doi:10.1016/j.jmb.2005.06.015 J. Mol. Biol. (2005) 351, 291–298

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Structural Mechanism for Coordination of Proofreading and Polymerase Activities in Archaeal DNA Polymerases

Toshihiro Kuroita[1]*, Hiroyoshi Matsumura[2], Naohiko Yokota[2] Masao Kitabayashi[1], Hiroshi Hashimoto[3], Tsuyoshi Inoue[2] Tadayuki Imanaka[4] and Yasushi Kai[2]

[1]*Tsuruga Institute of Biotechnology, Toyobo Co. Ltd., 10-24 Toyo-cho, Tsuruga Fukui 914-0047, Japan*

[2]*Department of Materials Chemistry, Graduate School of Engineering, Osaka University 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan*

[3]*Supramolecular Biology International Graduate School of Arts and Technology, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan*

[4]*Department of Synthetic Chemistry and Biological Chemistry, Graduate School of Engineering, Kyoto University Katsura Nishikyo-ku, Kyoto 615-8510, Japan*

**Corresponding author*

A novel mechanism for controlling the proofreading and polymerase activities of archaeal DNA polymerases was studied. The 3′-5′exonuclease (proofreading) activity and PCR performance of the family B DNA polymerase from *Thermococcus kodakaraensis* KOD1 (previously *Pyrococcus kodakaraensis* KOD1) were altered efficiently by mutation of a "unique loop" in the exonuclease domain. Interestingly, eight different H147 mutants showed considerable variations in respect to their 3′-5′exonuclease activity, from 9% to 276%, as against that of the wild-type (WT) enzyme. We determined the 2.75 Å crystal structure of the H147E mutant of KOD DNA polymerase that shows 30% of the 3′-5′exonuclease activity, excellent PCR performance and WT-like fidelity. The structural data indicate that the properties of the H147E mutant were altered by a conformational change of the Editing-cleft caused by an interaction between the unique loop and the Thumb domain. Our data suggest that electrostatic and hydrophobic interactions between the unique loop of the exonuclease domain and the tip of the Thumb domain are essential for determining the properties of these DNA polymerases.

Keywords: archaeral DNA polymerases; proofreading activity; PCR performance; structural change; domain interface

Introduction

Proliferation of cells requires faithful DNA replication by DNA polymerase. Different families of DNA polymerases are involved in different DNA polymerization processes including not only DNA replication but also exonuclease activity. Much is known about the replication of phages,[1] viruses,[2,3] *Procaryota*,[4] and *Eukaryota*.[2,5] Recently the study of archaeal replication, mostly family B DNA polymerases similar to eukaryotic replication enzymes pol α, δ and ε, has been actively pursued[6–9] because biotechnological applications such as high-fidelity PCR require proofreading activity and thermostability.[7]

Among archaeal DNA polymerases, the DNA polymerase from *Thermococcus kodakaraensis* KOD1 is widely utilized because of its high efficiency and extension rate. KOD DNA polymerase exhibits a fivefold higher extension rate (100–130 nucleotides/second) and 10–15-fold higher processivity (>300 bases) than the DNA polymerase from *Pyrococcus furiosus* (*Pfu* DNA polymerase).[7]

It was described that the editing of the DNA polymerase I Klenow fragment can be an intramolecular process or an intermolecular one involving dissociation and reassociation of the DNA.[10] In family B DNA polymerases, including bacteriophage, archaeal and most eukaryotic DNA

Abbreviations used: WT, wild-type; ss, single-stranded.

E-mail address of the corresponding author: toshihiro_kuroita@bio.toyobo.co.jp

polymerases, the control of the two catalytic activities was proposed to proceed intramolecularly, mediated by the flexible loop containing Y-GG/A motif located between the N-terminal 3′-5′exonuclease and the C-terminal polymerase domain.[11,12] In a study of the family B DNA polymerase from *Thermococcus aggregans*, an efficient change of the balance between synthesis and degradation of the DNA chain was achieved by mutations in the Y-GG/A motif,[12] and it was concluded that the conformation and flexibility of the loop might determine the activity balance.

Another mechanism for determination of the 3′-5′exonuclease activity has been suggested. The active center for polymerase activity is located on the antiparallel β-sheet structure of the Palm domain, and the reaction proceeds in an interspace that is formed by the Palm, Fingers, and Thumb domain. The Thumb domain contains the Thumb-1 and Thumb-2 subdomains. The exonuclease active site (Editing cleft (E-cleft)) is located just opposite the Thumb-2 subdomain. A study that structurally compared gp43 from bacteriophage RB69 and the family B DNA polymerase from *Thermococcus gorgonarius* (*Tgo*) DNA polymerase, indicated a structural change at the E-cleft possibly reflecting a transition between open and closed forms.[13] The E-cleft of *Tgo* DNA polymerase is constricted by a displacement of the tip of the Thumb-2 subdomain toward the exonuclease domain to prevent single-stranded (ss) DNA from entering in the E-cleft. The gp43 and *Tgo* DNA polymerases have loop structures that contain different sequences in the exonuclease domain. The metal-binding site (active center of the exonuclease activity) adjoins the loop. In *Tgo* DNA polymerase, the interaction between F152 at the root of the loop and F214 in the ExoII motif permits the formation of a unique shape of the loop, which is curved outward, away from the tip of the Thumb-2 subdomain. The angle of this unique loop enables the tip of the Thumb-2 subdomain to put a lid on the E-cleft. The sequences around the loop are highly conserved in a subset of the family B DNA polymerase, especially in Archaea belong to *Thermococcus sp.* and *Pyrococcus sp*. This loop seemed to coordinate the distance between the two domains.

