wwPDB EM Validation Summary Report (i)

May 26, 2024 – 12:10 PM EDT

PDB ID	:	7SHM
EMDB ID	:	EMD-25130
Title	:	Cryo-EM structure of ATP-bound human peroxisomal fatty acid transporter ABCD1
Authors	:	Wang, R.; Li, X.
Deposited on	:	2021-10-09
Resolution	:	3.14 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 (1)) were used in the production of this report:

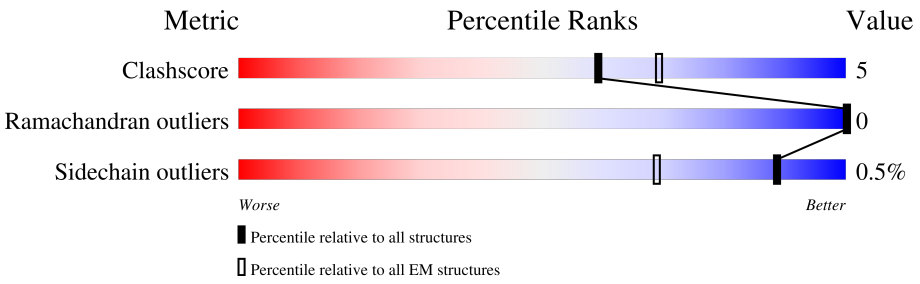
EMDB validation analysis	:	0.0.1.dev92
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.14 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	753	6% 66% 9% 24%
1	B	753	6% 66% 10% 24%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9186 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.

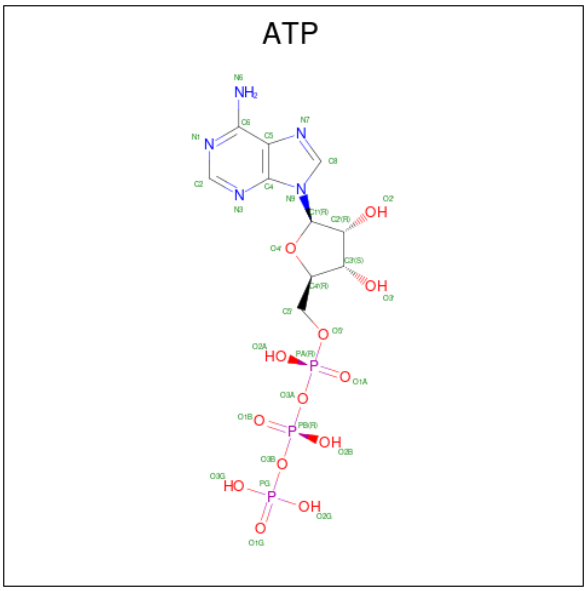
- Molecule 1 is a protein called ATP-binding cassette sub-family D member 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	570	4553	2927	805	799	22	0	0
1	B	572	4571	2938	808	803	22	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	expression tag	UNP P33897
A	-6	ASP	-	expression tag	UNP P33897
A	-5	TYR	-	expression tag	UNP P33897
A	-4	LYS	-	expression tag	UNP P33897
A	-3	ASP	-	expression tag	UNP P33897
A	-2	ASP	-	expression tag	UNP P33897
A	-1	ASP	-	expression tag	UNP P33897
A	0	ASP	-	expression tag	UNP P33897
A	1	LYS	-	expression tag	UNP P33897
A	630	GLN	GLU	engineered mutation	UNP P33897
B	-7	MET	-	expression tag	UNP P33897
B	-6	ASP	-	expression tag	UNP P33897
B	-5	TYR	-	expression tag	UNP P33897
B	-4	LYS	-	expression tag	UNP P33897
B	-3	ASP	-	expression tag	UNP P33897
B	-2	ASP	-	expression tag	UNP P33897
B	-1	ASP	-	expression tag	UNP P33897
B	0	ASP	-	expression tag	UNP P33897
B	1	LYS	-	expression tag	UNP P33897
B	630	GLN	GLU	engineered mutation	UNP P33897

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



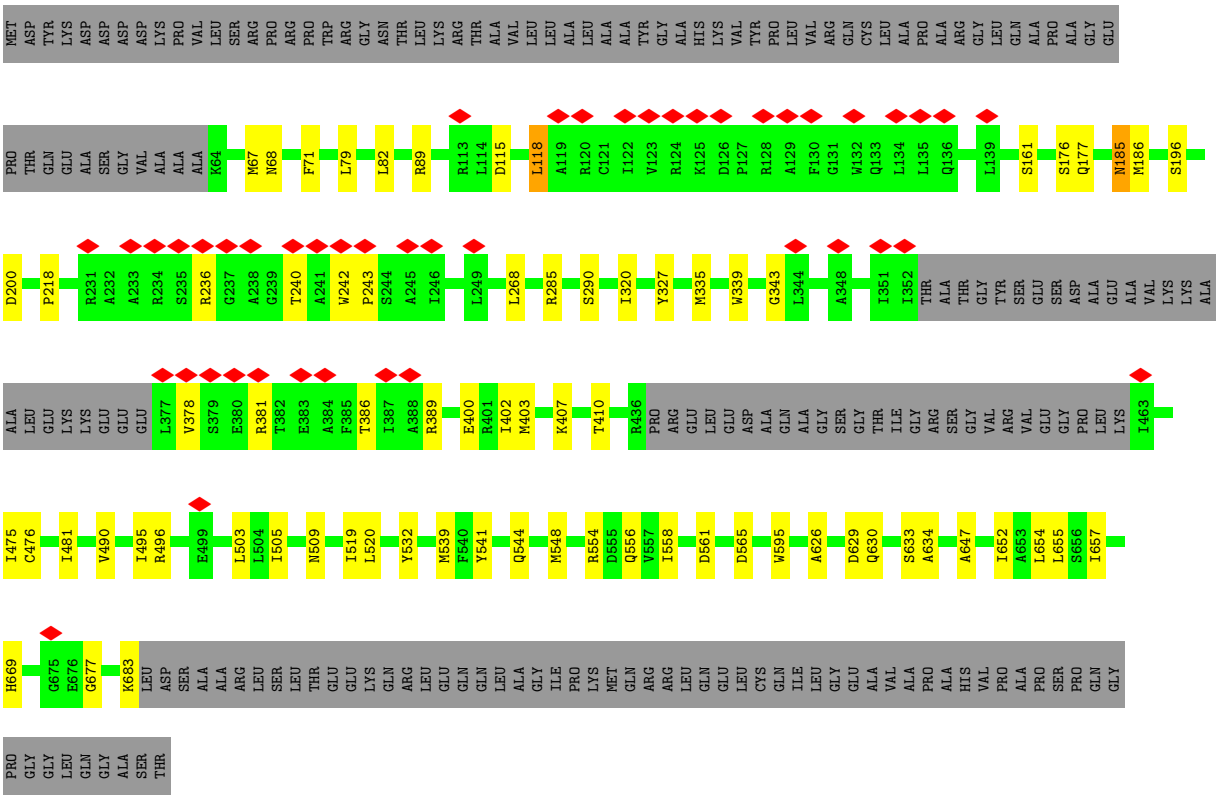
Mol	Chain	Residues	Atoms					AltConf
2	A	1	Total	C	N	O	P	0
			31	10	5	13	3	
2	B	1	Total	C	N	O	P	0
			31	10	5	13	3	

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.

