Structure-Function Analysis of PRDI DNA Polymerase

Guhung Jung
Departemt of Biology Education Seoul National University

INTRODUCTION

Since DNA contains the genetic information of an organism, accurate DNA replication is one of the most important events of the life cycle of an organism. DNA polymerases are the key enzymes catalyzing the accurate replication of DNA. However, none of the known DNA polymerases can initiate *de novo* synthesis of a DNA chain without a primer. Generally, small RNA molecules are used as primers. These provide a 3'-hydroxyl group for DNA chain elongation. However, a different type of DNA polymerase has been identified in various systems. These enzymes utilize proteins as primers for the initiation of DNA synthesis.

PRDl is a member of a group of lipid-containing phages and is a good model system for the protein primed DNA replication. The genome of PRDl is 14.7 kilobases long with terminal proteins covalently linked at both 5' ends.

DNA polymerases have been classified, based on their amino acid sequence similarities, into two major groups : family A and family B(Jung et al., 1987 ; Leavitt and Ito, 1989). The family A DNA polymerases, represented by *E. coli* DNA polymerase I, includes *Streptococcus pneumoniae* DNA polymerase I(Lopèz *et al.*, 1989), Taq DNA polymerase I(Lawyer *et al.*, 1989), and the DNA polymerases of bacteriophages T7(Dunn and Studier, 1983), T5(Leavitt and Ito, 1989), and SPO2(Raden and Rutberg, 1984 ; Iwabe and Miyata, 1989). The family B DNA polymerases include a wide variety of DNA polymerases consisting of eukaryotic DNA polymerases of cellular, viral, and plasmid origin, and of bacteriophage DNA polymerases, such as T4(Spicer *et al.*, 1988 ; Reha-Krantz, 1988) and ϕ29(Yoshikawa and Ito, 1982). The family B DNA polymerases share regions of highly conserved amino acid sequences(Jung *et al.*, 1987 ; Leavitt and Ito, 1989 ; Bernad *et al.*, 1989 ; Spicer *et al.*, 1988 ; Reha-Krantz, 1988 ; Yoshikawa and Ito, 1982 ; Argos *et al.*, 1986 ; Larder *et al.*, 1986 ; Wong *et al.*, 1987 ; Bernad *et al.*, 1987 ; Wong *et al.*, 1988 ; Gibbs *et al.*, 1985 ; Wang *et al.*, 1989). These conserved regions occur in the same order in all family B DNA polymerase molecules. Therefore, it is likely that these sequence conservations are a consequence of their contribution to the DNA polymerase function and structure. However, little is known at present about functional roles of these highly conserved regions.

PRDI DNA polymerase is the smallest member of the family B DNA polymerase(Jung *et al.*, 1987). This DNA polymerase is specified by bacteriophage PRDl which infects a wide variety of gram-negative bacteria(Mindich and Bamford, 1988). Because PRDl is highly amenable to genetic and biochemical manipulation, it is a convenient model system with which to study structure-function relationships of DNA polymerase molecules. To determine the functional roles of the highly conserved regions of the family B DNA polymerases, we have initiated site-directed mutagenesis with PRDl DNA polymerase, and our results show that mutations at the conserved regions within PRDl DNA polymerase inactivate polymerase complementing activity and catalytic activity.

MATERIALS AND METHODS

Bacteria, Phages and Plasmids

E. coli NM522(*lac-proAB*), *thi*, *hsd5*, *supE*, *F'* [*proAB*, *lacI*^{q}, *lacZ*ΔM15](Maniatis *et al.*, 1982)

and *E. coli* RZ1032[HfrKL16, PO/45{lysA(61-61)}, *dut*1, *ung*1, *thi*1, *rel*A1](Kunkel, 1985) were used as the hosts for recombinant plasmids. The phagemid expression vector pEMBLex3(Sollazzo *et al.*, 1987) was obtained from the European Molecular Biology Laboratory, *E. coli* HB94(pLM2), *E. coli* HB94(pLM2, pLM3) and a mutant PRDl(sus 2) were a generous gift of Dr. L. Mindich at the Public Health Research Institute of the city of New York. The helper phage M13KO7(Kunkel, 1987) was purchased from International Biotechnologies, Inc.

DNA sequencing

The plasmid pLM3 was cleaved with Pst I and the resulting DNA fragment containing genes 1 and 8 of PRDl was recloned into Pst I site of bacteriophage M13mp9(Messing and Viere, 1982). We then generated a nested set of deletions according to the method of Dale *et al.*(1985). Each set of deletions was sequenced by the dideoxy chain termination(Sanger *et al.*, 1977), some regions were sequenced by the method of Maxam and Gilbert(1980).

Computation

Protein data base searches were performed by a FASTP microcomputer program(Lipman and Pearson, 1985). DNA polymerase nucleotide sequences were obtained from GenBank database and translated and screened against consensus sequence with the use of the IBI(International Biotechnologies, Inc.) sequence analysis system program.

Cloning the PRDl DNA polymerase gene into phagemid expression vector

3kb DNA fragment containing genes 1 and 8 and 2kb DNA fragment containing gene 1 were cloned into the expression vector pEMBLex3(pEJ 3 and pEJ2, respectively). The gene product was analyzed on SDS-polyacrylamide gel. The expression of cloned genes 1 and 8 was also analyzed by dGMP-terminal protein complex formation.

