




1993

## Analysis of the Dna Repair/Ribosomal Protein Ap3/Po Gene in *Drosophila* (Homologue to Human Po)

David T. Grabowski  
*Loyola University Chicago*

Follow this and additional works at: [https://ecommons.luc.edu/luc\\_diss](https://ecommons.luc.edu/luc_diss)

 Part of the [Molecular Biology Commons](#)

---

### Recommended Citation

Grabowski, David T., "Analysis of the Dna Repair/Ribosomal Protein Ap3/Po Gene in *Drosophila* (Homologue to Human Po)" (1993). *Dissertations*. 3110.  
[https://ecommons.luc.edu/luc\\_diss/3110](https://ecommons.luc.edu/luc_diss/3110)

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License](#).  
Copyright © 1993 David T. Grabowski

LIBRARY-LOYOLA UNIVERSITY  
MEDICAL CENTER

ANALYSIS OF THE DNA REPAIR/RIBOSOMAL PROTEIN  
AP3/PO GENE IN DROSOPHILA (HOMOLOGUE TO HUMAN PO)

by

David T. Grabowski

A Dissertation submitted to  
the Faculty of the Graduate School  
of Loyola University of Chicago in Partial Fulfillment  
of the Requirements for the Degree of  
Doctor of Philosophy

January

1993

Copyright by David T. Grabowski, 1992

All rights reserved.

## ACKNOWLEDGEMENTS

I would like to express my gratitude to my director, Dr. Mark R. Kelley for his guidance and his excellent example as a research scientist. My thanks are also extended to members of my committee, Drs. Allen Frankfater, Bill Simmons, Leonard Erickson, and Deborah Hoshizaki for their advice, technical help and constructive criticism of these studies.

I wish to express my deepest appreciation to my fellow graduate students in the lab: David Wilson, Jim Carney, John Tentler, and Peg Halloran for their friendship, support and willingness to help whenever possible.

I would like to gratefully acknowledge Dr. John Lopes for his assistance in preparing some of my figures, and Dr. Sally Amero for much of the photography presented in this dissertation. I additionally extend much thanks to Dr. Dennis Derda for his support and contributions to the project.

Special thanks are extended to Janet Flores and Bridget Borja for their excellent secretarial skills and assistance throughout my stay here at Loyola.

Finally, I would like to express my most sincere gratitude to my family and to my wife, who have given me their constant support and encouragement throughout my educational endeavors.

I would like to dedicate my work and this dissertation to my parents, Ted and Barbara, my wife Peggy, my brother Steve and his family Patty, Stephanie and Kristen.

# TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
TABLE OF CONTENTS.....	iv
LIST OF FIGURES.....	vii
LIST OF ABBREVIATIONS.....	ix
INTRODUCTION.....	1
CHAPTER	
I. REVIEW OF CITED LITERATURE.....	4
Sources of DNA Damage.....	4
Types of Repair.....	7
Nucleotide-Excision Repair.....	10
Base-Excision Repair.....	13
AP Damage.....	19
AP Endonucleases.....	20
Ribosomal Proteins and DNA Repair.....	33
DNA Repair and Autoimmune Deficiencies.....	35
II. MATERIALS AND METHODS.....	37
Screening a Genomic Library by Plaque Colony Hybridization.....	37
Screening a cDNA Library for a Human Homologue to AP3.....	38
Purification of $\lambda$ Phage DNA.....	39
Generation of Single-Stranded DNA for Sequencing	40
DNA Sequencing.....	41
Labelling Double-Stranded DNA with Random Hexamers.....	41

Isolation of <i>Drosophila</i> DNA.....	42
Southern Blot Analysis.....	43
Total <i>Drosophila</i> RNA Isolation.....	43
Northern Blot Analysis.....	44
Construction of $\beta$ -galactosidase-AP3	
Overexpression Vectors.....	44
Thermal Cycle Amplification of AP3 cDNA.....	46
Overexpression of Fusion and Non-fusion Proteins	
in <i>E.coli</i> .....	47
Purification of Proteins from Inclusion Bodies..	48
Purification of Antigens for Antibody Production	49
Production of Antibodies.....	50
Purification of Antibodies.....	51
Subcellular Fractionation of Adult <i>Drosophila</i>	
Cells.....	52
Nuclei Purification.....	52
Chromatin Purification.....	53
Subnuclear Fractionation.....	54
Western Blot Analysis.....	55
5'-End Labelling of DNA Oligonucleotides.....	56
5'-End Mapping of AP3 Transcript.....	56
Tissue Culture of <i>Drosophila</i> Schneider II Cells.	58
III. RESULTS.....	59
Mapping the other Transcript Near AP3.....	68
Cloning and Characterization of <i>Drosopa</i> .....	74
Cloning of P0 and Comparison to AP3.....	82

Role of P0 as a DNA Repair Protein.....	88
Production of Anti-AP3 Antibodies.....	98
Overexpression of AP3 in <i>E.coli</i> .....	102
Purification of AP3.....	106
Subcellular Localization of AP3.....	109
AP Endonuclease Activity of AP3.....	115
IV. DISCUSSION.....	120
Cloning and Characterization of <i>Drosopa</i> .....	121
The 1.1kb Transcript.....	124
Genomic Map of AP3.....	127
Cloning of P0 and Comparison to AP3.....	129
P0 as a DNA-Repair Protein.....	132
Localization of AP3.....	136
AP3 is a Class II AP Endonuclease.....	139
LITERATURE CITED.....	143
VITA.....	161



## LIST OF FIGURES

Figure	Page
1. Diagrammatic representation of steps involved in excision repair.....	14
2. The four theoretical cleavage sites for AP endonucleases.....	22
3. List of some AP endonucleases from various sources.....	26
4. Schematic diagram of the genomic map of the <i>AP3</i> gene.....	61
5. Nucleotide and amino acid sequence of the <i>AP3</i> cDNA.....	64
6. Plot of open reading frames found in the 2.2kb <i>Bam</i> HI genomic fragment.....	66
7. Nucleotide and amino acid sequences of the postulated promoter region of <i>AP3</i> .....	69
8. Developmental northern blot of total RNA from various <i>Drosophila</i> stages.....	72
9. Nucleotide and predicted amino acid sequence of the <i>Dromsopa</i> cDNA.....	75
10. Northern blot analysis of <i>Drosophila</i> developmental, tissue, and sex specific RNA.....	78
11. Southern blot of <i>Drosophila</i> genomic DNA digested with restriction endonucleases.....	81
12. Comparison of the deduced amino acid sequences of human P0, <i>Drosophila</i> AP3, and yeast A0 proteins..	84
13. Comparison of a family of acidic proteins at the amino acid level.....	86
14. Northern blot analysis of the effects of L-PAM 4HC, and HN2 on P0 gene expression.....	90

15. Densitometric analysis of the northern blots shown in figure 14.....	92
16. Northern blot analysis of P0 expression in Mer <sup>+</sup> and Mer <sup>-</sup> human tumor cell lines.....	95
17. MGMT activity and densitometric analysis of the northern blot shown in figure 16.....	97
18. Schematic diagram of regions of AP3 subcloned into overexpression vectors.....	100
19. Developmental Western blot analysis of purified nuclei and post nuclear supernatants of adult Drosophila using anti-d1 and anti-d8 antibodies..	103
20. Coomassie blue stained gel with aliquots from the inclusion body purification of pET11d-AP3.....	107
21. Developmental Western blot of purified nuclei and post nuclear supernatants of Drosophila using purified anti-AP3 antibody.....	111
22. Western blot analysis of subcellular and subnuclear components of adult Drosophila.....	114
23. Schematic diagram of Class I and Class II AP endonuclease assays.....	117

## LIST OF ABBREVIATIONS

$\mu\text{Ci}$	microcurie
$\mu\text{g}$	microgram
$\mu\text{l}$	microliter
$\mu\text{M}$	micromolar
4-HC	4-hydroperoxy cyclophosphamid
AP	apurinic/apyrimidinic
AP3	Drosophila AP endonuclease/ribosomal protein
AT	ataxia telangiectasia
BCIP	5-bromo-4-chloro-3-indolyl phosphate
bp	base pair
D	daltons
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
DTT	dithiolthreitol
EDTA	ethylenediamine tetraacetic acid
HN2	mechlorethamine
kb	kilobases
kD	kilodaltons
L-PAM	L-phenylalanine mustard
M	molar
mg	milligram
MGMT	O <sup>6</sup> -methylguanine DNA methyltransferase

ml	milliliter
mM	millimolar
mRNA	messenger RNA
MMS	methyl methanesulfonate
MNNG	N-methyl-N'-nitro-N-nitrosoguanidine
MNU	N-methyl-N-nitrosourea
NBT	nitro blue tetrazolium chloride
O.D.	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
RNA	ribonucleic acid
Rrp1	recombination repair protein 1
SAM	S-adenosyl-L-methionine
TBE	tris-boric acid-EDTA electrophoresis buffer
UV	ultraviolet
Xg	times gravity

## INTRODUCTION

The presence of DNA damaging agents, such as carcinogens and mutagens, and the spontaneous loss of bases from DNA, has made it imperative that organisms develop mechanisms to repair damaged DNA. DNA repair is a complex biochemical process by which cells attempt to preserve genetic information from the damaging effects of exposure to environmental insults, such as ultraviolet (UV) and ionizing radiation and chemicals, as well as from spontaneous errors in recombination and replication. One common type of DNA damage, baseless or apurinic/apyrimidinic (AP) sites, is recognized by AP endonucleases, which initiate the process of nucleotide excision repair. The goal of this dissertation project was to characterize a *Drosophila* AP endonuclease, AP3, at the molecular and biochemical level. To accomplish this, a detailed genomic map of AP3 was constructed and the product of AP3 was overexpressed and purified to near homogeneity. In a related portion of the project, the human homologue to AP3 (PO) was also cloned and analyzed.

A cDNA encoding an AP endonuclease in *Drosophila melanogaster*, AP3, was previously cloned by screening a *Drosophila*  $\lambda$ gt11 expression library with an antibody against a purified human AP endonuclease. The AP3 cDNA was used to isolate genomic clones encompassing the region containing AP3.

PO, was isolated. This gene was also cloned in the laboratory of Dr. J. Steitz by screening a human  $\lambda$ gt11 expression library with anti-sera from patients with systemic lupus erythematosus (SLE).

## CHAPTER I

### REVIEW OF CITED LITERATURE

All living organisms are continuously exposed to endogenous as well as exogenous damaging agents capable of altering DNA. DNA is inherently a highly unstable molecule targeted by various physical and chemical agents and subject to spontaneous damage. Being the carrier of the genetic information, damage to DNA can be deleterious to the survival of an organism. The failure to repair lesions or errors can lead to inherited mutations (germ line mutations) or cell death (somatic tissue). To cope with structural DNA alterations, all organisms have developed a variety of DNA repair systems. This review will attempt to define some of the DNA damages that can occur, describe some related repair pathways and, finally, focus on one particular form of lesion (baseless sites) and the enzymes involved in the repair of that lesion.

#### Sources of DNA Damage

There exist several endogenous sources of DNA-damaging agents. The most common natural damage is the heat induced deamination of cytosine to uracil and the deamination of adenine and guanine to hypoxanthine and xanthine, respectively (Lindahl and Nyberg, 1974; Karran and Lindahl, 1980). The

heat-induced deamination of adenine to hypoxanthine would result in A:T to G:C transition mutations if the hypoxanthine is not removed prior to DNA replication. Similarly, the deamination of guanine to xanthine is only potentially mutagenic in that xanthine will still base pair with cytosine. However, a tautomeric shift in the structure of xanthine would allow for a mismatch pairing with thymine. This situation would result in a G:C to A:T transition as described above. The biological consequences of cytosine deamination can be direct mispairing, which will also result in G:C to A:T mutagenic transitions. The presence of uracil in DNA is not necessarily mutagenic. The misincorporation of dUTP instead of dTTP into DNA is not mutagenic as uracil will also pair with adenine. Uracil, hypoxanthine and xanthine can be removed from DNA prior to replication by DNA repair enzymes collectively called glycosylases. Glycosylases, which cleave the N-glycosylic bond releasing the altered base resulting in an apurinic or apyrimidinic (AP) site, will be discussed later in this review. The cleavage of the N-glycosylic bond between bases and the deoxyribose of DNA can also occur spontaneously at physiological temperatures, giving rise to AP sites. It has been estimated that at 37°C approximately 10,000 depurinations can occur per cell per day (Lindahl and Nyberg, 1972). These depurinations, can lead to single strand breaks in the DNA molecule (Lindahl and Andersson, 1972).

Normal oxidative metabolism is a source of DNA damaging



agents such as the superoxide radical ( $O_2^{\cdot-}$ ), the hydroxyl radical ( $OH\cdot$ ), and hydrogen peroxide ( $H_2O_2$ ) (Fridovich, 1978; Ames, 1983) and other free radical byproducts (Fridovich, 1978; Freeman and Crapo, 1982). Loeb (1989) has estimated that each human cell sustains an average of 10,000 "oxidative hits" per day. The various types of radicals formed produces a broad spectrum of DNA damage including single strand breaks, damages to the deoxyribose moiety, and damages to purine and pyrimidine bases (Hagen, 1986; Hutchinson, 1985; von Sonntag and Schuchman, 1986; Teoule, 1987).

Yet another endogenous cause of damage is the alkylation of DNA by the normal intracellular methyl group donor, S-adenosyl-L-methionine (SAM) (Rydberg and Lindahl, 1982). The resulting products of alkylation can be  $N^7$ -methylguanine,  $N^3$ -methyladenine,  $N^3$ -methylthymine and  $O^6$ -methylguanine. Methylation causes a further destabilization of the N-glycosylic bond, shown previously to be susceptible to cleavage forming baseless sites (Lindahl and Nyberg, 1972). Methylation products can also arise as a result of exogenous insults from various methylating agents. Monofunctional methylating agents, such as N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) and N-methyl-N-nitrosourea (MNU) are mutagenic and carcinogenic in mammalian cells (Lindahl, and Sedgwick 1988). Methylating agents covalently modify the exposed oxygens and nitrogens of the DNA helix resulting in base derivatives and replication blocking lesions.

Methylating agents produce O<sup>6</sup>-methylguanine, a major mutagenic lesion which results in G·C to A·T transitions (Snow *et al.*, 1984; Loechler *et al.*, 1984). 3-methyladenine is also an important product of methylating agents as it leads to cell-killing by blocking DNA replication (Lindahl and Sedgwick, 1988). O<sup>6</sup>-methylguanine and 3-methyladenine are corrected by separate DNA repair pathways discussed later in this review.

The normal diet of an organism can contain mutagenic and carcinogenic agents such as alcohol, nitrosamines, and phorbol esters (Ames, 1983). Agents found in the diet are able to react with DNA, generating single-strand breaks, double-strand breaks and bulky adducts. The ultraviolet (UV) component of sunlight has a number of DNA damaging effects: the formation of pyrimidine dimers and 6-4 photoproducts (Setlow, 1982); the production of cytosine hydrate and thymine glycols; single strand breaks, protein-DNA crosslinks, and intrastrand crosslinks (Friedberg, 1985). Ionizing radiation can perturb the ring structures of the bases resulting in ring opening, base fragmentation and strand breaks (Sancar and Sancar, 1988).

### Types of Repair

DNA repair is a complex biochemical process by which cells attempt to preserve genetic information from the damaging effects of spontaneous chemical transformations and harmful endogenous and exogenous agents, (discussed above), as well as from spontaneous errors in recombination and

replication. Generally, the repair systems can be divided into three categories i.e., reversal of damage, post-replication repair and excision repair. The first two repair systems will be briefly discussed followed by an extensive look at excision repair.

Direct repair does not include any structural changes to the DNA except for the removal of the lesion itself. The covalent modification of pyrimidines by UV light can be simply reversed by enzymatic photoreactivation. The enzyme actually uses energy from light (300-500nm) to break the covalent bond joining the cyclobutane pyrimidine dimers (Rupert et al., 1958; Eker, 1983). Another form of direct repair is the direct removal of methyl groups by a methyltransferase (Lindahl, 1982). O<sup>6</sup>-Methylguanine-DNA methyltransferase (MGMT) directly transfers the methyl group from the O<sup>6</sup> position of guanine to a cysteine residue in the MGMT molecule (Olsson and Lindahl, 1980), to regenerate an unmodified guanine (Foote et al., 1980). The O<sup>6</sup>-methylguanine-DNA methyltransferase can also recognize higher forms of O<sup>6</sup>-alkylguanine (O<sup>6</sup>-ethylguanine, O<sup>6</sup>-hydroxyguanine) but the repair rate is much slower (Lindahl and Sedgwick, 1988). A third form of direct repair involves the action of a purine base insertase which catalyses the direct re-insertion of bases into depurinated DNA. This activity was first identified in cultured human fibroblasts by Deutsch and Linn (1979a). Deutsch and Spiering (1985) later isolated a related

purine-base-insertase from *Drosophila*. Both the human and *Drosophila* insertases were found to be extremely labile, disallowing further characterization of this unique activity.

Direct repair processes are extremely important although only a few types of DNA lesions are corrected via this pathway. These lesions include UV-induced photoproducts, some alkylated bases (mainly alkylated guanine) and some apurinic sites. These lesions, along with the plethora of other DNA lesions, are also recognized by other forms of DNA repair.

In post-replication repair, as the name implies, the DNA lesion is bypassed by DNA polymerase, and is repaired later after replication has been allowed to continue. Two examples of post-replication repair have been well characterized in *E.coli*. In the first example, error-prone mutagenic replication, (SOS response) the 3' to 5' proof-reading function of the DNA polymerase is altered by RecA and UmuC proteins. These proteins, when bound to the replication complex, decrease the fidelity of the polymerase allowing it to place any nucleotide opposite the lesion as well as continue to misincorporate nucleotides during continued replication (Bridges and Woodgate, 1985). The errors introduced by mutagenic replication as well as the original DNA lesion are eventually corrected by the repair systems induced during the increased SOS response of the cell (Sancar and Sancar, 1988). One of these systems, excision repair, will be discussed later. The other type of repair is also the

second example of post-replication repair namely recombinational repair.

During recombinational repair, when the DNA polymerase encounters a pyrimidine dimer, it stops replicating and reinitiates replication about 1000 nucleotides downstream. The postreplication gap is filled in by the RecA protein, which transfers the complementary strand from the sister duplex into the gap (Rupp and Howard-Flanders, 1968; Rupp et al., 1971).

The third, general repair system, DNA excision repair, involves a multi-enzyme pathway described by these steps: 1) recognition of the damage; 2) nicking of the damage containing strand; 3) removal of the damaged region; 4) filling in of the resulting gap; and 5) ligation of the gap (Haseltine, 1983). The excision repair system can be further broken down into two similar, yet distinct pathways; 1) nucleotide-excision repair, and 2) base-excision repair.

#### Nucleotide-Excision Repair

One of the most important and best studied repair process in prokaryotes and eukaryotes is the nucleotide-excision repair pathway. A unique feature of this type of repair is the removal of damaged bases from DNA as part of an oligonucleotide (Setlow and Carrier, 1964; Boyce and Howard-Flanders, 1964). Although deficiencies in this type of repair have been isolated in mutant cells from *E.coli*, yeast and humans, the best characterized pathway is that found in

*E.coli*. Briefly, this pathway involves 5 steps: damage recognition, incision, excision, repair synthesis, and ligation. In *E.coli*, these basic steps are carried out by the UvrABC nuclease complex.

A remarkable feature of the UvrABC complex is its broad substrate diversity, being capable of acting on a plethora of DNA damage. The UvrABC complex can recognize pyrimidine dimers, 6-4 photoproducts, as well as base adducts of cisplatin and mitomycin C, and other modifications such as thymine glycols and apurinic sites (van Houten, 1990). It has been proposed that the complex recognizes damage-induced conformational changes of DNA and not the specific chemical modifications (Weiss and Grossman, 1987; Sancar and Rupp, 1983).

The mechanism of the UvrABC complex has been well characterized due to the fact that the genes for the subunits *uvrA*, *uvrB* and *uvrC*, have been cloned (Auerbach and Howard-Flanders, 1979), Pannekoek *et al.*, 1978), overexpressed, purified and used in reconstitution systems (Sancar and Rupp, 1983). Numerous experiments utilizing the cloned genes and purified products have led to a general model of nucleotide-excision repair. Prior to the recognition of damage, a primary complex is formed in the presence of ATP containing only subunits A and B, the Uvr A<sub>2</sub>B complex. The A<sub>2</sub>B complex binds to DNA around the damaged site leading to the dissociation of the UvrA dimer. The subsequent, stable Uvr B-

DNA complex is joined by UvrC. Once UvrC has bound, two phosphodiester bonds are hydrolyzed, one 7 nucleotides 5' and the other 3 or 4 nucleotides 3' of the modified nucleotide (Sancar and Sancar, 1988). The actual removal of the complex along with the damage containing oligonucleotide requires the joint action of UvrD (helicase II) and PolI which dislodges the complex with concomitant filling in of the gap.

Although *E.coli* nucleotide-excision repair has been extensively characterized on both the molecular and biochemical levels, much less information is available regarding nucleotide-excision repair in eukaryotes. DNA repair systems in higher eukaryotes, especially mammalian cells, have been most difficult to elucidate, resulting in a limited understanding at this level. A number of naturally occurring human diseases including xeroderma pigmentosum (XP), ataxia telangiectasis (AT) and Bloom's syndrome, have been associated with defects in DNA repair. Cells from patients with XP are defective in performing a variety of functions of nucleotide-excision repair (Friedberg, 1988). Cell fusion studies have identified seven XP complementation groups (Keijzer, et al., 1979; Lehmann et al., 1992), indicating that at least seven genes may be involved in damage recognition and repair. Genetic studies suggest that similar complexes to the UvrABC nuclease complex also exist in eukaryotic cells (Friedberg, 1985). In recent years significant progress has been made in cloning both yeast and human DNA repair genes

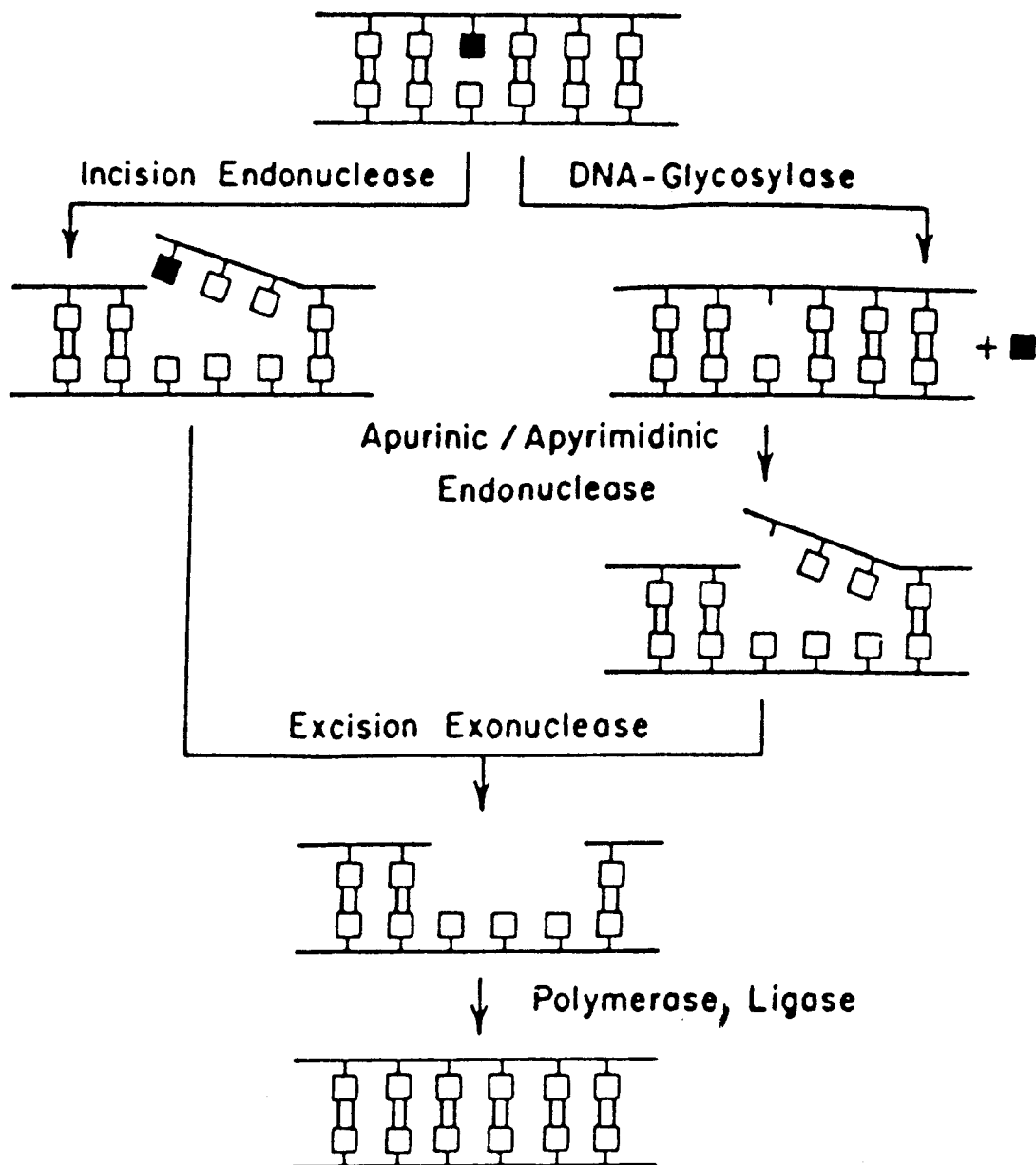
having significant homology with the *E.coli* Uvr genes. For instance, the yeast RAD3 protein has strong homology with both UvrA and UvrD proteins (Reynolds *et al.*, 1985). Recently, there have been several human repair genes cloned through the use of phenotypic complementation, namely the ERCC (excision repair cross complementing) genes (ERCC-1, -2, -3, -5 and -6) (Westerveld *et al.*, 1984; Weber *et al.*, 1988; Weeda *et al.*, 1990; Mudgett and MacInnes, 1990 and Troelstra, 1990). Analysis of the ERCC1 and ERCC2 genes has revealed a striking evolutionary sequence conservation at the protein level with yeast RAD10 (van Duin *et al.*, 1986) and RAD3 (Weber *et al.*, 1990), respectively. Presently, no biochemical functions can be assigned to any of the human proteins.

#### Base-Excision Repair

Base-excision repair differs from nucleotide-excision repair in that the damaged or modified base is removed by the concerted action of several sequentially acting enzymes not necessarily associated as a complex. The general steps of this pathway are the recognition of the modified base, release of the modified base by a glycosylase, cleavage of the phosphodiester backbone of DNA by an AP endonuclease, resynthesis of the gap by DNA polymerase and sealing of the nick by DNA ligase. An illustration of the base-excision repair pathway is shown in Figure 1. Base-excision repair, like nucleotide-excision repair is capable of recognizing a variety of damages, including modified bases resulting from



Figure 1. Diagrammatic representation of steps involved in Excision repair. The black box represents a damaged base. (From Friedberg, 1985).



exposure to alkylating agents ( $N^7$ -methylguanine,  $N^3$ -methyladenine,  $O^6$ -methylguanine,  $O^6$ -ethylguanine,  $O^6$ -hydroxyguanine), exposure to ionizing radiation (hydroxymethyluracil, formamidopyrimidines) and modified bases resulting from the spontaneous deaminations of cytosine (to uracil), adenine (to hypoxanthine) and guanine (to xanthine). The biological consequences of these lesions were discussed earlier in Sources of DNA damage. The formamidopyrimidines (ring open form of  $N^7$ -methylguanine) have been shown to block DNA polymerase I in *E.coli* (Boiteux and Laval, 1983) suggesting the potential lethality of this lesion.

A general group of repair nucleases, known as N-glycosylases, recognize a particular lesion and releases the modified base by hydrolyzing the N-glycosylic bond generating a hallmark feature of the base-excision repair pathway, an apurinic or apyrimidinic (AP) site. AP sites are then subject to attack by AP endonucleases which cleave the phosphodiester backbone of the DNA strand containing the baseless site. A detailed discussion of AP sites and AP endonucleases will follow a more detailed description of N-glycosylases.

The first of the glycosylases identified was uracil-DNA glycosylase from *E.coli* (Lindahl, 1974) and *B.subtilis* (Friedberg *et al.*, 1975). These enzymes have been extensively purified and characterized. Bacterial uracil DNA glycosylases all have similar molecular weights (20,000 - 30,000) and have no requirement for cofactors. Uracil-DNA glycosylase-

deficient mutants in *E.coli* (Duncan *et al.*, 1978) were shown to contain point mutations, insertions and deletions. Because these mutants are unable to remove uracil from DNA, they exhibit an increased frequency of G:C to A:T mutagenic transitions (Duncan and Weiss, 1982; Fix and Glickman, 1986, 1987). Eukaryotic uracil-DNA glycosylases have been purified and/or identified from a number of sources (Wallace, 1988). Yeast uracil-DNA glycosylase (Crosby *et al.*, 1981) shares the general properties of the bacterial uracil-DNA glycosylases. A yeast mutant was isolated and shown to have increased sensitivity to sodium bisulfite (Burgers and Klein, 1986) which causes a deamination of cytosine to uracil *in vitro* (Hayatsu, 1976). The human uracil-DNA glycosylase has also been purified and characterized (Caradonna and Cheng, 1980) and has been reported to be associated with DNA polymerase  $\alpha$  (Seal and Sirover, 1986). This finding leads to the speculation that base-excision repair may be carried out by a multiprotein complex associated with the replication process.

Specific glycosylases for the other nucleotide lesions have also been isolated. Hypoxanthine DNA glycosylases have been purified from *E.coli*, calf thymus (Lindahl, 1982; Harosh and Sperling, 1988) and HeLa cells (Myrnes *et al.*, 1982). 3-methyladenine DNA glycosylases which recognize some methylated bases (3-methyladenine, 3-methylguanine, O<sup>2</sup>-methylthymine, O<sup>2</sup>-methylcytosine, and 7-methylguanine) have been isolated in *E.coli* (Sakumi *et al.*, 1986, Bjelland and Seeberg, 1987).

Hydroxymethyluracil DNA glycosylase recognizes the product of hydroxyl radical attack on thymine. The mammalian form of this glycosylase has been purified and partially characterized from mouse plasmacytoma cells (Hollstein *et al.*, 1984). The short list of glycosylases described above is by no means exhaustive. Many glycosylases have been isolated, purified and cloned from a variety of organisms. As mentioned earlier, a main characteristic of glycosylases is the end product of their activity, an AP site. However, several glycosylases possess a concomitant AP endonuclease activity which are capable of carrying out the first three steps of base excision repair. One of the best known and best characterized DNA glycosylase-AP endonucleases is the pyrimidine dimer DNA glycosylase. Originally, this enzyme was identified in *M.luteus* cells and T4 phage-infected *E.coli* (Lindahl, 1982). More recently, a glycosylase-AP endonuclease was found in yeast (Hamilton *et al.*, 1992) possessing the same characteristics as the *M. luteus* and T4 phage proteins. These glycosylase-AP endonucleases recognize pyrimidine dimers generated by UV light.

