wwPDB EM Validation Summary Report

Aug 5, 2025 – 01:28 pm BST

PDB ID	:	9GD7 / pdb_00009gd7
EMDB ID	:	EMD-51249
Title	:	DNA-PK Ku80 mediated dimer bound to DNA polymerase Lambda and DNA ligase 4/XRCC4
Authors	:	Chaplin, A.K.; Amin, H.; Zahid, S.; Hardwick, S.W.
Deposited on	:	2024-08-05
Resolution	:	4.25 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:

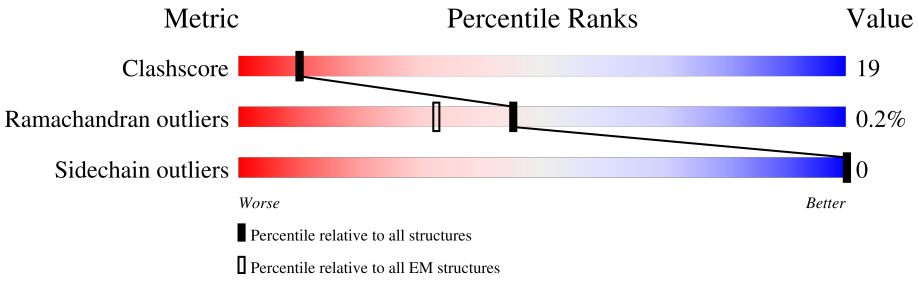
EMDB validation analysis	:	0.0.1.dev126
MolProbity	:	4-5-2 with Phenix2.0rc1
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.45.1

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 4.25 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.



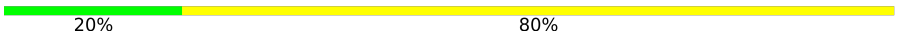
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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	E	911	17% 16% 10% 73%
2	L	732	19% 53% 35% 12%
3	M	204	8% 89%
4	P	336	16% 9% 8% 83%
4	Q	336	15% 10% 7% 83%
5	S	4128	50% 35% 14%
6	T	609	48% 35% 18%
7	i	25	40% 60%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	j	25	
9	A	575	

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 41717 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 ligase 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	E	246	Total	C	N	O	S	0	0
			1956	1240	332	371	13		

- Molecule 2 is a protein called X-ray repair cross-complementing protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	L	641	Total	C	N	O	S	0	0
			5084	3250	859	950	25		

- Molecule 3 is a protein called Protein PAXX.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	M	23	Total	C	N	O	S	0	0
			160	102	27	30	1		

- Molecule 4 is a protein called DNA repair protein XRCC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	P	57	Total	C	N	O	S	0	0
			480	303	90	86	1		
4	Q	57	Total	C	N	O	S	0	0
			475	301	89	84	1		

- Molecule 5 is a protein called DNA-dependent protein kinase catalytic subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	S	3535	Total	C	N	O	S	0	0
			27773	17865	4705	5020	183		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	T	501	Total	C	N	O	S	0	0
			4001	2561	679	743	18		

- Molecule 7 is a DNA chain called DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	i	25	Total	C	N	O	P	0	0
			510	247	92	146	25		

- Molecule 8 is a DNA chain called DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	j	25	Total	C	N	O	P	0	0
			514	250	86	154	24		

- Molecule 9 is a protein called DNA polymerase lambda.

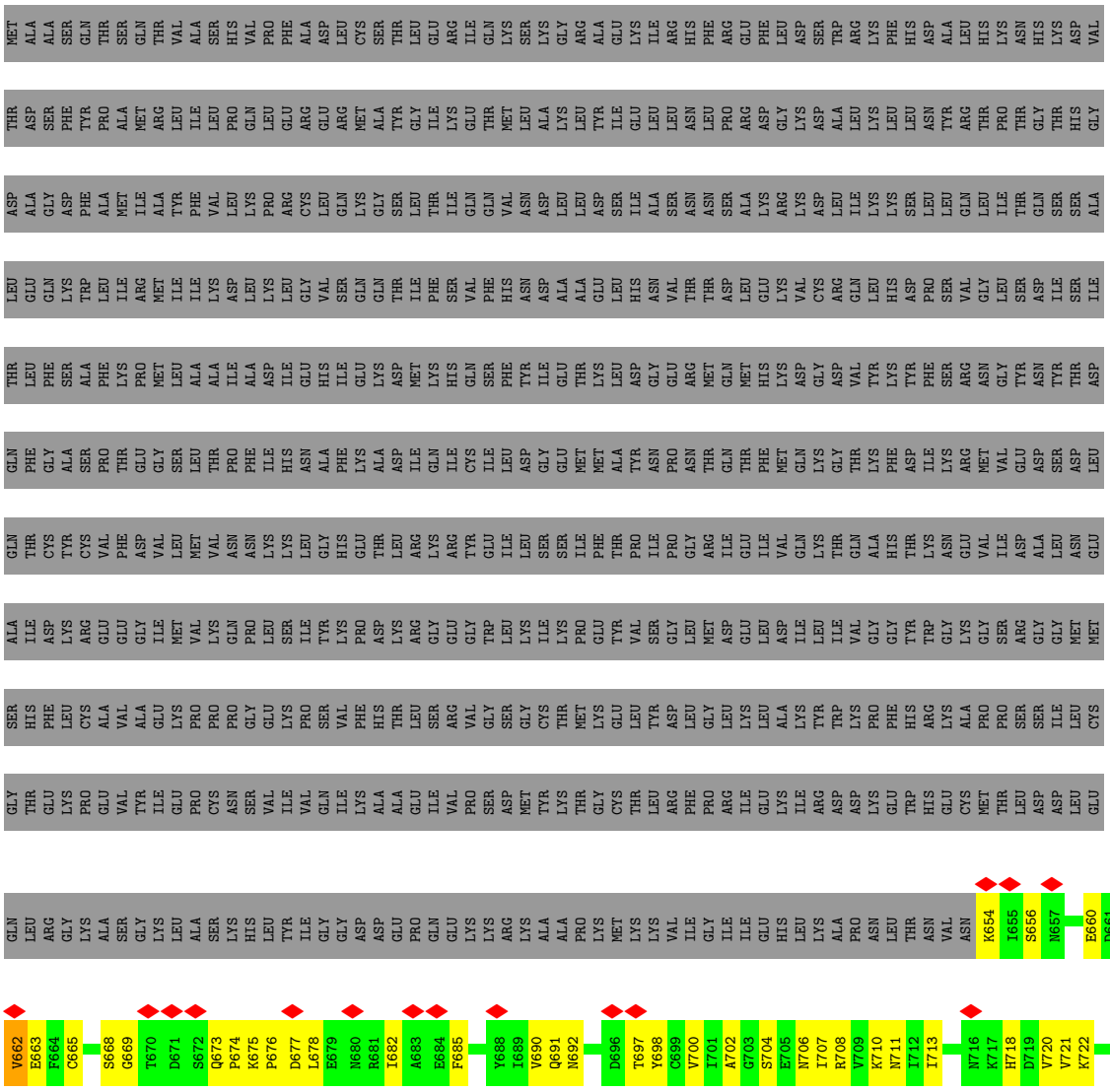
Mol	Chain	Residues	Atoms					AltConf	Trace
9	A	98	Total	C	N	O	S	0	0
			764	487	142	132	3		

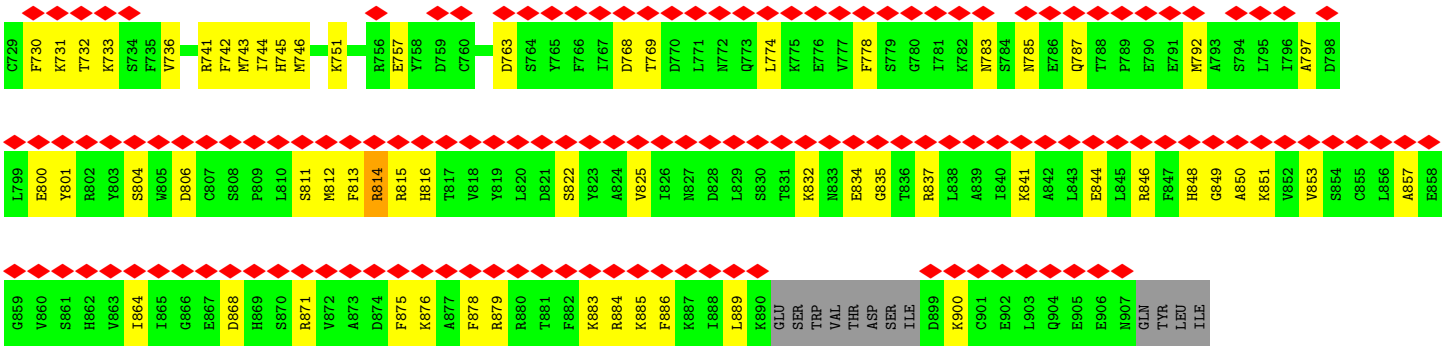
3 Residue-property plots (i)

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 ligase 4

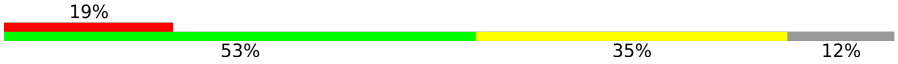
Chain E:

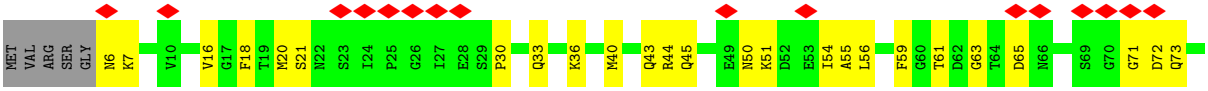


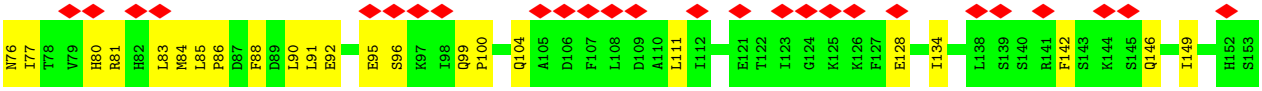


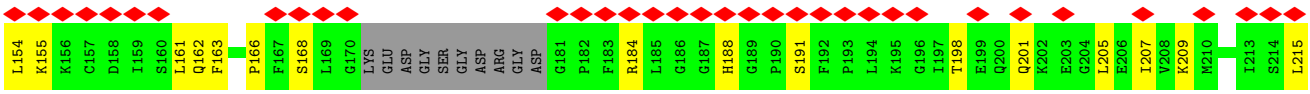
• Molecule 2: X-ray repair cross-complementing protein 5

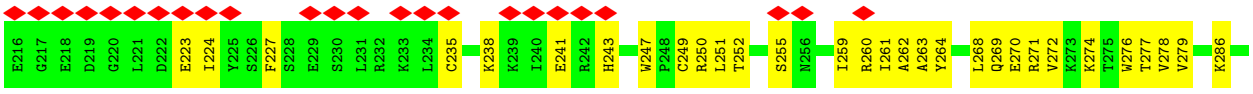
Chain L:

