This is to certify that the

dissertation entitled

MOLECULAR CLONING, BACULOVIRUS EXPRESSION AND RECONSTITUTION OF *DROSOPHILA* MITOCHONDRIAL DNA POLYMERASE

presented by

Yuxun Wang

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Genetics

Laurie S. Kaguni
Major professor

Date 1/15/99

# MOLECULAR CLONING, BACULOVIRUS EXPRESSION AND RECONSTIUTION OF *DROSOPHILA* MITOCHONDRIAL DNA POLYMERASE

By

Yuxun Wang

A DISSERTATION

submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Genetics Program

1998

# ABSTRACT

## MOLECULAR CLONING, BACULOVIRUS EXPRESSION AND RECONSTITUTION OF *DROSOPHILA* MITOCHONDRIAL DNA POLYMERASE

By

Yuxun Wang

Molecular cloning of the accessory subunit of *Drosophila* mitochondrial DNA polymerase (pol γ) was accomplished, and its mammalian homologs were identified. Sequence similarity and structural features that are shared between the accessory subunit of *Drosophila* pol γ and its mammalian homologs suggest that the heterodimeric structure of pol γ is conserved from insect to man. Co-expression of the catalytic and accessory subunits encoded by recombinant baculoviruses reconstitutes the heterodimeric pol γ holoenzyme in both the mitochondria and cytoplasm of Sf-9 insect cells. Purified recombinant pol γ exhibits nearly identical physical and biochemical properties as the native enzyme from *Drosophila* embryos.

The roles of both subunits of *Drosophila* pol γ in template-primer DNA recognition and binding were studied by photochemical cross-linking and DNase I footprinting. The catalytic subunit contains the DNA binding activity and the DNA-binding domain was mapped to a 65 kDa fragment. Native pol γ was found to protect two helical turns (20 nt) of the primer strand when it forms a discrete and stable complex with template-primer

DNA. Mitochondrial single-stranded DNA-binding protein was found to enhance the specific interaction between pol γ and template-primer DNA.

To my wife, Yueying Cao and my son, Ruotao Wang

# ACKNOWLEDGEMENTS

I would like to start by thanking my thesis advisor and mentor, Dr. Laurie S. Kaguni, for her guidance, encouragement, and support during the past five years of my graduate studies. I have learned not only science from her but also parenting and other very important things. I would also like to thank members of my committee: Drs. Zachary Burton, Tom Friedman, Shelagh Ferguson-Miller, Suzanne Thiem, and Vilma Yuzbasiyan-Gurkan for their continuing support and scientific input.

I also want to thank Dr. Jon Kaguni for many helpful suggestions in our regular research group meeting, and Dr. John Wang for allowing me to use the cell culture equipment in his laboratory.

Many thanks to members of both Kaguni laboratories: Dr. Anuradha Alahari, Carol Farr, Li Fan, Inmaculada Ruiz de Mena, Kevin Carr, Matias Vicente, and Lyle Simmons for correcting my English, sharing experimental materials and techniques, and interesting discussions. I am also very grateful to many former students of the Kaguni laboratories, including Dave Lewis, Mark Sutton, Carla Margulies, Jianjun Wang, Andrea Williams, Anthony Lagina, and Phillip Kiefer.

I thank many fellow graduate students in the department for everything, especially Lei Lei for sharing success and frustration in research, and other interests.

Finally, I thank my wife, Yueying Cao and my son, Ruotao Wang for their love, support, encouragement, and understanding.

# TABLE OF CONTENTS

# LIST OF TABLES

Page

## Chapter IV

# LIST OF FIGURES

## Chapter I

## Chapter II



## Chapter III

## Chapter IV

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AP | apurinic/apyrimidinic |
| ATP | adenosine triphosphate |
| ATPase | adenosine triphosphatase |
| BrU | 5-bromouracil |
| CDK | cyclin-dependent kinase |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytosine triphosphate |
| dGTP | deoxyguanosine triphosphate |
| *D. melanogaster* | *Drosophila melanogaster* |
| DNA | deoxyribonucleic acid |
| DNase I | deoxyribonuclease I |
| dNTP | deoxynucleoside triphosphate |
| dRP | deoxyribose phosphate |
| dTTP | deoxythymidine triphosphate |
| *E. coli* | *Escherichia coli* |
| gp | gene product |
| kDa | kilodalton |
| mRNA | messenger RNA |
| mtDNA | mitochondrial DNA |
| mtSSB | mitochondrial SSB |
| NCBI | National Center for Biotechnology Information |
| Ni-NTA | nickle nitrilo-tri-acetic acid |
| nt | nucleotide |
| PCNA | proliferating nuclear cell antigen |

| | |
|---|---|
| pfu | plaque forming unit(s) |
| PMSF | phenylmethylsulfonyl fluoride |
| RF-A | replication factor C |
| RNA | ribonucleic acid |
| RP-A | replication protein A |
| *S. cerevisiae* | *Saccharomyces cerevisiae* |
| *S. pombe* | *Schizosaccharomyces pombe* |
| SSB | single-stranded DNA-binding protein |
| ssDNA | single-stranded DNA |
| SV40 | simian virus 40 |
| tRNA | transfer RNA |
| UV | ultra-violet |
| *X. laevis* | *Xenopus laevis* |

# CHAPTER I

# INTRODUCTION

# MITOCHONDRIA

Mitochondria are eukaryotic cellular organelles that contain a double membrane, their own DNA genomes, and an independent protein synthesis system. Their main biological function is energy production through oxidative phosphorylation and ATP production (Wallace, 1992). They also contribute to the biosynthesis of pyrimidines, amino acids, phospholipids, nucleotides, folate coenzymes, heme, urea and many other metabolites (Whittakr, 1978). Recently, mitochondria have been demonstrated to play an important role in the regulation of apoptosis (Mignotte and Vayssiere, 1998).

The observation that modern mitochondria possess many characteristics of eubacteria has led to the hypothesis that they have evolved from eubacteria-like endosymbionts in a primitive eukaryotic cell as the host (Gray, 1989a). It is still not clear whether a eukaryotic or another prokaryotic cell served as the host for the ancestor of mitochondria. An alternative hypothesis is that an anaerobic and autotrophic archaebacterium served as the host for a eubacterium (the symbiont) and it was the common ancestor of all present day eukaryotic cells (Martin and Muller, 1998). Present day unicellular eukaryotes that do not contain mitochondria arose later by losing their mitochondria during evolution. Since taking residence in the host cell, mitochondria have diversified greatly and lost their genetic independence, and now requiring nuclear encoded proteins for oxidative phosphorylation, mitochondrial DNA replication, transcription, translation, and other biological functions (Schatz, 1996). However, mitochondria still contain their own genome, the mitochondrial DNA (mtDNA), which encodes a limited number of proteins.

Several genetic and degenerative human diseases have been linked to specific mutations in mtDNA (Wallace, 1995a), or mutations in nuclear genes encoding

mitochondrial proteins (Casari, et al., 1998). Accumulation of mtDNA mutations in somatic cells has been hypothesized to be one of the causes of cellular senescence and aging because mitochondria have been shown to possess relatively limited DNA repair functions and mtDNA is physically in closer contact to DNA damaging reagents (Shadel and Clayton, 1997).

**Mitochondrial DNA**

More than 70 complete mitochondrial genomes that represent many diverse species including mammals, birds, and insects have been sequenced and are available from the public domain database (GenBank, NCBI). Mitochondrial genomes have evolved extraordinarily and they show remarkable differences in their sizes and structures (Gray, 1989b; Gray, et al., 1998). This likely derives from a variety of different evolutionary constraints. However, mtDNA from animals as diverse as *Drosophila* and mammals share many common structural features, such as (1) a circular double-stranded DNA of 16 to 19 kb, (2) compact genome organization, (3) polycistronic transcriptional units, and (4) an identical gene content (Leblanc, et al., 1997). Available data from plant mtDNA sequences show that they contain more respiratory protein genes, a set of ribosomal protein genes, are larger in size, and appear to be more dynamic in structure (Fauron, et al., 1995). Data accumulated during the past 15 years have revealed that linear mitochondrial DNAs exist in a large number of lower eukaryotes, mostly in ciliata and fungi (Nosek, et al., 1998). Because of the diversity of mtDNAs in terms of their size and structure and consequently different replicative mechanisms, this presentation will focus on studies of animal mtDNA replication, including insects (*Drosophila*) and mammals (man and mouse).

Animal mitochondrial genomes are much smaller (approximately 16 to 19 kb) as compared to nuclear genomes that are usually close to or exceed a billion base pairs in size.

However, this second non-nuclear genome is equally important for the viability of animals. Animal mtDNA usually contains one noncoding region which shows the most diversity in sequence and length, and contains the origin of replication and the transcriptional promoters (Larsson and Clayton, 1995). The protein coding capacity of animal mtDNA is well conserved from insects to mammals. The oxidative phosphorylation system requires more than 90 protein components, but only 13 polypeptides are encoded by the mitochondrial genome. These 13 mitochondrial genes encode the same set of protein subunits of the oxidative phosphorylation complex in animals (Leblanc, et al., 1997). In addition, animal mtDNAs all contain genes for two ribosomal RNAs and 22 tRNAs which are required for the translation of mtDNA-encoded mRNA (Wallace, 1995b). The tRNA genes are usually located between protein-coding genes and distributed almost randomly over the entire mtDNA, with some degree of clustering (Shadel and Clayton, 1997).

**Mitochondrial DNA Replication**

Duplication of any DNA genome is the result of the collaboration of many enzymes and other proteins. The basic mechanisms of this complicated event are well conserved from prokaryotes to eukaryotes, despite the evolutionary distance. From *E. coli*, phage T4, to mammals (both nuclear and mitochondrial), DNA polymerases are assisted by other enzymes and replication proteins to complete DNA replication with high speed, fidelity and processivity in a regulated fashion (Kornberg and Baker, 1992).

In a typical animal cell that usually contains many mitochondria, mtDNA represents less than 1% of the total DNA, even though mtDNA molecules greatly outnumber the nuclear chromosomal DNA molecules. The total number of mtDNA molecules in a particular cell may vary according to the respiratory requirement, or the developmental stage, or the growth rate of the cell (Robin and Wong, 1988). The mtDNA copy number of mtDNA per mitochondrion is relatively constant. Thus, the variability in the total number

of mtDNA molecules in different tissues is primarily the result of mitochondrial density (Veltri, et al., 1990). This suggests that mtDNA replication is a regulated event during mitochondrial biogenesis. However, replication of mtDNA is not coupled to cell cycle control in contrast with strictly regulated chromosomal DNA replication (Clayton, 1992). Data regarding the cis-acting DNA elements and trans-acting protein factors required for the initiation of animal mtDNA replication are mostly from vertebrates (Lee and Clayton, 1997; Lee and Clayton, 1998), especially mammals, although more evidence in other species is accumulating. Mitochondrial DNA replication in mammals, and likely in other animals, is asymmetric (Shadel and Clayton, 1997). Both leading and lagging DNA strands are synthesized continuously, and the origins for the two strands are mapped to different genomic locations. In mammals, the origin of leading strand synthesis is located within the noncoding region while the lagging DNA strand initiation site is located approximately two thirds the genomic distance away from the leading strand origin. The initiation of lagging strand replication does not occur until the replication fork originated from the leading strand origin has passed the lagging strand initiation site so that the lagging strand origin is in the single-stranded state. In *Drosophila*, lagging strand initiation occurs when leading strand replication is almost complete (Goddard and Wolstenholme, 1980). The current model for mammalian mtDNA replication is based upon biochemical analysis of early replication intermediates and mitochondrial RNA processing (Lee and Clayton, 1998).

## Mitochondrial DNA Mutations and Their Consequences

Because each mitochondrion contains multiple copies of mtDNA and each cell contains many mitochondria, a single mutation does not cause immediate effect in cellular function. The deficiency of mitochondrial function will be observed first in those tissues requiring high energy production when the mutant form of mtDNA accumulates to a critical

portion of the total mtDNA (Wallace, 1995b). However, animal mtDNA is so compact that almost every mutational event will have a high probability of occurring in either the coding region or essential regulatory sequences of mitochondrial genes. Eukaryotic cells have evolved mechanisms to maintain the genetic stability of mitochondrial genomes.

Mutations involving large deletions and base substitutions have been identified in the mtDNA of humans and have been shown to associate with several maternally inherited and ageing-related neuro-muscular degenerative diseases ( MITOMAP data base; http://www.gen.emory.edu/mitomap.html). Large deletions in the non-coding region of mtDNA which contains the replication origin and promoters for transcription would be lethal in both germline and somatic cells. That is presumably the reason why these mutants have never been detected. However, more than 90 large deletions at different sites have been reported (Kogelnik, et al., 1998), and they fall into two classes. The first type of deletions always involve direct repeat sequences at the deletion junctions and deletions in the second class do not contain an obvious repeated sequence (Mita, et al., 1990). Mitochondrial DNA with large deletions apparently can not be transmitted through the female germline (Larsson, et al., 1992) because virtually all cases reported are sporadic. It is likely that a mitochondrion containing mutant mtDNA with large deletion is lethal during embryogenesis or female germline cell development. Deleted mtDNA was also observed to accumulate with age in many mammalian somatic cells (Cortopassi, et al., 1992; Cortopassi and Arnheim, 1990; Tanhauser and Laipis, 1995). Pathogenic point mutations have been reported at 28 sites in 9 protein-coding genes and 34 sites in 15 tRNA genes. They are associated with at least 16 clinical syndromes. Only two rRNA gene point mutations leading to disease have been observed (Kogelnik, et al., 1998). The presence of mtDNA pseudogenes in the nuclear genome makes the detection and interpretation of mitochondrial point mutation more difficult (Tsuzuki, et al., 1983). For example, whether point mutations in cytochrome c oxidase subunit genes (CO1 and CO2) can be linked to

late-onset Alzheimer disease requires further studies (Davis, et al., 1997; Wallace, et al., 1997; Hirano, et al., 1997; Hutchin, et al., 1997). The fact that same mtDNA mutation in a protein-coding gene can produce distinct clinical syndromes in different families or individuals suggests that other factors likely contribute to the pathogenesis as well. These factors may include (1) the tissue-specific distribution of mutant mtDNA, (2) the severity of the deficiency in individual oxidative phosphorylation complexes, and (3) protein-protein interactions between mitochondrial and nuclear encoded subunits within a specific complex.

Any structural change or lesion in DNA requires fixation to yield a mutation, and mutations can occur before or during DNA replication. Mitochondrial DNA polymerase (pol $\gamma$) is the key enzyme in the replication of mtDNA and is possibly also involved in mtDNA repair (Longley, et al., 1998a). It certainly plays an important role in maintaining a low mutation rate in mtDNA.

## Mitochondrial DNA Repair

Although DNA repair in animal mtDNA is not as well documented as that in the nuclear genome, accumulating evidence indicates clearly that DNA repair does occur in animal mitochondria (Driggers, et al., 1993; Ikeda and Ozaki, 1997; LeDoux, et al., 1992; LeDoux, et al., 1993). The apparently higher mutational rate (or evolution rate) in mtDNA is not necessarily the result of the lack of DNA repair functions. However, deficiency in DNA repair functions will certainly increase the overall mutation rate. Many different factors are likely to contribute to the overall higher mutation rate in mtDNA, such as multiple rounds of replication of mtDNA per cell cycle, a higher rate of fixation of mutation due to rapid genetic drift, close contact with DNA damaging agents such as oxygen free radicals, a relative higher frequency of direct repeat sequences in mtDNA, and the existence

of single-stranded DNA over a relatively long period of time during mtDNA replication.

Recently, a number of laboratories have reported the identification and purification of DNA repair and recombination activities/enzymes from mitochondria of several evolutionally divergent organisms (Chi and Kolodner, 1994; Takao, et al., 1998). The human recombinant catalytic subunit of pol γ has been shown to contain a DNA repair activity (Longley, et al., 1998a). Thus, earlier speculation that mammalian mitochondria may lack DNA repair function (Clayton, 1982) was probably the result of insufficient experimentation.

## DNA POLYMERASES

Many enzymes, including but not limited to DNA polymerase, primase, telomerase, DNA ligase, topoisomerase, helicase, nuclease, and other replication accessory proteins such as single-stranded DNA-binding protein, initiation factors, processivity factors, and termination factors are indispensable for the cell to replicate its genome completely (Kornberg and Baker, 1992). However, the major enzyme responsible for cellular DNA replication is DNA-dependent DNA polymerase (DNA pol). DNA polymerases also provide the pivotal function of DNA synthesis in the cellular processes of DNA repair and recombination. A large number of DNA polymerases have been studied and they exhibit a rich variety in their structures and functions (Kornberg and Baker, 1992; Wang, 1991). For example, the mammalian repair DNA polymerase β (pol β) is a single subunit enzyme of 39 kDa, and the replicative polymerase of *E. coli,* pol III, has 18 subunits, a molecular mass of close to 900 kDa, and multiple enzymatic activities (McHenry, 1988). However as a superfamily, they are unified by exhibiting an essential catalytic function:

template/primer-dependent nucleotidyl transferase (Kornberg and Baker, 1992). The fundamental function of DNA pols is to extend the 3' end of a primer strand by adding a single deoxynucleotide, and the difference in their structure is determined by requirements for other enzymatic activities such as primase and proofreading 3'-5' exonuclease, for high processivity in replication, for interacting with other replication proteins, and for regulatory control, *e.g.* cell cycle control. Prokaryotic organisms encode at least one DNA polymerase, and as many as seven DNA polymerases have been discovered in higher eukaryotes (Wang, 1991). Many viruses encode their own DNA polymerase to bypass the host replicative machinery. DNA pols are grouped into four families (family A, B, C, X) based upon their amino acid sequence similarity (Ito and Braithwaite, 1991).

Typically, replicative DNA pols work in complex with other accessory proteins, which provide additional functions that help meet a number of requirements and overcome a variety of constraints inherent in the semiconservative duplication of long supercoiled and condensed double-helical DNA genomes.

**Catalytic Mechanisms**

Despite the diversity of DNA polymerases in their protein sequences, all of them contain a small number of conserved crucial amino acid residues that form the active site. DNA polymerases can not synthesize DNA *de novo*; they all require a template and a primer, synthesize DNA from 5' to 3' using deoxynucleoside triphosphate substrates. Although detailed structural data are available only from a small number of DNA polymerases (Kiefer, et al., 1997; Beese, et al., 1993; Wang, et al., 1997a; Wang, et al., 1996; Kim, et al., 1995; Eom, et al., 1996; Sawaya, et al., 1994; Sawaya, et al., 1997), the conservation of chemistry of DNA synthesis is well established.

The reaction pathway of DNA polymerization has been characterized for several prokaryotic enzymes (Patel, et al., 1991; Kuchta, et al., 1987; Dahlberg and Benkovic,

1991; Capson, et al., 1992). Kinetic data were derived mainly from bacteria and phage enzymes because of the availability of large quantities of purified enzymes in these systems. However based upon large number of mutagenesis studies, the basic chemical mechanisms can be expected to be conserved in eukaryotic enzymes (Copeland and Wang, 1993; Jung, et al., 1990). The enzyme binds template-primer DNA first to form a binary complex that is able to productively bind a dNTP to form a catalytically competent ternary complex. A conformational change in the ternary complex directs both substrate DNA and nucleotide to the position for phosphodiester bond formation. A second conformational change following the chemical step releases the pyrophosphate and repositions the binary complex for the next incoming deoxynucleotide. Crystal structures of binary or ternary complexes containing DNA polymerase, template-primer DNA, and deoxynucleotide (Doublie, et al., 1998; Li, et al., 1998) have identified both the specific acidic residues that bind divalent metal and form the catalytic active site, and other residues participating in DNA and nucleotide binding.

**Fidelity of DNA Strand Synthesis**

Two separate processes ensure the high fidelity of the DNA synthesis reaction. High selectivity for the correct deoxynucleoside triphosphate by the DNA polymerase and proofreading activity by the 3'-5' exonuclease combine to limit the error rate to $10^{-6}$ (Kunkel and Mosbaugh, 1989; Kunkel, 1992). The high selectivity comes from (1) discrimination against mispairs in dNTP binding, (2) higher resistance in the conformational change step when an incorrect nucleotide is bound, and (3) reduced efficiency in the chemical step of phosphodiester bond formation with a mispaired 3'-end . Again most data are from prokaryotic enzymes and the relative importance of each step varies between the polymerases studied (Dahlberg and Benkovic, 1991; Capson, et al., 1992). The contribution of kinetic selectivity in eukaryotic DNA polymerases remains to be

studied.

Most, if not all, replicative DNA polymerases contain a proofreading 3'-5' exonuclease activity either in a separate subunit or in the catalytic subunit. If an incorrect nucleotide is added to the growing 3'-end of primer strand, then the template-primer DNA becomes a better substrate for the 3'-5' exonuclease instead of the 5'-3' DNA polymerase. In studies with *E. coli* pol I, the second conformational change following the incorporation of an incorrect nucleotide is much slower, thus providing an opportunity for the enzyme to move the 3'-end containing the mismatch to the exonuclease site (Kuchta, et al., 1987). If the wrong nucleotide survives each of the error-reducing steps and the enzyme enters the next cycle of polymerization, the binding of the next dNTP is further reduced when the template-primer DNA has a mismatch at its 3'-end. This gives the enzyme another opportunity to transfer the mispaired DNA substrate to the proofreading active site. Alternatively, the enzyme dissociates from the DNA containing mispaired primer end, and the DNA becomes a substrate for another exonuclease.

In addition to the DNA synthetic cycle *per se*, post-replication mismatch DNA repair and maturation of Okazaki fragments in lagging strand synthesis also contribute to the overall fidelity of DNA replication (Barnes, et al., 1996).

## Prokaryotic DNA Polymerases

### DNA polymerase I

DNA polymerase I (pol I) from *E. coli* is a single-chain monomer with multiple functional domains which exhibit 5'-3' DNA polymerase, and 3'-5' and 5'-3' exonuclease activities (Kornberg and Baker, 1992). Pol I is the most abundant DNA polymerase in bacterial cells, and was the first DNA polymerase discovered. It is the most extensively studied DNA polymerase and serves roles in DNA replication, recombination, and repair.

It is the only DNA polymerase that has 5'-3' exonuclease activity and is able to utilize nicked circular duplex DNA as a template, unwinding the parental DNA strand from its template as it polymerizes.

Several crystal structures of the Klenow fragment of *E. coli* pol I bound to substrate DNA (Beese, et al., 1993) and structures of pol Is from other bacteria are now available (Kim, et al., 1995), providing insights into the enzymatic mechanism of both DNA polymerase and exonuclease activities. *E. coli* pol I is the prototype of family A DNA polymerases.

## DNA polymerase II

DNA polymerase II (pol II) from *E. coli* was discovered many years ago, yet its precise role in pathways of DNA replication and repair remains undetermined. It has been established that *dinA* (*polB*), the structural gene encoding pol II, is repressed by LexA as a part of the SOS regulon, and that the cellular levels of pol II are increased by a factor of seven following induction of the SOS regulon (Bonner, et al., 1988). In spite of data suggesting that pol II could be involved in repair of UV-induced damage (Masker, et al., 1973), null mutants of *polB* are viable and do not appear to exhibit an increased sensitivity to UV light (Escarceller, et al., 1994). Pol II can synthesize DNA with a high processivity in the presence of the *E. coli* pol III accessory proteins, the β sliding clamp and the clamp loading complex *in vitro*. Recent genetic data suggest that pol II could be involved in DNA synthesis on the chromosome and on the F episome in actively dividing cells. Rangarajan *et al.* replaced the wild-type *polB* gene on the chromosome with a proofreading-defective *polB* mutant (*polBexl*) and found significant increases in mutation frequencies of reporter genes on both the chromosome and on F(*lac*) episomes (Rangarajan, et al., 1997). DNA pol II is a family B polymerase sharing sequence

similarity with eukaryotic pol α catalytic subunit and bacteriophage T4 DNA polymerase. The entire genome from the archaeon *M. jannaschii* has recently been sequenced (Bult, et al., 1996) and found to contain a single DNA polymerase gene which exhibits high degree of amino acid identity with *E. coli* pol II. Thus it appears that this *E. coli* pol II homolog must catalyze both replicative and repair DNA synthesis in *M. jannaschii*.

## DNA polymerase III

Pol III holoenzyme from *E. coli* is responsible for the replication of the bacterial chromosome. Pol III holoenzyme is distinguished from other DNA polymerases in the cell by its high processivity (>50 kbp) and rapid rate of synthesis (750 nucleotides/s) (Kornberg and Baker, 1992; Kelman and O'Donnell, 1995). Pol III holoenzyme is a large complex composed of 10 functional subunits encoded by 9 different genes, some of which are present in multiple copies for a total of 18 polypeptide chains (Onrust, et al., 1995; McHenry, 1988). The holoenzyme can be resolved into three primary functional units. (1) The DNA polymerase core contains three tightly bound subunits as follows: the polymerase or α subunit, the 3'-5' exonuclease or ε subunit, and the θ Subunit, with no known function. (2) The ring-shaped sliding clamp is a processivity factor consisting of a dimer of the β subunit. (3) The five protein γ complex is the "clamp loader" that couples ATP hydrolysis to assembly of β clamps around the DNA substrate (Onrust, et al., 1995). The γ complex consists of five different subunits ($\gamma_2\delta_1\delta'_1\chi_1\psi_1$). The δ subunit is the major touch point to the β clamp and leads to ring opening, but is buried within the γ complex such that contact with β is prevented (Naktinis, et al., 1995). The γ subunit binds ATP but

is not an ATPase by itself (Tsuchihashi and Kornberg, 1989). The δ' subunit bridges the δ and γ subunits, resulting in a γδδ' complex that exhibits DNA-dependent ATPase activity and is competent to assemble clamps on DNA (Onrust, et al., 1991). Upon binding of ATP, a change in the conformation of the γ complex exposes δ for interaction with β subunit (Naktinis, et al., 1995). The χ and ψ subunits form a 1:1 χψ complex and the χ subunit makes direct contact to single-stranded binding protein to facilitate replication of an SSB-coated template (Kelman, et al., 1998; Glover and McHenry, 1998). A dimer of the τ subunit acts as a "macromolecular organizer," holding together two molecules of core and one molecule of γ complex to form the pol III* subassembly (Onrust, et al., 1995).

