6114–6123 *Nucleic Acids Research, 2011, Vol. 39, No. 14* *Published online 31 March 2011*
doi:10.1093/nar/gkr181

A novel DNA nuclease is stimulated by association with the GINS complex

Zhuo Li[1,2], Miao Pan[1,2], Thomas J. Santangelo[3], Wiebke Chemnitz[4], Wei Yuan[1,5], James L. Edwards[1,5], Jerard Hurwitz[4], John N. Reeve[3] and Zvi Kelman[1,2,*]

[1]Institute for Bioscience and Biotechnology Research, 9600 Gudelsky Drive, Rockville, MD 20850, [2]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, [3]Department of Microbiology, Ohio State University, Columbus, OH 43210, [4]Program in Molecular Biology, Memorial Sloan Kettering Cancer Center, New York, NY 10065 and [5]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742, USA

Received February 16, 2011; Revised March 11, 2011; Accepted March 14, 2011

ABSTRACT

Chromosomal DNA replication requires the spatial and temporal coordination of the activities of several complexes that constitute the replisome. A previously uncharacterized protein, encoded by *TK1252* in the archaeon *Thermococcus kodakaraensis*, was shown to stably interact with the archaeal GINS complex *in vivo*, a central component of the archaeal replisome. Here, we document that this protein (TK1252p) is a processive, single-strand DNA-specific exonuclease that degrades DNA in the 5′ → 3′ direction. TK1252p binds specifically to the GINS15 subunit of *T. kodakaraensis* GINS complex and this interaction stimulates the exonuclease activity *in vitro*. This novel archaeal nuclease, designated GINS-associated nuclease (GAN), also forms a complex *in vivo* with the euryarchaeal-specific DNA polymerase D. Roles for GAN in replisome assembly and DNA replication are discussed.

INTRODUCTION

Chromosomal DNA replication has many universally conserved features, but there are differences in the proteins and complexes that initiate and maintain DNA replication forks in *Bacteria*, *Archaea* and *Eukarya*. Many of the proteins required for bacterial and eukaryal replication have been isolated and characterized extensively, while most of the components of the archaeal replication machinery have only been putatively identified by bioinformatics. When archaeal proteins with sequences in common with bacterial and/or eukaryal replisome proteins have been investigated, the results obtained have generally confirmed their predicted replication functions (1–3). This *in silico* approach, however, does not readily identify archaeal-specific replisome proteins. To address this limitation, ~20 *Thermococcus kodakaraensis* strains were constructed, each of which synthesized an established archaeal replication protein with an amino- or carboxy-terminal hexahistidine extension (His_6-tag). These proteins were purified directly by nickel-affinity from *T. kodakaraensis* cell lysates, and all proteins that were consistently co-isolated with each His_6-tagged replication protein were identified (4). Here, we report the characterization of a novel archaeal nuclease, encoded by *TK1252*, that was present in the complexes isolated by nickel-binding of His_6-tagged subunits of the *T. kodakaraensis* GINS complex and the archaeal-specific DNA polymerase D (Pol D).

In *Eukarya*, the heterotetramer GINS complex (containing Sld5, Psf1, Psf2 and Psf3) associates with the mini-chromosome maintenance (MCM) proteins, Mcm2-7 and with Cdc45 to form the Cdc45, Mcm2-7, GINS (CMG) complex. This complex has a 3′ → 5′ DNA helicase activity and is thought to function as the replicative helicase (5–7). The GINS complex is required to establish and maintain replication forks (8–10) and also interacts with the Polα-primase complex that synthesizes primers on the lagging strand (7,11). With these features, the eukaryal GINS complex appears to be the functional homologue of the τ subunit (DnaX) of the *Escherichia coli* replisome that binds to the bacterial replicative DNA polymerase (Pol III), DNA helicase (DnaB) and primase (DnaG) at the replication fork and coordinates leading and lagging strand syntheses (12). Sequence homologies predict that many *Archaea*, including *T. kodakaraensis*, have a GINS complex assembled from two molecules each of GINS15 (TK0536p) and GINS23 (TK1619p), proteins most closely related to the eukaryal Psf1 and

*To whom correspondence should be addressed. Tel: +1 240 314 6294; Fax: +1 240 314 6255; Email: zkelman@umd.edu

Sld5, and Psf2 and Psf3 proteins, respectively (11,13). Consistent with GINS being an archaeal replisome component, investigations of [$GINS15_2$–$GINS23_2$] complexes from several *Archaea* have documented interactions with the archaeal primase, MCM, Pol D and PCNA (4,14–16).

In the *T. kodakaraensis* genome annotation, the protein encoded by *TK1252* is predicted to be a single-strand specific nuclease (17). The results reported here confirm that this protein does associate with the GINS complex, specifically with the GINS15 component, and demonstrate that it is a single-strand (ss) DNA-specific $5' \rightarrow 3'$ exonuclease. The exonuclease activity of this protein, designated GINS-associated nuclease (GAN), is stimulated by its interaction with GINS15. Possible roles for the GAN–GINS association during archaeal DNA replication are discussed.

MATERIALS AND METHODS

Nuclease substrates

[γ-^{32}P]ATP was purchased from Perkin Elmer. Unlabeled, Cy3- and Cy5-labeled deoxy- and ribo-oligonucleotides, with the sequences listed in Supplementary Table S1, were obtained from the NIST/UMD nucleic acids synthesis facility. Double-strand (ds) DNA substrates were generated by annealing complementary oligonucleotides followed by PAGE purification, as previously described (18). To obtain linear and circular 200-mer substrates, 1.5 nmol of the 100-mer oligonucleotides A and B (Supplementary Table S1) were phosphorylated by incubation with 40 U of T4 polynucleotide kinase for 1 h at 37°C. The phosphorylation reaction mixture for oligonucleotide B also contained 71 pmol of [γ-^{32}P]ATP. To construct the linear substrate, 0.5 nmol of phosphorylated oligonucleotides A and B plus 2.5 nmol of the bridge oligonucleotide AB (Supplementary Table S1) were mixed in 20 mM HEPES (pH 7.5), 150 mM NaCl, heated to 100°C and the mixture was then allowed to cool slowly to 22°C. This procedure was also used to generate the circular substrate, except that the reaction mixture also contained 2.5 nmol of the bridge oligonucleotide BA (Supplementary Table S1). The reaction mixtures were placed at 16°C, 8000 U of T4 DNA ligase were added and incubation continued for 14 h. The reaction products were separated by electrophoresis at 15 W for 75 min through 10% (w/v) polyacrylamide–8 M urea gels run in TBE. The regions of the gel containing the desired 200-mer linear and circular ssDNAs were excised and the DNAs eluted from the gel into 0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, ethanol precipitated and dissolved in 30 µl of TE (pH 8). The resulting solution was passed through a S300 mini-column filter (GE Healthcare).

