wwPDB EM Validation Summary Report (i)

May 26, 2025 – 04:44 PM EDT

PDB ID	:	6VXJ / pdb_00006vxj
EMDB ID	:	EMD-21440
Title	:	Structure of ABCG2 bound to SN38
Authors	:	Orlando, B.J.; Liao, M.
Deposited on	:	2020-02-22
Resolution	:	4.10 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:

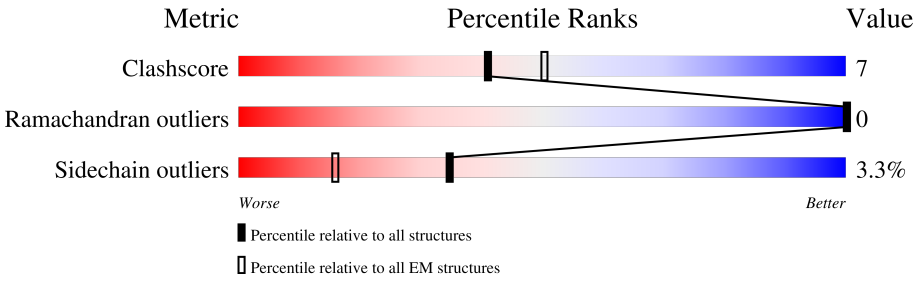
EMDB validation analysis	:	0.0.1.dev118
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4-5-2 with Phenix2.0rc1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 4.10 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	655	45% 68% 17% • 14%
1	B	655	47% 70% 15% • 14%

2 Entry composition

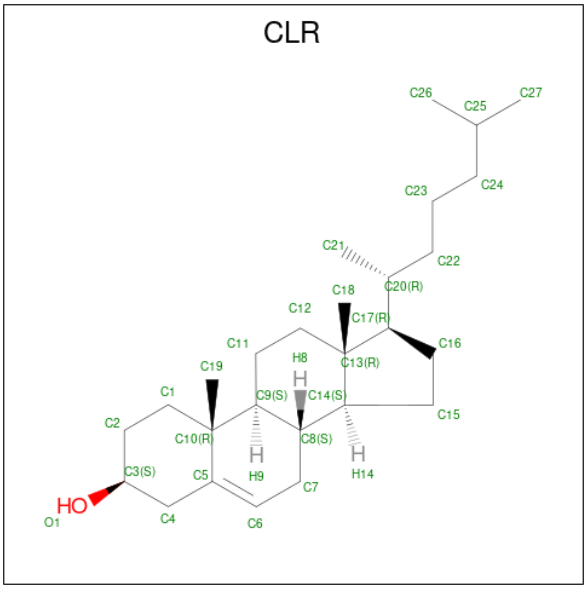
There are 3 unique types of molecules in this entry. The entry contains 8895 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Broad substrate specificity ATP-binding cassette transporter ABCG2.

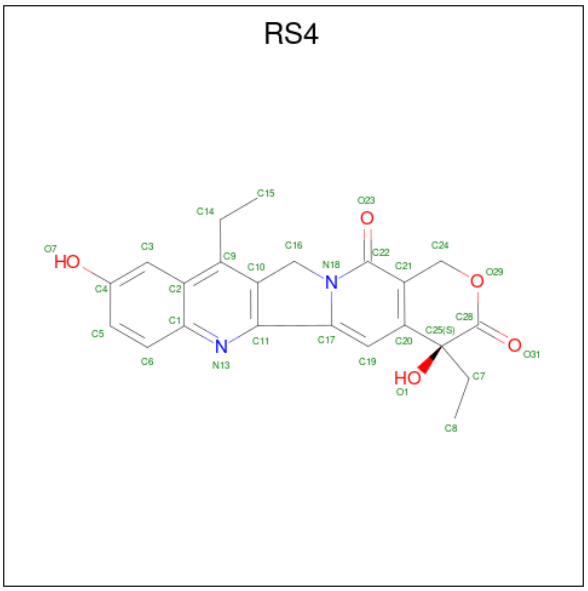
Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	566	Total	C	N	O	S	0	0
			4405	2859	726	791	29		
1	B	566	Total	C	N	O	S	0	0
			4405	2859	726	791	29		

- Molecule 2 is CHOLESTEROL (CCD ID: CLR) (formula: $C_{27}H_{46}O$).



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

- Molecule 3 is 7-ethyl-10-hydroxycamptothecin (CCD ID: RS4) (formula: $C_{22}H_{20}N_2O_5$) (labeled as "Ligand of Interest" by depositor).



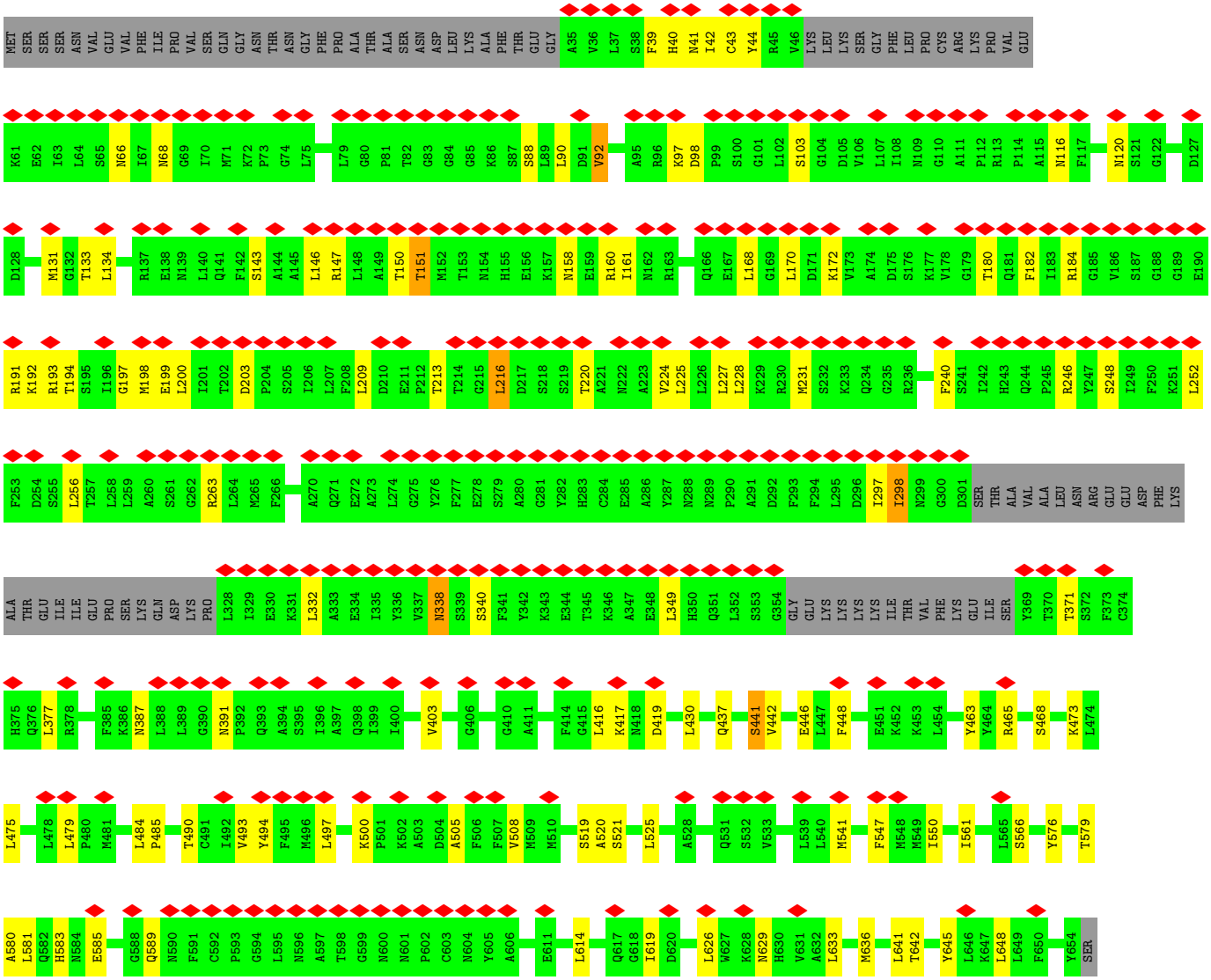
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
3	B	1	29	22	2	5	0

