wwPDB EM Validation Summary Report (i)

Nov 19, 2022 – 01:24 PM EST

PDB ID	:	7R8E
EMDB ID	:	EMD-24317
Title	:	The structure of human ABCG1 E242Q complexed with ATP
Authors	:	Sun, Y.; Li, X.; Long, T.
Deposited on	:	2021-06-26
Resolution	:	3.68 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:

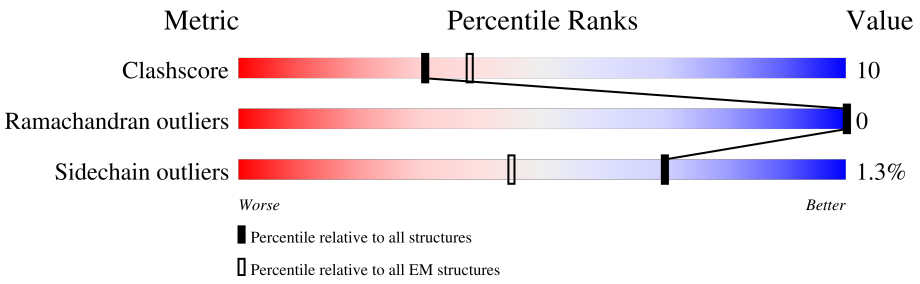
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.68 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	666	10% 68% 14% 18%
1	B	666	13% 69% 13% 18%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 8700 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.

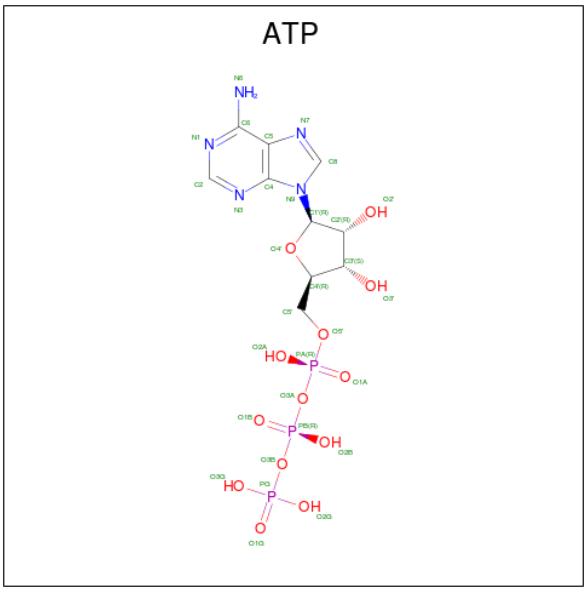
- Molecule 1 is a protein called Isoform 4 of ATP-binding cassette sub-family G member 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	548	4290	2781	711	764	34	0	0
1	B	548	4290	2781	711	764	34	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	242	GLN	GLU	engineered mutation	UNP P45844
B	242	GLN	GLU	engineered mutation	UNP P45844

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).



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

Continued on next page...

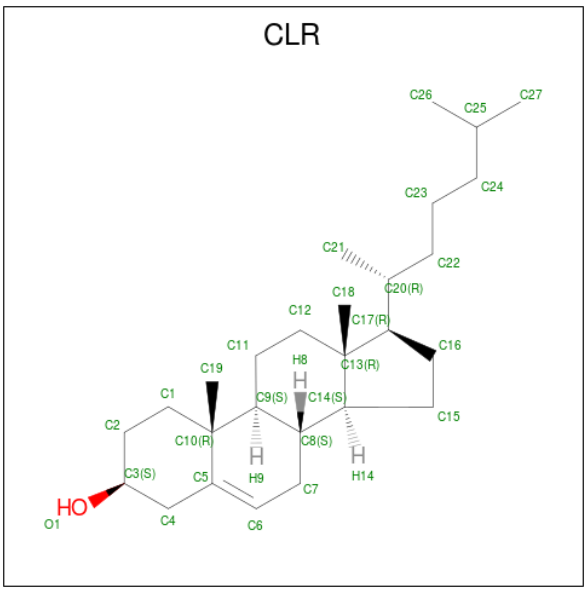
Continued from previous page...

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

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
3	A	1	1	1	0
3	B	1	1	1	0

- Molecule 4 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$) (labeled as "Ligand of Interest" by depositor).



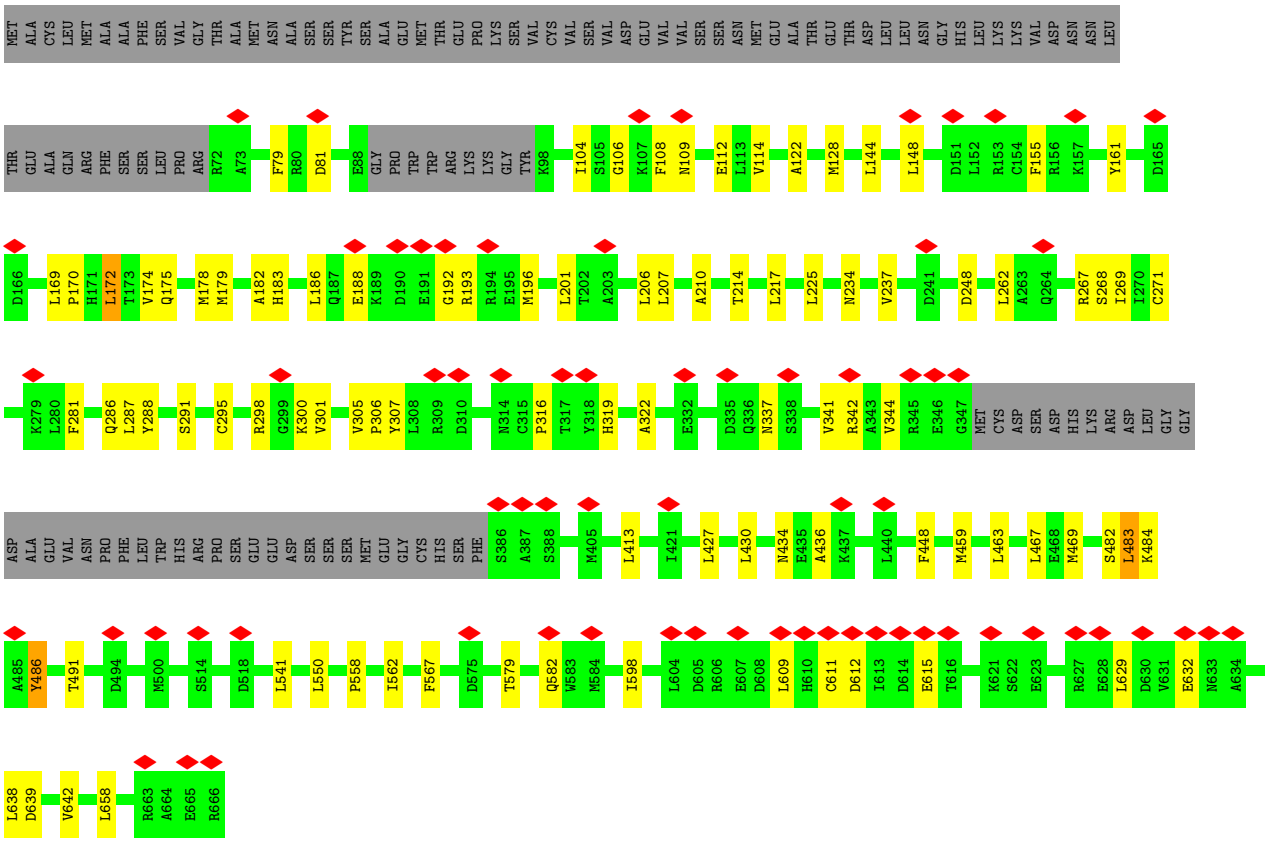
Mol	Chain	Residues	Atoms			AltConf
			Total	C	O	
4	A	1	28	27	1	0
4	B	1	28	27	1	0

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.