Plasmids construction

For protein expression in *E. coli*, the genes encoding GAN (*TK1252*), GINS15 (*TK0536*) and GINS23 (*TK1619*) were PCR-amplified from *T. kodakaraensis* genomic DNA using primers (listed in Supplementary Table S1) that added an in-frame His_6-encoding sequence to the 3′-terminus of the amplified gene. The amplified DNAs were ligated with pET15b (*TK1252*) or pET21a (*TK0536* and *TK1619*) linearized by digestion with the restriction enzyme listed in Supplementary Table S1. A plasmid that directed the synthesis of GAN (D34A) was generated by site-specific mutagenesis from the plasmid that expressed *TK1252* by using a QuikChange mutagenesis kit (Stratagene) using oligonucleotides with the sequences listed in Supplementary Table S1.

To construct a *T. kodakaraensis* strain that synthesized GAN-His_6 *in vivo*, *TK1252* and DNA from immediately upstream and downstream of *TK1252* were separately amplified from *T. kodakaraensis* genomic DNA. An overlapping PCR was used to add an His_6-encoding sequence in-frame to the 5′-terminus of *TK1252* (19). The three amplified DNAs were cloned into pUMT2 (20) adjacent to *trpE* (*TK0254*) to generate plasmid pZLE034 (Supplementary Figure S1). In pZLE034, *TK1252-His$_6$* is positioned between genomic sequences that are homologous to the DNA immediately upstream and downstream of *TK1252* in the *T. kodakaraensis* KW128 genome. An aliquot of pZLE034 DNA was used to transform *T. kodakaraensis* KW128 (Δ*pyrF*; Δ*trpE::pyrF*) as previously described (20,22) and transformants were selected by growth on plates lacking tryptophan. The desired replacement of TK1252 with the GAN-His6 encoding gene was confirmed in a representative transformant, designated *T. kodakaraensis* 34-5, by diagnostic PCR and sequencing (4).

Recombinant protein purification

The plasmids encoding GAN, GAN (D34A), GINS15 or GINS23 were transformed into *E. coli* BL21 (DE3)-CodonPlus-RIL (Stratagene). Isopropyl-β-D-thiogalactopyranoside induction, expression at 16°C for 16 h and purification of the recombinant N-terminal His_6-tagged GAN and GAN (D34A), and C-terminal His_6-tagged GINS15 and GINS23 from *E. coli* cell lysates by Ni^{2+}-affinity chromatography were carried out as previously described (21). Aliquots of the purified proteins were stored at −80°C.

Size exclusion chromatography

Aliquots of each experimental protein (100 µg) or protein mixture and Gel Filtration standards (Bio-Rad) were dissolved in 200 µl of 25 mM Tris–HCl (pH 7.5), 100 mM NaCl, 10% (v/v) glycerol and loaded onto a Superdex-200 column (HR10/30; GE Healthcare) pre-equilibrated in the same buffer. Fractions (250 µl) were collected from the column at a flow rate of 0.5 ml/min. The proteins present in aliquots (80 µl) of each fraction were separated by electrophoresis through a 12% (w/v) polyacrylamide–SDS gel and stained with Coomassie brilliant blue (R250).

Nuclease assays

Unless otherwise noted in the figure legends, the nuclease assay reaction mixtures (20 µl) containing the DNA substrate, BSA (125 µg/ml), 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$ and GAN, were incubated at 70°C for 20 min.

Nuclease digestion was stopped by adding 20 µl of 95% formamide, 0.1× TBE, 10 mM EDTA and incubation at 100°C for 2 min. The digestion products were visualized and quantified by phosphorimaging after electrophoretic separation through 20% (w/v) polyacrylamide–8 M urea gels run in TBE for 1.25 h at 15 W. For native gels nuclease reactions were stopped by adding 5 µl of 50% glycerol, 20 mM EDTA. The digestion products were visualized and quantified by phosphorimaging after electrophoretic separation through 20% (w/v) polyacrylamide gels run in TBE for 2 h at 300 V.

Liquid chromatography-mass spectrometry

Aliquots (10 µM) of the DNA templates, 5′-AAAAAAGG and 5′-GGAAAAAA, were incubated in reaction mixtures (50 µl), with or without 20 pmol GAN, for 1 h at 70°C in a buffer containing 5 mM ammonium formate (pH 6.5), 2 mM $MnCl_2$. The products were subjected to liquid chromatography (LC)/mass spectrometry (MS) analyses using the negative ion mode with a Finnigan LTQ ion trap mass spectrometer (San Jose, CA, USA) equipped with nanospray ionization (NSI) interface coupled to an Agilent 1200 HPLC system (Palo Alto, CA, USA). The flow from the Aligent pump was split from 0.85 ml to 25 nl/min using a 75 µm internal diameter (ID) silica capillary as the flow splitter. Separations were performed using 50 µm ID silica capillary columns (Polymicro Technology, Phoenix, AZ, USA) with in-house made frit packed with 15 cm of 3 µm Atlantis T3 C18 aqueous reversed phase particles (Waters, Milford, MA, USA). The mobile phase A was 5 mM ammonium formate (pH 6.0) in water, and the mobile phase B was 5 mM ammonium formate in methanol. Analytes were eluted over 30 min using a 0–95% linear gradient of solvent B. The heated capillary was at 200°C. Fragmentation was activated by collision-induced dissociation of 35%. Selective reaction monitoring was implemented with the following transitions: dAMP: 330.1 to 195.1 m/z; dGMP: 346.1 to 195.1 m/z; dAMP-dAMP: 643.1 to 330.1 m/z; and dGMP-dGMP: 675.1 to 346.1 m/z. The instrument control, data acquisition, and data analysis were performed by Xcalibur software (Thermo Electron Corporation, version 2.0.7 SP1).

Isolation and identification of His6-tagged GAN and associated proteins from T. kodakaraensis

Thermococcus kodakaraensis 34-5 cultures (5 l) were grown to late exponential phase (OD_{600} of ~0.8) at 80°C in MA-YT medium supplemented with 5 g sodium pyruvate/l in a BioFlow 415 fermentor (New Brunswick Scientific). The cells were harvested by centrifugation, resuspended in 30 ml of buffer A [25 mM Tris–HCl (pH 8), 500 mM NaCl, 10 mM imidazole and 10% glycerol] and lysed by sonication. After centrifugation, the resulting clarified lysate was loaded onto a 1 ml HiTrap chelating column (GE Healthcare) pre-equilibrated with $NiSO_4$. The column was washed with buffer A and proteins were eluted using a linear imidazole gradient from buffer A to 67% buffer B [25 mM Tris–HCl (pH 8), 100 mM NaCl, 150mM imidazole and 10% glycerol]. Fractions that contained the GAN protein were pooled and dialyzed against buffer C [25mM Tris–HCl (pH 8), 500 mM NaCl, 0.5 mM EDTA, 2 mM DTT]. Aliquots (30 µg) of the protein were precipitated by adding trichloroacetic acid (TCA; 15% final concentration). The TCA-precipitated proteins were identified by multidimensional protein identification technology at the Ohio State University mass spectrometry facility (http://www.ccic.ohio-state.edu/MS/proteomics.htm) using the MASCOT search engine. The protein isolation and mass spectrometry analyses were also repeated twice using lysates from two independent cultures of *T. kodakaraensis* KW128. These control experiments identified the *T. kodakaraensis* proteins that bound and eluted from a Ni^{2+}-charged matrix in the absence of a His_6-tagged protein. All proteins co-isolated with His_6-GAN by binding to Ni^{2+}-matrix from *T. kodakaraensis* 34-5 cell lysates that had high MASCOT scores, and that were not present in the control samples, are listed in Table 1.

