wwPDB EM Validation Summary Report (i)

Apr 28, 2024 – 10:14 am BST

PDB ID	:	4AQW
EMDB ID	:	EMD-2078
Title	:	Model of human kinesin-5 motor domain (1II6, 3HQD) and mammalian tubulin heterodimer (1JFF) docked into the 9.5-angstrom cryo-EM map of microtubule-bound kinesin-5 motor domain in the rigor state.
Authors	:	Goulet, A.; Behnke-Parks, W.M.; Sindelar, C.V.; Rosenfeld, S.S.; Moores, C.A.
Deposited on	:	2012-04-19
Resolution	:	9.50 Å(reported)
Based on initial models	:	1II6, 3HQD, 1JFF

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:

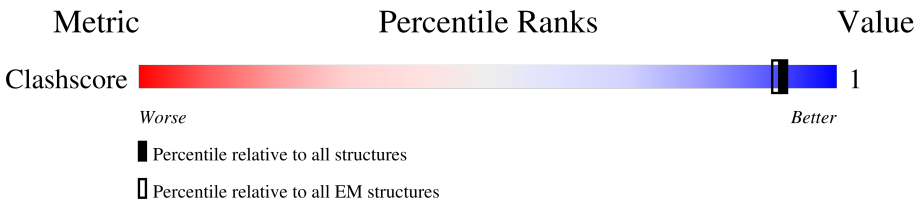
EMDB validation analysis	:	0.0.1.dev92
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 9.50 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

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	452	14% 91% 9%
2	B	445	16% 96% .
3	C	373	25% 94% 6%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 1311 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 TUBULIN ALPHA-1D CHAIN.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
1	A	412	412	412	0	412

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ILE	VAL	conflict	UNP Q2HJ86
A	114	ILE	LEU	conflict	UNP Q2HJ86
A	136	SER	LEU	conflict	UNP Q2HJ86
A	137	VAL	ILE	conflict	UNP Q2HJ86
A	265	GLY	ILE	conflict	UNP Q2HJ86
A	358	GLU	GLN	conflict	UNP Q2HJ86
A	437	VAL	MET	conflict	UNP Q2HJ86

- Molecule 2 is a protein called TUBULIN BETA-2B CHAIN.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
2	B	426	426	426	0	426

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	57	ALA	THR	conflict	UNP Q6B856
B	172	VAL	MET	conflict	UNP Q6B856
B	298	ALA	SER	conflict	UNP Q6B856
B	318	VAL	ILE	conflict	UNP Q6B856

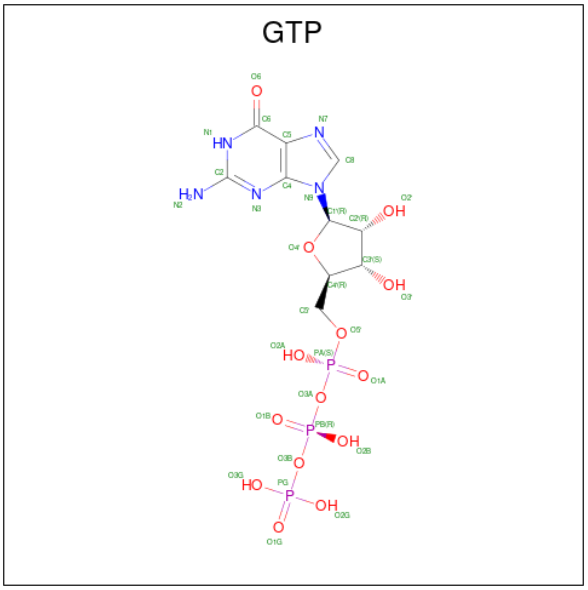
- Molecule 3 is a protein called KINESIN-LIKE PROTEIN KIF11.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
3	C	350	350	350	0	350

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	25	VAL	CYS	engineered mutation	UNP P52732
C	43	SER	CYS	engineered mutation	UNP P52732
C	87	ALA	CYS	engineered mutation	UNP P52732
C	99	ALA	CYS	engineered mutation	UNP P52732
C	126	CYS	THR	engineered mutation	UNP P52732
C	368	HIS	-	expression tag	UNP P52732
C	369	HIS	-	expression tag	UNP P52732
C	370	HIS	-	expression tag	UNP P52732
C	371	HIS	-	expression tag	UNP P52732
C	372	HIS	-	expression tag	UNP P52732
C	373	HIS	-	expression tag	UNP P52732

- Molecule 4 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



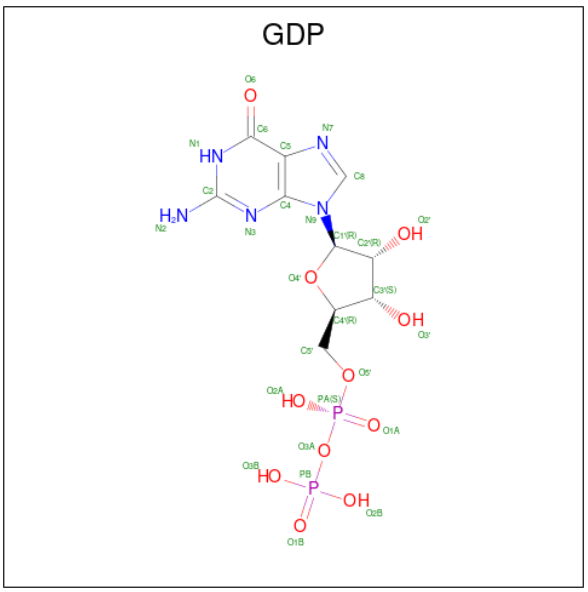
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
4	A	1	32	10	5	14	3	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
5	A	1	1	1	0

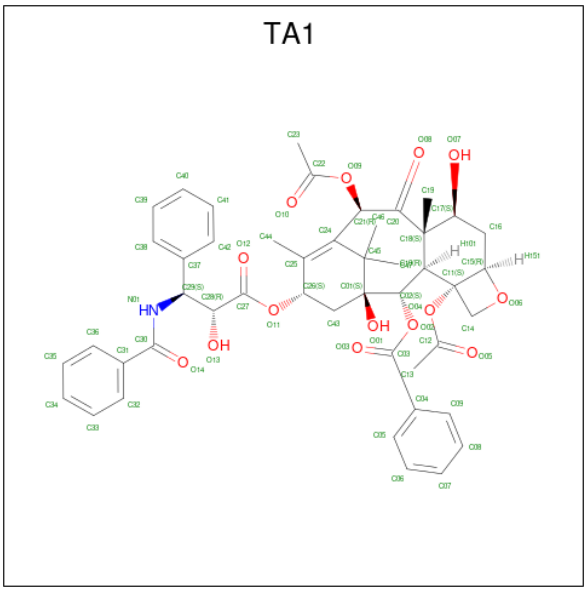
- Molecule 6 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula:

$C_{10}H_{15}N_5O_{11}P_2$).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
6	B	1	28	10	5	11	2	0

- Molecule 7 is TAXOL (three-letter code: TA1) (formula: $C_{47}H_{51}NO_{14}$).