Several glycosylases which also have AP endonuclease activity have also been identified in mammals. These glycosylase-endonucleases have substrate specificities and repair activities similar to *E.coli* endonuclease III (Doetsch *et al.*, 1986; Doetsch *et al.*, 1987). The substrates of endonuclease III include a variety of thymine and cytosine

lesions (pyrimidine dimers, thymine glycols and urea residues) caused by ionizing radiation and oxidative agents such as osmium tetroxide and potassium permanganate (Breimer and Lindahl, 1984). Endonuclease III also cleaves DNA at AP sites. The repair activity involves the hydrolysis of the glycosylic bond of the damaged base with the concerted cleavage of the phosphodiester bond 3' to the resulting baseless site (Demple and Linn, 1980). Most recently there have been reports of another endonuclease detected in human cells which also possesses glycosylase and AP endonuclease activity (S. Linn, personal communication). This UV endonuclease, like the other glycosylases, exhibits a substrate specificity for thymine glycols or cytosine hydrates generated by the damaging effects of UV light and is capable of cleaving DNA at AP sites. All known glycosylase-AP endonucleases have been classified as Class I AP endonucleases. Class I AP endonucleases are one of four theoretical Classes of AP endonucleases. These Classes will be discussed in some detail after a review of the generation and consequences of AP sites.

#### AP Damage

Apurinic and apyrimidinic (AP) sites are among the most common DNA lesions. Therefore, AP endonucleases play an extremely important role in the repair of DNA damage. An AP site is produced by the hydrolysis of the N-glycosylic bond that links a base to the deoxyribose of DNA. Under normal

physiological conditions, the rates of spontaneous hydrolysis of the glycosylic bond are  $10^{-6}$ /day for purines and  $5 \times 10^{-8}$ /day for pyrimidines (Lindahl, 1979). This results in the daily formation of  $10^5$  AP sites/cell/day. In addition, AP sites can be produced by the effects of radiation and various chemicals. Ionizing radiation leads to ring opening increasing the lability of the glycosylic bond (Friedberg, 1985). The treatment of DNA with alkylating agents also results in base loss (Hemminki, 1983). Finally, as discussed above, AP sites are generated as intermediates in the repair initiated by DNA glycosylases (Weiss and Grossman, 1987; Sakumi and Sekiguchi, 1990).

The repair of AP sites in DNA is most likely through the base-excision repair pathway described above. However, for completeness, the action of a purine base insertase (described earlier) may also be involved in the repair of AP sites. This activity involves the direct re-insertion of bases into single baseless sites (Deutsch and Linn, 1979a,b). Although Deutsch and Linn were able to demonstrate this activity was potassium dependent, associated with a large protein, and specific for purines, further characterization was unattainable due to the extreme lability of the insertase. The extent of insertase involvement in the repair of AP sites must await more stable preparations of insertase.

#### AP Endonucleases

AP endonucleases, however, have been isolated and well

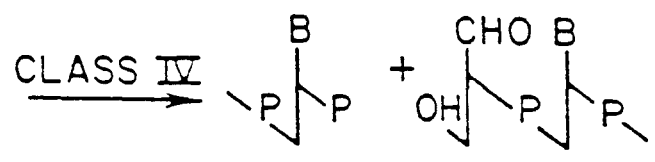
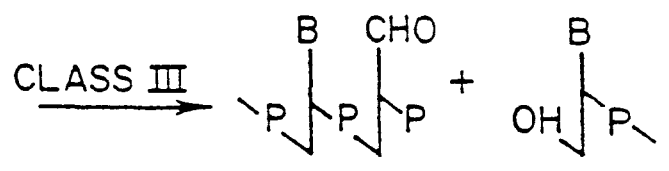
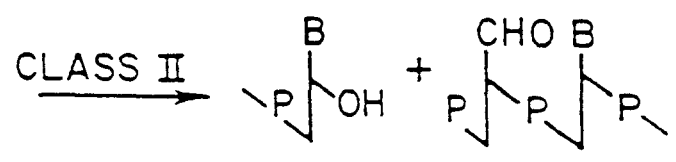
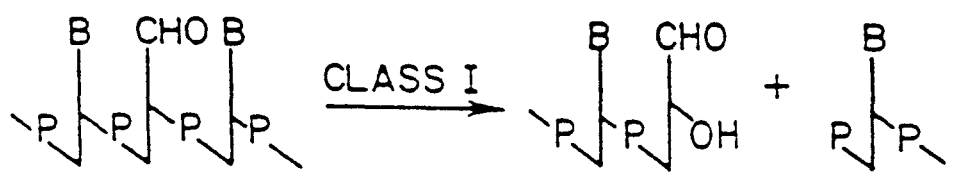
characterized from a number of organisms including but not limited to bacteria, *S.cerevisiae*, *D.melanogaster*, bovine and humans. Examples from each of these organisms will follow a more general description of AP endonucleases.

The action of AP endonucleases covers the third step of base-excision repair; the cleavage of the phosphodiester backbone of DNA. The phosphodiester bonds adjacent to an AP site can in theory be cleaved in four different ways. Therefore, AP endonucleases have been placed into four classes (Figure 2, Doetsch and Cunningham, 1990; Linn *et al.*, 1981). Class I AP endonucleases, also known as AP lyases, are considered to be beta-elimination catalysts, cleaving 3' to the AP site producing an unsaturated aldehyde as the 3' terminus and a 5'-phosphate as the 5' terminus (Bailly and Verly, 1987; Kim and Linn, 1988). The 3' terminus is not readily recognized by the 3' to 5' exonuclease activity of DNA polymerase and is not a good primer for DNA synthesis. A 3'-repair diesterase is required which will cleave 5' to the deoxyribose unsaturated aldehyde generating a 3'-hydroxyl termini. This termini is a good primer for DNA synthesis.

Class II AP endonucleases cleave hydrolytically 5' to the AP site producing a normal 3'-terminal 3'-hydroxyl nucleotide and a 5'-terminal deoxyribose-5-phosphate (Warner *et al.*, 1980; Mosbaugh and Linn, 1980). Although the 3' hydroxyl termini is a good primer for DNA synthesis, the deoxyribose-5-phosphate cannot be removed by the 5' to 3' exonuclease



Figure 2. The four theoretical cleavage sites for AP endonucleases. CHO denotes a baseless site and B denotes a normal base. (From Linn et al., 1981).



activity of DNA polymerase (Franklin and Lindahl, 1988). A deoxyribosephosphodiesterase must cleave 3' to the deoxyribose-5-phosphate to generate a one nucleotide gap. Now DNA polymerase can efficiently repair this site. A unique feature of Class II AP endonucleases is that they can remove several 3'-terminal blocking groups, including the 3' unsaturated aldehydes produced by Class I AP endonucleases, to produce efficient termini for DNA synthesis (Johnson and Demple, 1988a, 1988b). Therefore, Class I and Class II AP endonucleases can work in concert to generate a single nucleotide gap. Class I and Class II AP endonucleases are the two major Classes of enzymes known to cleave DNA at AP sites. Additionally, Class II AP endonucleases are the major AP endonucleases in most organisms (Lindahl, 1982).

A third Class of AP endonucleases, Class III, cleaves 3' to an AP site, via an unknown mechanism, to yield an abasic sugar with a 3'-phosphoryl group and a 5'-hydroxyl termini. The 3'-phosphoryl group is not recognized by the 3' to 5' exonuclease activity of DNA polymerase and, as in the case for Class I AP endonucleases, must be removed by a 3'-repair diesterase to generate an efficient termini for DNA synthesis. Only one example of a Class III AP endonuclease has been isolated (Spiering and Deutsch, 1986).

The final, theoretical Class of AP endonucleases, Class IV, would cleave 5' to the AP site producing a 3'-phosphoryl terminus and a 5' deoxyribose-5-hydroxyl terminus. Both of

these termini would have to be further modified to allow for efficient DNA synthesis to occur. The 3'-phosphoryl group would be cleaved by the 3'-diesterase discussed above (Class I) or by a Class II AP endonuclease. The 5'-deoxyribose-5-hydroxyl termini would also have to be removed perhaps by the deoxyribosephosphodiesterase required to assist Class II AP endonucleases. No examples of Class IV AP endonucleases have been isolated. However, Doetsch et al. (1986) isolated a calf thymus endonuclease which was shown to cleave both phosphodiester bonds that flank the abasic site, completely excising the deoxyribose. The resulting termini were 3'-phosphoryl and 5'-phosphoryl groups. The novel activity of the calf thymus endonuclease is unique and difficult to attribute to any of the previously described classes of AP endonucleases.

The following section will review some of the best characterized AP endonucleases. Figure 3 summarizes some of the properties of the AP endonucleases discussed in the next section.

Exonuclease III (Exo III) of *E.coli*, (Class II) was first identified as a byproduct of DNA polymerase I (Richardson and Kornberg, 1964) and subsequently shown to have 3'-phosphatase activity (Richardson et al., 1964) and RNAase H activity (Keller and Crouch, 1972). Exonuclease III was also shown to have endonuclease activity in two independent laboratories (Friedberg and Goldthwait, 1969; Verly and Rassart, 1975;

Figure 3. List of some AP endonucleases from a variety of sources.

- References:
- a. Lindahl, T., 1979; Lindahl, T. 1982.
  - b. Pierre and Laval 1980a,b.
  - c. Popoff *et al.*, 1990; Ramotar *et al.*, 1991.
  - d. Spiering and Deutsch, 1981, 1986.
  - e. Kelley *et al.*, 1989.
  - f. Sander *et al.*, 1991.
  - g. Kim and Linn, 1989.
  - h. Seki *et al.*, 1991a,b.
  - i. Sanderson *et al.*, 1989; Robson *et al.*, 1991.
  - j. Kane and Linn, 1981; Demple *et al.*, 1991.
  - k. Grafstrom *et al.*, 1982; Shaper *et al.*, 1982.

**AP endonucleases from a variety of sources**

Source	Mol. wt. (kD)	Cofactor Requirement	Class	Reference
<i>E. coli</i>				
Exonuclease III	28	Mg <sup>++</sup>	II	a
Endonuclease III	25	None	I	a
Endonuclease IV	33	None	II	a
<i>M. luteus</i>				
AP endonuclease A	35	None	II	b
AP endonuclease B	35	None	II	b
<i>S. cerevisiae</i>				
APN1	40	Co <sup>++</sup>	II	c
<i>D. melanogaster</i>				
AP endo I	66	None	III	d
AP endo II	63	Mg <sup>++</sup>	I	d
AP3	35	Mg <sup>++</sup>	II	e
Rrpl	74	Mg <sup>++</sup>	II	f
Mouse				
UV endonuclease I	43	None	I	g
UV endonuclease II	28	None	I	g
APEX	35	Mg <sup>++</sup>	II	h
Bovine				
BAP1	38	Mg <sup>++</sup>	II	i
Human				
HeLa, APE	37	Mg <sup>++</sup>	II	j
Placenta	37	Mg <sup>++</sup>	II	k

Gossard and Verly, 1978). Gossard and Verly (1978) characterized Exo III as having a molecular weight of 28,000, and a requirement for magnesium. Warner et al., (1980), later identified the endonuclease activity as Class II. The list of activities associated with Exo III was further expanded to include phosphomonoesterase and phosphodiesterase activities recognizing 3'-phosphoryl groups and 3'-phosphoglycolates, respectively (Demple et al., 1986). Exonuclease III is encoded by the *xth* gene and mutants in this gene are sensitive to killing by MMS (Lindahl, 1982; Friedberg; 1985), H<sub>2</sub>O<sub>2</sub> (Demple et al., 1983) and exposure to UV (Sammartano et al., 1986). These mutants (*xth*) also demonstrate a lack of 85% of the normal cellular AP endonuclease activity. The residual AP endonuclease activity led to the identification and isolation of the other *E.coli* AP endonuclease, endonuclease IV (Endo IV).

Endonuclease IV (Class II) was purified by Ljungquist et al., (1976) from *xth* (Exo III) mutants found to have residual AP activity. The gene encoding endonuclease IV (*nfo*) has been cloned (Cunningham et al., 1986) and shown to be inducible (Chan and Weiss, 1987). Endo IV, having several similar activities to Exo III will remove blocking groups from the 3' termini of duplex DNA (Demple et al., 1986; Mosbaugh and Linn 1982). The AP endonuclease activity has also been shown to be Class II (Bailly and Verly 1989). The big difference between Endo IV and Exo III is the additional exonuclease activity of

Exo III. A third prokaryotic AP endonuclease, endonuclease III (Class I) has an associated DNA N-glycosylase activity which recognizes urea residues and thymine glycols (discussed earlier in DNA glycosylases). The gene encoding endonuclease III, the *nth* gene has been cloned (Cunningham and Weiss, 1985) and sequenced (Asahara *et al.*, 1989). After some ambiguity in its classification Manoharan *et al.* (1988) have shown the mechanism of cleavage to be that of Class I.

A wide variety of eukaryotic AP endonucleases have been identified and characterized to varying degrees from yeast to mammals. Many of the recently characterized AP endonucleases show extensive similarity to the prokaryotic AP endonucleases discussed above (mainly Exo III and Endo IV). A few examples of eukaryotic AP endonucleases are discussed below (Also see Figure 3).

A large number of AP endonucleases have been partially or extensively purified from yeast. Armel and Wallace (1978, 1984) identified five distinct yeast AP endonucleases. All of these AP endonucleases were identified as having Class II activity like *E.coli* endonuclease IV. Recently, Johnson and Demple (1988a, 1988b), purified a 3'-diesterase activity shown to be the major AP endonuclease in yeast. The structural gene, APN1, was cloned and shown to have extensive homology to *E.coli* endonuclease IV (Popoff *et al.*; 1990). Popoff *et al.* (1990), using gene disruption experiments, confirmed that APN1 accounted for greater than 97% of both AP endonuclease and DNA



3'-repair diesterase activities in yeast cell-free extracts. The APN1 product also replaced the Endo IV activity in *nfo*-mutant strains (Ramotar *et al.*, 1991) conferring resistance to oxidizing and alkylating agents in these mutants.

As with yeast, the number of mammalian AP endonucleases is large but can be divided into two groups. In general, the AP endonucleases are either divalent cation-dependent and hydrolyze 5' to AP sites (Class II) or are divalent cation-independent and cleave 3' to AP sites (Class I). As discussed earlier, the Class II, divalent cation-dependent AP endonucleases, appear to represent the major mammalian AP endonucleases. The Class I, divalent cation-independent AP endonucleases also function as DNA N-glycosylases which initiate the first step of base-excision repair.

Kim and Linn (1989) purified and characterized two UV endonucleases (I and II) from murine plasmacytoma cells which recognize thymine glycols and possess Class I AP endonuclease activity. Both UV endonuclease I and II possess thymine glycol DNA glycosylase activity and cleave AP sites by  $\beta$ -elimination (Class I). However, UV endonuclease I is larger and possesses different parameters for optimal activity than UV endonuclease II. Kim and Linn suggest that these findings are analogous to a situation present in *E.coli* where endonucleases III and VII are non-identical but function in a similar role (Wallace, 1988). Another mouse AP endonuclease, APEX (Class II), was purified by Seki *et al.* (1991) and shown

to recognize bleomycin-damaged DNA and apurinic DNA. APEX is a magnesium dependent, multi-functional enzyme with Class II AP endonuclease, DNA phosphatase, 3'-diesterase and 3' to 5' exonuclease activities. The cDNA for APEX was cloned and the deduced amino acid sequence was shown to have significant homology to exonuclease III of *E.coli*.

A similar protein to APEX was isolated from bovine thymus and investigated by several groups (Ljungquist *et al.*, 1975; Henner *et al.*, 1987; Sanderson *et al.*, 1988). Sanderson *et al.* (1989) successfully determined the cleavage mechanism of the bovine AP endonuclease (BAP1) as a Class II mechanism. The cDNA clone encoding BAP1 was isolated and the deduced amino acid sequence showed strong homology to APEX of mouse and Exo III of *E.coli* (Robson *et al.*, 1991). Robson *et al.* (1991) have also recently reported the isolation of a cDNA encoding the human homologue of the BAP1 enzyme. The predicted amino acid sequence is homologous with that of BAP1 and ExoIII. These homologies suggest an evolutionarily conserved requirement for certain DNA repair enzymes. Comparing many of the properties of the bovine thymus enzyme with the major human cellular AP endonucleases also indicates a functional conservation between human and bovine cells.

The major human AP endonucleases have been purified from HeLa cells by Kane and Linn, (1981) and shown to be multifunctional in their activities like their bacterial counterparts, Exo III and Endo IV of *E.coli* and APN1 of *S.*

*cerevisiae* (Levin et al., 1988; Johnson and Demple, 1988). All of these enzymes cleave AP sites as Class II AP endonucleases. Demple et al. (1991) have recently cloned the cDNA (APE) encoding the major human AP endonuclease. It is not clear whether this cDNA encodes the human homologue of BAP1. This group identifies APE as a member of a family of DNA repair enzymes which includes Exo III and BAPI. Another member of this family, Rrpl from *Drosophila melanogaster*, will be discussed in the next section.

AP endonucleases have also been purified from *Drosophila melanogaster*. Two chromatographically distinct AP endonucleases, I and II (Class III and Class I, respectively), have been isolated from *Drosophila* embryos (Spiering and Deutsch, 1981, 1986). Both enzymes are larger in size (63-66 kd) than other eukaryotic AP endonucleases. Spiering and Deutsch (1986) characterized the mechanism of action for these two enzymes demonstrating that AP endonuclease I is a novel Class III enzyme and AP endonuclease II is a Class I enzyme. The cleaving mechanism described for AP endonuclease I (cleaving 3' to the AP site leaving a deoxyribose 3'-phosphate and a 5'-hydroxyl) is unique for all known endonucleases. Margulies and Wallace (1984) characterized a partially purified AP endonuclease activity from various developmental stages. The class of activity, however, was not reported. More recently, Sander et al. (1991) characterized an AP endonuclease with homologous recombination activities. This

combination of activities has not been observed previously. The protein, Rrpl (for recombination repair protein 1), was shown to have Class II AP endonuclease and 3' exonuclease activities. Rrpl also carries out single-stranded DNA renaturation in a magnesium-dependent manner. Analysis of the deduced 252-amino acid C-terminal region of Rrpl revealed homology to Exo III of *E.coli* and Exonuclease A from *S.pneumoniae*. All three of these enzymes belong to the larger family of DNA repair enzymes (based on homology) described earlier by Demple *et al.* (1991). The characteristics of Rrpl demonstrate yet another unique DNA repair enzyme found in *Drosophila*.

Recently, Kelley *et al.* (1989) reported the cloning of a cDNA (AP3) that encodes a *Drosophila* AP endonuclease. A human HeLa cell AP endonuclease antibody (Kane and Linn, 1981) was used to isolate the clone from a  $\lambda$ gt11 expression library. The cDNA encodes a protein with a molecular weight of 35,000 in much closer agreement in size to other eukaryotic AP endonucleases than AP Endo I and II previously isolated by Spiering and Deutsch (1981, 1986). Further, molecular and biochemical characterization will better define its relationship with previously identified *Drosophila* AP endonucleases and is the goal of this project.

#### Ribosomal Proteins Involved in DNA Repair

The *Drosophila* gene (AP3), having been isolated using antibody to a human AP endonuclease (discussed above; Kelley

et al. 1989), was used to isolate its human homologue (discussed in Results). A human gene, PO, was isolated and shown to be 66% identical and 79% similar at the amino acid level to the *Drosophila* AP3 gene (Grabowski et al., 1991; Results and Discussion). The PO gene, encoding a ribosomal protein, was also cloned by a separate group (Rich and Steitz, 1987) using sera from patients with systemic lupus erythematosus (SLE). SLE is an autoimmune disease generally characterized by serum antibodies directed against nuclear and cytoplasmic proteins as well as nucleic acids (Christian and Elkon, 1986). Anti-ribosomal antibodies from SLE patients show almost exclusive reactivity against three 60S ribosomal subunit phosphoproteins, PO, P1 and P2. Uchiumi et al. (1990), showed that these P-proteins participate in the binding of factors necessary for proper ribosomal function (i.e. protein synthesis-elongation cycle). Whether PO, and the *Drosophila* AP3, have a dual function as a ribosomal protein and as a DNA repair protein remains to be clarified. There is recent evidence however, linking ribosomal proteins to DNA repair.

Unpublished results from the laboratory of Dr. Stuart Linn (Univ. of California, Berkley) indicate they have isolated and characterized UV endonucleases capable of recognizing thymine glycol or cytosine hydrates induced by UV light. One of the UV endonucleases, possessing glycosylase/AP endonuclease activity, was purified to homogeneity from human

fibroblasts. Microsequence data establishes the identity of the UV endonuclease to be the human ribosomal protein S3.

Dr. Linn has also isolated a defective human S3 protein from XP-D cells suggesting that a defective S3 protein is the cause for the repair deficiency in these cells (personal communication). Pogue-Geile *et al.* (1991) isolated a cDNA clone encoding the human S3 ribosomal protein from a normal human colon cDNA library. The intriguing aspect of their work is that S3 was identified as one of several ribosomal proteins, including PO, whose level of expression was increased (4-10 fold) in colorectal tumors and polyps. Expression of PO has also been shown to be elevated (20-50 fold) in Mer human cells (deficient in O<sup>6</sup>-methyl-guanine methyltransferase) defective in the repair of DNA lesions generated by alkylating agents (Grabowski *et al.*, 1992). All of this information strongly supports a correlation between some ribosomal proteins and DNA repair.

#### DNA Repair and Autoimmune Deficiencies

As mentioned above, PO was cloned using sera from patients with SLE. Since SLE is an autoimmune disease, the possible link between DNA repair and immune deficiency syndromes is intriguing.

Several mammalian DNA-repair-defective mutants have been characterized that also have an associated immunodeficiency. In humans, ataxia telangiectasia (AT) is associated with an immunodeficiency, an increased incidence of cancer and defects

in DNA repair (McKinon, 1987). The analogous *scid* mouse, which lacks a functional immune system, has been shown to exhibit a profound hypersensitivity to DNA-damaging agents that cause double-strand breaks (Hendrickson et al., 1991). The data from Hendrickson et al. suggests that the *scid* gene product is involved in two pathways: repair of double strand breaks and the variable-(diversity)-joining [V(D)J] DNA rearrangement process.

## CHAPTER II

### MATERIALS AND METHODS

#### Screening Genomic Library by Plaque Colony Hybridization

A Maniatis genomic library from *Drosophila* was screened with radiolabeled AP3 cDNA by standard procedures (Maniatis, *et al.*, 1982). Approximately  $5 \times 10^6$  plaques were screened by plating  $2 \times 10^4$  plaques on 150 mm plates. Duplicate transfers to nitrocellulose filters were made from each plate. The filters were pre-hybridized for several hours at 42°C in 50% deionized formamide, 5X Denhardt's solution, 1% sodium dodecyl sulfate (SDS), 1 M NaCl, 10 Mm NaPO<sub>4</sub> (Ph 6.5), 0.1% pyrophosphate, and 2.5 mg/ml salmon sperm DNA (heat denatured by boiling 10 min.) and then hybridized overnight at 42°C with labeled probe in 50 ml of same solution. After hybridization, the filters were rinsed initially with 400 mls of 2 X SSC (1 X SSC is 0.15 M NaCl, 15 Mm Na Citrate, Ph 7.2) and 2% SDS at room temperature for 15 min. Final washes were in 1XSSC, 1% SDS for 30 min (3 times) at 37°C.

Autoradiographies of the blotted, wrapped filters were produced using XAR-5 film and intensifying screens at -70°C overnight. If autoradiograph was too dark, filters were washed at higher stringencies using increased temperature and decreased concentrations of SSC and SDS.



### Screening a cDNA Library for Human Homologue to AP3

A HeLa cDNA library was screened with radiolabeled AP3 cDNA essentially as described in Screening Genomic Library above. The stringency of hybridization was determined by the stringencies used during Southern analysis of AP3 on human genomic DNA. The pre-hybridization and hybridization solutions contained 30% formamide and were kept at 37°C throughout the procedure. The wash conditions were essentially as described above.

### Purification of $\lambda$ phage DNA

After positive phage clones have been isolated and purified to homogeneity through successive screens, the  $\lambda$  DNA must be purified to obtain the insert DNA. The procedure used to isolate the  $\lambda$  DNA was a modified version of Maniatis *et al.* (1982). An inoculum of  $10^5$  plaque-forming units (PFU) was introduced into 0.2 ml of the appropriate bacterial host strain (phage from Maniatis and  $\lambda$ gt10 libraries use *E.coli* C600 cells,  $\lambda$ gt11 libraries use *E.coli* Y1090 cells) and incubated at 37°C for 20 min. Approximately 8 mls of melted top agar (45°C) were added to the adsorbed phage/bacteria and spread on a 135 mm agar plate containing no antibiotics (C600 cells) or 100  $\mu$ g/ml ampicillin (Y1090 cells). After the top agar hardened the plates were inverted, and incubated at 37°C overnight. The amount of PFU's introduced should give confluent lysis of cells. The plates were cooled for 15 min. at 4°C and 10 mls of 1 X SM (100 mM NaCl, 1 mM MgSO<sub>4</sub>, 50 mM

Tris-HCl, pH 7.5, and 2% gelatin) was added to collect the phage during an overnight incubation at 4°C with gentle shaking. After incubation the fluid containing the phage was collected and clarified by centrifugation at 10,000 X g for 10 min. The supernatant was filtered through a 0.45 micron filter (Whatman) to remove residual bacteria. Nuclease mix (10 mg/ml RNAase A and DNAase I) was added to a final concentration of 5 ug/ml. Phage drop buffer (40% Polyethylene glycol and 4M NaCl) was added to a final concentration of 25% (eg: 2.5 ml phage drop buffer/7.5 ml supernatant). The precipitating phage was incubated at 37°C for 30 min and then overnight at 4°C. The phage were pelleted by centrifugation at 12,000 X g for 10 min and resuspended in 150  $\mu$ l TE. An equal volume of Phage Lysis/Binding Buffer (BIO101 INC.) was added to the resuspended phage, vortexed briefly and heated to 70°C for 10 min. The DNA was purified using  $\lambda$ DNA glassmilk from BIO101 (binding capacity is 1  $\mu$ g DNA/1  $\mu$ l glassmilk). The solution containing 5-10  $\mu$ l glassmilk was incubated at room temperature for 10 min with intermittent mixing allowing the  $\lambda$ DNA to bind. The  $\lambda$ DNA, bound to glassmilk was pelleted by centrifugation, washed with 500  $\mu$ l 1 X SM three times by resuspending and pelleting. Finally, the  $\lambda$ DNA was eluted in 100  $\mu$ l TE heated to 50°C for 5 min. The insert DNA was released by digestion with restriction enzymes, isolated, and cloned into the pBluescript SK<sup>+</sup> vector.

### Generation of Single-Stranded DNA for Sequencing

Following the procedures described by Russel *et al.* (1986) patches of cells (HB101/F<sup>+</sup>) containing the plasmid of interest were grown on LB agar plates with 100 µg/ml ampicillin and 50 µg/ml kanamycin overnight at 37°C. A loop of cells was transferred to 1.5 mls of 2X-YT medium (16 g/L Tryptone, 10g/L yeast extract, 5 g/L NaCl) containing 0.1% glucose, 100 µg/ml ampicillin and 10<sup>10</sup> PFU R408 helper phage and were allowed to grow with vigorous shaking at 37°C for 4 h. The cells were pelleted by centrifugation at 10,000 X g for 5 min and the supernatant was transferred to sterile microcentrifuge tubes containing 200 µl 2.5 M NaCl and 20% polyethylene glycol (PEG) solution. The solutions were mixed well and incubated at room temperature for at least 15 min. The particles containing single stranded DNA were pelleted by centrifugation at 10,000 X g for 10 min. All traces of PEG solution were removed by aspiration. The pellets were resuspended in 100 µl of TE<sub>2</sub> (10 mM Tris HCl, pH 8.0, 2 mM ethylenediamine tetraacetic acid (EDTA)) and extracted with 0.5 volume phenol. The aqueous phase was mixed with 250 µl of 25:1 mixture of ethanol and 3 M sodium acetate (pH 5.2) in a fresh tube. After incubating at -70°C for 30 min the precipitates were collected by centrifugation, washed with 100% ethanol, dried at 65°C for 5 min and resuspended in 20 µl TE (10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA). 5 µl of this solution contained sufficient single-stranded DNA for DNA

sequencing.

### DNA Sequencing

Nucleotide sequences of genomic fragments and cDNAs were determined by the dideoxynucleotide chain termination method (Sanger *et al.*, 1977) utilizing synthetic oligonucleotides as primers. Concurrently, sequencing of larger genomic fragments was accomplished using the overlapping deletion technique described by Henikoff *et al.* (1984). Briefly, nested sets of deletions lacking progressively more sequence from one end were generated using exonuclease III (Exo III). Exo III catalyses the stepwise uniform removal of 5' hydroxyl ends of double-stranded DNA (Weiss 1976). However, Exo III fails to initiate digestion at a 4-nucleotide protruding 3' terminus (Rogers and Weiss, 1980). To generate deletions, plasmid DNA was digested with two restriction enzymes in the polylinker region of BSK<sup>+</sup> between the target DNA and the sequencing primer binding site. One enzyme protects the vector from Exo III attack, allowing unidirectional digestion of the target sequence. Aliquots were removed from the reaction mixture at uniform intervals. The exposed remaining single strands were removed by digestion with exonuclease VII. The ends were made blunt using Klenow DNA polymerase, the vectors were closed using T4 DNA ligase and transfected into competent *E.coli* HB101/F<sup>+</sup> cells.

### Labelling Double-Stranded DNA with Random Hexamers

Double stranded DNA to be used as probes for

hybridization were labeled to high specific activity with alpha-<sup>32</sup>P-dCTP (New England Nuclear, specific activity 3000 Ci/mmol) by the random hexamer primer method (Feinberg and Vogelstein, 1984). Approximately 25 ng of DNA in sterile H<sub>2</sub>O was denatured by heating to 95°-100°C for 2 min and rapidly chilled on ice. The labeling reaction was assembled in a 1.5 ml microcentrifuge tube by the addition of 10 μl 5X buffer with oligonucleotides [250 mM Tris-HCl, pH 8.0, 25 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM HEPES, pH 6.6, 27 A<sub>260</sub> U/ml pd (N)<sub>6</sub> hexamers (Pharmacia)], 2 μl acetylated BSA (1mg/ml), 2 μl dNTP's (1.5 mM each), 50 μCi alpha <sup>32</sup>P-dCTP (New England Nuclear) and 5 U klenow.

The final reaction volume was 50 μl. The reaction was incubated at room temperature for 1 h, denatured by heating to 95°-100°C for 2 min and chilled immediately on ice. EDTA is added to a final concentration of 20 mM (2.0 μl of 0.5 M EDTA) and brought up to a final volume of 100 μl with sterile H<sub>2</sub>O. Unincorporated <sup>32</sup>P-dCTP was removed by centrifugation through a spin column containing Sephadex G-50.

#### Isolation of Drosophila DNA

Following the procedures of Kidd *et al.* (1983), approximately 120 flies were homogenized in 300 μl of solution A (Sol. A = 10 mM Tris-HCl, pH 7.5, 60 mM NaCl, 10 mM EDTA, 0.15 mM Spermine, 0.15 mM Spermidine and 5% Sucrose). Next, 300 μl of solution B (sol. B = 1.25% SDS, 0.3 M Tris - HCl, pH 9.0, 0.1 mM EDTA, 5% Sucrose) was added to the tube and

incubated at 65°C for 45 min. To this solution was added 90  $\mu$ l of 8 M KOAC (pH 9.0) incubated on ice for 1 h. Following a 1 min centrifugation at 10,000 X g the supernatant was extracted with 2 volumes of chloroform and then precipitated with 2 volumes of ethanol. The nucleic acid was resuspended in TE (TE is 10 mM Tris-HCl, pH 8.0, 0.1 mM EDTA) and stored at 4°C.

