wwPDB EM Validation Summary Report

Mar 5, 2025 – 02:26 AM EST

PDB ID	:	9DCH
EMDB ID	:	EMD-46751
Title	:	Single-stranded RNA-mediated PRC2 dimer
Authors	:	Song, J.S.; Kasinath, V.K.
Deposited on	:	2024-08-26
Resolution	:	3.40 Å(reported)
Based on initial model	:	8FYH

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:

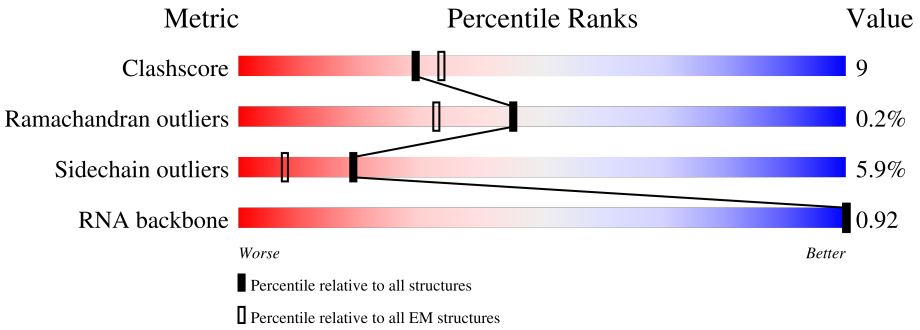
EMDB validation analysis	:	0.0.1.dev117
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

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.



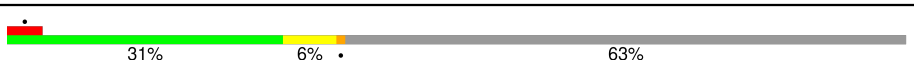
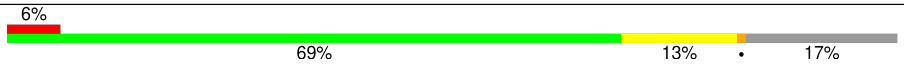
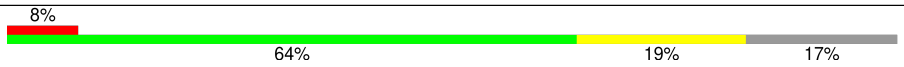
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	G	10	80%, 100%
2	A	750	12%, 51%, 9%, •, 39%
2	H	750	8%, 53%, 8%, •, 39%
3	B	727	9%, 48%, 10%, 42%
3	I	727	11%, 46%, 11%, 42%
4	D	425	7%, 72%, 22%, • 5%
4	K	425	8%, 72%, 22%, • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	F	294	
5	M	294	
6	E	332	
6	L	332	
7	C	441	
7	J	441	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 24433 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 RNA chain called TERRAmut RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	G	10	Total	C	N	O	P	0	0
			219	97	41	71	10		

- Molecule 2 is a protein called Isoform 2 of Histone-lysine N-methyltransferase EZH2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	460	Total	C	N	O	S	0	0
			3124	1967	558	568	31		
2	H	460	Total	C	N	O	S	0	0
			3128	1969	559	569	31		

- Molecule 3 is a protein called Polycomb protein SUZ12.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	423	Total	C	N	O	S	0	0
			2786	1755	529	493	9		
3	I	423	Total	C	N	O	S	0	0
			2789	1756	529	495	9		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	GLY	deletion	UNP Q15022
B	?	-	SER	deletion	UNP Q15022
B	?	-	GLY	deletion	UNP Q15022
B	?	-	PRO	deletion	UNP Q15022
B	?	-	SER	deletion	UNP Q15022
B	?	-	ALA	deletion	UNP Q15022
B	?	-	GLY	deletion	UNP Q15022
B	?	-	SER	deletion	UNP Q15022
B	?	-	GLY	deletion	UNP Q15022
B	?	-	GLY	deletion	UNP Q15022
B	?	-	GLY	deletion	UNP Q15022

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	GLY	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022
I	?	-	SER	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022
I	?	-	PRO	deletion	UNP Q15022
I	?	-	SER	deletion	UNP Q15022
I	?	-	ALA	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022
I	?	-	SER	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022
I	?	-	GLY	deletion	UNP Q15022

- Molecule 4 is a protein called RBAP48.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	405	Total	C	N	O	S	0	0
			2715	1714	496	504	1		
4	K	405	Total	C	N	O	S	0	0
			2718	1717	496	504	1		

- Molecule 5 is a protein called Zinc finger protein AEBP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	109	Total	C	N	O	S	0	0
			695	438	133	123	1		
5	M	109	Total	C	N	O	S	0	0
			695	438	133	123	1		

- Molecule 6 is a protein called Protein Jumonji.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	E	31	Total	C	N	O	S	0	0
			187	114	35	37	1		
6	L	31	Total	C	N	O	S	0	0
			187	114	35	37	1		

- Molecule 7 is a protein called Polycomb protein EED.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	C	365	Total 2589	C 1645	N 466	O 468	S 10	0	0
7	J	365	Total 2587	C 1643	N 466	O 468	S 10	0	0

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

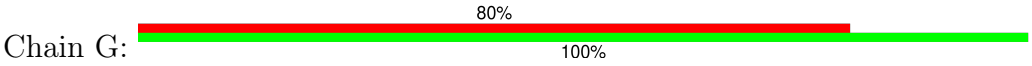
Mol	Chain	Residues	Atoms		AltConf
8	A	7	Total 7	Zn 7	0
8	H	7	Total 7	Zn 7	0

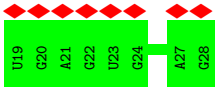
3 Residue-property plots (i)

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: TERRAmut RNA

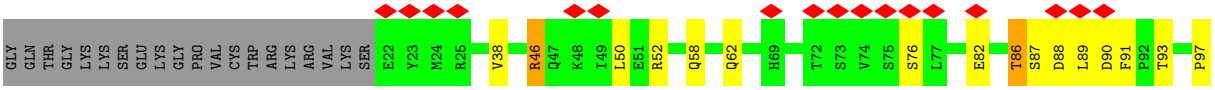
Chain G:

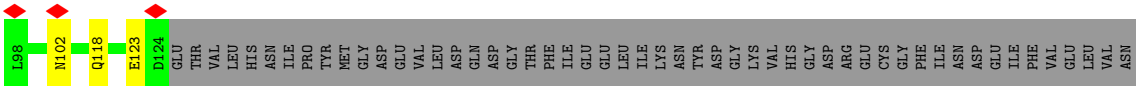


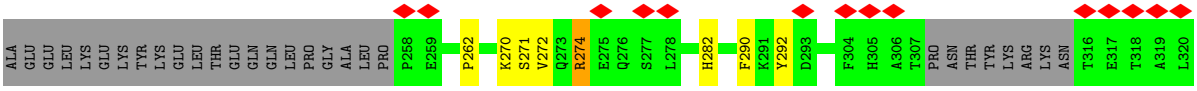


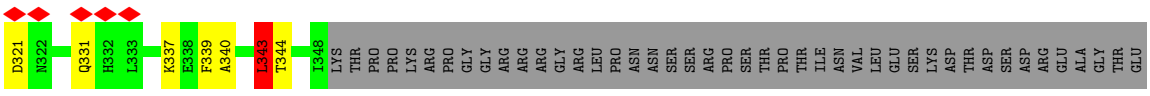
- Molecule 2: Isoform 2 of Histone-lysine N-methyltransferase EZH2

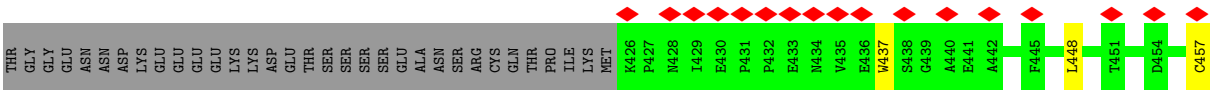
Chain A:

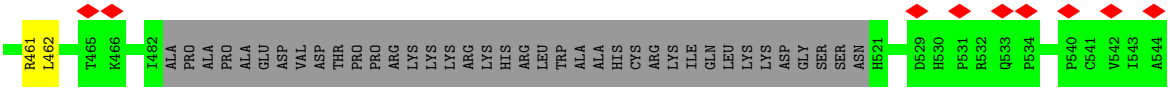


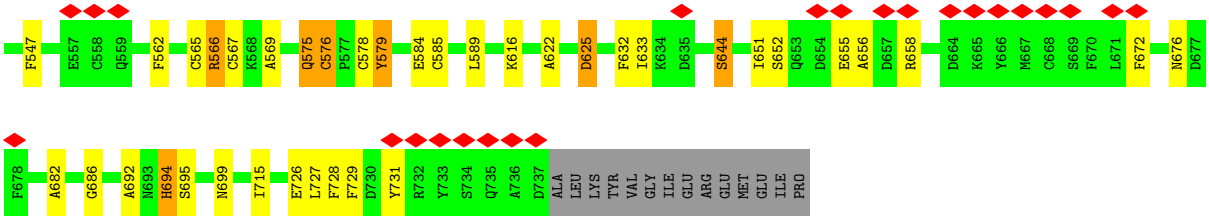






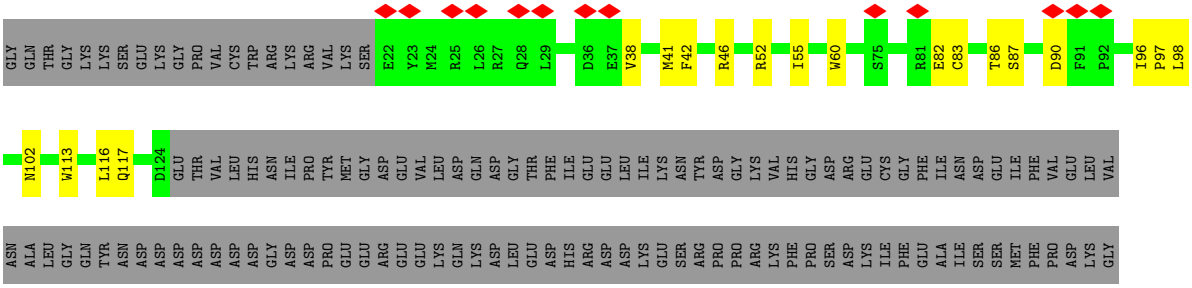


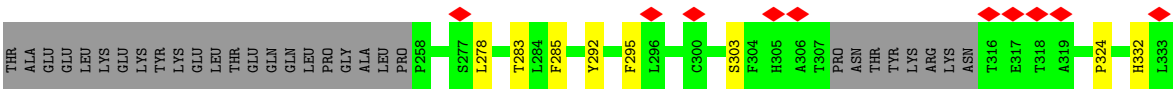


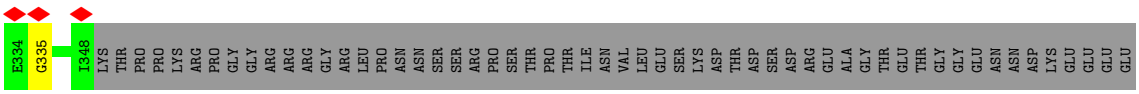


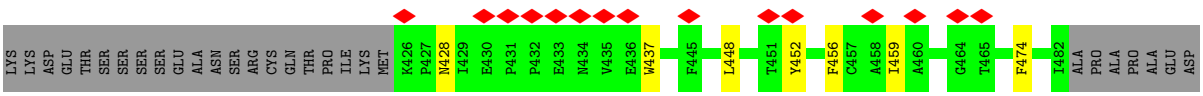
• Molecule 2: Isoform 2 of Histone-lysine N-methyltransferase EZH2

