



wwPDB EM Validation Summary Report ⓘ

Nov 13, 2022 – 06:11 AM EST

PDB ID : 6WJV
EMDB ID : EMD-21701
Title : Structure of the *Saccharomyces cerevisiae* polymerase epsilon holoenzyme
Authors : Yuan, Z.; Georgescu, R.; Schauer, G.D.; O'Donnell, M.; Li, H.
Deposited on : 2020-04-14
Resolution : 3.50 Å (reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : 0.0.1.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

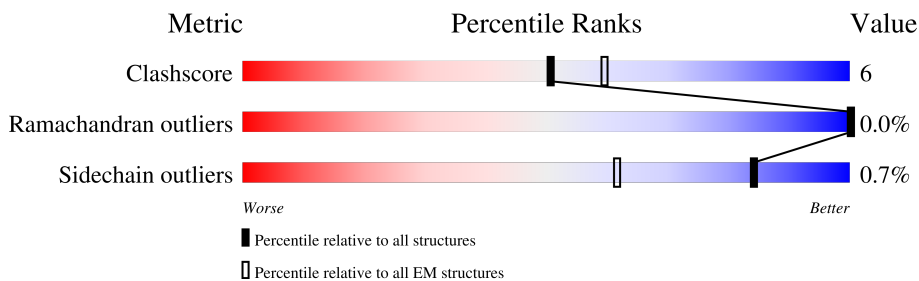
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion $< 40\%$). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	2222	55% (Poor fit) 77% (0 outliers), 10% (1 outlier), 12% (2+ outliers) 50% (Not modelled)
2	2	689	52% (0 outliers), 9% (1 outlier), 38% (2+ outliers) 15% (Poor fit) 58% (Not modelled)
3	3	201	40% (0 outliers), 58% (2+ outliers) 23% (Poor fit) 49% (Not modelled)
4	4	196	49% (0 outliers), 6% (1 outlier), 45% (2+ outliers) 23% (Poor fit) 45% (Not modelled)

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 19593 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 catalytic subunit A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1951	14886	9527	2476	2813	70	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	2	424	3225	2055	545	610	15	0	0

- Molecule 3 is a protein called DNA polymerase epsilon subunit C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	3	85	652	417	107	126	2	0	0

- Molecule 4 is a protein called DNA polymerase epsilon subunit D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	4	107	829	521	146	161	1	0	0

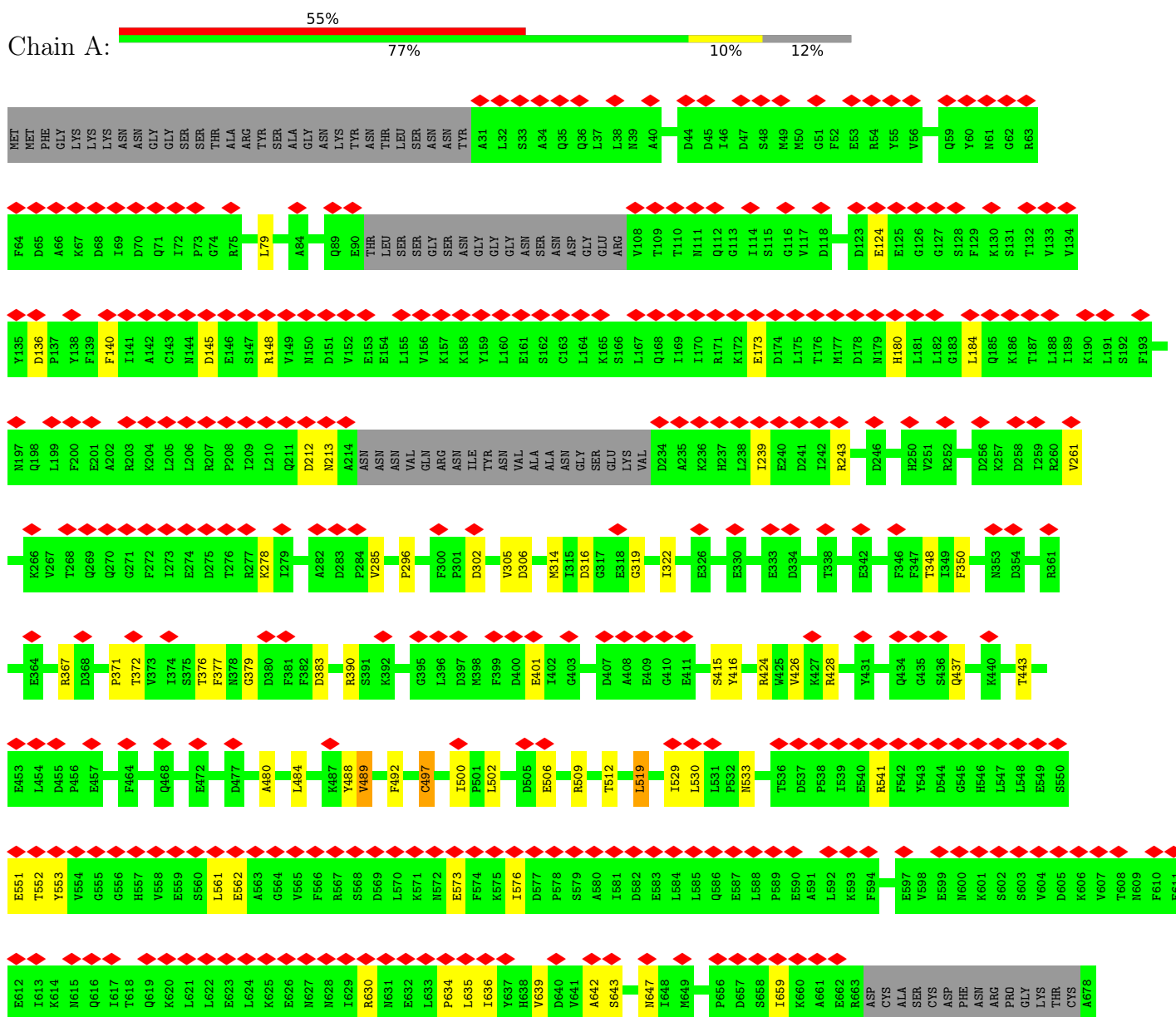
- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
5	A	1	1	1	0

