



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 16, 2018 – 09:44 AM EST

PDB ID : 3JC6  
EMDB ID: : EMD-6534  
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 : 3.70 Å(reported)  
Based on PDB ID : 2Q9Q

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report  
for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736



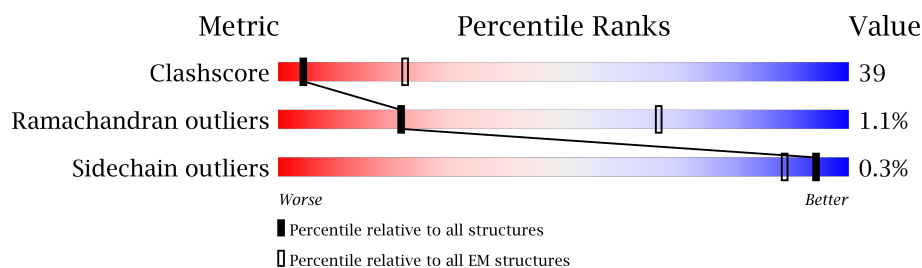
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.70 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)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	2	868	
2	3	971	
3	4	933	
4	5	775	
5	6	1017	
6	7	845	
7	E	672	
8	D	294	
9	B	213	

*Continued on next page...*



Continued from previous page...

Mol	Chain	Length	Quality of chain
10	A	208	<div><div></div><div>40%</div><div>58%</div><div></div></div>
11	C	194	<div><div></div><div>39%</div><div>42%</div><div>18%</div><div></div></div>



## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 23732 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
1	2	241	Total	C	N	O	S	0	0
			1911	1214	338	354	5		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	3	269	Total	C	N	O	S	0	0
			2130	1354	368	404	4		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	4	275	Total	C	N	O	S	0	0
			2203	1391	382	413	17		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	5	254	Total	C	N	O	S	0	0
			2028	1284	347	388	9		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	6	267	Total	C	N	O	S	0	0
			2049	1296	366	381	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	7	325	Total	C	N	O	S	0	0
			2611	1653	455	491	12		



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

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

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	651	ASP	-	EXPRESSION TAG	UNP Q08032
E	652	TYR	-	EXPRESSION TAG	UNP Q08032
E	653	LYS	-	EXPRESSION TAG	UNP Q08032
E	654	ASP	-	EXPRESSION TAG	UNP Q08032
E	655	HIS	-	EXPRESSION TAG	UNP Q08032
E	656	ASP	-	EXPRESSION TAG	UNP Q08032
E	657	GLY	-	EXPRESSION TAG	UNP Q08032
E	658	ASP	-	EXPRESSION TAG	UNP Q08032
E	659	TYR	-	EXPRESSION TAG	UNP Q08032
E	660	LYS	-	EXPRESSION TAG	UNP Q08032
E	661	ASP	-	EXPRESSION TAG	UNP Q08032
E	662	HIS	-	EXPRESSION TAG	UNP Q08032
E	663	ASP	-	EXPRESSION TAG	UNP Q08032
E	664	ILE	-	EXPRESSION TAG	UNP Q08032
E	665	ASP	-	EXPRESSION TAG	UNP Q08032
E	666	TYR	-	EXPRESSION TAG	UNP Q08032
E	667	LYS	-	EXPRESSION TAG	UNP Q08032
E	668	ASP	-	EXPRESSION TAG	UNP Q08032
E	669	ASP	-	EXPRESSION TAG	UNP Q08032
E	670	ASP	-	EXPRESSION TAG	UNP Q08032
E	671	ASP	-	EXPRESSION TAG	UNP Q08032
E	672	LYS	-	EXPRESSION TAG	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
			1696	1065	290	331	10		

