Full wwPDB EM Validation Report (i)

Feb 6, 2022 – 06:11 am GMT

PDB ID	:	7PEO
EMDB ID	:	EMD-13355
Title	:	Structure of the Caulobacter crescentus S-layer protein RsaA N-terminal domain bound to LPS and soaked with Holmium
Authors	:	von Kugelgen, A.; Bharat, T.A.M.
Deposited on	:	2021-08-11
Resolution	:	4.37 Å(reported)
Based on initial model	:	6T72

This is a Full wwPDB EM Validation 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 following versions of software and data (see references (i)) were used in the production of this report:

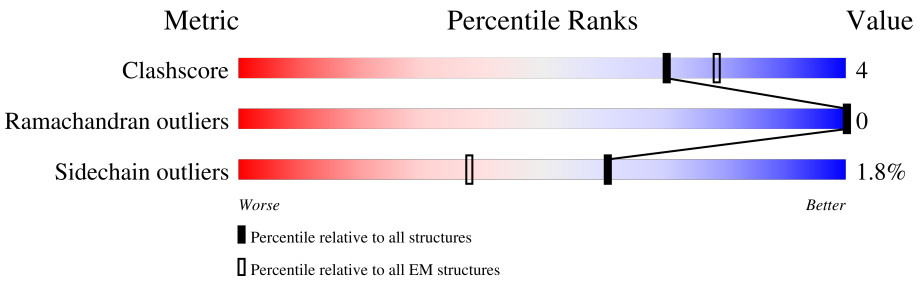
EMDB validation analysis	:	0.0.0.dev97
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 4.37 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	255	84% 11% 5%
2	B	12	42% 58%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1877 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.

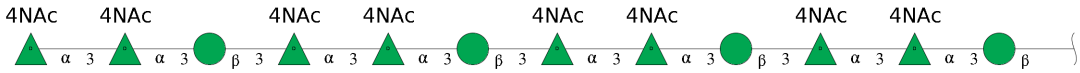
- Molecule 1 is a protein called S-layer protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	242	1726	1088	281	356	1	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	251	GLU	-	expression tag	UNP P35828
A	252	ASN	-	expression tag	UNP P35828
A	253	LEU	-	expression tag	UNP P35828
A	254	TYR	-	expression tag	UNP P35828
A	255	PHE	-	expression tag	UNP P35828
A	256	GLN	-	expression tag	UNP P35828

- Molecule 2 is an oligosaccharide called 4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose.



Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	B	12	148	88	8	52	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Ca	
3	A	2	2	2	0

- Molecule 4 is HOLMIUM ATOM (three-letter code: HO) (formula: Ho) (labeled as "Ligand of Interest" by depositor).

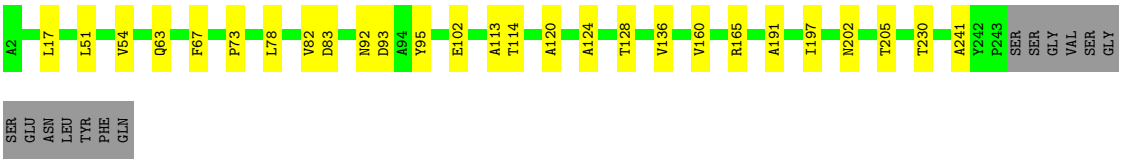
Mol	Chain	Residues	Atoms		AltConf
			Total	Ho	
4	A	1	1	1	0

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: S-layer protein

Chain A: 84% 11% 5%



• Molecule 2: 4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-4-acetamido-4,6-dideoxy-alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose

Chain B: 42% 58%

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	158430	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; RELION refinement with in-built CTF correction. The function is similar to a Wiener filter, so amplitude correction included.	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	44.8	Depositor
Minimum defocus (nm)	-1000	Depositor
Maximum defocus (nm)	-4000	Depositor
Magnification	130000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.031	Depositor
Minimum map value	-0.006	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.00765	Depositor
Map size (Å)	324.0, 324.0, 324.0	wwPDB
Map dimensions	300, 300, 300	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)

Bond lengths and bond angles in the following residue types are not validated in this section: HO, MRH, BMA, CA

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	0.30	0/1752	0.65	0/2407

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1726	0	1684	13	0
2	B	148	0	36	0	0
3	A	2	0	0	0	0
4	A	1	0	0	0	0
All	All	1877	0	1720	13	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 4.

