ISOLATION AND CHARACTERIZATION OF NOVEL INSERTIONAL MUTANTS IN THE EARLY REGION 1A OF ADENOVIRUS TYPE 5

By

DIOSDADO S. BAUTISTA, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

For the Degree

Doctor of Philosophy

McMaster University

Hamilton, Ontario, Canada

December 1989

Novel Mutants in Early Region 1A of Adenovirus 5

DOCTOR OF PHILOSOPHY (1989) McMASTER UNIVERSITY

(Biology) Hamilton, Ontario, Canada

TITLE: **Isolation and Characterization of Insertional Mutants in the Early Region 1A of Adenovirus 5**

AUTHOR: **Diosdado S. Bautista, B.Sc.** (McMaster University)

SUPERVISOR: **Frank L. Graham, Ph.D.**

NUMBER OF PAGES: xiv, 241

ABSTRACT

The early region 1A (E1A) of adenovirus 5 codes for oncoproteins whose functions are central to the life cycle of the virus. Its involvement in the initiation of the oncogenic transformation process has been a subject of intense study to understand the mechanism(s) of oncogenesis. To gain a better understanding of the structures of E1A proteins and their relationship to their various functions such as gene regulation and oncogenic transformation, a genetic study was undertaken. This involved mutagenesis of E1A by introducing a defined mutator DNA duplex into various sites in an E1A-containing plasmid specified by several multi-cut restriction enzymes. To facilitate the isolation and genotypic characterization of such mutants, a general method for introducing oligomers in the form of a cassette into multi-copy plasmids was developed. This system had several attractive features. First, the mutator cassette contained the bacterial *lac* operator sequence as a phenotypic label permitting easy identification of clones with inserts. Second, the cassette was designed for sufficient flexibility to introduce any desired sequence. Third, it allowed fast and efficient sequencing of the plasmid with an insert to determine its orientation as well as the structure around the insert.

To mutate E1A, a 39-bp cassette was used that was capable of coding for 13aa residues in one orientation but was designed to terminate translation in the

other orientation due to stop codons in each of the reading frames. It contained flanking *Bam*HI sites that allowed 'collapsing' the insert to result in a net insert of two residues. A total of 18 sites in E1A have been mutated, each containing these three types of mutations. To determine the phenotypes of such mutants with respect to E1A's transcriptional regulatory functions (*trans*-activation and *trans*-repression), reporter genes with the bacterial ß-galactosidase gene under the control of the adenovirus E1A, E1B and E3 promoters were constructed and used in co-transfection studies with the mutants in HeLa cells. The *trans*-activation assays showed that the unique region of the 13S E1A product was the only important domain for the activation of the E1A, E1B and E3 promoters as indicated by the finding that only insertions (both long and short) in this region abolished or diminished the expression of the reporter plasmids. However, while all long insertions in the unique region affected function, the shorter 2aa residue insertions had variable effects since one such mutant did not affect function, indicating that not only size but also the sequence of the insert could affect the *trans*-activation function.

The *trans*-repression assays were done using similar mutants modified so that only the 12S E1A product was expressed. This was necessary to prevent the interference in the assay by the *trans*-activation function found in the larger protein. The results showed that two regions flanking the unique region were sensitive to insertion mutations with respect to this function. Again perturbation of the sequence by size alone was not sufficient to destroy protein function but more important was the sequence of the insert. These regions are positioned adjacent to each other in the 12S product and most likely form a single domain for the *trans*-repression function.

ACKNOWLEDGEMENTS

My deepest gratitude to Dr. Frank L. Graham whose supervision of this project was in itself a lesson to learn: strict adherence to the main goal and ample flexibility and independence to explore other things of interest. I am most appreciative of his ability to create an atmosphere in the laboratory most condusive to scientific research.

I would like to thank the members of the McMaster Cancer Research Group especially those in the Graham lab who have made doing research most enjoyable and memorable: the students Gary Wilson, Frank Lee, Joel McGrory, Roger Lippe and Yousef Haj-Ahmad; Pamela Brinkley for finding ways to make times livelier; John Rudy for making things in the lab organized; and the summer students Christine Mielke, Nadia Pece, and Paula Stanford who were most helpful in the isolation and maintenance of mutants described herein.

I am indebted to the Government of Ontario for an OGS Scholarship (1983-84), to McMaster University for the Harry Lyman Hooker Scholarship (1984-88), to the Department of Biology for the sustained teaching assistantship, and most especially to the National Cancer Institute of Canada for the Steve Fonyo Studentship (1984-88).

My sincerest thanks to my parents, Dr. and Mrs. Samuel F. Bautista, for being most supportive of my ambition and aspiration, and equally to the members of my family especially Marc for his computer and printing this thesis and Meynard for helping me with the photographic works.

"For the Lord giveth wisdom: out of His mouth cometh knowledge and understanding." *Proverbs 2:6*

*To my wife **Filomena** and our son **Julian Francis***

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µg	microgram
µl	microliter
^{32}P	phosphorus-32
^{35}S	sulphur-35
A_{420}	absorbancy at wavelength 420nm
Ad	adenovirus
amp	ampicillin
ATP	adenosine triphosphate
BIS	*N,N'*-methylene-bis-acrylamide
bp	base pairs
BRL	Bethesda Research Laboratories
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
CAT	chloramphenicol acetyl transferase
cm	centimeter
CsCl	cesium chloride
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
ddATP	dideoxyadenosine triphosphate
ddCTP	dideoxycytosine triphosphate
ddGTP	dideoxyguanosine triphosphate
ddNTP	dideoxynucleotide triphosphate
ddTTP	dideoxythymidine triphosphate
dGTP	deoxyguanosine triphosphate
DMF	dimethylformamide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
E1A	early region 1A
E1B	early region 1B
E2A(e)	early region 2A
E3	early region 3
E4	early region 4
et al.	and co-workers
EtBr	ethidum bromide
EtOH	ethanol
FT buffer	freeze/thaw buffer
h	hour(s)

H_2O	water
HAT	hypoxanthine, aminopterin, thymidine
HeBS	HEPES-buffered saline
KCl	potassium chloride
kDa	kilodaltons
KH_2PO_4	potassium dihydrogen monophosphate
LB	Luria-Bertani
M	molar
MEM	minimal essential medium
$MgSO_4$	magnesium sulfate
min	minute(s)
ml	milliliter
MLP	major late promoter
mm	millimeter
mM	millimolar
mmole	millimole
mu	map units
MW	molecular weight
Na_2CO_3	sodium bicarbonate
Na_2EDTA	disodium ethylenediamine tetracetic acid
Na_2HPO_4	disodium hydrogen monophosphate
Na_2HPO_4	disodium hydrogen monophosphate
NaCl	sodium chloride
NaH_2PO_4	sodium dihydrogen monophosphate
NaOH	sodium hydroxide
NEB	New England Biolabs
NEN	New England Nuclear
ng	nanogram
NH_4Cl	ammonium chloride
nm	nanometer
OD_{nn}	optical density at wavelength nn
ONPG	o-nitrophenyl-ß-D-galactopyranoside
PAGE	polyacrylamide gel electrophoresis
PBS^{2-}	phosphate-buffered saline
pmole	picomole
PNK	polynucleotide kinase
SDS	sodium dodecyl sulfate
sec	second(s)
SSC	saline sodium citrate
TAE	Tris.acetate EDTA
TBE	Tris.borate EDTA
TE	Tris.EDTA
TEMED	*N,N,N',N'*-tetramethylethylenediamide
tk	thymidine kinase
Tris	tris(hydroxymethyl) aminomethane
v/v	volume per volume
V	volts
w/v	weight per volume
Xgal	5-bromo-4-chloro-3-indoyl-ß-D-galactopyranoside
Z buffer	*lac*-Z buffer

CHAPTER I

INTRODUCTION

A. ADENOVIRUS TUMOUR VIROLOGY

The emergence of adenovirus as one of the major tools in elucidating the basic mechanisms of tumour induction can be traced to the late 1950s and early 1960s when intensive research was being carried out to determine the efficacy of adenovirus vaccines. The development of such vaccines was sparked by a remarkably widespread incidence of adenovirus infection, especially among U.S. military recruits, by adenovirus serotypes 4 and 7, which inflicted symptoms characteristic of acute respiratory disease with resultant ineffectiveness among personnel (reviewed in Flint, 1981). Formalin-killed adenovirus vaccines from infected cultures of monkey kidney epithelium were developed (Hilleman *et al.*, 1956) and subsequently shown to be effective in preventing respiratory disease among British military recruits (Wilson *et al.*, 1960). However, the use of the vaccines in the general population, where it was believed that adenoviruses were extensively involved in respiratory disease, was never fully implemented owing at least partly to the observation that human Ad12 (Trentin *et al.*, 1962) and Ad18 (Huebner *et al.*, 1962) could induce malignant tumours when inoculated into newborn hamsters. These were the first demonstrations that a family of viruses of human origin was potentially oncogenic, raising speculations that at least some forms of human cancer might be caused by a viral agent. More than a decade later, however, extensive surveys of different human malignancies using DNA

hybridization techniques have failed to detect any evidence for the presence of adenovirus sequences (Mackey *et al.*, 1976; Green *et al.*, 1979).

The initial demonstration that at least some common human adenoviruses possess oncogenic capability shifted the emphasis of adenovirus research from an agent of sporadic outbreaks of acute respiratory diseases, which were generally not life-threatening, to an important model system for studying tumour induction at the molecular level. Of course, adenoviruses were not the first viruses discovered with oncogenic property. Long before adenoviruses were first discovered by Rowe *et al.* (1953) as the "adenoid degenerating" agent, Peyton Rous discovered in 1911 the first tumour virus, now called the Rous sarcoma virus, and established it as the etiological cause of a spontaneous sarcoma in chicken (reviewed in Weiss *et al.*, 1982), and papilloma viruses were already believed in the 1920s and 1930s to be the causative agents of benign growths (warts) in many species of mammals including man (reviewed in Grodzicker and Hopkins, 1981). However, it was in the late 1950s and early 1960s when various important discoveries (see Grodzicker and Hopkins for details) implicating many other viruses in tumour induction under laboratory conditions were made. To name a few, Stewart *et al.* (1958) showed that polyoma virus (then named as the parotid agent) induced various neoplasms when inoculated in mice and transformation of mouse and hamster cells in culture was demonstrated (Sachs and Winocour, 1959; Vogt and Dulbecco, 1960); SV40, the contaminant in many of the early batches of polio vaccine, was shown to induce tumours when injected into newborn hamsters (Eddy *et al.*, 1962; Girardi *et al.*, 1962); a poxvirus, the Yaba monkey virus, produced subcutaneous tumours in rhesus monkeys when injected (Niven *et al.*, 1961); and in 1964, Epstein *et al.* reported that a herpesvirus called Epstein-Barr virus was found in a line of Burkitt's lymphoma cells maintained in culture (Burkitt's lymphoma is a geographically restricted disease which commonly

afflicts children in regions of East Africa). These and similar discoveries initiated an era in virology that is characterized partly by keen interest in elucidating the etiological role of viruses in human cancer, and especially by an interest in elucidating the mechanism by which tumour viruses induce tumours and transform cells. Until now that interest has not abated. For adenoviruses, it was found that all serotypes have the capacity to transform non-established cells *in vitro*, although only members of certain groups (A, B and D) have been shown to induce tumours when inoculated into experimental animals (reviewed by Graham, 1984). Although it is now widely believed that adenoviruses do not play a significant role in human malignancies (reviewed in Branton *et al.*, 1985), their oncogenic properties continue to be the object of intensive research at the molecular level pertaining to mechanisms of tumour induction. By clearly defining the genes that are involved in the transformation process, and by understanding in full the biological and biochemical properties of the product of such genes, it may be possible to identify a common theme that could account for the origins of a wide variety of malignancies that are collectively called cancer.

The interest in adenoviruses is certainly not confined to the study of transformation mechanisms. It has become clear that a full understanding of transformation itself would require the elucidation of some more fundamental biological phenomena associated with cellular processes, most important of which is gene regulation. In this respect, certain adenovirus serotypes have served as useful tools in uncovering complex mammalian cell processes in specific areas such as mRNA metabolism, DNA replication, and temporal control of gene expression.

B. ONCOGENES OF ADENOVIRUS

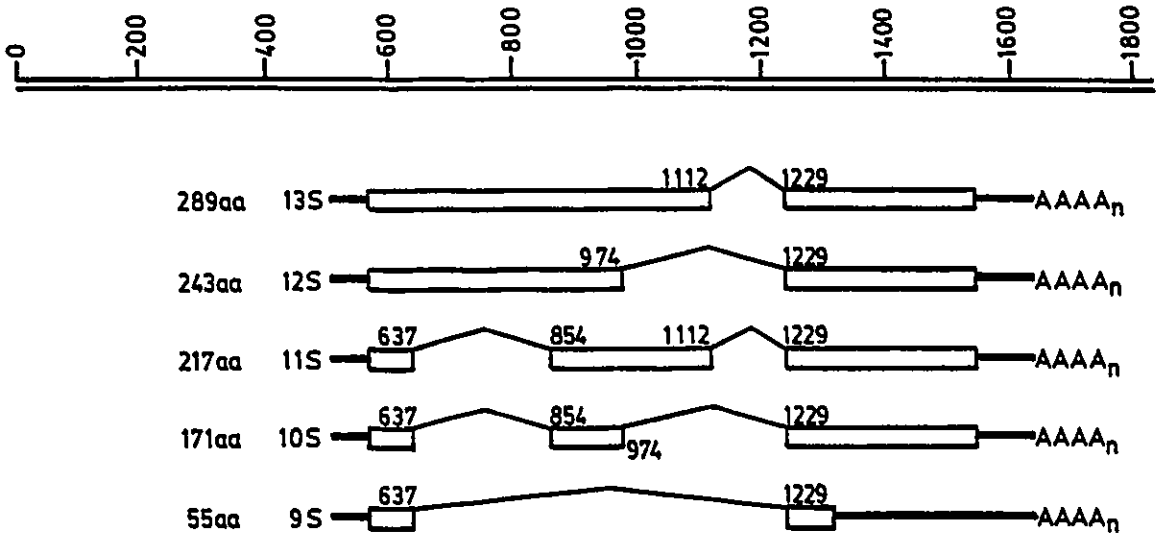
1. Early Region 1A

The E1A region of adenovirus encodes at least 5 mRNAs (**Figure 1-1**) that are derived from a common precursor RNA through differential splicing. Three of these mRNAs, which are referred to according to their sedimentation values of 13S, 12S and 9S species, were the first to have been characterized (Berk and Sharp, 1978; Chow *et al.*, 1979a; Perricaudet *et al.*, 1979). Both 13S and 12S mRNAs are synthesized throughout the adenovirus lytic cycle but the 9S mRNA is expressed preferentially during the late phase of infection (Chow *et al.*, 1979a; Esche *et al.*, 1980; Spector *et al.*, 1980). Two other E1A-encoded mRNA species have recently been detected, 11S and 12S, both being made preferentially at late times in infection by the highly related Ad5 (Stephens and Harlow, 1987) and Ad2 (Ulfendahl *et al.*, 1987) serotypes. These mRNA species utilize the splice junction sequences at nucleotides 637 and 854, so that E1A mRNAs of Ad5 reveal three potential splice donor sites (nucleotides 637, 974 and 1112) and two acceptor sites (nucleotides 854 and 1229). Utilization of the splicing sites appears to be phase-regulated, as judged from the appearance of the mRNA species. That is, while the splice donors at nt974 and nt1112 are used at both early and late phases of infection, the donor at nt637 is used at late times only. Similarly, the splice acceptor at nt1229 is used at all times while that at nt854 is only used again only at late times. Except for the 9S mRNA, all E1A mRNAs are encoded in the same translational reading frame so that their translational products are highly related and differ only in regions where splicing has occurred. For example, the 13S and 12S protein products differ only by 46aa residues, in what is referred to as the unique region. The 11S and 12S products differ from the 13S and 12S products, respectively, by a common 72aa residue-long sequence. All 4 mRNA products have a common 104aa residue-long sequence at the

Figure 1-1. **Early Region 1A of Adenovirus Type 5**

All E1A mRNAs are derived from a single precursor transcript through differential splicing. The cap site is at nt499 and the polyadenylation site is at nt1632. The 5 mRNAs (13S, 12S, 11S, 10S and 9S) produce polypeptides that are 289aa, 243aa, 217aa, 171aa and 55aa residues, respectively, illustrated as blank blocks. The numbers above the blank blocks refer to the splice donor/acceptor sites.

