



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

Feb 6, 2017 – 05:09 PM EST

PDB ID : 5U8T  
EMDB ID: : EMD-8519  
Title : Structure of Eukaryotic CMG Helicase at a Replication Fork and Implications  
Authors : Li, B.; Georgescu, R.; Yuan, Z.; Santos, R.; Sun, J.; Zhang, D.; Yurieva, O.;  
Li, H.; O'Donnell, M.E.  
Deposited on : 2016-12-15  
Resolution : 4.90 Å(reported)

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  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442

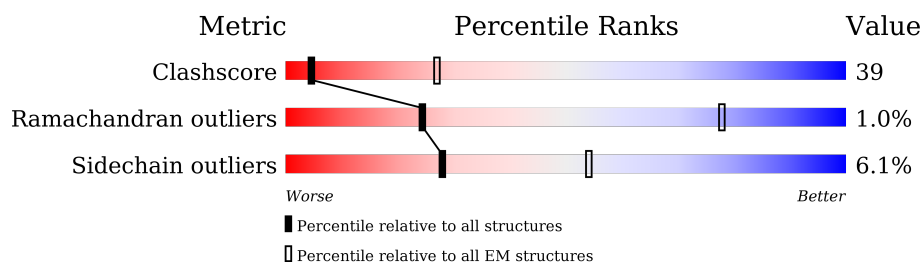
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.90 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	23% 39% . 33%
2	3	971	23% 34% . 39%
3	4	933	27% 41% . 28%
4	5	775	27% 47% . . 22%
5	6	1017	21% 40% . 35%
6	7	845	31% 44% . 22%
7	A	208	39% 56% 5%
8	B	213	30% 52% . 15%
9	C	194	32% 46% . 18%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
10	D	294	
11	E	650	
12	F	14	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
13	ANP	2	901	-	-	X	-
13	ANP	3	1001	-	-	X	-
13	ANP	5	801	-	-	X	-

## 2 Entry composition

There are 13 unique types of molecules in this entry. The entry contains 40788 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	583	Total	C	N	O	S	0	0
			4591	2899	818	858	16		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	3	589	Total	C	N	O	S	0	0
			4624	2915	824	872	13		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	4	672	Total	C	N	O	S	0	0
			5318	3340	929	1021	28		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	5	602	Total	C	N	O	S	0	0
			4740	2980	815	921	24		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
5	6	661	Total	C	N	O	S	0	0
			5142	3247	905	967	23		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	7	660	Total	C	N	O	S	0	0
			5201	3278	903	991	29		

- Molecule 7 is a protein called DNA replication complex GINS protein PSF1.

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

- Molecule 8 is a protein called DNA replication complex GINS protein PSF2.

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

- Molecule 9 is a protein called DNA replication complex GINS protein PSF3.

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

- Molecule 10 is a protein called DNA replication complex GINS protein SLD5.

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

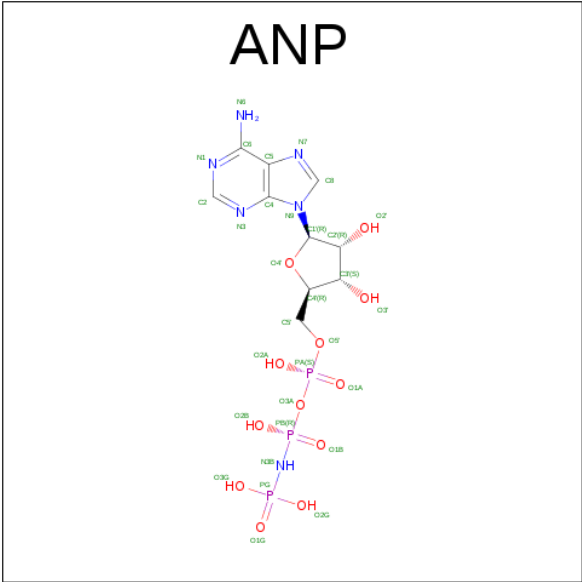
- Molecule 11 is a protein called Cell division control protein 45.

