Family A and family B DNA polymerases are structurally related: evolutionary implications

Weiguo Zhu and Junetsu Ito*

Department of Microbiology and Immunology, College of Medicine, The University of Arizona, Tucson, AZ 85724, USA

Received October 15, 1994; Accepted October 27, 1994

ABSTRACT

In order to establish the evolutionary relationship between the family A and B DNA polymerases, we have closely compared the 3'→5' exonuclease domains between the Klenow fragment of *E.coli* DNA polymerase I (a family A DNA polymerase) and the bacteriophage PRD1 DNA polymerase, the smallest member of the DNA polymerase family B. Although the PRD1 DNA polymerase has a smaller 3'→5' exonuclease domain, its active sites appear to be very similar to those of the Klenow fragment. Site-directed mutagenesis studies revealed that the residues important for the 3'→5' exonuclease activity, particularly metal binding ligands for the Klenow fragment, are all conserved in the PRD1 DNA polymerase as well. The metal binding ligands are also essential for the strand-displacement activity of the PRD1 DNA polymerase. Based on these results and the studies by others in various systems, we conclude that family A and B DNA polymerases, at least in the 3'→5' exonuclease domain, are structurally as well as evolutionarily related.

INTRODUCTION

Ever since the DNA molecule became the stable genetic informational carrier (DNA world), its replication has been central to the process of evolution. DNA polymerases are the key enzymes responsible for accurate replication of DNA. Generally, the fidelity of DNA replication is remarkably high; mutation rates are typically of the order of 10^{-8} to 10^{-11} per base replicated for bacteria (1). Three enzymatic processes are thought to be responsible for the high fidelity of DNA replication (2, 3). These are: (i) base selection, (ii) 3'→5' exonucleolytic proofreading and (iii) post-replicative mismatch repair (2, 3). The first two processes are often called 'error avoidance mechanisms' for which DNA polymerases generally are responsible.

Recently, more than 70 different primary sequences of DNA polymerases have been described. These DNA polymerases, including some putative DNA polymerases deduced from nucleotide sequence data, have been classified into four families, A, B, C and X, based on amino acid sequence similarities (4, 5).

At present, the large fragment (Klenow fragment) of *E.coli* DNA polymerase I (family A DNA polymerase) is the only replicative DNA polymerase for which three-dimensional structural information is available (6). Structural, biochemical and site-directed mutagenesis studies clearly established that the DNA polymerase and the 3'→5' exonuclease activities of the Klenow fragment reside on different domains that are separated by about 35 Å in the crystal structure (6–12). Based on structural studies of the Klenow fragment complexed with substrate DNA and dNMP at the 3'→5' exonuclease active site, Steitz and his colleagues (9, 11) have proposed a two metal ion mechanism for the hydrolytic phosphoryltransfer reaction. According to this mechanism, a nucleophilic attack on the phosphorous atom of the scissile phosphate is carried out by a hydroxide ion that is activated by metal A, while the second metal ion B stabilizes the transition state (11). Thus, the primary roles of the amino acid side chains in the exonuclease domain are to bind and orient the two metal ions (at A and B sites), the primer single-stranded DNA and the attacking water molecule in a proper configuration (11). Since 3'→5' exonuclease activities of all known DNA polymerases require Mg^{2+} (2), it is reasonable to expect that the other proofreading DNA polymerases have metal ion binding ligands, similar to the Klenow fragment.

But can the structure and the mechanism of the Klenow fragment exonuclease be applied to the 3'→5' exonuclease of the family B and C DNA polymerases? Are they truly structurally and evolutionarily related? While evidence supporting the notion that family A, B and C DNA polymerases are all related and share a common ancestor is accumulating (13–25), sequence similarities between the different families are at best very weak, as was pointed out by Beese and Steitz (11). Therefore, the possibility of convergent evolution cannot be ruled out. To address this question, we have applied systematic site-directed mutagenesis to a DNA polymerase encoded by a small lipid-containing *E.coli* phage PRD1 (26). The PRD1 DNA polymerase, consisting of 553 amino acid residues, is the smallest replicative DNA polymerase belonging to the family B DNA polymerases, which include eukaryotic replicative DNA polymerases α , δ and ϵ , archaeobacterial DNA polymerases and viral DNA polymerases (4, 5). Our results show that the 3'→5' exonuclease domain of the PRD1 DNA polymerase is smaller

*To whom correspondence should be addressed

than that of the Klenow fragment, but the active sites are co-linearly arranged with those of the Klenow exonuclease domain. Furthermore, putative metal binding residues are totally conserved, suggesting that these exonucleases share a common progenitor. In addition, we present evidence suggesting the metal binding ligands are involved in strand-displacement DNA synthesis.

