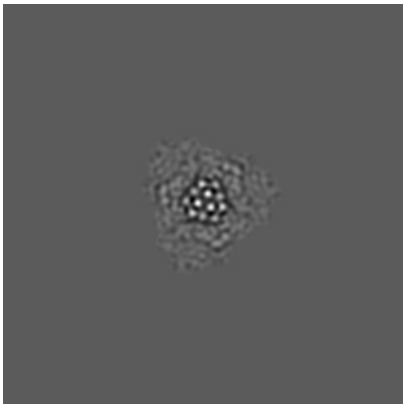
6.2 Central slices

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: 133

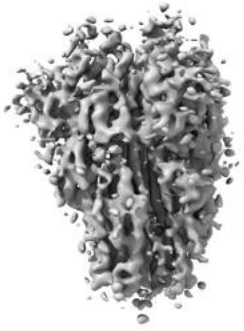
Y Index: 127

Z Index: 124

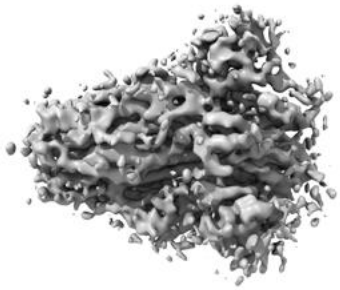
The images above show the largest variance slices of the map in three orthogonal directions.

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

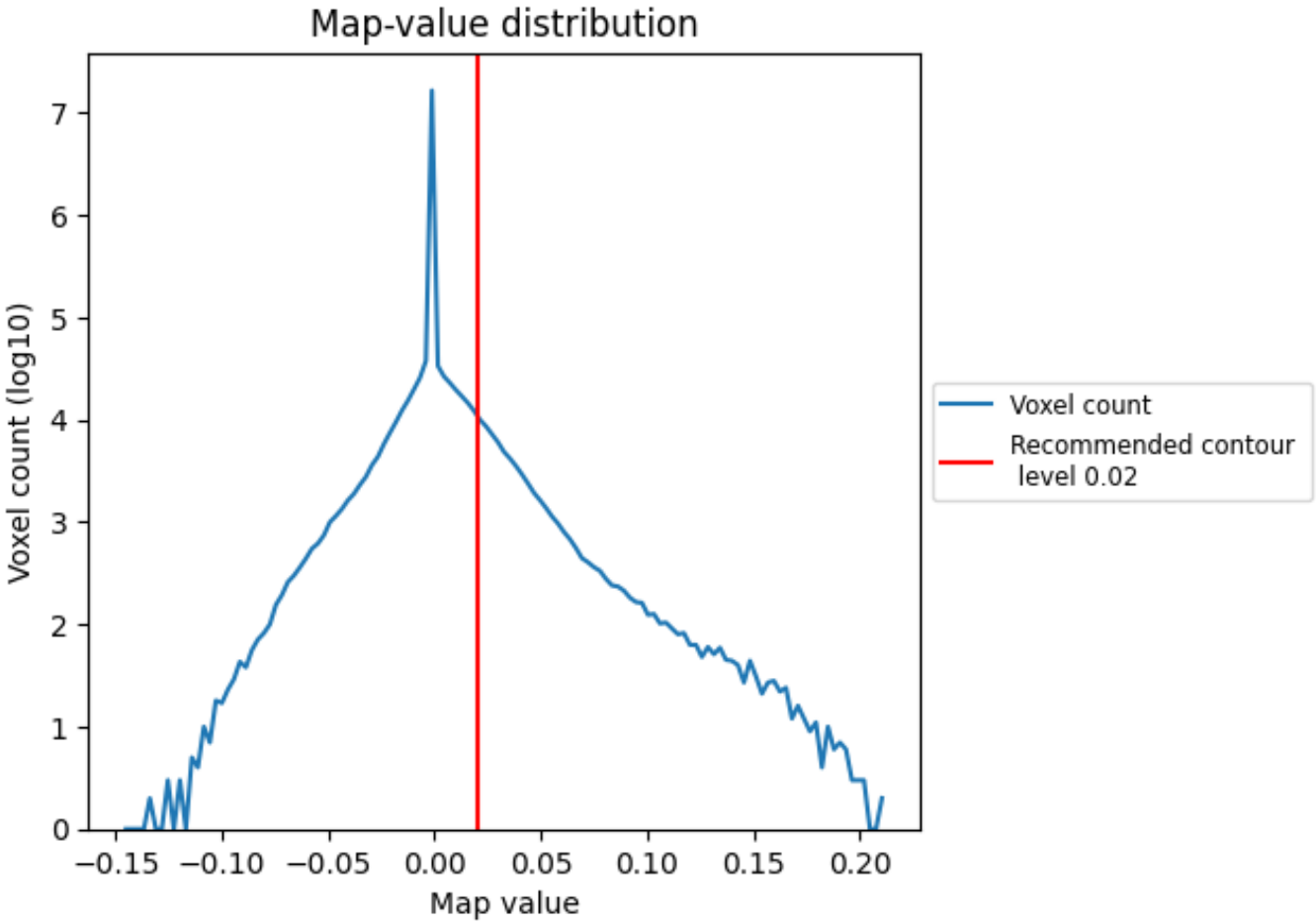
6.5 Mask visualisation (i)

