Published online 5 September 2016

Nucleic Acids Research, 2016, Vol. 44, No. 19 ***9505–9517***
doi: 10.1093/nar/gkw789

Atomic structure of an archaeal GAN suggests its dual roles as an exonuclease in DNA repair and a CMG component in DNA replication

Takuji Oyama[1,*], Sonoko Ishino[2], Tsuyoshi Shirai[3], Takeshi Yamagami[2], Mariko Nagata[2], Hiromi Ogino[2], Masami Kusunoki[1] and Yoshizumi Ishino[2,*]

[1]Faculty of Life and Environmental Sciences, University of Yamanashi, 4-4-37 Takeda, Kofu, Yamanashi 400-8510, Japan, [2]Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka, Fukuoka 812-8581, Japan and [3]Department of Computer Bioscience, Nagahama Institute of Bio-Science and Technology, Nagahama, Shiga 526-0829, Japan

Received July 04, 2016; Revised August 29, 2016; Accepted August 29, 2016

ABSTRACT

In eukaryotic DNA replication initiation, hexameric MCM (mini-chromosome maintenance) unwinds the template double-stranded DNA to form the replication fork. MCM is activated by two proteins, Cdc45 and GINS, which constitute the 'CMG' unwindosome complex together with the MCM core. The archaeal DNA replication system is quite similar to that of eukaryotes, but only limited knowledge about the DNA unwinding mechanism is available, from a structural point of view. Here, we describe the crystal structure of an archaeal GAN (GINS-associated nuclease) from *Thermococcus kodakaraensis*, the homolog of eukaryotic Cdc45, in both the free form and the complex with the C-terminal domain of the cognate Gins51 subunit (Gins51C). This first archaeal GAN structure exhibits a unique, 'hybrid' structure between the bacterial RecJ and the eukaryotic Cdc45. GAN possesses the conserved DHH and DHH1 domains responsible for the exonuclease activity, and an inserted CID (CMG interacting domain)-like domain structurally comparable to that in Cdc45, suggesting its dual roles as an exonuclease in DNA repair and a CMG component in DNA replication. A structural comparison of the GAN–Gins51C complex with the GINS tetramer suggests that GINS uses the mobile Gins51C as a hook to bind GAN for CMG formation.

INTRODUCTION

DNA replication is essential for all living organisms, and the basic mechanism is conserved across the three domains of life, Bacteria, Archaea, and Eukarya. DNA replication must occur accurately in a highly coordinated manner regulated by numerous proteins (1). At the initiation of DNA replication, the parental double-stranded DNA (dsDNA) is unwound to generate two single-stranded DNAs (ssDNAs), which form a replication fork. In eukaryotes, the hetero-hexameric MCM (mini-chromosome maintenance) comprising MCM2–7 acts as the helicase core to unwind the template DNA (2,3). Although isolated MCM generally exhibits weak helicase activity, two protein factors, Cdc45 and GINS, have been identified as MCM activators, and they form the CMG complex holoenzyme together with MCM (4,5). CMG is constructed on the DNA by the sequential loading of MCM (as the double-hexameric ring at the formation of Pre-RC; pre-replicating complex), Cdc45 (at the formation of Pre-IC; pre-initiation complex) and GINS (at the Pre-LC; pre-loading complex after pre-IC), rather than the loading of the pre-assembled complex (3,6,7). CMG formation is controlled by two protein kinases, CDK (cyclin-dependent kinase) and DDK (dbf4-dependent kinase), and is accompanied by the co-recruitment of numerous 'firing factors' such as Cdt1, Sld3, Sld7, Sld2, Dpb11, some of which are targets of phosphorylation by CDK and DDK. After the pre-IC formation, CMG becomes activated on a larger assembly, which is formed by the further recruitment of factors such as Mcm10 and DNA polymerase ϵ, and others.

The recent development of high-resolution cryo-electron microscopy (cryo-EM) is remarkable, and in the best cases protein structures can be determined at near-atomic resolution (8). This advanced technique has been applied

*To whom correspondence should be addressed. Tel: +81 92 642 4217; Fax: +81 92 642 3085; Email: ishino@agr.kyushu-u.ac.jp
Correspondence may also be addressed to Takuji Oyama. Tel: +81 55 220 8828; Fax: +81 55 220 8828; Email: takujio@yamanashi.ac.jp
Present address: Hiromi Ogino, Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan.

to the three-dimensional (3D) structure analyses of yeast MCM and CMG, which accelerated the elucidation of the structure-function relationship of CMG (9,10). The cryo-EM CMG structure revealed the possible interaction modes among the protein factors, in addition to the detailed architecture of each factor. The crystal structures of the CMG components provided important key information to build the entire atomic models into the EM maps, and thus the combinatorial approach of cryo-EM and X-ray crystallography is a powerful technique to investigate huge macromolecular structures at atomic resolution.

The archaeal proteins involved in DNA replication and repair are more similar to those from Eukarya than from Bacteria. Furthermore, archaeal protein complexes generally possess simplified compositions, as compared to the eukaryotic systems. Therefore, studies of the archaeal DNA replication system could provide good models to understand the more complicated eukaryotic mechanisms (11–14). This correspondence of archaeal proteins with their eukaryotic counterparts also occurs in the case of the CMG helicase (Figure 1). For example, the archaeal MCM is the homo-hexamer (15,16), while the eukaryotic counterpart is the hetero-hexamer, and the crystal structures of archaeal MCM have been used as the starting models to analyze the yeast cryo-EM replisomes containing MCM. Similarly, the archaeal GINS have been found as the $\alpha_2\beta_2$-type hetero-tetramer (17–19) and the α_4-type homo-tetramer (20), and constitute architectures similar to that of the eukaryotic GINS, composed of four different subunits (21–23). Although it has been a curious mystery that a Cdc45 homolog has not been identified in archaeal genomes, bioinformatics analyses proposed that RecJ, an ssDNA-specific 5′-3′ exonuclease responsible for DNA repair, could be the Cdc45 homolog (24,25). This hypothesis has been strongly supported by recent functional and structural studies. First, an archaeal GINS from *Thermococcus kodakaraensis* stimulated the activity of its cognate exonuclease (TK_RS06185), which was designated as GAN (GINS-associated nuclease) (26). Second, the crystal structure of human Cdc45 (27) partly resembles those of the bacterial RecJ enzymes from *Thermus thermophilus* (28,29) and *Deinococcus radiodurans* (30).

As in the case of eukaryotic MCM, archaeal MCM alone exhibits weak helicase activity, suggesting that accessory factors, like Cdc45 and GINS in Eukarya, should also work in Archaea to stimulate the MCM helicase activities (31). As mentioned above, there is strong evidence that archaeal RecJ/GAN, primarily working in DNA repair, could also be involved in DNA replication initiation, as an activator of MCM together with GINS. Nonetheless, our knowledge of archaeal RecJ/GAN is limited, particularly due to the lack of the 3D structure. Here we report the crystal structure of an archaeal GAN from *T. kodakaraensis*. The structure of GAN is similar to those of *T. thermophilus* RecJ (TthRecJ) and *D. radiodurans* RecJ (DraRecJ), with strict conservation of the catalytic residues, indicating that the molecular mechanism of GAN as an exonuclease is similar to that of bacterial RecJ. In addition, GAN and eukaryotic Cdc45 share a domain that probably functions specifically in these proteins as an MCM activator. Therefore, GAN could be an intermediate between bacterial type RecJ and eukaryotic Cdc45, and possess dual roles as an exonuclease in DNA repair and a CMG component in DNA replication. We also determined the atomic structure of GAN in complex with the C-terminal domain of Gins51 (Gins51C) of the cognate $\alpha_2\beta_2$-type GINS tetramer. In this complex structure, Gins51C is bound to the N-terminal domain of GAN. This binding feature is comparable to that observed in the yeast CMG helicase cryo-EM structure. Gins51C contacts GAN using a partly common surface that is also used when it is attached to the GINS tetramer. This binding mode suggests that GINS may catch GAN by using Gins51C as a hook, for the formation of the CMG helicase.

MATERIALS AND METHODS

Overexpression and purification of the recombinant GAN–Gins51C

The recombinant GAN–Gins51C complex was produced by co-expression of the two genes for GAN and Gins51C (Supplementary Methods) in *Escherichia coli* BL21-CodonPlus(DE3)-RIL cells (Agilent Technologies). The cells were grown in LB medium, containing 50 μg/ml kanamycin, 50 μg/ml ampicillin, and 34 μg/ml chloramphenicol at 37°C to an OD_{600} of 0.5, and then the gene expression was induced by 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) with further cultivation at 18°C for 18 h. The cells were harvested and were disrupted by sonication in buffer A (50 mM Tris-HCl, pH 8.0, and 0.5 M NaCl) including 20 mM imidazole. The soluble extracts were heated at 80°C for 20 min. The heat-resistant fraction was subjected to chromatography on a HisTrap HP column (GE Healthcare), which was developed with a linear gradient of 20–500 mM imidazole in buffer A. The eluted protein fractions were dialyzed against buffer B (50 mM Tris-HCl, pH 8.0, and 0.5 mM EDTA), and were applied to a HiTrap Heparin HP column (GE Healthcare). The column was developed with a linear gradient of 0–1 M NaCl in buffer B. The eluted complex fractions were applied to a Superdex200 16/60 column (GE Healthcare), and was eluted with buffer B containing 0.5 M NaCl. The protein concentration was determined by measuring the absorbance at 280 nm. The theoretical extinction coefficient is 38 850 M^{-1} cm^{-1}. To produce the selenomethionine (Se-Met) derivative of GAN–Gins51C, the *E. coli* cells harboring the two expression plasmids were grown in a minimal medium containing seleno-L-methionine at a final concentration of 25 μg/ml. The Se-Met protein was expressed and was purified using the same procedure as for the wild type complex.

