wwPDB EM Validation Summary Report (i)

Mar 9, 2024 – 07:45 AM EST

PDB ID	:	6OF4
EMDB ID	:	EMD-20042
Title	:	Precursor ribosomal RNA processing complex, apo-state.
Authors	:	Pillon, M.C.; Hsu, A.L.; Krahn, J.M.; Williams, J.G.; Goslen, K.H.; Sobhany, M.; Borgnia, M.J.; Stanley, R.E.
Deposited on	:	2019-03-28
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:

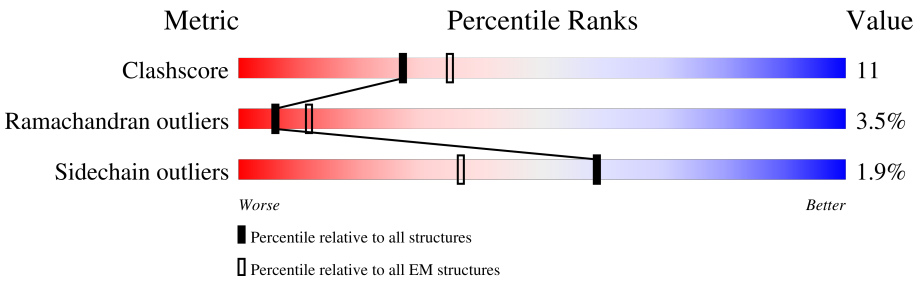
EMDB validation analysis	:	0.0.1.dev70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.20 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 <=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	391	• 42% 5% • 53%
1	D	391	• 41% 5% • 53%
2	B	640	9% 60% 20% • 18%
2	E	640	9% 59% 21% • 18%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 10998 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 Ribonuclease.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	183	1451	928	258	259	6	0	0
1	D	183	1451	928	258	259	6	0	0

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	initiating methionine	UNP G0SGE9
A	-26	SER	-	expression tag	UNP G0SGE9
A	-25	TYR	-	expression tag	UNP G0SGE9
A	-24	TYR	-	expression tag	UNP G0SGE9
A	-23	HIS	-	expression tag	UNP G0SGE9
A	-22	HIS	-	expression tag	UNP G0SGE9
A	-21	HIS	-	expression tag	UNP G0SGE9
A	-20	HIS	-	expression tag	UNP G0SGE9
A	-19	HIS	-	expression tag	UNP G0SGE9
A	-18	HIS	-	expression tag	UNP G0SGE9
A	-17	ASP	-	expression tag	UNP G0SGE9
A	-16	TYR	-	expression tag	UNP G0SGE9
A	-15	ASP	-	expression tag	UNP G0SGE9
A	-14	ILE	-	expression tag	UNP G0SGE9
A	-13	PRO	-	expression tag	UNP G0SGE9
A	-12	THR	-	expression tag	UNP G0SGE9
A	-11	THR	-	expression tag	UNP G0SGE9
A	-10	GLU	-	expression tag	UNP G0SGE9
A	-9	ASN	-	expression tag	UNP G0SGE9
A	-8	LEU	-	expression tag	UNP G0SGE9
A	-7	TYR	-	expression tag	UNP G0SGE9
A	-6	PHE	-	expression tag	UNP G0SGE9
A	-5	GLN	-	expression tag	UNP G0SGE9
A	-4	GLY	-	expression tag	UNP G0SGE9
A	-3	ALA	-	expression tag	UNP G0SGE9
A	-2	MET	-	expression tag	UNP G0SGE9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP G0SGE9
A	0	SER	-	expression tag	UNP G0SGE9
D	-27	MET	-	initiating methionine	UNP G0SGE9
D	-26	SER	-	expression tag	UNP G0SGE9
D	-25	TYR	-	expression tag	UNP G0SGE9
D	-24	TYR	-	expression tag	UNP G0SGE9
D	-23	HIS	-	expression tag	UNP G0SGE9
D	-22	HIS	-	expression tag	UNP G0SGE9
D	-21	HIS	-	expression tag	UNP G0SGE9
D	-20	HIS	-	expression tag	UNP G0SGE9
D	-19	HIS	-	expression tag	UNP G0SGE9
D	-18	HIS	-	expression tag	UNP G0SGE9
D	-17	ASP	-	expression tag	UNP G0SGE9
D	-16	TYR	-	expression tag	UNP G0SGE9
D	-15	ASP	-	expression tag	UNP G0SGE9
D	-14	ILE	-	expression tag	UNP G0SGE9
D	-13	PRO	-	expression tag	UNP G0SGE9
D	-12	THR	-	expression tag	UNP G0SGE9
D	-11	THR	-	expression tag	UNP G0SGE9
D	-10	GLU	-	expression tag	UNP G0SGE9
D	-9	ASN	-	expression tag	UNP G0SGE9
D	-8	LEU	-	expression tag	UNP G0SGE9
D	-7	TYR	-	expression tag	UNP G0SGE9
D	-6	PHE	-	expression tag	UNP G0SGE9
D	-5	GLN	-	expression tag	UNP G0SGE9
D	-4	GLY	-	expression tag	UNP G0SGE9
D	-3	ALA	-	expression tag	UNP G0SGE9
D	-2	MET	-	expression tag	UNP G0SGE9
D	-1	GLY	-	expression tag	UNP G0SGE9
D	0	SER	-	expression tag	UNP G0SGE9

- Molecule 2 is a protein called CLP1_P domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	528	4048	2586	702	739	21	0	0
2	E	528	4048	2586	702	739	21	0	0

There are 2 discrepancies between the modelled and reference sequences:

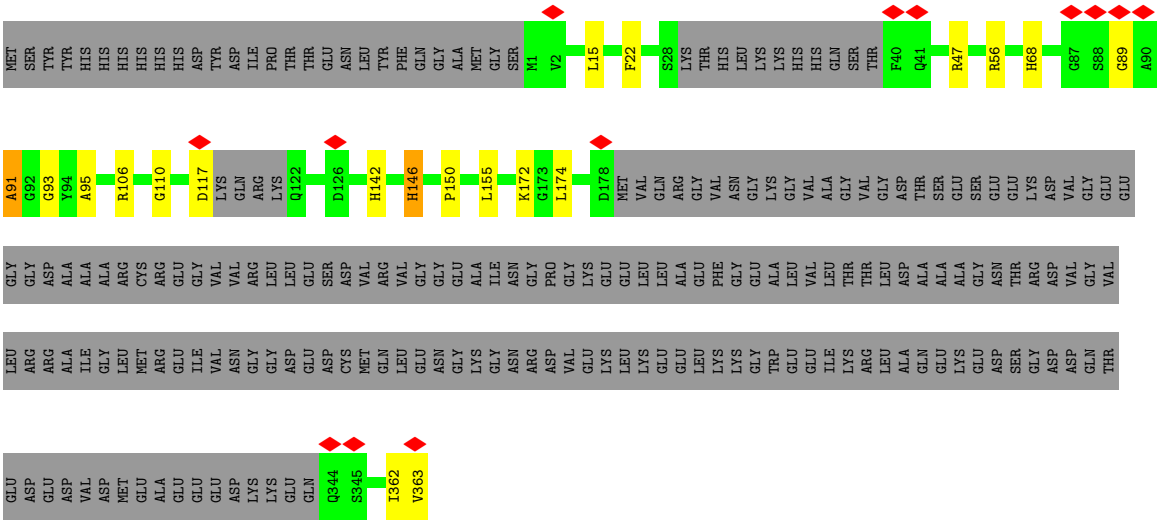
Chain	Residue	Modelled	Actual	Comment	Reference
B	109	MET	-	initiating methionine	UNP G0S263
E	109	MET	-	initiating methionine	UNP G0S263

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.

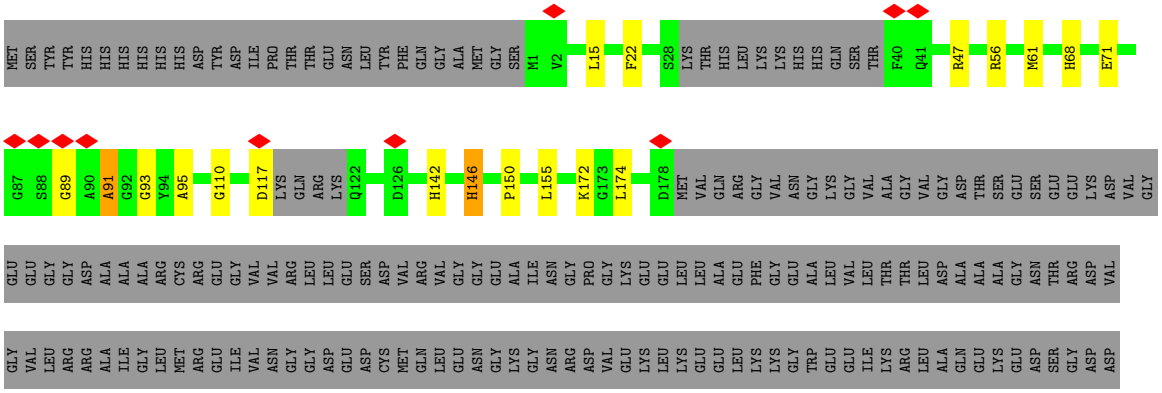
- Molecule 1: Ribonuclease

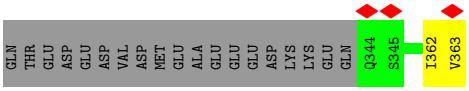
Chain A:



- Molecule 1: Ribonuclease

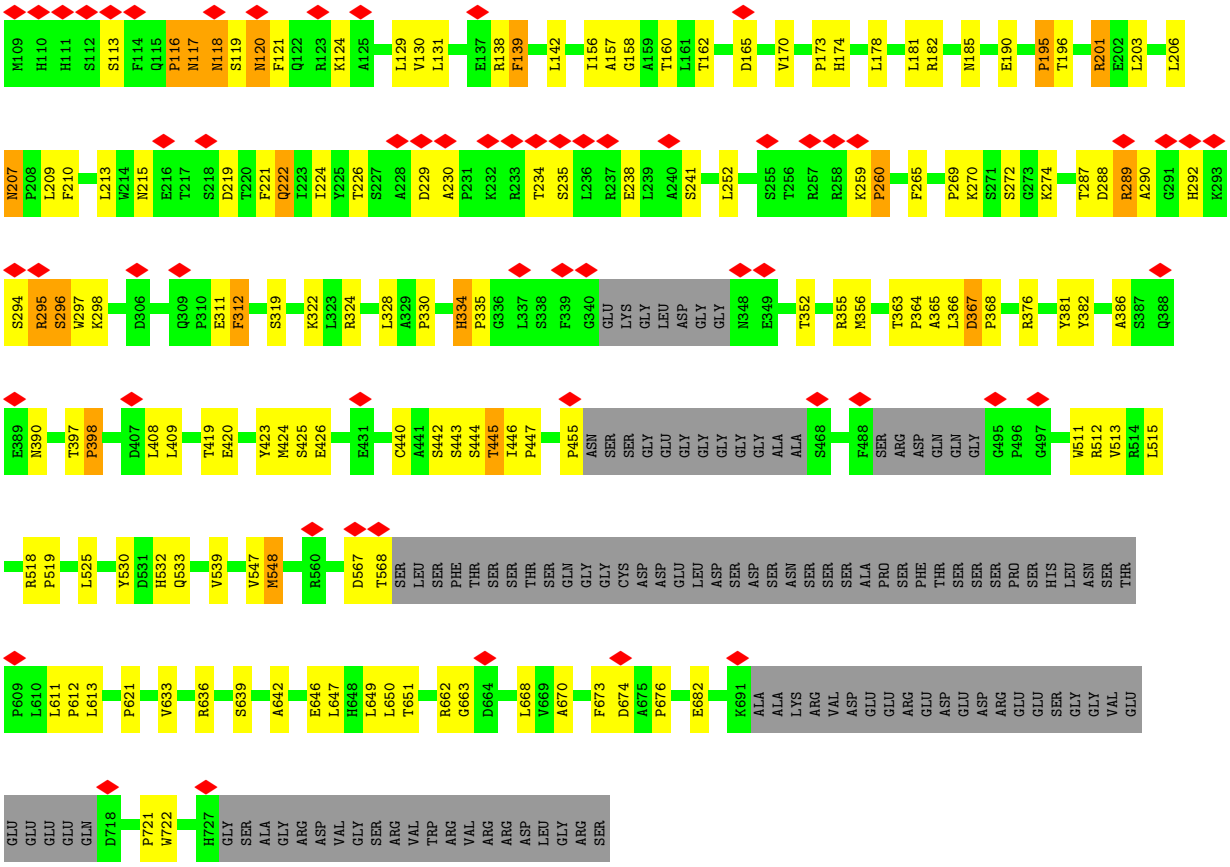
Chain D:





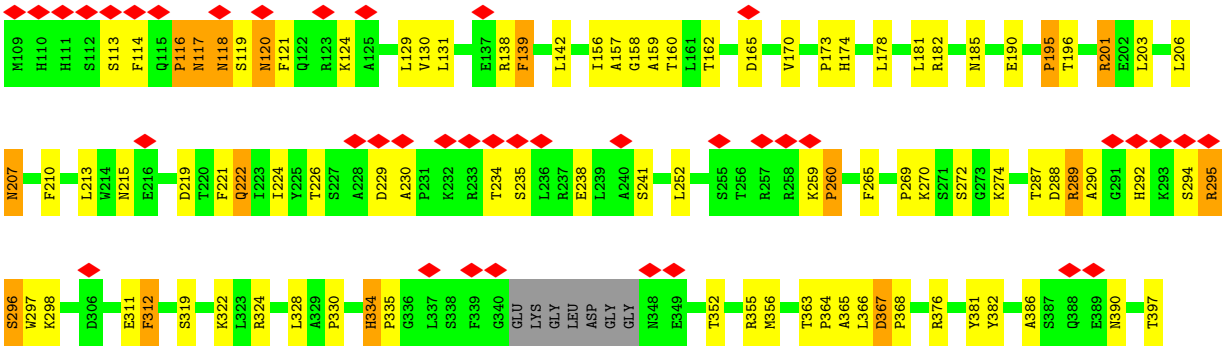
• Molecule 2: CLP1_P domain-containing protein

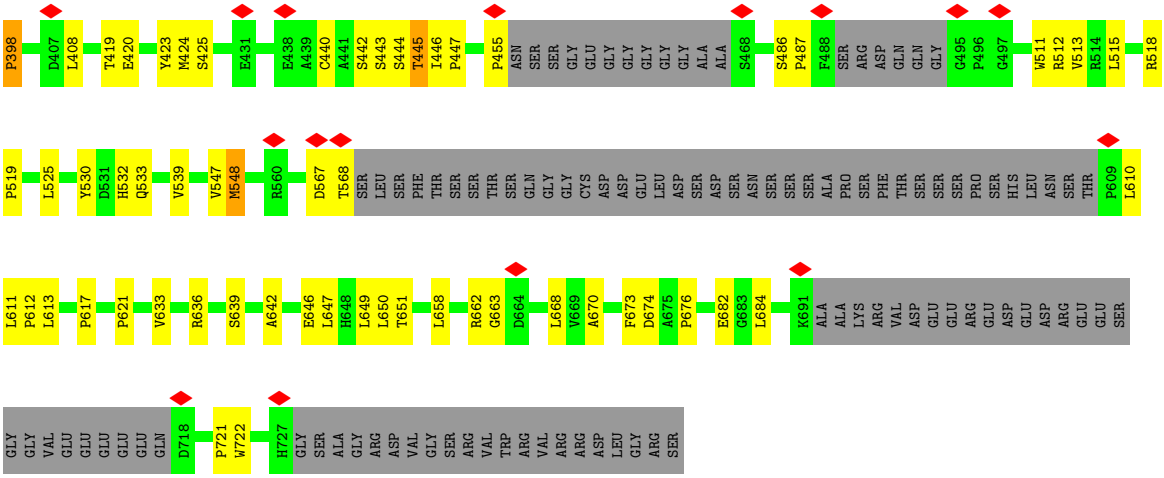
Chain B:



• Molecule 2: CLP1_P domain-containing protein

Chain E:




P398
D407
L408
T419
E420
Y423
M424
S425
E431
E438
A439
C440
A441
S442
S443
S444
T445
I446
P447
P455
ASN
SER
SER
GLY
GLU
GLY
GLY
GLY
GLY
GLY
ALA
ALA
S468
S486
P487
F488
SER
ARG
ASP
GLN
GLN
GLY
G495
P496
G497
W511
R512
V513
R514
L515
R518
P519
L525
Y530
D531
H532
Q533
V539
V547
M548
R560
D567
T568
SER
LEU
SER
PHE
THR
SER
SER
THR
SER
GLN
GLY
GLY
CYS
ASP
ASP
GLU
LEU
ASP
SER
ASP
SER
ASN
SER
SER
SER
ALA
PRO
SER
PHE
THR
SER
SER
SER
PRO
SER
HIS
LEU
ASN
SER
THR
P609
L610
L611
P612
L613
P617
P621
V633
R636
S639
A642
E646
L647
H648
L649
L650
T651
L658
R662
G663
D664
L668
V669
A670
F673
D674
A675
P676
E682
G683
L684
K691
ALA
ALA
LYS
ARG
VAL
ASP
GLU
GLU
ARG
GLU
ASP
GLU
ASP
ARG
GLU
GLU
SER
GLY
GLY
VAL
GLU
GLU
GLU
GLU
GLU
GLN
D718
P721
W722
H727
GLY
SER
ALA
GLY
ARG
ASP
VAL
GLY
SER
ARG
VAL
TRP
ARG
VAL
ARG
ARG
ASP
LEU
GLY
ARG
SER