## T7 DNA polymerase

T7 DNA polymerase (T7 pol) is a two subunit enzyme. The catalytic subunit is the 80 kDa product of gene 5 of the phage, and it interacts physically with one host protein, *E. coli* thioredoxin to form a complex. The consequence of this interaction is to convert gene 5 protein from a polymerase with low processivity to one of high processivity (Tabor, et al., 1987). A 76-amino acid sequence (residues 258-334), unique to T7 pol, has been implicated in this interaction . In fact, insertion of this thioredoxin binding sequence into the homologous region of *E. coli* pol I has converted it into a polymerase of high processivity (Bedford, 1997). T7 pol does not require other accessory proteins to bind to the template-primer DNA substrate. T7 pol also interacts with the hexameric gene 4 protein of the phage, a protein that provides both helicase and primase activity at the replication fork to coordinate both leading and lagging strand synthesis (Park, et al., 1998). Both T7

pol and gene 4 protein in turn interact with the T7 gene 2.5 protein, a single-stranded DNA-binding protein that stimulates both polymerase and primase activities (Kim and Richardson, 1993). T7 pol also contains a N-terminal 3'-5' exonuclease domain, and is a member of family A DNA polymerase (Ito and Braithwaite, 1991). The three-dimensional structure of T7 DNA polymerase has been solved recently (Doublie, et al., 1998), and it provides additional support for a universal DNA polymerization mechanism.

## T4 DNA polymerase

Bacteriophage T4 gene 43 encodes the viral T4 DNA polymerase (T4 pol) with a molecular weight of 103 kDa. T4 pol has regions of high similarity with animal virus DNA polymerases and human DNA polymerase alpha catalytic subunit, and shares very limited similarity with *E. coli* pol I and no detectable similarity with T7 pol (Spicer, et al., 1988). It is therefore a member of the family B DNA polymerase group (Ito and Braithwaite, 1991). It resembles T7 pol in that it also contains both DNA polymerase and 3'- 5' exonuclease activity in its C- and N-terminal regions, respectively. The enzyme is also a sequence-specific RNA-binding autogenous translational repressor (Young, et al., 1992). Unlike T7 pol, T4 pol requires other phage-encoded replication accessory proteins to recognize effectively the template-primer DNA substrate. The product of gene 45 acts as sliding clamp and the products of genes 44 and 62 form a complex with DNA-dependent ATPase activity that functions as a clamp loader protein (Kroutil, et al., 1998). This is very similar to *E. coli* pol III, although in *E. coli* pol III the functionally analogous proteins are subunits of the pol III holoenzyme.

### Eukaryotic DNA Polymerases

Many years of biochemical and genetic studies have led to the identification of at

least 6 distinct DNA polymerases in eukaryotic cells, DNA pols α, β, γ, δ, ε, and ζ (Wang, 1991). The *S. cerevisiae* genome contains an open reading frame that could potentially encode another DNA polymerase (Sugino, 1995). Recently a human gene that potentially encodes a seventh DNA polymerase has been identified (GenBank accession number AF052573, Sharief,F.S., Ropp,P.A., and Copeland,W.C.) and it is shown to be a homolog of the *Drosophila* Mus308 protein. The *Drosophila mus308* gene has been implicated in DNA cross-link repair and has sequence similarity to both DNA polymerase and DNA helicase (Harris, et al., 1996). Whether these two genes actually encode a new DNA polymerase remains to be confirmed by biochemical analysis of their gene products. The only DNA polymerase located in mitochondria of animal cells is DNA pol γ, although a β-like DNA polymerase has been reported in the protozoan *Crithidia* (Torri and Englund, 1995). In addition, plant cells contain a γ-like DNA polymerase for the replication of chloroplast DNA (Sala, et al., 1980).

## DNA polymerase α

The highly conserved DNA polymerase α-primase complex (pol α) is the only eukaryotic DNA polymerase that can initiate DNA synthesis *de novo* (Lehman and Kaguni, 1989). It is required both for the initiation of DNA replication at chromosomal origins and for the discontinuous synthesis of Okazaki fragments on the lagging strand at the replication fork (Wang, et al., 1989). Pol αs isolated from yeast to mammals contain 4 subunits of 165-180, 70-86, 60, 50 kDa (Kaguni, et al., 1983b; Plevani, et al., 1988; Miyazawa, et al., 1993). Pol α provides RNA-DNA primers for the initiation of both

leading and lagging strand synthesis based upon the evidence from reconstituted SV40 DNA replication. The large catalytic subunit contains the DNA polymerase activity (Copeland and Wang, 1991), and the two small subunits form the primase. The p48 subunit contains the primase activity (Schneider, et al., 1998) while the p58 subunit is necessary for the stability and enzymatic activity of p48 subunit. The second-largest subunit of the mouse DNA polymerase α has been shown to facilitate both production and nuclear translocation of the catalytic subunit of pol α (Mizuno, et al., 1998). Genes for the catalytic subunit as well as those for other subunits have been cloned and mapped from many species including *Drosophila*, human, mouse, and yeast.

Post-translational modification of pol α has been observed in human and *S. cerevisiae* cells. Both the human p180 and p70 subunits are phosphorylated in a cell cycle-dependent manner and the replication activity of pol α is modified by phosphorylation (Nasheuer, et al., 1991; Voitenleitner, et al., 1999). These data suggest that the mitotic phosphorylation of pol α may affect its physical interaction with other replicative proteins and/or with DNA at the replication fork. Because pol α is the only DNA polymerase that can synthesize DNA *de novo*, the phosphorylation of pol α may play critical regulatory role in the initiation of DNA replication.

## DNA polymerase β

Pol β from man, mouse, rat, and cow, is a single-subunit enzyme of 39 kDa and belongs to the family X DNA polymerase group (Ito and Braithwaite, 1991). The yeast pol II gene encodes a protein with predicted molecular mass of 68 kDa which is a homolog of

mammalian pol β. Mammalian pol β consists of 2 domains linked by a protease-sensitive hinge region. This enzyme is an integral part of the base excision repair pathway (Sobol, et al., 1996). The C-terminal 31 kDa domain possesses the polymerase catalytic activity. The N-terminal 8-kDa domain is required for single-stranded template DNA binding and has been recently shown to also contain an activity to excise deoxyribose phosphate (dRP) groups from 5'-incised apurinic/apyrimidinic (AP) sites during base excision repair (Prasad, et al., 1998; Matsumoto, et al., 1998).

Pol β is so far the only eukaryotic DNA polymerase for which a crystal structure is available (Sawaya, et al., 1997). The polymerase active site of pol β is structurally similar to that of prokaryotic DNA polymerases, suggesting a universal DNA polymerization mechanism(Sawaya, et al., 1994).

Expression of the pol β gene is constitutive under normal growth conditions because pol β enzymatic activity and mRNA levels are independent of cell growth and cell cycle regulation. There have been reports that pol β gene expression can be induced by DNA damaging agents, consistent with its role in DNA repair.

## DNA polymerase δ

Pol δ has been shown to play important roles in DNA replication, nucleotide excision repair, base excision repair and V-D-J recombination of immunoglobin genes by both genetic and biochemical approaches. The catalytic subunit of pol δ is well conserved from yeast to human, although the precise subunit structure of pol δ is still an unresolved

issue. Pol δ is usually purified as a heterodimeric enzyme of a 125 kDa catalytic subunit and a 50 kDa small subunit from mammalian cells, but budding yeast pol δ has been shown to contain a third subunit (Burgers and Gerik, 1998). Pol δ purified from the fission yeast *S. pombe* contains five distinct subunits, three of which are encoded by essential genes (Zuo, et al., 1997). The p125 subunit contains both DNA polymerase and 3'-5' exonuclease activity. An active human recombinant pol δ heterodimer has been reconstituted in the baculovirus expression system. The small subunit of human pol δ is required for the interaction and stimulation of DNA polymerase activity by proliferating cell nuclear antigen (PCNA) (Zhou, et al., 1997). Genes encoding both subunits have been cloned from many species.

## DNA polymerase ε

Like pol δ, the exact subunit composition of pol ε is not yet resolved. Purified human pol ε contains at least two subunits of 255 kDa and 55 kDa. In *S. cerevisiae,* pol ε is reported to consist of a 256 kDa catalytic subunit and four additional subunits of 80, 34, 31, and 29 kDa. The 34 and 31 kDa subunits are encoded by the same gene and may result from post-translational processing. The exact role of pol ε in DNA replication is not clear (Bambara, et al., 1997). In yeast, the gene encoding the catalytic subunit is essential for cell growth (Sugino, et al., 1998), but pol ε is not required for *in vitro* replication of SV40 virus DNA (Zlotkin, et al., 1996).

## DNA polymerase ζ

The genes encoding pol ζ, *rev3* and *rev7*, have been cloned based on their phenotypes in a UV sensitivity assay in budding yeast *S. cerevisiae* (Morrison, et al., 1989) . Pol ζ has been shown to be a nonessential DNA pol in yeast, and is involved in bypass DNA synthesis on DNA templates containing lesions (Nelson, et al., 1996). The *rev3* gene product is the catalytic subunit of pol ζ and *rev7* protein constitutes an accessory subunit which stimulate the overall DNA polymerase activity. Mammalian homologs of the catalytic subunit have been identified (Gibbs, et al., 1998), and it is likely that the pathway and mechanisms of DNA-damage induced mutagenic DNA synthesis are conserved.

## DNA polymerase γ

Pol γ is a low abundance enzyme in somatic cells, usually representing less than 1% of total cellular DNA polymerase activity. Nevertheless , it is an essential enzyme for the cell viability. Pol γ purified from many species shares many biochemical features including, the size of the catalytic subunits, intrinsic proofreading 3'-5' exonuclease activity, the capacity to use homoribopolymers as substrate *in vitro*, sensitivity to N-ethylmaleimide and dideoxynucleotides, and resistance to aphidicolin (Wernette and Kaguni, 1986). Molecular cloning of catalytic subunit genes has shown that the catalytic subunit of pol γ from yeast to man is well conserved in amino acid sequence, and that it belongs to the family A DNA polymerase group (Lecrenier, et al., 1997). However, the sequence similarity between pol γ and other family A DNA polymerases is restricted to the catalytic domains of both polymerase and exonuclease, and γ pols share unique conserved sequences among them, thus the γ pols represent a distinct subfamily.

*Drosophila* pol γ in its native form purified from embryos is a heterodimer of 125 kDa (α) and 35 kDa (β) subunits (Wernette and Kaguni, 1986). It has been shown to synthesize DNA efficiently on a variety of DNA substrates, and is stimulated by high salt concentration and by mitochondrial single-stranded DNA-binding protein (mtSSB). The large 125 kDa catalytic subunit contains both 5'-3' DNA polymerase and 3'-5' exonuclease activities (Lewis, et al., 1996) while no specific biochemical function has been assigned to the 35 kDa subunit. It has been shown that the β subunit is important to maintain the catalytic efficiency and/or the structural integrity of the holoenzyme (Olson, et al., 1995). In one report pol γ purified from human Hela cells contained polypeptides of 140 and 54 kDa (Gray and Wong, 1992). In another, the catalytic subunit (140 kDa) was shown to copurify with a complex of four polypeptides (77, 52, 49, and 45 kDa) (Longley, et al., 1998b). The purified porcine liver pol γ contains four polypeptides of 120, 55, 50, and 48 kDa (Kunkel and Mosbaugh, 1989), and a preparation from *Xenopus laevis* embryos possesses polypeptides of 100, 85, 55, 40, and 31 kDa in association with the 140 kDa catalytic subunit (Insdorf and Bogenhagen, 1989a).

## DNA Polymerase Accessory Proteins

DNA polymerase processivity factors have quite unique names in different systems because of their initial characterization. In *E. coli,* the β subunit of pol III holoenzyme serves as the processivity factor. For bacteriophage T4, gp45 protein provides the same biological function. In eukaryotes, PCNA is conserved from yeast to man and it increases the processivity of pol δ even though it was originally discovered as a cell cycle-regulated

nuclear protein. The crystal structures of PCNA and the β subunit have been solved (Kong, et al., 1992; Krishna, et al., 1994). These three proteins do not share significant sequence homology. However, their closed circular structures are very similar. While the β subunit of *E. coli* pol III is a dimer, PCNA and T4 gp45 protein form homotrimers. All three proteins contain a hole in the center that encircles the duplex DNA thus allowing them to track along the DNA molecule. These sliding clamps do not contain DNA-binding activity but depend upon clamp-loading proteins (replication factor C in eukaryotic systems, γ complex of pol III in E. coli, and gp42/66 complex in phage T4) to be loaded on DNA. Once they are loaded on DNA, these sliding clamp proteins are able to physically interact with replicative DNA polymerases thus conferring high processivity to the holoenzymes at the replication fork. PCNA has also been shown to bind many other proteins, including FEN-1 nuclease , DNA ligase I, cyclin-dependent kinase(CDK) inhibitor protein p21, nucleotide excision repair protein XPG, DNA-(cytosine 5) methyltransferase, mismatch repair proteins MLH1 and MSH2, and cyclin D (Kelman and Hurwitz, 1998). These interactions with DNA modifying proteins and cell cycle control proteins suggest that PCNA also plays important roles in other cellular functions such as DNA repair and cell-cycle progression in addition to its roles in DNA replication.

The single-stranded DNA binding protein of *E. coli* is a homotetrameric protein. In human and other eukaryotes it is called replication protein A (RP-A) or replication factor A. RP-A is a heterotrimer of 70 kD, 32-34 kD, and 11-14 kD subunits (Wold and Kelly, 1988; Marton, et al., 1994), and none of three subunits share any sequence homology to *E. coli* SSB. Interestingly, mitochondrial SSBs are similar to *E. coli* SSB in sequence and structure (Thommes, et al., 1995). All SSBs show preferred binding affinity to single-

stranded DNA over duplex DNA.

DNA helicases are enzymes that use the energy of NTP hydrolysis to processively denature duplex DNA to provide single-stranded DNA template for DNA replication. Helicases also play essential roles during DNA repair, recombination and transcription (Lohman, 1993). The DNA unwinding requirement in these processes vary considerably and so do diverse interactions between helicase and other proteins. It is not surprising that many DNA helicases have been purified and characterized and numerous orphan helicases have been isolated either by sequencing projects or by genetic cloning approaches. Only those helicases that have been shown to interact or associate (co-purification) with DNA replication proteins or DNA polymerases are now considered to be required for DNA replication.

## OVERVIEW

The overall sequence similarity between DNA pols from different families is very weak and they show great diversity in their subunit structures and biochemical activities. However, experimental evidence is accumulating from genetic, biochemical and structural studies that supports the hypothesis that family A, B, and C DNA polymerases are all related and share a common ancestor (Zhu and Ito, 1994). One fundamental feature of all DNA polymerases that is absolutely conserved from bacteria to man is the chemical mechanism of the template-primer DNA dependent nucleotidyl transfer reaction. Their substrates in various DNA metabolic pathways vary from a single gap in base excision DNA repair to an entire chromosome of over a 100 million nucleotides in mammalian cells. DNA synthesis in various pathways is also regulated in coordination with other aspects of the life of a particular cell. Thus, the requirements of each specific pathway determine the

structural and mechanistic diversity in DNA pols.

It is not surprising that DNA polymerases perform basically the same chemical reaction whether in DNA replication or DNA repair but show great diversity in their structures. However, I find it almost amazing that replicative DNA polymerase holoenzymes in different systems have evolved to perform other biochemical functions required for DNA synthesis with non-homologous proteins and even by completely different mechanisms. For example, the complex for high processivity in *E. coli* DNA replication is a dimer of β subunits in pol III(Kong, et al., 1992), while PCNA used by DNA pols in eukaryotic cells is a homotrimer (Krishna, et al., 1994). The β subunit and PCNA show no sequence similarity yet both proteins fold to nearly superimposable 3-dimensional structures and increase the processivity of DNA pols by a 'sliding clamp' mechanism. It is equally interesting that bacteriophage T7 DNA polymerase binds a host cellular protein, thioredoxin, to increase its processivity (Tabor, et al., 1987). The structural data by Doublie et al. (Doublie, et al., 1998) have demonstrated that thioredoxin binds to the tip of the thumb domain of T7 pol and forms a flexible structure which could close the DNA-binding groove of T7 pol. The herpes virus DNA pol might employ the same mechanism although structural data are not available (Hernandez and Lehman, 1990). Another example is the priming of DNA synthesis. The mechanism of elongating established primer chains by all DNA polymerases appears to be absolutely conserved based upon many lines of evidence. However, the processes of starting a new DNA strand are quite diverse. Initiation of chromosomal DNA replication in both bacteria and eukaryotes involves the loading of a primase to the unwound DNA. Again, different strategies are employed to locate the primase to the origin. The primase can be a separate enzyme as in *E. coli*, or contained in the same protein complex functioning as helicase as in herpes simplex virus (HSV) (Crute and Lehman, 1991), or in complex with a DNA

polymerase (pol α) as in eukaryotes (Lehman and Kaguni, 1989). In eukaryotic cells, both nuclear and mitochondrial DNA synthesis are generally primed by short RNA primers, but the enzymatic mechanisms of primer synthesis are different. Pol α contains a primase subunit to provide short RNA primers (Kaguni, et al., 1983a; Schneider, et al., 1998). In contrast, leading strand initiation in mtDNA replication uses RNA synthesized by mitochondrial RNA polymerase (Lee and Clayton, 1998). Some viruses have evolved mechanisms that do not even require RNA synthesis; they use either an origin-binding protein (King, et al., 1997; Mendez, et al., 1997) or self-complementary hairpin termini (adeno-associated viruses) to supply the 3' hydroxyl ends (Ward and Berns, 1995). The unique mechanism used by HIV reverse transcriptase is recruitment of a tRNA as the primer.

From the results of previous work in our laboratory it is clear that the β subunit of *Drosophila* pol γ is a distinct protein, not a proteolysis product of the catalytic subunit derived during the purification (Olson, et al., 1995), and that the β subunit binds to the α subunit with high affinity. We have not determined the functional role of the β subunit in biochemical studies. Many protein factors required for mtDNA replication have not been identified, and accessory subunits of different DNA pols can exhibit several biochemical activities and/or structural roles. In fact, the β subunit could perform one or several functions required for efficient and high fidelity mtDNA synthesis. My thesis research aims to determine (1) whether or not the β subunit is a *bone fide* subunit of the enzyme, (2) the sequence identity of this subunit, (3) whether pol γs from other sources also contain this subunit, and finally (4) the biochemical function of the β subunit in mtDNA

replication. The focus of my research is to clone the gene encoding the β subunit, and establish an expression system to produce, purify, and characterize recombinant pol γ.

The low abundance of native pol γ limits the physical amount of enzyme that can be purified, thus prevent us from performing interesting structure-function studies which require a large quantity of protein. The successful expression and purification of recombinant enzyme will not only provide us materials, but also an important methodology to study mutant forms of the enzyme and its individual subunits.

# CHAPTER II

# INTERACTION BETWEEN *DROSOPHILA* POL γ AND TEMPLATE—PRIMER DNA

# INTRODUCTION

Many enzymes are required to replicate mtDNA completely. However, DNA polymerase γ (pol γ) is the key enzyme in mtDNA synthesis. It has been documented that *Drosophila* pol γ can utilize efficiently a variety of DNA substrates including gaped DNA, poly(rA).oligo(dT), and singly-primed ϕX174 DNA (Wernette and Kaguni, 1986). Whether the template-primer DNA is presented by an RNA-processing enzyme acting on an RNA transcript created by mtRNA polymerase, or synthesized by a primase activity, specific interactions between pol γ and the template-primer DNA is a required step in the transition from initiation to processive elongation in both leading and lagging strand DNA replication. The functional interaction of pol γ and template-primer DNA is also a good target for the regulation of mtDNA replication, considering the proposed continuous asymmetric model of animal mitochondrial DNA replication. According to this model, for both leading and lagging DNA strand synthesis, pol γ needs to recognize and interact with the primer terminus only once if processive DNA polymerization is maintained. The large α subunit of pol γ contains both DNA polymerase and exonuclease activities (Lewis, et al., 1996), and the accessory β subunit is required for efficient catalysis (Olson, et al 1995). It is not known whether the β subunit is involved in DNA-protein interactions when pol γ binds to template-primer DNA. It is also not known whether the β subunit plays different

roles in the initiation and elongation of continuous mtDNA synthesis.

A number of experimental approaches have been developed to study the specific interactions between DNA-binding proteins and substrate DNAs. Gel mobility shift electrophoresis is a very effective method to probe the DNA sequence specificity and stability of DNA-protein complexes. Many sequence-specific DNA binding proteins have been shown to form covalent DNA-protein complexes when exposed to UV light (Safer, et al., 1988; Capson, et al., 1991). Substituting thymidine monophosphate with 5-bromodeoxyuridine monophosphate in the binding site sequence of a DNA substrate increases the photoreactivity of the DNA and better cross-linking efficiency is obtained (Blatter, et al., 1992). The structural change introduced by this substitution is minimal since 5-bromouracil (BrU) and thymine have very similar van der Waals radii (Blatter, et al., 1992). Most sequence-specific DNA binding proteins also do not distinguish between BrU and thymine so their DNA binding specificity and affinity are not greatly affected. BrU itself is not active in the dark. However, a highly reactive uracilyl radical is generated upon its exposure to UV light. If there is a hydrogen donor in direct contact with the radical, such as when the DNA is bound to a protein, covalent bond formation may occur between them. When long wavelength UV is applied, no DNA strand break or DNA-DNA cross-linking or protein-protein cross-linking is observed (Hockensmith, et al., 1986; Lin and Riggs, 1974). This technique has been applied to identify the catalytic subunit of *X. laevis* pol $\gamma$ (Insdorf and Bogenhagen, 1989a). Combined with proteolytic mapping, the specific DNA binding domain and even the specific amino acid residues involved can be identified. Enzymes with multiple activities such as DNA polymerases are usually structured as a series of independent functional domains connected by short polypeptide linkers which are more sensitive to proteolytic attack. Partial proteolysis of such proteins can yield active fragments that can define the functions of multiple domains.

The DNase I footprinting technique has been applied to study the specific contact between a DNA polymerase and its DNA substrate as well as the interactions between the DNA polymerase and other accessory factors at the template-primer junction. *E. coli* DNA polymerase III holoenzyme has been shown to footprint three helical turns on the primer strand. (Reems and McHenry, 1994). T4 DNA polymerase and four other replication proteins interact at the template-primer junction and form different complexes when various complex members are present (Munn and Alberts, 1991). One complication of footprinting with DNA polymerase containing an intrinsic 3'-5' exonuclease activity is that the exonuclease activity will act on the 3'-hydroxyl ends generated by the DNase I digestion, thus destroying the initial footprint, because the DNA polymerase is presented in molar excess in order to form stable DNA-protein complexes. The DNA substrate used must be labeled at the 3' end of the primer strand, so the DNase I fragments with newly-created 3' ends will not be labeled and not observed on the gel. Incorporation of a dideoxynucleotide at the 3' end of the primer strand renders the substrate resistant to hydrolysis by the 3'-5' exonuclease activity of pol γ because dideoxynucleotides are potent inhibitors of pol γ.

To elucidate structure-function relationships in *Drosophila* pol γ, we employed the gel mobility shift assay combined with UV-induced cross-linking, and DNase I footprinting to study the molecular interactions between the pol γ holoenzyme and template-primer DNA. In addition, we examined the effect of mitochondrial single-stranded DNA-binding protein (mtSSB) on these interactions.

# EXPERIMENTAL PROCEDURES

## Materials

*Nucleotides and Nucleic Acids-* The DNA substrate used for gel electrophoretic mobility shift and photocross-linking experiments is a synthetic 36 mer oligonucleotide (LSK-7, Figure 1C ) which will form a hairpin structure with a 5'-overhang under native conditions, thus creating a primed template for pol γ action. DNase I footprint substrate DNA is constructed by annealing a 50-mer oligonucleotide complementary to M13mp19 single stranded DNA at position 991 to 1041. Both oligonucleotides were synthesized in an Applied Biosystems mode 477 oligonucleotide synthesizer. M13mp19 single stranded DNA was prepared by standard laboratory methods. Unlabeled deoxyribonucleoside triphosphates were purchased from P-L Biochemicals. [$\alpha$-$^{32}$P]dATP was purchased from ICN.

*Enzymes and Proteins* - *Drosophila* pol γ Fraction VI (glycerol gradient fraction) was prepared as described by Wernette and Kaguni (Wernette and Kaguni, 1986). *Drosophila* mtSSB was purified from embryonic mitochondria by a variation of the published method [Thommes, 1995 #923 Farr and Kaguni, unpublished data]. Trypsin and DNase I were purchased from Sigma.

*Chemicals-* Protease inhibitor PMSF was purchased from Sigma, sodium metabisulfite was from J. T. Baker Chemical Co., and leupeptin was from Peptide Institute, Minoh-Shi, Japan.

## Methods

*Binding and Photochemical Cross-linking of pol γ to Template-primer DNA* -- A radiolabeled deoxyoligomer (40 nt) containing a 3'-terminal bromodeoxyuridylate residue was prepared by incubating the 36-mer DNA (LSK7, 22 pmol) described under "Materials" with *E. coli* DNA polymerase I Klenow fragment (0.4 units) for 10 minutes at 37 $^0$C in the presence of BrdUTP (16 pmol) and [$\alpha$-$^{32}$P]dATP (26 pmol), in a reaction mixture (0.2 ml) containing 50 mM Tris·HCl pH 8.0, 10 mM $MgCl_2$, 4 mM dithiothreitol, and 200 μg/ml bovine serum albumin. The mixture was incubated further for 30 minutes at 37 $^0$C after addition of unlabeled dATP (5 nmol). The reaction was terminated by addition of EDTA to 10 mM, and heating for 15 minutes at 65 $^0$C. The resulting radiolabeled 40-mer was precipitated with ethanol, resuspended in 10 mM Tris·HCl (pH 8.0), 1 mM EDTA, and its length and purity were analyzed by 18% polyacrylamide gel electrophoresis in the presence of 7 M urea (data not shown).

