wwPDB EM Validation Summary Report

Nov 20, 2022 – 03:25 PM EST

PDB ID	:	3JC5
EMDB ID	:	EMD-6535
Title	:	Structure of the eukaryotic replicative CMG helicase and pumpjack motion
Authors	:	Li, H.; Bai, L.; Yuan, Z.; Sun, J.; Georgescu, R.E.; Liu, J.; O'Donnell, M.E.
Deposited on	:	2015-11-24
Resolution	:	4.70 Å(reported)
Based on initial model	:	2Q9Q

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:

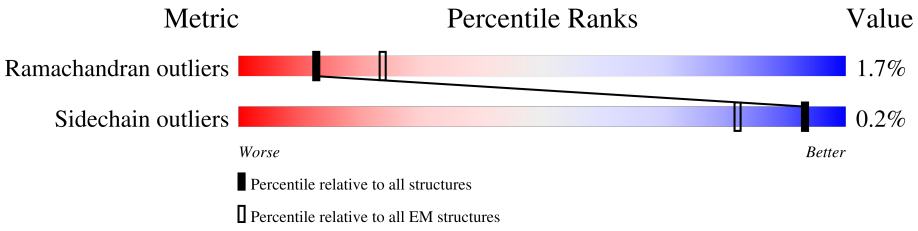
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 4.70 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.



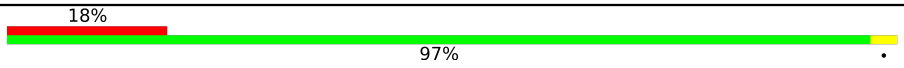
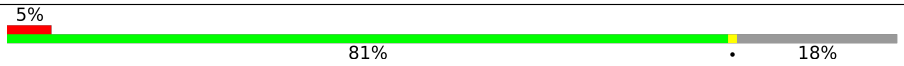
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	2	868	25% 66% · 33%
2	3	971	18% 58% · 39%
3	4	933	31% 59% · 39%
4	5	775	25% 80% · 16%
5	6	1017	29% 64% · 34%
6	7	845	30% 76% · 23%
7	c	650	11% 84% · 15%
8	D	294	· 73% · 25%
9	B	213	· 85% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	A	208	
11	C	194	

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 40041 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 licensing factor MCM2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	2	584	4600	2904	819	861	16	0	0

- Molecule 2 is a protein called DNA replication licensing factor MCM3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	3	588	4613	2909	820	871	13	0	0

- Molecule 3 is a protein called DNA replication licensing factor MCM4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	4	569	4516	2842	783	864	27	0	0

- Molecule 4 is a protein called Minichromosome maintenance protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	5	653	5171	3251	896	1001	23	0	0

- Molecule 5 is a protein called DNA replication licensing factor MCM6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	6	671	5211	3291	916	981	23	0	0

- Molecule 6 is a protein called DNA replication licensing factor MCM7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	7	652	5148	3249	895	977	27	0	0

- Molecule 7 is a protein called Cell division control protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	c	553	Total	C	N	O	S	0	0
			4470	2852	759	846	13		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
c	22	ALA	HIS	CONFLICT	UNP Q08032
c	155	GLU	GLN	CONFLICT	UNP Q08032
c	551	THR	TRP	CONFLICT	UNP Q08032

- Molecule 8 is a protein called DNA replication complex GINS protein SLD5.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	D	221	Total	C	N	O	S	0	0
			1820	1159	300	348	13		

- Molecule 9 is a protein called DNA replication complex GINS protein PSF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	B	181	Total	C	N	O	S	0	0
			1513	978	261	270	4		

- Molecule 10 is a protein called DNA replication complex GINS protein PSF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	A	208	Total	C	N	O	S	0	0
			1691	1062	287	332	10		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	161	ALA	VAL	CONFLICT	UNP Q12488
A	192	GLN	ARG	CONFLICT	UNP Q12488
A	207	LEU	LYS	CONFLICT	UNP Q12488

- Molecule 11 is a protein called DNA replication complex GINS protein PSF3.

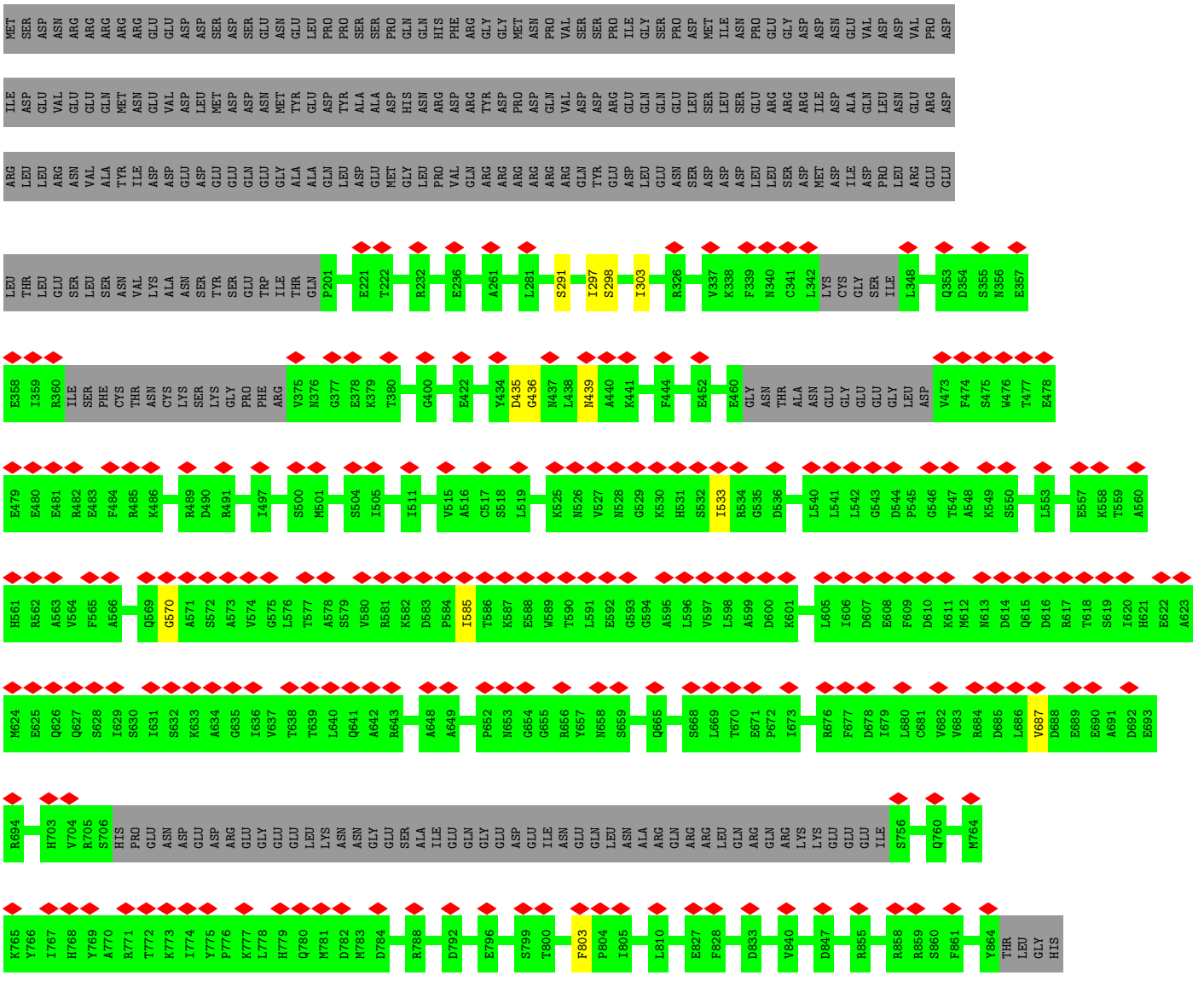
Mol	Chain	Residues	Atoms					AltConf	Trace
11	C	159	Total	C	N	O	S	0	0
			1288	843	207	232	6		

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.

