wwPDB EM Validation Summary Report (i)

Nov 8, 2022 – 09:15 AM JST

PDB ID	:	5ZAM
EMDB ID	:	EMD-6906
Title	:	Cryo-EM structure of human Dicer and its complexes with a pre-miRNA substrate
Authors	:	Liu, Z.; Wang, J.; Cheng, H.; Ke, X.; Sun, L.; Zhang, Q.C.; Wang, H.-W.
Deposited on	:	2018-02-07
Resolution	:	5.70 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:

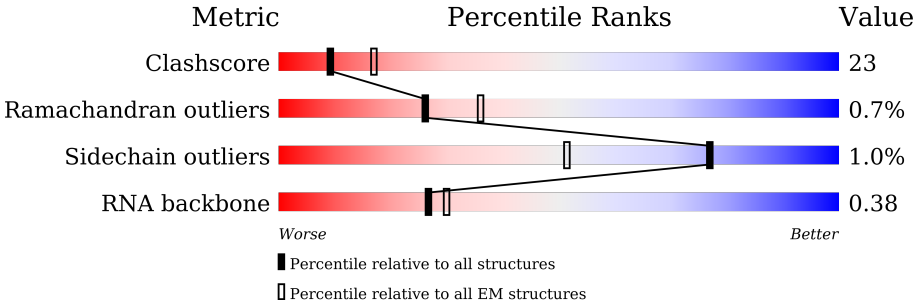
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 5.70 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	A	1922	8%; 35%; 31%; •; 32%
2	B	366	7%; 13%; 7%; 80%
3	C	73	•; 16%; 29%; 14%; •; 40%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12045 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 Endoribonuclease Dicer.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1314	10552	6742	1807	1943	60	0	0

- Molecule 2 is a protein called RISC-loading complex subunit TARBP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	75	560	345	99	110	6	0	0

- Molecule 3 is a RNA chain called RNA (73-mer).

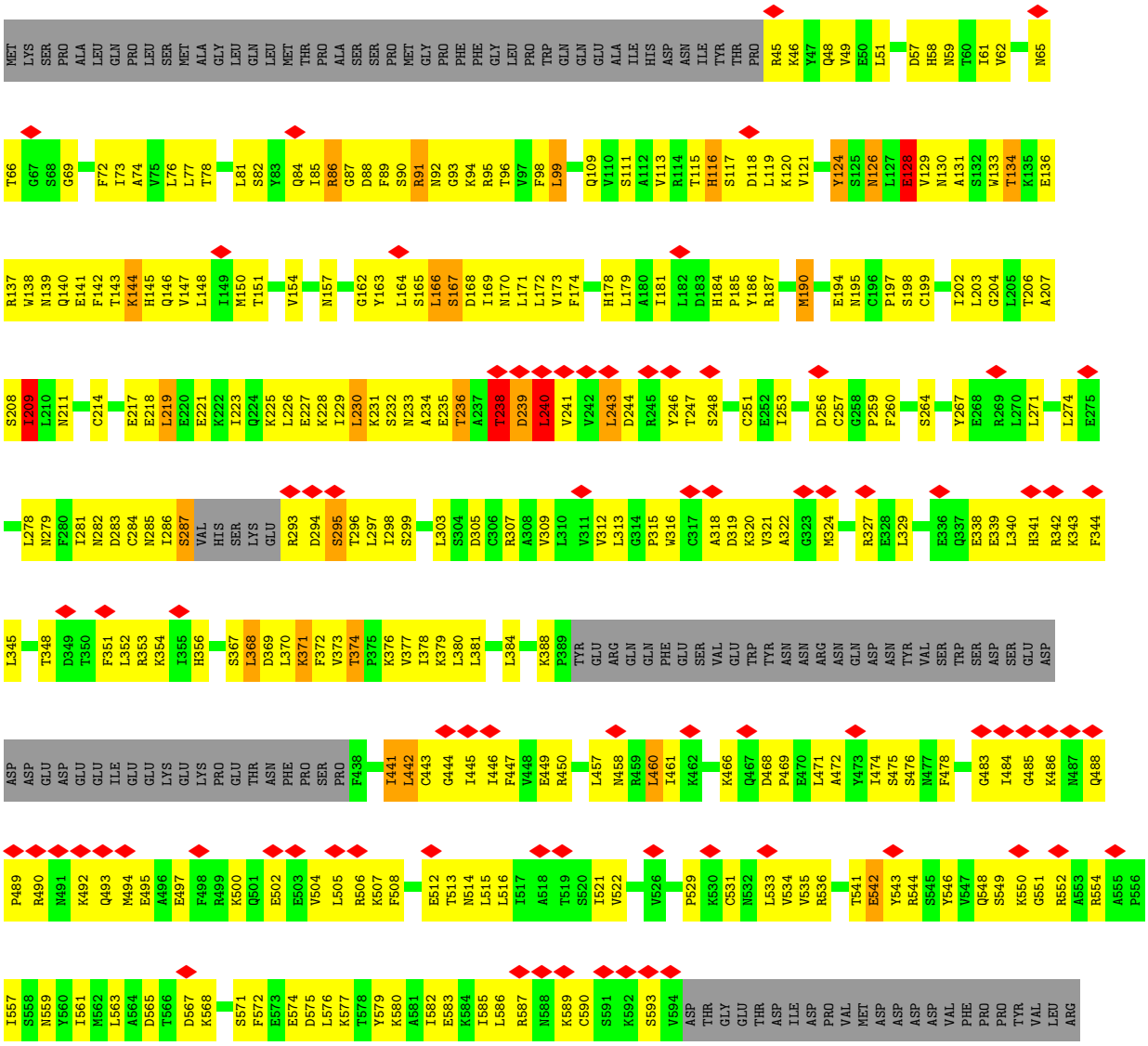
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	C	44	933	419	164	307	43	0	0

