wwPDB EM Validation Summary Report

Nov 9, 2022 – 10:56 AM EST

PDB ID	:	6PTJ
EMDB ID	:	EMD-20471
Title	:	Structure of Ctf4 trimer in complex with one CMG helicase
Authors	:	Yuan, Z.; Georgescu, R.; Bai, L.; Santos, R.; Donnell, M.; Li, H.
Deposited on	:	2019-07-15
Resolution	:	3.80 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:

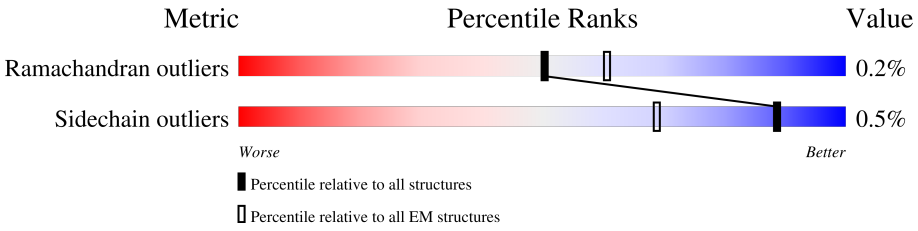
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance

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

The reported resolution of this entry is 3.80 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.



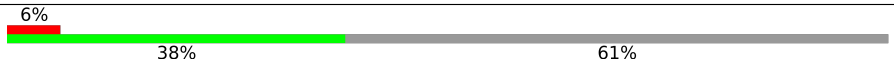
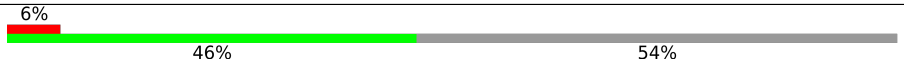
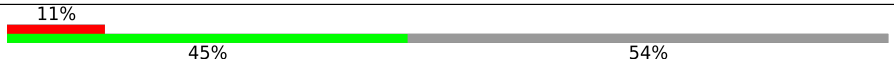
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	208	9%, 100%
2	B	213	86%, 13%
3	C	194	80%, 18%
4	D	294	77%, 22%
5	c	650	84%, 15%
6	2	868	30%, 70%
7	3	971	27%, 72%
8	4	933	30%, 69%
9	5	775	32%, 67%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	6	1017	
11	7	845	
12	E	927	
12	F	927	
12	G	927	

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 34361 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 replication complex GINS protein PSF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	208	1696	1065	290	331	10	0	0

- Molecule 2 is a protein called DNA replication complex GINS protein PSF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	186	1557	1005	271	277	4	0	0

- Molecule 3 is a protein called DNA replication complex GINS protein PSF3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	159	1288	843	207	232	6	0	0

- Molecule 4 is a protein called DNA replication complex GINS protein SLD5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	230	1891	1206	309	364	12	0	0

- Molecule 5 is a protein called Cell division control protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	c	553	4482	2862	763	844	13	0	0

- Molecule 6 is a protein called DNA replication licensing factor MCM2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	2	260	2025	1284	361	374	6	0	0

- Molecule 7 is a protein called DNA replication licensing factor MCM3.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	3	270	Total	C	N	O	S	0	0
			2139	1361	371	403	4		

- Molecule 8 is a protein called DNA replication licensing factor MCM4.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	4	285	Total	C	N	O	S	0	0
			2299	1450	400	432	17		

- Molecule 9 is a protein called Minichromosome maintenance protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	5	252	Total	C	N	O	S	0	0
			2006	1271	344	382	9		

- Molecule 10 is a protein called DNA replication licensing factor MCM6.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	6	273	Total	C	N	O	S	0	0
			2081	1314	371	390	6		

- Molecule 11 is a protein called DNA replication licensing factor MCM7.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	7	326	Total	C	N	O	S	0	0
			2615	1655	454	493	13		

- Molecule 12 is a protein called DNA polymerase alpha-binding protein.

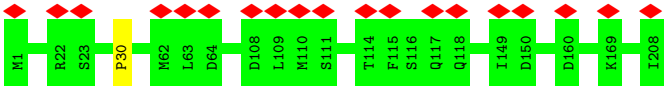
Mol	Chain	Residues	Atoms					AltConf	Trace
12	E	424	Total	C	N	O	S	0	0
			3408	2189	564	640	15		
12	F	431	Total	C	N	O	S	1	0
			3472	2227	576	653	16		
12	G	423	Total	C	N	O	S	1	0
			3402	2186	563	638	15		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

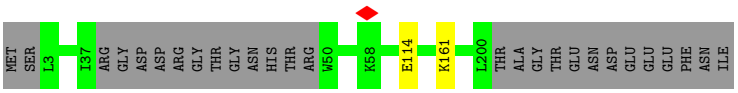
• Molecule 1: DNA replication complex GINS protein PSF1

Chain A:



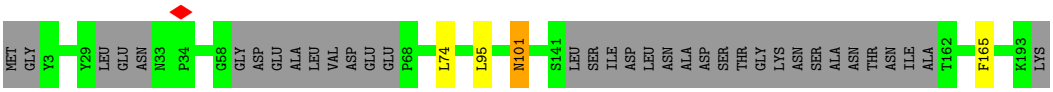
• Molecule 2: DNA replication complex GINS protein PSF2

Chain B:



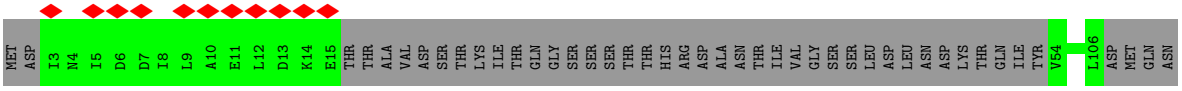
• Molecule 3: DNA replication complex GINS protein PSF3

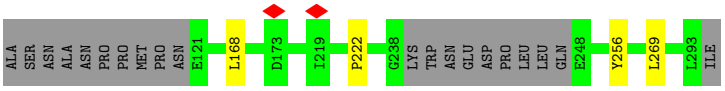
Chain C:



• Molecule 4: DNA replication complex GINS protein SLD5

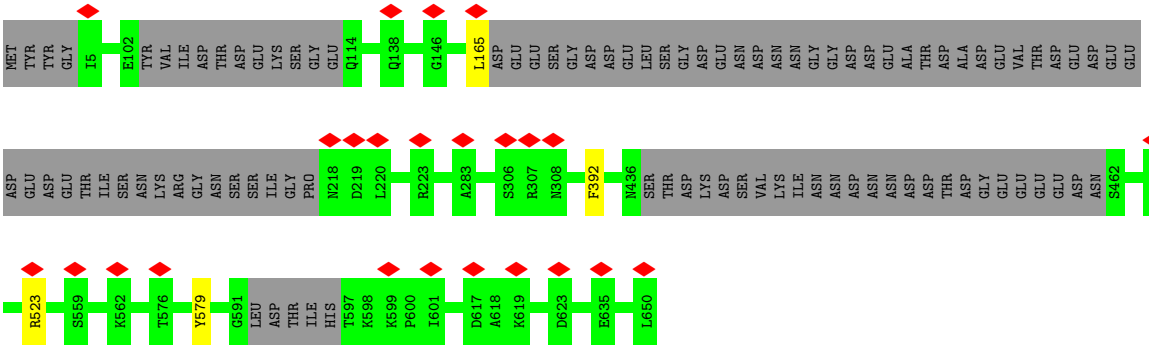
Chain D:





