wwPDB EM Validation Summary Report

Aug 12, 2024 – 09:49 am BST

PDB ID	:	8QH5
EMDB ID	:	EMD-18398
Title	:	CryoEM structure of UVSSA(VHS)-CSA-DDB1-DDA1
Authors	:	Lee, S.-H.; Sixma, T.K.
Deposited on	:	2023-09-06
Resolution	:	3.40 Å(reported)
Based on initial models	:	7OO3, ., 6ZX9, 6UD7

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 ⓘ 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 ⓘ) were used in the production of this report:

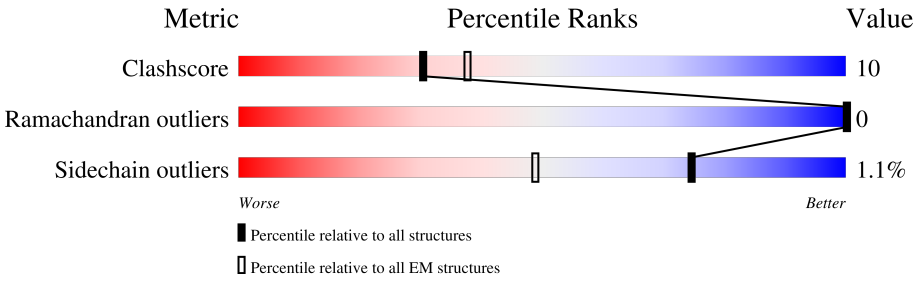
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.40 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	D	729	16% 16% · 79%
2	A	408	· 70% 17% · 12%
3	B	1160	9% 77% 20% · ·
4	C	152	5% 27% 7% 66%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 22744 atoms, of which 9435 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 UV-stimulated scaffold protein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	D	150	1230	779	220	227	4	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-19	MET	-	initiating methionine	UNP Q2YD98
D	-18	ALA	-	expression tag	UNP Q2YD98
D	-17	HIS	-	expression tag	UNP Q2YD98
D	-16	HIS	-	expression tag	UNP Q2YD98
D	-15	HIS	-	expression tag	UNP Q2YD98
D	-14	HIS	-	expression tag	UNP Q2YD98
D	-13	HIS	-	expression tag	UNP Q2YD98
D	-12	HIS	-	expression tag	UNP Q2YD98
D	-11	SER	-	expression tag	UNP Q2YD98
D	-10	ALA	-	expression tag	UNP Q2YD98
D	-9	ALA	-	expression tag	UNP Q2YD98
D	-8	LEU	-	expression tag	UNP Q2YD98
D	-7	GLU	-	expression tag	UNP Q2YD98
D	-6	VAL	-	expression tag	UNP Q2YD98
D	-5	LEU	-	expression tag	UNP Q2YD98
D	-4	PHE	-	expression tag	UNP Q2YD98
D	-3	GLN	-	expression tag	UNP Q2YD98
D	-2	GLY	-	expression tag	UNP Q2YD98
D	-1	PRO	-	expression tag	UNP Q2YD98
D	0	GLY	-	expression tag	UNP Q2YD98

- Molecule 2 is a protein called DNA excision repair protein ERCC-8.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	H	N	O	S		
2	A	359	5546	1754	2736	499	538	19	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	397	GLY	-	expression tag	UNP Q13216
A	398	THR	-	expression tag	UNP Q13216
A	399	SER	-	expression tag	UNP Q13216
A	400	ALA	-	expression tag	UNP Q13216
A	401	TRP	-	expression tag	UNP Q13216
A	402	SER	-	expression tag	UNP Q13216
A	403	HIS	-	expression tag	UNP Q13216
A	404	PRO	-	expression tag	UNP Q13216
A	405	GLN	-	expression tag	UNP Q13216
A	406	PHE	-	expression tag	UNP Q13216
A	407	GLU	-	expression tag	UNP Q13216
A	408	LYS	-	expression tag	UNP Q13216

- Molecule 3 is a protein called DNA damage-binding protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	B	1126	Total	C	H	N	O	S	0	0
			15083	5593	6254	1488	1699	49		

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	initiating methionine	UNP Q16531
B	-18	ALA	-	expression tag	UNP Q16531
B	-17	HIS	-	expression tag	UNP Q16531
B	-16	HIS	-	expression tag	UNP Q16531
B	-15	HIS	-	expression tag	UNP Q16531
B	-14	HIS	-	expression tag	UNP Q16531
B	-13	HIS	-	expression tag	UNP Q16531
B	-12	HIS	-	expression tag	UNP Q16531
B	-11	SER	-	expression tag	UNP Q16531
B	-10	ALA	-	expression tag	UNP Q16531
B	-9	ALA	-	expression tag	UNP Q16531
B	-8	LEU	-	expression tag	UNP Q16531
B	-7	GLU	-	expression tag	UNP Q16531
B	-6	VAL	-	expression tag	UNP Q16531
B	-5	LEU	-	expression tag	UNP Q16531
B	-4	PHE	-	expression tag	UNP Q16531
B	-3	GLN	-	expression tag	UNP Q16531
B	-2	GLY	-	expression tag	UNP Q16531
B	-1	PRO	-	expression tag	UNP Q16531

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLY	-	expression tag	UNP Q16531

- Molecule 4 is a protein called DET1- and DDB1-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	C	52	Total	C	H	N	O	0	0
			885	285	445	79	76		

