wwPDB EM Validation Summary Report

Apr 22, 2024 – 05:37 pm BST

PDB ID	:	6ZHA
EMDB ID	:	EMD-11217
Title	:	Cryo-EM structure of DNA-PK monomer
Authors	:	Chaplin, A.K.; Hardwick, S.W.; Chirgadze, D.Y.; Blundell, T.L.
Deposited on	:	2020-06-21
Resolution	:	3.91 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:

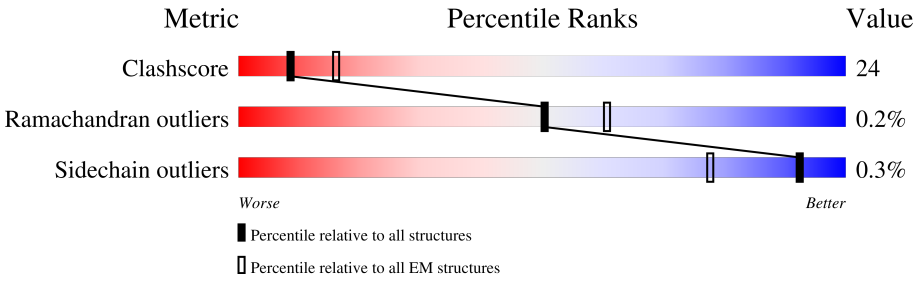
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 3.91 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 <=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	4156	51% 38% 11%
2	B	609	36% 44% 20%
3	C	732	35% 38% 26%
4	D	25	8% 92%
5	E	24	12% 83%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 38289 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,DNA-dependent protein kinase catalytic subunit,DNA-dependent protein kinase catalytic subunit,DNA-PKcs.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	3707	Total	C	N	O	S	0	0
			29124	18685	4920	5334	185		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	489	Total	C	N	O	S	0	0
			3908	2504	663	723	18		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	539	Total	C	N	O	S	0	0
			4260	2733	717	787	23		

- Molecule 4 is a DNA chain called DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	25	Total	C	N	O	P	0	0
			510	248	94	144	24		

- Molecule 5 is a DNA chain called DNA.

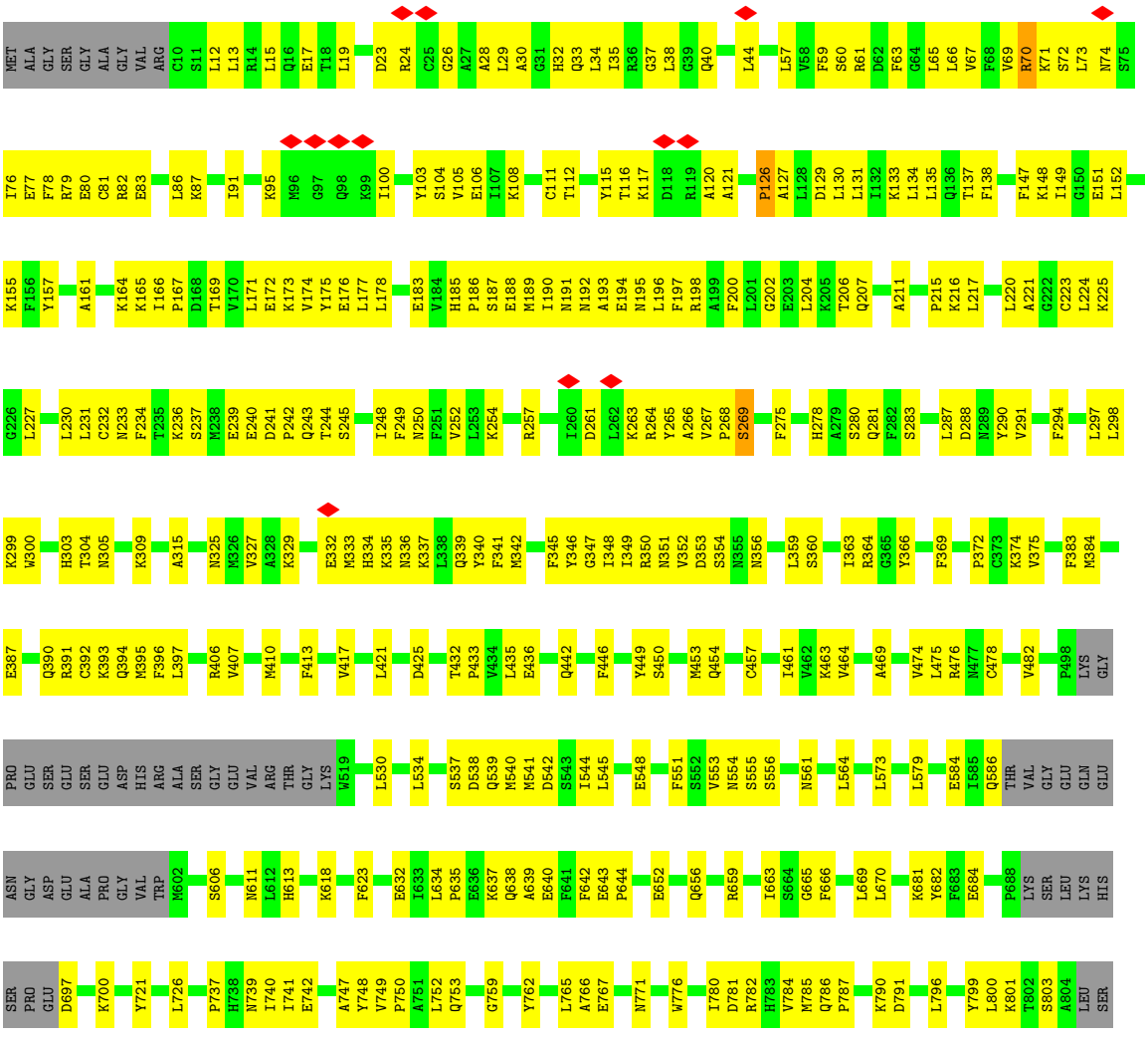
Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	24	Total	C	N	O	P	0	0
			487	240	80	145	22		

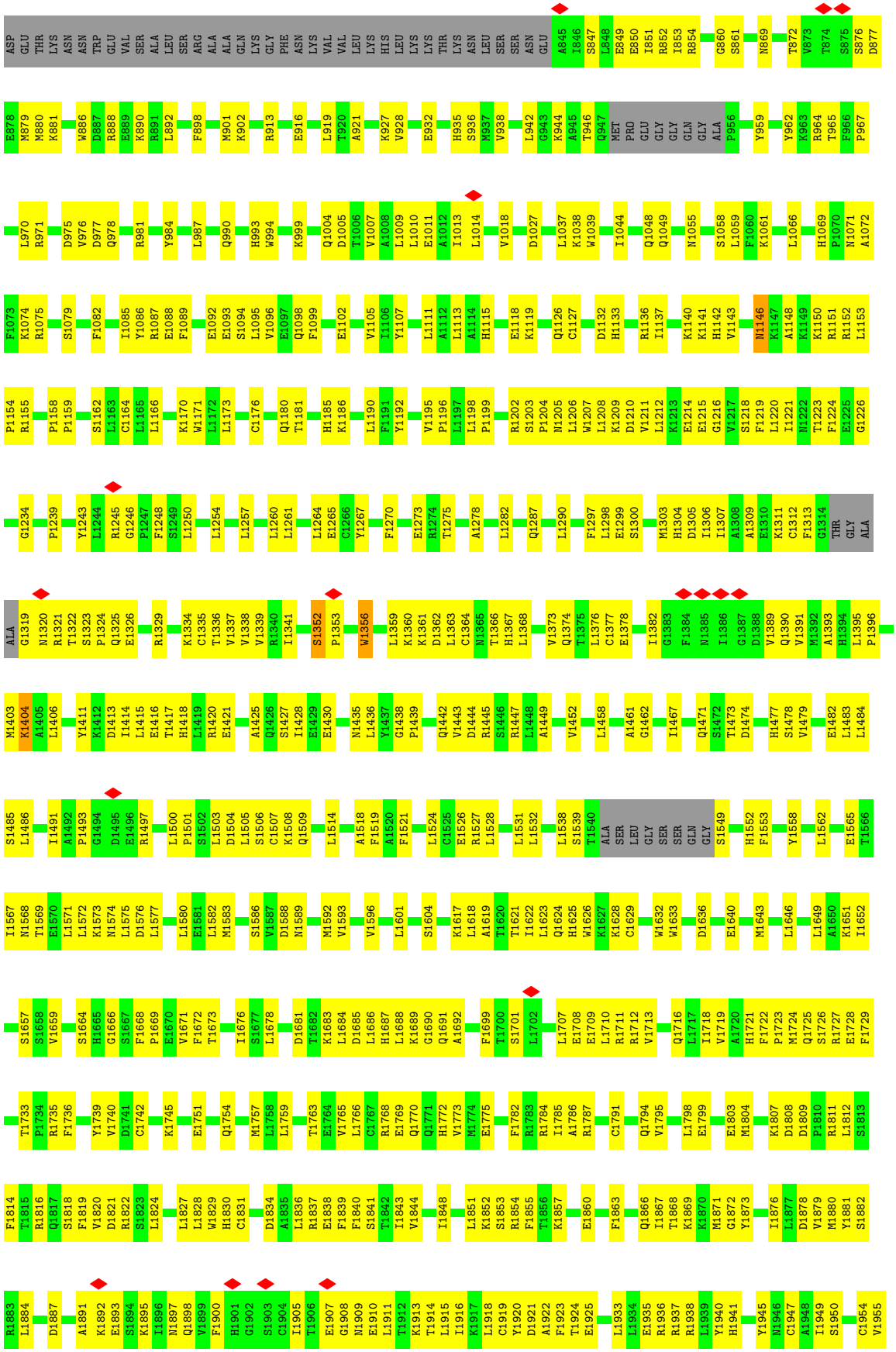
3 Residue-property plots (i)