### Southern Blot Analysis

Southern Blot analysis was used to construct the genomic map of AP3 (Southern, 1975). Briefly, DNA agarose gels to be transferred were washed in an acid depurination solution (2% HCl) for 10 min. and rinsed in denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 30 min. The gel was neutralized with two washes of 1.5 M NaCl and 1.0 M Tris-HCl, pH 7.4 for 30 min each. The DNA was transferred to Nytran (Schleicher and Schuell) by capillary action with 10 X SSC (1.5 M NaCl, 0.15 M Na citrate, pH 7.0) overnight. Following transfer the nucleic acids were crosslinked to the membrane using a UV Crosslinker (Stratagene).

The membranes were hybridized with various probes following the protocol described for Screening Genomic Library. Post-hybridization washes were at high stringency (0.5X SSC, 0.5% SDS and 65°C).

### Total Drosophila RNA Isolation

Total RNA was isolated from embryonic, larval and adult stages and fractionated on formaldehyde-agarose gels

(Maniatis, et al., 1982). The procedure follows that of Kelley et al. (1989). Approximately 50 mg of flies were homogenized in 500  $\mu$ l of Guanidinium- thiocyanate solution (GIT = 4 M Guanidine isothiocyanate, 25 mM Na citrate, pH 7.0, 0.5% Sarcosyl and 0.1 M 2-mercaptoethanol) in a microcentrifuge tube. After incubating at room temperature for 5 min, 50  $\mu$ l of 2 M NaOAC (pH 4.0), 500  $\mu$ l phenol and 100  $\mu$ l chloroform: isoamyl alcohol (49:1) were added, with mixing after each addition, and incubated on ice for 15 min. The tubes were centrifuged at 10,000 X g for 20 min at 4°C. The aqueous (top layer) phase was mixed with an equal volume of ethanol to precipitate nucleic acids. The RNA was pelleted by centrifugation as above and resuspended in 150  $\mu$ l of GIT solution. The RNA was precipitated again and pelleted by centrifugation, washed with 70% ethanol and resuspended in diethyl pyrocarbonate treated water and stored at -70°C.

#### Northern Blot Analysis

Following electrophoresis of RNA in agarose gels, the gels were prepared for transfer by soaking them for two 20 min periods in 10 X SSC at room temperature. The RNA was transferred to Nytran (Schleicher and Schuell) as described for Southern blot analysis. Hybridization with labelled probe was as described for Southern Blot Analysis.

#### Construction of $\beta$ -galactosidase-AP3 Overexpression Vectors

To overexpress the protein encoded by AP3, portions of the cDNA were subcloned into the pWR590 series of

overexpression vectors (gifts of Dr. R. Wu, Cornell University) which generate  $\beta$ -galactosidase fusions with the protein of interest. The pWR590 series of vectors contains the isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) inducible *E.coli lac* promoter and a portion of the coding sequence for  $\beta$ -galactosidase. The coding region ends with a polylinker region which allows for the subcloning of the gene of interest into any of eight restriction endonuclease sites. The pWR590 series includes three vectors that enable fusion to be produced in all three translational reading frames.

The first of two  $\beta$ -galactosidase-AP3 fusions (d1) was constructed by subcloning a *Sac*II-*Bam*HI fragment from the AP3 cDNA into the *Sma*I-*Bam*HI sites of the pWR590 vectors. Both *Sac*II and *Sma*I cleave DNA leaving a blunt end, therefore these sites are complimentary and compatible for ligation. The *Sac*II-*Bam*HI fragment includes the coding region for amino acids 14-317 and a short 3' untranslated region.

The second  $\beta$ -galactosidase-AP3 fusion protein (d8) was constructed similarly using a *Sac*II fragment which includes the coding region for amino acids 67 to 285. Before the *Sac*II fragment was subcloned into the *Sma*I site of the pWR590 series, the ends of the fragment were made blunt ended with Klenow. The blunt ended fragment was ligated to the blunt ends of the pWR590 vectors cleaved with *Sma*I using T4 DNA ligase.



### Thermal Cycle Amplification of AP3 cDNA

To overexpress the protein encoded by AP3, the cDNA was prepared for subcloning into overexpression vectors by the polymerase chain reaction (PCR). Oligonucleotide primers were designed to recognize the N-terminal and C-terminal four amino acids. Two sets of oligonucleotides were designed; each having the appropriate restriction endonuclease site enabling the directional subcloning of the amplified product into the appropriate expression vector (pGEX-3X (Pharmacia LKB) or pET11d (Novagen)). pGEX-3X vectors with insert produce a fusion protein with glutathione S-transferase (GST, 26kd) at the amino end of the protein with a Factor Xa cleavage site between the GST portion of the fusion protein and the protein of interest, AP3. The pET-11d vector, with insert, allows for the overexpression of the protein of interest, AP3, as a non-fusion protein. The N-terminal primer for the pGEX vector (5'-CGTGGGATCCCATGCCAGGG-3') contains a *Bam*HI site (underlined). The C-terminal primer for the pGEX vector (5'-CGATGAATTCGGCAGATGGATC-3') contains a *Eco*R1 site (underlined). The N-terminal primer for the pET11d vector (5' TATACCATGGTTAGGGAG-3') contains a *Nco*I site (underlined). The C-terminal primer for the pET11d vector (5' GGATCCAGCTTAGTCGAACAG-3') contains a *Bam*HI site (underlined). The AP3 cDNA, previously subcloned into pBluescript SK<sup>+</sup> (Kelley et al. 1989) was linearized by restriction endonuclease digestion. Approximately 1 $\mu$ g of linearized DNA

was diluted into a final volume of 100 $\mu$ l in 10mM Tris-HCl (pH 8.3), 50mM KCl<sub>2</sub>, 1.5mM MgCl<sub>2</sub>, 200 $\mu$ M each of deoxynucleotide triphosphate, 0.01% gelatin, 0.01% Tween-20, 0.01% Nonidet P-40, 2 $\mu$ M of each oligonucleotide and 2U Taq polymerase (Bethesda Research Laboratories or Promega). The polymerase amplification was carried out using an Eppendorf MicroCycler programmable heating/cooling dry block for 30 cycles of amplification (94°C, 30 sec, 60°C, 1 min, 72°C, 2 min), followed by 10 min at 72°C. The resulting product was purified, cleaved with the appropriate restriction endonucleases and subcloned into the appropriate vector.

#### Overexpression of Fusion and Non-fusion Proteins in *E.coli*

All of the overexpression vectors used in this project (pWR590, pGEX-3X, and pET11d) were induced following the same procedure. An overnight culture of LB-broth containing ampicillin (100  $\mu$ g/ml) was diluted 1:10 into fresh culture medium. The culture was grown with vigorous shaking at 37°C for 1 h. The inducing agent, IPTG, was added to a final concentration of 0.1 mM and the culture was allowed to grow for another 2-4 hours. The cells were collected by gentle centrifugation (~500 X g) and resuspended in Buffer A (see below) for rescue of overexpressed proteins. Crude extracts were run on sodium dodecyl sulfate (SDS) -polyacrylamide gels with subsequent staining to determine the extent of over expression and molecular weight.

## Purification of Proteins from Inclusion Bodies

The overexpressed proteins from the pGEX (Pharmacia) and pET11d (Novagen) vectors containing AP3 or PO were found to be insoluble and trapped in inclusion bodies. Many overexpressed proteins accumulate in the cytoplasm of host cells in the form of insoluble inclusion bodies that account for a major fraction of the total expressed proteins in *E.coli* (Kane and Hartley, 1988). Purification of inclusion bodies leads to fewer subsequent steps for purifying the overexpressed proteins.

The purification of inclusion bodies is a modified version of that described by Lin and Cheng (1991). Cells collected after induction and growth of 2-4 hours were resuspended in 1/10 volume of Buffer A (20 mM Tris-HCl, pH 7.5, 20% sucrose, 1 mM EDTA) and incubated on ice for 10 min. The cells are pelleted by centrifugation at 4000 X g and resuspended in same volume of ice cold H<sub>2</sub>O. After being incubated for 10 min, the spheroplasts were separated from the contaminating outer cell wall by centrifugation at 8000 X g. The pelleted spheroplasts were resuspended in the same volume of buffer P (150 mM NaCl, 16 mM Na<sub>2</sub>HPO<sub>4</sub>, 4 mM NaH<sub>2</sub>PO<sub>4</sub>, 5 mM EDTA, pH 7.3) containing protease inhibitor [leupeptin (1 µg/ml), aprotinin (20 µg/ml), phenylmethylsulfonyl flouride (PMSF) (0.5 mM)]. After brief sonication (50W, 3 times 15 seconds) with a 30 second pause in between pulses) RNase T1 and DNase were added to the sonicated cell suspension, (1. 3

$\times 10^3$  U/10 ml and 400  $\mu$ g/10 ml, respectively) and incubated at room temperature for 10 min. The cell suspension was diluted by the addition of 0.4 volumes of buffer P and the crude inclusion bodies were pelleted by centrifugation at 13,000 X g for 30 min. The pellet was suspended in 4.0 ml buffer W (buffer P with 25% sucrose, 1% Triton X-100), incubated on ice for 10 min and centrifuged at 25000 X g for 10 min. This wash step was repeated two more times. The clean inclusion bodies were denatured by resuspending in 2 mls of buffer D (50 mM Tris-HCl, pH 8.0, 5 M guanidine chloride, 5 mM EDTA) and brief sonication (50W, 5 sec. pulse). The suspension was incubated on ice for at least one hour up to overnight. The suspension was clarified by centrifugation at 12,000 X g for 30 min. The supernatant was added to 10 volumes buffer R (50 mM Tris-HCl, pH 8.0, 1 mM DTT, 20% glycerol and protease inhibitors) and mixed gently on a rocking platform at 4°C overnight. The supernatant was again clarified by centrifugation at 13,500 X g for 30 min. Purity of expressed protein was determined by silver-staining of polyacrylamide gels.

Alternatively, the clean inclusion bodies after the wash steps were used as the source of protein to be purified by electroelution from SDS-polyacrylamide gels.

#### Purification of Antigens for Antibody Production

There were two sources of protein for the purification of antigens. The first was cells containing  $\beta$ -galactosidase fusion proteins. These were induced with IPTG as described

above. The cells were collected by centrifugation (1000 X g), resuspended in 3-5 ml 2X SDS-gel sample buffer (120 mM Tris-HCl, pH 6.8, 20% glycerol, 70 mM SDS, 2% 2-mercaptoethanol, 0.5 mg Bromophenol Blue), and boiled for 5 min. The suspension was clarified by centrifugation at 12000 X g for 10 min. An appropriate volume of the supernatant (10-20 $\mu$ l/well) was loaded on a 1.5 mm thick 12% preparative SDS-polyacrylamide gel. The gels were run at 100 volts for 3-5 hrs. The gel was stained with coomassie blue for 30 min and destained in several washes of 20% methanol, 10% acetic acid. The band containing the expressed protein was cut out of the gel and placed into dialysis tubing containing 1X SDS running buffer (25 mM Tris-HCl, pH 8.0, 200 mM glycine, 0.1% SDS). The dialysis tubing containing the gel slices was placed in a horizontal gel apparatus and the proteins electroeluted for 2 hours at 30mA. After elution, the gel slice was removed and the tubing containing the protein was placed in a flask for dialysis against PBS (150 mM NaCl, 10 mM NaPO<sub>4</sub>, pH 7.4) overnight at 4°C.

The second source of protein for antigen purification was the washed inclusion bodies (See Purification of Proteins from Inclusion Bodies). The pellets containing washed inclusion body were resuspended in 2X SDS sample loading buffer and treated as described for  $\beta$ -galactosidase fusion proteins.

#### Production of Antibodies

Each of the overexpressed fusion-proteins, either  $\beta$ -

galactosidase fusions or glutathione-S-transferase (GST) fusions from the pWR590 series or pGEX-3X vectors, respectively, were purified from polyacrylamide gels by electroelution (see above). Approximately 100  $\mu$ g of antigen, emulsified in Freund's complete adjuvant were injected intradermally into female rabbits. Separate rabbits were used for each fusion protein. After three weeks, with subsequent boosts every 2-3 weeks, about 100  $\mu$ g of antigen in Freund's incomplete adjuvant was injected. Sera were collected at each boost with a determination of the integrity and titre of antibodies by Western blot analysis.

#### Purification of Antibodies

Immunopurification was carried out using antigen immobilized on nitrocellulose filters according to Maniatis et al. (1989). Briefly, purified antigen from inclusion bodies was electroblotted onto nitrocellulose from SDS-polyacrylamide gels. The filters were incubated in blocking buffer (see Western Blot Analysis) for 1 h at room temperature. Crude antisera was added to the filters and were incubated overnight at 4°C with gentle shaking. The crude antisera was removed and saved for subsequent purifications. The filters were rinsed with 1X TBST (see Western Blot Analysis) three times 20 min. One lane of the filter was cut off the filter to determine the position of the target antigen using an alkaline-phosphate conjugated anti-rabbit IgG antibody. Matching the stained lane with the original filter, a strip of

nitrocellulose carrying the target antigen was cut from the filters. Several nitrocellulose strips containing bound antigen were generated in this manner and could be used for several purifications. These strips were arranged on parafilm pressed onto the bottoms of petri dishes containing moist paper towels. Approximately 250  $\mu$ l of elution buffer (0.2 M glycine/pH 2.8, 1 mM EGTA) was applied to each strip. The petri dishes were incubated for 20 min at room temperature with gentle shaking. The elution buffer containing the antibodies of interest were collected and neutralized with 0.1 volume of 1 M Tris-HCl, pH 9.5 and made to a final concentration of 1X PBS (see Western Blot Analysis). The antibodies are stored at 4°C.

#### Subcellular Fractionation of Adult Drosophila Cells

To determine the subcellular localization of AP3 by Western Blot Analysis, the nuclei and chromatin of adult *Drosophila* were purified following Wu *et al.* (1979) and Kelly *et al.* (1989) respectively. Subnuclear fractions were generated using protocols established by Fisher and Blobel (1983). Each of the protocols is briefly described below.

#### Nuclei Purification

Following the procedure of Fisher and Blobel (1983), approximately 2 g of adult flies were homogenized in 10 mls of Buffer A (60 mM KCl, 15 mM NaCl, 1 mM EDTA, 0.2 mM EGTA, 0.15 mM spermine, 0.5 mM spermidine, 15 mM Tris-HCl (pH 7.4), 0.5 mM DTT and 0.1 mM PMSF) plus 1 M sucrose at 4°C using a Dounce

homogenizer. The homogenate was filtered through nylon mesh (TETKO INC.) and the filtrate clarified by centrifugation at 480 X g for 9 min. The pellet was discarded. The supernatant (S1) was made up to 0.2% Nonidet P-40 (NP-40), vortexed and centrifuged at 4300 X g for 10 min. The supernatant was saved and labeled post-nuclear supernatant (PNS). The pellet (P2) containing crude nuclei was resuspended in Buffer A1 (A1= Buffer A without the EDTA or EGTA) plus 1 M sucrose and centrifuged through a step gradient made up of Buffer A1 plus 1.5 M sucrose and Buffer A1 plus 1.8 M sucrose. Centrifugation was at 19,600 X g for 20 min. The supernatant was discarded and the purified nuclear pellet was resuspended in Buffer S (50 mM Tris-HCl, pH 7.8, 1 mM DTT, 1 mM EDTA and 0.5% Aprotinin) and at stored 0°C or resuspended in Buffer C (10 mM Tris-HCl, pH 7.5, 290 mM Sucrose, 0.1 mM MgCl<sub>2</sub>, and 2% Triton X-100) for purification of the chromatin. The purified nuclei were also stored in Buffer N (20 mM Tris-HCl, pH 7.5, and 5 mM MgCl<sub>2</sub>) for subsequent subnuclear fractionation.

#### Chromatin Purification

Following the procedure of Fisher and Blobel (1983), purified nuclei in Buffer C were incubated on ice for 10 min followed by centrifugation at 1100 X g for 10 min. The supernatant was saved as the TX-100 fraction. The pellet was resuspended and washed twice in 75 mM NaCl and 24 mM EDTA (pH 8.0). After subsequent centrifugations at 1100 X g the supernatants were saved as the NaCl/EDTA fraction. The



pellet, now containing disrupted nuclei, was resuspended and twice washed in 10 mM Tris-HCl (pH 8.0) and then twice washed in 5 mM Tris-HCl (pH 8.0). The supernatants were saved as the 10 mM and 5 mM fractions, respectively. Chromatin was prepared from the pellet material by overlaying it onto 5 mM Tris-HCl (pH 8.0) and 1.7 M sucrose and centrifuging at 50,000 X g for 3 h. Again the supernatant was saved as the Tris/Sucrose fraction. The resulting chromatin pellet was resuspended in 10 mM Tris-HCl (pH 8.0) and washed twice by centrifugation at 12,000 X g for 10 min.

#### Subnuclear Fractionation

Following the procedure of Fisher and Blobel (1983), purified nuclei in Buffer N were nuclease treated by adding DNase I and RNase A to final concentrations of 10  $\mu\text{g/ml}$  and 8  $\mu\text{g/ml}$ , respectively, and incubated at 37°C for 15 min. Nuclease supernatant was generated by centrifugation at 1000 X g for 10 min. The pellet was resuspended in 0.9 volumes Buffer C and 0.1 volumes 20% Triton X-100 and incubated at 0°C for 10 min. After centrifugation at 1000 X g for 10 min the supernatant was saved as the TX fraction. The pellet was resuspended in 0.5 volumes Buffer C and 0.5 volumes 2 M NaCl and incubated at 0°C for 10 min. After another centrifugation at 10,000 X g for 10 min the supernatant was saved as the SS-1 fraction. The 2 M NaCl step was repeated to generate the SS-2 fraction and the final pellet enriched in nuclear matrix, pore, and lamina components was stored in Buffer S at 0°C.

## Western Blot Analysis

The procedure for Western blot analysis follows the protocol of Kelley et al. (1989) with some modifications. Protein samples in 2X loading buffer (2X is 120 mM Tris-HCl, pH 6.8, 70 mM SDS, 20% glycerol and 2% 2-mercaptoethanol) were loaded on and electrophoresed through a 10% SDS-polyacrylamide gel at 100-150 volts in 1X electrophoresis buffer (5X electrophoresis buffer is 125 mM Tris-HCl, pH 8.3, 0.96 M glycine and 17 mM SDS) for about 1 h. The gels were prepared for transfer by soaking in transfer buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine and 20% methanol) for 15 min. Proteins are electroblotted to nitrocellulose (Schleicher and Schuell) with 150 mA of current for 1 h. The filters were prepared for antibody analysis by soaking in blocking solution (5% powdered milk and 1X TBST [10 X TBST = 0.1 M Tris-HCl, pH 8.0, 0.5% Tween 20 and 1.5 M NaCl]) for 1 h. The primary antibody solution (diluted 1:1000) was added to filters and incubated at 4°C overnight. The filters were washed three times for 10 min with 1X TBST. The secondary antibody, anti-rabbit alkaline-phosphatase conjugate in blocking solution (diluted 1:10,000) was added to the filter and incubated at room temperature for 2-3 h. The filters were washed as above and developed in alkaline phosphatase buffer (100 mM Tris-HCl, pH 9.5, 100 mM NaCl at 5 mM MgCl<sub>2</sub>) containing the chromogenic substrate 5-bromo-4-chloro-3-indoyl phosphate (BCIP) and nitro blue tetrazolium chloride (NBT).

The colorimetric reaction was stopped by placing the filter in stop solution (20 mM Tris-HCl, pH 8.0 and 5 mM EDTA). Alternatively, the secondary antibody was  $^{125}\text{I}$ -labeled anti-rabbit-IgG. After washes the filters are placed on film as described in Screening Genomic Libraries.

#### 5'-End Labelling of DNA Oligonucleotides

Synthetic oligonucleotides used as primers in mapping the 5' end of AP3 were labeled by the transfer of the gamma  $^{32}\text{P}$  from ATP to a 5' terminus by bacteriophage T4 polynucleotide kinase (Gibco/BRL) according to Maniatis *et al.* (1982). The reaction mixture containing 10 ng of oligonucleotide, 1X T4 kinase buffer (10X buffer = 0.5 M Tris-HCl, pH 7.6, 0.1 M  $\text{MgCl}_2$ , 50 mM DTT, 1 mM spermidine, 1 mM EDTA, pH 8.0) 50  $\mu\text{Ci}$  gamma  $^{32}\text{P}$ -ATP and 8 units of T4 polynucleotide kinase in a final volume of 20  $\mu\text{l}$  were mixed in a 1.5 ml microcentrifuge tube. The reaction mixture was incubated at 37°C for 45 min with subsequent heat killing of the kinase by incubation at 70°C for 10 min. The labeled oligonucleotide was precipitated by the addition of 1/10 volume 3 M NaOAc (pH 5.2) and 2 volumes 100% ethanol. This precipitation removes the bulk of unincorporated isotope. The labeled oligonucleotide was resuspended in 100  $\mu\text{l}$  of 0.3 M NaOAc and 1  $\mu\text{l}$  was removed for Cerenkov counting to estimate specific activity (dpm/ $\mu\text{g}$  oligonucleotide) of the probe.

#### 5'-End Mapping of AP3 Transcript

The precise 5' termini of mRNA can be determined through

primer extension, in which a radiolabeled oligonucleotide in hybridized to the mRNA and used as a primer for reverse transcriptase. The generation of extension products follows the standard procedures described by Maniatis *et al.* (1982). Radiolabeled primer ( $10^4 - 10^5$  cpm) was mixed with 5-10  $\mu\text{g}$  of RNA and precipitated by adding 0.1 volume of 3 M NaOAc (pH 5.2) and 2.0 volumes of ethanol. After storing the solution at  $-70^\circ\text{C}$  for 30 min the nucleic acids were recovered by centrifugation at 10,000 X g for 20 min at  $4^\circ\text{C}$ . The resulting pellet was washed with 70% ethanol, dried at room temperature and resuspended in 30  $\mu\text{l}$  of hybridization buffer (40 mM PIPES, pH 6.4, 1 mM EDTA, pH 8.0, 0.4 M NaCl, 60% formamide). The hybridization mixture was incubated at  $85^\circ\text{C}$  for 10 min and transferred to a  $40^\circ\text{C}$  water bath. The oligonucleotide was allowed to anneal to the transcript for 8-12 h. The hybrids were precipitated with 170  $\mu\text{l}$  of  $\text{H}_2\text{O}$  and 400  $\mu\text{l}$  of ethanol. After washing, the pellet was redissolved in 20  $\mu\text{l}$  reverse transcriptase buffer (RT Buffer is 50 mM Tris HCl, pH 7.6, 60 mM KCl, 10 mM  $\text{MgCl}_2$ , 1 mM each dNTP, 1 mM DTT, 1U/ $\mu\text{l}$  RNAase inhibitor, 50  $\mu\text{g}/\text{ml}$  actinomycin D). Approximately 50 units of Superscriptase (Gibco/BRL) reverse transcriptase was added to the reaction and incubated for 2 h at  $37^\circ\text{C}$ . After incubation, the RNA was removed by adding 1  $\mu\text{l}$  RNAse (5  $\mu\text{g}/\text{ml}$ ) and incubating for 30 min at  $37^\circ\text{C}$ . The solution was subjected to phenol/chloroform extraction, the aqueous phase transferred to a new tube and the nucleic acids recovered by precipitation.

The washed pellet was dissolved in formamide loading buffer (60% formamide, 10 mM EDTA pH 8.0, 1 mg/ml xylene cyanol, 1 mg/ml bromophenol blue) and analyzed by electrophoresis through a 6% polyacrylamide/7 M urea gel in 1X sequencing TBE (20X is 2M Tris-HCl, pH 8.0, 1.66M boric acid, 0.02M EDTA).

#### Tissue Culture of Drosophila Schneider II Cells

Schneider II cells are a nontransformed embryonic cell line. These cells were grown in 1X Schneider's Drosophila medium (Gibco) supplemented with 10% fetal calf serum and 100  $\mu$ g/ml penicillin and streptomycin. The cells were grown in T75 or T25 cell culture flasks at room temperature until near confluency. Dilutions of 1:10 to 1:15 mls of cells to fresh media were sufficient to propagate the cell line. Those cells used in Western analysis were collected by centrifugation at 5,000 X g for 10 min, rinsed in 1X PBS and stored frozen in 1/100 volume 1X PBS at -70°C.

## CHAPTER III

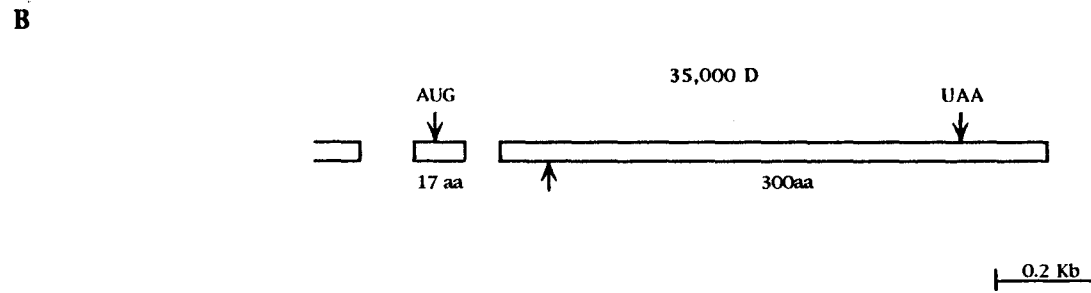
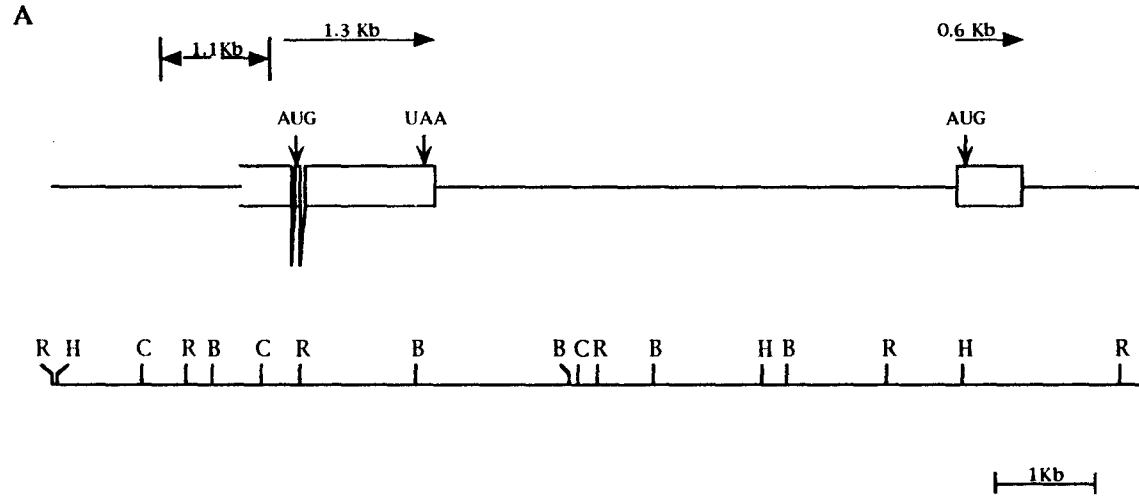
### RESULTS

The cDNA encoding a *Drosophila* AP endonuclease, *AP3*, was previously isolated by Kelley et al., (1989) as discussed earlier. *AP3* encodes a 35,000 dalton protein identical to the molecular size of an AP endonuclease activity recovered from *Drosophila* embryonic extracts, and is expressed throughout development. To characterize *AP3* on the genomic level, the *AP3* cDNA was used to screen a *D.melanogaster* genomic library to isolate genomic clones containing *AP3*. Several positive phage clones were purified to homogeneity through successive screening. The  $\lambda$  DNA was isolated from these clones and digested with restriction enzymes (*EcoRI* or *BamHI*). Three clones, G4, G5 and G6 all had the same restriction enzyme digested banding pattern and contained the same genomic fragments recognized by radiolabeled *AP3* cDNA on Southern blot analysis (data not shown). The restriction digests revealed bands ranging in size from 0.8kb to > 4.0kb. The larger bands were attributed to the arms of the  $\lambda$  vector or incomplete digestion. Southern analysis showed that the *AP3* gene is single copy, contained within two *BamHI* fragments (2.0kb and 1.6kb) or within two *EcoRI* fragments (1.2kb and 3.0kb).

To ease the manipulation of the genomic fragments for further analysis, each of the genomic clones was subcloned into pBluescript SK<sup>+</sup> vectors (Stratagene; La Jolla, CA). This vector was chosen because of the unique restriction sites in the multiple cloning site (MCS). Vectors containing each genomic fragment were transformed into HB101/F<sup>+</sup> *E.coli* cells for propagation and storage. Southern analysis and sequence data from portions of the genomic fragments resulted in the positioning of each fragment on a linear genomic restriction map of AP3 shown in Figure 4. The position of each fragment was confirmed by several different restriction digests utilized during Southern analysis. Sequences from the ends of the genomic fragments which overlapped either *EcoRI* or *BamHI* restriction sites confirmed the position of each genomic fragment with respect to the AP3 coding region. Sequence data of the *BamHI* fragments, B1.6 and B2.0, when compared to the cDNA sequence, revealed the existence of two short introns separating the AP3 gene into two small exons at the 5' end of the gene and one large exon encoding the remainder of AP3. The first exon contains the first 25 base pairs (bp) of the cDNA and is separated from the second exon by 111 bp. The second exon, 94 bp in length, contains the translational start codon, AUG, and encodes the first 17 amino acids. The remainder of the AP3, 300 amino acids, is encoded by the third exon (1078 bp in length) containing the UAA stop codon. Exons two and three are separated by 74 bp. Figure 4B illustrates

Figure 4. Schematic diagram of the genomic map of the AP3 gene. A) Restriction endonuclease map of the genomic region encompassing AP3. (B = *Bam*HI, C = *Cla*I, H = *Hind*III, R = *Eco*RI) The position of the AP3 transcript (1.3kb) is indicated by an arrow above the exonic regions of AP3 (open boxes). A second transcript (1.1kb) lies near the 5' end of AP3. Also indicated is the 0.6kb transcript encoded by the *Dromsopa* gene. The translational start sites for AP3 and *Dromsopa* are indicated by AUG over arrows. The stop codon, UAA, for AP3 is also indicated by an arrow. The 2.3kb *Eco*RI genomic fragment used to clone the *Dromsopa* cDNA is at the far right side of the restriction endonuclease map. B) Map of the exonic regions of AP3. The upward pointing arrow indicates the region of four putative alternative translational start sites.