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	102753	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 ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	28.728	Depositor
Minimum map value	-0.572	Depositor
Average map value	0.018	Depositor
Map value standard deviation	0.673	Depositor
Recommended contour level	4.5	Depositor
Map size (Å)	241.92001, 241.92001, 241.92001	wwPDB
Map dimensions	224, 224, 224	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.08, 1.08, 1.08	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	0.60	0/1490	0.57	1/2021 (0.0%)
1	D	0.60	0/1490	0.57	1/2021 (0.0%)
2	B	0.50	0/4158	0.63	1/5683 (0.0%)
2	E	0.50	0/4158	0.63	1/5683 (0.0%)
All	All	0.53	0/11296	0.62	4/15408 (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	1
1	D	0	1
2	B	0	2
2	E	0	2
All	All	0	6

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	146	HIS	C-N-CA	-6.07	106.53	121.70
1	D	146	HIS	C-N-CA	-6.07	106.54	121.70
2	E	455	PRO	N-CA-CB	6.01	110.52	103.30
2	B	455	PRO	N-CA-CB	5.98	110.48	103.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	91	ALA	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
2	B	295	ARG	Peptide
2	B	642	ALA	Peptide
1	D	91	ALA	Peptide
2	E	295	ARG	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	1451	0	1402	15	0
1	D	1451	0	1402	16	0
2	B	4048	0	4025	103	0
2	E	4048	0	4025	108	0
All	All	10998	0	10854	238	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 11.

The worst 5 of 238 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)
2:B:419:THR:HG22	2:B:420:GLU:HG3	1.54	0.89
2:E:419:THR:HG22	2:E:420:GLU:HG3	1.54	0.86
1:A:363:VAL:HG11	2:B:639:SER:HB3	1.58	0.84
2:B:130:VAL:HG22	2:B:190:GLU:HG2	1.63	0.80
2:E:130:VAL:HG22	2:E:190:GLU:HG2	1.63	0.80

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	175/391 (45%)	149 (85%)	26 (15%)	0	100	100
1	D	175/391 (45%)	149 (85%)	26 (15%)	0	100	100
2	B	516/640 (81%)	382 (74%)	110 (21%)	24 (5%)	2	17
2	E	516/640 (81%)	382 (74%)	109 (21%)	25 (5%)	2	17
All	All	1382/2062 (67%)	1062 (77%)	271 (20%)	49 (4%)	6	24

5 of 49 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	118	ASN
2	B	157	ALA
2	B	398	PRO
2	B	662	ARG
2	E	118	ASN

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	144/316 (46%)	142 (99%)	2 (1%)	67	86
1	D	144/316 (46%)	142 (99%)	2 (1%)	67	86
2	B	446/542 (82%)	437 (98%)	9 (2%)	55	80
2	E	446/542 (82%)	437 (98%)	9 (2%)	55	80
All	All	1180/1716 (69%)	1158 (98%)	22 (2%)	59	81

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	201	ARG
2	E	229	ASP
2	E	222	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	289	ARG
2	B	229	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	146	HIS
2	E	115	GLN
2	E	327	ASN
2	E	207	ASN
2	E	222	GLN

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.

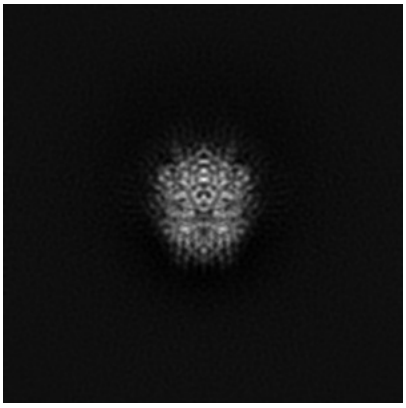
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-20042. These allow visual inspection of the internal detail of the map and identification of artifacts.

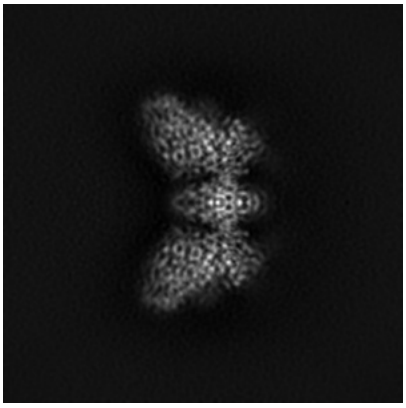
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

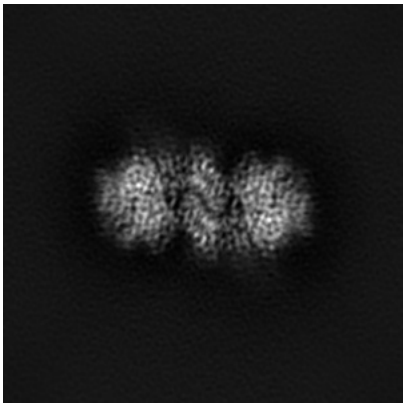
6.1.1 Primary map



X



Y



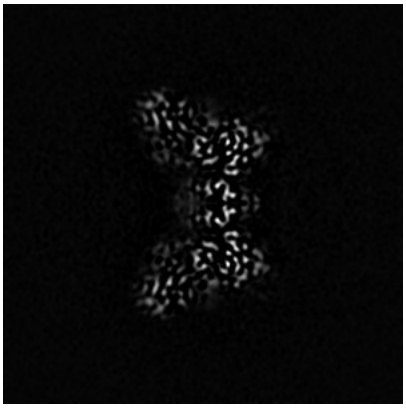
Z

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

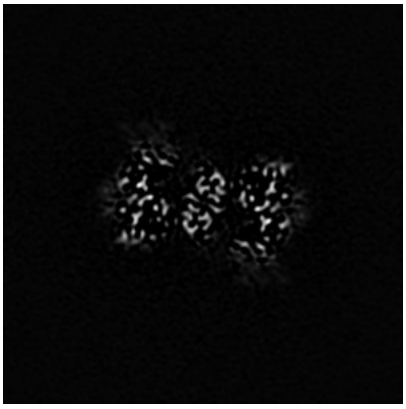
6.2 Central slices (i)

