IDENTIFICATION OF A NOVEL BINDING MOTIF IN
PYROCOCCUS FURIOSUS DNA LIGASE FOR THE FUNCTIONAL
INTERACTION WITH PROLIFERATING CELL NUCLEAR
ANTIGEN

Shinichi Kiyonari, Kohei Takayama, Hirokazu Nishida, and Yoshizumi Ishino

From the \(^1\)Department of Genetic Resources Technology, Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Fukuoka-shi, Fukuoka 812-8581, Japan, \(^2\)Central Research Laboratories, Hitachi Ltd., 1-280 Higashi-koigakubo, Kokubunji, Tokyo 185-8581, Japan

Running Title: DNA ligase-PCNA interaction

** Address correspondence to: Yoshizumi ISHINO, Department of Genetic Resources Technology, Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, Fukuoka 812-8581, Japan
Tel: +81-92-642-4217. Fax: +81-92-642-3051 E-mail: ishino@agr.kyushu-u.ac.jp

DNA ligase is an essential enzyme for all organisms that catalyzes a nick-joining reaction in the final step of the DNA replication, repair, and recombination processes. Herein we show the physical and functional interaction between DNA ligase and proliferating cell nuclear antigen (PCNA) from the hyperthermophilic euryarchaeon, \textit{Pyrococcus furiosus}. The stimulatory effect of \textit{P. furiosus} PCNA (PfuPCNA) on the enzyme activity of DNA ligase (PfuLig) was observed not at a low ionic strength, but at a high salt concentration, at which a DNA ligase alone cannot bind to a nicked DNA substrate. Based on mutational analyses, we identified the amino acid residues that are critical for the PCNA binding in a loop structure located in the N-terminal DNA binding domain (DBD) of PfuLig. We propose that the pentapeptide motif QKSFF is involved in the PCNA interacting motifs, in which Gln and the first Phe are especially important for the stable binding with PCNA.

DNA ligases catalyze nick sealing reactions via three nucleotidyl transfer steps, as described in recent review articles (1-4). In the first step, DNA ligases form a covalent enzyme-AMP intermediate, by reacting with ATP or NAD\(^+\) as a cofactor (step 1). In the second step, DNA ligases recognize the substrate DNA, and the AMP is subsequently transferred from the ligases to the 5’-phosphate terminus of the DNA, to form a DNA-adenylate intermediate (AppDNA) (step 2). Then, in the final step, the 5’-AppDNA is attacked by the adjacent 3’-hydroxy group of the DNA to form a phosphodiester bond (step 3). Three genes (\textit{LIG1}, \textit{LIG3} and \textit{LIG4}) encoding ATP-dependent DNA ligases have been identified in the human genome to date. Human DNA ligase I (Lig I) is a replicative enzyme that joins Okazaki fragments during the DNA replication process. It is well known that many eukaryotic proteins involved in DNA replication, DNA repair, and cell cycle control interact with a DNA sliding clamp, proliferating cell nuclear antigen (PCNA) (reviewed in 5-7). Extensive studies of the PCNA interacting proteins revealed the existence of a consensus sequence, called the PCNA interacting protein box (PIP box) (5). The PIP box consists of the sequence “Qxxhxxaa”, where “x” represents any amino acid, “h” represents hydrophobic residues (e.g. L, I or M), and “a” represents aromatic residues (e.g. F, Y or W). Furthermore, it has been proposed that other sites of the interacting protein can participate in PCNA binding. For example, a conserved pair of Lys and Ala residues was identified as a PCNA binding motif (KA box) by using a random peptide display library (8). However, the importance of the KA box is not obvious, because a detailed biochemical analysis
of the motif has not been performed to date. In *Escherichia coli*, the corresponding DNA sliding clamp is the β subunit of DNA polymerase III holoenzyme (henceforth referred to as the β clamp), which forms a toroidal dimeric structure (9). A bioinformatics approach revealed that a pentapeptide motif (consensus QL[SD]LF) plays an important role in binding to the β clamp (10).

In higher eukaryotes, human Lig I reportedly forms a stable complex with a PCNA trimer that is topologically linked to duplex DNA via an N-terminal PIP box motif (11, 12). The structures of eukaryotic DNA ligases can be divided into two major domains, an N-terminal noncatalytic domain (NCD) and a C-terminal catalytic domain (CD), which consists of an adenylation domain and an OB-fold domain (3, 13). The crystal structure of human Lig I in complex with a nicked DNA and biochemical analyses of the enzyme revealed that the NCD provides most of the DNA binding affinity (14), and therefore, this domain is called the N-terminal DNA-binding domain (DBD). Although several groups have characterized the physical interactions between human DNA ligase I and PCNA, no stimulatory or inhibitory effect on nick-joining activities has been observed in *vitro* (11, 15). In contrast, a stimulatory effect of PCNA was also reported (16). Thus, the detailed interaction mode between human PCNA and Lig I is somewhat unclear.

In Archaea, the third domain of life, a single homolog of the eukaryotic DNA ligase I has been identified (17-21). Interestingly, although most of the archaeal replicative enzymes have a eukaryotic PIP box at their C-terminus, no clear PCNA binding motif has been observed in the sequences of archaeal DNA ligases (5, 22). Recently, a physical and functional interaction between PCNA and DNA ligase from *Sulfolobus solfataricus* was reported (23). In the *S. solfataricus* DNA ligase (SsoLig), PIP box-like motifs were proposed to exist in the N-terminal region. Furthermore, a mutant SsoLig, which lacks 30 amino acids at the N-terminus, cannot interact with SsoPCNA (23, 24). However, it has not been determined whether the proposed motifs are actually important for the interaction with PCNA.