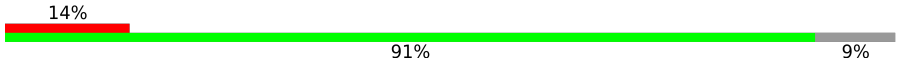
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
7	B	1	62	47	1	14	0

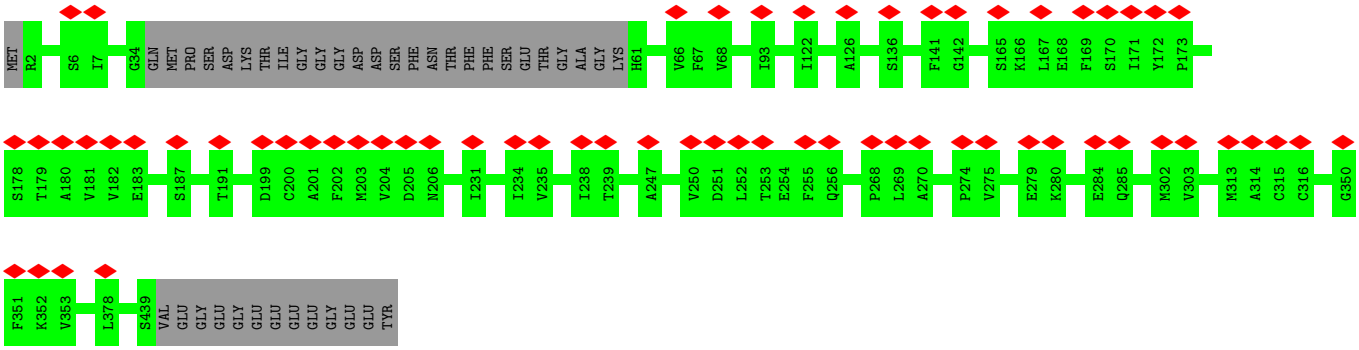
3 Residue-property plots (i)

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: TUBULIN ALPHA-1D CHAIN

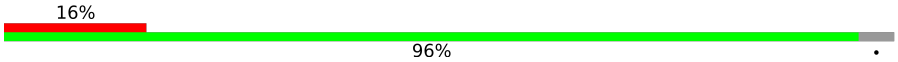
Chain A:

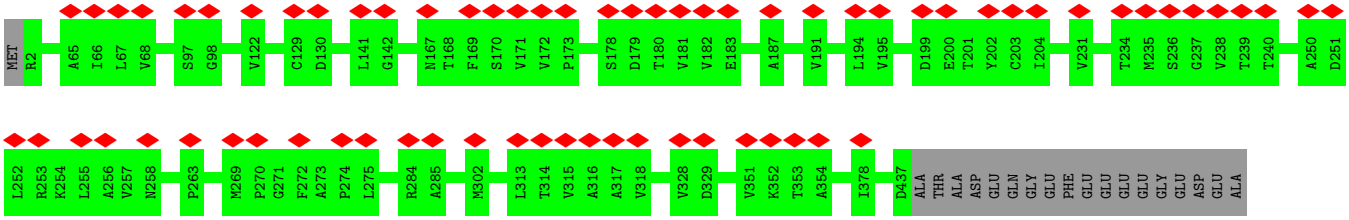




- Molecule 2: TUBULIN BETA-2B CHAIN

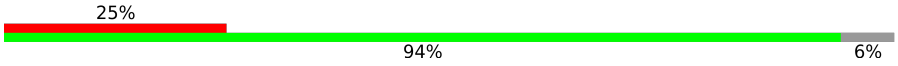
Chain B:

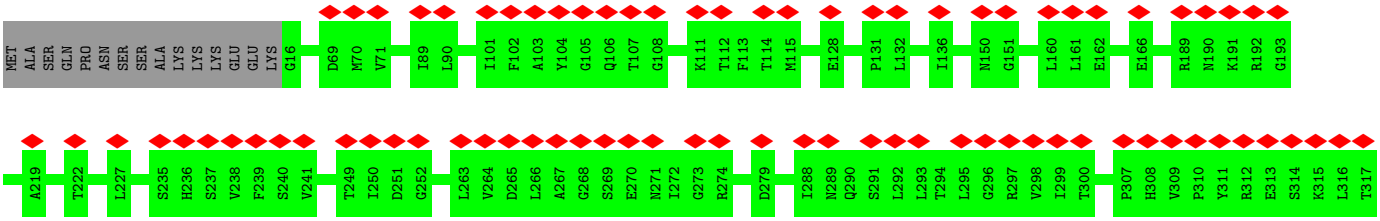


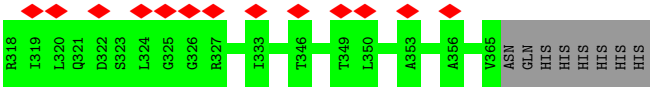


- Molecule 3: KINESIN-LIKE PROTEIN KIF11

Chain C:






R318
I319
L320
Q321
D322
S323
L324
G325
G326
R327
I333
T346
T349
L350
A353
A356
V365
ASN
GLN
HIS
HIS
HIS
HIS
HIS
HIS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	HELICAL, twist=Not provided°, rise=Not provided Å, axial sym=Not provided	Depositor
Number of particles used	6748	Depositor
Resolution determination method	Not provided	
CTF correction method	FREALIGN	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	18	Depositor
Minimum defocus (nm)	700	Depositor
Maximum defocus (nm)	2300	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 FILM	Depositor
Maximum map value	8.230	Depositor
Minimum map value	-6.370	Depositor
Average map value	0.075	Depositor
Map value standard deviation	1.896	Depositor
Recommended contour level	1.0	Depositor
Map size (Å)	126.0, 126.0, 126.0	wwPDB
Map dimensions	45, 45, 45	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.8, 2.8, 2.8	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: GTP, GDP, MG, TA1

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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

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	412	0	0	0	0
2	B	426	0	0	0	0
3	C	350	0	0	0	0
4	A	32	0	12	0	0
5	A	1	0	0	0	0
6	B	28	0	12	0	0
7	B	62	0	51	2	0
All	All	1311	0	75	2	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 1.