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: Broad substrate specificity ATP-binding cassette transporter ABCG2

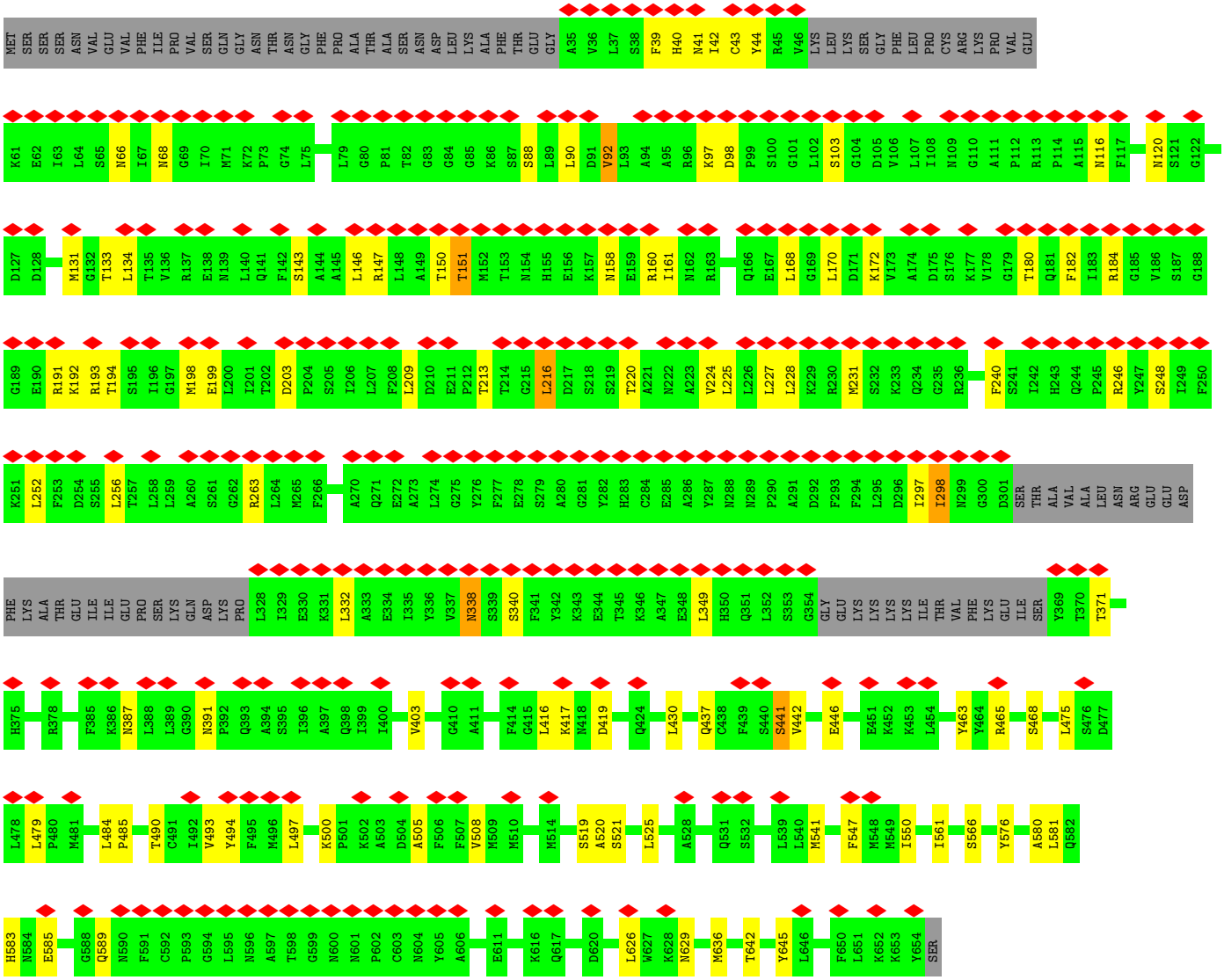
Chain A:



• Molecule 1: Broad substrate specificity ATP-binding cassette transporter ABCG2

Chain B:


47%
70%
15%
14%


MET SER SER SER ASN VAL GLU VAL PHE ILE PRO VAL SER GLN GLY ASN THR ASN GLY PHE PRO ALA THR ALA SER ASN ASP LEU LYS ALA PHE THR GLU GLY A35 V36 L37 S38 F39 H40 N41 I42 C43 Y44 R45 V46 LYS LEU LYS SER GLY PHE LEU PRO CYS ARG LYS PRO VAL GLU
K61 E62 I63 L64 S65 N66 I67 N68 G69 I70 M71 K72 P73 G74 L75 L79 G80 P81 T82 G83 G84 G85 K86 S87 S88 L89 L90 D91 V92 L93 A94 A95 R96 K97 D98 P99 S100 G101 L102 S103 G104 D105 V106 L107 I108 N109 G110 A111 P112 R113 P114 A115 N116 F117 N120 S121 G122
D127 D128 M131 G132 T133 L134 T135 V136 R137 E138 N139 L140 Q141 F142 S143 A144 A145 L146 R147 L148 A149 T150 T151 M152 T153 N154 H155 E156 K157 N158 E159 R160 I161 N162 R163 Q166 E167 L168 G169 L170 D171 K172 V173 A174 D175 S176 K177 V178 G179 T180 Q181 F182 I183 R184 G185 V186 S187 G188
G189 E190 R191 K192 R193 T194 S195 I196 G197 M198 E199 L200 I201 T202 D203 P204 S205 I206 L207 F208 L209 D210 E211 P212 T213 T214 G215 L216 D217 S218 S219 T220 A221 N222 A223 V224 L225 L226 L227 L228 K229 R230 M231 S232 K233 Q234 G235 R236 F240 S241 I242 H243 Q244 P245 R246 Y247 S248 I249 F250
K251 L252 F253 D254 S255 L256 T257 L258 L259 A260 S261 G262 R263 L264 M265 F266 A270 Q271 E272 A273 L274 G275 Y276 F277 E278 S279 A280 G281 Y282 H283 C284 E285 A286 Y287 N288 N289 P290 A291 D292 F293 F294 L295 D296 I297 I298 N299 G300 D301 SER THR ALA VAL ALA LEU ASN ARG GLU GLU ASP
PHE LYS ALA THR GLU ILE ILE GLU PRO SER LYS GLN ASP LYS PRO L328 I329 E330 K331 L332 A333 E334 I335 Y336 V337 N338 S339 S340 F341 Y342 K343 E344 T345 K346 A347 E348 L349 H350 Q351 L352 S353 G354 GLY GLU LYS LYS LYS ILE THR VAL PHE LYS GLU ILE SER Y369 T370 T371
H375 R378 F385 K386 N387 L388 L389 G390 N391 P392 Q393 A394 S395 I396 A397 Q398 I399 I400 V403 G410 A411 F414 G415 L416 K417 N418 D419 Q424 L430 Q437 C438 F439 S440 S441 V442 E446 E451 K452 K453 L454 Y463 Y464 R465 S468 L475 S476 D477
L478 L479 P480 M481 L484 P485 T490 C491 I492 V493 Y494 F495 M496 L497 K500 P501 K502 A503 D504 A505 F506 F507 V508 M509 M510 M514 S519 A520 S521 L525 A528 Q531 S532 L539 L540 M541 F547 M548 M549 I550 I561 S566 Y576 A580 L581 Q582
H583 N584 E585 G588 Q589 N590 F591 C592 P593 G594 L595 N596 A597 T598 G599 N600 N601 P602 C603 N604 Y605 A606 E611 K616 Q617 D620 L626 W627 K628 N629 M636 T642 Y645 L646 F650 L651 K652 K653 Y654 SER