Here, we show a physical and functional interaction between PCNA and DNA ligase from the hyperthermophilic euryarchaeon, *Pyrococcus furiosus*. The stimulatory effect of *P. furiosus* PCNA (PfuPCNA) on the nick-joining reaction of DNA ligase (PfuLig) was observed under physiological conditions with an extremely high salt concentration (0.5-0.6 M). Furthermore, we show that the pentapeptide sequence QKSFF, in the DBD of PfuLig, plays an important role in binding to PfuPCNA. Interestingly, this motif is located in a loop connecting two α-helices in the DBD, but is not at the N- or C-terminus of PfuLig, based on our crystal structure (Nishida et al. Submitted). We propose a novel PCNA binding motif, which may be located inside, but not at the terminus, of the PCNA-interacting proteins.

**EXPERIMENTAL PROCEDURES**

**Cloning the Pfu DNA ligase gene and its mutants --** The DNA ligase gene (*lig*) was amplified by PCR directly from *P. furiosus* genomic DNA, using the oligonucleotides 5'-GGCCATGGGTTATCTGGAGCTTGCTCAAC-3' and 5'-GCGGATCCTTAGCTTTCCACTTTTCCTTTCATC-3' as the forward and reverse primers, respectively. The amplified gene was cloned into the pGEM-T easy vector (Promega), and its nucleotide sequence was confirmed. The cloned gene was digested by NcoI-BamHI and was inserted into the corresponding sites of pET21d (Novagen). The resultant plasmid was designated as pET-Lig. Amino acid substituted mutations and N-terminally truncated mutations were introduced into the *lig* gene on the pET-Lig plasmid by PCR, using the appropriate primers. Their sequences are available on request.

**Overproduction and purification of the Pfu DNA ligase proteins --** To obtain the
recombinant PfuLig. *E. coli* BL21 codonPlus™(DE3)-RIL (Stratagene) carrying pET-Lig was grown in 1 liter of LB medium, containing 50 µg/ml ampicillin and 34 µg/ml chloramphenicol, at 37 °C. The cells were cultured to an A600 of 0.40, and then the expression of the lig gene was induced by adding isopropyl β-D-thiogalactopyranoside to a final concentration of 1 mM and continuing the culture for 5 h at 37 °C. After cultivation, the cells were harvested and disrupted by sonication in buffer A, containing 50 mM Tris-HCl, pH 8.0, 0.5 mM DTT, 0.1 mM EDTA, and 10% glycerol. The soluble cell extract, obtained by centrifugation (12,000 x g, 20 min), was heated at 80 °C for 20 min. The heat-resistant fraction obtained by centrifugation was treated with 0.15% polyethyleneimine to remove the nucleic acids. The soluble proteins were precipitated by 80% saturated ammonium sulfate precipitation. The precipitate was resuspended in buffer B, containing 50 mM Tris-HCl, pH 8.0, 1 M (NH₄)₂SO₄, 0.5 mM DTT, 0.1 mM EDTA, and 10% glycerol, and was subjected to chromatography on a Hitrap Phenyl column (Amersham Biosciences). The proteins were eluted at 0 M ammonium sulfate, and the eluted proteins were dialyzed against buffer A. The dialysate was loaded onto a Hitrap Heparin column (Amersham Biosciences), and the proteins were eluted at 0.3-0.35 M sodium chloride. The eluted proteins were dialyzed against buffer A, and the dialysate was subjected to chromatography on a Hitrap Q column (Amersham Biosciences). The proteins were eluted at 0.1-0.15 M sodium chloride, pooled and stored at 4°C. The mutated PfuLig proteins prepared in this study were purified by the same procedures. The purity of each protein used in this study was evaluated by SDS-PAGE. No extra band was detected by Comassie brilliant blue staining of the gel containing 2 µg of each purified protein. The protein concentrations were calculated by measuring the absorbance at 280 nm. The theoretical molecular absorption coefficient of the molecule was calculated based on its triptophan and tyrosine content.

**Preparation of the deadenylated enzyme** -- Purified wild type PfuLig protein (1.5 nmol) was incubated with an excess amount of pyrophosphoric acid (10 mM) in 150 µl of the reaction buffer, containing 20 mM Tris-HCl, pH 7.5, 20 mM KCl, 10 mM MgCl₂, 1 mM DTT, and 0.1% Tween 20, at 37 °C for 1 hour. The deadenylated enzyme was purified by applying the reaction mixture to a MicroSpin G-25 column (Amersham Biosciences), which was pre-equilibrated with buffer A.

**Adenyltransferase assay** -- Purified PfuLigs (15 pmol of the wild type or K249A mutant protein) was incubated with 0.05 µCi [α-32P]ATP in a reaction mixture in 20 µl of the reaction buffer, containing 20 mM Tris-HCl, pH 7.5, 20 mM KCl, 10 mM MgCl₂, 1 mM DTT, and 0.1% Tween 20, at 55 °C for 20 min. T4 DNA ligase (9 unit) (Promega) was incubated with 0.05 µCi [α-32P]ATP in a reaction mixture (20 µl), containing 30 mM Tris-HCl, pH 7.8, 10 mM MgCl₂, 10 mM DTT, and 5% polyethylene glycol (MW8000), at 35 °C for 20 min. The products were analyzed by 10% SDS-polyacrylamide gel electrophoresis (SDS-PAGE). After electrophoresis, the gels were dried and were scanned by using an image analyzer, FLA5000 (FUJIFILM), to detect the Ado-32P-DNA Ligase adducts.