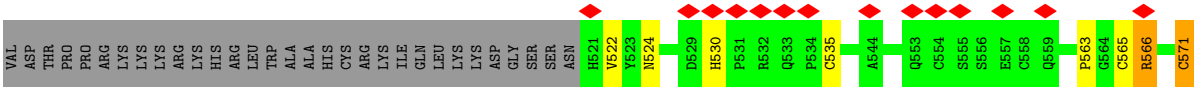
Chain H:

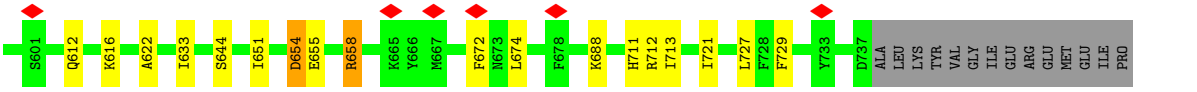






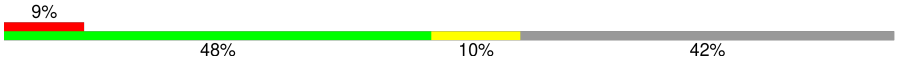


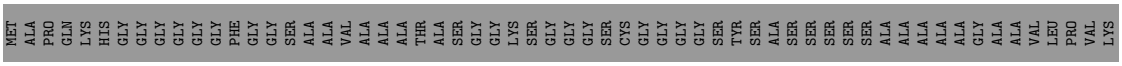


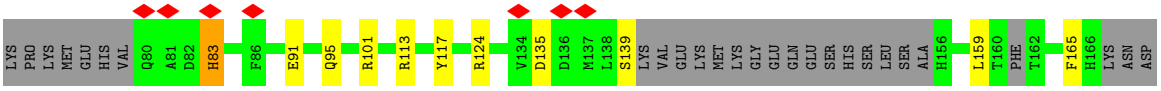


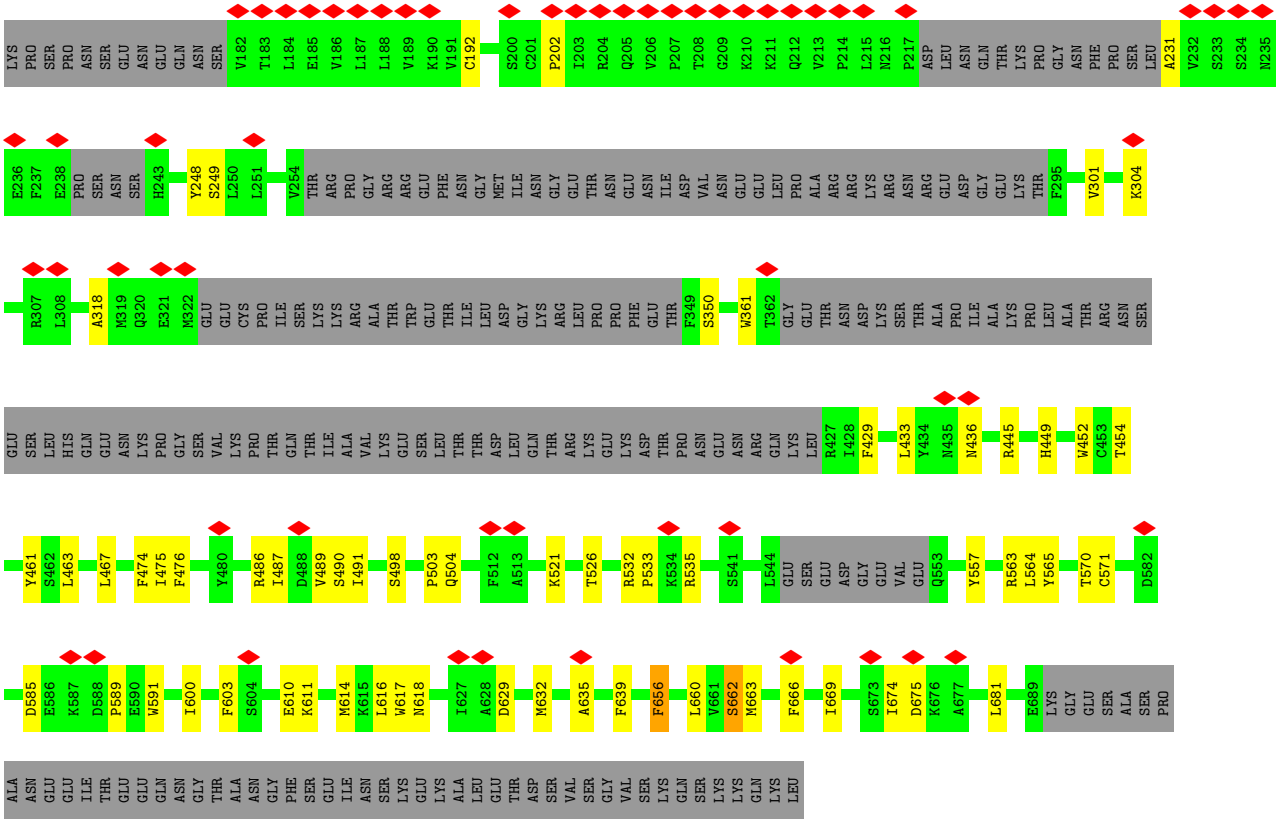
• Molecule 3: Polycomb protein SUZ12

Chain B:



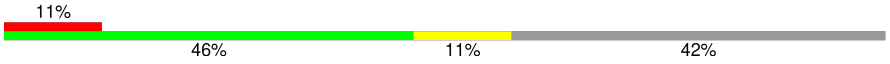


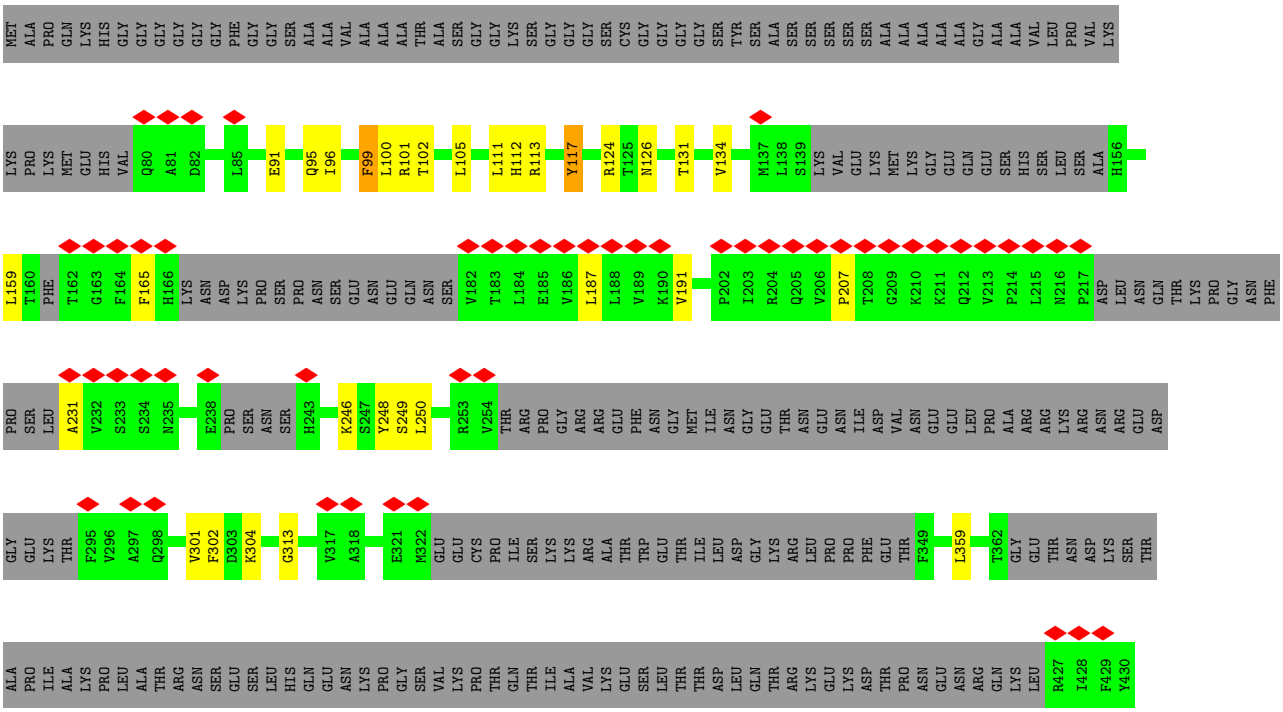


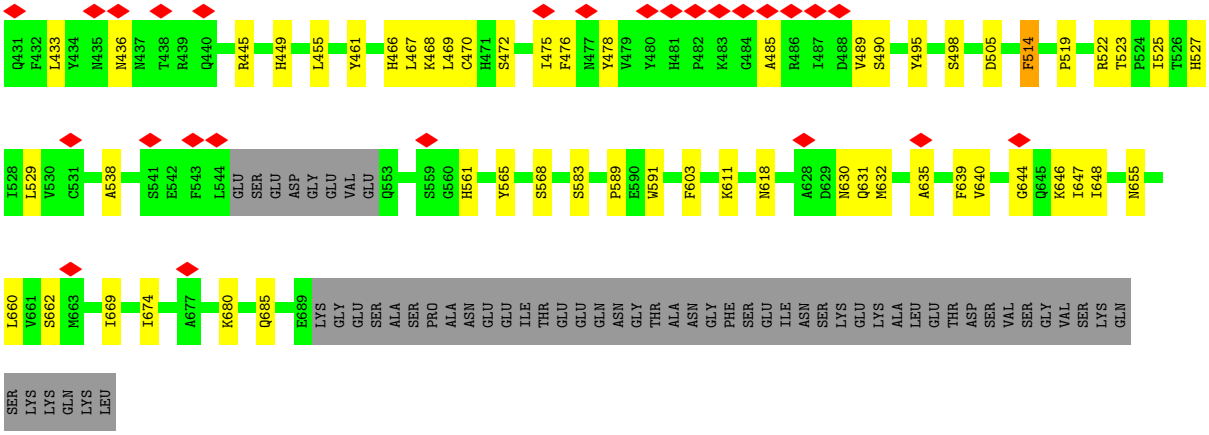


• Molecule 3: Polycomb protein SUZ12

Chain I:

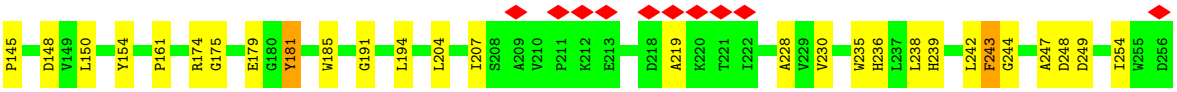


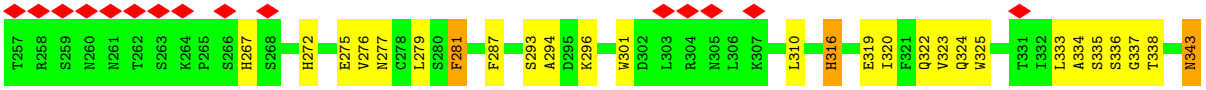


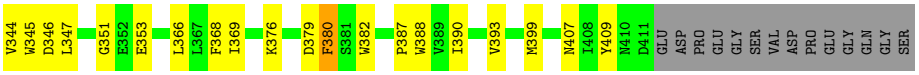


• Molecule 4: RBAP48

Chain D:

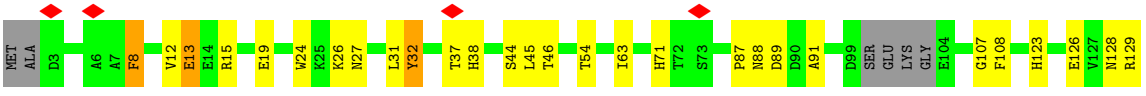


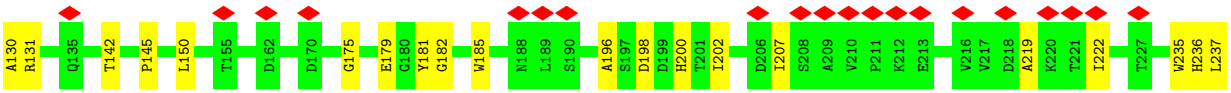


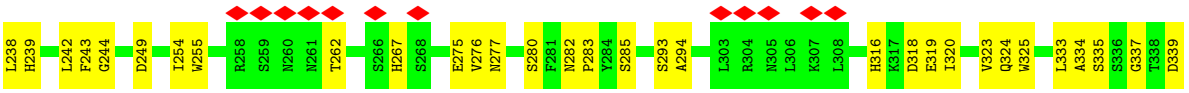


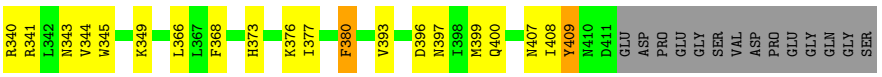
• Molecule 4: RBAP48

Chain K:



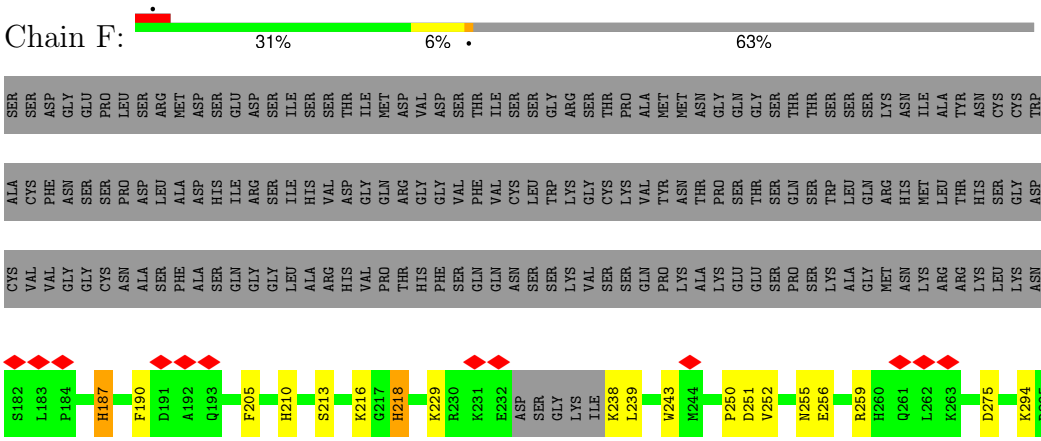






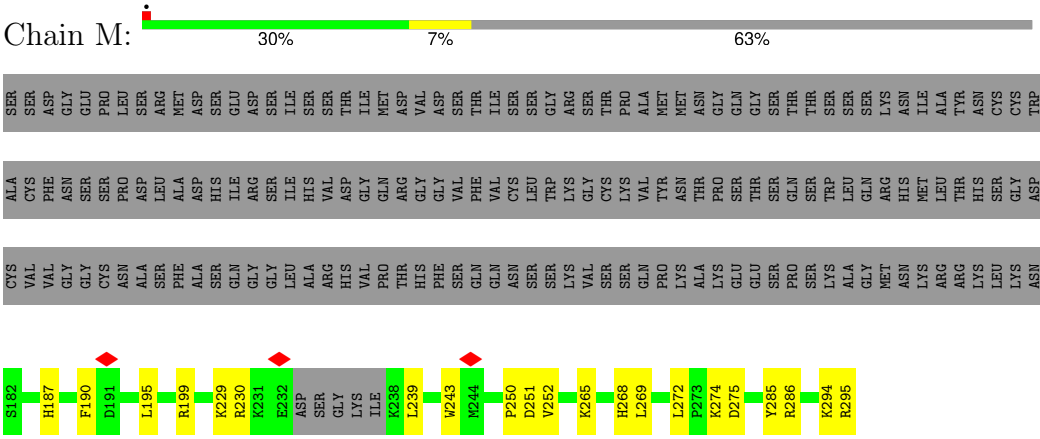
• Molecule 5: Zinc finger protein AEBP2

Chain F:



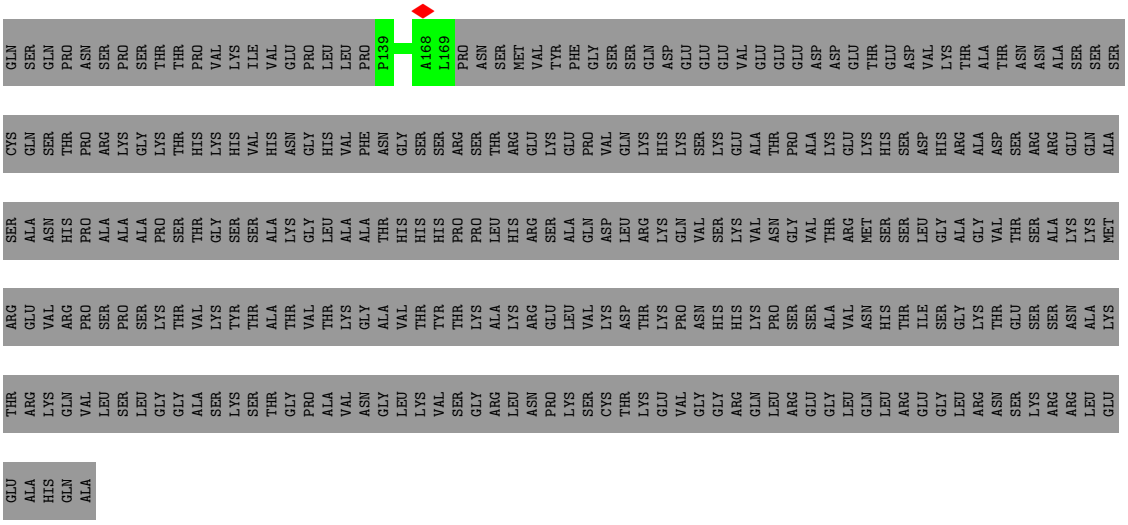
• Molecule 5: Zinc finger protein AEBP2

Chain M:



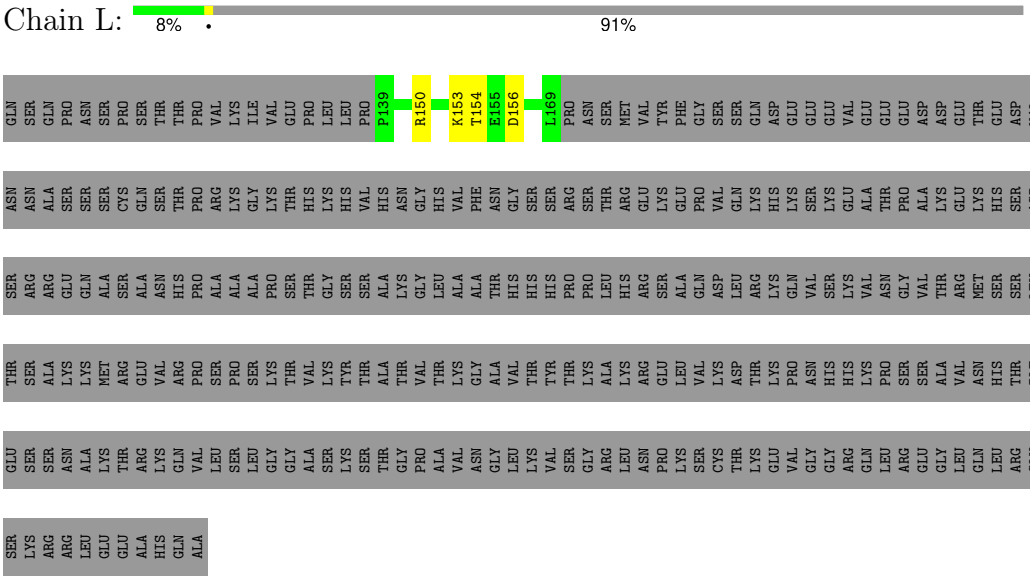
• Molecule 6: Protein Jumonji

Chain E:



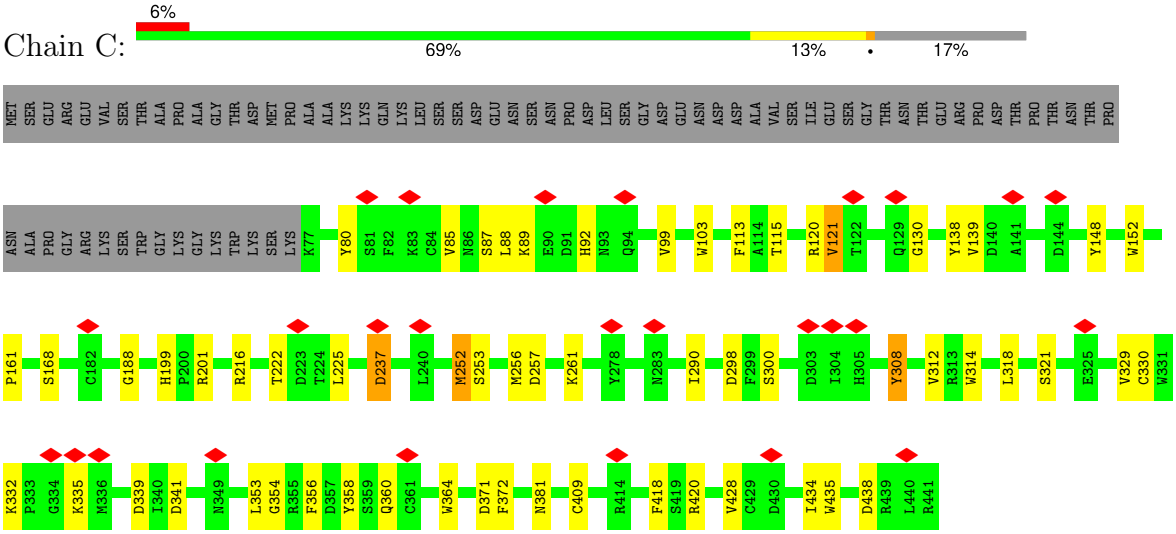
• Molecule 6: Protein Jumonji

Chain L:



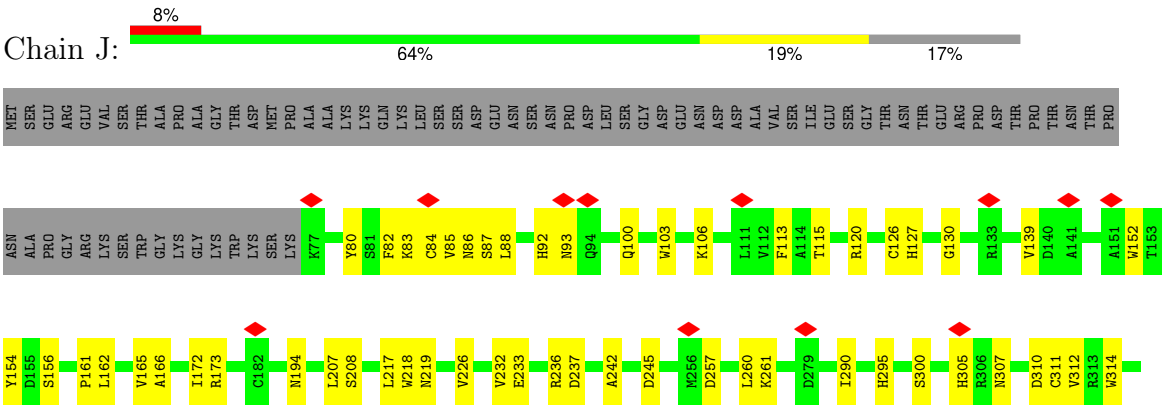
• Molecule 7: Polycomb protein EED

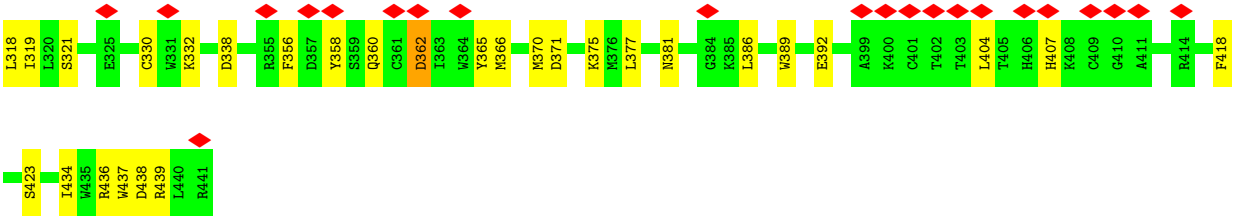
Chain C:



• Molecule 7: Polycomb protein EED

Chain J:




L318
I319
L320
S321
E325
C330
W331
K332
D338
R355
F356
D357
Y358
S359
Q360
C361
D362
I363
W364
Y365
M366
M370
D371
K375
M376
L377
N381
G384
K385
L386
W389
E392
A399
K400
C401
T402
T403
L404
T405
H406
H407
K408
C409
G410
A411
R414
F418
S423
I434
W435
R436
W437
D438
R439
L440
R441

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	120658	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION; CTFind	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	1900	Depositor
Magnification	130000	Depositor
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	0.046	Depositor
Minimum map value	-0.015	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.001	Depositor
Recommended contour level	0.0052	Depositor
Map size (Å)	310.40002, 310.40002, 310.40002	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.9700001, 0.9700001, 0.9700001	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	G	0.38	0/245	0.69	0/381
2	A	0.27	0/3193	0.51	2/4363 (0.0%)
2	H	0.27	0/3197	0.47	0/4368
3	B	0.24	0/2844	0.46	0/3893
3	I	0.24	0/2847	0.46	0/3897
4	D	0.24	0/2793	0.46	0/3849
4	K	0.24	0/2796	0.46	0/3853
5	F	0.23	0/714	0.43	0/986
5	M	0.23	0/714	0.44	0/986
6	E	0.25	0/190	0.51	0/261
6	L	0.24	0/190	0.52	0/261
7	C	0.24	0/2660	0.49	0/3649
7	J	0.24	0/2658	0.48	1/3646 (0.0%)
All	All	0.25	0/25041	0.47	3/34393 (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	1
2	H	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	575	GLN	CB-CA-C	-10.06	90.27	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	343	LEU	CA-CB-CG	5.85	128.75	115.30
7	J	362	ASP	CB-CG-OD1	5.14	122.93	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	566	ARG	Sidechain
2	H	566	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	G	219	0	108	0	0
2	A	3124	0	2523	61	0
2	H	3128	0	2529	43	0
3	B	2786	0	2067	50	0
3	I	2789	0	2075	53	0
4	D	2715	0	2082	63	0
4	K	2718	0	2091	68	0
5	F	695	0	496	15	0
5	M	695	0	496	15	0
6	E	187	0	148	0	0
6	L	187	0	148	2	0
7	C	2589	0	2145	34	0
7	J	2587	0	2138	54	0
8	A	7	0	0	0	0
8	H	7	0	0	0	0
All	All	24433	0	19046	397	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 397 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:575:GLN:HB2	2:H:563:PRO:HB2	1.38	1.02
2:H:530:HIS:HE1	2:H:535:CYS:SG	2.03	0.81
7:C:312:VAL:HA	7:C:321:SER:HA	1.63	0.79
2:A:565:CYS:HB3	2:A:567:CYS:SG	2.24	0.77
2:A:575:GLN:HB2	2:H:563:PRO:CB	2.15	0.77

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	
2	A	450/750 (60%)	419 (93%)	30 (7%)	1 (0%)	44	72
2	H	450/750 (60%)	413 (92%)	36 (8%)	1 (0%)	44	72
3	B	403/727 (55%)	378 (94%)	24 (6%)	1 (0%)	44	72
3	I	403/727 (55%)	371 (92%)	31 (8%)	1 (0%)	44	72
4	D	401/425 (94%)	359 (90%)	41 (10%)	1 (0%)	44	72
4	K	401/425 (94%)	364 (91%)	36 (9%)	1 (0%)	44	72
5	F	105/294 (36%)	95 (90%)	10 (10%)	0	100	100
5	M	105/294 (36%)	95 (90%)	10 (10%)	0	100	100
6	E	29/332 (9%)	22 (76%)	7 (24%)	0	100	100
6	L	29/332 (9%)	19 (66%)	10 (34%)	0	100	100
7	C	363/441 (82%)	338 (93%)	25 (7%)	0	100	100
7	J	363/441 (82%)	337 (93%)	26 (7%)	0	100	100
All	All	3502/5938 (59%)	3210 (92%)	286 (8%)	6 (0%)	45	72

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	669	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	I	669	ILE
4	D	207	ILE
4	K	207	ILE
2	H	571	CYS

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
2	A	249/671 (37%)	233 (94%)	16 (6%)	14 39
2	H	250/671 (37%)	240 (96%)	10 (4%)	27 52
3	B	179/642 (28%)	168 (94%)	11 (6%)	15 40
3	I	181/642 (28%)	169 (93%)	12 (7%)	14 39
4	D	194/375 (52%)	181 (93%)	13 (7%)	13 38
4	K	195/375 (52%)	184 (94%)	11 (6%)	17 43
5	F	42/262 (16%)	38 (90%)	4 (10%)	7 25
5	M	42/262 (16%)	39 (93%)	3 (7%)	12 37
6	E	14/284 (5%)	14 (100%)	0	100 100
6	L	14/284 (5%)	14 (100%)	0	100 100
7	C	213/392 (54%)	199 (93%)	14 (7%)	14 39
7	J	212/392 (54%)	201 (95%)	11 (5%)	19 45
All	All	1785/5252 (34%)	1680 (94%)	105 (6%)	19 41

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	C	438	ASP
3	I	248	TYR
7	J	307	ASN
2	H	46	ARG
2	H	654	ASP

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

Mol	Chain	Res	Type
3	I	618	ASN
3	I	624	HIS
4	K	27	ASN
2	A	575	GLN
2	A	297	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	G	9/10 (90%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 14 ligands modelled in this entry, 14 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)

There are no chain breaks in this entry.