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	97737	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	52	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	31000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.347	Depositor
Minimum map value	-0.214	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.014	Depositor
Recommended contour level	0.085	Depositor
Map size (Å)	236.16, 236.16, 236.16	wwPDB
Map dimensions	192, 192, 192	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.23, 1.23, 1.23	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CLR, RS4

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	0/4497	0.58	0/6085
1	B	0.30	0/4497	0.57	0/6085
All	All	0.30	0/8994	0.58	0/12170

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	B	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	484	LEU	Peptide
1	B	484	LEU	Peptide

5.2 Too-close contacts

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	4405	0	4482	64	0
1	B	4405	0	4482	56	0
2	A	28	0	46	1	0
2	B	28	0	46	2	0
3	B	29	0	0	0	0
All	All	8895	0	9056	121	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 7.

The worst 5 of 121 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:41:ASN:H	1:A:68:ASN:HB3	1.60	0.67
1:A:338:ASN:OD1	1:A:338:ASN:N	2.27	0.67
1:A:430:LEU:HD22	1:A:581:LEU:HD23	1.77	0.67
1:B:41:ASN:H	1:B:68:ASN:HB3	1.60	0.67
1:B:338:ASN:N	1:B:338:ASN:OD1	2.27	0.65

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	558/655 (85%)	505 (90%)	53 (10%)	0	100	100
1	B	558/655 (85%)	505 (90%)	53 (10%)	0	100	100
All	All	1116/1310 (85%)	1010 (90%)	106 (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	482/560 (86%)	466 (97%)	16 (3%)	33 55
1	B	482/560 (86%)	466 (97%)	16 (3%)	33 55
All	All	964/1120 (86%)	932 (97%)	32 (3%)	35 55

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	340	SER
1	B	403	VAL
1	A	340	SER
1	A	338	ASN
1	B	441	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	76	ASN
1	B	222	ASN
1	B	457	HIS
1	B	162	ASN
1	B	244	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

3 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	$\#\|Z\| > 2$
3	RS4	B	701	-	30,33,33	2.89	10 (33%)	32,52,52	2.11	9 (28%)
2	CLR	A	701	-	31,31,31	0.84	2 (6%)	48,48,48	1.61	7 (14%)
2	CLR	B	702	-	31,31,31	0.83	2 (6%)	48,48,48	1.59	7 (14%)

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
3	RS4	B	701	-	-	4/5/30/30	0/5/5/5
2	CLR	A	701	-	-	7/10/68/68	0/4/4/4
2	CLR	B	702	-	-	7/10/68/68	0/4/4/4

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	701	RS4	O29-C28	8.15	1.44	1.34
3	B	701	RS4	C21-C20	7.12	1.47	1.36
3	B	701	RS4	C10-C9	6.00	1.48	1.38
3	B	701	RS4	C2-C1	4.18	1.49	1.42
3	B	701	RS4	C22-N18	-4.07	1.33	1.40

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	701	RS4	C17-C11-N13	6.46	129.37	125.32
3	B	701	RS4	C10-C11-C17	-5.59	105.21	108.38
2	A	701	CLR	C13-C17-C20	-4.96	111.84	119.50
2	B	702	CLR	C13-C17-C20	-4.91	111.92	119.50
2	A	701	CLR	C13-C14-C8	-4.48	108.06	114.41

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	701	RS4	C20-C25-C7-C8
3	B	701	RS4	C28-C25-C7-C8
3	B	701	RS4	O1-C25-C7-C8
2	A	701	CLR	C21-C20-C22-C23
2	B	702	CLR	C21-C20-C22-C23

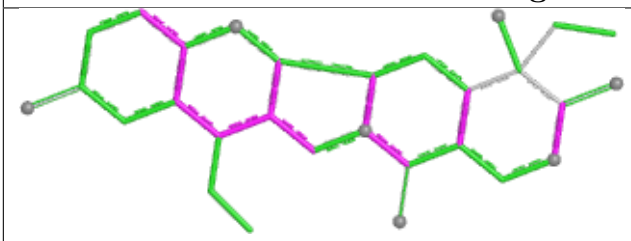
There are no ring outliers.

2 monomers are involved in 3 short contacts:

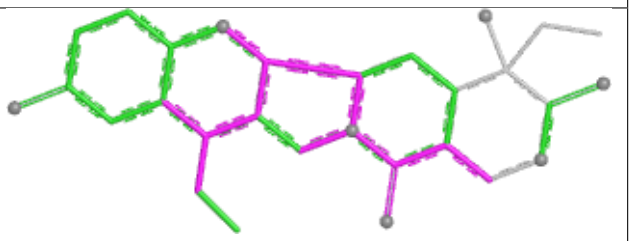
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	701	CLR	1	0
2	B	702	CLR	2	0

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.