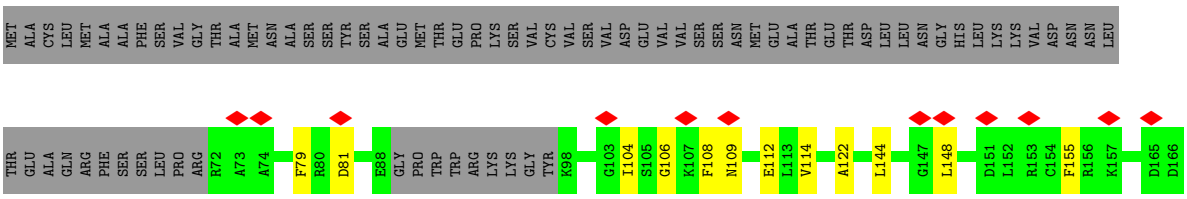
• Molecule 1: Isoform 4 of ATP-binding cassette sub-family G member 1

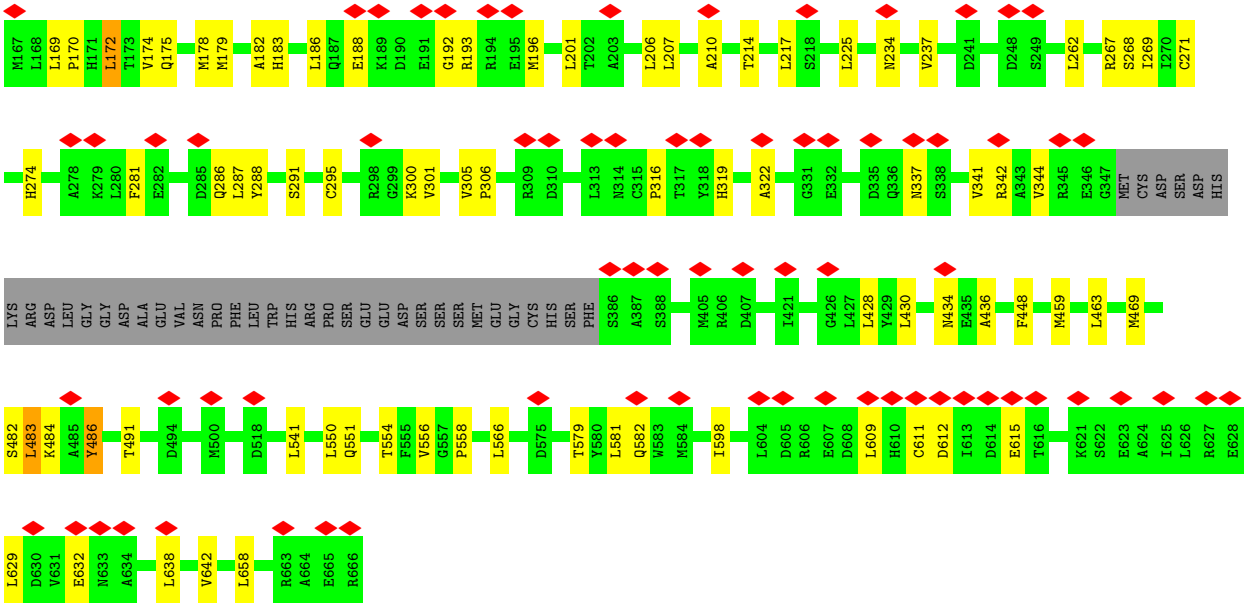
Chain A:



• Molecule 1: Isoform 4 of ATP-binding cassette sub-family G member 1

Chain B:




M167 L168 L169 P170 H171 L172 T173 V174 Q175 M178 M179 A182 H183 L186 Q187 E188 K189 D190 E191 G192 R193 R194 E195 M196 L201 T202 A203 L206 L207 A210 T214 L217 S218 L225 N234 V237 D241 D248 S249 L262 R267 S268 I269 I270 C271
H274 A278 K279 L280 F281 E282 D285 Q286 L287 Y288 S291 C295 R298 G299 K300 V301 V305 P306 R309 D310 L313 N314 C315 P316 T317 Y318 H319 A322 G331 E332 D335 Q336 N337 S338 V341 R342 A343 V344 R345 E346 G347 MET CYS ASP SER ASP HIS
LYS ARG ASP LEU GLY GLY ASP ALA GLU VAL ASN PRO PHE LEU TRP HIS ARG PRO SER GLU GLU ASP SER SER SER MET GLU GLY CYS HIS SER PHE S386 A387 S388 M405 R406 D407 I421 G426 L427 L428 Y429 L430 N434 E435 A436 F448 M459 L463 M469
S482 L483 K484 A485 Y486 T491 D494 M500 D518 L541 L550 Q551 T554 F555 V556 G557 P558 L566 D575 T579 Y580 L581 Q582 W583 M584 I598 L604 D605 R606 E607 D608 L609 H610 C611 D612 I613 D614 E615 T616 K621 S622 E623 A624 I625 L626 R627 E628
L629 D630 V631 E632 N633 A634 L638 V642 L658 R663 A664 E665 R666

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	72323	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 ($e^-/Å^2$)	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	0.088	Depositor
Minimum map value	-0.053	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	213.248, 213.248, 213.248	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.833, 0.833, 0.833	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: CLR, MG, 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	0.71	0/4378	0.79	0/5918
1	B	0.71	0/4378	0.79	0/5918
All	All	0.71	0/8756	0.79	0/11836

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	4290	0	4377	87	0
1	B	4290	0	4377	83	0
2	A	31	0	12	0	0
2	B	31	0	12	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	28	0	46	11	0
4	B	28	0	46	13	0
All	All	8700	0	8870	172	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 10.