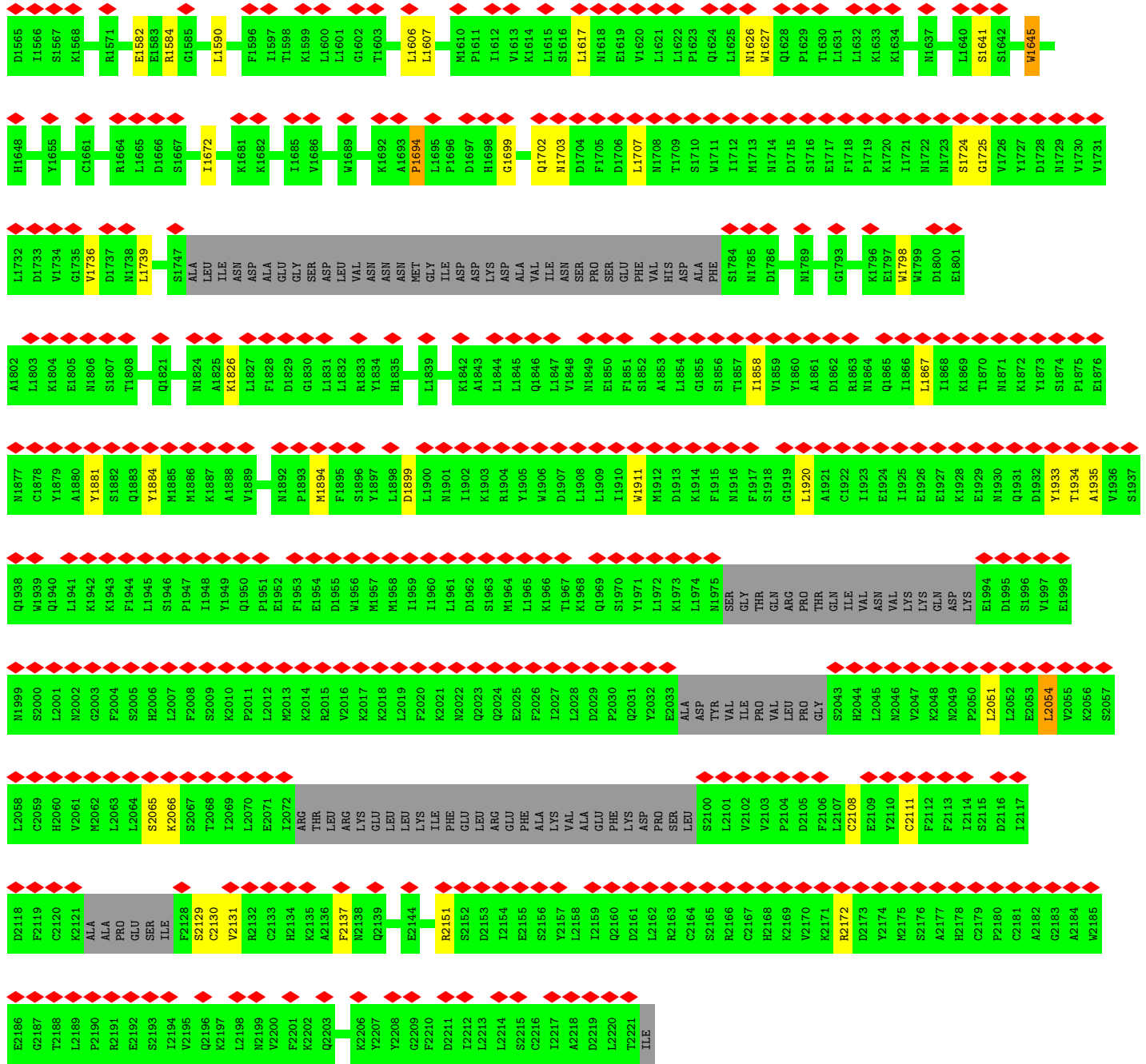
3 Residue-property plots [i](#)

