



wwPDB EM Validation Summary Report ⓘ

Jun 20, 2024 – 10:24 AM EDT

PDB ID : 9ASL
EMDB ID : EMD-43818
Title : Human DNA polymerase theta helicase domain tetramer, apo-form
Authors : Ito, F.; Li, Z.; Chen, X.S.
Deposited on : 2024-02-26
Resolution : 3.50 Å (reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ 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

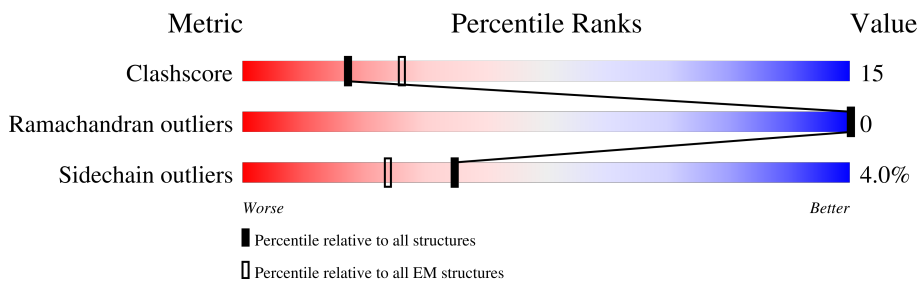
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	A	894	
1	B	894	
1	C	894	
1	D	894	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 23128 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.

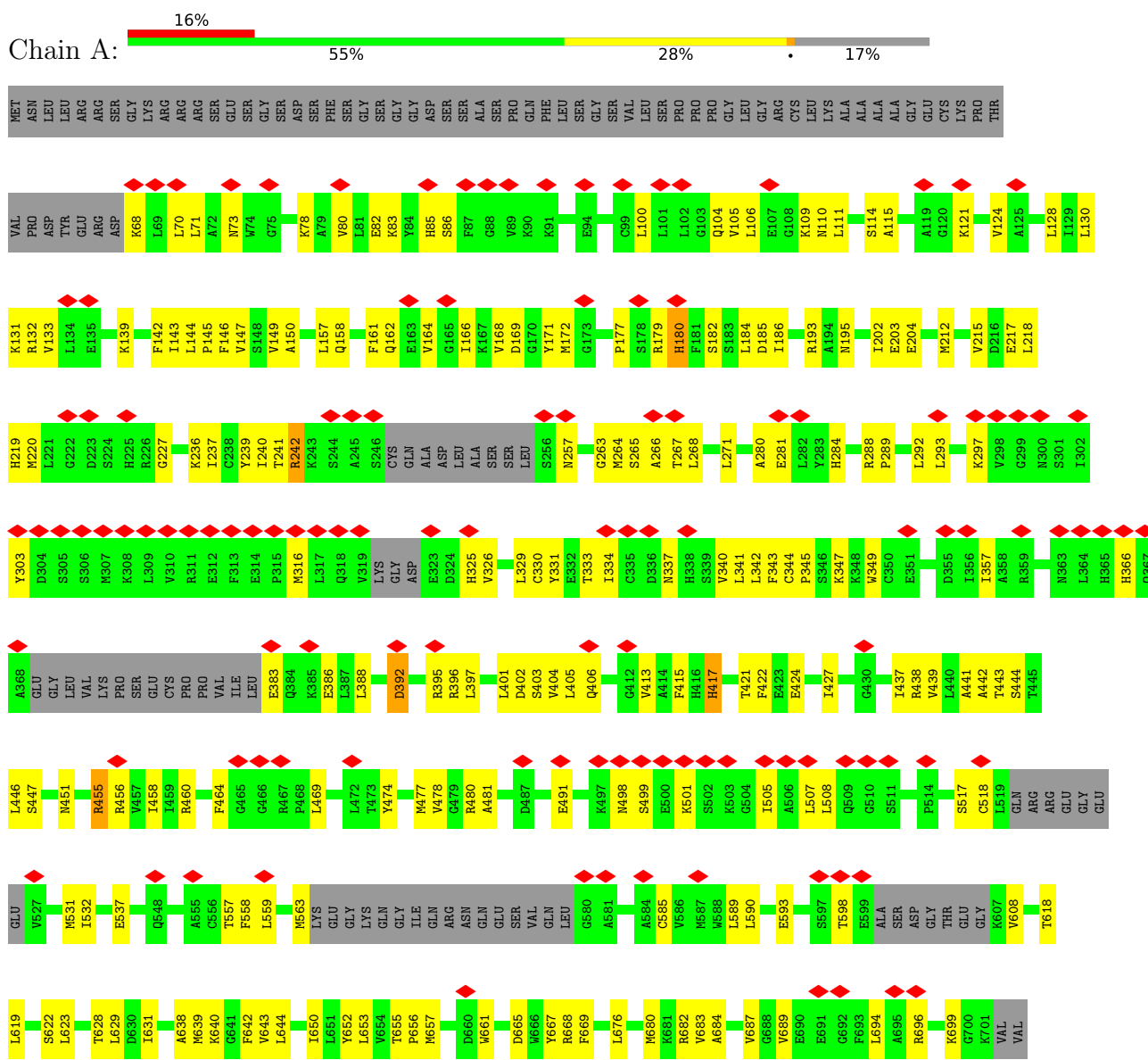
- Molecule 1 is a protein called DNA polymerase theta.

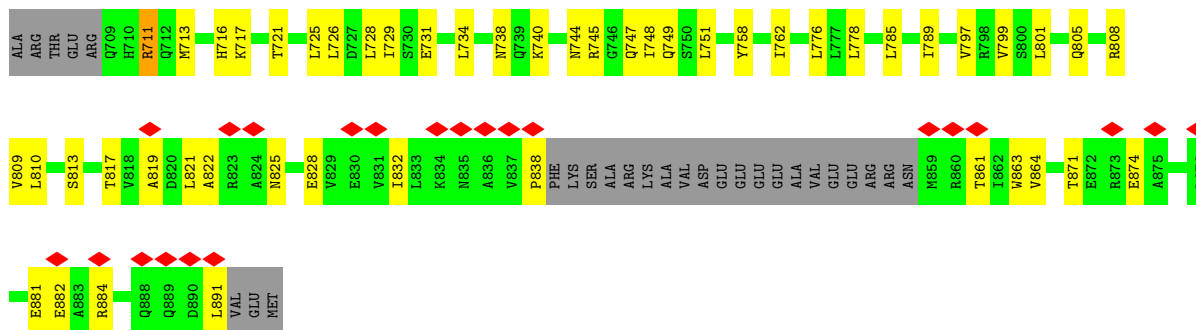
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	741	Total 5782	C 3707	N 989	O 1049	S 37	0	0
1	B	741	Total 5782	C 3707	N 989	O 1049	S 37	0	0
1	C	741	Total 5782	C 3707	N 989	O 1049	S 37	0	0
1	D	741	Total 5782	C 3707	N 989	O 1049	S 37	0	0