- 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		

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
12	7	1	Total	Zn	0
			1	1	



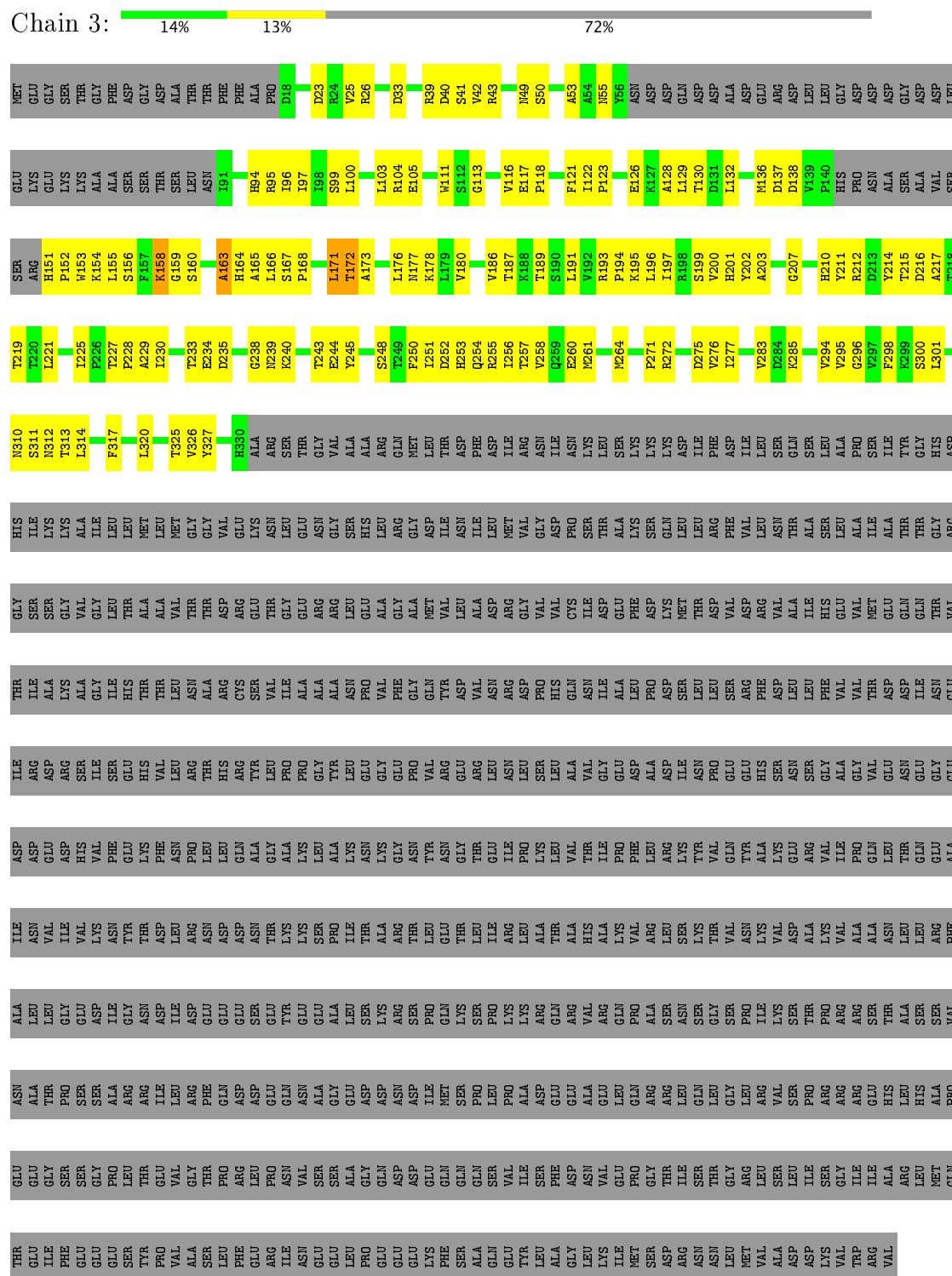
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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. 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.

