wwPDB EM Validation Summary Report (i)

Oct 6, 2024 – 02:42 pm BST

PDB ID	:	6HV8
EMDB ID	:	EMD-0287
Title	:	Cryo-EM structure of S. cerevisiae Polymerase epsilon deltacat mutant
Authors	:	Goswami, P.; Purkiss, A.; Cheung, A.; Costa, A.
Deposited on	:	2018-10-10
Resolution	:	4.40 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:

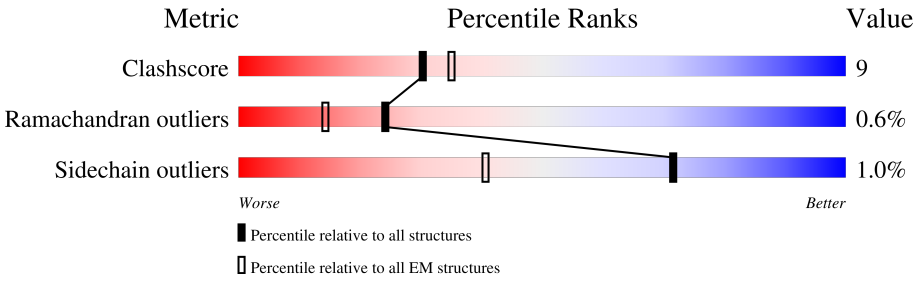
EMDB validation analysis	:	0.0.1.dev113
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 4.40 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	B	689	7% 47% 13% 40%
2	A	914	12% 64% 19% • 15%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9206 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA polymerase epsilon subunit B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	B	410	3256	2085	554	603	14	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	266	ASN	MET	conflict	UNP P24482

- Molecule 2 is a protein called DNA polymerase epsilon catalytic subunit A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	774	5948	3819	992	1104	33	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

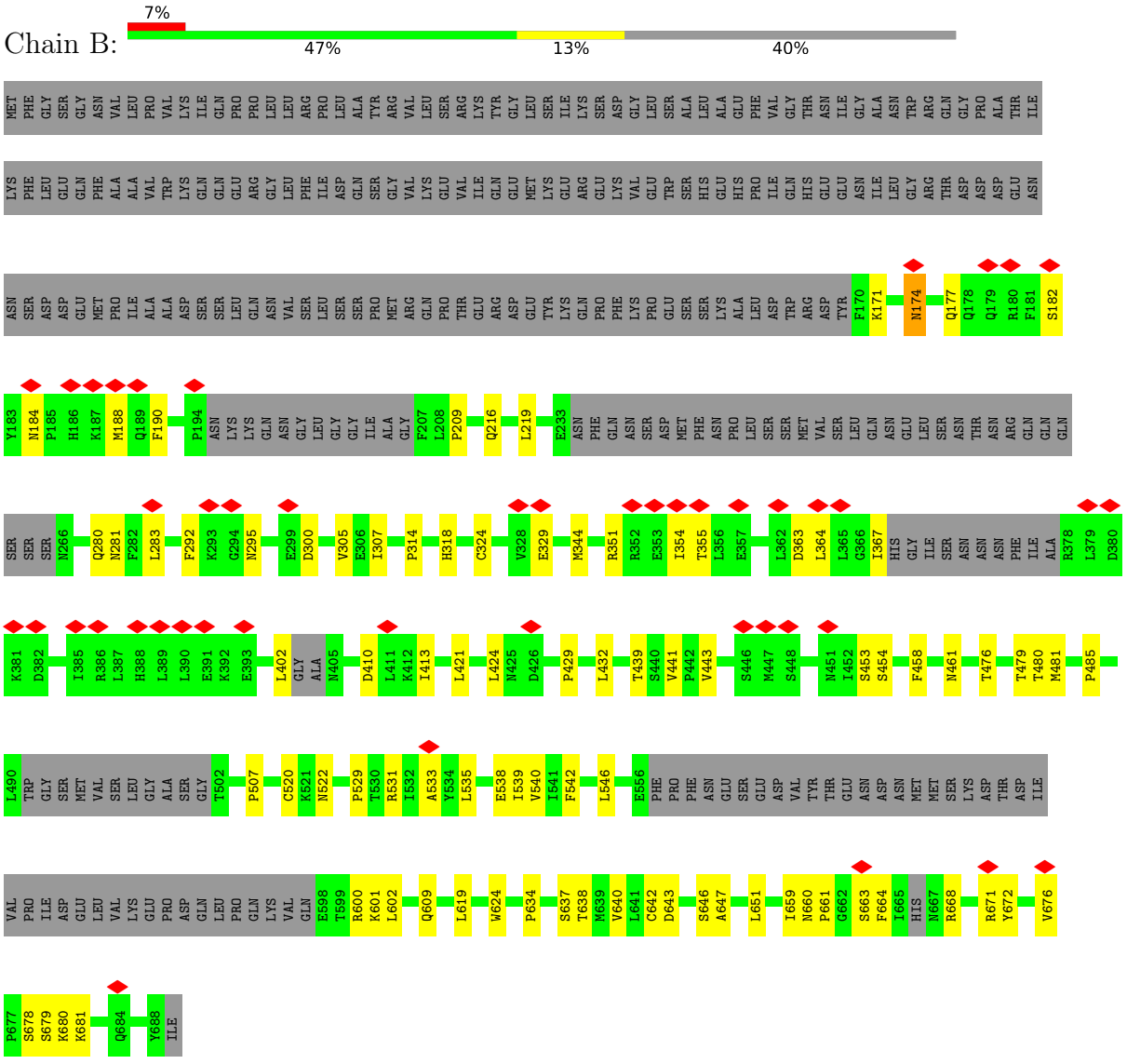
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
3	A	2	2	2	0