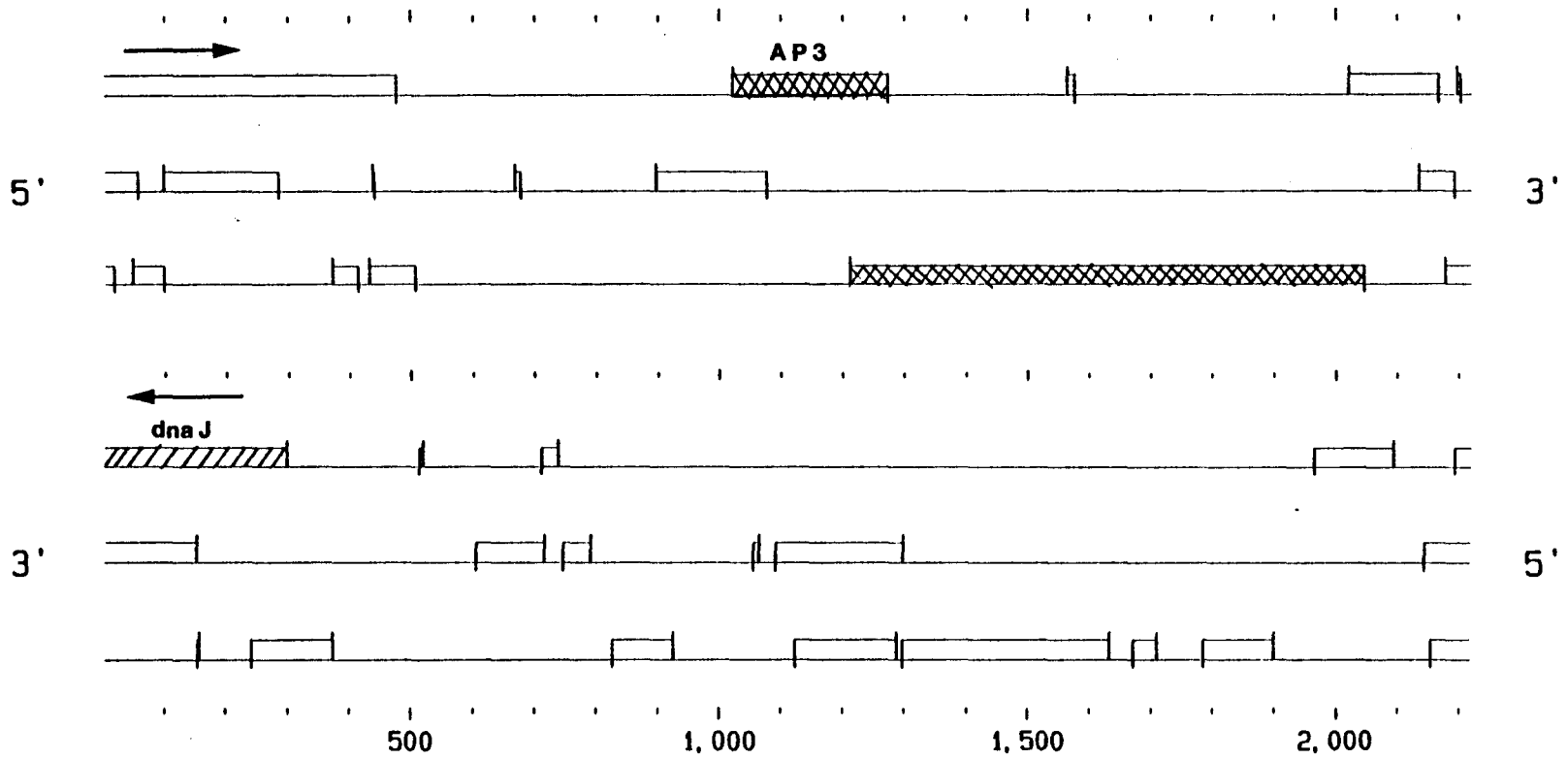
the exonic regions depicted by the open boxes. An upward pointing arrow in Figure 4B marks the position of possible secondary translational start sites approximately 175bp downstream of the 5' end of exon three. Nucleotide and protein sequence analyses were carried out using the University of Wisconsin Genetics Computer Group (UWGCG) software package (Devereux et al., 1984). Analyses of putative translational start sites found in the AP3 cDNA revealed the presence of four possible sites. Each of the four possible alternative translational start sites are in frame with the remainder of AP3 (Figure 5). The sizes of the encoded proteins for each methionine at positions 40, 55, 60 and 61 would be 29,700, 28,000, 27,500 and 27,400, respectively. These molecular weights are based on the amino acid composition from the deduced sequences. Methionine 60 is probably the only translational start site to be utilized as it is the only methionine of these four to fall within the consensus sequence for eukaryotic translation initiators (CCA/GCCATG) (Kozak, 1984).

Further analysis of the genomic sequences revealed the existence of two novel open reading frames found in a *Bam*HI-*Cla*I fragment (0.5kb) near the 5' end of AP3. These open reading frames are depicted as the open boxes at the left of Figure 6. This cartoon illustrates the presence of the open reading frames on both strands of DNA. The open reading frames for AP3 are represented by the double cross hatched

Figure 5. Nucleotide and amino acid sequence of AP3 cDNA. The authentic translational start site is at position 1. The other possible translational start sites are overlined and underlined. The possible products could be 29,700 D (M-40), 28,000 D (M-55), 27,500 D (M-60) and 27,400 D (M-61).

1 ATGGTTAGGGAGAACAAGGCAGCGTGAAGGCTCAGTACTTCATCAAGGTTGTGGAAGCTG 60  
 1 M V R E N K A A W K A Q Y F I K V V E L 20  
 61 TTCGATGAGTTCCCAAAGTCTTCATCGTGGGCGCCGACAACGTGGGCTCCAAGCAGATG 120  
 21 F D E F P K C F I V G A D N V G S K Q M 40  
 121 CAGAACATCCGTACCAGCCTGCGTGGACTGGCCGTCGTGCTTATGGGCAAGAACACCATG 180  
 41 Q N I R T S L R G L A V V L M G K N T M 60  
 181 ATGCGCAAGGCCATCCGCGGTCTCTGGAGAACAACCCGCGAGCTGGAGAAGCTGCTACCC 240  
 61 M R K A I R G H L E N N P Q L E K L L P 80  
 241 CACATCAAGGGCAACGTGGGATTCGTGTTACCAAGGGCGATCTCGCCGAGGTGCGCGAC 300  
 81 H I K G N V G F V F T K G D L A E V R D 100  
 301 AAGCTGCTGGAGTCCAAGGTGCGCGCCCCGCGCTCCCGGCGCTATTGCCCTCTGCAC 360  
 101 K L L E S K V R A P A R P G A I A P L H 120  
 361 GTCATCATCCCGGCGCAGAACACCGGCTGGGACCCGAGAAGACCAGTTTCTTCCAGGCC 420  
 121 V I I P A Q N T G L G P E K T S F F Q A 140  
 421 CTGTCCATCCCGACCAAAATTTCCAAGGGAACAATTGAAATCATCAACGATGTGCCCATC 480  
 141 L S I P T K I S K G T I E I I N D V P I 160  
 481 CTGAAGCCTGGCGACAAGGTGCGCGCCTCCGAGGCGACACTGCTCAACATGTTGAACATC 540  
 161 L K P G D K V G A S E A T L L N M L N I 180  
 541 TCGCCCTTCTCGTACGGTCTGATTGTCAACCAGGTCTACGACTCCGGCTCGATCTTTTCG 600  
 181 S P F S Y G L I V N Q V Y D S G S I F S 200  
 601 CCGGAGATCCTGGACATCAAGCCCGAGGATCTGCGCGCCAAGTTCCAACAGGGAGTGGCC 660  
 201 P E I L D I K P E D L R A K F Q Q G V A 220  
 661 AACTTGGCCGCCGTTTGTGTTGTCGGTGGGCTACCCACCATCGCCTCGGCCCGCACAGC 720  
 221 N L A A V C L S V G Y P T I A S A P H S 240  
 721 ATTGCCAACGGATTCAAGAATCTGCTGGCCATTGCTGCCACCACCGAGGTGGAGTTCAAG 780  
 241 I A N G F K N L L A I A A T T E V E F K 260  
 781 GAGGCGACCACCATCAAGGAGTACATCAAGGACCCCGAGTTCGCGCGAGCTGCTTCG 840  
 261 E A T T I K E Y I K D P S K F A A A A S 280  
 841 GCTTCGGCTGCCCGCGCGCGGAGCTACCGAGAAGAAGGAGGAGGCCAAGAAGCCC 900  
 281 A S A A P A A G G A T E K K E E A K K P 300  
 901 GAGTCCGAATCAGAGGAGGAGGACGATGATATGGGTTTCGGTCTGTTGACTAA  
 301 E S E S E E E D D D M G F G L F D \*

Figure 6. Plot of open reading frames found in the 2.2kb *Bam*HI genomic fragment. Hatch marks on top of the line indicate translational start sites. Hatchmarks below the line indicate translational stop sites. All three reading frames for both strands of DNA are shown. The cross hatched box represents the open reading frame that is homologous to the human homologue of the *E.coli* *dnaJ* protein. The double cross hatched boxes (top set of reading frames) represents the open reading frame for *AP3*. Splicing of the *AP3* transcript causes a shift in the open reading frame. The open ended box at the left of the upper reading frames (top line) represents an open reading frame on the opposite strand of the *dnaJ* open reading frame. The arrows indicate the direction of translation for the open reading frames.



boxes on the upper strand of DNA. The excision of the second intron during splicing of the AP3 transcript causes a frame shift in the open reading frame. Searches of databases using the GCG software identified a match to the nucleotide and deduced protein sequences from the lower strand of Figure 7. The match was to the human homologue of the *E.coli dnaJ* gene. There is a 67% identity on the nucleotide level and a 35% similarity on the amino acid level between the matched sequences. Further analysis of the matched sequences reveals that the open reading frame, possibly encoding the *Drosophila dnaJ* protein, is on the strand opposite to that of AP3 and therefore, is transcribed in the opposite direction of AP3 (Figure 7, pg. 69). This analysis would suggest that the promoter regions for the gene possibly encoding the *Drosophila dnaJ* protein and AP3 overlap or lie very close to one another. The *Bam*HI-*Cla*I genomic fragment (shown in Figure 7) was used to detect the presence of a transcript which would be transcribed from the gene encoding the homologous *dnaJ* open reading frame.

#### Mapping The Other Transcript Near AP3

The *Bam*HI-*Cla*I genomic fragment previously described (Figure 7), was used as a probe on a developmental northern blot of *Drosophila* RNA and detected two transcripts (Figure 8). One of these transcripts, 1.3kb, is transcribed from the AP3 gene. The other novel transcript, 1.1kb, is also detected by the *Bam*HI-*Cla*I fragment and is expressed in a

Figure 7. Nucleotide and amino acid sequences of the postulated promoter region of AP3. The amino acid sequence homologous to the human homologue to the *E.coli* dnaJ protein is shown under the lower strand (pg. 70). The first 17 amino acids of AP3 are shown above the top strand near the 3' end (pg. 71). The *Bam*HI and *Cla*I restriction sites are indicated (pg. 70). Arrows indicate the direction of translation.



## BamHI

5' GGATCCATCGCCGTGGGTAGCCTCTTTTTTCCGGTTGAAGAAGCTGGAATGTCTTGTAGGC  
 10 20 30 40 50 60  
 3' CCTAGGTAGCGGCACCCATCGGAGAAAAAGGCCAACTTCTTGACCTTACAGAACATCCG  
 S G D G H T A E K K R N F F Q F T K Y A

GGCGAACGCAAAGAGCACGGTTCCTCCGACGGCGCACATGAAGGGGAGCATGTCTGGCGT  
 70 80 90 100 110 120  
 CCGCTTGCCTTCTCGTGCCAAGGAGGCTGCCGCGTGTACTTCCCCTCGTACAGACCGCA  
 A F A F L V T G G V A C M F P L M D P T

GGGCTGGGCGAAGGTCGCAGCAGGCTCGTCATCACATTTGAGACCCTCCTCGCCGTGCTG  
 130 140 150 160 170 180  
 CCCGACCCGCTTCCAGCGTCGTCCGAGCAGTAGTGTAAGCTCTGGGAGGAGCGGCACGAC  
 P Q A F T A A P E D D C K L G E E G H Q

GTCGTATATCTCGCGCTTTTCTTATCAAAGAGCACTTCGAAGGCGGCCACCACCTCCCT  
 190 200 210 220 230 240  
 CAGCATATAGAGCGCGAAAAGGAATAGTTTCTCGTGAAGCTTCCGCGGTGGTGGAGGGA  
 D Y I E R K E K D F L V E F A A V V E R

AAACTGCTCCTCGGCCTGCGGATGGTCGTTCTTGTCCGGATGGTAGCGGAGAGCCATCCG  
 250 260 270 280 290 300  
 TTTGACGAGGAGCCGGACGCCTACCAGCAAGAACAGGCCTACCATCGCCTCTCGGTAGGC  
 F Q E E A Q P H D N K D P H Y R L A M R

GCGTATCCCTTCTTGACGTCTTCGCTGGACGCATTCCCTCTCGATGCCAGAATCTGTAGT  
 310 320 330 340 350 360  
 CGCATAGGGAAGAAGCTGCAGAAGCGACCTGCGTAAGGAGAGCTACGGGTCTTAGACATCA  
 R I G E Q R R R Q V C E E R H G S D T T

AATCCTTACCCATGTGTGAAATGTATTCTCCTGTTCTGGCTGAGATGGTCGTCTAGCGTG  
 370 380 390 400 410 420  
 TTAGGAATGGGTACACACTTTACATAAGAGGACAAGACCGACTCTACCAGCAGATCGCAC  
 I R V W T H F T N E Q E P Q S P R R A H

TTGCGCCGTGTATGGAATGTGAGGCAGCGTCTCGACTTGTGGTGCTTATGTGGAATGC  
 430 440 450 460 470 480  
 AACGCGGCACATACCTTACACTCCGTCGCAGAGCTGAACAACCAGGAATACACCATTACG  
 Q A T Y P I H P L T E V Q Q H K H P L A

GAGTATATATTGCTGATTGATTGCTGATTTAAACATTATTACAAATAAACAGTCAAT  
 490 500 510 520 530 540  
 CTCATATATAACGACTAAGCTAAGCGACTAAATTTTGTAAATAATGTTTATTTGTCAGTTA  
 L I Y Q Q N S E S I

## ClaI

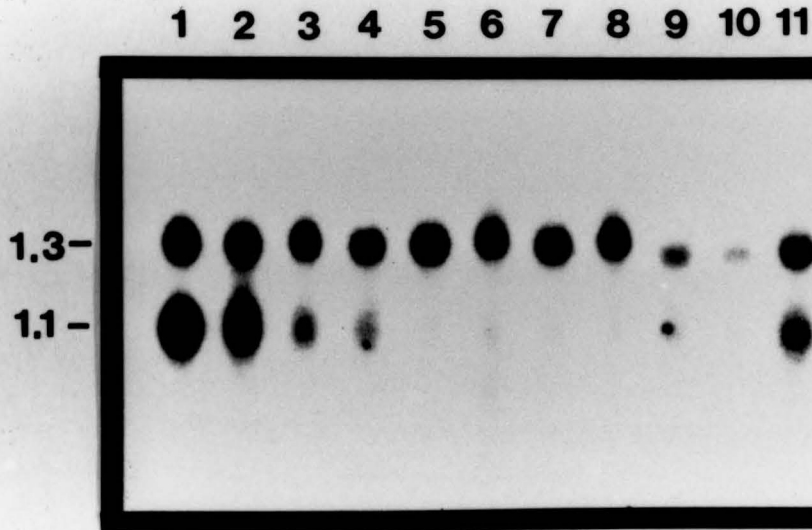
CGATAAATTATCGGCCGCCAGTGGCCAGAGTGACCCTACCTTACCACCAGAAGCACTGTT  
 550 560 570 580 590 600  
 GCTATTTAATAGCCGGCGGTACCCGGTCTCACTGGGATGGAATGGTGGTCTTCGTGACAA

AGTTAACAGAAGAGATTCTGACGATCGTAGCAGCTGTATAGGAGCACTATTAGCCTATTA  
 610 620 630 - 640 650 660  
 TCAATTGTCTTCTCTAAGACTGCTAGCATCGTCGACATATCCTCGTGATAATCGGATAAT  
  
 GTAATTAATGATCTGTTAAATCTATCTTTTTTGAGTACTTTAAAGGAATTACACATTTTCG  
 670 680 690 700 710 720  
 CATTAATTACTAGACAATTTAGATAGAAAAACTCATGAAATTTCTTAATGTGTAAAGC  
  
 AAAATAGAAAATATTCATATCGCTTACCGGTTTTGGTATTTTTTCAGACCGTGCGGTATC  
 730 740 750 760 770 780  
 TTTTATCTTTTATAAGTATAGCGAATGGCCAAAACCATAAAAAAGTCTGGCAGCCATAG  
  
 TTATTCGCCATCGAAGCGGTCACTGGGTGCCGCCGCCAACTTCACTCTTCCGTTCTG  
 790 800 810 820 830 840  
 AATAAGCGGTAGCTTCGCCAGTGTGACCCACGGCGCGGTTGAAGTGAGAAAGGCAAGAC  
  
 TGAGCGAAAACCGAAAAGTCTGTGCTTTGGTAAGTGTGCTAAAAGTTCGGAATAATGTT  
 850 860 870 880 890 900  
 ACTCGCTTTTGGCTTTTCAGACACGAAACCATTCAACGATTTTCAAGCCTTATTACAA  
  
 GCATCCCGAGCATTTTCGGGTACATAACTGTTCCACGGCGGTGGTCCAGCAAAGACTAAT  
 910 920 930 940 950 960  
 CGTAGGGCTCGTAAAAGCCCATGTATTGACAAGGTGCCGCCACCAGGTGTTTTCTGATTA  
  
 CGTTATCACGCCTTTTCGCAGTTCTTAAATTCACCCGACGAGTCCCTAATACACAATTA  
 970 980 990 1000 1010 1020  
 GCAATAGTGC GGAAAGCGTCAAGAATTTAAGTGGGCTGCTCAGGGATTATGTGTTAATTT

→  
 M V R E N K A A W K A Q Y F I K V 17  
 ATGGTTAGGGAGAACAAGGCAGCGTGAAGGCTCAGTACTTCATCAAGGTTGT 1073 3'

Figure 8. Developmental northern blot of total RNA from various *Drosophila* stages. Lanes: 1) 0-4 hr embryos; 2) 4-8 hr embryos; 3) 8-12 hr embryos; 4) 12-16 hr embryos; 5) 16-24 hr embryos; 6) 1st instar larvae; 7) 2nd instar larvae; 8) 3rd instar larvae; 9) pupae; 10) adult males; 11) adult females. The probe was a radiolabeled 0.5kb *Bam*HI-*Cla*I genomic fragment near the 5' end of AP3. (Figure 4). The 1.3kb band is the AP3 transcript. The 1.1kb band is presumably the *Drosophila dnaJ* transcript. Overexposure of the blot did not reveal any other transcripts. RNA loading was normalized using an actin gene probe as control (data not shown).

developmentally regulated fashion. The 1.1kb transcript is expressed only during early embryonic stages and in adult females. The gene most likely to encode the 1.1kb transcript is the *Drosophila* homologue to the human and *E.coli* *dnaJ* genes. This possibility is based on the premise that the molecular weight of the *Drosophila* *dnaJ* protein encoded by a gene expressing a 1.1kb transcript would be in close agreement



2.3kb genomic fragment (RI 2.3). The cDNA, named *Drosopa*, (*Drosophila* male specific *egg* containing) was subcloned into pBluescript SK<sup>+</sup> (Stratagene) and sequenced (Figure 9). The cDNA has an open reading frame of 69 amino acids and would encode for a protein of 8100 daltons. The initiation codon falls within the consensus sequence for eukaryotic translational initiators (CCA/GCCATG) [Kozak, 1981]. The 3' end of the cDNA

developmentally regulated fashion. The 1.1kb transcript is expressed only during early embryonic stages and in adult females. The gene most likely to encode the 1.1kb transcript is the *Drosophila* homologue to the human and *E.coli dnaJ* genes. This possibility is based on the premise that the molecular weight of the *Drosophila dnaJ* protein encoded by a gene expressing a 1.1kb transcript would be in close agreement with the molecular weight of the *E.coli dnaJ* protein (37,000 D, Bardwell et al., 1986). Isolation and characterization of a full length cDNA clone encoding the *Drosophila dnaJ* protein could verify the origin of the 1.1kb transcript.

#### Cloning and Characterization of Dromsopa

Another, developmentally regulated transcript was recognized by a 2.3kb *EcoRI* genomic fragment approximately 4.5kb downstream from the 3' end of AP3 (Figure 4A). This transcript is very small, 0.6kb, and is present only in adult males.

The cDNA encoding this 0.6kb transcript was cloned by screening a *Drosophila*  $\lambda$ gt10 library with a radiolabeled *EcoRI* 2.3kb genomic fragment (RI 2.3). The cDNA, named *Dromsopa*, (Drosophila male specific opa containing) was subcloned into pBluescript SK<sup>+</sup> (Stratagene) and sequenced (Figure 9). The cDNA has an open reading frame of 69 amino acids and would encode for a protein of 8100 daltons. The initiation codon falls within the consensus sequence for eukaryotic translation initiators (CCA/GCCATG) (Kozak, 1984). The 3' end of the cDNA

Figure 9. Nucleotide and predicted amino acid sequence of the *Dromsopa* cDNA. The cDNA was isolated from a  $\lambda$ gt10 cDNA library using a 2.3kb genomic fragment downstream of the AP3 gene (Figure 4). The poly (A) tail is not shown, but the two, AATAAA presumptive polyadenylation signals are underlined. The glutamine rich, *opa* (CAX) repeat is overlined. The molecular weight of the predicted protein is 8,100.

													Met Asn Phe	3			
1	CCC	CCC	CCC	CCC	CCC	CCC	CCC	CTT	CAT	ACC	ATA	CTC	GCC	ATG	AAC	TTC	48
4	Ile	Gln	Ile	Ala	Val	Ala	Val	Arg	Pro	Gly	Arg	Ser	Gly	Leu	Gly	Gln	19
49	ATA	CAG	ATC	GCC	GTG	GCT	GTT	CGT	CCT	GGT	CGC	AGT	GGC	CTT	GGC	CAG	96
20	Thr	Thr	Gly	Arg	Ser	Gly	Asn	Leu	Pro	Ala	Pro	Glu	Gln	Gln	Gln	His	35
97	ACC	ACA	GGA	AGA	TCC	GGC	AAT	CTG	CCA	GCT	CCA	GAG	CAG	CAG	CAG	CAC	144
36	His	Gln	Gln	Gln	Gln	Gln	His	Gln	Gln	Gln	Gln	Gln	Gln	His	His	Gln	51
145	CAC	CAG	CAG	CAG	CAG	CAG	CAC	CAG	CAG	CAG	CAG	CAG	CAG	CAC	CAC	CAG	192
52	His	His	Gln	His	His	Gln	Leu	Gln	His	Leu	Lys	Arg	Ser	Ser	Gly	Trp	67
193	CAC	CAC	CAG	CAC	CAC	CAG	CTG	CAG	CAC	CTC	AAG	CGC	TCC	AGC	GGG	TGG	240
68	Phe	Arg	***														
241	TTC	CGG	TAG	AAA	GAA	AAA	TGT	CAA	TCA	CAA	CGT	CAT	AAC	CAT	TGG	ATA	288
289	AAA	TTC	TTC	GAA	AGG	AAG	GAG	GCC	ACA	ATC	TTG	ATT	TAA	GAA	CCC	ATT	336
337	GGT	ACA	CCC	ATA	TTA	CAC	GCA	CTG	CGT	<u>TAA</u>	<u>TAA</u>	ATT	ATT	TTC	ATT	TGC	384
385	GTC	GCA	ATG	TGA	GGT	ATT	TAT	TTC	ACT	TCA	ATA	TAC	ACA	AGC	AGT	TTA	432
433	CTG	CTT	GTT	GTA	GAA	CTC	GCA	GAA	CAT	CAT	CTC	TCT	GAC	ACG	GAT	TTT	480
481	AGC	<u>TAA</u>	<u>ATA</u>	<u>AAG</u>	GGT	ATC	TGA	TAG	TCG	CGG	CAC	TCC					

contains a poly (A) tail. Two possible polyadenylation signals are present near the 3' end with one signal being 26bp upstream from the poly (A) tail. A unique feature of *Dromsopa* is the *opa* repeat found near the middle of the coding region. *Opa* repeats are defined by the sequence  $CAX_n$  where X is usually a G or C and encodes either glutamine or histidine and  $n =$  approximately 28. The *opa* repeat in *Dromsopa* consists of 26 CAX repeats. The cloned cDNA was used as a probe on a northern blot with RNA isolated from embryos, larvae and adults (Figure 10A) with a 0.6kb transcript detected in adults only. To further define the tissue specificity of *Dromsopa* expression, RNA was isolated from male adult abdomens, thoraxes and heads and used in a third northern blot (Figure 10B). A gender specific blot (Figure 10C) shows that *Dromsopa* is expressed only in adult males. Collectively, the results shown in Figure 10 reveals *Dromsopa* is expressed only in adult, male abdomens. Longer exposures of the northern blots did not reveal the presence of the 0.6kb transcript or any other transcript in any other developmental stage or tissue.

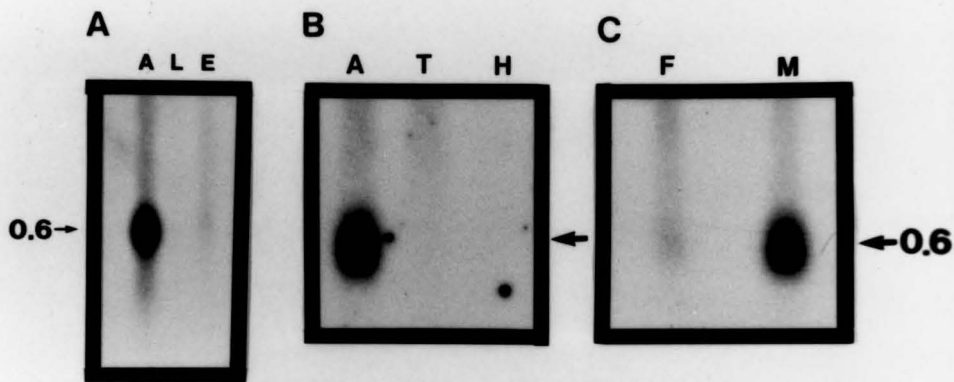
The tissue specific expression of *Dromsopa* was further defined from results of northern blot analysis using RNA from separate regions of the male reproductive system. The northern blot was borrowed from the laboratory of Dr. D. Cavener (Vanderbilt University), and probed with the *Dromsopa* cDNA . A 0.6kb transcript was weakly detected only in the anterior ejaculatory duct (data not shown). These results



Figure 10. Northern blot analysis of developmental, tissue, and sex specific RNA. (A) Lanes A, L and E contain RNA from adults(A), 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> instar larvae(L), and 0-20 hr embryos (E), respectively. (B) RNA isolated from adult males; lane A, abdomens; lane T, thoraxes and lane H, heads. (C) RNA isolated from whole adult females and males (lanes F and M, respectively). The probe was radiolabeled *Drosopa* cDNA. The size of the only detectable transcript (in kb) is indicated by the arrows. RNA loading was normalized using an actin gene probe as a control (data not shown).

indicates that *Drossops* may be expressed in an organ-specific manner.

Southern blot analysis using the *Drossops* cDNA as a probe revealed the existence of many *opa* containing genes (Figure 11). The Southern blot in Figure 11 was hybridized and washed under highly stringent conditions (see Materials and Methods) underscoring the extensive homology in the *opa* repeat regions, shared by different genes. On a separate Southern blot with *Dr*



repeats resulted in no matches. Therefore, *Drossops* is a novel, highly developmentally regulated, *opa* containing gene in *Drosophila*.

#### Cloning of P9 and Comparison to AP3

Since the AP3 cDNA was originally cloned by screening a *Drosophila*  $\lambda$ gt11 expression library with an antibody prepared against a purified human AP endonuclease (Holley et al., 1989; Kane and Linn, 1981), a *Helix* cDNA library was screened using

indicate that *Dromsopa* may be expressed in an organ-specific manner.

Southern blot analysis using the *Dromsopa* cDNA as a probe revealed the existence of many *opa* containing genes (Figure 11). The Southern blot in Figure 11 was hybridized and washed under highly stringent conditions (see Materials and Methods) underscoring the extensive homology in the *opa* repeat regions, shared by different genes. On a separate Southern blot with *Drosophila* genomic DNA digested with *EcoRI*, the main band recognized by *Dromsopa* was a 2.3kb fragment (data not shown). This fragment is presumably the same as the fragment used to isolate the *Dromsopa* cDNA. Sequence data from portions of the *EcoRI* 2.3kb genomic fragment and the cDNA indicate the direction of transcription of *Dromsopa* is equivalent to AP3.

Using the *Dromsopa* cDNA sequence in searches of the GenBank, EMBL and NBRF databases, the only sequences identified resulted from matches to the *opa* repeat. Searches of these databases with the cDNA sequence minus the *opa* repeats resulted in no matches. Therefore, *Dromsopa* is a novel, highly developmentally regulated, *opa* containing gene in *Drosophila*.

#### Cloning of P0 and Comparison to AP3

Since the AP3 cDNA was originally cloned by screening a *Drosophila*  $\lambda$ gt11 expression library with an antibody prepared against a purified human AP endonuclease (Kelley et al., 1989, Kane and Linn, 1981), a HeLa cDNA library was screened (using

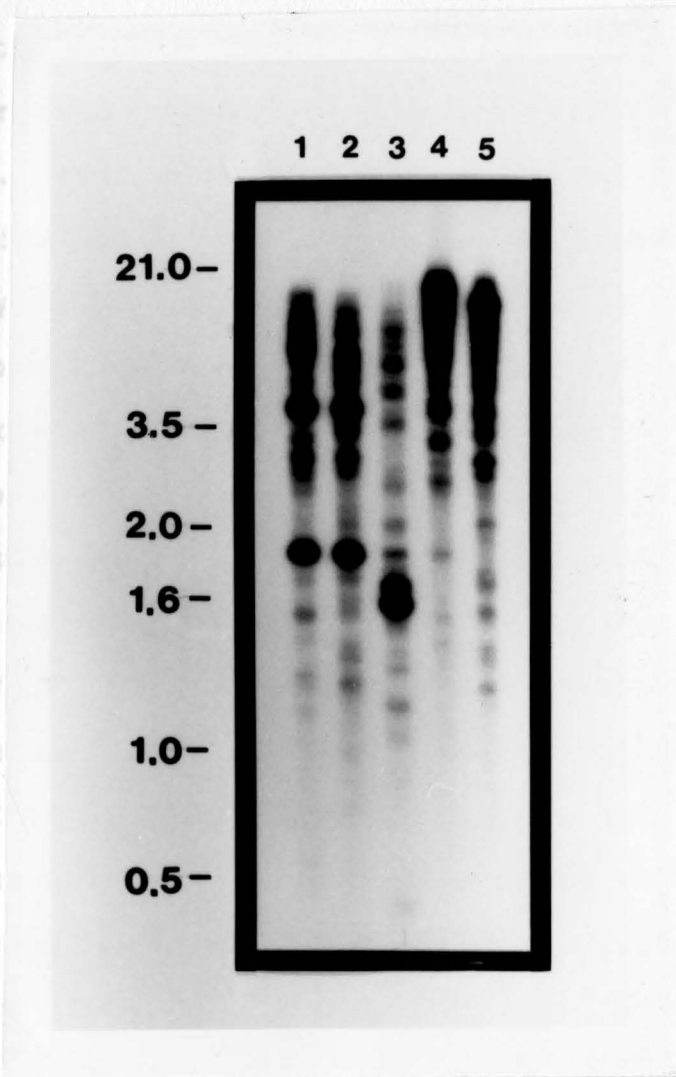
Figure 11. Southern blot of *Drosophila* genomic DNA digested with restriction endonucleases. Lanes: 1) *PvuI* and *HindIII*, 2) *ClaI* and *HindIII*, 3) *PstI*, 4) *PvuI* and 5) *ClaI*. Sizes of standard DNA markers, ( $\lambda$  DNA digested with *EcoRI* and *HindIII*) are positioned on the left. The probe was radiolabeled *Drosopa* cDNA.

the AP3 cDNA as a probe) for the human homologue to AP3. Two clones 2a and 1a were isolated and purified to homogeneity through successive screening. The cDNA inserts were isolated from the  $\lambda$  clones and subcloned into pBluescript SK<sup>+</sup>. Clone

2a was slightly longer in length, in fact (also 1.2kb). Southern analysis indicated further analysis

The comparison by dideoxynucleotide truncated subsequence techniques of the 1a cDNA with the human PO gene antisera from various sources (and Steitz, 1972). For this reason to be as PO. Comparison reveals a 68% (Figure 12)

identical in length, 317 amino acids, and contain a characteristic acidic carboxy tail seen in PO proteins from other organisms such as rat (Chan et al., 1988), mouse (Krowczynski et al., 1988) and yeast (AP; Mitoul et al., 1988) (Figure 13, Top). Another dramatic similarity between AP3 and



approximately 1.2kb the AP3 cDNA and subsequent analysis of 1a. All clones.

was determined by length cDNA and subsequent deletion analysis of 1a. Analysis of 1a is identical to the human PO gene cloned using antisera from various sources (Richman et al., 1988) (MW= 34,000) and can be referred to as PO. Comparison reveals a 68%

between the two proteins are identical in length, 317 amino acids, and contain a characteristic acidic carboxy tail seen in PO proteins from other organisms such as rat (Chan et al., 1988), mouse (Krowczynski et al., 1988) and yeast (AP; Mitoul et al., 1988) (Figure 13, Top). Another dramatic similarity between AP3 and

the AP3 cDNA as a probe) for the human homologue to AP3. Two clones 2a and 1e were isolated and purified to homogeneity through successive screening. The cDNA inserts were isolated from the  $\lambda$  clones and subcloned into pBluescript SK<sup>+</sup>. Clone 2a was slightly smaller than 1e which was approximately 1.2kb in length, in close agreement with the size of the AP3 cDNA (also 1.2kb). Both 2a and 1e were sequenced and subsequent analysis indicated 2a was a truncated version of 1e. All further analysis was carried out using the 1e clone.