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

7 Map analysis

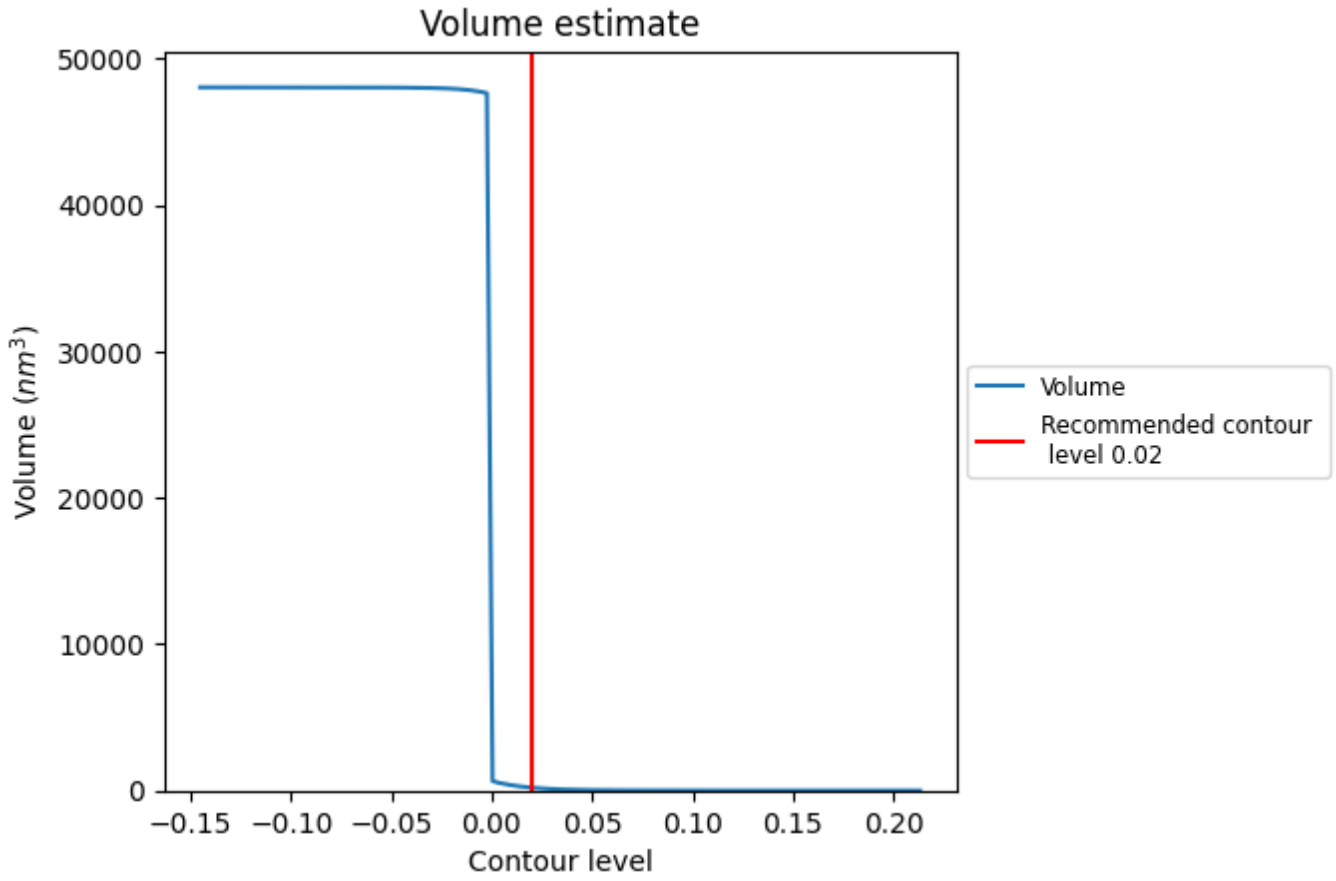
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