Additionally, the balance between ′3-5′exonuclease and polymerase activities was altered by mutation of H123 and nearby residues, which are located in the exonuclease domain but not in any of the important motifs in T7 DNA polymerase, a family A DNA polymerase.[14] One possible explanation given here is that the region in the vicinity of H123 primarily has a structural role for the exonuclease active site. However, this mechanism remains unsolved and has not yet been reported in family B DNA polymerases.

The first crystal structure of a family B DNA polymerase to be obtained was that of bacteriophage RB69 DNA polymerase.[15] The first crystal structure of archaeal DNA polymerase was *Tgo* DNA polymerase.[13] Further crystal structures of archaeal family B DNA polymerases have recently been reported: Tok DNA polymerase from *Desulfurococcus sp*. Tok,[16] 9°N-7 DNA polymerase from *Thermococcus sp.* 9°N-7,[17] and KOD DNA polymerase.[18] Also, the crystal structure of the family A DNA polymerase from bacteriophage T7 has recently been reported.[19] The 3′-5′exonuclease domain of the family B DNA polymerase is structurally equivalent to that of the family A DNA polymerase, however, it is bound at the opposite side of the polymerase unit by non-covalent contacts to the Thumb domain at the E-cleft, on one side, and by covalent and non-covalent contact to the N-terminal and Palm domain, on the other side.

Here, we present data on mutations in the unique loop of the exonuclease domain away from the Thumb-2 subdomain of KOD DNA polymerase. Further, we demonstrate the structural mechanism for the coordination of proofreading and polymerase activities in DNA polymerase.

Results and Discussion

Site-directed mutagenesis of the loop

H123 of T7 DNA polymerase, described in Introduction, is estimated to be located in a loop facing toward the tip of the Thumb domain according to a 3-D structure model obtained using a polymerase lacking six residues (118–123).[19] The E-cleft is located just opposite the tip of the Thumb domain. Similarly, the 3-D structure of the family B

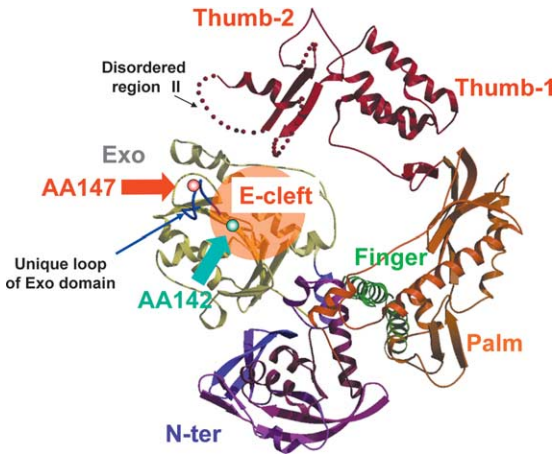

Figure 1. Overall structure of KOD DNA polymerase. The structure is composed of domains and subdomains, namely the N-terminal domain (N-ter, violet), exonuclease domain (Exo, gray) including the unique loop (blue), polymerase domain (Pol) including the Palm (brown) and Fingers (green) subdomains, and the Thumb domain (red), including the Thumb-1 and Thumb-2 subdomains and disordered regions (dotted lines). The orange region represents the E-cleft. The positions of two amino acids in the exonuclease domain are shown as colored balls (red, AA147; light blue, AA142).

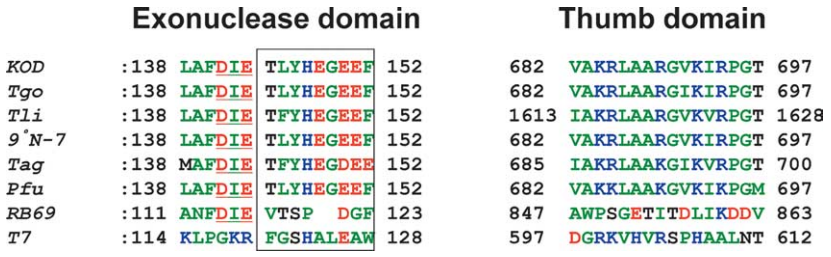

Figure 2. Amino acid sequence alignments of proposed structurally corresponding residues of KOD DNA polymerases (138–152, exonuclease domain; and 682–697, Thumb domain). Acidic, basic and hydrophobic residues are shown in red, blue and green, respectively. The boxed amino acid residues comprise the loop structure in the exonuclease domain. The underlined amino acid residues are the core sequence of the Exo I motif. The abbreviations used are as follows: *KOD*, *Thermococcus kodakaraensis* DNA polymerase; *Tgo*, *Thermococcus gorgonarius* DNA polymerase; *Tli*, *Thermococcus litoralis* DNA polymerase; *9°N-7*, *Thermococcus* sp. *9°N-7* DNA polymerase; *Tag*, *Thermococcus aggregans* DNA polymerase; *Pfu*, *Pyrococcus furiosus* DNA polymerase; *RB69*, bacteriophage RB69 DNA polymerase; *T7*, bacteriophage T7 DNA polymerase.

