wwPDB EM Validation Summary Report (i)

Dec 17, 2022 – 10:18 pm GMT

PDB ID	:	6ZPI
EMDB ID	:	EMD-11340
Title	:	Microtubule complexed with Kif15 motor domain. Symmetrised asymmetric unit
Authors	:	Atherton, J.; Hummel, J.J.A.; Olieric, N.; Locke, J.; Pena, A.; Rosenfeld, S.S.; Steinmetz, M.O.; Hoogenraad, C.C.; Moores, C.A.
Deposited on	:	2020-07-08
Resolution	:	4.50 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:

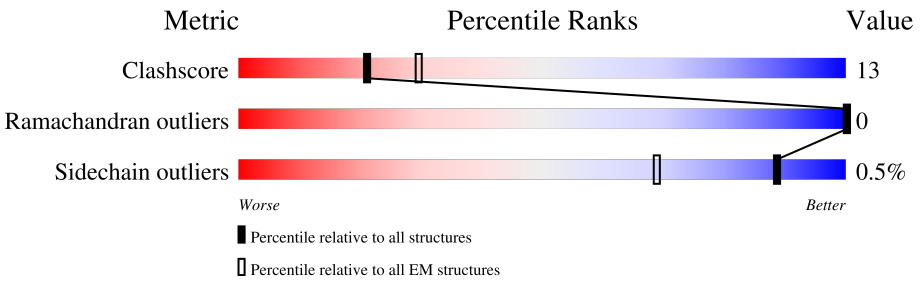
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 4.50 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
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	C	377	42% 58% 29% 12%
2	A	437	85% 12% •
3	B	431	81% 17% •

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 9412 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 Kinesin-like protein KIF15.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	C	332	2563	1601	457	497	8	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	5	SER	CYS	conflict	UNP Q9NS87
C	50	SER	CYS	conflict	UNP Q9NS87
C	162	SER	CYS	conflict	UNP Q9NS87
C	250	CYS	SER	conflict	UNP Q9NS87
C	294	THR	CYS	conflict	UNP Q9NS87
C	314	SER	CYS	conflict	UNP Q9NS87
C	346	SER	CYS	conflict	UNP Q9NS87
C	375	CYS	-	expression tag	UNP Q9NS87
C	376	LEU	-	expression tag	UNP Q9NS87
C	377	GLU	-	expression tag	UNP Q9NS87

- Molecule 2 is a protein called Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	426	3343	2121	569	631	22	0	0

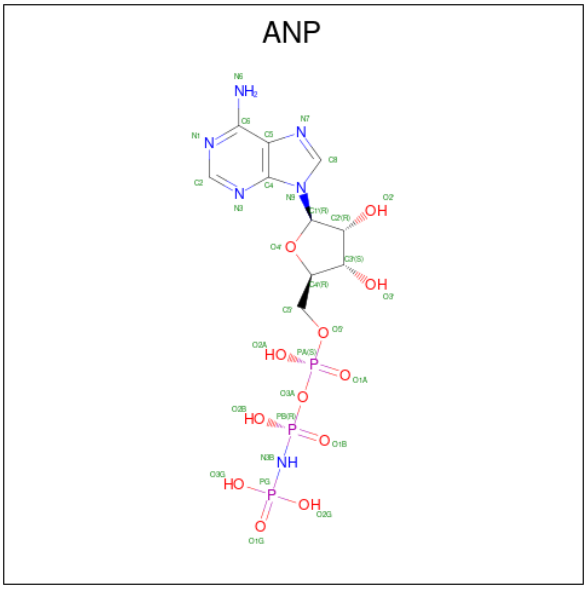
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	340	THR	SER	conflict	UNP Q2XVP4

- Molecule 3 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	B	426	3351	2106	575	644	26	0	0

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).

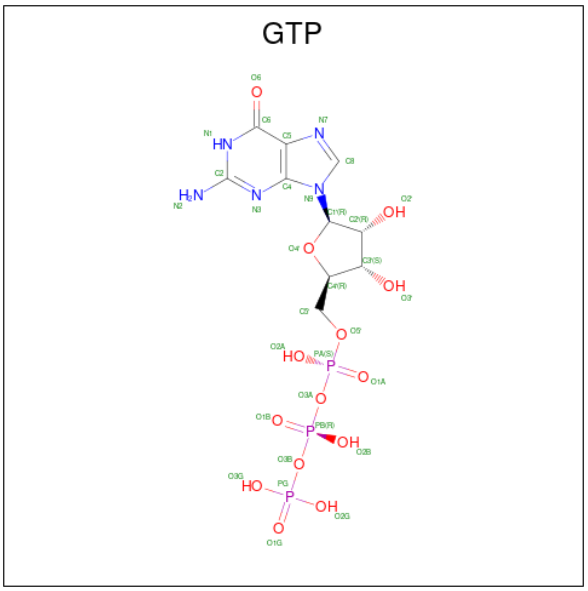


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
4	C	1	31	10	6	12	3	0

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

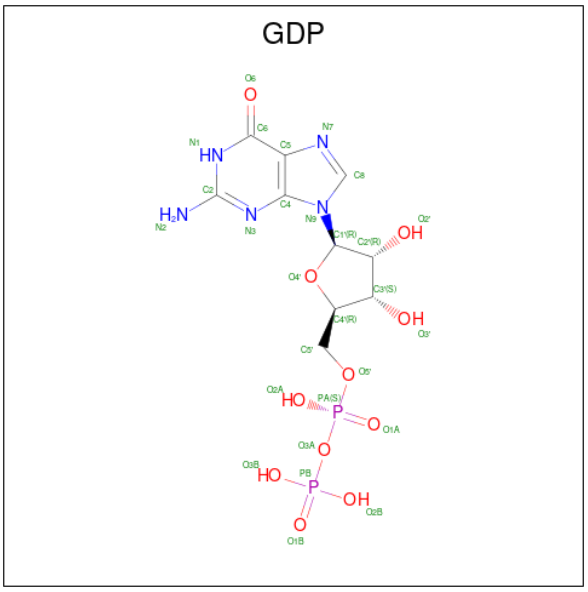
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
5	C	1	1	1	0
5	A	1	Total 1	Mg 1	0

- Molecule 6 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



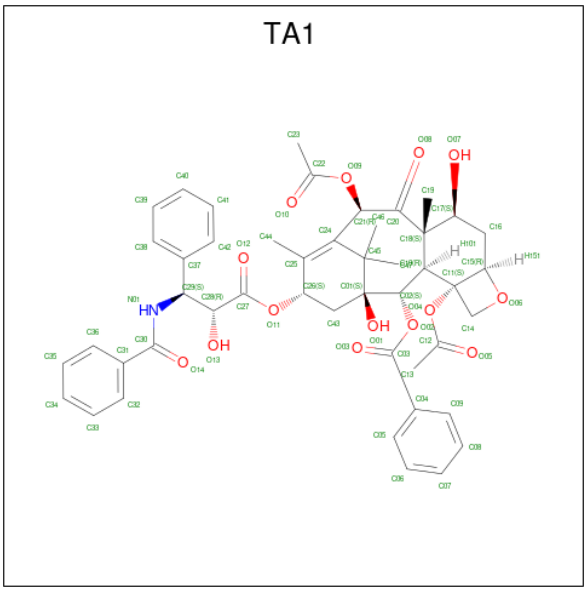
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
6	A	1	32	10	5	14	3	0

- Molecule 7 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	
7	B	1	28	10	5	11	2	0

- Molecule 8 is TAXOL (three-letter code: TA1) (formula: $C_{47}H_{51}NO_{14}$).



