THES!S
2
2000

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
JUN 2 8 2004		

HOMOLOGY MODELING OF THE ACCESSORY SUBUNIT OF
DROSOPHILA MITOCHONDRIAL DNA POLYMERASE AND
SUBUNIT INTERACTION STUDIES

By

Li Fan

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

2000

ABSTRACT

HOMOLOGY MODELING OF THE ACCESSORY SUBUNIT OF
DROSOPHILA MITOCHONDRIA DNA POLYMERASE AND
SUBUNIT INTERACTION STUDIES

By

Li Fan

Mitochondrial DNA (mtDNA) replication is essential for the
maintenance of the mitochondrial genome and the functions of
mitochondria. The key enzyme in mtDNA replication is mitochondrial
DNA polymerase or pol γ. The native enzyme isolated from embryos of
the fruit fly *Drosophila* is a heterodimer of a catalytic subunit (125 kDa)
and an accessory subunit (35 kDa). The catalytic subunit contains DNA
polymerase and 3'-5' exonuclease activities. Little is known about the
function or properties of the accessory subunit, although available data
suggest that the accessory subunit is important for the catalytic and
structural integrity of pol γ.

A computational modeling method was used to predict the structure
of the C-terminus (residues 254-361) of the accessory subunit based on
sequence homology with the anticodon binding domain of class IIa
aminoacyl-tRNA synthetases. The property of binding to the stem-loop of
cognate tRNA molecules in the latter implies that the accessory subunit may
have the potential of binding to the unusual RNA primers at mtDNA

replication initiation sites and serving a role as a primer recognition factor. This structural model also shows similarity to the structure of *E. coli* thioredoxin, the processivity factor of T7 DNA polymerase, suggesting that the accessory subunit may be the processivity factor of pol γ as well.

Physical interactions between the two subunits of *Drosophila* pol γ have been investigated by both *in vivo* reconstitution and *in vitro* protein overlay assays. Results with various deletion mutants suggest that the accessory subunit consists of three domains, termed the N-, M-, and C-domains. Both the M- and C-domains appear to comprise the main contacts involved in subunit interaction, possibly through multiple sites with the exonuclease (exo) domain and part of the spacer between the exo and polymerase (pol) domains in the catalytic subunit. Furthermore, the N-domain may modulate the subunit assembly through weak interaction with the pol domain. Sequence comparison suggests that the M-domain of the accessory subunit shares significant similarity with the RNase H domain of HIV-1 reverse transcriptase (RT). Based on these results and previous observations that pol γ is capable of synthesizing DNA using an RNA template as is a property of RT, we propose that the overall conformation and arrangement of functional domains in the heterodimeric pol γ complex resemble those of HIV-RT.

To my family

ACKNOWLEDGMENTS

I acknowledge the Department of Biochemistry, the Graduate Program Director Dr. Pamela Fraker, and Graduate Programs secretary Julie Oesterle for their support and help.

Finally, I would like to thank my family for their love, support and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aaRS	aminoacyl-tRNA synthetase
A+T	deoxyadenylate and thymidylate
ATP	adenosine triphosphate
bp(s)	base pair(s)
C-terminus	carboxyl-terminus
C-domain	residues 254-362 of pol γ-β
D-loop	displacement loop
D. melanogaster	*Drosophila melanogaster*
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
Gly-RS	glycyl-tRNA synthetase
HIV	human immunodeficiency virus
HSV	herpes simplex virus
kbp	kilobase pairs
kDa	kiloDalton
M-domain	residues 128-240 of pol γ-β
mRNA	messenger RNA

mtDNA	mitochondrial DNA
mtSSB	mitochondrial single-stranded-DNA binding protein
N-terminus	amino terminus
N-domain	residues 1-120 of pol γ–β
Ni-NTA	Nickel chelated nitrilo-tri-acetic acid
nt	nucleotide
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
pol γ	mitochondrial DNA polymerase
pol γ–α	catalytic subunit of pol γ
pol γ–β	accessory subunit of pol γ
Pro-RS	prolyl-tRNA synthetase
RNA	ribonucleic acid
RF-C	replication factor C
RT	reverse transcriptase
SSB	single-stranded DNA-binding protein
ssDNA	single-stranded DNA
SV40	simian virus 40
tRNA	transfer RNA
UV	ultraviolet light
X. laevis	*Xenopus laevis*

CHAPTER I

INTRODUCTION

Mitochondria are the central organelles of cellular energy metabolism. They are intimately involved with most of the major metabolic pathways used by a cell to build, break down, and recycle its molecular building blocks such as purines and pyrimidines for RNA and DNA synthesis. One principal mitochondrial function is to produce ATP (adenosine triphosphate) through the process of oxidative phosphorylation (OXPHOS) (Wallace, 1992). Most human cells contain hundreds of mitochondria. These cytoplasmic organelles are believed to have evolved about 1.5 billion years ago from eubacteria-like endosymbionts in a primitive eukaryotic cell as the host (Gray, 1989). Modern mitochondria still retain some features that reflect their endosymbiotic origin. These include a double-membrane structure and a circular mitochondrial DNA genome with mitochondrial-specific replication, transcription, translation, and protein assembly systems (for reviews see (Attardi, 1985; Gray, 1989; Clayton, 1991)). Mitochondria are the only cellular organelles to have their own DNA other than the nuclei in animal cells. Mitochondrial DNA (mtDNA) encodes proteins critical for oxidative phosphorylation (Wallace, 1992). Therefore, the proper maintenance and faithful replication of mtDNA molecules are very important for mitochondria to achieve their biological functions.

Mitochondrial defects were first noticed as rare childhood disorders (for review see (Luft, 1994)). It is now clear that mitochondrial diseases

are manifested as an extraordinarily broad spectrum of clinical problems, commonly involving the most energy-dependent organs such as brain, heart, skeletal muscle, kidney, endocrine glands and bone marrow (Wallace, 1994). Many common diseases such as diabetes and ischemic heart disease have a defective mitochondrial basis (Wallace, 1999). Aging-related diseases such as Parkinson's disease and Alzheimer's disease may result in part from mitochondrial dysfunction. Mitochondrial medicine has achieved great progress in the past years (Luft, 1994). A breakthrough in mitochondrial pathophysiology occurred in 1988 when mutations in mtDNA were reported to be associated with a form of young-adult blindness, Leber's hereditary optic neuropathy (LHON) (Wallace et al., 1988), and progressive muscle disorders (Holt et al., 1988). Since then, more than 50 pathogenic mtDNA base substitution mutations and hundreds of mtDNA rearrangement mutations (deletions and insertions) have been identified in a variety of degenerative diseases (MITOMAP: http://www.gen.emory.edu/mitomap.html). Some mitochondrial diseases involving mtDNA mutations may be inherited through the maternal side of the family as almost all mitochondria come from the mother (Wallace, 1995). In addition, mitochondrial toxicity caused by long term treatment with antiviral nucleoside analogue (ANA) drugs was found to resemble those of genetic mitochondrial diseases. These ANAs include zidovudine (AZT), zalcitabine (ddC), didanosine (ddI), stavudine (D4T), 2'-deoxy-3'-

thiacytidine (3TC), and 1-(2'-deoxy-2'fluoro-1-β-D-arabinofuranosyl 5-iodo)uracil (FIAU) for treatment of human immunodeficiency virus type-1 (HIV-1) and hepatitis B virus (for review see (Lewis and Dalakas, 1995)). The mitochondrial toxicity was at least partially due to their inhibition of mitochondrial DNA polymerase, a key enzyme for mtDNA replication (Chen and Cheng, 1989; Chen et al., 1991). These results gave evidence for direct connections between mtDNA replication defects and mitochondrial dysfunction. Therefore, studies on mtDNA replication and related enzymes and proteins have major medical relevance.

Mitochondrial DNA

Mitochondrial DNA molecules represent less than 1% of the total cellular DNA in a typical animal cell, but the copy number of mtDNA molecules is in great excess over the nuclear chromosomal DNA molecules (Clayton, 1982). Each mitochondrion has several mtDNA molecules. Mitochondrial DNA is located in the matrix which is encircled by a double-membrane structure. The inner membrane is folded into cristae toward the matrix. The outer membrane separates the organelle from the cytosol. There is an intermembrane space between the outer and inner membranes. There are about 3000 proteins in each mitochondrion, virtually all of them are encoded by nuclear genes (Lill et al., 1996). These proteins are synthesized on cytosolic ribosomes and have to be imported into the

mitochondrial subcompartments (Schatz, 1996). Many of these proteins have a mitochondrial target sequence at their N-terminus and the target sequences are usually removed after the proteins are delivered to the correct locations (Lill et al., 1996). Proteins encoded by mtDNA are expressed inside the mitochondrion which has its own machinery for gene expression and DNA replication (Clayton, 1982; Clayton, 1992).

Many mitochondrial DNA molecules from animals, plants, fungi and protozoa have been well characterized over the years (Attardi and Schatz, 1988; Gray, 1989; Leblanc et al., 1997). Sequences for more than 70 complete mitochondrial DNA genomes are now available from the GenBank database. Mitochondrial DNA has the same basic role in all eukaryotes: it encodes rRNA and tRNA components of a mitochondrial protein synthesizing machinery (Attardi and Schatz, 1988; Gray, 1989). However, proteins for gene expression and mtDNA replication are encoded by nuclear genes. The mtDNA genome also contains genes encoding proteins involved in the electron transport chain for oxidative phosphorylation. Although conservative in basic genetic function, mtDNA differs radically in structure among and even within the four traditional eukaryotic kingdoms (*Animalia, Plantae, Fungi and Protista*). Some mitochondrial genomes also contain additional genes encoding proteins associated with mitochondrial ribosomes and/or proteins involved in mitochondrial RNA processing or intron transposition (Gray, 1989).

Mitochondrial genome size varies over more than a 150-fold range, from 14.3 kbp in the nematode worm, *Ascaris suum* (Wolstenholme et al., 1987), to an estimated high of 2400 kbp in the cucurbit, *Cucumis melo* (muskmelon) (Gray, 1989). Most mtDNAs are circular but linear mtDNAs have been found in some species (Nosek et al., 1998). Nevertheless, with a few notable exceptions, the overall genetic content and gene order in mtDNA are remarkably conserved among animal species (Wolstenholme, 1992).

Animal mitochondrial genomes range in size from 16 to 19 kbp and can be isolated as closed circular double-stranded DNA molecules (Clayton, 1991). Because the mtDNA is supercoiled, its replication and transcription can be inhibited by ethidium bromide. The two strands of vertebrate mtDNA can be separated in denaturing cesium chloride gradients, due to a strand bias in guanine and thymidine (G+T) content, and are thus defined as the heavy strand (H-strand) and the light strand (L-strand), respectively. A major form observed for vertebrate mtDNA is the displacement-loop (D-loop) molecule, which contains a short piece of nascent H strand at the H-strand origin (O_H) of DNA replication (Clayton, 1982).

The human mitochondrial genome is about 16.5 kbp which encodes 13 essential polypeptides of oxidative phosphorylation plus two rRNAs and twenty-two tRNAs necessary for gene expression within the mitochondrion (Wallace, 1992) (Figure 1.1). The proteins encoded are seven subunits of

Figure 1.1   Human mtDNA map. The mtDNA comprises 16,569 bps with numbering start at O_H and proceeding counterclockwise around the circular map. The function of each gene is identified by shading according to the designations within the circle. The tRNA genes are indicated by the letter of their cognate amino acid. Transcription directions at the promoters (P_H and P_L) as well as the directions at the two DNA replication initiation sites (O_H and O_L) are indicated by arrows. Mitochondrial disease-related mutations including one 5 kb deletion and several point mutations are presented. Point mutations are named by the associated diseases followed by DNA sequence number and the mutated amino acid residue. Diseases abbreviations are: LHON, Leber's hereditary Optic Neuropath; NARP, Neurogenic muscle weakness, Ataxia, and Retinitis Pigmentosa; MERRF, Myoclonic Epilepsy Ragged-Red Fiber Disease; MELAS, Mitochondrial Encephalomyopathy, Lactic Acidosis, and Strokelike epsodes; LDYT, LHON and dystonia; DEAF, deafness.

The Human Mitchondrial DNA
From MITOMAP http://www.gen.emory.edu/mitomap.html

5 kb deletion
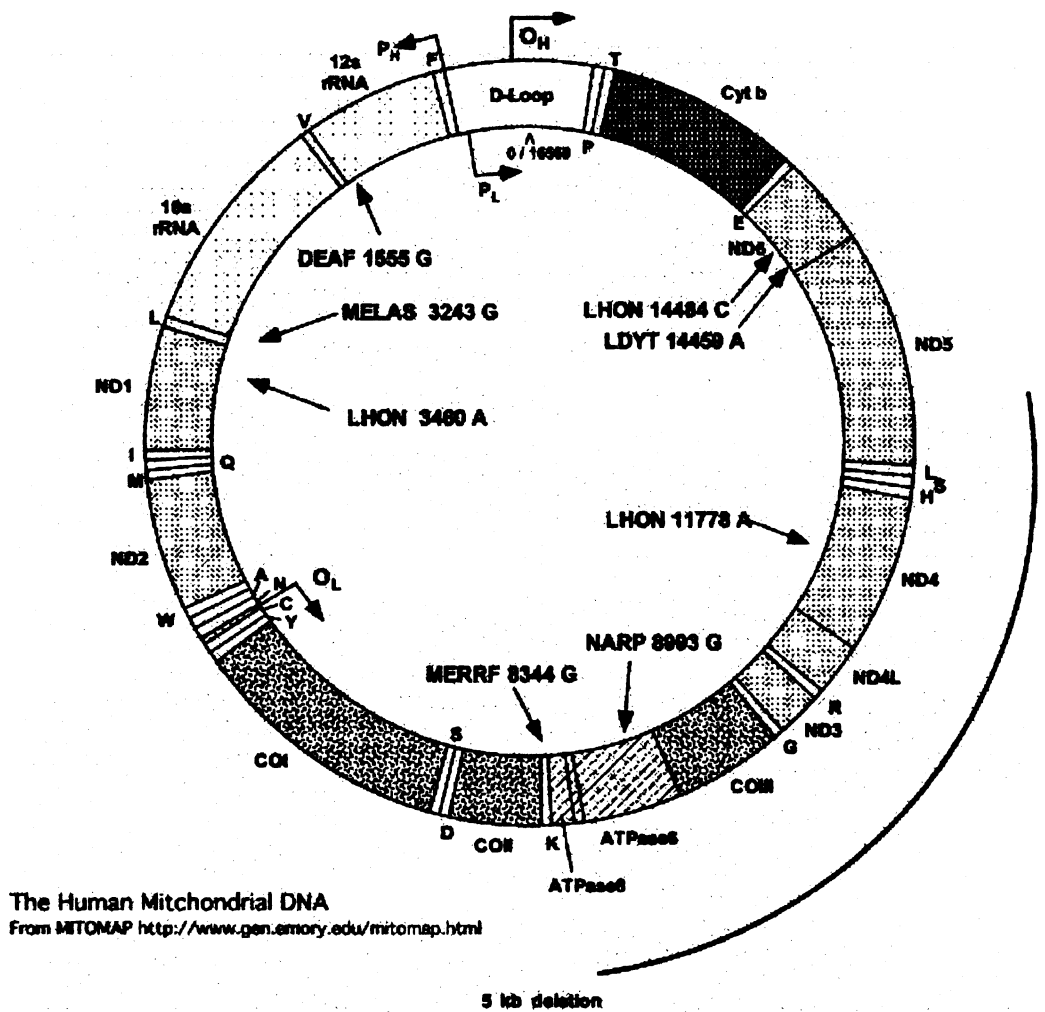
Figure 1.1

NADH dehydrogenase (ND1, ND2, ND3, ND4, ND4L, ND5, and ND6), the apocytochrome b component (Cyt b) of ubiquinol cytochrome-c reductase, three subunits of cytochrome-c oxidase (COI, COII, and COIII), and two subunits of ATP synthase (ATPase 6 and 8). Gene Cyt b and COI are common to all the mitochondrial genomes whose coding capacity has been defined completely (Okimoto and Wolstenholme, 1990). The H-strand functions as the template for the small (12S) and large (16S) rRNAs, 12 of the polypeptides, and 14 of the tRNAs. The L-strand is the template for one protein and 8 tRNAs (for details see Figure 1.1). The D-loop region contains two transcription promoters, one for the H-strand (P_H) and the other for the L-strand (P_L) in the opposite direction. MtDNA is transcribed symmetrically from these two promoters (for reviews see (Clayton, 1984; Clayton, 1992)). The two transcripts encompass the entire genome (Montoya et al., 1981). As the tRNAs are transcribed, they fold into their three-dimensional structures and are cleaved out of the transcript by a processing enzyme comparable to RNase P. Because most protein coding genes and the two rRNA genes are flanked by tRNA genes, cleavage of the tRNAs also releases the rRNAs and the mRNA molecules from the transcript (Ojala et al., 1981). Most mRNAs lack 5' and 3' nontranslated sequences and polyA tails are added to the mRNAs posttranscriptionally (Clayton, 1984). There is a weak bi-directional transcriptional terminator (nps 3237-3249) in the tRNA-leu(UUR) gene downstream of the 16S rRNA

9

gene. Hence, an excess of rRNA transcripts from H-strand transcription is produced over mRNAs and most L-strand transcripts terminate before the rRNA genes are transcribed (Clayton, 1992). Mitochondrial protein synthesis, like bacterial protein synthesis, is initiated with formylmethionine, and is sensitive to the bacterial ribosome inhibitor chloramphenicol. However, mitochondria use a different genetic code from the universal genetic code: the opal stop codon (UGA) and UGG are used for tryptophan while the arginine codons AGA and AGG are used for stop codons. This special mtDNA genetic code limits the expression of mtDNA to within the mitochondrion because most mitochondrial mRNAs contain multiple UGA codons and can not be translated in the cytoplasmic compartment.

Mitochondrial DNA replication

I.   General rules for DNA replication

DNA replication has been well studied in many systems including both prokaryotes and eukaryotes (Kornberg and Baker, 1992). Various DNA replication mechanisms have some basic common features.

DNA replication is always a semiconservative process. This means that two progeny molecules are produced from one parental DNA molecule and each progeny molecule has one strand derived from the parental DNA duplex and one novel strand with complementary nucleotide sequence.

10

DNA replication usually starts at specific sites within the genome, called the origins of DNA replication. Initiation of DNA replication is usually the major control point for DNA replication just as the initiation from the promoter in transcription. In bacteria, a single origin is used for initiation of DNA replication. Reinitiation at the same origin can occur before the first round of replication is complete for rapid growth. In eukaryotic chromosomes, replication begins at many origins located at different positions along the chromosomal DNA. The initiation of DNA replication at each origin is controlled for once per cell cycle. Origins of replication serve to import specificity and increase the efficiency of initiation of DNA replication by providing loci for assembly of multiprotein complexes to open locally the DNA duplex for access to the template strands.

Replication fork movement is uni- or bi-directional. The initial opening of the DNA duplex at the origin allows the establishment of one or two replication forks (the growing ends of the nascent DNA strand including the replicative enzyme complex) loaded with replicative proteins. If one replication fork is established, replication proceeds from the origin in one direction (unidirectional DNA replication). If two replication forks are produced, replication proceeds in both directions with the two forks moving away from each other.

A primer is essential to start DNA synthesis by DNA polymerases. All of the known DNA polymerases are incapable of starting chains *de novo*. In most DNA replication systems, DNA replication is primed by short RNA chains. The initiator RNA is later excised and replaced with DNA.

DNA strands are elongated in the 5'-3' direction by DNA polymerases through the addition of nucleotide monomers.

Many enzymes and proteins other than DNA polymerases are required for the completion of DNA synthesis. In general, these include primases, helicases, topoisomerases, ligases, and single-stranded DNA-binding proteins (SSB).

II.   Overview of mtDNA replication

The total number of mtDNA molecules in a particular cell may vary depending on the type of the cell and the developmental stage of the cell, suggesting that mtDNA replication is somehow regulated. However, it is believed that mtDNA replication is not directly coupled to cell cycle regulation, in contrast with the cell-cycle controlled nuclear DNA replication process (Shadel and Clayton, 1997). MtDNA molecules can replicate in any and all phases of the cell cycle in mouse L cells (Bogenhagen and Clayton, 1977).

Our current understanding of mtDNA replication is based largely on early electron microscopic study of mtDNA replication intermediates from mammalian mitochondria (for reviews see (Clayton, 1982; Clayton, 1991; Clayton, 1992)). There are two initiation sites of mtDNA replication, one for each DNA strand. H-strand replication starts from the origin (O_H) in the D-loop region. The replication fork extends to one direction using the L-strand as the template. The parental H-strand is displaced when the replication fork moves. The displaced H-strand is coated by mitochondrial single-stranded DNA-binding protein (mtSSB) for protection. The initiation of the L-strand replication does not start until the H-strand replication fork passes by the initiation site of L-strand (O_L), which is two-thirds of the genome away from O_H. The L-strand replication fork moves in the opposite direction back toward the D-loop using the displaced H-strand as the template. The completion of mtDNA replication results in two semiconserved supercoiled closed circular mtDNA molecules (Figure 1.2). In sum, mtDNA replication is asynchronous in firing at the two initiation sites, asymmetric at the replication fork, and unidirectional.

III. Specialized mechanisms for initiation of mtDNA replication.

Two different mechanisms of initiation are believed to occur at the two origins in mtDNA replication (Clayton, 1982; Shadel and Clayton, 1997). The initiation of the H-strand replication starts in the D-loop region
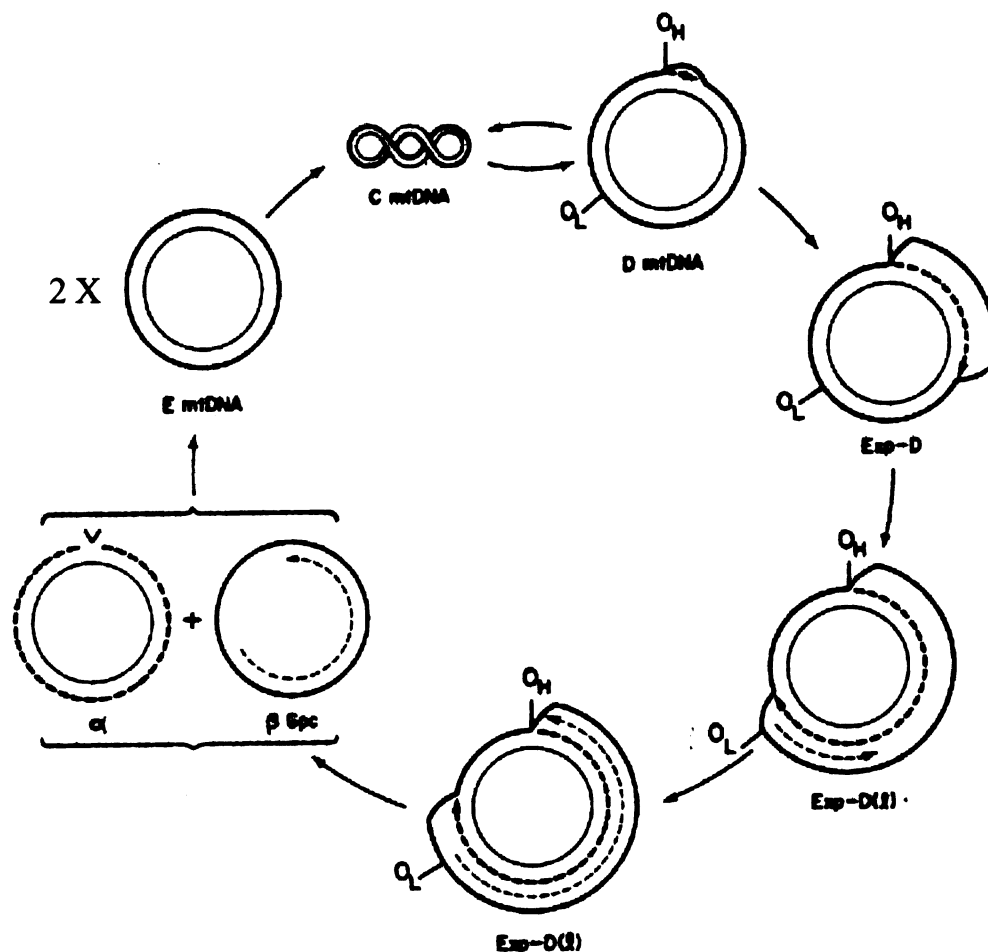
Figure 1.2   Replication model for mammalian mtDNA. Thick
solid lines: parental heavy (H) strands. Thin solid lines: parental light (L)
strands. Thick dashed lines: progeny H strands. Thin dashed lines: progeny
L strands. The order of replication is clockwise starting at D mtDNA, the
D-loop containing molecule. C mtDNA: supercoiled mtDNA molecule. O_H
and O_L : initiation sites for H- and L-strand synthesis, respectively. The
double arrows reflect the equilibrium between D mtDNA and C mtDNA.
Carat: interruption of at least one phosphodiester bond in the H strand of
the α progeny molecule. β Gpc: gapped circular progeny molecule. E
mtDNA: complete progeny molecule

14

(Chang and Clayton, 1985; Chang et al., 1985). A specialized primer is synthesized from the L-promoter by mitochondrial RNA polymerase. Transcription factors mtTFA and/or mtTFB are believed to be important for the transcription from the L-promoter (Shadel and Clayton, 1993; Jang and Jaehning, 1991). It is not clear if the transcriptional event is specific for DNA replication initiation or simply a normal L-strand transcription. In either case, the transcript is processed for mtDNA replication at specific sites in the D-loop region to produce a correct 3'-OH end for mtDNA polymerase to extend. The transcripts form a specialized structure involving the two separated DNA strands in the D-loop (Lee and Clayton, 1997). Folded structural features and three conserved sequence blocks (CBS I, II, and III) are essential for the processing of the transcripts into primers (Lee, 1996). The mitochondrial RNA processing nuclease (MRP), an enzyme similar to RNase P, is needed for this process (Chang and Clayton, 1987a; 1987b).

The involvement of processed RNA transcripts in initiation of DNA replication has been observed in other systems (Kornberg and Baker, 1992). Initiation of bacteriophage T7 DNA replication (Fuller et al., 1983) and *E. coli* plasmid Col E1 replication are two other examples of systems which utilize processed transcripts synthesized by RNA polymerase as primers for DNA synthesis. It is quite interesting that the secondary structure of the transcripts plays a role in the regulation of Col E1 DNA

replication (Kornberg and Baker, 1992). Col E1 is a 6.6-kbp, high-copy-number *E. coli* plasmid (about 20 copies per cell) containing resistance to colicin E1. Col E1 replicates unidirectionally from a unique origin and relies entirely on host-encoded proteins. The host RNA polymerase synthesizes the primer for DNA replication. The transcription of the primer (termed RNA II) is initiated at a promoter which is 555 bp upstream from the origin. RNA II spontaneously forms an RNA-DNA hybrid with the template in the origin region. RNA-DNA hybrid formation opens the duplex DNA, allowing the replication enzymes to be loaded. The formation of a specific RNA secondary structure is required for RNA-DNA hybrid formation, and the template sequence at the origin affects the efficiency with which this hybrid is utilized for initiation. RNase H, a specific enzyme for cleavage of RNA in RNA-DNA hybrids, generates the 3'-OH ends in the origin that serve as primers for DNA synthesis by *E. coli* DNA polymerase III (pol III). The proper RNA secondary structure is crucial for efficient cleavage by RNase H and for primer utilization. In addition, modulation of the RNA II secondary structure by a trans-acting factor, RNA I, regulates replication. RNA I is a 108 nt RNA molecule transcribed in the RNA II sequence region from the strand opposite that used to produce the RNA II primer. The antisense RNA inhibits primer formation by RNA II by complexing with RNA II and blocking its folding into the conformation required for RNA-DNA hybrid formation.

