This is to certify that the

dissertation entitled

A+T Regulatory Region of the Drosophila melanogaster Mitochondrial Genome: Organization, Evolution and Protein : DNA Interactions

presented by

David Lawrence Lewis

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Biochemistry

Laurie S. Vaguri
Major professor

Date April 1, 1994

A+T REGULATORY REGION OF THE *Drosophila melanogaster* MITOCHONDRIAL GENOME: ORGANIZATION, EVOLUTION AND PROTEIN : DNA INTERACTIONS

By

David Lawrence Lewis

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1994

ABSTRACT

A+T REGULATORY REGION OF THE *Drosophila melanogaster* MITOCHONDRIAL GENOME: ORGANIZATION, EVOLUTION AND PROTEIN : DNA INTERACTIONS

By

David Lawrence Lewis

Animal mitochondrial DNA contains a single noncoding region of substantial length. In *Drosophila* mitochondrial DNA, this region is called the "A+T region" because it contains 90- 96% deoxyadenylate and thymidylate nucleotides. The A+T region contains the origin of replication and most likely the promoters of transcription. Molecular cloning and DNA sequence analysis has revealed that the *D. melanogaster* A+T region is 4.6 kb in size and is organized in two different arrays of tandemly repeated DNA sequence elements with nonrepetitive intervening and flanking sequences comprising only 22% of its length. We conclude that the length heterogeneity observed in the A+T regulatory region in mitochondrial DNAs from the genus *Drosophila* results from the expansion and contraction of the repeat arrays.

Three DNA sequence elements are found to be highly conserved in *D. melanogaster* and in several *Drosophila* species with short A+T regions. These include a 300-bp DNA sequence element that overlaps the DNA replication origin and two thymidylate stretches identified on opposite DNA strands. We propose that the 300-bp conserved DNA sequence element, in conjunction with another nucleotide sequence determinant, perhaps the adjacent thymidylate stretch, functions in the regulation of mitochondrial DNA replication.

The DNA sequence of the coding regions flanking the A+T region was also determined. The flanking regions contain the genes encoding the 12S rRNA, 16S

rRNA, $tRNA^{val}$, $tRNA^{ile}$, $tRNA^{gln}$, $tRNA^{f\text{-}met}$ and the 5' coding region of subunit 2 of NADH dehydrogenase. These sequences, together with the DNA sequence of the A+T region and mtDNA sequences obtained in other laboratories, complete the nucleotide sequence of the *D. melanogaster* mitochondrial genome.

DNA binding proteins that interact with the A+T region were also identified. One such protein, mtDBP-26, preferentially binds to A+T region DNA with high affinity. Competition experiments using different mitochondrial A+T region DNA fragments did not show mtDBP-26 to have a pronounced preference for a particular A+T region sequence. These results might suggest that mtDBP-26 functions in the formation and maintenance of mitochondrial nucleoprotein structure.

To

Mom

Dad

and Bubby

ACKNOWLEDGMENTS

I would like to thank all the people who made this work possible and who shared so many experiences with me during these past few years. Laurie Kaguni, mentor and comrade, who bestowed on me her characteristics of perseverance and perfectionism (and a keen sense of fashion). None of this work could have been accomplished without her support, imagination and willingness to explore different avenues of research. I am greatly indebted to her. I also would like to thank other members of the lab for their efforts and companionship: Andrea Williams, Carol Farr - technician extraordinaire, Jianjun Wang and former members, Angie Kolhoff, Matt Olson and Cathy Wernette. I also acknowledge and thank the members of my thesis committee: Drs. Tom Friedman, Lee McIntosh, Steve Triezenberg and John Wilson who provided many thought provoking discussions and were always a source of useful suggestions.

No one can be a member of Laurie Kaguni's lab and not be a part of Jon Kaguni's lab. The interactions between the two have always been dynamic. I thank Jon Kaguni for many things, his Zen-like attitude towards golf, his laser printer, his restriction enzymes, etc. I thank Kevin Carr, friend and political opposite, for his unshakable belief that government can indeed make a difference. Carla Margulies, friend and alter ego, always inspired me with her enthusiasm for science and instilled in me the belief that suffering can be a good thing. I also thank Jarek Marszalek for his love of ice cream, knowledge of birds and for never letting inhibitions stop him from giving his opinions. Others members of the Jon Kaguni laboratory, past and present, including Ted Hupp, Cindy Petersen, Mark Sutton and Wenge Zhang have also contributed positively to my experiences over these past few years.

Among the many other people I have met and known in graduate school, I would like to thank Chuck Campbell for his unconditional friendship and for sharing with me his family, his beer and his insights on the games played both on and off the court, and Jim Otto, for his friendship and for showing me the importance of the Rose Bowl, a good jump shot and fine food. I also appreciate Julie Oesterle for her friendship and laughter and Carol Smith for teaching me about the nutritional benefits of cereal. In addition, I thank Julie and Carol for their efforts towards making the administrative aspects of being a graduate student less painful. I especially want to thank Diane Cox, whose patience was tried but not broken, who gave but not always received, and who I will always cherish.

Finally, I wish to express my love and gratitude to my parents, who always provided me with encouragement and support in whatever I pursued and to my brother, Kevin, for showing me things previously unknown.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER III

CHAPTER IV

LIST OF FIGURES

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

LIST OF ABBREVIATIONS

A+T	deoxyadenylate and thymidylate
ATPase	ATP synthase
bp	base pairs
CO	cytochrome oxidase
CSB	conserved sequence block
cyt	cytochrome
D-loop	displacement loop
Da	Dalton
H-strand	heavy strand
HMG	high-mobility group
HSP	heavy-strand promoter
kb	kilobase pairs
L-strand	light strand
LSP	light-strand promoter
LSU	large subunit
mtDNA	mitochondrial DNA
mRNA	messenger RNA
nt	nucleotides
ND	NADH dehydrogenase
O_H	origin of heavy-strand DNA synthesis
O_L	origin of light-strand DNA synthesis
rRNA	ribosomal RNA
SSU	small subunit
TAS	termination associated sequence
tRNA	transfer RNA

CHAPTER I

LITERATURE REVIEW

MITOCHONDRIA: BASIC CONCEPTS

Mitochondria are subcellular organelles found in all eukaryotic organisms (Tzagoloff 1982). The size of the mitochondrion varies between 1-2 μm in length and 0.5-1 μm in width, and is dependent on the type and respiratory state of the cell (Pollak and Sutton 1980). Mitochondria consist of an outer and inner membrane and two aqueous subcompartments, the inter-membrane space and the matrix. The inner membrane is folded into cristae, actually increasing its total surface area. The mitochondrion contains its own DNA (Nass and Nass 1963), consistent with the hypothesis that mitochondria arose from a bacterial symbiont (Wallin 1922).

The principal function of mitochondria is oxidative phosphorylation, although mitochondria also contribute to the biosynthesis of pyrimidines, amino acids, phospholipids, nucleotides, folate coenzymes, heme, urea and many other metabolites (Whittaker and Danks 1978). The mitochondrial energy-producing pathway consists of five enzyme complexes located in the inner mitochondrial membrane (Wallace 1992) . Complexes I to IV make up the electron transport chain in which electrons are transferred from NADH + H^+ and $FADH_2$ to the terminal electron acceptor, oxygen. As electrons traverse the transport chain, protons are pumped out of the mitochondrial matrix, across the mitochondrial inner membrane, and into the inter membrane space creating an electrochemical gradient. The gradient is utilized by the ATP synthase (complex V) to synthesize ATP from ADP and inorganic phosphate. Finally, an adenine nucleotide translocator exchanges mitochondrial ATP for cytosolic ADP where it is utilized in a multitude of cellular processes. All of the proteins encoded in the animal mitochondrial DNA (mtDNA) are components of this energy producing pathway (Anderson et al. 1981, Chomyn et al. 1983, Chomyn et al. 1985, Chomyn et al. 1986).

GENERAL FEATURES OF MITOCHONDRIAL DNA

The mtDNA fulfills the same basic role in all eukaryotic cells. It encodes rRNA and tRNA molecules utilized in the translation of the small number of mtDNA-derived mRNAs specifying protein subunits of the mitochondrial electron transport chain (Tzagoloff and Myers 1986, Attardi and Schatz 1988). In animal mtDNA, the proteins encoded are seven subunits of NADH dehydrogenase (ND1, ND2, ND3, ND4, ND4L, ND5 and ND6), the apocytochrome *b* component 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) (Fig. 1). Not all of these genes are found on the mtDNA of all eukaryotic organisms. Indeed, only the large subunit (LSU) rRNA and the small subunit (SSU) rRNA and the genes encoding COI and cytochrome *b* are common to all mitochondrial genomes whose coding capacity has been defined completely (Okimoto et al. 1992). In contrast, the mitochondrial genomes of some organisms contain additional genes; for example, those encoding proteins associated with mitochondrial ribosomes, or proteins involved in mitochondrial RNA processing or intron transposition (Chomyn and Attardi 1987). All other proteins involved in mitochondrial biogenesis and function are products of nuclear genes, translated in the cytoplasm and then imported into the mitochondria in a process requiring ATP and a membrane potential across the inner membrane (Pfanner and Neupert 1990).

The mitochondrial genetic code differs from the standard genetic code (Attardi and Schatz 1988). Variations in the mitochondrial genetic code can also occur between species. For example, the triplets AGA and AGG, both of which specify arginine in the standard genetic code, are used only as rare termination codons or not at all in mammalian mitochondrial genetic codes, and specify serine in insect mitochondrial genetic codes. The triplet UGA, which is a translational stop codon in the standard genetic code, specifies tryptophan in the mitochondrial genetic codes of animals, fungi and protozoa, but not in those of higher plants.

Figure 1. **Structures and genetic maps of vertebrate and *Drosophila* mitochondrial DNAs.** O_H, origin and direction of H-strand mitochondrial DNA replication in vertebrates; O_L, origin and direction of L-strand mitochondrial DNA replication in vertebrates; D, D-loop region of vertebrate mitochondrial DNA; R, direction of leading strand mtDNA replication in *Drosophila* ; A+T, A+T region of *Drosophila* mitochondrial DNA; 12S rRNA, coding region for the small subunit ribosomal RNA; 16S rRNA, coding region for the large subunit ribosomal RNA; ND1, ND2, ND3, ND4, ND4L, ND5, ND6, coding regions for six subunits of the NADH dehydrogenase complex; cyt *b* , coding region for cytochrome b; COI, COII, COIII, coding regions for the cytochrome oxidase subunits I, II and III; A6, coding region for ATPase subunit 6; A8, coding region for ATPase subunit 8. Capital letters represent the standard nomenclature for the genes which encode tRNAs. Codon recognition groups of serine and leucine tRNA genes are also shown. The arrow associated with each coding region indicates the direction of transcription. Adapted from Jacobs, H. T., D. J. Elliott, V. B. Math, and A. Farquharson (1988) J. Mol. Biol. **202:**185-217.

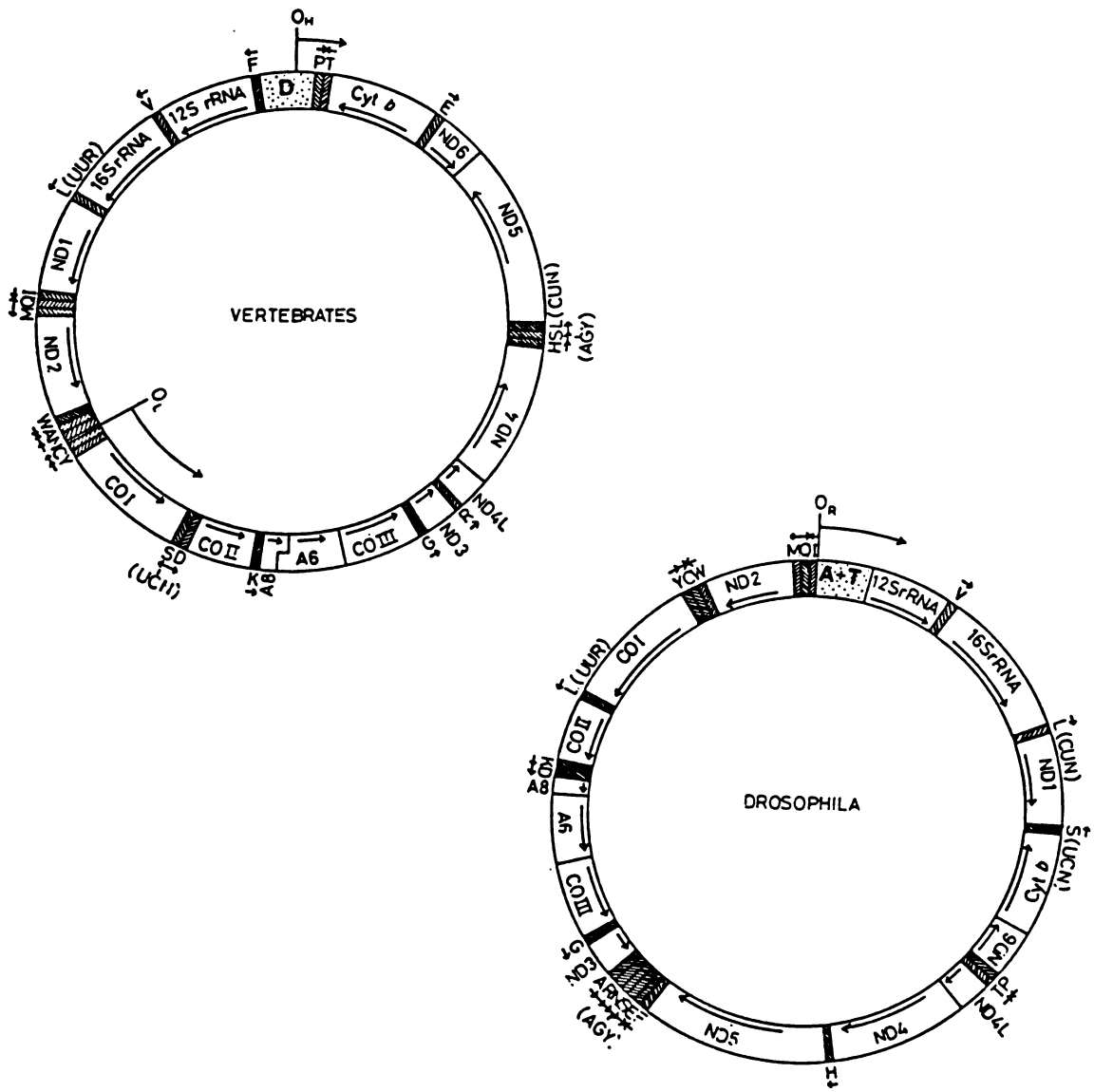


Figure 1

The mtDNA of eukaryotic organisms is present as a circular, double-stranded molecule although several examples of organisms with linear mitochondrial genomes have been noted (Grant and Chiang 1980; Wesolowski and Fukuhara 1981; Kovac et al. 1984; Warrior and Gall 1985; Suyama et al. 1985; Pritchard et al. 1986). In vertebrate mtDNA, the two DNA strands have different buoyant densities in alkaline cesium chloride gradients and have been termed the heavy (H) strand and the light (L) strand. The size of the mtDNA varies tremendously between species from 13.8 kb in the nematode, *Caenorhabditis elegans* (Okimoto et al. 1992) to an estimated 2400 kb in the cucurbit, *Cucumis melo* (muskmelon) (Ward et al. 1981). Most of the size variation is due to the size and number of noncoding regions and/or the presence of large duplications rather than increased mtDNA coding capacity (Clark-Walker 1985).

The mitochondrial genome of animals is genetically compact and contains only a single noncoding region of substantial length (Attardi 1985). The noncoding region varies in size from 121 bp in the sea urchin *Strongylocentrous purpuratus* (Jacobs et al. 1988) to approximately 6 kb in some *Drosophila* species (Fauron and Wolstenholme 1976), and contains the most divergent DNA sequences found in the animal mitochondrial genome (Moritz et al. 1987). In spite of the divergent nature of the noncoding region, the origin of DNA replication in both vertebrates and invertebrates has been located here (Clayton 1991a; Goddard and Wolstenholme 1978, 1980; Jacobs et al. 1989). In addition, the promoters of transcription for each DNA strand are located in the noncoding region of all vertebrate mtDNAs examined (Clayton 1991a) and by analogy, are most likely located here in invertebrate mtDNAs. In insect (Fauron and Wolstenholme 1976; Crozier and Crozier 1993) and nematode (Okimoto et al. 1992) mtDNA, the noncoding region is characterized by an extreme bias towards deoxyadenylate and thymidylate residues and is known as the A+T region. In vertebrate mtDNA, the non-coding region is characterized by the presence of a three-stranded structure in which a short piece of nascent H-strand DNA, termed

the displacement-loop (D-loop) strand, is located near the origin of leading-strand replication (Clayton 1982). Whether the D-loop strand 3'-hydroxyl end functions as a primer for continued replication of the mtDNA is not known, but a high level of D-loop strand turnover suggests that most do not (Bogenhagen and Clayton 1978).

There is evidence indicating that mtDNA D-loop strands exist in mitochondria as nucleoprotein complexes. Isolation of mtDNA: protein complexes from rat mitochondria revealed the presence of a 16 kDa protein, termed P16 (Van Tuyle and Pavco 1981, 1985a). P16 has a high affinity for single-stranded DNA suggesting that it is bound to D-loops in vivo (Van Tuyle and Pavco 1985b). A protein present in *Xenopus laevis* mitochondria also has an affinity for the D-loop structure in reconstitution experiments and could be recovered from mitochondrial nucleoids (Barat and Mignotte 1981). The protein was shown to have single-stranded DNA binding activity in vitro (Mignotte et al. 1985). Amino acid sequence analyses reveal both the rat and *Xenopus* proteins are homologous to *E. coli* SSB (Ghrir et al 1991a; Hoke et al. 1990).

Evidence also exists that suggests other noncoding region mtDNA, in addition to the D-loop strand, exists in a nucleoprotein complex. Electron micrographs of osmotically lysed mouse L-cell mitochondria revealed the presence of nucleoprotein complexes (Nass 1969). Proteins can also be seen in electron micrographs of rapidly sedimenting rat liver mitochondrial DNA (Van Tuyle and McPherson 1979). Similar studies using HeLa cell mitochondrial lysates showed the presence of a nucleoprotein complex within the D-loop region (Albring et al. 1977). Protection from a DNA crosslinking agent, trimethylpsoralen, was observed in the D-loop region of HeLa cell mtDNA (DeFrancesco and Attardi 1981) as well as the corresponding A+T region of *Drosophila* mtDNA (Potter et al. 1980, Pardue et al. 1984) suggesting that the protected areas are complexed with protein. In studies of *Xenopus laevis* mitochondria, several proteins have been identified in preparations of mitochondrial

nucleoids. These include a 28 kDa acid-soluble protein with unknown DNA-binding characteristics (Barat et al. 1985), and proteins that preferentially bind to supercoiled molecules containing the D-loop region in vitro (Mignotte et al. 1983). Another protein, termed mtDBP-C, is present in large amounts in *Xenopus laevis* mitochondria (Mignotte and Barat 1986). The protein has the ability to bind supercoiled, relaxed, or linear double-stranded DNA without apparent sequence specificity and in a cooperative manner (Mignotte and Barat 1986; Mignotte et al. 1988; Mignotte et al. 1990). Kinetoplast DNA, the mtDNA of trypanosomes, can also be isolated in complex with proteins (Xu and Ray 1993). The role of the nucleoprotein complex in the maintenance of mtDNA has not been established. One possibility is that the structure interacts with the inner mitochondrial membrane allowing proper segregation of the mtDNA during mitochondrial division (Albring et al. 1977).

REPLICATION OF THE MITOCHONDRIAL GENOME

Replication of Mitochondrial DNA in Vertebrates. The fact that mitochondria contain DNA that encodes essential gene products poses the question of how the DNA is stably inherited and its genes expressed. The observation that the mtDNA does not contain the genes essential for these purposes demands that nuclear gene products be involved. There is considerable evidence that the mtDNA copy number varies according to the respiratory state of the cell and the cell type indicating that mtDNA synthesis, and most likely the expression of nuclear gene products responsible for replicating mtDNA, are regulated events. For example, it has been demonstrated that the concentration of mtDNA in mammalian striated muscles increases in proportion to their oxidative capacity (Williams 1986, Williams et al. 1986), and the number of mitochondrial genomes in highly oxidative type I muscles is greater than in the glycolytic type II skeletal muscles (Annex and Williams 1990). The rate of cell growth also affects mitochondrial genome copy number as evidenced by the 2-fold

increase in the amount of mtDNA / cell in logarithmically growing mouse fibroblasts as compared to the amount in confluent cells (Shay et al. 1990). Also, mtDNA is amplified during oogenesis in sea urchin and frog, but does not replicate in the mature oocyte egg or early embryo (Matsumoto et al. 1974, Webb and Smith 1977). In addition, the amount of mtDNA in mammalian oocytes is approximately 400 times higher than that in somatic cells (Michaels et al. 1982). In HeLa cells, mtDNA replication occurs predominantly in the late S and G2 phases of the cell cycle (Pica-Mattoccia and Attardi 1972).

The mechanism of mtDNA replication has been studied most extensively in mammals although details of mtDNA replication in other systems are now emerging. The process of mammalian mtDNA replication initiation involves two distinct origins, one for each DNA strand, separated by two thirds of the genome (Clayton 1991a). The origin of H-strand synthesis (O_H) is located within the D-loop region. The origin of L-strand synthesis (O_L)is located within a small noncoding region near the middle of a cluster of five tRNA genes (Fig. 1). DNA synthesis initiates at O_H and proceeds in a unidirectional and asymmetric manner 67% of the way around the genome until O_L is exposed in a single-stranded form. L-strand synthesis is then initiated and proceeds in a direction opposite to that of H-strand replication. Thus, at least in the mammalian system, initiation at O_H dictates a replication event while initiation at O_L is a secondary but essential process.

The precise locations of the 5' ends of nascent H-strand replicative intermediates as well as the 5' ends of nascent D-loop strands have been determined (Tapper and Clayton 1981; Chang and Clayton 1985; Chang et al. 1985). The 5' ends of nascent H-strands are identical to three of the four major D-loop DNA species suggesting that a precursor-product relationship may exist between D-loop DNA strands and nascent H-strands. Given that the turnover of D-loop strands occurs at a much higher rate than initiation of mtDNA replication (Bogenhagen and Clayton 1978),

only a small subset of D-loop strands could function in replication. If elongation of D-loop strands leads to heavy strand synthesis, the control of D-loop strand termination could play a role in regulating mtDNA replication (Doda et al. 1981).

Conserved termination-associated sequences (TASs) at the downstream end of the D-loop region have been identified in proximity to the 3' ends of D-loop strands in a number of vertebrates (Fig. 2) (Doda et al. 1981; Mackay et al. 1986; Dunon-Bluteau and Brun 1987; Foran et al. 1988). TAS elements are sites of protein binding in situ as evidenced by protection from DNA methylating agents (Ghivizzani et al. 1993a; Madsen et al. 1993). Thus, the TAS-binding factor may regulate mtDNA replication by modulating termination of D-loop strands. However, it is still unknown whether mtDNA leading-strand synthesis proceeds by elongation of arrested D-loop DNA strands or whether a separate round of initiation and synthesis through the D-loop region is required. DNA-binding proteins potentially involved in the arrest of DNA synthesis have also been identified in extracts of sea urchin mitochondria (Roberti et al. 1991; Qureshi and Jacobs 1993). The protein identified by Qureshi and Jacobs (1993), termed mtPBP1, binds DNA at a sequence near a strong leading-strand DNA replication pause site outside the noncoding region, while the protein identified by Roberti et al. (1993) binds DNA sequences located in the noncoding region.

Mapping of the 5' and 3' ends of both DNA and RNA strands generated in vivo from the D-loop region has revealed that the 3' ends of some RNAs can be aligned with the 5' ends of some DNAs (Chang and Clayton 1985; Chang et al. 1985). The locations of the 5' DNA and 3' RNA ends are near three evolutionarily conserved DNA sequence blocks (CSBs I, II and III) (Fig. 2). The RNA molecules have a unique 5' end which maps to the light-strand promoter suggesting that RNAs produced from the light-strand promoter may function as primers for D-loop DNA synthesis. In support of this, a molecule consisting of RNA at its 5' portion and DNA throughout its distal portion was identified in mouse (Chang et al. 1985). Thus, it appears that transcriptional

Figure 2. **Schematic presentation of the known major features of the vertebrate mtDNA D-loop.** The dashed line represents RNA synthesis from the L-strand promoter (LSP) and the dashed/solid line represents an RNA-primed nascent DNA strand. The transition from RNA to DNA synthesis occurs within the boxed region indicated; there are three conserved sequence blocks (CSBs) at this location, I, II and III. The D-loop is bounded by the genes for $tRNA^{pro}$ and $tRNA^{phe}$. Mitochondrial RNA polymerase is shown at the transcriptional start sites at the bipartite LSP and at the H-strand promoter (HSP). The known binding sites for the mitochondria transcription factor mTF1 are indicated. The 3' ends of D-loop DNA strands map near conserved sequence elements that have been termed termination associated sequences (TAS). Taken from Clayton, D. A. (1991) Trends Biol. Sci. **16:**107-111.

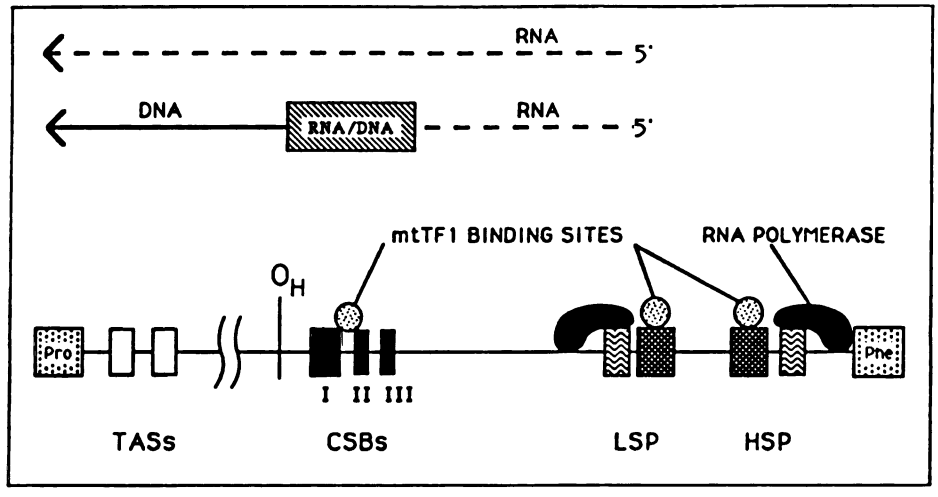

RNA
5'
DNA
RNA/DNA
RNA
5'
mtTF1 BINDING SITES
RNA POLYMERASE
O_H
Pro
Phe
I II III
TASs
CSBs
LSP
HSP

Figure 2

promoter function is involved at least in the production of D-loop DNA and possibly in replication of the mtDNA in its entirety.

The transition from RNA to DNA synthesis at O_H may be mediated by an RNA processing activity termed mitochondrial RNA-processing ribonuclease (RNase MRP) (Clayton 1991b). Originally identified in extracts of mouse mitochondria, it was found to have the ability to cleave an in vitro-generated RNA derived from D-loop region DNA in a site-specific manner (Chang and Clayton 1987a; Bennett and Clayton 1990). The location of cleavage lies between CSB II and CSB III, near one of the sites of a putative RNA-DNA transition. The endoribonuclease was later found to require an nuclear-encoded RNA component for its activity (Chang and Clayton 1987b). Molecular cloning and DNA sequencing of the gene encoding the RNA revealed it to contain a decamer sequence complementary to the sequence adjacent to the enzymatic cleavage site on D-loop RNA (Chang and Clayton 1989). The gene encoding the RNase MRP RNA has also been isolated from human and yeast (Topper and Clayton 1990; Schmitt and Clayton 1992). Human RNase MRP was also found to cleave RNA adjacent to CSB II near one of the sites of putative RNA-DNA transitions. Similarly, yeast RNase MRP cleaves an RNA derived from a putative origin near sequences resembling the CSB II in vertebrate mtDNA (Stohl and Clayton 1992).

A role for RNase MRP in mtDNA replication is based largely on circumstantial evidence and has been the subject of much debate. First, the location of RNase MRP cleavage sites in vitro do not correspond precisely with the positions of 5' D-loop DNA ends of the majority of in vivo generated species in either human or mouse mtDNA. The one exception is a minor RNase MRP cleavage site in mouse which correlates with one of the three major 5' D-loop DNA ends in that species. Second, the amount of full-length RNase MRP RNA found in mitochondrial fractions is less than 1% of that in nuclear fractions (Chang and Clayton 1987b). In fact, the amount of RNase MRP RNA associated with highly purified mitochondria was found to correspond to about

one RNA molecule per 100 mitochondria (Kiss and Filipowicz 1992). At the same time, a role of RNase MRP RNA in the nucleus has been demonstrated in yeast (Schmitt and Clayton 1993). Mutations in the yeast gene encoding the RNA lead to a defect in processing of nuclear-encoded pre-5.8S rRNA. In view of these facts, the question of whether or not a role exists for RNase MRP in mtDNA replication must be resolved.

In contrast to the complexity of events leading to the initiation of leading strand replication at O_H, initiation of light-strand replication at O_L is relatively simple. In human mtDNA, L-strand synthesis begins within a small noncoding region situated in a cluster of five tRNA genes (Tapper and Clayton 1981) (Fig. 1). This noncoding region contains DNA sequences with potential for forming a stem-loop structure. The putative structure is conserved in the mtDNAs of different mammalian species despite nucleotide sequence divergence. However, no noncoding region similar to O_L in mammalian mtDNA has been identified in chicken (Desjardins and Morais 1990) or sea urchin mtDNA (Jacobs et al. 1988, Cantatore et al. 1989). In these cases, it has been suggested that a stem-loop structure may be formed within one of the tRNA genes. L-strand synthesis begins only after O_L has been rendered single-stranded by displacement due to H-strand synthesis. The development of an in vitro priming system capable of L-strand synthesis allowed the identification of the critical DNA sequence elements required (Wong and Clayton 1981; Hixson et al. 1986). Included were an intact stem-loop structure and a short flanking sequence near the base of the stem where the transition from RNA to DNA synthesis occurs.

Replication of Mitochondrial DNA in *Drosophila* and Yeast.

Replication of insect mtDNA is best characterized in *Drosophila* species. Electron microscopic studies of *Drosophila* mtDNA replicative intermediates indicate that the origin of leading strand DNA replication is located in the A+T region (Goddard and Wolstenholme 1978, 1980). While the origin of lagging-strand DNA synthesis has not

yet been determined, lagging DNA strand initiation also likely occurs in the A+T region. *Drosophila* mtDNA replication is highly asymmetric: synthesis of the leading strand is up to 97% complete before synthesis of the lagging strand is initiated. No DNA sequence similarities between the conserved regions of vertebrate D-loop region and the *Drosophila* A+T region have been identified (Clary and Wolstenholme 1985b). Thus, the DNA sequence elements involved in initiation of *Drosophila* mtDNA replication are likely to differ from those in vertebrate mtDNA replication. However, this does not exclude the possibility that the general mechanism of mtDNA replication initiation is similar in both.

The organization of the yeast mitochondrial genome is quite different from that of animals. Genes are separated by large intergenic regions and gene order varies between yeast species (de Zamaroczy and Bernardi 1985). Three or four origins active for mtDNA replication have been identified and have several characteristics in common (de Zamaroczy et al. 1984; Faugeron-Fonty et al. 1984). All are approximately 280 bp in size and contain three GC clusters separated by A+T-rich stretches. In contrast to the relative locations of the O_H and O_L in vertebrate mtDNA, leading and lagging DNA strand synthesis in yeast mtDNA initiate in the immediate vicinity of each other (Schinkel and Tabak 1989) so that replication, although unidirectional, is symmetric. All of the yeast mtDNA replication origins are immediately downstream of an active transcriptional promoter (Christianson and Rabinowitz 1983). The existence of a functional promoter near the origins of mtDNA replication in yeast is reminiscent of the location of O_H downstream of the light-strand promoter in vertebrate mtDNA.

The Mitochondrial DNA Polymerase. The DNA polymerase responsible for the synthesis of mtDNA is DNA polymerase γ (Clayton 1982). Yeast strains containing a disruption in the gene encoding the catalytic subunit of Pol γ, *MIP1*, are completely deficient in both replication of mtDNA and DNA polymerase γ activity

(Foury 1989). Pol γ has been highly purified from five sources; chick embryos (Yamaguchi et al. 1980), *Drosophila* embryos (Wernette and Kaguni 1986), porcine liver (Mosbaugh 1988), *Xenopus laevis* oocytes (Insdorf and Bogenhagen 1989a) and HeLa cells (Gray and Wong 1992) and is distinguished from other eukaryotic polymerases by its sensitivity to N-ethylmaleimide and dideoxynucleoside triphosphates and optimal activity in high salt conditions. A proofreading, or 3'-5' exonuclease, activity has been shown to be associated with the DNA polymerase γ as purified from each source (Kunkel and Soni 1988; Kunkel and Mosbaugh 1989; Kaguni and Olson 1989; Insdorf and Bogenhagen 1989b; Gray and Wong 1992). DNA polymerase γ isolated from chick embryo has a sedimentation coefficient of 7.5S and elutes from Sephadex G-200 consistent with a native molecular weight of 180 kDa (Yamaguchi et al. 1980). Although Yamaguchi et al. proposed that the enzyme was a homotetramer of the 47 kDa polypeptide, SDS-polyacrylamide gel electrophoresis of the highly purified protein preparation revealed the presence of two polypeptides of 135 kDa and 47 kDa. The *Drosophila* Pol γ has a native molecular weight of 160 kDa as determined by velocity sedimentation and gel filtration analyses (Wernette and Kaguni 1986). Electrophoretic analyses of the purified enzyme on SDS-polyacrylamide gels indicated it is composed of a 125 kDa α-subunit and a 35 kDa β-subunit. Analysis of DNA polymerase activity in situ demonstrated that the 125 kDa subunit contains the DNA polymerase activity. The *Xenopus* and porcine liver γ polymerases have native molecular weights of approximately 180 kDa and 160 kDa, respectively, as determined by sedimentation analyses (Insdorf and Bogenhagen 1989; Kunkel and Mosbaugh 1989). Pol γ isolated from HeLa cells is composed of a 140 kDa subunit and a 54 kDa subunit as determined by SDS-PAGE (Gray and Wong 1992). Molecular cloning of the *MIP1* gene encoding the yeast DNA polymerase γ revealed an open reading frame consistent with a 143.5 kDa polypeptide (Foury 1989). Taken together, these results are consistent with a heterodimeric structure for

DNA polymerase γ consisting of a 125-145 kDa subunit containing the DNA polymerase catalytic activity, and a 35-47 kDa subunit, of unknown function.