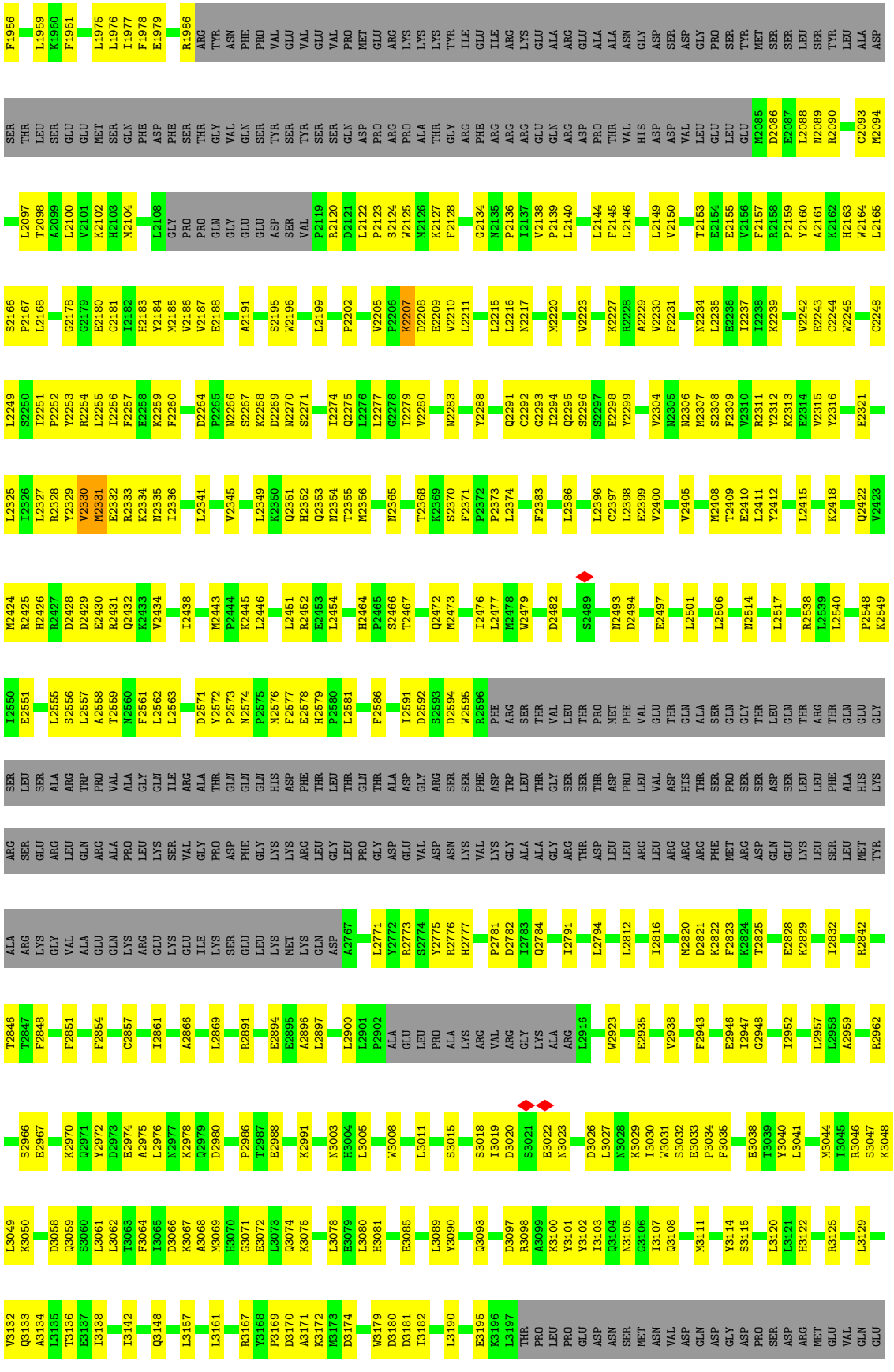
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion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

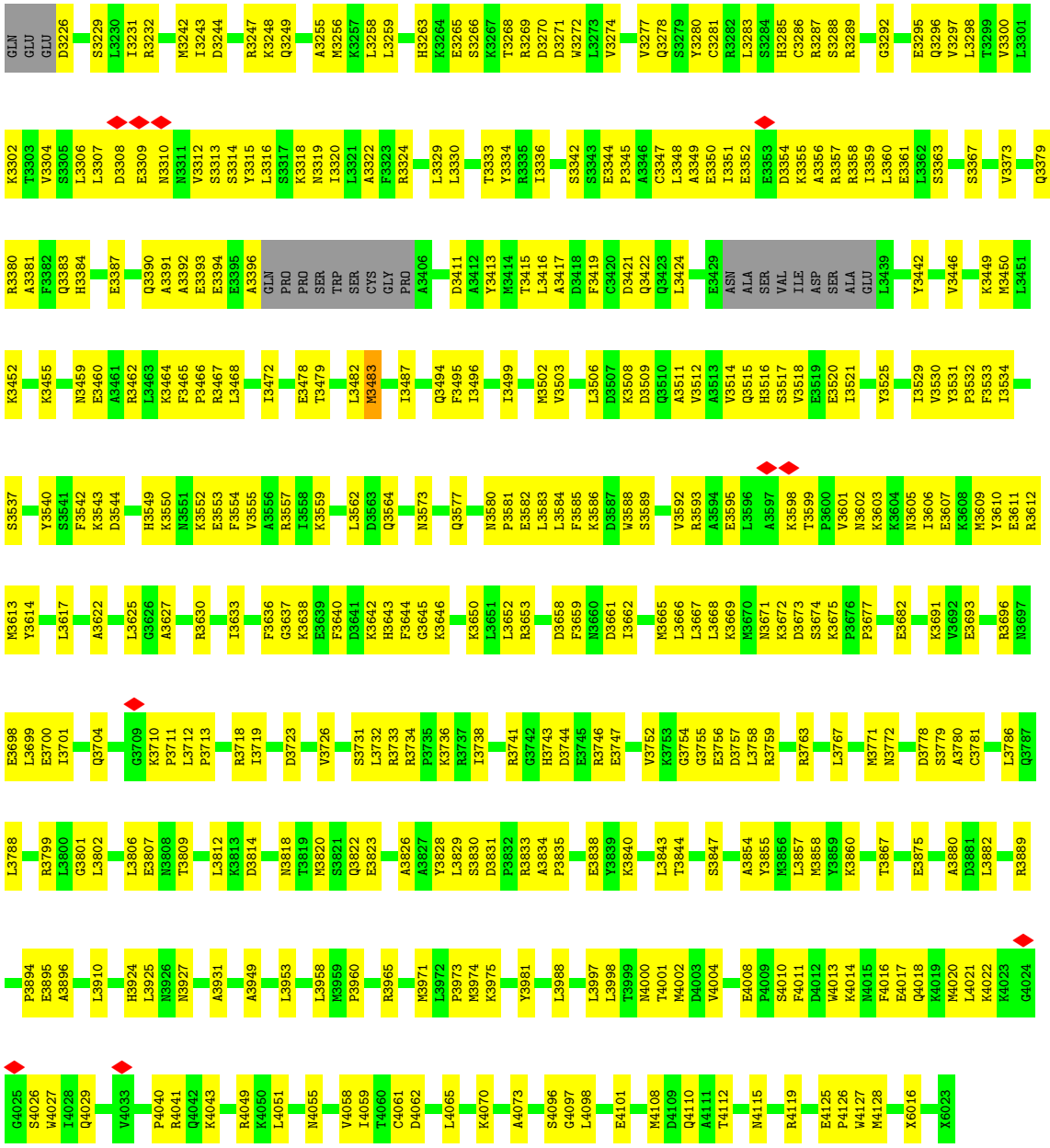
• Molecule 1: DNA-dependent protein kinase catalytic subunit,DNA-dependent protein kinase catalytic subunit,DNA-dependent protein kinase catalytic subunit,DNA-PKcs

Chain A:



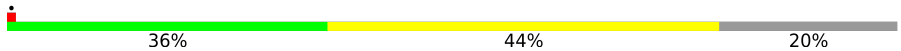

ASP GLU THR LYS ASN ASN TRP GLU VAL SER ALA LEU SER ARG ALA ALA GLN LYS GLY PHE ASN LYS VAL VAL LEU LYS HIS LEU LYS LYS THR LYS ASN LEU SER SER ASN GLU A845 I846 S847 L848 E849 E850 I851 R852 I853 R854 G860 S861 N869 T872 V873 T874 S875 S876 D877
E878 M879 M880 K881 W886 D887 R888 E889 K890 R891 L892 F898 M901 K902 R913 E916 L919 T920 A921 K927 V928 E932 H935 S936 M937 V938 L942 G943 K944 A945 T946 Q947 MET PRO GLU GLY GLY GLN GLY ALA P956 Y969 Y962 K963 R964 T965 F966 P967
L970 R971 D975 V976 D977 Q978 R981 Y984 L987 Q990 H993 W994 K999 Q1004 D1005 T1006 V1007 A1008 L1009 L1010 E1011 A1012 I1013 L1014 V1018 D1027 L1037 K1038 W1039 I1044 Q1048 Q1049 N1055 S1058 L1059 F1060 K1061 L1066 H1069 P1070 N1071 A1072
F1073 K1074 R1075 S1079 F1082 I1085 Y1086 R1087 E1088 F1089 E1092 E1093 S1094 L1095 V1096 E1097 Q1098 F1099 E1102 V1105 I1106 Y1107 L1111 A1112 L1113 A1114 H1115 E1118 K1119 Q1126 C1127 D1132 H1133 R1136 I1137 K1140 K1141 H1142 V1143 N1146 K1147 A1148 K1149 K1150 R1151 R1152 L1153
P1154 R1155 P1158 P1159 S1162 L1163 C1164 L1165 L1166 K1170 W1171 L1172 L1173 C1176 Q1180 T1181 H1185 K1186 L1190 F1191 Y1192 V1195 P1196 L1197 L1198 P1199 R1202 S1203 P1204 N1205 L1206 W1207 L1208 K1209 D1210 V1211 L1212 K1213 E1214 E1215 G1216 V1217 S1218 F1219 L1220 I1221 N1222 T1223 F1224 E1225 G1226
G1234 P1239 Y1243 L1244 R1245 G1246 P1247 F1248 S1249 L1250 L1254 L1257 L1260 L1261 L1264 E1265 C1266 Y1267 F1270 E1273 R1274 T1275 A1278 L1282 Q1287 L1290 F1297 L1298 E1299 S1300 M1303 H1304 D1305 I1306 I1307 A1308 A1309 E1310 K1311 C1312 F1313 G1314 THR GLY ALA
ALA G1319 N1320 R1321 T1322 S1323 P1324 Q1325 E1326 R1329 K1334 C1335 T1336 V1337 V1338 V1339 R1340 I1341 S1352 P1353 W1356 L1359 K1360 K1361 D1362 L1363 C1364 N1365 T1366 H1367 L1368 V1373 Q1374 T1375 L1376 C1377 E1378 I1382 G1383 F1384 N1385 I1386 G1387 D1388 V1389 Q1390 V1391 M1392 A1393 H1394 L1395 P1396
M1403 K1404 A1405 L1406 Y1411 K1412 D1413 I1414 L1415 E1416 T1417 H1418 L1419 R1420 E1421 A1425 Q1426 S1427 I1428 E1429 E1430 N1435 L1436 Y1437 G1438 P1439 Q1442 V1443 D1444 R1445 S1446 R1447 L1448 A1449 V1452 L1458 A1461 G1462 I1467 Q1471 S1472 T1473 D1474 H1477 S1478 V1479 E1482 L1483 L1484
S1485 L1486 I1491 A1492 P1493 G1494 D1495 E1496 R1497 L1500 P1501 S1502 L1503 D1504 L1505 S1506 C1507 K1508 Q1509 L1514 A1518 F1519 A1520 F1521 L1524 C1525 E1526 R1527 L1528 L1531 L1532 L1538 S1539 T1540 ALA SER LEU GLY SER SER GLN GLY S1549 H1552 F1553 Y1558 L1562 E1565 T1566
I1567 N1568 T1569 E1570 L1571 L1572 K1573 N1574 L1575 D1576 L1577 L1580 E1581 L1582 M1583 S1586 V1587 D1588 N1589 M1592 V1593 V1596 L1601 S1604 K1617 L1618 A1619 T1620 T1621 I1622 L1623 Q1624 H1625 W1626 K1627 K1628 C1629 W1632 W1633 D1636 E1640 M1643 L1646 L1649 A1650 K1651 I1652
S1657 S1658 V1659 S1664 H1665 G1666 S1667 F1668 P1669 E1670 V1671 F1672 T1673 I1676 S1677 L1678 D1681 T1682 K1683 L1684 D1685 L1686 H1687 L1688 K1689 G1690 Q1691 A1692 P1699 T1700 S1701 L1702 L1707 E1708 E1709 L1710 R1711 R1712 V1713 Q1716 L1717 I1718 V1719 A1720 H1721 F1722 P1723 M1724 Q1725 S1726 R1727 E1728 F1729
T1733 P1734 R1735 F1736 Y1739 V1740 D1741 C1742 K1745 E1751 Q1754 M1757 L1758 L1759 T1763 E1764 V1765 L1766 C1767 R1768 E1769 Q1770 Q1771 H1772 V1773 M1774 E1775 F1782 R1783 R1784 I1785 A1786 R1787 C1791 Q1794 V1795 L1798 E1799 E1803 M1804 K1807 D1808 D1809 P1810 R1811 L1812 S1813
F1814 T1815 R1816 Q1817 S1818 F1819 V1820 D1821 R1822 S1823 L1824 L1827 L1828 W1829 H1830 C1831 D1834 A1835 L1836 R1837 E1838 F1839 F1840 S1841 T1842 I1843 V1844 I1848 K1851 K1852 S1853 R1854 F1855 T1856 K1857 E1860 F1863 Q1866 I1867 T1868 K1869 K1870 M1871 G1872 Y1873 I1876 L1877 D1878 V1879 M1880 Y1881 S1882
R1883 L1884 D1887 A1891 K1892 E1893 S1894 K1895 I1896 N1897 Q1898 V1899 F1900 H1901 G1902 S1903 C1904 I1905 T1906 E1907 G1908 N1909 E1910 L1911 T1912 K1913 T1914 L1915 I1916 K1917 L1918 C1919 Y1920 D1921 A1922 F1923 T1924 E1925 L1933 L1934 E1935 R1936 R1937 R1938 L1939 Y1940 H1941 Y1945 N1946 C1947 A1948 I1949 S1950 C1954 V1955

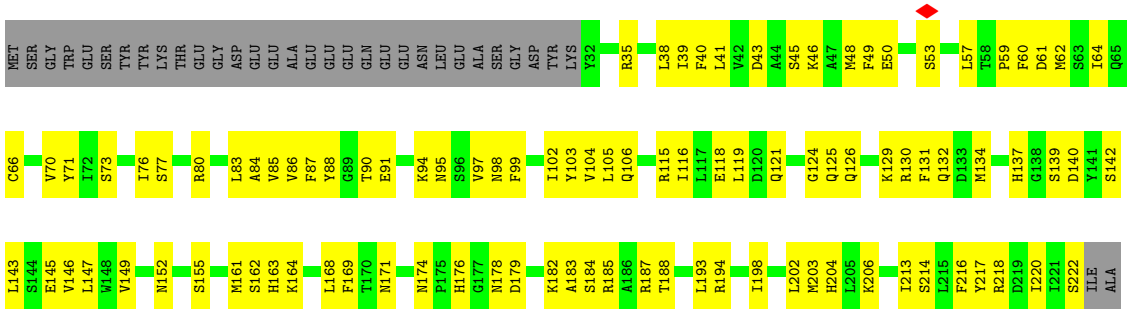


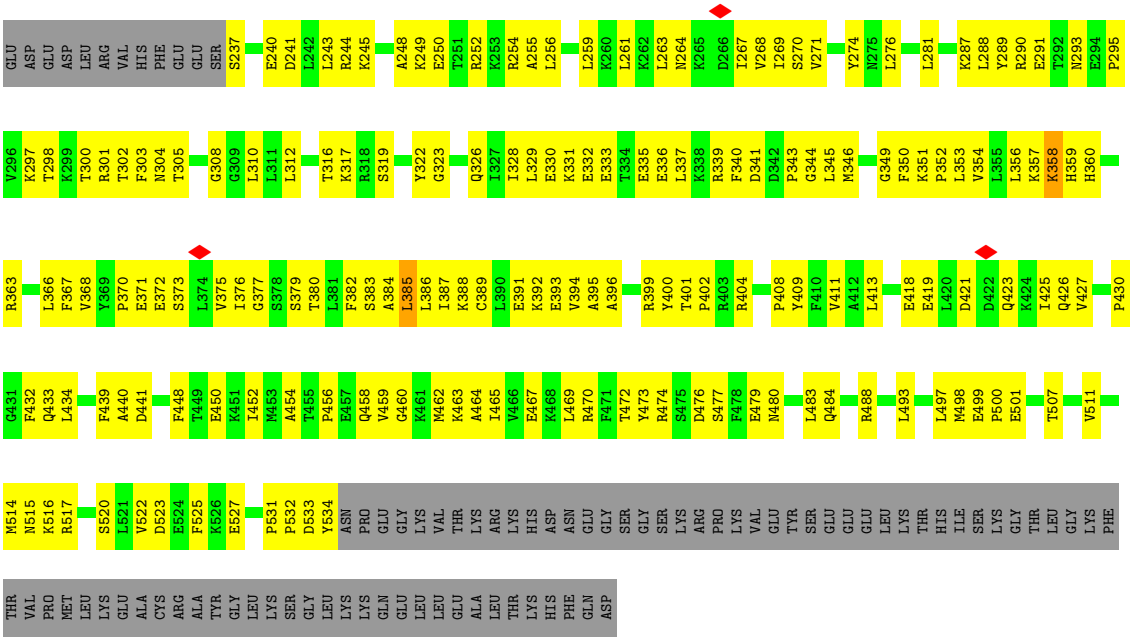


• Molecule 2: X-ray repair cross-complementing protein 6

Chain B:

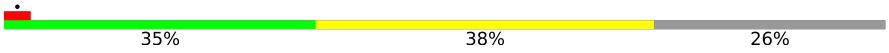


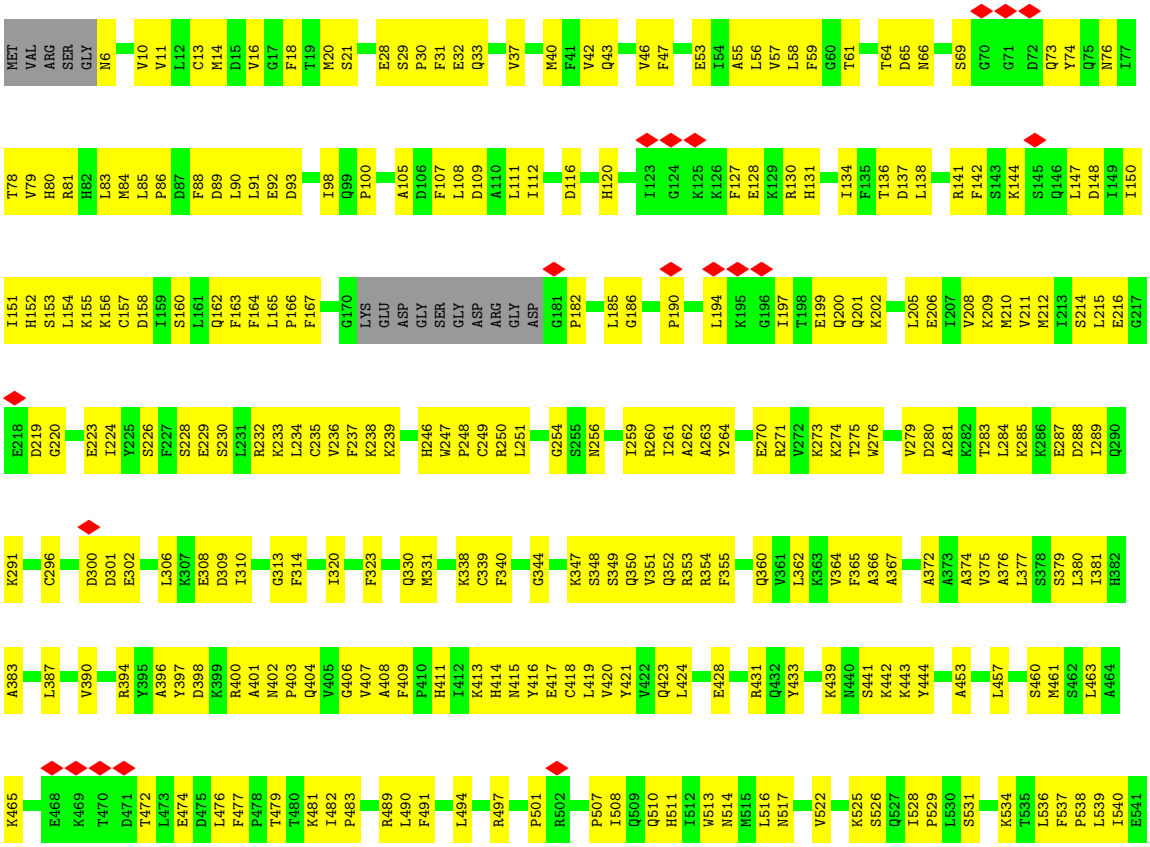


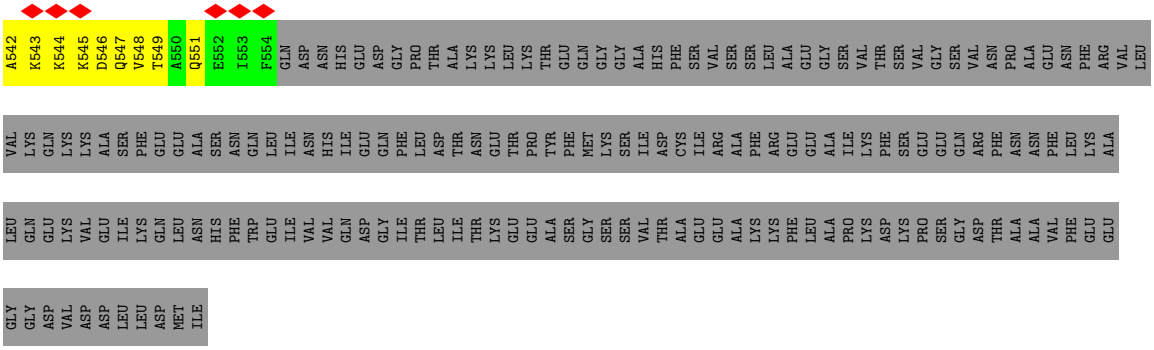


• Molecule 3: X-ray repair cross-complementing protein 5

Chain C:

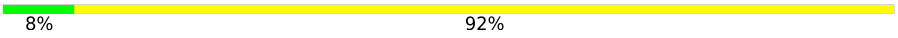


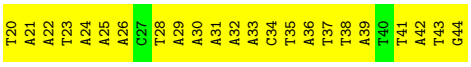




- Molecule 4: DNA

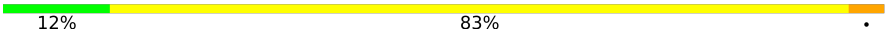
Chain D:

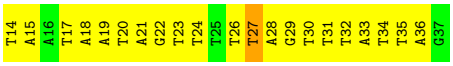




- Molecule 5: DNA

Chain E:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	59967	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 ($e^-/Å^2$)	52.97	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	0.525	Depositor
Minimum map value	-0.130	Depositor
Average map value	0.002	Depositor
Map value standard deviation	0.025	Depositor
Recommended contour level	0.125	Depositor
Map size (Å)	356.99997, 356.99997, 356.99997	wwPDB
Map dimensions	340, 340, 340	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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	0.44	0/29576	0.52	1/40030 (0.0%)
2	B	0.39	0/3984	0.53	0/5371
3	C	0.34	0/4346	0.50	0/5867
4	D	0.98	0/573	1.01	0/882
5	E	0.90	1/544 (0.2%)	1.11	0/837
All	All	0.45	1/39023 (0.0%)	0.55	1/52987 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	27	DT	C3'-O3'	-5.36	1.36	1.44

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	126	PRO	N-CA-CB	6.21	110.76	103.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	29124	0	29016	1292	0
2	B	3908	0	3962	253	0
3	C	4260	0	4262	260	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	510	0	286	24	0
5	E	487	0	279	23	0
All	All	38289	0	37805	1788	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 24.

The worst 5 of 1788 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:3472:ILE:HG21	1:A:3483:MET:CE	1.54	1.36
1:A:2330:VAL:CB	1:A:2336:ILE:HG22	1.55	1.35
1:A:2330:VAL:HG23	1:A:2335:ASN:O	1.31	1.25
2:B:370:PRO:HB3	2:B:382:PHE:CD2	1.76	1.21
1:A:2330:VAL:HB	1:A:2336:ILE:CG2	1.72	1.20

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	3649/4156 (88%)	3291 (90%)	351 (10%)	7 (0%)	47 79
2	B	485/609 (80%)	431 (89%)	53 (11%)	1 (0%)	47 79
3	C	535/732 (73%)	491 (92%)	44 (8%)	0	100 100
All	All	4669/5497 (85%)	4213 (90%)	448 (10%)	8 (0%)	50 79

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1352	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	2330	VAL
2	B	385	LEU
1	A	1390	GLN
1	A	2134	GLY

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	3153/3671 (86%)	3144 (100%)	9 (0%)	92 95
2	B	433/548 (79%)	432 (100%)	1 (0%)	93 96
3	C	466/649 (72%)	465 (100%)	1 (0%)	93 96
All	All	4052/4868 (83%)	4041 (100%)	11 (0%)	92 95

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2207	LYS
1	A	3483	MET
3	C	439	LYS
2	B	358	LYS
1	A	1146	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3927	ASN
3	C	547	GLN
1	A	1231	GLN
1	A	2365	ASN
1	A	2859	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	4128:MET	C	5009:UNK	N	93.14
1	A	5016:UNK	C	6004:UNK	N	49.17

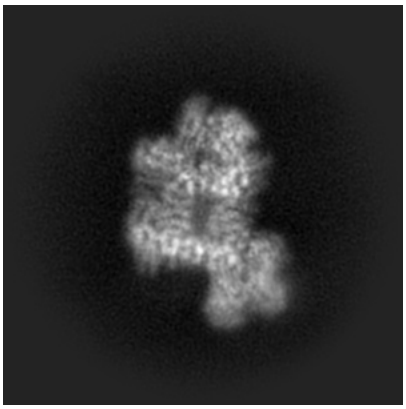
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-11217. These allow visual inspection of the internal detail of the map and identification of artifacts.

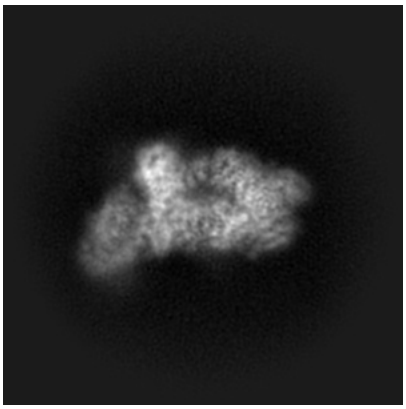
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