RESULTS

Purified GAN and GINS15 form complexes in solution

GAN (TK1252p) was co-isolated with His_6-GINS15 (TK0536p) from *T. kodakaraensis* cell lysates by His_6-GINS15 binding to a Ni^{2+}-charged matrix followed by imidazole elution consistent with GAN forming a stable complex with GINS15 *in vivo* (4). To determine if these proteins also interacted *in vitro,* recombinant GAN and GINS15 were mixed and the products examined by size exclusion chromatography. In the absence of GINS15, the GAN (52.9 kDa) elution profile was consistent with the presence of monomers and $(GAN)_2$ dimers (Figure 1A; elution peaks in fractions 53 and 59). GINS15 (21.5 kDa) alone eluted almost exclusively at a position consistent with a $(GINS15)_2$ dimer (Figure 1B, elution peak in fraction 56), as reported previously for GINS15 from *Sulfolobus solfataricus* (14). When incubated together, GAN and GINS15 interacted to form several complexes that eluted in fractions consistent with the formation of complexes larger than $(GAN)_2$ and $(GINS15)_2$ dimers (Figure 1E, elution peaks in fractions 38 and 47). Incubation of GAN with GINS23 (19.2 kDa) did not result in the formation of larger complexes (Figure 1F). GAN bound the $[GINS15_2–GINS23_2]$ complex (Figure 1G), suggesting that the GAN–GINS15 interactions did not disrupt the GINS complex (14,16).

GAN is a ssDNA nuclease

Based on limited sequence similarities between GAN and *E. coli* RecJ (Supplementary Figure S2), GAN was predicted to be a ssDNA nuclease (17). As shown in Figure 2A, this was confirmed. A Cy3-labeled single stranded deoxy-oligonucleotide (30-mer) was fully digested by GAN in the presence of Mn^{2+} (Figure 2A; lane 2) and limited activity was also observed with Mg^{2+} (Figure 2A; lane 8). There was no activity in the absence of metal ions, or with Zn^{2+}, Li^{+} or Ca^{2+} present. An aspartate, known to be important for RecJ activity (23), is

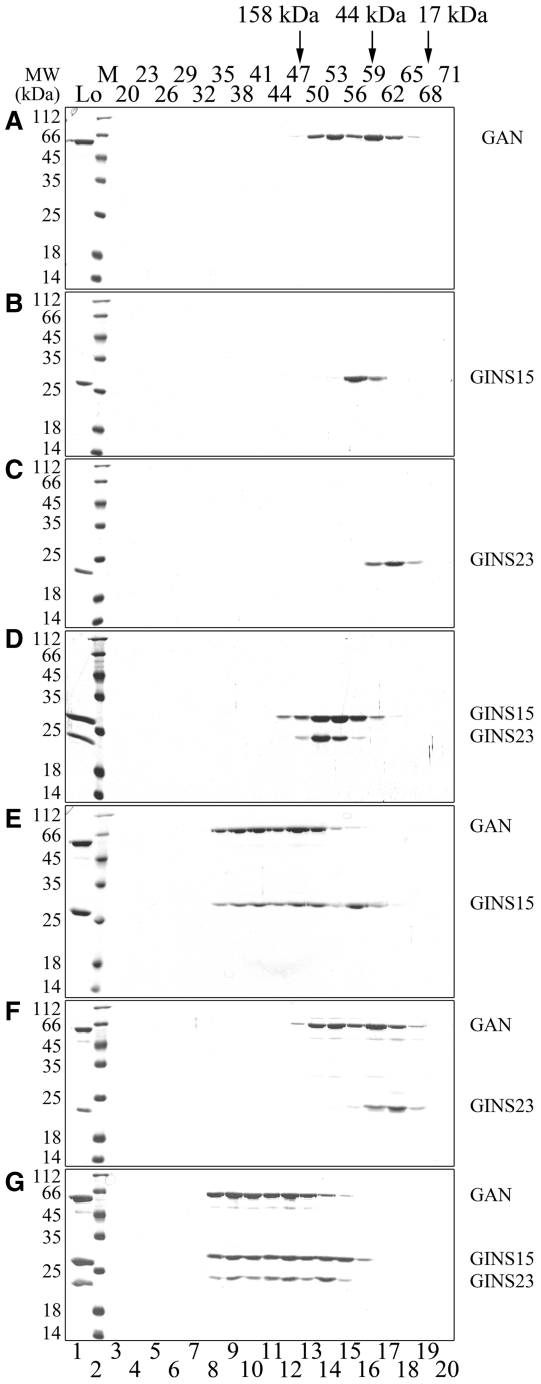

Figure 1. GAN interacts with GINS15. A sample (100 μg) of each protein listed to the right of the corresponding panels (**A** through **G**) was subjected to Superdex-200 gel filtration analysis. Aliquots (80 μl) from each fraction were separated by electrophoresis through 12% polyacrylamide–SDS gels and stained with Coomassie brilliant blue (R-250). The fractions in which γ-globulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa) eluted are noted at the top of the figure.

conserved in the GAN–RecJ alignment (Supplementary Figure S2, marked by an asterisk). This was replaced by alanine, and the GAN (D34A) variant had minimal nuclease activity with Mn^{2+} present (Figure 2A; lane 3) and no detectable activity with Mg^{2+} (Figure 2A; lane 9). An *E. coli* RecJ variant with the analogous alanine for aspartate replacement similarly retained a residual nuclease activity, but had ~400-fold lower activity than wild-type RecJ (23).

As expected, digestion of the 30-mer ssDNA substrate was dependent on the GAN concentration (Figure 2B and C) and time of incubation at 70°C (Figure 2D and E) and, under all reaction conditions, the ssDNA substrate was digested almost exclusively to mononucleotides as determined by MS analysis (Figure 3) and when 5′-^{32}P-labeled oligonucleotides were analysed using thin layer chromatography (data not shown). Consistent with GAN degrading the ssDNA molecules processively, there was no detectable accumulation of intermediate length oligonucleotides (Figure 2B and D). Providing further support for this conclusion, essentially the same pattern of digestion products was obtained when the 30-mer DNA was Cy3-labeled at the 3′-terminus or Cy5-labeled at the 5′-terminus (Figure 2C and E). As illustrated in Figure 2E, the 30-mer ssDNA was degraded by GAN at a rate of 36.9 ± 13.3 nt/mol/min. Although the monomer–dimer equilibrium remains to be determined, both GAN monomers and dimers appeared to be active (Supplementary Figure S3).