The worst 5 of 172 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:262:LEU:CD2	1:A:267:ARG:HD2	1.61	1.31
1:B:262:LEU:CD2	1:B:267:ARG:HD2	1.61	1.28
4:B:703:CLR:H221	4:B:703:CLR:H263	1.30	1.12
4:A:703:CLR:H263	4:A:703:CLR:H221	1.30	1.06
1:B:234:ASN:HB3	1:B:267:ARG:NH1	1.71	1.06

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	542/666 (81%)	490 (90%)	52 (10%)	0	100	100
1	B	542/666 (81%)	489 (90%)	53 (10%)	0	100	100
All	All	1084/1332 (81%)	979 (90%)	105 (10%)	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	469/571 (82%)	463 (99%)	6 (1%)	69	83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	469/571 (82%)	463 (99%)	6 (1%)	69 83
All	All	938/1142 (82%)	926 (99%)	12 (1%)	70 83

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	342	ARG
1	B	430	LEU
1	B	632	GLU
1	B	483	LEU
1	A	483	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	478	ASN
1	B	582	GLN
1	B	551	GLN
1	A	582	GLN
1	B	275	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)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

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	701	3	26,33,33	0.94	1 (3%)	31,52,52	1.61	5 (16%)
4	CLR	A	703	-	31,31,31	0.80	1 (3%)	48,48,48	1.54	6 (12%)
2	ATP	B	701	3	26,33,33	0.94	1 (3%)	31,52,52	1.61	5 (16%)
4	CLR	B	703	-	31,31,31	0.80	1 (3%)	48,48,48	1.54	6 (12%)

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	701	3	-	0/18/38/38	0/3/3/3
4	CLR	A	703	-	-	5/10/68/68	0/4/4/4
2	ATP	B	701	3	-	0/18/38/38	0/3/3/3
4	CLR	B	703	-	-	5/10/68/68	0/4/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	701	ATP	C5-C4	2.48	1.47	1.40
2	A	701	ATP	C5-C4	2.48	1.47	1.40
4	A	703	CLR	C13-C14	-2.03	1.51	1.55
4	B	703	CLR	C13-C14	-2.02	1.51	1.55

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	703	CLR	C13-C17-C20	-4.94	111.75	119.49
4	B	703	CLR	C13-C17-C20	-4.91	111.79	119.49
4	B	703	CLR	C13-C14-C8	-4.08	108.33	114.38
4	A	703	CLR	C13-C14-C8	-4.07	108.35	114.38
4	B	703	CLR	C17-C13-C14	3.73	104.50	100.07

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	703	CLR	C13-C17-C20-C22
4	B	703	CLR	C13-C17-C20-C22
4	A	703	CLR	C13-C17-C20-C21
4	B	703	CLR	C13-C17-C20-C21
4	A	703	CLR	C20-C22-C23-C24

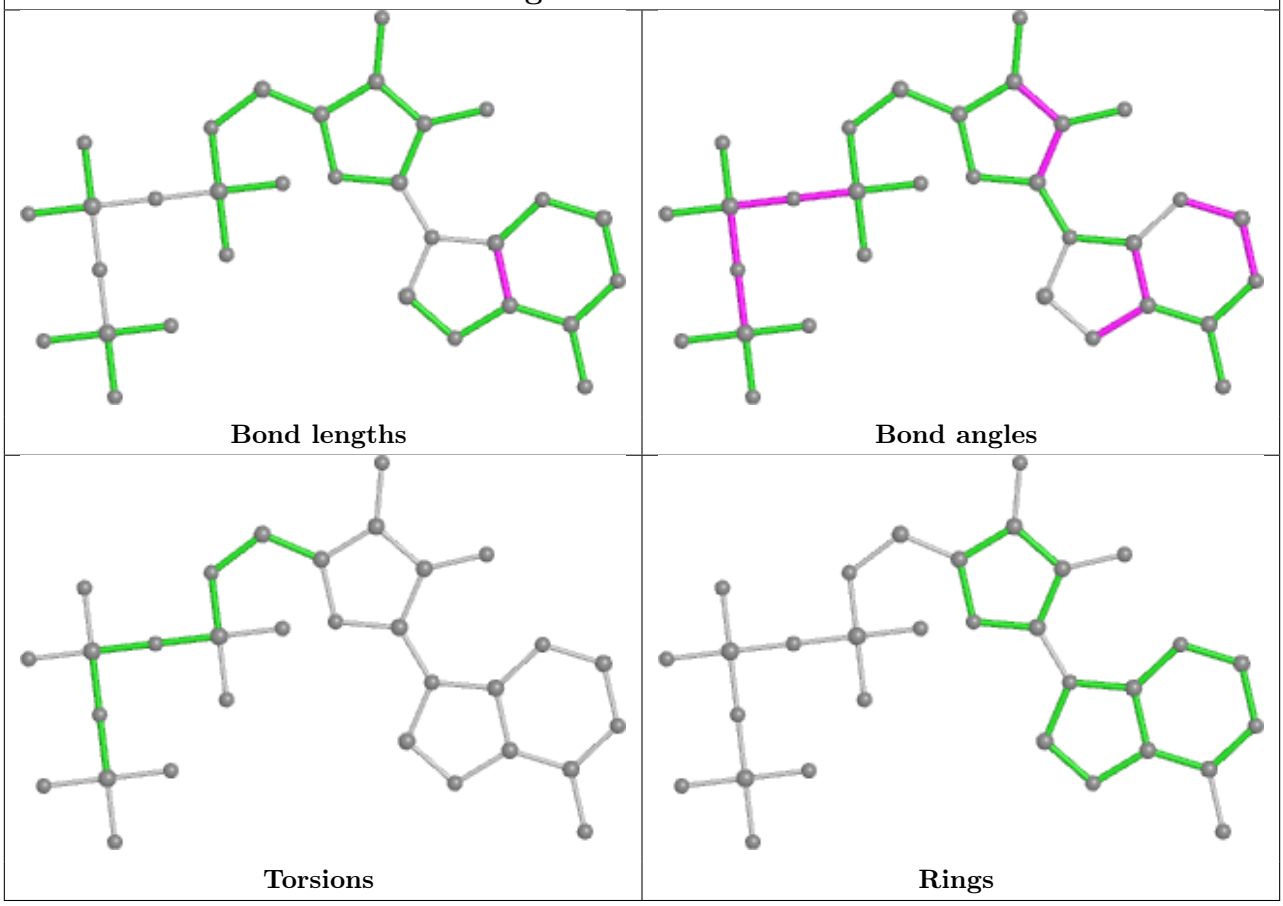
There are no ring outliers.

2 monomers are involved in 24 short contacts:

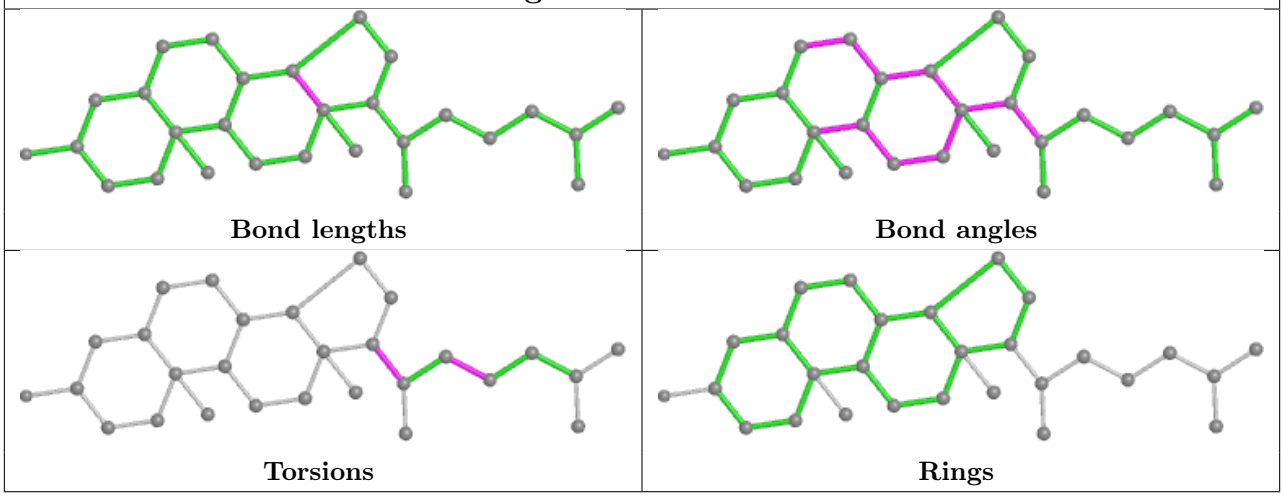
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	703	CLR	11	0
4	B	703	CLR	13	0

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.