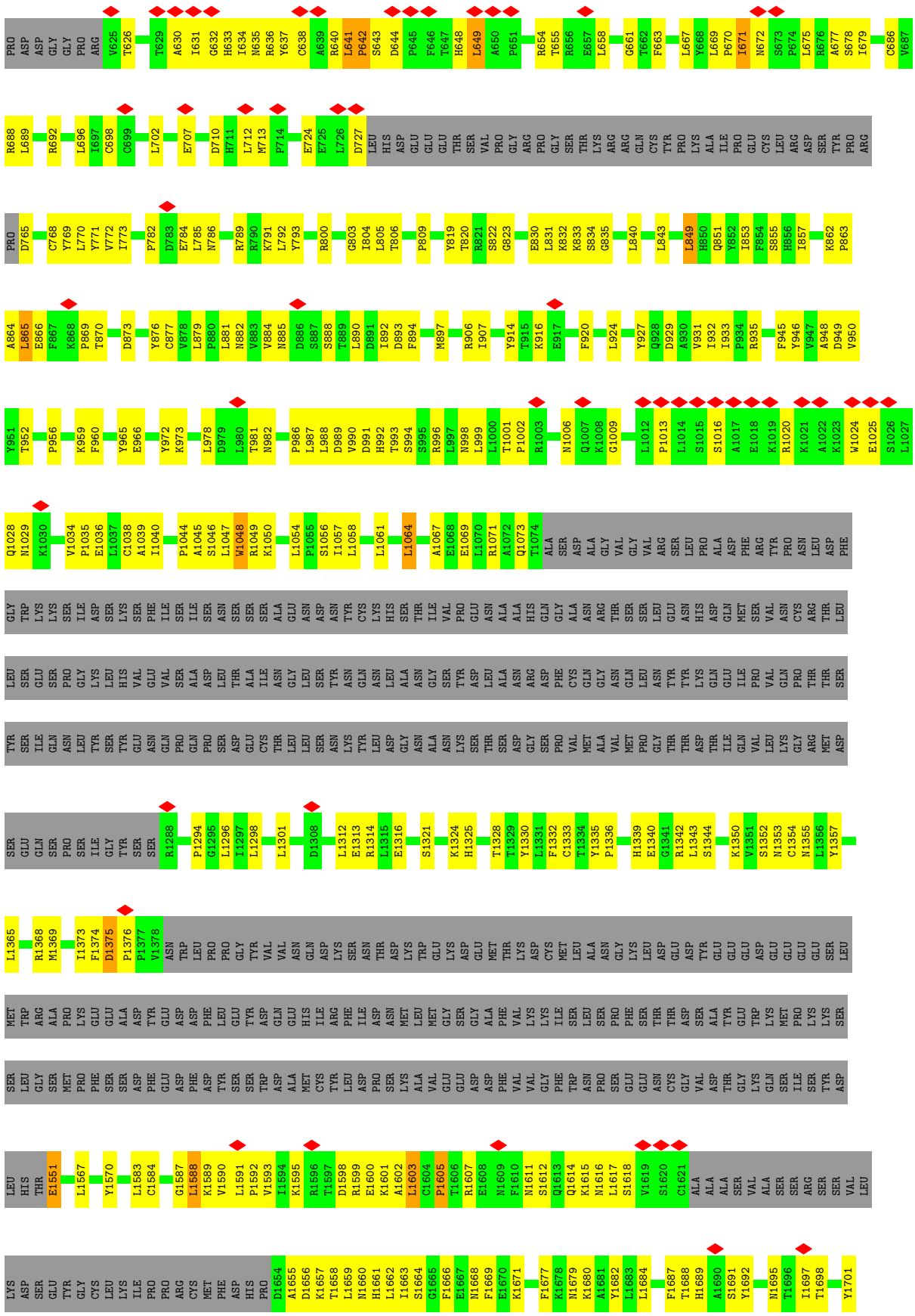
3 Residue-property plots (i)

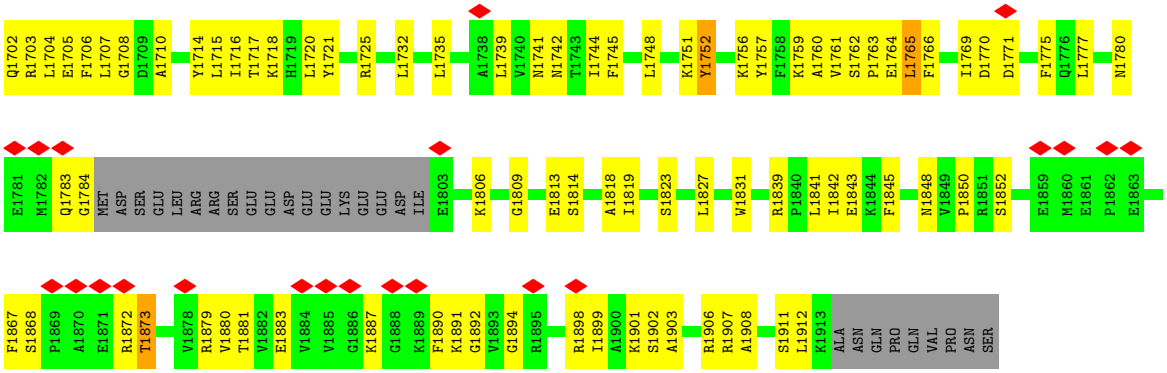
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: Endoribonuclease Dicer