- Chain 2:  16% 12% 72%





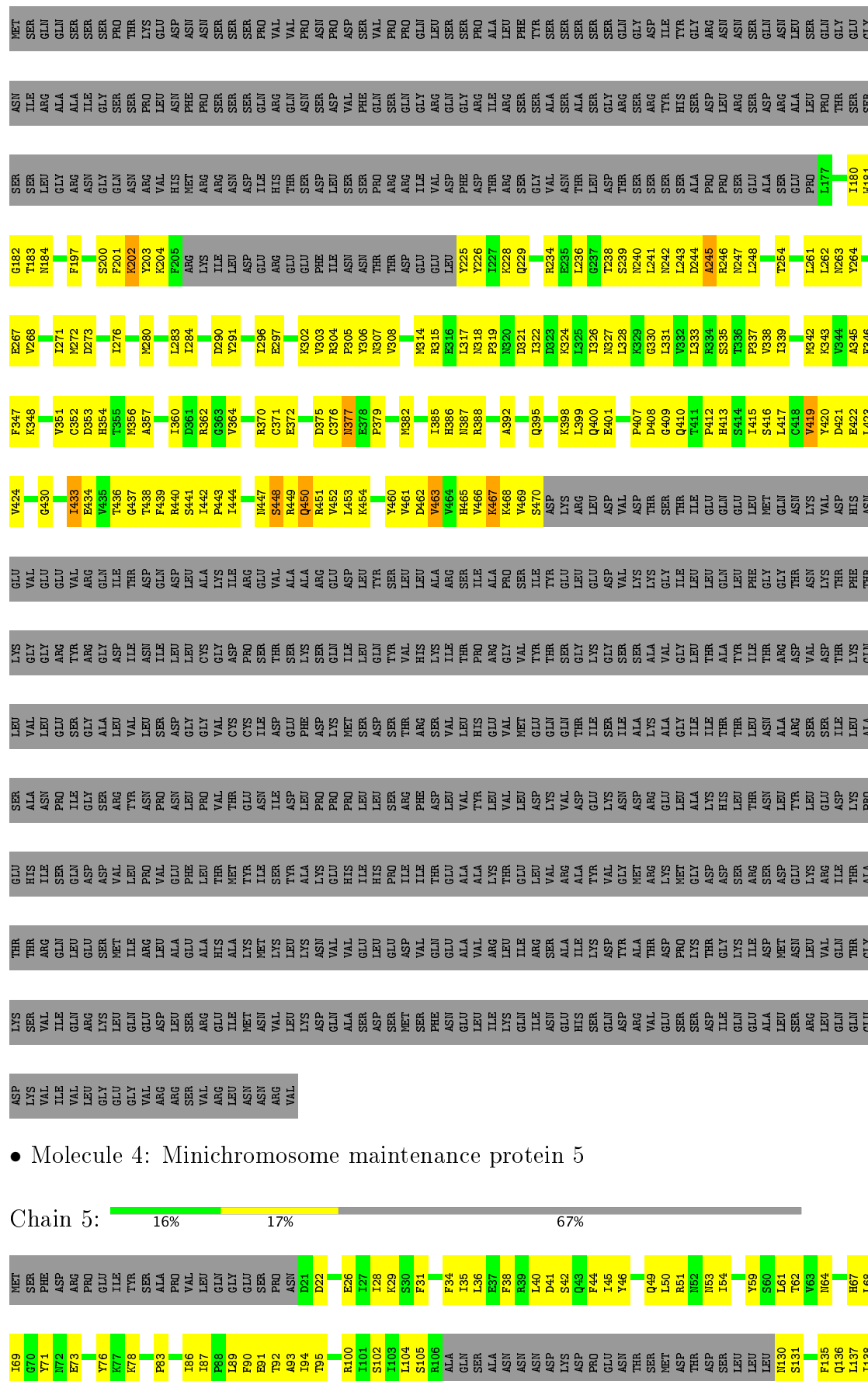
- Molecule 2: DNA replication licensing factor MCM3