6.2.1 Primary map

X Index: 112



Y Index: 112

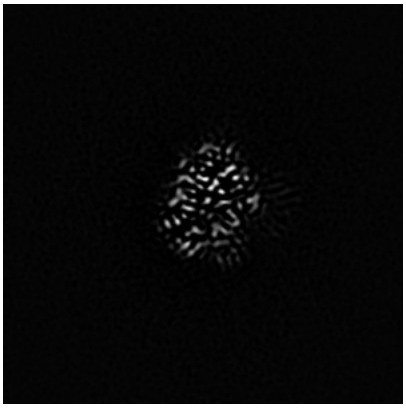


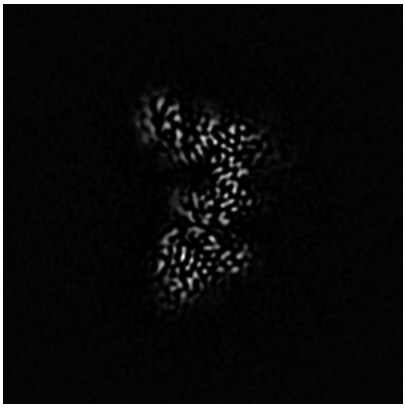
Z Index: 112

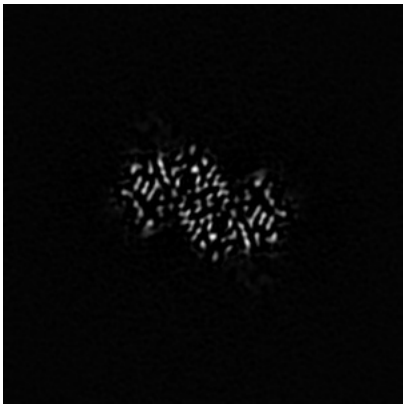
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map


X Index: 78

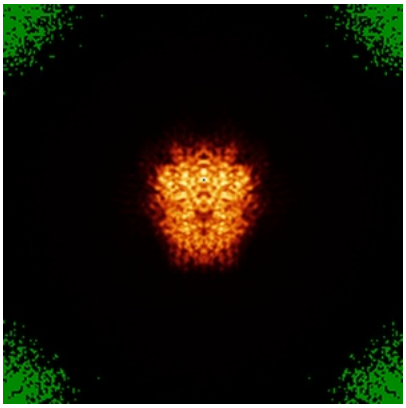

Y Index: 102

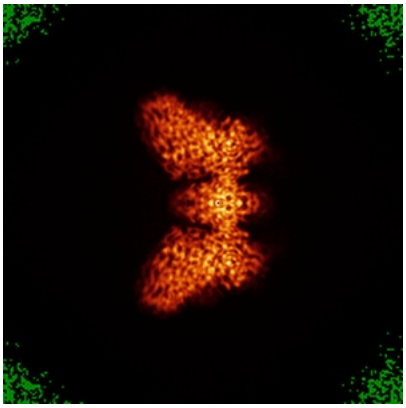

Z Index: 123

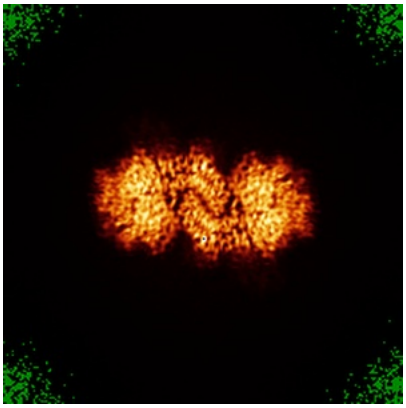
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) (i)

6.4.1 Primary map


X


Y


Z

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

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 4.5. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

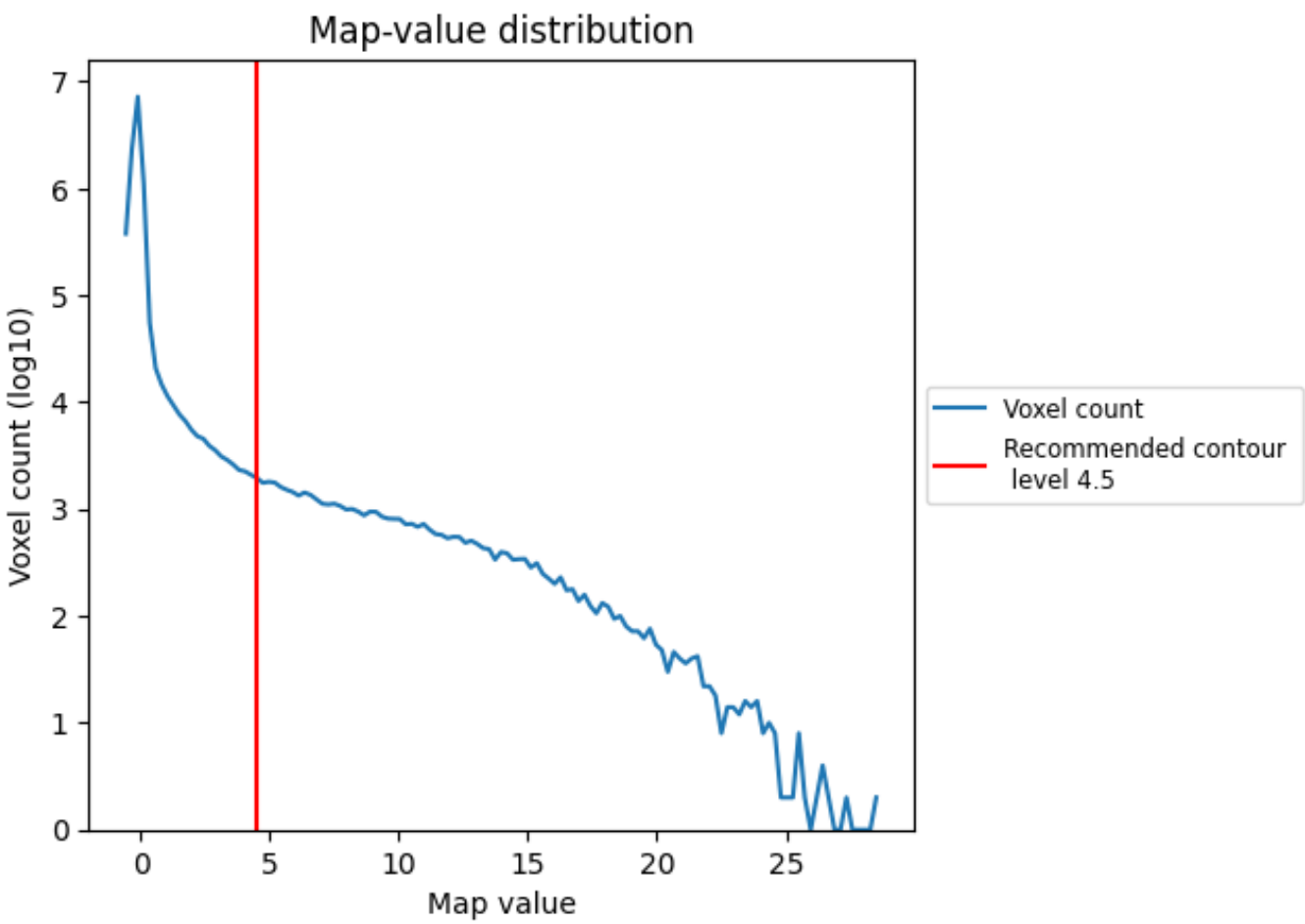
6.6 Mask visualisation (i)