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 catalytic subunit A



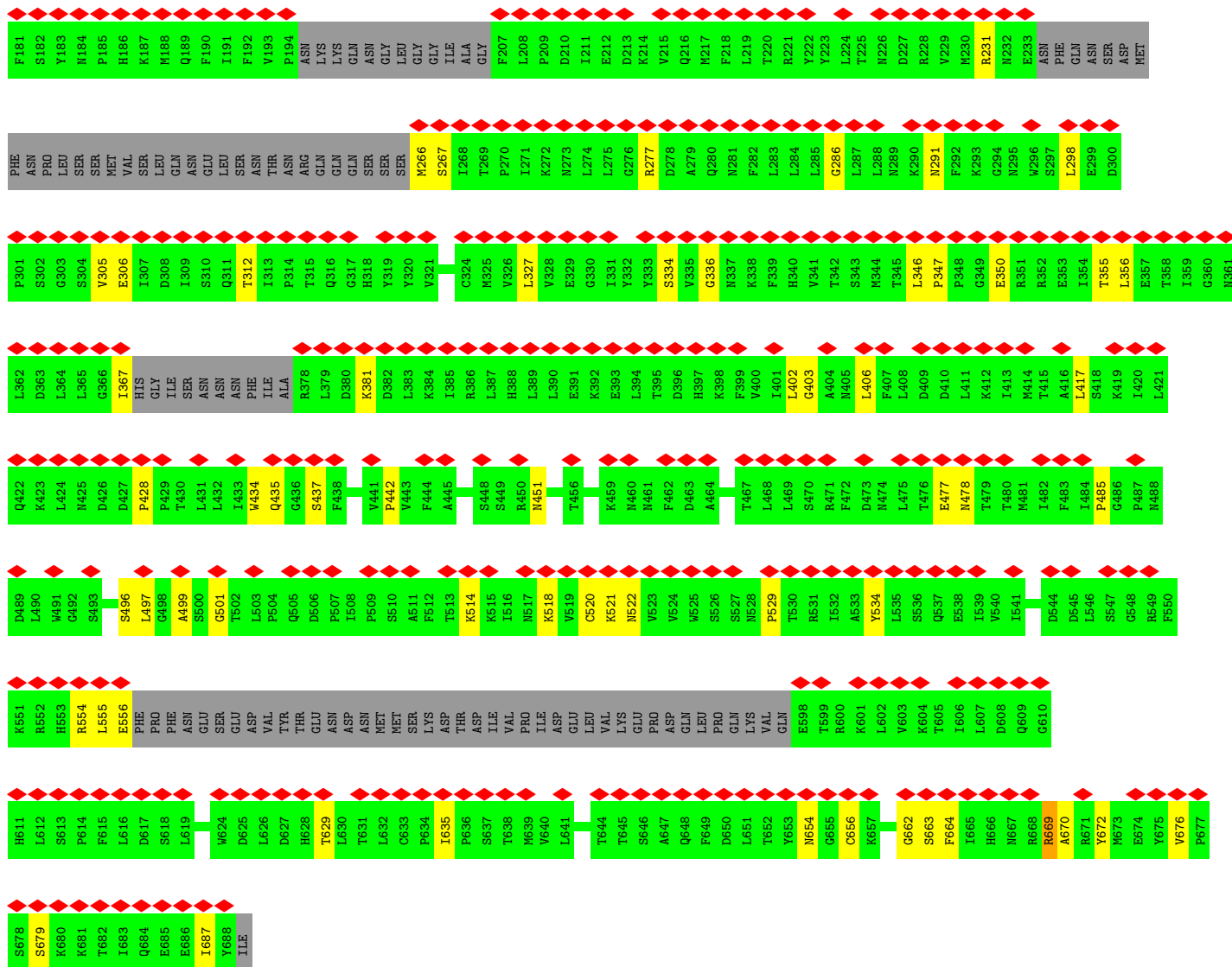
F1498	F1502	K1503	G1506	D1507	L1512	L1513	L1514	L1515	P1516	S1517	M1518	Q1519	A1520	Q1521	E1522	I1523	M1524	A1525	S1526	S1527	L1528	G1529	Q1530	I1531	Y1532	K1533	Q1534	M1535	F1536	E1537	K1538	K1539	K1540	G1541	K1542	I1543	I1544	T1545	Y1546	S1547	Y1548	Y1549	V1550	D1551	I1552	G1553	E1554	D1555	T1556	N1557	F1558	E1559	F1560	F1563	T1564				
D1420	E1421	M1422	V1423	L1424	G1425	V1426	F1427	E1428	T1432	P1433	H1434	Q1435	I1438	L1441	S1444	V1445	T1446	S1449	K1450	A1451	M1452	K1457	Q1460	Q1461	K1466	L1467	L1468	S1469	M1470	A1471	E1472	M1473	E1474	R1475	Y1476	S1481	M1482	D1483	I1484	G1485	Y1486	F1490	T1491	T1492	S1493	I1494	G1495	E1497											
M1359	K1360	F1361	K1362	S1363	Q1364	T1365	M1366	P1367	L1368	Q1369	K1370	I1371	K1372	M1373	C1374	L1375	I1376	E1377	K1378	S1379	S1380	A1381	S1382	L1383	P1384	M1385	M1386	P1387	K1388	T1389	S1390	M1391	P1392	ALA	GLY	GLN	LEU	PHE	LYS	ILE	THR	LEU	PRO	E1404	F1407	E1410	K1411	E1412	M1413	C1414	T1415	S1416	I1417	F1418	M1419				
L1273	M1274	K1277	I1278	M1280	K1281	R1285	D1286	M1287	K1288	R1289	M1290	D1291	Q1292	L1293	F1294	G1295	M1296	T1297	M1298	S1299	S1300	E1301	R1302	R1303	S1304	A1305	L1306	G1307	S1308	M1309	I1310	R1311	K1312	Q1313	A1314	E1315	A1318	M1322	L1325	Q1326	Y1327	K1328	D1329	S1330	E1331	E1332	V1340	M1343	K1344	G1345									
L1073	G1074	D1075	F1076	L1077	D1080	M1081	V1082	K1083	D1084	S1094	S1095	K1096	P1097	F1098	M1099	A1100	P1101	A1106	A1110	I1111	F1112	S1113	A1114	D1115	I1116	P1117	M1118	K1119	R1120	S1121	F1122	L1123	R1124	L1128	D1129	P1130	S1131	L1132	E1133	D1134	L1135	D1136	I1137	R1138	T1139	I1140	D1142	W1143	R1147	E1148	R1149	L1150							
G1151	S1152	A1153	I1154	Q1155	K1156	I1157	T1158	T1159	P1161	A1162	L1164	Q1165	G1166	V1167	S1168	M1169	F1170	V1171	P1172	R1173	W1174	E1175	H1176	P1177	D1178	M1179	L1180	K1181	A1182	K1183	I1184	A1185	T1186	LYS	GLU	ASP	LYS	PHE	LYS	GLN	THR	LEU	LEU	THR	LYS	THR	ASN	VAL	PRO	THR	MET	GLY							
I1002	F1003	K1004	V1005	F1006	L1007	E1008	G1009	D1010	P1011	L1012	E1013	G1014	C1015	Y1016	L1017	A1018	V1019	A1020	S1021	M1024	R1025	W1026	L1027	D1028	V1029	L1030	D1031	S1032	I1033	G1034	L1035	F1036	L1037	E1038	E1039	E1040	D1041	L1045	I1046	C1047	E1048	N1049	R1050	S1051	M1052	S1053	K1054	T1055	L1056	K1057	E1058	Y1059	G1061	R1071	R1072				
N940	T941	I942	F943	E944	E945	V946	D947	G948	P949	Y950	K951	A952	M953	I954	L955	P956	S957	S958	K959	E960	E961	G962	K963	G964	I965	K966	K967	R968	F972	N973	E974	D975	G976	S977	L978	A979	E980	L981	K982	F984	E985	L986	K987	R988	R989	G990	E991	L992	Q993	L994	I995	K996	N997	F998	Q999	S1000	D1001		
R806	D807	E808	A809	K810	K811	M812	I813	D817	Q820	L821	A822	I826	L827	N828	S829	F830	Y831	G832	Y833	V834	M835	R836	K837	G838	S839	Y842	S843	M844	E845	L852	A855	T856	I857	I858	Q859	M860	A861	R862	A863	L864	V865	E866	R867	G868	R870	P871	L872	D873	F874	D875	S876	D877							
G878	I879	W880	C881	I882	L883	P884	K885	S886	F887	E888	E889	T890	Y891	F892	F893	T894	L895	E896	N897	G898	K899	K900	L901	Y902	L903	S904	Y905	P906	C907	S908	N909	L910	N911	Y912	R913	I914	H915	Q916	K917	H921	Q922	Y923	Q924	E925	L926	K927	G928	D928	P929	L930	N931	Y932	L934	D935	Y936	H937	S938	E939	
L739	T740	E741	Y742	S743	R744	K745	V746	Y747	H748	R749	V750	K751	V752	S753	E754	I755	V756	E757	R758	E759	A760	I761	R765	E766	N767	P768	F769	Y770	Y771	D772	K775	S776	D779	T866	I867	I868	G869	M860	A861	R862	A863	L864	V865	E866	R867	G868	R870	P871	L872	D873	F874	D875	S876	D877					
R679	K680	L681	K682	W683	A684	W685	G686	R687	E688	F689	F690	P691	S692	K693	M694	D695	E696	Y697	M698	M699	I700	K701	R702	A703	L704	Q705	N706	E707	T708	F709	P710	N711	K712	N713	K714	F715	S716	K717	K718	K719	V720	L721	T722	F723	D724	E725	L726	S727	Y728	A729	D730	Q731	V732	I733	H734	K735	K736	K737	R738