**DNA ligation assay** -- The substrate DNA used in the ligation assay was a 49 bp DNA duplex containing a single nick at the center. The 22 mer deoxynucleotide (5'-AATTCGTGCAGGCACTTGCTGCT-3'), which was labeled with 32P at the 5'-terminus, and the 27 mer deoxynucleotide (5'-AGCTATGACCATGATCAGAAT TGCTT-3') were annealed to the 49 mer deoxynucleotidase with a complementary sequence in TAM buffer, containing 40 mM Tris-acetate, pH 7.8, and 0.5 mM magnesium acetate. The purified PfuLig proteins (at different concentrations for each experiment, as described in the figure legends) were
incubated with 5 nM of the nicked DNA substrate, prepared as described above, in 20 µl of the ligation buffer, containing 20 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 1 mM DTT, 0.01 mM ATP, 0.1% Tween 20 and 0.1 mg/ml BSA, at 60 °C for 15 min. Reactions were initiated by the addition of enzyme and were terminated with 5 µl of stop solution, containing 98% formamide, 10 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol. Samples were heated at 100 °C for 5 min and chilled rapidly on ice prior to loading onto a 10% polyacrylamide gel containing 8 M urea. After electrophoresis, the gels were dried and were scanned by using FLA5000 to detect the 32P-labeled DNA. Three independent experiments were carried out in succession for each ligation condition required in this study, and the standard error are shown as vertical lines on the plots in each graph.

Surface plasmon resonance analysis -- The BIAcore system (BIACORE) was used to study the physical interaction between PfuLig and PfuPCNA. Highly purified recombinant PfuLig or PfuPCNA (25) was fixed on a Sensor Chip CM5, research grade (BIACORE), according to the manufacturer’s recommendations. To measure the kinetic parameters, various concentrations of PfuPCNA were applied to the immobilized PfuLig. All measurements were performed at a continuous flow rate of 30 µL/min, in a buffer containing 10 mM HEPES, pH 7.4, 150 mM NaCl, and 0.005% Tween 20. At the end of each cycle, the bound protein was removed by washing with 2 M NaCl. The kinetic constants for PfuPCNA binding to PfuLig were determined from the association and dissociation curves of the sensorgrams, using the BIAevaluation program (BIACORE).

RESULTS

Biochemical properties of P. furiosus DNA ligase -- PfuLig has already been characterized and is commercially available, mainly as a reagent for ligase chain reactions (Stratagene, Patent# US 5506137). In this study, we cloned and purified PfuLig independently and constructed mutant proteins to analyze the structure-function relationships of this enzyme. It was predicted from the primary amino acid sequence similarity that PfuLig is an ATP-dependent DNA ligase. Therefore, we constructed the mutant lig gene encoding PfuLig K249A, in which the lysine at the predicted adenylylation site was substituted by alanine, in parallel with the gene for the wild type PfuLig, and tested their adenylyltransferase activities in the presence of ATP. As shown in Figure 1A, wild type PfuLig can form a covalent enzyme-AMP intermediate by reacting with [α-32P]ATP as a cofactor, but no adenylyltransferase activity was observed with the K249A mutant protein. It was reported that some thermophilic DNA ligases from archaea utilize ADP (26) or NAD+ (19, 27) as a cofactor. However, we detected a distinct activity of PfuLig in the presence of ATP, but not ADP, AMP, and NAD+ (Fig. 1B). A very small amount of ligation product was detected by the reaction with ADP. This result is the same as the case of DNA ligase from Pyrococcus horikoshii in the recent report (21). We think the ligation reaction may be derived from contaminant ATP (1.16 %) in our ADP reagent (Oriental yeast, Co, Osaka) according to the manufactures certificate, and therefore, we concluded that ADP is not an appropriate cofactor for PfuLig, as Shuman described for P. horikoshii DNA ligase (21).

Our crystallographic study of the wild type PfuLig revealed that the protein consists of three distinct domains, the N-terminal DNA binding domain (DBD), the middle adenylylation domain, and the C-terminal OB-fold (oligonucleotide binding fold) domain (Nishida et al., submitted). The last two domains are commonly called the catalytic core domain (CD), which is conserved in one branch of the nucleotidyltransferase superfamily, containing DNA ligase, RNA ligase, and mRNA capping enzymes (reviewed in 4). We made
truncated PfuLigs, including the DBD with amino acids from 1 to 218 (N-terminal domain) and the CD from 219 to 561 (the middle and C-terminal domains) to investigate the functions of each domain (Fig. 1C). A nick-joining assay was performed by using a wild type enzyme (WT) and the mutant PfuLigs, K249A, DBD, and CD (Fig. 1D). The CD protein could not complete the nick sealing reaction, even at high enzyme concentrations. A very small amount of ligation product was observed when the CD protein was added with concentration at 100 times higher than that of the wild type (Fig. 1D, lane CD). The accumulation of AppDNA products implied that the CD protein exhibits a lower activity at the “step 3” reaction. The CD protein shares a structural similarity with the full-length Chlorella virus DNA ligase, which is the smallest ATP-dependent DNA ligase with a distinct activity in vitro (28), and therefore, it is very interesting to investigate why the CD from PfuLig lacks most of the enzyme activity. This result indicates the importance of the DBD for the overall ligation activity of PfuLig, and it can be predicted from the structural similarity that the contribution of the DBD domain to the ligation reaction is also conserved in the eukaryotic DNA ligases.

Pfu DNA ligase can interact with both monomeric and trimeric PCNA proteins -- To determine the physical interaction between PfuLig and PfuPCNA, we first used an immunoprecipitation (IP) method. The PfuPCNA and PfuLig proteins were incubated together and then precipitated by each antiserum. However, significant interactions between them were not detected under several experimental conditions, probably because the protein-protein interaction is too weak to be detected by an IP method (data not shown). We therefore performed surface plasmon resonance (SPR) experiments to analyze the weak PCNA-DNA ligase interactions. The full-length PfuLig was immobilized onto the CM5 BIAcore sensor chip, and subsequently the wild type PfuPCNA and the mutant PfuPCNA D143A/D147A, which is unable to form a stable toroidal structure in solution and thus cannot stimulate P. furiosus DNA polymerase B activity (29), were injected at different concentrations. The physical interactions between the immobilized PfuLig and the two PCNA proteins were identified by the SPR sensorgram (Fig. 2). The calculated equilibrium constant K_D values for the wild type PCNA and the D143A/D147A mutant were 1.1 x 10^{-7} M and 1.4 x 10^{-6} M, respectively. The K_D values reported here are comparable to that of the human PCNA-p21 interaction determined by SPR analysis (K_D: 3.2x10^{-7} M) (30). These findings suggest that the toroidal structure is not necessarily required to form a stable PfuLig-PfuPCNA complex in vitro.