Drosophila mtDNA polymerase has been studied extensively in vitro with regard to its DNA synthetic mechanism. It is able to replicate efficiently both predominantly single-stranded and double-stranded DNA templates (Wernette et al. 1988), a characteristic consistent with the enzymatic requirements for mtDNA replication predicted by current models. Under moderate salt conditions, Pol γ is quasi-processive, incorporating 25 to 45 nucleotides per binding event (Wernette et al. 1988). Under low salt conditions, the enzyme is highly processive and is able to replicate the entire genome of the bacteriophage M13 in a single binding event (Williams et al. 1993). Interestingly, conditions which allow highly processive DNA synthesis are suboptimal for synthetic rate. This may suggest that other accessory protein factors that allow the enzyme to be highly processive while maintaining a high synthetic rate are involved in mtDNA replication.

TRANSCRIPTION OF THE MITOCHONDRIAL GENOME

Transcription of Mitochondrial DNA in Vertebrates. Studies of the molecular mechanism of mtDNA transcription were focused initially on mapping the 5' and 3' ends of mtDNA-encoded RNA molecules from human cells (Clayton 1984; Yoza and Bogenhagen 1984). Results from these studies demonstrated that transcription of both strands of human mtDNA is initiated in the D-loop region. Elongation of the transcripts results in the production of polycistronic RNAs encompassing nearly the entire mitochondrial genome (Montoya et al. 1981; Ojala et al. 1981). The position of the transcriptional initiation site at the heavy-strand promoter (HSP) was located near the $tRNA^{phe}$ gene and the initiation site at the light-strand promoter (LSP) approximately 150 bp upstream (Fig. 2). Both promoters appear to function in a

unidirectional manner. Similar mapping studies allowed assignment of 5' ends of in vivo generated RNAs to the D-loop region of *Xenopus laevis* (Yoza and Bogenhagen 1984; Bogenhagen et al. 1986), cow (King and Low 1987) and chicken mtDNA (L'Abbe et al 1991). In bovine mtDNA, the promoters function as in human mtDNA. In *Xenopus laevis* mtDNA, two promoters were identified and each was found to function in a bidirectional manner. In chicken, a single bidirectional promoter was identified. In all animal systems studied, endonucleolytic processing of the primary RNA transcripts results in the production of rRNAs, tRNAs and mRNAs.

The development of in vitro transcription systems has led to the identification of the DNA sequence elements and protein factors involved in mtDNA transcription (Clayton 1991a, Shadel and Clayton 1993). Walberg and Clayton (1983) utilized human mitochondrial extracts to demonstrate accurate in vitro transcription from cloned mtDNA. Consistent with the results obtained by mapping in vivo generated transcripts, the transcripts produced in vitro resulted in the identification of two major promoters, one for the L-strand and one for the H-strand. Further analysis led to the definition of the D-loop region DNA sequences responsible for supporting accurate initiation of both heavy- and light-strand transcripts (Chang and Clayton 1984; Hixson and Clayton 1985; Topper and Clayton 1989). Similar studies were performed with mouse mitochondrial extracts and recombinant mouse mtDNA (Chang and Clayton 1986a-c; Bhat et al. 1989). The data revealed the promoters to consist of two domains; a short sequence near the transcriptional start site necessary but not sufficient to direct transcription initiation and a region 25 to 35 bp upstream which, when present, greatly stimulated transcription initiation.

Fractionation and biochemical analysis of the transcriptional machinery in both the human and mouse mitochondrial extracts revealed the presence of a trans-acting factor separable from RNA polymerase and required for transcriptional initiation (Fisher and Clayton 1985; Fisher et al. 1989). The factor, mitochondrial transcription

factor A (mtTFA, formerly mtTF1), binds to the DNA sequence element upstream of the transcriptional start site of both the LSP and HSP (Fisher et al. 1987, 1989; Fisher and Clayton 1988). Mutations in the upstream element that result in the abolition of specific transcriptional activity also result in the loss of mtTFA binding (Topper and Clayton 1989). DNA binding studies performed with the purified mtTFA revealed it to have only weak sequence-specificity for the upstream DNA elements, especially at the HSP (Fisher and Clayton 1988). Mouse mtTFA is able to bind the human LSP and the human mtTFA is able to bind the mouse LSP, despite the divergent nature of the LSP DNA sequences (Fisher et al. 1989). However, the mouse mtTFA is only able to activate specific transcription of the human LSP in the presence of the human RNA polymerase preparation. Likewise, the human mtTFA is only able to activate transcription from the mouse LSP in the presence of the mouse RNA polymerase preparation. These results suggest that template-specific initiation of transcription requires a promoter-specific RNA polymerase or that the RNA polymerase preparations used in the assays contain an additional specificity factor. Further purification of the RNA polymerase-containing preparations will be required in order to distinguish between these two possibilities.

Molecular cloning of the human mtTFA revealed it to be related to the vertebrate, non-histone chromosomal high-mobility group protein HMG1 (Parisi and Clayton 1991). HMG proteins consist of three domains: the amino-terminal domain and central domains, that are closely related to each other and an acidic carboxy-terminal domain (Reeck et al. 1982). The amino-terminal and central domains have a net positive charge and are referred to as HMG boxes. The HMG boxes are believed to facilitate DNA-binding while the acidic carboxy-terminal domain may facilitate interaction with histones (Einck and Bustin 1985). Several HMG box-containing proteins have been implicated in the activation of transcription of specific nuclear genes (Jantzen et al. 1990; van de Wettering et al. 1991; Travis et al. 1991; Sinclair et

al. 1990; Kelly et al. 1988). mtTFA lacks the C-terminal domain but retains the putative DNA-binding domains.

A protein homologous to human mtTFA was identified in yeast mitochondria and the gene encoding it was cloned (Diffley and Stillman 1991). This protein, termed ABF2, interacts with yeast mtDNA regulatory regions in a nonspecific but phased manner and is present in large concentrations in yeast mitochondria (Diffley and Stillman 1992). Although ABF2 is clearly homologous to the mammalian mtTFA, ABF2 stimulates only slightly transcription from yeast mitochondrial promoters (Parisi et al. 1993). Further characterization of ABF2 revealed it to be indistinguishable from a previously described protein, HM, a basic protein with DNA-packaging capabilities characteristic of prokaryotic histone-like proteins (Caron et al. 1979; Certa et al. 1984). Facilitating mtDNA organization may be the primary role of ABF2. Consistent with this hypothesis, yeast strains harboring a disruption of ABF2 lose their mtDNA after several generations of growth, resulting in a phenotype known as *rho* ^{o} (Diffley and Stillman 1991). Yeast strains containing a disrupted ABF2 allele can be rescued by the gene encoding the human mtTFA, suggesting the two gene products share a similar role in the maintenance of the mitochondrial genome (Parisi et al. 1993). In addition to its role in mtDNA packaging, it is possible that the mammalian mtTFA has evolved to become a more potent transcriptional activator than the yeast homologue.

An in vitro transcription system has also been developed using *Xenopus laevis* mitochondrial proteins and DNA sequences (Bogenhagen and Yoza 1986). Unlike the promoter structure observed in mammalian mtDNA, specific transcription in *Xenopus* mtDNA requires only a core octanucleotide sequence surrounding the initiation site (Bogenhagen and Romanelli 1988). Purification of the 140 kDa core RNA polymerase resulted in the loss of specific transcription activity as assayed on *Xenopus* mtDNA promoters (Bogenhagen and Insdorf 1988). Specific transcription activity was restored upon addition of a factor separated from RNA polymerase during

purification. The *Xenopus* specificity factor is most likely not functioning as the mammalian mtTFA given the structure of the promoter region. Specifically, the lack of a requirement for an upstream DNA element for transcription of *Xenopus* mtDNA makes the latter resemble more the situation in yeast mtDNA transcription. Interestingly, N-terminal amino acid sequence analysis of mtDBP-C, a protein proposed to be a component of the mitochondrial nucleoid, is highly similar to HMG1 (Ghrir et al. 1991b). Whether or not mtDBP-C has a stimulatory effect on the transcription of *Xenopus laevis* mtDNA, like that of mtTFA on mammalian mtDNA transcription, is not known.

The process of transcription termination has also been investigated. It is known that the rRNAs are the most abundant RNAs in mammalian mitochondria (Gelfand and Attardi 1981). A significant proportion of the 3' ends of the LSU rRNA of human mitochondria are thought not to be generated by processing of long primary transcripts but rather by specific transcription termination. Development of an in vitro system for specific transcription termination allowed identification of a DNA binding protein, termed mitochondrial termination factor (mTERF), that binds specifically to DNA sequences immediately upstream of the transcription termination sites (Kruse et al. 1989; Daga et al. 1993). The factor is postulated to act by imposing a barrier to the transcription complex. Interestingly, the binding site for mTERF encompasses a point mutation in the mtDNA associated with the human disease MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes) ; (Wallace 1992). The mutation causes a marked decrease in affinity of the mTERF for the target sequence and impairs transcription termination in vitro (Hess et al. 1991). However, no change in the steady-state level of transcripts produced in vivo was detected (Chomyn et al. 1992), raising the question as to the function of mTERF in vivo.

Transcription of Mitochondrial DNA in *Drosophila* and Yeast. Studies on the transcription of *Drosophila* mtDNA are few in number and limited to

mapping steady-state RNA molecules on the mitochondrial genome (Merten and Pardue 1981; Berthier et al., 1986). Because the organization of the mitochondrial genomes of *Drosophila* and other animals is similar, it is likely that they are transcribed in a similar manner. A polycistronic message, initiated in the non-coding A+T region, would encompass the entire gene coding region on each DNA strand, and subsequent processing of the message would result in rRNAs, tRNAs and mRNAs. Identification of the precise promoter elements will require mapping the 5' ends of transcripts generated within the A+T region in vivo.

Unlike metazoan mtDNA which contains a single major promoter for transcription of each strand, yeast mtDNA harbors several promoters. All the yeast transcriptional promoters contain the nonanucleotide, 5'-ATATAAGTA-3' (Tzagoloff and Myers 1986). RNA transcripts are initiated at the A nucleotide at the 3' end of promoter. Although yeast mitochondrial genes share an identical promoter sequence, they are transcribed at different rates (Biswas and Getz 1986; Wettstein et al. 1988). The strength of the promoter is determined by the nucleotide at position +2 (Christianson and Rabinowitz 1983) and the low rate of transcription from weak promoters in vitro was shown to be due to the slow rate of formation of the first phosphodiester bond (Biswas 1990). Specific and efficient transcription from yeast mtDNA promoters in vitro requires only a 145 kDa core RNA polymerase and a 43 kDa specificity factor (Schinkel et al. 1988). The core RNA polymerase is encoded by the RPO41 gene (Kelly and Lehman 1986; Greenleaf et al. 1986) and DNA sequence analyses reveal it to be similar to the RNA polymerases of bacteriophage T3 and T7 (Masters et al. 1987). The specificity factor is encoded by the MTF1 gene. The gene encoding MTF1 was originally identified as a suppressor of a temperature-sensitive mutation in the RPO41 gene (Lisowsky and Michaelis 1988). DNA sequence comparisons reveal MTF1 is homologous to eubacterial sigma factors (Jang and Jaehning 1991). DNA binding studies revealed that the core RNA polymerase alone

bound DNA only weakly and without sequence specificity while MTF1 was unable to bind DNA at all. Combining both components resulted in specific binding to a promoter (Schinkel et al. 1988). Taken together, the data suggest transcription from yeast mtDNA promoters resembles transcription from bacterial promoters, with the sigma factor-like MTF1 imparting promoter recognition capability onto a largely non-specific RNA polymerase. A similar situation may also exist in transcription of Xenopus *laevis* mtDNA given the similarities to the promoters in yeast mtDNA and the general transcription properties of the core RNA polymerase. Also, although stimulation of specific transcription from mammalian mitochondrial promoters may require an upstream element and a protein factor which binds to it, an additional requirement for a sigma-like factor can not be ruled out. Evidence that a sigma-like factor has a role in mammalian mtDNA transcription is provided by the observation that transcription from human and mouse mitochondrial promoters require that a homologous and inhomogeneous RNA polymerase fraction be used (Fisher et al. 1989; Shadel and Clayton 1993).

PROCESSING AND TRANSLATION OF MITOCHONDRIAL TRANSCRIPTS

Transcription of animal mtDNA results in the production of polycistronic messages encompassing the entire coding region of each DNA stand (Montoya et al. 1981: Ojala et al. 1981). These primary transcripts are processed and give rise to the LSU and SSU rRNAs, tRNAs and a number of mRNAs (Clayton 1984). In most instances, protein coding genes as well as the genes encoding the two rRNAs abut genes encoding tRNAs. Comparisons of the mtDNA sequences with the mRNA and rRNA sequences indicated that the tRNA coding regions may be involved in processing of the primary transcripts (Montoya et al. 1981; Ojala et al. 1981). Thus, it is likely that processing primary transcripts involves the recognition of the 5' and 3' ends of tRNA sequences with subsequent cleavage at those two phosphodiester bond

positions. A processing activity with RNase P-like characteristics has been identified in HeLa cell (Doerson et al. 1985) and rat (Jayanthi and Van Tuyle 1992) mitochondrial extracts, and may be responsible for cleavage at the 5' ends of pre-tRNAs. In yeast mitochondria, both an RNase P (Hollingsworth and Martin 1986) and an endonuclease responsible for processing the 3' ends of tRNA precursors have been identified (Chen and Martin 1988).

With the exception of the SSU rRNA and the tRNAs, RNA sequence analysis has shown that mtDNA-encoded RNAs are polyadenylated (Ojala et al. 1981). In some cases, polyadenylation is required to generate translational stop codons at the 3' ends of mRNAs. The mechanism of polyadenylation is unknown and there is no common sequence at or near the 3' ends of the mRNAs that would represent a potential polyadenylation signal such as that found in nuclear DNA-encoded mRNAs (Clayton 1984). As stated above, mitochondrial tRNAs are not polyadenylated, but rather each receives CCA at the 3'-terminal end by the action of a CTP(ATP):tRNA nucleotidyltransferase (Roe et al. 1982).

Translation of mitochondrial mRNAs into proteins involves mechanisms homologous to those utilized in bacteria. In fact, inhibition of mitochondrial protein synthesis by chloramphenicol, and its resistance to cycloheximide, were early molecular indications of a possible evolutionary relationship between mitochondria and bacteria (Lamb et al. 1968; Freeman 1970). However, several distinctions do exist. Because of unusual wobble rules, the 22 tRNAs encoded in all metazoan mtDNAs are apparently sufficient to decode all of the mtDNA-encoded protein genes (Barrell et al. 1980). Mitochondrial tRNAs also lack many of the invariant or semi-invariant nucleotides present in the tRNAs in bacteria and in the eukaryotic cytosol (McClain 1993). Because many of the conserved nucleotides are thought to be essential for maintaining the L-shaped tertiary structure of eukaryotic cytoplasmic and bacterial tRNAs (Sampson et al. 1990), mitochondrial tRNAs have been postulated to

have distinct higher order structures (Barrell et al. 1980). One extreme example is that of the metazoan mitochondrial $tRNA^{ser}$ (AGY) which contains a severely truncated dihydrouridine arm (Garey and Wolstenholme 1989).

Like the properties of the mitochondrial tRNAs, several properties of mitochondrial ribosomes distinguish them from those of bacteria and the eukaryotic cytoplasm. They contain only half as much rRNA and nearly twice as many proteins, differences which affect their sedimentation coefficient and buoyant density (Hamilton 1974). Although a number of mitochondrial ribosomal proteins are clearly homologous to proteins present in the *E. coli* ribosome (Kitakawa and Isono 1991), other proteins of the mitochondrial ribosome appear not to have a bacterial homologue (Davis et al. 1992). The RNA components of mitochondrial ribosomes, the rRNAs, are much smaller than their bacterial counterparts but have retained most of the core structure and domains implicated in interactions with the other components of the translation apparatus (Gutell and Fox 1988; Noller 1991; Gutell 1993).

The observation that most metazoan mitochondrial mRNAs either do not contain a 5'-leader sequence, or contain only very short leader sequences, raises the question of how the mitochondrial ribosomes bind mRNA in order to initiate translation. In prokaryotes, most mRNAs contain a stretch of nucleotides upstream of the initiation codon that is complementary to a region near the 3' end of the SSU rRNA (Shine and Delgarno 1974; Steitz and Jakes 1975). The region of complementarity serves as a primary binding site for prokaryotic ribosomes and plays a role in positioning the initiation codon in the correct position for translation. Because the region of complementarity is missing in mitochondrial ribosomes, the initiation codon must be able to bind to the decoding site of ribosomes without any upstream interactions. However, mitochondrial ribosomes are not able to bind the isolated triplet AUG suggesting that the mechanism of ribosome binding to mitochondrial mRNAs involve

more than just recognition of an AUG near the 5' end of the mRNA (Denslow et al. 1989).

EVOLUTION OF MITOCHONDRIAL DNA

The Origin of Mitochondria and Mitochondrial DNA. Mitochondria are believed to have arisen from an eubacteria-like organism through a process of endosymbiosis, which occurred some time after the divergence of the eukaryotic nuclear lineage and the eubacterial lineage from their last common ancestor (Raven 1970; Schwartz and Dayhoff 1978; Magulis 1981; Gray and Doolittle 1982; Gray 1989). Molecular evidence in support of this view has arisen largely from DNA sequence comparison analyses. A high degree of similarity exists between homologous protein genes encoded by mitochondrial and bacterial DNA (*e.g.*, Raitio et al. 1987). However, because functional protein genes encoded in mtDNA are not found in eukaryotic nuclear DNA, comparison of mitochondrial protein genes alone does not exclude a nuclear origin for mtDNA. The evolutionary history of mtDNA is better studied by DNA sequence comparison analyses of genes encoding rRNA (Gray 1988). Not only are rRNA genes ubiquitous in all known genome types, the RNAs they encode are highly similar in their core secondary structure, thus providing a basis for accurate alignment of primary sequence (Woese 1987). Using this approach, Gray et al. (1984) provided strong evidence for a specific eubacterial ancestry for mitochondria of animals, plants and fungi.

Evolution of Animal Mitochondrial Genome Size and Structure. It is generally accepted that the relatively limited coding capacity of mtDNA observed in contemporary eukaryotic cells is the result of progressive transfer of genetic material from the mitochondrion to the nucleus (Wallace 1982). As discussed previously, the genetic content of the mtDNAs of different species is nearly identical. One explanation for this observation is that genetic transfer occurred before speciation of the progenitor

eukaryotic cell. That mitochondria have retained any functional DNA may be a result of adoption of a separate genetic code that effectively "froze" the remaining mtDNA in the mitochondrion of the eukaryotic progenitor (Attardi 1988). It has been suggested that in animals, smaller mtDNAs can have an advantage in transmission over larger mtDNAs (Solignac et al. 1984, Rand and Harrison 1986). Selection for smaller genome size, possibly accomplished by a "race for replication" may have been an important factor in the movement of mtDNA from the mitochondrion to the nucleus as well as controlling the size of present day animal mtDNAs (Rand 1993).

Although the genetic content of animal mitochondrial genomes is well conserved, the order of genes differs. Gene rearrangements are especially evident when the mitochondrial genomes of vertebrate and invertebrates are compared (Gray 1989). The rearrangements include both inversions and transpositions. The presence of tRNA genes near the boundaries of the rearrangements suggests tRNA gene sequences facilitate the recombination process (Moritz et al. 1987). Precisely how tRNA gene sequences mediate recombination is not known but the process is most likely intramolecular as no evidence that intermolecular recombination occurs in animal mtDNA has been provided (Moritz et al. 1987). However, the lack of evidence for intermolecular recombination may have a trivial explanation that informative markers are absent (Rand and Harrison 1986).

Duplications of DNA sequences followed by deletion events can lead to apparent inversion of the DNA sequence surrounding the duplication junction, and may explain at least some of the rearrangements observed in animal mtDNA. Evidence that duplications and deletions occur in the mtDNA of numerous different species has been well documented (Rand 1993). Duplications/deletions may include both coding and noncoding regions of the mtDNA and are observed not only between the mtDNA molecules of different species, but also between the mtDNA molecules in the same individual. In humans, mtDNA deletion mutations have been found to be the

cause of a number of diseases and may contribute to the process of aging as well (Wallace 1992). Slipped-strand mispairing between sites containing similar DNA sequences during DNA replication is one mechanism by which these types of mutations can occur (Streisinger et al. 1966). Because replication of animal mtDNA is unidirectional and highly asymmetric, long stretches of displaced lagging strand template are produced. The displaced strand would be expected to promote slippage-induced deletions between repeated DNA sequences.

Nucleotide Sequence Evolution of Animal Mitochondrial DNA. In mammals, the rate of nucleotide substitution in mtDNA is known to be higher than that in nuclear DNA, making the study of mtDNA sequences useful in determining the evolutionary relationships between closely related taxa (Brown et al. 1982). The pattern of nucleotide substitution is distinctly nonrandom, with sequences in the noncoding D-loop region evolving very rapidly (Foran et al. 1988) and the rRNA genes evolving more slowly (Hixson and Brown 1986). There is a strong transition bias in the pattern of base substitutions in mtDNA which is not observed in nuclear DNA (Brown et al. 1982). This may reflect the fact that transitions are thermodynamically the most likely types of base substitutions to occur during replication, and that an efficient DNA repair system is absent in mitochondria (Topal and Fresco 1976; Clayton 1982). The transition bias is most evident among closely related species but decreases as sequences diverge, most likely due to multiple substitutions at single sites.

Unlike vertebrate mtDNA sequences, *Drosophila* mtDNA sequences do not evolve more rapidly than nuclear DNA sequences (Caccone et al. 1988). However, this is most likely not due to a decreased substitution rate in *Drosophila* mtDNA but rather a decreased substitution rate in vertebrate nuclear DNA (Vawter et al. 1986, Tamura 1992). As in the noncoding D-loop region of vertebrate mtDNA, base substitutions in the noncoding A+T region of *Drosophila* mtDNA arise extremely rapidly as demonstrated by electron microscopic studies of heteroduplexes between

A+
of
be
sp
nu
co
is
po
Ta
pe
ab

A+T regions of closely related species (Fauron and Wolstenholme 1980a). Evidence of rapid A+T region evolution also can be found by comparing mtDNA sequences between *Drosophila yakuba* and *Drosophila teissieri*, two of the most closely related species in the *melanogaster* subgroup (Monnerot et al. 1990). In this case, the nucleotide substitution rate was found to be 8-fold greater in the A+T region than in the coding regions. The transition bias for substitutions occurring in mtDNA of vertebrates is observed only in *Drosophila* mtDNA when closely related species are compared, possibly indicating a low saturation ceiling for base substitutions (DeSalle et al. 1987; Tamura 1992). It seems likely that in *Drosophila* mtDNA an additional constraint, perhaps a selective pressure for maintaining A+T-rich DNA, limits the number of sites able to fix a base substitution at any given time.

CHAPTER II

SEQUENCE, ORGANIZATION AND EVOLUTION OF THE A+T REGION OF *Drosophila melanogaster* MITOCHONDRIAL DNA

INTRODUCTION

In the genus *Drosophila*, mitochondrial DNA (mtDNA) is present as a circular double-stranded molecule varying in size from 16-19.5 kb (Fauron and Wolstenholme 1976). The *Drosophila* genome is identical to vertebrate mtDNA genomes with regard to coding capacity and the compact packaging of genetic information. It encodes 13 polypeptides, all of which are involved in oxidative phosphorylation, and the 22 transfer RNAs (tRNAs) and two ribosomal RNAs (rRNAs) required for mitochondrial translation, with little or no intergenic spacing (Clary and Wolstenholme 1985b). A single noncoding region, called the "A+T-rich region" because it contains 90-96% deoxyadenylate and thymidylate residues (Fauron and Wolstenholme 1976), has been shown to contain the origin of DNA replication (Goddard and Wolstenholme 1978, 1980). In fact, in all animal mtDNAs whose replication properties have been examined to date, replication proceeds unidirectionally and asynchronously from a replication origin situated in the noncoding region (Wolstenholme et al. 1983; Clayton 1991a). Further, although steady-state transcripts have not been detected in the A+T region (Merten and Pardue 1981; Berthier et al. 1986), by analogy to the vertebrate noncoding region that contains promoters for both DNA strands (Clayton 1991a), it seems likely that promoters for transcription of *Drosophila* mtDNA are located in the A+T region, and that precursor RNAs are initiated there.

The noncoding region contains the most highly divergent sequence in the vertebrate mtDNA genome (Moritz et al. 1987). Several hot spots for nucleotide substitution have been identified in evolutionary studies of the human noncoding region (Ward et al. 1991). In mammals, the noncoding region averages about 1000 bp in size and generally contains two to three regions of conserved sequence termed "conserved sequence blocks" that are in the region of the DNA replication origin, and a conserved central core of unknown function (Walberg and Clayton 1981; Foran et al. 1988). It has been proposed that one or more of the conserved sequence blocks functions in the

transition from RNA to DNA synthesis in mtDNA replication (Chang et al. 1985; Chang and Clayton 1987), and that transcription from one of the two mtDNA promoters provides an RNA for use in replication priming via an RNA processing event (Clayton 1991b). Although the locations of the promoters for transcription of the two DNA strands are generally conserved, there is little, if any, nucleotide sequence conservation among vertebrate promoters.

mtDNA heteroplasmy is widely observed among animals, and can be attributed for the most part to length heterogeneity in the noncoding region (Moritz et al. 1987). Length variability frequently results from the presence of tandemly repeated elements. For example, in monkeys (Hayasaka et al. 1991) and rabbits (Mignotte et al. 1990), tandemly repeated elements have been identified in the region known to contain the transcriptional promoters in other vertebrate species. In *Drosophila*, the A+T region varies greatly in size from 1-5 kb among species and may vary among populations in a given species or in a single fly (Fauron and Wolstenholme 1976,1980a, 1980b; Solignac et al. 1983, 1986; Hale and Singh 1986; Monforte et al. 1993; Pissios and Scouras 1993). With one notable exception (Volz-Lingenhohl et al. 1992), size variation in the A+T region accounts for the size differences observed in mtDNAs among species. A+T region length polymorphisms are in part due to the presence of varying copy numbers of repeated elements of 470 bp, as determined by DNA restriction analyses (Solignac et al. 1986; Monforte et al. 1993). However, the lack of indicator restriction enzyme sites in the *D. melanogaster* A+T region precluded demonstration of repeated elements in this species. Nevertheless, based on its size, and on the pairing in heteroduplex studies in the repeated region of the mtDNA molecules of two closely related species, *D. simulans* and *D. mauritiana* (Fauron and Wolstenholme 1980a), *D. melanogaster* mtDNA also likely contains the repeated elements .

The short A+T regions (1 kb) of *D. yakuba* , *D. teissieri*, and *D. virilis* have been cloned and sequenced (Clary and Wolstenholme 1985b,1987; Monnerot et al. 1990).

Comparative DNA sequence analyses reveal that the majority of the A+T region has diverged significantly, reflecting its rapid rate of nucleotide substitution relative to the conserved coding sequences in *Drosophila* mtDNAs (Clary and Wolstenholme 1987). However, a 450 bp region of conserved sequence was identified in the A+T region adjacent to the gene for $tRNA^{ile}$ (Monnerot et al. 1990). More recently, an expanded DNA sequence comparison including *obscura* subgroup species showed that the 150 bp immediately adjacent to the $tRNA^{ile}$ gene (excluding a stretch of thymidylate residues) exhibit a relatively high substitution rate (Monforte et al. 1993). In contrast, the remaining 300 bp are highly conserved and contain the region to which the origin of DNA replication has been mapped by electron microscopic studies (Wolstenholme et al. 1983). None of the conserved DNA sequence elements identified in the noncoding region of vertebrate mtDNAs were identified in the *Drosophila* A+T regions.

The largest variation in A+T region length occurs in the *melanogaster* subgroup of the *melanogaster* species group: A+T region size ranges from 1.1 kb in *D. yakuba* to approximately 5.5 kb in *D. mauritiana* and *D. simulans* (Wolstenholme et al. 1979; Clary and Wolstenholme 1985b; Solignac et al. 1986.) Despite multifarious approaches, intact clones of the 5 kb A+T region of *D. melanogaster* mtDNA have not been obtained (Mason and Bishop 1980; Garesse 1988; this laboratory). We have cloned the A+T region of *D. melanogaster* mtDNA in overlapping *Ssp* I and *Pac* I restriction fragments, and report here the first nucleotide sequence of a long A+T region from the genus *Drosophila*. The structural organization, function and evolution of various DNA sequence elements are discussed.

EXPERIMENTAL PROCEDURES

Materials

Enzymes and Chemicals. Micrococcal nuclease and T4 DNA polymerase were purchased from Boehringer Mannheim. Restriction endonucleases were from Gibco BRL and New England Biolabs. *E. coli* DNA polymerase I Klenow fragment and T4 DNA ligase were from New England Biolabs, and Sequenase version 2.0 was from United States Biochemical. Hoechst dye 33258 was purchased from Sigma. 5-bromo-4-chloro-3-indolyl-beta-D-galactoside (X-gal) and isopropyl-beta-D-thiogalactopyranoside (IPTG) were from Research Organics. Low-melting-point agarose was obtained from FMC Bioproducts.

Nucleic Acids and Nucleotides. Forward and reverse universal M13/pUC DNA sequencing primers and DNA primers used in direct sequencing of the *D. melanogaster* mtDNA *Hin*d III-B fragment (5' ACAAATTTTTAAGCC 3', small rRNA primer; 5' TTTATCAGGCAATTC 3', tRNA primer; and 5' ATTAAATAAAATCTATTC 3', A+T region primer) were synthesized by the Macromolecular Structure and Synthesis Facility at Michigan State University using an Applied Biosystems model 477 oligonucleotide synthesizer. *Eco*R I linkers and the vectors pNEB 193 and pUC 119 were purchased from New England Biolabs. [α-^{32}P]dATP and [α-^{35}S]dATP were from New England Nuclear.

Methods

Preparation of D. melanogaster *mtDNA.* Partially purified mitochondria were prepared from *D. melanogaster* (Oregon R) embryos as described previously (Wernette and Kaguni 1986). Mitochondria obtained from 100 g of 0-16 hour old embryos were suspended in Buffer N (30 mM Tris-HCl, pH 8.0, 2 mM $CaCl_2$, 10% (w/v) sucrose) at a ratio of 0.4 ml/g starting embryo at 0^0C and then pelleted by centrifugation at 17,500 x g

for 10 min at 4^{0}C. The mitochondrial pellet was resuspended in Buffer N (0.3 ml/g) at 0^{0}C. 10 units of micrococcal nuclease/g embryo were added, and the mixture incubated at 20^{0}C for 1 h to degrade contaminating nuclear DNA and RNA. The reaction was terminated by the addition of EDTA to 10 mM, and the mitochondria were pelleted by centrifugation, washed by resuspension at a ratio of 0.5 ml/g embryo in 30 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.15 M NaCl, 10% (w/v) sucrose and then pelleted by centrifugation. The washing step was repeated three times, and the mitochondria were then resuspended in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA (0.2 ml/g). Mitochondrial lysis was accomplished by the addition of sodium dodecyl sulfate to 1% (w/v) and EDTA to 10 mM. The mixture was incubated for 10 min at room temperature with gentle agitation. NaCl was added to 1 M and the incubation was continued for 10 min on ice. Insoluble material was pelleted by centrifugation at 4000 x g for 15 min at 4^{0}C. The mtDNA-containing supernatant was transferred to a clean tube, extracted once with phenol, once with phenol/chloroform, precipitated by the addition of 2.2 volumes of absolute ethanol, and purified further by equilibrium sedimentation in density gradients containing CsCl (density=1.61 g/ml) and 11.1 ug Hoechst dye 33258/ml . The band corresponding to mtDNA was removed by side puncture and the dye removed by extraction with distilled water/NaCl-saturated n-butanol. The aqueous phase was diluted three-fold with 10 mM Tris-HCl, pH 8.0, 1 mM EDTA and the mtDNA precipitated by addition of 2.2 volumes absolute ethanol. The yield was approximately 1 ug of homogeneous mtDNA/g embryo.

Cloning and Sequencing of the D. melanogaster *mtDNA A+T Region. D. melanogaster* mtDNA was digested with *Hin*d III and *Hae* III restriction enzymes and the *Hin*d III-B fragment containing the entire A+T region was purified by gel electrophoresis in low-melting-point agarose. An *Ssp* I restriction digest of the *Hin*d III-B fragment was prepared and cloned: *Eco*R I linkers were attached and the linkered DNA fragments were ligated into the *Eco*R I site of the vector M13 G*ori* 1 (Kaguni and Ray 1979). A *Pac*

I digest of the *Hin*d III-B fragment was prepared and then cloned using two strategies. In the first, *Pac* I DNA fragments were ligated directly into the *Pac* I site of a pUC 119 phagemid vector derivative containing the polylinker region of pNEB 193. In the second strategy, the *Pac* I DNA fragments were rendered blunt-ended by the 3'-5' exonuclease of T4 DNA polymerase (Kaguni and Kaguni 1992), *Eco*R I linkers were attached and the linkered DNA fragments were ligated into the *Eco*R I site of M13 G*ori* 1. The *E. coli* strain "Sure" (Stratagene) was used in both transfection and transformation procedures, and pUC 119-derived recombinants screened on plates containing X-gal and IPTG at 0.05% and 0.4 mM final concentration, respectively. Plaques or colonies harboring recombinant DNA were screened for insert size and orientation of the recombinant DNA by restriction analyses. DNA sequencing was performed on double- or single-stranded DNA by the dideoxy chain termination method of Sanger et al. (1977) using Sequenase version 2.0. At sites in the A+T region where overlapping recombinants were not obtained, direct sequencing of the *D. melanogaster* mtDNA *Hin*d III-B fragment was performed using the primers described under "Materials." Sequence analyses were performed using the GCG Program Package (Genetics Computer Group 1991).

