wwPDB EM Validation Summary Report (i)

Dec 17, 2022 – 12:19 pm GMT

PDB ID	:	6Z6P
EMDB ID	:	EMD-11102
Title	:	HDAC-PC-Nuc
Authors	:	Lee, J.-H.; Bollschweiler, D.; Schaefer, T.; Huber, R.
Deposited on	:	2020-05-28
Resolution	:	4.43 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:

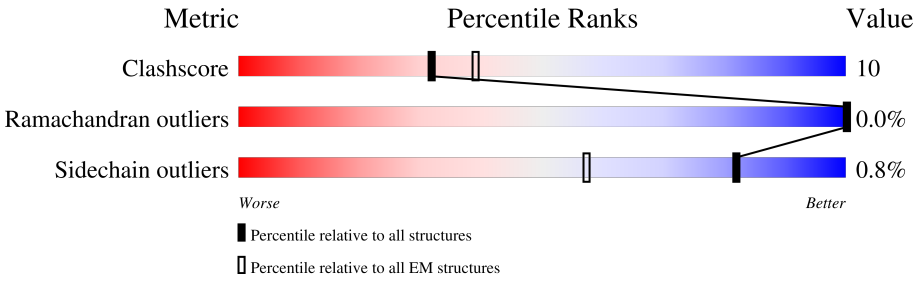
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 4.43 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	K	661	75% 23%
2	L	672	68% 29%
3	M	543	14% 79% 19%
4	N	629	11% 64% 22% 14%
5	A	97	67% 31%
6	B	83	67% 31%
7	C	103	72% 27%
8	D	95	65% 31%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	E	97	
10	F	78	
11	G	105	
12	H	93	
13	I	145	
14	J	145	

2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 31115 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 deacetylase HDA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	K	651	5178	3303	876	972	27	0	0

- Molecule 2 is a protein called Histone deacetylase HDA1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	L	661	5245	3339	883	995	28	0	0

- Molecule 3 is a protein called HDA1 complex subunit 3,HDA1 complex subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	M	534	4376	2765	737	854	20	0	0

- Molecule 4 is a protein called HDA1 complex subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	N	541	4416	2801	759	843	13	0	0

- Molecule 5 is a protein called Histone H3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	A	97	802	506	155	138	3	0	0

- Molecule 6 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	B	83	662	418	129	114	1	0	0

- Molecule 7 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	C	103	Total	C	N	O	0	0
			795	501	155	139		

- Molecule 8 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	D	95	Total	C	N	O	S	0	0
			745	469	134	140	2		

- Molecule 9 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	E	97	Total	C	N	O	S	0	0
			801	504	155	139	3		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	102	ALA	GLY	conflict	UNP P84233

- Molecule 10 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	F	78	Total	C	N	O	S	0	0
			619	391	120	107	1		

- Molecule 11 is a protein called Histone H2A type 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
11	G	105	Total	C	N	O	0	0
			809	510	158	141		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	99	ARG	GLY	conflict	UNP P06897

- Molecule 12 is a protein called Histone H2B.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	H	93	Total	C	N	O	S	0	0
			726	457	130	137	2		

- Molecule 13 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
13	I	145	Total	C	N	O	P	0	0
			2952	1404	537	867	144		

- Molecule 14 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
14	J	145	Total	C	N	O	P	0	0
			2987	1416	558	869	144		

- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

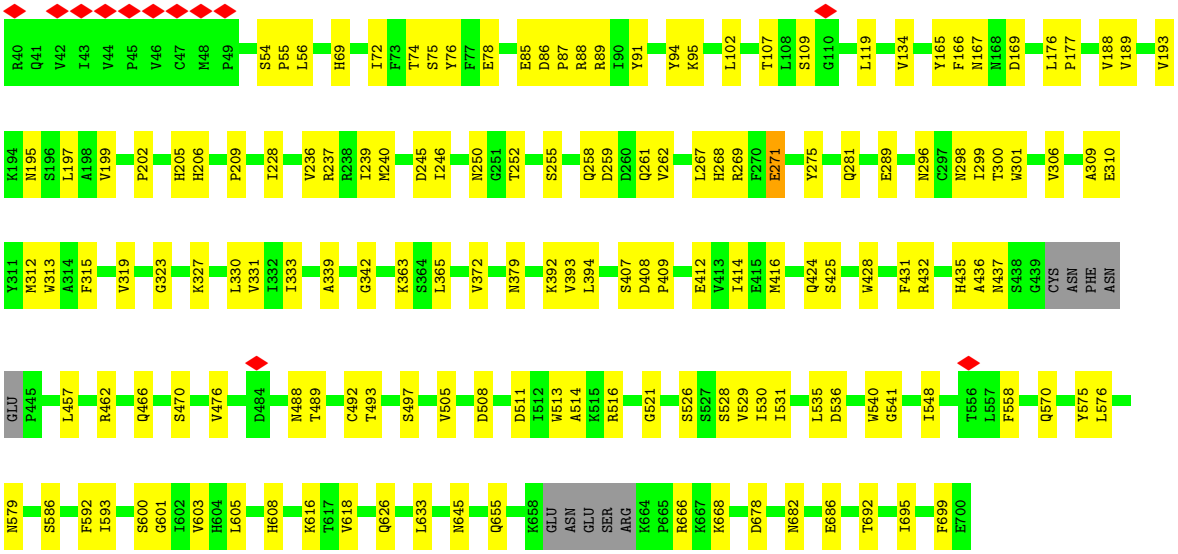
Mol	Chain	Residues	Atoms		AltConf
15	K	1	Total	Zn	0
			1	1	
15	L	1	Total	Zn	0
			1	1	

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.