GAN acts as 5′-exonuclease on ssDNA

GAN was incubated with the nucleic acids illustrated in Figure 4A to determine the direction of exonuclease digestion and substrate specificity. No nuclease activity was detected with linear dsDNA (substrate III; Figure 4A and B) or ds circular plasmid DNA (data not shown). This inability to degrade dsDNA provided an assay to determine the direction of GAN digestion of ssDNA. There was no digestion of a 3′-ssDNA extension from a dsDNA molecule (substrate II; Figure 4A and B) whereas a 5′-ssDNA extension was rapidly hydrolyzed (substrate I; Figure 4A and B). Digestion of the 5′-ssDNA extension was then followed by a much slower separation and degradation of the two strands of the dsDNA. In all experiments, DNA hydrolysis yielded mononucleotides (Figure 4A; lane 4; Figure 4C and D). Essentially the same results were obtained when the substrates were [^{32}P]-labeled rather than dye-labeled. Based on the results obtained, GAN is a 5′ → 3′ ssDNA-specific exonuclease, and given the almost simultaneous release of a label attached to either the 5′- or 3′-terminus of a ssDNA substrate (Figure 2C and E), the initial binding of GAN to the 5′-terminus is most likely the rate limiting step in the degradation of these substrates.

To further confirm the requirements for nuclease activity, experiments were undertaken using linear and circular versions of a 200-mer ssDNA substrate (Figure 5A and B). When linear, this oligonucleotide was readily degraded by GAN but, when circular, remained intact. A small amount of linear 200-mer present in the circular 200-mer preparations was degraded by exposure to GAN (Figure 5A). As controls, the 200-mer substrates were incubated with a mixture of *E. coli* exonucleases I and III and, as expected, these

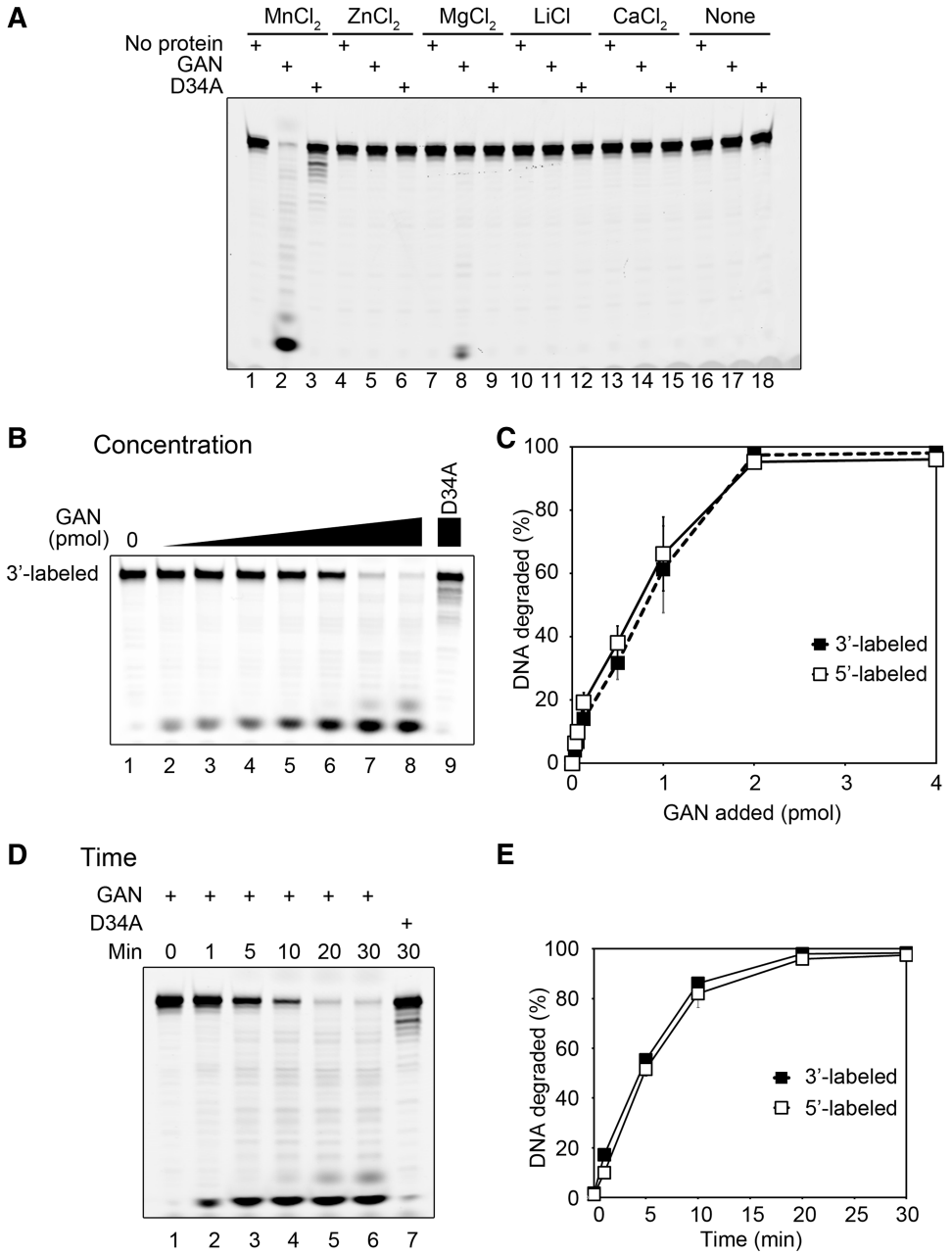

Figure 2. GAN is a Mn^{2+}-dependent exonuclease. (**A**) Electrophoretic separation of the products of reaction mixtures (20 µl) that contained 7.5 pmol 3′-Cy3-labeled oligonucleotides (A4, Supplementary Table S1), 4 pmol of GAN or GAN (D34A), 25 mM Tris–HCl (pH 7.5), 125 µg BSA/ml and 2 mM $MnCl_2$, $ZnCl_2$, $MgCl_2$, LiCl or $CaCl_2$, that were incubated at 70°C for 20 min. (**B**) Reaction mixtures (20 µl) containing 7.5 pmol of 3′-Cy3 labeled oligonucleotides (A4, Supplementary Table S1) in 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$, 125 µg BSA/ml and 0, 0.06, 0.125, 0.25, 0.5, 1, 2 and 4 pmol (lanes 1–8) of GAN protein. In lane 9, 4 pmol of GAN (D34A) was added in place of wild-type GAN. The reaction mixtures were incubated at 70°C for 20 min and the products were separated by electrophoresis, visualized and quantified by phosphorimaging. (**C**) Quantification of the A4 digestion products shown in (B), and of digestion products generated from 5′-Cy5 labeled oligonucleotide (A1, Supplementary Table S1) under the same reaction conditions. The data shown are the averages, with standard deviations, from three independent experiments with each substrate. (**D**) Separation of the reaction products generated in reaction mixtures (20 µl) that contained 10 pmol of 3′-Cy3-labeled oligonucleotide (A4, Supplementary Table S1) 1 pmol GAN or GAN (D34A), 2 mM $MnCl_2$ and 125 µg/ml BSA in 25 mM Tris–HCl (pH 7.5) incubated at 70°C for 0, 1, 5, 10, 20 and 30 min. (**E**) Quantification of the A4 digestion products shown in (D), and of digestion products generated from 5′-Cy5 labeled oligonucleotide (A1, Supplementary Table S1) under the same reaction conditions. The data shown are the averages, with standard deviations, from three independent experiments with each substrate.

well-characterized enzymes degraded only the linear substrate (Figure 5A). GAN appears to specifically degrade DNA as no degradation of RNA could be observed under all conditions tested (Figure 5C).