Preparation of the wild type and mutant GAN proteins

E. coli BL21-CodonPlus(DE3)-RIL cells transformed with pET-GAN (Supplementary Methods) were cultured in LB medium, containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol, at 37°C to an OD_{600} of 0.4, and the gene expression was induced by adding IPTG to a final concentration of 1 mM, followed by further cultivation for 16 h at 25°C. The cells were resuspended in buffer C (50 mM Tris-HCl, pH 8.0, 0.5 mM DTT and 0.1 mM EDTA) containing 1 M NaCl, and sonicated. The soluble cell extract

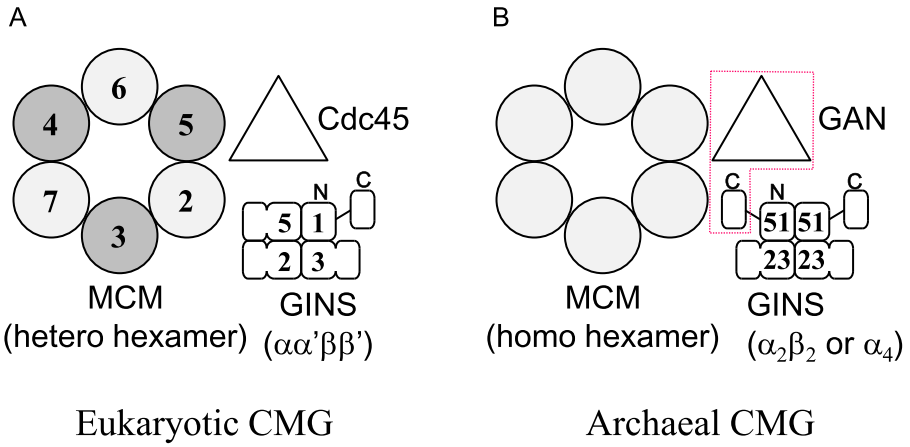

Figure 1. Schematic representations showing the subunit compositions of eukaryotic and archaeal CMG helicases. (**A**) Eukaryotic MCM. Each MCM subunit shown by a circle is generally composed of the N-terminal helical domain and the C-terminal AAA+ ATPase domain. Each MCM subunit is numbered. Cdc45 and GINS are considered to activate MCM by binding to the 'MCM2–MCM5 gate'. GINS comprises Sld5, Psf1, Psf2 and Psf3, and all of them are divided into two domains. The C-terminal domain of Psf1 is mobile. Cdc45 is shown by a triangle. (**B**) Archaeal CMG. MCM is a homo-hexamer, and GINS is either an $\alpha_2\beta_2$-type hetero-tetramer composed of Gins51 and Gins23, or a Gins51-only α_4-type homo-tetramer (the figure shows the $\alpha_2\beta_2$-type). Cdc45 is replaced with GAN. The target of the present study is indicated by the red dotted box.

was heated at 80°C for 30 min. The heat-resistant fraction was treated with 0.15% (v/v) polyethyleneimine. The soluble fraction was precipitated by 80% saturated ammonium sulfate, and resuspended in buffer C containing 1 M ammonium sulfate. The soluble fraction was applied to a HiTrap Phenyl HP column (GE Healthcare), and eluted with a linear gradient of 1–0 M ammonium sulfate in buffer C. The fractions containing GAN were dialyzed against buffer C, and were applied to a HiTrap Q HP column (GE Healthcare). The column was developed with a linear gradient of 0–1 M NaCl in buffer C. The protein concentration was calculated by measuring the absorbance at 280 nm. The theoretical molar extinction coefficient is 34 380 M^{-1} cm^{-1}.

Limited proteolysis of GAN–GINS full-length complex

Trypsin was added in a 1:200 weight ratio to a solution of the purified GAN–GINS full-length complex, and the mixture was incubated at room temperature. Aliquots were removed at various time intervals and analyzed by gel filtration chromatography using a SMART system and a Superdex 200 PC 3.2/30 column (GE Healthcare) (Figure 2A). Molecular mass markers used were thyroglobulin (670 K), immunoglobulin (158 K), ovalbumin (44 K) and myoglobin (17 K) (Bio-Rad). Aliquots of the fractions were subjected to 5–20% gradient SDS-PAGE followed by Coomassie Brilliant Blue (CBB) staining. For crystallization, the digestion reaction was stopped at 10 min by adding an excess amount of PMSF, and then the truncated GAN–Gins51C complex was purified by anion exchange column chromatography. The Hitrap Q column (5 ml) was equilibrated with buffer D (50 mM Tris-HCl, pH 8.0, 0.5 mM DTT, 0.1 mM EDTA, and 10% (v/v) glycerol). The trypsin-digested complex was loaded onto the column, which was developed with a linear gradient of 0.2–0.8 M NaCl in buffer D. Fractions containing the GAN–Gins51C complex were pooled and concentrated to about 10 mg/ml, using an Amicon Ultra 4 centrifugal concentrator (molecular-weight cut-off 3000, Millipore).

Surface plasmon resonance analysis

A Biacore J (GE Healthcare) system was used to study the physical interaction between GAN and GINS component proteins. All experiments were conducted at 25°C. Purified recombinant GAN and His-tagged Gins51 were immobilized to CM5 sensor chip and Ni-NTA sensor chip, respectively, according to the manufacturer's recommendation. To analyze the binding properties of GINS component proteins, 2 μM GINS (as a heterotetramer), 4 μM Gins51N (as a monomer), and 4 μM Gins23 (as a monomer) were applied to the immobilized GAN for 2 min in a buffer containing 10 mM HEPES, pH 7.4, 0.15 M NaCl, 3 mM EDTA, and 0.2% Tween20. To estimate the binding affinities, various concentrations (5, 10, 25, 50, 100, 250 and 500 nM) of GAN wild-type (WT) and its mutant (F55A/L57A) were applied to the immobilized Gins51 for 2 min in a buffer containing 10 mM HEPES, pH7.4, 0.15 M NaCl, and 0.2% Tween20. The apparent affinity was evaluated with BIAevaluation software (GE Healthcare).

Analysis of the nuclease activity of GAN

The recombinant GAN protein (10 nM) and the various GINS component proteins (0–40 nM as a monomer) was incubated with 1 μM 3′-FITC-labeled deoxyoligonucleotide (dCGAACTGCCTGGAATCCTGACGAACTGTAG) in a reaction mixture, containing 25 mM Bis-Tris, pH 6.9, 2 mM $MnCl_2$, 20 mM NaCl, 0.125 mg/ml BSA, and 0.05% NP-40 at 60°C for 2 min. The reactions were stopped with 10 mM EDTA, and the samples were subjected to 15% PAGE in 0.5× TBE (45 mM Tris, 45 mM boric acid and 1.25 mM EDTA, pH 8.3). The products were visualized and quantified with an image analyzer, Typhoon Trio+ and ImageQuant TL (GE Healthcare).

