THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 18, Issue of May 3, pp. 16179–16188, 2002
Printed in U.S.A.


# Elucidation of an Archaeal Replication Protein Network to Generate Enhanced PCR Enzymes*

Received for publication, August 14, 2001, and in revised form, December 21, 2001
Published, JBC Papers in Press, January 22, 2002, DOI 10.1074/jbc.M107793200

**Michael Motz‡§, Ingo Kober§¶, Charles Girardot¶, Eva Loeser¶, Ulrike Bauer‡, Michael Albers‡, Gerd Moeckel‖, Eric Minch‖, Hartmut Voss‡, Christian Kilger‡, and Manfred Koegl¶‖****

*From the ‡Exploratory Research, ¶Proteins and Assays, and ‖Life Science Informatics, LION Bioscience Ktiengesellschaft, D-69120 Heidelberg, Germany*

**Thermostable DNA polymerases are an important tool in molecular biology. To exploit the archaeal repertoire of proteins involved in DNA replication for use in PCR, we elucidated the network of proteins implicated in this process in *Archaeoglobus fulgidus*. To this end, we performed extensive yeast two-hybrid screens using putative archaeal replication factors as starting points. This approach yielded a protein network involving 30 proteins potentially implicated in archaeal DNA replication including several novel factors. Based on these results, we were able to improve PCR reactions catalyzed by archaeal DNA polymerases by supplementing the reaction with predicted polymerase co-factors. In this approach we concentrated on the archaeal proliferating cell nuclear antigen (PCNA) homologue. This protein is known to encircle DNA as a ring in eukaryotes, tethering other proteins to DNA. Indeed, addition of *A. fulgidus* PCNA resulted in marked stimulation of PCR product generation. The PCNA-binding domain was determined, and a hybrid DNA polymerase was constructed by grafting this domain onto the classical PCR enzyme from *Thermus aquaticus*, *Taq* DNA polymerase. Addition of PCNA to PCR reactions catalyzed by the fusion protein greatly stimulated product generation, most likely by tethering the enzyme to DNA. This sliding clamp-induced increase of PCR performance implies a promising novel micromechanical principle for the development of PCR enzymes with enhanced processivity.**

For all forms of life, the process of DNA-replication is essential in the propagation of genetic information. A complex multiprotein machinery including DNA polymerases, processivity factors, proof-reading, repair, and regulatory activities (1–3) has evolved to handle the tasks associated with this process. The importance of replication is demonstrated by the fact that its central features are highly conserved among all cellular organisms, while the protein sequences of some of the factors are not. This is exemplified by the processivity factors of DNA polymerases. Processivity is defined as the number of polymerization events during a single contact between polymerase and template. To prevent dissociation off the template DNA, many polymerases are bound by a protein that encircles the DNA in a ring-like structure. Together with the polymerase, the ring appears to move along the DNA as replication proceeds. Such "sliding clamps" exist both for eubacteria (the β-subunit of the polymerase) and for eukaryotes (PCNA),[1] and the structures of these proteins are almost superimposable (4). However, the proteins are highly deviant in primary sequence. This situation of sequence divergence and functional conservation is also evident for the proteins responsible for loading the sliding clamp onto the DNA. Whereas in eukaryotes the clamp loader is a heteropentameric complex called RFC (2, 5), the eubacterial clamp loader is represented by the pentameric so called γ-complex (6, 7). Until recently very little was known about replication in the third domain of life, the archaebacteria. In part, this is probably due to the difficult culture conditions for these often extremophile organisms. Only with the recent availability of the genomes of several archaea (8) have these organisms become amenable for bioinformatic and functional genomic analyses.

Such analyses support the notion that archaebacteria, although possessing metabolic features that are quite similar to bacterial processes, are more closely related to eukaryotes when translation, transcription, and replication proteins are compared. For example a clear homologue of PCNA can be found in all published archaebacterial genomes, (9) but no homologues for the eubacterial β-clamp or other eubacterial replication factors have been discovered. Furthermore, *bona fide* homologues of the eukaryotic DNA polymerase δ can be identified in archaea (10).

In other cases, the situation is not as straightforward. For the eukaryotic five-subunit clamp loader RFC, only two homologues could be identified so far in archaea (10, 11). Furthermore, relatively recent work detected a completely novel two-subunit DNA polymerase (DP1 and DP2) with assumed replicative function in thermophilic archaea, the large subunit of which displays no significant homology to any other DNA polymerase (12, 13).

Due to their use in DNA-sequencing and PCR applications, heat-stable DNA polymerases are technically and economically important enzymes. Current PCR reactions rely on single enzymes, which like the DNA polymerase from *Thermus aquaticus* (*Taq*) often originate from cellular DNA repair polymerases (14). Others represent potential replicative enzymes such as *Pfu* and *Pwo* (15), but are employed in the absence of any accessory replication factor. For longer PCR templates, artificial enzyme mixes with compromised fidelity are the state of the art (16). Remarkably, however, these reactions are orders

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Both authors contributed equally to this work.

** To whom correspondence should be addressed: LION bioscience AG, Im Neuenheimer Feld 515-519, D-69120 Heidelberg, Germany. Tel.: 49-6221-40-38-140; Fax: 49-6221-40-38-401; E-mail: manfred.koegl@lionbioscience.com.

[1] The abbreviations used are: PCNA, proliferating cell nuclear antigen; RFC, replication factor C; Y2H, yeast two-hybrid; S, small fraction; L, large fraction; YPDA, yeast extract/peptone/dextrose with adenine; SD-HUL, synthetic medium-histidine, uracil, and leucine.

of magnitude less efficient than the *in vivo* DNA replication process with respect to fidelity, speed, and product size. The replacement of current monomeric PCR enzymes by a selected subset of proteins taken from the replication machinery of a thermophilic archaebacterium may hence open new chances to improve PCR.

In this report a biochemical, bioinformatic, and functional proteomic approach was undertaken to deepen our understanding of archaeal replication to use this knowledge for the improvement of PCR applications. We demonstrate that processivity factors can enhance PCR catalyzed by archaeal polymerases. Finally, we show that by tethering *Taq* polymerase to a sliding clamp via a PCNA-interaction domain one can generate PCR enzymes with increased performance.

## EXPERIMENTAL PROCEDURES

*Bioinformatic Analysis of Replication Proteins in Archaeoglobus fulgidus*—In the first step of the bioinformatics analysis most of the key components of the DNA replication machinery in humans were identified. The following sequences were retrieved from the SWISSPROT database (26): AC11_HUMAN, AC12_HUMAN, AC13_HUMAN, AC14_HUMAN, AC15_HUMAN, PCNA_HUMAN, DPD2_HUMAN, and DPOD_HUMAN. In a first step, these sequences were compared with BLASTP (27) against a non redundant database of all publicly available protein sequences. In a next step, homologous sequences from related organisms including *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* from the blast result were used to generate a multiple-sequence alignment using the ClustalW program (28). Based on the obtained alignment a Hidden Markov Model was generated with the HMMER software package (version 1.8, Sean Eddy, Dept. of Genetics, Washington University School of Medicine, St. Louis, MO). This model includes the highly conserved regions but also the less homologous parts of the sequence alignment in between. Now the non-redundant sequence database was again searched with the HMMER program for distantly related homologues especially in *A. fulgidus*. Based on the sequence information of TIGR (8), a database of all *A. fulgidus* sequences was generated. This database was also searched in parallel. The following sequences were identified: AF2060 homologous to AC11_HUMAN, AC12_HUMAN, AC13_HUMAN, and AC14_HUMAN; AF1195 homologous to AC15_HUMAN; AF0335 homologous to PCNA_HUMAN; AF1790 homologous to DPD2_HUMAN; and AF0497 homologous to DPOD_HUMAN. Another protein with polymerase activity was identified by applying the same methods starting from a recently published sequence in *P. furiosus*: DP2L_PYRFU (12). This protein is homologous to AF1722.

Using these strategies we retrieved sequences from *A. fulgidus*, which are likely part of the replication machinery. The analysis was performed using the described bioinformatics tools, which were implemented in bioSCOUT (LION bioscience AG, Heidelberg, Germany, lionbioscience.com/solutions/bioscout), a multifunctional sequence analysis program package. bioSCOUT also includes all publicly available databases as well as proprietary and in-house databases. All sequence analysis runs were made on LION's in-house servers on database content of about one terabyte. An automatic alert service implemented in bioSCOUT was used to retrieve all newly entered sequences in all major sequence databases (*e.g.* NCBI, EMBL, SWISSPROT, PIR), related to the described project.