GINS15 stimulates the GAN nuclease activity

To determine if GINS affected the activity of GAN, its nuclease activity was assayed in the presence or absence of GINS15, GINS23 or the [$GINS15_2$–$GINS23_2$] complex.

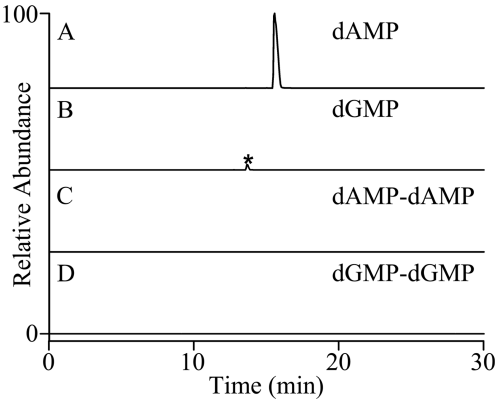

Figure 3. GAN nuclease produces mononucleotides. Reconstructed ion chromatograms of the products of GAN digestion of 5′-AAAAAAGG identified by HPLC–MS/MS analysis. The products were evaluated specifically for the presence of (**A**) dAMP, (**B**) dGMP, (**C**) dAMP–dAMP and (**D**) dGMP–dGMP. The ratio of dAMP to dGMP was 9:1 and deviated from the expected 3:1 ratio due to differences in the response factor/sensitivities of dAMP and dGMP. Asterisk indicates peak detection with signal to noise ratio > 20. The absence of dinucleotides, dAMP–dAMP (C) and dGMP–dGMP (D), are indicated by the traces. MS/MS transitions are: dAMP (m/z 330.1 → 195.1), dGMP (m/z 346.1 → 195.1), dAMP–dAMP (m/z 643.1 → 330.1) and dGMP–dGMP (m/z 675.1 → 346.1). Essentially, the same results were obtained when the substrate was 5′-GGAAAAAA.

The presence of GINS15 stimulated GAN nuclease activity (Figure 6A; lanes 1–5; Supplementary Figure S4) but had no stimulatory effects on GAN (D34A) (Figure 6A; lanes 11 and 12) although GINS15 and GAN (D34A) were found to interact (Supplementary Figure S5). GINS15 alone had no nuclease activity (Figure 6A; lane 6). Incubation with GINS23 did not stimulate the GAN nuclease activity (Figure 6A; lane10) and the presence of [$GINS15_2$–$GINS23_2$] complexes had the same stimulatory effect as the presence of GINS15 alone (Figure 6A; lanes 5 and 8). Incubation with GINS15 did not change the substrate specificity of GAN; with GINS15 present GAN still did not degrade dsDNA or RNA (Supplementary Figure 6; data not shown).

DNA polymerase D interacts with GAN *in vivo*

GAN was identified as a protein that formed a stable complex *in vivo* with His_6-tagged GINS15 (4). To determine if additional proteins formed complexes *in vivo* with GAN, a *T. kodakaraensis* strain was constructed that synthesized His_6-tagged GAN (Supplementary Figure S1) and proteins that co-purified with this tagged protein from cell lysates were identified by

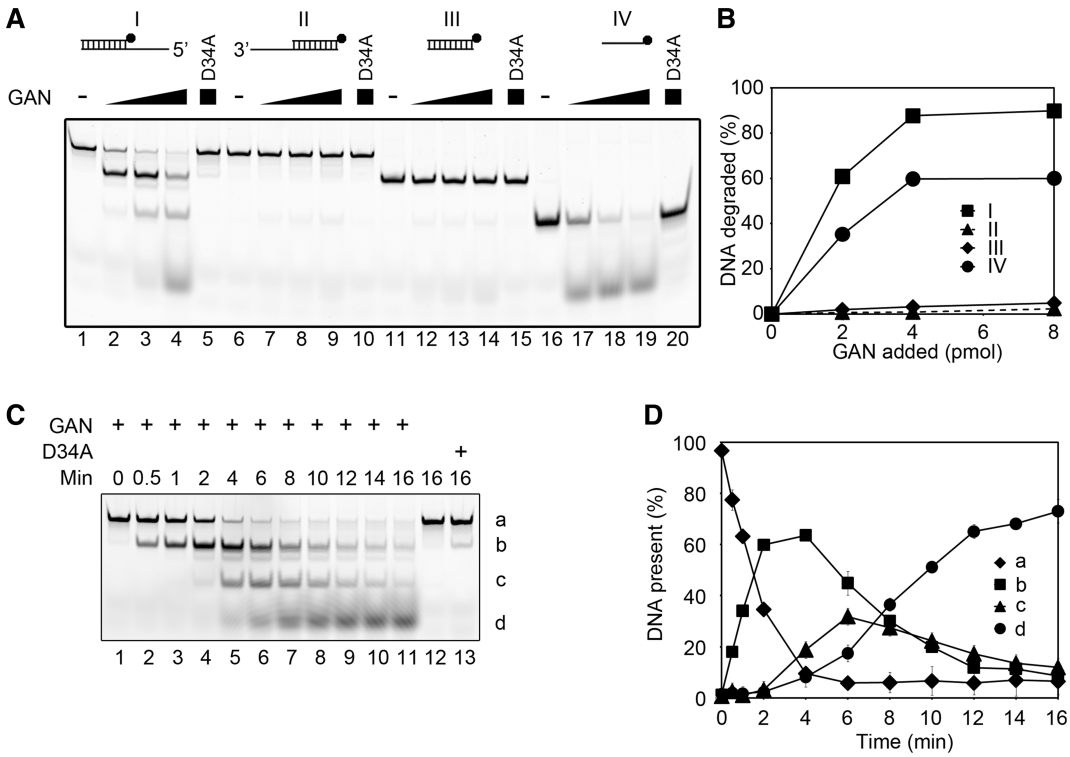

Figure 4. GAN is a 5′ → 3′ exonuclease. (**A**) The structures of the DNA substrates, with the position of the Cy3-label indicated (dark circle), are illustrated above the electrophoretic separations of the GAN digestion products. Reaction mixtures (20 µl) that contained 7.5 pmol of Cy3-labeled substrate, 125 µg BSA/ml, 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$ and 0 (−), 2, 4 or 8 pmol of GAN or 8 pmol of GAN (D34A) were incubated for 15 min at 55°C. The reaction products were separated by electrophoresis through 20% native polyacrylamide gels, visualized and quantified by phosphorimaging. (**B**) Average values from three independent repetitions of the experiments shown in (A). (**C**) Native gel electrophoretic separation of the products of digestion of substrate I in reaction mixtures (20 µl) that contained 7.5 pmol of substrate I, 4 pmol of GAN (lanes 1–11) or GAN (D34A) (lane 13), 125 µg BSA/ml, 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$. The reaction mixtures were incubated at 55°C for 0, 0.5, 1, 2, 4, 6, 8, 10, 12, 14 or 16 min. (**D**) Average values, with standard deviations, from three independent repetitions of the experiments shown in **C**. The Cy3-labeled molecules indicated and quantified were a: substrate I (filled diamond); b: ds DNA product (filled square); c: ss oligonucleotides from substrate I (filled triangle); d: mononucleotides (filled circle).

