wwPDB EM Validation Summary Report (i)

May 26, 2025 – 06:37 AM EDT

PDB ID	:	7K0Y / pdb_00007k0y
EMDB ID	:	EMD-22618
Title	:	Cryo-EM structure of activated-form DNA-PK (complex VI)
Authors	:	Chen, X.; Gellert, M.; Yang, W.
Deposited on	:	2020-09-06
Resolution	:	3.70 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 (i) 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 (i)) were used in the production of this report:

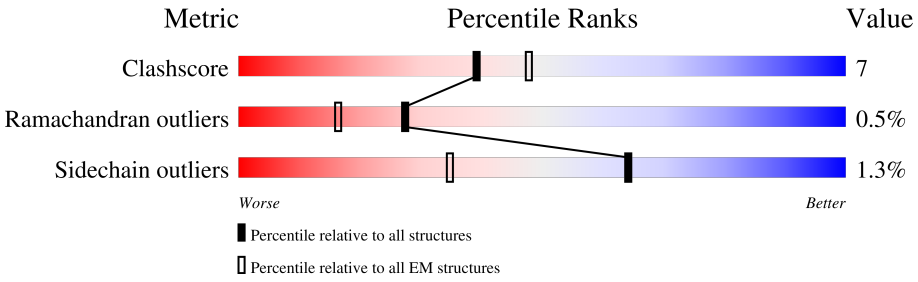
EMDB validation analysis	:	0.0.1.dev118
MolProbity	:	4-5-2 with Phenix2.0rc1
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.70 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

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	A	4128	74% 14% 11%
2	B	609	63% 16% 20%
3	C	732	13% 71% 18% 11%
4	D	24	92% 8%
4	F	24	88% 12%
5	E	16	6% 75% 6% 19%
5	G	16	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 39835 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	3669	29107	18680	4932	5306	189	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	490	3973	2545	677	733	18	2	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	655	5247	3357	876	988	26	0	0

- Molecule 4 is a DNA chain called DNA (5'-D(*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		
4	D	24	484	233	82	146	23	0	0
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(*AP*AP*GP*CP*AP*GP*TP*AP*GP*AP*GP*CP*A)-3').

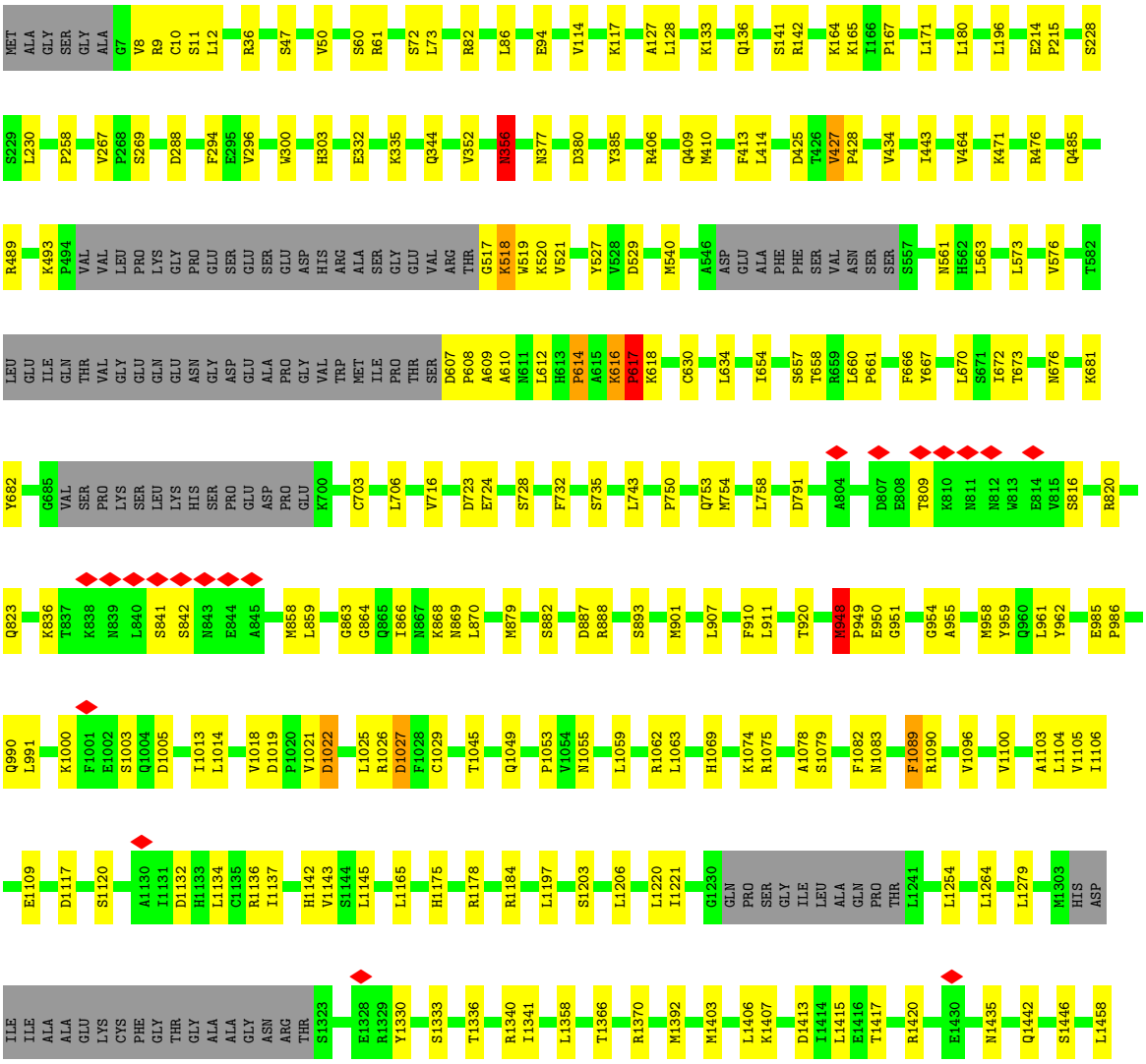
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	E	13	269	128	58	71	12	0	0
5	G	16	Total	C	N	O	P	0	0
			330	157	68	90	15		