All (2) 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)
7:B:601:TA1:H463	7:B:601:TA1:H261	1.80	0.63
7:B:601:TA1:H463	7:B:601:TA1:C26	2.46	0.44

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	$\#\|Z\| > 2$
4	GTP	A	500	5	26,34,34	1.32	4 (15%)	32,54,54	1.10	3 (9%)
6	GDP	B	600	-	24,30,30	2.72	9 (37%)	30,47,47	2.93	9 (30%)
7	TA1	B	601	-	68,68,68	2.01	19 (27%)	105,105,105	1.39	12 (11%)

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
4	GTP	A	500	5	-	3/18/38/38	0/3/3/3
6	GDP	B	600	-	-	4/12/32/32	0/3/3/3
7	TA1	B	601	-	-	9/41/127/127	0/7/7/7

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	600	GDP	O4'-C1'	6.20	1.49	1.41
6	B	600	GDP	O6-C6	5.71	1.34	1.23
7	B	601	TA1	C08-C09	5.29	1.50	1.38
7	B	601	TA1	C18-C10	5.10	1.69	1.57
6	B	600	GDP	C2-N1	4.65	1.49	1.37

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	600	GDP	C8-N7-C5	9.26	120.64	102.99
6	B	600	GDP	N2-C2-N3	6.29	131.97	119.74
6	B	600	GDP	C5-C6-N1	6.08	124.69	113.95
7	B	601	TA1	C08-C09-C04	-4.82	114.63	120.34
7	B	601	TA1	C07-C06-C05	4.75	127.43	120.19

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	600	GDP	PA-O3A-PB-O2B
6	B	600	GDP	C5'-O5'-PA-O3A
7	B	601	TA1	O02-C03-C04-C09
7	B	601	TA1	O02-C03-C04-C05

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	B	601	TA1	O03-C03-C04-C05

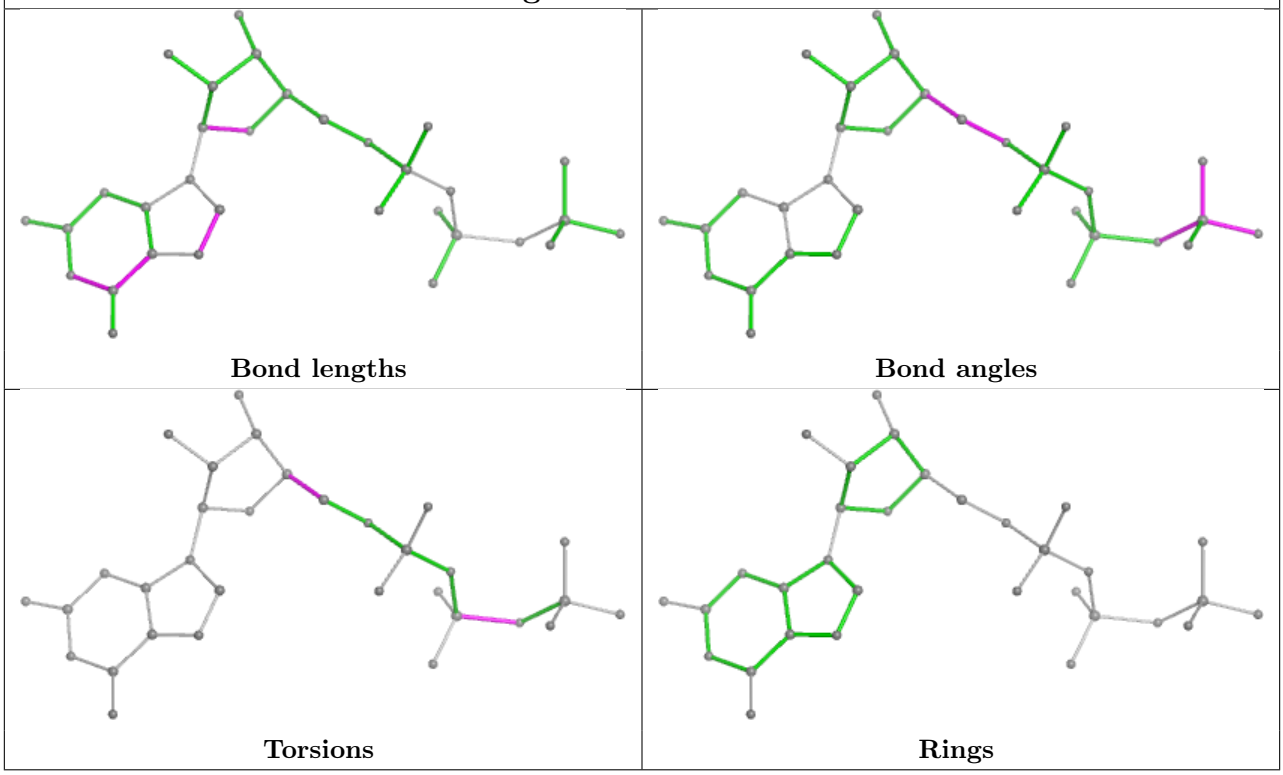
There are no ring outliers.

1 monomer is involved in 2 short contacts:

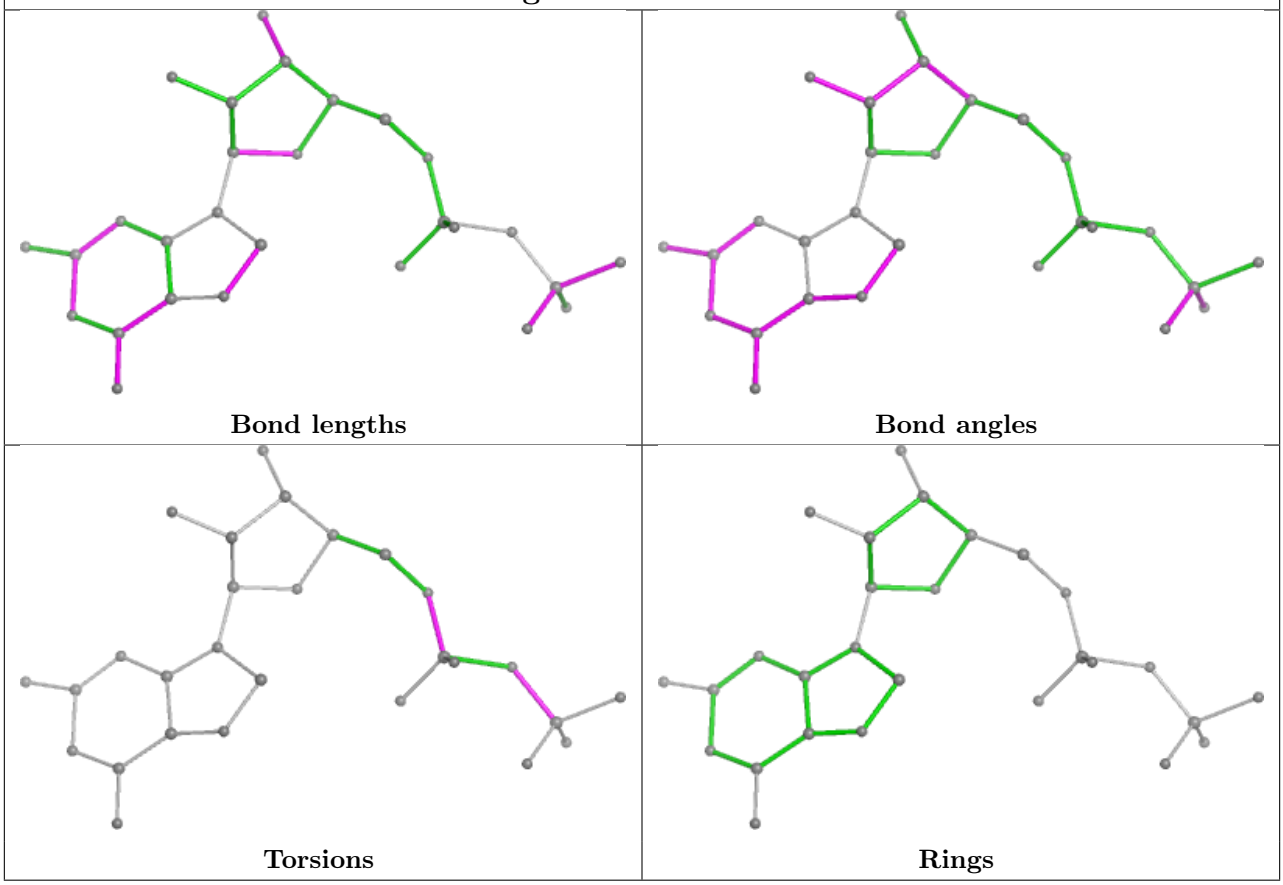
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	601	TA1	2	0

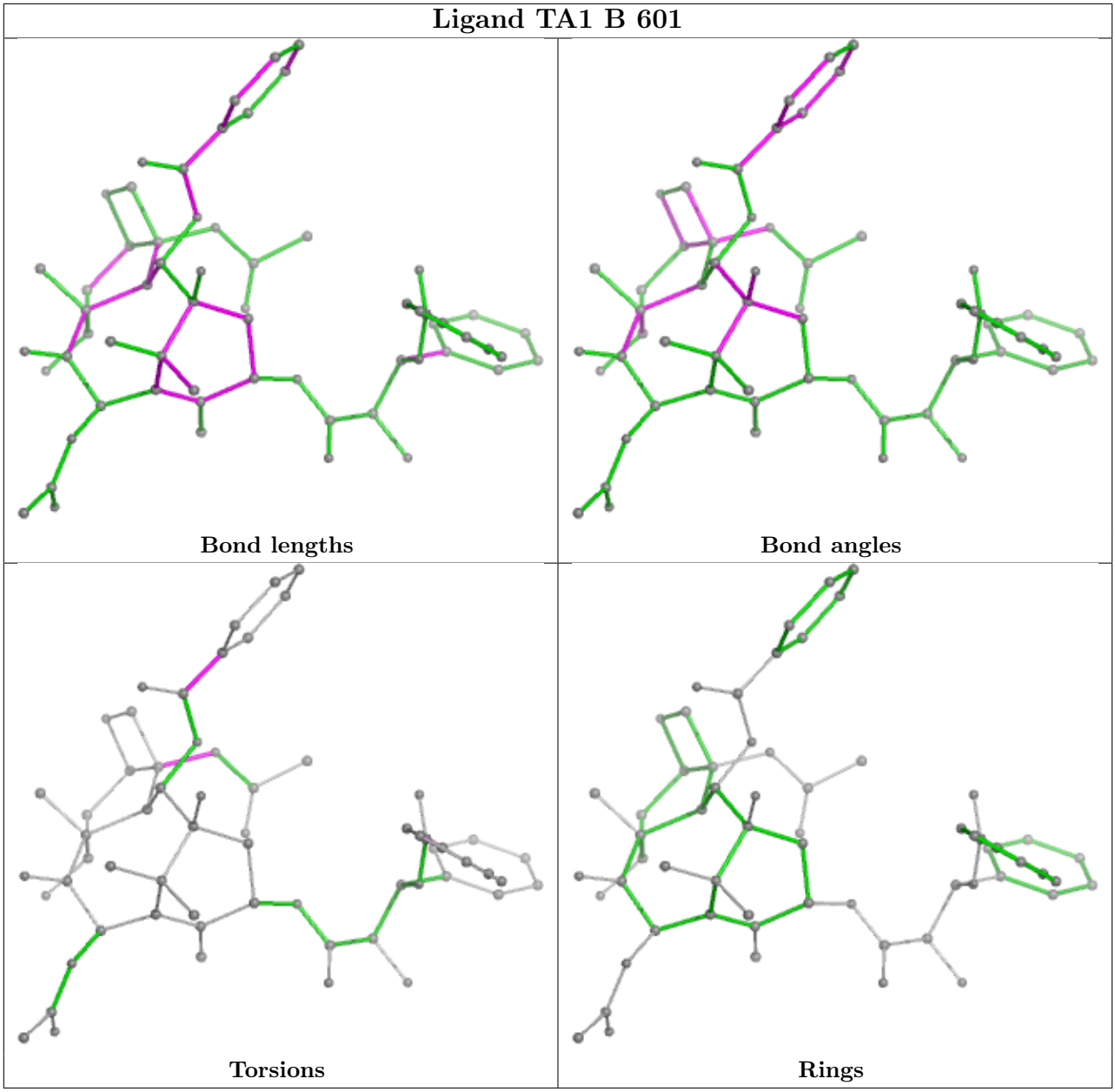
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 GTP A 500



Ligand GDP B 600





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.

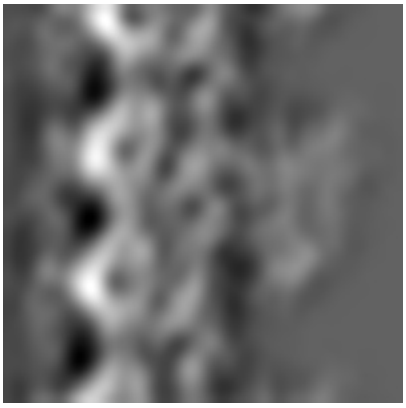
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-2078. These allow visual inspection of the internal detail of the map and identification of artifacts.