MATERIALS AND METHODS

Materials

Ultrapure, unlabeled dNTPs and poly(dC)–oligo(dG)_{12–18} were purchased from Pharmacia P-L Biochemicals. Radionucleotides, [α -³²P]dGTP, [α -³⁵S]dATP, [³H]dGTP and [³H]dCTP were from New England Nuclear. T7 DNA Sequencing kit (Version 2.0), T4 polynucleotide kinase, T4 DNA polymerase, T4 DNA ligase and T4 gp32 were purchased from United States Biochemicals. All the synthetic oligonucleotides used for the site-specific mutagenesis and DNA sequencing were synthesized on a Cyclone DNA synthesizer (Milligen/Bioscience). Beta-amidite synthesis kits used for the synthesis were purchased from Milligen/Bioscience Co.

Bacteriophage PRD1 was grown in *E. coli* strain HB94 carrying the kanamycin resistance plasmid pLM2 (26). Phage particles were prepared by 12% PEG/0.4 M NaCl precipitation and 5–20% sucrose gradient centrifugation. The PRD1 DNA–terminal protein complex was prepared from purified phage particles. Purified phage particles were heated at 60°C for 2 min and were disrupted with phenol. After centrifugation, the interface between the aqueous layer and phenol was collected. The sample was then dialyzed against Tris–HCl buffer (20 mM, pH 7.6) containing 1 mM EDTA. Deproteinized PRD1 DNA was prepared as described previously (27).

Site-specific mutagenesis

A plasmid pEJG containing PRD1 DNA polymerase gene was derived from phagemid pEMBLex3 which carries a strong λ phage pR promoter and a temperature-sensitive cI857 repressor gene (28). *In vitro* site-specific mutagenesis was carried out according to Kunkel *et al.* (29). *E. coli* CJ 236 [*dut-1*, *ung-1*, *thi-1*, *relA-1*, *pCJ107* (*Cm*^r)] (Bio-Rad) was used to prepare the uracil ssDNA template. *E. coli* JL2443 [Δ *lac-proAB*], *thi*, *hsd5*, *supE*/ *F'*(*traD36*, *proAB*, *lacI^q*, *lacZ* Δ M15)], from Dr John W. Little of the University of Arizona, was used for the transformation of the uracil-containing plasmid DNA and as the host for the expression of recombinant PRD1 DNA polymerase gene. The standard site-specific mutagenesis reaction conditions were described previously (30). The resulting double-stranded DNA was used to transform *E. coli* JL2443 competent cells. Colonies were selected randomly and directly sequenced to screen for the mutations.

The oligonucleotides designed to generate site-specific mutations are as follows: D17A: 5'-GCCGCTTTGCCTTTG-AAACTGAC-3'; D19A: 5'-TTTGACTTTGCAACTGACCC-TTTC-3'; N71D: 5'-TACGCTCATGACGGCGCAAGT-TTG-3'; F75Y: 5'-GGCGGCAAGTATGATTTTCTTTTCTC-3'; D76A: 5'-GGCAAGTTTGCTTTTCTTTTCTC-3'; D107H: 5'-AATTCCGCCATAGTTATG-3'; D107A: 5'-AAATCCGCGCTAGTTATGCA-3'; D107N: 5'-AATTCCGCAATAGTTATG-3'; D107E: 5'-TTCCGCGAAAGTTATGCA-3'; Y145F: 5'-ATTTTAGAATTCCTGAAAGGCGAT-3'; D149A: 5'-CTGAAAGGCGCTTGTGTAACC-CTG-3';

Protein purification

Both wild-type and mutant PRD1 DNA polymerases were purified to near homogeneity from the *E. coli* expression vector pEJG by using a guanidine hydrochloride denaturation–renaturation procedure as well as heparin–agarose and hydroxyapatite chromatographies (31). Protein concentrations were determined according to Bradford (32). The purity of the proteins were determined by SDS–polyacrylamide gel electrophoresis followed by Coomassie blue staining.

Assay conditions

The DNA polymerase activity on poly(dC)–oligo(dG) synthetic template was measured as described (33). One unit of DNA polymerase activity was defined as the amount of DNA polymerase that incorporates 1 nmol of labeled dGMP into acid-insoluble DNA at 30°C in 30 min under the standard assay conditions. 3'→5' exonuclease activity was determined as described (31) using [³H]dGMP-labeled single-stranded DNA. One unit of exonuclease activity was defined as the amount of DNA polymerase required to catalyze the release of 1 nmol of labeled dGMP to the ethanol-soluble form in 30 min at 30°C under the standard conditions. The conditions for PRD1 terminal protein–dGMP complex formation were as described (30). The activity was determined by excising the radioactive bands from the gel and counting the Cerenkov radiation. One unit of the initiation activity was defined as the formation of 1 pmol of terminal protein–dGMP complex. *In vitro* replication of PRD1 DNA with purified components was carried out as described previously (31) except that [³H]dCTP was used and the incubation temperature was 30°C. The unit definition of DNA polymerase activity on the terminal protein–DNA complex was the same as that on the poly(dC)–oligo(dG) synthetic template. The DNA polymerase fill-in reaction was carried out according to Bernad *et al.* (16). A 25 μ l reaction mixture contained 50 mM Tris–HCl, pH 7.6, 10 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 5% glycerol, 150 μ M each of dATP, dCTP and TTP, 30 μ M of [α -³²P]dGTP (5 μ Ci), 0.5 μ g of *Ban*I-digested PRD1 DNA and 100 ng of PRD1 DNA polymerase. After incubation at 30°C for 30 min, the reaction was stopped and the samples were filtered through Sephadex G-50 spin columns. The Cerenkov radiation was counted and one unit of fill-in activity represented the incorporation of 1 nmol of dGMP to the 3' ends of the molecules. The conditions for DNA polymerase strand-displacement activity was as described (34). A 25 μ l reaction mixture contained 50 mM Tris–HCl, pH 7.6, 10 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 5% glycerol, 150 μ M of dATP, dCTP and TTP, 30 μ M of [α -³²P]dGTP (5 μ Ci), 0.5 μ g primed-M13mp18 ssDNA and 100 ng of purified DNA polymerase. After incubation for 15 min at 30°C, the reaction was stopped by adding 20 mM EDTA. The samples were ethanol precipitated and subjected to electrophoresis in a 0.7% alkaline agarose gel and the labeled DNA was visualized by autoradiography. The activity was judged by comparing the size of replication products of the PRD1 DNA polymerase mutants to those of the Klenow fragment and the wild-type PRD1 DNA polymerase.