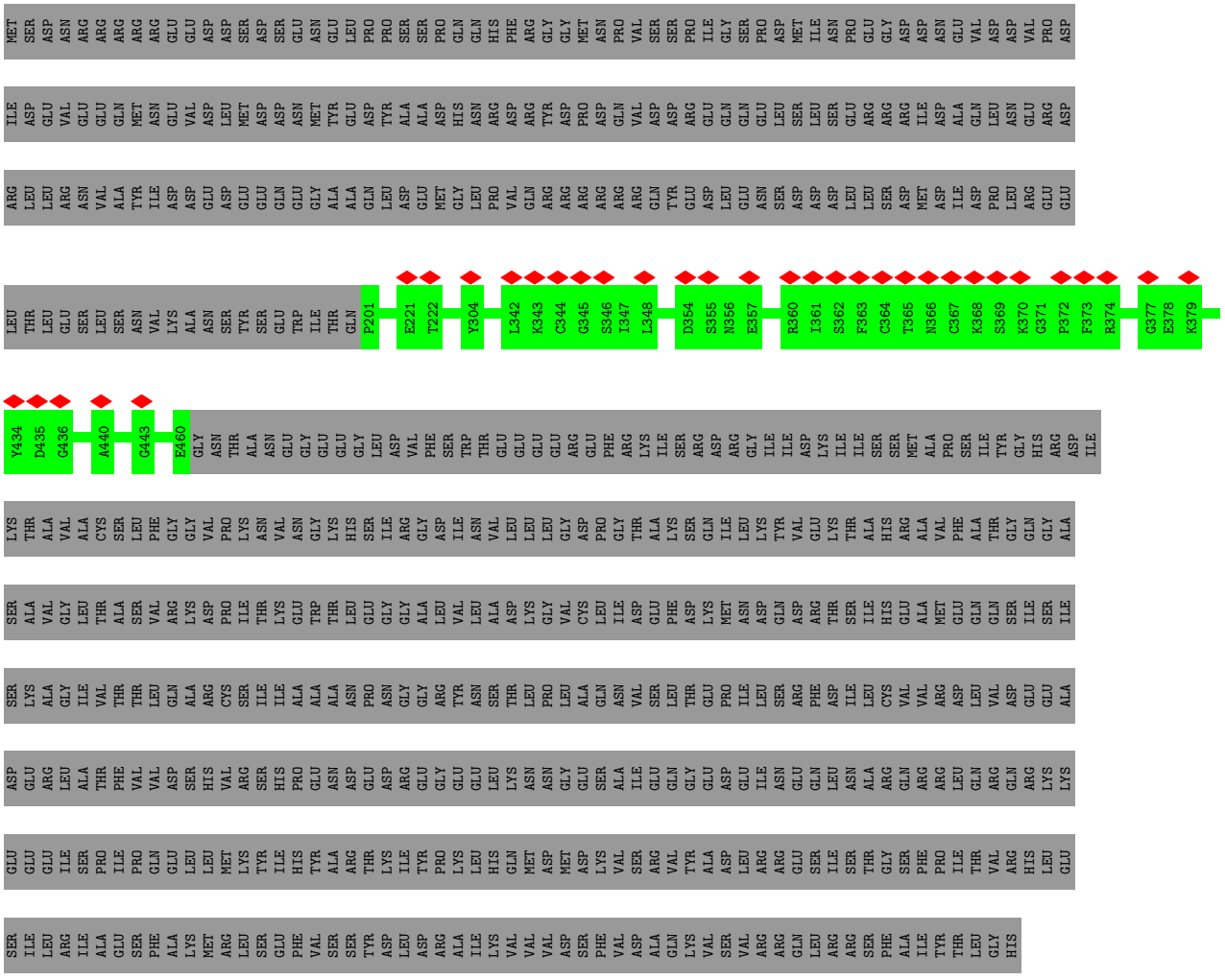
• Molecule 5: Cell division control protein 45

Chain c:



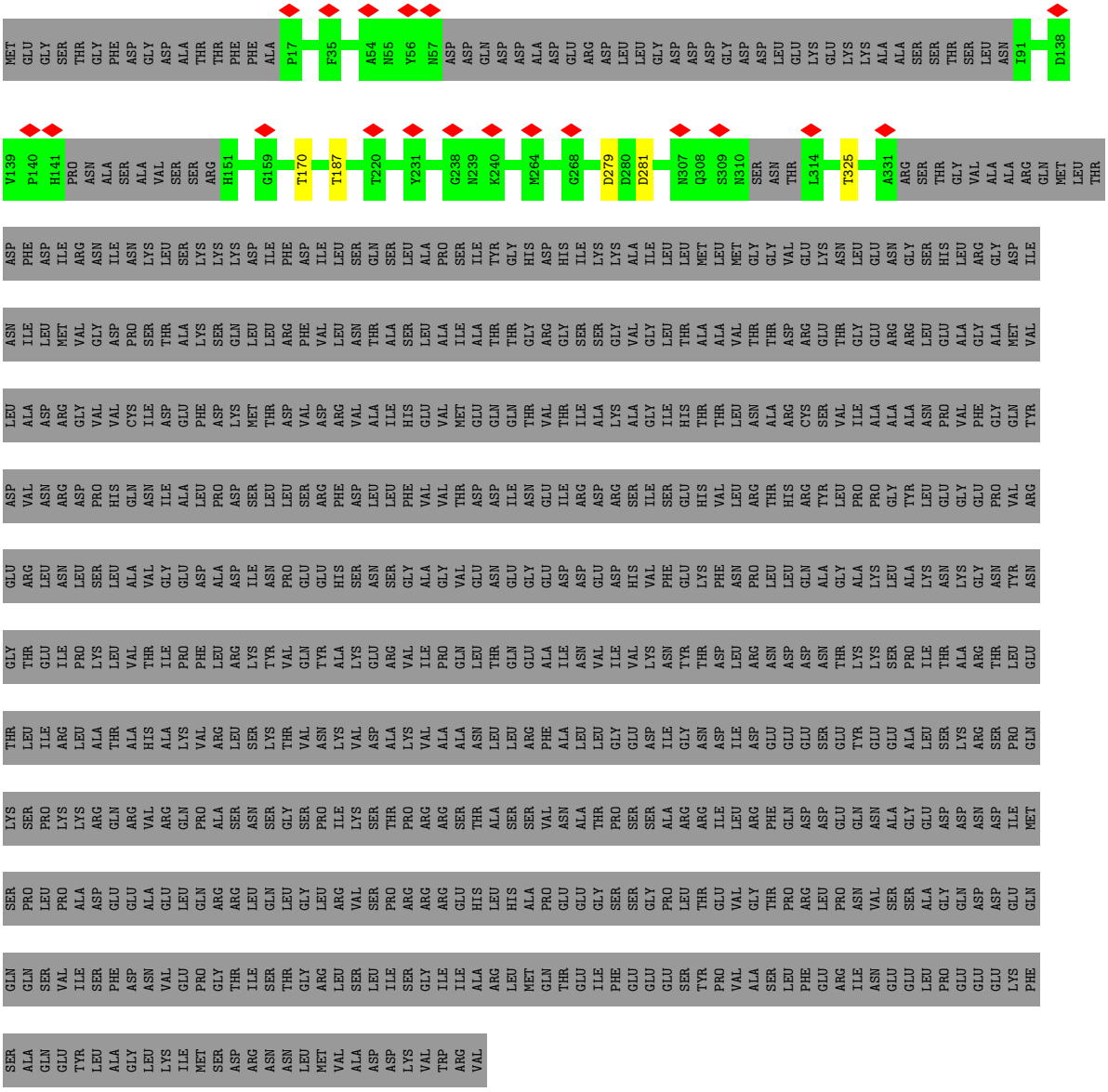
• Molecule 6: DNA replication licensing factor MCM2

Chain 2:



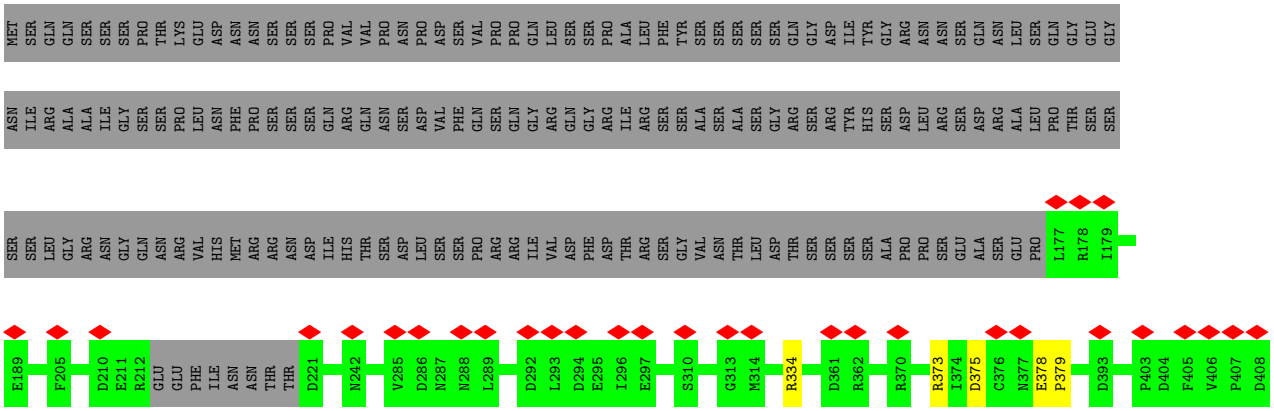
• Molecule 7: DNA replication licensing factor MCM3

Chain 3:



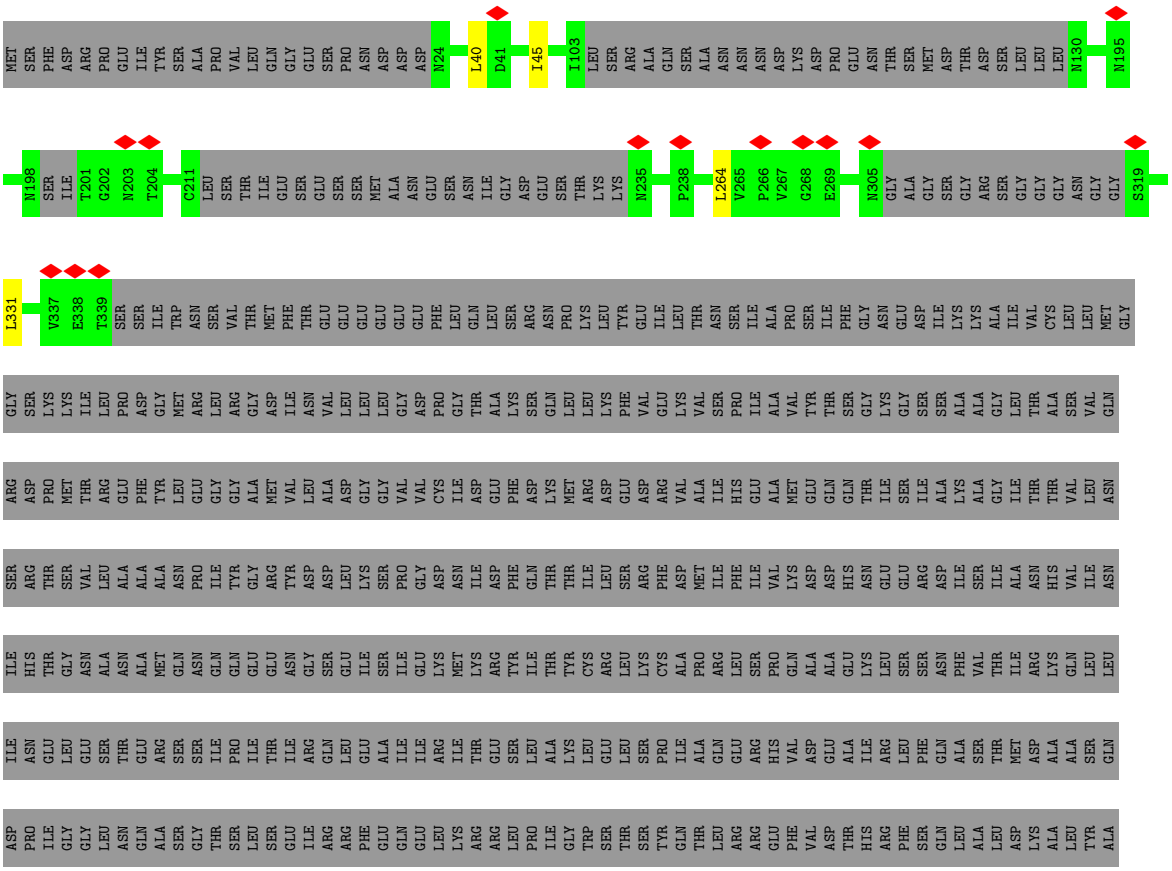
• Molecule 8: DNA replication licensing factor MCM4

