DNA replication in the hyperthermophilic archaeon *Sulfolobus solfataricus*

I. Dionne, N.P. Robinson, A.T. McGeoch, V.L. Marsh, A. Reddish and S.D. Bell[1]

The Medical Research Council Cancer Cell Unit, Hutchison MRC Centre, Hills Road, Cambridge CB2 2XZ, U.K.

Abstract

Studies of the DNA-replication machinery of Archaea have revealed striking similarities to that of eukaryotes. Indeed, it appears that in most cases Archaea possess a simplified version of the eukaryotic replication apparatus. Studies of Archaea are therefore shedding light on the fundamental processes of DNA replication in both domains of life.

DNA replication in eukaryotes is a vastly complex procedure involving interaction between many tens of proteins in a highly orchestrated manner to ensure precise and timely firing of replication origins and completion of synthesis prior to cell division [1]. The complexity of the eukaryotic apparatus presents a considerable technical challenge to biochemical analysis. However, in recent years it has become apparent that a simpler model system exists that can help to elucidate the core set of interactions at the heart of the replication machinery. This has come about with the recognition that the DNA-replication apparatus of Archaea resembles a simplified version of that found in eukaryotes [2,3]. As can be seen in Figure 1, the Archaea possess homologues of many eukaryotic replication-associated proteins and in many cases have a simplified version of the eukaryotic counterpart. A perfect example of this is the presumptive archaeal replicative helicase, the mini-chromosome maintenance (MCM) complex. As is discussed in the accompanying article by Jenkinson and Chong [3a] this is a homo-multimeric helicase complex (intriguingly, whereas it is a double hexamer in some species, the MCM complex in other species is a single hexamer) [4–6]. However, in contrast with the homo-multimeric archaeal MCM, the eukaryotic MCM is a heteromultimer of six distinct subunits. The six eukaryotic MCM subunits all possess sequence homology with each other and have presumably arisen via diversification of an Archaea-like progenitor MCM. The simplicity of composition of the MCM complex in Archaea has greatly facilitated biochemical dissection of the molecular events involved in the helicase action of this protein [4–7].

Another case where Archaea possess a simplified version of the eukaryotic machinery lies in the origin recognition complex (ORC). In eukaryotes, this contains six separate subunits (Orc1–6) and binds to origins of replication via direct protein–DNA interactions [1]. Archaea only possess homologues of the Orc1 subunit. Intriguingly the archaeal Orc1 is also related to eukaryotic Cdc6. In eukaryotes, Cdc6 plays a pivotal role in the recruitment of the MCM complex to origins. Some Archaea (e.g. *Pyrococcus abyssi*) only have a single Orc1/Cdc6 homologue and it has been proposed that this plays the dual roles of origin recognition and MCM recruitment. In contrast, other Archaea, e.g. *Sulfolobus solfataricus*, have multiple Orc1/Cdc6 homologues and it is possible that in these species the individual homologues have diversified to play unique roles in origin recognition or MCM recruitment. Alternatively, the exciting possibility exists that some archaeal species may possess multiple origins of replication and that each origin is recognized by a different Orc1/Cdc6 protein. We have recently found that one of the three Orc1/Cdc6 homologues in *S. solfataricus* plays a direct role in origin recognition, binding in an ATP-dependent manner to three conserved origin recognition boxes (ORB1, 2 and 3). Intriguingly, the ORB sequence is conserved in most other Archaea, suggesting that this is a common mechanism for origin definition in this domain of life (N.P. Robinson and S.D. Bell, unpublished work).

While much remains to be discovered about the molecular details of origin recognition and firing in Archaea, a considerable body of research has been generated on the elongation stage of replication. This work has centred around the processivity factor for the DNA polymerase, the sliding clamp, proliferating cell nuclear antigen (PCNA) [8–12]. In most eukaryotic and archaeal species PCNA is a homotrimeric molecule that forms a toroidal structure with a central cavity, through which DNA is proposed to pass. However, in order to load on DNA, the PCNA ring must be opened to allow admission of DNA to the central cavity. This loading of PCNA is mediated by the clamp loader, replication factor C (RFC), in a reaction that involves the hydrolysis of ATP (reviewed in [13,14]).