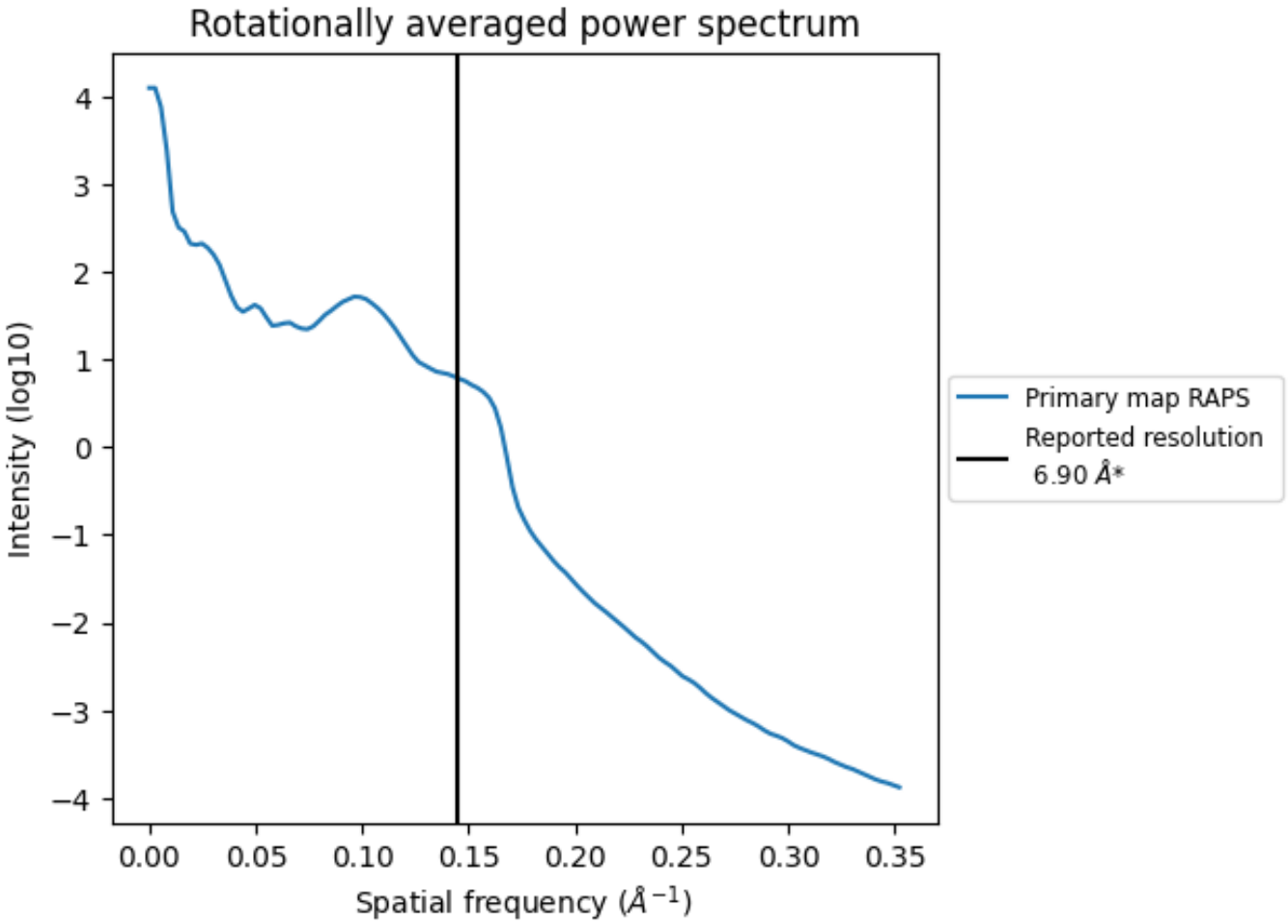
7.2 Volume estimate (i)