Site-directed mutagenesis

Site-specific deletion and site-directed mutagenesis of pEJG were done according to the method of Kunkel(1985).

Complementation Test

Phage PRDl infects a variety of gram-negative bacteria that harbor plasmids belonging to the P, N or W incompatibility group(Bradley and Rutherford, 1975). To test the ability of a mutant PRDl DNA polymerase to complement PRDl sus2 mutant, defective in the DNA polymerase gene, various recombinant plasmids were transferred into *E. coli* HB94(pLM2) cells. Cells containing both pLM2 and the recombinant pEMBLex3 clone were selected on plates containing kanamycin(100 μg/m*l*) and ampicillin(100 μg/m*l*). For the complementation test, the PRDl sus2 was plated with these bacteria using plates containing both antibiotics.

Preparation of Cell Extracts

Cell extracts were prepared as described previously(Yoo and Ito, 1989 ; Watabe *et al.*, 1984). The purification of PRDl DNA polymerase was essentially accomplished as described by Watabe *et al*(1984).

Assay conditions for DNA polymerase and for the Exonuclease

The DNA polymerase and the 3' to 5' exonuclease activities were assayed as described previously(Watabe *et al.*, 1984).

1. DNA polymerase assay : The reaction mixture(50 μ*l*) contained 50 mM Tris-HCl(pH 7.6), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 15 μM ^3H-dTTP(600 cpm/pmole). 5% glycerol, 0.5 mg/m*l* BSA, 14 μg/m*l* poly(dA)-oligo$(dT)_{12-18}$ and 6 μ*l* polymerase fraction. The mixture was incubated for 30 min at 30℃. Then, the acid insoluble materials were collected on Whatman GF/C glass paper and counted for radioactivity.

2. 3' to 5' exonuclease assay : The standard reaction mixture(50 μ*l*) contained 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 5% glycerol, 0.5 mg/m*l* BSA, 0.1-0.5 μg of heat denatured DNA and an appropriate amount of the nuclease fraction. After incubation of the reaction mixture at 30℃ for 20 min, the acid soluble materials were measured for their radioactivity.

RESULTS

Nucleotide Sequence and Predicted Amino Acid Sequence of PRDI DNA Polymerase.

In agreement with genetic studies by McGraw *et al.*(1983), nucleotide sequence analysis revealed that there are only two major open reading frames at the left end of the PRDI genome. One open reading frame corresponds to gene 8, the terminal protein gene consisting of 260 codons. The other one is gene 1, the DNA polymerase gene that consists of 554 condons. The nucleotide sequence of the PRDI DNA polymerase gene along with the predicted amino acid sequence and the flanking sequences are shown in Fig. 1. The calculated molecular weight of PRDI DNA polymerase is 63,300, which is in rough agreement with the reported size(Mindich *et al.*, 1982). This protein contains approximately equal numbers of acidic(14.2%) and basic(15.7%) amino acid residues and a high percentage(41.2%) of hydrophobic residues.

The initiation codon ATG for the PRDI DNA polymerase gene is precedes by a region complementary to the 3'-terminus of *Escherichia coil* 16S rRNA.