DNA polymerase from *T. gorgonarius* reveals that a loop toward the edge of the Thumb domain can play an important role in the conformational change of the two domains that is essential for switching between the "editing mode" and the "polymerase mode".[13] The 3-D structure of KOD DNA polymerase implies that movement of the loop is required to interact with the primer bound at the E-cleft.[18] This loop is located adjacent to the Exo I motif and contains a histidine residue at the tip of the loop.[13,17,18] Figure 1 shows the three-dimensional relationship between the unique loop and other domains of KOD DNA polymerase. As shown in Figure 2, the loops, which are presumed to be facing toward the Thumb-2 subdomain, are characterized by negatively charged and hydrophobic amino acids. In particular, the sequences of the loops are highly conserved among the family B DNA polymerases from *Thermococcus sp*. and *Pyrococcus sp.*, while T4 (data not shown) and RB69 DNA polymerases lack this loop. The family B DNA polymerases (e.g. KOD DNA polymerase, *Pfu* DNA polymerase, Vent DNA polymerase and Deep Vent DNA polymerase) from *Thermococcus sp*. and *Pyrococcus sp.* are the DNA polymerases that have commonly been used as high-fidelity PCR enzymes.

On the other hand, the edges of the Thumb-2 subdomain of these DNA polymerases are characterized by positively charged and hydrophobic amino acids. As shown in Figure 1, the edge of the Thumb-2 subdomain facing the unique loop is estimated to correspond to disordered region II in KOD DNA polymerase. Interestingly, the edge of the Thumb domain (^{597}DGRKVHVRSPHAALNT) of T7 DNA polymerase, a family A DNA polymerase, is also characterized by positively charged and hydrophobic amino acids. These observations suggest the possibility that electrostatic and hydrophobic interactions between the loop of the exonuclease domain and the tip of the Thumb-2 subdomain are essential for determining the properties of these DNA polymerases. Therefore, we started by exchanging H147 to basic amino acids (Lys and Arg), acidic amino acids (Glu and Asp) and neutral amino acids (Ala, Tyr, Ser and Gln). For a comparison with exonuclease-deficient mutants,

Table 1. Comparison of the polymerase activities, 3′-5′exonuclease activities, PCR fidelities and elongation rates of wild-type and mutated KOD DNA polymerases

		Relative polymerase activity (%)[a]		Exo/Pol			
DNA polymerase	Relative exonuclease activity (%)[a]	Calf thymus DNA	Primed M13 DNA	Calf thymus DNA	Primed M13 DNA	Mutation frequency (%)	Elongation rate (bp s^{-1})
H147K	276	69	93	4.0	3.0	0.12	133
H147R	234	78	99	3.0	2.4	0.17	131
WT	100	100	100	1.0	1.0	0.47	130
H147Y	124	138	195	0.90	0.64	0.39	N.D.[b]
H147Q	66	131	N.D.[b]	0.50	–	N.D.[b]	N.D.[b]
H147S	60	119	N.D.[b]	0.50	–	N.D.[b]	N.D.[b]
H147A	41	136	149	0.30	0.28	0.45	N.D.[b]
H147E	30	120	160	0.25	0.19	0.47	133
I142Q	24	97	114	0.25	0.21	0.47	N.D.[b]
H147D	9	155	210	0.06	0.04	0.65	132
I142K	7	110	118	0.06	0.06	0.95	127
Taq	N.D.[b]	N.D.[b]	N.D.[b]	–	–	7.91	N.D.[b]
Pfu	N.D.[b]	N.D.[b]	N.D.[b]	–	–	1.27	20

[a] Relative activity to the wild-type enzyme.
[b] N.D., not determined.

Exo I motif mutants (I142Q and I142K)[20] of KOD DNA polymerase were prepared and examined.

The properties of the mutated KOD DNA polymerases are summarized in Table 1. The polymerase activities of the WT and mutated enzymes were measured using activated calf thymus DNA and primed M13 DNA. The analysis revealed that the polymerase activity of each enzyme did not change dramatically, although H147D showed an increase in activity to above 150% (calf thymus DNA) and 200% (primed M13 DNA). The exonuclease activities of the WT and mutated enzymes were assayed using ^{3}H-labeled *Escherichia coli* DNA. Regarding the exonuclease activity, the mutants were divided into three groups: (i) mutants having WT-like exonuclease activity (50–150%): H147Y, H147Q and H147S; (ii) mutants with enhanced exonuclease activity (>150%): H147K and H147R; and (iii) mutants with decreased exonuclease activity (<50%): H147D, H147E, H147A, I142Q and I142K. In the H147 mutants, basic amino acids at position 147 increased the 3′-5′exonuclease activity, while acidic amino acids decreased the activity. On the other hand, it has been demonstrated that both basic and acidic amino acids decrease the 3′-5′exonuclease activity of Ile142 mutants.[20] Increases in the 3′-5′exonuclease activity have not been reported after mutagenesis of the core sequences of the exonuclease motifs. This suggests that the charge of the amino acid at position 147 is essential for regulating the 3′-5′exonuclease activity of family B DNA polymerases.

