



## wwPDB EM Validation Summary Report ⓘ

Nov 14, 2022 – 06:48 AM EST

PDB ID : 7K1N  
EMDB ID : EMD-22626  
Title : CryoEM structure of inactivated-form DNA-PK (Complex V)  
Authors : Chen, X.; Gellert, M.; Yang, W.  
Deposited on : 2020-09-08  
Resolution : 3.90 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](mailto: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.dev43  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.2

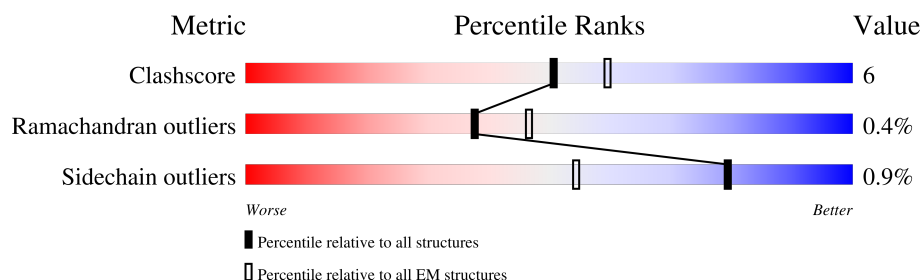
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.90 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  $\geq 3$ , 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	B	609	
3	C	732	
4	D	24	
4	F	24	
5	E	16	
5	G	16	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 39521 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
1	A	3634	Total	C	N	O	S	0	0
			28808	18485	4884	5252	187		

- Molecule 2 is a protein called X-ray repair cross-complementing protein 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	490	Total	C	N	O	S	0	0
			3954	2533	671	733	17		

- Molecule 3 is a protein called X-ray repair cross-complementing protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	655	Total	C	N	O	S	0	0
			5251	3359	877	989	26		

- Molecule 4 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
4	D	24	Total	C	N	O	P	0	0
			484	233	82	146	23		
4	F	21	Total	C	N	O	P	0	0
			425	204	69	131	21		

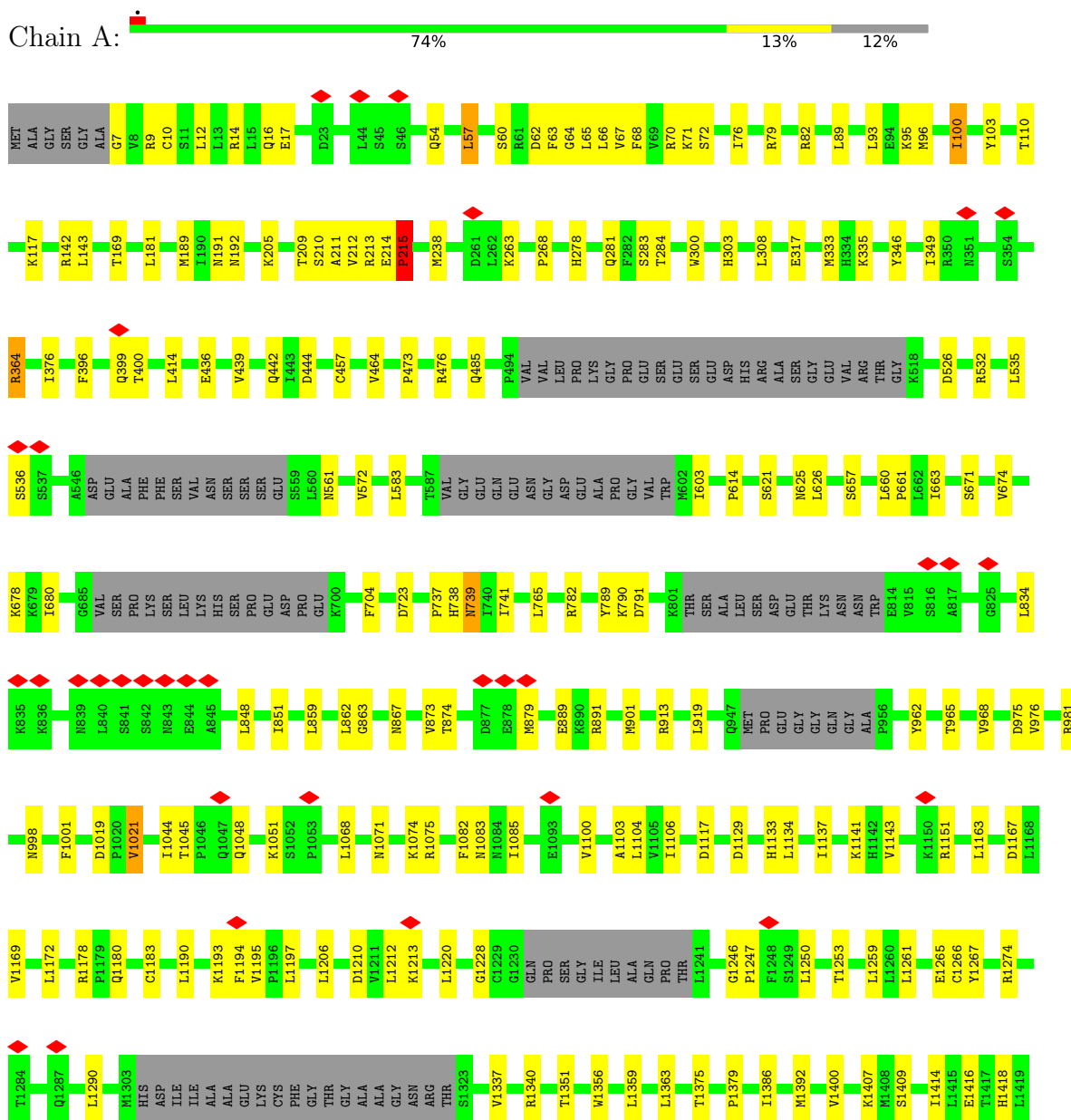
- Molecule 5 is a DNA chain called DNA (5'-D(P\*AP\*AP\*GP\*CP\*AP\*GP\*TP\*AP\*GP\*AP\*GP\*CP\*A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	13	Total	C	N	O	P	0	0
			269	128	58	71	12		
5	G	16	Total	C	N	O	P	0	0
			330	157	68	90	15		