In eukaryotes and Archaea, DNA polymerase interacts via a conserved PCNA-interaction motif with the inter-domain connector loop on the outer surface of the PCNA ring [15]. Thus PCNA acts as a means of tethering the DNA polymerase to DNA, thereby enhancing the processivity of the polymerase. Intriguingly, DNA polymerase is not the

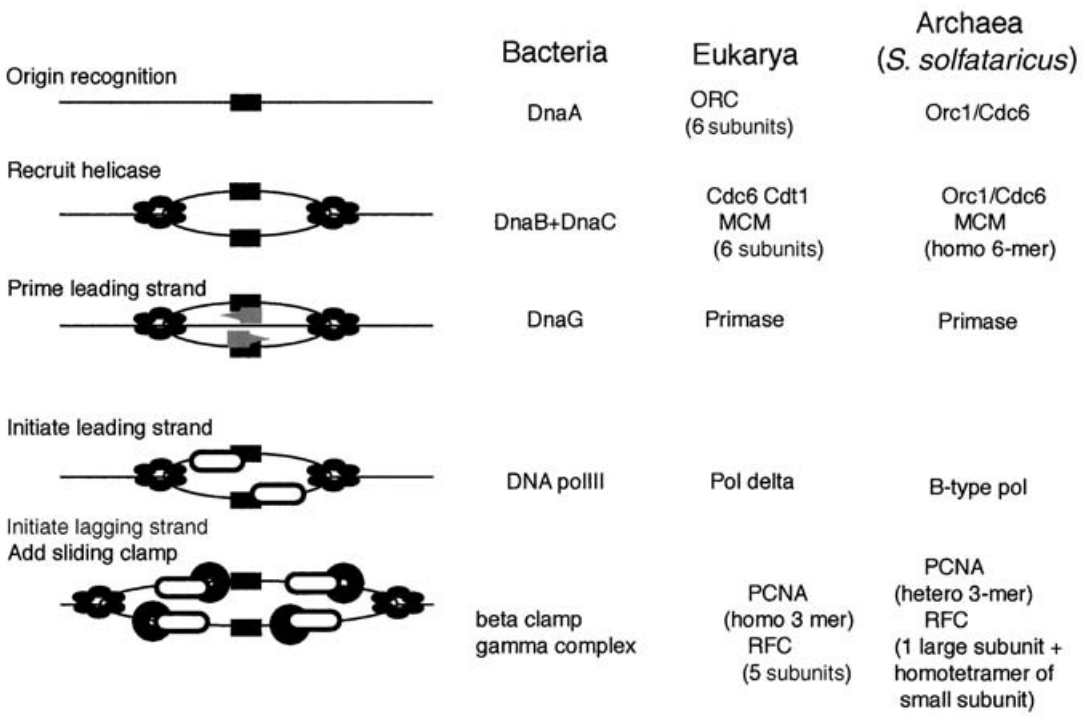
Key words: Archaea, evolution, lagging strand, proliferating cell nuclear antigen (PCNA), replication.

Abbreviations used: MCM, mini-chromosome maintenance; ORC, origin recognition complex; PCNA, proliferating cell nuclear antigen; RFC(-L/-S) replication factor C (large/small subunit); FEN1, flap endonuclease 1.

[1]To whom correspondence should be addressed (e-mail sb419@hutchison-mrc.cam.ac.uk).

Figure 1 | The stages in DNA replication, demonstrating the relatedness between archaeal and eukaryotic replication machineries

The molecules responsible for catalysing the various stages in the processes of origin opening and firing are named. Note that the archaeal column corresponds to proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*; in contrast to *S. solfataricus*, most Archaea have a homo-trimeric PCNA. In addition, euryarchaeal species contain an additional second family of DNA polymerases, the D-type polymerases [18].



only factor to interact with PCNA, for example enzymes involved in Okazaki fragment maturation such as DNA ligase 1 and the flap endonuclease 1 (FEN1) also interact with, and have their activity stimulated by, PCNA. In addition, a variety of DNA-repair factors have been found to interact with PCNA [15].

The interaction of DNA polymerase, FEN1 and ligase 1 with PCNA has led to the generation of a model in which PCNA acts as a scaffold for the assembly of these factors in a stepwise manner during Okazaki-fragment processing. However, the fact that there are three factors potentially interacting with the three identical subunits of PCNA also suggests the possibility that these factors could simultaneously interact with PCNA. As discussed below, our recent data provide support for this latter model.

