DNA Polymerases of Bacterial Origin

Hari Ram, Lebin Thomas, Alok Kumar and Ved Pal Singh*

Applied Microbiology and Biotechnology Laboratory, Department of Botany, University of Delhi, Delhi-110 007, India.

(Received: 11 September 2013; accepted: 12 November 2013)

In nature, prolonged cell survival is governed by the accurate transmission of DNA material from one generation to the next. Accuracy of the transmission is achieved by the concerted action of several enzymes and proteins of DNA replication and DNA repair mechanisms. In both prokaryotes and eukaryotes, various enzymes contribute to the fidelity of replication. Bacteria contain five DNA polymerase enzymes (I, II, III, IV and V) which show active contribution in the fidelity of DNA replication. DNA polymerase I is the first identified DNA polymerase from *Escherichia coli* which is involved in the DNA replication by Okazaki fragment processing and nucleotide excision repair. DNA polymerase III is the main chromosomal replicative polymerase with high processivity and fidelity. DNA polymerase II, IV and V are SOS regulon regulated DNA repair enzymes, which are induced in response to DNA damage.

Key words: Processivity, Fidelity, Trans-lesion synthesis (TLS).

The prolonged cell survival which is dependent on the accurate transmission of DNA from one cell to daughter cells, is accomplished by the combined action of a group of diverse DNA polymerases. All these polymerases ensure the faithful copy of the information packed genome sequences into the progeny cell nucleic acid. Several mechanisms ensure fidelity during replication process, such as base selection, proofreading and mismatch repair, which involve sequential activity of a diverse group of DNA polymerases. The combined efforts of all these mechanisms maintain a low error rate of $10^{-9}$ to $10^{-11}$ per base pair. Various models that support this remarkable accuracy have been proposed. An earlier model by Watson and Crick suggests the basis of selection of correct nucleotides from a structurally similar pool of nucleotides, which is depend on the structure of DNA, in which hydrogen bond mediated A-T and G-C base pairing could be responsible for the selection of correct nucleotides. However, another model was based on the geometry of base pairs which involves the selection of correct nucleotides over the incorrect nucleotides. In this model, the active site of DNA polymerase could accept and accommodate the correctly fitting base pairs of Watson-Crick (A-T and G-C) model, while incorrect base pairs would be rejected.

Ever since the discovery of DNA polymerase I in *Escherichia coli*, it has served as a prototype, which has resulted in the discovery of many more DNA polymerases within a single cell, along with the information on their biological roles during replication processes. As a consequence of all these efforts in the present time, a large amount of information about the biological role, along with the structure and composition of various DNA polymerases has been gathered from organisms occupying different habitats. In our laboratory, we have been working on the comparative analysis of the alpha subunit of DNA polymerase III among extremophilic (psychrophilic, mesophilic and thermophilic) bacteria.
In both prokaryotes and eukaryotes, the replication of chromosomal DNA is ensured by a group of DNA polymerase activity. In general, polymerase enzymes have been classified into seven families on the basis of amino acid sequence homology and analysis of the crystal structure, viz., Family A, Family B, Family C, Family D, Family X, Family Y and RT Family\(^8,9,10\).

**Family A**

Family A of DNA polymerases encompasses polymerases that show amino acid sequence homology to the \textit{pol A} gene encoded DNA polymerase I of \textit{E. coli}\(^8\). This family includes the bacterial DNA polymerase I, DNA polymerases of T3, T5 and T7 bacteriophages and eukaryotic (mitochondrial DNA polymerase \(\gamma\)), which function in the replication and repair of DNA. They ensure accuracy of replication through the nucleotide excision repair and through removal of ribonucleotides during Okazaki fragment processing during lagging strand synthesis\(^11\).