3 Residue-property plots

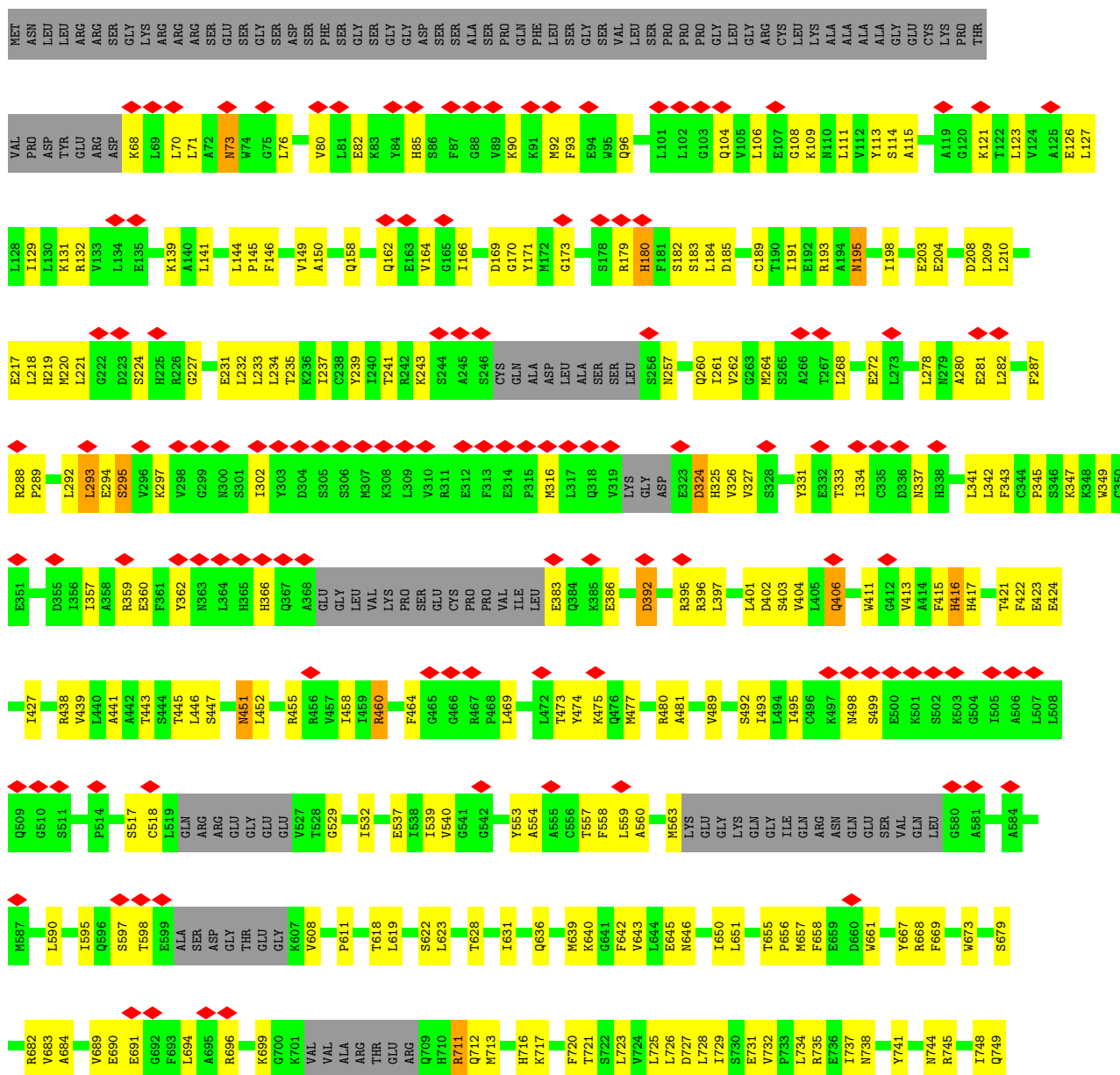
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

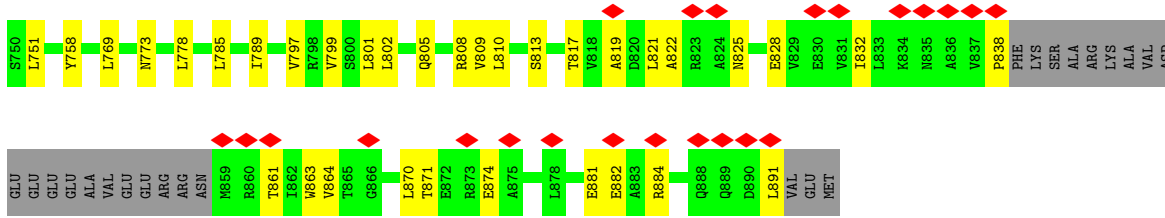
- Molecule 1: DNA polymerase theta





• Molecule 1: DNA polymerase theta





• Molecule 1: DNA polymerase theta



MET

Molecule 1: DNA polymerase theta



MET	ASN	LEU	LEU	ARG	ARG	SER	GLY	LYS	ARG	ARG	ARG	ARG	GLU	SER	GLU	SER	GLY	GLY	GLY	GLY	ASP	SER	SER	SER	ALA	ALA	SER	PRO	PRO	GLN	PHE	GLN	PHE	LEU	SER	GLY	SER	VAL	LEU	SER	PRO	PRO	PRO	PRO	GLY	GLY	LEU	LEU	ARG	CYS	LEU	LYS	ALA	ALA	ALA	ALA	ALA	GLY	GLY	CYS	LYS	PRO	THR																																																																																																																																																																																																																																																																																																																																																																																																														
VAL	PRO	ASP	TYR	GLU	ARG	ASP	K68	L69	L70	L71	A72	N73	G74	G75	L76	F77	K78	A79	V80	L81	Y84	H85	S86	F87	G88	H89	K90	R91	G92	F93	E94	N95	G96	A97	E98	C99	L100	L101	L102	G103	Q104	V105	L106	E107	G108	K109	N110	L111	A115	A119	G120	K121	T122	L123	V124	A125																																																																																																																																																																																																																																																																																																																																																																																																																					
E126	L127	L128	I129	K131	R132	V133	L134	E135	K139	A140	L141	G142	I143	L144	P145	F146	V149	A150	Q158	Q162	E163	V164	G165	I166	D169	G170	Y171	M172	G173	S178	R179	H180	F181	S182	S183	L184	C189	R193	A194	M195	I198	E203	N205	M207	D208	L209	L210	G211	M212	E217	H219	M220	L221	Q222	D223	S224	H225	L230	E231	I143	L144	P145	F146	T235	A236	C238	Y239	I240	T241	R242	K243	S244	A245	S246	CYS	GLN	ALA	ASP	LEU	ALA	SER	LEU	SER	S256	M257	Q260	I261	V262	G263	A266	T267	L268	L273	L276	M279	A280	E281	L282	Y283	F287	R289	L292	D355	I356	I357	A358	R359	E360	F361	Y362	N363	L364	H365	H366	Q367	A368	GLY	GLY	LEU	VAL	LYS	PRO	SER	GLU	CYS	F313	P315	M316	L317	Q318	V319	LYS	ASP	E323	D324	H325	V326	V327	S328	L329	C330	Y331	E332	T333	I334	C335	D336	N337	H338	S339	V340	L341	L342	F343	C344	P345	S346	K347	K348	C350	E351	D355	I356	I357	A358	R359	E360	F361	Y362	N363	L364	H365	H366	Q367	A368	GLY	GLY	LEU	VAL	LYS	PRO	SER	GLU	CYS	F313	P315	M316	L317	Q318	V319	LYS	ASP	E323	D324	H325	V326	V327	S328	L329	C330	Y331	E332	T333	I334	C335	D336	N337	H338	S339	V340	L341	L342	F343	C344	P345	S346	I427	I428	E429	R438	V439	L440	A441	A442	T443	S444	T445	L446	S447	N451	L452	R455	R456	V457	I458	I459	R460	F464	G465	G466	R467	L468	L469	L472	K475	R480	A481	E491	S492	I493	L494	I495	C496	K497	N498	A499	E500	K501	S502	K503	A506	L507	L508	Q509	G510	S511	P514	S517	C518	L519	GLN	ARG	ARG	ARG	GLY	GLY	GLY	V527	T528	G529	I532	L536	E537	I538	I539	V540	G541	G542	D549	Y553	A554	A555	C556	T557	F558	L559	A560	M563	LYS	GLY	LYS	LYS	GLN	GLY	ILE	GLN	ARG	ASN	GLN	SER	VAL	GLN	GLY	G580	A581	A584	M587	L590	E593	F594	I595	Q596	S597	T598	E599	ALA	SER	ASP	GLY	THR	GLY	K607	V608	P611	T618	L619	S622	L623	T628	L629	R630	I631	Q636	M639	K640	F642	V643	M646	D647	L648	H649	I650	T655	P656	M657	F658	E659	D660	T662	Y667	R668	F669	S679	R682	V683	A684	E685	L686	V689	E690	E691	G692	F693	L694	A695	R696	K699	G700	K701	VAL	VAL	ALA	ARG	THR	GLY	ARG	GLU	ARG	Q709	H710	R711	A822	M713	H716	K717	F720	T721	L725	L726	D727	L728	V732	P733	L734	T737	N738	Y741	N744	R745	G746	Q747	I748	Q749	S750	L751	Y758	N773	L778	L785	I789	V797	R798	V799	S800	L801	L802	Q805	V809	L810	S813	V817	V818	A819	D820	L821	A822	R823	A824	N825	E828	V829	E830	V831	L832	L833	K834	N835	A836	V837	P838	PHE	LYS	SER	ALA	ARG	LYS	ALA	VAL	ASP	GLU	GLU

