THE UNIVERSITY OF CALGARY

ANALYSIS OF *DROSOPHILA Pkc53E* VIA NOVEL GENE TARGETING

by

Kenton G. Hecker

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCES

CALGARY, ALBERTA

JANUARY, 1997

Abstract

Drosophila melanogaster has long been used as a model organism to answer *in vivo* genetic and developmental questions. Of the many molecular tools used, the characterization and use of P-elements for transgenic studies is by far the most powerful. This study adapts the P-element mediated gene targeting technique to the *white* locus by creating a nontransposable reporter vector that is used to perform *in vivo* transcriptional analysis of the *Drosophila* gene, *Pkc53E*.

The temporal and spatial transcriptional regulation is still poorly understood. To determine the basal promoter region of *Drosophila Pkc53E*, the upstream regulatory region was first subcloned and sequenced. Putative regulatory elements were identified through sequence analysis. The putative basal promoter fragment was subcloned into the newly constructed nontransposable reporter template and gene targeting experiments were performed. Of the 4508 male progeny screened, one contained the revertant eye colour indicating an insertion event.

Acknowledgments

I would first like to thank Dr. Mike Bentley for allowing me to experience Science at my own speed and to learn in my very own unique style. Not many people would have put up with my antics and my eccentricities, they are what keep me sane, thanks Mike. There are numerous people who have helped me over the years, who haved laughed with me and at me, for their support and advice I am forever grateful. They are Lauryl Nutter, Pat Ouimet, Mike Carpenter, Jill Johnston, Alisa Piekny and many others. There has also been the help of many supportive professors, who, I found out are human too! Thanks to Fred Lohka, Manju Kapoor, Paul Mains, Ken Sanderson, Lois Edwards and Leon Browder.

My friends, ahhhh, without whom I would have never made it through. They prodded me and pushed me to experience life outside of the lab. They would harass me when my flies would not work! To my two best friends Phil and Jamie, who kept me in check for so long and who kept me sane. I was fortunate enough to experience Calgary and the mountains, what a beautiful area, thanks guys!

Finally, to my family, there support, love and laughter kept me going throughout. This is dedicated to you!!! LOVE YOU ALL.

Table of Contents

List of Figures

List of Tables

Abbreviations

Adh	alcohol dehydrogenase
Antp	*Antennapedia*
Ap1	activator protein 1
Ap2	activator protein 2
bp	base pair
Ca^{2+}	calcium ion
CAT	chloramphenicol acetyltransferase
CIAP	calf intestinal alkaline phosphatase
CNS	central nervous system
DG	diacylglycerol
DIG	digoxigenin
DSB	double stranded break
DNA	deoxyribonucleic acid
DPE	downstream promoter element
en	*engrailed*
EtBr	ethidium bromide
f	*forked*: bristle mutation in *Drosophila*
g	gravitational force
GD	gonadal dysgenic sterility
Inr	initiator sequence
LacZ	β-galactosidase
kb	kilobase
PKC	Protein kinase C
aPKC	atypical PKC
cPKC	classical PKC
nPKC	novel PKC
mPKC	mammalian PKC
ml	milliliter
MCS	multiple cloning site
ng	nanogram
p.e.v.	position effect variegation
pfu	plaque forming units
PS	phosphatidylserine

psi	pounds per square inch
Pkc53E	*Drosophila* PKC
ry	*rosy;* eye colour mutation in *Drosophila*
SDSA	synthesis dependent strand annealing
Sp1	stimulatory protein 1
TAF	TATA box box-binding asscociated factors
TBP	TATA box binding protein
TPA	12-*O*-tetradecanoylphorbol-13-acetate
TRE	TPA- responsive element
T_3	T_3 bacteriophage RNA polymerase promoter
T_7	T_7 bacteriophage RNA polymerase promoter
P[ry$^+$: Δ2-3(99B)]	stock which has a stable P-element transposase on the right arm of chromosome 3
PNS	peripheral nervous system
w	*white;* eye colour mutation in *Drosophila*
w^{hd}	$w^{hd80k17}$: white allele resulting from a P-element insertion in exon 6 of the *white* gene
y	*yellow*; body colour mutation in *Drosophila*
µg	microgram
µl	microliter
v/v	volume/volume
w/v	weight/volume

Introduction

Transcriptional regulation studies have given important clues as to which genes are necessary for the development of organisms. Mammalian promoter sequences have been primarily characterized using promoter-reporter constructs in cell culture. While these studies answer some basic questions about transcriptional controls of specific genes they do not fully elucidate the developmental regulation of these genes.

Drosophila melanogaster, *Mus musculus* and *Caenorhabditis elegans* have been used as model organisms to study developmental questions from a molecular standpoint using molecular tools such as promoter-reporter constructs, enhancer traps, plasmid rescue vectors and transposable elements. In *Drosophila* the characterization and utilization of transposable elements has allowed researchers to answer many important developmental questions. Initial transposable element studies examined phenotypic changes that were due to altered gene expression as a result of the mobilization and insertion of endogenous P-elements (reviewed by Engels, 1989). In such studies the quasirandom P-element integration was perceived as an advantage. Researchers were then able to further modify transposable elements by adding known DNA sequences and introduce these constructs into the genome of *Drosophila* creating transgenic flies (Rubin and Spradling, 1982; Spradling and Rubin, 1982). A major drawback of the transposable elements for *in vivo* promoter studies is their quasirandom integration into the genome. This trait leads to the problem of position effect variegation (p.e.v: location of insert can determine the level of expression) and the effects of endogenous silencers and enhancers on the promoter-reporter construct, thus placing constraints on the interpretation of results and the type of experiments that can be performed.

In an attempt to understand how P-elements excise and the mechanisms by which the resulting DNA gap repairs itself, investigators discovered that P-element excision is

nonreplicative and repair is dependent upon the sequences flanking the break site (Engels *et al.*, 1990; Gloor *et al.*, 1991; Nassif *et al.*, 1994). Repair of these gaps can occur using template DNA sequences that are homologous to the region surrounding the site of excision. Eighty-five percent of the repair events use the sister chromatid as a template resulting in the net gain of one P-element (Engels *et al.*, 1990). The remaining 15% of repair events use sequences present within the genome that are homologous to the flanking sequence where the break occurred (Engels *et al.*, 1990; Gloor *et al.*, 1991; Nassif *et al.*, 1994). These findings lead to the possibility that constructed templates containing sequences homologous to the regions surrounding endogenous P-elements could be targeted to these known endogenous P-element sites after excision. This gene targeting mechanism has the potential to be a powerful tool for doing *in vivo* studies because of the ability to consistently target a construct to an exact known chromosomal location. For *in vivo* transcriptional studies, this would allow for consistent, reproducible results with a promoter fragment of interest and for direct comparison of different promoter sequences.

This project is aimed at the identification of the basal promoter of the *Drosophila* gene *Pkc53E*. To determine the basal promoter *in vivo*, the gene targeting method to the *white* locus was employed. Gene targeting allows for the specific targeting of a new promoter reporter construct, created and defined in this project, to a defined region within the *Drosophila* genome. This provides a mechanism to directly compare the effects of putative *Pkc53E* promoter fragments on expression of a reporter gene. This technique has not been used for *in vivo* transcriptional analyses to date. Therefore, part of this project involves the creation of a reporter vector that is specific for targeting constructs to the *white* locus in *Drosophila*.

Structure of P-elements

P-elements were first recognized as the cause of P-M hybrid dysgenesis (reviewed by Craig, 1990; Rio, 1990; Engels, 1996). P-M hybrid dysgenesis is a result of the mobilization of the P-elements in offspring when P strain males (bearing both autonomous and defective elements) are crossed to females lacking any transposable elements (M strain). Hybrid dysgenesis is characterized by temperature-dependent sterility, chromosome rearrangement, recombination and elevated rates of mutation (Kidwell and Sved, 1977). Autonomous P-elements are 2907 bp transposable elements composed of 31 base pair (bp) repeats flanking four exons which encodes for two proteins. A 66 kDa repressor protein prevents mobilization of the P-element in the soma of P-strains and an 87 kDa transposase protein is expressed in the gonads of P-strains (Karess and Rubin, 1986; Misra and Rio, 1990). Nonautonomous P-elements are believed to have arisen naturally from internal deletions disrupting the transposase source but retaining the sequences needed for transposition. P-elements utilized in transgenic research lack the internal sequences needed for the production of transposase but maintain the 31 bp inverted repeats and adjacent protein binding sites. These crippled P-elements rely on an exogenous source of transposase for mobilization and insertion.

Germline transformation vectors (first reported by Rubin and Spradling, 1982 and Spradling and Rubin, 1982) use defective P-elements with visible markers to cotransform DNA sequences of interest into the genome of *Drosophila*. Visible markers allow for the identification of construct integration. A visible marker that is commonly used is the wild type sequence of the *white* gene (w^+). When used in a *white* (w) genotypic background, integration of the vector is identified by a reversion in eye colour from the w background (white eye colour) to a wild type red eye colour. Vectors are injected into syncytial embryos that contain either an endogenous transposase source, for instance the stable insert P[ry^+; Δ2-3(99B)] on the third chromosome identified by a dominant marker (Robertson *et al.*, 1988), or coinjected with a transposase making "helper" plasmid that cannot itself

integrate into the chromosome (Karess and Rubin 1984). If the transposase source is endogenous, transformant progeny are selected that do not contain the transposase encoding chromosome.

Δ2-3 refers to the loss of the germ-line specific intron from an intact P-element resulting in a functional transposase gene in germ-line and somatic cells causing mobilization in all tissues (Laski *et al.*, 1986). This Δ2-3 construct, as mentioned above can be introduced as the transposase source in two ways: either as a stable endogenous source, or as an extrachromosomal plasmid that does not integrate. Success using Δ2-3 as a transposase source for transformations is well documented (reviewed by Engels, 1987) and the effects of somatic mobilization of P-elements have been well studied (Engels *et al.*, 1987). The mobilization of P-elements in the somatic cells has been shown to have a number of detrimental results ranging from pupal lethality and gonadal dysgenic (GD) sterility to a reduced lifespan. So, while there are numerous benefits for using Δ2-3 as a transposase source there are also a number of drawbacks. These drawbacks encourage the rapid removal of the Δ2-3 bearing chromosome from any experimental stock before the flies can be studied further.

Transcriptional Studies using P-elements

Transcriptional regulation of *Drosophila* genes has been extensively studied *in vitro* (Soeller *et al.*, 1988; Garber *et al.*, 1994; Thummel, 1989; Chao and Pellegrini, 1993; Gieser *et al.*, 1993; Hu *et al.*, 1995) and *in vivo* (Meyerowitz *et al.*, 1987; Fischer and Maniatis, 1988; Hawley *et al.*, 1992; Fridell and Searles, 1992; Ludwig *et al.*, 1993; Schier and Gehring, 1993; Buttgeriet and Renkawitz-Pohl, 1993; Vallett *et al.*, 1993; Gunaratne *et al.*, 1994; Lankenau *et al.*, 1996). The majority of the transcriptional studies that have been performed *in vivo* have utilized P-element based non autonomous vectors that contain a reporter gene, either β-galactosidase (LacZ) or chloramphenicol acetyl transferase (CAT).

and a phenotypic marker (e.g., w^+). The process by which these constructs are incorporated in the genome is as follows: The promoter-reporter P-element construct is microinjected into embryos with the appropriate genetic background and a source of transposase. Insertion of the P-element construct is determined by a visible phenotypic change of the adult fly, e.g., white eyes to red eyes. The revertant flies are mated with a stock that contained an endogenous transposase source (Δ2-3) in order to mobilize the P-elements and have them integrate randomly into other sites in the genome. Revertant progeny are collected that contains the construct but not the transposase source (identified as the absence of the dominant marker for the endogenous source). These are subsequently mated and lines are established. In order to determine the physical location of the P-element, *in situ* hybridizations to the polytene chromosomes are performed. Flies containing unique P-element insertions are then bred to create strains that have one unique P-element insertion. Once a P-element has been localized in each strain, then experiments analyzing the transcriptional regulation of the gene of interest could begin.

Numerous examples exist in the literature of studies using promoter fragments and reporter constructs that display great variability within samples containing the same construct. Ludwig *et al.* (1993) studied six promoter-reporter constructs. For each construct, 6-12 independently transformed stocks were generated that contained one construct insertion. To control for position effect on *lacZ* expression, studies were performed on at least three different strains, each strain containing the construct on a different chromosome. Even with all the controls, *lacZ* expression varied between strains containing the same construct with different insert sites. These researchers justified this observation by stating that they were only interested in the qualitative expression pattern and not the amount of expression of the reporter gene (Ludwig *et al.*, 1993). Buttgereit *et al.* (1993) studied 11 promoter reporter constructs, the number of strains that were examined for each of the constructs were between 2/construct to as many as 11/construct.

There studies were focused on studying the expression patterns of various deletions of the *Drosophila* β1 tubulin promoter. The expression of the *lacZ* reporter gene varied for each construct depending upon location of insert. They measured *lacZ* expression between constructs on a very subjective scale ranging from very strong expression (15 fold increase in staining of tissues) to variable expression (1-2 fold increase in staining of tissues). These two studies as well as others (Meyerowitz *et al.*, 1987; Fischer and Maniatis, 1988; Hawley *et al.*, 1992; Fridell and Searles, 1992; Schier and Gehring, 1993; Vallett *et al.*, 1993; Gunaratne *et al.*, 1994) all cite that randomness of integration leading to p.e.v. does not allow for a complete understanding of the expression patterns observed.

Targeted Gene Replacement

Genetic recombination in eukaryotes allows for the movement of genetic information between chromosomes and thus diversifies the genetic pool. There are three fundamental models of genetic recombination in eukaryotes: The Holliday model (Holliday 1964, 1968); the Meselson-Radding Model (Meselson and Radding 1975); and the Double-Strand-Break Repair Model (Szostak, 1983). While the first two models explain homologous recombination, the passing of genetic material between two intact chromosomes, the last model explains the mechanism by which a DNA gap repairs itself via genetic recombination. Studies undertaken to understand this last model have been called the "genetics of genetics", and have been extensively characterized in *Drosophila* (Suzuki *et al.*, 1986; reviewed by Lankenau, 1995).

Transposable elements move in either a replicative or a non-replicative manner. Replicative transposition leads to insertion of a transposable element at a new genomic site without removing the element from its original site. Nonreplicative transposition involves the excision and subsequent reinsertion into a new genomic location, creating a double stranded break (DSB) at the site of excision. Experiments have now shown that P-

elements move in a non-replicative or a "cut and paste" manner thus leaving a DSB at the point of excision (Engels *et al.*, 1990). This break is repaired by the molecular processes of homologous recombination. Focus on the repairing of this DSB has lead to a new model called the "Synthesis Dependent Strand Annealing" model (SDSA model; Nassif *et al.*, 1994). The SDSA model (Figure 1) is an adaptation of the DNA repair model proposed by Formosa and Alberts (1986) which addresses the end dependent repair of bacteriophage T4 in T4-infected *Escherichia coli* cells. Repair of DSBs is initiated by each of the two 3' termini locating and invading an homologous DNA sequence. The majority of the repairs, 85%, use the sister chromatid as the template while the remaining 15% use the homolog and/or ectopic sequence (Engels et al., 1990). According to this model the ends can identify and invade two different homologous sequences and synthesis proceeds with the newly formed strand being displaced from the template. The new strands produced will contain regions of homology and will combine in this region of overlap with non-homologous sequences being removed beyond the overlap. Finally, localized DNA synthesis occurs to fill in any gaps using the sister strand as a template.

Requirements for successful repair of the DSBs using introduced ectopic templates at a high frequency are two fold. The amount of homologous sequence needed for strand invasion and synthesis consists of a few hundred base pairs and, more importantly, the sequence should be identical to the genomic sequence where the break occurred (Nassif and Engels, 1993). Secondly, ectopic templates located in cis configuration to the break site are copied at higher frequencies than those found in trans or on other chromosomes (Gloor *et al.*, 1991; Engels *et al.*, 1994; Keeler *et al.*, 1996).

P-element mediated targeted gene replacement (Engels *et al.*, 1990; Gloor *et al.*, 1991) was initially performed at the X-linked *white* locus. The genotype, $w^{hd80k17}$ (w^{hd}) is a result of a 629 bp P-element insertion into the coding region of exon 6 of the *white* gene giving a bleached *white* eye phenotype (Rubin *et al.*, 1982). Gloor *et al.*

(1991) introduced ectopic *white* templates (Figure 2) through P-element transposition into the genome of *Drosophila* containing the w^{hd} allele. The location of the P-element based vector template was determined by *in situ* hybridization to polytene chromosomes. Subsequent crosses were performed to induce a double stranded break at the w^{hd} locus by mobilizing this 629 bp P-element. Gene targeting and repair of the DSB is described in Figure 1. In the context of the DSB created at the *white* locus, the P-element shown corresonds to the 629 bp P-element and the template shown for repair corresponds to the *white* template shown in Figure 2. For repair using the*white* template, the ends of the genomic DNA, after undergoing exonuclease digestion, must identify regions of homology in the template and use this sequence for the synthesis and repair of the DSB thereby incorporating the template sequence into the genome. Progeny of these crosses showing a revertant red eye phenotype were analyzed by Southern blots and polymerase chain reaction (PCR) to determine if the sequence from the ectopic *white* template was used to repair the DSB. The results showed that the template can be copied at a high frequency (1% of all the progeny had the template) with a low mutation rate, almost the same as a single-base substitution at the same location.

Further alteration of the template so it contained nearly 8000 bp of non homologous DNA was found to be copied into the DNA double stranded gap at the *white* locus at the same frequency and same mutation rate (Gloor *et al.*, 1991; Nassif *et al.*, 1994; Keeler *et al.*, 1996). Heterologous DNA was cloned into a multiple cloning site (MCS) in the region of the *white* gene template sequence corresponding to the poly-A tail (Figure 2; Figure 1 n). The mechanism by which incorporation occurs is believed to be a result of the exonuclease digestion of the genomic DNA ends (Figure 1). If there is sufficient digestion of the genomic DNA then synthesis can occur because the 3' genomic end will recognize homologous *white* sequence further downstream from the insertion site of the non

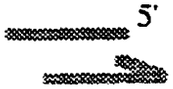
Figure 1. The Synthesis-Dependent Strand Annealing (SDSA) model for DNA gap repair at the *white* locus. This model of repair proposes that after a P-element inserts (A) it can be mobilized leaving a double stranded break (B). The 3' ends after exonuclease degradation can independently search, invade, displace and act as primers for DNA synthesis from a template that can contain non homologous DNA (C) through a bubble migration mechanism. The bubble collapses behind the DNA polymerase by displacement of the newly formed strand. If this displacement is slower than or is at the same rate as the synthesis then the majority of the gap will be repaired from the template sequence and there will be incorporation of the non homologous DNA. The strands will then reanneal (D) and the remaining gap be filled by using the opposite strand as a template (E). Not drawn to scale (modified from Nassif *et al.*, 1994)

- 629 bp w^{hd} P-element in the *white* locus

Drosophila chromosome

extrachromosomal template

region of extrachromosomal template homologous to the site of the double stranded break used for repair



3' vs 5' ends

▲ site corresponding to insertion site of non homologous DNA in the *Drosophila* chromosome (refer to Figure 2)