RESULTS

Sequence similarities and site-directed mutagenesis in the 3'→5' exonuclease domain of PRD1 DNA polymerase

The amino acid residues 1–168 of PRD1 DNA polymerase can be aligned with residues 328–531 of *E. coli* DNA polymerase

PRD1 DNA polymerase

Figure 1. Schematic representation of the PRD1 DNA polymerase domain structure and the sequence alignment of the N-terminal portion of PRD1 DNA polymerase with that of the Klenow fragment of *E. coli* DNA polymerase I. The top bar represents PRD1 DNA polymerase. The N-terminal portion is enlarged and amino acid residues 1–168 are aligned to amino acid residues 338–531 of *E. coli* DNA polymerase I. Conserved amino acid residues are indicated by black boxes. The amino acid residues of PRD1 DNA polymerase subjected to mutagenesis are indicated by arrows. The amino acid residues involved in the 3'→5' exonuclease active site of the Klenow fragment are indicated by asterisks and the positions are noted.

I (Pol I) (Figure 1). Clearly, the 3'→5' exonuclease domain of PRD1 DNA polymerase is smaller (about 9%) than that of Pol I. The crystal structure of the Klenow fragment showed that this region contains anti-parallel β -pleated sheets 2–5 as well as α -helices B–F (6). The structural studies also showed that a divalent metal ion is oriented by the carboxylate groups of Asp355, Glu357 and Asp501 of the Klenow fragment, together with the phosphate of dNMP (A site) (9, 11). The second metal ion was located between the dNMP phosphate and the carboxylate groups of Asp424 and Asp355 (9, 11). As can be seen from Figure 1, the amino acid residues corresponding to those ligands of the Klenow exonuclease domain all appear to be conserved in PRD1 DNA polymerase as well. To test the possibility that these conserved amino acid residues are the active sites for 3'→5' exonuclease in PRD1 DNA polymerase, we have constructed a series of mutant proteins by site-specific mutagenesis. All the amino acid residues that are changed in this study are indicated in Figure 1. Each mutant DNA polymerase was purified to

apparent homogeneity and assayed for various enzymatic activities. Table 1 summarizes the effects of mutations on the various catalytic activities of PRD1 DNA polymerase. It is evident from Table 1 that mutations D17A, E19A, D76A, D107H and D149A severely affected the 3'→5' exonuclease activity. These mutations reduced the exonuclease activity by at least five orders of magnitude. On the other hand, these mutants, except for D107H which showed about 7.5% of the wild-type level, showed nearly normal levels of the DNA polymerase activity, suggesting that these proteins are not drastically affected structurally by the mutations. Assuming that the Klenow fragment and the PRD1 DNA polymerase are structurally homologous for this region, residues Asp17, Glu19 and Asp149 of PRD1 DNA polymerase are the metal A binding ligands, while Asp17 and Asp76 are the metal B binding ligands.

Asp107 appears to be highly conserved among the family A and family B DNA polymerases (5) and is aligned to Asp459 of the Klenow fragment (Figure 1). Histidine substitution for

