Crystal structure of \textit{Pfu}, the high fidelity DNA polymerase from \textit{Pyrococcus furiosus}

Suhng Wook Kim\textsuperscript{a,1}, Dong-Uk Kim\textsuperscript{b,c,1}, Jin Kwang Kim\textsuperscript{d}, Lin-Woo Kang\textsuperscript{d,**}, Hyun-Soo Cho\textsuperscript{b,c,*}

\textsuperscript{a} Department of Clinical Laboratory Science, College of Health Sciences, Korea University, Seoul 136-703, Republic of Korea
\textsuperscript{b} Department of Biology, College of Science, Yonsei University, 134 Shinchon-Dong, Seodaemoon-Gu, Seoul 120-749, Republic of Korea
\textsuperscript{c} Protein Network Research Center, Yonsei University, Seoul 120-749, Republic of Korea
\textsuperscript{d} Department of Advanced Technology Fusion, Konkuk University, 1 Hwayang-dong, Gwangjin-gu, Seoul 143-701, Republic of Korea

Received 22 October 2007; received in revised form 18 January 2008; accepted 18 January 2008
Available online 12 February 2008

Abstract

We have determined a 2.6 Å resolution crystal structure of \textit{Pfu} DNA polymerase, the most commonly used high fidelity PCR enzyme, from \textit{Pyrococcus furiosus}. Although the structures of \textit{Pfu} and KOD1 are highly similar, the structure of \textit{Pfu} elucidates the electron density of the interface between the exonuclease and thumb domains, which has not been previously observed in the KOD1 structure. The interaction of these two domains is known to coordinate the proofreading and polymerization activity of DNA polymerases, especially via H147 that is present within the loop (residues 144–158) of the exonuclease domain. In our structure of \textit{Pfu}, however, E148 rather than H147 is located at better position to interact with the thumb domain. In addition, the structural analysis of \textit{Pfu} and KOD1 shows that both the Y-GG/A and β-hairpin motifs of \textit{Pfu} are found to differ with that of KOD1, and may explain differences in processivity. This information enables us to better understand the mechanisms of polymerization and proofreading of DNA polymerases.

© 2008 Elsevier B.V. All rights reserved.

Keywords: DNA polymerase; Pfu; Crystal structure

1. Introduction

DNA polymerases can be classified into seven main groups based upon sequence homology and phylogenetic relationships; these groups consist of family A (e.g., \textit{E. coli} Pol I), family B (e.g., \textit{E. coli} Pol II), family C (e.g., \textit{E. coli} Pol III), family D (e.g., Euryarchaeotic Pol II), family X (e.g., human Pol β), family Y, and family RT (e.g., reverse transcriptase) [1]. Each DNA polymerase can also be classified into replicative or repairing, as well as error-free or error-prone, based on its own characteristics. At present, many DNA polymerase structures have been determined and their structures are well conserved overall [2–7]. The structural components of DNA polymerases have been generally divided into three separate domains: the fingers, palm, and thumb. Various aspects of DNA replication, such as substrate binding, nucleotide transfer, fidelity, and processivity, have been proposed from the binary/tertiary structures of DNA polymerases in complex with diverse DNA substrates [8]. The fingers and thumb domains change positions depending on whether the polymerase is bound to a substrate [9]. Unbound DNA polymerases form open conformations of the fingers and thumb domains, and when substrate is bound, the two domains move toward the palm domain to hold the template/primer strand more tightly [8].
The existence of additional domains, such as an exonuclease for error-free replication, varies among DNA polymerases. In the case of archaeal DNA polymerases, including *Pfu*, there are five domains that consist of the fingers, palm, thumb, exonuclease, and N-terminal domains [4,5]. The exonuclease domain also changes its conformation depending on the status of DNA replication or editing [5]. When a mismatched nucleotide is incorporated into the newly synthesized DNA strand, the template/primer strand binds to the polymerase more weakly or is misaligned with respect to the polymerase active site. Eventually, the double helix unwinds and the mismatched nucleotide is moved to the active site of the exonuclease domain and is excised [3]. At the DNA-bound, closed conformation of the thumb domain, the 3'-end of the primer strand cannot bind to the exonuclease domain due to steric hindrance, primarily caused by the edge of the thumb domain. Only open conformations of the thumb domain allow the binding of single-stranded primer DNA (ssDNA) to the exonuclease active site. The overall conformations of each domain are tightly coordinated with the others to carry out the DNA replication process.

