Published online 27 September 2021

Nucleic Acids Research, 2022, Vol. 50, No. 7 ***3601–3615***
https://doi.org/10.1093/nar/gkab799

NAR Breakthrough Article

Family D DNA polymerase interacts with GINS to promote CMG-helicase in the archaeal replisome

Keisuke Oki[1], Mariko Nagata[1], Takeshi Yamagami[1], Tomoyuki Numata[1], Sonoko Ishino[1], Takuji Oyama[2,*] and Yoshizumi Ishino[1,*]

[1]Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 819-0395, Japan and [2]Faculty of Life and Environmental Sciences, University of Yamanashi, Kofu, Yamanashi 400-8510, Japan

Received June 24, 2021; Revised August 29, 2021; Editorial Decision September 01, 2021; Accepted September 06, 2021

ABSTRACT

Genomic DNA replication requires replisome assembly. We show here the molecular mechanism by which CMG (GAN–MCM–GINS)-like helicase cooperates with the family D DNA polymerase (PolD) in *Thermococcus kodakarensis*. The archaeal GINS contains two Gins51 subunits, the C-terminal domain of which (Gins51C) interacts with GAN. We discovered that Gins51C also interacts with the N-terminal domain of PolD's DP1 subunit (DP1N) to connect two PolDs in GINS. The two replicases in the replisome should be responsible for leading- and lagging-strand synthesis, respectively. Crystal structure analysis of the DP1N–Gins51C–GAN ternary complex was provided to understand the structural basis of the connection between the helicase and DNA polymerase. Site-directed mutagenesis analysis supported the interaction mode obtained from the crystal structure. Furthermore, the assembly of helicase and replicase identified in this study is also conserved in Eukarya. PolD enhances the parental strand unwinding via stimulation of ATPase activity of the CMG-complex. This is the first evidence of the functional connection between replicase and helicase in Archaea. These results suggest that the direct interaction of PolD with CMG-helicase is critical for synchronizing strand unwinding and nascent strand synthesis and possibly provide a functional machinery for the effective progression of the replication fork.

INTRODUCTION

Genomic DNA replication is essential for all living organisms, and its molecular mechanisms have been studied in the three domains of life, namely, Bacteria, Archaea and Eukarya (1–3). This accurate and tightly regulated process requires the replisome assembly, in which helicase unwinds parental double-stranded DNA (dsDNA) into single-stranded DNA (ssDNA). DNA polymerases then synthesize nascent strands (4). The eukaryotic and archaeal replicative helicases are named CMG-helicases, which come from their eukaryotic components Cdc45 (cell division cycle 45), MCM2–7 (mini-chromosome maintenance) and GINS (go–ichi–ni–san, which means 5–1–2–3 in Japanese as it consists of Sld5, Psf1, Psf2 and Psf3) (5,6). MCM2–7, helicase's catalytic core, belongs to the AAA^+ ATPase superfamily. ATP hydrolysis drives CMG-helicase to translocate on the leading-strand template DNA in the 3′–5′ direction (5,7,8). In the initial DNA replication step, the MCM2–7 double hexamer is loaded onto the dsDNA at the replication origin. It remains inactive during the assembly of other numerous replication factors, including Cdc45 and GINS (6,9,10). The Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK) phosphorylate MCM2–7, which then recruits Cdc45, followed by the pre-loading complex (pre-LC) comprising GINS and DNA polymerase ε (Polε) (11). Psf1 subunit of GINS stably binds Polε via its B-subunit Dpb2 to connect CMG-helicase to DNA polymerase (12). Polε is critical for CMG activation (13), and Dpb2 and the C-terminal domain of Pol2, the catalytic subunit of Polε, probably interact with MCM2–7 (14,15). CMG-helicase activation allows lagging-strand ejection, double-hexamer separation and bidirectional replication-fork progression (16). GINS also interacts with DNA polymerase α (Polα) via Ctf4 (17). Thus,

*To whom correspondence should be addressed. Tel: +81 92 802 4715; Fax: +81 92 802 4696; 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

the eukaryotic replisome retains at least two different DNA polymerases.

In Archaea, the replisome assembly mechanism, including helicase and polymerases, is still poorly understood. In contrast to the eukaryotic MCM2–7, most archaeal genomes contain only a single *mcm* homolog (18). Although *Thermococcus kodakarensis* has three *mcm* genes (*mcm1–3*) on its genome, only *mcm3* is essential for cell viability, and thereby, the MCM3 homo-hexamer appears to be the replicative helicase (19,20). Indeed, we previously demonstrated that GINS stimulates the ATPase and helicase activities of MCM3 (hereafter referred to simply as MCM), and GINS-MCM forms an archaeal CMG-complex with GINS-associated nuclease (GAN), a homolog of Cdc45 (21). The similar CMG-helicase complexes have been reported from *Sulfolobus* and *Saccharolobus* species (22,23). The archaeal GINS is a homotetramer (α_4) (24) or a tetramer formed by two types of subunits, Gins51 and Gins23 ($\alpha_2\beta_2$) (23,25,26). *T. kodakarensis* GINS belongs to the latter category. Gins51 is related to the eukaryotic Sld5 and Psf1, which possess large α-helical 'A domains' at the N-termini, small β-strand-rich 'B domains' at the C-termini and flexible linkers in between (26–29). By contrast, Gins23 resembles eukaryotic Psf2 and Psf3 and has a permuted domain arrangement (B–A instead of A–B). Our previous study revealed a striking structural similarity between *T. kodakarensis* and human GINS despite their low sequence identity (26). GAN is a Cdc45/RecJ homolog belonging to the DHH hydrolase superfamily (30). Although *T. kodakarensis* GAN does not exhibit nuclease activity in physiological conditions nor stimulates MCM activity, it forms a complex with MCM and GINS and is required for cell growth at high temperatures (21).

T. kodakarensis retains two DNA polymerases, namely, family B DNA polymerase (PolB) and family D DNA polymerase (PolD). PolB was first identified in *Pyrococcus furiosus* as a eukaryotic Polα-like family B DNA polymerase (31), and PolD was then found as a novel family based on its distinct amino acid sequence (32–34). In contrast to PolB, which possesses both 3′–5′ exonuclease and polymerase active sites in one polypeptide, PolD separates them in the DP1 and DP2 subunits, respectively. DP1 belongs to the calcineurin-like phosphodiesterase superfamily, like the Mre11 nuclease (35,36). The second subunit of all eukaryotic replicative polymerases (POLA2, POLE2 and POLD2) possess the conserved structures of DP1's core region, but they lost their 3′–5′ exonuclease activities. The catalytic subunits of the eukaryotic replicases also possess a similar structure to DP2's C-terminal domain, which binds DP1 (37,38). By contrast, DP2's catalytic core is not similar to other DNA polymerases but rather homologous to the double-psi beta-barrel (DPBB) family of RNA polymerases (39). Recent electron-microscopic (EM) studies unveiled 3D structures of PolD from *Pyrococcus abyssi* and *T. kodakarensis*, in complex with DNA and the proliferating cell nuclear antigen (PCNA), a processivity factor (40,41). We reported very recently that PolD connects primase to the archaeal replisome and that partner exchange from primase to PCNA on DP2 switches PolD's function from *de novo* synthesis to processive elongation (42).

Despite accumulating knowledge on the archaeal replisome, the mechanism by which archaeal replicative polymerase coordinates with CMG-helicase remains unclear. We show here the molecular basis for archaeal replisome assembly, in which GINS connects PolD to CMG-helicase. The N-terminal domain of DP1 subunit (DP1N) bound to the C-terminal domain of Gins51 subunit (Gins51C), which we previously reported to bind GAN. Thereby, we determined an atomic structure of the DP1N–Gins51C–GAN ternary complex, which should act as the replisome core in *T. kodakarensis*. A structural comparison of *T. kodakarensis* DP1N–Gins51C–GAN with a corresponding region of the yeast replisome revealed that the structural basis for the interaction with CMG-helicase is conserved between PolD and Polε. We also provide functional insights into the assembly of PolD and CMG-helicase. GINS bound to two PolD molecules, which presumably contributes to synchronization of DNA replication in both leading and lagging strands. PolD-binding increased the helicase activity and the ATPase activity of the archaeal CMG-helicase by up to one order of magnitude, completely changing its substrate preference from dsDNA to ssDNA. These findings will provide insight into the archaeal replisome assembly for the effective progression of replication fork.

MATERIALS AND METHODS

Recombinant proteins preparation

We prepared recombinant PolB, encoded by TK_RS00010 (43); DP1, by TK_RS09520 (44); DP2, by TK_RS09525 (44); PolD (DP1 and DP2) (42); MCM (Mcm3), by TK_RS08085 (21); GINS, by TK_RS02640 (Gins51) and TK_RS08080 (Gins23) (21); Gins51C (amino acids 138–188 of Gins51) (45); GAN wild type and D36A (nuclease deficient mutant), by TK_RS06185 (46); and PCNA (PCNA1), by TK_RS02635 (43) as described previously.

The pET-TaqLF expressing the large fragment of *Thermus aquaticus* DNA polymerase (PolI) (referred to as 'TaqLF' in this study; it is a truncated mutant of the 5′–3′ exonuclease domain of Taq polymerase by TTHTAQP1A, with residues 1–867 deleted) was prepared. The gene for TaqLF fragment was amplified from pTV-Taq (47) using Pfu DNA polymerase with the primer set 'TaqLF-F/R' (Supplementary Table S1). The amplified fragment was digested using NdeI and NotI restriction enzymes and ligated using the T4 DNA ligase into the corresponding sites of the pET-28a(+) expression vectors (Novagen). The nucleotide sequence was confirmed by sequencing using CEQ 2000XL (Beckman Coulter). *Escherichia coli* BL21-CodonPlus (DE3)-RIL cells (Agilent Technologies) containing recombinant plasmids were cultivated at 37°C in LB medium containing 50 μg/ml streptomycin and 34 μg/ml chloramphenicol at 37°C until OD_{600} reached 0.3. The cloned gene expression was induced by adding IPTG to a final concentration of 1 mM and further cultivated at 25°C for 16 h to overproduce TaqLF. The protein was purified using heat treatment at 80°C for 20 min, followed by polyethyleneimine treatment and ammonium sulfate precipitation. Sequential chromatography on HiTrap Butyl FF, HiTrap Heparin HP

and HiTrap Q HP columns (GE Healthcare) were conducted to obtain the homogeneous protein. The gene encoding the DP1N (amino acids 1–64 of DP1) was amplified by PCR using pET24a-DP1 as a template, and the primer set, DP1-F/DP1N-R (Supplementary Table S1). The amplified DNA was digested using NdeI-NotI and was inserted into the corresponding sites of the modified pCDF-1b, resulting pCDF-DP1N. NcoI-recognition sequence was changed to NdeI-recognition sequence in the modified pCDF-1b. The DP1N protein was produced in *E. coli* BL21 CodonPlus (DE3)-RIL cells, which were cultivated at 37°C in LB medium containing 50 μg/ml streptomycin and 34 μg/ml chloramphenicol at 37°C until OD_{600} reached 0.3. Gene expression was induced by adding IPTG to a final concentration of 1 mM and the cells were further cultivated at 25°C for 16 h. The protein was purified using heat treatment at 80°C for 20 min, followed by polyethyleneimine treatment and ammonium sulfate precipitation. Column chromatography was conducted sequentially on HiTrap Butyl HP and HiTrap SP HP columns (GE Healthcare).