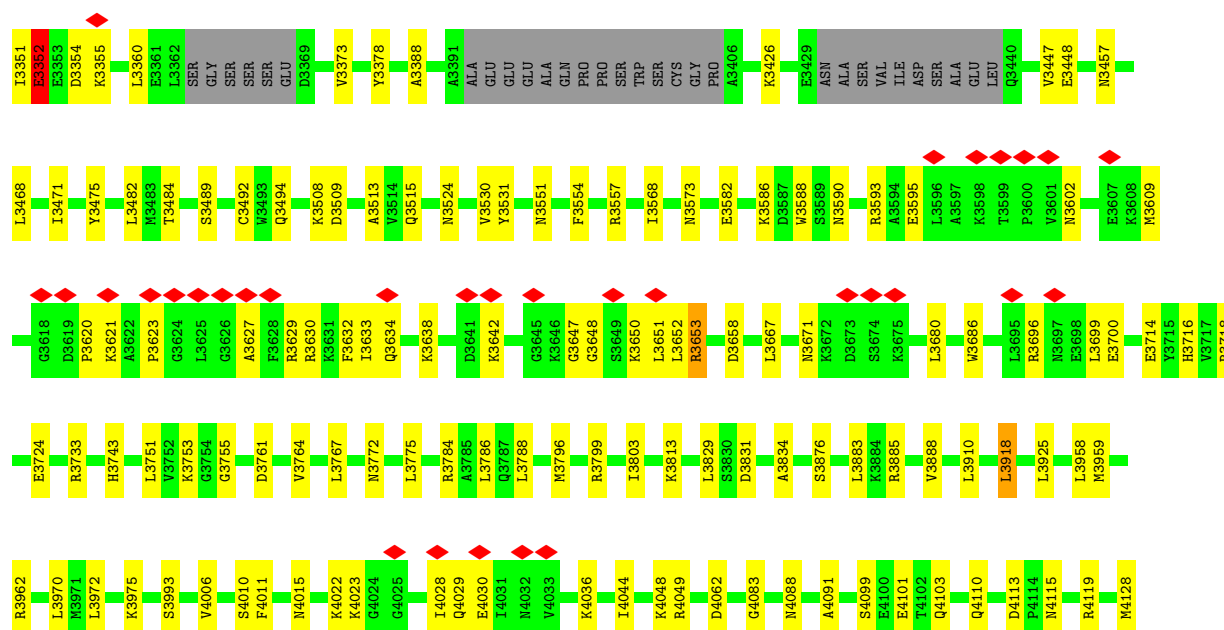
### 3 Residue-property plots

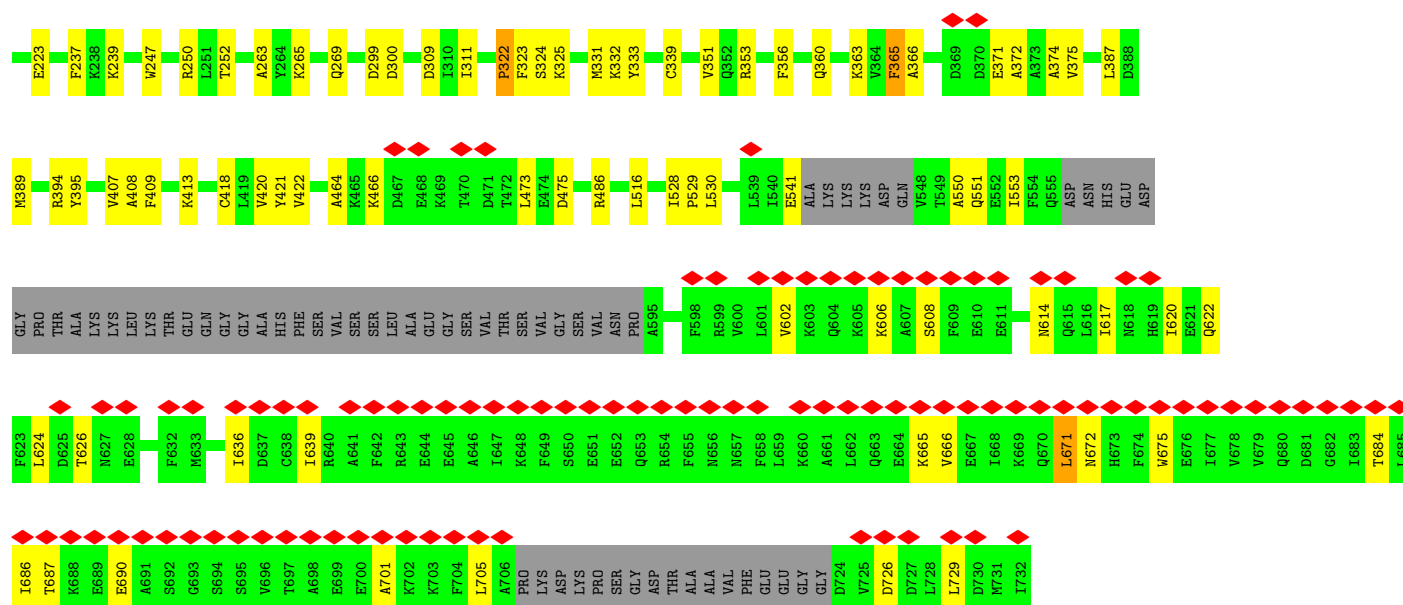
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