Chain A: 8% 35% 31% • 32%

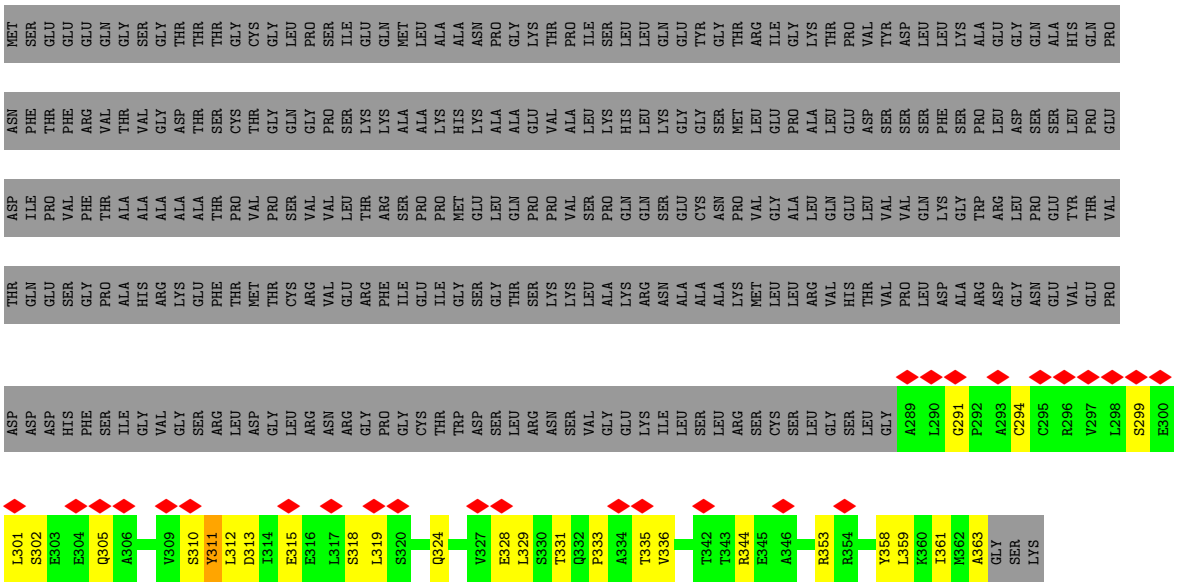



PRO ASP ASP GLY GLY PRO ARG V625 T626 T629 A630 I631 G632 H633 I634 N635 R636 Y637 C638 A639 R640 L641 P642 S643 D644 F645 F646 T647 H648 L649 A650 P651 R654 T655 R656 E657 L658 G661 T662 F663 L667 Y668 L669 P670 I671 N672 S673 P674 L675 R676 A677 S678 I679 C686 V687
R688 L689 R692 L696 I697 C698 C699 L702 E707 D710 H711 L712 M713 P714 E724 E725 L726 D727 LEU HIS ASP GLU GLU GLU THR SER VAL PRO GLY ARG PRO GLY SER THR LYS ARG ARG GLN CYS TYR PRO LYS ALA ILE PRO GLU CYS LEU ARG ASP SER TYR PRO ARG
PRO D765 C768 Y769 L770 Y771 V772 I773 P782 D783 E784 L785 N786 R789 R790 K791 L792 Y793 R800 G803 I804 L805 T806 P809 Y819 T820 R821 S822 G823 E830 L831 K832 K833 S834 G835 L840 L843 L849 H850 Q851 Y852 I853 F854 S855 H856 I857 K862 P863
A864 L865 E866 F867 K868 P869 T870 D873 Y876 C877 V878 L879 P880 L881 N882 V883 V884 N885 D886 S887 S888 T889 L890 D891 I892 D893 F894 M897 R906 I907 Y914 T915 K916 E917 F920 L924 Y927 Q928 D929 A930 V931 I932 I933 P934 R935 F945 Y946 V947 A948 D949 V950
Y951 T952 P956 K959 F960 Y965 E966 Y972 K973 L978 D979 L980 T981 N982 P986 L987 L988 D989 V990 D991 H992 T993 S994 S995 R996 L997 N998 L999 L1000 T1001 P1002 R1003 N1006 Q1007 K1008 G1009 L1012 P1013 L1014 S1015 S1016 A1017 E1018 K1019 R1020 K1021 A1022 K1023 W1024 E1025 S1026 L1027
Q1028 N1029 K1030 V1034 P1035 E1036 L1037 C1038 A1039 I1040 P1044 A1045 S1046 L1047 W1048 R1049 K1050 L1054 P1055 S1056 I1057 L1058 L1061 L1064 A1067 E1068 E1069 L1070 R1071 A1072 Q1073 T1074 ALA SER ASP ALA GLY VAL GLY VAL ARG SER LEU PRO PRO ASP PHE ARG TYR PRO ASN LEU ASP PHE
GLY TRP LYS LYS SER ILE ASP SER LYS SER PHE ILE SER ILE SER ASN SER SER SER ALA GLU ASN ASP ASN TYR CYS LYS HIS SER THR ILE VAL PRO GLU ASN ALA ALA HIS GLN GLY ALA ASN ARG THR SER SER LEU GLU ASN HIS ASP GLN MET SER VAL ASN CYS ARG THR LEU
LEU SER GLU SER PRO GLY LYS LEU HIS VAL GLU VAL SER ALA ASP LEU THR ALA ILE ASN GLY LEU SER TYR ASN GLN ASN LEU ALA ASN GLY SER TYR ASP LEU LEU ALA ASN ASP ARG ASP PHE CYS GLN GLY ASN GLN TYR TYR LYS GLN GLU ILE GLN VAL PRO VAL GLN PRO THR THR SER
TYR SER ILE GLN ASN LEU TYR SER TYR GLU ASN GLN PRO GLN PRO SER ASP GLU CYS THR LEU LEU SER ASN LYS TYR LEU ASP GLY ASN ALA ASN LYS SER THR SER ASP GLY SER PRO VAL MET ALA VAL MET PRO GLY THR THR ASP THR ILE GLN VAL LEU LYS GLY ARG MET ASP
SER GLU GLN SER PRO SER ILE GLY TYR SER SER R1288 P1294 G1295 L1296 I1297 L1298 L1301 D1308 L1312 E1313 R1314 L1315 E1316 S1321 K1324 H1325 T1328 T1329 Y1330 L1331 F1332 C1333 T1334 Y1335 P1336 H1339 E1340 G1341 R1342 L1343 S1344 K1350 V1351 S1352 N1353 C1354 N1355 L1356 Y1357
L1365 R1368 M1369 I1373 F1374 D1375 P1376 P1377 V1378 ASN TRP LEU PRO PRO GLY TYR VAL VAL ASN GLN ASP LYS SER ASN THR ASP LYS TRP GLU LYS ASP GLU MET THR LYS ASP CYS MET LEU ALA ASN GLY LYS LEU ASP GLU GLU ASP TYR GLU GLU GLU ASP GLU GLU GLU GLU SER LEU
MET TRP ARG ALA PRO LYS GLU GLU ALA ASP TYR GLU ASP ASP PHE LEU GLU TYR ASP GLN GLU HIS ILE ARG PHE ILE ASP ASN MET LEU MET GLY SER GLY ALA PHE VAL LYS LYS ILE SER LEU ALA PRO PHE SER THR THR ASP SER ALA TYR GLU TRP LYS MET PRO LYS LYS SER
SER LEU GLY SER MET PRO PHE SER SER SER ASP PHE GLU ASP PHE ASP TYR SER SER TRP ASP ALA MET CYS TYR LEU ASP PRO SER LYS ALA VAL GLU GLU GLU ASP ASP PHE VAL VAL GLY PHE TRP ASN PRO SER GLU GLU ASN CYS GLY VAL ASP THR GLY LYS GLN SER ILE SER TYR ASP
LEU HIS THR E1551 L1567 Y1570 L1583 C1584 G1587 L1588 K1589 V1590 L1591 P1592 V1593 I1594 K1595 R1596 T1597 D1598 R1599 E1600 K1601 A1602 L1603 C1604 P1605 T1606 R1607 E1608 N1609 F1610 N1611 S1612 Q1613 Q1614 K1615 N1616 L1617 S1618 V1619 S1620 C1621 ALA ALA ALA SER VAL ALA SER SER ARG SER SER VAL LEU
LYS ASP SER GLU TYR GLY CYS LEU LYS ILE PRO PRO ARG CYS MET PHE ASP HIS PRO D1654 A1655 D1656 K1657 T1658 L1659 N1660 H1661 L1662 I1663 S1664 G1665 F1666 E1667 N1668 F1669 E1670 K1671 F1677 K1678 N1679 K1680 A1681 Y1682 L1683 L1684 F1687 T1688 H1689 A1690 S1691 Y1692 N1695 T1696 I1697 T1698 Y1701