Homology to the other DNA Polymerase.

```
     AAA CGC GGC TAT GGC AGC AAG GGG GTT TAA GAT ATG CCG CGC CGT TCC CGT AAA AAG GTG GAA TAT AAA ATT GCC
                                                  M   P   R   R   S   R   K   K   V   E   Y   K   I   A
15   GCC TTT GAC TTT GAA ACT GAC CCT TTC AAG CAT GAC CGA ATC CCT AAA CCG TTT TCA TGG GGT TTT TAT AAT GGC
      A   F   D   F   E   T   D   P   F   K   H   D   R   I   P   K   P   F   S   W   G   F   Y   N   G
40   GAA ATT TAT AAA GAC TAT TGG GGC GAT GAT TGC ATA GAA CAG TTT ATT TAC TGG CTG GAT ACC ATA GAA GAA CCG
      E   I   Y   K   D   Y   W   G   D   D   C   I   E   Q   F   I   Y   W   L   D   T   I   E   E   P
65   CAC GTT ATA TAC GCT CAT AAC GGC GGC AAG TTT GAT TTT CTT TTT CTC ATG AAA TAC TTT CGC GGG AAA TTG AAA
      H   V   I   Y   A   H   N   G   G   K   F   D   F   L   F   L   M   K   Y   F   R   G   K   L   K
90   ATA GTT AAT GGG CGT ATT TTG GAA GTA GAA CAC GGC ATC CAT AAA TTC CGC GAT AGT TAT GCA ATC CTG CCG GTG
      I   V   N   G   R   I   L   E   V   E   H   G   I   H   K   F   R   D   S   Y   A   I   L   P   V
115  CCG CTT GCT GCC AGC GAT GAA AAG ATA GAA ATT GAT TAT GGC AAG ATG GAA AGG GAA ACA CGC GAA CAG CAC AAG
      P   L   A   A   S   D   E   K   I   E   I   D   Y   G   K   M   E   R   E   T   R   E   Q   H   K
140  GCG GAA ATT TTA GAA TAC CTG AAA GGC GAT TGT GTA ACC CTG CAT AAA ATG GTT TCT TTA TTT ATT GCT GAA TTT
      A   E   I   L   E   Y   L   K   G   D   C   V   T   L   H   K   M   V   S   L   F   I   A   E   F
165  GGA ATG CGC CTA ACC ATA GGC GGT ACG GCA ATG AAT GAA TTA AAA CAG TTC CAC CCT TAT GAC CCT GTG CGC AAA
      G   M   R   L   T   I   G   G   T   A   M   N   E   L   K   Q   F   H   P   Y   D   P   V   R   K
190  GGC TTT GAT GAA GCC ATG CGC CCC TTT TAT TTT GGC GGA AGG TGC CAA GCA TTC GAG AAA GGA ATA ATT GAA GAT
      G   F   D   E   A   M   R   P   F   Y   F   G   G   R   C   Q   A   F   E   K   G   I   I   E   D
215  GAT ATA AAA GTT TAT GAT GTT AAT AGT ATG TAC CCC CAT GCT ATG CGA AAT TTC CGC CAT CCT TTC AGC GAT GAA
      D   I   K   V   Y   D   V   N   S   M   Y   P   H   A   M   R   N   F   R   H   P   F   S   D   E
240  TTT TAT GAA GCC AAT GAA ATA ACA GAA GAA ACT TAT TTT ATT GAA TGG GAA GGC GAG AAT AAC GGC GCG GTG CCT
      F   Y   E   A   N   E   I   T   E   E   T   Y   F   I   E   W   E   G   E   N   N   G   A   V   P
265  GTT AGG ACT AAA ACA GGT TTA GAC TTT AAT CAG CGT AGC GGC ATT TTC CAT ACG TCA ATC CAT GAA TGG CGG GCG
      V   R   T   K   T   G   L   D   F   N   Q   R   S   G   I   F   H   T   S   I   H   E   W   R   A
290  GGT ATT GAT ACC GGC ACG ATT AAA CCT AAT CGG ATT ATA AGG ACA ATC AAT TTT ACT GAA ACA ACC ACT TTC GGC
      G   I   D   T   G   T   I   K   P   N   R   I   I   R   T   I   N   F   T   E   T   T   T   F   G
315  GCA TTC ATT GAC CAT TTC TTT AGC AAG CGT GAC GCT GCC AAA AAG GCG GGT GAT TTA TTC CAC AAT ATT TTT TAT
      A   F   I   D   H   F   F   S   K   R   D   A   A   K   K   A   G   D   L   F   H   N   I   F   Y
340  AAA CTG ATT TTA AAT AGC AGT TAT GGG AAG TTT GCA CAA AAC CCC GAA AAT TAT AAA GAG TGG TGC ATA ACG GAA
      K   L   I   L   N   S   S   Y   G   K   F   A   Q   N   P   E   N   Y   K   E   W   C   I   T   E
365  GGC GGC ATT TAT TTA GAA GGC TAT GAC GGC GAA GGG TGC GAA GTA CAG GAA CAT TTA GAC TAT ATT TTA TGG GGT
      G   G   I   Y   L   E   G   Y   D   G   E   G   C   E   V   Q   E   H   L   D   Y   I   L   W   G
390  AGG CCC GCT GAA ATG TTT AAT TAT TTT AAC GTG GCA GTG GCG GCA AGT ATT ACA GGC GCG GCC CGT TCC GTT TTA
      R   P   A   E   M   F   N   Y   F   N   V   A   V   A   A   S   I   T   G   A   A   R   S   V   L
415  TTG CGG GCA TTG GCG CAA GCG GAA AGG CCG CTT TAT TGT GAC ACT GAT TCT ATT ATT TGC CGT GAT TTA AAA AAT
      L   R   A   L   A   Q   A   E   R   P   L   Y   C   D   T   D   S   I   I   C   R   D   L   K   N
440  GTT CCG CTT GAC GCT TAC CAG CTA GGC GCG TGG GAT TTG GAA GCA ACC GGC GAT AAA ATA GCG ATT GCC GGT AAA
      V   P   L   D   A   Y   Q   L   G   A   W   D   L   E   A   T   G   D   K   I   A   I   A   G   K
465  AAA TTA TAT GCG CTT TAC GCT GGT GAT AAT TGC GTT AAA ATT GCA AGT AAG GGG GCT AGT CTG GTT CCG CGT GAT
      K   L   Y   A   L   Y   A   G   D   N   C   V   K   I   A   S   K   G   A   S   L   V   P   R   D
490  ATT GGG TTT TTA ATG CCC CCG GAT ATG GAA CCG AAA GCC GCC AAA AAG GTA GCG CAA CAA AAG GCT AAA AAT ATT
      I   G   F   L   M   P   P   D   M   E   P   K   A   A   K   K   V   A   Q   Q   K   A   K   N   I
515  GGT GGC GAG AAA ATT TTA AAG GTG GCT AAT GGC GGC GTG TAT GAT TTT GTA AAT GAT GCC CCG TCA TTT AAG CTA
      G   G   E   K   I   L   K   V   A   N   G   G   V   Y   D   F   V   N   D   A   P   S   F   K   L
540  AAT GGC AAC GTG CAA TTT ATC AAG CGC ACA ATC AAA GGA ACA TAA AAT GCA ATA TAC ACT TTG GGA TAT TAT CAG
      N   G   N   V   Q   F   I   K   R   T   I   K   G   T   -
```

Fig. 1. The nucleotide sequence and predicted amino acid sequence of the DNA polymerase. The Shine-Dalgarno sequence is underlined. Start and termination codons are boxed. Amino acid number is designated in the left side of the Figure.