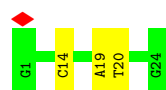


K2978	Q2979	D2980	E2985	E2990	F2993	L3011	D3026	L3027	N3028	S3047	K3048	L3049	K3050	Q3054	D3058	L3061	N3069	H3070	S3083	D3095	R3098	Y3102	V3119	R3125	E3137	Q3139	E3140	K3147	N3150	W3164	T3165	N3166	R3167	F3189	L3190	L3197							
PRO	LEU	PRO	GLU	ASN	SER	ASN	VAL	ASP	GLN	PRO	SER	ASP	ARG	MET	GLU	VAL	GLN	GLU	GLN	GLU	ALA	GLY	VAL	GLY	GLY	ALA	ALA	LEU	LEU	ALA	ARG	GLN	GLY	SER	GLN	VAL	I1552	H1574	D1588				
Q2834	K2835	L2836	L2837	N2841	R2842	T2846	S2849	F2854	C2857	L2858	Q2859	C2880	P2887	L2897	L2898	R2899	LEU	PRO	VAL	ALA	GLU	LEU	PRO	VAL	GLY	ARG	GLY	ALA	ARG	L2922	A2927	D2937	T2942	S2955	R2962	S2963	Q2971						
GLY	SER	LEU	SER	ALA	TRP	ALA	GLY	ILE	ARG	ALA	THR	GLN	GLN	HIS	ASP	THR	GLN	GLY	PRO	GLY	LEU	PRO	VAL	GLY	THR	LEU	THR	VAL	PRO	THR	GLN	LEU	SER	GLN	LEU	THR	ALA						
L2327	R2328	Y2329	E2332	I2336	L2337	E2338	C2342	V2345	H2352	F2360	D2376	F2383	F2384	P2387	K2388	R2404	M2408	D2428	Q2432	V2458	Y2459	H2464	C2469	L2476	P2487	E2488	S2489	E2490	T2491	D2492	N2493	Q2496	V2505	L2506	T2507	D2512	F2524						
S2124	W2125	W2126	G2131	W2135	R2143	L2149	T2153	K2162	H2163	W2164	L2169	Q2170	L2171	W2177	G2178	G2179	T2182	H2183	W2184	V2187	V2190	T2193	A2200	R2214	V2230	N2234	R2254	L2255	L2256	F2257	P2265	P2266	P2290	Q2295	S2296	F2309							
F2043	D2044	T2047	G2048	VAL	GLN	SER	TYR	SER	TYR	SER	GLN	ASP	PRO	ARG	PRO	ALA	THR	GLY	ARG	PHE	ARG	ILE	GLU	ARG	THR	VAL	HIS	ASP	ASP	VAL	GLY	ASP	SER	PRO	GLN	GLY	ASP	S2034	T2035	L2036	S2037	M2040	R2120
E1969	K1973	N1974	L1975	L1984	K1985	R1986	R1987	Y1988	N1989	V1994	GLU	VAL	PRO	PRO	MET	GLU	THR	LYS	LYS	LYS	TYR	ILE	GLU	ARG	GLU	VAL	HIS	ALA	ASN	GLY	ASP	R1883	D1888	K1892	I1896	F1900	H1901	I1905	G1908	K1917	F1956	G1964	S1968
N1589	T1590	F1605	R1608	L1623	W1626	W1632	W1633	A1634	L1639	K1642	V1645	T1663	D1681	K1682	K1683	L1684	D1685	L1686	H1687	L1688	P1697	G1704	L1707	Q1725	S1726	R1727	E1728	F1729	R1735	F1736	H1737	N1738	E1751	L1752	M1762	V1765	M1774	L1777					
R1420	E1421	K1422	E1430	D1440	R1445	C1455	L1458	H1459	R1460	G1462	L1463	D1474	L1475	H1476	L1484	G1494	ASP	GLU	ARG	GLN	CYS	LEU	P1501	S1502	L1503	C1507	K1508	L1524	C1525	L1532	A1541	SER	LEU	GLY	SER	SER	GLN	SER	VAL	I1552	H1574	D1588	

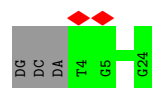
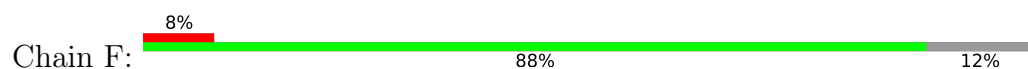




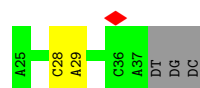
- Molecule 4: DNA (5'-D(P\*GP\*CP\*AP\*TP\*GP\*CP\*TP\*CP\*TP\*AP\*CP\*TP\*GP\*CP\*TP\*T P\*CP\*GP\*AP\*TP\*AP\*TP\*CP\*G)-3')



- Molecule 4: 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 5: DNA (5'-D(P\*AP\*AP\*GP\*CP\*AP\*GP\*TP\*AP\*GP\*AP\*GP\*CP\*A)-3')



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



There are no outlier residues recorded for this chain.

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	138985	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	0.104	Depositor
Minimum map value	-0.055	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.013	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	0.37	1/29394 (0.0%)	0.59	11/39755 (0.0%)
2	B	0.36	0/4031	0.60	2/5429 (0.0%)
3	C	0.32	0/5351	0.60	2/7213 (0.0%)
4	D	0.80	0/540	1.06	0/831
4	F	0.77	0/473	1.05	0/727
5	E	0.60	0/304	0.87	0/468
5	G	0.77	0/372	0.90	0/573
All	All	0.39	1/40465 (0.0%)	0.62	15/54996 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1195	VAL	C-N	5.33	1.44	1.34

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	3311	ASN	N-CA-C	-8.40	88.33	111.00
3	C	671	LEU	CA-CB-CG	7.85	133.36	115.30
1	A	2887	PRO	CB-CA-C	7.76	131.41	112.00
1	A	1984	LEU	CA-CB-CG	6.89	131.15	115.30
2	B	497	LEU	N-CA-C	-5.85	95.20	111.00

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	28808	0	29022	330	0
2	B	3954	0	4042	66	0
3	C	5251	0	5269	78	0
4	D	484	0	274	2	0
4	F	425	0	240	0	0
5	E	269	0	146	1	0
5	G	330	0	180	0	0
All	All	39521	0	39173	460	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 6.