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 polymerase epsilon subunit B

Chain B:



• Molecule 2: DNA polymerase epsilon catalytic subunit A

Chain A:


S1308 M1309 I1310 R1311 K1312 Q1313 A1314 E1315 S1316 Y1317 A1318 N1319 S1320 T1321 W1322 L1325 S1330 G1331 V1335 L1336 E1337 V1338 F1339 V1340 T1341 I1342 N1343 G1344 I1349 T1350 F1351 H1352 Y1358 M1359 K1362 P1367 L1368 I1371 K1372 N1373 C1374 L1375 I1376 A1381 S1382 N1385 N1386 P1387 K1388 T1389
S1390 N1391 P1392 ALA GLY GLY GLN LEU PHE LYS ILE THR LEU PRO E1404 F1407 K1411 C1414 T1415 E1421 N1422 V1423 L1424 G1425 G1429 T1430 I1431 T1432 P1433 M1439 ASP LEU GLY ALA SER VAL THR PHE ARG SER LYS ALA MET GLY ALA LEU GLY LYS GLY ILE GLN GLN GLY PHE
GLU MET LYS ASP L1468 E1472 N1473 E1474 R1475 Y1476 L1477 M1482 G1485 H1489 T1492 S1493 I1494 G1495 Y1496 E1497 F1498 K1503 T1508 L1512 V1513 L1514 K1515 P1516 Q1519 E1522 I1523 N1524 A1525 S1526 S1527 I1531 Y1532 K1533 F1536 K1540 I1543 E1544 L1549 V1550 D1551
I1552 K1553 E1559 T1564 K1568 L1569 Y1570 R1571 R1572 Q1575 L1580 K1581 G1585 L1586 L1589 L1590 L1591 L1592 Q1593 S1594 P1595 F1596 I1597 T1598 K1599 L1600 T1603 L1606 L1607 N1608 P1611 I1612 V1613 K1614 L1615 S1616 L1617 N1618 E1619 V1620 L1621 L1622 Q1628 P1629 T1630 L1640 I1646
L1649 N1657 I1660 C1661 N1662 L1663 R1664 L1665 D1666 S1667 M1668 L1675 R1678 I1685 V1686 L1687 W1688 W1689 N1690 L1695 P1696 G1699 G1700 I1701 Q1702 N1703 ASP PHE ASP LEU ASN THR SER W1711 I1712 M1713 S1716 E1717 F1718 P1719 K1720 I1721 V1726 Y1727 D1728 N1729 V1730 V1731 L1732 D1733
D1737 N1742 T1743 I1744 L1745 T1746 S1747 ALA LEU ILE ASN ASP ALA GLU GLY SER ASP LEU VAL ASN ASN ASN MET GLY ILE ASP ASP LYS ASP ALA VAL ILE ASN SER PRO SER GLU PHE VAL HIS ASP ALA PHE S1784 G1793 M1794 L1795 K1796 K1804 K1826 L1827 F1828 L1832
V1836 T1840 E1850 Y1860 A1861 D1862 R1863 N1864 Q1865 I1866 L1867 I1868 K1869 T1870 Y1873 S1874 P1875 Y1879 M1886 R1890 T1891 N1892 P1893 M1894 F1895 S1896 Y1897 L1898 D1899 Y1905 W1906 L1909 I1910 W1911 M1912 D1913 K1914 F1915 N1916 F1917 A1921 C1922 I1923 E1924 I1925 N1930 Q1931 D1932 Y1933
V1936 S1937 Q1938 L1941 P1947 I1948 M1957 N1975 SER GLY THR GLN ARG PRO THR GLN ILE VAL ASN VAL LYS LYS GLN ASP K1993 E1994 D1996 S1996 V1997 E1998 D2029 P2030 A2034 ASP TYR VAL ILE PRO VAL LEU PRO GLY S2043 H2044 L2045 N2046 V2047 K2048 N2049 P2050 K2056 S2057
L2058 C2059 H2060 V2061 M2062 L2063 L2064 S2065 K2066 S2067 T2068 E2071 I2072 ARG THR LEU ARG LYS GLU LEU LEU LYS ILE PHE GLU LEU GLU ARG PHE ALA LYS VAL ALA GLU PHE LYS ASP PRO SER LEU S2100 L2101 V2102 C2108 C2111 S2115 D2118 F2119 C2120 K2121 ALA ALA PRO GLU SER
ILE F2128 S2129 L2146 K2149 L2158 R2163 C2164 S2165 R2166 C2167 H2168 D2173 S2176 A2182 G2183 A2184 W2185 L2189 P2190 R2191 T2221

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	161376	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 ($e^-/Å^2$)	30	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	0.290	Depositor
Minimum map value	-0.158	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.012	Depositor
Recommended contour level	0.0576	Depositor
Map size (Å)	200.56001, 200.56001, 200.56001	wwPDB
Map dimensions	184, 184, 184	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.09, 1.09, 1.09	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	B	0.44	0/3318	0.71	1/4493 (0.0%)
2	A	0.45	0/6061	0.85	17/8223 (0.2%)
All	All	0.45	0/9379	0.80	18/12716 (0.1%)

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1687	LEU	CA-CB-CG	7.72	133.06	115.30
2	A	1622	LEU	CA-CB-CG	7.21	131.88	115.30
2	A	2030	PRO	N-CA-CB	6.60	111.22	103.30
2	A	2063	LEU	CA-CB-CG	6.26	129.71	115.30
1	B	300	ASP	CB-CG-OD1	6.04	123.73	118.30

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	B	3256	0	3258	51	0
2	A	5948	0	5712	110	0
3	A	2	0	0	0	0
All	All	9206	0	8970	160	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 9.