Restriction Analyses of the D. melanogaster Hin*d III-B Fragment.* Restriction site mapping by partial restriction enzyme digestion and gel electrophoresis was accomplished by first end-filling the 3' overhangs of the purified *Hin*d III-B fragment by using *E. coli* DNA polymerase I Klenow fragment and [α-^{32}P]dATP. The radiolabeled *Hin*d III-B fragment was cleaved with *Ava* II or *Bsl* I restriction enzyme that recognize single sites at its opposite ends. Partial restriction- enzyme digestions were carried out using 15 ng of the appropriately digested *Hin*d III-B fragment. *Pac* I partial digestion reactions also contained 0.6 ug of pNEB 193 DNA as a bulk substrate and reactions with *Ssp* I, *Ase* I and *Dra* I contained 1 ug of lambda DNA. *Pac* I (1 unit), *Ssp* I (1.25 units), *Ase* I (1.25 units) and *Dra* I (1.25 units) reactions were incubated at 37^{0}C for 15

min, 75 s, 70 s and 3 min, respectively, and terminated by the addition of EDTA to 20 mM. Aliquots were then electrophoresed in a 1.2% agarose slab gel (13x18x0.7 cm) in 89 mM Tris, 89 mM boric acid, 2 mM EDTA for 750 volt-hours. The gel was dried under vacuum and exposed at -80^{0}C to Kodak X-OMAT AR film using a Dupont Cronex Quanta III intensifying screen.

RESULTS

Cloning and Nucleotide Sequence of the A+T Region of *D. melanogaster* Mitochondrial DNA. The 5.8 kb *Hin*d III-B fragment of *D. melanogaster* mtDNA was cloned as *Ssp* I and *Pac* I restriction digests into M13 bacteriophage and phagemid vectors as described under "Experimental Procedures." The A+T region was sequenced in its entirety using the strategy presented in Figure 1. The nucleotide sequence is shown in Figure 2. The A+T region comprises 4601 bp of which 96% are deoxyadenylate and thymidylate residues. The size of the A+T region and its high A+T content determined by DNA sequence analysis correlate well with estimates derived by denaturation mapping and electron microscopy (Klukas and Dawid 1976; Fauron and Wolstenholme 1976). The DNA sequence indicates two types of tandemly repeated elements that constitute 78% of the A+T region. Five type I repeated elements with an average length of 346 bp are present within the half of the A+T region adjacent to the gene for the small rRNA. Four type II repeats of 464 bp are present adjacent to the gene for $tRNA^{ile}$, and a fifth partial repeat is present at the end of the repeat array distal to the $tRNA^{ile}$ gene.

Structure and Organization of Repeated DNA Sequence Elements. To demonstrate conclusively the number and organization of repeated DNA sequence elements in the A+T region of *D. melanogaster* mtDNA, restriction enzyme mapping was performed on the *Hin*d III-B fragment that was isolated directly from mtDNA and end-labeled. Electrophoretic analysis of the products of partial digestion produced by each of four enzymes from each end is presented in Figure 3. The restriction patterns confirm the number and organization of repeated DNA sequence elements that we deduced from the recombinant DNA sequence.

The autoradiograph on the left-hand side of Figure 3 shows the results obtained when mapping was performed on the *Hin*d III-B fragment labeled at the end containing the small rRNA gene after cleavage with *Bsl* I. Five type I repeats are evident, each

Figure 1. ***D. melanogaster* mtDNA *Hin*d III-B fragment and DNA sequencing strategy.** The upper schematic shows the 5.8 kb *Hin*d III-B fragment containing the central A+T region (4601 bp) and flanking genes encoding the small rRNA, $tRNA^{ile}$ (I), $tRNA^{gln}$ (Q), $tRNA^{f\text{-}met}$ (M), and the 5' portion of NADH dehydrogenase subunit 2 (ND2): solid boxes, tRNA genes; box with rightward leaning slashes, 5'-coding region for NADH dehydrogenase subunit 2; dotted box, small rRNA gene; solid line, non-coding sequences including the central A+T region. Solid lines above the schematic indicate *Pac* I recognition sites; dashed lines below the schematic indicate *Ssp* I recognition sites. Arrows shown below the schematic indicate the extent and direction of individual DNA sequences obtained. Bold arrows indicate DNA sequences obtained by direct sequencing of mtDNA.

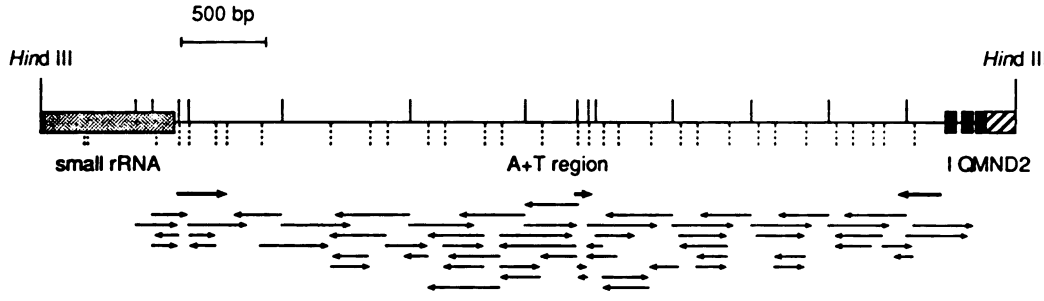


Figure 1

Figure 2. **Nucleotide sequence of the 4601 bp A+T region of *D. melanogaster* mtDNA.** Numbering begins at the first nucleotide of the A+T region nearest the gene encoding the small rRNA and continues to the last nucleotide adjacent to the gene encoding $tRNA^{ile}$. Repeated DNA sequence elements are bracketed; the longest thymidylate stretch on each DNA strand is boxed.

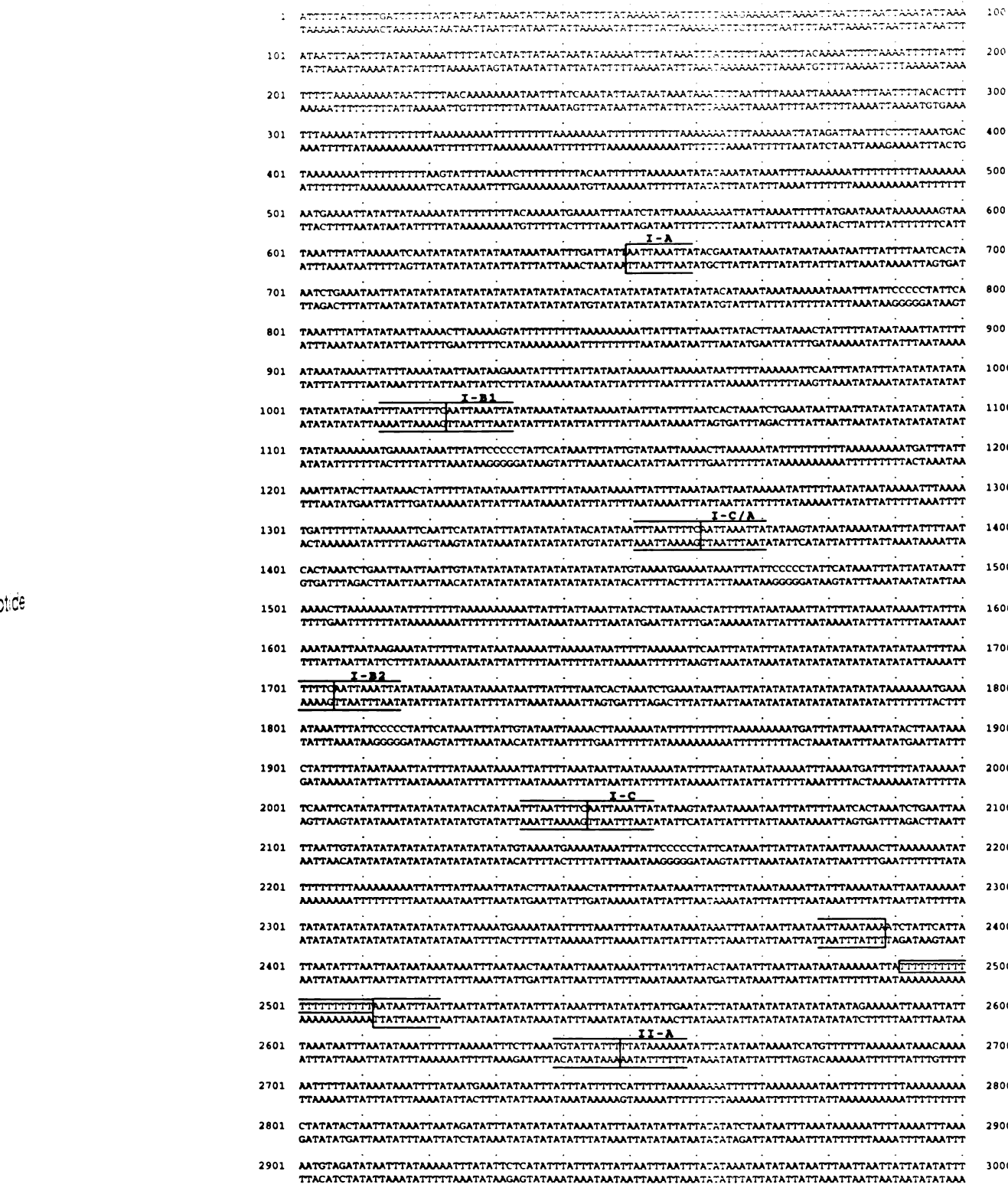


Figure 2

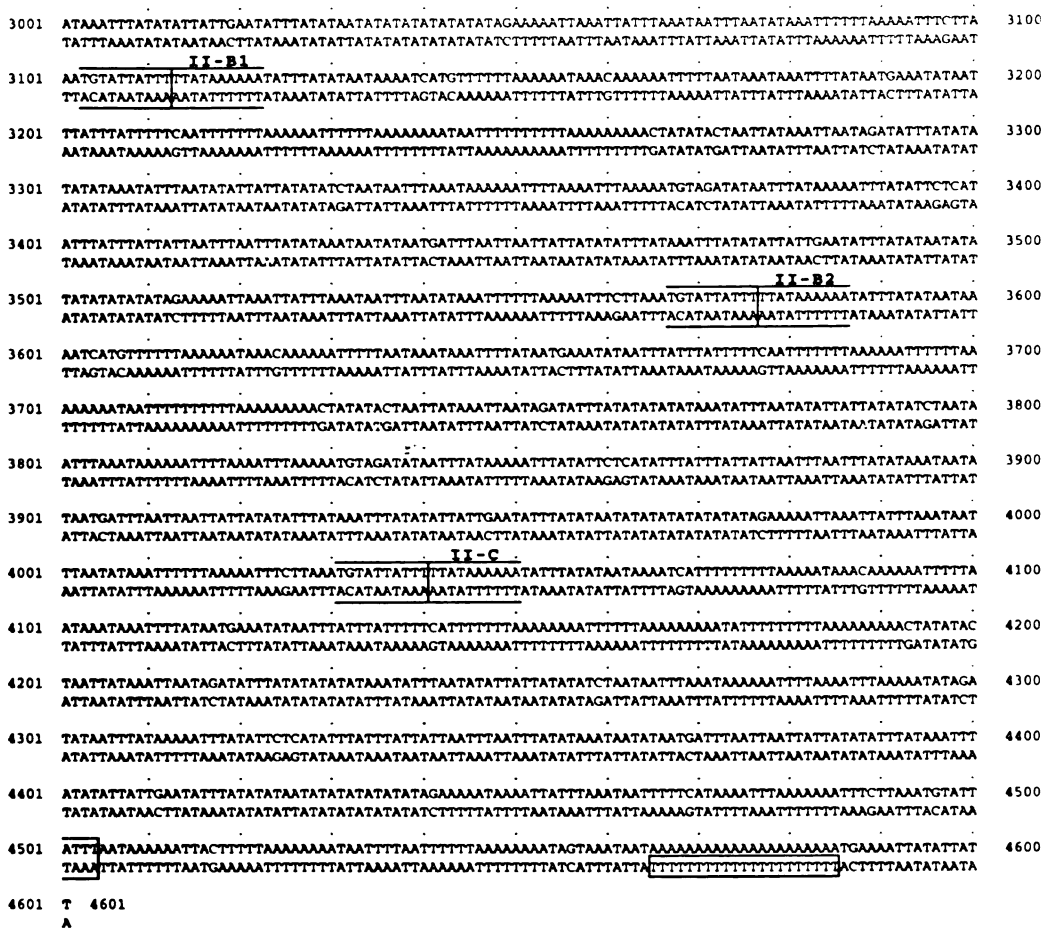

II-B1
II-B2
II-C

Figure 3. **Restriction site mapping of the *D. melanogaster* mtDNA *Hin*d III-B fragment.** Partial restriction enzyme digestion of the mtDNA *Hin*d III-B fragment was carried out as described under "Material and Methods." Left panel, Restriction site mapping of the *Hin*d III-B fragment radiolabeled at the end containing the gene for the small rRNA. M, lambda DNA digested with *Hin*d III ; lane 1, gel-purified mtDNA *Hin*d III-B fragment; lane 2, *Hin*d III-B fragment digested to completion at the *Bsl* I site in the gene for $tRNA^{ile}$; lanes 3-6, *Pac* I, *Ssp* I, *Ase* I and *Dra* I partial digestions of the *Bsl* I-digested *Hin*d III-B fragment, respectively. Repeated restriction-fragment patterns nearest the gene for the small rRNA are bracketed. Right panel, Restriction site mapping of the *Hin*d III-B fragment radiolabeled at the end containing the gene for NADH dehydrogenase subunit 2. M, lambda DNA digested with *Hin*d III; lane 1, gel-purified mtDNA *Hin*d III-B fragment; lane 2, *Hin*d III-B fragment digested to completion at the *Ava* II site in the gene for the small rRNA; lanes 3-6, *Pac* I, *Ssp* I, *Ase* I and *Dra* I partial digestions of the *Ava* II-digested *Hin*d III-B fragment, respectively. Repeated restriction-fragment patterns nearest the gene encoding $tRNA^{ile}$ are bracketed. The additional *Ssp* I fragment in the first repeated pattern is indicated by an arrow.

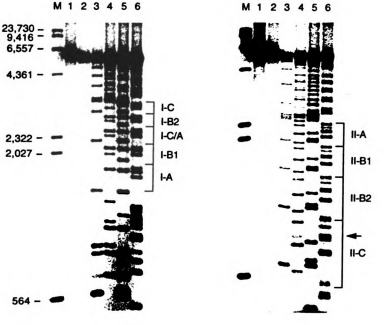

M 1 2 3 4 5 6
23,730
9,416
6,557
4,361
2,322
2,027
564
I-C
I-B2
I-C/A
I-B1
I-A
M 1 2 3 4 5 6
II-A
II-B1
II-B2
II-C

Figure 3

containing *Ase* I and *Dra* I sites at identical locations. *Pac* I and *Ssp* I sites in some of the repeats allowed verification of the order in which the elements occur. The two type I-B repeats contain two *Ssp* I sites, while types I-A and I-C contain only the second or first site, respectively. *Pac* I sites are found only near C-B junctions within the repeated region. The C/A repeat is a hybrid element (see below).

The autoradiograph on the right-hand side of figure 3 shows the results obtained when mapping was performed on the *Hin*d III-B fragment labeled at the end containing the gene for nicotinamide adenine dinucleotide (NADH) dehydrogenase subunit 2 after cleavage with *Ava* II. Four type II repeated elements are evident. *Pac* I, *Ase* I and *Dra* I sites are located at identical positions in each repeat, reflecting their nearly identical nucleotide sequences. The presence of an additional *Ssp* I site in the type II-C repeat allows its unambiguous assignment nearest the $tRNA^{ile}$ gene. Direct sequencing of the *Hin*d III-B fragment isolated from mtDNA also confirmed the position of the type II-C repeat (Fig. 1 and data not shown). Further, our isolation of a single recombinant DNA with an internal *Ssp* I site that resulted from partial digestion, allows the placement of the type II-A repeat at the innermost position in the A+T region. Finally, the DNA fragment containing the type II-B element migrates anomalously and as a doublet in polyacrylamide gel electrophoretic analyses of *Pac* I digests of the *Hin*d III-B fragment, indicating that there are two type B elements (data not shown). Thus, these may be assigned to the internal positions in the repeat unit.

Figure 4 shows the organization of the repeated DNA sequence elements in the A+T region. A nucleotide sequence alignment of the type I repeated elements is presented in Figure 5. They range from 338 to 373 bp in length and contain 134 nucleotide substitutions relative to the repeat consensus shown, representing an average 7.7% divergence. The type C/A repeat aligns clearly as a hybrid element, containing type I-C sequence in the first two-thirds and I-A sequence in the last one-third. Because of the multiple substitutions throughout, all of the type I repeats can be

Figure 4. **Repeated DNA sequence elements in the A+T-rich region of *D. melanogaster* mtDNA.** The *D. melanogaster Hin*d III-B fragment is shown as in Figure 1. The five type I repeats are shown as boxes with wavy lines. The four type II repeats and the partial repeat at the left end of the type II repeat array are shown as boxes with leftward leaning slashes. The arrow indicates the additional *Ssp* I recognition site found only in the type II-C repeat (see Fig. 3, right panel).

f *D.*
in
ype II
n as

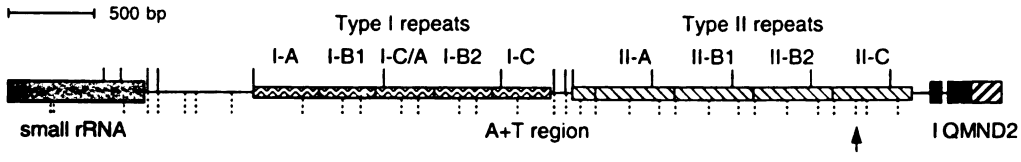


Figure 4

Figure 5. **Nucleotide sequence comparison of the type I repeats.** The nucleotide position of each repeat in the DNA sequence presented in Fig. 2 is indicated. Only one of the two type I-B repeats is shown because the I-B1 and I-B2 repeats have an identical nucleotide sequence. The consensus sequence is also shown. A dot in the sequence indicates a nucleotide that is the same as in the consensus sequence. A dash indicates a nucleotide that is absent. A letter indicates a substitution. Sequence alignments were produced using the "Pileup" program of the Genetics Computer Group software package with a gap weight of 5.0 and a gap length weight of 0.3. The consensus sequence was generated using the program "Pretty."

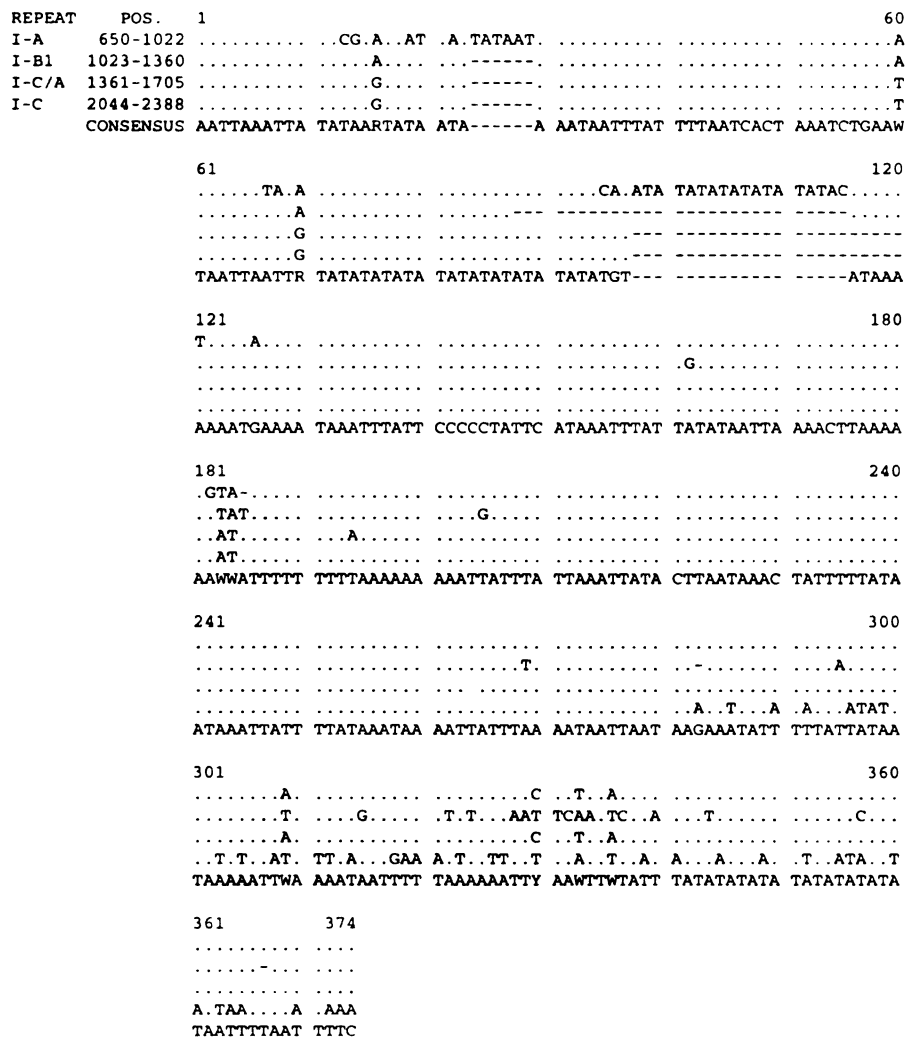

REPEAT POS. 1 60
I-A 650-1022 ....CG.A..AT .A.TATAAT. ....A
I-B1 1023-1360 ....A.... ...------. ....A
I-C/A 1361-1705 ....G.... ...------. ....T
I-C 2044-2388 ....G.... ...------. ....T
CONSENSUS AATTAAATTA TATAARTATA ATA------A AATAATTTAT TTTAATCACT AAATCTGAAW
61 120
.....TA.A ....CA.ATA TATATATATA TATAC.....
.....A ....--- ---------- ---------- -----.....
.....G ....--- ---------- ----------
.....G ....--- ---------- ----------
TAATTAATTR TATATATATA TATATATATA TATATGT--- ---------- -----ATAAA
121 180
T....A.... ....
.... .G..... .....
....
....
AAAATGAAAA TAAATTTATT CCCCCTATTC ATAAATTTAT TATATAATTA AAACTTAAAA
181 240
.GTA-..... ....
..TAT..... ....G..... ....
..AT..... ...A..... ....
..AT..... ....
AAWWATTTTT TTTTAAAAAA AAATTATTTA TTAAATTATA CTTAATAAAC TATTTTTATA
241 300
....
....T. ....-..... ....A.....
....
....A..T...A .A...ATAT.
ATAAATTATT TTATAAATAA AATTATTTAA AATAATTAAT AAGAAATATT TTTATTATAA
301 360
.....A. ....C ..T..A.... ....
.....T. ....G..... .T.T...AAT TCAA.TC..A ...T..... .....C...
.....A. ....C ..T..A.... ....
..T.T..AT. TT.A...GAA A.T..TT..T ..A..T..A. A...A...A. .T..ATA..T
TAAAAATTWA AAATAATTTT TAAAAAATTY AAWTTWTATT TATATATATA TATATATATA
361 374
....
.....-... ....
....
A.TAA....A .AAA
TAATTTTAAT TTTC

Figure 5

assigned unambiguous map positions. In contrast, nucleotide sequence alignment of the four type II repeated elements shows they are all 464 bp in length and have a nearly identical sequence (Figure 6). A fifth partial repeat aligns with the final one-third of the DNA sequence of element II-A. The type II elements contain only 27 nucleotide substitutions relative to the repeat consensus shown, representing an average 1.5% divergence. Nevertheless, for the reasons indicated above, we can assign clearly the map positions presented in Figure 4.

Conserved DNA Sequence Elements Among *Drosophila* Species. Two types of highly conserved DNA sequence elements are present in both the A+T region of *D. melanogaster* mtDNA and the shorter A+T regions of *D. yakuba*, *D. virilis* and *D. teissieri* mtDNA. The type II repeated element of *D. melanogaster* contains the only highly conserved region of substantial size between the four species. It is 281 bp in length and corresponds to the highly conserved DNA sequence element identified previously in several species with short A+T regions (Clary and Wolstenholme 1987; Monnerot et al. 1990; Monforte et al. 1993). The nucleotide sequence alignment shown in Figure 7 shows that the conserved element present in each of the four type II repeats in *D. melanogaster* bears 86% sequence similarity with *D. yakuba* and *D. teissieri*, and 80% with *D. virilis*. In both alignments, G or C to A or T substitutions are biased toward increased A+T content in *D. melanogaster*. Of the 27 substitutions of G or C in *D. yakuba*/*D. virilis*, 22 yield A or T in *D. melanogaster*. The conserved DNA sequence element overlaps the DNA replication origin in the three species in which it has been mapped by electron microscopic studies (Wolstenholme et al. 1983; Figure 8).

The second type of conserved DNA sequence element present in the A+T regions of the mtDNAs of the four species is a thymidylate stretch, present in two locations on opposite DNA strands in each mtDNA (Figure 7). The thymidylate stretches map to similar positions among the four species, and represent the longest such stretch on each strand in the A+T region. The first, ranging in size from 19 nt in *D.*

Figure 6. **Nucleotide sequence comparison of the type II repeats.** The nucleotide position of each repeat in the DNA sequence presented in Fig. 2 is indicated. Only one of the two type II-B repeats is shown because the II-B1 and II-B2 have an identical nucleotide sequence. The consensus sequence is also shown. A dot in the sequence indicates a nucleotide that is the same in the consensus sequence. A dash indicates a nucleotide that is absent. A letter indicates a substitution. Sequence alignments and the consensus were generated as described in the legend to Fig. 5.

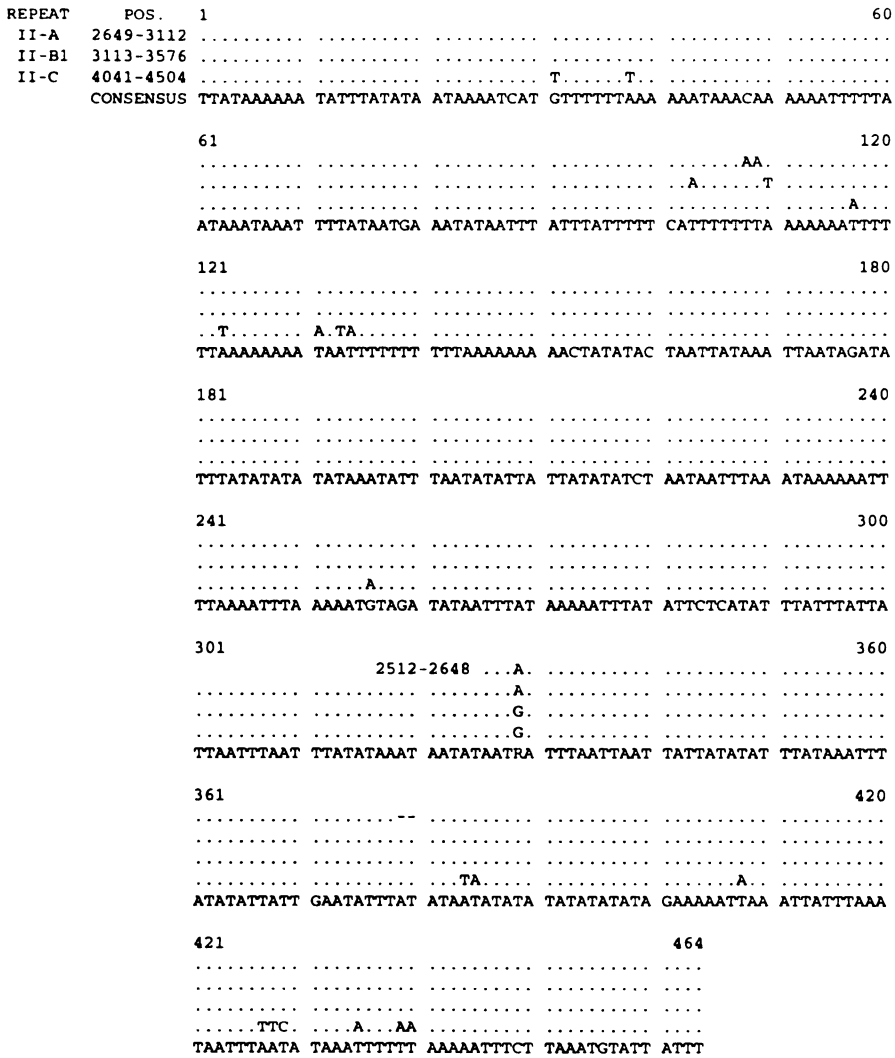

REPEAT POS. 1 60
II-A 2649-3112
II-B1 3113-3576
II-C 4041-4504
CONSENSUS TTATAAAAAA TATTTATATA ATAAAATCAT GTTTTTTAAA AAATAAACAA AAAATTTTTA
61 120
ATAAATAAAT TTTATAATGA AATATAATTT ATTTATTTTT CATTTTTTTA AAAAAATTTT
121 180
TTAAAAAAAA TAATTTTTTT TTTAAAAAAA AACTATATAC TAATTATAAA TTAATAGATA
181 240
TTTATATATA TATAAATATT TAATATATTA TTATATATCT AATAATTTAA ATAAAAAATT
241 300
TTAAAATTTA AAAATGTAGA TATAATTTAT AAAAATTTAT ATTCTCATAT TTATTTATTA
301 360
2512-2648
TTAATTTAAT TTATATAAAT AATATAATRA TTTAATTAAT TATTATATAT TTATAAATTT
361 420
ATATATTATT GAATATTTAT ATAATATATA TATATATATA GAAAAATTAA ATTATTTAAA
421 464
TAATTTAATA TAAATTTTTT AAAAATTTCT TAAATGTATT ATTT

Figure 6

Figure 7. **Comparison of the conserved DNA sequence elements in the A+T region of *D. melanogaster*, *D. yakuba*, *D. teissieri* and *D. virilis* mtDNA.** The top sequence is *D. melanogaster* (Dm) followed by that of *D. yakuba* (Dy; Clary and Wolstenholme 1985), *D. teissieri* (Dt; Monnerot et al. 1990), and *D. virilis* (Dv; Clary and Wolstenholme 1987). The consensus sequence is also shown. A dot in the sequence indicates a nucleotide that is the same in the consensus sequence. A dash indicates a nucleotide that is absent. A letter indicates a substitution. The conserved DNA sequence element is bracketed. Boxes indicate the position of the longest thymidylate stretch in each DNA strand. The *D. yakuba*, *D. teissieri* and *D. virilis* sequences are numbered according to the reports published previously. Sequence alignments and the consensus were generated as described in the legend to Fig. 5.

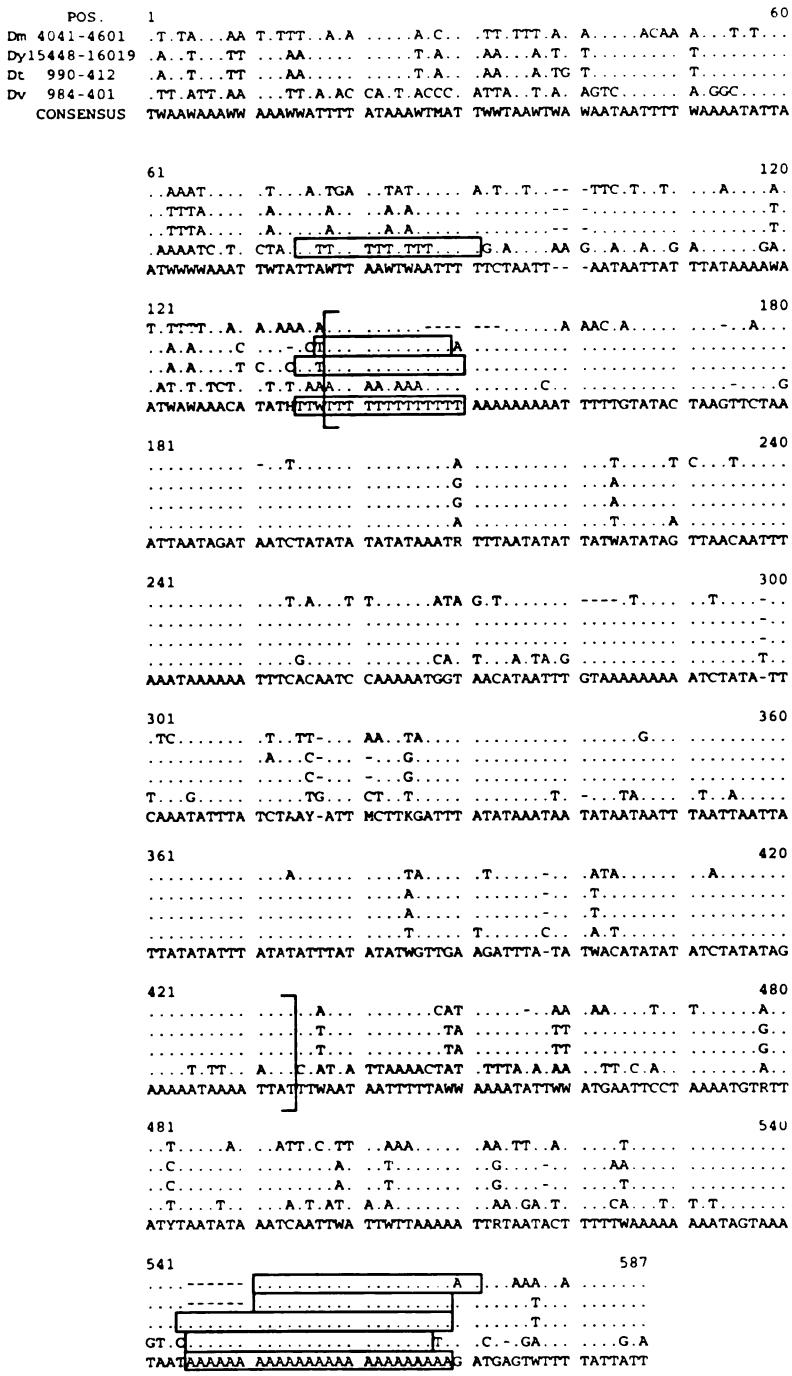

POS. 1 60
Dm 4041-4601 .T.TA...AA T.TTT..A.A .....A.C.. .TT.TTT.A. A.....ACAA A...T.T...
Dy15448-16019 .A..T...TT ...AA..... .....T.A.. .AA...A.T. T..... T.....
Dt 990-412 .A..T...TT ...AA..... .....T.A.. .AA...A.TG T..... T.....
Dv 984-401 .TT.ATT.AA ...TT.A.AC CA.T.ACCC. ATTA..T.A. AGTC..... A.GGC.....
CONSENSUS TWAAWAAAWW AAAWWATTTT ATAAAWTMAT TWWTAAWTWA WAATAATTTT WAAAATATTA
61 120
..AAAT.... .T...A.TGA ..TAT..... A.T..T..-- -TTC.T..T. ...A....A.
..TTTA.... .A.....A.. ..A.A..... .....-- -..... .....T.
..TTTA.... .A.....A.. ..A.A..... .....-- -..... .....T.
.AAAATC.T. CTA..TT..TTT.TTT....G.A....AA G..A..A..G A.....GA.
ATWWWWAAAT TWTATTAWTT AAWTWAATTT TTCTAATT-- -AATAATTAT TTATAAAAWA
121 180
T.TTTT..A. A.AAA.A..... .....---- ---.....A AAC.A..... ...-..A...
..A.A....C ...-.T..... .....A ....
..A.A....T C..C..T..... ....
.AT.T.TCT. .T.T.AAA.. AA.AAA.... .....C.. ....-....G
ATWAWAAACA TATHTTWTTT TTTTTTTTTT AAAAAAAAAT TTTTGTATAC TAAGTTCTAA
181 240
..... -..T..... .....A ....T.....T C...T.....
....G ....A..... .....
....G ....A..... .....
....A ....T.....A .....
ATTAATAGAT AATCTATATA TATATAAATR TTTAATATAT TATWATATAG TTAACAATTT
241 300
....T.A...T T.....ATA G.T..... -----T..... ..T....-..
....-..
....-..
....G..... .....CA. T...A.TA.G ....T..
AAATAAAAAA TTTCACAATC CAAAAATGGT AACATAATTT GTAAAAAAAA ATCTATA-TT
301 360
.TC..... .T..TT-... AA..TA.... ....G... .....
..... .A...C-... -...G..... ....
....C-... -...G..... ....
T...G..... .....TG... CT..T..... .....T. -...TA.... .T..A.....
CAAATATTTA TCTAAY-ATT MCTTKGATTT ATATAAATAA TATAATAATT TAATTAATTA
361 420
....A..... ....TA.... .T.....-.. .ATA..... ..A.....
....A..... .....-.. .T..... .....
....A..... .....-.. .T..... .....
....T..... T.....C.. .A.T..... .....
TTATATATTT ATATATTTAT ATATWGTTGA AGATTTA-TA TWACATATAT ATCTATATAG
421 480
....A... .....CAT .....-..AA .AA....T.. T.....A..
....T... .....TA .....TT ....G..
....T... .....TA .....TT ....G..
....T.TT.. A...C.AT.A TTAAAACTAT .TTTA.A.AA ..TT.C.A.. .....A..
AAAAATAAAA TTATTTWAAT AATTTTTAWW AAAATATTWW ATGAATTCCT AAAATGTRTT
481 540
..T.....A. ..ATT.C.TT ..AAA..... .AA.TT..A. ....T..... .....
..C..... .....A. ..T..... ..G....-.. ...AA..... .....
..C..... .....A. ..T..... ..G....-.. ....T..... .....
..T....T.. ...A.T.AT. A.A..... ..AA.GA.T. ...CA...T. T.T.....
ATYTAATATA AATCAATTWA TTWTTAAAAA TTRTAATACT TTTTWAAAAA AAATAGTAAA
541 587
....------ .....A ...AAA..A .....
....------ ....T... .....
....T... .....
GT.C..... .....T.. .C.-.GA... ....G.A
TAATAAAAAA AAAAAAAAAA AAAAAAAAAG ATGAGTWTTT TATTATT

Figure 7

Figure 8. **A+T regions and flanking genes in the mtDNAs of four species of *Drosophila*.** Depicted are the *Hin*d III-B fragment of *D. melanogaster* mtDNA and the corresponding DNA fragments of *D. yakuba*, *D. teissieri*, and *D. virilis*. The DNA fragments have been aligned with the common *Hin*d III site in the gene for NADH dehydrogenase subunit 2. Gene designations are as in the legend to figure 1. Arrows above the coding regions indicate the direction of transcription. The DNA replication origin region and direction of leading DNA strand synthesis in *D. melanogaster*, *D. yakuba* and *D. virilis* mtDNA (Wolstenholme et al. 1983) are indicated by bracketed arrows. The bold lines below the A+T region indicate the 300 bp conserved DNA sequence elements (see Fig. 7). Open boxes in the A+T region indicate the position of the longest thymidylate stretches on each DNA strand. In the *D. melanogaster* schematic, the type I repeats are indicated as boxes with wavy lines and the type II repeats as boxes with leftward leaning slashes.