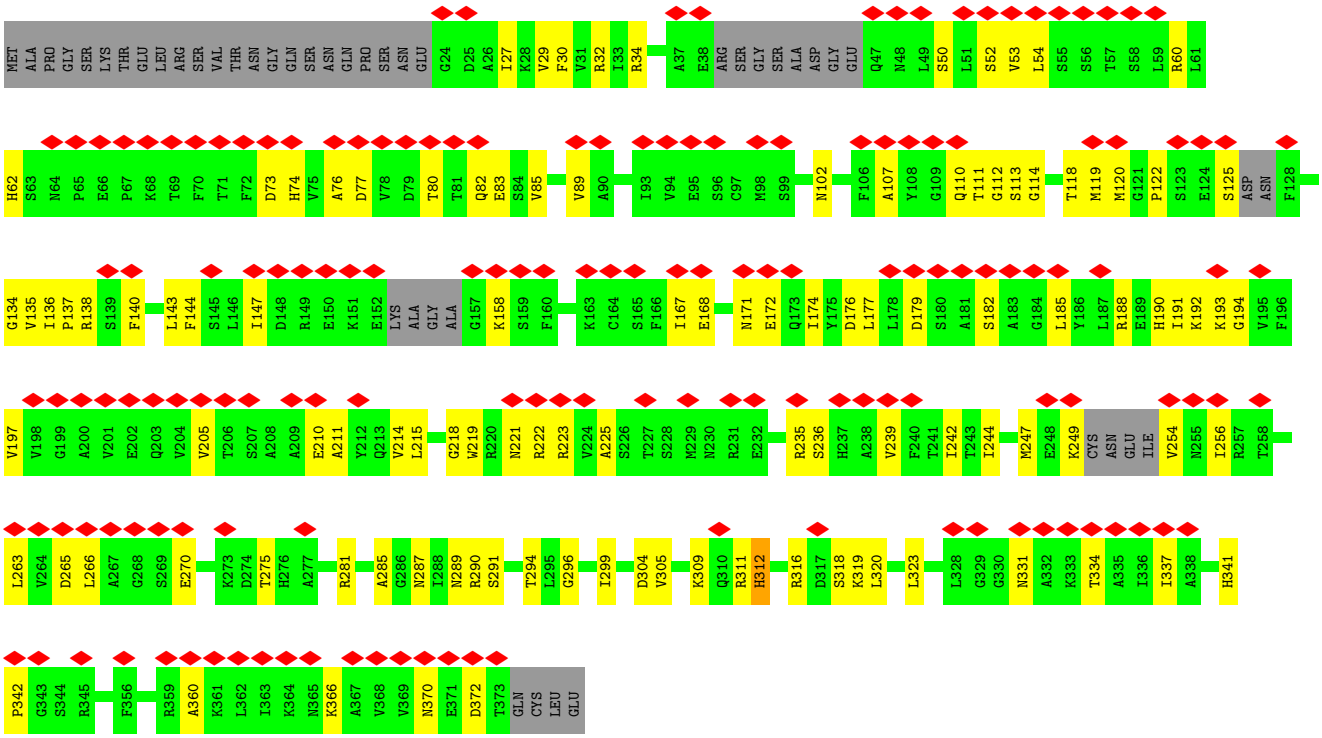
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
8	B	1	62	47	1	14	0

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.

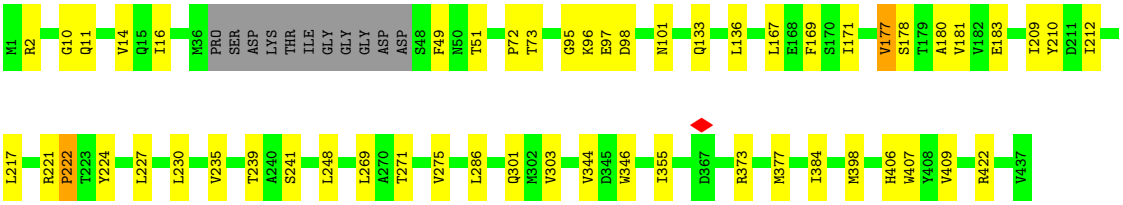
- Molecule 1: Kinesin-like protein KIF15

Chain C: 42% 58% 29% 12%



- Molecule 2: Tubulin alpha-1B chain

Chain A: 85% 12%



- Molecule 3: Tubulin beta chain

Chain B: 81% 17%

M1 G13 F20 I24 I30 D31 P32 R48 I49 L67 V68 D69 T74 V78 R79 I86 F87 F92 V93 F94 D116 T138 T145 M149 G150 T151 F169 Q193 I204 D205 N206 E207 A208 L209 D226 H229 L230 V231 T234 M235
S236 G237 R243 Q247 L248 N249 A250 D251 L252 R253 K254 L255 A256 V257 N258 M259 V260 P263 R264 L265 P270 V288 Q294 A304 C305 D306 P307 R308 H309 T314 Y315 A316 S324 M325 K326 E327 V328 L333 W346 I347 P348 N349 K352 T376 N416
E420 Y432 Q433 Q434 Y435 Q436 ASP ALA THR ALA ASP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	12674	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 ($e^-/Å^2$)	42	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.149	Depositor
Minimum map value	-0.051	Depositor
Average map value	0.005	Depositor
Map value standard deviation	0.014	Depositor
Recommended contour level	0.027	Depositor
Map size (Å)	95.909996, 137.61, 140.39	wwPDB
Map dimensions	69, 99, 101	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.39, 1.39, 1.39	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: GDP, GTP, ANP, 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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >5$	RMSZ	$\#\|Z\| >5$
1	C	0.31	0/2599	0.54	1/3499 (0.0%)
2	A	0.39	2/3419 (0.1%)	0.57	5/4639 (0.1%)
3	B	0.38	1/3426 (0.0%)	0.60	3/4641 (0.1%)
All	All	0.37	3/9444 (0.0%)	0.57	9/12779 (0.1%)