The DP1N–Gins51C–GAN ternary complex was produced by co-expressing its three genes (DP1N, Gins51C and GAN) in *E. coli* BL21-CodonPlus (DE3)-RIL cells. The cells carrying pCDF-DP1N, pET28TEV-Gins51C (with an N-terminal His6-tag) and pET21a-GAN were cultivated in combination in LB medium containing 50 μg/ml streptomycin, 50 μg/ml kanamycin, 50 μg/ml ampicillin and 34 μg/ml chloramphenicol at 37°C until OD_{600} reached 0.3. The gene expression was induced by adding IPTG to a final concentration of 1 mM and the cells were further cultivated at 25°C for 18 h. The complex was purified using heat treatment at 80°C for 20 min, followed by sequential chromatography on HisTrap HP, HiTrap Q HP and Hiload 26/60 Superdex 200 pg columns (GE Healthcare). The complex-containing fraction was pooled and concentrated to 47 mg/ml for crystallization using a HiTrap Q HP column.

The GINS mutants (V148A/V170A and V148D/V170D) and DP1N mutants (L7A/L12A and L7D/L12D) were prepared using *E. coli* containing the corresponding genes. Each mutation was introduced by PCR-mediated site-directed mutagenesis and the produced proteins were purified in the same way as wild type proteins.

DNA substrate preparation

The pBlueScript II SK(+) (pBS) plasmid was used as a primary substrate for obtaining the fork-structured DNA with a long dsDNA region. The oligonucleotide 'Xho1-NH2' (Supplementary Table S1) was annealed to pBS and digested using Xho1 to linearize pBS. After the purification of linear DNA by Wizard SV Gel and PCR Clean-Up System (Promega), the primer 'polyApri45' (Supplementary Table S1) was annealed to linear pBS and then the primer extension was handled to make fork-structured dsDNA. The products have a 2855 bp dsDNA region with 106 nt of ssDNA at the 3′-terminus to anneal primers or loading enzymes. After cleaning up the DNA, 5′-Cy5-labeled priHel2 (Supplementary Table S1) was annealed and used as the substrate.

Surface plasmon resonance (SPR) analysis

The physical interaction of PolD with GINS was assessed using a Biacore J system (GE Healthcare). PolD was immobilized on a CM5 sensor chip (GE Healthcare) according to the manufacturer's recommendation. To measure the kinetic parameters, 1000 nM of GINS was passed through a running buffer (50 mM Tris–HCl, pH 8.0, 0.3 M NaCl, 0.1% IGEPAL) through the PolD-immobilized chip at a continuous flow of 30 μl/min at 25°C for 120 s. The regeneration buffer (50 mM Tris–HCl, pH 8.0, 1 M NaCl, 0.1% IGEPAL) failed to eliminate the bound analytes.

Yeast two-hybrid assay

The interaction of the indicated proteins was detected using a Y2H detection system (Matchmaker Gold Yeast Two-hybrid System, Matchmaker GAL4 Two-Hybrid System 3, Clontech). The plasmid pGBKT7, encoding the GAL4 DNA-binding region, and the plasmid pGADT7, encoding the activation domains, were used to prepare plasmids containing the indicated genes from *T. kodakarensis*. According to the manufacturer's protocol, the yeast Y2H Gold cells were co-transformed with pGBKT7 and pGADT7 (Clontech Matchmaker manual). Cell suspensions of each strain (3 μl of 2×10^6 cell/ml) were spotted onto synthetic defined plates without Leu and Trp for the non-selection and Leu, Trp and His or Leu, Trp, His, and Adenine for two different selection strengths. The agar plates were incubated at 30°C for 4 days, and cells showing the interactions of the two proteins produced from the two co-transformed plasmids were grown.

Native-PAGE gel shift assay

The protein-protein interaction was also detected using native-PAGE. Each protein and mixture were incubated for 2 min at 60°C, and then mixed with 5× gel-loading buffer (15% Ficoll and 0.1% bromophenol blue) and subjected to native-5% PAGE/TBE and Coomassie Brilliant Blue (CBB) staining.

Gel filtration chromatography

Gel filtration chromatography was performed using the SMART system (Amersham Pharmacia). In the experiments testing the PolD–GINS and PolD–GINS–GAN complex formation, each protein and mixture were incubated for 3 min at 60°C. Aliquots (20 μl) of the protein solution was applied on a Superose 6 PC 3.2/30 column (GE Healthcare) and eluted with a buffer containing 50 mM Tris–HCl, pH 8.0 and 0.15 M NaCl. Aliquots (0.1 μl) of the applied solution and the aliquots (4 μl) of each fraction from the eluates were subjected to 10% SDS-PAGE containing WIDE RANGE Gel Preparation Buffer (Nacalai Tesque), followed by silver staining. In these experiments testing the DP1N–Gins51C interaction, each protein (120 μM) and mixture were incubated for 5 min at 60°C, and were applied on a Superdex 200 PC 3.2/30 column (GE Healthcare). The aliquots (4 μl) of each fractionated solution eluted with a buffer containing 50 mM Tris–HCl, pH 8.0 and 0.3 M NaCl were applied Tricine–SDS–

10%T, 2.6%C PAGE, followed by CBB staining. The standard marker proteins, including thyroglobulin (670 000), γ-globulin (158 000), ovalbumin (44 000) and myoglobin (17 000), were also subjected to the same gel filtration as controls.

Pull-down assay

A pull-down experiment was conducted using DP1N-producing *E. coli* cell extract and purified His-tagged Gins51C. DP1N-producing BL21-CodonPlus (DE3)-RIL cells (from 150 ml of LB medium) were re-suspended in 3 ml of binding buffer (50 mM Tris–HCl, pH 8.0, 0.5 M NaCl) and sonicated to prepare the cell extract. Gins51C (22 μl, 240 μM) was bound to 40 μl of Ni-NTA agarose at 25°C for 1 h. After incubation, a 400 μl aliquot of cell lysate was added. After 5 min of incubation at 60°C, the unbound fraction was separated. The resin was washed using 600 μl of binding buffer thrice, and the bound protein was eluted with 40 μl of binding buffer containing 0.5 M imidazole. The eluates were subjected to Tricine–SDS–10%T, 2.6%C PAGE, followed by CBB staining.

Crystallization and structure determination

We crystallized the DP1N–Gins51C–GAN ternary complex using the hanging-drop vapor diffusion method. The initial screening was performed using the semi-automatic dispensing machine Mosquito (TTP Labtech) with the commercially available crystallization screening reagents Crystal Screen HT and the INDEX HT (Hampton Research). Crystallization drops were prepared by mixing equal volumes (200 nl) of the protein solution and a reservoir solution and then were incubated against the reservoir at 293 or 277 K. Long and rectangular-shaped single crystals were obtained from a reservoir INDEX H07 containing 0.15 M DL-malic acid, pH 7.0 and 20% (w/v) PEG3350 at 293 K. The crystals were thin and fragile. To get thicker and more stable crystals, optimizing the crystallization condition was set up manually. The second screening successfully found larger crystals by mixing 2 μl of the protein solution (17 mg/ml) with 1 μl of a reservoir containing 0.15 M DL-malic acid, pH 7.0, 1 mM tris (2-carboxyethyl) phosphine and 14% (w/v) PEG3350 at 293 K. Although the crystals were still thin, their quality was ehough for high-resolution structure determination. X-ray diffraction data were collected on BL26B1 of SPring-8, Harima, Japan. The crystals were soaked into a reservoir solution supplemented with 20% (v/v) glycerol for seconds, and flash-frozen in liquid nitrogen and maintained the cryogenic temperature using a cold nitrogen gas stream (100 K) on the beamline. The diffraction spots were recorded on an EIGER X 4M detector with a fine-slicing (0.1°/frame, a total of 1800 frames) collection strategy. The data were processed using XDS (48), POINTLESS (49) and AIMLESS (50) programs (Supplementary Table S2). There were two ternary complexes in the asymmetric unit, having a Matthews coefficient of 2.67 and a solvent content of 54%. The crystal structure of DP1N–Gins51C–GAN was determined by molecular replacement using the PHENIX program package (51). A polyaniline model of the previously determined GAN–Gins51C binary complex crystal structure (PDB code: 5GHS) was used as a probe and the two complexes were correctly located in the asymmetric unit. The position of the C-terminal DHH1 domain of GAN, which is connected to the N-terminal DHH domain through the long linker helix, was drastically different (Supplementary Figure S1). Therefore, DHH1 was manually fitted into the electron density map. Furthermore, there were clear electron density maps corresponding to DP1N, and the atomic models of *Pyrococcus horikoshii* DP1N (PDB code 2KXE (52)) were fitted into the map to complete the ternary complex model construction. The overall atomic model was built using the COOT program (53) and refined using PHENIX. Manual model correction and refinement were iterated until convergence (Supplementary Table S2).

Leading-strand replication assay

A leading-strand replication assay was conducted by detecting the primer extension activities of polymerases using alkaline agarose gel electrophoresis. The reactions were performed as follows unless otherwise indicated. For the helicase loading reaction, the substrate DNA (10 nM) was mixed with 100 nM MCM (as the hexamer), 200 nM GINS and 200 nM GAN (D36A) in 20 μl of a buffer containing 25 mM Tris–HCl at pH 8.0, 5 mM $MgCl_2$, 5 mM DTT and 50 nM 5′-Cy5-labeled priHel2 DNA. After 5 min of incubation at 70°C, each of 100 nM PolB, 100 nM TaqLF or 200 nM PolD and 200 nM PCNA (as the trimer) were added and further incubated at 70°C for 3 min. Next, the reaction was started by adding dNTP and ATP to achieve a final concentration of 0.125 mM dNTP (each) and 5 mM ATP, followed by incubation at 70°C for 10 min. The aliquots (8 μl) of the reaction mixture were added to 3 μl of stop solution (12.5% Ficoll, 100 mM EDTA and 0.1% Orange G). The products were analyzed on a 1.5% alkaline agarose gel in 50 mM NaOH and 1 mM EDTA and then visualized by using a Typhoon Trio+ image analyzer.

ATPase assay

The reaction mixtures (50 μl) containing 25 mM Bis–Tris, pH 7.0, 50 mM NaCl, 5 mM $MgCl_2$, 2.5 mM ATP, 2 mM DTT and 100 nM MCM (as the hexamer) were incubated for 10, 20 and 30 min at 70°C in the presence or absence of 40 μM (in nucleotides) DNA (M13 mp18 ssDNA or M13 mp18 RF DNA). The effects of GAN and GAN–GINS, and PolD were analyzed by adding 0–0.4 μM GINS, GAN and PolD (as the tetramer, monomer and heterodimer, respectively) to the reaction. According to the manufacturer's protocol, the orthophosphate produced in the reactions was quantified with an EnzChek Phosphate Assay kit (Thermo Fisher Scientific). The standard error of the mean (SEM) was calculated from three independent experiments.