**Family B**

Family B of DNA polymerases includes those polymerases that show amino acid sequence homology to the \textit{pol B} gene encoded DNA polymerase II of \textit{E. coli}\(^8\). There is amino acid sequence homology between \textit{E. coli} DNA polymerase II and alpha-like DNA polymerase of the eukaryotes replicative DNA polymerases\(^12\). This family includes eukaryotic replicative DNA polymerase \(\alpha\), bacteriophage DNA polymerases (PRD1, \(\phi 29\), M2, T4) and viral DNA polymerases\(^8\). Family B members in prokaryotes function as a part of the SOS regulon; whereas in eukaryotes, these polymerases function as a processive replicative DNA polymerase for chromosomal DNA replication\(^10\).

**Family C**

Family C encompasses those DNA polymerases that exhibit amino acid sequence homology to the \textit{pol C} gene encoded DNA polymerase III alpha subunit of \textit{E. coli}\(^9\). DNA polymerase III holoenzyme functions in the accurate and efficient duplication of chromosomal DNA and it consists of ten different subunits\(^13\), which through their cooperative action, assures the fidelity of replication. The amino acid sequence of the alpha subunit of DNA polymerase III from \textit{E. coli}, \textit{Salmonella typhimurium} and \textit{Bacillus subtilis} have shown significant sequence homology among them\(^8\).

Several decades long research work has revealed the structure, composition and arrangement of subunits in the holoenzyme complex. Today, different genes encoding the different subunits of the holoenzyme complex have been isolated, over-expressed, purified and through their reconstitution, the mode of holoenzyme assembling and polymerization process has also been revealed. The multisubunit DNA polymerase III holoenzyme structure is different from the single subunit enzymes of the family A and Family B DNA polymerases\(^14\).

**Family D**

Family D polymerases include the euryarchaea (subdomain of archaea) replicative DNA polymerase\(^10\). A heterodimer DNA polymerase pol II (Pol D) was discovered during the elucidation of DNA replication mechanism of euryarcheota, which consists of two subunits (DP1 and DP2). Archaea exhibits similarity to the eukaryotes in the information transmission machinery; whereas, on the other hand, it exhibits metabolic closeness to bacteria\(^15\). Small subunit DP 1 of DNA Pol D shows sequence homology to the eukaryotic DNA polymerase \(\delta\) small subunit, which supports the hypothesis that archaea and eukarya are sister branches, which earlier got branched from bacteria\(^16\).

**Family X**

Family X of polymerases has been defined by the eukaryotic DNA polymerase \(\xi\) that does not have significant sequence homology with DNA polymerases from other families. Therefore, this group was designated as X family\(^8\). DNA polymerase \(\xi\) is predominantly involved in DNA repair through the base excision repair pathway at abasic sites. This family includes a wide range of DNA polymerases with different specific activities, such as DNA polymerase (\(\beta\), \(\sigma\), \(\mu\), \(\lambda\)), yeast polymerase IV and virus polymerase X\(^10\).

**Family Y**

Family Y of DNA polymerases has been defined by super family of \(\text{Umu C, Din B, Rev 1 and Rad 30}\)\(^17\). These enzymes commit DNA synthesis with low fidelity on undamaged DNA and bypass DNA lesion caused by ultraviolet radiation and other mutagenic agents, which normally would stall replication by DNA polymerases of other families (A, B, C, D or X) due...
to fidelity of replicative DNA polymerase\(^{10,17}\).

Phylogenetic analysis reveals the distribution of the members of the family Y DNA polymerases among the three kingdoms of life. One of the superfamiliation genes, \textit{umu} \textit{C}, codes for the \textit{E. coli} Pol V, with different forms present in Gram-positive and Gram-negative bacteria. The \textit{umu} \textit{C} genes found only in prokaryotes. The gene \textit{din} \textit{B} is present among bacteria, eukaryotes and archaea, and codes for DNA polymerase IV in bacteria, whereas in eukaryotes it codes for pol \(\phi\) and pol k. The Rev1 and Rad 30 proteins are found only in eukaryotes, but absent in bacteria and archaea\(^{17}\).