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
3	B	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	133	GLN	CB-CG	-6.30	1.35	1.52
2	A	222	PRO	CG-CD	-5.43	1.32	1.50
3	B	326	LYS	CG-CD	-5.03	1.35	1.52

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	294	GLN	CA-CB-CG	10.50	136.51	113.40
3	B	294	GLN	CB-CG-CD	7.91	132.16	111.60
2	A	301	GLN	CA-CB-CG	7.07	128.95	113.40
2	A	301	GLN	N-CA-CB	-6.28	99.30	110.60
1	C	312	HIS	N-CA-CB	6.12	121.62	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	309	HIS	Sidechain

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	C	2563	0	2563	98	0
2	A	3343	0	3263	111	0
3	B	3351	0	3237	145	0
4	C	31	0	13	3	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
6	A	32	0	12	0	0
7	B	28	0	12	0	0
8	B	62	0	51	5	0
All	All	9412	0	9151	245	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 13.

The worst 5 of 245 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)
2:A:97:GLU:OE2	3:B:253:ARG:NH2	1.67	1.25
2:A:96:LYS:HA	3:B:1:MET:SD	1.80	1.21
2:A:177:VAL:HG22	3:B:333:LEU:HD11	1.23	1.21
2:A:407:TRP:NE1	3:B:257:VAL:HA	1.58	1.19
2:A:222:PRO:HG2	3:B:326:LYS:CG	1.87	1.04

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	C	322/377 (85%)	315 (98%)	7 (2%)	0	100	100
2	A	422/437 (97%)	388 (92%)	34 (8%)	0	100	100
3	B	424/431 (98%)	411 (97%)	13 (3%)	0	100	100
All	All	1168/1245 (94%)	1114 (95%)	54 (5%)	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	C	284/320 (89%)	283 (100%)	1 (0%)	91	94
2	A	359/368 (98%)	357 (99%)	2 (1%)	86	92
3	B	367/370 (99%)	365 (100%)	2 (0%)	88	93
All	All	1010/1058 (96%)	1005 (100%)	5 (0%)	89	93

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

Mol	Chain	Res	Type
1	C	223	ARG
2	A	49	PHE
2	A	422	ARG
3	B	116	ASP
3	B	309	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	249	ASN
3	B	349	ASN
3	B	434	GLN
1	C	370	ASN
1	C	221	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

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	$\#\|Z\| > 2$
4	ANP	C	601	5	29,33,33	1.35	3 (10%)	31,52,52	2.04	10 (32%)
7	GDP	B	501	-	24,30,30	1.00	2 (8%)	30,47,47	1.16	3 (10%)
6	GTP	A	501	5	26,34,34	0.69	0	32,54,54	0.73	1 (3%)
8	TA1	B	502	3	68,68,68	0.58	0	105,105,105	0.97	9 (8%)

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	ANP	C	601	5	-	4/14/38/38	0/3/3/3
7	GDP	B	501	-	-	3/12/32/32	0/3/3/3
6	GTP	A	501	5	-	5/18/38/38	0/3/3/3
8	TA1	B	502	3	-	8/41/127/127	0/7/7/7

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	601	ANP	PB-O3A	-4.49	1.53	1.59
4	C	601	ANP	O5'-C5'	-3.03	1.33	1.44
7	B	501	GDP	C6-N1	-2.47	1.34	1.37
4	C	601	ANP	PG-O1G	2.37	1.49	1.46
7	B	501	GDP	O4'-C1'	2.02	1.43	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	601	ANP	C5'-C4'-C3'	-4.51	98.29	115.18
4	C	601	ANP	O5'-C5'-C4'	4.24	123.57	108.99
4	C	601	ANP	PA-O5'-C5'	-4.23	96.85	121.68
4	C	601	ANP	O3A-PB-N3B	-4.02	95.45	106.59
7	B	501	GDP	PA-O3A-PB	-3.25	121.67	132.83

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	601	ANP	PB-N3B-PG-O1G
4	C	601	ANP	PG-N3B-PB-O1B
4	C	601	ANP	PG-N3B-PB-O3A
6	A	501	GTP	PB-O3B-PG-O3G
6	A	501	GTP	C5'-O5'-PA-O1A

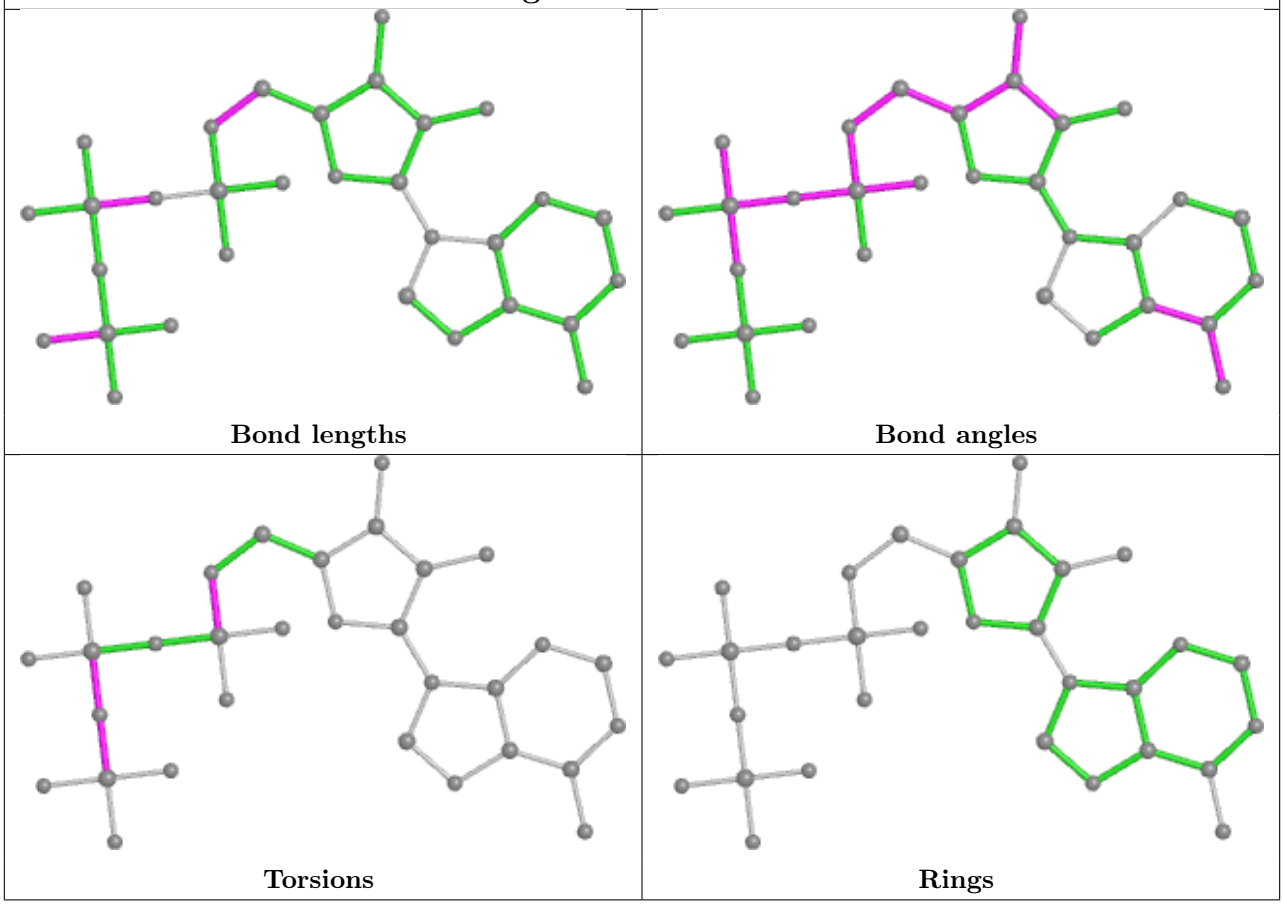
There are no ring outliers.