Chain 4:



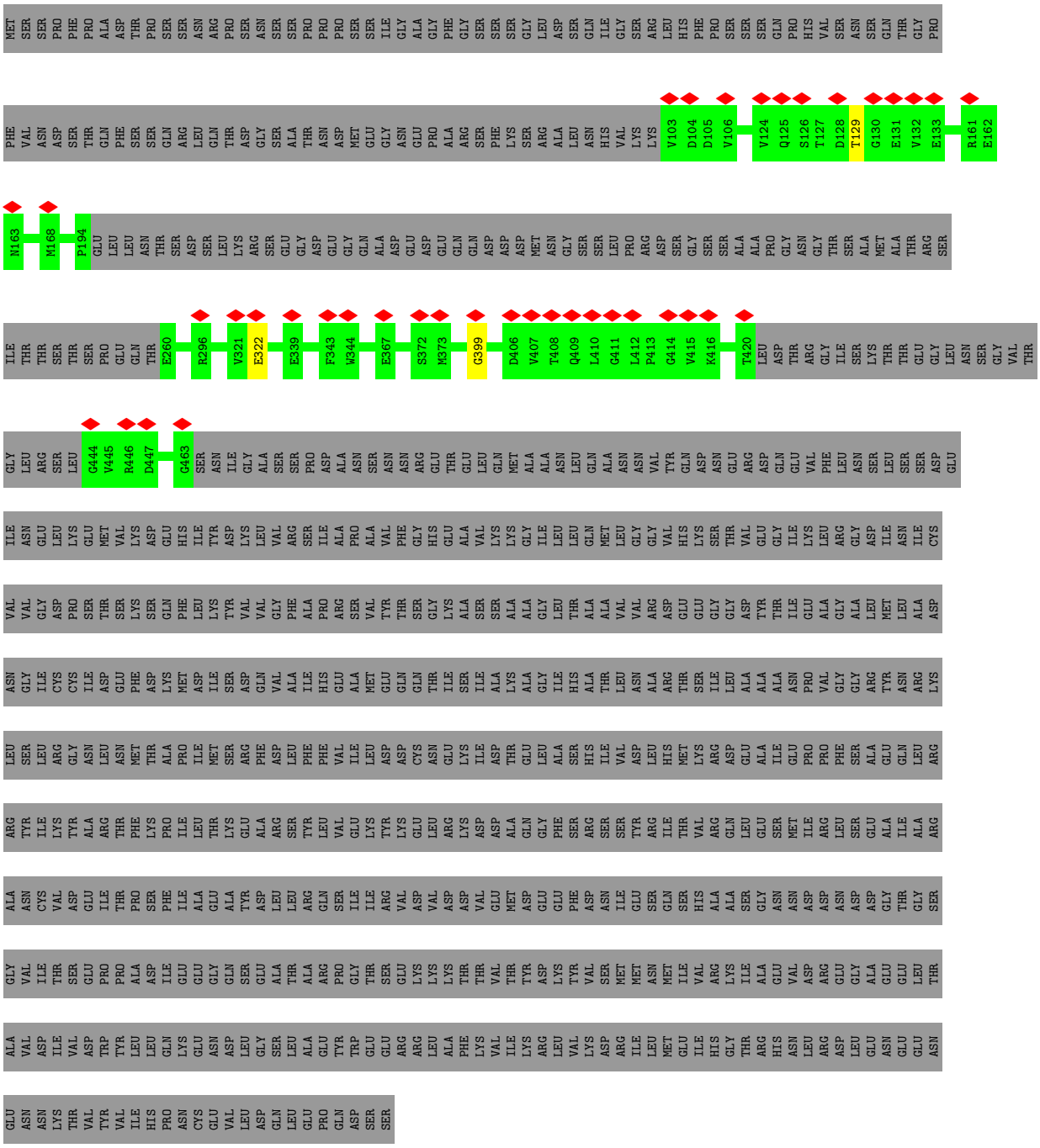
• Molecule 9: Minichromosome maintenance protein 5

Chain 5:



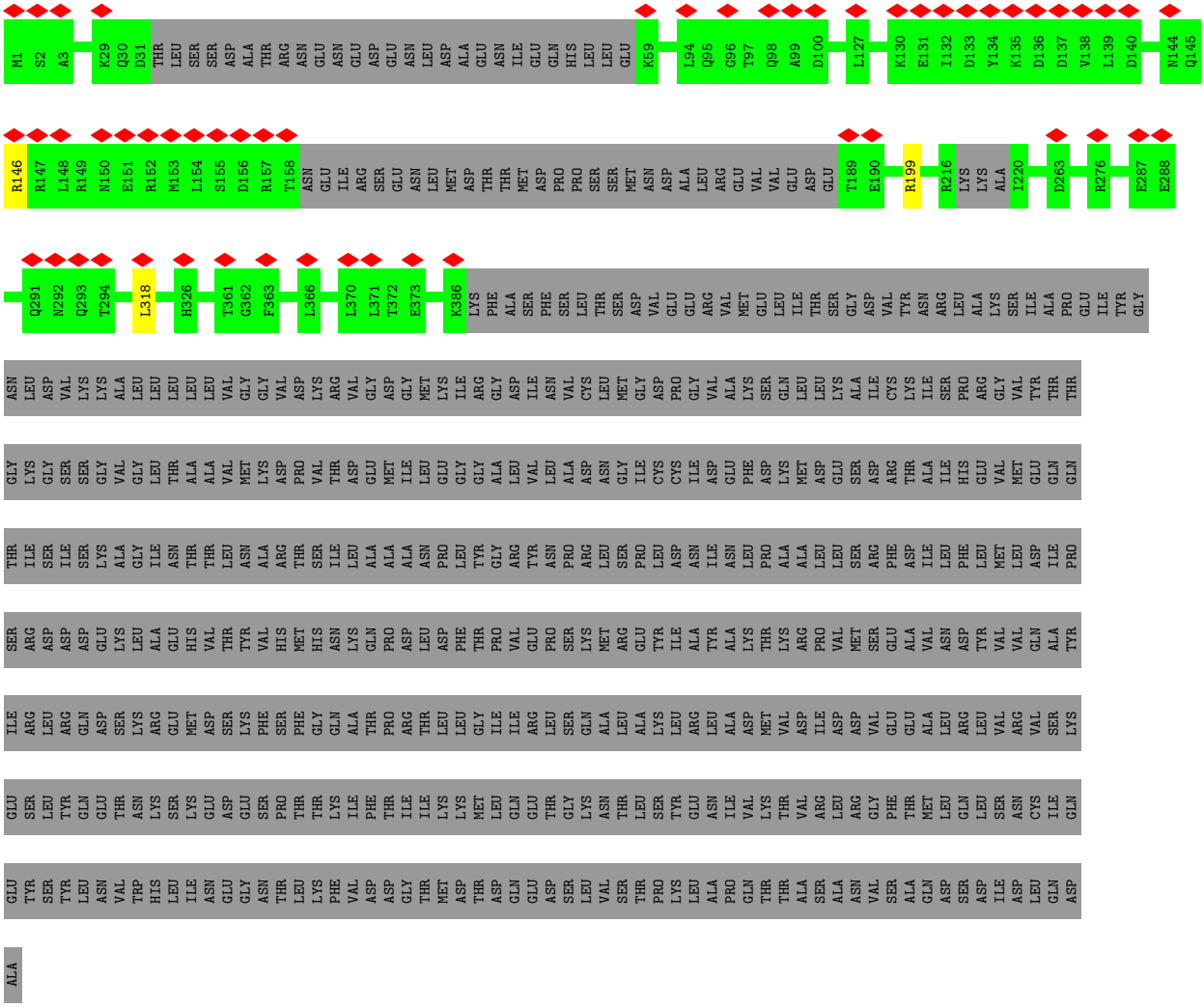
• Molecule 10: DNA replication licensing factor MCM6