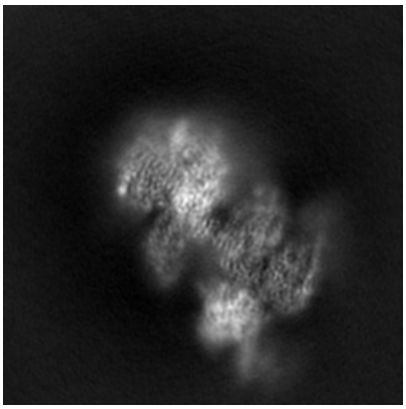
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-46751. These allow visual inspection of the internal detail of the map and identification of artifacts.

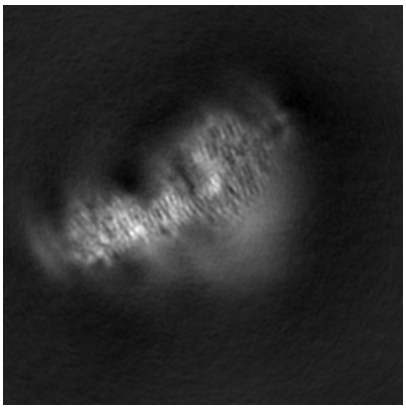
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

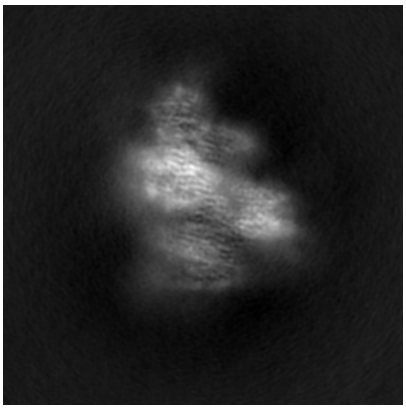
6.1.1 Primary map



X

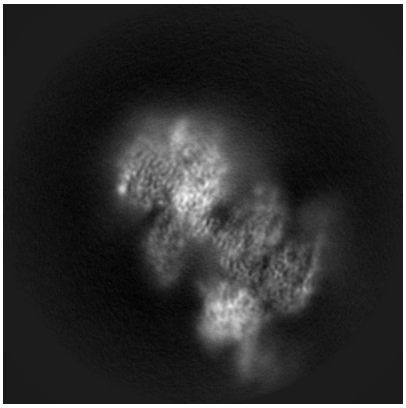


Y

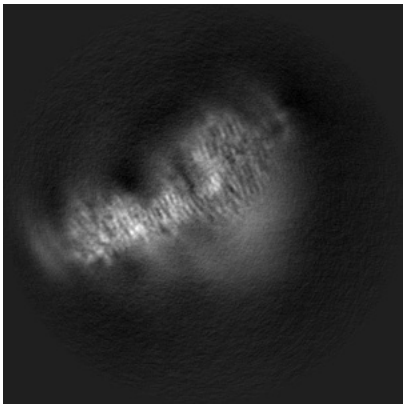


Z

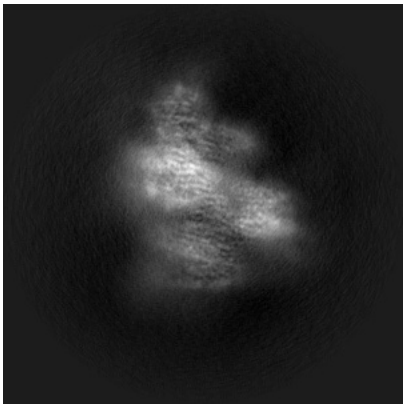
6.1.2 Raw map



X



Y

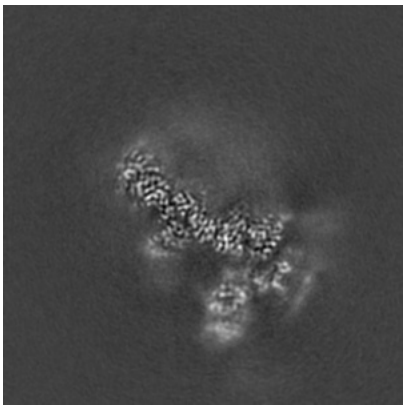


Z

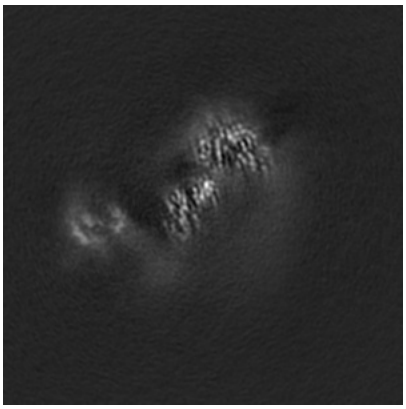
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

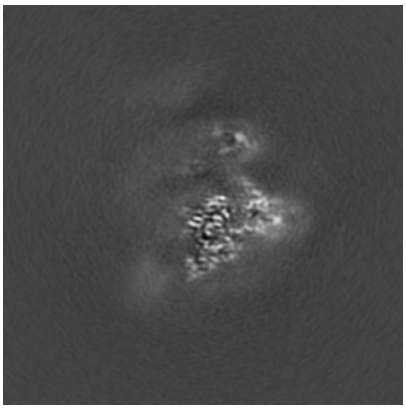
6.2.1 Primary map



X Index: 160

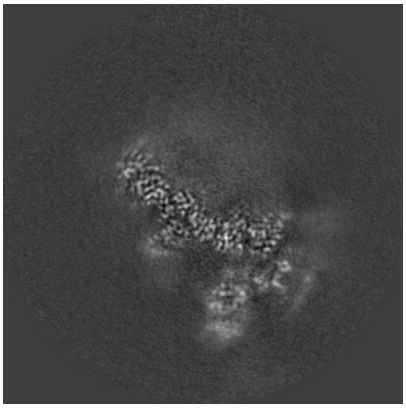


Y Index: 160

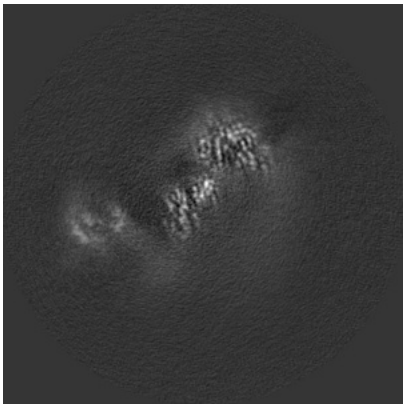


Z Index: 160

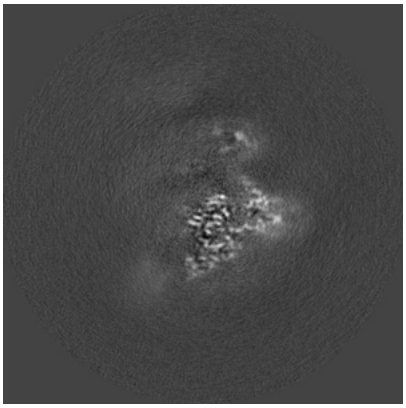
6.2.2 Raw map



X Index: 160



Y Index: 160

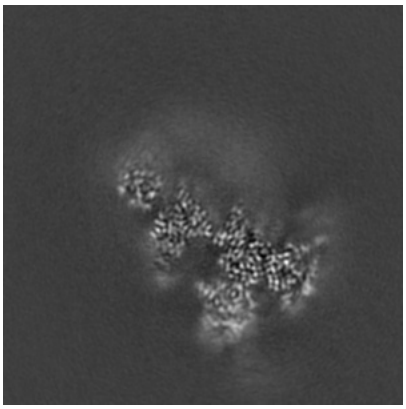


Z Index: 160

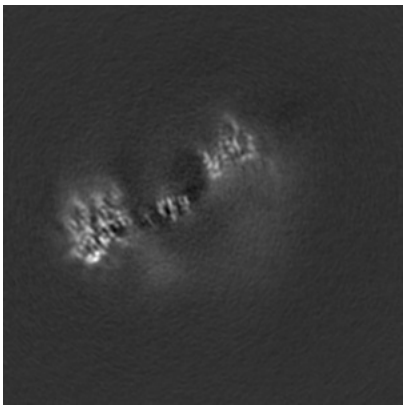
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