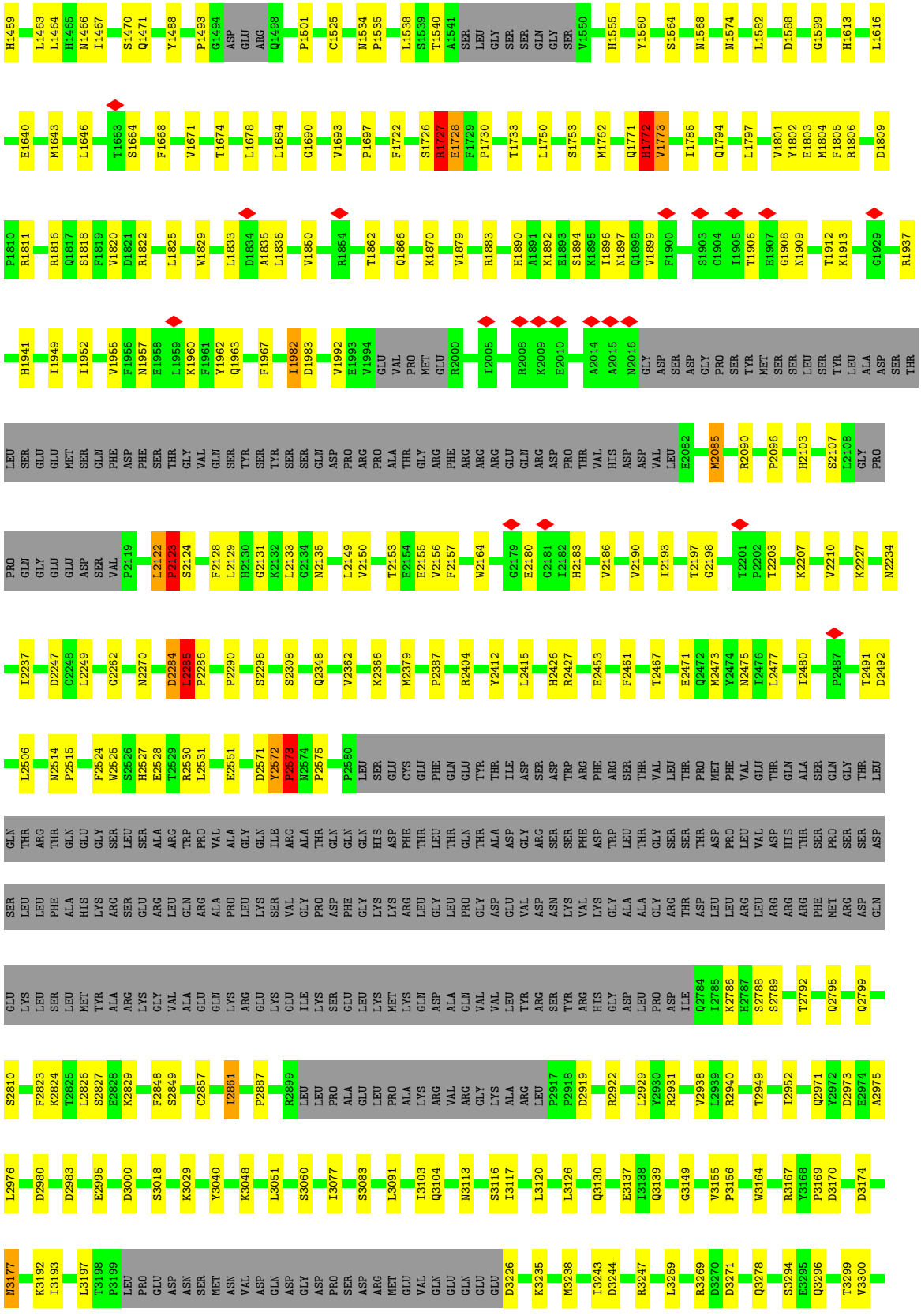
3 Residue-property plots (i)

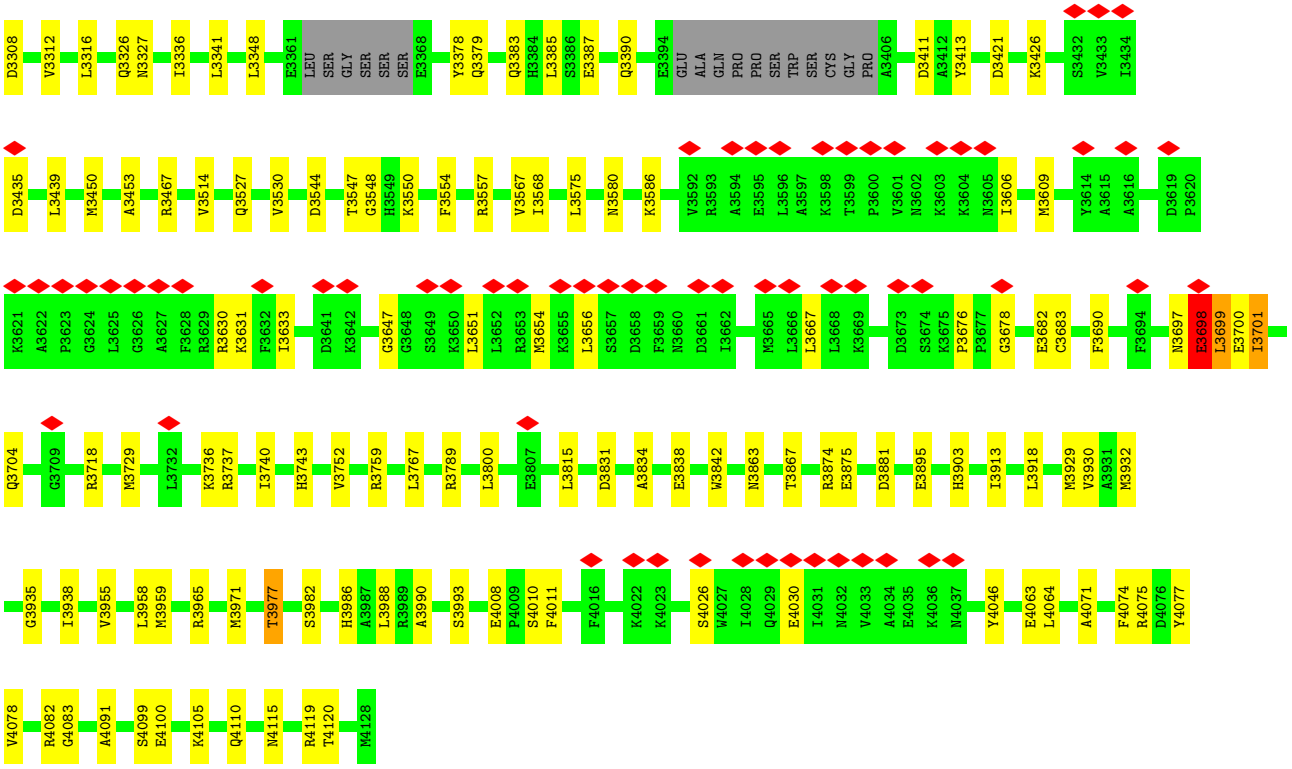
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

Chain A:

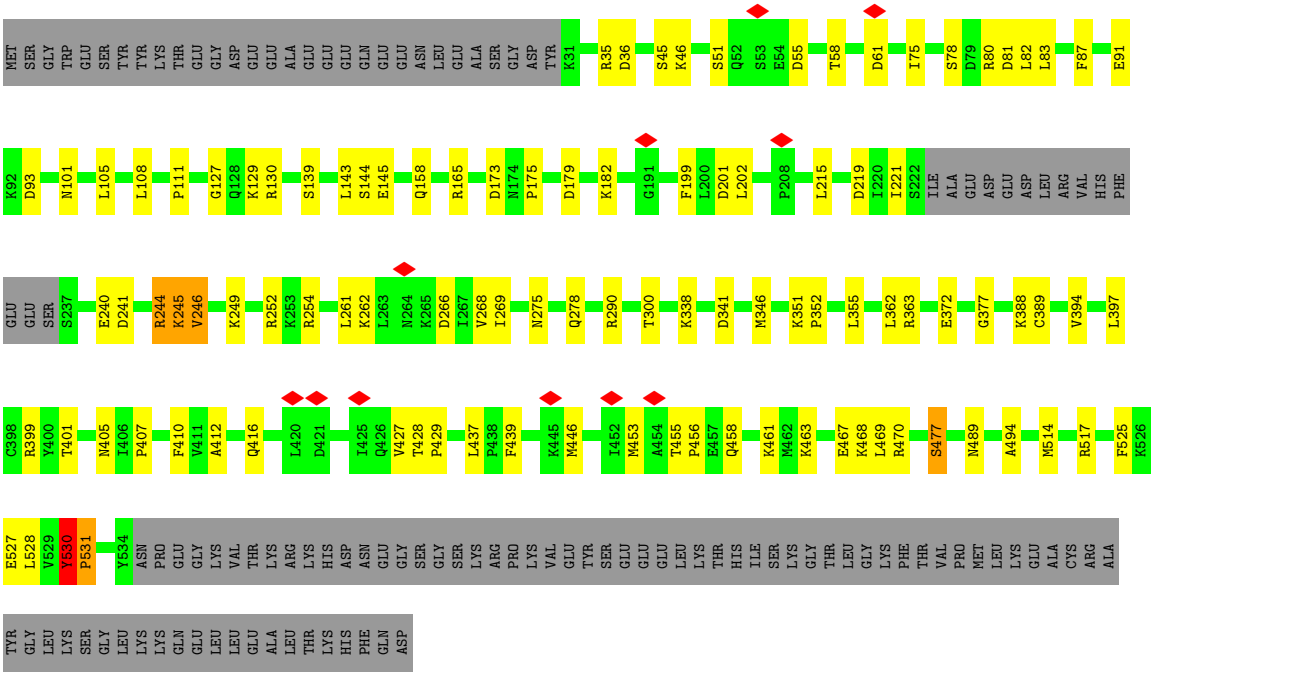



H1459 L1463 L1464 H1465 N1466 I1467 S1470 Q1471 Y1488 P1493 G1494 ASP GLU ARG Q1498 P1501 C1525 N1534 P1535 L1538 S1539 T1540 A1541 SER LEU GLY SER SER GLN GLY SER V1550 H1555 Y1560 S1564 N1568 N1574 L1582 D1588 G1599 H1613 L1616
E1640 M1643 L1646 T1663 S1664 F1668 V1671 T1674 L1678 L1684 G1690 V1693 P1697 F1722 S1726 R1727 E1728 F1729 P1730 T1733 L1750 S1753 M1762 Q1771 H1772 V1773 I1785 Q1794 L1797 V1801 Y1802 E1803 M1804 F1805 R1806 D1809
P1810 R1811 R1816 Q1817 S1818 F1819 V1820 D1821 R1822 L1825 W1829 L1833 D1834 A1835 L1836 V1850 R1854 T1862 Q1866 K1870 V1879 R1883 H1890 A1891 K1892 E1893 S1894 K1895 I1896 N1897 Q1898 V1899 F1900 S1903 C1904 I1905 T1906 E1907 G1908 N1909 T1912 K1913 G1929 R1937
H1941 I1949 I1952 V1955 F1956 N1957 E1958 L1959 K1960 F1961 Y1962 Q1963 F1967 I1982 D1983 V1992 E1993 V1994 GLU VAL PRO MET GLU R2000 I2005 R2008 K2009 E2010 A2014 A2015 N2016 GLY ASP SER ASP GLY PRO SER TYR MET SER SER LEU SER TYR LEU ALA ASP SER THR
LEU SER GLU GLU MET SER GLN PHE ASP PHE SER THR GLY VAL GLN SER TYR SER TYR SER SER GLN ASP PRO ARG PRO ALA THR GLY ARG PHE ARG ARG ARG GLU GLN GLU ARG ASP PRO THR VAL HIS ASP ASP VAL LEU E2082 M2085 R2090 P2096 H2103 S2107 I2108 GLY PRO
PRO GLN GLY GLU GLU ASP SER VAL P2119 L2122 P2123 S2124 F2128 L2129 H2130 G2131 K2132 L2133 G2134 N2135 L2149 V2150 T2153 E2154 E2155 V2156 F2157 W2164 G2179 E2180 G2181 I2182 H2183 V2186 V2190 I2193 T2197 G2198 T2201 P2202 T2203 K2207 V2210 K2227 N2234
I2237 D2247 C2248 L2249 G2262 N2270 D2284 L2285 P2286 P2290 S2296 S2308 Q2348 V2362 K2366 M2379 P2387 R2404 Y2412 L2415 H2426 R2427 E2453 F2461 T2467 E2471 Q2472 M2473 Y2474 N2475 I2476 L2477 I2480 P2487 T2491 D2492
L2506 N2514 P2515 F2524 W2525 S2526 H2527 E2528 T2529 R2530 L2531 E2551 D2571 Y2572 P2573 N2574 P2575 P2580 LEU SER GLU CYS GLU PHE GLN GLU TYR THR ILE ASP SER ASP TRP ARG PHE ARG SER THR VAL LEU THR PRO MET PHE VAL GLU THR GLN ALA SER GLN GLY THR LEU
GLN THR ARG THR GLN GLU GLY SER LEU SER ALA ARG TRP PRO VAL ALA GLY GLN ILE ARG ALA THR GLN GLN GLN HIS ASP PHE THR LEU GLN THR ALA ASP GLY ARG SER SER PHE ASP TRP LEU THR GLY SER SER THR ASP PRO LEU VAL ASP HIS THR SER PRO SER SER ASP
SER LEU LEU PHE ALA HIS LYS ARG SER GLU ARG LEU GLN ARG ALA PRO LEU LYS SER VAL GLY PRO ASP PHE GLY HIS LYS LYS ARG THR LEU GLY LEU PRO GLY ASP GLU VAL ASP ASN LYS VAL LYS GLY ALA ALA GLY ARG THR ASP LEU PRO ARG ARG ARG ARG PHE MET ARG ASP GLN
GLU LYS LEU SER LEU MET TYR ALA ARG LYS GLY VAL ALA GLU GLN VAL LEU TYR ARG SER TYR ARG HIS GLY ASP LEU PRO ASP ILE Q2764 I2785 K2786 H2787 S2788 S2789 T2792 Q2795 Q2799
S2810 F2823 K2824 T2825 L2826 S2827 E2828 K2829 F2848 S2849 C2857 I2861 P2887 R2899 LEU LEU PRO ALA GLU LEU PRO ALA LYS ARG VAL ARG GLY LYS ALA ARG LEU P2917 P2918 D2919 R2922 L2929 Y2930 R2931 V2938 L2939 R2940 T2949 I2952 Q2971 Y2972 D2973 E2974 A2975
L2976 D2980 D2983 E2995 D3000 S3018 K3029 Y3040 K3048 L3061 S3060 I3077 S3083 L3091 I3103 Q3104 N3113 S3116 I3117 L3120 L3126 Q3130 E3137 I3138 Q3139 G3149 V3155 P3156 W3164 R3167 Y3168 P3169 D3170 D3174
N3177 K3192 I3193 L3197 T3198 P3199 LEU PRO GLU ASP ASN SER MET ASN VAL ASP GLN ASP GLY ASP PRO SER ASP ARG MET GLU VAL GLN GLU GLN GLU GLU D3226 K3235 M3238 I3243 D3244 R3247 L3259 R3269 D3270 D3271 Q3278 S3294 E3295 Q3296 T3299 V3300