2 monomers are involved in 8 short contacts:

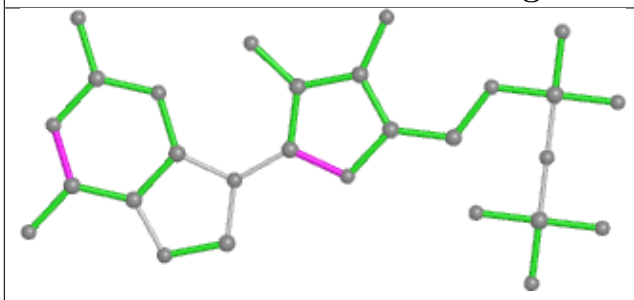
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	601	ANP	3	0
8	B	502	TA1	5	0

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.

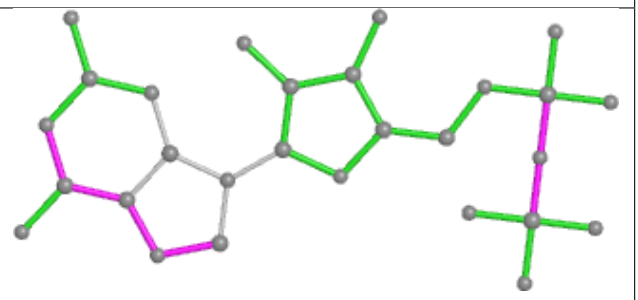
Ligand ANP C 601



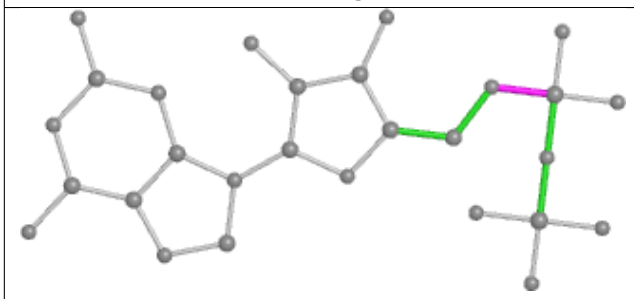
Ligand GDP B 501



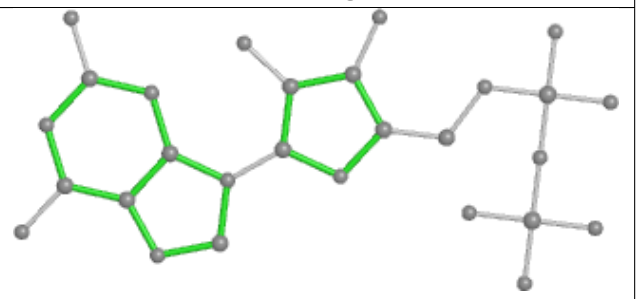
Bond lengths



Bond angles

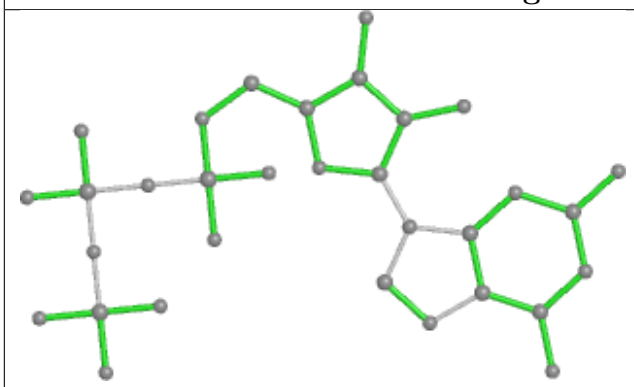


Torsions

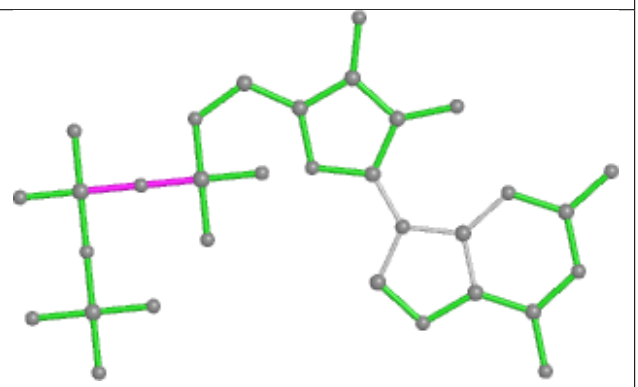


Rings

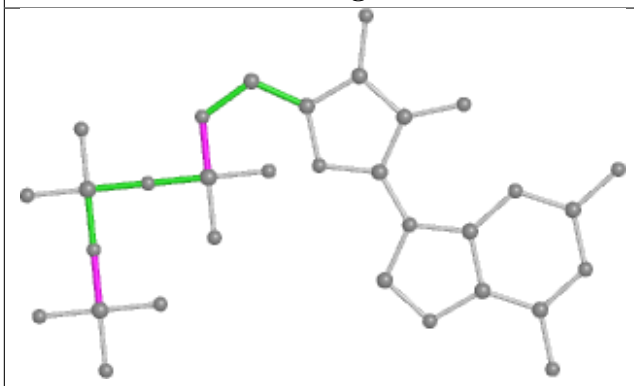
Ligand GTP A 501



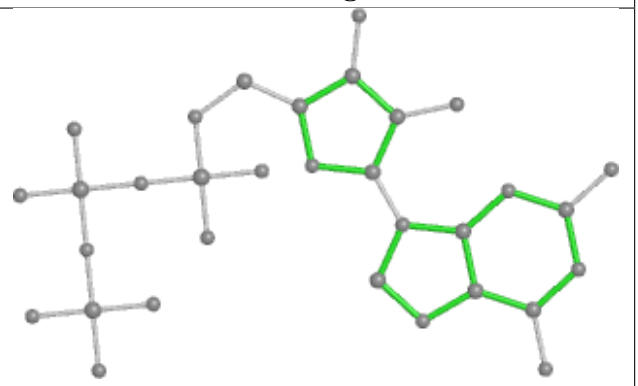
Bond lengths