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

7 Map analysis (i)

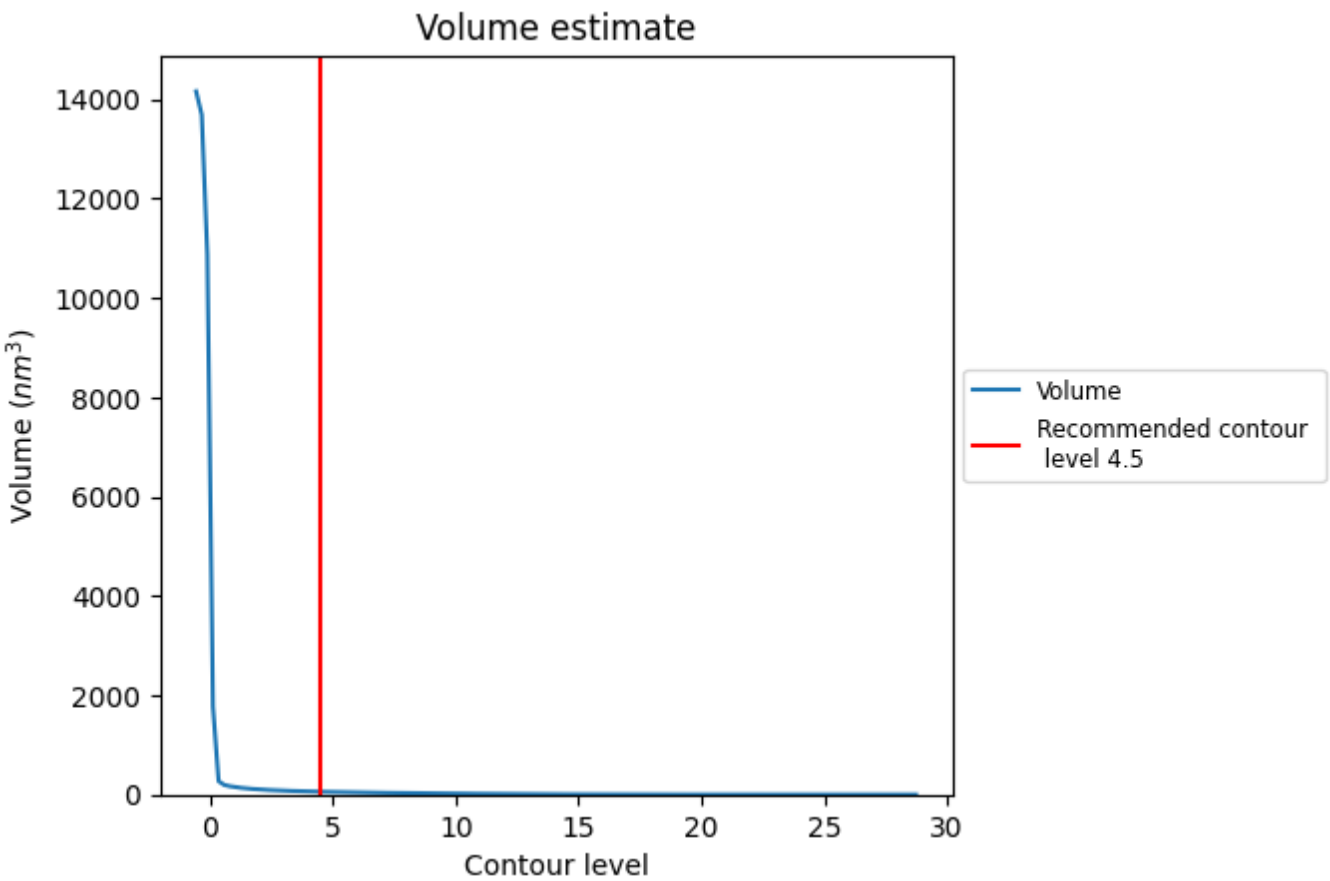
This section contains the results of statistical analysis of the map.

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.