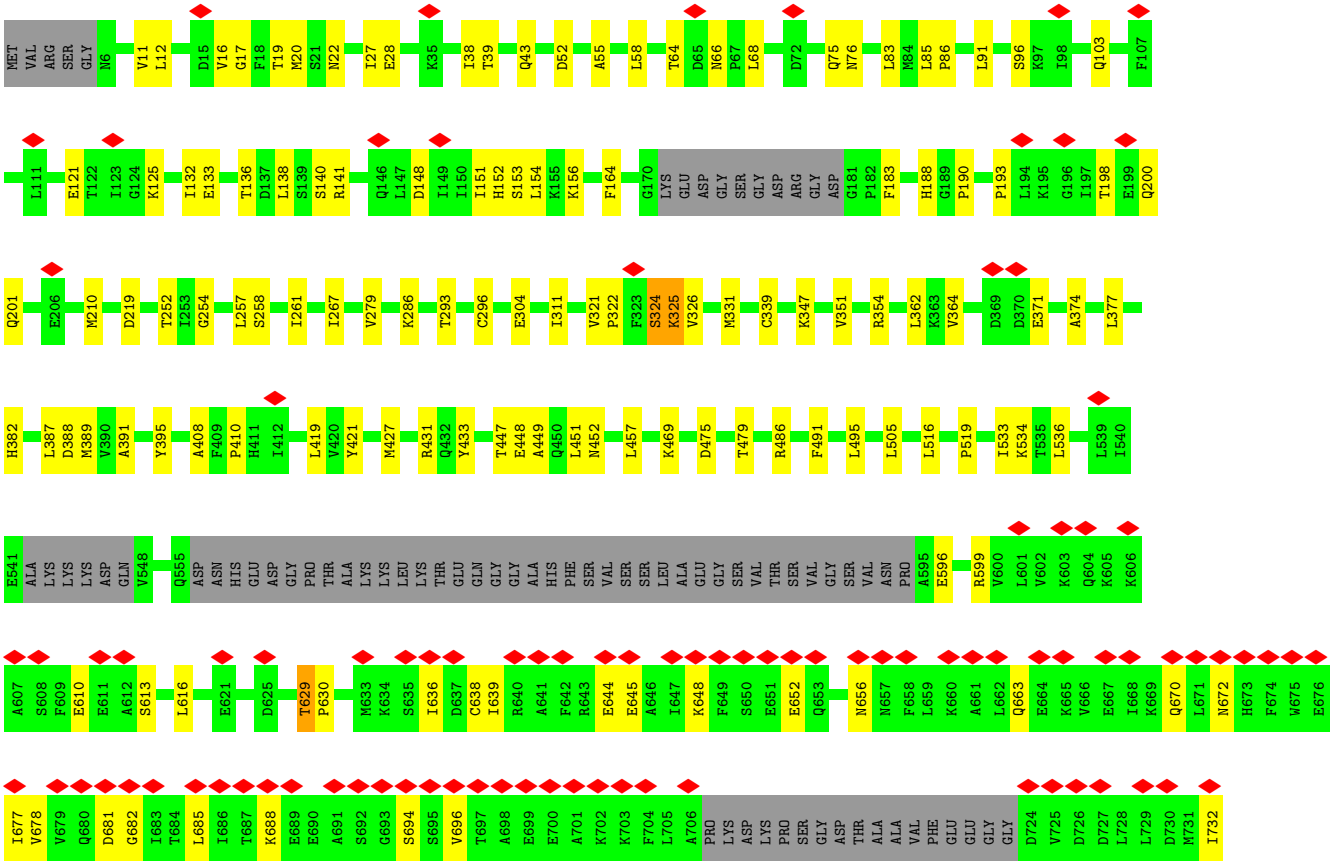
• Molecule 2: X-ray repair cross-complementing protein 6

Chain B:



• Molecule 3: X-ray repair cross-complementing protein 5

Chain C:

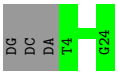


• Molecule 4: DNA (5'-D(*GP*CP*AP*TP*GP*CP*TP*CP*TP*AP*CP*TP*GP*CP*TP*TP*CP*GP*AP*TP*AP*TP*CP*G)-3')

Chain D: 92% 8%

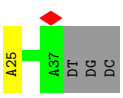
• Molecule 4: DNA (5'-D(*GP*CP*AP*TP*GP*CP*TP*CP*TP*AP*CP*TP*GP*CP*TP*TP*CP*GP*AP*TP*AP*TP*CP*G)-3')

Chain F: 88% 12%

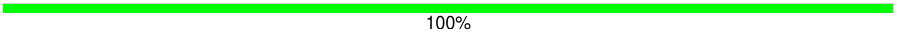


• Molecule 5: DNA (5'-D(*AP*AP*GP*CP*AP*GP*TP*AP*GP*AP*GP*CP*A)-3')

Chain E: 6% 75% 6% 19%



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

Chain G: 

There are no outlier residues recorded for this chain.

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	254646	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	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	0.135	Depositor
Minimum map value	-0.072	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.012	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

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.41	1/29703 (0.0%)	0.69	21/40166 (0.1%)
2	B	0.36	0/4056	0.73	6/5460 (0.1%)
3	C	0.29	0/5347	0.63	0/7208
4	D	0.45	0/540	0.57	0/831
4	F	0.38	0/473	0.54	0/727
5	E	0.27	0/304	0.40	0/468
5	G	0.42	0/372	0.53	0/573
All	All	0.39	1/40795 (0.0%)	0.68	27/55433 (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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	356	ASN	CG-OD1	5.10	1.33	1.23

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	2122	LEU	CA-C-N	15.17	138.80	119.84
1	A	2122	LEU	C-N-CA	15.17	138.80	119.84
1	A	3895	GLU	CB-CA-C	10.29	127.15	110.90
2	B	530	TYR	CA-C-N	-10.07	110.00	120.38
2	B	530	TYR	C-N-CA	-10.07	110.00	120.38

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1727	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	A	29107	0	29371	389	0
2	B	3973	0	4075	89	0
3	C	5247	0	5263	90	0
4	D	484	0	274	3	0
4	F	425	0	240	0	0
5	E	269	0	146	2	0
5	G	330	0	180	0	0
All	All	39835	0	39549	553	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 7.