Chain 6:



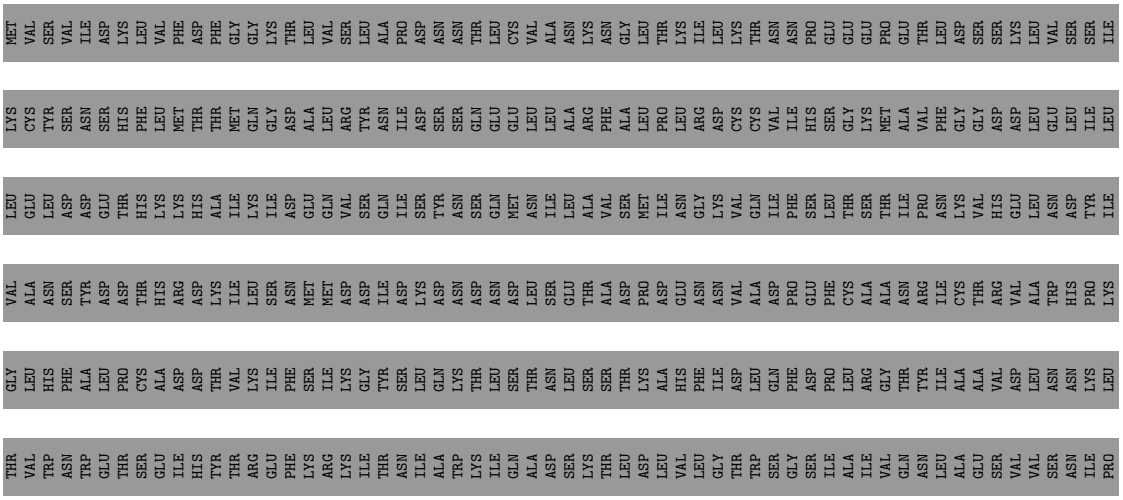
• Molecule 11: DNA replication licensing factor MCM7

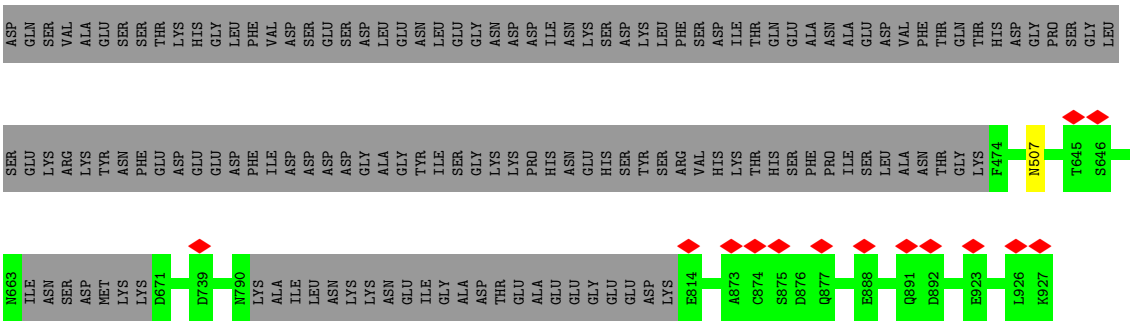
Chain 7:



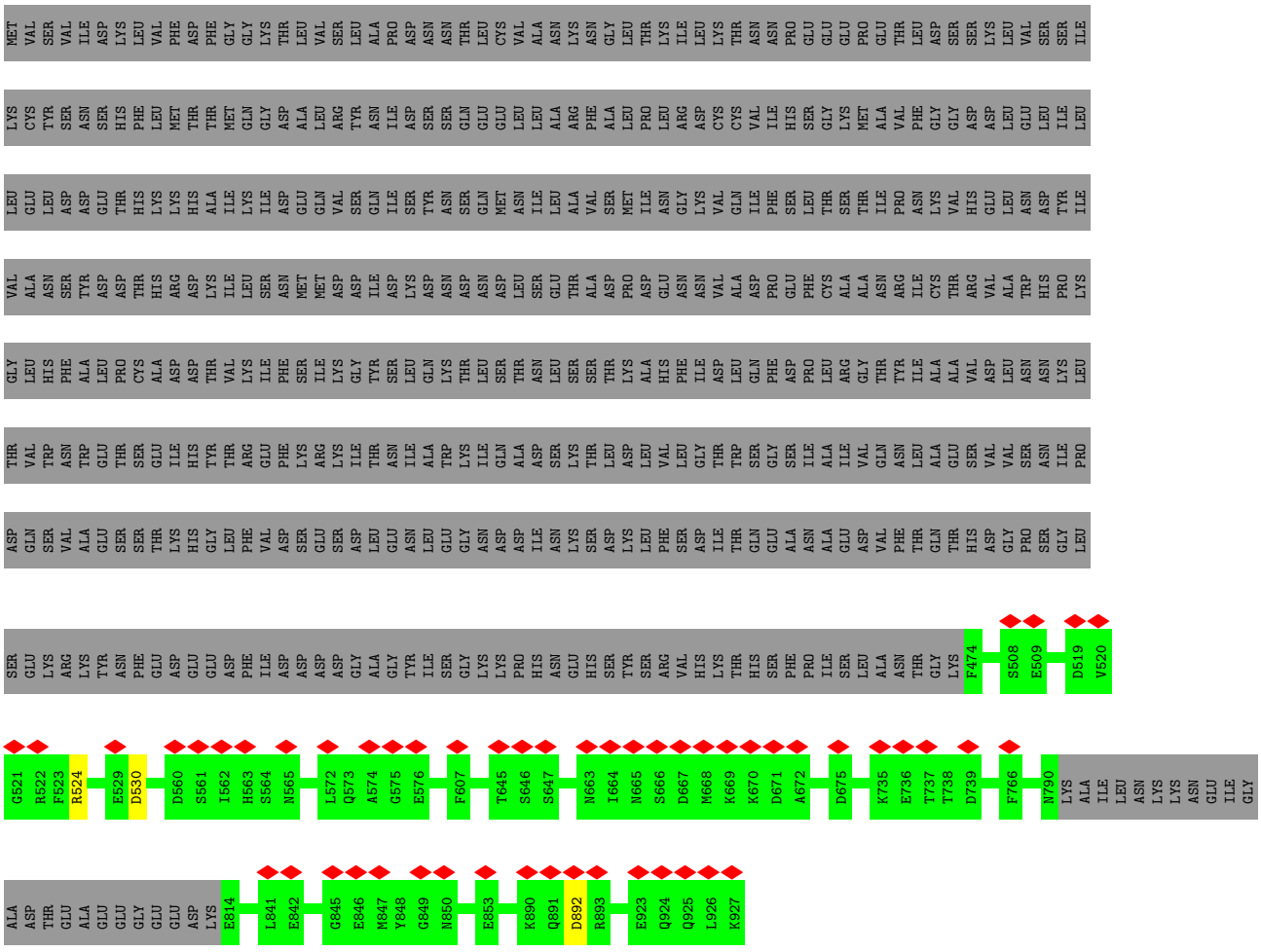
• Molecule 12: DNA polymerase alpha-binding protein

Chain E:

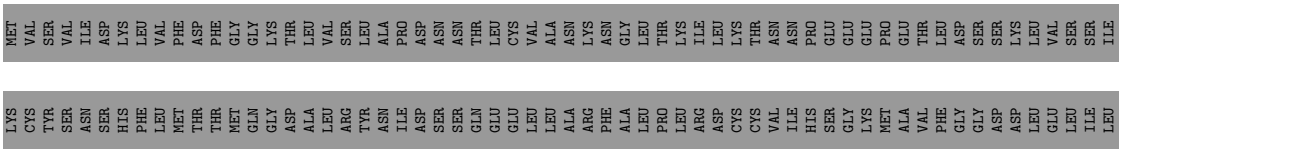


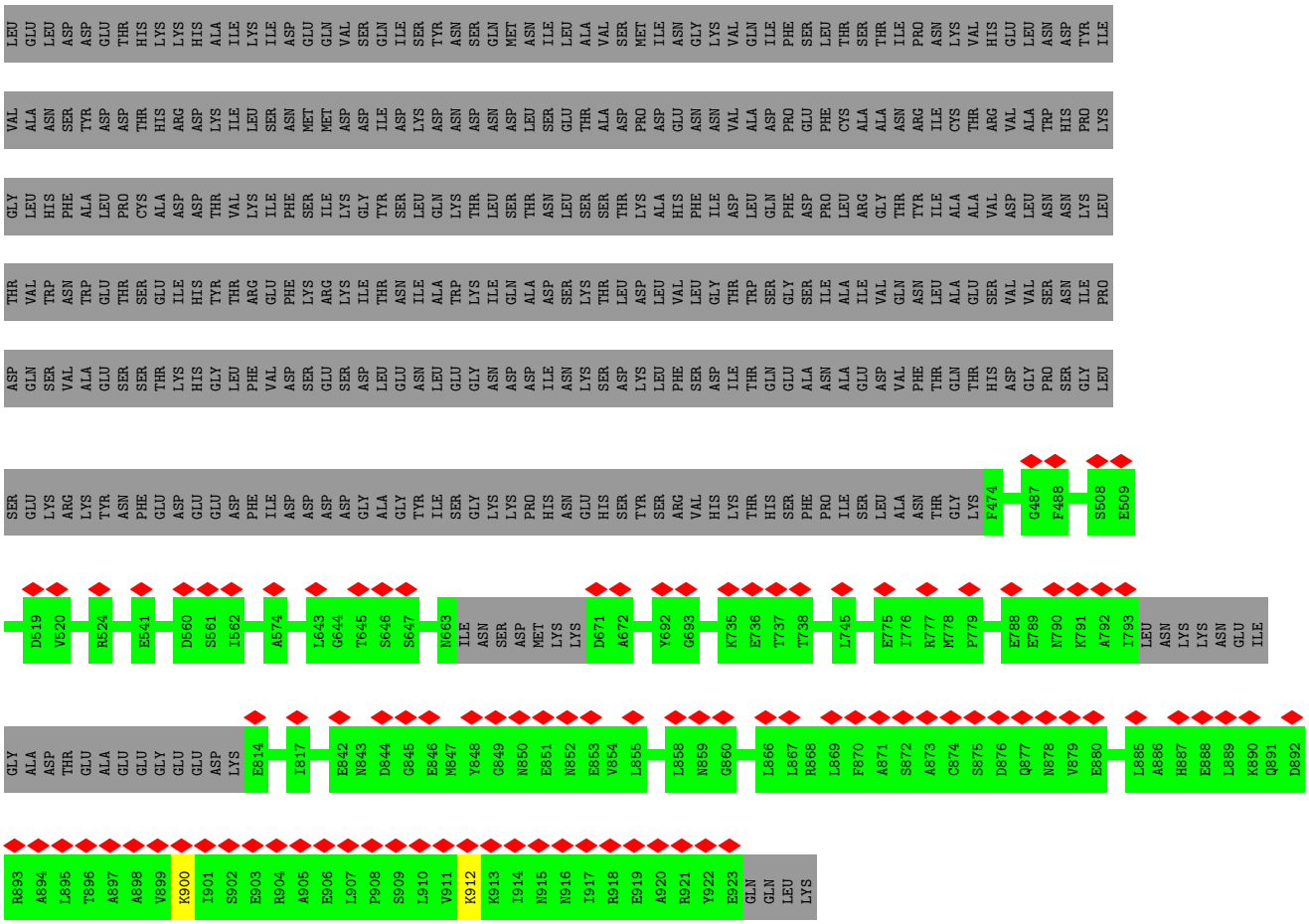


• Molecule 12: DNA polymerase alpha-binding protein



• Molecule 12: DNA polymerase alpha-binding protein




LEU GLU LEU ASP ASP GLU THR HIS LYS LYS HIS ALA ILE LYS ILE ASP GLU GLN VAL SER GLN ILE SER TYR ASN SER GLN MET ASN ILE LEU ALA VAL SER MET ILE ASN GLY LYS VAL GLN ILE PHE SER LEU THR SER THR ILE PRO ASN LYS VAL HIS GLU LEU ASN ASP TYR ILE
VAL ALA ASN SER TYR ASP ASP THR HIS ARG ASP LYS ILE LEU SER ASN MET MET ASP ASP ILE ASP LYS ASP ASN ASP ASP LEU SER GLU THR ALA ASP PRO ASP GLU ASN ASN VAL ALA ASP PRO GLU PHE CYS ALA ALA ASN ARG ILE CYS THR ARG VAL ALA TRP HIS PRO LYS
GLY LEU HIS PHE ALA LEU PRO CYS ALA ASP ASP THR VAL LYS ILE PHE SER ILE LYS GLY TYR SER LEU GLN LYS THR LEU SER THR ASN LEU SER SER THR LYS ALA HIS PHE ILE ASP LEU GLN PHE ASP PRO LEU ARG GLY THR TYR ILE ALA ILE ALA VAL GLN ASN LEU ASN LEU VAL LEU LYS ASN ASN LYS LEU
THR VAL TRP ASN TRP GLU THR SER GLU ILE HIS TYR THR ARG GLU PHE LYS ARG LYS ILE THR ASN ILE ALA TRP LYS ILE GLN ALA ASP SER LYS THR LEU ASP LEU VAL LEU GLY THR TRP SER GLY SER ILE ALA ILE VAL GLN ASN LEU ALA GLU SER VAL SER ASN ILE PRO
ASP GLN SER VAL ALA GLU SER SER THR LYS HIS GLY LEU PHE VAL ASP SER GLU SER ASP LEU GLU ASN LEU GLU GLY ASN ASP ASP ILE ASN LYS SER ASP LYS LEU PHE SER ASP ILE THR GLN GLU ASN ALA ASN ALA GLU ASP VAL PHE THR GLN THR HIS ASP GLY PRO SER GLY LEU
SER GLU LYS ARG LYS TYR ASN PHE GLU ASP GLU GLU ASP PHE ILE ASP ASP ASP ASP GLY ALA GLY TYR ILE SER GLY LYS LYS PRO HIS ASN GLU HIS SER TYR SER ARG VAL HIS LYS THR HIS SER PHE PRO ILE SER LEU ALA ASN THR GLY LYS F474 G487 F488 S508 E509
D519 V520 R524 E541 D560 S561 I562 A574 L643 G644 T645 S646 S647 N663 ILE ASN SER ASP MET LYS LYS D671 A672 Y692 G693 K735 E736 T737 T738 L745 E775 I776 R777 M778 P779 E788 E789 N790 K791 A792 I793 LEU ASN LYS LYS ASN GLU ILE
GLY ALA ASP THR GLU ALA GLU GLU GLY GLU GLU ASP LYS E814 I817 E842 N843 D844 G845 E846 M847 Y848 G849 N850 E851 N852 E853 V854 L855 L858 N859 G860 L866 L867 R868 L869 F870 A871 S872 A873 C874 S875 D876 Q877 N878 V879 E880 L885 A886 H887 E888 L889 K890 Q891 D892
R893 A894 L895 T896 A897 A898 V899 K900 I901 S902 E903 R904 A905 E906 L907 P908 S909 L910 V911 K912 K913 I914 N915 N916 I917 R918 E919 A920 R921 Y922 E923 GLN GLN LEU LYS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	200491	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 ($e^-/Å^2$)	50	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	0.230	Depositor
Minimum map value	-0.128	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.034	Depositor
Map size (Å)	300.72, 300.72, 300.72	wwPDB
Map dimensions	280, 280, 280	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.074, 1.074, 1.074	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	0.40	0/1718	0.58	0/2314
2	B	0.47	0/1589	0.66	0/2150
3	C	0.41	0/1320	0.58	1/1784 (0.1%)
4	D	0.44	0/1923	0.67	2/2594 (0.1%)
5	c	0.43	0/4563	0.59	1/6173 (0.0%)
6	2	0.38	0/2063	0.60	0/2794
7	3	0.40	0/2189	0.63	2/2975 (0.1%)
8	4	0.31	0/2336	0.59	1/3155 (0.0%)
9	5	0.46	0/2035	0.66	3/2752 (0.1%)
10	6	0.32	0/2114	0.58	1/2862 (0.0%)
11	7	0.33	0/2660	0.60	1/3595 (0.0%)
12	E	0.46	0/3493	0.55	1/4730 (0.0%)
12	F	0.34	0/3558	0.54	2/4817 (0.0%)
12	G	0.35	0/3487	0.56	0/4724
All	All	0.39	0/35048	0.59	15/47419 (0.0%)

There are no bond length outliers.

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	4	373	ARG	N-CA-C	-7.67	90.28	111.00
10	6	399	GLY	N-CA-C	7.62	132.16	113.10
9	5	264	LEU	N-CA-C	-7.01	92.08	111.00
9	5	331	LEU	CA-CB-CG	6.88	131.12	115.30
7	3	325	THR	CA-C-O	-6.85	105.72	120.10

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	206/208 (99%)	180 (87%)	25 (12%)	1 (0%)	29	66
2	B	182/213 (85%)	149 (82%)	32 (18%)	1 (0%)	29	66
3	C	151/194 (78%)	136 (90%)	14 (9%)	1 (1%)	22	60
4	D	222/294 (76%)	197 (89%)	24 (11%)	1 (0%)	29	66
5	c	543/650 (84%)	482 (89%)	61 (11%)	0	100	100
6	2	258/868 (30%)	218 (84%)	40 (16%)	0	100	100
7	3	262/971 (27%)	221 (84%)	40 (15%)	1 (0%)	34	70
8	4	281/933 (30%)	247 (88%)	31 (11%)	3 (1%)	14	51
9	5	242/775 (31%)	214 (88%)	28 (12%)	0	100	100
10	6	267/1017 (26%)	221 (83%)	45 (17%)	1 (0%)	34	70
11	7	318/845 (38%)	282 (89%)	36 (11%)	0	100	100
12	E	418/927 (45%)	392 (94%)	26 (6%)	0	100	100
12	F	428/927 (46%)	404 (94%)	24 (6%)	0	100	100
12	G	418/927 (45%)	398 (95%)	20 (5%)	0	100	100
All	All	4196/9749 (43%)	3741 (89%)	446 (11%)	9 (0%)	50	79

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	114	GLU
8	4	375	ASP
8	4	378	GLU
7	3	281	ASP
3	C	101	ASN

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	193/193 (100%)	193 (100%)	0	100 100
2	B	176/198 (89%)	175 (99%)	1 (1%)	86 92
3	C	144/173 (83%)	141 (98%)	3 (2%)	53 74
4	D	221/279 (79%)	220 (100%)	1 (0%)	88 94
5	c	499/586 (85%)	496 (99%)	3 (1%)	86 92
6	2	212/770 (28%)	212 (100%)	0	100 100
7	3	236/835 (28%)	234 (99%)	2 (1%)	81 89
8	4	260/848 (31%)	259 (100%)	1 (0%)	91 95
9	5	237/688 (34%)	236 (100%)	1 (0%)	91 95
10	6	210/886 (24%)	209 (100%)	1 (0%)	88 94
11	7	296/753 (39%)	294 (99%)	2 (1%)	84 91
12	E	376/825 (46%)	376 (100%)	0	100 100
12	F	384/825 (46%)	383 (100%)	1 (0%)	92 96
12	G	375/825 (46%)	373 (100%)	2 (0%)	88 94
All	All	3819/8684 (44%)	3801 (100%)	18 (0%)	89 94

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	7	199	ARG
12	G	912	LYS
12	G	900	LYS
7	3	170	THR
11	7	146	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
12	F	553	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
12	F	877	GLN
7	3	164	HIS
6	2	294	HIS
12	F	915	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 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

There are no chain breaks in this entry.