All (13) 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:102:GLU:HA	1:A:202:ASN:HD22	1.53	0.72
1:A:230:THR:HG21	1:A:241:ALA:HB1	1.76	0.66
1:A:113:ALA:O	1:A:120:ALA:HA	1.95	0.65
1:A:83:ASP:HB2	1:A:92:ASN:HD21	1.69	0.58
1:A:124:ALA:O	1:A:128:THR:OG1	2.21	0.58
1:A:63:GLN:HE22	1:A:205:THR:HB	1.70	0.57
1:A:51:LEU:O	1:A:54:VAL:HG12	2.04	0.57
1:A:102:GLU:HA	1:A:202:ASN:ND2	2.22	0.53
1:A:67:PHE:CE2	1:A:197:ILE:HG21	2.49	0.47
1:A:17:LEU:HD12	1:A:82:VAL:HB	1.97	0.46
1:A:114:THR:HG21	1:A:191:ALA:HA	1.99	0.45
1:A:136:VAL:HG12	1:A:160:VAL:HG13	2.01	0.42
1:A:73:PRO:HG2	1:A:78:LEU:HG	2.03	0.41

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	240/255 (94%)	234 (98%)	6 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	168/179 (94%)	165 (98%)	3 (2%)	59 77

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	93	ASP
1	A	95	TYR
1	A	165	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	49	ASN
1	A	63	GLN
1	A	107	ASN
1	A	111	ASN
1	A	134	GLN
1	A	202	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)

12 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	BMA	B	1	2	11,11,12	0.93	0	15,15,17	1.58	4 (26%)
2	BMA	B	10	2	11,11,12	1.04	1 (9%)	15,15,17	1.71	2 (13%)
2	MRH	B	11	2	12,13,14	0.26	0	15,18,20	0.76	0
2	MRH	B	12	2	12,13,14	0.27	0	15,18,20	0.88	1 (6%)
2	MRH	B	2	2	12,13,14	0.27	0	15,18,20	0.52	0
2	MRH	B	3	2	12,13,14	0.25	0	15,18,20	0.67	0
2	BMA	B	4	2	11,11,12	0.93	0	15,15,17	1.51	3 (20%)
2	MRH	B	5	2	12,13,14	0.25	0	15,18,20	0.73	0
2	MRH	B	6	2	12,13,14	0.24	0	15,18,20	0.71	0
2	BMA	B	7	2	11,11,12	1.32	1 (9%)	15,15,17	1.99	4 (26%)
2	MRH	B	8	2	12,13,14	0.29	0	15,18,20	0.81	1 (6%)
2	MRH	B	9	2	12,13,14	0.29	0	15,18,20	1.00	1 (6%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BMA	B	1	2	-	0/2/19/22	0/1/1/1
2	BMA	B	10	2	-	2/2/19/22	0/1/1/1
2	MRH	B	11	2	-	0/4/21/24	0/1/1/1
2	MRH	B	12	2	-	3/4/21/24	0/1/1/1
2	MRH	B	2	2	-	0/4/21/24	0/1/1/1
2	MRH	B	3	2	-	0/4/21/24	0/1/1/1
2	BMA	B	4	2	-	2/2/19/22	0/1/1/1
2	MRH	B	5	2	-	2/4/21/24	0/1/1/1
2	MRH	B	6	2	-	0/4/21/24	0/1/1/1
2	BMA	B	7	2	-	2/2/19/22	0/1/1/1
2	MRH	B	8	2	-	2/4/21/24	0/1/1/1
2	MRH	B	9	2	-	0/4/21/24	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	7	BMA	C2-C3	3.43	1.57	1.52
2	B	10	BMA	C2-C3	2.58	1.56	1.52

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	7	BMA	C2-C3-C4	4.25	118.25	110.89
2	B	10	BMA	C2-C3-C4	4.03	117.88	110.89
2	B	10	BMA	C1-O5-C5	3.84	117.40	112.19
2	B	7	BMA	C1-C2-C3	3.73	114.25	109.67
2	B	4	BMA	C1-O5-C5	3.71	117.22	112.19
2	B	1	BMA	C1-O5-C5	3.55	117.00	112.19
2	B	7	BMA	C1-O5-C5	3.12	116.42	112.19
2	B	7	BMA	O2-C2-C3	2.98	116.11	110.14
2	B	9	MRH	C2-C3-C4	2.82	113.11	110.63
2	B	1	BMA	C2-C3-C4	2.64	115.47	110.89
2	B	4	BMA	C2-C3-C4	2.55	115.31	110.89
2	B	1	BMA	C3-C4-C5	2.39	114.50	110.24
2	B	12	MRH	C2-C3-C4	2.37	112.71	110.63
2	B	1	BMA	C1-C2-C3	2.19	112.36	109.67
2	B	4	BMA	C1-C2-C3	2.08	112.22	109.67
2	B	8	MRH	C2-C3-C4	2.05	112.44	110.63

There are no chirality outliers.