Δ site corresponding to w^{hd} location in the extrachromosomal template (refer to Figure 2)

non homologous DNA in the template that is copied into the *white* locus

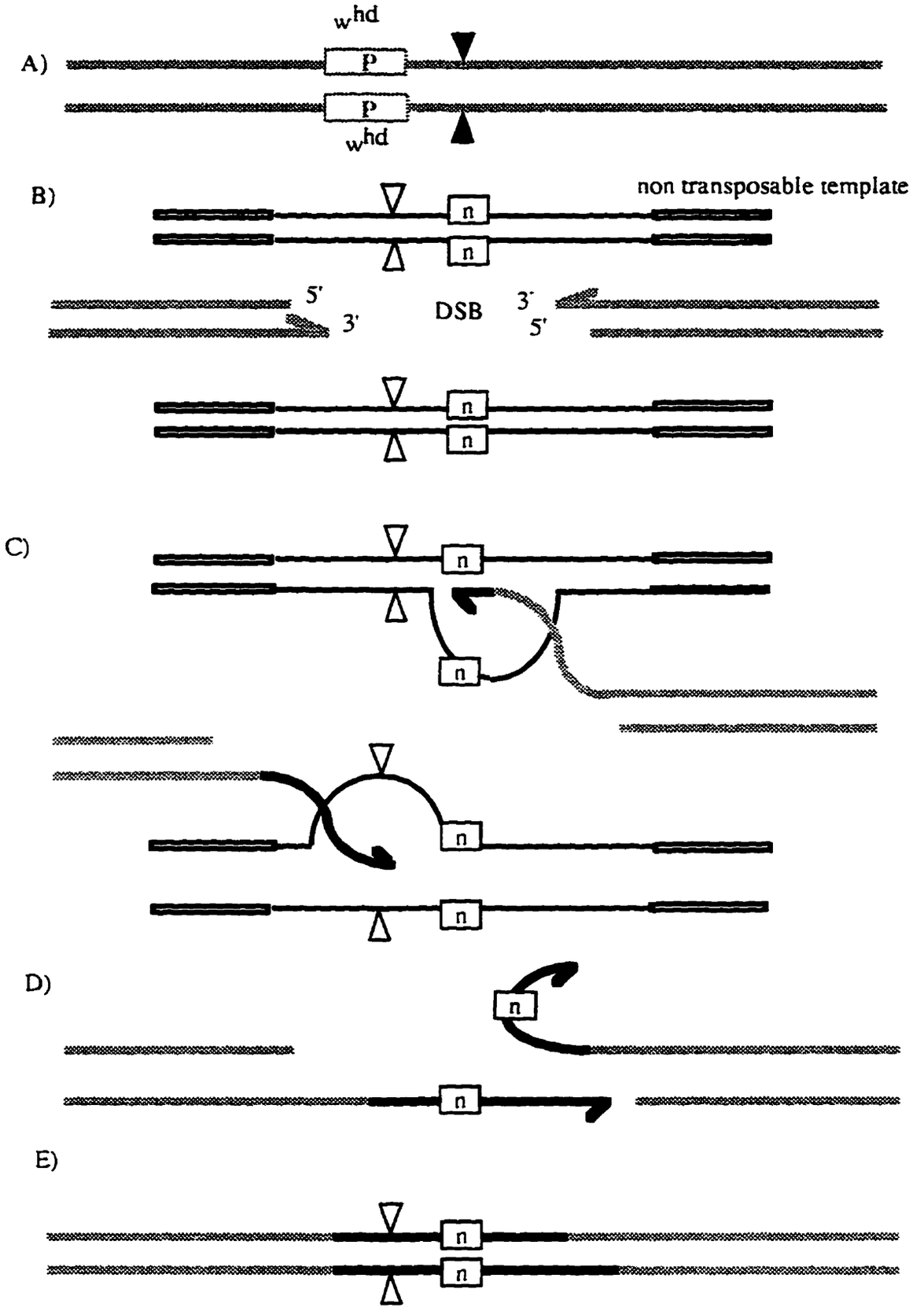
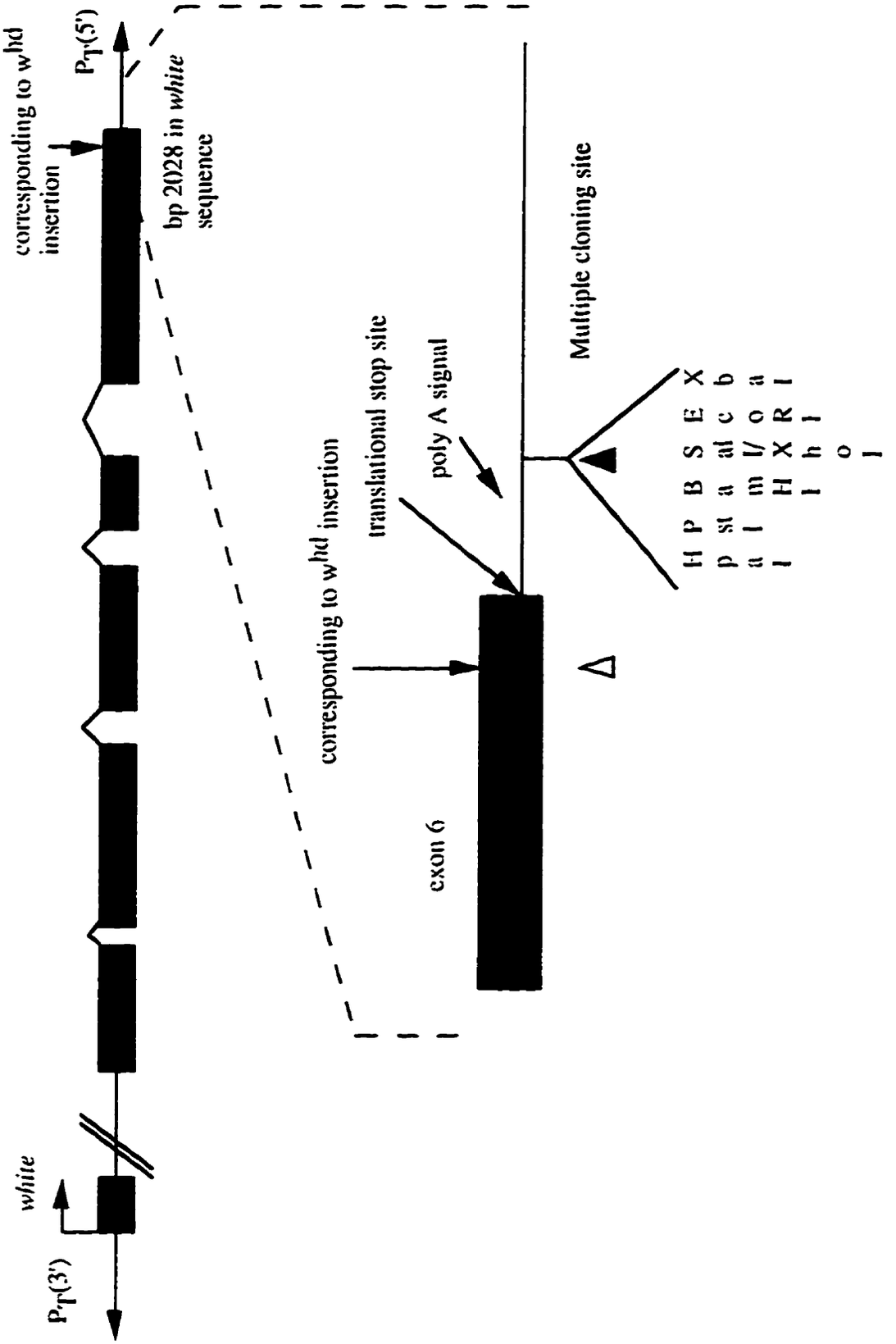

A)
w^{hd}
P
P
w^{hd}
B)
non transposable template
n
n
5'
3'
DSB
3'
5'
n
n
C)
n
n
n
n
D)
n
n
E)
n
n

Figure 2. The *white* gene targeting template (4,292 bp) for the *white* locus. The black boxes indicate exons of the *white* gene. The enlarged region depicts the w^{hd} P-element insertion site in exon 6 of the *white* template. There is a 30 bp MCS inserted at postion 238 incorporating a number of unique restriction cut sites that are 34 nucleotides downstream of the 3' end of the *white* gene. This multiple cloning site was used to insert heterologous fragments of DNA into the *white* locus, without disrupting *white* function. P_T(3') and P_T(5') are the 3' and 5' P element ends respectively (modified from Nassif *et al.*, 1994; Gloor *et al.*, 1991).

Δ site corresponding to w^{hd} in the *white* template

▲ insertion site (MCS) of non homologous DNA in to the *white* template


white
$P_T(3')$
corresponding to w^{hd} insertion
$P_T(5')$
bp 2028 in white sequence
exon 6
corresponding to w^{hd} insertion
translational stop site
poly A signal
Multiple cloning site
H P B S E X
p st a al c b
a I m I/ o a
I H X R I
I h I
o
I

homologous DNA. This allows for the wildtype template and the non homolgous DNA to be used as a template for repair and incorporation. This template and technique targets heterologous DNA consistently and reproducibly to a specific locus, effectively negating the randomness of insertion exhibited by P-element vectors.

Templates for gene conversion have been of two types, transposable and non transposable. Transposable templates are non autonomous P-element based vectors that insert randomly into the genome. Once the templates has been localized by *in situ* hybridization then gene targeting is induced by a series of crosses. As with the typical *in vivo* P-element transcriptional studies, a number of crosses have to be performed before the actual experiments begin. Non transposable templates do not contain the 31-bp inverted terminal repeats and do not insert into the genome (Banga and Boyd, 1992; Keeler *et al.*, 1996). This method has only been used recently to target heterologous fragments of DNA to the *white* locus. Keeler *et al.* (1996) and Banga and Boyd (1992) reported that this type of template inserts at a lower, but still detectable, frequency. The benefit of using a non transposable template is that it reduces the number of crosses that need to be performed to get integration at the locus. The one drawback that has been reported is the presence of duplications of the template DNA in the targeted site (Keeler *et al.*, 1996). This is believed to be the result of the repair machinery reading around the repair template. All duplications had the same structure of template-vector DNA-template.

The frequency of repair at the *white* locus using exogenous templates was dependent upon the location of the template. If the template was autosomal, reversion rates were 0.9% (Gloor *et al.*, 1991); if the template was X-linked the reversion rate was 5.5% (Engels *et al.*, 1994); and when the template was extrachromosomal, the reversion rate was 0.5% (Banga and Boyd, 1992; Keeler *et al.*, 1996). There have been successful gene replacement studies done at the *forked* locus (Lankenau *et al.*, 1996), at the *white* locus

with two different P-elements (Gloor *et al.*, 1991; Lackenau *et al.*, 1996) the *singed* locus, (Gloor *et al.*, 1991) and the *oaf* locus (Merli *et al.*, 1996).

The findings suggest that gene targeting is a powerful tool in performing *in vivo* transgenic studies. The *white* model and the non transposable template that has been characterized, has been chosen to perform transcriptional analysis for a number of reasons. The location of the P-element in the *white* gene is in exon 6, reducing the likelihood that the transcriptional regulatory domains of the *white* gene will affect the promoter region under study. The non transposable vector has been tested with a number of constructs and tit could be modified to make it a reporter construct. The number of genetic crosses that are needed to identify and isolate insertions are fewer than the original P-element mediated transposition studies and those studies using transposable templates. The possibility that this could lead to a much more quantitative and reproducible mechanism for studying transcriptional regulation made this system an ideal candidate to perform the initial promoter analysis studies on *Pkc53E* in *Drosophila*.

The Role of PKCs in Growth and Development

Initially protein kinase C (PKC) was characterized by Nishizuka's group as a Ca^{2+}, diacyl glycerol (DG), and phospholipid dependent, serine and threonine kinase (Takai *et al.*, 1977). Since then, interest has flourished in understanding PKC function and regulation due to its role as a key second messenger enzyme in signal transduction cascades. Many of these isoenzymes are known cellular receptors for tumor promoting phorbol esters (12-*O* -tetradecanoylphorbol-13-acetate: TPA), thereby linking PKC with tumorigenesis.

PKC is now known to be a family of at least 12 isoenzymes characterized as having two to four conserved domains (C1-C4) and five variable isoenzyme specific domains (V1-V5). The PKC family has been divided into three categories. The classical PKCs

(cPKCs), α, βI, βII, and γ are activated by Ca^{2+}, phosphatidylserine (PS) and DG or tumor promoting phorbol esters. The novel PKCs (nPKC) δ, ε, η, θ, and μ are activated by DG or tumor promoting phorbol esters and PS. These isoenzymes lack the C2 region and are therefore unresponsive to Ca^{2+}. The atypical PKCs (aPKC) ζ, λ, and ι are not responsive to DG, tumor promoting phorbol esters, or Ca^{2+}. The signal that activates the atypical isoenzymes is still unknown.

Recent research has shown that the abnormal regulation, altered expression of one of the isoforms, or the expression of one of the isoforms in tissues where it is normally not expressed, may result in oncogenic or antioncogenic effects. Despite these findings, few studies have examined the transcriptional regulation of the mammalian PKC isoforms in different tissues.

Promoter analyses have been performed on two mammalian cPKCs to determine the regulatory elements that control the tissue-specific and developmental stage specific expression of the protein. The 5' flanking regions of rat PKCγ, mouse PKCγ, human PKCγ and human PKCβ were isolated, sequenced, and analyzed (Chen *et al.*, 1990; Niino *et al.*, 1992; Obeid *et al.*, 1992; Takanaga *et al.*, 1995; Mahajna *et al.*, 1995). Transcriptional studies with the 5' region of human PKCβ examined the regulation of a reporter construct in 5 different human cell lines and 3 different rodent cell lines (Obied *et al.*, 1992; Niino *et al.*, 1992; Mahajna *et al.*, 1995). Transcriptional start sites were first identified by S1 nuclease analysis and primer extension. The murine studies showed one transcriptional start site 287 nucleotides 5' from that determined with S1 nuclease studies performed using human erythroleukemia K562 human cells (Niino *et al.*, 1992; Obeid *et al.*, 1992). This difference was postulated to be a result of some minor promoter or a promoter that is used in a species other than humans. Mahajna *et al.* (1995) performed S1 nuclease studies in a human histiocytic lymphoma U937 cell line to determine the transcriptional start site. Their results indicated that there were multiple transcripts resulting

from different transcriptional start sites, one of which corresponded to the site reported by Obeid *et al.* (1992). The different transcriptional start sites were believed to be the result of cell-specific transcriptional machinery. Basal promoter regions from the human cell line promoter studies showed two different sizes of 5' fragments needed for transcription to occur. Obeid *et al.* (1992) reported that in K562 cells the minimal promoter fragment was between -111 to +43, whereas Mahajna *et al.*(1995), reported that the minimal promoter needed in U937 cells was between -690 to +353. Negative and positive regulatory elements that alter the expression of the reporter gene are present in the 5' region of PKCβ reporter constructs from Mahajna *et al.* (1995) and Niino *et al.*, (1992) with the negative elements being distal from the basal promoter fragment.

PKCγ transcriptional studies were performed using the promoter regions from mouse, rat and human PKCγ genes. Sequence analysis of the 5' flanking region demonstrated high similarity among the three species (Takanaga *et al.*, 1995) . Although the similarity in the 5' flanking region was high, the minimal sequence that was necessary to drive promoter-reporter constructs varied with the different promoters and cell lines used. Mouse PKCγ needed 87 bp upstream from the determined transcriptional start site for expression in rat pheochromocytoma PC12 cells. However basal level transcription from rat PKCγ in 293 cells required a 160 bp fragment upstream of the transcriptional start site. The studies on human PKCγ did not determine the minimal fragment needed for transcription. Mahajna *et al.* (1995) used a 5.5 kb fragment containing part of the first exon and the adjoining 5' region from human PKCγ in a promoter reporter construct which was expressed in promyelocyitc leukemia HL60 and U937 cells. This PKC isoform is expressed strictly in the central nervous system (CNS) in all three organisms, suggesting a high level of transcriptional regulation. The elements that are necessary for this specificity have not determined.

Putative regulatory elements were found in all studies (Table 1). Of particular interest is the frequency of Ap1 (activator protein 1), Ap2 (activator protein 2) and Sp1 (stimulatory protein 1) binding sites. TPA-inducible genes, *c-fos*, *c-myc*, collagenase, stromelysin and SV40 all contain a conserved 9 bp motif which was named a TPA Responsive Element (TREs; Angel *et al.*, 1987a; Angel *et al.*, 1987b). These elements bind Ap1 (Angel *et al.*, 1987a), a transcriptional factor that is either heterodimeric (cFos and cJun) or homodimeric (cJun and cJun) and has been found to phosphorylated directly by specific PKC enzymes (Pulverer *et al.*, 1991; Adler *et al.*, 1992) or dephosphorylated in response to PKC activation (Boyle *et al.*, 1992; Hirai *et al.*, 1994). TPA is now believed to also activate a number of transcription factors that bind to a variety of sequences, termed non-Ap1 TREs (Kim *et al.*, 1994; Tseng *et al.*, 1994; Garber *et al.*, 1994; Mar *et al.*, 1995; Johnson, 1996; Hawker *et al.*, 1996).

PKC β and γ promoter-reporter constructs were induced by TPA suggesting an autoregulatory feedback system. Obied *et al.* (1992) examined the regulation of PKCβ and observed an 8-20 fold increase in transcriptional activity when cells were treated with TPA. The basal promoter was defined from -111 to +43. This construct does not contain an Ap1 site implying that there is an Ap1 independent mechanism that has not yet been identified. The human PKCγ promoter was also potently induced by TPA and there were no Ap1 sites identified, only Ap2 sites were present in the fragment that was sequenced (Mahajna *et al.*, 1995). Ap2 is also known to be responsive to TPA suggesting that there are more that one type of transcriptional activator that is upregulated in response to treatment by TPA (Johnson *et al.*, 1996) .

The other similarity that all the PKC promoter sequences share is that they do not have the canonical TATA sequence from -25 to -30 bp. All known PKC genes have the consensus TATA sequence present roughly 500 to 900 bps upstream of the identified

Table 1. Potential transcription factor binding sites within mammalian PKC β and PKC γ sequences. The locations of the elements are given in relation to the published transcriptional start sites (summarized from data in: Chen et al., 1990; Obied *et al.*, 1992; Niino *et al.*, 1992; Mahajna *et al.*, 1995; Takanaga et al., 1995). The numbers in brackets represent the location of the sequences relative to the transcriptional start site. Underlined sequences are bases that differ from the consensus sequences.

Transcription Factor	Consensus Sequence	PKCγ Human	PKCγ Rat	PKCγ Mouse	PKCβ Human
Ap1	TGAGTCA	n.p.	TGAGTCG (-264)	n.p.	TGAGTGA (-442)
Ap2	CCCCACCC	(-38, -189)	(-37, -96)	(+55, -41)	(-320)
Sp1	CCCGCC	(-53, -82)	(-69)	(-76)	(-94,-63)
c-myc	GAAAGGG	n.p.	(-205)	n.p.	n.p.
Oct BP	ATGCAAAT	n.p.	n.p.	n.p.	(-76)
CREB	GAGACGT CAG	n.p.	GAGACGT GAG (-573)	n.p.	n.p.

AP1, AP2- activator proteins 1 and 2; Sp1- stimulatory protein 1; Creb- cAMP regulatory element-binding protein; Oct BP- octamer binding motif; n.p. not present.

transcriptional start site (Chen *et al.*, 1990; Niino *et al.*, 1992; Obeid *et al.*, 1992; Takanaga *et al.*, 1995; Mahajna *et al.*, 1995).

Studies of basal promoter elements in cell culture cannot fully elucidate the temporal and/or spatial regulatory elements that regulate expression during organismal or tissue development. *In vivo* transcriptional studies in *Drosophila* can overcome this barrier. Studying similar genes *in vivo* will allow for a better understanding of the *in vitro* results and what is really happening in the organism.

Drosophila melanogaster Pkc53E

Rosenthal *et al.* (1987) and Schaeffer *et al.* (1989) were the first to isolate and characterize 2 *Drosophila* genes that were analogous to those that encode cPKCs in mammals and one encoding a protein that is analogous to the nPKCs. The three genes isolated to date, the transcripts that they encode, and their cytogenetic locations are shown in Table 2. The PKC53E protein is 64% homologous to the classical PKCs (α,β and γ) and is believed to be Ca^{2+} and DAG dependent for activation (Rosenthal *et al* , 1987). The gene that encodes InaC protein is located 50 kbp from *Pkc53E* and it also encodes a classical PKC that is DAG and Ca^{2+} dependent. This protein is expresssed specifically in the photreceptor cells and is important in light adaptation (Hardle *et al.*, 1993). Little is known about the protein encoded by *Pkc98E*. It is 61% homologous to the novel PKC. PKCδ,and it is expressed specifically in the the cell bodies in the brain but the method of activation and function is still unknown (Schaeffer *et al.*, 1989; Kozoza *et al.*, 1991). No promoter studies have been reported to date for any of these genes.

Rosenthal *et al.* (1987) determined the genetic structure of the *Pkc53E* gene (located at 53E on chromosome number 2) by probing a *Drosophila* EcoRI partial digest genomic library with a bovine PKCα (bPKCα) probe. Two EcoRI fragments from two of the 4 positive genomic clones were further used to probe three *Drosophila* cDNA libraries

Table 2. Transcript sizes (kb) from *Drosophila* PKC encoding genes. O indicates that there has been no transcripts found at that developmental stage. (summarized from the data presented in Rosenthal *et al.*, 1987; Natesan, 1991; Schaeffer *et al.*, 1989).