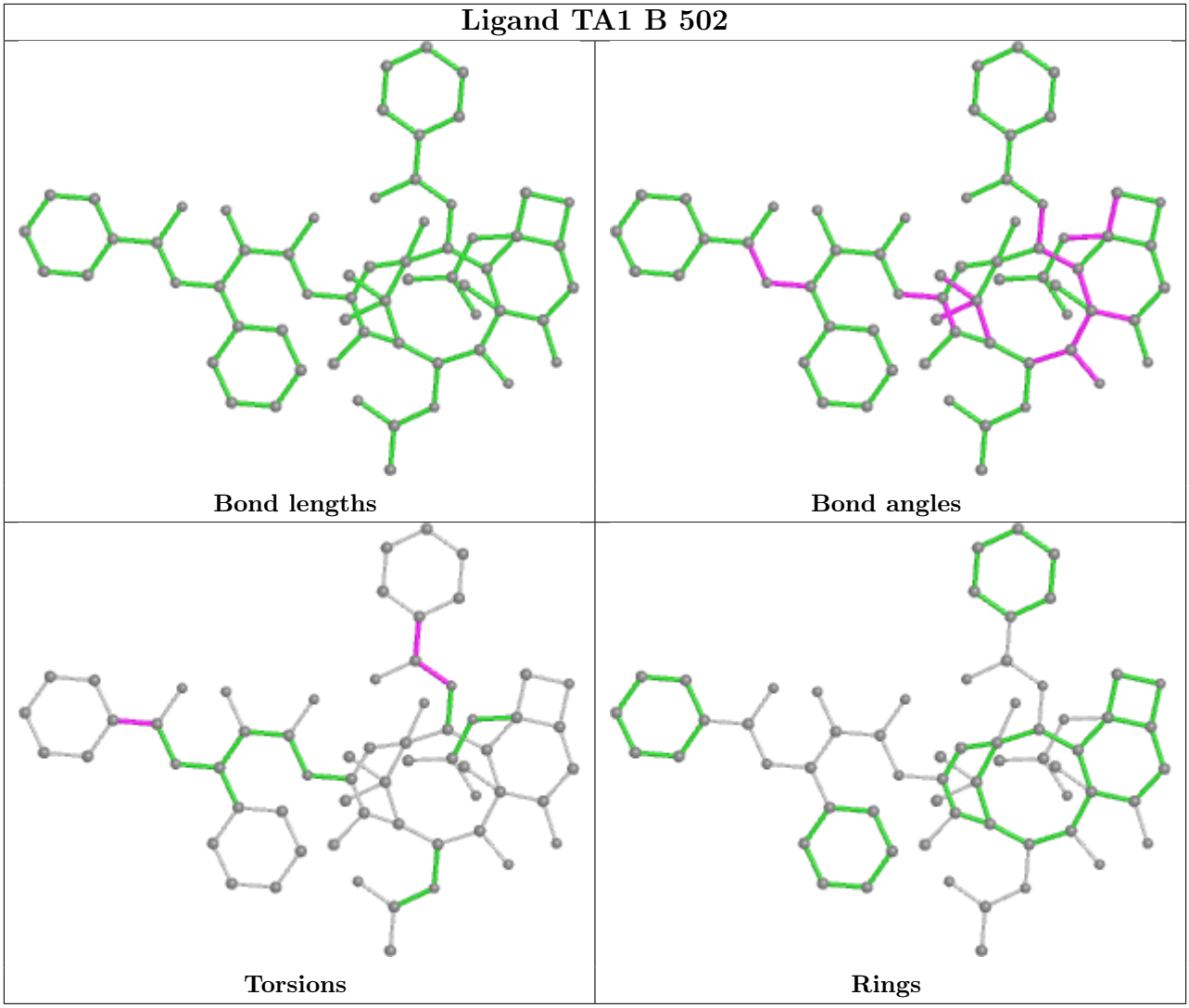
Bond angles



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.

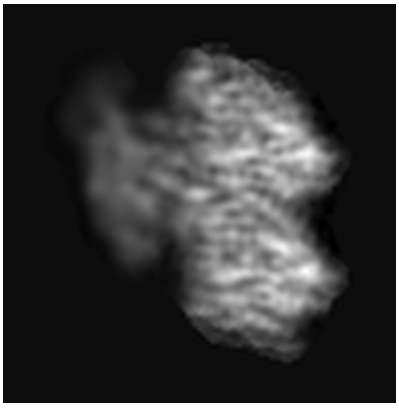
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-11340. These allow visual inspection of the internal detail of the map and identification of artifacts.

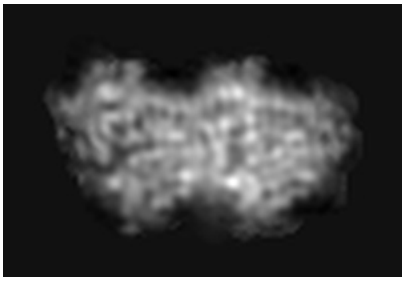
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

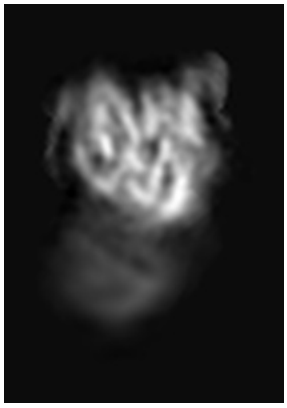
6.1.1 Primary map



X



Y

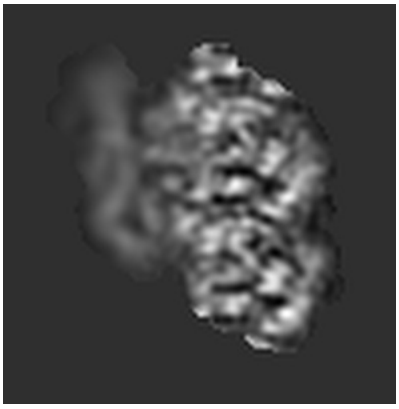


Z

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

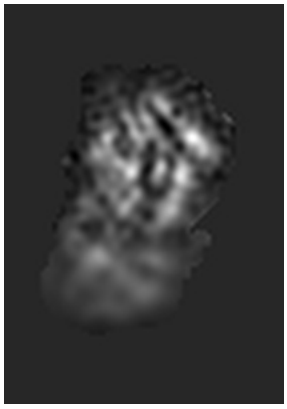
6.2 Central slices (i)

