

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# DNA Replication in Archaea, the Third Domain of Life

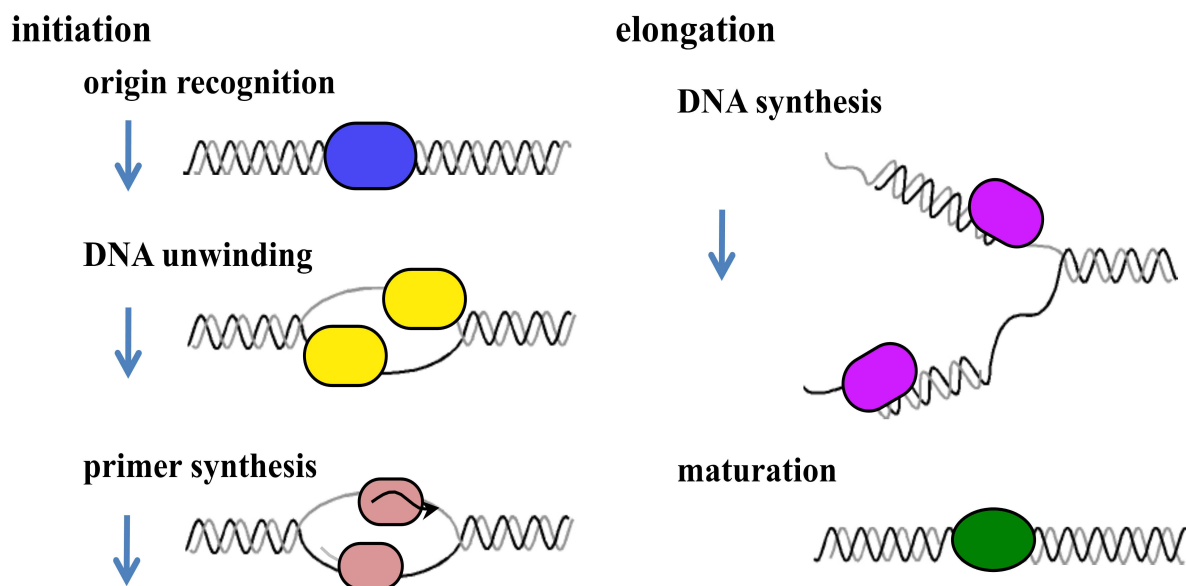
Yoshizumi Ishino and Sonoko Ishino

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/53986>

## 1. Introduction

The accurate duplication and transmission of genetic information are essential and crucially important for living organisms. The molecular mechanism of DNA replication has been one of the central themes of molecular biology, and continuous efforts to elucidate the precise molecular mechanism of DNA replication have been made since the discovery of the double helix DNA structure in 1953 [1]. The protein factors that function in the DNA replication process, have been identified to date in the three domains of life (Figure 1).



**Figure 1.** Stage of DNA replication

|                    | Archaea                               | Eukaryota   | Bacteria                             |
|--------------------|---------------------------------------|---|--------------------------------------|
| initiation         |                                       |   |                                      |
| origin recognition | Cdc6/Orc1                             | ORC   | DnaA                                 |
| DNA unwinding      | Cdc6/Orc1                             | Cdc6  | DnaC                                 |
|                    | MCM                                   | Cdt1  | DnaB                                 |
|                    | GIN5                                  | MCM<br>GIN5<br>Cdc45                                |                                      |
| primer synthesis   | DNA primase                           | Pol $\alpha$ / primase                              | DnaG                                 |
| elongation         |                                       |   |                                      |
| DNA synthesis      | family B                              | family B  | family C                             |
|                    | DNA polymerase<br>(Pol B)             | DNA polymerase<br>(Pol $\delta$ ) (Pol $\epsilon$ ) | DNA polymerase<br>(Pol III)          |
|                    | family D<br>DNA polymerase<br>(Pol D) |   |                                      |
|                    | clamp loader<br>(RFC)                 | clamp loader<br>(RFC)                               | clamp loader<br>( $\gamma$ -complex) |
|                    | clamp<br>(PCNA)                       | clamp<br>(PCNA)                                     | clamp<br>( $\beta$ -clamp)           |
|                    |                                       |   |                                      |
| maturation         | Fen1                                  | FEN1  | Pol I                                |
|                    | Dna2                                  | DNA2  | RNaseH                               |
|                    | DNA ligase                            | DNA ligase  | DNA ligase                           |

**Table 1.** The proteins involved in DNA replication from the three domains of life

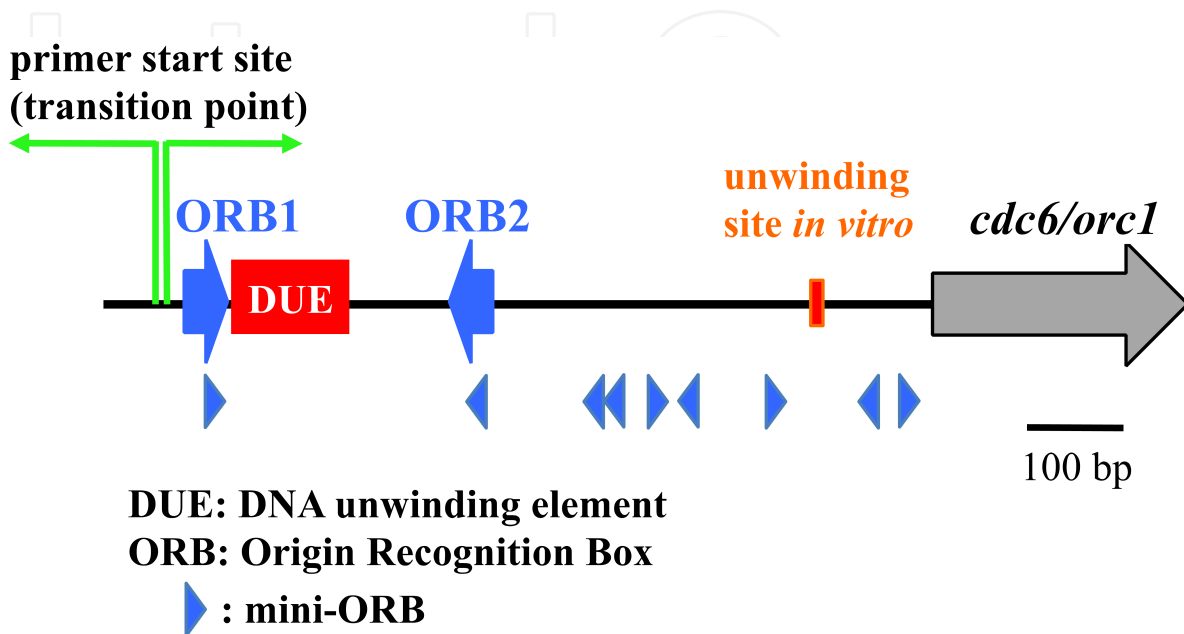
Archaea, the third domain of life, is a very interesting living organism to study from the aspects of molecular and evolutionary biology. Rapid progress of whole genome sequence analyses has allowed us to perform comparative genomic studies. In addition, recent microbial ecology has revealed that archaeal organisms inhabit not only extreme environments, but also more ordinary habitats. In these situations, archaeal biology is among the most exciting of research fields. Archaeal cells have a unicellular ultrastructure without a nucleus, resembling bacterial cells, but the proteins involved in the genetic information processing pathways, including DNA replication, transcription, and translation, share strong similarities with those of eukaryotes. Therefore, most of the archaeal proteins were identified as homologues of many eukaryotic replication proteins, including ORC (origin recognition complex), Cdc6, GINS (Sld5-Psf1-Psf2-Psf3), MCM (minichromosome maintenance), RPA (replication protein A), PCNA (proliferating cell nuclear antigen), RFC (replication factor C), FEN1 (flap endonuclease 1), in addition to the eukaryotic primase, DNA polymerase, and DNA ligase; these are obviously different from bacterial proteins (Table 1) and these proteins were biochemically characterized [2-4]. Their similarities indicate that the DNA replication machi-

neries of Archaea and Eukaryota evolved from a common ancestor, which was different from that of Bacteria [5]. Therefore, the archaeal organisms are good models to elucidate the functions of each component of the eukaryotic type replication machinery complex. Genomic and comparative genomic research with archaea is made easier by the fact that the genome size and the number of genes of archaea are much smaller than those of eukaryotes. The archaeal replication machinery is probably a simplified form of that in eukaryotes. On the other hand, it is also interesting that the circular genome structure is conserved in Bacteria and Archaea and is different from the linear form of eukaryotic genomes. These features have encouraged us to study archaeal DNA replication, in the hopes of gaining fundamental insights into this molecular mechanism and its machinery from an evolutionary perspective. The study of bacterial DNA replication at a molecular level started in about 1960, and then eukaryotic studies followed since 1980. Because Archaea was recognized as the third domain of life later, the archaeal DNA replication research became active after 1990. With increasing the available total genome sequences, the progress of research on archaeal DNA replication has been rapid, and the depth of our knowledge of archaeal DNA replication has almost caught up with those of the bacterial and eukaryotic research fields. In this chapter, we will summarize the current knowledge of DNA replication in Archaea.

## 2. Replication origin

The basic mechanism of DNA replication was predicted as “replicon theory” by Jacob et al. [6]. They proposed that an initiation factor recognizes the replicator, now referred to as a replication origin, to start replication of the chromosomal DNA. Then, the replication origin of *E. coli* DNA was identified as *oriC* (origin of chromosome). The archaeal replication origin was identified in the *Pyrococcus abyssi* in 2001 as the first archaeal replication origin. The origin was located just upstream of the gene encoding the Cdc6 and Orc1-like sequences in the *Pyrococcus* genome [7]. We discovered a gene encoding an amino acid sequence that bore similarity to those of both eukaryotic Cdc6 and Orc1, which are the eukaryotic initiators. After confirming that this protein actually binds to the *oriC* region on the chromosomal DNA we named the gene product Cdc6/Orc1 due to its roughly equal homology with regions of eukaryotic Orc1 and Cdc6, [7]. The gene consists of an operon with the gene encoding DNA polymerase D (it was originally called Pol II, as the second DNA polymerase from *Pyrococcus furiosus*) in the genome [8]. A characteristic of the *oriC* is the conserved 13 bp repeats, as predicted earlier by bioinformatics [9], and two of the repeats are longer and surround a predicted DUE (DNA unwinding element) with an AT-rich sequence in *Pyrococcus* genomes (Figure 2) [10]. The longer repeated sequence was designated as an ORB (Origin Recognition Box), and it was actually recognized by Cdc6/Orc1 in a *Sulfolobus solfataricus* study [11]. The 13 base repeat is called a miniORB, as a minimal version of ORB. A whole genome microarray analysis of *P. abyssi* showed that the Cdc6/Orc1 binds to the *oriC* region with extreme specificity, and the specific binding of the highly purified *P. furiosus* Cdc6/Orc1 to ORB and miniORB was confirmed *in vitro* [12]. It has to be noted that multiple origins were identified in the *Sulfolobus* genomes. It is now well recognized that *Sulfolobus* has three origins and

they work at the same time in the cell cycle [11, 13-16]. Analysis of the mechanism of how the multiple origins are utilized for genome replication is an interesting subject in the research field of archaeal DNA replication. The main questions are how the initiation of replication from multiple origins is regulated and how the replication forks progress after the collision of two forks from opposite directions.



**Figure 2. The *oriC* region in *Pyrococcus* genome.** The region surrounding *oriC* is presented schematically. The ORB1 and ORB2 are indicated by large arrow, and the mini-ORB repeats are indicated by small arrowheads. DUE is indicated in red. The unwinding site, determined by *in vitro* analysis, is indicated in orange. The transition site is indicated by green arrows. The *cdc6/orc1* gene located in downstream is drawn by gray arrow.

### 3. How does Cdc6/Orc1 recognize *oriC*?

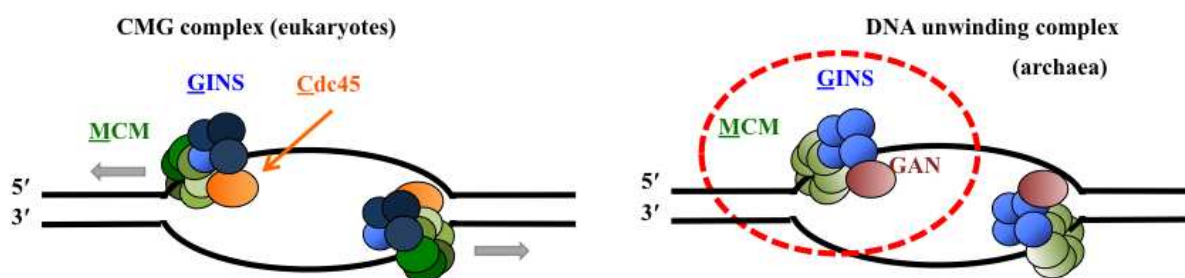
An important step in characterizing the initiation of DNA replication in Archaea is to understand how the Cdc6/Orc1 protein recognizes the *oriC* region. Based upon amino acid sequence alignments, the archaeal Cdc6/Orc1 proteins belong to the AAA<sup>+</sup> family of proteins. The crystal structures of the Cdc6/Orc1 protein from *Pyrobaculum aerophilum* [17] and one of the two Cdc6/Orc1 proteins, ORC2 from *Aeropyrum pernix* (the two homologs in this organism are called ORC1 and ORC2 by the authors) [18] were determined. These Cdc6/Orc1 proteins consist of three structural domains. Domains I and II adopt a fold found in the AAA<sup>+</sup> family proteins. A winged helix (WH) fold, which is present in a number of DNA binding proteins, is found in the domain III. There are four ORBs arranged in pairs on both sides of the DUE in the *oriC* region of *A. pernix*, and ORC1 binds to each ORB as a dimer. A mechanism was proposed in which ORC1 binds to all four ORBs to introduce a higher-order assembly for unwinding of the DUE with alterations in both topology and superhelicity [19]. Furthermore, the crystal structures of *S. solfataricus* Cdc6-1 and Cdc6-3 (two of the three

Cdc6/Orc1 proteins in this organism) forming a heterodimer bound to *ori2* DNA (one of the three origins in this organism) [20], and that of *A. pernix* ORC1 bound to an origin sequence [21] were determined. These studies revealed that both the N-terminal AAA<sup>+</sup> ATPase domain (domain I+II) and C-terminal WH domain (domain III) contribute to origin DNA binding, and the structural information not only defined the polarity of initiator assembly on the origin but also indicated the induction of substantial distortion, which probably triggers the unwinding of the duplex DNA to start replication, into the DNA strands. These structural data also provided the detailed interaction mode between the initiator protein and the *oriC* DNA. Mutational analyses of the *Methanothermobacter thermautotrophicus* Cdc6-1 protein revealed the essential interaction between an arginine residue conserved in the archaeal Cdc6/Orc1 and an invariant guanine in the ORB sequence [22].