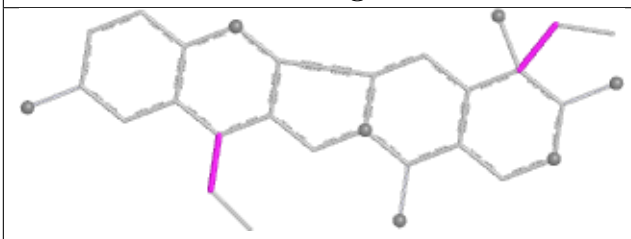
Ligand RS4 B 701



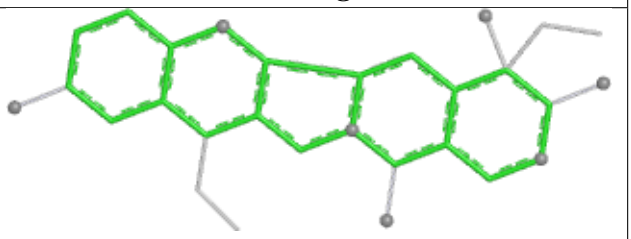
Bond lengths



Bond angles

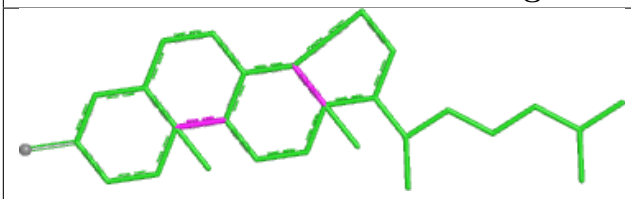


Torsions

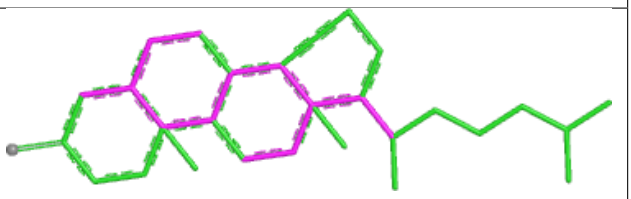


Rings

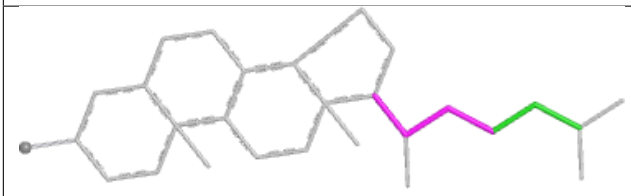
Ligand CLR A 701



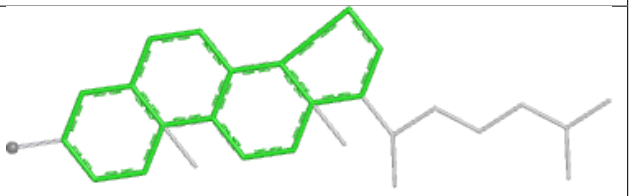
Bond lengths



Bond angles

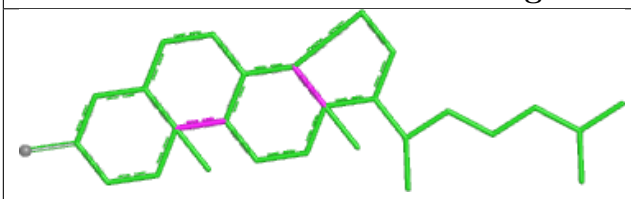


Torsions

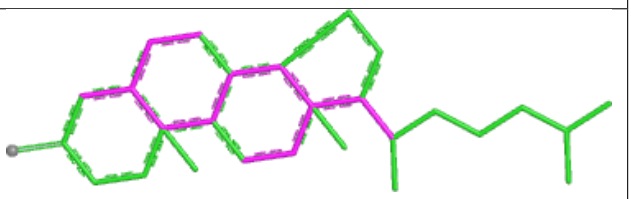


Rings

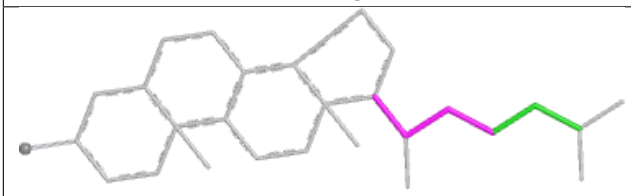
Ligand CLR B 702



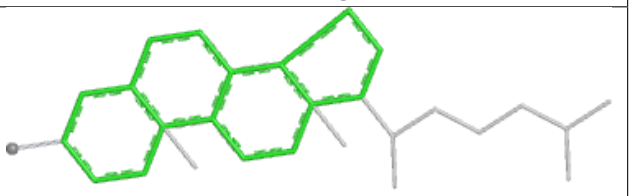
Bond lengths



Bond angles



Torsions



Rings

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.

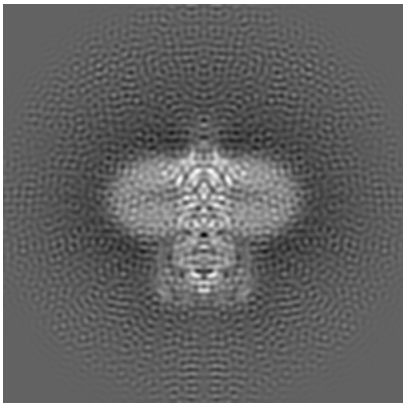
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-21440. These allow visual inspection of the internal detail of the map and identification of artifacts.

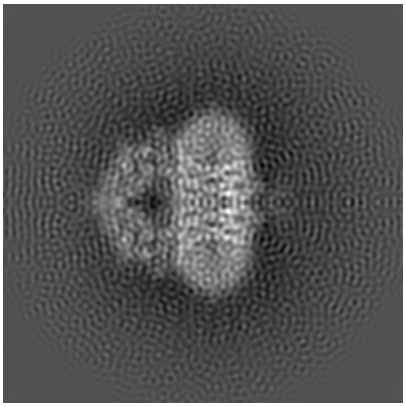
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

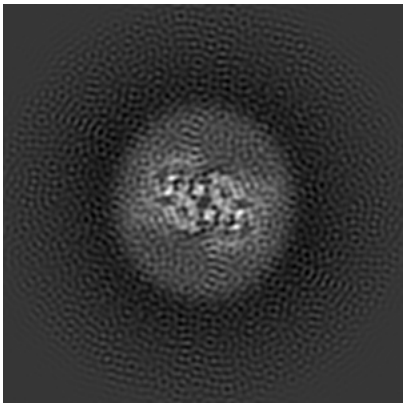
6.1.1 Primary map



X



Y

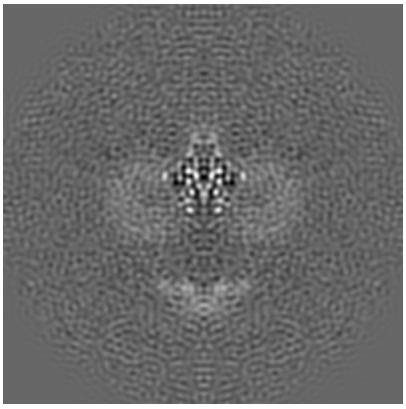


Z

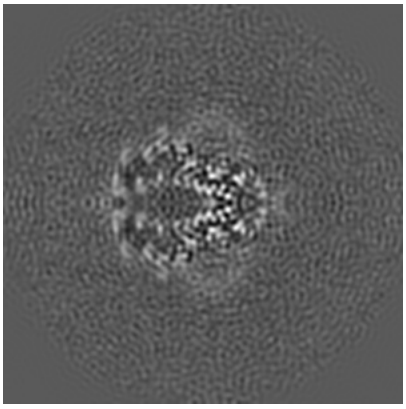
The images above show the map projected in three orthogonal directions.