These results for KOD DNA polymerase have some analogy with those for T7 DNA polymerase. In the site-directed mutagenesis of residue 123 in T7 DNA polymerase, substitution to an acidic amino acid (Glu) decreased the 3′-5′exonuclease activity more than substitution to a neutral amino acid (Ser). Moreover, it was also confirmed that the substitution affected the 3′-5′exonuclease activity more than the polymerase activity.[14]

The elongation rate of each enzyme was measured using primed M13 DNA. As shown in Table 1, all the mutants showed a WT-like elongation rate. The mutation frequency of each enzyme was measured using a PCR-based fidelity assay. Mutants were divided into the following groups: (i) mutants showing WT-like mutation frequency (0.2–0.5%): H147Y, H147A and H147E; (ii) mutants with decreased mutation frequency (< 0.2%): H147K and H147R; and (iii) mutants with increased mutation frequency (>0.5%): H147D. The mutation frequency of each mutant showed a correlation with the 3′-5′exonuclease activity to some extent. Even H147D showed a higher fidelity (12-fold) than that of Taq DNA polymerase.

PCR efficiency of the mutants

It has been indicated that insufficient template DNA can sometimes cause poor amplification that produces an indistinct non-specific band, especially for PCR of large targets of over 2 kb using KOD DNA polymerase. The strong 3′-5′exonuclease activity of KOD DNA polymerase has been thought to be the cause of the poor efficiency and the indistinct non-specific band.

At first, a fragment of the human β-globin gene (3.6 kb) was amplified from different concentrations of human genomic DNA (final concentrations 2 ng/μl and 0.2 ng/μl) by each mutated enzyme. Under the high template DNA condition (2 ng/μl), each mutant showed a distinct band at the expected position upon gel analysis (data not shown). The change in template concentration from 2 ng/μl to 0.2 ng/μl greatly increased the frequency of failed reactions. Only four mutants (i.e. H147D, H147E, H147Y and H147A) resulted in successful amplification. Although I142K also showed a faint band, conspicuous unexpected bands were amplified at the same time. The other mutants generated only indistinct non-specific bands (Figure 3(a)). This experiment indicates that the 3′-5′exonuclease activity is not the only cause of PCR failure, because some mutants exhibiting similar Exo/Pol ratios (e.g. H147E and I142Q) produced different results. From these experiments, it is concluded that the negative charge or hydrophobicity of the amino acid at position 147 plays an important role for the sensitivity of PCR.

Next, the mutants that showed successful amplification in the above experiments (H147D, H147E, H147Y and H147A) were applied to "long PCR". A DNA fragment of the myosin heavy chain (6.2 kb) was amplified from human genomic DNA (final concentration, 1 ng/μl). As shown in Figure 3(b), H147D and H147E successfully amplified 6.2 kb products. The yield with H147D was higher than that with H147E. The target was not amplified by H147Y and H147A. PCR with the other mutants and the WT enzyme also ended in failure (data not shown). These results indicate that a negative charge at residue 147 of KOD DNA

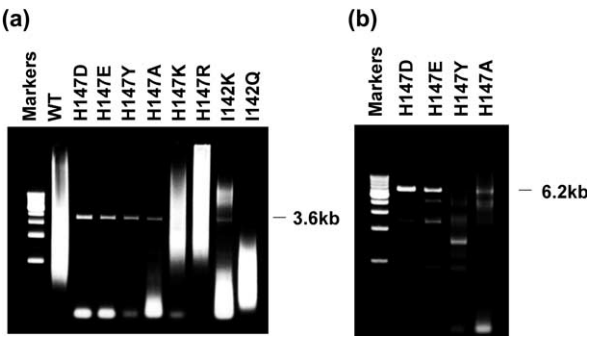

Figure 3. PCR with KOD polymerase mutants. (a) Agarose gel (1%) showing 3.6 kb PCR products. One unit of each mutant or WT enzyme was added to a mixture of 10 ng human genomic DNA and a primer pair designed to yield a 3.6 kb DNA fragment. (b) Long PCR with each mutant. One unit of each mutant was added to a mixture of 50 ng human genomic DNA and a primer pair designed to yield a 6.2 kb DNA fragment.

polymerase plays an essential role for efficient amplification of a long target. Moreover, these results suggest that interaction between the two domains through the intermediary of the loop determines the PCR performance.