Gene	*Pkc53E*	*inaC*	*Pkc98E*
Cytogenetic Location	53E	53E	98E
Embryo	0	0	4.3, 4.5, 5.5
Larva	2.4, 3.0, 3.4	0	5.5
Pupa	2.4, 4.0, 4.3	2.5	5.5
Adult	2.4, 4.0, 4.3	2.5	5.5

representing RNAs from 0-3 hour embryo, 5.5-7.5 day early pupal, and adult *Drosophila*. Interestingly, the reported transcriptional +1 site from the cDNA sequence corresponds exactly with an EcoRI cut site, the same EcoRI site reported for one of their genomic fragments they used as a probe.

Natesan (1991) and Rosenthal *et al.* (1987) determined that *Pkc53E* encodes at least five transcripts (2.4,3.0, 3.4, 4.0, and 4.3 kb). Northern hybridization analyses have shown that *Pkc53E* mRNAs are not normally expressed until the first instar larval stage (Rosenthal *et al.*, 1987; Natesan, 1991). The 3.4, 3.0 and 2.4 transcripts are detected in second and third instar larvae, while the 4.3 and a 4.0 kb transcripts appear in the late third instar larval stage. The adult head contains three known transcripts: 2.4; 4.0 and 4.3 kb which are localized primarily in the brain and eyes (Hughes, 1992). A 2.4 kb transcript is present predominantly in the adult gonads (Hughes, 1992).

The 3.4 and 2.4 kb transcripts of *Pkc53E* are transcriptionally induced by TPA in 0-3 hour old embryos and 1st instar larvae (Natesan, 1991). This induction and the finding that PKC is activated by TPA (Nishizuka, 1988), suggests that TPA can activate *Pkc53E* at both the enzymatic and transcriptional levels. *Pkc53E* mRNA are normally not found at these developmental stages, leading to the speculation that maternal Pkc53E could be the target for TPA or there could be another intracellular target for TPA that activates TRE binding proteins. *Pkc98E* is transcribed in both the embryonic and adult stage, therefore *Pkc98E* may be the TPA receptor in embryonic cells (Schaeffer et al., 1989).

These findings, coupled with the transcriptional studies performed on the mammalian cPKCs, suggest that PKC is transcriptionally tightly regulated. An understanding of the transcriptional machinery controlling these genes may allow insight into development.

<u>*Drosophila* Basal Promoter Structures</u>

The promoter structures of several *Drosophila* genes have been characterized by *in vivo* and *in vitro* studies and by sequence analysis (examples: Perkins *et al.*, 1988; Arkhipova, 1995; Lankenau *et al., 1996*). The *Drosophila* promoter database contains 230 entries of 5' sequence (Arkhipova, 1995). Statistical analysis comparing the composition of the promoter regions relative to the known transcriptional start sites of the genes reveals a basic tripartite structure in the region surrounding the transcriptional +1 site (Arkhipova, 1995). The TATA box located -25 to -35 bp upstream of the start of transcription was found in half of the *Drosophila* promoter sequences. The TATA sequence is a binding site for the TATA-box-binding protein (TBP) component of the multisubunit TFIID complex. The cap-site or initiator (Inr) sequence (^{T}/A/$_{G}$TCA^{G}/$_{T}$T^{T}/$_{C}$G) is located +/- 5 bp from the RNA start site and this element can initiate transcription through the TATA-box-binding proteins (TAFs) in the absence of a TATA box (Martinez *et al.*, 1994; Martinez *et al.*, 1995). The Inr sequence element has been identified in one third of the promoter sequences. A downstream promoter element (DPE) located at +20/+30, is present in three quarters of the TATA-less promoter sequences and is necessary for transcription in conjunction with the Inr (Thummel, 1989; Arkhipova, 1995; Burke and Kadonga, 1996). Sequences that were found in high frequency in mammalian promoters, such as the sequence for transcription factor Sp1 (GGGCGG) were not as evident in *Drosophila* promoters.

The 5' proximal flanking region of genes that are developmentally regulated and genes which are expressed in the central nervous system (CNS) and the peripheral nervous system (PNS) in *Drosophila* show a TATA-less structure and do not have a high GC rich concentration [e.g. *Antennapedia* (*Antp*) (Perkins *et al.*, 1988); *engrailed* (*en*) (Soeller *et al.*, 1988); *E74* (Thummel, 1989); *elav* (Yao and White, 1994)]. These findings contradict the canonical promoter sequence that have thus far been studied. It is believed that

transcriptional studies will reveal that this is more and more the norm because the original studies all focused on the structure of house keeping genes.

The study of transcriptional regulation has given important insights into gene expression. The molecular tools used in *Drosophila* make it an ideal candidate to study the effects of transcriptional regulation over time. Introduction of P-element based constructs into the genome is a very powerful tool. The major drawback with P-element based reporter vectors is the randomness of integration. With the characterization of P-element excision and the subsequent repair, a gene targeting mechanism is now possible (Engels *et al.*, 1990; Gloor *et al.*, 1991). The aim of this project is to identify the basal promoter region of the gene, *Pkc53E*, using a newly constructed promoter-reporter vector for gene targeting to the *white* locus. This technique should allow for consistent reproducible results for identifying the basal promoter of a gene and for further studies identifying specific regulatory elements used for transcription of the gene of interest. The 5' region of the tightly regulated *Drosophila* gene, *Pkc53E* was first identified by restriction and sequence analysis. The gene targeting technique was then adapted by constructing a new nontransposable reporter vector and then creating promoter reporter vectors that were used for the identification of the basal promoter region of *Pkc53E*.

Materials and Methods

Chemical Suppliers

Chemicals and supplies were purchased from BDH chemicals, Canlab, Fisher Scientific, GIBCO/Bethesda Research Laboratories Canadian Life Technologies Inc. (BRL), or Sigma. Radioactive nucleotides were purchased from Amersham or ICN Biomedicals. The nucleic acid metabolizing enzymes were obtained from BRL or Pharmacia. The Glassmax DNA spin cartridge system and Deletion Factory II System were purchased from BRL. The DIG Labeling and Detection Kit was purchased from Boehringer Mannheim. Table 3 lists the solutions and buffers that were used for the experiments described.

Drosophila stocks

Drosophila stocks used were maintained on standard agar based cornmeal/yeast/ sucrose/dextrose media (Lewis, 1960). The stocks were kept at room temperature (18-22°C) or at 25°C. Stock descriptions are from Lindsley and Zimm (1992) and Flybase (1996). Fly strains were obtained from the Mid-American Drosophila Stock Center or were generous gift from Dr. Greg Gloor.

y w f : Females of this stock contain a compound X chromosome with three recessive visible mutations. *y* (*yellow*, 1-0.0) causes a yellow body with brown bristles and yellow tip. *w* (*white*, 1-1.5) causes a white eye colour and the allele in this stock is w^1, a spontaneous mutation causing the white phenotype. *f* (*forked*, 1-56.7) causes the bristles to be shorter than wild type and bent at the tip. The autosomes are Canton-S. The males are phenotypically wild type.

Table 3. Solutions and buffers used for experiments. Percentages are weight per volume for solids and volume per volume for liquids. pH in brackets indicates pH of solution being added, pH at the bottom of the reagents means the solution has to be altered.

Name	Composition
cell lysis buffer	0.2 M NaOH 1% SDS
cell neutralization buffer	5M KAc (pH4.8) 29.5% glacial acetic acid KOH to pH 4.8
cell resuspension buffer	50 mM Tris (pH 7.5) 10 mM EDTA 1% RNaseA
DNA loading dye	0.25% bromophenol blue 0.25% xylene cyanol FF 30% glycerol in H_2O
egg collection media	10% glucose 10% dried yeast 2% agar
GET	50 mM glucose 10 mM EDTA (pH 8.0) 25 mM Tris-HCl (pH 8.0)
injection buffer	0.1 mM phosphate buffer (pH6.8) 5 mM KCl
Lambda medium	10% tryptone 42mM NaCl 0.6 mM maltose 10 mM $MgSO_4$-$7H_2O$
Lambda top agar	10% tryptone 42mM NaCl 7% agar
LB medium	10% tryptone 5% yeast extract 0.17M NaCl pH 7.0
LB top agar	10% tryptone 5% yeast extract 0.17M NaCl 7% agar

Name	Composition
NZC	10% caesin enzymatic hydrolysate 8.6 mM NaCl 10 mM $MgCl\text{-}6H_2O$ autoclaved 0.1% casamino acids
SM	10 mM NaCl 8 mM $MgSO_4\text{-}7H_2O$ 0.05 mM Tris-HCl (pH 7.5) 0.01% gelatin
TAE	0.04 M Tris-acetate 0.001 M EDTA 5.71% glacial acetic acid pH 8.5
TE	10 mM Tris-HCl 1 mM EDTA pH 8.0
20X SSC	3 M NaCl 0.3 M sodium citrate pH 7.0

y w^{hd80k17} f : This strain has three recessive visible mutations on the X chromosome and the autosomes are from Canton-S. *w^{hd80k17}* (*w^{hd}*; white, 1-1.5) causes a bleach white eye colour, this allele is the result of a 629 bp P-element insertion into the coding portion of exon 6 (O'Hare and Rubin, 1983).

y w; Ki rv^{506} P[ry^{+};Δ 2-3](99B): This stock has two visible recessive mutations on the X chromosome and a visible dominant marker indicating the presence of the transposase source on the third chromosome. *Ki* (*Kinked*; 3-47.6) causes all the bristles to become shortened and twisted. P[ry^{+},Δ2-3](99B) is a plasmid that has become stabilized on the third chromosome, this construct expresses the transposase enzyme in both the germ line and soma (Robertson *et al.*, 1988)

Plasmids and Growth Conditions

Table 3 shows the cloning vectors, competent cell strains, growth media, and antibiotics used for the subcloning techniques. Figure 3 is a schematic of the nontransposable vector, pBS[walL] (given to us by Dr. Greg Gloor), used for targeted transposition (Keeler et al., *1996*). Figure 4 is a schematic representation of the transposable reporter vector, pCaSpeR-AUG-βgal (Thummel *et al.*, 1988).

Lambda Phage Maintenance

Plating of lambda phage. To maintain stocks of lambda bacteriophage 5 ml of lambda medium with 0.2% maltose and 10mM $MgSO_4$ were inoculated with 1 colony of an appropriate *E. coli* host (Y1090) and incubated at 37°C with shaking until growth to saturation. Lambda top agar was prepared and placed in a 50°C water bath and allowed to cool to the water bath temperature. From the saturated culture, 0.3 ml was aliquoted to five

Table 4. Plasmid vectors, competent cells, growth media and selectable markers used. All subcloning procedures (described in text) used these materials. Antibiotic concentration varied according to protocol.

Vectors	Competent Cells **	Growth Media	Selectable Markers*
pdelta1 (BRL)	DH10B	LB medium	Ampicillin
	DF1	LB medium	Tetracycline Kanamycin
pBluescript SK+ (Stratagene)	DH5α	LB medium	Ampilcillin
	DH10B	LB medium	Ampicillin
pGem 7zf+ (promega)	DH10B	LB medium	Ampicillin
	DH5α	LB medium	Ampicillin
pCaSpeR-AUG-βgal	DH10B	LB medium	Ampicillin

*all selectable markers were purchased from Sigma.
** all competent cells were purchased from BRL.

Figure 3. Schematic representation of a non-transposable template. This Bluescript based vector contains an altered 4.4 kb *white* gene homologous sequence (refer to Figure 2) with a 30 bp unique MCS used to incorporate heterologous DNA into the *white* sequence. The *white* gene transcriptional start site is shown. This vector and subsequent modifications were introduced into the genome of *Drosophila* through microinjections followed by gap repair (Keeler *et al.*, 1996). ori- origin of replication, amp^{+}- gene encoding for ampicillin resistance, T7 and T3 promoters- T7 and T3 RNA polymerase binding sites. Not drawn to scale.

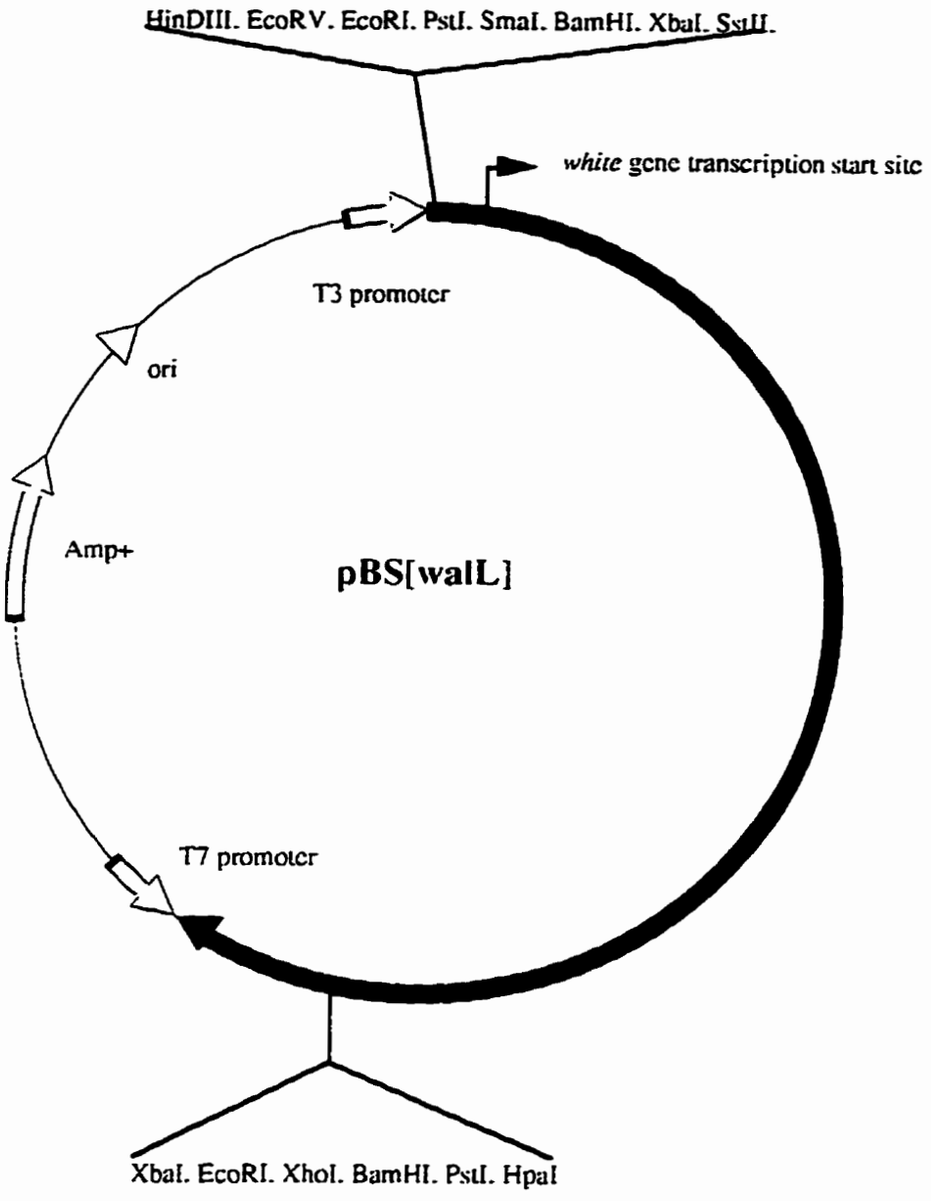
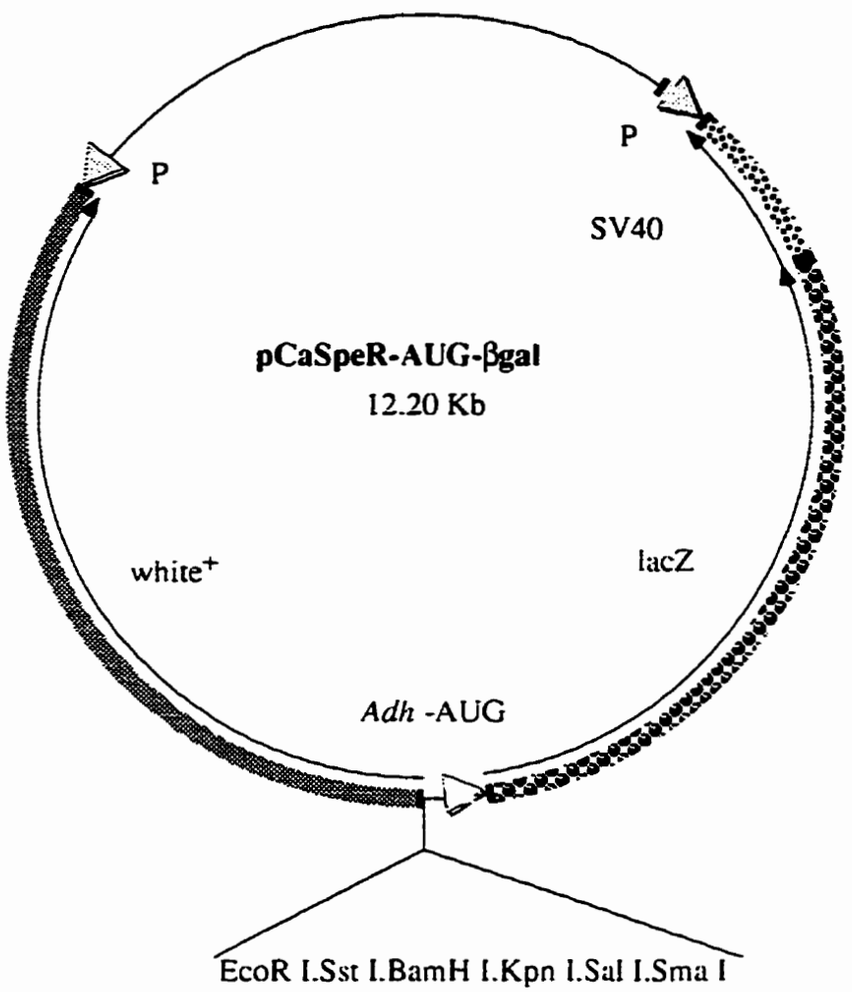

HinDIII. EcoRV. EcoRI. PstI. SmaI. BamHI. XbaI. SstII.
white gene transcription start site
T3 promoter
ori
Amp+
pBS[walL]
T7 promoter
XbaI. EcoRI. XhoI. BamHI. PstI. HpaI

Figure 4. The structure of the P-element vector, pCaSpeR-AUG-βgal. This construct contains its own ATG upstream of the reporter gene *lacZ*. The unique cloning sites are EcoRI, BamHI and KpnI. P: 31 base pair inverted repeats from the P-elements: *Adh*-AUG: known *Drosophila* start codon for the translationa of the *lacZ* gene (modified for Figure 1, Thummel et al., 1988).


P
P
SV40
pCaSpeR-AUG-βgal
12.20 Kb
white+
lacZ
Adh -AUG
EcoR I.Sst I.BamH I.Kpn I.Sal I.Sma I

8 x 80 mm tubes. Hundred fold serial dilutions were made of the phage lysate in SM and 0.1 ml of the first dilution was added to one tube of *E. coli*, 0.1 ml of the next dilution was added to the next tube until all 5 tubes were inoculated. The tubes were incubated at room temperature for 20 minutes and then moved to a 37°C water bath for 10 minutes. While the tubes were in the water bath, 5 LB plates were labeled corresponding to the dilution tubes. The tubes were removed from the water bath, and 2.5 ml of the top agar was added to each of the tubes, swirled and plated separately. The plates were placed at 37°C and left for 12 hours. The titer (pfu: plaque forming units) of the phage lysate was calculated by counting the number of plaques and dividing by the dilution that was used. In order to save a plaque a single colony was picked with a sterile capillary tube and blown into a centrifuge tube containing 1 ml of SM (modified from Sambrook *et al.*, 1989)

Lambda plate lysis. 1 ml of LB containing 0.2% maltose and 10 mM $MgSO_4$ was inoculated with an appropriate *E. coli* host (Y1090) and grown overnight at 37°C with shaking. When saturated, 10^5 pfu of bacteriophage (usually 1/20 of a resuspended plaque) was mixed with 0.1 ml of bacteria from a saturated culture and incubated at 37°C for 20 minutes. Top LB agar was prepared and kept molten in a 50°C water bath. 0.3 ml was added to the bacteria/phage mixture, and poured onto an 82-mm plate containing 30-35 ml of hardened bottom LB agar. Higher yields were obtained from bottom agar containing 0.3% glucose, 0.075 mM $CaCl_2$, 0.004 mM $FeCl_2$, and 2 mM $MgSO_4$. The plates were incubated for 6-8 hours at 37°C. At the time of harvesting the plaques were touching one another, and the only visible bacterial growth was a gauzy webbing that marks the junction between the adjacent plaques. Five ml of SM was added to each plate and the plates were stored at 4°C for 4 hours with gentle agitation. Using a sterile Pasteur pipette, the SM was harvested and placed in a polypropylene tube (13x100 mm). An extra 1 ml of SM was added and the plates were stored for 15 minutes at room temperature in a tilted position.