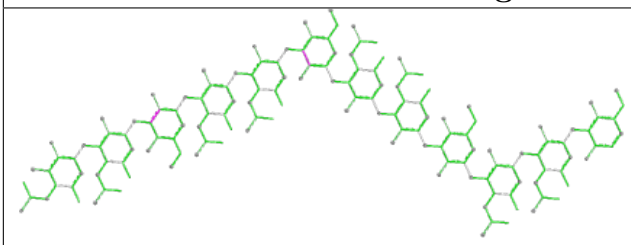
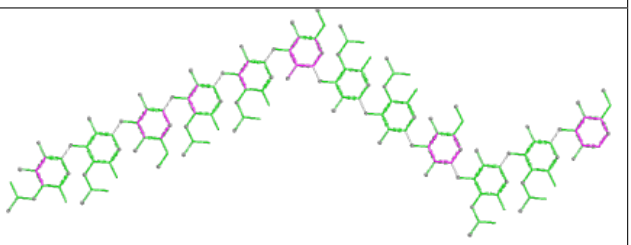
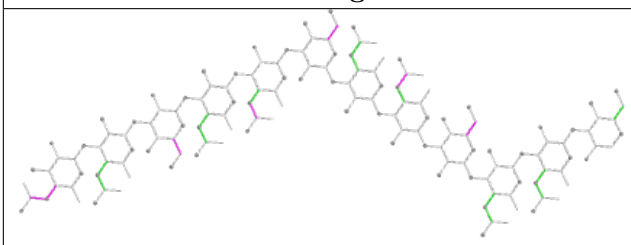
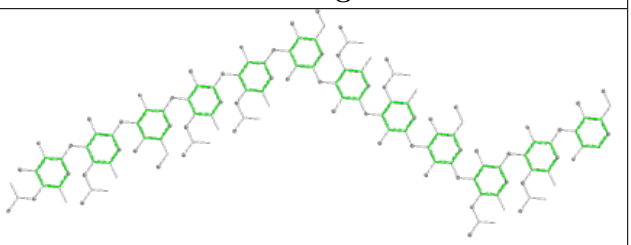
All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	12	MRH	C5-C4-N4-C7
2	B	12	MRH	C8-C7-N4-C4
2	B	12	MRH	O7-C7-N4-C4
2	B	10	BMA	C4-C5-C6-O6
2	B	10	BMA	O5-C5-C6-O6
2	B	7	BMA	O5-C5-C6-O6
2	B	4	BMA	O5-C5-C6-O6
2	B	8	MRH	C8-C7-N4-C4
2	B	7	BMA	C4-C5-C6-O6
2	B	8	MRH	O7-C7-N4-C4
2	B	4	BMA	C4-C5-C6-O6
2	B	5	MRH	C8-C7-N4-C4
2	B	5	MRH	O7-C7-N4-C4

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain B	
 Bond lengths	 Bond angles
 Torsions	 Rings

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.

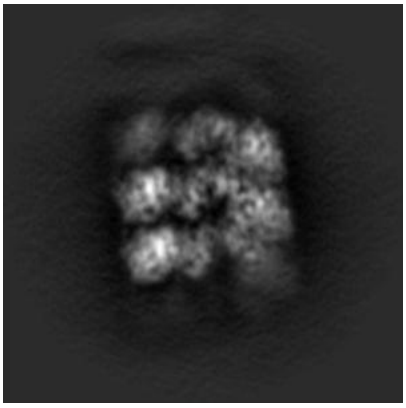
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-13355. These allow visual inspection of the internal detail of the map and identification of artifacts.

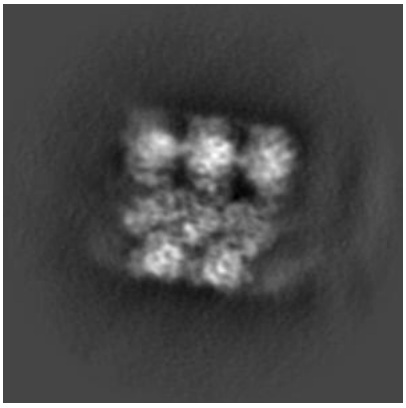
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

