



wwPDB EM Validation Summary Report ⓘ

Feb 24, 2024 – 10:20 AM EST

PDB ID : 7K1B
EMDB ID : EMD-22623
Title : CryoEM structure of DNA-PK catalytic subunit complexed with DNA (Complex II)
Authors : Chen, X.; Gellert, M.; Yang, W.
Deposited on : 2020-09-07
Resolution : 4.30 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.dev70
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) : 2.36

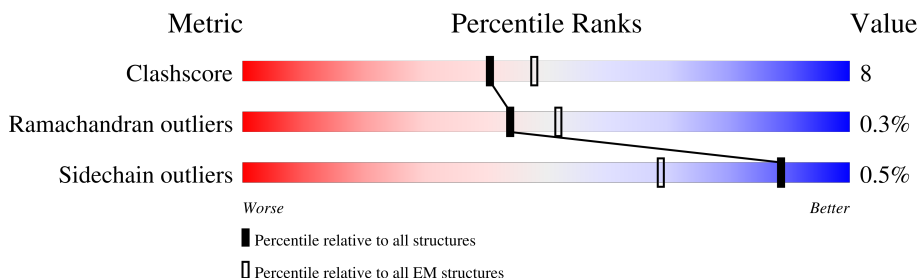
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.30 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	4128	
2	D	24	
2	F	24	
3	G	16	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 29321 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-dependent protein kinase catalytic subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	3589	28343	18183	4807	5168	185	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(P*GP*CP*AP*TP*GP*CP*TP*CP*TP*AP*CP*TP*GP*CP*TP*TP*CP*GP*AP*TP*AP*TP*CP*G)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	D	24	484	233	82	146	23	0	0
2	F	8	164	78	30	48	8	0	0

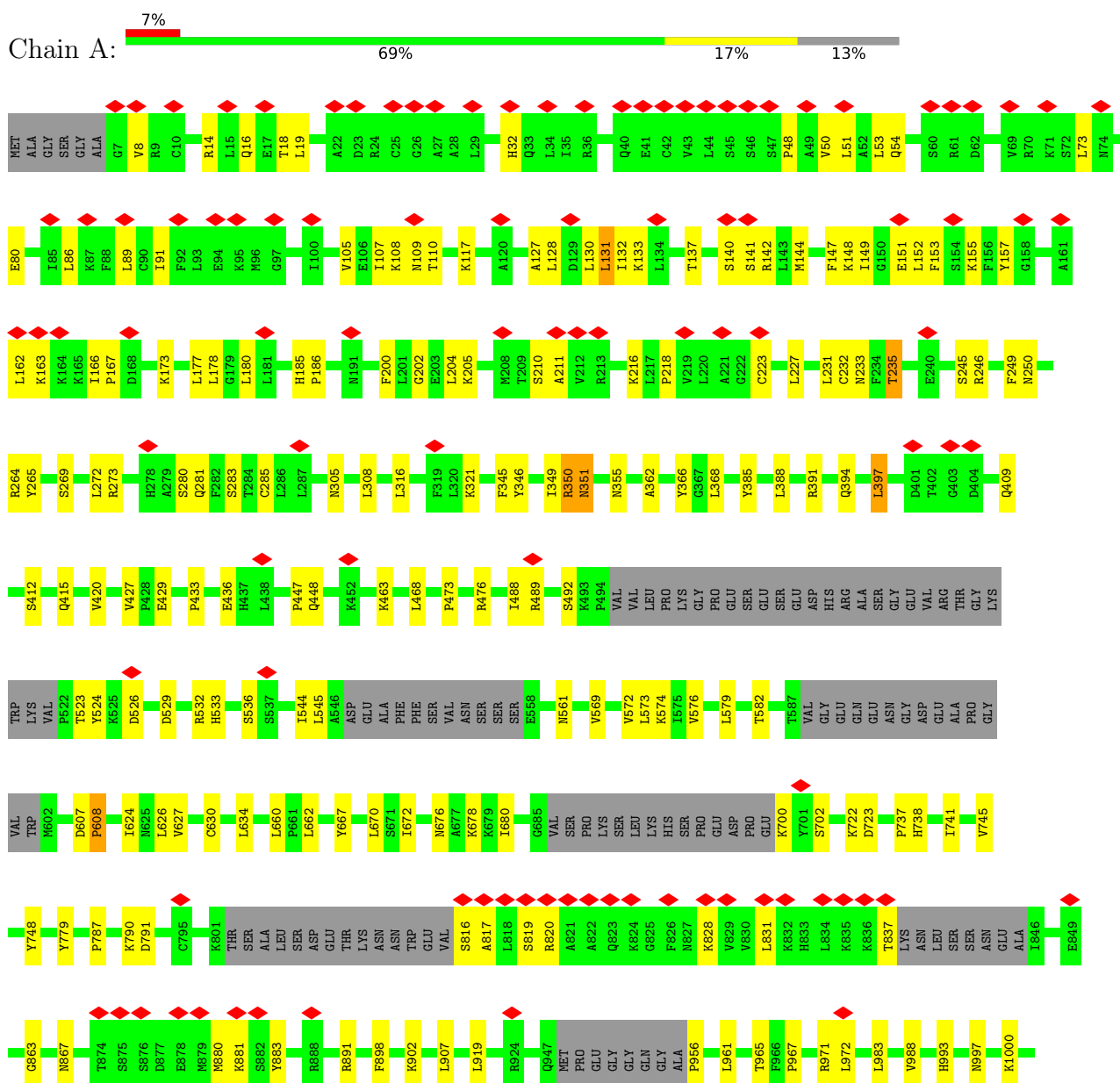
- Molecule 3 is a DNA chain called DNA (5'-D(P*AP*AP*GP*CP*AP*GP*TP*AP*GP*AP*GP*CP*AP*TP*GP*C)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	G	16	330	157	68	90	15	0	0