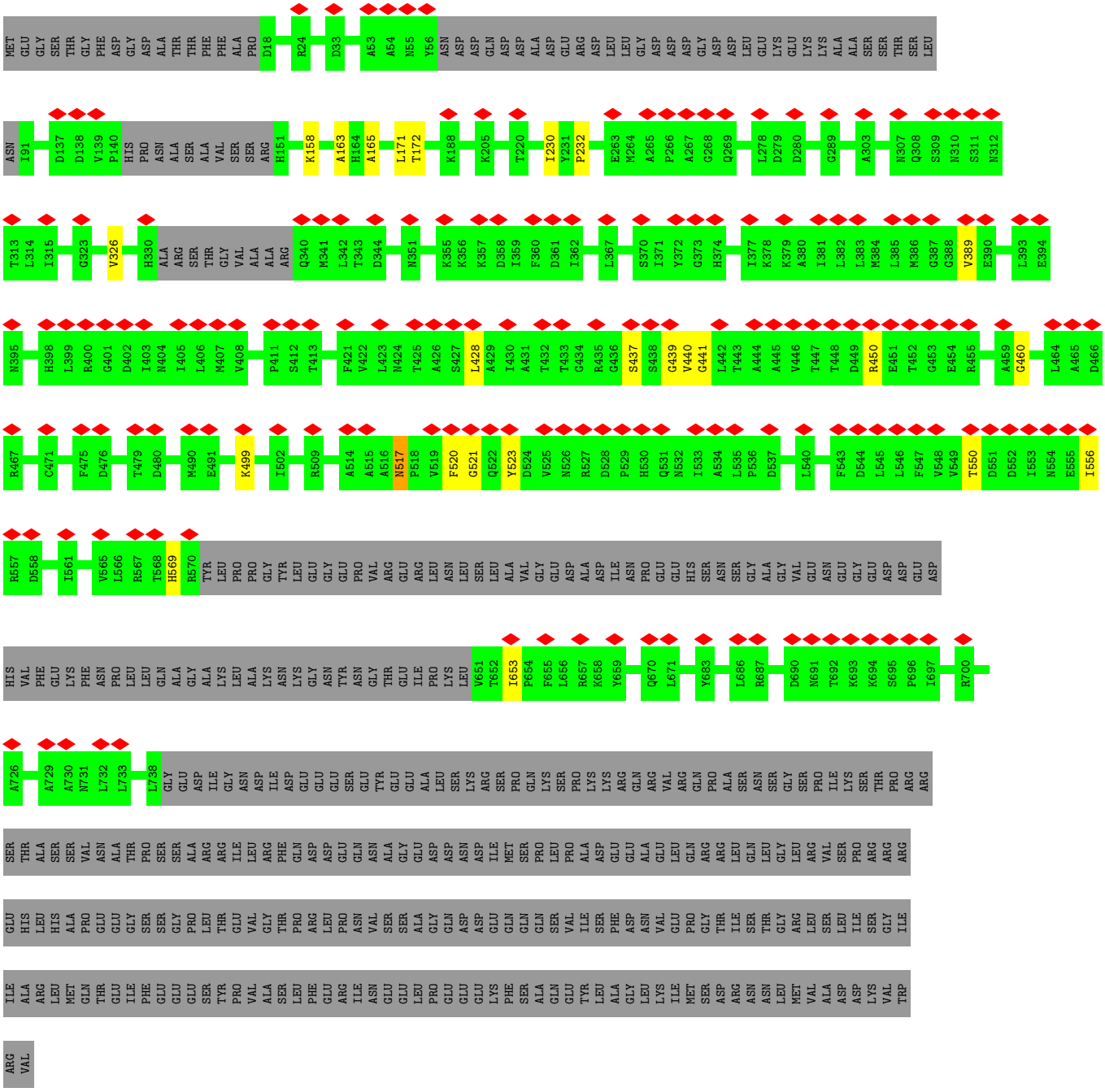
• Molecule 1: DNA replication licensing factor MCM2

Chain 2:



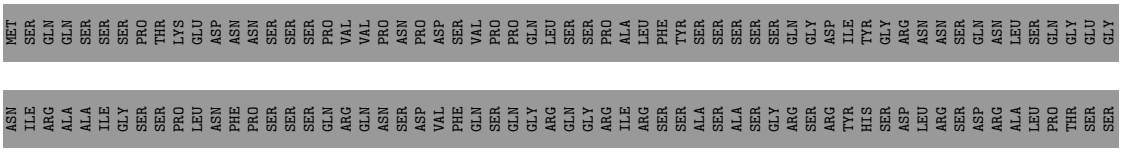
• Molecule 2: DNA replication licensing factor MCM3

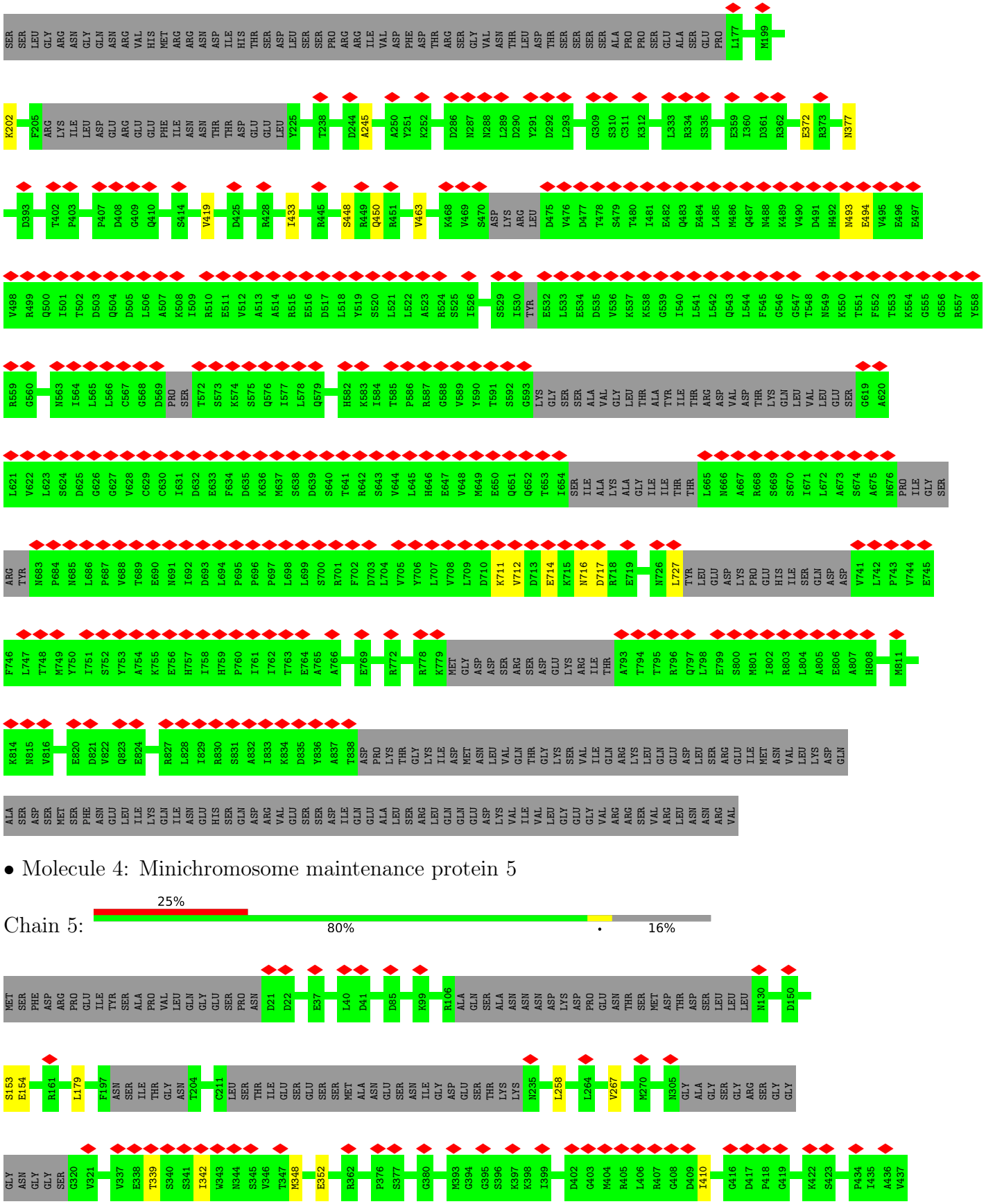
Chain 3:



• Molecule 3: DNA replication licensing factor MCM4

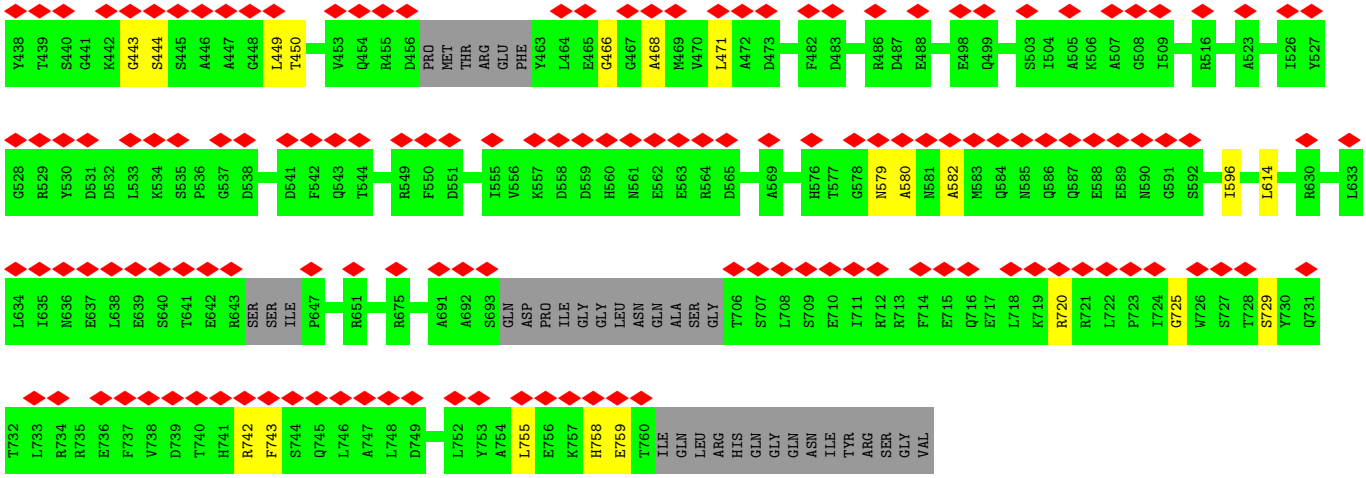
Chain 4:





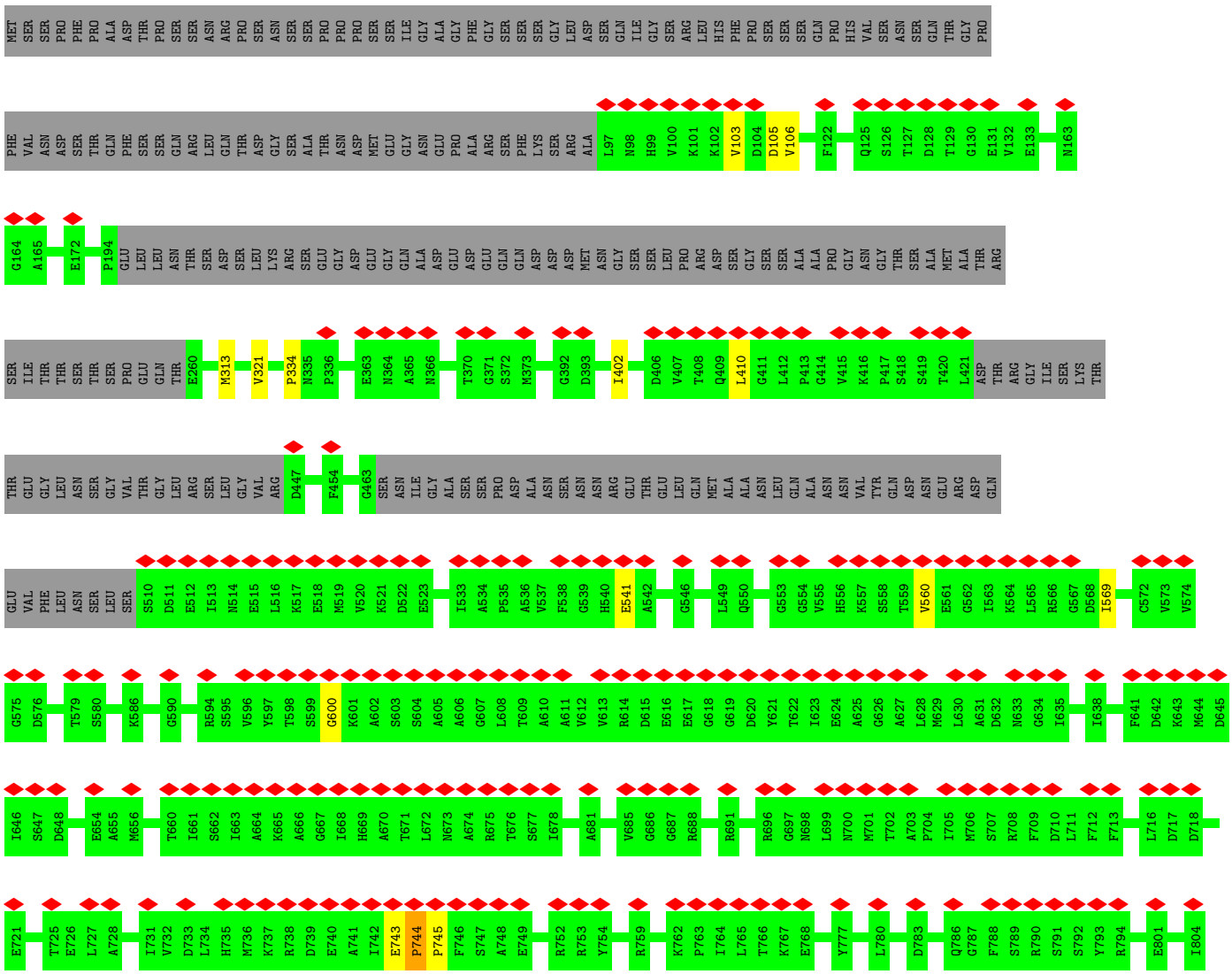
• Molecule 4: Minichromosome maintenance protein 5

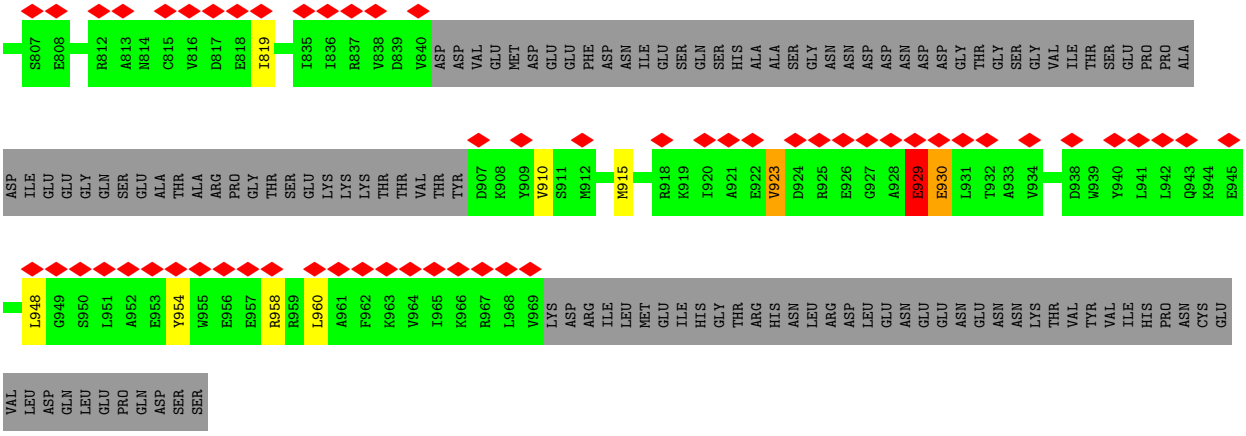
Chain 5:



• Molecule 5: DNA replication licensing factor MCM6

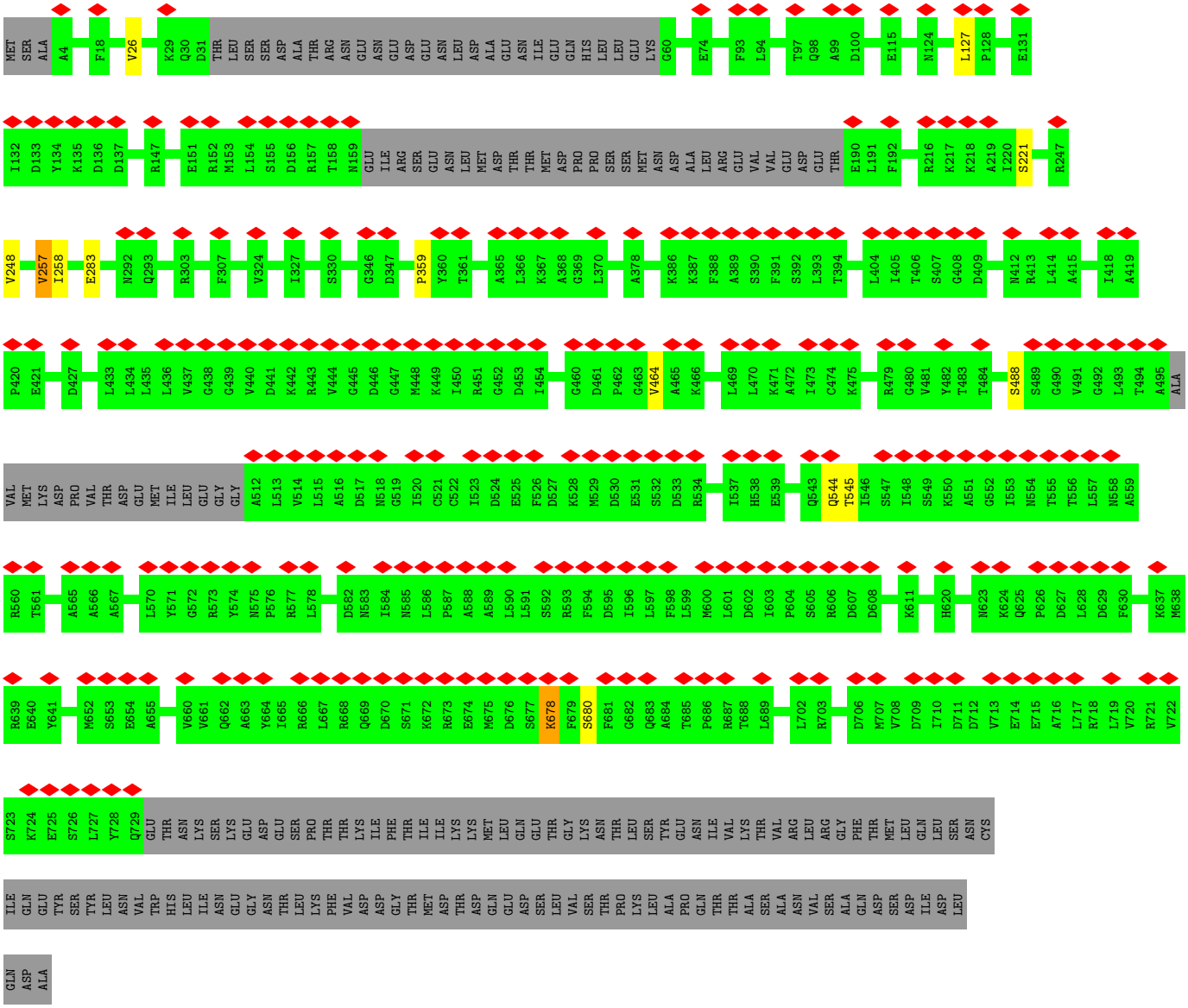
Chain 6:





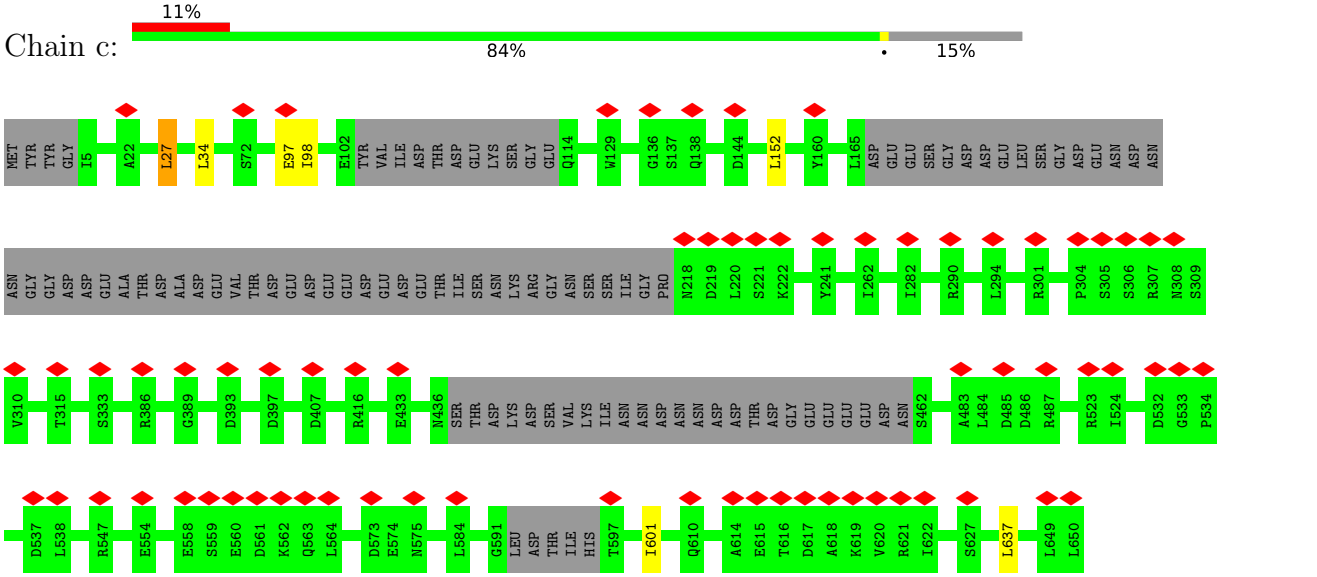
• Molecule 6: DNA replication licensing factor MCM7

Chain 7:



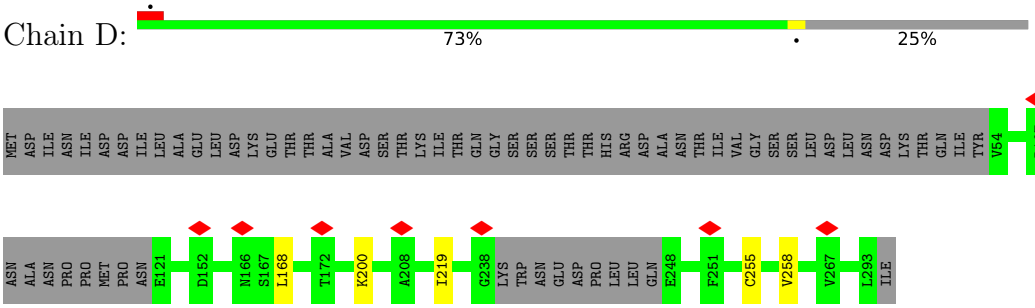
• Molecule 7: Cell division control protein 45

Chain c:



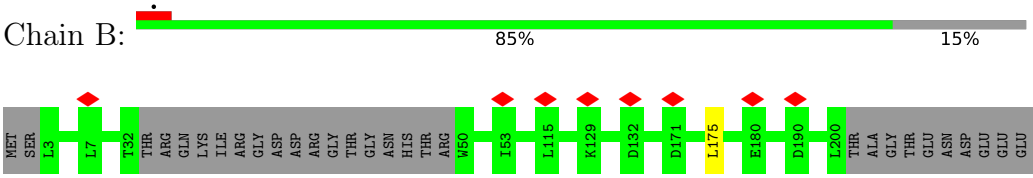
• Molecule 8: DNA replication complex GINS protein SLD5

Chain D:



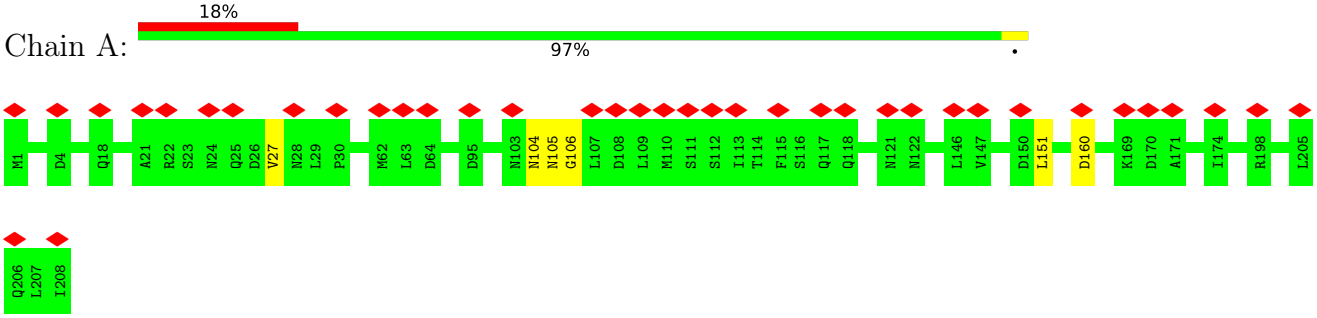
• Molecule 9: DNA replication complex GINS protein PSF2