The SM was combined with the first harvest and the plates were then discarded (modified from Sambrook *et al.*, 1989).

Nucleic Acid Isolation

Bacterial cultures. Liquid bacterial cultures were grown in a nutrient rich medium (LB medium, NZC, or Lambda medium) at 37°C with shaking. Where required, ampicillin was added to a final concentration of 10 µg/ml. For single colony selection, bacteria were plated on LB plates and placed at 37°C. To maintain a line of bacteria, they were grown in LB medium to saturation and placed at -70°C in 15% glycerol.

Lambda DNA isolation by liquid lysate. In order to purify Lambda DNA, the Qiagen system (Qiagen Inc.) was utilized and the procedure was carried out as per instructions provided by Qiagen. For a liquid lysate, 1 ml of LB containing 0.2% maltose and 10 mM $MgSO_4$ was inoculated with a single colony of an appropriate *E. coli* (Y1090) host and grown overnight at 37°C with shaking. In a centrifuge tube, 0.1 ml of eluted phage, 0.1 ml of saturated culture and 0.1 ml of 10 mM $MgCl_2$/10 mM $CaCl_2$ solution were mixed and placed at 37°C for 15 minutes. The solution was transferred to 100 ml of NZC solution and shaken vigorously at 37°C for 14-18 hours until lysis occurred. A few drops of chloroform were added to lyse any remaining cells. The solution was transferred to 50 ml Nalgene tubes and centrifuged for 10 minutes at 11000 x g, at 4°C. The supernatant was transferred to a screw-cap tube, a few drops of chloroform were added and the lysate was stored at 4°C. Purification of lambda DNA from 100 ml of a liquid lysate yielded up to 60 µg of DNA.

Alkaline lysis plasmid miniprep. In order to screen bacterial colonies, tubes of 5 ml of LB broth were inoculated with single colonies and grown to saturation overnight. From these starter cultures, 1 ml was transferred to a microcentrifuge tube and pelleted at 12000 x g in a bench top microcentrifuge for 30 seconds. The supernatant was removed and the pellet was resuspended in 200 µl of cell resuspension buffer. the cells were then lysed by the addition of 200 µl cell lysis buffer and 100 µl cell neutralization buffer. The mixture was placed on ice for 5 minutes and centrifuged at 12000 x g for 10 minutes to pellet the cellular debris and chromosomal DNA. The supernatant was transferred to a new microfuge tube and mixed with 2 volumes of 95% ethanol in order to precipitate the plasmid DNA. The solution was then centrifuged at 12000 x g for 5 minutes, the supernatant was poured off, and the pellet was washed with 200 µl of 70% ethanol. The pellet was air dried and resuspended in 30 µl of TE.

Alkaline lysis plasmid miniprep for sequencing. To obtain supercoiled plasmid DNA for sequencing the following protocol was used. Three ml of LB broth with the appropriate antibiotic (10 mg/ml) was inoculated with a single bacterial colony and grown to saturation. From these cultures 1.5 ml were aliquoted into microfuge tubes and centrifuged at 12000 x g for 30 seconds to pellet the cells. The supernatant was resuspended in 200 µl cell resuspension buffer and the bacterial cells were lysed with 200 µl cell lysis buffer and 300 µl cell neutralization buffer. The mixture was centrifuged at 12000 x g for 5 minutes and the supernatant was transferred to a new microfuge tube and centrifuged again for 5 minutes at 12000 x g to remove all cellular debris. The supernatant was transferred to a microfuge tube and 70 µl of RNAse A (10 mg/ml: Sigma) was added and the solution was stored at 37°C for 2 hours. Two, 1:1 volume phenol:chloroform extractions and two, 1:1 volume chloroform extractions were then performed and the DNA was then precipitated with isopropanol (0.7%) at room temperature for 10 minutes. The

solution was then centrifuged at 12000 x g for 10 minutes, the supernatant poured off, the pellet washed with 200 µl of 70% ethanol and air dried. The pellet was resuspended in 32 µl of sterile double distilled water, 8 µl 4 M NaCl, and 40 µl 13% polyethylene glycol-8000 (w/v) and the tube was placed on ice for 20 minutes. The solution was centrifuged at 4°C at 12000 x g for 15 minutes. The supernatant was removed, the pellet was washed with 200 µl of 70% ethanol and air dried. The pellet was resuspended in 20 µl of sterile double distilled water (Ausubel *et al.*, 1995).

Alkaline lysis plasmid maxipreps. In order to purify large quantities of plasmid DNA from bacterial cultures of 100 ml and 500 ml, alkaline lysis was performed (Sambrook *et al.*, 1989).

DNA purification by CsCl banding. Plasmid DNA that was used for microinjection in *Drosophila* embryos was purified by CsCl banding. Bacteria containing the plasmid of interest were grown in 250 ml of LB medium overnight at 37°C with shaking. The saturated sample was then transferred to a 250 ml centrifuge tube and the DNA pelleted by centrifugation at 6000 x g. The supernatant was removed and the cellular pellet was resuspended in 6 ml of GET. The slurry was transferred to 50 ml centrifuge tube. 1ml of 20 mg/ml of lysozyme (Sigma) was added and the solution incubated at room temperature for 10 minutes. The bacteria were then lysed by adding 14 ml of cell lysis buffer and placed on ice for 5 minutes. Seven ml of cell neutralization buffer was added and the solution placed back on ice for 15 minutes. To remove the insolubles, the solution was centrifuged at 11000 x g for ten minutes. The supernatant was placed into a sterile 50 ml centrifuge tube and extracted once with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1). The plasmid DNA in the aqueous phase was precipitated with 0.6 volume isopropyl alcohol for 10 minutes at room temperature, and centrifuged at 11000 x g

for 15 minutes. The supernatant was removed, the DNA pellet air dried for 10 minutes, and then resuspended in 1 ml of TE.

To the nucleic acid-TE mixture, 1.1 g of CsCl was added per ml of solution and inverted several times in order to dissolve the CsCl. For every 1 ml of CsCl-nucleic acid-TE mixture, 800 µl of ethidium bromide (EtBr) solution (10 mg/ml) and 10 µl of a 1:100 dilution of Triton X-100 per ml of tube volume was added. The solution was then transferred to a 1.2 ml Quick Seal® tube specific for a TLN-120 rotor (Beckman). The tube was then balanced against another tube containing CsCl of the same density. The tubes were heat sealed using a Beckman heat sealing device and placed in the TLN-120 rotor. Spacers and plugs were put over the tubes and the plugs were torqued to 13 Newton meters (N-m). The rotor was placed in an Optima™ TLX ultracentrifuge (Beckman) and the samples were centrifuged at 120000 x g for 1 hour 15 minutes at 20°C with the deceleration setting at 5.

The tube was placed in a tube rack and the plasmid band was identified using long wave UV light. To relieve pressure, the top of the tube was punctured with an 18-gauge needle. An 18-gauge needle on a 1-ml syringe is inserted bevel up below the band and the plasmid was withdrawn slowly. The plasmid containing solution was placed in a 15 ml polypropylene tube and the EtBr was removed by adding alcohol (η–butanol saturated with CsCl and water). Upon mixing, the top layer turned red, this was removed and the process repeated until the top layer was colourless. The bottom layer, which contains the plasmid DNA, was collected by dialysis (performed according to Ausubel *et al.*, 1995).

***Drosophila* genomic DNA.** DNAzol was used to isolate *Drosophila* genomic DNA according to manufacturer's instructions (GIBCO-BRL) with the following modifications. Between 10 and 20 *Drosophila* were placed in a 1.5 ml tube with 250 µl of DNAzol and homogenized. The homogenate was centrifuged at 12000 x g for 5 minutes. The

supernatant was removed and placed into a sterile 1.5 ml microcentrifuge tube, the DNA was precipitated by adding 125 µl of 100% ethanol, and left at room temperature for 3 minutes. The solution was then centrifuged at 12000 x g for 5 minutes. The ethanol was removed, the pellet was washed twice with 500 µl of 95% ethanol, and centrifuged at 12000 x g for 5 minutes. The ethanol was removed and the DNA pellet was resuspended in 100 µl of 8mM NaOH or TE and placed at 37°C for 12 hours. The solution was centrifuged at 12000 x g for 5 minutes to remove the insolubles and the supernatant was transferred to a sterile tube.

Manipulation of DNA

Restriction endonuclease digestion. Digestion of DNA with restriction enzymes were carried out according to manufacturers instructions (BRL, Pharmacia).

Agarose gel electrophoresis of DNA. Agarose gel electrophoresis was used to quantitate, size and purify fragments of DNA. The concentration of agarose in the gel varied from 0.7% to 1.2% depending upon the size of the DNA fragment of interest. Gel electrophoresis was performed as described by Ausubel et al. (1995)

Dephosphorylation. Dephosphorylation of plasmids that were cut with a single enzyme were performed in order to remove the 5' phosphate group to prevent recircularization or concatamerization of the plasmids. Dephosphorylation was performed according to GIBCO-BRL specifications for Calf Intestinal Alkaline Phosphatase (CIAP) with the following modifications. Approximately 5µg of plasmid (in a volume of 10 µl distilled water) was mixed with 5µl 10 x CIAP buffer, 10 µl CIAP (600units/ ml) and 25 µl of water and placed at 37°C for 30 minutes. To inactivate the CIAP 2 µl of 0.5 M EDTA was added and the solution was placed at 70°C for 10 minutes. A 1:1

phenol:chloroform:isoamyl alcohol (25:24:1) extraction and a 1:1 phenol:chloroform extraction where then performed. The top aqueous solution that contains the CIAP treated plasmid was precipitated with 1/10 volume 3M NaAc pH 4.8 and 2 volumes of 95% ethanol, and placed at -20°C for 20 minutes. The solution was then centrifuged at 12000 x g for 10 minutes the supernatant removed and washed with 200 µl of 70% ethanol. The pellet was allowed to air dry and was resuspended in 10 µl TE.

Ligation reactions. Generally, a 5:1 insert to vector ratio was used for each ligation reaction. A 10 µl ligation reaction volume was achieved by mixing vector and insert with sterile double distilled water, 2 µl of 5x T4 DNA ligase buffer (0.25 M Tris-Cl pH 7.6, 50 mM MgCl2, 5mM ATP, 5 mM DTT, 25% (w/v) polyethlyene glycol-8000; BRL) and 2 µl T4 DNA ligase (2 unit/ml). The ligation reactions were left overnight at room temperature and competent cells were then transformed as described below.

Calcium chloride competent cells. Preparation of competent DH10B and DH5α cells were done according to Sambrook *et al.* (1989)with the following modifications. The growth of the cells were monitored every half an hour and the cells were harvested when the OD_{600} reached 0.375. To the cells that were resuspended in 2 ml of 0.1M $CaCl_2$, 300 µl of glycerol were added and mixed thoroughly and then the cells were aliquoted (200µl) to microcentrifuge tubes and placed at -70°C.

Transformation of bacterial cells. Transformation of $CaCl_2$ competent cells were carried out according to Sambrook *et al.* (1989).

Isolation of DNA fragments from agarose gels. Two protocols were used to isolate DNA fragments from agarose gels . The DNA of interest was first digested with

restriction enzyme(s) and then electrophoresed through an agarose gel. The fragment was then cut out of the gel using a razor blade. If further modifications were required for the fragment (i.e. dephosphorylation). Gene Clean (BioCAN) was used according to the manufacturer's instructions. If the DNA fragment did not have to be modified and was to be subcloned then the fragment was purified as follows. The gel fragment was transferred to a 500 µl microcentrifuge tube that was punctured in the bottom with an 18 gauge needle and contained 2-3mm of siliconized glass wool This tube was then placed in a 1.5 ml microcentrifuge tube and centrifuged for ten minutes at 6000 x g. The eluate which contains the DNA fragment was collected in the 1.5ml microcentrifuge tube. Recovery depended upon the size of the fragment. Yields were as high as 90% for fragments between 0.1 Kb and 3 Kb (Boyle and Lew, 1995).

Deletions. Unidirectional nested deletions were performed according to manufacturer's specifications for the Deletion Factory System-2 (BRL). The cloning, minipreps, and maxipreps are described above.

Dideoxy sequencing

Dideoxy sequencing reactions were performed according to manufacturer's specifications for the T7 Sequencing™ Kit (Pharmacia). Denaturing gel electrophoresis was performed according to Ausubel *et al.* (1995) with the following modifications. The acrylamide concentration for all the sequencing gels was 6% and the fragments sequenced ranged is size from 200 bp to 700 bp. For the 60 cm gel sandwich, 80 ml of the denaturing acrylamide gel solution was made instead of the 60 ml that was recommended.

Southern analysis

Southern transfer. Genomic or plasmid DNA was purified, digested with the appropriate restriction enzyme, and electrophoresed on an agarose gel. The gel was then photographed, depurinated, denatured, and neutralized before transferring the DNA to a Hybond N+ nylon membrane as recommended by the manufacturer of the membrane (Amersham). Transfer of DNA took up to 24 hours depending upon the size of the DNA fragments. Once transfer was complete, the membrane was washed in 2 x SSC for 2 minutes, dried DNA side up, and the DNA was fixed onto the membrane by exposing it to short wave UV light for 5 minutes. The membrane was stored between clean paper towels if not used right away.

Probe synthesis, hybridization and detection. Probe synthesis, probe hybridization to Southern blots, and detection of the probe, were performed using the non-radioactive DIG (digoxigenin) DNA Labeling and Detection Kit from Boehringer Mannhiem. Probe size ranged from 1.2 Kb to 2.9 Kb with concentrations of 5-20 ng/ml. Colourimetric detection was used to detect the presence or absence of hybridization of the probe to DNA fragments on the membrane.

Transformation of *Drosophila*

Injection conditions. Injections were performed at 18-20°C because the reduced temperature (compared to 25°C) slows down development of the embryos allowing more time to do the injections and the embryos are more viscous reducing cytoplasmic leakage.

Preparation of DNA. The DNA that was injected was twice CsCl banded as described above. After dialysis the DNA was ethanol precipitated with 1/10 volume 3 M NaAc (pH 4.8) and 0.7 volume isopropanol and placed at -20°C for 12 hours. The DNA was washed

three times with 70% ethanol - 0.2 M NaCl and then dissolved in injection buffer. The DNA was then quantitated on a Beckman Du® 640 Spectrophotometer.

Preparation of the injection needles. Injection needles were made from siliconized capillary tubes (Kimax-51®) that had an inner diameter of 0.75mm and an outer diameter of 1.0 mm. The needles were pulled on an Ultrafine Micropipette Puller (Frederick Haer and Co,) with a velocity index of 100, a pull index of 300, a heat index of 370 and a pressure (pounds per square inch: psi) of 50. The very tip of the needle was broken using tungsten forceps in order to make a sharper edge. Injections were done with a PLI-100 picoinjector (Medical Systems Corp). To backfill the needle, the needle was placed on an angle of greater than 45° with the tip in the injection solution. The picoinjector then applied a vacuum to draw the solution into the taper of the tip.

Collection of embryos and dechorionation. Injections were performed before the pole cells in the developing embryo were formed from the early syncytial blastoderm. Pole cell formation normally occurs in the first two hours after being laid. In order to synchronize egg development, 100-300 male and female flies were placed in an egg collection chamber (a medium sized tupperware container taped shut with a circular hole cut out of the lid the circumference of which fit a petri dish) with a petri plate containing egg collection media with yeast paste spread on the surface. Embryos were collected every hour and the first two collections were discarded because the embryos were in different developmental stages (Ashburner, 1989).

Embryos were manually dechorionated on double sided sticky tape and transferred to a 22 x 40 mm coverslip, which had a thin piece of double sided tape attached. Embryos were aligned with the posterior end protruding over the edge of the tape. Anterior ends were identified by the presence of the micropyle. The embryos were then dessicated in a glass

petri dish containing Dri-rite for 3-5 minutes. Halocarbon oil-700 was used to cover the embryos immediately after dessication (Ashburner, 1989).

Injections. Injection were performed using a PLI-100 picoinjector (Medical Systems Corp), a micromanipulator (Narishige Co.) and a Nikon microscope. The coverslip with the dessicated embryos under halocarbon oil, was placed on a slide and mounted on the microscope stage. The embryos were brought into rough focus and the needle was then lowered down into the oil and brought into the same plane of focus. The embryos were then moved towards the needle so that the needle would inject the posterior end of the embryos. Once the needle had pierced the embryo, it was manipulated in the embryo so that the tip was as close to the posterior end as possible. Injection time was 0.2 ms, injecting roughly 1/20 to 1/10 the volume of the embryo. Embryos that were past the first two stages of development were destroyed.

After the embryos were injected, the coverslip was placed in a petri dish that was kept humid by a piece of damp Whatman paper taped to the inside of the lid. The embryos were checked regularly to make sure they were not desiccating. If the oil had run off the embryos, more was added. The larvae were collected after the first day and placed in food vials where they were left to develop for 8-9 days at 25°C.

Results

Subcloning and Sequencing

In order to identify the basal promoter used to regulate the transcription of *Pkc53E*, the 5' region first had to be isolated, cloned, subcloned, and sequenced. The genomic clone, λ8 (a gift from Drs. A. Rosenthal and D.V. Goeddel; Genentech Inc, U.S.A.), contains a 17 kb genomic DNA fragment from the *Drosophila melanogaster* 53E locus packaged in charon 4A (Rosenthal *et al.*, 1987). This clone spans the upstream regulatory regions, the published transcriptional +1 site, and first two exons (non-coding exon 1 and coding exon 2) of *Pkc53E*. Figure 5 is a schematic diagram representing the 17 kb fragment and the fragments that were further subcloned into various vectors for identification and sequencing purposes. pPKC5'-1.2E is a 1245 bp EcoRI fragment subcloned into the EcoRI site of pGem 7Zf+ vector. This fragment contains the sequence immediately 5' of the published +1 site. pPKC5'-1.4H is a 1401 bp HinDIII fragment subcloned into the HinDIII site of pBluescript SK+. This fragment spans the published +1 site. p5'-B is a roughly 9 kb BamHI fragment subcloned into the BamHI site of the pDelta1 vector (Gibco BRL).

Sequence strategies are shown in Figure 6. For pPKC5'-1.2E (Figure 6A) three subclones were made using the unique cut sites, HinDIII and XhoI, which were initially confirmed by sequencing from both the 3' and 5' ends of pPKC5'-1.2E. The three new fragments were then sequenced and double stranded sequence was obtained (Figure 7). In order to check if this sequence was contiguous through the published transcriptional +1 site, a 1.4 kb fragment containing part of the 1st non-coding exon extending 5' was subcloned into pBluescript (Figure 6B). This pPKC5'-1.4H plasmid was further modified by digesting the construct with XhoI, removing the 5' most sequence, creating a 866 bp fragment that spans the published +1 site. The sequence underlined is double stranded sequence that was generated from this fragment, which shows that the sequence from pPKC5'-1.2E contains the region that is immediately 5' to the +1 site (Figure 7).

Figure 5. Schematic representation of λ8 and derivatives: pSB29: pPKC5'-1.2E; pPKC5'-1.4H; and p5'-B. λ8 is a 17 kb genomic fragment containing the putative transcriptional regulatory regions and first two exons of *Pkc53E* packaged in charon 4A (Rosenthal *et al.*, 1987). Fragments of this genomic clone were subcloned into either pGem, pBluescript or pDelta1. B-BamHI, C-ClaI, E-EcoRI, H-HinDIII. +1- published transcriptional start site.