*P. furiosus* Cdc6/Orc1 is difficult to purify in a soluble form. A specific site in the *oriC* to start unwinding *in vitro*, was identified using the protein prepared by a denaturation-renaturation procedure recently [23]. As shown in Figure 2, the local unwinding site is about 670 bp away from the transition site between leading and lagging syntheses, which was determined earlier by an *in vivo* replication initiation point (RIP) assay [10]. Although the details of the replication machinery that must be established at the unwound site are not fully understood in Archaea, it is expected to minimally include MCM, GINS, primase, PCNA, DNA polymerase, and RPA, as described below. The following *P. furiosus* studies revealed that the ATPase activity of the Cdc6/Orc1 protein was completely suppressed by binding to DNA containing the ORB. Limited proteolysis and DNase I-footprint experiments suggested that the Cdc6/Orc1 protein changes its conformation on the ORB sequence in the presence of ATP. The physiological meaning of this conformational change has not been solved, but it should have an important function to start the initiation process [24] as in the case of bacterial DnaA protein. In addition, results from an *in vitro* recruiting assay indicated that MCM (Mcm protein complex), the replicative DNA helicase, is recruited onto the *oriC* region in a Cdc6/Orc1-dependent, but not ATP-dependent, manner [24], as described below. However, this recruitment is not sufficient for the unwinding function of MCM, and some other function remains to be identified for the functional loading of this helicase to promote the progression of the DNA replication fork.

#### 4. MCM helicase

After unwinding of the *oriC* region, the replicative helicase needs to remain loaded to provide continuous unwinding of double stranded DNA (dsDNA) as the replication forks progress bidirectionally. The MCM protein complex, consisting of six subunits (Mcm2, 3, 4, 5, 6, and 7), is known to be the replicative helicase “core” in eukaryotic cells [25]. The MCM further interacts with Cdc45 and GINS, to form a ternary assembly referred to as the “CMG complex”, that is believed to be the functional helicase in eukaryotic cells (Figure 3) [26]. However, this idea is still not universal for the eukaryotic replicative helicase.



**Figure 3. DNA-Unwinding complex in eukaryotes and archaea.** The CMG complex is the replicative helicase for the template DNA unwinding reaction in eukaryotes. The archaeal genomes contain the homologs of the Mcm and Gins proteins, but a Cdc45 homolog has not been identified. Recent research suggests that a RecJ-like exonuclease GAN, which has weak sequence homology to that of Cdc45, may work as a helicase complex with MCM and GINS.

Most archaeal genomes appear to encode at least one Mcm homologue, and the helicase activities of these proteins from several archaeal organisms have been confirmed *in vitro* [27-31]. In contrast to the eukaryotic MCM, the archaeal MCMs, consist of a homohexamer or homo double hexamer, having distinct DNA helicase activity by themselves *in vitro*, and therefore, these MCMs on their own may function as the replicative helicase *in vivo*. The structure-function relationships of the archaeal Mcms have been aggressively studied using purified proteins and site-directed mutagenesis [32]. An early report using the ChIP method showed that the *P. abyssi* Mcm protein preferentially binds to the origin *in vivo* in exponentially growing cells [7, 12]. The *P. furiosus* MCM helicase does not display significant helicase activity *in vitro*. However, the DNA helicase activity was clearly stimulated by the addition of GINS (the Gins23-Gins51 complex), which is the homolog of the eukaryotic GINS complex (described below in more detail). This result suggests that MCM works with other accessory factors to form a core complex in *P. furiosus* similar to the eukaryotic CMG complex as described above [31].

Some archaeal organisms have more than two Cdc6/Orc1 homologs. It was found that the two Cdc6/Orc1 homologs, Cdc6-1 and Cdc6-2, both inhibit the helicase activity of MCM in *M. thermautotrophicus* [33, 34]. Similarly, Cdc6-1 inhibits MCM activity in *S. solfataricus* [35]. In contrast, the Cdc6-2 protein stimulates the helicase activity of MCM in *Thermoplasma acidophilum* [36]. Functional interactions between Cdc6/Orc1 and Mcm proteins need to be investigated in greater detail to achieve a more comprehensive understanding of the conservation and diversity of the initiation mechanism in archaeal DNA replication.

Another interesting feature of DNA replication initiation is that several archaea have multiple genes encoding Mcm homologs in their genomes. Based on the recent comprehensive genomic analyses, thirteen archaeal species have more than one *mcm* gene. However, many of the *mcm* genes in the archaeal genomes seem to reside within mobile elements, originating from viruses [37]. For example, two of the three genes in the *Thermococcus kodakarensis* genome are located in regions where genetic elements have presumably been integrated [38]. The establishment of a genetic manipulation system for *T. kodakarensis*, is the first for a hyperthermophilic euryarchaeon [39, 40], and is advantageous for investigating the function of these Mcm proteins. Two groups have recently performed gene disruption experiments for each *mcm* gene [41, 42]. These experiments revealed that the knock-out strains for *mcm1*

and *mcm2* were easily isolated, but *mcm3* could not be disrupted. Mcm3 is relatively abundant in the *T. kodakarensis* cells. Furthermore, an *in vitro* experiment using purified Mcm proteins showed that only Mcm3 forms a stable hexameric structure in solution. These results support the contention that Mcm3 is the main helicase core protein in the normal DNA replication process in *T. kodakarensis*.

The functions of the other two Mcm proteins remain to be elucidated. The genes for Mcm1 and Mcm2 are stably inherited, and their gene products may perform some important functions in the DNA metabolism in *T. kodakarensis*. The DNA helicase activity of the recombinant Mcm1 protein is strong *in vitro*, and a distinct amount of the Mcm1 protein is present in *T. kodakarensis* cells. Moreover, Mcm1 functionally interacts with the GINS complex from *T. kodakarensis* [42]. These observations strongly suggest that Mcm1 does participate in some aspect of DNA transactions, and may be substituted with Mcm3. Our immunoprecipitation experiments showed that Mcm1 co-precipitated with Mcm3 and GINS, although they did not form a heterohexameric complex [42], suggesting that Mcm1 is involved in the replisome or repairsome and shares some function in *T. kodakarensis* cells. Although western blot analysis could not detect Mcm2 in the extract from exponentially growing *T. kodakarensis* cells [42], a RT-PCR experiment detected the transcript of the *mcm2* gene in the cells (Ishino et al., unpublished). The recombinant Mcm2 protein also has ATPase and helicase activities *in vitro*. [41] Therefore, the *mcm2* gene is expressed under normal growth conditions and may work in some process with a rapid turn over. Further experiments to measure the efficiency of *mcm2* gene transcription by quantitative PCR, as well as to assess the stability of the Mcm2 protein in the cell extract, are needed. Phenotypic analyses investigating the sensitivities of the  $\Delta mcm1$  and  $\Delta mcm2$  mutant strains to DNA damage caused by various mutagens, as reported for other DNA repair-related genes in *T. kodakarensis* [43], may provide a clue to elucidate the functions of these Mcm proteins.

*Methanococcus maripaludis* S2 harbors four *mcm* genes in its genome, three of which seem to be derived from phage, a shotgun proteomics study detected peptides originating from three out of the four *mcm* gene products [44]. Furthermore, the four gene products co-expressed in *E. coli* cells were co-purified in the same fraction [45]. These results suggest that multiple Mcm proteins are functional in the *M. maripaludis* cells.

## 5. Recruitment of Mcm to the *oriC* region

Another important question is how MCM is recruited onto the unwound region of *oriC*. The detailed loading mechanism of the MCM helicase has not been elucidated. It is believed that archaea utilize divergent mechanisms of MCM helicase assembly at the *oriC* [46].

An *in vitro* recruiting assay showed that *P. furiosus* MCM is recruited to the *oriC* DNA in a Cdc6/Orc1-dependent manner [24]. This assay revealed that preloading Cdc6/Orc1 onto the ORB DNA resulted in a clear reduction in MCM recruitment to the *oriC* region, suggesting that free Cdc6/Orc1 is preferable as a helicase recruiter, to associate with MCM and bring it to *oriC*. It would be interesting to understand how the two tasks, origin recognition and



MCM recruiting, are performed by the Cdc6/Orc1 protein, because the WH domain, which primarily recognizes and binds ORB, also has strong affinity for the Mcm protein. The assembly of the Mcm protein onto the ORB DNA by the Walker A-motif mutant of *P. furiosus* Cdc6/Orc1 occurred with the same efficiency as the wild type Cdc6/Orc1. The DNA binding of *P. furiosus* Cdc6/Orc1 was not drastically different in the presence and absence of ATP, as in the case of the initiator proteins from *Archaeoglobus fulgidus* [28], *S. solfataricus* [11], and *A. pernix* [19]. Therefore, it is still not known whether the ATP binding and hydrolysis activity of Cdc6/Orc1 regulates the Mcm protein recruitment onto *oriC* in the cells.

One more important issue is the very low efficiency of the Mcm protein recruitment in the reported *in vitro* assay [24]. Quantification of the recruited Mcm protein by the *in vitro* assay showed that less than one Mcm hexamer was recruited to the ORB. The linear DNA containing ORB1 and ORB2, used in the recruiting assay, may not be suitable to reconstitute the archaeal DNA replication machinery and a template that more closely mimics the chromosomal DNA may be required. Additionally, it may be that as yet unidentified proteins are required to achieve efficient *in vitro* helicase loading in the *P. furiosus* cells. Finally, it will ultimately be necessary to construct a more defined *in vitro* replication system to analyze the regulatory functions of Cdc6/Orc1 precisely during replication initiation.

In *M. thermautotrophicus*, the Cdc6-2 proteins can dissociate the Mcm multimers [47]. The activity of Cdc6-2 might be required as the MCM helicase loader in this organism. The interaction between Cdc6/Orc1 and Mcm is probably general. However, the effect of Cdc6/Orc1 on the MCM helicase activity differs among various organisms, as described above. Some other protein factors may function in various archaea, for example a protein that is distantly related to eukaryotic Cdt1, which plays a crucial role during MCM loading in Eukaryota, exists in some archaeal organisms, although its function has not been characterized yet [14].

## 6. GINS

The eukaryotic GINS complex was originally identified in *Saccharomyces cerevisiae* as essential protein factor for the initiation of DNA replication [48]. GINS consists of four different proteins, Sld5, Psf1, Psf2, and Psf3 (therefore, GINS is an acronym for Japanese go-ichi-nisan, meaning 5-1-2-3, after these four subunits). The amino acid sequences of the four subunits in the GINS complex share some conservation, suggesting that they are ancestral paralogs [49]. However, most of the archaeal genomes have only one gene encoding this family protein, and more interestingly, the Crenarchaeota and Euryarchaeota (the two major subdomains of Archaea) characteristically have two genes with sequences similar to Psf2 and Psf3, and Sld5 and Psf1, respectively referred to as Gins23 and Gins51 [31, 49]. A Gins homolog, designated as Gins23, was biochemically detected in *S. solfataricus* as the first Gins protein in Archaea, in a yeast two-hybrid screening for interaction partners of the Mcm protein, and another subunit, designated as Gins15, was identified by mass-spectrometry analysis of an immunoaffinity-purified native GINS from an *S. solfataricus* cell extract. [50]. The *S. solfataricus* GINS, composed of two proteins, Gins23 and Gins15, forms a tetrameric struc-

ture with a 2:2 molar ratio [50]. The GINS from *P. furiosus*, a complex of Gins23 and Gins51 with a 2:2 ratio, was identified as the first euryarchaeal GINS [31]. Gins51 was preferred over Gins15 because of the order of the name of GINS.