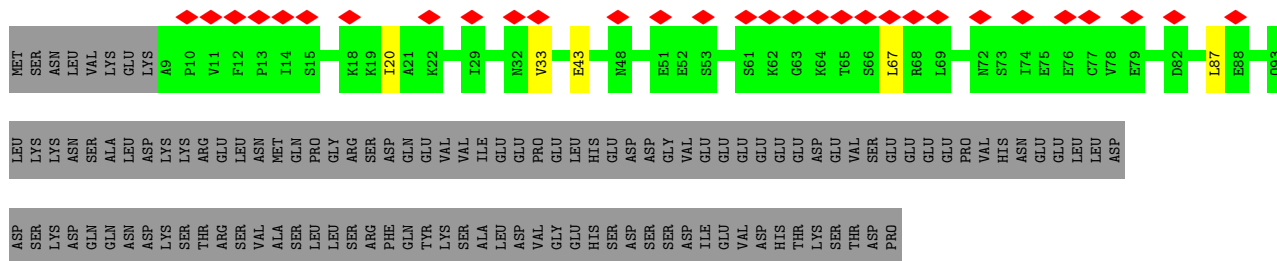
• Molecule 2: DNA polymerase epsilon subunit B



MET	LYS	ASN
PHE	PHE	SER
GLY	LEU	ASP
SER	GLU	ASP
GLY	GLN	GLU
ASN	PHE	MET
VAL	ALA	PRO
LEU	ALA	ILE
PRO	VAL	ALA
VAL	TRP	ASP
LYS	VAL	SER
ILE	GLN	SER
GLN	GLU	LEU
PRO	ARG	LEU
PRO	ARG	GLN
LEU	GLY	ASN
LEU	GLY	VAL
LEU	PHE	ARG
ARG	PHE	PRO
PRO	ILE	PRO
LEU	ASP	SER
ALA	GLN	SER
TYR	SER	PRO
ARG	GLY	GLN
VAL	VAL	ARG
LEU	GLU	PRO
SER	VAL	THR
ARG	VAL	GLU
LYS	ILE	GLU
ILE	GLN	ARG
TYR	GLU	ASP
GLY	GLU	GLY
LEU	GLY	GLU
LEU	ILE	LYS
ILE	GLN	LYS
PHE	GLU	GLN
GLU	GLU	ARG
LEU	GLY	PRO
LEU	VAL	PHE
LEU	GLY	ASP
LEU	GLU	ASP
ALA	GLU	ASP
THR	GLU	TRP
THR	LEU	SER
ASN	HIS	SER
ASN	GLU	LYS
ILE	ILE	ALA
GLY	ILE	LEU
ALA	ALA	ALA
ASN	GLU	ASP
GLY	GLY	TYR
ILE	ILE	F170
ALA	ALA	K171
ASP	ASP	V172
ASP	ASP	I173
THR	THR	M174
GLY	GLY	A175
PRO	PRO	S176
ALA	ALA	Q177
THR	THR	Q178
ILE	ILE	Q179
ILE	ILE	R180

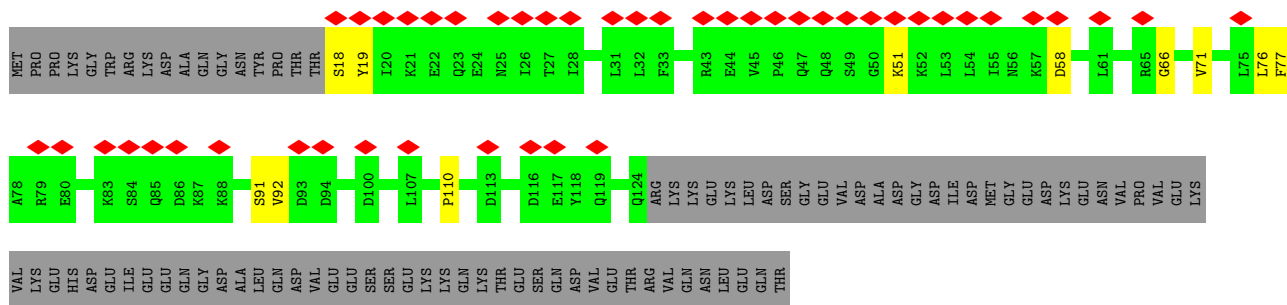


• Molecule 3: DNA polymerase epsilon subunit C



• Molecule 4: DNA polymerase epsilon subunit D





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	187298	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 ($e^-/\text{\AA}^2$)	60	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	0.223	Depositor
Minimum map value	-0.112	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.06	Depositor
Map size (Å)	263.424, 263.424, 263.424	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.029, 1.029, 1.029	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	A	0.47	4/15201 (0.0%)	0.73	15/20663 (0.1%)
2	2	0.44	0/3286	0.80	7/4467 (0.2%)
3	3	0.42	0/660	0.60	0/891
4	4	0.54	1/839 (0.1%)	0.74	0/1132
All	All	0.47	5/19986 (0.0%)	0.74	22/27153 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1161	PRO	N-CD	-8.85	1.35	1.47
4	4	110	PRO	N-CD	-7.39	1.37	1.47
1	A	1645	TRP	CB-CG	-6.58	1.38	1.50
1	A	497	CYS	CB-SG	-5.92	1.72	1.81
1	A	489	VAL	CB-CG1	-5.36	1.41	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	1884	TYR	CA-CB-CG	8.09	128.78	113.40
2	2	403	GLY	N-CA-C	7.99	133.07	113.10
1	A	1707	LEU	CA-CB-CG	7.95	133.59	115.30
1	A	1933	TYR	CA-CB-CG	7.02	126.75	113.40
2	2	662	GLY	N-CA-C	6.39	129.09	113.10

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	14886	0	13968	179	0
2	2	3225	0	3134	43	0
3	3	652	0	664	4	0
4	4	829	0	842	10	0
5	A	1	0	0	0	0
All	All	19593	0	18608	226	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 226 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:305:VAL:HG13	1:A:1301:ARG:NH1	1.48	1.26
1:A:1115:ASP:OD1	1:A:1118:ILE:CG1	1.99	1.08
1:A:1540:LYS:HA	1:A:1543:ILE:HG22	1.32	1.05
1:A:306:ASP:OD1	1:A:1301:ARG:NH1	1.93	1.00
1:A:1115:ASP:OD1	1:A:1118:ILE:HG12	1.60	0.98