Table 1. Effect of mutations on the catalytic activities of PRD1 DNA polymerase

	3'→5' Exo ^a U/mg (%)	Replication on poly(dC)-oligo(dG) ^b U/mg (%)	Ratio of 3'→5' exo to replication %	Filling-in ^c U/mg (%)	Initiation ^d U/mg (%)	Replication on TP-DNA complex ^e	Strand displacement ^f U/mg (%)
Wild-type	177.4 (100)	1307.1 (100)	100.0	2.70 (100)	6.05 (100)	519.0 (100)	+++
D17A	<0.01 (<0.01)	1529.3 (117)	<0.01	2.78 (103)	5.26 (86.9)	17.6 (3.4)	±
E19A	<0.01 (<0.01)	1261.4 (96.5)	<0.01	2.38 (88.1)	4.15 (68.6)	14.0 (2.7)	—
N71D	23.1 (13.0)	1288.8 (98.6)	13.2	3.83 (88.1)	1.28 (21.2)	325.9 (62.8)	+++
F75Y	168.2 (94.8)	1269.2 (97.1)	97.6	4.64 (172)	3.61 (59.6)	353.4 (68.1)	++
D76A	<0.01 (<0.01)	1211.7 (92.7)	<0.01	2.24 (83.0)	4.73 (78.2)	9.3 (1.8)	—
D107H	<0.01 (<0.01)	98.0 (7.5)	<13.3	0.08 (3.0)	0.25 (4.1)	0.5 (0.1)	—
D107A	6.6 (3.7)	474.5 (36.3)	10.2	ND	ND	ND	+++
D107N	21.3 (12.0)	766.0 (58.6)	20.5	ND	ND	ND	+++
D107E	44.4 (25.0)	1141.1 (87.3)	28.6	ND	ND	ND	+++
Y145F	83.6 (47.1)	1281.0 (98.0)	48.1	2.35 (87.1)	7.50 (124)	183.7 (35.4)	++
D149A	<0.01 (<0.01)	949.0 (72.6)	<0.01	4.94 (183)	6.96 (115)	3.6(0.7)	—

^aThe 50 μ l reaction mixture contained 50 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 5% glycerol, 10 μ g/ml ³H-labeled poly(dG) single-stranded DNA (30,000 cpm) and 100 ng of the PRD1 DNA polymerase. After 30 min incubation at 30°C, the ethanol-soluble products were measured for their radioactivities.

^bThe standard reactions were performed using poly(dC)-oligo(dG)₁₂₋₁₈ (1:1 ratio) as template in 50 mM Tris-HCl, pH 7.6, 5 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 30 μ M [³H]dGTP (500 cpm/pmol), 0.3 μ M of template and 100 ng of the PRD1 DNA polymerase.

^cThe reaction mixture (25 μ l) contained 50 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 5% glycerol, 150 μ M each of dATP, dCTP, TTP, 30 μ M [³H]dGTP (5 μ Ci), 0.5 μ g of the *Ban*I-digested PRD1 DNA and 100 ng of the PRD1 DNA polymerase. After incubation of the reaction mixture at 30°C for 20 min, the sample was run through a Sephadex G-50 column and subjected to 1.5% agarose gel electrophoresis as well as autoradiography.

^dThe assay conditions for the PRD1 terminal protein-dGMP complex formation were the same as described (30).

^eThe reaction mixture (50 μ l) contained 50 mM Tris-HCl pH, 7.6, 7.5 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 10 mM (NH₄)₂SO₄, 100 μ M each of dATP, dGTP and TTP, 10 μ M [³H]dCTP, 500 ng of the purified PRD1 terminal protein, 100 ng of PRD1 DNA polymerase and 1 μ g of the PRD1 DNA-terminal protein complex as template. The reaction was carried out at 37°C for 30 min. The acid-insoluble products were collected on Whatman GF/C filter and counted for radioactivity.

^fThe reaction conditions for the DNA polymerase strand-displacement activity were the same as that of the filling-in assay except that primed M13mp18 single-stranded DNA was used as the template. The activity was judged by comparing the size of the replication products of the PRD1 DNA polymerase mutants to those of the Klenow fragment and wild-type PRD1 DNA polymerase. The Klenow fragment was a negative reference and the wild-type PRD1 DNA polymerase was a positive control.

ND, not determined.

Asp107 severely affected not only the 3'→5' exonuclease activity but also the entire DNA polymerase function. Since the histidine substitution changes the negative charged side chain to a positive one, it is likely to cause disruption of the secondary structure of the protein. Therefore, we have constructed three additional mutants by substituting the Asp107 of PRD1 DNA polymerase with alanine, asparagine and glutamic acid, respectively. Our functional analysis of these mutant DNA polymerases indicated that glutamic acid and asparagine substitutions were better tolerated than alanine substitution, suggesting that charge and ligand geometry are as important as the side chain at position 107 of PRD1 DNA polymerase (Table 1).

Mutations N71D and Y145F reduced the exonuclease activity to 13% and 47% of the wild-type level, respectively (Table 1). Asn71 is highly conserved and aligned to Asn420 of the Klenow fragment. It is known that the δ -amino group of Asn420 forms a hydrogen bond with the 4'-oxygen of the deoxyribose in the 3' penultimate nucleotide of the DNA template (11). Therefore, this residue is involved in the interaction between the single-stranded substrate DNA and the enzyme. Replacement of Asn71 with aspartic acid would disrupt the hydrogen bond, thus affecting substrate binding. The Phe75 of PRD1 DNA polymerase is highly conserved among the family B DNA polymerases. However, since it corresponds to Tyr423 of the Klenow fragment, mutation F75Y does not appreciably affect the exonuclease activity of PRD1 DNA polymerase, although it affected the protein-primed initiation activity considerably (Table 1).