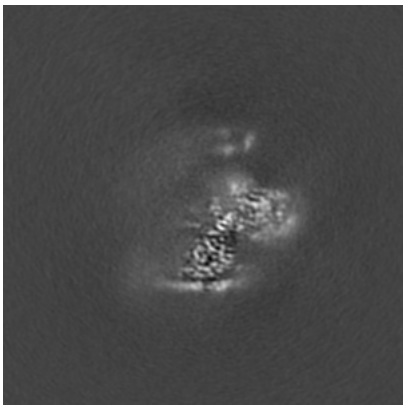
6.3.1 Primary map



X Index: 151

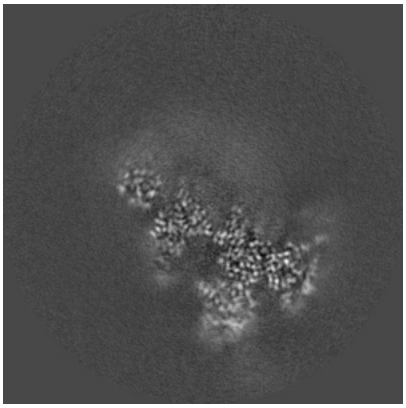


Y Index: 172

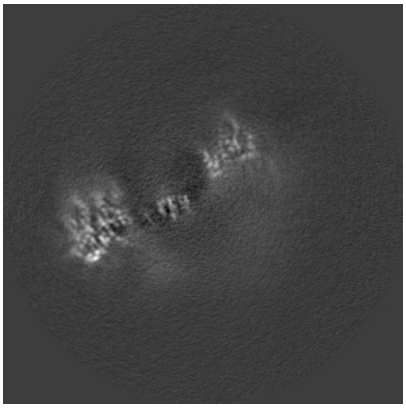


Z Index: 172

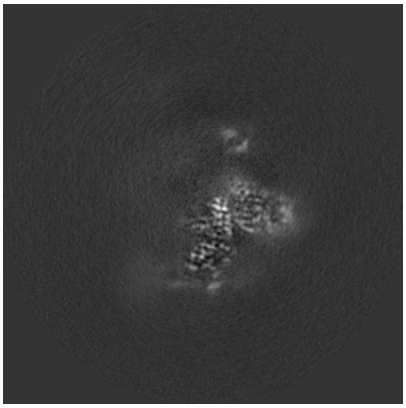
6.3.2 Raw map



X Index: 151



Y Index: 172

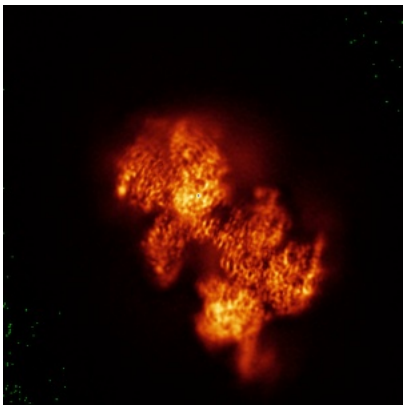


Z Index: 167

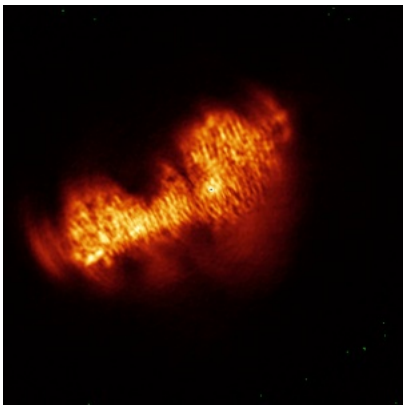
The images above show the largest variance slices of the map in three orthogonal directions.

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

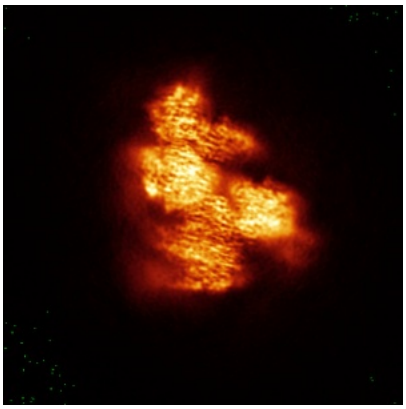
6.4.1 Primary map



X

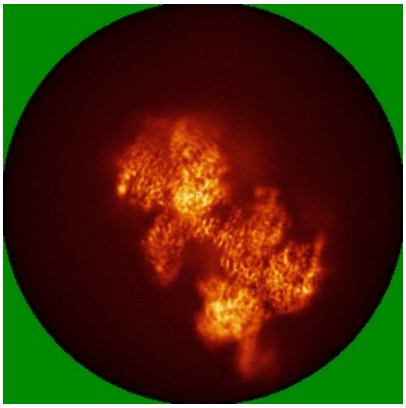


Y

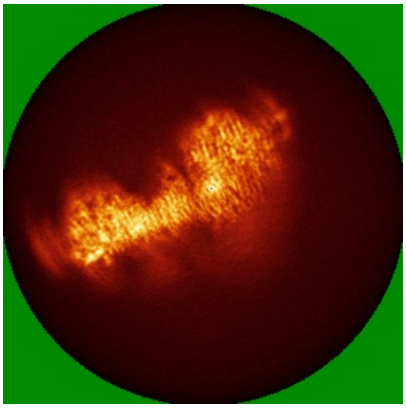


Z

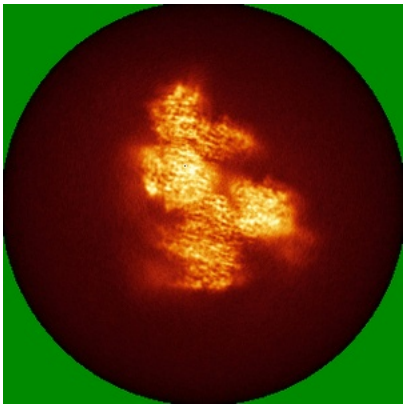
6.4.2 Raw map



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 0.0052. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

X Y Z

These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

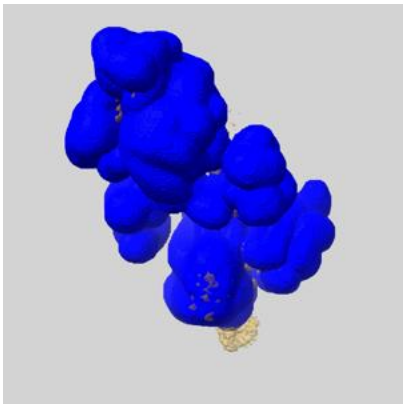
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

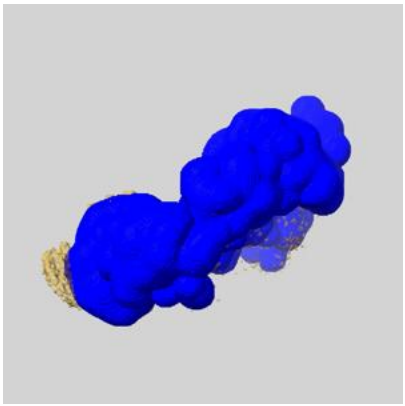
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

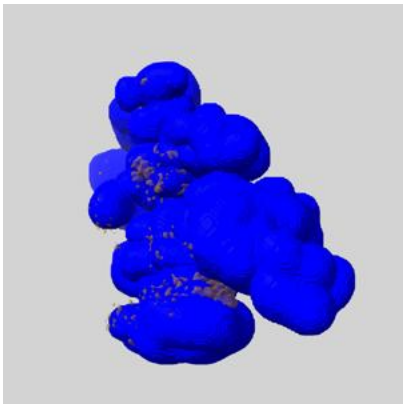
6.6.1 emd_46751_msk_1.map (i)



