

wwPDB EM Validation Summary Report (i)

Nov 1, 2022 – 06:46 PM EDT

PDB ID : 5KGF EMDB ID : EMD-8246

Title: Structural model of 53BP1 bound to a ubiquitylated and methylated nucleo-

some, at 4.5 A resolution

Authors : Wilson, M.D.; Benlekbir, S.; Sicheri, F.; Rubinstein, J.L.; Durocher, D.

Deposited on : 2016-06-13

Resolution : 4.54 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.dev43

Mogul : 1.8.5 (274361), CSD as541be (2020)

MolProbity: 4.02b-467

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)

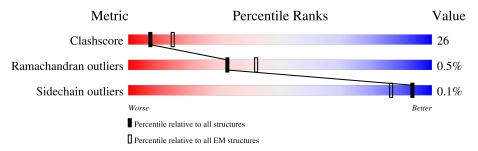
Validation Pipeline (wwPDB-VP) : 2.31.2

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 4.54 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 $(\# \mathrm{Entries})$	${ m EM\ structures} \ (\#{ m Entries})$
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	Qual	lity of chain	
1	A	136	43%	32%	26%
1	Е	136	43%	30% •	26%
2	В	103	7% 45%	35%	• 18%
2	F	103	48%	34%	18%
3	С	130	46%	40%	• 13%
3	G	130	45%	39%	• 13%
4	D	126	43%	33% •	21%
4	Н	126	49%	26% •	23%



 $Continued\ from\ previous\ page...$

Mol		Length		lity of chain	
5	I	145	42%	41%	17%
6	J	145	39%	49%	12%
7	K	21	38% 43%	57%	
7	L	21	48% 48%	52%	
8	M	76	33%	66%	
8	О	76	30%	64%	



2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 13782 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 Histone H3.2.

\mathbf{Mol}	Chain	Residues		At	oms			AltConf	Trace	
1	A	101	Total 832		N 161			0	0	
1	E	100	Total 825			O 142	D	0	0	

• Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues		At	oms			AltConf	Trace
2	В	84		425	133	115	2	0	0
2	F	84	Total 675		N 133			0	0

• Molecule 3 is a protein called Histone H2A type 1.

Mol	Chain	Residues		Ato	ms		AltConf	Trace
2	C	113	Total	С	N	О	0	0
3		110	874	549	174	151	0	U
2	С	113	Total	С	N	О	0	0
)	G	110	874	549	174	151	U	U

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
С	13	ARG	LYS	engineered mutation	UNP P0C0S8
С	16	SER	THR	engineered mutation	UNP P0C0S8
С	36	ARG	LYS	engineered mutation	UNP P0C0S8
G	13	ARG	LYS	engineered mutation	UNP P0C0S8
G	16	SER	THR	engineered mutation	UNP P0C0S8
G	36	ARG	LYS	engineered mutation	UNP P0C0S8

• Molecule 4 is a protein called Histone H2B type 1-C/E/F/G/I.



Mol	Chain	Residues		At	oms			AltConf	Trace
1	D	100	Total	С	N	О	S	0	0
4	ט	100	788	493	147	146	2	0	U
1	П	97	Total	С	N	О	S	0	0
4	11	91	766	479	142	143	2	U	

• Molecule 5 is a DNA chain called DNA (145-MER).

Mo	l C	Chain	Residues		\mathbf{A}^{1}	toms			AltConf	Trace
5		I	145	Total 2952	C 1404	N 537	O 867	P 144	0	0

• Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues		\mathbf{A}^{1}	toms			AltConf	Trace
6	J	145	Total 2987	C 1416	N 558	O 869	P 144	0	0

• Molecule 7 is a protein called Tumor suppressor p53-binding protein 1.