The volume at the recommended contour level is 197 nm^3; this corresponds to an approximate mass of 178 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)

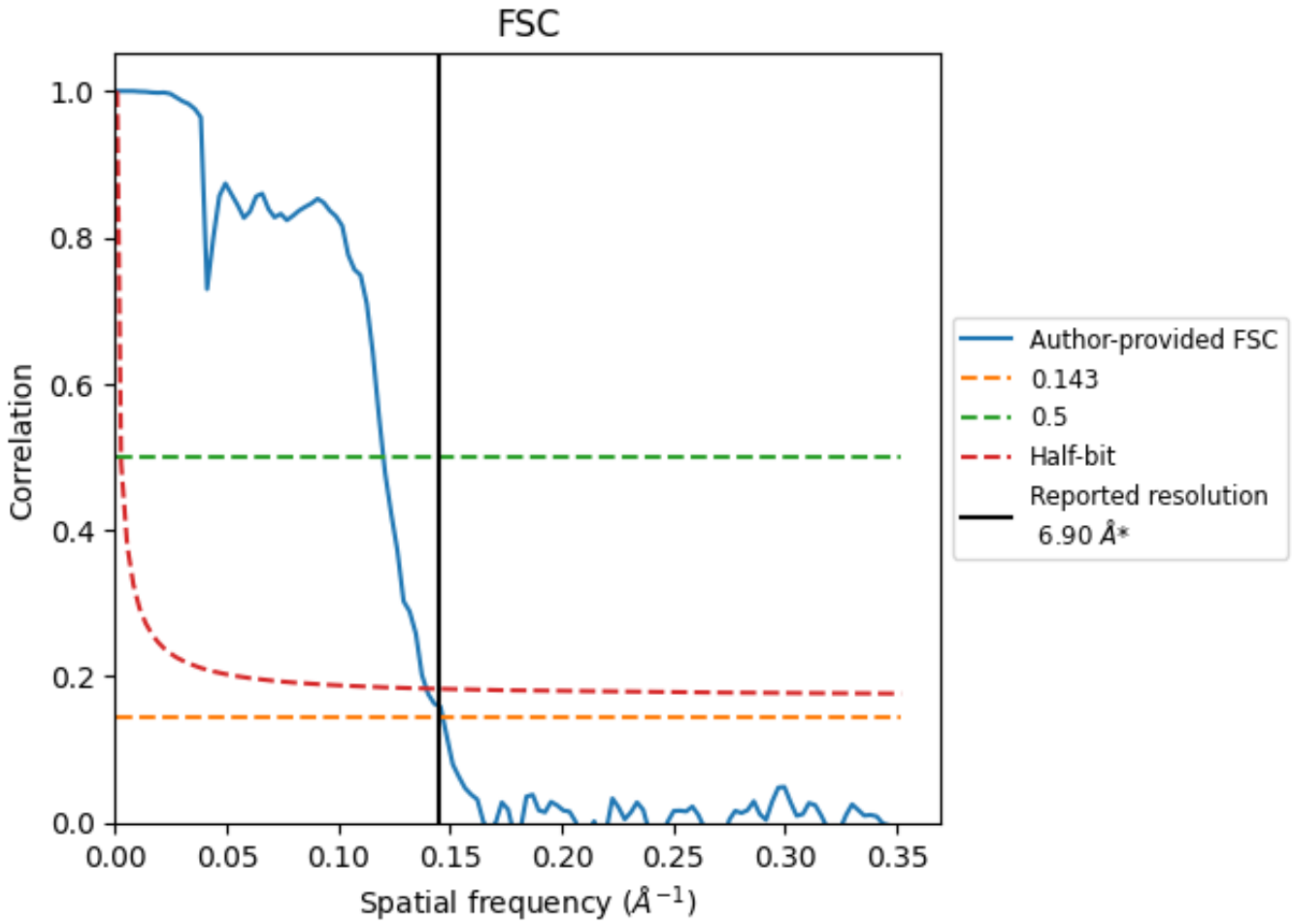


*Reported resolution corresponds to spatial frequency of 0.145 Å^{-1}

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.145 Å^{-1}

8.2 Resolution estimates (i)

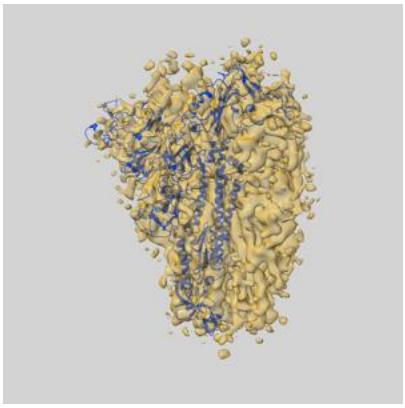
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	6.90	-	-
Author-provided FSC curve	6.81	8.32	7.17
Unmasked-calculated*	-	-	-