Template-primer DNA binding by pol γ was examined in a gel electrophoretic mobility shift assay as follows. Pol γ Fraction VI (28 fmol) was incubated with the [$^{32}$P]dAMP-labeled BrdUMP-substituted 40-mer (0.22 pmol) in the absence or presence of competitor DNA for 10 min at 30 $^0$C in standard pol γ reaction buffer containing 50 mM Tris·HCl (pH 8.5), 4 mM $MgCl_2$, 5 mM dithiothreitol, and 30 mM KCl, followed by the addition of bromphenol blue and glycerol to 0.001 and 5%, respectively, and

electrophoresis in a 4.5% native polyacrylamide gel (13x13x.15 cm) in 45 mM Tris-borate (pH 8.3) and 1 mM EDTA. After electrophoresis, the gel was dried under vacuum and exposed at -80 $^0$C to Kodak X-Omat AR X-ray film using a DuPont NEN Quanta III intensifying screen.

Photochemical cross-linking with UV light was performed after incubation of pol γ Fraction VI with template-primer DNA as described above. The reaction mixtures were irradiated for 15 min at 0 $^0$C with UV light (300 nm) from a germicidal bulb (Fotodyne, 4 X 15 watts) at a distance of 8 cm. After irradiation, the samples were made 1X in Laemmli sample buffer (Laemmli, 1970), denatured, and electrophoresed in a 7.5% SDS-polyacrylamide gel (13x13x0.15 cm). The gel was dried under vacuum and autoradiographed as described above. In experiments where pol γ was subjected to digestion with trypsin prior to photochemical cross-linking, the DNA binding mixtures contained higher levels of pol γ (0.6 pmol), template-primer DNA (1.5 pmol ), and trypsin (200 or 400 ng) and were incubated for 15 min at 20 $^0$C in a digestion buffer (25 μl) containing 40 mM potassium phosphate pH 7.6, 160 mM ammonium sulfate, 0.012% Triton X-100, 1.6 mM EDTA, 5mM DTT, 0.8 mM PMSF, 8 mM sodium metabisulfite, 1.6 μg/ml Leupeptin, and 16 % glycerol. The digestion was terminated by addition of sodium metabisulfite to 20 mM and leupeptin to 20 μg/ml. Aliquots of samples were then added to the DNA binding reaction mixture, processed, irradiated and electrophoresed as described above. The trypsin cleavage products of pol γ were also resolved on 7.5 % SDS-polyacrylamide gel and identified by silver staining.

*DNase I Footprinting* -- The footprinting substrate DNA was constructed by

annealing a 50-mer oligonucleotide complementary to M13mp19 single stranded DNA at position 991 to 1041 in a buffer containing 40 mM Tris pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl to the circular single-stranded M13mp19 DNA. The primer strand was then further extended by Sequenase Version 2.0 with [$\alpha$-$^{32}$P]dATP (3000 Ci/mmol, 0.13 $\mu$M) and ddTTP (0.8 $\mu$M). Excess unannealed primer DNA and unincorporated nucleotides were removed by gel filtration on a Sephadex G-50 column equilibrated with 10 mM Tris-HCl pH 8.0, 1 mM EDTA. The DNA concentration of purified template-primer DNA was determined by measuring the absorbance at 260 nm. The specific activity of labeled DNA was determined by liquid scintillation. The final DNA product was also analyzed on 15 % polyacrylamide-urea gel to assess the size of primer strand DNA. The DNase I digestion reaction was performed in a reaction buffer of 50 mM Tris pH 8.5, 4 mM $MgCl_2$, 5 mM DTT, 30 mM KCl. The amount of DNase I needed was determined by titration to be 0.25 ng/$\mu$l. 17 fmol of DNA was used for all of the reactions, but various amounts of pol $\gamma$ and mtSSB were added in different sets of experiments. After mixing all the reagents except the DNase I, the reaction mixtures were incubated at 37°C for 5 minutes to reach equilibrium. An additional 2 minutes of incubation at 37°C was performed after the DNase I was added. The digestion was terminated by adding 150 $\mu$l of stop buffer (0.5 % SDS, 0.2 M sodium acetate, 30 mM EDTA, 25 $\mu$g/ml tRNA). The quenched reaction mixtures were then extracted with an equal volume of phenol/chloroform (1:1), and precipitated with ethanol at - 80 °C for 30 minutes. The DNA pellets were washed with 70 % ethanol and dried in a vacuum desiccator. The DNAs were redissolved in 6 $\mu$l of 95 % formamide containing 10 mM NaOH, 1 mM EDTA, and 0.3 mg/ml each of bromophenol blue and xylene cyanol as tracking dyes. Aliquots from each reaction were denatured for 2 min at 100 °C, chilled on ice and electrophoresed in a 10 % polyacrylamide DNA sequencing gel

containing 7 M urea in 90 mM Tris-borate pH 8.3 and 2 mM EDTA. Autoradiographs were obtained by exposing the dried gel to Kodak X-Omat AR x-ray film.

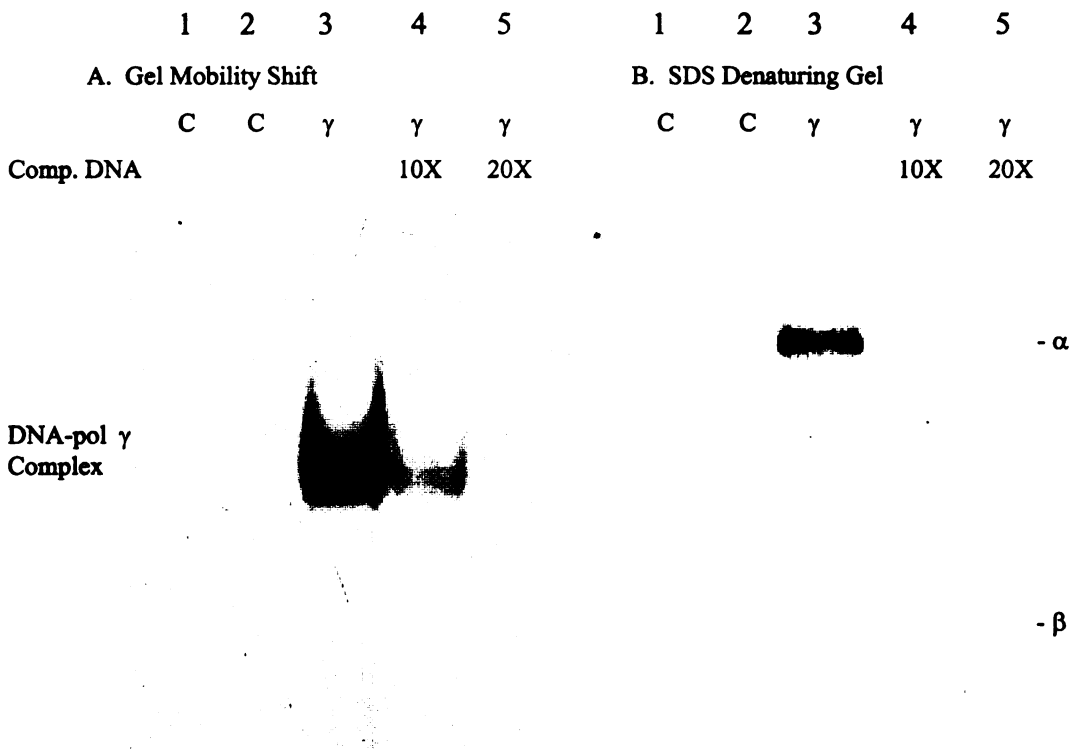
## RESULTS

### *Drosophila* pol γ forms a discrete complex with template-primer DNA

*Drosophila* pol γ catalyzes efficient DNA synthesis on a variety of template-primer DNAs (Wernette, et al., 1988). To begin to elucidate the roles of its two subunits in enzyme function, template-primer DNA binding by native pol γ was examined by gel electrophoretic mobility shift assay and UV-induced cross-linking experiments. A radiolabeled, partially double-stranded oligomer of 40 nt that contains a 3'-terminal bromodeoxyuridylate residue (see Figure 1C) was synthesized and used as the substrate for the DNA binding assay. A single discrete enzyme-DNA complex was observed and the presence of both DNA substrate and *Drosophila* pol γ was necessary and sufficient for the formation of this complex ( Figure 1A). The incorporation of BrdU at the 3' end of the primer strand of the template-primer substrate does not affect the recognition and binding by pol γ. Furthermore, the enzyme-DNA complex formed by *Drosophila* pol γ was detected over a 10-fold range of pol γ to DNA molecules (data not shown), and the stable complex formation could be competed away nearly completely by a 20-fold excess of unlabeled DNA substrate (Figure 1A).

Fig. 1 **Binding and photochemical cross-linking of *Drosophila* pol γ to template-primer DNA.** A, template-primer DNA binding by pol γ. Pol γ Fraction VI (0.36 units, 28 fmol) was incubated with radiolabeled bromodeoxyuridylate-substituted template-primer DNA (shown in C, 0.22 pmol) for 10 min at 30°C as described under "Methods," and the reaction products were electrophoresed in a 4.5% native polyacrylamide gel and the gel autoradiographed. *Lanes 1* and *2* represent no protein and no DNA controls, respectively. *Lanes 3-5* represent samples containing both pol γ and radiolabeled DNA substrate in the absence (*lane 3*) or presence of unlabeled competitor DNA (*lane 4*, 2.2 pmol; *lane 5*, 4.4 pmol). B, photochemical cross-linking of pol γ to template-primer DNA. Pol γ was incubated with template-primer DNA as in A. The samples were then irradiated with UV light (300 nm) for 15 min at 0°C, processed, and electrophoresed in a 7.5% SDS-polyacrylamide gel, and the gel was autoradiographed as described under "Methods." *Lanes 1* and *2* represent no protein and no irradiation controls, respectively. *Lanes 3-5* represent irradiated samples containing both pol γ and radiolabeled DNA substrate in the in the absence (*lane 3*) or presence of unlabeled competitor DNA (*lane 4*, 2.2 pmol; *lane 5*, 4.4 pmol). C, template-primer DNA. A partially double-stranded deoxyoligomer (40 nt) containing [$^{32}$P]-dAMP (*) and bromodeoxyuridylate (*B* in sequence) at its 3'-terminus was prepared as described under "Methods."

Binding and photochemical cross-linking of *Drosophila* pol γ to template-primer DNA




C. DNA Substrate

```
      A
C     C  CAGGCCGTCAAAB-3'
T     A  GTCCGGCAGTTTAATACCTCT-5'
      C
```

FIGURE 1

**The catalytic subunit of *Drosophila* pol γ makes close contact with template-primer DNA**

To probe the involvement of the two subunits of *Drosophila* pol γ in template-primer DNA binding, the enzyme-DNA complex was subjected to photocross-linking in the presence of UV light and the covalently cross-linked enzyme-DNA complex was analyzed by denaturing SDS-polyacrylamide gel electrophoresis and autoradiography (Figure 1B). The catalytic α subunit but not the β subunit in native pol γ can be cross-linked to the radiolabeled BrdUMP-substituted template-primer DNA, and the complex can be competed away by 20-fold excess of unlabeled DNA substrate as well (Figure 1B). Control experiments were also performed to show that UV irradiation did not cause DNA-DNA cross-linking, that the enzyme-DNA complex was stably maintained after UV irradiation, and that the formation of cross-linked products was UV dose dependent (data not shown). The catalytic subunit of native pol γ has a molecular weight of 125 kDa, and the covalently cross-linked enzyme-DNA complex migrated on the denaturing SDS-polyacrylamide gel with an apparent molecular mass of 140 kDa (the calculated molecular weight of substrate DNA is 13 kDa). Both micrococcal nuclease and DNase I could cleave the DNA in the cross-linked complex but yielded products differing in size presumably resulting from the cleavage at different sites on the DNA substrate. Combined action of both nucleases produced a final product very close in size to the molecular weight of the α subunit itself (Figure 2). There was no detectable cross-linking product of 50 kDa, the expected protein-DNA complex if the β subunit alone were cross-linked to the substrate DNA. In addition, we did not detect a cross-linked product of 170 kDa which would be the expected protein-DNA complex if both subunits were covalently cross-linked to the DNA. From the results

Fig. 2 **Photochemical cross-linking of *Drosophila* pol γ to template-primer DNA covalently links the catalytic subunit to primer DNA.** Pol γ (0.36 units, 28 fmol) was incubated with radiolabeled and BrdUMP-substituted template-primer DNA (0,22 pmol), and irradiated with UV light (300 nm) for 15 min at 0°C, and the DNA-enzyme complex was digested with micrococcal nuclease (*lane2*, 1 μg) or DNase I (*lane 3*, 2 μg and *lane4* 10 μg), and electrophoresed in a 7.5% SDS-polyacrylamide gel, and the gel was autoradiographed as described under "Methods." *Lanes 1* represent samples without nuclease digestion. The molecular masses of DNA-protein complexes were determined by the protein standard electrophoresed in the same gel and stained with silver.

Catalytic subunit of *Drosophila* pol γ is cross-linked to the template -primer DNA

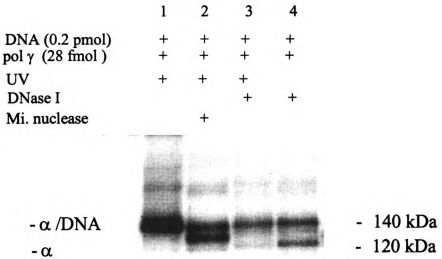



FIGURE 2

of cross-linking experiments we concluded that the template-primer DNA-binding domain(s) of native pol γ resides in the α subunit. Meanwhile, this result does not provide conclusion about the role of the β subunit in DNA-binding. However, these results do indicate that either the β subunit does not bind to DNA directly or that it does not make close contact at the template-primer junction of the DNA substrate when native pol γ holoenzyme binds to DNA.

**The recombinant catalytic subunit of *Drosophila* pol γ is active in template-primer DNA binding**

The same gel electrophoretic mobility shift and UV-induced cross-linking assays were performed using recombinant catalytic subunit of pol γ expressed and purified from *E. coli* (Lewis, et al., 1996) to study the DNA-binding activity of the catalytic subunit. In the absence of the β subunit, the catalytic subunit is also able to recognize the DNA substrate and form a stable discrete complex (Figure 3). Further, photochemical cross-linking of the catalytic subunit-DNA complex induced by UV light demonstrates a specific association of the catalytic subunit with the template-primer DNA. (Figure 3)

**A proteolytic fragment of the catalytic subunit retains DNA binding activity**

To dissect functional domains in *Drosophila* pol γ, the purified native enzyme was subjected to limited tryptic digestion in the presence of template-primer DNA, followed by UV cross-linking and then denaturing SDS-polyacrylamide gel electrophoresis analysis and autoradiography. Limited tryptic digestion of native pol γ produces a form of the

Fig. 3 **Binding and photochemical cross-linking of the recombinant catalytic subunit of *Drosophila* pol γ to template-primer DNA.** A, template-primer DNA binding by the recombinant catalytic subunit of *Drosophila* pol γ. Bacterially expressed catalytic subunit (urea-extracted insoluble fraction, 70% pure) was incubated for 10 min at 30°C with radiolabeled bromodeoxyuridylate-substituted template-primer DNA as described under "Methods," and the reaction products were electrophoresed in a 4.5% native polyacrylamide gel and the gel autoradiographed. *Lanes 1* represents a no protein control. *Lane 2* represents a sample containing Pol γ Fraction VI (0.09 units, 7 fmol) and radiolabeled substrate DNA (0.11 pmol). *Lanes 3-5* represent samples containing0.06 units (96 fmol) of recombinant catalytic subunit and radiolabeled substrate DNA (0.22 pmol) in the absence (*lane 3*) and presence of unlabeled competitor DNA (*lane 4*, 1.1 pmol; *lane 5*, 2.2 pmol). B, photochemical cross-linking of the recombinant catalytic subunit to template-primer DNA. The samples were incubated with labeled substrate DNA as in A and irradiated with UV light (300 nm) for 15 min at 0°C. The DNA-protein complexes were processed and electrophoresed in a 7.5% SDS-polyacrylamide gel, and the gel was autoradiographed. Protein, labeled substrate DNA, unlabeled competitor DNA were as in A. *Lane 1* represents a no protein control. *Lane 2* represents a sample containing Pol γ Fraction VI. *Lanes 3-5* represent irradiated samples containing the recombinant catalytic subunit.

Binding and photochemical cross-linking of the recombinant catalytic subunit of *Drosophila* pol γ to template--primer DNA

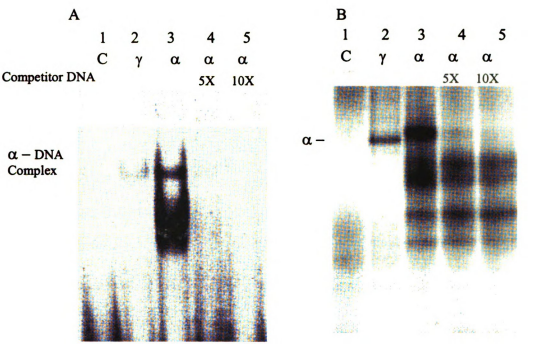



FIGURE 3

Fig. 4 **Photochemical cross-linking of trypsin-digested *Drosophila* pol γ to template-primer DNA.** Pol γ Fraction VI (7 units,0.6 pmol) was incubated for 15 min at 20°C with radiolabeled BrdUMP-substituted template-primer DNA (0.2 pmol), in the absence (*lane 1* and *2*) or presence of 200 ng (*lane 3*) or 400 ng (*lane 4-6*) of trypsin, and the digestion was terminated by addition of a large excess of protease inhibitors as described under "Methods." Following incubation, the samples were irradiated with UV light for 15 min at 0°C, processed, and electrophoresed in a 7.5% SDS-polyacrylamide gel, and the gel was autoradiographed. *Lane 1*, a UV-irradiated, no protein control; *lane 2*, pol γ irradiated without prior trypsin digestion; *lane 3* and *4*, pol γ digested with 200 or 400 ng of trypsin, respectively, prior to UV irradiation; *lane 5*, a UV-irradiated, trypsin only control; *lane 6*, as in lane 2 except that bovine serum albumin was substituted for pol γ.

Photochemical cross-linking of trypsin-digested *Drosophila* pol γ to template-primer DNA

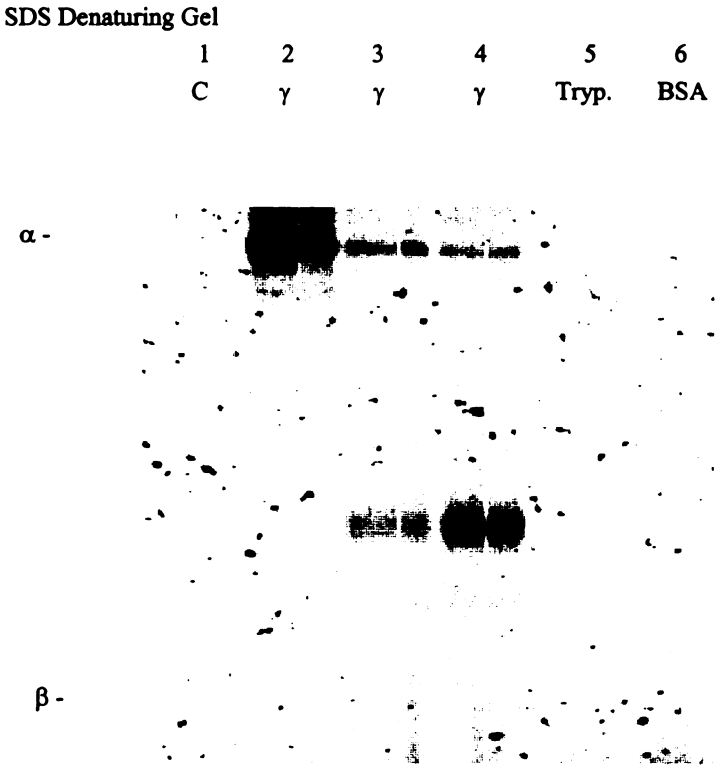



FIGURE 4

enzyme that retains DNA binding activity, at a level that produces a cross-linked product comparable in intensity to the intact enzyme (Figure 4). In the proteolyzed form, the pol γ catalytic subunit is trimmed from a 125- to a 65 kDa DNA-binding polypeptide. Staining of the SDS-polyacrylamide gel with silver nitrate identifies two predominant digestion products of the catalytic subunit of 65 and 55 kDa, and an intact β subunit (data not shown). Notably, the same results are obtained when template-primer DNA is added before or after digestion with trypsin. However, the binding of template-primer DNA appears to protect the enzyme from further degradation, while nearly quantitative conversion of the α subunit from a 125- to a 65-kDa DNA-binding polypeptide is observed in the presence of template-primer DNA (Figure 4), it is cleaved further to yield smaller polypeptides in the absence of template-primer DNA, at a point where about 50% of the α subunit remains intact (data not shown). The data presented here suggest that the 65-kDa polypeptide represents both a structural and a functional domain of the α subunit with respect to DNA binding activity. Whether or not this form of the enzyme exhibits either DNA polymerase or 3'-5' exonuclease activity or if the β subunit remains associated, remains to be determined.

### Cross-linked *Drosophila* pol γ retains DNA polymerase activity

The cross-linked pol γ holoenzyme apparently retains DNA polymerase activity as determined by first producing the covalently cross-linked enzyme-DNA complex with unlabeled DNA substrate and then initiating DNA synthesis in the presence of radiolabeled nucleoside triphosphate. A radiolabeled enzyme-DNA complex can be detected on a denaturing SDS-polyacrylamide gel (Figure 5 lane 3) indicating that pol γ can add an

Fig. 5 **Cross-linked *Drosophila* pol γ retains DNA polymerase activity.** Pol γ Fraction VI (0.35units, 0.03 pmol) was incubated for 10 min at 20 $^0$C with BrdUMP-substituted template-primer DNA (0.2 pmol). Following incubation, the samples were irradiated with UV light for 15 min at 0 $^0$C, and further incubated with [$^{32}$P]-dTTP (*lane 3*) at 25 $^0$C for additional 20 min, processed, and electrophoresed in a 7.5% SDS-polyacrylamide gel, and the gel was autoradiographed. *Lane 2* represents no [$^{32}$P]-dTTP control. *Lane 1* represents a cross-linked DNA-enzmye complex using radiolabeled DNA substrate as control.

Cross-linked *Drosophila* pol γ catalytic subunit retains DNA polymerase activity

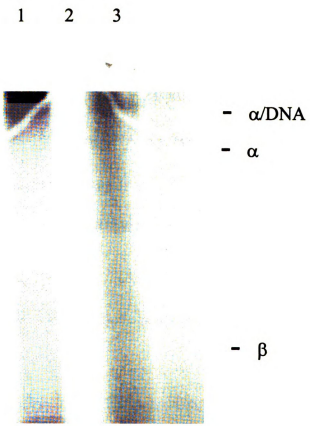



FIGURE 5

additional nucleoside monophosphate to the 3' terminus of the primer even when it is covalently linked to the substrate DNA. This result also suggests that at least a significant amount of DNA-pol γ complex is maintained in a state such that the 3' end of the primer strand is located in the DNA polymerase active site.

### *Drosophila* pol γ footprints two helical turns of its primer

Pol γ holoenzyme purified from *Drosophila* embryos forms a highly stable complex with template-primer DNA. When enzyme/template-primer DNA complexes were subjected to DNase I footprint analysis, about 20 nucleotides from the 3' end on the primer strand were protected against DNase I digestion. The minimum amount of pol γ required to observe a stable footprint on 17 fmol of DNA substrate was 135 fmol in the absence of any other DNA replication accessory proteins (Figure 6A). This result indicates that the pol γ holoenzyme makes close contact with 20 base pairs of duplex DNA, corresponding to two helical turns.

### Mitochondrial single-stranded DNA binding protein enhances template-primer DNA binding by pol γ

Mitochondrial single-stranded DNA-binding protein stimulates pol γ activity (Williams and Kaguni, 1995). This may occur by enhancing primer recognition and binding, by enhancing formation of stable and/or productive pol γ : template-primer DNA complexes, and/or by stimulating nucleotide polymerization *per se*. To evaluate these possible mechanisms of mtSSB stimulation, enzyme/template-primer DNA complexes were subjected to DNase I footprint analysis in the presence of a saturating amount of

Fig. 6 ***Drosophila* pol γ footprints two helical turns of primer DNA and mitochondrial single-stranded DNA-binding protein enhances primer DNA binding by pol γ.** Singly-primed M13 DNA was radiolabeled at the 3'-end of the primer strand and a ddTMP was incorporated to produce a 3'-5' exonuclease-resistant template-primer DNA substrate. A, DNase I digestion reactions were performed by incubating 17 fmol of substrate DNA with 7 to 135 fmol of pol γ fraction VI at 30°C for 5 minutes and incubating in the presence of 0.25 ng/μl of DNase I at 37 °C for an additional 2 minutes. The reactions were quenched and DNAs were recovered by phenol/chloroform extraction and ethanol precipitation. The footprinting products were analyzed by 10% polyacrylamide-7M urea gel electrophoresis and autoradiography. *Lane 1* and *8* represent no protein controls. *Lane 2, 3, 4, 5, 6,* and *7* represent samples containing 7, 14, 27, 55, 83, 135 fmol of pol γ respectively. B, DNase I digestion reactions were performed as in A except saturating amount (15 pmol) of mtSSB was also added in the reaction. *Lane 1* and *8* represent no protein controls. Lane 2 represents a mtSSB only control. *Lanes 3, 4, 5, 6, and 7* represent samples containing 7, 14, 27, 55, 83 fmol of pol γ respectively.