Structural change of the exonuclease-Thumb interface

To confirm the interaction between the exonuclease domain and the Thumb-2 subdomain, the crystal structure of the H147E mutant of KOD DNA polymerase was determined at 2.75 Å resolution and compared with that of the WT enzyme. H147E is a clone that shows higher PCR efficiency than the exonuclease active site-deficient mutant (I142Q), which exhibits similar 3′-5′exonuclease activity to the H147E mutant and has identical fidelity to the WT enzyme. It is thought that the structures of KOD DNA polymerases described here provide information for the DNA-free, most relaxed conformations.

As shown in Figure 4, three disordered regions (disordered region I, R667-G677; disordered region II, R689-696; and disordered region III, K705-D712) were observed at the tip of the Thumb-2 subdomain in the H147E mutant as well as in the WT enzyme. Disordered region II is presumed to be able to interact with the loop of the exonuclease domain. The average *B*-factor of the Thumb domain (588–758) of the H147E mutant was similar to that of the WT enzyme (WT, 99.4; H147E mutant, 91.4). The *B*-factors of protein crystal structures reflect the fluctuation of atoms about their average positions and provide important information about protein dynamics. The slight decrease in the *B*-factor of the H147E mutant's Thumb domain might indicate a measure of stabilization for the Thumb domain. The relative shift of the tip of the Thumb-2 subdomain (helical region (678–688)) toward the loop was estimated to be 1.5 Å. On the other hand, structural change was not observed in the main-chain of the exonuclease domain (r.m.s. deviation 0.59 Å for the C^{α} of 196 residues), and the structures of the side-chains of the active center of the 3′-5′exonuclease domain were very similar (shown in Figure 4). In the H147E mutant, the decrease in the 3′-5′exonuclease activity is estimated to be caused by a constriction of the E-cleft depending on displacement of the tip of the Thumb-2 subdomain toward the exonuclease domain that prohibits ssDNA binding.

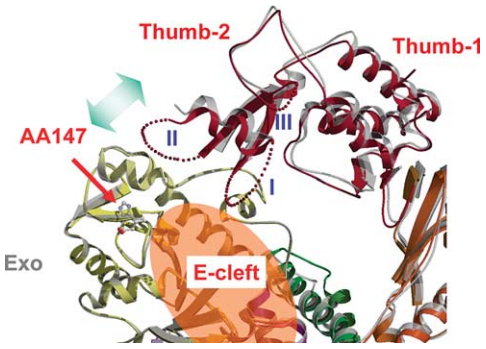

Figure 4. Comparison of the three-dimensional structures of the H147E mutant and wild-type KOD DNA polymerase. The main-chains of the exonuclease domains of the WT enzyme and H147E mutant are overlaid. Gray, WT KOD DNA polymerase; colored in red and yellow, H147E mutant; dotted line, assumed disordered regions in the H147E mutant (the corresponding regions of the WT enzyme are not shown). AA147 is shown as a ball-and-stick model. In H147E mutant, the Thumb domain is estimated to move 1.5 Å closer to the exonuclease domain.

The movement of the Thumb-2 subdomain toward the exonuclease domain side seems to be led by an interaction between the negatively charged residue (E147) and the tip of the Thumb-2 subdomain, which is characterized by positively charged and hydrophobic residues. Similar effects are analogized about the mutants that have negatively charged or hydrophobic amino acids at position 147. On the other hand, opposite results are expected for the mutants that have a positively charged amino acid at residue 147. These conformation changes can be connected to the so-called "closed" and "open" conformations of the E-cleft. It has been reported that the Thumb domain of KOD DNA polymerase is shifted to produce an open conformation and that a portion of the Palm domain neighboring the root of the Thumb domain is slightly shifted as a result of the large movement of the Thumb domain in comparison to other archaeal DNA polymerases.[18] Our results suggest that the mechanisms of the changes in the enzymatic properties of H147 mutants are fundamentally different from those of active site-deficient mutants (i.e. I142 mutants). For the active site-deficient mutants, ssDNAs, which need to be proofread, have to undergo extra movement on average compared with the H147 mutants. This extra movement can be inhibitory for PCR. For long PCR in particular, the mutants having an acidic amino acid at position 147 (i.e. H147D and H147E) are thought to have an advantage over the other mutants, because of their reduced E-cleft. It would appear that the active site-deficient mutants are not accompanied by structural changes like the H147 mutants.

A correlation between the unique loop and the Thumb domain has been reported.[13] Here, we demonstrated the structure-function relationships in the family B DNA polymerase from *T. kodakaraensis* KOD1. Simultaneously, we reported the same observations for T7 DNA polymerase, which is categorized as a family A DNA polymerase. Interestingly, DNA polymerases that have the loop characterized by acidic and hydrophobic amino acids always have basic and hydrophobic amino acids at the tip of the Thumb-2 subdomain. This suggests that electrostatic and hydrophobic interactions between the loop and the tip of the Thumb domain also play important roles in

determining the enzymatic properties of WT DNA polymerases. The evolution of the interaction between the two domains is interesting because some family B DNA polymerases (e.g. RB69) lacking the loop have no basic or hydrophobic amino acids at the tip of the Thumb domain. We can consider the possibility that this kind of structural relationship between the unique loop and the Thumb domain in a subset of the family B DNA polymerases has evolved in parallel with that in several family A DNA polymerases containing T7 DNA polymerase independent of the Y-GG/A motif, although the sequences and overall structures of the two groups are reasonably different. Furthermore, it is likely that the histidine residue in the unique loop has evolved as the most appropriate residue to regulate the overall performance of the enzyme. Moreover, it is interesting to discover that such useful polymerases have obtained this structural relationship.