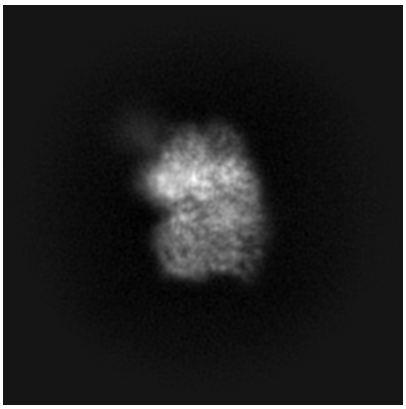
6.1.1 Primary map



X

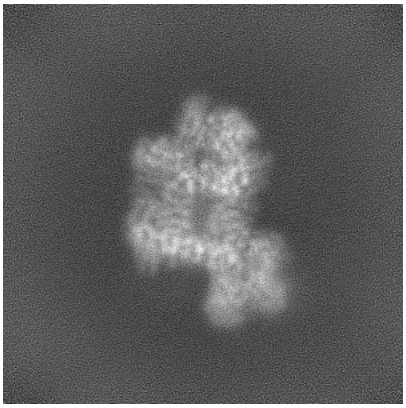


Y

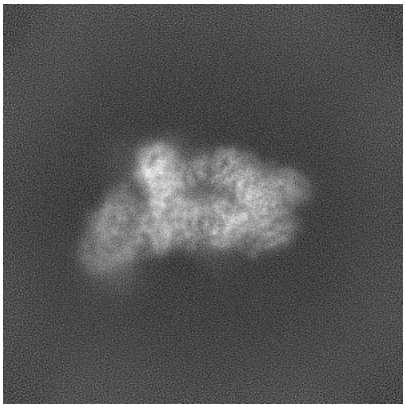


Z

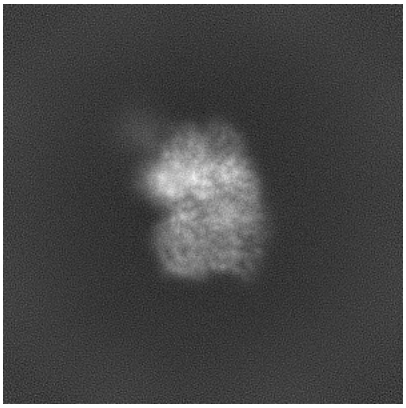
6.1.2 Raw map



X



Y

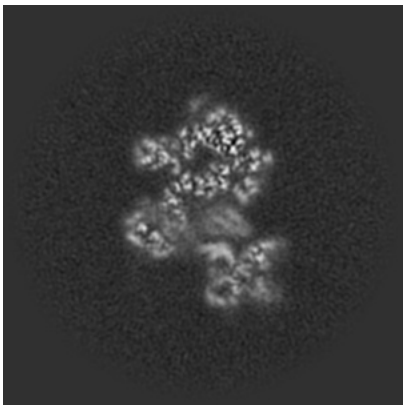


Z

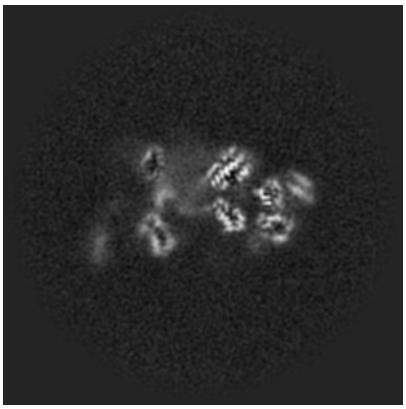
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

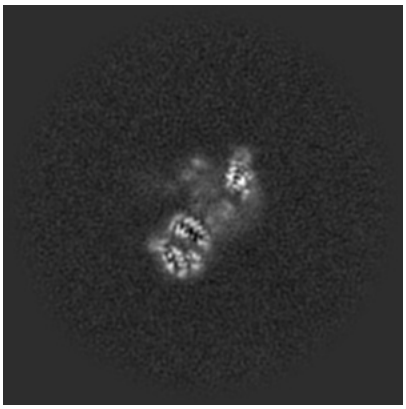
6.2.1 Primary map



X Index: 170

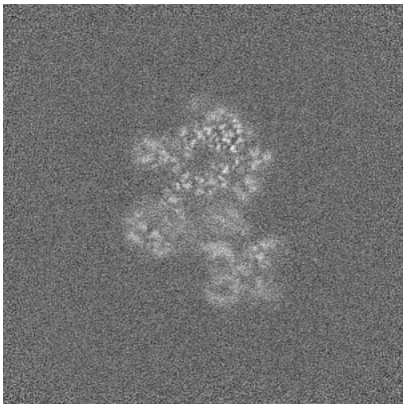


Y Index: 170

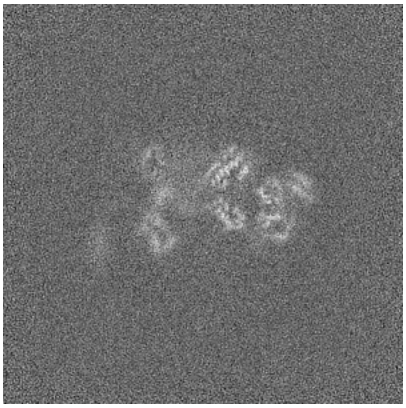


Z Index: 170

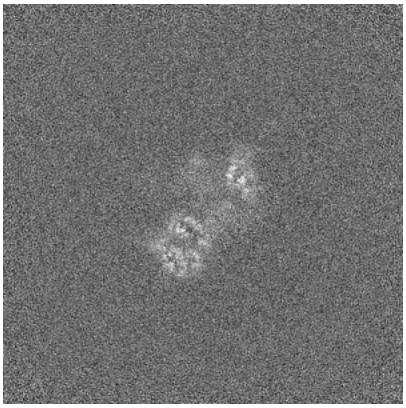
6.2.2 Raw map



X Index: 170



Y Index: 170

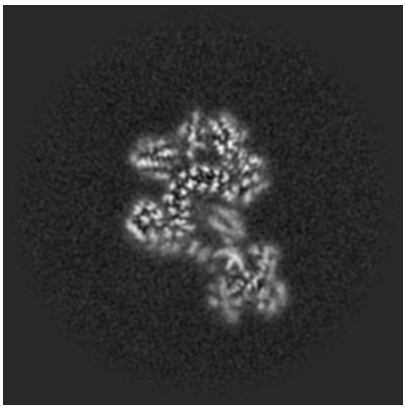


Z Index: 170

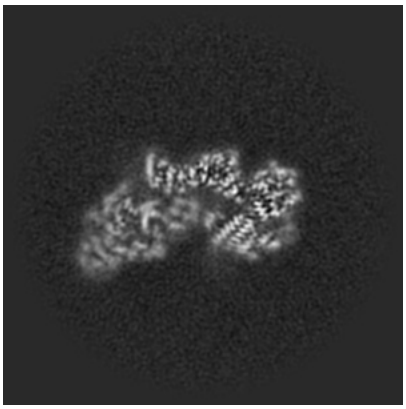
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

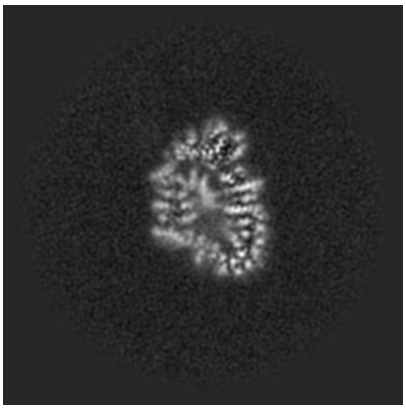
6.3.1 Primary map



X Index: 161

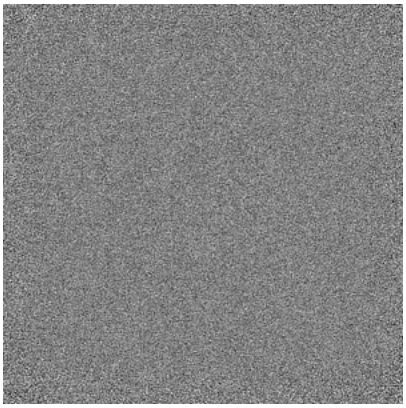


Y Index: 194

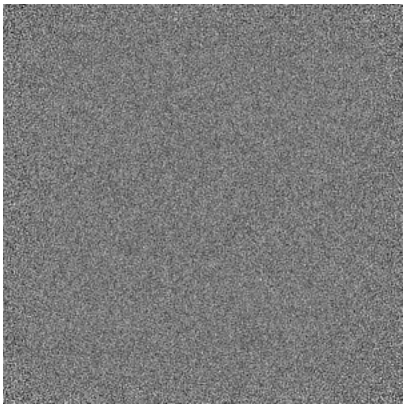


Z Index: 131

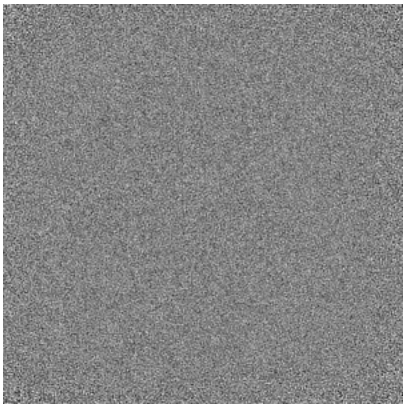
6.3.2 Raw map



X Index: 0



Y Index: 0

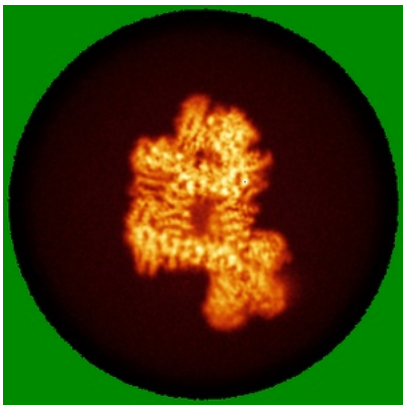


Z Index: 0

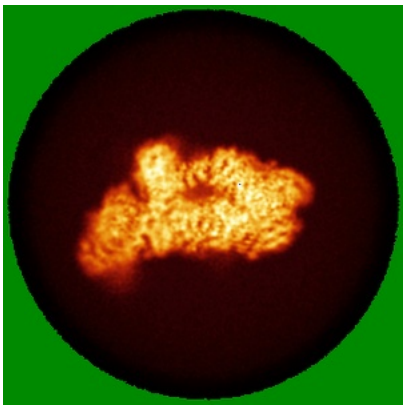
The images above show the largest variance slices of the map in three orthogonal directions.

