wwPDB EM Validation Summary Report (i)

Mar 10, 2024 – 01:49 PM EDT

PDB ID	:	6PPJ
EMDB ID	:	EMD-20440
Title	:	Cryo-EM structure of AdnA(D934A)-AdnB(D1014A) in complex with AMPPNP
Authors	:	Jia, N.; Unciuleac, M.; Shuman, S.; Patel, D.J.
Deposited on	:	2019-07-07
Resolution	:	3.50 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:

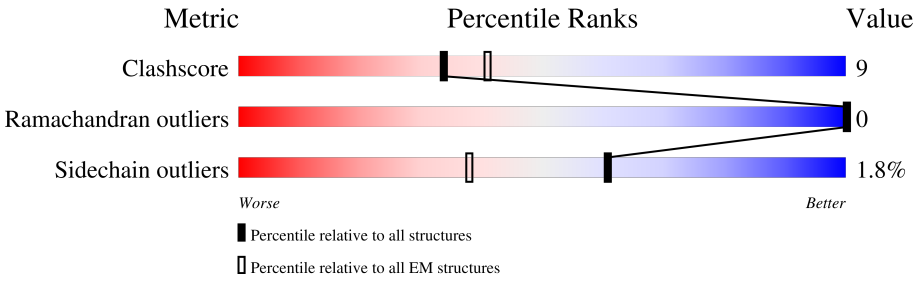
EMDB validation analysis	:	0.0.1.dev70
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.50 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	B	1095	68% 9% 23%
2	A	1045	75% 11% 14%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12004 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 UvrD/REP helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	B	848	5756	3668	1063	1011	14	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	537	THR	SER	conflict	UNP A0A0D6HIW1
B	540	VAL	LEU	conflict	UNP A0A0D6HIW1
B	559	GLU	GLN	conflict	UNP A0A0D6HIW1
B	599	ALA	VAL	conflict	UNP A0A0D6HIW1
B	627	GLY	SER	conflict	UNP A0A0D6HIW1
B	1014	ALA	ASP	engineered mutation	UNP A0A0D6HIW1

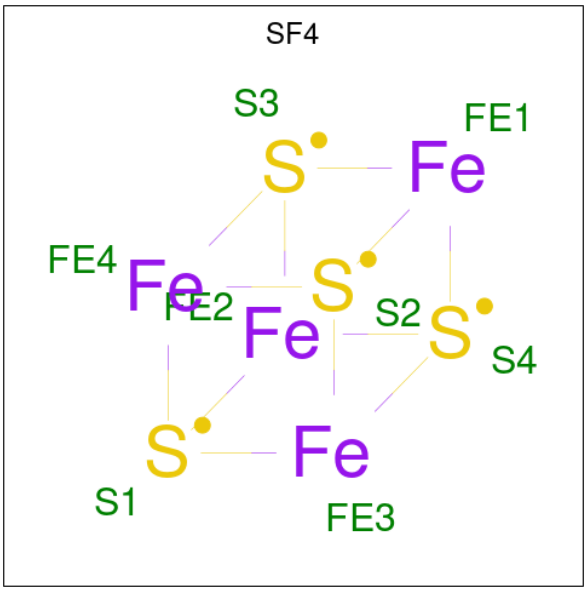
- Molecule 2 is a protein called ATP-dependent DNA helicase (UvrD/REP).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	901	6209	3943	1161	1086	19	0	0

There is a discrepancy between the modelled and reference sequences:

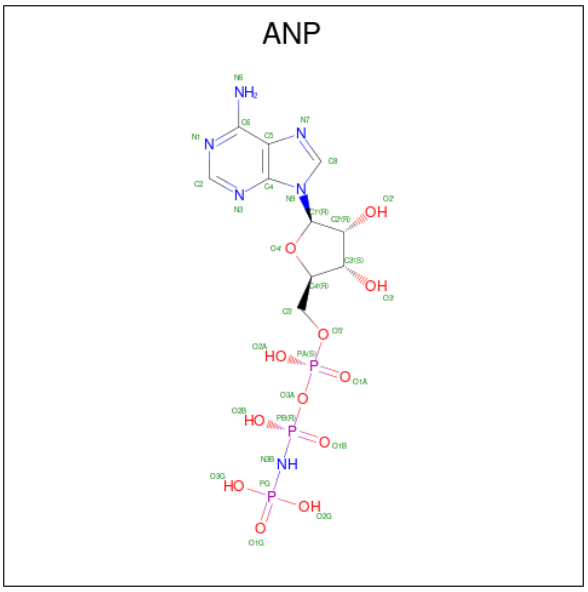
Chain	Residue	Modelled	Actual	Comment	Reference
A	934	ALA	ASP	engineered mutation	UNP A0A0D6HKQ2

- Molecule 3 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
3	A	1	8	4	4	0

- Molecule 4 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$) (labeled as "Ligand of Interest" by depositor).



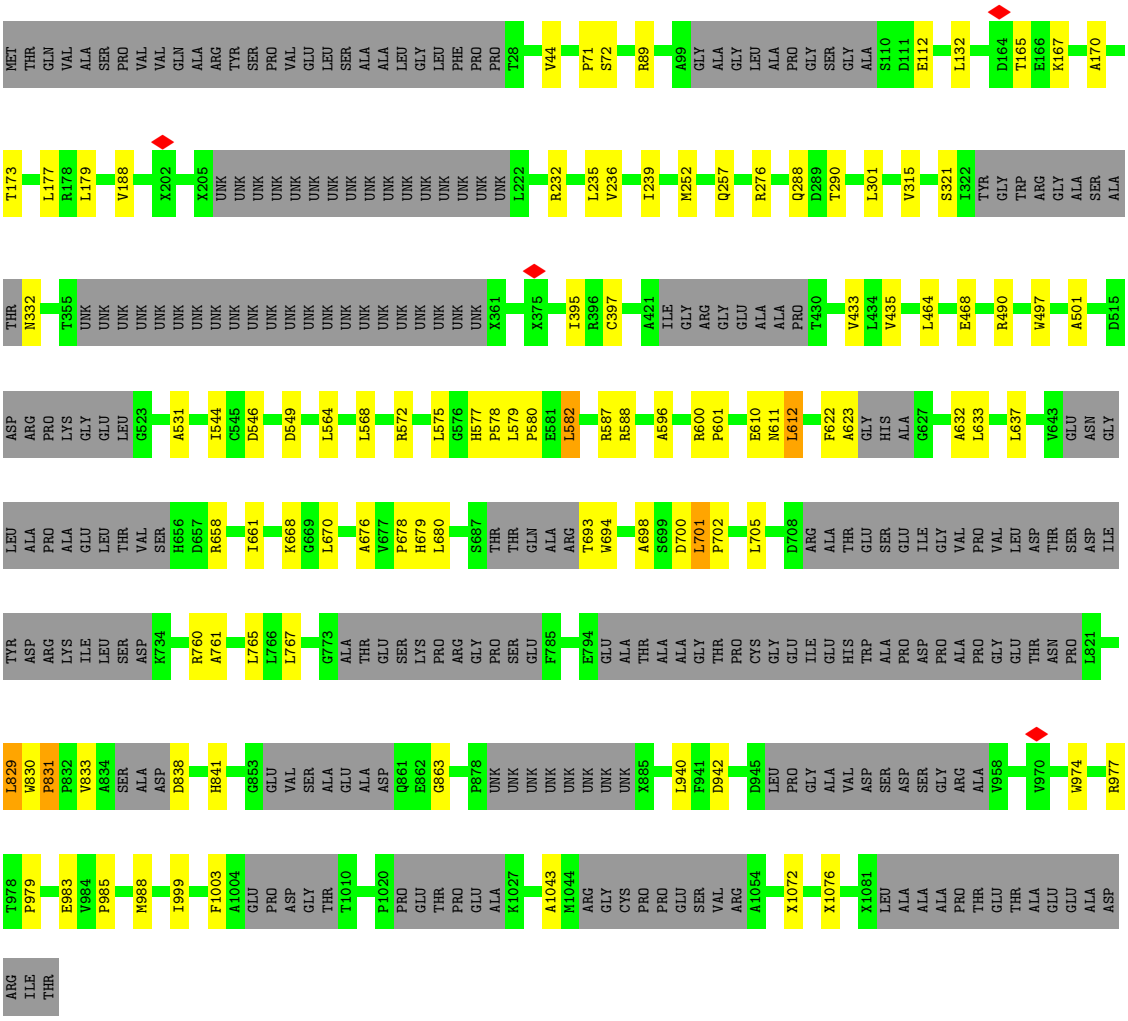
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
4	A	1	31	10	6	12	3	0