The MCM2-7 hexamer was copurified in complex with Cdc45 and GINS from *Drosophila melanogaster* embryo extracts and *S. cerevisiae* lysates, and the "CMG (Cdc45-MCM2-7-GINS) complex" (Figure 3), as described above, should be important for the function of the replicative helicase. The CMG complex was also associated with the replication fork in *Xenopus laevis* egg extracts, and a large molecular machine, containing Cdc45, GINS, and MCM2-7, was proposed as the unwindosome to separate the DNA strands at the replication fork [51]. Therefore, GINS must be a critical factor for not only the initiation process, but also the elongation process in eukaryotic DNA replication. *S. solfataricus* GINS interacts with MCM and primase, suggesting that GINS is involved in the replisome. The concrete function of GINS in the replisome remains to be determined. No stimulation or inhibition of either the helicase or primase activity was observed by the interaction with *S. solfataricus* GINS *in vitro* [50]. On the other hand, the DNA helicase activity of *P. furiosus* MCM is clearly stimulated by the addition of the *P. furiosus* GINS complex, as described above [31].

In contrast to *S. solfataricus* and *P. furiosus*, which each express a Gins23 and Gins51, *Thermoplasma acidophilum* has a single Gins homolog, Gins51. The recombinant Gins51 protein from *T. acidophilum* was confirmed to form a homotetramer by gel filtration and electron microscopy analyses. Furthermore, a physical interaction between *T. acidophilum* Gins51 and Mcm was detected by a surface plasmon resonance analysis (SPR). Although the *T. acidophilum* Gins51 did not affect the helicase activity of its cognate MCM, when the equal ratio of each molecule was tested *in vitro* [52], an excess amount of Gins51 clearly stimulated the helicase activity (Ogino et al., unpublished). In the case of *T. kodakarensis*, the ATPase and helicase activities of MCM1 and MCM3 were clearly stimulated by *T. kodakarensis* GINS *in vitro*. It is interesting that the helicase activity of MCM1 was stimulated more than that of MCM3. Physical interactions between the *T. kodakarensis* Gins and Mcm proteins were also detected [53]. These reports suggested that the MCM-GINS complex is a common part of the replicative helicase in Archaea (Figure 3).

Recently, the crystal structure of the *T. kodakarensis* GINS tetramer, composed of Gins51 and Gins23 was determined, and the structure was conserved with the reported human GINS structures [53]. Each subunit of human GINS shares a similar fold, and assembles into the heterotetramer of a unique trapezoidal shape [54-56]. Sld5 and Psf1 possess the  $\alpha$ -helical (A) domain at the N-terminus and the  $\beta$ -stranded domain (B) at the C-terminus (AB-type). On the other hand, Psf2 and Psf3 are the permuted version (BA-type). The backbone structure of each subunit and the tetrameric assembly of *T. kodakarensis* GINS are similar to those of human GINS. However, the location of the C-terminal B domain of Gins51 is remarkably different between the two GINS structures [53]. A homology model of the homotetrameric GINS from *T. acidophilum* was performed using the *T. kodakarensis* GINS crystal structure as a template. The Gins 51 protein has a long disordered region inserted between the A and B domains and this allows the conformation of the C-terminal domains to be more flexible.

This domain arrangement leads to the formation of an asymmetric homotetramer, rather than a symmetrical assembly, of the *T. kodakarensis* GINS [53].

The Cdc45 protein is ubiquitously distributed from yeast to human, supporting the notion that the formation of the CMG complex is universal in the eukaryotic DNA replication process. However, no archaeal homologue of Cdc45 has been identified. A recent report of bioinformatic analysis showed that the primary structure of eukaryotic Cdc45 and prokaryotic RecJ share a common ancestry [57]. Indeed, a homolog of the DNA binding domain of RecJ has been co-purified with GINS from *S. solfataricus* [50]. Our experiment detected the stimulation of the 5'-3' exonuclease activity of the RecJ homologs from *P. furiosus* and *T. kodakarensis* by the cognate GINS complexes (Ishino et al., unpublished). The RecJ homolog from *T. kodakarensis* forms a stable complex with the GINS, and the 5'-3' exonuclease activity is enhanced *in vitro*; therefore, the RecJ homolog was designated as GAN, from GINS-Associated Nuclease in a very recent paper [58]. Another related report found that the human Cdc45 structure obtained by the small angle X-ray scattering analysis (SAXS) is consistent with the crystallographic structure of the RecJ family members [59]. These current findings will promote further research on the structures and functions of the higher-order unwindosome in archaeal and eukaryotic cells (Figure 3).

## 7. Primase

To initiate DNA strand synthesis, a primase is required for the synthesis of a short oligonucleotide, as a primer. The DnaG and p48-p58 proteins are the primases in Bacteria and Eukaryota, respectively. The p48-p58 primase is further complexed with p180 and p70, to form DNA polymerase  $\alpha$ -primase complex. The catalytic subunits of the eukaryotic (p48) and archaeal primases, share a little, but distinct sequence homology with those of the family X DNA polymerases [60]. The first archaeal primase was identified from *Methanococcus jannaschii*, as an ORF with a sequence similar to that of the eukaryotic p48. The gene product exhibited DNA polymerase activity and was able to synthesize oligonucleotides on the template DNA [61]. We characterized the p48-like protein (p41) from *P. furiosus*. Unexpectedly, the archaeal p41 protein did not synthesize short RNA by itself, but preferentially utilized deoxynucleotides to synthesize DNA strands up to several kilobases in length [62]. Furthermore, the gene neighboring the p41 gene encodes a protein with very weak similarity to the p58 subunit of the eukaryotic primase. The gene product, designated p46, actually forms a stable complex with p41, and the complex can synthesize a short RNA primer, as well as DNA strands of several hundred nucleotides *in vitro* [63]. The short RNA but not DNA primers were identified in *Pyrococcus* cells, and therefore, some mechanism to dominantly use RNA primers exists in the cells [10].

Further research on the primase homologs from *S. solfataricus* [64-66], *Pyrococcus horikoshii* [67-69], and *P. abyssi* [70] showed similar properties *in vitro*. Notably, p41 is the catalytic subunit, and the large one modulates the activity in the heterodimeric archaeal primases. The small and large subunits are also called PriS and PriL, respectively. The crystal structure of

the N-terminal domain of PriL complexed with PriS of *S. solfataricus* primase revealed that PriL does not directly contact the active site of PriS, and therefore, the large subunit may interact with the synthesized primer, to adjust its length to a 7-14 mer. The structure of the catalytic center is similar to those of the family X DNA polymerases. The 3'-terminal nucleotidyl transferase activity, detected in the *S. solfataricus* primase [64, 66], and the gap-filling and strand-displacement activities in the *P. abyssi* primase [70] also support the structural similarity between PriS and the family X DNA polymerases.

A unique activity, named PADT (template-dependent Polymerization Across Discontinuous Template), in the *S. solfataricus* PriSL complex was published very recently [71]. The activity may be involved in double-strand break repair in Archaea.

The archaeal genomes also encode a sequence similar to the bacterial type DnaG primase. The DnaG homolog from the *P. furiosus* genome was expressed in *E. coli*, but the protein did not show any primer synthesis activity *in vitro*, and thus the archaeal DnaG-like protein may not act as a primase in *Pyrococcus* cells (Fujikane et al. unpublished). The DnaG-like protein was shown to participate in RNA degradation, as an exosome component [72, 73]. However, a recent paper reported that a DnaG homolog from *S. solfataricus* actually synthesizes primers with a 13 nucleotide length [74]. It would be interesting to investigate if the two different primases share the primer synthesis for leading and lagging strand replication, respectively, in the *Sulfolobus* cells, as the authors suggested [74]. A proposed hypothesis about the evolution of PriSL and DnaG from the last universal common ancestor (LUCA) is interesting [71].

The *Sulfolobus* PriSL protein was shown to interact with Mcm through Gins23 [50]. This primase-helicase interaction probably ensures the coupling of DNA unwinding and priming during the replication fork progression [50]. Furthermore, the direct interaction between PriSL and the clamp loader RFC (described below) in *S. solfataricus* may regulate the primer synthesis and its transfer to DNA polymerase in archaeal cells [75].

## 8. Single-stranded DNA binding protein

The single-stranded DNA binding protein, which is called SSB in Bacteria and RPA in Archaea and Eukaryota, is an important factor to protect the unwound single-stranded DNA from nuclease attack, chemical modification, and other disruptions during the DNA replication and repair processes. SSB and RPA have a structurally similar domain containing a common fold, called the OB (oligonucleotide/oligosaccharide binding)-fold, although there is little amino acid sequence similarity between them [76]. The common structure suggests that the mechanism of single-stranded DNA binding is conserved in living organisms despite the lack of sequence similarity. *E. coli* SSB is a homotetramer of a 20 kDa peptide with one OB-fold, and the SSBs from *Deinococcus radiodurans* and *Thermus aquaticus* consist of a homodimer of the peptide containing two OB-folds. The eukaryotic RPA is a stable heterotrimer, composed of 70, 32, and 14 kDa proteins. RPA70 contains two tandem repeats of an OB-fold, which are responsible for the major interaction with a single-stranded DNA in its

central region. The N-terminal and C-terminal regions of RPA70 mediate interactions with RPA32 and also with many cellular or viral proteins [77, 78]. RPA32 contains an OB-fold in the central region [79-81], and the C-terminal region interacts with other RPA subunits and various cellular proteins [77, 78, 82, 83]. RPA14 also contains an OB-fold [77]. The eukaryotic RPA interacts with the SV40 T-antigen and the DNA polymerase  $\alpha$ -primase complex, and thus forms part of the initiation complex at the replication origin [84]. The RPA also stimulates Pol $\alpha$ -primase activity and PCNA-dependent Pol  $\delta$  activity [85, 86].

The RPAs from *M. jannaschii* and *M. thermautotrophicus* were reported in 1998, as the first archaeal single-stranded DNA binding proteins [87-89]. These proteins share amino acid sequence similarity with the eukaryotic RPA70, and contain four or five repeated OB-fold and one zinc-finger motif. The *M. jannaschii* RPA exists as a monomer in solution, and has single-strand DNA binding activity. On the other hand, *P. furiosus* RPA forms a complex consisting of three distinct subunits, RPA41, RPA32, and RPA14, similar to the eukaryotic RPA [90]. The *P. furiosus* RPA strikingly stimulates the RadA-promoted strand-exchange reaction *in vitro* [90].

While the euryarchaeal organisms have a eukaryotic-type RPA homologue, the crenarchaeal SSB proteins appear to be much more related to the bacterial proteins, with a single OB fold and a flexible C-terminal tail. However, the crystal structure of the SSB protein from *S. solfataricus* showed that the OB-fold domain is more similar to that of the eukaryotic RPAs, supporting the close relationship between Archaea and Eukaryota [91].

The RPA from *Methanosarcina acetivorans* displays a unique property. Unlike the multiple RPA proteins found in other archaea and eukaryotes, each subunit of the *M. acetivorans* RPAs, RPA1, RPA2, and RPA3, have 4, 2, and 2 OB-folds, respectively, and can act as a distinct single-stranded DNA-binding proteins. Furthermore, each of the three RPA proteins, as well as their combinations, clearly stimulates the primer extension activity of *M. acetivorans* DNA polymerase BI *in vitro*, as shown previously for bacterial SSB and eukaryotic RPA [92]. Architectures of SSB and RPA suggested that they are composed of different combinations of the OB fold. Bacterial and eukaryotic organisms contain one type of SSB or RPA, respectively. In contrast, archaeal organisms have various RPAs, composed of different organizations of OB-folds. A hypothesis that homologous recombination might play an important role in generating this diversity of OB-folds in archaeal cells was proposed, based on experiments characterizing the engineered RPAs with various OB-folds [93].

## 9. DNA polymerase

DNA polymerase catalyzes phosphodiester bond formation between the terminal 3'-OH of the primer and the  $\alpha$ -phosphate of the incoming triphosphate to extend the short primer, and is therefore the main player of the DNA replication process. Based on the amino acid sequence similarity, DNA polymerases have been classified into seven families, A, B, C, D, E, X, and Y (Table 2) [94-98].

The fundamental ability of DNA polymerases to synthesize a deoxyribonucleotide chain is widely conserved, but more specific properties, including processivity, synthesis accuracy,

and substrate nucleotide selectivity, differ depending on the family. The enzymes within the same family have basically similar properties. *E. coli* has five DNA polymerases, and Pol I, Pol II, and Pol III belong to families A, B, and C, respectively. Pol IV and Pol V are classified in family Y, as the DNA polymerases for translesion synthesis (TLS). In eukaryotes, the replicative DNA polymerases, Pol  $\alpha$ , Pol  $\delta$ , and Pol  $\epsilon$ , belong to family B, and the translesion DNA polymerases,  $\eta$ ,  $\iota$ , and  $\kappa$ , belong to family Y [99].