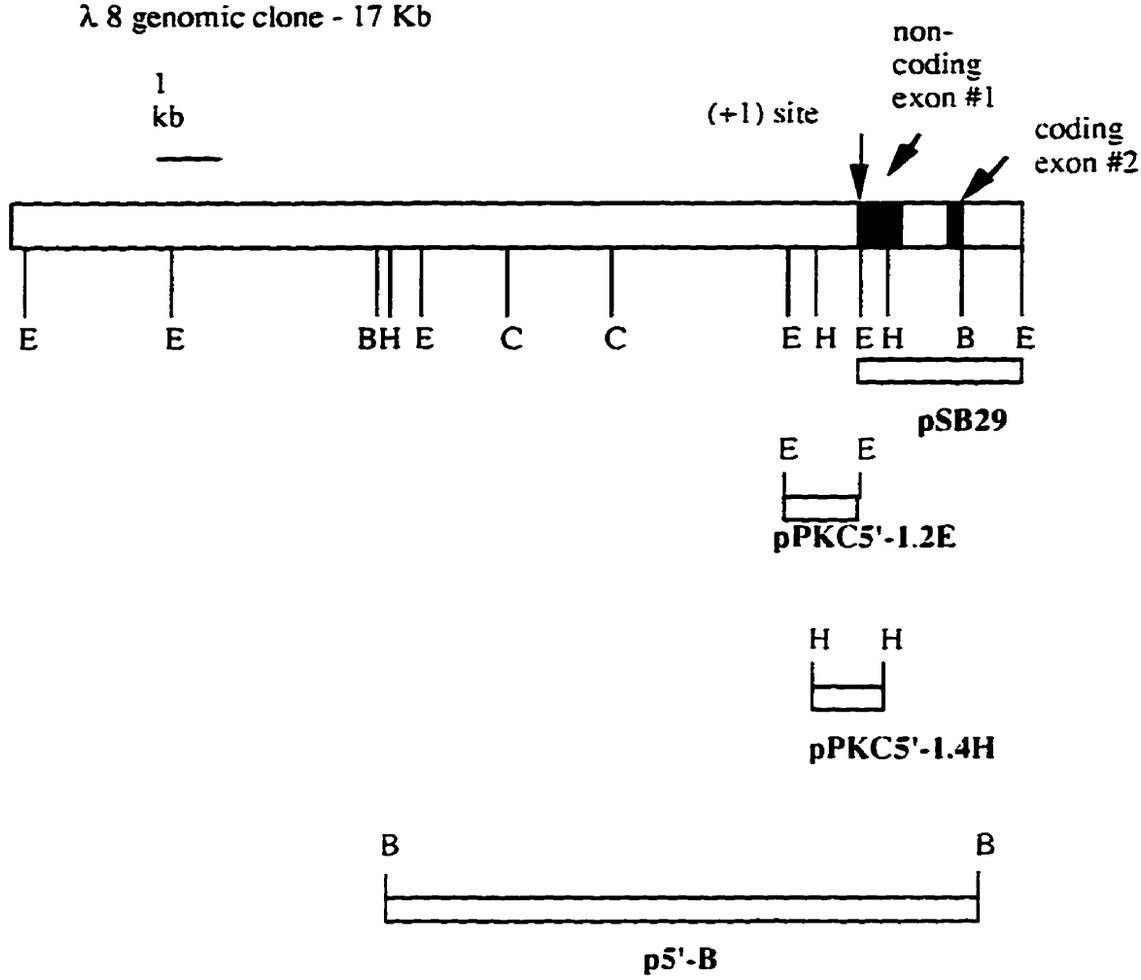
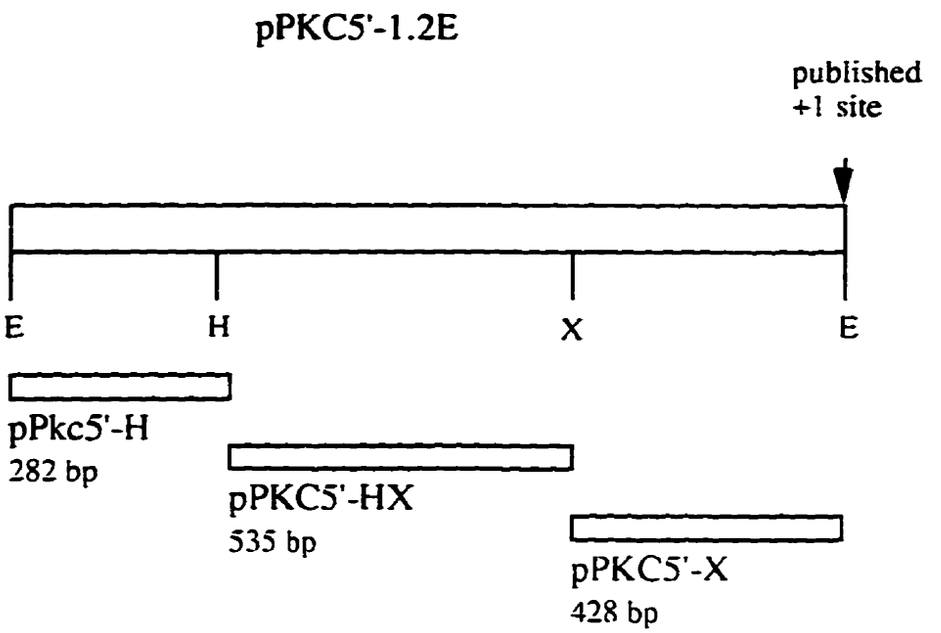

λ 8 genomic clone - 17 Kb
1
kb
non-
coding
exon #1
(+1) site
coding
exon #2
E E BH E C C E H E H B E
pSB29
E E
pPKC5'-1.2E
H H
pPKC5'-1.4H
B B
p5'-B

Figure 6. Subcloning and Sequencing strategies for pPKC5'-1.2E and pPKC5'-1.4H. A) pPKC5'-1.2E. Restriction digest analysis and preliminary sequence data revealed the presence of two unique cut sites, HinDIII and XhoI. Three plasmids were constructed using these cut sites, and sequenced. B) pPKC5'-1.4H. This fragment spans the published transcriptional start site. pPKC5'-0.8XH was created to confirm that the sequence from pPKC5'-1.2E was adjacent to the published sequence for the transcriptional start site of *Pkc53E*. B-BamHI, E-EcoRI, H-HinDIII, and X-XhoI.

A)


pPKC5'-1.2E
published
+1 site
E
H
X
E
pPkc5'-H
282 bp
pPKC5'-HX
535 bp
pPKC5'-X
428 bp

B)

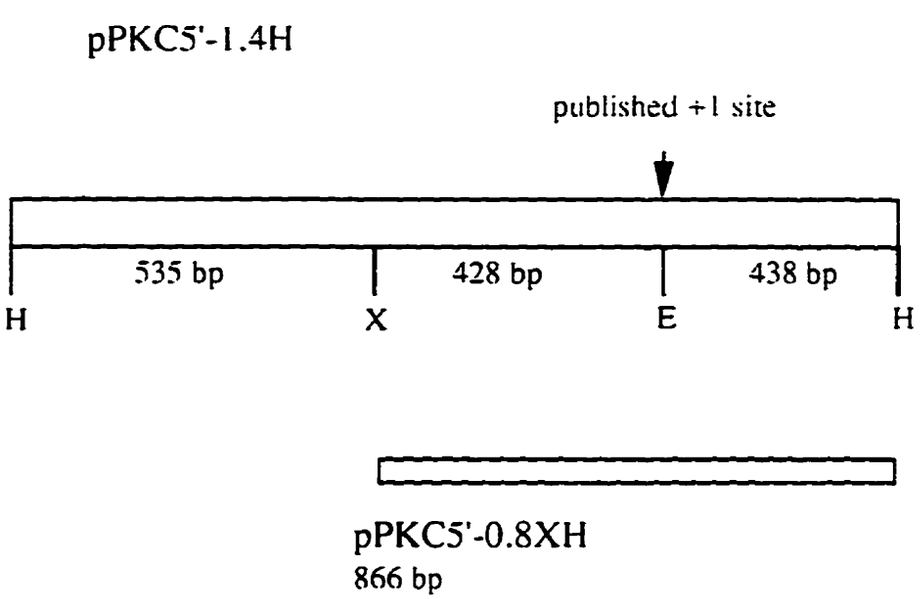

pPKC5'-1.4H
published +1 site
535 bp
428 bp
438 bp
H
X
E
H
pPKC5'-0.8XH
866 bp

Figure 7. Compiled sequence of pPKC5'-1.2E and pPKC5'-1.4H. Shown is 1646 bp of DNA sequence from the 5' region of *Pkc53E*. DNA sequence searches using TFsearch ver 3.1, Genmoredata:TFsite.Dat., and GCG show some sequence similarity to known transcriptional binding sites. Selected restriction cut sites are shown. +1 published transcriptional start site. Underlined sequence corresponds to sequence generated from the region surrounding the published +1 site (Rosenthal *et al.*, 1987)

———— restriction enzyme cut sites. E-EcoRI, H-HinDIII, X-XhoI

═══ consensus sequence for TATA boxes

▬▬▬▬ consensus sequence for insect transcriptional start site and Initiator sequence

——► Ap1 binding sequence (consensus sequence- TGAGTCA)

```
-1245   GAATTCCAAA GTTATTTTAT GATAACGTCG ATTTAGCAAC ATTAATAAAA
-1195   TAAAATATAT AACCTTATAG CTTGATACCC GAACAAAATT TAAAAATTCT
-1145   ACATATAATT AAATAGTCCA TAATTCCAGA GAATTTTCCA TGTTTTTAAT
-1095   ACATTAATTT TTTGCATTGA GTTGCCTTTT TTTAGGGCTT GGATAGTTAT
-1045   CTACTTTTTA TGAGGCGCAA CAGTATTTAT AAGGCGCGTC TCAGTGTTTC
-995    GCAATAATGT GATAAATGCG TCAAGGCAAT TAAAGCTTAT GCAAAATATA
-945    CATTGCTATT TTCATAAAGG CCTTTGTTTC ATTTCATTAA CATCTTGAGT
-895    TTTATATTAA TCATTGAAGT CTTGAGATTC AATTTTATTC GCTAAAGTTG
-845    TTAGAAAGTT GAGAGCCATT CAAATTGATA AGTATATACT AATTTCCCTA
-795    TCTCACAATC CGATCAGCCA CGAGGGGGAT TCCATTTATA GATCAAATCA
-745    CTTCTTGTTC AACGGCTAAT CGAATGGCTA GAGGCAGCCT TTGCAAACTC
-695    GGGCCCCAAA ACCGGCTGGT ATCGTGGCCA TTTAACAAGT TTCCGTTCAG
-645    TACGGTATCA ATTAGCGGCC AAGAGGTCAT CGGGTGCTCC AGGGTCTGAC
-595    GTTGGTCGGT ACGCTATGGG TTTACTTATG TTAATCAGTT GGCCGCAATG
-545    CAAATGAGGC GTTAACAGCG CTGTTGCTTT TGGCTAACAG TGCGCCCAAA
-495    AGCTCGCTAC GAATATGGCG CTCCAGAAAA GACGGCCATG TGCCAGTTGG
-445    TCAGAACGCG TTTGTGTCTC GAGCGAGACA GCGACGTGAC CGGGAAAGCG
-395    ATACAAATTC GCAGTGCCAG GAGGCCACAA TAACAACAAC AACAGATGTG
-345    TGTGAGTGTT AGTGCGTGCG AGAAGAAATC TTGTGTAACC CGGAGAGCGA
-295    GAAAGAGGGA GACAGGCAAC ACATGGGAAA TTCACCACAG TCCATGGACA
-245    TGGATTTTCC GCCTCGTGTG AAAATGCTCG CACGCTCCAT ACGTTTCGGT
-195    TCGCCAGGCA AAGGCAGTTG CTCCTCGCAT TTTCGCAGCC CGTGCGTGTA
-145    GTGTAGGGTA GTCTAAATAA GTACCAATCA CACTAAGTCA TCGGTCAAAA
-95     AGTAAACAAA ATAAACGCCA CATGTCGTTA TGAACAAGTG CTTGAAACTA
```

E H X

```
                                                                    +1
-45     GACAAGTTAC TCCAAATTCA AAGCACTATT AACCACAAAG AATTCTAATA
                                                              E
+6      ATACCCCTCG GGATAATAAC GAAAAAGCAG TCATAGCCAT TATGGCCATT

+56     AAGCACAAAA AAGTGACATG AAAACGTGAA GGGAAATGTG TTCCACAAAT

+116    GGAGAAGAAA AAAAAACGCG CTAACGCAAG TCGTTCGGAA ATGATAGTAA

+166    CTGTTGCATT TCGCTCACAA AGACTATGAT AGTTTTGCTA ACACAGAGTT

+206    TCAGTCGAGT CCTGAATTGT CCTTTTTGTG GTCTAGTTTG CGGTTGCAAT

+256    CAATGCTGAA TATGTTGTAT GCAGGGTAGC AATAGAGTTG TACTTACTGC

+316    TGATTAGATA ATTAAATGTA CATCTTGGAT AGATGAAAAT GTTCTTGCTT

+366    TAAAATGTTT CAATTAGTTC ATTGATTCAG TTGAACCAAT TATTTTACAA

+406    TTAGCCAACA TTTCTTGATA TAATTAAAGC TT
                                          H
```

The pDelta1 vector is a 7.9 kb plasmid that is specifically made for making nested deletions using the transposon γδ and selectable markers (BRL). A 9 kb fragment that encompasses the first two exons, the +1 site and the 5' regulatory regions was subcloned into pDelta1 and nested deletions were made. Seven deletion clones were made from the original plasmid, p5'-B. The plasmids range in size from 16 kb to 9 kb, therefore the deletions of the original fragment range in size from 8 kb to 2 kb. All the constructs that were made are listed in Table 5.

Sequence analyses to identify known regulatory sites were performed using TFsearch ver 1.3 ((c)1995 Yutaka Akiyama (Kyoto Univ.): http://www.genome.ad.jp/SIT/TFSEARCH.html) and Genmoredata.Tfsite.Dat. (A. Hessel, personal communication). These databases are set up exclusively for the identification of sequences that correspond to known mammalian and arthropod transcriptional regulatory binding sites. The "Best Fit" and "Fasta" programs of GCG (NIH Genetics Computer Group) were also used to compare the 1245 bp sequence against known *Drosophila* promoter sequences and the sequences in the NCBI (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform=1) databases respectively. All sequences reported are relative to the published transcriptional +1 site of *Pkc53E*. Analysis of the 1245 bp fragment showed the presence of a conserved cap site/Inr *Drosophila* sequence -555 (TAATCAGTT). Other sequence elements found that could be functional are the consensus TATA boxes starting at -1000 bp (Figure) and the Ap1 site (TAAGTCA) which differs from the consensus Ap1 site, TGAGTCA, by one base pair. The published +1 site was determined by sequence analysis of a cDNA clone but multiple transcripts were found at varying developmental stages, suggesting that more that one transcriptional start site was present (Rosenthal *et al.*, 1987; Natesan, 1991). The sequence analysis presented here would suggest that the sequence corresponding to a conserved +1 site could be a previously undefined transcriptional start site for one of the transcripts. The sequence was also compared against 250 known basal *Drosophila* promoter sequences (compiled by Arkhipova, 1995) using the "Fasta" and "Align" program (NIH Genetics Computer Group). With different stringencies (word size ranging from 6 letters to 3 letters) used there was only minimal

Table 5. Plasmids constructed from the genomic clone, λ8. Fragments of genomic DNA from *Pkc53E* were subcloned into vectors that were either sequenced or analyzed for size.

Plasmid	Vector	Origin and insert
pPKC5'-1.2E	pGem7zf+	1245 bp fragment from the 5' upstream regulatory region of *Pkc53E*
pPKC5'-H	pGem7zf+	282 bp EcoRI-HinDIII fragment from pPKC5'-1.2E
pPKC5'-HX	pGem7zf+	535 bp XhoI-HinDIII fragment from pPKC5'-1.2E
pPKC5'-X	pGem7zf+	428 bp XhoI-EcoRI fragment from pPKC5'-1.2E
pPKC5'-1.4H	pBluescript SK+	1401 bp fragment that includes 438 bp of the 5' untranslated region and the published +1 site
pPKC5'-0.8XH	pBluescript SK+	866 bp fragment from pPKC5'-1.4H
p5'-B	pdelta1 (BRL)	a 9 kb fragment containing the first two exons, the published +1 and the upstream regulatory regions of *Pkc53E*. Deletions were made from this construct.
p5'Δ-B1	pdelta1 (BRL)	The smallest deletion of p5'-B. Plasmid size is 17 kb
p5'Δ-B2	pdelta1 (BRL)	Plasmid size 16 kb
p5'Δ-B3	pdelta1 (BRL)	Plasmid size 14 kb
p5'Δ-B4	pdelta1 (BRL)	Plasmid size 13 kb
p5'Δ-B5	pdelta1 (BRL)	Plasmid size 11 kb
p5'Δ-B6	pdelta1 (BRL)	Plasmid size 10 kb
p5'Δ-B7	pdelta1 (BRL)	Plasmid size 9 kb. The largest deletion of p5'-B

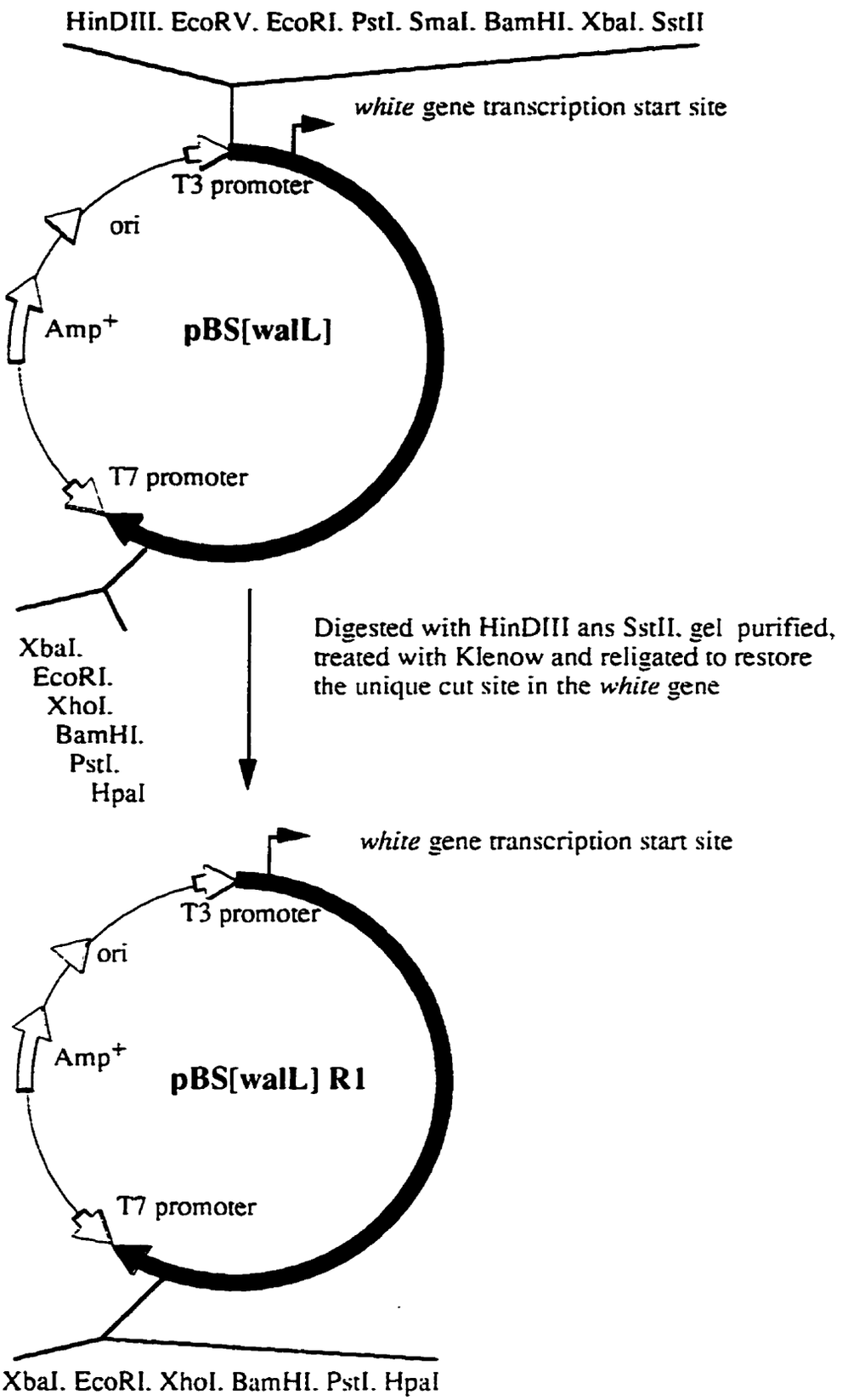
regions of homology with the 5' regions of the *Drosophila* promoter sequences. Sequence comparisons between the 1.2 kb fragment and the sequences from the mammalian cPKC 5' regions revealed no significant regions of homology. Comparison between the 1245 bp fragment and the basal promoter sequence structure (Arkhipova, 1995) identified the Inr sequence, but did not confirm the presence of a TATA box in the canonical location nor the presence of a down stream promoter element (DPE). The "Fasta" sequence queries compared the 1245 bp sequence against sequences in the nucleotide database at NCBI (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform=1). This search revealed no significant regions of homology with any other organism.