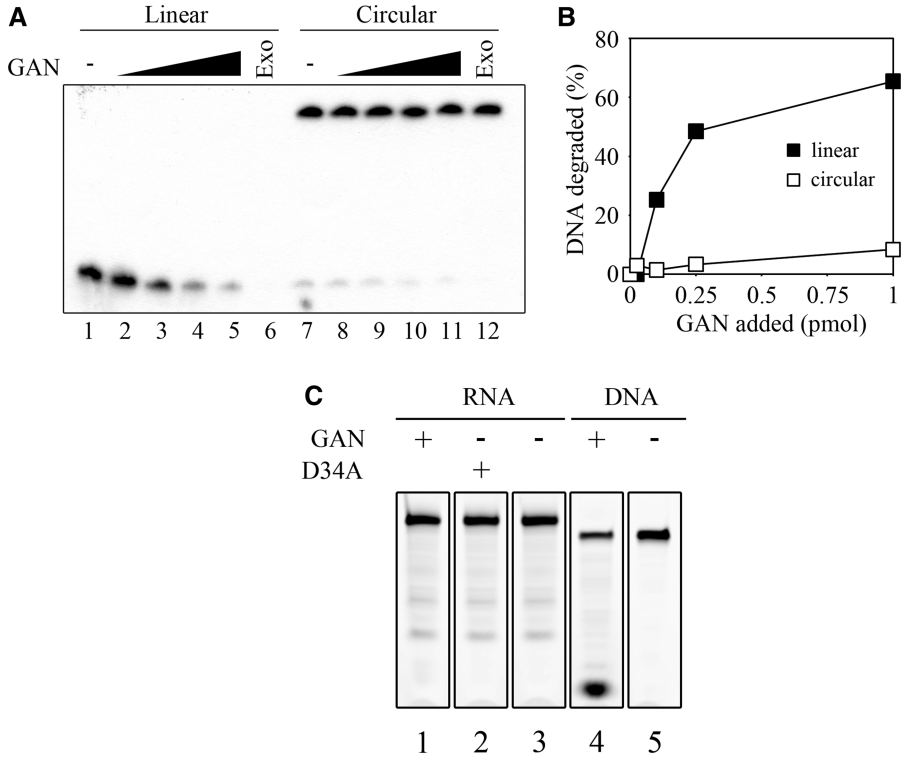

Figure 5. GAN is an ssDNA-specific exonuclease. (**A**) Electrophoretic separation of the products of reaction mixtures (20 µl) that contained 0.75 pmol of [^{32}P]-labeled 200 nt linear or circular ssDNA, 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$ and 125 µg/ml BSA, and 0 (−), 0.025, 0.1, 0.25 and 1 pmol GAN incubated at 60°C for 10 min. The products of incubation of these substrates with a mixture of *E. coli* exonucleases I and III (10 and 50 units, respectively; Exo) for 30 min at 37°C were separated in lanes 6 and 12. (**B**) Quantification of the GAN digestion shown in (A). (**C**) Electrophoretic separation of the products of reaction mixtures (20 µl) that contained 7.5 pmol of Cy3-labled RNA or DNA (A4R and A4, respectively, Supplementary Table S1), 2 pmol of GAN or GAN (D34A), 125 µg/ml BSA, 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$ incubated at 55°C for 10 min. These results were obtained when the experiment was performed with GINS15 and GINS23 present (Supplementary Figure S4).

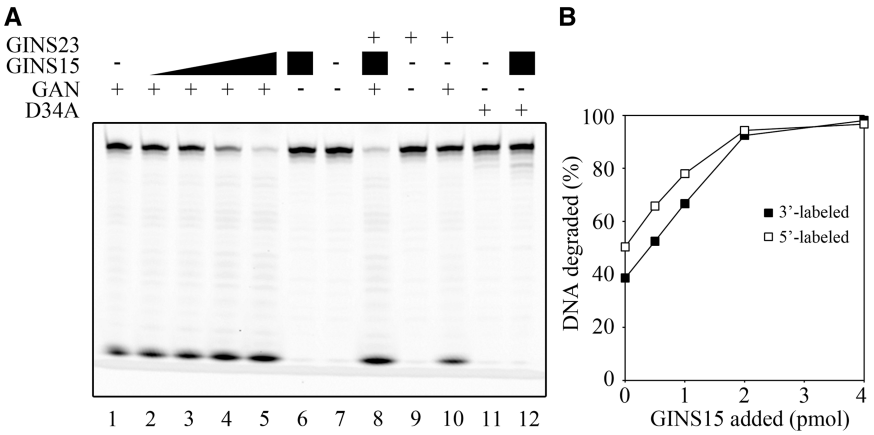

Figure 6. GAN nuclease activity is stimulated by GINS15. (**A**) Electrophoretic separation of the products of digestion of a 3′-Cy3-labeled substrate (A4, Supplementary Table S1) in reaction mixtures (20 µl) that contained 7.5 pmol of substrate, 125 µg BSA/ml, 25 mM Tris–HCl (pH 7.5), 2 mM $MnCl_2$ and, when indicated, 2 pmol GAN or GAN (D34A) in the presence of 0.5, 1, 2 (lanes 2–4) or 4 (lanes 5, 6, 8 and 12) pmol of GINS15, or 4 pmol GINS23 (lanes 8–10 and 12). The reaction mixtures were incubated at 70°C for 4 min. (**B**) Quantification of the A4 digestion products shown in (A), and of digestion products generated from 5′-Cy5-labeled oligonucleotides (A1, Supplementary Table S1) under the same reaction conditions. The data shown are the averages from three independent experiments with each substrate.

MS (Table 1). Both subunits of the GINS complex (GINS15 and GINS23) and both subunits of the euryarchaeal specific Pol D were predominant among the co-purified proteins. This provides strong reciprocal evidence for the presence of a GAN–GINS complex *in vivo*, and is consistent with a larger replisome structure in which the GAN–GINS complex also interacts with Pol D.

Table 1. Proteins that co-purified with GAN

Gene #	Score	MW (Da)	Peptide matches	Percent coverage	Function
TK1903	1354	150 190	54	19.5	Pol D-L
TK1902	947	80 848	43	28.4	Pol D-S
TK1252	746	52 858	26	37.1	GAN
TK1619	155	19 154	6	14.6	GINS23
TK1637	105	29 250	4	13.8	Proteasome subunit alpha
TK1496	97	22 992	3	12.9	30S ribosomal protein S2
TK0536	87	21 583	2	14.9	GINS15
TK1748	81	125 718	3	3.5	Isoleucyl-tRNA synthetase
TK0847	79	8671	4	38.6	Hypothetical protein TK0847
TK2217	79	44 001	2	3.3	2-amino-3-ketobutyrate coenzyme A ligase
TK0593	73	45 867	3	6.7	Unknown
TK2157	54	36 667	3	6	Unknown
TK2106	50	46 763	2	3.5	Phosphopyruvate hydratase
TK1940	48	39 764	3	11.2	Small-conductance mechanosensitive channel
TK0171	48	41 290	2	2.8	Unknown
TK0566	45	96 249	2	1.6	DEAD/DEAH box RNA helicase
TK0714	43	74 602	3	2.7	Iron(II) transport protein B
TK2253	41	28 743	13	3.2	Unknown
TK0470	40	198 006	2	0.6	Reverse gyrase
TK2276	39	23 395	2	16.9	Orotidine 5′-phosphate decarboxylase
TK1448	37	39 518	2	2	5,10-methylenetetrahydrofolate reductase
TK2270	29	7456	2	30.2	Unknown
TK2255	28	48 783	2	3	Bifunctional phosphatase/dolichol-phosphate glucosyltransferase
TK0263	27	43 206	2	3.5	3-phosphoshikimate 1-carboxyvinyltransferase

Proteins with at least two peptides matches are listed with their molecular weight, MASCOT score, and the percentage of the amino acid sequence covered by the matching peptides (see text for further details).