Strand-displacement DNA synthesis

PRD1 DNA is replicated by a terminal protein-primed, strand-displacement mechanism (35, 36). Thus, replication is initiated

Figure 2. Comparison of strand-displacement activity of PRD1 DNA polymerase with that of other DNA polymerases. The strand-displacement assay was carried as described under Materials and Methods. The reaction mixture (25 μ l) contained 50 mM Tris-HCl, pH 7.6, 10 mM MgCl₂, 1 mM DTT, 0.5 mg/ml BSA, 5% glycerol, 150 μ M each of dATP, dCTP and TTP, 30 μ M [³H]dGTP (5 μ Ci), 0.5 μ g of primed M13mp18 ssDNA and 100 ng of the purified DNA polymerase. After incubating for 15 min at 30°C, the reaction was stopped by adding EDTA to 20 mM. Samples were processed and analyzed by 0.7% alkaline agarose gel electrophoresis and autoradiography. Lane 1: Klenow fragment of *E. coli* DNA polymerase I; lane 2: Exo⁻ mutant Klenow fragment; lane 3: T7 DNA Sequenase version 1.0 (chemically modified); lane 4: T7 DNA Sequenase version 2.0 (Exo deletion mutant); lane 5: Taq DNA polymerase; lane 6: T4 DNA polymerase; lane 7: wild-type PRD1 DNA polymerase. Both wild-type and exo⁻ mutant Klenow enzymes, T7 DNA Sequenases and T4 DNA polymerase were from United State Biochemical. Taq DNA polymerase was from Perkin-Elmer.

Figure 3. Mutational effect on the strand-displacement activity of PRD1 DNA polymerase. The strand-displacement assay was carried as described (34). Lane 1 represents the Klenow fragment of *E. coli* DNA polymerase I. Lane 2 represents the wild-type PRD1 DNA polymerase. Lanes 3–8 represent PRD1 DNA polymerase exonuclease mutants D17A, E19A, N71D, F75Y, D76A and D149A, respectively. The DNA length markers are indicated as well.

at either end of the terminal redundant DNA molecule with the covalent linkage of dGMP to Tyr190 of the PRD1 terminal protein (30). The subsequent DNA chain elongation concomitantly displaces one parental DNA strand. While PRD1 DNA polymerase is the smallest known enzyme containing editing 3'→5' exonuclease activity, it is remarkably capable of performing strand-displacement DNA synthesis, unaided by other proteins (31, 36). Evidently, the ability to catalyze strand-displacement DNA synthesis without other protein factors is a common intrinsic property of several protein-primed DNA polymerases, as was described for the *Bacillus* phage ϕ 29 DNA polymerase (34). Other DNA polymerases which are known to be able to catalyze strand-displacement synthesis without the aid of other proteins on nicked, duplex DNA templates are modified T7 DNA polymerase (37, 38), T5 DNA polymerase (39) and the Klenow fragment of *E. coli* DNA polymerase I (38). Contrary to earlier reports (40, 41), Lechner and Richardson (37) showed that *E. coli* DNA polymerase I cannot catalyze strand-displacement synthesis at a nick. To compare the strand-displacement activity of PRD1 DNA polymerase with other DNA polymerases, we employed an oligonucleotide-primed M13 DNA system described by Blanco and Salas (34). The reaction products were analyzed by denatured gel electrophoresis. As shown in Figure 2, only the PRD1 DNA polymerase is capable of performing extensive strand-displacement DNA synthesis in this assay. Under the conditions of lower ionic strength, the Taq and T4 DNA polymerases are not processive (42). Our results also agree with Blanco and Salas's (34) in that the Klenow fragment is incapable of catalyzing strand-displacement DNA synthesis in the M13 DNA strand-displacement assay. It is somewhat surprising that the modified T7 DNA polymerase (Sequenase) cannot perform strand-displacement synthesis in this assay.

Soengas *et al.* (22) have recently presented evidence that ϕ 29 DNA polymerase mutants altered at Tyr165 and Asp169, located in the region of the 'Exo III motif' of the 3'→5' exonuclease domain, cannot catalyze strand-displacement DNA synthesis. We tested all the mutants deficient in 3'→5' exonuclease described in this study for their abilities to catalyze strand-displacement synthesis. Our results showed that PRD1 DNA polymerase mutants E19A, D76A and D149A were incapable of catalyzing strand-displacement DNA synthesis (Figure 3, Table 1). In addition, a mutant D17A was only partially active in the strand-

displacement assay. All these mutants are believed to be involved in divalent metal ion binding. While mutants N71D, D107A and D107N have lower 3'→5' exonuclease activity (3.7–13%) compared to the wild-type enzyme, they appear to be capable of promoting normal strand-displacement DNA synthesis (Table 1).

To test the possibility that the inability of the 3'→5' exonuclease mutants to catalyze strand-displacement DNA synthesis after one round of M13 DNA synthesis is derived from their inability to unwind the 5' end of primed DNA, we used preformed 5' end displaced oligonucleotide primer with M13 DNA as template. However, this was not the case, because these mutants were unable to perform strand-displacement synthesis even when presented with preformed strand-displaced DNA substrate (data not shown).