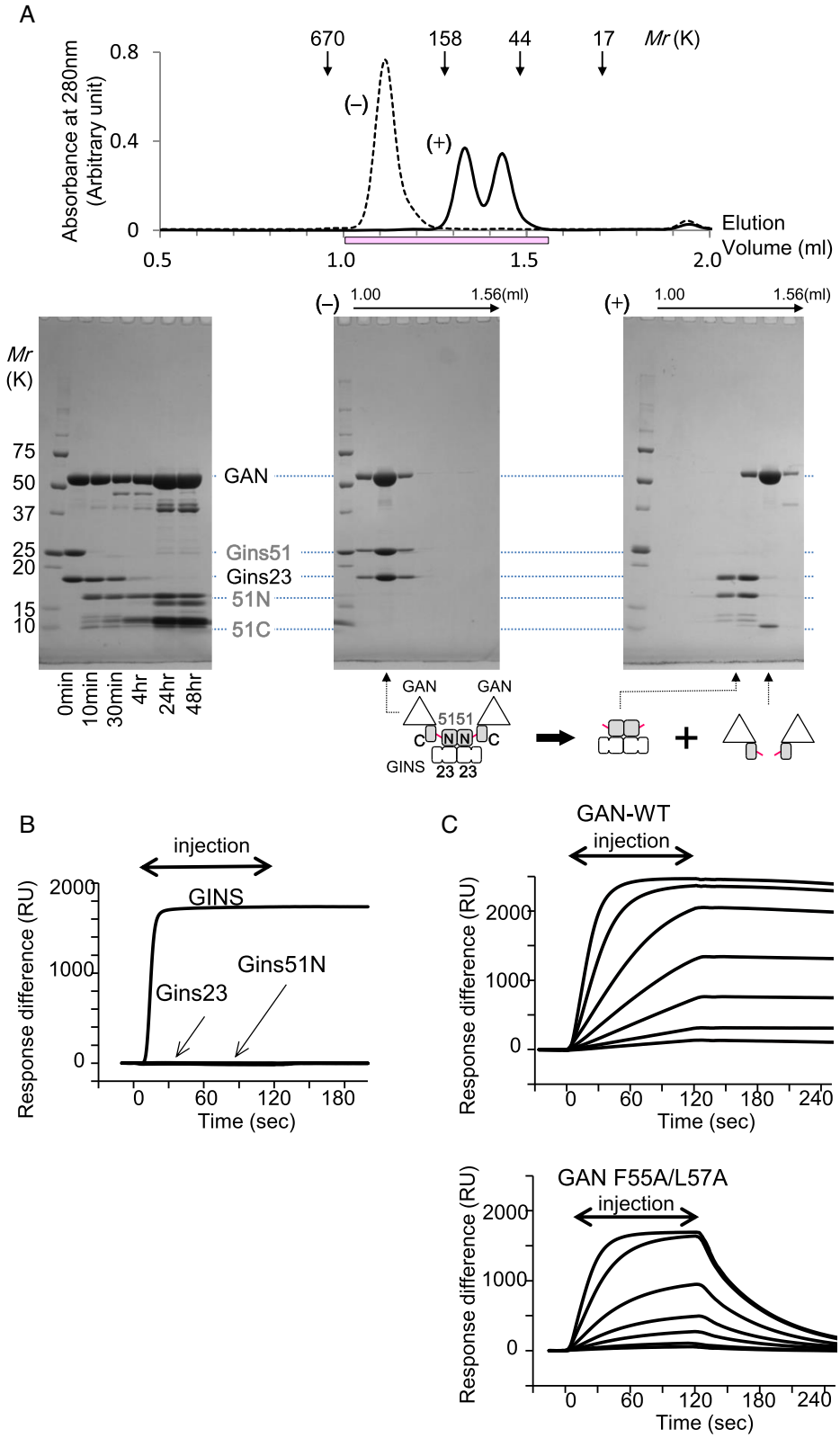

Figure 2. Biochemical analyses of the GAN–GINS complex. (**A**) Trypsin digestion of the GAN–GINS full-length complex. Bottom left: The complex was digested with a 1:200 ratio of trypsin to the complex, and aliquots were taken at the indicated times and analyzed by 5–20% gradient SDS-PAGE. Molecular mass standards are shown on the left. Top: The intact full-length complex ('−', dotted line) and the trypsin-cleaved complex ('+', thick line) were analyzed by gel filtration chromatography. The elution volumes of the marker proteins are indicated by arrows on the top. The corresponding SDS-PAGE analyses are shown at the bottom middle (−) and right (+) panels. Fractions are indicated by a bar. (**B**) SPR analyses of the GINS proteins with the GAN-bound chip. The loading proteins were indicated on the sensorgrams. (**C**) SPR analyses of the GAN-WT (upper panel) and the mutant (F55A/L57A) (lower panel) on the Gins51-immobilized chip. Seven different concentrations (5, 10, 25, 50, 100, 250 and 500 nM) of proteins were loaded.

Crystallization and structure determination

Crystallization screening was performed by the hanging-drop vapor diffusion method, using a semi-automated nano-liter dispenser Mosquito (TTP Labotech) with two screening kits (Crystal Screen HT and INDEX HT, Hampton Research). The initial crystals of trypsin-derived GAN–Gins51C were obtained from a few conditions with high molecular weight polyethylene glycols (PEG) as a precipitant at around neutral pH, and the optimization procedure yielded single crystals. Equal volumes of the protein solution and a reservoir solution, containing 100 mM MES, pH 7.0, 0.2 M ammonium sulfate, and 30% (w/v) PEG monomethyl ether 5000, were mixed and equilibrated against the reservoir solution at 293 K. Thin plate-shaped crystals grew within a week. X-ray diffraction data were collected using the synchrotron radiation source on BL38B1 of SPring-8 (Harima, Japan). The crystals were cryo-protected using 10% (v/v) glycerol supplemented to the reservoir solution, and diffracted X-rays to 3.2 Å resolution. The space group and unit cell constants indicated that the crystal contains one to two GAN–Gins51C complexes within the asymmetric unit. We attempted to determine the structure by the molecular replacement method, but failed probably due to the heterogeneity of the protein sample.

Both the recombinant wild-type and Se-Met derivatized GAN–Gins51C complexes were crystallized under almost the same conditions as those for the trypsin-digested complex. The crystals of the recombinant complexes diffracted X-rays up to 2.5 Å resolution on the beamlines PF AR-NW12 and AR-NE3A of the Photon Factory (Tsukuba, Japan). Diffraction data were processed by the program HKL2000 (32). The structure of the complex was determined by the single wavelength anomalous dispersion (SAD) method, using the Se-Met derivative data. The SAD calculations were performed by PHENIX.autosol, followed by density modification by Phenix.autobuild, in the PHENIX software suite (33). The initial atomic model was built with the program COOT (34). The atomic model was refined using the Phenix.refine program and rebuilt manually on COOT iteratively, and the refinement was converged with R/free-R = 21.2/24.9%. The crystal contains two GAN–Gins51C complexes located side-by-side in the asymmetric unit, with no rotational symmetry. While the first complex is well-ordered in the crystal, the second one contains the less-ordered C-terminal domain of GAN, which was removed from the final atomic model. The structure of the wild type complex was then determined, using the Se-Met derivative as the starting model. Overall, the Se-Met derivative had slightly better quality, and therefore the description of the structure is based on the Se-Met derivative complex.

The best crystals of GAN were obtained when equal volumes of the protein solution (12.8 mg/ml in buffer containing 50 mM Tris-HCl, pH 8.0, 400 mM NaCl and 0.1 mM EDTA) and a reservoir solution, containing 100 mM Bis-Tris, pH 5.5, 0.2 M ammonium acetate, and 30% (w/v) PEG 3350, were mixed and equilibrated against the reservoir at 293K. Although polygonal-shaped crystals grew to a length of 150–200 μm within a week, dissolution of the crystals was observed after further incubation. Therefore, the crystals were harvested before the dissolution started. The crystals were soaked in the reservoir solution for several seconds, transferred to the same buffer supplemented with 10% (v/v) PEG 400 for cryoprotection, and then quickly flash-cooled in liquid nitrogen. X-ray diffraction data were collected on BL38B1 of SPring-8 to 2.8 Å resolution. The GAN crystals contained one molecule in the asymmetric unit. The structure was determined by molecular replacement, using the GAN structure in the GAN–Gins51C complex as a probe. The correct solution was obtained when only the N-terminal DHH and CID-like domains were used. The atomic model of the C-terminal DHH1 domain was manually located on the electron density map during the refinement. The atomic model of GAN alone was refined until the convergence with R/free-R of 20.3/25.1% was attained. The statistics are summarized in Table 1. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB), under the accession codes 5GHR for the Se-Met derivative of GAN–Gins51C, 5GHS for the wild-type complex and 5GHT for GAN.

RESULTS AND DISCUSSION

Domain organization of GAN and GINS

RecJ, GAN, and Cdc45 are characterized by two catalytic domains, the N-terminal DHH domain and the following DHH1 domain, with seven conserved motifs, I to IV on DHH and VI and VII (or DHHA1 motif) on DHH1, and the active site is formed between the two domains (28–30). Cdc45 lacks nuclease activity, due to replacements of the catalytic residues (Supplementary Figure S1).

The eukaryotic GINS tetramer is composed of four different subunits, Sld5, Psf1, Psf2 and Psf3 (35), with a unique relationship among them, despite their low identity. Sld5 and Psf1 possess large α-helical domains at the N-termini and small β-stranded domains at the C-termini, and the order of these domains is inverted in Psf2 and Psf3. On the other hand, there are two types of GINS in Archaea, the $\alpha_2\beta_2$-type GINS contains two each of Gins51 and Gins23, which correspond to Sld5/Psf1 and Psf2/Psf3, respectively. The homotetrameric GINS comprises the Gins51 subunit only, but it forms the functional tetramer by using a long disordered loop between the α-helical and β-stranded domains (19,20). The crystal structures of the human and archaeal GINS indicated that the C-terminal domains in Psf1 and Gins51 are mobile, while the corresponding N-terminal domains of Psf2/Psf3/Gins23 and the C-terminal domain of Sld5 are required for the stabilization of the tetramer. The mobile C-terminal domains of eukaryotic Psf1 and archaeal Gins51, including the interdomain connecting loops, are responsible for the GINS functions in replication and chromatin binding (21,36,37).