6.2.1 Primary map



X Index: 34

Y Index: 49

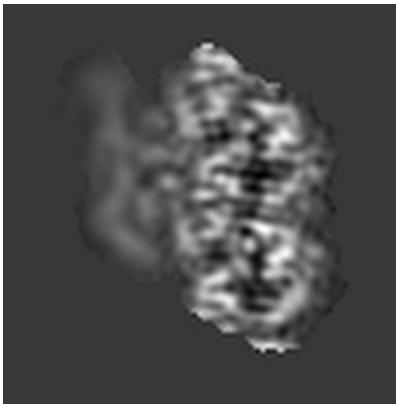


Z Index: 50

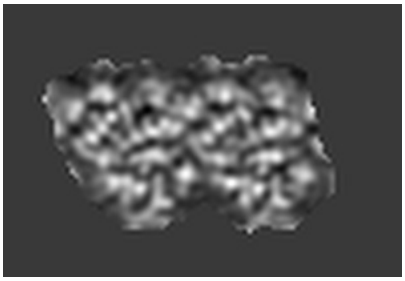
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

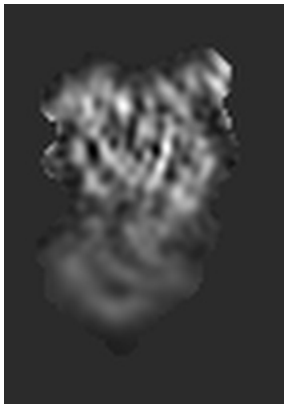
6.3.1 Primary map



X Index: 36



Y Index: 70

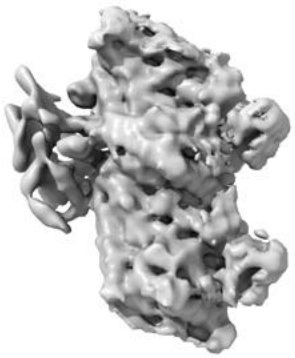


Z Index: 62

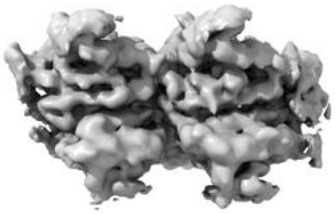
The images above show the largest variance slices of the map in three orthogonal directions.

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

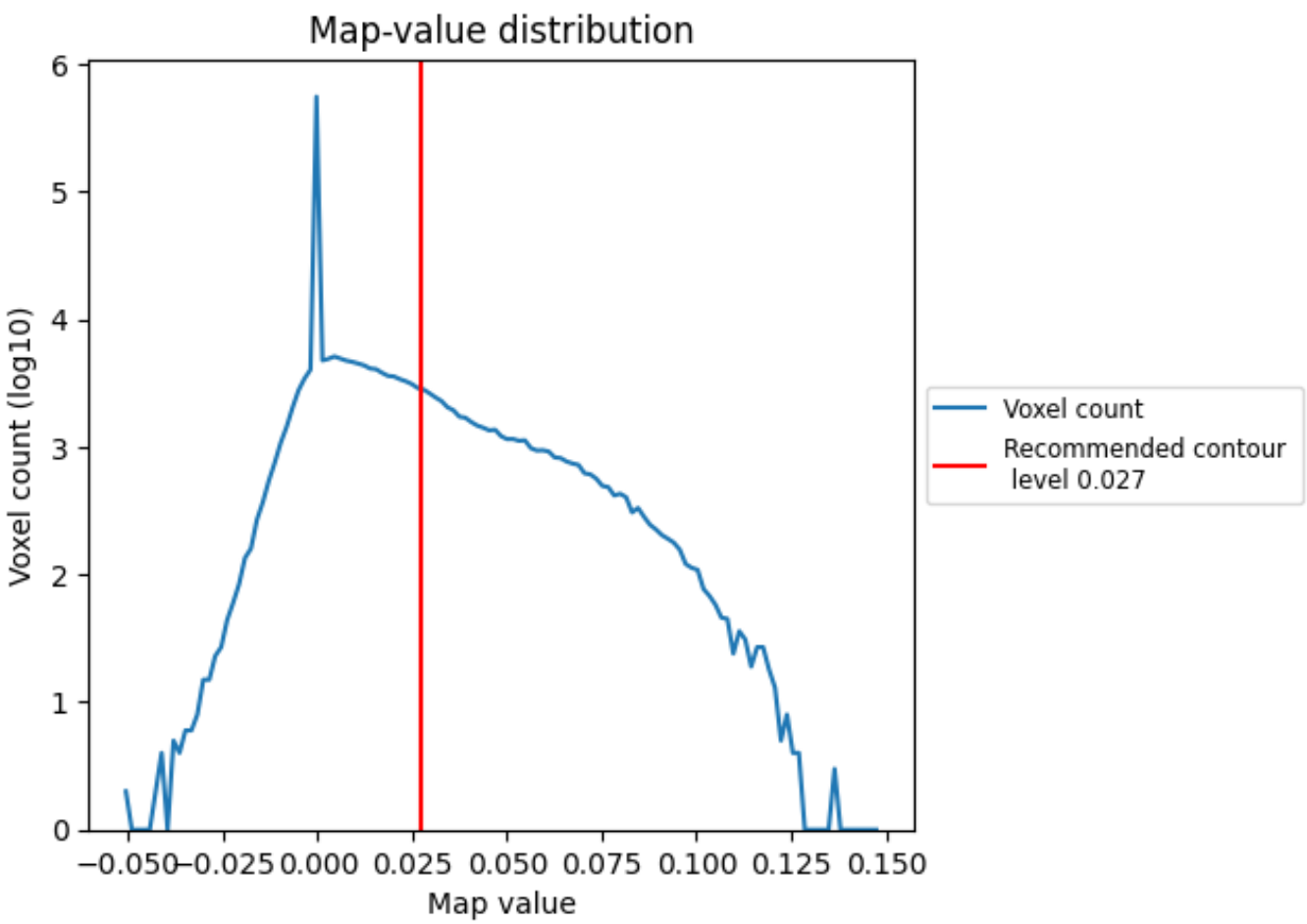
6.5 Mask visualisation (i)