*Two-hybrid Methods*—For the construction of a genomic library of the 2,178,400-bp sized *A. fulgidus* genome for yeast two-hybrid (Y2H) screening, genomic DNA was fragmented by sonication and cloned into pGAD424; 16 $\mu$g of genomic DNA of *A. fulgidus* (obtained from the German Collection of Microorganisms and Cell Cultures) were sonicated for 8 s, such that fragments ranging from 0.1 to 5 kb were generated. To minimize the cloning bias, the ligations and transformations were done in two size classes in separate reactions. Size classes were 0.3–0.7 kb (small, the S fraction) and 0.7–2.5 kb (large, the L fraction). DNA fragments of the two size classes were isolated by cutting out the respective areas of a 1% agarose gel after electrophoretic isolation in TBE. 300 ng and 450 ng of sonicated DNA of the S and L fractions, respectively, were filled in with 2.5 units of *Pwo* polymerase (Roche Molecular Biochemical) and 35 $\mu$M of an equimolar dNTP mix (Roche Molecular Biochemical) for 30 min at 72 °C in the buffer provided with the *Pwo* polymerase. After the reaction, the DNA was purified using the QIAquick PCR purification kit (Qiagen). 5 $\mu$g of pGAD424 were digested with *Sma*I (New England Biolabs) and dephosphorylated with calf intestine phosphatase, (Roche Molecular Biochemical) and purified from an agarose gel after electrophoresis. The DNA fragments were ligated into the linearized vector at 16 °C overnight using 5 units of T4 DNA ligase (United States Biochemicals) each time in 50 $\mu$l using the buffer provided by the manufacturer and a total amount of DNA of 170 or 300 ng for the S and L fraction, respectively. The optimal ratio of vector to insert mass had been determined beforehand and was 1:1.5 for the small fraction, and 1:2 for the large fraction. The library was transformed into DH10B *Escherichia coli* cells by electroporation, yielding $2 \times 10^5$ and $3.5 \times 10^5$ independent colonies for the S and L fractions, respectively. Colonies were washed off, DNA was prepared, and the library was transformed into haploid yeast PJ69–4$\alpha$ (25) using 200 and 300 15-cm dishes for the small and large insert fractions, each plate containing 1000–2000 colonies. Each plate was then washed off with YPDA containing 20% glycerol and stored separately in the wells of 2-ml deep well microtiterplates at −80 °C. Thus, one copy of the library was contained in a set of five microtiterplates. As working stocks, the library was diluted to an optical density at 600 nm ($OD_{600}$) of 5.0. The final library included roughly 500,000 independent clones. Thus, fusion proteins will be generated every 4.4 bases, resulting in an in-frame fusion every nine amino acids, on average.

For the generation of a library of fragments of the AF0497/PolB gene, we followed the same protocol as for the generation of the library with the following alterations. We subjected a PCR fragment representing the coding region of the gene to sonication for 20 s. Fragment size was 0.05–0.3 kb. Blunt fragment ends were generated using *Pwo* polymerase as described above, and fragments were cloned into *Sma*I-linearized pGBDU and pGAD424. $10^4$ independent colonies were collected for both vectors after transformation in *E. coli*. DNA was prepared and transformed into PJ69- and PJ69–4$\alpha$, to yield $5 \times 10^5$ yeast colonies. Cells were washed off the plates as described above and frozen. These libraries were used to isolate clones coding for protein fragments interacting with AF0335/PCNA by mating of yeast cells as described below.

Bait and prey vectors were generated by the two-step PCR protocol as described in Hudson *et al.* (17). 5′ extension for the 2nd PCR was GAATTCGGTACCACCACCATG, and 3′ extension was GATCCCCGGGAATTGCCATGTC.

For the mating of yeast cells, bait proteins were cultured overnight to an $OD_{600}$ of 1–2 in selective medium. For each bait construct, four independent clones were cultured and mixed in equal amounts of cells, to minimize the chances of working with a clone carrying detrimental mutations from the PCR reactions. Library cells were thawed out and allowed to recover for 1 h at ambient temperature. For screens, an 0.5-OD solution of the bait mix was prepared in YPDA. 10 $\mu$l of the library were transferred from the library plates to new 96 deep well plates that contained 10 $\mu$l of YPDA and allowed to recover for 2 h at 30 °C. After the incubation, 90 $\mu$l of the bait solution were added. Mating was allowed overnight, shaking at 1000 rpm in a microtiterplate shaker. The mating mixture was then diluted with 1.1 ml of SD-HUL containing varying amounts of 3-amino triazole and cultivated for 6–10 days at 30 °C.

For the pairwise mating, the bait mix and the prey mix were prepared at a concentration of 1 OD/ml in SD-HUL and SD-L, respectively. 50 $\mu$l of each were mixed together for the mating. The mating mixture was then diluted with 1.1 ml of SD-HUL containing varying amounts of 3-amino triazole and cultivated for 6–10 days at 30 °C.

However, mating efficiencies were not satisfying in microtiterplates, and mating in flasks proved to be the superior method in our hands. Thus, for the second round of library screens, the small and large fractions of the library were combined into two separate pools, and mating was done in 50-ml Erlenmeyer flasks in YPDA at an OD of 1.0 in a total volume of 10 ml, shaking for 5 h at 100 rpm. In this case, after the mating cells were washed in selective medium lacking histidine, uracil, and leucine (SD-HUL), and aliquoted into microtiterplates in the same medium.

For both protocols as well as for library screens, after 6–10 days the cells were passaged once to microtiterplates containing fresh selection medium and allowed to grow for 2–4 days. Potential positives were identified from wells displaying cell growth. The activation of the secondary reporter genes *ADE2*, *lacZ*, and *Mel1* was determined, and double-positive cells were collected. For ease of measurement, activation of *Mel1* was determined using 5-bromo-4-chloro-3-indoyl-a-D-galactopyranoside (X-$\alpha$-Gal, CLONTECH) and scoring for blue color. The inserts were amplified by PCR from the yeast cells, sequenced, and analyzed using standard bioinformatic methods as well as bioSCOUT for annotation. Interaction networks were visualized using the automated viewer piSCOUT® and reviewed for sticky proteins identified with multiple unrelated baits.

*DNA Polymerization Assays Using Activated DNA in Filter-binding Assays*—Polymerization reactions were set up in a volume of 50 μl of 10 mM Tris/Cl, pH 7.5, 50 mM KCl, 4 mM $MgCl_2$, 0.2 μg/μl bovine serum albumin, 0,5 mM digoxigenin dUTP (Roche Molecular Biochemical), 12 ng/μl activated calf thymus DNA, and 100 μM each of the four dNTPs. Incubation after addition of proteins as indicated in the figures was for 30 min at 68 °C. After precipitation by ethanol/sodium acetate samples were resuspended in 20 μl of 100 mM Tris/Cl, pH 7.9, and 10 μl of each sample were dotted into the wells of a 96-well silent screen plate with Nylon 66 Biodyne B 0.45 μM pore membrane (Nunc). Nucleic acids were fixed on the membrane by baking at 70 °C for 10 min. Detection of incorporated digoxigenin was performed with the DIG luminescent detection kit for nucleic acids (Roche Molecular Biochemical) as recommended by the manufacturer.

*DNA Polymerization Assays Using Primer Extension Analysis*—For primer extension analysis with the *A. fulgidus* proteins, a CGCGCGGGGAGAGGCGGTTTGC primer was annealed to single-stranded circular M13mp18 DNA, and 1 μg of the primed DNA were incubated for 4 min at 68 °C with recombinant proteins as described in the figure legends in a buffer containing 50 mM KCl, 4 mM $MgCl_2$, 10 mM Tris, pH 7.5, 20 μM digoxigenin dUTP, and 25 μM each of the four dNTPs. Total reaction volume was 50 μl, and reactions were terminated by phenol/chloroform extraction. Nucleic acids were then precipitated with ethanol/sodium acetate, and pellets were resuspended in 5 μl of 20 mM NaOH, 20% glycerol, 0,1% bromphenol blue. Samples were loaded onto an 1% agarose gel containing 25 mM NaOH and 2 mM EDTA. After electrophoresis the gel was blotted onto a Biodyne A membrane and detection of digoxigenin labeled DNA was performed with the DIG luminescent detection kit for nucleic acids (Roche Molecular Biochemical) as recommended by the manufacturer.

For primer extension analysis with the chimeric *Taq* polymerase, a IRD 700-labeled M13 primer was annealed to 1 μg of M13 plasmid construct at 53 °C for 10 min after 5 min of denaturation at 95 °C in a buffer containing 50 mM KCl, 2 mM $MgCl_2$, 10 mM Tris, pH 7.5, and 25 μM each of the four dNTPs. At the start of the annealing heat-purified chimeric *Taq* polymerase was added with different amounts of *A. fulgidus* PCNA and then incubated for 15 min at 68 °C. The reactions at a volume of 50 μl were terminated by phenol/chloroform extraction. Nucleic acids were then precipitated with ethanol/sodium acetate, and pellets were resuspended in 2 μl of formamide. 0.2 μl of the concentrated solution was then resolved by 4% PAGE on a LICOR DNA sequencer Long READIR 4200 (LICOR, Lincoln, NE) under standard electrophoresis conditions.

*PCR Reactions Using Polymerases of A. fulgidus*—PCR reactions were performed in a Stratagene robocycler gradient. If not stated otherwise in the figure legends, reactions were performed in a total volume of 50 μl of 1 × PCR buffer containing 2 mM $MgCl_2$, 50 mM KCl, and 10 mM Tris/Cl (pH 8.3 for *Taq* reactions and pH 7.5 for all reactions performed by *A. fulgidus* enzymes) at a deoxynucleotide concentration of 200 μM each. Because stimulation of the polymerase by PCNA is magnesium-dependent (data not shown), reactions were supplemented with $MgCl_2$ to a final concentration of 4 mM where appropriate. Primer concentrations were 300 nM in all cases, and for each reaction 20 ng of template DNA were employed. Two amplicons were used for the experiments presented here. For amplification of a 420-bp product from M13 mp18 closed circular single-stranded DNA, primer sequences were GGATTGACCGTAATGGGATAGGTTACGTT and AGCGGATAACAATTTCACACAGGAAACAG, respectively. Cycle number was as indicated in the figures, and individual cycles were 30 s at 95 °C, 30 s at 59 °C, and 60 s at 68 °C. For amplification of a 3395-bp construct from a pQE30 plasmid harboring the coding sequence of AF 1722 cloned into the *Bam*HI/*Kpn*I site of pQE 30, the same primers were employed, which were previously used to amplify the insert DNA from a genomic *A. fulgidus* template (ACGCGCGGATCCGATGCAACTCTTGACAGGTTC and ACGGGGTACCAAATCCTTCTCGTCCACAGG). For amplification of this construct, an initial denaturation step of 180 s at 95 °C was followed by 35 cycles of 30 s at 95 °C, 30 s at 55 °C, and 360 s at 68 °C, and the last cycle was extended for 10 min at 68 °C. After cycling 5 μl of all PCR reactions were subjected to electrophoresis on 0.8% agarose gels run in 1 × TBE.