DISCUSSION

Genes encoding well-conserved GAN homologues are widely distributed in the *Euryarchaea* but are not present in *Crenarchaea*, *Bacteria* or in the *Nanoarchaeum equitans* genome (Supplementary Figure 7). The most parsimonious interpretation is therefore that the GAN nuclease evolved in the archaeal domain, and specifically in the euryarchaeal lineage after separation from the crenarchaeal lineage. It seems likely that GAN plays a conserved role in euryarchaeal DNA metabolism, and given its robust $5' \rightarrow 3'$ exonuclease activity, stimulated by association with GINS15, GAN could play a role directly in euryarchaeal genome replication. In *Bacteria*, the 5′ exonuclease activity of DNA polymerase I (Pol I) is required to remove the RNA primers from Okazaki fragments prior to gap filling and ligation of the lagging strand [Figure 7A; (24)] but eukaryal and archaeal DNA polymerases do not have 5′ exonuclease activity. Instead, the flap endonuclease 1 (Fen1) removes the RNA primers on the lagging strand. The strand displacement activity of eukaryal DNA polymerase δ (Polδ) can generate flap structures that contain the RNA primer (Figure 7B) and are recognized and cleaved by Fen1 (25,26). Dna2 helicase/nuclease may also participate in the resection of the flap structures (27). However, as Fen1 is not essential for viability (28,29), alternative mechanism(s) must exist to mature the eukaryal Okazaki fragments. In the *Euryarchaea,* which includes *T. kodakaraensis*, Pol D may synthesize the lagging strand (30). This polymerase does have strand displacement activity (30,31) which, on encountering an Okazaki fragment, could generate the 5′ ssDNA required for archaeal Fen1 digestion (Figure 7C–I). Since Fen1 is also not essential for *Euryrachaea* viability (32) it is possible that the 5′ ssDNA generated by the Pol D displacement reaction can be removed by GAN (Figure 7C-II). While bacterial and eukaryal Okazaki fragments are initiated by short ribonucleotides, lagging strand synthesis in *Archaea* may involve only deoxynucleotide primers (33). If this was the case, their degradation could be catalyzed by the $5' \rightarrow 3'$ exonuclease activity of GAN. Here, we have established that a 5′ ssDNA extension is required to initiate GAN activity *in vitro* but, *in vivo*, as a component of a replisome complex, GAN might have the ability to initiate DNA degradation from an unsealed nick in dsDNA. If this were the case, then GAN could provide a function equivalent to the $5' \rightarrow 3'$ exonuclease activity of bacterial Pol I (Figure 7C-III).

The protein encoded by SSO0295 in *S. solfataricus* co-purified with the GINS complex from this crenarchaeon (14) and, in common with GAN, SSO0295p also has some limited sequence similarity to the DNA-binding domain of *E. coli* RecJ (14). One possibility for this similarity may be that GAN and SSO0295, via their DNA-binding domains, recruit the GINS complex to the replication fork. GAN could also recruit Pol D (Table 1) and help maintain the structure of a larger replisome complex containing Pol D, helicase and/or primase components. As part of the CMG complex, GINS and Cdc45 are likely to play important structural roles in the eukayal replication machinery. To date, no archaeal homologues of Cdc45 have been identified. Though the structure of Cdc45 is not yet known, a prediction of its structure by the HHpred program (34) revealed a significant structural similarity to RecJ and

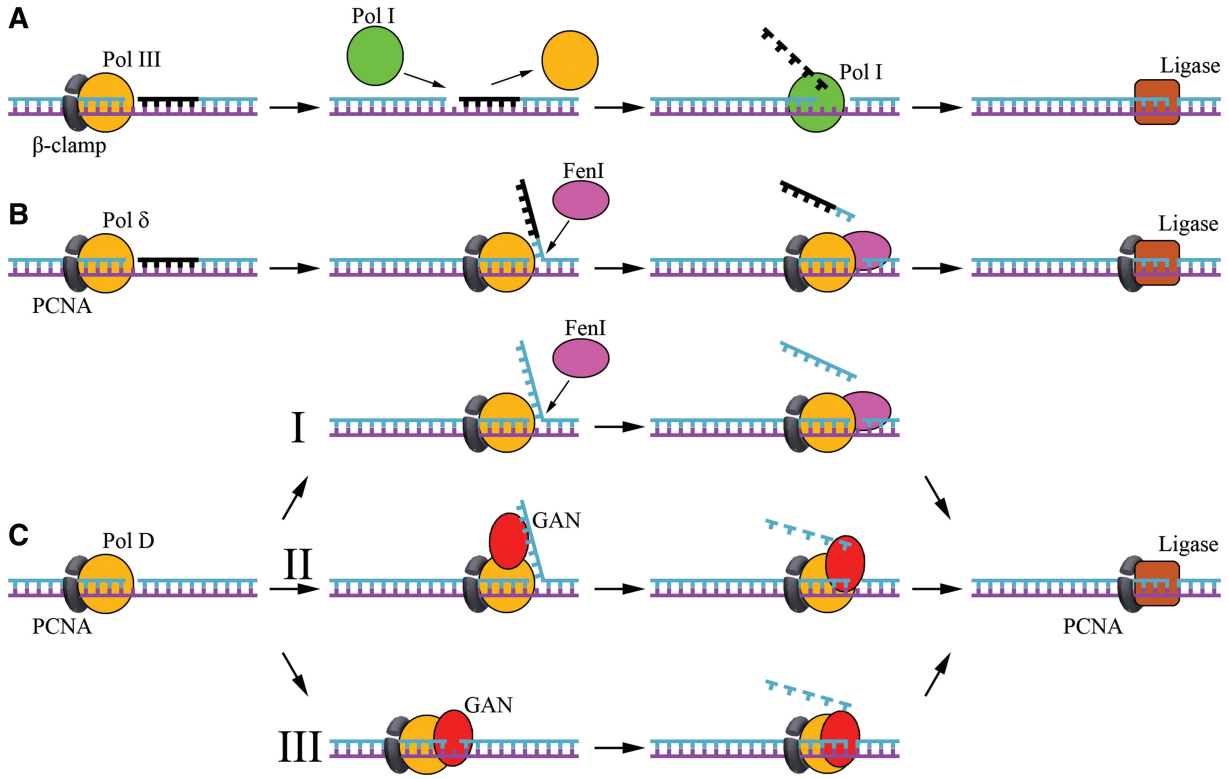

Figure 7. Proposed role for GAN during archaeal DNA replication. Models for lagging strand DNA maturation in the (**A**) *Bacteria*, (**B**) *Eukarya* and (**C**) *Archaea*. See text for details.

hence to GAN. We conjecture that GAN might be the archaeal homologue of Cdc45 and also play a structural role in archaeal replication.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

Specialty oligonucleotide synthesis was supported as part of a NIST/UMD Joint Program in Bioscience Measurements (70NANB9H9196).