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

7 Map analysis (i)

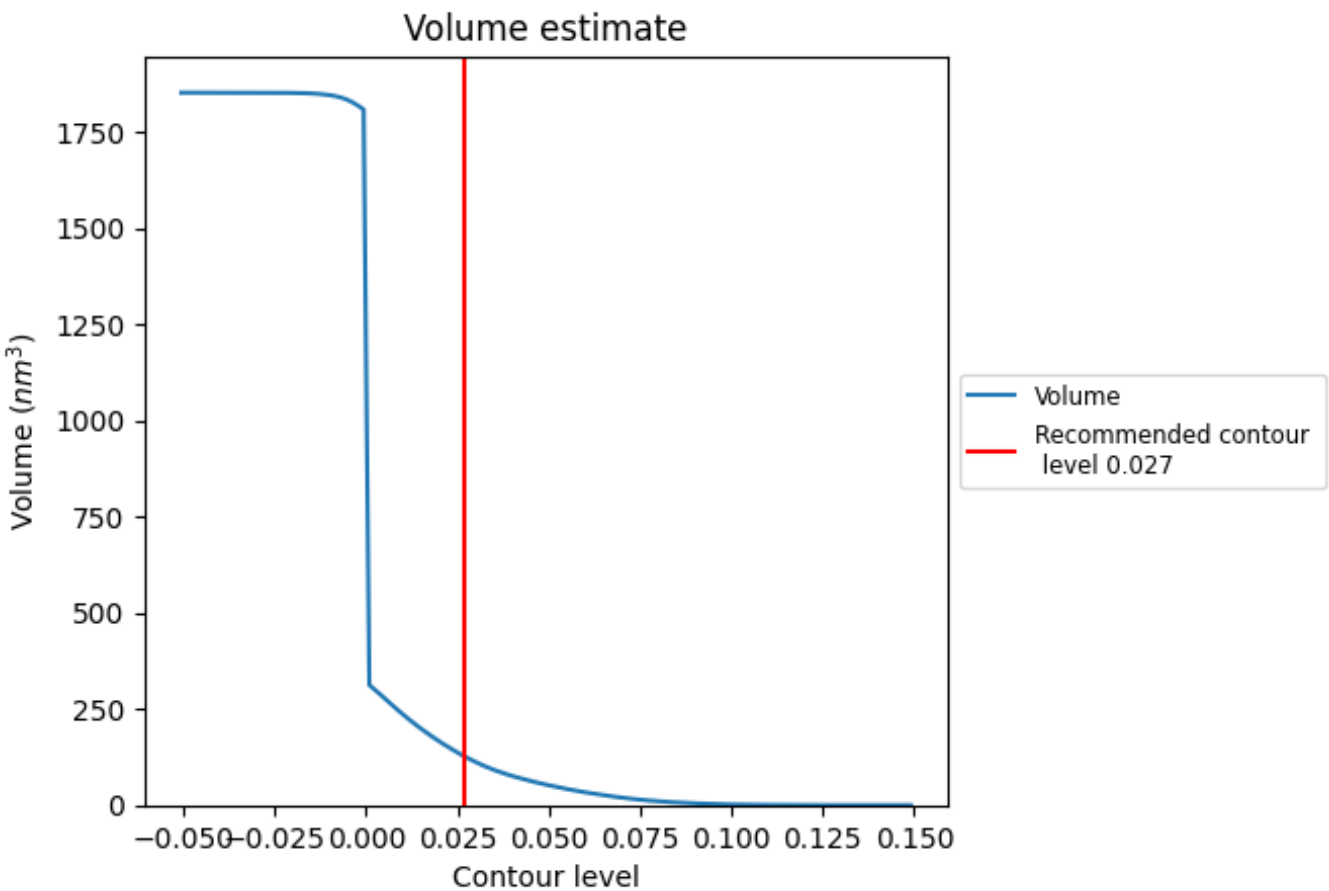
This section contains the results of statistical analysis of the map.

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 126 nm^3; this corresponds to an approximate mass of 114 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)

This section was not generated. The rotationally averaged power spectrum is only generated for cubic maps.

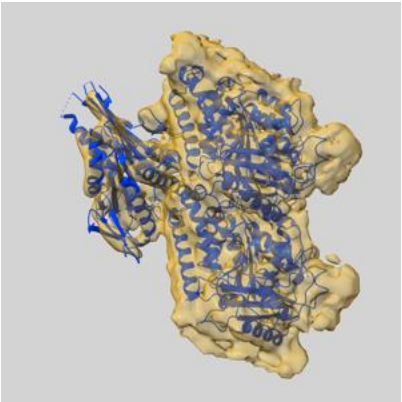
8 Fourier-Shell correlation (i)

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

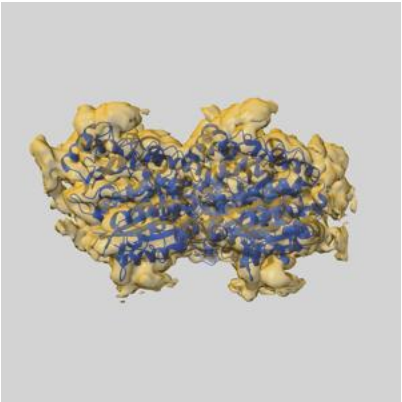
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-11340 and PDB model 6ZPI. Per-residue inclusion information can be found in section 3 on page 7.

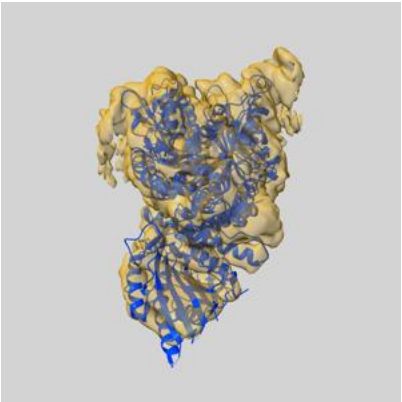
9.1 Map-model overlay (i)



X



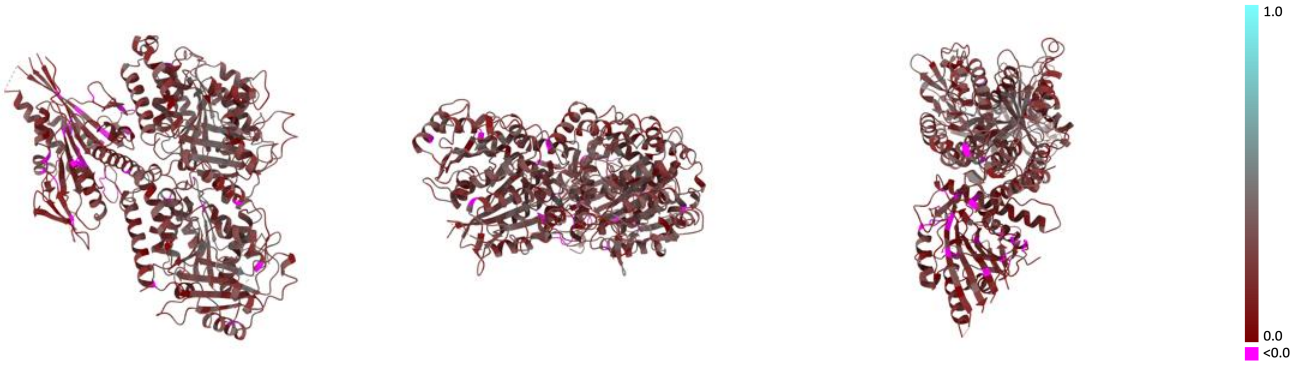
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.027 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.