*Generation of the Chimeric Taq DNA Polymerase*—In an overlapping PCR reaction, a fusion construct was generated comprising the DNA sequence of the *Taq* DNA polymerase and the *Afu* polB carboxyl-terminal terminal 50 amino acids. First, via PCR the gene of *Taq* DNA polymerase (GenBank™ accession no. J04639) was amplified from a *T. aquaticus* strain YT-1 (ATCC-25104, Manassas, VA). For amplification of this construct with the primers Taqfw, GGATGCTGCCCCTCTTTG, and Taqrev, TGGCCGCGGCCGCGGTGGTCACTCCTTGGCGGAGAGC, an initial denaturation step of 180 s at 94 °C was followed by 30 cycles of 30 s at 94 °C, 30 s at 55 °C, and 120 s at 72 °C, and the last cycle was extended for 7 min at 72 °C. The primer at the 3′ end was designed with an overhang sequence comprising the linker sequence (TGGCCGCGGCCGCGGTGG). Additionally in a second PCR reaction the region of the 50 carboxyl-terminal amino acids of *Afu* polB has been amplified with a primer at the 3′ end having the complementary linker sequence. For amplification of this fragment with the primers AfupolBfw (CCACCGCGGCCGCGGCCAAAGAGCGGAATAGAGATA) and

TABLE I
*Interactions of proteins as determined by Y2H*

Bait proteins are generated as fusion of the respective complete open reading frame with the sequence encoding the Gal4 DNA-binding domain in pGBDU (25). Prey proteins are encoded by inserts from the genomic library (Y2H-screen) or contain the complete open reading frame encoded by a fusion gene with the activation domain of Gal4 in pGAD424 (matrix mating).

| Bait | Prey | Method |
|---|---|---|
| AF0130/HDAC[a] | AF0472/Prp | Y2H-screen |
| AF0335/PCNA[b] | AF0264/RAD2 | Y2H-screen |
| AF0335/PCNA[b] | AF0335/PCNA | matrix-mating |
| AF0335/PCNA[b] | AF0497/PolB | matrix-mating |
| AF0335/PCNA[b] | AF0621/RNAseHII | Y2H-screen |
| AF0335/PCNA[b] | AF1195/RFC-53 | Y2H-screen |
| AF0335/PCNA[b] | AF1347 | Y2H-screen |
| AF0335/PCNA[b] | AF1722/DP2-I | matrix-mating |
| AF0382/RPA-26[b] | AF0111/IMPDH | Y2H-screen |
| AF0382/RPA-26[b] | AF0699/RecJ-2 | Y2H-screen |
| AF0382/RPA-26[b] | AF0735/RecJ-1 | Y2H-screen |
| AF0382/RPA-26[b] | AF0735/RecJ-1 | matrix-mating |
| AF0382/RPA-26[b] | AF0780/RPA-36 | matrix-mating |
| AF0382/RPA-26[b] | AF1692/ENDO-III | Y2H-screen |
| AF0472/Prp | AF0472/Prp | matrix-mating |
| AF0497/PolB[b] | AF0335/PCNA | matrix-mating |
| AF0699/RecJ-2 | AF0103/HisK-1 | Y2H-screen |
| AF0699/RecJ-2 | AF1194 | Y2H-screen |
| AF0699/RecJ-2 | AF1194 | matrix-mating |
| AF0699/RecJ-2 | AF1332/HisK-2 | Y2H-screen |
| AF0699/RecJ-2 | AF1452/HisK-3 | Y2H-screen |
| AF0735/RecJ-1[a] | AF0382/RPA-26 | matrix-mating |
| AF0780/RPA-36[b] | AF0335/PCNA | matrix-mating |
| AF0780/RPA-36[b] | AF0382/RPA-26 | Y2H-screen |
| AF0780/RPA-36[b] | AF0382/RPA-26 | matrix-mating |
| AF1194[a] | AF0699/RecJ-2 | Y2H-screen |
| AF1194[a] | AF0699/RecJ-2 | matrix-mating |
| AF1194[a] | AF2215/mmCoA-mut | Y2H-screen |
| AF1195/RFC-53[b] | AF0335/PCNA | matrix-mating |
| AF1195/RFC-53[b] | AF0382/RPA-26 | matrix-mating |
| AF1195/RFC-53[b] | AF2060/RFC-35[b] | Y2H-screen |
| AF1195/RFC-53[b] | AF2060/RFC-35[b] | matrix-mating |
| AF1195/RFC-53[b] | AFtransposases | Y2H-screen |
| AF1347 | AF0335/PCNA | matrix-mating |
| AF1347 | AF0472/Prp | Y2H-screen |
| AF1347 | AF1559 | matrix-mating |
| AF1347 | AF2111 | Y2H-screen |
| AF1558/SMC-1[a] | AF0382/RPA-26 | matrix-mating |
| AF1558/SMC-1[a] | AF1559 | Y2H-screen |
| AF1559 | AF0225/braE-1 | Y2H-screen |
| AF1650/fwdBB-1[a] | AF1045/tplC-2 | Y2H-screen |
| AF1692/ENDO-III | AF0382/RPA-26 | matrix-mating |
| AF1692/ENDO-III | AF0382/RPA-26 | Y2H-screen |
| AF1692/ENDO-III | AF1272/purC | Y2H-screen |
| AF1692/ENDO-III | AF1741/pyrE | Y2H-screen |
| AF1692/ENDO-III | AFtransposase | Y2H-screen |
| AF1722/DP-2[b] | AF0335/PCNA | matrix-mating |
| AF1722/DP-2[b] | AF1790/DP-1 | matrix-mating |
| AF1790/DP-1[b] | AF0335/PCNA | matrix-mating |
| AF1790/DP-1[b] | AF1722/DP-2 | matrix-mating |
| AF1790/DP-1[b] | AF1722/DP-2 | Y2H-screen |
| AF2215/mmCoA-mut | AF0472/Prp | matrix-mating |

[a] Proteins included in the screen that had been published to interact with the DP2 DNA polymerase of *P. furiosus* (see Ref. 19). The other proteins were used as baits because they had been identified in screens with the first set of baits. AF2060/RFC-35 was included as a bait in the screens but failed to yield interacting proteins.

[b] The set of baits that were identified by bioinformatics and used for the first round of screens. This set includes the proteins that had previously been annotated as replication factors on The Institute for Genomic Research web site.