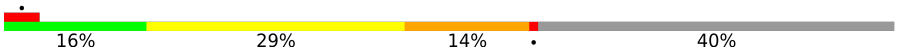
- Molecule 2: RISC-loading complex subunit TARBP2

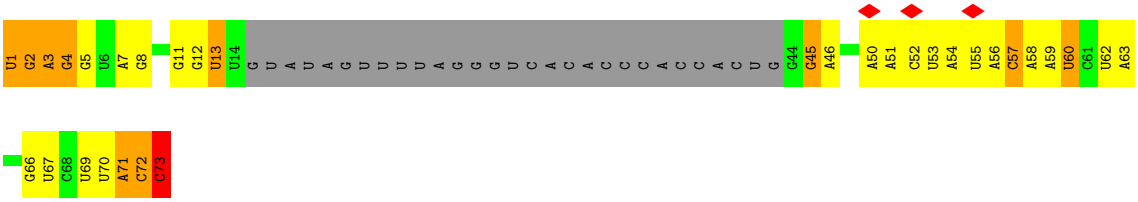
Chain B:



- Molecule 3: RNA (73-mer)

Chain C:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	60000	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)	2500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	36496	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.162	Depositor
Minimum map value	-0.062	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.04	Depositor
Map size (Å)	246.6, 246.6, 246.6	wwPDB
Map dimensions	180, 180, 180	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.37, 1.37, 1.37	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	0.51	1/10766 (0.0%)	0.83	23/14558 (0.2%)
2	B	0.44	0/566	0.65	0/765
3	C	0.70	0/1042	1.56	19/1619 (1.2%)
All	All	0.52	1/12374 (0.0%)	0.92	42/16942 (0.2%)

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	39
2	B	0	1
All	All	0	40

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1048	TRP	CB-CG	-5.26	1.40	1.50

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	72	C	C2-N1-C1'	13.47	133.62	118.80
3	C	72	C	N1-C2-O2	12.00	126.10	118.90
3	C	4	G	N7-C8-N9	11.57	118.89	113.10
3	C	72	C	N3-C2-O2	-9.79	115.05	121.90
3	C	72	C	C6-N1-C1'	-9.50	109.40	120.80

There are no chirality outliers.

5 of 40 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	116	HIS	Peptide
1	A	124	TYR	Peptide
1	A	126	ASN	Peptide
1	A	128	GLU	Peptide
1	A	93	GLY	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	10552	0	10711	496	0
2	B	560	0	559	26	0
3	C	933	0	473	32	0
All	All	12045	0	11743	536	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 23.

The worst 5 of 536 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:641:LEU:HD21	1:A:670:PRO:CG	1.41	1.46
1:A:641:LEU:CD2	1:A:670:PRO:HG3	1.55	1.37
1:A:295:SER:OG	1:A:344:PHE:HD2	1.03	1.28
1:A:286:ILE:HD11	2:B:315:GLU:CD	1.64	1.17
3:C:2:G:O6	3:C:70:U:C4	2.00	1.13

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	1296/1922 (67%)	1035 (80%)	252 (19%)	9 (1%)	22	62
2	B	73/366 (20%)	69 (94%)	4 (6%)	0	100	100
All	All	1369/2288 (60%)	1104 (81%)	256 (19%)	9 (1%)	26	62

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	239	ASP
1	A	240	LEU
1	A	295	SER
1	A	238	THR
1	A	1593	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	1182/1730 (68%)	1169 (99%)	13 (1%)	73	84
2	B	60/301 (20%)	60 (100%)	0	100	100
All	All	1242/2031 (61%)	1229 (99%)	13 (1%)	77	86

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	371	LYS
1	A	506	ARG
1	A	1765	LEU
1	A	1588	LEU
1	A	1590	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1552	GLN
1	A	1695	ASN
2	B	324	GLN
1	A	1679	ASN
1	A	548	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	C	42/73 (57%)	11 (26%)	0

5 of 11 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	C	2	G
3	C	45	G
3	C	46	A
3	C	50	A
3	C	51	A

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.

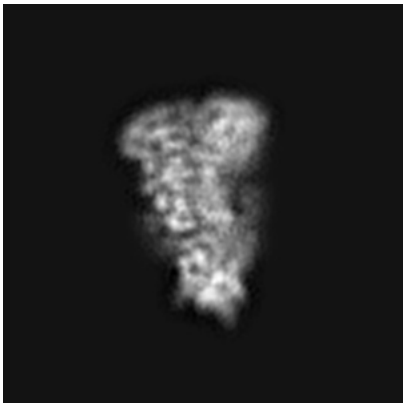
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-6906. These allow visual inspection of the internal detail of the map and identification of artifacts.