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

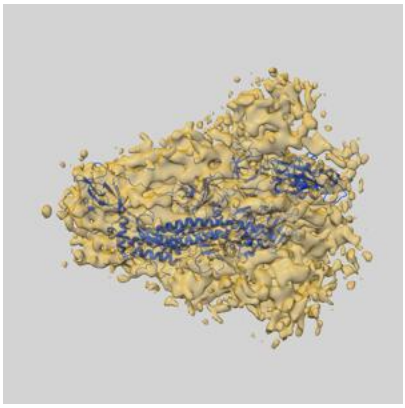
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-31578 and PDB model 7FG7. Per-residue inclusion information can be found in section 3 on page 4.

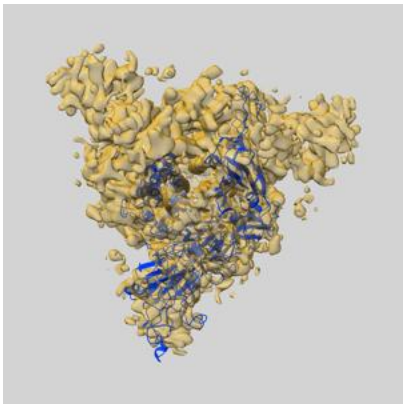
9.1 Map-model overlay



X



Y



Z

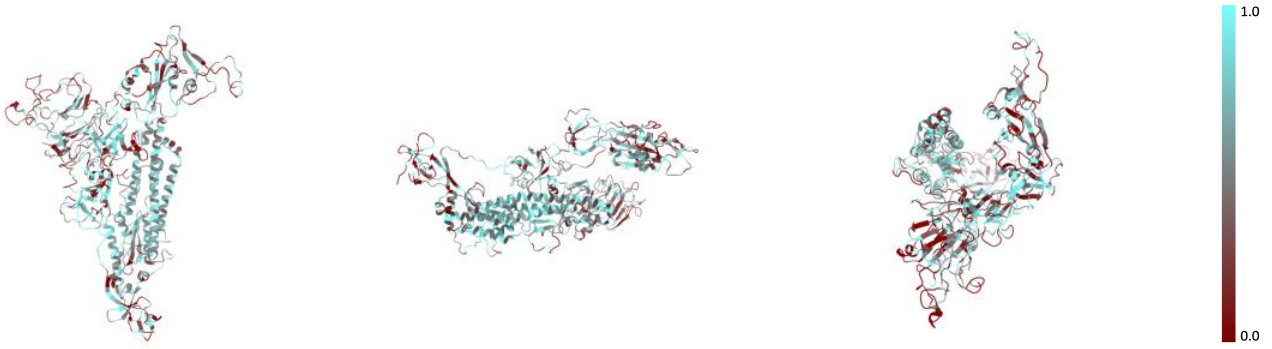
The images above show the 3D surface view of the map at the recommended contour level 0.02 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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.