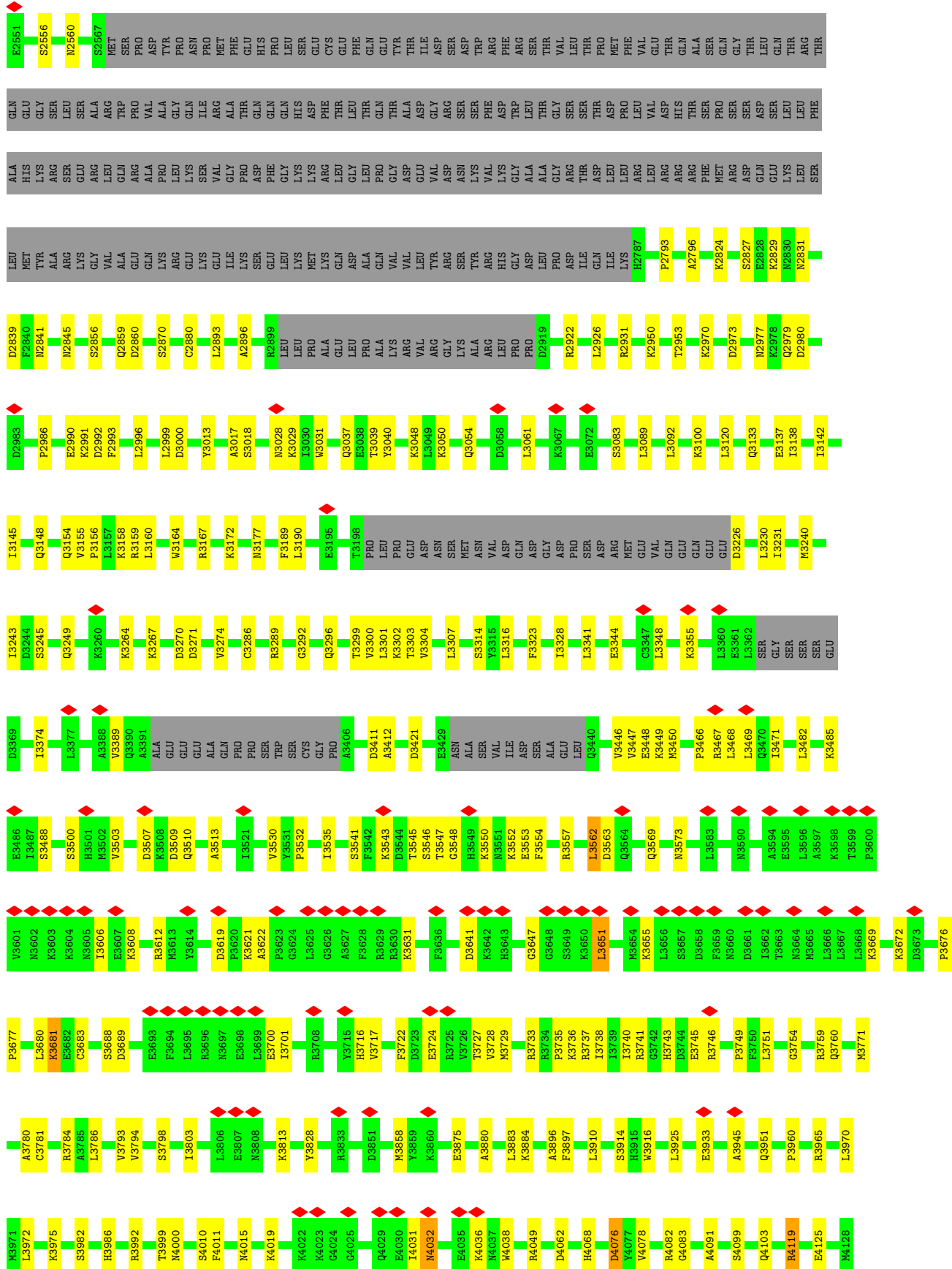
3 Residue-property plots

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.

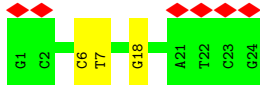
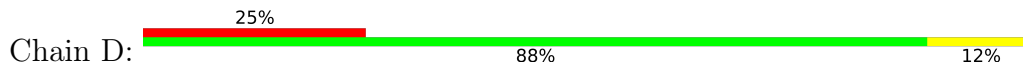
- Molecule 1: DNA-dependent protein kinase catalytic subunit