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

- Molecule 12 is a DNA chain called DNA (5'-D(P\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*T  
P\*TP\*TP\*TP\*T)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
12	F	14	Total	C	N	O	P	0	0
			280	140	28	98	14		

- Molecule 13 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>6</sub>O<sub>12</sub>P<sub>3</sub>).

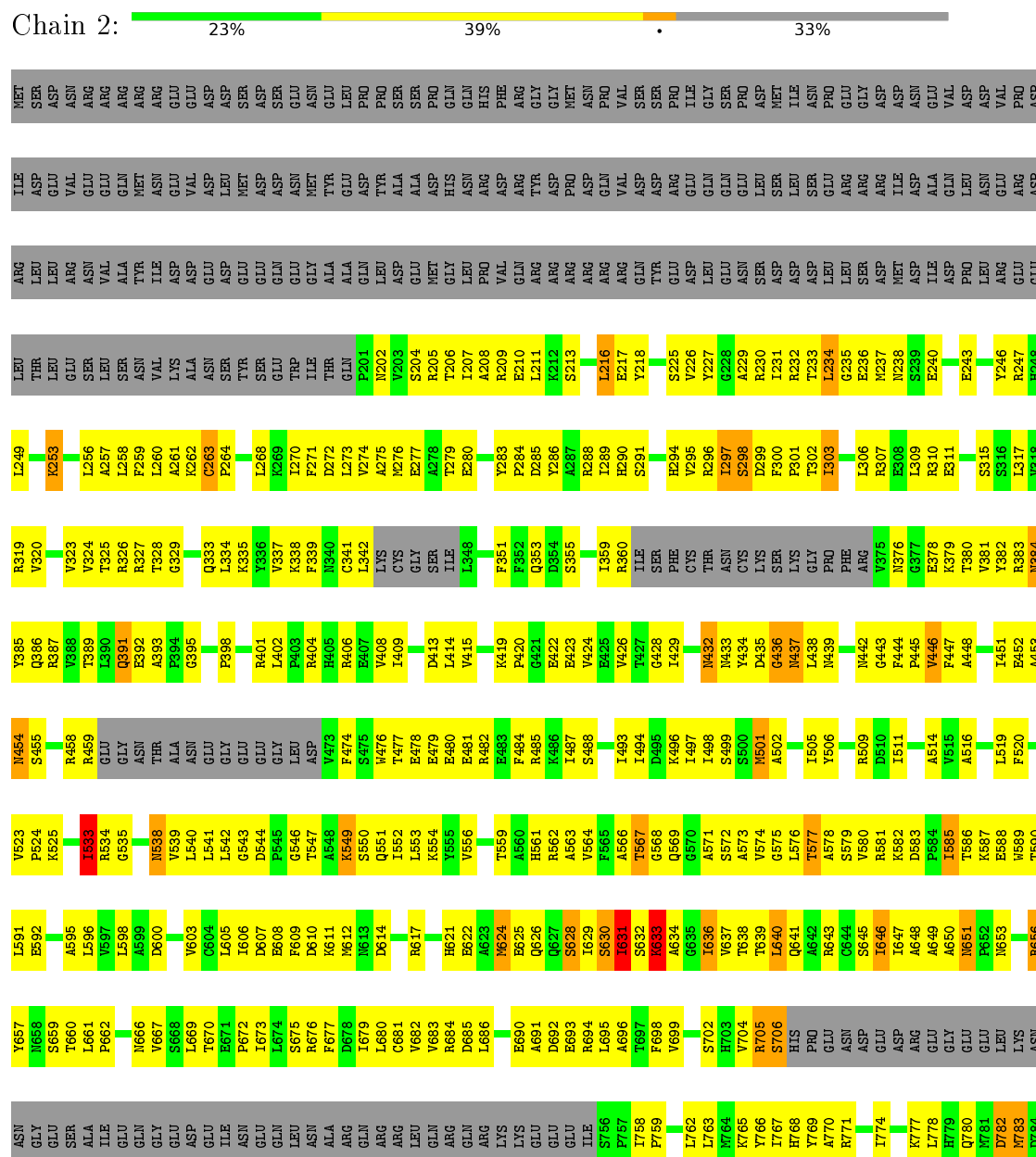


Mol	Chain	Residues	Atoms					AltConf
13	2	1	Total	C	N	O	P	0
			31	10	6	12	3	
13	3	1	Total	C	N	O	P	0
			31	10	6	12	3	
13	5	1	Total	C	N	O	P	0
			31	10	6	12	3	