The most interesting feature discovered at the inception of this research area was that the archaea indeed have the eukaryotic Pol  $\alpha$ -like (Family B) DNA polymerases [100-102]. Members of the Crenarchaeota have at least two family B DNA polymerases [103, 104]. On the other hand, there is only one family B DNA polymerase in the Euryarchaeota. Instead, the euryarchaeal genomes encode a family D DNA polymerase, proposed as Pol D, which seems to be specific for these archaeal organisms and has never been found in other domains [95, 105]. The genes for family Y-like DNA polymerases are conserved in several, but not all, archaeal genomes. The role of each DNA polymerase in the archaeal cells is still not known, although the distribution of the DNA polymerases is getting clearer (Table 2) [106].

|                  | families of DNA polymerases       |   |         |       |   |  |                 |
|------------------|-----------------------------------|---|---------|-------|---|--|-----------------|
|                  | A                                 | B   | C       | D     | E   | X  | Y               |
| <b>Archaea</b>   |                                   |   |         |       |   |  |                 |
| Crenarchaeota    |                                   | Pol BI, Pol BII<br>Pol BIII                                 |         |       | Pol E*  |  | Pol Y           |
| Euryarchaeota    |                                   | Pol BI  |         | Pol D | Pol E*  |  | Pol Y           |
| Korarchaeota     |                                   | Pol BI, Pol BII   |         | Pol D |   |  |                 |
| Aigarchaeota     |                                   | Pol BI, Pol BII   |         | Pol D |   |  | Pol Y           |
| Thaumarchaeota   |                                   | Pol BI  |         | Pol D |   |  | Pol Y           |
| <b>Bacteria</b>  |                                   |   |         |       |   |  |                 |
|                  | Pol I                             | Pol II  | Pol III |       |   |  | Pol IV<br>Pol V |
| <b>Eukaryota</b> |                                   |   |         |       |   |  |                 |
|                  | Pol $\theta$<br>Pol $\gamma^{**}$ | Pol $\alpha$ , Pol $\delta$<br>Pol $\epsilon$ , Pol $\zeta$ |         |       | Pol $\beta$ , Pol $\lambda$<br>Pol $\mu$ , Pol $\sigma$ | Pol $\eta$ , Pol $\iota$<br>Pol $\kappa$ |                 |

\* plasmid-encoded

\*\* mitochondrial

**Table 2.** Distribution of DNA polymerases from seven families in the three domains of life.

The first family D DNA polymerase was identified from *P. furiosus*, by screening for DNA polymerase activity in the cell extract [107]. The corresponding gene was cloned, revealing that this new DNA polymerase consists of two proteins, named DP1 and DP2, and that the deduced amino acid sequences of these proteins were not conserved in the DNA polymerase families [8]. *P. furiosus* Pol D exhibits efficient strand extension activity and strong proof-reading activity [8, 108]. Other family D DNA polymerases were also characterized by several groups [109-115]. The Pol D genes had been found only in Euryarchaeota. However, recent environmental genomics and cultivation efforts revealed novel phyla in Archaea: Thaumarchaeota, Korarchaeota, and Aigarchaeota, and their genome sequences harbor the genes encoding Pol D.

A genetic study on *Halobacterium* sp. NRC-1 showed that both Pol B and Pol D are essential for viability [116]. An interesting issue is to elucidate whether Pol B and Pol D work together at the replication fork for the synthesis of the leading and lagging strands, respectively. According to the usage of an RNA primer and the presence of strand displacement activity, Pol D may catalyze lagging strand synthesis [106, 114].

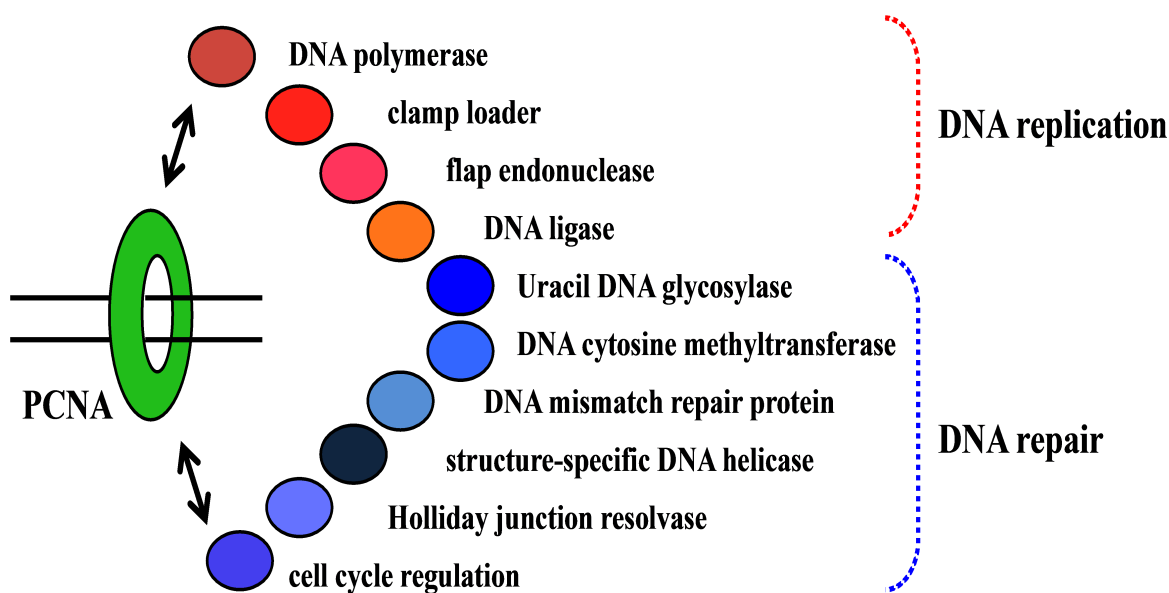
Thaumarchaeota and Aigarchaeota harbor the genes encoding Pol D and crenarchaeal Pol BII [117, 118], while Korarchaeota encodes Pol BI, Pol BII and Pol D [119]. Biochemical characterization of these gene products will contribute to research on the evolution of DNA polymerases in living organisms. A hypothesis that the archaeal ancestor of eukaryotes encoded three DNA polymerases, two distinct family B DNA polymerases and a family D DNA polymerase, which all contributed to the evolution of the eukaryotic replication machinery, consisting of Pol  $\alpha$ ,  $\delta$ , and  $\epsilon$ , has been proposed [120].

A protein is encoded in the plasmid pRN1 isolated from a *Sulfolobus* strain [121]. This protein, ORF904 (named RepA), has primase and DNA polymerase activities in the N-terminal domain and helicase activity in the C-terminal domain, and is likely to be essential for the replication of pRN1 [122, 123]. The amino acid sequence of the N-terminal domain lacks homology to any known DNA polymerases or primases, and therefore, family E is proposed. Similar proteins are encoded by various archaeal and bacterial plasmids, as well as by some bacterial viruses [124]. Recently, one protein, tn2-12p, encoded in the plasmid pTN2 isolated from *Thermococcus nautilus*, was experimentally identified as a DNA polymerase in this family [125]. This enzyme is likely responsible for the replication of the plasmids. Further investigations of this family of DNA polymerases will be interesting from an evolutionary perspective.

## 10. PCNA and RFC

The sliding clamp with the doughnut-shaped ring structure is conserved among living organisms, and functions as a platform or scaffold for proteins to work on the DNA strands. The eukaryotic and archaeal PCNAs form a homotrimeric ring structure [126, 127], which encircles the DNA strand and anchors many important proteins involved in DNA replication and repair (Figure 4). PCNA works as a processivity factor that retains the DNA poly-

merase on the DNA by binding it on one surface (front side) of the ring for continuous DNA strand synthesis in DNA replication (Figure 5). To introduce the DNA strand into the central hole of the clamp ring, a clamp loader is required to interact with the clamp and open its ring. The archaeal and eukaryotic clamp loader is called RFC (Figure 5). The most studied archaeal PCNA and RFC molecules to date are *P. furiosus* PCNA [128-132] and RFC [133-136]. The PCNA and RFC molecules are essential for DNA polymerase to perform processive DNA synthesis. The molecular mechanism of the clamp loading process has been actively investigated [137] (Figure 5). An intermediate PCNA-RFC-DNA complex, in which the PCNA ring is opened with out-of plane mode, was detected by a single particle analysis of electron microscopic images using *P. furiosus* proteins (Figure 6) [138]. The crystal structure of the complex, including the ATP-bound clamp loader, the ring-opened clamp, and the template-primer DNA, using proteins from bacteriophage T4, has recently been published [139], and our knowledge about the clamp loading mechanism is continuously progressing.

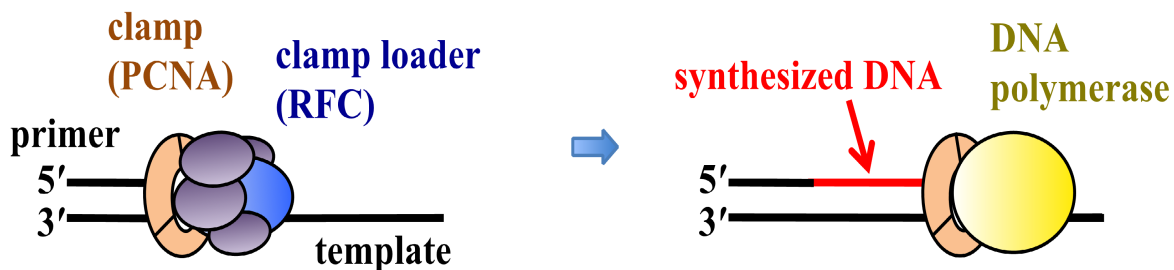


**Figure 4.** PCNA-interacting proteins

After clamp loading, DNA polymerase accesses the clamp and the polymerase-clamp complex performs processive DNA synthesis. Therefore, structural and functional analyses of the DNA polymerase-PCNA complex is the next target to elucidate the overall mechanisms of replication fork progression. The PCNA interacting proteins contain a small conserved sequence motif, called the PIP box, which binds to a common site on PCNA [140]. The PIP box consists of the sequence “Qxxhxxaa”, where “x” represents any amino acid, “h” represents a hydrophobic residue (e.g. L, I or M), and “a” represents an aromatic residue (e.g. F, Y or W). Archaeal DNA polymerases have PIP box-like motifs in their sequences [141]). However, only a few studies have experimentally investigated the function of the motifs. The crystal structure of *P. furiosus* Pol B complexed with a monomeric PCNA mutant was determined,



and a convincing model of the polymerase-PCNA ring interaction was constructed [142]. This study revealed that a novel interaction is formed between a stretched loop of PCNA and the thumb domain of Pol B, in addition to the authentic PIP box. A comparison of the model structure with the previously reported structures of a family B DNA polymerase from RB69 phage, complexed with DNA [143, 144], suggested that the second interaction site plays a crucial role in switching between the polymerase and exonuclease modes, by inducing a PCNA-polymerase complex configuration that favors synthesis over editing. This putative mechanism for the fidelity control of replicative DNA polymerases is supported by experiments, in which mutations at the second interaction site enhanced the exonuclease activity in the presence of PCNA [144]. Furthermore, the three-dimensional structure of the DNA polymerase-PCNA-DNA ternary complex was analyzed by electron microscopic (EM) single particle analysis. This structural view revealed the entire domain configuration of the trimeric ring of PCNA and DNA polymerase, including the protein-protein or protein-DNA contacts. This architecture provides clearer insights into the switching mechanism between the editing and synthesis modes [145].

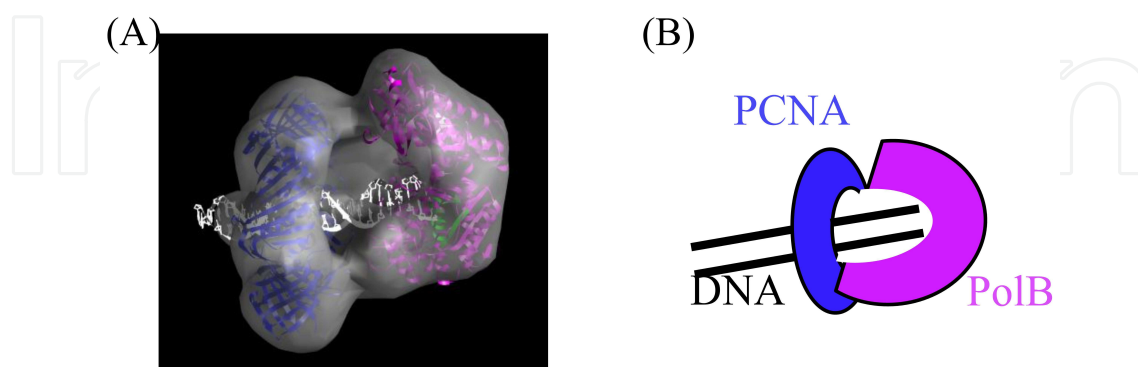


**Figure 5. Mechanisms of processive DNA synthesis** The clamp loader (RFC) tethers the clamp (PCNA) onto the primer terminus of the DNA strand. The clamp loader is then replaced by DNA polymerase, which can synthesize the DNA strand processively without falling off.

In contrast to most euryarchaeal organisms, which have a single PCNA homolog forming a homotrimeric ring structure, the majority of crenarchaea have multiple PCNA homologues, and they are capable of forming heterotrimeric rings for their functions [146, 147]. It is especially interesting that the three PCNAs, PCNA1, PCNA2, and PCNA3, specifically bind PCNA binding proteins, including DNA polymerases, DNA ligases, and FEN-1 endonuclease [147, 148]. Detailed structural studies of the heterologous PCNA from *S. solfataricus* revealed that the interaction modes between the subunits are conserved with those of the homotrimeric PCNAs [149, 150].

*T. kodakarensis* is the only euryarchaeal species that has two genes encoding PCNA homologs on the genome [38]. These two genes from the *T. kodakarensis* genome, and the highly purified gene products, PCNA1 and PCNA2, were characterized [151]. PCNA1 stimulated the DNA synthesis reactions of the two DNA polymerases, Pol B and Pol D, from *T. kodakarensis* *in vitro*. PCNA2 however only had an effect on Pol B. The *T. kodakarensis* strain with *pcna2* disruption was isolated, whereas gene disruption for *pcna1* was not possible. These results suggested that PCNA1 is essential for DNA replication, and PCNA2 may play a different role in *T. kodakarensis* cells. The sensitivities of the  $\Delta pcna2$  mutant strain to ultraviolet

irradiation (UV), methyl methanesulfonate (MMS) and mitomycin C (MMC) were indistinguishable to those of the wild type strain. Both PCNA1 and PCNA2 form a stable ring structure and work as a processivity factor for *T. kodakarensis* Pol B *in vitro*. The crystal structures of the two PCNAs revealed the different interactions at the subunit-subunit interfaces [152].