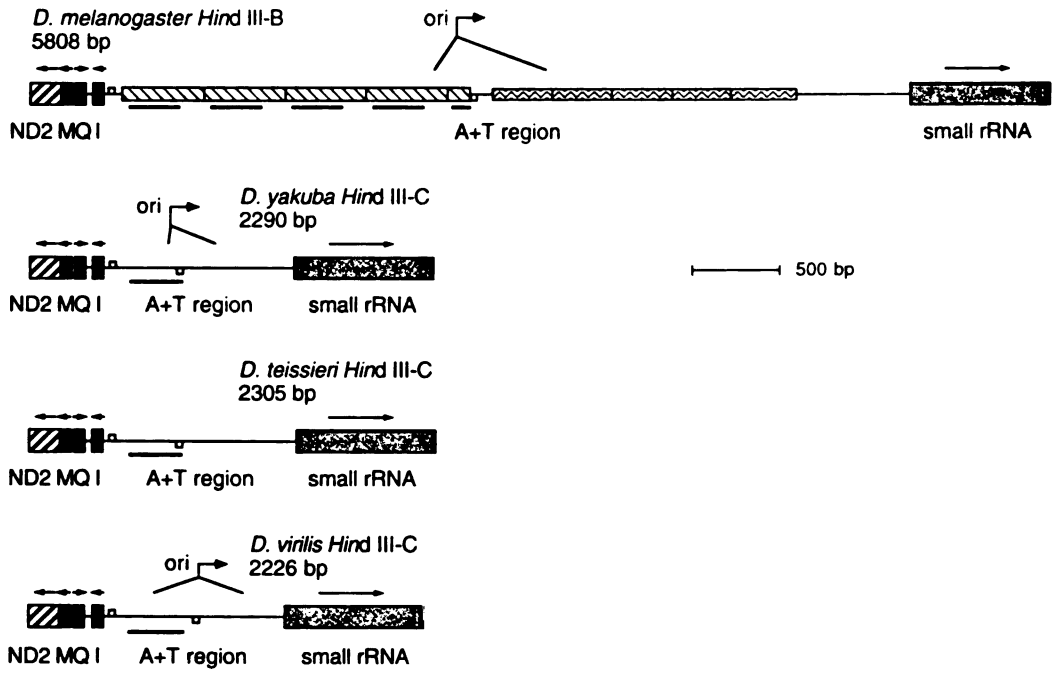

D. melanogaster Hind III-B
5808 bp
ori
ND2 MQ I
A+T region
small rRNA
ori
D. yakuba Hind III-C
2290 bp
ND2 MQ I
A+T region
small rRNA
500 bp
D. teissieri Hind III-C
2305 bp
ND2 MQ I
A+T region
small rRNA
D. virilis Hind III-C
2226 bp
ori
ND2 MQ I
A+T region
small rRNA

Figure 8

yakuba to 25 nt in *D. teissieri*, is present in the region adjacent to the gene for $tRNA^{ile}$ and corresponds to that identified previously in several species with short A+T regions (Clary and Wolstenholme 1987; Monnerot et al. 1990; Monforte et al. 1993). The second thymidylate stretch is present on the opposite DNA strand, overlapping the 300 bp conserved DNA sequence element in *D. yakuba* and *D. teissieri* and adjacent to it in *D. virilis*, and ranges in size from 13-17 nucleotides. In *D. melanogaster*, the second thymidylate stretch is 21 nt in length and is present only adjacent to the conserved element at the innermost position in the A+T region (Figure 8).

DISCUSSION

A+T Content in *Drosophila* mtDNA. *Drosophila* mtDNAs, and those of nematodes and bees, exhibit an extreme bias towards high A+T content, particularly in the noncoding region (Fauron and Wolstenholme 1976, 1980a, 1980b, Okimoto et al. 1992; Crozier and Crozier 1993). Sequence analysis indicates that the A+T region of *D. melanogaster* mtDNA contains 96% deoxyadenylate and thymidylate residues as compared to the *D. virilis* A+T region containing 90% (Clary and Wolstenholme 1987), and those of *D. yakuba* (Clary and Wolstenholme 1985b) and *D. teissieri* (Monnerot et al. 1990) containing 93%. Nucleotide sequence comparison of the conserved region shows that nucleotide substitutions between the four species are biased toward A or T in *D. melanogaster*.

Wolstenholme and Clary (1985b) have proposed that there is a continuous selection for A+T nucleotides at all sites in *Drosophila* mtDNA where it is compatible with function, and that *Drosophila* mtDNAs may have become A+T rich during long-term evolution as a result of a functional requirement (or preference) of mtDNA and/or RNA polymerase for A+T-rich DNA. We have studied extensively the mechanism of nucleotide polymerization by *D. melanogaster* mtDNA polymerase and have shown that it is a remarkably versatile enzyme with regard to template-primer usage, that it exhibits no preference for dATP or TTP as substrates, and that it is highly accurate, polymerizing nucleotides with an in vitro error rate of approximately one misincorporated nucleotide per million bases replicated (Wernette et al. 1988; Kaguni et al. 1988; Williams et al. 1993). Notably, mtDNA polymerases from both vertebrate and invertebrate sources are among the most highly accurate DNA polymerases (Kunkel 1985; Kunkel and Alexander 1986; Wernette et al. 1988; Kunkel and Mosbaugh 1989). However, the fidelity of *Drosophila* mtDNA polymerase (and that of both procaryotic and nuclear DNA polymerases) is affected greatly by in vitro reaction conditions, and in particular, by nucleotide pool biases (Wernette et al. 1988; Olson and Kaguni 1992; C.

L. Farr and L. S. Kaguni, unpublished data). Because neither deoxyribonucleotide synthesis nor the regulation of deoxynucleotide pools have been well studied in mitochondria, it is not possible to estimate the effects of nucleotide pool imbalances on the fidelity of mtDNA replication in vivo. Although the molecular mechanism responsible for the high A+T content in *Drosophila* mtDNAs remains unknown, it will be of interest to evaluate further the possible correlation between A+T content and A+T region length that is suggested by our current data.

Evolution of the A+T Region Repeat Arrays. The A+T region of *D. melanogaster* mtDNA contains two types of tandemly repeated elements: the type I repeats located in the variable region, and the type II repeats in the conserved region. The presence of tandemly repeated elements in the noncoding region of mtDNA appears to be ubiquitous in the animal kingdom; it has been documented in other invertebrates including sea scallops (Snyder et al. 1987; La Roche et al. 1990; Gjetvaj et al. 1992), cricket (Rand and Harrison 1989), bark weevils (Boyce et al. 1989), honeybee (Cornuet et. al. 1991) and nematodes (Okimoto et al. 1991), and in vertebrates including lizards (Densmore et al. 1985), monkeys (Karawya and Martin 1987; Hayasaka et al. 1991), American shad (Bentzen et al. 1988), birds (Avise and Zinc 1988), sturgeon (Buroker et al. 1990), Atlantic cod (Johansen et al. 1990), rabbit (Mignotte et al. 1990), evening bat (Wilkinson and Chapman 1991), seals (Arnason and Johnsson 1992; Hoelzel et al. 1993; Arnason et al. 1993) and pig (Ghivizzani et al. 1993b).

The 464 bp type II repeats are present in four copies and are located in the half of the *D. melanogaster* A+T region adjacent to the gene for $tRNA^{ile}$. By DNA restriction site analysis, repeats of nearly identical size were identified in three other species in the *melanogaster* subgroup, *D. simulans, D. sechellia* and *D. mauritiana* (Solignac et al. 1986). However, recognition sites for the enzymes used in that study (*Acc* I and *Hpa* I) are not present in the *D. melanogaster* repeats. The data presented here establish the

presence of the repeats in *D. melanogaster* mtDNA, and indicate that a single nucleotide substitution propagated throughout the repeats resulted in the loss of the *Acc* I restriction sites found in the repeats of other *melanogaster* subgroup species. Two copies of a repeat containing an *Acc* I restriction site have also been reported in *D. tristis*, a member of the *obscura* species subgroup (Monforte et al. 1993), providing evidence for convergent evolution of repeated elements in *Drosophila* mtDNA.

The type I repeats in *D. melanogaster* mtDNA are present in five copies and occur in the variable half of the A+T region near the gene for the small rRNA. They differ in size from 338 to 373 bp. The 373 bp type I-A repeat contains additional AT dinucleotide repeats that account for most of the size difference. Based on the length of the variable region in eight species of the *melanogaster* subgroup, the presence of repeated elements was predicted but not confirmed, again due to the lack of indicator restriction enzyme sites (Solignac et al. 1986). In fact, earlier heteroduplex mapping studies of *D. melanogaster* mtDNA revealed pairs of deletion loops in this region, indicative of repeated DNA sequence elements (Merten and Pardue 1981). Interestingly, length polymorphisms in the variable region are also found in members of the *obscura* group (Monforte et al. 1993). Based on our findings in *D. melanogaster* and the documented length polymorphism of the region among *Drosophila* species, it seems likely that the organization if not the sequence of the variable half of the A+T region is conserved among species. That is, length polymorphisms are likely due to the presence of tandemly repeated elements whose sequence is not conserved among species. Indeed, heteroduplex studies of the A+T regions of *D. melanogaster*, *D. simulans* and *D. mauritiana* mtDNA revealed base pairing only in the region of the type II repeats (Fauron and Wolstenholme 1980a, 1980b). Thus, the mechanism of selection of mitochondrial genomes with variable region repeats may be dependent on DNA secondary structure rather than on primary sequence determinants. Further, based on our data and on DNA restriction site analyses of the A+T region in mtDNAs from both

the *melanogaster* (Solignac et al. 1986) and *obscura* groups (Monforte et al. 1993), it seems likely that expansion (or contraction) of both the conserved and variable halves of the A+T region occur in concert, as a result of duplication (or deletion) of at least two types of repeated DNA sequence elements.

The structural organization of the type I repeat array in *D. melanogaster* mtDNA indicates that it is the product of at least two separate events involving different duplication endpoints. The first involved the duplication of sequence in an ancestral fly leading to the type I repeat array ABC. A subsequent event, duplicating approximately one-third of repeat A and complete copies of B and C, resulted in the repeat structure A B C/A B C observed in *D. melanogaster*. The latter duplication event and subsequent deletion events probably occur at a relatively high frequency, given the 100% identity of the two type I-B repeats and the >99% identity of the repeated portions of the type I-A and I-C repeats.

Although the molecular mechanisms involved in the generation of repeated DNA sequence elements remain undetermined, both replication slippage (Streisinger et al. 1966) and homologous recombination could yield the observed duplications. Homologous recombination has been demonstrated in fungal mtDNA (Clark-Walker 1989; Almasan and Mishra 1991), but has not been documented in animal mtDNA. Alternatively, replication slippage would be promoted if the displaced portion of a nascent DNA strand could form a secondary structure, which would further stabilize the misaligned strand on the template DNA. In this regard, DNA capable of forming a cruciform structure has been identified in repeats in scallop mtDNA (La Roche et al. 1990), and a short inverted DNA sequence was identified in the cricket mtDNA repeat (Rand and Harrison 1989). Computer analysis reveals extensive intrastrand base pairing is also possible within the repeats of *D. melanogaster* mtDNA (D. L. Lewis and L. S. Kaguni, unpublished data). In either mechanism, recurring duplications and deletions would result in the homogenization of the repeated DNA sequences. Indeed, the high

nucleotide sequence similarity of the repeated DNA sequence elements in *D. melanogaster* provides evidence that such a process occurs in *Drosophila* mtDNA.

Conserved DNA Sequence Elements in the A+T Region. The type II repeats in *D. melanogaster* mtDNA contain a 300 bp region highly conserved in several *Drosophila* species with short A+T regions (Clary and Wolstenholme 1987; Monnerot et al. 1990; Monforte et al. 1993). Although the function of the conserved region in mtDNA metabolism has not been established, it appears to be involved in site-specific protein : DNA interactions. The 300 bp conserved DNA sequence elements identified here in the *D. melanogaster* A+T region correspond in number, size and location to DNA sequence elements that were protected from trimethylpsoralen crosslinking *in situ* and mapped by electron microscopy (Potter et al. 1980). Likewise, protection from crosslinking was also observed in the single 300 bp conserved region in *D. virilis* mtDNA (Pardue et al. 1984). Further, the DNA replication origin in *D. melanogaster*, *D. yakuba* and *D. virilis* mtDNA overlaps the conserved DNA sequence element (Wolstenholme et al. 1983; Figure 8). Notably, only the innermost repeat of the conserved element in the *D. melanogaster* A+T region maps to the origin region, suggesting an additional DNA sequence determinant is required for DNA replication origin function. In this regard, we would propose that the conserved thymidylate stretch adjacent to the conserved DNA sequence element is involved in *Drosophila* mtDNA replication. Interestingly, thymidylate stretches have been demonstrated in nuclear and viral systems to function both as core elements in DNA replication origins and as transcriptional activators (Campbell 1986; Delucia et. al., 1986).

Although the relationship between transcription and replication remains to be elucidated in *Drosophila* mtDNA, the positions and orientations of the two conserved thymidylate stretches suggest roles in promoting both transcription and replication. Unidirectional promoters for transcription of each DNA strand have been identified in mammalian mtDNA (Chang and Clayton 1986; Topper and Clayton 1989), and a role

has been proposed for the light-strand promoter in the production of an RNA that is processed to yield a primer for initiation of mtDNA replication (Clayton 1991b). Indeed, if the thymidylate stretch located near the origin of DNA replication is necessary for origin function as proposed above, its apparent downstream position relative to leading DNA strand synthesis would rule out an involvement in RNA primer synthesis. Alternatively, it could function by another mechanism such as transcriptional activation, implicated in a variety of non-mitochondrial systems. The lack of DNA sequence conservation in the regulatory regions of *Drosophila* and vertebrate mtDNAs may in fact reflect a difference in the mechanisms of DNA replication and/or transcriptional initiation and its regulation.

CHAPTER III

Drosophila melanogaster MITOCHONDRIAL DNA: COMPLETION OF THE NUCLEOTIDE SEQUENCE AND EVOLUTIONARY COMPARISONS

INTRODUCTION

The metazoan mitochondrial DNA (mtDNA) occurs as a double stranded, circular molecule ranging in size from 14 to 39 kb (Moritz et al. 1987; Snyder et al. 1987). The complete nucleotide sequence of the mtDNA molecule has been determined for several mammalian species: human, cow, mouse, rat, fin whale, harbor seal, blue whale and grey seal (Anderson et al. 1981, 1982; Bibb et al. 1981; Gadaleta et al. 1989; Arnason et al. 1991; Arnason and Johnsson 1992; Arnason and Gullberg 1993; Arnason et al. 1993): an amphibian, *Xenopus laevis* (Roe et al. 1985): a bird, *Gallus domesticus* (Desjardins and Morais 1990): a fish, *Crossostoma lacustre* (Tzeng et al. 1992): two sea urchins, *Strongylocentrotus purpuratus* and *Paracentrotus lividus* (Jacobs et al. 1988; Cantatore et al. 1989): and two nematodes, *Caenorhabditis elegans* and *Ascaris suum* (Okimoto et al. 1992). The nucleotide sequence has also been determined for the mitochondrial genomes of two insect species, *Drosophila yakuba* (Clary and Wolstenholme 1985b) and *Apis mellifera* (Crozier and Crozier 1993).

All metazoan mtDNAs contain the genes for 12-13 polypeptides involved in oxidative phosphorylation as well as the 22 tRNAs and two rRNAs of the mitochondrial translational apparatus (Moritz et al. 1987; Okimoto et al. 1992). Gene arrangement varies somewhat among vertebrates, with that of birds differing from that of mammals and amphibians (Desjardins and Morais 1990). Gene arrangement in the mtDNA of invertebrates varies more substantially. Unique arrangements as a result of multiple inversions and translocations are evident when comparisons are made between the mitochondrial genomes of insects, sea urchins and nematodes (Clary and Wolstenholme 1985b; Crozier and Crozier 1993; Jacobs et al. 1988; Cantatore et al. 1989). Among insects, a segment containing three tRNAs varies in position between *Drosophila* species and honeybees (Clary and Wolstenholme 1985b; Garesse 1988; Crozier et al. 1989).

Drosophila mtDNA has an unusually high A+T content (Moritz et al. 1987). Extreme A+T bias is especially evident in the major noncoding region, called the A+T-rich region because it contains 90-96% deoxyadenylate and thymidylate residues (Fauron and Wolstenholme 1976). The A+T region varies greatly in size from 1 to 5 kb among species and may vary among populations in a given species or in a single fly (Fauron and Wolstenholme 1976,1980a, 1980b; Solignac et al. 1983, 1986; Hale and Singh 1986; Monforte et al. 1993; Pissios and Scouras 1993). A+T region length polymorphisms are in part due to the presence of varying copy numbers of repeated elements (Solignac et al. 1986; Monforte et al. 1993; Chapter II). In *D. melanogaster*, two types of tandemly repeated elements are present: five copies of a 338 to 373 bp repeat element located in the part of the A+T region, displaying a high substitution frequency between different species, and four copies of a 464 bp repeated element located in the part of the A+T region containing conserved sequence elements (Chapter II).

Segments of the *D. melanogaster* mitochondrial genome have been sequenced previously (Clary et al. 1982; Clary et al. 1983; de Bruijn 1983; Satta et al. 1987; Garesse 1988; Satta and Takahata 1990; Chapter II) and the cDNA sequence of the 16S rRNA reported (Benkel et al 1988; Kobayashi and Okada 1990). Nucleotide sequence comparisons between *D. yakuba* and *D. virilis* and between *D. melanogaster* and *D. yakuba* have revealed that in *Drosophila* mtDNA transversions and transitions occur at approximately equal frequencies (Clary and Wolstenholme 1987; Garesse 1988). This contrasts with the pattern of nucleotide substitutions in vertebrate mtDNA where transitions greatly outnumber transversions (Moritz et al. 1987) and may reflect a functional constraint of *Drosophila* mtDNA to maintain a high deoxyadenylate and thymidylate content thereby limiting the number of hypervariable sites (DeSalle et al. 1987). Indeed, nucleotide comparisons of more closely related

Drosophila species reveal a bias towards transitions (de Bruijn 1983; DeSalle et al. 1987; Satta et al. 1987; Tamura 1992).

In this chapter, the nucleotide sequences of the genes encoding four tRNAs and the 12S rRNA, and the 5' ends of the genes encoding NADH dehydrogenase subunit II (ND-2) and 16S rRNA in the *D. melanogaster* mtDNA are reported, thus completing the nucleotide sequence for the *D. melanogaster* mitochondrial genome.

EXPERIMENTAL PROCEDURES

Materials

Enzymes and Chemicals. Restriction endonucleases were from Gibco BRL and New England Biolabs. T4 DNA ligase was from New England Biolabs and Sequenase version 2.0 was from United States Biochemical. 5-bromo-4-chloro-3-indolyl-beta-D-galactoside (X-gal) and isopropyl-beta-D-thiogalactopyranoside (IPTG) were from Research Organics. Low melting point agarose was obtained from FMC Bioproducts.

Nucleic Acids and Nucleotides. Forward and reverse universal M13/pUC DNA sequencing primers were synthesized by the Macromolecular Structure and Synthesis Facility at Michigan State University using an Applied Biosystems model 477 oligonucleotide synthesizer. The vector pNEB 193 was purchased from New England Biolabs. [α-^{35}S]dATP was from New England Nuclear.

Methods

Cloning and Sequencing of the D. melanogaster *mtDNA. D. melanogaster* (Oregon R) mtDNA was extracted from embryos as described previously (Chapter II). The *D. melanogaster* mtDNA *Hin*d III-B fragment containing the A+T region, the genes encoding the 12S rRNA, $tRNA^{ile}$, $tRNA^{gln}$, $tRNA^{f\text{-}met}$ and the 5' coding regions of $tRNA^{val}$ and NADH dehydrogenase subunit II was cloned as *Pac* I fragments or *Ssp* I fragments as described previously (Chapter II). In order to obtain clones containing the 3' coding region of $tRNA^{val}$ and the 5' coding region of the 16S rRNA, *D. melanogaster* mtDNA was digested using the *Hae* III restriction enzyme and the *Hae* III-A fragment purified by gel electrophoresis in low melting point agarose. An *Xba* I restriction digest of the *Hae* III-A fragment was prepared and ligated using T4 DNA ligase into the vector pNEB 193 digested with *Xba* I and *Hinc* II. The *E. coli* strain Sure

(Stratagene) was used in transformation procedures, and recombinants screened on plates containing X-Gal and IPTG at 0.05% and 0.4 mM final concentration, respectively. Colonies harboring recombinant DNA were screened for insert size and orientation of the recombinant DNA by restriction analysis. DNA sequencing was performed using the dideoxy chain termination method of Sanger et al. (1977) and Sequenase version 2.0. Sequence analyses were performed using the GCG Package, Version 7, April 1991, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711.

RESULTS AND DISCUSSION

Genome Structure

The mtDNA sequence presented in this chapter overlaps the previously published *D. melanogaster* sequence at the conserved *Hin*d III site in the gene encoding NADH dehydrogenase subunit 2 (ND2) (de Bruijn et al.) and is contiguous with the sequence reported by Garesse (1988) beginning in the 5' coding region of the 16S rRNA. The mitochondrial genome of *D. melanogaster* is represented schematically in Figure 1. The circular mtDNA is 19,517 bp in length as compared to 16,019 bp in *D. yakuba* mtDNA (Clary and Wolstenholme, 1985b). The relative arrangement and orientation of the genes encoded on both mtDNAs are identical. The coding regions of *Drosophila* mitochondrial genomes are highly homologous with few insertions/deletions. In contrast, the noncoding A+T region varies greatly in both size and nucleotide sequence (Fauron and Wolstenholme 1976). The relatively large A+T region of *D. melanogaster* mtDNA is 4601 bp in length (Chapter II) as compared to 1077 bp in *D. yakuba* (Clary and Wolstenholme, 1985b), 1091 bp in *D. teissieri* (Monnerot et al. 1990) and 1029 bp in *D. virilis* mtDNA (Clary and Wolstenholme, 1987). The large size of the *D. melanogaster* A+T region is almost entirely due to the presence of two types of tandemly repeated DNA sequence elements (Chapter II). Overall, the *D. melanogaster* mitochondrial genome has an A+T content of 82.2% versus 78.6% in *D. yakuba* (Clary and Wolstenholme 1985b). Of the mitochondrial genomes completely sequenced, only that of the honeybee, *Apis mellifera* (Crozier and Crozier 1993) has a higher A+T content (84.9%) than that of *D. melanogaster*.

Nucleotide Substitutions

NADH dehydrogenase subunit 2. In Figure 2, the *D. melanogaster* mtDNA sequences flanking the A+T region including the complete coding sequences for the

Figure 1. **Structure and genetic map of *D. melanogaster* mitochondrial DNA.** The replication origin is shown as Ori and the direction of leading strand replication is denoted by the arrow. The control region is labeled A+T region (shaded). Transfer RNA genes (hatched) are indicated by the one-letter amino acid code and individual tRNA codon recognition sequences for the tRNAs specifying serine and leucine are also shown. 12S rRNA, coding region for the small subunit rRNA; 16S rRNA, coding region for the large subunit rRNA; ND1, ND2, ND3, ND4, ND4L, ND5, ND6, coding regions for six subunits of the NADH dehydrogenase complex; cyt b, coding region for cytochrome b; COI, COII, COIII, coding regions for cytochrome oxidase subunits; ATPase 6, ATPase 8, coding regions for ATPase subunits. The arrow associated with each coding region indicates the direction of transcription. The locations of the DNA sequences cloned in this Chapter and in Chapter II are indicated by the arc.

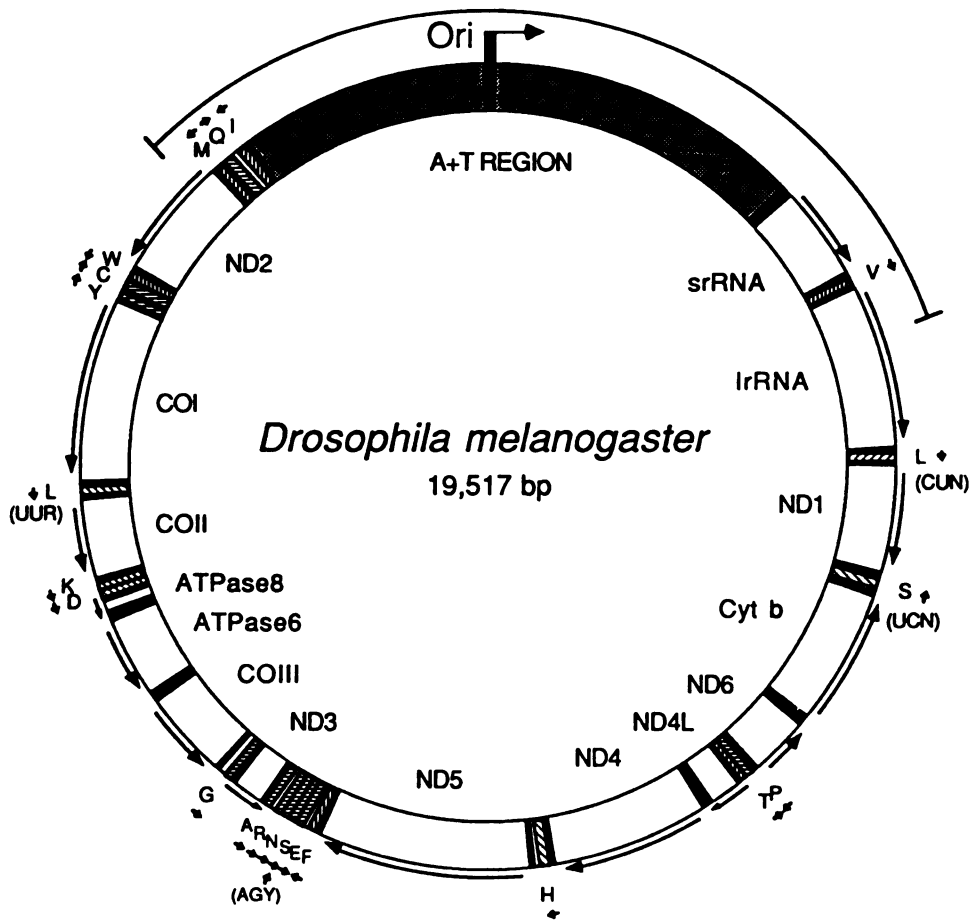

Ori
A+T REGION
srRNA
lrRNA
ND1
L (CUN)
S (UCN)
Cyt b
ND6
ND4L
ND4
ND5
H
TP
ARNSEF
(AGY)
G
ND3
COIII
ATPase6
ATPase8
K
D
COII
L (UUR)
COI
YC
W
ND2
Q
M
I
V
Drosophila melanogaster
19,517 bp

Figure 1

Figure 2. **Nucleotide sequence comparison of A+T region flanking genes of four species of *Drosophila*.** The DNA sequences encoding the $tRNA^{ile}$, $tRNA^{gln}$, $tRNA^{f\text{-}met}$, the 5' 13 nucleotides of $tRNA^{val}$, the 12S rRNA and the 5' 169 nucleotides of ND2 are compared between *D. melanogaster* (Dm), *D. yakuba* (Dy), *D. teissieri* (Dt) and *D. virilis* (Dv). The 5' 439 nucleotides of the gene encoding the 16S rRNA and 3' 59 nucleotides of $tRNA^{val}$ are compared between *D. melanogaster* and *D. yakuba* only. Each nucleotide sequence is numbered beginning at the 5' nucleotide of that sequence. Genes encoding tRNAs are boxed by solid lines and genes encoding rRNAs are boxed by broken lines. An arrow indicates the direction of transcription of each gene. The central A+T region of each species mtDNA sequence is indicated by bracketed arrows. The consensus sequence is shown (Con). A dot in the individual mtDNA sequences indicates a nucleotide that is the same as in the consensus sequence and a dash indicates a nucleotide that is absent. The predicted amino acid sequence of the 5' 169 nucleotide sequence of the gene encoding ND2 of *D. melanogaster* is shown and the predicted amino acid substitutions in *D. yakuba*/*D. teissieri* and *D. virilis* are shown in parentheses. Nucleotide differences in the sequence of a cDNA clone of the *D. melanogaster* 16S rRNA are shown in parentheses above the genomic sequence of the *D. melanogaster* 16S rRNA.