PfuPCNA enhances the ligation activity of PfuLig at a physiological salt concentration -- It is well known that some hyperthermophilic archaea contain strikingly high intracellular potassium ion concentrations. Based on the study of a euryarchaeon, Pyrococcus woesei, which was later proved to be a subspecies of P. furiosus (31), the potassium ion concentration in the hyperthermophilic archaeanal cells was determined to range between 0.5 to 0.6 M (32). We initially examined the effect of increasing salt concentrations on the nick-joining activity of PfuLig by supplementing the reaction with KCl and K-Glu (potassium glutamate) salts. A reduction in the ligation activity was observed with each of these monovalent salts in a concentration-dependent manner, and about 90% inhibition was seen at a 200 mM salt concentration (Fig. 3A). A similar result was obtained in the enzyme assay using NaCl (data not shown). These observations are not specific for PfuLig, as the same phenomena were reported in the characterizations of other DNA and RNA ligases (18, 27, 33).

To determine whether PfuPCNA can stimulate the ligation activity of PfuLig at the physiological ionic strength, the proteins were assayed in a broad range of salt concentrations. The stimulation effect of PfuPCNA on PfuLig
was observed at 0.05-0.2 M KCl, but the effect was decreased over 0.2 M KCl (Fig. 3B). In the same manner, we performed the enzyme assay by adding K-Glu salt, which was reported to be an important factor contributing to the thermostability of archaeal proteins (34). As shown in Figure 3C, the stimulation effect was observed over 0.2 M K-Glu and peaked at 0.3-0.4 M, and residual activity was detected near the physiological potassium concentrations (0.5-0.6 M). Before comparison of the effects of salt concentration between KCl and K-Glu on the PfuPCNA-dependent ligation activity of PfuLig, the ligation reactions were performed with increasing concentrations of PfuPCNA at constant salt concentrations, and appropriate concentrations of PfuPCNA for each reaction in the presence of KCl and K-Glu, respectively, were determined (Supplementary Fig. 1). These results show that a chloride ion (Cl\(^-\)) concentration over 0.2 M, but not this concentration of potassium ion (K\(^+\)) had an inhibitory effect on the enzyme activity of PfuLig. The same result was observed in the characterization of a Holliday junction resolving enzyme, Hjc, from *P. furiosus*, which exhibited the maximum enzyme activity at 0.2 M KCl (35). For the *P. furiosus* enzymes that catalyze nucleic acid modification reactions, high Cl\(^-\) concentrations may affect their activity.

A novel PCNA binding site in the N-terminal DBD of PfuLig -- To determine the region responsible for PCNA binding in PfuLig, we utilized the two truncated mutants, DBD and CD, as shown in Figure 1C. The interactions between PfuPCNA and these truncated PfuLig proteins were examined qualitatively by an SPR analysis. The wild type PfuPCNA was immobilized onto a CM5 BIACore sensor chip, and the two truncated DNA ligase mutants were then injected. The wild type PfuLig and DBD interacted with the immobilized PfuPCNA to almost the same extent, but CD had no binding ability (Fig. 4). This SPR analysis using immobilized PfuPCNA showed very low resonance units as compared with that shown in Figure 2, in which PfuLig was immobilized. This phenomenon often happens in our experience of the SPR analyses using PfuPCNA and its binding proteins. The difference probably depends on the direction of the proteins fixed on the sensor chip. Due to the relatively low resonance units (< 250 RU) observed in this experiment, the equilibrium constant K\(_a\) was not determined. These findings suggest that the N-terminal DBD plays a critical role in the PCNA binding of PfuLig.