Our work arose from the observation that *S. solfataricus* encodes three homologues of PCNA (PCNA 1, 2 and 3). To our considerable surprise, we found by affinity chromatography, gel filtration, immunoprecipitation and surface-plasmon-resonance experiments that the individual PCNA subunits did not homo-multimerize but rather formed a unique heterotrimeric PCNA. Furthermore, the heterotrimer had a distinct order of assembly; first, PCNA1 and PCNA2 form a heterodimer. Formation of this heterodimer is a prerequisite for the subsequent recruitment of the third subunit, PCNA3 [16]. We have made point mutations in the N- and C-terminal intersubunit-interaction motifs and found that the polarity of the subunits in the ring is head-to-tail PCNA1–PCNA2–PCNA3 (I. Dionne and S.D. Bell, unpublished work).

Figure 2 | Diagram showing the simultaneous binding of DNA polymerase, flap endonuclease and DNA ligase to PCNA

PCNA shown is based on the structure of human PCNA from [19].

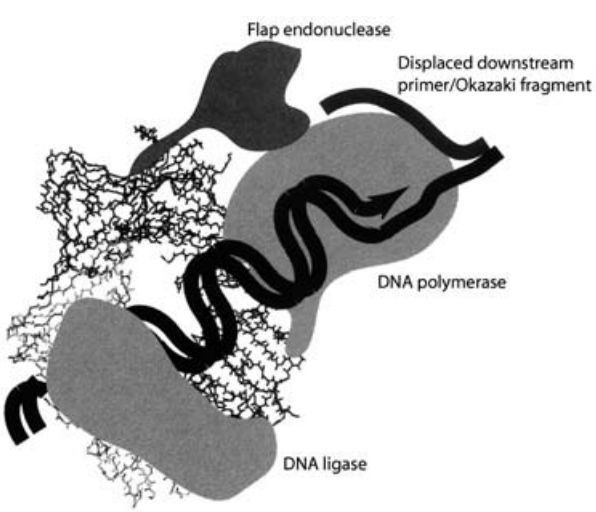
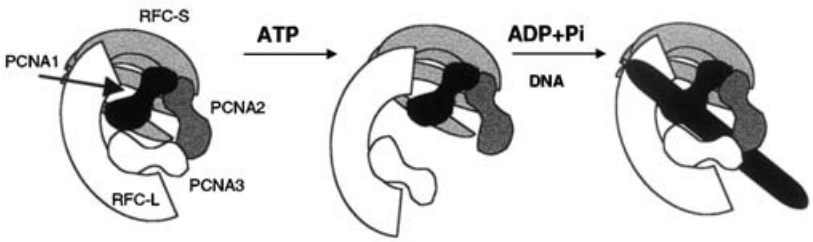


Figure 3 | Model for the opening and loading on to DNA of PCNA by the clamp loader RFC

The large subunit of RFC (RFC-L) is shown as a white crescent and the tetrameric small subunit (RFC-S) shown as four grey crescents. The three subunits of PCNA are coloured black (PCNA1), grey (PCNA2) and white (PCNA3). RFC-L binds PCNA3 and the PCNA1–2 dimer is bound by the RFC-S tetramer. Binding of ATP is proposed to open the RFC structure, thereby levering PCNA3 off the PCNA1–2 dimer. Upon addition of DNA, ATP is hydrolysed, resealing PCNA around DNA.



This unique situation of a PCNA with three distinct subunits led us to speculate that individual PCNA subunits might have distinct interaction partners. Using affinity chromatography and yeast two-hybrid assays we found highly preferred interactions between PCNA1 and FEN1, PCNA2 and DNA polymerase B1, and PCNA3 and ligase 1. This suggested that these three factors could bind simultaneously to the PCNA ring and indeed we were able to demonstrate that the PCNA heterotrimer could bridge between FEN1, ligase 1 and DNA polymerase B1. This therefore provides evidence for a model for Okazaki-fragment processing in which a constitutive complex of ligase, FEN1 and PCNA translocates with the polymerase and allows an extremely tight coupling of synthesis and processing events (Figure 2). Although the unique asymmetry of the *S. solfataricus* PCNA facilitated detection of this complex, we feel that it is likely that the model is generally applicable to homotrimeric PCNAs too [16].