The complete sequence of the 1.2kb 1e cDNA was determined by dideoxynucleotide sequencing of the full length cDNA and truncated subclones generated by ExoIII/ExoVII nested deletion techniques described in Materials and Methods. Analysis of the 1e cDNA sequence revealed that 1e was identical to the human PO gene. The PO gene was originally cloned using antisera from patients with systemic lupus erythematosus (Rich and Steitz, 1987), and encodes a ribosomal protein (MW= 34, 272). For the sake of clarity, our 1e clone will be referred to as PO. Comparison of PO and AP3 at the amino acid level reveals a 66% identity and 79% similarity between the two (Figure 12). Furthermore, the predicted proteins are identical in length, 317 amino acids, and contain a characteristic acidic carboxy tail seen in PO proteins from other organisms such as rat (Chan *et al.*, 1989), mouse (Krowczynska *et al.*, 1989) and yeast (AO; Mitsui *et al.*, 1989) (Figure 13, Top). Another dramatic similarity between AP3 and

Figure 12. Comparison of the deduced amino acid sequences of human PO, *Drosophila* AP3, and yeast AO proteins. PO and AP3 are identical in length; the AO protein has a gap near the carboxy terminus. Amino acid substitutions in the *Drosophila* and yeast sequences are indicated below the human PO sequence.

PO M P R E D R A T W K S N Y F L K I I Q L L D D Y P K C F I V G A D N V G  
 AP3 V N K A A Q I V V E F E F  
 AO - - M G G I R E K A E A L R E Y E E K S L F V V S

PO S K Q M Q Q I R M S L R G K A V V L M G K N T M M R K A I R G H L E N N  
 AP3 N T L  
 AO Q H E V K E R V R F S D L

PO P A L E K L L P H I R G N V G F V F T K E D L T E I R D M L L A N K V P  
 AP3 Q K A V K E S R  
 AO D F F V K Y N P K N V I V S R A

PO A A A R A G A I A P C E V T V P A Q N T G L G P E K T S F F Q A L G I T  
 AP3 P P L H I I S P  
 AO P V E D I W R V M E G V P

PO T K I S R G T I E I L S D V Q L I K T G D K V G A S E A T L L N M L N I  
 AP3 K I N P I L P  
 AO A V K V V D A N Q S L

PO S P F S F G L V I Q Q V F D N G S I Y N P E V L D I T E E T L H S R F L  
 AP3 Y I V N Y S F S I K P D R A K Q  
 AO T T V V Y Q V F P S S I D E V H V

PO E G V R N V A S V C L Q I G Y P T V A S V P H S I I N G Y K R V L A L S  
 AP3 Q A L A S V I A A F N L I A  
 AO S A S T I I S A L P G T L N D L V A

PO V E T D Y T F P L A E K V K A F L A D P S A F V A A A P V A A A T T A A  
 AP3 A T E V E K E T T I E Y I K K A S A S P A G  
 AO I A A S H Y E I D L V D R I E N E K Y A A T S A A S G

PO P A A A A A P A K V E A K E E S E E S D E D M G F G L F D  
 AP3 G T E K K E E A K K P E S E D  
 AO - - - D E E A A E D



Figure 13. Comparison of a family of acidic proteins at the amino acid level. The top shows the existence of the highly antigenic, highly acidic carboxy terminus. The middle comparison illustrates the existence of a potential 28s rRNA interaction region common to all five proteins. The numbers in parantheses indicates the first amino acid of the common regions. Amino acid substitutions in the Drosophila, rat, mouse or yeast sequences are indicated below the human PO sequence. The bottom comparison demonstrates the existence of the casein kinase I (S(P)-X<sub>2</sub>-S\*/T\*) and II (S\*/T\*-(D/E/S(P))<sub>1,2</sub>, X<sub>2,0</sub>) potential phosphorylation sites where the single letter amino acid abbreviations are used. The letter P in parentheses denotes a phosphoamino acid and an asterisk denotes the phosphoacceptor (Kennelly and Krebs, 1991).

## ACIDIC CARBOXY-TERMINAL AMINO ACIDS

HUMAN PO (302)--EESEESDEDMGFGLFD  
 DROS. AP3 (302)--S E D  
 RAT PO (302)--  
 MOUSE PO (302)--  
 YEAST AO (297)-- E D

## POTENTIAL 28s rRNA INTERACTION REGION

HUMAN PO (44)--RMSLRGKAVVLMGKNTMMRKAIR  
 DROS. AP3 (44)-- T L  
 RAT PO (44)--  
 MOUSE PO (44)--  
 YEAST AO (42)-- KE R V R

## CASEIN KINASE I OR II SITES

HUMAN PO EESEESDED  
 DROS. AP3 ESESEEDDD  
 RAT PO EESEESDED  
 MOUSE PO EESEESDED  
 YEAST AO EEEEEESDDD

human P0, as well as the other P0 proteins, is a region of 23 amino acids near the amino terminus which are presumed to be involved in the binding of P0 proteins to the ribosomes (Mitsui *et al.*, 1989). Figure 13 (middle) shows the comparison of this binding region found in the other P0 proteins to the similar region in *Drosophila* AP3. All the proteins appear to contain the 28S rRNA interaction region. They also contain well defined casein kinase I (CKI) and/or II (CKII) phosphorylation sites [CKI = S(P)-X<sub>2</sub>-S\*/T\* and CKII = S\*/T\*-(D/E/S(P))<sub>1,2</sub>, X<sub>2,0</sub>], Kennelly and Krebs, 1991 (Figure 13, Bottom)] which are most likely involved in regulating proper ribosomal protein function (Hasler *et al.*, 1991). Although this data will be discussed later, antibodies produced to fusion proteins of AP3 detected a single protein in the nuclear matrix, as well as one associated with ribosomes in the cytoplasm of *Drosophila* extracts (data not shown). Whether these proteins have dual function as DNA repair proteins and ribosomal proteins remains to be clarified.

#### Role of P0 as DNA Repair Protein

P0 may be involved in providing a protective role against damaged bases in DNA as suggested by the evidence of Grabowski *et al.* (1992). A summary of their results will be presented here. To determine if the expression of P0 was inducible, a human colon carcinoma cell line (Colo 320HSR) was exposed to three clinically useful DNA-damaging antitumor agents, melphalan (L-PAM), an active form of cyclophosphamide or 4-

hydroperoxycylcophosphamide (4HC), and mechlorethamine (HN2). Conditions were chosen that produced a 1 or 3 log kill of the Colo 320HSR cells (Futscher and Erickson, 1990). The effects on PO expression were measured by northern blot analysis (Figure 14). The results of densitometric analyses of the northern blots are shown in Figure 15. At a concentration producing a 1 log kill, each of the agents produced an initial increase in the steady-state levels of the PO transcript (Figures 14 and 15). For HN2, expression of PO was initially induced 1.3-fold immediately after treatment and fell below the control level at 6, 12 and 24 h post-treatment. L-PAM also induced PO but declined to control levels at 12 h post-treatment. 4HC induced PO expression greater than 2.0-fold and caused the level of expression to remain at this level beyond the 24 h time point (Figures 14 and 15). At concentrations producing a 3 log kill, however, all three compounds caused a slight decrease in the steady-state expression of PO, followed by a recovery exceeding control levels by 24 h (Figures 14 and 15). For HN2 and L-PAM at the 3 log kill dose the induction of PO expression by 24 h was 8-fold and 2.5-fold, respectively. Levels of PO expression affected by this dose of 4HC generally remained unchanged aside from the initial decline in expression.

In further support of PO involvement in DNA-repair, the levels of PO expression were found to be constitutively overexpressed in Mer<sup>-</sup> cells (Grabowski et al., 1992) which are

Figure 14. Northern blot analysis of the effects of L-PAM (a), 4HC (b) and HN2 (c) on P0 gene expression. Total RNA was isolated at various time points after chemical treatment. The left side of the panel represents 12.5 $\mu$ M L-PAM, 40 $\mu$ M 4HC and 2.1 $\mu$ M HN2 treatments (1 log kill, lanes 1-5). The right side (3 log kill, lanes 6-10) represents 25 $\mu$ M L-PAM, 80 $\mu$ M 4HC and 6.3 $\mu$ M HN2 levels. The lanes marked C are the untreated controls for each experiment. 0, 6, 12 and 24 are times in h post-treatment. (Drug treatment, RNA isolation and northern blot analyses were performed by B. Futscher).

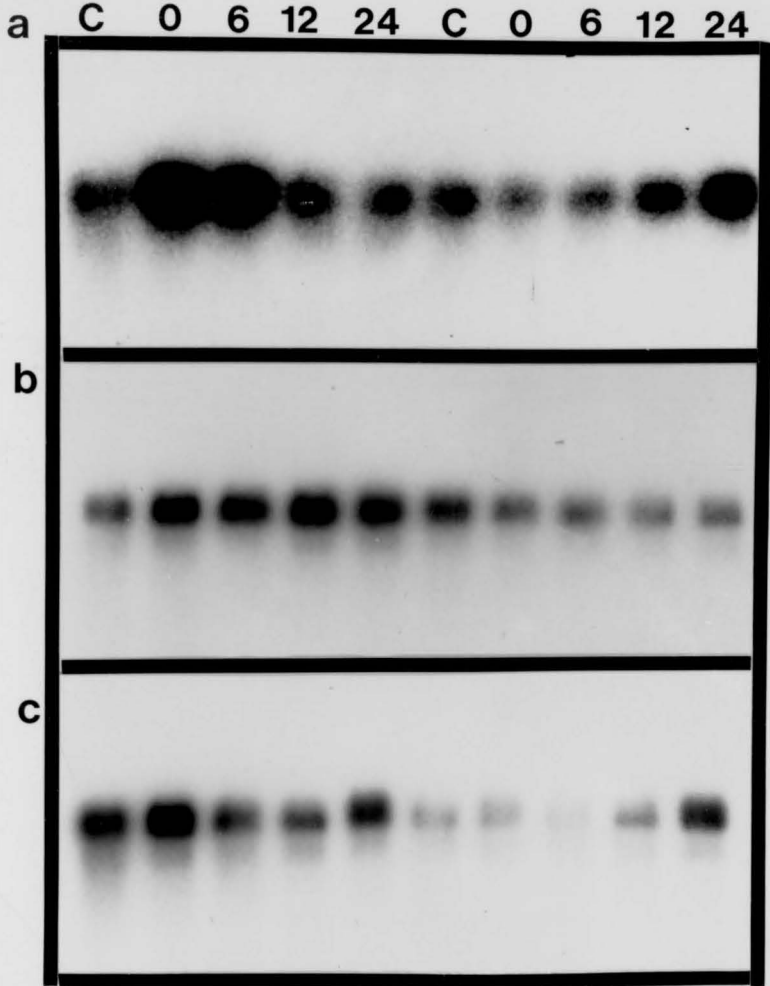
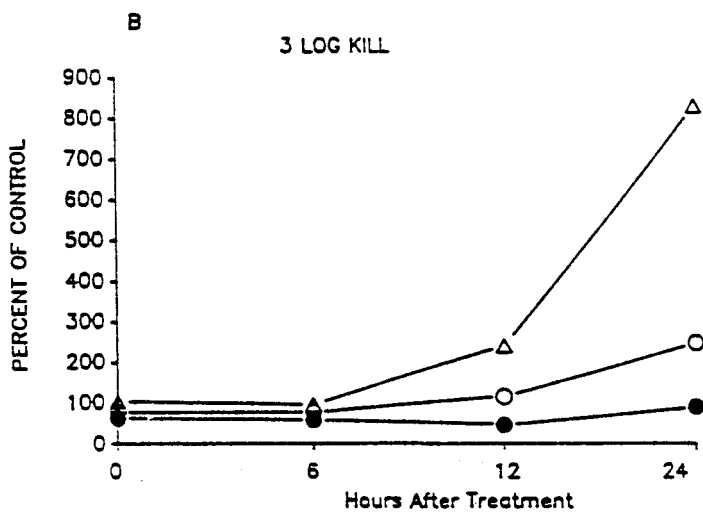
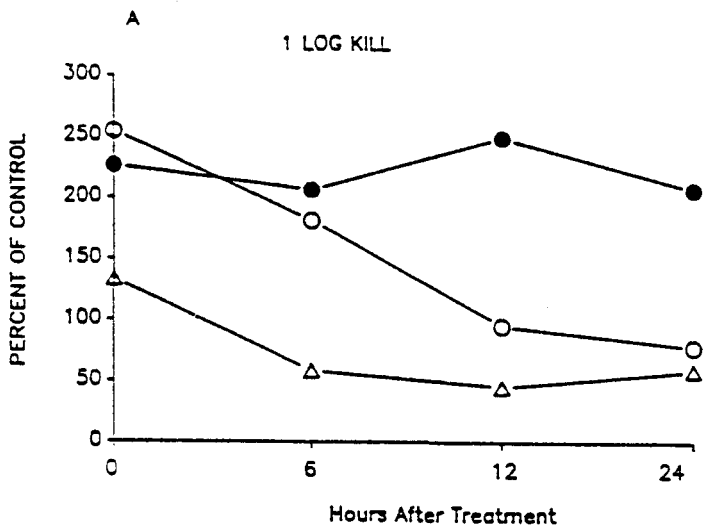


Figure  
in Fig  
and 2.  
(\*) a  
percent

shown  
C (●)  
or 48C  
as a

Figure 15. Densitometric analysis of the Northern blots shown in Figure 13. (A) 1 log kill or 12.5 $\mu$ M L-PAM (○) 40 $\mu$ M 4HC (●) and 2.1 $\mu$ M HN2 ( $\Delta$ ). (B) 3 log kill or 25 $\mu$ M L-PAM (○), 80 $\mu$ M 4HC (●) and 6.3 $\mu$ M HN2 ( $\Delta$ ). All values are represented as a percentage of the untreated control cells.





DNA repair defective for O<sup>6</sup>-alkylguanine lesions. Mer<sup>-</sup> cells are deficient in methylguanine methyl-transferase (MGMT). The expression of PO was measured in Mer<sup>-</sup> and Mer<sup>+</sup> cell lines by northern blot analysis (Figure 16). The MGMT activity profiles and the densitometric analysis of the northern blots are illustrated in Figure 17. All six Mer<sup>-</sup> cell lines have little or no detectable MGMT activity when assayed with protocols described by Pieper et al. (1990). The remaining Mer<sup>+</sup> cell lines all have some degree of MGMT activity. PO steady-state transcript levels were 30-to 50-fold higher in Mer<sup>-</sup> cells than the Mer<sup>+</sup> cells (Figures 16 and 17). Interestingly, three Mer<sup>+</sup> cell lines also had increased levels of PO expression. Two of these cell lines (W98VA and WI26VA4) were transformed by SV40 and showed an induced level of PO expression (Figures 16 and 17). MR1, a Mer<sup>-</sup> cell line (A1235) transfected with Mer<sup>+</sup> HT-29 DNA and selected for resistance to N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), did not have an altered level of PO expression compared to the parental A1235 Mer<sup>-</sup> cell line. Southern analysis demonstrated that the differences in PO expression were not due to gene copy number (data not shown).

The overexpression of the PO gene in Mer<sup>-</sup> cells is not a global phenomenon of ribosomal protein genes. The human ribosomal protein S17 cDNA clone (Chen and Roufa, 1988), was used as a probe on the Northern blots. There was no difference in the steady-state levels of the S17 transcript

Figure 16. Northern blot analysis showing constitutive level of P0 expression in Mer<sup>+</sup> or Mer<sup>-</sup> human tumor cell lines. Lanes 1-6 are the Mer<sup>-</sup> cell lines and lanes 7-13 are Mer<sup>+</sup>. Lane 1, A427; lane 2, A1235; lane 3, BE; lane 4, Colo 320HSR; lane 5, VA13; lane 6, HeLaMR; lane 7, HT-29; lane 8, W98VA1; lane 9, IMR90; lane 10, NAT; lane 11, MR1; lane 12 A2182; lane 13, W126VA4. The blot is slightly overexposed to show the lower level of P0 mRNA in lanes 7, 9, 10 and 12. Loading differences were corrected using the histone H3.3 probe. (Northern blot provided by R. Pieper.)

1 2 3 4 5 6 7 8 9 10 11 12 13

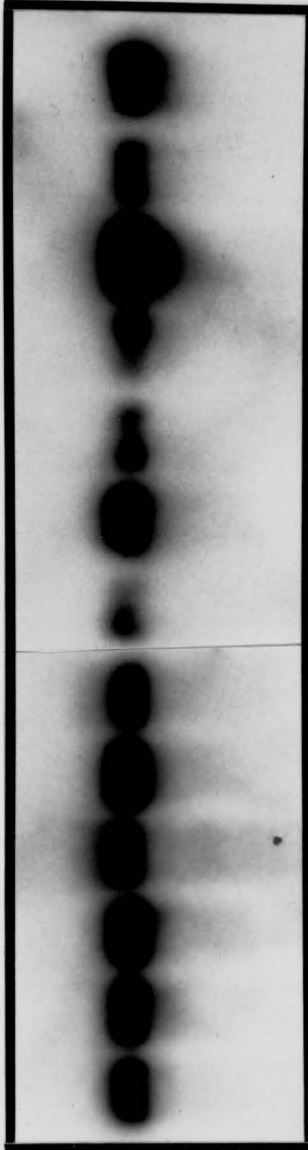


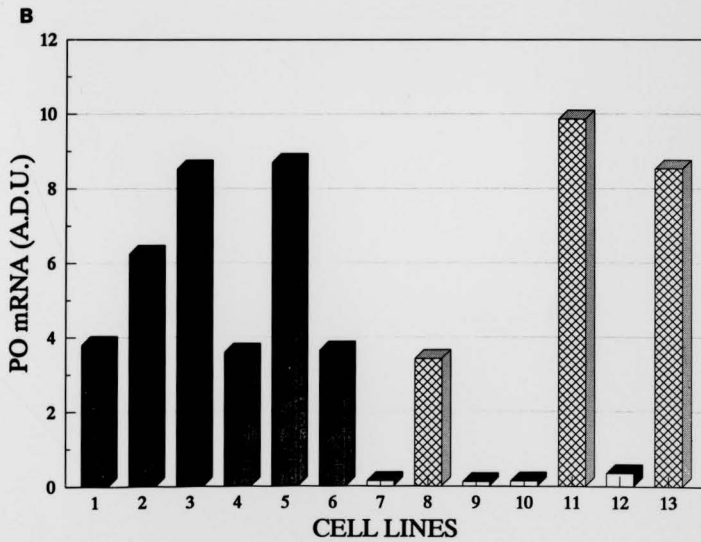
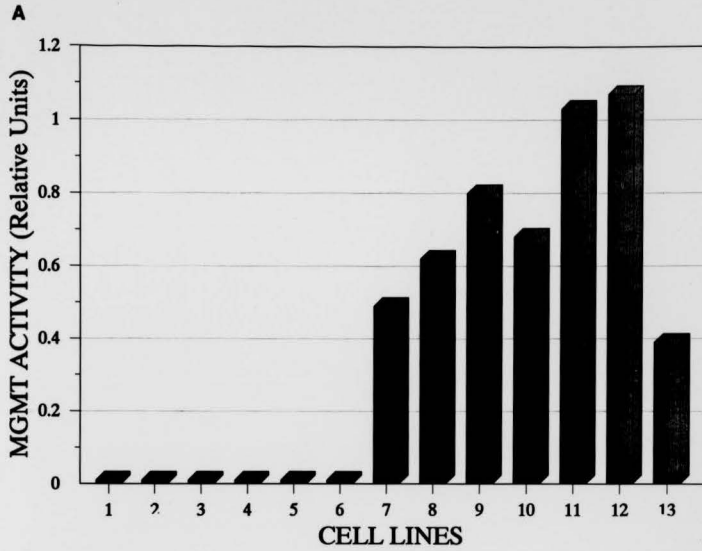
Figure  
North  
respe  
North  
Figure  
(ADU)  
are M  
lane  
lane  
lane  
(NGMT

the  
the  
the  
as in  
units  
s 1-6  
427;  
VH13/  
VH90/  
VH4.  
er.)

Figure 17. MGMT activity and densitometric analysis of the Northern blot shown in Figure 15. (A) MGMT activities for the respective cell lines. (B) Densitometric analysis of the Northern blot shown in Figure 15. The order is the same as in Figure 15 and data expressed as arbitrary densitometric units (ADU) after normalizing for the histone H3.3 gene. Lanes 1-6 are Mer<sup>-</sup> and Lanes 7-13 are Mer<sup>+</sup> cell lines. Lane 1, A427; lane 2, A1235; lane 3, BE; lane 4, Colo 320HSR; lane 5, VA13; lane 6, HeLa MR; lane 7, HT-29; lane 8, W98VA1; lane 9, IMR90; lane 10, NAT; lane 11, MR1; lane 12, A2182; lane 13, WI26VA4. (MGMT activity data provided by L. Erickson and R. Pieper.)

between Mar<sup>+</sup> or Mar<sup>-</sup> cell lines (data not shown).

### Production of Anti-AP3 Antibodies



proteins of 98,000 and 29,000 daltons (41 and 48, respectively) in extracts from the *E. coli* cells (data not shown). Western blot analysis of extracts of cells carrying

between Mer<sup>+</sup> or Mer<sup>-</sup> cell lines (data not shown).

### Production of Anti-AP3 Antibodies

To analyze AP3 on the biochemical level, antibodies directed against fusion proteins of AP3 were generated in rabbits. Initially, portions of AP3 cDNA were subcloned into vectors of the pWR590 series (gifts from Dr. Ray Wu, Cornell University). These vectors generate fusion proteins with truncated  $\beta$ -galactosidase ( $\beta$ -gal) at the amino-terminus contributing approximately 60,000 daltons to the expressed fusion. The additional  $\beta$ -gal portion of the fusion protein lends stability to otherwise rapidly degraded low molecular weight protein. Two  $\beta$ -gal-AP3 fusion protein constructs were generated. The top of Figure 18 illustrates the portions of AP3 that were subcloned into the appropriate pWR590 vector keeping AP3 in frame with the  $\beta$ -gal portion of the fusion. Construct d1 contains almost all of AP3 leaving out the first 13 amino acids. The second construct, d8, contains an internal portion of AP3 (amino acids 67-285). d8 was constructed in attempts to generate antibodies that did not recognize the highly antigenic acidic carboxy tail (amino acids 301-310) found in AP3 and other PO proteins (Rich and Steitz, 1987). Both constructs were overexpressed in *E.coli* cells, as determined by the presence of two novel inducible proteins of 98,000 and 88,000 daltons (d1 and d8, respectively) in extracts from the *E.coli* cells (data not shown). Western blot analysis of extracts of cells carrying

Figure 18. Schematic diagram of regions of AP3 subcloned into overexpression vectors. Open boxes indicate the portions of AP3 coding region subcloned; d1) aa 14-317; d8) aa 67-285. The angled arrows indicate the position of oligonucleotides used to PCR amplify AP3 for subcloning into the pGEX-3X vector (1 and 2) or pET11d vector (3 and 4). The sequences for each oligonucleotide is presented below the diagram.

14 317 d1

67 285 d8

1 317 AP3

1 ↘ →

← ↙ 2 pGEX-AP3

3 ↘ →

← ↙ 4 pET11d-AP3

BamHI

1) 5'-CGTGGGATCCCATGCCAGGG-3'

EcoRI

2) 5'-CGATGAATTCGGCAGATGGATC-3'

NcoI

3) 5'-TATACCATGGTTAGGGAG-3'

BamHI

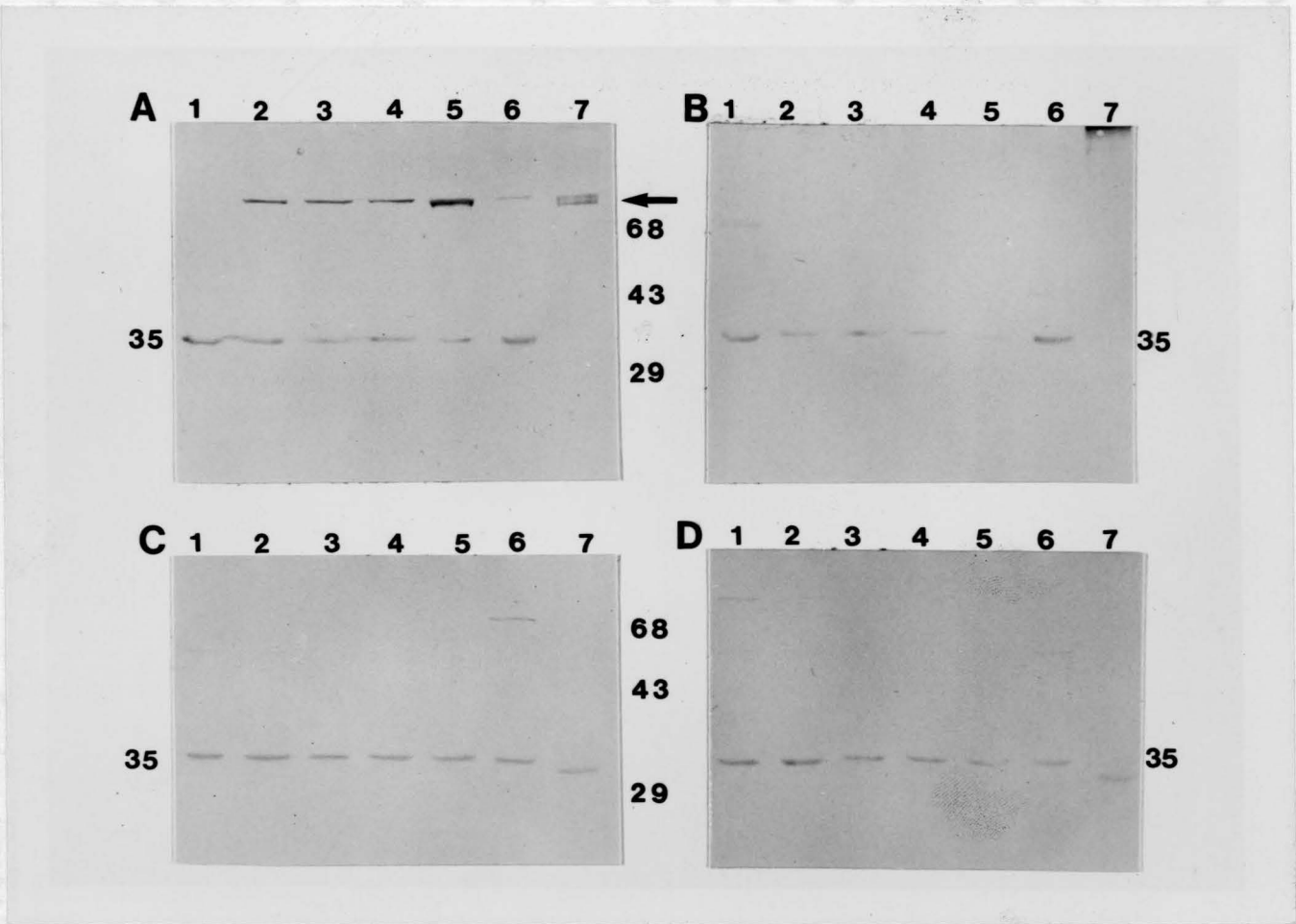
4) 5'-GGATCCAGCTTAGTCGAACAG-3'



the expression vectors containing d1 or d8 showed that both overexpressed constructs were recognized by the anti-human AP endonuclease antibody (Kane and Linn, 1981) originally used to clone the AP3 cDNA (data not shown). These results indicated that the overexpressed products of d1 and d8 were from the proper reading frame of AP3.

The overexpressed proteins were extracted from SDS-polyacrylamide gels and used as antigens to produce anti-d1 and anti-d8 antibodies in rabbits (see Methods and Materials). A third antibody, anti-Ud1, was generated against non-denatured d1 fusion protein. Dr. D. Derda purified the d1 fusion protein by passing extracts of cells overexpressing the fusion over an anti- $\beta$ -gal affinity column. The eluted, purified, non-denatured d1 fusion protein was used as antigen in a rabbit to generate the anti-Ud1 antibodies. The integrity and specificity of each antibody was determined by Western analysis. Figure 19 depicts a typical developmental Western blot where a 35,000 dalton protein is recognized by anti-d1 (panels A and C) and by anti-d8 (panels B and D) antibodies throughout development. It is not surprising to find a DNA repair protein associated with the nucleus as that is where DNA repair is expected to occur. The recognition of AP3 in the post nuclear supernatant was surprising at first, but these results are reconciled when the homology between AP3 and PO are included in the analysis. If AP3 has a dual function and is associated with the ribosomes, it stands to

Figure 19. Developmental Western blot analysis of purified nuclei (A and B) and post nuclear supernatant (C and D). Equal amounts of protein from developmental stages were run on a 10% SDS-polyacrylamide gel, electroblotted to nitrocellulose and assayed with anti-d1 (blots A and C) or anti-d8 (blots B and D) along with  $^{125}\text{I}$ -anti-IgG for detection. Lanes 1-7: 1) 0-16 hr embryos, 2) 16-24hr embryos, 3) 1st instar larvae, 4) 2nd instar larvae, 5) 3rd instar larvae, 6) pupae and 7) adults. The molecular weights corresponding to marker proteins are given in the middle of the blots. The position of AP3 is indicated by 35. Arrow indicates an unknown protein recognized by anti-d1 antibody.



reason that some of the protein would be recognized in fractions containing the ribosomes (i.e. post nuclear supernatant). The presence of a larger band (near 110kd) recognized by anti-d1 in the nuclei remains a mystery. Other than this larger band, anti-d1 and anti-d8 both recognized the same protein.