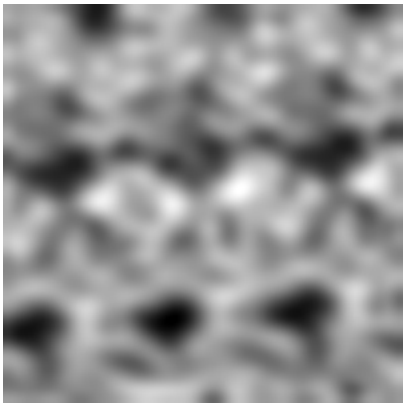
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

6.1.1 Primary map



X



Y

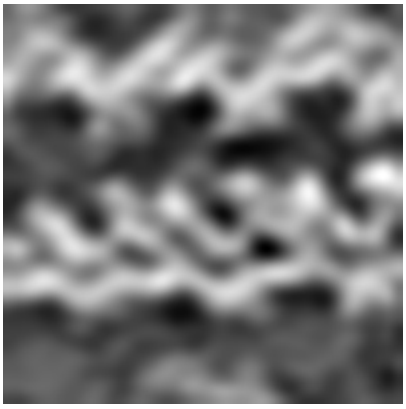
Z

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

6.2 Central slices (i)

6.2.1 Primary map

X Index: 22



Y Index: 22

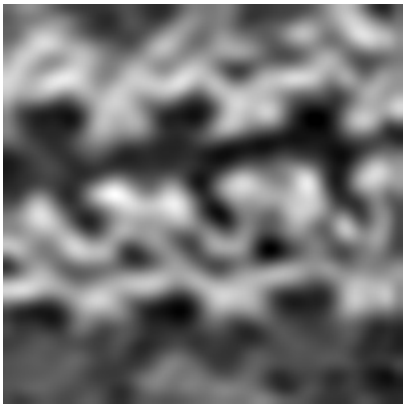
Z Index: 22

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: 8



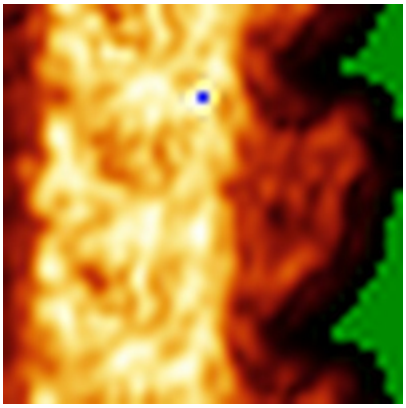
Y Index: 21

Z Index: 18

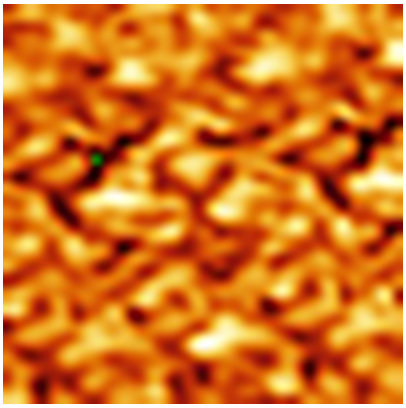
The images above show the largest variance slices of the map in three orthogonal directions.