Helicase assay

The helicase activity was measured with trap DNA in 20 μl of reaction mixture containing 25 mM Tris–HCl at pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 0.1% Triton X-100, 2.5 mM ATP, 10 nM DNA substrate (Cy5-hel1/hel2, Supplementary Table S1), 250 nM trap DNA,

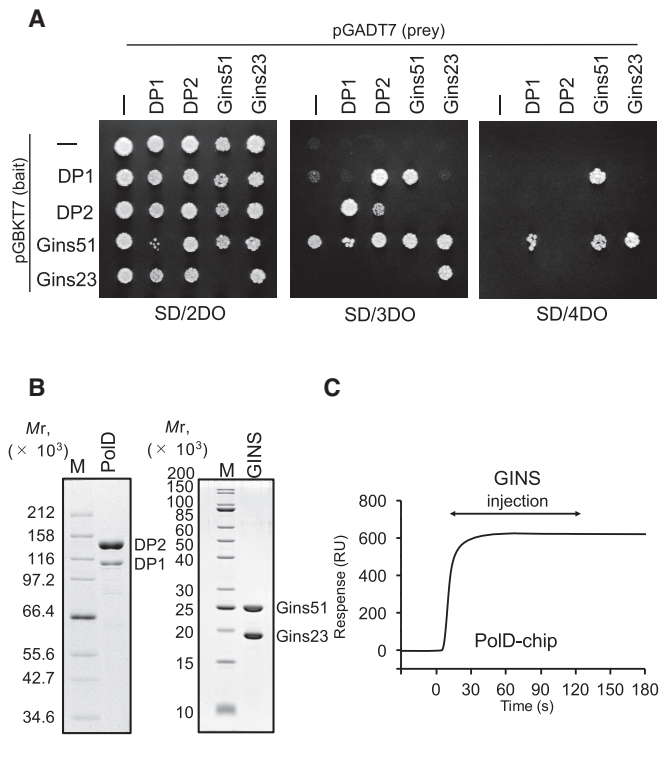

Figure 1. PolD stably binds GINS via DP1 and Gins51. (**A**) Yeast two-hybrid assay on PolD and GINS showing the interaction between PolD (DP1 and DP2) and GINS (Gins51 and Gins23). Cell suspensions of each strain were spotted onto SD plates without Leu and Trp (SD/2DO, left) for the non-selection plate or Leu, Trp and His (SD/3DO, middle), and Leu, Trp, His and Adenine (SD/4DO, right) for the selection plates with two different selection strength. The minus sign indicates the transformants with the bait or prey plasmid without insert DNA. Notably, no matter how many attempts that were made, the combination of pGBKT7-Gins23 and pGADT7-Gins51 did not produce any transformant even from non-selected plates. (**B**) Analysis of purified recombinant protein by SDS-PAGE. Purified PolD (2.8 μg) and GINS (2 μg) were subjected to SDS-7.5% PAGE and SDS-12% PAGE, respectively, followed by CBB staining. Protein size markers were run in lane M, and their sizes are indicated on the left side of the gel. (**C**) SPR analyses of GINS with the PolD-bound chip. 1000 nM of GINS was loaded onto the PolD-immobilized chip for 120 s and washed with a running buffer.

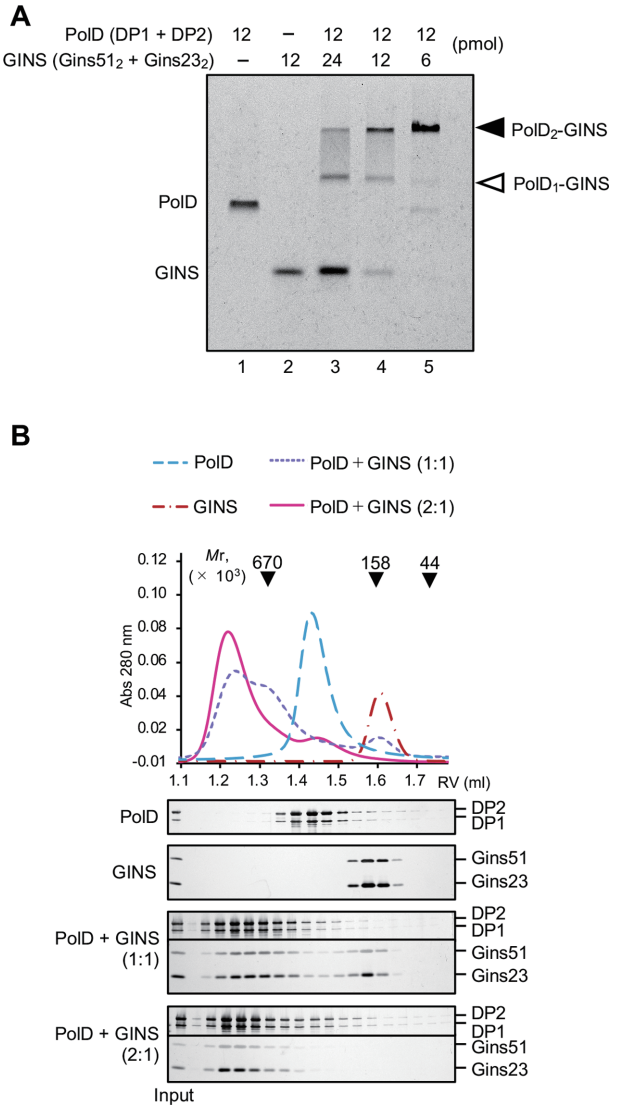

Figure 2. The GINS complex binds to two PolD molecules. (**A**) Stoichiometric ratio of PolD and GINS analyzed by native-PAGE. The amounts of proteins are indicated at the top of the gel. Black and white arrowheads indicate the $PolD_2$–GINS and $PolD_1$–GINS complexes, respectively. (**B**) Complex formation of PolD and GINS analyzed by gel filtration chromatography. A fixed amount of PolD (2.4 μM) was mixed with 2.4 and 1.2 μM GINS. The elution profiles, monitored by the absorbance at 280 nm, are shown. The peak positions of the marker proteins are indicated on the top. Aliquots of each fraction were subjected to SDS-PAGE, followed by silver staining.

100 nM MCM (as the hexamer), 400 nM GINS (as the tetramer), 400 nM GAN (D36A, exonuclease deficient mutant), 400 nM PolD and 400 nM DP1. After incubation at 65°C for 20 min, the reactions were stopped by adding 5 μl of stop solution (12.5% Ficoll, 100 mM EDTA, 0.5% SDS and 0.1% Orange G) and the aliquots (5 μl) were separated by 10% PAGE in TBE supplemented with 0.1% SDS. The results of the electrophoresis were visualized by using a Typhoon Trio+ image analyzer.

RESULTS

PolD, but not PolB, interacts with GINS

Genetic analyses showed that PolD is essential for the cell viability of *T. kodakarensis* whereas PolB is not, and disrupting the *polB* gene does not affect cell growth (54,55). These reports imply that PolD, but not PolB, works in the replisome as a replicative DNA polymerase. In fact, affinity purification experiments co-isolated PolD and GINS from *P. abyssi* (56) and *T. kodakarensis* (57) cell extracts. More recently, an *in vitro* analysis showed that PolD's DP1 subunit and GINS' Gins51 subunit from *Thermococcus* sp. 4557 interact with themselves (58). Our yeast two-hybrid (Y2H) experiment supports the earlier report of DP1–Gins51 interaction, as well as the DP1–DP2 interaction for PolD and the Gins51–Gins23 interaction for GINS complexes (Figure 1A). By contrast, the same Y2H analysis did not reveal interactions between PolB and GINS subunits (Supplementary Figure S2). We next purified the recombinant PolD and GINS (Figure 1B) and analyzed them using SPR.