SI003	SI004	DI005	TI006	L1014	II017	VI018	DI019	SI023	RI026	DI027	F1028	RI034	E1035	Q1043	TI044	TI045	K1051	RI067	LI068	HI069	PI070	MI071	RI075	LI076	GI077	AI078	SI079	F1082	MI083	NI084	LI085	YI086	F1089	L1095	VI096	EI097	QI098	F1099	LI104	DI117	LI134	CI135	II138																																																
A1148	R1151	R1162	L1163	II017	DI167	W1171	F1194	VI195	PI196	LI197	LI198	LI198	GI200	P1204	V1211	K1213	V1217	GI230	GLN	PRO	SER	GLY	ILE	ALA	ALA	GLN	THR	L1241	F1246	S1249	L1261	A1262	A1263	L1264	I1271	L1279	L1282	GI283	THR	GLU	ALA	GLN	S1288																																																
LI298	MI303	HIS	ASP	ILE	ILE	ALA	GLU	LYS	CYS	PHE	GLY	THR	GLY	ALA	GLY	ASN	ARG	THR	ALA	ALA	P1324	Q1325	E1328	K1334	V1337	I1341	TI347	L1348	L1349	MI350	TI351	L1358	L1359	D1362	L1363	TI366	HI367	RI370	II382	GI387	Q1390	DI397	DI397																																																
K1404	SI409	DI413	II414	LI415	E1416	TI417	HI418	LI419	RI420	E1421	K1422	L1431	VI434	MI435	LI436	DI440	AI441	Q1442	VI443	DI444	RI445	RI447	K1456	Q1457	LI458	HI459	RI460	M1466	II467	TI473	DI474	LI475	HI476	VI487	II491	AI492	P1493	GLY	ASP	GLU	ARG	GLN	CYS	LEU	P1501	DI504	LI505																																												
SI506	CI507	K1508	SI512	L1515	F1519	A1520	RI527	SI530	NI534	PRO	ALA	VAL	LEU	SER	THR	ALA	ALA	SER	LEU	GLY	SER	SER	GLN	GLY	SER	VAL	ILE	HIS	PHE	SI554	L1562	F1563	SI564	M1568	LI572	Q1614	TI620	L1623	Q1624	W1632	SI637	P1638	L1639	E1640	TI641	K1642																																													
MI643	AI644	VI645	K1651	Q1654	II655	DI656	SI657	VI659	SI660	L1678	L1679	A1680	K1683	MI684	L1684	DI685	LI686	HI687	GI690	HI691	AI692	VI693	TI694	LI695	LI696	F1699	GI704	L1707	P1723	RI727	TI733	P1734	RI735	C1742	L1747	Q1754	L1758	M1762	RI768	Q1771																																																			
LI777	F1778	SI780	SI781	F1782	RI783	RI787	RI788	GI789	E1799	SI800	VI801	RI802	MI803	MI804	RI806	DI809	P1810	F1819	SI823	LI824	LI825	TI826	LI827	LI828	HI829	HI830	C1831	SI832	LI833	DI834	AI835	LI836	RI837	II843	RI854	F1855	L1858	F1863	DI864	II867	RI883	LI884	RI885	K1886	DI887	DI888																																													
VI889	HI890	K1895	VI899	F1900	HI901	GI902	SI903	CI904	TI905	TI906	E1907	GI908	MI909	E1910	LI911	TI912	K1913	K1917	GI929	E1930	MI931	LI934	AI943	MI946	F1956	MI957	E1958	LI959	Q1963	GI964	F1965	SI968	E1969	L1975	L1976	E1979	RI986	RI987	E1993	VI994	GLU	VAL	PRO	MET	GLU																																														
ARG	LYS	LYS	LYS	TYR	ILE	ILE	ARG	LYS	GLU	VAL	HIS	ASP	GLU	ALA	ALA	ASN	GLY	ASP	ASP	SER	ASP	GLY	PRO	TYR	MET	SER	SER	LEU	SER	TYR	LEU	ALA	ASP	S2034	T2035	L2036	S2037	P2044	F2045	G2048	VAL	GLN	SER	TYR	TYR	SER	GLN	ASP	PRO	PRO	ARG	ALA	ALA	GLY	ARG	GLY	VAL	TYR	PRO	GLU	VAL	TYR	PRO	PRO	THR	GLY	ARG																								
PHE	ARG	ARG	ARG	GLU	GLN	ARG	ASP	PRO	THR	VAL	HIS	ASP	ASP	VAL	ALA	E2082	L2083	E2084	M2085	D2086	E2087	L2088	M2089	E2092	A2095	T2098	V2101	K2102	H2103	L2108	GLY	PRO	PRO	GLN	GLY	GLU	GLU	ASP	VAL	PRO	E2120	S2124	W2125	L2129	H2130	G2134	M2135	P2136	T2137	L2143	L2144	F2145	K2148	R2158	K2162	L2169	A2173	N2176	N2177	G2178	G2179	E2180	G2181	T2182	H2183	V2186	V2187	E2188	T2189	V2190	L2193	L2194	T2197	A2200	L2211	A2229	R2232	H2233	N2234	I2237	L2251	R2254	F2257	E2258	K2259	F2260	G2134	M2135	P2136	T2137	L2276
R2452	Y2288	D2289	P2290	Q2291	C2292	Q2295	Q2301	V2304	N2305	N2306	V2310	R2311	M2331	R2335	E2343	Q2351	H2352	T2355	K2359	K2366	A2375	F2389	H2390	G2391	K2394	T2409	E2410	F2413	K2416	V2423	H2426	R2431	Q2432	C2435	I2438	K2445	R2452	H2464	S2466	T2467	R2470	E2471	N2475	N2476	L2477	I2480	D2482	N2483	V2484	R2485	D2486	P2487	E2488	S2489	E2490	N2493	D2494	S2495	Q2496	E2497	K2500	N2514	F2515	G2516	L2517	Q2518	L2519	N2523	F2524	E2528	T2529	R2530	L2531	F2532	S2533	D2537	R2538	A2541	L2542	N2543	S2544										



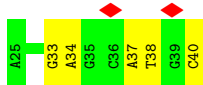
- Molecule 2: DNA (5'-D(P*GP*CP*AP*TP*GP*CP*TP*CP*TP*AP*CP*TP*GP*CP*TP*T
P*CP*GP*AP*TP*AP*TP*CP*G)-3')



- Molecule 2: DNA (5'-D(P*GP*CP*AP*TP*GP*CP*TP*CP*TP*AP*CP*TP*GP*CP*TP*T
P*CP*GP*AP*TP*AP*TP*CP*G)-3')



- Molecule 3: DNA (5'-D(P*AP*AP*GP*CP*AP*GP*TP*AP*GP*AP*GP*CP*AP*TP*GP*C
) -3')



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	46407	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 ($e^-/\text{\AA}^2$)	45	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.058	Depositor
Minimum map value	-0.032	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.002	Depositor
Recommended contour level	0.014	Depositor
Map size (Å)	408.31998, 408.31998, 408.31998	wwPDB
Map dimensions	352, 352, 352	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.16, 1.16, 1.16	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	0.32	0/28905	0.58	9/39096 (0.0%)
2	D	0.66	0/540	1.15	1/831 (0.1%)
2	F	0.65	0/183	1.08	0/280
3	G	0.66	0/372	0.91	0/573
All	All	0.34	0/30000	0.61	10/40780 (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
1	A	0	2

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	3676	PRO	C-N-CD	6.57	142.20	128.40
1	A	468	LEU	CA-CB-CG	6.29	129.77	115.30
1	A	3651	LEU	CA-CB-CG	6.18	129.52	115.30
1	A	162	LEU	CA-CB-CG	5.99	129.08	115.30
1	A	3562	LEU	CA-CB-CG	5.74	128.51	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	607	ASP	Peptide
1	A	608	PRO	Peptide

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	28343	0	28518	437	0
2	D	484	0	274	1	0
2	F	164	0	91	0	0
3	G	330	0	180	5	0
All	All	29321	0	29063	439	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 8.