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

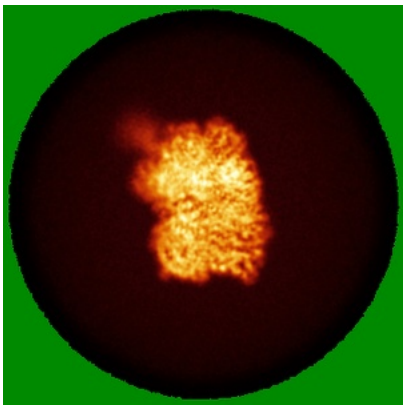
6.4.1 Primary map



X

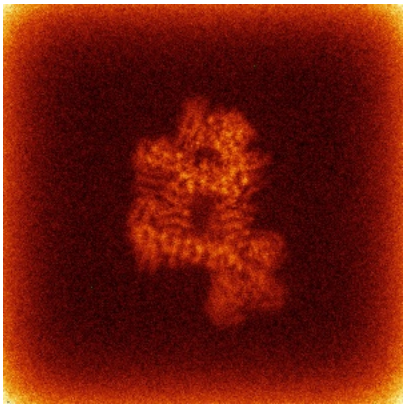


Y

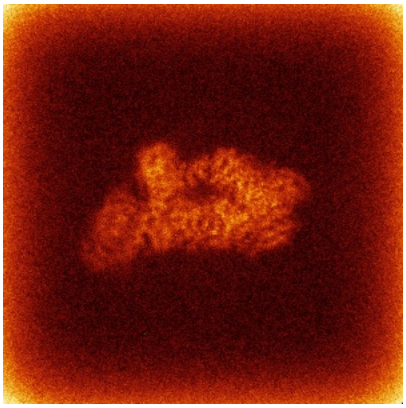


Z

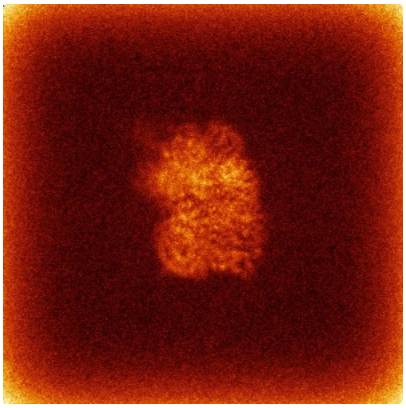
6.4.2 Raw map



X



Y

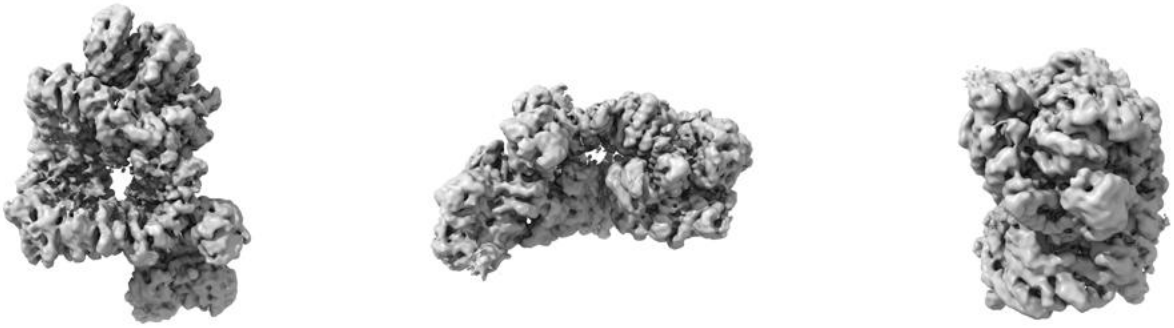


Z

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

6.5 Orthogonal surface views (i)

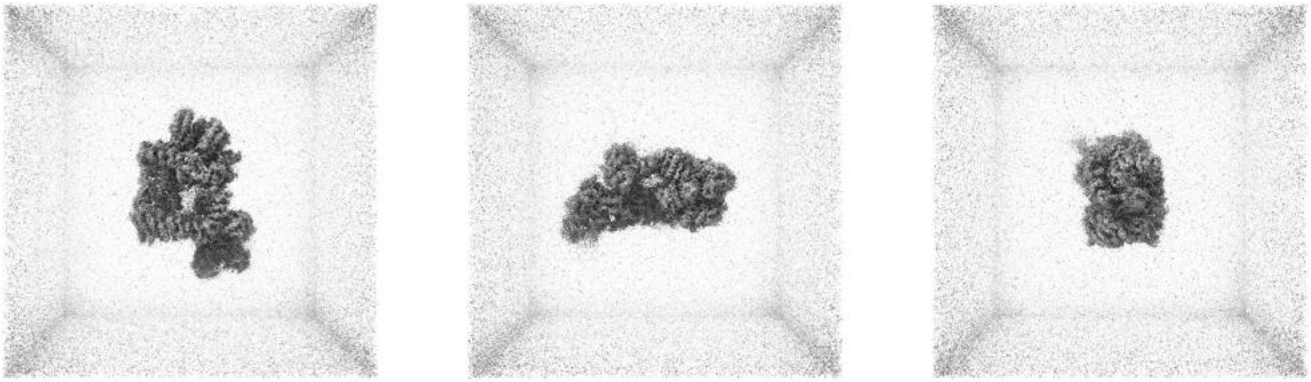
6.5.1 Primary map



X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.125. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map



X Y Z

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

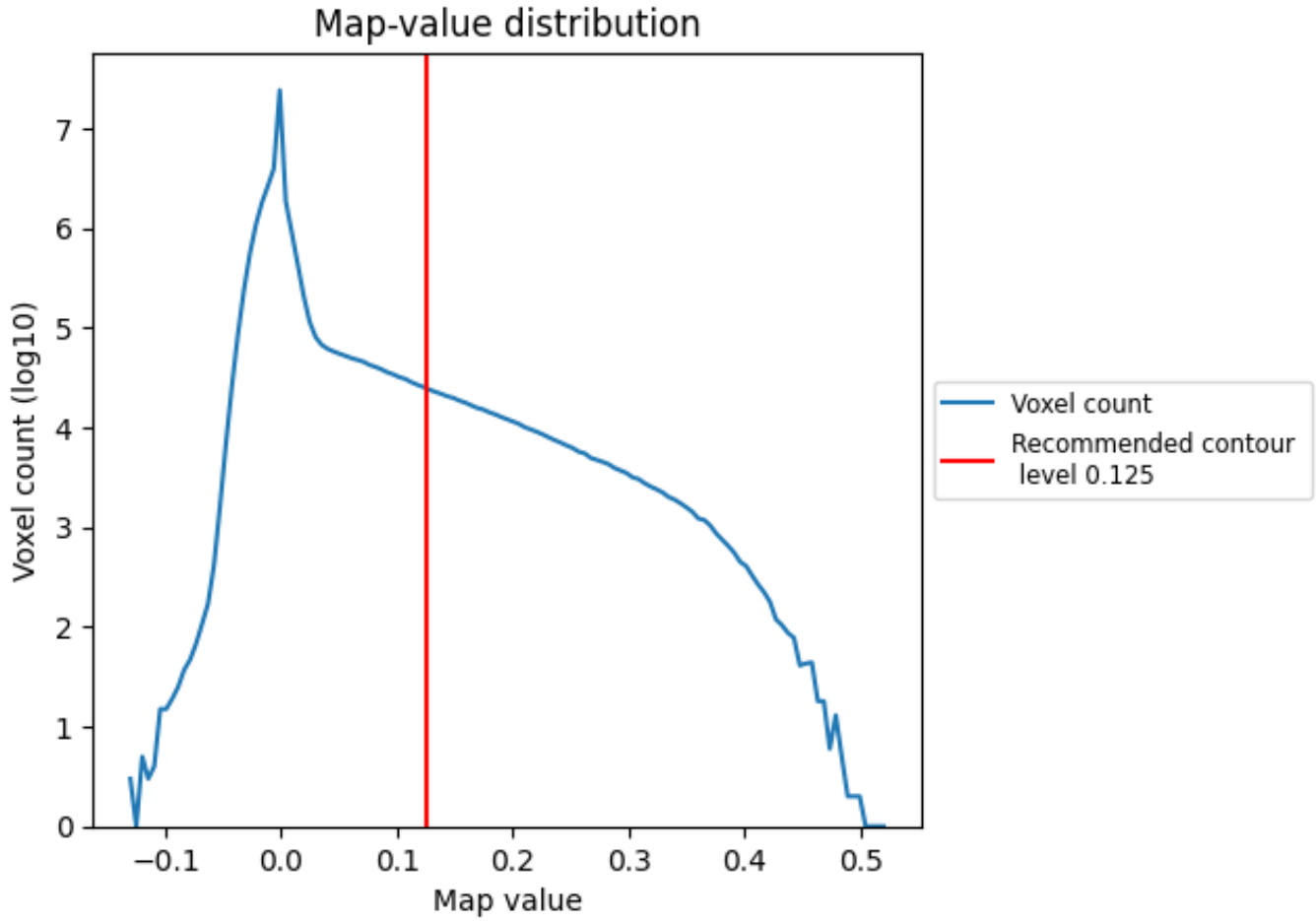
6.6 Mask visualisation (i)