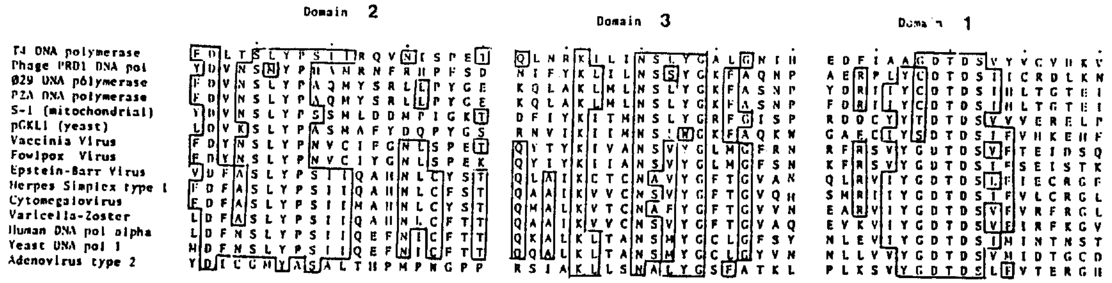

Fig. 2. Regions of homology between the PRDl DNA polymerase and other DNA polymerases. Regions were designated 1, 2, and 3 according to Larder *et al.*(1987).

Recently, several laboratories have noted partial homologies among various viral DNA polymerase (Jung *et al.*, 1987 ; Bernard *et al.*, 1987 ; Wong *et al.*, 1988 ; Larder *et al.*, 1987). When the amino acid sequence of PRDl DNA polymerase is compared with that of other viral DNA polymerases, three partially homologus regions are evident. Fig. 2 shows the amino acid sequence alignment of the three conserved regions appear to be conserved in all DNA polymerase compared, regardless of size. Our computer search revealed that phage T4 DNA polymerase also contains three conserved regions.

Since the DNA polymerases of T4, herpes virus, and vaccinia virus are not the protein-primed DNA polymerase, these conserved regions are clearly not unique to the DNA polymerases which use protein primers. Rather, it is likely that these 3 conserved regions represent domains which perform important functions common to many DNA polymerases. These function could include DNA chain elongation, the 3' to 5' exonuclease activity, DNA binding, nucleotide triphosphate binding or pyrophosphate binding.

Cloning PRDl DNA Polymerase Gene into the Expression Vector.

Total protein of cell was analyzed on SDS-polyacrylamide gels. Cells transformed with the recombinant plasmid pEJ3 produce two new protein bands with approximate moleular weights of 68 kD and 28 kD. These are consistent with reported values for PRDl DNA Polymerase and terminal protein(Mindich *et al.*, 1982), and roughly agree with the 63 kD and 29 kD value deduced from the nucleotide sequence(Jung *et al.*, 1987 ; Hsieh *et al.*, 1987). The amount of DNA synthesis occurring in *E. coli* extracts harboring pEJ3 was found to be five times higher than that found in *E. coli* extracts harboring pLM3 and five hundred times higher than PRDl infected cells(Bamford and Mindich, 1984).

The expression of cloned genes was also analyzed by characterization of dGMD-terminal protein complex formation because this reaction is very specific to the protein-primed PRDl DNA polymerase. Based on autoradiographic results, *E. coli* harboring pEJ3 has much more PRD1 terminal protein than *E. coli* harboring pLM3. These results indicate that PRDl DNA polymerase and terminal protein are functionally expressed in cells containing pEJ3.

In order to overexpress only PRDl DNA polymerase, 2 kb DNA fragment was cloned into pEMBLex3(pEJ2). Expression of PRDl DNA polymerase was very low, based on analysis of coomassie blue-stained SDS polyacrylamide protein gel. To remove a small segment of 3' end coding sequence for terminal protein, the ribosome binding site of the DNA polymerase gene, and the small N-terminal sequence of the lac Z gene of pEJ2, 57 bp of pEJ2 was deleted site-specifically(pEJG). Then, PRDl DNA polymerase gene remained under the control of lambda pR promotor and utilized the ribosome-binding site of the lambda cro gene. The expression level of DNA polymerase from pEJG was slightly better than that observed in pEJ3.

Site-specific Mutagenesis of Region 1, 2 and 3

within the PRDl DNA Polymerase.

In order to evaluate the function of the conserved regions, one of the most highly conserved amino acid residues within each domain was changed ; (1) In domain 2, aspartic acid and tyrosine residues are present in all members of the group. Initially, we focused on the negatively charged aspartic acid-219. The codon for this amino acid is GAT. Alternation of a nucleotide at any of the three possible positions can generate 8 different amino acid substitutions. Among these 8 amino acids, only valine is hydrophoic, Conversion of aspartic acid to valine will result in a change in charge and hydrophobicity. To make this change, GAT was converted to GTT in domain 2(Fig. 3).