The worst 5 of 439 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:A:231:LEU:O	1:A:235:THR:HA	1.53	1.06
1:A:4078:VAL:O	1:A:4082:ARG:HB2	1.73	0.87
1:A:700:LYS:N	1:A:702:SER:HG	1.76	0.84
1:A:1366:THR:O	1:A:1370:ARG:HB2	1.83	0.78
1:A:350:ARG:O	1:A:351:ASN:HB3	1.86	0.75

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	3545/4128 (86%)	3103 (88%)	433 (12%)	9 (0%)	41 76

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	351	ASN
1	A	3677	PRO
1	A	167	PRO
1	A	350	ARG
1	A	1000	LYS

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	3111/3671 (85%)	3094 (100%)	17 (0%)	88 93

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3794	VAL
1	A	4119	ARG
1	A	1213	LYS
1	A	1787	ARG
1	A	1837	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3422	GLN
1	A	3423	GLN
1	A	3590	ASN
1	A	1946	ASN
1	A	1611	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.

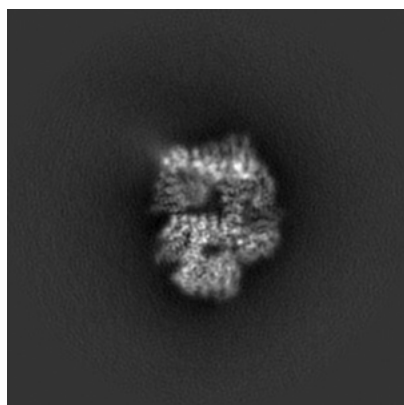
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22623. These allow visual inspection of the internal detail of the map and identification of artifacts.

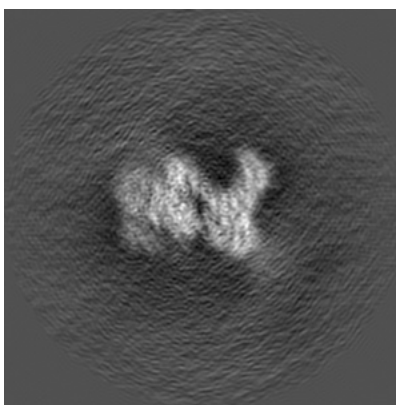
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

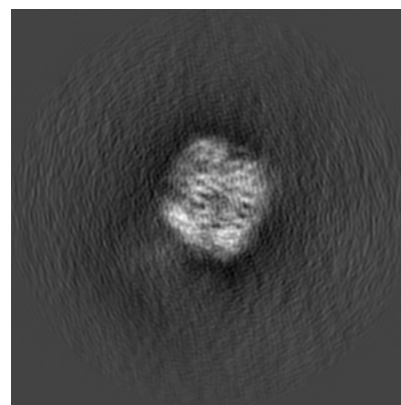
6.1.1 Primary map



X



Y

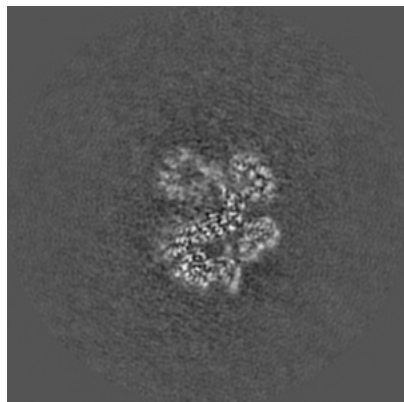


Z

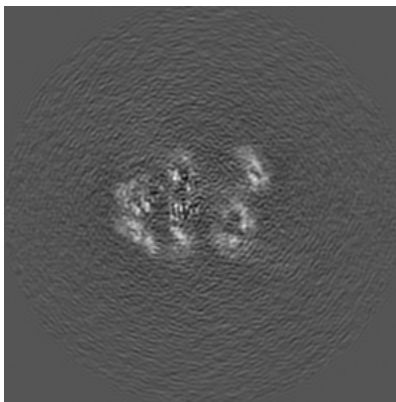
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

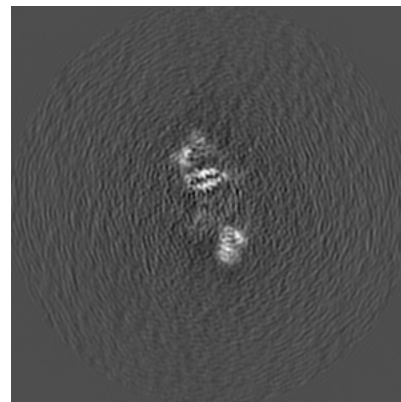
6.2.1 Primary map



X Index: 176



Y Index: 176

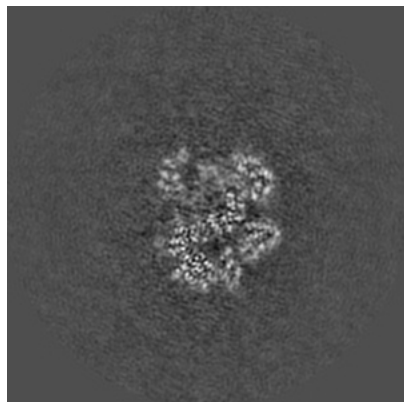


Z Index: 176

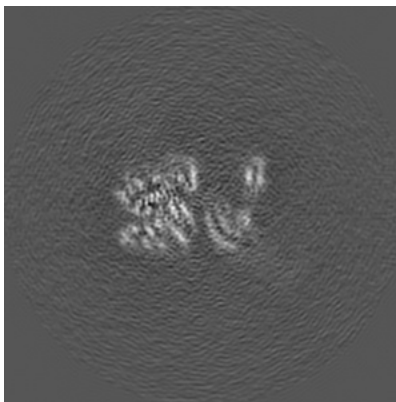
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