#### Overexpression of AP3 in *E.coli*

Because we desired full length overexpressed AP3, two other overexpression vectors were employed to generate non-truncated AP3. The overexpression vectors chosen for this task were pGEX-3X (Pharmacia) and pET-11d (Novagen). Each has its own unique features. pGEX-3X produces a fusion protein with glutathione S-transferase (GST, 26kd) at the amino end of the fusion with a Factor Xa cleavage site between the GST portion of the fusion protein and the protein of interest, AP3. The GST portion of the fusion will bind to a glutathione 4B sepharose column enabling the purification of the fusion from the remainder of cellular proteins. The protein of interest can then be released from the fusion using Factor Xa to cleave at the specific Xa site. A requirement for this system to function properly is that the overexpressed protein must be soluble.

The other vector, pET-11d allows for the overexpression of the protein of interest, AP3, as a non-fusion protein. A drawback of this system is the more difficult task of purifying the native protein.

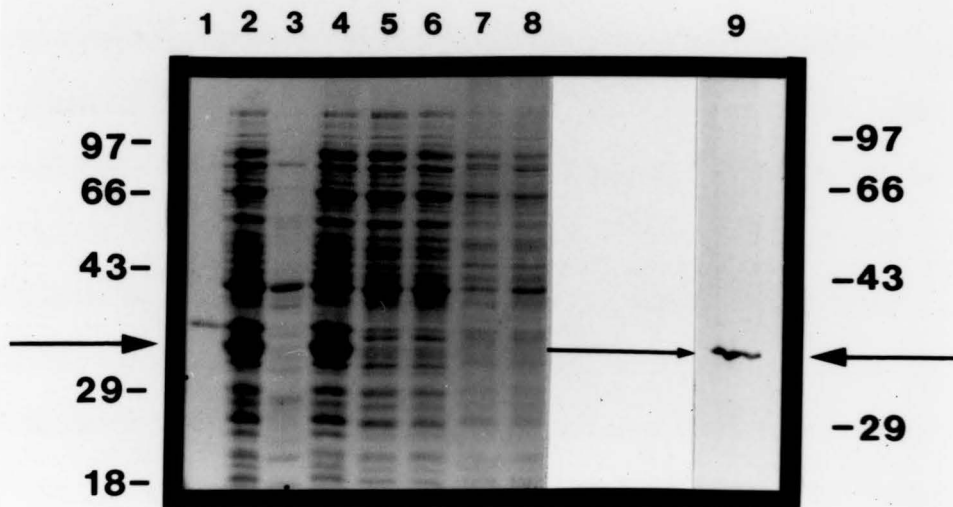
Oligonucleotides recognizing the nucleotide sequences of the amino terminal and carboxy terminal 4 amino acids of AP3, and containing appropriate restriction sites were obtained from National Biosciences (Minnesota). AP3 cDNA was PCR amplified using the appropriate pair of oligonucleotides (Figure 18) and digested with *Bam*HI and *Eco*RI for insertion into the pGEX-3X vector or with *Nco*I and *Bam*HI for insertion into the pET-11d vector. Overexpression of each protein was determined by coomassie blue staining of SDS-polyacrylamide gels containing extracts from cells carrying the expression vectors with insert (data not shown). The pGEX-AP3 vector generated an overexpressed protein of 61,000 daltons, the expected size for this fusion protein. Extracts of cells carrying the pGEX-3X vector only, generated an overexpressed protein of 26,000 dalton, the expected size of the GST domain (data not shown). The pET11d-AP3 vector, as expected generated an overexpressed protein of 35,000 daltons (Figure 20). The vector alone did not generate any observable overexpressed protein (data not shown). The extracts containing the overexpressed protein were separated by centrifugation into supernatant and pellet. This was done to determine if the overexpressed proteins were soluble (contained in the supernatant) or insoluble (trapped in the pellet). Many overexpressed proteins accumulate in the cytoplasm of host cells in the form of insoluble inclusion bodies that account for a major fraction of the total

Figure 20. A 10% SDS-polyacrylamide gel with aliquots from the inclusion body purification of pET11d-AP3. Lanes 1-8: 1 and 3) Supernatant from first wash of collected cells; 2 and 4) Crude extract containing pET11d-AP3; 5-8) supernatants from successive washes of inclusion bodies. The gel of lanes 1-8 was stained with coomassie blue. Lane 9) Silver stain of purified pET11d-AP3. The large arrows indicate the overexpressed (left) and purified (right) pET11d-AP3 protein. The molecular weights corresponding to marker proteins are given at the left and right side of each gel.

expressed proteins (Kane and Hartley, 1988). Both pGEX-AP3 and pET11d-AP3 produced insoluble protein. Although this was first considered a drawback, the purification of inclusion bodies leads to fewer contaminants which facilitates the purification of the overexpressed proteins.

### Purification of AP3

Both overexpressed proteins were subsequently purified



of this, which has been detected by silver staining techniques. This lane clearly demonstrates that the purified AP3 protein is greater than 95% pure. This protein can now be utilized in various biochemical analyses of AP3.

### Subcellular Localization of AP3

Some of the purified protein from the inclusion body preparations were used as antigens in rabbits to generate new anti-AP3 antibodies (for both pGEX-AP3 and pET11d-AP3). These antibodies along with the previously generated anti-d1 and

expressed proteins (Kane and Hartley, 1988). Both pGEX-AP3 and pET11d-AP3 produced insoluble proteins. Although this was first considered a drawback, the purification of inclusion bodies leads to fewer subsequent steps for purifying the overexpressed proteins.

### Purification of AP3

Both overexpressed proteins were subsequently purified from inclusion bodies (see Materials and Methods). The purified protein from inclusion bodies can be greater than 95% pure. Figure 20 shows the results of the various steps of purification. In lanes 1 thru 8, aliquots from each step of purification of pET11d-AP3 were separated on SDS-polyacrylamide gels and stained with coomassie blue. The arrow indicates the overexpressed AP3 protein. This protein is absent from lanes 1, 3, and 5-8 because these lanes contain solubilized proteins that have been washed from the inclusion bodies. Lane 9 shows an aliquot of purified pET11d-AP3 (10 $\mu$ l of 8mls) which has been detected by silver staining techniques. This lane clearly demonstrates that the pET11d-AP3 protein is greater than 95% pure. This protein can now be utilized in various biochemical analyses of AP3.

### Subcellular Localization of AP3

Some of the purified protein from the inclusion body preparations were used as antigens in rabbits to generate new anti-AP3 antibodies (for both pGEX-AP3 and pET11d-AP3). These antibodies along with the previously generated anti-d1 and



anti-d8 can eventually be affinity purified using purified AP3 bound to nitrocellulose strips (see Materials and Methods). Anti-Ud1 antibodies were immunoselected by this method generating purified anti-AP3 antibody. This purified antibody was subsequently used in Western blot analysis attempting to further isolate the subcellular localization of AP3. Figure 21 shows a typical Western blot of proteins from three separate origins (adult, embryo, and embryonic tissue culture) separated into purified nuclei (N) or post-nuclear supernatant (P). As expected the purified anti-Ud1 recognizes a 35,000 dalton protein both in the nuclei and post nuclear supernatant. A subtle difference in relative amounts of AP3 can be seen in the adult and embryo lanes. These two stages appear to have twice as much AP3 found in the post nuclear supernatant compared to the nucleus. The Schneider cells, however, appear to have near equal amounts of AP3 in both locations. An additional protein was seen in the nuclei of the Schneider tissue culture cells. The origin of this protein is unknown, however, this may be the product of the alternate translational start site found in the AP3 sequence (See above; Figure 5). A more intense analysis of the association of AP3 with the nucleus was undertaken to determine how tightly AP3 is bound to the nucleus. Using the purified anti-Ud1 antibody in Western analysis of purified nuclei treated with detergents and high salt washes, we see that AP3 is tightly associated with the nuclear matrix (Figure

Figure 21. Developmental Western blot of purified nuclei (N) and post nuclear supernatant (P). Equal amounts of protein from three different sources, [A adults; E) embryos; S) Schneider II tissue culture cells] were run on a 10% SDS-polyacrylamide gel, electroblotted to nitrocellulose and assayed with purified anti-AP3 antibody along with <sup>125</sup>I-anti IgG for detection. The molecular weights corresponding to marker proteins are given at the right side of the blot.

22). The apparent decrease in amount of AP3 in the last three lanes (T, S, and N Figure 22) is due to underloading. The purified anti-Udi did recognize large amounts of AP3 in the filtered crude homogenate, post-nuclear supernatant, and purified nuclei of adult *Drosophila* (Figures 22, Lanes C, P and N respectively).

Western analyses of the post nuclear supernatant, carried out in

association

ribosomes

35,000

shown)

surpris

*Drosoph*

eukary

not

phosph

kinase II

(Figure 13, bottom). Phosphorylation could affect

the enzymatic activity of AP3 or the localization of AP3. AP3

has been shown to be phosphorylated *in vitro* by both casein

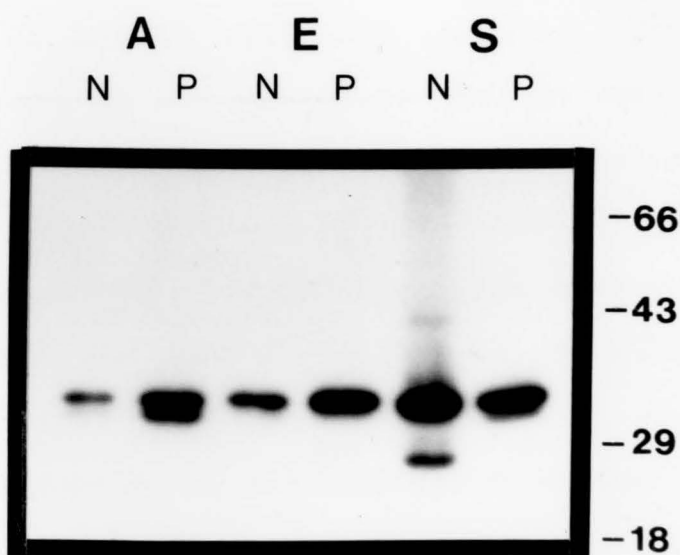
kinase I and casein kinase II but not by protein kinase A or

protein kinase C (data not shown). Using purified pET11d-AP3

protein, Dr. M. Tao (U.I.C.) *in vitro* phosphorylated the

protein using casein kinase I and casein kinase II. The

effect of this phosphorylation on *in vivo* localization is yet



22). The apparent decrease in amount of AP3 in the last three lanes (T,S, and M Figure 22) is due to underloading. The purified anti-Udl did recognize large amounts of AP3 in the filtered crude homogenate, post-nuclear supernatant, and purified nuclei of adult *Drosophila* (Figure 22, Lanes C, P and N respectively).

Western analyses of the post nuclear supernatant, carried out in our laboratory by Dr. D. Derda, revealed a close association of AP3 to partially purified mitochondria and ribosomes. Anti-d1 and anti-d8 antibodies both recognized a 35,000 dalton protein in both of these fractions (data not shown). Finding AP3 associated with ribosomes is not surprising considering the extensive similarities between the *Drosophila* AP3 and the ribosomal P0 proteins from other eukaryotes.

The localization of AP3 could be dependent on whether or not it is phosphorylated. AP3 contains potential phosphorylation sites recognized by casein kinase I and casein kinase II (Figure 13, Bottom). Phosphorylation could effect the enzymatic activity of AP3 or the localization of AP3. AP3 has been shown to be phosphorylated *in vitro* by both casein kinase I and casein kinase II but not by protein kinase A or protein kinase C (data not shown). Using purified pET11d-AP3 protein, Dr. M. Tao (U.I.C.) *in vitro* phosphorylated the protein using casein kinase I and casein kinase II. The effect of this phosphorylation on *in vivo* localization is yet

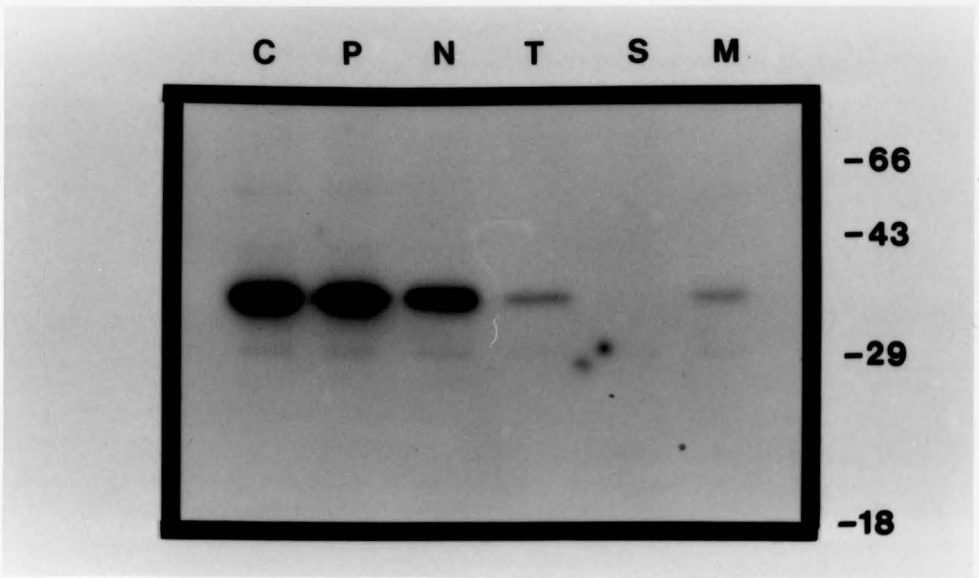
Figure 22. Western blot analysis of subcellular and subnuclear components of adult *Drosophila*. Proteins from crude filtered homogenate (C), post nuclear supernatant (P), purified nuclei (N), nuclei treated with Triton X-100 (T), 1M NaCl (S), and nuclear matrix (M) fractions were run on 10% SDS-polyacrylamide gels, electroblotted to nitrocellulose and assayed with purified anti-Udl antibody. The decreased level of AP3 detected in the Triton X-100 and nuclear matrix lanes (T and M, respectively) are due to lesser amounts of total protein loaded in these lanes. The molecular weights corresponding to marker proteins are given at the right side of the blot.

to be determined.

### AP Endonuclease Activity of AP3

The activity of AP3 as an AP endonuclease was clearly demonstrated in our laboratory (J. Carey, personal communication). This was accomplished using purified AP3 in an assay which clearly distinguishes between Class I and Class II AP endonuclease activity. The distinction is

to  
of  
T  
(  
I  
e  
P  
R  
t  
t



treatment of alkali and heat. The 3' radiolabel was released only through the action of a 3'-diesterase such as endonuclease IV. Therefore, the Class of an endonuclease can be determined as either Class I or Class II depending on the treatment of the reaction product to release the radiolabel.

Using the method described above, the presence of AP3 clearly demonstrated a Class II activity. This activity was further demonstrated in

to be determined.

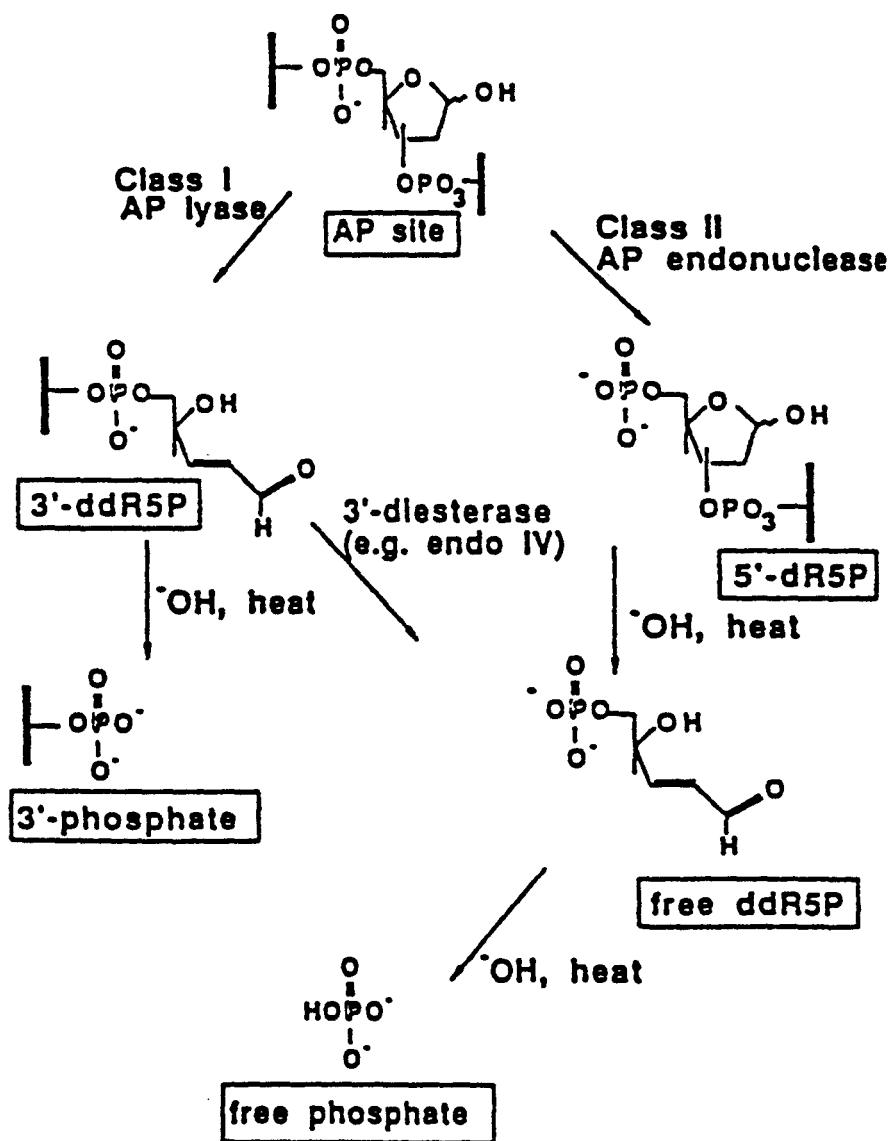
### AP Endonuclease Activity of AP3

The activity of AP3 as an AP endonuclease was recently demonstrated in our laboratory (J. Carney, personal communication). This was accomplished using purified pET11d-AP3 in an assay which clearly distinguishes between Class I and Class II AP endonuclease activity. The distinction of the two classes depends on the use of a synthetic polymer that contains AP sites with 5'-radiolabeled phosphate residues. The assay is based on a method developed by Clements *et al.* (1978) and modified by Levin and Demple (1990), as diagrammed in Figure 23. A Class I AP endonuclease would catalyze the  $\beta$ -elimination reaction at the AP site leaving the radiolabeled phosphate attached to the 3'-end of DNA in the form of a 3'-ddR5P. A Class II AP endonuclease, cleaving hydrolytically 5' to the AP site, would produce a radiolabeled 5' end of DNA in the form of 5'-dR5P. This label can be released by subsequent treatment of alkali and heat. The 3' radiolabel can be released only through the action of a 3'-diesterase such as endonuclease IV. Therefore, the Class of an unknown AP endonuclease can be determined as either Class I or Class II depending on the treatment of the reaction product needed to release the radiolabel.

Using the method described above, the pET11d-AP3 protein clearly demonstrated a Class II activity (data not shown). This activity was further demonstrated to be magnesium-

Figure 23. Principle of class I and class II AP endonuclease assays. The  $^{32}\text{P}$ -labelled phosphate is on the 5' side of the AP site. Cleavage of AP sites by class II AP endonucleases follows the right-hand pathway, the action of class I AP lyases the left-hand pathway. The products expected from secondary reactions are shown. (From Levin and Demple, 1990).





dependent (J. Carney, personal communication). However, the parameters for optimal endonuclease activity have not yet been established. The possible effect of phosphorylation on endonuclease activity was also examined. Using *in vitro* phosphorylated pET11d-AP3 (from Dr. M. Tao, University of Illinois, Chicago) in the Class I/Class II activity assay, no significant change was detected in enzyme activity. A change may be detected when the parameters for optimal activity of AP3 are established.

## CHAPTER IV

### DISCUSSION

The subject of this dissertation has been the molecular and biochemical analysis of AP3, a DNA repair enzyme in *Drosophila melanogaster* and its surrounding genomic region. The original cloning of the AP3 cDNA and subsequent analysis of the *in vitro* translated product by Kelley et al. (1989) demonstrated compelling evidence that AP3 encodes an AP endonuclease. The cDNA was cloned by screening a *Drosophila*  $\lambda$ gt11 expression library using an antibody directed against a purified human AP endonuclease (Kelley et al., 1989, Kane and Linn, 1981). The antibody also recognized the *in vitro* synthesized product of the AP3 cDNA, which is identical in molecular weight to a protein with AP endonuclease activity recovered from *Drosophila* embryonic extract. Although this evidence is quite compelling, more direct evidence establishing a definitive function for AP3 was sought. Molecular analysis would offer a better understanding of the regulation of the AP3 gene while biochemical analysis would support and/or define the role of the AP3 protein.

Genomic clones of *Drosophila* DNA were isolated from a Maniatis library using the AP3 cDNA as a probe. Sequence data and Southern blot analysis of the isolated genomic clones

established the composition of the genomic map of AP3 (Figure 4). At least three transcripts were discovered to be generated by three separate genes residing in this region. One transcript, from AP3, is 1.3kb in length and encodes a protein of 35,000 daltons. The two other transcripts, 1.1kb and 0.6kb, were detected by northern analysis using genomic fragments from regions that did not hybridize to AP3. The possibility that either of these two transcripts could affect the expression of AP3 made it necessary to further characterize the transcripts.

#### Cloning and Characterization of *Dromsopa*

The smaller of the two transcripts found to be expressed near AP3, was detected by a 2.3kb *EcoRI* genomic fragment which lies approximately 4.0kb downstream from AP3 (Figure 4A). The cDNA expressing the 0.6kb transcript was cloned and named *Dromsopa* (for *Drosophila*, male specific, opa containing gene). The only stage of *Drosophila* in which *Dromsopa* is expressed is the adult male (Figure 10). Overexposure of the blot did not reveal any other transcripts. The expression is further restricted to the anterior ejaculatory duct of the male reproductive system (data not shown). This highly regulated expression of *Dromsopa* indicates that its product is involved in a highly specific function which is yet to be determined. This function might be an interaction with other proteins which regulate the expression of other organ specific genes. Generating mutants deficient in *Dromsopa* would begin to define

any interactive function it may have with other genes in the male reproductive system. Since *Dromsopa* is expressed in such a developmentally and organ specific manner, it is unlikely that *Dromsopa* plays any role in the regulation of *AP3* expression.

From the sequence analysis of the *Dromsopa* cDNA, the computer generated translation results in a protein of 8,104 daltons. This was generated based on the predicted amino acid sequence of the longest open reading frame (Figure 9). Whether the transcript is actually translated is yet to be determined. However, a translational start site, ATG, falls within the consensus sequence for eukaryotic translation initiators CCA/GCCATG (Kozak, M 1984). The most interesting aspect of the coding region is the existence of a poly-glutamine stretch encoded by a previously characterized repeat, *opa*. The *opa* repetitive element was initially identified in the Notch gene of *Drosophila* (Kidd et al., 1983; Kidd et al., 1986, Wharton et al., 1985). The *opa* repeat has subsequently been identified in a number of genes in various organisms from *Drosophila* to humans. *Opa* repeats are defined by the sequence  $CAX_n$  where X is usually a G, C or A and encodes either glutamine (Gln: CAG and CAA) or histidine (His:CAC) and n is usually less than 30 (Steward, 1987). *Dromsopa* contains 26 CAX repeats. This pattern codes for a poly(glutamine) stretch of amino acids (typically interrupted with histidine residues) known to be associated with homeobox

genes in *Drosophila*. Although *opa* is a short repetitive element, it occurs in a number of gene products and is apparently translated. These genes include, but are not limited to, the *Drosophila* gene *Antennapedia* (Laughon *et al.*, 1986), *cut* (Blochinger *et al.*, 1988), *dorsal* (Steward, 1987) and the previously mentioned *Notch*. In mammals, the human androgen receptor gene (Lubahn *et al.*, 1988), the rat ventral prostate glucocorticoid receptor gene (Chang *et al.*, 1987; Danielson *et al.*, 1987) and the *Mopa* genes in mouse (Duboule, *et al.*, 1987) all contain the *opa* element. Previously identified *opa* containing genes have been shown to be involved in cell-type-specific functions in vertebrates and can be found in mRNAs transcribed in an adult tissue-specific manner (Duboule *et al.*, 1987). It has been suggested that the *opa* repeats encode linking or spacing units between separate domains of a protein (Laughon *et al.*, 1985). Therefore, a possible function for *Drosophila* would be as a small regulatory factor needed to assist in the positioning of multiple protein domains. The amino- and carboxyl-termini of *Drosophila* may interact with two separate protein domains of a complex linking them through the *opa* repeat region. This tethering could decrease or block an otherwise normal freedom of movement between the separate domains altering the function of the protein complex. The expanded linking regions may allow aberrant complexes to form with diminished or deficient

function. Recently, the expansion of *opa* repeats has been associated with human genetic diseases such as Fragile-X syndrome (Fu, et al., 1991), Kennedy's disease (LaSpada et al., 1991) and myotonic dystrophy (Harley et al., 1992). The actual mechanisms involved in establishing the disease states due to the expanded *opa* repeats is yet to be determined. However, one can postulate that the expanded repeats would alter the linking or spacing between the domains of the proteins encoded by affected genes in the above diseases.

#### The 1.1kb Transcript.

The other transcript found to reside near *AP3* is also expressed in a developmentally regulated fashion. The 1.1kb transcript (Figure 4A) was detected on northern blot analysis by a genomic fragment *Bam*HI-*Cla*I which lies immediately 5' to the *AP3* gene. This *Bam*HI-*Cla*I fragment was chosen as a probe for two reasons. First, this fragment was used to assist in defining the transcriptional start site for *AP3*. Sequence data of this portion of the genomic map revealed that it was not part of the *AP3* cDNA but resided close enough to the 5' end of the cDNA to perhaps recognize the 5' untranslated region of the *AP3* transcript. From the northern blot showing the detection of the 1.1kb transcript (Figure 8) we see that *Bam*HI-*Cla*I fragment also recognized the 1.3kb *AP3* transcript.

The second reason for using the *Bam*HI-*Cla*I fragment stems from computer analysis of the sequence of this fragment (GCG

software, University of Wisconsin; Devereux, et al., 1984). This analysis revealed the existence of two open reading frames almost entirely throughout the *Bam*HI-*Cla*I fragment (Figure 6).

A search of the Genbank database using nucleotide sequence from the genomic region (nucleotides 1-850 Figure 7), found a match with the mRNA sequence for the *dnaJ* gene from *E.coli* and the recently cloned *dnaJ* human homologue (Raabe and Manley, 1991). This match had a 67% identity over approximately 230 bp in the 5' upstream region of AP3. Also, analysis of the match predicts that the transcript encoding the protein is transcribed away from AP3. That is, the promoter region of both genes appear to overlap. The human *dnaJ* protein is the human homologue to the *E.coli* *dnaJ* protein. In *E.coli* the *dnaJ* protein along with two other proteins (*dnaK* and *grpE*) function in a variety of cellular processes including protein folding (Gaitanaris et al., 1990) proteolysis (Strauss et al. 1988), and also replication of  $\lambda$  (Zylicz et al., 1985) and P1 phage (Wickner, 1990). All three proteins are from the family of heat shock proteins known to be involved in protein folding, transport and assembly. During  $\lambda$  replication the *dnaJ* and *dnaK* proteins dissociate a protein complex at the origin of replication allowing for the proper initiation of replication. In yeast, *dnaJ* and *dnaK* along with other heat shock proteins act as "chaperones" by



binding to proteins while they are being transported to a proper site of assembly (Pelham, 1988). The sequence of the mRNA for the human *dnaJ* protein is unpublished, found only in the database. The human homologue appears to be preferentially expressed in neurones and was cloned from a lambda gt11 cDNA library of Alzheimer frontal brain tissue (Cheetham *et al.*, 1991). The *E.coli dnaJ* gene has been cloned and shown to encode a protein of  $M_r$  37,000 (Bardwell *et al.*, 1986). If a protein were encoded by the 1.1kb transcript, its size would be close to that reported for the *E.coli dnaJ* protein. The cloning of the full length cDNA for the *Drosophila dnaJ* gene will greatly assist in determining the expression pattern of the gene and proper size of its product. The role of the *Drosophila dnaJ* protein would probably be similar to the other *dnaJ* proteins previously described.

Whatever the 1.1kb transcript may encode, the fact remains that its gene resides very close to the 5' end of *AP3* and is developmentally regulated. The pattern of expression of the 1.1kb transcript (early embryos and adult females, Figure 8) suggests that it may be of maternal origin. The similarity of the presumptive *Drosophila dnaJ* protein with the *E.coli dnaJ* protein and the apparent maternal expression of the *Drosophila dnaJ* gene suggests that the *Drosophila dnaJ* protein would most likely be involved in the assembly and

transport of nascent polypeptides produced early in development. The recent evidence of the *E.coli* dnaJ protein interacting with the heat shock transcription factor sigma-32 (Gamer *et al.*, 1992) suggests that the *Drosophila* dnaJ protein would also function in the transcriptional regulation of other heat shock genes expressed during early *Drosophila* development. Because AP3 is expressed throughout the development of *Drosophila*, the 1.1kb transcript may not and probably does not play an integral role in any regulation of AP3. However, the converse of AP3 affecting the expression of the 1.1kb transcript may be quite a different story. The regulation of expression of either of these genes must await the fine structure mapping of the 5' regions of both genes.

#### Genomic Map of AP3

The exonic and intronic regions of AP3 were established by comparing the sequences of the genomic fragments and the AP3 cDNA. At present, AP3 is separated into three exons and two rather small introns (Figure 4B). The third and largest exon encodes 300 of the 317 amino acids of AP3. It also contains the stop codon UAA, which lies just 5' to a *Bam*HI restriction endonuclease site (Figure 4A). Nucleotide and protein sequence analysis of AP3 revealed the positions of four possible alternative translational start sites (Figure 5). All four start sites are in frame with the rest of AP3 and would encode truncated versions of AP3 containing the

carboxyl terminal portion of AP3. The *in vitro* transcription and translation of the AP3 cDNA (Kelley et al., 1989) failed to generate any products other than the 35,000 dalton full length AP3 protein. Whether proteins are actually generated from these putative alternate translational start sites is yet to be determined. We have shown data which suggests that the truncated proteins may indeed be produced. Western blot analysis of nuclear and cytoplasmic fractions from Schneider II cells revealed the existence of a protein with a molecular weight very similar to the possible truncated AP3 proteins. This protein was detected with anti-AP3 antibodies strongly suggesting the 28,000 dalton protein could very well be the product of one of the four alternate translational start sites. The fact that this protein was not detected in any other fractions or stages rules out the possibility that it might be a degradative product from AP3.