Figure 1-1. **Early Region 1A of Adenovirus Type 5**



COOH-terminus. For the 9S product, only the first 26aa residues are in common to the other products since the nt637/nt1229 splice which creates the 9S mRNA results in a shift in the translational reading frame that produces a new sequence and a translational termination to give a protein product that is predicted to be only 55aa residue-long. This product, if made at all, has no known function (Moran *et al.*, 1986b).

Translation of the various E1A mRNAs would lead to at least 5 different proteins. In fact, the E1A encoded products appear even more heterogeneous on SDS-PAGE (reviewed by Branton *et al.*, 1985). E1A proteins also migrate anomalously resulting in lower than predicted mobilities. The early E1A proteins undergo extensive post-translational modification, mainly by phosphorylation, that in part accounts for the observed heterogeneity of E1A species in gel electrophoresis (Yee and Branton, 1985a; Tsukamoto *et al.*, 1986; Stephens *et al.*, 1986).

In virus-infected cells, E1A is primarily involved in the activation of several early promoters that otherwise have low basal levels of activity (reviewed in Berk, 1986a). This function of E1A is required for productive infection. However, deletion of E1A sequences, as in the mutant *dl*312 (Shenk *et al.*, 1979), results in a much delayed onset of early events and the subsequent appearance of viral progenies rather than a total block in viral replication (Nevins, 1981; Gaynor and Berk, 1983), demonstrating that the requirement for E1A is not absolute. These and other studies (reviewed in Sharp, 1984) demonstrate that a) E1A activities are mainly concerned with facilitating the expression of other early genes whose adequate level of expression is required to ensure the completion of the viral cycle, and b) the basal level of expression of E1A-dependent viral genes is inadequate except at high multiplicities of infection or at late times for replication of viruses lacking E1A, presumably through basal expression levels which eventually become sufficient to

Table 1-1. Partial List of Adenovirus-Responsive Transcription Units

Transcription Unit	References
Ad E1A	Borrelli *et al.* (1984); Smith *et al.* (1985)
Ad E1B	Rossini (1983)
Ad E2A(early)	Rossini (1983); Guilfoyle *et al.* (1985)
Ad E2A(late)	Rossini (1983)
Ad VA-I; VA-II	Berger and Folk (1985); Gaynor *et al.* (1985); Hoeffler and Roeder (1985)
Ad MLP	Lewis *et al.*, (1985); Nevins (1981)
Ad IVa_2	Natarajan and Salzman (1985)
epsilon-globin	Allan *et al.* (1984)
human beta-globin	Green *et al.* (1983)
rabbit beta-globin	Svensson and Akusjarvi (1984)
rat pre-proinsulin I	Gaynor *et al.* (1984)
HTLV I & II LTR	Chen *et al.* (1985)
HSV-1 gD	Everett and Dunlop (1984)
SV40 early	Borrelli *et al.* (1984); Velcich and Ziff (1985)
hsp70	Nevins (1981)
hsp 89-alpha	Simon *et al.* (1987)
c-*fos*, c-*myc*	Sassone-Corsi and Borrelli (1987)
insulin	Stein and Ziff (1987)
immunoglobulin heavy chain	Hen *et al.* (1985)
interferon	Toth *et al.* (1987)
polyoma early/late genes	Velcich *et al.* (1986)
rabbit beta-globin	Everett and Dunlop (1984)
Drosophila tRNA	Gaynor *et al.* (1985)
topoisomerase	Chow and Pearson (1985)
MHC	Vasavada *et al.* (1985)
thymidylate synthase	Zerler *et al.* (1987)
HeLa beta-tubulin	Stein and Ziff (1984)
immunoglobulin kappa-chain	Bergman and Shavit (1988)

release from E1A-dependence.

The transcriptional regulatory functions of E1A are well-documented and fall under 2 classes: 1) transcriptional activation of promoters, viral and cellular, an activity referred to as *trans*-activation (reviewed in Berk, 1986a), and 2) *trans*-repression, especially of promoters linked to enhancers (Borrelli *et al.*, 1984; Velcich and Ziff, 1985; Hen *et al.*, 1985). **Table 1-1** is a partial list of documented E1A-responsive genes, both viral and cellular in origin. It has been reported that both 289R and 243R E1A proteins appear to have *trans*-activating function (reviewed by Berk, 1986a), but it is generally believed that the 289R protein is the major *trans*-activator and the 243R protein is primarily responsible for repression. If indeed the smaller E1A protein can act as a *trans*-activator (see, for example, Leff *et al.*, 1985; Montell *et al.*, 1984; Winberg and Shenk, 1984), it would probably occur through

some mechanism(s) that differ both quantitatively and qualitatively from that of the larger protein since the 243R protein lacks the unique region domain to which the *trans*-activation function of the 289R protein is ascribed. Studies using the bacterially-produced 243R protein microinjected into Vero cells (Ferguson *et al.*, 1985) or *Xenopus* oocytes (Richter *et al.*, 1986) also demonstrate the *trans*-activating potential of the smaller protein, although they perhaps suggest that high levels of the protein attained by microinjection are needed to see the effect. E1A *trans*-activation presently appears to occur via multiple mechanisms all of which are related in that common cellular factors are involved. A discussion of these mechanisms is found in a later section.

One important characteristic of E1A regulatory functions is that they are not limited to the control of viral transcriptional units since cellular and other elements from other viral systems respond as well. Conversely, E1A regulation, particularly *trans*-activation, is not universal since not all promoters tested respond to its effects (reviewed in Berk, 1986a). In this respect, the E1A products may be classified as true modulators of gene expression in contrast to some ubiquitous factors that make up the core transcriptional machinery needed by all promoters.

The demonstration that E1A is intimately involved in the transforming capabilities of adenovirus (Graham *et al.*, 1974; Gallimore *et al.*, 1974; reviewed in Graham, 1984 and in Branton *et al.*, 1985) and that it is the primary regulator of viral gene expression during infection (reviewed in Sharp, 1984) have been the main impetus for the intensive search of a specific causal relationship, if any, between control of gene expression and oncogenic transformation. Transformation by adenovirus requires the action of both E1A and E1B (reviewed in Graham, 1984; Branton *et al.*, 1985). In the absence of E1B, E1A induces incomplete transformation (referred to as immortalization) of primary cells in culture (Houweling *et al.*, 1980;

van den Elsen *et al.*, 1983c). Interestingly, either E1A or E1B may be substituted in these transformation assays by other oncogenes, suggesting that induction of transformation by other genes may occur through similar mechanisms. For example, an activated H-*ras* oncogene or the polyoma middle T antigen can cooperate with E1A for the full transformation of primary cells in culture (Ruley, 1983). By the same token, E1B may cooperate with any of the following E1A substitutes: polyoma large T, *myc* oncogene, or the p53 oncogene (Land *et al.*, 1983; Ruley, 1983; Eliyaju *et al.*, 1984; Jenkins *et al.*, 1984; Parada *et al.*, 1984; Schwab *et al.*, 1985; Yancopoulos *et al.*, 1985).

Alignment of the primary sequence (Kimmelman *et al.*, 1985) of the large E1A proteins from various adenovirus serotypes (Ad5, Ad7, and Ad12) has been used to identify regions of homology that are called conserved regions (reviewed in Moran and Mathews, 1987). **Figure 1-2** is an updated alignment of E1A proteins from several human adenoviruses (Ad2, Ad4, Ad5, Ad7, Ad12, and the enteric adenoviruses Ad40 and Ad41), simian adenovirus 7 (SA7), the tree shrew adenovirus (AdTS), and mouse adenovirus 1 (MAV1, Ball *et al.*, 1988). It can be seen from this and other alignments (Kimmelman *et al.*, 1985, Figge *et al.*, 1988) that regions with high homology alternate with regions of nonhomology. Various mutational studies have suggested that most activities of E1A may be assigned to one or more of these conserved regions that have sometimes been called functional domains (reviewed in Moran and Mathews, 1987). The region between residues 40 and 78 (numbering is according to the Ad5 13S product) is highly conserved and is often called **region 1.** Only the two largest products contain this region since it is removed by the 637/854 splice in both 11S and 10S products. Another conserved region (**region 2**) is found between residues 108 and 139, and is present in all four largest E1A mRNAs. Both regions 1 and 2 are important in the immortalization of primary cells (Lillie *et al.*,

Figure 1-2. **Alignment of E1A Proteins from Various Adenovirus Species**

The primary structure of largest E1A protein from various adenovirus species was deduced based on information found in GenBank (release no. 58) except for MAV1 (Ball *et al.*, 1988). The alignment was generated using the GAP and PRETTY programs of University of Wisconsin GCG Sequence Analysis Software Package (Version 5.3).

Figure 1-2. Alignment of E1A Proteins from Various Adenovirus Species

```
                                                      40   REGION 1
Ad2        MRHII C. HGGVI TEEMAASLLDQLIEEVLADNLP. PPSHFEPPTLHELYDLDVT 52
Ad4        MRHLR DLPDEEI I I ASGSEI LELVVNATMGDDHPEPPTPFGTPSLHDLYDLEVD
Ad5        MRHII C. HGGVI TEEMAASLLDQLIEEVLADNLP. PPSHFEPPTLHELYDLDVT
Ad7        MRHLR FLPQEI I SSETGI EI LEFVVNTLMGDDPEPPVQPFDPPTLHDLYDLEVD
Ad12       MRTEMT. . PLVLSYQEADDI LEHLVDNFF. . NEVPSDDDLYVPSLYELYDLDVE
Ad40       MRMLP DFFTGNWDD. . . . . MFQGLLETEYVFDFPEPSEASEEMSLHDLFDVEVD
Ad41       MRMLP DFFTGNWDD. . . . . MFQGLLEAEHPFDFPEPSQAFEEI SLHNLFDVELD
AdTS       MRN. . . . . . . WELSVSPSFLELCDQYV. . . SLCESPSFSGPSCLNDMAPDDI !.
SA7        MRHLL . . . ALEMI SELLDLGLDTI DGWLHTEFRPVPAGVSHNMSLHEMYDLDV.
MAV1       MSRLLR. LSLSSRVWLAA. . . QEATRNVSEDPVV. QRTPWDGSPTCTAVRVVRA
Consensus  MRHL- - - - - - - - - - - - - - - LE- LVE- - - D- LP- P- - - FE- PSLH- LYDLDVD

              REGION 1                80
Ad2        APE. DPNEEAVSQI FPDSVMLAVQEGI DLLTFPPAPGSPEPPHLSRQPEQPEQR 105
Ad4        VPEDDPNEKAVNDLFSDAALLAAEEA. . . . . . . . . . . . SSPSSDSDSSLHTPRH
Ad5        APE. DPNEEAVSQI FPDSVMLAVQEGI DLLTFPPAPGSPEPPHLSRQPEQPEQR
Ad7        GPE. DPNEGAVNGFFTDSMLLAADEGLDI . . . . . . . . NPPPETLVTPGVVVES
Ad12       SAGEDNNEQAVNEFFPESLI LAASEGLFL . . . . . . . . PEPPVLSPVCE. . . . . .
Ad40       GFEEDANQEAVDGMFPERL. . . . . . . . . . . . . . . . . . . . . . LSEAESAAESG
Ad41       ESEGDPNEEAVDGMFPNWMLSE. . . . . . . . . . . . . . . . . . . . DHSADSGAA
AdTS       G. NCDLFAEAADALFPDCLLEEVEAASGL. . . . . . . . . . . . . . . . . . . .
SA7        GQE. DENEEAVDGVFSDAMLLAAEEGI EMPNLYSPG. . . . . . . . . . . . . . . P
MAV1       EVL. ADGTMDLDI VFPEA- - - AVQAVFSRTPWQDSTTATSAEEPSASTDSI SSD
Consensus  - - E- DPNEEAV- - - FPD- - LLA- - EG- - L- - - - - - - - - - - P- - - S- - - - - - - -

Ad2        ALGPVSMPNLVPEVI DLTCHEAGFPPSDDED. . . . . . EE. G. . . . . . . . . . . . . 139
Ad4        DRGEKEI PGLKWEKMDLRCYEECLPPSDDEDEQAI QNAASHG. . . . . . . . . . . . V
Ad5        ALGPVSMPNLVPEVI DLTCHEAGFPPSDDED. . . . . . EE. G. . . . . . . . . . . . .
Ad7        GRGGKKLPDLGAAEMDLRCYEEGFPPSDDEDGETEQSI HTAV. . . . . . . . . NEGV
Ad12       PI GGECMPQLHPEDMDLLCYEMGFPCSDSED. . . . EQDENGMAHVSASAAAAAA
Ad40       SGDSGVGEELLPVDLDLKCYEDGLPPSDPET. . DEATEAEE. . . . . EAAMPTYV
Ad41       SGDSGVGEDLVEVNLDLKCYEEGLPPSGSEA. . . DEAEERAEEEETAVSNYVNI
AdTS       . . . . . . . . . . . . . . . . . . . . . . AFETNEEVE. . . GF. . . . . . . . . . . . . . . .
SA7        LVGGGEMPELQPEEEDLFCYEDGFPPSDSEEGEHSQVETERKMAEAAAAGAAAAA
MAV1       PLPI SCVESF. . EDMDLRCYEQ. LSPSPESI . . . . . . ET. I . . . . . . . . . . . . .
Consensus  - - G- - - MP- L- PE- MDL- CYEEGFPPSDDED- - - - - - - - - - - - - - - - - - -
             108      REGION 2              136

             140     REGION 3 - UNIQUE REGION          184
Ad2        . . . . EEFVLDYVEHPGHGCRSCHYHRRNTGDPDI MCSLCYMRTCGMFVYSPVSE 189
Ad4        QAVSESFALDCPPLPGHGCKSCEFHRI NTGDKAVLCALCYMRAYNHCVYSPVSD
Ad5        . . . . EEFVLDYVEHPGHGCRSCHYHRRNTGDPDI MCSLCYMRTCGMFVYSPVSE
Ad7        KAASDVFKLDCPELPGHGCKSCEFHRNNTGMKELLCSLCYMRMHCHFI YSPVSD
Ad12       DRERE EFQLDHPELPGHNCKSCEHHRNSTGNTDLMCSLCYLRAYNMFI YSPVSD
Ad40       NENENELVLDCPENPGRGCRACDFHRGTSGNPEAMCALCYMRLTGHCI YSPI SD
Ad41       AEGASQLVLDCPENPGRGCRACDFHRGSSGNPEAMCALCYMRLTGHCI YSPI SD
AdTS       . . . . . . VFPDCPERPGQECRSCKQHREMSGDPSI LCSLCYMRLTACFVYSPVSD
SA7        RREQDDFRLDCPSVPGHGCSSCDYHRKTSGCPEI LCSLCYLRANSMFI YSPVSD
MAV1       . . . . EVF. . . . . . . . PPCSTCGGH. . . . . EVNGFCSLCYLRGLTGKVF. . . . .
Consensus  - - - - EEF- LDCPE- PGHGCRSC- FHR- - TG- PEI MCSLCYMR- - - - FI YSPVSD

Ad2        PEPEP EPEPEPARPTRRPKLVPAI LRRPTSPVSRECNSSTDSCDSGPSNTPPE  242
Ad4        ADDETPT. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . TESTLSPPE
Ad5        PEPEP EPEPEPARPTRRPKMAPAI LRRPTSPVSRECNSSTDSCDSGPSNTPPE
Ad7        DE. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . SPSPDSTTSPPE
Ad12       NEPEP NSTLDGDERPSPPKLGSA. . . . . . . . . . . . . . . . . . . . . . . .
Ad40       AEGES E. . . . . . . . . . . . . . . . . . . . . . . . . . . . . SGSPEDTDPPHP
Ad41       AEGE . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . CELGSNEETEL
AdTS       VEDE . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
SA7        SEPD . . . . . . . . . . . . . . . . . . . . . EPDSTTADSNHGS. . . PPT
MAV1       . . . . . . . . . . . . . . . . . . LLRM . . . . . . . . . . . . . . . . . . .
Consensus  - E- E - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - PP-

Ad2        I HPVVPLCPI KPVAVRV. GGRRQAVECI EDLLNESGQ. . . . PLDLSC. KRPRP  289
Ad4        I GTSPSDNI VRPVPVRA. TGRRAAVECLDDLLQGGDE. . . . PLDLCTRKRPRH
Ad5        I HPVVPLCPI KPVAVRV. GGRRQAVECI EDLLNEPGQ. . . . PLDLSC. KRPRP
Ad7        I QAPAPANVCKPI PVKPKPGKRPAVDKLEDLLEGGDG. . . . PLDLSTRKLPRQ
Ad12       . . . . VPEGVI KPVPQRVTGRRRCAVESI LDLI QEEEREQTVPVDLSVKRPRCN
Ad40       LTATPPHGI VRTI PCRVSCRRRPAVECI EDLLEE. DPTDEPLNLSLKRPKCS
Ad41       PCSLTATAPVRPTPCRVSCRRRPAVDCI EDLLEE. DPTDEPLNLSLKRPKSS
AdTS       . . . . . . . . . . . . . . . . . . . . . EPTEGVAENSLKRQADSSLCSSSPKRFC
SA7        LRCTPPRDLPRPVPVKASPGKRPAVNSLHDLI EEVEQ. . TVPLDLSLKRSRSN
MAV1       . . . . . . . . . . . . . . . . . . . . . MSFYGSLFVSNSP. . . . DLFFRS. AASGG
Consensus  - - - - - P- - - - - PVPVRV- - GRR- AVECI EDLLEE- - - - - - PLDLS- - - - - -
```