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	103	ASP	-	expression tag	UNP Q9BW61
C	104	VAL	-	expression tag	UNP Q9BW61
C	105	LEU	-	expression tag	UNP Q9BW61
C	106	PHE	-	expression tag	UNP Q9BW61
C	107	GLN	-	expression tag	UNP Q9BW61
C	108	GLY	-	expression tag	UNP Q9BW61
C	109	PRO	-	expression tag	UNP Q9BW61
C	110	GLY	-	expression tag	UNP Q9BW61
C	111	ALA	-	expression tag	UNP Q9BW61
C	112	TRP	-	expression tag	UNP Q9BW61
C	113	SER	-	expression tag	UNP Q9BW61
C	114	HIS	-	expression tag	UNP Q9BW61
C	115	PRO	-	expression tag	UNP Q9BW61
C	116	GLN	-	expression tag	UNP Q9BW61
C	117	PHE	-	expression tag	UNP Q9BW61
C	118	GLU	-	expression tag	UNP Q9BW61
C	119	LYS	-	expression tag	UNP Q9BW61
C	120	GLY	-	expression tag	UNP Q9BW61
C	121	GLY	-	expression tag	UNP Q9BW61
C	122	GLY	-	expression tag	UNP Q9BW61
C	123	SER	-	expression tag	UNP Q9BW61
C	124	GLY	-	expression tag	UNP Q9BW61
C	125	GLY	-	expression tag	UNP Q9BW61
C	126	GLY	-	expression tag	UNP Q9BW61
C	127	SER	-	expression tag	UNP Q9BW61
C	128	GLY	-	expression tag	UNP Q9BW61
C	129	GLY	-	expression tag	UNP Q9BW61
C	130	GLY	-	expression tag	UNP Q9BW61
C	131	SER	-	expression tag	UNP Q9BW61
C	132	TRP	-	expression tag	UNP Q9BW61
C	133	SER	-	expression tag	UNP Q9BW61

Continued on next page...

Continued from previous page...

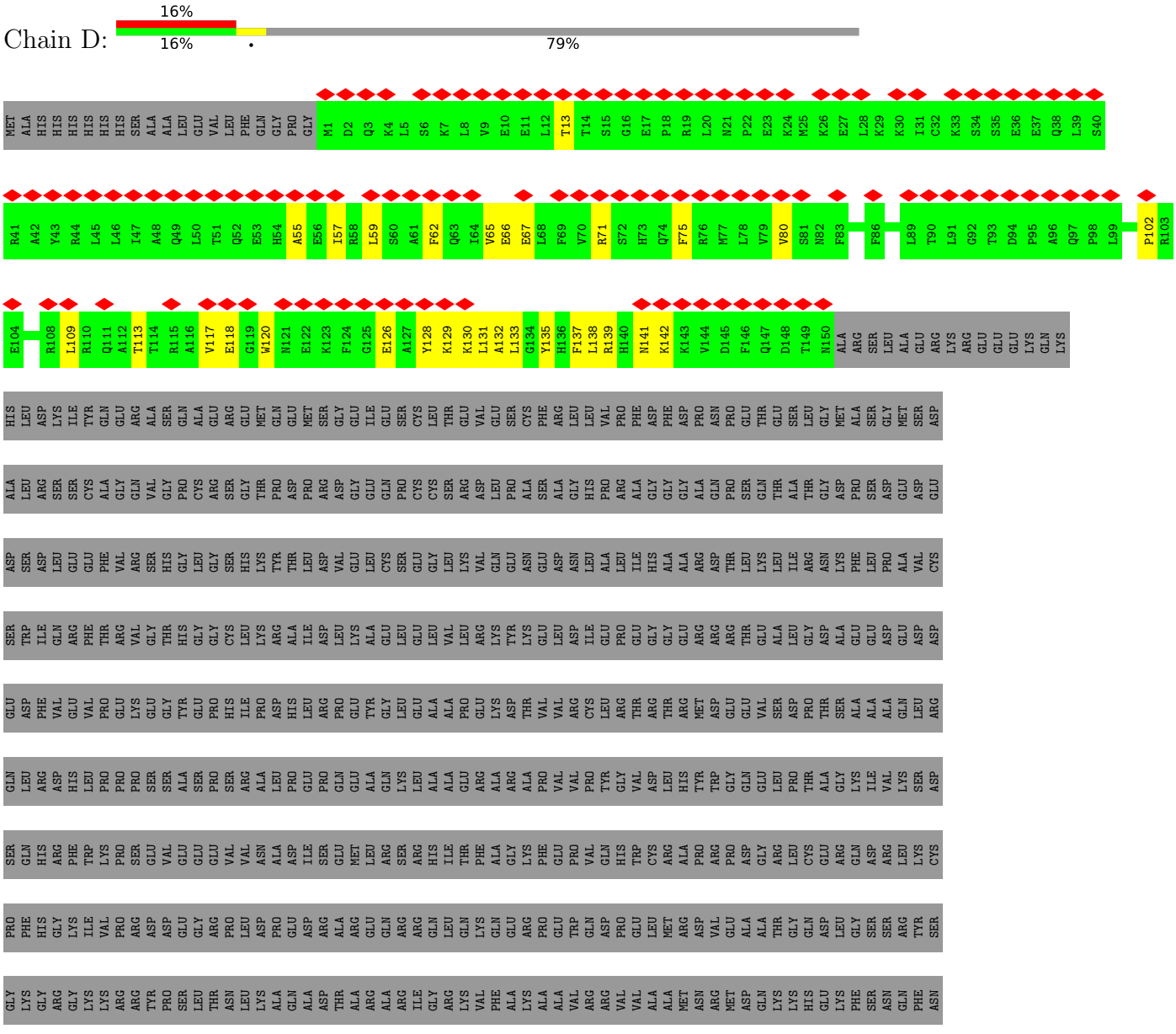
Chain	Residue	Modelled	Actual	Comment	Reference
C	134	HIS	-	expression tag	UNP Q9BW61
C	135	PRO	-	expression tag	UNP Q9BW61
C	136	GLN	-	expression tag	UNP Q9BW61
C	137	PHE	-	expression tag	UNP Q9BW61
C	138	GLU	-	expression tag	UNP Q9BW61
C	139	LYS	-	expression tag	UNP Q9BW61
C	140	GLY	-	expression tag	UNP Q9BW61
C	141	ALA	-	expression tag	UNP Q9BW61
C	142	SER	-	expression tag	UNP Q9BW61
C	143	GLY	-	expression tag	UNP Q9BW61
C	144	GLU	-	expression tag	UNP Q9BW61
C	145	ASP	-	expression tag	UNP Q9BW61
C	146	TYR	-	expression tag	UNP Q9BW61
C	147	LYS	-	expression tag	UNP Q9BW61
C	148	ASP	-	expression tag	UNP Q9BW61
C	149	ASP	-	expression tag	UNP Q9BW61
C	150	ASP	-	expression tag	UNP Q9BW61
C	151	ASP	-	expression tag	UNP Q9BW61
C	152	LYS	-	expression tag	UNP Q9BW61

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.