6.2 Central slices

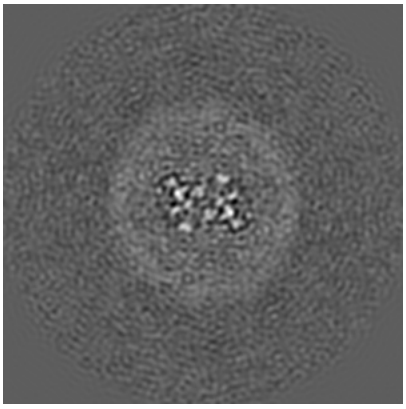
6.2.1 Primary map



X Index: 96



Y Index: 96

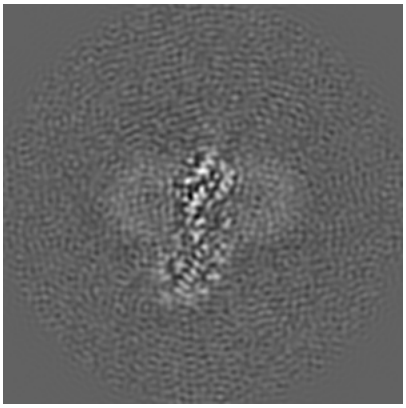


Z Index: 96

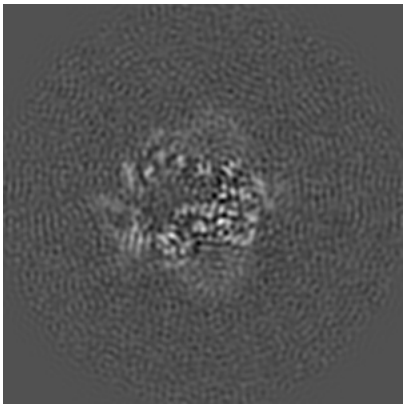
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

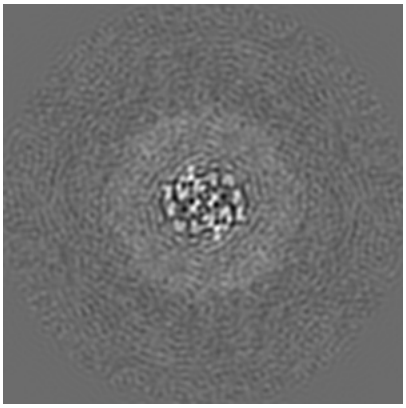
6.3.1 Primary map



X Index: 108



Y Index: 101

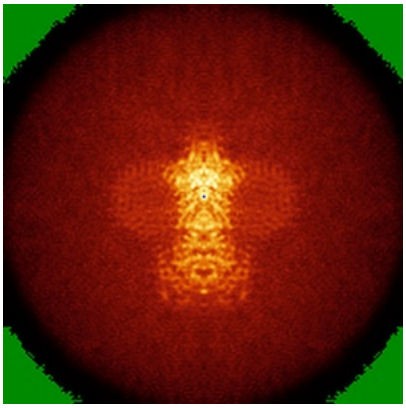


Z Index: 107

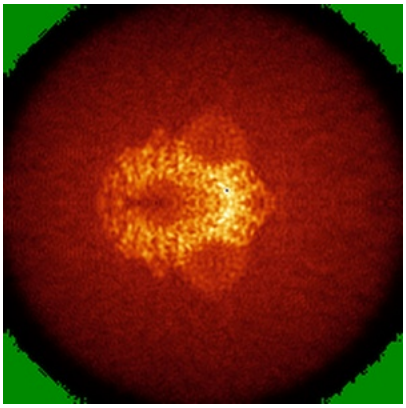
The images above show the largest variance slices of the map in three orthogonal directions.

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

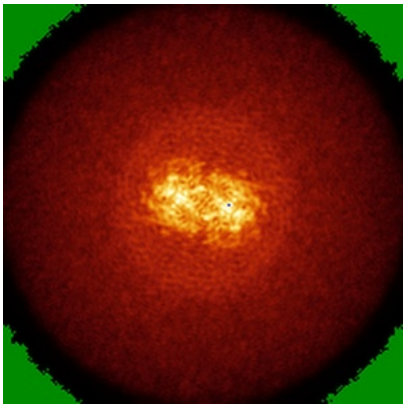
6.4.1 Primary map



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

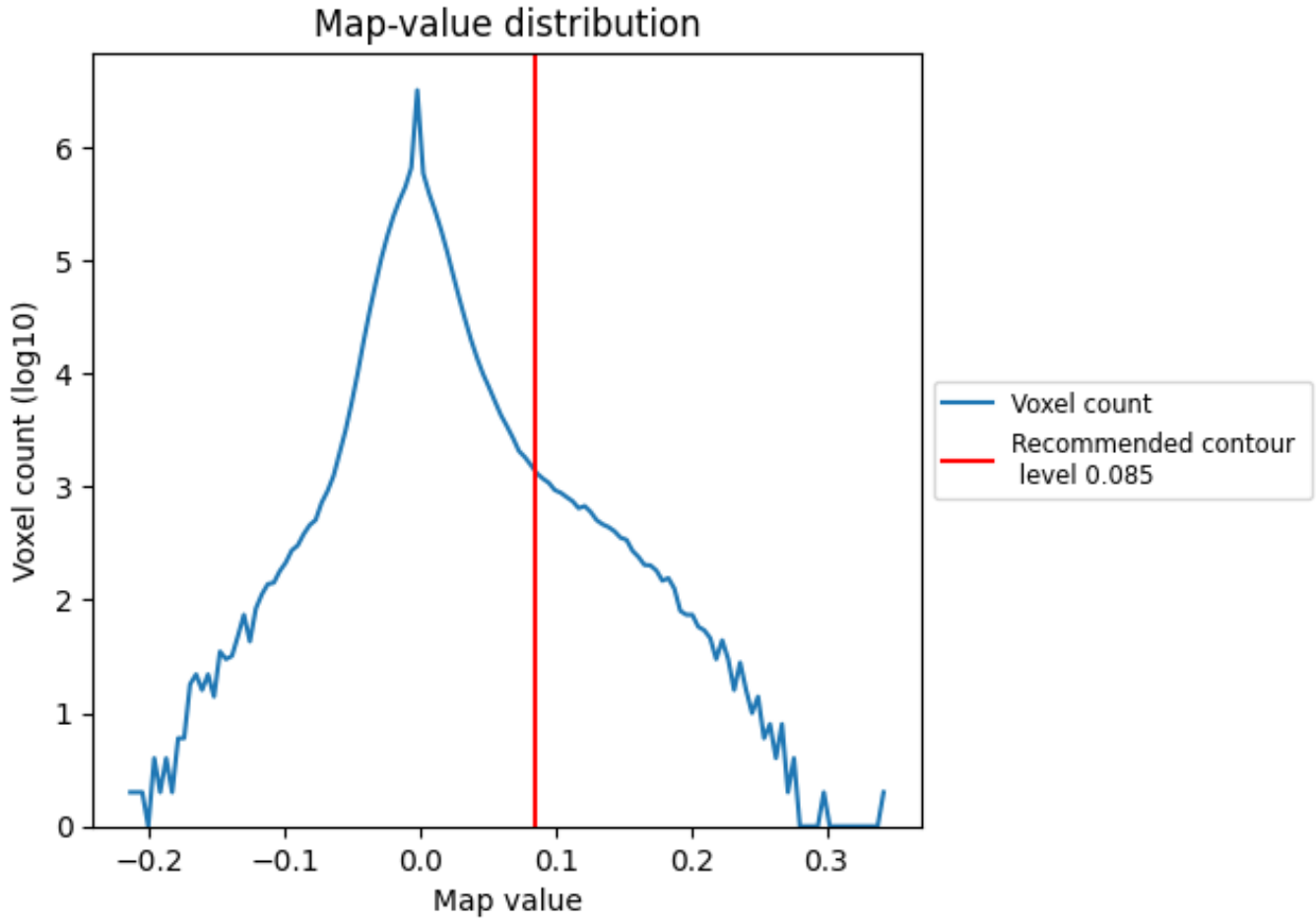
6.6 Mask visualisation (i)