1986), in transformation in cooperation with *ras* (Moran *et al.*, 1986a), and in induction of cellular DNA synthesis (Lillie *et al.*, 1987). Region 2 shows similarity tonuclear products of unrelated viral oncogenes of the family of papovaviruses (Stabel *et al.*, 1985; Phelps *et al.*, 1988). At least two reports indicate that the region of similarity between E1A and that encoding the SV40 large T antigen is involved in similar functions. Kalderon and Smith (1984) showed that such region is within the portion of the SV40 large T antigen region that is required for transformation. More recently, Moran (1988) showed that a chimeric E1A protein in which the E1A region is replaced by a portion of the SV40 product containing the region 2 similarity was able to transform with efficiency comparable to wild-type. The unique region (**region 3**) is found only in the 13S and 11S products, both of which are able to *trans*-activate the Ad5 E3 promoter in co-transfection experiments using plasmid DNA, but the 11S product functions only inefficiently as a *trans*-activator in lytic infections as shown by using a mutant virus expressing 11S alone (Ulfendahl *et al.*, 1987).

The alignment of E1A proteins in **Figure 1-2** reveals other conserved regions which are relatively shorter than those described above. Three such regions are found before **region 1**, the first being made up of amino acid residues 1 to 4. This is followed by a region between residues 19 and 26 and a region between residues 29 and 35. **Region 1** should be extended to include residues 38 to 80 inclusive, and **region 2** should include residues 108 to 136 only instead of up to residue 139. The region of nonhomology between 2 and 3 varies extensively among the large E1A proteins. **Region 3**, which is the unique region, is the longest stretch of homology in the E1A proteins and also with the highest degree of homology.

Extensive homologies are also found in exon 2. In fact, immediately downstream from the unique region is a stretch of 9 residues with high degree of homology. These are probably important to both the 13S and 12S products since in

either case these residues are closely located to the unique region of the 13S product and conserved region 2 in the 12S product. Another region with considerable homology is located between residues 253 and 276, and still another between 279 and 284. To what extent these relatively short conserved regions contribute to E1A functions is not known.

The conserved region 3 (unique region) is rich in cysteine residues that are conserved in all adenovirus species (**Figure 1-2**). The structure formed by the stretch of residues that include the 4 cysteines in the unique region of Ad5 and Ad2 (C-X_2-C-X_{13}-C-X_2-C) might resemble the "zinc finger" structural motif (see reviews by Klug and Rhodes, 1987 and Evans and Hollenberg, 1987) which was originally found in the DNA-binding *Xenopus* transcription factor IIIA (Miller *et al.*, 1985; Brown *et al.*, 1985). Culp *et al.* (1988) recently demonstrated using bacterially produced 289aa E1A protein that the consensus zinc finger motif in the unique region does bind zinc in a 1:1 molar ratio. Substitution of any of the cysteine residues (at positions 154, 157, 171 and 174) with glycine resulted in abrogation of the *trans*-activation activity but the effect on zinc-binding was not shown (Culp *et al.*, 1988). However, it was shown that zinc-binding is not sufficient for *trans*-activation since substitution of cys_{154} for a histidine residue, a mutation that abolished *trans*-activation function, did not decrease zinc-binding (Culp *et al.*, 1988). There is little doubt that the zinc finger motif contributes significantly to the *trans*-activation function but it appears that the original role attributed to zinc fingers to mediate specific protein-DNA binding may not be applicable to the E1A protein since its DNA-binding activity is at best nonspecific (Ferguson *et al.*, 1985). However, because of the role played by these proteins as transcriptional regulators, it is possible that by forming complexes with other cellular factors, the potential of the putative zinc fingers to bind DNA might be realized. Alternatively, the zinc finger might be a structure involved in forming a

stable complex with a factor (Culp *et al.*, 1988).

The E1A proteins have been produced in bacteria (Ferguson *et al.*, 1985) and appear to retain their biological properties as activators of transcription. Proteins denatured by boiling quickly regain their activity (Ferguson *et al.*, 1985) most probably by virtue of the proline residues that comprise 16% (for 289R protein) to 18% (for 243R) of the proteins. Proline residues provide a rigid structure to the protein backbone, hence the E1A proteins may be fairly rigid, which explains in part why denaturation by boiling appears to have little effect on activity (at least as assayed by *trans*-activation).

Mutational analysis of E1A has shown that certain E1A activities are distinct from others, and can more or less be assigned to different conserved regions in the protein (reviewed in Moran and Mathews, 1987). Of the various functional domain assignments, perhaps what is least contested is the identification of the unique region (conserved region 3) as the main *trans*-activation domain in the larger protein (Carlock and Jones, 1981; Montell *et al.*, 1982; Moran *et al.*, 1986a; Lillie *et al.*, 1986 and 1987; Jelsma *et al.*, 1988). Perhaps most convincing is the demonstration that a 49aa residue synthetic peptide encoding mainly conserved region 3 can act autonomously as a *trans*-activator of the E2A promoter when microinjected into HeLa cells (Green *et al.*, 1988).

A relatively less understood E1A function is *trans*-repression of promoters, especially those linked to transcriptional enhancers (Borrelli *et al.*, 1984; Velcich and Ziff, 1985; Hen *et al.*, 1985). Some enhancers, however, are not repressed but are activated by E1A, as in the case of the adenovirus E2A enhancer (Imperiale *et al.*, 1985; Jalinot and Kedinger, 1986).

Attempts to correlate E1A regulatory and transforming functions have led to intriguing conclusions. First, it is clear from transformation assays using activated

H-*ras* in cooperation with E1A that the *trans*-activation function can be separated from those involved in transformation and immortalization (reviewed in Moran and Mathews, 1987). Whether this applies to transformation using E1A plus E1B remains to be clarified, as this involves a situation that is complicated by the fact that products from either region affect the expression of the other at the transcriptional level (see below). It would appear from transformation assays with H-*ras* that *trans*-repression of enhancer-dependent transcription is more likely involved in transformation (Lillie *et al.*, 1986; Moran *et al.*, 1986b). However, recent reports have cast doubt upon these early conclusions since mutants have been described which are defective in their transforming ability but not *trans*-repression, and still others in which only the *trans*-repression function is affected but not transformation (Subramanian *et al.*, 1988; Velcich and Ziff, 1988). These reports would then suggest that the transcriptional regulatory functions of the E1A proteins are not responsible for transformation.

It has long been recognized that the E1A proteins stimulate DNA synthesis in adenovirus-infected, serum-starved, G_0-arrested rodent primary or secondary cells (Shimojo and Yamashita, 1968). It appears that induction of DNA synthesis is mediated by both E1A proteins, either of which can function efficiently (Nakajima *et al.*, 1987; Zerler *et al.*, 1987; Kaczmarek *et al.*, 1986; Bruner *et al.*, 1988). It is most likely that this function has appeared in the evolution of the virus and is probably required when adenovirus infects cells of a normal animal host in which the most common cell types in tissues are G_0-arrested and are terminally differentiated. Thus, in natural infections, quiescent cells of the epithelial type (membranes lining both respiratory and digestive tracts) may be triggered to undergo DNA synthesis in order to support viral infection and subsequent proliferation. It would also seem likely that in transformation, for some reason, the infecting virus fails to complete its cycle so

that the host cell that is now primed to undergo DNA synthesis would continue its cycle in a manner similar to that of undifferentiated cells. The dedifferentiated cell may now continue to proliferate indefinitely and in the process acquire the various characteristics of an oncogenically-transformed phenotype. These acquired characteristics, in fact, could simply be due to direct effects by viral infection which could cause chromosomal aberrations (shift from G_1 diploid DNA content to G_2 diploid, aneuploid, and polyploid amounts of DNA), possibly due to premature reconstitution of the interphase nucleus and reinitiation of DNA replication prior to chromosome segregation and cytokinesis (Murray *et al.*, 1982).

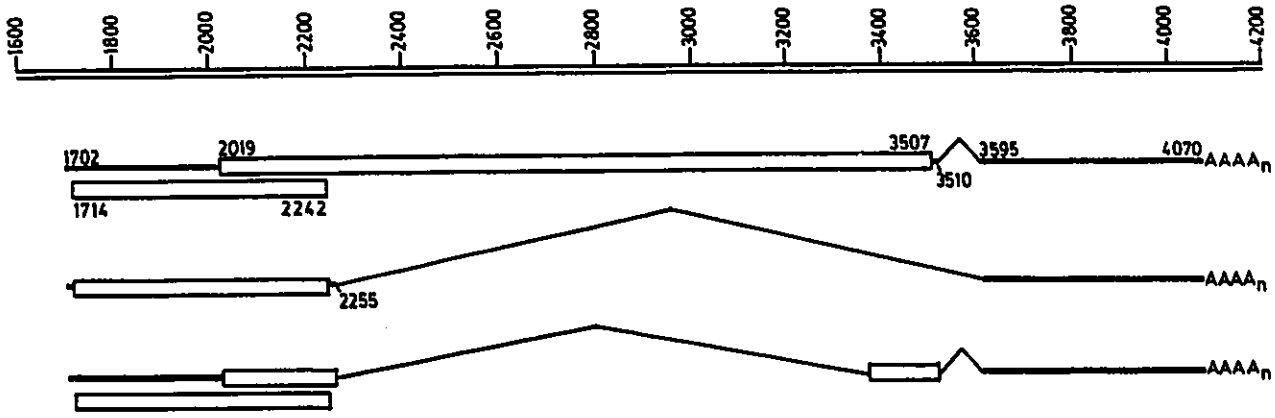
How are DNA synthesis and mitosis induced? It appears that growth factor(s) produced during infection could greatly stimulate proliferation of non-established epithelial cells (Quinlan *et al.*, 1987). This would explain, in part, the observation that BRK cells are induced to synthesize cellular DNA and proliferate in the absence of serum (Quinlan and Grodzicker, 1987) following infection by adenovirus containing E1A that expresses 12S alone. In fact, such growth factors are required during E1A-mediated immortalization of primary epithelial cells (Quinlan *et al.*, 1988), demonstrating a molecular basis for the E1A-induced physiological change in the E1A-expressing cell.

The E1A proteins also play an important role in the immune response of the host to adenovirus-infected and -transformed cells. This is treated in a later section.

2. Early Region 1B

The adenovirus type 5 E1B region and its products have not been studied as extensively as E1A, hence its role in infection and transformation is less clear. The region codes for a precursor RNA whose length is in excess of 2000 bases to give rise to 3 mRNAs (more in other serotypes such as Ad12) through differential splicing (**Figure 1-3**). The three mRNAs (22S, 14S and 13S) encode 2 major tumour antigens,

Figure 1-3. **Early Region 1B of Adenovirus Type 5**



the E1B 19 kDa and 55 kDa proteins, which are found in both virus-infected and -transformed cells.

Adenovirus mutants have been useful in uncovering the roles of various E1B-encoded products in both infection and transformation. The 55 kDa protein has been implicated in the shut-off of host cell protein synthesis (Babiss and Ginsberg, 1984a) and in the efficient nucleus-to-cytoplasm transport of late viral mRNAs in the infected cell (Babiss *et al.*, 1985; Pilder *et al.*, 1984 and 1986). In addition, in co-infection with adeno-associated virus (AAV), efficient and timely accumulation of AAV mRNA, proteins and DNA requires the large E1B protein (Samulski and Shenk, 1988). Its role in transformation is less clear, however, since it appears that the full protein is required in virus-mediated transformation but only a certain portion in the COOH-terminus is required in DNA-mediated transformation (reviewed in Branton *et al.*, 1985; Graham, 1984).

The 19 kDa E1B protein, like the E1A proteins, appears to be multi-functional. Being acylated it has preference for membrane association (McGlade *et al.*, 1987). It contains transcriptional regulatory activities (see below) and appears to be indispensable in E1-mediated transformation by Ad5 (or Ad2) (reviewed in Graham, 1984; Branton *et al.*, 1985). In addition, lesions in the E1B 19 kDa gene exhibit pleiotropic phenotypes during infection by adenovirus mutants. One of the earliest known phenotypes is the induction of abnormal cytopathic or cytocidal (*cyt*) effect (Takemori *et al.*, 1968; Takemori *et al.*, 1984). Second, E1B 19 kDa mutants induce degradation (*deg* phenotype) of host cell chromosomal and viral DNA (D'Halluin *et al.*, 1979; Enzo *et al.*, 1981; Lai Fatt and Mak, 1982; Pilder *et al.*, 1984; White *et al.*, 1984). Third, E1B 19 kDa mutants form large plaques (*lp* phenotype) (Chinnadurai, 1983; Subramanian *et al.*, 1984a and 1984b). Finally, they exhibit a host range (*hr*) phenotype that allows such mutants to have growth

advantage over wild-type when grown in human W138 cells (White *et al.*, 1986; White and Stillman, 1987).

Several reports have suggested that the E1B-encoded 19 kDa tumour antigen possesses transcriptional regulatory functions. Yoshida *et al.* (1987) have demonstrated a cell-type-specific activation of promoters linked to SV40, polyoma, E1A and immunoglobulin heavy chain enhancers in fibroblasts but not in myeloma cells using transient expression assays. In a different experimental test, Jochemsen *et al.* (1987) have also shown that E1B enhances levels of E1A transcription as determined by nuclear run-on assays of Ad5 E1A- and Ad5 E1-transformed cells. The stimulatory effect does not seem to be confined to enhancer-linked promoters since in transient expression assays, all adenovirus early promoters (E1A, E1B, E2A, E3 and E4) and the cellular heat shock gene (hsp70) promoter can be activated (Hermann *et al.*, 1987). In addition, White *et al.* (1988) have concluded from studies using mutant viruses that the 19 kDa protein has both positive and negative regulatory activities since it can stimulate viral gene expression and DNA synthesis in the absence of E1A, but displays *trans*-repression activity of E1A-dependent gene expression in the presence of E1A. Also, the Ad12 E1B 19 kDa protein has been shown to activate the human ß-interferon promoter in CAT (chloramphenicol acetyl transferase) assays (Shiroki and Toth, 1988). In contrast to these observations, Herbst *et al.* (1988) have used mutant viruses to demonstrate that the E1B 19 kDa protein does not enhance the rate of transcription initiation from the mouse immunoglobulin heavy chain promoter nor does it diminish the ability of E1A proteins to *trans*-repress the rate of enhancer-dependent transcription. However, they have not excluded the possibility that their observations might have differed from those mentioned because of the complexity of virus infections.