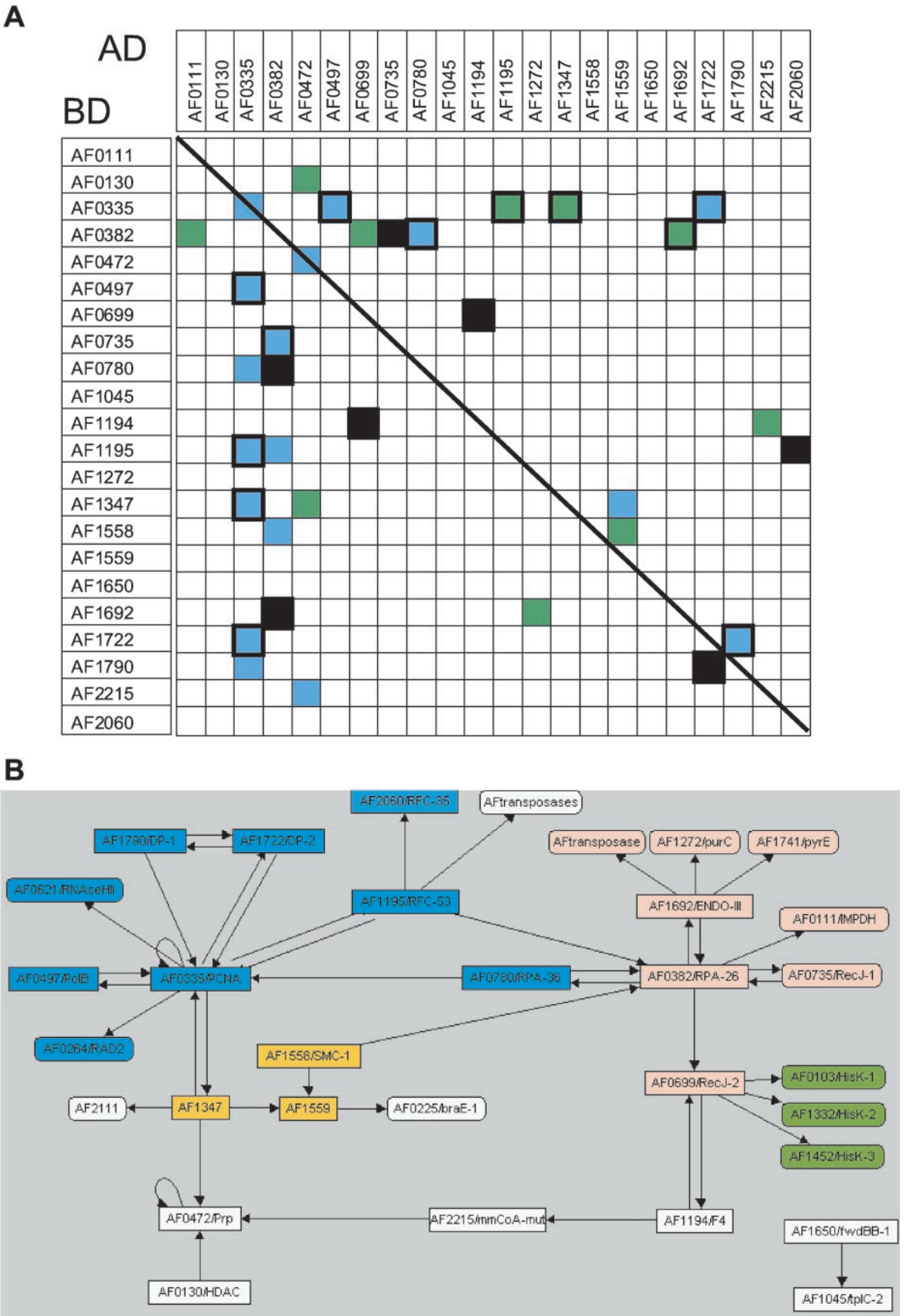


FIG. 1. **Graphical representation of the interactions.** *A*, representation in matrix. Rows show proteins fused to the Gal4 DNA-binding domain; columns show proteins fused to the Gal4 activation domain. Interactions found in screens against the genomic library are represented as *green squares*, interactions detected in the matrix mating are represented in *blue*, and interactions found with both methods are indicated in *black*. Reciprocal interactions, *i.e.* interactions detected irrespective of which partner was fused to the DNA-binding or activation domains, are marked

TABLE II
*Functional assignments to A. fulgidus proteins*

Annotations were taken from The Institute for Genomic Research database of *A. fulgidus*, or are based on bioSCOUT analysis (indicated by an asterisk). Abbreviations used throughout this publication are indicated in parentheses. Proteins 1–22 have been used for screens and direct mating assays. Proteins 23–32 were identified as preys in screens but not used as baits for another round of screening.

| Gene | Function |
|---|---|
| 1. AF0111 | Related to inosine-5′-monophosphate dehydrogenase (IMPDH)* |
| 2. AF0130 | Putative histone deacetylase, complexes with DP-2 in *P. furiosus* (HDAC)* |
| 3. AF0335 | PCNA, sliding clamp to tether polymerases to DNA |
| 4. AF0382 | Replication protein A, 26-kD subunit single-stranded DNA binding complex (RPA-26)* |
| 5. AF0472 | Unknown function |
| 6. AF0497 | DNA polymerase B1 (polB) |
| 7. AF0699 | Putative ssDNA-specific exonuclease, homology to AF0735, RecJ related, (RecJ-2)* |
| 8. AF0735 | recJ related, homology to AF0699, complexes with DP-2, (RecJ-1)* |
| 9. AF0780 | Replication protein A, 36-kD subunit single-stranded DNA binding complex (RPA-36)* |
| 10. AF1045 | Methyl-accepting chemotaxis protein (tlpC-2) |
| 11. AF1194 | Unknown, complexes with DP-2 in *P. furiosus* |
| 12. AF1195 | Replication factor C, 53-kD subunit (RFC-53) |
| 13. AF1272 | Phosphoribosylaminoimidazolesuccinocarboxamide synthase (purC) |
| 14. AF1347 | Unknown function, homology to AF1590* |
| 15. AF1558 | Chromosome segregation protein (smc1), complexes with DP-2 (SMC-1) |
| 16. AF1559 | Unknown function, putative same operon as 1558 |
| 17. AF1650 | Tungsten formylmethanofuran dehydrogenase (fwdB-1), complexes with DP-2 in *P. furiosus* |
| 18. AF1692 | Endonuclease III (Endo-III) |
| 19. AF1722 | DNA polymerase, archaeal type II, large subunit, (DP-2) |
| 20. AF1790 | DNA polymerase, archaeal type II, small subunit, (DP-1) |
| 21. AF2060 | Replication factor C, 35-kD subunit (RFC-35) |
| 22. AF2215 | Methyl-malonyl-CoA-mutase (mmCoA-mut) |
| 23. AF0103 | Tentative homology to histidine kinases (HisK-1)* |
| 24. AF0225 | Branched-chain amino acid ABC transporter, permease protein (braE-1) |
| 25. AF0264 | DNA repair protein RAD2 (RAD2) |
| 26. AF0621 | Ribonuclease HII (RNAseHII) |
| 27. AF1332 | Histidine kinase (HisK-2) |
| 28. AF1452 | Histidine kinase (HisK-3) |
| 29. AF1741 | Orotate phosphoribosyl transferase (pyrE) |
| 30. AF2111 | Function unknown |
| 31. | Transposases (tp1, tp2)) |

AfupolBrev (TTATGCGAATATTCCAG), an initial denaturation step of 180 s at 94 °C was followed by 30 cycles of 30 s at 94 °C, 30 s at 55 °C, and 120 s at 72 °C, and the last cycle was extended for 7 min at 72 °C. By adding now both constructs in a third PCR reaction with the primers Taqfw and AfupolBrev, the total chimeric *Taq* DNA polymerase was assembled. PCR was carried out with an initial denaturation step of 180 s at 94 °C and was followed by 30 cycles of 30 s at 94 °C, 30 s at 55 °C, and 360 s at 72 °C, and the last cycle was extended for 7 min at 72 °C. Cloning into the pQE expression vector (Qiagen), DNA sequencing, and expression (see "Protein Expression") of this construct revealed a functionally active variant with no sequence variations.

*PCR Reactions Using the PCNA Guided Derivative of Taq DNA Polymerase*—PCR reactions showing the stimulating effect of AF0335/PCNA on chimeric *Taq* polymerase were performed as follows. The reactions took place in a total volume of 50 $\mu$l of 10 mM Tris-HCl, 50 mM KCl, pH 8.3, 8% glycerol, and 5 mM $MgCl_2$ at 20 °C, at a deoxynucleotide concentration of 200 $\mu$M each. Primer concentrations were 20 pmol for each primer. Three different amplicons were used for the experiments presented. (i) For amplification of the 463-bp product from pCR 2.1 plasmid DNA, 200 ng of plasmid DNA were employed to the reaction. Primer sequences were AGGGCGTGGTGCGGAGGGCGGT and TCGAGCGGCCGCCCGGGCAGGT, respectively. Cycles were performed as follows: 5 min at 95 °C denaturing, 30 cycles of 30 s at 95 °C, 30 s at 55 °C, and 40 s at 72 °C, and an additional 7 min at 72 °C. (ii) For the amplification of the 500-bp and 1-kb fragment of the human p53 gene from human genomic DNA (Roche Diagnostics, Mannheim, Germany) 200 ng of the genomic DNA were employed for the reaction using the following primers: p53fw1 (CTTGTGCCCTGACTTTCAACTCT) and p53rev1 (CTTTGCACATCTCATGGGGTTAT) (500 bp), and p53fw2 (CTCATCTTGGGCCTGTGTTATC) and p53rev2 TGGTATAAGTTGGTGTTCTGAAGTTAG) (1 kb). For amplification of this construct, an initial denaturation step of 180 s at 95 °C was followed by 30 cycles of 20 s at 95 °C, 30 s at 60 °C, and 60 s at 72 °C, and the last cycle was extended for 7 min at 72 °C. (iii) For the amplification of a 5-kb fragment of the mitochondrial genome, 200 ng human genomic DNA (Roche Diagnostics) were employed in the reaction using the following primers mtDNAfw (AGGAACAACATATGACGCACTCT) and mtDNArev (TAGGTGGCCTGCAGTAATGTTAG). For amplification of this construct, an initial denaturation step of 180 s at 95 °C as followed by 30 cycles of 30 s at 95 °C, 30 s at 55 °C, and 300 s at 72 °C, and the last cycle was extended for 7 min at 72 °C. In all cases 5 $\mu$l of all PCR reactions were subjected to electrophoresis after the cycling on 0.8–1% agarose gels run in 1 × TBE.

*Protein Expression*—After amplification from genomic *A. fulgidus* DNA, expression constructs were cloned in frame into the polylinker of pQE 30 (Qiagen), and sequence integrity was verified by redundant sequencing. Expression was performed in *E. coli* M15 (prep 4) (Qiagen), a strain that overexpresses the LAC repressor and hence efficiently prevents expression of pQE 30 constructs in the absence of an inductor. Transformation, growth, and induction were performed as depicted in the third edition of the QIAexpressionist (Qiagen). Induction by 1 mM isopropyl-1-thio-$\beta$-D-galactopyranoside was performed for 12–18 h at 37 °C. Cells were harvested by centrifugation for 15 min at 5000 × *g*. For the preparation of heat-purified protein extracts, pellets were resuspended in 100 ml/liter culture of 50 mM Tris/Cl, pH 7.9, 50 mM glucose, and 1 mM EDTA, washed once with the same buffer, and then resuspended in 50 ml/l culture of the same buffer supplemented with 4 mg/ml lysozyme. After 15 min at room temperature an equal amount of 10 mM Tris/Cl, pH 7.9, 50 mM KCl, 1 mM EDTA, 0,5% Tween 20, and 0.5% IGPAL was added, and *E. coli* proteins were denatured by incubation at 75 °C for 60 min. Denatured proteins were removed by centrifugation at 27000 × *g* for 15 min, and the supernatant was dialyzed 2 × 8 h against 50 volumes of 10 mM Tris/Cl, pH 7.9, 50 mM KCl, 1 mM EDTA, 50% glycerol, 1 mM dithiothreitol, and 0.5 mM phenylmethylsulfonyl fluoride. Optionally proteins were further purified by Ni-NTA agarose making use of the His tag provided by the pQE vector. Purifi-

by a *bold outline*. Proteins that have not been used as baits are not included as well as their binding partners. *B*, representation as a network. *Double arrows* indicate reciprocal interactions. The group of proteins centered around PCNA, tentatively implicated in DNA replication, are marked in *blue* and the group of proteins tentatively implicated in DNA repair are stained in *pink*. Putative histidine kinases are stained *green*. Proteins that were used in screens are depicted as *normal rectangles* and proteins that have not been used as baits are depicted as *rectangles with rounded corners*.