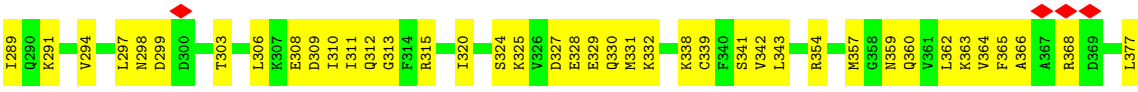


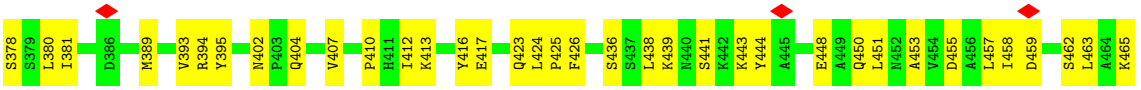


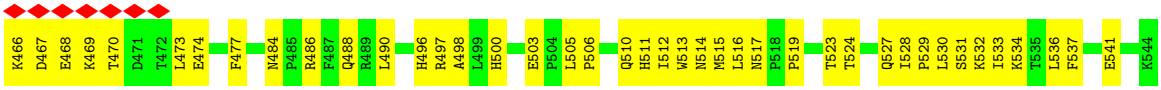


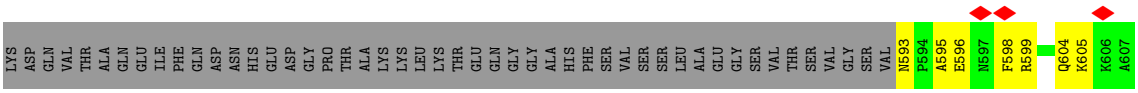


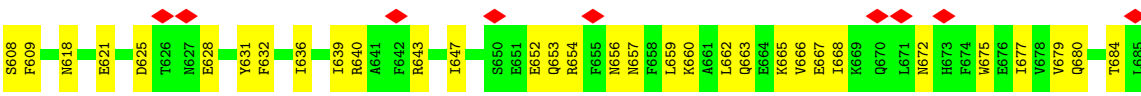


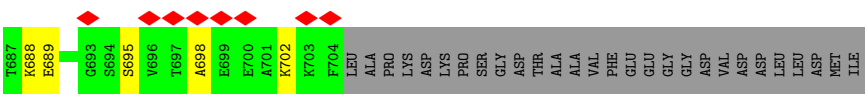






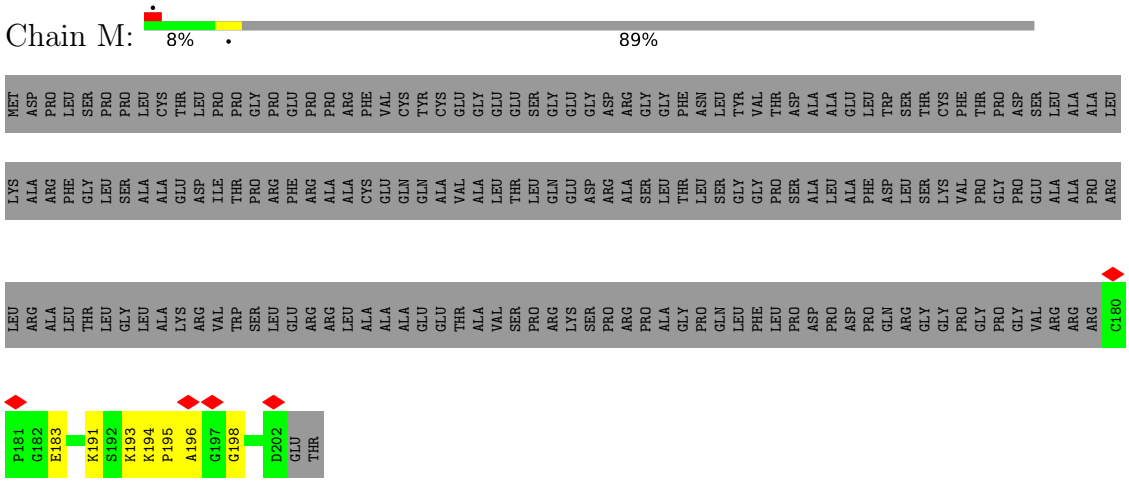






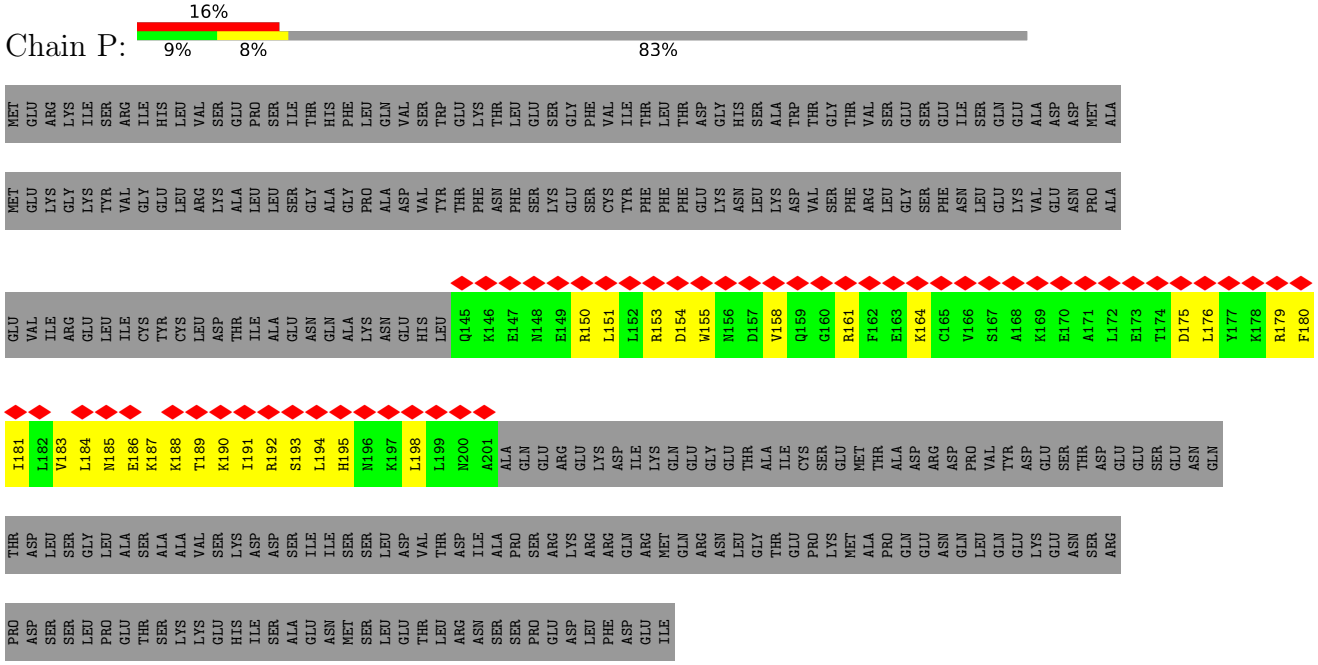
• Molecule 3: Protein PAXX

Chain M:



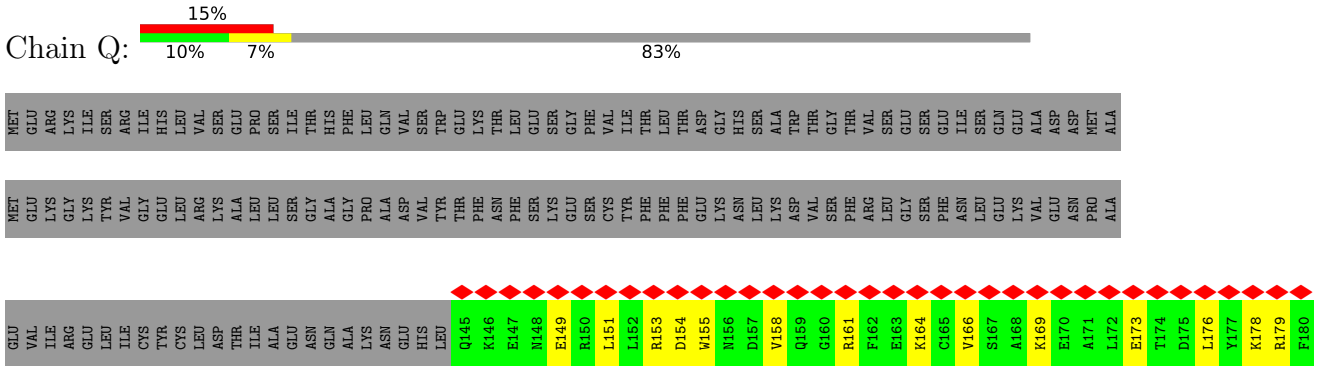
• Molecule 4: DNA repair protein XRCC4

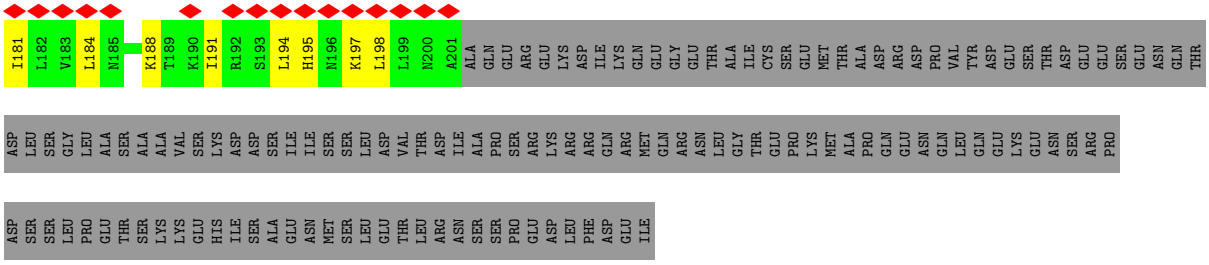
Chain P:



• Molecule 4: DNA repair protein XRCC4

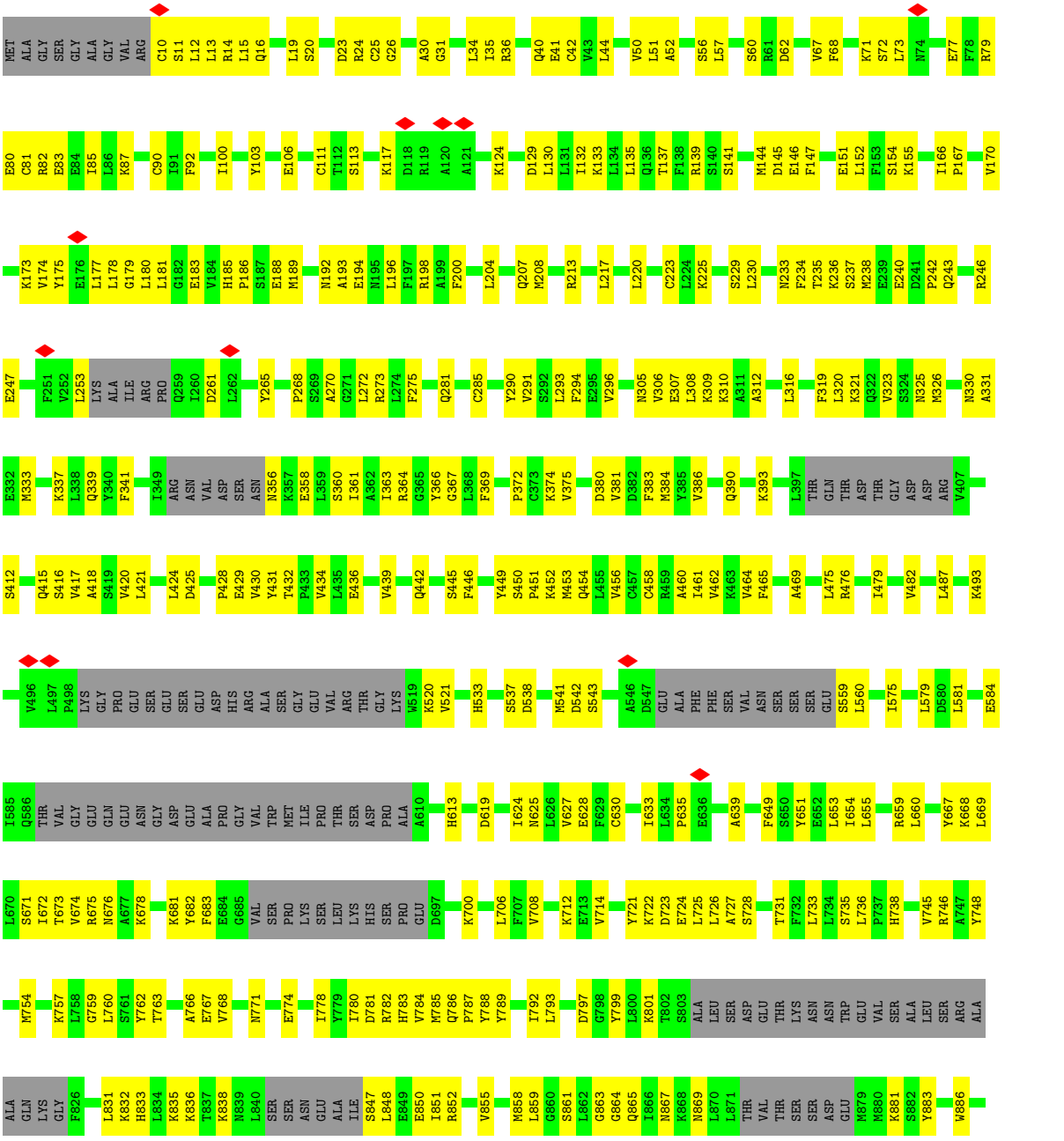
Chain Q:

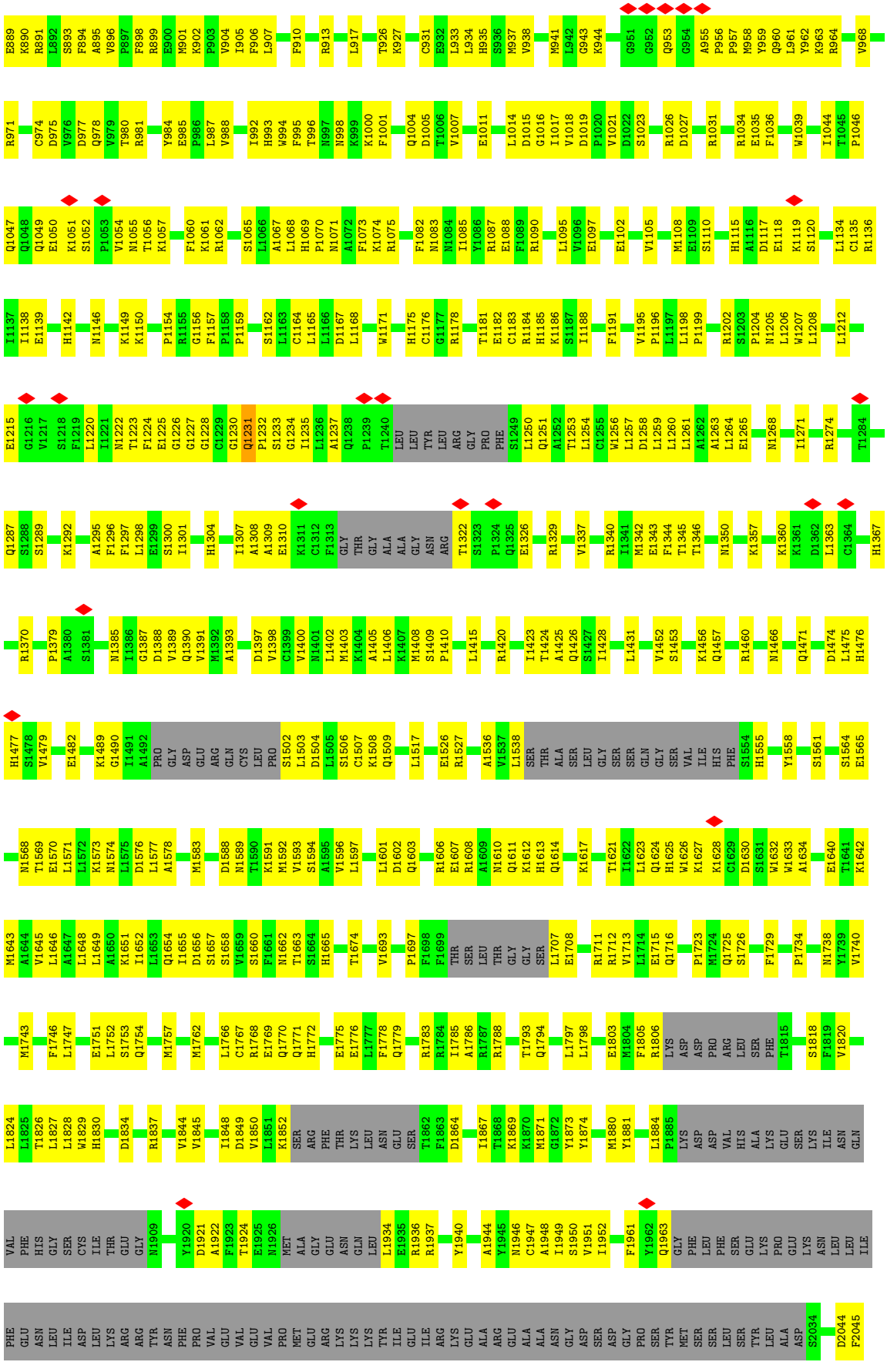


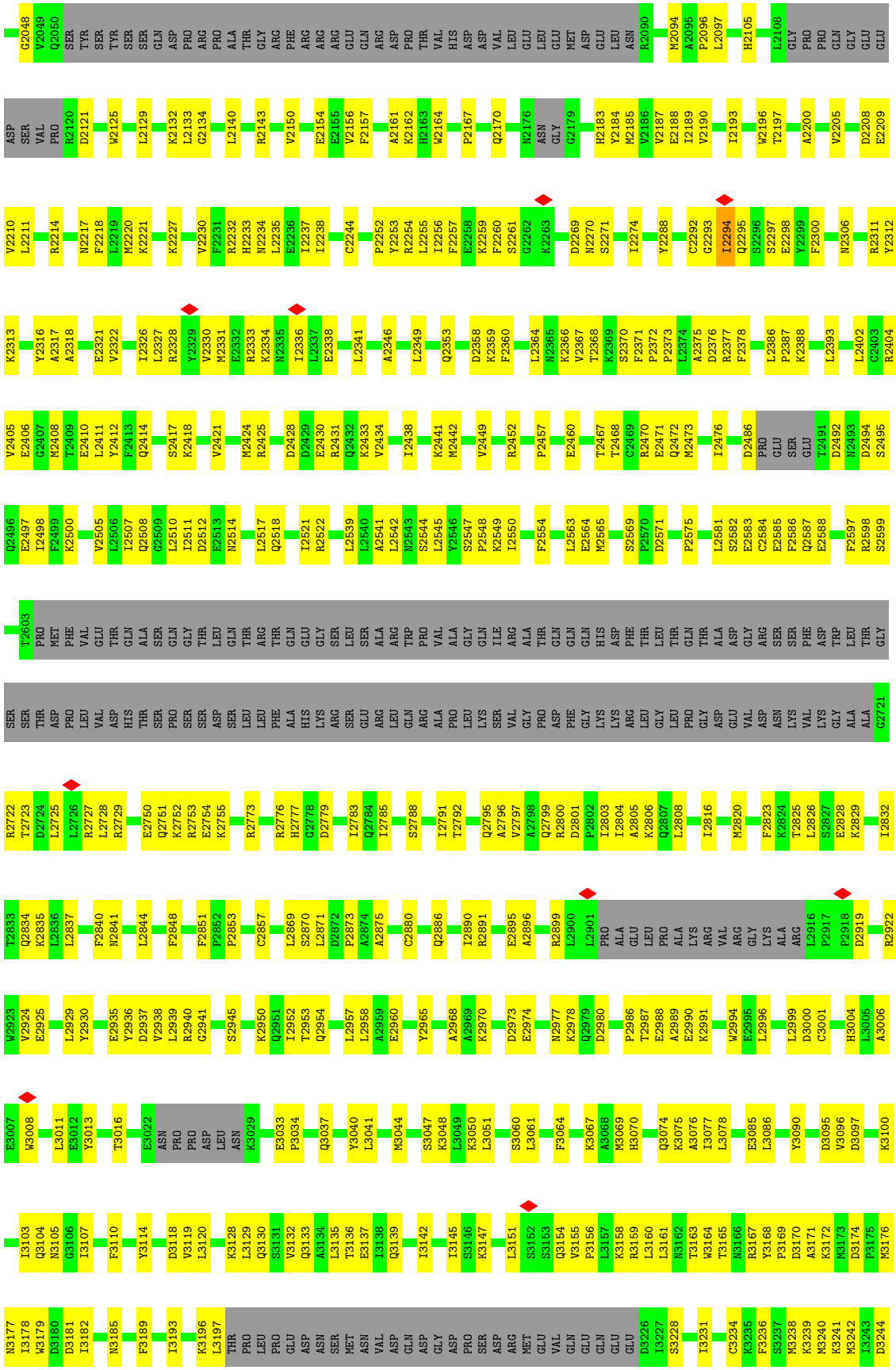


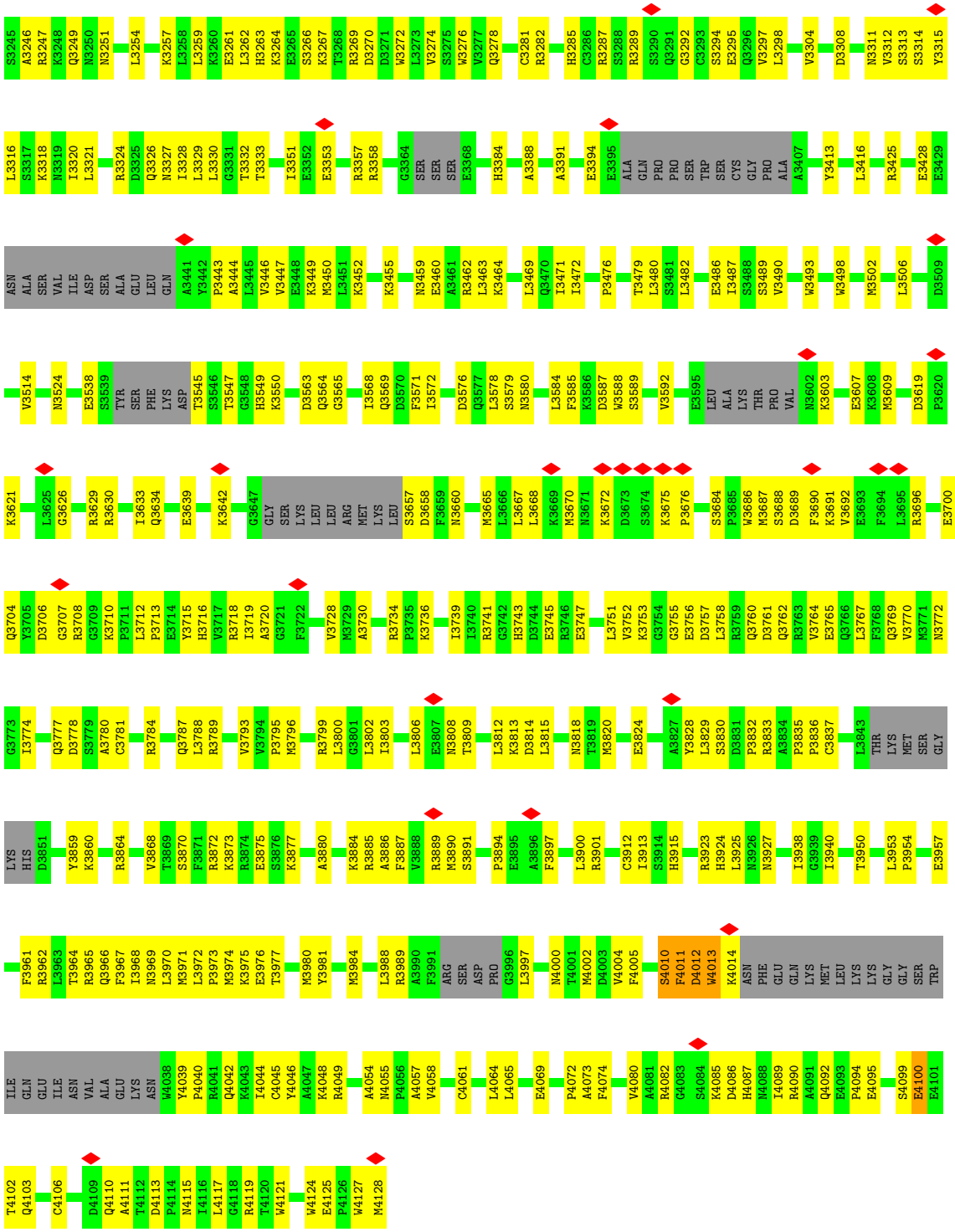
• Molecule 5: DNA-dependent protein kinase catalytic subunit