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

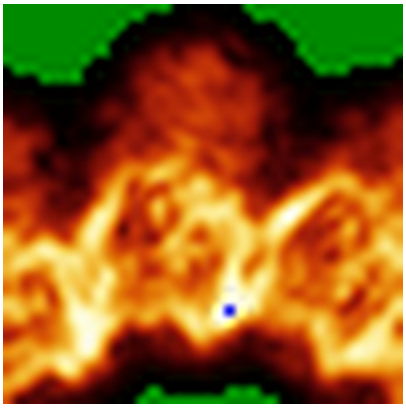
6.4.1 Primary map



X



Y

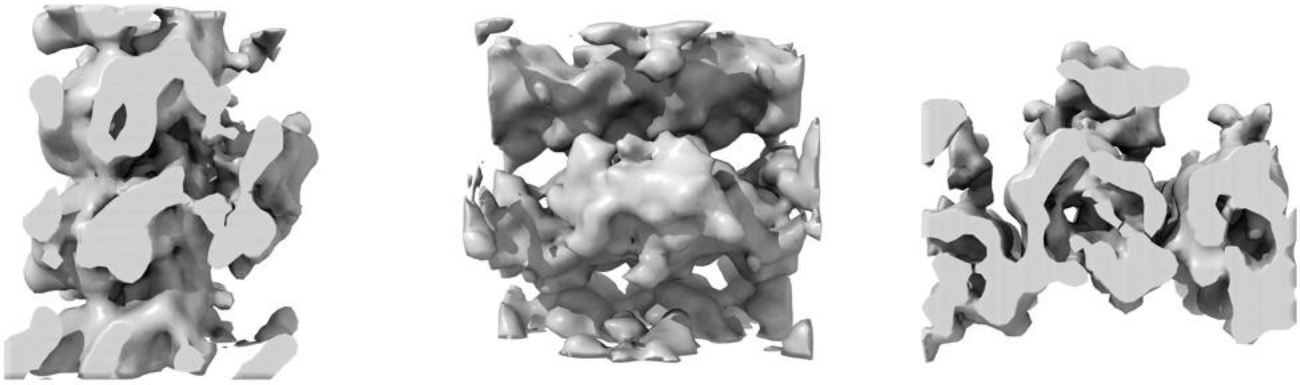


Z

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

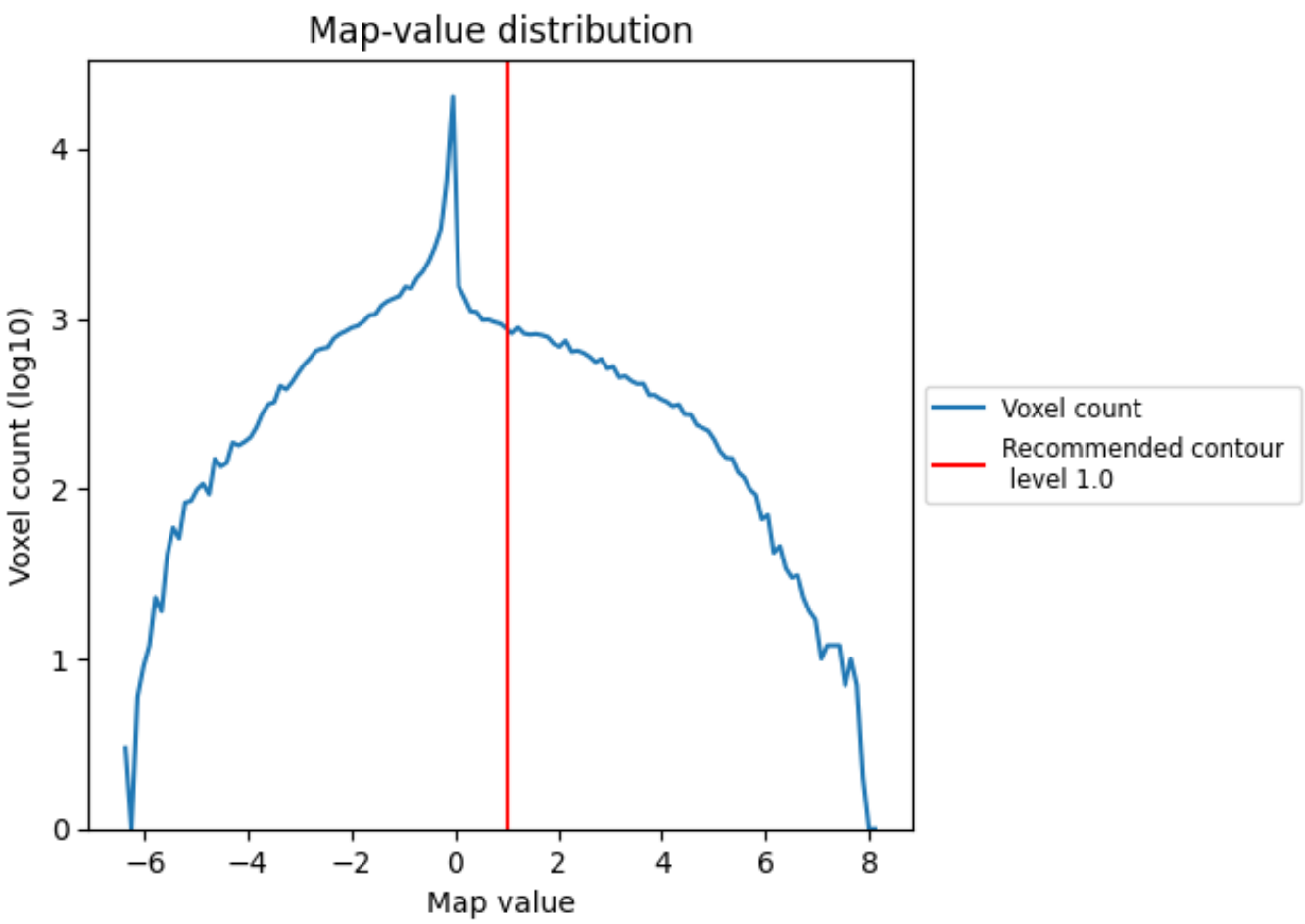
6.6 Mask visualisation (i)

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

7 Map analysis (i)

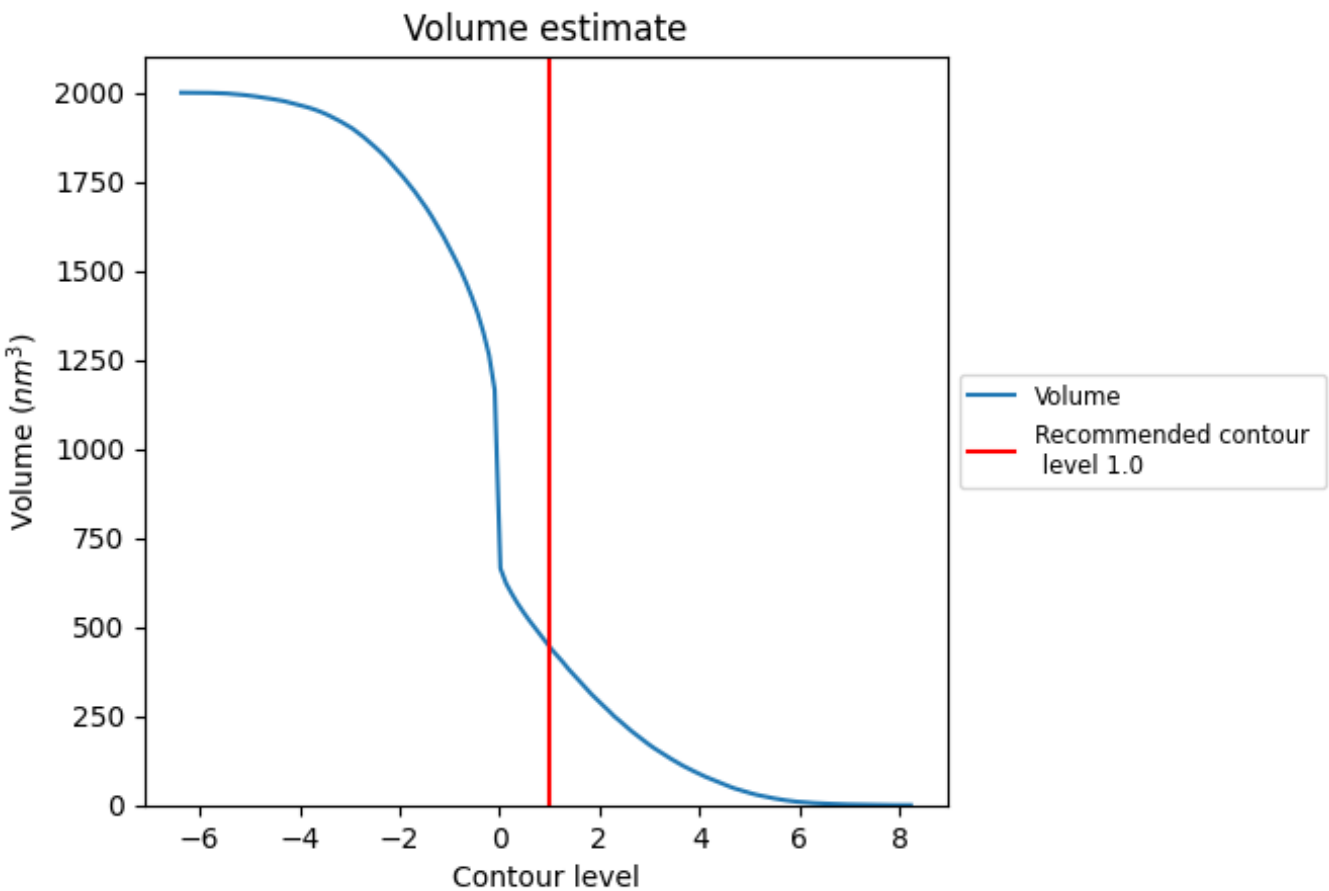
This section contains the results of statistical analysis of the map.

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.