The worst 5 of 160 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:182:SER:O	1:B:190:PHE:HA	1.76	0.85
2:A:2068:THR:HA	2:A:2071:GLU:HB3	1.59	0.84
2:A:1906:TRP:HB3	2:A:1909:LEU:HD11	1.69	0.74
2:A:1879:TYR:OH	2:A:1886:MET:SD	2.49	0.69
2:A:1712:ILE:HG12	2:A:1712:ILE:O	1.96	0.65

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	B	394/689 (57%)	339 (86%)	53 (14%)	2 (0%)	25	63
2	A	754/914 (82%)	548 (73%)	201 (27%)	5 (1%)	19	56
All	All	1148/1603 (72%)	887 (77%)	254 (22%)	7 (1%)	24	59

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	1657	ASN
2	A	1712	ILE
1	B	443	VAL
1	B	535	LEU
2	A	1668	MET

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	B	368/629 (58%)	367 (100%)	1 (0%)	91 92
2	A	627/837 (75%)	618 (99%)	9 (1%)	62 76
All	All	995/1466 (68%)	985 (99%)	10 (1%)	71 81

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	1668	MET
2	A	1794	MET
2	A	2062	MET
2	A	1325	LEU
2	A	1411	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	1742	ASN
2	A	1657	ASN
2	A	1608	ASN
2	A	1373	ASN
2	A	1609	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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	A	1
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	2060:HIS	C	2061:VAL	N	14.34
1	B	280:GLN	C	281:ASN	N	5.02

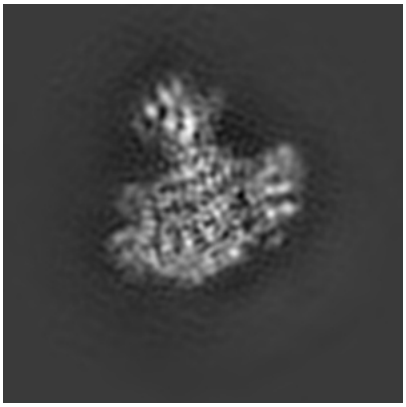
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-0287. These allow visual inspection of the internal detail of the map and identification of artifacts.

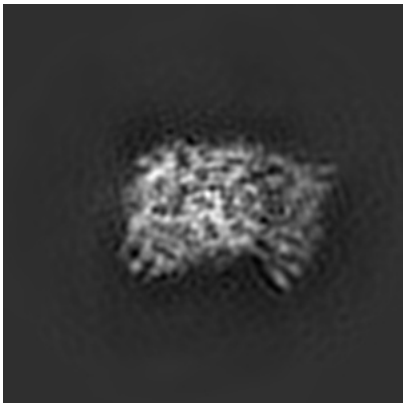
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

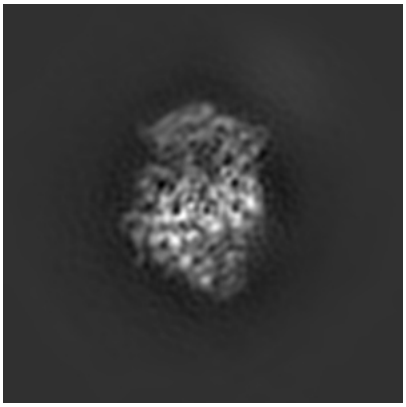
6.1.1 Primary map



X



Y

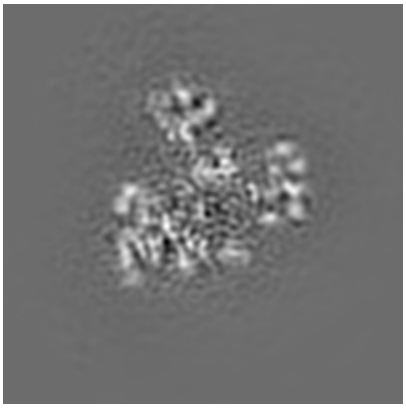


Z

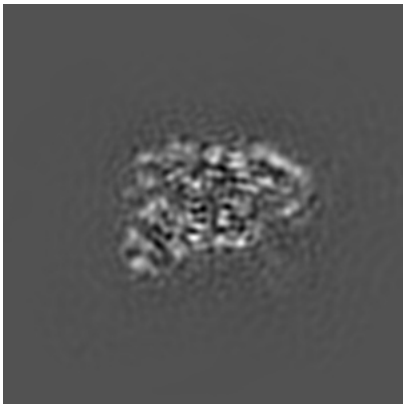
The images above show the map projected in three orthogonal directions.