X



Y

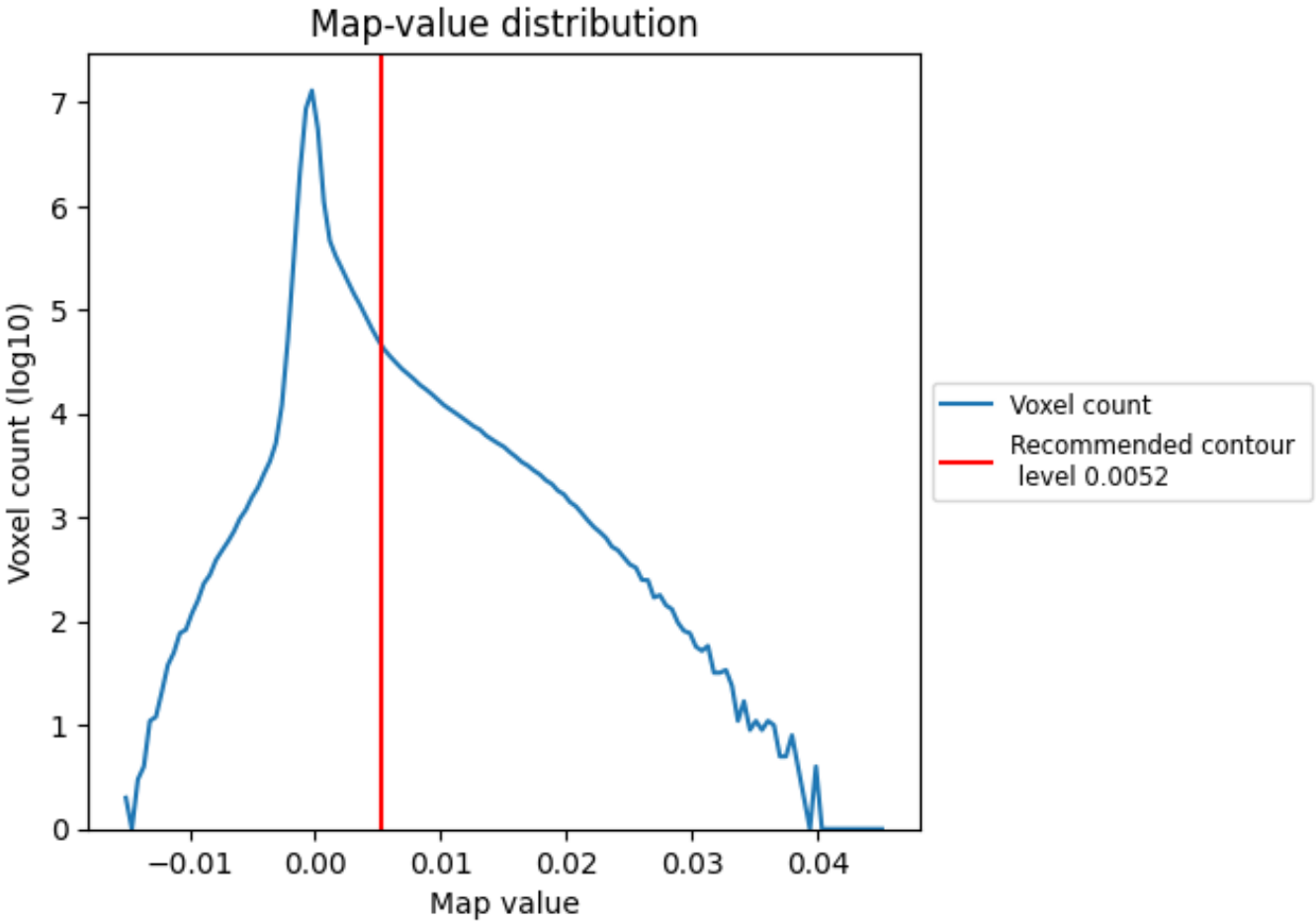


Z

7 Map analysis

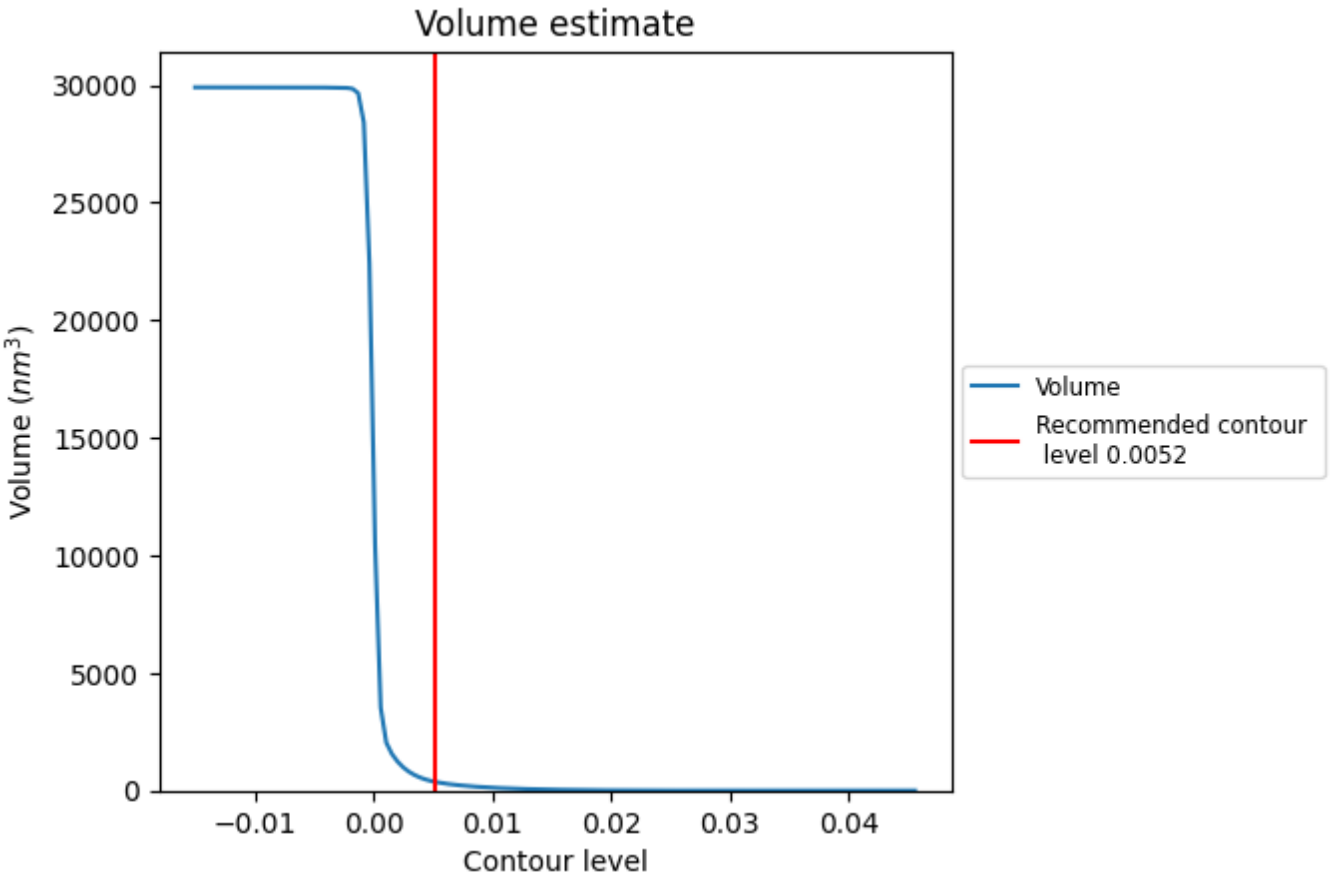
This section contains the results of statistical analysis of the map.

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.

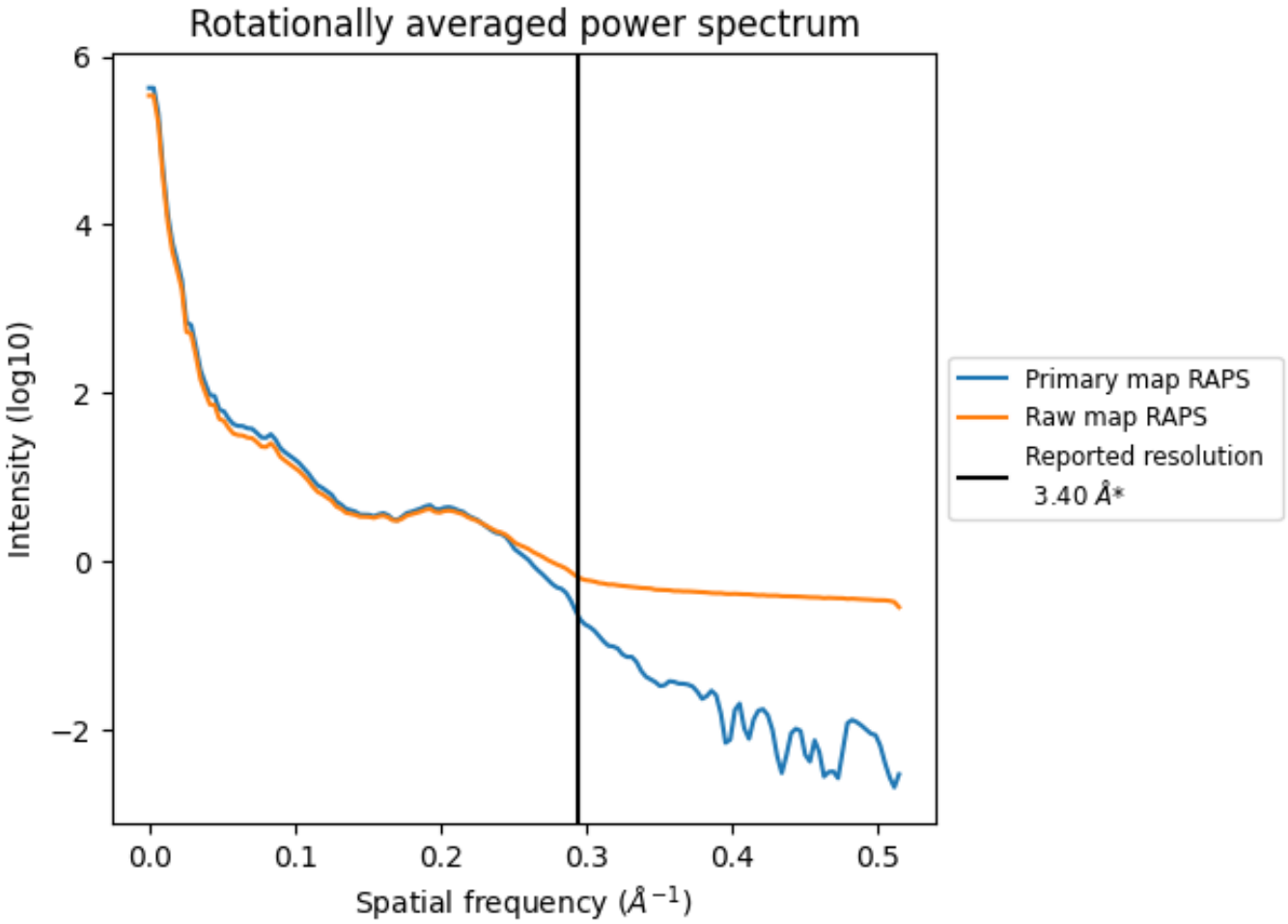
7.2 Volume estimate (i)



The volume at the recommended contour level is 367 nm^3; this corresponds to an approximate mass of 332 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)

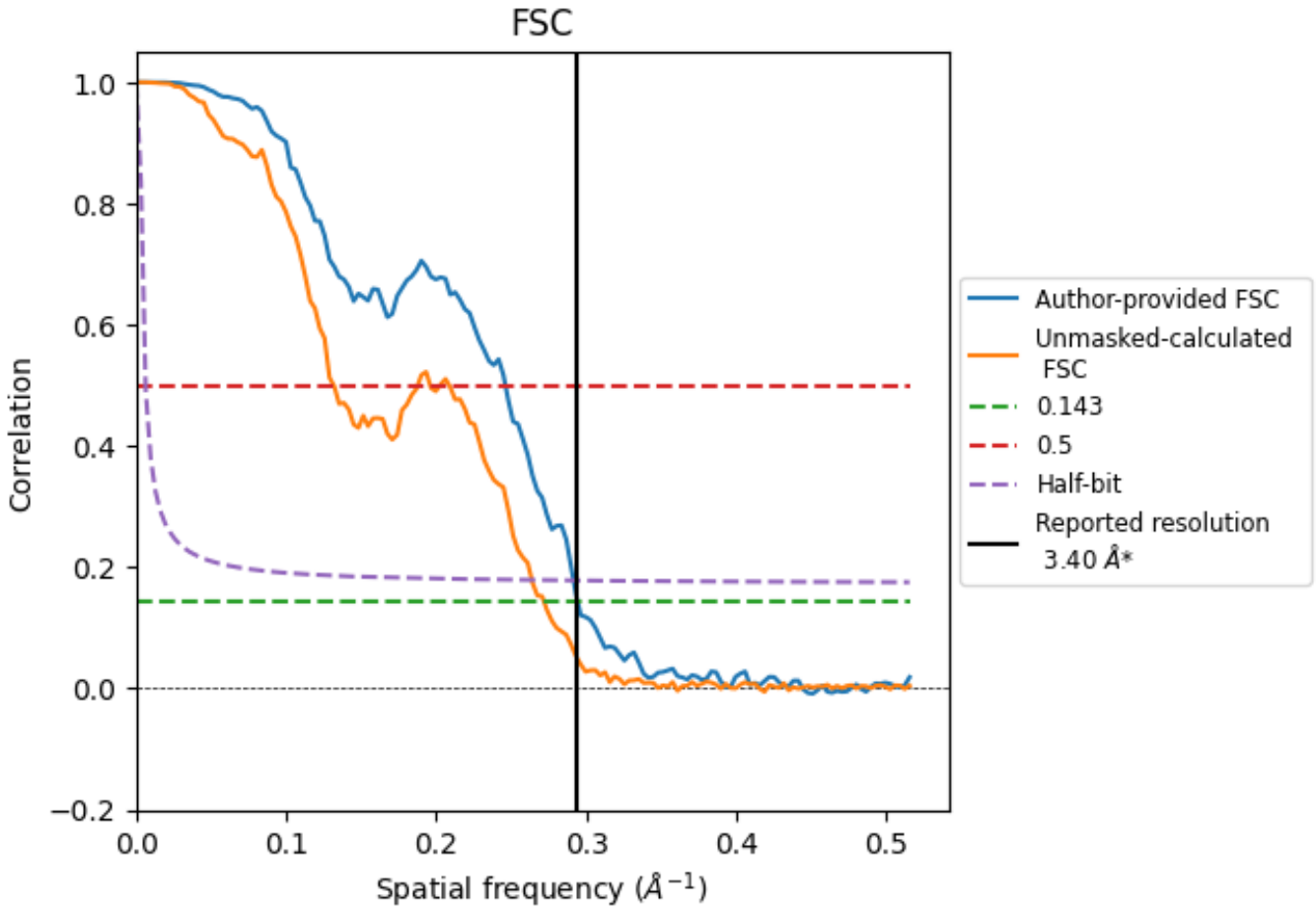


*Reported resolution corresponds to spatial frequency of 0.294 $Å^{-1}$

8 Fourier-Shell correlation

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



*Reported resolution corresponds to spatial frequency of 0.294 Å^{-1}

8.2 Resolution estimates (i)

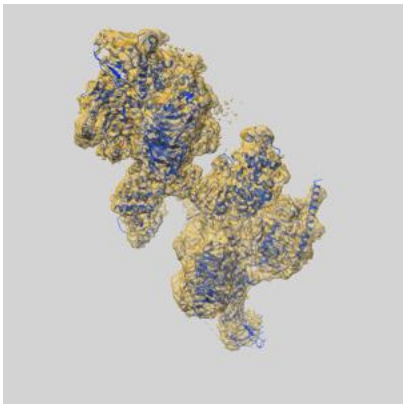
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.40	-	-
Author-provided FSC curve	3.40	4.06	3.43
Unmasked-calculated*	3.68	7.59	3.79