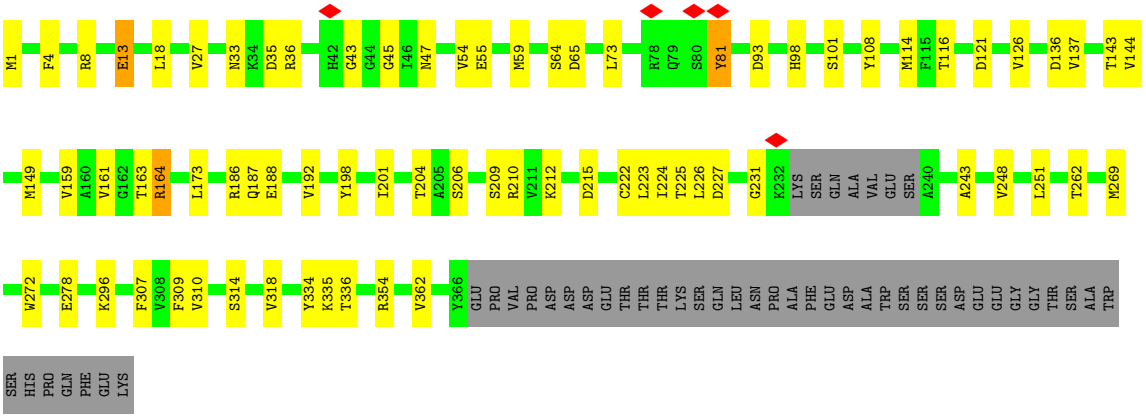
- Molecule 1: UV-stimulated scaffold protein A

Chain D:



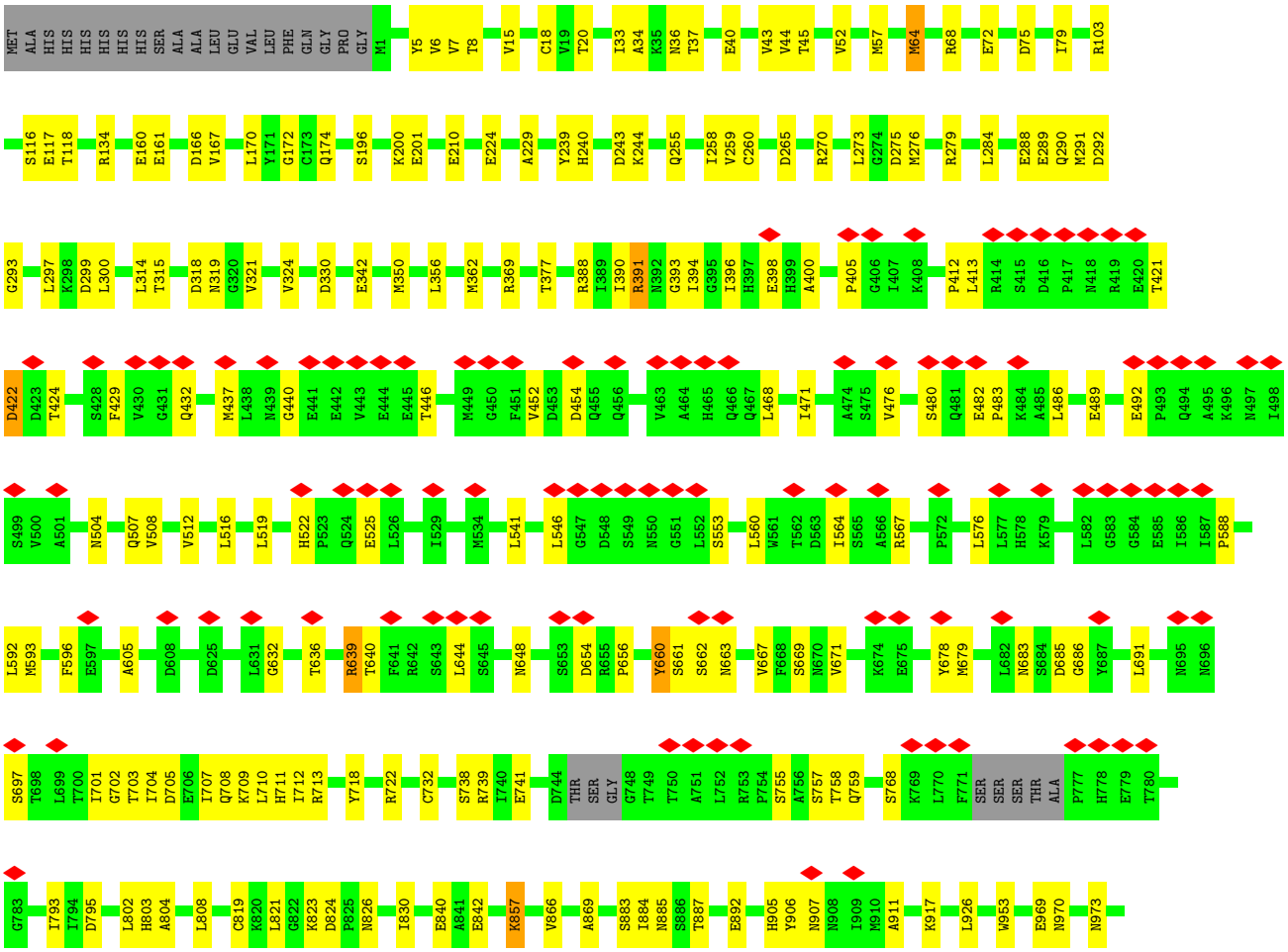
• Molecule 2: DNA excision repair protein ERCC-8

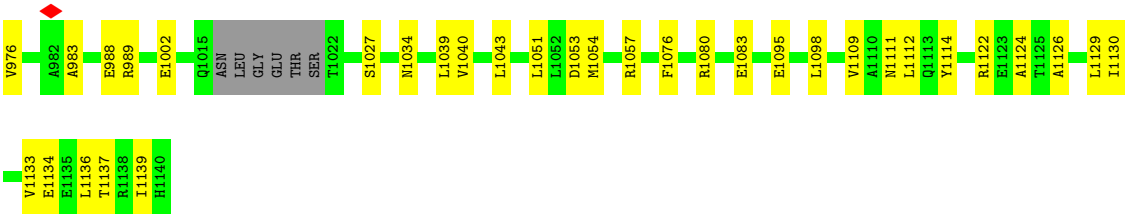
Chain A:



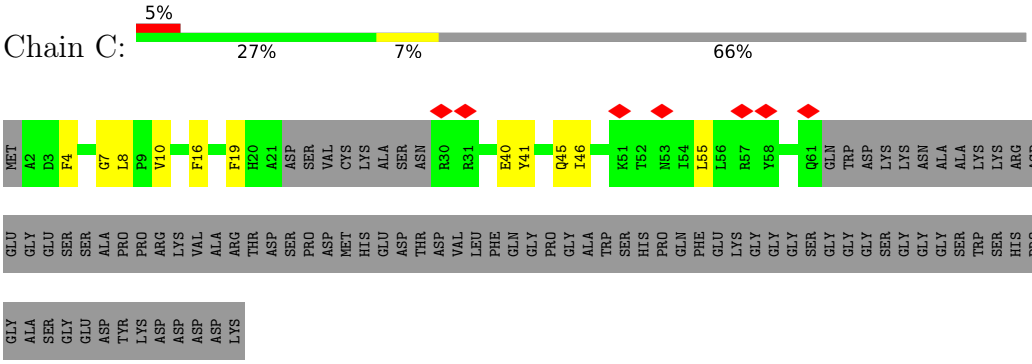
• Molecule 3: DNA damage-binding protein 1

Chain B:





• Molecule 4: DET1- and DDB1-associated protein 1



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	294142	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 ($e^-/Å^2$)	60	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	81000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	85.662	Depositor
Minimum map value	-1.899	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.948	Depositor
Recommended contour level	0.828	Depositor
Map size (Å)	338.24, 338.24, 338.24	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.0569999, 1.0569999, 1.0569999	Depositor

5 Model quality

5.1 Standard geometry

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	D	0.23	0/1253	0.43	0/1685
2	A	0.29	0/2869	0.54	0/3886
3	B	0.30	0/8991	0.52	2/12175 (0.0%)
4	C	0.28	0/451	0.59	0/609
All	All	0.29	0/13564	0.52	2/18355 (0.0%)

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
2	A	0	4
3	B	0	6
All	All	0	10

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	685	ASP	CB-CG-OD2	5.26	123.03	118.30
3	B	422	ASP	CB-CG-OD2	5.09	122.88	118.30

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	164	ARG	Sidechain
2	A	186	ARG	Sidechain
2	A	210	ARG	Sidechain
2	A	354	ARG	Sidechain
3	B	369	ARG	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	D	1230	0	1248	32	0
2	A	2810	2736	2735	57	0
3	B	8829	6254	8799	177	0
4	C	440	445	443	14	0
All	All	13309	9435	13225	263	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 263 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)
3:B:471:ILE:HG23	3:B:476:VAL:HG22	1.23	1.12
3:B:656:PRO:HB2	3:B:671:VAL:HB	1.45	0.98
3:B:1109:VAL:HG12	3:B:1129:LEU:HD12	1.47	0.95
3:B:398:GLU:HA	3:B:702:GLY:HA3	1.52	0.89
3:B:18:CYS:SG	3:B:315:THR:OG1	2.27	0.89

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	D	148/729 (20%)	147 (99%)	1 (1%)	0	100	100
2	A	355/408 (87%)	335 (94%)	20 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	B	1118/1160 (96%)	1055 (94%)	63 (6%)	0	100	100
4	C	48/152 (32%)	47 (98%)	1 (2%)	0	100	100
All	All	1669/2449 (68%)	1584 (95%)	85 (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	D	134/623 (22%)	134 (100%)	0	100	100
2	A	315/358 (88%)	309 (98%)	6 (2%)	57	78
3	B	988/1014 (97%)	978 (99%)	10 (1%)	76	88
4	C	49/128 (38%)	49 (100%)	0	100	100
All	All	1486/2123 (70%)	1470 (99%)	16 (1%)	74	86

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	917	LYS
3	B	857	LYS
3	B	356	LEU
3	B	819	CYS
3	B	260	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	796	GLN
3	B	952	ASN
4	C	20	HIS
3	B	455	GLN
3	B	465	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

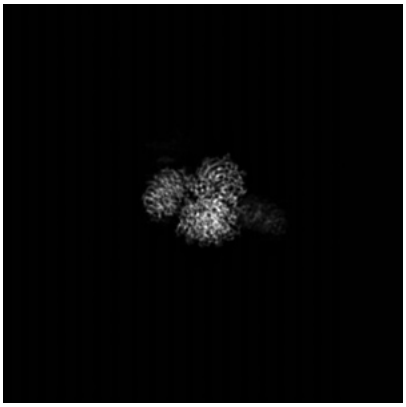
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-18398. These allow visual inspection of the internal detail of the map and identification of artifacts.

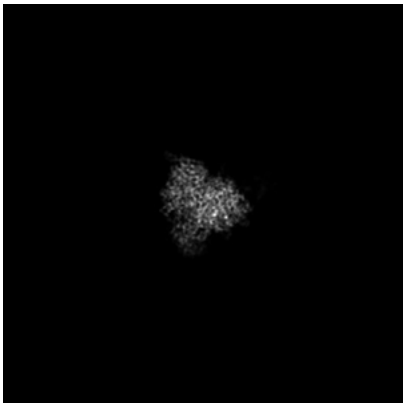
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