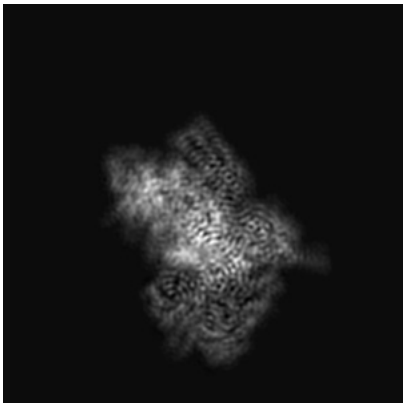
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-20471. These allow visual inspection of the internal detail of the map and identification of artifacts.

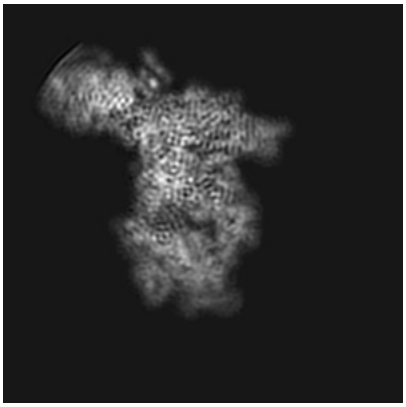
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

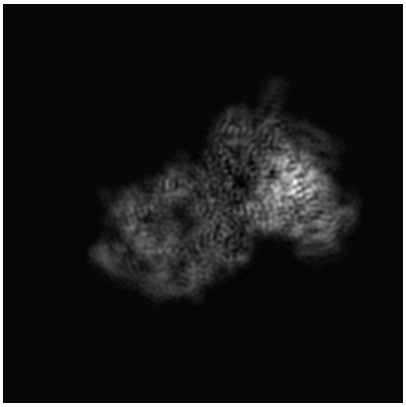
6.1.1 Primary map



X



Y

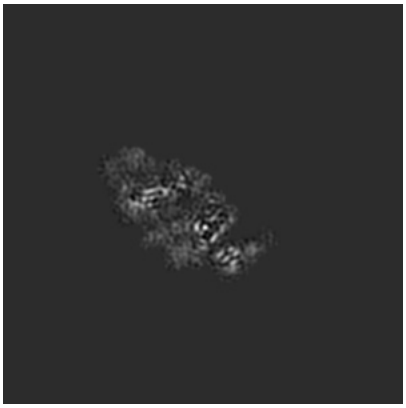


Z

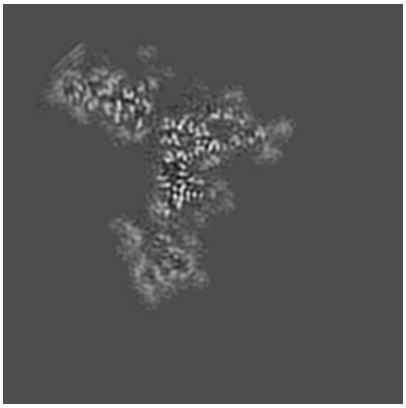
The images above show the map projected in three orthogonal directions.

6.2 Central slices

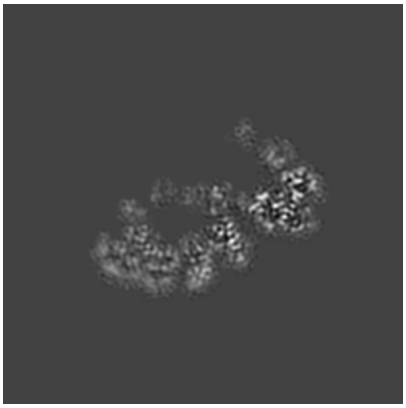
6.2.1 Primary map



X Index: 140



Y Index: 140

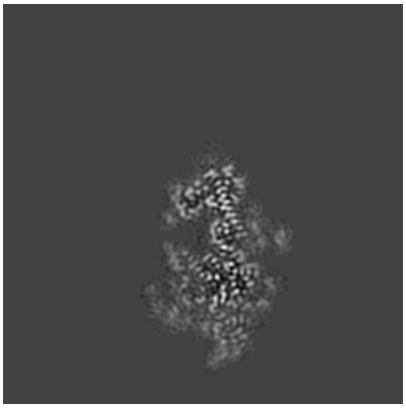


Z Index: 140

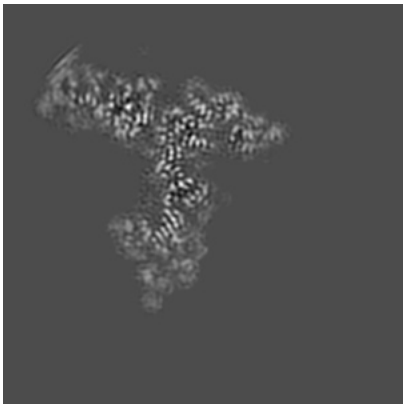
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

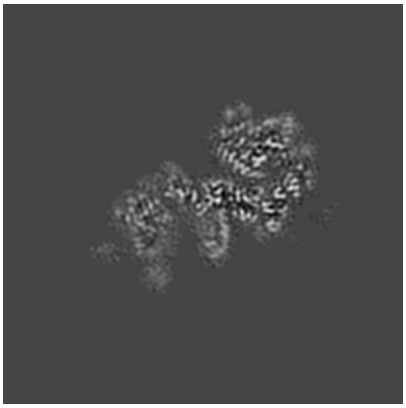
6.3.1 Primary map



X Index: 204



Y Index: 146



Z Index: 121

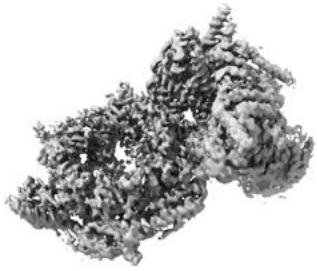
The images above show the largest variance slices of the map in three orthogonal directions.

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

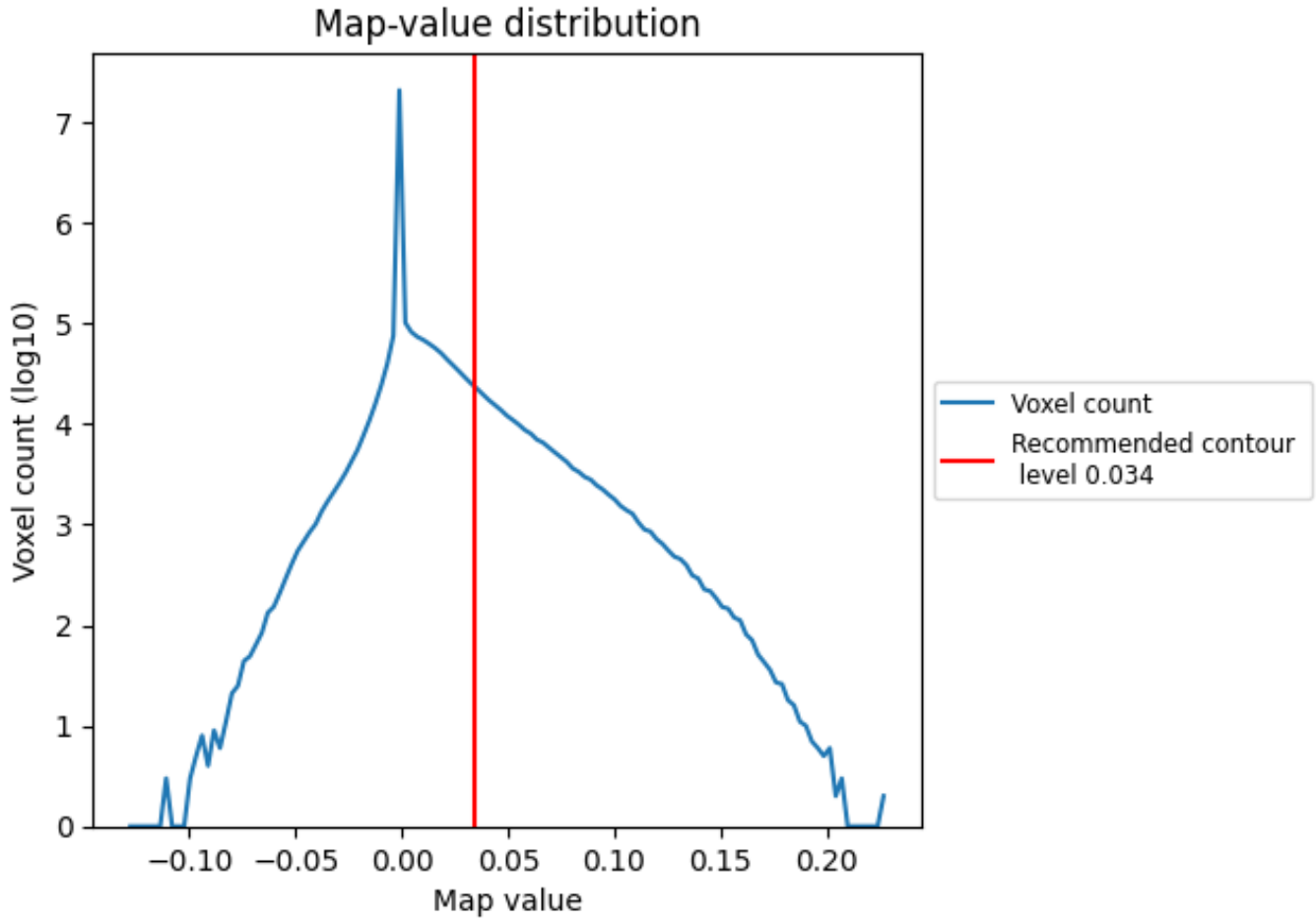
6.5 Mask visualisation (i)