In Archaea, it was reported that *S. solfataricus* DNA ligase has a PCNA binding site in its N-terminal region (23). To determine whether the same region of PfuLig is responsible for PCNA binding, we cloned and purified two N-terminally truncated mutants, PfuLigΔN14 (15-561) and PfuLigΔN32 (33-561), on the basis of the crystal structure of PfuLig (Fig. 5A). In addition, we carefully examined the amino acid sequence of the DBD, and found some regions that may be involved in interactions with PfuPCNA. Single amino acid substituted mutants in these regions were prepared to examine their effects on the PCNA-interaction. These regions included a candidate KA box and a PIP box-like motif found in the DBD (Fig. 5B). The Lys\(_{867}\) in the candidate KA box and the two aromatic residues, Phe\(_{106}\) and Phe\(_{107}\), in the PIP box-like sequence were examined by alanine substitutions. To test the stimulation activity of PCNA in the nick sealing reaction (described above) under equivalent conditions, the relative activities of these mutant proteins were determined without PfuPCNA. The F106A/F107A mutant exhibited almost the same activity as that of the wild type PfuLig (Fig. 5C). The decreased activity observed in the two N-terminally truncated mutants, ΔN14 and ΔN32, revealed that the integrity of the DBD is important for the overall ligation activity. In order to determine the region responsible for PCNA binding, we tested the stimulation effect of PfuPCNA on these mutants. As a result, only the ligation activity of F106A/F107A was not
stimulated by PfuPCNA (Fig. 5D). Furthermore, no effect of PfuPCNA was observed with increasing concentrations of the F106A/F107A mutant PfuLig (data not shown). The PfuPCNA-dependent ligation ability of the K67A mutant was not different from that of the wild type (Fig. 5E). We concluded that at least one of the two aromatic residues, Phe$_{106}$ or Phe$_{107}$, plays a crucial role in PCNA binding via a hydrophobic interaction. Gln$_{103}$ and Phe$_{106}$ of PfuLig are critical for the functional interaction with PfuPCNA -- The PIP box-like sequence, $^{103}$QKSFF$^{107}$ described above, is located in a loop structure in the DBD, based on the crystal structure of PfuLig (Fig. 6A). Using three single amino acid substituted mutants, Q103A, F106A, and F107A, we examined the detailed roles of each of these amino acid residues in the $^{103}$QKSFF$^{107}$ sequence. The specific activities of these mutant PfuLigs were confirmed to be the same (Supplementary Fig. 2). The physical interactions between the mutant PfuLig proteins and the immobilized PfuPCNA were analyzed by SPR analyses. The F106A mutant was not able to bind to PCNA, but the Q103A and F107A mutants showed very weak responses (50 RU) as compared to the with wild type PfuLig (240 RU) (Fig. 6B). Next, the stimulatory effect of PCNA on the ligation activity of these mutants was examined in vitro. As a result, the activities of the Q103A and F106A mutants were slightly stimulated by PfuPCNA, whereas the F107A mutant exhibited an intermediate response to PfuPCNA (Fig. 6C). These analyses indicate that Phe$_{106}$ in the $^{103}$QKSFF$^{107}$ sequences is the most important residue in the physical and functional interactions with PfuPCNA. The Gln$_{103}$ residue may stabilize the PCNA-DNA ligase complex after connecting PfuLig to PfuPCNA with Phe$_{106}$ residue. The F107A mutant PfuLig, which possesses the Gln$_{103}$ and Phe$_{106}$ residues, showed very weak binding activity to PfuPCNA in the SPR analysis, comparable to that of the Q103A mutant, but a distinct response to PCNA was retained in the ligation assay. Further analyses will be required to understand the detailed role of Phe$_{107}$ in the functional interaction between PfuLig and PfuPCNA.

**DISCUSSION**

Functional roles of the conserved DBD in eukaryotic DNA ligases -- As shown in our mutational analyses, the integrity of the DBD is important for the overall ligation activity of PfuLig itself (Fig 1D). Furthermore, the other important function of the DBD is to interact with PCNA. The functional interaction between PfuLig and PfuPCNA seems to be stoichiometric. However, much excess amounts of PfuPCNA is required for stimulation of ligation reaction by PfuLig (Fig. 5D). This inconsistency could be explained by the difficulty of PCNA loading onto the DNA fragment in the assay mixture. The PCNA trimer loads by diffusion onto the double-stranded DNA fragment over the ends without clamp loader (RFC) in this case. This is probably a limited process, and efficient loading requires a large stoichiometric excess of PCNA as discussed previously for human Lig I (16).

In Eukarya and Archaea, PCNA binding proteins generally interact with PCNA via a conserved PIP box motif (e.g. archaeal DNA polymerase B, and flap endonuclease 1 have a typical PIP box motif at their C-terminus (reviewed in 7). Human Lig I has a typical PIP box at the N-terminal tail, but PfuLig lacks a long N-terminal tail. We identified the PCNA-interaction motif of PfuLig in a loop structure, which connects two $\alpha$-helices in the N-terminal DBD. Based on the information from the crystal structure of human Lig I (N-terminal truncated mutant) complexed with a nicked duplex DNA, a model structure of the DNA ligase-PCNA complex with 1:1 stoichiometry was proposed (14). This interaction is likely to involve “face to face” binding because of the proteins’ similar sizes and toroidal structure of PCNA. After binding to the PCNA-DNA complex via the PCNA
binding motif in the DBD, the conformations of the CD may change freely to encircle a nicked DNA, because it has no interactive region with PCNA (Fig. 4), and subsequently the enzyme catalyzes the nick-joining reaction.

*PCNA is a scaffold protein for binding to DNA under physiological ionic conditions* -- There have been some contradictory observations about the stimulatory effect of human PCNA on DNA ligase I activity. One group suggested that these discrepancies are due to differences in the experimental conditions (16). As shown in Figure 3B and 3C, the PfuLig activity was inhibited by PCNA at low salt concentrations (0-0.05 M KCl and 0-0.1 M K-Glu, but the activity was enhanced in a salt concentration-dependent manner over 0.05 M and peaked at 0.2 M for KCl, over 0.1 M and peaked at 0.3-0.4 M for K-Glu, respectively. In a previous report on the inhibitory effect of PCNA on human DNA ligase I, the inhibitory effect was observed at 0 and 50 mM NaCl and no effect of PCNA on ligation was observed at 100 mM NaCl (in pH 6.5 buffer) (15). Based on our findings, the stimulatory effect may be observed at 150 mM NaCl, which is near the physiological ionic strength within human cells. However, it is not easy to discuss the differences in the assay conditions, because the salt concentration of each fraction containing the purified recombinant protein is not always obvious from the presented information.

It can be predicted that, by themselves, eukaryotic DNA ligases cannot bind to DNA to catalyze the nick-joining reaction at a physiological salt concentration, but they can recognize the substrate DNA by interacting with PCNA on a nicked DNA. Most of the DNA modification enzymes can interact with substrate DNAs to express their function at low salt concentrations, but lose their activities at high salt concentrations in vitro. The DNA binding abilities of these enzymes themselves are probably inhibited by salt in the cells. Each protein involved in DNA replication and repair has to work at a certain time in the successive processes at their appropriate sites. To control the specific timing and the position for each related protein factor to access the target DNA in vivo, the salt concentration, which prevents non-specific binding of protein factors in the cells, is especially important, and in the case of replication fork progression, for example, PCNA probably functions as a platform to control the order and the sites of interacting proteins involved in this successive reaction process.