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	202004	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS GLACIOS	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	58	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	150000	Depositor
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	0.925	Depositor
Minimum map value	-0.564	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.015	Depositor
Recommended contour level	0.12	Depositor
Map size (Å)	368.0, 368.0, 368.0	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.92, 0.92, 0.92	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	A	0.26	0/5882	0.51	0/7956
1	B	0.26	0/5882	0.52	0/7956
1	C	0.26	0/5882	0.52	0/7956
1	D	0.26	0/5882	0.52	0/7956
All	All	0.26	0/23528	0.52	0/31824

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5782	0	5868	174	0
1	B	5782	0	5868	192	0
1	C	5782	0	5868	172	0
1	D	5782	0	5868	180	0
All	All	23128	0	23472	709	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 15.

The worst 5 of 709 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)
1:B:93:PHE:H	1:B:96:GLN:HE21	1.23	0.84
1:B:132:ARG:HH21	1:B:260:GLN:HE21	1.23	0.82
1:C:326:VAL:HG21	1:C:460:ARG:HH12	1.44	0.82
1:C:109:LYS:NZ	1:C:281:GLU:OE2	2.16	0.79
1:D:219:HIS:HB3	1:D:268:LEU:HD11	1.65	0.79

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	A	723/894 (81%)	691 (96%)	32 (4%)	0	100	100
1	B	723/894 (81%)	682 (94%)	41 (6%)	0	100	100
1	C	723/894 (81%)	691 (96%)	32 (4%)	0	100	100
1	D	723/894 (81%)	681 (94%)	42 (6%)	0	100	100
All	All	2892/3576 (81%)	2745 (95%)	147 (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	A	626/766 (82%)	604 (96%)	22 (4%)	36	67
1	B	626/766 (82%)	601 (96%)	25 (4%)	31	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	626/766 (82%)	599 (96%)	27 (4%)	29	62
1	D	626/766 (82%)	601 (96%)	25 (4%)	31	64
All	All	2504/3064 (82%)	2405 (96%)	99 (4%)	35	64

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	343	PHE
1	C	623	LEU
1	C	362	TYR
1	C	455	ARG
1	D	73	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	498	ASN
1	D	96	GLN
1	C	712	GLN
1	D	180	HIS
1	B	96	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

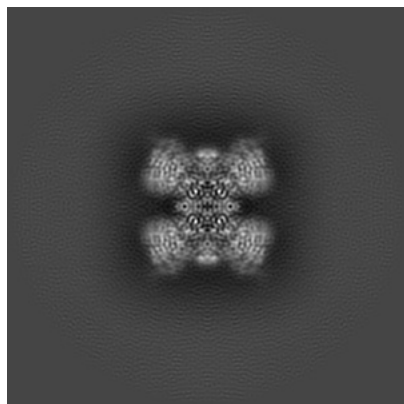
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-43818. These allow visual inspection of the internal detail of the map and identification of artifacts.

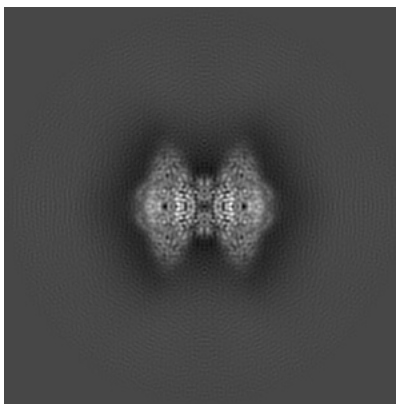
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections [i](#)

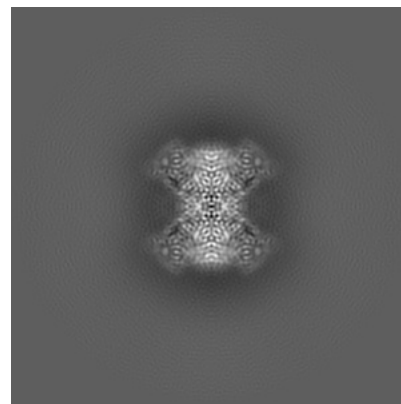
6.1.1 Primary map



X

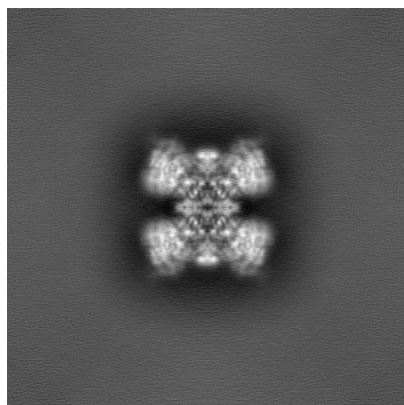


Y

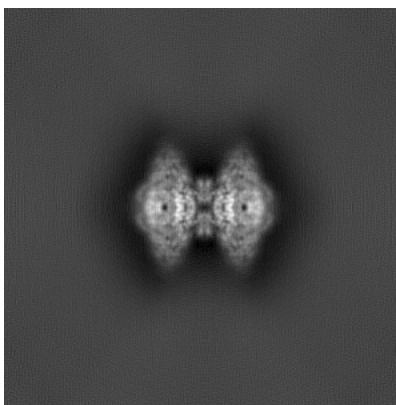


Z

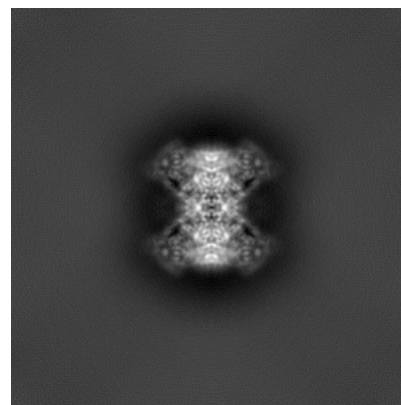
6.1.2 Raw map



X



Y

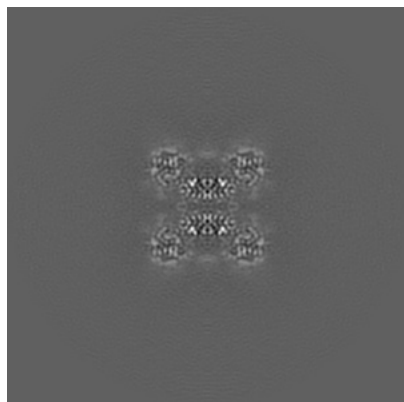


Z

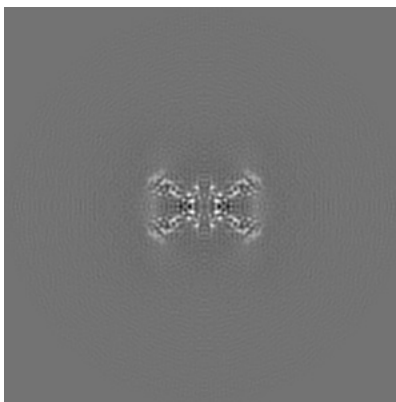
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