**Figure 6. Electron Microscopic Analysis of *P. furious* DNA polymerase-PCNA-DNA complex.** The complex in the editing mode of the DNA polymerase-PCNA-DNA ternary complex was shown. (A) Electron microscopic (EM) map of the complex is depicted by gray surface. DNA polymerase and PCNA are shown in a ribbon representation colored purple and blue, respectively. The DNA is colored white. The exonuclease active site is shown in a green ribbon. (B) Schematic view of the complex.

The RFC molecule is conserved as a pentameric complex in Eukaryota and Archaea. However, the eukaryotic RFC is a heteropentameric complex, consisting of five different proteins, RFC1 to 5, in which RFC1 is larger than the other four RFCs. On the other hand, the archaeal RFC consists of two proteins, RFCS (small) and RFCL (large), in a 4 to 1 ratio. A different form of RFC, consisting of three subunits, RFCS1, RFCS2, and RFCL, in a 3 to 1 to 1 ratio, was also identified from *M. acetivorans* [153]. The three subunits of RFC may represent an intermediate stage in the evolution of the more complex RFC in Eukaryota from the less complex RFC in Archaea [153, 154]. The subunit organization and the spatial distribution of the subunits in the *M. acetivorans* RFC complex were analyzed and compared with those of the *E. coli*  $\gamma$ -complex, which is also a pentamer consisting of three different proteins. These two clamp loaders adopt similar subunit organizations and spatial distributions, but the functions of the individual subunits are likely to be diverse [154].

## 11. DNA ligase

DNA ligase is essential to connect the Okazaki fragments of the discontinuous strand synthesis during DNA replication, and therefore, it universally exists in all living organisms. This enzyme catalyzes phosphodiester bond formation via three nucleotidyl transfer steps [155, 156]. In the first step, DNA ligase forms a covalent enzyme-AMP intermediate, by reacting with ATP or NAD<sup>+</sup> as a cofactor. In the second step, DNA ligase recognizes the substrate DNA, and the AMP is subsequently transferred from the ligase to the 5'-phosphate terminus of the DNA, to form a DNA-adenylate intermediate (AppDNA). In the final step,

the 5'-AppDNA is attacked by the adjacent 3'-hydroxy group of the DNA and a phosphodiester bond is formed. DNA ligases are grouped into two families, according to their requirement for ATP or NAD<sup>+</sup> as a nucleotide cofactor in the first step reaction. ATP-dependent DNA ligases are widely found in all three domains of life, whereas NAD<sup>+</sup>-dependent DNA ligases exist mostly in Bacteria. Some halophilic archaea [157] and eukaryotic viruses [158] also have NAD<sup>+</sup>-dependent enzymes.

Three genes (*LIG1*, *LIG3* and *LIG4*) encoding ATP-dependent DNA ligases have been identified in the human genome to date and DNA ligase I (Lig I), encoded by *LIG1*, is a replicative enzyme that joins Okazaki fragments during DNA replication [156]. The first gene encoding a eukaryotic-like ATP-dependent DNA ligase was found in the thermophilic archaeon, *Desulfolobus ambivalens* [159]. Subsequent identifications of the DNA ligases from archaeal organisms revealed that these enzymes primarily use ATP as a cofactor. However, this classification may not be so strict. The utilization of NAD<sup>+</sup>, as well as ATP, as a cofactor has been observed in several DNA ligases, including those from *T. kodakarensis* [160], *T. fumicolans*, *P. abyssi* [161]), *Thermococcus* sp. NA1 [162], *T. acidophilum*, *Picrophilus torridus*, and *Ferroplasma acidophilum*, although ATP is evidently preferable in all of the cases [163] (Table 3). The dual co-factor specificity (ATP/NAD<sup>+</sup>) is an interesting feature of these DNA ligase enzymes and it will be enlightening to investigate the structural basis for this. Another dual co-factor specificity exists in the archaeal DNA ligases, which use ADP as well as ATP, as found in the enzymes from *A. pernix* [164] and *Staphylothermus marinus* [165], and in the case of *Sulfobococcus zilligii*, GTP is also the functional cofactor [166]. The DNA ligases from *P. horikoshii* [167] and *P. furiosus* [168] have a strict ATP preference (Table 3). Sufficient biochemical data have not been obtained to resolve the issue of dual co-factor specificity, and further biochemical and structural analyses are required.

| <b>cofactor</b>                                   |                                |                                  |                               |
|---|--------------------------------|----------------------------------|-------------------------------|
| <b>ATP</b>  | <b>ATP and ADP</b>             | <b>ATP and NAD<sup>+</sup></b>   | <b>ATP, ADP, and GTP</b>      |
| <i>Acidithiobacillus ferrooxidans</i>             | <i>Aeropyrum pernix</i>        | <i>Ferroplasma acidophilum</i>   | <i>Sulfobococcus zilligii</i> |
| <i>Ferroplasma acidarmanus</i>                    | <i>Staphylothermus marinus</i> | <i>Picrophilus torridus</i>      |                               |
| <i>Methanothermobacterium thermoautotrophicum</i> |                                | <i>Thermococcus fumicolans</i>   |                               |
| <i>Pyrococcus horikoshii</i>                      |                                | <i>Thermococcus kodakarensis</i> |                               |
| <i>Pyrococcus furiosus</i>                        |                                | <i>Thermococcus</i> sp.          |                               |
| <i>Sulfolobus acidocaldarius</i>                  |                                | <i>Thermoplasma acidophilum</i>  |                               |
| <i>Sulfolobus shibatae</i>                        |                                |                                  |                               |
| <i>Thermococcus</i> sp. 1519                      |                                |                                  |                               |

**Table 3.** Cofactor dependency of the archaeal DNA ligases

The crystal structure of *P. furiosus* DNA ligase [169] was solved and the physical and functional interactions between the DNA ligase and PCNA was shown [168]. The detailed interaction mode between human Lig I and PCNA is somewhat unclear, because of several controversial reports [170-172]. The stimulatory effect of *P. furiosus* PCNA on the enzyme activity of the cognate DNA ligase was observed at a high salt concentration, at which a DNA ligase alone cannot bind to a nicked DNA substrate. Interestingly, the PCNA-binding site is located in the middle of the N-terminal DNA binding domain (DBD) of the *P. furiosus* DNA ligase, and the binding motif, QKSFF, which is proposed as a shorter version of the PIP box, is actually looped out from the protein surface [168]. Interestingly, this motif is located in the middle of the protein chain, rather than the N- or C-terminal region, where the PIP boxes are usually located. To confirm that this motif is conserved in the archaeal/eukaryotic DNA ligases, the physical and functional interactions between *A. pernix* DNA ligase and PCNA was analyzed and the interaction was shown to mainly depend on the phenylalanine 132 residue, which is located in the predicted region from the multiple sequence alignment of the ATP-dependent DNA ligases [173].

The crystal structure of the human Lig I, complexed with DNA, was solved as the first ATP-dependent mammalian DNA ligase, although the ligase was an N-terminal truncated form [174]. The structure comprises the N-terminal DNA binding domain, the middle adenylation domain, and the C-terminal OB-fold domain. The crystal structure of Lig I (residues 233 to 919) in complex with a nicked, 5'-adenylated DNA intermediate revealed that the enzyme redirects the path of the dsDNA, to expose the nick termini for the strand-joining reaction. The N-terminal DNA-binding domain works to encircle the DNA substrate like PCNA and to stabilize the DNA in a distorted structure, positioning the catalytic core on the nick. The crystal structure of the full length DNA ligase from *P. furiosus* revealed that the architecture of each domain resembles those of Lig I, but the domain arrangements strikingly differ between the two enzymes [168]. This domain rearrangement is probably derived from the "domain-connecting" role of the helical extension conserved at the C-termini in the archaeal and eukaryotic DNA ligases. The DNA substrate in the open form of Lig I is replaced by motif VI at the C-terminus, in the closed form of *P. furiosus* DNA ligase. Both the shapes and electrostatic distributions are similar between motif VI and the DNA substrate, suggesting that motif VI in the closed state mimics the incoming substrate DNA. The subsequently solved crystal structure of *S. solfataricus* DNA ligase is the fully open structure, in which the three domains are highly extended [175]. In this work, the *S. solfataricus* ligase-PCNA complex was also analyzed by SAXS. *S. solfataricus* DNA ligase bound to the PCNA ring still retains an open, extended conformation. The closed, ring-shaped conformation observed in the Lig I structure as described above is probably the active form to catalyze a DNA end-joining reaction, and therefore, it is proposed that the open-to-closed movement occurs for ligation, and the switch in the conformational change is accommodated by a malleable interface with PCNA, which serves as an efficient platform for DNA ligation [175]. After the publication of these crystal structures, the three-dimensional structure of the ternary complex, consisting of DNA ligase-PCNA-DNA, using the *P. furiosus* proteins was obtained by EM single particle analysis [176]. In the complex structure, the three domains of the crescent-shaped *P. furiosus* DNA ligase surround the central DNA duplex, encircled by the closed PCNA ring. The

relative orientations of the ligase domains remarkably differ from those of the crystal structures, and therefore, a large domain rearrangement occurs upon ternary complex formation. In the EM image model, the DNA ligase contacts PCNA at two sites, the conventional PIP box and a novel second contact in the middle adenylation domain. It is also interesting that a substantial DNA tilt from the PCNA ring axis is observed. Based on these structural analyses, a mechanism in which the PCNA binding proteins are bound and released sequentially. In fact, most of the PCNA binding proteins share the same binding sites in the interdomain connecting loop (IDCL) and the C-terminal tail of the PCNA. The structural features exclude the possibility that the three proteins contact the single PCNA ring simultaneously, because DNA ligase occupies two of the three subunits of the PCNA trimer. In the case of the RFC-PCNA-DNA complex structure obtained by the same EM technique, RFC entirely covers the PCNA ring, thus blocking the access of other proteins [138]. These ternary complexes appear to favor a mechanism involving the sequential binding and release of replication factors.

## 12. Flap endonuclease 1 (FEN1)

Efficient processing of Okazaki fragments to make a continuous DNA strand is essential for the lagging strand synthesis in asymmetric DNA replication. The primase-synthesized RNA/DNA primers need to be removed to join the Okazaki fragments into an intact continuous strand DNA. Flap endonuclease 1 (FEN1) is mainly responsible for this task. Okazaki fragment maturation is highly coordinated with continuous DNA synthesis, and the interactions of DNA polymerase, FEN1, and DNA ligase with PCNA allow these enzymes to act sequentially during the maturation process, as described above.

FEN1, a structure-specific 5'-endonuclease, specifically recognizes a dsDNA with an unannealed 5'-flap [177, 178]. In the eukaryotic Okazaki fragment processing system, 5'-flap DNA structures are formed by the strand displacement activity of DNA polymerase  $\delta$ . Lig I seals the nick after the flapped DNA is cleaved by FEN1. These processing steps are facilitated by PCNA [179]. The interactions between eukaryotic FEN1 and PCNA have been well characterized [140, 171], and the stimulatory effect of PCNA on the FEN1 activity was also shown [180]. The crystal structure of the human FEN1-PCNA complex revealed three FEN1 molecules bound to each PCNA subunit of the trimer ring in different configurations [181]. Based on these structural analyses together with the description in the DNA ligase section, a flip-flop transition mechanism, which enables proteins to internally switch for different functions on the same DNA clamp are currently being considered.

The eukaryotic homologs of FEN1 were found in Archaea [182]. The crystal structures of FEN1 from *M. jannaschii* [183], *P. furiosus* [184], *P. horikoshii* [185], *A. fulgidus* [186], and *S. solfataricus* [150] have been determined. In addition, detailed biochemical studies were performed on *P. horikoshii* FEN1 [187, 188]. Thus, studies of the archaeal FEN1 proteins have provided important insights into the structural basis of the cleavage reaction of the flapped DNA. Our recent research showed that the flap endonuclease activity of *P. furiosus* FEN1 was stimulated by PCNA. Furthermore, the stimulatory effect of PCNA on the sequential

action of FEN1 and DNA ligase was observed *in vitro* (Kiyonari et al., unpublished). Based on these results, a model of the molecular switching mechanisms of the last steps of Okazaki-fragment maturation was constructed. The quaternary complex of FEN1-Lig-PCNA-DNA was also isolated for the EM single particle analysis. These studies will provide more concrete image of the molecular mechanism.

### 13. Summary and perspectives

Research on the molecular mechanism of DNA replication has been a central theme of molecular biology. Archaeal organisms became popular in the total genome sequencing age, as described above, and most of the DNA replication proteins are now equally understood by biochemical characterizations. In addition, the archaeal studies are especially interesting to understand the mechanisms by which cells live in extreme environmental conditions. Furthermore, it is also noteworthy that the proteins from the hyperthermophilic archaea are more stable than those from mesophilic organisms, and they are advantageous for the structural and functional analyses of higher-ordered complexes, such as the replisome. Studies on the higher-ordered complexes, rather than single proteins, are essential for understanding each of the events involved in DNA metabolism, and the archaeal research will continuously contribute to the development and advancement of the DNA replication research field, as summarized in part in a recent review [189, 190].