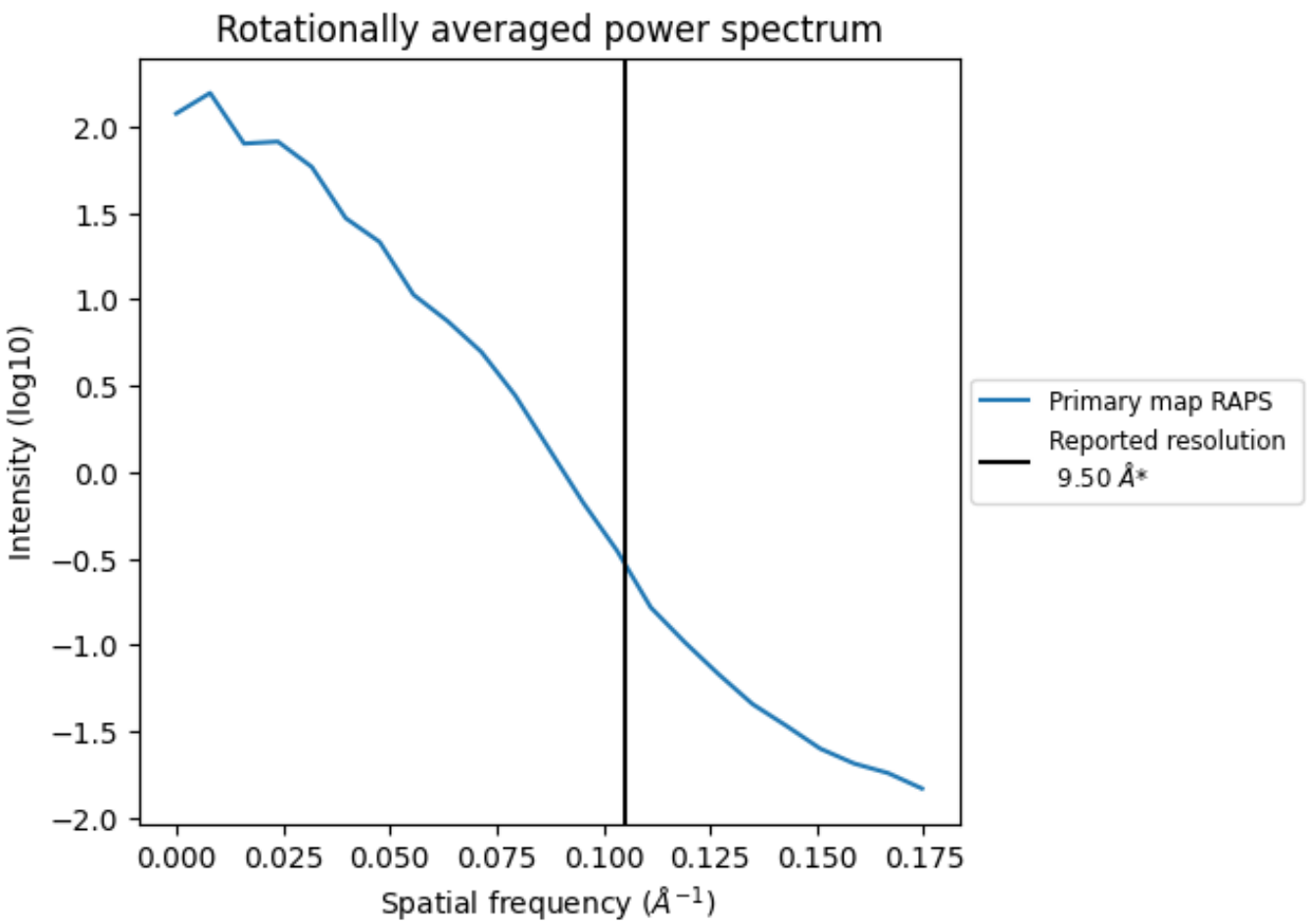
7.2 Volume estimate (i)



The volume at the recommended contour level is 445 nm^3; this corresponds to an approximate mass of 402 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 0.105 Å^{-1}

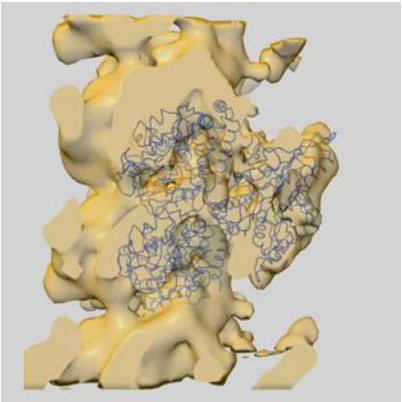
8 Fourier-Shell correlation (i)

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

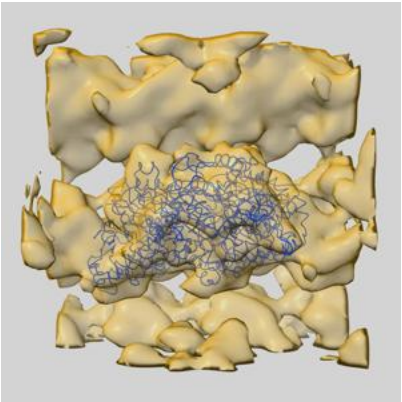
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-2078 and PDB model 4AQW. Per-residue inclusion information can be found in section 3 on page 6.

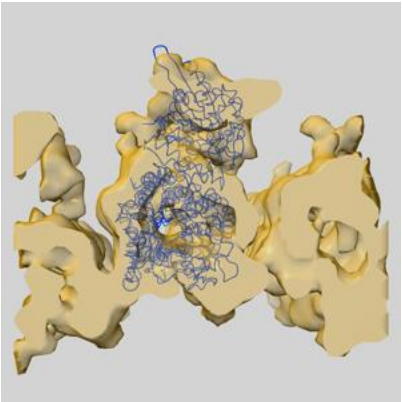
9.1 Map-model overlay (i)



X



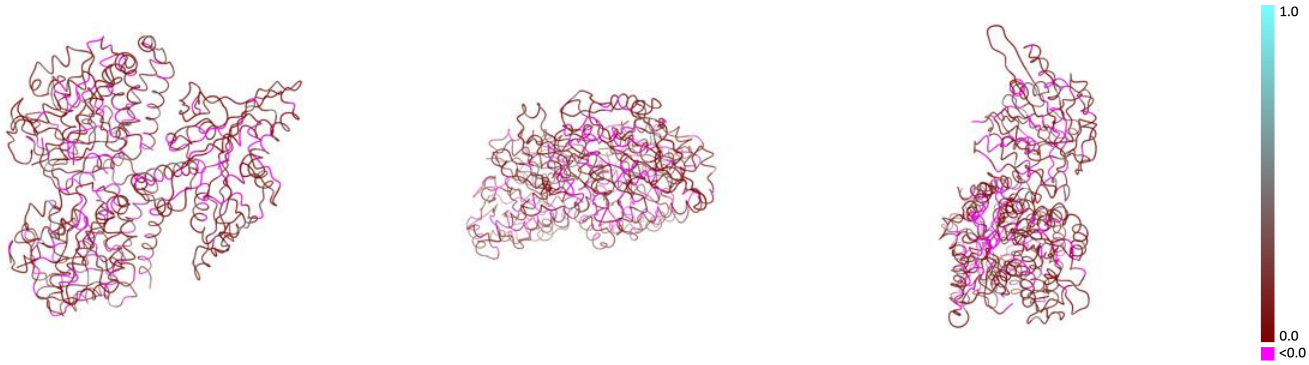
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 1.0 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.