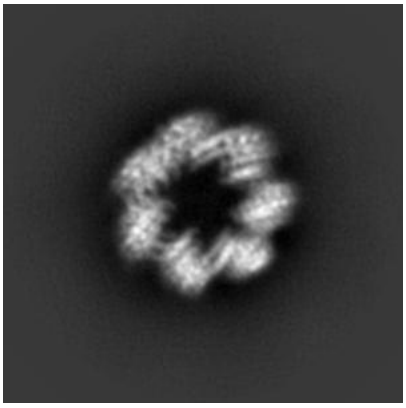
6.1.1 Primary map



X



Y

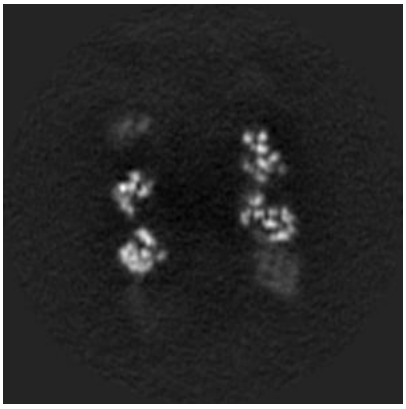


Z

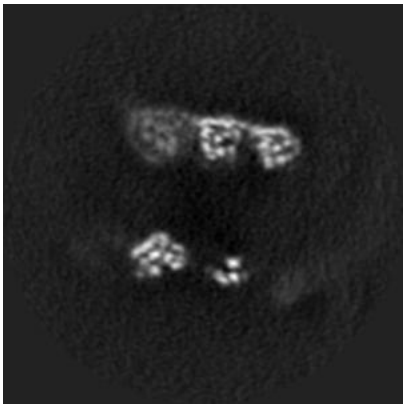
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

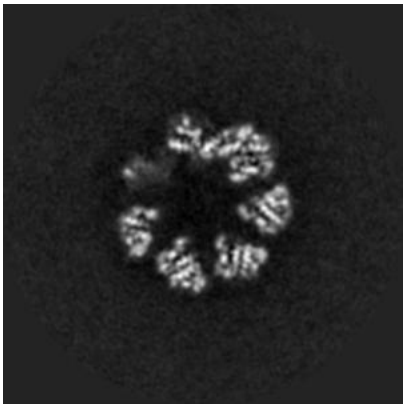
6.2.1 Primary map



X Index: 150



Y Index: 150

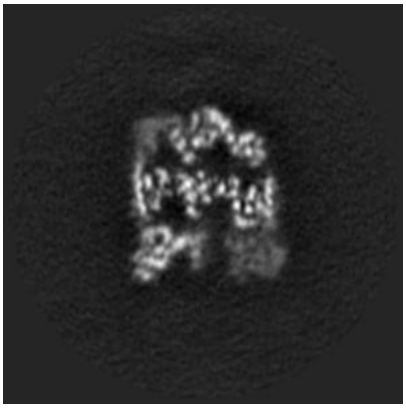


Z Index: 150

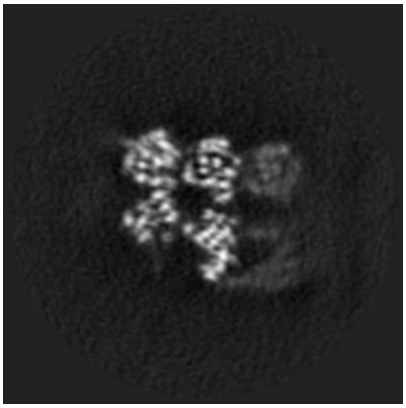
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

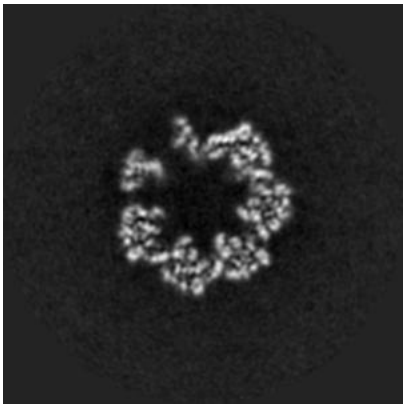
6.3.1 Primary map



X Index: 188



Y Index: 111

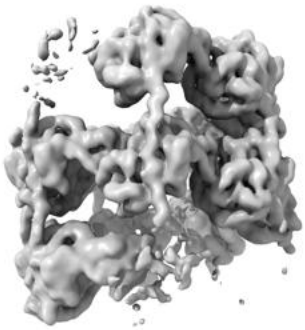


Z Index: 156

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

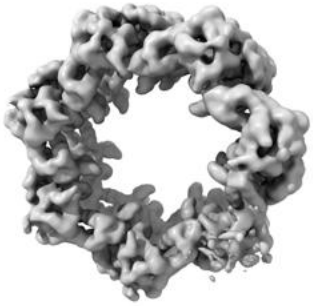
6.4 Orthogonal surface views (i)

6.4.1 Primary map



X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.00765. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