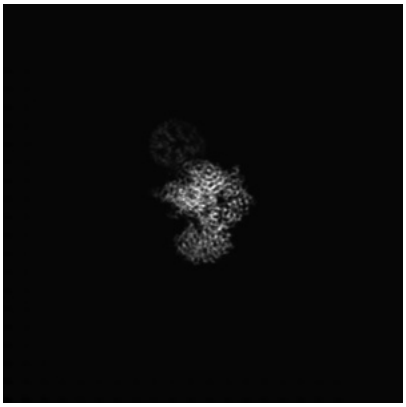
6.1.1 Primary map



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: 160

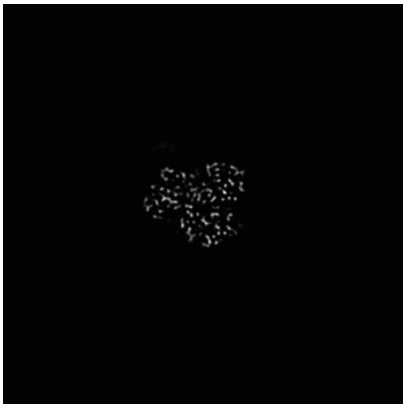
Y Index: 160

Z Index: 160

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: 166

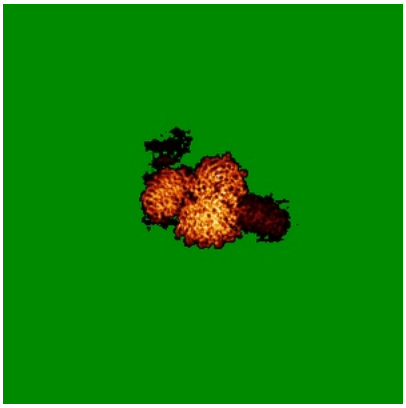
Y Index: 168

Z Index: 157

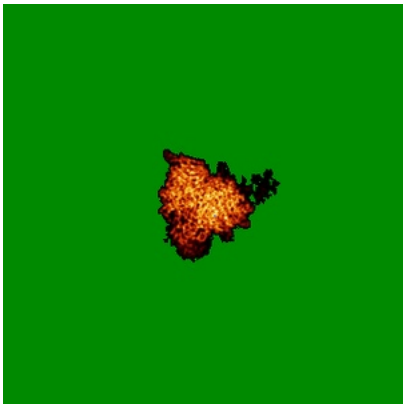
The images above show the largest variance slices of the map in three orthogonal directions.

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

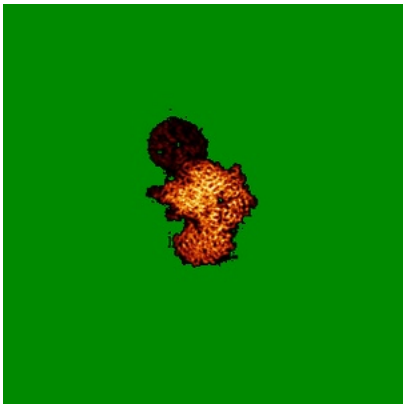
6.4.1 Primary map



X



Y

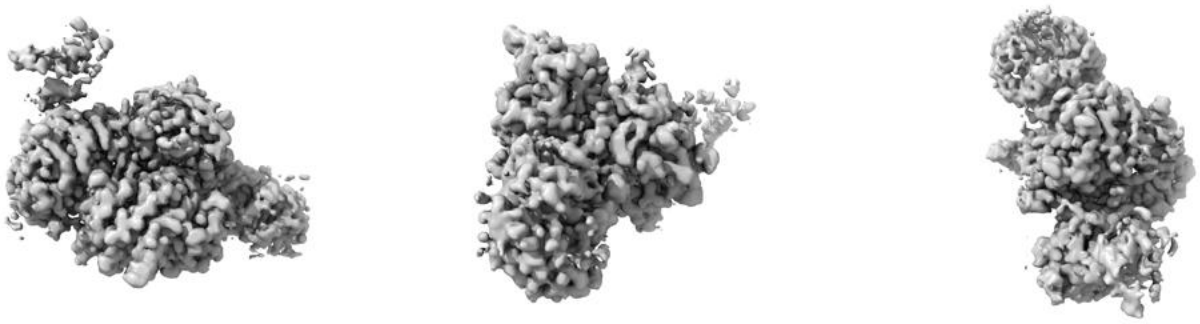


Z

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

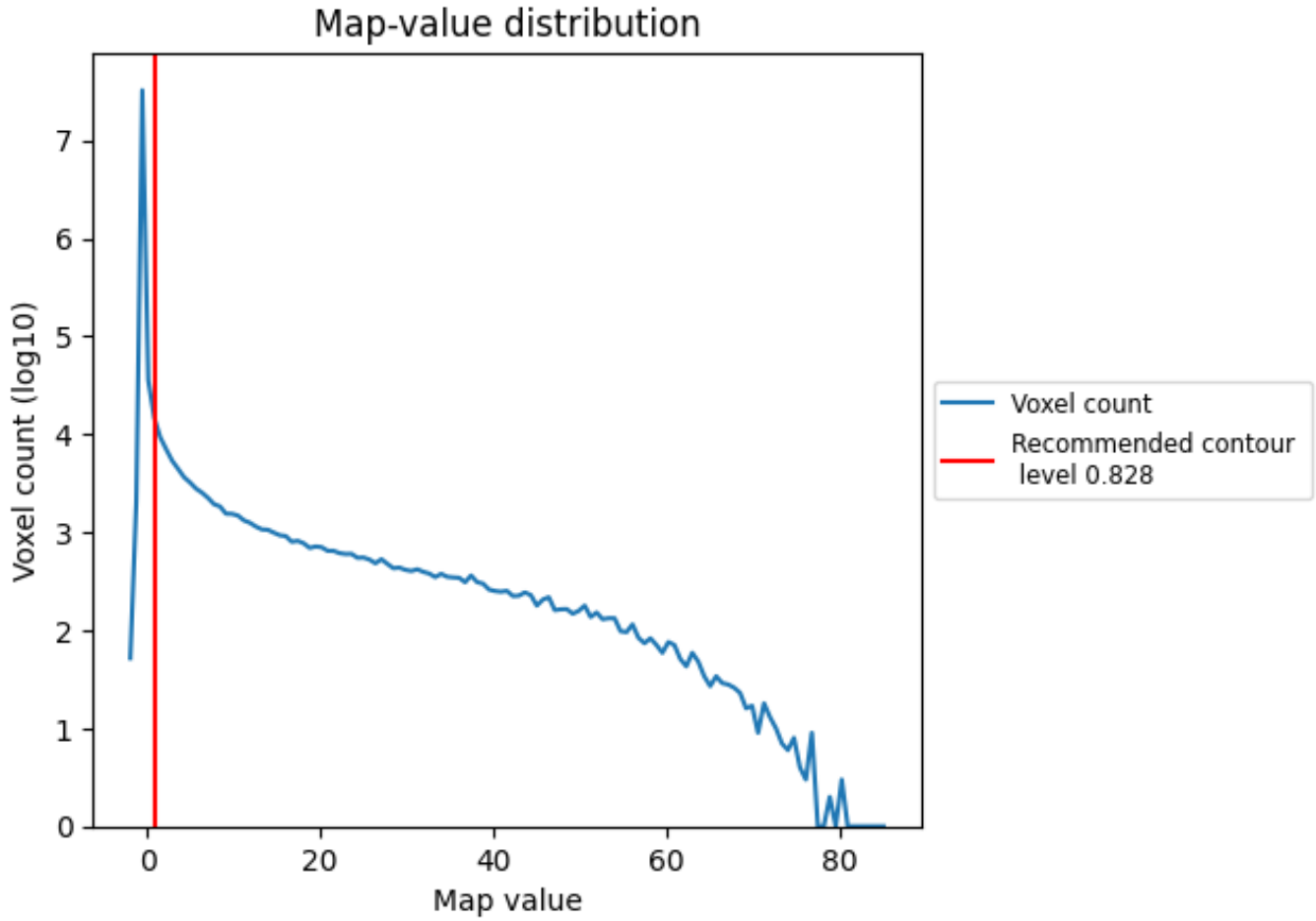
6.6 Mask visualisation (i)