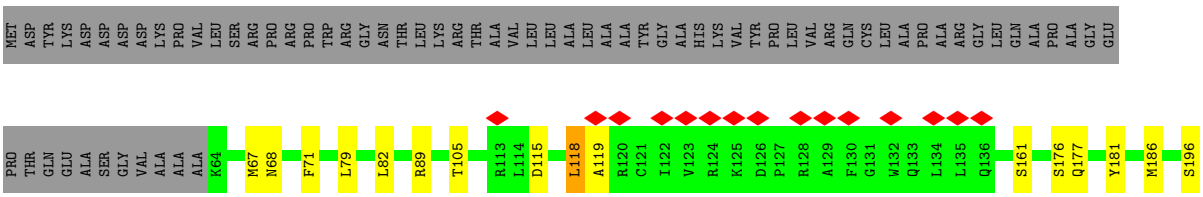
• Molecule 1: ATP-binding cassette sub-family D member 1

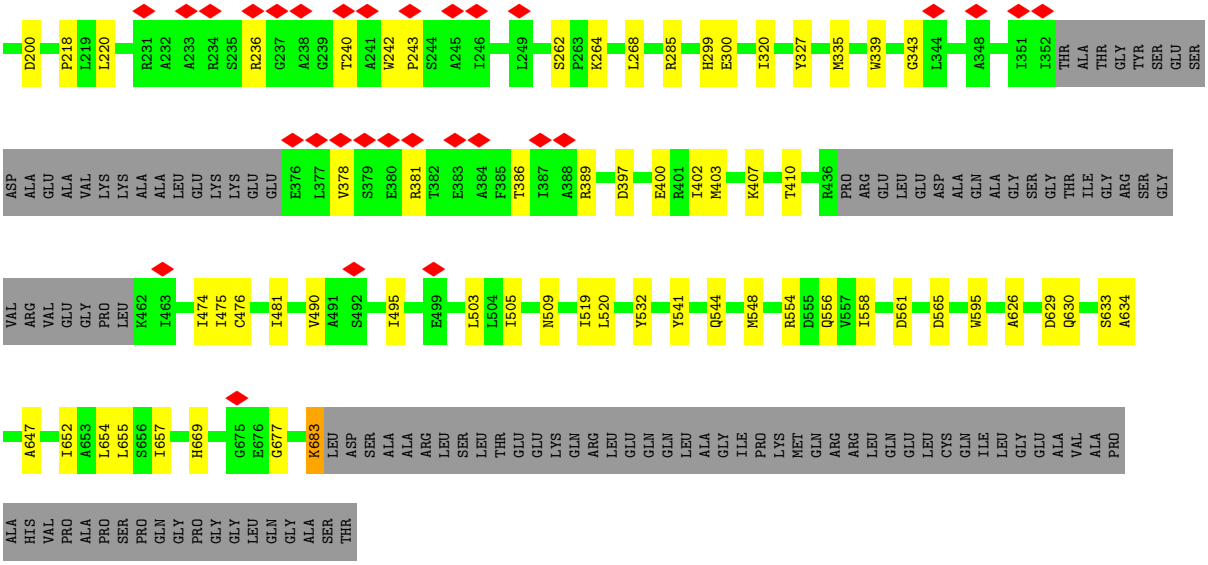
Chain A:



• Molecule 1: ATP-binding cassette sub-family D member 1

Chain B:




D200
P218
L219
L220
R231
A232
A233
R234
S235
R236
G237
A238
G239
T240
A241
W242
P243
S244
A245
I246
L249
S262
P263
K264
L268
R285
H299
E300
I320
Y327
M335
W339
G343
L344
A348
I351
I352
THR ALA THR GLY TYR SER GLU SER
ASP ALA GLU ALA VAL LYS LYS ALA ALA LEU GLU GLU LYS LYS GLU GLU
E376
L377
V378
S379
E380
R381
T382
E383
A384
F385
T386
I387
A388
R389
D397
E400
R401
I402
M403
K407
T410
R436
PRO ARG GLU LEU GLU ASP ALA GLN ALA GLY SER GLY THR ILE GLY ARG SER GLY
VAL ARG VAL GLU GLY PRO LEU
K462
I463
I474
I475
C476
I481
V490
A491
S492
I495
E499
L503
L504
I505
N509
I519
L520
Y532
Y541
Q544
M548
R554
D555
Q556
V557
I558
D561
D565
W595
A626
D629
Q630
S633
A634
A647
I652
A653
L654
L655
S656
I657
H669
G675
E676
G677
K683
LEU ASP SER ALA ALA ARG LEU SER LEU THR GLU GLU LYS GLN ARG LEU GLU GLN GLN LEU ALA GLY ILE PRO LYS MET GLN ARG ARG LEU GLN GLU LEU CYS GLN ILE LEU GLY GLU ALA VAL ALA PRO
ALA HIS VAL PRO ALA PRO SER PRO GLN GLY PRO GLY GLY LEU GLN GLY ALA SER THR

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	247022	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	4.215	Depositor
Minimum map value	-2.436	Depositor
Average map value	0.006	Depositor
Map value standard deviation	0.092	Depositor
Recommended contour level	0.32	Depositor
Map size (Å)	229.568, 229.568, 229.568	wwPDB
Map dimensions	272, 272, 272	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.844, 0.844, 0.844	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: ATP

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	0.33	0/4654	0.56	3/6302 (0.0%)
1	B	0.33	0/4672	0.56	2/6325 (0.0%)
All	All	0.33	0/9326	0.56	5/12627 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	185	ASN	CB-CA-C	-5.64	99.12	110.40
1	B	655	LEU	CA-CB-CG	5.25	127.38	115.30
1	A	655	LEU	CA-CB-CG	5.25	127.37	115.30
1	A	558	ILE	CG1-CB-CG2	-5.22	99.91	111.40
1	B	558	ILE	CG1-CB-CG2	-5.22	99.92	111.40

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	4553	0	4612	44	0
1	B	4571	0	4631	50	0
2	A	31	0	12	0	0
2	B	31	0	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9186	0	9267	88	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 5.