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	1929/2222 (87%)	1802 (93%)	127 (7%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	2	414/689 (60%)	370 (89%)	43 (10%)	1 (0%)	47	81
3	3	83/201 (41%)	80 (96%)	3 (4%)	0	100	100
4	4	105/196 (54%)	96 (91%)	9 (9%)	0	100	100
All	All	2531/3308 (76%)	2348 (93%)	182 (7%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	2	434	TRP

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	1522/2014 (76%)	1510 (99%)	12 (1%)	81	91
2	2	351/629 (56%)	349 (99%)	2 (1%)	86	94
3	3	72/185 (39%)	71 (99%)	1 (1%)	67	85
4	4	91/173 (53%)	91 (100%)	0	100	100
All	All	2036/3001 (68%)	2021 (99%)	15 (1%)	84	93

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1047	CYS
2	2	669	ARG
1	A	1081	MET
3	3	43	GLU
1	A	2111	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	2	537	GLN
2	2	517	ASN
1	A	1461	GLN
2	2	174	ASN
1	A	1176	HIS

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](#)

Of 1 ligands modelled in this entry, 1 is 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](#)

There are no chain breaks in this entry.

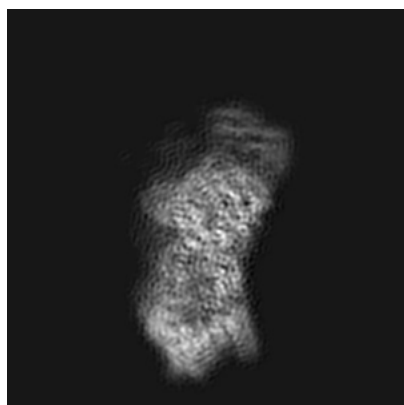
6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-21701. These allow visual inspection of the internal detail of the map and identification of artifacts.

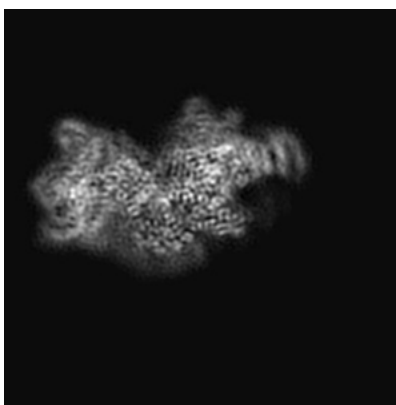
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections [i](#)

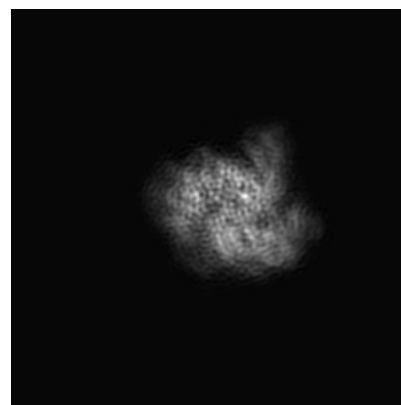
6.1.1 Primary map



X



Y



Z

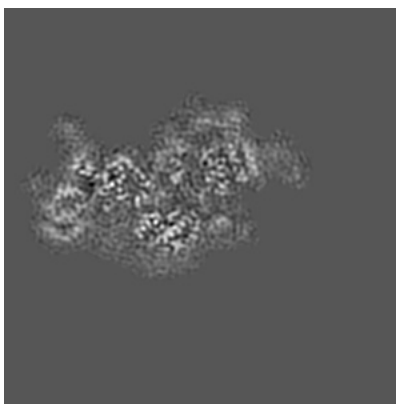
The images above show the map projected in three orthogonal directions.

6.2 Central slices [i](#)

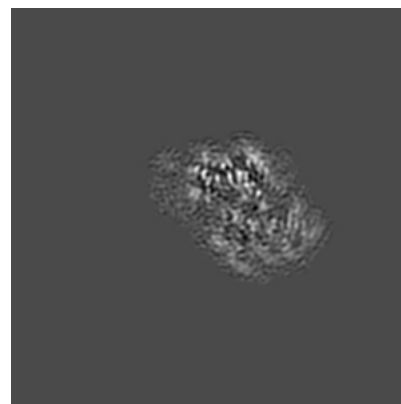
6.2.1 Primary map



X Index: 128



Y Index: 128

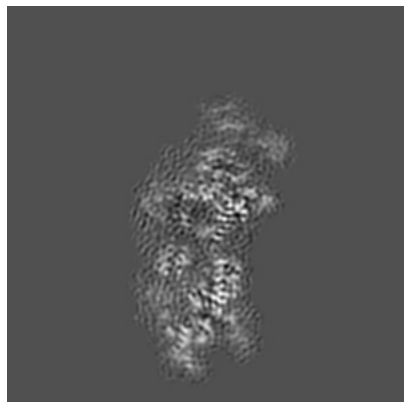


Z Index: 128

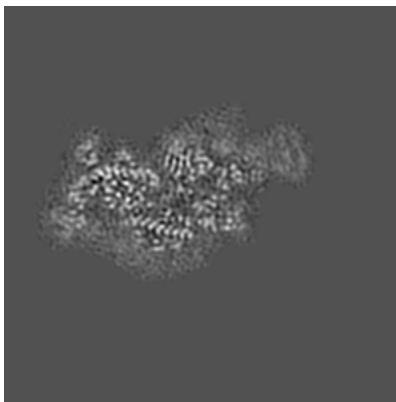
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices [i](#)