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.

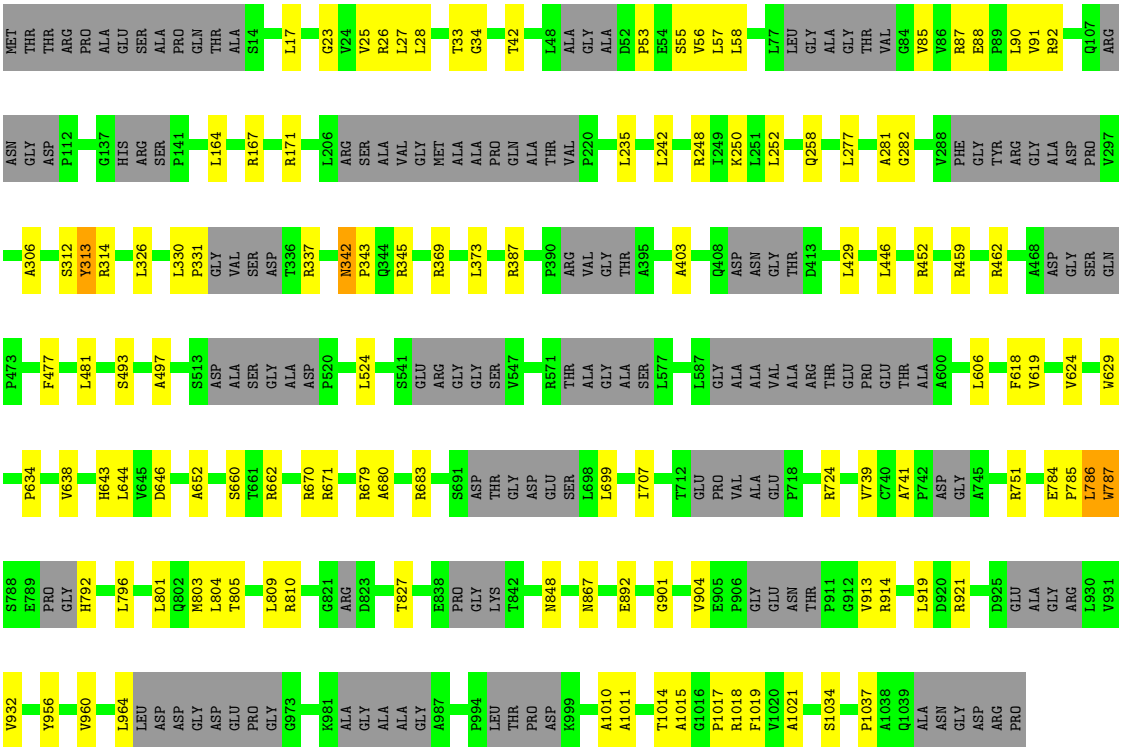
- Molecule 1: UvrD/REP helicase

Chain B: 68% 9% 23%



- Molecule 2: ATP-dependent DNA helicase (UvrD/REP)

Chain A: 75% 11% 14%


MET THR THR ARG PRO ALA GLU SER ALA PRO GLN THR ALA S14 L17 G23 V24 V25 R26 L27 L28 T33 G34 T42 L48 ALA GLY ALA D52 P53 E54 S55 V56 L57 L58 L77 LEU GLY ALA GLY THR VAL G84 V85 V86 R87 E88 P89 L90 V91 R92 Q107 ARG
ASN GLY ASP P112 G137 HIS ARG SER P141 L164 R167 R171 L206 ARG SER ALA VAL GLY MET ALA ALA PRO GLN ALA THR VAL P220 L235 L242 R248 I249 K250 L251 L252 Q258 L277 A281 G282 V288 PHE GLY TYR ARG GLY ALA ASP PRO V297
A306 S312 Y313 R314 L326 L330 P331 GLY VAL SER ASP T336 R337 N342 P343 Q344 R345 R369 L373 R387 P390 ARG VAL GLY THR A395 A403 Q408 ASP ASN GLY THR D413 L429 L446 R452 R459 R462 A468 ASP GLY SER GLN
P473 F477 L481 S493 A497 S513 ASP ALA SER GLY ALA ASP P520 L524 S541 GLU ARG GLY GLY SER V547 R571 THR ALA GLY ALA SER L577 L587 GLY ALA ALA VAL ALA ARG THR GLU PRO GLU THR ALA A600 L606 F618 V619 V624 W629
P634 V638 H643 L644 V645 D646 A652 S660 T661 R662 R670 R671 R679 A680 R683 S691 ASP THR GLY ASP GLU SER L698 L699 I707 T712 GLU PRO VAL ALA GLU P718 R724 V739 C740 A741 P742 ASP GLY A745 R751 E784 P785 L786 W787
S788 E789 PRO GLY H792 L796 L801 Q802 M803 L804 T805 L809 R810 G821 ARG D823 T827 E838 PRO GLY LYS T842 N848 N867 E892 G901 V904 E905 P906 GLY GLU ASN THR P911 G912 V913 R914 L919 D920 R921 D925 GLU ALA GLY ARG L930 V931
V932 Y956 V960 L964 LEU ASP ASP GLY ASP GLU PRO GLY G973 K981 ALA GLY ALA ALA GLY A987 P994 LEU THR PRO ASP K999 A1010 A1011 T1014 A1015 G1016 P1017 R1018 F1019 V1020 A1021 S1034 P1037 A1038 Q1039 ALA ASN GLY ASP ARG PRO

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	80396	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 ($e^-/Å^2$)	2.16	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	0.173	Depositor
Minimum map value	-0.068	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.01	Depositor
Map size (Å)	241.164, 241.164, 241.164	wwPDB
Map dimensions	280, 280, 280	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.8613, 0.8613, 0.8613	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: ANP, SF4

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	B	0.42	0/5466	0.61	0/7494
2	A	0.40	0/6302	0.61	0/8626
All	All	0.41	0/11768	0.61	0/16120

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	B	5756	0	5147	116	0
2	A	6209	0	5972	111	0
3	A	8	0	0	0	0
4	A	31	0	13	2	0
All	All	12004	0	11132	207	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 9.