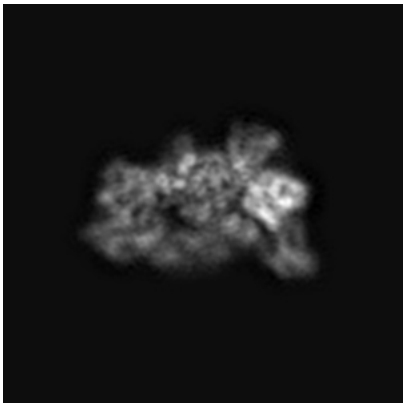
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

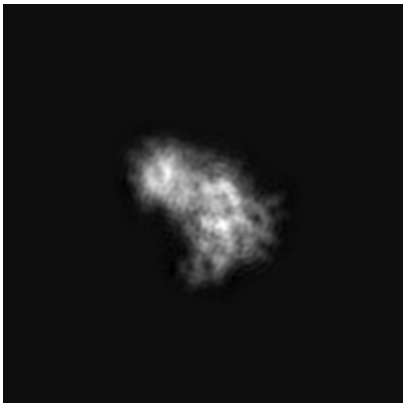
6.1.1 Primary map



X



Y



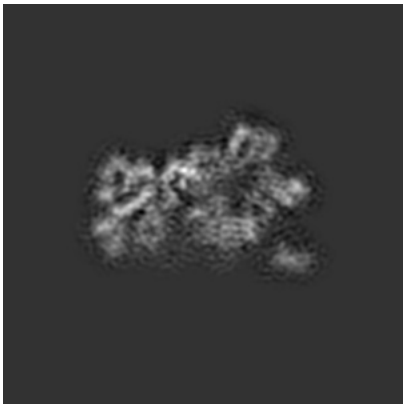
Z

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

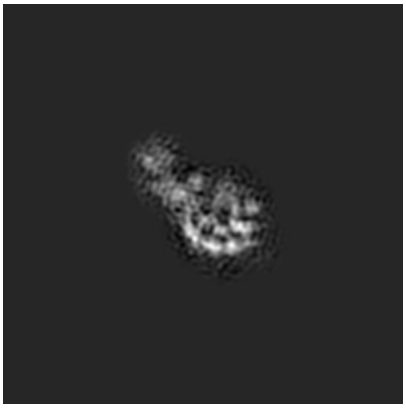
6.2 Central slices (i)

6.2.1 Primary map

X Index: 90



Y Index: 90



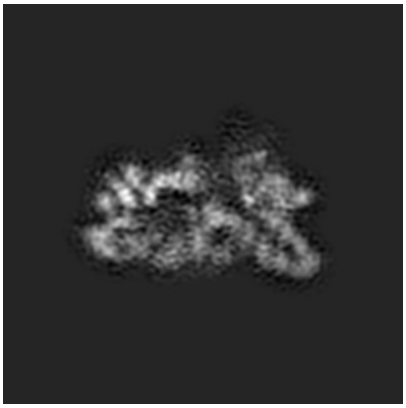
Z Index: 90

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

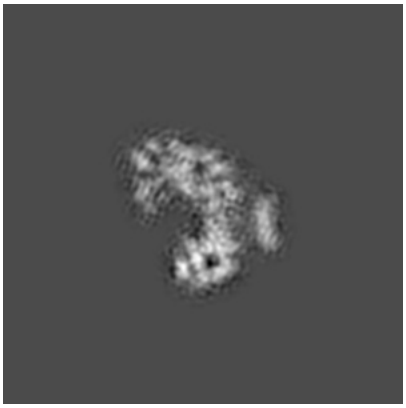
6.3 Largest variance slices (i)

6.3.1 Primary map

X Index: 98



Y Index: 97



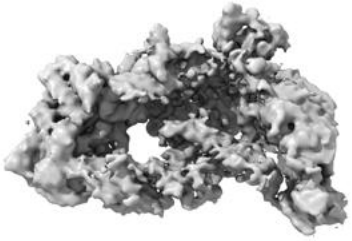
Z Index: 122

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

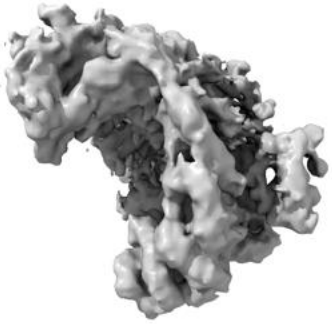
6.4 Orthogonal surface views (i)

6.4.1 Primary map

X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.04. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

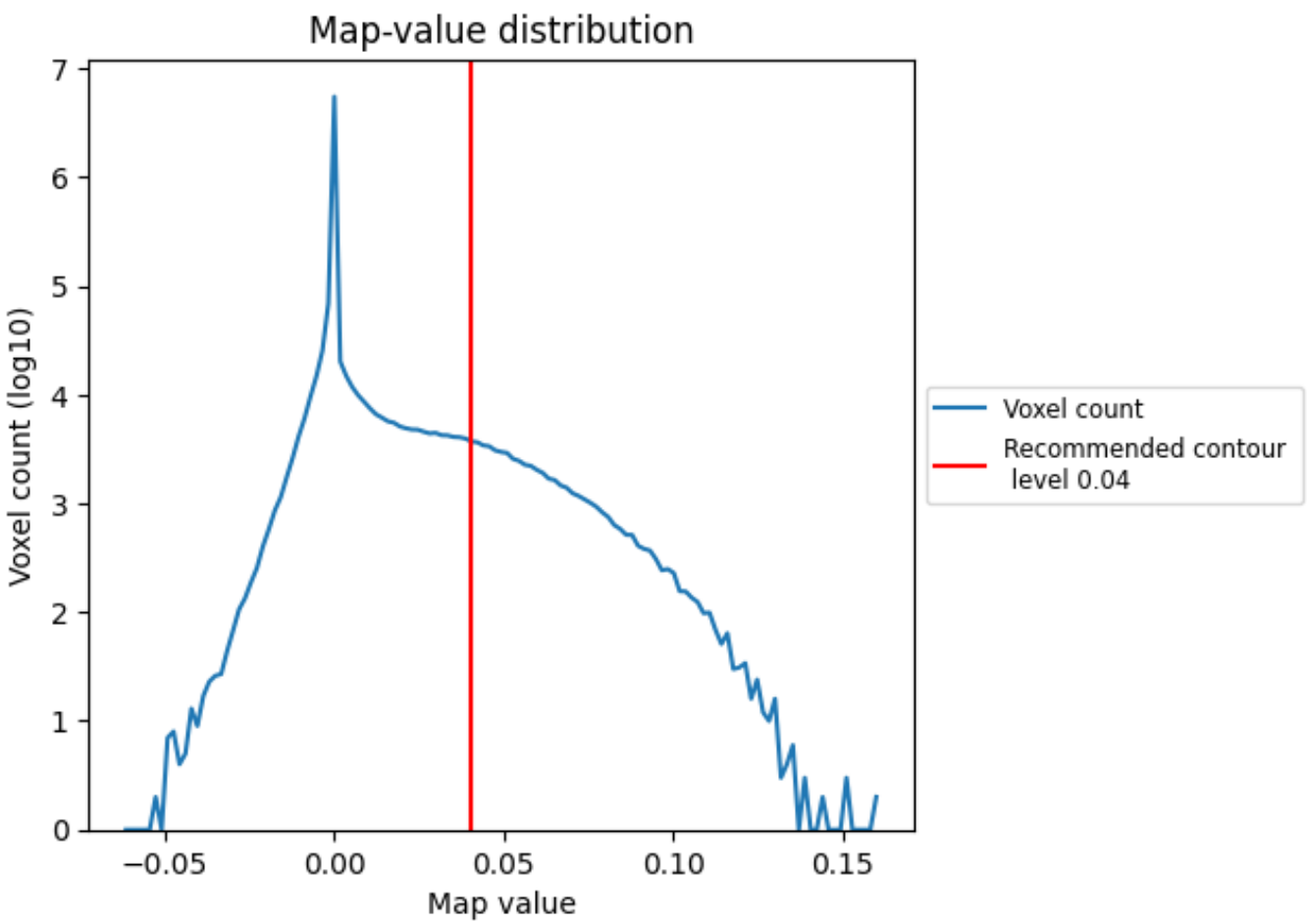
6.5 Mask visualisation (i)