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

7 Map analysis

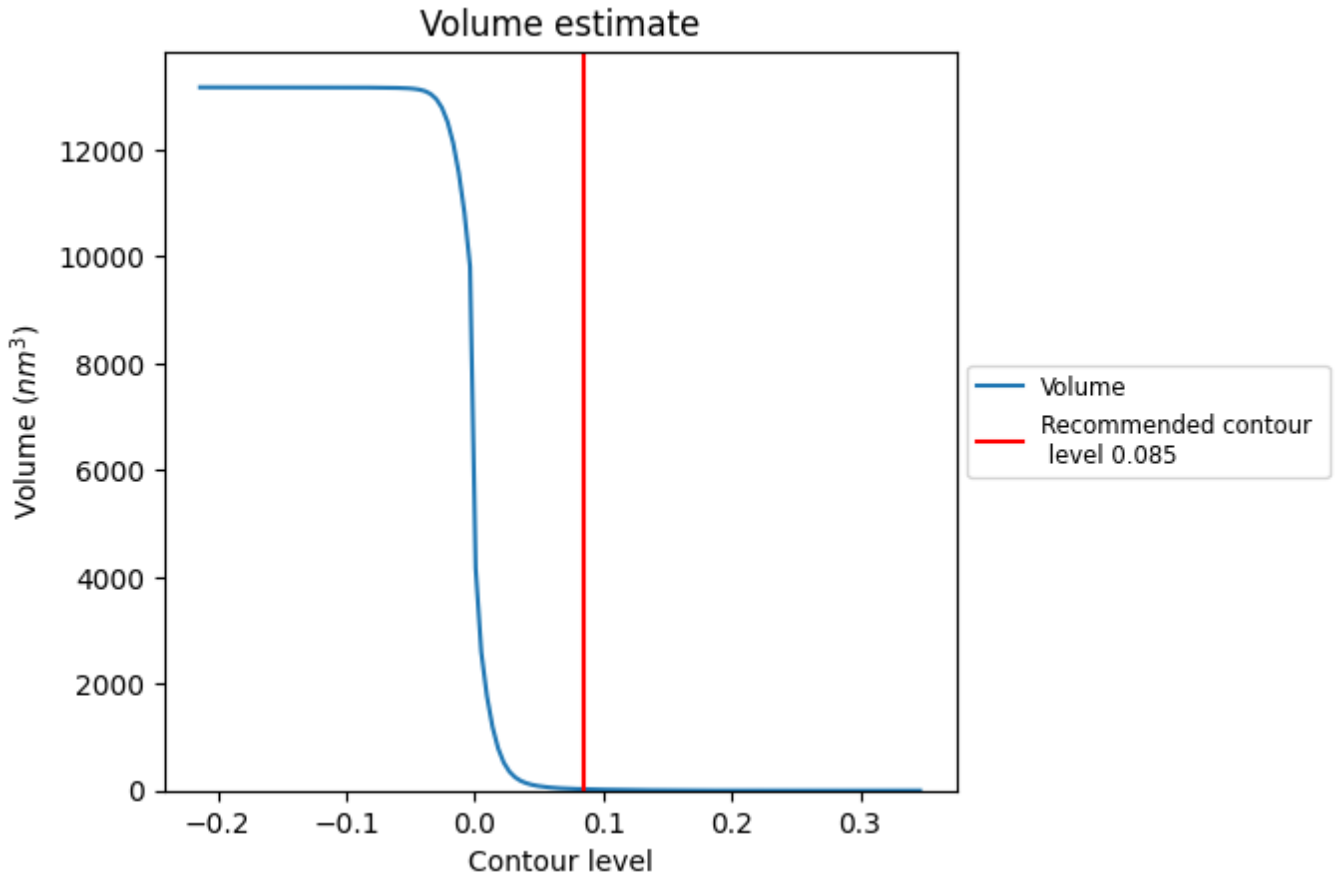
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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.

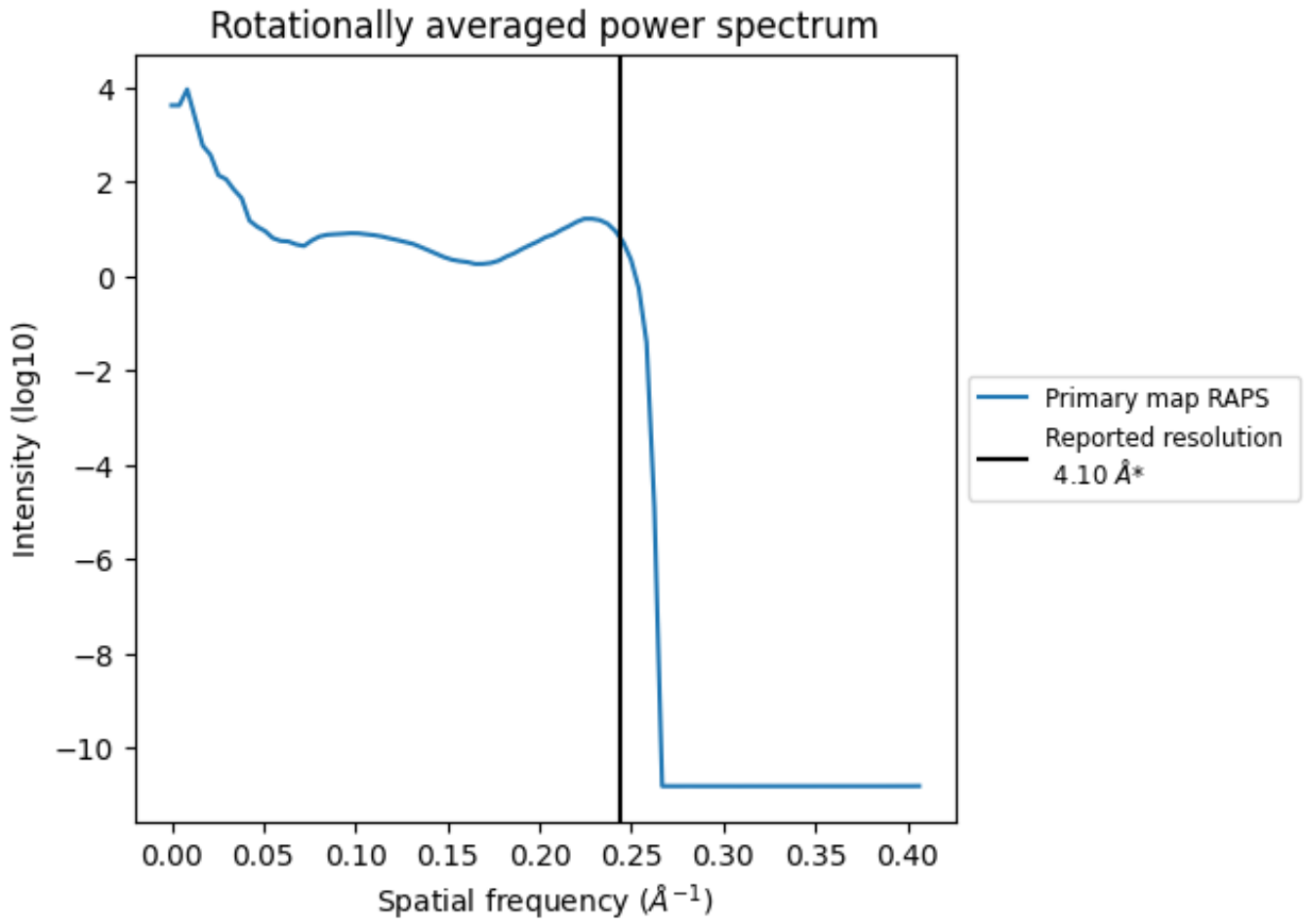
7.2 Volume estimate (i)