*Drosophila* pol γ footprints two helical turns of primer DNA and mtSSB enhances primer DNA binding by pol γ

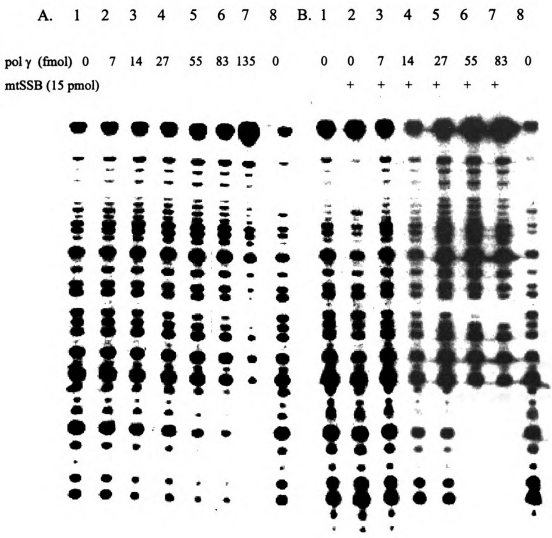



FIGURE 6

mtSSB (15 pmol; 7.5 pmol of mtSSB is required to coat completely the circular M13mp19 single-stranded DNA because each mtSSB monomer binds 17 nucleotides). mtSSB reduced 3- to 4- fold the amount of pol γ required to observe a stable footprint on the primer terminus (Figure 6B). mtSSB binds single-stranded DNA preferentially, and it did not itself yield an observable footprint on the primer DNA strand. The coating of the single-stranded region of the substrate DNA by mtSSB would eliminate potential non-specific DNA binding by pol γ. When non-specific single-stranded DNA-binding of pol γ is greatly reduced by mtSSB, formation of a stable pol γ/template-primer DNA complex is able to reach equilibrium at a lower concentration of enzyme.

## DISCUSSION

Identification of the molecular interactions that occur between a replicative holoenzyme and the template-primer is a prerequisite for understanding the mechanistic contributions of each subunit of the enzyme and the effects of other replication accessory proteins. Experimental data presented here demonstrate clearly that *Drosophila* pol γ can recognize and bind to template-primer DNA, forming a very stable enzyme-DNA complex at a molar ratio as low as 0.1 pol γ molecule/template-primer DNA, without the assistance of any other replication proteins. In contrast, several other DNA polymerases known to associate with accessory proteins for catalytic function, including bacteriophage T4 DNA polymerase (Munn and Alberts, 1991), *E. coli* DNA polymerase III (Reems and McHenry, 1994), and calf thymus DNA polymerase δ, do not form a stable complex with

template-primer DNA in the absence of accessory proteins. That *Drosophila* pol γ does so, and catalyzes relatively efficient DNA synthesis on a variety of template-primer DNAs in the absence of accessory proteins (Wernette, et al., 1988), suggests that such factors are not required for mitochondrial DNA replication. However, all of our experiments were performed *in vitro* and on a model DNA substrate. mtSSB stimulates DNA synthesis by pol γ *in vitro* (Thommes, et al 1995), raising the possibility that it and potentially other accessory proteins might enhance the enzyme-DNA interactions of pol γ *in vivo,* particularly because the concentration of primer termini in mitochondria is much lower. Indeed, the DNase I footprint analysis reveals that mtSSB enhances the specific interaction between pol γ and template-primer DNA. In fact, the DNase I footprint analysis of primer binding may represent an underestimate of the effect of increased binding on catalysis because this experimental approach only measures stable enzyme-DNA complexes. mtSSB also exhibits stimulation by additional mechanisms, such as eliminating secondary DNA structures in the template DNA.

Photochemical cross-linking of protein-DNA complexes formed with either pol γ holoenzyme or recombinant catalytic subunit revealed that the catalytic subunit but not the β subunit makes close contact with the DNA. At the same time, a super-shift assay using subunit-specific antibody in the DNA-binding reaction suggests that the β subunit is physically present in the pol γ-DNA complex (data not shown). Furthermore, the recombinant catalytic subunit alone exhibits very similar DNA-binding activity as compared to the native holoenzyme, and limited proteolysis of the complexes with trypsin identified a 65 kDa proteolytic intermediate of the α subunit that retains DNA binding activity. The 65

kDa proteolytic intermediate of the α subunit is stabilized by association with DNA during protease digestion. These results suggest that native pol γ assumes a different conformation when complexed with the DNA substrate and that the accessory subunit is neither required for nor does it affect the DNA binding activity. Bacteriophage T4 DNA polymerase can bind DNA in three different modes. The location of the primer terminus can be either in the DNA polymerase active site, 3'-5' exonuclease active site, or in an intermediate transition site (Baker and Reha-Krantz, 1998). The exact location of the 3'-end of the primer DNA terminus can not be determined conclusively in our experimental approach, but it is clear that pol γ can initiate DNA synthesis efficiently only when the primer terminus is directed to the polymerase active site. We found that the cross-linked pol γ holoenzyme retains the ability to add additional nucleotides to the 3'-end of primer strand (Figure 4). This result suggests that the 3'-end of primer strand is oriented in the polymerase active site.

mtSSB may facilitate productive template-primer DNA binding and nucleotide polymerization by pol γ by the combined actions of eliminating non-specific interaction between pol γ and single-stranded DNA, eliminating secondary DNA structures in template DNA, promoting DNA-synthesis competent template-primer binding and promoting the transition from initiation to elongation. *E. coli* Pol III holoenzyme is stimulated by SSB protein, but this stimulation is dependent on the specific protein-protein interaction between the SSB and the χ subunit in the γ complex (Kelman, et al., 1998; Glover and McHenry, 1998). Other DNA polymerases are stimulated by their cognate SSBs and interact with them. For example, T4 gp32 protein and T7 gene 2.5 protein are both SSBs that bind to and stimulate T4 and T7 DNA polymerases respectively. In addition, human SSB (RP-A)

has been shown to interact directly with DNA pol α. Although we could not determine the biochemical function of the accessory subunit in these experiments, it is possible that the β subunit is required for pol γ to interact with other replication proteins including mtSSB. Other possible functions of the β subunit might include mediating the interaction between pol γ and the RNA primer terminus at mitochondrial DNA replication origins, and conferring processivity upon pol γ.

## FUTURE RESEARCH

We have shown previously that the association of the two subunits of *Drosophila* pol γ is so strong that the conditions required to physically separate them results in the denaturation of the protein and complete loss of enzymatic activity (Olson, et al., 1995). Here we demonstrated that the catalytic subunit makes close contact with the template-primer DNA. Protein-protein cross-linking reagents can be added to the DNA/pol γ/mtSSB complex to detect potential physical interactions between pol γ and mtSSB and determine which subunit contacts mtSSB. Whether the accessory subunit contains DNA-binding activity can also be assayed with a recombinant form of the β subunit. We have performed photochemical cross-linking studies with a photo-reactive nucleotide incorporated into the primer strand of a DNA substrate. Interaction between the β subunit and template-primer DNA can also be examined with DNA containing a photo-reactive nucleotide in the template

strand or at upstream postitions in the primer strand. When a large quantity of recombinant pol γ is available, photochemical cross-linking and limited proteolysis experiments will become very useful to determine the specific DNA binding region of pol γ.

We performed our footprinting analysis with a labeled primer strand. A footprinting study to observe the overall DNA binding by pol γ will require further experiments to determine the protected region on the template strand DNA. We have demonstrated that mtSSB protein enhances the template-primer DNA binding activity of pol γ. Footprinting studies can also provide us an assay to probe the effect of other replication proteins in their actions at the replication fork.

Reha-Krantz *et al.* used a base analog, 2-aminopurine, inserted at the 3' terminus of the primer strand as a fluorescent reporter to analyze the partitioning of primer DNA strands between the exonuclease and polymerase active site of bacteriophage T4 DNA polymerase (Reha-Krantz, et al., 1998, Beechem, et al., 1998). The 3'-5' exonuclease activity of pol γ is highly mismatch specific (Olson and Kaguni, 1992). We could use this fluorescent reporter substrate to examine the template-primer DNA binding mode of pol γ and determine how the mtSSB protein stimulates the polymerase and exonuclease activities of pol γ.

# CHAPTER III

# MOLECULAR CLONING OF THE ACCESSORY SUBUNIT OF *DROSOPHILA* POL γ

# INTRODUCTION

It is now well established that animal mitochondria contain a nuclear-encoded DNA polymerase, which is capable of performing high fidelity mtDNA replication (Kaguni and Olson, 1989) and is very likely also involved in mtDNA repair (Pinz and Bogenhagen, 1998; Longley, et al., 1998a). A second DNA polymerase activity has been reported in yeast mitochondria (Lucas, et al., 1997), and a DNA polymerase bearing sequence homology to pol β has been identified in mitochondria of the protozoan, *Crithidia* (Torri and Englund, 1995). However, it remains to be determined whether animal mitochondria also contain another DNA polymerase. The genome sequencing projects in progress and subsequent biochemical studies are likely to answer this question in the future, because all of the proteins required for mitochondrial DNA replication are encoded by nuclear genes.

*D. melanogaster* pol γ in its native form purified from embryos is a heterodimer of 125 kDa and 35 kDa subunits (Wernette and Kaguni, 1986). The large 125 kDa catalytic subunit contains both 5'-3' DNA polymerase and 3'-5' exonuclease activities (Lewis, et al., 1996), while no specific biochemical function has been assigned to the 35 kDa subunit. It has been shown that the *Drosophila* β subunit is a distinct polypeptide and is important to maintain the catalytic efficiency and/or the structural integrity of the holoenzyme (Olson, et al., 1995). The catalytic subunits of pol γ from frog, man, mouse, *S. cerevisiae, S. pombe,* and *P. pastoris* have also been cloned (Lecrenier, et al., 1997). In one report, pol γ purified from human Hela cells contained a polypeptide of 54 kDa in addition to the catalytic subunit (Gray and Wong, 1992). The purified porcine liver pol γ contains four polypeptides of 120, 55, 50, and 48 kDa (Kunkel and Mosbaugh, 1989), and a preparation

of pol γ from *Xenopus laevis* embryos possesses several polypeptides of 100, 85, 55, 40, and 31 kDa in association with the 140 kDa catalytic subunit (Insdorf and Bogenhagen, 1989a). Some of these smaller polypeptides are likely proteolytic products of the catalytic subunit that are derived during the purification. Whether pol γ is a single polypeptide enzyme, a heterodimer, or has a variable subunit structure among different species remains an unresolved issue.

The study presented here documents the molecular cloning of the β subunit of pol γ from *D. melanogaster*, and the subsequent identification of its homologs in man and mouse. The determination of the genomic structure of the *Drosophila* pol γ genes and identification of potential transcriptional regulatory elements in the promoter sequence of both genes is also presented.

# EXPERIMENTAL PROCEDURES

## Materials

*Nucleotides and Nucleic Acids*- Unlabeled deoxyribonucleoside triphosphates were purchased from P-L Biochemicals. [$\gamma$-$^{32}$P]ATP was purchased from ICN. Plasmid pUC119, pET-11a, pET-16b and $\lambda$gt11 DNAs were prepared by standard laboratory methods. Synthetic oligodeoxynucleotides as indicated in the text were synthesized in an Applied Biosystems model 477 oligonucleotide synthesizer.

*Enzymes and Proteins- Drosophila* pol $\gamma$ Fraction VI was prepared as described by Wernette and Kaguni (Wernette and Kaguni, 1986). T4 polynucleotide kinase, T4 DNA ligase and nuclease S1 were purchased from Boehringer Mannheim. *Taq* DNA polymerase was from GIBCO-BRL. Exonuclease III was from Life Technologies, Inc.

*Bacterial strains-E.coli* LE392 *hsdR514, hsdM, supE44, supF58, lac Y1 or (lacIZY)6, galK2, galT22, metB1, trpR55* was used for screening a $\lambda$gt11 ovarian cDNA library from *Drosophila melanogaster* (the generous gift of Dr. Chuen-Sheue Chiang, Stanford University). *E. coli* XL-1 Blue *recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac, (F'proAB, lacI$^q$Z M15,* Tn10(tet$^r$)) was used to subclone the $\beta$–subunit cDNA for sequence analysis. *E. coli* BL21 lDE3 *ompT*, $r_B^-$ $M_B^-$ was used for the expression of pET-11a and pET-16b (Novagen) constructs.

*Chemicals*- Isopropylthio-β-D-galactoside, nitro blue tetrazolium, and 5-bromo-4-chloro-3-indolyl phosphate were purchased from Sigma. Sodium metabisulfite and leupeptin were purchased from the J. T. Baker Chemical Co. and the Peptide Institute, Minoh-Shi, Japan, respectively.

## Methods

*Sequence Analysis of D. melanogaster pol γ- D.melanogaster* pol γ (255 pmol), purified from 1120 g embryos of average age 9 hours, was processed for peptide sequencing as described by Lewis et al (Lewis, et al, 1996). Protein sequence analysis was performed at Harvard MicroChem (Harvard University).

*Cloning of the β subunit of D. melanogaster pol γ-* One of the peptide sequences IVPHDLAEDLNPNDYQAIDIR was used to generate two degenerate oligonucleotide primers for amplifying relevant sequences from a λgt11 cDNA library derived from *D. melanogaster* ovarian mRNA by polymerase chain reaction (PCR). The degenerate-sense forward primer coding for the peptide sequence DLAEDL was 5'-GAT(CT)T(CG)GC(ACGT)GA(AG)GA(CT)(CT)T-3'. The degenerate-antisense reverse primer corresponding to the peptide sequence YQAIDI was 5'-AT(AG)TC(AGT)AT(ACGT)GC(CT)TG(AG)TA-3'. The PCR amplification reaction was performed for 30 cycles of 94°C at 90 s, 50°C for 90 s, and 72°C for 60 s in a reaction volume of 0.1 ml containing 20 mM Tris-HCl, pH 8.4, 50 mM KCl, 4 mM $MgCl_2$, 1 mM dithiothreitol, 200 μM each dATP, dCTP, dGTP, dTTP, 0.5 μg of λgt11 DNA, 50 pmol of forward oligonucleotide primer, and 4 pmol of 5'-$^{32}$P-labeled reverse oligonucleotide

primer ($1.1x10^6$ cpm/pmol), and 1.25 units of *Taq* DNA polymerase. The expected PCR product of 47 bp was obtained, and purified by electrophoresis in a native 15% polyacrylamide gel. The DNA extracted from the gel was subjected to sequencing by the Maxam and Gilbert method (Maxam and Gilbert, 1980). Based on the nucleotide sequence obtained, a 23-nucleotide deoxyoligomer, 5'-AGGACCTCAATCCCAACGACTAC-3', was synthesized and used to screen the λgt11 ovarian cDNA library by the Benton and Davis method (Benton and Davis, 1977) using *E. coli* LE392 as the bacterial host.

Seven positive clones were obtained in the teriary-screen, and one plaque pure isolate with a 1.7-kilobase pair insert was subcloned and sequenced in its entirety on both strands. The insert cDNA was PCR amplified and subcloned into plasmid pUC119 at its unique *EcoRI* site using *E. coli* XL1-Blue as the bacterial host. Fourteen nested deletion plasmids were generated using Exonuclease III and S1 nuclease (Henikoff, 1984). A nucleotide sequence of 1569 bp was determined on both strands by automated fluorescent DNA sequencing using the Applied Biosystems Catalyst 800 for *Taq* cycle sequencing and model 373 DNA Sequencer for the analysis of products. The complete cDNA sequence was assembled using the Sequencher version 2.1.1 software package.

*Identification of the human and mouse homologs of the β subunit of mitochondrial DNA polymerase-* Multiple human and mouse EST cDNA sequences were detected from the dbEST (Expressed Sequences Tags) database of the National Center for Biotechnology Information (NCBI) by the BLAST program using the peptide sequence of the β subunit of *Drosophila* pol γ. The partial C-terminal sequence of the putative human and mouse β subunit homologs showed more than 40% identity to the corresponding region of *Drosophila* protein at the amino acid level. One EST cDNA clone containing the putative

human (accession number H05453) and another EST cDNA clone containing the potential mouse β subunit homolog (accession number AA153407) were purchased from Genome System Inc. (8620 Pennell Drive, St. Louis, MO63114). The complete DNA sequences of these two clones were determined by automated fluorescent DNA sequencing using the Applied Biosystems Catalyst 800 for *Taq* cycle sequencing and model 373 DNA Sequencer for the analysis of products. The complete cDNA sequences for both human and mouse homologs were assembled using the Sequencher version 2.1.1 software package.

# RESULTS

## Cloning of the β Subunit of *Drosophila* DNA Pol γ

Five amino-terminal sequences were obtained from the purified native β-subunit polypeptide following tryptic digestion and fractionation of the resulting peptides by microbore HPLC (Figure 1). Many attempts were made to screen a λgt11 cDNA library derived from *D. melanogaster* ovarian mRNA (gift of Dr. Chuen-Sheue Chiang, Stanford University) by using degenerate oligonucleotides as probes. However, a large number of false positive clones made this experimental approach impractical (data not shown). Degenerate oligonucleotides corresponding to the amino- and carboxyl-terminal amino acid residues of one of the peptides were constructed, and used to amplify *D. melanogaster* cDNA fragments by PCR as described under "*Methods*." A correctly-sized PCR product was recovered by gel purification, and DNA sequence analysis confirmed that it corresponded to the amino acid sequence of the tryptic peptide. A 23 nt DNA probe was then synthesized, and used to screen a λgt11 cDNA library derived from *D. melanogaster* ovarian mRNA.

Tertiary-screen positive clones were obtained at a frequency of $1.5X10^{-6}$ phage screened, a frequency comparable to that found in screening the same library for catalytic subunit cDNA clones ($5X10^{-6}$) (Lewis, et al., 1996). The largest cDNA was 1.7 Kb in length and the nucleotide sequence of clone # 210-1 was determined in its entirety on both DNA strands. DNA sequences of 5'- and 3'- ends of four additional independently-isolated cDNA clones were obtained. All contained an identical 3'- end sequence, and differed only in the length of the 5'-untranslated region (data not shown).

Fig. 1 **Nucleotide and deduced amino acid sequences of the β subunit of *Drosophila* DNA polymerase γ.** Amino acid residues enclosed in *boxes* correspond to those for which the amino acid sequence was determined by peptide sequencing as described under "*Methods*". Nucleotides are numbered from the putative transcriptional initiation site. The putative polyadenylation signal sequence is indicated by a *solid line* above the nucleotide sequence. (GenBank accession number: U94702)

# Nucleotide and deduced amino acid sequences of the β subunit of *Drosophila* DNA polymerase γ

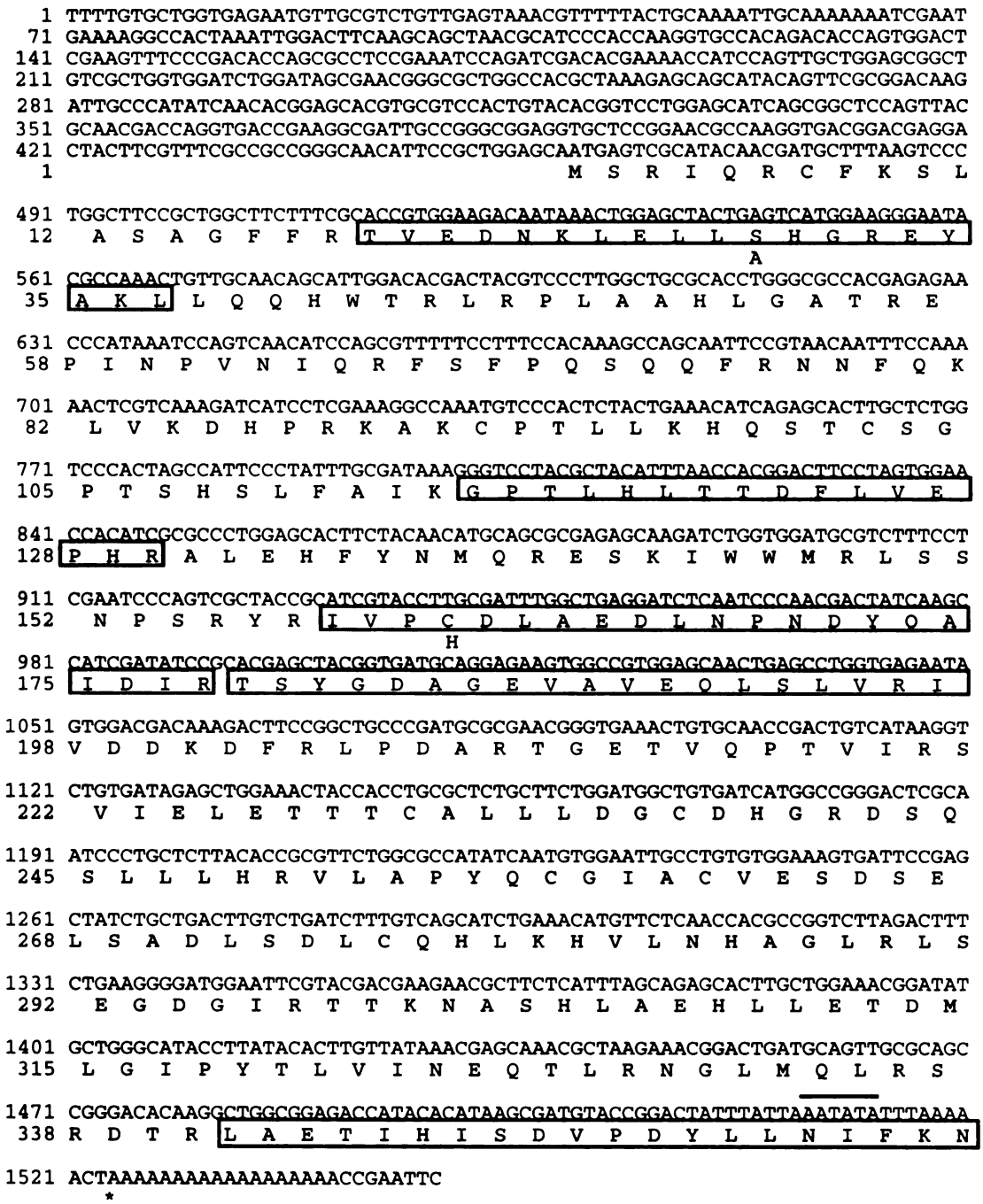


FIGURE 1

An open reading frame of 1083 nucleotides was identified which encodes the β-subunit polypeptide of 361 amino acid residues with a predicted molecular mass of 41 kDa. This is consistent with the size of 35 kDa determined for the small subunit in purified native pol γ. The amino-terminal peptide sequences of all five tryptic peptides derived from the β subunit of the purified native enzyme are located within the deduced animo acid sequence of the β-subunit cDNA. However, there are two amino acid residues showing conflicting sequences between the deduced amino acid sequence and the amino acid sequence obtained from the purified native β subunit (Figure 1). There is no other in-frame ATG codon between the 5' end of cDNA and the translational initiation codon, but multiple in-frame stop codons are present in the 5' untranslated region. The amino-terminal sequence of the mature β subunit was not determined experimentally because of the limited amount of β-subunit polypeptide derived from native pol γ. However, because the amino terminus of one of the tryptic peptides is located at position +19 in the deduced amino acid sequence, the mitochondrial presequence peptide of β subunit is no longer than 18 amino acid residues. This is comparable with the mitochondrial presequence peptide of the α subunit, which was determined experimentally to be 9 residues (Lewis, et al., 1996).

To investigate further the discrepancy between the peptide sequence and deduced amino acid sequence of the β-subunit cDNA, DNA sequences from nt # 950 to 1250 were also determined for all seven β-subunit cDNA clones, and the entire cDNA sequence of clone # 211-5 was determined on both strands. Comparison of cDNA and genomic sequence demonstrates that two classes of cDNA clones are isolated and they all represent

normal transcripts from the β-subunit gene. These two classes of cDNAs all encode the 35 kDa subunit with substitutions in residue # 188 and 213 in the deduced amino acid sequences. The first class is represented by clone # 210-1, and the amino acid residues at # 188 and 213 are A and T respectively. The second class is represented by clone # 211-5, and the amino acid residues at # 188 and 213 are T and I respectively. The genomic DNA sequence provided by the Berkeley *Drosophila* Genome Project (BDGP, http://www.fruitfly.edu) matches well with the cDNA clones represented by #211-5. This may reflect the normal polymorphism among biologically functional alleles.

**Identification of human and mouse homologs of the β subunit of *Drosophila* pol γ**

No homolog of the β subunit of *Drosophila* pol γ was identified upon extensive searching of the latest release of the GenBank and EMBL databases (releases 107 and 57 respectively), or the *Saccharomyces* Genome Database. However, multiple human, mouse, and rat partial cDNA sequences were detected in the dbEST (Expressed Sequences Tags) database of the National Center for Biotechnology Information (NCBI) using the BLAST searching program with the deduced amino acid sequence of the β subunit of the *Drosophila* pol γ as query. For determining both human and mouse homolog DNA sequences, we obtained EST clones (Genome Systems Inc. 8620 Pennell Drive, St. Louis, MO63114) containing possible human and mouse homologs respectively. The human cDNA sequence is 1594 bp in length and contains a complete open reading frame. The human β subunit homolog has a deduced amino acid sequence of 372 residues (Figure 2), with a calculated molecular mass of 42.4 kDa. The mouse cDNA sequence is only 1127

Fig. 2 **Sequence alignment of the β subunit of *Drosophila* pol γ and its mammlian homologs.** Amino acid residues that are conserved and similar are represented in uppercase. For mouse and rat sequences, only the C-terminal partial cDNA sequences have been determined. Human and mouse sequences have been deposited in GenBank and their accession numbers are U94703 and AF006072 respectively. The rat sequence is deduced from the EST cDNA sequence obtained from GenBank dbEST database (accession number AA892950).