Clearly, in the case of the *S. solfataricus* PCNA-interacting factors, partner specificity is exerted in the interaction between the PCNA interaction motif and the interdomain connector loop. We are currently employing selection methodologies to determine these rules for subunit specificity.

It was also possible to exploit the asymmetry of *S. solfataricus* PCNA to examine the interactions between RFC and PCNA during clamp loading. In Archaea, RFC possesses a single large subunit (RFC-L) in complex with a homotetramer of a small subunit (RFC-S). We found that PCNA3 (the last subunit to enter the PCNA ring) interacted with RFC-L and that the RFC-S tetramer bound to the dimer of PCNA1 and PCNA2 [16]. In conjunction with data on ATP binding and hydrolysis by RFC [17] these interaction studies suggest the model for clamp-loading shown in Figure 3.

It is generally accepted that the archaeal replication machinery by virtue of its simplicity coupled with its position in the tree of life will shed insight into the more complex yet fundamentally related eukaryotic machinery. However, an unanticipated complication in the *S. solfataricus* machinery, the presence of a heterotrimeric PCNA, has also proven to be highly informative. It is anticipated that the experimental tractability of the archaeal machinery will be of great value in establishing the biochemical bases of the fundamental paradigms in the process of DNA replication. Key goals lie in unravelling the molecular processes whereby the MCM complex is loaded at replication origins and in understanding the biochemical basis of origin firing and its regulation in the archaeal cell cycle.

This work was funded by the Medical Research Council.

References

1 Bell, S.P. and Dutta, A. (2002) Annu. Rev. Biochem. **71**, 333–374
2 Kelman, Z. (2000) Trends Biochem. Sci. **25**, 521–523
3 Bernander, R. (2000) Trends Microbiol. **8**, 278–283
3a Jenkinson, E.R. and Chong, J.P.J. (2003) Biochem. Soc. Trans. **31**, 669–673
4 Kelman, Z., Lee, J.K. and Hurwitz, J. (1999) Proc. Natl. Acad. Sci. U.S.A. **96**, 14783–14788
5 Chong, J.P.J., Hayashi, M.K., Simon, M.N., Xu, R.M. and Stillman, B. (1999) Proc. Natl. Acad. Sci. U.S.A. **97**, 1530–1535
6 Shechter, D.F., Ying, C.Y. and Gautier, J. (1999) J. Biol. Chem. **275**, 15049–15059
7 Poplawski, A., Grabowski, B., Long, S.F. and Kelman, Z. (2001) J. Biol. Chem. **276**, 49371–49377
8 Kelman, Z. and Hurwitz, J. (2000) J. Biol. Chem. **275**, 7327–7336
9 Daimon, K., Kawarabayasi, Y., Kikuchi, H., Sako, Y. and Ishino, Y. (2002) J. Bacteriol. **184**, 687–694
10 Cann, I.K.O., Ishino, S., Hayashi, I., Komori, K., Toh, H., Morikawa, K. and Ishino, Y. (1999) J. Bacteriol. **181**, 6591–6599
11 Cann, I.K.O., Ishino, S., Yuasa, M., Daiyasu, H., Toh, H. and Ishino, Y. (2001) J. Bacteriol. **183**, 2614–2623
12 Oyama, T., Ishino, Y., Cann, I.K.O., Ishino, S. and Morikawa, K. (2001) Mol. Cell **8**, 455–463
13 Ellison, V. and Stillman, B. (2001) Cell **106**, 655–660
14 Jeruzalmi, D., O'Donnell, M. and Kuriyan, J. (2002) Curr. Opin. Struct. Biol. **12**, 217–224
15 Warbrick, E. (2000) Bioessays **22**, 997–1006
16 Dionne, I., Nookala, R.K., Jackson, S.P., Doherty, A.J. and Bell, S.D. (2003) Mol. Cell **11**, 275–282
17 Shiomi, Y., Usukura, J., Masamura, Y., Takeyasu, K., Nakayama, Y., Obuse, C., Yoshikawa, H. and Tsurimoto, T. (2000) Proc. Natl. Acad. Sci. U.S.A. **97**, 14127–14132
18 Cann, I.K.O. and Ishino, Y. (1999) Genetics **152**, 1249–1267
19 Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M. and Kuriyan, J. (1996) Cell **87**, 297–306

Received 31 January 2003