(2) In domain 3, lysine, aspartic acid and glycine are conserved in each polymerase. Based on the above idea, lysine-340 was converted to isoleucine in domain 3(Fig. 4).

(3) In domain 1, serine, and aspartic acids at two positons are 100 percent conserved. Aspartic

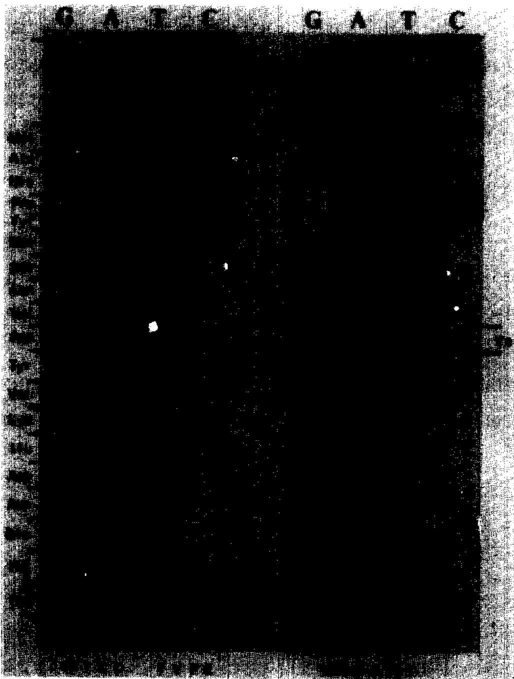

Fig. 3. DNA sequence of mutant(pEJG2) and wild-type clones in the region of mutation. Arrow indicates the mutated sequence.

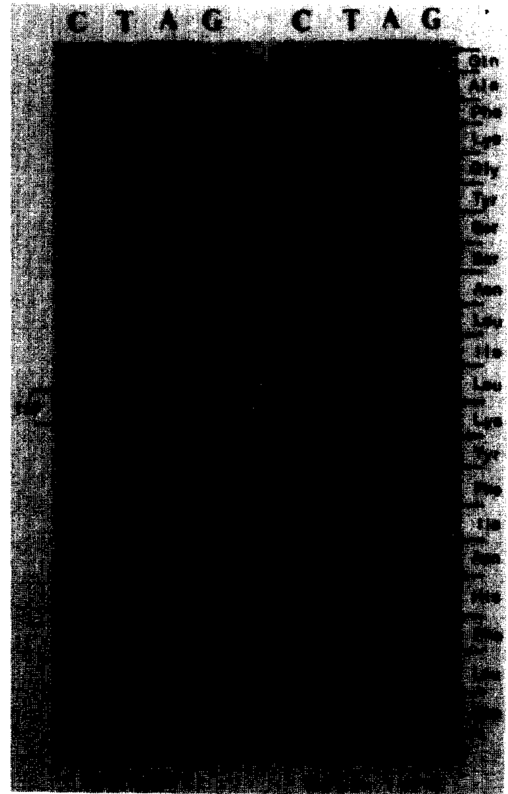

Fig. 4. DNA sequence of mutant(pEJG3) and wild-type clones in the region of mutation. Arrow indicates the mutated sequence.

acid-428 and serine-432 were changed to tyrosine and threonine, respectively, in mutant pEJGl(Fig. 5). In mutant pEJGam, tyrosine was changed to an amber codon(Fig. 6).

Complementation

PRDl can infect only those host cells which contain the plasmid pLM2. The receptor for the PRDl virion is believed to be encoded on this plasmid. To test the function of the cloned PRDl DNA polymerase generated in this work, the ability of various recombinant pEMBLex3 plasmids to complement PRDl(sus2) mutants was tested. This mutant strain of PRDl has an amber mutation which has been mapped within the DNA polymerase gene(Mindich and Mcgraw, 1983). In a host which does not contain a tRNA suppressor for this phage, viral replication will only occur if PRDl DNA polymerase is supplied *in trans*. The *E. coli*

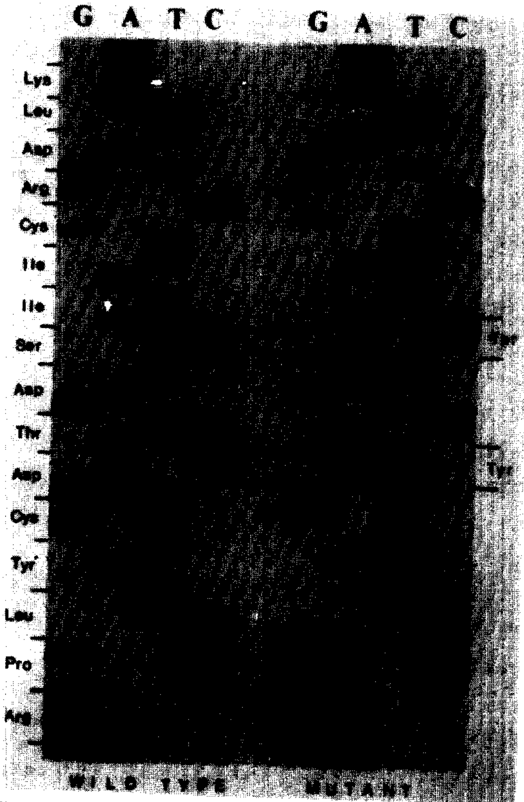

Fig. 5. DNA sequence of mutant(pEJG1) and wild-type clones in the region of mutation. Arrow indicates the mutated sequence.