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

7 Map analysis

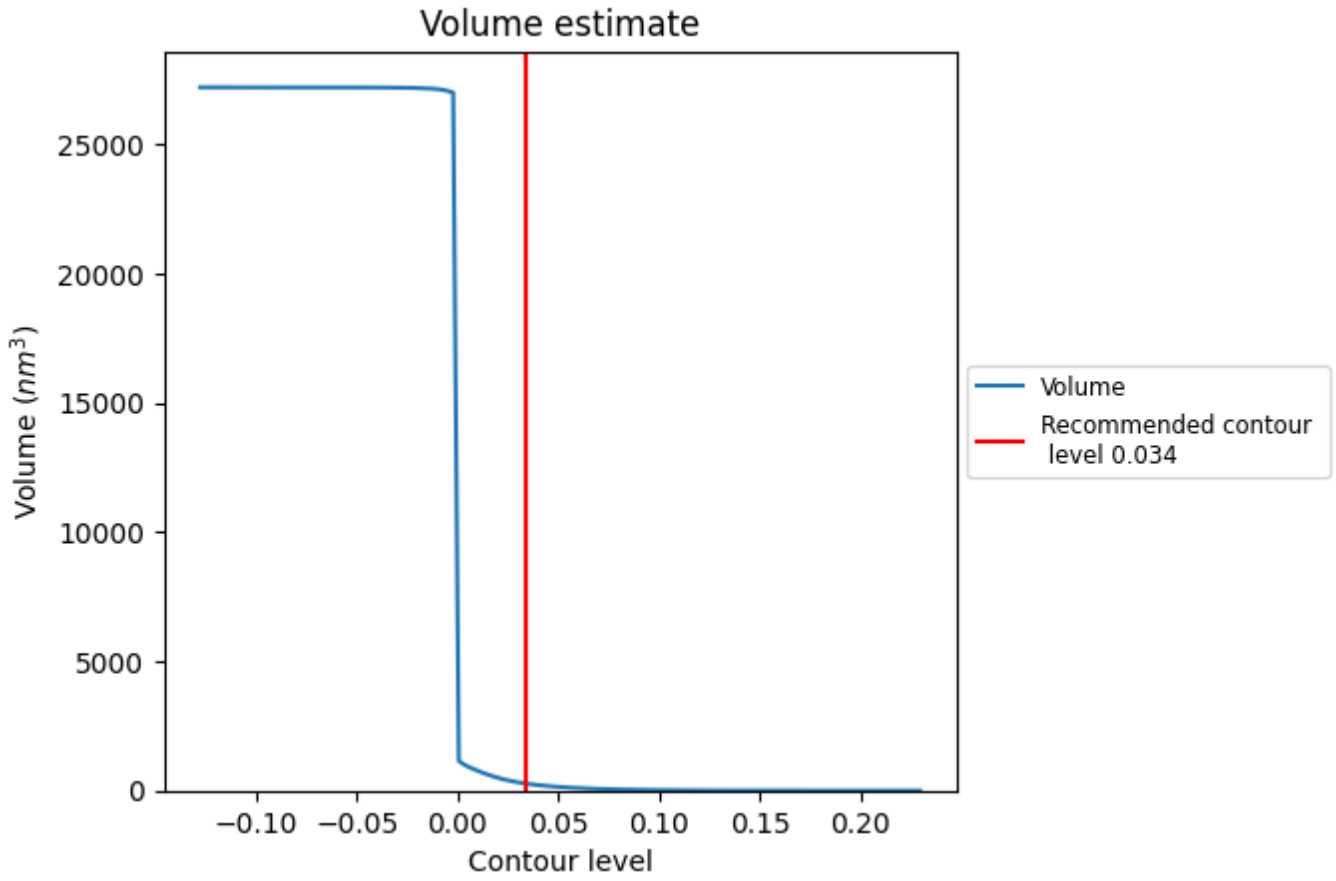
This section contains the results of statistical analysis of the map.

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.

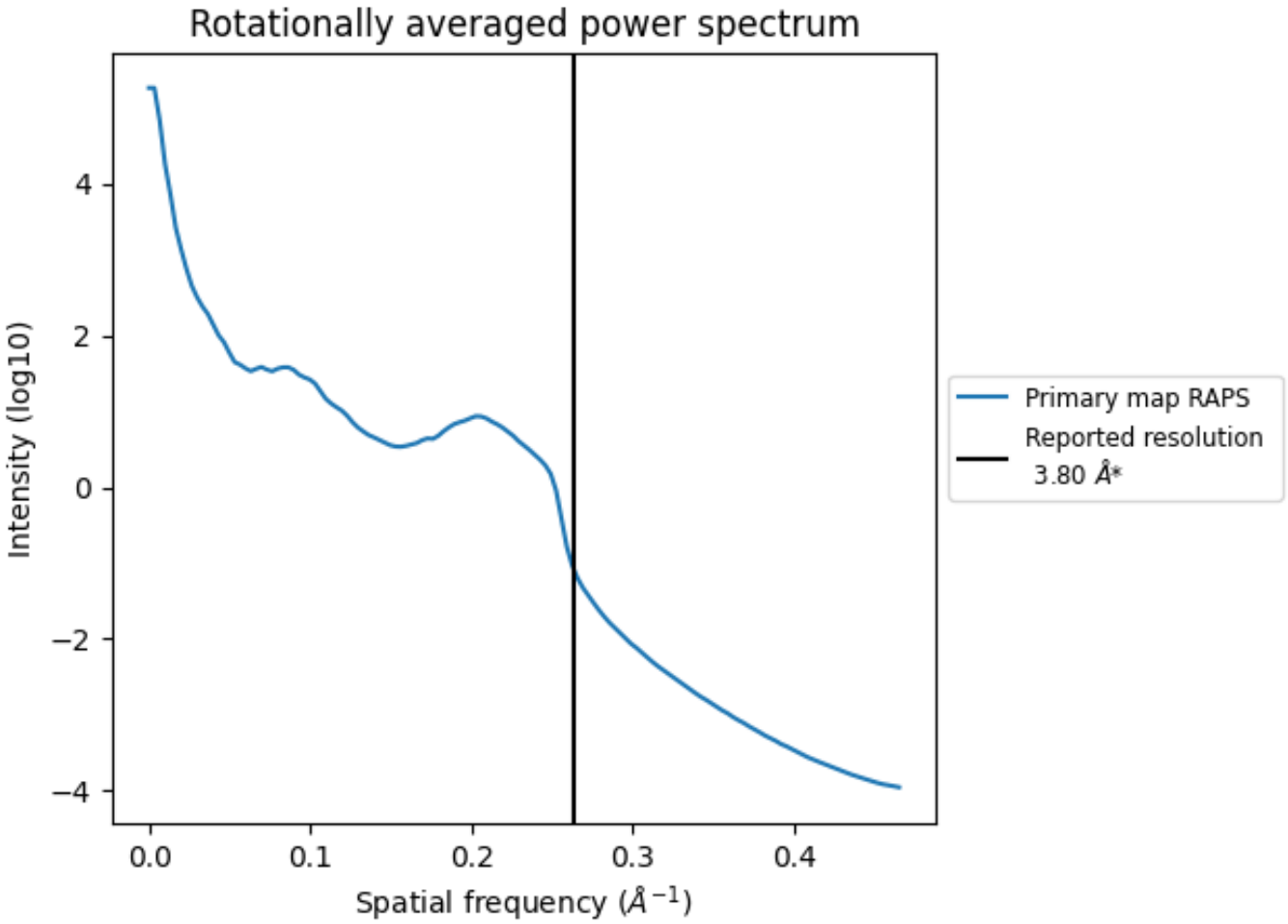
7.2 Volume estimate (i)



The volume at the recommended contour level is 270 nm^3; this corresponds to an approximate mass of 244 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 0.263 Å^{-1}

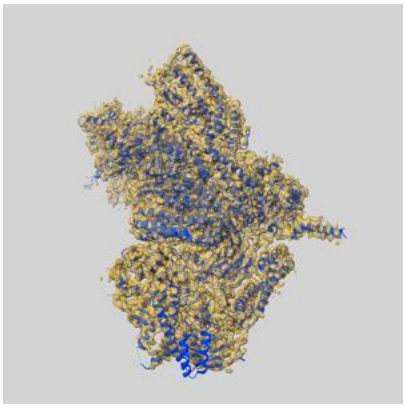
8 Fourier-Shell correlation

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

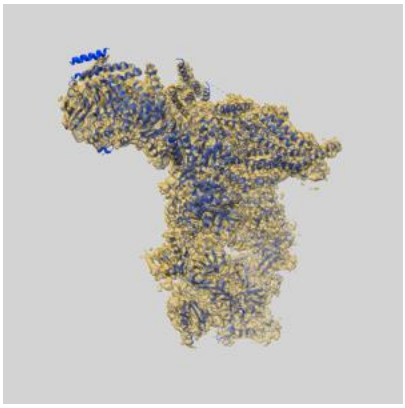
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-20471 and PDB model 6PTJ. Per-residue inclusion information can be found in section 3 on page 6.