The worst 5 of 460 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:3352:GLU:HG3	1:A:3355:LYS:HG3	1.38	1.00
1:A:7:GLY:N	1:A:10:CYS:HG	1.69	0.90
1:A:3352:GLU:HG3	1:A:3355:LYS:CG	2.13	0.77
1:A:1416:GLU:O	1:A:1420:ARG:HB2	1.86	0.74
2:B:495:LEU:HB3	2:B:497:LEU:CD2	2.24	0.68

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	3594/4128 (87%)	3123 (87%)	459 (13%)	12 (0%)	41	75
2	B	486/609 (80%)	409 (84%)	74 (15%)	3 (1%)	25	63
3	C	645/732 (88%)	549 (85%)	92 (14%)	4 (1%)	25	63
All	All	4725/5469 (86%)	4081 (86%)	625 (13%)	19 (0%)	38	71

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1476	HIS
1	A	2826	LEU
1	A	3352	GLU
2	B	144	SER
1	A	215	PRO

### 5.3.2 Protein sidechains ⓘ

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	3174/3671 (86%)	3146 (99%)	28 (1%)	78	87
2	B	444/548 (81%)	436 (98%)	8 (2%)	59	77
3	C	587/649 (90%)	586 (100%)	1 (0%)	93	96
All	All	4205/4868 (86%)	4168 (99%)	37 (1%)	79	87

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	36	ASP
2	B	497	LEU
2	B	217	TYR
2	B	409	TYR
1	A	1752	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3524	ASN
2	B	152	ASN
1	A	3573	ASN
1	A	3822	GLN
2	B	326	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.

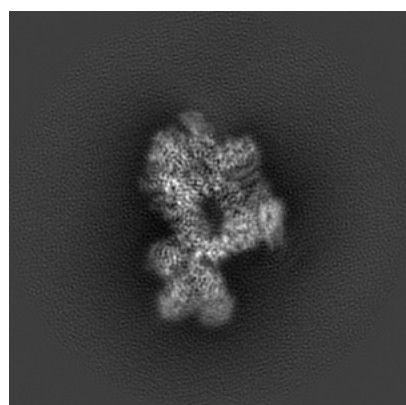
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-22626. These allow visual inspection of the internal detail of the map and identification of artifacts.

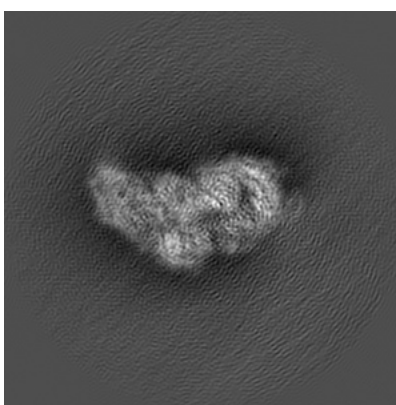
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

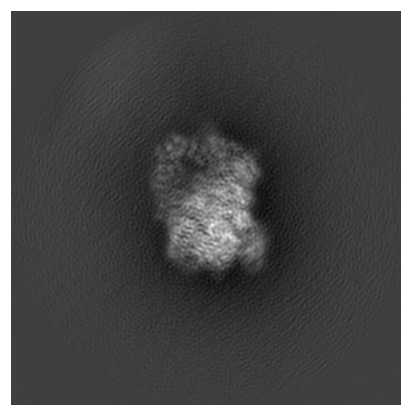
#### 6.1.1 Primary map



X



Y

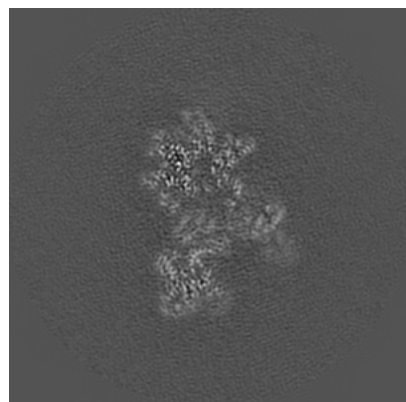


Z

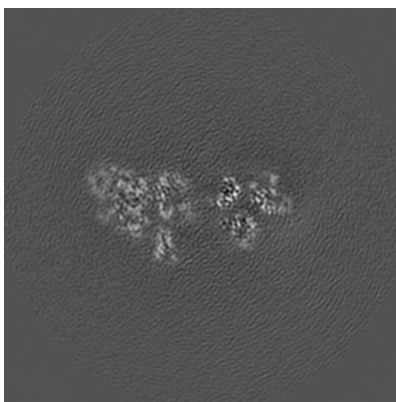
The images above show the map projected in three orthogonal directions.

### 6.2 Central slices [i](#)

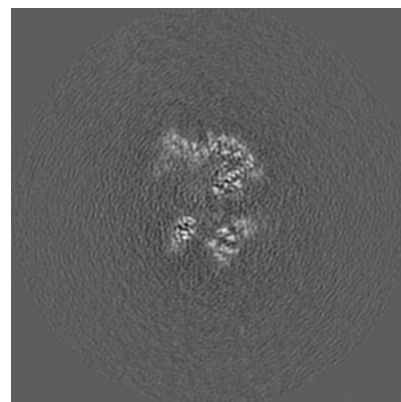
#### 6.2.1 Primary map



X Index: 176



Y Index: 176

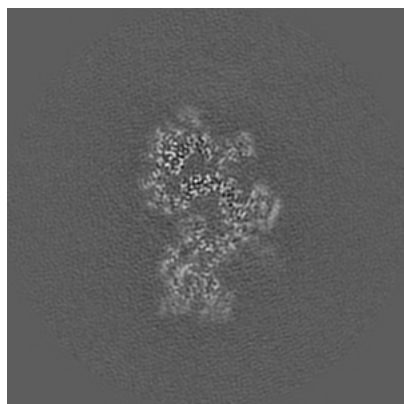


Z Index: 176

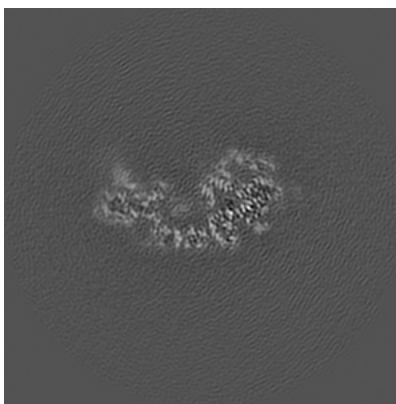
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