Chain S:



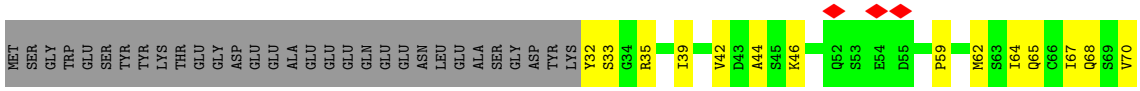


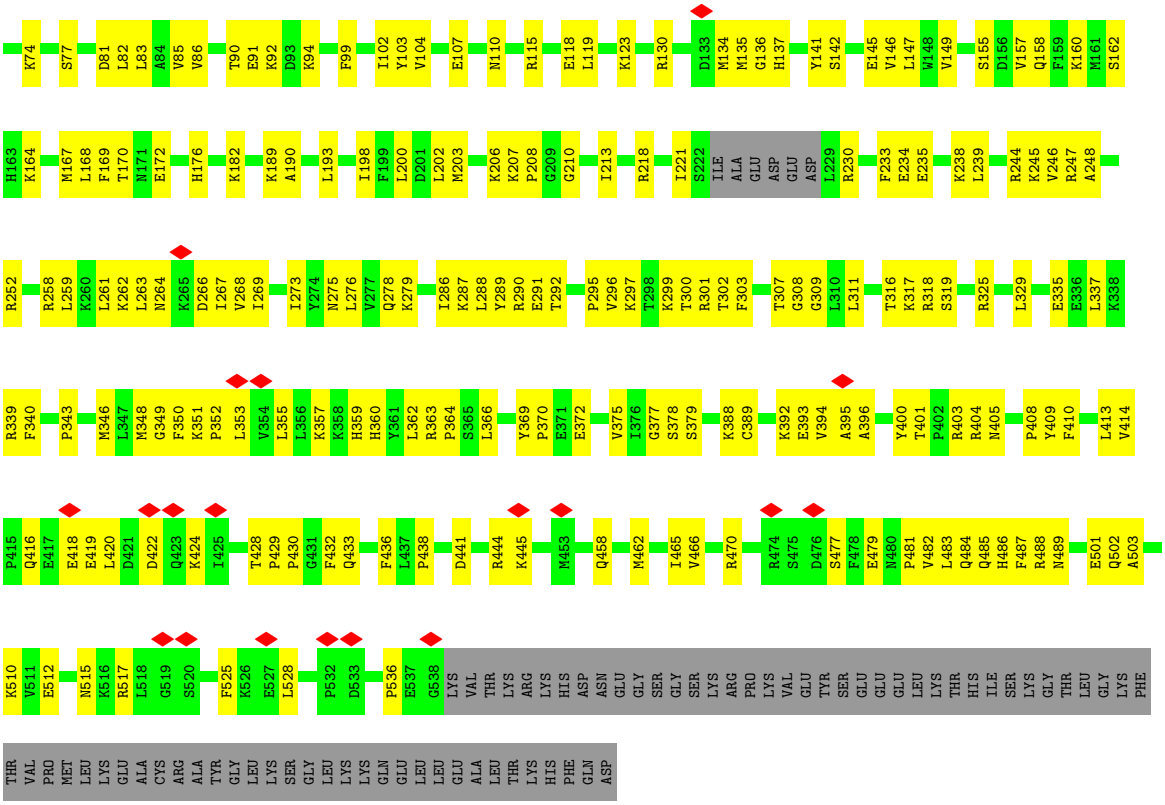




• Molecule 6: X-ray repair cross-complementing protein 6

Chain T:





• Molecule 7: DNA

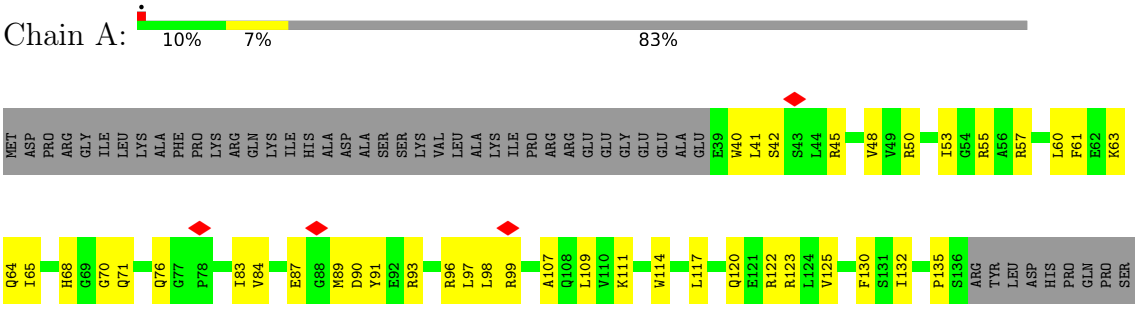
Chain i:

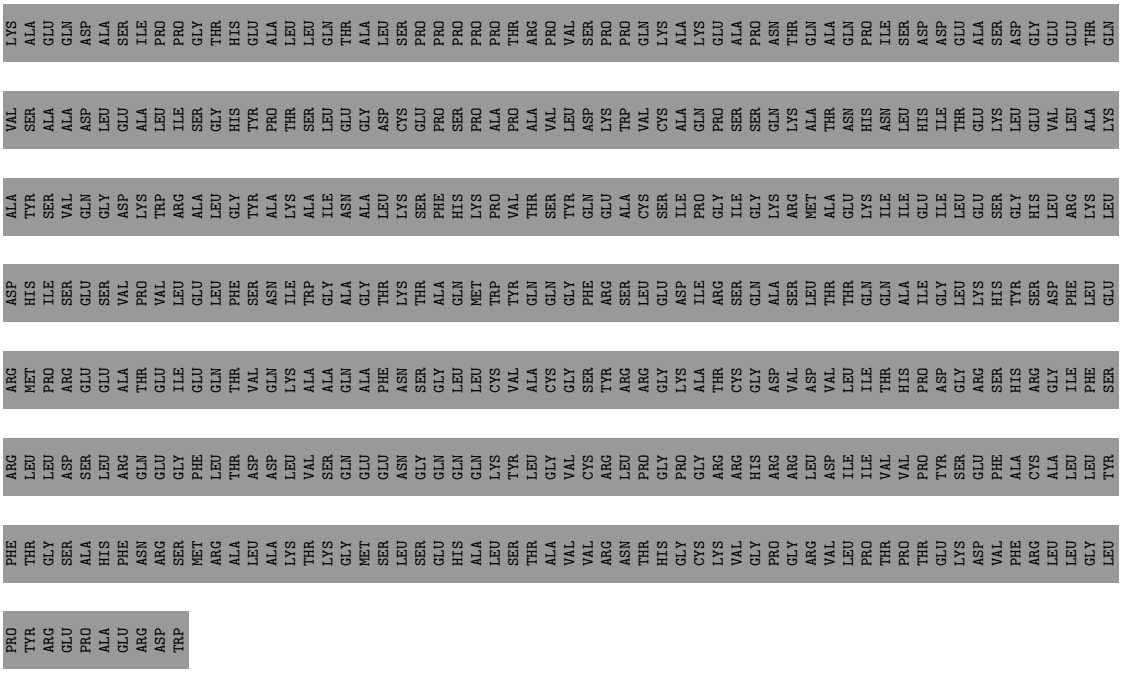
• Molecule 8: DNA

Chain j:

• Molecule 9: DNA polymerase lambda

Chain A:




LYS ALA GLU GLN ASP ALA SER ILE PRO PRO GLY THR HIS GLU ALA LEU LEU GLN THR ALA LEU SER PRO PRO PRO PRO THR ARG PRO VAL SER PRO PRO GLN LYS ALA LYS GLU ALA PRO ASN THR GLN ALA GLN PRO ILE SER ASP ASP GLU ALA SER ASP GLY GLU GLU THR GLN
VAL SER ALA ALA ASP LEU GLU ALA LEU ILE SER GLY HIS TYR PRO THR SER LEU GLU GLY ASP PRO SER PRO ALA PRO ALA VAL LEU ASP LYS TRP VAL CYS ALA GLN PRO SER SER GLN LYS ALA THR ASN HIS ASN LEU HIS ILE THR GLU LYS LEU GLU VAL LEU ALA LYS
ALA TYR SER VAL GLN GLY ASP LYS TRP ARG ALA LEU GLY TYR ALA LYS ALA ILE ASN ALA LEU LYS SER PHE HIS LYS PRO VAL THR SER TYR GLN GLU ALA CYS SER ILE PRO GLY ILE GLY LYS ARG MET ALA GLU LYS ILE ILE GLU ILE LEU GLU SER GLY HIS LEU ARG LYS LEU
ASP HIS ILE SER GLU SER VAL PRO VAL LEU GLU GLU LEU PHE SER ASN ILE TRP GLY ALA GLY THR LYS THR ALA GLN MET TRP TYR GLN GLN GLY PHE ARG SER LEU GLU ASP ILE ARG SER GLN ALA SER LEU THR THR GLN GLN ALA ILE GLY LEU LYS HIS TYR SER ASP PHE LEU GLU
ARG MET PRO ARG GLU GLU ILE GLU GLN THR VAL GLN LYS ALA ALA GLN ALA PHE ASN SER GLY LEU LEU CYS VAL ALA CYS GLY SER TYR ARG ARG GLY LYS ALA THR CYS GLY ASP VAL ASP VAL LEU ILE THR HIS PRO ASP GLY ARG SER HIS ARG GLY ILE PHE SER
ARG LEU LEU ASP SER LEU ARG GLN GLU GLY PHE LEU THR ASP ASP LEU VAL SER GLN GLU GLU ASN GLY GLN GLN GLN LYS TYR LEU GLY VAL CYS ARG LEU PRO GLY PRO GLY ARG HIS ARG ARG LEU ASP ILE ILE VAL VAL PRO TYR SER GLU PHE ALA CYS ALA LEU LEU TYR
PHE THR GLY SER ALA HIS PHE ASN ARG SER MET ARG ALA LEU ALA LYS THR LYS GLY MET SER LEU SER GLU HIS ALA LEU SER THR ALA VAL VAL ARG ASN THR HIS GLY CYS LYS VAL GLY PRO GLY ARG VAL LEU PRO THR PRO THR GLU LYS ASP VAL PHE ARG LEU LEU GLY LEU
PRO TYR ARG GLU PRO ALA GLU ARG ASP TRP

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	14964	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 ($e^-/Å^2$)	51.96	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	130000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.696	Depositor
Minimum map value	-0.178	Depositor
Average map value	0.002	Depositor
Map value standard deviation	0.034	Depositor
Recommended contour level	0.211	Depositor
Map size (Å)	417.28, 417.28, 417.28	wwPDB
Map dimensions	320, 320, 320	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.304, 1.304, 1.304	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	E	0.14	0/1997	0.40	0/2695
2	L	0.14	0/5182	0.36	0/6990
3	M	0.17	0/164	0.44	0/220
4	P	0.11	0/485	0.34	0/647
4	Q	0.13	0/480	0.35	0/640
5	S	0.19	0/28315	0.48	0/38300
6	T	0.17	0/4079	0.41	0/5499
7	i	0.28	0/572	0.50	0/879
8	j	0.32	0/575	0.58	0/888
9	A	0.16	0/780	0.44	0/1057
All	All	0.18	0/42629	0.46	0/57815

There are no bond length outliers.