Mol	Chain	Residues	Atoms	AltConf	Trace
7	L	21	Total C N 166 100 34	0	0
7	K	21	Total C N 166 100 34	0	0

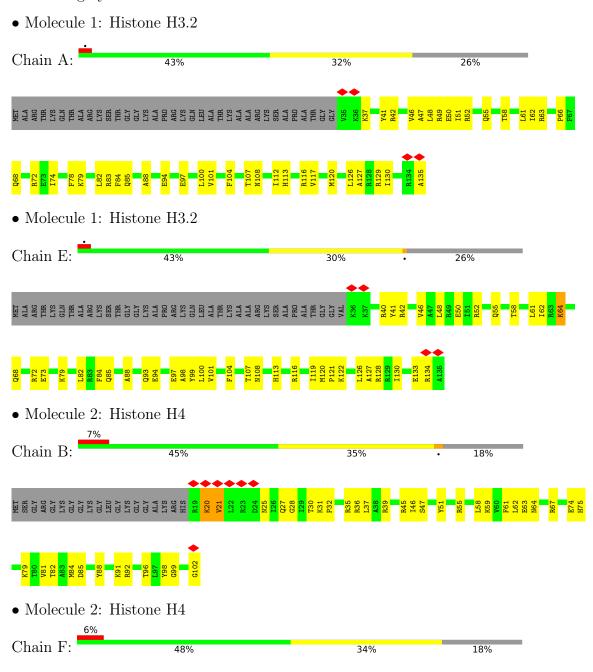
• Molecule 8 is a protein called Ubiquitin.

Mol	Chain	Residues		At	oms			AltConf	Trace
Q	0	76	Total	С	N	О	S	0	0
0		70	601	378	105	117	1	0	0
Q	М	76	Total	С	N	О	S	0	0
0	1V1	10	601	378	105	117	1	0	U