- Molecule 3: DNA replication licensing factor MCM4

Chain 4:  14% 15% . 71%

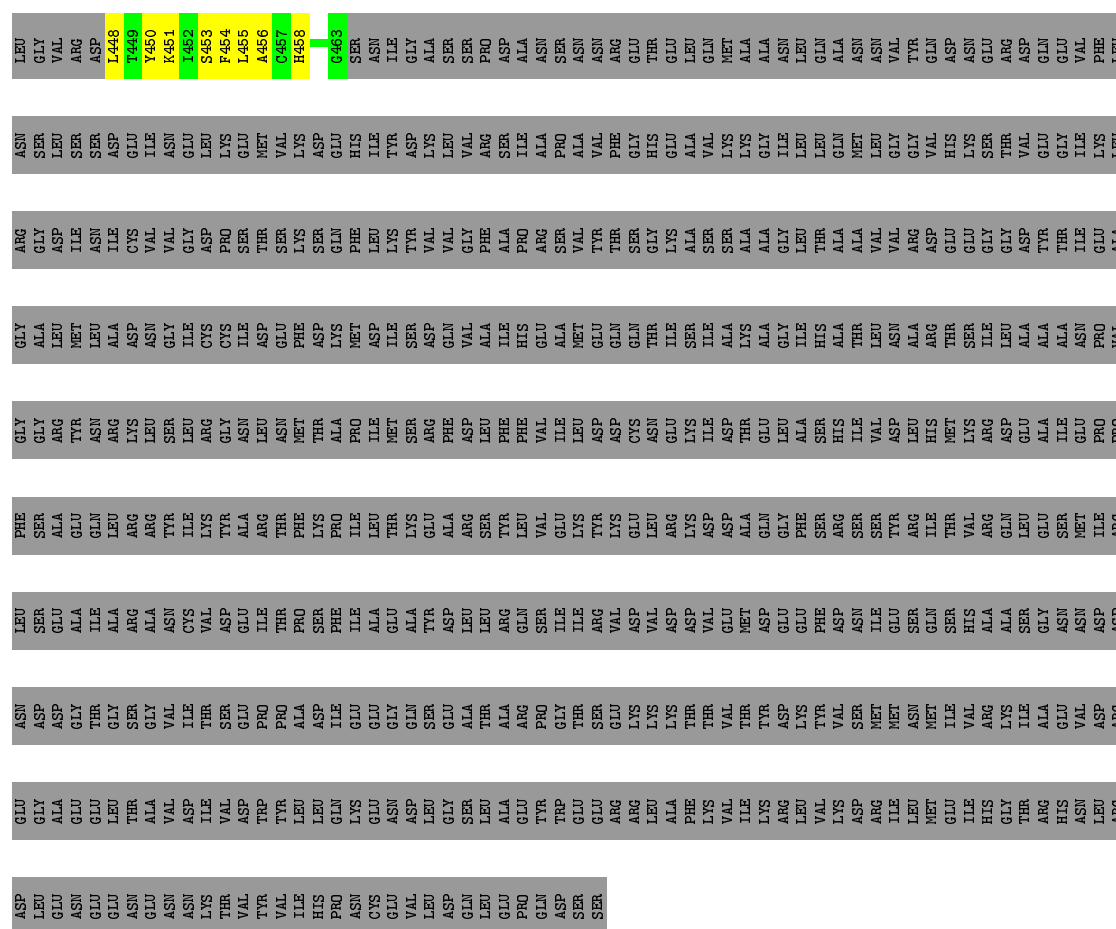