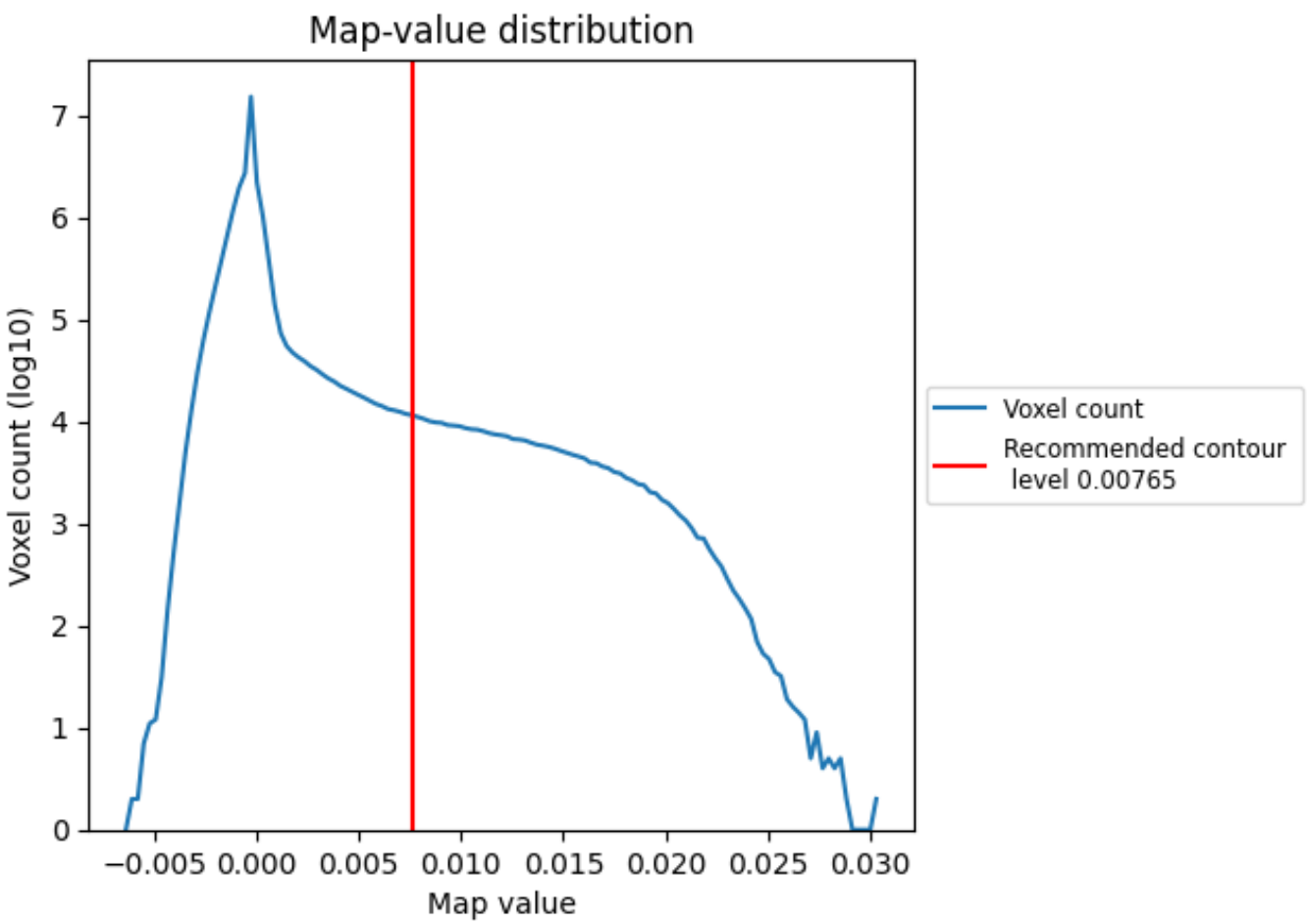
6.5 Mask visualisation (i)