The worst 5 of 88 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:B:683:LYS:O	1:B:683:LYS:HD2	1.78	0.84
1:A:544:GLN:NE2	1:A:630:GLN:HE21	1.77	0.81
1:B:544:GLN:NE2	1:B:630:GLN:HE21	1.77	0.81
1:B:240:THR:O	1:B:243:PRO:HD2	1.90	0.71
1:B:397:ASP:O	1:B:400:GLU:HB3	1.90	0.71

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	564/753 (75%)	547 (97%)	17 (3%)	0	100	100
1	B	566/753 (75%)	549 (97%)	17 (3%)	0	100	100
All	All	1130/1506 (75%)	1096 (97%)	34 (3%)	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	480/618 (78%)	478 (100%)	2 (0%)	91 96
1	B	482/618 (78%)	479 (99%)	3 (1%)	86 94
All	All	962/1236 (78%)	957 (100%)	5 (0%)	89 95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	118	LEU
1	A	236	ARG
1	B	118	LEU
1	B	236	ARG
1	B	683	LYS

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	B	318	ASN
1	B	544	GLN
1	B	659	HIS
1	B	466	GLN
1	A	659	HIS

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)

2 ligands 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	ATP	A	801	-	26,33,33	1.10	2 (7%)	31,52,52	1.92	5 (16%)
2	ATP	B	801	-	26,33,33	1.08	2 (7%)	31,52,52	1.90	5 (16%)

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	ATP	A	801	-	-	5/18/38/38	0/3/3/3
2	ATP	B	801	-	-	7/18/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	801	ATP	O4'-C1'	2.60	1.44	1.41
2	B	801	ATP	O4'-C1'	2.48	1.44	1.41
2	B	801	ATP	C8-N7	-2.48	1.30	1.34
2	A	801	ATP	C8-N7	-2.47	1.30	1.34

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	801	ATP	PA-O3A-PB	-6.17	111.65	132.83
2	B	801	ATP	PA-O3A-PB	-5.99	112.25	132.83
2	A	801	ATP	PB-O3B-PG	-5.57	113.71	132.83
2	B	801	ATP	PB-O3B-PG	-5.50	113.94	132.83
2	B	801	ATP	N3-C2-N1	-4.26	122.02	128.68

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

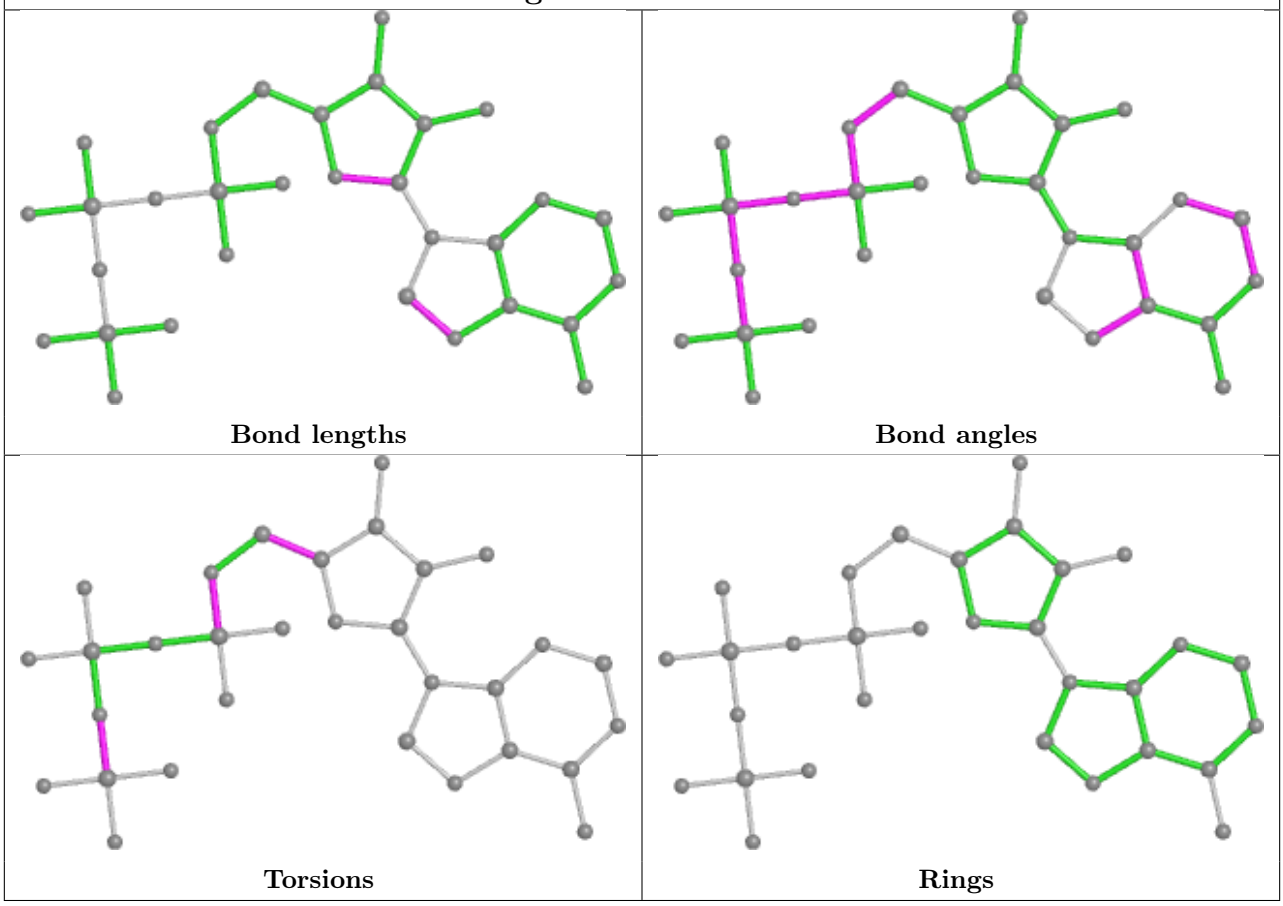
Mol	Chain	Res	Type	Atoms
2	A	801	ATP	PB-O3B-PG-O3G
2	A	801	ATP	C5'-O5'-PA-O1A
2	A	801	ATP	C5'-O5'-PA-O3A
2	B	801	ATP	C5'-O5'-PA-O1A
2	B	801	ATP	C5'-O5'-PA-O3A