C. HOST IMMUNE RESPONSE TO ADENOVIRUS INFECTED AND TRANSFORMED CELLS

In response to the presence of foreign antigens such as viruses and other pathogens, the host mounts an immune reaction, both humoral and cellular, for the purpose of eliminating the pathogen and host cells that have been infected by the foreign agent. The humoral aspect of the immune response acts directly by eliminating extracellular viral particles using virus-specific antibodies in order to prevent its spread *in vivo*. On the other hand, the cellular response is much more complex as it involves interaction between cells of the immune system and cells expressing foreign antigens.

The cellular component of the immune response is responsible mainly for eliminating cells that have been virally infected or transformed. The recognition of neoplastic cells is believed to be mediated by two immune responses that differ in their time of appearance: an early-appearing response which is immunologically non-specific of which natural killer (NK) lymphocytes and activated macrophages are the main effectors; and a relatively late-appearing immunologically specific response of which the cytotoxic thymus-derived lymphocytes (CTLs) are the effector cells (Heberman and Ortaldo, 1981). Recognition of target cells by CTLs requires expression of class I major histocompatibility complex (MHC) antigens in association with which the virus-specific antigen is presented.

There are at least 4 reported closely related aspects of interaction between the host immune system and adenovirus infected and transformed cells, involving both E1A and E3 regions of adenovirus. First, a product of early region 3 (E3) is a glycoprotein with a predicted size of 19 kDa which has been shown to bind specifically to class I antigens of MHC (Kvest *et al.*, 1978; Signas *et al.*, 1982; Paabo *et al.*, 1983; Kampe *et al.*, 1983). Structurally, the E3 E19 glycoprotein

contains a 17 aa residue signal peptide that is followed by an intraluminal domain of 104 aa residues (Wold *et al.*, 1985), and at the COOH-terminus resides a stretch of about 8 aa residues that comprise the endoplasmic reticulum retention signal (Paabo *et al.*, 1987). Severinsson and Peterson (1985) have demonstrated that translation of mRNA specific for the E19 glycoprotein in *Xenopus laevis* abrogates the intracellular transport of class I MHC antigens. Similarly, Andersson *et al.* (1985) used adenovirus-infected HeLa cells to demonstrate that the association between E19 gp with class I MHC antigens also impaired intracellular transport of class I antigens, resulting in the accumulation of both in the perinuclear region. The intracellular accumulation of both class I antigens and E19 gp was paralleled by a concomitant loss of class I antigens from the cell surface. This has led to the notion that virally-infected cells expressing E19 gp evade host CTL immune surveillance by sequestering MHC class I products and viral antigens from the cell surface (Burgert and Kvist, 1987). It appears, however, that the modulation of cell-surface expression of class I MHC antigens by sequestration is confined to adenoviruses of subgenera B, C, D and E (Paabo *et al.*, 1986a,b), but not viruses of the highly oncogenic subgenus A, which have been shown to abolish expression by reducing mRNA levels at the post-transcriptional level possibly at splicing and polyadenylation via a different route (Vaessen *et al.*, 1987a). The E19 gp exhibits differential association with human MHC class I antigens since binding to HLA-A2 molecules is twice as high as with HLA-B7 antigens (Severinsson *et al.*, 1986). In the mouse, the H-2 K^d antigen associates with the E19 gp whereas the allelic product H-2 K^k does not (Burgert and Kvist, 1987); and expression of H-2 D^b but not H-2 K^b is affected by the E3 protein (Tanaka and Tevethia, 1988). These observations might lead one to suggest that the pathogenicity of the virus is dependent on the HLA-type of the infected individual. The role, therefore, of the E19 gp is to facilitate the *in vivo* replication of the virus

by evading the host's immune surveillance, thus leading to latent and persistent infections. That CTL-mediated lysis of infected cells is inhibited by E19 gp by blocking cell-surface expression of class I antigens has been demonstrated (Andersson *et al.*, 1987; Burgert *et al.*, 1987), which is consistent with the model described above. Similar E3 products from group B adenoviruses, Ad7 and Ad35, appear to function in an identical manner (Flomemberg *et al.*, 1987), hence the process is a general adenovirus-specific activity, except perhaps for those in subgenus A.

It is evident from the previous discussion that the interaction between a viral protein and class I MHC antigen expression provides a mechanism by which a virus can successfully infect a host by evading its immune surveillance. However, recognition of virally infected and transformed cells by the host is specified by certain transplantation antigens that are of viral origin, so that the fate of such cells is determined by the dominance of one mechanism over the other. As shall be seen, the adenovirus E1A region plays a major role in the immune interaction between infected and transformed cells and the host.

The two major products of E1A, 12S and 13S mRNAs, specify susceptibility of transformed cells to cytolysis by the tumour necrosis factor α or TNF-α (Chen *et al.*, 1987). TNF-α is a hormone secreted by monocytes primarily in response to various invasive agents, and it also possesses a wide variety of biological activities involving regulation of cell growth and differentiation, selective killing of some tumour cells, and growth stimulation of some normal fibroblasts (Chen *et al.*, 1987 and references therein). Thus, bacterially-produced TNF-α is able to lyse cloned NIH 3T3 and Fisher BRK-derived cells that are expressing exogenously introduced oncogenes. Both of the E1A products, 12S and 13S, are capable of inducing TNF-α susceptibility, although it is not clear whether the process requires cell-surface expression of E1A antigens. Surveillance against adenovirus-transformed cells through

the action of TNF-α-secreting monocytes represents the second mode of interaction between the host immune system and adenovirus-transformed cells. A strategy employed by the virus to overcome this particular host surveillance is another viral function. It appears that a 14.7 kDa protein coded in the E3 region inhibits the cytolytic effect of TNF-α (Gooding *et al.*, 1988). The role of the inhibitory 14.7 kDa protein is still not clear since it implies that in natural cases, cells that have been transformed by the virus can evade TNF-α-mediated cytolysis if they express this particular E3 protein. It would be of interest then to examine if the expression of the 14.7 kDa protein can be correlated with the degree of tumourigenicity of various transformed cells.

The interaction between immunologically specific cells (CTL-mediated, MHC-restricted) and adenovirus-transformed or -infected cells appears to differ in some subtle aspects from the interactions involving non-specific host cells (NK- and activated macrophage-mediated). CTL-mediated killing of virus-infected or -transformed cells requires at least two factors, according to a current concept of class I MHC antigen function (Doherty *et al.*, 1984): 1) sufficient level of class I MHC antigens on the cell surface, and 2) expression of virus-specific cell surface antigens. Early observations that highly oncogenic adenovirus-transformed, such as those by Ad12, have greatly reduced levels of class I MHC antigen expression (Bernards *et al.*, 1983; Schrier *et al.*, 1983) have been interpreted as the basis for the differential tumourigenicity of different groups of adenoviruses. Despite demonstrations that indeed an E1A epitope(s) is present on the surface of cells immortalized by E1A from a nononcogenic group and which is able to generate specific CTLs (Bellgrau *et al.*,1988), CTL-mediated immune surveillance does not seem to be the major mode of defence employed by the host against transformed cells. On the other hand, expression of either 12S or 13S E1A products is sufficient for induction of

susceptibility of E1-transformed rodent cells to lysis by non-specific inflammatory cells (Cook *et al.*, 1986). In the same manner, infected cells are made susceptible to lysis by NK and activated macrophages mainly by expression of either E1A product (Cook *et al.*, 1987). These results were obtained using nononcogenic Ad2 or Ad5 to infect cells, or their E1A genes to transform cells. Interestingly, induction of susceptibility is absent in Ad12-infected cells (Cook *et al.*, 1987; Cook and Lewis, 1984; Sawada *et al.*, 1985).

D. CELLULAR FACTORS INVOLVED IN TRANSCRIPTION

Genetic analysis of many eukaryotic transcriptional units revealed the importance of some *cis*-acting elements in maintaining the efficiency, accuracy and regulation of expression. Promoters, enhancers and silencers have been localized in many genes; their subsequent analysis using deletion and/or linker scanning mutants has revealed several regions or "boxes" that are important for function. In genes transcribed by RNA polymerase II, many promoters are found to contain the CCAAT box, which is found around -80 relative to the cap site, and more importantly the TATA box, which is located around -30. Many other promoters contain other elements that are either unique to that particular gene or are found in only a few other genes. It was thought that such boxes are the sites of interaction with cellular factors other than the RNA polymerase, which are most certainly involved in the modulation of expression.

The following is a short review of some relevant cellular factors that are known to interact with many transcriptional elements particularly those found in adenovirus early promoter elements. Identification of these factors and their interaction sites in promoter elements is the first step in understanding how the various early promoters are modulated during viral replication.

1. CCAAT-Binding Factors

The CCAAT box has long been recognized as an important region within the promoter of many mammalian genes (Benoist *et al.*, 1980; Efstratiadis *et al.*, 1980; McKnight and Tjian, 1986). It is now recognized that there are several different and, undoubtedly related, cellular proteins that can interact with the CCAAT motif and are capable of distinguishing between various CCAAT-containing elements (Dorn *et al.*, 1987; Oikarinen *et al.*, 1987; Chodosh *et al.*, 1988a,b).

a) CTF/NFI

CCAAT binding transcription factor (CTF) was the name given to a HeLa factor that binds to the HSV tk promoter *in vitro* (Jones *et al.*, 1985). Subsequent purification by sequence-specific DNA affinity chromatography and analysis of its biochemical properties (i.e., polypeptide composition, DNA binding properties, immunological cross-reactivity and *in vitro* stimulation of DNA replication and transcription initiation) have revealed that this factor is identical to NF1 (Jones *et al.*, 1987), the nuclear factor that is required for initiation of adenovirus DNA replication (Nagata *et al.*, 1983a). Gel analysis of purified CTF/NF1, as the factor is now commonly called, shows that it consists of a family of polypeptides that range in size between 52-66 kDa, suggesting that these are either products of post-translational modification or products of different but related genes that belong to the same family (Jones *et al.*, 1987). Comparison of several binding sites of the factor reveals 5'-TGG(A/C)N5GCCAA-3' as the consensus binding sequence.

The participation of CTF/NF1 in both gene transcription and DNA replication processes is an interesting combination of activities of a multi-functional protein since it is suggestive of some linkage between the two processes in which CTF/NF1 might play an important role. In several documented cases, eukaryotic origins of DNA replication contain transcriptional elements as important components (reviewed in

DePamphilis, 1988). Accordingly, eukaryotic origins of replication (ori) frequently contain a) a core component which is involved exclusively in DNA replication, as it determines where replication begins; and b) an auxiliary component comprising promoter or enhancer elements that, depending on certain conditions, may be required to determine both efficiency and/or cell specificity. Some examples that have been cited include the ori of SV40 and polyoma virus which contain enhancers affecting replication rate as well as cell-type specificity. Also, there are indications that DNA replication plays a role in the activation of β-globin gene expression (Enver *et al.*, 1988). It appears that adenovirus falls into the same category since the inverted terminal repeats (ITR's) found at the ends of the genome, and containing the ori's, are closely associated with transcriptional elements from E1A on the left-end, and E4 on the right-end. However, evidence is yet to be presented to support the idea that transcription from these terminally-located regions actually affects replication. But in analogy to both SV40 and polyoma virus ori's, enhancer elements are found on both E1A and E4 transcriptional elements. Also, aside from CTF/NF1, another replication factor NFIII which is required for adenovirus replication is known to bind to several promoters and enhancers (reviewed in DePamphilis, 1988).

One speculation as to the role of transcriptional elements in ori functions is that cellular factors that recognize transcriptional elements may interact with replication factors to stabilize their binding to ori's either by promoting localized strand separation or by increasing their binding activity. Another is that they may be involved in modifying chromatic structure perhaps by making it more accessible to initiation factors (DePamphilis, 1988). In adenovirus, CTF/NFI might play an important role in regulating transcription and replication at the termini.

b) CBF

Another CCAAT binding factor, CBF, was first identified by its binding to

the mouse α2(I) collagen promoter in the -80 region (Hatamochi *et al.*, 1986). Binding sites to ß-actin, RSV-LTR but not to α1(III) and SV40 promoters have been identified (Hatamochi *et al.*, 1986). *In vitro* studies using purified CBF have shown that this factor also stimulates the adenovirus major late promoter (Maity *et al.*, 1988).

c) C/EBP

Yet another CCAAT-binding protein was called CBP, which was partially purified from rat liver (Graves *et al.*, 1986). An insight into the activity of CCAAT binding factors was recently obtained in that when a C-to-G transversion in the first C of the CCAAT sequence occurred, the promoter activity of both MSV-LTR and HSV-tk was severely affected but at the same time increases the affinity of CBP (Graves *et al.*, 1986). Thus, it is believed that dissociation of this factor from its binding site is required for optimal function resulting in the induction of conformational change that might allow another factor to enter (Graves *et al.*, 1986). This protein also binds to the core homology common to many viral enhancers, hence the new name C/EBP to reflect the fact that both CCAAT and enhancer sequences are recognized (Johnson *et al.*, 1987). More recently, the gene for C/EBP has been cloned and the predicted sequence of the protein product appears to contain a DNA-binding domain with similarity to the products of *myc* and *fos* oncogenes (Landschulz *et al.*, 1988).

d) NF-Y

This CCAAT-binding factor was first identified as a sequence-specific protein that recognizes the Y-box, a promoter element common to all MHC class II genes (Dorn *et al.*, 1987). This factor differs from CTF and C/EBP as demonstrated by cross-competition experiments. For example, a mutation in the CCATT sequence which increased C/EBP binding, abolished binding by NF-Y (Dorn *et al.*, 1987). The

consensus NF-Y binding site is ATTGG/CCAAT. A similar factor, called ACF or albumin CCAAT factor purified from rat cells, was identified recently (Raymondjean *et al.*, 1988).

2. Sp1

The Sp1 factor was first detected in HeLa cells on the basis of its ability to activate the SV40 early promoter (Dynan and Tjian, 1983a). Its ability to recognize and bind selectively to the so-called GC box (a GC-rich decanucleotide sequence) which is present in six tandem copies in the SV40 promoter was demonstrated by Dynan and Tjian (1983b). Purified Sp1 appears as two prominent polypeptides, 105 and 95 kDa, respectively, the latter forming a dimer via intermolecular disulfide bonds to form a 190 kDa species (Briggs *et al.*, 1986). The prerequisite for activation by this factor appears quite simple: that at least one properly positioned GC-box element is contained in the promoter element (Briggs *et al.*, 1986). From a comparison of 36 different binding sites, a consensus binding site for Sp1 has been derived: 5' GGGGCGGGG 3'
TA TAAT

Recent mutational analysis of Sp1 (Courey and Tjian, 1988) indicates that its transcriptional activation domain, which is distinct from the DNA binding domain composed of three zinc fingers, contains glutamine-rich stretches, representing a novel structural motif for transcriptional activation. These highly basic domains are in sharp contrast to the better known acidic domains that are found in eukaryotic activators, including the unique region of E1A.

3. TATA-box Factor

Fractionation of whole cell extracts into components that could support *in vitro* transcription of the adenovirus major late promoter resulted in the identification three protein components, TFIIB, TFIID, and TFIIE, in addition to the RNA

polymerase II (Parker and Topol, 1984; Sawadogo and Roeder, 1985). Of these, TFIID binds to the TATA box, and probably comprises the basic transcriptional machinery since the TATA box is found in most class II promoters. The TATA box is most likely one of the targets for adenovirus E1A induction. This is discussed in a later section.