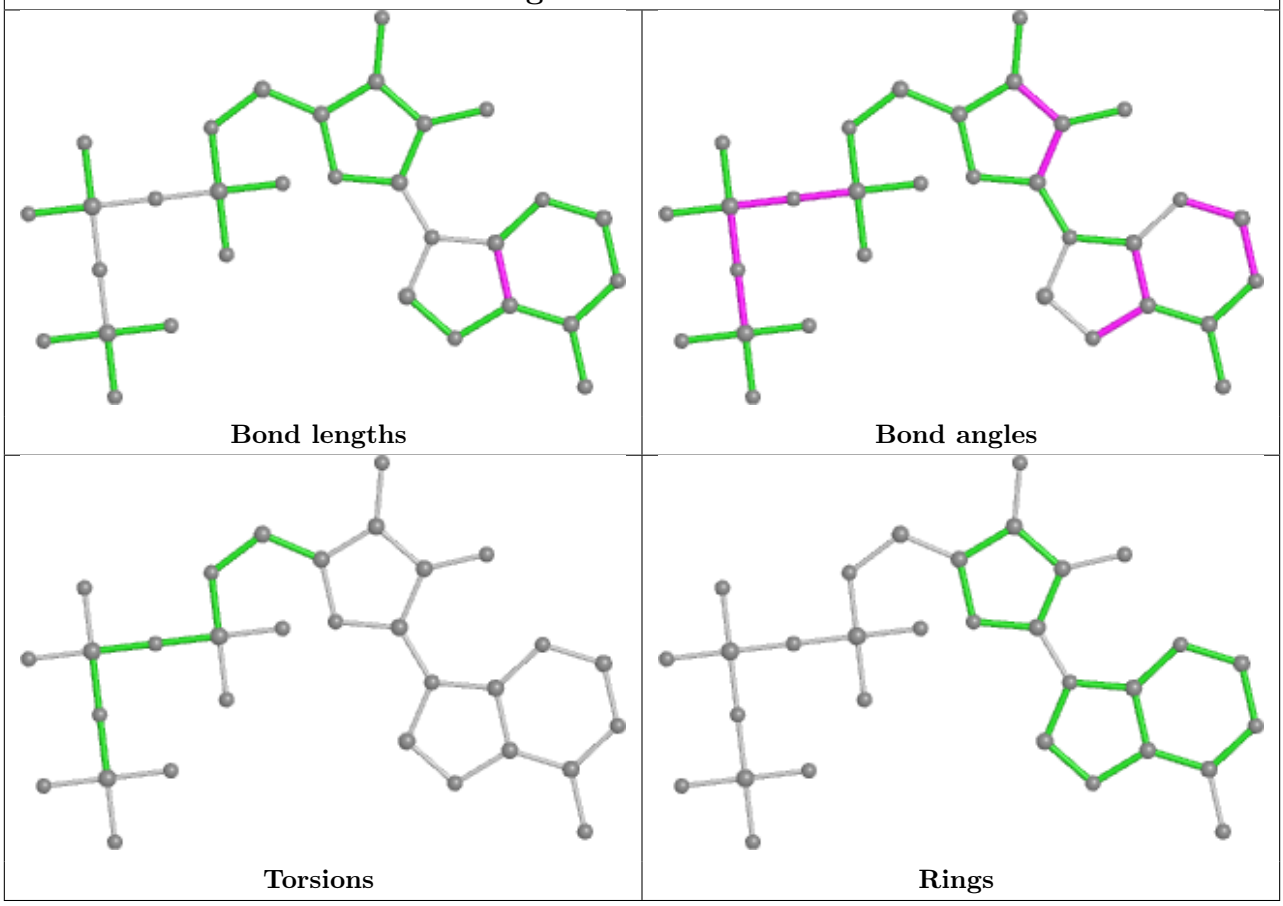
Ligand ATP A 701



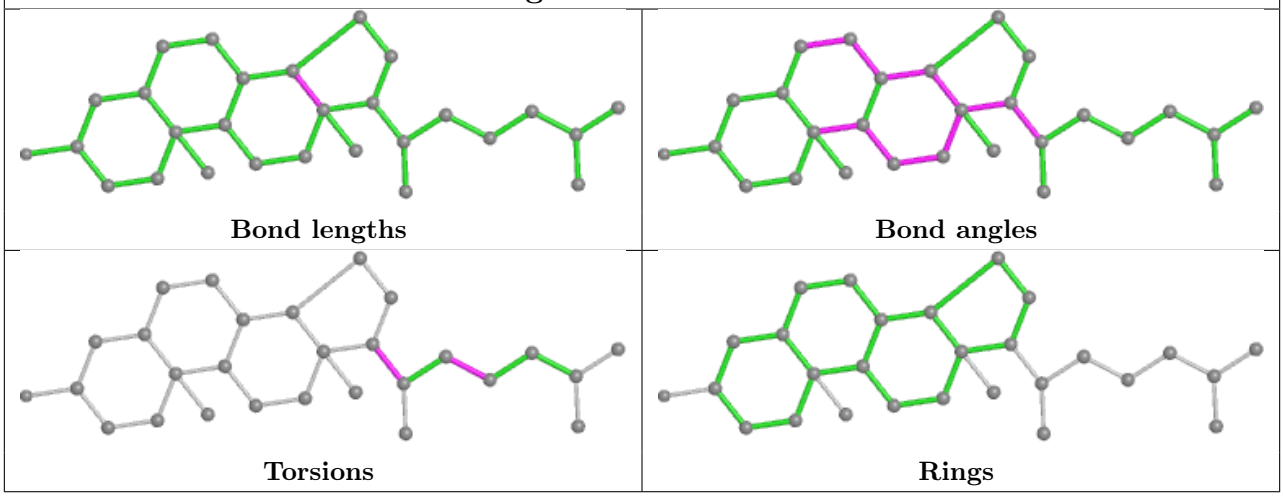
Ligand CLR A 703



Ligand ATP B 701



Ligand CLR B 703



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.

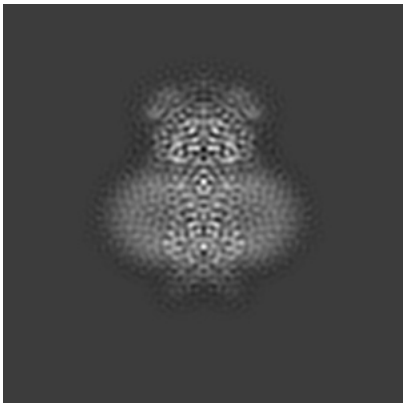
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-24317. These allow visual inspection of the internal detail of the map and identification of artifacts.