There are no bond angle outliers.

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	E	1956	0	1905	76	0
2	L	5084	0	5080	206	0
3	M	160	0	148	6	0
4	P	480	0	501	27	0
4	Q	475	0	492	25	0
5	S	27773	0	27732	1057	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	T	4001	0	4050	188	0
7	i	510	0	285	18	0
8	j	514	0	290	20	0
9	A	764	0	785	35	0
All	All	41717	0	41268	1564	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 19.

The worst 5 of 1564 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)
5:S:2806:LYS:HG3	5:S:2857:CYS:HB3	1.59	0.84
1:E:757:GLU:HB3	1:E:763:ASP:H	1.42	0.83
5:S:835:LYS:HA	5:S:838:LYS:HG2	1.62	0.81
5:S:1300:SER:HA	5:S:1304:HIS:HB2	1.65	0.79
5:S:1820:VAL:HA	5:S:1824:LEU:HD23	1.64	0.79

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	E	242/911 (27%)	208 (86%)	32 (13%)	2 (1%)	16 53
2	L	635/732 (87%)	598 (94%)	36 (6%)	1 (0%)	44 78
3	M	21/204 (10%)	21 (100%)	0	0	100 100
4	P	55/336 (16%)	55 (100%)	0	0	100 100
4	Q	55/336 (16%)	55 (100%)	0	0	100 100
5	S	3459/4128 (84%)	3188 (92%)	263 (8%)	8 (0%)	44 78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	T	497/609 (82%)	461 (93%)	36 (7%)	0	100	100
9	A	96/575 (17%)	96 (100%)	0	0	100	100
All	All	5060/7831 (65%)	4682 (92%)	367 (7%)	11 (0%)	45	78

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	814	ARG
5	S	4012	ASP
5	S	4013	TRP
5	S	1231	GLN
5	S	4010	SER

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	E	217/808 (27%)	217 (100%)	0	100	100
2	L	559/649 (86%)	559 (100%)	0	100	100
3	M	16/160 (10%)	16 (100%)	0	100	100
4	P	53/303 (18%)	53 (100%)	0	100	100
4	Q	51/303 (17%)	51 (100%)	0	100	100
5	S	2998/3671 (82%)	2998 (100%)	0	100	100
6	T	444/548 (81%)	444 (100%)	0	100	100
9	A	81/480 (17%)	81 (100%)	0	100	100
All	All	4419/6922 (64%)	4419 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
5	S	1611	GLN
6	T	484	GLN
5	S	2493	ASN
6	T	433	GLN
5	S	4055	ASN

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 oligosaccharides 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

There are no chain breaks in this entry.

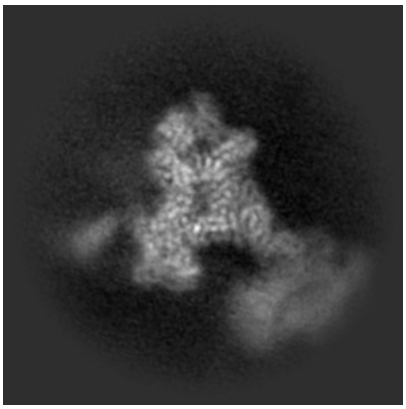
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-51249. These allow visual inspection of the internal detail of the map and identification of artifacts.

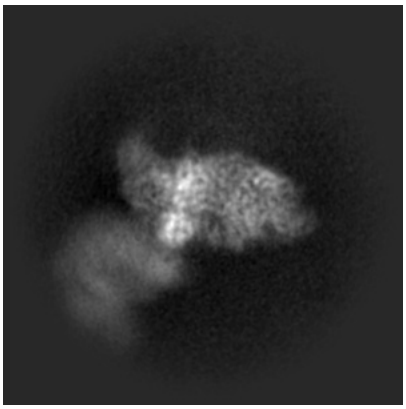
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

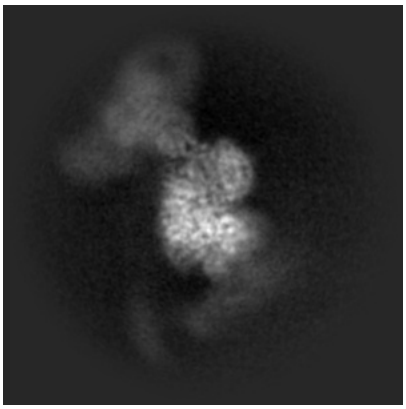
6.1.1 Primary map



X

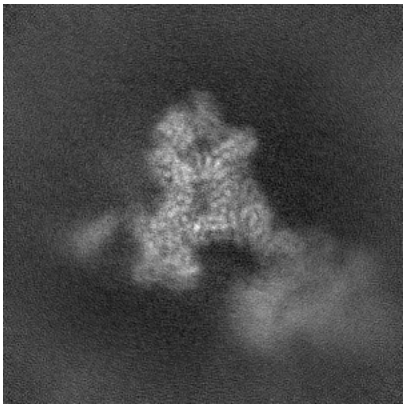


Y

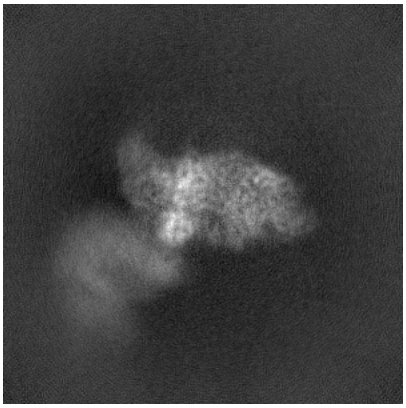


Z

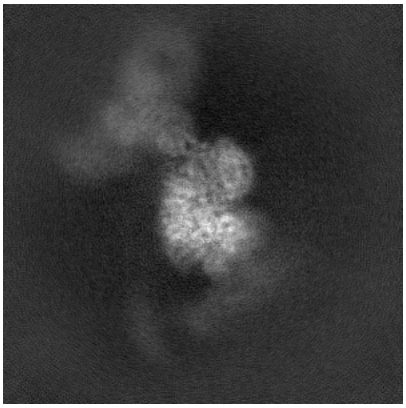
6.1.2 Raw map



X



Y

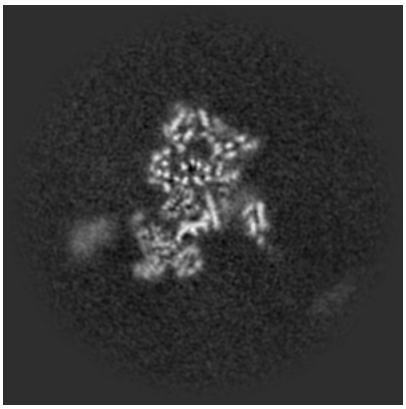


Z

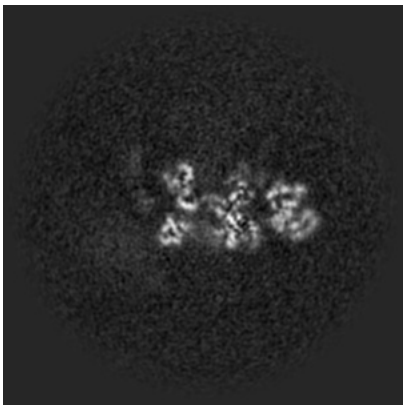
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

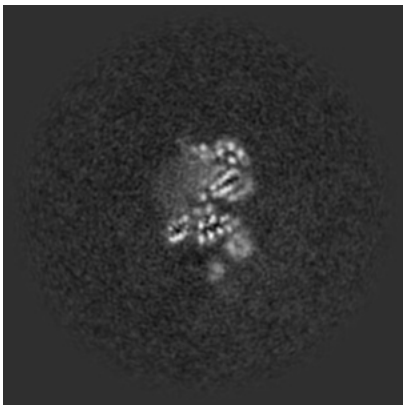
6.2.1 Primary map



X Index: 160

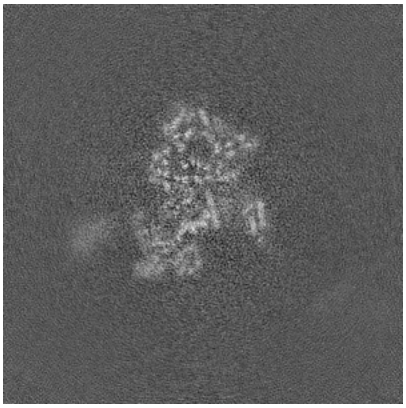


Y Index: 160

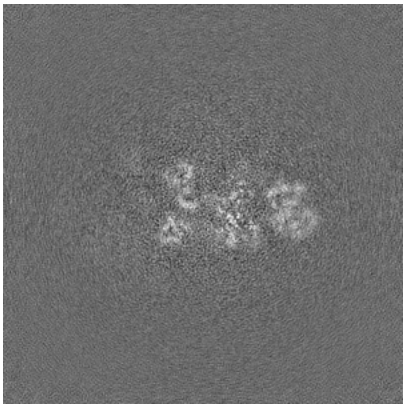


Z Index: 160

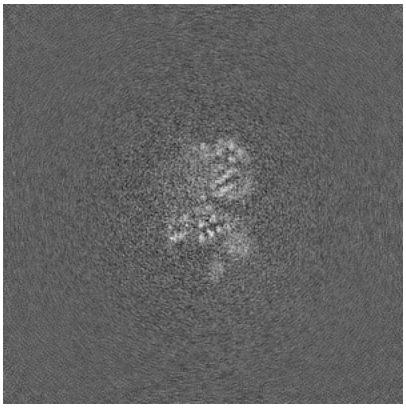
6.2.2 Raw map



X Index: 160



Y Index: 160

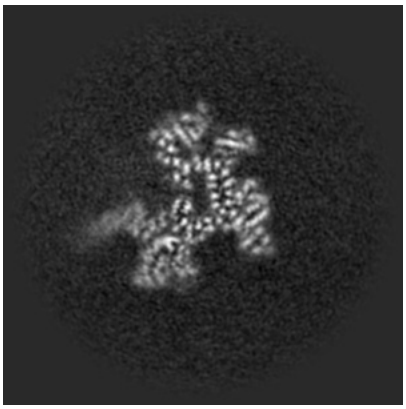


Z Index: 160

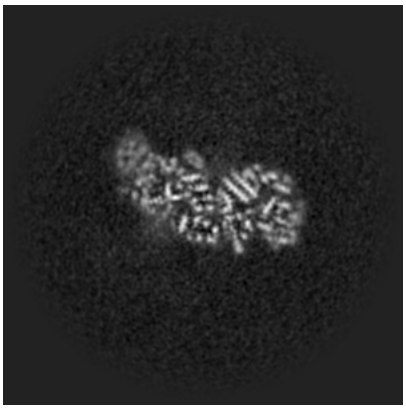
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