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

7 Map analysis

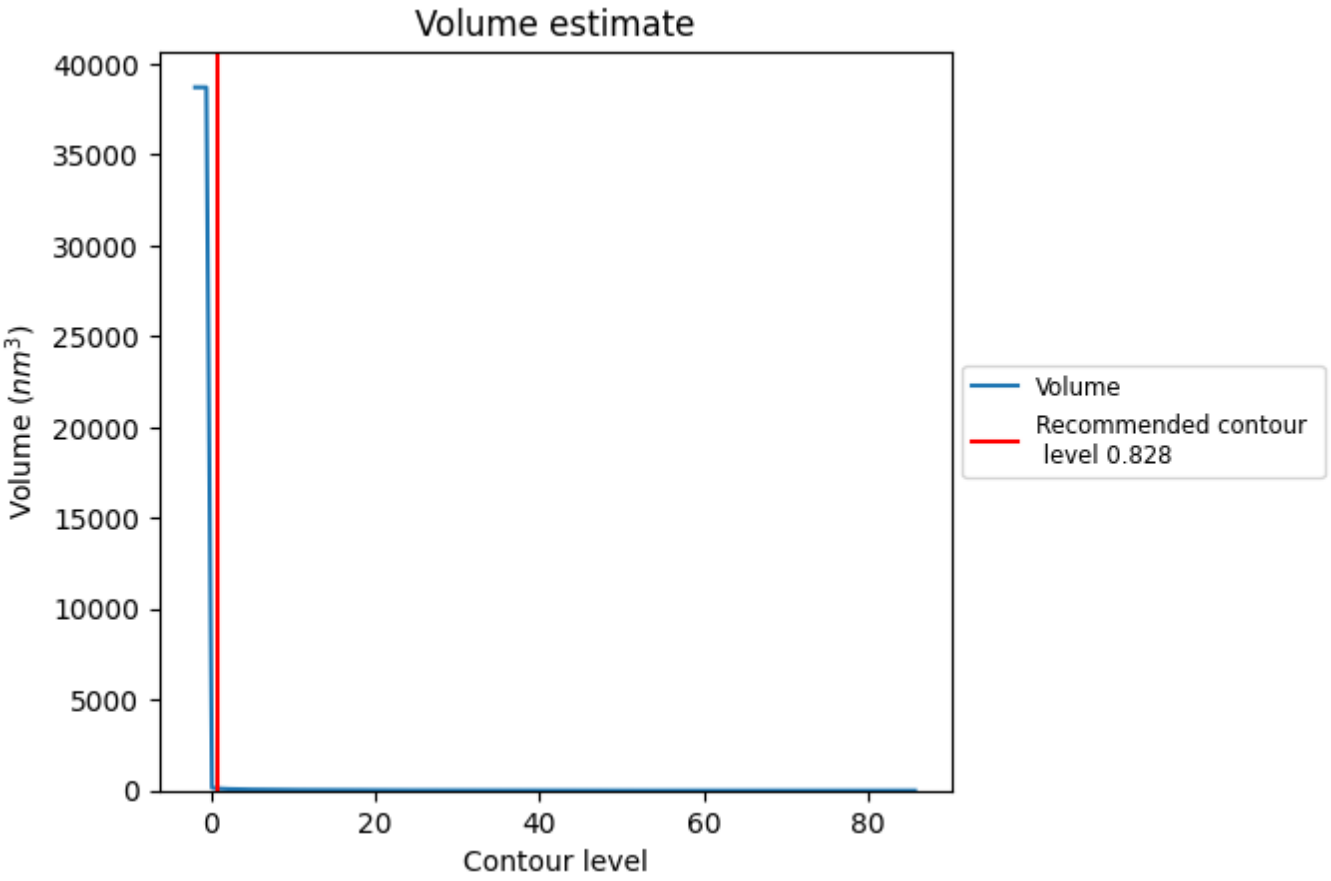
This section contains the results of statistical analysis of the map.

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.