The conformational change in the Thumb domain can also affect the properties of the polymerase activity since some mutants showed increased polymerase activity, especially on the primed M13 DNA (H147D, H147E, H147A and H147Y). Newly synthesized duplex DNA is grasped by the polymerase and Thumb domains. The tip of the Thumb-2 subdomain of KOD DNA polymerase is estimated to be flexible. The average *B*-factor of the Thumb domain of KOD DNA polymerase is markedly higher than the other domains.[18] The Thumb domains of H147 mutants that showed improved PCR performance might be stabilized by the exonuclease-Thumb interaction because the rates of increase in polymerase activity on primed M13 DNA polymerase for H147D and H147E were more significant than those for I142K and I142Q, which showed similar 3′-5′exonuclease activities.

This theory suggests the possibility that some mutation in the unique loop or Thumb domain enables the generation of a more useful DNA polymerase exhibiting high efficiency and excellent fidelity. Furthermore, methods for altering the structure of a domain interface, that includes the active center or binding site, can be important strategies for the improvement of enzymes.

Materials and Methods

Purification and mutagenesis of KOD DNA polymerase

A QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used to introduce mutations into the KOD DNA polymerase gene. Enzymes were overexpressed in *E. coli* BL21(DE3) and purified using a previously reported method.[7] The purification involved a heat treatment step and chromatography on heparin-Sepharose. All protein samples were >95% pure, as determined by Coomassie blue-stained SDS-PAGE.

Assay for DNA polymerase activity and exonuclease activity

DNA polymerase activity was measured by a TCA precipitation assay.[21] The reaction mixture contained 20 mM Tris–HCl (pH 7.5), 8 mM $MgCl_2$, 50 μg ml^{-1} BSA, 0.15 mM each dNTP, [*methyl*-^{3}H]TTP (0.13 mCi $nmol^{-1}$ final concentration), 150 mg ml^{-1} activated calf thymus DNA, and 7.5 mM dithiothreitol. One unit of activity was defined as the amount of enzyme required to incorporate 10 nmol of dNTP into an acid-insoluble fraction at 75 °C in 30 minutes. For DNA polymerase activity assays with single-stranded (ss) DNA, the reaction mixture contained 20 mM Tris–HCl (pH 7.5), 8 mM $MgCl^2$, 50 μg ml^{-1} BSA, 0.15 mM each dNTP, [*methyl*-^{3}H]TTP (0.13 mCi $nmol^{-1}$ final concentration), M13 ssDNA primed with the P7 primer (5′-CGCCAGGGTTTTCCCAGTCACGAC-3′), and 7.5 mM dithiothreitol.

Uniformly ^{3}H-labeled *E. coli* DNA was used as a substrate for measurement of the exonuclease activity.[20] Reaction mixtures for the exonuclease assay (50 μl) contained 120 mM Tris–HCl (pH 8.0), 1.2 mM $MgCl_2$, 6 mM $(NH_4)_2SO_4$, 10 mM KCl, 0.1% (v/v) Triton X-100, 0.01% (w/v) BSA, 5 μg of ^{3}H-labeled DNA (7.4 kBq μg^{-1}) and serially diluted enzymes. After incubation at 75 °C for ten minutes, the reactions were stopped by the addition of 50 μl of BSA and 100 μl of 10% trichloroacetic acid. After incubation on ice for ten minutes, the precipitated DNA was collected by centrifugation at 12,000*g* for ten minutes. The acid-soluble radioactivity was measured in 100 μl of the supernatant. The relative exonuclease activity of each enzyme was estimated using the radioactivity counts of the serially diluted enzymes.

The elongation rate of DNA polymerase was calculated from the length of DNA synthesized in a fixed time (30 s, 60 s or 90 s) using M13 ssDNA primed with the P7 primer (5′-CGCCAGGGTTTTCCCAGTCACGAC-3′) as a substrate.[20] The reaction mixture for the elongation rate assay (20 μl) contained 20 mM Tris–HCl (pH 8.8), 2 mM $MgSO_4$, 6 mM $(NH4)_2SO_4$, 10 mM KCl, 0.1% Triton X-100, 0.1 mg ml^{-1} BSA, 0.4 μg M13 ssDNA, 4 pmol P7 primer, 0.2 mM dATP, dGTP and TTP, 0.1 mM [α-^{32}P]dCTP (1.11 MBq), and 1.25 units of DNA polymerase. The mixtures were incubated at 75 °C for various times. The reactions were stopped by the addition of 20 μl of 60 mM EDTA, 60 mM NaOH and 0.1% (w/v) bromophenol blue. The samples were electrophoresed in a 1% (w/v) agarose gel containing 50 mM NaOH and 1 mM EDTA, and then visualized by autoradiography.