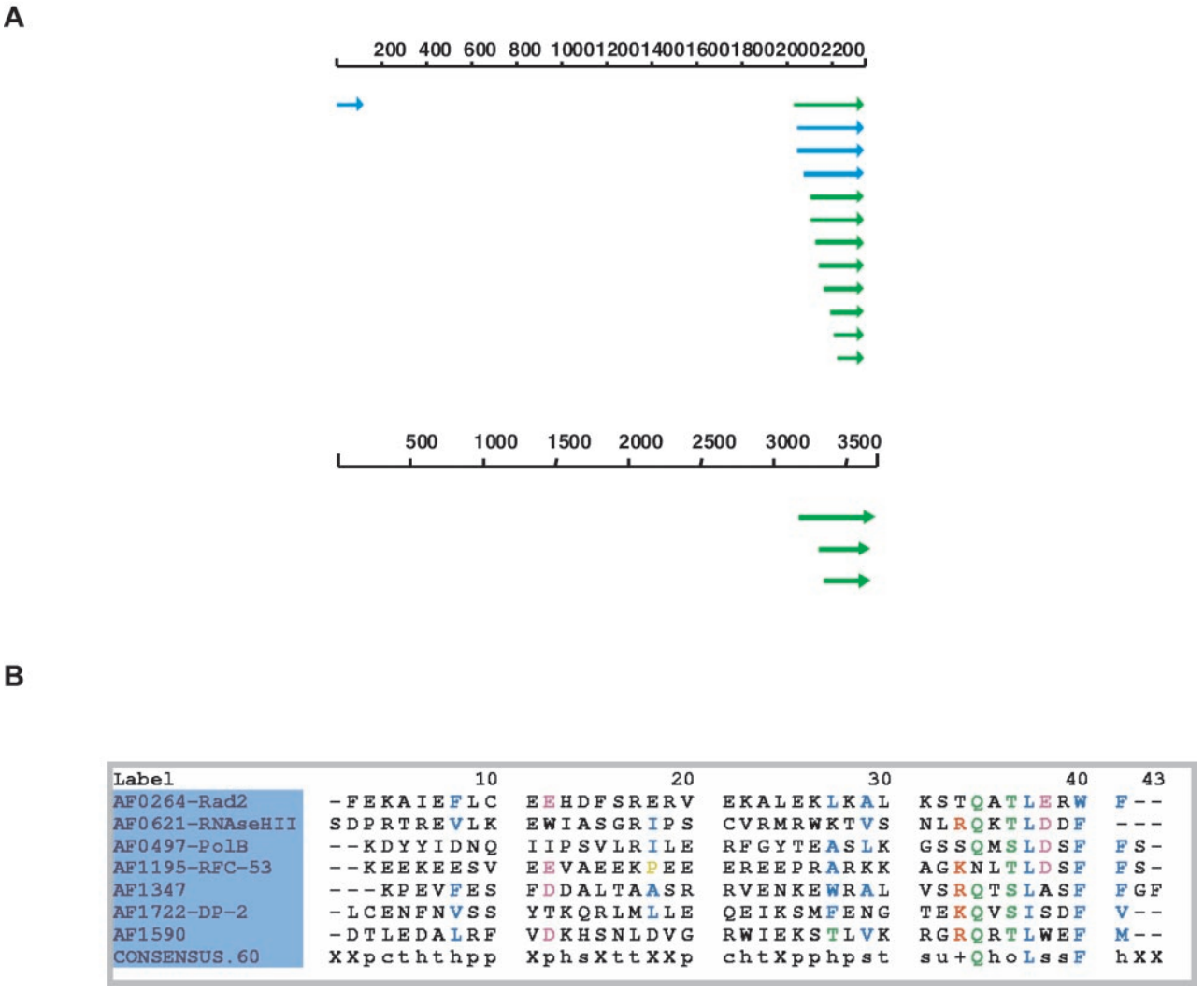


FIG. 2. **The carboxyl terminus of archaebacterial PolB interacts with PCNA.** *A*, a library directing expression of fragments of the PolB proteins of *A. fulgidus* (*top*) or *P. horikoshii* (*bottom*) was screened to isolate the domains that interacted with PCNA from the respective species. *Green arrows* indicate fragments obtained from a library of fragments fused to the DNA-binding domain of Gal4 and *blue arrows* indicate fusion with the activation domain. *B*, the PCNA binding motif of *A. fulgidus* proteins. The last 40 amino acids of the indicated proteins were aligned using the ClustalW program within bioSCOUT. The consensus motif can be found at residues 33–41. A consensus sequence based on a 60% threshold of occurrence is included. Color code: *blue*, aliphatic; *pink*, negatively charged; *red*, positively charged; *green*, polar; and *yellow*, helix breaking. The motif of the protein AF1590 was predicted by homology searches based on the motif present in the other proteins (see "Discussion").

cation was performed according to the instructions of the manufacturer (Qiagen).

## RESULTS

*Elucidation of the Network of Replication Factors in A. fulgidus*—We wanted to adapt archaebacterial proteins involved in DNA-replication for the development of improved PCR protocols. As a first step, we were interested in determining all the cofactors potentially involved in DNA-replication in an archaebacterial genome/proteome. *A. fulgidus* was chosen as a model organism because its genome has been completely sequenced, there are no inteins in *A. fulgidus* proteins, and the patent situation is not prohibitive. First, we identified homologues to proteins known to be involved in DNA replication by an extensive bioinformatic analysis (see "Experimental Procedures" for details). However, bioinformatic predictions of protein functions based on homology are limited in scope because they are extrapolations from existing data and depend on detectable conservation of primary protein sequence. To get a more complete picture of the set of proteins involved in DNA-replication in *A. fulgidus*, we approached the problem from a different angle; mapping out the network of protein-protein interactions surrounding the central proteins of DNA-synthesis should unravel the relationships of known factors and identify novel archaebacterial replication factors. To this end, we made use of the Y2H system.

*Interaction Mapping Strategy*—Our approach consisted of two parts. (i) In a systematic library screening approach, we used predicted replication factors as baits to try and identify interacting proteins in a genomic DNA library. (ii) In a second, independent approach, interactions of all candidate proteins were tested directly in all possible combinations and orientations. This latter strategy of matrix mating turned out to be useful both to detect interactions that were missed in the screens, as well as to allow verification of interactions found in the screen.

For the systematic screening approach, a genomic *A. fulgidus* Y2H library was constructed (see "Experimental Procedures" for details). Bait and prey expression plasmids were generated using a two-step PCR strategy followed by gap repair in yeast cells (17). Each bait was then mated to the yeast cells containing the library, and clones displaying reporter gene activation were isolated. Most steps of the Y2H protocols were

done in microtiterplates, allowing straightforward automation of the method. To further extend the network, identified interactors were in turn used as baits for a second round of screens (see Table I for details). This procedure of "proteome walking" allowed the generation of a local protein interaction map centered around the well known replication factors. Details on which bait was used in the first and second rounds of screening can be found in Table I.

In the matrix-mating approach, full-length versions of the proteins were used as bait and prey and directly tested for interaction in the Y2H system. This was done for all possible combinations of the available proteins. Only protein pairs that caused reporter activation repeatedly in multiple independent trials were considered as interacting partners.

In addition to the use of multiple reporter systems and the retesting of protein-protein interactions in the matrix mating approach, we excluded all clones that were found only once with a given bait, as well as clones that were found with multiple, non-related baits ("sticky" proteins). Only proteins that were found as multiple independent clones with a given bait were considered as potentially interacting proteins.

*Predicted Protein Interactions Involving A. fulgidus Replication Factors*—The combined results of the two independent approaches (screening and matrix-mating) are listed in Table I. The source of information is indicated, as well as the reason as to why the proteins were included as baits. In short, we started with a set of screens including nine putative replication factors. In a second set of screens, we included proteins identified in the first set of screens, as well as proteins that were implicated in DNA replication by homology to proteins binding to the DP2 DNA polymerase of *Pyrococcus furiosus* (19).

A graphical representation of the interactions can be seen in Fig. 1, *A* and *B*. Our efforts resulted in a network connecting 32 protein nodules, 30 of which are part of a single coherent network. Tentative functional annotations are listed in Table II. The majority of proteins are directly or indirectly implicated in DNA-replication. Interestingly, two of the proteins identified in our initial screens (AF1194 and AF0735) have been found in a physical complex with the DP2 DNA polymerase of *P. furiosus* (19). This complex includes three more proteins, which were also used as baits (AF1558, AF0130, and AF1650). Of these, AF1558 and AF0130 both interacted with proteins of the main replication network (see also "Discussion"). For AF1650, however, we did not find additional evidence for its involvement in DNA replication.