*The conserved residues in the novel PCNA binding motif “QKSFF”* -- The well-known PIP box is generally located in the N-terminal or C-terminal region within the peptide chain of the PCNA interacting proteins. However, the PCNA-binding motif, QKSFF, found in this study is in the middle of the PfuLig protein. This novel PCNA binding motif resembles a putative bacterial β clamp binding motif, QL[SD]LF, which is located not only at the terminus but also in the middle of some β clamp interacting proteins. In this bacterial motif, the pair of hydrophobic residues, LF, is important for binding to the β clamp (10). In PfuLig, the corresponding hydrophobic residues, FF, are also important for binding to PCNA, and furthermore, our work showed that the former residue, Phe106, is more critical than Phe107 (Fig. 6C). Moreover, the importance of Gln103 was revealed in our experiments. A structural comparison of the novel motif in the PfuLig crystal with that of the PIP box in the RFC large subunit in the cocrystal with PfuPCNA (36) is shown in Figure 7. Interestingly, the locations of the amino acid residues responsible for the hydrophobic and ionic interactions, respectively, clearly correspond to each other, and especially, the positions of Gln470 and Phe476 in the PIP box of RFCL, corresponding to Gln103 and Phe107 in PfuLig, are remarkably conserved among the PIP box sequences. This new motif may represent a shorter version of the original PIP box. To determine the detailed role of each amino acid in PCNA binding, an X-ray crystallographic structure of the
PfuPCNA-PfuLigase complex will be required.

Interestingly, this novel PIP box motif is widely conserved in the same region of other archaeal DNA ligases (Fig. 8A). The QKSFF sequence is completely conserved, especially in Thermococccals (Pyrococcus and Thermococcus species). In addition, the basic residues located in the region upstream of the motif are also conserved, and especially, the remarkable cluster of basic residues is conserved in the DNA ligases from Thermococccals and some Methanogens. As we proposed previously, based on mutational analyses of the RFC large subunit (RFCL) from P. furiosus (37), these basic residues may function for the formation of the stable Lig-PCNA-DNA complex in these organisms.

We examined the sequences of the eukaryotic DNA ligases, and found that they also have the archaea-type PIP-box in the middle of the peptide chain (Fig. 8B). Eukaryotic Lig I may bind to PCNA at the site corresponding to the motif that we found in this study as discussed above, in addition to the N-terminal PIP box. These analyses show that the PCNA-DNA ligase interaction mode is also interesting from an evolutionary perspective, and we plan to investigate this possibility by introducing mutations into the conserved Glu in the human Lig I protein.

Acknowledgements -- Y. I. was supported by the Human Frontier Science Program. This work was supported in part by the Japan Science and Technology Agency. This work was also supported in part by Grants-in Aid from the Ministry of Education, Culture, Sports, Science, and Technology, Japan, to Y. I. Y. I also acknowledge the Hou-ansha foundation for the support of our work.

REFERENCES

1. Tomkinson, A. E., and Mackey, Z. B. (1998) Mutat. Res. 407, 1-9
2. Doherty, A. J., and Suh, S. W. (2000) Nucleic Acids Res. 28, 4051-4058
3. Wilkinson, A., Day, J., and Bowater, R. (2001) Mol. Microbiol. 40, 1241-1248
4. Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., and Ellenberger, T. (2006) Chem. Rev. 106, 687-699
5. Warbrick, E. (2000) Bioessays 22, 997-1006
6. Maga, G., and Hubscber, U. (2003) J. Cell Sci. 116, 3051-3060
7. Vivona, J. B., and Kelman, Z. (2003) FEBS Lett. 546, 167-172
8. Xu, H., Zhang, P., Liu, L., and Lee, M. Y. (2001) Biochemistry 40, 4512-4520
9. Kong, X. P., Öhrust, R., O'Donnell, M., and Kuriyan, J. (1992) Cell 69, 425-437
10. Dalrymple, B. P., Kongsuwan, K., Wijffels, G., Dixon, N. E., and Jennings, P. A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 11627-11632
11. Levin, D. S., Bai, W., Yao, N., O'Donnell, M., and Tomkinson, A. E. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 12863-12868
12. Montecucco, A., Rossi, R., Levin, D. S., Gary, R., Park, M. S., Motycka, T. A., Ciarrocchi, G., Villa, A., Biamonti, G., and Tomkinson, A. E. (1998) EMBO J. 17, 3786-3795
13. Martin, I. V., and MacNeill, S. A. (2002) Genome Biol. 3, REVIEWS3005.1-7
14. Pascal, J. M., O'Brien, P. J., Tomkinson, A. E., and Ellenberger, T. (2004) Nature 432, 473-478
15. Jonsson, Z. O., Hindges, R., and Hubscber, U. (1998) EMBO J. 17, 2412-2425
16. Tom, S., Henriksen, L. A., Park, M. S., and Bambara, R. A. (2001) J. Biol. Chem. 276, 24817-24825
17. Kletzin, A. (1992) Nucleic Acids Res. 20, 5389-5396
18. Sriskanda, V., Kelman, Z., Hurwitz, J., and Shuman, S. (2000) Nucleic Acids Res. 28, 2221-2228
19. Nakatani, M., Ezaki, S., Atomi, H., and Imanaka, T. (2000) J. Bacteriol. 182, 6424-6433
20. Lai, X., Shao, H., Hao, F., and Huang, L. (2002) Extremophiles 6, 469-477
21. Keppetipola, N., and Shuman, S. (2005) J. Bacteriol. 187, 6902-6908
22. Grabowski, B., and Kelman, Z. (2003) Annu. Rev. Microbiol. 57, 487-516
23. Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J., and Bell, S. D. (2003) Mol. Cell 11, 275-282
24. Dionne, I., and Bell, S. D. (2005) Biochem. J. 387, 859-863
25. Cann, I. K., Ishino, S., Hayashi, I., Komori, K., Toh, H., Morikawa, K., and Ishino, Y. (1999) J. Bacteriol. 181, 6591-6599
26. Jeon, S. J., and Ishikawa, K. (2003) FEBS Lett. 550, 69-73
27. Rolland, J. L., Gueguen, Y., Persillon, C., Masson, J. M., and Dietrich, J. (2004) FEMS Microbiol. Lett. 236, 267-273
28. Ho, C. K., Van Etten, J. L., and Shuman, S. (1997) J. Virol. 71, 1931-1937
29. Matsumiya, S., Ishino, S., Ishino, Y., and Morikawa, K. (2003) Protein Sci. 12, 823-831
30. Knibiehler, M., Goubin, F., Escalas, N., Jonsson, Z. O., Mazarguil, H., Hubscher, U., and Ducommun, B. (1996) FEBS Lett. 391, 66-70
31. Kanoksilapatham, W., González, J. M., Maeder D. L., DiRuggiero, J., and Robb F. T. (2004) Archaea 1, 277-283
32. Scholz, S., Sonnenbichler, J., Schafer, W., and Hensel, R. (1992) FEBS Lett. 306, 239-242
33. Martins, A., and Shuman, S. (2004) J. Biol. Chem. 279, 50654-50661
34. Hettké, C., Bergerat, A., Hausner, W., Forterre, P., and Thomm, M. (1999) Genetics 152, 1325-1333
35. Komori, K., Sakae, S., Fujikane, R., Morikawa, K., Shinagawa, H., and Ishino, Y. (2000) Nucleic Acids Res. 28, 4544-4551
36. Matsumiya, S., Ishino, S., Ishino, Y., and Morikawa, K. (2002) Genes Cells 7, 911-922
37. Nishida, H., Ishino, S., Miyata, T., Morikawa, K., and Ishino, Y. (2005) Genes Genet. Syst. 80, 83-93