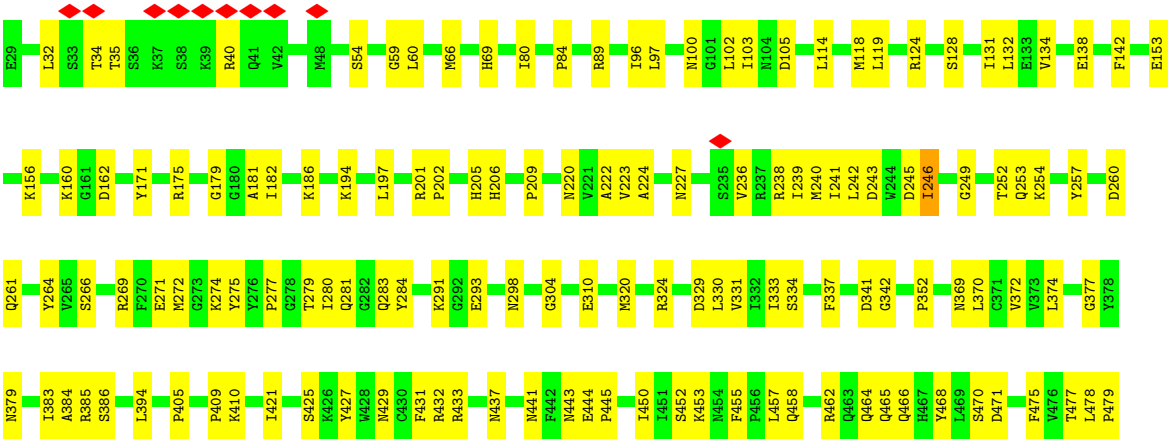
• Molecule 1: Histone deacetylase HDA1

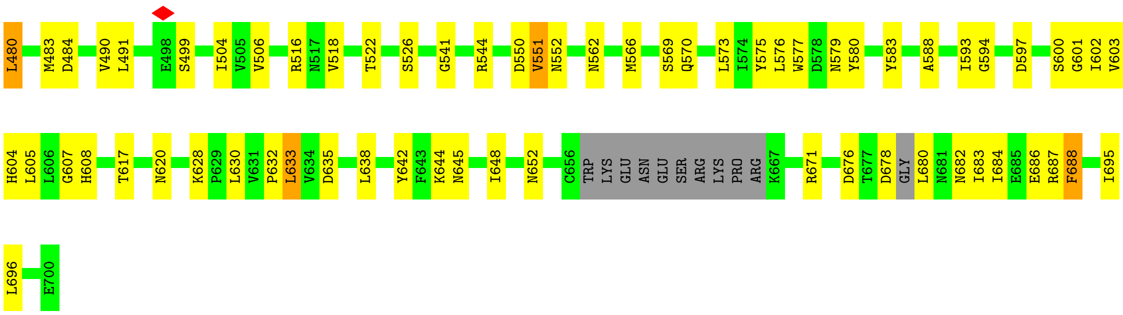
Chain K:



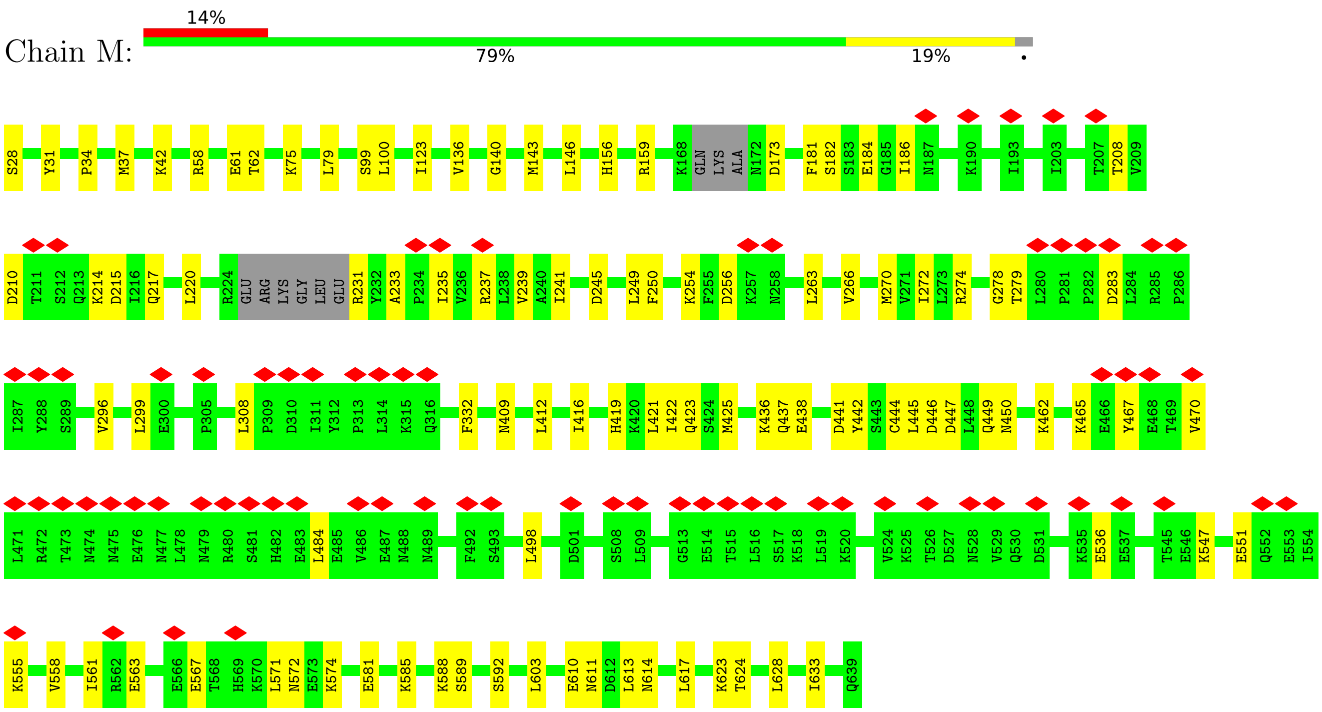
• Molecule 2: Histone deacetylase HDA1

Chain L:

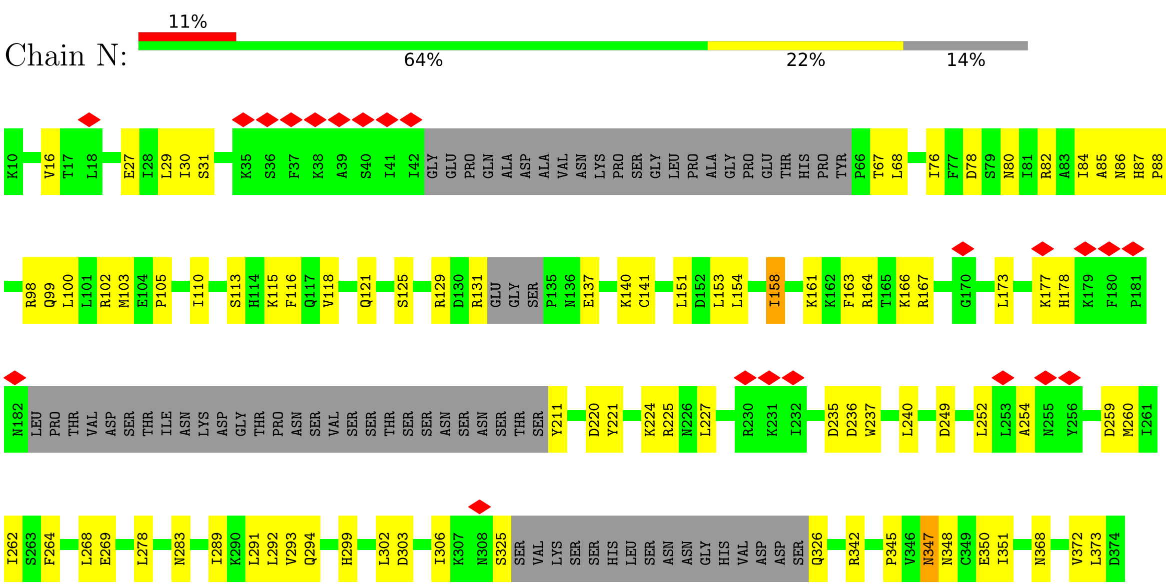


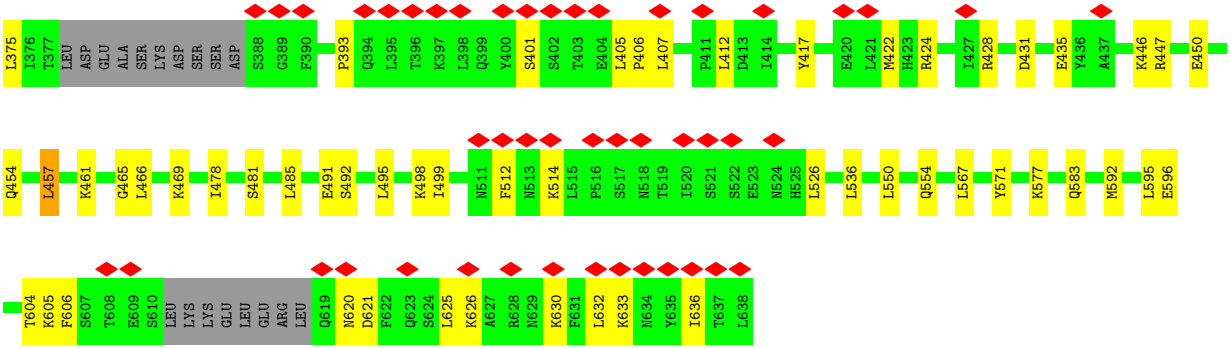


• Molecule 3: HDA1 complex subunit 3,HDA1 complex subunit 3



• Molecule 4: HDA1 complex subunit 2



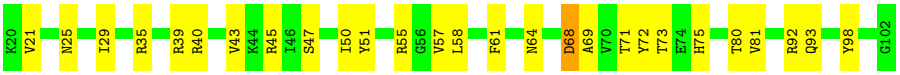


- Molecule 5: Histone H3

Chain A:

- Molecule 6: Histone H4

Chain B:

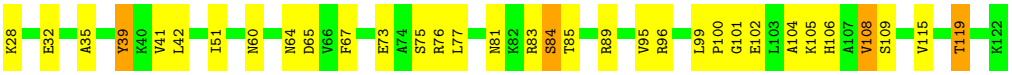


- Molecule 7: Histone H2A

Chain C:

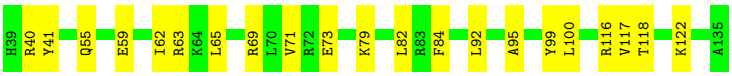
- Molecule 8: Histone H2B

Chain D:



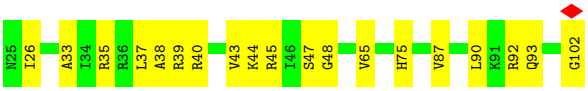
- Molecule 9: Histone H3.2

Chain E:



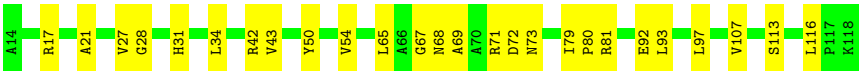
- Molecule 10: Histone H4

Chain F:



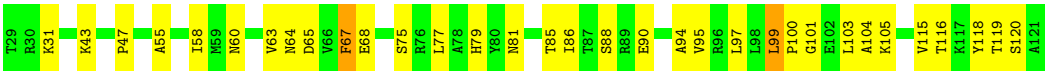
• Molecule 11: Histone H2A type 1

Chain G:



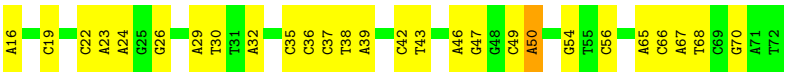
• Molecule 12: Histone H2B

Chain H:



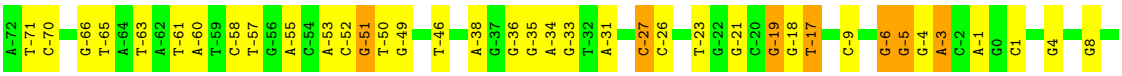
• Molecule 13: DNA (145-MER)

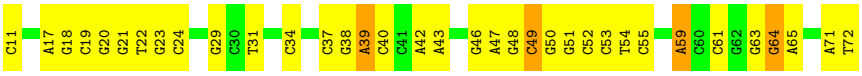
Chain I:



• Molecule 14: DNA (145-MER)

Chain J:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	41279	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	77.2	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.271	Depositor
Minimum map value	-0.061	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.008	Depositor
Recommended contour level	0.035	Depositor
Map size (Å)	459.64798, 459.64798, 459.64798	wwPDB
Map dimensions	540, 540, 540	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.8512, 0.8512, 0.8512	Depositor