The worst 5 of 207 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:787:TRP:NE1	2:A:1015:ALA:HA	1.37	1.40
2:A:787:TRP:HE1	2:A:1015:ALA:CA	1.43	1.29
1:B:579:LEU:HD23	1:B:633:LEU:CD1	1.71	1.19
1:B:188:VAL:CG1	1:B:705:LEU:HD21	1.78	1.14
1:B:579:LEU:HD23	1:B:633:LEU:HD11	1.25	1.11

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	B	733/1095 (67%)	684 (93%)	49 (7%)	0	100	100
2	A	849/1045 (81%)	783 (92%)	66 (8%)	0	100	100
All	All	1582/2140 (74%)	1467 (93%)	115 (7%)	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	B	451/743 (61%)	443 (98%)	8 (2%)	59	81
2	A	540/796 (68%)	530 (98%)	10 (2%)	57	80
All	All	991/1539 (64%)	973 (98%)	18 (2%)	61	81

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	786	LEU
2	A	964	LEU
2	A	848	ASN
2	A	258	GLN
2	A	624	VAL

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

Mol	Chain	Res	Type
2	A	46	HIS
2	A	342	ASN
2	A	848	ASN
1	B	577	HIS
1	B	257	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)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	ANP	A	1102	-	29,33,33	1.18	5 (17%)	31,52,52	1.09	2 (6%)
3	SF4	A	1101	-	0,12,12	-	-	-		

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
4	ANP	A	1102	-	-	7/14/38/38	0/3/3/3
3	SF4	A	1101	-	-	-	0/6/5/5

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1102	ANP	PG-O1G	2.97	1.50	1.46
4	A	1102	ANP	PB-O1B	2.84	1.50	1.46
4	A	1102	ANP	PB-O2B	-2.28	1.50	1.56
4	A	1102	ANP	PG-O2G	-2.16	1.51	1.56
4	A	1102	ANP	PG-O3G	-2.13	1.51	1.56

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1102	ANP	O2B-PB-O1B	4.16	118.65	109.92
4	A	1102	ANP	C5-C6-N6	2.29	123.83	120.35

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1102	ANP	PB-N3B-PG-O1G
4	A	1102	ANP	PG-N3B-PB-O1B
4	A	1102	ANP	C5'-O5'-PA-O2A
4	A	1102	ANP	O4'-C4'-C5'-O5'
4	A	1102	ANP	C3'-C4'-C5'-O5'

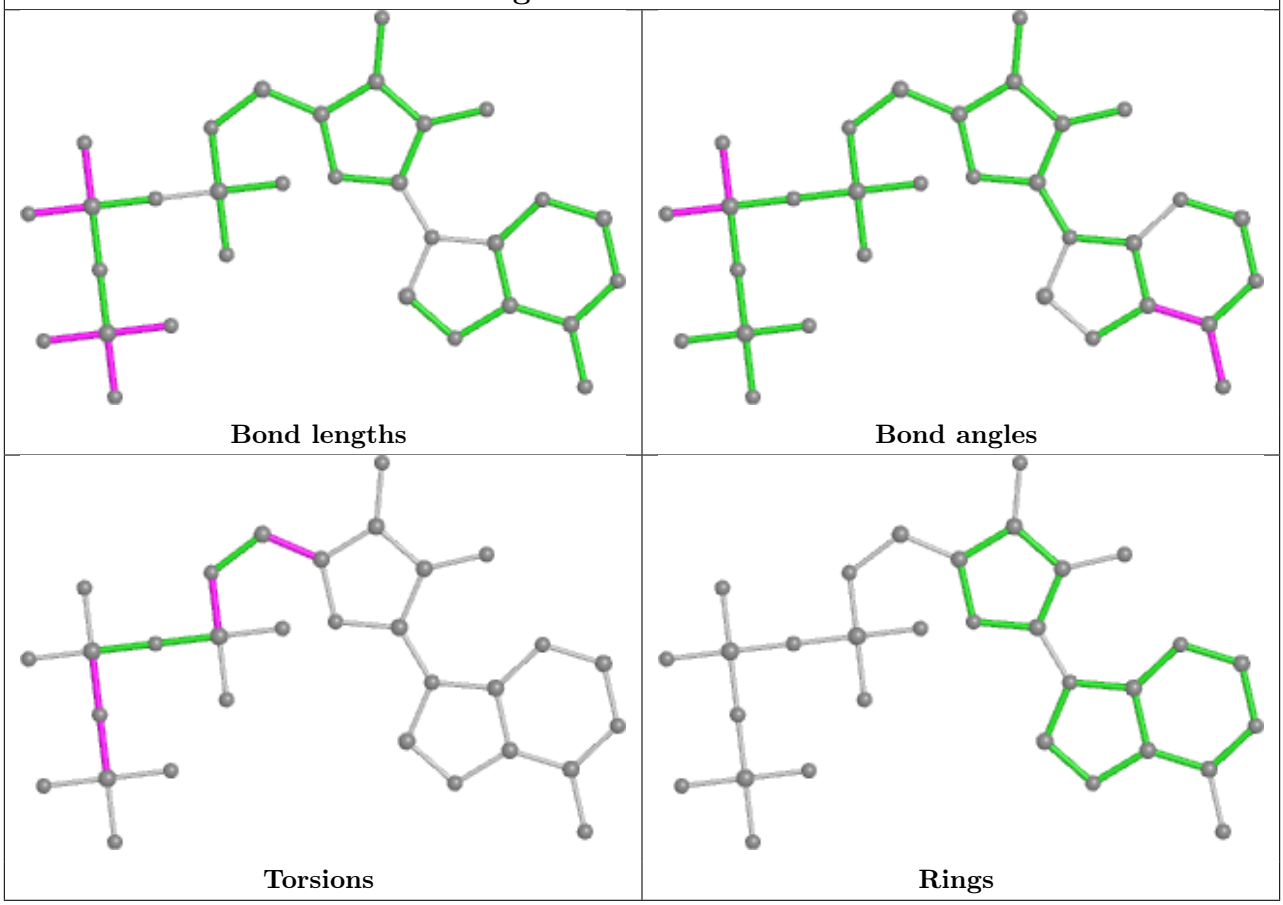
There are no ring outliers.