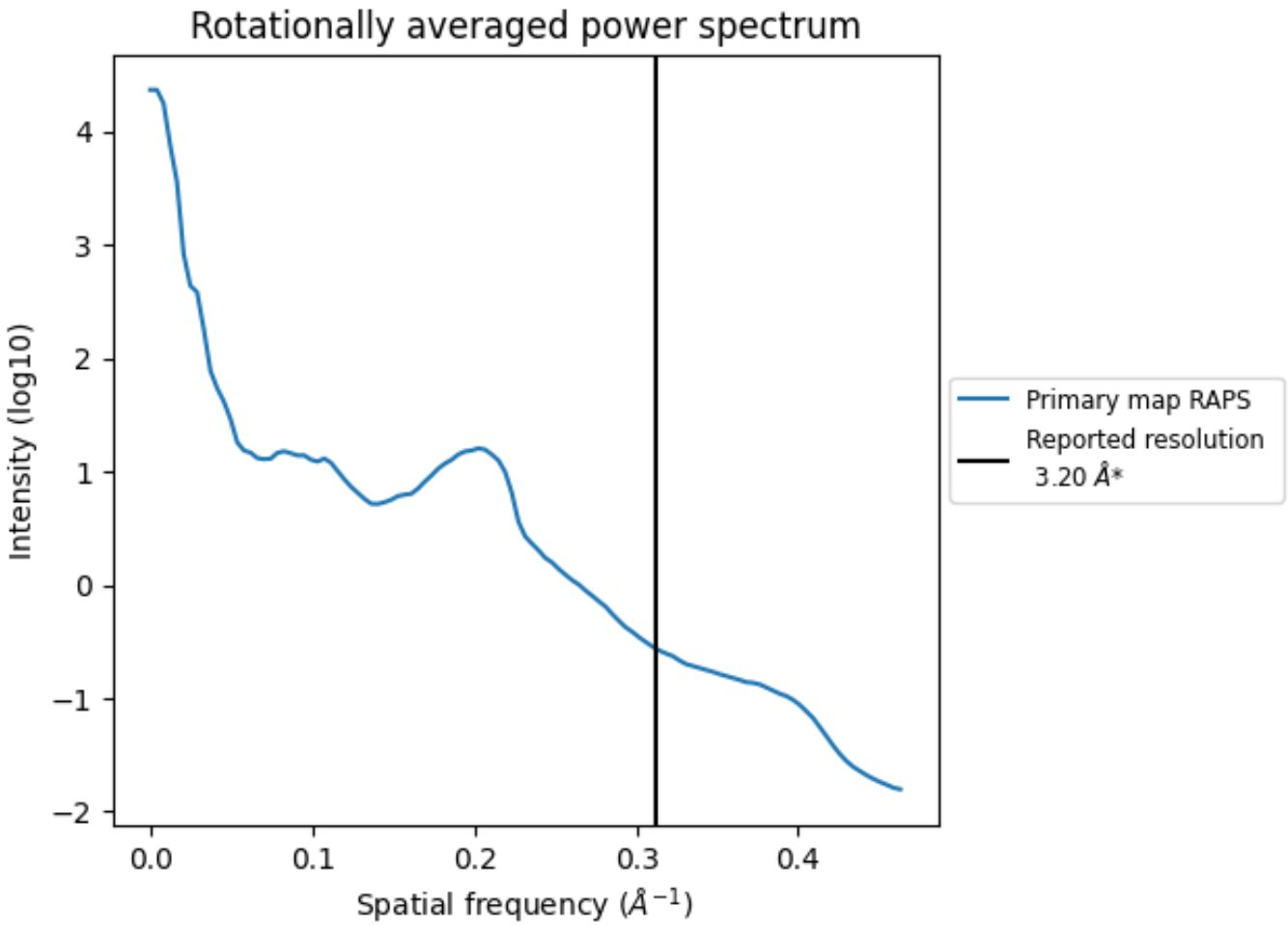
7.2 Volume estimate (i)



The volume at the recommended contour level is 57 nm^3; this corresponds to an approximate mass of 51 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}

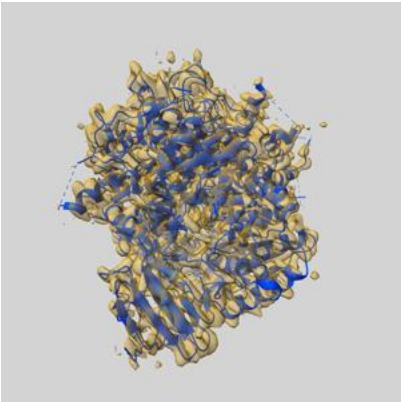
8 Fourier-Shell correlation (i)

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

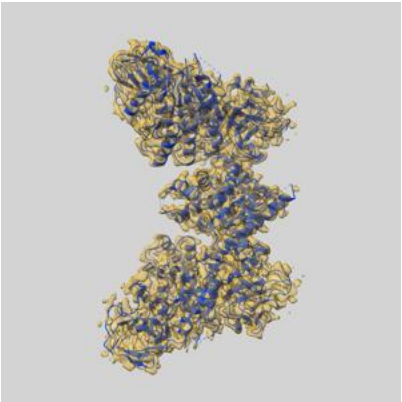
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-20042 and PDB model 6OF4. Per-residue inclusion information can be found in section 3 on page 6.

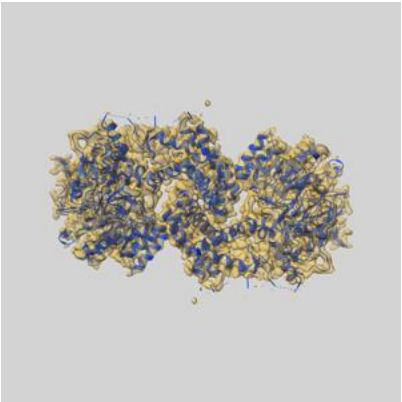
9.1 Map-model overlay (i)



X



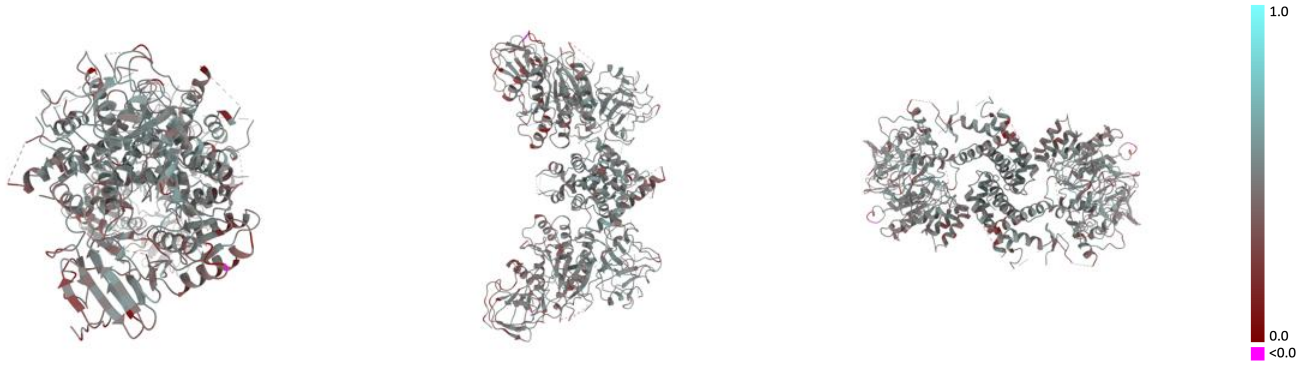
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 4.5 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.