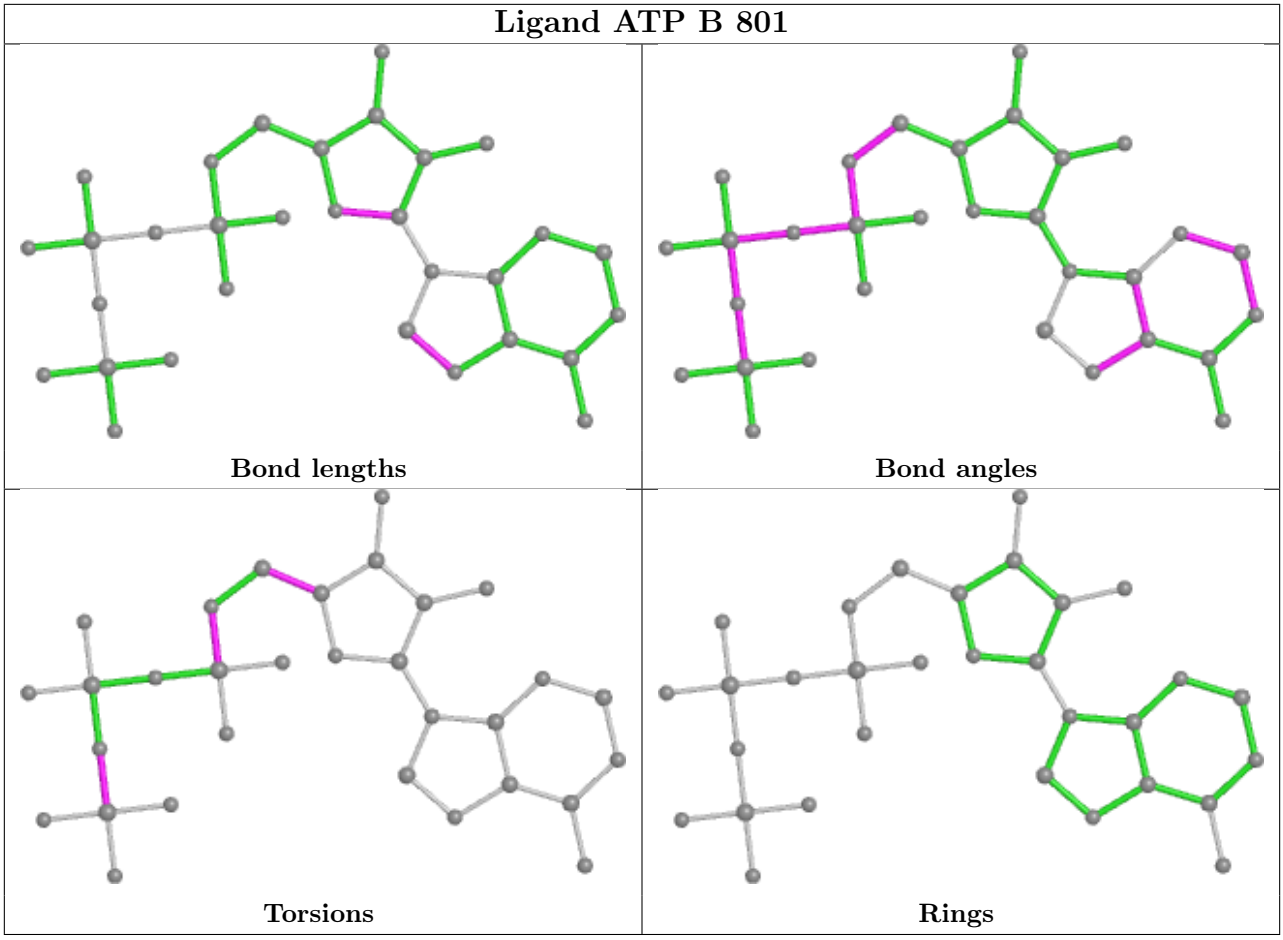
There are no ring outliers.

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 all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand ATP A 801





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.

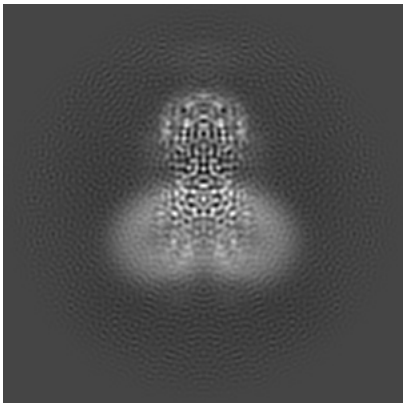
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-25130. These allow visual inspection of the internal detail of the map and identification of artifacts.

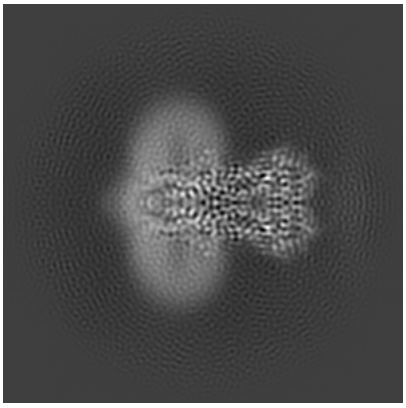
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

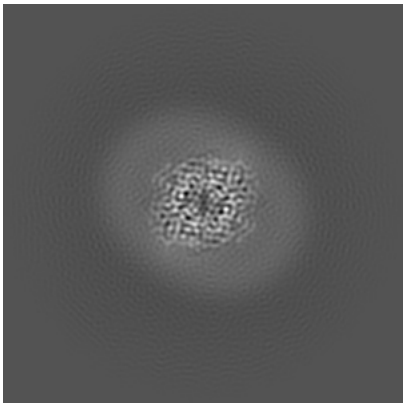
6.1.1 Primary map



X



Y

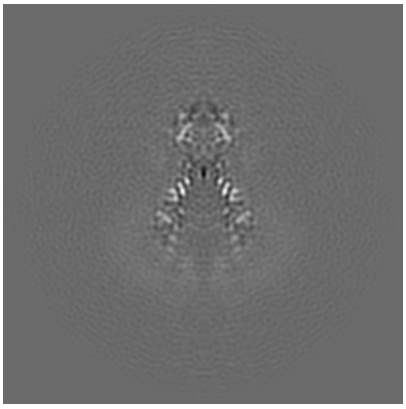


Z

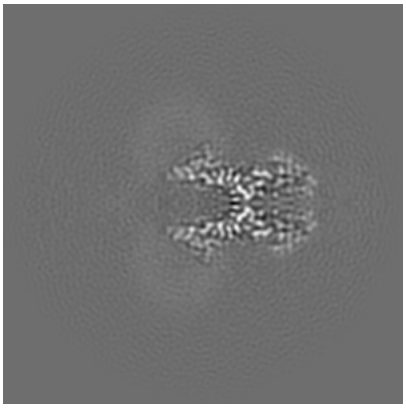
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