Identification and biochemical analysis of the GAN–Gins51C complex

We first attempted to solve the structure of the GAN–GINS complex using the full-length protein subunits of *T. kodakarensis* Gins23, Gins51, and GAN. However, the crystals of the full-length complex diffracted X-rays poorly, in

Table 1. Data collection and refinement

Complex	GAN–Gins51C (Se-Met derivative)	GAN–Gins51C (wild type)	GAN
Data collection			
Beamline	PF AR-NW12	PF AR-NE3A	SPring-8 BL38B1
Space group	$P2_12_12_1$	$P2_12_12_1$	$I23$
Unit cell constants			
a (Å)	48.10	48.04	158.13
b (Å)	116.00	116.28	158.13
c (Å)	234.60	235.40	158.13
β (deg)	90.00	90.00	90.00
Wavelength (Å)	0.97970	1.0000	1.0000
Resolution (Å)[a]	50.0–2.50 (2.54–2.50)	50.0–2.60 (2.64–2.60)	50.0–2.80 (2.85–2.80)
No. of total reflections	323 254 (14 839)	301 862 (14 990)	228 756 (10 800)
No. of unique reflections	4 5946 (2090)	42104 (2082)	16 452 (777)
Completeness (%)	98.9 (92.5)	100.0 (100.0)	99.9 (98.0)
I/σ(I)	19.3 (3.4)	18.6 (3.9)	45.0 (6.2)
Redundancy	6.3 (7.1)	7.0 (7.2)	13.7 (13.9)
Mosaicity	0.19–0.41	0.50–0.83	0.58–0.88
R_{merge} (%) [b]	12.9 (47.3)	12.5 (56.1)	10.0 (55.5)
Refinement			
Resolution range (Å)	25.0–2.51	37.1–2.59	42.3–2.80
R_{work} [c] / R_{free} [d] (%)	21.2/24.9	22.8/26.6	20.3/25.1
Number of atoms			
Protein	7089	7183	3630
Water	60	44	–
Ligand (SO_4 ions)	25	25	–
Average B-factor (Å)			
Protein	42.1	35.6	48.5
Water	30.5	24.0	–
Ligand	68.6	60.1	–
r.m.s.d.			
Bond lengths (Å)	0.002	0.003	0.004
Angles (°)	0.51	0.71	0.67
PDB code	5GHR	5GHS	5GHT

[a]Values in parentheses correspond to the last shell.
[b]$R_{merge} = (\Sigma|I_i - < I_i >|) / \Sigma_i |I_i|$, where $< I_i >$ is the mean I_i over symmetry-equivalent reflections.
[c]$R_{work} = \Sigma||F_O| - |F_C|| / \Sigma|F_O|$ for all data, excluding data used to calculate R_{free}.
[d]$R_{free} = \Sigma||F_O| - |F_C|| / \Sigma|F_O|$ was calculated using 5% of the total reflections chosen randomly and omitted from the refinement.

spite of many trials to improve the crystal quality. We then searched for a truncated version of the GAN–GINS complex that retained the functional contact between them, by limited proteolysis using trypsin. Under controlled conditions, trypsin cleaved the interdomain loop of Gins51 to separate the GAN–Gins51C complex from the GINS tetramer (Figure 2A). The cleavage site of Gins51C by trypsin was detected at Arg130 by determination of the N-terminal sequence of the cleaved Gins51C. This corresponds to the domain organization determined by X-ray crystallography (19). We prepared the recombinant GINS component proteins (Supplementary Figures S2A and S2B). As shown in SPR analyses in Figure 2B, GINS complex bound to GAN tightly. On the other hand, Gins23 and Gins51N showed no responses to the immobilized GAN. These results clearly indicated that Gins51C is essential for the interaction of GINS with GAN. To estimate the binding affinities, we applied various concentrations (5, 10, 25, 50, 100, 250 and 500 nM) of GAN-WT and its mutant onto Gins51-immobilized chip (Figure 2C). In the dissociation phase (after 120-s injection), the curves of GAN-WT were almost horizontal, indicating strong interactions between GAN-WT and Gins51. Even 10 nM GAN caused tight binding with little dissociation. In the case of GAN mutant (F55A/L57A, described below) as a control, a substantial dissociation was observed. Then we purified and crystallized the trypsin-cleaved GAN–Gins51C complex. Although the crystals diffracted X-rays to 3.2 Å resolution, the structure could not be determined. Accordingly, we prepared the complex with recombinant Gins51C protein and crystallized it. The crystal of the recombinant complex diffracted X-rays up to 2.5 Å resolution, and the structure was determined (Table 1).

Overall view of the GAN–Gins51C complex

GAN adopts an unbalanced dumbbell-like shape, in which the two N-terminal domains, DHH (residues 1–168) and the following 'CID (CMG Interacting Domain)-like' domain (residues 169–288), together form a larger globule connected to a smaller one formed by the C-terminal DHH1 domain (residues 355–477) via a long α-helical linker region (residues 289–354) (Figure 3A). Both DHH and DHH1 fold into the α/β structure, in which the central five-stranded parallel β-sheet (DHH) or the mixed β-sheet (DHH1) is flanked by α-helices. The middle domain of GAN is topologically comparable to the corresponding CID domain of human Cdc45 (27), and herein is designated as the 'CID-like' domain. The CID-like domain possesses a four helix bundle and a two-stranded antiparallel β-sheet, which form

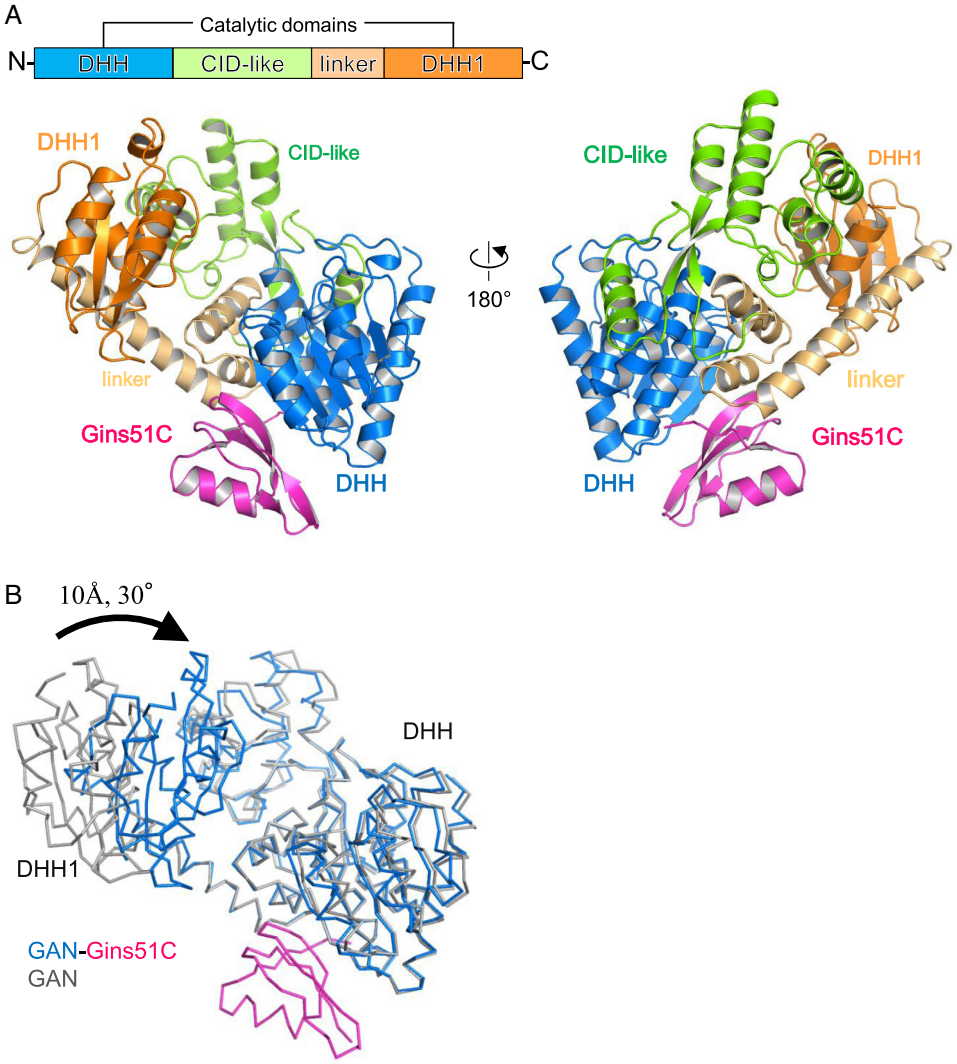

Figure 3. Crystal structures of the GAN–Gins51C complex and the isolated GAN. (**A**) Structure of GAN–Gins51C. Two orthogonal views are drawn by a ribbon representation, with the domain organization shown above the structure. The DHH, CID-like, and DHH1 domains of GAN are colored sky blue, green and orange, respectively, and the linker region is light orange. Gins51C is colored magenta. (**B**) Structure comparison between GAN in the free form (grey) and the Gins51C (magenta)-bound form (sky blue). DHH1 in the complex shifts toward the N-terminal DHH by approximately 10 Å, accompanied by a 30° rotation. Structural graphics were rendered using the PyMOL molecular graphics package (39).