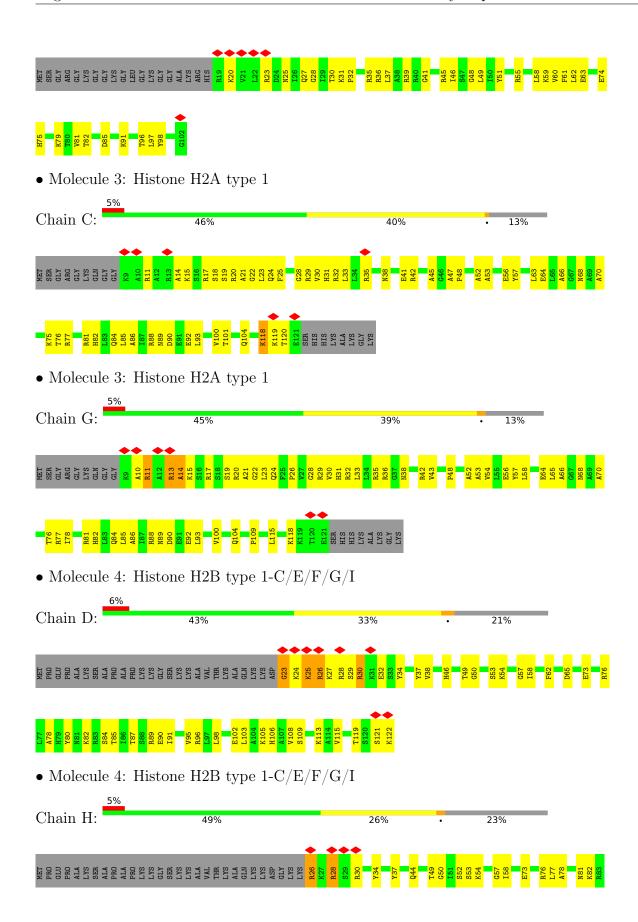


3 Residue-property plots (i)

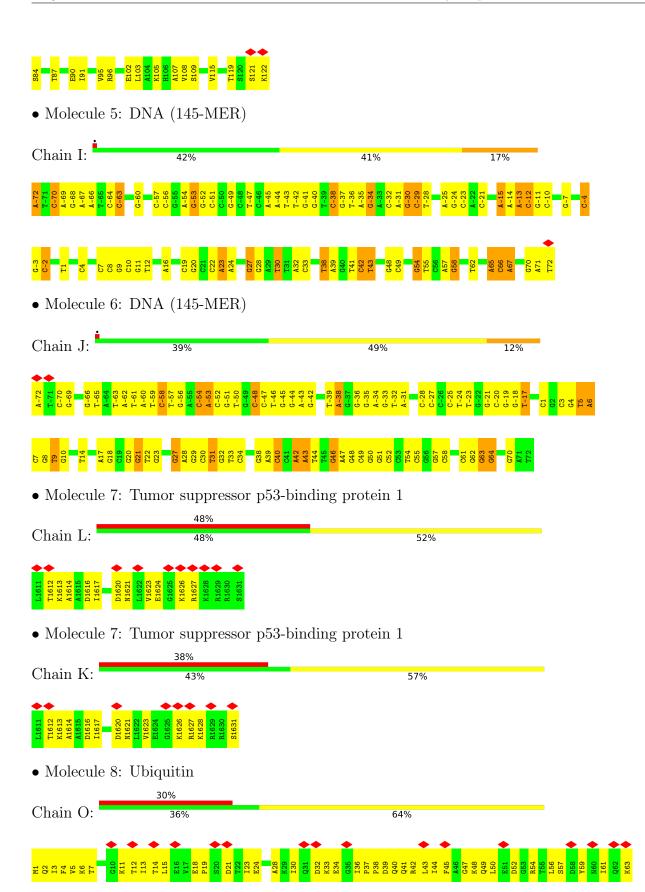
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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.







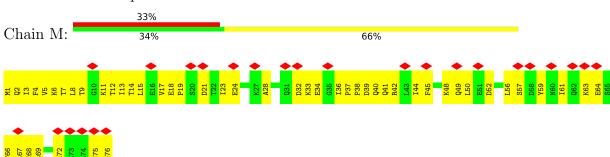








• Molecule 8: Ubiquitin





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	45361	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE	Depositor
	CORRECTION	
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose $(e^-/\text{Å}^2)$	36	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	34483	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.177	Depositor
Minimum map value	-0.087	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.013	Depositor
Recommended contour level	0.04	Depositor
Map size (Å)	185.6, 185.6, 185.6	wwPDB
Map dimensions	128, 128, 128	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.45, 1.45, 1.45	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: M2L

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	Bo	nd lengths	В	ond angles
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	A	0.46	0/844	0.65	0/1130
1	Е	0.46	0/837	0.62	0/1120
2	В	0.48	0/670	0.63	0/894
2	F	0.48	0/670	0.58	0/894
3	С	0.45	0/884	0.74	0/1190
3	G	0.44	0/884	0.69	0/1190
4	D	0.48	0/799	0.62	0/1067
4	Н	0.48	0/777	0.56	0/1040
5	I	0.74	1/3308 (0.0%)	1.43	43/5099 (0.8%)
6	J	0.73	0/3354	1.45	48/5180 (0.9%)
7	K	0.33	0/165	0.98	1/216 (0.5%)
7	L	0.35	0/165	0.95	1/216 (0.5%)
8	M	0.32	0/607	0.54	0/816
8	O	0.32	0/607	0.54	0/816
All	All	0.60	1/14571 (0.0%)	1.11	93/20868 (0.4%)

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 maintain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	F	0	1
3	С	0	1
3	G	0	3
4	D	0	3
4	Н	0	2
7	K	0	1
7	L	0	1
All	All	0	12



All (1) bond length outliers are listed below:

Mol			V -			$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}(ext{\AA})$
5	I	-72	DA	O5'-C5'	9.26	1.65	1.42

The worst 5 of 93 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
6	J	20	DG	O3'-P-O5'	-11.97	81.26	104.00
6	J	20	DG	OP2-P-O3'	-10.03	83.14	105.20
6	J	14	DT	O4'-C1'-N1	8.79	114.15	108.00
6	J	21	DG	O4'-C1'-N9	8.60	114.02	108.00
6	J	34	DC	P-O3'-C3'	8.10	129.41	119.70

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	С	118	LYS	Peptide
4	D	23	GLY	Peptide
4	D	25	LYS	Peptide
4	D	26	ARG	Mainchain
2	F	23	ARG	Peptide

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	832	0	878	67	0
1	Е	825	0	867	46	0
2	В	675	0	722	58	0
2	F	675	0	722	37	0
3	С	874	0	939	100	0
3	G	874	0	939	95	0
4	D	788	0	824	84	0
4	Н	766	0	795	62	0
5	I	2952	0	1629	119	0
6	J	2987	0	1630	176	0
7	K	166	0	180	26	0



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	L	166	0	180	14	0
8	M	601	0	629	49	0
8	О	601	0	629	51	0
All	All	13782	0	11563	665	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 26.