Chain B:



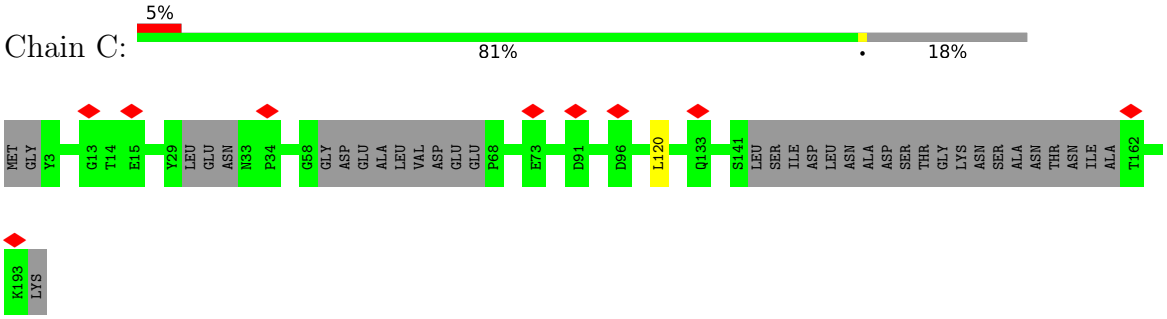
• Molecule 10: DNA replication complex GINS protein PSF1

Chain A:



• Molecule 11: DNA replication complex GINS protein PSF3

Chain C:


5%
81%
18%
MET GLY Y3 G13 T14 E15 Y29 LEU GLU ASN N33 P34 G58 GLY ASP GLU ALA LEU VAL ASP GLU GLU P68 E73 D91 D96 L120 Q133 S141 LEU SER ILE ASP LEU ASN ALA ASP SER THR GLY LYS ASN SER ALA ASN THR ASN ILE ALA T162
K193 LYS

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	178530	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	CTFFIND4	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	49505	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.064	Depositor
Minimum map value	-0.029	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.018	Depositor
Map size (Å)	258.56, 258.56, 258.56	wwPDB
Map dimensions	256, 256, 256	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.01, 1.01, 1.01	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	2	0.41	0/4677	0.62	2/6318 (0.0%)
2	3	0.48	1/4691 (0.0%)	0.66	5/6360 (0.1%)
3	4	0.40	0/4574	0.65	1/6172 (0.0%)
4	5	0.48	1/5242 (0.0%)	0.73	7/7075 (0.1%)
5	6	0.48	1/5289 (0.0%)	0.78	13/7139 (0.2%)
6	7	0.42	0/5228	0.64	1/7062 (0.0%)
7	c	0.39	0/4548	0.63	2/6152 (0.0%)
8	D	0.43	0/1853	0.64	0/2500
9	B	0.43	0/1545	0.69	0/2092
10	A	0.38	0/1713	0.62	0/2309
11	C	0.46	0/1320	0.64	1/1784 (0.1%)
All	All	0.44	3/40680 (0.0%)	0.67	32/54963 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	2	0	1
2	3	0	3
3	4	0	5
5	6	0	6
6	7	0	7
7	c	0	1
8	D	0	2
10	A	0	4
All	All	0	29

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	3	232	PRO	N-CD	16.13	1.70	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	6	929	GLU	C-O	7.76	1.38	1.23
4	5	720	ARG	C-N	-7.10	1.17	1.34

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	6	954	TYR	CB-CG-CD2	-16.55	111.07	121.00
4	5	742	ARG	CA-CB-CG	10.90	137.38	113.40
5	6	954	TYR	CB-CG-CD1	9.43	126.66	121.00
2	3	550	THR	N-CA-C	-9.05	86.57	111.00
4	5	742	ARG	N-CA-C	8.60	134.21	111.00

There are no chirality outliers.

5 of 29 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	2	803	PHE	Peptide
2	3	163	ALA	Peptide
2	3	165	ALA	Peptide
2	3	428	LEU	Peptide
3	4	202	LYS	Peptide

5.2 Too-close contacts (i)

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	2	574/868 (66%)	521 (91%)	44 (8%)	9 (2%)	9 45
2	3	578/971 (60%)	503 (87%)	57 (10%)	18 (3%)	4 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	4	547/933 (59%)	475 (87%)	61 (11%)	11 (2%)	7 40
4	5	637/775 (82%)	555 (87%)	61 (10%)	21 (3%)	4 29
5	6	661/1017 (65%)	591 (89%)	58 (9%)	12 (2%)	8 41
6	7	642/845 (76%)	561 (87%)	73 (11%)	8 (1%)	13 50
7	c	543/650 (84%)	496 (91%)	45 (8%)	2 (0%)	34 72
8	D	215/294 (73%)	197 (92%)	16 (7%)	2 (1%)	17 56
9	B	177/213 (83%)	163 (92%)	14 (8%)	0	100 100
10	A	206/208 (99%)	179 (87%)	26 (13%)	1 (0%)	29 68
11	C	151/194 (78%)	144 (95%)	7 (5%)	0	100 100
All	All	4931/6968 (71%)	4385 (89%)	462 (9%)	84 (2%)	13 43

5 of 84 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	2	291	SER
2	3	499	LYS
2	3	517	ASN
2	3	569	HIS
3	4	450	GLN

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	2	503/770 (65%)	503 (100%)	0	100 100
2	3	511/835 (61%)	511 (100%)	0	100 100
3	4	509/848 (60%)	509 (100%)	0	100 100
4	5	588/688 (86%)	586 (100%)	2 (0%)	92 95
5	6	547/886 (62%)	547 (100%)	0	100 100
6	7	576/753 (76%)	576 (100%)	0	100 100
7	c	498/585 (85%)	495 (99%)	3 (1%)	86 92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	D	213/279 (76%)	212 (100%)	1 (0%)	88 93
9	B	171/198 (86%)	170 (99%)	1 (1%)	86 92
10	A	192/192 (100%)	191 (100%)	1 (0%)	88 93
11	C	144/173 (83%)	144 (100%)	0	100 100
All	All	4452/6207 (72%)	4444 (100%)	8 (0%)	93 96

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	A	151	LEU
9	B	175	LEU
7	c	152	LEU
7	c	34	LEU
8	D	168	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
4	5	259	GLN
9	B	62	ASN
5	6	653	HIS
8	D	110	ASN
10	A	175	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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
6	7	1
3	4	1
4	5	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	7	386:LYS	C	387:LYS	N	7.90
1	4	467:LYS	C	468:LYS	N	3.57
1	5	720:ARG	C	721:ARG	N	1.17

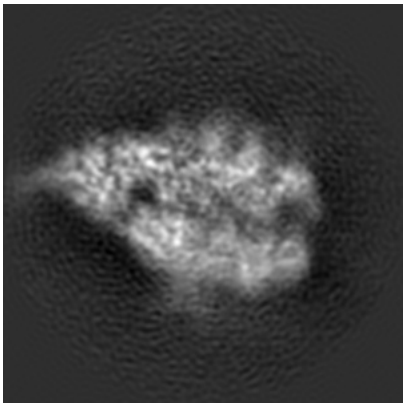
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-6535. These allow visual inspection of the internal detail of the map and identification of artifacts.

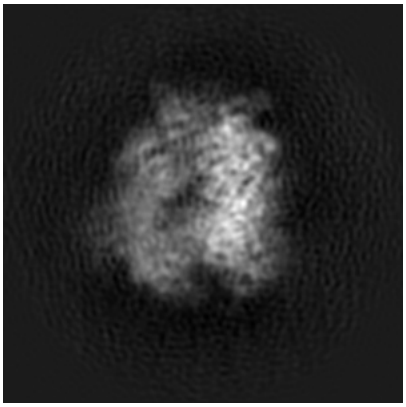
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

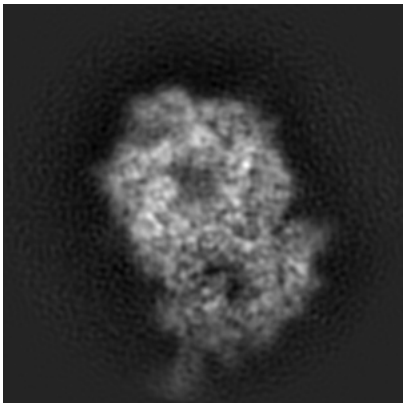
6.1.1 Primary map



X



Y

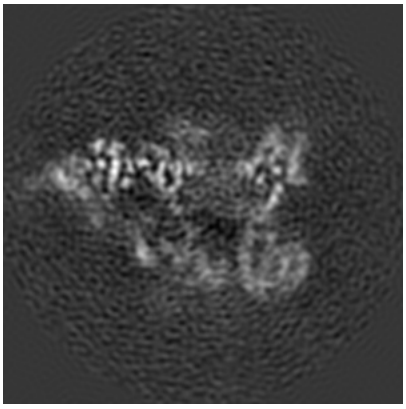


Z

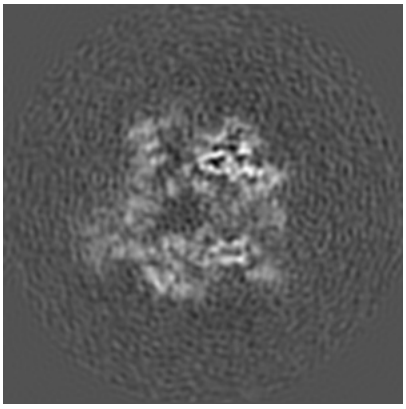
The images above show the map projected in three orthogonal directions.