The volume at the recommended contour level is 26 nm^3; this corresponds to an approximate mass of 23 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)

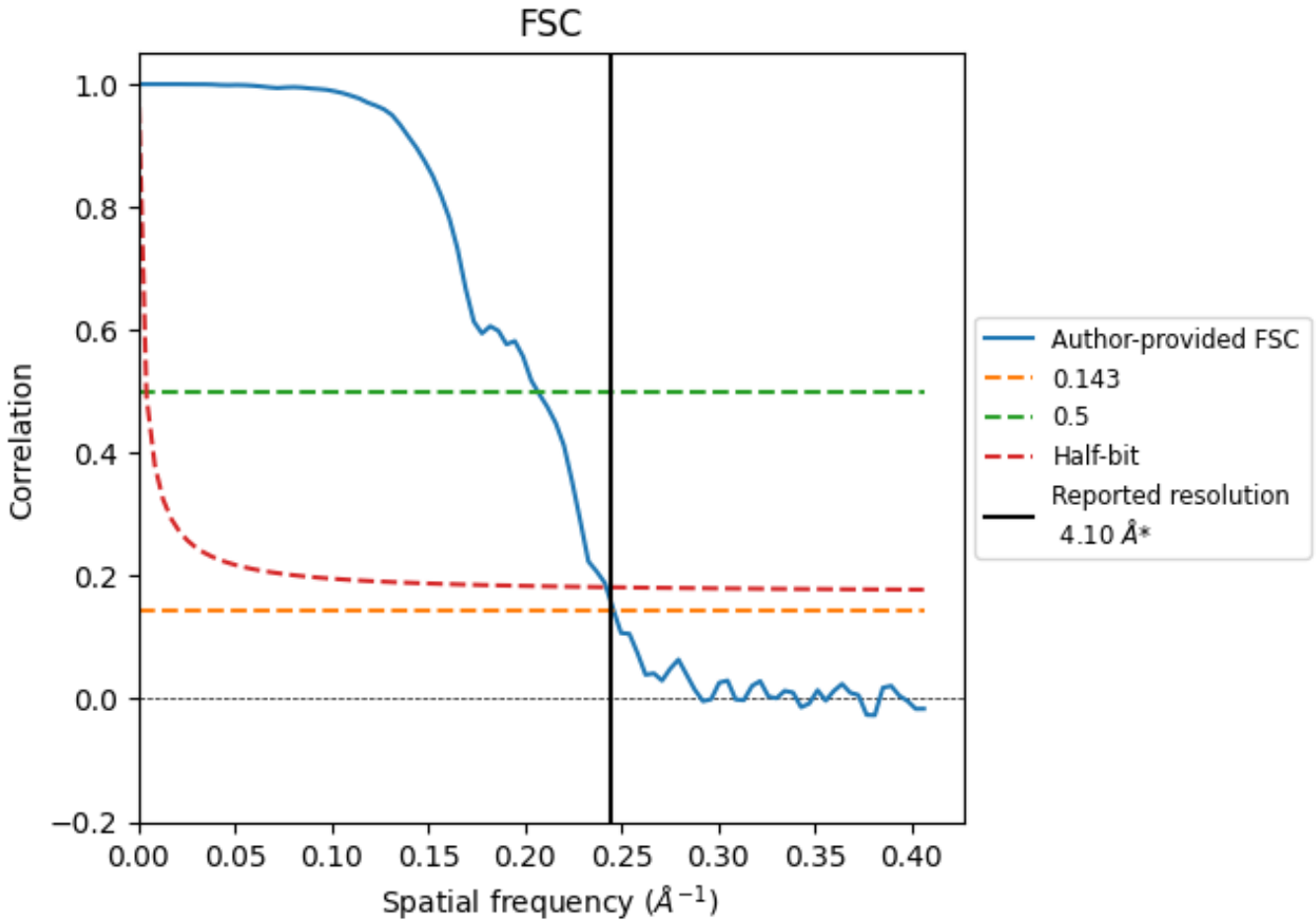


*Reported resolution corresponds to spatial frequency of 0.244 Å^{-1}

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.244 Å^{-1}

8.2 Resolution estimates (i)

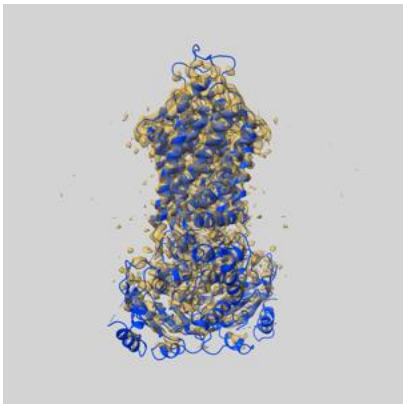
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.10	-	-
Author-provided FSC curve	4.07	4.84	4.13
Unmasked-calculated*	-	-	-

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

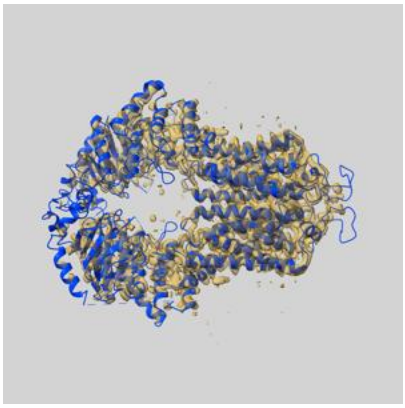
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-21440 and PDB model 6VXJ. Per-residue inclusion information can be found in section 3 on page 5.