The worst 5 of 553 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:B:245:LYS:O	2:B:246:VAL:CG2	2.04	1.06
1:A:948:MET:H	1:A:949:PRO:HD2	1.20	1.00
1:A:948:MET:HG3	1:A:949:PRO:HD3	1.40	1.00
1:A:948:MET:H	1:A:949:PRO:CD	1.78	0.96
1:A:630:CYS:O	1:A:634:LEU:HB2	1.69	0.93

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	3635/4128 (88%)	3076 (85%)	543 (15%)	16 (0%)	30	62
2	B	488/609 (80%)	410 (84%)	73 (15%)	5 (1%)	13	44
3	C	645/732 (88%)	547 (85%)	94 (15%)	4 (1%)	22	54
All	All	4768/5469 (87%)	4033 (85%)	710 (15%)	25 (0%)	27	57

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	948	MET
1	A	1772	HIS
1	A	2123	PRO
1	A	2285	LEU
1	A	2573	PRO

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	3207/3671 (87%)	3160 (98%)	47 (2%)	60	75
2	B	447/548 (82%)	444 (99%)	3 (1%)	81	88
3	C	586/649 (90%)	582 (99%)	4 (1%)	81	88
All	All	4240/4868 (87%)	4186 (99%)	54 (1%)	64	77

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1982	ILE
1	A	2827	SER
2	B	530	TYR
1	A	1992	VAL
1	A	2285	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	98	ASN
3	C	627	ASN
2	B	360	HIS
3	C	298	ASN
1	A	1794	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 oligosaccharides 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.

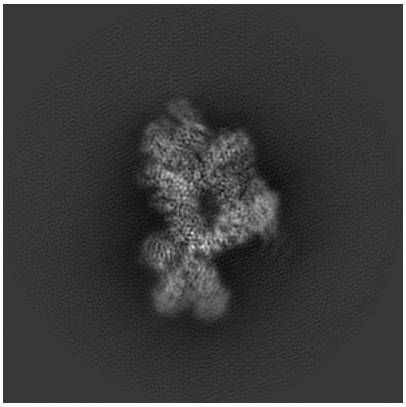
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-22618. These allow visual inspection of the internal detail of the map and identification of artifacts.

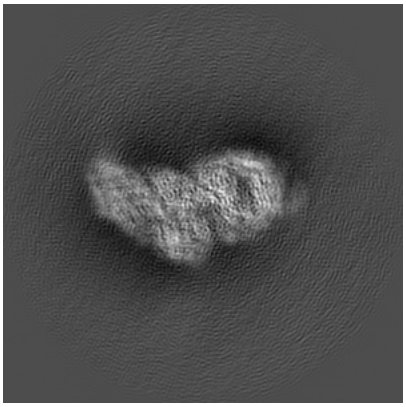
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

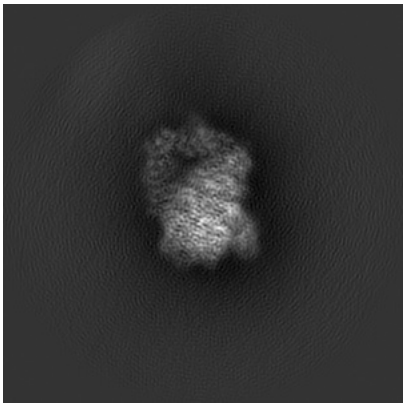
6.1.1 Primary map



X



Y

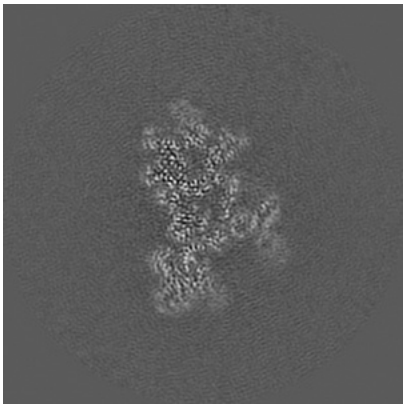


Z

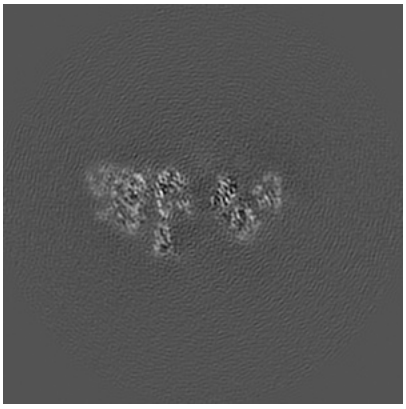
The images above show the map projected in three orthogonal directions.

6.2 Central slices

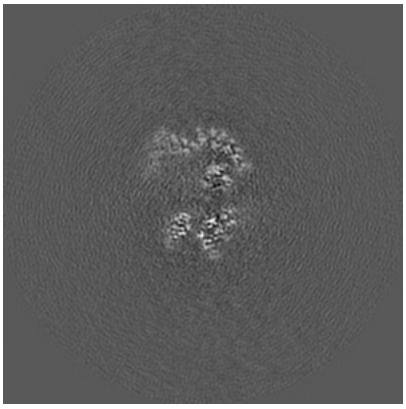
6.2.1 Primary map



X Index: 176



Y Index: 176

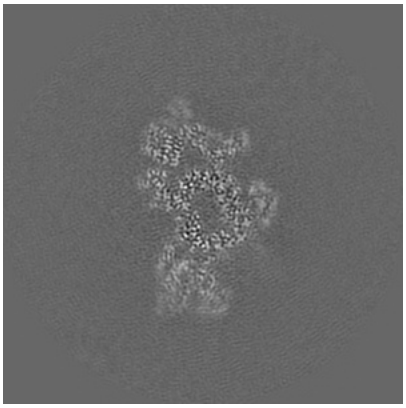


Z Index: 176

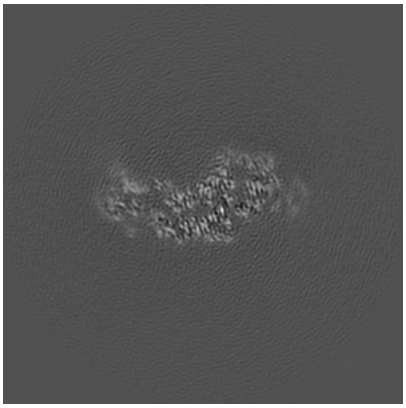
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

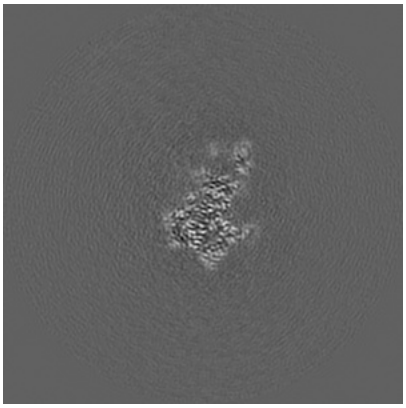
6.3.1 Primary map



X Index: 186



Y Index: 154

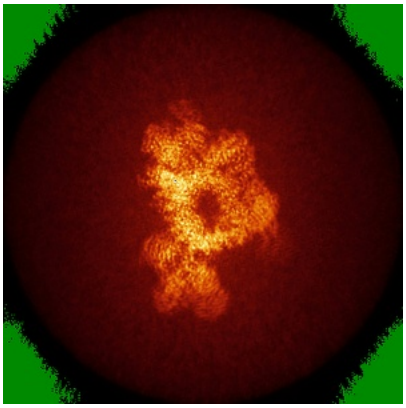


Z Index: 195

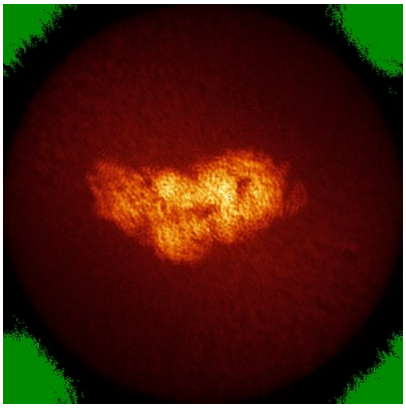
The images above show the largest variance slices of the map in three orthogonal directions.