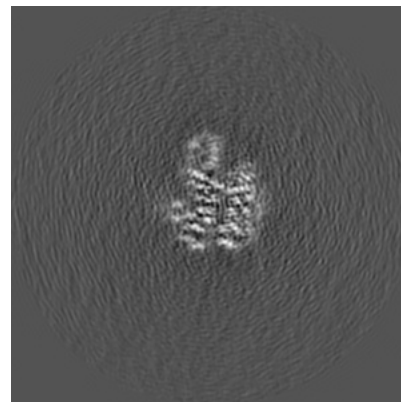
6.3.1 Primary map



X Index: 178



Y Index: 164

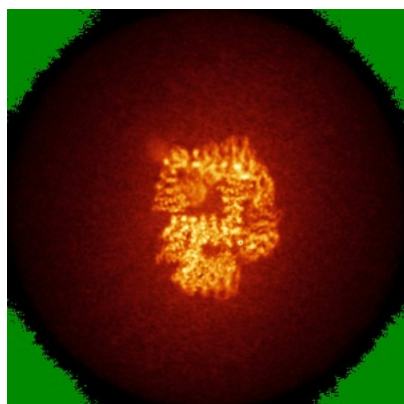


Z Index: 158

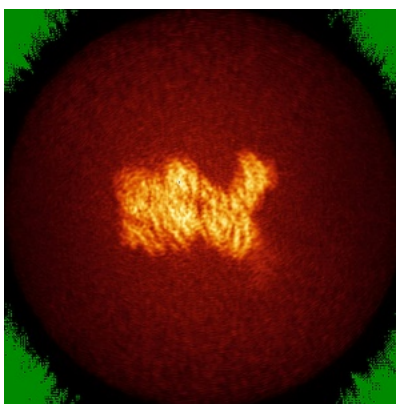
The images above show the largest variance slices of the map in three orthogonal directions.

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

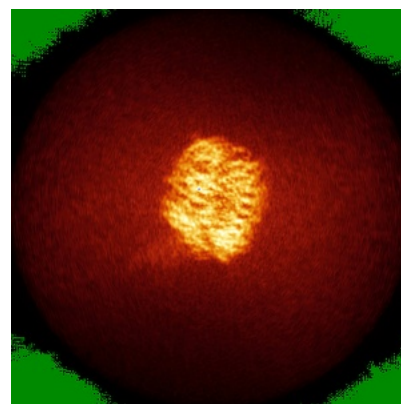
6.4.1 Primary map



X



Y

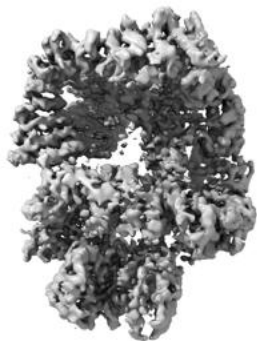


Z

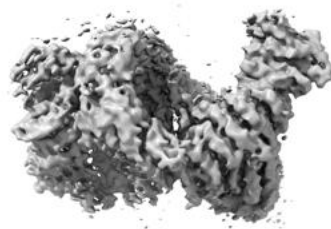
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views [i](#)

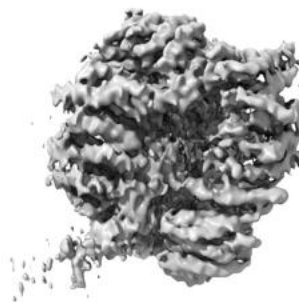
6.5.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.014. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

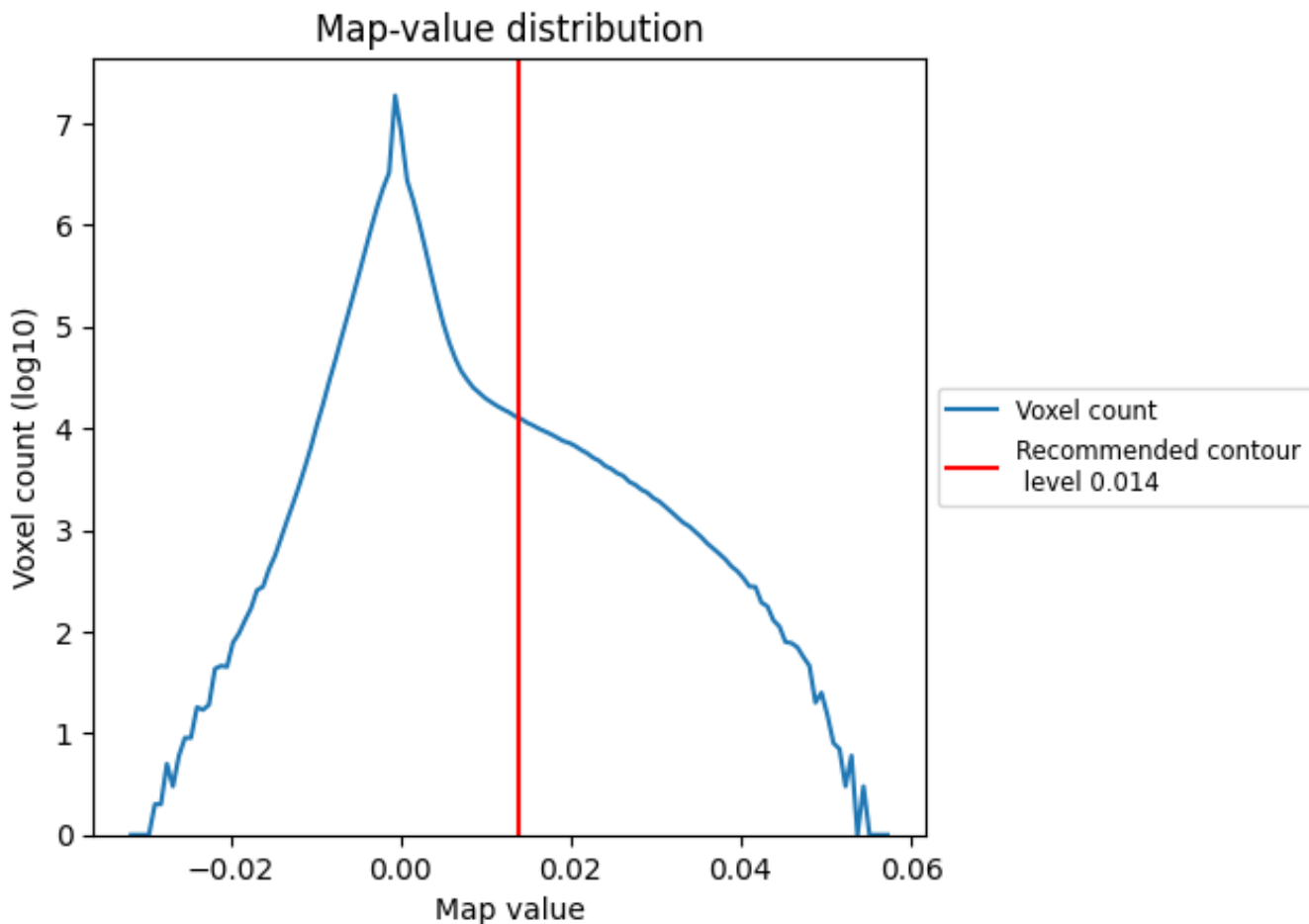
6.6 Mask visualisation [i](#)