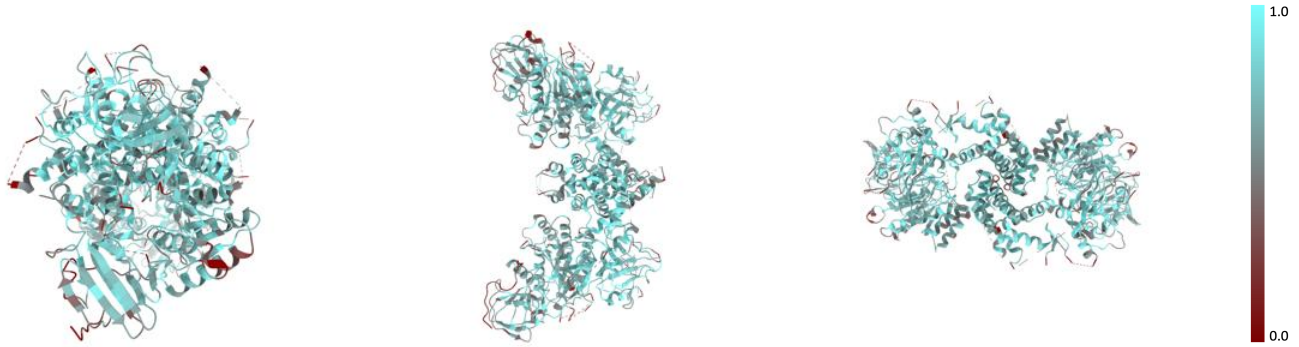
9.2 Q-score mapped to coordinate model (i)



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.

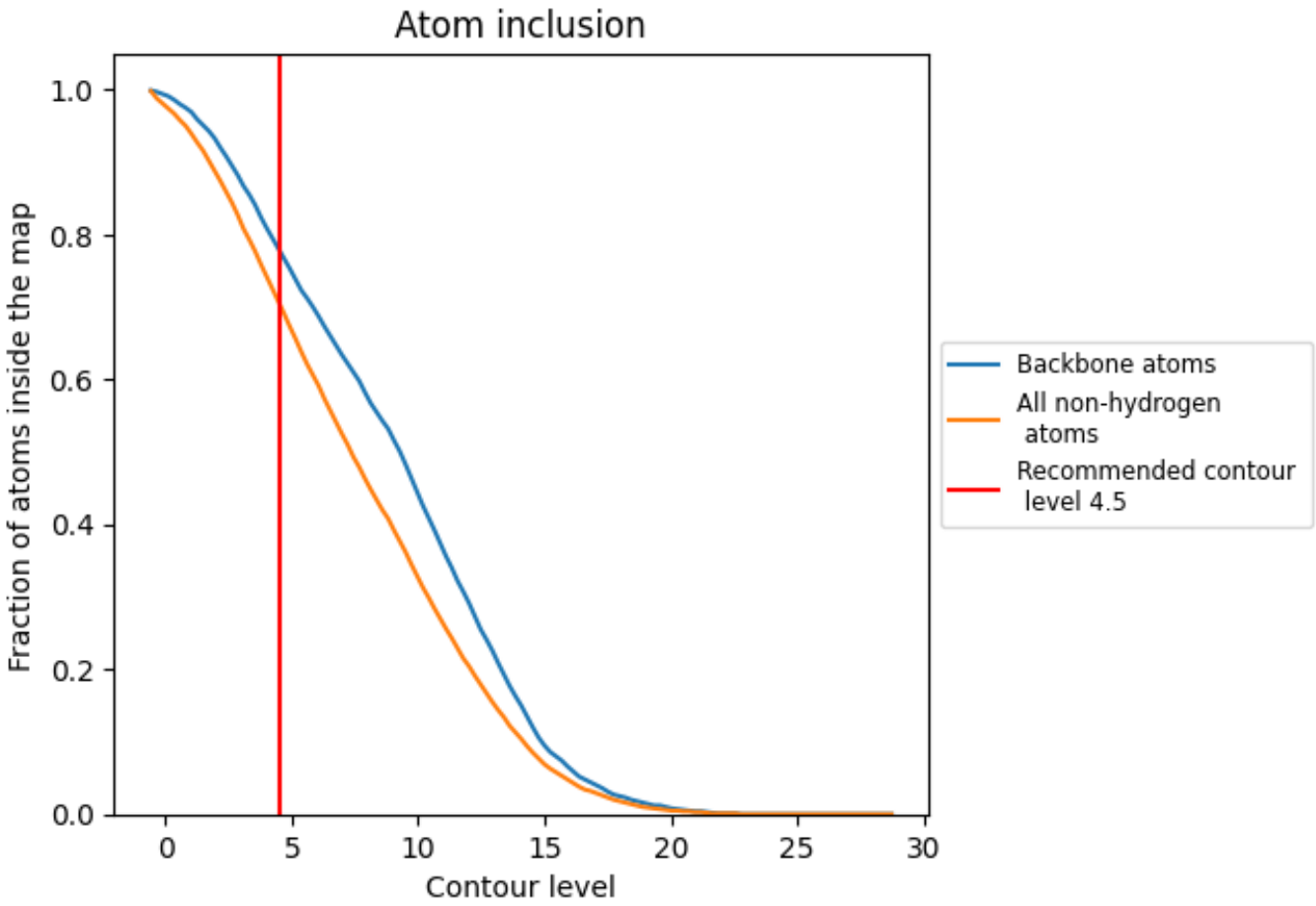
9.3 Atom inclusion mapped to coordinate model (i)



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 (4.5).

9.4 Atom inclusion (i)



At the recommended contour level, 78% of all backbone atoms, 70% 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 (4.5) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7050	0.4630
A	0.7470	0.4920
B	0.6900	0.4540
D	0.7470	0.4900
E	0.6910	0.4530