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

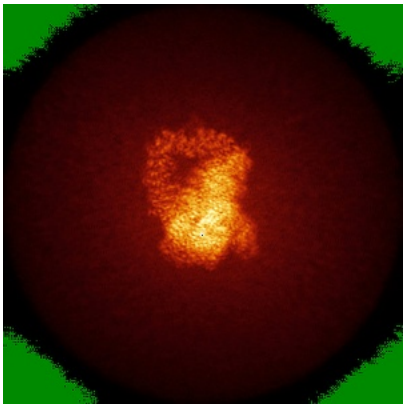
6.4.1 Primary map



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

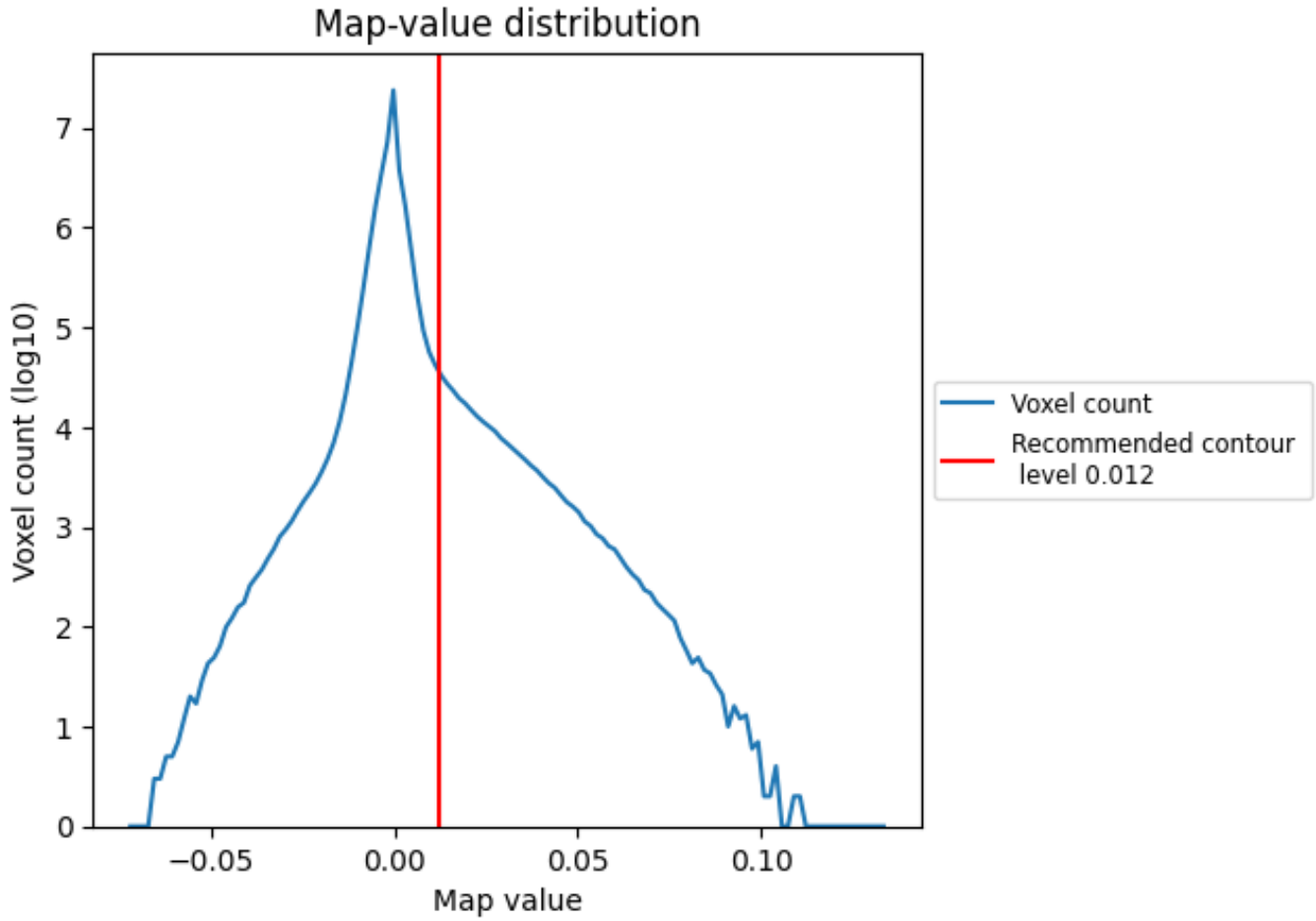
6.6 Mask visualisation (i)