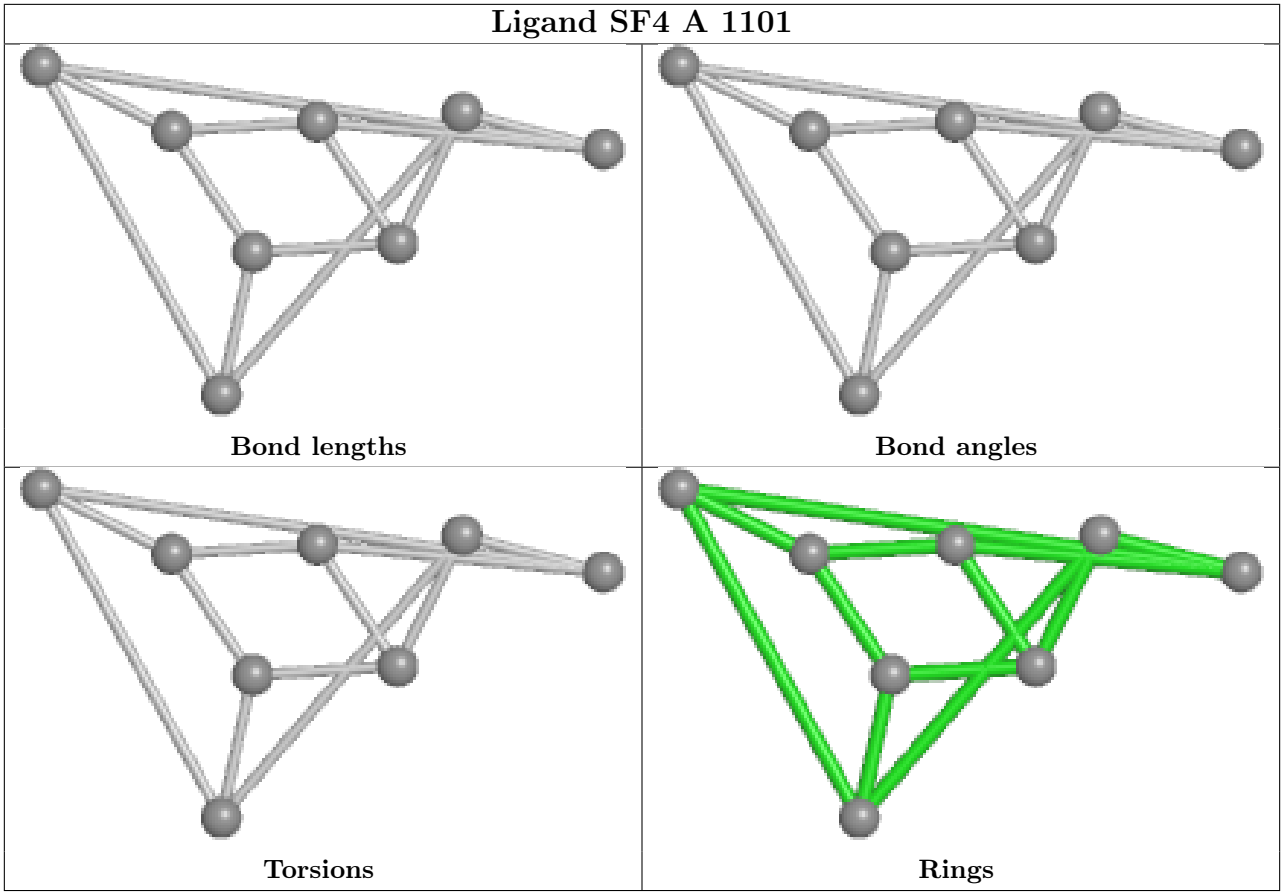
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1102	ANP	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 ANP A 1102





5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	376:UNK	C	394:THR	N	25.56
1	B	917:UNK	C	920:HIS	N	10.39
1	B	1067:UNK	C	1070:UNK	N	5.56

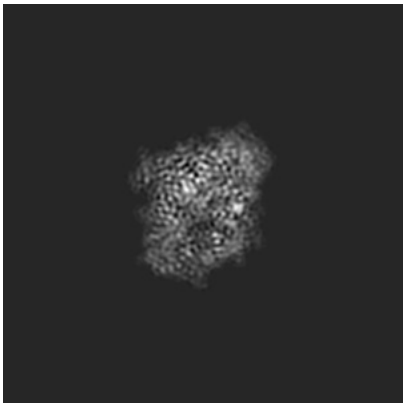
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-20440. These allow visual inspection of the internal detail of the map and identification of artifacts.

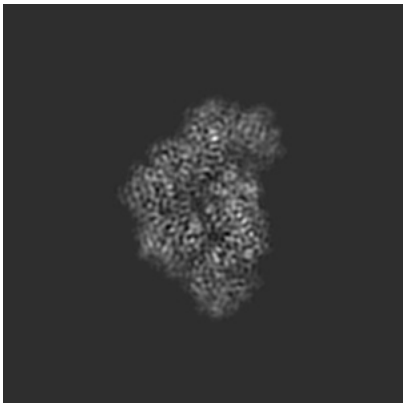
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

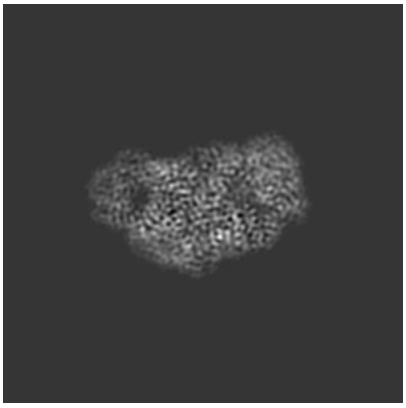
6.1.1 Primary map



X



Y



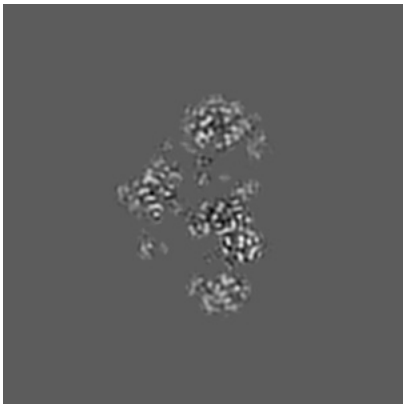
Z

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

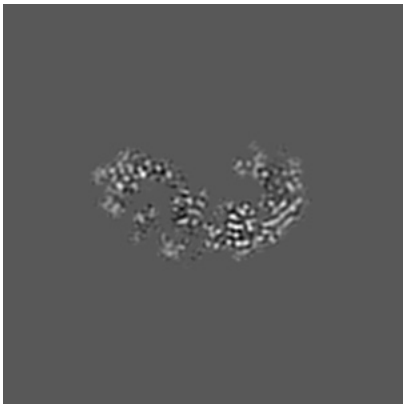
6.2 Central slices (i)

6.2.1 Primary map

X Index: 140



Y Index: 140



Z Index: 140

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: 121

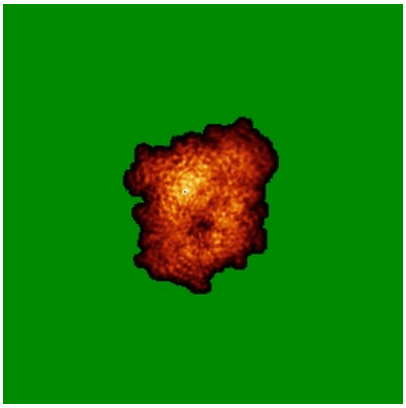
Y Index: 127

Z Index: 148

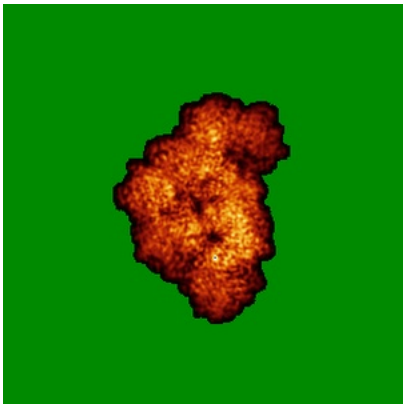
The images above show the largest variance slices of the map in three orthogonal directions.