5 Model quality

5.1 Standard geometry

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

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	K	0.45	0/5301	0.83	6/7189 (0.1%)
2	L	0.44	0/5366	0.84	14/7277 (0.2%)
3	M	0.39	0/4449	0.79	6/5997 (0.1%)
4	N	0.39	0/4486	0.84	12/6037 (0.2%)
5	A	0.43	0/814	0.82	2/1092 (0.2%)
6	B	0.42	0/669	0.81	0/894
7	C	0.40	0/805	0.73	0/1088
8	D	0.48	0/756	0.92	4/1015 (0.4%)
9	E	0.39	0/812	0.79	1/1088 (0.1%)
10	F	0.42	0/626	0.83	0/837
11	G	0.40	0/819	0.77	0/1106
12	H	0.43	0/737	0.84	3/993 (0.3%)
13	I	1.24	10/3308 (0.3%)	1.47	54/5099 (1.1%)
14	J	1.24	6/3354 (0.2%)	1.50	78/5180 (1.5%)
All	All	0.68	16/32302 (0.0%)	1.01	180/44892 (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 mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	K	0	3
2	L	0	5
3	M	0	3
4	N	0	7
6	B	0	2
7	C	0	1
8	D	0	1
10	F	0	1
All	All	0	23

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	J	-55	DA	C3'-O3'	-8.28	1.33	1.44
13	I	66	DC	C3'-O3'	6.41	1.52	1.44
13	I	-63	DC	C3'-O3'	5.99	1.51	1.44
14	J	64	DG	C3'-O3'	5.99	1.51	1.44
13	I	-48	DC	N1-C6	-5.80	1.33	1.37

The worst 5 of 180 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	I	-35	DA	O4'-C1'-N9	-11.78	99.75	108.00
8	D	65	ASP	CB-CG-OD1	10.11	127.40	118.30
14	J	-63	DT	O4'-C1'-N1	9.11	114.38	108.00
14	J	46	DG	OP1-P-O3'	9.10	125.23	105.20
13	I	-53	DG	OP2-P-O3'	8.54	123.98	105.20

There are no chirality outliers.

5 of 23 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	K	271	GLU	Peptide
1	K	408	ASP	Peptide
1	K	54	SER	Peptide
2	L	246	ILE	Peptide
2	L	54	SER	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	K	5178	0	5104	101	0
2	L	5245	0	5155	130	0
3	M	4376	0	4380	64	0
4	N	4416	0	4493	75	0
5	A	802	0	841	24	0
6	B	662	0	709	21	0
7	C	795	0	846	25	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	D	745	0	773	22	0
9	E	801	0	838	16	0
10	F	619	0	659	15	0
11	G	809	0	864	15	0
12	H	726	0	747	22	0
13	I	2952	0	1629	25	0
14	J	2987	0	1630	19	0
15	K	1	0	0	0	0
15	L	1	0	0	0	0
All	All	31115	0	28668	491	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 10.

The worst 5 of 491 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)
13:I:3:DT:H3	14:J:-4:DG:H1	1.12	0.92
2:L:575:TYR:O	2:L:579:ASN:HB2	1.76	0.85
2:L:242:LEU:O	2:L:334:SER:HB3	1.76	0.84
2:L:156:LYS:O	2:L:160:LYS:HB2	1.81	0.81
3:M:28:SER:N	3:M:31:TYR:HH	1.80	0.80

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	K	645/661 (98%)	585 (91%)	60 (9%)	0	100	100
2	L	655/672 (98%)	590 (90%)	65 (10%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	M	526/543 (97%)	499 (95%)	27 (5%)	0	100	100
4	N	527/629 (84%)	485 (92%)	41 (8%)	1 (0%)	47	81
5	A	95/97 (98%)	92 (97%)	3 (3%)	0	100	100
6	B	81/83 (98%)	75 (93%)	6 (7%)	0	100	100
7	C	101/103 (98%)	96 (95%)	5 (5%)	0	100	100
8	D	93/95 (98%)	87 (94%)	6 (6%)	0	100	100
9	E	95/97 (98%)	92 (97%)	3 (3%)	0	100	100
10	F	76/78 (97%)	73 (96%)	3 (4%)	0	100	100
11	G	103/105 (98%)	98 (95%)	5 (5%)	0	100	100
12	H	91/93 (98%)	84 (92%)	7 (8%)	0	100	100
All	All	3088/3256 (95%)	2856 (92%)	231 (8%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	N	87	HIS

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	K	575/585 (98%)	575 (100%)	0	100	100
2	L	586/596 (98%)	585 (100%)	1 (0%)	93	96
3	M	504/511 (99%)	502 (100%)	2 (0%)	91	94
4	N	504/582 (87%)	504 (100%)	0	100	100
5	A	85/85 (100%)	82 (96%)	3 (4%)	36	60
6	B	68/68 (100%)	66 (97%)	2 (3%)	42	64
7	C	82/82 (100%)	80 (98%)	2 (2%)	49	69
8	D	81/81 (100%)	76 (94%)	5 (6%)	18	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
9	E	84/84 (100%)	82 (98%)	2 (2%)	49 69
10	F	63/63 (100%)	62 (98%)	1 (2%)	62 79
11	G	83/83 (100%)	81 (98%)	2 (2%)	49 69
12	H	79/79 (100%)	77 (98%)	2 (2%)	47 68
All	All	2794/2899 (96%)	2772 (99%)	22 (1%)	82 89

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	D	119	THR
10	F	26	ILE
9	E	117	VAL
11	G	50	TYR
6	B	21	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
4	N	368	ASN
4	N	620	ASN
4	N	454	GLN
4	N	518	ASN
5	A	108	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 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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
3	M	1
4	N	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	M	334:LYS	C	404:ASN	N	81.01
1	N	308:ASN	C	325:SER	N	26.70

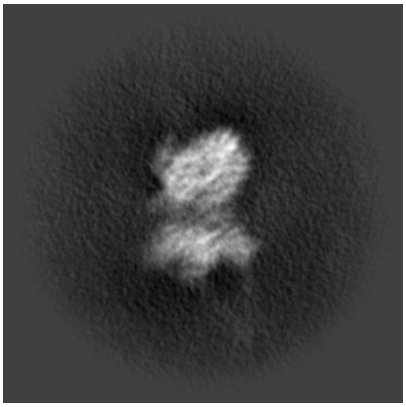
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-11102. These allow visual inspection of the internal detail of the map and identification of artifacts.