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

7 Map analysis

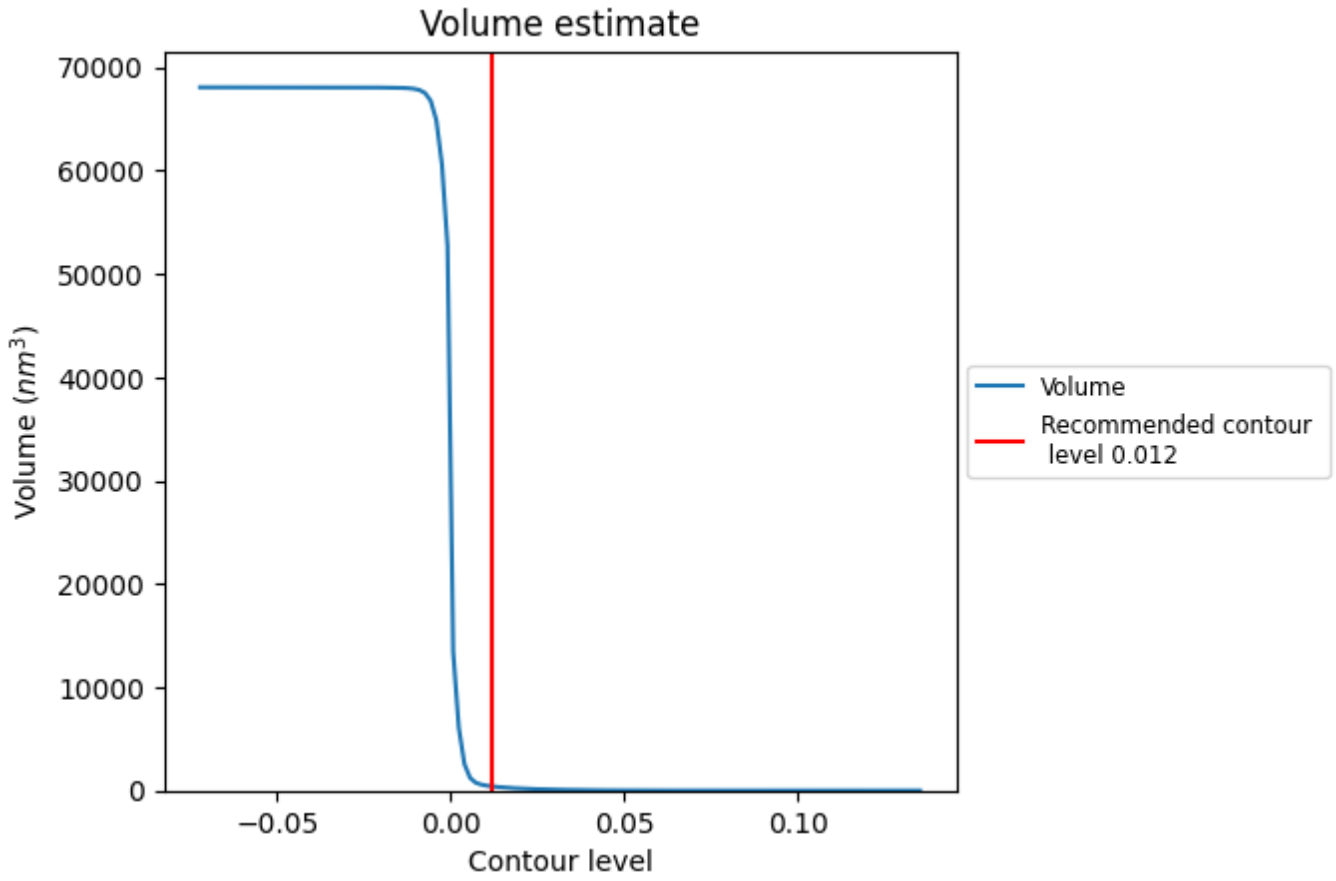
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution



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

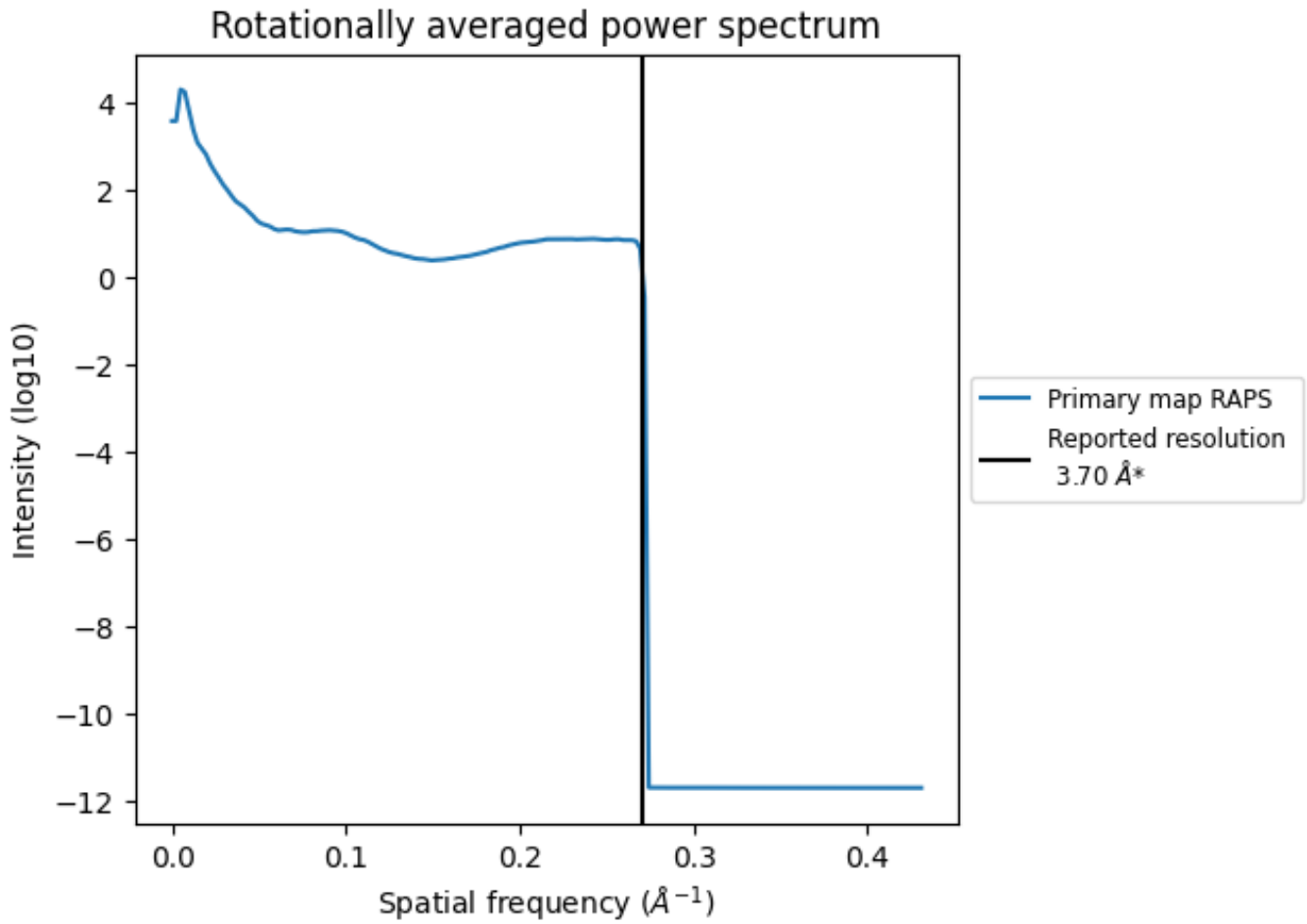
7.2 Volume estimate (i)



The volume at the recommended contour level is 405 nm^3; this corresponds to an approximate mass of 366 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 0.270 Å^{-1}

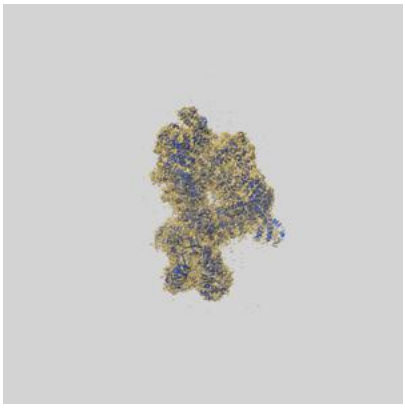
8 Fourier-Shell correlation (i)

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

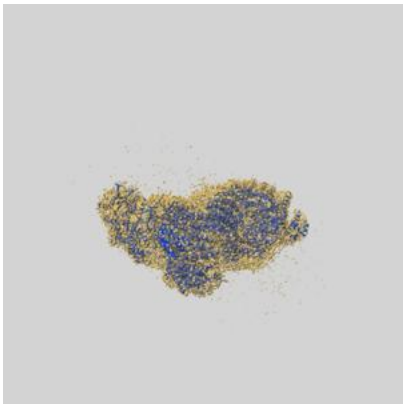
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-22618 and PDB model 7K0Y. Per-residue inclusion information can be found in section 3 on page 4.