The Kuroita group previously proposed a novel role for the unique loop of the exonuclease domain, which is exclusively conserved within archaeal DNA polymerases [10]. The unique loop, especially H147, has been shown by site-directed mutagenesis and activity assays to interact directly with the edge of the thumb domain. When H147 is mutated into a glutamate residue, the negative charge of the glutamate reinforces the electrostatic attraction for the positive charge of the thumb domain, shifting the thumb domain 1.5 Å closer to the exonuclease domain. The closer position of the thumb domain to the exonuclease domain prevents the binding of the 3'-end of the ssDNA to the exonuclease active site, explaining the low 3'-5' exonuclease activity of the H147E mutant. The wide-open conformation of the thumb domain is essential for the editing function of the polymerase [10,11]. All characteristics of each DNA polymerase, such as fidelity, processivity and the ability to replicate damaged DNA, can be explained from small differences in the amino acid sequences of local regions which are conserved within archaeal DNA polymerases [10]. The Kuroita group previously proposed a novel role for the exonuclease and thumb domains [10]. But the detailed interaction was not provided because of the unseen edge of the thumb domain, which interacts with the unique loop of the exonuclease domain. Here, we determined the three-dimensional crystal structure of *Pfu*, the family B DNA polymerase from *Pyrococcus furiosus*, which shows clearly the edge in of the thumb domain which was unseen in KOD1 structure. We also found important differences between the *Pfu* and KOD1 structures. This information enables us to better understand the coordinated mechanisms of proofreading and polymerization that can be used to generate a high-performance DNA polymerase for PCR.

2. Material and methods

2.1. Cloning, protein expression, and purification

*P. furiosus* cells were obtained from the Korean Culture Center of Microorganisms (KCCM) and genomic DNA was isolated by standard procedures [14]. Based on the DNA sequence of the *Pfu* DNA polymerase (accession no. D12983), two primers (5'-GGG AGC CAT ATG ATT TTA GAT GTG GAT GAT ATA-3' and 5'-CTA TCG GTC GAC TAG GAT TTT TTA ATG TTA AGC CA-3') were synthesized and used for the amplification of DNA templates of *P. furiosus*. DNA amplification was performed using 2.5 units of *Pfu* DNA polymerase (Stratagene) in a 50 µL reaction volume of PCR buffer, 0.5 µM each primer, 0.2 µM each dNTP, and 0.2 µg genomic DNA. The cycling conditions consisted of an initial denaturation step for 7 min at 95 °C, followed by 35 cycles of 1 min at 95 °C, 1 min at 55 °C, 3 min at 72 °C, and an additional 10 min for final elongation at 72 °C. The PCR products were cloned into the expression vector pET21(b) (Novagen). Clones with the correct orientation were selected and designated as pET21-pfu.

Overexpression and purification of the *Pfu* DNA polymerase were carried out with modifications of methods described previously [15]. The *Pfu* DNA polymerase was expressed in *E. coli* strain BL21(DE3)pLysS carrying the pET21-pfu plasmid. For the seed culture, this bacterial colony was cultured in 10 mL Luria–Bertani (LB) medium containing 0.1 mg/mL ampicillin at 37 °C for 16 h. This seed culture was transferred to 1 L LB medium containing 0.1 mg/mL ampicillin and cultured at 37 °C to an optical density of 0.5 at 600 nm. For induction of protein production, isopropyl-β-D-thiogalactoside (IPTG) was added to the culture at a final concentration of 1 mM and the culture was incubated for 5 h at 37 °C. After protein induction, cells were harvested by centrifugation at 4000 rpm for 15 min at 4 °C. For protein purification, the cell pellet was resuspended in 50 mM Tris–HCl (pH 8.0), 50 mM NaCl, and 1 mM EDTA. After sonication on ice, samples were immediately heated at 75 °C for 30 min and then centrifuged at 16,000 × g at 4 °C for 20 min. The supernatant was loaded onto a DEAE-Sephalac column, equilibrated, and eluted with Tris–HCl (pH 8.0) plus 50 mM NaCl. The flow-through fraction was collected and immediately applied to a HiTrap Heparin HP column (Amersham Biosciences) and equilibrated with Tris–HCl (pH 8.0) plus 50 mM NaCl. The column was eluted by a 50–500 mM NaCl gradient in Tris–HCl (pH 8.0). Each fraction was assayed
by SDS-PAGE and those containing a 90 kDa protein were pooled and concentrated using a Centriprep 50 column (Amicon).