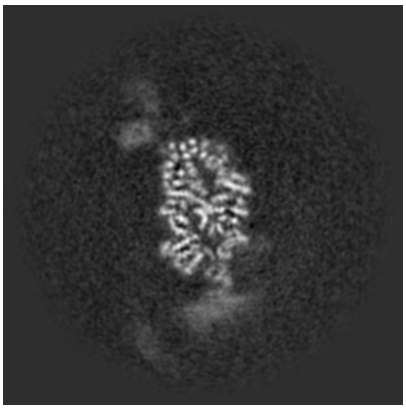
6.3.1 Primary map



X Index: 175

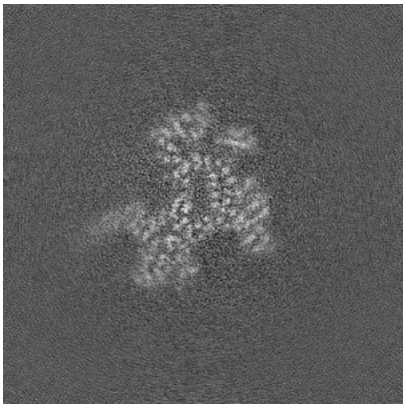


Y Index: 138

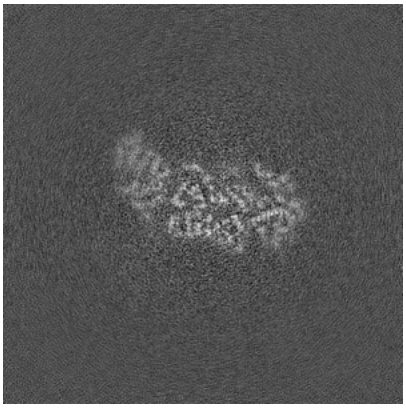


Z Index: 142

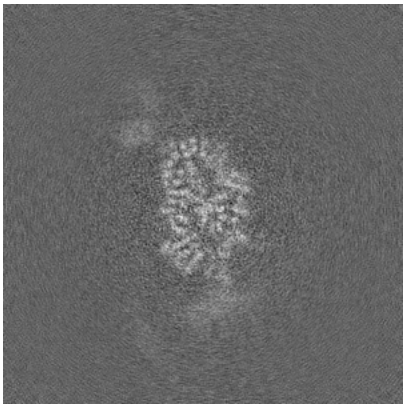
6.3.2 Raw map



X Index: 174



Y Index: 143

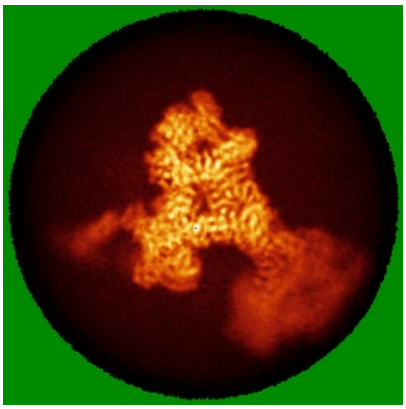


Z Index: 142

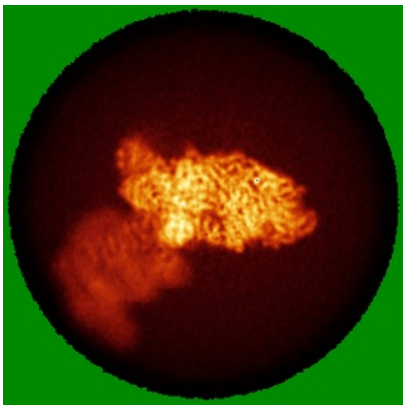
The images above show the largest variance slices of the map in three orthogonal directions.

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

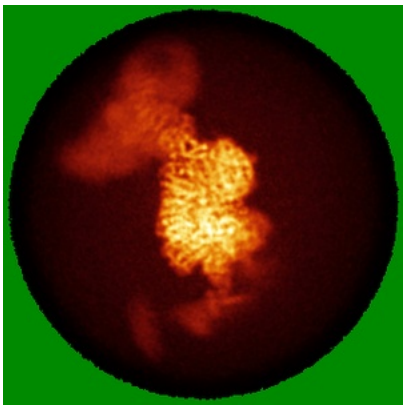
6.4.1 Primary map



X

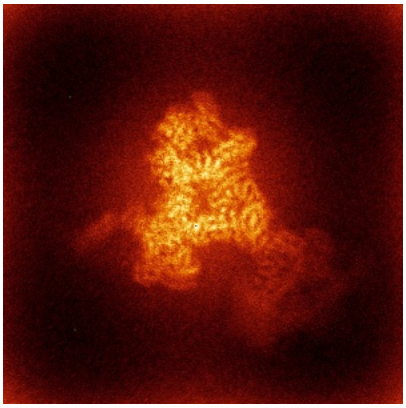


Y

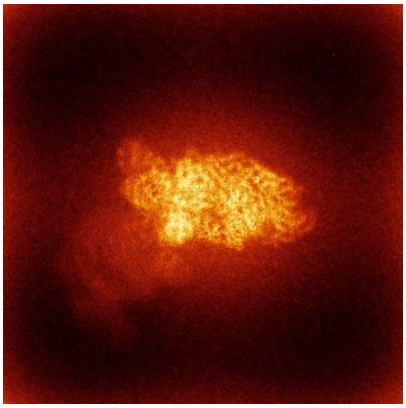


Z

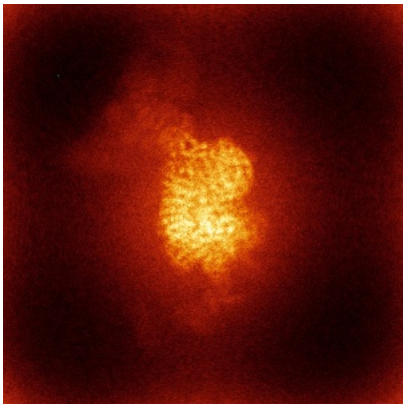
6.4.2 Raw map



X



Y

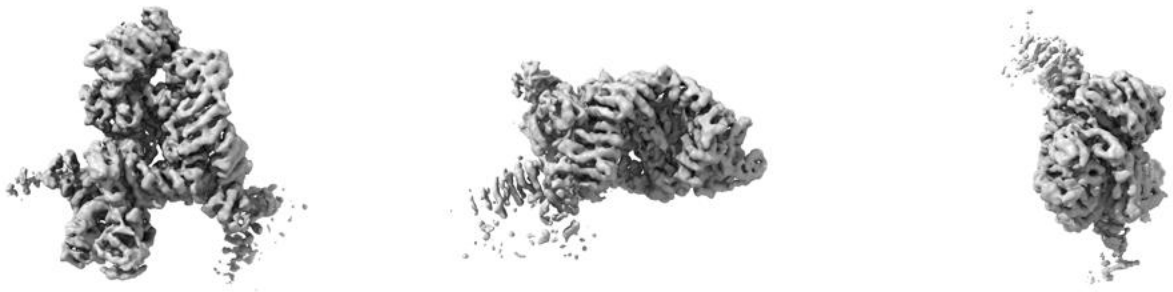


Z

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 0.211. 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.

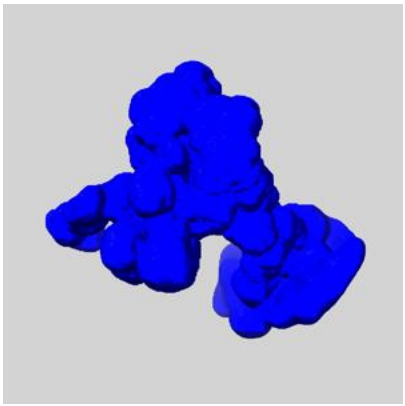
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

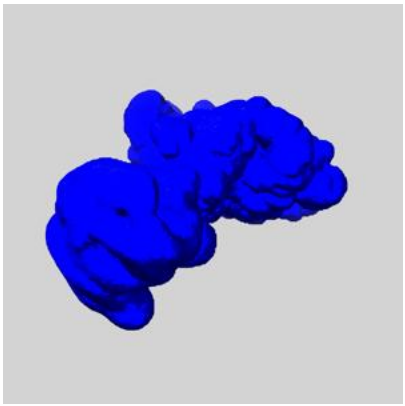
A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

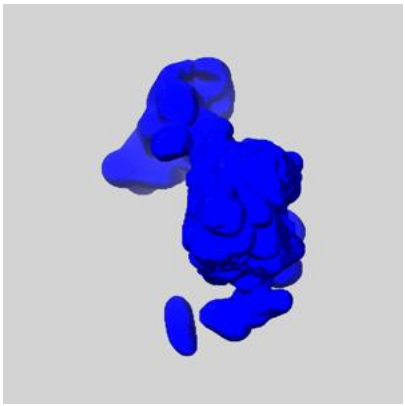
6.6.1 emd_51249_msk_1.map (i)