However, if the RNA-DNA hybrid is formed, the inhibition of RNA I is not effective any more.

The initiation of the L-strand replication of mtDNA synthesis is quite different from initiation at O_H (Clayton, 1982). The O_L sequence is approximately 30 nucleotides in size and is nested in a cluster of five tRNA genes. The origin itself has the potential to form a stem-loop structure once the region is single-stranded during H-strand replication (Hixson et al., 1986). This specialized secondary structure plays a role in the initiation of the L-strand DNA replication. Early studies revealed that a putative human mtDNA primase could recognize O_L and initiate priming and DNA synthesis (Wong and Clayton, 1985a; 1985b). Daughter L-strand synthesis is primed by RNA synthesis complementary to the T-rich loop of O_L. It has been proposed that the presumptive primase produces a short RNA primer which is then extended by mitochondrial DNA polymerase to complete the L-strand synthesis. The transition of RNA synthesis to DNA synthesis occurs near the base of the stem in O_L. However, in chicken mtDNA molecules, the lack of this O_L region in the cluster of tRNA genes implies that the tRNA genes themselves may play a role in the initiation of L-strand synthesis (Desjardins and Morais, 1990).

Drosophila as an animal model for mtDNA replication studies

Drosophila mtDNA is very similar to the human counterpart in gene organization and structure (Figure 1.3). There are no introns in *Drosophila* mtDNA, and few intergenic nucleotides. The genes found in *Drosophila* mtDNA are the same as those in human mtDNA, but there are differences in gene arrangement and in the relative proportions of the two strands of the molecule that serve as templates for transcription (for details see Figure 1.3). MtDNAs isolated from flies and *Drosophila* tissue culture cells are supercoiled circular molecules (Rubenstein et al., 1977) but contain no detectable D-loop forms (less than 1%) (Goddard and Wolstenholme, 1978; Goddard and Wolstenholme, 1980) which, in contrast, is the major form of mtDNA in vertebrates (>75%) (Robberson and Clayton, 1972). A noncoding "A+T-rich" region, so called because it is rich (90-96%) in adenine (A) and thymine (T), substitutes for the function of the D-loop region in the vertebrate mtDNA (Goddard and Wolstenholme, 1978; Goddard and Wolstenholme, 1980). This A+T-rich region shows species-specific size variation from 1.0 kbp in *D. virilis* (Goddard and Wolstenholme, 1978) to 4.8 kbp in *D. simulans* (Goddard and Wolstenholme, 1980). This region also shows great variation in nucleotide sequence compared to the rest of mitochondrial genome. The mtDNA from *D. melanogaster* is about 19.6 kbp with a 4.6 kbp A+T-rich region (Lewis
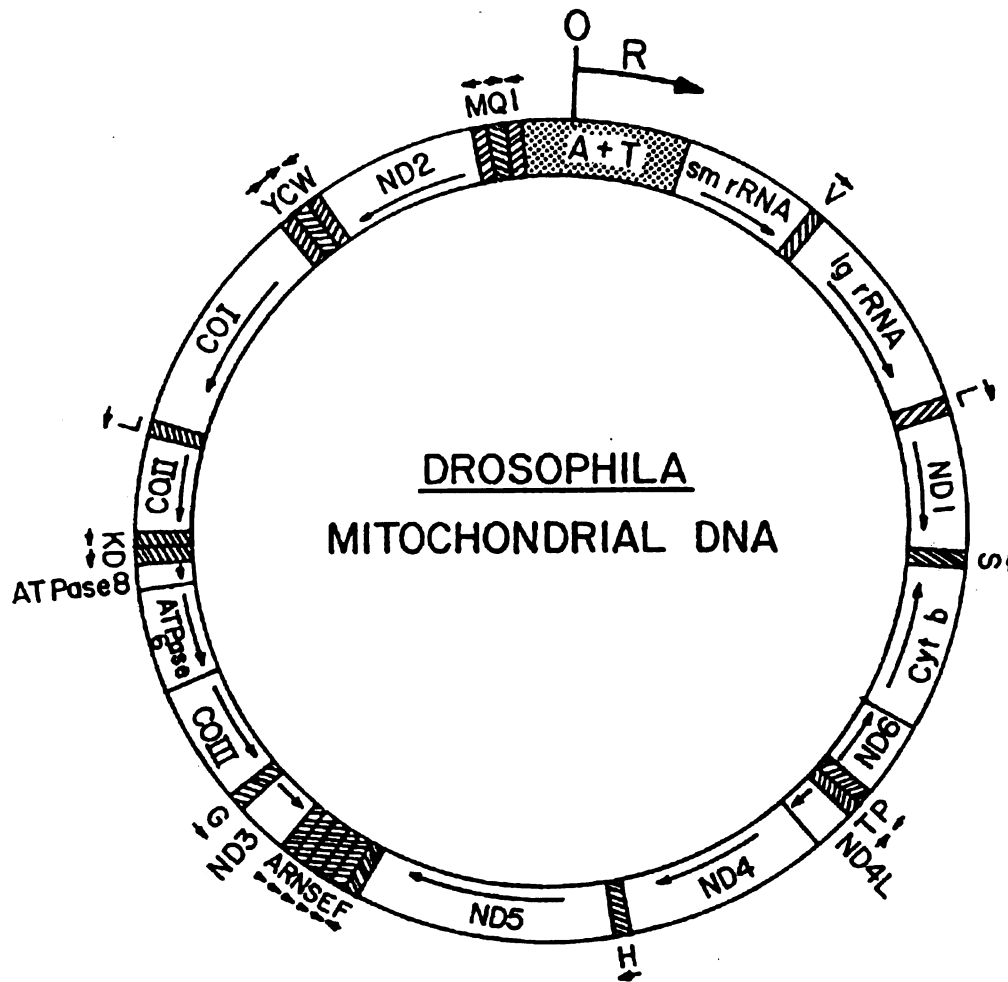
Figure 1.3 *Drosophila* **mtDNA map.** *Drosophila melanogaster* mtDNA contains 19,517 bps including a control region (A+T). The genes are named as those in human mtDNA with arrows in the coding regions indicating the direction of transcription. The direction of mtDNA leading strand replication is indicated by the arrow at O_R in the control region.

et al., 1994). Early studies including electron microscopic analysis of mtDNA replication intermediates suggested that the mechanism of *Drosophila* mtDNA replication is very similar to that of human mtDNA replication (Goddard and Wolstenholme, 1978; Goddard and Wolstenholme, 1980). The basic features of the unidirectional, asymmetric and asynchronous mechanism of mammalian mtDNA replication are well conserved in *Drosophila* mtDNA replication. *Drosophila* has been a very useful model for studies of developmental biology and other biological studies for a long time. The well-developed genetic tools in *Drosophila* will be very useful to assess in the context of the organism as a whole the results from *in vitro* biochemical studies on mtDNA replication, and to study cellular phenotypic consequences of deficiencies in mtDNA replication and genome maintenance.

Mitochondrial DNA polymerase and other DNA polymerases

Many enzymes are believed to be involved in mtDNA replication but only few have been identified. The key enzyme is mitochondrial DNA polymerase, generally called pol γ. Other known enzymes involved in mtDNA synthesis include mtRNA polymerase and primase for priming at either O_H or O_L, and mtSSB which helps to protect the displaced single strand during mtDNA synthesis. Other enzymes such as helicases, topoisomerases, ligases are also expected to contribute to mtDNA

replication based on studies of nuclear DNA replication and prokaryotic DNA replication (Kornberg and Baker, 1992).

Mitochondrial DNA polymerase is a member of a large superfamily of DNA polymerases. DNA polymerases play an essential role in nucleic acid metabolism, including DNA replication, repair and recombination. A great number of DNA polymerases have been identified from various species including eukaryotes, prokaryotes and viruses (Kornberg and Baker, 1992). These DNA polymerases show both sequence diversity and variety in subunit components. For example, the human DNA polymerase β (pol β) is a single polypeptide of 40 kDa, while *E. coli* pol III holoenzyme has at least 20 subunits consisting of 10 distinct polypeptides with sizes varying from 9 kDa to 130 kDa. However, all DNA polymerases share a basic common mechanism that is to add deoxynucleotides onto the growing end of a DNA primer strand. Some general rules apply to all DNA polymerase action (Kornberg and Baker, 1992):

- DNA polymerase can not start DNA synthesis *de novo*: a primer, usually a short RNA with a 3'-OH terminus, annealed to the template is essential;
- The added nucleotide is a monomer, activated as a 5'-triphosphate;
- The phosphodiester bond is formed by the nucleophilic attack of the 3'-OH group of the primer terminus on the α-phosphorus of the activated dNTP;

- Correct base pairing is required at the primer terminus as well as in matching of the incoming nucleotide to the template;

- Chain elongation proceeds in the 5' to 3' direction;

- Polarity of the growing chain is opposite to that of the template.

The DNA polymerase often serves as the core of a holoenzyme complex such as the α subunit in *E. coli* DNA pol III. All other subunits are added to this core subunit to facilitate its function. The complexity of a particular DNA polymerase depends largely on the role of that DNA polymerase. In general, DNA polymerases for genomic DNA replication (e.g. *E. coli* pol III) are much more complex than those for DNA repair (e.g. human pol β). Two activities are added to replicative DNA polymerases: one is the proofreading 3'-5' exonuclease activity such as the ε subunit in *E. coli* pol III to enhance the fidelity of genome duplication, and the other is a processivity factor such as the β subunit of *E. coli* pol III which enables the DNA polymerase to add many deoxynucleotides to the growing DNA strand without dissociation to ensure the completion of DNA replication within permitted time. In most cases, the 3'-5' exonuclease activity is contained in the same polypeptide with the polymerase activity but in different structural domains.

The availability of DNA polymerase sequence data in the past years allows the compilation and alignment of all the DNA polymerase protein

sequences. According to their similarities to the three *E. coli* DNA polymerases (pol I, II, and III), Ito and Braithwaite classified the DNA-dependent DNA polymerases into four families: A, B, C, and X (Ito and Braithwaite, 1991; Braithwaite and Ito, 1993). Despite the sequence diversity of DNA polymerases, there are a small number of crucial active-site residues conserved for all DNA polymerases or within an individual polymerase family (Braithwaite and Ito, 1993). Crystal structures for several DNA polymerases including that of the Klenow fragment (Ollis, 1985; Beese et al., 1993a; 1993b) of *E. coli* pol I, T7 DNA polymerase (Doublie et al., 1998), *Thermus aquaticus* DNA polymerase (Taq DNA polymerase) (Kim et al., 1995), *Bacillus stearothermophilus* DNA polymerase (Kiefer et al., 1997), and rat DNA pol β (Sawaya et al., 1994; Pelletier et al., 1994) also suggest that they have a similar overall structure, exemplified by the structure of the polymerase domain of the Klenow fragment of *E. coli* pol I. The pol domain of the Klenow fragment can be divided into three subdomains, named "palm," "fingers," and "thumb" for their anatomical analogy to a right hand. The structure has a cleft in the middle to hold the template and the primer strand. The palm subdomain contains a β-sheet that forms the base of the cleft and usually contains the catalytic residues; the fingers subdomain forms one wall of the cleft, while the thumb subdomain forms the other side of the cleft. This structural

23

similarity reflects very well the conserved polymerase activity of all DNA polymerases (Steitz, 1998).

However, different DNA polymerases are distinct enzymes: different proteins from different organisms are different; even within a particular cell, different DNA polymerases are distinguished from each other biochemically and structurally. The following sections will describe some details about some well studied DNA polymerases which will provide information on how different DNA polymerases fulfill their unique functions in a particular system. Particular attention will be devoted to mitochondrial DNA polymerase and to DNA polymerases that are similar to mitochondrial DNA polymerase.

I.   Mitochondrial DNA polymerase

Mitochondrial DNA polymerase, pol γ, usually accounts for less than 1% of total cellular DNA polymerase activity in somatic cells (Weissbach, 1979). Pol γ is distinguished from other cellular DNA polymerases by its sensitivity to N-ethylmaleimide and dideoxynucleoside triphosphates and its optimal activity under high salt conditions. Pol γ has been purified from several sources including chick embryos (Yamaguchi et al., 1980), *Drosophila melanogaster* embryos (Wernette and Kaguni, 1986), porcine liver (Mosbaugh, 1988), *Xenopus laevis* oocytes (Insdorf and Bogenhagen, 1989a; 1989b), and HeLa cells (Gray and Wong, 1992). The purified

enzyme from all these sources has a proofreading 3'-5' exonuclease activity (Kunkel and Soni, 1988; Kaguni and Olson, 1989; Kunkel and Mosbaugh, 1989; Insdorf and Bogenhagen, 1989b; Gray and Wong, 1992). Pol γ purified from *D. melanogaster* embryos is a heterodimer comprising subunits of 125 kDa and 35 kDa. Human pol γ isolated from HeLa cells consists of two subunits of 140 kDa and 54 kDa (Gray and Wong, 1992). Pol γ isolated from other sources indicated a native enzyme of 160-180 kDa (Yamaguchi et al., 1980; Insdorf and Bogenhagen, 1989a; Mosbaugh, 1988). These results together suggest that animal pol γ is a heterodimer complex of two subunits of 125-140 kDa and 35-54 kDa. The large subunit of pol γ contains both polymerase and 3'-5' exonuclease activities (Lewis et al., 1996; Longley et al., 1998).

The gene, MIP1, for the catalytic subunit of mitochondrial DNA polymerase was initially identified from *S. cerevisae* (Foury, 1989). It encodes a protein of 143.5 kDa. The cDNAs of the catalytic subunit from fission yeast (Ropp and Copeland, 1995), fly (Lewis et al., 1996), and mammals (Ropp, 1996; Lecrenier et al., 1997) have been isolated. Consistent with the biochemical data, these catalytic subunits have conserved motifs for 3'-5' exonuclease and polymerase activities identified in *E. coli* pol I. The cDNA of the small subunit of *D. melanogaster* pol γ and its mammalian homologs were isolated more recently (Wang et al., 1997). The cDNA for the small subunit pol γ of *X. laevis* is now also

available (Carrodeguas et al., 1999). However, no homologues have been identified in the complete yeast or nematode genomes.

The polymerization and exonucleolytic mechanism of pol γ from *D. melanogaster* embryos has been investigated in detail (Wernette et al., 1988; Kaguni et al., 1988; Olson and Kaguni, 1992; Williams et al., 1993; Williams and Kaguni, 1995). Recent development of reconstitution of human pol γ with recombinant proteins will definitely help biochemical characterizations of human pol γ (Lim, et al., 1999). The enzyme catalyzes DNA synthesis on a variety of DNA substrates. Interestingly, pol γ can also catalyze DNA synthesis using poly(rA)•oligo(dT) as a substrate (Wernette and Kaguni, 1986; Longley et al., 1998), a characteristic of reverse transcriptase. *Drosophila* pol γ replicates on single-stranded DNA of natural DNA sequence with the highest substrates specificity in a quasiprocessive manner, and with a high degree of nucleotide insertion fidelity. Pol γ misincorporates 1 nucleotide per 100-500 thousand nucleotides incorporated (Wernette et al., 1988), similar to the error rates of the nuclear enzymes pol ε and pol δ in the presence of proliferating cell nuclear antigen (PCNA) (see below). The intrinsic 3'-5' exonuclease activity of *Drosophila* pol γ is mispair-specific, showing about 15-fold higher efficiency of hydrolysis of 3'-terminal mispair than 3'-terminal base pairs under nonpolymerization conditions (Olson and Kaguni, 1992). Under DNA polymerizing conditions, pol γ does not extend a 3'-mispair

26

even in the presence of a large excess of the next correct nucleotide (Olson and Kaguni, 1992). Maximal activity of pol γ is obtained when assayed under moderate salt while the highest processivity is observed under low salt conditions (Williams et al., 1993). In the presence of mtSSB, pol γ exhibits maximal DNA polymerase activity and processivity. The enhancement of pol γ processivity by mtSSB is optimal at high salt conditions (Williams and Kaguni, 1995). Mitochondrial SSB stimulates 30-fold the rate of initiation of DNA strands by enhancing the primer recognition and binding by pol γ in addition to its role for clearing secondary structures of the template (Farr et al., 1999). The rate of DNA synthesis mediated by *D. melanogaster* pol γ is about 2 nucleotides per second (Wernette et al., 1988), comparable to the rate of 4.5 nucleotides per second observed *in vivo* (Clayton, 1982).

II.   *E coli* pol I and pol III

As the first DNA polymerase discovered, *E. coli* DNA polymerase I is the most abundant DNA polymerase in bacterial cells and serves roles in DNA replication, recombination, and repair (Kornberg and Baker, 1992). It is the most well studied DNA polymerase. The achievements of biochemical characterization and structural determination of *E. coli* pol I have provided tremendously helpful information for studies of other DNA polymerases. The main function of pol I is to fill short gaps and it shows

27

moderate processivity. Pol I is a single-subunit enzyme of 103 kDa. It can be separated into two fragments by partial proteolysis. The large fragment (67 kDa), termed the Klenow fragment or enzyme, contains 3'-5' exonuclease and DNA polymerase activities, while the small fragment (36 kDa) contains a 5'-3' exonuclease activity. The 5'-3' exonuclease functions both in replication and repair for the removal of the RNA primers that initiate nascent DNA synthesis and the removal of damaged bases, respectively. The overall structure of Klenow fragment determined by crystallography (Beese et al., 1993a; 1993b) showed very clearly that the 67-kDa molecule consists of two domains: a larger carboxyl (C)-terminal domain with a very prominent cleft containing the polymerase active site, and a smaller globular amino (N)-terminal domain for the 3'-5' exonuclease activity. The 3'-5' exonuclease site is located sufficiently close to the polymerase site to allow the 3' terminus of the growing chain to shuttle back and forth between them. Thus, polymerization and editing can both proceed without release of DNA from the enzyme (Joyce and Steitz, 1994).

Pol I is the prototype of DNA polymerase family A (Ito and Braithwaite, 1991). Sequence comparison has indicated that the important amino acid residues for both the 3'-5' exonuclease and polymerase active centers are conserved from pol I to pol γ enzymes (Lecrenier et al., 1997). The established structural details of the Klenow fragment of *E. coli* pol I

are a valuable asset for structure-function study on pol γ enzymes.

However, the spacer region separating the exo and pol domains is much

larger in pol γ enzymes than that in the Klenow enzyme (Lecrenier et al.,

1997). This lengthy spacer of pol γ may provide the enzyme with special

properties which are lacking in the Klenow enzyme.

The replicative pol III is the most complex holoenzyme among all

the DNA polymerases. Fortunately, tremendous progress has been achieved

in the past years to resolve the subunit composition and to understand in

great detail the interactions between subunits as well as the functions of

most subunits. Therefore, pol III provides an excellent model for studies of

other DNA polymerases. Pol III is a complex of about two copies each of

10 distinct polypeptides organized as an asymmetric dimer of about 900

kDa (McHenry, 1988; Kornberg and Baker, 1992; Kelman and O'Donnell,

1995). The core of pol III consists of three subunits (α, ε, and θ). The 130-

kDa α subunit is the polymerase (Maki et al., 1985), and the 27.5-kDa ε

subunit is a 3'-5' proofreading exonuclease (Scheuermann and Echols,

1984). The α subunit encoded by gene *polC* in *E. coli* is the prototype of

DNA polymerase family C (Ito and Braithwaite, 1991). The α and ε

subunits form a tight 1:1 complex, resulting in 2-fold increase in

polymerase but 50- to 100-fold increase in 3'-5' exonuclease (Maki and

Kornberg, 1987). The function of the 8.6-kDa θ subunit is not defined yet.

The θ subunit interacts with the ε but not α subunit (Studwell-Vaughan and

O'Donnell, 1993). Pol III holoenzyme contains two cores which act coordinately at a replication fork to replicate both strands of the duplex bacterial chromosome (Maki et al., 1988). Each core has a dimer of the β subunit as a processivity factor (Burgers et al., 1981). The structure of the β subunit reveals how it functions as the processivity factor. In the crystal structure, two β subunits appear as a ring-shaped head-to-tail dimer with a central cavity of sufficient diameter to accommodate duplex DNA (Kong et al., 1992). There is room in the cavity for 1-2 layers of water molecules between the DNA and the β subunit dimer which may insulate β from local interaction with DNA so that the β dimer can move along the DNA as a sliding clamp. The β dimer does not assemble onto DNA by itself. A clamp loader, the γ complex, is needed to load β clamps onto DNA at the cost of ATP hydrolysis (Onrust et al., 1991). The γ complex is composed of five distinct subunits in the stoichiometry $\gamma_2\delta_1\delta'_1\chi_1\psi_1$ (Maki and Kornberg, 1988). The crystal structure of the δ' subunit has also been determined (Guenther et al., 1997). The δ' subunit consists of three domains which are arranged in a C-shaped architecture. The δ' (36.9 kDa) amino acid sequence (Dong et al., 1993) shows homology to γ (47 kDa) (Flower and McHenry, 1990). The γ subunit is the catalytic component of the γ complex which has weak DNA-dependent ATPase activity and is stimulated by β clamps (Onrust et al., 1991). Nucleotides are believed to bind to γ at an interdomain interface within the inner surface of the predicted C shape

structure of the γ subunit (Guenther et al., 1997). A dimer of τ subunit (71 kDa) is believed to tether the γ complex to two $\alpha\epsilon\theta$ cores (McHenry, 1982). The γ complex and the β clamps together mediate the grip of pol III holoenzyme to primed DNA in the presence of ATP. The effects of holoenzyme assembly on the biochemical properties of the enzyme are remarkable (for review see (Kelman and O'Donnell, 1995)). The rate of DNA synthesis by the enzyme increases from 8 nucleotides per second for the α subunit to 20 nucleotides per second for the three-subunit core and to 750 nucleotides per second for pol III holoenzyme. The processivity of the enzyme increases from 11 nucleotides per association for the core to several thousand nucleotides by pol III holoenzyme.

III.   Bacteriophage T7 DNA polymerase

T7 DNA polymerase appears to be very similar to mitochondrial DNA polymerase. It is a two-subunit hybrid: the catalytic subunit is the 80-kDa polypeptide of T7 gene 5, and the accessory subunit is *E. coli* thioredoxin from the host cell. Sequence analysis has indicated that T7 g5p, like the catalytic subunit of mitochondrial DNA polymerase, belongs to DNA polymerase family A (Ito and Braithwaite, 1991). T7 g5p contains both 3'-5' exonuclease and polymerase but shows only modest polymerase activity and low processivity (1-15 nucleotides/association) (Tabor et al., 1987). After bacteriophage T7 infects a bacterial cell, T7 g5p annexes

thioredoxin from the host cell to form a 1:1 complex (K_d~5 nM) (Modrich and Richardson, 1975), which immediately converts T7 g5p into a polymerase of high processivity (thousands of nucleotides/association) (Huber et al., 1987). *E. coli* thioredoxin is a single polypeptide of 108 amino acid residues (11.7 kDa) and functions mainly as a protein disulfide oxido-reductase or a substrate for reductive enzymes (Holmgren, 1985). However, the redox capacity of thioredoxin is not required for stimulation of DNA polymerase activity but the oxidized thioredoxin (-Cys32-S-S-Cys35-) can not bind to T7 g5p well (Adler, 1979). The X-ray structure (Holmgren et al., 1975) of thioredoxin was determined and it is not a ring structure like β dimer. Therefore, there is no clamp and clamp loader involved, which is quite different from the mechanisms observed in *E. coli* pol III holoenzyme. The crystal structure of the T7 DNA polymerase-DNA complex (Figure 1.4) has been determined (Doublie et al., 1998). The structure of T7 g5p is very similar to that of the Klenow fragment of *E. coli* pol I, with an amino-terminal 3'-5' proofreading exonuclease domain and a carboxyl-terminal polymerase domain. Sequence alignment with the Klenow enzyme has shown that there is a 71-residue insertion in T7 g5p in the region between the 3'-5' exonuclease and the polymerase domain. This insertion is located in the tip of the thumb of the pol domain and forms an extended loop that binds to thioredoxin (Doublie et al., 1998). Interestingly, when the same 71-residue fragment was inserted into the
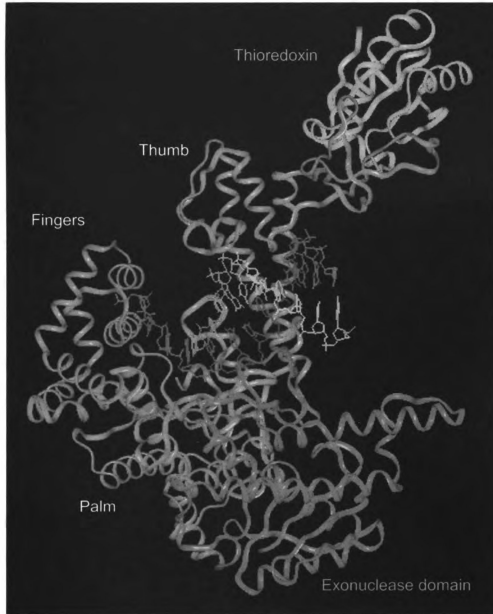
Figure 1.4 Topology and overall fold of T7 DNA polymerase.
T7 DNA polymerase complex (1t7p) bound with a primer (yellow)
and template (red) DNA is represented. The polymerase domain is
represented in magenta with fingers, thumb, and palm subdomains
labeled separately, and exonuclease domain in blue. Thioredoxin is
represented in golden yellow. Image is presented in color.

corresponding region in the Klenow enzyme by protein engineering, the mutated Klenow enzyme also interacted with thioredoxin and showed enhanced processivity (Bedford, 1997). Biochemical data indicate that thioredoxin increases by 80-fold the affinity of T7 DNA polymerase binding to the primer-template junction (Huber et al., 1987) but how this happens is not clear. In the crystal structure, thioredoxin and the extended binding loop in T7 g5p appear to be flexibly tethered to the thumb of T7 g5p, and they could presumably swing across the DNA-binding groove to encircle the primer-template exiting the polymerase. In this way, they could secure the primer-template in the polymerase, either by forming a lid over the active site that functions like a slide clamp within the polymerase, or by making electrostatic interactions with the DNA (Doublie et al., 1998).