4. AP-1 Factors

AP-1 was first described by Lee *et al.* (1987a) as a transcriptional factor that is required for optimal expression of the metallothionein IIA and SV40 promoters, and later shown to interact with TPA-inducible enhancer elements (Lee *et al.*, 1987b). It binds to the consensus sequence 5'TGACTCA3'. Several polypeptides with apparent sizes ranging from 40-44 kDa can be detected in PAGE when purified AP-1 is analyzed (Angel *et al.*, 1987; Lee *et al.*, 1987b; Bohmann *et al.*, 1987; Angel *et al.*, 1988a; Chiu *et al.*, 1988). The major 44 kDa AP-1 is the product of proto-oncogene c-*jun* (Bohmann *et al.*, 1987; Angel *et al.*, 1988a; Chiu *et al.*, 1988) which was originally identified as the cellular counterpart of the oncogene of avian sarcoma virus (Maki *et al.*, 1987). The Jun/AP-1 product, as the 44 kDa AP-1 factor is referred to now, appears to play a central role in cellular signal transduction and regulation of proliferation. Specific *cis*-acting elements that mediate the response of some genes to phorbol ester tumour promoters such as TPA are recognized by the AP-1 polypeptides (Angel *et al.*, 1987; Lee *et al.*, 1987b; Chiu *et al.*, 1987). Several transforming oncogene products induce various genes using pathways that involve the AP-1 recognition site (Imler *et al.*, 1988; Wasylyk *et al.*, 1988; Schonthal *et al.*, 1988). The connection is important because the target of TPA is protein kinase C (Nishizuka, 1986), which is normally activated when ligands such as epidermal growth factor and tumour necrosis factor α bind to their cell receptors (Angel *et al.*, 1988b).

It has been shown that AP-1 responsive genes are stimulated at a higher rate when the product of another proto-oncogene, c-*fos*, associates with Jun/AP-1 (Chiu *et al.*,1988; reviewed in Curran and Franza, 1988). In the same manner, the induction of AP-1 responsive genes by TPA and transforming oncogenes (Schonthal *et al.*, 1988) is inhibited when Fos expression is prevented by antisense RNA. The formation of a complex between Fos and Jun/AP-1 and other AP-1 polypeptides enhances binding to AP-1 sites which presumably stimulates transcription (Nakabeppu, 1988). Jun/AP-1 products are able to stimulate the c-*jun* promoter due to the presence of AP-1 binding sites (Angel *et al.* 1988b). Fos and Jun/AP-1 form a heterodimer complex in a region that has sequence similarity with the DNA binding domain of the yeast transcription factor GCN4 (Vogt *et al.*, 1987). At the same time, Jun/AP-1 can form a homodimer by itself (Nakabeppu *et al.*, 1988; Halazonetis *et al.*, 1988) through the same sequences. It has been postulated that over-expression of the transforming v-*jun* product could lead to homodimer formation which could abrogate the normal mechanisms controlling gene expression (Halazonetis *et al.*, 1988) leading to virus-mediated transformation.

E. ADENOVIRUS EARLY PROMOTERS

The precise mechanisms of E1A-mediated *trans*-activation and *trans*-repression are not known despite the identification of most, if not all, *trans*-acting (cellular and viral) factors, and *cis*-acting elements found in the regulatory sequences of responsive genes. The adenovirus early promoters that respond to E1A regulatory effects are the transcriptional elements most extensively studied in attempts to elucidate these particular E1A functions (reviewed in Berk, 1986b). In the last few years, new information dealing with the problem of *trans*-activation mechanism has accumulated: DNA sequences to which various factors bind and, presumably, are the sites of

DNA-protein interactions have been precisely mapped, and the identity of such factors has been determined. Understanding these DNA-protein interactions must precede the elucidation of the mechanism(s) involved in E1A-mediated transcriptional control. The following discussion is an attempt to provide a current description of the adenovirus early promoters.

1. E1A Promoter

As discussed earlier, the products of the E1A region represent the master control for adenovirus infectivity since they are the first to be expressed, and efficient expression of the rest of the viral genome is E1A-dependent. As in many other viral systems, the expression of all viral functions is under strict temporal control to achieve maximum efficiency in viral growth and replication. Under the same scheme, E1A expression is also under both positive and negative regulation by its own products, and, as shall be appreciated, this is achieved via E1A's transcriptional control elements.

The expression of E1A is autoregulated by its own products: activation by the E1A 289 aa-residue protein (Hearing and Shenk, 1985) and repression by the 243 aa-residue protein (Borrelli *et al.*, 1984). *Trans*-activation by the large E1A protein, however, appears to be cell type dependent (Dery *et al.*, 1987). Extensive deletion mutagenesis of the E1A control element identified a duplicated E1A enhancer located around -300 and -200, respectively, relative to the E1A cap site (Hearing and Shenk, 1983a,b). These enhancer sequences contain one copy of the core enhancer element common to a number of DNA viruses such as the polyoma virus: nt199 to nt209 (AGGAAGTGACA) and nt297 to nt307 (AGGAAGTGAA). The activity of the E1A enhancer has been assayed using mutant viruses (Hearing and Shenk, 1983). Similar deletion analysis was carried out and the effect on E1A transcription was determined by transient expression assays (Sassone-Corsi *et al.*, 1983) and the results were in

agreement with those found for intact viruses. When assayed in a heterologous system, however, Hen *et al.* (1983) found another important promoter element in the -349 and -302 regions where the sequence GGATGTGTAA (for Ad2, or GGATGTGGCA for Ad5) is located; this sequence has a strong resemblance to the SV40 core enhancer region (GGGTGTGAAA; Weiher *et al.*, 1983). This region appears to be the binding site of a cellular factor present in uninfected HeLa cells (Barrett *et al.*, 1987). Another region which appears to have an enhancer-like activity when assayed with the E2A early promoter is located between -45 to -188 (Imperiale *et al.*, 1983) and possibly another in the ITR (Weeks and Jones, 1985).

The TATA box of E1A is typical (TATTTATA) being structurally similar to other TATA boxes and located in the -30 position relative to the start site. Although it is primarily involved in establishing a precise start site (Hearing and Shenk, 1983a; Osborne *et al.*, 1982), the TATA box shows a significant effect on the rate of transcription as demonstrated by various deletion studies (Hearing and Shenk, 1983a; Osborne and Berk, 1983; Osborne *et al.*, 1982).

An interesting finding was made previously in that a *cis*-acting element which is located within the protein coding sequence of E1A (at +399) appears to influence E1A transcription (Osborne *et al.*, 1984). The effect was detected in a mutant having a single base pair deletion and causing a reduction of transcription rate to 2% that of wild type.

The adenovirus transcription factor EIIA-EF (or ATF as it is presently known) binds to upstream promoter sequences of E1A (SivaRaman *et al.*, 1986). The consensus binding site for this factor (A/T)CGTCA (Lee *et al.*, 1987a) can be found in at least 4 positions upstream from the E1A start site, at -435, -403, -329 and -43, respectively (Hardy and Shenk, 1988). The binding site is a part of the bilaterally symmetrical 8-bp CRE sequence (TGACGTCA; Montminy *et al.*, 1986). As discussed

earlier, it appears that ATF and the CRE-binding protein are identical protein factors. Thus, some or all of these CRE-like binding sites, which are yet to be demonstrated by DNAse foot-printing assays, might be the sequences responsible for the synergistic induction by cAMP and E1A (Engel *et al.*, 1988). Identifiable Sp1 binding sites (GC boxes) are found at positions nt88-nt95 and nt356-nt361, the binding site for CTF/NF1 at position nt25-nt38, and lastly for the TATA-box factor TFIID in the -30 position.

2. E1B Promoter

The E1B promoter appears to be a relatively simple one being composed of a TATA box (-30/-23), a nearby GC box (-48/-39) and possibly a CCAAT box (-65/-57) (reviewed in Berk, 1986b). In contrast to the Ad12 E1B promoter which had been mapped between -135 and -122 relative to the start site (Bos and ten Wold-Kraamwinkel, 1983), the Ad2 E1B promoter appears confined within 65 bp upstream from the CAP site (Wu *et al.*, 1987). Mutational analysis revealed that only mutations in the TATA box interfere with E1A *trans*-activation (Wu *et al.*, 1987), raising speculations that one mechanism for E1A *trans*-activation involves the TATA box factor only. More recently, insertional mutagenesis has revealed that both the Sp1 binding site and the TATA box are under strict spatial constraints for the maintenance of efficient transcription (Wu and Berk, 1988a). This suggests that there is a specific interaction between Sp1 and factors that bind to the TATA box which is easily disrupted by alteration of the distance between them. This interaction, however, did not appear critical for induction of the promoter by either E1A 13S or the pseudorabies virus immediate early protein, suggesting that E1A *trans*-activation in this case is not mediated through facilitating the physical interaction between cellular factors, but rather through the increase in the activity of the cellular factors that bind to the TATA box (Wu and Berk, 1988b). In the herpes virus tk gene, similar

separation of the Sp1 binding site and the TATA box did not seem to affect basal level of expression (McKnight, 1982).

The E1B promoter is also stimulated by cAMP when tested in intact viruses through a mechanism that differs from the cAMP-stimulation of both E1A and E4 promoters since viruses lacking E1A were stimulated by cAMP as efficiently as those with intact E1A gene (Engel *et al.*, 1988). The stimulation by cAMP comes as a surprise because CREB-binding sites are not found upstream from the E1B promoter (Engel *et al.*, 1988). This raises the possibility that perhaps transcription from the upstream E1A promoter, which contains numerous ATF sites, positively influences transcription from the E1B promoter (Engel *et al.*, 1988).

3. E2A(e) Promoter

The early transcription region II (E2) is transcribed by RNA polymerase II and exhibits a relatively complex organization. At early stages of infection, three messages are transcribed from a promoter at 75.1 mu designated E2A early or E2A(e) (Stillman *et al.*, 1981; Baker and Ziff, 1981). Three messages are transcribed from this promoter, one coding for the 72 kDa DNA binding protein (designated E2A) and two coding for polypeptides involved in DNA synthesis (designated E2B). A second promoter at 72.0 mu operates at intermediate stages after infection (Baker and Ziff, 1981; Zajchowski *et al.*, 1985) and therefore designated E2A(late). Little is known about the switch from one promoter to another in the E2 region (Guilfoyle *et al.*, 1985; Leff and Chambon, 1986). It therefore serves as a convenient system for studying multi-promoter control of a transcriptional block.

The E2A(e) promoter responds to stimulation by E1A but is unusual in that its TATA box (TTAAGA) at around -30 position has a poor homology to the canonical TATA box sequence (reviewed by Breathnach and Chambon, 1981). Within this box is another start site (-26) which is used by about 20% of the total transcripts

(Mathis *et al.*, 1981), and an AT-rich region with better homology to the consensus TATA box is found approximately 30 bp away from the minor start site. It appears from linker-scanning studies that the poor TATA box as well as a region between -82 and -66 are important for expression (Murthy *et al.*, 1985). In these studies, however, no sequence was found to be responsible for induction by E1A.

Multiple factors interact with the E2A(e) promoter. One factor, E2F, was identified to interact with the promoter (Kovesdi *et al.*, 1986a), its level increasing markedly upon infection, and was dependent upon E1A (Kovesdi *et al.*, 1986b). DNase foot-printing has identified two non-overlapping regions protected by the E2F factor: -33 to -49 and -53 to -71 (Yee *et al.*, 1987). The recognition site, when present in a heterologous promoter could stimulate transcription but only when transfected in the presence of E1A (Kovesdi *et al.*, 1987).

Another factor, called E2A-EF, identified in HeLa cell nuclear extracts, binds to a region between -66 to -82 upstream from the major cap site (SivaRaman *et al.*, 1986) but, in contrast to the E2F factor, its level did not increase during infection (SivaRaman and Thimmappaya, 1987). This factor may represent a distinct cellular polypeptide whose binding activity and levels appear to be independent of E1A. Comparison of its binding site to that of other factors shows a strong homology to the Jun/AP-1 binding site. A potential AP-1 binding site is also found around the -125 position. As mentioned earlier, AP-1 binding sites were originally found in metallothionein and SV40 promoters and TPA-responsive enhancer elements. Although these AP-1 sites have not been confirmed, it is likely that they belong to the enhancer-like elements found in the E2 promoter (Imperiale and Nevins, 1984; Imperiale *et al.*, 1985).

A third factor appears to be the TATA-box-specific polypeptide which recognizes the region at -30 (Boeuf *et al.*, 1987) having a poor homology with the

canonical sequence. The binding sites for other well-characterized factors such as Sp1 and the CCAAT binding proteins cannot be found, indicating that these may not be involved in the expression of the E2 region.

4. E3 Promoter

Among the five early adenovirus promoters, the early region 3 (E3) promoter is perhaps the most strongly induced by E1A (reviewed in Sharp, 1984; Berk, 1986b). Located at 76.6 mu and transcribed in a left-to-right direction (Cladaras *et al.*, 1985a,b), the E3 promoter is positioned to the right of the E2 promoter at 75 mu. The back-to-back arrangement of these two E1A-inducible promoters, whose respective start sites are separated by approximately 500 bases, is interesting in that their regulation probably represents a level of control where competition for limited factors determines the relative transcription rate of each promoter.

Two independent cellular factor binding studies have revealed virtually identical maps for binding of such factors in the E3 promoter (Garcia *et al.*, 1987; Hurst and Jones, 1987). Detailed analysis of these binding sites show that 4 known transcriptional factors are most likely to be involved: TFIID (-7 to -37), ATF/CREB (-44 to -68), Jun/AP-1 (-81 to -103) and CTF/NF1 (-154 to -183). Garcia *et al.* (1987) concluded that binding of either ATF or Jun/AP-1 is sufficient to stabilize binding of the TATA box factor (TFIID), and that these 3 factors are needed for complete basal and E1A-induced transcription of the promoter.

The E1A-dependent enhancer in the E2A promoter (Imperiale *et al.*, 1985; Jalinot and Kedinger, 1986) appears to contribute to maximal E1A responsiveness of the E3 promoter (Zajchowski *et al.*, 1988). An element that maps between -82 to -71 within the enhancer appears to be responsible for this effect. As noted earlier, this corresponds to the overlapping ATF and E2F binding sites, suggesting that a common E1A-inducible enhancer coordinately controls the expression of both E2A and E3

promoters.

5. E4 Promoter

The E4 region is located at the right-end of the genome and is transcribed in a right-to-left direction (reviewed in Sharp, 1984; Berk, 1986b). Genetic studies with a mutant in which 98% of the E4 region was deleted have indicated that the progress of productive infection in permissive cell lines is delayed in events such as DNA synthesis, induction of the cellular gene for dihydrofolate reductase, and onset of switch from viral early to late gene expression (Yoder and Berget, 1986). The delay seems temporary since these events in the lytic cycle eventually approach wild-type levels, indicating that some functions encoded by E4 are needed to facilitate the switch from early-to-late (Yoder and Berget, 1986). Other events such as translation of E2A mRNA, accumulation of late viral RNAs and host cell shut-off are also perturbed in another E4 deletion mutant (Halbert *et al.*, 1985). Using another deletion mutant which was isolated with the aid of cell line W162 that is able to complement E4 defects, Weinberg and Ketner (1986) have shown that there is a 5 to 7-fold decrease in the rate of DNA synthesis which eventually reaches wild type levels after 24 hours. Finally, the E4 region is also important for efficient virus particle assembly (Falgout and Ketner, 1987).

An as yet unidentified E4 product appears to stimulate transcription from the E2A(e) promoter in cotransfection experiments (Goding *et al.*, 1985). When assayed in HeLa cells, the effect is 5 to 15-fold as measured by S1 analysis and requires specific DNA sequences between -48 to -19 in the E2A(e) promoter.