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

7 Map analysis (i)

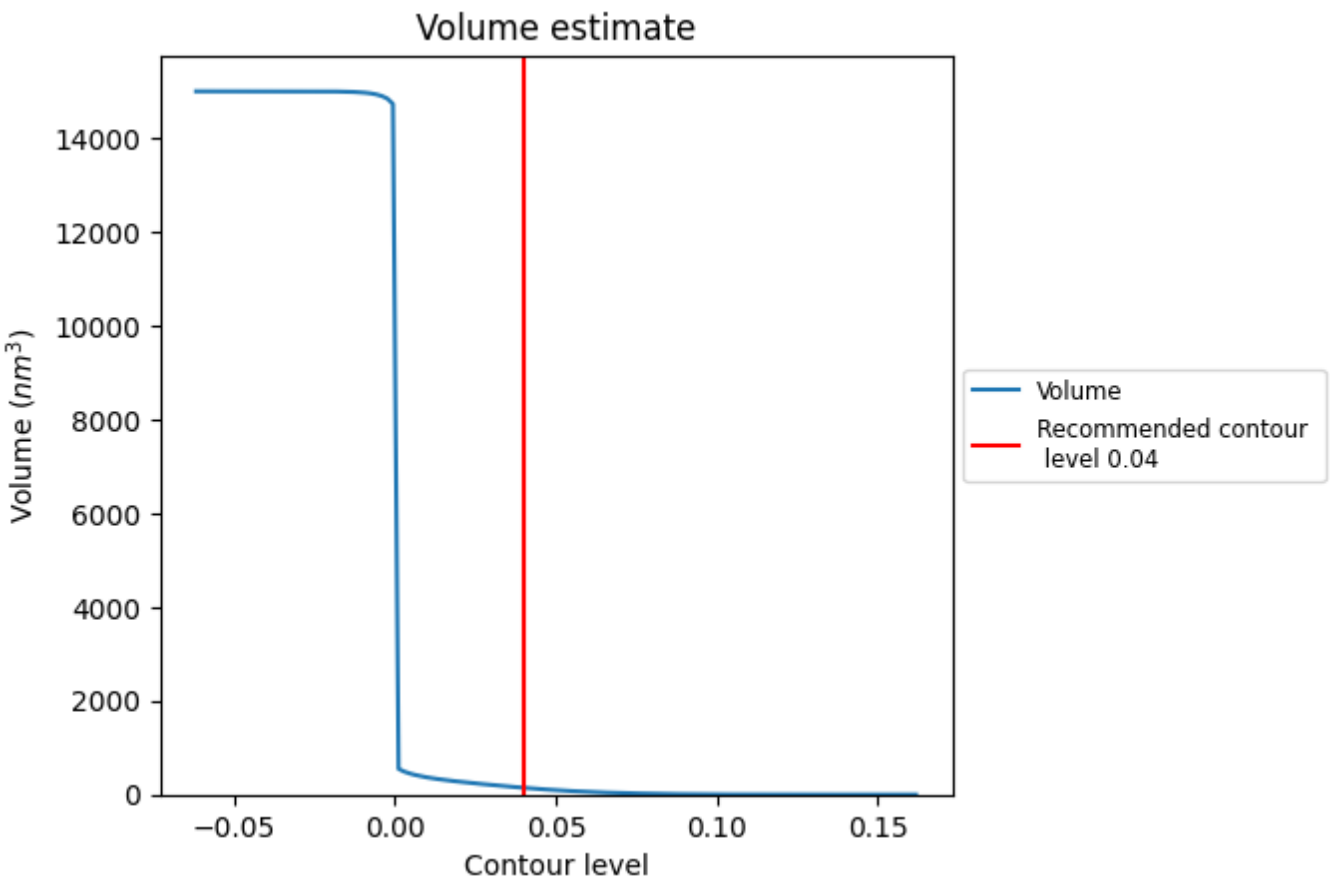
This section contains the results of statistical analysis of the map.

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.