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

- Molecule 1: DNA replication licensing factor MCM2

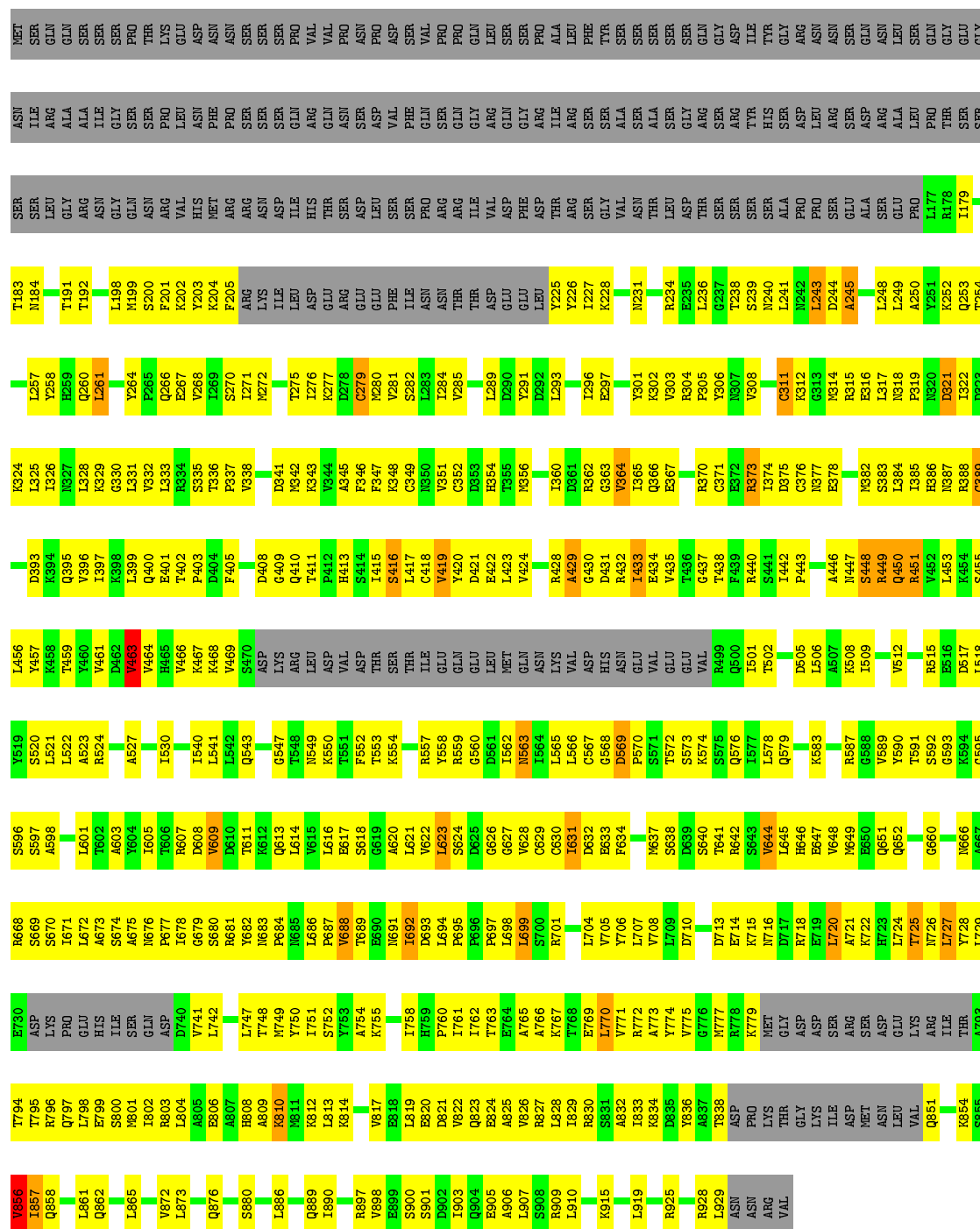






- Molecule 3: DNA replication licensing factor MCM4

Chain 4:  27% 41% . 28%

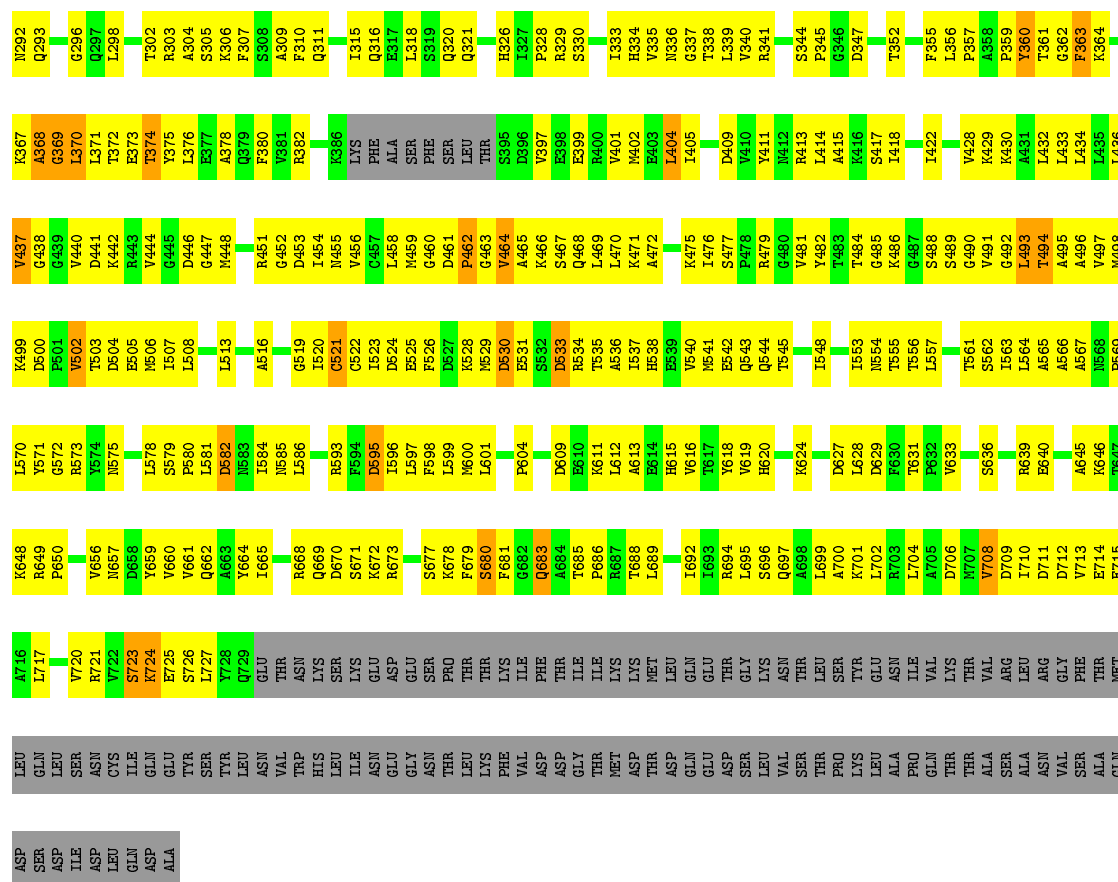


- Molecule 4: Minichromosome maintenance protein 5

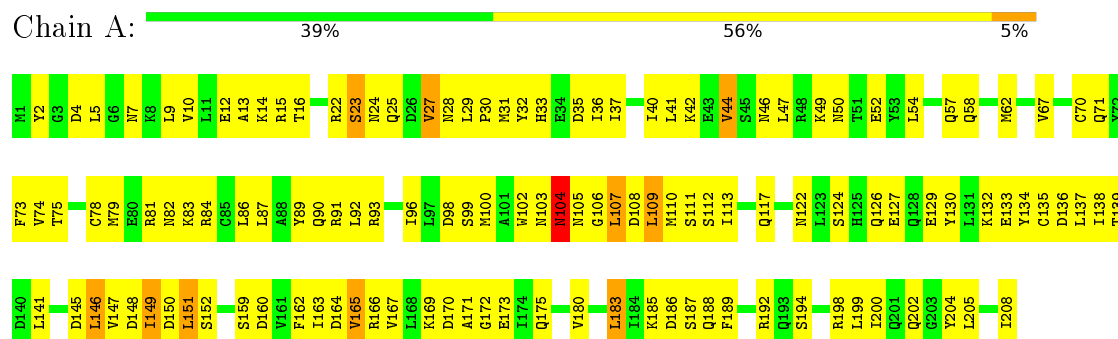
Chain 5:  27% 47% .. 22%



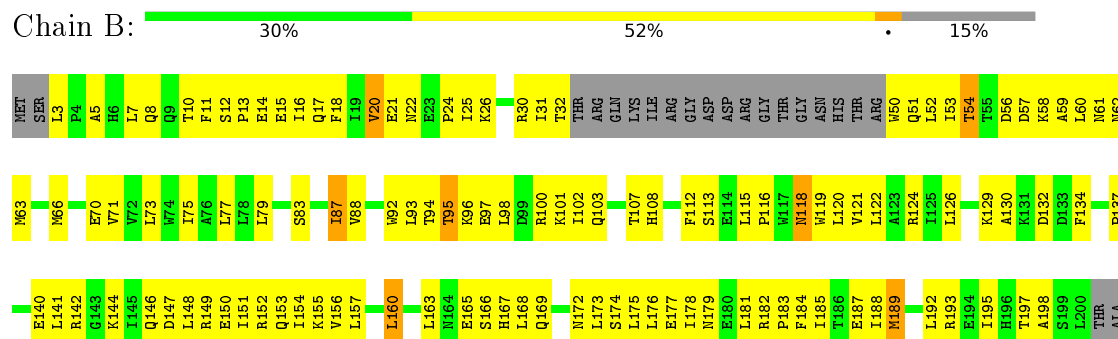




• Molecule 7: DNA replication complex GINS protein PSF1

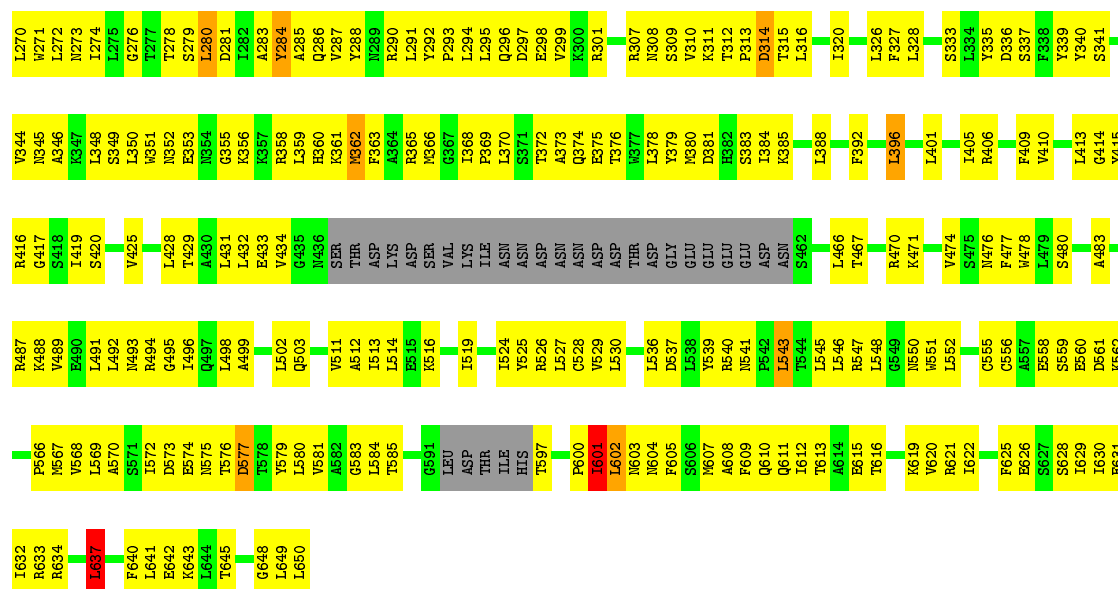


• Molecule 8: DNA replication complex GINS protein PSF2



WORLDWIDE  
**PDB**  
PROTEIN DATA BANK

 **EMDataBank**  
Unified Data Resource for 3DEM