The E4 region, while positively regulated by E1A, is negatively regulated by the E2 region. During infection, the rate of transcription from the E1A region appears constant but rates from E2 and E4 decline between 4 (for E4) and 6 (for E2) hours post-infection (Nevins and Winkler, 1980). An E2 ts mutant, H5ts125, fails to exhibit

the normal decline in E4 transcription when grown in non-permissive temperatures; a similar effect was not seen for E2 transcription (Nevins and Winkler, 1980). The defect in H5ts125 was shown to be the 72 kDa DBP (Blanton and Carter, 1979). From competition studies, the action of DBP appears to be to directly suppress transcription from the E4 promoter, but not from E1A, E1B and MLP promoters (Handa *et al.*, 1983). To date, regulation of the E4 promoter by the 72 kDa DBP has not been shown to require specific DNA sequences, which is perhaps not surprising since DBP binds nonspecifically. The possibility exists that binding of DBP during initiation of replication is involved in the regulation of E4 transcription.

The E4 promoter has a complex organization that requires several *cis*-acting elements and sequence specific factors for expression and E1A activation (Nishigaki *et al.*, 1988). The TATA box of this promoter, whose deletion has greater effect on expression *in vitro* than *in vivo* (Hanaka *et al.*, 1987), is structurally redundant (TATATATA located at -32 to-26) which probably explains the unusually high number of cap sites situated in a run of T's. Watanabe *et al.* (1988) have identified at least 3 cellular factors, E4TF1, E4TF2, and E4TF3, that specifically bind to *cis*-acting elements identified previously by deletion studies (Hanaka *et al.*, 1987). E4TF1 binds to sequences between -134 and -149 in the E4 promoter which contain the sequence 5'-GGAAGTGACG-3' that is duplicated at -168 and shown to be a component of the E1A enhancer (Hearing and Shenk, 1983b). E4TF2 binds to the region between -118 and -94. E4TF3 binds to 4 specific regions around -40, -155, -255, and -260, all of which, except the region at -40, contain the sequence 5'-GTGACGT-3'; region -40 contains 5'-GTTACGT-3' instead. Leza and Hearing (1988) have identified the same factor, which they referred to as ETF-A, and assigned its binding sites to similar regions. The binding site of ETF-A/E4TF3 shows similarity to the cAMP response element (CRE; Short *et al.*, 1986) and, indeed, the E4 promoter responds to cAMP

activation (Leza and Hearing, 1988), thereby suggesting that this cellular factor might be the same as CREB, the protein that binds to the CRE of the somatostatin gene (Montminy and Bilezikjian, 1987). Also, Lee and Green (1987) have identified a cellular factor, called E4F1, with similar binding site sequence to ETF-A/E4TF3, making it very likely that E4F1 is the same as ETF-A/E4TF3. Furthermore, based on DNA-binding and transcription activation experiments with E4F1, Lee *et al.* (1987) demonstrated that this factor is the same as E2A-EF, the E2 factor whose level is not elevated by E1A (SivaRaman *et al.*, 1986), and competition studies and foot-printing analyses have indicated that E4F3 also binds to the E2A promoter indicating that an E2A factor is similar to E4TF3 (Watanabe *et al.*, 1988). Lee *et al.* (1987) call the E4F1/E2A-EF cellular factor the activating transcription factor, or ATF, which appears to bind to an important element of the E3 promoter (see below). It is apparent from this discussion that E4TF3, E4F1, E2A-EF and CREB are one and the same factor. This factor will be referred to as ATF/CREB henceforth. Aside from these factors, 3 sequences with high similarity to the binding sites for Sp1 can be identified between -253 and -247 (CGCGGGG), between -238 and -230 (GGGGCGGG), and between -215 and -210 (GGGCGG). Lastly, the binding site of CTF/NF1 is found around the -300 region.

It appears that the E4 promoter contains sites for various cellular factors whose binding to those sites determines the temporal regulation of the region. The complexity of its organization, however, makes it difficult to speculate as to the possible mechanism of its induction by E1A.

F. MECHANISMS FOR E1A-MEDIATED *TRANS*-ACTIVATION

The complexity of E1A *trans*-activation is perhaps indicative of several mechanisms, rather than a single one, through which the function is probably

mediated. Through the identification and characterization of several cellular factors that bind to specific sites in adenovirus early promoters, there has been a steady progress towards understanding some of these mechanisms. The following is an attempt to bring into focus three mechanisms for E1A *trans*-activation which are rather surprisingly different in view of the relatively small size of the E1A proteins.

1. *Trans*-activation Via the E2F Factor

The transcription factor E2F (Kovesdi *et al.*, 1986a,b) is one of two presently known cellular transcriptional factors whose binding activity increases as a function of E1A stimulatory effects; the other factor, TFIIIC, is involved in the transcription of tRNA and adenovirus VAI by RNA polymerase III (Hoeffler and Roeder, 1985; Yoshinaga *et al.*, 1986). E2F was first detected as factor that binds to the E2A promoter (Kovesdi *et al.*, 1986a,b). Presently, it is not known how E2F is induced by E1A, whether the E1A-induced increase in binding activity is caused by post-translational modifications or due to an increase in quantity of the factor. Alterations of the protein that could affect activity could occur in several ways but phosphorylation may be a likely post-translational modification in analogy to the activation of TFIIIC by the E1A protein (Hoeffler *et al.*, 1988). Regardless, it was evident that an increase in the binding of E2F to its sites on the E2A promoter correlated with the latter's inducibility by E1A. Therefore, in the presence of E1A, a responsive gene is transcribed because a cellular factor which possesses an otherwise low binding activity is activated via a post-translational modification.

2. *Trans*-Activation Via the TFIID (TATA box) Factor

As mentioned previously, most class II genes contain an AT-rich element referred to as the TATA box which when present specifies accurate initiation of transcription *in vivo* and is indispensable *in vitro*; hence, the TATA box and the associated cap site sequences (transcription initiation sites), are considered core

promoter elements (Breathnach and Chambon, 1981). These core elements in association with RNA polymerase II make up the basic transcriptional machinery. The factor that recognizes and binds to the TATA box is called TFIID (Parker and Topol, 1984; Sawadogo and Roeder, 1985). Binding of the factor to the TATA box can therefore serve as a point where transcription regulation may be modulated by regulatory factors either directly (factor facilitating or preventing binding to the TATA box) or indirectly (factor associates with other factors to mediate activity). Elements (including enhancers) and their associated factors that are located upstream from the TATA box have the potential to play the role of intermediary factors that specify the modulation of a particular promoter. In this context, disruption of the interactions between these upstream factors and the TATA box could result in the loss of regulation and/or efficiency of transcription.

There have been a number of studies in support of these ideas. One of the first was carried out using the adenovirus major late promoter (MLP) with which it was shown that the binding of the factor MLTF to an upstream element of MLP was greatly stabilized by binding of TFIID to the TATA box (Sawadogo and Roeder, 1985). Perhaps the most dramatic effect of an interaction between an upstream-binding factor and TFIID was observed in the adenovirus E1B promoter where the basal rate of transcription was greatly reduced when the distance between the binding site for Sp1 and the TATA box was increased by a half-turn of DNA helix but only slightly by a full-turn (Wu and Berk, 1988a). Since the basal promoter activity of this promoter was contained within a short span containing the Sp1 and TFIID binding sites, it is believed that the interaction between these factors, most possibly by protein-to-protein contacts when bound in the DNA, is critical for transcription (Wu and Berk, 1988a). Similar results were obtained when more complex promoters such as the adenovirus E4 promoter were used: interaction

between the transcription factors ATF and TFIID is necessary to facilitate the establishment of a pre-initiation complex (Horikoshi *et al.*, 1988a). Finally, when the binding sites for the yeast activator GAL4 were located upstream from the E4 TATA box to produce a hybrid promoter that responds to GAL4, it was noted that GAL4-responsiveness required an interaction with bound TFIID to facilitate a stable initiation complex (Horikoshi *et al.*, 1988b). All of these examples illustrate that interaction between the TATA box and at least one other activating factor is sufficient to activate transcription. But where does E1A-mediated *trans*-activation fit into this scheme?

In the case of the simple E1B promoter where only Sp1 and TATA factors are necessary for basal expression, mutations that affect E1A responsiveness always affected the basal rate (Wu *et al.*, 1987). As with other adenovirus early promoters, the E1B promoter is also activated by the immediate early (IE) protein of the pseudorabies virus, probably via a similar, if not identical, mechanism since it can functionally substitute for the E1A protein (reviewed in Berk, 1986b). Abmayr *et al.* (1988) have used *in vitro* transcription assays to demonstrate that the IE *trans*-activation of the adenovirus MLP could be mimicked in two ways: 1) by adding nonpromoter-containing DNA which presumably titrated out nonspecific DNA-binding proteins in the crude extract and, 2) by pre-incubating the template with purified TFIID. When such conditions were used to stimulate transcription, addition of IE did not stimulate transcription further. These results would suggest that any mechanism that could facilitate interaction between TFIID and its binding site may be sufficient to stimulate transcription. Because of the functional similarity between the pseudorabies IE and the adenovirus E1A protein (for E1B stimulation, see Wu and Berk, 1988 for example), one mechanism for E1A *trans*-activation could simply be to facilitate TFIID-promoter interaction at least in certain cases such as the E1B

promoter. This could perhaps be achieved by forming an active complex between E1A and TFIID to increase the DNA binding activity of TFIID through the potential DNA-binding zinc finger in the unique region of the large E1A.

Trans-activation through the TATA box alone has been demonstrated using promoters lacking sequences located upstream from the TATA box (Simon *et al.*, 1988). In these studies, promoters from various sources containing only the TATA element were used to control the expression of the reporter CAT gene and were transfected into HeLa that had been infected with adenovirus with or without E1A. An interesting observation was made in that inducibility of the various promoters by E1A occurred only when the TATA element contains the sequence 5'-TATAA-3' although the degree of inducibility varied considerably from one promoter to another. The variation in the TATA box sequence found in several promoters suggests the existence of several TATA-box binding factors, but it is likely that only one might be involved in E1A-induction, in particular that which recognizes 5'-TATAA-3'.

3. The Adenylate Cyclase Pathway and the ATF Connection

Cyclic-AMP acts as a second messenger through which a wide variety of intracellular responses are elicited by diverse extracellular stimuli (Sutherland, 1972). Many extracellular ligands, such as hormones, bind to a variety of specific cell-surface receptors; the binding activates membrane-bound GTP-binding proteins that either stimulate or inhibit adenylate cyclase, the catalyst for the synthesis of cAMP (Gilman, 1984). Sufficient intracellular levels of cAMP result in the activation of the cAMP-dependent protein kinase holoenzyme by the binding of cAMP to the holoenzyme's regulatory subunit, resulting in the dissociation of the activated catalytic subunits that contain the phosphorylation function (Nimmo and Cohen, 1977). The phosphorylation of one or more appropriate protein substrates elicits the particular intracellular response (Krebs and Beavo, 1979). In eukaryotes, it appears that

activation of the cAMP-dependent protein kinase is virtually the only direct response to intracellular cAMP signals (Kuo and Greengard, 1969; Coffino *et al.*, 1976; Walter *et al.*, 1977; Toda *et al.*, 1987). Among the specific physiological roles for cAMP-mediated signal transduction is the selective transcriptional activation of genes (Lamers *et al.*, 1982; Hashimoto *et al.*, 1984; Wynshaw-Boris *et al.*, 1984; Comb *et al.*, 1986; Montminy *et al.*, 1986; Ran *et al.*, 1986; Deutsch *et al.*, 1987; Montminy and Bilesikjian, 1987; Nakagawa *et al.*, 1988), which has attracted considerable attention in recent years. Recent studies have strongly implicated the catalytic subunit of cAMP-dependent protein kinase II as the mediator of cAMP action on gene expression (Van Buskirk *et al.*, 1985). Analysis of several cAMP-inducible promoters has revealed sequences that are similar to TGACGTCA, a core sequence of what is referred to as CRE (for cAMP-responsive element) such as is found in the transcriptional control elements of genes like somatostatin (Montminy *et al.*, 1986), tyrosine hydroxylase (Lewis *et al.*, 1987), c-*fos* (Greenberg *et al.*, 1985), phosphoenolpyruvate carboxykinase (Short *et al.*, 1986), and the α-subunit of human chorionic gonadotropin (Silver *et al.*, 1987). Equally important is the identification of a protein factor which links kinase activation with transcription of cAMP-responsive genes (Montminy and Bilezikjian, 1987), and is referred to as CREB (CRE nuclear binding protein). This protein is similar, if not identical, to the nuclear factor ATF, which has been found to bind homologous sequences in most early adenovirus promoters (Hurst and Jones, 1987; Lin and Green, 1988; Sassone-Corsi, 1988).

In adenovirus, *in vitro* binding assays have revealed CREB binding sites in the promoters of the E1A, E2A, E3 and E4 genes (Hardy and Shenk, 1988). These CRE-like sequences are located in critical domains for the optimal expression of the E2, E3 and E4 promoters (Imperiale *et al.*, 1985; Leff *et al.*, 1985, Murthy *et al.*, 1985; Zajchowski *et al.*, 1985; Gilardi and Perricaudet, 1986; Garcia *et al.*, 1987;

Hanaka *et al.*, 1987; Lee and Green 1987). That these CRE-like sequences confer responsiveness to exogenously added cAMP was demonstrated recently (Engel *et al.*, 1988). When cells infected with wild type adenovirus were treated with dibutyryl cAMP, transcription from E4, E1A and E1B was greater than when cells were infected with a mutant virus lacking E1A--the levels of expression of both E1A and E4 when both cAMP and E1A were present being greater than when either inducer is present alone (Engel *et al.*, 1988). Such results demonstrate a synergistic action between cAMP and the E1A protein. Furthermore, cAMP induction of E1A and E4 was functionally divisible into E1A-dependent and -independent components, the former being absent in the response of E1B expression to cAMP (Engel *et al.*, 1988). As these experiments demonstrate a strong link between cAMP and E1A responsiveness, one is led to speculate as to the nature of such interactions since in normal infection by adenovirus, cAMP levels have not been reported to increase. There exists the possibility that a particular mechanism for E1A-mediated *trans*-activation utilizes components of the adenylate cyclase pathway. Engel *et al.* (1988) have proposed, on the basis of their finding that cAMP acts synergistically with E1A to activate some early promoters, that the activation of the cAMP-dependent protein kinase results in the phosphorylation of the E1A proteins, thereby increasing their *trans*-activation activity. Alternatively, the kinase acts on some cellular factors, thereby altering their interaction with E1A. These proposals, however, do not address the normal situation in which cells are infected in the absence of exogenously added cAMP. Nonetheless, there are strong indications that in an as yet undetermined manner E1A-mediated *trans*-activation is related to the adenylate cyclase pathway. The relationship might be important in view of the fact that the adenylate cyclase pathway is involved in the activation of a protein kinase which could have a plethora of effects on cellular activities. Clues which suggest that such a relation might indeed

exist include the fact that CREB-binding sites required for E1A *trans*-activation (Sassone-Corsi, 1988), are the same sequences that confer cAMP-responsiveness. Another clue might be the possibility that the factors CREB and ATF are the same, and even if they are not, the fact that they are very similar is intriguing. There are many questions that have remained unanswered. First, studies on E1A-mediated *trans*-activation involving the ATF pathway have not demonstrated any change in the constitutive level or DNA-binding activity of ATF following E1A expression (SivaRaman *et al.*, 1986) in contrast to the E2F factor whose level responds to E1A induction (Kovesdi *et al.*, 1986a,b). It is very likely then that the ATF pathway involves one or several additional factors through which the E1A-induction is presented. Second, what role is played by the cAMP-dependent protein kinase? There is strong evidence that it is directly involved in gene activation by cAMP since promoters that are otherwise responsive to cAMP do not show induction in cells lacking this particular protein (Montminy and Bilezikjian, 1987). Similarly, the synergistic effect of cAMP with E1A is also absent in cells lacking this protein (Engel *et al.*, 1988). This would suggest that E1A activates this protein kinase through some as yet unknown mechanism.