IV.   HSV DNA polymerase

With the exception of the papovaviruses and parvoviruses, which utilize cellular DNA polymerases for their replication, all known DNA viruses that infect animal cells encode their own DNA polymerases. In general, viral DNA polymerases share considerable sequence homology with eukaryotic DNA polymerase α, δ, and ε, and they belong to DNA polymerase family B (Ito and Braithwaite, 1991). Herpes simplex virus (HSV) DNA polymerase is the most well studied viral DNA polymerase.

HSV DNA polymerase is distinguished from most cellular polymerases by its stimulation by relatively high concentrations of salt, its relatively low K_m values for dNTPs, and its sensitivity to a variety of inhibitors (Weissbach et al., 1973). HSV DNA polymerase is a heterodimer which consists of the products of the viral UL30 and UL42 genes (Hernandez and Lehman, 1990). The polypeptide encoded by UL30 is the polymerase catalytic subunit which also contains an intrinsic 3'-5' exonuclease activity and an intrinsic RNase H activity (Marcy et al., 1990; Crute and Lehman, 1989). Limited proteolysis showed that the amino-terminal half of the pol molecule is sufficient for its 3'-5' exonuclease and RNase H activities (Weisshart et al., 1994). An approximately 12-kDa carboxyl-terminal structural domain has DNA-binding activity required for polymerase activity (Digard et al., 1993; Weisshart et al., 1994). The catalytic subunit alone is much less active than the holoenzyme for polymerization on certain primer templates (Gottlieb et al., 1990). Mutational analysis indicates that the catalytic subunit can suffer deletions as large as 227 amino acid residues at the amino terminus and 58 amino acid residues at the carboxyl terminus without loss of polymerase activity *in vitro* (Dorsky and Crumpacker, 1988; Haffey et al., 1990). However, deletion of the 58 amino acid residues at the carboxyl terminus impairs association with UL42 (Digard et al., 1993). The smaller subunit encoded by UL42 is a processivity factor (Gottlieb et al., 1990). The UL42 subunit has intrinsic

affinity for double-stranded DNA (Gallo et al., 1988; Chow and Coen, 1995). The current model suggests that UL42 simultaneously binds pol and DNA and thereby increases affinity for primer template. The interaction of HSV pol and UL42 is not only required for processive DNA synthesis *in vitro*, it is also required for viral DNA replication *in vivo* (Digard et al., 1993).

V. HIV reverse transcriptase

RNA viruses contain a specialized DNA polymerase which can mediate DNA synthesis on either RNA or DNA templates. A well studied example is the human immunodeficiency virus type 1 reverse transcriptase (HIV-1 RT). HIV-1 RT consists of two subunits of 66 kDa (p66) and 51 kDa (p51) in 1:1 ratio (Joyce and Steitz, 1994). P66 has a polymerase domain at the N-terminus and an RNase H domain at the C-terminus. The two domains are connected by a flexible nonessential region. The pol domain has three subdomains similar to the fingers, palm, and thumb subdomains in the Klenow enzyme of *E. coli* pol I. P51 is derived from p66 by loss of the RNase H domain. Even though their amino acid sequences are identical, the polymerase domains are arranged differently in the two subunits, with p66 forming a large active-site cleft and p51 forming an inactive closed structure. The polymerase active site of p66 contains a catalytic triad of aspartic acid residues conserved in many DNA

polymerases (Joyce and Steitz, 1994). HIV-RT has no 3'-5' exonuclease activity for proofreading and shows a very high error rate of polymerization: a misincorporation rate *in vitro* of 1 per 2000 to 4000 nucleotides polymerized (Preston et al., 1988).

HIV-1 RT has an essential role in the viral replication and has been a major target for drug design. Antiviral nucleoside analogs such AZT and ddI are presumed to bind directly to the polymerase active site as competitive inhibitors (Kohlstaedt et al., 1992). However, long term clinical usage of these drugs have shown sever mitochondrial defects due to the inhibitory effects on pol γ. This may suggest that HIV-RT shares more similarities with pol γ than previously expected.

VI. Nuclear DNA polymerases

In eukaryotic cells, DNA polymerases are required to maintain the integrity of the genome during processes such as DNA replication, DNA repair, and DNA recombination. It has been known for a long time that there are four distinct DNA polymerases, named DNA polymerase α, β, δ, ϵ (Wang, 1991), in the nuclei of eukaryotes. The number of known nuclear DNA polymerases has increased to eight in recent years (for reviews see (Burgers, 1998; Hubscher et al., 2000)). Genetic studies of yeast have demonstrated that pols α, δ, and ϵ are all essential for cell viability and nuclear DNA replication (Campbell, 1993). In contrast, strains harboring a

deletion or mutation of POL4 (the pol β gene) are viable (Leem et al., 1994). The newly identified members (termed pols ζ (zeta), η (eta), θ (theta), and ι (iota)) are involved in various pathways for DNA repair (Baynton and Fuchs, 2000; Hubscher et al., 2000).

DNA pol α is the only enzyme with an associated DNA primase (Wang, 1991). It is required for both the initiation of DNA replication and priming for the discontinuous synthesis of Okazaki fragments on the lagging strand (Waga et al., 1994). Pol α is a moderately processive enzyme in Mg^{++}-catalyzed reactions (Copeland and Wang, 1991) and has no 3'-5' exonuclease activity for proofreading. The human pol α consists of four subunits: a catalytic subunit of 180 kDa, a 70-kDa subunit, and two subunits of 49 kDa and 58 kDa that contain the DNA primase activity. Similar subunit components are found in all eukaryotes examined (Wang, 1991). Although no enzymatic activity has been associated with p70 subunit, biochemical studies have shown that it may play a role in the assembly of the primosome (Collins et al., 1993). Both the human p180 and p70 subunits are phosphorylated predominantly during late G2 and M phases of the cell cycle (Nasheuer et al., 1991), suggesting that the phosphorylation of pol α might play a role either in coordinating DNA replication and mitosis or in resetting the replication apparatus for the next S phase.

DNA polymerase β is a single polypeptide of 39 kDa (Wang, 1991).

It is the only eukaryotic cellular DNA polymerase with determined three-

dimensional structure (Sawaya et al., 1994; Pelletier et al., 1994). Pol β is

a distributive enzyme with long gapped DNA as a substrate but can

synthesize DNA processively in short gapped substrates (Prasad et al.,

1994), in agreement with its role in base excision repair (Matsumoto et al.,

1998). Pol β consists of two domains connected by a protease-sensitive

hinge region. The amino-terminal 8-kDa domain functions in binding

single-stranded DNA (Kumar et al., 1990), recognition of a 5'-phosphate in

gapped DNA structures, and as a 5'-deoxyribose phosphate (dRP) lyase

(Prasad et al., 1998a; 1998b). The carboxyl-terminal 31-kDa domain has

approximately 5% of the polymerase activity of the intact pol β (Kumar et

al., 1990). This domain has fingers, palm, and thumb subdomains similar

to those of the Klenow fragment of *E. coli* pol I (Pelletier et al., 1994).

The 8-kDa domain is composed of two pairs of antiparallel helices and is

attached to the 31-kDa domain at the tip of the finger subdomain by a

flexible hinge. The 8-kDa domain moves closer to the palm subdomain

when pol β binds to DNA primer template in a closed conformation

(Pelletier et al., 1994). DNA pol β has no intrinsic 3'-5' exonuclease

activity. Structural data suggest that pol β may enhance fidelity by an

induced fit mechanism in which correct base pairing between template and

incoming dNTP induces alignment of catalytic groups for catalysis (via thumb closure), but incorrect base pairing will not (Sawaya et al., 1997).

DNA polymerase δ plays important roles in DNA replication, nucleotide excision repair, base excision repair and VDJ recombination (Hindges and Hubscher, 1997). Pol δ is required for both *in vivo* DNA replication and SV40 cell-free DNA replication (Waga et al., 1994). There is considerable subunit diversity between enzymes from different sources. Pol δ from mammalian cells and *D. melanogaster* consists of two subunits (Wang, 1991) while pol δ from *Schizosaccharomyces pombe* has at least four subunits (Zuo et al., 1997) and pol δ *from S. cerevisae* is a hexamer formed by two heterotrimers (Burgers and Gerik, 1998). All these enzymes have a low processivity in the absence of proliferating cell nuclear antigen. PCNA functions as a processivity factor. It exists as a trimer forming a ring structure similar to the β clamp of *E. coli* pol III even though there is no significant sequence similarity between them (Krishna et al., 1994).

DNA polymerase ε is distinguished from pol δ because of its high processivity in the absence of PCNA (Wang, 1991). The purified human pol ε consists of two subunits (Kesti et al., 1993). Mutational analyses suggest that pol ε is essential for completion of S phase (Araki et al., 1992) but the role of pol ε in DNA replication is not clear.

Overview

In sum, mitochondrial DNA polymerase appears more similar overall to bacteriophage T7 DNA polymerase and viral DNA polymerases than cellular DNA polymerases from both prokaryotes and eukaryotes. This could be the result of natural selection. The endosymbiotic origin of mitochondria in evolution parallels the infection of cells by viruses. Therefore, mitochondria may utilize enzymes for their biogenesis just as effectively as those viruses. Mitochondrial DNA polymerase is a two-subunit enzyme, and so are T7 DNA polymerase, HSV DNA polymerase, and HIV-RT, as compared to the multiple-subunit complexes like *E. coli* pol III, pol α and pol δ. The catalytic subunit of pol γ and T7 DNA polymerase belong to the same DNA polymerase family A. Biochemically, mitochondrial DNA polymerase is stimulated by high salt concentration as is HSV DNA polymerase. Mitochondrial DNA polymerase can also catalyze DNA synthesis on an RNA template, a characteristic of reverse transcriptase. However, mitochondrial DNA polymerase is the least well studied in terms of structure-function relationships. Even the subunit composition is not completely resolved in many species. Little was known about the function of the accessory subunit when this project began. Data about protein-protein interactions between the two subunits in mitochondrial DNA polymerase are completely absent.

My thesis research aims to explore the function(s) of the accessory subunit in *Drosophila* mitochondrial DNA polymerase and how it interacts with the catalytic subunit. Different strategies have been adopted to fulfill the aims. In chapter II, a computational model of the C-terminal domain of the accessory subunit is described. The structural model was developed based on the sequence homology between the C-terminal domain of the accessory subunit and the anticodon binding domain of prolyl-tRNA synthetase. According to the characteristics of this structural model, we have predicted a dual role of the accessory subunit as a processivity factor and a primer recognition factor during mtDNA replication. In chapter III, physical interactions between the two subunits of *Drosophila* mitochondrial DNA polymerase were investigated with recombinant proteins of various mutations both by *in vivo* reconstitution in Sf9 cells and by *in vitro* protein overlay assays. The results indicate that the assembly of the heterodimer of pol γ occurs through multiple contact sites between the two subunits and results in a complex with overall similarity to the structure of HIV-RT holoenzyme. Chapter IV is the brief description of some preliminary results on the biochemical properties of the accessory subunit, and outlines future plans to extend the results described in this thesis.

CHAPTER II

HOMOLOGY MODELING OF

THE CARBOXYL-TERMINAL DOMAIN OF

THE ACCESSORY SUBUNIT OF

DROSOPHILA MITOCHONDRIAL DNA POLYMERASE

The work described in this chapter was carried out in collaboration with

Paul Sanschagrin and Dr. Leslie A. Kuhn

Department of Biochemistry, Michigan State Univeristy

INTRODUCTION

Mitochondrial DNA polymerase (pol γ) is the key enzyme involved in the replication of the mtDNA genome, and its catalytic features appear to be conserved from yeast to man. Previous studies from this laboratory have shown that *Drosophila* pol γ is a heterodimer of catalytic and accessory subunits of 125 and 35 kDa, respectively (Olson et al., 1995; Wernette and Kaguni, 1986). The large catalytic subunit contains both 5'-3' DNA polymerase and 3'-5' exonuclease activities (Lewis et al., 1996). Typical of replicative DNA polymerases, the two-subunit *Drosophila* pol γ is both highly processive and highly accurate in nucleotide polymerization (Wernette et al., 1988). With the high processivity determined in *in vitro* biochemical studies, *Drosophila* pol γ could replicate completely mtDNA upon binding of the initiating primer on each DNA strand. Thus, the recruitment of pol γ to the initiating primer and maintenance of enzyme processivity may be critical for efficient and faithful mitochondrial DNA replication.

The accessory subunit of the heterodimeric pol γ from *Drosophila* embryos has been shown to be required to maintain the structural integrity and/or catalytic efficiency of the holoenzyme (Lewis et al., 1996; Wang et al., 1997). Based on amino acid sequence analysis, the catalytic subunit shares a high degree of sequence similarity with other DNA polymerase

44

catalytic cores, both from the pol γ family (Lecrenier et al., 1997; Lewis et al., 1996) and among other members of the family A class of DNA polymerases that includes the two-subunit bacteriophage T7 DNA polymerase (Braithwaite and Ito, 1993). In contrast, the cDNA for the accessory subunit encodes a novel protein (Wang et al., 1997). However, human, mouse and rat homologs of the *Drosophila* cDNA have been identified from the EST (Expressed Sequence Tag) databases, providing evidence for a common heterodimeric structure among animal pol γs (Wang et al., 1997).

In parallel with biochemical studies of pol γ and its isolated subunits, we have pursued molecular modeling of the accessory subunit to elucidate its possible functions. Although protein function is best determined experimentally, it can also be predicted by sequence alignment of a query protein with proteins of known functions (Dujon, 1996; Koonin and Mushegian, 1996; Oliver, 1996). This is possible because similar protein sequences tend to have similar functions, although exceptions occur (Orengo et al., 1994). One way to enhance sequence-based predictions of function is to determine or predict the three-dimensional (3D) structure of a protein. The 3D structure of a protein generally provides more information about its function than its sequence, because function is more directly dependent on structure than on sequence, and structure is more conserved in evolution than sequence (Chothia and Lesk, 1986). In this

case, the structure, rather than the amino acid sequence, is compared against those of the protein structure database (Protein Data Bank). If there are one or more significant structural homologs, the novel protein is predicted to have molecular properties similar to the homologs (Zarembinski et al., 1998). Recent advances in molecular and structural biology and in computer simulation have made it possible to predict 3D structure of a protein with satisfactory accuracy without experimental determination. In recent years, fold recognition or homology modeling of the 3D structure of a protein sequence has become particularly popular (Abagyan et al., 1994; Bowie et al., 1991; Bryant and Lawrence, 1993; Fischer and Eisenberg, 1996; Flockner et al., 1995; Godzik and Skolnick, 1992; Jones et al., 1992; Jones et al., 1999; Russell et al., 1996). In this regard, blind testing has shown that fold recognition methods can be very effective (Flockner et al., 1997; Shortle, 1997; Shortle, 1999).

Here we present a 3D structural model for the carboxyl-terminal domain of the accessory subunit of *Drosophila* pol γ based on the fold recognition method developed by Fischer and Eisenberg (Fischer and Eisenberg, 1996). This structural model shares a striking structural similarity with the anticodon binding domain of several class IIa aminoacyl tRNA synthetases. That this domain has been shown to be involved in tRNA binding in the latter suggests that the accessory subunit of pol γ plays a role in recognition of the unusual RNA primers generated at the

mitochondrial DNA replication origin (Lee and Clayton, 1998), to guide

the catalytic subunit to the primer terminus. In addition, the accessory

subunit model shows structural homology with thioredoxin, the accessory

subunit of T7 DNA polymerase, despite their lack of amino acid sequence

homology. Thus, we predict a dual role for the accessory subunit as a

processivity clamp and as a primer recognition factor/clamp loader in

mtDNA replication.

MATERIALS AND METHODS

Sequence Analysis of the Accessory Subunit of *Drosophila* pol γ

The SwissProt protein sequence database was searched using the *BLAST* algorithm (Altschul et al., 1990) through the NCBI web server (http://www.ncbi.nlm.nih.gov/) for homologs to the sequence of the accessory subunit of *Drosophila* pol γ (Wang et al., 1997). Sequence alignments were obtained with the *Gap* and *Bestfit* programs of the University of Wisconsin Genetics Computer Group software package (*GCG*; (Devereux et al., 1984)) using the Needleman and Wunsch algorithm (with gap weight 5 and length weight 4). Residues 254-361 of the accessory subunit of *Drosophila* pol γ (pol γ-β) were aligned using *Gap* to residues 288-386 of the sequence from the crystal structure of the *Thermus thermophilus* prolyl tRNA synthetase (Pro-RS; generously provided by S. Cusack of EMBL-Grenoble; (Cusack et al., 1998)), corresponding to an independent folding domain in *E. coli* pro-RS identified by the *BLAST* search as being similar to the pol γ-β sequence. Minor modifications of this alignment were made so that insertions and deletions fell into regions between secondary structures.

48

Structural Modeling of the Accessory Subunit

Initial prediction of the three-dimensional (3D) fold of the accessory subunit of *Drosophila* pol γ was obtained from the *Protein Fold Recognition Server* (http://www.doe-mbi.ucla.edu/people/frsvr/preds.html), based on the algorithm of Fischer and Eisenberg (Fischer and Eisenberg, 1996). This approach uses both the secondary structure assigned to the sequence by *PhDsec* ((Rost and Sander, 1994), http://www.embl-heidelberg.de/predictprotein/predictprotein.html) and the Gonnet amino acid substitution matrix to assess the similarity between the probe sequence (*e.g.*, the accessory subunit) and sequences from the Protein Data Bank (PDB; (Bernstein et al., 1977; Sussman et al., 1998)) of three-dimensional protein structures. The use of both primary and predicted secondary structures in the search enhances its sensitivity in identifying the protein with the most similar 3D fold to that of the probe. When tested on a set of 68 proteins whose 3D structures were known but ignored, the fold recognition server can recognize successfully the 3D fold of a new sequence in 70% of the cases (Fischer and Eisenberg, 1996).

Three-dimensional modeling was carried out using Molecular Simulations' *InsightII* and *Biopolymer* software on a Silicon Graphics Indigo2 graphics workstation. Using the sequence alignment between the accessory subunit of *D. melanogaster* pol γ (pol γ-β) and the 2.4 Angstrom

crystal structure of residues 288-386 of *T. thermophilus* pro-RS, the

backbone of pro-RS was used as a structural template, and side chains

differing in pol γ-β were substituted individually. For two of the three

short loops between regular secondary structures that differ between pol γ-

β and pro-RS, the *Searchloop* function in *Biopolymer* was used to identify

a loop of known structure with appropriate length, geometry, and packing

to use as a structural template; for the third loop, between β strands 4 and

5, the corresponding loop in the structure of *T. thermophilus* histidyl

tRNA synthetase (His-RS; PDB entry 1adj; (Aberg et al., 1997)) was used

as a template, followed by side-chain substitution where necessary. When

side-chain substitutions resulted in steric collisions, they were resolved by

minimal torsional angle adjustments in *InsightII*, followed by energy

minimization on the loop regions only, by 100 steps of steepest descents

minimization with the CVFF forcefield using *Discover* software

(Molecular Simulations; San Diego, CA). The stereochemistry of the final

structural model of the C-terminal domain of pol γ-β was validated by

Procheck (Laskowski et al., 1993; Morris et al., 1992), and the favorability

of amino acid environments within this structure was assessed using the 3D

Profile method (Luthy et al., 1992) via the *Verify3D* server

(http://www.doe-mbi.ucla.edu/Services/Verify3D.html). Three-

dimensional structural comparisons between the accessory subunit model

and the available crystallographic structures of thioredoxin and the δ'

subunit of pol III employed the *Dali* server ((Holm and Sander, 1993),

http://www.embl-ebi.ac.uk/dali).

RESULTS AND DISCUSSION

Identification of Sequence Homologs of the Accessory Subunit of Pol γ

The accessory subunit of *Drosophila* pol γ represents a novel protein in the protein and DNA databases (Wang et al., 1997). However, *BLAST* search in the SwissProt database suggested moderate sequence homology between a region of the accessory subunit of *Drosophila* pol γ (pol γ-β) and aminoacyl-tRNA synthetases (aaRS) from both prokaryotes and eukaryotes (Table 2.1). The top ten matches, with up to 35% identity over ~40 residues (data not shown), were all members of the class II aminoacyl-tRNA synthetase family (Cusack, 1995). Although the sequence similarities are not so great, these results suggest that the accessory subunit of pol γ may have evolved from the same ancient as those aminoacyl-tRNA synthetases, implying that they may still have some function and structural similarities. The most similar was a probable glycyl-tRNA synthetase from *Methanococcus jannaschii* (SwissProt accession code Q57681), with a score of 84 and likelihood of 0.0018 over two aligned segments, reflecting 31% identity to residues 211-254 of pol γ-β and 25% identity to residues 290-355 (data not shown). The *BLAST* sequence alignments (data not shown) between the accessory subunit of *Drosophila* pol γ and the rest aaRSs in Table 2.1 were mostly limited to the carboxy (C)-terminus (residues 290-360) of pol γ-β.

Table 2.1 BLAST SEARCH RESULTS

(Database: non-redundant SwissProt sequences)

Sequences producing High-scoring Segment Pairs:		High Score	Smallest Sum Probability P (N)	N
sp|Q57681|SYG_METJA	PROBABLE GLYCYL-TRNA SYNTHETASE...........84		0.0018	2
sp|P47493|SYG_MYCGE	GLYCYLTRNA SYNTHETASE 72		0.15	1
sp|P36431|SYP_CHLTR	PROBABLE PROLYL-TRNA SYNTHETASE...........72		0.15	1
sp|P18255|SYT1_BACSU	THREONYL-TRNA SYNTHETASE 1................72		0.15	1
sp|P70076|SYH_FUGRU	HISTIDYL-TRNA SYNTHETASE..................71		0.20	1
sp|P16659|SYP_ECOLI	PROLYL-TRNA SYNTHETASE....................70		0.26	1
sp|P49590|SYHH_HUMAN	HISTIDYL-TRNA SYNTHETASE HOMOLOG..........67		0.55	1
sp|P75000|SYP_ZYMMO	PUTATIVE PROLYL-TRNA SYNTHETASE...........66		0.67	1
sp|P10179|SYG_SHPO	PUTATIVE GLYCYL-TRNA SYNTHETASE...........64		0.88	1
sp|P12081|SYH_HUMAN	HISTIDYL-TRNA SYNTHETASE..................64		0.88	1
sp|P07178|SYH_MESAU	HISTIDYL-TRNA SYNTHETASE..................63		0.95	1
sp|Q61035|SYH_MOUSE	HISTIDYL-TRNA SYNTHETASE..................63		0.95	1

The first column indicates the database (SwissProt (sp)), accession code, and the symbol of protein connected by "_" with the source (species). The second column indicates the full name (or part of it due to space) of individual proteins. High score is the parameter used by BLAST to rank the results. N is the number of segments aligned by sequence comparison. P(N) is the smallest probability of sequence alignment of N segments.

Modeling the Accessory Subunit of pol γ Based on Structural Homology with tRNA Synthetases

To obtain insights into the structure and function of the accessory subunit of pol γ, three-dimensional fold recognition (Fischer and Eisenberg, 1996) was used to identify the best structural match for pol γ-β, based on the best alignment between its sequence and secondary structures with those of all the distinct structural folds in the Protein Data Bank. For 71% of test sequences, this fold recognition algorithm is able to identify the correct structural fold (Fischer and Eisenberg, 1996). However, the major benefit of the fold recognition server in the case of pol γ-β was the identification of a structural homolog of the pol γ-β protein that was not yet included in sequence databases and therefore not identified by *BLAST* search. Using the sequence of the accessory subunit, the server indicated 3D structural similarity for pol γ-β residues 24-359 with residues 73-504 of the glycyl-tRNA synthetase from *T. thermophilus* (PDB entry 1ati; (Logan et al., 1995)) (Figure 2.1). The Z-score was 11.5, more than twice the threshold value (4.8+/-1.0) for significant similarity (Table 2.2).

Aminoacyl-tRNA synthetases are a group of 20 enzymes that catalyze the attachment of each amino acid to its cognate tRNA (Arnez and Moras, 1997; Cusack, 1995). The 20 aaRS enzymes can be divided into two different families of 10 members each, based on their amino acid sequence and structural similarity, and crystallographic structures have been

Figure 2.1   Structure-based sequence alignment between the accessory subunit of *Drosophila* pol γ and *Thermus thermophilus* glycyl-tRNA synthetase. The sequence alignment between the accessory subunit (pol γ–β) and glycyl-tRNA synthetase (Gly-RS) was provided by fold recognition service (http://www.doe-mbi.ucla.edu/people/frsvr). Secondary structural elements are indicated by *hhhh* for alpha helices and *bbb* for beta strands. The structural elements for pol γ–β were based on PHD predication while the structural elements for Gly-RS were assigned by X-ray structural determination (PDB entry 1ati). Conserved motif II and III in class II aminoacyl-tRNA synthetases are boxed in doted line and dashed line, respectively. Conserved residues in aminoacyl-tRNA synthetases are in bold letters. The region corresponding to anticodon binding domain is shadowed for both sequences.

```
                     hhhhhhhhh hhhhhhhhhhhhhhh          bbbb        hh hhhhhhhhhhhhh
Pol γ-β  24  KLELLSHGREYAKLLQQHWTRLRPLAAHLGATREPINPVNIQRFSFPQS.QQFRNNPVQKLVKDHPRKAKCPLL   96
                     |||       ||             |
Gly-RS   73  VLHYSGHEATFADPMVDNWTPPRYFNMMFVQDLRGPRGGRGLLAYLRPETAQGGIFVNFKNVLDATSRKLGFGIAQ  214
             hhhhhhhhhbbbbbbb        bbbbb bbbbb    hhhbbbbbb  hhhhhhhhhhhhhhhh    bbbb

             :b             bbbbb  bbbbb    hhhhhhhhhhhhhhhhhhhh          bbbbb  hhh
Pol γ-β  97  IGKAFRNEITPRNFIFVREFEQMEIEYFVRPGEDEYWHRYWVEERLKWWQEM..GLSRENLVPYVQQPPESSAHYA  170
             :b
Gly-RS  215  KHQSTCSGPTSHSLFAIKGPTLHLTTDFLVEPHRALEHFYNMQRESKIWWMRLSSNPSRYRIVPCDLAEDLNPN..  288
             :bbbbb.hhhhh.bbbbbbbbbbb hhhhhhhhhhhhhhhhhhhh    hhhbbbbbb

             bbbbb          hhhhhhhhhbbb    bbbbb                a bbbbbbbbb
Pol γ-β 171  ...........DYQAIDIRTSYGDAGEVAVEQLSL...VRIVDDKD.....FRLFDARTGEVWQPTVIRSVIE  224
                                                              |||
Gly-RS 289  KATVDILYRFPHGSLELEGIAQRTDF.DLGSHTKDQEALGITARVLRNEHSTQRLAYRDPETQKWFVPYVIEPSAG  363
             bbbbbbbbb  bbbbbbbbbbb hh hhhhhh                    bbb  bbbbbbbbb

             bb  bbb        hhhhhhhhhh   b bbbbbb              hhhhhhhhhhhh  bb
Pol γ-β 225  LETTTCALLLDQC.....DHGRDSQSLLHRVLAPYQGCGIACVESDSELSADLSDLCQHLKHVLNHAGLRLSEG   293
                                       |||                     |||
Gly-RS 364  VDRGVLALLAEAPTREELPNGEERIVLKLKPQLAPIKVAVIFPLVKN..RPEITEYAKRLKARLLALGL....   429
             hhhhhhhhhhbbbbb bbbbbb hhhh  bbbbbbb    hhhhhhhhhhhhhhhh

             bbb    hhhhhhhhh     bbbbb    hhhhhhh     hh       b bbbbb    hhhhhhhh
Pol γ-β 294  DGIRTTKNASHLAEHLLEFTDMLGIPYTLVINEQTL...RNGLMQL......RSRDTRLAETIHISDVPDYLLNIFK  360
             bbbbb  hhhhhhhhhh  bbbbbbhhhhh
Gly-RS 430  GRVLYEDTGNIGKAYRRHDEVGTPPAVTVDYIDPILGQSKDGTTRLKDPVTVVRDRDTMEQIRLRVDELEGFLRERLR  504
             bbbbb  hhhhhhhhhh  bbbbbbhhhhh    bbbbbbhhhhhhhhhhhhhhhhh
```

Figure 2.1

Table 2.2 SUMMARY OF FOLD RECOGNITION RESULTS

METHOD	CONFIDENCE LEVEL	FIRST HIT	SCORE
BLAST SEQ	e-10 to e-5	SW: SYG_METJA	0.0084
FASTA SEQ	?	SW:SYG_METJA	
MOMENT	No. OF TMHS:	0	
GON+PREDSS	4.8 +/- 1.0	1atia (ttGly-RS)	11.80
GON+PREDSS+MULT	5.0 +/- 1.0	1atia	9.19
GON+PREDSS+LOCAL	6.0 +/- 1.0	1atia	14.75
GON	4.8 +/- 1.0	1atia	9.68

METHOD, method used for analysis. MOMENT, a method of transmembrane helices (TMHS) prediction. GON, a Gonnet algorithm of sequence comparison. GON+PREDSS, analysis based on sequence and predicted secondary structure by *PHD* program. MULT, multiple sequence alignment. GON+PREDSS+LOCAL, a Smith-Waterman alignment algorithm. CONFIDENCE LEVEL, a threshold value for reliable results by indicated methods. FIRST HIT, the best candidate presented by PDB code or SwissProt (SW) Symbol.

determined for members of both classes ((Cusack, 1995; Cusack et al., 1998), and refs therein). Besides Gly-RS, class II aaRSs include Ser-, Thr-, Pro-, His-, Gly-, Asp-, Asn-, Lys-, Phe-, and Ala-RS. All are homodimers, with the catalytic domain built around a six-stranded antiparallel β sheet. Most class II aaRSs have three conserved, class-defined motifs: motifs 2 and 3 are responsible for ATP binding and catalysis, whereas motif 1 is involved in dimerization essential for the enzymatic activity. Based on amino acid sequence comparison, Thr-, Pro-, His-, and most Gly-RS (including Gly-RS from *T. thermophilus*) have a homologous C-terminal domain (shadowed in Figure 2.1) attached to the catalytic core, identifying them as members of subclass IIa. This domain interacts with the anticodon loop and stem of the tRNA substrate, as elucidated by crystallography (Cusack et al., 1998).