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

7 Map analysis

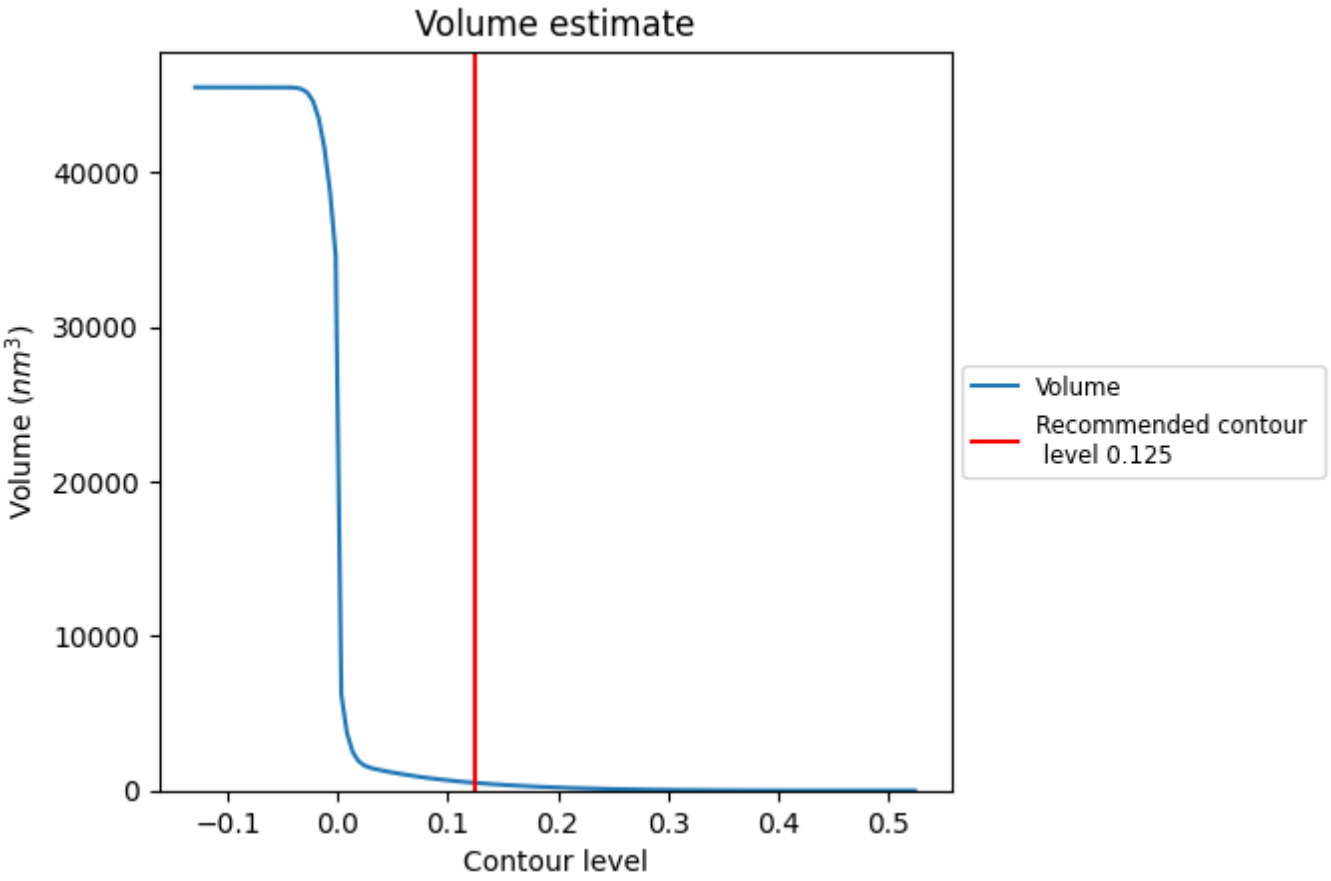
This section contains the results of statistical analysis of the map.

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.

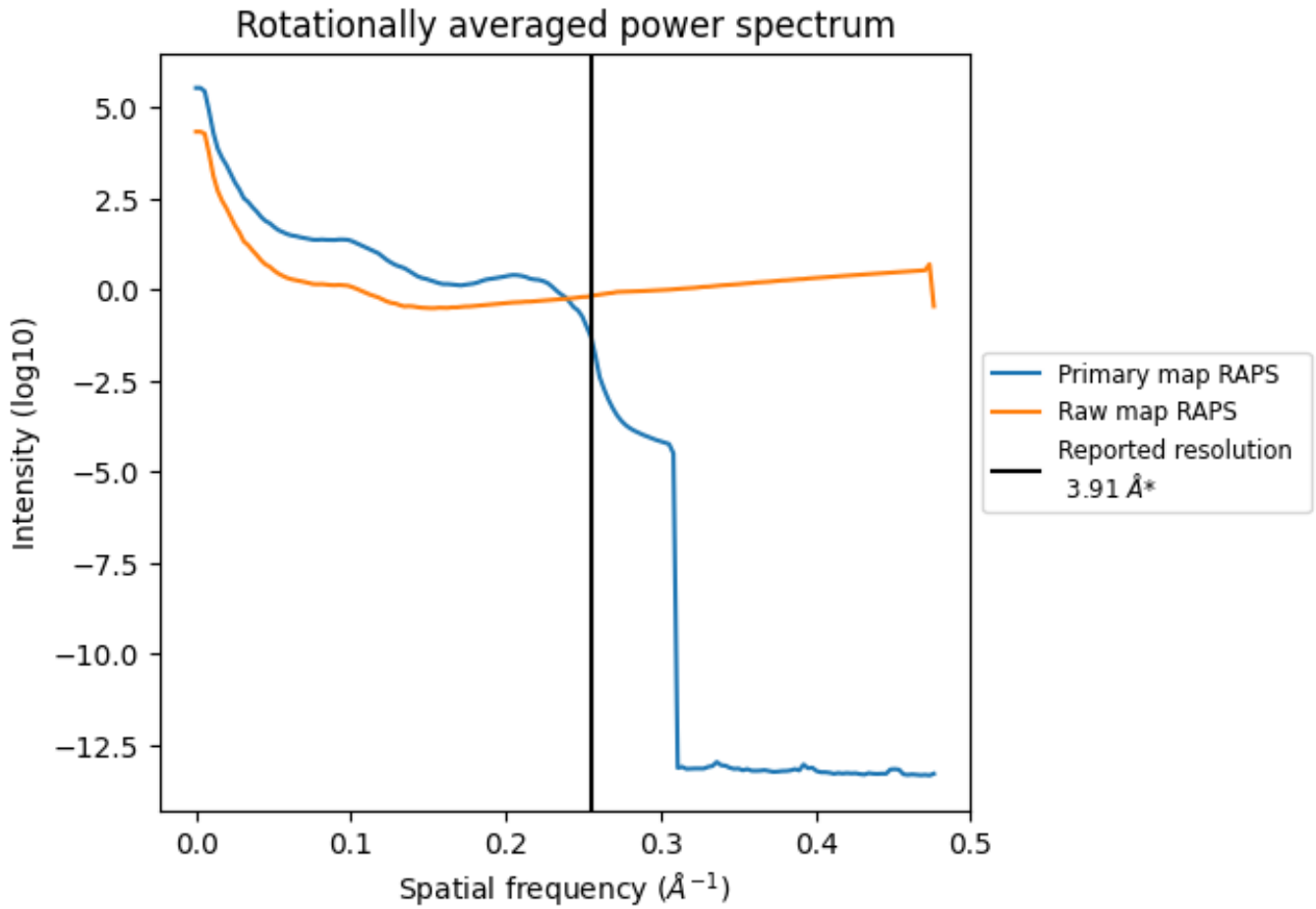
7.2 Volume estimate



The volume at the recommended contour level is 497 nm^3; this corresponds to an approximate mass of 449 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)

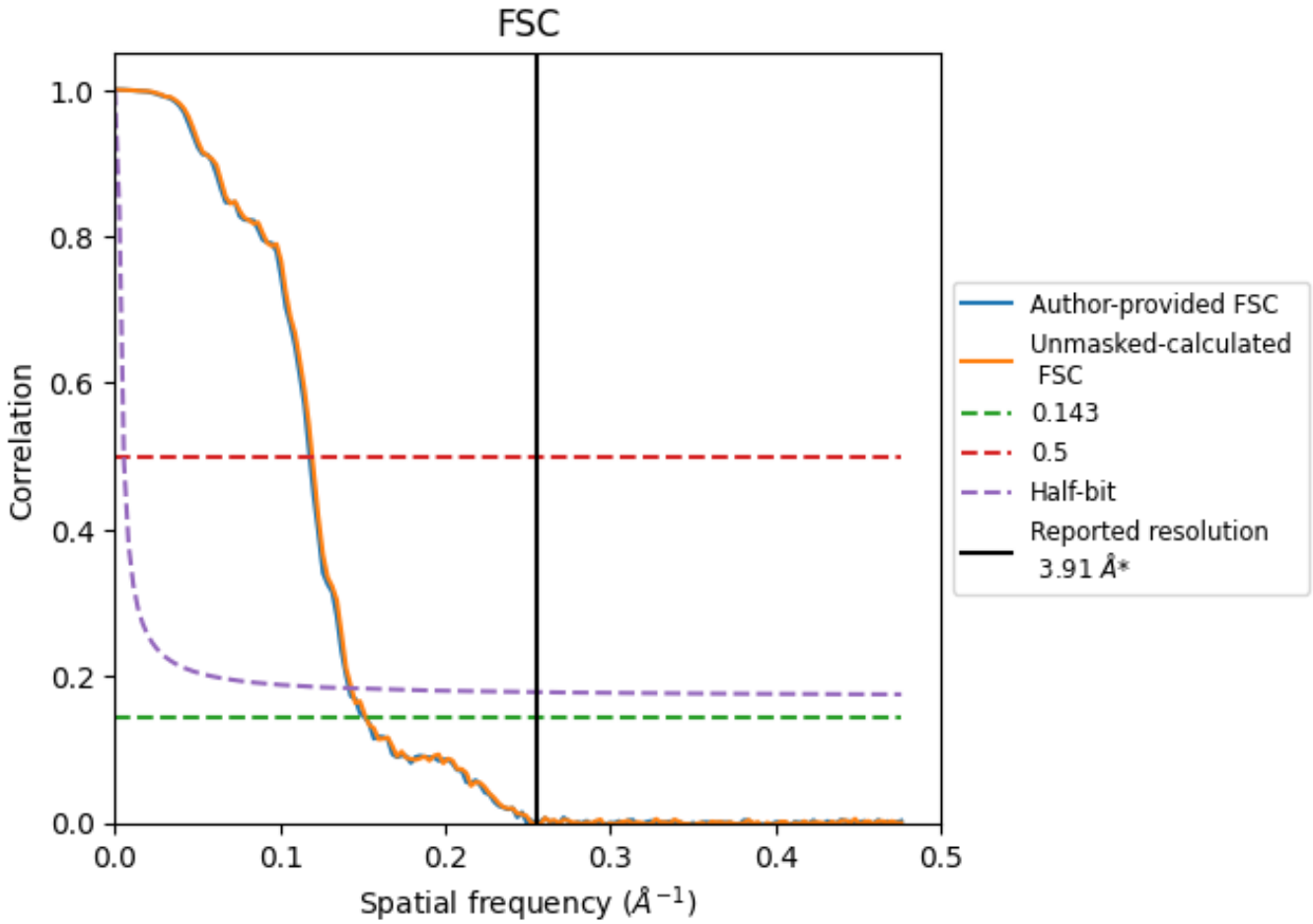


*Reported resolution corresponds to spatial frequency of 0.256 $Å^{-1}$

8 Fourier-Shell correlation

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

8.1 FSC



*Reported resolution corresponds to spatial frequency of 0.256 Å^{-1}

8.2 Resolution estimates (i)

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.91	-	-
Author-provided FSC curve	6.60	8.47	7.05
Unmasked-calculated*	6.57	8.37	6.98