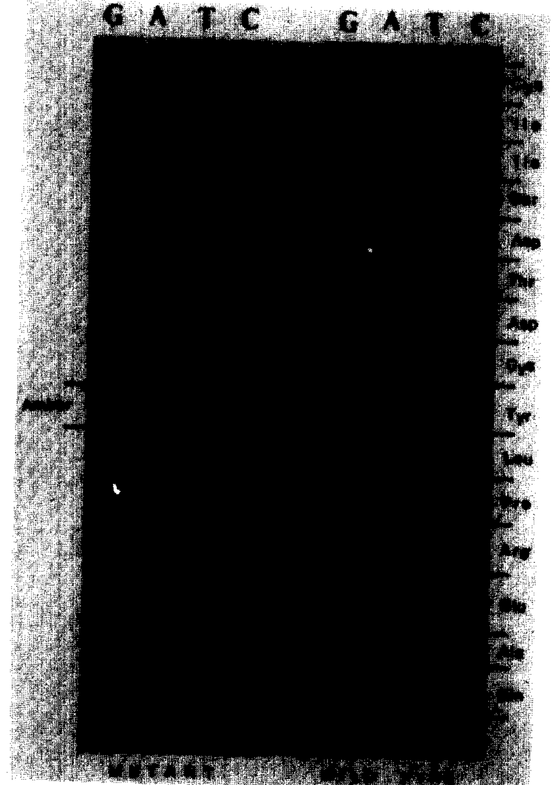

Fig. 6. DNA sequence of mutant(pEJGam) and wild-type clones in the region of mutation. Arrow indicates the mutated sequence.

host cell HB94(pLM2) was transformed with various pEMBLex3 plasmids isolated from plates containing kanamycin and ampicillin at 42℃. In initial control experiments, it was determined that PRDl(sus2) could produce plaques efficiently on a sus^{+} host(*S. Typhimurium* DB7156) but gave only a very low background of revertant plaque formation on the non-permissive sus^{-} host HB94 (pLM2). The PRDl(sus2) mutant formed a large number of plaques only on *E. coli* HB94(pLM2) which contained pEJG and pEJG2, but not pLM3, pEJG1, pEJG3 or pEJGam. Although pEJG2 can complement PRDl(sus2), the number of plaques obtained was 1% of that obtained with pEJG. To confirm the presence of the pLM2 plasmids, wild type PRDl phage was also used. Wild type PRDl formed plagues efficiently on all host cells which were used in these experiment.

Assay for DNA polymerase Activity

To characterize PRDl mutant DNA polymerases, we have partially purified each enzyme. During the course of PRDl DNA polymerase can be separated from the bulk of host cell DNA polymerase by DEAE-cellulose column chromatography. Figure 7 shows DEAE-cellose chromatography of DNA polymerases in extracts of the *E. coli* cells containing wild type and mutant PRDl DNA polymerase gene. The activity of the wild type PRDl DNA polymerase was eluted from the column after 0.3 M NaCl. Only trace polymerase activities were detected with 3 mutant extracts(Fig. 7). To determine the effect of mutations in the conserved regions on the 3' to 5' exonuclease activity of the PRDl DNA polymerase, we assayed each column fractions for exonuclease activity using the method described previously(Yoo and Ito, 1989).

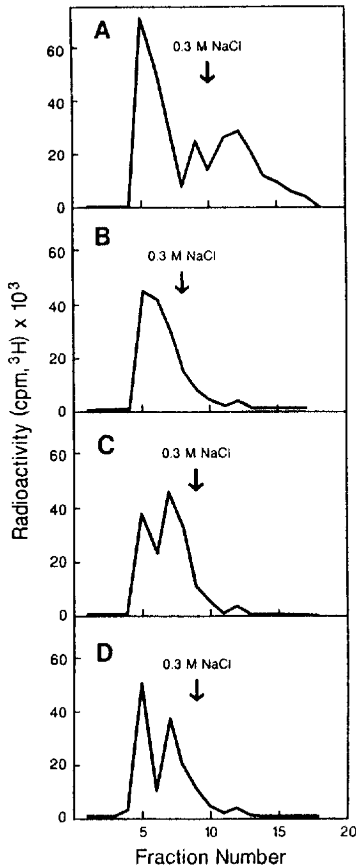

Fig. 7. DEAE-Cellulose chromatography. Cell extract (1 m*l*) was dialyzed for 8 to 10 hrs against 250 m*l* of Tris-HCl buffer(20 mM, pH 7.4) containing 1 mM EDTA. 1 mM DTT. A column of whatman DE-52 DEAE-Cellulose(10×0.9 cm) was prepared and equilibrated with the same buffer. The dialyzate was applied to the column, which was then rinsed 20 m*l* of the Tris-HCl buffer and eluted with buffer with increasing NaCl concentration. Fractions(1 m*l* each) were collected and assayed for DNA polymerase activity. The DNA polymerase activity was determined using synthetic poly (dA)-oligo$(dT)_{12-18}$ as described in Materials and Methods.