2.2. Crystallization, data collection, and determination of structure

Crystals of Pfu were grown in a mixture of 1.5 μL protein sample and 1.5 μL reservoir solution containing 0.2 M ammonium sulfate, 0.1 M Na-cacodylate (pH 6.5), 5 mM DTT, 50 mM MnCl₂, and 15% (w/v) PEG 8000, and equilibrated against 1 mL reservoir solution over 3–5 days. The X-ray diffraction data were collected to 2.6 Å at the 6B and 4A beamlines of the Pohang Light Source (PLS, Korea). Prior to data collection, a Pfu crystal was soaked in mother liquor with 20% (v/v) glycerol added as a cryoprotectant. Collected data were processed using DENZO and the scale adjusted by the HKL 2000 program.

Data collection statistics from crystallographic analysis are summarized in Table 1. Refined coordinates and structure factors have been deposited in the RCSB Protein Data Bank under the accession code 2JGU.

2.3. Modeling of the closed conformation of Pfu

The model of the DNA-bound Pfu polymerase was based on the structure of DNA-bound Gp43 from bacteriophage RB69 (Protein Data Bank ID 1Q9Y). Each domain of Pfu was separately aligned with the Cα positions of each corresponding domain of Gp43. Alignment of each domain was performed by program CNS [16,17]. After fitting the positions of each domain of Pfu into those of Gp43, the template and primer DNA strands were added into the closed conformation model of Pfu.

3. Results and discussion

3.1. Structure of Pfu polymerase

The 2.6-Å resolution crystal structure of Pfu was determined (Table 1). Pfu is a donut-shaped molecule with overall dimensions of approximately 50 Å × 80 Å × 100 Å. The structure is reminiscent of the canonical structures of other known family B DNA polymerases. A single polypeptide chain of 775 amino acids is folded into five distinct structural domains: the N-terminal domain (residues 1–130, 327–368), the 3′–5′ exonuclease domain (131–326), the palm domain (369–450 and 501–588), the fingers (451–500), and the thumb domain (589–775). The Pfu structure shows an open conformation compared with the editing complex structure of Gp43 from bacteriophage RB69. The fingers and thumb domains are rotated outward by 33° and 24°, respectively, and its overall conformation is similar to that of KOD1 DNA polymerase [4] (Fig. 1A).

3.2. Structure of the interface between the exonuclease domain and the edge of the thumb domain

The structure of Pfu contains a previously disordered electron density in the edge of the KOD1 thumb domain (Fig. 1B). The KOD1 DNA polymerase has as much fidelity as Pfu and even higher processivity. The Kuroita group proposed that the mechanisms of proofreading and polymerization are coordinated from the interactions between the loop (residues 144–158) of the exonuclease domain and the positively charged edge of the thumb domain [10]. The H147 residue was speculated to exist at the tip of the unique loop based on the structure of T7 DNA polymerase in complex with template/ primer DNA substrates. The mutation of H147 into negatively, neutrally, and even higher processivity. The Kuroita group proposed that the mechanisms of proofreading and polymerization are coordinated from the interactions between the loop (residues 144–158) of the exonuclease domain and the positively charged edge of the thumb domain [10]. The H147 residue was speculated to exist at the tip of the unique loop based on the structure of T7 DNA polymerase in complex with template/primer DNA substrates. The mutation of H147 into negatively, neutrally, and even higher processivity. The Kuroita group proposed that the mechanisms of proofreading and polymerization are coordinated from the interactions between the loop (residues 144–158) of the exonuclease domain and the positively charged edge of the thumb domain [10]. The H147 residue was speculated to exist at the tip of the unique loop based on the structure of T7 DNA polymerase in complex with template/primer DNA substrates.
thumb domains. It was found that the higher exonuclease and lower polymerase activities were obtained in KOD1 having a positively charged residue at the position of H147. This result suggests that the disruption of binding between the two domains allows $Pfu$ to have an open conformation that is more suitable for the editing function rather than polymerization.