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

7 Map analysis [i](#)

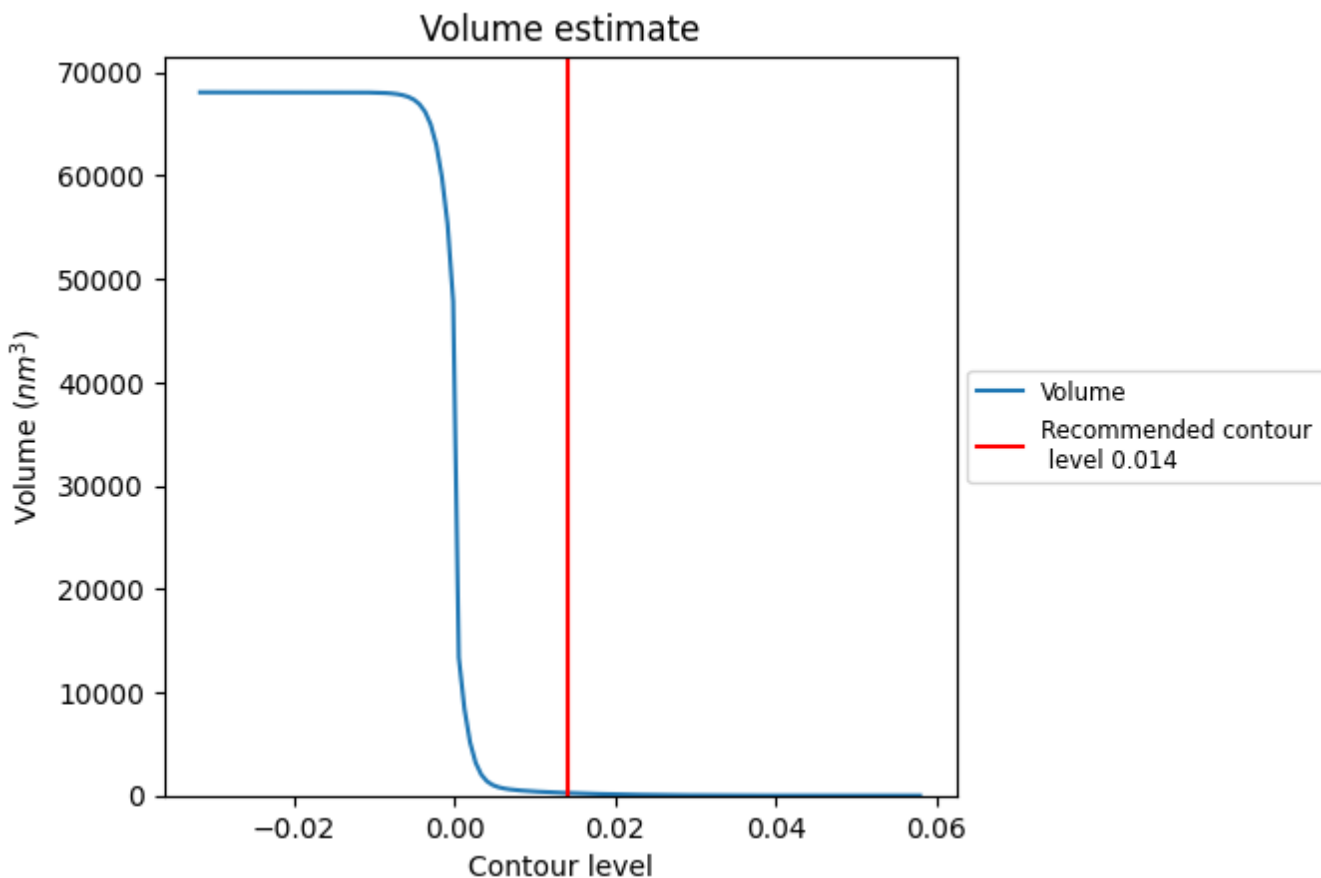
This section contains the results of statistical analysis of the map.

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.