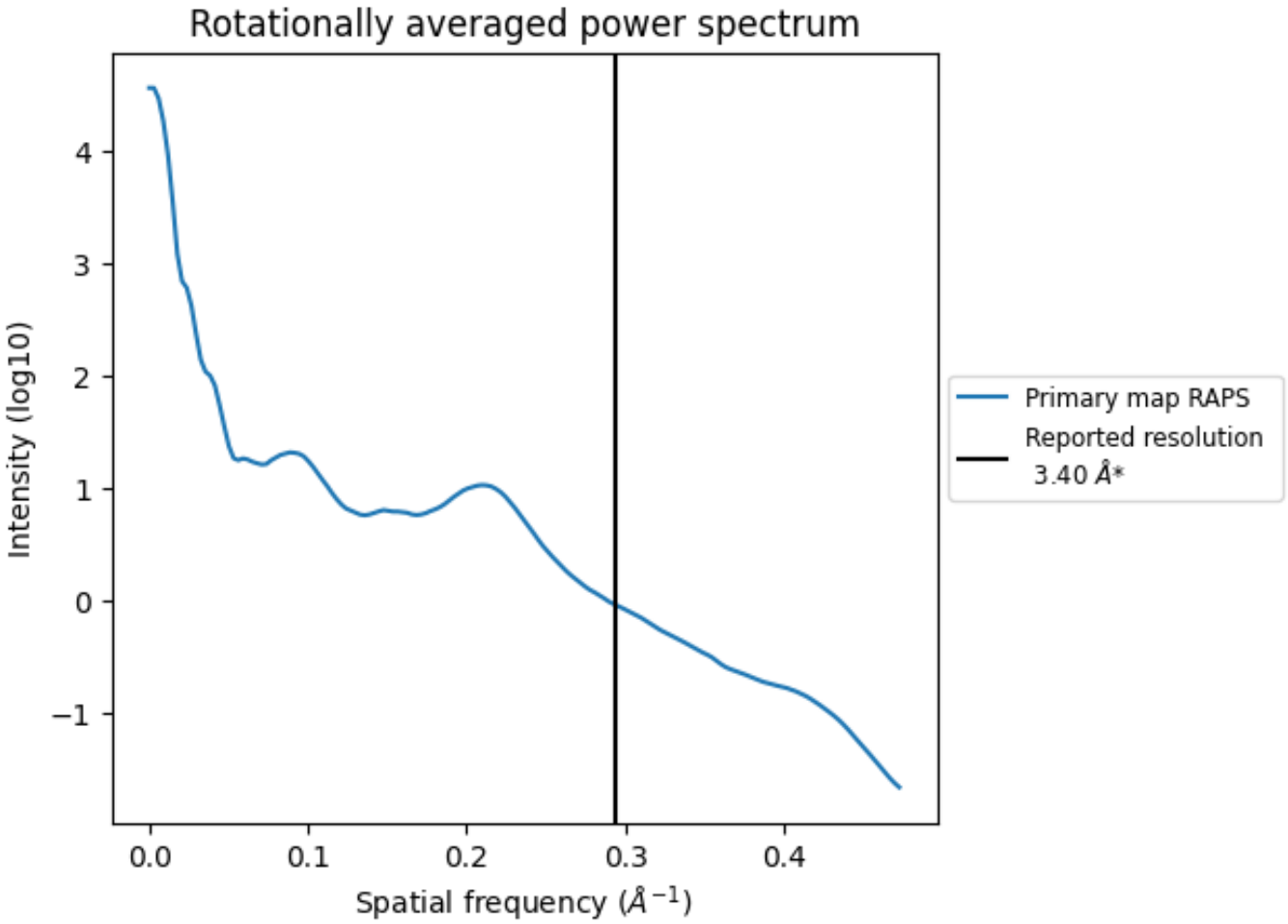
7.2 Volume estimate (i)



The volume at the recommended contour level is 113 nm^3; this corresponds to an approximate mass of 102 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 0.294 $Å^{-1}$

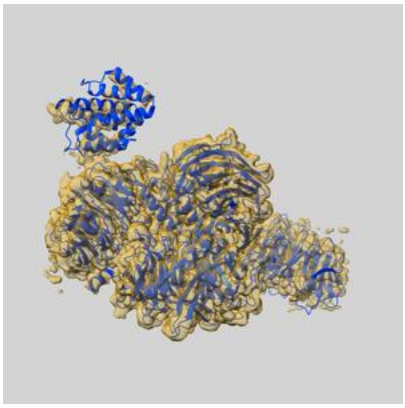
8 Fourier-Shell correlation (i)

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

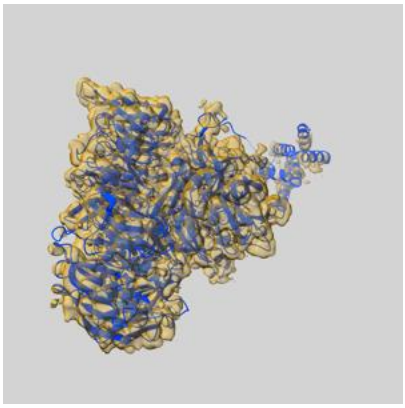
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-18398 and PDB model 8QH5. Per-residue inclusion information can be found in section 3 on page 7.

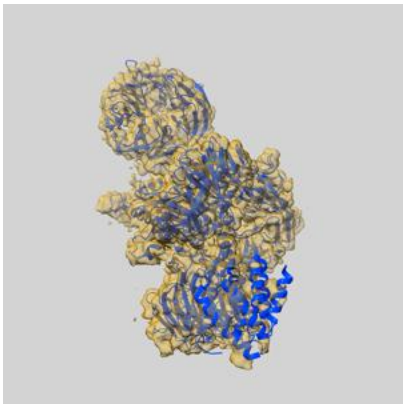
9.1 Map-model overlay



X



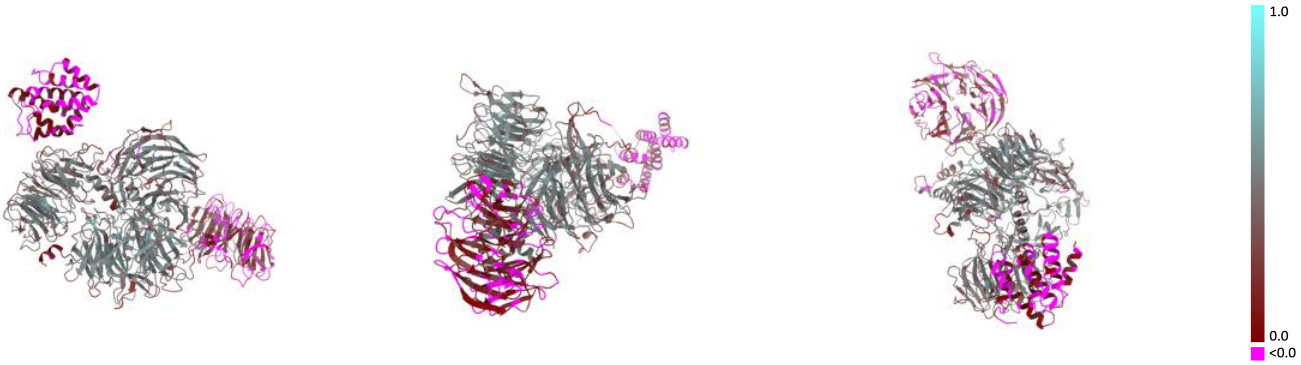
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.828 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.