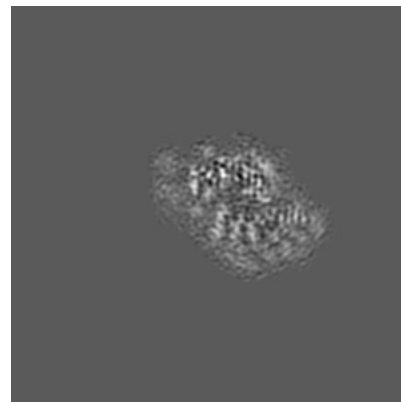
6.3.1 Primary map



X Index: 150



Y Index: 136



Z Index: 125

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

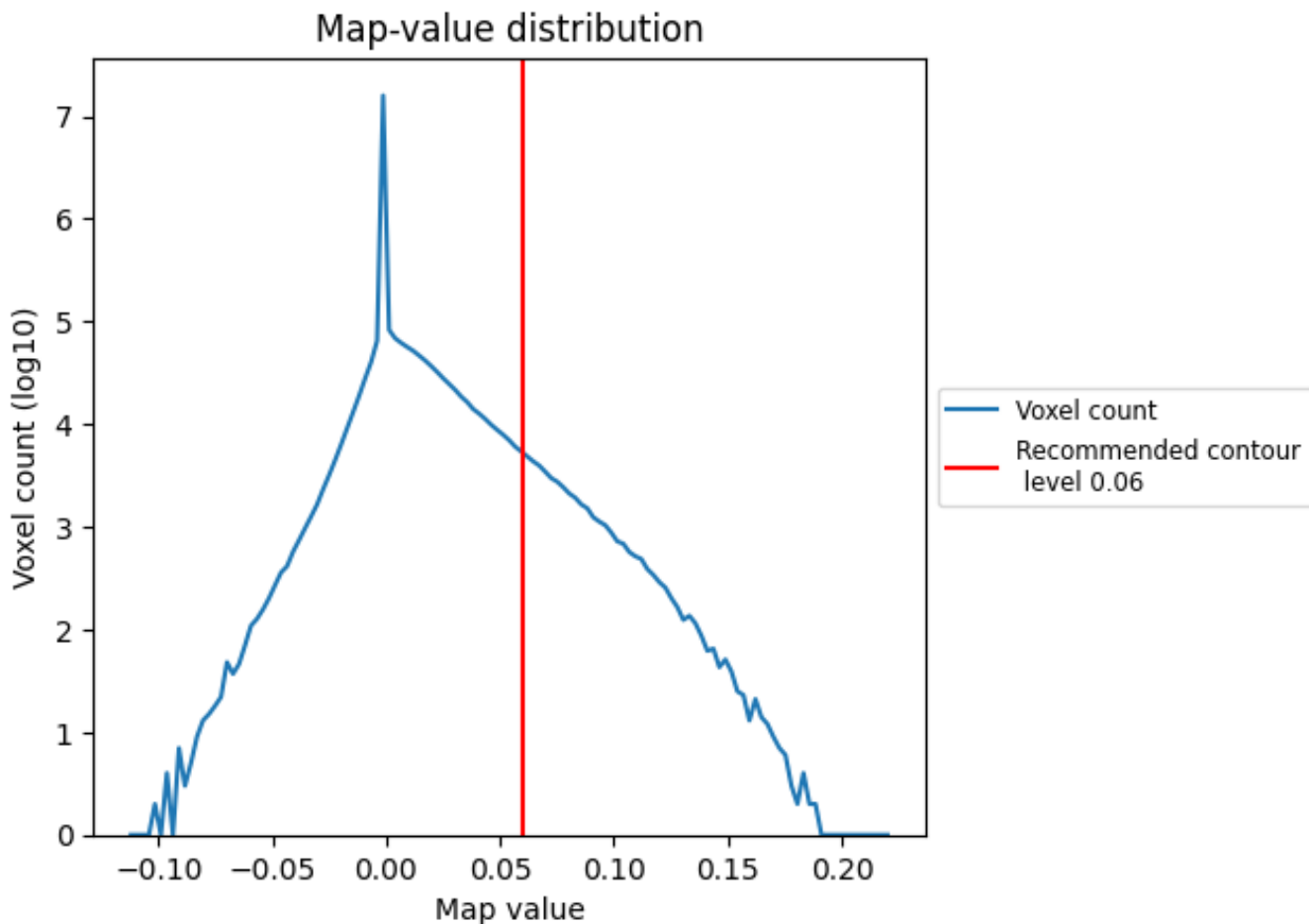
6.5 Mask visualisation

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

7 Map analysis [i](#)

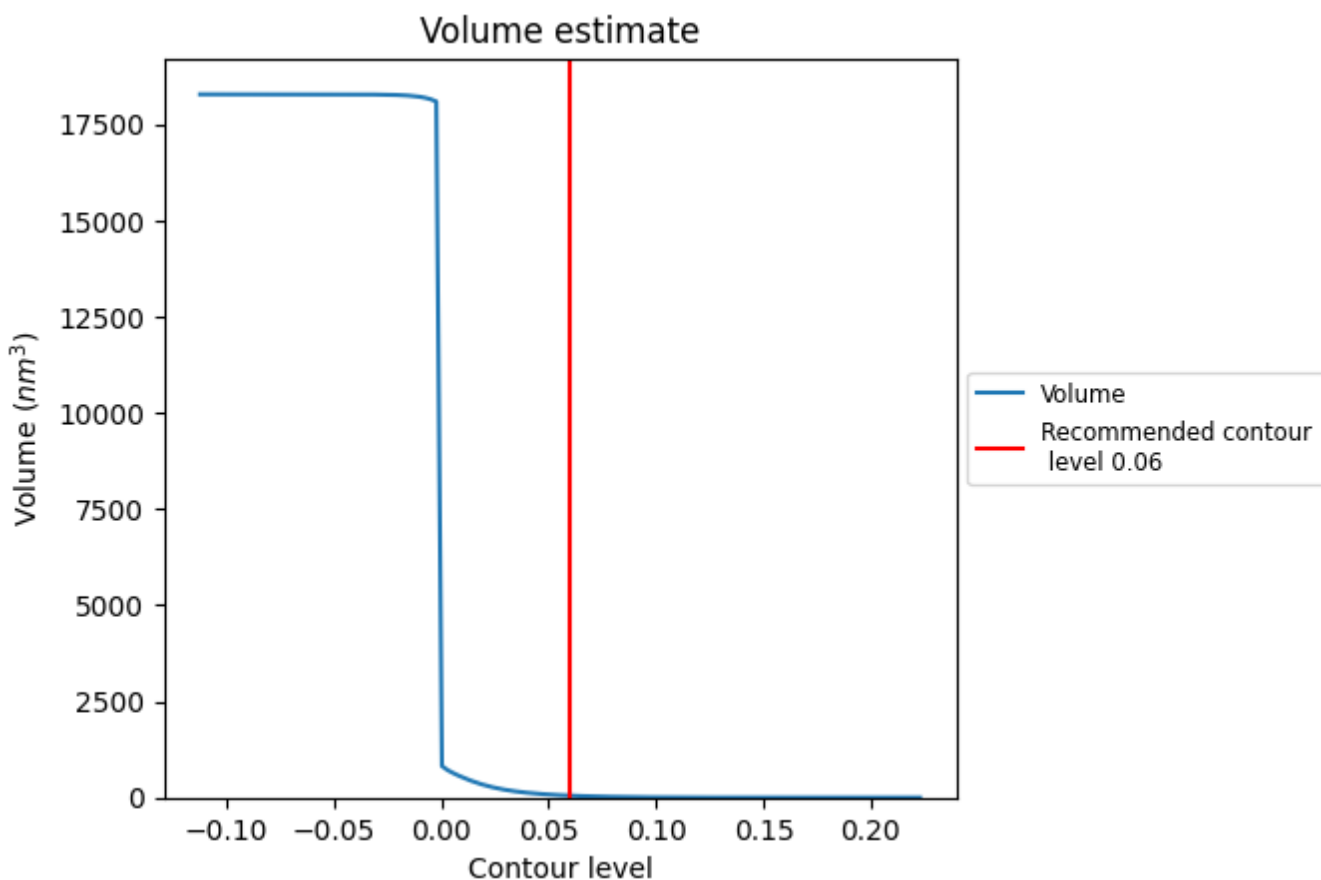
This section contains the results of statistical analysis of the map.