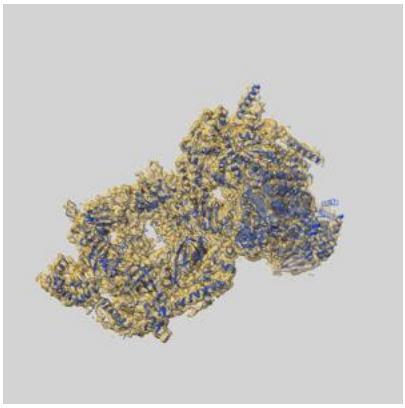
9.1 Map-model overlay



X



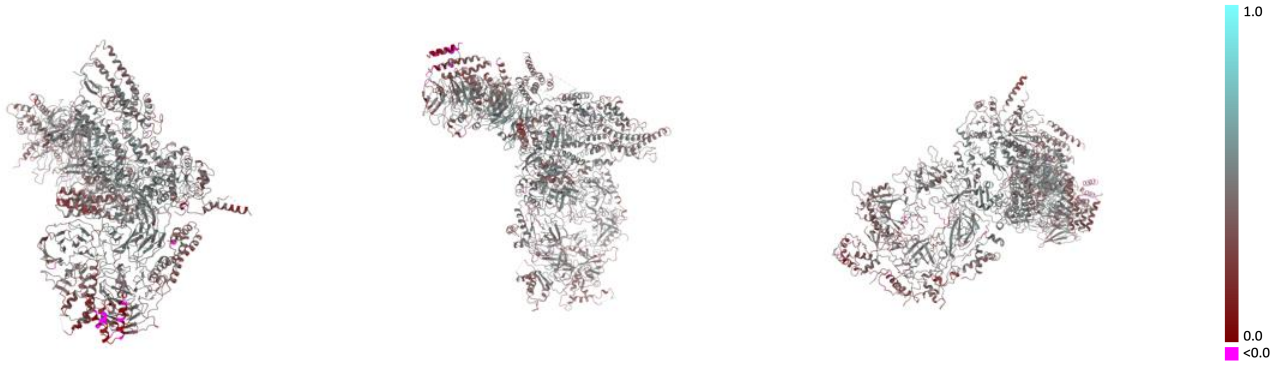
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.034 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

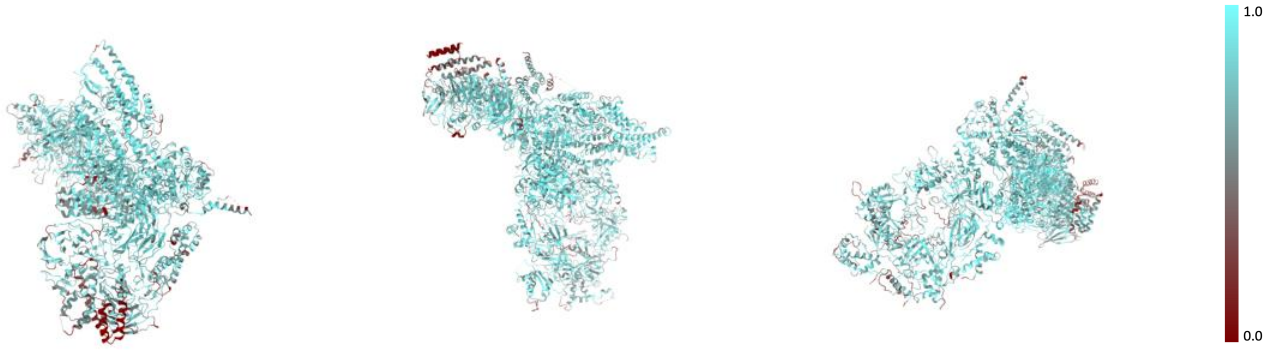
9.2 Q-score mapped to coordinate model (i)



X Y Z

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

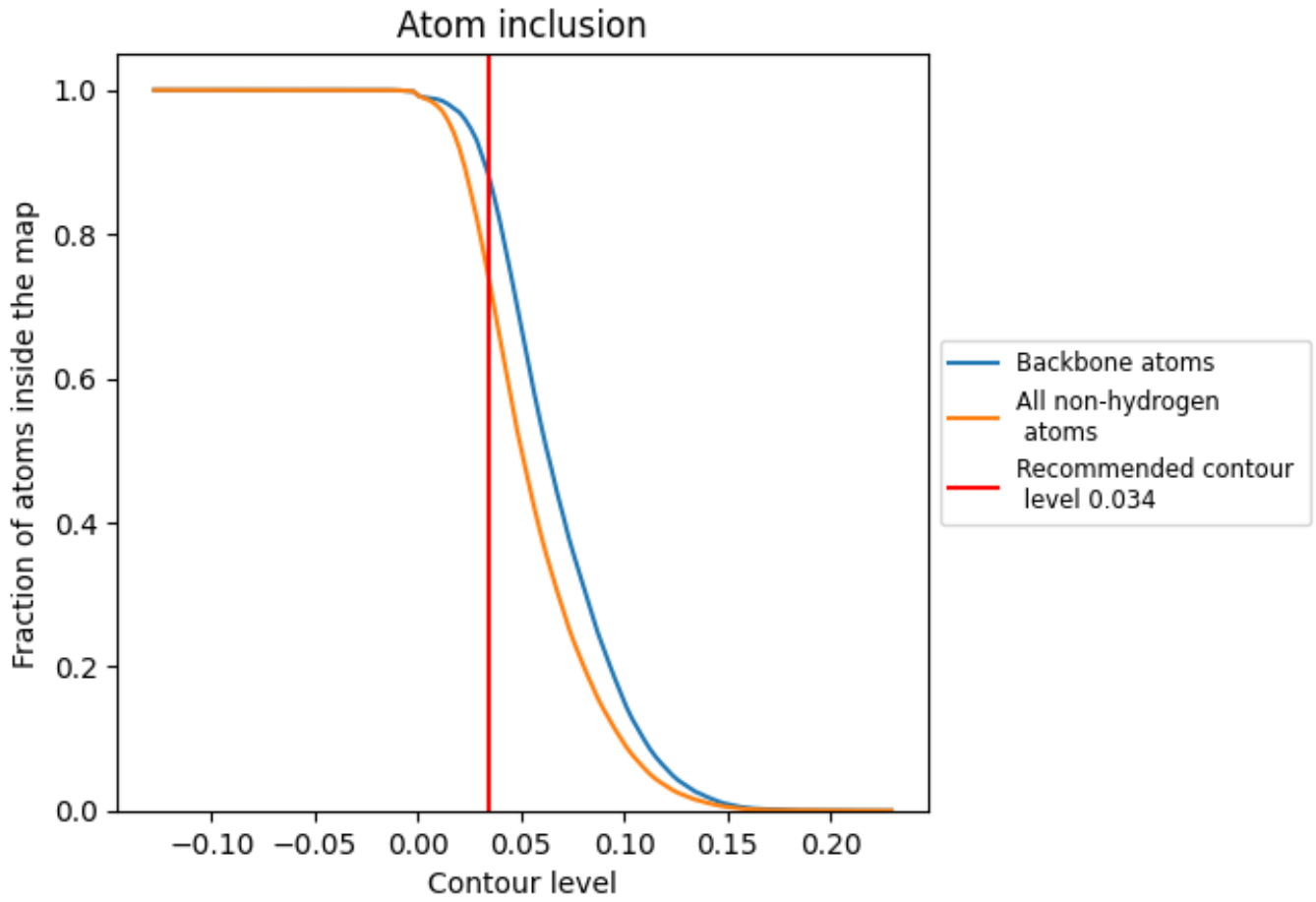
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.034).

9.4 Atom inclusion (i)



At the recommended contour level, 88% of all backbone atoms, 74% 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 (0.034) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7444	0.4140
2	0.7353	0.4080
3	0.7820	0.4310
4	0.6833	0.3360
5	0.8303	0.4740
6	0.7023	0.3820
7	0.6646	0.3640
A	0.7462	0.4030
B	0.8416	0.4790
C	0.8434	0.4660
D	0.7885	0.4380
E	0.8069	0.4620
F	0.6927	0.3990
G	0.6117	0.3620
c	0.8003	0.4380