Construction of a Non-Transposable Reporter Vector

The original concept for the *in vivo* studies was to use a known P-element based reporter vector to introduce the promoter-reporter construct into the genome. With the characterization of the gene targeting method to the white locus (Engels *et al.*, 1990; Gloor *et al.*, 1991; Nassif *et al.*, 1994; Keeler *et al.*, 1996), it was believed that we could take this application one step further and use it as a tool for transcriptional studies. The non-transposable vector, pBS[walL] had to be modified to contain a known reporter gene used in *Drosophila* transcriptional studies (Figure 8). To remove the MCS of the original Bluescript vector from pBS[walL] (7252 bp), the vector was digested with HinDIII and SstII. The trailing ends were blunt ended and the plasmid ligated creating a unique MCS in the white gene of pBS[walL] (pBS[walL]:R1). The reporter construct containing the LacZ reporter gene was taken from the transposable vector, pCaSpeR-AUG-bgal (Thummel *et al.*, 1988). This vector was digested with EcoRI and PstI to release a 4.4 kb fragment that contained a known *Drosophila* start codon, *Adh* -AUG, upstream of the reporter gene, LacZ, and an SV40 polyadenylation recognition sequence. pBS[walL]:R1 vector was digested with EcoRI and PstI and gel purified. The 4.4 kb reporter fragment and digested pBS[walL]:R1 vector were then ligated together to create pBS[walL]:LacZ. The non-transposable

Figure 8. Subicloning strategy to create the non-transposable reporter vector, pBS[walL]:LacZ. pBS[walL] was modified as follows: A) The vector was digested with HinDIII and SstI to remove the multiple cloning site of pBluescript and to restore the unique cut sites in the 3' end of the white gene creating pBS[walL]R1. B) To insert the reporter gene, lacZ, pBS[walL]:R1 was digested with EcoR1 and PstI. The P-element vector pCaSpeR-AUG-βgal was digested with EcoR1 and PstI to release a 4.4 kb fragment containing the Drosophila *Adh*-AUG, the lacZ gene and the SV40 sequence. The digested pBS[walL]R1 vector and 4.4 kb fragment were ligated creating pBS[walL]:LacZ with three unique cut sites (outlined).

A)


HinDIII. EcoRV. EcoRI. PstI. SmaI. BamHI. XbaI. SstII
white gene transcription start site
T3 promoter
ori
Amp+
pBS[walL]
T7 promoter
XbaI.
EcoRI.
XhoI.
BamHI.
PstI.
HpaI
Digested with HinDIII ans SstII. gel purified, treated with Klenow and religated to restore the unique cut site in the white gene
white gene transcription start site
T3 promoter
ori
Amp+
pBS[walL] R1
T7 promoter
XbaI. EcoRI. XhoI. BamHI. PstI. HpaI

B)

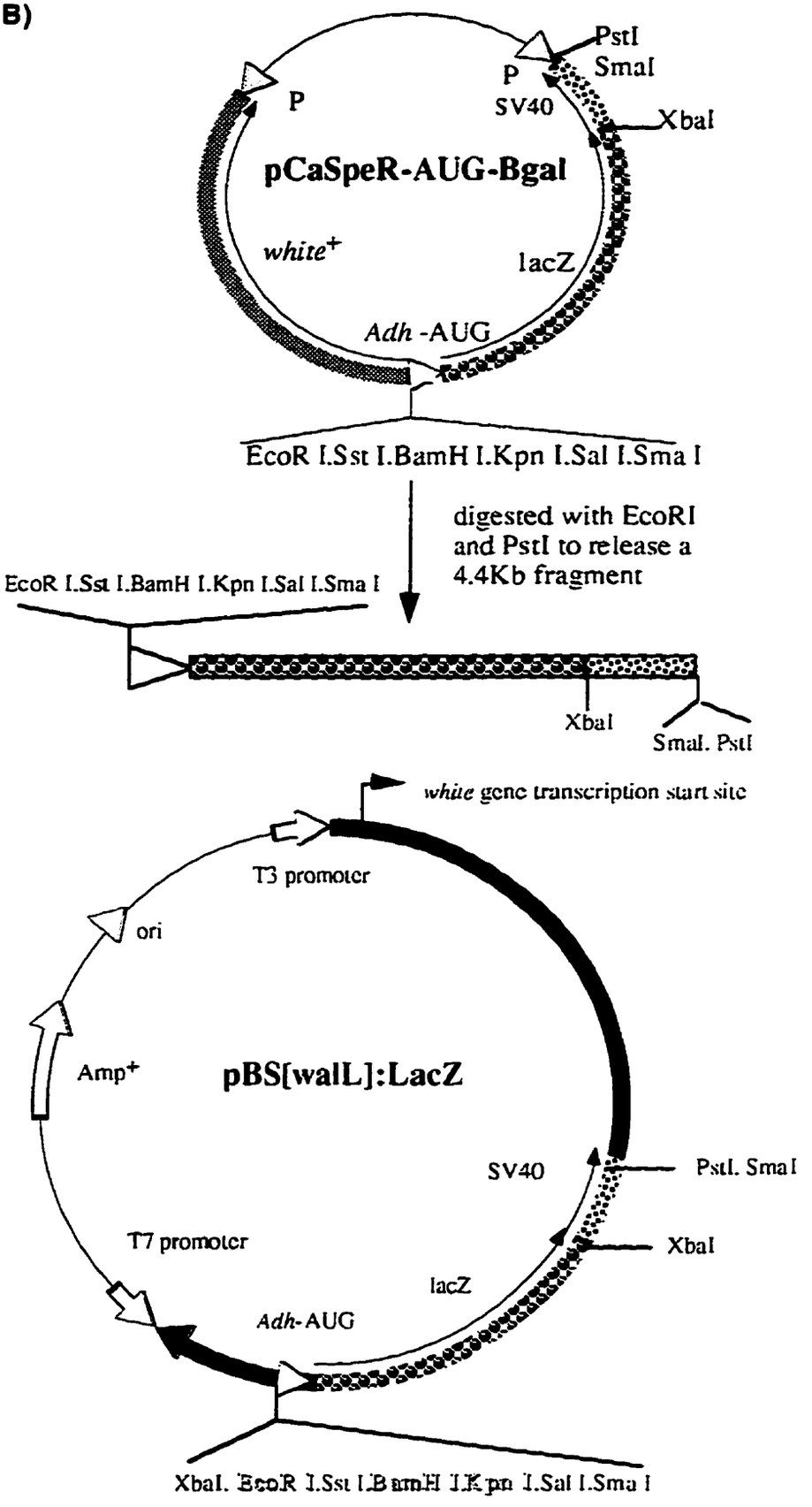
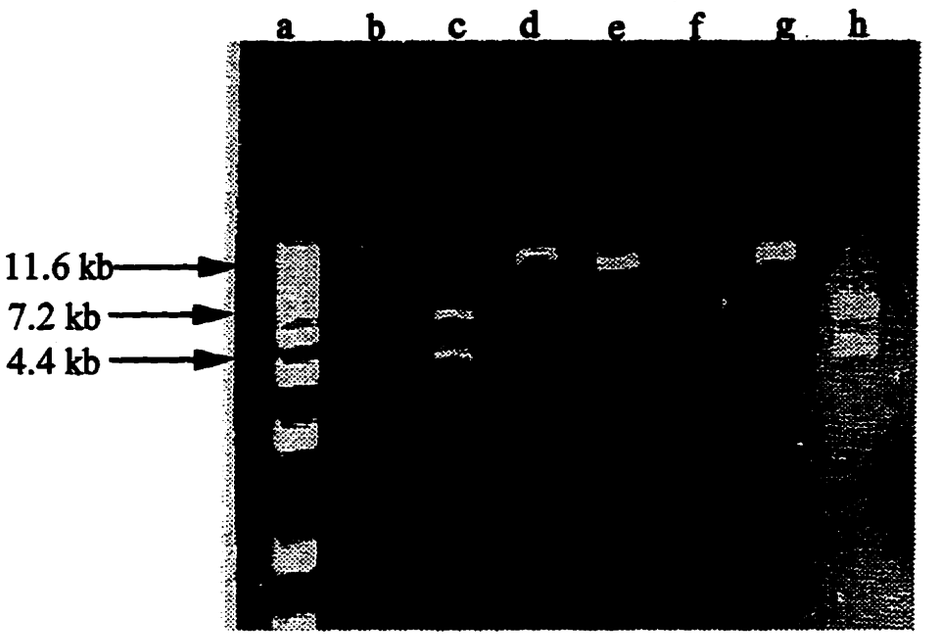
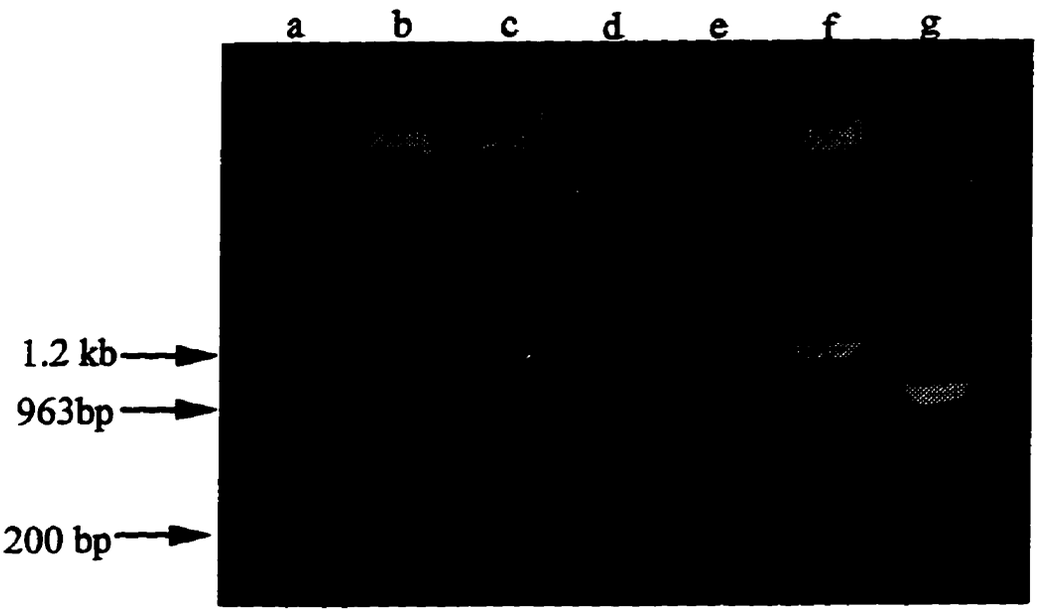

PstI
SmaI
P
P
SV40
XbaI
pCaSpeR-AUG-BgaI
white+
lacZ
Adh -AUG
EcoR I.Sst I.BamH I.Kpn I.Sal I.Sma I
digested with EcoRI
and PstI to release a
4.4Kb fragment
EcoR I.Sst I.BamH I.Kpn I.Sal I.Sma I
XbaI
SmaI. PstI
white gene transcription start site
T3 promoter
ori
Amp+
pBS[walL]:LacZ
SV40
PstI. SmaI
XbaI
T7 promoter
lacZ
Adh-AUG
XbaI. EcoR I.Sst I.BamH I.Kpn I.Sal I.Sma I

Figure 9. Restriction analysis of pBS[walL]:LacZ. The samples were run on a 1% agarose, 1x TAE gel. Lane c shows the respective fragments from pBS[walL]:LacZ R1 (7.2 kb) and the reporter gene, LacZ (4.4 kb) when pBS[walL]:LacZ is digested with EcoRI and PstI. Lanes d, e, and g confirm the presence of the unique cut sites, EcoRI, BamHI, and KpnI in the MCS of pBS[walL]:LacZ.



a- 1 kb DNA size ladder
b- undigested pBS[walL]:LacZ
c- EcoRI/PstI digested pBS[walL]:LacZ
d- EcoRI digested pBS[walL]:LacZ
e- BamHI digested pBS[walL]:LacZ
f- XhoI digested pBS[walL]:LacZ
g- KpnI digested pBS[walL]:LacZ
h- SmaI digested pBS[walL]:LacZ

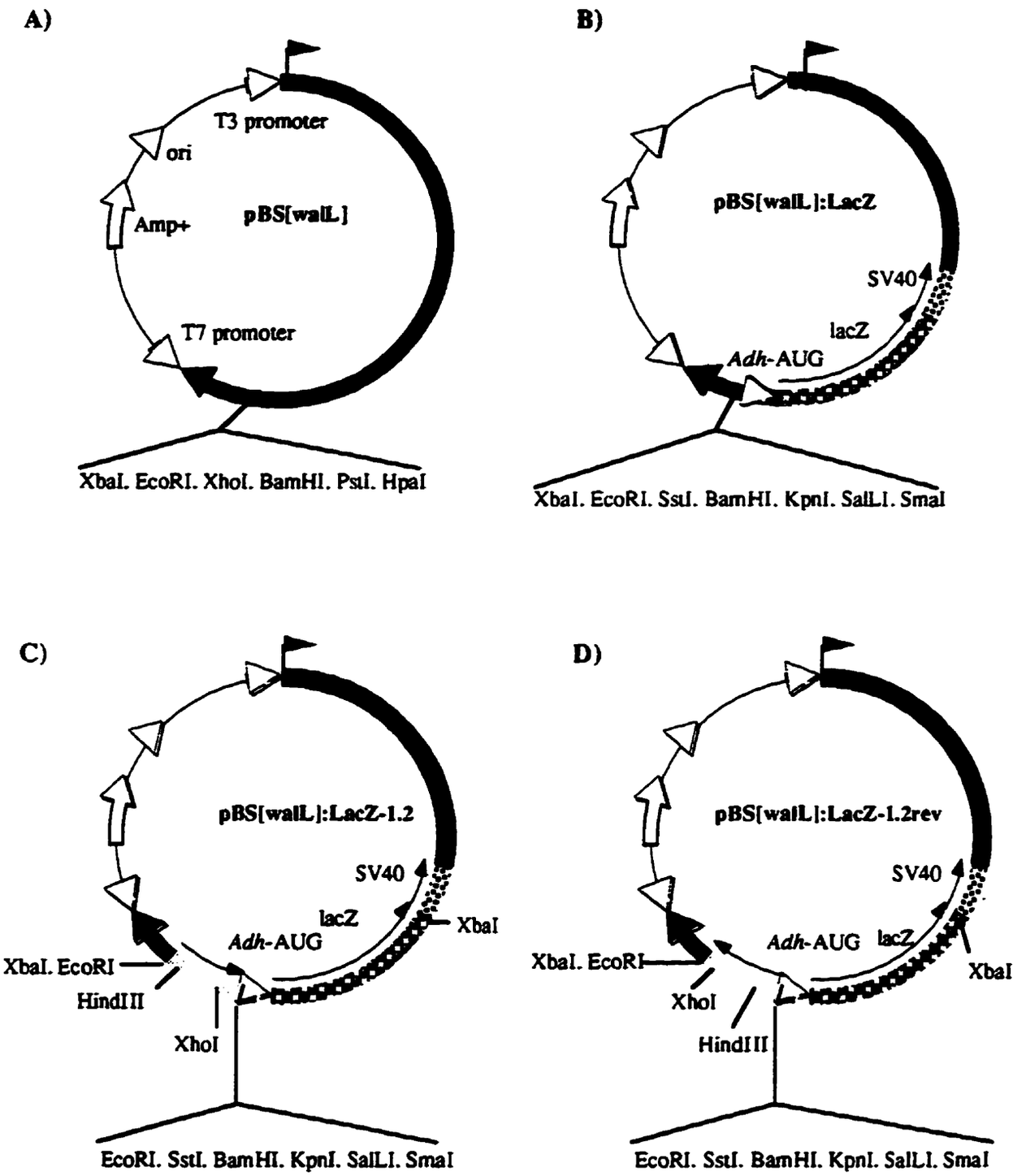
Figure 10. Restriction analysis of pBS[walL]:LacZ-1.2 and pBS[walL]:LacZ-1.2rev. To show that the 1.2 kb promoter fragment was subcloned into the EcoRI site, the two plasmids were digested with EcoRI to release the 1.2 kb fragment (c and f). To show orientation, the two plasmids were digested with HinDIII and XbaI. There is one HinDIII sight present in the 1.2 kb fragment and two XbaI sites present in the vector (shown in Figure 11 c and d). When the 1.2 kb fragment is in the correct orientation the vector will release a 200 bp fragment (d). When the 1.2 kb fragment in the reverse orientation then the vector will release a 963 bp fragment (g).



a- 1 kb DNA size ladder
b- undigested pBS[walL]:LacZ-1.2
c- EcoRI digested pBS[walL]:LacZ-1.2
d- HinDIII/XbaI digested pBS[walL]:LacZ-1.2
e- undigested pBS[walL]:LacZ-1.2rev
f- EcoRI digested pBS[walL]:LacZ-1.2rev
g- HinDIII/XbaI digested pBS[walL]:LacZ-1.2rev

Figure 11. Schematics of the 4 injected templates. A) pBS[walL], the template used by Nassif *et al.* (1996). B) pBS[walL]:LacZ, the non-transposable template with the reporter construct. C)pBS[walL]:LacZ-1.2, contains the sequenced 1.2 kb EcoRI 5' fragment of *Pkc53E* in the proper orientation. D) pBS[walL]:LacZ-1.2rev, contains the sequenced 1.2 kb EcoRI 5' fragment of *Pkc53E* in the opposite orientation..

white transcription start site


A)
T3 promoter
ori
Amp+
pBS[walL]
T7 promoter
XbaI. EcoRI. XhoI. BamHI. PstI. HpaI
B)
pBS[walL]:LacZ
SV40
lacZ
Adh-AUG
XbaI. EcoRI. SstI. BamHI. KpnI. SalLI. SmaI
C)
pBS[walL]:LacZ-1.2
SV40
lacZ
XbaI
Adh-AUG
XbaI. EcoRI
HindIII
XhoI
EcoRI. SstI. BamHI. KpnI. SalLI. SmaI
D)
pBS[walL]:LacZ-1.2rev
SV40
Adh-AUG
lacZ
XbaI
XbaI. EcoRI
XhoI
HindIII
EcoRI. SstI. BamHI. KpnI. SalLI. SmaI

reporter construct pBS[walL]:LacZ contains three unique cut sites, EcoRI, BamHI, and KpnI that are adjacent to the *Adh* -AUG start codon. Figure 9 is a restriction digest of pBS[walL]:LacZ to confirm the size and the presence of the unique cut sites that were created in the plasmid

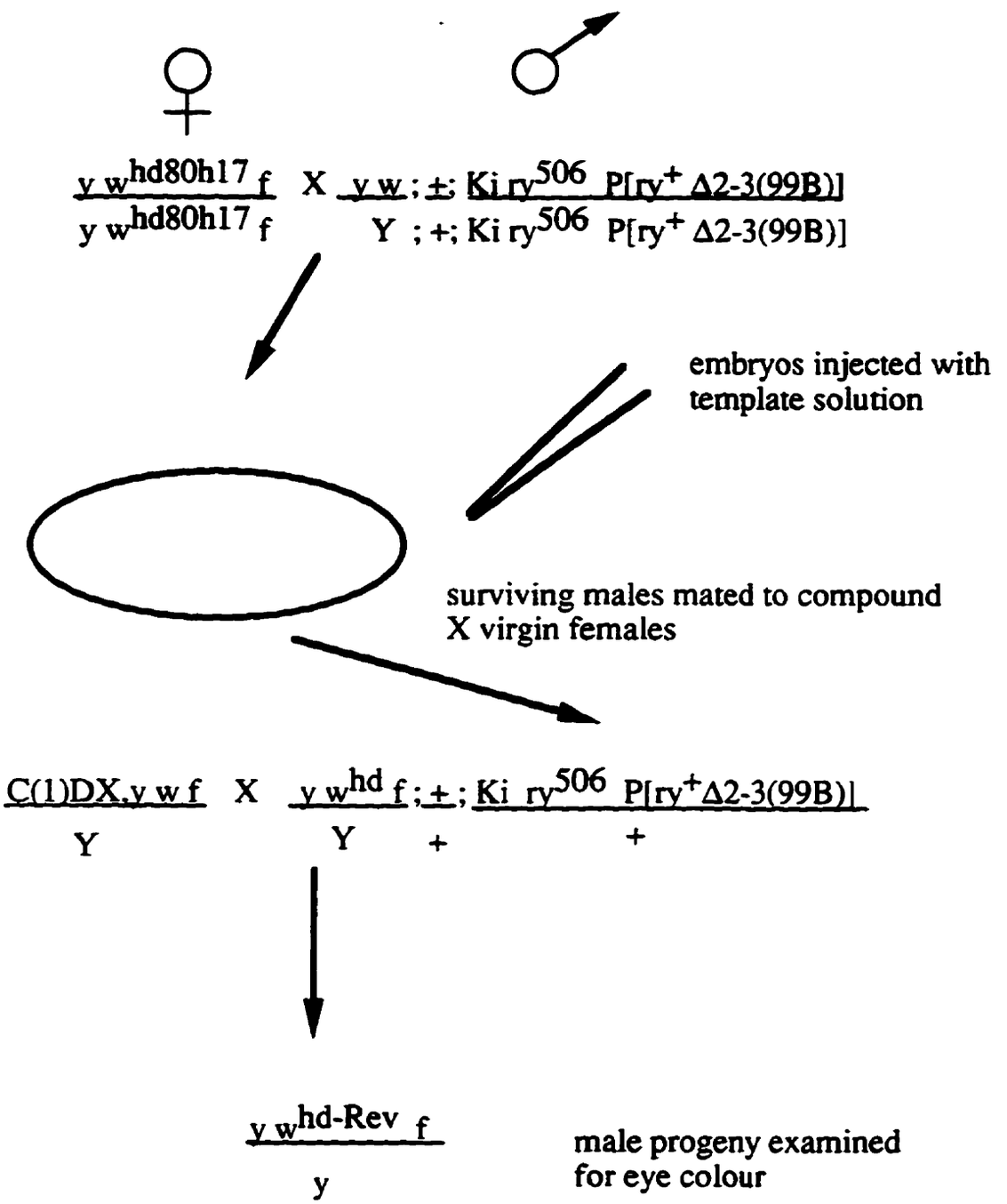
The 1245 bp putative promoter fragment was subcloned from pPKC5'-1.2E into the EcoRI site in both the correct and opposite orientation. The orientation of the promoter fragment was determined by restriction analysis using HinDIII and XbaI (Figure 10). These two constructs, as well as the original vector pBS[walL], and pBS[walL]:LacZ were used as constructs for microinjection (Figure 11) of *Drosophila* embryos.