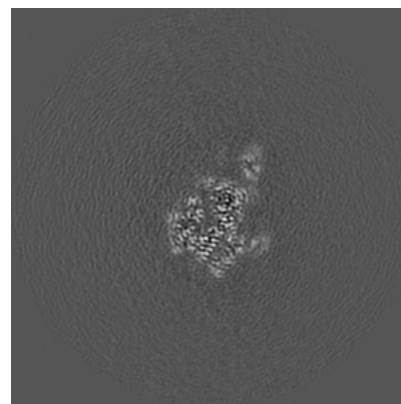
### 6.3.1 Primary map



X Index: 187



Y Index: 149



Z Index: 199

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

## 6.4 Orthogonal surface views [i](#)

### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.013. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

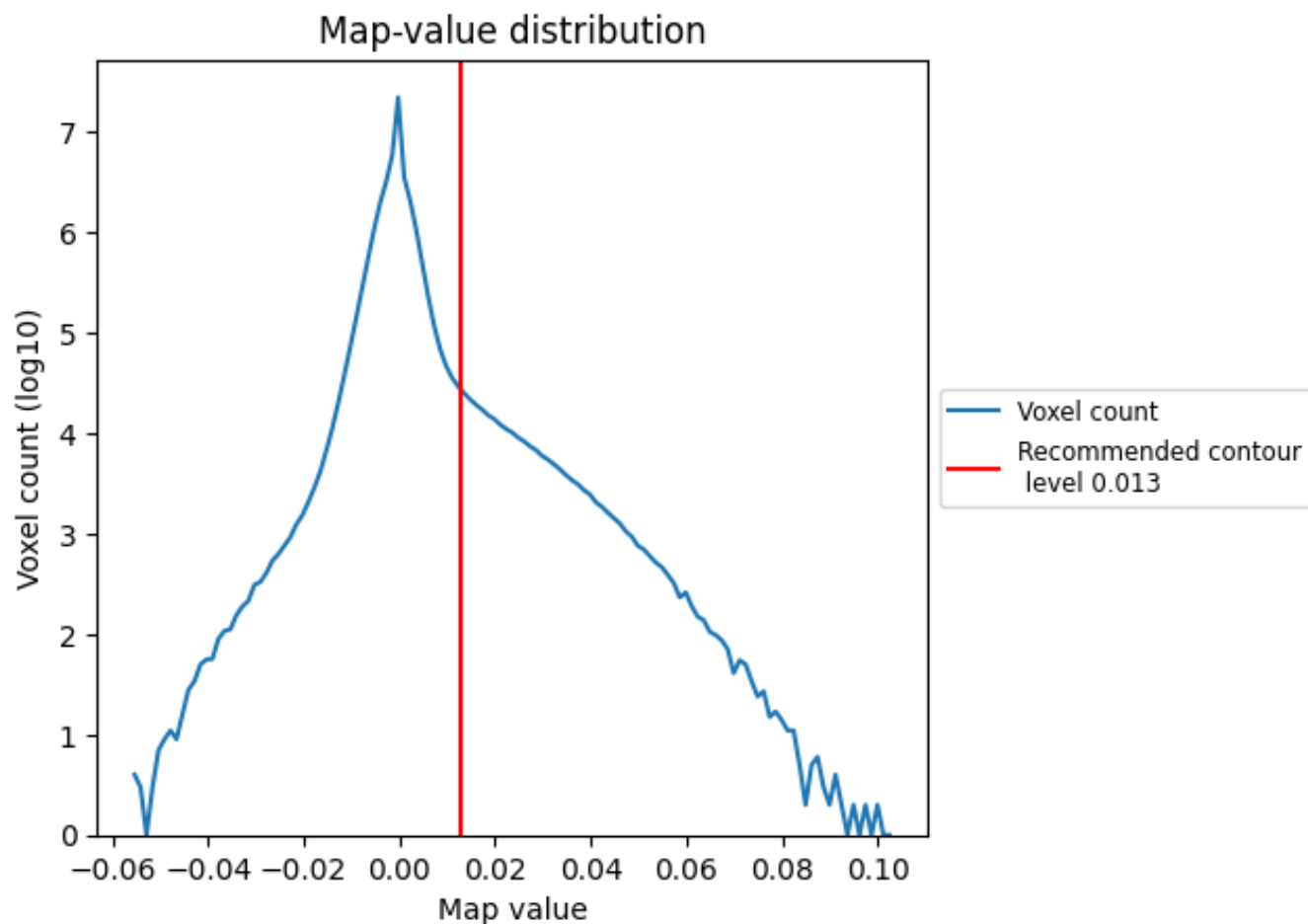
## 6.5 Mask visualisation

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.