**RT Family**

RT Family includes retrovirus Reverse Transcriptase that converts single stranded RNA genome into double-stranded DNA genome\(^{18,10}\). Telomerase is also assigned to the RT family, which uses the integral RNA strand as template for the synthesis of telomere\(^{19}\). Telomerase plays an important role in the stability of the chromosome by addition of telomeric repeats to the chromosomal ends in eukaryotes.

**Bacterial DNA polymerases**

Among the seven families of DNA polymerases, bacterial DNA polymerases belong to only four families (A, B, C and Y). The discovery of DNA polymerase I of \textit{E. coli} has served as the prototype that has lead to the elucidation of the structural and biochemical evaluation of various other DNA polymerases of both prokaryotic and eukaryotic origin.

**(A) DNA Polymerase I**

Family A of DNA polymerases has been defined by the homology of the DNA polymerase I of \textit{E. coli} because this served as a prototype in revealing the mechanism of DNA polymerase enzymes\(^{1}\). DNA pol I represents the most abundant DNA polymerase of \textit{E. coli} (approximately 400 per cell). DNA pol I, encoded by \textit{pol} \textit{A} gene (approximately 3000 base pairs) is approximately 1000 amino acids long single polypeptide chain (109,000 MW). Partial proteolysis of DNA pol I has revealed that it is composed of a single polypeptide chain, which contains two distinct and independently functioning structural domains of 76,000 and 34,000 molecular weights\(^{20}\). Their C-terminal domain (called as Klenow fragment) has 5'-3' polymerase and 3'-5' proofreading activity, while the N-terminal domain has 5'-3' exonuclease activity\(^{21}\). These two independently functioning domains work in a cooperative strategy to fill-in the gaps formed during the Okazaki fragment processing of lagging strand of DNA synthesis\(^{13}\). DNA pol I uses exonuclease activity to remove ribonucleotides from lagging strand and then fills in the resulting gaps by its polymerase activity\(^4\). The multifunctional nature of DNA pol I also plays a significant role in DNA repair through excision repair by removing the DNA lesions caused by UV radiations (thymidine dimer), oxidative lesions (8-oxo guanine) and alkalyting lesions (4-methyl adenine)\(^{22}\). DNA pol I plays a central role in the restoration of inactive replication fork by participating in the recombination-dependent DNA replication which has been recognized as a central mechanism for the repair of double strand breaks as well as lesions caused by various DNA damaging agents\(^{23}\).

Crystal structures of DNA polymerase I from distantly related prokaryotes have been determined\(^{24}\) and their comparative analysis from these distantly related prokaryotes has shown an overall morphological resemblance, despite having less sequence homology\(^{1,24}\). This structural resemblance and sequence heterogeneity between enzymes in turn suggest that they have a common mechanism for the incorporation of nucleotides by producing identical folding patterns. DNA polymerase I is able to retro-transcribe RNA into DNA with accuracy, but is limited by low processivity\(^{25}\).

**(B) DNA Polymerase II**

Family B of DNA polymerases were defined by \textit{din} \textit{A (pol} \textit{B) gene encoded DNA polymerase II of \textit{E. coli}}\(^{9}\). This gene is damage inducible, and forms a part of the SOS response system, regulated by the repressor \textit{Lex A} at the transcriptional level\(^{26}\). Ever since its discovery in 1970 from the cell free extracts of \textit{E. coli} mutant cells (\textit{pol} \textit{A}-), the cellular role of DNA polymerase II has remained an enigma for a long period of time as compared to DNA polymerase I and III (27). Knippers\(^{27}\) found a cellular component, which was able to synthesize new DNA semi-conservatively at a rate similar to the DNA replication \textit{in vivo} by replicative DNA Polymerase.