the amino- and carboxyl- terminal strands in this domain (Figure 4B). The overall shape of GAN is similar to that of Cdc45, rather than the enzymatically related TthRecJ and DraRecJ, because of their different domain architectures (Figure 4). The locations of the conserved domains (DHH and DHH1) also differ among them, although these domains can be superimposed independently. DHH of GAN superimposed on those of TthRecJ, DraRecJ and Cdc45, with root mean square deviations (RMSDs) of 1.9, 1.8 and 2.0 Å for 161, 194 and 130 corresponding C_α atoms, respectively. DHH1 of GAN aligned with those of TthRecJ, DraRecJ and Cdc45, with RMSDs of 1.8, 1.8 and 2.1 Å for 89, 96 and 89 corresponding C_α atoms, respectively. In the GAN–Gins51C complex structure, Gins51C is bound to the N-terminal DHH of GAN via a mixture of hydrophobic and hydrophilic interactions. The structure of Gins51C is almost the same as that in the GINS tetramer. The two Gins51C structures superimposed with an RMSD of 0.7 Å for 53 C_α atoms.

Structural basis for the nuclease activity of GAN

The active site of RecJ/GAN is formed by several key residues on the conserved motifs of DHH, which coordinates one or two divalent metal ions such as Mg^{2+} or Mn^{2+} to stabilize the scissile phosphate and activate the nucleophilic-attacking water (28–30). These amino acid residues are conserved in GAN, suggesting that it cleaves DNA in a similar manner to other RecJ enzymes (Supplementary Figure S1). Although no metal ion was observed in the active site of the present GAN structure, the conserved residues are located suitable for Mg^{2+} or Mn^{2+} binding.

The crystal structure of DraRecJ has been solved in complexes with a mononucleotide or an ssDNA, which provided the detailed interactions between the enzyme and

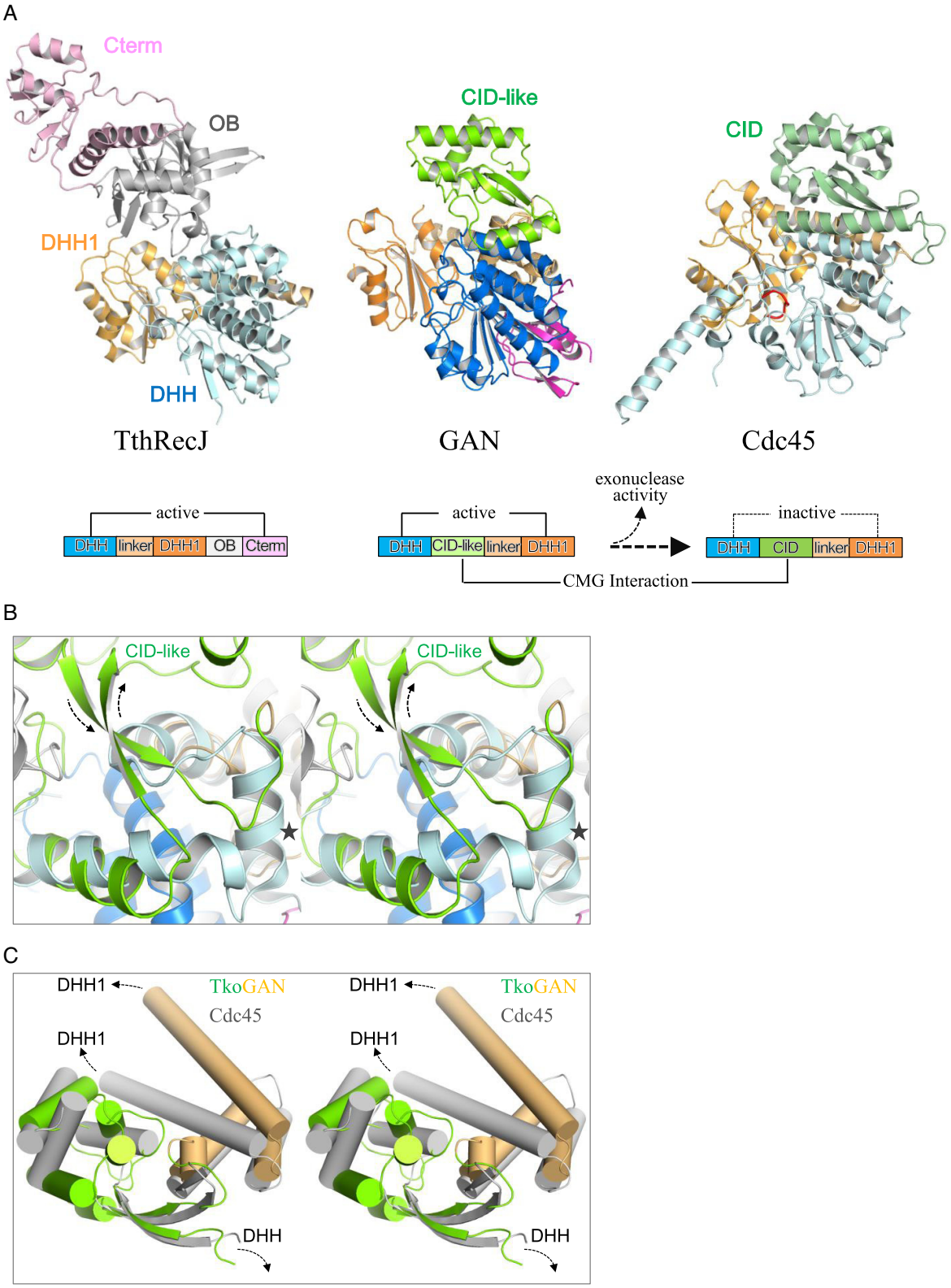

Figure 4. Structure comparison among RecJ, GAN, and Cdc45. (**A**) The three crystal structures are shown aligned with the N-terminal DHH domain. Domain arrangements are shown below the structure. The OB-fold domain (OB) and the C-terminal domain (Cterm) in TthRecJ (left) were colored grey and light pink, respectively. The DHH-binding loop on DHH1 of Cdc45 is highlighted in red. The CID/CID-like domain was inserted to the common ancestor of GAN (middle) and Cdc45 (right). During the evolution of GAN to eukaryotic Cdc45, mutations of the catalytic residues led to the loss of the enzymatic activity, and those on the CID-like domain have been remodeled to fit the eukaryotic CMG system. (**B**) Close-up view showing the position of the inserted CID-like domain in GAN. The entrance and exit segments of the CID-like domain form a twisted two-stranded antiparallel β-sheet. The corresponding region in TthRecJ is a single α-helix (marked by a star). (**C**) Comparison of the GAN CID-like domain and Cdc45 CID structures. α-Helices are shown by cylinders.

DNA (30). A structural comparison between this bacterial enzyme–DNA complex with GAN indicated that the DNA binding mode of GAN is partly similar to that of DraRecJ, particularly around DHH and DHH1. For example, three amino acid residues, Arg280 (motif V of the linker region), Arg373 (motif VI of DHH1) and His397 (motif VII of DHH1), are involved in DNA binding in the DraRecJ structure, and the corresponding residues, Arg307, Arg410 and His440 of GAN, are located in similar positions. Notably, a sulfate ion is bound to Arg410 of GAN, which may mimic the DNA backbone binding. The structure of the DraRecJ–DNA complex also revealed that its OB fold domain holds the DNA at the entrance of the active site, and the truncation of the OB fold domain impaired its activity (30). GAN lacks an OB fold domain (Supplementary Figure S1), and instead the CID-like domain occupies a similar spatial position to the OB fold domain of DraRecJ. It is unknown whether the CID-like domain participates in binding the substrate DNA.

A comparison of the structures of GAN in the free-form and the Gins51C-bound form revealed that DHH1 in the complex shifts toward DHH by approximately 10 Å with a 30 degree rotation (Figure 3B). The structures of DraRecJ in the DNA-bound and unbound forms showed that the enzyme exhibits the most closed state between DHH and DHH1 when bound to DNA (30). As Gins51C contacts the N-terminus side of the linker helix (Figure 5A), it is tempting to speculate that Gins51C binding to GAN causes the closure of the DHH–DHH1 gap, leading to the increased activity. However, the DHH1 movement seems to be caused by a kink in the middle of the linker helix (around Tyr335), because there is less structural change on the contact region of GAN. At the present time, we cannot rule out the possibility that this domain movement is due to a crystal packing effect. It should be noted that the distance between DHH–DHH1 in the GAN–Gins51C structure is still about 10 Å longer than that in the DraRecJ–DNA complex.