While levels of the exonuclease activities fluctuated somewhat, all mutant DNA polymerases were active. From these results we concluded that mutations introduced into highly conserved regions inactivate DNA polymerase catalytic activities, but not exonuclease activities.

Very recently, Dorsky and Crumpacker(1990) reported that point mutations introduce into region 1 of the Herpes simplex virus type 1 DNA polymerase abolished the enzyme activity. During the preparation of this paper, we also learned that Bernad *et al.*(1990) have performed similar studies with ϕ29 DNA polymerase. They have suggested that region 1 of the family B DNA polymerases is involved in metal binding at the polymerase active site.

DISCUSSION

The predicted amino acid sequence indicates that this polymerase contains 553 amino acids. To our knowledge, this is the smallest DNA polymerase ever isolated from prokaryotic cells.

Several groups have recently reported partial amino acid homologies among various DNA polymerases. A comparative amino acid sequence analysis between the PRD1 DNA polymerase and other DNA polymerases was performed. These results are consistent with others and show that PRD1 DNA polymerase shares three highly conserved homologous regions with the other DNA polymerases. This suggests that neither insertion nor frameshift mutation have occurred within these regions since the separation of eukaryotes from prokaryotes more than 2 billion years ago. This also implies that both protein-primed and RNA-primed DNA polymerases share a common ancestor.

Based on evolutionary relatedness, we propose to classify DNA polymerases as follows. Adopting Doolittle's terminology(Doolittle, 1981), we designated *E. coli* DNA polymerase I and T7 DNA polymerase to be DNA polymerase family A. All DNA polymerases carrying the three conserved regions described above are designated family B. It is possible that there are other families of DNA polymerases, which do not have any homology with family A or B enzymes. Such DNA polymerases can be classified in families C, D, and so on.

The 3 kd PRD1 insert which was used in the construction of pEJ3 also cloned into other expression vectors such as pPLc236, pKK233-2, and pT

7-6. pPLc236 is a derivative of pBR322 and carries the lambda pL promoter(Remaut *et al.*, 1981). pKK 233-2 is also a pBR322 derivative carring the tac promoter and a lac Z' ribosome binding site (Amam and Brosius, 1985). The pT7-6 plasmid contains a T7 RNA promoter.

Although the 3 kb PRDl insert was cloned into several expression vectors, pEMBLex3 was the only vector which could produce enough PRDl DNA polymerase to visualize distinctly as a protein band on SDS-polyacrylamide protein gels. High expression of PRDl DNA polymerase in pEMBLex3 than other plasmid vectors may be due to the difference in copy number. pEMBLex3 is a pUC plasmid derivative and the others are pBR 322 derivatives. The copy number of pUC is approximately 200 plasmids per cell while that of pBR 322 is a approximately 50 copies per cell(Balbas *et al.*, 1988). Despite the importance of DNA polymerase, only a few DNA polymerase gene have been sequenced, and very few DNA polymerase have been cloned into expression systems for subsequent study. Up to this time only the Klenow fragment of *E. coli* DNA Pol I(Joyce and Grindly, 1983), phage subunit of T7 DNA polymerase(Tabor and Richardson, 1985 ; Reutimann *et al.*, 19 85), phage T4 DNA polymerase(Lin *et al.*, 1987) and phage φ29 DNA polymerase(Blanco *et al.*, 1984) have been cloned and the proteins purified for study from overexpressing cells.

As a result of several studies, two tentative conclusions related to localizing the enzymatics activities of the DNA polymerase family B have been drawn : one is that conserved region 2 and 3 may represent dNTP binding sites.

The second conclusion drawn from past studies is that the polymerase activity appears to reside in the carboxy-terminus and the 3'-5' exonuclease function is located within the amino-terminus. A new conserved region between DNA polymerase family A and B in the N-terminal portion has been identified which contains four highly conserved amino acids known to be involved in the 3'-5' exonuclease active site of Klenow fragment of *E. coli* polymerase I(Ollis *et al.*, 1985). This strongly supports the idea that the 3'-5' exonuclease activity is located within the amino terminus.

There exists no experimental evidence concerning the function of region 1. It may be that mutant cannot be isolated within this region because any amino acid substitution is lethal. It seems likely that the function of regions 1 will be different from that of regions 2 and 3 because more than ten mutants have been isolated in region 2 and 3 from random screening but none were found which mapped to region 1.

Argos(1988) proposes that the amino acid sequence(YGDTDS) of region 1 is very similar to the RNA directed polymerase core hydrophobic (YGDD-), including the two five-residues hydrophobic flanking spans and secondary structure of this region. The major difference between two aspartic acids in the conserved region 1 sequence of DNA polymerase family B. He also suggests that the two conserved aspartates may bind magnesium cation as well as act catalytically in the polymerization reaction.