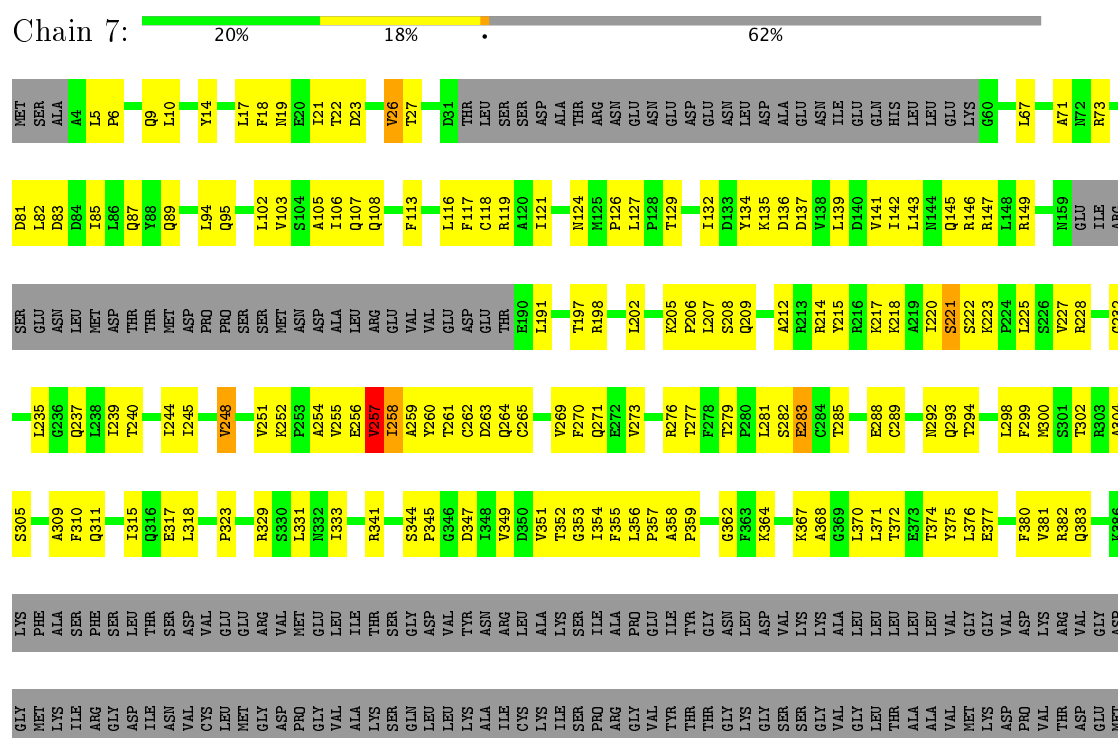








- Molecule 6: DNA replication licensing factor MCM7

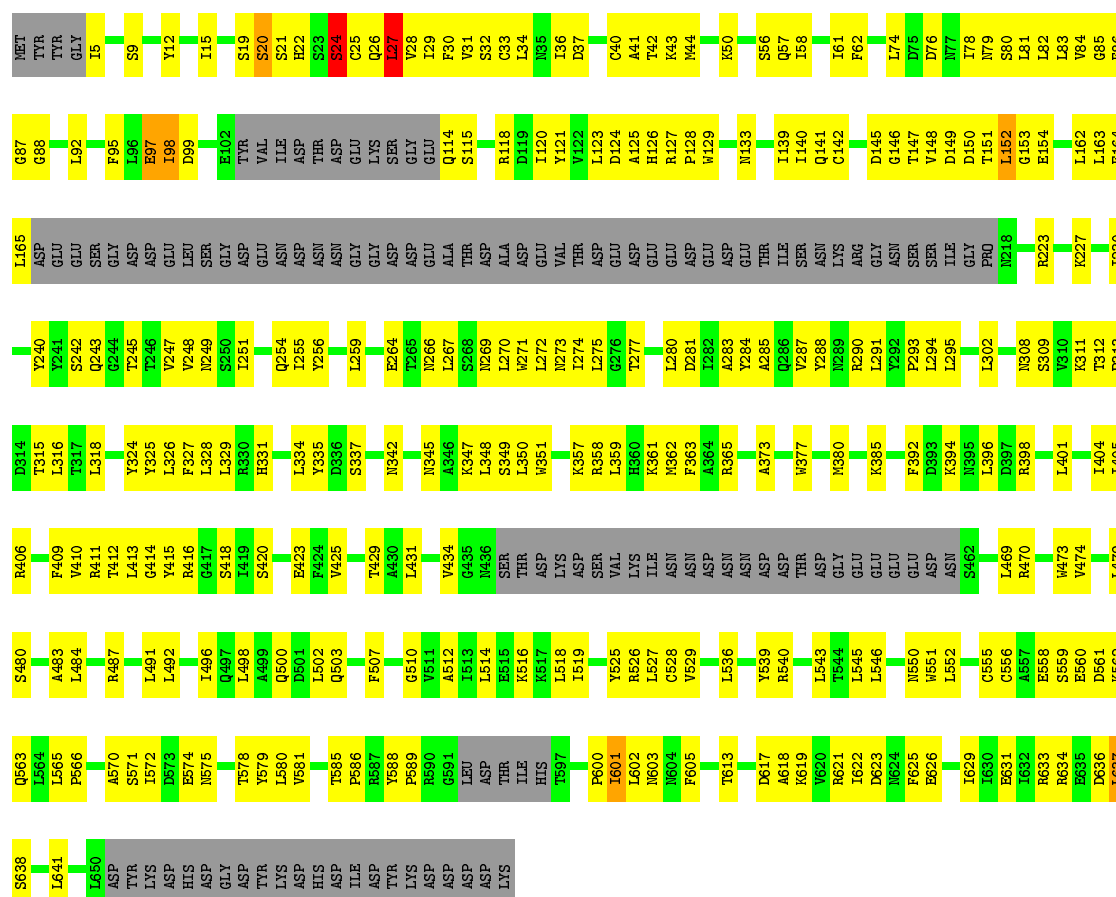





[illegible]