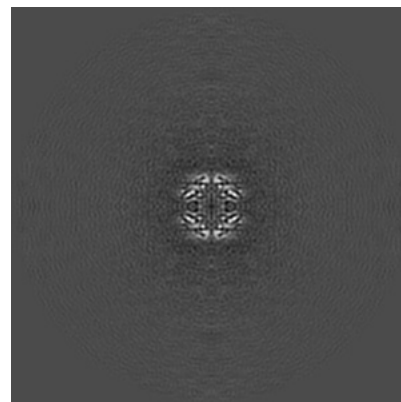
6.2.1 Primary map



X Index: 200

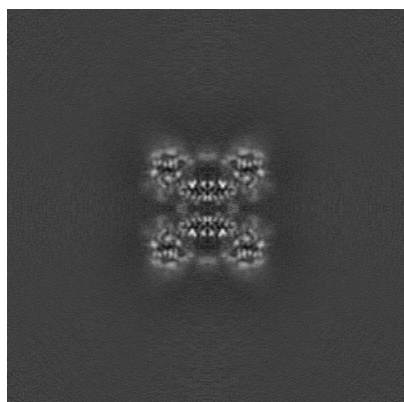


Y Index: 200

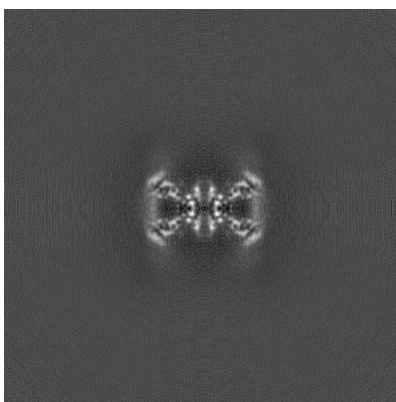


Z Index: 200

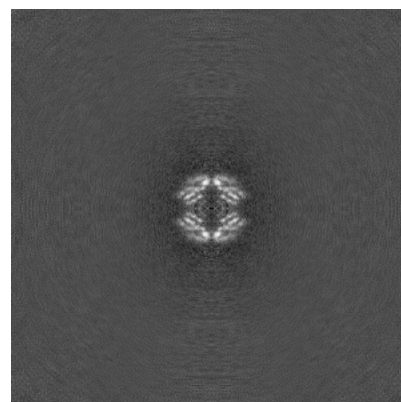
6.2.2 Raw map



X Index: 200



Y Index: 200

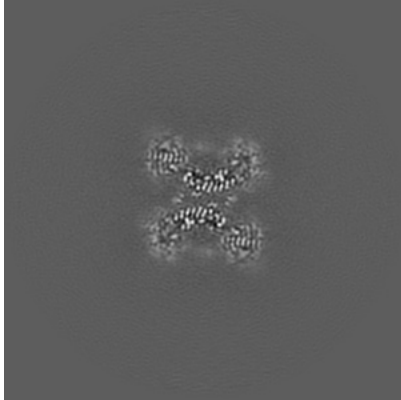


Z Index: 200

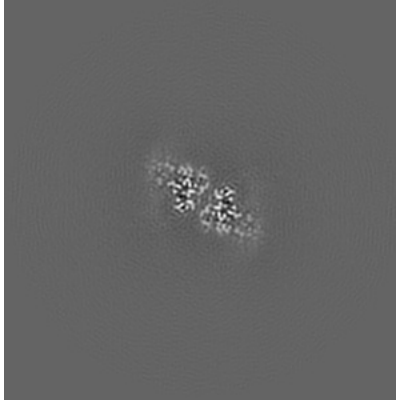
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

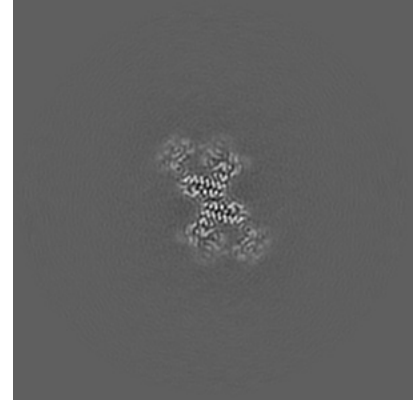
6.3.1 Primary map



X Index: 195

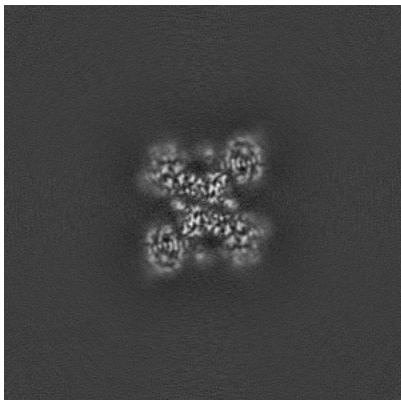


Y Index: 188

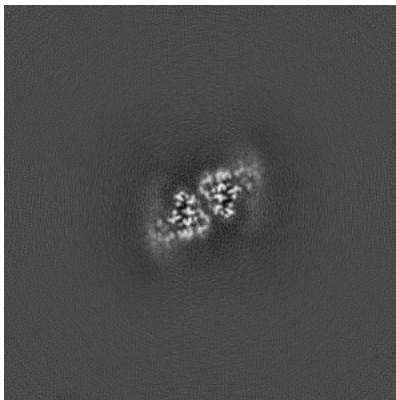


Z Index: 176

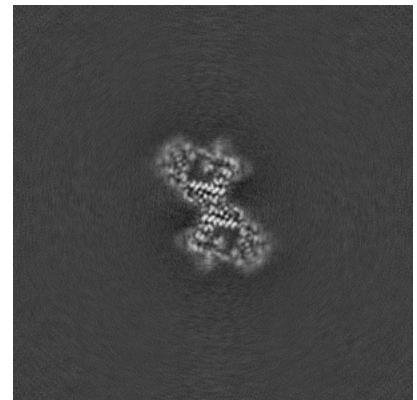
6.3.2 Raw map



X Index: 207



Y Index: 212

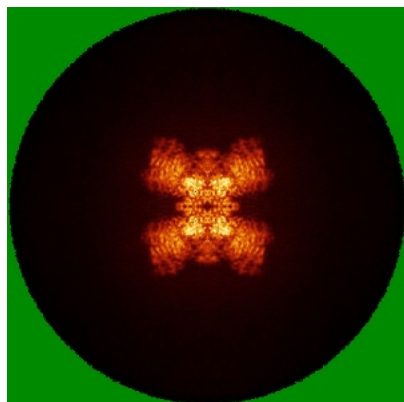


Z Index: 174

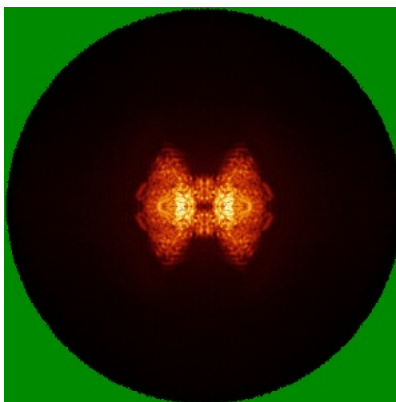
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) [i](#)