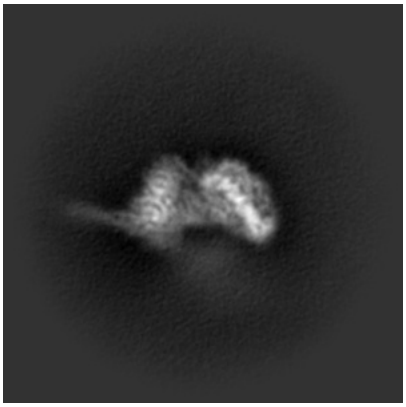
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

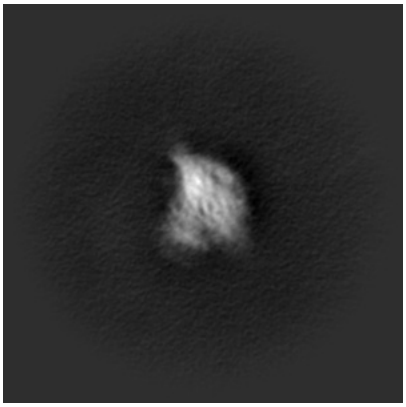
6.1.1 Primary map



X



Y

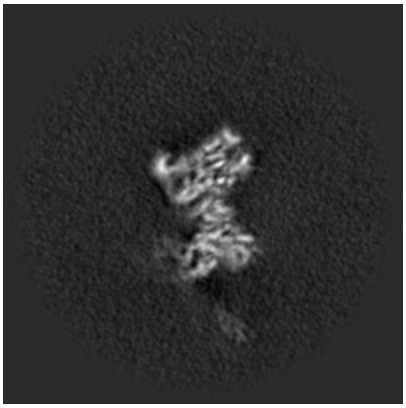


Z

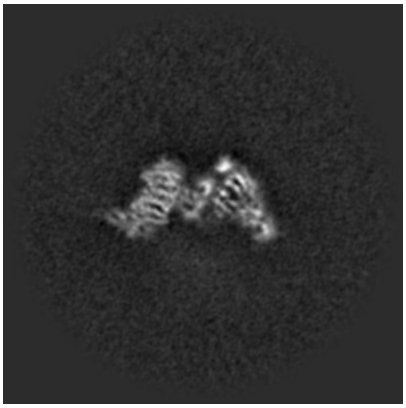
The images above show the map projected in three orthogonal directions.

6.2 Central slices

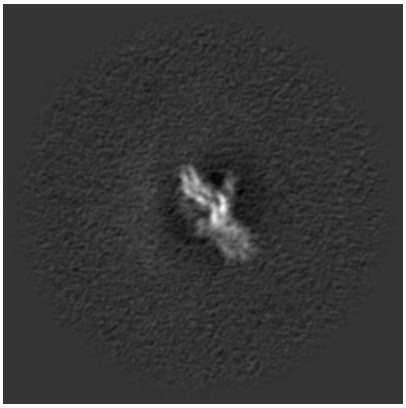
6.2.1 Primary map



X Index: 270



Y Index: 270

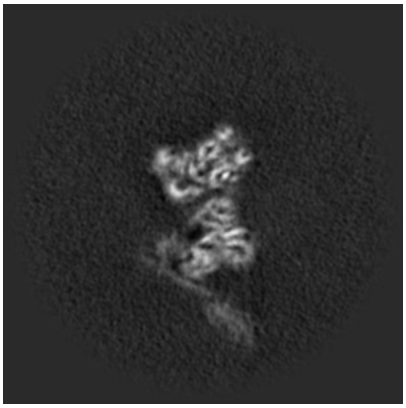


Z Index: 270

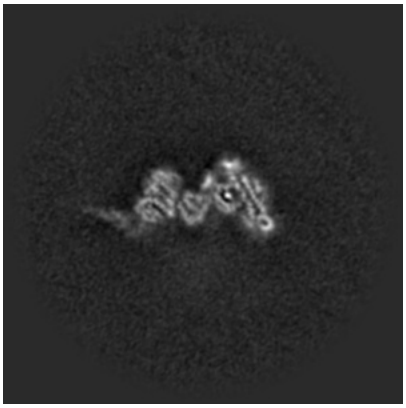
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

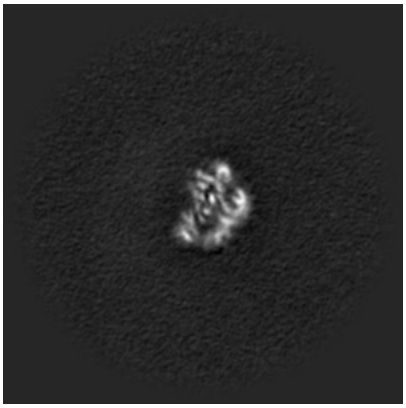
6.3.1 Primary map



X Index: 261



Y Index: 280

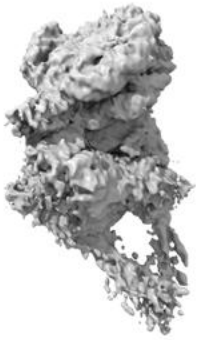


Z Index: 301

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

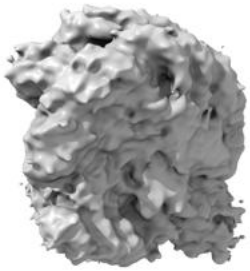
6.4 Orthogonal surface views (i)

6.4.1 Primary map



X

Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.035. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