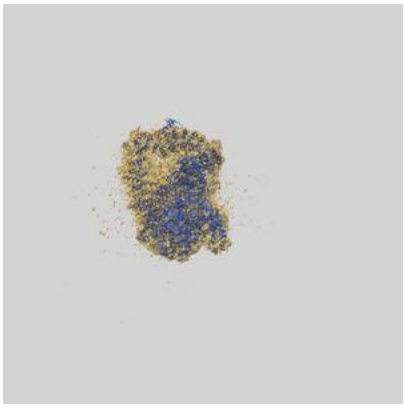
9.1 Map-model overlay



X



Y



Z

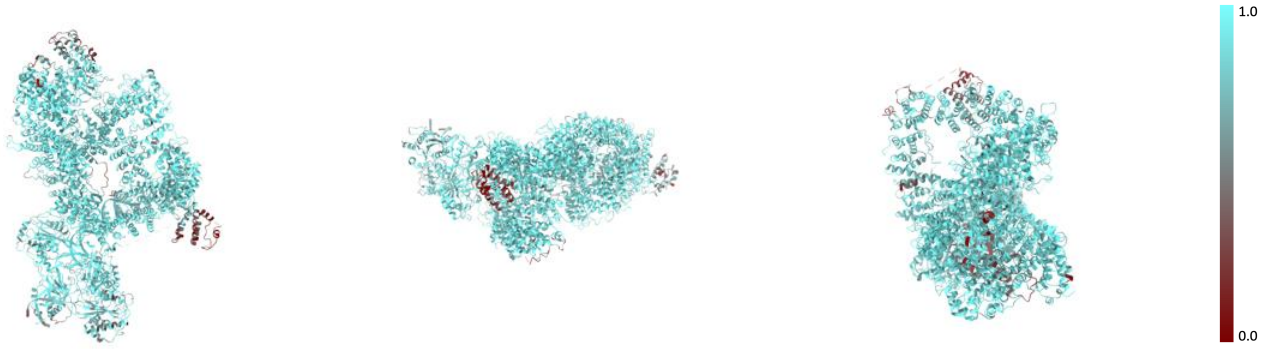
The images above show the 3D surface view of the map at the recommended contour level 0.012 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2 Q-score mapped to coordinate model (i)

X Y Z

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

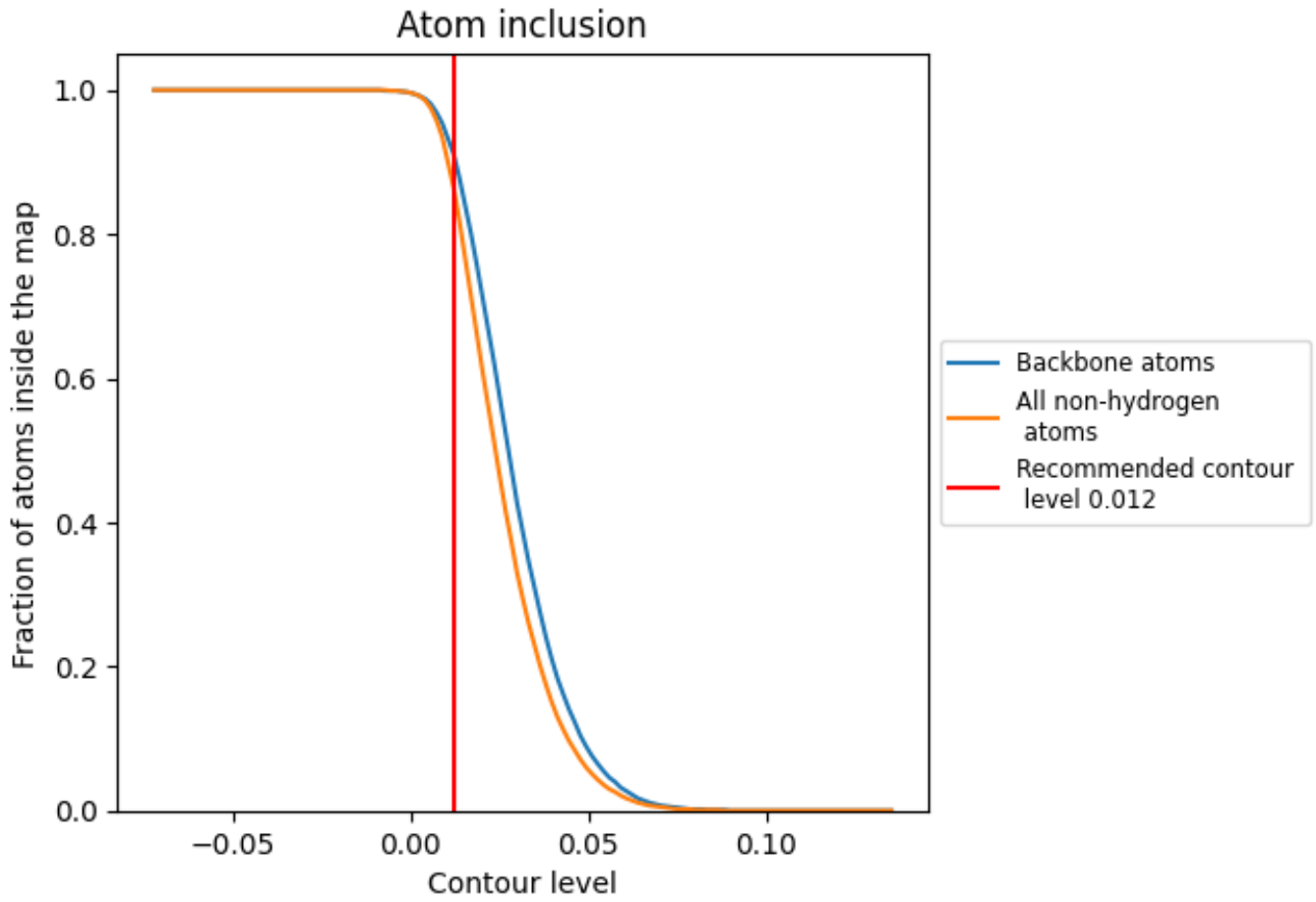
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.012).

9.4 Atom inclusion (i)



At the recommended contour level, 91% of all backbone atoms, 87% 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 (0.012) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8680	0.4410
A	0.8900	0.4600
B	0.8650	0.4240
C	0.7390	0.3700
D	0.9360	0.4050
E	0.7810	0.2720
F	0.8960	0.3270
G	0.9580	0.4090