6.2 Central slices

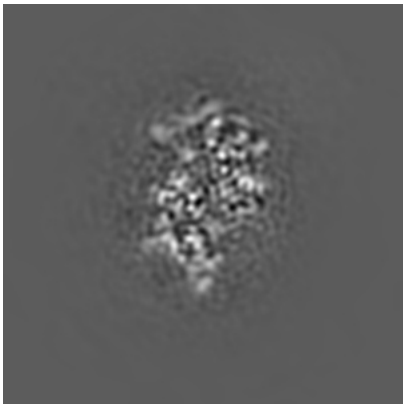
6.2.1 Primary map



X Index: 92



Y Index: 92

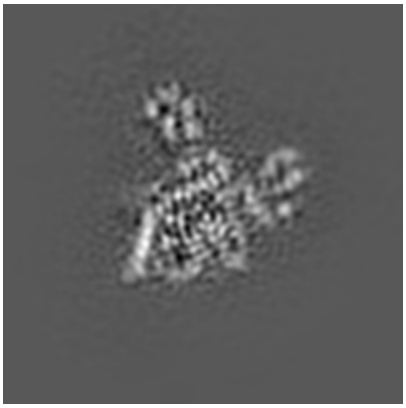


Z Index: 92

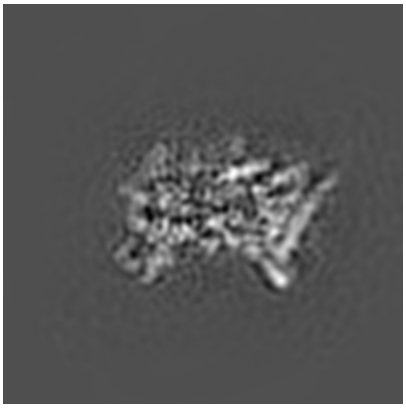
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

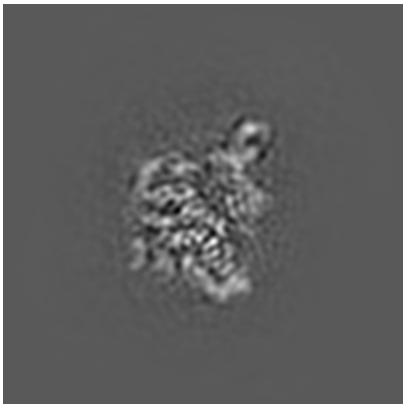
6.3.1 Primary map



X Index: 84



Y Index: 85

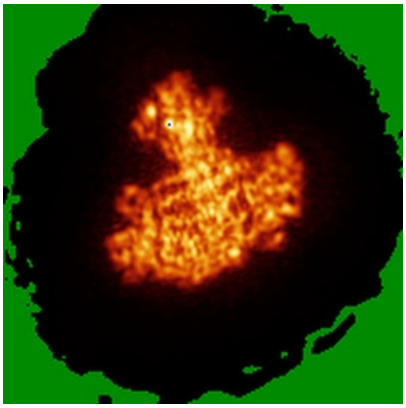


Z Index: 76

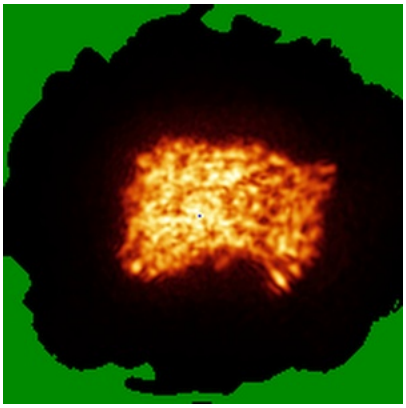
The images above show the largest variance slices of the map in three orthogonal directions.

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

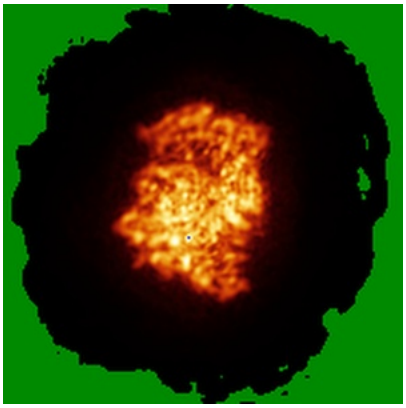
6.4.1 Primary map



X



Y

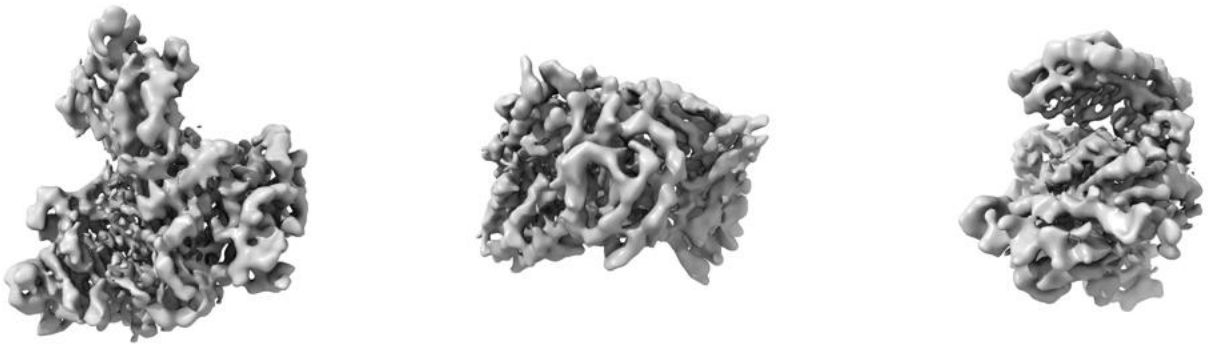


Z

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 0.0576. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