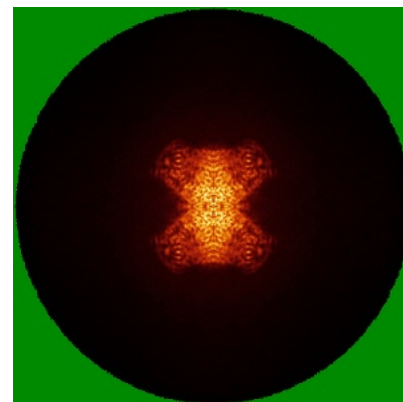
6.4.1 Primary map



X

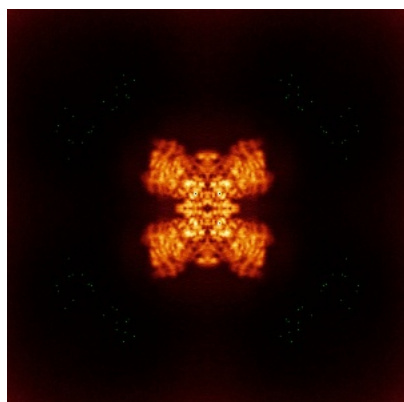


Y

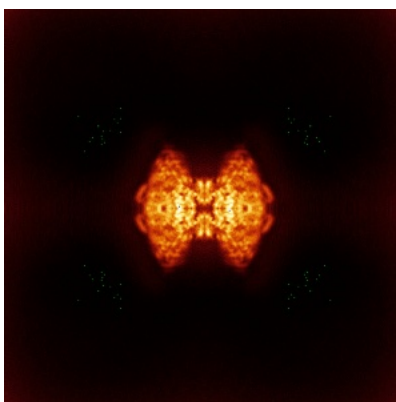


Z

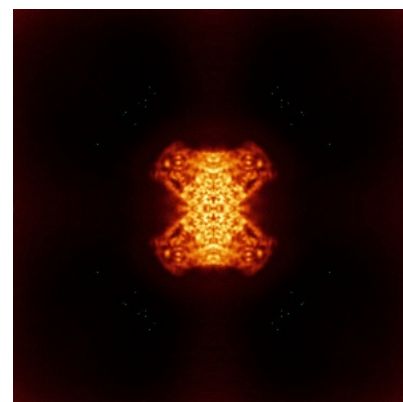
6.4.2 Raw map



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.

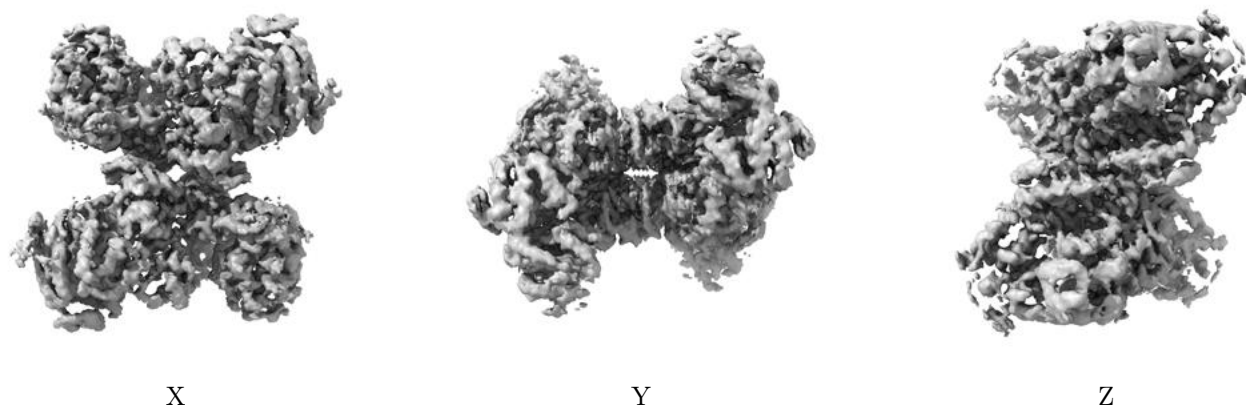
6.5 Orthogonal surface views [i](#)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.12. 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



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.

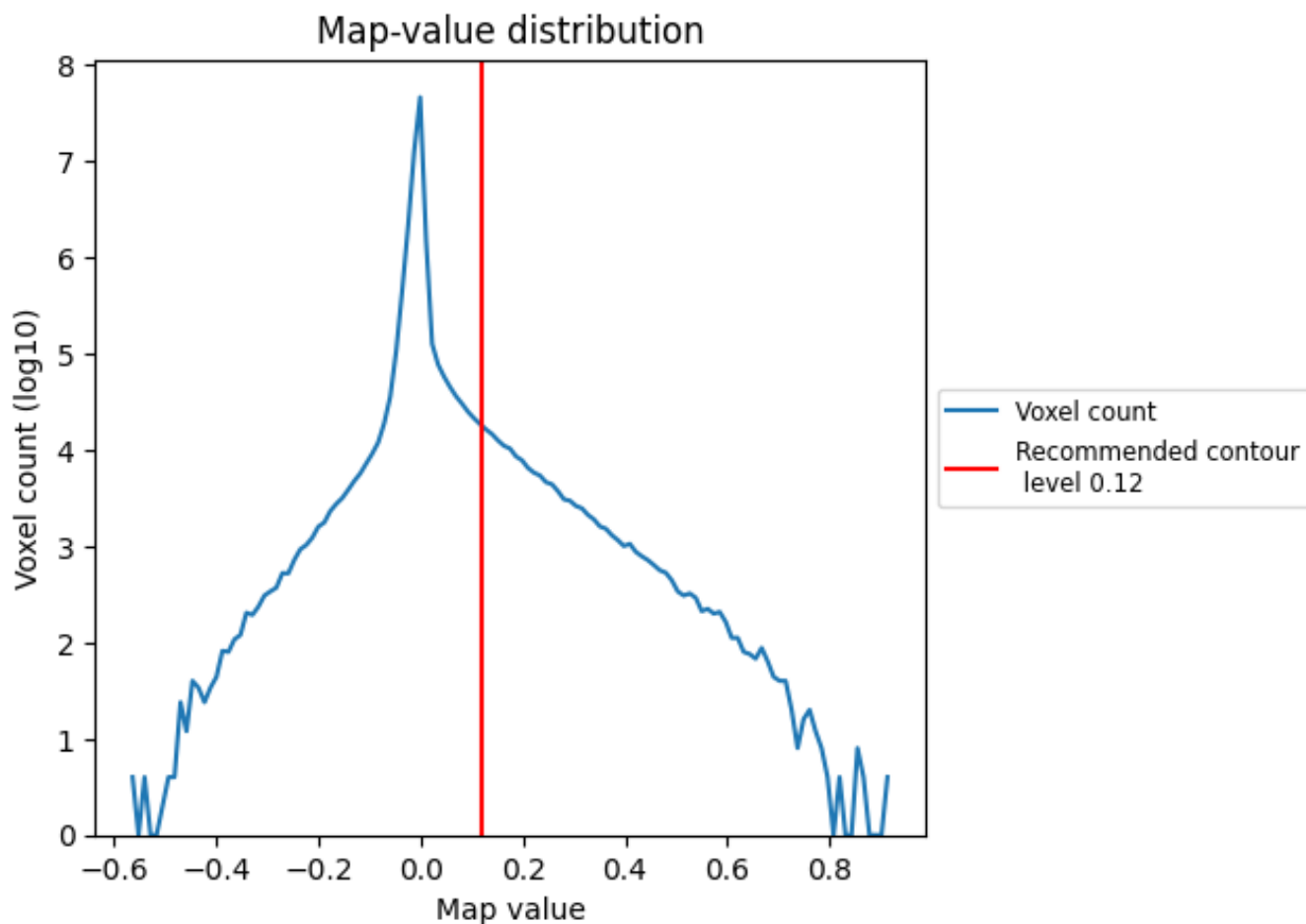
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