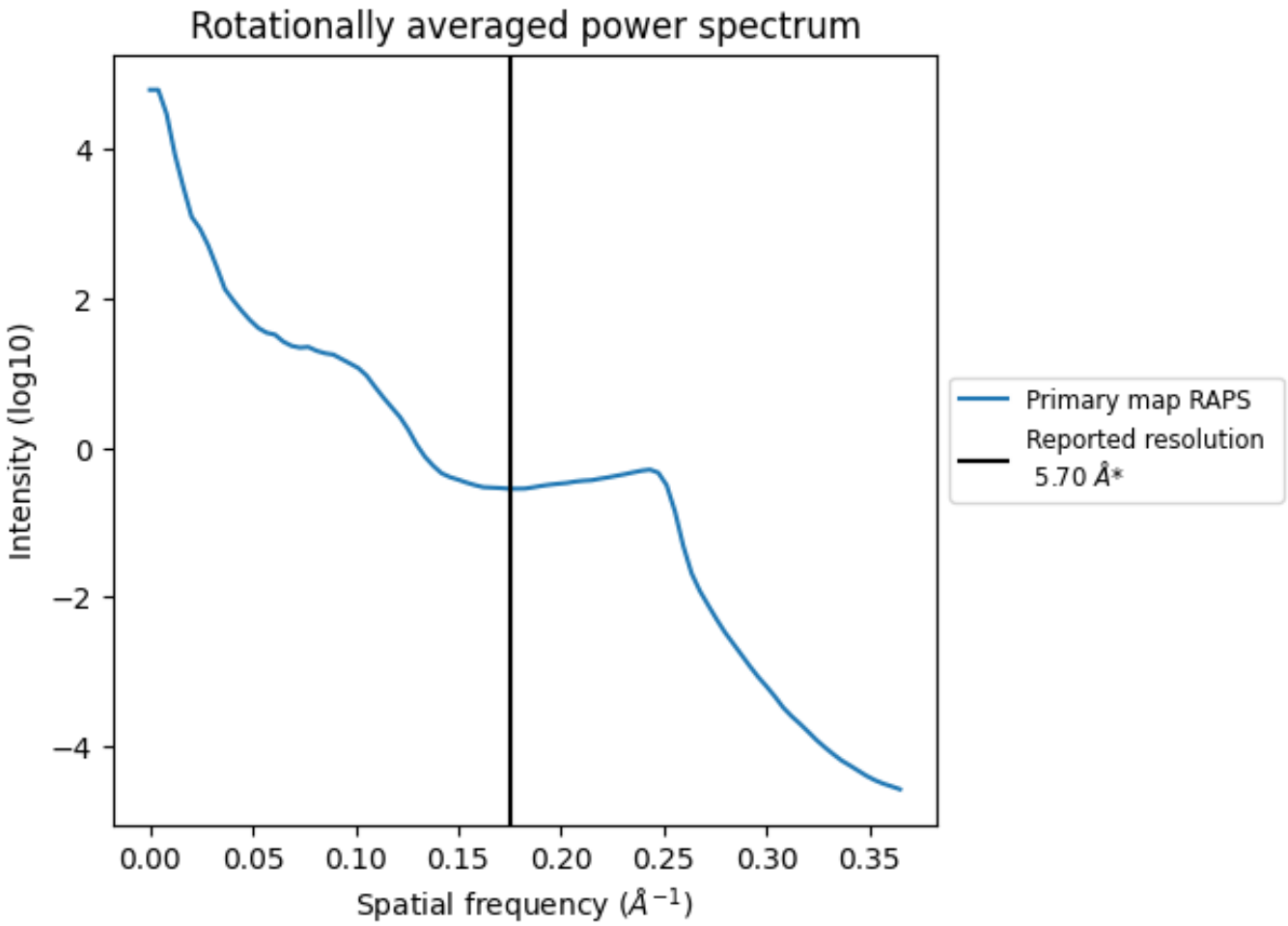
7.2 Volume estimate (i)



The volume at the recommended contour level is 145 nm^3; this corresponds to an approximate mass of 131 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 0.175 Å^{-1}

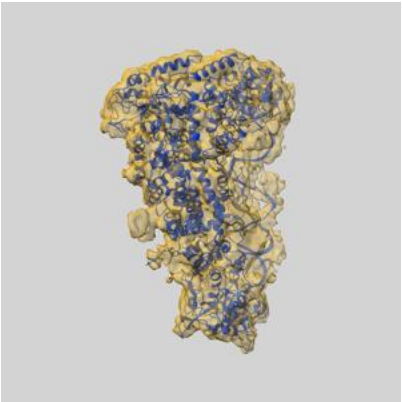
8 Fourier-Shell correlation (i)

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

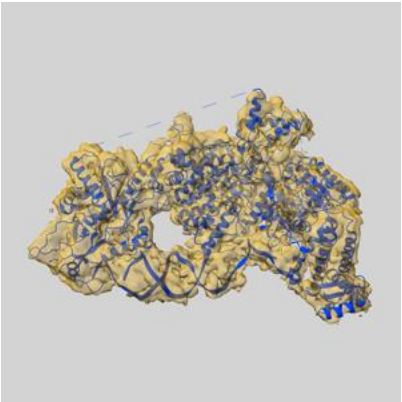
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-6906 and PDB model 5ZAM. Per-residue inclusion information can be found in section 3 on page 4.

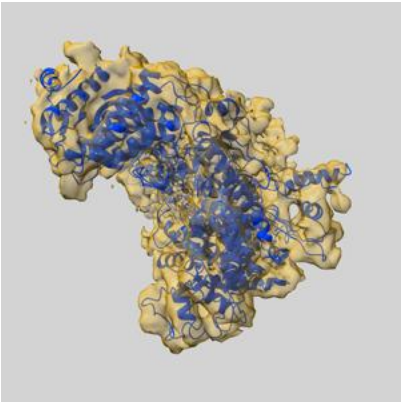
9.1 Map-model overlay (i)