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

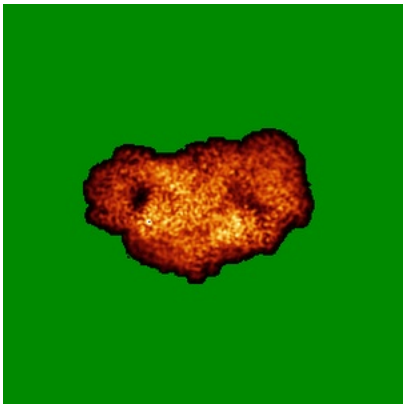
6.4.1 Primary map



X



Y

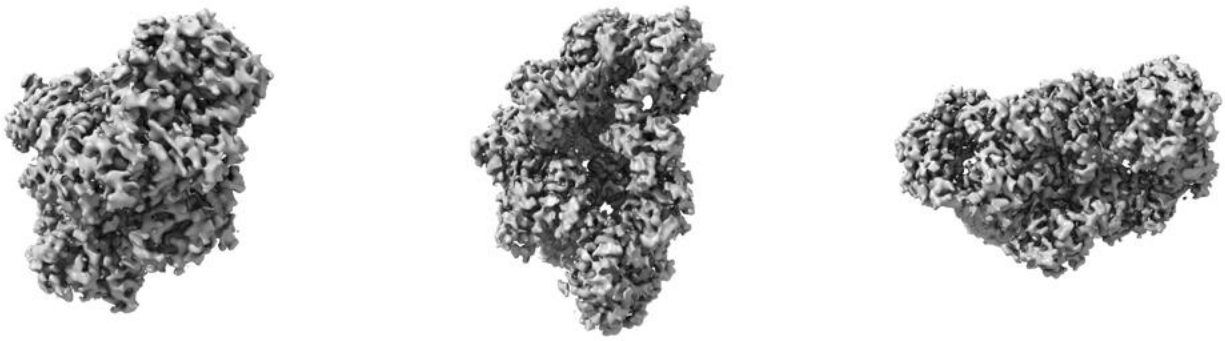


Z

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

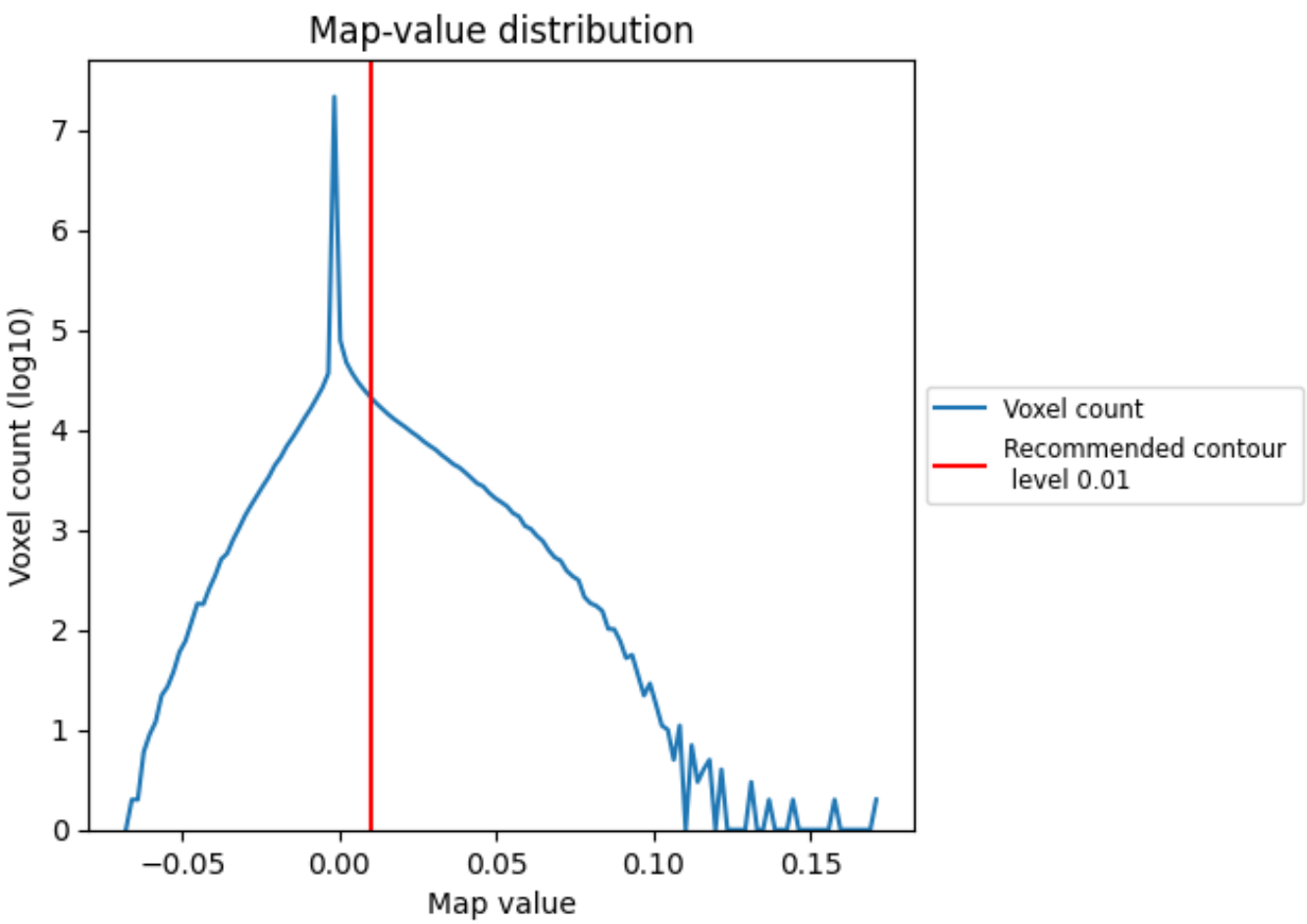
6.6 Mask visualisation (i)