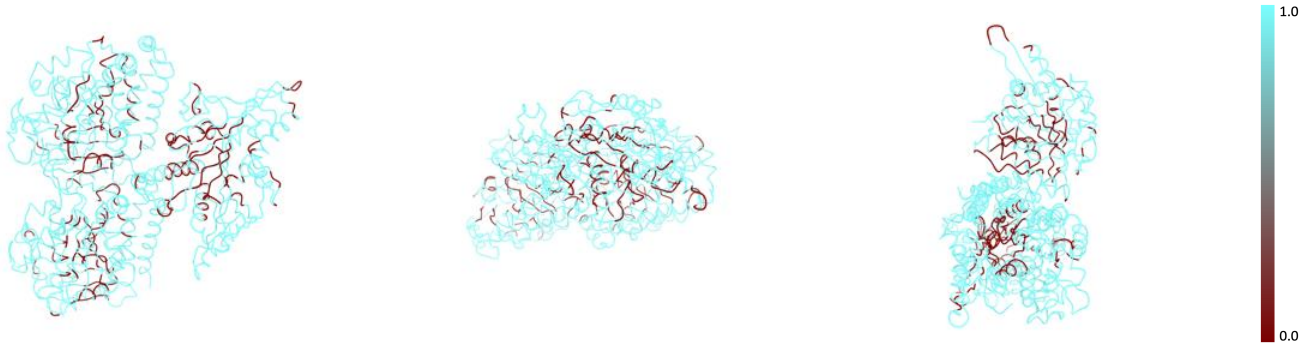
9.2 Q-score mapped to coordinate model (i)



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.

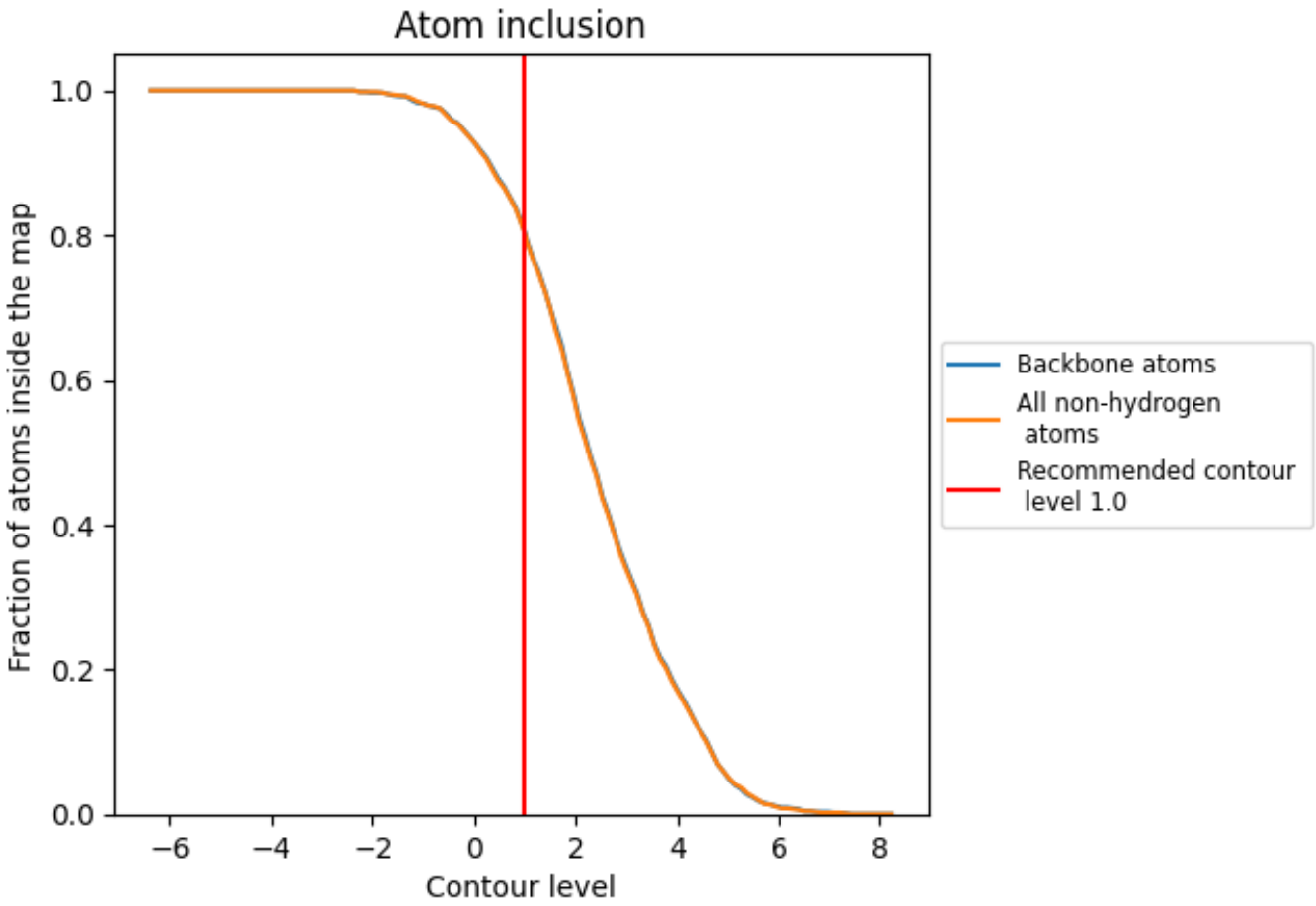
9.3 Atom inclusion mapped to coordinate model (i)



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 (1.0).

9.4 Atom inclusion (i)



At the recommended contour level, 80% 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 (1.0) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8020	0.1030
A	0.8270	0.1070
B	0.8260	0.1040
C	0.7340	0.0970