## Sequence alignment of the β subunit of *Drosophila* pol γ and its mammalian homologs

```
Dm                msriq rcFkslAsag ffrtVe..Dn kleLLShgre yakLlqqhwT
Hs             mvdlgggvhg avFpvdAlhh kpspLlpgDs afrLVSaetl reiLqdkelS
Mm
Rn
CONSENSUS      ---------- --F---A--- --------D- ---L-S---- ---L------

Dm             RlrpLAAhLg atrepinpVn iQrfsfpqsQ qFrnNfqkLV .KdhPrkAkc
Hs             K.eqLVAfLe nvlktsgkLr .EnllhgaLE HYV.NCLDLV NKRLPYGLAQ
Mm                                       LE HYV.NCLDLV NRKLPFGLAQ
Rn
CONSENSUS      ----L-A-L- ---------- --------LE HYV-NCLDLV NK-LP-GLAQ

Dm             ptLlkHqsTc SgpTshSLfa IkGp.Tlhlt tdFlvePhRA lehFyNmq.R
Hs             IGVCFHPVfd TkQirngVks I.GEKTEASL VWFT..PpRT SnQWlDFWLR
Mm             IGVCFHPVSn SnQTpsSVtr V.GEKTEASL VWFT..PtRT SsQWsDFWLR
Rn
CONSENSUS      IGVCFHPV-- S-QT--SV-- I-GEKTEASL VWFT--P-RT S-QW-DFWLR

Dm             eskIWWmRls snPSRYrivp CDlaEdln.p NdYqaidirt syGDaGevAV
Hs             HRLqWWRKFA MSPSNFSSsD CQDEEGRKGT NfttiFPWGK ELIETLWNLG
Mm             HRLLWWRKFA MSPSNFSSaD CQDElGRKGS klYysFPWGK EPIETLWNLG
Rn
CONSENSUS      HRL-WWRKFA MSPSNFSS-D CQDEEGRKG- N-Y--FPWGK E-IETLWNLG

Dm             EQlsLvrivd DkdfrlpdaR tGet.VqPtV IrSV.IELEt ttcAlLlDgc
Hs             DhELLHmYPG NVSkLhG..R DGRKNVVPCV L.SVnGDLDr GmLAYLYDSF
Mm             DQELLHtYPG NVStIqG..R DGRKNVVPCV L.SVsGDVDl GtLAYLYDSF
Rn
CONSENSUS      D-ELLH-YPG NVS---G--R DGRKNVVPCV L-SV-GDLD- GTLAYLYDSF

Dm             DhgrDS...q slLlHRVLap yqCGI.AcVE sDsel..SAD LsDLCQhLkh
Hs             QLtENSFtRK KnLhRKVLKL HPCLAPIKVA LDVGRGPTLE LRQVCQGLfN
Mm             QLaENSFaRK KsLqRKVLKL HPCLAPIKVA LDVGKGPTVE LRQVCQGLLN
Rn                                  HPCLAPIKVA LDVGKGPTVE LRQVCQGLLN
CONSENSUS      QL-ENSF-RK K-L-RKVLKL HPCLAPIKVA LDVGKGPTVE LRQVCQGLLN

Dm             vLnhaGlrLS egdGirtTkn aShlAEHLle t.DmLGIPYT LVInEqTLrN
Hs             ELLENG..IS VWPGYLETmQ SS..LEQLYS KYDEMSILFT VLVTETTLEN
Mm             ELLENG..IS VWPGYsETvh SS..LEQLhS KYDEMSVLFS VLVTETTLEN
Rn             ELLENG..Ia VWPGYLETaQ SS..LEQLYS KYDEMSILFT VLITETTLEs
CONSENSUS      ELLENG--IS VWPGYLET-Q SS--LEQLYS KYDEMSILFT VL-TETTLEN

Dm             GLMQLRSRDT rLaEtIHISd VpDYLLnifk n*
Hs             GLIHLRSRDT TMKEMMHISK LKDFLIKYIs SAkNV*
Mm             GLIQLRSRDT TMKEMMHISK LRDFLVKYLA SAsNV*
Rn             GLIQLRSRDT TMKEMMHISR LRDFLVKYLA SAgkA*
CONSENSUS      GLIQLRSRDT TMKEMMHISK LRDFLVKYLA SA-NV-
```

FIGURE 2

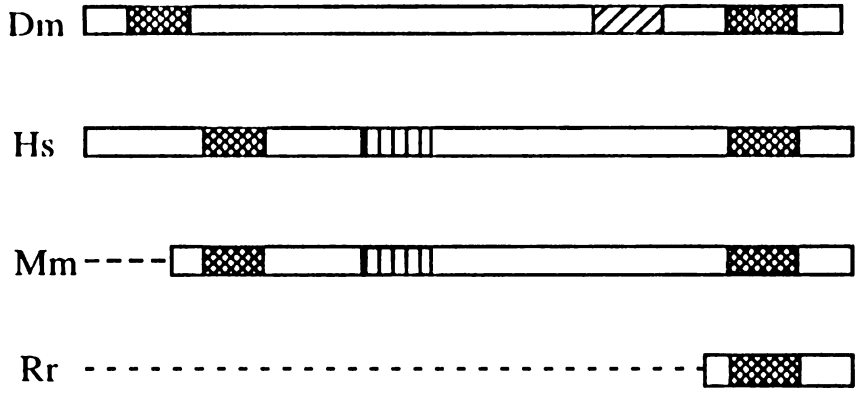
bp in length and represents the C-terminal region of the gene; it contains an open reading frame with a deduced amino acid sequence of 296 residues (Figure 2). The mouse β-subunit homolog likely has a similar primary structure because of its high sequence identity (81%) with the human homolog, although its N-terminal sequence remains to be determined. Alignment of the complete sequences of *Drosophila* and human, and partial sequences of mouse and rat using the Gap program of the GCG program package (version 7, Genetics Computer Group) show 30% amino acid sequence identity and 60% sequence similarity overall.

The C-terminal region is the most conserved of all three proteins. Two putative leucine zipper motifs of the sequence L(X)6L(X)6L(X)6I and L(X)6L(X)6L(X)6L (*Drosophila*, residue #38 - 59, #321 - 342) and L(X)6L(X)6L(X)6I and V(X)6L(X)6L(X)6M (human, residue #77 - 98, # 328 - 349) are identified at the N- and C-terminal region of the *Drosophila* and human sequences (Figure 3). The leucine zipper motifs in the mouse sequence are almost identical to the human sequence with only one or two conserved amino acid residue substitutions in each motif (Figure 3). Further, the *Drosophila* β subunit contains a zinc finger motif of the C2H2 class located between the two leucine zipper motifs. Although the human and mouse β subunits lack this zinc finger motif, they contain a zinc-binding signature sequence (LGDHELLHNMY) found in the zinc-dependent neutral metallopeptidase family proteins (Murphy, et al., 1991). Another notable feature of both the fly and human proteins is that the C-terminal regions are very leucine rich: 18% and 17%, respectively, of the 120 C-terminal residues are leucines.

Protein structure prediction was performed by the BENCHMARK method (http://www.mbi.ucla.edu/people/fischer/BENCH/benchmark1.html). The fly and human proteins were shown to have similarity with tRNA synthetases based on predicted

Fig. 3 **Putative structural features of the β subunit of *Drosophila* pol γ and its human, mouse, and rat homologs.** The leucine zipper motifs in the *Drosophila* and mammalian sequences are located in the homologous region at the N- and C-termini. Also identified are a zinc finger motif in the *Drosophila* sequence and a zinc-binding signature sequence (ZBS) in the human and mouse sequences.

Putative structural features of the β subunit of *Drosophila* pol γ and its human, mouse, and rat homologs




| | | | |
|---|---|---|---|
| **Zinc finger** | | | $C(X)_3C(X)_{19}H(X)_3H$ |
| **ZBS** | Hs | | LGDHELLHMY |
| | Mm | | LGDQELLHTY |
| **Leu Zipper 1** | Dm | | $L(X)_6L(X)_6L(X)_6I$ |
| | Hs & Mm | | $L(X)_6L(X)_6L(X)_6I$ |
| **Leu Zipper 2** | Dm | | $L(X)_6L(X)_6L(X)_6L$ |
| | Hs & Mm &Rr | | $V(X)_6L(X)_6L(X)_6M$ |
| **Not determined** | | | |

FIGURE 3

secondary structures. The *Drosophila* β subunit is also predicted to contain a three-dimensional fold found in moloney murine leukemia virus reverse transcriptase while the human β-subunit protein is predicted to resemble part of HIV reverse transcriptase. These predictions are potentially significant because all the proteins described above are involved in enzymatic interactions with RNA or DNA molecules.

**Genomic DNA structure of the *Drosophila* β subunit gene**

Two genomic sequences containing the β-subunit gene were deposited into the GenBank by the *Drosophila* Genome center at UC, Berkeley (accession # L39626 and L39627). These two genomic sequences are from a P1 genomic clone (DS00941) which is mapped to the left arm of chromosome 2 (subdivision 34D2-34D7) near the region containing alcohol dehydrogenase gene (subdivision 35B). Remarkably, the same P1 clone contains the catalytic-subunit gene of *Drosophila* pol γ (Lewis, et al., 1996). Further, the genes encoding the two subunits of native pol γ holoenzyme are separated only by an interval sequence of 3.8 kilobase pairs, with the β-subunit gene 5'- to the α-subunit gene (Figure 4).

Comparison of the genomic and cDNA sequences indicates that the β-subunit gene contains one small intron of 54 nt, that splits the codon specifying amino acid residue A231 in the deduced amino acid sequence. The intron/exon boundaries are GTAAGT (donor site) and TAG (acceptor site) respectively, consistent with the GT.....AG rule of intron/exon border sequence. The coding sequence of the exon in the genomic clones is 99% identical to that of the corresponding regions of the cDNA. There are two single

Fig. 4 **Genomic DNA structure of *Drosophila* pol γ genes**. The organization and structure of the *Drosophila* pol γ genes are represented schematically. The *boxes* represent the exons: *closed boxes* represent the protein-coding sequences, and *open boxes* represent either 5'- or 3'- untranslated sequences. The *thick line* represents the introns and intragenic sequences. Genomic sequences upstream of the putative transcriptional initiation sites of both genes are enlarged to indicate locations of potential DNA-binding sites for transcription factors. The *numbers* indicate nucleotide positions relative to putative transcriptional initiation sites. *KB*, kilobase pair.

Genomic DNA structure of *Drosophila* pol γ genes

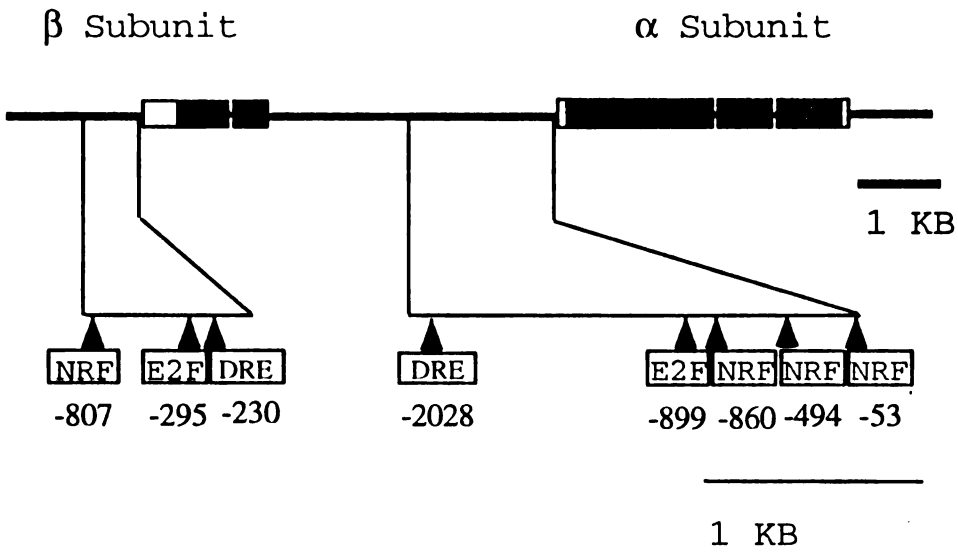



FIGURE 4

nucleotide deletions in the genomic sequence that disrupt the reading frame, and 11 mismatches between the genomic and cDNA sequences represented by clone # 210-1. Of the latter, only 5 of those mismatches would result in amino acid residue substitutions. Interestingly, the deduced amino acid sequence of the second class of β-subunit cDNA represented by clone # 211-5 matched well with that of the genomic sequence. Thus, the differences between them may well represent normal genetic polymorphism.

An additional ORF encoding a protein with predicted molecular mass of 42 kDa is identified in the genomic sequence between the β-subunit gene and the α-subunit gene. This gene may also encode a mitochondrial protein ( Dr. Rafael Garesse, Departmento de Bioquimica, Facultad de Medicina, Universidad Autonoma de Madrid, Spain, unpublished results).

A search for RNA polymerase II promoter elements in the genomic DNA sequence in the region upstream of the putative transcription initiation site (5'- end of the longest cDNA sequence) did not reveal a consensus TATA sequence. A sequence matching the consensus transcriptional initiator TCAGT (Cherbas and Cherbas, 1993) is located in the genomic sequence 16 bp upstream of the 5'- end of the cDNA sequence, and the translational initiation ATG codon is located at a position 459 bp downstream. (Figure 1) Two additional consensus transcriptional initiator sequences are located within 500 bp in the upstream region. Interestingly, a *Drosophila* promoter-activating element, DRE (DNA replication-related element), previously identified in a number of genes involved in nuclear DNA replication (Matsukage, et al., 1995), is located 230 bp upstream of the putative transcriptional starting site (Figure 4). Further, potential binding sites for nuclear respiratory factor 1 (NRF-1) (Virbasius, et al., 1993) and transcription factor E2F (Ohtani and Nevins, 1994) are located at position -807 and -295, respectively (Figure 4). Notably, consensus DRE element (position -2028), and potential NRF-1 (positions -860, -494, -53)

and E2F (position -899) binding sites, are also present in the upstream region of the catalytic subunit gene of *Drosophila* pol γ (Figure 4 ).

The 3'- end of the β-subunit cDNA clone contains an 18 nt poly (A) sequence. Surprisingly, the poly (A) sequence begins within the translational termination codon, TAA (Figure 1). Based on the comparison of the cDNA and genomic sequences, the polyadenylation site is located no more than six nucleotides downstream from the termination codon. This feature is common in mitochondrially-encoded genes, and may be of evolutionary significance. As with the catalytic subunit cDNA, no match to the consensus poly(A) signal sequence (AATAAA) is present in the β-subunit cDNA, although 23/24 of the nucleotides immediately upstream of the termination codon are deoxyadenylate or thymidylate residues. Interestingly, both the catalytic- and β-subunit cDNAs contain the sequence of AATATA (at positions -51 and -16, respectively, relative to the polyadenylation site), which might serve as the poly(A) signal. This alternative poly(A) signal has been implicated in directing polyadenylation of mRNAs of *Drosophila* cyclin A (Lehner and O'Farrell, 1989), *bicoid* (Berleth, et al., 1988), *orb* (oo18 RNA-binding protein) (Lantz, et al., 1992), and *Pgc* (*Polar granule component*) genes (Nakamura, et al., 1996). The expression of these genes have been shown to be developmentally regulated.

Taken together, the presence and locations of common transcriptional and post-transcriptional regulatory signals in the catalytic and accessory subunit genes of *Drosophila* pol γ suggest the possibility of coordinate gene regulation of the heterodimeric mitochondrial DNA polymerase.

# DISCUSSION

The redundant nature of mitochondrial DNA does not eliminate the requirement for high fidelity DNA replication, because the accumulation of mutations in mtDNA is linked to many degenerative diseases in humans (Wallace, 1995a). The most important replicative enzyme involved in mtDNA replication is pol γ. *Drosophila* pol γ (Wernette and Kaguni, 1986) and pol γ from many species including yeast (Foury, 1989) , frog (Insdorf and Bogenhagen, 1989b), chick, pig (Kunkel and Mosbaugh, 1989), and human (Gray and Wong, 1992) have been demonstrated to contain dual enzymatic activities. They possess both highly accurate DNA polymerase and mispair-specific 3' - 5' exonuclease activities. While we have shown previously that *Drosophila* pol γ is a heterodimer of 125 kDa and 35kDa subunits (Olson, et al., 1995), the structure of other mitochondrial DNA polymerases remains a controversial issue.

In a previous report, physical and immunological evidence had demonstrated clearly that the 35kDa β subunit of *Drosophila* pol γ is structurally distinct from the 125 kDa catalytic subunit (Olson, et al., 1995). The experimental data presented here demonstrate that the β subunit is encoded by a novel nuclear gene, and bears no similarity to the catalytic subunit in its amino acid sequence. Further, subunit-specific rabbit antisera directed against recombinant proteins recognize specifically and immunoprecipitate the native enzyme (Wang, et al., 1997b). Antibodies raised against native pol γ also recognize the *E. coli* expressed recombinant β subunit (Wang, et al., 1997b). We conclude that the 35 kDa polypeptide is a *bona fide* subunit of the *Drosophila* pol γ.

The nuclear genes encoding both the catalytic and accessory subunits are mapped to the same genomic region on the left arm of chromosome 2, and they are closely linked with a physical distance of 3.8 kb. This close linkage may be of evolutionary significance, given the current hypotheses regarding the bacterial origin of mitochondria as a cellular organelle. The physical location of these two genes may also provide advantages for the regulation of pol γ gene expression. The genomic location, similar gene structure ( *i.e.* TATA-less promoter, presence of consensus transcriptional initiator and alternative poly(A) signal sequences, small introns and short 3'- untranslated region), and the presence of the DRE elements and potential binding sites for transcription factors NRF-1 and E2F in the 5' untranslated regions of both genes, suggest that the two genes share common regulatory mechanisms, and the possibility that they are coordinately regulated. The DRE element is a recognition sequence for an 80 kDa DNA-binding protein, DREF (Hirose, et al., 1993), and is responsible for activating promoters of nuclear replication genes in *Drosophila*, including PCNA (proliferating cell nuclear antigen), cyclin A and DNA polymerase α-encoding genes (Ohno, et al., 1996). Functional NRF-1 sites have been identified in many mammalian nuclear genes that encode mitochondrial proteins. These include cytochrome c, mitochondrial transcription factor A, and at least one subunit of each of the respiratory complexes III, IV, and V (Virbasius, et al., 1993; Virbasius and Scarpulla, 1994). The E2F transcription factor was first identified as a cellular DNA-binding activity responsible for activation of the adenoviral E2A gene promoter. Eight members of E2F family in mammals have been shown to form heterodimeric complexes that are crucial for transcriptional activation of genes regulating S phase entry (C-myc and cyclin E) and genes functioning to engage in DNA synthesis (dihydrofolate reductase, thymidine kinase, PCNA and DNA pol α) (Zwicker and Muller, 1997). Activity of E2Fs as a transcription activator

or repressor is regulated by their interaction with other cell cycle controlling proteins including Retinoblastoma (Rb) tumor suppressor, other Rb family proteins (p107, p130), and cyclin-dependent kinase. *Drosophila* E2F-1 is essential for activating DNA pol α gene expression, and is required for the G1 to S phase transition during embryogenesis (Duronio, et al., 1995). Further, Sawado *et al* have recently demonstrated that the DREF transcription factor also regulates the expression of the E2F gene in *Drosophila* (Sawado, et al., 1998). The promoter region of the *Drosophila* mtSSB gene also contains DRE elements, and these potential binding sites of transcription factor DREF have been shown to be required for promoter activity in an *in vivo* transfection assay (I. Ruiz de Mena and L.S. Kaguni, unpublished results). That the DRE, NRF-1 and E2F sequence elements are present in the *Drosophila* pol γ genes is consistent with the fact that mitochondrial DNA polymerase is required for mtDNA replication and hence for cell proliferation, and that pol γ activity varies greatly during development, and is highest in embryos (Wernette and Kaguni, 1986).

Molecular cloning, bacterial overexpression and biochemical analysis of the catalytic subunit has allowed the precise assignment of both 5' - 3' DNA polymerase and 3' -5' exonuclease functions to the 125 kDa subunit of *Drosophila* pol γ (Lewis, et al., 1996). Likewise, Foury has shown via its overexpression in mitochondrial extracts, that the 140 kDa polypeptide of the MIP1 gene contains both enzymatic activities in *S. cerevisiae* pol γ (Foury and Vanderstraeten, 1992). Graves *et al.* (Graves, et al., 1998) and (Longley, et al., 1998b) have also shown that the recombinant 140 kDa subunit of human pol γ exhibits both activities when overexpressed in baculovirus infected insect cells. Cloning and sequence alignments have demonstrated that a large polypeptide of 115-140 kDa contains

both the conserved DNA polymerase and exonuclease domains in pol γ of several yeasts, *Xenopus*, and mouse (Lecrenier, et al., 1997). However, pol γ purified from *Xenopus* embryos, pig, and human Hela cells has been shown to contain a large catalytic subunit co-purified with several smaller polypeptides, leaving it difficult to determine the native structure of these enzymes. The identification of human and mouse homologs of the β subunit of *Drosophila* pol γ suggests that animal mitochondrial DNA polymerases may share a common subunit structure.

The accessory subunit of *Drosophila* pol γ and its mammalian homologs represent novel proteins of 41 and 43 kDa, respectively. They both contain two conserved leucine zipper motif sequences. The catalytic subunits of *Drosophila* , *Xenopus*, human and mouse pol γ also contain leucine zipper motifs located between the conserved DNA polymerase and exonuclease domains (Ye, et al., 1996). One class of zinc finger domain has been shown to participate in protein-protein interactions. It is also possible that the putative zinc finger domain in the β subunit may play significant role in its association with the catalytic subunit. We suggest that the putative leucine zipper domains in both subunits participate in subunit interactions because of the genetic, biochemical and structural evidence for an interaction function in other leucine zipper-containing proteins. If so, we would also propose that both the heterodimeric composition and the structural basis for subunit interaction are conserved among animal pol γ holoenzymes.

Data base searching against the complete yeast genome did not identify any open reading frame with significant homology to the β subunit of *D. melanogaster* pol γ. However, this does not exclude the existence of a yeast homolog of the β subunit. It is

possible that the product of an analogous gene in yeast serves the biochemical function of the β subunit. The yeast MIP1 gene encodes a polypeptide of 140 kDa, but it has 200 amino acid residues at its C-terminus that are absent in all other pol γ catalytic subunits. Further, the less highly conserved region between the DNA polymerase and exonuclease domains is substantially shorter. If the C-terminal domain of yeast enzyme serves the same function as the β subunit of *Drosophila* enzyme, then yeast may not require a β subunit.

The molecular cloning of the β subunit of *D. melanogaster* pol γ and its mammalian homologs does not suggest a biochemical function for the accessory subunit because it represents a novel gene. However, other DNA polymerases also contain additional subunits to increase the processivity of the enzyme, or to mediate protein-protein interactions between the polymerase and other replication proteins, thereby enhancing the overall efficiency and/or accuracy of DNA synthesis. For example, T7 DNA polymerase recruits *E. coli* thioredoxin as a processivity factor. The herpes simplex virus UL42 protein functions as a processivity factor for the viral-encoded herpes DNA polymerase. Human pol δ contains a 50 kDa subunit which is responsible for interaction with PCNA to increase processivity. The newest member of yeast DNA polymerase (pol ζ) is a two subunit enzyme. Gene Rev3 encodes the catalytic subunit while gene Rev7 encodes a smaller subunit. The association of Rev7 protein with Rev3 protein increases the overall enzymatic activity of pol ζ (Nelson, et al., 1996). The protein-protein interaction domain of T7 DNA polymerase recently has been identified between the exonuclease domain and the DNA polymerase domain. We are currently investigating the function(s) of the β subunit of *D. melanogaster* pol γ. We propose that the β subunit of *D. melanogaster* pol

γ could serve a similar function(s) as the smaller subunits in other DNA polymerases discussed above.

# FUTURE RESEARCH

The availability of both cDNA and genomic sequences for both subunits provides us opportunities to explore not only the functions of the individual subunits but also the regulation of pol γ gene expression. Studies of the recombinant forms of pol γ holoenzyme with either wild type sequence or with mutants containing site-specific mutations in specific regions will define the roles played by conserved sequences. The biochemical functions of the accessory subunit can also be elucidated by comparing the biochemical properties of the holoenzyme with those of the catalytic subunit. Other possible roles of pol γ, such as in mtDNA repair can also be tested with model DNA substrates using purified recombinant pol γ. Overproduction of recombinant pol γ will allow us to apply protein affinity chromatography to identify other mitochondrial replication factors in addition to allowing further structure-function experiments.

The functional significance of potential transcription factor binding sites located in the promoters of both genes can be studied by both *in vivo* and *in vitro* approaches. The *Drosophila* DRE-binding factor (DREF), and E2F/DP transcription factor have been identified and cloned. Gel mobility shift assay and DNase I footprinting using recombinant DREF and E2F/DP transcription factors will determine whether these transcription factors recognize the DRE element sequences and E2F binding sites in the promoter region of the pol γ genes. We can also design *in vivo* experiments to test whether promoter sequences containing these transcription factor binding sites will activate reporter gene expression in cultured *Drosophila* cells.

We will also take advantage of the excellent genetic methods available in *Drosophila* to study the tissue and developmental specific effects of mutations in pol γ genes by

creating transgenic flies overexpressing additional copies of pol γ gene(s).

In summary, identification of both cDNA and genomic clones of the pol γ genes will extend greatly our ability to perform biochemical and genetic studies to further define the critical functions of pol γ.

# CHAPTER IV

# PRODUCTION AND BIOCHEMICAL CHARACTERIZATION OF RECOMBINANT *DROSOPHILA* POL γ

# INTRODUCTION

Recent progress in genetic, biochemical and structural studies of DNA-dependent DNA polymerases has expanded our understanding of their structure and mechanism. At least six distinct DNA polymerases have been identified in eukaryotic cells and they are all encoded by nuclear genes (Wang, 1991). A human gene that is homologous to the *Drosophila* mus 308 gene and potentially encodes a seventh DNA pol (pol η) (Harris, et al., 1996), has been recently cloned (GenBank accession number AF052573, Sharief,F.S., Ropp,P.A., and Copeland,W.C.). All DNA polymerases perform the same enzymatic reaction of adding a single nucleotide to the 3'-end of the primer. Eukaryotic cells have evolved a whole set of DNA polymerases to carry out the same DNA synthetic reaction on different substrates in replication, recombination, or repair. However, pol γ is the only DNA pol involved in the replication of animal mitochondrial DNA, and it has been suggested to also carry out base excision DNA repair because it contains a 5'-deoxyribose phosphate lyase activity in the catalytic subunit and participates in DNA repair at abasic sites *in vitro* (Pinz and Bogenhagen, 1998; Longley, et al., 1998a). DNA pols α, β, δ, ε, and ζ have been demonstrated to be nuclear polymerases and involved either in nuclear genome DNA replication or DNA repair.