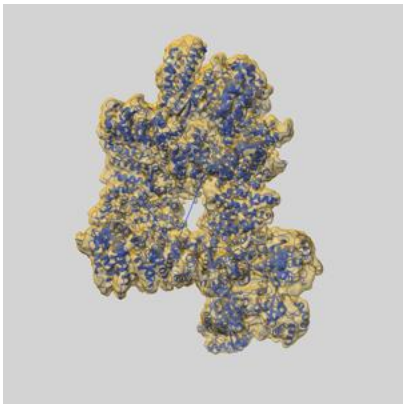
*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from author-provided FSC intersecting FSC 0.143 CUT-OFF 6.60 differs from the reported value 3.91 by more than 10 %

The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 6.57 differs from the reported value 3.91 by more than 10 %

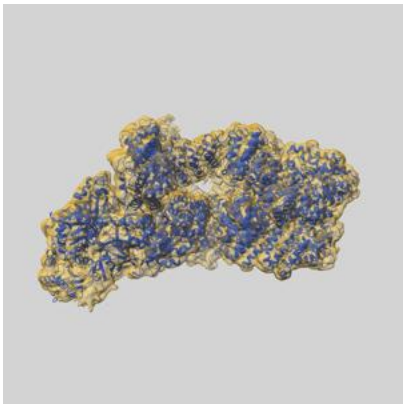
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-11217 and PDB model 6ZHA. Per-residue inclusion information can be found in section 3 on page 4.

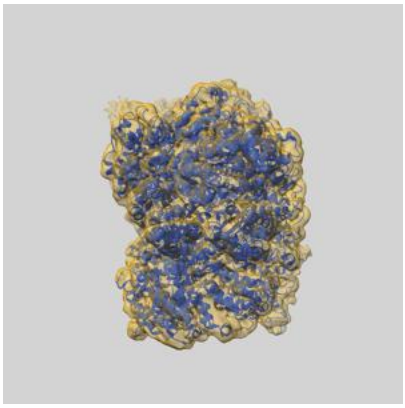
9.1 Map-model overlay



X



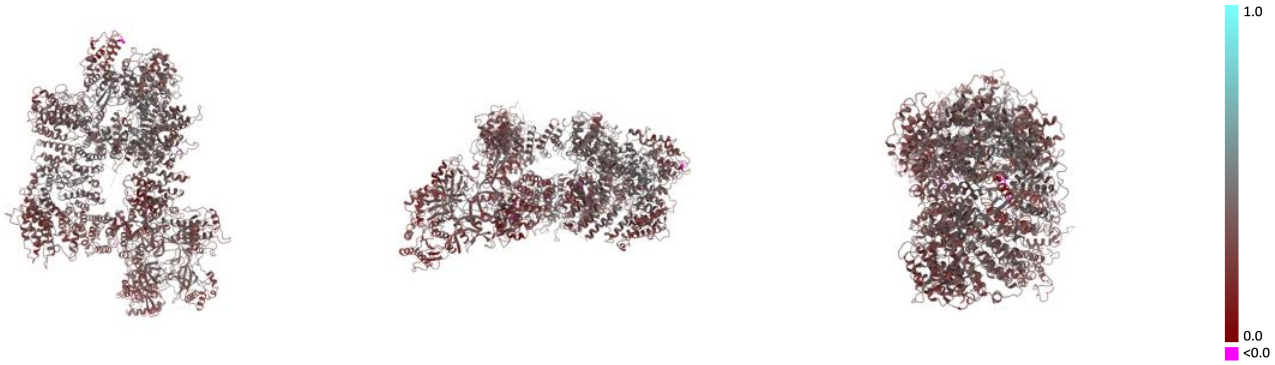
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.125 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

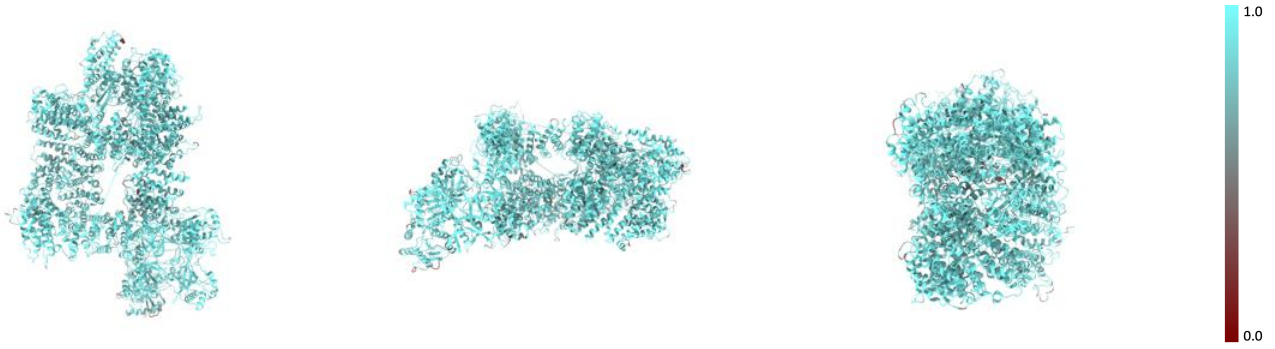
9.2 Q-score mapped to coordinate model



X Y Z

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

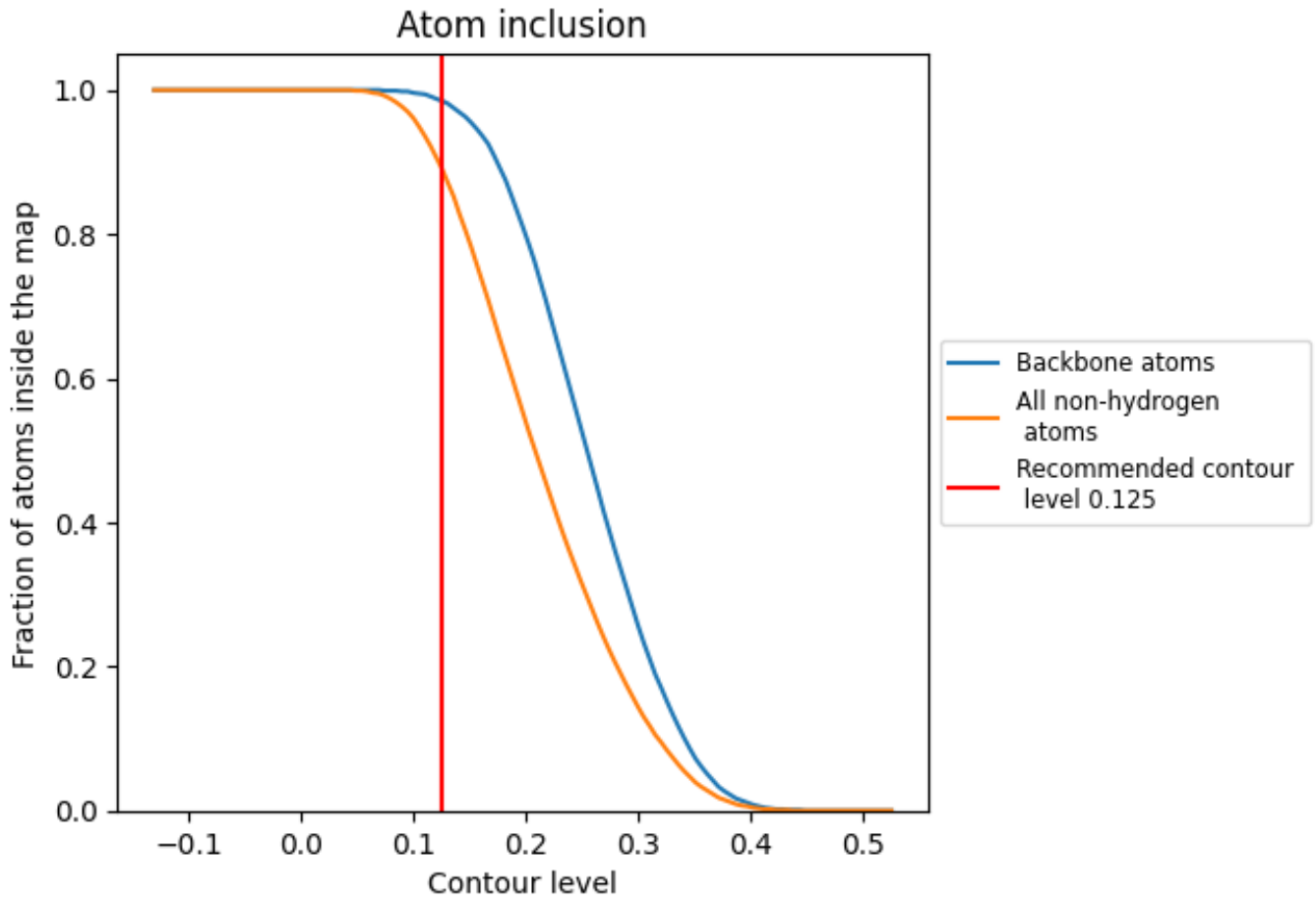
9.3 Atom inclusion mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.125).

9.4 Atom inclusion (i)



At the recommended contour level, 98% of all backbone atoms, 89% 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 (0.125) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8930	0.3380
A	0.8960	0.3460
B	0.8940	0.3370
C	0.8550	0.2850
D	0.9940	0.3340
E	0.9820	0.3100