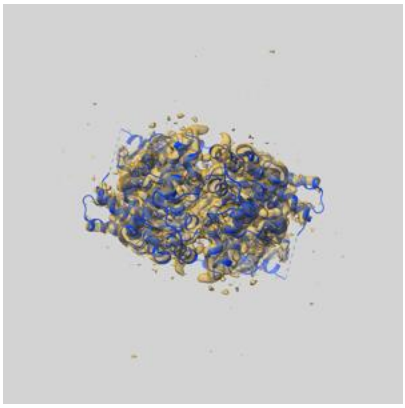
9.1 Map-model overlay



X



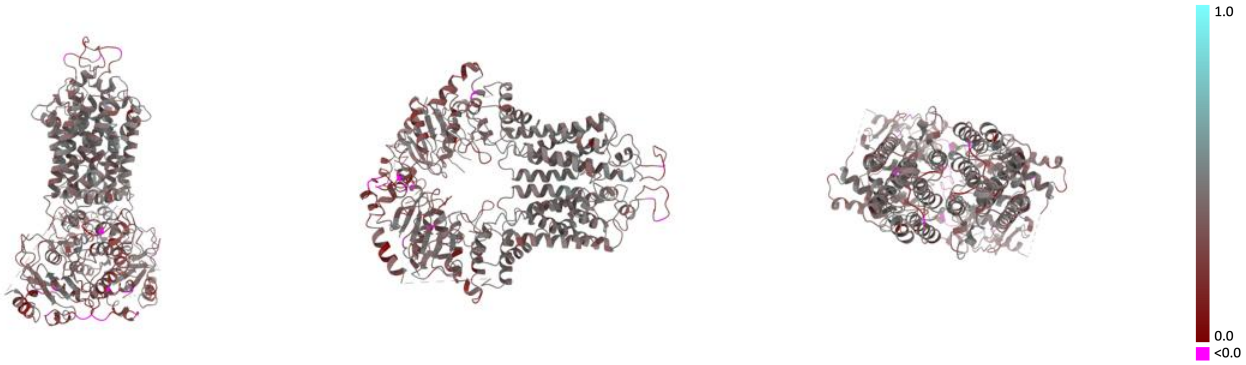
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.085 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.

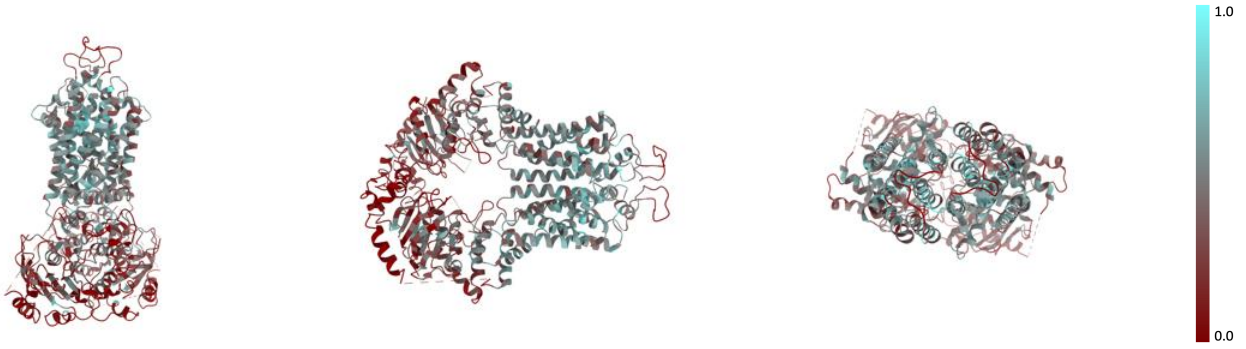
9.2 Q-score mapped to coordinate model (i)



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.

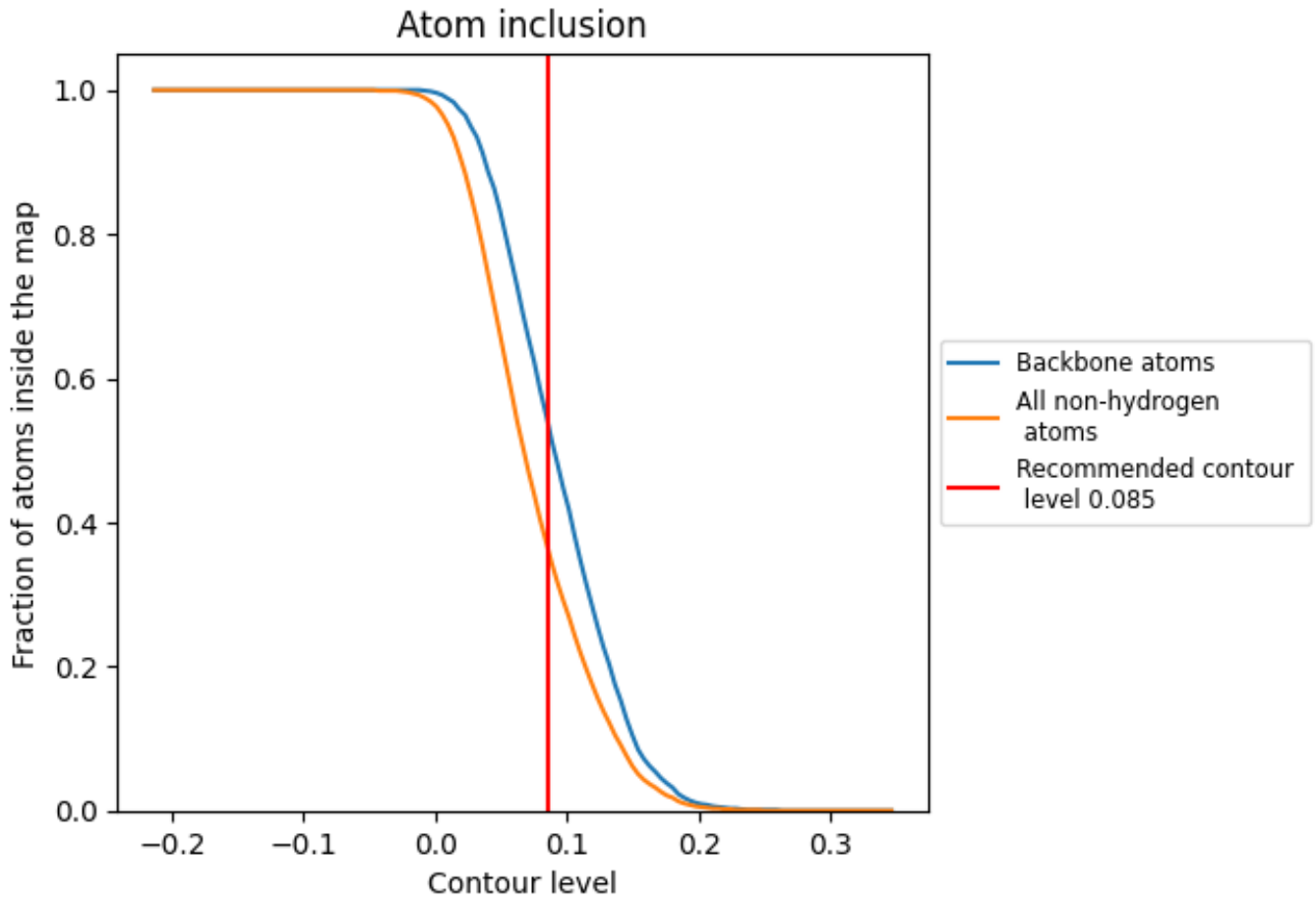
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.085).

9.4 Atom inclusion (i)



At the recommended contour level, 54% of all backbone atoms, 37% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.085) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.3670	0.3760
A	0.3670	0.3740
B	0.3660	0.3780