The mutation frequency rates of the DNA polymerases on PCR were measured by the positive selection system using the *rpsL* gene as reported.[20] Twenty-five cycles were carried out as follows: 20 s at 99 °C, 2 s at 65 °C, and three minutes at 68 °C. The PCR products included the full-length DNA (4.0 kbp) from the plasmid pMol21,[22] carrying the *bla* gene for ampicillin resistance and the *rpsL* gene for streptomycin sensitivity.

Polymerase chain reaction

The reaction mixture (50 μl) consisted of 1 mM $MgSO_4$, 200 mM dNTPs, 0.3 μM each primer, one unit of enzyme, 10 ng, 50 ng or 100 ng of human genomic DNA and buffer for KOD-Plus (TOYOB, Japan). For the cycling reaction, a GeneAmp 2400 (Perkin-Elmer) was used. The thermal conditions were two minutes at 94 °C, followed by 35 cycles of 15 s at 94 °C, 30 s at 55 °C and one minute/kb at

68 °C. The primers were as follows (5′to 3′): bg-F: GGT GTTCCCTTGATGTAGCACA; bg-R: ACATGTATTTGC ATGGAAAACAACTC; my-F: AGTGCTTCGTGCCCGA TGAC; and my-R: TGCCCCTTGGTGACATACTCG.

Crystallization, data collection and structure determination

Crystals were grown by the hanging-drop vapor-diffusion method[23] at 20 °C. Drops containing 2 μl of protein solution at 6.0 mg ml^{-1} in 50 mM Tris–HCl (pH 8.0), 0.1 mM EDTA, 1.0 mM DTT, 0.001% (v/v) Tween 20, 50% (v/v) glycerol, and 2 μl of precipitating buffer were equilibrated against 500 μl of precipitating buffer containing 3% (w/v) polyethylene glycol 1450, 10 mM $NiCl_2$, and 17.5% glycerol. Crystals took one day to appear and up to three days to grow to a size suitable for diffraction analysis. The crystals belonged to space group $P2_12_12_1$, with cell dimensions of $a=74.34$ Å, $b=111.13$ Å and $c=111.56$ Å.

X-ray diffraction data were measured at the SPring-8 in Japan. The crystal-to-detector distance and the crystal oscillation angle per image were set to 220 mm and 1°, respectively. Reflections were indexed, integrated and scaled using DENZO and SCALEPACK[24] (Table 2). Assuming monomeric DNA polymerase in the asymmetric unit, the crystal volume per enzyme mass (V_m) and the solvent content were calculated to be 2.57 $Å^3$/Da and 51.9%, respectively.[25] The refinement data statistics are summarized in Table 2.

The crystal structure was determined using a KOD DNA polymerase structure (Protein Data Bank entry 1GCX) as the initial phasing model. The model was manually modified using the program O[26] and subjected to further rounds of refinement using data in the resolution range 33.3–2.75 Å with the program CNS.[27] The R_{free} set contained a random sample of 5% of all data. The refinement statistics are shown in Table 2. MOLSCRIPT[28] and Raster3D[29] were used for drawing Figure 4. Secondary structural elements were determined with DSSP.[30]

Table 2. Crystallographic data statistics

	KOD DNA polymerase(H147E)
Diffraction data	
Space group	$P2_12_12_1$
Cell dimensions (Å)	$a=74.34$
	$b=111.13$
	$c=111.56$
Resolution range (Å)	33.3–2.75
No. of observations	90,637
No. of unique observations	22,669
Completeness (%)	90.0 (78.2)
R_{merge} (%)	8.2 (27.8)
$I/\sigma(I)$	10.0
Structure refinement	
Resolution range (Å)	33.3–2.75
R-factor (%)	22.7
No. of reflections used	21,228
Free *R*-factor (%)	30.6
No. of reflections used	1092
No. of protein atoms	5877
No. of solvent atoms	278
No. of heterogen atoms	7
Average *B* ($Å^2$)	53.7
r.m.s.d. in bond lengths (Å)	0.008
r.m.s.d. in bond angles (°)	1.2

Protein Data Bank accession numbers

The coordinates of the structure have been deposited in the RCSB Protein Data Bank (accession code 1WN7).

Acknowledgements

We thank Drs K. Miura and K. Hasegawa of SPring-8 for valuable help with data collection using synchrotron radiation at BL38B1. In addition, we are grateful to Drs Y. Tozawa, Y. Kawamura, B. Kawakami, M. Oka, and Y. Nishiya for useful discussions and to Mr H. Komatsubara, Mr Y. Ishida and Ms A. Tanabe for technical advice. The synchrotron radiation experiments were performed at the SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal No. 2002B0845-RL1-np). This work was partly supported by a grant from the Ministry of Education, Science, Sports and Technology of Japan.