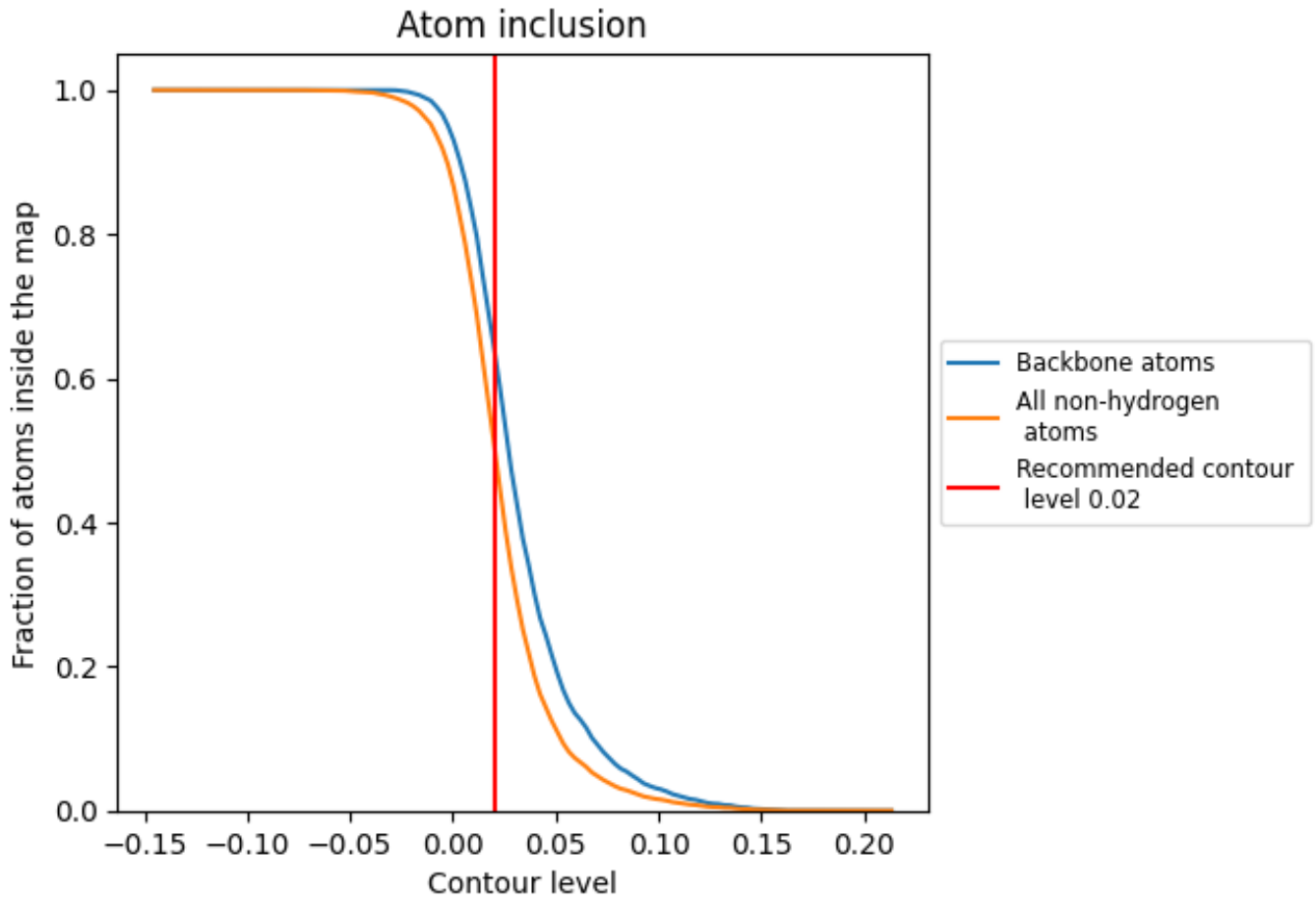
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.02).

9.4 Atom inclusion (i)



At the recommended contour level, 64% of all backbone atoms, 51% 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 (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5099	0.1530
A	0.5099	0.1530