As suggested, the edge of the thumb domain is located proximally to the unique loop of the exonuclease domain in our crystal structure; however, in the wide-open conformation, there is no direct contact between these domains. Interestingly, from our crystal structure E148 (among residues in the unique loop) is located at a better position at which to have direct contact with the edge of the thumb compared with H147 (Fig. 1C). The distances between the terminal oxygen atoms of the carboxylic group in the E148 residue in the exonuclease loop and the terminal nitrogen atom of the 4-aminobutyl side chain in both the K693 and K695 residues at the edge of the thumb domain are 5.6 Å and 3.7 Å, respectively, in the closest rotamer conformation of each residue (the density of the side chains is not shown in the 2Fo-Fc map). However, E148 does not directly contact any part of the thumb domain. These weak interactions can be explained from the wide-open conformation of this DNA poly-

![Fig. 1. Overall structure of $Pfu$ DNA polymerase and comparison with KOD1. (A) The structure of $Pfu$ consists of five domains, which are exonuclease domain (blue), N-terminal domain (purple), palm domain (orange), finger domain (green), and thumb domain (red). The structure has a wide-open conformation compared with DNA-bound close conformation. Superimposed KOD1 structure is shown as silver. (B) The $2\sigma_F - 2\sigma_C$ (green at 1.0σ) density of $Pfu$ at 2.6 Å resolution is overlaid on the refined model in the region around interface between the unique loop of exonuclease and the edge of thumb domains. (C) The interface between the exonuclease and thumb domain is represented, and key residues are shown as sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)](image)

![Fig. 2. Structural comparison of $Pfu$ with KOD1 and open and close conformations of $Pfu$. (A) Structures of $Pfu$ (green) and KOD1 (orange) at the Y-GG/A motif and the helix are represented as ribbons. Y-GG/A motif and inserted leusine residue are represented as sticks. (B) The DNA-bound structure of $Pfu$, shown as pink, is modeled based on the DNA-bound Gp43 structure from bacteriophage RB69. Template DNA strand is shown as blue; primer DNA strand, red. Crystal structure of open conformation of $Pfu$ is shown as green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)](image)
merase while the closer interactions between the exonuclease domain and the thumb domain are found in the model structure of the DNA-bound Pfu polymerase (Fig. 2B). Based on our crystal structure, H147 appears to affect both editing and polymerization functions indirectly rather than by having a direct contact with the edge of the thumb domain.

3.3. Fidelity and processivity of the Pfu polymerase

Although several structures of archaeal DNA polymerases have been determined, including the structure of the high fidelity KOD1 DNA polymerase, the detailed mechanisms controlling fidelity and processivity of DNA replication are not fully understood. Archaeal DNA polymerases generally have high fidelity; among them, KOD1 and Pfu have the highest fidelity. KOD1 also has the highest processivity and elongation rate. We compared both the structures and sequences of the Pfu and KOD1 DNA polymerases to explain the differences in processivity. Overall sequence identity between the two polymerases was 79.6%, and the RMS deviation between the two structures was 1.56 Å.