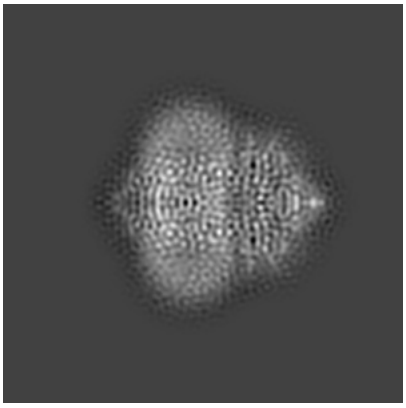
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

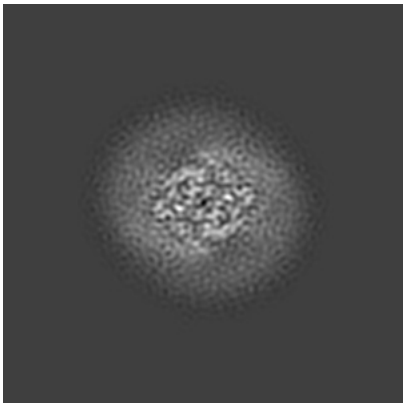
6.1.1 Primary map



X



Y



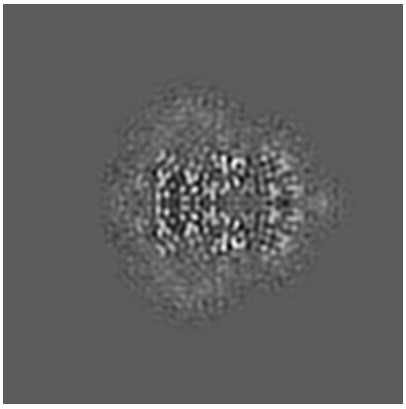
Z

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

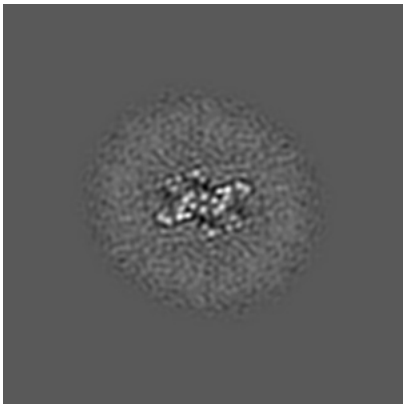
6.2 Central slices (i)

6.2.1 Primary map

X Index: 128



Y Index: 128

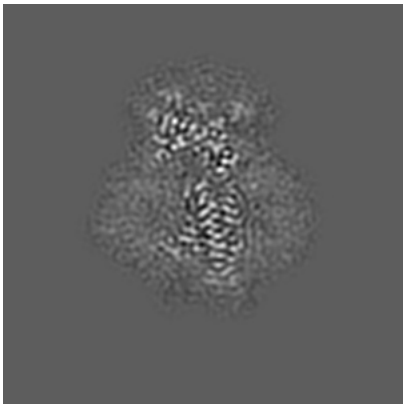


Z Index: 128

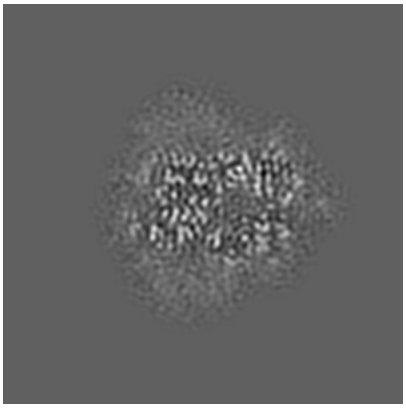
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map



X Index: 120



Y Index: 132

Z Index: 163

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

6.4 Orthogonal surface views (i)

6.4.1 Primary map

X

Y

Z

The images above show the 3D surface view of the map at the recommended contour level 0.02. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

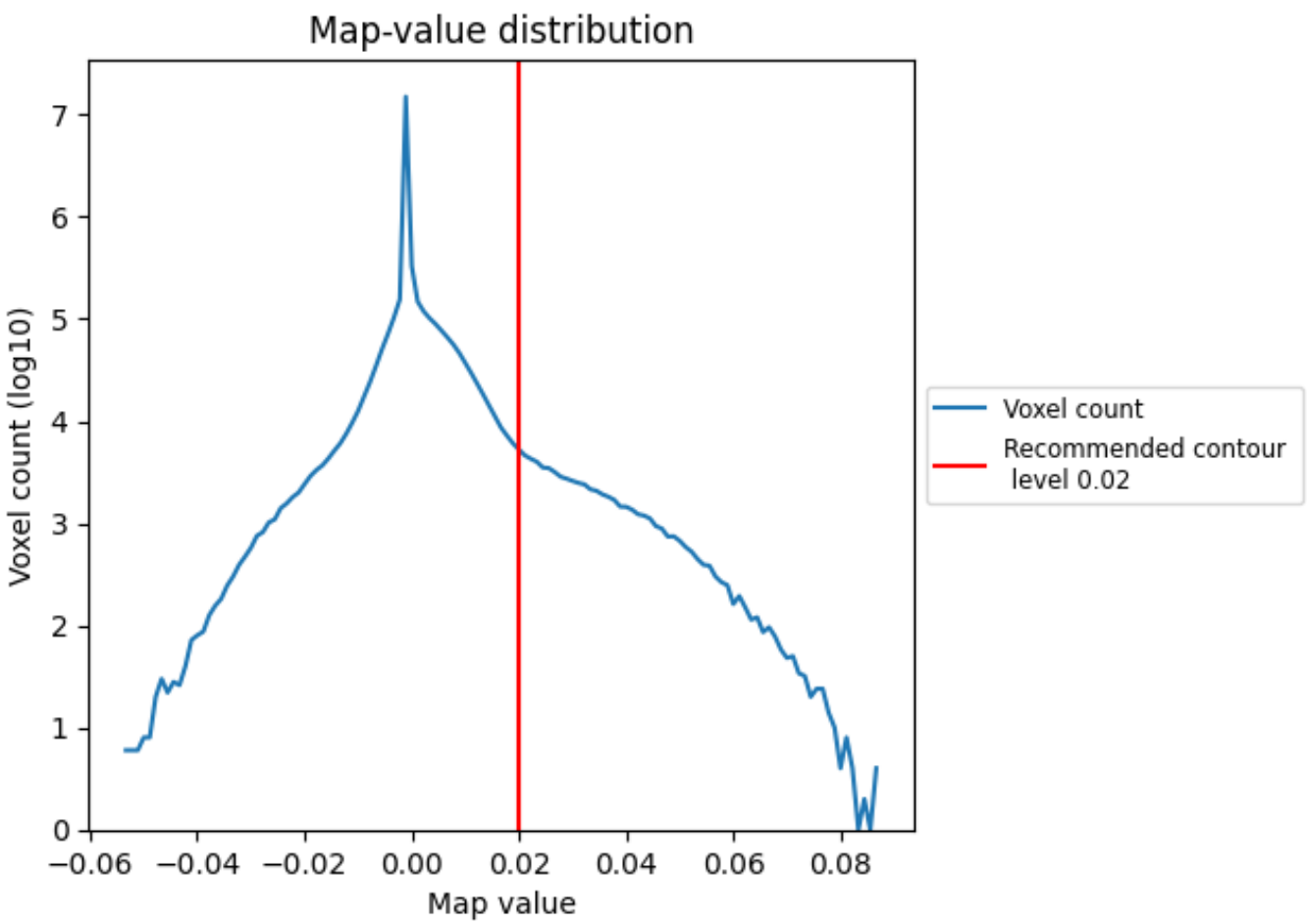
6.5 Mask visualisation (i)