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

7 Map analysis (i)

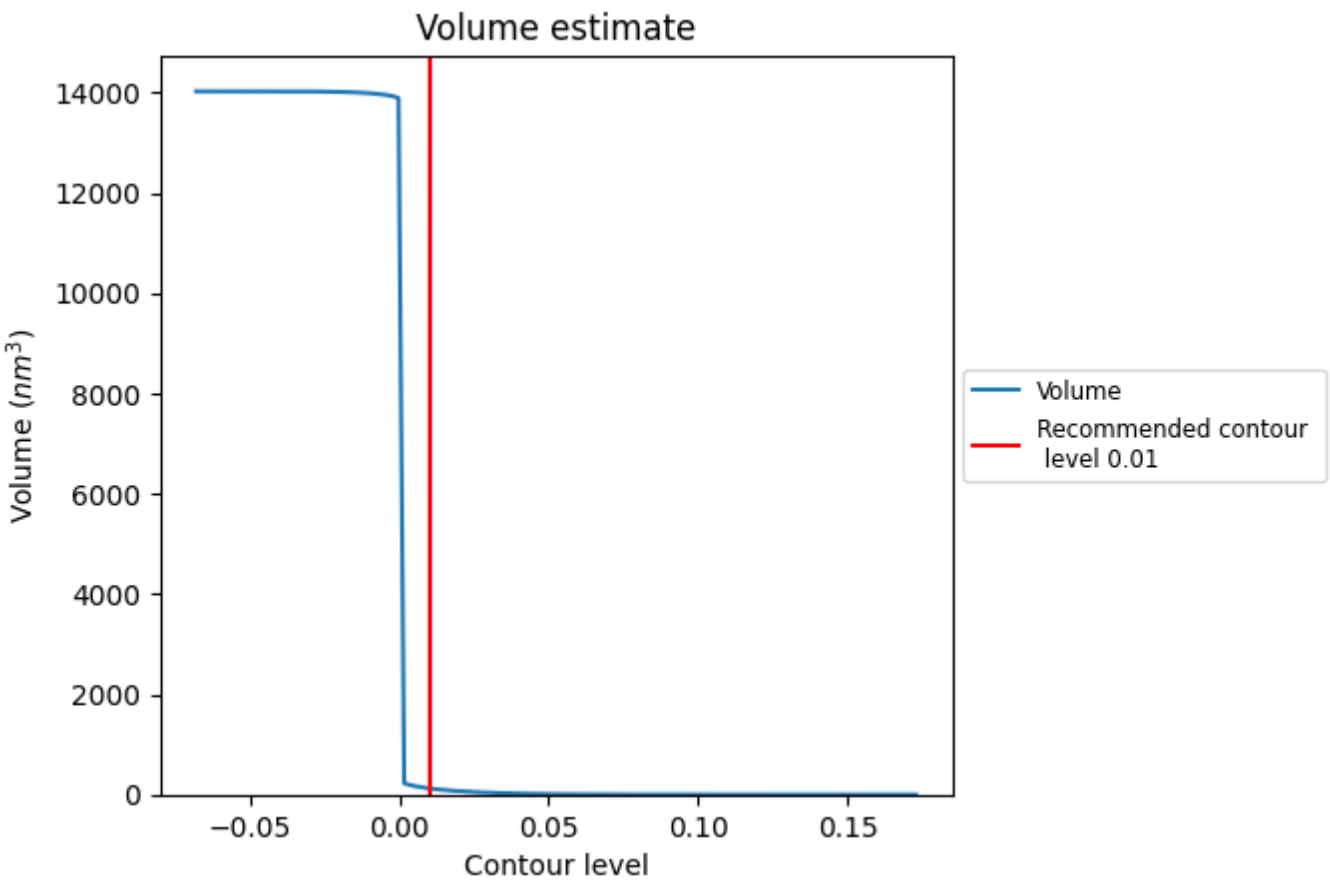
This section contains the results of statistical analysis of the map.

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.

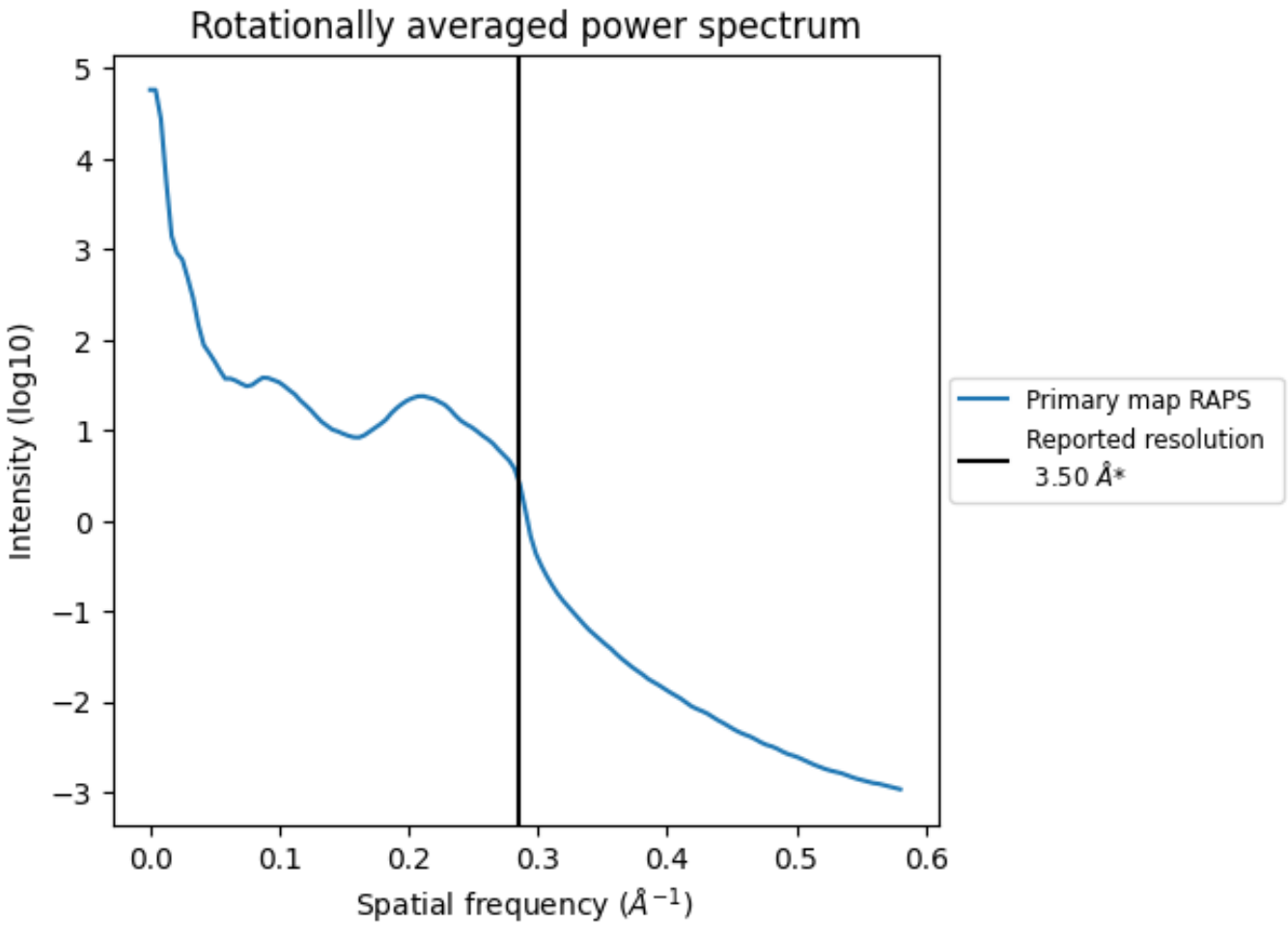
7.2 Volume estimate (i)