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

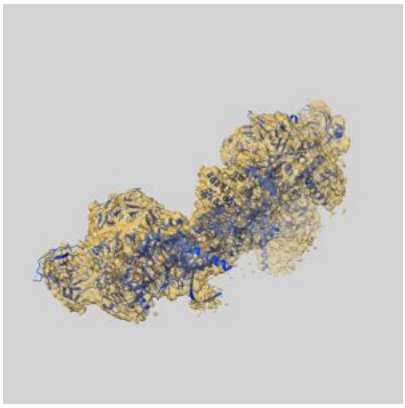
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-46751 and PDB model 9DCH. Per-residue inclusion information can be found in section 3 on page 7.

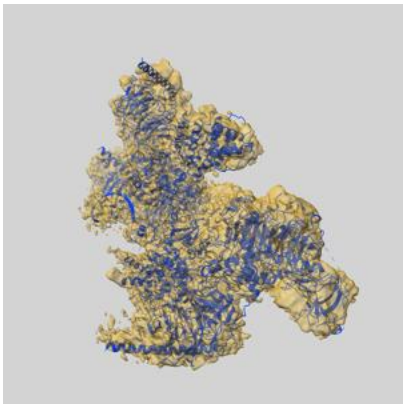
9.1 Map-model overlay



X



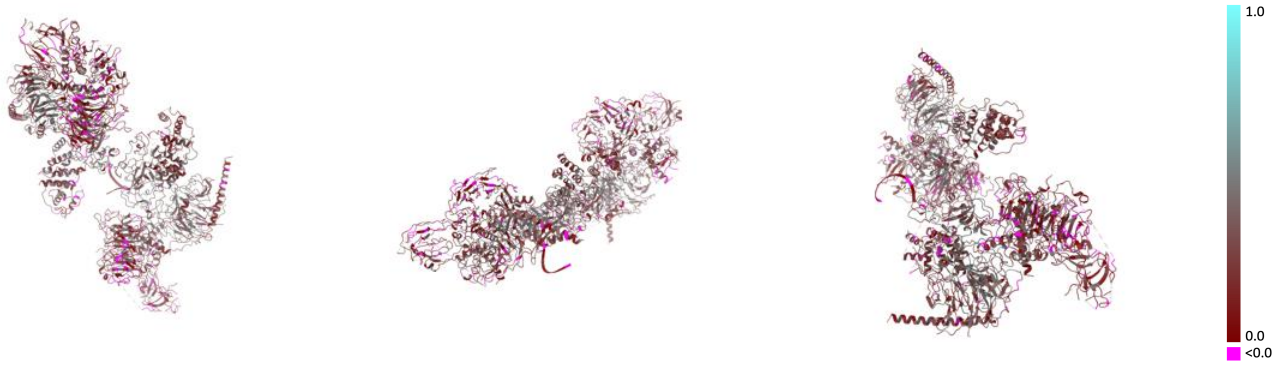
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0052 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.

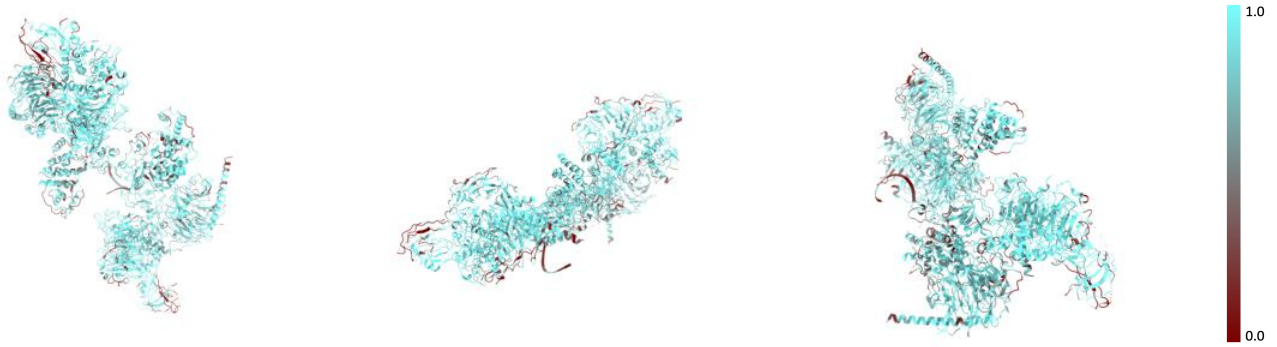
9.2 Q-score mapped to coordinate model



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.

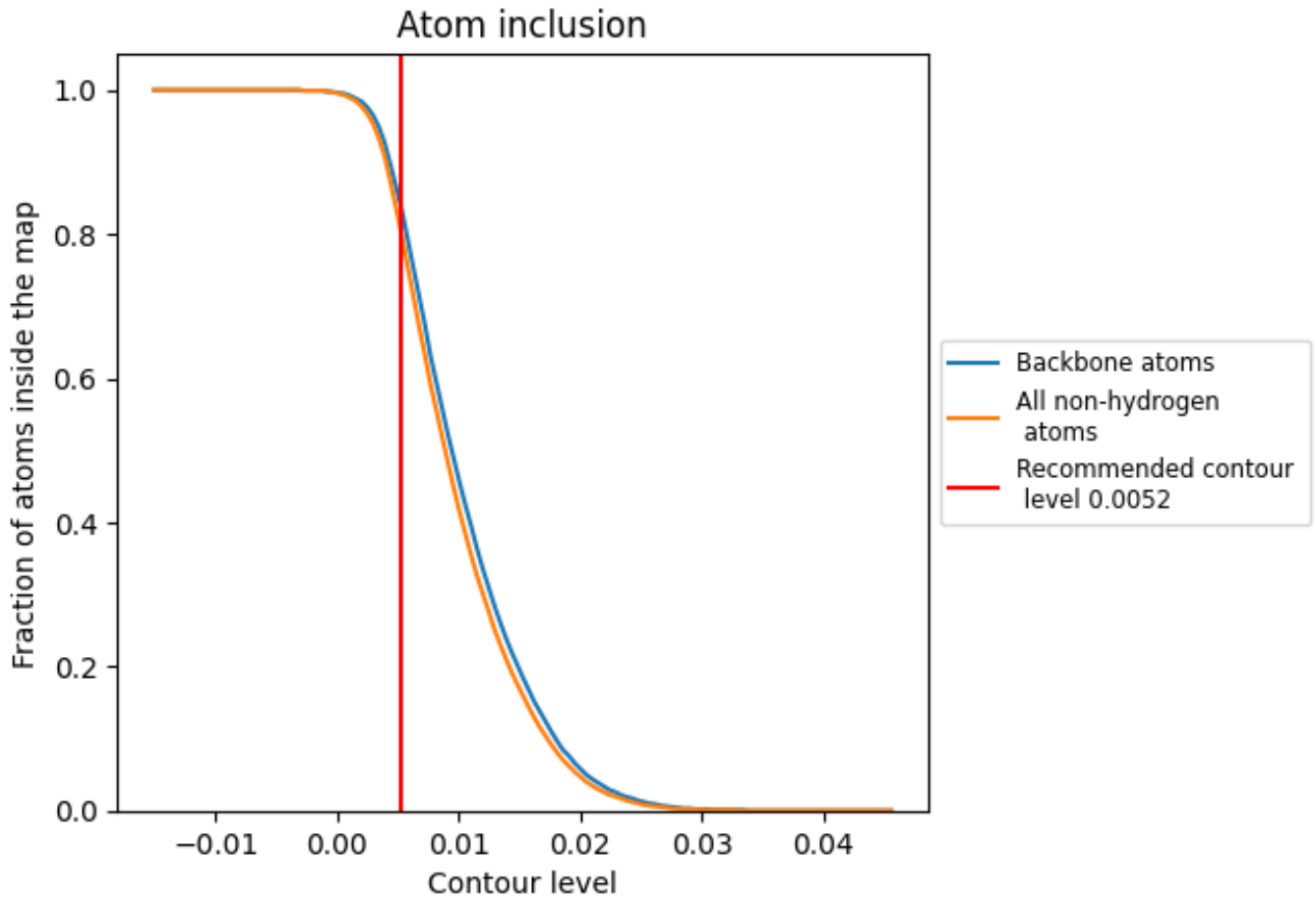
9.3 Atom inclusion mapped to coordinate model



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 (0.0052).

9.4 Atom inclusion (i)



At the recommended contour level, 84% of all backbone atoms, 81% 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 (0.0052) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8070	0.2670
A	0.7190	0.2950
B	0.7760	0.2370
C	0.7990	0.3450
D	0.9010	0.1980
E	0.9130	0.2010
F	0.8840	0.2580
G	0.2790	0.0600
H	0.7840	0.3150
I	0.7710	0.2540
J	0.8230	0.3330
K	0.8790	0.1870
L	0.8920	0.2110
M	0.9450	0.2600