The sequence-structure alignments from fold recognition (Figure 2.1) reveal that residues 254-360 of *Drosophila* pol γ-β map onto the anticodon binding domain of class II aaRSs. The match of residues 24-253 of pol γ-β to regions of the catalytic core of the synthetase proved more difficult to interpret, as critical motifs 2 and 3 (boxed in Figure 2.1) in the synthetase were not matched, and generally the alignment was substantially fragmented. Two β strands in the catalytic core of Gly-RS are absent from pol γ-β, which suggests that it would be difficult for the two to fold the same in this region. In addition, the motif 1 structure in Gly-RS is missing

in pol γ-β. These results indicate that pol γ-β lacks the central core structure for catalysis of tRNA aminoacylation.

While these results suggest that the accessory subunit of pol γ is not enzymatically homologous to the aaRS family, it is likely to share structural homology with the anticodon binding domain of the synthetases, based on significant sequence identity in this region. In order to select the best template for homology modeling, we compared the sequence of the C-terminal region of *Drosophila* pol γ–β with the sequences of different anticodon binding domain of class IIa aminoacyl-tRNA synthetases with known structures (data not shown). Of the known structures of aaRS's, *T. thermophilus* (tt) Pro-RS has the highest sequence identity to pol γ-β, with 28% identity (42.7% similarity) between the 81 C-terminal residues of *Drosophila* pol γ-β and residues 302-381, which form the anticodon binding domain in ttPro-RS (Cusack et al., 1998). The 3D structure of this C-terminal region of pol γ-β is therefore expected to have a closely superimposible 3D structure with that of the anticodon-binding domain. This expectation is based on exceeding the degree of sequence identity required for structural similarity found from thousands of pairwise sequence and structural alignments between known protein structures in the PDB (Sander and Schneider, 1991). Once an appropriate structural template has been identified based on sequence identity, a structural

homology model may be built by using this template as the main chain, and substituting side chains according to the sequence alignment.

We utilized the anticodon binding domain of *T. thermophilus* Pro-RS (coordinates provided by S. Cusack, EMBL-Grenoble) as the structural template to build by homology modeling the three-dimensional structure of the C-terminal domain of the accessory subunit of pol γ. Amino acid sequence comparison indicated that the C-terminal domain of the accessory subunit of *Drosophila* pol γ is also the most conserved region as compared with its human and murine homologs, supporting its functional importance (Wang, 1998). For example, the human and *Drosophila* amino acid sequences of pol γ–β have less than 20% identity over the whole length of the polypeptide but have more than 31% identity over the last 120 amino acid residues (Figure 2.2). Modeling was based on the *Bestfit* sequence alignment between pol γ-β and Pro-RS (Figure 2.3), adjusted so that deletions or insertions fell in loop regions rather than in regular secondary structures. Residues 288-301 of Pro-RS were aligned manually to pol γ-β based on the fold recognition alignment between Gly-RS and pol γ-β. The aligned region shows 23-29% pairwise sequence identity between the aaRS sequences, which is the same as the identity percentage between pol γ–β and ttPro-RS (Figure 2.3) Based on this alignment of ttPro-RS to pol γ-β, *InsightII* was used to substitute pol γ-β side chains on the ttPro-RS main-chain structure, followed by side-chain and loop optimization (as

```
man     1 MVDLGGGVHGAVFPVDALHHKPSPLLPGDSAFRLVS.AETLREILQDKEL 49
             .              .|         |.    |.|      .:||
fly     1 .......MSRIQRCFKSLASAGFFRTVEDNKLELLSHGREYAKLLQQHWT 43

man    50 SKEQLVAFLENVLKTSGKLR.ENLLHGALEHYVNCLDLVNKRLPYGLAQI 98
           |  |  |    .   . :      : :  |    . |   |
fly    44 RLRPLAAHLGATREPINPVNIQRFSFPQSQQFRNNFQKLVKDHPRKAKCP 93

man    99 GVCFHPVFDTKQIRNGVKSI.GEKTEASLVWFTPPRTSNQWLDFWLRHRL 147
           .  |        . . .|  |     .  .    |   . :        |
fly    94 TLLKHQSTCSGPTSHSLFAIKGPTLHLTTDFLVEPHRALEHFYNMQRESK 143

man   148 QWWRKFAMSPSNFSSSDCQDEEGRKGTNFTTIFPWGKELIETLWNLGDHE 197
           ||  :  .  .||  :        |     |          .:      |
fly   144 IWWMRLSSNPSRYRIVPCDLAEDLNPNDYQAIDIRTSYGDAGEVAVEQLS 193

man   198 LLHMYPGNVSKLHGRDGRKNVVPCVLSVNGDLDRGMLAYLYDSFQLTENS 247
           |..       :|    . | | |:    :|:    | | |       .|
fly   194 LVRIVDDKDFRLPDARTGETVQPTVIRSVIELETTTCALLLDGCDHGRDS 243

man   248 FTRKKNLHRKVLKLHPCLAPIKVALDVGRGPTLELRQVCQGLFNELLENG 297
             .    |||.        :|  :.    :.      .:|  .|| | .| |
fly   244 ..QSLLLHRVLAPYQCGIACVESDSEL....SADLSDLCQHLKHVLNHAG 287

man   298 ..ISVWPGYLETMQSS.LEQLYSKYDEMSILFTVLVTETTLENGLIHLRS 344
           :|   |   |   .| |:    . |   | |  :|..: | || |||. |||
fly   288 LRLSEGDGIRTTKNASHLAEHLLETDMLGIPYTLVINEQTLRNGLMQLRS 337

man   345 RDTTMKEMMHISKLKDFLIKYISSAKNV 372
           |||  : |  .|||  . |:|:      .
fly   338 RDTRLAETIHISDVPDYLLNIFKN.... 361
```

Figure 2.2   Sequence alignment between the accessory subunit of *Drosophila* pol γ (fly) and human pol γ (man). The alignment was obtained using the *Gap* program in the UW-GCG package (Devereux et al., 1984). The amino acid sequence identity and similarity over the entire length of the polypeptide are 19.7% and 26.5%, respectively. Identity and similarity for the most conserved region (underlined for the fly sequence) are 32% and 42%, respectively. Identical residues are connected by bars between the two sequences while similar residues are connected by dots. Every 10 residues are marked by dots at the top of the human sequence.

```
Dmpolγ-β  255  YQCG AC ESDSELSADLSDLCQH KHV NHAG - LSEGDGIRTTKNASHLAEHDLSTDML

TTProRS   288  IQVV VP YKD-ESRERVLEAAQG RQA LAQG - VHLDDRDQH------TPGYKFHEWELK
TTGlyRS   399  IKVA IP VK---NRPEITEYAKR KAR LALG G VLYED-TG------NIGKAYRRHDEV
TTHisRS   329  PDIY LP T-----EEAVIEAFY AEA RPR- AEYALAPR-------KPAKGLEEALKR
               |β1|              |α1|                |β2|      |α2|

Dmpolγ-β  316  IP TLV NSQTDR ---------NGLMQ RSRD RLAET HISD PDY LNIEKN*   361

TTProRS   343  VP RVE GPKDLE-------GGQAV ASRLGG-KET PLAA PEA PGKEDA          387
TTGlyRS   451  TP AVT -DYD IGQSKDGTTRLKDTVT RDRD MEQIR HVDE EGF RERLRW*    506
TTHisRS   376  AA AGF GEDEER-------AGEVT KRLADGEQVR SREE PGYA LQADG*        421
               |β3| |α3|            |β4|          |β5|       |α4|
```

■ Conserved in all   ▧ Conserved in all but one RS   □ Conserved only in RSs

Figure 2.3   **Sequence comparison of the accessory subunit (pol γ–β) of *Drosophila* pol γ with class IIa aminoacyl-tRNA synthetases.** *Drosophila* pol γ–β sequence was first aligned with *Thermus thermophilus* (TT) prolyl-tRNA synthetase (ProRS)(Cusack et al., 1998) using *Bestfit* program of the UW-GCG package (Devereux et al., 1984). Sequences of glycyl-tRNA synthetase (GlyRS) (PDB entry 1ati, (Logan et al., 1995)) and histidyl-tRNA synthetase (HisRS) (PDB entry 1adj, (Aberg et al., 1997)) were then aligned with that of TTProRS according to structural conservation (Cusack et al., 1998). Secondary structural elements are numbered and indicated as α for helices and β for beta strands.

62

described in Methods). The structural model was finally evaluated using two techniques, *Procheck*, which evaluates stereochemistry (favorable bond lengths, angles, and atomic packing), and *Verify3D*, which assesses the favorability of residue environments within the folded structure. *Procheck* found all residues to have favored or highly favored main-chain dihedral (Φ,Ψ) angles, and the model to have significantly better overall stereochemistry (G-factor of 0.07) than typical crystal structures solved at the resolution of the Pro-RS structure (2.4 Angstroems). Evaluation of residue environments by *Verify3D* gave the accessory subunit model structure a score matching those of higher resolution (2.0 Angstroem) crystal structures.

The structural model of the C-terminal domain of the accessory subunit of *Drosophila* pol γ reveals a rare α/β fold, comprising a five-stranded mixed β sheet surrounded by four α helices (Figure 2.4A), found only in the anticodon binding domains of class IIa aaRSs (Cusack et al., 1998). The main-chain of the resulting model of the C-terminal region of pol γ-β (residues 255-361) is the same as that of the anticodon binding domain of ttPro-RS (0.21 Angstroem RMSD), except for three loop insertions: loop 1 between strand β1 strand and helix α1, loop 2 between β2 and α2, and loop 3 between β4 and β5 (Figure 2.3 and Figure 2.4B). The anticodon binding pocket of the class II tRNA synthetases, which faces the viewer in the lower half of Figure 2.4B, is formed predominantly by

Figure 2.4   Homology modeling of the C-terminal domain of the accessory subunit of *Drosophila* pol γ. A) Molecular model of the C-terminal domain of the accessory subunit of *Drosophila* pol γ. The main-chain ribbon is shown along with protein side-chain atoms colored according to chemistry: Negatively-charged residues and derivatives (Asp, Asn, Glu, Gln), red; positively-charged residues (His, Lys, Arg), blue; small hydrophobic residues (Leu, Ile, Val, Met), green; large hydrophobic residues (Phe, Tyr, Trp), purple; cysteine, yellow; and other residues (Pro, Ala, Gly, Ser, Thr), grey. The termini of the α helices and β strands are numbered according to the *Drosophila* pol γ-β sequence in Figure 2.3, with the N- and C- termini labeled as (N) and (C). B) Comparison of the structures of the modeled C-terminal domain of the accessory subunit of *Drosophila* pol γ and the anticodon binding domain of *T. thermophilus* prolyl-tRNA synthetase, used as a template for the pol γ-β model. The accessory subunit is colored according to sequence conservation between the *Drosophila* and human sequences shown in Figure 2.2; residues conserved between the two sequences are in yellow with surface-accessible side chains shown, conservatively substituted residues are in green, non-conserved residues are in blue, and insertions in the fly sequence relative to the human sequence are in red. (There are no deletions in the fly sequence relative to human.). The Pro-RS main-chain ribbon is colored by sequence conservation among the three related aminoacyl-tRNA synthetases shown in Figure 2.3, with conserved residues in dark pink, similar residues in light pink, and non-conserved residues in grey.

A.
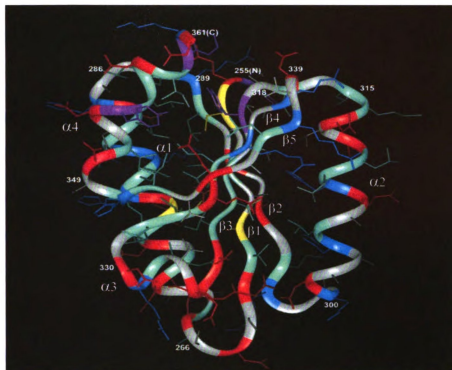
B)

Figure 2.4 (images are presented in color)

the central β sheet and by α-helices 2 and 3 (Cusack et al., 1998); this β sheet appears at center, with α helix 2 appearing as the long helix at right and helix 3 as the short helix at lower left. Whereas residue conservation is often associated with side chains that bind substrates or ligands, and thus the anticodon binding pocket might be expected to be conserved amongst aaRSs and between aaRSs and the accessory subunit, this is explicitly not the case. Different tRNA synthetases must recognize different anticodons, and the accessory subunit of pol γ must recognize at least two distinct primers; thus their ligand binding residues are expected to vary to confer this specificity. Of the residues identical in the aligned region between pol γ-β and ttPro-RS (Figure 2.3), half are surface-exposed and half are somewhat or completely buried in the protein, and do not colocalize to a specific region of the structure; for the additional residues that are chemically similar but not identical, ~2/3 are exposed and the remainder are buried (Figure 2.4). Conserved residues at the surface of this domain in ttPro-RS are not associated with an interface with the region N-terminal region to this domain, as the two domains are connected only by a loose tether in class IIa synthetases (Cusack et al., 1998). Conserved surface-exposed residues of pol γ-β (yellow side chains in Figure 2.4B) may be involved in interactions with the N-terminal domain of the accessory subunit, or with the catalytic subunit or nucleic acid substrates. An RNA-binding structural fold in the C-terminus of the accessory subunit

corresponding to that in class IIa aaRSs was unanticipated, but is consistent with a function of the accessory subunit in interacting with RNA molecules containing tRNA-like structures.

The modeled region of the accessory subunit also includes two previously noted sequence motifs: a putative zinc finger and a putative leucine zipper (Wang et al., 1997). The segment containing the leucine zipper sequence in the model structure (residues 319-347) consists of one α-helical segment and two short β strands (Figure 2.4A), and thus it can not function as a leucine zipper. The segment for the putative Zn finger motif in pol γ-β (residues 256-286) consists of a β strand, a loop, and an α helix (Figure 2.4A), which is similar to the structure observed for some known Zn fingers (Pavletich and Pabo, 1991). However, the side chains of Cys and His residues in the pol γ-β structural model are not positioned such that they could bind zinc.

Biochemical Implications of the Structural Model

The accessory subunit of pol γ as a specialized primer recognition factor

Mitochondrial DNA replication is thought to involve different priming mechanisms at the two origins. The leading DNA strand origin is located within the major non-coding region of the mtDNA genome in all animals examined to date (Clayton, 1982). In man and mouse, transcripts initiated from the L-strand promoter are processed for use as primers for

H-strand synthesis; RNAse MRP cleaves the transcripts at one of several discrete sites at O_H, approximately 100 nucleotides from the promoter (Chang and Clayton, 1987; Lee and Clayton, 1997). The 3'-termini of these processed transcripts are then presumably elongated by pol γ, but data regarding the mechanism of pol γ binding at the template-primer junction are lacking. Our molecular model suggests that the accessory subunit may play a role in the recognition of these unusual RNA-DNA hybrids, to recruit pol γ to the template-primer junction. In particular, we propose that by virtue of its conserved C-terminal domain, homologous to the anticodon binding domain in class IIa aaRSs, the accessory subunit recognizes secondary structural features in the processed RNA, thereby enhancing the template-primer binding affinity of pol γ.

A similar mode of RNA-DNA hybrid recognition by pol γ may occur at the lagging DNA strand origin. Priming of L-strand synthesis in mammals appears to be catalyzed by a primase (Wong and Clayton, 1985a) that recognizes a specific stem-loop structure formed in the displaced parental H-strand (Wong and Clayton, 1985b). Because the lagging DNA strand initiation site is nested within a cluster of five tRNA genes, the involvement of tRNA molecules cannot be excluded. Bird mtDNA genomes lack the non-coding sequence within the tRNA cluster that contains the stem-loop structure, suggesting a replication function for the tRNA genes (or a corresponding tRNA) *per se* (Desjardins and Morais,

1990). In *Drosophila*, the lagging DNA strand initiation site has been mapped within the non-coding A+T region, and in several species a stem-loop structure has been identified that may serve a replication function (Clary and Wolstenholme, 1987). Thus, although only in mammals has the stem-loop structure been documented to function in lagging DNA strand initiation, the situation may be similar in birds and in *Drosophila*. If so, the function of the accessory subunit of pol γ in specialized primer recognition is likely very similar in both leading and lagging strand initiation in mtDNA replication, and generally conserved in animals.

The postulated role of the accessory subunit of pol γ as a primer recognition factor parallels the function of the non-homologous γ complex of *E. coli* DNA polymerase III and eukaryotic replication factor C (RFC). In contrast to the unidirectional and asymmetric mode of mtDNA replication, chromosomal replication in both prokaryotes and eucaryotes proceeds bidirectionally with symmetric synthesis of leading and lagging DNA strands. This involves multiple cycles of recognition of short oligonucleotide primers by the γ complex and RFC, respectively. The γ complex and RFC are multi-subunit proteins that share both sequence and structural similarity (Guenther et al., 1997; O'Donnell et al., 1993). Our structural model of the C-terminus of the accessory subunit of pol γ shows some similarity with the crystal structure determined for the δ' subunit of the γ complex, which forms a C-shaped structure comprising three domains

(Guenther et al., 1997). While the N-terminal domain of δ' is not highly similar to the C-terminal domain of pol γ-β based on *Dali* comparison (statistical Z-score of 1.0, below the significance threshold of 2.0), they are qualitatively similar, sharing an α/β fold with a central β sheet (Figure 2.5, *left*). This is consistent with our hypothesis of the accessory subunit as a primer recognition factor, and may also indicate a general structure-function theme in primer recognition factors in various replicative systems.

The accessory subunit of pol γ as a processivity clamp

The structural and functional features of *Drosophila* pol γ are similar to those of the bacteriophage T7 DNA polymerase. The T7 DNA polymerase and pol γ catalytic cores are members of the family A group of DNA polymerases (Braithwaite and Ito, 1993). T7 DNA polymerase has a low intrinsic processivity, dissociating from the DNA after incorporation of only a few nucleotides (Huber et al., 1987). Upon infection, T7 DNA polymerase recruits a host-encoded protein, *E. coli* thioredoxin (12 kDa), to form a heterodimer that is capable of polymerizing thousands of nucleotides without dissociation (Huber et al., 1987). Likewise, we have found that the heterodimeric pol γ is a highly processive enzyme (Wernette et al., 1988). In the *E. coli* and eucaryotic replicative systems, pol III and pol δ use ring-shaped accessory subunit complexes, β (Kong et al., 1992) and PCNA (Krishna et al., 1994), respectively, to convert them into highly

Figure 2.5   Comparison of the structures of the accessory subunit of pol γ, the N-terminal domain of δ' subunit of *E. coli* DNA polymerase III , and *E. coli* thioredoxin.   The N-terminal domain of DNA polymerase III δ' (PDB entry 1a5t (Guenther et al., 1997)) is shown in blue, pol γ-β in magenta, and thioredoxin (PDB entry 1t7p, chain B (Doublie et al., 1998)) in green, with residues interacting with the catalytic subunit of T7 polymerase shown in yellow. Optimally similar orientations were produced using *Dali* (Holm and Sander, 1994; Holm and Sander, 1993).

DNA polymerase γ–β

DNA polymerase III–δ'

E. coli thioredoxin
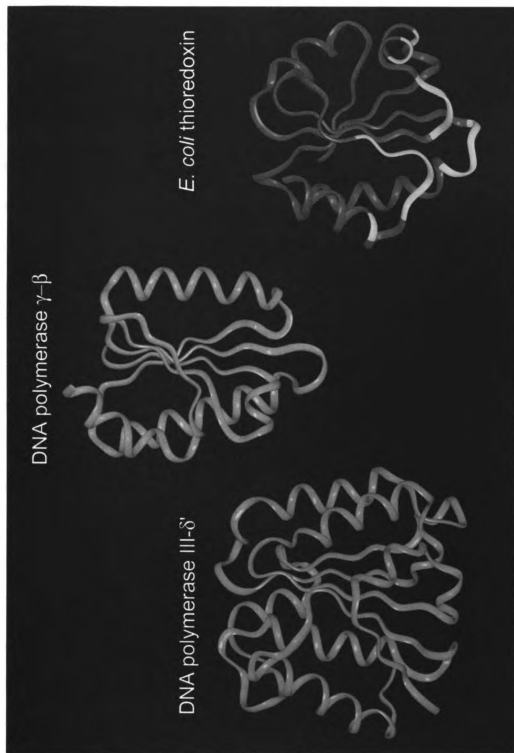
Figure 2.5 (Images are presented in color)

72

processive enzymes. These "processivity clamps" are directed to the template-primer by the γ complex and RFC, which actually serve dual roles in primer recognition and clamp loading.

The recently determined crystal structure of a T7 DNA polymerase - thioredoxin - DNA complex revealed that thioredoxin is not a ring-shaped protein (Doublie et al., 1998). Instead, thioredoxin binds to the tip of the thumb in the T7 DNA polymerase structure, and from this position presumably clamps the template-primer. Corresponding biochemical studies indicate that this association results in an ~80-fold increase in the affinity of T7 DNA polymerase for the primer terminus (Huber et al., 1987). Thus, the accessory subunit of T7 DNA polymerase apparently serves a dual function in primer binding and processivity enhancement. In fact, the mechanism of initiation at the T7 primary DNA replication origin shares similarity with that at the leading DNA strand origin in mitochondrial replication. In T7 replication, T7 RNA polymerase initiates transcription at a site upstream of the origin and proceeds through it; T7 DNA polymerase then displaces the RNA polymerase and extends the 3'-end of the transcript (Kornberg and Baker, 1992). We find that although the accessory subunit of pol γ and thioredoxin have low sequence similarity, our structural model of the C-terminus of the accessory subunit is highly similar to the structure of thioredoxin (Figure 2.6, *right*), which consists of a central core of five β strands surrounded by four helices.

While the thioredoxin (PBD entry 1t7p) and the C-terminal domain of pol γ-β do not have identical fold topologies, *Dali* structural alignment indicates significant structural similarity between them (Z-score of 2.7 over 71 residues).

Herpes simplex virus type 1 DNA polymerase is also a two-subunit enzyme (Gottlieb et al., 1990; Hernandez and Lehman, 1990). The catalytic subunit is a member of the family B DNA polymerase class (Braithwaite and Ito, 1993). Its accessory subunit, the processivity factor UL42, binds double-stranded DNA (Gallo et al., 1988). Comparative sequence analysis reveals that residues 115-174 of UL42 bear 35% identity with the anticodon binding domain (residues 329-383) of *Thermus thermophilus* ProRS (data not shown). Furthermore, UL42 residues 80-289 aligned with pol γ-β residues 105-300, and of the 45 of these residues contained in the structural model, 31% are identical. Residues 129-163 and 202-337 in UL42 are required for stimulation of DNA polymerase activity (Monahan et al., 1993), and the former and several specific residues within the latter are also required for DNA binding by UL42 (Chow and Coen, 1995). Taken together, these findings support our hypothesis of the role of the accessory subunit in pol γ function.

In sum, the C-terminal domain of the accessory subunit of *Drosophila* pol γ reveals a rare α/β fold topology, comprising a five-stranded mixed β sheet surrounded by four α helices, that is found only in

74

the anticodon binding domains of class IIa aaRSs (Cusack et al., 1998).

This structure may endow pol γ-β with the capacity to bind RNA molecules with tRNA-like structures, as in aminoacyl tRNA synthetases. This and the structural similarity with the δ' subunit of *E. coli* DNA polymerase III and particularly with thioredoxin suggest that the accessory subunit may play a dual role in loading pol γ onto the primer terminus and in enhancing its processivity. Considering the fact that all of the proteins required for its replication are the products of nuclear genes, it is not surprising that the mitochondrion has evolved to require only a single polypeptide to perform the functions provided by both the multi-subunit clamp loaders and processivity clamps in procaryotic and nuclear replicative systems. Furthermore, this conservative mechanism for a replicative cellular DNA polymerase can likely be generalized to animal pol γs, because our sequence-based structural comparisons indicate that the proposed C-terminal structure is conserved in the accessory subunits from fly to man.