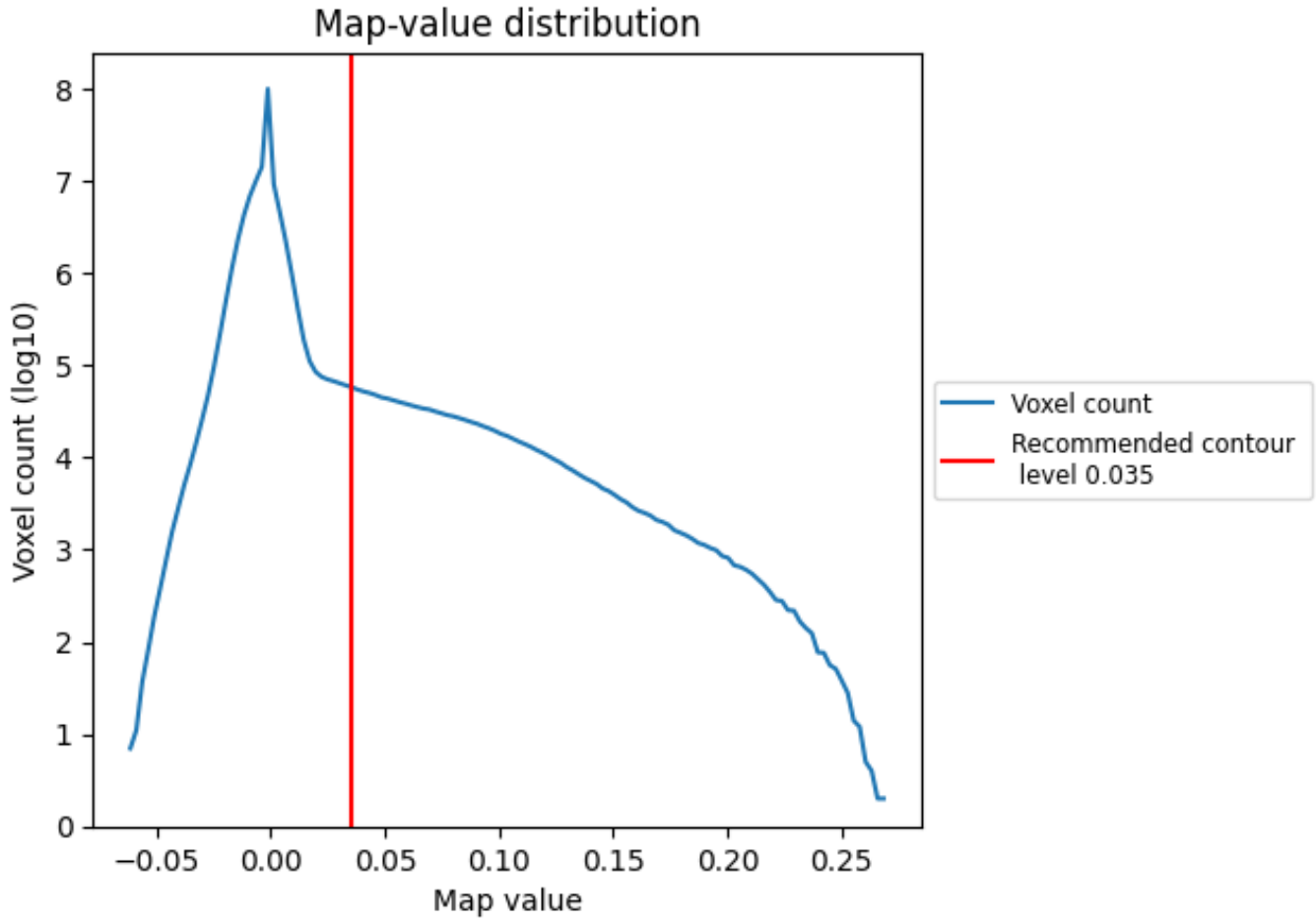
6.5 Mask visualisation (i)

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

7 Map analysis

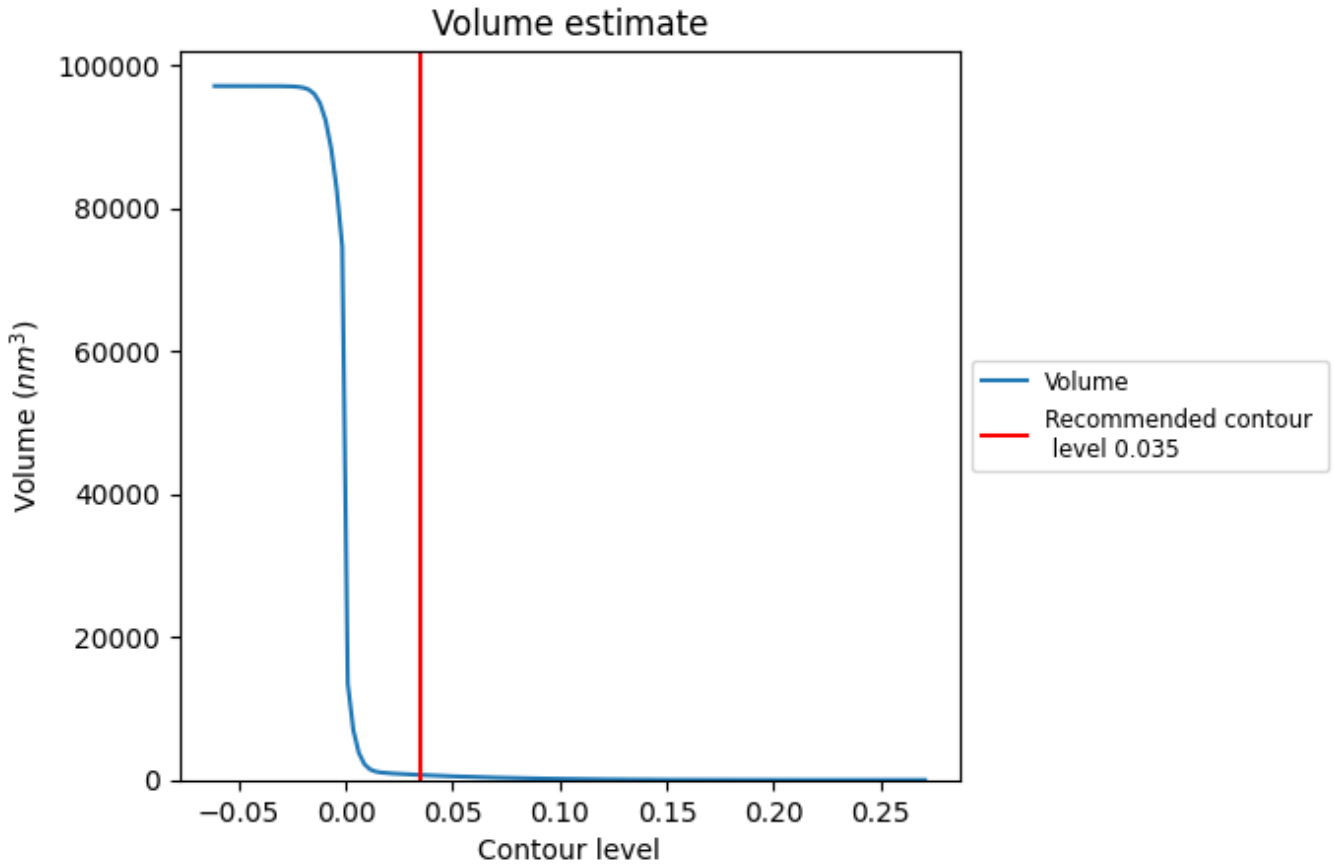
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