The worst 5 of 665 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$egin{aligned} ext{Clash} \ ext{overlap } (ext{Å}) \end{aligned}$
5:I:-72:DA:C5'	5:I:-72:DA:O5'	1.65	1.41
4:D:30:ARG:CB	6:J:49:DC:H4'	1.46	1.41
4:D:30:ARG:HB2	6:J:49:DC:C4'	1.54	1.36
1:E:41:TYR:OH	5:I:-66:DA:H5'	1.19	1.33
4:D:29:SER:CB	4:D:30:ARG:HH21	1.43	1.31

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	Perce	entiles
1	A	99/136~(73%)	91 (92%)	8 (8%)	0	100	100
1	Е	98/136 (72%)	92 (94%)	5 (5%)	1 (1%)	15	54
2	В	81/103 (79%)	72 (89%)	8 (10%)	1 (1%)	13	50
2	F	81/103 (79%)	74 (91%)	7 (9%)	0	100	100
3	С	111/130 (85%)	100 (90%)	10 (9%)	1 (1%)	17	56
3	G	111/130 (85%)	104 (94%)	5 (4%)	2 (2%)	8	42
4	D	98/126~(78%)	93 (95%)	5 (5%)	0	100	100



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
4	Н	95/126 (75%)	91 (96%)	4 (4%)	0	100 100
7	K	19/21~(90%)	15 (79%)	4 (21%)	0	100 100
7	L	19/21 (90%)	15 (79%)	4 (21%)	0	100 100
8	M	74/76~(97%)	73 (99%)	1 (1%)	0	100 100
8	О	74/76 (97%)	73 (99%)	1 (1%)	0	100 100
All	All	960/1184 (81%)	893 (93%)	62 (6%)	5 (0%)	32 68

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	G	14	ALA
3	С	14	ALA
3	G	13	ARG
2	В	21	VAL
1	Е	64	LYS

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	Perce	ntiles
1	A	88/111 (79%)	88 (100%)	0	100	100
1	E	87/111 (78%)	87 (100%)	0	100	100
2	В	68/78~(87%)	68 (100%)	0	100	100
2	F	68/78~(87%)	68 (100%)	0	100	100
3	C	88/99~(89%)	88 (100%)	0	100	100
3	G	88/99 (89%)	88 (100%)	0	100	100
4	D	86/106~(81%)	85 (99%)	1 (1%)	71	84
4	Н	84/106 (79%)	84 (100%)	0	100	100
7	K	18/18 (100%)	18 (100%)	0	100	100
7	L	18/18 (100%)	18 (100%)	0	100	100



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Perce	ntiles
8	M	68/68 (100%)	68 (100%)	0	100	100
8	О	68/68 (100%)	68 (100%)	0	100	100
All	All	829/960 (86%)	828 (100%)	1 (0%)	93	96

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	30	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	75	HIS
2	F	93	GLN
8	M	41	GLN
3	G	84	GLN
8	O	41	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)

2 non-standard protein/DNA/RNA residues 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	Tuno	Chain	Dag	Link	B	Bond lengths			Bond angles		
MIOI	Type	Chain	Res	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2	
2	M2L	F	20	2	9,10,11	3.47	2 (22%)	6,11,13	2.41	4 (66%)	
2	M2L	В	20	2	9,10,11	3.47	2 (22%)	6,11,13	2.43	4 (66%)	



In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	M2L	F	20	2	-	2/7/9/11	-
2	M2L	В	20	2	-	2/7/9/11	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	$Ideal(\AA)$
2	F	20	M2L	CD-SG	-7.57	1.52	1.81
2	В	20	M2L	CD-SG	-7.54	1.53	1.81
2	В	20	M2L	CB-SG	-6.99	1.52	1.80
2	F	20	M2L	CB-SG	-6.95	1.53	1.80

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
2	F	20	M2L	CM2-NZ-CM1	4.02	120.11	109.73
2	В	20	M2L	CM2-NZ-CM1	3.96	119.98	109.73
2	F	20	M2L	CB-SG-CD	2.42	109.52	102.27
2	В	20	M2L	CB-SG-CD	2.40	109.47	102.27
2	В	20	M2L	CM2-NZ-CE	2.35	120.04	110.74

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	В	20	M2L	CD-CE-NZ-CM2
2	F	20	M2L	CD-CE-NZ-CM1
2	F	20	M2L	CD-CE-NZ-CM2
2	В	20	M2L	CD-CE-NZ-CM1

There are no ring outliers.