CHAPTER III

SUBUNIT INTERACTIONS IN

DROSOPHILA MITOCHONDRIAL DNA POLYMERASE

INTRODUCTION

Replicative DNA polymerases are generally multiprotein complexes with a polymerase catalytic subunit associated with one or more accessory subunits that bring additional functions for effective DNA synthesis (Kornberg and Baker, 1992). It is well known that protein-protein interactions play an important role in the overall function of a DNA polymerase complex. For example, the association of *E. coli* thioredoxin with T7 DNA polymerase not only increases tremendously the processivity of the holoenzyme but also enhances significantly the overall activity (Tabor et al., 1987; Huber et al., 1987). In HIV-RT, the presence of p51 is essential for the enzymatic activity although the active sites reside in p66 (Kohlstaedt et al., 1992; Jacobo-Molina et al., 1993; Joyce and Steitz, 1994).

The heterodimer of *Drosophila* mitochondrial DNA polymerase provides an excellent model for the studies of the function and structure of a DNA polymerase complex. The large catalytic subunit of pol γ contains both DNA polymerase and 3'-5' exonuclease activities (Kaguni and Olson, 1989; Lewis et al., 1996). The molecular model presented in Chapter II suggests that the accessory subunit may have dual roles as a primer recognition factor and a processivity factor. The *Drosophila* pol γ heterodimer has been reconstituted with recombinant proteins expressed in

baculovirus-infected Sf9 cells, and exhibits biochemical properties indistinguishable by *in vitro* biochemical assays from those of the native enzyme isolated from fly embryos (Wang and Kaguni, 1999). However, the recombinant catalytic subunit alone is nearly devoid of DNA polymerase activity on gapped DNA substrates, demonstrating that the accessory subunit may be important for the overall activity of pol γ.

In this chapter, I have investigated the subunit interactions in the heterodimer of *Drosophila* pol γ. I first monitored subunit assembly *in vivo* with recombinant subunits containing various deletion mutations in Sf9 cells. The results show that the accessory subunit interacts with the catalytic subunit through multiple contact sites in different regions. Subunit interactions were investigated further by *in vitro* protein overlay assays using purified recombinant proteins expressed in *E. coli*, and proteins prepared by coupled *in vitro* transcription and translation. These results together suggest that the accessory subunit of *Drosophila* pol γ consists of three domains: an N-terminal domain (aa1-120), an intermediate (M) domain (aa128-240), and a C-terminal domain (aa254-361). Sequence analysis indicates that the M-domain shares significant sequence similarity with the RNase H domain of HIV-RT. One or more motifs in the M-domain and the C-domain interact mainly with the exo domain (aa1-430) and part of the spacer region (aa431-750) in the catalytic subunit. Deletion of the first 100 amino acid residues in the N-domain appears to enhance the

interaction between the remainder of the accessory subunit and the catalytic

subunit, suggesting that the N-domain may modulate overall subunit

interactions in *Drosophila* pol γ.

MATERIALS AND METHODS

Materials

Enzymes and Proteins - Restriction enzymes were purchased from

GIBCO-BRL and New England BioLabs. Deep Vent PCR DNA polymerase

was purchased from New England BioLabs. *Pfu* DNA polymerase was

purchased from Stratagene. Polyclonal antibodies were raised against

recombinant pol γ–α or β expressed in bacteria and were purified by the

method of Olson et al. (Olson et at., 1995). The *in vitro* transcription and

translation (TNT) kit was purchased from Promega.

Nucleotides and Nucleic Acids - Unlabeled deoxyribonucleoside

triphosphates were purchased from P-L Biochemicals. [γ-^{32}P] ATP was

purchased from ICN. Plasmid pET-11a, pET-16b, and pET-23d DNAs

were purchased from Novagen. The cDNAs of the two subunits (pol γ-α

and β) of *Drosophila* pol γ were cloned previously into pET-11a and pET-

16b vectors (Wang, 1998). Plasmid pETGA-11a and pETGA-16b are pol

γ-α expression vectors while pETGB-11a and pETGB-16b are pol γ–β

expression vectors. The pETGA-11a(NL) plasmid was used for expression

of a non-leader (the deletion of amino acid residues 2-9) pol γ-α subunit.

Baculovirus transfer vector pVL1392/1393, and linear modified

baculovirus AcMNPV DNA (BaculoGold) were purchased from

PharMingen. Wild type cDNAs pol γ–α and pol γ–β were cloned

80

separately into transfer vectors and named as pVLGA and pVLGB, respectively (Wang, 1998). Synthetic oligodeoxynucleotides as indicated in the text were synthesized in an Applied Biosystems Model 477 oligonucleotide synthesizer.

Bacterial Strains - *E. coli* XL-1 Blue (*recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac,* (*F'poAB, lacI^{q}Z M15, Tn10* (tet^r))) was used for PCR cloning. *E. coli* BL21 DE3 (*ompT, r_B M_B^-*) was used for the expression of recombinant proteins cloned in pET-11a, pET-16b, and pET-23d vectors.

Insect Cells and Tissue Culture Medium - Sf-9 (*Spodoptera frugiperda*) cells were passed from the stored stocks (Wang, 1998). TC-100 insect cell culture medium and fetal bovine serum were purchased from GIBCO-BL.

Chemicals - Isopropylthio-β-D-galactoside (IPTG), nitro blue tetrazolium (NTB), 5-bromo-4-chlorol-3-indolyl phosphate (BCIP), and tryptose broth 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. SeaKem ME low melting agarose was purchased from FMC Bioproducts. Antibiotics amphotericine, penicillin-G, streptomycin, and the protease inhibitor phenylmethylsulfonyl fluoride (PMSF) were purchased from Sigma. Ni-NTA resin was

purchased from Qiagen. [35S]-L-methionine was purchased from ICN

Pharmaceuticals.

Methods

<u>Construction of plasmids for expression of pol γ-β mutants in *E. coli*:</u>

Four pol γ-β mutants were constructed from pol γ-β cDNA in

pETGB-11a by polymerase chain reactions (PCR) using different primers

as follows:

pol γ-β(1-318):

<div align="center">

5'-cgctgga<u>catATG</u>AGTCGCATACAACG-3'
NdeI
3'-TATACGACCCGTATGG<u>Aatcctagg</u>acc-5'
BamHI

</div>

pol γ-β(1-262)

<div align="center">

5'-cgctgga<u>catATG</u>AGTCGCATACAACG-3'
NdeI
3'-CACCTTAACGGACACA<u>Catcctagg</u>atc-5'
BamH I

</div>

pol γ-β(41-355):

<div align="center">

5'-aga<u>ccATGg</u>AGCATTGGACACGACTACG-3'
NcoI
3'---GCTACATGGCCTGATAAAT<u>gagctcg</u>tg---5'
XhoI

</div>

pol γ-β(101-355):

<div align="center">

5'-aga<u>ccATGg</u>ctACTTGCTCTGGTCCCAC-3'
NcoI
3'---GCTACATGGCCTGATAAAT<u>gagctcg</u>tg---5'
XhoI

</div>

Capitalized sequences are the cDNA sequence of pol γ–β while lower case letters are sequences added as linkers for cloning. A typical PCR was carried out with 50-100 ng pETGB-11a plasmid DNA in a 50 μl reaction mixture under conditions recommended by New England BioLabs for the Deep Vent enzyme. PCR products of pol γ–β(1-318) and pol γ–β(1-262) were gel purified, digested with NdeI and BamHI and inserted into the vector pET-11a or pET-16b to produce N-terminal His-tagged fusion pol γ–β mutants. Pol γ–β(41-355) and pol γ–β(101-355) were subcloned into pET-23d at its NcoI and XhoI sites to produce C-terminal His-tagged fusion proteins. A fifth mutant pol γ–β(101-295) was prepared by digestion of pol γ–β(101-355) cDNA with NcoI and EcoRI and then inserted into pET-23 at its NcoI and HincII sites by standard recombinant DNA methods.

Bacterial overexpression and preparation of cell lysates:

Recombinant plasmid containing BL21(DE3) cells were grown at 37 °C with aeration in 1 liter of Luria broth containing 100 μg/ml ampicillin. When the optical density at 600 nm reached 0.8-1.0, IPTG was added to 1 mM, and the culture was incubated for additional four hours. Cells were harvested by centrifugation at 5000 x g for 15 minutes. Cell pellets were either stored at -80 °C or were processed directly for Ni-NTA affinity purification.

Purification of His-Tag fusion proteins with Ni-NTA affinity chromatography:

The purification followed the procedure for denaturing purification of insoluble proteins recommended by Qiagen. Briefly, cell pellets were resuspended in Buffer B (8 M urea, 0.1 M Na-phosphate, 0.01 M Tris-HCl, pH 8.0) at 5 ml per gram of pellet. Cells were lysed by stirring 1 hour at room temperature. The lysate was collected by centrifugation at 10,000 x g for 15 min at 4 °C. The lysate was loaded onto a 2 ml Ni-NTA column at a flow rate of 10-15 ml/hr. The column was then washed with 10 column volumes of Buffer B, followed by 5 column volumes of Buffer C (8 M urea, 0.1 M Na-phosphate, 0.01 M Tris-HCl, pH 6.3). The His-tag fusion proteins were eluted with 300 mM imidazole in Buffer C.

Protein overlay assays:

Recombinant pol γ–β proteins were expressed in *E. coli* and purified by Ni-NTA affinity chromatography as described in the above. Samples containing 1-2 μg of pol γ–β proteins were subjected to 12% SDS-PAGE and transferred to nitrocellulose membranes PROTRAN (Schleicher & Schuell, Keene, NH) for 4 to 6 hours at 200 mA in a TE Series Transphor Electrophoresis Unit (Hoefer Scientific Instruments, San Francisco, CA). Prestained protein standards (New England BioLabs) were used as molecular weight markers. The nitrocellulose membranes were trimmed to

minimal sizes and blocked with 0.5% BSA (Sigma) in TBST containing 2.5

mM $MgCl_2$ and 2 mM DTT for 45 to 60 minutes at room temperature

followed by a TBST wash for 5 minutes. The blots were then incubated at

4 °C overnight with ^{35}S-Met labeled protein probes as indicated in the text

in TBST containing 1 mM PMSF, 10 mM sodium metabisulfite, 2 mg/ml

leupeptin, and 2 mM DTT. The blots were washed five times with TBST

for 10-15 minutes each at room temperature, and were then dried and

scanned by Phosphor Imager(Molecular Dynamics, CA). Quantitation of

individual bands was performed with Imagequant computer software

(Molecular Dynamics, CA).

Construction of Recombinant Baculoviruses:

Baculovirus transfer vectors containing various pol γ–α and pol γ–β

mutants were prepared mainly by QuickChange mutagenesis (see Figure

3.1). There are at least three important advantages in QuickChange

mutagenesis. First, it is a one-step cloning procedure; second, it requires

only one PCR cycle per clone to avoid err accumulation; finally, it uses a

highly accurate PCR polymerase, P*fu* DNA polymerase, to ensure a very

low intrinsic PCR error rate.

Step 1
Plasmid Preparation

Gene in plasmid with
target site (⊚) for mutation

Step 2
Temperature Cycling

Denature the plasmid and anneal the
oligonucleotide primers (◢) containing
the desired mutation (×)

Using the nonstrand-displacing
action of PfuTurbo DNA polymerase,
extend and incorporate the
mutagenic primers resulting
in nicked circular strands

Step 3
Digestion

Digest the methylated, nonmutated
parental DNA template with Dpn I

Step 4
Transformation

Transform the circular, nicked dsDNA
into XL1-Blue supercompetent cells

After transformation, the XL1-Blue
supercompetent cells repair the
nicks in the mutated plasmid

(from online http://www.stratagene.com/manuals/200518.pdf)

Figure 3.1   Schematic diagram of QuickChange mutagenesis

A typical PCR was carried out in a 50 μl volume with 2 units of P*fu* polymerase in 1x P*fu* polymerase buffer (Stratagene, CA) in the presence of 200 μM of the four dNTPs and 50 ng of DNA template (pVLGA or pVLGB). A specific primer pair was used for each mutant as indicated in Table 3.1 and 3.2. The DNA template was first denatured at 95 °C for 45 seconds, followed by 20-25 three-step cycles: 95 °C for 45 seconds, 50 °C for 1 minute, and 68 °C for 2 minutes per kbp DNA template. The reaction mixture was then digested with 10 units of DpnI for 2 hours at 37 °C to eliminate the methylated parental DNA template. A 2 μl aliquot was used for transformation of *E. coli* competent cells (XL-1 blue) by electroporation using an *E. coli* Pulser (BioRad, CA). The cloned plasmid DNA was confirmed by restriction enzyme mapping and DNA sequencing. Mutant pol γ–β(1-200) was cloned by standard recombinant DNA methods into pVL1392 at restriction sites of PstI and BamHI using the PCR product of the following primer pair:

```
5'--gatctgcagATGAGTCGCATACAACGA--3'  PstI site is underlined
3'--CACCTGCTGTTTCTGAtcctaggacc--5'   BamHI site is underlined
```

Lower case letters are added sequences for purpose of cloning.

Table 3.1   **PCR PRIMERS FOR CONSTRUCTION OF POL** γ–α
MUTATIONS BY QUICKCHANGE MUTAGENESIS


```
pol γ-αΔ(483-533):
5'--GGAGATTAAGGATTCTGGA^CTAGAGGATGACGAAGAGCCG--3'
3'--CCTCTAATTCCTAAGACCT^GATCTCCTACTGCTTCTCGGC--5'


pol γ-αΔ(483-674):
5'--GGAGATTAAGGATTCTGGA^AATGGACCTTCCTTTCGAGTG--3'
3'--CCTCTAATTCCTAAGACCT^TTACCTGGAAGGAAAGCTCAC--5'


pol γ-αΔ(536-581):
5'--GGGAAGATGAAATTCTAGAG^TTTCGCAGCGATTCAGAAGG--3'
3'--CCCTTCTACTTTAAGATCTC^AAAGCGTCGCTAAGTCTTCC--5'


pol γ-αΔ(413-470):
5'--CCAACTGGGAGCGGTACATA^AAGCCTCTTCCTACAGTGG--3'
3'--GGTTGACCCTCGCCATGTAT^TTCGGAGAAGGATGTCACC--5'


pol γ-αΔ(27-126):
5'--GCAGTAGCGTGAAGATCTTT^ATTGCCAAGGAACAAGTGCAGC--3'
3'--CGTCATCGCACTTCTAGAAA^TAACGGTTCCTTGTTCACGTCG--5'


pol γ-αΔ(27-492):
5'--GCAGTAGCGTGAAGATCTTT^AAGTTTCAGCACCTCTATGAT--3'
3'--CGTCATCGCACTTCTAGAAA^TTCAAAGTCGTGGAGATACTA--5'


pol γ-αΔ(128-404):
5'--TTGAAGAGCACTTCCACAACATT^AATTCCAACTGGGAGCGG--3'
3'--AACTTCTCGTGAAGGTGTTGTAA^TTAAGGTTGACCCTCGCC--5'


pol γ-αΔ(27-742):
5'--GCAGTAGCGTGAAGATCTTT^AATGAATTCACTGGCGAAAAGTGCC--3'
3'--CGTCATCGCACTTCTAGAAA^TTACTTAAGTGACCGCTTTTCACGG--5'
```


^ indicates deletions.

Table 3.2   PCR PRIMERS FOR CONSTRUCTION OF POL γ-β
             MUTATIONS BY QUICKCHANGE MUTAGENESIS

pol γ-βΔ(213-246):
5'--CCGATGCGCGAACGGGTGAA^CTCTTACACCGCGTTCTG--3'
3'--GGCTACGCGCTTGCCCACTT^GAGAATGTGGCGCAAGAC--5'

pol γ-βΔ(37-242):
5'--CATGGAAGGGAATACGCCAAA^TCGCAATCCCTGCTCTTACAC--3'
3'--GTACCTTCCCTTATGCGGTTT^AGCGTTAGGGACGAGAATGTG--5'

pol γ-βΔ(61-104):
5'--GCCACGAGAGAACCCATAAAT^CCCACTAGCCATTCCCTATTTG--3'
3'--CGGTGCTCTCTTGGGTATTTA^GGGTGATCGGTAAGGGATAAAC--5'

pol γ-βΔ(129-160):
5'--CGGACTTCCTAGTGGAACCA^TGCGATTTGGCTGAGGATCTC--3'
3'--GCCTGAAGGATCACCTTGGT^ACGCTAAACCGACTCCTAGAG--5'

pol γ-βΔ(264-274):
5'--TGTGGAATTGCCTGTGTGGAA^CTTTGTCAGCATCTGAAACAT--3'
3'--ACACCTTAACGGACACACCTT^GAAACAGTCGTAGACTTTGTA--5'

pol γ-βΔ(327-346):
5'--ACACTTGTTATAAACGAGCAA^CACATAAGCGATGTACCGGAC--3'
3'--TGTGAACAATATTTGCTCGTT^GTGTATTCGCTACATGGCCTG--5'

pol γ-β(1-349):
        5'--CCATACACATAAGCtGATGTACCGGACT--3'
        3'--GGTATGTGTATTCGaCTACATGGCCTGA--5'

L52P:
    5'-- CCTTGGCTGCGCACCC(T)GGGCGCCACGAG--3'
    3'-- GGAACCGACGCGTGGG(A)CCCGCGGTGCTC--5'

L38VL45V:
    5'-- CCAAACTGG(T)TGCAACAGCATTGGACACGAG(C)TACGTCCC--3'
    3'-- GGTTTGACC(A)ACGTTGTCGTAACCTGTGCTC(G)ATGCAGGG--5'

^ indicates deletions, bold letters replace the letters in parenthesis, and lower case letters are insertion.

Transfer vectors containing various mutants were prepared and purified by standard CsCl sedimentation methods. Aliquots (5 µg) of pure plasmid DNA were used with BaculoGold™ DNA to prepare recombinant viruses following the procedure in the commercial manuals (Pharmingen, CA). Individual viruses were amplified from a single plaque to make recombinant viral stocks for expression and other applications.

Phosphocellulose chromatography for monitoring the assembly of pol γ–α and pol γ–β mutants in SF9 cells:

Cell extracts of ~ 6×10^7 Sf9 cells infected with a pair of pol γ–α and pol γ–β mutant containing recombinant viruses as indicated in the text were prepared as described by Wang (Wang, 1998) with some modifications. Briefly, 60 ml of Sf9 cells in two 150 cm² flasks were co-infected with recombinant viruses at a multiplicity of infection of five each at 27 °C for 45-50 hours. The cells were harvested by centrifugation of 5 minutes at 2 krpm in a Beckman TH-4 rotor at room temperature. The pellets were washed with 30 ml of cold phosphate buffer saline (pH 7.5) and recentrifuged. The pellets were resuspended on ice in 5 ml HB.28 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) and lysed by 20 strokes in a Dounce homogenizer. The homogenate was centrifuged at 3000 rpm for 7 minutes at 3 °C in an SS34

rotor. This step was repeated twice. The supernatants were combined and centrifuged at 10 krpm for 15 minutes at 3 °C in an SS34 rotor. The supernatant was centrifuged further at 100,000 x g for 30 minutes at 3 °C to obtain a soluble cell extract. After the salt concentration of the cell extract was adjusted to 60-80 mM (measured by conductivity) by addition of 1M potassium phosphate buffer (pH7.6) (about 0.6-1.0 ml), the extract was applied to a 2.5 ml phosphocellulose column packed in a 3 ml syringe in the cold room.

The phosphocellulose column was pre-equilibrated with 60 mM potassium phosphate buffer (60 mM potassium phosphate, pH7.6, 10% glycerol, 2 mM EDTA, 1 mM PMSF, 0.5 mM DTT, 10 mM Na bisulfite, 2 µg/ml leupeptin). The soluble extract was loaded at 1 column volume per hour, followed by washes with 5-10 column volumes of 60 mM potassium phosphate buffer. Proteins were eluted with 5 column volumes of 300 mM potassium phosphate buffer and collected in 0.4 ml fractions. Aliquots of various fractions were analyzed by SDS-PAGE and immunoblot analysis with pol γ subunit-specific polyclonal antibodies. The salt concentrations of the various fractions were monitored by conductivity measurement with CDM80 Conductivity Meter (Radiometer, OH).

Gel filtration chromatography:

Aliquots (0.2-0.4 ml) of the pool of peak fractions from phosphocellulose chromatography above were applied to a gel filtration column which was prepared by packing Sephacryl-S00 Superfine resins (Pharmacia, Sweden) into a glass column (~1 cm x 80 cm). The column was equilibrated with running buffer (50 mM Na phosphate, pH 7.5, 10% glycerol, 1 mM PMSF, 0.5 mM DTT, 10 mM Na bisulfite, 2 μg/ml leupeptin). The column was calibrated with a mixture of Dextran blue 2000, yeast alcohol dehydrogenase, phosphorylase b, bovine serum albumin, carbonic anhydratase, and cytochrome C (100-200 μg each). Eluted fractions were collected in volume of 0.2-0.3 ml. Aliquots of 0.2 ml were analyzed by SDS-PAGE and immunoblot analysis after precipitation with TCA in the presence of 2 μg of cytochrome C as a carrier.

RESULTS

The accessory subunit interacts through multiple sites with the catalytic subunit.

In order to identify the regions important for physical interaction between the two subunits of *Drosophila* mitochondrial DNA polymerase, we monitored complex assembly between individual mutants of the accessory subunit and the wild type catalytic subunit coexpressed in Sf9 cells by infection with recombinant viruses. An example is presented in Figure 3.2 to show the interaction between mutant pol γ-β(1-318) and pol γ-α. When an extract from Sf9 cells expressing the two recombinant proteins was applied to a phosphocellulose column, mutant pol γ-β(1-318) was coeluted with the catalytic subunit at ~200 mM KPO_4 (fraction 10 in Figure 3.2A). Gel filtration chromatography of this peak fraction confirmed the formation of a complex (fractions 19-23 in Figure 3.2B) larger than the free catalytic subunit (fractions 25-29 in Figure 3.2B). A relatively large portion of free catalytic subunit indicates a dissociation of the two proteins during chromatography, suggesting a weak interaction between mutant pol γ-β(1-318) and pol γ-α. As demonstrated in Figure 3.2C, the interaction between pol γ-β(1-318) and pol γ-α is much weaker than the interaction of pol γ-β with pol γ-α. In the presence of a 2 fold
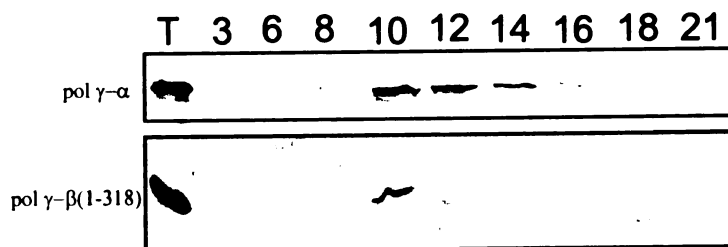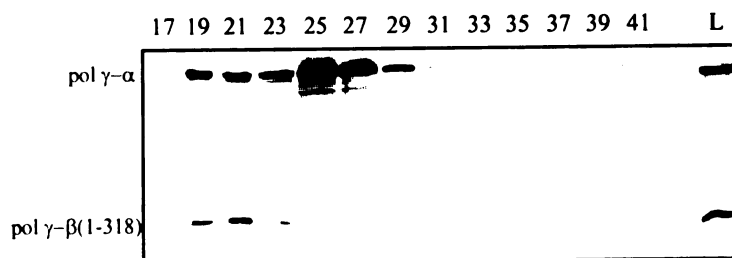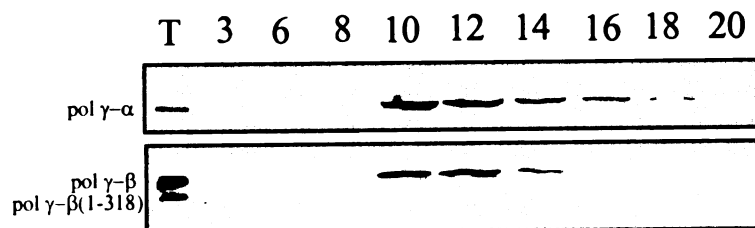
Figure 3.2. Interaction between pol γ–α and pol γ–β(1-318) in Sf9 cells. Cell extracts from about 60 millions of Sf9 cells infected with two recombinant viruses expressing pol γ–α and pol γ–β(1-318) (A and B) or three recombinant viruses expressing pol γ–α, pol γ–β, and pol γ–β(1-318) (C) were analyzed by phosphocellulose chromatography as described in Methods. The peak sample (fractions 9-11) in (A) was analyzed by S200 gel filtration chromatography (B). Aliquots from elution fractions were analyzed by 10% SDS-PAGE and immunoblot with subunit-specific polyclonal antibodies. Numbers on the top of each lane are the number of individual elution fractions. T is for total proteins from 10,000 Sf9 cells.

excess pol γ–β, pol γ–β(1-318) no longer coeluted with pol γ–α upon

phosphocellulose chromatography whereas wild type pol γ–β was coeluted

with pol γ–α at ~200 mM KPO_4, suggesting the interaction between pol γ–

α and mutant pol γ–β(1-318) is reduced to less than 30%. These results

also indicate that the elution profile upon phosphocellulose chromatography

can be used to monitor directly the interactions between an individual pol

γ–β mutant and pol γ–α when they are coexpressed in Sf9 cells. In fact, in

the absence of the catalytic subunit, neither the wild type nor any of the

mutants of the accessory subunit bound to the phosphocellulose column

(data not shown). In this way, if the expression level of various pol γ–β

mutants is kept at the same level, their relative amounts in the peak

fractions of pol γ–α coeluted from phosphocellulose columns will represent

the relative strength of a particular mutant interacting with pol γ–α, when

the level of pol γ–α in the elution fractions is also the same.