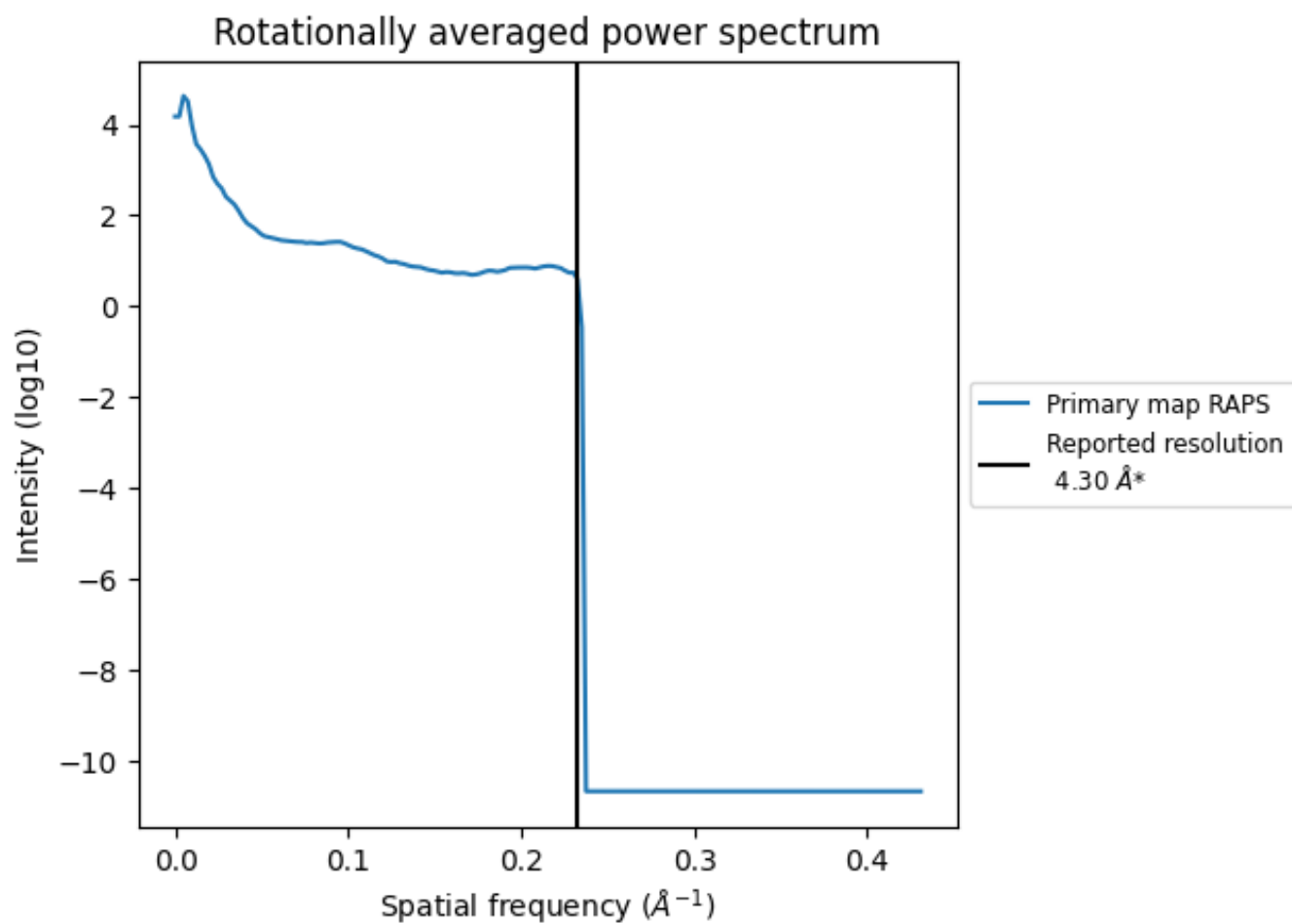
7.2 Volume estimate [i](#)



The volume at the recommended contour level is 250 nm³; this corresponds to an approximate mass of 226 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 0.233 Å⁻¹

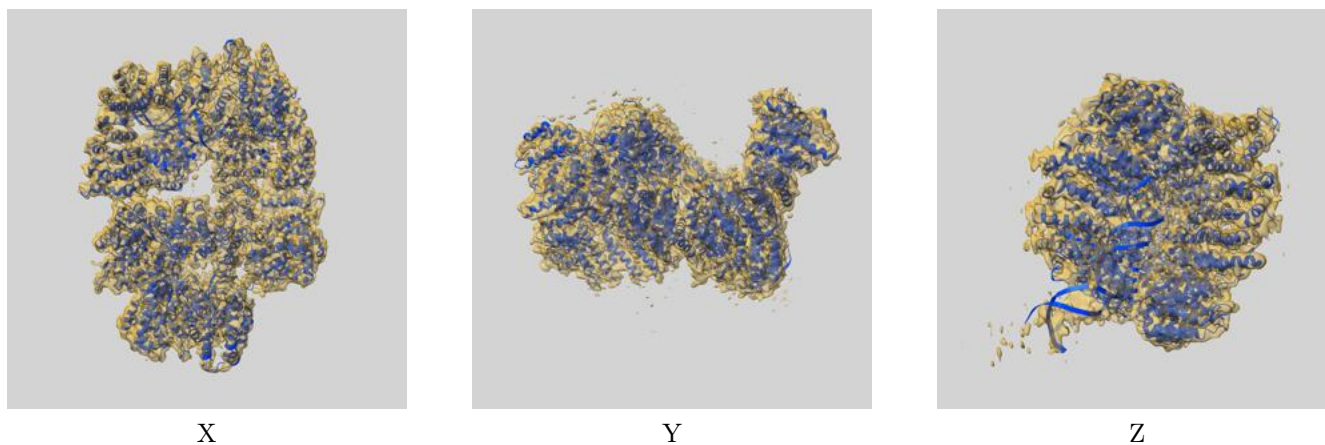
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

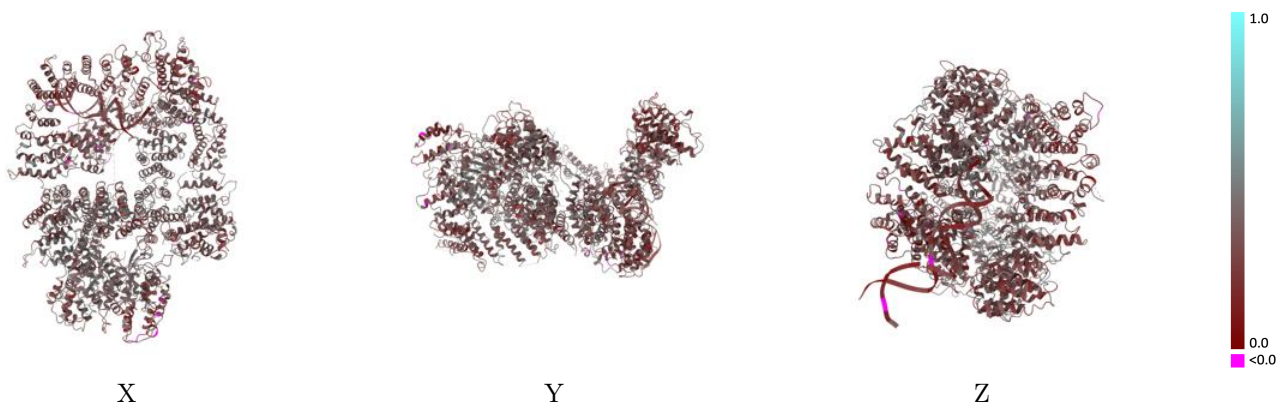
This section contains information regarding the fit between EMDB map EMD-22623 and PDB model 7K1B. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlay [i](#)