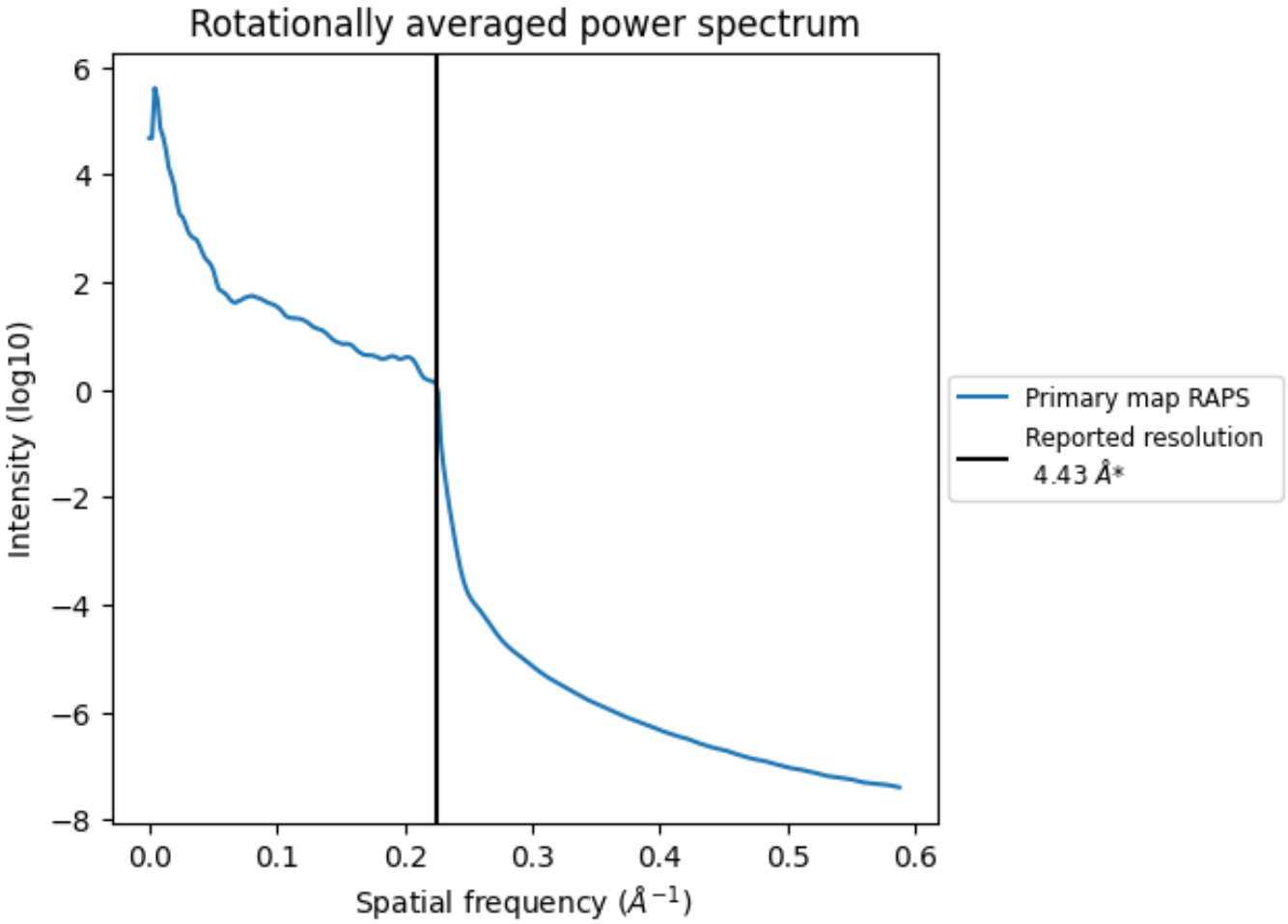
7.2 Volume estimate (i)



The volume at the recommended contour level is 699 nm^3; this corresponds to an approximate mass of 632 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 0.226 Å^{-1}

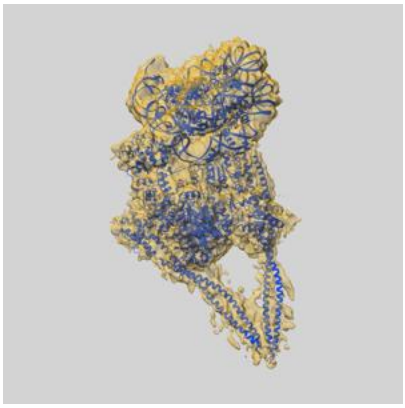
8 Fourier-Shell correlation

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

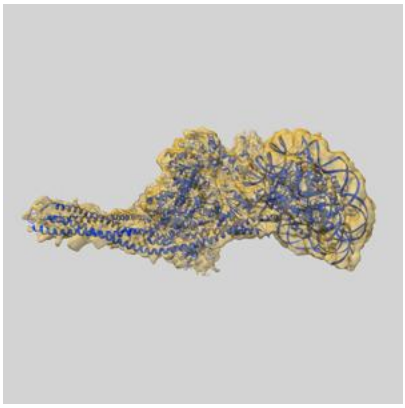
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-11102 and PDB model 6Z6P. Per-residue inclusion information can be found in section 3 on page 7.