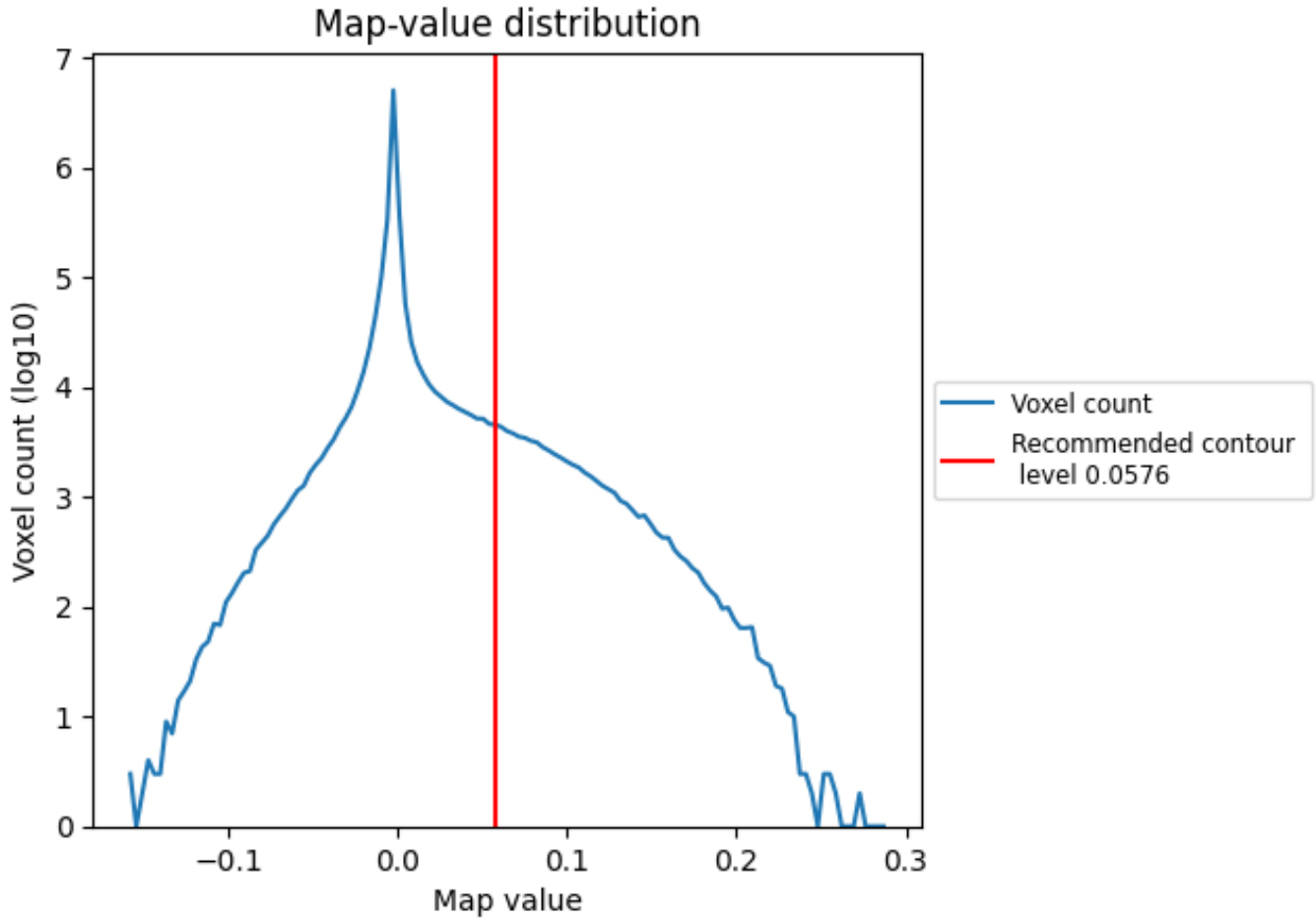
6.6 Mask visualisation (i)