Various mutants of pol γ–β were expressed pairwise with the

catalytic subunit in Sf9 cells. In most cases, the expression level of these

recombinant protein pairs was within 2-fold variation except for the

mutant pol γ–βΔ(37-242) (lane 2 in Figure 3.3A). The interactions of these

pol γ–β mutants with pol γ–α were monitored by phosphocellulose

chromatography. The immunoblotting results (Figure 3.3B) of the peak

fractions eluted from phosphocellulose columns show various levels of the

A) Immunoblot of recombinant proteins expressed in Sf9 cells
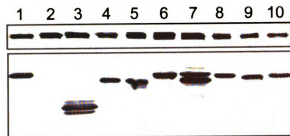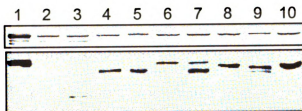
B) Immunoblot of phosphocellulose peak fractions

C)
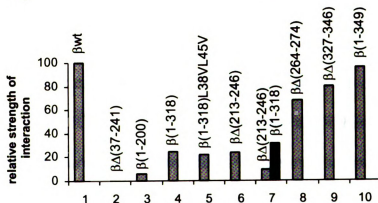
Figure 3.3   Interactions of pol γ–α with pol γ–β mutants in Sf9 cells. Soluble cytoplasmic extracts from Sf9 cells coinfected with recombinant baculoviruses expressing pol γ–α and various pol γ–β mutants as indicated were chromatographed on phosphocellulose as described under "Methods". The expression level of individual recombinant proteins was monitored by 12% SDS-PAGE and immunoblotting of total cell lysates (A). The peak fractions eluted at ~200 mM KPO_4 were pooled and subjected to 10% SDS-PAGE and immunoblot analysis (B). The relative strength of individual pol γ–β mutant interacting with pol γ–α is plotted according to the relative intensities of individual pol γ–β bands after normalization of the pol γ–α bands (C). Immunoblotting assays were conducted separately with subunit-specific polyclonal antibodies.

pol γ–β mutants in the presence of similar levels of pol γ–α, indicating

different affinities for interaction of pol γ–β mutants with pol γ–α. Mutant

pol γ–βΔ(37-242) (lane 2 in Figure 3.3B) was not observed in the elution

fractions because of its reproducibly very low level of expression, which

may suggest that subunit interaction is also important for the stability of

pol γ–β *per se*. Mutant pol γ–β(1-200) (lane 3 in Figure 3.3) shows a very

weak interaction (<10%) with pol γ–α. Other mutants exhibit interactions

from 20% to 80% as compared to wild type pol γ–β, while pol γ–β(1-349)

(lane 10 in Figure 3.3B &C) shows the strongest interaction with pol γ–α

(96%). The result that mutant pol γ–β(1-318) retains only 20-30% of wild

type interaction with pol γ–α (lanes 4 and 7 in Figure 3.3B &C) is in

agreement with the estimation of <30% by the competitive interaction assay

shown in Figure 3.2C. Results for these mutants and others tested are

summarized in Figure 3.4. Mutant βΔ(2-9), having the putative

mitochondrial target sequence interrupted, shows no interaction with the

catalytic subunit, indicating that protein sorting is important for *in vivo*

subunit interaction in mitochondrial DNA polymerase. The extremely weak

interaction between mutant pol γ–β(1-200) and the catalytic subunit

indicates that the deleted region (residues 201-361) is important for

protein-protein interaction whereas the region of residues 1-200 provides

limited subunit-subunit interaction (<10%). Furthermore, mutant pol γ–
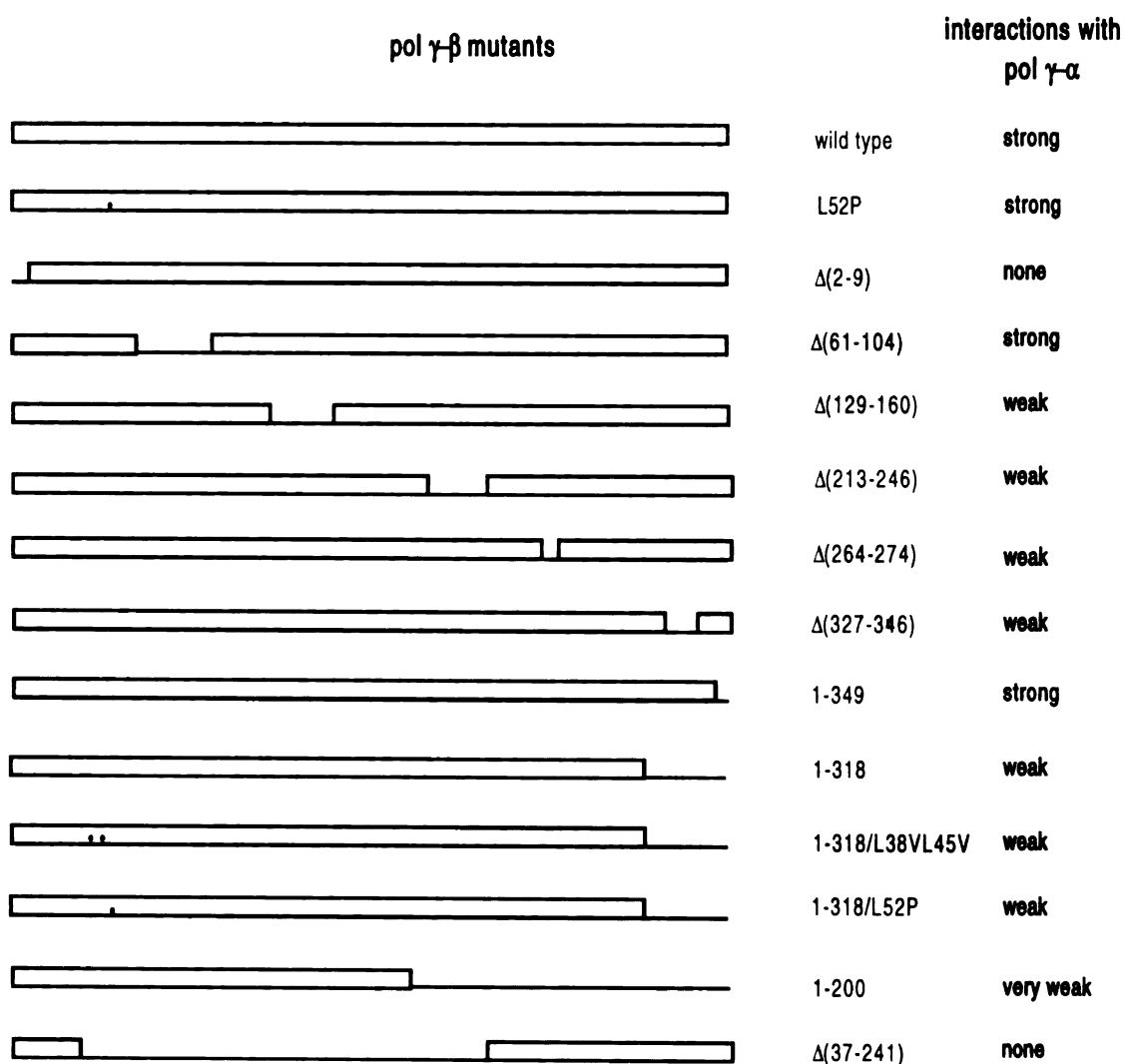
Figure 3.4   Summary of interactions between pol γ–α and pol γ–β mutants in Sf9 cells. The interactions were monitored by phosphocellulose chromatography as described in Methods. Individual pol γ–β mutants are presented in schematic boxes. Dots indicate residue substitution mutations. Lines are for deletion mutations.

βΔ(129-160) also shows reduced subunit interaction. In contrast, both mutant pol γ–βΔ(61-104) and mutant pol γ–β(1-349) retain strong interaction with pol γ–α. These results suggest that the interaction motifs are localized mainly within residues 129-349 of the accessory subunit. Within this region of 221 amino acid residues, deletions such as Δ(213-246), Δ(264-274), Δ(327-346), and Δ(319-361) (β(1-318) show various reductions in subunit interaction, indicating multiple sites are involved in subunit-subunit interaction in the *Drosophila* pol γ heterodimer.

In order to test if the putative leucine zipper (Lx6Lx6Lx6I) at the position of residues 38-59 is involved in subunit-subunit interaction, the interaction of pol γ–α with several mutants (pol γ–β/L52P, pol γ–β(1-318)/L52P, pol γ–β(1-318)/L38VL45V) with leucine residues changed into proline or valine was monitored. These amino acid residue-substitution mutations show no significant effects on protein-protein interaction (Figure 3.4 and Figure 3.3), indicating that the leucine zipper motif plays no role in subunit-subunit interaction in the *Drosophila* pol γ complex.

The accessory subunit of *Drosophila* pol γ possibly consists of three domains with different functions.

The above subunit interaction results have shown that the accessory subunit interacts with the catalytic subunit through several contact sites located largely in the region of residues 129-349. Based on these results

and the molecular modeling presented in Chapter II, I divide the accessory subunit into three domains shown in Figure 3.5. The N-terminal region (aa1-120, N-domain) in its entirety may not be important for subunit interaction because a deletion of 45 amino acid residues in the middle of this domain in mutant $\beta\Delta(61-104)$ did not affect the subunit interaction (Figure 3.4). The C-terminal region (aa254-361, C-domain) shows sequence homology (~30% identity) with the anticodon binding domain of class IIa aminoacyl-tRNA synthetases and was predicted to fold similarly to provide the accessory subunit with the property of binding to specialized RNA primers (Fan et al., 1999). The intermediate region (aa128-240) forms the M-domain. Both the C- and M-domains are involved in subunit interactions. The secondary structures predicted by the PHD program within different domains fit quite well with the three-domainal assignment based on the results of the biochemical subunit interaction studies (Figure 3.5). According to the PHD prediction, the N-domain comprises three helices (H1, H2, and H3) and three beta strands (S1, S2, and S3). The first helix H1 may function as the mitochondrial targeting signal whereas H2 and H3 possibly form a helix-loop-helix (HLH) motif which is a common DNA binding structure. The M-domain consists of two helices and four beta strands. The M-domain is connected with the C-domain by a short helix H6.
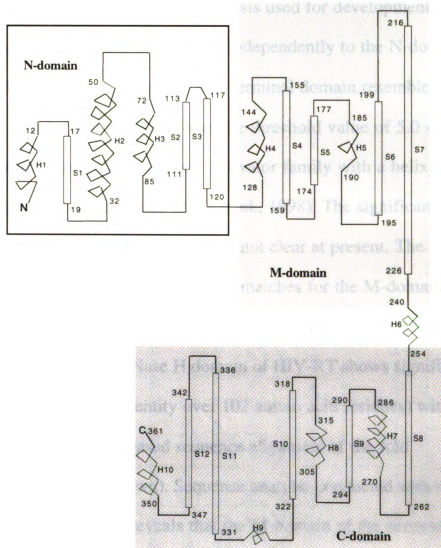
Figure 3.5   Proposed structural domains of *Drosophila* pol γ–β.
Three structural domains are proposed based on results from Figure 3.4
and previous computer-modeling (Fan, et at., 1999). The amino terminal
region of 120 amino acid residues forms the N-domain. Residues 128-240
form the M-domain in the middle of the polypeptide. The carboxyl
terminal region of residues 254-361 is the C-domain as predicted by
computer modeling. The secondary structural elements are based on results
from PHD predication. Bars represent β-strands and helices are for alpha
helices. Lines represent undefined structural elements or loops.

In order to test if the assigned domains can form independent structures, the same fold recognition analysis used for development of the C-terminal structural model was applied independently to the N-domain and the M-domain. The folding of the N-terminal domain resembles the structure of MarA (Z score of 5.6 over the threshold value of 5.0 +/- 1.0), a member of the AraC transcriptional activator family with a helix-loop-helix (HLH) DNA binding motif (Rhee et al., 1998). The significance of the HLH motif in the accessory subunit is not clear at present. The fold recognition analysis did not provide good matches for the M-domain. The top ten candidates have Z-scores close to the threshold value but not higher (Table 3.3). However, the RNase H domain of HIV-RT shows significant sequence homology (30% identity over 102 amino acid residues) with the M-domain by the structure-based sequence alignment of the fold recognition analysis (not shown). Sequence analysis conducted with the GCG program (Figure 3.6) reveals that the M-domain of the accessory subunit shares 27% amino acid sequence identity and 39% similarity with the whole RNase domain of HIV-RT. In addition, three acidic amino acid residues (shadowed in Figure 3.6) which are important for RNase H activity are also conserved or replaced by similar residues in the M-domain. This level of sequence homology between the M-domain and the RNase H domain of HIV-RT is comparable to the sequence homology conserved among known RNase H homologs. For example, a

Table 3.3 FOLD RECOGNITION RESULTS OF THE M-DOMAIN OF POL γ–β

(provided by online: http://www.doe-mbi.ucla.edu/frsv)

RANK	FOLD	CHAIN	DESCRIPTION	Z-SCORE	LENGTH	IDENTITY
1	1UMU	A	STRUCTURE DETERMINATION OF UMUD' BY MAD	4.93	90	24
2	1BGL	A	BETA-GALACTOSIDASE (CHAINS A-H)	4.15	81	27
3	1BGL	A	BETA-GALACTOSIDASE (CHAINS A-H)	4.01	83	28
4	1JVR		STRUCTURE OF THE HTLV-II MATRIX PROTEIN	3.75	84	29
5	1DUT	A	FIV DUTP PYROPHOSPHATASE	3.66	88	22
6	1HRH	A	RIBONUCLEASE H DOMAIN OF HIV REVERSE	3.58	102	30
7	1BRO		THREE DIMENSIONAL STRUCTURE OF THE N-TERM	3.52	72	13
8	1DLH	A	HLA-DR1 (DRA, DRB1 0101) HUMAN CLASS II	3.31	78	12
9	1FMT	A	METHIONYL-TRNFMET FORMYLTRANSFERASE	2.93	65	17
10	1COM	A	CHORISMATE MUTASE (E.C.5.4.99.5) COMPLEX	2.92	79	16

FOLD   Protein Data Bank codes for the coordinates of the 3D structures
CHAIN   Protein chain identifier
Z-SCORE   The Z-score is computed using the distribution of raw scores (not shown) of all
          folds. With this method, the confidence threshold is a Z-score of 5.0 +/- 1.0.
LENGTH   The number of residues from query sequence that were aligned to the fold
IDENTITY Percentage of identical residues in the alignment

structure-based sequence alignment has shown previously that *E. coli*

RNase H shares 28% identity with a yeast RNase H enzyme, 27% with

RNase H of Moloney murine leukemia virus, 19% with RNase H

of Rous sarcoma virus, and 24% with the RNase H domain of HIV-RT

(Yang, et al., 1990). The sequence homology suggests that the M-domain

may have a structure similar to that of the RNase H domain of HIV-RT.

Thus it is plausible that the accessory subunit is a multifunctional protein

with different functions for different domains.


```
                     *          *          *          *          *
RT-RH    432  EPIVGAETFYV GAANRETKLGKAGYVTNKGRQKVVP..LTNTTN....Q  475
              ||    | ||          ||.|:      .|   |  ::||   |     |       |
pol γ-β  127  EPHRALEHFY. M..QRESKIWWMRLSSNPSRYRIVPCDLAEDLNPNDYQ  173

                     *          *          *          *          *
RT-RH    476  KT LQAIY.....LALQDSGLEVNIVT SQYALGIIQAQPDKSESELVN.  519
              :.  |      .|.:    |  | ||     : |      ||     | |
pol γ-β  174  AI IRTSYGDAGEVAVEQLSL.VRIVD KDFRL......PDARTGETVQP  216

                     *          *          *          *
RT-RH    520  QIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKIL   560
              :|  .|. |  . | .        | . |   |.   :
pol γ-β  217  TVIRSVIELETTTCALLLDGCDHGRDSQ.SLLLHRVLA      253
```

Figure 3.6   Sequence homology between the M-domain of the accessory subunit (pol γ–β) of *Drosophila* mitochondrial DNA polymerase and the RNase H domain (RT-RH) of HIV reverse transcriptase. The alignment was based on the *Gap* alignment of GCG program and adjusted manually. I) indicates identical amino acid residues. :) indicates highly similar amino acid residues. .) is for similar amino acid residues. The two sequences share 26.6% identity and 38.4% similarity over the alignmed region. Every ten residues are marked on the top by (*). Three acidic residues important for RT-RH activity are shadowed. The sequences are numbered as in pol γ–β and HIV reverse transcriptase, respectively.

The interactions between residues 1-490 of the catalytic subunit and the C- and M-domains of the accessory subunit are the main force of subunit interactions in *Drosophila* pol γ
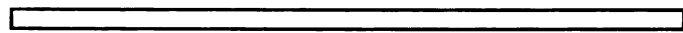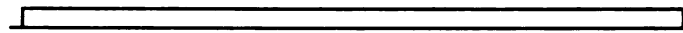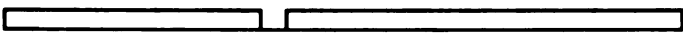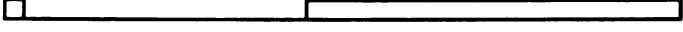
Reconstitution of altered holoenzymes with various mutants of the catalytic subunit and the wild type accessory subunit in Sf9 cells was also monitored by phosphocellulose chromatography as described in the above section. The results are summarized in Figure 3.7B. Several examples of these interactions are shown in Figure 3.7A and in lanes indicated by "W" in Figure 3.8. The relative strength of individual pol γ–α mutants interacting with the accessory subunit has not been tested directly because all of pol γ–α mutants were eluted at overlapping salt concentrations in the presence or absence of the accessory subunit during phosphocellulose chromatography (data not shown). All of the mutants including large deletions at different regions spanning the entire pol γ–α polypeptide show interaction with the accessory subunit except mutant αΔ(2-9) which has the mitochondrial target sequence deleted in Figure 3.7. This re-enforces the fact that protein sorting plays an important role in *in vivo* reconstitution of *Drosophila* mitochondrial DNA polymerase. These results argue that the catalytic subunit interacts with the accessory subunit through multiple contacts.

Based on sequence comparison with the Klenow fragment of *E. coli* DNA pol I, the catalytic subunit of *Drosophila* pol γ can be divided

105

A) Immunoblot data of elution fractions from phosphocellulose chromatography
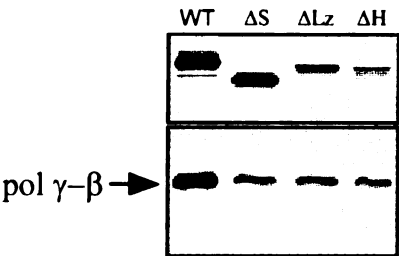
B) Summary of interactions between pol γ–β and pol γ–α mutants in Sf9 cells

Figure 3.7   Interactions of various pol γ–α mutants with pol γ–β in Sf9 cells. Wild type pol γ–β and pol γ–α mutants were expressed in pairs in Sf9 cells. The interactions were monitored by phosphocellulose chromatography as described in the text. Representative peak elution fractions are shown in (A). The results of various pol γ–α mutants interacting with pol γ–β are summarized in (B). Individual pol γ–α mutants are represented by schematic boxes, lines represent deletion mutations.

106

approximately into three domains: the exonuclease (exo) domain (aa1-430), the spacer domain (aa431-750), and the polymerase (pol) domain (aa751-1145). Several mutants with deletions in the spacer domain were investigated to test if the spacer domain is important for subunit interactions, because the corresponding part in T7 DNA polymerase provides the motif for interaction with thioredoxin. The putative leucine zipper at position 483-53 in the spacer and the residues around this putative leucine zipper are not essential for subunit interactions because mutants pol γ–$\alpha\Delta$(483-533), Δ(423-470), Δ(536-581) and Δ(483-674) show interaction with pol γ–β as shown in Figure 3.7. The results are consistent with the above result that the leucine zipper of pol γ–β is not involved in subunit interactions. Furthermore, residues 1-419 at the N-terminus of pol γ–α are sufficient to allow complex formation with pol γ–β. Interestingly, mutant pol γ–$\alpha\Delta$(27-492) also forms a complex with pol γ–β. The interactions of pol γ–β with pol γ–α(1-490) and pol γ–$\alpha\Delta$(27-492) were investigated further (see below).

From the results described in the previous section, we have located the interaction regions in the accessory subunit to the M- and C-domains. In order to understand how the domains of pol γ–α interact with the domains of pol γ–β, we have monitored by phosphocellulose chromatography the reconstitution of three mutated catalytic subunit proteins with the wild type and two mutated accessory subunit proteins

expressed in pairs in Sf9 cells. Mutant pol γ–βΔ(129-160) with a deletion

of 32 amino acid residues in the M-domain and mutant pol γ–β(1-318) with

a deletion of 43 residues in the C-domain were chosen for the tests because

these deletions likely disrupt the conformation of the individual domain

where the deletion occurs. Therefore, we can exclude the interactions by

the M-domain and monitor only the interactions by the C-domain of mutant

pol γ–βΔ(129-160) with the pol γ–α proteins. Similarly, we monitor only

the interactions mediated by the M-domain of mutant pol γ–β(1-318). As a

control, wild type pol γ–α was tested pairwise with pol γ–β (W), pol γ–

βΔ(129-160) (ΔM), and pol γ–β(1-318) (ΔC) (Figure 3.8B). As indicated

by the relative intensities of the bands for the three pol γ–β proteins, pol

γ–βΔ(129-160) and pol γ–β(1-318) show weaker interaction with pol γ–α

than pol γ–β does, about 60% and 23%, respectively (Figure 3.8C).

Similarly, the interaction between βΔ(129-161) or β(1-318) and pol γ–

α(1-490) is much weaker, only about 20% of that between pol γ–α(1-490)

and pol γ–β (Figure 3.8B & C). These results indicate that the exo domain

and part of the spacer (aa431-490) of the catalytic subunit interact with

both the M- and C-domains of the accessory subunit because subunit

interaction was reduced by 80% through disruption of either domain in the

accessory subunit. Similar results were obtained for pol γ–αΔ(483-674),

which deletes most of the spacer of the catalytic subunit. The results

suggest that the region of aa675-1145, which contains mainly the pol
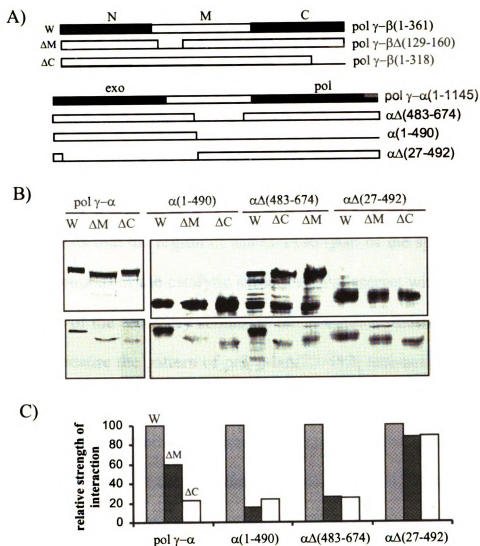
Figure 3.8   Domain interactions between *Drosophila* pol γ–α and pol γ–β coexpressed in Sf9 cells. Recombinant proteins pol γ–β (W), pol γ–βΔ(129-160) (ΔM), pol γ–β(1-318) (ΔC) were examined for interaction with pol γ–α, pol γ–α(1-490), pol γ–αΔ(483-674), and pol γ–αΔ(27-492) by pairwise coinfections in Sf9 cells. Reconstitution of heterodimeric complex was monitored by phosphocellulose chromatography. Aliquots of elution peak fractions were analyzed by immunoblotting. The bands shown by immunoblot were scanned. The relative intensities of the bands of pol γ–β proteins are converted into relative strength of protein-protein interaction after normalization to the pol γ–α bands within each group assuming 100% interaction for wild type pol γ–β. The interactions are not compared between different groups.

domain of the catalytic subunit, may not provide additional contacts with the accessory subunit. These data argue that the interactions between the C- and M-domains of pol γ–β and the region of aa1-490 of pol γ–α are the dominant force for subunit interactions. In contrast, the same mutations in the M- and C- domains did not show significant differences (about 10% reduction) for the interactions of these pol γ–β mutants with pol γ–αΔ(27-492). This suggests that the region of aa493-1146 (part of the spacer and the whole pol domain) of the catalytic subunit might interact with residues in the N-domain of the accessory subunit in the absence of the exo domain. Alternatively, because the pattern of pol γ–αΔ(27-492) interaction with these three pol γ–β proteins is different from that of the wild type pol γ–α interaction with those proteins , the deletion of residues 27-492 might change the conformation of the remainder of the catalytic subunit, and result in different interactions than those which occur in the native *Drosophila* pol γ complex. In either case, the results suggest that the interactions between the region of aa1-490 of the catalytic subunit and the C- and M-domains of the accessory subunit are the dominant force for subunit interactions in *Drosophila* pol γ.

The N-domain of the accessory subunit may modulate subunit interactions through weak interaction with the pol domain of the catalytic subunit.

The above results are based on *in vivo* reconstitution of altered holoenzymes with recombinant proteins in Sf9 cells. I have also tested *in vitro* physical interactions with purified recombinant proteins by protein overlay assays, in order to confirm the *in vivo* results and to investigate further the involvement of the N-domain of the accessory subunit in subunit interactions. As shown in Figure 3.9, different pol γ–β mutants were expressed in *E. coli* as His-tagged fusion proteins and were purified by Ni-NTA affinity chromatography. These purified recombinant proteins were then resolved by 12% SDS-PAGE and transferred to nitrocellulose membranes. After the proteins were renatured as described under "Methods", the membranes were probed with ^{35}S-labeled pol γ–α(9–1146) protein prepared by *in vitro* transcription/translation with rabbit reticulocyte lysate. After phosphor image scanning, bands detected reveal the interactions between the corresponding pol γ–β protein and the probe. The relative intensities of different bands resulting from interaction between the ^{35}S-labeled pol γ–α probe and the pol γ–β proteins on the membrane indicate the relative strength of the interactions. As shown in Figure 3.9B, the nonleader pol γ–α(9–1146) interacts with his-tagged wild

A) Proteins on the membranes

1 ●━━━━━━━━● pol γ–β(1–361)
2 ●━━━━━━━●   pol γ–β(41–355)
3 ━━━━━━●   pol γ–β(101–355)
4 ━━━━━●   pol γ–β(101–295)
5 ●━━━━━━   pol γ–β(1–318)
6 ●━━━━━   pol γ–β(1–262)
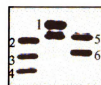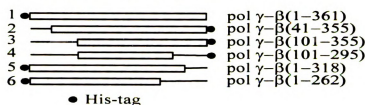
● His-tag

Immunoblot

B) Protein overlay assay
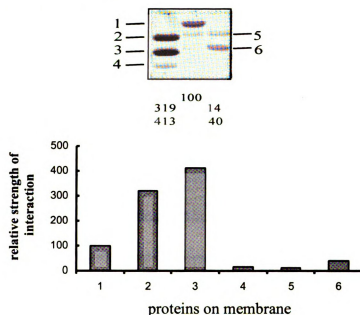
Figure 3.9 *In vitro* interactions of pol γ–α(9-1145) with pol γ–β mutants by protein overlay assay. One microgram of each purified His-tagged pol γ–β proteins was resolved by 12% SDS-PAGE and transferred to nitrocellulose membrane. The proteins were then renatured and blotted with ^{35}S-labeled pol γ–α(9-1145) prepared by *in vitro* transcription/translation as described in "Methods". The membrane was scanned by Phosphor Imager. A) schematic diagram of His-tagged pol γ–β proteins and immunoblot results. Lines represent deletions. B) results of protein overlay assay. The numbers under each lane represent the relative intensities of individual bands in the same order. Data analysis was performed with ImagQuant software. The relative strength of protein-protein interactions is plotted based on band intensity.

type pol γ–β on the membrane, indicating that the mitochondrial targeting sequence is not directly involved in subunit-subunit interaction, which might be misinterpreted by *in vivo* reconstitution-based results shown in Figure 3.4 and Figure 3.7. In agreement with *in vivo* subunit interaction results described in the above, the interaction of pol γ–α(9-1145) with pol γ–β (1-318) is only about 14% of that of pol γ–α(9-1145) with wild type pol γ–β, confirming the conclusion that the C-terminal domain is important for subunit-subunit interaction. Furthermore, the interaction between pol γ–β(101-295) and pol γ–α(9-1145) (18% in Figure 3.9B) is nearly the same as that of pol γ–β(1-318) with pol γ–α(9–1145), indicating that the N-domain (at least residues 1-100) is not involved directly in subunit interactions. The results also indicate that the M-domain (here including residues 101-295) is sufficient for limited subunit interaction (18%). This is also in agreement with the earlier conclusion that the M-domain is important for subunit interactions. Mutant pol γ–β(1-262) also shows reduced interaction (40%) with pol γ–α(9-1145); however, it is not clear why deletion of 100 residues at the C-terminus has less effect than deletion of 43 residues in mutant pol γ–β(1-318).