REFERENCE

1. Amann, E. & Brosius, J.(1985) *Gene* **40**, 183-190
2. Argos, P., Tucker, A.D. & Philipson, I.(1986) *J. Virol.* **149**, 208-216
3. Balbas, P., Soberon, X., Bolivar, F. & Rodriguez, R.L.(1988) In "Vectors"(Rodriguez, R.L. & Denhardt, D.T. eds.) pp. 5-42. Butlerworths, Stoneham, MA.
4. Bamford, D.H. & Mindich, L.(1984) *J. Virol.* **50**, 309-315
5. Bernad, A., Blanco, L., Lazaaro, J.M. Martin, G. & Salas, M.(1989) *Cell* **59**, 219-228
6. Bernad, A., Lazaro, J.M., Salas, M. & Blanco, L.(1990) *Proc. Natl. Acad. Sci.* U.S.A. **87**, 4610-4614
7. Blanco, L., Garcia, J.A. & Salas, M.(1984) *Gene* **29**, 33-40
8. Dale, R.M. K., McClyre, B.A. & Houchins, J.P. (1985) *Plasmid* **13**, 31-40

9. Doolittle, R.F.(1981) *Science* **214**, 140-159
10. Dorsky, D.I. & Crumpacker, C.S.(1990) *J. Virol.* **64**, 1394-1397
11. Dunn, J.J. & Studier, F.W.(1983) *J. Mol. Biol.* **166**, 477-535
12. Gibbs, J.S. Chiou, H.C., Hall, J.D., Mount, D.W., Retondo, M. J., Weller, D.K. & Coen, D.M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7969-7973
13. Hsieh, J., Jung, G., Leavitt, M.C. & Ito, J. (1987) *Nucl. Acids Res.* **15**, 8999-9009
14. Iwabe, N., Kuma, K., & Miyata, T.(1989) *Nucl. Acids Res.* **7**, 8866
15. Joyce, C.M. & Grindly, N.D.F.(1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1830-1834
16. Jung, G., Leavitt, M.C., Hsieh, J.-C. & Ito, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8287-8291
17. Kunkel, T.A.(1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488-492
18. Kunkel, T.A., Roberts, J.D. & Zakour, R.A. (1987) *Methods in Enzymol.* **154**, 367-382
19. Larder, B.A., Kemp, S.D. & Darby, G.(1987) *EMBO J.* **6**, 169-175
20. Laywer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummind, R. & Gelfand, D.H.(1989) *J. Biol. Chem.* **264**, 6424-6437
21. Leavitt, M.C. & Ito, J.(1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4465-4469
22. Lin, T.-C., Rush, J., Spicer, E.K. & Konigsberg, W.H.(1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7000-7004
23. Lipman, D.J. & Pearson, W.R.(1985) *Science* **227**, 1435-1441
24. Lopéz, P., Martinez, S., Diaz, A., Espinoza, M. & Lacks, S.A.(1989) *J. Biol. Chem.* **264**, 4255-4263
25. Maxam, A.M. & Gilbert, W.(1980) *Methods in Enzymol.* **65**, 499-5600
26. McGraw, T., Yang, H. & Mindich, L.(1983) *Mol. Gen. Genet.* **190**, 237-244
27. Messing, J. & Vieira, J.(1982) *Gene* **19**, 269-276
28. Mindich, L., Bamford, D.H., Goldthwarte, D., Laverty, M. & Mackenzie, G.(1982) *J. Virol.* **44**, 1013-1020
29. Mindich, L. & McGraw, T.(1983) *Mol. Gen. Genet.* **190**, 233-236
30. Mindich, L. & Bamford, D.H.(1988) In R. Calender(eds), The bacteriophages, vol II, p 149-262, Plenum Publishing corp., New York
31. Ollis, D.L., Klijne, C. & Steitz, T.A.(1985) *Nature* **313**, 818-820
32. Raden, B. & Rutberg, L.(1984) *J. Virol.* **52**, 9-15
33. Reha-Lrantz, L.J.(1988) *J. Mol. Biol.* **202**, 711-724
34. Remaut, E., Stanssen, P. & Fiers, W.(1981) *Gene* **15**, 81-93
35. Reutiman, H., Sjoberg, B.-M. & Holmgren, A. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6783-6787
36. Sanger, F., Micklen, S. & Coulson, A.R.(1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467
37. Sollazo, M., Frank, R. & Cesareni, G.(1985) *Gene* **37**, 199-206
38. Spicer, E.K., Rush, J., Fung, C., Reha-Krantz, L. J., Karan, J.D. & Konigsberg, W.H.(1988) *J. Biol. Chem.* **263**, 7478-7486
39. Tabor, S. & Richardson, C.C.(1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1074-1978
40. Wang, T.S, -F., Wong, S.W. & Korn, D.(1989) *FASEB J.* **3**, 14-21
41. Watabe, K., Lensch, M. & Ito, J.(1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5734-5378
42. Watabe, K., Lensch, M. & Ito, J.(1984) *Biochem. Biophys. Res. Comm.* **123**, 1019-1026
43. Wong, S.W., Wall, A.F., Yuan, P.M., Arai, N., Pearson, B.E., Arai, K., Korn, D., Hunkapiller, M.W. & Wang, T.S. -F.(1988) *EMBO J.* **7**, 37-470
44. Yoo, S. & Ito, J.(1989) *J. Virol.* **170**, 442-449
45. Yoshikawa, H. & Ito, J.(1982) *Gene* **17**, 323-335