In addition to basic molecular biology research, DNA replication proteins from thermophiles have been quite useful reagents for gene manipulations, including genetic diagnosis, forensic DNA typing, and detection of bacterial and virus infections, as well as basic research. Numerous enzymes have been commercialized around the world, and are utilized daily. An example of the successful engineering of an archaeal DNA polymerase for PCR is the creation of the fusion protein between *P. furiosus* Pol B and a nonspecific dsDNA binding protein, Sso7d, from *S. solfataricus*, by genetic engineering techniques [191]. The fusion DNA polymerase overcame the low processivity of the wild type Pol B by the high affinity Sso7d to the DNA strand. As another example, we successfully developed a novel processive PCR method, using the archaeal Pol B with the help of a mutant PCNA [192, 193]. Several DNA sequencing technologies, referred to as “next-generation sequencing”, have been developed [194, 195], and are now commercially available. Single-molecule detection, using dye-labeled modified nucleotides and longer read lengths, is now known as “third-generation DNA sequencing” [196]. These technologies apply DNA polymerases or DNA ligases from various sources, indicating that these DNA replication enzymes are indispensable for the development of DNA manipulation technology. These facts prove that the progress of the basic research on the molecular biology of archaeal DNA replication will promote the development of the new technologies for genetic engineering.

## Author details

Yoshizumi Ishino and Sonoko Ishino

Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Japan

## References

- [1] Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 1953;171(4356) 737-738.
- [2] Kelman Z, White MF. Archaeal DNA replication and repair. *Curr Opin Microbiol*, 2005;(6) 669-676.
- [3] Barry ER, Bell SD. DNA replication in the archaea. *Microbiol Mol Biol Rev*, 2006;70(4) 876-887.
- [4] Wigley DB ORC proteins: marking the start. *Curr Opin Struct Biol*, 2009;19(1) 72-78.
- [5] Leipe DD, Aravind L, Koonin EV. Did DNA replication evolve twice independently? *Nucleic Acids Res*, 1999;27(17) 3389-3401.
- [6] Jacob F, Brenner SCR. *Hebd Seances Acad Sci*. 1963;256 298-300.
- [7] Matsunaga F, Forterre P, Ishino Y, Myllykallio H. In vivo interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin. *Proc. Natl. Acad. Sci. USA*, 2001;98(20) 11152-11157.
- [8] Uemori T, Sato Y, Kato I, Doi H, Ishino Y. A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization. *Genes Cells*, 1997;2(8) 499-512.
- [9] Lopez P, Philippe H, Myllykallio H, Forterre P. Identification of putative chromosomal origins of replication in Archaea. *Mol Microbiol*, 1999;32(4) 883-886.
- [10] Matsunaga F, Norais C, Forterre P, Myllykallio H. Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Rep*, 2003;4(2) 154-158.
- [11] Robinson NP, Dionne I, Lundgren M, Marsh VL, Bernander R, Bell SD. Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus* Cell, 2004;116(1) 25-38.
- [12] Matsunaga F, Glatigny A, Mucchielli-Giorgi MH, Agier N, Delacroix H, Marisa L, et al. Genomewide and Biochemical Analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in *Pyrococcus* sp. *Nucleic Acids Res*, 2007;35(10)3214-3222.

- [13] Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R. Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci USA*, 2004;101(18):7046-7051.
- [14] Robinson NP, Bell SD. Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes. *Proc Natl Acad Sci USA*, 2007;104(14) 5806-5811.
- [15] Norais C, Hawkins M, Hartman AL, Eisen JA, Myllykallio H, Allers T. Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet*, 2007;3(5) e77.
- [16] Capes MD, Coker JA, Gessler R, Grinblat-Huse V, DasSarma SL, Jacob CG, et al. The information transfer system of halophilic archaea. *Plasmid*, 2011;65(2) 77-101.
- [17] Liu J, Smith CL, DeRyckere D, DeAngelis K, Martin GS, Berger JM. Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol Cell*, 2000;6(3) 637-648.
- [18] Singleton MR, Morales R, Grainge I, Cook N, Isupov MN, Wigley DB. Conformational changes induced by nucleotide binding in Cdc6/ORC from *Aeropyrum pernix*. *J Mol Biol*, 2004;343(3) 547-557.
- [19] Grainge I, Gaudier M, Schuwirth BS, Westcott SL, Sandall J, Atanassova N, et al. Biochemical analysis of a DNA replication origin in the archaeon *Aeropyrum pernix*. *J Mol Biol*, 2006;363(2) 355-369.
- [20] Dueber EL, Corn JE, Bell, SD, Berger JM. Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science*, 2007;317(5842) 1210-1213.
- [21] Gaudier M, Schuwirth BS, Westcott SL, Wigley DB. Structural basis of DNA replication origin recognition by an ORC protein. *Science*, 2007;317(5842) 1213-1216.
- [22] Majernik AI, Chong JP. A conserved mechanism for replication origin recognition and binding in archaea. *Biochem J*, 2008;409(2) 511-518.
- [23] Matsunaga F, Takemura K, Akita M, Adachi A, Yamagami T, Ishino Y. Localized melting of duplex DNA by Cdc6/Orc1 at the DNA replication origin in the hyperthermophilic archaeon *Pyrococcus furiosus*. *Extremophiles*, 2010;14(1) 21-31.
- [24] Akita M, Adachi A, Takemura K, Yamagami T, Matsunaga F, Ishino Y. Cdc6/Orc1 from *Pyrococcus furiosus* may act as the origin recognition protein and Mcm helicase recruiter. *Gens Cells*, 2010;15(5) 537-552.
- [25] Masai H, You Z, Arai K. Control of DNA replication: regulation and activation of eukaryotic replicative helicase, MCM. *IUBMB Life*, 2005;57(4-5) 323-335.
- [26] Ilves I, Petojevic T, Pesavento JJ, Botchan MR. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell*, 2010;37(2) 247-258,



- [27] Chong JP, Hayashi MK, Simon MN, Xu RM, Stillman B. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci USA*, 2000;97(4) 1530-1535.
- [28] Grainge I, Scaife S, Wigley DB. Biochemical analysis of components of the pre-replication complex of *Archaeoglobus fulgidus*. *Nucleic Acids Res*, 2003;31(16) 4888-4898.
- [29] Kelman Z, Lee JK, Hurwitz J. The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum*  $\Delta$ H contains DNA helicase activity. *Proc Natl Acad Sci USA*, 1999;96(26) 14783-14788.
- [30] Shechter DF, Ying CY, Gautier J. The intrinsic DNA helicase activity of *Methanobacterium thermoautotrophicum*  $\Delta$ H minichromosome maintenance protein. *J Biol Chem*, 2000;275(20) 15049-15059.
- [31] Yoshimochi T, Fujikane R, Kawanami M, Matsunaga F, Ishino Y. The GINS complex from *Pyrococcus furiosus* stimulates the MCM helicase activity. *J Biol Chem*, 2008;283(3) 1601-1609.
- [32] Sakakibara N, Kelman LM, Kelman Z. Unwinding the structure and function of the archaeal MCM helicase. *Mol Microbiol*, 2009;72(2) 286-296.
- [33] Shin JH, Grabowski B, Kasiviswanathan R, Bell SD, Kelman Z. Regulation of minichromosome maintenance helicase activity by Cdc6. *J Biol Chem*, 2003;278(39) 38059-38067.
- [34] Kasiviswanathan R, Shin JH, Kelman Z. Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties. *Nucleic Acids Res*, 2005;33(15) 4940-4950.
- [35] De Felice M, Esposito L, Pucci B, Carpentieri F, De Falco M, Rossi M, et al. Biochemical characterization of a CDC6-like protein from the crenarchaeon *Sulfolobus solfataricus*. *J Biol Chem*, 2003;278(47) 46424-46431.
- [36] Haugland GT, Shin JH, Birkeland NK, Kelman Z. Stimulation of MCM helicase activity by a Cdc6 protein in the archaeon *Thermoplasma acidophilum*. *Nucleic Acids Res*, 2006, 34(21) 6337-6344.
- [37] Krupovic M, Gribaldo S, Bamford DH, Forterre P. The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. *Mol Biol Evol*, 2010;27(12) 2716-2732.
- [38] Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S, Imanaka T. Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res*, 2005;15(3) 352-363.
- [39] Sato T, Fukui T, Atomi H, Imanaka T. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J Bacteriol*, 2003;185(1) 210-220.

- [40] Sato T, Fukui T, Atomi H, Imanaka T. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *Appl Environ Microbiol*, 2005;71(7) 3889-3899.
- [41] Pan M, Santangelo TJ, Li Z, Reeve JN, Kelman Z. *Thermococcus kodakarensis* encodes three MCM homologs but only one is essential. *Nucleic Acids Res*, 2011;39(22) 9671-9680.
- [42] Ishino S, Fujino S, Tomita H, Ogino H, Takao K, Daiyasu H, et al. Biochemical and genetical analyses of the three mcm genes from the hyperthermophilic archaeon, *Thermococcus kodakarensis*. *Genes Cells*, 2011;16(12) 1176-1189.
- [43] Fujikane R, Ishino S, Ishino Y, Forterre, P. Genetic analysis of DNA repair in the hyperthermophilic archaeon, *Thermococcus kodakarensis*. *Genes Genet Syst*, 2010; 85(4)243-257.
- [44] Xia Q, Hendrickson E L, Zhang Y, Wang T, Taub F, Moore BC, et al. Quantitative proteomics of the archaeon *Methanococcus maripaludis* validated by microarray analysis and real time PCR. *Mol Cell Proteomics*, 2006;5(5) 868-881.
- [45] Walters AD, Chong JP. An archaeal order with multiple minichromosome maintenance genes. *Microbiology*, 2010;156(Pt 5):1405-1414.
- [46] Sakakibara N, Kelman LM, Kelman Z. How is the archaeal MCM helicase assembled at the origin? Possible mechanisms *Biochem. Soc. Trans.* 2007;37(Pt 1):7-11.
- [47] Shin J-H, Heo GY, Kelman Z. The *Methanothermobacter thermautotrophicus* Cdc6-2 protein, the putative helicase loader, dissociates the minichromosome maintenance helicase. *J Bacteriol*, 2008;190(11) 4091-4094.
- [48] Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H. et al. GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev*, 2003;17(9)1153-1165.
- [49] Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Res*, 2005;33(14) 4626-4638.
- [50] Marinsek N, Barry ER, Makarova KS, Dionne I, Koonin EV, Bell SD. GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep*, 2006;7(5) 539-545.
- [51] Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC. Localization of MCM2-7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol Cell*, 2006;21(4) 581-587.
- [52] Ogino H, Ishino S, Mayanagi K, Haugland GT, Birkeland NK, Yamagishi A, et al. The GINS complex from the thermophilic archaeon, *Thermoplasma acidophilum* may function as a homotetramer in DNA replication. *Extremophiles*, 2011;15(4) 529-539.

- [53] Oyama T, Ishino S, Fujino S, Ogino H, Shirai T, Mayanagi K, et al. Architectures of archaeal GINS complexes, essential DNA replication initiation factors. *BMC Biol*, 2011;9 28.
- [54] Kamada K, Kubota Y, Arata T, Shindo Y, Hanaoka F. Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nat Struct Mol Biol*, 2007;14(5)388-396.
- [55] Choi JM, Lim HS, Kim JJ, Song OK, Cho Y. Crystal structure of the human GINS complex. *Genes Dev*, 2007;21(11) 1316-1321.
- [56] Chang, YP, Wang G, BermudezV, Hurwitz J, Chen XS, et al. Crystal structure of the GINS complex and functional insight into its role in DNA replication. *Proc Natl Acad Sci USA*, 2007;104(31) 12685-12690.
- [57] Sanchez-Pulido L, Ponting CP. Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics*, 2011;27(14) 1885-1888.
- [58] Li Z, Pan M, Santangelo T J, Chemnitz W, Yuan W, Edwards JL, et al. A novel DNA nuclease is stimulated by association with the GINS complex. *Nucleic Acids Res*, 2011;39(14) 6114-6123.
- [59] Krastanova I, Sannino V, Amenitsch H, Gileadi O, Pisani FM, Onesti S., et al. Structural and functional insights into the DNA replication factor Cdc45 reveal an evolutionary relationship to the DHH family of phosphoesterases. *J Biol Chem*, 2011;287(6) 4121-4128.
- [60] Kirk BW, Kuchta RD. Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. *Biochemistry*, 1999;38(31) 7727-7736
- [61] Desogus G, Onesti S, Brick P, Rossi M, Pisani FM. Identification and characterization of a DNA primase from the hyperthermophilic archaeon *Methanococcus jannaschii*. *Nucleic Acids Res*, 1999;27(22) 4444-4450.
- [62] Bocquier A, Liu L, Cann I, Komori K, Kohda D, Ishino Y. Archaeal primase: bridging the gap between RNA and DNA polymerase. *Curr Biol*, 2001;11(6) 452-456.
- [63] Liu L, Komori K, Ishino S, Bocquier AA, Cann IK, Kohda D, et al. The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J Biol Chem*, 2001;276(48) 45484-45490.
- [64] Lao-Sirieix SH, Bell SD. The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J Mol Biol*, 2004;344(5) 1251-1263.
- [65] Lao-Sirieix SH, Nookala RK, Roversi P, Bell SD, Pellegrini L. et al. Structure of the heterodimeric core primase. *Nat Struct Mol Biol*, 2005;12(12) 1137-1144.