The nuclease activity of GAN is stimulated by both the GINS tetramer and the Gins51 subunit, indicating the functional interactions between them (26). Our experiments also showed that the nuclease activity of GAN was stimulated by GINS and Gins51, but not by Gins23 (Supplementary Figure S2C). Interestingly, neither Gins51N, Gins51C, nor a mixture of Gins51N and Gins51C stimulated the GAN nuclease activity, indicating Gins51N and Gins51C facilitates GAN in a cooperative manner, although Gins51C possesses a stronger binding affinity for GAN (Supplementary Figure S2D).

GAN could be an evolutionary intermediate from bacterial type RecJ to eukaryotic Cdc45

The first crystal structure of archaeal GAN revealed that the architecture of a long insertion between DHH and DHH1, with an unpredictable structure, was that of the CID-like domain. This fold is unique to the archaeal GAN and eukaryotic Cdc45, and is not present in bacterial RecJ (Supplementary Figure S1). The comparison of the domain arrangements among these proteins led to the hypothesis that archaeal GAN is an intermediate between bacterial type RecJ and Cdc45 (Figure 4A). The molecular phylogeny suggested that CID/CID-like domain was inserted to the common ancestor of GAN and Cdc45 (Supplementary Figure S1B). From the multiple sequence alignment, the common ancestor likely retained nuclease activity, since the catalytic residues required for metal-binding were conserved among RecJ and GAN. In the evolutionary process to form Cdc45, however, suppression of nuclease activity seemed to have occurred. The catalytic residues, Asp83, His106, and Asp166 (in numbering system of *T. kodakaraensis* GAN) have been replaced into Asn, Ser and Gln, respectively, on the branch leading to the Cdc45 (Supplementary Figure S1). A previous bioinformatics study also pointed out that several archaea possess inactivated versions of RecJ/GAN in their genomes, and they proposed a scenario for the evolution of RecJ homologs in Archaea and the origin of Cdc45 (24). The last common ancestor of all archaea probably had a GAN ortholog with active DHH domain, and archaeal ancestor of eukarya also retained the GAN ortholog, in which some of the catalytic residues of the DHH domain might be conserved. It is not easy to presume why the nuclease activity went lost to become an essential activator of MCM to form the CMG unwindosome in Eukarya during the evolution of the RecJ/GAN/Cdc45 family proteins toward Cdc45. Interestingly, a similar enzymatic dysfunction was hypothesized for DNA replisome in generating eukaryotic DNA polymerase ϵ from a tandem repeat of two different ancestral DNA polymerase Bs, where the essential metal-binding residues in one of the repeated DNA polymerase B were substituted in the eukaryotic DNA polymerase ϵ (38).

Although the fold of the CID-like domain of GAN was conserved in the CID domain of Cdc45, the difference in structural detail made a significant shift of relative spatial position between DHH and DHH1 domains (Figure 4C), suggesting that this domain also took a significant part in converting GAN into Cdc45. Another notable difference is observed between the GAN and Cdc45 DHH1 domains. Cdc45 possesses a unique loop (residues 539–544) in DHH1 which interacts with DHH, and the contact causes an occlusion of the DHH–DHH1 gap corresponding to the ssDNA binding site of RecJ (Figure 4A) (27). On the other hand, the corresponding region of GAN is involved in the β-sheet structure, as in the case of RecJ, and has no contact with DHH. Therefore, Cdc45 might have acquired the DHH contacting loop in DHH1, leading to the locked conformation observed in the crystal structure.

Gins51C interacts with the N-terminal DHH domain of GAN

The detailed interactions between GAN and Gins51C in the complex are shown in Figure 5A. The interactions are mainly hydrophobic, with an 839 $Å^2$ interface. The hydrophobic interactions involve Ile140, Met164, Ile166, Val184 and Ile186 of Gins51C with Leu313, Val315, Ala316, and Leu319 of GAN. The GAN–Gins51C contact also contains prominent hydrophilic interactions. In particular, the fourth β-strand of Gins51C hydrogen bonds with the second β-strand of GAN DHH in an antiparallel manner, to form an intermolecular seven-stranded mixed β-sheet structure. Asp163 on the fourth β-strand of Gins51C forms hydrogen bonds with both Gln56 and Ser58 on the neighboring β-strand of GAN. A similar interaction

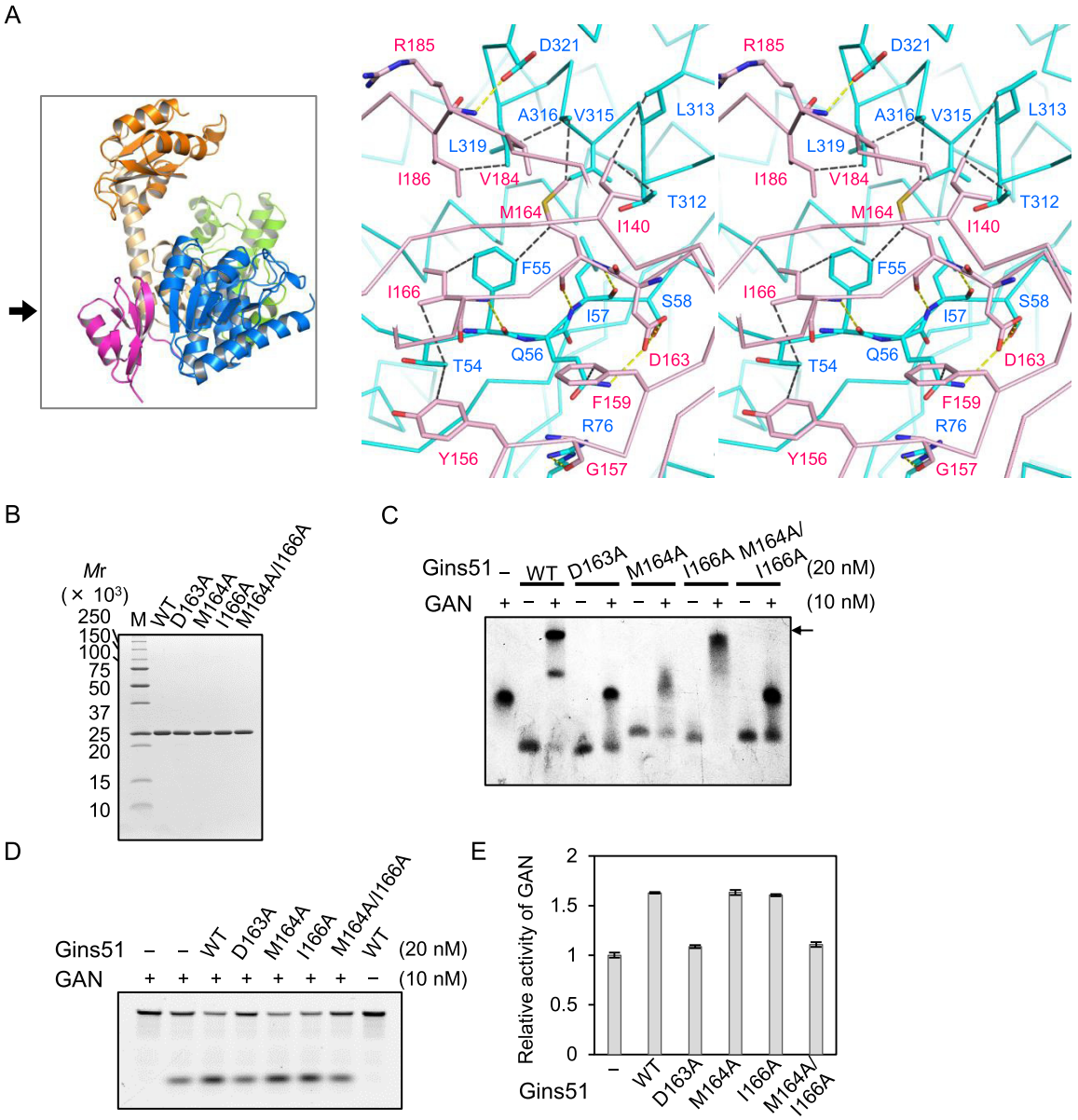

Figure 5. GAN–Gins51C interaction and mutational analyses. (**A**) Stereo view showing the detailed interactions between GAN and Gins51C. The viewpoint is indicated by the arrow in the overall structure. GAN is colored cyan and Gins51C is pink. Residues involved in the contacts are labeled. Grey and yellow dashed lines indicate hydrophobic and hydrophilic interactions, respectively. (**B**) The purified Gins51 wild-type (WT) and its mutant (D163A, M164A, I166A, and M164A/I166A) proteins (1 μg) were subjected to 12.5% SDS-PAGE, and were stained by CBB. M, molecular mass standards (Bio-Rad). (**C**) Physical interaction between GAN and Gins51 mutants. GAN-WT was incubated with Gins51-WT or its mutant at 60°C for 2 min, and analyzed by 6% PAGE, followed by silver staining. The names of the proteins were indicated on the top. The arrow indicates the 2:2 complex of GAN and Gins51. (**D**) The nuclease activities of GAN in the presence of Gins51-WT or its mutants are shown. GAN-WT was incubated with 3′-FITC-labeled 30-nt DNA in the presence of Gins51-WT or its mutants at 60°C for 2 min, and analyzed by 15% PAGE. (**E**) Quantification of the products shown in (D). Data are averages of three independent experiments with the standard error of the mean.

has been proposed for the yeast cryo-EM CMG structure, where the Psf1 C-terminal domain contacts Cdc45 (27).