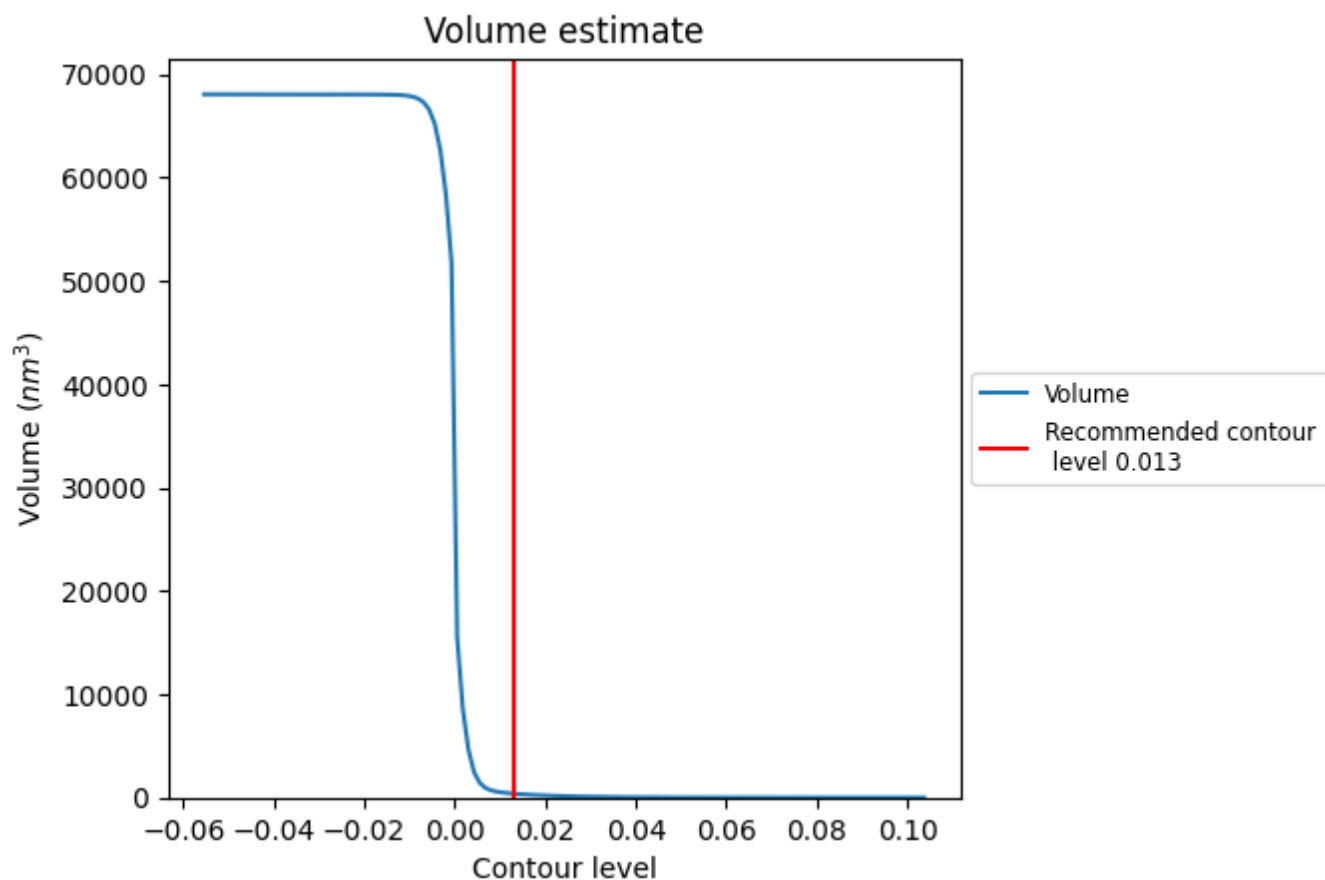
### 7.1 Map-value distribution [i](#)



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.



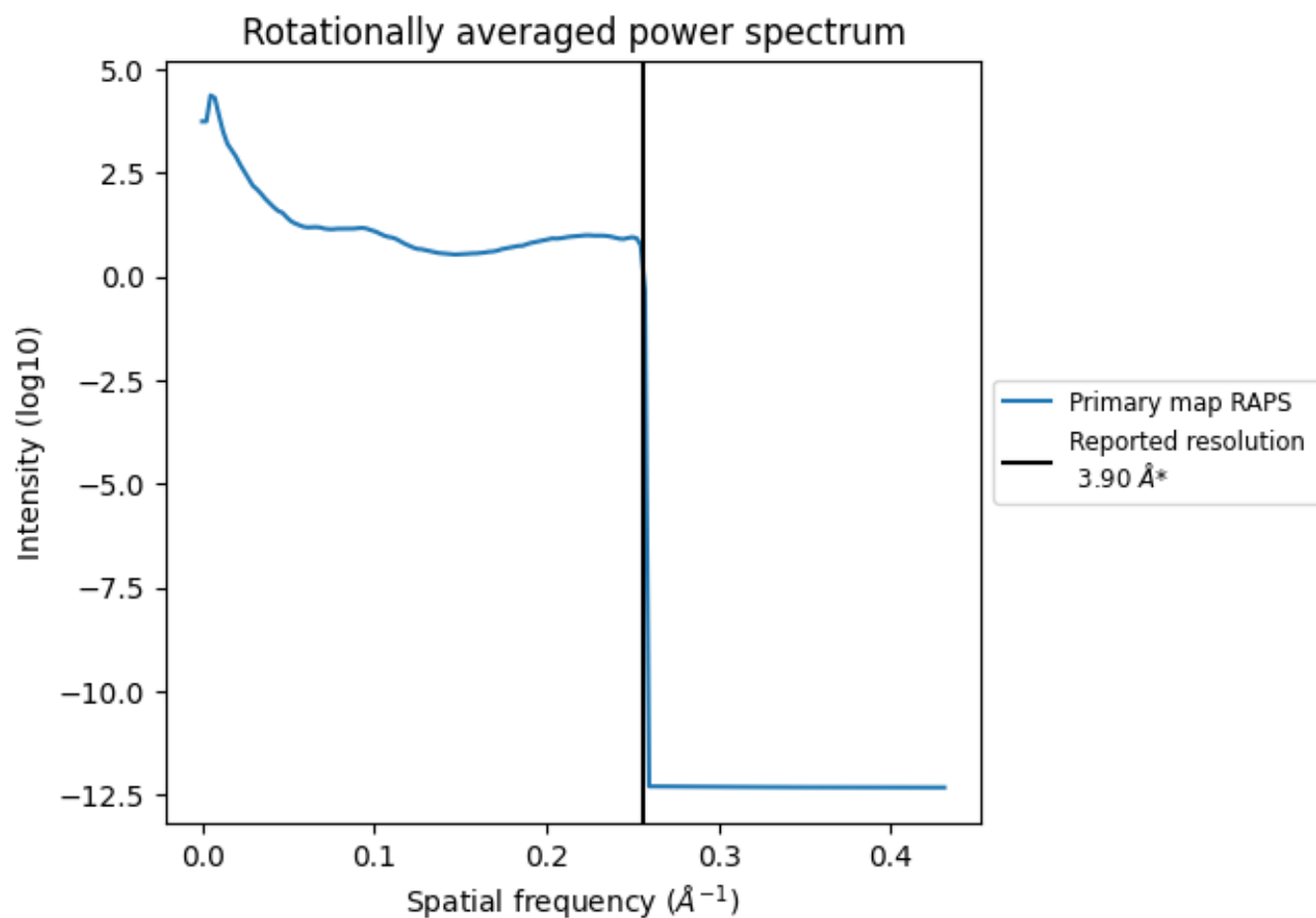
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 378 nm<sup>3</sup>; this corresponds to an approximate mass of 341 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 ⓘ