G. PROJECT DESCRIPTION

The genetic analysis of E1A using both classical and "reverse genetics" methods, has been extensive and continues to be pursued. This reflects the great interest in understanding the biological and biochemical activities of the E1A proteins that are still only beginning to be elucidated. In this study, in-phase insertional mutagenesis was the approach taken to analyze structure-function relationships of the E1A proteins. By inserting a mutator DNA duplex at random in the E1A coding region, it was hoped that a large set of insertion mutants could be generated, the use

of which might provide a different perspective in elucidating E1A proteins both in function and structure.

The number of E1A insertion mutants which have been constructed and characterized to date is small in contrast to the large number of other mutants, the most common of which are deletion and point mutations. As a group, the E1A insertion mutants currently available are too few and too diverse to define the functional domains of the E1A proteins. In the studies described in this report, an attempt was made to generate a relatively large library of insertional mutants to better define the domains involved in the various functions of the E1A proteins. There were four distinct stages in this study. The first stage was the development of an efficient method for inserting oligonucleotides into plasmids and the subsequent characterization of each mutant at the sequence level. The second stage involved the actual construction and structural characterization of the mutants. In the third stage, assays for E1A *trans*-activation and enhancer repression based on expression of the bacterial ß-galactosidase (*lac* Z) gene were developed. The fourth stage was the actual phenotypic characterization of the mutations using the newly developed assays. The results confirmed previous observations which mapped *trans*-activation to the unique region of E1A 13S product and indicated that this particular E1A function was very resistant to insertional mutations outside this region. It was also observed that the unique region of E1A showed differential sensitivity to insertions of 2 amino acid residues. Secondly, the *trans*-repression function could be inactivated by mutations in regions previously undefined by other mutational studies. The relationship between these newly introduced structural changes in the E1A gene and the observed phenotypes of the various mutants is discussed.

CHAPTER II

MATERIALS AND METHODS

A. BACTERIA, MEDIA, GROWTH AND STORAGE

Several strains of *Escherichia coli* K12 were used as hosts in various cloning experiments described in this report. These include **LE392** (F⁻ *hsd*R17($r_K^-m_K^-$) *sup*E44 *sup*F58 *lac*Y1 or Δ(*lac*IZY)6 *gal*K2 *gal*T22 *met*B1 *trp*R55 λ⁻; Murray *et al.*, 1977), **E5014** (Δ[*lac-pro*A]XIII, Spc^R, SuE [F'lac⁺ *pro*AB]; Reznicoff, *pers. comm.*), **HB101** (F⁻ *hsd*S20 ($r_B^-m_B^-$) *rec*A 13 *leu*B6 *ara*-14 *pro*A2 *lac*Y1 *gal*K2 *rps*L20(str^r) *xyl*-5 *mtl*-1 *sup*E44 λ⁻; Boyer and Roulland-Dussoix, 1969), and **JM83** (F⁻ *ara* Δ(*lac-pro*AB) *rps*L ϕ80 Δ*lac*ZDM15; Vieira and Messing, 1982). These cells were typically grown in LB broth (per liter: 10 g bacto-tryptone, 5 g yeast extract, 10 g NaCl, NaOH to pH 7.4) at 37°C with continuous shaking. In certain applications as noted, either B broth and M9 minimal media were also used. B broth contained 10 g of bacto-tryptone and 8 g of NaCl for every 1 liter in distilled water and was sterilized by autoclaving. M9 minimal medium was prepared as follows: M9 salts were prepared as 10X concentrate (60 g Na_2HPO_4, 30 g KH_2PO_4, 5 g NaCl and 10 g NH_4Cl for every 1 liter, autoclaved). To make 1 liter of M9 medium, the following reagents were mixed together under sterile conditions: 100 ml of 10X M9 salts, 1 ml of 1 M $MgSO_4.7H_2O$, 1 ml of 1% vitamin B1, 10 ml 10 mM $CaCl_2$, a carbon source (eg., 10 ml of 20% glucose or 20% glycerol), and enough autoclaved distilled H_2O to make 1 liter.

Table 2-1. List of Plasmids obtained from Various Sources

Name of Plasmid	Reference
pMC1871	Casadaban *et al.* (1983)
pXC1	McKinnon *et al.* (1982)
pHB1	F.L. Graham (*pers. comm.*)
pUC18/19	Yanish-Perron *et al.* (1985)
pLOP	F.L. Graham (*pers. comm.*)
pHE1	McKinnon *et al.* (1985)
pSV2CAT	Gorman *et al.* (1982)
pKCAT23	Weeks and Jones (1983)
pSV2NEO	Southern and Berg (1982)

Bacterial cells in liquid media were kept at 4°C for temporary storage, or in 20% glycerol at -70°C for extended storage as stocks. To recover bacterial cultures from frozen stocks a small volume (about 2 ml) of a growth medium was inoculated with a loopful of frozen cells.

B. PLASMIDS

Several plasmids (**Table 2-1**) that were used in assays and as sources of restriction fragments for cloning purposes were obtained from various sources. Other plasmids that were constructed are described in detail in the **RESULTS** section and **APPENDIX.**

C. ENZYMES

1. Restriction Enzymes

The restriction enzymes used in this study were obtained from commercial suppliers, namely Bethesda Research Laboratories, Inc. (BRL), New England Biolabs (NEB), Pharmacia, and International Biotechnologies, Inc. (IBI). All enzymes were received in solutions of various storage buffers containing 50% glycerol. Unless otherwise noted, restriction enzyme digests were carried out in a

standard buffer containing 10 mM Tris.Cl (pH 7.5), 50 mM KCl, and 10 mM $MgCl_2$. The stock solution prepared as 10X concentrate was stored at 4°C where it was stable for several months.

Although this restriction enzyme buffer was compatible with most enzymes, the activities of some were less than optimal. In most cases there was a reduction of activity (eg., *Kpn*I, *Sma*I and *Xba*I) and only rarely (eg., *Eco*RI, *Pvu*II, *Bam*HI) was there an apparent increase in activity relative to that obtained when the recommended buffer was used. This problem, however, was not critical in many applications since it could often be remedied by adjusting the amount of enzyme and the length of incubation as required. In extreme cases, there were some enzymes that required other buffers, usually differing only in NaCl concentration, to maximize their activity (eg., *Xma*III, *Nla*IV) and that adjustments in amount of enzyme and/or incubation time was not suitable for other reasons. In such cases the buffer recommended by the manufacturer, usually accompanying the enzyme when purchased, was used. Inhibition of enzyme activity due to high glycerol concentration as a result of multiple enzyme digests was avoided by ensuring that the total volume of enzyme used did not exceed 10% of the final reaction volume, or the equivalent of about 5% glycerol.

The reaction mixture was usually in a 25 µl volume when the digest was used for analysis in gel electrophoresis (**section F**) or in appropriate volumes for other purposes. Digestions using crude DNA extract, as opposed to CsCl/ethidium bromide-banded preparations (**section G2**), required the use of a two-fold excess of enzyme and/or longer incubation time. Thus, in the digestion of DNA prepared by the alkaline-SDS method (**section F3**) from a 1.5 ml bacterial culture, about 1/50th of total DNA yield was digested with 2 units (usually defined as the amount

needed to digest 1 μg of a test DNA at 37°C for 1 h) and incubated at 37°C for 2 hours or more.

a. Multiple-enzyme digests

In several instances simultaneous digestions with two or more different enzymes were required. The use of a single enzyme buffer was particularly convenient in such cases even if the activity of the enzymes being used was less than optimal in the universal buffer. Again, the variation in the activities of the enzymes was often not critical in carrying out many simultaneous digests so that an equal amount of enzymes was usually sufficient to obtain complete digests. However, when the problem occasionally arose, similar adjustments in amount of enzyme and incubation time were made.

b. Inactivation of enzyme

In some instances where, for example, a cloning experiment required multiple steps, it was necessary to inactivate the enzymes used in one step before proceeding to the next. Most enzymes used in this project were labile when subjected to heat-treatment at 70°C for 15 minutes. Others required higher temperatures (around 80°C) to effectively destroy the activity, while still others required phenol/chloroform-extraction followed by ethanol-precipitation. The phenol/chloroform reagent was prepared by mixing 1 volume of redistilled phenol, 1 volume of chloroform, and 0.1% 8-hydroxyquinoline (w/v). This was fully saturated in TE buffer (pH 8.0) and stored in the dark at 4°C indefinitely.

c. Partial digestion with multi-cut enzymes

An important technique in the generation of the set of insertion mutants was the partial digestion of plasmid DNA with restriction enzymes that recognize multiple sites on the same plasmid. Enzymes used had recognition sequences that

are 4bp long and, therefore, cut relatively more frequently than those with longer recognition sequences.

Using such enzymes to linearize a supercoiled plasmid was accomplished using the method of Parker *et al.* (1977) which is based on the differential binding of ethidium bromide to supercoiled and linear plasmid DNA. In a typical reaction volume of 100 µl, 20 µg of supercoiled DNA, normal concentration of the universal buffer, ethidium bromide at a final concentration of 2.6 µg/ml, and 1 unit of enzyme were added together. The incubation time in which approximately 10% linear DNA was generated was empirically determined for each enzyme. For example, 10 µl aliquots were removed from the incubation reaction every 1 min and the enzymatic reaction was stopped by adding 5 µl of 6X gel loading buffer (**section F4**). Using gel electrophoresis, the extent of reaction was estimated by the concentration of the different forms of plasmid DNA (see **RESULTS**). Keeping the fraction of linear form to low levels (at about 10%) was desirable since it reduced the levels of linear DNA that contained more than one cut to less than about 1%.

2. Klenow Fragment of DNA polymerase I

The large peptide of DNA polymerase I due to cleavage by subtilisin (Jacobsen *et al.*, 1974), commonly called the Klenow polymerase, was used to remove protruding single stranded regions at the termini of DNA fragments produced by some restriction enzymes. In the presence of deoxynucleotide triphosphates (collectively called dNTPs which include dGTP, dATP, dTTP, dCTP; all purchased in powder form from Boehringer Mannheim, prepared as 10 mM solutions in 10 mM Tris.Cl pH 8, stored at -20°C), termini with recessed 3' ends were filled by the polymerase activity of the fragment. However, the enzyme also possesses a 3'-to-5' exonuclease activity which can be utilized to remove a

3'-overhang. This procedure was used primarily to prepare DNA fragments that have incompatible termini for ligation by T4 DNA ligase.

Treatment with Klenow DNA polymerase was typically carried out using a nick-translation buffer (Maniatis *et al.*, 1982) that contained 50 mM Tris.Cl (pH 7.2), 10 mM $MgSO_4$, and 0.1 mM DTT (BSA was omitted). A 10X concentrate of this buffer was prepared and stored in small aliquots at -20°C. The reaction mixture (50 µl) also contained 2-3 µg of DNA and 0.5 mM each of the dNTPs. This was incubated at room temperature for 30-40 minutes. The reaction was terminated by heat inactivation (70°C, 15 min) and the unincorporated dNTPs were removed by either microdialysis or spin-column dialysis (Maniatis *et al.*, 1982). If there was a need to concentrate the DNA, it was either ethanol-precipitated or dried in SpeedVac and redissolved in appropriate volumes; otherwise, the solution was used directly in molecular cloning.

3. Alkaline Phosphatase

In some applications either for cloning or for end-labeling DNA fragments, enzymatic removal of the free 5'-phosphate was necessary. In cloning, the procedure was used to prevent the self-ligation of vector DNA, thus promoting the ligation between the vector and the fragment intended to be cloned, which was usually left with the 5'-phosphate attached. In other situations, DNA fragments were needed to be end-labeled at the 5'-end by T4 DNA kinase and γ-^{32}P-ATP. Usually the DNA fragment was in a restriction enzyme reaction mixture whose buffer system was compatible with alkaline phosphatase, so that changing the buffer was not necessary.

The reaction mixture was prepared by adjusting the concentration of Tris.Cl to 50 mM (pH 8.0) through the addition of appropriate amounts of a 1 M stock.

Between 1 and 2 U of enzyme per 1 µg DNA were used. Incubation was at 37°C for 30 minutes. Then a single phenol/chloroform extraction followed by ethanol precipitation was performed to remove traces of the enzyme and to make a change in buffer. Usually this purification step was sufficient for end-labeling dephosphorylated DNA. However, if the purchased enzyme was stored in a buffer containing ammonium ions, it was necessary to purify the fragment by gel electrophoresis or by GeneClean® in order to be re-phosphorylated successfully (section C4).

4. T4 Polynucleotide Kinase

The enzyme T4 DNA kinase was used to label the 5'-ends of DNA fragments (dephosphorylated restriction fragments or single-stranded synthetic oligonucleotides) with γ-^{32}P-ATP (from NEN, sp. act. 3000 Ci/mmole, 10 mCi/ml), or to simply phosphorylate synthetic oligonucleotides with cold ATP before ligation. A typical reaction mixture contained a polynucleotide kinase (PNK) buffer with 50 mM Tris.Cl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM Na_2EDTA which was prepared as a 10X concentrate and stored at -20°C. The amount of enzyme (supplied by either BRL or Pharmacia) added was typically 1-2 U per 20 µl reaction volume but was varied with different applications.

a. Labeling of primers for DNA sequencing

Since γ-^{32}P-ATP was relatively less expensive than α-^{32}P-labeled deoxynucleotide triphosphates, in the dideoxy-sequencing reactions (section H) 5'-end-labeled primers were used instead of the usual method of using one of the deoxynucleotides labeled with ^{32}P or ^{35}S. In a typical reaction of 100 µl, 60 µl primer (120 ng), 10 µl 10X PNK buffer, 28 µl γ-^{32}P-ATP, and 2 µl T4 PNK (10 units) were mixed together. This was incubated at 37°C for 45 min and used

directly for sequencing without purification from the unincorporated ATP.

b. 5'-end labeling of restriction fragments

Restriction fragments that were to be end-labeled with γ-^{32}P-ATP and T4 DNA kinase were dephosphorylated first since kinasing using the forward reaction (as opposed to replacement reaction) was more efficient by several orders of magnitude. As noted previously, phosphatased DNA fragments in solution must be free of ammonium ions in order to be successfully phosphorylated. The conditions used were similar to those in section C4a.