In contrast, deletions of the N-terminal domain appear to enhance protein-protein interaction because mutant pol γ–β(41–355) and pol γ–β(101-355) show 3 to 4 fold stronger interactions with pol γ–α(9–1145) than does the wild type pol γ–β (Figure 3.9B). This implies that the

presence of the N-domain prevents the two subunits from approaching each other closely, possibly by space exclusion with parts other than the interaction regions in pol γ–α. If this is true, the negative effects (or the enhancement by deletions) of the N-terminal domain will be less significant for smaller mutated pol γ–α probes which still retain the interaction regions. The results shown in Figure 3.10 meet this expectation for probes pol γ–α (460-1145), pol γ–α (1-360), and pol γ–α (460-743). The enhancement of Δ(1-100) in pol γ–β gradually decreases from 4-fold to 3-fold to 2-fold to 1-fold as the masses of these probes decrease. The enhancement of Δ(1-40) decreases from 3-fold to 2-fold to 1-fold in a similar way. These results imply that the N-domain of pol γ–β may play a role to modulate subunit-subunit interaction in *Drosophila* pol γ, either to keep the two subunits in the required orientations for productive protein-protein interaction or to maintain the complex in an active conformation. In addition, these three probes (pol γ–α (460-1146), pol γ–α (1-360) and pol γ–α (460-743)) also show reduced interactions with pol γ–β(1–318) and pol γ–$\beta\Delta$(1-262), once again supporting the conclusion that the C-domain of pol γ–β is important for subunit interactions.

Probe pol γ–α (742-1145) shows an entirely different pattern of interactions with these pol γ–β proteins. The deletions at the C-domain of pol γ–β show little effect on protein-protein interactions between pol γ–α (742-1145) and the mutants pol γ–β(1-262) or pol γ–β(1-318) on the

114

A) ^{35}S-labeled probes

α-exo   pol γ–α(9-360)
α-pol   pol γ–α(742-1145)
α-s/pol   pol γ–α(460-1145)
α-s   pol γ–α(460-743)

B)   Protein overlay assays

	α-s/pol		α-exo		α-s		α-pol	
	100		100		100		100	
	217	21	101	43	127	10	160	90
	255	22	158	67	130	22	86	105
	16		57		8		41	
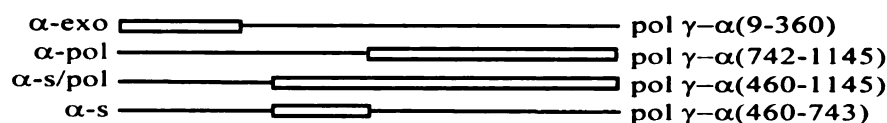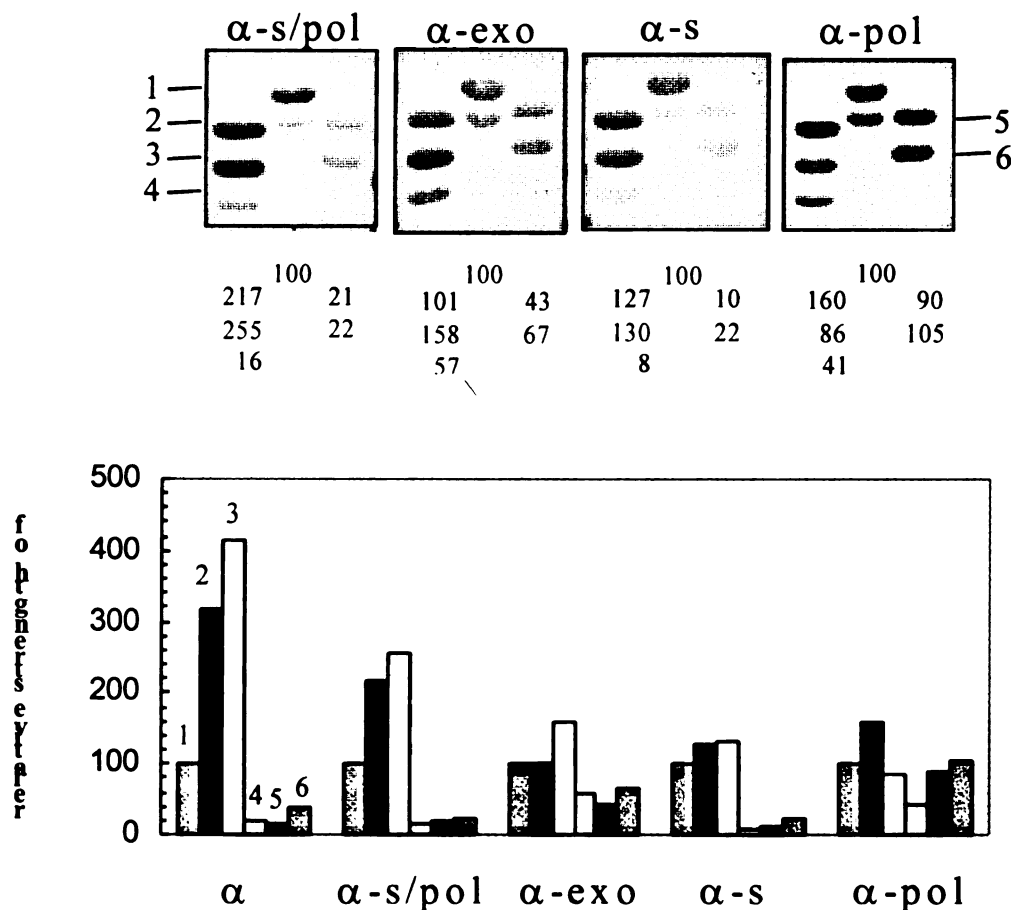
Figure 3.10   *In vitro* interactions of pol γ–α mutants with pol γ–β and mutants. Four pol γ–α mutants were used as probes to interact with His tagged pol γ–β proteins (labeled as 1-6) on nitrocellulose membrane as described in Figure-3.9. A) schematic diagram of four pol γ–α proteins. Lines represent deletions. B) results of protein overlay assays. As a control, the results of pol γ–α(9-1145)(α) is also presented in the plot.

115

membrane, indicating that there is little physical interaction between the C-domain of pol γ–β and the pol domain of pol γ–α, in agreement with the *in vivo* results obtained for pol γ–αΔ(27-492) and pol γ–β(1-318) in Sf9 cells. Pol γ–α(742-1145) shows weak interaction with residues 41-100 at the N-terminus of pol γ–β, because its interaction with pol γ–β(41-355) drops from 160% to 86% for interaction with pol γ–β(101-355). Pol γ–α(742-1145) may also interact weakly with pol γ–β(101-295) (41% in Figure 3.10B). However, the different pattern of pol γ–α(742-1145) interactions with the pol γ–β mutants as compared to pol γ–α(9-1145) might again indicate that the deletion of residues 1-742 may change the conformation of the rest of the polypeptide, and result in interactions different from those observed in the native *Drosophila* pol γ.

DISCUSSION

The results presented here demonstrate the complexity of subunit interactions in *Drosophila* mitochondrial DNA polymerase. The physical interaction between the two subunits involves multiple contact sites, which may explain the biochemical finding that subunit association in the native pol γ is so tight that separation can not be achieved without denaturation of the enzyme (Olson et al., 1995). These physical interactions are likely required to keep the catalytic subunit in an active conformation, because the recombinant catalytic subunit purified either from *E. coli* or from baculovirus-infected Sf9 cells exhibits very low catalytic activity, whereas the recombinant holoenzyme complex isolated from baculovirus-infected Sf9 cells is indistinguishable from the native pol γ (Wang, 1998).

The current studies on subunit interactions have provided some insight about the structure of the accessory subunit, which is emerging as one of the hot topics in this area of research. The results suggest that the accessory subunit can be divided into three domains: an N-domain of residues 1-120, an M-domain of residues 128-240, and a C-domain of residues 254-361. Both the M- and C-domains are important for interactions with the exo domain and part of the spacer of the catalytic subunit. The N-terminal domain is not involved in direct interactions with these regions in pol γ–α but its presence may play a role in coordinating

the major interactions of the C- and M-domains of pol γ–β with pol γ–α. The N-domain may interact with the pol domain of pol γ–α in a way different from the interactions of the C- and M-domains with the exo domain and spacer in pol γ–α. The latter likely contributes the dominant force in subunit interactions in the pol γ complex, whereas the former may hold the two subunits in the required orientation or conformation for formation of an active enzyme complex. These hypotheses are in agreement with preliminary data from this laboratory that a complex formed by pol γ–α and pol γ–βΔ(61-104) in Sf9 cells has virtually no DNA polymerase activity. Other mutants described in this study with deletions in the M- and C-domains of pol γ–β show weak interaction with pol γ–α, and can not be purified as functional complexes with pol γ–α for further biochemical characterization.

It was not a surprise to us that the C-terminal domain of the accessory subunit contributes to the subunit association. Previously, we used computational modeling to predict that the C-domain likely folds similarly to the anticodon domain of class IIa amino acyl-tRNA synthetases (Fan et al., 1999). In class IIa aminoacyl-tRNA synthetases, the anticodon binding domain is also involved in the dimerization which is essential in most cases for the activity of class IIa aminoacyl-tRNA synthetases (for reviews see (Arnez and Morase, 1997; Cusack, 1995)). Several deletions in the C-domain as well as in the M-domain reduce the subunit interaction,

suggesting that each domain has multiple regions interacting with pol γ–α, or that the conformation of the individual domain is very important for subunit interactions. Unfortunately, the nature of most mutations in the accessory subunit makes it impossible to distinguish the direct involvement of the deleted regions in subunit-subunit interaction from the effects of the structural changes of the domains caused by these deletions. According to the predicted secondary structures in Figure 3.5, most mutations in the accessory subunit include deletions of one or two beta strands and alpha helices which most likely interrupt the domainal structure. An exception is deletion Δ(264-274) which includes a part of a loop plus less than two turns of an alpha helix (H7 in Figure 3.5) in the C-domain. So mutant $\beta\Delta$(264-274) may still have a C-terminal domain folded similarly but nonetheless shows a weak interaction with the catalytic subunit, suggesting that residues aa264-274 are involved directly in subunit association. Interestingly, a structurally corresponding region (aa31-36) in thioredoxin, presumably the counterpart of pol γ–β in T7 DNA polymerase complex, is actually involved in interaction with T7 DNA polymerase as shown in the determined X-ray structure.

It is noteworthy that the sequence homology between the M-domain and RNase H domain of HIV-RT is an unique property of the accessory subunit of *Drosophila* mitochondrial DNA polymerase; this similarity is not conserved in other accessory subunit homologs of animal mitochondrial

DNA polymerases. As noted earlier, the most conserved region (or domain) is the C-domain among all the identified accessory subunit homologs (Wang, 1998; Fan et al., 1999). This may suggest that the mechanisms of mtDNA replication in different systems are similar but not identical. This prediction is in agreement with the different structures of mtDNA molecules from different sources (for review see (Gray, 1989)). MtDNA molecules with a D-loop form have been isolated from mammals but not from fruit flies, even though *Drosophila* mtDNA molecules contain an 'A+T' region which is believed to function similarly to the D-loop region of mammalian mtDNA (Clayton, 1982; Goddard and Wolstenholme, 1980). The position and structure of the initiation site for the lagging strand mtDNA synthesis is also different in flies than in mammals.

The results presented in this chapter allows parallels to be drawn between *Drosophila* mitochondrial DNA polymerase and HIV reverse transcriptase. HIV-RT is a heterodimer of p66 and p51, and both proteins are encoded by the same *pol* gene. The p66 subunit contains both a polymerase and an RNase H domain (Kohlstaedt et al., 1992; Jacobo-Molina et al., 1993). The RNase H domain itself does not show activity even though it has a stable independent structure (Davies et al., 1991). However, p66 or the reconstitution of p51 and RNase H domain polypeptide provides RNase H activity (Hostomsky et al., 1991; Prasad et al., 1991). By comparison, the M-domain of pol γ–β itself may not be

sufficient for RNase H activity. Other domains in the accessory subunit of pol γ or even some parts from the catalytic subunit may be essential for the RNase H activity if it exists in the heterodimeric mitochondrial DNA polymerase. In the heterodimer of HIV-RT, p66 and p51 have quite different structures with p66 providing both the pol and RNase H activities. However, the presence of p51 is essential for the polymerase activity. HIV-RT starts minus strand DNA synthesis with human tRNA(Lys-3) whose 3' end partially unfolds and forms 18 base pairs of duplex with the viral RNA primer binding site. In order to allow plus strand DNA synthesis, the RNA template is then removed by its RNase H activity that leaves an oligoribonucleotide to serve as a primer for plus strand DNA synthesis. These features in both minus and plus strand DNA synthesis share some similarities with the initiation of mtDNA replication. In fact, mitochondrial DNA polymerases from fly and man have been shown to have reverse transcriptase activity (Wernette and Kaguni, 1986; Longley et al., 1998). Based on these similarities, I propose a model of the heterodimeric pol γ complex (Figure 3.11) which resembles the structure of the HIV-RT complex (Kohlstaedt, et al., 1992). The catalytic subunit of pol γ is divided into the exo, spacer and pol domains in the model. Because pol γ has reverse transcriptase activity, the pol domain of pol γ–α may also fold like the pol domain of p66 in HIV-RT. The exo domain together with the
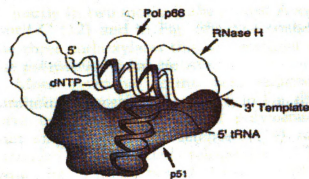
(from Science 256: 1787 (1992))
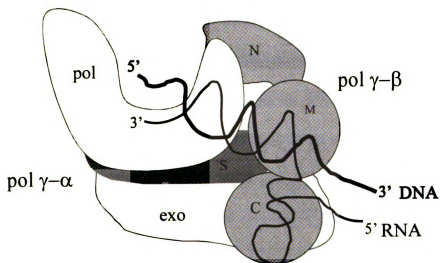
**Figure 3.11   Proposed complex for *Drosophila* pol γ at
initiation site of mtDNA replication.** In the working model, the
catalytic subunit of pol γ is divided into three domains: exonuclease (exo),
spacer (S), and polymerase (pol). The accessory subunit comprises N-, M-,
and C-domains as described in the text. The domain arrangement is based
on the results of subunit interactions, on structural modeling, and on
similarities between pol γ and HIV-RT. The substrates RNA (thin line) and
DNA(thick line) are positioned to fit the hypothetical functions of
individual domains. The proposed model resembles the initiation complex
of HIV-RT at the initiation site of minus strand replication (top) proposed
by Kohlstaedt, et al.(1992).

spacer region provide the corresponding regions of p51 in HIV-RT. The M-domain of pol γ–β corresponds to the RNase H domain of p66 in HIV-RT. Thus, in this model, the pol domain and the M-domain are contributed by different subunits in *Drosophila* pol γ but have a similar orientation as the pol domain and the RNase H domain in p66 of HIV-RT. The C-domain of pol γ–β is connected to the bottom of the M-domain to keep both domains in contact with the exo domain and the spacer of pol γ–α. It is less clear how the N-domain of pol γ–β is oriented with respect to pol γ–α but the biochemical studies suggest that there is weak interaction between the N-domain and the pol domain, which may maintain the six domains in the two subunits in the orientation required for an active pol γ enzyme. It seems plausible that *Drosophila* pol γ may initiate mtDNA replication by a mechanism similar to that of HIV minus strand initiation (Kohlstaedt, et al., 1992). As proposed previously by the structural model (Fan, et al., 1999), the C-terminal domain of pol γ–β may bind to a stem-loop structure of RNA transcribed by mitochondrial RNA polymerase, thereby placing the holoenzyme of *Drosophila* pol γ at the initiation site in a correct conformation. Once this happens, the M-domain could function as an RNase H to process the RNA into a primer with an extendible 3'-OH end. Pol γ–α then extends the primer to start DNA synthesis. In this proposed model, the M-domain would interact with the novel duplex DNA during

strand elongation, which may increase the interaction between the

holoenzyme and DNA substrate and thereby enhance its processivity.

CHAPTER IV

FUTURE RESEARCH

Biochemical analysis and computer-aided modeling have been applied in the research described in this thesis to investigate the structure-function relationships of *Drosophila* mitochondrial DNA polymerase. The results presented here reveal for the first time subunit interactions in a mitochondrial DNA polymerase complex from any source. The results suggest that the accessory subunit of *Drosophila* pol γ comprises three domains: termed N-, M-, and C-domains. The two subunits of *Drosophila* pol γ interact through multiple sites in different regions. The interactions of the M- and C-domains of the accessory subunit with the exo domain and spacer of the catalytic subunit are the main forces for the formation of the heterodimeric complex of *Drosophila* pol γ. The N-domain of the accessory subunit may coordinate the overall subunit-subunit interaction to keep the complex in an active conformation. These results provide valuable information on the overall domain interactions in *Drosophila* pol γ. Future research should focus on extending these results to identify interaction motifs or specific residues involved in subunit-subunit interaction.

Because of the complexity of the subunit interactions in *Drosophila* pol γ as discussed in chapter III, the deletion mutations described in this thesis may not distinguish the effects of the mutations on subunit interaction from structural defects caused by these mutations. Different strategies may be used for future research. I would propose to set up a phage display library for initial screening of possible peptides which are the favored

substrates of the individual subunit. In the past decade, phage-display has developed into one of the most popular and powerful methods for protein-protein interaction studies. A strategy similar to that used for mapping the protein-protein interactions between c-myb and its coactivator CBP (Kiewitz and Wolfes, 1997) would suit perfectly our purposes here. The approach would be to clone random cDNA segments from each subunit of *Drosophila* pol γ into a modified phagemid to construct a mini-library for expression of many fragments in both subunits of pol γ. This mini-library will then be propagated as phages and screened with either individual subunit or the domains of each subunit as the bait. A pool of peptides binding to the bait would be obtained. By analyzing the sequences of these peptides, we can then identify the regions in individual subunits which have sequences matching those of the peptides, and pursue selective site-directed mutagenesis to study subunit-subunit interaction by the methods described in this thesis. In this way, we will limit the number of mutations to a reasonable level and find most of the important motifs and residues for subunit interactions.

Sequence and structural analysis presented in this thesis predict that the accessory subunit may have properties of binding to DNA or RNA molecules, and in particular, RNA with stem-loop like secondary structures. These predictions provide useful information for further characterization of the accessory subunit. I suggest to use two systems to

explore the biochemical properties of the accessory subunit of *Drosophila* pol γ.

In one approach, native pol γ or recombinant pol γ complex would be the main source for biochemical characterization. Although the two subunits are associated in the complex, UV-crosslinking experiments such as those shown in Figure 4.1 will be very useful to test if the accessory subunit itself can bind to RNA or DNA molecules. As indicated in Figure 4.1, the pol γ complex was incubated with an RNA molecule (p60S), or p60S annealed to an oligodeoxyribonucleotide DK114 as a mimic of unprocessed RNA-DNA hybrid at the initiation site of mtDNA replication. Under the standard conditions for DNA polymerase activity assay, the accessory subunit but not the catalytic subunit can be crosslinked to p60S by UV irradiation, indicating strong interactions between the accessory subunit and the p60S molecule. The results confirm our prediction that the accessory subunit can bind to RNA molecules. Annealing oligonucleotide DK114 to p60S did not affect the binding. In addition, in the presence of dATP, a new crosslinked complex with slower mobility was formed as indicated by the filled arrow in Figure 4.1, suggesting that an RNA molecule extended by the pol γ was crosslinked to the accessory subunit. This preliminary result supports our prediction about the role of the accessory subunit in the initiation of mtDNA replication. Two different strategies can be used for future research. One is to study the substrate

A) Sequence and structure of primer-template substrate:

Sequence of RNA primer (p60S) (142 nt)

```
5'-gggcgaauug gguaccggaa uugccuuugc caaacuuuua
   gaaguaaagu augcuuaugg aauuauggau uuuuuguugu
   uuuuguuguu uguuuagauc gaauuccugc agcccggggg
   auccacu aguucuagagcggcc-3'          142
```

Sequence of template (DK114) (27 nt):

3'-TCAAGATCTCGCCGGTTATAATATATA-5'

Predicted secondary structure of nt1-127 in p60S:

```
              10
-G       AAUU         CCG
   GGCG        GGGUA
   CCGU        UCCGU    G
UC        ---U         UAA
   30                20


              40
   AAAC         G       AA
      UUUUA  AAGUA
      AAGGU  UUCGU   G
   ----          A     AU
      60          50


       70          80              90
   UUA          UUGUUU     ---U       UUU
      UGGAUUUUUUG      UUGU     GUUUG
      ACCUAGGGGGC      GACG     UAAGC   A
   ---              ----CC   UCCU        UAG
      120              110         100
```
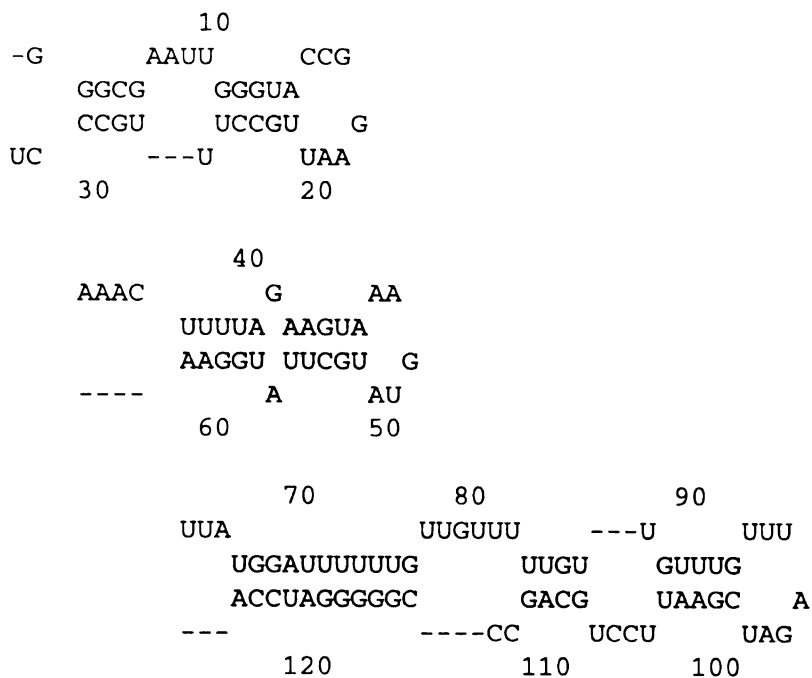
Figure 4.1

B) UV-crosslinking of pol γ–β to ^{32}P-labeled p60S
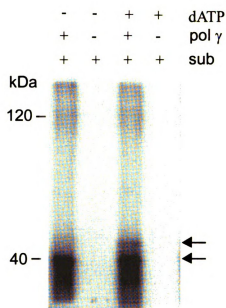
Figure 4.1   **The accessory subunit of *Drosophila* pol γ binds to RNA
primer in the presence of the catalytic subunit**. Purified recombinant
pol γ heterodimer from baculovirus-infected Sf9 cells was incubated
under conditions of standard DNA polymerase activity assay with a
primer (p60S)-template (DK114) subustrate (sub). RNA p60S was labeled
with ^{32}P at 5'-end and annealed at 3'-end with DNA DK114 where
underlined nucleotides form base-pairs (A). The secondary structure of un-
annealed sequence of p60S was predicted by *Foldrna* program of *GCG*.

The reaction mixture was incubated for 15 minutes at 30 $^{\circ}$C, and then
irradiated under UV light (305 nm) for 15 minutes on ice. The mixture
was finally resolved by 10% SDS-PAGE (B). The gel was dried and
subjected to Phosphor Imager analysis. +/- indicate the presence or absence
of reaction components. Protein sizes are based on prestain markers from
BioRad. Arrows point to crosslinked RNA-pol γ–β products.

specificity in this protein-RNA interaction by varying either the sequence or the secondary structure of the RNA substrate. Another one is to map the binding regions in the accessory subunit. One way to do this is to increase the scale of the UV-crosslinking reaction so that the amount of the crosslinked RNA-protein complex is suitable for protease digestion. The digestion products can then be analyzed by two-dimensional gel electrophoresis and mass spectrum analysis. The interaction regions of the accessory subunit will be tracked by the attached ^{32}P-labeled RNA.

A second approach to study biochemical activity of the accessory subunit, independent of the catalytic subunit, would prepare ^{35}S-labeled accessory subunit by *in vitro* transcription /translation using rabbit reticulocyte lysate. During the course of this project, I found that a significant portion of the accessory subunit prepared in this way is soluble, whereas other methods of expressing the accessory subunit alone failed. As shown in Figure 4.2, after the gradient sedimentation, the accessory subunit appears as a clear peak at a position corresponding to its size of 40 kDa although there is some unfolded or aggregated material pelleted at the bottom of the gradient. The accessory subunit prepared in this way shows weak binding to a ssDNA column (Figure 4.2). In the future, various mutated forms of the accessory subunit can be prepared and examined for their capacity to bind to ssDNA in order to map the interaction regions or residues.

A) Glycerol Gradient Sedimentation of pol γ–β
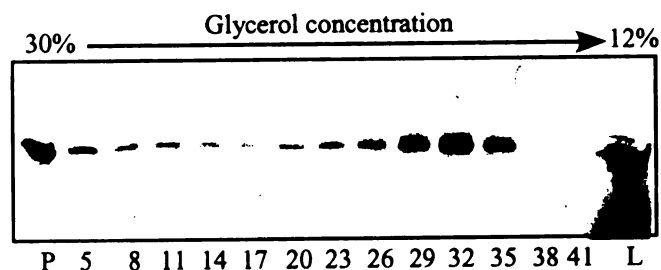
B) Binding of pol γ–β to an ssDNA-cellulose column
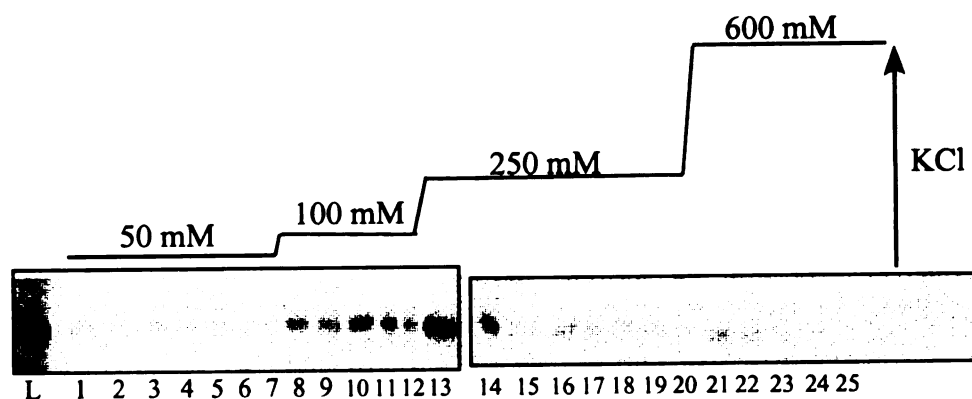
Figure 4.2 **The accessory subunit binds to ssDNA independent of the catalytic subunit.** The accessory subunit protein was prepared by *in vitro* transcription/translation with rabbit reticulocyte lysate (Promega, CA) in the presence of ^{35}S-labeled methionine. A sample of 200 μl reaction mixture was subjected to glycerol gradient sedimentation (A) or ssDNA-cellulose chromatography as described by Wang (Wang, 1998). The fractions were analyzed by 12% SDS-PAGE and autoradiography.