References

1. Debyser, Z., Tabor, S. & Richardson, C. C. (1994). Coordination of leading and lagging strand DNA synthesis at the replication fork of bacteriophage T7. *Cell*, **77**, 157–166.
2. Waga, S. & Stillman, B. (1998). The DNA replication fork in eukaryotic cells. *Annu. Rev. Biochem*, **67**, 721–751.
3. Fanning, E. & Knippers, R. (1992). Structure and function of simian virus 40 large tumor antigen. *Annu. Rev. Biochem.* **61**, 55–85.
4. Kelman, Z. & O'Donnell, M. (1995). DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.* **64**, 171–200.
5. Wang, T. S. F. (1991). Eukaryotic DNA polymerases. *Annu. Rev. Biochem.* **60**, 513–552.
6. Uemori, T., Sato, Y., Kato, I., Doi, H. & Ishino, Y. (1997). A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization. *Genes Cells*, **2**, 499–512.
7. Takagi, M., Nishioka, M., Kakihara, H., Kitabayashi, M., Inoue, H., Kawakami, B. *et al*. (1997). Characterization of DNA polymerase from *Pyrococcus sp.* strain KOD1 and its application to PCR. *Appl. Environ. Microbiol*. **63**, 4504–4510.
8. Braithwaite, D. K. & Ito, J. (1993). Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucl. Acids Res.* **21**, 787–802.
9. Bult, C. J., White, O., Olsen, G. J., Zhou, L. M., Clayton, R. A., Gocayre, J. D. & Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science*, **273**, 1058–1073.
10. Joyce, C. M. (1989). How DNA travels between the separate polymerase and 3′-5′-exonuclease sites of DNA polymerase I (Klenow fragment). *J. Biol. Chem.* **264**, 10858–10866.
11. Truniger, V., Lázaro, J., Salas, M. & Blanco, L. (1996). A

DNA binding motif coordinating synthesis and degradation in proofreading DNA polymerases. *EMBO J.* **15**, 3430–3441.
12. Böhlke, K., Pisani, F. M., Vorgias, C. E., Frey, B., Sobek, H., Rossi, M. & Antranikian, G. (2000). PCR performance of the *B*-type DNA polymerase from the thermophilic euryarchaeon *Thermococcus aggregans* improved by mutations in the Y-GG/A motif. *Nucl. Acids Res.* **28**, 3910–3917.
13. Hopfner, K., Eichinger, A., Engh, R. A., Laue, F., Ankenbauer, W., Huber, R. & Angerer, B. (1999). Crystal structure of a thermostable type *B* DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl Acad. Sci. USA*, **96**, 3600–3605.
14. Tabor, S. & Richardson, C. C. (1989). Selective inactivation of the exonuclease activity of bacteriophage T7 DNA polymerase by *in vitro* mutagenesis. *J. Biol. Chem.* **264**, 6447–6458.
15. Wang, J., Sattar, A. K., Wang, C. C., Karam, J. D., Konigsberg, W. H. & Steitz, T. A. (1997). Crystal structure of a pola family replication DNA polymerase from bacteriophage RB69. *Cell*, **89**, 1087–1099.
16. Zhao, Y., Jeruzalmi, D., Moarefi, I., Leighton, L., Lasken, R. & Kuriyan, J. (1999). Crystal structure of an archaebacterial DNA polymerase. *Structure*, **7**, 1189–1199.
17. Rodriguez, A. C., Park, H., Mao, C. & Beese, L. S. (2000). Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon *Thermococcus sp.* 9° N-7. *J. Mol. Biol.* **299**, 447–462.
18. Hashimoto, H., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., Inoue, T. & Kai, Y. (2001). Crystal structure of DNA polymerase from hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1. *J. Mol. Biol.* **306**, 469–477.
19. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature*, **391**, 251–257.
20. Nishioka, M., Mizuguchi, H., Fujiwara, S., Komatsubara, S., Kitabayashi, M., Uemura, H. *et al.* (2001). Long and accurate PCR with a mixture of KOD DNA polymerase and its exonuclease deficient mutant enzyme. *J. Biotechnol.* **88**, 141–149.
21. Richardson, C. C. (1966). DNA polymerase from *Escherichia coli*. In *Procedures in Nucleic Acid Research* (Cantoni, G. L. & Davies, D. R., eds), pp. 263–276, Harper & Row, New York, NY.
22. Mo, J., Maki, H. & Sekiguchi, M. (1991). Mutational specificity of the dnaE173 mutator associated with a defect in the catalytic subunit of DNA polymerase III of *Escherichia coli*. *J. Mol. Biol.* **222**, 925–936.
23. Hashimoto, H., Matsumoto, T., Nishioka, M., Yuasa, T., Takeuchi, S., Inoue, T. *et al.* (1999). Crystallographic studies on a family B DNA polymerase from hyperthermophilic archaeon *Pyrococcus kodakaraensis* strain KOD1. *J. Biochem.* **125**, 983–986.
24. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
25. Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.
26. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1994). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
27. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
28. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
29. Merritt, E. A. & Bacon, D. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.
30. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

Edited by M. Guss

(Received 27 November 2004; received in revised form 6 March 2005; accepted 7 June 2005)

Available online 23 June 2005