FUNDING

National Science Foundation (MCB-0815646 to Z.K.); National Institutes of Health (GM034559 to J.H.); GM53185 to J.N.R.; 1F32-GM073336 to T.J.S.). Funding for open access charge: National Science Foundation.

Conflict of interest statement. None declared.

REFERENCES

1. Grabowski,B. and Kelman,Z. (2003) Archaeal DNA replication: eukaryal proteins in a bacterial context. *Annu. Rev. Microbiol.*, **57**, 487–516.
2. Kelman,L.M. and Kelman,Z. (2003) Archaea: an archetype for replication initiation studies? *Mol. Microbiol.*, **48**, 605–615.
3. Barry,E.R. and Bell,S.D. (2006) DNA replication in the archaea. *Microbiol. Mol. Biol. Rev.*, **70**, 876–887.
4. Li,Z., Santangelo,T.J., Čuboňová,L., Reeve,J.N. and Kelman,Z. (2010) Affinity purification of an archaeal DNA replication protein network. *MBio.*, **1**, e00221–e00210.
5. Moyer,S.E., Lewis,P.W. and Botchan,M.R. (2006) Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl Acad. Sci. USA*, **103**, 10236–10241.
6. 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.
7. Labib,K. and Gambus,A. (2007) A key role for the GINS complex at DNA replication forks. *Trends Cell Biol.*, **17**, 271–278.
8. 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.
9. Kubota,Y., Takase,Y., Komori,Y., Hashimoto,Y., Arata,T., Kamimura,Y., Araki,H. and Takisawa,H. (2003) A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes Dev.*, **17**, 1141–1152.
10. Yang,X., Gregan,J., Lindner,K., Young,H. and Kearsey,S.E. (2005) Nuclear distribution and chromatin association of DNA polymerase α-primase is affected by TEV protease cleavage of Cdc23 (Mcm10) in fission yeast. *BMC Mol. Biol.*, **6**, 13.
11. MacNeill,S.A. (2010) Structure and function of the GINS complex, a key component of the eukaryotic replisome. *Biochem. J.*, **425**, 489–500.
12. Pomerantz,R.T. and O'Donnell,M. (2007) Replisome mechanics: insights into a twin DNA polymerase machine. *Trends Microbiol.*, **15**, 156–164.
13. Swiatek,A. and Macneill,S.A. (2010) The archaeo-eukaryotic GINS proteins and the archaeal primase catalytic subunit PriS share a common domain. *Biol. Direct.*, **5**, 17.

14. 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.
15. 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. USA*, **104**, 12685–12690.
16. 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.
17. Fukui,T., Atomi,H., Kanai,T., Matsumi,R., Fujiwara,S. and Imanaka,T. (2005) Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res.*, **15**, 352–363.
18. Shin,J.H., Jiang,Y., Grabowski,B., Hurwitz,J. and Kelman,Z. (2003) Substrate requirements for duplex DNA translocation by the eukaryal and archaeal minichromosome maintenance helicases. *J. Biol. Chem.*, **278**, 49053–49062.
19. Kasiviswanathan,R., Shin,J.H., Melamud,E. and Kelman,Z. (2004) Biochemical characterization of the *Methanothermobacter thermautotrophicus* minichromosome maintenance (MCM) helicase N-terminal domains. *J. Biol. Chem.*, **279**, 28358–28366.
20. Sato,T., Fukui,T., Atomi,H. and Imanaka,T. (2003) Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.*, **185**, 210–220.
21. Kasiviswanathan,R., Shin,J.H. and Kelman,Z. (2005) Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic Acids Res.*, **33**, 4940–4950.
22. Santangelo,T.J., Cubonova,L., James,C.L. and Reeve,J.N. (2007) TFB1 or TFB2 is sufficient for *Thermococcus kodakaraensis* viability and for basal transcription *in vitro*. *J. Mol. Biol.*, **367**, 344–357.
23. Sutera,V.A. Jr, Han,E.S., Rajman,L.A. and Lovett,S.T. (1999) Mutational analysis of the RecJ exonuclease of *Escherichia coli*: identification of phosphoesterase motifs. *J. Bacteriol.*, **181**, 6098–6102.
24. Kornberg,A. and Baker,T.A. (1992) *DNA replication*, 2nd edn. W.H. Freeman, New York.
25. Hubscher,U. and Seo,Y.S. (2001) Replication of the lagging strand: a concert of at least 23 polypeptides. *Mol. Cells*, **12**, 149–157.
26. Burgers,P.M. (2009) Polymerase dynamics at the eukaryotic DNA replication fork. *J. Biol. Chem.*, **284**, 4041–4045.
27. Bae,S.H. and Seo,Y.S. (2000) Characterization of the enzymatic properties of the yeast Dna2 Helicase/endonuclease suggests a new model for Okazaki fragment processing. *J. Biol. Chem.*, **275**, 38022–38031.
28. Reagan,M.S., Pittenger,C., Siede,W. and Friedberg,E.C. (1995) Characterization of a mutant strain of *Saccharomyces cerevisiae* with a deletion of the RAD27 gene, a structural homolog of the RAD2 nucleotide excision repair gene. *J. Bacteriol.*, **177**, 364–371.
29. Zheng,L., Jia,J., Finger,L.D., Guo,Z., Zer,C. and Shen,B. (2011) Functional regulation of FEN1 nuclease and its link to cancer. *Nucleic Acids Res.*, **39**, 781–794.
30. Henneke,G., Flament,D., Hubscher,U., Querellou,J. and Raffin,J.P. (2005) The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication. *J. Mol. Biol.*, **350**, 53–64.
31. Rouillon,C., Henneke,G., Flament,D., Querellou,J. and Raffin,J.P. (2007) DNA polymerase switching on homotrimeric PCNA at the replication fork of the euryarchaea *Pyrococcus abyssi*. *J. Mol. Biol.*, **369**, 343–355.
32. Meslet-Cladiere,L., Norais,C., Kuhn,J., Briffotaux,J., Sloostra,J.W., Ferrari,E., Hubscher,U., Flament,D. and Myllykallio,H. (2007) A novel proteomic approach identifies new interaction partners for proliferating cell nuclear antigen. *J. Mol. Biol.*, **372**, 1137–1148.
33. Lao-Sirieix,S.H., Pellegrini,L. and Bell,S.D. (2005) The promiscuous primase. *Trends Genet.*, **21**, 568–572.
34. Soding,J., Biegert,A. and Lupas,A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, **33**, W244–W248.