X



Y

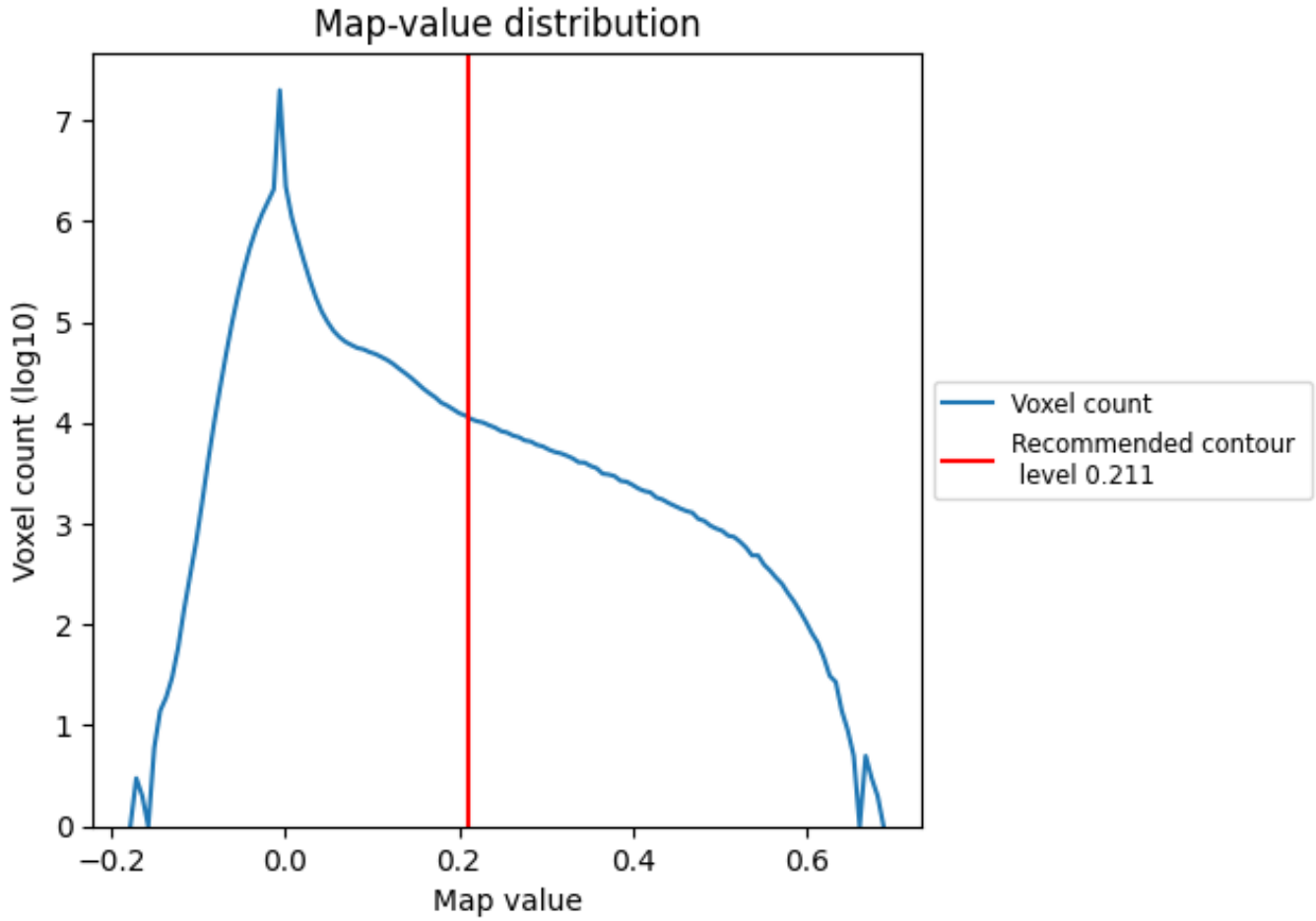


Z

7 Map analysis

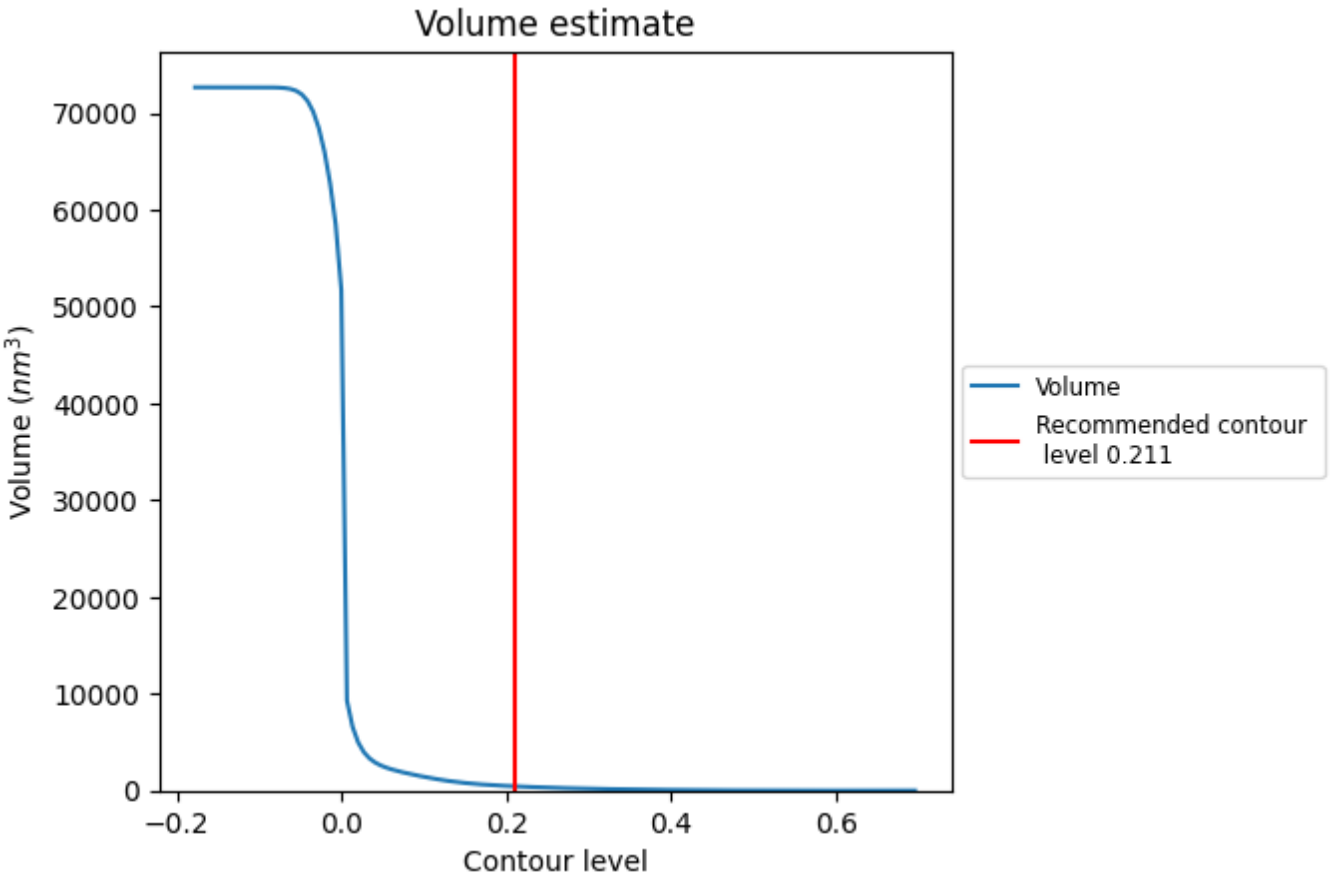
This section contains the results of statistical analysis of the map.

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.

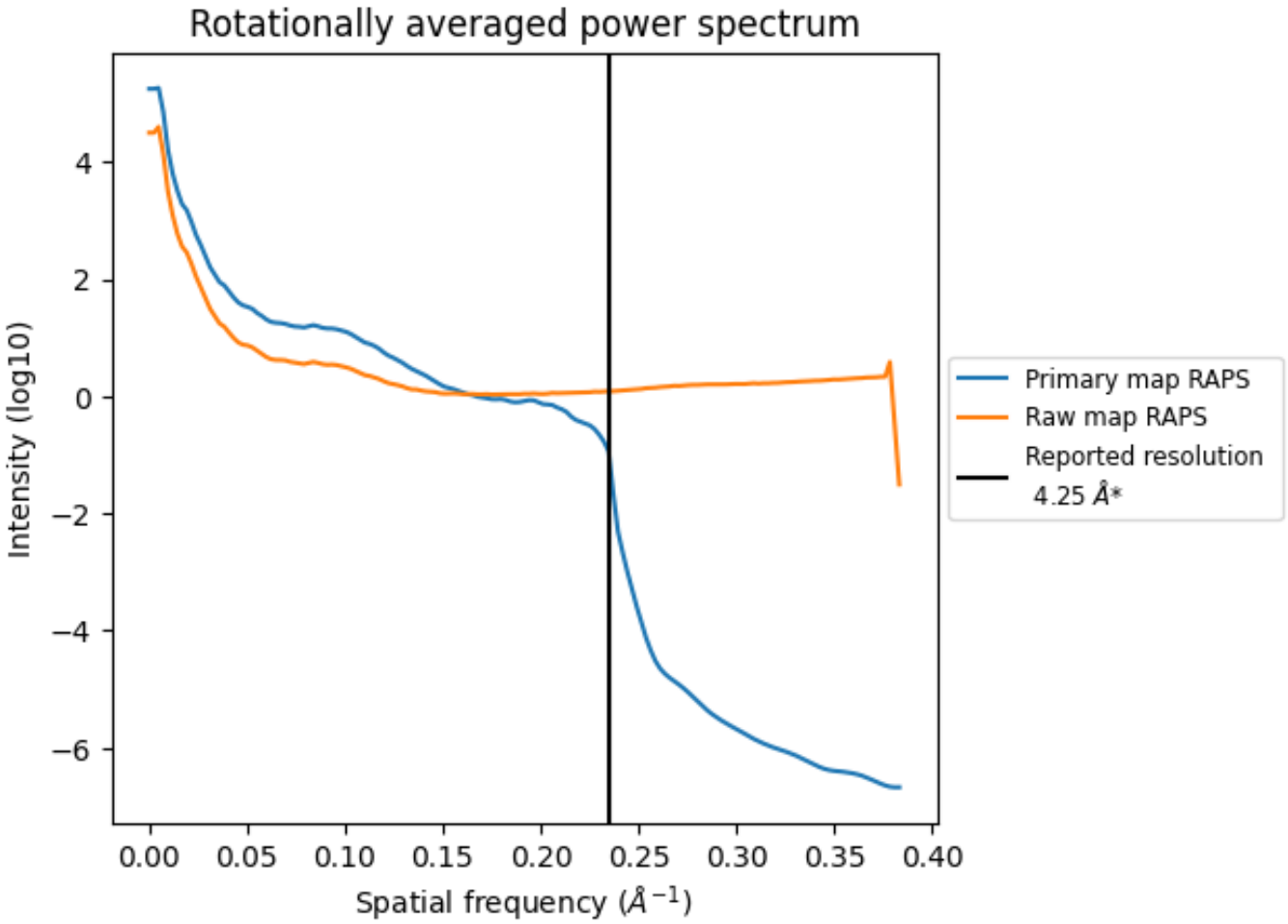
7.2 Volume estimate (i)



The volume at the recommended contour level is 434 nm^3; this corresponds to an approximate mass of 392 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)