The second exon (94 bp), is separated from the third exon by a 74 bp intron. The second exon contains the authentic translational start site and encodes the first 17 amino acids of AP3. A small portion of the 5' untranslated region also resides in this exon. The first exon, depicted as an open ended box (Figure 4A and 4B), does not contain any coding region. The box is open at the 5' end because the transcriptional start site has not yet been definitively mapped. Primer extension experiments using three separate

oligonucleotide primers (all residing near the 5' end of the cDNA) indicate that the 5' end of the transcript could be greater than 500 nucleotides upstream from the end of the AP3 cDNA (data not shown). These results are in conflict with the expected size of the primer extension products based on the size of the AP3 transcript (1.3kb). The position of the oligonucleotide primers used in the extension experiments are near the 5' end of the AP3 cDNA (1.2kb). Therefore, the extension products should be no more than 200bp in length. A possible explanation for the unexpected size of the extension products could be the interference of an undetected transcript generated from a gene upstream and on the same strand of DNA as AP3. One piece of evidence suggesting the existence of this unknown gene is the fact that a long open reading frame, revealed during analysis of the genomic sequences upstream of AP3, exists on the same strand of DNA as AP3 (Figure 6).

#### Cloning of P0 and Comparison to AP3

An antibody prepared against a homogenous preparation of a human apurinic/apyrimidinic endonuclease (Kane and Linn, 1981) was used to clone the *Drosophila* AP3 cDNA (Kelley et al., 1989), which in turn was used to clone a human cDNA subsequently identified as P0 (Grabowski et al., 1991; Rich and Steitz, 1987). Rich and Steitz cloned the P0 gene using sera from patients with systemic lupus erythematosus (SLE), an autoimmune disease manifested by antibodies directed against

nuclear and cytoplasmic components of the cell (Christian and Elkon, 1986). P0 encodes a ribosomal protein considered to be a member of the A-proteins (acidic) or P-proteins (the phosphorylated eukaryotic A-proteins) family. These proteins are generally present in multiple copies on the ribosome and have a hydrophobic amino acid composition (notably about 20% alanine). The *Drosophila* protein also has a large number of alanines (about 12%) and unlike P0 has a net positive charge as it has three more basic than acidic amino acids. Both P0 and AP3 proteins are the same length, 317 amino acids, and have the characteristic acidic carboxyl tail.

This carboxyl tail, found in P0 proteins from various organisms (Figure 13), is considered to be the antigenic region which the autoantibodies recognize in patients with SLE (Elkon *et al.*, 1986). In fact, the specificity of lupus autoantibodies for 3 proteins out of a total of approximately 80 ribosomal proteins (Elkon *et al.*, 1985) was shown to be further restricted to the carboxyl acidic tail of all three proteins, namely P0, P1, and P2 (Elkon *et al.*, 1986). These three proteins are located within the 60S ribosomal subunit attached to each other in a 1:2:2 stoichiometry (P0, P1 and P2 respectively), and play an essential role in protein synthesis. This role is dependent on the phosphorylation states of the ribosomal proteins. Hasler *et al.* (1991) demonstrated that all three P proteins (P0, P1 and P2) are

phosphorylated by casein kinase II (CKII) and a ribosomally associated CKII-like enzyme but were not substrates for protein kinase A which phosphorylates many other ribosomal proteins. The extent of phosphorylation of ribosomal proteins appears to affect the functional role of these proteins as an increase in phosphorylation causes an increase in the affinity of ribosomal proteins for ribosomes (Sanchez-Madrid *et al.*, 1981).

Both P0 and AP3 proteins contain casein kinase I and II potential phosphorylation sites (Figure 13). Although AP3 has been shown to be phosphorylated by casein kinase I and II (see Results), the effect of this phosphorylation on the affinity of AP3 for ribosomes has not been studied. The homology between AP3 and P0 on the amino acid level, along with several similar structural features indicates that AP3 may have a dual function, one of which might involve an association with ribosomes.

This potential association is supported by the fact that both AP3 and P0 also contain potential 28s rRNA interaction regions (Figure 13). This 23 amino acid region is very similar to a region found in the L10 acidic ribosomal protein of *E.coli*. It is this region which has been implicated in the binding of L10 to the 28s rRNA molecules in *E.coli* (Mitsui *et al.*, 1989). P0 is the presumptive mammalian equivalent of L10. Since AP3 is the *Drosophila* homologue to P0 it stands to

reason that AP3 is strongly related to *E.coli* L10 and could be the *Drosophila* equivalent to L10. In *E.coli*, L10 is responsible for complexing with two L7/L12 dimers (equivalents to mammalian P1 and P2) and binding the complex directly to 23S rRNA (Brot and Weissbach, 1981). The equivalent complex is attached to the ribosomes, interacts with elongation factors, and are required for aminoacyl-tRNA binding and GTPase activity associated with the reactions of polypeptide synthesis (MacConnell and Kaplan, 1982, Sanchez-Madrid et al., 1979). Thus, P0 is an integral part of a necessary complex required for protein synthesis. The role of AP3 in *Drosophila* when associated with ribosomes would have a similar role in complexing with the *Drosophila* equivalents of P1 and P2, assisting in polypeptide elongation. Although not presented in this dissertation, data from Western analysis of blots containing partially purified ribosomes probed with anti-AP3 antibodies indicate that AP3 is associated with ribosomes (Dr. D. Derda, personal communication). This evidence strongly supports a role for AP3 in polypeptide synthesis. However, the initial characterization of AP3 indicates that it is a DNA repair enzyme. AP3 may in fact have dual roles, one in DNA repair and one involving ribosome function. Whether P0 has a dual role as well has yet to be determined.

#### P0 as a DNA-Repair Protein.

Although we did not investigate the role of AP3 as a

ribosomal protein, we did attempt to determine whether P0 may play a role in DNA-repair. To this end, we investigated the ability of three structurally related bifunctional alkylating agents to induce the expression of a presumptive human DNA repair gene, P0 (Grabowski et al., 1992). The three alkylating agents, melphalan (L-PAM), 4-hydroperoxycyclophosphamide (4HC), and mechlorethamine (HN2) are all clinically useful DNA-damaging antitumor agents. These agents interact with numerous cellular macromolecules, but are thought to exert their cytotoxic effects through the production of DNA alkylations and interstrand and intrastrand crosslinks (Erickson and Zlotogorski, 1984). Previous studies by Futscher et al. (1990) have shown these agents to increase the expression of *c-myc* and *c-fos* mRNAs but have no effect on the expression of  $\beta$ -actin or *N-ras*. Their work suggests that these agents may initiate a cellular response to DNA damage, not merely a global increase in expression of all genes. In our investigation, low doses of each of the alkylating agents produced an immediate increase in the expression of P0 (Figure 15A). Higher doses also induced the expression of P0 but only after a lag period (Figure 15B). If P0 were exclusively a ribosomal protein, one would not expect an induction of its expression in response to DNA-damaging agents. The results described in this dissertation pertaining to the expression of P0 in response to DNA damaging agents or DNA-repair



deficiencies leads us to conclude that P0 must act as a DNA repair protein. The constitutive overexpression of P0 in Mer<sup>-</sup> human tumor cell lines, versus Mer<sup>+</sup> cells lines strongly suggests that P0 is being overexpressed to compensate for the lack of O<sup>6</sup>-methylguanine methyltransferase activity (MGMT). Mer<sup>-</sup> cells lack any detectable MGMT activity. The constitutive expression of P0 is 40- to 50-fold higher, without an increase in copy number, in Mer<sup>-</sup> human tumor cell lines (Grabowski *et al.*, 1992). This is important in that overexpressed P0 could act as a DNA repair enzyme in a repair pathway different than that of MGMT. The alkylated bases are recognized by glycosylases of the base-excision repair pathway. The resulting AP sites generated by the glycosylases would then be recognized by AP endonucleases, in this case possibly P0. P0, being equivalent to the AP endonuclease AP3 would continue the process of base-excision repair.

This elevation in mRNA was not a global effect of other ribosomal associated genes as our results with the human ribosomal gene S17 show no change in the steady-state expression of this true ribosomal protein gene. However, the changes in gene expression of ribosomal genes is not completely limited to P0 either. Pogue-Geile *et al.* (1991) isolated an overexpressed ribosomal protein gene, S3, as the result of investigating changes in gene expression in

adenocarcinoma of the colon compared to normal colonic mucosa. Their investigation demonstrated increased levels of mRNAs from several other ribosomal proteins as well, including S6, S8, S12, L5 and P0. An explanation is offered describing the increase of ribosomal protein expression as a result of the proliferative activity of the cell. Many cell types undergoing active proliferation show increased levels of ribosome proteins (Kief and Warner, 1981; Tushinski and Warner, 1982). However, the proliferative indices of colorectal carcinomas and normal colonic mucosa are not significantly different (Bleiberg and Galand, 1976). Thus, it is unlikely that increased levels of ribosomal protein mRNAs is due to a higher percentage of dividing cells in a tumor. The increased level of ribosomal protein gene expression may be the result of neoplasia and not a causal event of carcinogenesis. However, the ribosomal protein gene S17 was not overexpressed in Mer<sup>-</sup> versus Mer<sup>+</sup> cell lines, and this explanation does not explain the differences in the level of P0 expression in Mer<sup>-</sup> versus Mer<sup>+</sup> cell lines. It is unknown whether any of the colorectal tumors, used to isolate the S3 ribosomal gene and demonstrate increased expressions of several other ribosomal genes, including P0, were Mer<sup>-</sup> or Mer<sup>+</sup>.

More direct evidence supporting a dual role of ribosomal proteins involved in DNA-repair, comes from the laboratory of

Dr. S. Linn (personal communication) which demonstrates ribosomal protein S3 having DNA glycosylase/AP endonuclease activity to UV damaged DNA. Dr. Linn's group isolated and characterized a UV endonuclease, which was identified, from microsequence data of the purified protein, as the human ribosomal protein, S3. Concurrently, this group isolated a defective S3 protein from the repair-deficient xeroderma pigmentosum group-D cells suggesting that a defective S3 protein could be the cause of the repair deficiency (personal communication). Dr. Linn's work directly demonstrates that some ribosomal proteins are capable of DNA-repair activity. This evidence, along with the previously discussed observations of P0 and its relationship to AP3 leads us to conclude that P0 can be a multifunctional protein.

#### Localization of AP3

A goal of this dissertation was to analyze the AP3 protein on the biochemical level. To this end, polyclonal antibodies directed against fusion proteins containing all of AP3 or truncated portions of AP3 (Figure 18) were generated in rabbits. These antibodies were affinity purified using AP3 antigen immobilized on nitrocellulose strips and subsequently used in Western analysis to define a subcellular localization of AP3. Nuclear and cytoplasmic extracts from various developmental stages of *Drosophila* were analyzed using SDS-PAGE followed by Western blot analysis using the affinity

purified anti-AP3 antibody. The 35,000 dalton AP3 protein was detected in both nuclear and cytoplasmic fractions in all stages of development. Following a protocol for the purification of the nuclear matrix (Fisher and Blobel, 1983), we have determined that AP3 is not only located in the nucleus but is associated, at least partially, with purified nuclear matrix (Figure 22). The localization of AP3 to the nuclear matrix was not surprising as other DNA metabolism proteins, such as topoisomerase II and DNA ligases are also thought to be localized to the nuclear matrix. Berrios *et al.*, (1985) have identified topoisomerase II as a major polypeptide component of the *Drosophila* nuclear matrix-pore complex-lamina. The nuclear matrix has also been implicated in many processes involving interactions with nucleic acids including DNA replication, transcription, and of course DNA repair (Hakes and Berezney, 1991; Mullenders *et al.*, 1990). There are many proteins interactive with DNA located within the internal nuclear matrix as well as at the nuclear periphery. Our results indicating the subnuclear localization of AP3 to the nuclear matrix demonstrates that AP3 is not merely associated with the nuclear periphery.

The detection of AP3 in the cytoplasmic fractions at first came as a surprise. Why would a DNA repair protein reside outside of the nucleus? The answer to this question can only be that AP3 is a multifunctional protein. The

function of AP3 when residing in the nucleus is AP endonucleolytic activity. As discussed earlier, the function of AP3 when not in the nucleus is to form a complex with the *Drosophila* P1 and P2 proteins and bind to the ribosomes. This complex is necessary for the elongation step during protein synthesis.

The phosphorylation of P0 increases its affinity for the ribosome and both P0 and AP3 have been shown to contain potential casein kinase I and II phosphorylation sites (Figure 13). AP3, overexpressed in *E.coli* and purified to homogeneity (Figure 20), was phosphorylated *in vitro* at both the S residues, by casein kinase I and II (Dr. M. Tao, personal communication). This phosphorylation may be the necessary signal to traffic AP3 to the ribosome. What other effects the phosphorylation may have is unknown. However, the phosphorylation of AP3 may direct it to the cytoplasm unidirectionally. To test the role of phosphorylation, *Drosophila* tissue culture cells could be grown in the presence of radiolabeled inorganic phosphate. All proteins that are phosphorylated would become labeled. Immunoprecipitating AP3 from nuclear and post nuclear fractions of these cells using purified anti-AP3 and running the immunoprecipitants through a SDS-polyacrylamide gel followed by auto-radiography would indicate which fractions contained phosphorylated AP3. If the phosphorylation directs AP3 to the cytoplasm to be associated

with ribosomes, then the radiolabeled AP3 from the immunoprecipitation experiment should all be in the post nuclear fraction.

The possibility that portions of AP3 are directed to the cytoplasm is supported by the observation that approximately twice the amount of AP3 was detected in the cytoplasmic fractions from adults and embryos compared to the nuclear fractions (Figure 21). Since equal amounts of protein were loaded in each lane, the difference in amount of AP3 detected cannot be due to loading differences. Results of the suggested immunoprecipitation experiment may explain why different amounts of AP3 reside in separate portions of the cell.

#### AP3 is a Class II AP Endonuclease.

The most intriguing question regarding the biochemical characterization of AP3 is whether it displays actual AP endonuclease activity. To determine this, full length AP3 was overexpressed in *E.coli* from the pET11d-AP3 vector and purified from inclusion bodies (Figure 20) following a modified version of the protocol described by Lin and Cheng (1991). This protocol resulted in AP3 being purified to a single silver-staining band (Figure 20). Purified AP3 was assayed for AP endonuclease activity using an assay which discriminates between Class I ( $\beta$ -lyase) and Class II (hydrolytic) AP endonuclease activity (Levin and Demple,

1990). AP3 was determined to have specific Class II AP endonuclease activity that is inhibited by EDTA (J. Carney, personal communication). The parameters ensuring optimal activity for AP3 have not been determined. However, the effect of phosphorylation on the activity of AP3 was investigated. Using *in vitro* phosphorylated AP3 (from Dr. M. Tao), no significant change was detected in enzyme activity. One possible explanation for this might be the fact that the optimal conditions for AP3 endonuclease activity have not yet been determined. Perhaps under optimal conditions an increase (or decrease) of AP3 activity may be observed. However, the phosphorylation state of AP3 may not play any role in the AP endonuclease activity of AP3. As described earlier, the phosphorylation of AP3 may only serve to determine the subcellular localization of AP3 and perhaps increase its affinity with ribosomes.

Linn *et al.* (1981) have described four theoretical classes of AP endonucleases (Figure 2). To date, nearly all known AP endonucleases have been identified as either Class I AP lyases or Class II AP endonucleases. A putative Class III AP endonuclease (AP Endonuclease I) was isolated from *Drosophila* (Spiering and Deutsch, 1986). Spiering and Deutsch also identified a Class I AP endonuclease from *Drosophila* (AP Endonuclease II). AP Endonuclease I (63,000 D) and II (66,000 D) are chromatographically distinct and cross-react with the

antibody prepared toward a human AP endonuclease (Kane and Linn, 1981) used to isolate the cDNA for AP3 (Kelley et al., 1989). The cross-reactivity shared by these three *Drosophila* AP endonucleases when using the anti-human AP endonuclease antibody, appears to be the only similarity between them. It is important to note that purified anti-AP3 antibodies do not cross react with any other *Drosophila* proteins. AP3 is further distinguished from AP Endonuclease I and II by its size (35,000 daltons). Finally, the activity of AP3 is unlike AP Endonuclease I and II as AP3 is a Class II AP endonuclease. Therefore, AP3 can compliment the Class I activity of AP Endonuclease II to create an efficient terminus for DNA synthesis.

The cDNA of another Class II AP endonuclease encoding a recombination repair protein, *Rrpl*, was recently isolated by Sander et al. (1991). *Rrpl* has an N-terminal 427-amino acid region unrelated to known proteins, and a 252-amino acid C-terminal region with homology to two AP endonucleases. Although initially characterized as having strand transfer activity (Lowenhaupt et al., 1989), *Rrpl* demonstrates 3' exonuclease, AP endonuclease, and single-stranded DNA renaturation activities. This combination of repair and recombination activities has not been observed previously. The relationships between *Rrpl* and AP3 is based only on the fact that both exhibit Class II AP endonuclease activity. The



sequences of *AP3* and *Rrpl* share no known homology. Although both *AP3* and *Rrpl* are Class II AP endonucleases, neither has been determined to be the major AP endonuclease in *Drosophila*.

In summary, a genomic map of *AP3* was constructed based on sequence data and Southern blot analysis of genomic clones isolated from a Maniatis library using the *AP3* cDNA as a probe. Two other genes were discovered, which reside near *AP3*. One of the genes, *Dromsopa*, lies approximately 4.0kb downstream from *AP3*. The promoter of the other gene, not yet cloned, overlaps the promoter region of *AP3*. *AP3* has been detected in two subcellular locations, the nucleus and associated with ribosomes, supporting the theory that *AP3* is both a DNA-repair enzyme and a ribosomal protein. The DNA-repair activity of purified *AP3* was determined to be Class II. The role of *AP3* as a ribosomal proteins is supported by its extensive homology to the human ribosomal protein P0. P0 has been implicated in DNA repair through induction of P0 in cells treated with DNA-damaging agents and through the constitutive overexpression of P0 in DNA-repair deficient cell lines. The data presented in this dissertation makes a strong case for both *AP3* and P0 being dual function proteins involved in DNA-repair and ribosomal function.

LITERATURE CITED

- Ames, B.N. 1983. Dietary carcinogens and anticarcinogens: Oxygen radicals and degenerative diseases. *Science* 221:1256-1264.
- Armel, P.R., and S.S. Wallace. 1978. Apurinic endonucleases from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 5:3347-3356.
- Armel, P.R., and S.S. Wallace. 1984. DNA repair in *Saccharomyces cerevisiae*: purification and characterization of apurinic endonucleases. *J. Bacteriol.* 160:895-902.
- Asahara, H., Wistort, P.M., Bank, J.F., Bakerian, R.H., and R.P. Cunningham. 1989. Purification and characterization of *Escherichia coli* endonuclease III from the cloned *nth* gene. *Biochemistry* 28:4444-4449.
- Auerbach, J., and P. Howard-Flanders. 1979. The isolation and genetic characteristics of lambda transducing phages of the *uvrA*<sup>+</sup> and *uvrC*<sup>+</sup> genes of *E.coli* K12. *Mol. Gen. Genet.* 168:341-344.
- Bailly, V., Derydt, M., and W.G. Verly. 1989.  $\delta$ -Elimination in the repair of AP(apurinic/apyrimidinic) sites in DNA. *Biochem. J.* 261:707-713.
- Bailly, V., and W.G. Verly. 1987. *Escherichia coli* endonuclease III is not an endonuclease but a  $\beta$ -elimination catalyst. *Biochem. J.* 242:565-572.
- Bardwell, J.C.A., Tilly, K., Craig, E., King, J., Zylicz, M. and C. Georgopoulos. 1986. The nucleotide sequence of the *Escherichia coli* K12 *dnaJ*<sup>+</sup> gene. *J. Biol. Chem.* 261:1782-1785.
- Berrios, M., Osheroff, N. and P.A. Fisher. 1985. *In situ* localization of DNA topoisomerase II, a major polypeptide component of the *Drosophila* nuclear matrix fraction. *Proc. Natl. Acad. Sci. USA* 82:4142-4146.
- Bjelland, S., and E. Seeberg. 1987. Purification and characterization of 3-methyladenine DNA glycosylase I from *Escherichia coli*. *Nucleic Acids Res.* 15:2787-2801.
- Bleiberg, H., and P. Galand. 1976. *In vitro* autoradiographic determination of cell kinetic parameters in adenocarcinomas and adjacent healthy mucosa of the human colon and rectum. *Cancer Res.* 36:325-328.

Blochlinger, K., Bodmer, R., Jack, J., Jan, L.Y., and Y.N. Jan. 1988. Primary structure and expression of a product from *cut*, a locus involved in specifying sensory organ identity in *Drosophila*. *Nature* (London) 333:629-635

Boiteux S., and J. Laval. 1982. Coding properties of poly (deoxycytidylic acid) templates containing uracil or apyrimidinic sites: in vitro modulation of mutagenesis by DNA repair enzymes. *Biochemistry* 21:6746-6751.

Boiteux, S., and J. Laval. 1983. Imidazole open ring 7-methylguanine: An inhibitor of DNA synthesis. *Biochem. Biophys. Res. Commun.* 110:552-558.

Boyce, R., and P. Howard-Flanders. 1964. Release of ultraviolet light-induced thymine dimers from DNA in *E.coli* K-12. *Proc. Natl. Acad. Sci. USA* 51:293-300.

Breimer, L.H., and T. Lindahl. 1984. DNA glycosylase activities for thymine residues damaged by ring saturation, fragmentation, or ring contraction are functions of Endonuclease III in *Escherichia coli*. *J. Biol. Chem.* 259:5543-5548.

Bridges, B.A., and R. Woodgate. 1985. Mutagenic repair in *Escherichia coli*: Products of the *recA* gene and of the *umuD* and *umuC* genes act at different steps in UV-induced mutagenesis. *Proc. Natl. Acad. Sci. USA* 82:4193-4197.

Brot, N., and H. Weissbach. 1981. Chemistry and biology of *E.coli* ribosomal protein L12. *Mol. Cell. Biochem.* 36:47-63.

Burgers, P.M.J. and M.B. Klein. 1986. Selection by genetic transformation of a *Saccharomyces cerevisiae* mutant defective for the nuclear uracil-DNA-glycosylase. *J. Bacteriol.* 166:905-913.

Caradonna, S.J., and Y.-C. Cheng. 1980. Uracil DNA-glycosylase. *J. Biol. Chem.* 255:2293-2300.

Chan, P.-K., Paz, V. and I.G. Wool. 1989. Submitted to EMBL/Genbank data base under accession number P19945.

Chan, E., and B. Weiss. 1987. Endonuclease IV of *Escherichia coli* is induced by paraquat. *Proc. Natl. Acad. Sci. USA* 84:3189-3193.

Chang, C., Kokontis, J., Chang, C-T., and S. Liao. 1987. Cloning and sequence analysis of the rat ventral prostate glucocorticoid receptor cDNA. *Nucleic Acids Res.* 15:9603.

- Cheetham, M.E., Brian, J.P., and B.H. Anderton. 1991. GCG Database, unpublished data.
- Chen, I.-T., and D.J. Roufa. 1988. The transcriptionally active human ribosomal protein S17 gene. *Gene* 70:107-116.
- Christian, C.L., and K.B. Elkon. 1986. Autoantibodies to intracellular proteins: clinical and biological implications. *Am. J. Med.* 80:53-61.
- Clements, J.E., Rogers, S.G., and B. Weiss. 1978. A DNase for apurinic/apyrimidinic sites associated with exonuclease III of *Hemophilus influenzae*. *J. Biol. Chem.* 253:2990-2999.
- Crosby, B., Prakash, L., Davis, H., and D.C. Hinkle. 1981. Purification and characterization of a uracil-DNA glycosylase from yeast, *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 9:5797-5809.
- Cunningham, R.P., Saporito, S.M., Spitzer, S.G., and B. Weiss. 1986. Endonuclease IV (*nfo*) mutant of *Escherichia coli*. *J. Bacteriol.* 168:1120-1127.
- Cunningham, R.P., and B. Weiss. 1985. Endonuclease III (*nth*) mutants of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 82:474-478.
- Danielson, M., Northrop, J.P., Jonklaas, J., and G.M. Ringold. 1987. Domains of the glucocorticoid receptor involved in specific and nonspecific deoxyribonucleic acid binding, hormone activation, and transcriptional enhancement. *Mol. Endocrinol.* 1:816-822.
- Demple, B., Halbrook, J., and S. Linn. 1983. *Escherichia coli xth* mutants are hypersensitive to hydrogen peroxide. *J. Bacteriol.* 153:1079-1082.
- Demple, B., Herman, T., and D.S. Chen. 1991. Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: Definition of a family of DNA repair enzymes. *Proc. Natl. Acad. Sci. USA* 88:11450-11454.
- Demple, B., Johnson, A., and D. Fung. 1986. Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H<sub>2</sub>O<sub>2</sub>-damaged *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 83:7731-7735.
- Demple, B., and S. Linn. 1980. DNA N-glycosylases and UV repair. *Nature* 287:203-208.

Deutsch, W.A., and S. Linn. 1979a. Further characterization of a depurinated DNA purine base insertion activity from cultured human fibroblasts. *J. Biol. Chem.* 254:12099-12103.

Deutsch, W.A., and S. Linn. 1979b. DNA binding activity from cultured human fibroblasts that is specific for partially depurinated DNA and that inserts purines into apurinic sites. *Proc. Natl. Acad. Sci. USA* 76:141-144.

Deutsch, W.A., and A.L. Spiering. 1985. Characterization of a depurinated-DNA purine-base-insertion activity from *Drosophila*. *Biochem. J.* 232:285-288.

Devereux, J., Haeberli, P. and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387-395.

Doetsch, P.W., and R.P. Cunningham. 1990. The enzymology of apurinic/apyrimidinic endonucleases. *Mutat. Res.* 236:173-201.

Doetsch, P.W., Helland, D.E., and W.A. Haseltine. 1986. Mechanism of action of a mammalian DNA repair endonuclease. *Biochemistry* 25:2212-2220.

Doetsch, P.W., Henner, W.D., Cunningham, R.P., Toney, J.H., and D.E. Helland. 1987. A highly conserved endonuclease activity present in *Escherichia coli*, bovine, and human cells recognizes oxidative DNA damage at sites of pyrimidines. *Mol. Cell Biol.* 7:26-32.

Duboule, D., Haenlin, M., Galliot, B., and E. Mohier. 1987. DNA sequences homologous to the *Drosophila opa* repeat are present in murine mRNAs that are differentially expressed in fetuses and adult tissues. *Mol. Cell. Biol.* 7:2003-2006.

Duncan, B.K., Rockstroh, P.A., and H.R. Warner. 1978. *Escherichia coli* K-12 mutants deficient in uracil-DNA glycosylase. *J. Bacteriol.* 134:1039-1045.

Duncan, B.K., and B. Weiss. 1982. Specific mutator effects of *ung* (uracil-DNA glycosylase) mutations in *Escherichia coli*. *J. Bacteriol.* 151:750-755.

Eker, A.R.M. 1983. Molecular Models of Photoresponsiveness Plenum Press, New York, NY.

Elkon, K.B., Parnassa, A.P. and C.L. Foster. 1985. Lupus autoantibodies target ribosomal P proteins. *J. Exp. Med.* 162:459-471.

Elkon, K., Skelly, S., Parnassa, A., Moller, W., Danho, W., Weissbach, H. and N. Brot. 1986. Identification and chemical synthesis of a ribosomal protein antigenic determinant in systemic lupus erythematosus. *Proc. Natl. Acad. Sci. USA* 83:7419-7423.

Erickson, L., and C. Zlotogorski. 1984. Induction and repair of macromolecular damage by alkylating agents. *Cancer Treat. Rev.* 11:(Suppl. A) 25-35.

Feinberg, P. and B. Vogelstein. 1984. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* 137:266-267.

Fisher, P.A., and G. Blobel. 1983. Preparation of a nuclear matrix - pore complex - lamina fraction from embryos of *Drosophila melanogaster*. *Methods Enzymol.* 96:589-596.

Fisher, E.F., and M.H. Caruthers. 1979. Studies on gene control regions XII. The functional significance of a *Lac* operator constitutive mutation. *Nucleic Acids Res.* 7:401-416.

Fix, D.F., and B.W. Glickman. 1986. Differential enhancement of spontaneous transition mutations in the *lacI* gene of an *Ung* strain of *Escherichia coli*. *Mutat. Res.* 175:41-45.

Fix, D.F., and B.W. Glickman. 1987. Asymmetric cytosine deamination revealed by spontaneous mutational specificity in an *Ung* strain of *Escherichia coli*. *Mol. Gen. Genet.* 209:78-82.

Flamee, P.-A. and W.G. Verly. 1985. Action of intact AP (apurinic/apyrimidinic) sites and AP sites associated with breaks on the transcription of T7 coliphage DNA by *Escherichia coli* RNA polymerase. *Biochem. J.* 229:173-181.

Foote, R.S., Mitra, S., and B.C. Pal. 1980. Demethylation of O<sup>6</sup>-methylguanine in a synthetic DNA polymer by an inducible activity in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 97:654-659.

Franklin, W.A., and T. Lindahl. 1988. DNA deoxyribosephosphodiesterase. *EMBO J.* 7:3617-3622.

Freeman, B.A., and J.D. Crapo. 1982. Biology of disease. Free radicals and tissue injury. *Lab. Invest.* 47:412-426.

Friedberg, E.C., and D.A. Goldthwait. 1969. Endonuclease II of *E. coli*. 1. Isolation and purification. *Proc. Natl. Acad. Sci. USA* 62:934-940.

- Fridovich, I. 1978. The biology of oxygen radicals. *Science* 201:875-880.
- Friedberg, E.C. 1985. DNA Repair, W.H. Freeman Co., San Francisco, CA.
- Friedberg, E.C. 1988. Deoxyribonucleic acid repair in the yeast *Saccharomyces cerevisiae*. *Microbiol. Rev.* 52:70-102.
- Friedberg, E.C., Ganesan, A.K., and K. Minton. 1975. N-glycosidase activity in extracts of *Bacillus subtilis* and its inhibition after infection with bacteriophage PBS2. *J. Virol.* 16:315-321.
- Fu, Y.-H., Kuhl, D.P.A., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M.H., Holden, J.J.A., Fenwick, R.G., Jr., Warren, S.T., Oostra, B.A., Nelson, D.L. and C.T. Caskey. 1991. Variation of the CGG repeat at the fragile-X site results in genetic instability: Resolution of the Sherman paradox. *Cell* 67:1047-1058.
- Futscher, B.W., and L. C. Erikson. 1990. Changes in c-myc and c-fos expression in a human tumor cell line following exposure to bifunctional alkylating agents. *Cancer Res.* 50:62-66.
- Gaitanaris, G.A., Papavassiliou, A.G., Rubock, P., Silverstein, S.J., and M.E. Gottesman. 1990. Renaturation of denatured  $\lambda$  repressor requires heat shock proteins. *Cell* 61:1013-1020.
- Gamer, J., Bujard, H., and B. Bukau. 1992. Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma-32. *Cell* 69:833-842.
- Gossard, F., and W.G. Verly. 1978. Properties of the main endonuclease specific for apurinic sites of *Escherichia coli* (endonuclease VI). *Eur. J. Biochem.* 82:321-332.
- Grabowski, D.T., Deutsch, W.A., Derda, D., and M.R. Kelley. 1991. Drosophila AP3, a presumptive DNA repair protein, is homologous to human ribosomal associated protein P0. *Nucleic Acids Res.* 19:4297.
- Grabowski, D.T., Pieper, R.O., Futscher, B.W., Deutsch, W.A., Erickson, L.C., and M.R. Kelley. 1992. Expression of ribosomal phosphoprotein P0 is induced by antitumor agents and increased in Mer<sup>-</sup> human tumor cell lines. *Carcinogenesis* 13:259-263.