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

7 Map analysis (i)

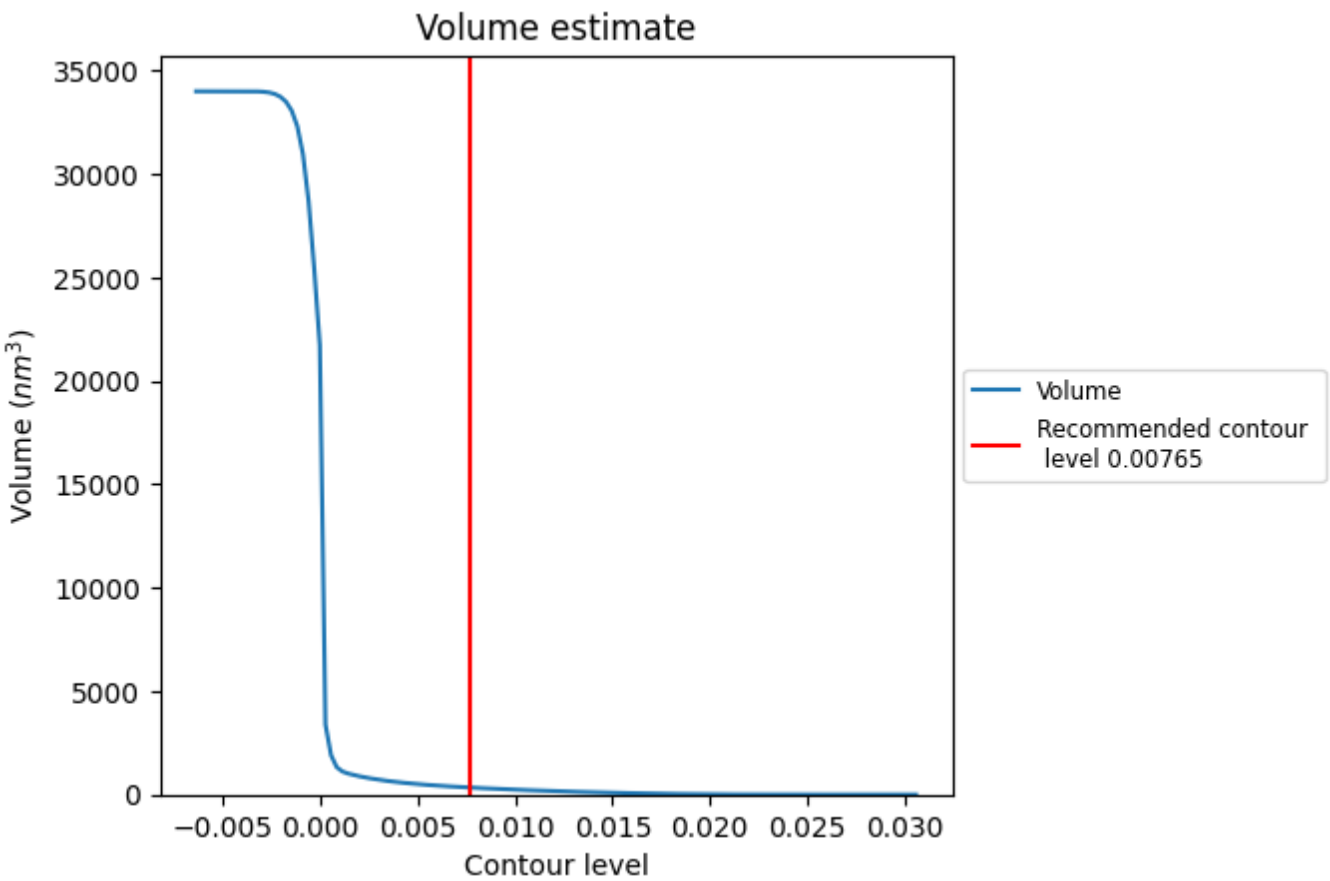
This section contains the results of statistical analysis of the map.

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.

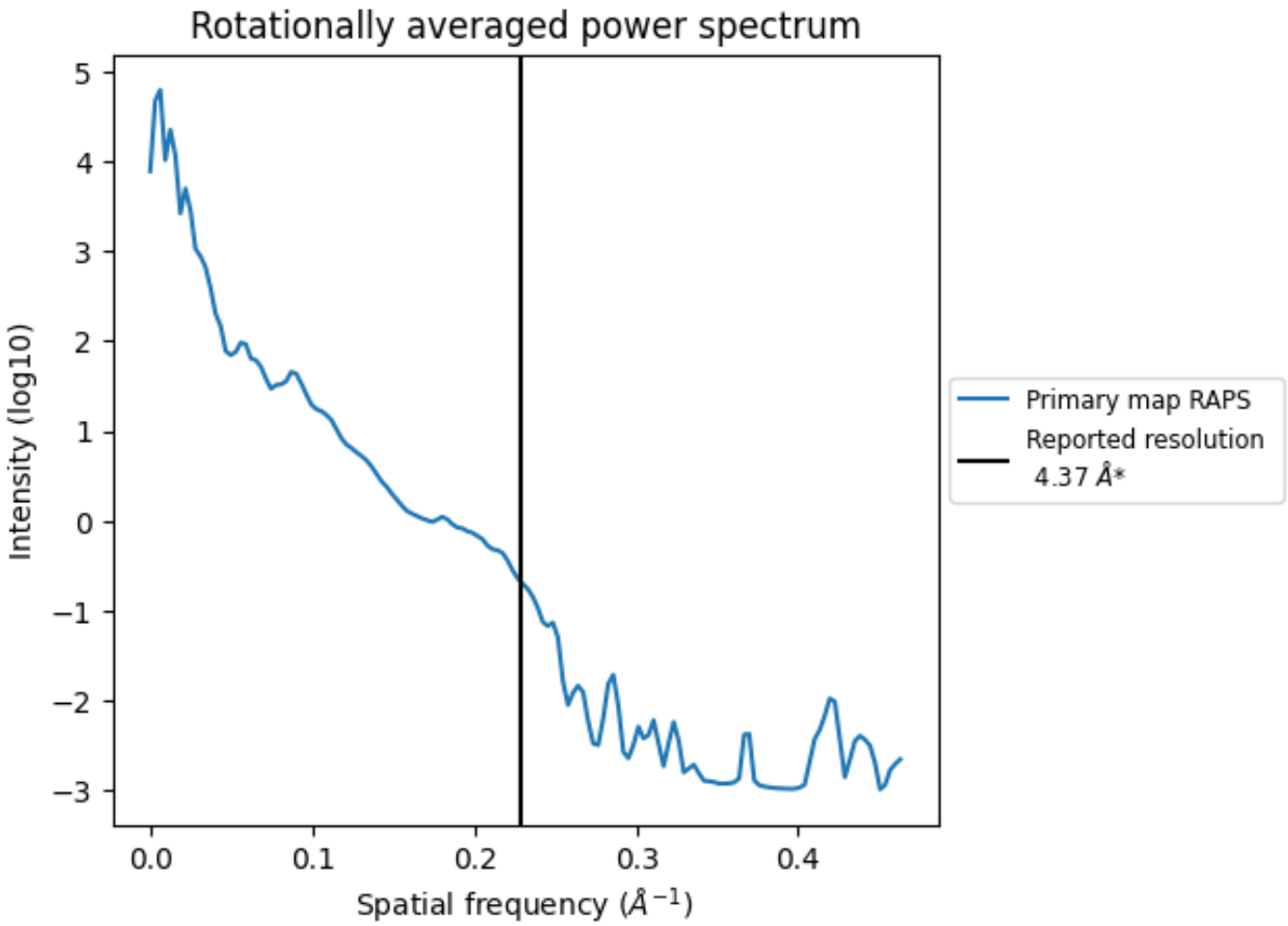
7.2 Volume estimate (i)