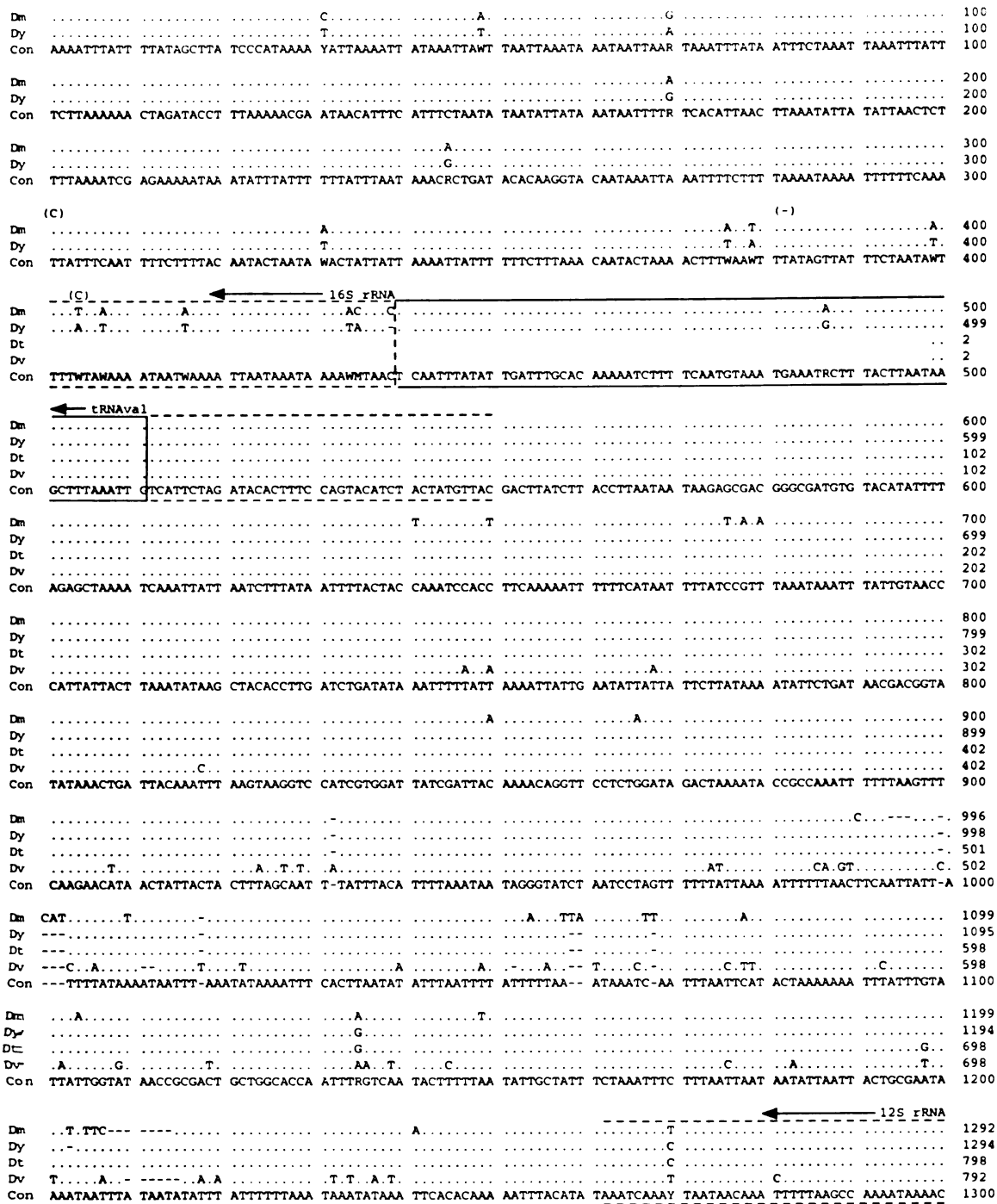
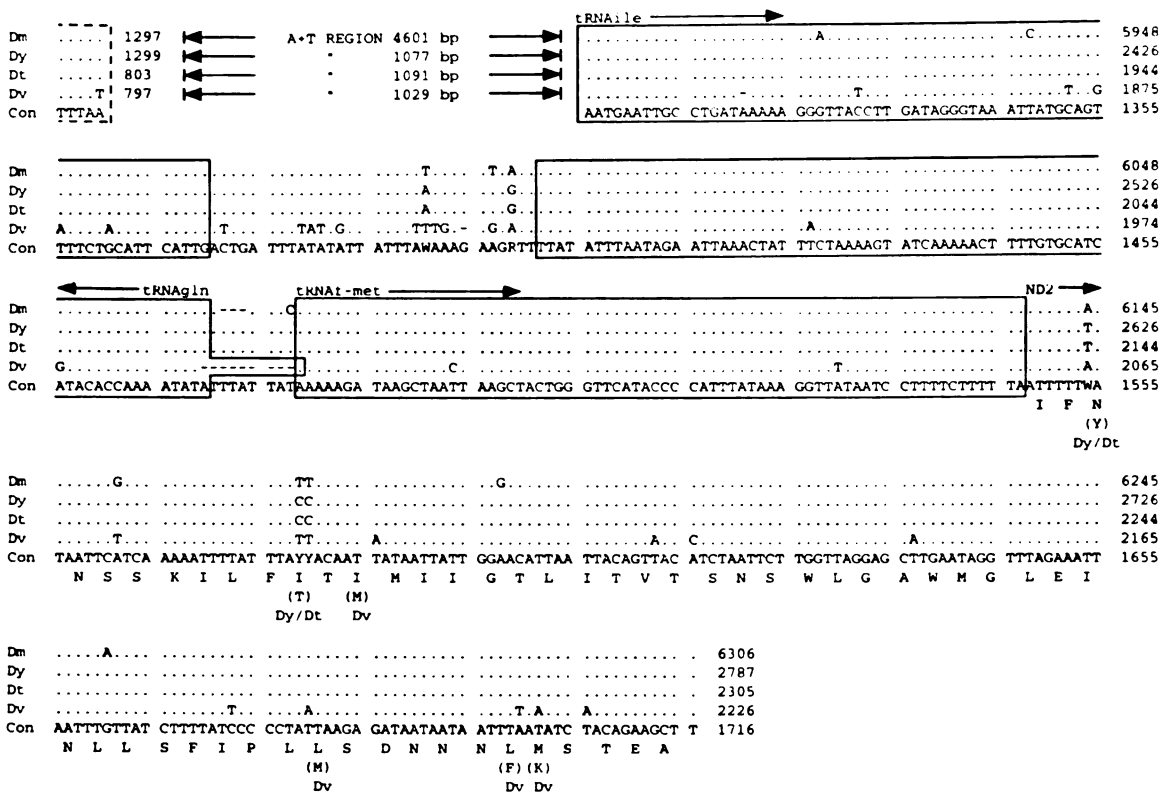


Figure 2


tRNAile
A+T REGION
4601 bp
1077 bp
1091 bp
1029 bp
tRNAgln
tRNAt-met
ND2
Dm
Dy
Dt
Dv
Con

12S rRNA and the tRNAs specifying valine, isoleucine, glutamine and methionine as well as the 5' coding regions for ND2 and the 16S rRNA are compared to those of other *Drosophila* species whose corresponding sequences are known. The nucleotide substitution frequencies between the mitochondrial sequences of different species are given in Table 1. The 5' 169 nucleotides of the gene encoding ND2 is less divergent (3.6%) than the remainder of the gene (7.6%) as revealed by nucleotide sequence comparison between corresponding regions of *D. melanogaster* and *D. yakuba / D. teissieri*. The corresponding 5' region of the *D. virilis* ND2 gene contains a greater percentage of substitutions when compared to that of *D. melanogaster* (7.1%) than that of *D. yakuba / D. teissieri* reflective of the greater time of divergence from *D. melanogaster*. The majority of substitutions are conservative changes in the third position of codons as has been observed in comparisons of other *Drosophila* mitochondrial protein coding genes (Wolstenholme and Clary 1985b; Clary and Wolstenholme 1987; Garesse 1988).

Mitochondrial large subunit ribosomal RNA. The gene encoding the 16S rRNA of *D. melanogaster* contains 1325 bp of which 82.9% are deoxyadenylate and thymidylate residues, which is slightly more A+T-rich than the mtDNA coding region as a whole (78.1%). The complete nucleotide sequence of a cDNA for the gene has been determined (Kobayashi and Okada 1990) and differs in three positions from the 439 bp genomic sequence of the 5' coding region presented here (Fig. 2). The differences are likely to be strain dependent as in other mtDNA sequences in *D. melanogaster* (de Bruijn et al. 1983). Nucleotide substitution frequencies of the 16S rRNA genes of *D. melanogaster* and *D. yakuba* mtDNA are given in Table 1. Overall, nucleotide sequence divergence is 3.2%, 69.8% of which are A↔T transversions. In the 5' coding region of the gene encoding the 16S rRNA presented here, there is a single insertion of a deoxycytidylate residue at the 5' terminus of the *D. melanogaster* gene. The cDNA sequence of the *D. melanogaster* 16S rRNA also contains the

Table 1. **Nucleotide substitution frequencies among corresponding genes encoding 16S rRNA, 12S rRNA, four tRNAs AND ND2 of *D. melanogaster*, *D. yakuba* and *D.virilis***

	Gene (Coding Strand)					
	16S rRNA[a]	12S rRNA		tRNAs		ND2[b]
	yakuba	*yakuba*	*virilis*	*yakuba*[c]	*virilis*[d]	*yakuba*
Transitions						
A↔G	7	7	17	1	2	14
C↔T	4	3	5	2	5	18
Total	11	10	24	3	7	32
Transversions						
A↔C	0	0	0	0	0	6
A↔T	30	10	37	0	3	33
G↔T	1	2	1	0	1	0
G↔C	0	0	0	0	0	0
Total	31	12	38	0	4	39
Total substitutions	42	22	62	3	11	71
Total nucleotides (*D. melanogaster*)	1325	786		275	203	1023
Nucleotide sequence divergence	3.2%	2.8%	7.9%	1.1%	5.4%	6.9%

[a] Sequence data for the 3' portion of the *D. melanogaster* 16S rRNA gene was taken from Garesse (1988).

[b] Sequence data for the 3' coding region of the *D. melanogaster* ND2 gene was taken from Wolstenholme and Clary (1985).

[c] Data includes tRNAs for isoleucine, glutamine, methionine and valine.

[d] Data includes tRNAs for isoleucine, glutamine and methionine.

insertion and it is therefore likely to be present on the mature RNA in this species (Kobayashi and Okada 1990).

The secondary structure proposed for the *Drosophila yakuba* 16S rRNA is based on nucleotide sequence comparisons of homologous rRNAs (Gutell and Fox 1988). The substitution pattern observed between the *D. melanogaster* and *D. yakuba* genes is consistent with the secondary structure model. No substitutions occur in the proposed "peptidyl transferase center" in the 3' portion of the molecule underscoring the importance of both sequence and structure here (Noller et al. 1992). Indeed, nucleotides protected by A and P site tRNA from chemical cleavage in the corresponding region of *E. coli* 23S rRNA (Noller and Moazed 1989) are conserved in *D. melanogaster* and *D. yakuba* 16S rRNA.

Mitochondrial small subunit ribosomal RNA. The gene encoding the 12S rRNA of *D. melanogaster* contains 786 bp of which 80.2 % are deoxyadenylate and thymidylate residues, which is slightly less A+T-rich than the 16S rRNA (82.9%) but higher than the mtDNA coding region as a whole (78.1%). The mtDNA sequence of the 12S rRNA has also been determined for *D. yakuba* (Clary and Wolstenholme 1985b), *D. teissieri* (Monnerot et al. 1990) and *D. virilis* (Clary and Wolstenholme 1987). The nucleotide sequence comparison shown in Figure 2 reveals that substitutions occur mainly in the 5' 350 bp of the coding sequence. All occurrences of insertions/deletions have also arisen in this portion of the gene. Frequencies of nucleotide substitutions are given in Table 1. Overall, the nucleotide sequence of the 12S rRNA gene is 2.8% divergent from that of *D. yakuba* and 7.9% divergent from that of *D. virilis*. The *D. teissieri* sequence differs from that of *D. yakuba* only at position 697 (T↔C) and 698 (insertion/deletion).

A secondary structure model of the *D. yakuba* 12S rRNA has been constructed based on that of the mouse 12S rRNA (Clary and Wolstenholme 1985a) and other species (Neefs et al. 1993) and is presented in Figure 3. Effects of nucleotide

Figure 3. **Secondary structure of the *D. yakuba* mitochondrial SSU rRNA and differences in the *D. melanogaster* mitochondrial SSU rRNA gene.** The *D. yakuba* sequence is shown in the proposed secondary structure. Large letters near filled arrows indicate nucleotide substitutions. An X indicates a deletion in the *D. melanogaster* sequence. Large letters near winged arrows indicate insertions in the *D. melanogaster* sequence. Helices are numbered as described in the text. The *D. yakuba* secondary structure model was kindly provided by Dr. Rupert de Wachter with updates by Birgitta Winnepenninckx, Antwerp University, Belgium.

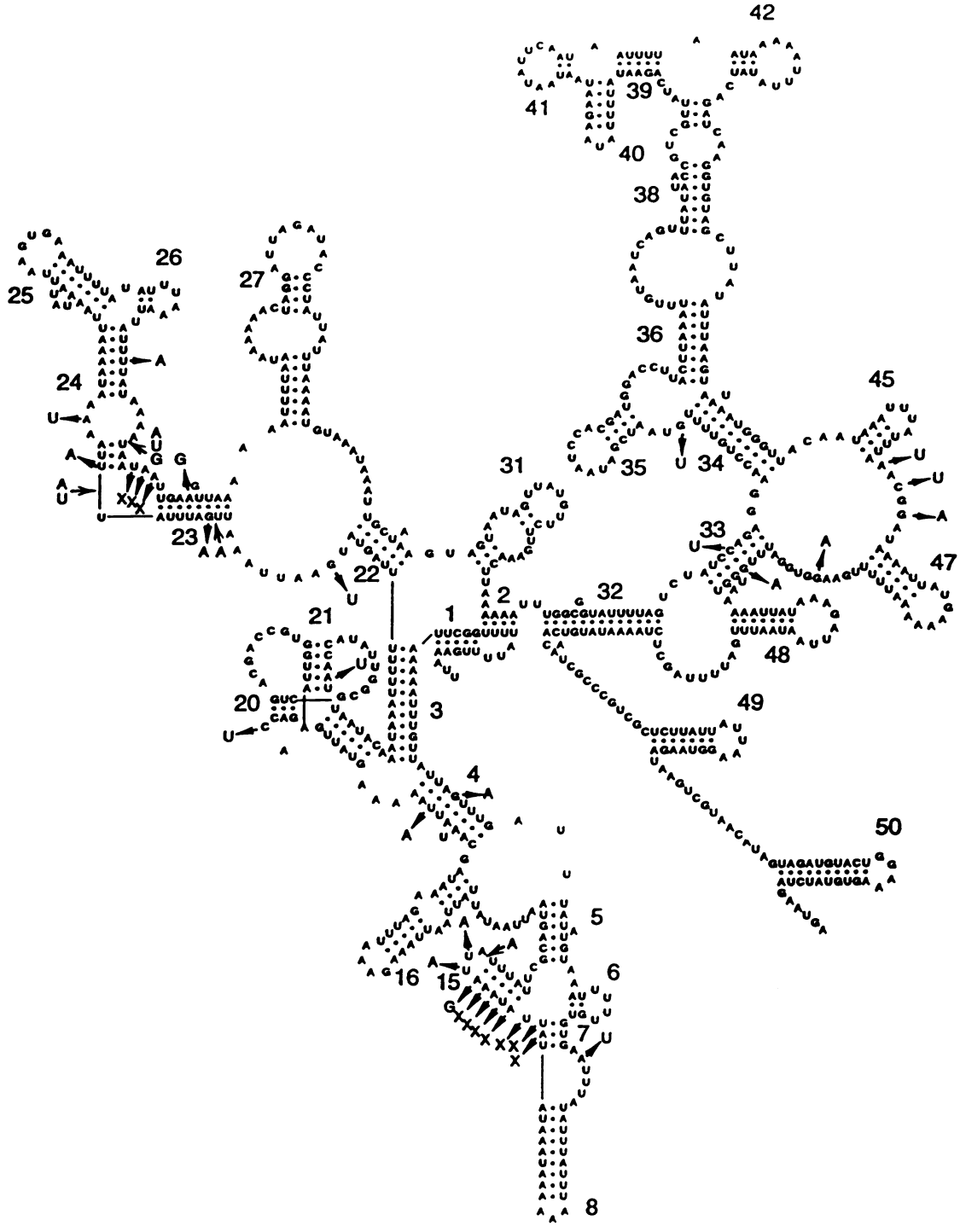

42
41
39
40
38
36
45
35
34
31
33
47
32
2
1
22
21
48
20
3
49
4
50
5
16
15
6
7
8
26
25
27
24
23

Figure 3

substitutions in the *D. virilis* molecule on the secondary structure proposed for that of *D. yakuba* were examined previously (Clary and Wolstenholme 1987). Nucleotide changes in the 12S rRNA sequence between *D. melanogaster* and *D. yakuba* occur in three main clusters when viewed in secondary structure. The first cluster is present in or near helices 7 and 15 (numbers according to Neefs et al. 1993) and consists of a deletion of 7 nucleotides, an insertion and three substitutions in the *D. melanogaster* sequence. Overall, the deletions and substitutions destroy four and one base pairings, respectively, and would result in the disruption of helices 7 and 15. A number of deletions occur in the corresponding segment of *D. virilis* (Clary and Wolstenholme 1987) and the only nucleotide differences found in the 12S rRNA genes of *D. yakuba* and *D. teissieri* also occur in this region (Monnerot et al. 1990). The region containing these mutations varies greatly among mitochondrial 12S rRNAs of other metazoans. According to the current secondary structure model, the corresponding region in the *C. elegans* rRNA contains only a 46 bp loop (Gutell 1993). In *Homo sapiens,* the region contains a larger number of nucleotides than either *C. elegans* or *D. melanogaster* concomitant with a number of stem-loop structures (Neefs et al. 1993). The region in *D. melanogaster* 12S rRNA seems to be intermediate in regard to size and secondary structure content.

The second cluster of nucleotide substitutions occurs in the region of helices 23 and 24 in the proposed secondary structure. A total of five base substitutions, one insertion each of one, two and three nucleotides and a deletion of three nucleotides occur in this segment of the *D. melanogaster* sequence as compared to that of *D. yakuba*. In helix 23, a single nucleotide bulge (U) is created while changing a G·U wobble pairing to an A-T Watson-Crick pairing. The remaining substitution in helix 23 changes an A-T Watson-Crick pairing to a G·U wobble pairing. The unnumbered three bp helix between helix 23 and 24 in the *D. yakuba* molecule could be maintained in the *D. melanogaster* molecule by base pairing the U and G nucleotides in the proximal

insertion to the A and U nucleotides, respectively, on the opposite strand of the helix. The third base pair of the helix would consist of the substituted A nucleotide and the U nucleotide formerly paired with the A at the opposite end of the helix. Only the substitution in helix 24 disrupts a base pairing. Thus, all changes are compatible with the overall secondary structure proposed for the region with only minor variations in the number of nucleotides contained in internal loops.

The third cluster of nucleotide substitutions occurs in the 3' region of the 12S rRNA. None of the seven nucleotide substitutions in the cluster affect base pairing: two are covariant in helix 33 and the remainder occur in loops. Most substitutions in this region occur in a secondary structure presented as a large multibranched loop in current models of SSU rRNAs in organisms as diverse as *E. coli* and *Homo sapiens* (Neefs et al. 1993).

The *Drosophila* 12S rRNA gene encodes conserved nucleotides shown to be protected by A and P site bound tRNA in studies using the *E. coli* SSU rRNA (Noller 1991). In the *E. coli* SSU rRNA 530 loop, the GGT sequence (positions 529-531 in the *E. coli* sequence) is protected from chemical cleavage by A site tRNA (Noller 1991). In addition, substitution of A for G529 results in impaired EF-Tu dependent binding of aminoacyl-tRNA in vitro, suggesting that the defect involves a function related to EF-Tu (Powers and Noller 1993). Protection from chemical cleavage by P-site tRNA was noted at A532 (Noller 1991) which is also conserved in known *Drosophila* sequences. In addition, many of the nucleotides in the *E. coli* SSU rRNA flanking helix 49 that are protected by A-site and P-site tRNAs are conserved in the *Drosophila* SSU rRNA. The fact that most of the protected nucleotides are invariant between *E. coli* and *D. melanogaster* mt-rRNAs suggests the basic molecular machinery that drives translation is highly similar.

Mitochondrial transfer RNAs. The segment of the mitochondrial genome sequenced in this work includes four of the 22 tRNAs encoded for on the mtDNA,

specifically $tRNA^{val}$, $tRNA^{ile}$, $tRNA^{gln}$ and $tRNA^{f\text{-}met}$. The $tRNA^{gln}$ and $tRNA^{ile}$ genes, which are transcribed in opposite directions, are separated by five and eight nucleotides in *D. melanogaster* and *D. yakuba* mtDNA, respectively, but overlap by one nucleotide in the *D. virilis* mtDNA molecule (Clary and Wolstenholme 1987). As shown in Table 1 and Figure 2, the four tRNAs of *D. melanogaster* are 1.1% divergent from those of *D. yakuba* with substitutions occurring only in $tRNA^{val}$ (one change) and $tRNA^{ile}$ (two changes). No insertions/deletions occur in the four tRNA genes. All four tRNA gene sequences are shown in proposed secondary structure form in Figure 4. The substitution in the $tRNA^{val}$ of *D. melanogaster* occurs at position 25 (numbering system in Steinberg et al. 1993) in the base of the dihydrouridine stem and results in a G·T wobble pairing as opposed to the G-C base pair in the $tRNA^{val}$ of *D. yakuba*. $tRNA^{val}$ in both species contains a T-T mismatch in the TΨC stem at position 50. Substitutions in $tRNA^{ile}$ occur in the variable loop and the dihydrouridine loop. Combined with the results of Garesse (1988) and Wolstenholme and Clary (1985b), the average frequency of substitution per nucleotide is 3.2 % for all 22 tRNA genes encoded on the *D. melanogaster* mtDNA compared to 3.2% and 2.8% for the 16S and 12S rRNA genes, respectively, and 7.2% for the protein coding genes.

As expected given the greater time of divergence between *D. melanogaster* and *D. virilis*, the nucleotide substitution frequency of the three tRNAs compared between these two species is larger (5.4%) than that between *D. melanogaster* and *D. yakuba* (0.95%). The gene encoding $tRNA^{val}$ in *D. virilis* mtDNA has not been sequenced. The genes encoding $tRNA^{gln}$ and $tRNA^{f\text{-}met}$ both contain two substitutions while the $tRNA^{ile}$ gene contains the remaining seven. As shown in Figure 4, eight of the substitutions occur in loops and three occur in stems in the proposed secondary structures, none of which disrupt base pairing. The only insertion/deletion occurs in the dihydrouridine loop of $tRNA^{ile}$.

Figure 4. **Proposed secondary structures of the *D. melanogaster* mitochondrial $tRNA^{val}$, $tRNA^{ile}$, $tRNA^{gln}$ and $tRNA^{f\text{-}met}$ based on the corresponding mtDNA sequences.** Bold letters near circles indicate nucleotide differences from the *D. melanogaster* sequence in the corresponding tRNA genes of *D. yakuba* ($tRNA^{val}$, $tRNA^{ile}$, $tRNA^{gln}$ and $tRNA^{f\text{-}met}$) and *D. teissieri* ($tRNA^{ile}$, $tRNA^{gln}$ and $tRNA^{f\text{-}met}$). Bold letters near arrows indicate nucleotide differences in the corresponding tRNA genes of *D. virilis* ($tRNA^{ile}$, $tRNA^{gln}$ and $tRNA^{f\text{-}met}$). The X accompanying the arrow in the $tRNA^{ile}$ structure indicates that nucleotide is missing in the *D. virilis* gene.

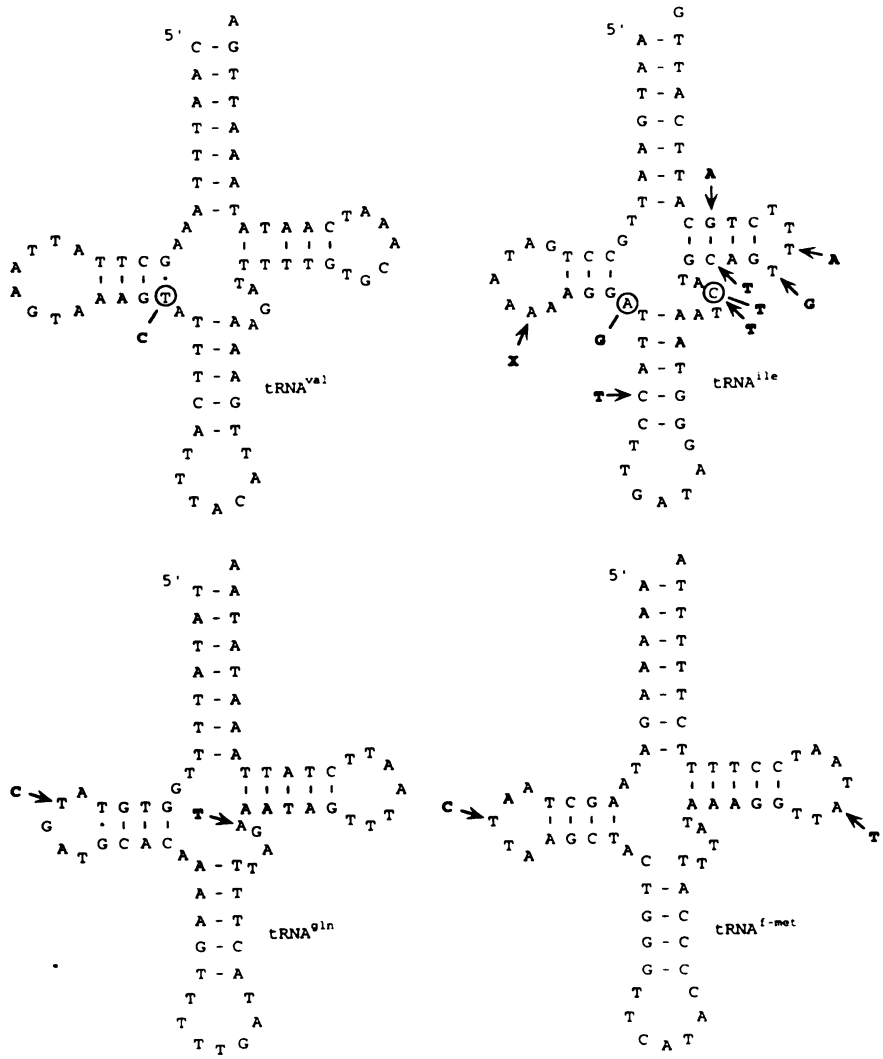


Figure 4

The nucleotide sequence of the tRNAs reported here are similar to other mitochondrial tRNAs in that they lack many of the invariant and semi-invariant nucleotides found in tRNAs of prokaryotes and the non-organellar tRNAs of eukaryotes. Only the conserved T33 and Pu37 are present in all sequenced *Drosophila* mitochondrial tRNAs (de Bruijn 1983; Clary and Wolstenholme 1985b, 1987; Garesse 1988). Further, one or more of the universally conserved nucleotide pairs responsible for tertiary interactions in non-organellar tRNAs, for example G18-T55, G19-C56, A58-T54 T8-A14 (Sampson et al. 1990), are missing in the tRNAs of *Drosophila* mitochondria suggesting that overall tertiary bonding is weaker. In fact, bovine mt $tRNA^{phe}$ has been shown to have a much lower melting temperature than the yeast cytosolic $tRNA^{phe}$ and *E. coli* $tRNA^{phe}$ suggesting the bovine mt $tRNA^{phe}$ has a less stable tertiary structure (Kumazawa et al. 1989,1991).

A less stable mitochondrial tRNA tertiary structure may limit the variety of mechanisms that could be utilized by mitochondrial tRNA synthetases to recognize their cognate tRNAs. Nucleotides which contribute to the folding of yeast $tRNA^{phe}$ have been shown to be important for recognition by the tRNA's cognate aminoacyl-tRNA synthetase, but are not involved in sequence-specific interactions with the synthetase (Sampson et al. 1990). Potential sites of recognition of mitochondrial tRNAs by their cognate mitochondrial aminoacyl tRNA synthetases have not been examined. However, the finding that bovine mitochondrial phenylalanyl-, threonyl-, arginyl- and lysyl-tRNA synthetases were able to unilaterally charge the cognate tRNAs from *E. coli* suggests they recognize a conserved feature of these tRNAs (Kumazawa et al. 1991). The most likely conserved feature is the anticodon itself suggesting it is a major determinant for defining the specificity of these mitochondrial tRNA synthetases. The fact that *E. coli* aminoacyl synthetases were unable to charge the mitochondrial tRNAs suggests additional determinants are required by the *E. coli* enzymes for cognate tRNA recognition. With regard to the mitochondrial tRNAs sequenced in this report,

evolution of the tRNA recognition mechanisms utilized by the *E. coli* and mitochondrial aminoacyl tRNA synthetases may not have involved radical differences. It has been shown that the anticodon is the major determinant for establishing the identities of *E. coli* $tRNA^{met}$, $tRNA^{val}$ and $tRNA^{ile}$ by their cognate aminoacyl tRNA synthetases (Schulman and Pelka 1988, Muramatsu et al. 1988). It is also one of the determinants for the identity of $tRNA^{gln}$ (Rogers and Soll 1988). If the mechanism of recognition has been conserved, it is plausible that the recognition sequences in the four mt-tRNAs presented in this report include the anticodons as well.

CHAPTER IV

A MITOCHONDRIAL DNA BINDING PROTEIN FROM *Drosophila melanogaster* EMBRYOS: SPECIFIC INTERACTIONS WITH THE A+T CONTROL REGION

INTRODUCTION

Drosophila melanogaster mitochondrial DNA (mtDNA) is a double-stranded, circular molecule of approximately 19.5 kb. Its complete nucleotide sequence is now known (Clary et al. 1982; Clary et al. 1983; de Bruijn 1983; Garesse 1988; Chapters II, III). *Drosophila* mtDNAs like those of other metazoans except nematodes, encode 13 polypeptides involved in oxidative phosphorylation and the 2 ribosomal RNAs and 22 tRNAs required for mitochondrial translation (Moritz et al. 1987; Okimoto et al. 1992). The *Drosophila* mtDNA genome is arranged in a compact manner with few or no intergenic nucleotides.

A single noncoding region, termed the A+T region due to its extreme bias towards deoxyadenylate and thymidylate residues, is present in *Drosophila* mtDNAs and varies in size from 1 to 5 kb (Fauron and Wolstenholme 1976). Size differences can be attributed for the most part to varying copy numbers of repeated DNA sequence elements (Chapter II; Fauron and Wolstenholme 1980b; Solignac et al. 1983; Hale and Singh 1986; Solignac et al. 1986; Monforte et al. 1993). In *D. melanogaster* mtDNA, two types of tandemly repeated elements have been identified (Chapter II). The first type averages 346 bp in length and is present in five copies located in the half of the A+T region nearest the gene encoding the small rRNA. The second type is approximately 460 bp in length and occurs in four copies adjacent to the gene encoding $tRNA^{ile}$.

Electron micrographic studies of DNA heteroduplexes and comparative DNA sequence analyses of the A+T regions of different *Drosophila* species indicate it is highly divergent relative to the coding region (Chapter II; Fauron and Wolstenholme 1980b; Fauron and Wolstenholme 1980a; Clary and Wolstenholme 1985b; Clary and Wolstenholme 1987). Despite this, conserved DNA sequence elements have been identified. Two thymidylate stretches, one on each DNA strand and comprising 13-25 bp, occur at similar positions within the A+T regions of different species (Chapter II). A

conserved element of approximately 300 bp is also present in all *Drosophila* mtDNA A+T regions whose DNA sequence is known (Clary and Wolstenholme 1985b; Clary and Wolstenholme 1987; Monnerot et al. 1990; Chapter II). The A+T regions from *Drosophila* species with short A+T regions contain a single copy whereas the long A+T region of *D. melanogaster* contains four copies, present within the tandemly repeated DNA sequence elements near the gene encoding $tRNA^{ile}$.

The role of the conserved DNA sequence elements in mtDNA metabolism has not been established. We have proposed the possibility that a 300 bp conserved element, perhaps in conjunction with an adjacent thymidylate stretch, functions in DNA replication. Evidence to support this suggestion derives from electron micrographic studies of replicative intermediates that map near these sequence elements (Wolstenholme et al. 1983). Alternatively, the conserved DNA sequence elements may function in the transcription of *Drosophila* mtDNA. Although steady-state transcripts have not been detected in the A+T region (Merten and Pardue 1981), it seems likely that precursor RNAs originate here as in the noncoding regions of vertebrate systems (Berthier et al. 1986).

Our knowledge of proteins involved in the regulation of animal mtDNA replication and transcription has derived largely from the study of vertebrates (Clayton 1991a). A protein identified in human and mouse, termed mitochondrial transcription factor A (mtTFA), is necessary for specific transcription from mitochondrial promoters in vitro (Fisher and Clayton 1985; Fisher and Clayton 1988; Fisher et al. 1989). mtTFA binds to upstream regulatory elements of promoters, albeit with low sequence specificity (Fisher and Clayton 1988; Fisher et al. 1989; Fisher et al. 1987). Because initiation of both DNA replicative intermediates and RNA transcripts occur at the light-strand promoter (Chang et al. 1985; Chang and Clayton 1985), it has been proposed that mtTFA is involved in both processes (Clayton 1991a; Parisi et al. 1993). Molecular cloning of the gene encoding mtTFA and subsequent DNA sequence

analysis has shown it to be a member of the high-mobility-group (HMG) proteins found in the nucleus (Parisi and Clayton 1991). Originally identified as general DNA-binding proteins in chromatin (Johns et al. 1977), several examples of HMG-box containing proteins have been implicated in the activation of specific transcription in the nucleus (Jantzen et al. 1990; van de Wertering et al. 1991; Travis et al. 1991; Sinclair et al. 1990; Kelly et al. 1988). An HMG protein, mtDBP-C, has also been identified in mitochondria of *Xenopus* oocytes (Mignotte and Barat 1986; Ghrir et al. 1991b). Like mtTFA, mtDBP-C binds DNA with low sequence specificity and is able to induce superhelical turns into relaxed circular DNA in the presence of DNA topoisomerase I (Mignotte and Barat 1986; Mignotte et al. 1988; Fisher et al. 1992). Although the physiological role of mtDBP-C is unknown, it was proposed to be involved in DNA compaction of the mitochondrial nucleoid (Mignotte et al. 1988; Mignotte et al. 1990).

Little is known about the proteins involved in the replication and transcription of *Drosophila* mtDNA. To date only *Drosophila* mitochondrial DNA and RNA polymerases have been described (Wernette and Kaguni 1986; Goldenthal and Nishiura 1987). We have sought to identify mitochondrial proteins that interact with the noncoding region of mtDNA in order to better understand the factors which regulate mitochondrial replication and transcription in *Drosophila*. We report here the identification and biochemical characterization of an A+T region-specific DNA-binding protein from mitochondria of *Drosophila* embryos.

EXPERIMENTAL PROCEDURES

Materials

Chemicals. Sodium metabisulfite was purchased from J. T. Baker Chemical Co. and prepared as a 1.0 M stock solution at pH 7.5 and stored at -20 °C. Leupeptin purchased from the Peptide Institute, Minoh-Shi, Japan, was prepared as a 1 mg/ml stock solution in 0.1 M potassium phosphate buffer, pH 7.5, and stored at -20 °C. Dithiothreitol (Sigma) was prepared as 1.0 M stock solution in glass distilled water and stored at -20 °C. Triton X-100 was purchased from Sigma. Low melting point agarose was obtained from FMC Bioproducts.

Nucleic Acids and Nucleotides. The synthetic polymers poly(dI-dC) · poly(dI-dC), poly(dA-dT) · poly(dA-dT), poly(dA) and poly(dT)$_{15}$ were purchased from Pharmacia LKB Biotechnology, Inc. Each was resuspended in 10 mM Tris, pH 8.0, 1 mM EDTA and the concentrations determined spectrophotometrically. Unlabeled deoxy- and ribo-nucleotides were purchased from P-L Biochemicals; [α-^{32}P]dATP was from New England Nuclear.