Each eukaryotic DNA polymerase has evolved structural features for specific functions in diverse DNA synthesis pathways. Pol α is exclusively required for the initiation of nuclear chromosomal DNA replication and the priming of Okazaki fragments during lagging strand synthesis, so it also possesses a primase activity. Pol α purified

from different sources consists of four subunits in every organism examined. The role of pol β in base-excision repair is well established in mammalian systems and it is conserved as a single polypeptide enzyme from yeast to man. However, there is a discrepancy in the subunit structure of pol δ and ε, despite accumulating evidence that they likely perform the same function in various systems. Pol δ purified from mammalian cells is a heterodimer, but budding yeast *Saccharomyces cerevisiae* pol δ may contain a third subunit (Burgers and Gerik, 1998). Pol δ purified from the fission yeast *Schizosaccharomyces pombe* contains five subunits (Zuo, et al., 1997). Pol ε derived from human cells has been shown to contain only two subunits of 255 and 55 kDa, while budding yeast pol ε contains four polypeptides. It is also not known whether the subunit structure of pol γ is conserved. Pol γ from *Drosophila* is a heterodimer of 125 and 35 kDa subunits. In budding yeast, the MIP1 gene has been shown to encode the catalytic subunit of pol γ and no homolog of *Drosophila* β subunit has been identified. Pol γ catalytic subunit from human and other mammalian sources also co-purifies with smaller polypeptides. Whether one of these small polypeptides with apparent molecular mass of 40 to 50 kDa is the gene product of *Drosophila* β-subunit homolog in mammals remains to be determined. If the *Drosophila* β-subunit homolog in mammals does encode an accessory subunit of pol γ, what specific function does it possess and is its function conserved?

Animal mtDNA replication is not coupled to cell cycle, yet, the activity of pol γ and

the expression of the pol γ genes are clearly developmentally regulated (Wernette and Kaguni, 1986; Ye, et al., 1996). The expression of the pol γ genes, mitochondrial targeting of precursors of both subunits, and holoenzyme assembly are important steps to produce the functional pol γ and very likely regulated in response to the cellular requirement for mitochondrial function. The accuracy and efficiency of each step certainly affects the enzymatic activity of pol γ, and thus has impact on genomic stability of mtDNA.

DNA pol γ represents less than 1% of the total DNA polymerase activity in cells, and the low abundance of pol γ makes it very difficult to obtain sufficient native enzyme to conduct structure-function studies. The baculovirus expression system provides a eukaryotic environment for recombinant protein production. The advantageous features of this system include high-level expression, capacity for post-translational modification, simultaneous expression of multiple genes, and subcellular compartmentalization of recombinant proteins. Many examples of successful production of active enzymes have been reported, including the expression of human pol α (Copeland and Wang, 1991), replication factor C complex , and recently the catalytic subunit of human pol γ (Longley, et al., 1998b). However, few mitochondrial proteins have been expressed using this system, and no multi-subunit mitochondrial enzyme has been expressed and purified.

We have achieved the baculovirus overexpression, purification and biochemical characterization of the recombinant holoenzyme of *Drosophila* pol γ from insect cells. We also examined holoenzyme reconstitution in both the mitochondria and cytoplasm of infected Sf-9 insect cells.

# EXPERIMENTAL PROCEDURES

## Materials

*Enzymes and Proteins* - native *Drosophila* pol γ Fraction VI was prepared as described by Wernette and Kaguni (Wernette and Kaguni, 1986). Polyclonal antibodies raised against bacterial produced recombinant α or β subunit were prepared and purified by the method of Olson et al (Olson, et al., 1995).

*Nucleotides and Nucleic Acids* - Baculovirus transfer vector pVL1392 /1393, and linearized wild type baculovirus AcMNPV DNA (BaculoGold) were purchased from PharMingen. Wild type baculovirus AcMNPV was generous gift from of Dr. Suzanne Thiem, Department of Entomology, Michigan State University. Synthetic oligodeoxynucleotides as indicated below were synthesized in an Applied Biosystems Model 477 oligonucleotide synthesizer.

*Insect Cells and Tissue Culture Medium* - Sf-9 ( *Spodoptera frugiperda* ) cells were kind gift of Dr. Suzanne Thiem, Department of Entomology, Michigan State University. TC-100 insect cell culture medium and Fetal bovine serum were from GIBCO-BRL. Insect cell transfection buffer and Grace's medium were from PharMingen.

*Chemicals* - SeaKem ME low melting agarose was purchased from FMC. Amphotericine, Penicillin-G, Streptomycin, tryptose broth were from Sigma. Protease inhibitors PMSF, sodium metabisulfite, and leupeptin were purchased from Sigma, J. T. Baker Chemical Co., and Peptide Institute, Minoh-Shi, Japan respectively.

## Methods

*Construction of Recombinant Baculoviruses* - Transfer plasmid vectors containing either the complete coding sequence or coding sequence with N-terminal deletion of the mitochondrial import presequence of both subunit of pol γ were prepared by standard DNA manipulation. An Nde I fragment containing the complete coding sequence of the α subunit with or without N-terminal deletion was released from *E. coli* expression pET plasmid vector DNA (Lewis, et al., 1996) and subsequently was subcloned into the EcoR I site of the transfer vector pVL1393 by blunt end ligation after both the vector and insert DNA were repaired with Klenow to generate pVL93α and pVL93α-NL. A restriction fragment of Xba I and BamH I from *E. coli* expression pET plasmid vector containing either the complete coding sequences or coding sequence with N-terminal deletion of mitochondrial import presequence of the β subunit was directionally cloned into the Xba I and BamH I sites of the transfer vector pVL1392 to obtain pVL92β and pVL92β-NL. The transfer plasmid containing an N-terminal His-tag and deletion of mitochondrial presequence of the α subunit (pVL93α-NLNHis) was obtained by replacing the Xba I fragment of pVL93α (representing the N-terminal portion of the coding sequence) with an Xba I fragment from an *E. coli* expression vector (pET28a-αNL). The transfer plasmid containing C-terminal His-tagged α subunit (pVL93α-CHis) was prepared by PCR-mediated site-directed mutagenesis to insert the His-tag coding sequence to the C-terminus. Partial DNA sequence of the plasmid constructs were obtained to confirm the correct

Fig. 1 **Baculovirus transfer vectors.** The complete coding sequences of α and β subunits of *Drosophila* pol γ were subcloned from their respective bacterial expression plasmids into the transfer vector pVL1392/1393 to obtain pVL93α and pVL92β. The mitochondrial presequences of both subunits were deleted from their N-termini and a translation initiation codon was added to obtain pVL93α-NL and pVL92β-NL respectively. A $(His)_6$ coding sequence was inserted at the N-terminus of pVL93α-NL sequence or C-terminus of pVL93α sequence to obtain pVL93α-NLNHis and pVL93α-CHis.

Baculovirus transfer vectors

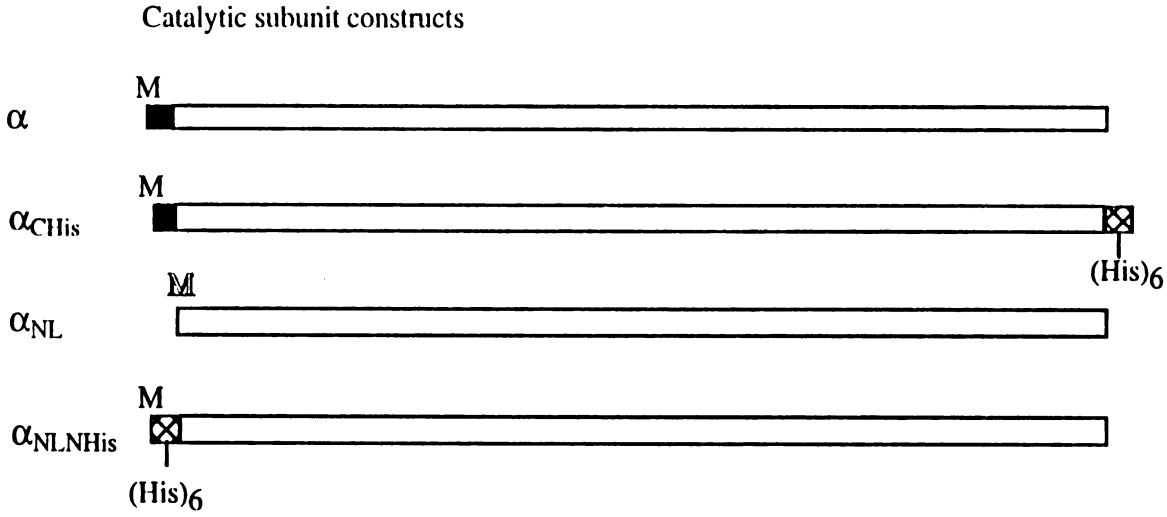



FIGURE 1

orientation of the insertion and the preservation of correct reading frame and translational initiation and stop codons.

Linearized wild type baculovirus DNA (BaculoGold, 0.5 μg, from PharMingen) and purified transfer plasmid DNA containing different constructs encoding α or β subunit (2 μg) were co-transfected in transfection buffer ( 25mM Hepes pH 7.1, 125mM $CaCl_2$, 140mM NaCl ) at 27 $^0$C for 4 hours following manufacture's recommendation. The putative recombinant viruses were plaque purified. The resulting recombinant viruses were designated as AcNPVα, AcNPVβ, AcNPVα-NL, AcNPVβ-NL, AcNPVα-CHis, AcNPVα-NLNHis for viruses encoding full length α subunit (p125), full length β subunit (p35), α subunit without mitochondrial import presequence (p125NL), β subunit without mitochondrial import presequence (p125NL), α subunit with C-terminal His tag (p125CHis), α subunit with N-terminal His tag and presequence deletion respectively (p125NLNHis, see Figure 1). All viruses were amplified in Sf-9 cells to titers of $5x10^7$ to $1x10^8$ plaque forming units/ml.

*Cell Culture and Production of Recombinant α and β subunit-* Sf-9 (*Spodoptera frugiperda* ) cells are grown in TC-100 insect cell culture medium containing 10% fetal bovine serum at 27$^0$C. Sf-9 cells were infected with recombinant virus with the multiplicity of infection (MOI) of 5 in TC-100 medium supplemented with 10% fetal bovine serum. Cells ( about 1 $X10^6$ cells/ml ) were collected by centrifugation at low speed (1x1000g) and

lysed in 1X SDS Laemmli gel loading buffer or washed with PBS buffer and frozen in liquid nitrogen and stored at -80$^0$ C. The overproduction of recombinant form of α and β subunit was analyzed by immunoblotting with subunit-specific antibodies.

*Mitochondrial Extraction* - Cells collected from 50 ml culture were homogenized in HB buffer ( 15 mM HEPES pH 8.0, 2 mM $CaCl_2$, 5 mM KCl, 0.5 mM EDTA, 280 mM Sucrose, 1 mM PMSF, 0.5 mM DTT, 10 mM Na bisulfite, 2 μg/ml leupeptin) or EB buffer (50mM Tris-HCl pH 7.5, 50 mM KCl, 280 mM Sucrose, 1 mM PMSF, 5 mM β-mercaptoethanol, 10 mM Na bisulfite, 2 μg/ml leupeptin). Mitochondria were pelleted at 7,800xg after removing nuclei and unbroken cells. Mitochondrial fraction was extracted with 2% Na cholate in extraction buffer ( 25 mM HEPES pH 8.0, 10% glycerol, 2 mM EDTA, 300 mM NaCl, 0.5 mM DTT, 1 mM PMSF, 10 mM Na bisulfite, and 2μg/ml leupeptin ). DTT was replaced by 5 mM β-mercaptoethanol and EDTA was omitted in Ni-NTA agarose affinity purification protocol used for His-tagged protein. Both cytoplasmic and mitochondrial fractions were centrifuged at 100,000 xg to remove insoluble materials.

*Purification of Recombinant Drosophila pol γfrom the cytoplasm of Sf-9 cells* - All purification operations were all performed at 0-4 $^o$C. Sf-9 cells infected with recombinant baculovirus encoding the catalytic and accessory subunit were harvested 48 hours post infection. The cells were pelleted and washed with equal volume of cold PBS buffer. The cell pellet ($1x10^9$ cells) was resuspended in 20 ml of extraction buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 280 mM Sucrose, 5 mM EDTA, 2 mM DTT, 10 mM Na metabisulfite, 1 mM PMSF, and 2μg/ml leupeptin) and lysed by 20 strokes in a Dounce

homogenizer. The homogenate was centrifuged at 3000 rpm, 3 °C, in a SS-34 rotor for 7 min. The pellet was resuspended in 14 ml of extraction buffer and re-homogenized and centrifuged as above. The combined supernatant (Fraction Ia1, 34 ml) was centrifuged at 10,000 rpm, 3 °C, for 15 min to pellet the mitochondria. The supernatant (Fraction Ia2) was centrifuged in a Beckman 60Ti rotor at 100,000 X g to obtain cytoplasmic soluble fraction (Fraction Ia3, 2.1 mg/ml of protein) and the mitochondria pellet was washed twice with 50 ml of extraction buffer and frozen in liquid nitrogen and stored at -80 °C.

Purification of recombinant pol γ holoenzyme was conducted as described (Wernette and Kaguni, 1986) with the following modifications. Soluble (S100) cytoplasmic fraction (Fraction Ia3) was adjusted to an ionic equivalent of 100 mM potassium phosphate and loaded onto a phosphocellulose column (15 ml, 6 mg of protein per packed ml of phosphocellulose resin) equilibrated with 80 mM potassium phosphate buffer at a flow rate of 12 ml/h. The column was washed with 45 ml of 100 mM potassium phosphate buffer at a flow rate of 30 ml/h. Proteins were eluted with 44-ml linear gradient from 150 to 350 mM potassium phosphate buffer at a flow rate of 30 ml/h and active fractions (#17 - 28) were pooled (Fraction II, 20 ml) and adjusted with 80% sucrose to a final concentration of 10%. After addition of 1.2 volumes of saturated ammonium sulfate, pH 7.5, and incubation on ice for 2 h, the precipitate was collected by centrifugation at 96,000 X g for 30 min at 3 °C. The protein was resuspended in 2.0 ml of 10 mM K phosphate buffer containing 45% glycerol, and stored at -20 °C (Fraction IIb). To achieve better precipitation, BSA (0.5 mg/ml) was added to the supernatant and the above procedure was repeated to obtain 4 ml of Fraction IIb (0.46 mg/ml). The Fraction IIb was dialyzed in 250 ml dialysis buffer (10 mM K phosphate, pH 7.6, 20% glycerol, 2

mM EDTA) in a collodion bag (Mr cutoff, 25,000 kDa) until an ionic equivalent of 85 mM KCl was reached. It was then loaded onto a single-stranded DNA-cellulose column (1.8 ml, 1mg of protein per packed ml of resin) equilibrated with 20 mM K phosphate buffer at a flow rate of 1.3 ml/h. The column was washed with 4 ml 20 mM K phosphate buffer containing 100 mM KCl at 2.7 ml/h followed by successive elution with 20 mM K phosphate buffer containing 250 mM KCl (8 ml at 4ml/h), 600 mM KCl (6 ml at 4ml/h), and 1 M KCl (4 ml at 5ml/h). Active fractions were pooled as Fraction III (# 17 - 27, 5 ml). Solid ammonium sulfate was added to Fraction III from the ssDNA-cellulose column to final concentration of 0.4 mg/ml. After stirring for 20 min on ice, the suspensions were centrifuged at 20,000 rpm, 3 $^{o}$C, for 10 min in a Beckman Ti60 rotor. The supernatant was loaded onto an Octyl-Sepharose column (0.5ml) which was equilibrated with 20 mM K phosphate buffer at a flow rate of 0.5 ml/h. The Octyl-Sepharose column was washed with 2 ml of equilibration buffer at 2 ml/h and then eluted at the same flow rate with 1.75 ml of buffer containing 0.3% Triton X-100, 2 ml of buffer containing 1% Triton X-100, and 1.5 ml of buffer containing 2% Triton X-100. Fraction IV was then loaded on 12-30% glycerol gradient and the peak fractions (Fraction V) containing recombinant pol $\gamma$ were collected, assayed, pooled, stabilized, and stored at -20 $^{o}$C or frozen with liquid nitrogen and stored at -80 $^{o}$C.

*Ni-NTA agarose affinity purification of recombinant pol $\gamma$* - Soluble (S100) cytoplasmic fraction (10 ml) of Sf-9 cells co-infected with viruses was mixed with 50 µl of pre-charged Ni-NTA (nitrilo-tri-acetic acid) agarose resin and incubated on ice with gentle shaking for 4 hours. The resin was washed 2 times with 10 ml of buffer (50 mM Tris-HCl pH 7.5,200 mM KCl, 10% glycerol, 1 mM PMSF, 5 mM β-mercaptoethanol, 10 mM Na

bisulfite, 2 μg/ml leupeptin, and 5 mM imidazole). Proteins retained on the resin were eluted first with 1 ml of washing buffer (with 50 mM imidazole) followed by second elution with 0.3 ml washing buffer (with 300 mM imidazole). All the washing and elution steps were done on ice and for 30 minutes with gentle shaking. Mitochondria from 150 ml of Sf-9 cells infected with recombinant virus expressing α and β subunits were extracted in 4 ml of EB-mito extraction buffer (25mM HEPES pH 8.0, 10% glycerol, 300mM NaCl, 1 mM PMSF, 10 mM Na bisulfite, 2 μg/ml leupeptin). Recombinant pol γ was purified with 100 μl of pre-charged Ni-NTA agarose resin following the same procedure as for the cytoplasmic extract except the washing and elution buffer was changed to EB-mito buffer.

*DNA polymerase assay* - DNA polymerase activities were assayed on DNase I-activated calf thymus DNA or singly-primed M13 DNA at 30 - 200 mM KCl as described by Olson (Olson, et al., 1995). The reaction mixtures (0.05 ml) contained 50 mM Tris-HCl pH 8.5, 4 mM $MgCl_2$, 10 mM DTT, 400 μg/ml bovine serum albumin, and 20 or 30 μM each of dGTP, dATP, dCTP, and [$^3$H]dTTP(1000 cpm/pmol). Specific modifications are described in the figure legends. One unit of DNA polymerase activity is defined as the amount that catalyzes the incorporation of 1 nmol of deoxyribonucleoside triphosphate into acid insoluble material in 60 min at 30 °C using DNase I-activated calf thymus DNA as the substrate.

*3' -5' exonuclease assay* - Reaction mixtures (0.05 ml) containing 50 mM Tris-HCl pH 8.5, 4 mM $MgCl_2$, 10 mM DTT, 400 μg/ml bovine serum albumin, and KCl as indicated, 4 μM 5'-end labeled singly-primed M13 DNA containing a 3'-terminal mispair,

and 0.1 unit of Fraction VI recombinant pol γ were incubated for 30 min at 30 °C in the presence or absence of saturating amount of recombinant mtSSB protein. Samples were then made 1% in SDS and 10 mM in EDTA, heated for 10 min at 65 °C and precipitated with ethanol in the presence of 1 μg sonicated salmon sperm DNA as carrier. The ethanol precipitates were resuspended in 80% formamide and 90 mM Tris-borate. Aliquots were denatured for 2 min at 100 °C, chilled on ice and electrophoresed in an 18% polyacrylamide gel (13x13x0.075 cm) containing 7 M urea in 90 mM Tris-borate pH 8.3 and 2 mM EDTA. After electrophoresis, the gel was washed in 15% glycerol for 20 min and exposed at - 80 °C to Kodak X-Omat AR x-ray film using a DuPont NEN Quanta III intensifying screen.

*Other Methods* - Protein concentration was determined by the method of Bradford with Bovine serum albumin as the protein standard. Gel electrophoresis and protein transfer and immunoblotting were performed as described by D. L. Lewis et al. (Lewis, et al., 1996).

# RESULTS

## Construction of recombinant baculoviruses and overexpression of recombinant subunits of *Drosophila* pol γ in cultured insect cells

Co-transfection of linear baculovirus DNA and transfer vector containing α-subunit or β-subunit coding sequence with or without their mitochondrial presequence (leaders) generated recombinant baculoviruses very efficiently (Figure 1). All recombinant viruses were purified by isolating a single viral plaque and amplified in Sf-9 cells to high titer ( 1X $10^8$ pfu/ml). For each recombinant virus, ten to twelve independently isolated clones were amplified and tested for expression of recombinant protein. All of these produced recombinant α or β subunit polypeptides at similar levels (data not shown). Analysis of whole cell extracts of infected Sf9 cells showed that the expression levels of recombinant α- or β-subunit polypeptides are similar to that of viral polyhedrin protein. Figure 2 shows an SDS-polyacrylamide gel analysis of whole cell extracts from Sf-9 cells, cells infected with wild-type baculovirus, and cells infected with recombinant baculoviruses encoding α or β subunits in different forms. Cells infected with AcNPVα, AcNPVα-NL, AcNPVα-CHis, AcNPVα-NLNHis all produced a protein of 125 kDa that was immunoblotted by α subunit-specific antibody. Similarly, cells infected with AcNPVβ or AcNPVβ-NL produced a protein of 35 kDa that reacted with β subunit-specific antibody. Extracts from uninfected Sf-9 cells and cells infected with wild type virus did not contain any protein that cross-reacted with either α or β subunit-specific antibody. Production of total recombinant

Fig. 2 **Overexpression of recombinant α and β subunits in Sf-9 cells.** Sf-9 cells were infected with recombinant baculovirus expressing α or β subunit as full-length protein or with modifications as indicated at the MOI of 5, and harvested 48 hours after infection and total cell extract from $1X10^5$ cells (for immunoblotting in panel B) or $2X10^5$ cells (for staining in panel A) was resolved on 10% SDS-PAGE gels. *Lane 1* represents a control. *Lane 2* and *3* represent cell extracts from uninfected cells (C) and cells infected with wild type baculovirus (W). *Lanes 4, 5, 6,* and*7* represent cells expressing α subunit as full length (α), C-terminal His-tag fusion (αCHis), N-terminal presequence deletion (αNL), and N-terminal presequence deletion with His-tag fusion (αNHis). *Lanes 8* and*9* represent cells expressing β subunit as full length (β) and N-terminal presequence deletion (βNL).

Overexpression of recombinant α and β subunits in Sf-9 cells.

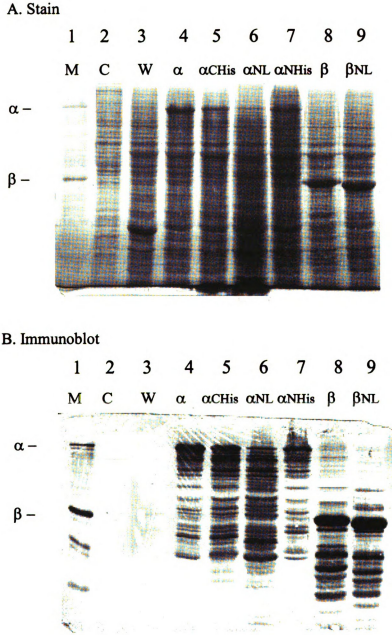



FIGURE 2

protein was calculated to be about 2 μg/ml of cultured cells. Protease degradation of recombinant proteins was observed, and subcellular fractionation experiments indicated that only about 10 to 20 % of recombinant protein was recovered in the soluble fraction. Overexpression in other insect cell lines (Sf21 or High Five) did not improve the solubility or eliminate the proteolysis of recombinant proteins (data not shown). Experiments performed by varying MOI (multiplicity of infection) or adding the protease inhibitor phenylmethylsulfonyl fluoride (PMSF) at 1mM into the culture medium also failed to prevent proteolysis (data not shown).

**Mitochondrial import of recombinant α and β subunits**

Mitochondria were prepared as indicated in *Method*s and the crude soluble S-100 extracts were analyzed by immunoblotting. Recombinant α and β subunit are clearly detected in the mitochondrial fraction (Figure 3). Full-length recombinant α or β subunit are competent for mitochondrial import; however, the targeting and import efficiency appears to be suboptimal. Insect cell mitochondria isolated 48 hours post infection contain mostly full length recombinant proteins with minimal levels of degradation (Figure 3), indicating that the mature recombinant α and β polypeptides are protected from protease(s) once they are imported into the mitochondrial matrix. The mitochondrial presequence is necessary for proper mitochondrial import. No recombinant protein is detected in mitochondrial extracts when they are expressed without a mitochondrial presequence at their N-termini (p125NL or p35NL); instead, the recombinant α and β polypeptides remained in the insect cell cytoplasm (Figure 3). The estimated yield of recombinant α and

Fig.3 **Mitochondrial import of recombinant α and β subunits.** Sf-9 cells were infected with recombinant baculovirus expressing α and/or β subunits as full-length proteins (α or β) or lacking presequences (αNL or βNL) at the MOI of 5, and harvested 48 hours after infection. Soluble mitochondrial (Mi) and cytoplasmic (Cy) fractions were extracted from $3X10^7$ cells and aliquots of these fractions were resolved on 10% SDS-PAGE gels. The recombinant polypeptides in each fraction were detected by immunoblotting using subunit-specific antibodies. Lane *1* and *2* represent samples from co-expression of full-length α and β subunits. Lane *3* and *4* represent samples from co-expression of α and β subunits lacking presequences. Lane *5*, *6*, *7*, and *8* represent samples from α and β subunits expressed individually and lacking presequences.