132

ACKNOWLEDGMENT

BIBLIOGRAPHY

134

Abagyan, R., Frishman, D., and Argos, P. (1994). Recognition of distantly related proteins through energy calculations. Proteins *19*, 132-40.

Aberg, A., Yaremchuk, A., Tukalo, M., Rasmussen, B., and Cusack, S. (1997). Crystal structure analysis of the activation of histidine by *Thermus thermophilus* histidyl-tRNA synthetase. Biochemistry *36*, 3084-94.

Adler, S. a. M., P. (1979). T7-induced DNA polymerase. J. Biol. Chem. *245*, 11605-11614.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403-10.

Araki, H., Ropp, P. A., Johnson, A. L., Johnston, L. H., Morrison, A., and Sugino, A. (1992). DNA polymerase II, the probable homolog of mammalian DNA polymerase epsilon, replicates chromosomal DNA in the yeast *Saccharomyces cerevisiae*. EMBO J. *11*, 733-40.

Arnez, J. G., and Moras, D. (1997). Structural and functional considerations of the aminoacylation reaction. Trends Biochem. Sci. *22*, 211-6.

Attardi, G. (1985). Animal mitochondrial DNA: an extreme example of genetic economy. Int. Rev. Cytol., 93-145.

Attardi, G., and Schatz, G. (1988). Biogenesis of mitochondria. Annu. Rev. Cell. Biol. *4*, 289-333.

Baynton, I., and Fuchs, R. P. (2000). Lesions in DNA: hurdles for polymerases. Trends Biochem. Sci. *25*, 74-9.

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.

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

Beese, L. S., Friedman, J. M., and Steitz, T. A. (1993b). Crystal structures of the Klenow fragment of DNA polymerase I complexed with

deoxynucleoside triphosphate and pyrophosphate. Biochemistry *32*, 14095-101.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977). The Protein Data Bank. A computer-based archival file for macromolecular structures. Eur. J. Biochem. *80*, 319-24.

Bogenhagen, D., and Clayton, D. A. (1977). Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. Cell *11*, 719-27.

Bowie, J. U., Luthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three- dimensional structure. Science *253*, 164-70.

Braithwaite, D. K., and Ito, J. (1993). Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res. *21*, 787-802.

Bryant, S. H., and Lawrence, C. E. (1993). An empirical energy function for threading protein sequence through the folding motif. Proteins *16*, 92-112.

Burgers, P. M. (1998). Eukaryotic DNA polymerases in DNA replication and DNA repair. Chromosoma *107*, 218-27.

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.

Burgers, P. M., Kornberg, A., and Sakakibara, Y. (1981). The dnaN gene codes for the beta subunit of DNA polymerase III holoenzyme of escherichia coli. Proc. Natl. Acad. Sci. USA *78*, 5391-5.

Campbell, J. L. (1993). Yeast DNA replication. J. Biol. Chem. *268*, 25261-25264.

Carrodeguas, J. A., and Bogenhagen, D. F. (2000). Protein sequences conserved in prokaryotic aminoacyl-tRNA synthetases are important for the activity of the processivity factor of human mitochondrial DNA polymerase. Nucleic Acids Res. *28*, 1237-1244.

Carrodeguas, J. A., Kobayashi, R., Lim, S. E., Copeland, W. C., and Bogenhagen, D. F. (1999). The accessory subunit of *Xenopus laevis* mitochondrial DNA polymerase gamma increases processivity of the catalytic subunit of human DNA polymerase gamma and is related to class II aminoacyl-tRNA synthetases. Mol. Cell. Biol. *19*, 4039-46.

Chang, D. D., and Clayton, D. A. (1987a). A mammalian mitochondrial RNA processing activity contains nucleus- encoded RNA. Science *235*, 1178-84.

Chang, D. D., and Clayton, D. A. (1987b). A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. EMBO J. *6*, 409-17.

Chang, D. D., and Clayton, D. A. (1985). Priming of human mitochondrial DNA replication occurs at the light-strand promoter. Proc. Natl. Acad. Sci. USA *82*, 351-355.

Chang, D. D., Hauswirth, W. W., and Clayton, D. A. (1985). Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. EMBO J. *4*, 1559-67.

Chen, C. H., and Cheng, Y. C. (1989). Delayed cytotoxicity and selective loss of mitochondrial DNA in cells treated with the anti-human immunodeficiency virus compound 2',3'- dideoxycytidine. J. Biol. Chem. *264*, 11934-7.

Chen, C. H., Vazquez-Padua, M., and Cheng, Y. C. (1991). Effect of anti-human immunodeficiency virus nucleoside analogs on mitochondrial DNA and its implication for delayed toxicity. Mol. Pharmacol. *39*, 625-8.

Chothia, C., and Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. EMBO J. *5*, 823-6.

Chow, C. S., and Coen, D. M. (1995). Mutations that specifically impair the DNA binding activity of the herpes simplex virus protein UL42. J. Virol. *69*, 6965-71.

Clary, D. O., and Wolstenholme, D. R. (1987). *Drosophila* mitochondrial DNA: conserved sequences in the A + T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. J. Mol. Evol. *25*, 116-125.

137

Clayton, D. A. (1991). Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell. Biol. *7*, 453-78.

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.

Clayton, D. A. (1984). Transcription of the mammalian mitochondrial genome. Annu. Rev. Biochem., 573-94.

Collins, K. L., Russo, A. A., Tseng, B. Y., and Kelly, T. J. (1993). The role of the 70 kDa subunit of human DNA polymerase alpha in DNA replication. EMBO J. *12*, 4555-4566.

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.

Crute, J. J., and Lehman, I. R. (1989). Herpes simplex-1 DNA polymerase. Identification of an intrinsic 5'-3' exonuclease with ribonuclease H activity. J. Biol. Chem. *264*, 19266-70.

Cusack, S. (1997). Aminoacyl-tRNA synthetases. Curr. Opin. Struct. Biol. *7*, 881-9.

Cusack, S. (1995). Eleven down and nine to go. Nat. Struct. Biol. *2*, 824-31.

Cusack, S., Yaremchuk, A., Krikliviy, I., and Tukalo, M. (1998). tRNA(Pro) anticodon recognition by *Thermus thermophilus* prolyl-tRNA synthetase. Structure *6*, 101-8.

Desjardins, P., and Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J. Mol. Biol. *212*, 599-634.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. *12*, 387-95.

Digard, P., Bebrin, W. R., Weisshart, K., and Coen, D. M. (1993). The extreme C terminus of herpes simplex virus DNA polymerase is crucial for functional interaction with processivity factor UL42 and for viral replication. J. Virol. *67*, 398-406.

Dong, Z., Onrust, R., Skangalis, M., and O'Donnell, M. (1993). DNA polymerase III accessory proteins. I. *holA* and *holB* encoding delta and delta'. J. Biol. Chem. *268*, 11758-65.

Dorsky, D. I., and Crumpacker, C. S. (1988). Expression of herpes simplex virus type 1 DNA polymerase gene by *in vitro* translation and effects of gene deletions on activity. J. Virol. *62*, 3224-32.

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. Nature *391*, 251-8.

Dujon, B. (1996). The yeast genome project: what did we learn? Trends Genet. *12*, 263-70.

Fan, L., Sanschagrin, P. C., Kaguni, L. S., and Kuhn, L. A. (1999). The accessory subunit of mtDNA polymerase shares structural homology with aminoacyl-tRNA synthetases: Implications for a dual role as a primer recognition factor and processivity clamp. Proc. Natl. Acad. Sci. USA *96*, 9527-9532.

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. *274*, 14779-85.

Fischer, D., and Eisenberg, D. (1996). Protein fold recognition using sequence-derived predictions. Protein Sci. *5*, 947-55.

Flockner, H., Braxenthaler, M., Lackner, P., Jaritz, M., Ortner, M., and Sippl, M. J. (1995). Progress in fold recognition. Proteins *23*, 376-86.

Flockner, H., Domingues, F. S., and Sippl, M. J. (1997). Protein folds from pair interactions: a blind test in fold recognition. Proteins *Suppl. 1*, 129-33.

Flower, A. M., and McHenry, C. S. (1990). The gamma subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. Proc. Natl. Acad. Sci. USA *87*, 3713-7.

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.

Fuller, C. W., Beauchamp, B. B., Engler, M. J., Lechner, R. L., Matson, S. W., Tabor, S., White, J. H., and Richardson, C. C. (1983). Mechanisms for the initiation of bacteriophage T7 DNA replication. Cold Spring Harb. Symp. Quant. Biol. *47 Pt 2*, 669-79.

Gallo, M. L., Jackwood, D. H., Murphy, M., Marsden, H. S., and Parris, D. S. (1988). Purification of the herpes simplex virus type 1 65-kilodalton DNA-binding protein: properties of the protein and evidence of its association with the virus-encoded DNA polymerase. J. Virol. *62*, 2874-83.

Goddard, J. M., and Wolstenholme, D. R. (1978). Origin and direction of replication in mitochondrial DNA molecules from *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA *75*, 3886-90.

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.

Godzik, A., and Skolnick, J. (1992). Sequence-structure matching in globular proteins: application to supersecondary and tertiary structure determination. Proc. Natl. Acad. Sci. USA *89*, 12098-102.

Gottlieb, J., Marcy, A. I., Coen, D. M., and Challberg, M. D. (1990). The herpes simplex virus type 1 UL42 gene product: a subunit of DNA polymerase that functions to increase processivity. J. Virol. *64*, 5976-87.

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

Guenther, B., Onrust, R., Sali, A., O'Donnell, M., and Kuriyan, J. (1997). Crystal structure of the delta' subunit of the clamp-loader complex of *E. coli* DNA polymerase III. Cell *91*, 335-45.

Haffey, M. L., Novotny, J., Bruccoleri, R. E., Carroll, R. D., Stevens, J. T., and Matthews, J. T. (1990). Structure-function studies of the herpes simplex virus type 1 DNA polymerase. J. Virol. *64*, 5008-18.

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.

Hindges, R., and Hubscher, U. (1997). DNA polymerase delta, an essential enzyme for DNA transactions. Biol. Chem. *378*, 345-62.

Hixson, J. E., Wong, T. W., and Clayton, D. A. (1986). Both the conserved stem-loop and divergent 5'-flanking sequences are required for initiation at the human mitochondrial origin of light- strand DNA replication. J. Biol. Chem. *261*, 2384-90.

Holm, L., and Sander, C. (1994). The FSSP database of structurally aligned protein fold families. Nucleic Acids Res. *22*, 3600-9.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123-38.

Holmgren, A. (1985). Thioredoxin. Annu. Rev. Biochem. *54*, 237-71.

Holmgren, A., Soderberg, B. O., Eklund, H., and Branden, C. I. (1975). Three-dimensional structure of *Escherichia coli* thioredoxin-S2 to 2.8 A resolution. Proc. Natl. Acad. Sci. USA *72*, 2305-9.

Holt, I. J., Harding, A. E., and Morgan-Hughes, J. A. (1988). Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature *331*, 717-9.

Huber, H. E., Tabor, S., and Richardson, C. C. (1987). *Escherichia coli* thioredoxin stabilizes complexes of bacteriophage T7 DNA polymerase and primed templates. J. Biol. Chem. *262*, 16224-16232.

Hubscher, I., Nasheuer, I., and Syvaoja, J. E. (2000). Eukaryotic DNA polymerases, a growing family. Trends Biochem. Sci. *25*, 143-147.

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*, 21498-21503.

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

Ito, J., and Braithwaite, D. K. (1990). Yeast mitochondrial DNA polymerase is related to the family A DNA polymerases. Nucleic Acids Res. *18*, 6716.

Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., and arnold, E. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proc. Natl. Acad. Sci. USA *90*, 6320-6324.

Jang, S. H., and Jaehning, J. A. (1991). The yeast mitochondrial RNA polymerase specificity factor, MTF1, is similar to bacterial sigma factors. J. Biol. Chem. *266*, 22671-7.

Johnson, A. A., Tsai, Y., Graves, S. W., and Johnson, K. A. (2000). Human mitochondrial DNA polymerase holoenzyme: reconstitution and characterization. Biochemistry *39*, 1702-1708.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). A new approach to protein fold recognition. Nature *358*, 86-9.

Jones, D. T., Tress, M., Bryson, K., and Hadley, C. (1999). Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure. Proteins *37*, 104-111.

Joyce, C. M., and Steitz, T. A. (1994). Function and structure relationships in DNA polymerases. Annu. Rev. Biochem. *63*, 777-822.

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. USA *86*, 6469-6473.

Kaguni, L. S., Wernette, C. M., Conway, M. C., and Yang-Cashman, P. (1988). Structural and catalytic features of the mitochondrial DNA polymerase from *Drosophila melanogaster* embryos. In Cancer Cells: Eukaryotic DNA Replication, T. J. Kelly and B. Stillman, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory), pp. 425-432.

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.

Kesti, T., Frantti, H., and Syvaoja, J. E. (1993). Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase epsilon. J. Biol. Chem. *268*, 10238-45.

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.

Kiewitz, A., and Wolfes, H. (1997). Mapping of protein-protein interactions between c-myb and its coactivator CBP by a new phage display technique. FEBS Lett. *415*, 258-62.

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.

Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992). Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science *256*, 1783-1790.

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.

Koonin, E. V., and Mushegian, A. R. (1996). Complete genome sequences of cellular life forms: glimpses of theoretical evolutionary genomics. Curr. Opin. Genet. Dev. *6*, 757-62.

Kornberg, A., and Baker, T. A. (1992). DNA Replication, Second Edition (New York: W. H. Freeman and Company).

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.

Kumar, A., Abbotts, J., Karawya, E. M., and Wilson, S. H. (1990). Identification and properties of the catalytic domain of mammalian DNA polymerase beta. Biochemistry *29*, 7156-7159.

Kumar, A., Widen, S. G., Williams, K. R., Kedar, P., Karpel, R. L., and Wilson, S. H. (1990). Studies of the domain structure of mammalian DNA polymerase beta. Identification of a discrete template binding domain. J. Biol. Chem. *265*, 2124-2131.

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

Kunkel, T. A., and Soni, A. (1988). Exonucleolytic proofreading enhances the fidelity of DNA synthesis by chick embryo DNA polymerase-gamma. J. Biol. Chem. *263*, 4450-4459.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. *26*, 283-91.

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.

Lee, D. Y., and Clayton, D. A. (1996). Properties of a primer RNA-DNA hybrid at the mouse mitochondrial DNA leading-strand origin of replication. J. Biol. Chem. *271*, 24262-24269.

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

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.

Leem, S. H., Ropp, P. A., and Sugino, A. (1994). The yeast *Saccharomyces cerevisiae* DNA polymerase IV: possible involvement in double strand break DNA repair. Nucleic Acids Res. *22*, 3011-7.

Lewis, D. L., Farr, C. L., Farquhar, A. L., and Kaguni, L. S. (1994). Sequence, organization, and evolution of the A+T region of *Drosophila melanogaster* mitochondrial DNA. Mol. Biol. Evol. *11*, 523-538.

Lewis, D. L., Farr, C. L., Wang, Y., Lagina, A. T. Jr., 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.

Lewis, W., and Dalakas, M. C. (1995). Mitochondrial toxicity of antiviral drugs. Nat. Med. *1*, 417-22.

Lill, R., Nargang, F. E., and Neupert, W. (1996). Biogenesis of mitochondrial proteins. Curr. Opin. Cell. Biol. *8*, 505-12.

Lim, S. E., Longley, M. J., and Copeland, W. C. (1999). The Mitochondrial p55 Accessory Subunit of Human DNA Polymerase gamma Enhances DNA Binding, Promotes Processive DNA Synthesis, and Confers N- Ethylmaleimide Resistance. J. Biol. Chem. *274*, 38197-38203.

Logan, D. T., Mazauric, M. H., Kern, D., and Moras, D. (1995). Crystal structure of glycyl-tRNA synthetase from *Thermus thermophilus*. EMBO J. *14*, 4156-67.

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

Luft, R. (1994). The development of mitochondrial medicine. Proc. Natl. Acad. Sci. USA *91*, 8731-8738.

Luthy, R., Bowie, J. U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature *356*, 83-5.

Maki, H., Horiuchi, T., and Kornberg, A. (1985). The polymerase subunit of DNA polymerase III of *Escherichia coli*. I. Amplification of the *dnaE* gene product and polymerase activity of the alpha subunit. J. Biol. Chem. *260*, 12982-6.

Maki, H., and Kornberg, A. (1987). Proofreading by DNA polymerase III of *Escherichia coli* depends on cooperative interaction of the polymerase and exonuclease subunits. Proc. Natl. Acad. Sci. USA *84*, 4389-92.

Maki, H., Maki, S., and Kornberg, A. (1988). DNA Polymerase III holoenzyme of *Escherichia coli*. IV. The holoenzyme is an asymmetric dimer with twin active sites. J. Biol. Chem. *263*, 6570-8.

Maki, S., and Kornberg, A. (1988). DNA polymerase III holoenzyme of *Escherichia coli*. II. A novel complex including the gamma subunit essential for processive synthesis. J. Biol. Chem. *263*, 6555-60.

Marcy, A. I., Olivo, P. D., Challberg, M. D., and Coen, D. M. (1990). Enzymatic activities of overexpressed herpes simplex virus DNA polymerase purified from recombinant baculovirus-infected insect cells. Nucleic Acids Res. *18*, 1207-15.

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

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

McHenry, C. S. (1982). Purification and characterization of DNA polymerase III'. Identification of *tau* as a subunit of the DNA polymerase III holoenzyme. J. Biol. Chem. *257*, 2657-63.

Modrich, P., and Richardson, C. C. (1975). Bacteriophage T7 Deoxyribonucleic acid replication in vitro. A protein of *Escherichia coli* required for bacteriophage T7 DNA polymerase activity. J. Biol. Chem. *250*, 5508-14.

Monahan, S. J., Barlam, T. F., Crumpacker, C. S., and Parris, D. S. (1993). Two regions of the herpes simplex virus type 1 UL42 protein are required for its functional interaction with the viral DNA polymerase. J. Virol. *67*, 5922-31.

Montoya, J., Ojala, D., and Attardi, G. (1981). Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs. Nature *290*, 465-70.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. Proteins *12*, 345-64.

Mosbaugh, D. W. (1988). *Porcine* pol gammma. Nucleic Acids Res. *16*, 5645-5659.

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.

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.

O'Donnell, M., Onrust, R., Dean, F. B., Chen, M., and Hurwitz, J. (1993). Homology in accessory proteins of replicative polymerases--*E. coli* to humans. Nucleic Acids Res. *21*, 1-3.

Ojala, D., Montoya, J., and Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. Nature *290*, 470-474.

Okimoto, R., and Wolstenholme, D. R. (1990). A set of tRNAs that lack either the T psi C arm or the dihydrouridine arm: towards a minimal tRNA adaptor. EMBO J. *9*, 3405-11.

Oliver, S. G. (1996). From DNA sequence to biological. Nature *379*, 597-600.

Ollis, D. L. B., P., Hamlin, R. Xuong, N. G. and Steitz, T. A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. Nature *313*, 762-766.

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.

147

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., 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.

Orengo, C. A., Jones, D. T., and Thornton, J. M. (1994). Protein superfamilies and domain superfolds. Nature *372*, 631-4.

Pavletich, N. P., and Pabo, C. O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809-17.

Pelletier, H., Sawaya, M. R., Kumar, A., Wilson, S. H., and Kraut, J. (1994). Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. Science *264*, 1891-1903.

Prasad, R., Beard, W. A., Chyan, J. Y., Maciejewski, M. W., Mullen, G. P., and Wilson, S. H. (1998a). Functional analysis of the amino-terminal 8-kDa domain of DNA polymerase beta as revealed by site-directed mutagenesis. DNA binding and 5'-deoxyribose phosphate lyase activities. J. Biol. Chem. *273*, 11121-6.

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

Prasad, R., Beard, W. A., and Wilson, S. H. (1994). Studies of gapped DNA substrate binding by mammalian DNA polymerase beta. Dependence on 5'-phosphate group. J. Biol. Chem. *269*, 18096-101.

Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988). Fidelity of HIV-1 reverse transcriptase. Science *242*, 1168-71.

Rhee, S., Martin, R. G., Rosner, J. L., and Davies, D. R. (1998). A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. Proc. Natl. Acad. Sci. USA *95*, 10413-8.

Robberson, D. L., and Clayton, D. A. (1972). Replication of mitochondrial DNA in mouse L cells and their thymidine kinase - derivatives: displacement replication on a covalently-closed circular template. Proc. Natl. Acad. Sci. USA *69*, 3810-4.

Ropp, P. A., and Copeland, W. C. (1996). Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase γ. Genomics *36*, 449-458.

Ropp, P. A., and Copeland, W. C. (1995). Characterization of a new DNA polymerase from *Schizosaccharomyces pombe*: a probable homologue of the *Saccharomyces cerevisiae* DNA polymerase gamma. Gene *165*, 103-7.

Rost, B., and Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. Proteins *19*, 55-72.

Rubenstein, J. L., Brutlag, D., and Clayton, D. A. (1977). The mitochondrial DNA of *Drosophila melanogaster* exists in two distinct and stable superhelical forms. Cell *12*, 471-82.

Russell, R. B., Copley, R. R., and Barton, G. J. (1996). Protein fold recognition by mapping predicted secondary structures. J. Mol. Biol. *259*, 349-65.

Sander, C., and Schneider, R. (1991). Database of homology-derived protein structures and the structural meaning of sequence alignment. Proteins *9*, 56-68.

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. Biochemistry *36*, 11205-15.

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

Scheuermann, R. H., and Echols, H. (1984). A separate editing exonuclease for DNA replication: the epsilon subunit of *Escherichia coli* DNA polymerase III holoenzyme. Proc. Natl. Acad. Sci. USA *81*, 7747-51.

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

Shadel, G. S., and Clayton, D. A. (1993). Mitochondrial transcription initiation. Variation and conservation. J. Biol. Chem. *268*, 16083-6.

Shortle, D. (1997). Structure prediction: folding proteins by pattern recognition. Curr. Biol. *7*, R151-4.

Shortle, D. (1999). Structure prediction: The state of the art. Curr. Biol. *9*, R205-9.

Steitz, T. A. (1998). A mechanism for all polymerases Nature *391*, 231-2.

Studwell-Vaughan, P. S., and O'Donnell, M. (1993). DNA polymerase III accessory proteins. V. Theta encoded by *holE*. J. Biol. Chem. *268*, 11785-91.

Sussman, J. L., Lin, D., Jiang, J., Manning, N. O., Prilusky, J., Ritter, O., and Abola, E. E. (1998). Protein Data Bank (PDB): Database of Three-Dimensional Structural Information of Biological Macromolecules. Acta Cryst. *D54*, 1078-1084.

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.

Waga, S., Bauer, G., and Stillman, B. (1994). Reconstitution of complete SV40 DNA replication with purified replication factors. J. Biol. Chem. *269*, 10923-34.

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

Wallace, D. C. (1994). Mitochondrial DNA mutations in diseases of energy metabolism. J. Bioenergy and Biomembranes *26*, 241-250.

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

Wallace, D. C. (1999). Mitochondrial diseases in man and mouse. Science *283*, 1482-8.

Wallace, D. C., Singh, G., Lott, M. T., Hodge, J. A., Schurr, T. G., Lezza, A. M., Elsas, L. J. d., and Nikoskelainen, E. K. (1988). Mitochondrial

DNA mutation associated with Leber's hereditary optic neuropathy. Science 242, 1427-30.

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

Wang, Y. (1998). Molecular cloning, baculovirus expression and reconstitution of Drosophila mitochondrial DNA polymerase. Ph. D. Dissertation. Michigan State University.

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, 13640-6.

Wang, Y., and Kaguni, L. S. (1999). Baculovirus expression reconstitutes Drosophila mitochondrial DNA polymerase. J. Biol. Chem. 274, 28972-7.

Weissbach, A. (1979). The functional roles of mammalian DNA polymerase. Arch. Biochem. Biophys. 198, 386-396.

Weissbach, A., Hong, S. C., Aucker, J., and Muller, R. (1973). Characterization of herpes simplex virus-induced deoxyribonucleic acid polymerase. J. Biol. Chem. 248, 6270-7.

Weisshart, K., Kuo, A. A., Hwang, C. B., Kumura, K., and Coen, D. M. (1994). Structural and functional organization of herpes simplex virus DNA polymerase investigated by limited proteolysis. J. Biol. Chem. 269, 22788-96.

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. Biochemistry 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.

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.

Williams, A. J., Wernette, C. M., and Kaguni, L. S. (1993). Processivity of mitochondrial DNA polymerase from *Drosophila* embryos. Effects of reaction conditions and enzyme purity. J. Biol. Chem. *268*, 24855-24862.

Wolstenholme, D. R. (1992). Animal mitochondrial DNA: structure and evolution. In Mitochondrial Genomes, D. R. Wolstenholme and K. W. Jeon, eds. (New York: Academic Press), pp. 173-216.

Wolstenholme, D. R., Macfarlane, J. L., Okimoto, R., Clary, D. O., and Wahleithner, J. A. (1987). Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. Proc. Natl. Acad. Sci. USA *84*, 1324-8.

Wong, T. W., and Clayton, D. A. (1985a). Isolation and characterization of a DNA primase from human mitochondria. J. Biol. Chem. *260*, 11530-5.

Wong, T. W., and Clayton, D. A. (1985b). *In vitro* replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. Cell *42*, 951-8.

Yamaguchi, M., Matsukage, A., and Takahashi, T. (1980). Chick embryo DNA polymerase gamma. Purification and structural analysis of nearly homogeneous enzyme. J. Biol. Chem. *255*, 7002-9.

Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990). Structure of ribonuclease H phased at 2 A resolution by MAD analysis of the selenomethionyl protein. Science *249*, 1398-405.

Zarembinski, T. I., Hung, L. W., Mueller-Dieckmann, H. J., Kim, K. K., Yokota, H., Kim, R., and Kim, S. H. (1998). Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. Proc. Natl. Acad. Sci. USA *95*, 15189-93.

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. USA *94*, 11244-9.