Figure legends

FIG 1. ATP dependency and domain structure of PfuLig. (A) Adenylyltransferase assay of wild type (WT) and Lys249 mutant PfuLig. Purified enzymes (64 kDa) were mixed with [α-32P]ATP in the ligation reaction buffer, as described in the experimental procedures. The reaction mixtures were fractionated by SDS-PAGE, and the adenylylated band was detected by autoradiography. T4 DNA ligase (62 kDa) was used as a positive control of a well-known ATP-dependent DNA ligase. The positions and sizes (kDa) of maker polypeptides are indicated on the left. (B) Cofactor dependence of the PfuLig ligation activity. The deadenylylated PfuLig (10 fmol) was subjected to the ligation reaction with 500 fmol of the nicked DNA described in the experimental procedures. The ligation efficiencies were compared with different cofactors, ATP, ADP, AMP, and NAD+. (C) Domain structure of PfuLig. The domain structure of PfuLig is schematically drawn. DBD, AdD, and OB represent DNA binding domain, adenylylation domain, and oligonucleotide binding fold domain, respectively. (D) Domain structure of PfuLig. The domain structure of PfuLig is schematically drawn. DBD, AdD, and OB represent DNA binding domain, adenylylation domain, and oligonucleotide binding fold domain, respectively. Two truncated mutant PfuLigs were constructed based on the domain structure. (D) The ligation abilities were compared among the wild type (10 fmol), K249A (10 fmol), DBD (1 pmol), and CD (AdD + OB) (1 pmol) proteins. The ligation reactions (containing 100 fmol of the nicked DNA) were analyzed by a denaturing PAGE, followed by autoradiography. The upper band on the 22 mer deoxynucleotide substrate comes from the adenylylated form (AppDNA), which is the
intermediate observed at the second step.

**Fig 2.** Physical interaction between PfuLig and PfuPCNA. SPR analysis was performed using a BIAcore system to detect a physical interaction between PfuLig and PfuPCNA. Purified PfuLig was immobilized on a BIAcore sensor chip, and purified PfuPCNA (3 µM) was loaded. The wild type PfuPCNA and a mutant PfuPCNA (D143A/D147A), which cannot form a stable trimeric ring structure, were used to investigate their affinities to PfuLig. The equilibrium constant $K_D$ was calculated from the obtained sensogram.

**Fig 3.** Salt sensitivity of PfuLig and the effect of PfuPCNA on ligation ability. (A) The ligation efficiencies were compared with various KCl and K-Glu concentrations in the reaction mixture, as described in the experimental procedures, with 20 fmol of PfuLig and 100 fmol of nicked DNA, in the absence of PfuPCNA. (B) and (C) Salt dependency of PfuPCNA-mediated stimulation of PfuLig activity. The ligation reactions of PfuLig (2 fmol) and the nicked DNA (100 fmol) were done with increasing concentrations of KCl (B) and K-Glu (C) in the presence or absence of PfuPCNA. The concentration of PfuPCNA is calculated as trimer form.

**Fig 4.** The N-terminal DBD of PfuLig is essential for PfuPCNA-binding. Physical interactions with PfuPCNA were compared among the wild type (WT) and two truncated PfuLigs, DBD and CD, using the BIAcore system. Purified PfuPCNA was immobilized on a BIAcore sensor chip, and purified PfuLig proteins (3 µM) were loaded, in reverse to the experiment shown in Fig. 2.