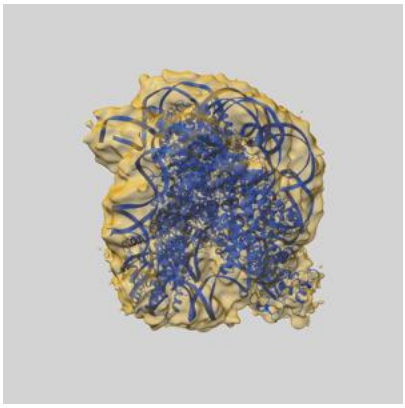
9.1 Map-model overlay



X



Y



Z

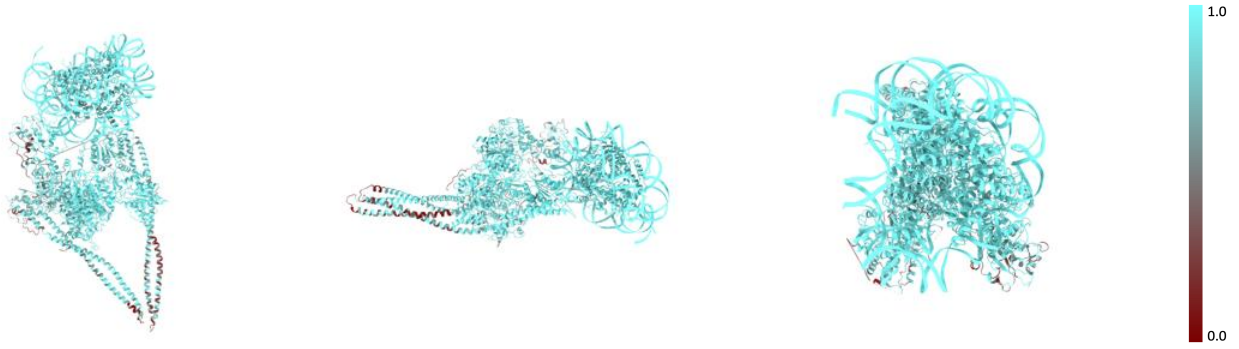
The images above show the 3D surface view of the map at the recommended contour level 0.035 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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.

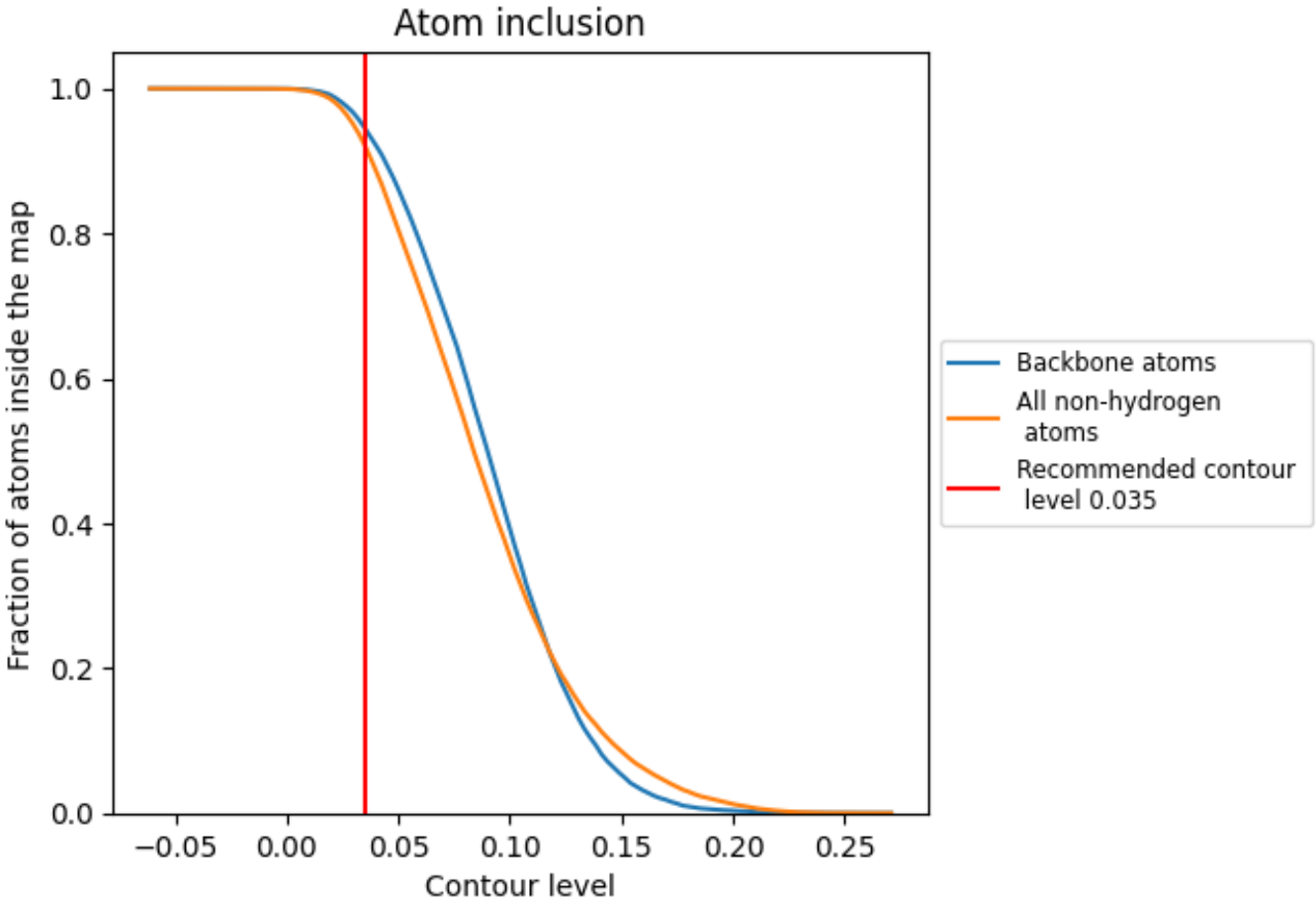
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.035).

9.4 Atom inclusion (i)



At the recommended contour level, 94% of all backbone atoms, 92% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary

The table lists the average atom inclusion at the recommended contour level (0.035) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.9209	0.1890
A	0.9922	0.2030
B	0.9858	0.2180
C	0.9857	0.2120
D	0.9835	0.2110
E	0.9649	0.1990
F	0.9445	0.1960
G	0.9783	0.2200
H	0.9690	0.2090
I	0.9997	0.2250
J	1.0000	0.2280
K	0.9536	0.1840
L	0.9465	0.1900
M	0.7808	0.1500
N	0.8098	0.1530