The volume at the recommended contour level is 340 nm^3; this corresponds to an approximate mass of 308 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 0.229 Å^{-1}

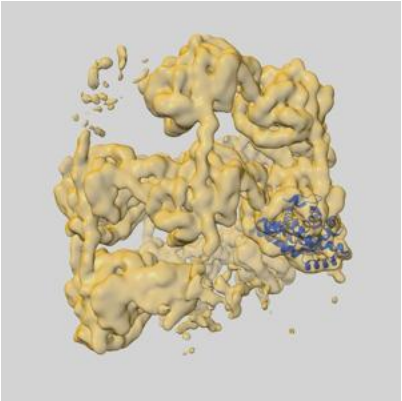
8 Fourier-Shell correlation (i)

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

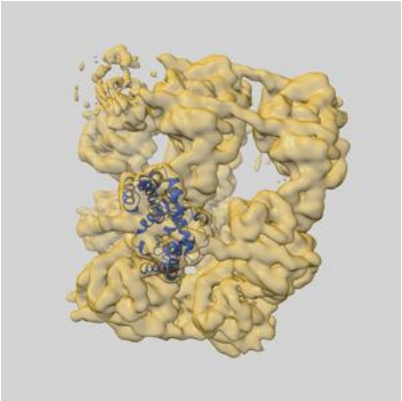
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-13355 and PDB model 7PEO. Per-residue inclusion information can be found in section 3 on page 5.

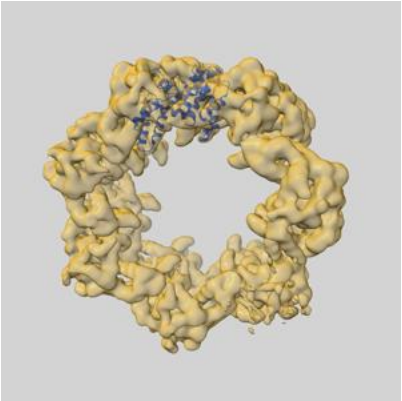
9.0.1 Map-model overlay (i)



X

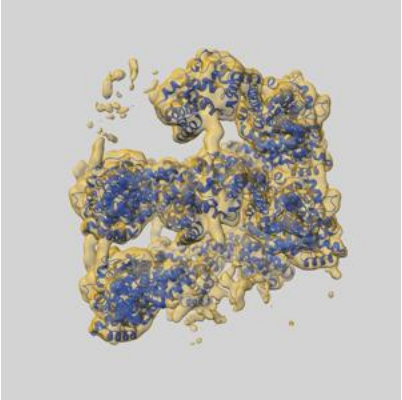


Y

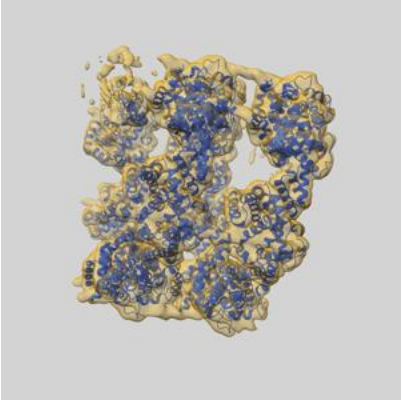


Z

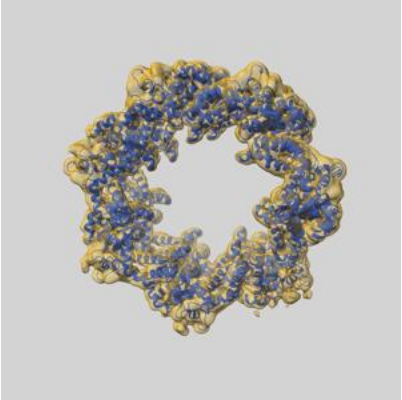
9.0.2 Map-model assembly overlay (i)