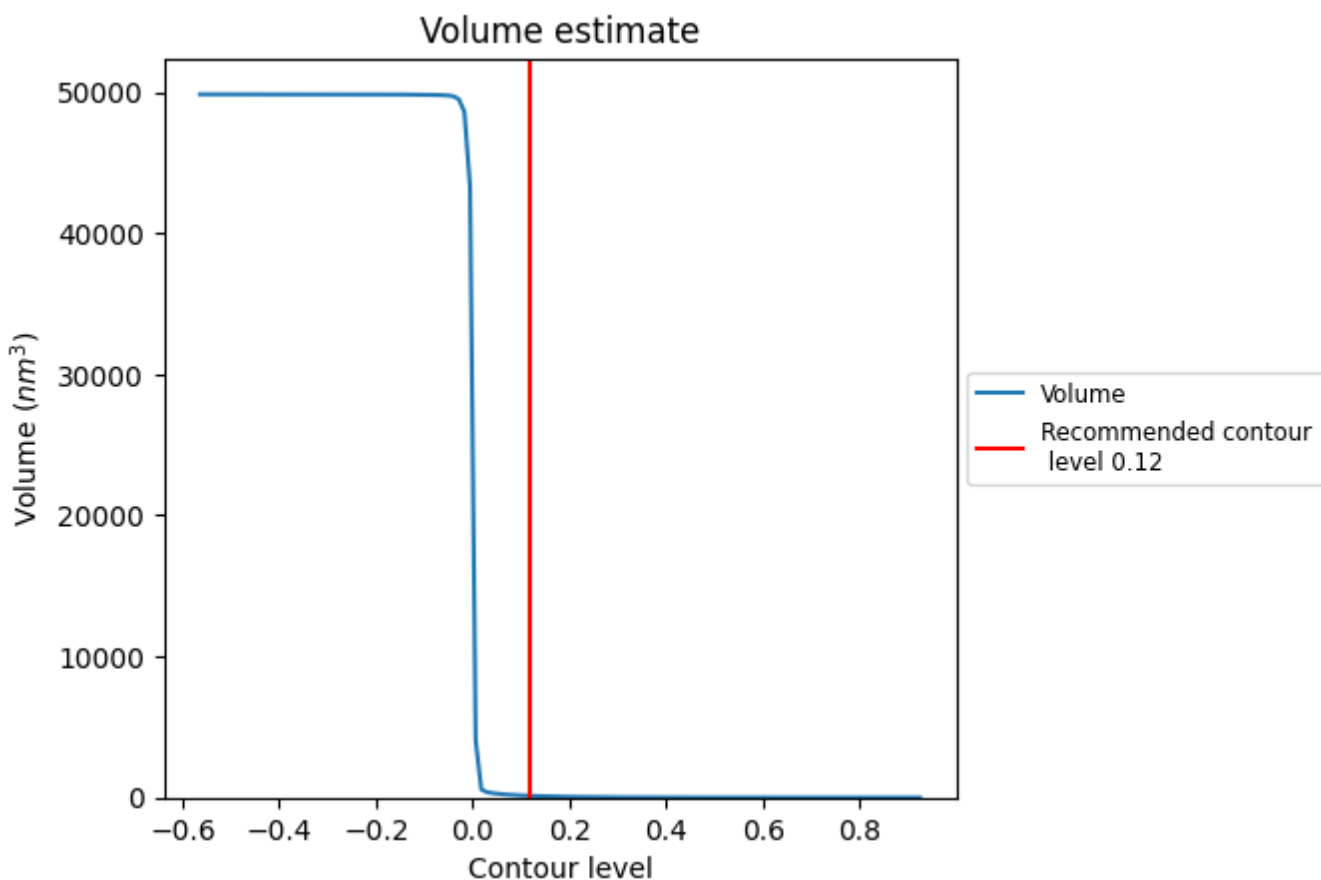
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution [i](#)



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

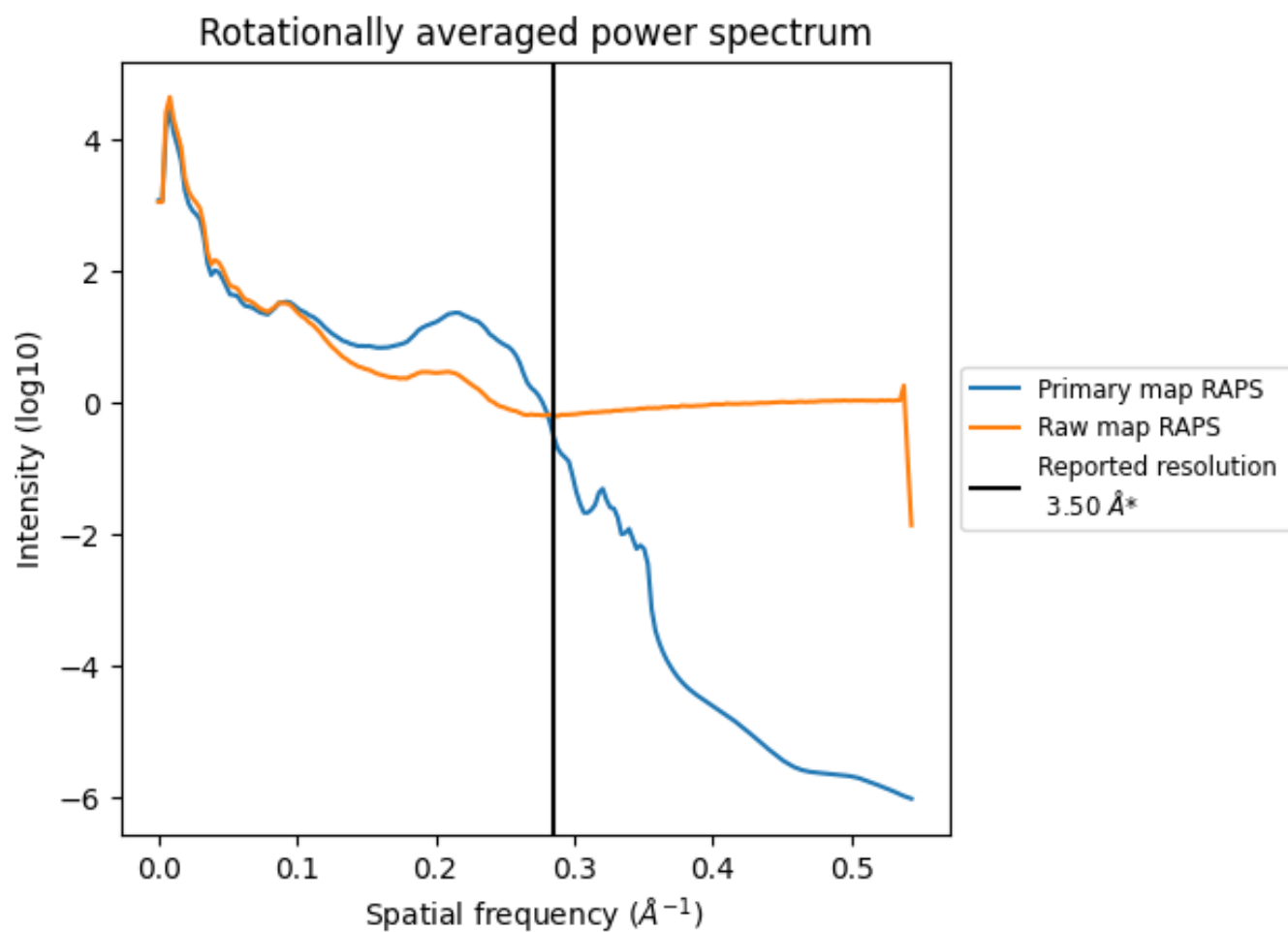
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 121 nm³; this corresponds to an approximate mass of 109 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum i