6.2 Central slices

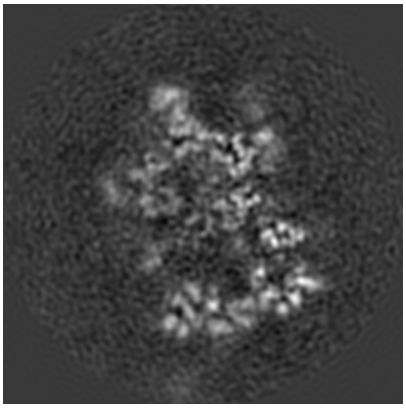
6.2.1 Primary map



X Index: 128



Y Index: 128

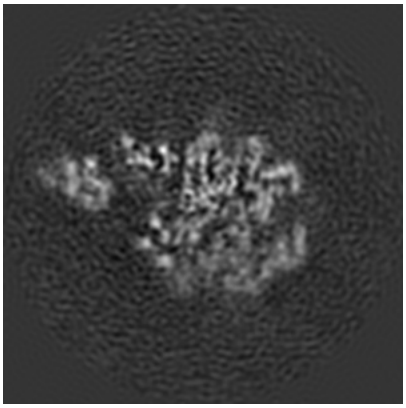


Z Index: 128

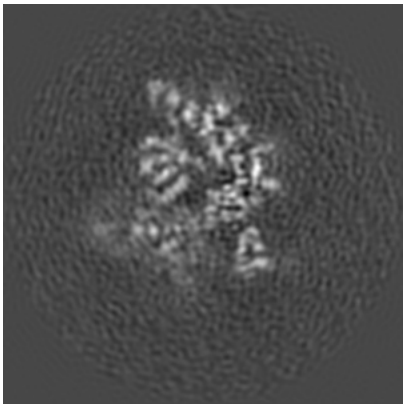
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

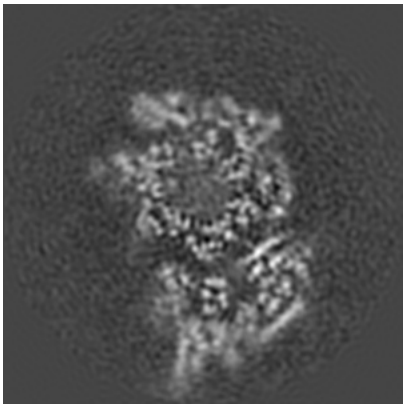
6.3.1 Primary map



X Index: 150



Y Index: 102

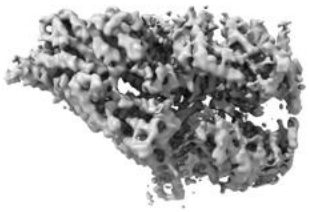


Z Index: 145

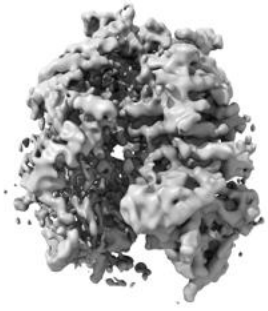
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views (i)

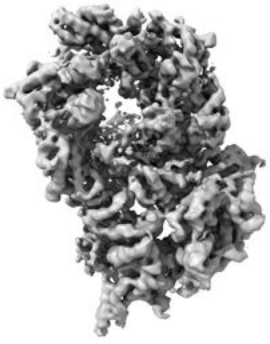
6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.018. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

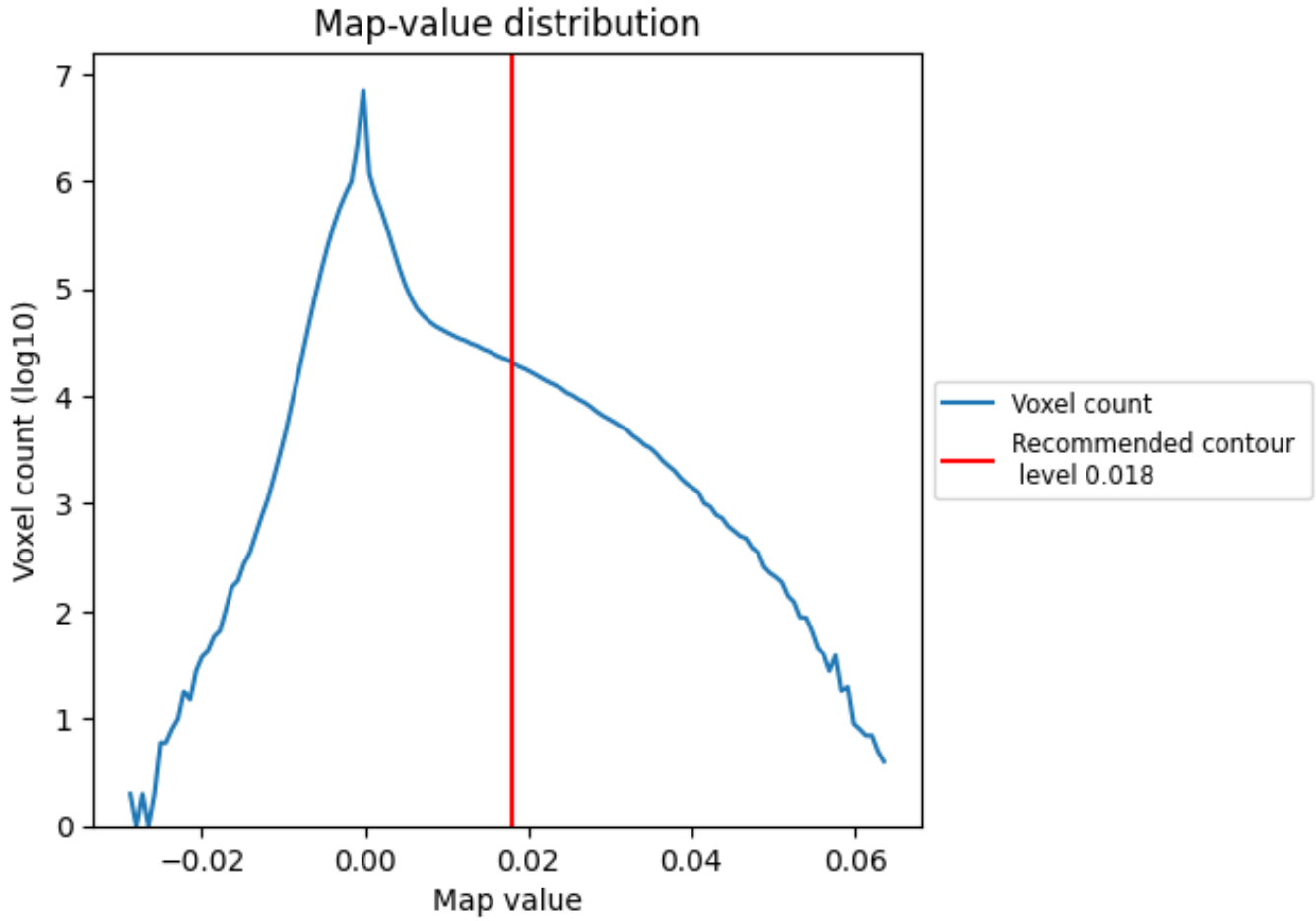
6.5 Mask visualisation (i)