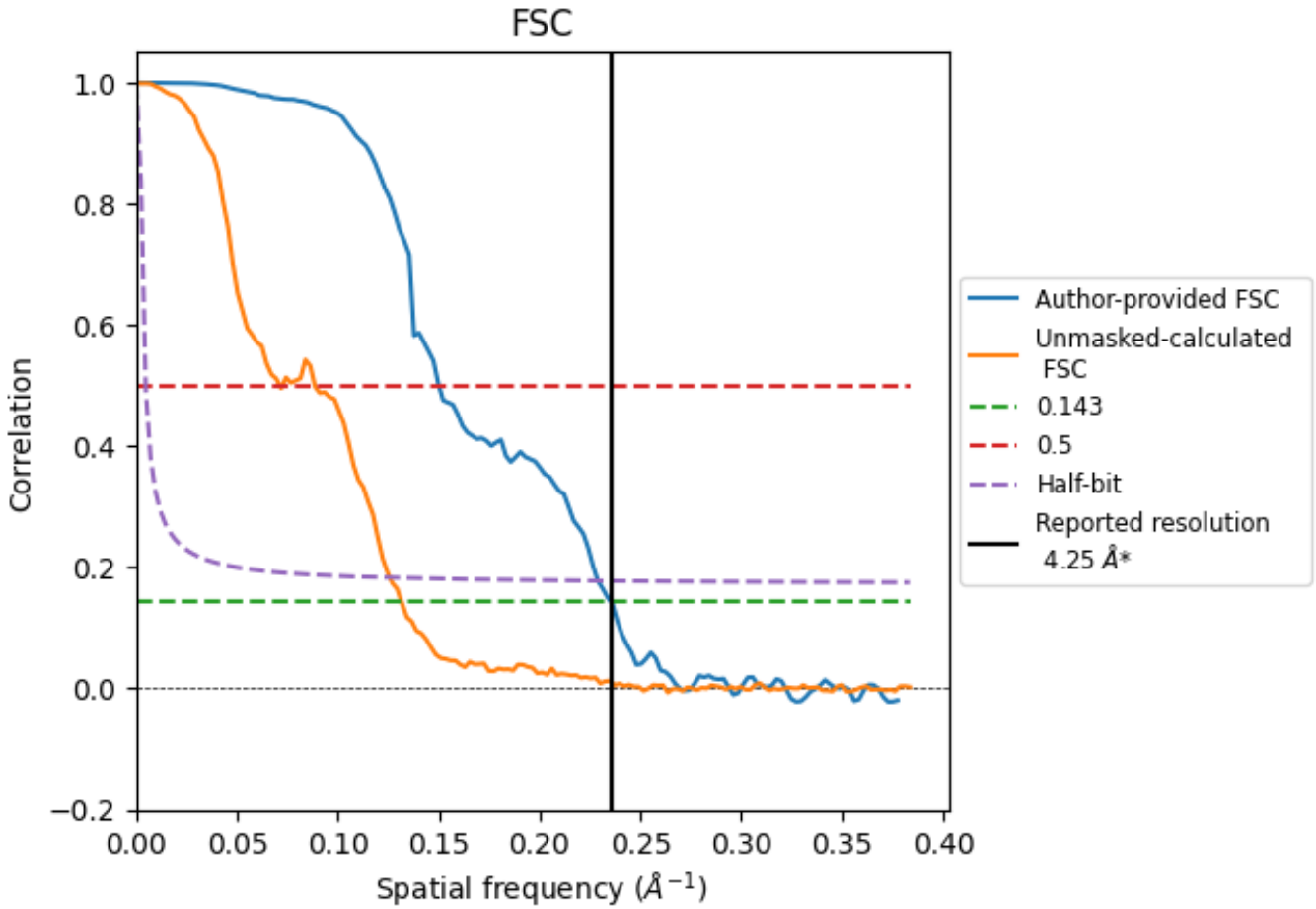


*Reported resolution corresponds to spatial frequency of 0.235 Å^{-1}

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 0.235 Å^{-1}

8.2 Resolution estimates (i)

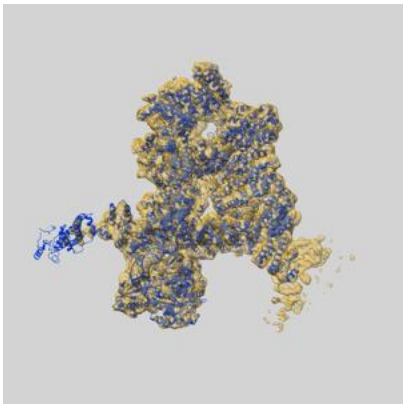
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	4.25	-	-
Author-provided FSC curve	4.25	6.66	4.37
Unmasked-calculated*	7.60	14.04	7.95

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 7.60 differs from the reported value 4.25 by more than 10 %

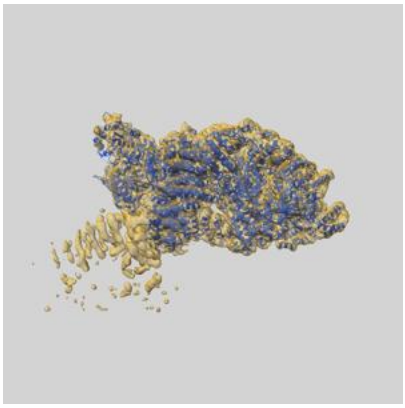
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-51249 and PDB model 9GD7. Per-residue inclusion information can be found in section 3 on page 6.

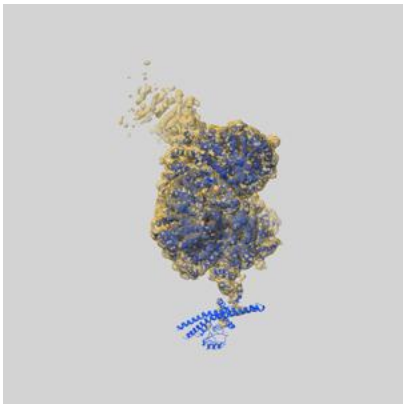
9.1 Map-model overlay



X



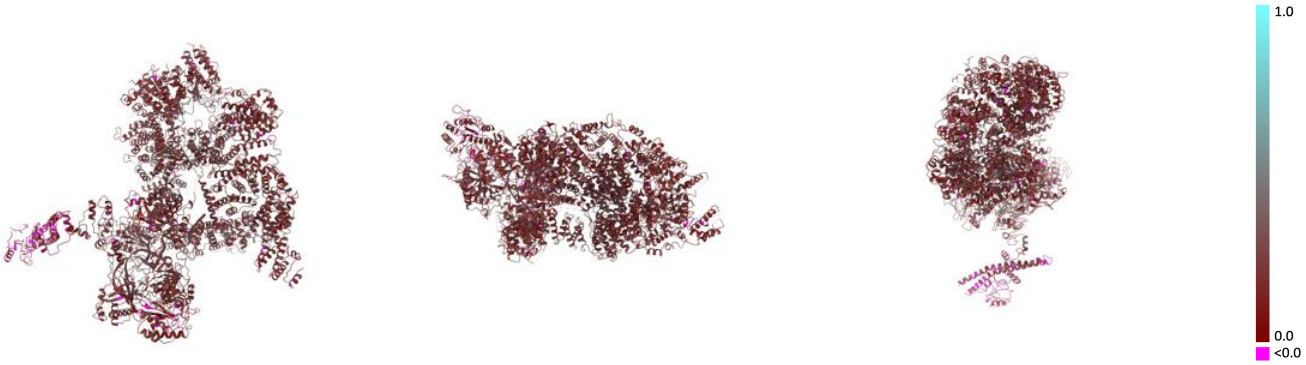
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.211 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.

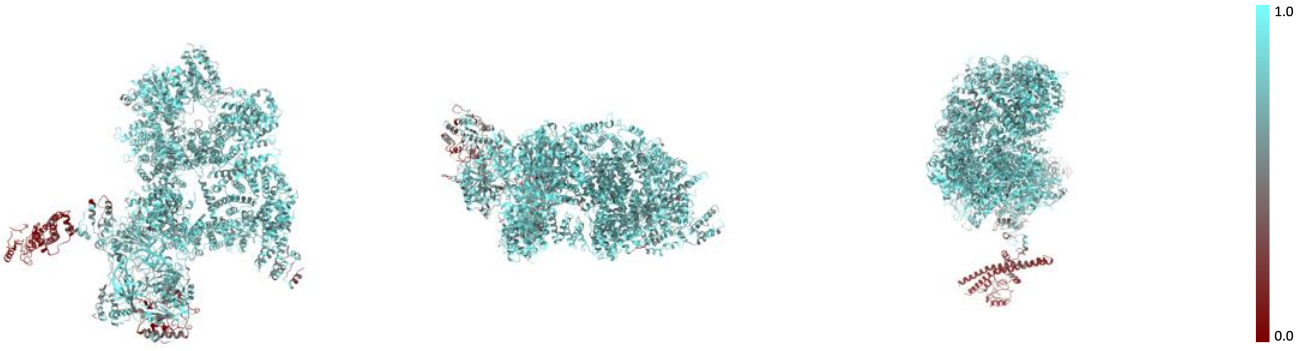
9.2 Q-score mapped to coordinate model



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.

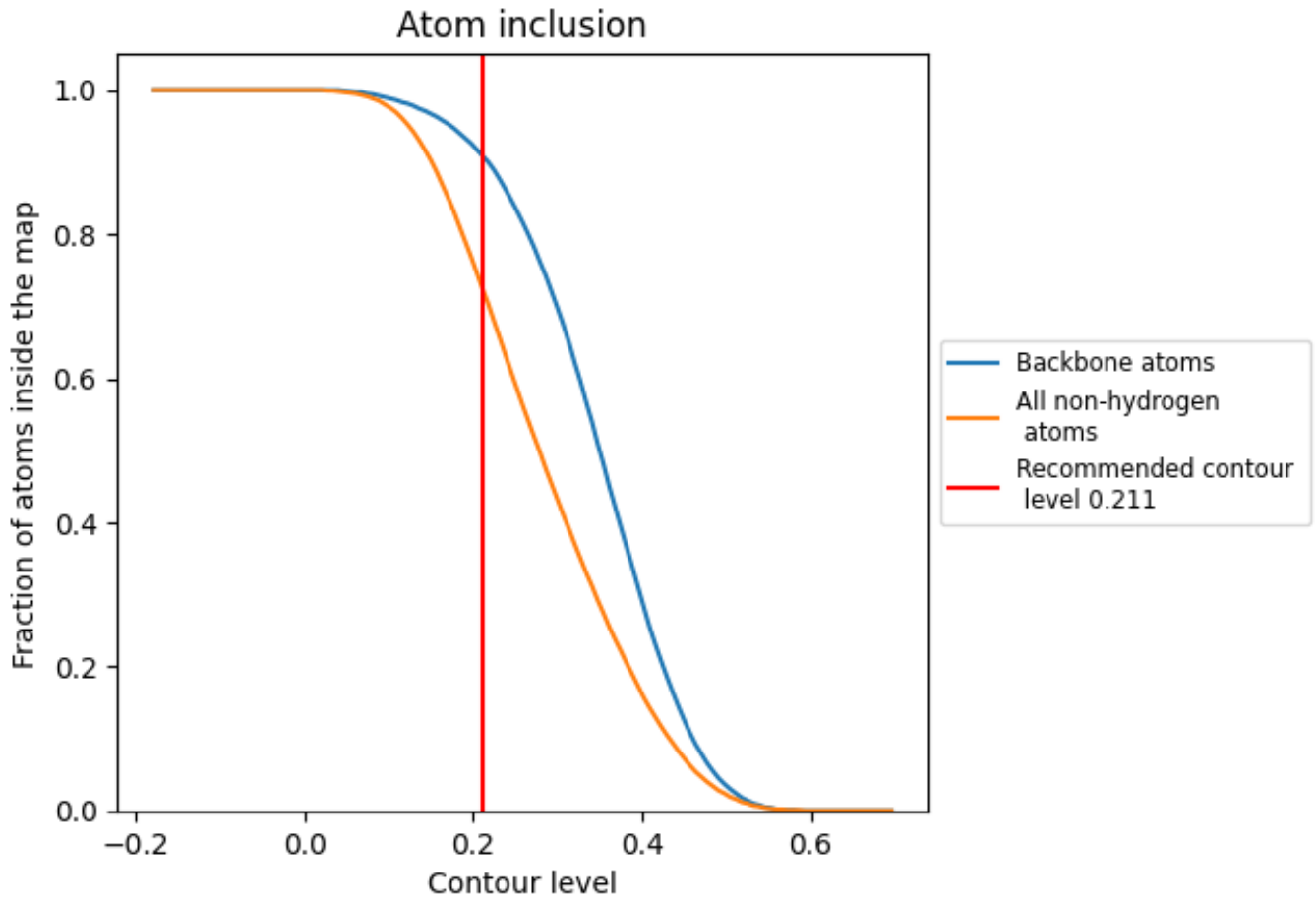
9.3 Atom inclusion mapped to coordinate model



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 (0.211).

9.4 Atom inclusion (i)



At the recommended contour level, 91% of all backbone atoms, 72% 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 (0.211) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7240	0.2280
A	0.7750	0.2160
E	0.2650	0.1390
L	0.6110	0.2070
M	0.6560	0.2950
P	0.0530	0.0970
Q	0.0970	0.0800
S	0.7870	0.2370
T	0.7420	0.2480
i	0.9770	0.3020
j	0.9810	0.2960