Mitochondrial import of recombinant α and β subunits

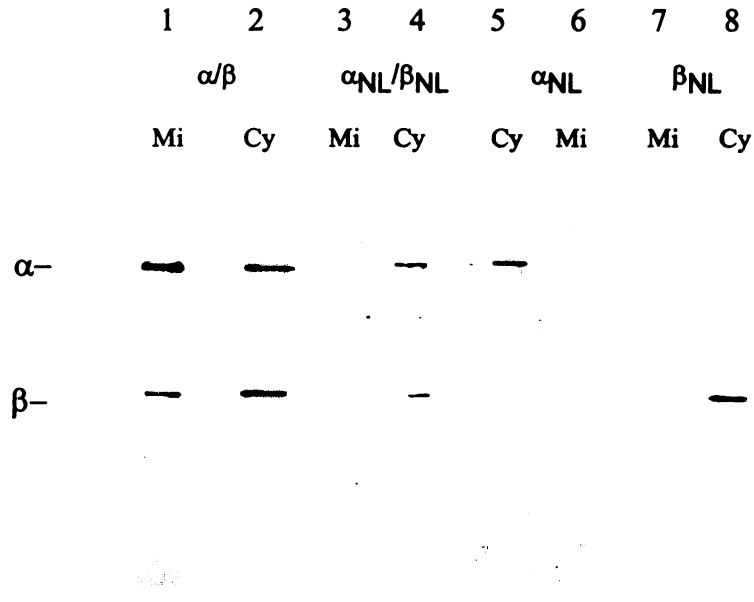



FIGURE 3

β polypeptides in the mitochondrial fraction is only about 40 ng per ml of cells collected 48 hours post infection, representing ~ 2% of total recombinant protein.

**Co-expression of recombinant α and β subunits reconstitutes pol γ holoenzyme** Whether the recombinant α and β subunits are properly folded and recombinant heterodimeric pol γ holoenzyme is formed in the mitochondria or cytoplasm was determined by glycerol gradient sedimentation. The crude mitochondrial soluble fraction derived from cells infected with both recombinant viruses AcNPVα and AcNPVβ was sedimented in a 12-30% glycerol gradient. The recombinant α and β subunits were found to co-sediment (Figure 4). Native pol γ purified from *Drosophila* embryos sedimented at the same glycerol density. The identity of recombinant β subunit in the peak fractions was confirmed by immunoblotting using β subunit-specific antibody (data not shown). Identical results were obtained whether the cells were harvested 36, 42, or 48 hours after co-infection (data not shown). Remarkably, recombinant full length α and β subunits are also assembled into heterodimer in the cytoplasm of Sf-9 cells, and the addition of a $(His)_6$ tag at the C-terminus of the α subunit does not disrupt the protein-protein interaction (data not shown). Recombinant catalytic subunit expressed individually sedimented at lower glycerol density. However, recombinant β subunit expressed individually was largely insoluble and failed to sediment as a discrete species (data not shown).

Fig. 4 **Reconstitution of recombinant *Drosophila* pol γ heterodimer in mitochondria of Sf-9 cells.** Soluble mitochondrial extract from Sf-9 cells infected with recombinant baculoviruses expressing α and β subunits was sedimented in a 12-30% glycerol gradient. Proteins were resolved on 10% SDS-PAGE gels and detected by immunoblotting using subunit-specific antibodies. Numbers labeling each lane represent the fraction numbers.

Reconstitution of a recombinant *Drosophila* pol γ heterodimer in mitochondria of Sf-9 cells

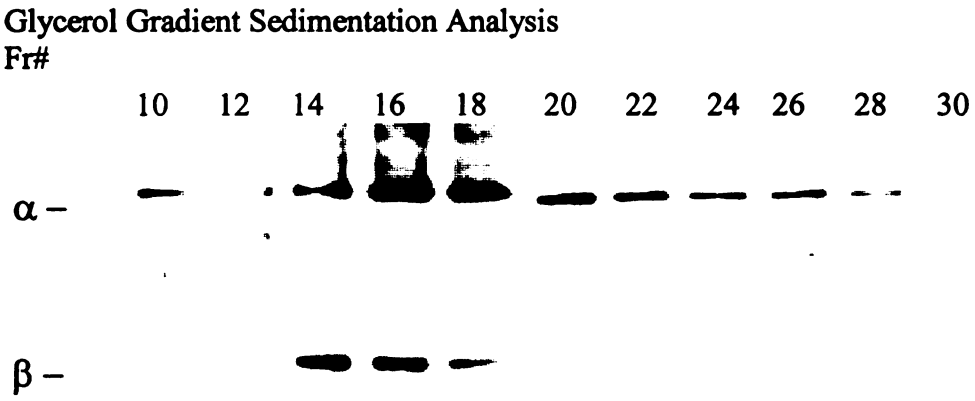



FIGURE 4

Interestingly, co-expression of recombinant α and β subunits as mature forms lacking their mitochondrial presequences (p125NL/p35NL) does not yield a pol γ heterodimer in the soluble cytoplasmic fraction (Figure 5). In fact, removal of the mitochondrial presequence of the α subunit alone (p125NL) in co-expression experiments is sufficient to disrupt holoenzyme assembly in the cytoplasm. Soluble α subunit without the leader sequence (p125NL) can be purified by chromatography on a phosphocellulose column (Figure 5), although the β subunit alone can not be recovered.

**Purification of recombinant *Drosophila* pol γ**

Near homogeneous recombinant *Drosophila* pol γ holoenzyme was purified directly from the soluble cytoplasmic extract of Sf-9 cells infected with baculoviruses encoding the complete coding sequences of the α and β subunits by sequential chromatography on phosphocellulose, ssDNA-cellulose, and octyl-Sepharose followed by glycerol gradient sedimentation as described in *Methods*. Table 1 and Figure 6 summarize the results of one purification. The two subunits of pol γ remained tightly associated in the same apparent stoichiometry throughout the purification. The purified recombinant holoenzyme is an active DNA polymerase with proofreading 3' - 5' exonuclease activity. Furthermore, the specific activity of the recombinant enzyme is 20,000 units/mg, which is similar to that of native pol γ. The heterodimeric form of recombinant pol γ with the catalytic subunit expressed as a C-terminal his-tag fusion protein can also be purified as an active enzyme using Ni-NTA agarose affinity chromatography (data not shown), although

Fig. 5 **Co-expression of α and β subunits lacking mitochondrial presequences.** Soluble cytoplasmic fraction from $3X10^7$ Sf-9 cells infected with recombinant baculoviruses expressing α and β subunits was chromatographed on a phosphocellulose column. The fractions collected from the column were resolved on 10% SDS-PAGE gels and the α and β subunits were detected by immunoblotting. *Mi* and *Cy* represent mitochondrial and cytoplasmic soluble fractions. *F* and *W* represent flowthough and wash fractions. Numbers above each lane represent the eluted fraction number. *In* and *M* represent the cytoplasmic insoluble fraction and a molecular weight control respectively.

Co-expression of α and β subunits lacking mitochondrial presequnces

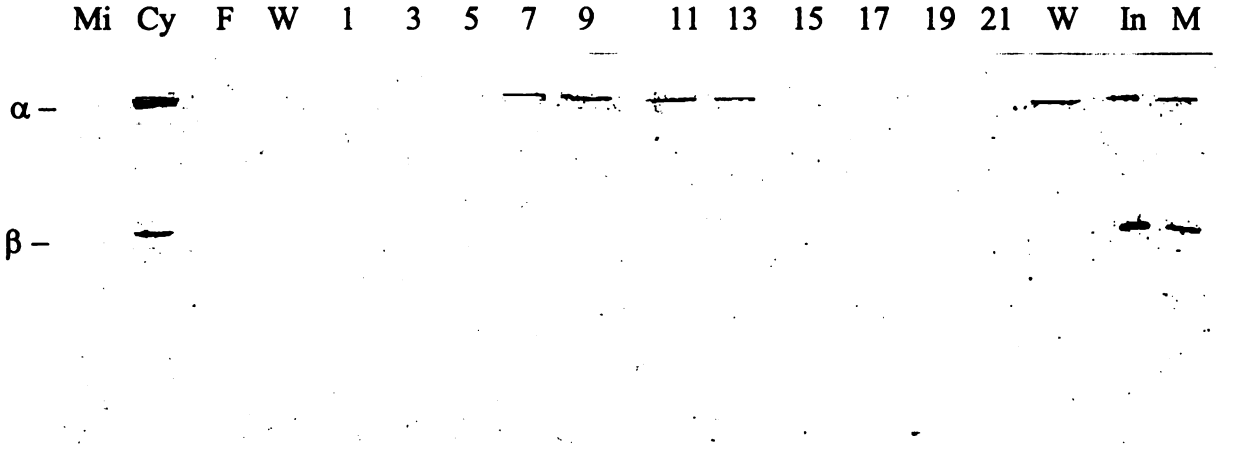



FIGURE 5

Table I Purification of recombinant *Drosophila* pol γ from baculovirus-infected Sf-9 cells

| Fraction | Volume (ml) | Protein (mg) | Activity (units) | Specific Activity (units/mg) | Yield (%) |
|---|---|---|---|---|---|
| I. Soluble cyto. extract[a] | 34 | 71.4 | 8092 | 113 | 100 |
| II. Phosphocellulose and ammonium sulfate | 4.5 | 1.84 | 974 | 529 | 12 |
| III. ssDNA-cellulose | 11.2 | 0.69 | 748 | 1084 | 9 |
| IV. Octyl-sepharose | 3.0 | ND | 159 | ND | 2 |
| V. Glycerol gradient | 9.0 | 0.0075 | 149 | 20,000 | 2 |

[a]Fraction I was prepared from 0.5 L of infected Sf-9 cells

Fig. 6 **Purification of recombinant *Drosophila* pol γ.** Fraction VI of native (*lane 1*, 10 units) and fraction V of recombinant (*lane 2*, 9 units ) pol γ were electrophoresed in a 10% SDS-PAGE gel and the proteins were stained with silver. Fraction V of recombinant pol γ (*lane 3*, 1.5 units) was also probed with subunit-specific antibodies.

Purification of recombinant *Drosophila* pol γ

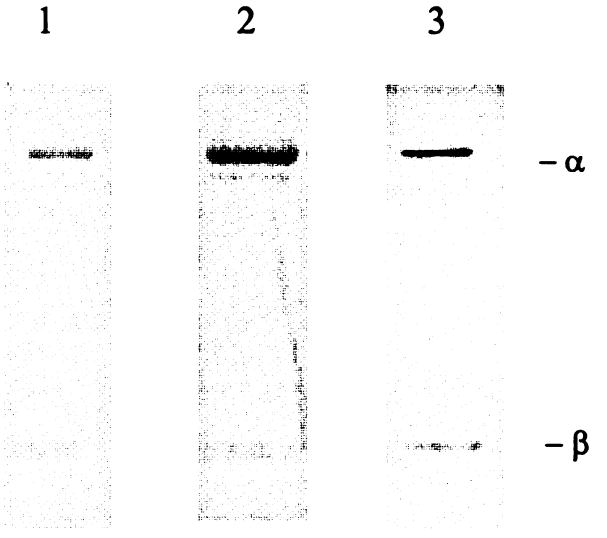



FIGURE 6

the specific activity is ~10-fold lower.

**Stimulation of both DNA polymerase and 3'-5' exonuclease activities of recombinant *Drosophila* pol γ by salt and by mtSSB**

The recombinant pol γ holoenzyme exhibits very similar biochemical properties to that of native pol γ purified from *Drosophila* embryos. Its DNA polymerase activity is stimulated 8 fold by elevated salt (200 mM KCl) on DNase I-activated calf thymus DNA and 30 fold by mtSSB on singly-primed M13 DNA. Its mispair-specific 3' - 5' exonuclease activity is also stimulated by salt and mtSSB (data not shown).

# DISCUSSION

An efficient and reliable system for overexpression of recombinant proteins, particularly those proteins with very low abundance, will provide the foundation for further structural and functional studies. We have pursued overexpression of recombinant *Drosophila* pol γ since we isolated the cDNAs for both subunits. The baculovirus expression system can produce both subunits of pol γ in soluble form and yield an active recombinant heterodimeric enzyme. (At the same time, we failed to achieve *in vivo* reconstitution of heterodimeric holoenzyme by bacterial expression (Wang, et al., 1997b). The recombinant subunits of pol γ are targeted and imported into mitochondria of Sf-9 cells whether they are expressed individually or simultaneously. The successful reconstitution of holoenzyme in the mitochondria of Sf-9 cells suggests that the cellular machinery for mitochondrial localization and processing is conserved between *Drosophila* and *Spodoptera frugiperda.* The low yield of recombinant pol γ in Sf-9 cell mitochondria is probably the result of the lytic nature of baculovirus infection. It is recognized that host protein synthesis is greatly reduced 24 hours after infection when viral late genes dominate mRNA transcription. The fact that the recombinant holoenzyme can be purified from the soluble cytoplasmic fraction and exhibits nearly identical chromatographic behavior and biochemical and physical properties suggests strongly that the recombinant holoenzyme is able to achieve proper folding in the cultured insect cells, and demonstrates further that the β subunit is an essential component of pol γ. The successful *in vivo* reconstitution of holoenzyme provides us with a useful system to study mutant forms of holoenzyme and to examine subunit interactions in pol γ.

We have constructed recombinant baculoviruses containing the corresponding coding sequences of the mature form of both subunits in an effort to produce and purify the recombinant enzyme without the N-terminal mitochondrial presequences, to compare the recombinant and native pol γ, because the latter purified from embryonic mitochondria lacks the presequences. However, when both subunits were co-expressed without their presequences, the solubility of both recombinant proteins and holoenzyme reconstitution was greatly reduced. The deletion of the presequence in the catalytic subunit alone is sufficient to disrupt the proper association of two subunits in the cytoplasm. The exact role of the presequence in promoting holoenzyme assembly is not clear. However, it is possible that the targeting of nascent recombinant protein to the mitochondria itself brings the two subunits in close proximity, thus aiding heterodimer formation. The low efficiency of mitochondrial import of recombinant pol γ and the discovery that the heterodimeric form of recombinant pol γ is reconstituted in the cytoplasm of insect cells prompted us to purify recombinant pol γ directly from the soluble cytoplasmic fraction. The recombinant pol γ purified in this study is thus slightly different structurally from native pol γ, as it contains the mitochondrial presequences in both subunits. The presequence in the catalytic subunit represents only 9 amino acid residues, and the predicted presequence in the accessory subunit is 11 amino acid residues. One modified human recombinant catalytic subunit expressed and purified from insect cells is an N-terminal his-tag fusion protein; there the addition of the his-tag does not change its enzymatic properties (Longley, et al., 1998b). We did not detect a significant alternation of the catalytic properties of our recombinant pol γ as compared with the native enzyme. In fact, the addition of a his-tag at the C-terminus of the catalytic subunit does not disrupt the subunit association. The lower than expected

specific activity of this form of enzyme was likely the result of its purification by Ni-NTA affinity chromatography.

The human pol γ catalytic subunit has been purified as recombinant enzyme from Sf-9 cells and from cultured Hela cells. They exhibit similar enzymatic properties. However, its DNA polymerase activity is salt sensitive (Longley, et al., 1998b) and it is not known whether mtSSB can stimulate its DNA polymerase and 3' - 5' exonuclease activities. Our initial characterization of recombinant *Drosophila* pol γ clearly shows that it exhibits very similar biochemical properties relative to native pol γ in terms of its response to stimulation by both elevated salt and mtSSB. Whether the difference observed in our study and the others reflects the functional significance of the β subunit remains to be determined. However, these results suggest the possibility that physical interaction of two subunits is important in modulation of the catalytic properties of the catalytic subunit.

In summary, we have established an effective eukaryotic expression system to produce recombinant pol γ and study holoenzyme assembly. To our knowledge, this work represents the first example of reconstitution of a multi-subunit mitochondrial enzyme using the baculovirus expression system.

# FUTURE RESEARCH

We are currently performing overexpression and purification of the catalytic subunit alone to compare the biochemical properties of the catalytic subunit versus holoenzyme. Whether the β subunit plays a role in modulating overall activity, and/or fidelity and processivity of DNA synthesis will be tested when purified catalytic subunit becomes available. Mutant holoenzyme lacking 3'-5' exonuclease activity will also be produced and its enzymatic properties will be examined. Structure-function studies with recombinant enzymes will be carried out following the experimental approaches described in Chapter II.

# LIST OF PUBLICATIONS--YUXUN WANG

Olson, M.W., **Wang, Y.**, Elder, R.H., and Kaguni, L.S. (1995) Subunit structure of mitochondrial DNA polymerase from *Drosophila* embryos: Physical and immunological studies. J. Biol. Chem. **270**, 28932-28937.

Lewis, D.L., Farr, C.L., **Wang, Y.**, Lagina, A.T. III, and Kaguni, L.S. (1996) Catalytic subunit of mitochondrial DNA polymerase from *Drosophila* embryos: Cloning, bacterial overexpression and biochemical analysis. J. Biol. Chem. **271**, 23389-23394.

**Wang, Y.**, Farr, C.L., and Kaguni, L.S. (1997) Accessory subunit of mitochondrial DNA polymerase from *Drosophila* embryos: Cloning, molecular analysis and association in the native enzyme. J. Biol. Chem. **272**, 13640-13646.

Farr, C.L., **Wang, Y.**, and Kaguni, L.S. (1999) Functional interactions of mitochondrial DNA polymerase and single-stranded DNA-binding protein: Template-primer DNA binding, and initiation and elongation of DNA strand synthesis. J. Biol. Chem., submitted for publication.

**Wang, Y.**, and Kaguni, L.S. (1999) Baculovirus expression reconstitutes native *Drosophila* mitochondrial DNA polymerase. J. Biol. Chem., manuscript in preparation.

# REFERENCES


Baker, R. P. and Reha-Krantz, L. J. (1998) Identification of a transient excision intermediate at the crossroads between DNA polymerase extension and proofreading pathways. Proc Natl Acad Sci U S A **95:** 3507-12.

Bambara, R. A., Murante, R. S. and Henricksen, L. A. (1997) Enzymes and reactions at the eukaryotic DNA replication fork. J Biol Chem **272:** 4647-50.

Barnes, C. J., Wahl, A. F., Shen, B., Park, M. S. and Bambara, R. A. (1996) Mechanism of tracking and cleavage of adduct-damaged DNA substrates by the mammalian 5'- to 3'-exonuclease/endonuclease RAD2 homologue 1 or flap endonuclease 1. J Biol Chem **271:** 29624-31.

Bedford, E., Tabor, S., and Richardson, C. C. (1997) The thioredoxin binding domain of bacteriophage T7 DNA polymerase confers processivity on *Escherichia coli* DNA polymerase I. Proc Natl Acad Sci USA **94:** 479-484.

Beechem, J. M., Otto, M. R., Bloom, L. B., Eritja, R., Reha-Krantz, L. J. and Goodman, M. F. (1998) Exonuclease-polymerase active site partitioning of primer-template DNA strands and equilibrium Mg2+ binding properties of bacteriophage T4 DNA polymerase. Biochem **37:** 10144-55.

Beese, L. S., Derbyshire, V. and Steitz, T. A. (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA. Science **260:** 352-355.

Benton, W. D. and Davis, R. W. (1977) Screening lambda gt recombinant clones by hybridization to single plaques *in situ*. Science **196:** 180-182.

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nusslein-Volhard, C. (1988) The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo. EMBO J **7:** 1749-56.

Blatter, E. E., Ebright, Y. W. and Ebright, R. H. (1992) Identification of an amino acid-base contact in the GCN4-DNA complex by bromouracil-mediated photocrosslinking. Nature **359:** 650-2.

Bonner, C. A., Randall, S. K., Rayssiguier, C., Radman, M., Eritja, R., Kaplan, B. E., McEntee, K. and Goodman, M. F. (1988) Purification and characterization of an inducible *Escherichia coli* DNA polymerase capable of insertion and bypass at abasic lesions in DNA. J Biol Chem **263:** 18946-52.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M. and Venter, J. C. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii* [see comments]. Science **273:** 1058-73.

Burgers, P. M. and Gerik, K. J. (1998) Structure and processivity of two forms of *Saccharomyces cerevisiae* DNA polymerase delta. J Biol Chem **273:** 19756-62.

Capson, T. L., Benkovic, S. J. and Nossal, N. G. (1991) Protein-DNA cross-linking demonstrates stepwise ATP-dependent assembly of T4 DNA polymerase and its accessory proteins on the primer-template. Cell **65:** 249-58.

Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M. and Benkovic, S. J. (1992) Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. Biochem **31:** 10984-94.

Casari, G., De Fusco, M., Ciarmatori, S., Zeviani, M., Mora, M., Fernandez, P., De Michele, G., Filla, A., Cocozza, S., Marconi, R., Durr, A., Fontaine, B. and Ballabio, A. (1998) Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. Cell **93:** 973-83.

Cherbas, L. and Cherbas, P. (1993) The arthropod initiator: the capsite consensus plays an important role in transcription. Insect Biochem Mol Biol **23:** 81-90.

Chi, N. W. and Kolodner, R. D. (1994) Purification and characterization of MSH1, a yeast mitochondrial protein that binds to DNA mismatches. J Biol Chem **269:** 29984-92.

Clayton, D. A. (1982) Replication of animal mitochondrial DNA. Cell **28:** 693-705.

Clayton, D. A. (1992) Transcription and replication of animal mitochondrial DNAs. Int Rev Cytol 217-32.

Copeland, W. C. and Wang, T. S. (1991) Catalytic subunit of human DNA polymerase alpha overproduced from baculovirus-infected insect cells. Structural and enzymological characterization. J Biol Chem **266:** 22739-22748.

Copeland, W. C. and Wang, T. S. (1993) Mutational analysis of the human DNA polymerase alpha. The most conserved region in alpha-like DNA polymerases is involved in metal- specific catalysis. J Biol Chem **268:** 11028-11040.

Cortopassi, G. A. and Arnheim, N. (1990) Detection of a specific mitochondrial DNA deletion in tissues of older humans. Nucleic Acids Res **18:** 6927-33.

Cortopassi, G. A., Shibata, D., Soong, N. W. and Arnheim, N. (1992) A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues. Proc Natl Acad Sci U S A **89:** 7370-4.

Crute, J. J. and Lehman, I. R. (1991) Herpes simplex virus-1 helicase-primase. Physical and catalytic properties. J Biol Chem **266:** 4484-8.

Dahlberg, M. E. and Benkovic, S. J. (1991) Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant. Biochem **30:** 4835-43.

Davis, R. E., Miller, S., Herrnstadt, C., Ghosh, S. S., Fahy, E., Shinobu, L. A., Galasko, D., Thal, L. J., Beal, M. F., Howell, N. and Parker, W. D., Jr. (1997) Mutations in mitochondrial cytochrome c oxidase genes segregate with late-onset Alzheimer disease. Proc Natl Acad Sci U S A **94:** 4526-31.

Doublie, S., Tabor, S., Long, A. M., Richardson, C. C. and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution [see comments]. Nature **391:** 251-8.

Driggers, W. J., LeDoux, S. P. and Wilson, G. L. (1993) Repair of oxidative damage within the mitochondrial DNA of RINr 38 cells. J Biol Chem **268:** 22042-22045.

Duronio, R. J., O'Farrell, P. H., Xie, J. E., Brook, A. and Dyson, N. (1995) The transcription factor E2F is required for S phase during *Drosophila* embryogenesis. Genes Dev **9:** 1445-55.

Eom, S. H., Wang, J. and Steitz, T. A. (1996) Structure of Taq ploymerase with DNA at the polymerase active site. Nature **382:** 278-81.

Escarceller, M., Hicks, J., Gudmundsson, G., Trump, G., Touati, D., Lovett, S., Foster, P. L., McEntee, K. and Goodman, M. F. (1994) Involvement of *Escherichia coli* DNA polymerase II in response to oxidative damage and adaptive mutation. J Bacteriol **176:** 6221-8.

Fauron, C., Casper, M., Gao, Y. and Moore, B. (1995) The maize mitochondrial genome: dynamic, yet functional [published erratum appears in Trends Genet 1995 Jul;11(7):293]. Trends Genet **11:** 228-35.

Foury, F. (1989) Cloning and sequencing of the nuclear gene MIP1 encoding the catalytic subunit of the yeast mitochondrial DNA polymerase. J Biol Chem **264:** 20552-20560.

Foury, F. and Vanderstraeten, S. (1992) Yeast mitochondrial DNA mutators with deficient proofreading exonucleolytic activity. EMBO J **11:** 2717-26.

Gibbs, P. E., McGregor, W. G., Maher, V. M., Nisson, P. and Lawrence, C. W. (1998) A human homolog of the *Saccharomyces cerevisiae* REV3 gene, which encodes the catalytic subunit of DNA polymerase zeta. Proc Natl Acad Sci U S A **95:** 6876-80.

Glover, B. P. and McHenry, C. S. (1998) The chi psi subunits of DNA polymerase III holoenzyme bind to single- stranded DNA-binding protein (SSB) and facilitate replication of an SSB- coated template. J Biol Chem **273:** 23476-84.

Goddard, J. M. and Wolstenholme, D. R. (1980) Origin and direction of replication in mitochondrial DNA molecules from the genus *Drosophila*. Nucleic Acids Res **8:** 741-57.

Graves, S. W., Johnson, A. A. and Johnson, K. A. (1998) Expression, purification, and initial kinetic characterization of the large subunit of the human mitochondrial DNA polymerase. Biochem **37:** 6050-8.

Gray, H. and Wong, T. W. (1992) Purification and identification of subunit structure of the human mitochondrial DNA polymerase. J Biol Chem **267:** 5835-5841.

Gray, M. W. (1989a) The evolutionary origins of organelles. Trends Genet **5:** 294-9.

Gray, M. W. (1989b) Origin and evolution of mitochondrial DNA. Annu Rev Cell Biol **5:** 25-50.

Gray, M. W., Lang, B. F., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T. G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, Y. and Burger, G. (1998) Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res **26:** 865-78.

Harris, P. V., Mazina, O. M., Leonhardt, E. A., Case, R. B., Boyd, J. B. and Burtis, K. C. (1996) Molecular cloning of *Drosophila* mus308, a gene involved in DNA cross- link repair with homology to prokaryotic DNA polymerase I genes. Mol Cell Biol **16:** 5764-71.

Henikoff, S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351-9.

Hernandez, T. R. and Lehman, I. R. (1990) Functional interaction between the herpes simplex-1 DNA polymerase and UL42 protein. J Biol Chem **265:** 11227-32.