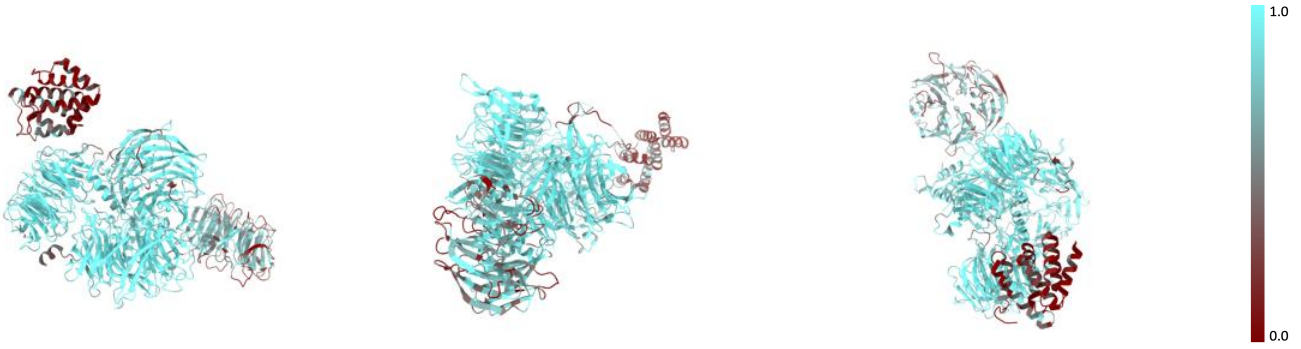
9.2 Q-score mapped to coordinate model (i)



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.

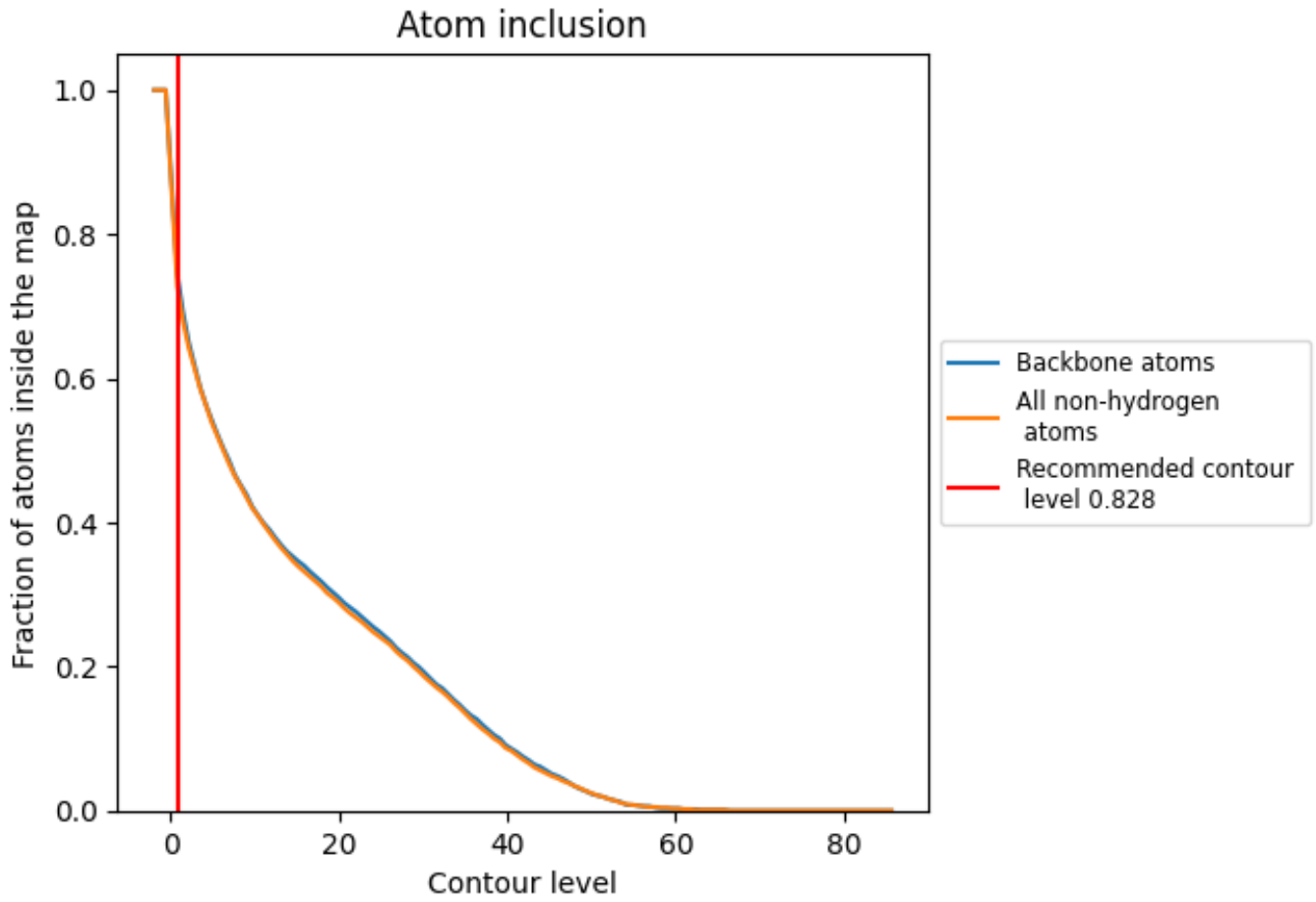
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.828).

9.4 Atom inclusion (i)



At the recommended contour level, 75% of all backbone atoms, 73% 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 (0.828) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7290	0.3200
A	0.9140	0.4010
B	0.8240	0.3400
C	0.7630	0.3070
D	0.2000	-0.0010