*The Binding Motif of PCNA-interacting Proteins in A. fulgidus*—Among the PCNA-binding proteins, we found subunits of archaeal DNA polymerases, a homologue of the replication factor RFC, AF1195/RFC-53, the putative single-stranded DNA binding factor AF0780/RPA-36, and AF0621/RNaseHII as well as AF0264/Rad2. In addition, a novel interactor, AF1347, was identified. Interestingly, AF1347 indirectly binds AF1558/SMC-1 via AF1559. AF1558/SMC-1 has previously been isolated in a physical complex with the DP2 polymerase of *P. furiosus*. We did not see direct binding of AF1558/SMC-1 to the DP2 polymerase. Based on our data, this interaction appears to be indirect, involving a putative complex of proteins including the polymerase AF1722/DP2, AF0335/PCNA, AF1347, and AF1559 as well as AF1558/SMC-1. The AF1558/SMC-1 and AF1559 genes are located next to each other on the *A. fulgidus* genome, separated by only 36 base pairs, and it appears likely that the two genes are part of the same operon and thus functionally related.

Given the diversity of proteins binding to PCNA, we were interested in determining the molecular basis for these interactions. To this end, we isolated PCNA-interacting clones from a library of gene fragments of AF0497/PolB (see "Experimental Procedures" for details). All but one clone correspond to the extreme carboxyl-terminal end of the protein (Fig. 2*A*). Identical results were obtained for the homologous proteins from *Pyrococcus horikoshii* (Fig. 2*B*). Thus, the PCNA-binding sequence is located within the last 50 amino acids of AF0497/PolB.

A conserved PCNA binding peptide can be found in a variety of proteins interacting with PCNA in different species ranging from yeast to man. Inspection of the carboxyl-terminal sequences of the isolated PCNA interacting proteins revealed a conserved nonapeptide similar to the canonical PCNA binding peptide (Ref. 21 and Fig. 2*B*). This motif is present in all PCNA-binding proteins of *A. fulgidus*, apart from AF1790 and AF0780.

A search in the *A. fulgidus* genome using a hidden Markov model based on the alignment of these genes revealed an additional protein that carries the motif within its carboxyl terminus, AF1590. For unknown reasons, we did not find AF1590 in our screens. Remarkably this protein is part of a protein family including the putative PCNA-binding protein AF1347, as well as AF1346. Given its homology to PCNA-binding proteins and its genomic co-localization with AF1347, AF1346 is likely to be functionally involved in DNA metabolism as well. Thus, we propose a novel group of proteins related to DNA replication, consisting of the AF1346, AF1347, and AF1590 family of proteins. This group of proteins seems to be restricted to archaebacteria, and their actual role in replication remains to be elucidated.

*PCR Reactions Using Polymerases of A. fulgidus*—Next, we established PCR reactions using the proteins predicted to be involved in DNA replication in *A. fulgidus*, with the aim to use processivity factors to stimulate the reaction. Open reading frame AF0497/PolB of *A. fulgidus* encodes a family B type DNA polymerase that displays significant homology to the largest subunit of eukaryotic DNA polymerase Δ and to the commercially available PCR enzymes *Pfu* and *Pwo*. Since these two enzymes can perform PCR in the absence of auxiliary proteins, we tested whether their recombinant *A. fulgidus* homologue AF0497/PolB can be employed for the same purpose. Fig. 3*A* shows that in fact AF0497/PolB protein can be successfully used for standard PCR reactions. In contrast, we failed to obtain PCR products with the *A. fulgidus* homologues of the recently described DP2/DP1 polymerase (AF1722 and AF 1790, data not shown). In our hands the only activity assignable to this enzyme was observed in a simpler one-cycle primer extension reaction (data not shown). In this case, both subunits, AF1722 and AF1790, were needed to obtain primer extension activity consistent with their interaction predicted by the Y2H analysis.

*Enhanced PCR-reactions Using Processivity Factors*—Our Y2H data predict that AF0497/PolB binds the sliding clamp AF0335/PCNA. To test whether PCNA can stimulate AF0497 polymerase activity, we analyzed nick-translated DNA or specific primer extension products in a filter-binding assay. Using nicked DNA as template, we observed a significant stimulation of AF0497 polymerase activity by PCNA (Fig. 3*B*). Similarly, we found that in primer extension experiments addition of PCNA not only increased the overall amount of product but also shifted the most prominent products to a higher molecular weight range (Fig. 3*C*). To analyze whether *A. fulgidus* PCNA was able to stimulate polymerase activity in PCR, we employed suboptimal amounts of Af0497/PolB for amplification of a 3.5-kb insert from plasmid DNA. Under standard PCR conditions, PCNA was able to dramatically stimulate polymerase activity (Fig. 3*D*).

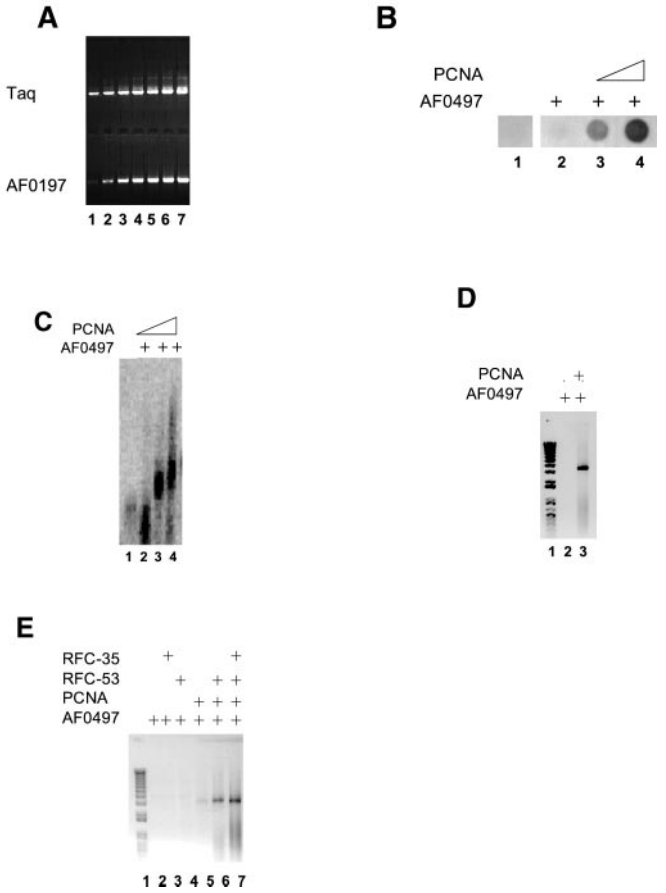


FIG. 3. ***A. fulgidus* replication proteins can be used in PCR.** *A*, PCR activity of AF0497/PolB using M13mp18 DNA as a template was assayed as described under "Experimental Procedures." For the *upper panel* 1 unit of *Taq* polymerase was employed, whereas the *lower panel* depicts PCR activity of 0.2 $\mu$g of a heat-purified fraction of *A. fulgidus* AF0497 in each lane. To compare performance of the two polymerases, 5-$\mu$l aliquots were taken from the reactions after 16, 21, 26, 28, 30, 32, and 34 cycles (*lanes 1–7*) and loaded into two rows of wells of a 1% agarose gel so that after electrophoresis the AF0497 products are located directly beneath the *Taq* products. Previous experiments proved that both polymerases yield products of the same size (not shown). For cycling conditions see "Experimental Procedures." *B*, stimulation of AF0497/PolB activity using a filter-binding assay. A non-radioactive filter-binding assay using activated calf thymus DNA as template was performed as described under "Experimental Procedures." *Dot 1* represents the background of an enzyme-free reaction; for *dots 3* and *4*, 0.4 and 0.8 $\mu$g, respectively, of affinity-purified *A. fulgidus* PCNA were added to the reaction in addition to the polymerase. *C*, stimulation of AF0497/PolB by AF0335/PCNA in primer extension analysis. Primer extension reaction for all lanes was performed with 0.2 $\mu$g of a heat-purified fraction of *A. fulgidus* AF0497. For *lanes 2–4*, 0.2, 0.4, and 0.8 $\mu$g of affinity-purified *A. fulgidus* PCNA were added supplementarily. AF0497 under these conditions produces mainly products of 400–600 nt. In the presence of PCNA, the average size of the products triples, and some products reach a length of more than 4000 nt. *D*, stimulation of AF0497/PolB PCR activity by PCNA. 0.2 $\mu$g of heat-purified AF0497 polymerase were employed in the absence (*lane 2*) or presence (*lane 3*) of 1.6 $\mu$g of affinity-purified *A. fulgidus* PCNA for standard PCR reactions from the AF1722 plasmid template as described under "Experimental Procedures". In *lane 1*, a DNA size marker is shown. *E*, stimulation of AF0497/PolB and AF0335/PCNA PCR activity by clamp loaders. PCR template and cycling conditions were as for *panel D* and are described in detail under "Experimental Procedures." All reactions contained 0.2 $\mu$g of heat-purified AF0497/PolB polymerase, which was supplemented for *lanes 5–7* with 0.6 $\mu$g of affinity-purified *A. fulgidus* PCNA. In addition to polymerase and PCNA, 0.8 $\mu$g of affinity-purified *A. fulgidus* RFC-53 were added to *lanes 6–7*, and the sample presented in *lane 7* was further supplemented by 0.4 $\mu$g of affinity-purified *A. fulgidus* RFC-35. *Lane 3* depicts the effect of 0.4 $\mu$g of affinity-purified *A. fulgidus* RFC-35 on polymerase activity in the absence of PCNA, and the same control is shown for 0.8 $\mu$g of affinity-purified *A. fulgidus* RFC-53 in *lane 4*. In *lane 1* a DNA size marker is shown.