- [66] De Falco M, Fusco A, DeFelice M, Rossi M, Pisani FM. The DNA primase is activated by substrates containing a thymine-rich bubble and has a 3'-terminal nucleotidyl-transferase activity. *Nucleic Acids Res*, 2004;32(17) 5223-5230.
- [67] Ito N, Nureki O, Shirouzu M, Yokoyama S, Hanaoka F. Crystal structure of the *Pyrococcus horikoshii* DNA primase-UTP complex: implications for the mechanism of primer synthesis. *Genes Cells*, 2003;8(12) 913-923.
- [68] Matsui E, Nishio M, Yokoyama H, Harata K, Damas S, Matsui I. Distinct domain functions regulating de novo DNA synthesis of thermostable DNA primase from hyperthermophile *Pyrococcus horikoshii*. *Biochemistry*. 2003;42(50) 14968-14976.
- [69] Ito N, Matsui I, Matsui E. Molecular basis for the subunit assembly of the primase from an archaeon *Pyrococcus horikoshii*. *FEBS J*, 2007;274(5):1340-1351.
- [70] Le Breton M, Henneke G, Norais C, Flament D, Myllykallio H, Querellou J, et al. The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair? *J Mol Biol*, 2007;374(5) 1172-1185.
- [71] Hu J, Guo L, Wu K, Lang S, Huang L. Template-dependent polymerization across discontinuous templates by the heterodimeric primase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res*, 2011;40(8) 3470-3483.
- [72] Evguenieva-Hackenberg E, Walter P, Hochleitner E, Lottspeich F, Klug G. An exosome-like complex in *Sulfolobus solfataricus*. *EMBO Rep*, 2003;4(9) 889-934.
- [73] Walter P, Klein F, Lorentzen E, Ilchmann A, Klug G, Evguenieva-Hackenberg E. et al. Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Microbiol*, 2006;62(4) 1076-1089.
- [74] Zuo Z, Rodgers CJ, Mikheikin AL, Trakselis MA. Characterization of a functional DnaG-type primase in archaea: implications for a dual-primase system. *J Mol Biol*, 2010;397(3) 664-676.
- [75] Wu K, Lai X, Guo X, Hu J, Xiang X, Huang L. Interplay between primase and replication factor C in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Microbiol*, 2007;63(3) 826-837.
- [76] Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J*, 1993;12(3) 861-867.
- [77] Braun KA, Lao Y, He Z, Ingles CJ, Wold MS. Wold MS Role of protein-protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry*, 1997;36(28) 8443-8454.
- [78] Lin YL, Chen C, Keshav KF, Winchester E, Dutta A. Dissection of functional domains of the human DNA replication protein complex replication protein A. *J Biol Chem*, 1996;271(29) 17190-17198.

- [79] Bochkarev A, Bochkareva E, Frappier L, Edwards AM. The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J*, 1999;18(16) 4498-4504.
- [80] Bochkareva E, Frappier L, Edwards A M, Bochkarev A. The RPA32 subunit of human replication protein A contains a single-stranded DNA-binding domain. *J Biol Chem*, 1998;273(7) 3932-3936.
- [81] Philipova D, Mullen JR, Maniar HS, Lu J, Gu C, Brill SJ, et al. A hierarchy of SSB promoters in replication protein A. *Genes Dev*, 1996;10(17) 2222-2233.
- [82] Gomes XV, Wold MS. Structural analysis of human replication protein A. Mapping functional domains of the 70-kDa subunit. *J Biol Chem*, 1995;270(9) 4534-4543.
- [83] Mer G, Bochkarev A, Gupta R, Bochkareva E, Frappier L, Ingles CJ. Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell*, 2000;103(3) 449-456.
- [84] Dornreiter I, Erdile LF, Gilbert IU, von Winkler D, Kelly TJ, Fanning E. Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J*, 1992;11(2) 769-776.
- [85] Tsurimoto T, Stillman B. Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. *EMBO J*, 1989;8(12) 3883-3889.
- [86] Kenny MK, Lee SH, Hurwitz J. Multiple functions of human single-stranded-DNA binding protein in simian virus 40 DNA replication: single-strand stabilization and stimulation of DNA polymerases alpha and delta. *Proc Natl Acad Sci USA*, 1989;86(24) 9757-9761.
- [87] Chedin F, Seitz EM, Kowalczykowski SC. Novel homologs of replication protein A in archaea: implications for the evolution of ssDNA-binding proteins. *Trends Biol Sci*, 1998;23(8) 273-277.
- [88] Kelly TJ, Simancek P, Brush GS. Identification and characterization of a single-stranded DNA-binding protein from the archaeon *Methanococcus jannaschii*. *Proc Natl Acad Sci USA*, 1998;95(25) 14634-14639.
- [89] Kelman Z, Pietrokovski S, Hurwitz J. Isolation and characterization of a split B-type DNA polymerase from the archaeon *Methanobacterium thermoautotrophicum* ΔH. *J Biol Chem*, 1999;274(40) 28751-28761.
- [90] Komori K, Ishino Y. Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination. *J Biol Chem*, 2001;276(28) 25654-25660.
- [91] Kerr ID, Wadsworth RI, Cubeddu L, Blankenfeldt W, Naismith JH, White MF. Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein. *EMBO J*, 2003;22(11) 2561-270.
- [92] Robbins JB, Murphy MC, White BA, Mackie RI, Ha T, Cann IK, et al. Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivor-*

- ans* and their effects on DNA synthesis by DNA polymerase  $\beta$ . *J Biol Chem*, 2004;279(8) 6315-6326.
- [93] Lin Y, Lin LJ, Sriratana P, Coleman K, Ha T, Spies M, et al. Engineering of functional replication protein a homologs based on insights into the evolution of oligonucleotide/oligosaccharide-binding folds. *J Bacteriol*, 2008;190(17) 5766-5780.
- [94] Braithwaite DK, Ito J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucleic Acids Res*, 1993, 1993;21(4) 787-802.
- [95] Cann I, Ishino Y. Archaeal DNA replication: Identifying the pieces to solve a puzzle. *Genetics*, 1999;152(4) 1249-1267.
- [96] Ishino, Y., Cann, I. The euryarchaeotes, a subdomain of archaea, survive on a single DNA polymerase: fact or farce? *Genes Genet Syst*, 1998;73(6) 323-336.
- [97] Lipps G, Röther S, Hart C, Krauss G. A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *EMBO J*, 2003;22(10) 2516-2525.
- [98] Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, Hinkle D, et al. The Y-family of DNA polymerases. *Mol Cell*;8(1) 7-8.
- [99] Guo C, Kosarek-Stancel JN, Tang TS, Friedberg EC. Y-family DNA polymerases in mammalian cells. *Cell Mol Life Sci*, 2009;66(14) 2363-2381.
- [100] Pisani FM, De Martino C, Rossi M. A DNA polymerase from the archaeon *Sulfolobus solfataricus* shows sequence similarity to family B DNA polymerases. *Nucleic Acids Res*, 1992;20(11) 2711-2716.
- [101] Perler, FB, Comb, DG, Jack, WE, Moran, LS, Qiang, B, Kucera, RB, et al. Intervening sequences in an Archaea DNA polymerase gene. *Proc. Natl. Acad. Sci. USA* 1992;89(12) 5577-5581.
- [102] Uemori T, Ishino Y, Toh H, Asada K, Kato I. Organization and nucleotide sequence of the DNA polymerase gene from the archaeon *Pyrococcus furiosus*. *Nucleic Acids Res*, 1992;21(2) 259-265.
- [103] Uemori T, Ishino Y, Doi H, Kato I. The hyperthermophilic archaeon *Pyrodictium occultum* has two alpha-like DNA polymerases. *J Bacteriol*, 1995;177(8) 2164-2177.
- [104] Cann I, Ishino S, Nomura N, Sako Y, Ishino Y. Two family B DNA polymerases in *Aeropyrum pernix*, an obligate aerobic hyperthermophilic crenarchaeote. *J Bacteriol*, 1999;181(19) 5984-5992.
- [105] Cann I, Komori K, Toh H, Kanai S, Ishino Y. A heterodimeric DNA polymerase: evidence that members of the euryarchaeota possess a novel DNA polymerase. *Proc. Natl. Acad. Sci. USA*, 1998;95(24) 14250-14255.
- [106] Ishino S, Ishino Y. Comprehensive search for DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*. *Nucleosides, Nucleotides, and Nucleic Acids*. 2006;25(4-6):681-691.

- [107] Imamura M, Uemori T, Kato I, Ishino Y. A non- $\alpha$ -like DNA polymerase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biol Pharm Bull*, 1995;18(12) 1647-1652.
- [108] Ishino Y, Ishino S. Novel DNA polymerases from Euryarchaeota. *Meth Enzymol*, 2001, 334 249-260
- [109] Gueguen Y, Rolland JL, Lecompte O, Azam P, Le Romancer G, Flament D, et al. Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon *Pyrococcus abyssi*. *Eur J Biochem*, 2001;268(22) 5961-5969.
- [110] Shen Y, Musti K, Hiramoto M, Kikuchi H, Kawarabayashi Y, Matsui I. Invariant Asp-1122 and Asp-1124 are essential residues for polymerization catalysis of family D DNA polymerase from *Pyrococcus horikoshii*. *J Biol Chem*, 2001;276(29) 27376-27383.
- [111] Shen Y, Tang X, Matsui I. Subunit interaction and regulation of activity through terminal domains of the family D DNA polymerase from *Pyrococcus horikoshii*. *J Biol Chem*, 2003;278(23) 21247-21257.
- [112] Tang XF, Shen Y, Matsui E, Matsui I. Domain topology of the DNA polymerase D complex from a hyperthermophilic archaeon *Pyrococcus horikoshii*. *Biochemistry*, 2004;43(37) 11818-11827.
- [113] Jokela M, Eskelinen A, Pospiech H, Rouvinen J, Syväoja JE. Characterization of the 3' exonuclease subunit DP1 of *Methanococcus jannaschii* replicative DNA polymerase D. *Nucleic Acids Res*, 2004;32(8)2430-2440.
- [114] Henneke G, Flament D, Hübscher U, Querellou J, Raffin JP. The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication. *J Mol Biol*, 2005;350(1) 53-64.
- [115] Castrec B, Laurent S, Henneke G, Flament D, Raffin JP. The glycine-rich motif of *Pyrococcus abyssi* DNA polymerase D is critical for protein stability. *J Mol Biol*, 2010;396(4) 840-848.
- [116] Berquist BR, DasSarma P, DasSarma S. Essential and non-essential DNA replication genes in the model halophilic Archaeon, *Halobacterium* sp. NRC-1. *BMC Genet*, 2007;8 31.
- [117] Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol*, 2008;6(3) 245-252.
- [118] Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, et al. Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res*, 2011;39(8) 3204-3223.
- [119] Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L, et al. A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci USA*, 2008;105(23) 8102-8107.

- [120] Tahirov TH, Makarova KS, Rogozin IB, Pavlov YI, Koonin EV. 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*, 2009;4 11.
- [121] Zillig W, Kletzin A, Schleper C, et al. Screening for *Sulfolobales*, their plasmids and their viruses in Icelandic solfataras. *Syst Appl Microbiol*, 1993, 16: 609–628
- [122] Lipps G, Röther S, Hart C, Krauss G. A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *EMBO J*, 2003;22(10) 2516-2525.
- [123] Lipps G. Molecular biology of the pRN1 plasmid from *Sulfolobus islandicus*. *Biochem Soc Trans*, 2009;37(Pt 1) 42-45.
- [124] Lipps G. The replication protein of the *Sulfolobus islandicus* plasmid pRN1. *Biochem Soc Trans*, 2004;32(Pt 2) 240-244.
- [125] Soler N, Marguet E, Cortez D, Desnoues N, Keller J, van Tilbeurgh H. Two novel families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. *Nucleic Acids Res*, 2010;38(15) 5088-5104.
- [126] Moldovan G-L, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. *Cell*, 2007;129(4)665-679.
- [127] Pan M, Kelman L, Kelman Z. The archaeal PCNA proteins. *Biochem Soc Trans*, 2011;39(1) 20-24.
- [128] Cann I, Ishino S, Hayashi I, Komori K, Toh H, Morikawa K, et al. Functional interactions of a homolog of proliferating cell nuclear antigen with DNA polymerases in Archaea. *J. Bacteriol*, 1999, Nov;181(21):6591-6599.
- [129] Ishino Y, Tsurimoto T, Ishino S, Cann IK. Functional interactions of an archaeal sliding clamp with mammalian clamp loader and DNA polymerase  $\delta$ . *Genes Cells*, 2001;6(8) 699-706.
- [130] Matsumiya S, Ishino S, Ishino Y, Morikawa K. Physical interaction between proliferating cell nuclear antigen and replication factor C from *Pyrococcus furiosus*. *Genes Cells*, 2002;7(9) 911-922.
- [131] Matsumiya S, Ishino S, Ishino Y, Morikaw K. Intermolecular ion pairs maintain toroidal structure of *Pyrococcus furiosus* PCNA. *Prot Sci*, 2003;12(4) 823-831.
- [132] Tori K, Kimizu M, Ishino S, Ishino Y. Both DNA polymerase BI and D from the hyperthermophilic archaeon, *Pyrococcus furiosus* bind to PCNA at the C-terminal PIP box motifs. *J Bacteriol*, 2007;189(15) 5652-5657.
- [133] Cann I, Ishino S, Yuasa M, Daiyasu H, Toh H, Ishino Y. Biochemical analysis of the replication factor C from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol*, 2001;183(8) 2614-2623.