The GAN–Gins51C interaction was analyzed by biochemical analyses using site-directed mutagenesis. Native-PAGE analysis showed that mutations at Asp163 and Met164/Ile166 on the contacting β-strand of Gins51C, abolished the complex formation with GAN (Figure 5B and C). Consistently, the same Gins51C mutants (D163A and M164A/I166A) lost the stimulating effect on the GAN nuclease activity (Figure 5D and E). Furthermore, mutations at residues, Phe55, Gln56, Leu57, Ser58, and Phe55/Leu57, the second β-strand of GAN DHH disrupted the complex formation (Supplementary Figure S3A, B). The SPR analysis showed the remarkable decreasing of the Gins51–GAN interaction by the mutation at Phe55/Leu57 in GAN (apparent KD was calculated to be 230 nM) (Figure 2C). On the other hand, the stimulation effects of Gins51 for the nuclease activity of mutant GANs were almost the same as wild-type GAN (Supplementary Figure S3C). However, the effect was completely lost when the Gins51C mutant

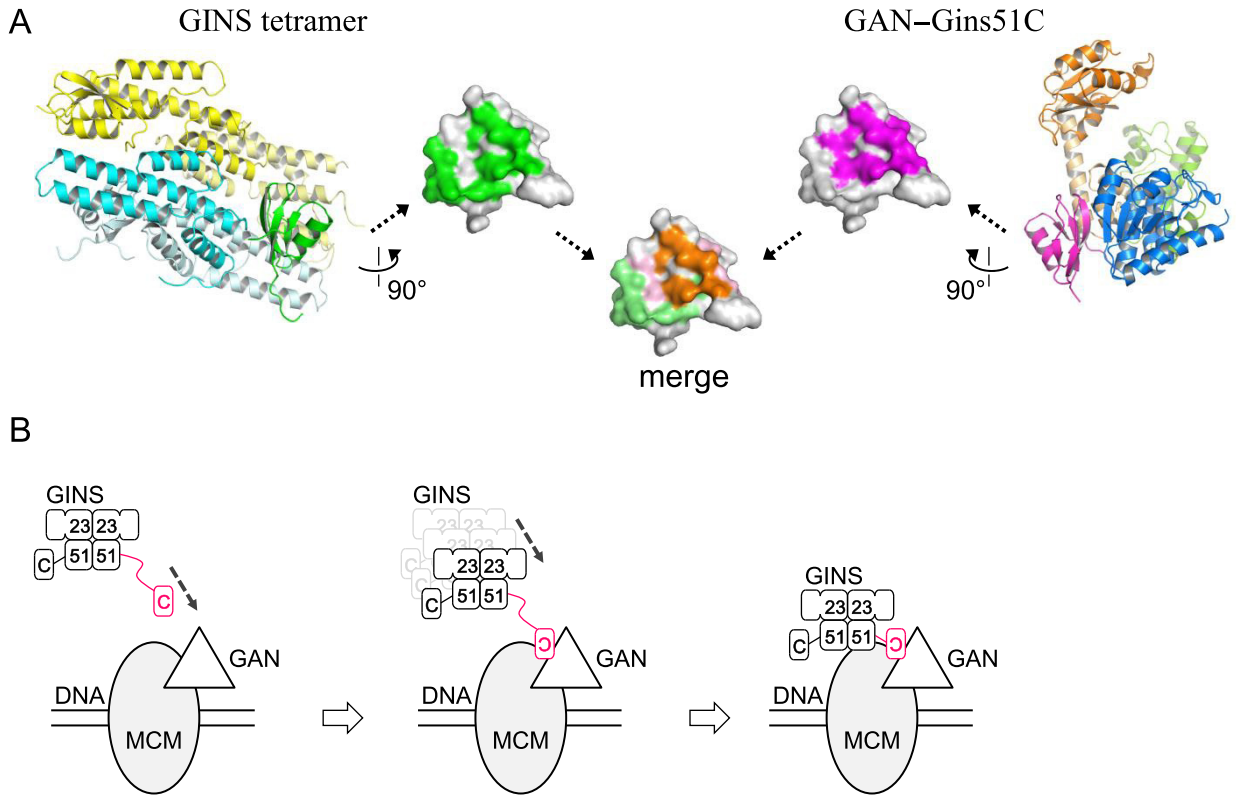

Figure 6. A model of GAN–GINS interaction on the formation of CMG. (**A**) GINS tetramer and GAN–Gins51C are shown as ribbon representations, in which the Gins51Cs in the two complexes are aligned by a 180° rotation about the Y-axis. Gins51 and Gins23 are colored cyan and yellow, respectively, and one Gins51C domain in the tetramer is highlighted in green. GAN–Gins51C is colored as in Figure 5A. Surface models of Gins51C, in which the residues involved in the contacts are colored. The central surface model (merge) shows the residues used commonly (orange) in both complexes. (**B**) A hypothetical model of GINS loading onto DNA to form CMG. Origin firing factors other than the CMG components are omitted for clarity. Gins51C detaches from GINS to bind to GAN, and then the GINS tetramer lands on the MCM–GAN complex by following the interdomain connecting loop.

(M164A) was used for the nuclease activity assays for the GAN mutants (Supplementary Figure S3D). These results collectively showed the functional importance of the observed interaction between GAN and Gins51C.

Gins51C could act as hook for the formation of CMG

Interestingly, several Gins51C residues involved in contacts with GAN also participate in the interaction with Gins51N in the GINS tetramer, with a contact area of 558 Å^2 (Figure 6A) (19). Although the detailed interaction scheme in the GINS tetramer is different from that observed in the complex with GAN, it is also mainly hydrophobic. The comparison of the Gins51C contacts in the two complexes suggested that a hydrophobic triangle could be formed by the key residues Met164, Ile166 and Ile186.

The contact modes of Gins51C in the two functional complexes indicated that Gins51C exclusively binds to either the GINS tetramer or GAN. Based on these observations, we propose a unique complex formation mode between GAN and GINS (Figure 6B). First, Gins51C detaches from the GINS tetramer to expose the buried hydrophobic surface for binding to GAN. Gins51C may bind only weakly to the GINS tetramer and move freely in solution while simply tethered to the tetramer by the linker loop, as in the case of the corresponding C-terminal domain of eukaryotic Psf1 (21). After hooking Gins51C onto GAN, the GINS tetramer reaches the contact site by following the linker loop and rebinding to Gins51C, leading to the formation of CMG. In the CMG complex, the GINS tetramer probably contacts Gins51C, in a manner similar to that observed in the yeast cryo-EM CMG structure, where GINS and GAN also interact with the MCM subunits (27). It is tempting to speculate that the functional importance of the disordered linker loop is to ensure the dynamic motion of Gins51C for the prior binding to GAN, in addition to the interactions with other proteins as previously reported (21,37).

Concluding Remarks

The present crystal structures of GAN alone and in its complex with Gins51C provide two structural insights into the GAN–GINS interaction toward CMG formation. First, the architecture of the archaeal GAN implies evolutionary relationships among RecJ, GAN and Cdc45. The acquisition of a CID-like domain led bacterial type RecJ to the archaeal GAN/eukaryotic Cdc45, which presumably allowed these family proteins to evolve to an activator of the MCM helicase, although it remains to be shown if GAN also functions as a member of the unwindosome in the archaeal cells. An interesting aspect is that if the separate evolution of

GAN and Cdc45 from a common ancestor may synchronize or follow the evolution of MCM, to form the functional homo-hexamer in Archaea and the hetero-hexamer in Eukarya. An evolutionary trace analysis of GAN and Cdc45 would be attractive, particularly if it is conducted by taking the co-evolution with MCM into account. Further detailed structure-function comparisons between GAN and Cdc45 will elucidate evolutional process of this protein family. Functional meaning of the 5′-3′ exonuclease in GAN, but not Cdc45, is an important issue to be clarified.

Second, Gins51C contacts the catalytic DHH domain of GAN, and part of this surface also interacts with the GINS tetramer. The facts that Gins51C binding to the GINS tetramer or GAN is exclusive and that Gins51C is connected to the GINS tetramer by a long loop suggest a unique GAN–GINS interaction, in which Gins51C is released to bind GAN, followed by the subsequent arrival of the GINS tetramer on the CMG unwindosome. Further studies such as mutational analyses of the GINS activity, the GAN–GINS interaction and the CMG formation, as well as structural studies of larger complexes, with CMG as the most attractive and challenging target, could provide more detailed and precise views of the formation and function of CMG in the archaeal DNA replication initiation system.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr. Tatsuya Nishino for encouragement during the course of this work and Ami Okamoto and Moena Watanabe for assistance with crystallization. We also thank Kohei Kobayashi, Eri Watatani, and Namiko Imai for the technical assistance of the mutational analyses. X-ray diffraction experiments were performed at the Photon Factory with the approval of the Photon Factory Program Advisory Committee (Proposal No. 2015G024), and at SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2015A1025, 2016A2543). This research was also partially supported by the NIG Collaborative Research Program (2014-A07, 2015-A101), and the Platform Project for Supporting in Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from the Japan Agency for Medical Research and Development (AMED). Finally, we are grateful to Dr. Patrick Forterre for the continuous encouragement in advancing this subject.