The proteins AF1195/RFC-35 and AF2060/RFC-53 are homologous to subunits of the eukaryotic clamp-loading complex and are predicted to interact with each other as well as PCNA (Fig. 1*B*). Therefore, we included both proteins in primer extension and PCR reactions performed by AF0497/PolB (Fig. 3*E*). Neither RFC-53 alone nor RFC-35 alone had a significant effect on the activity of the polymerase. However, in the presence of PCNA, RFC-35 alone had a dramatic effect on polymerase activity, which was further stimulated by the simultaneous addition of RFC-53. This effect was also apparent in PCR-reactions, where the addition of RFC-35 alone or RFC-53 and RFC-35 together stimulated product generation by the polymerase in the presence of suboptimal amounts of PCNA. In summary, we demonstrate that it is possible to enhance PCR reactions catalyzed by archaeal polymerases by supplementation with the appropriate replication factors.

*Design of a PCNA-guided Derivative of Taq Polymerase and Its Use in PCR*—*Taq* polymerase is one of the most used polymerases for PCR applications. However, despite its robustness and reliability in activity, *Taq* DNA polymerase has the considerable drawback of a high error rate (22) and the lack of the ability to synthesize products larger than 5 kb. We reasoned that it might be possible to enhance the activity of *Taq* by grafting onto it the PCNA binding peptide. The delineation of the PCNA-binding domain in AF0497/PolB allowed us to put this idea to the test. We constructed a fusion protein of *Taq* that has a carboxyl-terminal extension (23) based on the sequence of the AF0497/PolB protein, encompassing the sequence required for binding to PCNA as detailed in Fig. 4*A*. In PCR reaction, this protein was active, albeit slightly attenuated compared with the wild type counterpart (Fig. 4*B*). Addition of PCNA had no effect on wild type *Taq* activity. However, PCNA dramatically stimulated product generation by the hybrid protein. This was evident when plasmid or mitochondrial DNA was used as a template (Fig. 4*C*). We then tested if this increase in product generation was accompanied by an increase in processivity in primer extension experiments. As shown in Fig. 4*D*, the fusion protein had a slightly reduced processivity when compared with wild type *Taq*. The inclusion of PCNA in the reaction lead to an increase in the amount of product and also shifted the products to a higher molecular weight, indicating a gain in processivity. No effect of PCNA on wild type *Taq* could be observed. We then tested if the increased processivity of the PCNA-supplemented PCR reaction allowed us to amplify larger DNA products. Indeed, we observed that the PCNA-stimulated fusion protein was able to PCR out a 5-kb product from genomic DNA, whereas wild type *Taq* was not (Fig. 4*C* and data not shown).

Thus, the fusion of the last 50 amino acids of AF0497/PolB to the carboxyl terminus of *Taq* render the polymerase responsive to PCNA, most likely by tethering it to DNA (24). The success of this naive approach was somewhat surprising since we expected the localization of the domain on *Taq* to be crucial and that a graft at the carboxyl terminus might block enzymatic activity as we initially found for fusion of the PCNA interaction peptide to the amino terminus of *Taq*, which did not result in a functional polymerase (data not shown). We expect that optimization of the construction of similar hybrid proteins with respect to sequence and position of the grafted domain should allow us to significantly extend the capabilities of established PCR enzymes or other DNA modifying enzymes.

## DISCUSSION

*The DNA Replication Protein Network of A. fulgidus*—In this report, we have described the development of PCR protocols aiming at increased polymerase processivity. The selection of protein components for such reactions was based on the eluci-

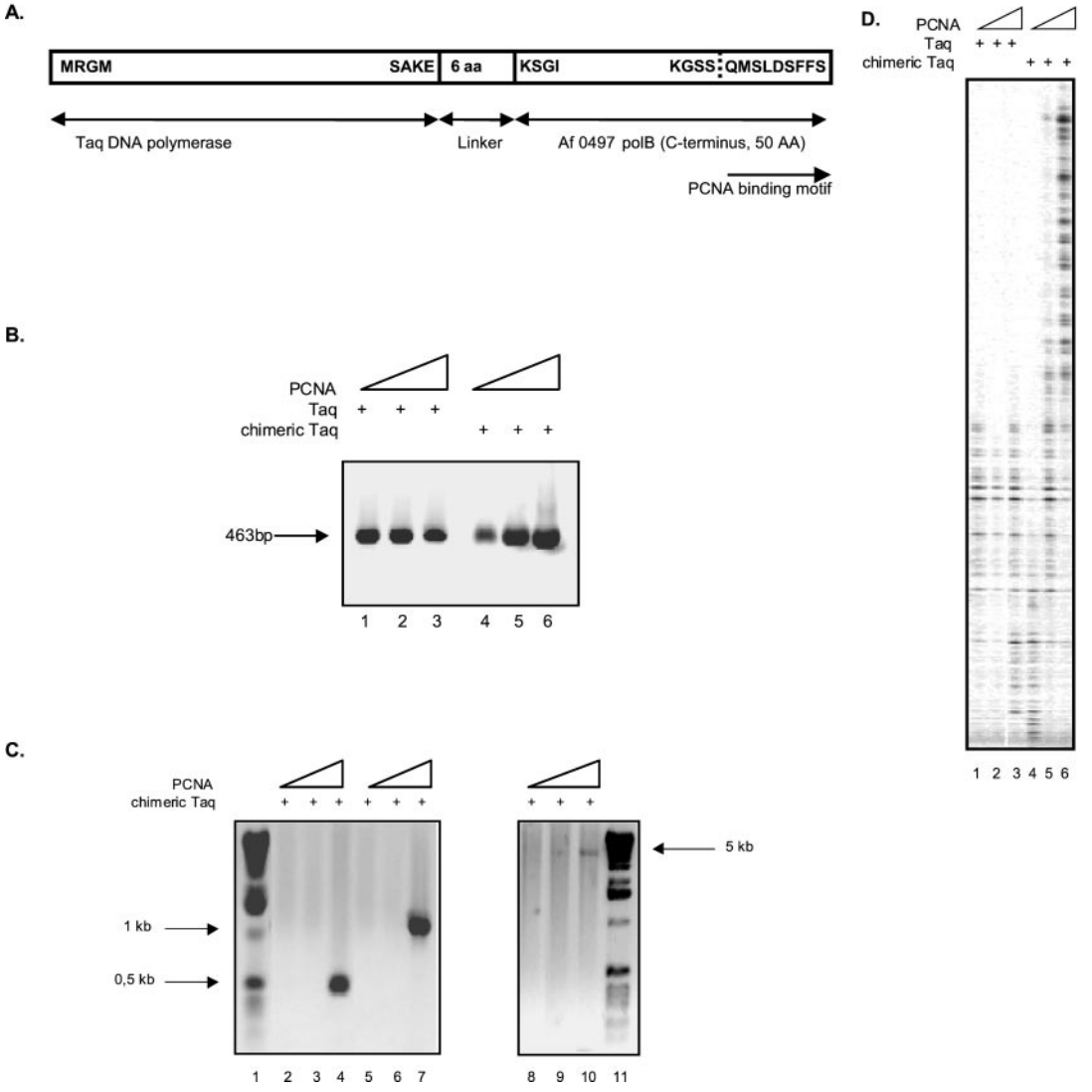


FIG. 4. **Processivity of a chimeric *Taq* can be stimulated by PCNA.** *A*, delineation of the chimeric *Taq* protein comprising the 832 amino acids of *Taq* DNA polymerase, with a linker comprising six amino acids attached at the carboxyl terminus of *Taq* followed by the 50 carboxyl-terminal amino acids of AF0497/PolB. *B*, activity and stimulation of *Taq* DNA polymerase and the chimeric variant in PCR. The activity of *Taq* DNA polymerase in comparison with the chimeric variant of *Taq* DNA polymerase in PCR and the influence of AF0335/PCNA on its activity was investigated. 1 unit of *Taq* DNA polymerase and 0.9 $\mu$g of heat-purified chimeric *Taq* DNA polymerase were comparatively employed for standard PCRs (150 ng of plasmid DNA pCR 2.1 vector), and addition of AF0335/PCNA had no stimulating effect on the wild type variant of *Taq* DNA polymerase (*lanes 1–3*). When applying the chimeric *Taq* DNA polymerase (0.9 $\mu$g) with 0, 0.4, and 0.8 $\mu$g of AF0335/PCNA (*lanes 4–6*), the amount of PCR product increased tremendously. *C*, stimulation of chimeric variant of *Taq* DNA polymerase in PCR on genomic DNA. We employed 0.3 $\mu$g of chimeric *Taq* DNA polymerase with 0, 0.4, and 0.8 $\mu$g of AF0335/PCNA for the amplification of a 0.5-kb fragment (*lane 2–4*) and a 1-kb (*lane 5–7*) fragment of the human p53 gene and 5-kb fragment of a mitochondrial gene (*lane 8–10*). Under standard PCR conditions, AF0335/PCNA was able to dramatically stimulate the activity of the chimeric protein in all cases. Sizes are indicated in kb. *D*, stimulation of chimeric *Taq* DNA polymerase by AF0335/PCNA in primer extension analysis. The primer extension reactions were performed with wild type *Taq* polymerase (1 unit), *lane 1–3*, or 0.9 $\mu$g of a heat-purified fraction of the chimeric *Taq* DNA polymerase (*lanes 4–6*). 0, 0.4, and 0.8 $\mu$g of affinity-purified *A. fulgidus* PCNA were added in *lanes 1–3* and *4–6*, resulting in an increased processivity of the chimeric *Taq* polymerase. The sizes of the products were roughly 90 nt for *Taq* DNA polymerase (*lanes 1–3*) and up to 250 nt for the chimeric *Taq* polymerase supplemented with *Afu* PCNA (*lanes 4–6*).

dation of the network of proteins involved in DNA replication in the archaebacterium *A. fulgidus*. To identify the individual components of this archaeal replication machinery two supplementary approaches were followed. First, proteins involved in DNA replication of *A. fulgidus* were predicted by homology searches based on known eukaryotic and prokaryotic replication proteins employing BLAST and hidden Markov model algorithms. To verify the results of these searches and to additionally identify potential archeal replication proteins without known homologues in other organisms, we performed yeast two hybrid analysis for *A. fulgidus*. To this end, we employed the proteins identified by *in silico* analysis as baits for a genome-wide screen. This approach revealed a network connecting 32 protein nodules, 30 of which are part of a single coherent network. The net contains multiple closed circles, involving 13 of the proteins. Two distinct domains can be considered within this net. One group comprises the PCNA-binding proteins, including the DNA polymerases and the clamp-loading proteins, and the other group centers around the putative single-stranded DNA-binding proteins AF0382/RPA-26 and AF0780/RPA-36 (Fig. 1*B*). Within the latter group, a conspicuous number of proteins with a potential role in DNA metabolism and repair are clustered together, including a putative DNA endonuclease, putative single-stranded DNA exonucleases, and three proteins implicated in nucleotide metabolism. Interestingly, this part of the network is linked via AF0699/RecJ-2 to three proteins with homology to histidine kinases. Proteins of this class play a role in signal transduction processes (20).