1 monomer is involved in 1 short contact:

\mathbf{Mol}	Chain	Res	Type	Clashes	Symm-Clashes
2	В	20	M2L	1	0



5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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.



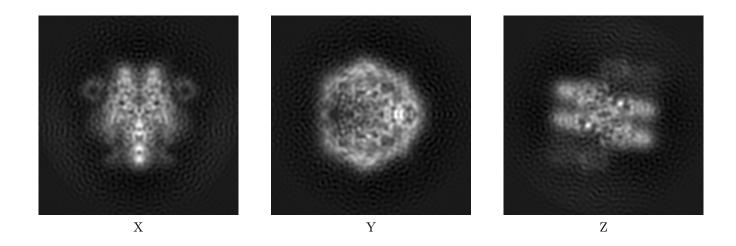
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-8246. 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)

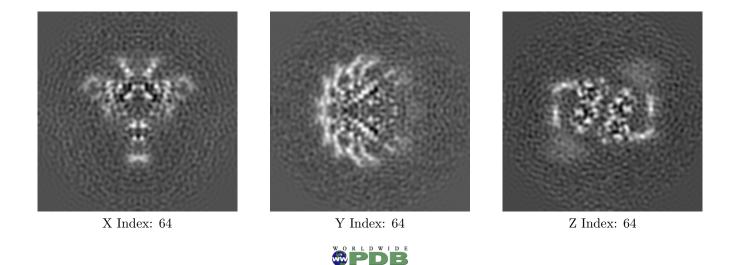
6.1.1 Primary map



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

6.2 Central slices (i)

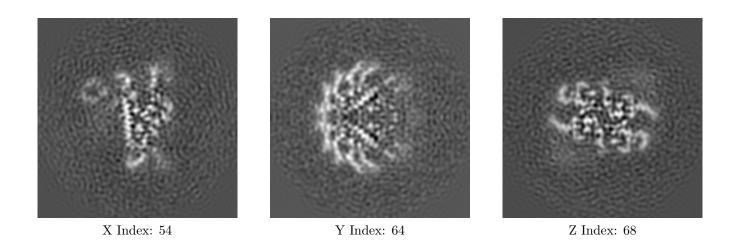
6.2.1 Primary map



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

6.3 Largest variance slices (i)

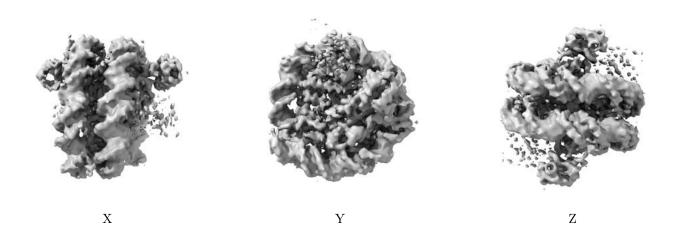
6.3.1 Primary map



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



The images above show the 3D surface view of the map at the recommended contour level 0.04. 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)

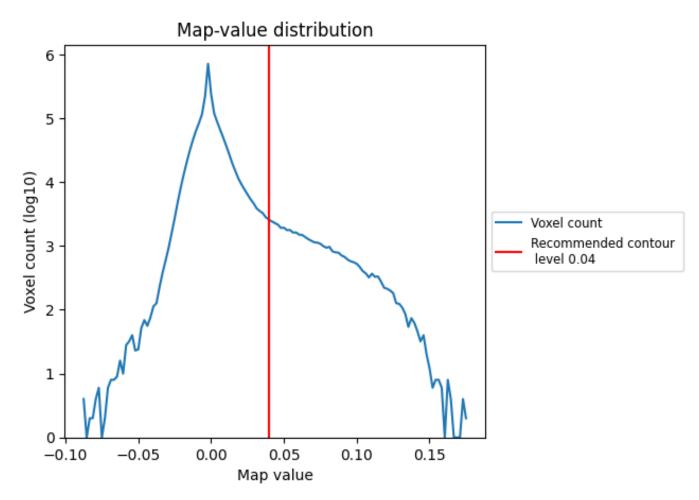
This section was not generated. No masks/segmentation were deposited.