\*Reported resolution corresponds to spatial frequency of 0.256 Å<sup>-1</sup>

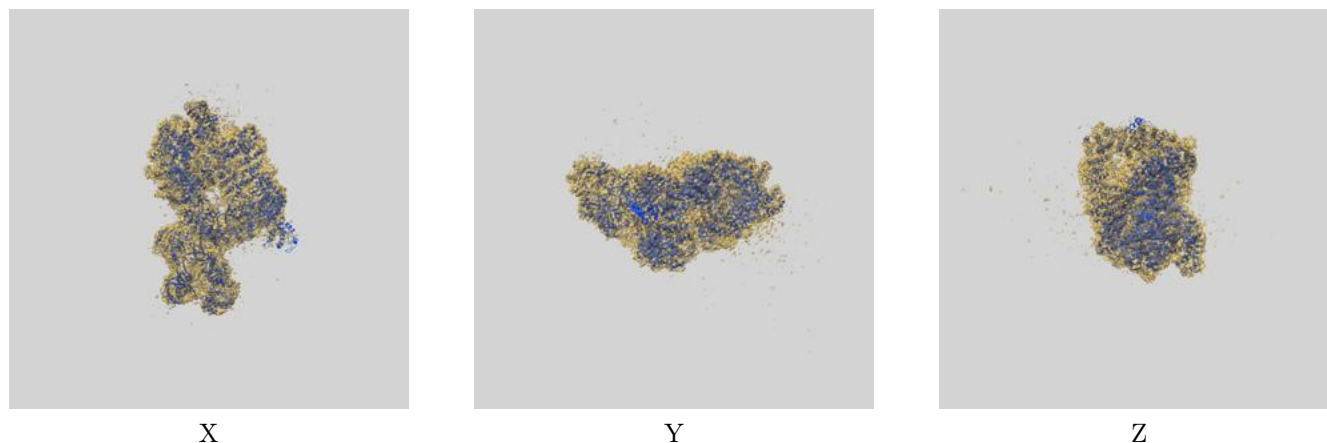
## 8 Fourier-Shell correlation

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

## 9 Map-model fit [i](#)

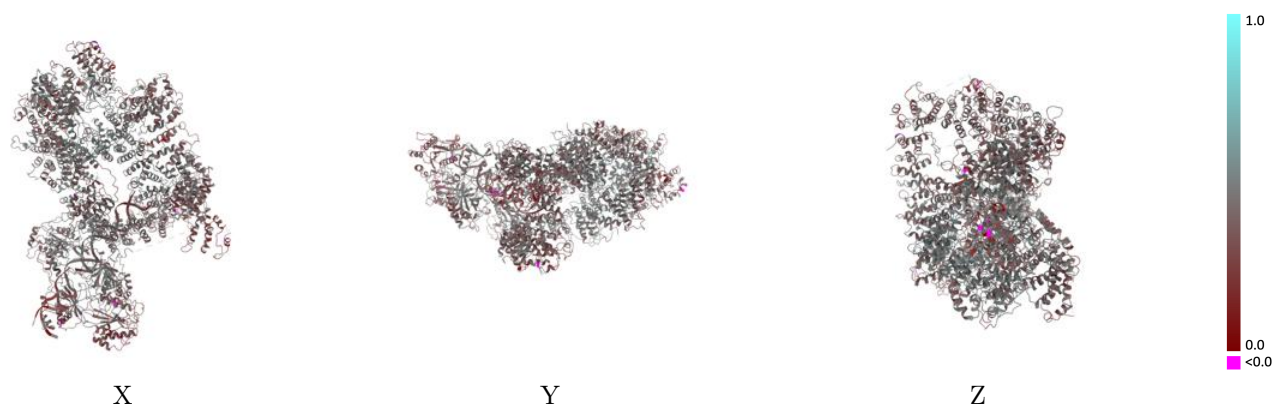
This section contains information regarding the fit between EMDB map EMD-22626 and PDB model 7K1N. Per-residue inclusion information can be found in [section 3](#) on [page 4](#).