- Grafstrom, R.H., Shaper, N.L., and L. Grossman. 1982. Human placental apurinic/aprimidinic endonuclease. *J. Biol. Chem.* 257:13459-13464.
- Hagen, U. 1986. Current aspects on the radiation induced base damage in DNA. *Radiat. Environ. Biophys.* 25:261-271.
- Hakes, D.J., and R. Berezney. 1991. DNA binding properties of the nuclear matrix and individual nuclear matrix proteins. *J. Biol. Chem.* 266:11131-11140.
- Hamilton, K.K., Kim, P.M.H., and P. Doetsch. 1992. A eukaryotic DNA glycosylase/lyase recognizing ultraviolet light-induced pyrimidine dimers. *Nature* 356:725-728.
- Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Reardon, W., Buckler, A.J., Harper, P.S., Housman, D.E., and D.J. Shaw. 1992. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 355:545-546.
- Harosh, I., and J. Sperling. 1988. Hypoxanthine-DNA glycosylase from *Eschericia coli*. *J. Biol. Chem.* 263:3328-3334.
- Haseltine, W.A. 1983. Ultraviolet light repair and mutagenesis revisited. *Cell* 33:13-17.
- Hasler, P., Brot, N., Weissbach, H., Parnassa, A.P., and K.B. Elkon. 1991. Ribosomal proteins P0, P1, and P2 are phosphorylated by casein kinase II at their conserved carboxyl termini. *J. Biol. Chem.* 266:13815-13820.
- Hayatsu, H. 1976. Bisulfite modification of nucleic acids and their constituents. *Prog. Nucl. Acids Res. Mol. Biol.* 16:75-124.
- Hemminki, K. 1983. Nucleic acid adducts of chemical carcinogens and mutagens. *Arch. Toxicol.* 52:249-285.
- Hendrickson, E.A., Qin, X.-Q., Bump, E.A., Schatz, D.G., Oettinger, M., and D.T. Weaver. 1991. A link between double-strand break-related repair and V(D)J recombination: the scid mutation. *Proc. Natl. Acad. Sci. USA* 88:4061-4065.
- Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
- Henner, W.D., Niker, N.P., Jorgensen, T.J., and J.N. Munck. 1987. Purification and amino acid terminal sequence of an apurinic/aprimidinic endonuclease from calf-thymus. *Nucleic Acids Res.* 15:5529-5544.



- Hollstein, M.C., Brooks, P., Linn, S., and B.N. Ames. 1984. Hydroxymethyluracil DNA glycosylase in mammalian cells. *Proc. Natl. Acad. Sci. USA* 81:4003-4007.
- Hutchinson, F. 1985. Chemical changes induced in DNA by ionizing radiation. *Prog. Nucleic Acids Res. Mol. Biol.* 32:115-154.
- Johnson, A.W., and B. Demple. 1988a. Yeast DNA diesterase for 3'-fragments of deoxyribose: purification and physical properties of a repair enzyme for oxidative DNA damage. *J. Biol. Chem.* 263:18009-18016.
- Johnson, A.W., and B. Demple. 1988b. Yeast 3'-repair diesterase is the major cellular apurinic/apyrimidinic endonuclease: substrate specificity and kinetics. *J. Biol. Chem.* 263:18017-18022.
- Kane, C.M., and S. Linn. 1981. Purification and characterization of an apurinic/apyrimidinic endonuclease from HeLa cells. *J. Biol. Chem.* 256:3405-3414.
- Kane, J.E., and D.L. Hartley. 1988. Formation of recombinant protein inclusion bodies in *E.coli*. *Trends Biotechnol.* 6:95-101.
- Karran, P., and T. Lindahl. 1980. Hypoxanthine in deoxyribonucleic acid: generation by heat-induced hydrolysis of adenine residues and release in free form by a deoxyribonucleic acid glycosylase from calf thymus. *Biochemistry* 19:6005-6011.
- Keijzer, W., Jaspers, N.G.J., Abrahams, P.J., Taylor, A.M.R., Arlett, C.F., Zelle, B., Takebe, H., Kinmont, P.D.S., and D. Bootsma. 1979. A seventh complementation group in excision-deficient xeroderma pigmentosum. *Mutat. Res.* 62:183-190.
- Keller, W., and R. Crouch. 1972. Degredation of DNA-RNA hybrids by ribonuclease H and DNA polymerases of cellular and viral origin. *Proc. Natl. Acad. Sci. USA* 69:3360-3364.
- Kelley, M.R., Venugopal, S., Harless, J., and W.A. Deutsch. 1989. Antibody to a human DNA repair protein allows for cloning of a *Drosophila* cDNA that encodes an apurinic endonuclease. *Mol. Cell. Biol.* 9:965-973.
- Kennelly, P.J. and E.G. Krebs. 1991. Consensus sequences as substrate spectificity determinants for protein kinases and protein phosphatases. *J. Biol. Chem.* 266:15555-15558.

Kidd, S., Kelley, M.R., and M.W. Young. 1986. Sequence of the Notch locus of *Drosophila melanogaster*: Relationship of the encoded protein to mammalian clotting and growth factors. *Mol. Cell. Biol.* 6:3094-3108.

Kidd, S., Lockett, T.J., and M.W. Young. 1983. The Notch locus of *Drosophila melanogaster*. *Cell* 34:421-433.

Kief, D.R., and J.R. Warner. 1981. Coordinate control of synthesis of ribosomal ribonucleic acid and ribosomal proteins during nutritional shift-up in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 1:1007-1015.

Kim, J., and S. Linn. 1988. The mechanisms of action of *E.coli* endonuclease III and T4 UV endonuclease (endonuclease V) at AP sites. *Nucleic Acids Res.* 16:1135-1141.

Kim, J., and S. Linn. 1989. Purification and characterization of UV endonucleases I and II from murine plasmacytoma cells. *J. Biol. Chem.* 264:2739-2745.

Kozak, M. 1984. Compilation and analysis of sequences ipstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* 12:857-872.

Krowczynska, A.M., Coutts, M., Makrides, S., and G. Brawerman. 1989. The mouse homologue of the human acidic ribosomal phosphorprotein PO: a highly conserved polypeptide that is under translational control. *Nucleic Acids Res.* 17:6408.

LaSpada, A.R., Wilson, E.M., Lubahn, D.B., Harding A.E., and K.H. Fishbeck. 1991. Androgen receptor gene mutations in X-linked spinal and balbar muscular atrophy. *Nature* 352:77-79.

Laughon, A., Boulet, A.M., Bermingham, J.R., Jr., Laymon, R.A., and M.P. Scott. 1986. Structure of transcripts from the homeotic antennapedia gene of *Drosophila melanogaster*: Two promoters control the major protein-coding region. *Mol. Cell. Biol.* 6:4676-4689.

Laughon, A., Carroll, S.B., Storfer, F.A., Riley, P.D., and M.P. Scott. 1985. Common properties of proteins incoded by the antennepedia complex genes of *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.* 50:253-262.

Lehmann, A.R., Hoeijmakers, J.H.J., van Zeeland, A.A., Backendorf, C.M.P., Bridges, B.A., Collins, A., Fuchs, R.P.O., Margison, G.P., Montesano, R., Moustacchi, E., Natarajan, A.T., Radman, M., Sarasin, A., Seeberg, E., Smith, C.A., Stefanini, M., Thompson, L.H., van der Schans, G.P., Weber, C.A., and M.Z. Zdzienicka. 1992. Workshop on DNA repair. *Mutat. Res.* 273:1-28.

Levin, J.D., and B. Demple. 1990. Analysis of Class II (hydrolytic) and Class I ( $\beta$ -lyase) apurinic/apyrimidinic endonucleases with a synthetic DNA substrate. *Nucleic Acids Res.* 18:5069-5075.

Levin, J.D., Johnson, A.W., and B. Demple. 1988. Homogeneous *Escherichia coli* endonuclease IV. *J. Biol. Chem.* 263:8066-8071.

Lin, K., and S. Cheng. 1991. An efficient method to purify active eukaryotic proteins from the inclusion bodies in *Escherichia coli*. *Biotechniques* 11:748-753.

Lindhahl, T. 1974. An N-glycosidase from *Escherichia coli* that releases uracil from DNA containing deaminated cytosine residues. *Proc. Natl. Acad. Sci. USA* 71:3649-3653.

Lindhahl, T. 1979. DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision repair. *Prog. Nucleic Acid Res. Mol. Biol.* 22:135-192.

Lindhahl, T. 1982. DNA repair enzymes. *Ann. Rev. Biochem.* 51:61-87.

Lindhahl, T., and A. Andersson. 1972. Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry* 11:3618-3623.

Lindhahl, T., and B. Nyberg. 1972. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* 11:3610-3618.

Lindhahl, T., and B. Nyberg. 1974. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* 13:3405-3410.

Lindhahl, T., and B. Sedgwick. 1988. Regulation and expression of the adaptive response to alkylating agents. *Ann. Rev. Biochem.* 57:133-157.

Linn, S., Demple, B., Mosbaugh, D.W., Warner, H.R., and W. Deutsch. 1981. In: *Chromosome Damage and Repair* (Seeberg, E. and Kleppe eds.) Plenum Press, New York. pp 97-112.

Ljungquist, S., Lindahl, T., and P. Howard-Flanders. 1976. Methyl methane sulfonate-sensitive mutant of *Escherichia coli* deficient in an endonuclease specific for apurinic sites in deoxyribonucleic acid. *J. Bacteriol.* 126:646-653.

Ljungquist, S., Nyberg, B., and T. Lindahl. 1975. Mammalian DNA endonuclease acting at apurinic sites: Absence of associated exonuclease activity. *FEBS Lett.* 57:169-171.

- Loeb, L.A. 1989. Endogenous carcinogenesis: Molecular oncology into the twenty-first century. Presidential address. *Cancer Res.*, 49:5489-5496.
- Loechler, E.L., Green, C.L., and J.M. Essignmann. 1984. *In vitro* mutagenesis by O<sup>6</sup>-methylguanine built into a unique site in a viral genome. *Proc. Natl. Acad. Sci. USA* 81:6271-6275.
- Lowenhaupt, K., Sander, M., Hauser, C., and A. Rich. 1989. *Drosophila melanogaster* strand transferase. *J. Biol. Chem.* 264:20568-20575.
- Lubahn, D.B., Joseph, D.R., Sar, M., Tan, J., Higgs, H.N., Larson, R.E., French, F.S., and E.M. Wilson. 1988. The human androgen receptor: Complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate. *Mol. Endocrinol.* 2:1265-1275.
- MacConnell, W.P., and N.O. Kaplan. 1982. The activity of the acidic phosphoproteins from the 80s rat liver ribosome. *J. Biol. Chem.* 257:5359-5366.
- Maniatis, T., Fritsch, E.F., and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Maniatis, T., Fritsch, E.F., and J. Sambrook. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. Vol. 2.
- Manoharan, M., Mazumder, A., Ransom, S.C., Gerlt, J.A., and P.H. Bolton. 1988. Mechanism of UV endonuclease V cleavage of abasic sites determined by <sup>13</sup>C labeling. *J. Am. Chem. Soc.* 110:2690-2691.
- Margulies, L., and S.S. Wallace. 1984. Apurinic endonuclease activity remains constant during early *Drosophila* development. *Cell Biol. Toxic.* 1:127-143.
- McKinnon, P.J. 1987. Ataxia-telangiectasia: an inherited disorder of ionizing-radiation sensitivity in man. *Hum. Gent.* 75:197-208.
- Mitsui, K., Nakagawa, T. and K. Tsurugi. 1989. The gene and the primary structure of acidic ribosomal protein A0 from yeast *Saccharomyces cerevisiae* which shows partial homology to bacterial ribosomal protein L10. *J. Biochem.* 106:223-227.
- Mosbaugh, D.W., and S. Linn. 1980. Further characterization of human fibroblast apurinic/apurimidinic DNA endonucleases. *J. Biol. Chem.* 255:11743-11752.

Mosbaugh, D.W., and S. Linn. 1982. Characterization of the action of *Escherichia coli* DNA polymerase I at incisions produced by repair endodeoxyribonucleases. *J. Biol. Chem.* 257:575-583.

Mudgett, J.S., and M.A. MacInnes. 1990. Cloning of the functional human excision repair gene ERCC-5: potential gene regulatory features conserved with other human repair genes. *Prog. Clin. Biol. Res.* 340:265-274.

Mullenders, L.H.F., Venema, J., Van Hoffen, A., Mayne, L.V., Natarajam, A.T., and A.A. van Zeeland. 1990. The role of the nuclear matrix in DNA repair. *Prog. Clin. Biol. Res.* 340A:223-32.

Myrnes, B., Guddal, P-H., and H. Krokan. 1982. Metabolism of dITP in HeLa cell extracts, incorporation into DNA by isolated nuclei and release of hypoxanthine from DNA by a hypoxanthine - DNA glycosylase activity. *Nucleic Acids Res.* 10:3693-3701.

Olsson, M., and T. Lindahl. 1980. Repair of alkylated DNA in *Escherichia coli*. *J. Biol. Chem.* 255:10569-10571.

Pannekoek, H., Noordermeer, I.A., van Sluis, C.A., and P. van de Putte. 1978. Expression of the *uvrB* gene of *Escherichia coli*: In vitro construction of a pMB9 *uvrB* plasmid. *J. Bacteriol.* 133:884-896.

Pelham, H.R.B. 1989. Heat shock and the sorting of luminal ER proteins. *EMBO J.* 8:3171-3176.

Pieper, R.O., Futscher, B.W., Dong, Q., Ellis, T.M., and L.C. Erickson. 1990. Comparison of O<sup>6</sup>-methylguanine DNA methyltransferase (MGMT) mRNA levels in Mer<sup>+</sup> and Mer<sup>-</sup> human tumor cell lines containing the MGMT gene by the polymerase chain reaction (PCR) technique. *Cancer Comm.* 2:13-20.

Pierre, J., and J. Laval. 1980a. *Micrococcus luteus* endonucleases for apurinic/apyrimidinic sites in deoxyribonucleic. 1. Purification and general properties. *Biochemistry* 19:5018-5023.

Pierre, J., and J. Laval. 1980b. *Micrococcus luteus* endonucleases for apurinic/apyrimidinic sites in deoxyribonucleic acid. 2. Further studies on the substrate specificity and mechanism of action. *Biochemistry* 19:5024-5029.

Pogue-Geile, K., Geiser, J.R., Shu, M., Miller, C., Wool, I, Meisler, A.I., and J.M. Pipas. 1991. Ribosomal protein genes are overexpressed in colorectal cancer: Isolation of a cDNA clone encoding the human S3 ribosomal protein. *Mol. Cell. Biol.* 11:3842-3849.

Popoff, S.C., Spira, A.I., Johnson, A.W. and B. Demple. 1990. Yeast structural gene (APN1) for the major apurinic endonuclease: Homology to *Escherichia coli* endonuclease IV. *Proc. Natl. Acad. Sci. USA* 87:4193-4197.

Raabe, T., and J.L. Manley. 1991. A human homologue of the *Escherichia coli* DnaJ heat-shock protein. *Nucleic Acids Res.* 19:6645.

Ramotar, D., Popoff, S.C., and B. Demple. 1991. Complementation of DNA repair-deficient *Escherichia coli* by the yeast APN1 apurinic/aprimidinic endonuclease gene. *Molec. Microbiol.* 5:149-155.

Reynolds, P., Higgins, D.R., Prakash, L., and S. Prakash. 1985. The nucleotide sequence of RAD3 gene of *Saccharomyces cerevisiae*: a potential adenine nucleotide binding amino acid sequence and a nonessential acidic carboxyl terminal region. *Nucleic Acids Res.* 13:2357-2372.

Rich, B.E., and J.A. Steitz. 1987. Human acidic ribosomal phosphoproteins P0, P1, and P2: analysis of cDNA clones, *in vitro* synthesis, and assembly. *Mol. Cell. Biol.* 7:4065-4074.

Richards, R.G., Sowers, L.C., Laszlo, J., and W.D. Sedwick. 1984. The occurrence and consequences of deoxy-uridine in DNA. *Adv. Enzyme Regul.* 22:157-185.

Richardson, C.C., and A. Kornberg. 1964. A deoxyribonucleic acid phosphatase-exonuclease from *Escherichia coli*. *J. Biol. Chem.* 239:242-250.

Richardson, C.C., Lehman, I.R., and A. Kornberg. 1964. A deoxyribonucleic acid phosphatase-exonuclease from *Escherichia coli*. *J. Biol. Chem.* 239:251-258.

Robson, C.N., Milne, A.M., Pappin, D.J.C., and I.D. Hickson. 1991. Isolation of cDNA clones encoding an enzyme from bovine cells that repairs oxidative DNA damage *in vitro*: homology with bacterial repair enzymes. *Nucleic Acids Res.* 19:1087-1092.

Rogers, S.G., and B. Weiss. 1980. Exonuclease III of *Escherichia coli* K-12, and AP endonuclease. *Methods Enzymol.* 65:201-211.

- Rupert, C.S., Goodgal, S.H. and R.M. Herriott. 1958. Photoreactivation *in vitro* of ultraviolet inactivated *Hemophilus influenzae* transforming factor. *J. Gen. Physiol.* 41:451-471.
- Rupp, W.D., and P. Howard-Flanders. 1968. Discontinuities in the DNA synthesized in an excision-deficient strain of *Escherichia coli* following ultraviolet irradiation. *J. Mol. Biol.* 31:291-304.
- Rupp, W.D., Wilde, C.E., Reno, D.L., and P. Howard-Flanders. 1971. Exchanges between DNA strands in ultraviolet irradiated *Escherichia coli*. *J. Mol. Biol.* 61:25-44.
- Russel, M., Kidd, S., and M.R. Kelley. 1986. An improved filamentous helper phage for generating single-stranded plasmid DNA. *Gene* 45:333-338.
- Rydberg, B., and T. Lindahl. 1982. Nonenzymatic methylation of DNA by intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* 1:211-216.
- Sakumi, K., Nakabeppu, Y., Yamamoto, Y., Kawabata, S., Iwanga, S., and M. Sekiguchi. 1986. Purification and structure of 3-methyladenine-DNA glycosylase I of *Escherichia coli*. *J. Biol. Chem.* 261:15761-15766.
- Sakumi, K., and M. Sekiguchi. 1990. Structures and functions of DNA glycosylases. *Mutat. Res.* 236:161-172.
- Sammartano, L.J., Tuveson, R.W., and R. Davenport. 1986. Control of sensitivity to inactivation by H<sub>2</sub>O<sub>2</sub> and broad-spectrum near-UV radiation by the *Escherichia coli* *katF* locus. *J. Bacteriol.* 168:13-21.
- Sancar, A., and W.D. Rupp. 1983. A novel repair enzyme: UVRABC excision nuclease of *Escherichia coli* cuts a DNA strand on both sides of the damaged region. *Cell* 33:249-260.
- Sancar, A., and G.B. Sancar. 1988. DNA repair enzymes. *Ann. Rev. Biochem.* 57:29-67.
- Sanchez-Madrid, F., Reyes, R., Conde, P., and J.P.G. Ballesta. 1979. Acidic ribosomal proteins from eukaryotic cells: effect on ribosomal functions. *Eur. J. Biochem.* 98:409-416.
- Sanchez-Madrid, F., Vidales, F.J. and J.P.G. Ballesta. 1981. Functional role of acidic ribosomal proteins. Interchangeability of proteins from bacterial and eukaryotic cells. *Biochemistry* 20:3263-3266.

Sander, M., Lowenhaupt, K., and A. Rich. 1991. *Drosophila* Rrp1 protein: an apurinic endonuclease with homologous recombination activities. *Proc. Natl. Acad. Sci. USA* 88:6780-6784.

Sanderson, B.J.S., Chang C.-N., Grollman, A.P., and W.D. Henner. 1989. Mechanism of DNA cleavage and substrate recognition by a bovine apurinic endonuclease. *Biochemistry* 28:3894-3901.

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

Seal, G., and M.A. Sirover. 1986. Physical association of the human base-excision repair enzyme uracil glycosylase with the 70,000-dalton catalytic subunit of DNA polymerase  $\alpha$ . *J. Biol. Chem.* 261:7608-7612.

Seki, S., Akiyama, K., Watanabe, S., Hatsushika, M., Ikeda, S. and K. Tsutsui. 1991a. cDNA and deduced amino acid sequence of a mouse DNA repair enzyme (APEX nuclease) with significant homology to *Escherichia coli* exonuclease III. *J. Biol. Chem.* 266:20797-20802.

Seki, S., Ikeda, S., Watanabe, S., Hatsushika, M., Tsutsui, K., Akiyama, K., and B. Zhang. 1991b. A mouse DNA repair enzyme (APEX nuclease) having exonuclease and apurinic/apyrimidinic endonuclease activities: purification and characterization. *Biochem. Biophys. Acta* 1079:57-64.

Setlow, R.B. 1982. DNA repair, aging and cancer. *Natl. Cancer Inst. Monogr.* 60:249-255.

Setlow, R.B., and W.L. Carrier. 1964. The disappearance of thymine dimers from DNA: an error-correcting mechanism. *Proc. Natl. Acad. Sci. USA* 51:226-231.

Shaper, N.L., Grafstrom, R.H., and L. Grossman. 1982. Human placental apurinic/apyrimidinic endonuclease. *J. Biol. Chem.* 257:13455-13458.

Snow, E.T., Foote, R.S., and S. Mitra. 1984. Base-pairing properties of O<sup>6</sup>-methylguanine in template DNA during *in vitro* DNA replication. *J. Biol. Chem.* 259:8095-8100.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.



Spiering, A.L., and W.A. Deutsch. 1981. Apurinic DNA endonucleases from *Drosophila melanogaster* embryos. *Mol. Gen. Genet.* 183:171-174.

Spiering, A.L., and W.A. Deutsch. 1986. *Drosophila* apurinic/aprimidinic DNA endonucleases. *J. Biol. Chem.* 261:3222-3228.

Steward, R. 1987. *Dorsal*, an embryonic polarity gene in *Drosophila*, is homologous to the vertebrate proto-oncogene *c-rel*. *Science* 238:692-694.

Strauss, D.B., Walter, W.A., and C.A. Gross. 1988. *Escherichia coli* heat shock gene mutants are defective in proteolysis. *Genes and Dev.* 2:1851-1858.

Teoule, R. 1987. Radiation-induced DNA damage and its repair. *Int. J. Radiat. Biol.* 51:573-589.

Troelstra, C., Odijk, H., De Wit, W.J., Westerveld, A., Thompson, L.H., Bootsma, D., and J.H.J. Hoeijmakers. 1990. Molecular cloning of the human DNA excision repair gene *ERCC-6*. *Mol. Cell. Biol.* 10:5806-5813.

Tushinski, R.J., J.R. Warner. 1982. Ribosomal proteins are synthesized preferentially in cells commencing growth. *J. Cell. Physiol.* 112:128-135.

Uchiumi, T., Traut, R.R., and R. Kominami. 1990. Monoclonal antibodies against acid phosphoproteins P0, P1, and P2 of eukaryotic ribosomes as functional probes. *J. Biol. Chem.* 265:89-95.

van Duin, M., de Wit, J., Odijk, H., Westerveld, A., Yasui, A., Koken, M.H.H., Hoeijmakers, J.H.J., and D. Bootsma. 1986. Molecular characterization of the human excision repair gene *ERCC-1*: cDNA cloning and amino acid homology with the yeast DNA repair gene *RAD10*. *Cell* 44:913-923.

van Houten, B. 1990. Nucleotide excision repair in *Escherichia coli*. *Microbiol. Rev.* 54:18-51.

Verly, W.G., and E. Rassart. 1975. Purification of *Escherichia coli* endonuclease specific for apurinic sites in DNA. *J. Biol. Chem.* 250:8214-8219.

Von Sonntag, C., and H.P. Schuchmann. 1986. The radiolysis of pyrimidines in aqueous solutions: an updating review. *Int. J. Radiat. Biol.* 49:1-34.

- Wallace, S.S. 1988. AP endonucleases and DNA glycosylases that recognize oxidative DNA damage. *Environ. Mol. Mut.* 12:431-477.
- Ward, J.F. 1989. DNA damage in mammalian cells. *Free Radic. Res. Commun.* 6:179-180.
- Warner, H.R., Demple, B.F., Deutsch, W.A., Kane, C.M., and S. Linn. 1980. Apurinic/aprimidinic endonucleases in repair of pyrimidine dimers and other lesions in DNA. *Proc. Natl. Acad. Sci. USA* 77:4602-4606.
- Weber, C.A., Salazar, E.P., Stewart, S.A. and L.H. Thompson. 1988. Molecular cloning and biological characterization of a human gene, *ERCC2*, that corrects the nucleotide excision repair defect in CHO UV5 cells. *Mol. Cell. Biol.* 8:1137-1146.
- Weber, C.A., Salazar, E.P., Stewart, S.A., and L.H. Thompson. 1990. *ERCC2*: cDNA cloning and molecular characterization of a human nucleotide excision repair gene with high homology to yeast *RAD3*. *EMBO J.* 9:1437-1447.
- Weeda, G., van Ham, R.C.A., Vermeulen, W., Bootsma, D., van der Eb, A.J. and J.H.J. Hoeijmakers. 1990. A presumed DNA helicase encoded by *ERCC-3* is involved in the human repair disorders xeroderma pigmentosum and Cockayne's syndrome. *Cell* 62:777-791.
- Weiss, B. 1976. Endonuclease II of *Eschericia coli* is exonuclease III. *J. Biol. Chem.* 251:1896-1901.
- Weiss, B., and L. Grossman. 1987. Phosphodiesterases involved in DNA repair. *Adv. Enzymol.* 60:1-33.
- Westerveld, A., Hoeijmakers, J.H.J., van Duin, M., de Wit, J., Odijk, H., Pastink, A., Wood, R.D., and D. bootsma. 1984. Molecular cloning of a human DNA repair gene. *Nature* 310:425-429.
- Wharton, K.A., Yedvobnick, B., Finnerty, V.G., and S. Artavanis-Tsakonas. 1985. *opa*: A novel family of transcribed repeats shared by the *Notch* locus and other developmentally regulated loci in *D.melanogaster*. *Cell* 40:55-62.
- Wickner, S.H. 1990. Three *Eschericia coli* heat shock proteins are required for P1 plasmid DNA replication: formation of an active complex between *E.coli* *dnaJ* protein and the P1 initiator protein. *Proc. Natl. Acad. Sci. USA* 87:2690-2694.

Wu, C., Bingham, P.M., Livak, K.J., Holmgren, R., and S.C.R. Elgin. 1979. The chromatin structure of specific genes: Evidence for higher order domains of defined DNA sequence. *Cell* 16:797-806.

Zylicz, M., Yamamoto, T., McKittrick, N., Sell, S. and C. Georgopoulos. 1985. Purification and properties of the dnaJ replication protein of *Escherichia coli*. *J. Biol. Chem.* 260:7591-7598.

## VITAE

David Thomas Grabowski is the son of Barbara A. and Ted C. Grabowski. He was born June 11, 1964 in Milwaukee, Wisconsin.

His elementary education was obtained at Blessed Sacrament School in Milwaukee Wisconsin and his secondary education was obtained at Thomas More High School in St. Francis, Wisconsin, where he graduated in 1982.

In August 1982, he entered Marquette University in Milwaukee, Wisconsin and graduated in May 1987, with a Bachelor of Science in Chemistry.

In August 1987, he joined the Department of Molecular and Cellular Biochemistry at Loyola University of Chicago, Maywood, Illinois and was awarded a Loyola University Basic Sciences Fellowship. In January of 1988, he joined the laboratory of Dr. Mark R. Kelley. In 1991, he was awarded a Dissertation Fellowship by the Arthur J. Schmitt Foundation.

In August 1992, he will begin a postdoctoral fellowship in the laboratory of Dr. Sankar Mitra in the John Sealy Center for Molecular Sciences at the University of Texas Medical Branch at Galveston in Texas.

ABSTRACTS

1. Kelley, M.R., Grabowski, D.T. and Deutsch, W.A. (1989) Molecular characterization of a cloned DNA repair gene in *Drosophila*, mice, and humans. 30th Annual *Drosophila* Research Conference, New Orleans, LA.
2. Grabowski, D.T., Derda, D., Kelley, M.R. (1989) DNA repair in *Drosophila*. 1989 Midwest *Drosophila* Conference, Allerton Park, IL.
3. Kelley, M.R., Grabowski, D.T., Derda, D. and Deutsch, W.A. (1990) Biochemical analysis of DNA repair gene AP3. 31st Annual *Drosophila* Research Conference, Asilomar, CA.
4. Grabowski, D.T., Derda, D., Deutsch, W.A., Kelley, M.R. (1991). Molecular and biochemical analysis of a presumptive DNA repair gene with homology to a human ribosomal protein, PO. 32nd Annual *Drosophila* Research Conference, Chicago, IL.
5. Grabowski, D.T., Carney, J.P., Kelley, M.R. (1991). Cloning and molecular characterization of an adult male specific gene in *Drosophila* containing the repetitive element "Opa". 32nd Annual *Drosophila* Research Conference, Chicago, IL

PUBLICATIONS

1. Grabowski, D.T., Carney, J.P. and Kelley, M.R. (1991) A *Drosophila* gene containing the opa repetitive element is exclusively expressed in adult male abdomens. *Nucleic Acids Res* 19: 1709.
2. Grabowski, D.T., Carney, J.P. and Kelley, M.R. (1991) An adult male specific gene in *Drosophila* containing the repetitive element opa. *Biochemica et Biophysica Acta* 1090: 115-118.
3. Grabowski, D.T., Deutsch, W.A., Derda, D. and Kelley, M.R. (1991) *Drosophila* AP3, a presumptive DNA repair protein, is homologous to human ribosomal associated protein PO. *Nucleic Acids Res* 19: 4297.
4. Grabowski, D.T., Pieper, R.O., Futscher, B.W., Deutsch, W.A., Erickson, L.C. and Kelley, M.R. Expression of ribosomal phosphoprotein PO is induced by antitumor agents and increased in Mer<sup>r</sup> human tumor cell lines. *Carcinogenesis* 13: 259-263.
5. Grabowski, D.T. and Kelley, M.R. Genomic organization and splicing pattern of the *Drosophila* AP endonuclease DNA repair gene. In preparation.

6. **Grabowski, D.T., Carney, J.P., Derda, D. and Kelley, M.R.** Cellular location and protein characterization of the *Drosophila* AP endonuclease gene AP3. In preparation.
7. **Grabowski, D.T., Wilson, D.M., Carney, J.P. and Kelley, M.R.** Cloning of the *Drosophila* homologue to the *E. coli* dnaJ heat shock and DNA replication gene. In preparation.

APPROVAL SHEET

The dissertation submitted by David Thomas Grabowski has been read and approved by the following committee:

Dr. Mark R. Kelley, Director  
Assistant Professor  
Department of Molecular and Cellular Biochemistry  
Assistant Professor  
Department of Medicine  
Loyola University of Chicago

Dr. Allen Frankfater  
Associate Professor  
Department of Molecular and Cellular Biochemistry  
Loyola University of Chicago

Dr. William H. Simmons  
Associate Professor  
Department of Molecular and Cellular Biochemistry  
Loyola University of Chicago

Dr. Leonard C. Erickson  
Professor  
Department of Medicine and Hematology/Oncology  
Loyola University of Chicago

Dr. Deborah K. Hoshizaki  
Assistant Professor  
Department of Biochemistry  
University of Illinois, Chicago

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is, therefore, accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Date

July 6, 1992

Director's Signature

Mark R Kelley