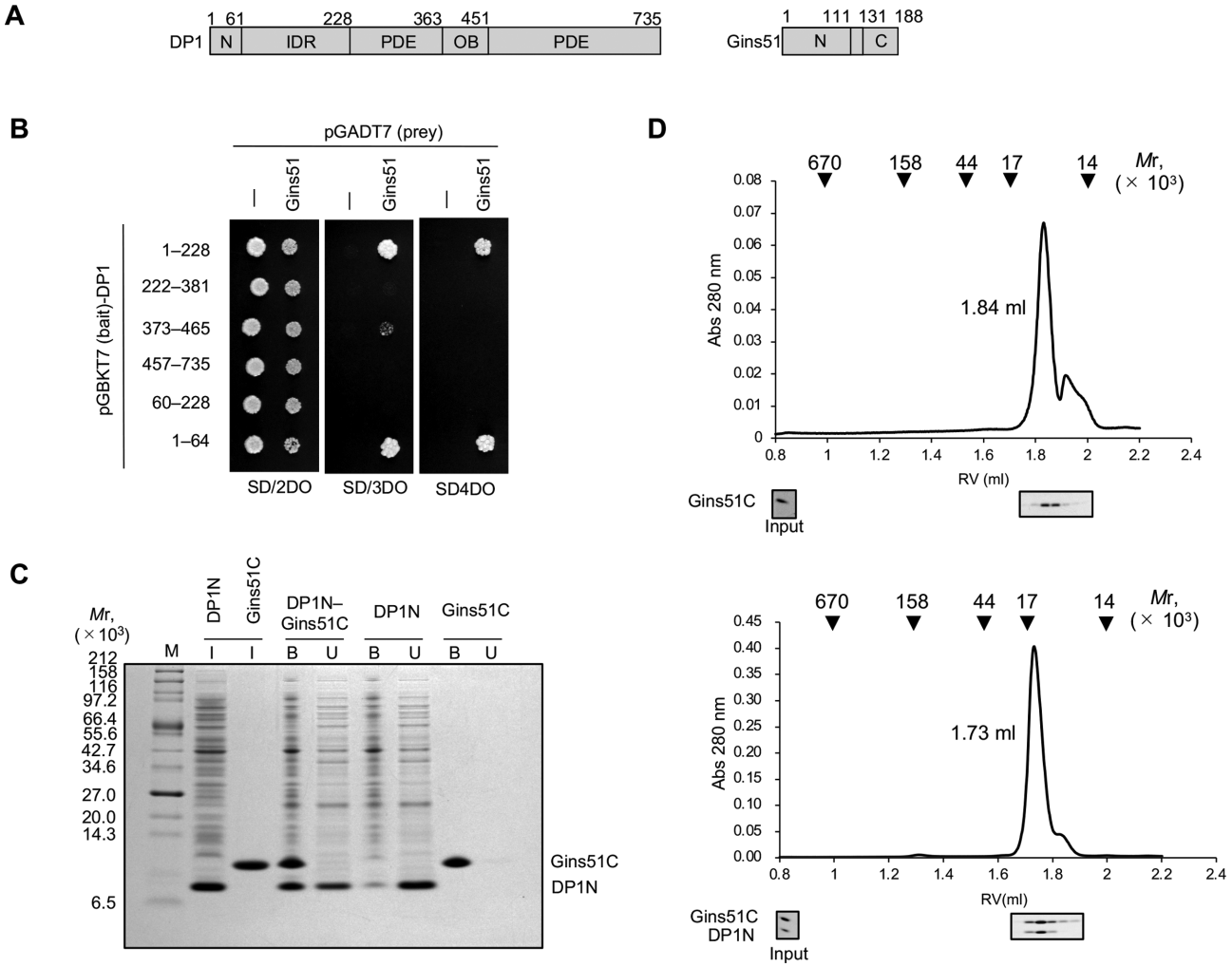

Figure 3. DP1's N-terminal domain interacts with Gins51's C-terminal domain. (**A**) Domain organization of DP1 and Gins51. N: N-terminal, IDR: intrinsically disordered region, PDE: phosphodiesterase, OB: oligonucleotide/oligosaccharide binding, C: C-terminal. (**B**) Yeast two-hybrid assay showing the interactions between Gins51 and DP1 fragments. (**C**) DP1N was tested for interaction with His-tagged Gins51C on Ni-NTA resin and analyzed by CBB-stained SDS-PAGE. Lane I is the input fraction containing DP1-producing cell lysate and purified Gins51C. Lane B and U are bound and unbound fractions, respectively. (**D**) Complex formation of Gins51C and DP1N analyzed by gel filtration chromatography. Purified Gins51C (top) and a mixture of DP1N and Gins51C (bottom) were analyzed. Aliquots of each fraction were subjected to Tricine-SDS-10%T,2.6%C PAGE, followed by CBB staining.

We loaded GINS onto the PolD-immobilized chip and obtained a sensorgram, which reached a plateau in a few seconds, and the continuous flow solution showed no dissociation (Figure 1C). These data strongly suggest that the PolD–GINS complex is distinctly stable, and PolD, but not PolB, forms a complex with CMG-helicase in the replisome.

Two PolD molecules bind to GINS

We next analyzed the stoichiometry of PolD and GINS in the complex. Similarly, the eukaryotic Polε directly binds GINS via the interaction of Dpb2 and Psf1 (12). However, contrary to the eukaryotic hetero-tetrameric GINS, *T. kodakarensis* GINS is a tetramer of two Gins51 subunits and two Gins23 subunits, conferring the potential for interacting with two PolD copies per GINS. As expected, native-PAGE analysis showed that the mixture of PolD and GINS formed two retarded bands (indicated by the black and white arrowheads in Figure 2A), considered to reflect the PolD–GINS complexes with 2:1 and 1:1 ratios, respectively. We further mixed PolD and GINS at 1:1 and 2:1 molar ratios (corresponding to lanes 4 and 5 in Figure 2A, respectively) and subjected them to analytical gel filtration chromatography. As shown in Figure 2B, PolD and GINS co-eluted earlier than individual proteins. The peak height corresponding to the PolD–GINS complex clearly increased when half amount of the GINS was mixed with PolD. The apparent molecular weight of the complex from the elution position is much higher than the size marker of 670,000, although the theoretical molecular weight of $PolD_2$–GINS from amino acid sequences is 543,451.4. This difference could be due to the intrinsically disordered regions in DP1 (44). Therefore, we conclude that GINS can connect two PolD molecules. We speculate that GINS links two PolD molecules to CMG-helicase to synthesize both leading and lagging strands. A weak positive signal showing the DP2–DP2 interaction was observed in the Y2H experiment (Figure 1A). Although our previous research identified that

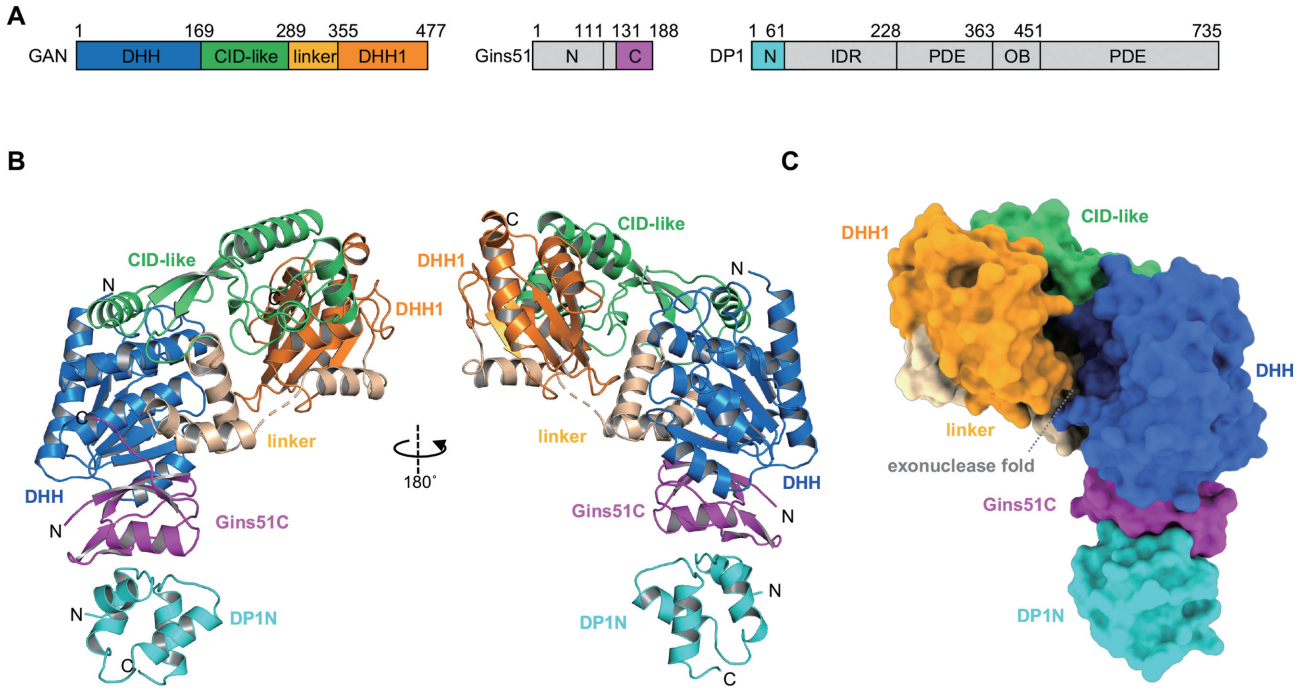

Figure 4. Crystal structure of the DP1N–Gins51C–GAN ternary complex. (**A**) Domain organizations of GAN, Gins51 and DP1. The DHH, CID-like, and DHH1 domains of GAN are colored blue, green and orange colors, respectively. The C-terminal domain of Gins51 (Gins51C) and N-terminal domain of DP1 (DP1N) are colored in magenta and cyan, respectively. (**B**) Cartoon representation of the DP1N-Gins51C-GAN ternary complex. The missing part is indicated with a dotted line. (**C**) Surface model of the DP1N–Gins51C–GAN ternary complex.

T. kodakarensis PolD exists as a heterodimer of DP1 and DP2 in solution (44), dimerization of DP2's N-terminal domain has been reported in *P. horikoshii* (59–61). The interaction possibly contributes in adjusting the positions of the two PolD molecules to synthesize both strands simultaneously.

N-terminal domain of DP1 interacts with C-terminal domain of Gins51

To elucidate how the complex's structure enables GINS to bind two PolD molecules without any steric hindrance, we identified the DP1 region that interacts with Gins51. As shown in Figure 3A, DP1 consists of an N-terminal globular domain (1–61) and a catalytic core domain (229–735) that are flexibly tethered by an unstructured (intrinsically disordered) region (62–228) (39,52). On the basis of this structural property, we divided DP1 into five parts and subjected them to Y2H analysis. This experiment clearly showed an interaction between the N-terminal domain of DP1 (1–64, DP1N) and Gins51 (Figure 3B).

We attempted to identify the PolD-interacting region in Gins51 but could not get a clear answer from a further Y2H experiment (Supplementary Figure S3). However, it is noteworthy that Gins51C (131–188) fused with GAL4's activation domain showed a self-activating signal. Nevertheless, cells producing it in combination with DP1N fused with GAL4's DNA-binding domain did not grow, suggesting that the DP1N–Gins51C interaction inhibited cell growth. To test whether DP1N interacts with Gins51C, we performed a pull-down experiment using purified His-tagged Gins51C and *E. coli* extracts containing recombinant DP1N. His-tagged Gins51C pulled down DP1N from the lysate (Figure 3C). Furthermore, gel filtration chromatography showed that Gins51C and DP1N co-eluted earlier than Gins51C alone (Figure 3D). These results show the DP1N–Gins51C complex formation.

Gins51C binds to both DP1N and GAN

The interaction between Gins51C and DP1N contributes to the stable binding of GINS and PolD, as shown in Figure 1C. Gins51C also binds to GAN with the same affinity as the PolD–GINS interaction (22,45). A recent report isolated a recombinant DP1–Gins51–GAN complex from *Thermococcus* sp. 4557 (58). We further tested simultaneous binding of PolD and GAN to GINS, and found that GAN co-eluted with PolD–GINS earlier than GAN alone from a gel filtration (Supplementary Figure S4A). Although GINS can bind two GAN molecules (22,45), a limited resolution of the gel filtration chromatography prevented us from determining stoichiometry of each component.