FUNDING

Ministry of Education, Culture, Sports, Science and Technology of Japan [60423133 to T.O. and 26242075 to Y.I.]; JSPS bilateral program (to Y.I.); Institute for Fermentation, Osaka (IFO) (to S.I. and H.O.); Grant-in-Aid for Japan Society for the Promotion of Science (JSPS) Fellows [201602633 to M.N.]; AMED Platform for Drug Discovery, Informatics, and Structural Life Science [16am0101042j0005 to T.S.]. Funding for open access charge: Ministry of Education, Culture, Sports, Science, and Technology [26242075].
Conflict of interest statement. None declared.

REFERENCES

1. Masai,H., Matsumoto,S., You,Z., Yoshizawa-Sugata,N. and Oda,M. (2010) Eukaryotic chromosome DNA replication: where, when, and how? *Annu. Rev. Biochem.*, **79**, 89–130.
2. Duderstadt,K.E. and Berger,J.M. (2013) A structural framework for replication origin opening by AAA+ initiation factors. *Curr. Opin. Struct. Biol.*, **23**, 144–153.
3. Tanaka,S. and Araki,H. (2013) Helicase activation and establishment of replication forks at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.*, **5**, a010371.
4. Gambus,A., Jones,R.C., Sanchez-Diaz,A., Kanemaki,M., van Deursen,F., Edmondson,R.D. and Labib,K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.*, **8**, 358–366.
5. Ilves,I., Petojevic,T., Pesavento,J.J. and Botchan,M.R. (2010) Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol. Cell*, **37**, 247–258.
6. Yelles,J.T., Deegan,T.D., Janska,A., Early,A. and Diffley,J.F. (2015) Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature*, **519**, 431–435.
7. Abid Ali,F. and Costa,A. (2016) The MCM helicase motor of the eukaryotic replisome. *J. Mol. Biol.*, **428**, 1822–1832.
8. Bartesaghi,A., Merk,A., Banerjee,S., Matthies,D., Wu,X., Milne,J.L. and Subramaniam,S. (2015) 2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor. *Science*, **348**, 1147–1151.
9. Li,N., Zhai,Y., Zhang,Y., Li,W., Yang,M., Lei,J., Tye,B.K. and Gao,N. (2015) Structure of the eukaryotic MCM complex at 3.8 Å. *Nature*, **524**, 186–191.
10. Yuan,Z., Bai,L., Sun,J., Georgescu,R., Liu,J., O'Donnell,M.E. and Li,H. (2016) Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat. Struct. Mol. Biol.*, **23**, 217–224.
11. Kelman,Z. and White,M.F. (2005) Archaeal DNA replication and repair. *Curr. Opin. Microbiol.*, **8**, 669–676.
12. Beattie,T.R. and Bell,S.D. (2011) Molecular machines in archaeal DNA replication. *Curr. Opin. Chem. Biol.*, **15**, 614–619.
13. Ishino,Y. and Ishino,S. (2012) Rapid progress of DNA replication studies in Archaea, the third domain of life. *Sci. China Life Sci.*, **55**, 386–403.
14. Makarova,K.S. and Koonin,E.V. (2013) Archaeology of eukaryotic DNA replication. *Cold Spring Harb. Perspect. Biol.*, **5**, a012963.
15. Brewster,A.S., Wang,G., Yu,X., Greenleaf,W.B., Carazo,J.M., Tjajadi,M., Klein,M.G. and Chen,X.S. (2008) Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 20191–20196.
16. Bae,B., Chen,Y.H., Costa,A., Onesti,S., Brunzelle,J.S., Lin,Y., Cann,I.K. and Nair,S.K. (2009) Insights into the architecture of the replicative helicase from the structure of an archaeal MCM homolog. *Structure*, **13**, 211–222.
17. Marinsek,N., Barry,E.R., Makarova,K.S., Dionne,I., Koonin,E.V. and Bell,S.D. (2006) GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep.*, **7**, 539–545.
18. Yoshimochi,T., Fujikane,R., Kawanami,M., Matsunaga,F. and Ishino,Y. (2008) The GINS complex from *Pyrococcus furiosus* stimulates the MCM helicase activity. *J. Biol. Chem.*, **283**, 1601–1609.
19. Oyama,T., Ishino,S., Fujino,S., Ogino,H., Shirai,T., Mayanagi,K., Saito,M., Nagasawa,N., Ishino,Y. and Morikawa,K. (2011) Architectures of archaeal GINS complexes, essential DNA replication initiation factors. *BMC Biol.*, **9**, 28.
20. Ogino,H., Ishino,S., Mayanagi,K., Haugland,G.T., Birkeland,N.K., Yamagishi,A. and Ishino,Y. (2011) The GINS complex from the thermophilic archaeon, *Thermoplasmaacidophilum* may function as a homotetramer in DNA replication. *Extremophiles.*, **15**, 529–539.
21. Kamada,K., Kubota,Y., Arata,T., Shindo,Y. and Hanaoka,F. (2007) Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nat. Struct. Mol. Biol.*, **14**, 388–396.

22. Choi,J.M., Lim,H.S., Kim,J.J., Song,O.K. and Cho,Y. (2007) Crystal structure of the human GINS complex. *Genes Dev.*, **21**, 1316–1321.
23. Chang,Y.P., Wang,G., Bermudez,V., Hurwitz,J. and Chen,X.S. (2007) Crystal structure of the GINS complex and functional insights into its role in DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 12685–12690.
24. Sanchez-Pulido,L. and Ponting,C.P. (2011) Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics*, **27**, 1885–1888.
25. Makarova,K.S., Koonin,E.V. and Kelman,Z. (2012) The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes. *Biol. Direct*, **7**, 7.
26. Li,Z., Pan,M., Santangelo,T.J., Chemnitz,W., Yuan,W., Edwards,J.L., Hurwitz,J., Reeve,J.N. and Kelman,Z. (2011) A novel DNA nuclease is stimulated by association with the GINS complex. *Nucleic Acids Res.*, **39**, 6114–6123.
27. Simon,A.C., Sannino,V., Costanzo,V. and Pellegrini,L. (2016) Structure of human Cdc45 and implications for CMG helicase function. *Nat. Commun.*, **18**, 11638.
28. Yamagata,A., Kakuta,Y., Masui,R. and Fukuyama,K. (2002) The crystal structure of exonuclease RecJ bound to Mn^{2+} ion suggests how its characteristic motifs are involved in exonuclease activity. *Proc. Natl. Acad. Sci.*, **99**, 5908–5912.
29. Wakamatsu,T., Kitamura,Y, Kotera,Y., Nakagawa,N., Kuramitsu,S. and Masui,R. (2010) Structure of RecJ exonuclease defines its specificity for single-stranded DNA. *J. Biol. Chem.*, **285**, 9762–9769.
30. Cheng,K., Xu,H., Chen,X., Wang,L., Tian,B., Zhao,Y. and Hua,Y. (2016) Structural basis for DNA 5′-end resection by RecJ. *eLife*, **5**, e14294.
31. Ishino,S., Fujino,S., Tomita,H., Ogino,H., Takao,K., Daiyasu,H., Kanai,T., Atomi,H. and Ishino,Y. (2011) Biochemical and genetical analyses of the three mcm genes from the hyperthermophilic archaeon, Thermococcus kodakarensis, *Genes Cells*, **16**, 1176–1189.
32. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
33. Adams,P.D., Afonine,P.V., Bunkoczi,G. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.*, **D66**, 213–221.
34. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Cryst.*, **D66**, 486–501.
35. Takayama,Y., Kamimura,Y., Okawa,M., Muramatsu,S., Sugino,A. and Araki,H. (2003) GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev.*, **17**, 1153–1165.
36. Sengupta,S., van Deursen,F., De Piccoli,G. and Labib,K. (2013) Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr. Biol.*, **23**, 543–552.
37. Ogino,H., Ishino,S., Oyama,T., Kohda,D. and Ishino,Y. (2014) Disordered interdomain region of Gins is important for functional tetramer formation to stimulate MCM helicase in *Thermoplasma acidophilum*. *Biosci. Biotechnol. Biochem.*, **79**, 432–438.
38. Tahirov,T.H., Makarova,K.S., Rogozin,I.B., Pavlov,Y.I. and Koonin,E.V. (2009). Evolution of DNA polymerases: an inactivated polymerase-exonuclease module in Pol epsilon and a chimeric origin of eukaryotic polymerases from two classes of archaeal ancestors. *Biol. Direct*, **4**, 11.
39. The PyMOL Molecular Graphics System, Version 1.6 Schrödinger, LLC.