Enzymes and Protein Standards. T4 DNA ligase was purchased from Gibco BRL and the Klenow fragment of *Escherichia coli* DNA polymerase I from New England Biolabs. Restriction enzymes were from New England Biolabs and Pharmacia LKB Biotechnology, Inc. Bovine serum albumin (fraction V), rabbit muscle L-lactate dehydrogenase and cytochrome *c* were purchased from Sigma. Horse heart myoglobin, chicken ovalbumin and bovine gamma globulin were from BioRad. Human serum albumin and bovine carbonic anhydrase were from Worthington. *Escherichia coli* DNA polymerase I was purchased from New England Biolabs.

Methods

Preparation of DNA and DNA substrates. Recombinant DNAs containing fragments of the mtDNA A+T region of *Drosophila* embryonic mitochondria were obtained as previously described (Chapter II). The recombinant plasmid pDMS 38 contains DNA from nucleotides 5213 in the A+T region to the *Hin*d III site in the gene encoding subunit 2 of the NADH dehydrogenase complex; pDMP 104 contains nucleotides 3273-3782; pDMP 2 contains nucleotides 2900-3210 and pDMP 63 contains nucleotides 1451-2217 (numbers as in Chapter II). Large-scale plasmid DNA preparations were purified by equilibrium CsCl density gradient centrifugation according to standard laboratory procedures. Plasmid DNA inserts were isolated by digestion with the appropriate restriction enzymes followed by gel electrophoresis in low melting point agarose. Poly(dA) · poly(dT) was prepared by annealing an equal molar amount as nucleotide of 5' phosphorylated dT_{15} to poly(dA) followed by ligation with T4 DNA ligase. The mixture was extracted once with phenol and once with chloroform:isoamyl alcohol (24:1) and the DNA precipitated by the addition of 2.5 volumes of ethanol. After centrifugation, the DNA pellet was resuspended in 10 mM Tris, pH 8.0, 1 mM EDTA and the concentration determined spectrophotometrically.

Electrophoretic mobility shift assay. DNA-binding activity was detected by gel mobility shift assay (Revzin 1989). DNA fragments were 5' end-filled with [α-^{32}P] dATP using Klenow fragment of *Escherichia coli* DNA polymerase I according to standard procedures (Sambrook et al. 1989). Polyacrylamide gels (4% total acrylamide, 29:1 acrylamide:*N N'* -methylene-bis-acrylamide) were prepared in 22 mM Tris base, 22 mM boric acid, 0.5 mM EDTA and allowed to cure at least two hours prior to use. The gels were pre-electrophoresed at 175V until constant current was achieved. The running buffer, 22 mM Tris base, 22 mM boric acid, 0.5 mM EDTA, was recirculated during both pre-electrophoresis and sample electrophoresis. Reaction mixtures (30 ul) contained 20 mM *N* -2-hydroxyethylpiperizine-*N'* -2-ethanesulfonic

acid (HEPES), pH 7.6, 50 mM KCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 2 mM dithiothreitol, 5 % glycerol, 0.15 mg/ml BSA, 100 ng of poly(dI-dC) poly(dI-dC), 2 fmol of radiolabeled DNA fragment and various amounts of DNA binding protein. Incubation was at 18°C for 15 min. The reaction mixture was immediately loaded onto a polyacrylamide gel running at 75V at room temperature. The samples were electrophoresed for 1 hr at 75V and then 2 hr 15 min at 175V. After electrophoresis, the gel was dried and autoradiographed on Kodak X-OMAT AR film. The amount of free DNA and DNA in complex with mtDBP-26 was quantitated using a PhosphorImager (Molecular Dynamics, Inc.) and related software.

Preparation of mtDBP-26 and Glycerol Gradient Sedimentation. The purification procedure for DNA polymerase γ from *D. melanogaster* embryos was performed as previously described (Wernette and Kaguni 1986). After glycerol gradient sedimentation, fractions were assayed for DNA-binding activity by electrophoretic mobility shift analysis using the radiolabeled 386 bp *Eco*R I-*Alu* I fragment of pDMS 38 as the DNA substrate in the presence of 100 ng of unlabeled poly(dI-dC) · poly (dI-dC). Protein in each fraction (50 ul) was precipitated by the addition of TCA to 12 % (w/v) final concentration and then subjected to SDS polyacrylamide gel electrophoresis according to the method of Laemmli (Laemmli 1970). The proteins were stained in the gel with silver (Oakley et al. 1980) and then quantitated by scanning the gel with a laser densitometer (Molecular Dynamics, Inc.). Protein fractions containing mtDBP-26 activity were pooled, an equal volume of 20 mM potassium phosphate, pH 7.6, 2 mM EDTA, 80% glycerol, 0.015% Triton X-100 was added, and the protein was stored at -20°C. A parallel glycerol gradient was calibrated with L-lactate dehydrogenase, *E.* coli DNA polymerase I, human serum albumin, carbonic anhydrase and cytochrome c.

Gel Filtration Chromatography of mtDBP-26. Bovine serum albumin (1 mg/ml final concentration) was added to the glycerol gradient-purified mtDBP-26. Gel

filtration chromatography of mtDBP-26 (200 ul) was performed at a flow rate of 0.1 ml/min on a Superdex 75 10/30 HR FPLC column equilibrated with 20 mM potassium phosphate, pH 7.6, 250 mM KCl, 2 mM EDTA, 20% glycerol, 2 mM dithiothreitol, 10 mM sodium metabisulfite, 2 ug/ml leupeptin and 0.015% Triton X-100. Fractions (0.2 ml) were collected and aliquots were assayed for DNA-binding activity by electrophoretic mobility shift assay. Protein in each fraction (185 ul) was precipitated by the addition of TCA to 12 % (w/v) final concentration and then subjected to SDS polyacrylamide gel electrophoresis according to the method of Laemmli (Laemmli 1970). The proteins were stained in the gel with silver (Oakley et al. 1980) and then quantitated by scanning the gel with a laser densitometer (Molecular Dynamics, Inc.). The gel filtration column was calibrated with blue dextran 2000, bovine gamma globulin, chicken ovalbumin and horse myoglobin.

Determination of the affinity constant of mtDBP-26 : DNA complexes. The equilibrium binding constant of mtDBP-26 : DNA complexes was determined by mobility shift analysis performed as described above. A constant amount of mtDBP-26 Fr. VI and increasing concentrations of the [^{32}P] - labeled 386 bp DNA (1.5 - 50 pM) were used. After resolution of the free DNA from that complexed with mtDBP-26 on mobility shift gels, the amount of DNA in each band was quantitated using a PhosphorImager (Molecular Dynamics, Inc.) and related software. The amount of DNA in fmol in each band was calculated from a standard curve generated by electrophoresis of varying amounts of DNA on mobility shift gels in the absence of mtDBP-26. The binding affinity was determined by plotting the data in a Woolf plot (Cressie et al. 1981) where the slope of the line is equal to $1/B_{max}$ and the y intercept is equal to K_D/B_{max}. A curve was fitted to the data using the method of least squares with the curve fitting function of the DeltaGraph Professional software (Deltapoint, Inc.). Alternatively, the affinity constant was determined by fitting the data to the equation, $BOUND = (n*K_{eq}*FREE*B_{max}) / (1+ K_{eq} * FREE)$ where n is the number of

independent DNA-binding sites on the protein, BOUND is the amount of DNA complexed with mtDBP-26 and FREE is the amount of free DNA. A computer program was written which varied K_{eq} and B_{max} until a best-fit curve was achieved (T. Deits, personal communication). In these calculations it was assumed n=1.

Determination of the dissociation rate constant of mtDBP-26 : DNA complexes. The [^{32}P] - labeled *Eco*R I-*Alu* I fragment of pDMS 38 was preincubated with mtDBP-26 Fr. VI for 15 min under standard assay conditions in a total volume of 80 ul. After the preincubation, a 20 ul aliquot was removed to a separate tube and 10 ul of this was loaded onto a mobility shift gel running at 100V. A 100-fold molar excess of the unlabeled DNA fragment was added to the remaining reaction mixture and a 10 ul aliquot was immediately loaded onto the gel. Additional 10 ul aliquots of the mixture were then loaded onto a mobility shift gel at various times after the addition of the unlabeled DNA. At the end of the time course, the remaining 10 ul of the mixture not containing the competitor was loaded onto the gel as a control. Electrophoresis continued at 100 V for 2 hr 30 min after the last sample was loaded. The amount of DNA bound by mtDBP-26 at each time point was quantitated using a PhosphorImager as above. Dissociation rate data was plotted and then curve-fitted according to the equation, $BOUND_t / BOUND_o = e^{(-k_{off}t)}$, where $BOUND_t$ is the amount of DNA bound at time t, $BOUND_o$ is the amount of DNA bound immediately prior to the addition of excess unlabeled competitor DNA, and k_{off} is the dissociation rate constant.

Generation of competitive displacement curves. The [^{32}P] - labeled *Eco*R I-*Alu* I fragment of pDMS 38 was incubated with mtDBP-26 Fr. VI under standard conditions in the presence of various concentrations of unlabeled competitor DNAs and subjected to mobility shift analysis as described above. Poly(dI-dC) · poly(dI-dC) was present at 100 ng per assay in each competition assay except for that in which poly(dI-dC) · poly(dI-dC) itself was titrated. After electrophoresis, the amount of DNA bound by mtDBP-26 was quantitated using a PhosphorImager as above. The amount of DNA

bound at each point was expressed as a percentage of DNA bound in the absence of competitor except in the case of the poly(dI-dC) · poly(dI-dC) titration were the amount of DNA bound was made relative to that amount bound at 100 ng added DNA.

RESULTS

Purification and Physical Properties.

In an effort to identify sequence-specific DNA-binding activities in *D. melanogaster* mitochondria, mobility shift assays were employed to test for the presence of mitochondrial protein(s) that could specifically interact with a DNA fragment containing sequences within the A+T region of *D. melanogaster* mtDNA. The 386 bp *Eco*R I-*Alu* I fragment from pDMS 38 was chosen as the DNA substrate. The fragment contains 145 bp of A+T region DNA including a conserved thymidylate stretch (Chapter II) and extends into the flanking coding region (Fig. 1). Initially, mobility shift assays were performed on protein fractions collected during the purification of DNA polymerase γ (Wernette and Kaguni 1986). Multiple A+T region DNA-binding activities were detected during the initial stages of the purification as evidenced by numerous shifted DNA species in the presence of excess poly(dI-dC) · poly(dI-dC) competitor (data not shown). However, one DNA-binding activity co-purified with Pol γ through Fraction V. Glycerol gradient sedimentation of Pol γ Fraction V resulted in the separation of Pol γ activity from the DNA-binding activity. The calculated sedimentation coefficient of the DNA-binding activity is 2.2 S (Fig. 2). Analysis of the protein contained in the fractions with DNA-binding activity by SDS-PAGE and subsequent densitometric scanning of the silver stained gel revealed a major polypeptide of 26,000 daltons which correlated with DNA-binding activity (data not shown). Superdex-75 FPLC gel filtration of the glycerol gradient fraction again resulted in the correlation of DNA-binding activity with a 26,000 dalton polypeptide as determined by SDS-PAGE analysis (Fig. 3). The observed Stokes radius of the DNA-binding activity as measured by gel filtration is 26 A. Calculation of the native molecular mass of the DNA-binding activity by utilizing the sedimentation coefficient and the Stokes radius (Siegel and Monty 1966) results in a value of 24,000 daltons.

Figure 1. **Schematic of the *Drosophila melanogaster* mtDNA A+T region and flanking genes and the locations of the DNA fragments used in the experiments.** DNA fragments containing A+T region DNA were isolated as described under "Methods." The schematic shows the 5.8 kb *Hin*d III-B fragment containing the central A+T region (4601 bp) and flanking genes encoding the small rRNA, $tRNA^{ile}$ (I), $tRNA^{gln}$ (Q), $tRNA^{f\text{-}met}$ (M), and the 5' portion of NADH dehydrogenase subunit 2 (ND2): solid boxes, tRNA genes; box with rightward leaning slashes, 5'-coding region for NADH dehydrogenase subunit 2; dotted box, small rRNA gene; solid line, non-coding sequences including the central A+T region; boxes with leftward leaning slashes, type II repeated DNA sequence elements; boxes with wavy lines, type I repeated DNA sequence elements. Open boxes in the A+T region indicate the position of the longest thymidylate stretches on each DNA strand. Bold lines below the A+T region indicate the position of 300 bp conserved DNA sequence elements. Arrows above the coding regions indicate the direction of transcription. The DNA replication origin region and direction of leading DNA strand synthesis in *D. melanogaster* mtDNA (Wolstenholme et al. 1983) is indicated by a bracketed arrow. Bracketed lines denote the extent and location of the DNA fragments used in the experiments.

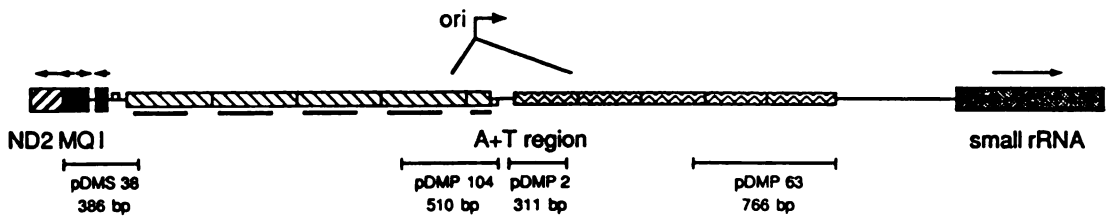

ori
ND2 MQ I
A+T region
small rRNA
pDMS 38
386 bp
pDMP 104
510 bp
pDMP 2
311 bp
pDMP 63
766 bp

Figure 1

Figure 2. **Glycerol gradient sedimentation profile of *Drosophila* mtDBP-26.** Glycerol gradient sedimentation of DNA polymerase γ Fr. V, which also contains the DNA-binding activity, was performed as described previously (Wernette and Kaguni 1986). DNA-binding activity (open circles) was detected by electrophoretic mobility shift assay and the relative amount of the 26 kDa protein in each fraction (closed circles) was quantitated as described under "Methods." The large arrow associated with the letter "S" indicates the direction of sedimentation. The glycerol gradient was calibrated with L-lactate dehydrogenase (LDH, 7.7S), *E. coli* DNA polymerase I (POL I, 5.5S), human serum albumin (HSA, 4.6S), carbonic anhydrase (CA, 3.2S) and cytochrome c (CYT, 1.7S) and the result shown in the inset. The small arrow indicates the sedimentation position of the mitochondrial DNA-binding activity.

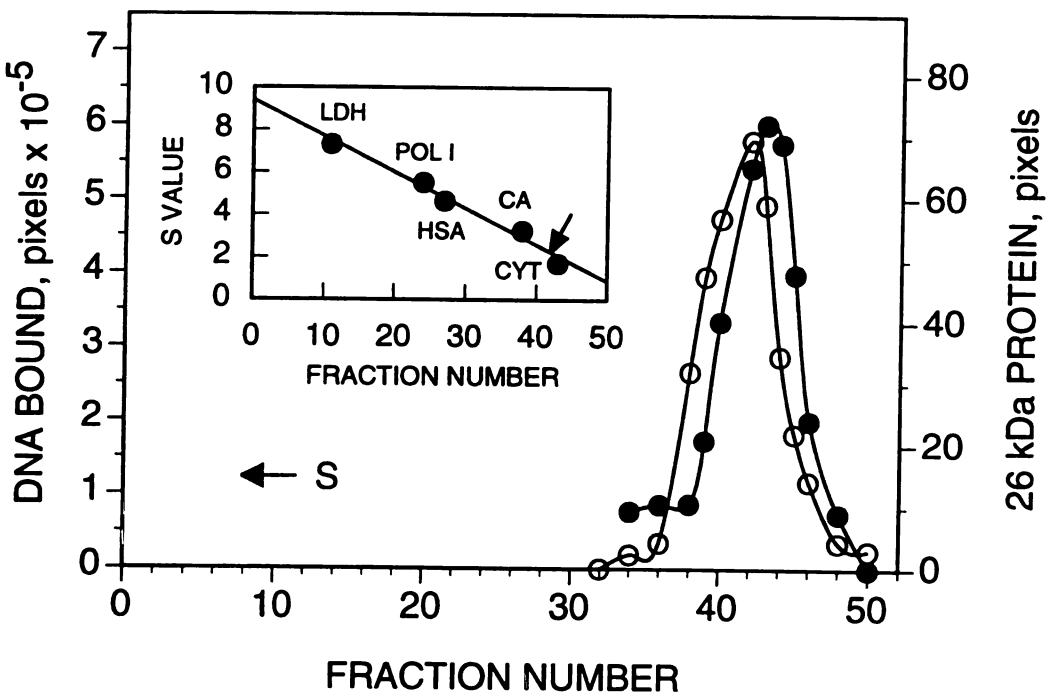

DNA BOUND, pixels x 10^{-5}
26 kDa PROTEIN, pixels
FRACTION NUMBER
S VALUE
LDH
POL I
CA
HSA
CYT
S
0
10
20
30
40
50
60
70
80

Figure 2

Figure 3. **Gel filtration of *Drosophila* mtDBP-26.** Gel filtration chromatography of glycerol gradient-purified DNA-binding activity was performed using a Superdex 75 10/30 HR FPLC column as described under "Methods." DNA-binding activity and the relative amount of the 26 kDa protein present in each fraction was determined as in Fig. 2. A. Graphical representation of the data presented in B and C; open circles, the relative amount of DNA in the complex; closed circles, the relative amount of the 26 kDa protein. The gel filtration column was calibrated with myoglobin (MYO, 19 A), ovalbumin (OVA, 27 A) and γ-globulin (GAM, 44 A) and the result shown in the inset. The small arrow indicates the elution position of the mitochondrial DNA-binding activity. B. Autoradiograph of the electrophoretic mobility shift gel; BOUND, position of the radiolabeled DNA fragment in complex with the mitochondrial DNA-binding activity. FREE, position of the radiolabeled DNA fragment not bound by protein. C. SDS polyacrylamide gel analysis of column fractions spanning the peak of mitochondrial DNA-binding activity.

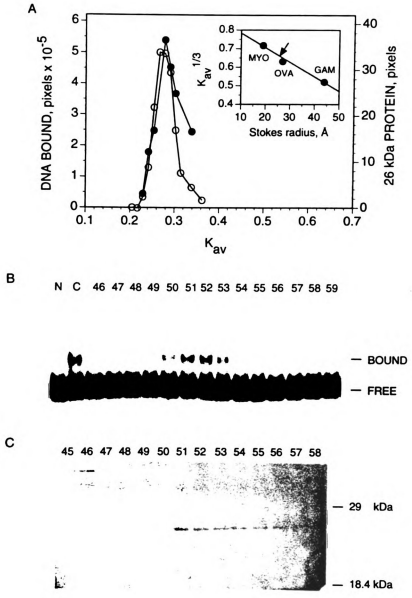

A
DNA BOUND, pixels x 10^-5
26 kDa PROTEIN, pixels
K_av
K_av 1/3
MYO
OVA
GAM
Stokes radius, Å
B
N C 46 47 48 49 50 51 52 53 54 55 56 57 58 59
BOUND
FREE
C
45 46 47 48 49 50 51 52 53 54 55 56 57 58
29 kDa
18.4 kDa

Figure 3

Based on these results, we have assigned the DNA-binding activity to the 26,000 dalton polypeptide and termed it mitochondrial DNA-binding protein-26 (mtDBP-26).

mtDBP-26 DNA Binding Properties.

Reaction Requirements. The effects of MgCl and monovalent salt concentrations and temperature on mtDBP-26 activity were investigated in electrophoretic mobility shift analyses. DNA-binding activity was absolutely dependent on the presence of Mg^{++} with optimal binding occurring in the presence of 2 to 5 mM MgCl. DNA-binding activity was unaffected by concentrations of KCl up to 250 mM and was to half-maximal at 500 mM. No alteration in the fraction of DNA bound by mtDBP-26 was observed when incubation temperature was varied between 0 and 37°C. The heat stability of mtDBP-26 in the absence of DNA was also tested. The protein retained near normal levels of DNA-binding activity after pre-incubation at temperatures up to 90°C for 10 min.

Determination of the equilibrium binding constant of the mtDBP-26 : DNA complex. In order to quantitate the affinity of mtDBP-26 for the 386 bp *Eco*R I-*Alu* I fragment of pDMS 38, the equilibrium dissociation constant (K_d) of the mtDBP-26 : DNA complex was determined. Three independent experiments were performed. The data was plotted in a Woolf plot (free/bound DNA vs. free DNA) or as bound DNA versus free DNA and curve fitted as described under "Methods" (Fig. 4). The K_D values derived from both plots were similar, 16 pM when the data was plotted in a Woolf plot and 19 pM plotted as bound DNA versus free DNA. The linearity of the Woolf plot is suggestive of a single DNA-binding species.

Dissociation kinetics of mtDBP-26 : DNA complexes. The stability of mtDBP-26 bound to the pDMS 38 fragment was measured using the mobility shift assay as described under "Methods." As shown in Fig. 5, dissociation of mtDBP-26 from the A+T DNA containing fragment displayed first order kinetics as expected. The resulting line is described by the equation $f(t) = 96\ e^{-0.22t}$. Thus, the first order dissociation rate

Figure 4. **Determination of the dissociation constant (K_D) of mtDBP-26 : DNA complexes.** Electrophoretic mobility shift assays were performed using a constant amount of the glycerol gradient-purified mtDBP-26 and increasing amounts of the radiolabeled 386 bp *EcoR* I-*Alu* I fragment of pDMS 38. The amount of free DNA and DNA bound by mtDBP-26 in each assay was quantitated as described under "Methods." A. DNA-binding data plotted in a Woolf plot and a curve-fitted using the method of least squares. B. DNA-binding data plotted as BOUND versus FREE and curve-fitted according to the equation BOUND = $(n*K_{eq}*FREE*B_{max}) / (1+ K_{eq} * FREE)$. C. Autoradiograph showing the results of a typical electrophoretic mobility shift experiment. Duplicates at each DNA concentration are as indicated by the brackets above the lanes. Increasing amounts of the DNA fragment are indicated below. BOUND, position of the DNA fragment in complex with mtDBP-26. FREE, position of the DNA not bound by protein.

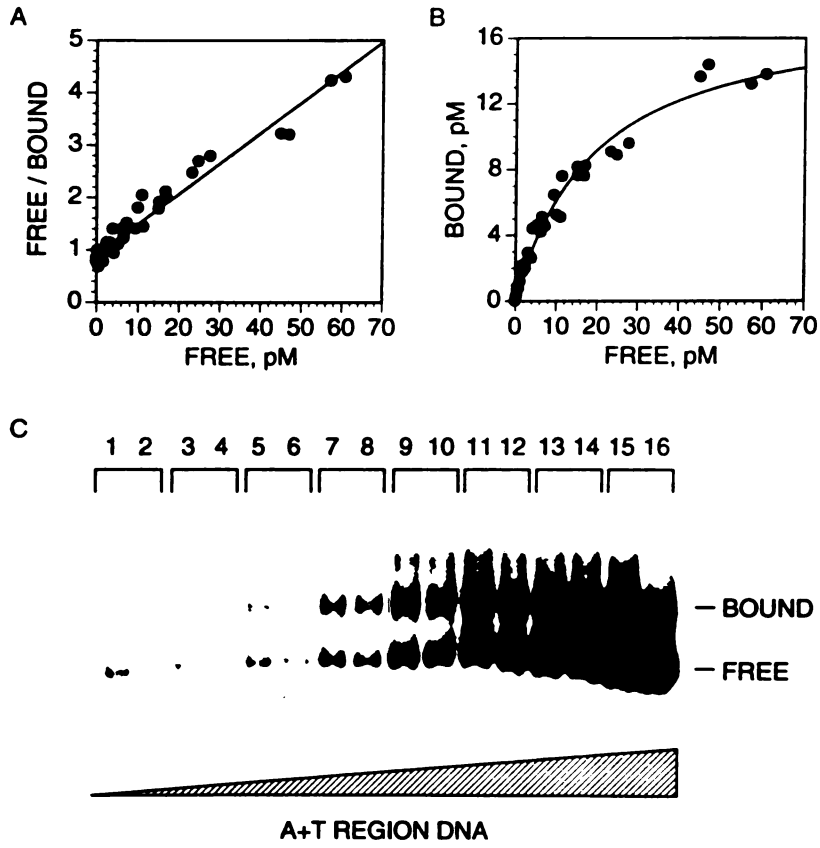

A
FREE / BOUND
0
1
2
3
4
5
0 10 20 30 40 50 60 70
FREE, pM
B
BOUND, pM
0
4
8
12
16
0 10 20 30 40 50 60 70
FREE, pM
C
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
— BOUND
— FREE
A+T REGION DNA

Figure 4

Figure 5. **Determination of the dissociation rate constant (k_{off}) of mtDBP-26 : DNA complexes.** Dissociation of mtDBP-26 : DNA complexes was measured by pre-incubating glycerol gradient purified mtDBP-26 with the radiolabeled 386 bp *EcoR* I-*Alu* I fragment of pDMS 38 then adding a 50-fold molar excess of the unlabeled DNA fragment as described under "Methods." Aliquots were removed and subjected to electrophoretic mobility shift analysis at the times indicated. A. Graphical representation of the data derived from the experiment below. $BOUND_t$, the amount of DNA in complex with mtDBP-26 at time t; $BOUND_0$, the amount of DNA in complex with mtDBP-26 immediately prior to the addition of unlabeled DNA. B. Autoradiograph of the electrophoretic mobility shift gel. Lanes 1 and 8 contain controls in which no DNA was added after pre-incubation. BOUND, position of the DNA fragment in complex with mtDBP-26. FREE, position of the DNA not bound by protein.

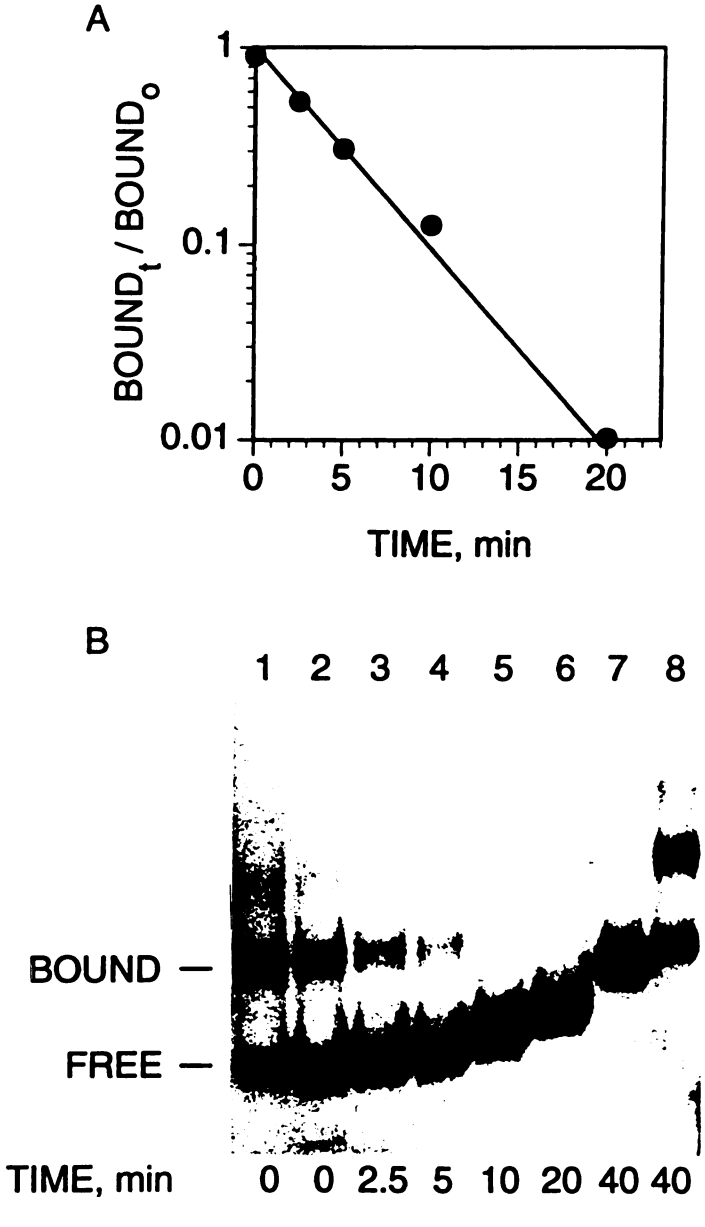

A
1
0.1
0.01
BOUND$_t$ / BOUND$_o$
0
5
10
15
20
TIME, min
B
1 2 3 4 5 6 7 8
BOUND —
FREE —
TIME, min
0 0 2.5 5 10 20 40 40

Figure 5

constant, k_{off}, is 0.22 min^{-1} corresponding to a half-life of 3.1 min for the protein-DNA complexes. From the calculated dissociation constant and equilibrium binding constant calculated above, the association rate constant can be derived. Using the values of 19 pM for the K_d and 3.7 x 10^{-3} s^{-1} for k_{off}, the association constant, k_{on}, is calculated to be 1.9 x 10^8 $M^{-1}s^{-1}$. This value is similar to the rate of diffusion calculated for binding of the lac repressor to the lac operator and suggests that the association rate of mtDBP-26 with the DNA target site is also limited by the diffusion rate (Riggs et al. 1970).

A similar experiment was performed using non-A+T region DNA, specifically the 375 bp *Eco*R I-*Bam*H I fragment of pBR 322 (data not shown). Although mtDBP-26 was able to shift the pBR 322 derived fragment in a similar manner to the A+T region-containing fragment in the absence of unlabeled A+T region DNA, no shifted fragment could be detected after loading the sample immediately after the addition of the unlabeled A+T DNA fragment. Thus, mtDBP-26 completely dissociates from the pBR 322 derived fragment within 30-60s, the estimated time it takes for the sample to enter the gel. Increased stability of mtDBP-29 bound to A+T region DNA relative to that of normal base content DNA might suggest that mtDBP-26 is stabilized by contacts to specific bases in the A+T region containing fragment.

Relative binding affinities of mtDBP-26 to various non-A+T region and A+T region DNAs. In order to characterize the DNA sequence-specificity of mtDBP-26, the relative affinities of the protein for various DNA fragments were determined. The amount of the pDMS 38 *Eco*R I-*Alu* I fragment bound by mtDBP-26 was effectively decreased in the presence of increasing concentrations of the indicated unlabeled non-A+T region DNAs. Quantitation of the amount of the mtDBP-26 : pDMS 38 complex formed in the presence of increasing concentrations of the indicated competitor DNA was determined and competitive displacement curves were generated (Fig. 6A). The relative binding affinities of the non-A+T region competitor

DNAs as compared to that for the pDMS 38 *Eco*R I-*Alu* I fragment can be determined by comparing the concentrations required to reduce mtDBP-26 binding to the [^{32}P] - labeled fragment by 50 % of that in the absence of competitor (IC_{50}). In the saturation binding experiments described above, the affinity of mtDBP-26 binding to the 386 bp fragment of pDMS 38 was determined to be 19 pM. Comparison of the relative IC_{50} values reveals that mtDBP-26 binds to linearized pUC 1193 DNA and the synthetic DNAs, poly(dA-dT) · poly(dA-dT) and poly(dA) · poly(dT), with a 9- to 15-fold lower affinity than to the A+T region containing fragment (Table 1). Binding to the synthetic DNA, poly (dI-dC) · poly (dI-dC), occurs with a 500-fold lower affinity. Competition experiments using M13 DNA revealed mtDBP-26 binds single-stranded DNA with at least a 100-fold lower affinity than the double-stranded A+T region containing fragment (data not shown). The fact that mtDBP-26 only has a slight preference for DNA sequences contained within the A+T region containing fragment over pUC 1193 DNA, poly(dA-dT) · poly(dA-dT) and poly(dA) · poly(dT) suggests these DNAs contain sequences resembling the binding site in the A+T region containing fragment. Because sequence-specific DNA-binding proteins are expected to exhibit at least a 100-fold preference for target sequences versus other sequences, mtDBP-26 does not appear to be a sequence-specific DNA-binding protein in the classical sense. However, it is possible that the 386 bp A+T region containing fragment used in the above experiments does not contain the sequence elements required for specific binding by mtDBP-26.

To test whether mtDBP-26 specifically recognizes other sequences in the A+T region, DNA fragments containing other A+T region sequences were used to generate competitive displacement curves as above (Fig. 6B). The DNAs tested included i) pDMP 104, which contains the central conserved thymidylate stretch and a major portion of one of the conserved 300 bp elements as well as the partial conserved element adjacent to the thymidylate stretch, ii) pDMP 2, which contains most of a

Figure 6. **A+T region binding of *Drosophila* mtDBP-26.** Glycerol gradient-purified mtDBP-26 was incubated with a constant amount of the radiolabeled 386 bp *Eco*R I-*Alu* I fragment of pDMS 38 in the presence of increasing amounts of various unlabeled competitor DNAs as described under "Methods". The percentage of DNA bound (% BOUND) is defined as the amount of radiolabeled DNA bound by mtDBP-26 at a given competitor concentration divided by the amount bound in the absence of competitor (x100). All DNA concentrations are expressed as if 386 bp represented a unique mtDBP-26 binding site in order to normalize the data. A. DNA-binding competition experiments with non-A+T region DNAs. closed circles, 386 bp *EcoR* I-*Alu* I fragment of pDMS 38; open circles, poly(dA-dT) · poly(dA-dT); closed squares, pUC 1193 linearized with *Eco*R I; open squares, poly(dA) · poly(dT); closed triangles, poly(dI-dC) · poly(dI-dC). B. DNA-binding competition experiments with A+T region DNA fragments. closed circles, *EcoR* I-*Alu* I fragment of pDMS 38; open circles, *Eco*R I-*Pst* I fragment of pDMS 2; closed squares, *Eco*R I-*Pst* I fragment of pDMP 104; open squares, *Eco*R I-*Pst* I fragment of pDMP 63.