X



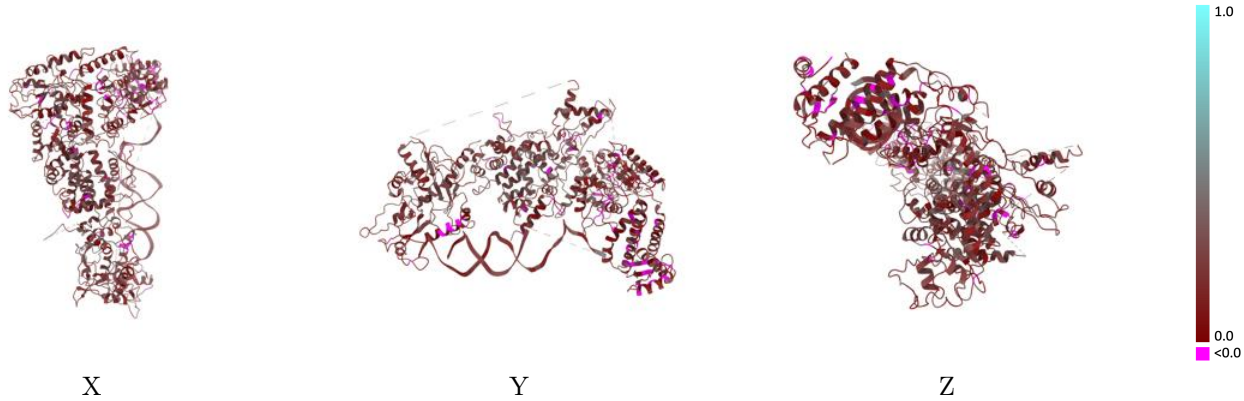
Y



Z

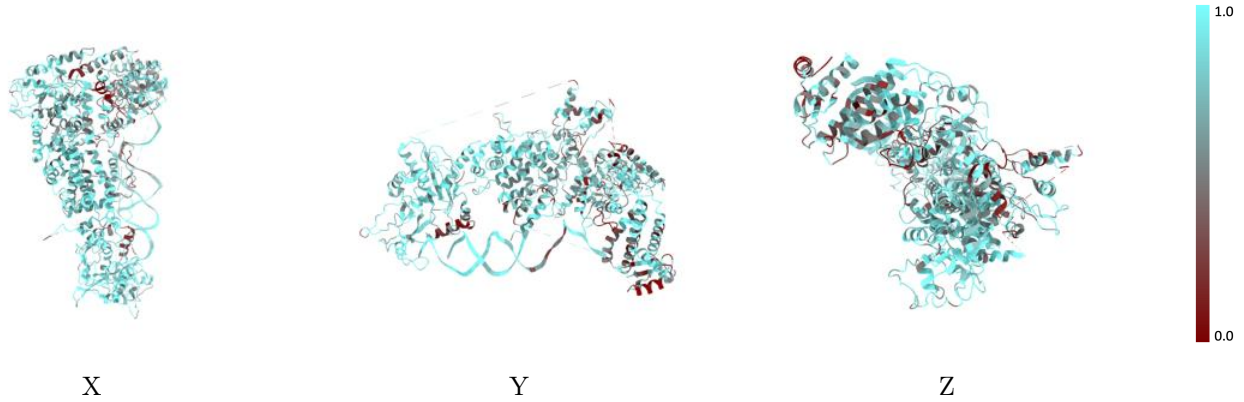
The images above show the 3D surface view of the map at the recommended contour level 0.04 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)



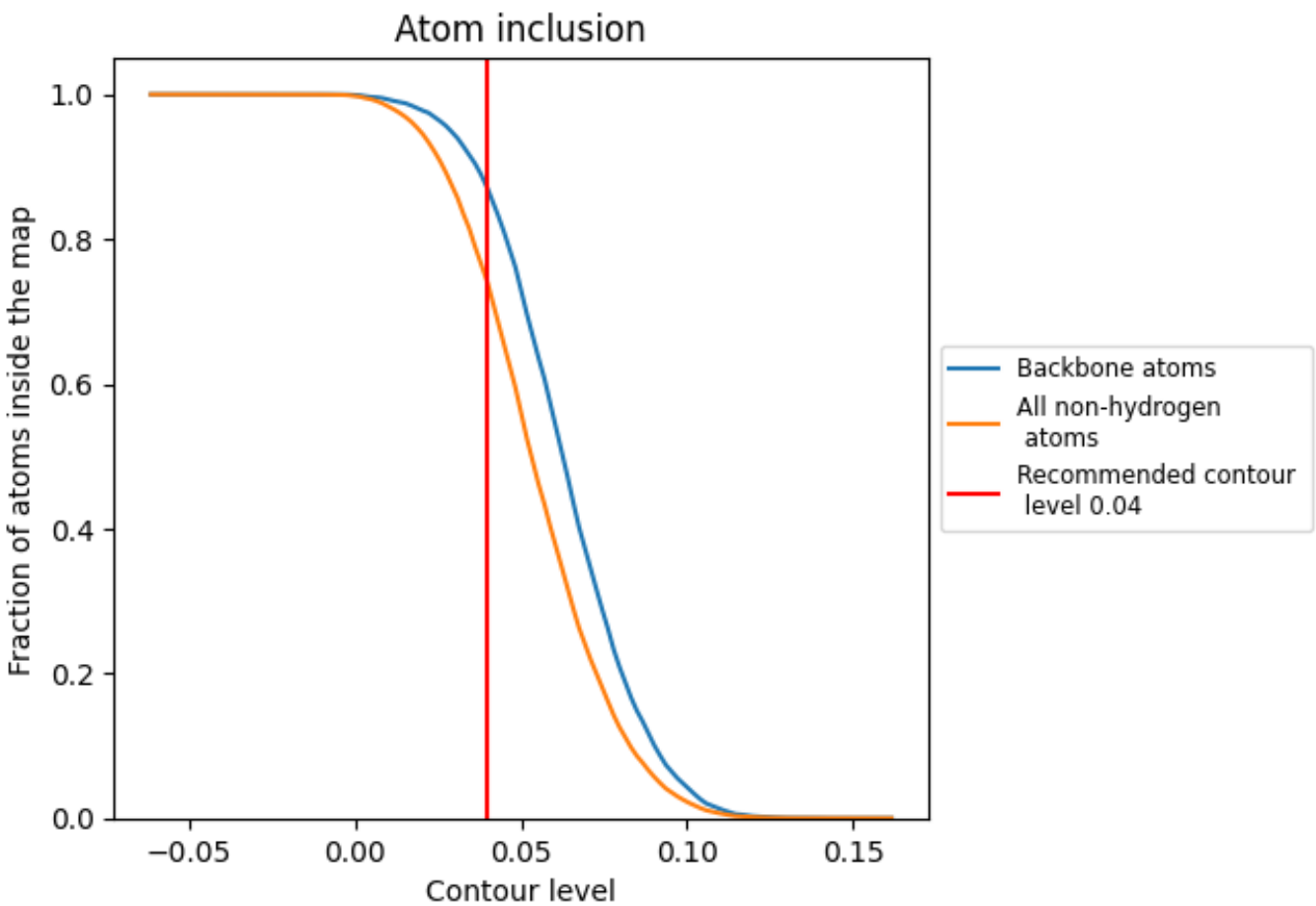
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.04).

9.4 Atom inclusion (i)



At the recommended contour level, 87% of all backbone atoms, 74% 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 (0.04) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7389	0.2150
A	0.7431	0.2180
B	0.5730	0.1620
C	0.7889	0.2160