The volume at the recommended contour level is 123 nm^3; this corresponds to an approximate mass of 111 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)

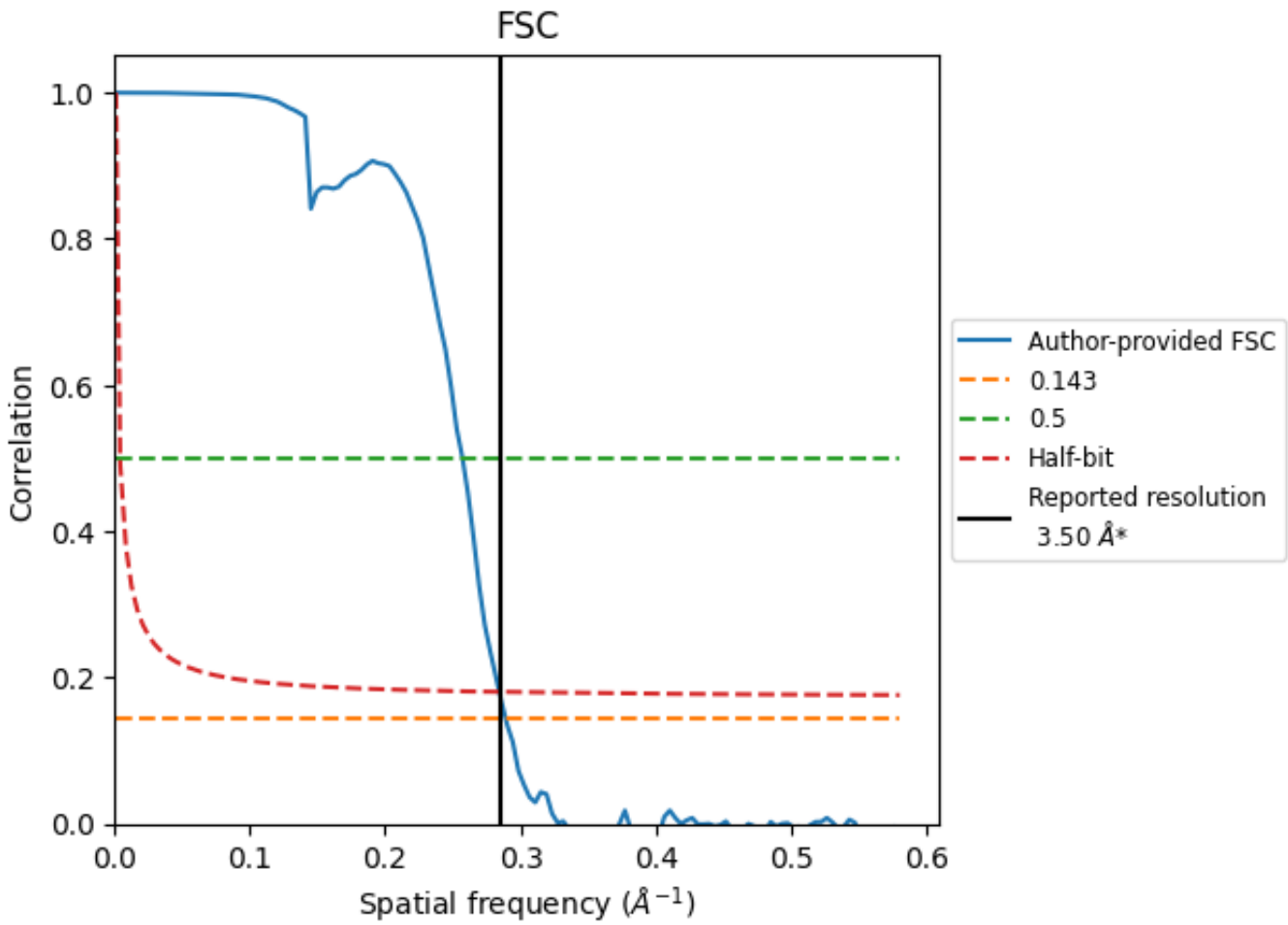


*Reported resolution corresponds to spatial frequency of 0.286 Å^{-1}

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 0.286 Å^{-1}

8.2 Resolution estimates (i)

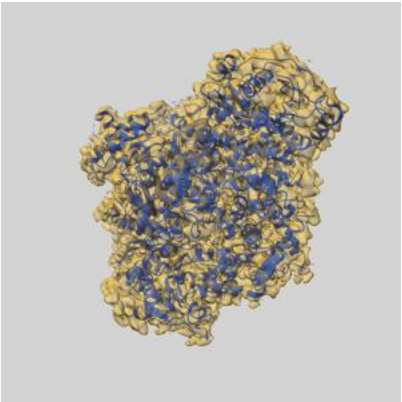
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.50	-	-
Author-provided FSC curve	3.46	3.89	3.52
Unmasked-calculated*	-	-	-

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

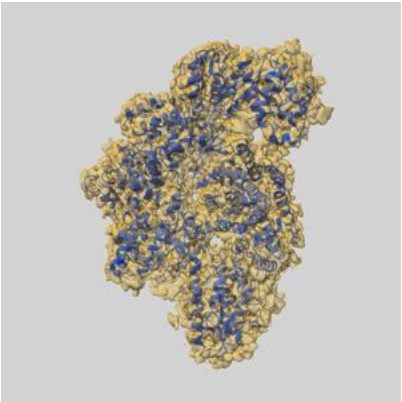
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-20440 and PDB model 6PPJ. Per-residue inclusion information can be found in section 3 on page 5.