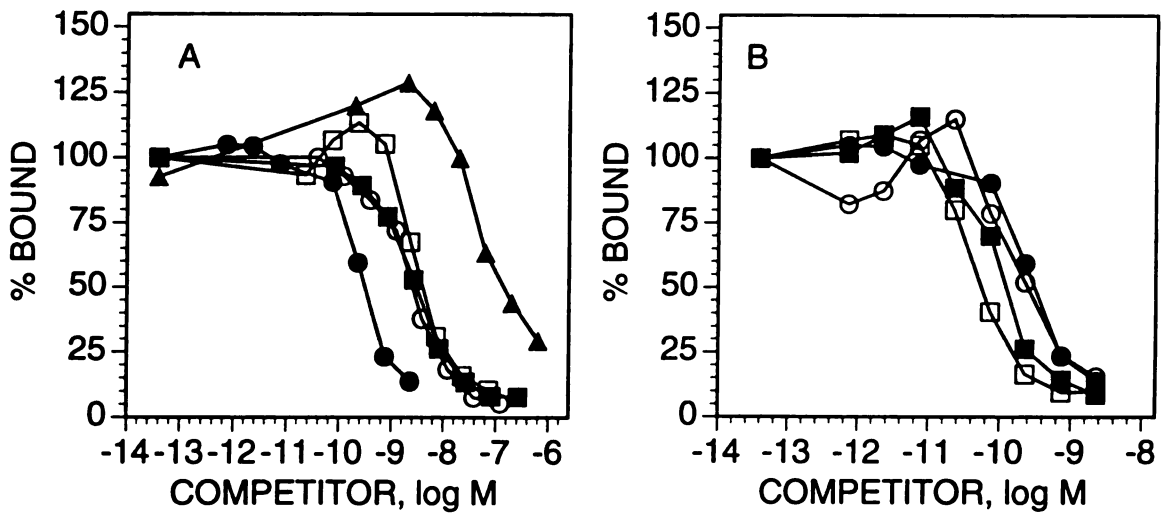

A
B
% BOUND
150
125
100
75
50
25
0
-14 -13 -12 -11 -10 -9 -8 -7 -6
-14 -13 -12 -11 -10 -9 -8
COMPETITOR, log M

Figure 6

Table 1. **Relative DNA binding affinity of *Drosophila* mtDBP-26 for various DNAs.**

Competitor	Relative DNA Binding Affinity, pM	K_D rel / K_D pDMS 38
pDMS 38	19	1.0
poly(dA-dT) · poly(dA-dT)	170	8.7
pUC 1193 / *Eco*R I	210	11
poly(dA) · poly(dT)	270	14
poly(dI-dC) · poly(dI-dC)	9300	490
A+T Region DNA		
pDMP 2	16	0.83
pDMP 104	8.0	0.42
pDMP 63	3.7	0.20

type I repeated element located in the half of the A+T region termed the variable region because of the poor sequence conservation among *Drosophila* species, and iii) pDMP 63, which contains two copies of the variable region repeat (Fig. 1). As shown in Figure 6, addition of increasing amounts of each A+T region DNA fragment effectively decreased the amount of the 386 bp fragment bound by mtDBP-26. IC_{50} values were obtained from the competitive displacement plot and used to calculate the relative binding affinities for each fragment (Table 1). Results reveal only slight differences in the binding affinity of mtDBP-26 for fragments containing different sequences of the A+T region. In addition, mtDBP-26 did not show a preference for A+T region DNA in a supercoiled versus linear form (data not shown), unlike the HMG-like proteins identified in *Xenopus* and yeast mitochondria (Mignotte and Barat 1986; Diffley and Stillman 1992). Thus, mtDBP-26 appears not to be a sequence-specific DNA-binding protein, at least in terms of its affinity towards different A+T region sequences. However, it should be noted that all fragments containing A+T region DNA were more effective at competing for mtDBP-26 binding than non-A+T region DNAs, including poly(dA-dT) · poly(dA-dT) and poly(dA) · poly(dT). The fact that the A+T region is composed of 96% deoxyadenylate and thymidylate nucleotides suggests that some feature of the A+T region DNA other than A+T-richness accounts for preferential binding. One possibility is that mtDBP-26 recognizes secondary structure elements present in A+T region DNA not found in poly(dA-dT) · poly(dA-dT) or poly(dA) · poly(dT). In fact, A+T region containing fragments display anomalous mobilities when electrophoresed in polyacrylamide gels (data not shown), suggesting curvature of these DNAs (Marini et al. 1982).

DISCUSSION

mtDBP-26 Binds to A+T Region DNA With Low Sequence Specificity. Data from experiments designed to test the DNA-binding properties of mtDBP-26 indicate that it binds A+T region DNA with a greater affinity than non-A+T region DNA. However, mtDBP-26 does not appear to bind to specific sequences within the A+T region raising the question as to the mechanism A+T region DNA recognition by the protein. Degenerate recognition of A+T-rich DNA has been described previously for a number of nuclear proteins including α-protein (Solomon et al. 1986), D1 protein (Levinger 1985), and histones (Richmond et al. 1984; McGhee and Felsenfeld 1980). Extensive minor groove contacts characterize α-protein and histones binding to DNA consistent with theoretical predictions for degenerate recognition of A+T-rich DNA (Seeman et al. 1976). In contrast, minor groove contacts by mtDBP-26 probably do not significantly contribute to DNA sequence recognition as evidenced by the inability of poly(dI-dC) · poly(dI-dC) to effectively compete for mtDBP-26 binding. This fact, in addition to the observation that poly(dA-dT) · poly(dA-dT) and poly(dA) · poly(dT) compete for mtDBP-29 binding 60- and 35-fold better, respectively, than poly(dI-dC) · poly(dI-dC), suggests major groove contacts play a larger role in sequence recognition. Alternatively, mtDBP-26 may recognize its binding sites through the use of hydrogen bonds to the phosphates, rather than to the base pairs, as in the case of the tryptophan repressor protein of *Escherichia coli* (Otwinoski et al. 1988). The sequences of bases in the DNA would provide for a specific phosphate conformation that can be bound by the protein.

The extremely high content of deoxyadenylate and thymidylate residues in the A+T region of *Drosophila melanogaster* mtDNA most likely results in an unusual DNA structure. Homopolymeric stretches of poly(dA) · poly(dT) occur many times in the A+T region (Chapter II). The longest stretches are located in the same relative positions in the mtDNA molecules of different *Drosophila* species. Poly(dA) · poly(dT) is

characterized by a narrow minor groove and a high degree of propeller twist and is recalcitrant to nucleosome formation in vitro (Ayami et al. 1989; Kunkel and Martinson 1981). In addition, sequences containing these tracts at specific intervals have been shown to generate curved DNA (reviewed in Hagerman 1990). The 15- and 9-fold lower affinity of mtDBP-26 for poly(dA) · poly(dT) and poly(dA-dT) · poly(dA-dT), respectively, as compared to the A+T region-containing fragment, in addition to degenerate A+T region binding, suggests high affinity mtDBP-26 binding depends on DNA conformation and/or bendability rather than a characteristic minor groove width or A+T richness per se. Curved DNA sequences have been identified in the rat mtDNA D-loop region in the vicinity of the origin of heavy-strand DNA synthesis (Pepe et al. 1989) and in human mtDNA near the origin of light-strand DNA synthesis (Welter et al. 1989). In the both cases, an activity was shown to bind to the curved DNA segment in vitro. Curved DNA has also been identified in a number of chromosomal DNA replication origins in both eukaryotes and prokaryotes and may be a universal feature of these specialized regions (Eckdahl and Anderson 1990).

Possible Roles of mtDBP-26 in Mitochondria. The precise role of mtDBP-26 in mtDNA metabolism in *D. melanogaster* remains to be determined. However, its A+T region specific DNA-binding properties suggests it may be involved in the condensation and packaging of the mtDNA in vivo. It has been shown that the mtDNAs of mouse (Nass 1969), rat (Van Tuyle and McPherson 1979), human (Albring et al. 1977; DeFrancesco and Attardi 1981), frog (Barat et al. 1985), and trypanosomes (Xu and Ray 1993) exist as nucleoprotein complexes. In most cases the complexes map to the noncoding region (Nass 1969; Van Tuyle and McPherson 1979; Albring et al. 1977; DeFrancesco and Attardi 1981). In *Drosophila* mtDNA, stretches of the A+T region are protected from the DNA crosslinking agent trimethylpsoralen in situ suggesting these sequences are bound by protein (Potter et al. 1980; Pardue et al. 1984). In *D. melanogaster*, the protected regions map to the 300 bp conserved DNA

elements located adjacent to the gene encoding $tRNA^{ile}$. Although mtDBP-26 binds to these sequences with high affinity in vitro, it also binds to other A+T region sequences. If mtDBP-26 is responsible for the observed pattern of protection of mtDNA, other factors not present in the in vitro experiments described here may influence the DNA-binding specificity of mtDBP-26 in vivo.

The HMG-like proteins identified in frog (Ghrir et al. 1991b), human (Parisi and Clayton 1991a), mouse (Fisher et al. 1989), and yeast (Diffley and Stillman 1992) mitochondria have in common the ability to condense and wrap the mtDNA and bind with low sequence specificity. The human and mouse proteins have been shown to activate transcription from mitochondrial promoters in vitro (Fisher and Clayton 1985; Fisher and Clayton 1988; Fisher et al. 1989). The precise mechanism for the activation of transcription by the HMG-like proteins remains unclear. One attractive hypothesis arises from the observation that binding to mtDNA regulatory elements occurs in a phased manner (Fisher et al. 1992; Diffley and Stillman 1992). Phased DNA binding could leave critical DNA sequences exposed for other trans-acting factors involved in mtDNA replication and transcription. In yeast, phased binding of ABF2 to the REP2 origin of mitochondrial DNA replication, which is composed entirely of deoxyadenylate and thymidylate residues, is accomplished by excluding ABF2 binding from sites containing thymidylate stretches (Diffley and Stillman 1992). Stretches of poly(dA) · poly(dT) are known to be relatively inflexible due to bifurcated hydrogen bonding from the N6 position of adenine (Nelson et al. 1987). The exclusion of the yeast ABF2 protein from deoxyadenylate (or thymidylate) stretches may be a result of a resistance of the DNA to bending. One of the two conserved thymidylate stretches in *Drosophila* mtDNA is located near the origin of leading-strand DNA synthesis. The lower binding affinity of mtDBP-26 for poly(dA) · poly(dT) relative to A+T region containing fragments and poly(dA-dT) · poly(dA-dT) in vitro might suggest mtDBP-26 is excluded from binding at the thymidylate stretch. Exclusion of mtDBP-26

would make the region more accessible to general replication proteins or general transcription proteins. A more detailed analysis of mtDBP-26 binding to A+T-region DNA must be undertaken to elucidate preferred mtDBP-26 binding sites.

CHAPTER V

SUMMARY AND PERSPECTIVES

Obtaining recombinant DNAs of the A+T region of *Drosophila melanogaster* mtDNA was an essential step towards elucidating the DNA sequence elements important for DNA replication and transcription of genes in *Drosophila* mitochondria. Sequence analysis of the A+T-region DNA revealed its extremely high content of deoxyadenylate and thymidylate nucleotides as well as the presence of the directly repeated DNA sequence elements that predominate. DNA sequence comparisons of the *D. melanogaster* mtDNA A+T region and the mtDNA A+T regions of other *Drosophila* species resulted in the definition of a 300 bp conserved DNA sequence element and the identification of conserved thymidylate stretches previously undetected. Combining the DNA sequence comparison results with electron micrographic studies performed previously allowed correlation of the conserved DNA sequence elements with the position of the mtDNA replication origin.

Repeated DNA sequence elements in the A+T region of *D. melanogaster* mtDNA had been proposed previously to exist based on restriction analysis of the mtDNA A+T regions of other closely related *Drosophila* species. However, a dearth of informative restriction enzyme recognition sites in the D. *melanogaster* A+T region prevented proof of their existence. DNA sequence analysis of the *D. melanogaster* mtDNA A+T region resulted in identification of the predicted repeated elements, termed Type II repeats, as well as identification of an additional repeat array, composed of Type I repeated DNA sequence elements. The Type I repeats are located in a different part of the A+T region and differ from the Type II repeats in both length and DNA sequence. The structure of the repeat arrays has provided insight into the possible mechanisms driving their evolution.

The DNA sequence of the genes flanking the A+T region of *Drosophila* mtDNA was also determined. The order and orientation of the genes was conserved when compared to those in the mtDNA of other *Drosophila* species whose mtDNA sequence is known. The types and positions of the nucleotide differences in the genes encoding

the tRNAs and rRNAs were found to be consistent with proposed secondary structure models of the tRNA and rRNA molecules. Obtaining the DNA sequences of these genes, as well as the DNA sequences of the A+T region and the 5' end of the gene encoding ND2, completed the nucleotide sequence of the *D. melanogaster* mitochondrial genome.

The availability of recombinant DNAs containing A+T region sequences will make possible numerous experiments designed to aid in the understanding of *Drosophila* mtDNA molecular biology and evolution. For example, slipped-strand mispairing between repeated sequences during DNA replication is one mechanism suggested to be responsible for generating polymorphic regions and large-scale deletions in animal mtDNA. The mechanism of slipped-strand mispairing during replication of *Drosophila* mtDNA sequences can now be examined in vitro. Initial experiments would utilize the *Drosophila* DNA polymerase γ and single-stranded DNA templates containing repeated segments of the *Drosophila* mtDNA A+T region. The products of in vitro replication would be analyzed by gel electrophoresis in order to detect the presence of any deletions or duplications in the region of interest. If duplications or deletions are detected, their endpoints would be mapped. The mapping data would provide a basis for the construction of DNAs containing variants of the repeated regions to be used in experiments designed to determine the DNA sequences involved in slipped-strand mispairing. In addition, it is known that the activity, processivity and fidelity of DNA polymerase γ is highly sensitive to reaction conditions. Reaction conditions would likely have an effect on the process of slipped-strand mispairing as well and should be investigated.

The recombinant DNAs of the A+T region are currently being used by others in the laboratory in mapping experiments designed to determine the 5' positions of in vivo generated transcripts. Initial results suggest that primary transcripts do in fact originate from the A+T region as hypothesized. It will be of interest to determine

whether the 5' ends of these transcripts map to the conserved DNA elements identified in the mtDNA A+T region. Knowledge of the positions of transcriptional initiation sites in vivo will expedite development of an in vitro transcription system and provide a basis for dissection of the *Drosophila* mtDNA transcription machinery.

A DNA binding activity, mtDBP-26, was identified in protein fractions obtained during the purification of DNA polymerase γ. Although the DNA binding activity co-purified with Pol γ through a number of purification steps, there is no evidence that mtDBP-26 is specifically associated with Pol γ. The DNA-binding protein was shown to have a high affinity for A+T region DNA but did not display a pronounced binding preference for specific DNA fragments of the A+T region. Its role in mtDNA metabolism remains unknown, but its size and DNA-binding characteristics suggest it could be a component of the mitochondrial nucleoid. Research that will be pursued in order to understand the role of mtDBP-26 in the mitochondrion will include purification of mtDBP-26 in large quantities and detailed examinations of the DNA binding properties of the protein on A+T region DNA.

Although mtDBP-26 co-purifies with the γ polymerase, it has been difficult to assess the yield of mtDBP-26 at the various steps of the purification due to the presence of contaminating nucleases that degrade the DNA substrate used to detect DNA binding activity. However, the observation that mtDBP-26 is heat stable suggests that contaminating nucleases could be inactivated by heating the fractions before they are assayed. Initial results using this approach have shown that mtDBP-26 activity is detectable in heat-treated fractions even at early stages of the purification and that a substantial amount of mtDBP-26 DNA-binding activity elutes outside the peak of Pol γ activity upon phosphocellulose chromatography. Eliminating large losses in yield by determining more accurately the elution profile of mtDBP-26 at each step in the purification will result a greater overall yield. A larger quantity of protein will make

possible experiments designed to identify unambiguously the protein responsible for DNA-binding, such as photoaffinity crosslinking of protein : DNA complexes.

The mechanism by which mtDBP-26 is able to bind with high affinity to A+T region DNA is not known. It is likely that, as a consequence of the A+T-rich character of the region, unusual DNA structures are present and that these structures are recognized by mtDBP-26. Experiments designed to test the effect of DNA structure on DNA binding will include competition experiments using distamycin. Distamycin specifically interacts with the minor grooves in A+T DNA sequences and is known to remove intrinsic curvature often associated DNA of this base composition. Inhibition of mtDBP-26 binding to DNA by distamycin would suggest that both recognize similar DNA conformations. Footprinting studies of mtDBP-26 bound to A+T region sequences will also be performed. These experiments will reveal whether mtDBP-26 binding to A+T region DNA occurs in a random manner or is phased. Phased binding by mtDBP-26 might indicate that it functions by leaving certain DNA sequences exposed to other trans-acting factors involved in the maintenance and expression of the *Drosophila* mitochondrial genome.

BIBLIOGRAPHY

BIBLIOGRAPHY

Albring, M., J. Griffith, and G. Attardi. 1977. Association of a protein structure of probable membrane derivation with HeLa cell mitochondrial DNA near its origin of replication. Proc. Natl. Acad. Sci. USA **74:**1348-1352.

Almasan, A., and N. C. Mishra. 1991. Recombination by sequence repeats with formation of suppressive or residual mitochondrial DNA in *Neurospora*. Proc. Natl. Acad. Sci. USA **88**:7684-7688.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden and I. G. Young. 1981. Sequence and organization of the human mitochondrial genome. Nature **290:**457-465.

Anderson, S., M. H. L. de Bruijn, A. R. Coulson, I. C. Eperon, F. Sanger and I. G. Young. 1982. Complete nucleotide sequence of bovine mitochondrial DNA. J. Mol. Biol. **156:**683-717.

Annex, B. H., and R. S. Williams. 1990. Mitochondrial DNA structure and expression in specialized subtypes of mammalian striated muscle. Mol. Cell. Biol. **10:**5671-5678.

Arnason, U., A. Gullberg and B. Widegren. 1991. The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. J. Mol. Evol. **33:**556-568.

Arnason, U., and E. Johnsson. 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. J. Mol. Evol. **34**:493-505.

Arnason, U., and A. Gullberg. 1993. Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. J. Mol. Evol. **37:**312-322.

Arnason, U., A. Gullberg, E. Johnsson, and C. Ledje. 1993. The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, *Halichoerus grypus*, and a comparison with mitochondrial sequences of other true seals. J. Mol. Evol. **37:** 323-330.

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

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

Avise, J. C., and R. M. Zinc. 1988. Molecular genetic divergence between avian sibling species: king and clapper rails, long-billed and short-billed dowitchers, boat-tailed and great-tailed grackles, and tufted and black-crested titmice. Auk **105**:516-528.

Ayami, J., M. Coll, C. A. Frederick, A. H.-J. Wang, and A. Rich. 1989. The propellar DNA conformation of poly(dA) · poly(dT). Nucl. Acids Res. **17:**3229-3245.

Barat, M., and B. Mignotte. 1981. A DNA-binding protein from *Xenopus laevis* oocyte mitochondria. Chromosoma **82:**583-593.

Barat, M., D. Rickwood, C. Dufresne, and J.-C. Mounolou. 1985. Characterization of DNA-protein complexes from the mitochondria of *Xenopus laevis* oocytes. Exp. Cell Res. **157:**207-217.

Barrell, B. G., S. Anderson, A. T. Bankier, M. H. L. de Bruijn, E. Chen, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schrier, A. J. H. Smith, R. Staden, and I. G. Young. 1980. Different pattern of codon recognition by mammalian mitochondrial tRNAs. Proc. Natl. Acad. Sci. USA **77:**3164-3166.

Benkel, B. F., P. Duschesnay, P. H. Boer, Y. Genest and D. A. Hickey. 1988. Mitochondrial large ribosomal RNA: an abundant polyadenylated sequence in *Drosophila*. Nucleic Acids Res. **16:**9880.

Bennett, J. L., and D. A. Clayton. 1990. Efficient site-specific cleavage by RNase MRP requires interaction with two evolutionarily conserved mitochondrial RNA sequences. Mol. Cell. Biol. **10:**2191-2201.

Bentzen, P., W. C. Leggett, and G. G. Brown. 1988. Length and restriction site heteroplasmy in the mitochondrial DNA of American shad (*Alosa sapidissima*). Genetics **118**:509-518.

Berthier, F., M. Renaud, S. Alziari, and R. Durand. 1986. RNA mapping on *Drosophila* mitochondrial DNA: precursors and template strands. Nucleic Acids Res. **14**:4519-4533.

Bhat, K. S., N. Avdalovic, N. G. Avadhami. 1989. Characterization of primary transcripts and identification of transcription initiation sites on the heavy and light strands of mouse mitochondrial DNA. Biochem. **28:**763-769.

Bibb, M. J., R. A. Van Etten, C. T. Wright, M. W. Walberg and D. A. Clayton. 1981. Sequence and gene organization of mouse mitochondrial DNA. Cell **26:**167-180.

Biswas, T. K., and G. S. Getz. 1986. Nucleotides flanking the promoter sequence influence the transcription of the yeast mitochondrial gene coding for ATPase subunit 9. Proc. Natl. Acad. Sci USA **83:**270-274.

Biswas, T. K. 1990. Control of mitochondrial gene expression in the yeast *Saccharomyces cerevisae*. Proc. Natl. Acad. Sci. USA **87:**9338-9342.

Bogenhagen, D., and D. A. Clayton. 1978. Mechanism of mitochondrial DNA replication in mouse L-cells: kinetics of synthesis and turnover of the initiation sequence. J. Mol. Biol. **119:**49-68.

Bogenhagen, D. F., E. F. Applegate, and B. K. Yoza. 1984. Identification of a promoter for transcription of the heavy strand of human mtDNA: In vitro transcription and deletion mutagenesis. Cell **36:**1105-1113.

Bogenhagen, D. F., B. K. Yoza, and S. S. Cairns. 1986. Identification of initiation sites for transcription of *Xenopus laevis* mitochondrial DNA. J. Biol. Chem. **261:**8488-8494.

Bogenhagen, D. F., and B. K. Yoza. 1986. Accurate in vitro transcription of *Xenopus laevis* mitochondrial DNA from two bidirectional promoters. Mol. Cell. Biol. **6:**2543-2550.

Bogenhagen, D. F., and M. F. Romanelli. 1988. Template sequences required for transcription of *Xenopus laevis* mitochondrial DNA from two bidirectional promoters. Mol. Cell. Biol. **8:**2917-2924.

Bogenhagen, D. F., and N. F. Insdorf. 1988. Purification of *Xenopus laevis* mitochondrial RNA polymerase and identification of a dissociable factor required for specific transcription. Mol. Cell. Biol. **8:**2910-2916.

Boyce, T. M., M. E. Zwick, and C. F. Aquadro. 1989. Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy. Genetics **123**:825-836.

Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18:**225-239.

Buroker, N. E., J. R. Brown, T. A. Gilbert, P. J. O'hara, A. T. Beckenbach, W. K. Thomas, and M. J. Smith. 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics **124**:157-163.

Caccone, A., G. D. Amato, and J. R. Powell. 1988. Rates and Patterns of scnDNA and mtDNA divergence within the *Drosophila melanogaster* subgroup. Genetics. **118:**671-683.

Campbell, J. L. 1986. Eukaryotic DNA replication. Annu. Rev. Biochem. **55**:733-771.

Cantatore, P., M. Roberti, G. Rainaldi, M. N. Gadaleta, and C. Saccone. 1989. The complete nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of *Paracentrotus lividus*. J. Biol. Chem. **264:**10965-10975.

Caron, F., C. Jacq, and J. Rouviere-Yaniv. 1979. Characterization of a histone-like protein extracted from yeast mitochondria. Proc. Natl. Acad. Sci. USA **76:**4265-4269.

Certa, U., M. Colavito-Shepanski, and M. Grustein. 1984. Yeast may not contain histone H1: the only known 'histone H1-like' protein in *Saccharomyces cerevisae* is a mitochondrial protein. Nucl. Acids Res. **12:**7975-7985.

Chang, D. D., and D. A. Clayton. 1984. Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell **36:**635-643.

Chang, D. D., W. W. Hauswirth, and D. A. Clayton. 1985. Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. EMBO J. **4**:1559-1567.

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

Chang, D. D., and D. A. Clayton. 1986a. Identification of primary transcriptional start sites of mouse mitochondrial DNA: accurate in vitro initiation of both heavy- and light-strand transcripts. Mol. Cell. Biol. **6**:1446-1453.

Chang, D. D., and D. A. Clayton. 1986b. Precise assignment of the light-strand promoter of mouse mitochondrial DNA: accurate in vitro initiation of both heavy- and light-strand transcripts. Mol. Cell. Biol. **6:**3253-3261.

Chang, D. D., and D. A. Clayton. 1986c. Precise assignment of the heavy-strand promoter of mouse mitochondrial DNA; cognate start sites are not required for transcriptional initiation. Mol. Cell. Biol. **6:**3262-3267.

Chang, D. D., and D. A. Clayton. 1987a. A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. EMBO J. **6:**409-417.

Chang, D. D., and D. A. Clayton. 1987b. A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. Science **235:**1178-1184.

Chang, D. D., and D. A. Clayton. 1989. Mouse RNAase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. Cell **56:**131-139.

Chen, J.-Y., and N. C. Martin. 1988. Biosynthesis of tRNA in yeast mitochondria: an endonuclease is responsible for the 3'-processing of tRNA precursors. Proc. Natl. Acad. Sci. USA **263:**13677-13682.

Chomyn, A., P. Mariottini, N. Gonzalez-Cadavid, G. Attardi, D. D. Strong, D. Trovato, M. Riley, and R. F. Doolittle. 1983. Identification of the polypeptides encoded in the ATPase 6 gene and in the unassigned reading frames 1 and 3 of human mtDNA. Proc. Natl. Acad. Sci. USA **80:**5535-5539.

Chomyn, A., P. Mariottini, M. W. J. Cleeter, C. I. Ragan, A. Matsuno-Yagi, Y. Hatefi, R. F. Doolittle, and G. Attardi. 1985. Six unidentified reading frames of human mitochondrial DNA encode components of the respiratory-chain NADH dehydrogenase. Nature **314:**592-597.

Chomyn, A., M. W. J. Cleeter, C. I. Ragan, M. Riley, R. F. Doolittle, and G. Attardi 1986. URF6, last unidentified reading frame of human mtDNA, codes for an NADH dehydrogenase subunit. Science **234:**614-618.

Chomyn, A., and G. Attardi. 1987. Mitochondrial gene products. Curr. Top. Bioener. **15:**295-239.

Chomyn, A., A. Martinuzzi, M. Yoneda, A. Daga, O. Hurko, D. Johns, S. T. Lai, I. Nonaka, C. Angelini, and G. Attardi. 1992. MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis but no change in levels of upstream and downstream mature transcripts. Proc. Natl. Acad. Sci. USA **89:**4221-4225.

Christianson, T., and M. Rabinowitz. 1983. Identification of multiple transcriptional initiation sites on the yeast mitochondrial genome by in vitro capping with guanylyltransferase. J. Biol. Chem. **258:**14025-14033.

Clark-Walker, G. D. 1985. Basis of diversity in mitochondrial DNAs. Pp. 277-297 *in* T. Cavalier-Smith, ed. The Evolution of Genome Size. Wiley, New York.

Clark-Walker, G. D. 1989. *In vivo* rearrangement of mitochondrial DNA in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **86**:8847-8851.

Clary, D. O., J. M. Goddard, S. C. Martin, C. M.-R. Fauron and D. R. Wolstenholme. 1982. *Drosophila* mitochondrial DNA: a novel gene order. Nucleic Acids Res. **10:**6619-6637.

Clary, D. O., J. A. Wahleithner and D. R. Wolstenholme. 1983. Transfer RNA genes in *Drosophila* mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial tRNA genes. Nucleic Acids Res. **11:**2411-2425.

Clary, D. O., and D. R. Wolstenholme. 1985a. The ribosomal RNA genes of *Drosophila* mitochondrial DNA. Nucleic Acids Res. **13:**4029-4045.

Clary, D. O., and D. R. Wolstenholme. 1985b. The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. **22**:252-271.

Clary, D. O., and D. R. Wolstenholme. 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.

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

Clayton, D. A. 1984. Transcription of the mammalian mitochondrial genome. Annu. Rev. Biochem. **53:**573-594.

Clayton, D. A. 1991a. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. **7**:453-478.

Clayton, D. A. 1991b. Nuclear gadgets in mitochondrial DNA replication and transcription. Trends Biol. Sci. **16:**107-111.

Cornuet, J.-M., L. Garnery, and M. Solignac. 1991. Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. Genetics **128**:393-403.

Cressie, N. A. C., and D. D. Keightley. 1981. Analysing data from hormone-receptor assays. Biometrics **37:**235-249.

Crozier, R. H., Y. C. Crozier and A. G. Mackinlay. 1989. The COI and COII region of honey-bee mitochondrial DNA: evidence for variation in insect mitochondrial rates. Mol. Biol. Evol. **6:**399-411.

Crozier, R. H., and Y. C. Crozier. 1993. The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization. Genetics **133**:97-117.

de Bruijn, M. H. L. 1983. *Drosophila melanogaster* mitochondrial DNA, a novel organization and genetic code. Nature **304:**234-241.

Daga, A., V. Michol, D. Hess, R. Abersold, and G. Attardi. 1993. Molecular characterization of the transcriptional termination factor from human mitochondria. J. Biol. Chem. **268:**8123-8130.

Davis, S. C., A. Tzagoloff, and S. R. Ellis. 1992. Characterization of a yeast mitochondrial ribosomal protein structurally related to the mammalian 68-kDa high affinity laminin receptor. J. Biol. Chem. **267:**5508-5514.

DeFrancesco, L., and G. Attardi. 1981. In situ photochemical crosslinking of HeLa cell mitochondrial DNA by a psoralen derivative reveals a protected region near the origin of replication. Nucl. Acids Res. **9:**6017-6030.

Delucia, A. L., D. Sumitra, K. Partin, and P. Tegtmeyer. 1986. Functional interactions of the simian virus 40 core origin of replication with flanking regulatory sequences. J. Virol. **57**:138-144.

Denslow, N. D., G. S. Michaels, J. Montoya, G. Attardi, and T. W. O'Brien. 1989. Mechanism of mRNA binding to bovine mitochondrial ribosomes. J. Biol. Chem. **264**:8328-8338.

Densmore, L. P., J. W. Wright, and W. M. Brown. 1985. Length variation and heteroplasmy are frequent in mitochondrial DNA from parthenogenetic and bisexual lizards (genus *Cnemidophorus*). Genetics **110**:689-707.

Diffley, J. F. X., and B. Stillman. 1991. A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. Proc. Natl. Acad. Sci. USA **88**:7864-7868.

Diffley, J. F. X., and B. Stillman. 1992. DNA binding properties of an HMG-related protein from yeast mitochondria. J. Biol. Chem. **267**:3368-3374.

DeSalle, R., T. Freedman, E. M. Prager, and A. C. Wilson. 1987. Tempo and Mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol. **26**:157-164.

Desjardins, P., and R. Morais. 1990. Sequence and gene organization of the chicken mitochondrial genome. J. Mol. Biol. **212**:599-634.

de Zamaroczy, M., and G. Bernardi. 1985. Sequence organization of the mitochondrial genome of yeast - a review. Gene **37**:1-17.

de Zamroczy, M., G. Faugeron-Fonty, G. Baldacci, R. Goursot, and G. Bernardi. 1984. The *ori* sequences of the mitochondrial genome of a wild-type yeast strain: number, location, orientation and structure. Gene **32**:439-457.

Doda, J. N., C. T. Wright, and D. A. Clayton. 1981. Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. USA **78**:6116-6120.

Doerson, C.-J., C. Guerrier-Takada, S. Altman, and G. Attardi. 1985. Characterization of an RNase P activity from HeLa cell mitochondria: comparison with the cytosol RNase P activity. J. Biol. Chem. **260**:5942-5249.

Dunon-Bluteau, D. C., and G. M. Brun. 1987. Mapping at the nucleotide level of *Xenopus laevis* mitochondrial D-loop H-strand: structural features of the 3' region. Biochem. Int. **14**:643-657.

Eckdahl, T. T., and J. N. Anderson. 1990. Conserved DNA structures in origins of replication. Nucl. Acids Res. **18**:1609-1612.

Einck, L., and M. Bustin. 1985. The intracellular distribution and function of the high mobility group chromosomal proteins. Exp. Cell Res. **156**:295-310.

Faugeron-Fonty, G., Le Van Kim, C., M. de Zamaroczy, R. Goursot, and G. Bernardi. 1984. A comparative study of the ori sequences from the mitochondrial genomes of twenty wild-type yeast strains. Gene **32**:459-473.

Fauron, C. M.-R., and D. R. Wolstenholme. 1976. Structural heterogeneity of mitochondrial DNA molecules within the genus *Drosophila*. Proc. Natl. Acad. Sci. USA **73**:3623-3627.

Fauron, C. M.-R., and D. R. Wolstenholme. 1980a. Extensive diversity among *Drosophila* species with respect to nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules. Nucleic Acids Res. **8**:2439-2452.

Fauron, C. M.-R., and D. R. Wolstenholme. 1980b. Intraspecific diversity of nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules of *Drosophila mauritiana*, *Drosophila melanogaster* and *Drosophila simulans*. Nucleic Acids Res. **8**:5391-5410.

Fisher, R. P., and D. A. Clayton. 1985. A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. J. Biol. Chem. **260**:11330-11338.

Fisher, R. P., J. N. Topper, and D. A. Clayton. 1987. Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. Cell **50**:247-258.

Fisher, R. P., and D. A. Clayton. 1988. Purification and characterization of human mitochondrial transcription factor 1. Mol. Cell. Biol. **8**:3496-3509.

Fisher, R. P., M. A. Parisi, and D. A. Clayton. 1989. Flexible recognition of rapidly evolving promoter sequences by mitochondrial transcription factor 1. Genes Dev. **3**:2202-2217.

Fisher, R. P., Lisowsky, T., M. A. Parisi, and D. A. Clayton. 1992. DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. J. Biol. Chem. **267**:3358-3367.

Foran, D. R., J. E. Hixson, and W. M. Brown. 1988. Comparisons of ape and human sequences that regulate mitochondrial DNA transcription and D-loop DNA synthesis. Nucleic Acids Res. **16**:5841-5861.

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.

Freeman, K. B. 1970. Inhibition of mitochondrial and bacterial protein synthesis by chloramphenicol. Can. J. Biochem. **48**:479-485.

Gadaleta, G., G. Pepe, G. De Candia, C. Quagliariello, E. Sbisa and C. Saccone. 1989. The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparitive analysis between vertebrates. J. Mol. Evol. **28:**497-516.

Garey, J. R., and D. R. Wolstenholme. 1989. Platyhelminth mitochondrial DNA: evidence for early origin of a tRNA-ser(AGN) that contains a dihydrouridine arm replacement-loop, and of serine-specifying AGA and AGG codons. J. Mol. Evol. **28:**374-387.

Garesse, R. 1988. *Drosophila melanogaster* mitochondrial DNA: Gene organization and evolutionary considerations. Genetics **118**:649-663.