Hirano, M., Shtilbans, A., Mayeux, R., Davidson, M. M., DiMauro, S., Knowles, J. A. and Schon, E. A. (1997) Apparent mtDNA heteroplasmy in Alzheimer's disease patients and in normals due to PCR amplification of nucleus-embedded mtDNA pseudogenes. Proc Natl Acad Sci U S A **94:** 14894-9.

Hirose, F., Yamaguchi, M., Handa, H., Inomata, Y. and Matsukage, A. (1993) Novel 8-base pair sequence (*Drosophila* DNA replication-related element) and specific binding factor involved in the expression of *Drosophila* genes for DNA polymerase alpha and proliferating cell nuclear antigen. J Biol Chem **268:** 2092-9.

Hockensmith, J. W., Kubasek, W. L., Vorachek, W. R. and von Hippel, P. H. (1986) Laser cross-linking of nucleic acids to proteins. Methodology and first applications to the phage T4 DNA replication system. J Biol Chem **261:** 3512-8.

Hutchin, T. P., Heath, P. R., Pearson, R. C. and Sinclair, A. J. (1997) Mitochondrial DNA mutations in Alzheimer's disease. Biochem Biophys Res Commun **241:** 221-5.

Ikeda, S. and Ozaki, K. (1997) Action of mitochondrial endonuclease G on DNA damaged by L-ascorbic acid, peplomycin, and cis-diamminedichloroplatinum (II). Biochem Biophys Res Commun **235:** 291-4.

Insdorf, N. F. and Bogenhagen, D. F. (1989a) DNA polymerase gamma from *Xenopus laevis*. I. The identification of a high molecular weight catalytic subunit by a novel DNA polymerase photolabeling procedure. J Biol Chem **264:** 21491-21497.

Insdorf, N. F. and Bogenhagen, D. F. (1989b) DNA polymerase gamma from *Xenopus laevis*. II. A 3'----5' exonuclease is tightly associated with the DNA polymerase activity. J Biol Chem **264:** 21498-503.

Ito, J. and Braithwaite, D. K. (1991) Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res **19:** 4045-4057.

Jung, G. H., Leavitt, M. C., Schultz, M. and Ito, J. (1990) Site-specific mutagenesis of PRD1 DNA polymerase: mutations in highly conserved regions of the family B DNA polymerase. Biochem Biophys Res Commun **170:** 1294-300.

Kaguni, L. S. and Olson, M. W. (1989) Mismatch-specific 3'-5' exonuclease associated with the mitochondrial DNA polymerase from *Drosophila* embryos. Proc Natl Acad Sci U S A **86:** 6469-6473.

Kaguni, L. S., Rossignol, J. M., Conaway, R. C., Banks, G. R. and Lehman, I. R. (1983a) Association of DNA primase with the beta/gamma subunits of DNA polymerase alpha from *Drosophila melanogaster* embryos. J Biol Chem **258:** 9037-9.

Kaguni, L. S., Rossignol, J. M., Conaway, R. C. and Lehman, I. R. (1983b) Isolation of an intact DNA polymerase-primase from embryos of *Drosophila melanogaster*. Proc Natl Acad Sci U S A **80:** 2221-2225.

Kelman, Z. and Hurwitz, J. (1998) Protein-PCNA interactions: a DNA-scanning mechanism? Trends Biochem Sci **23:** 236-8.

Kelman, Z. and O'Donnell, M. (1995) DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. Annu Rev Biochem **64:** 171-200.

Kelman, Z., Yuzhakov, A., Andjelkovic, J. and O'Donnell, M. (1998) Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. Embo J **17:** 2436-49.

Kiefer, J. R., Mao, C., Hansen, C. J., Basehore, S. L., Hogrefe, H. H., Braman, J. C. and Beese, L. S. (1997) Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 A resolution. Structure **5:** 95-108.

Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W. and Steitz, T. A. (1995) Crystal structure of *Thermus aquaticus* DNA polymerase. Nature **376:** 612-6.

Kim, Y. T. and Richardson, C. C. (1993) Bacteriophage T7 gene 2.5 protein: an essential protein for DNA replication. Proc Natl Acad Sci U S A **90:** 10173-7.

King, A. J., Teertstra, W. R. and van der Vliet, P. C. (1997) Dissociation of the protein primer and DNA polymerase after initiation of adenovirus DNA replication. J Biol Chem **272:** 24617-23.

Kogelnik, A. M., Lott, M. T., Brown, M. D., Navathe, S. B. and Wallace, D. C. (1998) MITOMAP: a human mitochondrial genome database--1998 update. Nucleic Acids Res **26:** 112-5.

Kong, X. P., Onrust, R., O'Donnell, M. and Kuriyan, J. (1992) Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. Cell **69:** 425-37.

Kornberg, A. and Baker, T. A. (1992) DNA Replication.

Krishna, T. S., Fenyo, D., Kong, X. P., Gary, S., Chait, B. T., Burgers, P. and Kuriyan, J. (1994) Crystallization of proliferating cell nuclear antigen (PCNA) from *Saccharomyces cerevisiae*. J Mol Biol **241:** 265-8.

Kroutil, L. C., Frey, M. W., Kaboord, B. F., Kunkel, T. A. and Benkovic, S. J. (1998) Effect of accessory proteins on T4 DNA polymerase replication fidelity. J Mol Biol **278:** 135-46.

Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A. and Benkovic, S. J. (1987) Kinetic mechanism of DNA polymerase I (Klenow). Biochem **26:** 8410-7.

Kunkel, T. A. (1992) DNA replication fidelity. J Biol Chem **267:** 18251-4.

Kunkel, T. A. and Mosbaugh, D. W. (1989) Exonucleolytic proofreading by a mammalian DNA polymerase. Biochem **28:** 988-995.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680-685.

Lantz, V., Ambrosio, L. and Schedl, P. (1992) The *Drosophila* orb gene is predicted to encode sex-specific germline RNA-binding proteins and has localized transcripts in ovaries and early embryos. Development **115:** 75-88.

Larsson, N. G. and Clayton, D. A. (1995) Molecular genetic aspects of human mitochondrial disorders. Annu Rev Genet **29:** 151-78.

Larsson, N. G., Eiken, H. G., Boman, H., Holme, E., Oldfors, A. and Tulinius, M. H. (1992) Lack of transmission of deleted mtDNA from a woman with Kearns-Sayre syndrome to her child. Am J Hum Genet **50:** 360-3.

Leblanc, C., Richard, O., Kloareg, B., Viehmann, S., Zetsche, K. and Boyen, C. (1997) Origin and evolution of mitochondria: what have we learnt from red algae? Curr Genet **31:** 193-207.

Lecrenier, N., Van Der Bruggen, P. and Foury, F. (1997) Mitochondrial DNA polymerases from yeast to man: a new family of polymerases. Gene **185:** 147-52.

LeDoux, S. P., Patton, N. J., Avery, L. J. and Wilson, G. L. (1993) Repair of N-methylpurines in the mitochondrial DNA of xeroderma pigmentosum complementation group D cells. Carcinogenesis **14:** 913-917.

LeDoux, S. P., Wilson, G. L., Beecham, E. J., Stevnsner, T., Wassermann, K. and Bohr, V. A. (1992) Repair of mitochondrial DNA after various types of DNA damage in Chinese hamster ovary cells. Carcinogenesis **13:** 1967-1973.

Lee, D. Y. and Clayton, D. A. (1997) RNase mitochondrial RNA processing correctly cleaves a novel R loop at the mitochondrial DNA leading-strand origin of replication. Genes Dev **11:** 582-92.

Lee, D. Y. and Clayton, D. A. (1998) Initiation of mitochondrial DNA replication by transcription and R-loop processing [In Process Citation]. J Biol Chem **273:** 30614-21.

Lehman, I. R. and Kaguni, L. S. (1989) DNA polymerase alpha. J Biol Chem **264:** 4265-4268.

Lehner, C. F. and O'Farrell, P. H. (1989) Expression and function of *Drosophila* cyclin A during embryonic cell cycle progression. Cell **56:** 957-68.

Lewis, D. L., Farr, C. L., Wang, Y., Lagina, A. T. r. and Kaguni, L. S. (1996) Catalytic subunit of mitochondrial DNA polymerase from *Drosophila* embryos. Cloning, bacterial overexpression, and biochemical characterization. J Biol Chem **271:** 23389-94.

Li, Y., Korolev, S. and Waksman, G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation. EMBO J **17:** 7514-7525.

Lin, S. Y. and Riggs, A. D. (1974) Photochemical attachment of lac repressor to bromodeoxyuridine- substituted lac operator by ultraviolet radiation. Proc Natl Acad Sci U S A **71:** 947-51.

Lohman, T. M. (1993) Helicase-catalyzed DNA unwinding. J Biol Chem **268:** 2269-72.

Longley, M. J., Prasad, R., Srivastava, D. K., Wilson, S. H. and Copeland, W. C. (1998a) Identification of 5'-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair in vitro. Proc Natl Acad Sci U S A **95:** 12244-8.

Longley, M. J., Ropp, P. A., Lim, S. E. and Copeland, W. C. (1998b) Characterization of the native and recombinant catalytic subunit of human DNA polymerase gamma: identification of residues critical for exonuclease activity and dideoxynucleotide sensitivity. Biochem **37:** 10529-39.

Lucas, P., Laquel-Robert, P., Plissonneau, J., Schaeffer, J., Tarrago-Litvak, L. and Castroviejo, M. (1997) A second DNA polymerase activity in yeast mitochondria. C R Acad Sci III **320:** 299-305.

Martin, W. and Muller, M. (1998) The hydrogen hypothesis for the first eukaryote [see comments]. Nature **392:** 37-41.

Marton, R. F., Thommes, P. and Cotterill, S. (1994) Purification and characterisation of dRP-A: a single-stranded DNA binding protein from *Drosophila melanogaster*. Febs Lett **342:** 139-144.

Masker, W., Hanawalt, P. and Shizuya, H. (1973) Role of DNA polymerase II in repair replication in *Escherichia coli*. Nat New Biol **244:** 242-3.

Matsukage, A., Hirose, F., Hayashi, Y., Hamada, K. and Yamaguchi, M. (1995) The DRE sequence TATCGATA, a putative promoter-activating element for *Drosophila melanogaster* cell-proliferation-related genes. Gene **166:** 233-6.

Matsumoto, Y., Kim, K., Katz, D. S. and Feng, J. A. (1998) Catalytic center of DNA polymerase beta for excision of deoxyribose phosphate groups. Biochem **37:** 6456-64.

Maxam, A. M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol **65:** 499-560.

McHenry, C. S. (1988) DNA polymerase III holoenzyme of *Escherichia coli*. Ann Rev Biochem **57:** 519-550.

Mendez, J., Blanco, L. and Salas, M. (1997) Protein-primed DNA replication: a transition between two modes of priming by a unique DNA polymerase. EMBO J **16:** 2519-27.

Mignotte, B. and Vayssiere, J. L. (1998) Mitochondria and apoptosis. Eur J Biochem **252:** 1-15.

Mita, S., Rizzuto, R., Moraes, C. T., Shanske, S., Arnaudo, E., Fabrizi, G. M., Koga, Y., DiMauro, S. and Schon, E. A. (1990) Recombination via flanking direct repeats is a major cause of large- scale deletions of human mitochondrial DNA. Nucleic Acids Res **18:** 561-7.

Miyazawa, H., Izumi, M., Tada, S., Takada, R., Masutani, M., Ui, M. and Hanaoka, F. (1993) Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase alpha-primase complex and their gene expression during cell proliferation and the cell cycle. J Biol Chem **268:** 8111-22.

Mizuno, T., Ito, N., Yokoi, M., Kobayashi, A., Tamai, K., Miyazawa, H. and Hanaoka, F. (1998) The second-largest subunit of the mouse DNA polymerase alpha-primase complex facilitates both production and nuclear translocation of the catalytic subunit of DNA polymerase alpha. Mol Cell Biol **18:** 3552-62.

Morrison, A., Christensen, R. B., Alley, J., Beck, A. K., Bernstine, E. G., Lemontt, J. F. and Lawrence, C. W. (1989) REV3, a *Saccharomyces cerevisiae* gene whose function is required for induced mutagenesis, is predicted to encode a nonessential DNA polymerase. J Bacteriol **171:** 5659-67.

Munn, M. M. and Alberts, B. M. (1991) DNA footprinting studies of the complex formed by the T4 DNA polymerase holoenzyme at a primer-template junction. J Biol Chem **266:** 20034-20044.

Murphy, G. J., Murphy, G. and Reynolds, J. J. (1991) The origin of matrix metalloproteinases and their familial relationships. FEBS Lett **289:** 4-7.

Nakamura, A., Amikura, R., Mukai, M., Kobayashi, S. and Lasko, P. F. (1996) Requirement for a noncoding RNA in *Drosophila* polar granules for germ cell establishment. Science **274:** 2075-9.

Naktinis, V., Onrust, R., Fang, L. and O'Donnell, M. (1995) Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. II. Intermediate complex between the clamp loader and its clamp. J Biol Chem **270:** 13358-65.

Nasheuer, H. P., Moore, A., Wahl, A. F. and Wang, T. S. (1991) Cell cycle-dependent phosphorylation of human DNA polymerase alpha. J Biol Chem **266:** 7893-903.

Nelson, J. R., Lawrence, C. W. and Hinkle, D. C. (1996) Thymine-thymine dimer bypass by yeast DNA polymerase zeta. Science **272:** 1646-9.

Nosek, J., Tomaska, L., Fukuhara, H., Suyama, Y. and Kovac, L. (1998) Linear mitochondrial genomes: 30 years down the line. Trends Genet **14:** 184-8.

Ohno, K., Hirose, F., Sakaguchi, K., Nishida, Y. and Matsukage, A. (1996) Transcriptional regulation of the *Drosophila* CycA gene by the DNA replication-related element (DRE) and DRE binding factor (DREF). Nucleic Acids Res **24:** 3942-6.

Ohtani, K. and Nevins, J. R. (1994) Functional properties of a *Drosophila* homolog of the E2F1 gene. Mol Cell Biol **14:** 1603-12.

Olson, M. W. and Kaguni, L. S. (1992) 3'-5' exonuclease in *Drosophila* mitochondrial DNA polymerase. Substrate specificity and functional coordination of nucleotide polymerization and mispair hydrolysis. J Biol Chem **267:** 23136-23142.

Olson, M. W., Wang, Y., Elder, R. H. and Kaguni, L. S. (1995) Subunit structure of mitochondrial DNA polymerase from *Drosophila* embryos. Physical and immunological studies. J Biol Chem **270:** 28932-7.

Onrust, R., Finkelstein, J., Turner, J., Naktinis, V. and O'Donnell, M. (1995) Assembly of a chromosomal replication machine: two DNA polymerases, a clamp loader, and sliding clamps in one holoenzyme particle. III. Interface between two polymerases and the clamp loader. J Biol Chem **270:** 13366-77.

Onrust, R., Stukenberg, P. T. and O'Donnell, M. (1991) Analysis of the ATPase subassembly which initiates processive DNA synthesis by DNA polymerase III holoenzyme. J Biol Chem **266:** 21681-6.

Park, K., Debyser, Z., Tabor, S., Richardson, C. C. and Griffith, J. D. (1998) Formation of a DNA loop at the replication fork generated by bacteriophage T7 replication proteins. J Biol Chem **273:** 5260-70.

Patel, S. S., Wong, I. and Johnson, K. A. (1991) Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. Biochem **30:** 511-25.

Pinz, K. G. and Bogenhagen, D. F. (1998) Efficient repair of abasic sites in DNA by mitochondrial enzymes. Mol Cell Biol **18:** 1257-65.

Plevani, P., Foiani, M., Muzi Falconi, M., Pizzagalli, A., Santocanale, C., Francesconi, S., Valsasnini, P., Comedini, A., Piatti, S. and Lucchini, G. (1988) The yeast DNA polymerase-primase complex: genes and proteins. Biochim Biophys Acta **951:** 268-73.

Prasad, R., Beard, W. A., Strauss, P. R. and Wilson, S. H. (1998) Human DNA polymerase beta deoxyribose phosphate lyase. Substrate specificity and catalytic mechanism. J Biol Chem **273:** 15263-70.

Rangarajan, S., Gudmundsson, G., Qiu, Z., Foster, P. L. and Goodman, M. F. (1997) *Escherichia coli* DNA polymerase II catalyzes chromosomal and episomal DNA synthesis in vivo. Proc Natl Acad Sci U S A **94:** 946-51.

Reems, J. and McHenry, C. S. (1994) *Escherichia coli* DNA polymerase III holoenzyme footprints three helical turns of its primer. J Biol Chem **269:** 33091-33096.

Reha-Krantz, L. J., Marquez, L. A., Elisseeva, E., Baker, R. P., Bloom, L. B., Dunford, H. B. and Goodman, M. F. (1998) The proofreading pathway of bacteriophage T4 DNA polymerase. J Biol Chem **273:** 22969-76.

Robin, E. D. and Wong, R. (1988) Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Physiol **136:** 507-13.

Safer, B., Cohen, R. B., Garfinkel, S. and Thompson, J. A. (1988) DNA affinity labeling of adenovirus type 2 upstream promoter sequence- binding factors identifies two distinct proteins. Mol Cell Biol **8:** 105-13.

Sala, F., Amileni, A. R., Parisi, B. and Spadari, S. (1980) A gamma-like DNA polymerase in spinach chloroplasts. Eur J Biochem **112:** 211-7.

Sawado, T., Hirose, F., Takahashi, Y., Sasaki, T., Shinomiya, T., Sakaguchi, K., Matsukage, A. and Yamaguchi, M. (1998) The DNA replication-related element (DRE)/DRE-binding factor system is a transcriptional regulator of the *Drosophila* E2F gene. J Biol Chem **273:** 26042-51.

Sawaya, M. R., Pelletier, H., Kumar, A., Wilson, S. H. and Kraut, J. (1994) Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. Science **264:** 1930-1935.

Sawaya, M. R., Prasad, R., Wilson, S. H., Kraut, J. and Pelletier, H. (1997) Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. Biochem **36:** 11205-15.

Schatz, G. (1996) The protein import system of mitochondria. J Biol Chem **271:** 31763-6.

Schneider, A., Smith, R. W., Kautz, A. R., Weisshart, K., Grosse, F. and Nasheuer, H. P. (1998) Primase activity of human DNA polymerase alpha-primase. Divalent cations stabilize the enzyme activity of the p48 subunit. J Biol Chem **273:** 21608-15.

Shadel, G. S. and Clayton, D. A. (1997) Mitochondrial DNA maintenance in vertebrates. Annu Rev Biochem **66:** 409-35.

Sobol, R. W., Horton, J. K., Kuhn, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K. and Wilson, S. H. (1996) Requirement of mammalian DNA polymerase-beta in base-excision repair [published errata appear in Nature 1996 Feb 29;379(6568):848 and 1996 Oct 3;383(6599):457]. Nature **379:** 183-6.

Spicer, E. K., Rush, J., Fung, C., Reha-Krantz, L. J., Karam, J. D. and Konigsberg, W. H. (1988) Primary structure of T4 DNA polymerase. Evolutionary relatedness to eucaryotic and other procaryotic DNA polymerases. J Biol Chem **263:** 7478-86.

Sugino, A. (1995) Yeast DNA polymerases and their role at the replication fork. Trends Biochem Sci **20:** 319-23.

Sugino, A., Ohara, T., Sebastian, J., Nakashima, N. and Araki, H. (1998) DNA polymerase epsilon encoded by cdc20+ is required for chromosomal DNA replication in the fission yeast *Schizosaccharomyces pombe*. Genes Cells **3:** 99-110.

Tabor, S., Huber, H. E. and Richardson, C. C. (1987) *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. J Biol Chem **262:** 16212-16223.

Takao, M., Aburatani, H., Kobayashi, K. and Yasui, A. (1998) Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. Nucleic Acids Res **26:** 2917-22.

Tanhauser, S. M. and Laipis, P. J. (1995) Multiple deletions are detectable in mitochondrial DNA of aging mice. J Biol Chem **270:** 24769-75.

Thommes, P., Farr, C. L., Marton, R. F., Kaguni, L. S. and Cotterill, S. (1995) Mitochondrial single-stranded DNA-binding protein from *Drosophila* embryos. Physical and biochemical characterization. J Biol Chem **270:** 21137-43.

Torri, A. F. and Englund, P. T. (1995) A DNA polymerase beta in the mitochondrion of the trypanosomatid *Crithidia fasciculata*. J Biol Chem **270:** 3495-7.

Tsuchihashi, Z. and Kornberg, A. (1989) ATP interactions of the tau and gamma subunits of DNA polymerase III holoenzyme of *Escherichia coli*. J Biol Chem **264:** 17790-5.

Tsuzuki, T., Nomiyama, H., Setoyama, C., Maeda, S. and Shimada, K. (1983) Presence of mitochondrial-DNA-like sequences in the human nuclear DNA. Gene **25:** 223-9.

Veltri, K. L., Espiritu, M. and Singh, G. (1990) Distinct genomic copy number in mitochondria of different mammalian organs. J Cell Physiol **143:** 160-4.

Virbasius, C. A., Virbasius, J. V. and Scarpulla, R. C. (1993) NRF-1, an activator involved in nuclear-mitochondrial interactions, utilizes a new DNA-binding domain conserved in a family of developmental regulators. Genes Dev **7:** 2431-45.

Virbasius, J. V. and Scarpulla, R. C. (1994) Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis. Proc Natl Acad Sci U S A **91:** 1309-13.

Voitenleitner, C., Rehfuess, C., Hilmes, M., O'Rear, L., Liao, P. C., Gage, D. A., Ott, R., Nasheuer, H. P. and Fanning, E. (1999) Cell Cycle-Dependent Regulation of Human DNA Polymerase alpha-Primase Activity by Phosphorylation. Mol Cell Biol **19:** 646-656.

Wallace, D. C. (1992) Diseases of the mitochondrial DNA. Annu Rev Biochem **61:** 1175-1212.

Wallace, D. C., Shoffner, J. M., Trounce, I., Brown, M. D., Ballinger, S. W., Corral-Debrinski, M., Horton, T., Jun, A. S., and Lott, M. T. (1995a) Mitochondrial DNA mutations in human degenerative diseases and aging. Biochim Biophys Acta **1271:** 141-151.

Wallace, D. C. (1995b) Mitochondrial DNA variation in human evolution, degenerative disease, and aging. Am J Hum Genet **57:** 201-223.

Wallace, D. C., Stugard, C., Murdock, D., Schurr, T. and Brown, M. D. (1997) Ancient mtDNA sequences in the human nuclear genome: a potential source of errors in identifying pathogenic mutations. Proc Natl Acad Sci U S A **94:** 14900-5.

Wang, J., Sattar, A. K., Wang, C. C., Karam, J. D., Konigsberg, W. H. and Steitz, T. A. (1997a) Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. Cell **89:** 1087-99.

Wang, J., Yu, P., Lin, T. C., Konigsberg, W. H. and Steitz, T. A. (1996) Crystal structures of an NH2-terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions. Biochem **35:** 8110-9.

Wang, T. S. (1991) Eukaryotic DNA polymerases. Annu Rev Biochem 513-552.

Wang, T. S., Wong, S. W. and Korn, D. (1989) Human DNA polymerase alpha: predicted functional domains and relationships with viral DNA polymerases. Faseb J **3:** 14-21.

Wang, Y., Farr, C. L. and Kaguni, L. S. (1997b) Accessory subunit of mitochondrial DNA polymerase from *Drosophila* embryos. Cloning, molecular analysis, and association in the native enzyme. J Biol Chem **272:** 13640-6.

Ward, P. and Berns, K. I. (1995) Minimum origin requirements for linear duplex AAV DNA replication in vitro. Virology **209:** 692-5.

Wernette, C. M., Conway, M. C. and Kaguni, L. S. (1988) Mitochondrial DNA polymerase from *Drosophila melanogaster* embryos: kinetics, processivity, and fidelity of DNA polymerization. Biochem **27:** 6046-6054.

Wernette, C. M. and Kaguni, L. S. (1986) A mitochondrial DNA polymerase from embryos of *Drosophila melanogaster.* Purification, subunit structure, and partial characterization. J Biol Chem **261:** 14764-14770.

Whittakr, P. A., and S. M. Danks (1978) Mitochonria: structure, function, and assembly.

Williams, A. J. and Kaguni, L. S. (1995) Stimulation of *Drosophila* mitochondrial DNA polymerase by single-stranded DNA-binding protein. J Biol Chem **270:** 860-865.

Wold, M. S. and Kelly, T. (1988) Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. Proc Natl Acad Sci U S A **85:** 2523-7.

Ye, F., Carrodeguas, J. A. and Bogenhagen, D. F. (1996) The gamma subfamily of DNA polymerases: cloning of a developmentally regulated cDNA encoding *Xenopus laevis* mitochondrial DNA polymerase gamma. Nucleic Acids Res **24:** 1481-8.

Young, M. C., Reddy, M. K. and von Hippel, P. H. (1992) Structure and function of the bacteriophage T4 DNA polymerase holoenzyme. Biochem **31:** 8675-90.

Zhou, J. Q., He, H., Tan, C. K., Downey, K. M. and So, A. G. (1997) The small subunit is required for functional interaction of DNA polymerase delta with the proliferating cell nuclear antigen. Nucleic Acids Res **25:** 1094-9.

Zhu, W. and Ito, J. (1994) Family A and family B DNA polymerases are structurally related: evolutionary implications. Nucleic Acids Res **22:** 5177-83.

Zlotkin, T., Kaufmann, G., Jiang, Y., Lee, M. Y., Uitto, L., Syvaoja, J., Dornreiter, I., Fanning, E. and Nethanel, T. (1996) DNA polymerase epsilon may be dispensable for SV40- but not cellular- DNA replication. Embo J **15:** 2298-305.

Zuo, S., Gibbs, E., Kelman, Z., Wang, T. S., O'Donnell, M., MacNeill, S. A. and Hurwitz, J. (1997) DNA polymerase delta isolated from *Schizosaccharomyces pombe* contains five subunits. Proc Natl Acad Sci U S A **94:** 11244-9.

Zwicker, J. and Muller, R. (1997) Cell-cycle regulation of gene expression by transcriptional repression. Trends Genet **13:** 3-6.