To investigate the structural basis of the DP1N–Gins51C–GAN interactions, we tried to determine the crystal structure of this ternary complex. We overexpressed the proteins in *E. coli* and successfully purified the DP1N–Gins51C–GAN complex by sequential chromatography, including gel filtration (Supplementary Figure S4B). This result indicated that Gins51C is sufficient for GINS to interact with both PolD and GAN concurrently. We obtained the crystals from this protein sample (Supplementary Figure S4C) and finally solved an atomic structure of the DP1N–Gins51C–GAN complex at 2.45 Å resolution (Figure 4 and Supplementary Table S2). DP1N and GAN bound oppo-

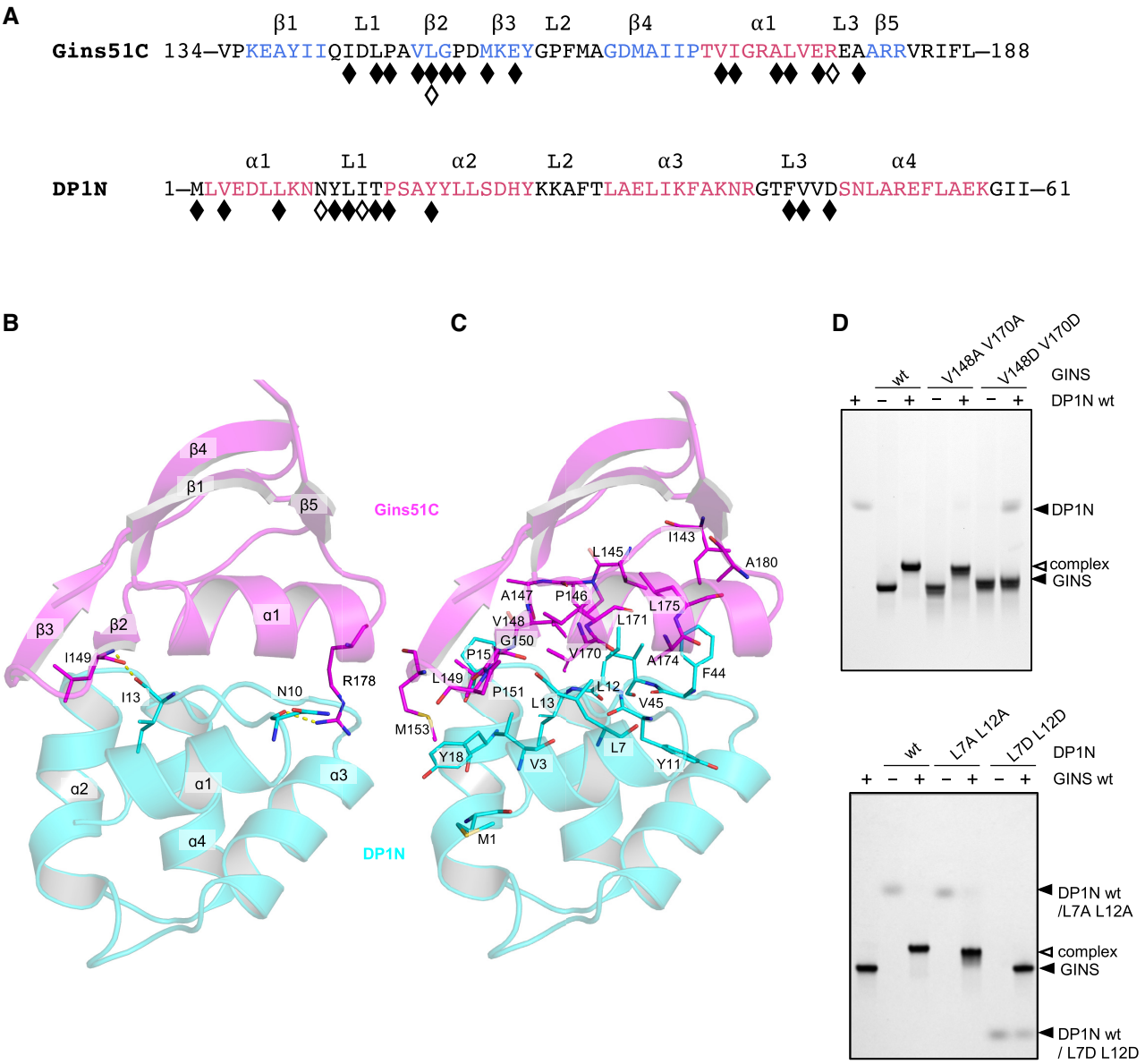

Figure 5. The interface between Gins51C and DP1N. (**A**) Amino acid sequences of Gins51C and DP1N. Amino acids are colored by secondary structure (red, α-helix; blue, β-sheet). The black and white rhombuses indicate the amino acids involved in inter-subunit hydrophobic and hydrophilic interactions, respectively. (B, C) Close-up view of interaction interface between DP1N and Gins51C. The amino acids involved in the hydrogen bonds (**B**) and hydrophobic interactions (**C**) are indicated. (**D**) Complex formation of DP1N-wt with GINS mutants (top) and DP1N mutants with GINS-wt (bottom) tested by native-PAGE. Each protein (100 pmol of DP1N, and 50 pmol of GINS) and mixtures were analyzed.

site sides of the globular Gins51C, and no contacts between DP1N and GAN were observed.

DP1N–Gins51C interface

DP1N is composed of four α-helices (Figure 4), as shown by the NMR-determined N-terminal structure of *P. horikoshii* DP1 (52). The Gins51C–DP1N interaction occurs between a loop (L1), β-sheets (β2 and β3) and an α-helix (α1) of Gins51C, and α1, L1, α2 and L3 of DP1N, with a 646.0 Å^2 interface (Figures 4 and 5). The interaction involves two hydrogen bonds (Figure 5A and B) and many hydrophobic interactions (Figure 5A and C). To verify functional validity of this structure, we analyzed the DP1N–Gins51C interaction by a gel shift assay using native-PAGE after introducing substitution mutations into these proteins. Replacing valine 148/170 of Gins51 or leucine 7/12 of DP1 with alanine had no apparent effect. However, substitution of these amino acids with aspartic acid, the hydrophilic amino acid, completely abolished the complex formation (Figure 5D and Supplementary Figure S5).

GAN–Gins51C interface

GAN bound Gins51C with an 872.4 Å^2 interface, mainly composed of β4 in Gins51C and β2 of DHH domain

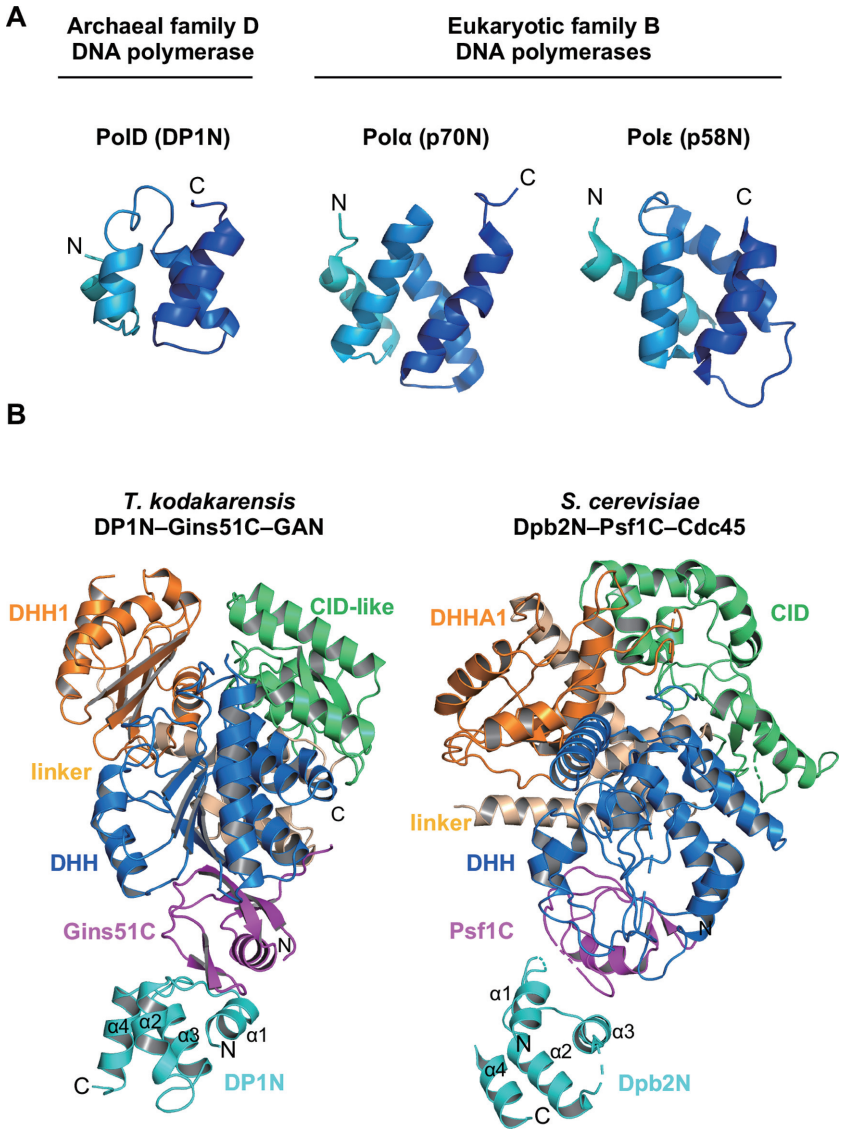

Figure 6. Structural basis for the replisome assembly is conserved between Archaea and Eukarya. (**A**) Structural comparison of DP1N from *T. kodakarensis* with p70N of Polα (PDB: 2KEB) and p58N of Polε (PDB: 2V6Z) from a human. Cartoon representation ranging from cyan (N-terminus) to blue (C-terminus). (**B**) Structural comparison of the DP1N–Gins51C–GAN complex from *T. kodakarensis* and Dpb2N–Psf1C–Cdc45 from *Saccharomyces cerevisiae* (PDB: 6HV9).

in GAN in an antiparallel manner, consistent with our previous report on the TkoGAN–Gins51C dimeric complex (Figure 4B, C, and Supplementary Figure S1) (45). Those two structures were well superimposed, except for the DHH1 domain of GAN, with a root mean square deviation (RMSD) of 0.27 Å over 297 corresponding Cα atoms. The current structure missed density for residues 332–337 in the middle of the linker in GAN, indicating the partial melting of the long α-helix, and the DHH1 domain shifted towards the CID-like domain by approximately 19 Å with a 35° rotation (Supplementary Figure S1). Comparisons of GAN's free form with the Gins51C-bound form (45) or GAN from *T. kodakarensis* and *P. furiosus* (62) revealed similar DHH1 domain movements. Because DP1N binding apparently did not affect either Gins51C or GAN's DHH1, a crystal packing effect may cause this domain movement. The DHH1 domain seems flexible, accompanying a binding or a partial melt in the middle of the linker's α-helix to form a substrate-binding cleft between DHH and DHH1 (62,63).

The structural basis for the replisome assembly is conserved between Archaea and Eukarya

Eukaryotic replicative DNA polymerases (Polα and ε) share the structure of DP1N (Figure 6A) (52). Among them, the N-terminal domain of Dpb2 interacts with the C-terminal domain of Psf1 to link Polε with CMG-helicase (12). Thus, we extracted Dpb2N–Psf1C–Cdc45 from the reported cryo-EM structure of *S. cerevisiae* DNA–CMG–Polε (14) and compared it with the corresponding structure obtained in this study (Figure 6B and Supplementary Figure S6). The overall arrangement was conserved, though DP1N and

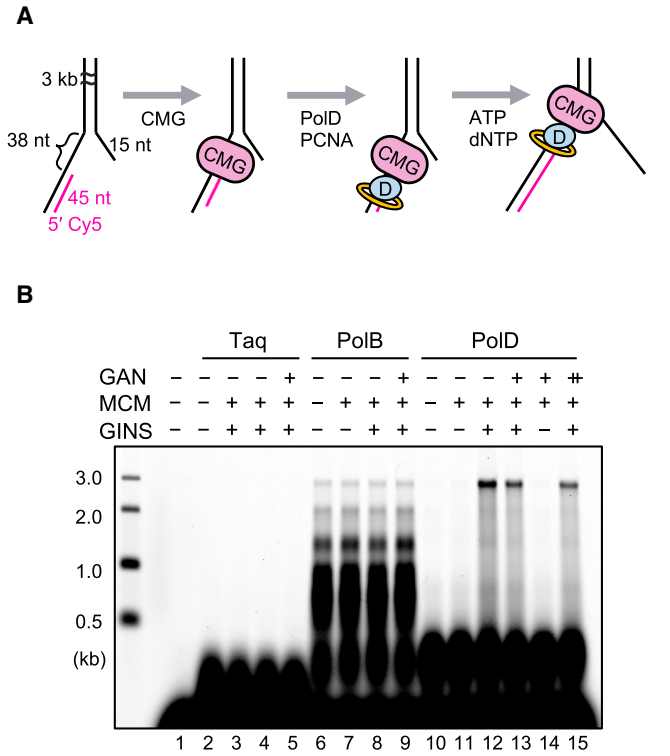

Figure 7. Only PolD binds the CMG helicase complex. (**A**) Flowchart of the leading-strand replication assay using CMG, PolD, and PCNA. The 3 kb of fork-structured dsDNA substrate was constructed as described in Material and Methods. (**B**) Reaction procedure and alkaline agarose gel of reaction products. '+' and '++' in the GAN line stand for equal (200 nM) and excess (800 nM) amounts of GAN (D36A) compared with GINS, respectively.