DISCUSSION

The purpose of the present study was to gain clear evidence indicating that the family A DNA polymerases, including *E. coli* DNA polymerase I, are structurally and evolutionarily related to the family B DNA polymerases.

Our sequence alignment suggested that the N-terminal portion of PRD1 DNA polymerase can be aligned with the 3'→5' exonuclease domain of the Klenow fragment. While the PRD1 DNA polymerase has a smaller 3'→5' exonuclease domain, all the active sites appear to be conserved. Site-directed mutagenesis studies showed that all the corresponding carboxylic amino acid ligands, Asp17, Glu19, Asp76 and Asp149, are functionally critical for the 3'→5' exonuclease activity of the PRD1 DNA polymerase. These results are very similar to those described by Derbyshire *et al.* (10), where mutations D355A, E357A, D424A and D501A reduced the exonuclease activity of the Klenow fragment by three–four orders of magnitude. Therefore, it is reasonable to assume that PRD1 DNA polymerase Glu19 and Asp149 correspond to Klenow Glu357 and Asp501 (metal A ligands), Asp76 of PRD1 DNA polymerase corresponds to Klenow Asp424 (a metal B ligand) and PRD1 Asp17 corresponds to Klenow Asp355 (metal A and B shared ligand).

There are some clear variations between the PRD1 DNA polymerase and the Klenow fragment, however. In the case of the Klenow fragment, the bases on either side of the scissile phosphate are positioned by interactions with the side chains of Phe473 and Leu361 (11). The PRD1 DNA polymerase has none of these amino acid residues at the corresponding positions. However, recent sequence alignment of more than 60 different 3'→5' exonuclease domains indicated that the Leu361 and Phe473 are not highly conserved amino acids (unpublished data).

Our site-directed mutagenesis studies showed that the Asp107 of PRD1 DNA polymerase could be an important residue in not only the 3'→5' exonuclease activity but for DNA polymerase activity as well. This amino acid residue appears to be highly conserved among family B and A DNA polymerases. In the case of the Klenow fragment, the corresponding residue is Asp459 which is located at the end of α -helix D (6). Amino acid substitution studies indicated that Asp107 of PRD1 DNA polymerase can be replaced at least partially with glutamic acid and asparagine, suggesting that charge and ligand geometry could be as important as the side chain at this position. It is of significance to note here that mutations in the 3'→5' exonuclease domain of various DNA polymerases impaired DNA polymerase activities or *in vivo* DNA replication (23, 25, 43–45). Tabor

and Richardson (43) have shown that T7 phage mutants defective in 3'→5' exonuclease activity have a drastically reduced burst size, although DNA polymerase activities are normal in *in vitro* assays. T7 DNA polymerase is a member of the family A DNA polymerases. In the case of Herpes simplex DNA polymerase which belongs to the family B DNA polymerases, mutations in the 3'→5' exonuclease domain severely impaired viral DNA replication *in vivo* (44), or significantly affected DNA polymerase sensitivities to various inhibitors (45). From these observations, it has been suggested that the domains of the Herpesvirus DNA polymerase may be more interconnected. Similar results have recently been reported for T4 DNA polymerase (25). Therefore, at a glance, all this evidence appears to be contradictory to the notion that the 3'→5' exonuclease domain and DNA polymerization domain are structurally separated. However, it seems only natural that two domains of DNA polymerase are functionally interdependent *in vivo*. Perhaps, it is relevant to note that a recent structure of Klenow fragment bound to duplex DNA showed much closer distance between the 3'→5' exonuclease domain and DNA polymerization domain than was previously thought (12).

The strand-displacement DNA synthesis described in this communication appears to be one clear example of the functional interdependence of the DNA polymerase domains. It has been shown that the $\phi 29$ DNA polymerase mutants located in the 'Exo III motif' of the 3'→5' exonuclease domain cannot perform strand-displacement DNA synthesis (22). Our present study shows further that in fact all the metal binding ligands may be required in strand-displacement DNA synthesis (Table 1, Figure 3). Although the mechanistic features required for strand-displacement DNA synthesis are not well understood, one can imagine the possible ways in which the DNA polymerase domain and the 3'→5' exonuclease domain could be interdependent. (i) Unlike the Klenow fragment, which is not very processive, the PRD1 DNA polymerase is highly processive (31). This strong processivity may be due to the 3'→5' exonuclease domain. It is well documented that the processivity of the *E. coli* DNA polymerase III α - ϵ subunit complex is strictly dependent on the 3'→5' exonuclease subunit, ϵ (46). It is also known that the 3'→5' exonuclease mutants of yeast mitochondrial DNA polymerase, MIP1, are less processive compared to the wild-type DNA polymerase (23). (ii) The PRD1 DNA polymerase may have an intrinsic 'DNA helicase' activity. Since the PRD1 DNA polymerase, unlike T4 or T7 DNA polymerase, can perform strand-displacement without aid of any other proteins, it is possible that this polymerase contains an intrinsic SV-40 T antigen-type DNA helicase activity which translocates in the 3'→5' direction along the template DNA and promotes strand-displacement of the 5' chain (47, 48). All the known DNA helicase activities require Mg^{2+} ions (49). (iii) The possibility that the PRD1 DNA polymerase works as a dimer has not yet been ruled out. Such a homodimer could play multiple roles in processivity, editing and strand-displacement DNA synthesis. The availability of point mutations defective in strand-displacement DNA synthesis may facilitate the detailed analysis of this mechanism.