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.

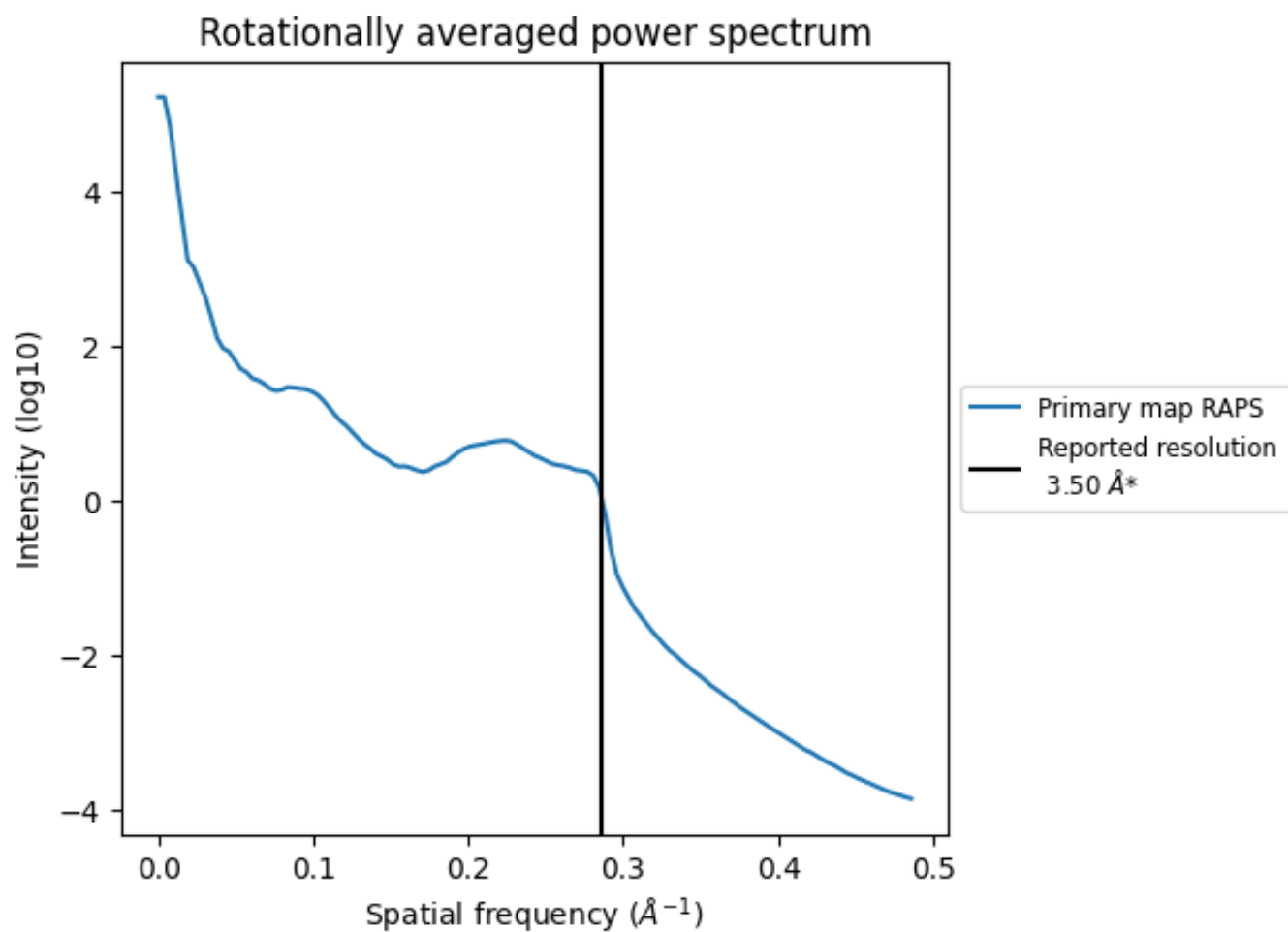
7.2 Volume estimate [\(i\)](#)



The volume at the recommended contour level is 50 nm³; this corresponds to an approximate mass of 45 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 0.286\AA^{-1}

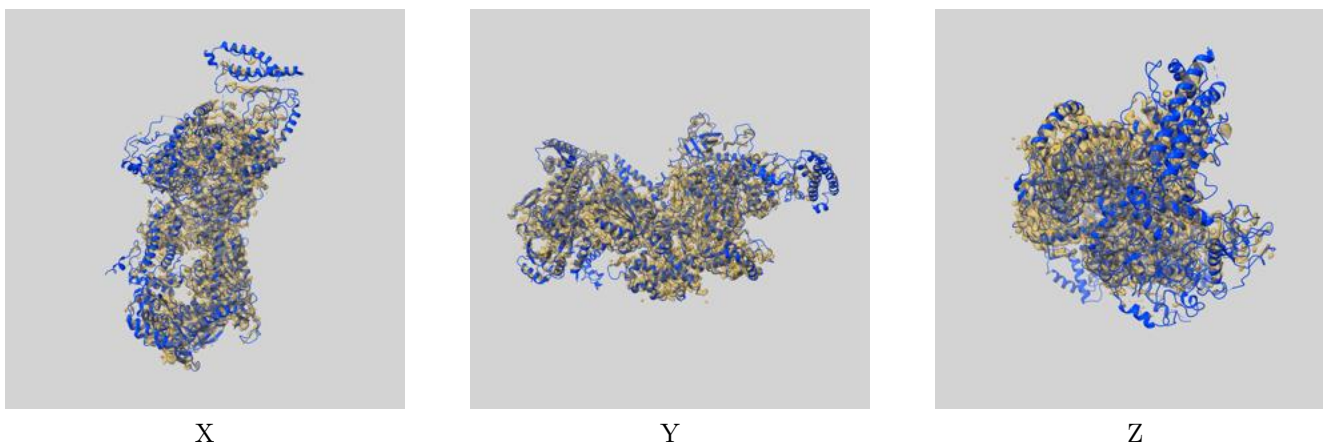
8 Fourier-Shell correlation

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

9 Map-model fit [i](#)

This section contains information regarding the fit between EMDB map EMD-21701 and PDB model 6WJV. Per-residue inclusion information can be found in section [3](#) on page [4](#).