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

7 Map analysis (i)

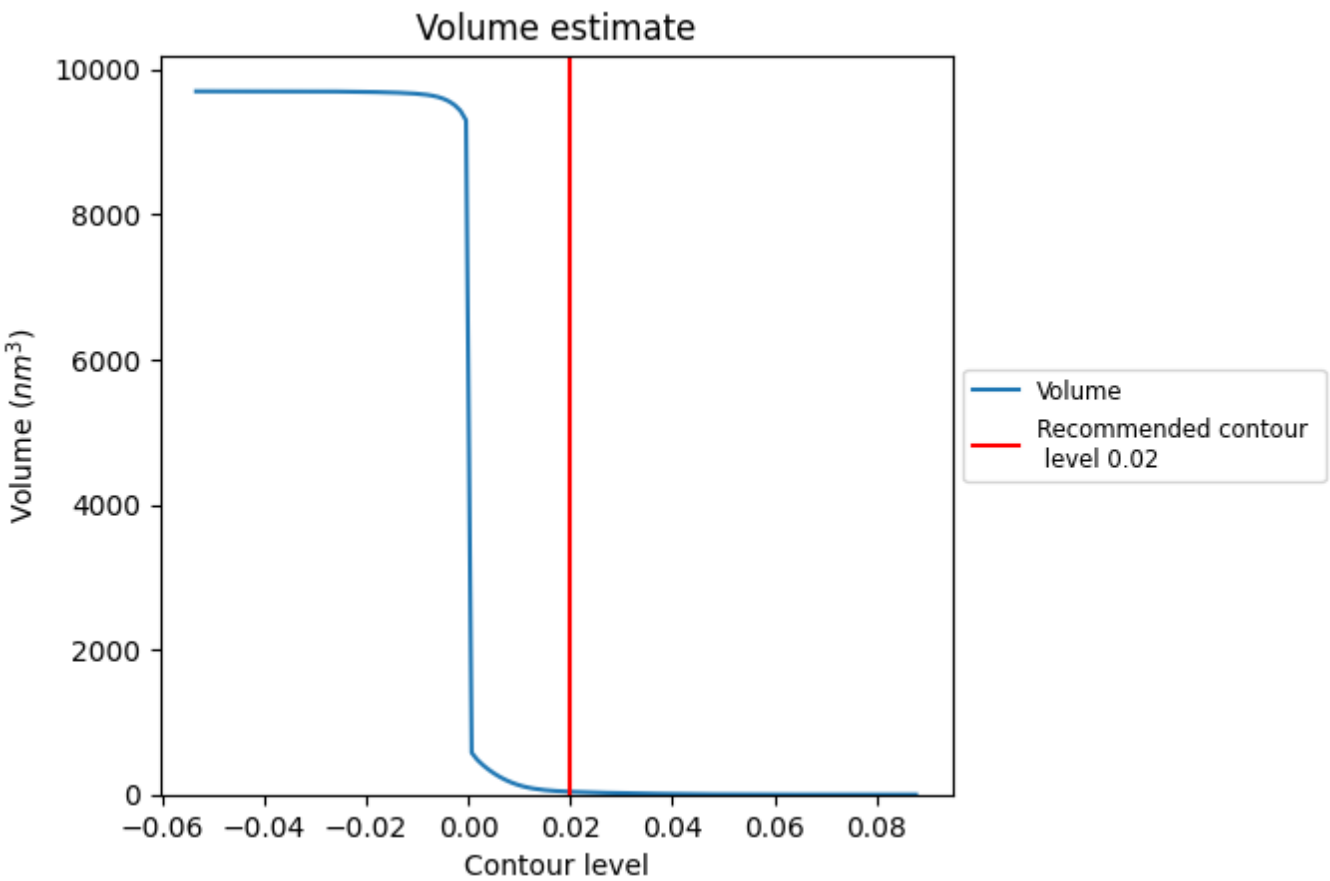
This section contains the results of statistical analysis of the map.

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.

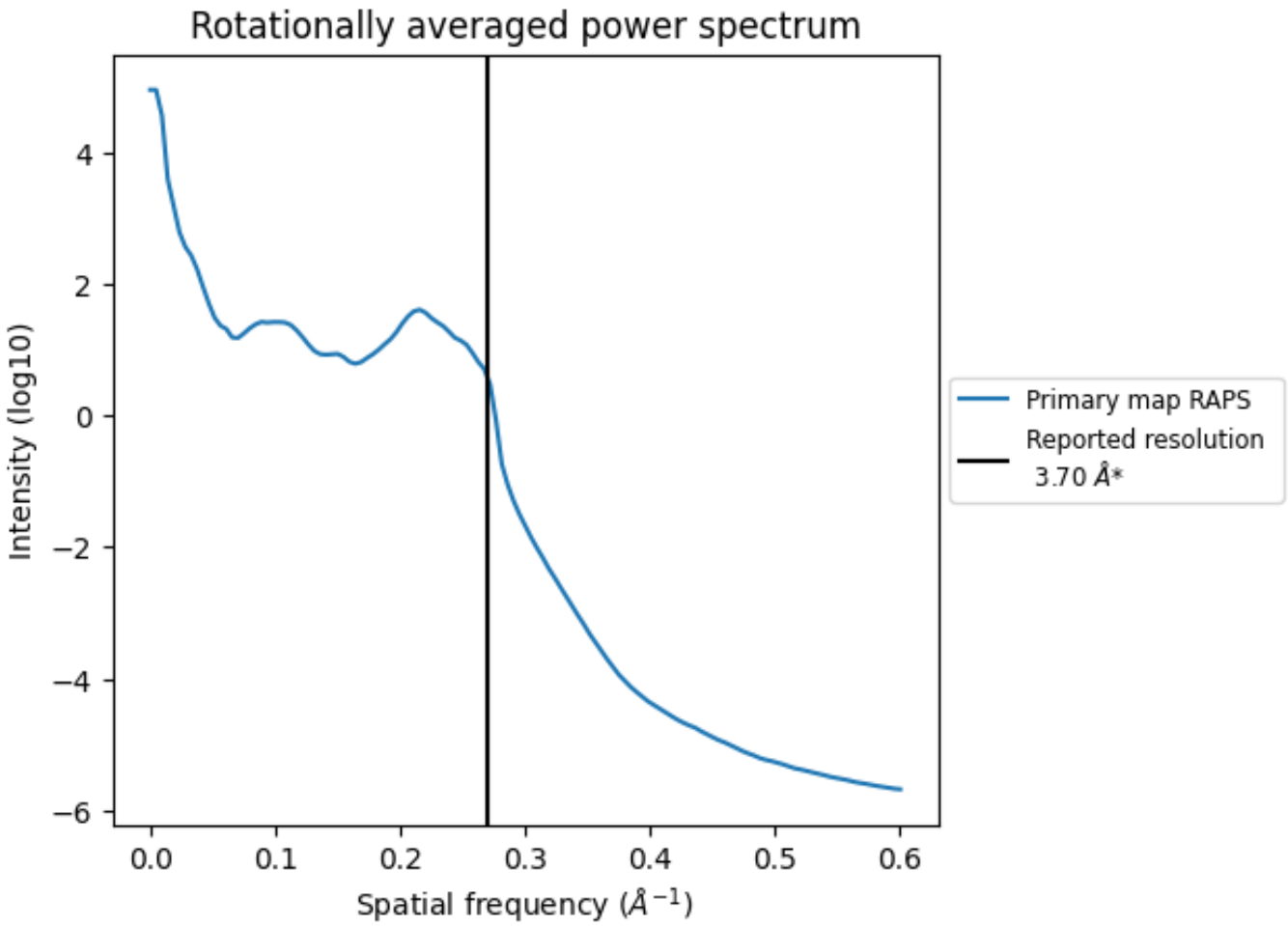
7.2 Volume estimate (i)