Thus, it is tempting to speculate that these signal transducing proteins link DNA replication and repair to external or internal regulatory signals. In summary, we believe that the close clustering of functionally related proteins provides confidence in the relevance of the predicted interactions (18).

However, since sliding clamps are known to be the most important processivity factors and not much is known about their functionality in archaea, we focused our interest on the PCNA-binding proteins and the molecular basis for these interactions. The sequence motif responsible for PCNA binding was found to be located at the carboxyl terminus of AF0497/PolB. An evolutionary conserved PCNA binding motif was found to be present in all PCNA-binding proteins of *A. fulgidus*, apart from AF1790 and AF0780. A search in the *A. fulgidus* genome using a hidden Markov model based on the alignment of these genes revealed an additional protein that also carries the motif within its carboxyl terminus, AF1590 (Fig. 2*B*). Remarkably, this protein is part of a family of sequence-related proteins including the novel PCNA-binding protein AF1347 as well as its genomic neighbor AF1346 (29). Given the presence of the conserved PCNA binding motif in AF1590 and the genomic co-localization of AF1346 with AF1347, this family of proteins is likely to be functionally involved in DNA metabolism. Thus, we propose a novel set of proteins related to DNA replication, consisting of the AF1346, AF1347, and AF1590 group of proteins. This class of proteins seems to be restricted to archaebacteria, and their actual role in replication will be an interesting area of future investigations.

*PCR Reactions Using Archaeal Replication Proteins*—To verify assumed functionalities, we employed the proteins predicted to be involved in DNA replication in *A. fulgidus* for PCR reactions. Again, emphasis was on processivity-related factors and the corresponding polymerases. We found that AF0497/PolB protein, a family B-type DNA polymerase with significant homology to the largest subunit of eukaryotic DNA polymerase Δ and to the commercially available PCR enzymes *Pfu* and *Pwo*, can be successfully used for standard PCR reactions. More unexpectedly, the *A. fulgidus* homologues of the recently described DP2/DP1 polymerase (AF1722 and AF1790) failed in PCR. Although the DP2/DP1 enzyme from *P. furiosus* were shown to work well for PCR (12, 19), our results match the data obtained with *Methanococcus jannaschii* DP2/DP1, for which only very weak primer extension activity could be shown.

Our experiments show that PCNA is able to improve the processivity of the AF0497 DNA polymerase. Surprisingly, this was even true when a closed circular DNA structure was used as a template in the absence of specific clamp loader proteins. Interestingly, similar stimulation of polymerases by archaebacterial PCNA on closed circles of DNA in the absence of clamp loaders has been reported by Cann *et al.* (9). Future experiments will have to reveal the mechanistic reasons for this unexpected phenomenon that stands in contrast to the observations made for eukaryotic PCNA.

The proteins AF1195/RFC-35 and AF2060/RFC-53 are homologous to subunits of the eukaryotic clamp-loading complex. In line with their interaction predicted by the Y2H study, in PCR and primer extension experiments, RFC-35 and RFC-53 had a dramatic effect on polymerase activity in the presence of PCNA. However, in conflict with results obtained for the homologous eukaryotic proteins, this influence was not ATP-dependent. We consider it likely that ATP will have an effect on the activity of the RFC-complex in experimental conditions that have yet to be defined. We would like to speculate that the higher temperatures used for the archaebacterial DNA replication reaction could partially relieve the requirement for energy in this process of protein rearrangement.

*Introduction of Functional Interaction Sites for A. fulgidus PCNA on a Heterologous Polymerase Can Increase Processivity in the Presence of A. fulgidus PCNA—Taq* DNA polymerase is a repair-type eubacterial enzyme. It is one of the most widely used PCR enzymes due to its robust activity, but it lacks the possibility to bind PCNA or any other known processivity factor in its wild type form. Aiming at the improvement of PCR performance of *Taq* polymerase, we have attached the PCNA binding peptide derived from the *A. fulgidus* analysis to the carboxyl terminus of this polymerase. Primer extension experiments show that addition of PCNA to the modified *Taq* polymerase results in an increased processivity. This effect can also be seen in PCR; amplification of a 5-kb product, which could not be yielded with *Taq* alone, was possible with the chimeric system in a PCNA-dependent way. Further optimizations of the chimeric protein with respect to the structural binding properties and the additional use of stabilizing and perhaps even further stimulating replication proteins should allow us to significantly extend the capabilities of processivity-driven PCR enzymes.

Taken together, we show here that the comprehension of cofactor proteins can stimulate PCR reactions catalyzed by both archaeal as well as chimeric prokaryotic DNA polymerases engineered to bind archeal PCNA. We expect that putting to practice the principles we have demonstrated in this report will facilitate the design of better PCR systems.

REFERENCES


1. Keck, J. L. Berger, J. M. (2000) *Chem. Biol.* **7,** R63–R71
2. Waga, S., Stillman, B. (1998) *Annu. Rev. Biochem.* **67,** 721–751
3. Zannis-Hadjopoulos, M., Price, G. B. (1999) *J. Cell. Biochem.* **32,** 1–14
4. Tsurimoto, T. (1998) *Biochim. Biophys. Acta* **1443,** 23–39
5. Mossi, R., and Hubscher, U. (1998) *Eur. J. Biochem.* **254,** 209–216
6. Maki, S., and Kornberg, A. (1998) *J. Biol. Chem.* **263,** 6547–6554
7. Onrust, R., Finkelstein, J., Turner, J., Naktinis, V., O'Donnell, M. (1995) *J. Biol. Chem.* **270,** 13348–13357
8. www.ncbi.nlm.nih.gov/Entrez/Genome/org.html
9. Cann, I. K., Ishino, S., Yuasa, M., Daiyasu, H., Toh, H., Ishino, Y. (1999) *J. Bacteriol.* **181,** 6591–6599
10. Cann, I. K., Ishino Y. (1999) *Genetics.* **152,** 1249–1267
11. Kelman, Z, Hurwitz, J. A (2000) *J. Biol. Chem.* **275,** 7327–7336
12. Uemori, T., Sato, Y., Kato, I., Doi, H., Ishino, Y. (1997) *Genes Cells* **2,** 499–512
13. Cann, I. K., Komori, K., Toh, H., Kanai, S., Ishino, Y. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14250–14255
14. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., Erlich, H. A. (1987) *Science.* **239,** 487–491
15. Lundberg, K. S., Shoemaker, D. D., Adams, M. W., Short, J. M., Sorge, J. A., Mathur, E. J. (1991) *Gene.* **108,** 1–6
16. Barnes, W. M. (1995) *Proc. Natl. Acad. Sci.* **91,** 2216–2220
17. Hudson, J. R., Jr., Dawson, E. P., Rushing, K. L., Jackson, C. H., Lockshon, D., Conover, D., Lanciault, C., Harris, J. R., Simmons, S. J., Rothstein, R., Fields, S. (1997) *Genome Res.* **7,** 1169–1173
18. Walhout, A. J., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., Vidal, M. (2000) *Science* **287,** 116–122
19. Uemori, T., *et al.* (August 1, 1999) Japanese Patent WO99/00506
20. Grebe, T. W., Stock, J. B. (1999) *Adv. Microb. Physiol.* **41,** 139–227
21. Warbrick, E. (2000) *Bioessays* **20,** 195–199
22. Eckert, K. A., Kunkel, T. A. (1991) *PCR Methods Applications* **1,** 17–24
23. Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., Kuriyan, J. (1996) *Cell.* **87,** 297–306
24. Hingorani, M. M., and O'Donnell, M. (2000) *Curr. Biol.* **10,** R25–R29
25. James, P., Halladay, J., and Craig, E. A. (1996) *Genetics* **144,** 1425–1436
26. Bairoch, A., Apweiler R. (2000) *Nucleic Acids Res.* **28,** 45–48
27. Altschul, S. F., Madden, T. L, Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
28. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
29. Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D., Richardson, D. L., Kerlavage, A. R., Graham, D. E., Kyrpides, N. C., Fleischmann, R. D., Quackenbush, J., Lee, N. H., Sutton, G. G., Gill, S., Kirkness, E. F., Dougherty, B. A., McKenney, K., Adams, M. D., Loftus, B. Venter, J., *et al.* (1997) *Nature* **390,** 364–370