- Molecule 7: Cell division control protein 45

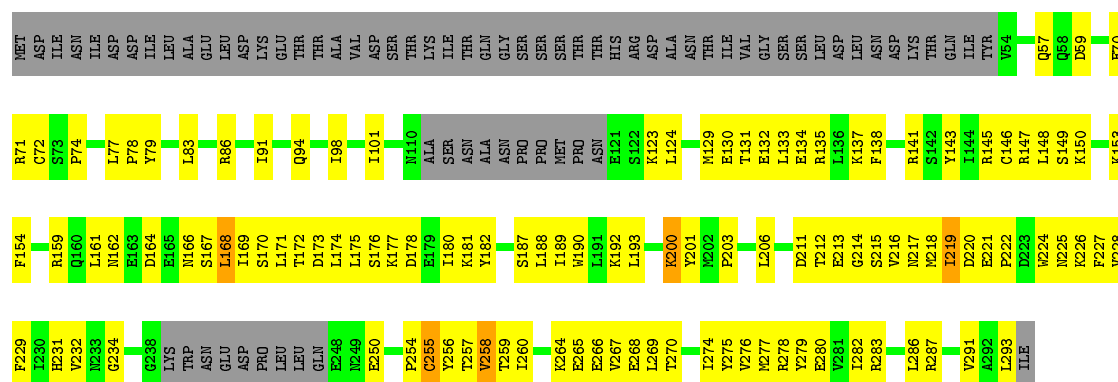
Chain E:  43% 38% . 18%



- Molecule 8: DNA replication complex GINS protein SLD5

Chain D:  36% 37% . 25%

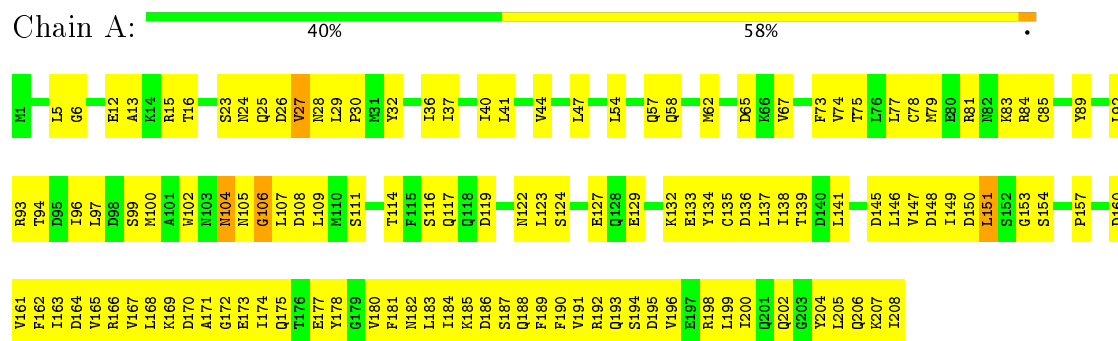




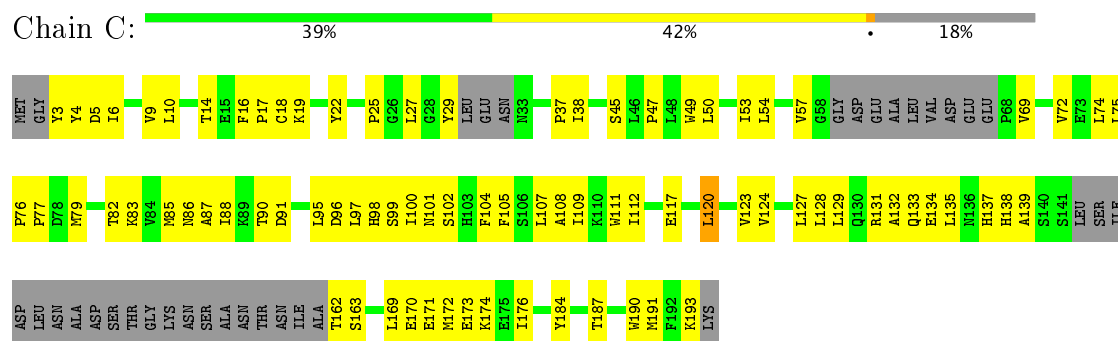
- Molecule 9: DNA replication complex GINS protein PSF2



- Molecule 10: DNA replication complex GINS protein PSF1



- Molecule 11: DNA replication complex GINS protein PSF3





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	469818	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 ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	49505	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor



## 5 Model quality

### 5.1 Standard geometry

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