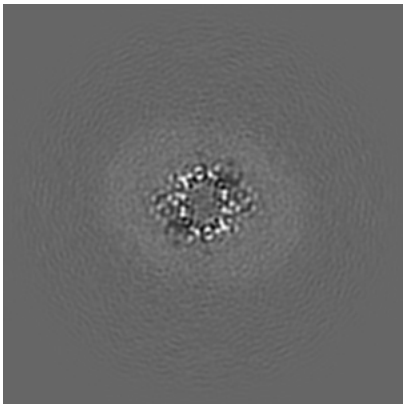
6.2.1 Primary map



X Index: 136



Y Index: 136

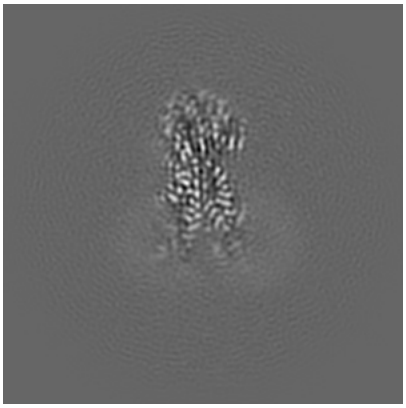


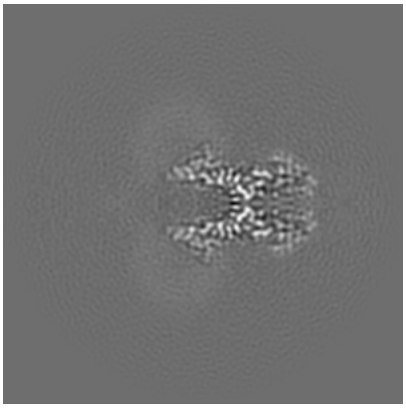
Z Index: 136

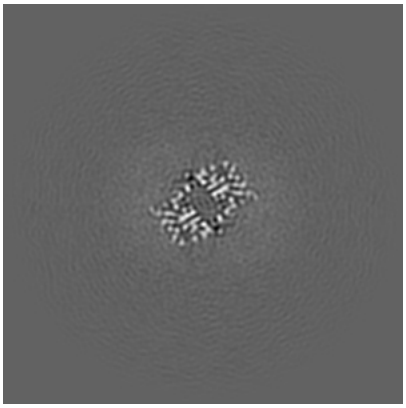
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map


X Index: 123

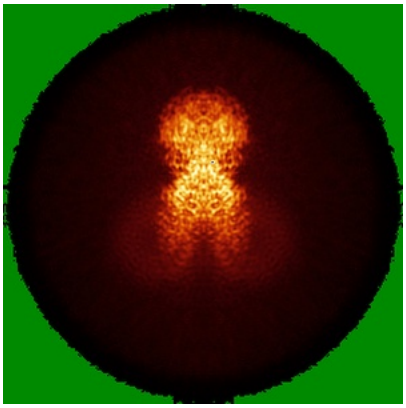

Y Index: 136

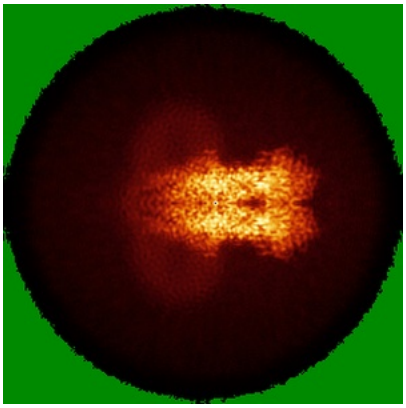

Z Index: 144

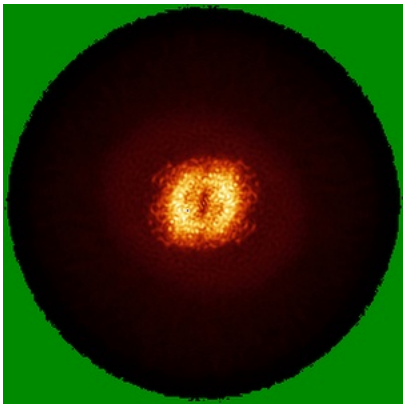
The images above show the largest variance slices of the map in three orthogonal directions.