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

7 Map analysis

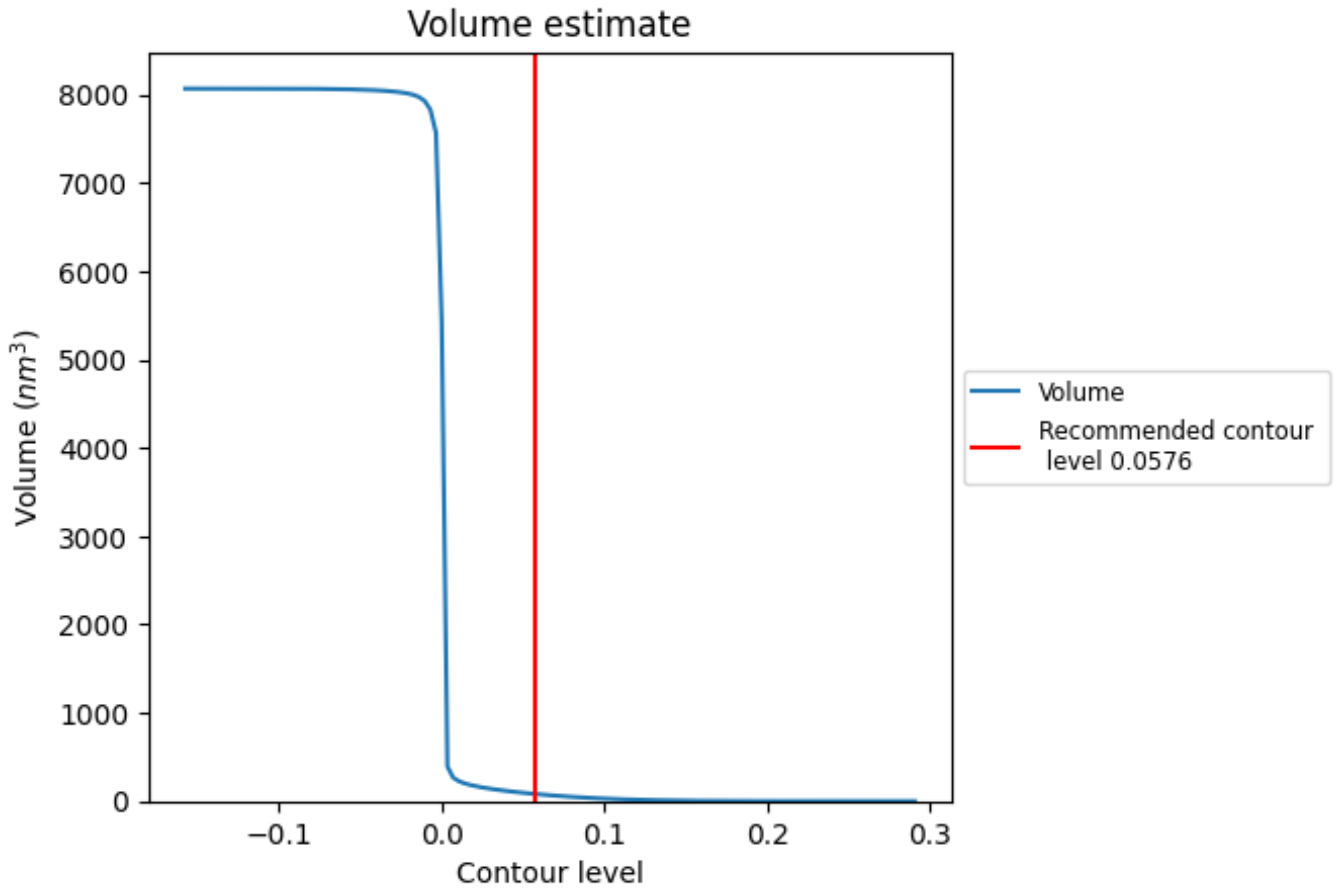
This section contains the results of statistical analysis of the map.

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.

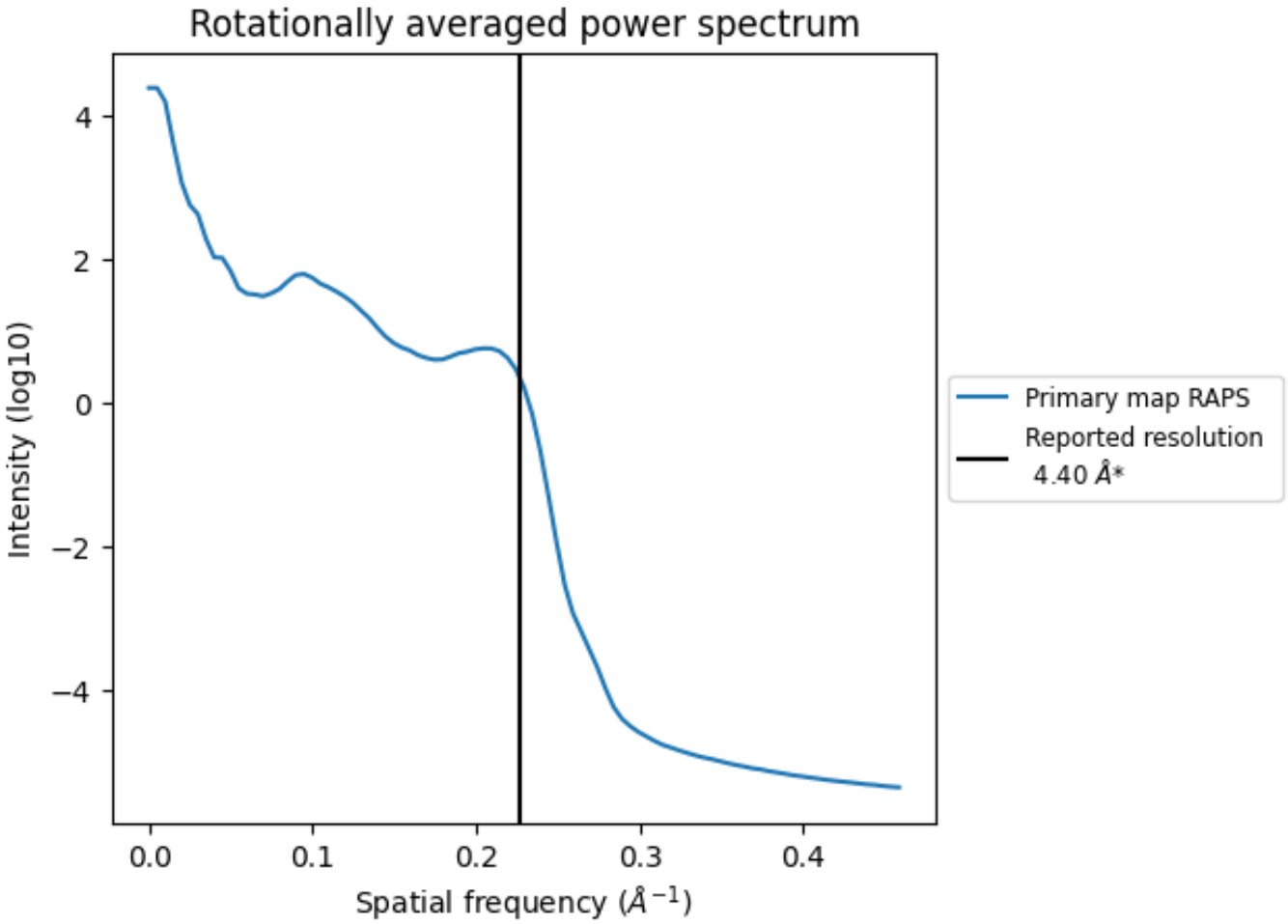
7.2 Volume estimate (i)