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

7 Map analysis

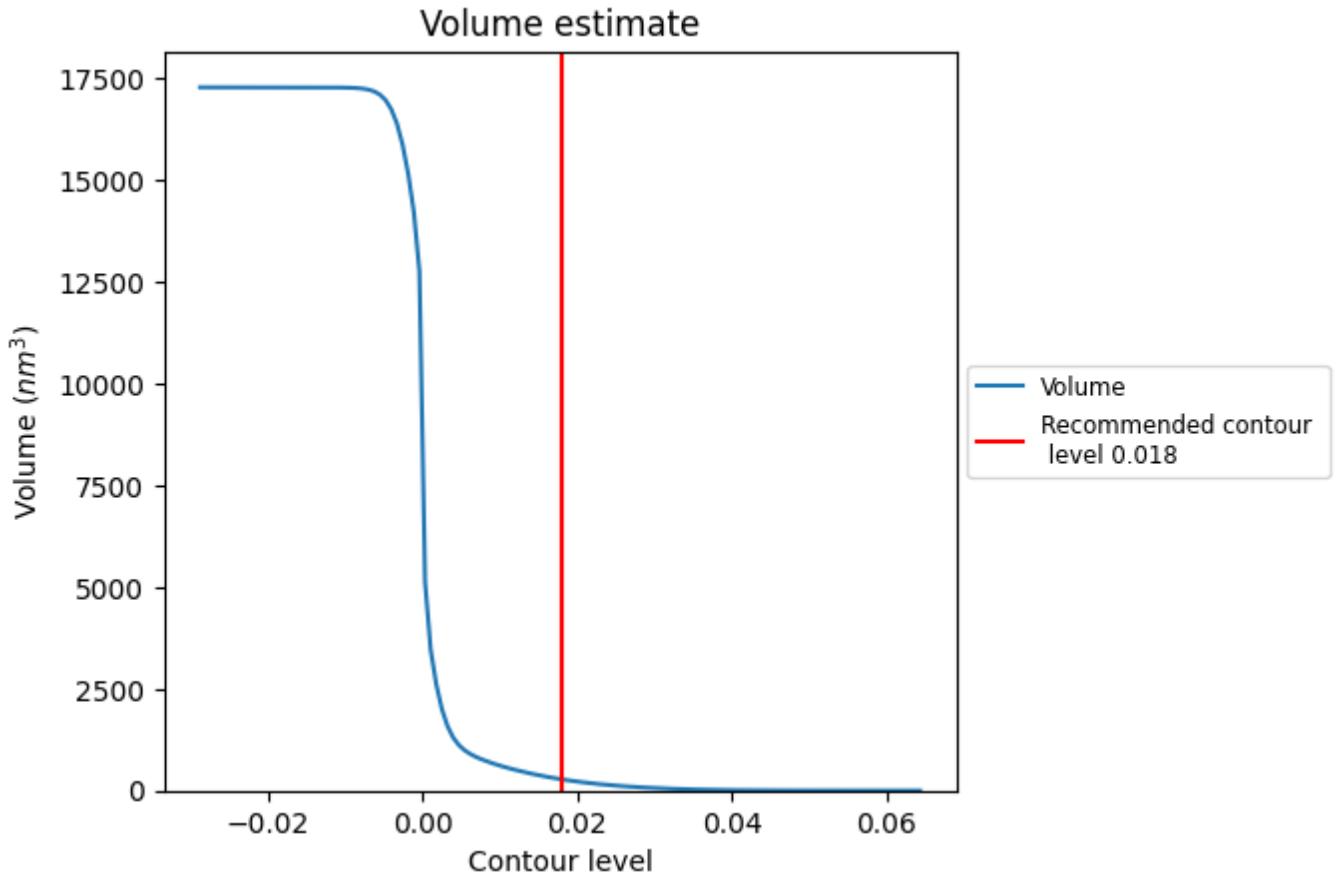
This section contains the results of statistical analysis of the map.

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.

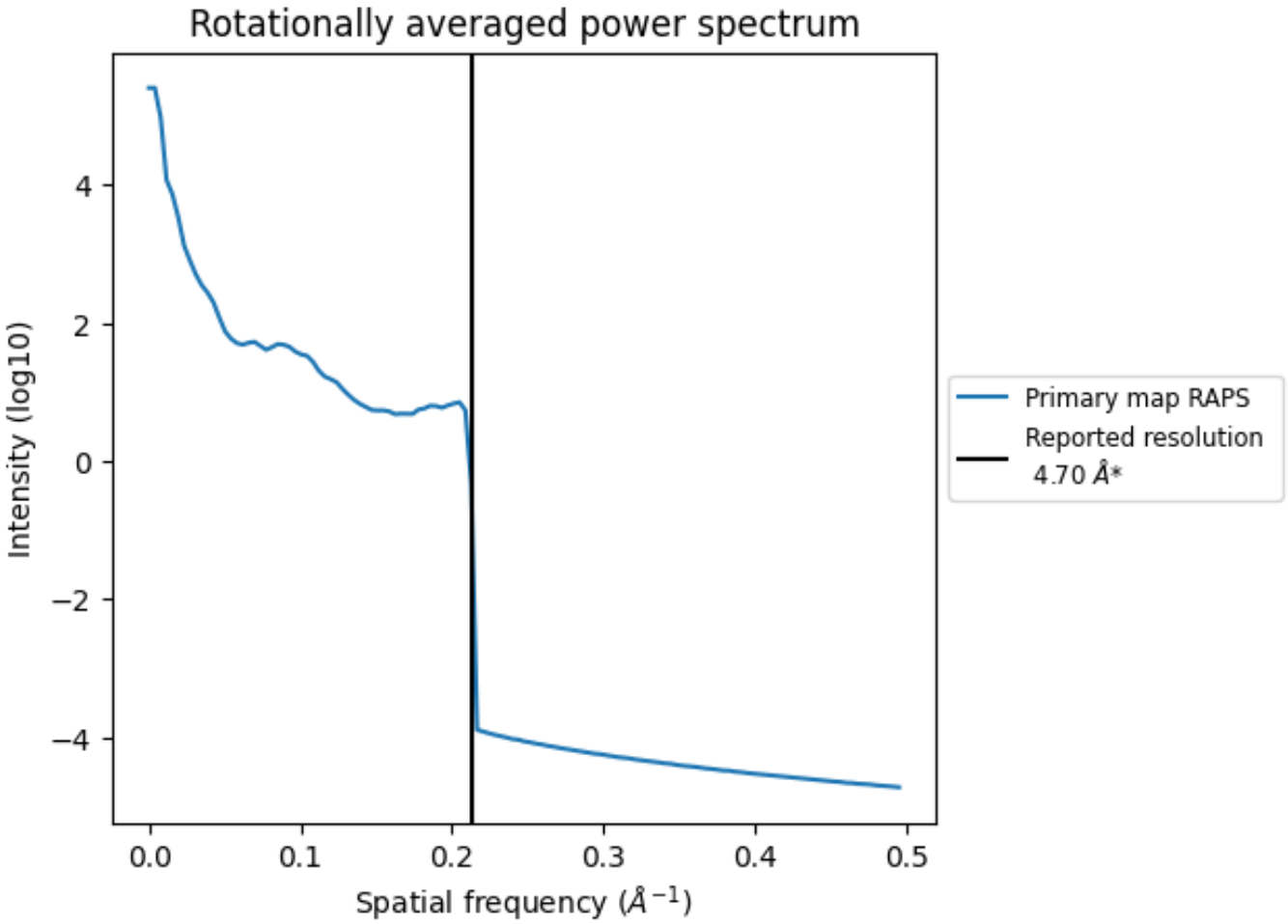
7.2 Volume estimate (i)



The volume at the recommended contour level is 275 nm^3; this corresponds to an approximate mass of 249 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 0.213 Å^{-1}

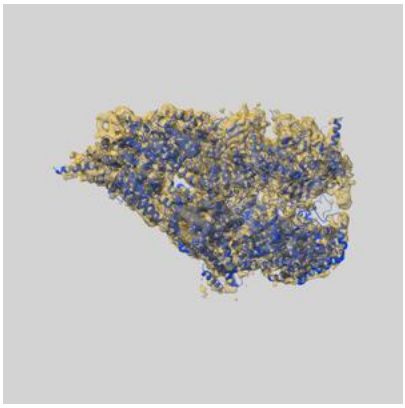
8 Fourier-Shell correlation (i)

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

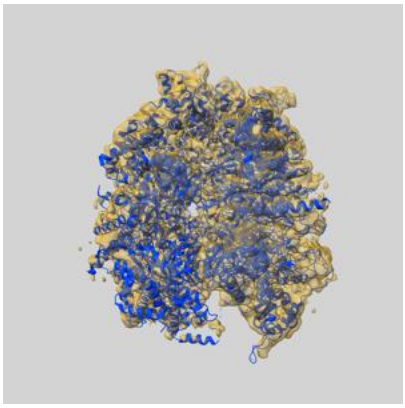
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-6535 and PDB model 3JC5. Per-residue inclusion information can be found in section 3 on page 6.