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

6.4.1 Primary map


X

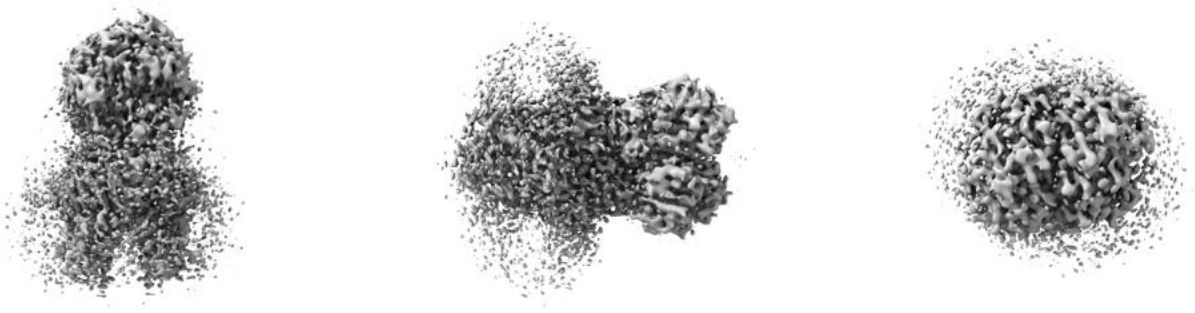

Y


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

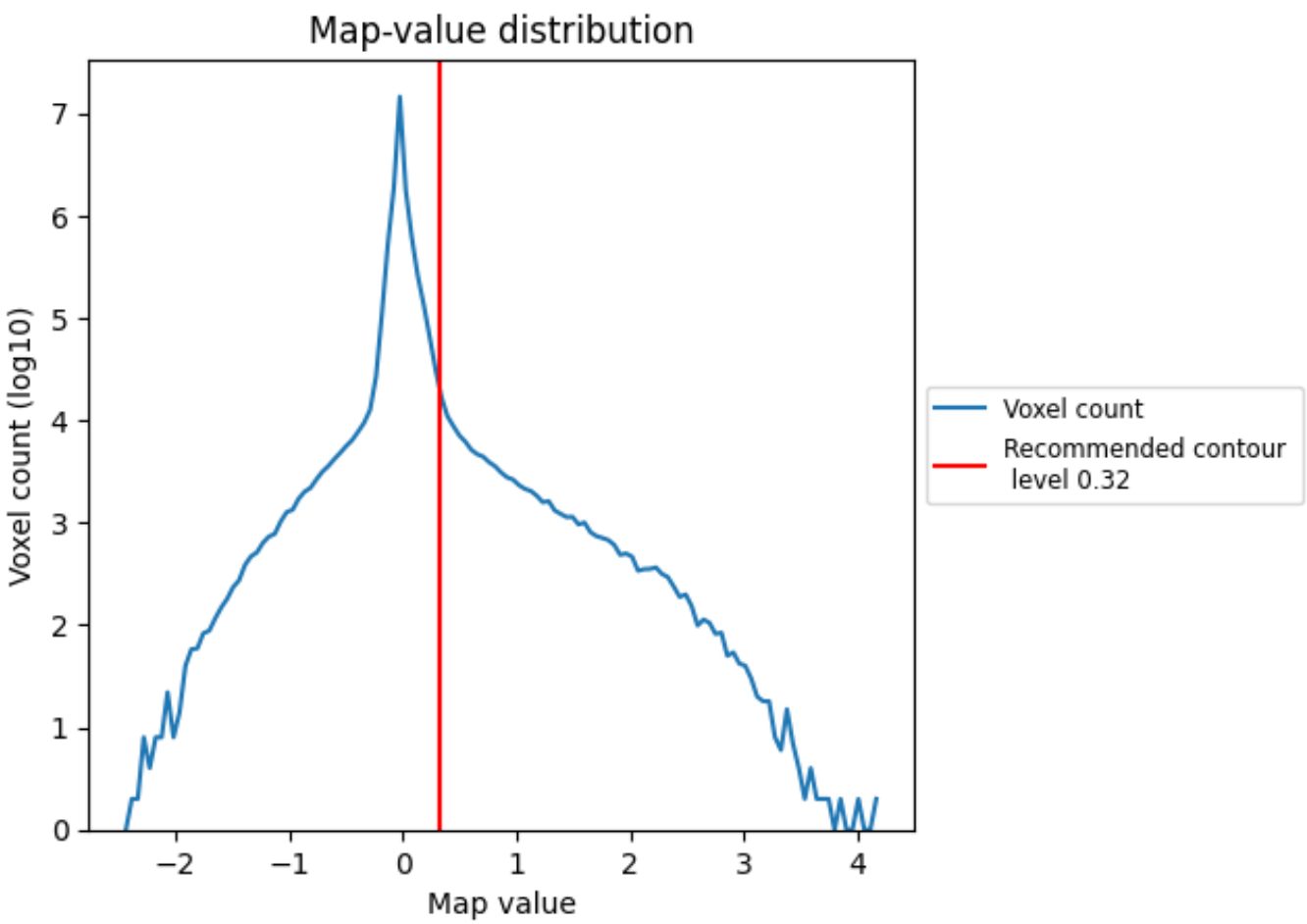
6.6 Mask visualisation (i)

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

7 Map analysis (i)

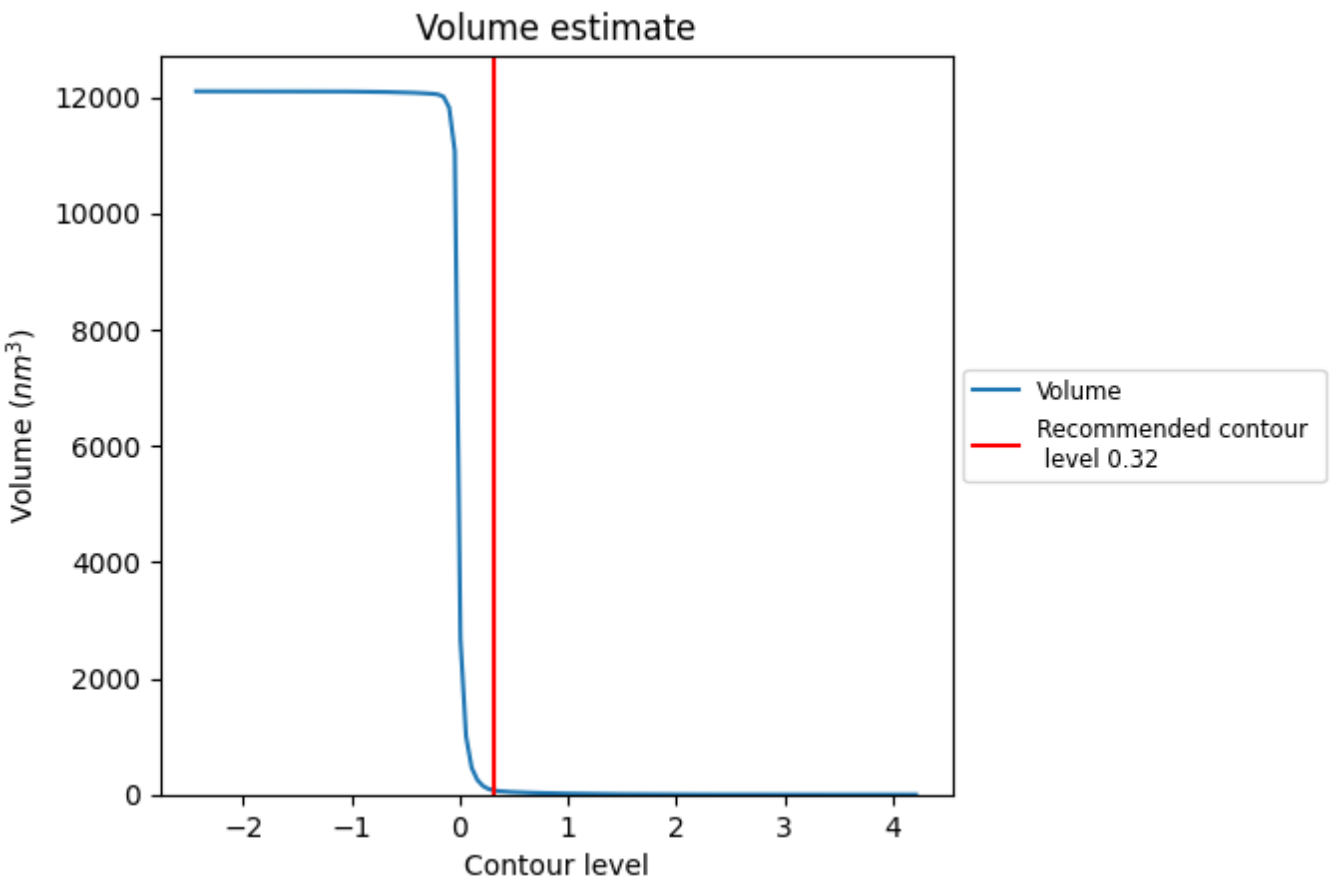
This section contains the results of statistical analysis of the map.