Dpb2N differ in their directions. Gins51C bound to DP1N's α1–α2 surface, but Psf1C bound to the α1–α3 surface.

The interaction of PolD with CMG-helicase allows efficient and coupled DNA replication

It is hypothesized that PolD follows CMG-helicase in a coordinated manner, like the replisome members for fork progression. To investigate functional relationship, we constructed a leading-strand replication assay using recombinant proteins (MCM, GINS, GAN, PolD and PCNA) and a synthetic DNA substrate mimicking a replication fork, as described in Figure 7A. *T. kodakarensis* PCNA can self-load onto DNA strand, and RFC is not necessarily needed to load PCNA onto DNA strand to stimulate DNA polymerase *in vitro* (43). Despite PCNA assistance, PoID poorly extended the primers and products shorter than 0.5 kb was accumulated in the absence of CMG-helicase (lane 10 in Figure 7B) due to its poor strand displacement ability. We confirmed that PolD could not extend this dsDNA substrate, but it fully extended 3 kb of pBlueScript II SK(+) ssDNA with a single primer (Supplementary Figure S7A). By contrast, full-length products were generated in lanes 12, 13, and 15 of Figure 7B, indicating that PolD completely extended the primer after CMG-helicase unwound 3 kb of the dsDNA region. The nascent strand synthesis seems so processive that we observed no intermediate products. GINS was critical for the reaction (compare lanes 12 and 14), whereas GAN did not promote it by itself (compare lanes 12–15).

Seeking to clarify the importance of the specific connection of PolD with CMG-helicase, we compared the reactions using PolB and *Taq* DNA polymerase (Figure 7B). PolB synthesized DNA strands mainly until ∼1 kb after unwinding dsDNA by its strand displacement activity, and CMG-helicase had no effect (lanes 6–9). On the other hand, *Taq* DNA polymerase did not extend the primer (lanes 2–5). We verified that these two polymerases could also synthesize full-length products on ssDNA (Supplementary Figure S7B). These data reveal that PolD tethering to CMG-helicase via GINS couples unwinding and synthesis and is essential for efficient DNA replication.

PolD stimulates ATPase and helicase activity of MCM–GINS complex

To investigate how PolD binding affects CMG's helicase activity, we directly measured it using a short splayed arm DNA. GINS and GINS–GAN can stimulate the helicase activity of MCM *in vitro* (21), but no effect was noticed in the condition with a lower protein concentration here (Figure 8A, lane 2–4). In this condition, PolD clearly stimulated the helicase activity of the MG and CMG complex (compare lanes 3, 4, 7 and 8). This stimulation was not observed from DP1 alone. We further measured ATPase activity of MCM in the presence and absence of GINS, GAN, and PolD (Figure 8B and Supplementary Figure S8). Consistently with previous reports, MCM showed a DNA-dependent ATPase activity, which can be stimulated by GINS (Figure 8B, lanes 1 and 2) (19,21). It is striking that PolD drastically stimulated the ssDNA-dependent ATPase activity in the presence of GINS by approximately 11-fold relative to just MCM alone (lanes 1 and 4) and 5-fold relative to MCM–GINS (lanes 2 and 4). In the absence of GINS, PolD did not stimulate but rather slightly suppressed the ATPase activity of MCM (lanes 1 and 3) presumably because of competitive DNA binding between MCM and PolD. The ssDNA was preferable over the dsDNA for this stimulation. We previously reported that GAN did not stimulate the activity (21). Consistently, little difference was observed in the presence and absence of GAN in both helicase and ATPase assays (compare lanes 7 and 8 in Figure 8A, and lanes 4 and 5 in Figure 8B), suggesting that GAN does not affect critically on the helicase activity. DP1 alone did not clearly stimulate helicase or ATPase activity in the presence of GINS, although a slight difference was observed between the ATPase reactions with and without DP1 (Figure 8B, lanes 2 and 7). DP2 alone did not promote ATPase activity (Figure 8B, lanes 9–11). These results indicate that GINS-tethered PolD stimulated the ATPase activities of MCM, resulting in the stimulation of the helicase activities of the archaeal CMG complex. DP1 is enough to connect PolD and GINS, but both DP1 and DP2 are required to stimulate the strand unwinding. The effective leading-strand synthesis by PolD in the presence of MG or CMG complexes, shown in Figure 7, seems to reflect the promotion of the helicase activity of CMG cooperated with PolD.

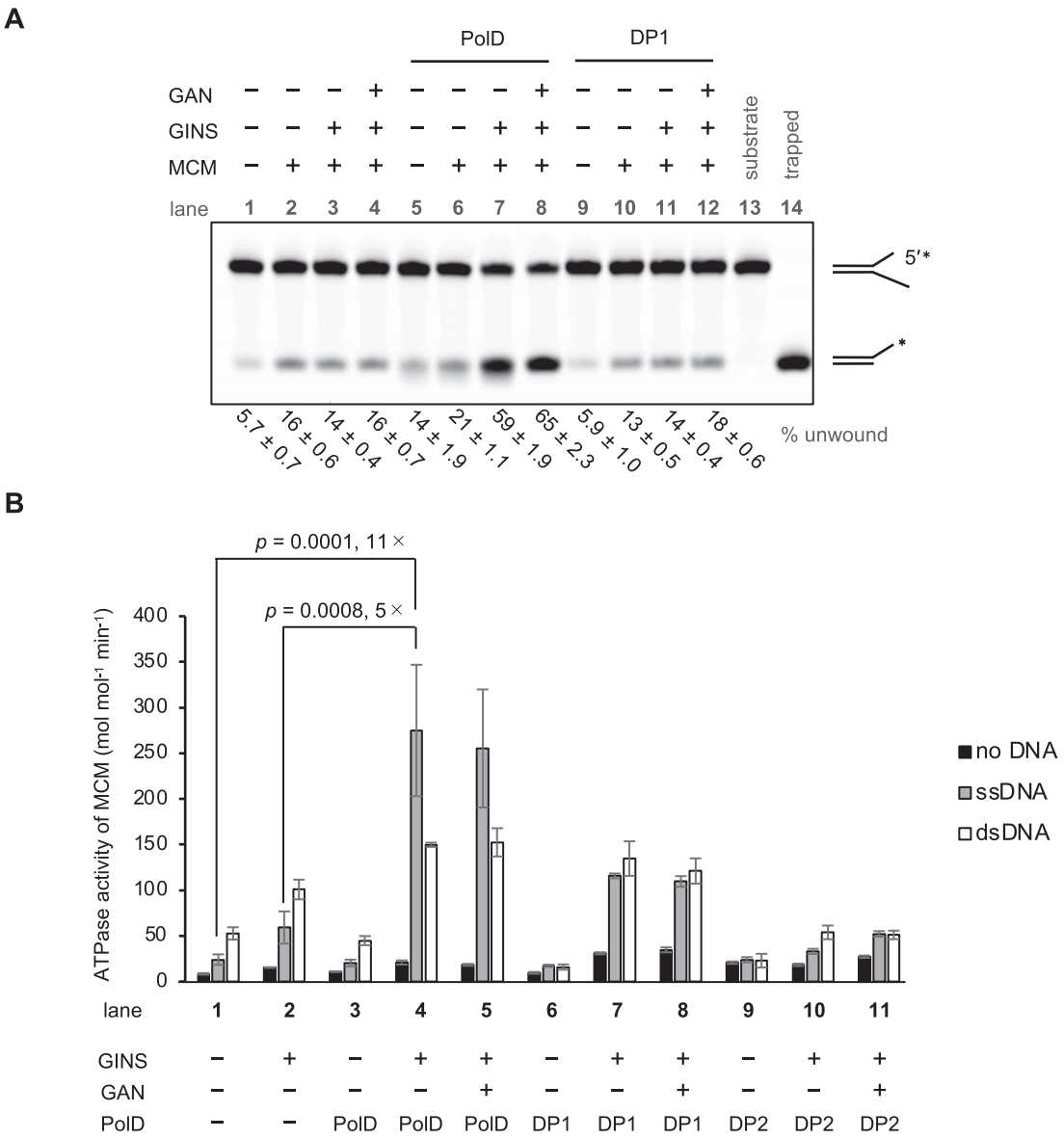

Figure 8. PolD stimulates the strand unwinding activity of MCM–GINS complex. (**A**) Helicase activity of MCM in the presence of GINS, PolD, DP1, and GAN. The 5′-Cy5-labeled splayed-arm DNA substrate was incubated with the indicated proteins. The helicase activity is expressed at the bottom of the panel as the relative amount of unwound DNA (%) with the standard error of the mean from the three independent experiments. (**B**) MCM's ATPase activity in the presence of a 4-fold excess of GINS, PolD and GAN. The ATPase activity is expressed as the amount of Pi released by a constant MCM amount (as the hexamer) in the absence or presence of circular ssDNA or dsDNA. The error bars represent the standard error of the mean. The effects of GINS and PolD (PolD, DP1, or DP2) on the ssDNA-dependent ATPase activity of MCM were assessed by two-way ANOVA Tukey's multiple comparison test. Lanes 5, 8 and 11, which focus on the effect of GAN, were excluded from the statistical analysis. The ATPase activity of MCM in the presence and absence of GINS was clearly stimulated by PolD (compare lane 4 with lanes 1 and 2), and the significant differences ($P < 0.05$) were observed in the effect of PolD. The calculated *P*-values were shown above the lanes.