### 9.1 Map-model overlay [i](#)



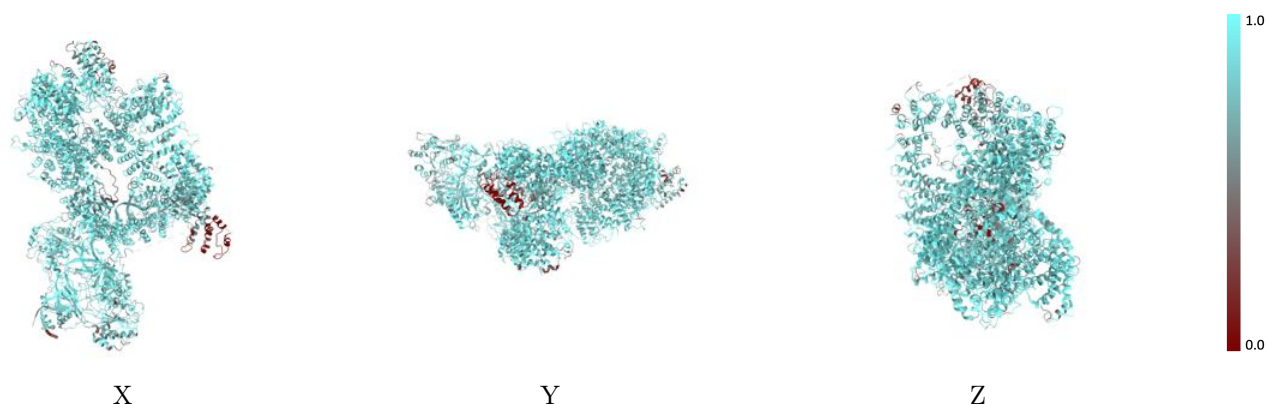
The images above show the 3D surface view of the map at the recommended contour level 0.013 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



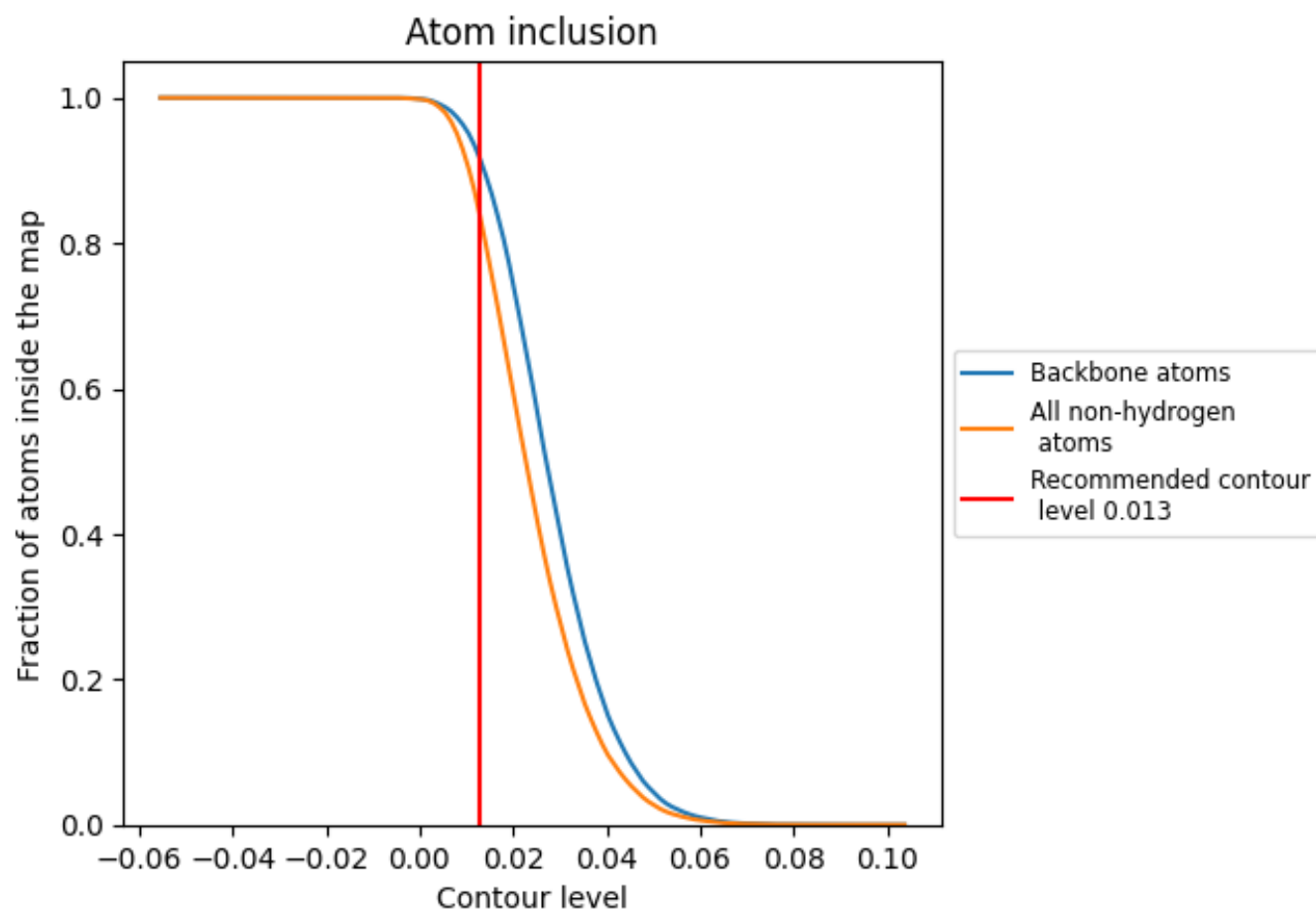
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.013).

## 9.4 Atom inclusion [i](#)



At the recommended contour level, 91% of all backbone atoms, 84% 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 (0.013) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	<div><div></div>0.8351</div>	<div><div></div>0.4170</div>
A	<div><div></div>0.8594</div>	<div><div></div>0.4300</div>
B	<div><div></div>0.8488</div>	<div><div></div>0.4160</div>
C	<div><div></div>0.6847</div>	<div><div></div>0.3710</div>
D	<div><div></div>0.8926</div>	<div><div></div>0.3510</div>
E	<div><div></div>0.7584</div>	<div><div></div>0.2840</div>
F	<div><div></div>0.8588</div>	<div><div></div>0.3510</div>
G	<div><div></div>0.9000</div>	<div><div></div>0.3430</div>

1.0

0.0

<0.0