- Molecule 12: DNA (5'-D(P\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*TP\*T)-3')



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	395443	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	0.37	1/4668 (0.0%)	0.65	2/6306 (0.0%)
10	D	0.34	0/1853	0.69	2/2500 (0.1%)
11	E	0.33	0/4563	0.63	5/6173 (0.1%)
12	F	0.88	1/307 (0.3%)	1.42	3/472 (0.6%)
2	3	0.32	0/4702	0.63	1/6374 (0.0%)
3	4	0.31	0/5388	0.63	0/7273
4	5	0.34	0/4805	0.65	1/6489 (0.0%)
5	6	0.36	0/5218	0.69	3/7039 (0.0%)
6	7	0.33	0/5281	0.65	2/7136 (0.0%)
7	A	0.36	0/1718	0.70	1/2314 (0.0%)
8	B	0.33	0/1545	0.62	0/2092
9	C	0.32	0/1320	0.60	1/1784 (0.1%)
All	All	0.34	2/41368 (0.0%)	0.66	21/55952 (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
1	2	0	2
10	D	0	1
2	3	0	5
3	4	0	3
4	5	0	1
5	6	0	1
6	7	0	5
7	A	0	2
All	All	0	20

All (2) bond length outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	2	633	LYS	CE-NZ	6.70	1.65	1.49
12	F	12	DT	O5'-C5'	-5.01	1.29	1.42

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	7	369	GLY	N-CA-C	8.34	133.96	113.10
12	F	7	DT	O4'-C4'-C3'	-7.76	101.35	106.00
12	F	12	DT	OP1-P-OP2	7.47	130.81	119.60
10	D	269	LEU	CA-CB-CG	7.46	132.46	115.30
5	6	948	LEU	CA-CB-CG	7.37	132.25	115.30

There are no chirality outliers.

5 of 20 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	2	355	SER	Peptide
1	2	633	LYS	Peptide
2	3	163	ALA	Peptide
2	3	172	THR	Peptide
2	3	428	LEU	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	4591	0	4639	428	0
2	3	4624	0	4689	376	0
3	4	5318	0	5390	424	0
4	5	4740	0	4798	439	0
5	6	5142	0	5093	499	0
6	7	5201	0	5276	422	0
7	A	1696	0	1698	141	0
8	B	1513	0	1558	124	0
9	C	1288	0	1298	92	0
10	D	1820	0	1824	201	0
11	E	4482	0	4499	367	0
12	F	280	0	169	8	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
13	2	31	0	13	9	0
13	3	31	0	13	10	0
13	5	31	0	13	9	0
All	All	40788	0	40970	3219	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 3219 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)
10:D:260:ILE:HG13	10:D:266:GLU:CG	1.46	1.45
6:7:221:SER:CA	6:7:222:SER:HB2	1.40	1.44
10:D:260:ILE:CG1	10:D:266:GLU:CD	1.87	1.42
10:D:260:ILE:HG12	10:D:266:GLU:CD	1.38	1.38
10:D:260:ILE:CG1	10:D:266:GLU:OE2	1.74	1.35