The volume at the recommended contour level is 39 nm^3; this corresponds to an approximate mass of 35 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)

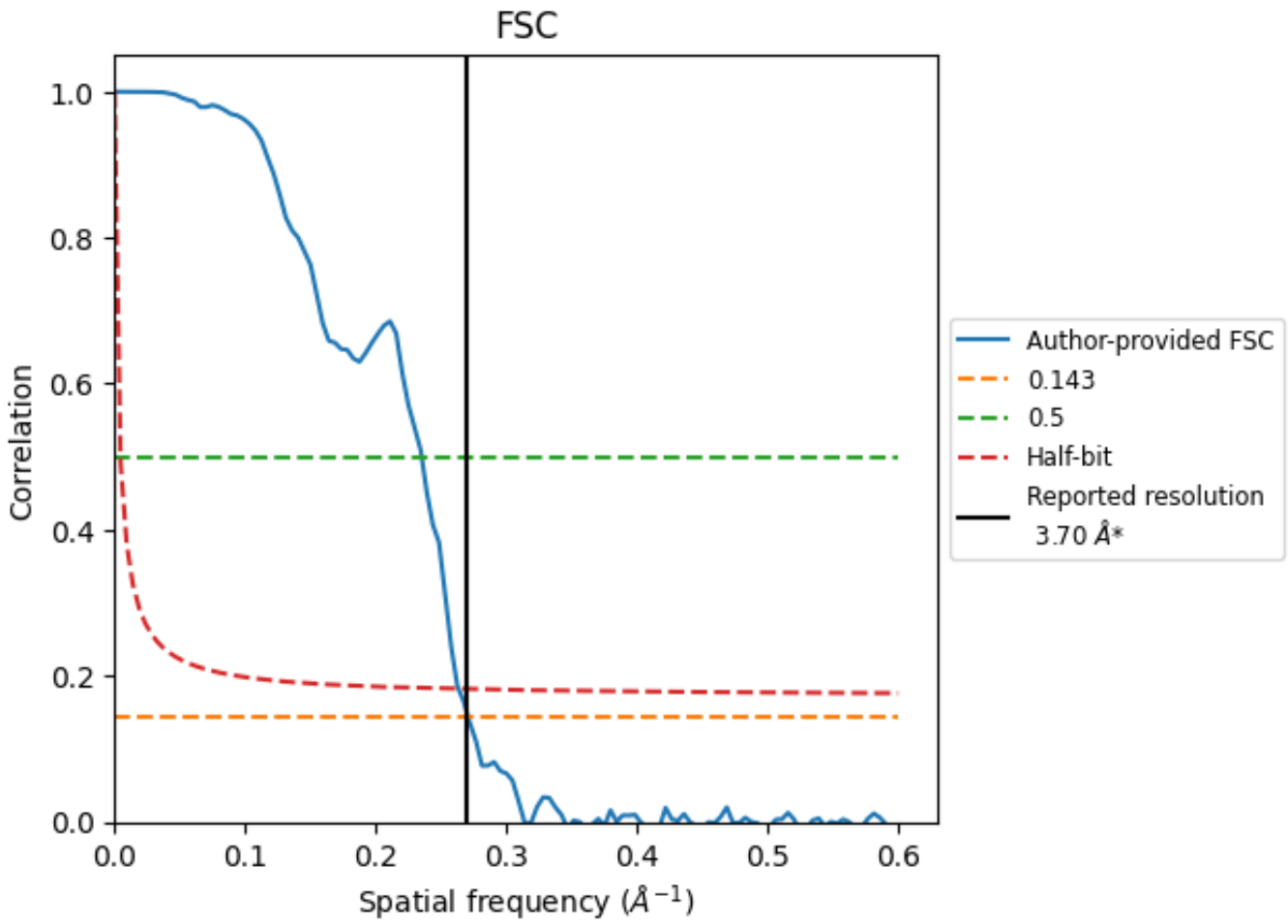


*Reported resolution corresponds to spatial frequency of 0.270 Å^{-1}

8 Fourier-Shell correlation (i)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC (i)



*Reported resolution corresponds to spatial frequency of 0.270 Å^{-1}

8.2 Resolution estimates (i)

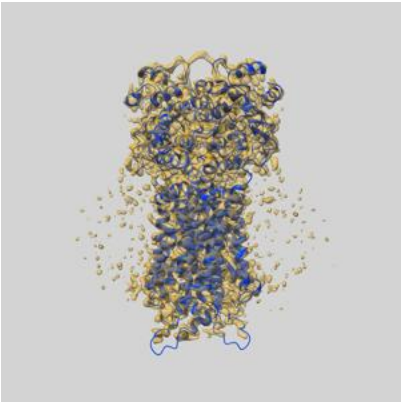
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.70	-	-
Author-provided FSC curve	3.69	4.25	3.80
Unmasked-calculated*	-	-	-

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

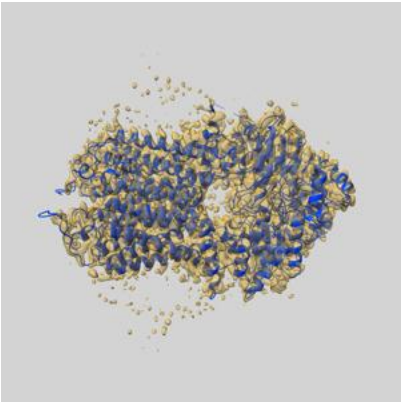
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-24317 and PDB model 7R8E. Per-residue inclusion information can be found in section 3 on page 5.