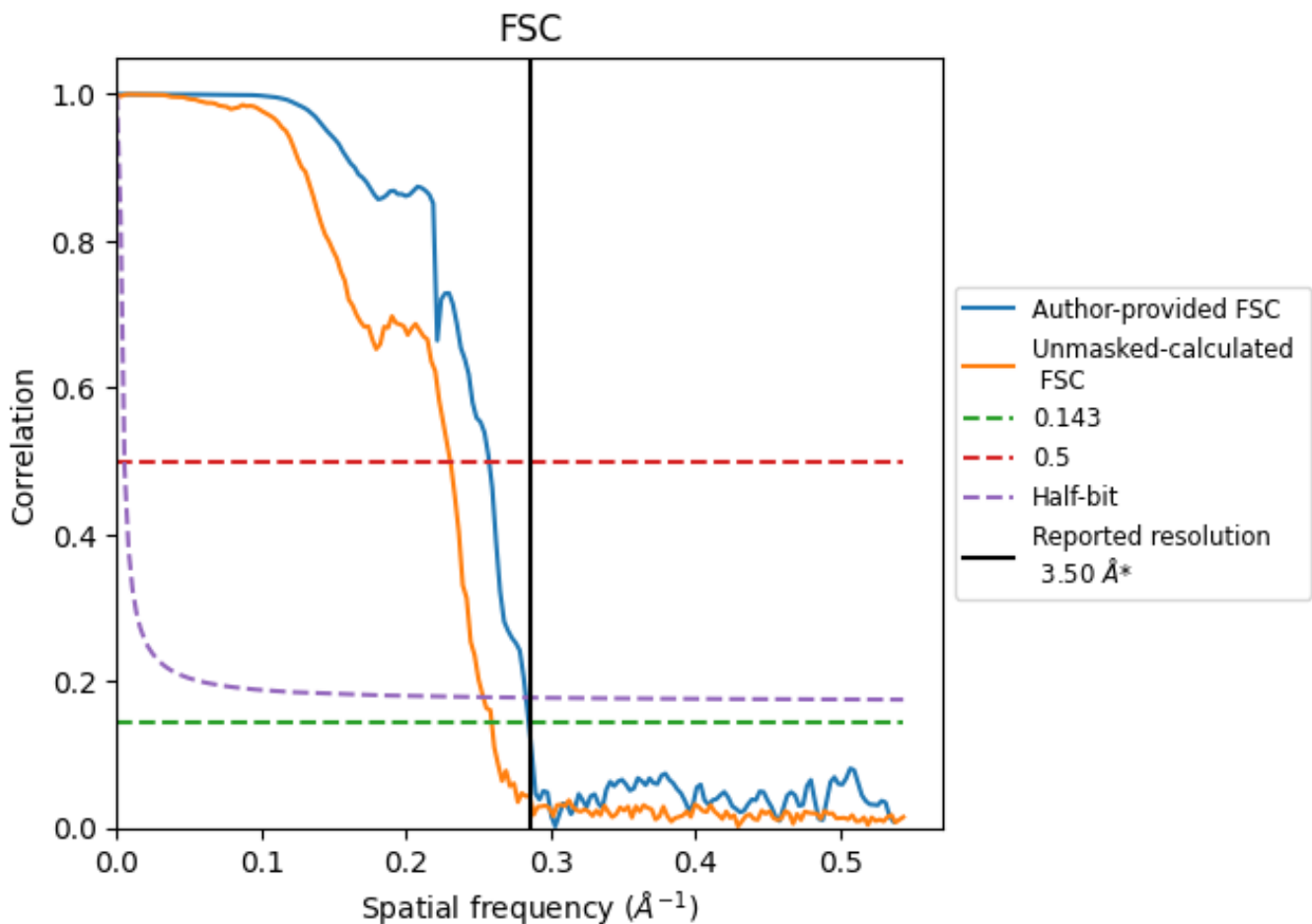


*Reported resolution corresponds to spatial frequency of 0.286 Å⁻¹

8 Fourier-Shell correlation [i](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

8.1 FSC [i](#)



*Reported resolution corresponds to spatial frequency of 0.286 Å⁻¹

8.2 Resolution estimates [i](#)

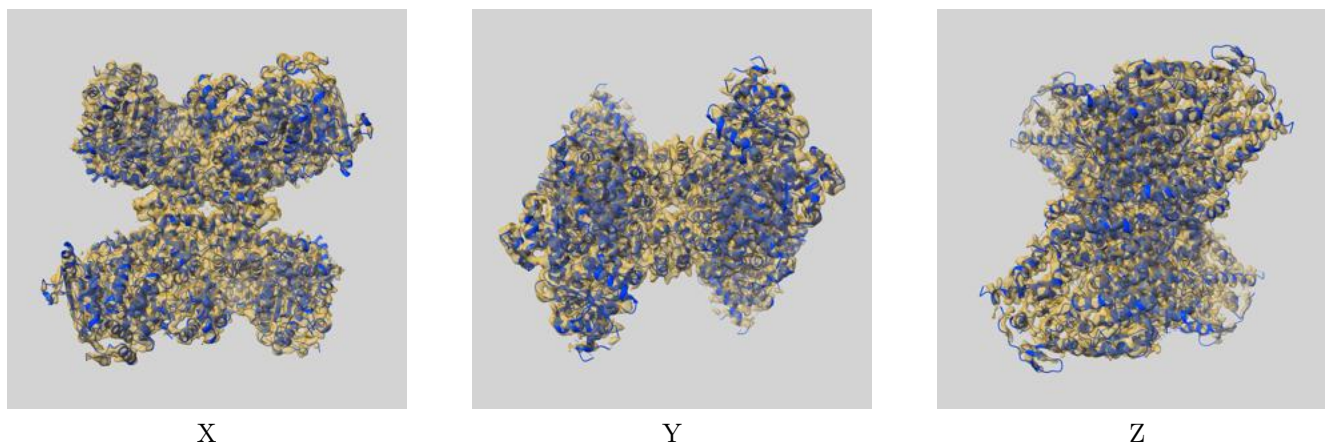
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.50	-	-
Author-provided FSC curve	3.51	3.89	3.54
Unmasked-calculated*	3.86	4.34	3.95

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.86 differs from the reported value 3.5 by more than 10 %