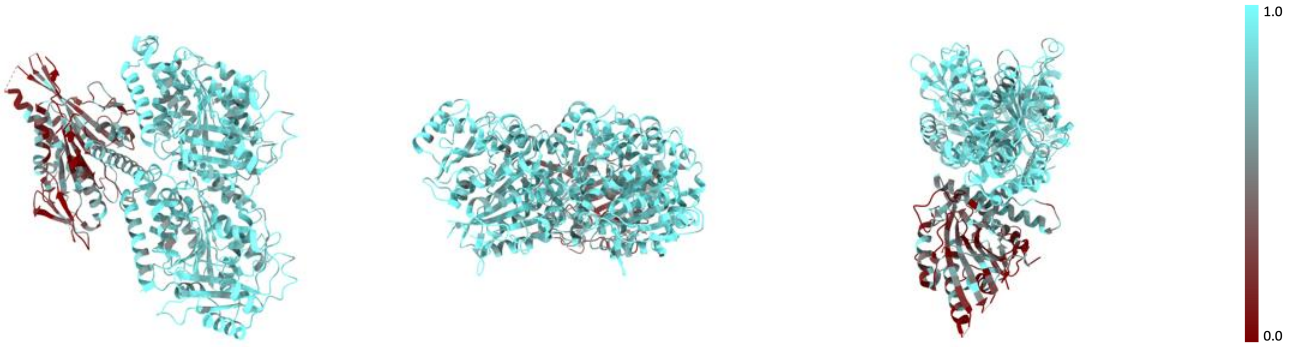
9.2 Q-score mapped to coordinate model (i)



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.

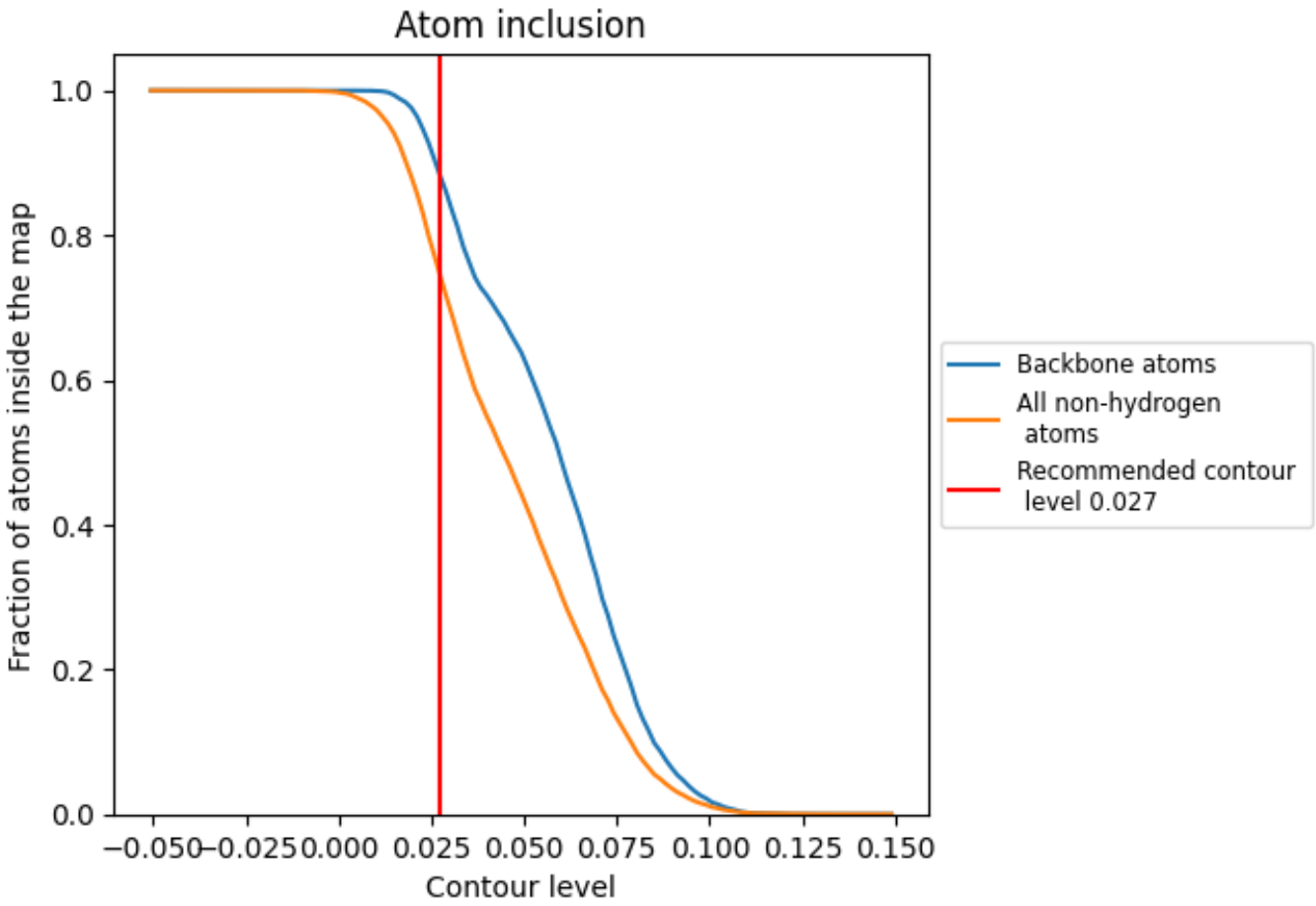
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.027).

9.4 Atom inclusion (i)



At the recommended contour level, 89% of all backbone atoms, 75% 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 (0.027) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7499	0.2380
A	0.8750	0.2680
B	0.8833	0.2710
C	0.4100	0.1550