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.

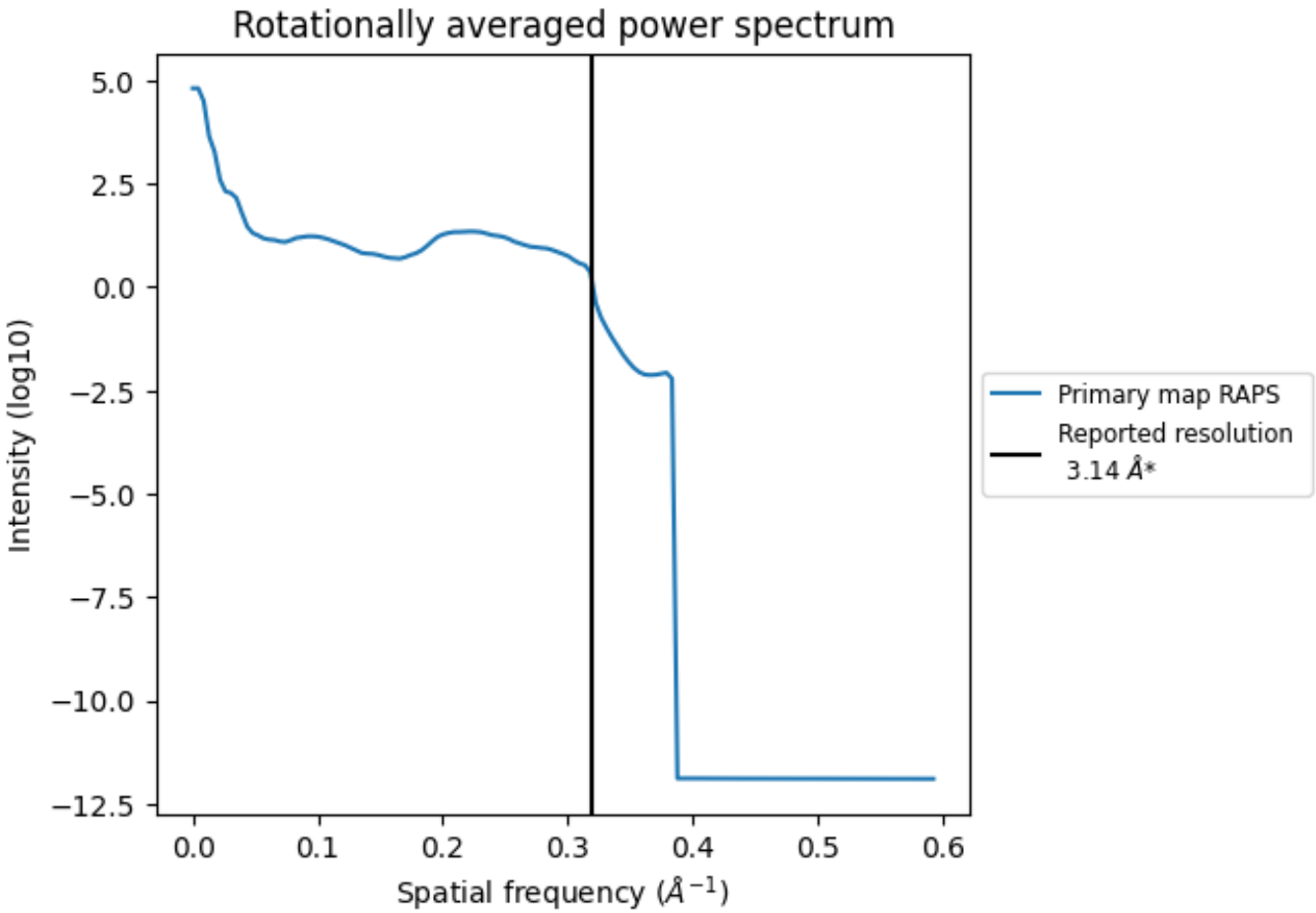
7.2 Volume estimate (i)



The volume at the recommended contour level is 70 nm^3; this corresponds to an approximate mass of 63 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 0.318 Å^{-1}

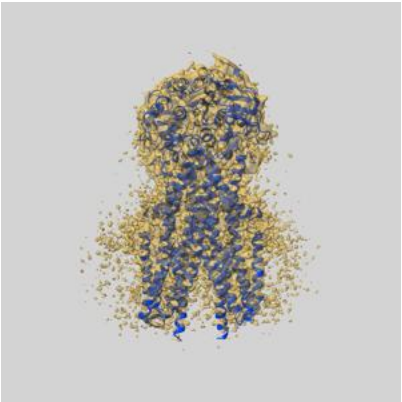
8 Fourier-Shell correlation (i)

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

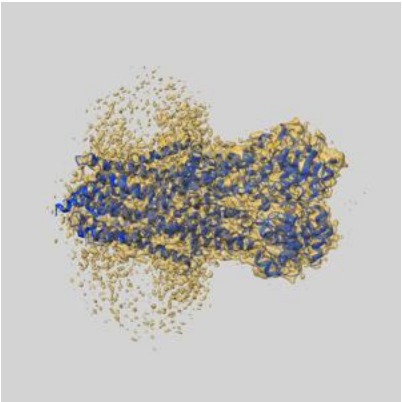
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-25130 and PDB model 7SHM. Per-residue inclusion information can be found in section 3 on page 5.

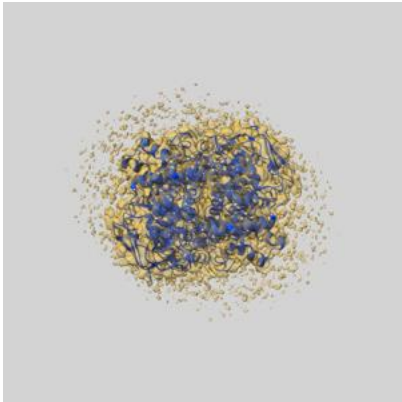
9.1 Map-model overlay (i)