Gene Targeting Studies

Gene targeting to the white locus is dependent upon three factors. The first is the presence of a transposable element at the white locus. In this experiment it is a 629 bp P-element found in exon 6 of the white gene (*Drosophila* stock: *y* $w^{hd80h17}$ *f*). The second is a transposase source that will cause non-replicative excision of the P-element (*Drosophila* stock: *y w; Ki* ry^{506} P[$ry^{+}\Delta$ 2-3(99B)]). For these experiments an endogenous transposase source (Δ2-3) was introduced in the initial cross. The last requirement for gene targeting to a DSB is a template in order for repair. In this experiment four different templates were introduced separately into the germline of *Drosophila* embryos.

As a control, the complete set of crosses (shown in Figure 12) were performed without microinjection of the embryos. Males from the initial cross were collected, group mated to compound X females and the progeny were examined. From the initial crosses 30 males, 3 of which had mosaicism in one of the eyes, were mated to compound X females. The male progeny of this control cross were scored in order to detect any abnormalities that might arise as a product of the cross. The results of this cross were not reported by Gloor *et al.* (1991) or Nassif *et al.* (1994) and there was no mention of the presence of the mosaicism that was observed in these studies.

Figure 12. Mating scheme for gene targeting. Females containing the w^{hd} allele were mated to males that contained the endogenous source of transposase (P[$ry^{+}\Delta 2$-3(99B)]) which allows P-w^{hd} to excise. The embryos were then injected with template DNA at a concentration of 1 mg/ml. The embryos were grown to adulthood and males were mated to compound X females and the male progeny examined for eye phenotype. Those that had the parental white eye phenotype were discarded. Males that were red eyed (indicating loss of the P-w^{hd} element) were collected and mated to compound X females creating a true breeding line. Males that had the mosaic eye phenotype were also collected and mated to compound X females.



Males with red eyes and mosaic eyes were mated to compound X-females to establish a convertant line.

Figure 12 is a schematic of the crosses and microinjections performed in order to fill all three requirements for gene targeting. Virgin females carrying the P-element based *white* allele, w^{hd}, were crossed with males carrying an endogenous source of transposase. The embryos were collected from this cross and injected with the template constructs. Larvae from the injections were collected and placed on *Drosophila* media. Male flies that eclosed were mated with females from a compound X stock and at least two broods were analyzed from each cross to determine if the males were fertile. The embryos injected, larvae collected, males mated and the number of fertile males for each injected template are shown in Table 6.

Progeny of the single male matings are shown in Table 7. There were four possible phenotypes that could have been seen: y w f; y w Ki; y w Ki with mosaicism in the eyes and y w^+ f which indicates the presence of an insertion of the template containing the w^+ gene. Flies with the y w f phenotype appeared in greater number and represented the majority of the flies that eclosed in the first few broods. Males that had white eye colour were counted and discarded. Males that were mosaic for the eye phenotype or had the revertant eye colour were mated to compound X females. The flies containing the eye mosaicism were present in all the crosses performed. Only one male fly had a revertant eye colour and the correct phenotype of y w^+ f. This male came from the embryos that were injected with the pBS[walL]:LacZ-1.2rev template. It was mated with compound X virgin females to form a true breeding stock in which the males express the revertant eye colour.

The male progeny were counted and eye mosaicism was found in the frequency of 0.016 (70/4508). The frequency with which the mosaicism appeared is a result of the cross and not a result of the injected template (chi squared test: $0.5>p>0.1$; degrees of freedom: 4). The expected number of mosaics for each cross was calculated by multiplying the total frequency (0.016) by the total number of progeny for separate experiments. All flies that had the mosaic eyes also had the endogenous transposase source, identified by the dominant phenotypic marker Ki, present in the genome. Flies expressing the mosaicism from all the crosses were mated with compound X

Table 6. Survival of fertile males following microinjections. The males were mated with compound X females. Each fertile stock was kept for 7 broods.

	- control (no template)	pBS[walL]	pBS[walL]: LacZ	pBS[walL]: LacZ-1.2	pBS[walL]: LacZ-1.2rev
# of injected embryos	n/i	253	235	575	390
# of larvae collected	n/i	24	32	147	118
# of males collected and mated	30	8	10	23	12
# of fertile males	30	5	7	13	5

n/i- not injected

Table 7. Male progeny collected from the crosses between fertile male progeny of injected embryos and compound X females. Four phenotypes were expected except for the control cross.

Construct	y w f	y w; Ki *	y w; Ki * mosaic eye (fertile)	y w ^{rev} f	Total number male of progeny
control cross	589	384	13 (2)	0	986
pBS[walL]	360	118	5 (0)	0	483
pBS[walL]: LacZ	441	235	9 (1)	0	685
pBS[walL]: LacZ-1.2	981	619	34 (3)	0	1634
pBS[walL]: LacZ-1.2rev	427	283	9 (1)	1	720
Total	2798	1630	70 (7)	1	4508

y- yellow, w- white, f- forked, Ki- Kinked
mosaic- eye phenotype was a partial mosaic expressing red in a white background
w^{rev}- a revertant red eye colour indicating a precise excision of P-w^{hd} and incorporation of the nontransposable template

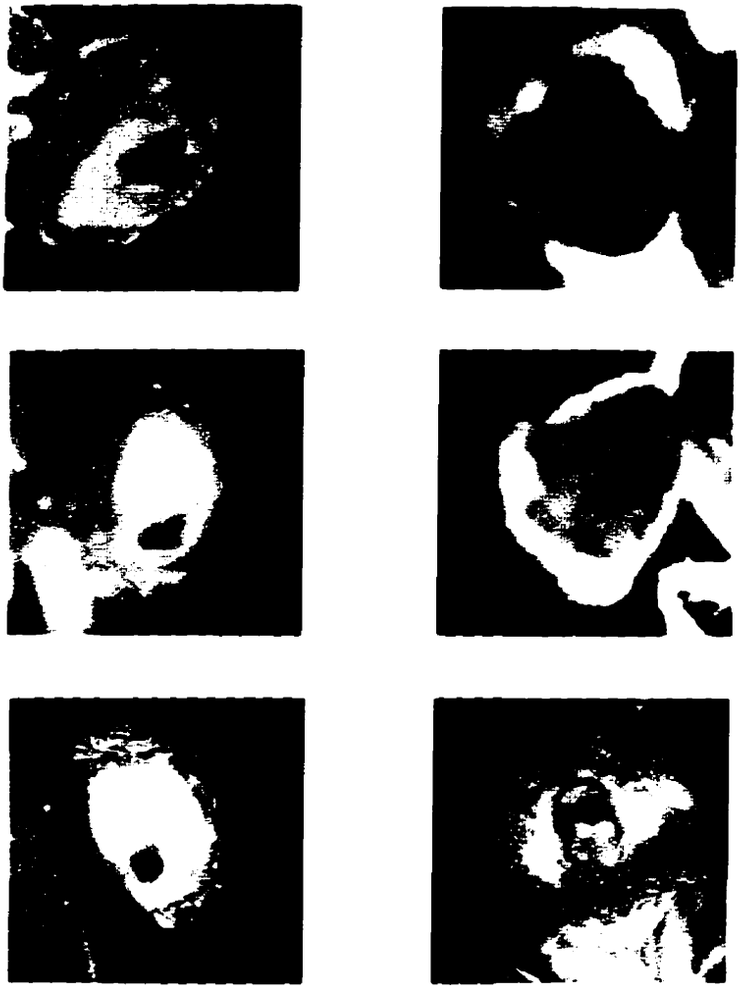
* genotypically forked is present but it is masked by Ki

females to determine if this mosaicism was a product of the cross (as in the control cross) or a product of the injected template being present. The male progeny of the seven fertile males were scored according to phenotypes: y w f; y w Ki; and y w Ki with mosaicism in the eye. There were no full revertants in the male progeny scored thus there were no reversions to red eyes. Figure 13A compares the phenotype of revertant flies (y w^{+}f) to the parental flies (y w f). Figure 13B depicts 6 typical examples of the observed mosaicism. Lankenau *et al.* (1996) reported this initial negative control cross with their experiments at the *forked* locus and reported the presence of mosaics in the macrocheate phenotype.

Figure 13. Comparison of male progeny phenotype. A) Comparison between a parental y w f male and a revertant y w^+ f male that contains the pBS[walL]LacZ-1.2rev insert. B) Six examples of mosaicism that were found in each of the crosses performed.

A)

B)



Discussion

In vivo transcriptional regulation studies performed using *Drosophila* have given insights into the importance of gene regulation throughout development. This study was aimed at defining the basal promoter needed for expression of *Drosophila Pkc53E*. Five different *Pkc53E* mRNA transcripts are temporally and spatially regulated (Rosenthal *et al.*, 1987; Natesan, 1991). This study was designed to take the first step in determining the promoter elements responsible for this specificity.

Sequence generated corresponding to the 5' end of *Pkc53E* shows the presence of many putative transcriptional regulatory sequences (Figure 7). Of primary interest is the conserved cap site/Inr *Drosophila* sequence (TAATCAGTT) at -555, that has been shown to be found adjacent to the start site of transcription for a number of *Drosophila* genes (Cherbas and Cherbas, 1993). The Inr sequence is found in one third of the *Drosophila* promoters sequenced to date (Arkhipova, 1995). This sequence binds the TATA box-binding protein associated factors (TAFs) in the absence of a TATA box and can act as the primary regulator of transcription (Martinez *et al.*, 1994; Martinez *et al.*, 1995; Burke and Kadonga, 1996). A TATA box was not found in the -25 to -30 bp range from either the published transcriptional +1 site or the Inr binding site sequence suggesting that this is a TATA less promoter region. Consensus TATA sequences were found starting at -1000 bp from the published +1 site. The published putative transcriptional +1 site was identified through the sequencing of one cDNA clone, the developmental stage of which was not reported (Rosenthal *et al.*, 1987). At least five *Pkc53E* mRNA transcripts (2.4, 3.0, 3.4, 4.0 and 4.3 kb) have been identified and are developmentally regulated and tissue specific (Table 2: Rosenthal *et al.*, 1987; Natesan 1991). The previous identification of several transcripts suggests the possibility of multiple transcriptional start sites. It is unclear which of the mRNA transcripts, if any,

correspond to the published +1. If it corresponds to the transcriptional start site of the 3.0 or 3.4 kb mRNA transcripts then the consensus Inr binding sequence and TATA boxes found are most likely non functional. But, if the complete 5' UTR has not yet been identified, then these consensus sequences could be important for the transcriptional regulation of this gene.

Another putative binding site present that could be important in the regulation of this gene is the TRE. This element is found -106 bp from the published +1 site (Figure 7). The TRE element is the recognition sequence for Ap1, a transcriptional factor that is either heterodimeric (cFos and cJun) or homodimeric (cJun) (Angel *et al.*, 1987). Ap1 becomes activated by many stimuli including cytokines, T cell activators, growth factors, and neurotransmitters (Angel and Marin, 1991). Ap1 is also activated by TPA, a known tumor promoting phorbol ester, through a PKC-dependent pathway (Hirai *et al.*, 1994). The transcriptional studies and sequence analyses performed on the mammalian cPKCs have shown the presence of Ap1 and Ap2 binding sites, which are known TREs (Chen *et al.*, 1990; Niino *et al.*, 1992; Obeid *et al.*, 1992; Takanaga *et al.*, 1995; Mahajna *et al.*, 1995). Human PKC β and human PKC γ genes were transcriptionally upregulated upon exposure to TPA, suggesting that certain PKC genes are responsive to tumor-promoting phorbol esters and that this response is mediated through a TRE (Obied *et al.*, 1992; Mahajna *et al.*, 1995). Previous studies of *Pkc53E* did not find mRNA transcripts in the 0-3 hour embryo (Rosenthal *et al.*, 1987; Natesan, 1991). Natesan (1991) found that exposure of the 0-3 hour embryos to TPA caused expression of the 2.4 kb and 3.4 kb transcripts. This finding coupled with the identification of a putative TRE sequence in the 5' region of *Pkc53E*, suggests that TPA regulates expression of *Pkc53E* through a TRE.

Sequence comparisons between the promoter sequence generated in this study and the transcriptional regulatory region of the mammalian cPKCs studied showed no regions of homology. Search queries using nucleotide databases (NCBI) showed that the

promoter sequence presented here is novel, demonstrating no significant homology with any of the sequences contained in the databases. Sequence analysis between the 5' sequence of *Pkc53E* and sequences in the *Drosophila* promoter database have shown that there were few or no regions of homology. The GC content of the PKC sequence was 40%, which was lower than that found on average for the entire *Drosophila* promoter database (Arkhipova, 1995). *Drosophila* genes that are expressed in a similar tissue specific and developmentally regulated way as *Pkc53E* [e.g. *Antp* (Perkins *et al.*, 1988); *en* (Soeller *et al.*, 1988); *elav* (Yao and White, 1994)], have promoter regions that have a low GC content and are TATA-less. This suggests that there is a subset of genes that contain promoter regions that are distinct for strict tissue specificity. Transcriptional studies and sequence analysis of other known PKC genes in *Drosophila*, yeast and *C. elegans* have not been performed, making it difficult to reach a conclusion about sequence homology for the promoter region between known PKC genes of lower eukaryotes.

In vivo Transcriptional Analysis via Gene Targeting

In order to determine the basal promoter region needed to drive expression of *Pkc53E*, *in vivo* gene targeting experiments were attempted. To produce a quantifiable, reproducible transcriptional assay system, we altered the gene targeting technique characterized by Engels *et al.* (1990) and Gloor *et al.* (1991). DSBs as a result of nonreplicative P-element excision result in a repair mechanism that can be utilized to target fragments of DNA to where the break occurred (Engels *et al.*, 1990; Gloor *et al.*, 1991). The requirements for the mechanism are three fold: A known P-element insertion site, a source of transposase and a template that can be used to copy DNA of interest into the break site. The P-element that was chosen for excision in this study was the 629 bp P-element insertion into exon 6 of the *white* gene, $w^{hd80k17}$. This construct was chosen

for a number of reasons for transcriptional studies. The P-element at this locus has been extensively characterized for P-element mediated gene targeting (Engels *et al.*, 1990; Gloor *et al.*, 1991; Nassif *et al.*, 1994; Keeler *et al.*, 1996). The P-element insert has a high rate of mobility, increasing the number of repair events that can occur at this locus (Engels *et al.*,1990). A reversion event, white eyes to red eyes, is easily identifiable indicating insertion. The location of the P-element is in exon 6, downstream of the promoter that drives the *white* gene thus reducing the likelihood of effects on *Pkc53E* transcriptional control elements.

The nontransposable construct was chosen for two reasons. The number of crosses that had to be performed to identify an insertion event were fewer than those needed for transposable templates (Banga and Boyd, 1994; Keeler *et al.*, 1996). The second benefit of this method was that the nontransposable template was an extrachromosomal array and, as such, alteration of the genetic sequence by the embryonic machinery should be minimal. If the template was recognized it would be copied directly into the break site without going through any intermediary changes. The drawback of using a nontransposable template is the lower rate of reversion observed. In the two studies done there was a 0.5% reversion rate (Banga and Boyd 1994; Keeler *et al.*, 1996) compared to a 6% reversion rate for transposable templates (Gloor *et al.*, 1991; Nassif *et al.*, 1994; Keeler *et al.*, 1996).

Table 8 is a comparison of the experiments done thus far using non-transposable templates for gene targeting to the *white* locus. Banga and Boyd (1992) used 50 base, single stranded oligonucleotides and 50 bp, double stranded oligonucleotides, to target to the *white* locus. The frequency of reversion produced in the progeny was 3.9×10^{-3}. Keeler *et al.* (1996) observed a reversion frequency of 6.1×10^{-3} with three nontransposable templates. Two templates had altered *white* gene sequence for identification purposes (pBSwalter and pBS[walL]) and the third, pBS[walL]y , contained

Table 8. Summary of the data using non-transposable templates for gene targeting. Previous to this there were only two other groups of researchers that had accomplished gene targeting with a non-transposable template (Banga and Boyd, 1994; Keeler *et al.*, 1996).

Template	Fertile males	Progeny	Convertants
pBSwalter[a]	37	2,229	5
pBS[walL][a]	24	2,229	9
pBS[walL]y[a]	12	970	21
single stranded oligo[b]	24	923	1
double stranded oligo[b]	30	2600	13
pBS[walL]	5	483	0
pBS[walL]:LacZ	7	685	0
pBS[walL]:LacZ-1.2	13	1634	0
pBS[walL]:LacZ-1.2rev	5	720	1

a- Keeler *et al.*, 1996
b- Banga and Boyd, 1994

heterologous DNA, a wild type copy of the yellow gene. The rate of reversion was 2.8 x 10^{-4} for the gene targeting experiments for *Pkc53E* performed with the non-transposable reporter constructs and the original pBS[walL] vector. The reasons for the lack of convertants could be due to a number of conditions. The microinjection temperatures for Keeler *et al.* (1996) and Banga and Boyd (1992) were at 27°C and 29°C, the injection temperature in this study was at 20°C. The reduced temperature was used in order to stabilize the embryos and the time frame to inject into the pole plasm before cellularization of the embryo. Although the number of injected embryos in this study was relatively large, 1453 embryos injected, the number of fertile males and their subsequent progeny was significantly smaller than those recorded from Banga and Boyd (1992) and Keeler *et al.* (1996) (Table 8). The crosses that were performed were identical to the crosses performed in both these experiments. Lethality has been reported to be a result of the microinjection technique and the transposase source used (Engels *et al.*, 1987). Microinjection is invasive, reducing the chance of survival. The transposase source, Δ2-3, results in pupal lethality and gonadal sterility (Engels *et al.*, 1987). Although microinjections and Δ2-3 have been used extensively in P-element transposition studies, there is still a low rate of survival for transgenic flies. This transposase source is stably inserted on the 3rd chromosome and several successful experiments have been performed using this transposase source and the gene targeting system (Gloor *et al.*, 1991; Nassif *et al.*, 1994; Keeler *et al.*, 1996).

The presence of mosaicism in the male flies from both the control cross and from the progeny can be explained by somatic mosaicism. The transposase source, Δ2-3, is expressed in both the soma and germline cells. The repair of the resulting gap, will, in some of the cells, result in a functional gene product leading to the expression of wild type colour in the phenotypic background. The observation that all the mosaic flies still had the transposase source present (indicated by the presence of the phenotypic marker,

Ki) and only 7 of the 30 males were fertile suggests that the transposase effected the mosaicism and fertility of these flies. The progeny of the control cross nor the presence of mosaics were presented in any of the papers using the *white* locus (Engels *et al.*, 1990; Gloor *et al.*, 1991; Nassif *et al.*, 1994; Banga and Boyd 1994; Keeler *et al.*, 1996). The studies performed with gene targeting to the forked locus reported progeny of the control cross and observed the presence of forked mosaics with the transposase source present (Lankenau *et al.*, 1996). They mated these flies and also found a reduced viability. Their conclusions were that the mosaicism was a result of somatic mosaicism and the transposase source used. They concluded that only full revertants to the wildtype phenotype should be analyzed.