9 Map-model fit [i](#)

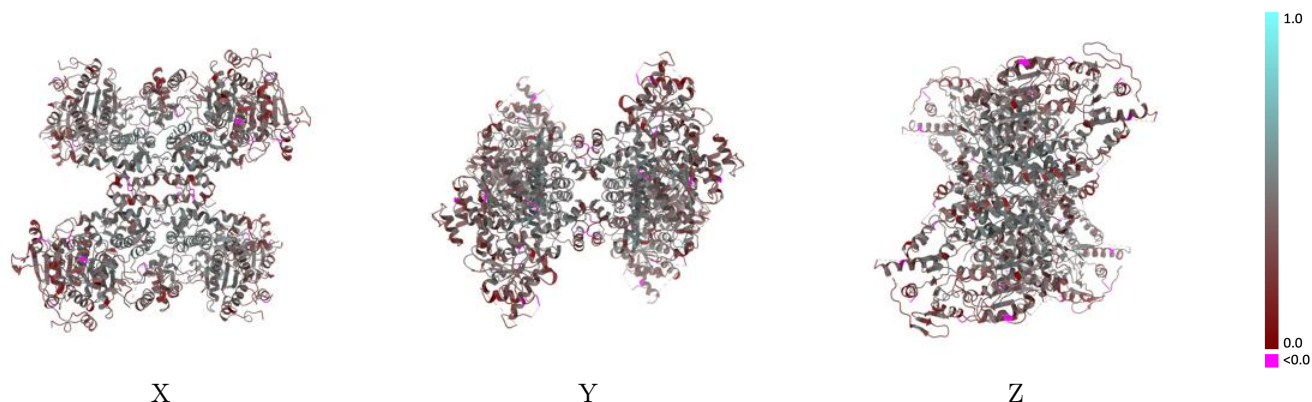
This section contains information regarding the fit between EMDB map EMD-43818 and PDB model 9ASL. Per-residue inclusion information can be found in section [3](#) on page [4](#).

9.1 Map-model overlay [i](#)



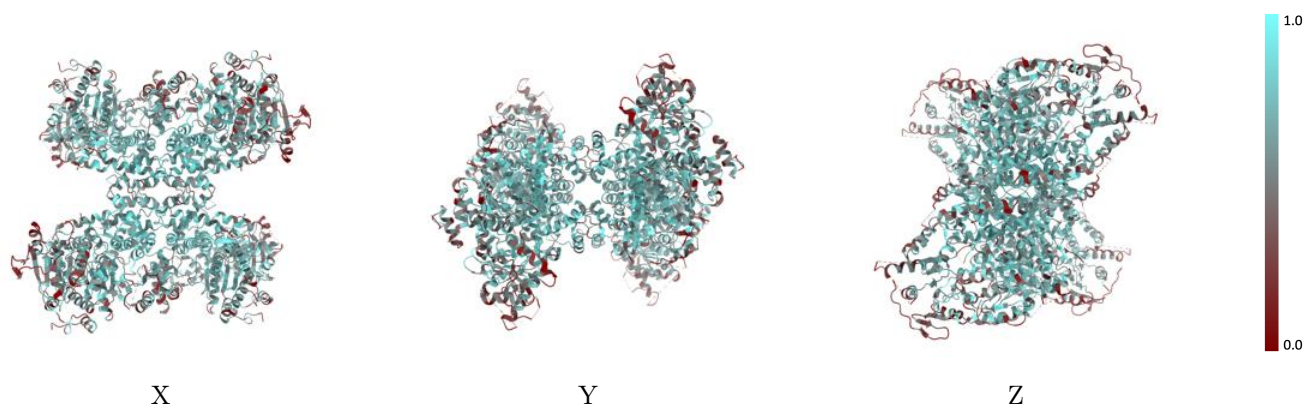
The images above show the 3D surface view of the map at the recommended contour level 0.12 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model [i](#)



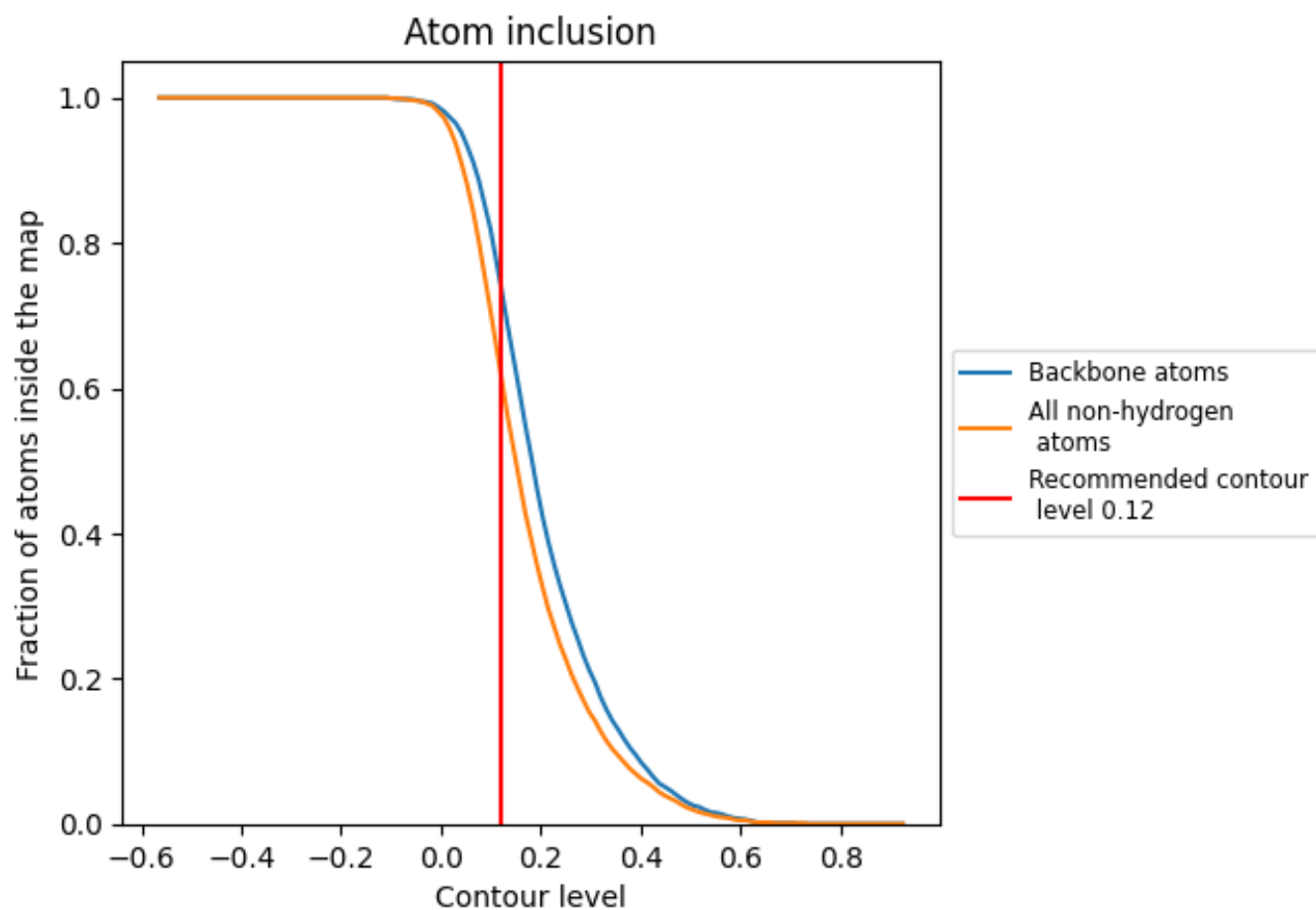
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

9.3 Atom inclusion mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.12).











9.4 Atom inclusion [i](#)



At the recommended contour level, 74% of all backbone atoms, 62% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary [i](#)

The table lists the average atom inclusion at the recommended contour level (0.12) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.6190	 0.3920
A	 0.6180	 0.3930
B	 0.6190	 0.3900
C	 0.6190	 0.3940
D	 0.6200	 0.3930