7 Map analysis (i)

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

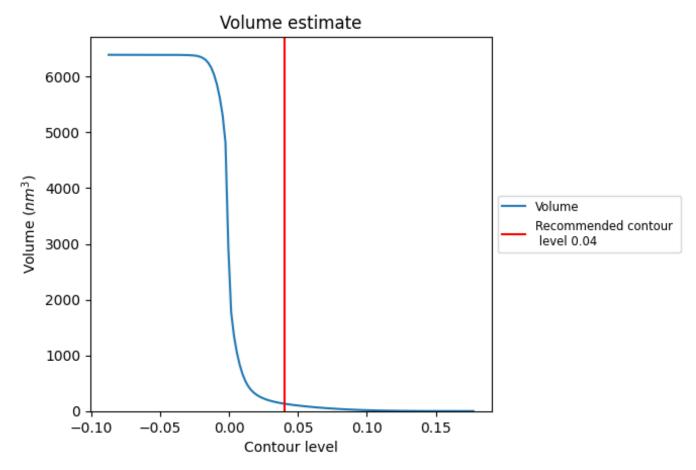
7.1 Map-value distribution (i)



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)

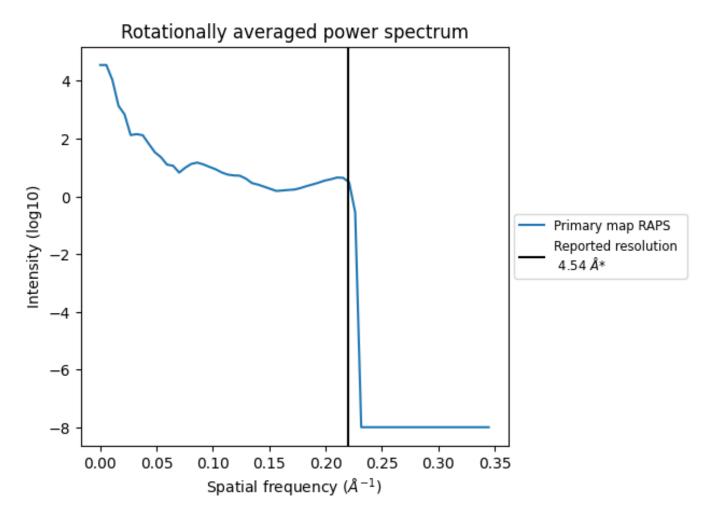


The volume at the recommended contour level is $133~\mathrm{nm}^3$; this corresponds to an approximate mass of $120~\mathrm{kDa}$.

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)



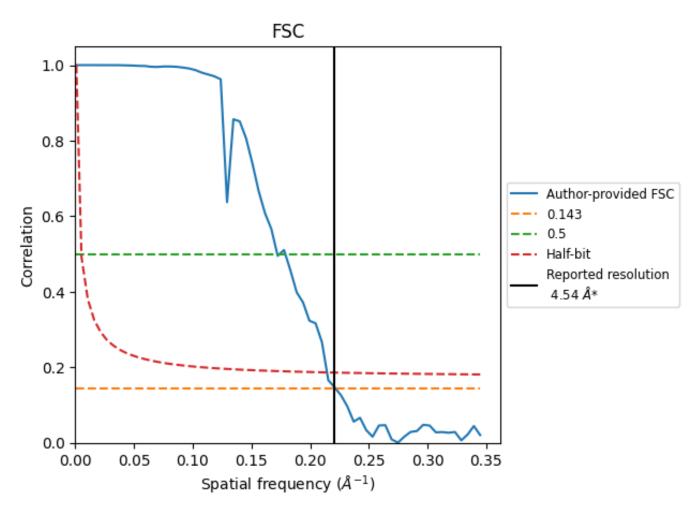
^{*}Reported resolution corresponds to spatial frequency of 0.220 $\rm \AA^{-1}$