Conclusions

Three fragments, 1245 bp, 1401 bp and 9 kb, corresponding to the 5' regulatory region and part of the transcribed region for *Pkc53E* were subcloned from the genomic clone λ8 and two fragments were sequenced. The plasmid, pPKC5'-1.2E contained the sequence that was immediately 5' of the published putative transcriptional start site. To confirm that this sequence was continuous with the published 5' UTR the plasmid pPKC5'-1.4H was sequenced. Sequence analyses performed identified various transcriptional regulatory elements, TATA boxes, a TRE, and an Inr sequence, that could play roles in the developmental and tissue specificity that has been observed (Rosenthal *et al.*, 1987; Natesan 1990).

To determine the basal promoter region by *in vivo* analysis, the gene targeting method to the *white* locus was employed. The nontransposable targeting template, pBS[walL], was altered to contain a reporter gene, LacZ, with unique cut sites that are located upstream of the reporter gene. The modification and subsequent utilization of the gene targeting method was initiated in order to circumvent the effects of endogenous

silencers and enhancers and p.e.v. cited by other reported *in vivo* P-element based transcriptional studies. The microinjections that were performed using the four templates did not produce a significant number of fertile male progeny, this was due to the invasiveness of the procedure and the source of transposase. The positive outcome was that there was a conversion event that occurred with the pBS[walL]:LacZ-1.2rev template, proving that targeting with this reporter construct is possible. It is still unknown as to whether the reporter construct, pBS[walL]:LacZ, will be functional using this technique at this locus. But the potential for gene targeting as a method to control for the effects of p.e.v and effects of endogenous silencers and enhancers as well as reducing the number of crosses that have to be performed and the number of flies that have to be screened leads to the exciting possibility that this technique could be a very powerful tool for doing *in vivo* transcription studies.

Further Experiments

To determine whether there are silencers that could suppress expression at the location of the w^{hd} insert in the *white* gene, an inducible promoter whose expression patterns have been characterized should be employed. For example, the *Drosophila* hsp70 promoter could be utilized. Reversions found in the injected males progeny with the hsp70-pBS[walL]:LacZ construct could be transcriptionally induced by increasing the temperature and the flies stained for β-galactosidase. Staining expression and concentration could then be compared to known studies to determine if this reporter construct will work.

The *Pkc53E* sequence presented has many putative transcriptional regulatory elements. To first determine the transcriptional start site(s), S1 nuclease protection assays, primer extension and 5' RACE have to be performed to identify where and if there are multiple transcriptional start sites. If transcriptional start sites extend further 5' then

the sequence presented here then the deletion clones , pΔ5'-B1 to pΔ5'-B7 can be used. To identify the basal promoter region of *Pkc53E*, the fragment that contains all the elements could then be subcloned into the nontransposable reporter vector and targeted to the *white* locus. The revertants could then be stained at different developmental stages for the expression of β-galactosidase. Once the basal promoter region has been identified then experiments determining the regulatory elements needed for developmental and tissue specificity could begin. To determine if the TRE sequence that has been identified mediates the response to TPA, 0-3 hour embryos containing the basal promoter-reporter fragment with the TRE could be exposed to TPA and stained for expression of β-galactosidase. The same promoter fragment with a mutated TRE should also be used to show that transcriptional activation of *Pkc53E* by TPA induction is dependent upon the presence of a functional TRE. This project took the first step in trying to characterize the basal promoter region of *Pkc53E*. It was also aimed at using the well characterized gene targeting technique to the *white* locus as a tool for doing *in vivo* transcriptional analysis.

Literature Cited

Angel, P., M. Imagawa, R. Chui, B, Stein, R.J. Imbra, H.J. Ramsdorf, C. Jonat, Pherrlich and M. Karin. 1987a. Phorbol ester-inducible genes contain a common *Cis* element recognized by a TPA modulated Trans-acting factor. Cell. 9: 729-739.

Angel, P., I. Baumann, B. Stein, H. Deluis, H.J. Ramsdorf and P. Herrlich. 1987b. 12-*0*-tetradecanoyl-phorbol-13-acetate induction of the human collagenase gene is mediated by an inducible enhancer element located in the 5'-flanking region. Mol. and Cell. Biol. 7:256-2266.

Angel, P and M. Karin. 1991 The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim Biophys Acta 1072, 129-57 (1991)

Arkhipova, I.R. 1995. Promoter elements on *Drosophila* melanogaster revealed by sequence analysis. Genetics. 139: 1359-1369.

Ashburner, M. 1989. *Drosophila* : a laboratory manual. Cold Spring Harbor Press; Cold Spring Harbor, New York.

Banga, S. S. and J. B. Boyd. 1992. Oligonucleotide-directed site-specific mutagenesis of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. 89:1735-1739.

Boyle, J.S. and A.M. Lew. 1995. An inexpensive alternative to glass milk for DNA purification. TIG 11: 8.

Burke, T.W. and J.T. Kadonga. 1996. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes and Devel. 10: 711-724.

Buttgeriet, D. and R. Renkawitz-Pohl. 1993. Expression of β1 tubulin (βTub56D) in *Drosophila* testes stem cells is regulated by a short upstream sequence while intron elements guide expression in somatic cells. Mol. Gen. Genet. 241: 263-270.

Chao, Y. and Pelligrini, M. 1993. In vitro transcription of *Drosophila* rRNA genes shows stimulation by a phorbol ester and serum. Mol. Cell. Biol. 13: 934:944

Chen, K-H., S.G. Widen, S.H. Wilson and K-P Huang. 1990. Characterization of the 5' flanking region of the rat protein kinase γ gene. J. Biol. Chem. 265: 19961-19965.

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

Craig, N.L. 1990. P-element transposition. Cell. 62: 399-402.

Engels, W.R. 1989. P-elements in *Drosophila*. pp. 437-484. in mobile DNA. D.Berg and M. Howe (eds). American Society of Microbiology, Washington D.C.

Engels, W. R., W. K. Benz, C. R. Preston, P. L. Graham, R. W. Phillis and H. M. Robertson. 1987. Somatic effects of P element activity in *Drosophila* melanogaster: Pupal lethality. Genetics 117:745-757.

Engels, W. R., D. M. Johnson-Schlitz, W. B. Eggleston and J. Sved. 1990. High-frequency P element loss in *Drosophila* is homolog-dependent. Cell 62:515-525.

Engels, W. R., C. R. Preston and D. M. Johnson-Schlitz. 1994. Long-range cis preference in DNA homology search extending over the length of a *Drosophila* chromosome. Science 263:1623-1625.

Engels, W.R. 1996. "P Elements in *Drosophila*", p 103-123. In H. Saedler and A. Gierl (Eds.), Transposable Elements. Springer-Verlag, Berlin. (accessed as an HTML document downloaded October, 1996, fromhttp://www.wisc.edu/genetics/CATG/engels/Pelements/ index/html)

Fischer, J.A., and T. Maniatis. 1988. *Drosophila* Adh promoter element expands the tissue specificity of an enhancer. Cell. 53: 451-461.

Flybase. 1996. The *Drosophila* genetic database. Nucl. Acids. Res. 24: 53-56. (Available from the flybase.bio.indiana.edu network server and Gopher site).

Formosa T. and B.M. Alberts. 1986. DNA Synthesis dependent upon genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. Cell. 47: 793-806.

Fridell, Y-W.C., and L.L. Searles. 1992. In vivo transcriptional analysis of the TATA-less promoter of the *Drosophila* melanogaster vermillion gene. Mol. Cell. Biol. 12: 4571-4577.

Garber, M.E., A. Vilalta and D.A. Johnson. 1994. Induction of *Drosophila* RNA polymerase III gene expression by the phorbol ester 12-*O*-tetradecanoyl-13-acetate (TPA) in mediated by the transcription factor IIIB. Mol. Cell. Biol. 14: 339-347.

Geiser, A.G., K.J. Boisam, S-J. Kim, R. Lafyatis, M.A. O'Reilly, R. Webbink, A.B. Roberts and M.B. Sporn. 1993. Regulation of the transforing growth factor β-1 and β-3 by transcription factor Sp1. Science. 129: 223-228.

Glaser, R.L., M.F. Wolfner and J.T. Lis. 1986. Spatial and temporal pattern of hsp26 expression during normal development. EMBO J. 5: 747-754.

Gloor, G. B., N. A. Nassif, D. M. Johnson-Schlitz, C. R. Preston and W. R. Engels. 1991. Targeted gene replacement in *Drosophila* via P element-induced gap repair. Science 253:1110-1117.

Gunaratne, P., J.L. Ross. Q. Zhang, E.L. Organ and D.R. Cavener. 1994. An evolutionary conserved palindrome conserved in the *Drosophila Gld* promoter directs tissue specific expression. Proc. Natl. Acad. Sci., USA 91: 2738-2742.

Hardle, R.C., A. Peretz, E. Suss-Toby, A. Rom-Glas, S.A. Bishop, Z. Sellinger, and B. Mike. 1993. Protein kinase C is required for light adaptation in *Drosophila* photoreceptors. Nature. 634-637.

Hawker, K.L., J.K. Vass and B. W. Ozanne. 1996. Isolation of novel, transcriptionally active Ap1 binding sties: implications for cellular transformation. Oncogene. 13: 283-292.

Hawley, S.A., W.W. Doane and R.A. Norman. 1992. Molecular analysis of *cis*-regulatory sequences at the α-amylase locus in *Drosophila* melanogaster. Biochem. Genet. 30: 257-277.

Hirai, S-I., Y. Izumi, K. Higa, Kaibuchi, K. Mizuno, S-I Osada, K. Suzuki and S. Ohno. 1994. Ras-dependent signal transduction is indispensible by not sufficient for the activation of Ap1/Jun by PKCδ Embo J. 13: 2331-2340.

Holliday, R. 1964. A mechanism for gene conversion in fungi. Genet. res. Camb. 5: 282-304.

Holliday, R. 1968. Genetic recombination in fungi. In: Peacock W.J. Brook R.D. (Eds) Replication and recombiation of genetic material. Australian Academy of Science, Canberra, pp 157-174.

Hu, J., H. Qazza and M.D. Brennan. 1995. A transcriptional role for conserved footprinting sequences within the larval promoter of a *Drosophila* alcohol dehydrogenase gene. J. Mol. Biol. 249: 259-269.

Hughes, S. 1992. Pkc1 in situs in *Drosophila*. MSc. Thesis. Department of Biological Sciences. University of Calgary.

Hultmark, D., R. Klemenz and W. J. Gehring. 1986. Translation and trascriptional control elements in the untranslated leader of the heat-shock gene hsp22. Cell. 44: 429-438.

Johnson, A. 1996. Activation of epidermal growth factor receptor gene transcription by phorbol 12-myristate 13-acetate is mediated by activator protein2. J. of Biol. Chem. 271: 3033-3038.

Johnson-Schlitz, D. M. and W. R. Engels. 1993. P element-induced interallelic gene conversion of insertions and deletions in *Drosophila*. Mol. Cell. Biol. 13:7006-7018.

Karess, R.E. and G.M. Rubin. 1986. Analysis of P transposable element functions on *Drosophila*. Cell. 38: 135-146.

Keeler, K.J., T. Dray, J.E. Penney and G.B. Gloor. 1996. Gene targeting of a plasmid borne sequence to a double strand break in *Drosophila* melanogaster. Mol. Cell. Biol. 16: 522-528.

Kidwell, M.G., J.F. Kidwell and J.A. Sved. 1977. Hybrid dysgenesis in *Drosophila* melanogaster: a syndrome of aberrant traits including mutation, sterility and male recombination. Genetics. 86: 813-833.

Kim, Y-J., H. Pan and A.K. Verma. 1994. Non-Ap1 tumor promoter 12-*0*-tetradecanoylphorbol-13-acetate-responsive sequences in the human ornithine decarboxylase gene. Mol. Carcinogenesis. 10: 169-179.

Kozoza, V.A, E.M. Baricheva, A.V. Katakhin. 1991. Cloning and cytogenetic localization of evolutionary conserved genomic sequences whcih are expressed in the head of *Drosophila melanogaster* imago. Genetik a. 27:51-60.

Lankenau, D-H. 1995. Genetics of genetics in *Drosophila*: P elements serving the study of homologous recombination, gene conversion and targeting. Chromosoma. 103:659-668.

Lankenau, D., V. Corces and W. R. Engels. 1996. Comparison of targeted gene replacement-frequencies in *Drosophila* at the *forked* and *white* loci. Mol. Cel. Biol. 16:3535-3544.

Laski, F. A., D. C. Rio and G. M. Rubin. 1986. Tissue specificity of *Drosophila* P element transposition is regulated at the level of mRNA splicing. Cell 44:7-19.

Lindsley, D.L., and G.G. Zimm. 1992. The genome of *Drosphila melanogaster*. Academic Press, San Diego, Calif.

Ludwig, M.Z., N.A. Tamarina and R.C. Richmond. 1993. Localization of sequence controlling the spatial, temporal and sex specific expression of the esterase 6 locus in *Drosophila* melanogaster adults. Proc. Natl. Acad. Sci., USA. 90: 6233-6237.

Mahajna, J., P. King, P. Parker and J. Haley. 1995. Autoregulation of cloned human protein kinase β and γ gene promoters in U937 cells. DNA and Cell Biol. 14: 213-222.

Martinez, E., C-M. Chiang, H. Ge and R.G. Roeder. 1994. TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. EMBO J. 13: 3115-3126.

Martinez, E., Q. Zhou, N. D. L'Etoile, T. Oelgeschlager, A. J. Berk & R. G.Roeder. 1995. Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding. Proc Natl Acad Sci U S A 92: 11864-11868

Merli, C., D.E. Bergstrom, J.A. Cygan and R.K. Blackman. 1996. Promoter specificity mediates the independent regulation of neighboring genes. Genes and Dev. 16: 1260-1270.

Messelson, M.S. and Radding, C.M. 1975. A general model for genetic recombination. Proc. Natl. Acad. Sci. USA. 78: 358-361.

Meyerowitz, E.M., K.V. Raghovan, P.H. Mathers and M. Roark. 1987. How *Drosophila* larvae makes glue: control of Sgs-3 gene expression. Trends in Genetics. 3: 288-293.

Misra, S., and D.C. Rio. 1990. Cytotype control of *Drosophila* P element transpostition: the 66 kDa protein is a repressor of transposase activity. Cell. 62: 269-284.

Nassif, N. A. and W. R. Engels. 1993. DNA homology requirements for mitotic gap repair in *Drosophila*. Proc. Nat. Acad. Sci. USA 90:1262-1266.

Nassif, N. A., J. Penney, S. Pal, W. R. Engels and G. B. Gloor. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P element-induced gap repair. Mol. Cell. Biol. 14:1613-1625.

Natesan, S. 1990. Protein Kinase C in *Drosophila*. PhD. Thesis. Department of Biological Sciences. University of Calgary.

Niino, Y.S., S. Ohno and K. Suzuki. 1992. Positive and negative regulation of the transcription of the human protein kinase C β gene. J. of Biol. Chem. 267: 6158-6163.

Nishizuka Y. 1988. The molecular heterogeneity of protein kinase and its implications for cellular regulation. Nature. 334:661-665.

Obeid, L.M., G.C. Blobe, L.A. Karalak and Y.A. Hunnan. 1992. Cloning and characterization of the major promoter of the human protein kinase C β gene. J. Biol. Chem. 267: 20804-20810.

O'Hare, K. and G. M. Rubin. 1983. Structure of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell 34:25-35.

Parker, P.J., L. Coussens, N. Tooty, L. Rhu, S. Young, E.Chen, S. Stabel, M.D. Waterfield and A. Ulbrich. 1986. The complete primary structure of protein kinase C--the major phorbol ester receptor Science. 233- 853-859.

Perkins, K.K. g.M. Dailey and R. Tijan. 1988. *In vitro* analysis of the *Antennapedia* P2 promoter: Identification of a new *Drosophila* transcription factor. Genes Dev. 2: 1615-1626.

Quivy, J-P. and P.B. Becker. 1996. The Architecture of the heat-inducible *Drosophila* hsp27 promoter nuclei. J. Mol. Biol. 256: 249-26 3.

Rio, D. C. and G. M. Rubin. 1988. Identification and purification of a *Drosophila* protein that binds to the terminal 31-base-pair repeats of the P transposable element. Proc. Natl. Acad. Sci. USA 85:8929-8933.

Rio, D.C. 1990. Molecular mechanism regulating *Drosophila* P element transposition. Annu. Rev. Genet. 24: 543-578.

Rosenthal, A., L. Rhee, R. Yadegari, R. Paro, A. Ullrich and D.V. Goedel. 1987. Structure and nucleotide sequence of a *Drosophila* melanogaster protein kinase C gene. Embo J. 6: 341-347.

Robertson, H.M., C.T. Preston, R.W. Phillis, D.M. Johnson-Schlitz, W.K. Benz and W.R. Engels. 1988. a stable source of P element transposase in *Drosophila* melanogaster. Genetics. 118: 461-470.

Rubin, G.M. and A.C. Spradling. 1982. Genetic transformation of *Drosophila* with transposable element vectors. Science. 218: 348-353.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Springs Harbour Laboratory Press, Cold Springs Harbour, New York.

Schaeffer, E., D. Smith, G. Mardon, W. Quinn and C. Zuker. 1989. Isolation and characterization of two new *Drosophila* protein kinase C genes, including one specifically expressed in photoreceptor cells. Cell. 57: 403-412.

Spradling, A. C. and G. M. Rubin. 1982. Transposition of cloned P elements into *Drosophila* germ line chromosomes. Science 218:341-347.

Schier, A.F. and W.J. Gehring. 1993. Analysis of a *fushi tarazu* autoregulatory element: multiple sequence elements contribute to enhancer activity. EMBO J. 12: 1111-1119.

Soeller, W.C., S.J. Poole and T. Kornberg. 1988. *In vitro* transcription of the *Drosophila engrailed* gene. Genes Devel. 2: 68-81.

Spradling, A.C., and G.M. Rubin. 1982. Transposition of cloned P elements into *Drosophila* germ line chromosomes. Science. 218: 341-347.

Suzuki D, A.J.V. Griffiths, J.H. Miller, R.C. Lewontin. 1986. An introduction to genetic analysis. Freeman, New York, p. 357.

Sved, J. A., L. M. Blackman, A. S. Gilchrist and W. R. Engels. 1991. High levels of recombination induced by homologous P elements in *Drosophila* melanogaster. Mol. Gen. Genet. 225:443-447.

Szostak, J.W., T.L. Orr-Weaver, R.J. Rothstein, and F.W. Stahl. 1983. The double-strand-break-repair model for recombination. Cell. 33: 25-35.

Takai, Y., A. Kushimoto, Y. Iwasa, Y. Kawakara, T. Mori and Y. Nishizuka. 1977. Calcium-dependent activation of a mutifunctional protein kinase by membrane phospholipids. J. Biol. Chem. 254: 3692-3695.

Takanaga, H., H. Mukai, M. Shimakawa, H. Konishi, U, Kikkawa, T. Kiozumi and Y. Ono. 1995. Functional characterization of the promoter region of the mouse protein kinase C γ gene. FEBS Letters. 368: 276-278.

Thummel, C.S. 1989. The *Drosophila* E74 promoter contains essential sequences downstream from the start site of transcription. Genes Devel. 3: 782-792.

Tseng, G-P., Y-J. Kim, R. Kumarm and A.K. Verma. 1994. Involvement of protein kinase C in the transcriptional regulation of 12-*O*-tetradecanoylphorbol-13-acetate inducible genes modulated by Ap1 or non Ap1 transacting factors. Carcinogenesis. 15:707-711.

Vallet, S.M., M. Brudnak, M. Pellegrini and H.W. Weber. 1993. In vivo regulation of rRNA transcription occurs rapidly in nondividing and dividing *Drosophila* cells in response to a phorbol ester and serum. Mol. Dell. Biol. 13: 928-933.

Yao, K-M. and K. White. 1994. Neural specificity of elav expression: defining a *Drosophila* promoter for directing expression to the nervous system. J. of Neurochemistry. 63: 41-51.