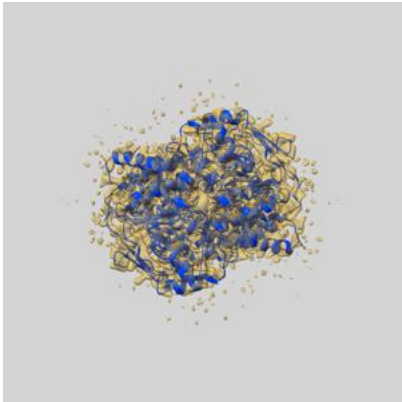
9.1 Map-model overlay (i)



X



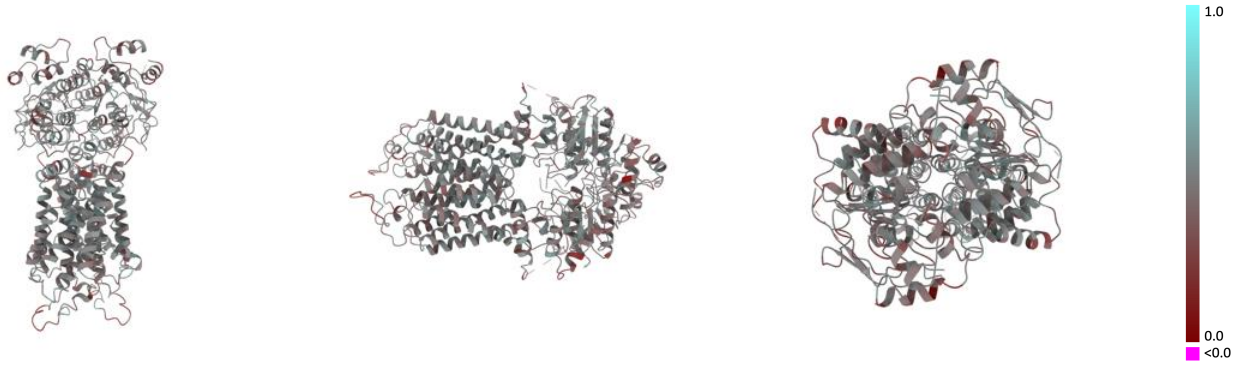
Y



Z

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

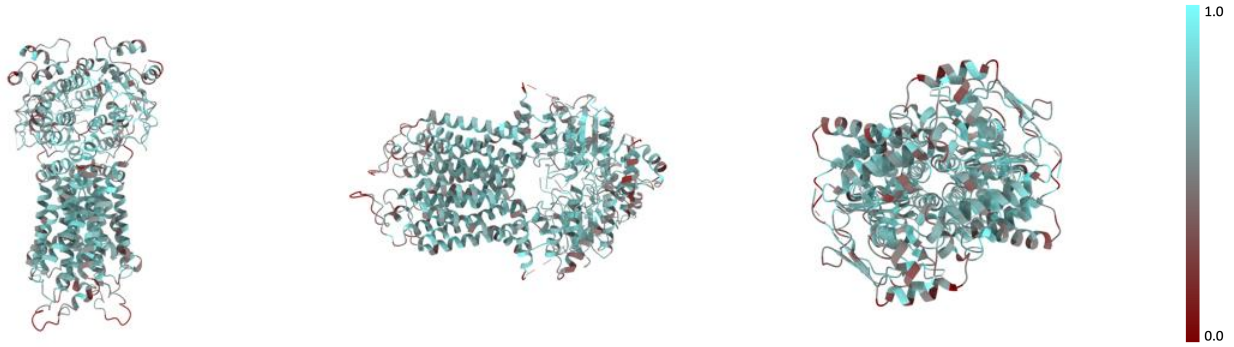
9.2 Q-score mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

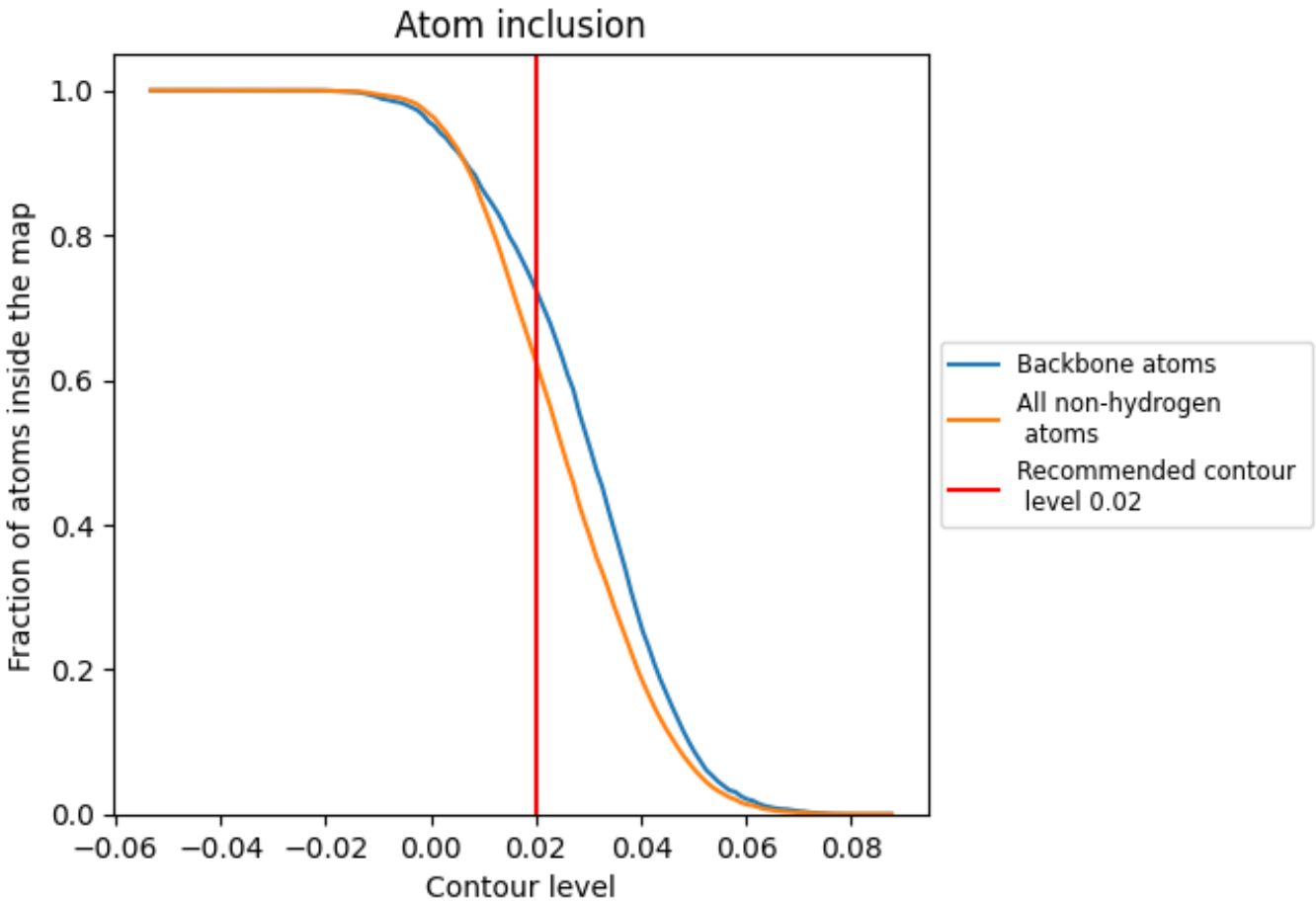
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.02).

9.4 Atom inclusion (i)



At the recommended contour level, 73% of all backbone atoms, 63% 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 (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6252	0.4510
A	0.6338	0.4590
B	0.6165	0.4430