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	573/868 (66%)	502 (88%)	63 (11%)	8 (1%)	14	58
2	3	579/971 (60%)	525 (91%)	52 (9%)	2 (0%)	46	83
3	4	660/933 (71%)	577 (87%)	72 (11%)	11 (2%)	11	56
4	5	590/775 (76%)	537 (91%)	48 (8%)	5 (1%)	24	69
5	6	649/1017 (64%)	571 (88%)	72 (11%)	6 (1%)	21	67
6	7	652/845 (77%)	566 (87%)	76 (12%)	10 (2%)	13	58
7	A	206/208 (99%)	185 (90%)	20 (10%)	1 (0%)	34	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
8	B	177/213 (83%)	156 (88%)	21 (12%)	0	100	100
9	C	151/194 (78%)	140 (93%)	11 (7%)	0	100	100
10	D	215/294 (73%)	193 (90%)	19 (9%)	3 (1%)	14	58
11	E	543/650 (84%)	488 (90%)	50 (9%)	5 (1%)	21	67
All	All	4995/6968 (72%)	4440 (89%)	504 (10%)	51 (1%)	24	65

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	4	429	ALA
4	5	596	ILE
6	7	26	VAL
6	7	222	SER
6	7	464	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	502/770 (65%)	460 (92%)	42 (8%)	14	50
2	3	512/835 (61%)	475 (93%)	37 (7%)	18	57
3	4	599/848 (71%)	570 (95%)	29 (5%)	31	68
4	5	542/688 (79%)	500 (92%)	42 (8%)	16	54
5	6	539/886 (61%)	497 (92%)	42 (8%)	16	53
6	7	582/753 (77%)	553 (95%)	29 (5%)	30	67
7	A	193/193 (100%)	182 (94%)	11 (6%)	25	64
8	B	171/198 (86%)	162 (95%)	9 (5%)	28	66
9	C	144/173 (83%)	136 (94%)	8 (6%)	26	65
10	D	213/279 (76%)	207 (97%)	6 (3%)	51	79
11	E	499/586 (85%)	479 (96%)	20 (4%)	38	72
All	All	4496/6209 (72%)	4221 (94%)	275 (6%)	28	62

5 of 275 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	5	293	THR
5	6	364	ASN
10	D	260	ILE
4	5	339	THR
4	5	454	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
4	5	411	ASN
4	5	581	ASN
11	E	269	ASN
4	5	454	GLN
4	5	676	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
13	ANP	2	901	-	29,33,33	2.84	6 (20%)	26,52,52	1.37	3 (11%)
13	ANP	3	1001	-	29,33,33	2.95	6 (20%)	26,52,52	1.20	3 (11%)
13	ANP	5	801	-	29,33,33	2.13	5 (17%)	26,52,52	1.14	1 (3%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
13	ANP	2	901	-	-	0/13/38/38	0/3/3/3
13	ANP	3	1001	-	-	1/13/38/38	0/3/3/3
13	ANP	5	801	-	-	0/13/38/38	0/3/3/3

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	2	901	ANP	PB-O3A	-3.00	1.55	1.59
13	3	1001	ANP	PB-O2B	-2.93	1.48	1.56
13	5	801	ANP	PB-O3A	-2.91	1.55	1.59
13	3	1001	ANP	PB-O3A	-2.82	1.55	1.59
13	3	1001	ANP	PG-O2G	-2.72	1.49	1.56

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	2	901	ANP	PA-O3A-PB	-5.62	112.30	132.71
13	5	801	ANP	PA-O3A-PB	-4.40	116.73	132.71
13	3	1001	ANP	PA-O3A-PB	-3.95	118.36	132.71
13	3	1001	ANP	O3G-PG-O1G	-2.56	106.85	113.58
13	2	901	ANP	O3G-PG-O1G	-2.19	107.81	113.58

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
13	3	1001	ANP	O1B-PB-N3B-PG

There are no ring outliers.

3 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	2	901	ANP	9	0
13	3	1001	ANP	10	0
13	5	801	ANP	9	0

## 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.