The volume at the recommended contour level is 81 nm^3; this corresponds to an approximate mass of 73 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 0.227 Å^{-1}

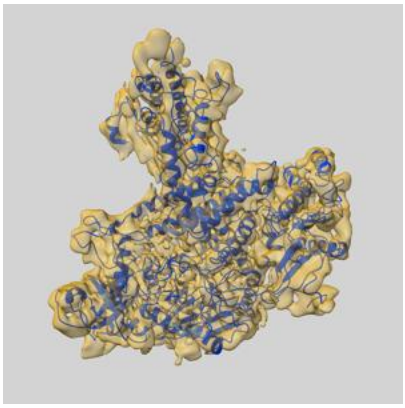
8 Fourier-Shell correlation (i)

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

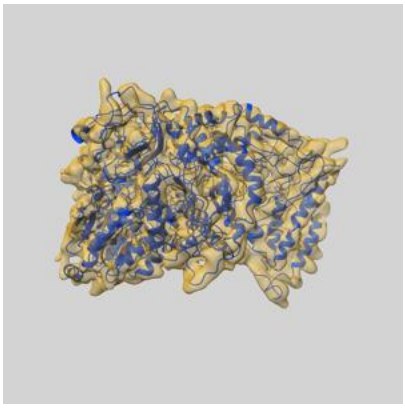
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-0287 and PDB model 6HV8. Per-residue inclusion information can be found in section 3 on page 4.

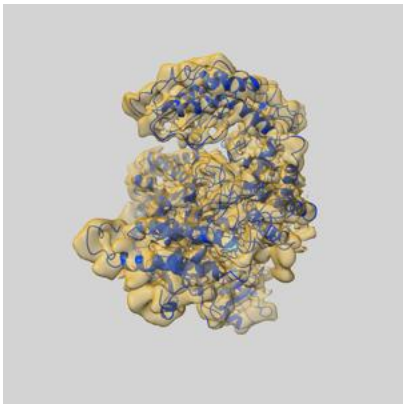
9.1 Map-model overlay



X



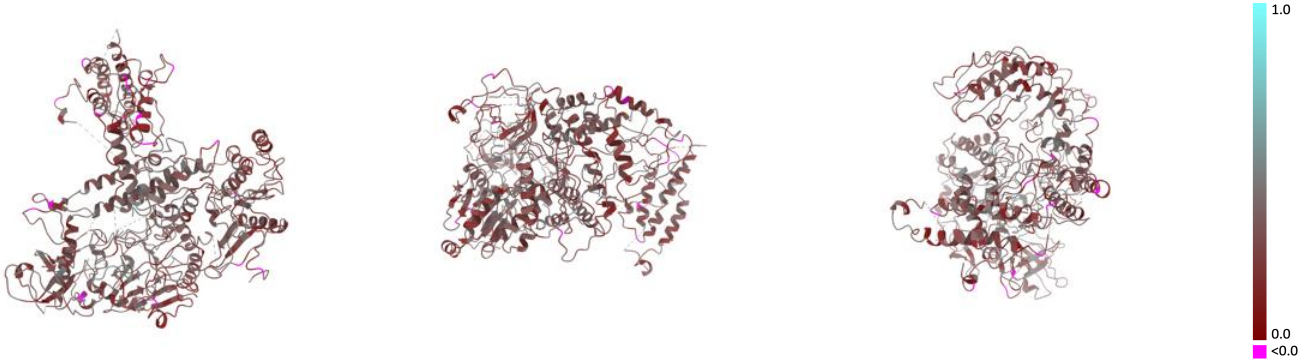
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.0576 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.