X



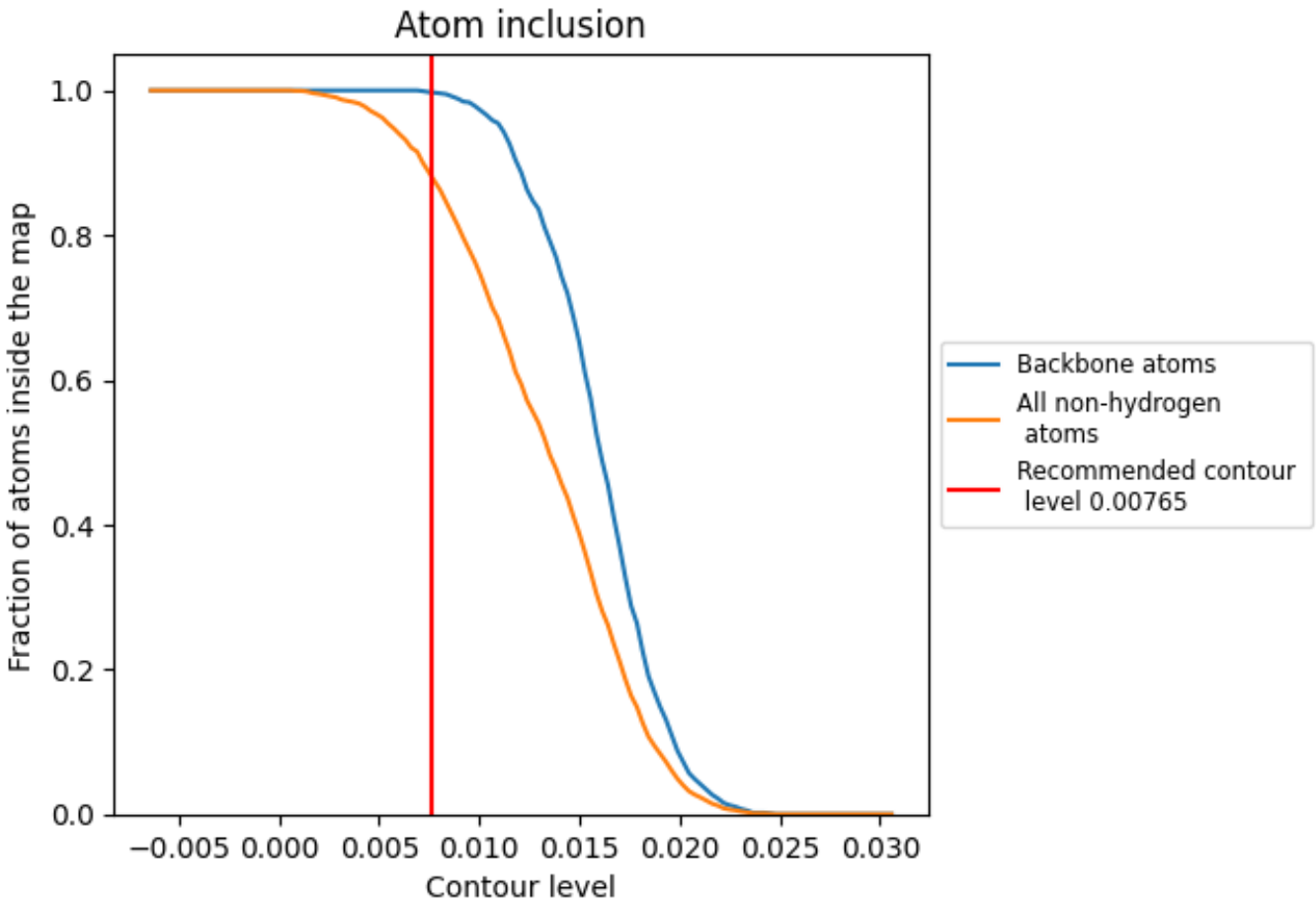
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.00765 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.

9.1 Atom inclusion (i)



At the recommended contour level, 100% of all backbone atoms, 88% of all non-hydrogen atoms, are inside the map.