DISCUSSION

We investigated the functional replisome formation in *T. kodakarensis.* A summary of the interactions revealed in this study is schematically shown in Figure 9. Connecting helicase and polymerases is a smart strategy and is an efficient DNA replication machinery. Here, we reported structural and functional insights into the interaction of the archaeal GINS with PolD to connect replicative helicase and polymerase. We found that GINS connects two PolD molecules via two Gins51 subunits (Figures 1 and 2), suggesting the assembly of a symmetric replisome, in which we presume that each PolD is responsible for either leading or lagging strands. The GINS complex of *T. kodakarensis,* like most crenarchaeal organisms, consists of two Gins51s and two Gins23s. However, in most euryarchaeal organisms like *Thermoplasma acidophilum*, GINS is a homo-tetramer of a single subunit, Gins51 (24,25). A notable difference is that *T. acidophilum*'s GINS requires disordered linkers between its A and B domains of its Gins51 forming a tetramer

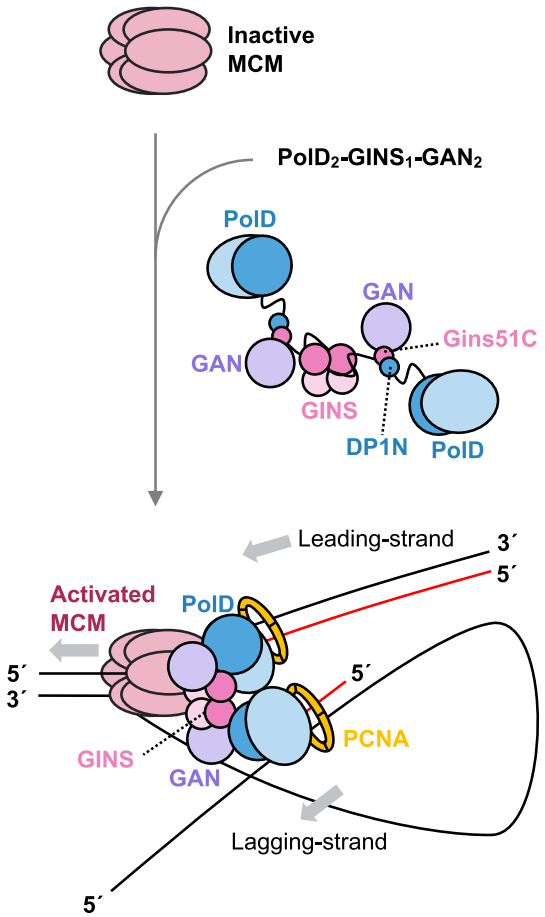

Figure 9. A model for the replisome activation in Archaea. Functional replisome is constructed and schematically drawn based on the interaction analyses in this study. Inactive MCM is loaded onto the DNA first, followed by the recruitment of $PolD_2$–$GINS_1$–GAN_2. The activated MCM shows the ATPase activity, resulting in translocation on leading-strand template DNA in the 3′–5′ direction. Each subunit is shown in a different color. Domain structures and the disordered flexible regions are represented by balls and lines, respectively. Two molecules of PolDs are connected by GINS, and synthesize nascent leading and lagging strands, respectively.

(64). We previously constructed a homology model of *T. acidophilum* GINS based on a negative-stain EM analysis (26). The 3D structure model of the *T. acidophilum* homotetramer GINS, in which the Gins51-type dimer was stacked onto the Gins23-type dimer, was to that of *T. kodakarensis* GINS. In this structure, the C-terminal B-domain of the two subunits corresponding to the Gins23-type subunit participated in the inter-subunit interaction and were unlikely to interact with PolD due to the steric hindrance. Therefore, we assume that the homotetrameric GINS from *T. acidophilum* binds two PolD molecules using two Gins51-type subunits out of four subunits.

We further demonstrated that DP1N and Gins51C (both tethered to the core regions by disordered linkers) mediate the PolD–CMG interaction (Figure 3). Gins51C simultaneously bound to DP1N and GAN, a CMG-helicase component (Figures 3 and 4). As seen in SPR and gel filtration chromatography analyses (Figure 1C and Supplementary Figure S4A), the PolD–GINS–GAN complex formation was highly stable compared to the MCM–GINS interaction ($K_D = 1.9 \times 10^{-7}$) (21). We assume that PolD, GINS, and GAN constitutively form a complex in cells and are jointly loaded onto or unloaded from MCM. This complex formation mode is different from that in eukaryotes, in which GINS and Polε are loaded onto MCM as a fragile pre-LC, following Cdc45.

Structural comparison between *T. kodakarensis* DP1N–Gins51C–GAN and *S. cerevisiae* Dpb2N–Psf1C–Cdc45 revealed a common basis for the replisome formation (Figure 6). In contrast to Psf1, the C-terminal domain of Sld5 (Sld5C) contributes to the eukaryotic GINS complex assembly (28) but not to that of the other replisome components. Previously, Xu and collaborators reported the interaction of Gins51C and Cdc45 from the *Sulfolobus* species and proposed that during the diversification of Psf1 and Sld5, which evolved from the common ancestral Gins51, the Cdc45-interaction site on Sld5C has diverged and lost the ability to interact with Cdc45 (22). Our data strongly support this scenario, and the same can be true for the interaction with DNA polymerases. Another diversified point is that Sld5 acquired an extra N-terminal α-helix containing a Ctf4 interacting peptide (CIP)-motif, which neither Psf1 nor Gins51 possessed, for the interaction with Polα via Ctf4 (17). A multiple sequence alignment of the archaeal Gind51s, eukaryotic Sld5s and Psf1s are exhibited in the Supplementary Figure S9. Interestingly, the N-terminal domain of the second subunit (p70) of human Polα is responsible for the interaction with CMG-helicase (via AND-1) (65), as well as with SV40 helicase T-antigen (66). A recent report further showed that the second subunit of *S. cerevisiae* Polδ also plays a critical role in tethering Polδ to the replisome during multiple cycles of Okazaki fragment synthesis, even though it lacks the corresponding N-terminal domain (67).

We also showed that PolD binding to CMG-helicase is critical for coupling unwinding and synthesis, allowing efficient DNA replication (Figure 7). Consistent with our data, the human Polε, but not Polδ, can follow eukaryotic CMG-helicase. Human Polε reportedly conducted a processive DNA strand synthesis only in the presence of CMG complex *in vitro* (68). Additionally, Polε regulates CMG-helicase, and it is involved in activation of CMG-helicase at the replication origin and lets it stall at the lesion. If *S. cerevisiae* Polε stops synthesizing at roadblocks, CMG stalls unwinding due to polymerase-helicase uncoupling (69,70). Furthermore, a recent single-molecule study revealed that *S. cerevisiae* CMG switches to a diffusive mode on ds-DNA when uncoupled from DNA polymerase. Upon encountering a new replication fork, the CMG can switch back to ssDNA to restart the replication (71). A study using *Xenopus laevis* egg extract showed that Polε and GINS leave the replisome when the replication fork collapses (72). Strikingly, we found that PolD promoted CMG-helicase activity and stimulated its ATPase activity by one order of magnitude in the presence of ssDNA (Figure 8). The functional switch observed in our reconstitution analyses may explain the underlying reason that DNA polymerases regulate CMG-helicase activation at origin or stalled forks.

At present, we found no direct interaction between MCM and PolD. Thus, the structural mechanism of the functional switch for stimulating the helicase activity of CMG using PolD remains unknown. Moreover, 3D structure of PolD is quite different from that of family B DNA polymerases but is rather similar to that of RNA polymerase II. Further structural and biochemical studies will unveil the molecular mechanism of the functional regulation in the archaeal replisome.

DATA AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structure have been deposited into the Protein Data Bank under accession number 7E15.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Yalu Tang for technical assistance. The authors would like to thank Enago (https://www.enago.jp/) for the English language review.

FUNDING

Grant-in-Aid for JSPS Fellows [JP20J12260 to K.O.]; JSPS KAKENHI [JP19K22289 to Y.I., JP21K05394 to S.I., JP18K06081 to T.O.]; JST Adaptable and Seamless Technology Transfer Program through Target-driven R&D (A-STEP) [JPMJTM19AT to T.O.]; T.O. was also supported by Platform Project for Supporting Drug Discovery and Life Science Research (Basis of Supporting Innovative Drug Discovery and Life Science Research) from AMED [support number 0670]. Funding for open access charge: Grant-in-Aid for JSPS Fellows.
Conflict of interest statement. None declared.