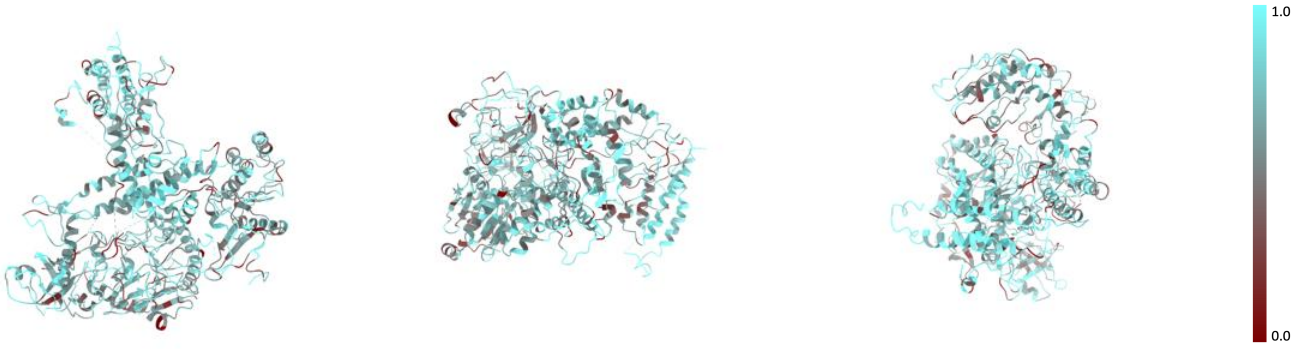
9.2 Q-score mapped to coordinate model



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.

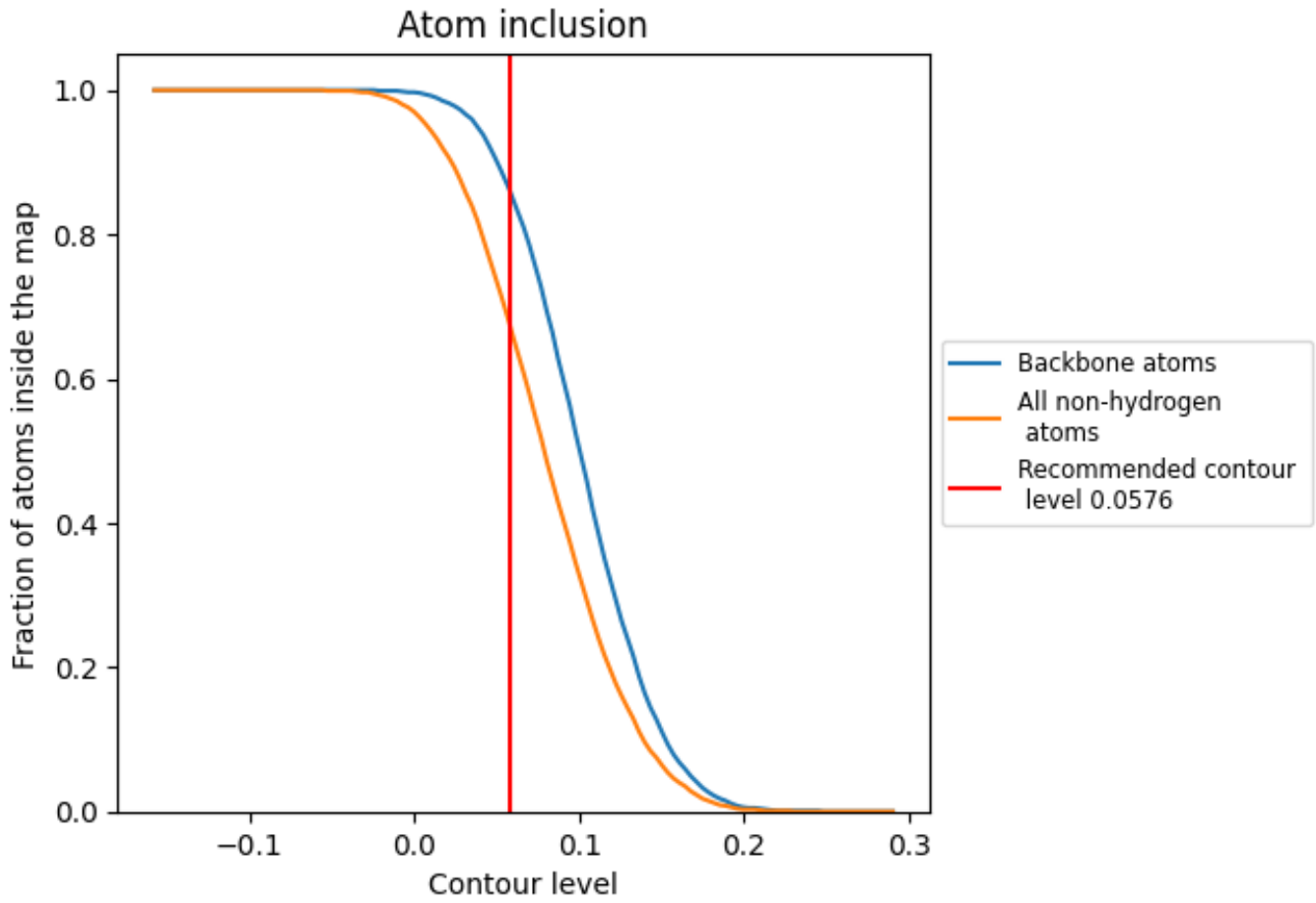
9.3 Atom inclusion mapped to coordinate model



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 (0.0576).

9.4 Atom inclusion (i)



At the recommended contour level, 86% of all backbone atoms, 68% 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 (0.0576) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6780	0.3150
A	0.6750	0.3050
B	0.6840	0.3320