1	2	0.41	0/1946	0.66	1/2633 (0.0%)
10	A	0.39	0/1718	0.65	0/2314
11	C	0.46	0/1320	0.64	1/1784 (0.1%)
2	3	0.46	0/2179	0.70	1/2963 (0.0%)
3	4	0.38	0/2240	0.67	0/3029
4	5	0.50	0/2057	0.73	2/2781 (0.1%)
5	6	0.39	0/2081	0.66	1/2813 (0.0%)
6	7	0.41	0/2657	0.66	1/3592 (0.0%)
7	E	0.39	0/4563	0.63	2/6173 (0.0%)
8	D	0.43	0/1853	0.64	0/2500
9	B	0.44	0/1545	0.69	0/2092
All	All	0.42	0/24159	0.66	9/32674 (0.0%)

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
10	A	0	4
2	3	0	2
3	4	0	5
5	6	0	3
6	7	0	4
7	E	0	1
8	D	0	2
All	All	0	21

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	E	27	LEU	CA-CB-CG	7.97	133.63	115.30
4	5	258	LEU	CA-CB-CG	7.60	132.78	115.30
2	3	171	LEU	CA-CB-CG	7.21	131.88	115.30
5	6	105	ASP	CB-CG-OD1	6.55	124.20	118.30
1	2	436	GLY	N-CA-C	5.96	128.01	113.10

There are no chirality outliers.