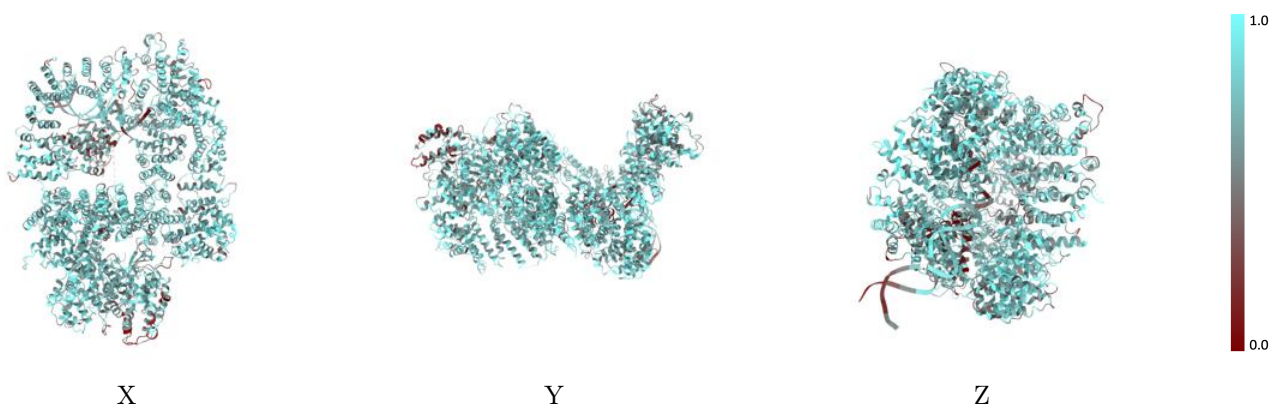
The images above show the 3D surface view of the map at the recommended contour level 0.014 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



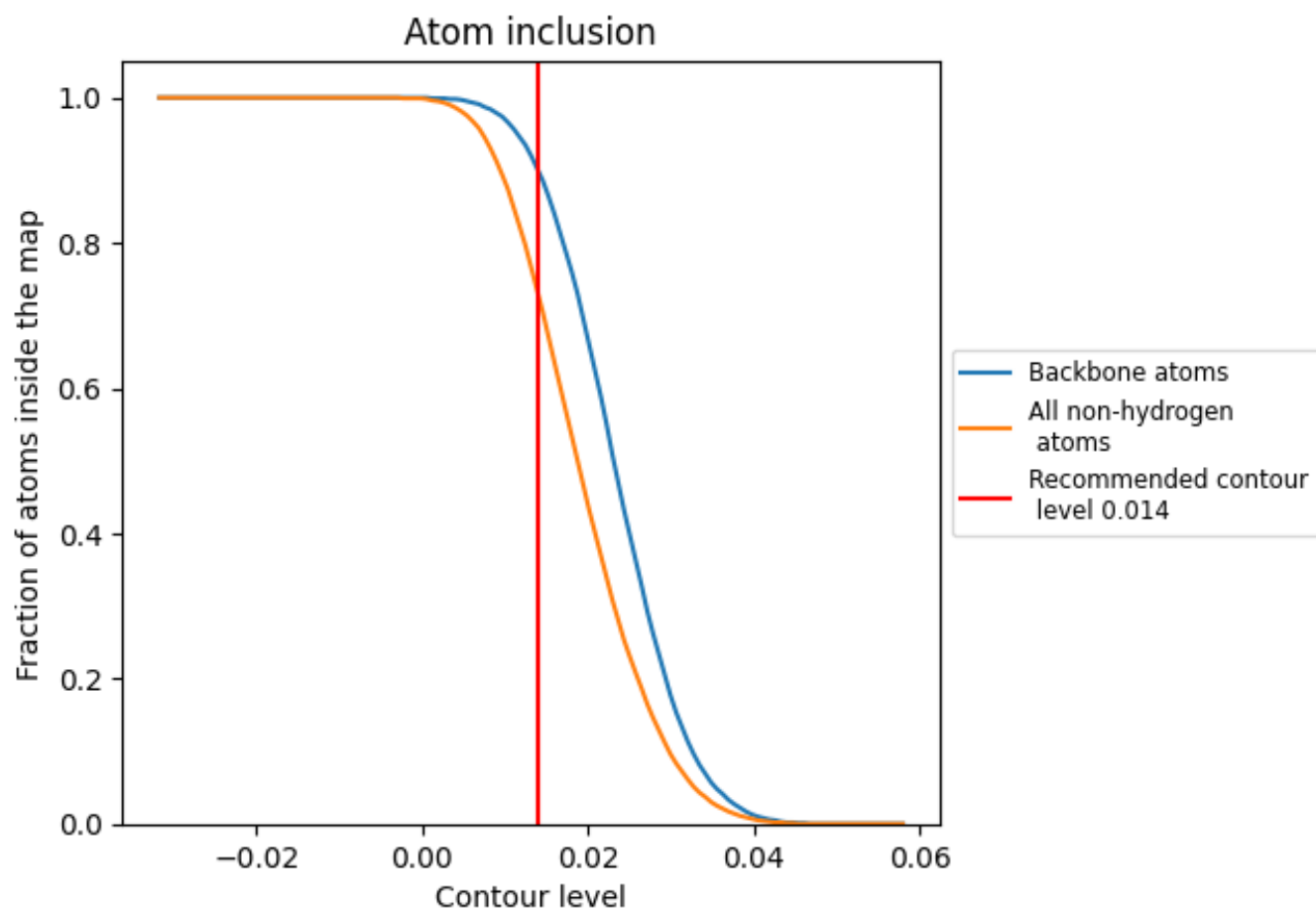
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.014).











9.4 Atom inclusion [i](#)



At the recommended contour level, 90% 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 (0.014) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.7290	 0.3360
A	 0.7340	 0.3410
D	 0.5950	 0.1810
F	 0.4020	 0.1350
G	 0.7090	 0.2210