Gelfand, R., and G. Attardi. 1981. Synthesis and turnover of mitochondrial ribonucleic acid in HeLa cells: the mature ribosomal and messenger ribonucleic acid species are metabolically unstable. Mol. Cell. Biol. **1:**497-511.

Genetics Computer Group (1991) Program manual for the GCG package, version 7. Genetics Computer Group, Madison.

Ghrir, R., J.-P. Lecaer, C. Dufresne, M. Gueride. 1991a. Primary structure of the two variants of *Xenopus laevis* mtSSB, a mitochondrial DNA binding protein. Arch. Biochem. Biophys. **291:**395-400.

Ghrir, R., B. Mignotte, and M. Gueride. 1991b. Amino terminal sequence of mitochondrial protein mtDBP-C: similarity with nonhistone chromosomal proteins HMG 1 and 2. Biochimie **73:**615-617.

Ghivizzani, S. C., C. S. Madsen, and W. W. Hauswirth. 1993a. In-organello footprinting: analysis of protein binding at regulatory regions in bovine mitochondrial DNA. J. Biol. Chem. **268:**8675-8682.

Ghivizzani, S. C., S. L. D. MacKay, C. S. Madsen, P. J. Laipis, and W. W. Hauswirth. 1993b. Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. J. Mol. Evol. **37**:36-47.

Gjetvaj, B., D. I. Cook, and E. Zouros. 1992. Repeated sequences and large-scale size variation of mitochondrial DNA: a common feature among scallops (*Bivalvia*: *Pectinidae*). Mol. Biol. Evol. **9**:106-124.

Goddard, J. M., and D. R. Wolstenholme. 1978. Origin and direction of replication in mitochondrial DNA molecules from *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **75**:3886-3890.

Goddard, J. M., and D. R. Wolstenholme. 1980. Origin and direction of replication in mitochondrial DNA molecules from the genus *Drosophila*. Nucleic Acids Res. **8**:741-757.

Goldenthal, M., and J. T. Nishiura. 1987. Isolation and characterization of a mitochondrial RNA polymerase from Drosophila melanogaster. Biochem. Cell Biol. **65:**173-182.

Grant, D., and K.-S. Chiang. 1980. Physical mapping and characterization of *Chlamydomonas* mitochondrial DNA molecules: their unique ends, sequence homogeneity, and conservation. Plasmid **4:**82-96.

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

Gray, M. W., D. Sankoff, and R. J. Cedergren. 1984. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. Nucl. Acids Res. **12:**5837-5852.

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

Gray, M. W., and W. F. Doolittle. 1982. Has the endosymbiont hypothesis been proven? Microbiol. Rev. **46:**1-42.

Gray, M. W. 1988. Organelle origins and ribosomal RNA. Biochem. Cell Biol. **66:**325-348.

Greenleaf, A. L., J. L. Kelly, and I. R. Lehman. 1986. Yeast *RPO41* gene product is required for transcription and maintenance of the mitochondrial genome. Proc. Natl. Acad. Sci. USA **83:**3391-3399.

Gutell, R. R., and G. E. Fox. 1988. A compilation of large subunit RNA sequences presented in a structural format. Nucleic Acids Res. **16:**r175-r269.

Gutell, R. R. 1993. Collection of small subunit (16S- and 16S-like) ribosomal RNA structures. Nucleic Acids Res. **21:**3051-3054.

Hagerman, P. J. 1990. Sequence-directed curvature of DNA. Annu. Rev. Biochem. **59:**755-781.

Hale L. R., and R. S. Singh. 1986. Extensive variation and heteroplasmy in size of mitochondrial DNA among geographic populations of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **83**:8813-8817.

Hamilton, M. G., and T. W. O'Brien. 1974. Ultracentrifugal characterization of the mitochondrial ribosome and subribosomal particles of bovine liver. Biochem. **13:**5400-5403.

Hayasaka, K., T. Ishida, and S. Horai. 1991. Heteroplasmy and polymorphism in the major noncoding region of mitochondrial DNA in japanese monkeys: association with tandemly repeated sequences. Mol. Biol. Evol. **8**:399-415.

Hess, J. F., M. A. Parisis, J. L. Bennett, and D. A. Clayton. Impairment of mitochondrial transcription termination by a point mutation associated with the MELAS subgroup of mitochondrial encephalopathies. Nature **351:**236-239.

Hixson, J. E., and D. A. Clayton. 1985. Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. Proc. Natl. Acad. Sci. USA **82:**2660-2664.

Hixson, J. E., T. W. Wong, and D. A. Clayton. 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:**2384-2390.

Hixson, J. E., and W. M. Brown. 1986. A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequence structure, evolution, and phylogenetic implications. Mol. Biol. Evol. **3:**1-18.

Hoelzel, A. R., J. M. Hancock, and G. Dover. 1993. Generation of VNTRs and heteroplasmy by sequence turnover in the mitochondrial control region of two elephant seal species. J. Mol. Evol. **37**:190-197.

Hoke, G. D., P. A. Pavco, B. J. Ledwith, and G. C. Van Tuyle. 1990. Structural and functional studies of the rat mitochondrial single strand DNA binding protein P16. Arch. Biochem. Biophys. **282:**116-124.

Hollingsworth, M. J., and N. C. Martin. 1986. RNase P activity in the mitochondria of *Saccharomyces cerevisae* depends on both mitochondria and nucleus-encoded components. Mol. Cell. Biol. **6:**1058-1064.

Insdorf, N. F., and D. F. Bogenhagen. 1989a. DNA polymerase γ 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 D. F. Bogenhagen. 1989b. DNA polymerase γ from *Xenopus laevis* II. A 3'-5' exonuclease is tightly associated with the DNA polymerase activity. J. Biol. Chem. **264:**21498-21503.

Jacobs, H. T., D. J. Elliott, V. B. Math, and A. Farquharson. 1988. Nucleotide sequence and gene organization of sea urchin mitochondrial DNA. J. Mol. Biol. **202:**185-217.

Jacobs, H. T., E. R. Herbert, and J. Rankine. 1989. Sea urchin egg mitochondrial DNA contains a short displacement loop (D-loop) in the replication origin. Nucl. Acids Res. **17:**8949-8965.

Jang, S. H., and J. A. Jaehning. 1991. The yeast mitochondrial RNA polymerase specificity factor, MTF1, is similar to bacterial σ factors. J. Biol. Chem. **266:**22671-22677.

Jantzen, H.-M., A. Admon, S. P. Bell, and R. Tijan. 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature **344:**830-836.

Jayanthi, G. P., and G. C. Van Tuyle. 1992. Characterization of ribonuclease P isolated from rat liver cytosol. Arch. Biochem. Biophys. 296:264-270.

Johansen, S., P. H. Guddal, and T. Johansen. 1990. Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua*. Nucleic Acids. Res. **18:**411-419.

Johns, E. W., G. H. Goodwin, J. R. B. Hastings, J. M. Walker. 1977. Organization and expression of the eukaryotic genome, pp. 3-19. E. M. Bradbury and K. Javaherian, Eds. Academic Press, London.

Kaguni, J. M., and D. S. Ray. 1979. Cloning of a functional replication origin of phage G4 into the genome of phage M13. J. Mol. Biol. **135**:863-878.

Kaguni, J. M., and L. S. Kaguni. 1992. Enzyme labeled probes for nucleic acid hybridization. Pp. 115-127 *in* C. H. Suelter and L. Kricka, eds. Methods of Biochemical Analysis. Vol. **36**. Bioanalytical applications of enzymes. John Wiley and Sons, Inc., New York, Chichester, Brisbane, Toronto, Singapore.

Kaguni, L. S., C. W. Wernette, M. C. Conway, and P. Yang-Cashman. 1988. Structural and catalytic features of the mitochondrial DNA polymerase from *Drosophila melanogaster* embryos. Pp. 425-432 *in* T. J. Kelly and B. W. Stillman, eds. Cancer Cells **6**: Eukaryotic DNA Replication. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

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

Karawya, E. M., and R. G. Martin. 1987. Monkey (CV-1) mitochondrial DNA contains a unique triplication of 108 bp in the origin region. Biochim. et Biophys. Acta **909**:30-34.

Kelly, J. L., and I. R. Lehman. 1986. Yeast mitochondrial RNA polymerase. J. Biol. Chem. **261:**10340-10347.

Kelly, M., J. Burke, M. Smith, A. Klar, and D. Beach. 1988. Four mating-type genes control sexual differentiation in the fission yeast. EMBO. J. **7:**1537-1547.

King, T. C., and R. L. Low. 1987. Mapping of control elements in the displacement loop region of bovine mitochondrial DNA. J. Biol. Chem. **262:**6204-6213.

Kiss, T., and W. Filipowicz. 1992. Evidence against a mitochondrial location of the 7-2/MRP RNA in mammalian cells. Cell **70:**11-20.

Kitakawa, M., and K. Isono. 1991. The mitochondrial ribosomes. Biochimie **73:**813-825.

Klukas, C. K., and I. B. Dawid. 1976. Characterization and mapping of mitochondrial ribosomal RNA and mitochondrial DNA in *Drosophila melanogaster.* Cell **9**:615-625.

Kobayashi, S., and M. Okada. 1990. Complete cDNA sequence encoding mitochondrial large ribosomal RNA of *Drosophila melanogaster.* Nucleic Acids Res. **18:**4592.

Kovac, L., J. Lazowska, P. P. Slonimski. 1984. A yeast with linear molecules of mitochondrial DNA. Mol. Gen. Genet. **197:**420-424.

Kruse, B., N. Narasimhan, and G. Attardi. 1989. Termination of transcription in human mitochondrial: identification and purification of a DNA binding protein factor that promotes termination. Cell **58:**391-397.

Kumazawa, Y., T. Yokogawa, E. Hasegawa, K. Miura and K. Watanabe. 1989. The aminoacylation of structurally variant phenylalanine tRNAs from mitochondria and various nonmitochondrial sources by bovine mitochondrial phenylalanyl-tRNA synthetase. J. Biol. Chem. **264:**13005-13011.

Kumazawa, Y., H. Himeno, K. Miura and K. Watanabe. 1991. Unilateral aminoacylation specificity between bovine mitochondria and eubacteria. J. Biochem. **109:**412-427.

Kunkel, G. R., and H. G. Martinson. 1981. Nucleosomes will not form on double-stranded RNA of over poly(dA) · poly(dT) tracts in recombinant DNA. Nucl. Acids Res. **9:**6869-6888.

Kunkel, T. A. 1985. The mutational specificity of DNA polymerase-α and -γ during *in vitro* DNA synthesis. J. Biol. Chem. **260**:12866-12874.

Kunkel, T. A., and P. S. Alexander. 1986. The base substitution fidelity of eucaryotic DNA polymerases. J. Biol. Chem. **261**:160-166.

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

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

L'Abbe, D., J.-F. Duhaime, B. F. Lang, and R. Morais. 1991. The transcription of DNA in chicken mitochondria initiates from one bidirectional promoter. J. Biol. Chem. **266:**10844-10850.

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

Lamb, A. J., G. D. Clark-Walker, and A. W. Linnane. 1968. The biogenesis of mitochondria. 4. The differentiation of mitochondrial and cytoplasmic protein synthesizing systems in vitro by antibiotics. Biochim. Biophys. Acta. **161:**415-427.

La Roche J., M. Snyder, D. I. Cook, K. Fuller, and E. Zouros. 1990. Molecular characterization of a repeat element causing large-scale size variation in the mitochondrial DNA of the sea scallop *Placopecten magellanicus*. Mol. Biol. Evol. **7**:45-64.

Levinger, L. F. 1985. D1 protein of *Drosophila melanogaster*: purification and AT-DNA binding properties. J. Biol. Chem. **260:**14311-14318.

Lisowsky, T., and G. Michaelis. 1988. A nuclear gene essential for mitochondrial replication suppresses a defect of mitochondrial transcription in *Saccharomyces cerevisae*. Mol. Gen. Genet. **214:**218-223.

Mackay, S. L. D., P. D. Olivo, P. J. Laipis, and W. W. Hauswirth. 1986. Template-directed arrest of mammalian mitochondrial DNA synthesis. Mol. Cell Biol. **6:**1261-1267.

Madsen, C. S., S. C. Ghivizzani, and W. W. Hauswirth. 1993. Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Mol. Cell Biol. **13:**2162-2171.

Margulis, L. 1981. Symbiosis in cell evolution. W. H. Freeman and Co., San Francisco, CA.

Marini, J. C., Levene, S. D., Crothers, D. M., and P. T. Englund. 1982. Bent helical structure in kinetoplast DNA. Proc. Natl. Acad. Sci. USA. **79:**7664-7668.

Mason, P. J., and J. O. Bishop. 1980. Molecular cloning of part of the mitochondrial DNA of *Drosophila melanogaster*. Biochem. Biophys. Res. Comm. **95**:1268-1274.

Masters, B. S., L. L. Stohl, and D. A. Clayton. Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. Cell **51:**89-99.

Matsumoto, L., H. Kasamatsu, L. Piko, and J. Vinograd. 1974. Mitochondrial DNA replication in sea urchin oocytes. J. Cell. Biol. **63:**146-159.

McClain, W. H. 1993. Rules that govern tRNA identity in protein synthesis. J. Mol. Biol. **234:**257-280.

McGhee, J. D., and G. Felsenfeld. 1980. Nucleosome structure. Annu. Rev. Biochem. **49:**1115-1156.

Merten, S. H., and M. L. Pardue. 1981. Mitochondrial DNA in *Drosophila*. An analysis of genome organization and transcription in *Drosophila melanogaster* and *Drosophila virilis*. J. Mol. Biol. **153**:1-21.

Michaels, G. S., W. W. Hauswirth, and P. J. Laipis. 1982. Mitochondrial copy number in bovine oocytes and somatic cells. Dev. Biol. **94:**246-251.

Mignotte, B., M. Barat, J. Marsault, J.-C. Mounolou. 1983. Mitochondrial DNA binding proteins that bind preferentially to supercoiled molecules containing the D-loop region of *Xenopus laevis* mtDNA. Biochem. Biophys. Res. Comm. **117:**99-107.

Mignotte, B., M. Barat, and J.-C. Mounolou. 1985 Characterization of a mitochondrial protein binding to single-stranded DNA. Nucl. Acids Res. **13:**1703-1716.

Mignotte, B., and M. Barat. 1986. Characterization of a *Xenopus laevis* mitochondrial protein with a high affinity for supercoiled DNA. Nucleic Acids Res. **14:**5969-5980.

Mignotte, B., Delain, E., Rickwood D., and M. Barat-Gueride. 1988. The *Xenopus laevis* mitochondrial protein mtDBP-C cooperatively folds the DNA in vitro. EMBO J. **7:**3873-3879.

Mignotte, F., M. Gueride, A.-M. Champagne, and J.-C. Mounolou. 1990. Direct repeats in the non-coding region of rabbit mitochondrial DNA. Involvement in the generation of intra- and inter-individual heterogeneity. Eur. J. Biochem. **194**:561-571.

Mignotte, B., B. Theveny, and B. Revet. 1990. Structural modifications induced by the mtDBP-C protein in the replication origin of *Xenopus laevis* mitochondrial DNA. Biochimie **72:**65-72.

Moazed, D., and H. F. Noller. 1989. Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. Cell **57:**585-597.

Monforte, A., E. Barrio, and A. Latorre. 1993. Characterization of the length polymorphism in the A+T-rich region of the *Drosophila obscura* group species. J. Mol. Evol. **36**:214-223.

Monnerot, M., M. Solignac, and D. R. Wolstenholme. 1990. Discrepancy in divergence of the mitochondrial and nuclear genomes of *Drosophila teissieri* and *Drosophila yakuba*. J. Mol. Evol. **30**:500-508.

Montoya, J., D. Ojala, and G. Attardi. 1981. Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs. Nature **290:**465-470.

Moritz, C., T. E. Dowling, and W. M. Brown. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. **18**:269-292.

Mosbaugh, D. W. 1988. Purification and characterization of porcine liver DNA polymerase γ: utilization of dUTP and dTTP during in vitro synthesis. Nucl. Acids Res. **16:**5645-5659.

Muramatsu, T., K. Nishikawa, F. Nemoto, Y. Kuchino, S. Nishimura, T. Miyazawa and S. Yokoyama. 1988. Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. Nature **336:**179-181.

Nass, M. M. K., and S. Nass. 1963. Intramitochondrial fibers with DNA characteristics. I. Fixation and electron staining reactions. J. Cell Biol. **19:**593-611.

Nass, M. M. K. 1969. Mitochondrial DNA: I. Intramitochondrial distribution and structural relations of single- and double-length circular DNA. J. Mol. Biol. **42:**521-528.

Nelson, H. C. M., J. T. Finch, B. F. Luisi, and A. Klug. 1987. The structure of an oligo(dA) . oligo(dT) tract and its biological implications. Nature **330:**221-226.

Neefs, J.-M., Y. Van de Peer, P. De Rijk, S. Chapelle and R. De Wachter. 1993. Compilation of small ribosomal subunit RNA structures. Nucleic Acids Res. **21:** 3025-3049.

Noller, H. F. 1991. Ribosomal RNA and translation. Annu. Rev. Biochem. **60:**191-227.

Noller, H. F., V. Hoffarth and L. Ziminiak. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. Science **256:**1416-1419.

Oakley, B. R., D. R. Kirsch, and N. R. Morris. 1980. A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. Anal. Biochem. **105:**361-363.

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

Okimoto, R., H. M. Chamberlin, J. L. Macfarlane, and D. R. Wolstenholme. 1991. Repeated sequence sets in mitochondrial DNA molecules of root knot nematodes (*Meloidogyne*): nucleotide sequences, genome location and potential for host-race identification. Nucleic Acids Res. **19**:1619-1626.

Okimoto, R., J. L. Macfarlane, D. O. Clary, and D. R. Wolstenholme. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. Genetics **130:** 471-498.

Olson, M. W., and L. S. Kaguni. 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.

Otwinoski, Z., R. W. Schevitz, R.-G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi, and P. B. Sigler. 1988. Crystal structure of the trp repressor-operator complex at near atomic resolution. Nature **335:**321-329.

Pardue, M. L., J. M. Fostel, and T. R. Cech. 1984. DNA-protein interactions in the *Drosophila virilis* mitochondrial chromosome. Nucleic Acids Res. **12:**1991-1999.

Parisi, M. A., and D. A. Clayton. 1991. Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science **252:**965-969.

Parisi, M. A., B. Xu, and D. A. Clayton. 1993. A human mitochondrial transcriptional activator can functionally replace a yeast mitochondrial HMG-box protein both in vitro and in vivo. Mol. Cell. Biol. **13:**1951-1961.

Pepe, G., G. Gadaleta, G. Palazzo, and C. Saccone. 1989. Sequence-dependent DNA curvature: conformational signal present in the main regulatory region of the rat mitochondrial genome. Nucl. Acids Res. **17:**8803-8819.

Pica-Mattoccia, L., and G. Attardi. 1972. Expression of the mitochondrial genome in HeLa cells. IX. Replication of mitochondrial DNA in relationship to the cell cycle in HeLa cells. J. Mol. Biol. **64:**465-484.

Pissios, P., and Z. G. Scouras. 1993. Mitochondrial DNA evolution in the *Montium*-species subgroup of *Drosophila*. Mol. Biol. Evol. **10**:375-382.

Pfanner, N., and W. Neupert. 1990. The mitochondrial protein import apparatus. Annu. Rev. Biochem. **59:**331-353.

Pollack, J. K., and R. Sutton. 1980. The differentiation of animal mitochondria during development. Trends Biochem. Sci. **5:**23-27.

Potter, D. A., J. M. Fostel, M. Berninger, M. L. Pardue, and T. R. Cech. 1980. DNA-protein interactions in the *Drosophila melanogaster* mitochondrial genome as deduced from trimethylpsoralen crosslinking patterns. Proc. Natl. Acad. Sci. USA **77:**4118-4122.

Powers, T., and H. F. Noller. 1993. Evidence for functional interaction between elongation factor Tu and 16S ribosomal RNA. Proc. Natl. Acad. Sci. USA **90:** 1364-1368.

Pritchard, A. E., J. J. Seilhammer, and D. J. Cummings. 1986. *Paramecium* mitochondrial DNA sequences and RNA transcripts for cytochrome oxidase subunit I, URFI, and three ORFs adjacent to the replication origin. Gene **44:**243-53.

Qureshi, S. A., and H. T. Jacobs. 1993. Characterization of a high-affinity binding site for a DNA binding protein form sea urchin embryo mitochondria. Nucl. Acids Res. **21:**811-816.

Raitio, M., T. Jalli, and M. Saraste. 1987. Isolation and analysis of the genes for cytochrome *c* oxidase in *Paracoccus dentrificans.* EMBO J. **6:**2825-2833.

Rand, D. M., and R. G. Harrison. 1986. Mitochondrial DNA transmission genetics in crickets. Genetics **114:**955-970.

Rand, D. R., and R. G. Harrison. 1989. Molecular population genetics of mtDNA size variation in crickets. Genetics **121**:551-569.

Rand, D. M. 1993. Endotherms, Ectotherms, and mitochondrial genome-size variation. J. Mol. Evol. **37:**281-295.

Raven, P. H. 1970. A multiple origin for plastids and mitochondria. Science **169:**641-646.

Reeck, G. R., P. J. Isackson, and D. C. Teller. 1982. Domain structure in high molecular weight high mobility group nonhistone chromatin proteins. Nature **300:**76-78.

Revzin, A. 1989. Gel electrophoresis assays for DNA-protein interactions. Biotechniques. **7:**346-355.

Richmond, T. J., J. T. Finch, B. Rushton, D. Rhodes, and A. Klug. 1984. Structure of the nucleosome core particle at 7Å resolution. Nature **311:**532-537.

Riggs, A. D., S. Bourgeois, and M. Cohn. 1970. The lac repressor-operator interaction: III. Kinetic studies. J. Mol. Biol. **53:**401-417.

Roberti M., A. Mustich, M. N. Gadaleta, and P. Cantatore. 1991. Identification of two homologous mitochondrial DNA sequences, which bind strongly and specifically to a mitochondrial protein of *Paracentroutus lividis.* Nucl. Acids Res. **19:**6249-6254.

Roe, B. A., J. F. H. Wong, E. Y. Chen, P. W. Armstrong, A. Stankiewicz, et al. 1982. Mitochondrial Genes. Pp. 45-49. Cold Spring Harbor, NY: Cold Spring Harbor Lab. 500 pp.

Roe, B. A., D.-P. Ma, R. K. Wilson, and J. F.-H. Wong. 1985. The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. J. Biol. Chem. **260:** 9759-9774.

Rogers, M. J., and D. Soll. 1988. Discrimination between glutaminyl-tRNA synthetase and seryl-tRNA synthetase involves nucleotides in the acceptor helix of tRNA. Proc. Natl. Acad. Sci. USA **85:**6627-6631.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sampson, J. R., A. B. DiRenzo, L. S. Behlen and O. C. Uhlenbeck. 1990. Role of the tertiary nucleotides in the interaction of yeast phenylalanine tRNA with its cognate synthetase. Biochem. **29:** 2523-2532.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463-5467.

Satta, Y., H. Ishiwa and S. I. Chigusa. 1987. Analysis of nucleotide substitutions of mitochondrial DNA in *Drosophila melanogaster* and its sibling species. Mol. Biol. Evol. **4**:638-650.

Satta, Y., and N. Takahata. 1990. Evolution of Drosophila mitochondrial DNA and the history of the *melanogaster* subgroup. Proc. Natl. Acad. Sci. USA **87:**9558-9562.

Schinkel, A. H., M. J. A. Groot Koerkamp, and H. F. Tabak. 1988. Mitochondrial RNA polymerase of *Saccharomyces cerevisae*: composition and mechanism of promoter recognition. EMBO J. **7:**3255-3262.

Schinkel, A. H., and H. F. Tabak. 1989. Mitochondrial RNA polymerase: dual role in transcription and replication. Trends. Genet. **5:**149-154.

Schmitt, M. E., and D. A. Clayton. 1992. Yeast site-specific ribonucleoprotein endoribonuclease MRP contains an RNA component homologous to mammalian RNase MRP RNA and essential for cell viability. Genes Dev. **6:**1975-1985.

Schmitt, M. E., and D. A. Clayton. 1993. Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisae*. Mol. Cell. Biol. **13:**7935-7941.

Schulman, L. H., and H. Pelka. 1988. Anticodon switching changes the identity of methionine and valine tRNAs. Science **242:**765-768.

Schwartz, R. M., and M. O. Dayhoff. 1978. Origins of prokaryotes, eukaryotes, mitochondria and chloroplasts. Science **199**:395-403.

Seeman, N. C., J. M. Rosenberg, and A. Rich. 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. USA. **73**:804-808.

Shadel, G. S., and D. A. Clayton. 1993. Mitochondrial transcription initiation: variation and conservation. J. Biol. Chem. **268**:16083-16086.

Shay, J. W., D. J. Pierce, and H. Werbin. 1990. Mitochondrial DNA copy number is proportional to total cell DNA under a variety of growth conditions. J. Biol. Chem. **265**:14802-14807.

Shine, J., and L. Delgarno. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342-1346.

Siegel, L. M., and K. J. Monty. 1966. Determination of molecular weights and frictional ratios of proteins in impure systems by use of gel filtration and density gradient centrifugation. Application to crude preparations of sulfite and hydroxylamine reductases. Biochim. et Biophys. Acta. **112**:346-362.

Sinclair, A. H., P. Berta, M. S. Palmer, J. R. Hawkins, B. L. Griffiths, M. J. Smith, J. W. Foster, A.-M. Frischauf, R. Lovell-Badge, and P. N. Goodfellow. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature **346**:240-244.

Snyder, M., A. R. Fraser, J. La Roche, K. E. Gartner-Kepkay, and E. Zouros. 1987. Atypical mitochondrial DNA from the deep-sea scallop *Placopecten magellanicus*. Proc. Natl. Acad. Sci. USA **84**:7595-7599.

Solignac, M., M. Monnerot, and J.-C. Mounolou. 1983. Mitochondrial DNA heteroplasmy in *Drosophila mauritiana*. Proc. Natl. Acad. Sci. USA **80**:6942-6946.

Solignac, M., Generemont, J., M. Monnerot, and J. C. Mounolou. 1984. Genetics of mitochondria in *Drosophila*; inheritance in heteroplasmic strains of *D. Mauritiana*. Mol. Gen. Genet. **197**:183-188.

Solignac, M., M. Monnerot, and J.-C. Mounolou. 1986. Concerted evolution of sequence repeats in *Drosophila* mitochondrial DNA. J. Mol. Evol. **24**:53-60.

Solomon, M. J., F. Strauss, and A. Varshavsky. 1986. A mammalian high mobility group protein recognizes any stretch of six A·T base pairs in duplex DNA. Proc. Natl. Acad. Sci. USA **83**:1276-1280.

Steinberg, S., A. Misch and M. Sprinzl. 1993. Compilation of tRNA sequences and sequences of tRNA genes. Nucleic Acids Res. **21**:3011-3015.

Steitz, J. A., and K. Jakes. 1975. How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli.* Proc. Natl. Acad. Sci. USA **72:**4734-4738.

Stohl, L. L., and D. A. Clayton. 1992. *Saccharomyces cerevisae* contains an RNase MRP that cleaves at a conserved mitochondrial RNA sequence implicated in replication priming. Mol. Cell. Biol. **12:**2561-2569.

Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye. 1966. Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. **31**:77-84.

Suyama, Y., H. Fukuhara, F. Sor. 1985. A fine resolution map of the linear mitochondrial DNA of *Tetrahymena pyriformis*: genome size, map locations of rRNA and tRNA genes, terminal inversion repeat, and restriction site polymorphism. Curr. Genet. **9:**479-493.

Tamura, K. 1992. The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA. Mol. Biol. Evol. **9:**814-825.

Tapper, D. P., and D. A. Clayton. 1981. Mechanism of replication of human mitochondrial DNA: localization of the 5' ends of nascent daughter strands. J. Biol. Chem. **256:**5109-5115.

Topal, M. D., and J. R. Fresco. 1976. Complementary base pairing and the origin of substitution mutations. Nature **263:**285-289.

Topper, J. N., and D. A. Clayton. 1989. Identification of transcriptional regulatory elements in human mitochondrial DNA by linker substitution analysis. Mol. Cell. Biol. **9:**1200-1211.

Topper, J. N., and D. A. Clayton. 1990. Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endonuclease when assembled as an RNP. Nucleic Acids. Res. **18:**793-799.

Travis, A., A. Amsterdam, C. Belanger, and R. Grosschedl. 1991. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α-enhancer function. Genes Devel. **5:**880-894.

Tzagoloff, A. 1982. The Mitochondrion. Plenum Press, New York, 342 pp.

Tzagoloff, A., and A. M. Myers. 1986. Genetics of mitochondrial biogenesis. Annu. Rev. Biochem. **55:**249-285.

Tzeng, C.-S., C.-F. Hui, S.-C. Shen and P. C. Huang. 1992. The complete nucleotide sequence of the *Crossostoma lucustre* mitochondrial genome: conservation and variations among vertebrates. Nucleic Acids Res. **20:**4853-4858.

van de Wertering, M., M. Oosterwegel, D. Dooijes, and H. Clevers. 1991. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-HMG box. EMBO J. **10:**123-132.

Van Tuyle, G. C., and M. L. McPherson. 1979. A compact form of rat liver mitochondrial DNA stabilized by bound proteins. J. Biol. Chem. **254:**6044-6053.

Van Tuyle, G. C., and P. A. Pavco. 1981. Characterization of a rat liver mitochondrial DNA-protein complex: replicative intermediates are protected against branch migrational loss. J. Biol. Chem. **256:**12772-12779.

Van Tuyle, G. C., and P. A. Pavco. 1985a. The rat liver mitochondrial DNA-protein complex: displaced single strands of replicative intermediates are protein coated. J. Cell Biol. **100:**251-257.

Van Tuyle, G. C., and P. A. Pavco. 1985b. Purification and general properties of the DNA-binding protein P16 from rat liver mitochondria. J. Cell Biol. **100:**258-264.

Vawter, L., and W. M. Brown. 1986. Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science **234:**194-196.

Volz-Lingenhohl A., M. Solignac, and D. Sperlich. 1992. Stable heteroplasmy for a large-scale deletion in the coding region of *Drosophila subobscura* mitochondrial DNA. Proc. Natl. Acad. Sci. USA **89:**11528-11532.

Walberg, M. W., and D. A. Clayton. 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res. **9:**5411-5421.

Walberg, M. W., and D. A. Clayton. 1983. In vitro transcription of human mitochondrial DNA. Identification of specific light-strand transcripts from the displacement loop region. J. Biol. Chem. **258:**1268-1275.

Wallace, D. C. 1982. Structure and evolution of organelle genomes. Microbiol. Rev. **46:**208-240.

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

Wallin, J. E. 1922. On the nature of mitochondria. Am. J. Anat. **30:**203-229, 451-471.

Ward, B. L., R. S. Anderson, and A. J. Bendich. 1981. The mitochondrial genome is large and variable in a family of plants (*Cucuribitaceae*). Cell **25:**793-803.

Ward, R. H., B. L. Frazier, K. Dew-Jager, and S. Paabo. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. Proc. Natl. Acad. Sci. USA **88**:8720-8724.

Warrior, R., and J. Gall. 1985. The mitochondrial DNA of *Hydra attenuata* and *Hydra littoralis* consists of two linear molecules. Arch. Sc. Geneve **38**:439-445.

Webb, A. C., and L. D. Smith. 1977. Accumulation of mitochondrial DNA during oogenesis in *Xenopus laevis*. Dev. Biol. **56**:219-225.

Welter, C., S. Dooley, K. D. Zang, and N. Blin. 1989. DNA curvature in front of the human mitochondrial L-strand replication origin with specific protein binding. Nucl. Acids Res. **17**:6077-6086.

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

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

Wesolowski, M., and H. Fukuhara. 1981. Linear mitochondrial deoxyribonucleic acid from the yeast *Hansenula mrakii*. Mol. Cell. Biol. **1**:387-393.

Wettstein, J., B. S. Ticho, N. C. Martin, D. Najarian, and G. S. Getz. 1986. In vitro transcriptional and promoter strength analysis of five mitochondrial tRNA promoters in yeast. J. Biol. Chem. **261**:2905-2911.

Whittaker, P. A., and S. M. Danks. 1978. Mitochondria: structure, function, and assembly. Longman, New York, 148 pp.

Wilkinson, G. S., and A. M. Chapman. 1991. Length and sequence variation in evening bat D-loop mtDNA. Genetics **128**:607-617.

Williams, R. S. 1986. Mitochondrial gene expression in mammalian striated muscle: evidence that variation in gene dosage is the major regulatory event. J. Biol. Chem. **261**:12390-12394.

Williams, R. S., S. Salmons, E. A. Newsholme, R. E. Kaufman. and J. Mellor. 1986. Regulation of nuclear and mitochondrial gene expression by contractile activity in skeletal muscle. J. Biol. Chem. **261**:376-380.

Williams, A. J., Wernette, C. M., and L. S. Kaguni. 1993. Processivity of mitochondrial DNA polymerase from *Drosophila* embryos. J. Biol. Chem. **268**:24855-24862.

Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. **51**:221-271.

Wolstenholme, D. R., J. M. Goddard, and C. M.-R. Fauron. 1979. Structure and replication of mitochondrial DNA from the genus *Drosophila*. Pp. 409-425 *in* D. J. Cummings, P. Borst, I. B. Dawid, S. M. Weissman, and C. F. Fox, eds. Extrachromosomal DNA. Academic Press Inc. New York, NY.

Wolstenholme, D. R., J. M. Goddard, and C. M.-R. Fauron. 1983. Replication of *Drosophila* mitochondrial DNA. Pp. 131-148 *in* Y. Becker ed. Replication of viral and cellular genomes. Martinus Nijhoff Publishing, Boston, MA.

Wolstenholme, D. R., and D. O. Clary. 1985. Sequence evolution of *Drosophila* mitochondrial DNA. Genetics **109**:725-744.

Wong, T. W., D. A. Clayton. 1985. In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. Cell **42**:951-958.

Xu, C., and D. S. Ray. 1993. Isolation of proteins associated with kinetoplast DNA networks in vivo. Proc. Natl. Acad. Sci. USA **90**:1786-1789.

Yamaguchi, M., A. Matsukage, and T. Takahashi. 1980. Chick embryo DNA polymerase γ: purification and structural analysis of nearly homogeneous enzyme. J. Biol. Chem. **255**:7002-7009.

Yoza, B. K., and D. F. Bogenhagen. 1984. Identification and in vitro capping of a primary transcript of human mitochondrial DNA. J. Biol. Chem. **259**:3909-3915.