X



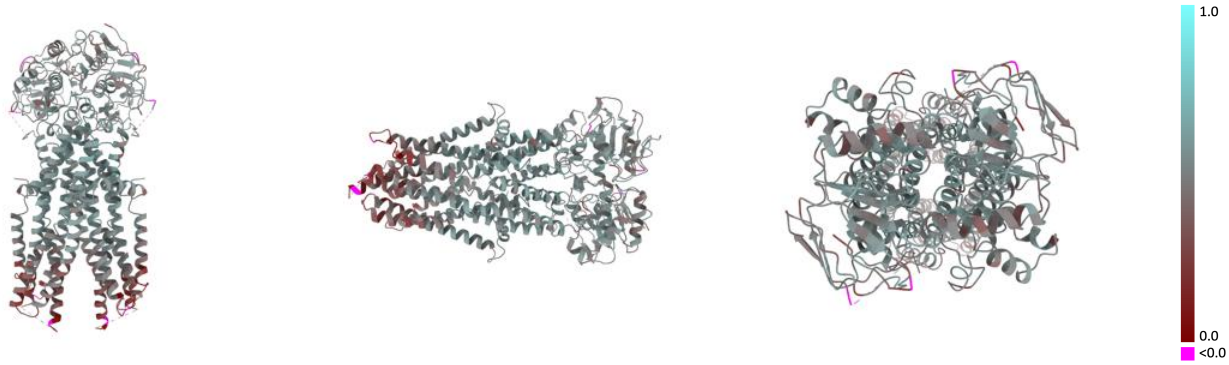
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.32 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.

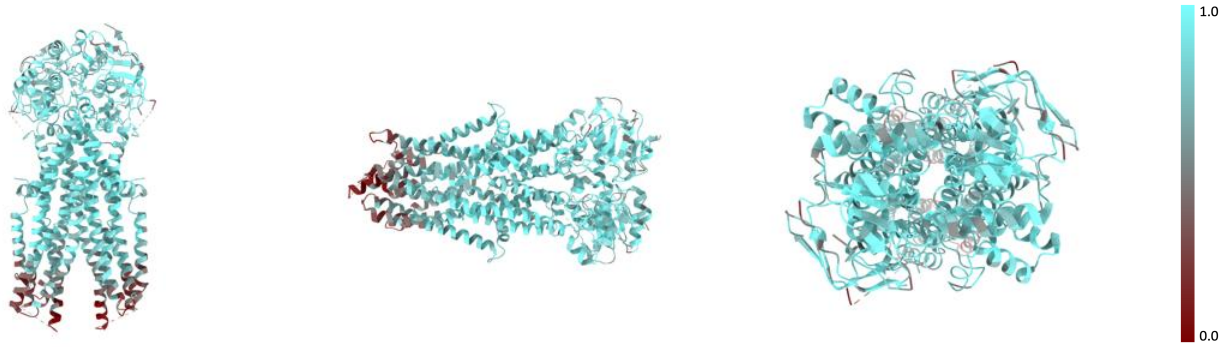
9.2 Q-score mapped to coordinate model (i)



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.

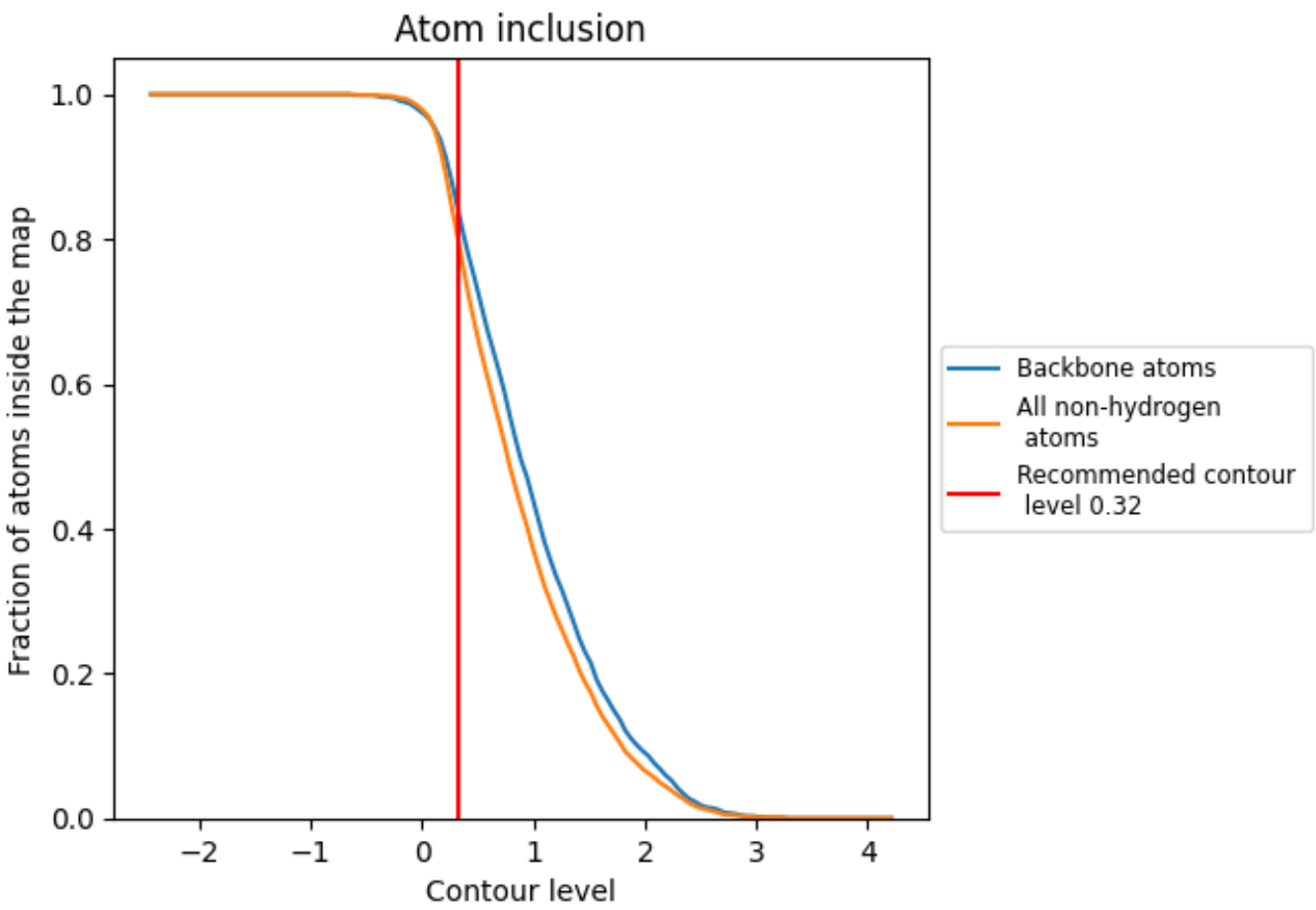
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.32).

9.4 Atom inclusion (i)



At the recommended contour level, 84% of all backbone atoms, 80% 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 (0.32) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8030	0.4880
A	0.8040	0.4890
B	0.8010	0.4870