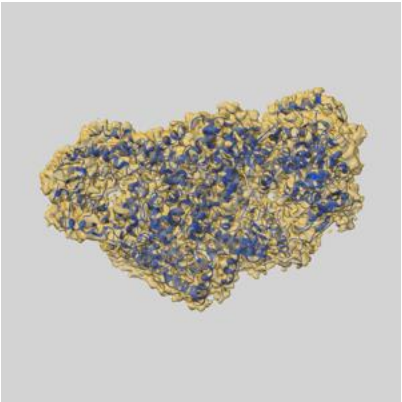
9.1 Map-model overlay (i)



X



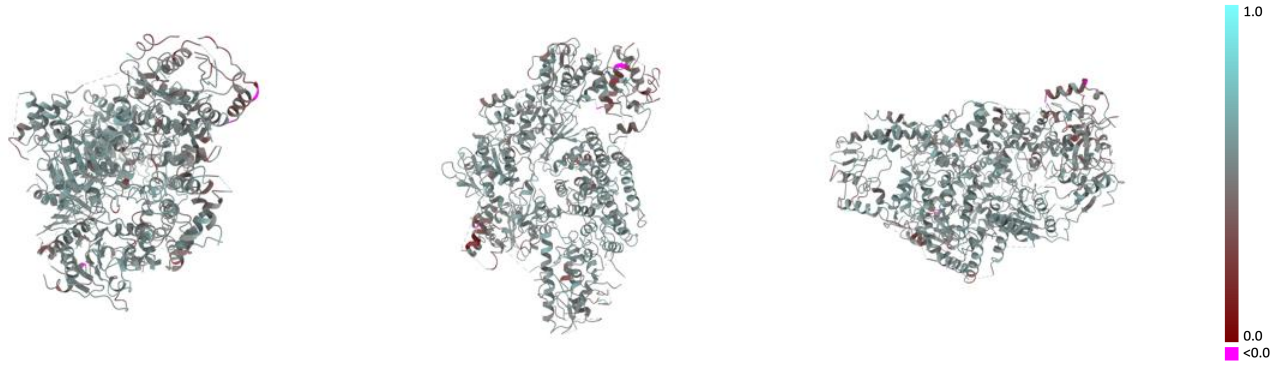
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.01 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.

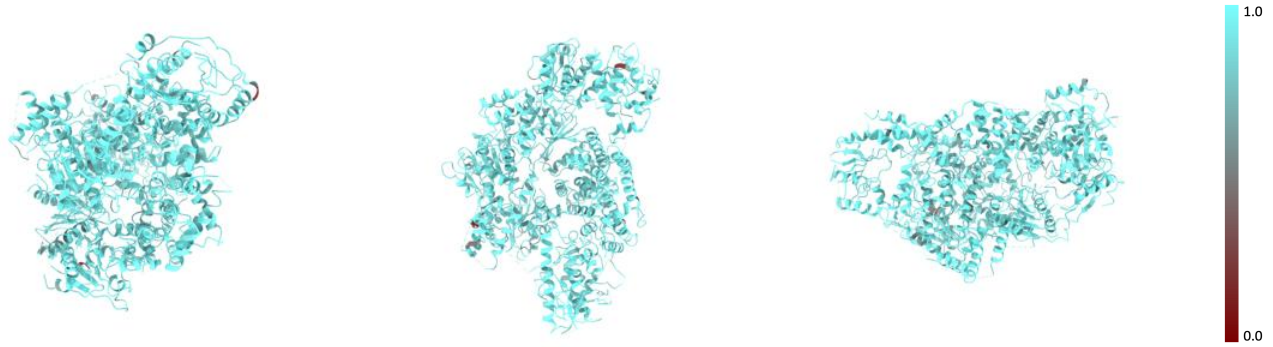
9.2 Q-score mapped to coordinate model (i)



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.

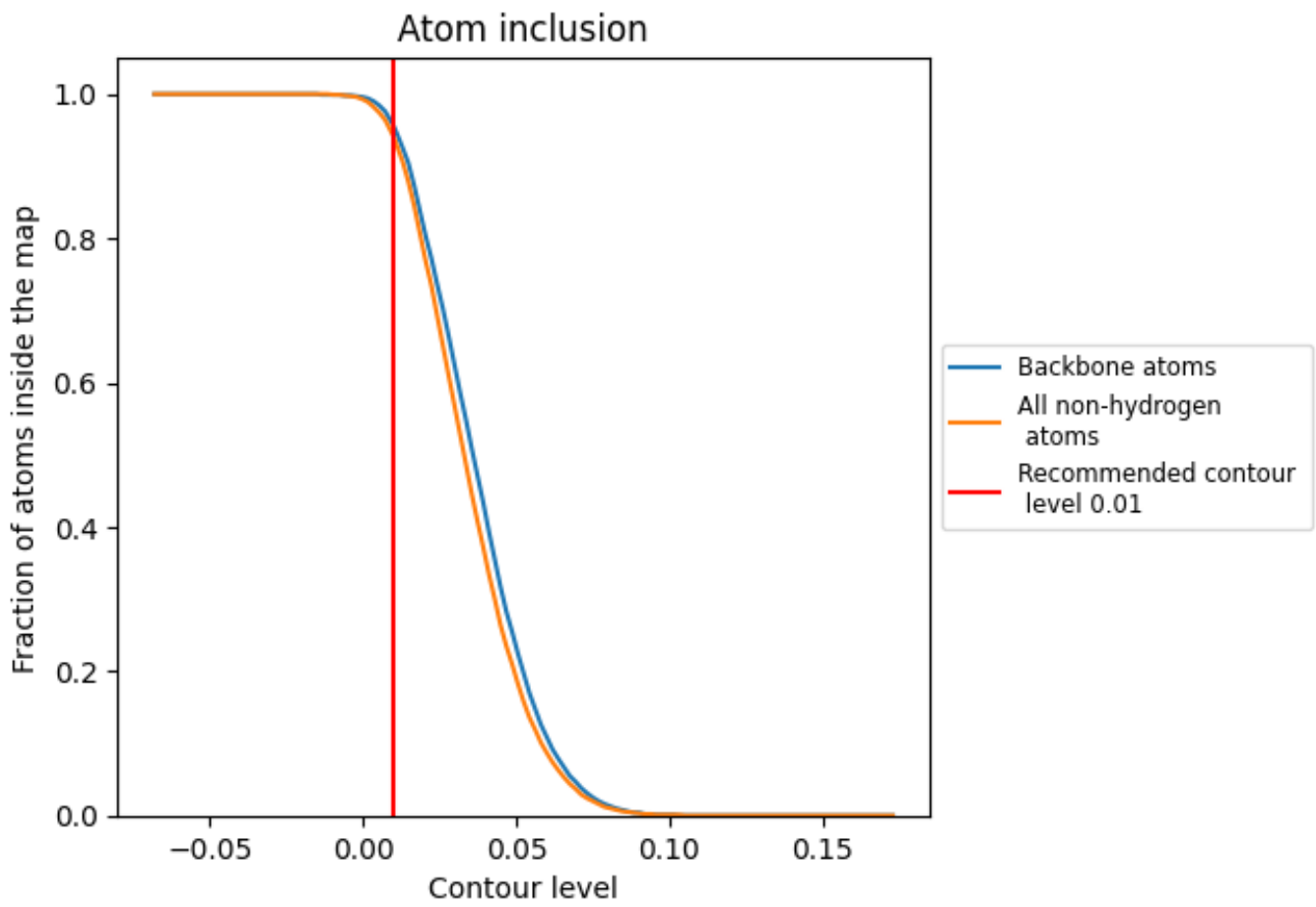
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.01).

9.4 Atom inclusion (i)



At the recommended contour level, 96% of all backbone atoms, 94% 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 (0.01) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9410	0.5190
A	0.9500	0.5270
B	0.9320	0.5100