9.1 Map-model overlay [i](#)



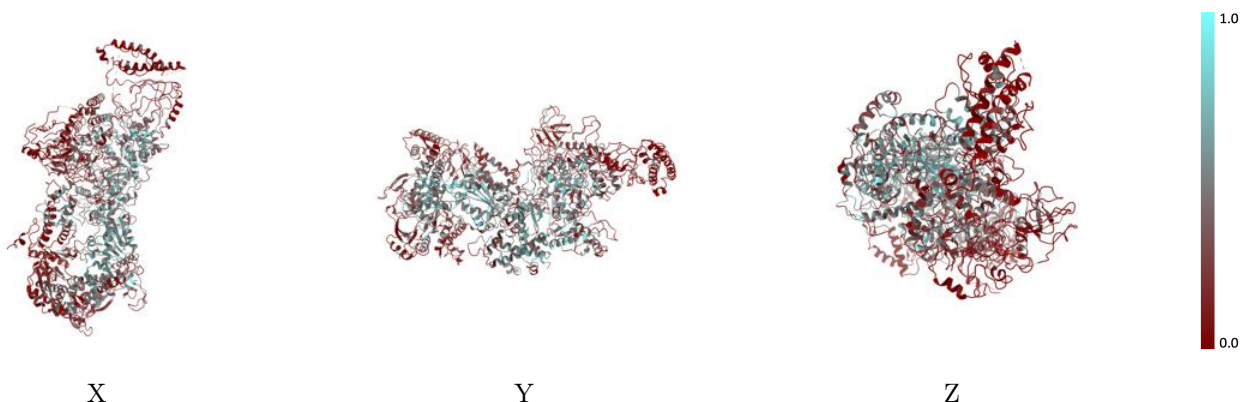
The images above show the 3D surface view of the map at the recommended contour level 0.06 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

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



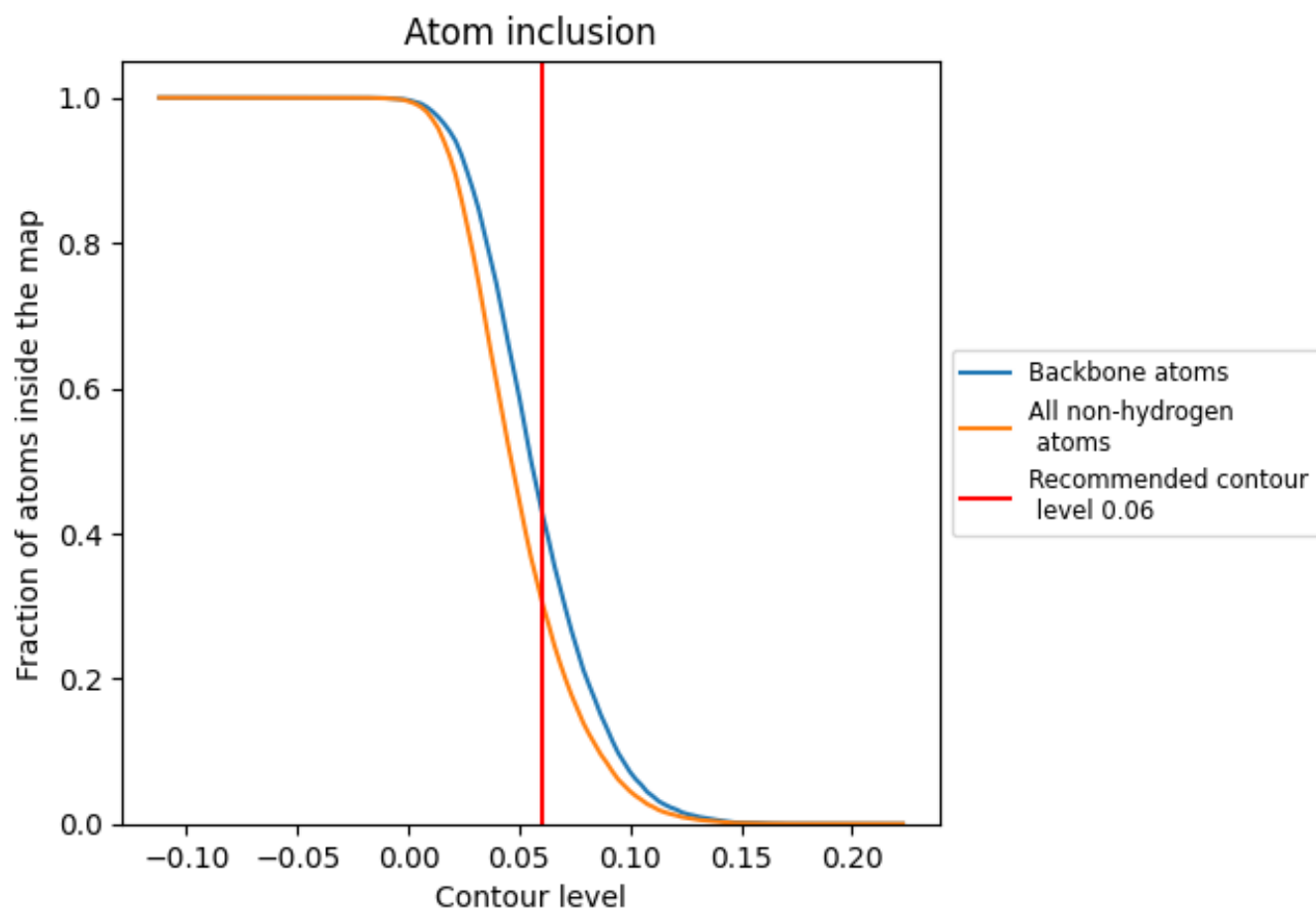
The images above show the model with each residue coloured according to its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 (0.06).











9.4 Atom inclusion [i](#)



At the recommended contour level, 43% of all backbone atoms, 31% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary [i](#)

The table lists the average atom inclusion at the recommended contour level (0.06) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.3085	 0.3490
2	 0.2000	 0.3340
3	 0.4405	 0.3700
4	 0.4347	 0.3910
A	 0.3192	 0.3490