Evolutionary implications

In 1988, it was first pointed out that there exist possible sequence similarities between the 3'→5' exonuclease regions of T4 DNA polymerase (a member of family B DNA polymerases) and *E. coli* DNA polymerase I (a family A DNA polymerase) (50).

Subsequent sequence alignments and site-directed mutagenesis studies with various systems (13–25) have led to the hypothesis that the family A and family B, and possibly family C, DNA polymerases are all related. Our present results provide strong support for this idea.

Very recently, a non-replicative DNA polymerase β (family X DNA polymerase) has been crystallized (51–53). This DNA polymerase does not contain a 3'→5' exonuclease domain and does not show sequence homology to any other replicative DNA polymerases. Interestingly, however, the crystal structure of rat DNA polymerase β suggested a two metal ion mechanism for the nucleotidyl transfer reaction, similar to that of the 3'→5' exonuclease domain of the Klenow fragment (9, 11). Although the geometries of metal sites A and B are similar, the orientations of the carboxylic amino acids are quite different between the polymerase β and the 3'→5' exonuclease of the Klenow fragment (51–53). Therefore, it seems unlikely that a similar arrangement of the conserved motif of the 3'→5' exonuclease domains could have evolved convergently in three evolutionary lineages. Rather, the 3'→5' exonuclease domains of the family A, B and C DNA polymerases must have evolved divergently from a common ancestor. Since all three primary lines of descent, bacteria, archaea and eukarya (54), contain family B DNA polymerases (5), it seems most likely that the last common ancestor cell should have had at least an ancestral family B DNA polymerase. Most of the present-day family B DNA polymerases contain a 3'→5' editing exonuclease. Therefore, the ancestral family B DNA polymerase probably possessed a 3'→5' exonuclease and would have exhibited high fidelity in replicating DNA. Doolittle *et al.* (55) have estimated the prokaryote–eukaryote divergence time from protein sequences. Their results indicated the average time of the divergence to be 1.8 ± 0.4 billion years. Therefore, it is likely that a self-correcting, highly accurate DNA polymerase existed more than 1.8 billion years ago.

ACKNOWLEDGEMENTS

We would like to thank Dr Russell F. Doolittle for helpful discussions. We also thank Harris Bernstein, Jennifer Hall, Vivian Gage and Mark Dodson for their critical reading of this manuscript. J.I. gratefully acknowledges a sabbatical supplement award for molecular studies of evolution from the Alfred P. Sloan Foundation. This research was supported by a grant from the N.I.H. (GM28013).

REFERENCES

1. Drake, J.W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7160–7164.
2. Kornberg, A. and Baker, T.A. (1992) *DNA Replication* (2nd Edition). Freeman, San Francisco, CA.
3. Echols, H. and Goodman, M.F. (1991) *Annu. Rev. Biochem.* **60**, 477–511.
4. Ito, J. and Braithwaite, D.K. (1991) *Nucleic Acids Res.* **19**, 4045–4057.
5. Braithwaite, D.K. and Ito, J. (1993) *Nucleic Acids Res.* **21**, 787–802.
6. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. and Steitz, T.A. (1985) *Nature (Lond.)* **313**, 762–766.
7. Joyce, C.M. and Steitz, T.A. (1987) *Trends Biochem. Sci.* **12**, 288–292.
8. Derbyshire, V., Freemont, P.S., Sanderson, M.R., Beese, L.S., Friedman, J., M. Joyce, C.M. and Steitz, T.A. (1988) *Science* **240**, 199–201.
9. Freemont, P.S., Friedman, J.M., Beese, L.S., Sanderson, M.R. and Steitz, T.A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8924–8928.
10. Derbyshire, V., Grindley, N.D.F. and Joyce, C.M. (1991) *EMBO J.* **10**, 17–24.
11. Beese, L.S. and Steitz, T.A. (1991) *EMBO J.* **10**, 25–33.
12. Beese, L.S., Derbyshire, V. and Steitz, T.A. (1993) *Science* **260**, 352–355.