Although these two structures are strikingly similar, we found several significant differences that provide clues for the sources of the difference in processivity. The high processivity of KOD1 is explained previously by seven arginine residues at the forked-point that stabilize melted DNA strands for editing [4]. In Pfu, three of the seven arginine residues are replaced with methionine, threonine, and lysine residues. There is the Y-GG/A motif, which plays an important role in processivity and fidelity [19], and the α-helix just in front of the Y-GG/A motif that has extensive contacts with the phosphate backbone of the template DNA strand. At this site, the template DNA strand is bent almost 117° at the reference position of the phosphate backbone. Pfu has an additional leucine residue inserted at the end of the helix, which shifts the sequence one residue toward the C-terminal end (Fig. 3). This shift makes the Y385 residue, the tyrosine residue of Y-GG/A motif, face inside the Pfu enzyme instead of facing outside and the two glycines of the Y-GG/A motif are thus located further away from the DNA backbone (Fig. 2A). Such a conformational change disrupts the conserved interactions between the phosphate backbone of the template DNA and the Y-GG/A motif of the archaean DNA polymerases. Another important difference exists at the β-hairpin motif of the exonuclease domain. The R247 residue in KOD1 is known to bind the penultimate base at the 3′ end of the template/primer for editing. R247 in KOD is replaced by methionine in Pfu, which is not as good an electron donor as arginine and may also slow down polymerization by Pfu.

3.4. Model of DNA-bound conformation of the Pfu polymerase

To mimic the detailed interactions between the exonuclease and thumb domains, the closed Pfu structure is modelled after the DNA-bound Gp43 structure from bacteriophage RB69 (Fig. 2B). The palm domain of Pfu is first aligned with that of Gp43 by superpositioning the conserved residues in the active site. Each of other four domains is aligned separately with its corresponding domain in Gp43, especially based on the structurally robust, conserved secondary structures of each domain. To minimize errors caused by the molecular modeling process, further optimizations were omitted, such as energy minimization of the closed Pfu structure. From the closed conformation model, the unique loop of the exonuclease domain and the edge of the thumb domain are within hydrogen bonding distance as we speculated.

Interestingly, in the closed conformation model a previously unrecognized motif within the exonuclease domain, the β-hairpin (residues 243–248), was found to be located sufficiently near the same edge of the thumb domain to allow direct contact. The amino acid sequence of the β-hairpin motif is also uniquely found in archaean DNA polymerases (Fig. 3). The β-hairpin motif is located at the junction of the template-binding and editing clefts and is known to play a key role in switching the 3′ end of the primer strand between the polymerization and editing active sites for rapid and accurate replication [4]. From our model it is possible that the β-hairpin has direct contacts with the edge of the thumb domain and affects the conformational change of the thumb domain (Fig. 2B). Additional experiments, such as site-directed mutagenesis and enzymatic activity assays, are necessary to clarify the proposed additional role of the β-hairpin motif.

Most DNA polymerase loops in each domain from archaea and thermophiles are shortened compared with those in other species that live either at lower temperatures or in milder environments (e.g., mesophiles and eukaryotes), but the edge of the thumb domain in archaea has retained a long loop length despite evolutionary pressure for thermostability. The conserved edge

![Fig. 3. Sequence alignment among archaeal DNA polymerases. The structure-based sequence alignments of the β-hairpin motif in exonuclease domain and the template DNA interacting α-helix and Y-GG/A motif in palm domain. The positions of M247 and L381 are marked with the reversed triangle. KOD, Thermococcus kodakaraensis DNA polymerase; D. TOK, Desulfurococcus tok DNA polymerase; TGO, Thermococcus gorgonarius DNA polymerase; 9′N-7, Thermococcus sp. DNA polymerase; Pfu, P. furiosus DNA polymerase.](image-url)
of the thumb domain enables close interactions with several different regions of the exonuclease domain, supporting the important coordinated roles of proofreading and polymerization. Taken together, these data help us to better understand the polymerization and proofreading activities of DNA polymerases and can be used to generate more efficient DNA polymerases that exhibit more stringent fidelity during PCR.

Acknowledgments

We thank Drs. Sun-Sin Cha and Kyunghwa Kim for assistance at beamline 6B and 4MX of the Pohang Light Source. This work was supported by Korea Research Foundation grants from the Korean Government (R08-2004-000-10403-02-004), by Korea Science and Engineering Foundation (KOSEF) grant (R112000078010010), by Korea Research Foundation Grant funded by the Korean Government (KRF-2004-005-J04502), and by a grant (Code # 20070501034003) from BioGreen 21 Program, Rural Development Administration, in Republic of Korea.

References