**Fig 5.** Search for the candidate amino acid sequence motifs in the DNA ligase proteins. (A) Based on the proposed binding region (boxed in black) in *S. solfataricus* DNA ligase (SsoLig) and the secondary structure of PfuLig, two truncated mutant PfuLigs were prepared to investigate the interaction site of PfuLig to PfuPCNA. An asterisk and a colon between the sequences indicate identity and similarity, respectively, of the aligned residues. (B) The amino acid sequence of the DBD was examined carefully, and a new PIP box-like motif was found, in addition to the candidate KA box, as shown with a black box. The regions containing these candidate motifs were aligned with the human Lig I sequence. This PIP box-like motif is located in the loop structure connecting 6th and 7th α-helices (residues 103 to 107) of the PfuLig crystal structure. In the crystal structure of human Lig I (huLig I), a part of the corresponding region is disordered (residues 385 to 392, indicated by dashed line) (14). (C) The ligation activities were compared among the wild type (WT), ΔN14, ΔN32, and F106A/F107A mutant PfuLigs. Various amount of PfuLigs, as indicated, were incubated with 100 fmol of nicked DNA in the reaction mixture as described in the Experimental procedures. The ligation efficiency is plotted as a function of the enzyme concentration (logarithmically shown on the horizontal axis). (D) Stimulatory effect of PfuPCNA on the ligation ability of mutant enzymes. PfuLig (2 fmol WT and F106A/F107A, 4 fmol ΔN14, and 60 fmol ΔN32) was incubated with 100 fmol of nicked DNA and 0-150 nM of PfuPCNA in the reaction mixture containing 0.2 M KCl as described in the Experimental procedures. The ligation activity of PfuLig without KCl in the absence of PfuPCNA is plotted on the Y-axis as a control to show the basic PfuLig activity used for this experiment. (E) The ligation abilities were compared between the wild type (WT) and K67A. The enzyme (2 fmol) was incubated with 100 fmol of nicked DNA and 0-150 nM of PfuPCNA in the reaction mixture containing 0.2 M KCl as described in the Experimental procedures, and the efficiencies were plotted as a function of PfuPCNA concentration. The plot on the Y-axis shows a control as explained in D.
The predicted motif in PfuLig is actually important for PfuPCNA-binding. (A) The position of the predicted PIP box-like motif in PfuLig (stick models surrounded in an oval) is shown in its overall crystal structure, flanked by a close-up stereo view in a box. (B) Using the PfuPCNA-immobilized BIacore sensor chip as shown in Fig. 4, the wild type (WT), Q103A, F106A, F107A, and F106A/F107A mutant PfuLigs (3 μM) were loaded to investigate their physical interactions with PfuPCNA. (C) Stimulatory effect of PfuPCNA on the ligation ability of mutant PfuLigs. PfuLig (2 fmol) was incubated with 100 fmol of nicked DNA and 0-150 nM of PfuPCNA in the reaction mixture containing 0.2 M KCl as described in the Experimental procedures. The ligation activity of PfuLig without KCl in the absence of PfuPCNA is plotted on the Y-axis as a control to show the basic PfuLig activity used for this experiment.

Structural comparison of the novel PCNA binding motif in the PfuLig crystal with that of the PIP box in the RFC large subunit in the cocrystal with PfuPCNA. The two motifs were extracted from each crystal structure (PfuLig: cyan; RFC-L: orange). The side chains were depicted by stick models, and overlaid with the surface representations. The atoms (nitrogen: blue; oxygen: red) were mapped onto the surface.

The novel PCNA binding motif is widely conserved in the archaeal and eukaryotic DNA ligases. (A) Alignment of the amino acid sequences of the putative loop regions containing the novel PIP motif in archaeal DNA ligases. To investigate the conservation of the PCNA binding motifs identified in this study, a protein-protein BLAST search was carried out at the web site (http://www.ncbi.nlm.nih.gov/BLAST/), using the deduced DNA ligase sequences from the 26 completely sequenced archaeal genomes, involving 20 Euryarchaeota (black), 5 Crenarchaeota (red), and 1 Nanoarchaeota 1 (blue), to date. The search results were manually evaluated and the sequences were manually realigned. The basic residues, glutamine (Q), and phenylalanine (F) are highlighted in blue, magenta, and yellow, respectively. (B) Multiple alignments of the amino acid sequences of the loop regions in eukaryotic DNA ligases.

PfuPCNA-dependent ligation activity of PfuLig in the presence of KCl or K-Glu. The ligation reactions containing 2 fmol of PfuLig, 100 fmol of nicked DNA, and increasing amounts of PfuPCNA were carried out as described in the experimental procedures, and the efficiencies were plotted as a function of PfuPCNA concentration. The ligation activity of PfuLig without salt in the absence of PfuPCNA is plotted on the Y-axis as a control to show the basic PfuLig activity used for this experiment.

The ligation abilities of PfuLigs, containing an alanine-substitution in one of the conserved residues of the QKSFF motif. The ligation activities were compared among the wild type (WT), Q103A, F106A, and F107A mutant PfuLigs. Various amounts of PfuLigs, as indicated, were incubated with 100 fmol of nicked DNA in the reaction mixture as described in the Experimental procedures. The ligation efficiency is plotted as a function of the enzyme concentration.
Figure 1 A-D
Figure 2
Figure 3A, 3B
Figure 3C

The graph shows the ligation (fmol) as a function of K-Glu (M) concentration. The data is presented for two conditions: + PCNA (100 nM) and - PCNA. Error bars indicate the standard error of the mean.
Figure 4
Figure 5 A-B
Figure 5C, 5D, 5E
Figure 6A
Figure 6 B
Figure 6 C
Figure 7
Figure 8 A-B
Supplementary Figure 1

PCNA (nM)

Ligation (fmol)

- KCl (0.2 M)
- K-Glu (0.4 M)
Supplementary Figure 2
Identification of a novel binding motif in pyrococcus furiosus DNA ligase for the functional interaction with proliferating cell nuclear antigen

Shinichi Kiyonari, Kohei Takayama, Hirokazu Nishida and Yoshizumi Ishino

J. Biol. Chem. published online July 6, 2006

Access the most updated version of this article at doi: 10.1074/jbc.M603403200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2006/09/18/M603403200.DC1