13. Reha-Krantz, L.J., Liesner, E.M., Parmaksizoglu, S. and Stocki, S. (1986) *J. Mol. Biol.* **189**, 261–272.
14. Jung, G., Leavitt, M.C., Hsieh, J.C. and Ito, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8287–8291.
15. Leavitt, M.C. and Ito, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4465–4469.
16. Bernad, A., Blanco, L., Lazaro, J.M., Martin, G. and Salas, M. (1989) *Cell* **59**, 219–228.
17. Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P. (1990) *Protein Engng* **3**, 461–467.
18. Morrison, A., Bell, J.B., Kunkel, T.A. and Sugino, A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9473–9477.
19. Simon, M., Giot, L. and Faye, G. (1991) *EMBO J.* **10**, 2165–2170.
20. Blanco, L., Bernad, A., Blasco, M.A. and Salas, M. (1991) *Gene* **100**, 27–38.
21. Patel, S.S., Wong, I. and Johnson, K.A. (1991) *Biochemistry* **30**, 511–525.
22. Soengas, M.S., Esteban, J.A., L zaro, J.M., Bernad, A., Blasco, M.A., Salas, M. and Blanco, L. (1992) *EMBO J.* **11**, 4227–4237.
23. Foury, F. and Vanderstraeten, S. (1992) *EMBO J.* **11**, 2717–2726.
24. Frey, M.W., Nossal, N.G., Capson, T.L. and Benkovic, S.J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2579–2583.
25. Reha-Krantz, L.J. and Nonay, R.L. (1993) *J. Biol. Chem.* **268**, 27100–27108.
26. Mindich, L. and McGraw, T. (1983) *Mol. Gen. Genet.* **190**, 233–236.
27. Ito, J. (1978) *J. Virol.* **28**, 895–904.
28. Jung, G., Leavitt, M.C., Schultz, M. and Ito, J. (1990) *Biochem. Biophys. Res. Commun.* **170**, 1294–1300.
29. Kunkel, T.A., Roberts, J.D. and Zákour, R.A. (1987) *Methods Enzymol. – Recombinant DNA* **154**, pp. 367–382.
30. Shiue, S., Hsieh, J.-C. and Ito, J. (1991) *Nucleic Acids Res.* **19**, 3805–3810.
31. Zhu, W. and Ito, J. (1994) *Biochim. Biophys. Acta* **1219**, 267–276.
32. Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248–254.
33. Watabe, K., Leusch, M. and Ito, J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5374–5378.
34. Blanco, L., Bernad, A., Lazaro, J.M., Martin, G., Germendia, C. and Salas, M. (1989) *J. Biol. Chem.* **264**, 8935–8940.
35. Yoo, S.-K. and Ito, J. (1989) *Virology* **170**, 442–449.
36. Savilhati, H., Caldentey, J., Lurdstrom, K., Syvaaja, J.E. and Bamford, D.H. (1991) *J. Biol. Chem.* **266**, 18737–18744.
37. Lechner, R.L. and Richardson, C.C. (1983) *J. Biol. Chem.* **258**, 11185–11196.
38. Lechner, R.L., Engler, M.J. and Richardson, C.C. (1983) *J. Biol. Chem.* **258**, 11174–11184.
39. Fujimura, R.K. and Roop, B.C. (1976) *J. Biol. Chem.* **251**, 2168–2175.
40. Schildkraut, C.L., Richardson, C.C. and Kornberg, A. (1964) *J. Mol. Biol.* **9**, 24–45.
41. Masamune, Y. and Richardson, C.C. (1971) *J. Biol. Chem.* **246**, 2692–2701.
42. Jarvis, T.C., Newport, J.W. and von Hippel, P.H. (1991) *J. Biol. Chem.* **266**, 1830–1840.
43. Tabor, S. and Richardson, C.C. (1989) *J. Biol. Chem.* **264**, 6447–6458.
44. Gibbs, J.S., Weisshart, K., Digard, P., DeBruynkops, A., Knipe, D.M. and Coen, D.M. (1991) *Mol. Cell. Biol.* **11**, 4786–4795.
45. Wang, Y., Woodward, S. and Hall, J.D. (1992) *J. Virol.* **66**, 1814–1816.
46. Studwell, P.S. and O'Donnel, M. (1990) *J. Biol. Chem.* **265**, 1171–1178.
47. Wiekowski, M., Schwarz, M.W. and Stahl, H. (1988) *J. Biol. Chem.* **263**, 436–442.
48. Goetz, G.S., Dean, F.B., Hurwitz, J. and Matson, S.W. (1988) *J. Biol. Chem.* **263**, 383–392.
49. Matson, S.W. and Kaiser-Rogers, K.A. (1990) *Ann. Rev. Biochem.* **59**, 289–329.
50. Reha-Krantz, L.J. (1988) *J. Mol. Biol.* **202**, 711–724.
51. Davies, J.F.II, Almassy, R.J., Hostomska, Z., Ferre, R.A. and Hostomsky, Z. (1994) *Cell* **76**, 1123–1133.
52. Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H. and Kraut, J. (1994) *Science* **264**, 1891–1903.
53. Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H. and Kraut, J. (1994) *Science* **264**, 1930–1935.
54. Woese, C.R., Kandler, O. and Wheelis, M.L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579.
55. Doolittle, R.F., Anderson, K.L. and Feng, D. (1989) In Fernholm, B., Bremer, K. and Jornvall, H. (eds), *The Hierarchy of Life*. Elsevier Science Publisher, pp. 73–85.