- [134] Mayanagi K, Miyata T, Oyama T, Ishino Y, Morikawa K. Three-dimensional electron microscopy of clamp loader small subunit from *Pyrococcus furiosus*. *J Struct Biol*, 2000;134(1) 35-45.
- [135] Ishino S, Oyama T, Yuasa M, Morikawa K, Ishino Y. Mutational Analysis of *Pyrococcus furiosus* Replication Factor C based on the Three-Dimensional Structure. *Extremophiles*, 2003;7(3) 169-175.
- [136] Oyama T, Ishino Y, Cann I, Ishino S, Morikawa K. Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus*. *Mol Cell*, 2001;8(2) 455-463.
- [137] Indiani C, O'Donnell M. The replication clamp-loading machine at work in the three domains of life. *Nat Rev Mol Cell Biol*, 2006;7(10) 751-761.
- [138] Miyata T, Suzuki H, Oyama T, Mayanagi K, Ishino Y, Morikawa K. Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. *Proc Natl Acad Sci USA*, 2005;102(39) 13795-13800.
- [139] Kelch BA, Makino DL, O'Donnell M, Kuriyan J. The replication clamp-loading machine at work in the three domains of life. *Science*, 2011;334(6063) 1675-1680.
- [140] Warbrick EM. The puzzle of PCNA's many partners. *BioEssays*, 2000;22(11) 997-1006.
- [141] Vivona JB, Kelman Z. The diverse spectrum of sliding clamp interacting proteins. *FEBS Lett*, 2003;546(2-3) 167-172.
- [142] Nishida H, Mayanagi K, Kiyonari S, Sato Y, Oyama T, Ishino Y, et al. Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex. *Proc Natl Acad Sci USA*, 2009;106(49) 20693-20698.
- [143] Shamoo Y, Steitz TA. Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell*, 1999;99(2) 155-166.
- [144] Franklin MC, Wang J, Steitz TA. Structure of the replicating complex of a pol  $\alpha$  family DNA polymerase. *Cell*, 2001;105(5) 657-667.
- [145] Mayanagi K, Kiyonari S, Nishida H, Saito M, Kohda D, Ishino Y, et al. Architecture of the DNA polymerase B-proliferating cell nuclear antigen (PCNA)-DNA ternary complex. *Proc Natl Acad Sci USA*, 2011;108(5):1845-1849
- [146] Daimon K, Kawarabayasi Y, Kikuchi H, Sako Y, Ishino Y. Three proliferating cell nuclear antigen-like proteins found in the hyperthermophilic archaeon *Aeropyrum pernix*: interactions with the two DNA polymerases. *J Bacteriol*, 2002;184(3) 687-694.
- [147] Dionne I, Nookala RK, Jackson SP, Doherty AJ, Bell SD. A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Cell*, 2003;11(1) 275-282.

- [148] Imamura K, Fukunaga K, Kawarabayasi Y, Ishino Y. Specific interactions of three proliferating cell nuclear antigens with replication-related proteins in *Aeropyrum pernix*. *Mol Microbiol*, 2007;64(2) 308-318.
- [149] Williams GJ, Johnson K, Rudolf J, McMahon SA, et al. Structure of the heterotrimeric PCNA from *Sulfolobus solfataricus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 2006;62(Pt 10) 944-948.
- [150] Doré AS, Kilkenny ML, Jones SA, Oliver AW, Roe SM, Bell SD, et al. Structure of an archaeal PCNA1-PCNA2-FEN1 complex: elucidating PCNA subunit and client enzyme specificity. *Nucleic Acids Res*, 2006;34(16) 4515-4526.
- [151] Kuba Y., Ishino, S., Yamagami, T., Tokuhara, M., Kanai, T., Fujikane, R., et al. Comparative analyses of the two PCNAs from the hyperthermophilic archaeon, *Thermococcus kodakarensis*. *Genes Cells*, 2012, 17, in press
- [152] Ladner JE, Pan M, Hurwitz J, Kelman Z. Crystal structures of two active proliferating cell nuclear antigens (PCNAs) encoded by *Thermococcus kodakaraensis*. *Proc Natl Acad Sci USA*, 2011;108(7) 2711-2716.
- [153] Chen YH, Kocherginskaya SA, Lin Y, Sriratana B, Lagunas AM, Robbins JB, et al. Biochemical and mutational analyses of a unique clamp loader complex in the archaeon *Methanosarcina acetivorans*. *J Biol Chem*, 2005;280(51) 41852-41863.
- [154] Chen YH, Lin Y, Yoshinaga A, Chhotani B, Lorenzini JL, Crofts AA, et al. Molecular analyses of a three-subunit euryarchaeal clamp loader complex from *Methanosarcina acetivorans*. *J Bacteriol*, 2009;191(21) 6539-6549.
- [155] Wilkinson A, Day J, Bowater R. Bacterial DNA ligases. *Mol Microbiol*, 2001;40(6) 1241-1248.
- [156] Tomkinson AE, Vijayakumar S, Pascal JM, Ellenberger T. DNA ligases: structure, reaction mechanism, and function. *Chem Rev*, 2006;106(2) 687-699.
- [157] Poidevin L, MacNeill SA. Biochemical characterisation of LigN, an NAD<sup>+</sup>-dependent DNA ligase from the halophilic euryarchaeon *Haloferax volcanii* that displays maximal in vitro activity at high salt concentrations. *BMC Mol Biol*, 2006;7 44.
- [158] Benarroch D, Shuman S. Characterization of mimivirus NAD<sup>+</sup>-dependent DNA ligase. *Virology*, 2006;353(1) 133-143.
- [159] Kletzin A. Molecular characterisation of a DNA ligase gene of the extremely thermophilic archaeon *Desulfurolobus ambivalens* shows close phylogenetic relationship to eukaryotic ligases. *Nucleic Acids Res*, 1992;20(20) 5389-5396.
- [160] Nakatani M, Ezaki S, Atomi H, Imanaka T. A DNA ligase from a hyperthermophilic archaeon with unique cofactor specificity. *J Bacteriol*, 2000;182(22) 6424-6433.
- [161] Rolland JL, Gueguen Y, Persillon C, Masson JM, Dietrich J. Characterization of a thermophilic DNA ligase from the archaeon *Thermococcus fumicolans*. *FEMS Microbiol Lett*, 2004;236(2)267-273.

- [162] Kim YJ, Lee HS, Bae SS., Jeon JH, Yang SH, Lim JK, et al. Cloning, expression, and characterization of a DNA ligase from a hyperthermophilic archaeon *Thermococcus* sp. *Biotechnol Lett*, 2006;28(6) 401-407.
- [163] Ferrer M, Golyshina OV, Belouqui A, Böttger LH, Andreu JM, Polaina J, et al. A purple acidophilic di-ferric DNA ligase from *Ferroplasma*. *Proc Natl Acad Sci USA*, 2008;105(26) 8878-8883.
- [164] Jeon SJ, Ishikawa K. A novel ADP-dependent DNA ligase from *Aeropyrum pernix* K1. *FEBS Lett*, 2003;550(1-3) 69-73.
- [165] Seo MS, Kim YJ, Choi JJ, Lee MS, Kim JH, Lee JH, et al. Cloning and expression of a DNA ligase from the hyperthermophilic archaeon *Staphylothermus marinus* and properties of the enzyme. *J Biotechnol*, 2007;128(3) 519-530.
- [166] Sun Y, Seo MS, Kim JH, Kim YJ, Kim GA, Lee JI, et al. Novel DNA ligase with broad nucleotide cofactor specificity from the hyperthermophilic crenarchaeon *Sulfolobococcus zilligii*: influence of ancestral DNA ligase on cofactor utilization. *Environ Microbiol*, 2008;10(12) 3212-3224.
- [167] Keppetipola N, Shuman S. Characterization of a thermophilic ATP-dependent DNA ligase from the euryarchaeon *Pyrococcus horikoshii*. *J Bacteriol*, 2005;187(20) 6902-6908.
- [168] Kiyonari S, Takayama K, Nishida H, Ishino Y. Identification of a novel binding motif in *Pyrococcus furiosus* DNA ligase for the functional interaction with proliferating cell nuclear antigen. *J Biol Chem*, 2006;281(38) 28023-28032.
- [169] Nishida H, Kiyonari S, Ishino Y, Morikawa K. The closed structure of an archaeal DNA ligase from *Pyrococcus furiosus*. *J Mol Biol*, 2006;360(5) 956-967.
- [170] Levin DS, Bai W, Yao N. An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. *Proc Natl Acad Sci USA*, 1997;94(24) 12863-12868.
- [171] Jónsson ZO, Hindges R, Hübscher U. Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *EMBO J*, 1998;17(8) 2412-2425.
- [172] Tom S, Henricksen LA, Park MS, Bambara RA. DNA ligase I and proliferating cell nuclear antigen form a functional complex. *J Biol Chem*, 2001;276(27) 24817-24825.
- [173] Kiyonari S, Kamigochi T, Ishino Y. A single amino acid substitution in the DNA-binding domain of *Aeropyrum pernix* DNA ligase impairs its interaction with proliferating cell nuclear antigen. *Extremophiles*, 2007;11(5) 675-684.
- [174] Pascal JM, O'Brien PJ, Tomkinson AE, Ellenberger T. Human DNA ligase I completely encircles and partially unwinds nicked DNA. *Nature*, 2004;432(7016) 473-478.
- [175] Pascal JM, Tsodikov OV, Hura GL, Song W, Cotner EA, Classen S, et al. A flexible interface between DNA ligase and PCNA supports conformational switching and efficient ligation of DNA. *Mol Cell*, 2006;24(2) 279-291.

- [176] Mayanagi K, Kiyonari S, Saito M, Shirai T, Ishino Y, Morikawa K. Mechanism of replication machinery assembly as revealed by the DNA ligase-PCNA-DNA complex architecture. *Proc Natl Acad Sci USA*, 2009;106(12) 4647-4652.
- [177] Liu Y, Kao HI, Bambara RA. Flap endonuclease 1: a central component of DNA metabolism. *Annu Rev Biochem*, 2004, 2004;73 589-615.
- [178] Zheng L, Jia J, Finger LD, Guo Z, Zer C, Shen B. Functional regulation of FEN1 nuclease and its link to cancer. *Nucleic Acids Res*, 2011;39(3) 781-794.
- [179] Rossi ML, Bambara RA. Reconstituted Okazaki fragment processing indicates two pathways of primer removal. *J Biol Chem*, 2006;281(36):26051-26061.
- [180] Tom S, Henricksen LA, Bambara RA. Mechanism whereby proliferating cell nuclear antigen stimulates flap endonuclease 1. *J Biol Chem*, 2000;275(14) 10498-10505.
- [181] Sakurai S, Kitano K, Yamaguchi H, Hamada K, Okada K, Fukuda K, et al. Structural basis for recruitment of human flap endonuclease 1 to PCNA. *EMBO J*, 2005;24(4) 683-693.
- [182] Shen B, Qiu J, Hosfield D, Tainer JA. Flap endonuclease homologs in archaebacteria exist as independent proteins. *Trends Biochem Sci*, 1998;23(5) 171-173.
- [183] Hwang KY, Baek K, Kim HY, Cho Y. The crystal structure of flap endonuclease-1 from *Methanococcus jannaschii*. *Nat Struct Biol*, 1998;5(8) 707-713.
- [184] Hosfield DJ, Mol CD, Shen B, Tainer JA. Structure of the DNA repair and replication endonuclease and exonuclease FEN-1: coupling DNA and PCNA binding to FEN-1 activity. *Cell*, 1998;95(1)135-146.
- [185] Matsui E, Musti KV, Abe J, Yamasaki K, Matsui I, Harata K. Molecular structure and novel DNA binding sites located in loops of flap endonuclease-1 from *Pyrococcus horikoshii*. *J Biol Chem*, 2002;277(40) 37840-37847.
- [186] Chapados BR, Hosfield DJ, Han S, Qiu J, Yelent B, Shen B, et al. Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, 2004;116(1) 39-50.
- [187] Matsui E, Kawasaki S, Ishida H, Ishikawa K, Kosugi Y, Kikuchi H, et al. Thermostable flap endonuclease from the archaeon, *Pyrococcus horikoshii*, cleaves the replication fork-like structure endo/exonucleolytically. *J Biol Chem*, 1999;274(26) 18297-18309
- [188] Matsui E, Abe J, Yokoyama H, Matsui I. Aromatic residues located close to the active center are essential for the catalytic reaction of flap endonuclease-1 from hyperthermophilic archaeon *Pyrococcus horikoshii*. *J Biol Chem*, 2004;279(16) 16687-16696.
- [189] Beattie TR, Bell SD. Molecular machines in archaeal DNA replication. *Curr Opin Chem Biol*, 2011;15(5) 614-619.
- [190] Ishino S and Ishino Y. Rapid progress of DNA replication studies in Archaea, the third domain of life. *Sci China Life Sci*, 2012;55(5) 386-403.

- [191] Wang Y, Prosen DE, Mei L, Sullivan JC, Finney M, Vander Horn PB. A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. *Nucleic Acids Res*, 2004;32(3) 1197-1207
- [192] Kawamura A, Ishino Y, Ishino S. Biophysical analysis of PCNA from *Pyrococcus furiosus*. *J Jap Soc Extremophiles*, 2012;11(1) 12-18
- [193] Ishino S, Kawamura, A, Ishino Y. Application of PCNA to processive PCR by reducing the stability of its ring structure. *J Jap Soc Extremophiles*, 2012;11(1) 19-25
- [194] Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*, 2008;26(10) 1135-1145.
- [195] Ansorge WJ. Next-generation DNA sequencing techniques. *Nat Biotechnol*, 2009;25(4) 195-203
- [196] Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*, 2010;11(1) 31-46.