REFERENCES

1. 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.
2. 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.
3. Mott,M.L. and Berger,J.M. (2007) DNA replication initiation: mechanisms and regulation in bacteria. *Nat. Rev. Microbiol.*, **5**, 343–354.
4. O'Donnell,M., Langston,L. and Stillman,B. (2013) Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.*, **5**, a010108.
5. 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.
6. 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.
7. 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.
8. 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.
9. Evrin,C., Clarke,P., Zech,J., Lurz,R., Sun,J., Uhle,S., Li,H., Stillman,B. and Speck,C. (2009) A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 20240–20245.
10. Remus,D., Beuron,F., Tolun,G., Griffith,J.D., Morris,E.P. and Diffley,J.F. (2009) Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*, **139**, 719–730.
11. Muramatsu,S., Hirai,K., Tak,Y.S., Kamimura,Y. and Araki,H. (2010) CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol (epsilon}, and GINS in budding yeast. *Genes Dev.*, **24**, 602–612.
12. 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.
13. Handa,T., Kanke,M., Takahashi,T.S., Nakagawa,T. and Masukata,H. (2012) DNA polymerization-independent functions of DNA polymerase epsilon in assembly and progression of the replisome in fission yeast. *Mol. Biol. Cell*, **23**, 3240–3253.
14. Goswami,P., Abid Ali,F., Douglas,M.E., Locke,J., Purkiss,A., Janska,A., Eickhoff,P., Early,A., Nans,A., Cheung,A.M.C. *et al.* (2018) Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nat. Commun.*, **9**, 5061.
15. Sun,J., Shi,Y., Georgescu,R.E., Yuan,Z., Chait,B.T., Li,H. and O'Donnell,M.E. (2015) The architecture of a eukaryotic replisome. *Nat. Struct. Mol. Biol.*, **22**, 976–982.
16. Ticau,S., Friedman,L.J., Ivica,N.A., Gelles,J. and Bell,S.P. (2015) Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell*, **161**, 513–525.
17. Simon,A.C., Zhou,J.C., Perera,R.L., van Deursen,F., Evrin,C., Ivanova,M.E., Kilkenny,M.L., Renault,L., Kjaer,S., Matak-Vinkovic,D. *et al.* (2014) A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. *Nature*, **510**, 293–297.
18. Sakakibara,N., Kelman,L.M. and Kelman,Z. (2009) Unwinding the structure and function of the archaeal MCM helicase. *Mol. Microbiol.*, **72**, 286–296.
19. 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.
20. Pan,M., Santangelo,T.J., Li,Z., Reeve,J.N. and Kelman,Z. (2011) Thermococcus kodakarensis encodes three MCM homologs but only one is essential. *Nucleic Acids Res.*, **39**, 9671–9680.
21. Nagata,M., Ishino,S., Yamagami,T., Ogino,H., Simons,J.R., Kanai,T., Atomi,H. and Ishino,Y. (2017) The Cdc45/RecJ-like protein forms a complex with GINS and MCM, and is important for DNA replication in Thermococcus kodakarensis. *Nucleic Acids Res.*, **45**, 10693–10705.
22. Xu,Y., Gristwood,T., Hodgson,B., Trinidad,J.C., Albers,S.V. and Bell,S.D. (2016) Archaeal orthologs of Cdc45 and GINS form a stable complex that stimulates the helicase activity of MCM. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 13390–13395.
23. 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.
24. 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, Thermoplasma acidophilum may function as a homotetramer in DNA replication. *Extremophiles*, **15**, 529–539.
25. 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.
26. 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.
27. 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.
28. 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.
29. 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.
30. 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.
31. Uemori,T., Ishino,Y., Toh,H., Asada,K. and Kato,I. (1993) Organization and nucleotide sequence of the DNA polymerase gene from the archaeon Pyrococcus furiosus. *Nucleic. Acids. Res.*, **21**, 259–265.
32. Uemori,T., Sato,Y., Kato,I., Doi,H. and Ishino,Y. (1997) A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization. *Genes Cells*, **2**, 499–512.
33. Cann,I.K., Komori,K., Toh,H., Kanai,S. and Ishino,Y. (1998) A heterodimeric DNA polymerase: evidence that members of Euryarchaeota possess a distinct DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 14250–14255.
34. Ishino,Y., Komori,K., Cann,I.K. and Koga,Y. (1998) A novel DNA polymerase family found in Archaea. *J. Bacteriol.*, **180**, 2232–2236.
35. Aravind,L. and Koonin,E.V. (1998) Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res.*, **26**, 3746–3752.
36. Makiniemi,M., Pospiech,H., Kilpelainen,S., Jokela,M., Vihinen,M. and Syvaoja,J.E. (1999) A novel family of DNA-polymerase-associated B subunits. *Trends Biochem. Sci.*, **24**, 14–16.
37. Raia,P., Carroni,M., Henry,E., Pehau-Arnaudet,G., Brule,S., Beguin,P., Henneke,G., Lindahl,E., Delarue,M. and Sauguet,L. (2019) Structure of the DP1-DP2 PolD complex bound with DNA and its implications for the evolutionary history of DNA and RNA polymerases. *PLoS Biol.*, **17**, e3000122.
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. Sauguet,L., Raia,P., Henneke,G. and Delarue,M. (2016) Shared active site architecture between archaeal PolD and multi-subunit RNA polymerases revealed by X-ray crystallography. *Nat. Commun.*, **7**, 12227.
40. Madru,C., Henneke,G., Raia,P., Hugonneau-Beaufet,I., Pehau-Arnaudet,G., England,P., Lindahl,E., Delarue,M., Carroni,M. and Sauguet,L. (2020) Structural basis for the increased processivity of D-family DNA polymerases in complex with PCNA. *Nat. Commun.*, **11**, 1591.
41. Mayanagi,K., Oki,K., Miyazaki,N., Ishino,S., Yamagami,T., Morikawa,K., Iwasaki,K., Kohda,D., Shirai,T. and Ishino,Y. (2020) Two conformations of DNA polymerase D-PCNA-DNA, an archaeal replisome complex, revealed by cryo-electron microscopy. *BMC Biol.*, **18**, 152.
42. Oki,K., Yamagami,T., Nagata,M., Mayanagi,K., Shirai,T., Adachi,N., Numata,T., Ishino,S. and Ishino,Y. (2021) DNA polymerase D temporarily connects primase to the CMG-like helicase before interacting with proliferating cell nuclear antigen. *Nucleic Acids Res.*, **49**, 4599–4612.
43. Kuba,Y., Ishino,S., Yamagami,T., Tokuhara,M., Kanai,T., Fujikane,R., Daiyasu,H., Atomi,H. and Ishino,Y. (2012) Comparative analyses of the two proliferating cell nuclear antigens from the hyperthermophilic archaeon, *Thermococcus kodakarensis*. *Genes Cells*, **17**, 923–937.
44. Takashima,N., Ishino,S., Oki,K., Takafuji,M., Yamagami,T., Matsuo,R., Mayanagi,K. and Ishino,Y. (2019) Elucidating functions of DP1 and DP2 subunits from the Thermococcus kodakarensis family D DNA polymerase. *Extremophiles*, **23**, 161–172.
45. Oyama,T., Ishino,S., Shirai,T., Yamagami,T., Nagata,M., Ogino,H., Kusunoki,M. and Ishino,Y. (2016) Atomic structure of an archaeal GAN suggests its dual roles as an exonuclease in DNA repair and a CMG component in DNA replication. *Nucleic Acids Res.*, **44**, 9505–9517.
46. Nagata,M., Ishino,S., Yamagami,T. and Ishino,Y. (2019) Replication protein A complex in Thermococcus kodakarensis interacts with DNA polymerases and helps their effective strand synthesis. *Biosci. Biotechnol. Biochem.*, **83**, 695–704.
47. Ishino,Y., Ueno,T., Miyagi,M., Uemori,T., Imamura,M., Tsunasawa,S. and Kato,I. (1994) Overproduction of Thermus aquaticus DNA polymerase and its structural analysis by ion-spray mass spectrometry. *J. Biochem.*, **116**, 1019–1024.
48. Kabsch,W. (2010) XDS. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 125–132.
49. Evans,P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D. Biol. Crystallogr.*, **62**, 72–82.
50. Evans,P.R. and Murshudov,G.N. (2013) How good are my data and what is the resolution?*Acta Crystallogr. D. Biol. Crystallogr.*, **69**, 1204–1214.
51. Adams,P.D., Afonine,P.V., Bunkóczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 213–221.
52. Yamasaki,K., Urushibata,Y., Yamasaki,T., Arisaka,F. and Matsui,I. (2010) Solution structure of the N-terminal domain of the archaeal D-family DNA polymerase small subunit reveals evolutionary relationship to eukaryotic B-family polymerases. *FEBS Lett.*, **584**, 3370–3375.
53. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.*, **66**, 486–501.
54. Cubonova,L., Richardson,T., Burkhart,B.W., Kelman,Z., Connolly,B.A., Reeve,J.N. and Santangelo,T.J. (2013) Archaeal DNA polymerase D but not DNA polymerase B is required for genome replication in *Thermococcus kodakarensis*. *J. Bacteriol.*, **195**, 2322–2328.
55. Kushida,T., Narumi,I., Ishino,S., Ishino,Y., Fujiwara,S., Imanaka,T. and Higashibata,H. (2019) Pol B, a family B DNA polymerase, in *Thermococcus kodakarensis* is important for DNA repair, but not DNA replication. *Microbes Environ.*, **34**, 316–326.
56. Pluchon,P.F., Fouqueau,T., Creze,C., Laurent,S., Briffotaux,J., Hogrel,G., Palud,A., Henneke,G., Godfroy,A., Hausner,W. *et al.* (2013) An extended network of genomic maintenance in the archaeon Pyrococcus abyssi highlights unexpected associations between eucaryotic homologs. *PLoS One*, **8**, e79707.
57. Li,Z., Santangelo,T.J., Cubonova,L., Reeve,J.N. and Kelman,Z. (2010) Affinity purification of an archaeal DNA replication protein network. *mBio*, **1**, e00221.
58. Lu,S., Zhang,X., Chen,K., Chen,Z., Li,Y., Qi,Z., Shen,Y. and Li,Z. (2019) The small subunit of DNA polymerase D (DP1) associates with GINS-GAN complex of the thermophilic archaea in *Thermococcus* sp. 4557. *Microbiologyopen*, **8**, e00848.
59. Matsui,I., Urushibata,Y., Shen,Y., Matsui,E. and Yokoyama,H. (2011) Novel structure of an N-terminal domain that is crucial for the dimeric assembly and DNA-binding of an archaeal DNA polymerase D large subunit from Pyrococcus horikoshii. *FEBS Lett.*, **585**, 452–458.
60. Matsui,I., Matsui,E., Yamasaki,K. and Yokoyama,H. (2013) Domain structures and inter-domain interactions defining the holoenzyme architecture of archaeal d-family DNA polymerase. *Life (Basel)*, **3**, 375–385.
61. Shen,Y., Musti,K., Hiramoto,M., Kikuchi,H., Kawarabayashi,Y. and Matsui,I. (2001) Invariant Asp-1122 and Asp-1124 are essential residues for polymerization catalysis of family D DNA polymerase from Pyrococcus horikoshii. *J. Biol. Chem.*, **276**, 27376–27383.
62. Li,M.J., Yi,G.S., Yu,F., Zhou,H., Chen,J.N., Xu,C.Y., Wang,F.P., Xiao,X., He,J.H. and Liu,X.P. (2017) The crystal structure of *Pyrococcus furiosus* RecJ implicates it as an ancestor of eukaryotic Cdc45. *Nucleic Acids Res.*, **45**, 12551–12564.
63. 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.
64. Ogino,H., Ishino,S., Oyama,T., Kohda,D. and Ishino,Y. (2015) Disordered interdomain region of Gins is important for functional

tetramer formation to stimulate MCM helicase in Thermoplasma acidophilum. *Biosci. Biotechnol. Biochem.*, **79**, 432–438.
65. Kilkenny,M.L., Simon,A.C., Mainwaring,J., Wirthensohn,D., Holzer,S. and Pellegrini,L. (2017) The human CTF4-orthologue AND-1 interacts with DNA polymerase alpha/primase via its unique C-terminal HMG box. *Open Biol*, **7**, 170217.
66. Zhou,B., Arnett,D.R., Yu,X., Brewster,A., Sowd,G.A., Xie,C.L., Vila,S., Gai,D., Fanning,E. and Chen,X.S. (2012) Structural basis for the interaction of a hexameric replicative helicase with the regulatory subunit of human DNA polymerase alpha-primase. *J. Biol. Chem.*, **287**, 26854–26866.
67. Lewis,J.S., Spenkelink,L.M., Schauer,G.D., Yurieva,O., Mueller,S.H., Natarajan,V., Kaur,G., Maher,C., Kay,C., O'Donnell,M.E. *et al.* (2020) Tunability of DNA polymerase stability during Eukaryotic DNA replication. *Mol. Cell*, **77**, 17–25.
68. Kang,Y.-H., Galal,W.C., Farina,A., Tappin,I. and Hurwitz,J. (2012) Properties of the human Cdc45/Mcm2-7/GINS helicase complex and its action with DNA polymerase ε in rolling circle DNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 6042–6047.
69. Taylor,M.R.G. and Yeeles,J.T.P. (2019) Dynamics of replication fork progression following helicase-polymerase uncoupling in Eukaryotes. *J. Mol. Biol.*, **431**, 2040–2049.
70. Hizume,K., Endo,S., Muramatsu,S., Kobayashi,T. and Araki,H. (2018) DNA polymerase epsilon-dependent modulation of the pausing property of the CMG helicase at the barrier. *Genes Dev.*, **32**, 1315–1320.
71. Wasserman,M.R., Schauer,G.D., O'Donnell,M.E. and Liu,S. (2019) Replication fork activation is enabled by a single-stranded DNA gate in CMG helicase. *Cell*, **178**, 600–611.
72. Hashimoto,Y., Puddu,F. and Costanzo,V. (2011) RAD51- and MRE11-dependent reassembly of uncoupled CMG helicase complex at collapsed replication forks. *Nat. Struct. Mol. Biol.*, **19**, 17–24.