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 0.220 $\rm \AA^{-1}$



8.2 Resolution estimates (i)

Resolution estimate (Å)	Estim	Estimation criterion (FSC cut-off)				
resolution estimate (A)	0.143	0.5	Half-bit			
Reported by author	4.54	-	-			
Author-provided FSC curve	4.51	5.81	4.66			
Unmasked-calculated*	-	-	-			

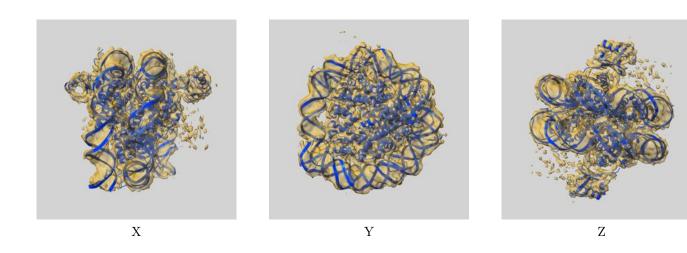
^{*}Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.



9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-8246 and PDB model 5KGF. Per-residue inclusion information can be found in section 3 on page 6.

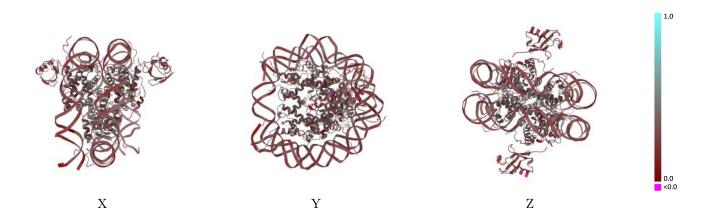
9.1 Map-model overlay (i)



The images above show the 3D surface view of the map at the recommended contour level 0.04 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.

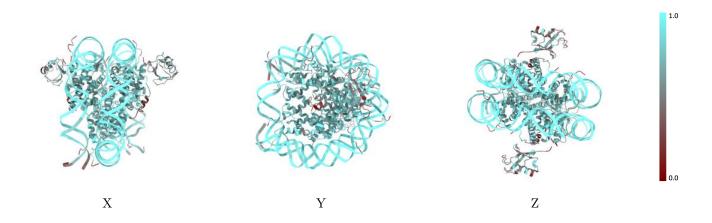


9.2 Q-score mapped to coordinate model (i)



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.

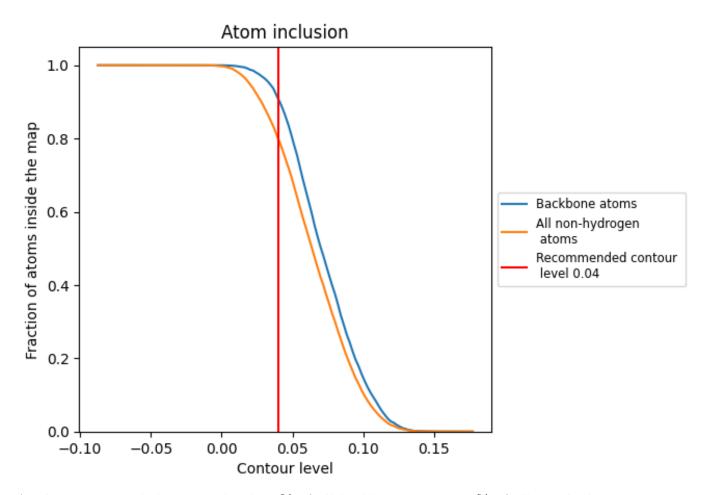
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.04).



9.4 Atom inclusion (i)



At the recommended contour level, 91% 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 (0.04) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7979	0.3010
A	0.7591	0.3520
В	0.7496	0.3500
С	0.7609	0.3370
D	0.7236	0.3210
E	0.7683	0.3590
F	0.7620	0.3550
G	0.7716	0.3420
Н	0.7356	0.3200
I	0.9211	0.2750
J	0.9150	0.2720
K	0.4688	0.2300
L	0.4437	0.2320
M	0.5051	0.2460
O	0.5203	0.2520