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	3	163	ALA	Peptide
2	3	165	ALA	Peptide
3	4	202	LYS	Peptide
3	4	245	ALA	Peptide
3	4	372	GLU	Peptide

## 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	2	1911	0	1908	126	0
2	3	2130	0	2105	172	0
3	4	2203	0	2188	198	0
4	5	2028	0	2055	177	0
5	6	2049	0	1959	132	0
6	7	2611	0	2623	165	0
7	E	4482	0	4497	322	0
8	D	1820	0	1823	214	0
9	B	1513	0	1558	150	0
10	A	1696	0	1698	310	0
11	C	1288	0	1298	104	0
12	7	1	0	0	0	0
All	All	23732	0	23712	1852	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 39.

The worst 5 of 1852 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)
4:5:197:PHE:CE2	4:5:329:LYS:HG2	1.19	1.65
6:7:17:LEU:HD11	6:7:102:LEU:CD2	1.37	1.55
5:6:290:ILE:HD13	5:6:454:PHE:CZ	1.40	1.51
3:4:342:MET:HB3	3:4:360:ILE:CD1	1.45	1.46
9:B:187:GLU:OE2	11:C:176:ILE:CG2	1.65	1.43

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	2	235/868 (27%)	205 (87%)	25 (11%)	5 (2%)	8	49
2	3	263/971 (27%)	236 (90%)	22 (8%)	5 (2%)	9	51
3	4	271/933 (29%)	230 (85%)	36 (13%)	5 (2%)	10	52
4	5	244/775 (32%)	222 (91%)	19 (8%)	3 (1%)	15	59
5	6	259/1017 (26%)	229 (88%)	27 (10%)	3 (1%)	15	59
6	7	319/845 (38%)	270 (85%)	45 (14%)	4 (1%)	14	57
7	E	543/672 (81%)	491 (90%)	47 (9%)	5 (1%)	20	64
8	D	215/294 (73%)	197 (92%)	16 (7%)	2 (1%)	20	64
9	B	177/213 (83%)	163 (92%)	14 (8%)	0	100	100
10	A	206/208 (99%)	180 (87%)	25 (12%)	1 (0%)	32	73
11	C	151/194 (78%)	144 (95%)	7 (5%)	0	100	100
All	All	2883/6990 (41%)	2567 (89%)	283 (10%)	33 (1%)	21	61

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	2	291	SER
3	4	450	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
6	7	26	VAL
7	E	601	ILE
3	4	419	VAL

### 5.3.2 Protein sidechains ⓘ

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	206/770 (27%)	206 (100%)	0	100	100
2	3	236/835 (28%)	236 (100%)	0	100	100
3	4	249/848 (29%)	249 (100%)	0	100	100
4	5	241/688 (35%)	241 (100%)	0	100	100
5	6	207/886 (23%)	207 (100%)	0	100	100
6	7	295/753 (39%)	295 (100%)	0	100	100
7	E	499/607 (82%)	495 (99%)	4 (1%)	85	93
8	D	213/279 (76%)	212 (100%)	1 (0%)	91	96
9	B	171/198 (86%)	170 (99%)	1 (1%)	89	95
10	A	193/193 (100%)	192 (100%)	1 (0%)	91	96
11	C	144/173 (83%)	144 (100%)	0	100	100
All	All	2654/6230 (43%)	2647 (100%)	7 (0%)	94	98

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	E	152	LEU
10	A	151	LEU
8	D	168	LEU
7	E	24	SER
9	B	175	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:



Mol	Chain	Res	Type
6	7	145	GLN
7	E	26	GLN
10	A	175	GLN
6	7	311	GLN
7	E	243	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 carbohydrates 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.