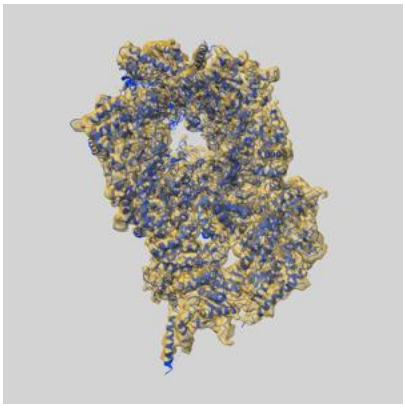
9.1 Map-model overlay



X



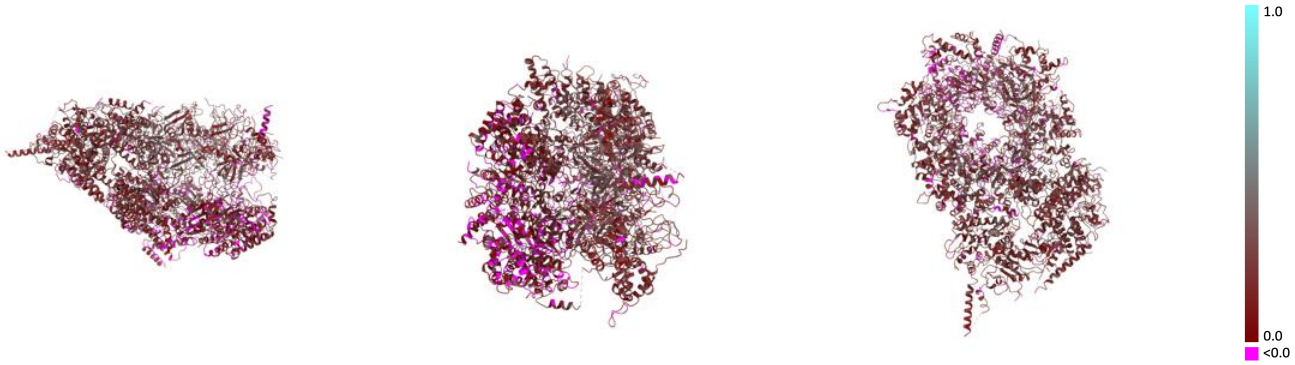
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.018 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.

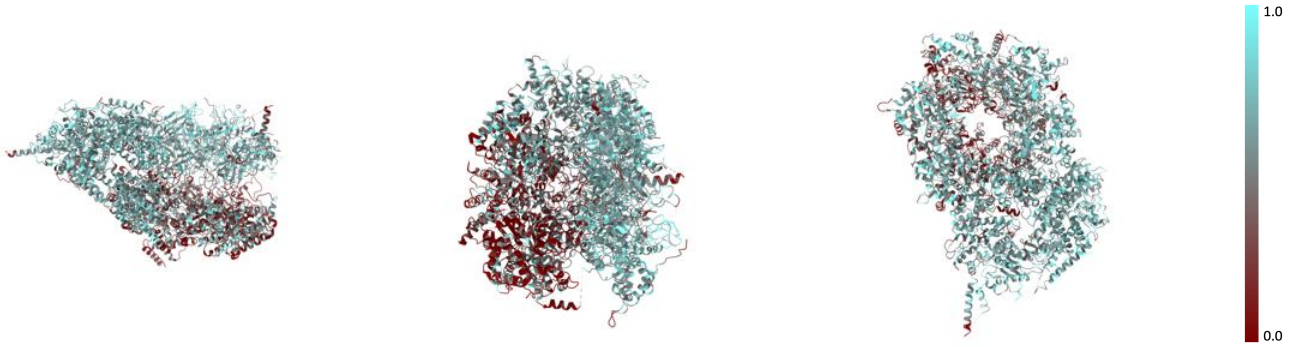
9.2 Q-score mapped to coordinate model



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.

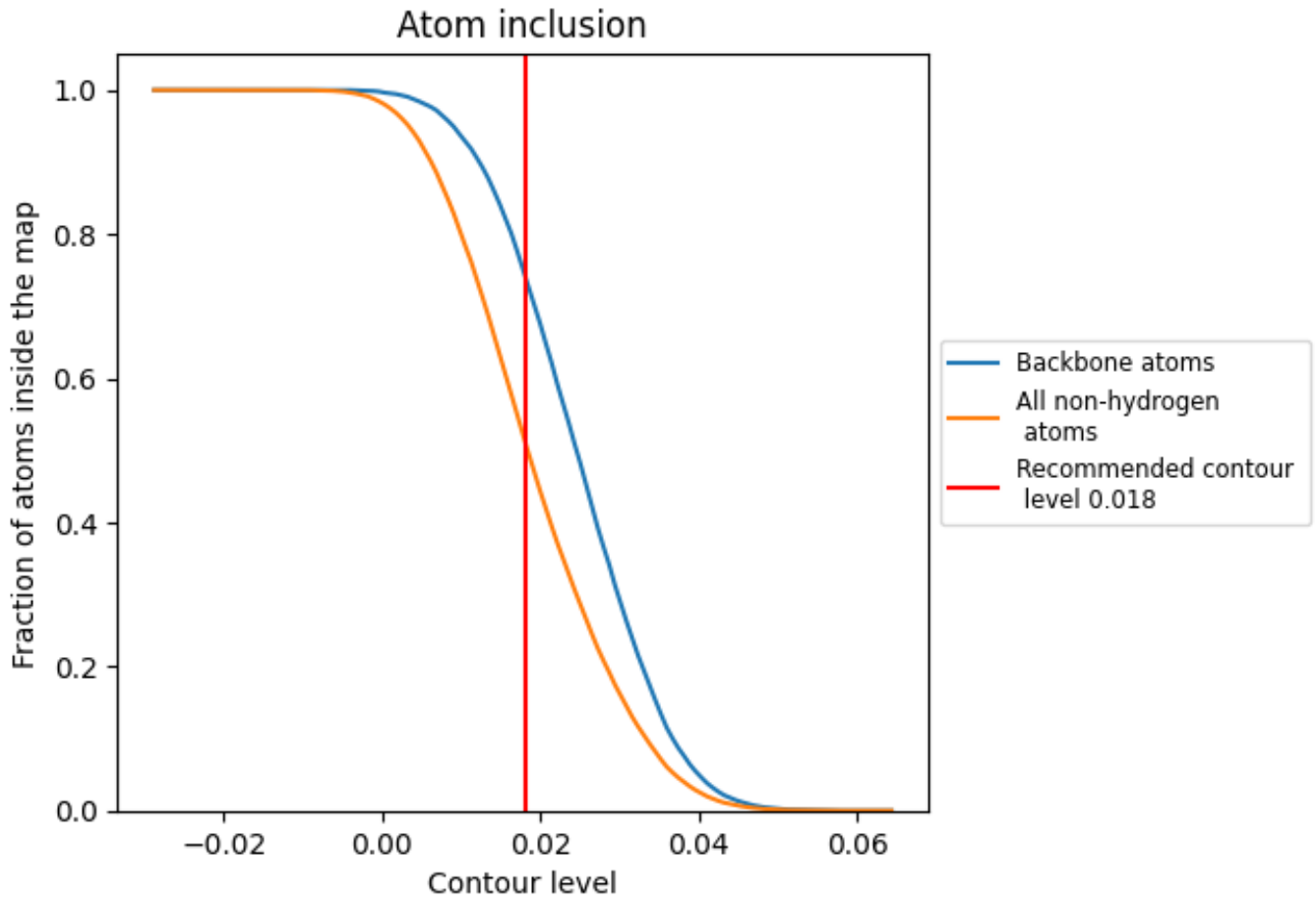
9.3 Atom inclusion mapped to coordinate model



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 (0.018).

9.4 Atom inclusion (i)



At the recommended contour level, 74% of all backbone atoms, 51% 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 (0.018) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5128	0.1910
2	0.4766	0.1960
3	0.5353	0.1980
4	0.3852	0.1210
5	0.5028	0.2050
6	0.4427	0.1690
7	0.4681	0.1770
A	0.5825	0.2070
B	0.6528	0.2270
C	0.6562	0.2580
D	0.6659	0.2260
c	0.6227	0.2200