Explosive molecular studies to date have revealed that cells have evolved different strategies to respond to challenges that cause damage to the
genetic material. Those challenges include the DNA lesions caused by numerous DNA damaging agents, such as UV-irradiation induced lesions \(^2\), oxidative lesions \(^2\), abasic lesions \(^26, 30\) and interstrand cross-links \(^2\), that may stall the replication fork when encountered by replicative DNA polymerases. After the replication fork encounters the DNA template lesion, it induces the expression of a set of damage inducible genes as a stress response (known as “SOS response”) to restart or/bypass the lesions for the recovery of DNA replication in an error free repair pathway. UV-irradiation transiently inhibits DNA replication by forming DNA lesions on DNA template, which is later restarted by the pivotal role of DNA polymerase II. Mutant lacking pol II exhibits delayed replication restart by 50 min, compared to the wild type, which shows that the replication restarts just after the quick stop \(^2\). DNA polymerase II under the SOS response bypasses the abasic site in vitro and also shows a seven fold increase in the enzyme level following the induction by nalidixic acid \(^26, 30\). DNA polymerase II plays an active role in the survival of cells in the oxidative damaged state. The pol B A1 mutant cells exhibited 5 to 10 fold higher sensitivity to oxidative lesion agents (H\(_2\)O\(_2\)), compared to the wild Pol B + cells \(^2\).

Cytotoxic lesions (such as inter-strand cross-links) are formed by antitumor agents (mustard, nitrogen mustard). These interstrand cross-links cause the cytotoxicity to the cell. Cells, to overcome this complication, rely on DNA repair through nucleotide excision repair (NER) and recombination. DNA polymerase II plays a significant role in repairing of inter-strand cross-links, in combination with the nucleotide excision repair pathway. This combination probably repairs the small fractions than the NER/recombination pathway \(^3\).

**(C) DNA Polymerase III**

Bacterial DNA polymerase III enzyme’s \(\alpha\)-subunit defines the family C of DNA polymerase \(^8\). DNA polymerase III is a major chromosomal replicase of \(E.\ coli\), which functions in the duplication of chromosomal DNA with high processivity and high speed \(^2\). DNA polymerase III is a holoenzyme complex and is composed of many multisubunits, including two polymerizing subunits with nine accessory subunits \(^14\). This multisubunit structure is responsible for the high fidelity of DNA replication, with an error rate of \(10^{-9}\) to \(10^{-10}\) per base pair replicated \(^6\). Holoenzyme shares amino acid sequence homology to replicases of most other systems, and also shows structural and functional similarities to the eukaryotic chromosomal replicase from yeast to human. DNA polymerase III holoenzyme is encoded by various dna genes. This multisubunit enzyme contains ten subunits, in which three subunits (\(\alpha\), \(\epsilon\) and \(\theta\)) form the core of the DNA polymerase \(^13\), dimer of \(\beta\) subunit forms the DNA sliding clamp, and \(g\) complex (\(\gamma\delta\gamma\psi\)) forms the clamp loader, which functions as a matchmaker \(^2\).

The three subunits (\(\alpha\), \(\epsilon\) and \(\theta\)) that form the polymerase core of the holoenzyme, are encoded by dna E, dna Q, mut D and hol E genes of \(E.\ coli\) respectively. The \(\alpha\)-subunit of the polymerase core shows the polymerase activity, while \(\beta\) subunit forms 3’-5’ proofreading exonuclease activity and \(\theta\) subunit stimulates the exonuclease activity of the subunit \(\epsilon\) \(^2\). The dimer of \(\beta\) subunit forms the \(\beta\) sliding clamp, which holds the holoenzyme on the DNA template for several thousand nucleotides polymerization events without dissociating from the template \(^5\). The loading of \(\beta\) sliding clamp is an ATP dependent pathway, which is catalyzed by the \(\gamma\) complex. \(\beta\) subunit dimer appears as a ring shaped structure with a central cavity, that can accommodate duplex DNA which confers the processivity of the DNA polymerase III \(^2\). The \(\gamma\) complex consists of five different subunits (\(\gamma\delta\gamma\psi\)), and they function as a molecular matchmaker in the loading of \(\beta\) sliding clamp by using energy obtained from the hydrolysis of ATP \(^3\). The genes encoding for the five different subunits have been isolated and identified. Subunit \(\gamma\) is encoded by \(dna\ X\) gene that binds to ATP molecule, \(\delta\) subunit is encoded by \(hol\ A\) gene, which binds to \(\beta\) subunit and \(\theta\) subunit is encoded by \(hol\ B\) gene, which act as cofactor for ATPase and stimulates clamp loading, while \(\psi\) and \(\psi\) subunits are encoded by \(hol\ C\) and \(hol\ D\), respectively, which help by binding to SSB, and by acting as a connecting bridge between \(\gamma\) and \(\gamma\) subunits for the formation of holoenzyme complex \(^4\).

**(D) DNA Polymerase IV**

The Y family of DNA polymerases is referred to as the error prone DNA polymerases, which are the products of a superfamily of \(umu\ C\),
DNA polymerase IV is reported to be a part of the SOS regulon, which has low fidelity owing to lack of its proofreading activity. The replication fork after the insertion of nucleotide on misaligned 3'-primer might get stalled, after which DNA polymerase IV might gain access and then rescue the replication fork by extending the bulged primer end. Exposure of cells to DNA damaging agents causes the induction of several proteins of the SOS regulon, and most of the genes of the SOS system are regulated by repression of Lex A protein, which is induced by self-cleavage after an association with Rec A molecule. This Rec A molecule is activated when it encounters a single-stranded DNA after DNA damage. The expression level of DNA pol IV gets increased after DNA damage, but it also results in an increased mutation rate. DNA pol IV participates in the DNA repair process by bypassing lesion and through trans-lesion synthesis of DNA. Since DNA pol IV is devoid of proofreading activity, it frequently generates mutations known as targeted mutagenesis. In addition to this, SOS induced activity also results in an enhanced mutagenesis in DNA that has not been exposed to DNA damaging agents. These mutations are called untargeted mutagenesis.

Environmental stress also induces genetic changes in organisms, as a response to ensure their survival in the prevailing environmental conditions. This genetic change is known as adaptive mutation. DNA pol IV has been shown to have a significant accountability in the adaptive mutation. There is relatively high number of DNA pol IV (about 250 molecules) as compared to replicative DNA polymerase III (about 10 to 20 molecules) in uninduced cells. Under adverse conditions, cells explore the use of error prone DNA polymerase to survive, such as during exponential phase, DNA pol IV has less effect on growth dependent mutation rate, but in stressed cells during adverse condition it has large effect on mutation. RNA polymerase sigma factor (RpoS) is a regulatory element in stress response of cell. The dependence of the DNA polymerase IV level, on the RpoS factor has been supported the role of DNA pol in adaptive mutation in stress conditions.

**CONCLUSION**

Bacteria are the wonderful system of the prokaryotic domain to study DNA polymerases for their role in DNA replication. Although, several studies have been done still, further detailed study on the role of DNA polymerases in the survival of organism and in the genetic stability during their life time needs to be carried out. Till today, five different DNA polymerases (I, II, III, IV and V) for the genetic integrity have been reported. The alpha subunit (a subunit of replicative DNA polymerase III), which plays a key role in the replication of DNA is being investigated in our research laboratory for its adaptability in bacteria thriving under extremophilic (psychrophilic, mesophilic and thermophilic) habitats.

**ACKNOWLEDGMENTS**

Financial supports in the form of J.R.F. from U.G.C. to Hari Ram, from C.S.I.R. to Lebin Thomas, S.R.F. from U.G.C. to Alok Kumar, and grant under R & D program from the University of Delhi, are gratefully acknowledged.
REFERENCES