5. T4 DNA Ligase

For cloning purposes DNA fragments were ligated together using T4 DNA ligase (BRL). The buffer used consisted of 50 mM Tris.Cl (pH 7.7), 20 mM $MgCl_2$, 20 mM DTT, and 0.2 mM ATP. This was prepared as a 10X concentrate and stored in small aliquots at -20°C. Each aliquot was thawed once and was used within a period of 2-3 week, after which the solution was discarded. As a rule this enzyme was used in all ligation experiments since it catalyses the covalent linkage between DNA fragments with blunt or recessed ends. Reaction mixtures were incubated at room temperature for 3-6 hours or at 15°C for 16 hours.

a. Ligation of linkers and vector DNA

The conditions used for efficient ligation between vector DNA and linkers were established empirically as described in the **RESULTS** section. Typically the reaction mixture was carried out in 20 µl volumes with 0.5 µg vector DNA and 3.5 pmoles of linker DNA.

b. Ligation of restriction fragments

Restriction fragments used in molecular cloning were usually gel-purified to enhance the efficiency of isolation of the desired clone. In a typical reaction of

Table 2-2. List of Oligonucleotides Used in this Work

NAME	SEQUENCE	DESCRIPTION
AB6	5'TCTAGA3'	XbaI/blunt
AB7	5'ATTGGGTCTAGA3'	XbaI/blunt
AB8	5'CCCAATTGTGAGCGGATA3'	CORE
AB9	5'ACAATTTCTAGA3'	XbaI/blunt
AB10	5'AATTGTTATCCGCTCACA3'	CORE
AB20	5'GGATCCTGA3'	BamHI/blunt
AB21	5'ACAATTGGATCC3'	BamHI/blunt
AB22	5'GGATCC3'	BamHI/blunt
AB23	5'ATTGGGTCAGGATCC3'	BamHI/blunt
AB24	5'CGGATCCTGA3'	BamHI/5'CG-overhang
AB25	5'ACAATTGGATC3'	BamHI/5'CG-overgang
AB26	5'CGGATCC3'	BamHI/5'CG-overhang
AB27	5'ATTGGGTCAGGATC3'	BamHI/5'CG-overhang
AB42	5'ATGGATCC3'	ATG/BamHI
AB43	5'ACAATTGGATCCAT3'	ATG/BamHI
AB44	5'ATTGGGGGATCCAT3'	ATG/BamHI
AB238	5'GATCCTGACCCAATTGTGAGCGGATAACAATTG3'	Full-length cassette
AB239	5'GATCCAATTGTTATCCGCTCACAATTGGGTCAG3'	Full-length cassette
AB289	5'GTGCAGGAAGGGATTGACTTA3'	nt779-nt799/sequencing
AB376	5'CCGGTTAAGCTTTAA3'	12S Mutants
AB377	5'CCGGTTAAAGCTTAA3'	12S Mutants
SAM67	5'ACAATTGGATCC3'	LOP-labeled *Bam*HI linker
SAM68	5'GGATCCAATTGTGAGCGGATA3'	LOP-labeled *Bam*HI linker
SAM69	5'GGATCCAATTGTTATCCGCTC3'	LOP-labeled *Bam*HI linker

25 µl, two fragments of approximately equal length were added at 0.5 µg each to achieve a molar ratio of 1:1 for the fragments involved. When fragments of different lengths were involved, the amount of the smaller fragment was adjusted to maintain the 1:1 molar ratio.

D. SYNTHETIC OLIGONUCLEOTIDES

All oligonucleotides (**Table 2-2**) were purchased from the Central Facility of the Institute for Molecular Biology and Biotechnology, McMaster University. The oligomers were dissolved in TE buffer (10 mM Tris.Cl pH 8.0, 1 mM Na_2EDTA) and stored at -20°C for long-term storage or at 4°C to avoid repeated freezing and thawing. The use of these oligonucleotides is described in appropriate sections.

E. RECOMBINANT DNA TECHNIQUES

1. *In vitro* Ligation of DNA Fragments

Ligation of DNA fragments was conducted as described in section C.

2. Transformation of *E. coli*

The method of Mandel and Higa (1970) for transforming *E. coli* cells using $CaCl_2$ as modified by Goodman and McDonald (1979) was used to introduce recombinant plasmids into bacterial cells.

a. Preparation of competent cells

An overnight inoculum of the appropriate bacterial cell line was grown in LB broth at 37°C. This was diluted 50-fold using fresh broth and grown with vigorous shaking until the OD_{660} was between 0.2 and 0.3. The culture was transferred into 50 ml plastic centrifuge tubes, cooled on ice for 15 min, and cells were sedimented by centrifugation using an IEC centrifuge (3000 rpm, 4°C). The cell pellet was resuspended in 20 ml of cold transformation buffer containing 75 mM $CaCl_2$ and 5 mM Tris.Cl, pH 7.5 and incubated on ice for 1 to 16 hours to make competent cells. The cells were collected again by centrifugation and resuspended in 2 ml of transformation buffer. At this stage, competent cells were either used directly for transformation or sometimes frozen in 10% glycerol at -70°C and stored in 200 µl aliquots in small glass vials. However, reduction in transformation efficiency by as much as 70% was observed when frozen cells were used so that when DNA was limiting, freshly prepared competent cells were preferred.

b. Transformation of competent cells

About 0.2 ml of a competent bacterial suspension was used for transformation with ligation reactions not exceeding 40 µl volumes. When frozen cells were used, the cells were thawed at 37°C and placed on ice immediately. The transformation

mixture was transferred to small glass vials and incubated on ice for 40 min to allow the cells to take up DNA. To facilitate uptake of DNA, cells were heat-shocked at 42°C for 2 minutes after which 3 ml of warm LB broth was added. The cells were incubated at 37°C for 1 hour without shaking to allow recovery and expression of antibiotic resistance gene(s) before plating on selection plates.

c. Selection

Selection plates contained the required medium with 1.8% agar (Difco) and ampicillin at a concentration of 50 µg/ml. The molten agar was allowed to cool to 44°C before adding antibiotics and then poured immediately after. The agar plates were stored at 4°C and pre-warmed at 37°C just prior to use. When Xgal (5-bromo-4-chloro-3-indoyl-ß-D-galactopyranoside) was required, 1 ml of 2% solution in DMF (dimethyl formamide) was added to every 500 ml molten agar at 44°C. Alternatively, 50 µl of the Xgal solution was spread onto agar plates before use.

Suspensions containing transformed bacteria were plated either directly (after pre-incubation at 37°C) or from appropriate dilutions. Dilutions were made so that transformed colonies on selector plates were well separated; that is, the colony count did not exceed about 150 colonies per plate. The plates were incubated at 37°C.

When M9 agar selection plates were used, it was necessary to change the medium of the transformed cells because the $CaCl_2$ present in the suspension and the phosphate salts present in the agar media were incompatible. The pellet of transformed cells was collected after centrifugation and resuspended in sterile water by mild vortexing prior to dilution in water (rather than M9 medium) and spreading onto plates.

3. Analysis of Transformants

Individual colonies were picked with a sterile wooden stick and were used to inoculate 2.0 ml fresh LB broth in a sterile glass test-tube. These were incubated at 37°C with vigorous shaking for 4-16 hours. About 1.5 ml of the culture was transferred to a fresh microcentrifuge tube and the rest was maintained at 4°C for temporary storage. The complete analysis of the transformants was carried out as quickly as possible (within two days) to ensure that the bacterial culture stored at 4°C remained viable. After analysis and verification, which in many cases was by restriction enzyme digestion and gel electrophoresis plus, in some occasions, sequencing, the desired transformant was recloned on agar to ensure the homogeneity of the bacterial stock and the plasmid that was to be used in further studies.

a. Extraction of Plasmid DNA

The procedure used for extracting plasmid DNA was the alkaline SDS lysis technique of Birnboim and Doly (1978). The protocol, with a number of minor modifications, was used to obtain relatively pure plasmid DNA preparations from small scale cultures for sequencing as well as molecular cloning.

A 1.5 ml culture transferred into a plastic microcentrifuge tube was pelleted using a bench-top microcentrifuge (Eppendorf) at 15,000 rpm for 15 seconds. The supernatant was removed by vacuum aspiration and the pellet was immediately put on ice. The pellet was resuspended by vortexing with 100 µl of buffered lysozyme solution (10 mM Tris.Cl pH 8.0, 50 mM glucose, 50 mM Na_2EDTA, stored at 4°C as a pre-mixed solution with the lysozyme added just prior to use at a concentration of 10 mg/ml w/v). The resuspended pellet was incubated on ice for 15 min then 200 µl of alkaline SDS (1% SDS and 0.2 N NaOH, mixed fresh from

10X concentrated stocks of SDS and NaOH, respectively) was added and mixed gently by inverting the tube several times. When properly mixed, the reaction mixture would turn relatively clear almost immediately or within 5 min during incubation on ice. The mixture was neutralized by adding 150 µl of 3 M sodium acetate (pH 4.8, pre-mixed and stored at 4°C) and mixed by quick vortexing which resulted in the precipitation of cellular components and denatured macromolecules within a few seconds. Vortexing the mixture aided in breaking the precipitate into finer particles which were easier to separate from the supernatant during centrifugation. To ensure complete precipitation of cellular protein and nucleic acid, the suspension was left on ice for at least 1 hour, and at times overnight if the DNA preparation was intended for use in sequencing. Prolonged incubation at this stage reproducibly reduced the amount of contaminating RNA in the crude preparation.

The pellicle was separated by centrifugation in a microcentrifuge for 5 min, and the clear liquid phase was transferred to another tube. To precipitate the DNA, 1 ml of 95% EtOH (-20°C) was added and mixed by inverting the tube. This was immediately centrifuged for 5 minutes (freezing to -70°C to aid in the rapid formation of DNA precipitate proved unnecessary). The ethanol was removed by aspiration and the nucleic acid pellet at the bottom of the tube was dissolved in 100 µl of autoclaved doubly-distilled water. This was precipitated a second time by adding 200 µl of cold ethanol and again immediately centrifuged (this did not require adding sodium acetate for precipitation). The pellet was dried for 15-30 min at 37°C and dissolved in 100 µl of autoclaved TE buffer (pH 8.0) and stored at 4°C.

4. Screening with Restriction Enzymes and Gel Electrophoresis

a. Restriction digests

Plasmid DNA extracted from a 1.5 ml inoculum was analyzed by restriction enzymes and agarose gel electrophoresis. The method for digesting crude plasmid DNA extract is described in **section C1.** The digestion (in 25 µl) was stopped by adding 5 µl of 6X stopper-gel loading buffer containing 10 mM Tris.Cl (pH 8.0), 1 mM Na_2EDTA, 20% sucrose, 0.1% bromophenol blue and 1% SDS (prepared by boiling for 15 minutes and stored at room temperature).

b. Gel Electrophoresis

Two types of gel electrophoresis buffer were used: a) Tris-borate EDTA (TBE) was used for electrophoresis at a relatively high voltage (6 V/cm and higher), and b) Tris.acetate EDTA (TAE) was used for lower voltage specifically when large fragments (greater than 6 kbp) were to be resolved. TBE buffer was prepared as a 10X concentrate and stored at room temperature for no more than 4 weeks. The final running buffer contained the following: 100 mM Tris.borate (pH 8.1), 83 mM boric acid, and 1 mM Na_2EDTA.

The TAE buffer gave better resolution for large fragments (greater than 4 kbp) when run at a relatively high voltage and was therefore the buffer of choice for analysis of large plasmids. The 10X stock was stored at room temperature for long periods. The final running buffer contained the following: 40 mM Tris.acetate (pH 7.9), 5 mM sodium acetate, and 1 mM Na_2EDTA.

Agarose Gels

Appropriate amounts of agarose (BRL) were mixed with 1X electrophoresis buffer containing ethidium bromide (0.1 µg/ml w/v) and boiled until completely dissolved. The solution was allowed to cool to 44°C before pouring. The gel was

run submerged in buffer which also contained ethidium bromide of equal concentration. To prevent formation of a pH gradient on the gel during electrophoresis, the buffer was recirculated by a pump when TAE, but not TBE, was used. Upon completion, the gel was viewed under UV (ultraviolet) illumination (long wave) and photographed with a Polaroid® camera using Type 57 film to produce a positive image in black and white.

Polyacrylamide Gels

Polyacrylamide gel electrophoresis (PAGE) for nucleic acid analysis were prepared at appropriate acrylamide concentrations (6-8% w/v) using a 30:1 ratio of acrylamide to BIS (BioRad). When lower gel concentrations were desired (eg., 4.5%) the ratio was changed to 19:1. Appropriate amounts acrylamide and BIS were dissolved in the buffer of choice and filtered using a 0.45 µm Nalgene filter. Immediately before pouring, ammonium persulfate (prepared as a 10% concentrate and stored at 4°C) was added to a final concentration of 0.1% (w/v), and 200 µl of TEMED to initiate polymerization. After pouring, the gel was allowed to polymerize completely for a period of 1 hour and was used immediately thereafter without pre-running. To obtain good resolution of small fragments (lower than 100 bp), the gel was run between 200-300 V. After electrophoresis, the gel was stained with ethidium bromide (0.2 µg/ml in electrophoresis buffer) and photographed as for agarose gels.

F. BATCH PREPARATION OF PLASMID DNA

1. Plasmid Extraction

Plasmid DNA, when needed in large quantities, was extracted using a scaled-up alkali-SDS procedure (Birnboim and Doly, 1979). A 9 ml overnight

culture from a frozen bacterial stock was prepared and used to inoculate 350 ml of LB broth with antibiotics and incubated overnight at 37°C with vigorous shaking. Plasmid amplification was usually carried out with chloramphenicol which was added when the optical density (OD_{660}) was between 0.5 and 0.6 resulting in increase in yield. After incubation overnight the bacterial culture was transferred into a 450 ml plastic tube and centrifuged at 4000 rpm in a Sorval GS3 rotor (4°C). After decanting the supernatant, the cell pellet was resuspended in 10 ml of lysozyme (5 g/ml w/v) solution using a pipet. After incubation for about 15 min, 20 ml of alkali-SDS was added and mixed by swirling. The mixture usually turned relatively clear and viscous indicating complete lysis of the cells. After 5 min incubation on ice, 15 ml of acidic sodium acetate was added and mixed by shaking the tube 1-2 times so that the pellicle that formed broke into fine pieces. This was incubated for at least 1 hour on ice. The relative purity of crude DNA extract appeared to be correlated with incubation time. The insoluble pellicle was removed by filtration through 4 layers of cheesecloth with no prior centrifugation of the suspension. The liquid portion was transferred to a fresh tube and 100 ml of cold 95% EtOH was added. After mixing thoroughly, the nucleic acids were pelleted by centrifugation at 7000 rpm in the Sorval (GS3 rotor, 4°C, 10 min). The pellet was resuspended in 10 ml of autoclaved doubly distilled H_2O and transferred into a fresh 50 ml plastic centrifuge tube. Cold EtOH was added to the 50 ml mark on the tube for the second precipitation. The tube was centrifuged at 4500 rpm for 10 min and then the supernatant was drained.

2. CsCl-Ethidium Bromide Isopycnic Centrifugation

The crude DNA preparation was further purified by cesium chloride and ethidium bromide centrifugation. Thus, the DNA pellet (drained of ethanol but

never dried) was dissolved by vortexing in 5 ml of 0.1X SSC (prepared as a 20X stock containing 3 M NaCl and 0.3 M sodium citrate) and 2 ml of 50 mM Tris.Cl pH 8.0 and 10 mM Na_2EDTA. To this solution, 9 g of cesium chloride was added and mixed thoroughly until completely dissolved. The density of this solution (about 1.64 g/ml) was checked by weighing 1 ml of the solution. The solution was transferred into a 12-ml Beckman ultracentrifuge tube (quick seal) and 0.5 ml of 5% (w/v) ethidium bromide was layered on top. The tube was completely filled with light paraffin oil before sealing and mixing and was centrifuged in a Beckman Ti50 rotor at 35,000 rpm for 40 hours at 15°C.

The DNA band that migrated at a higher density was extracted by puncturing the side of the tube with an 18G needle and collecting DNA band with a disposable syringe. This band was usually visible in ordinary light. The ethidium bromide was extracted three times with 2 volumes of 1-propanol saturated with TE buffer and cesium chloride. Finally, the DNA solution was extensively dialysed against TE buffer (three times in 4-liter flask, 4 hours each for the first and second and overnight for the third, at 4°C with constant stirring using a magnetic bar). After dialysis, the DNA preparation was stored at -70°C for extended storage or at 4°C if to be used frequently (to avoid repeated freezing and thawing).

G. DNA SEQUENCE ANALYSIS

1. Sanger Technique

The chain termination procedure for DNA sequencing dideoxy-nucleotides (Sanger *et al.*, 1977) was routinely used for sequencing of recombinant DNA molecules. Templates were prepared from double-stranded plasmid DNA by linearizing plasmid DNA with a restriction enzyme and denaturing the sample by

boiling (Korneluk *et al.*, 1985). In most cases, the primer was end-labeled with γ-^{32}P-ATP and T4 polynucleotide kinase (**section C4a**).

In a standard reaction, 2-3 µg of crude plasmid DNA (**section F3a**) was digested with 5 units of an appropriate restriction enzyme in a 20 µl volume. The incubation time was kept at a minimum since as observed by Korneluk *et al.* (1985) and as found in this work excessively long incubation resulted in the appearance of non-specific bands on the sequencing ladder. This was presumably due to some contaminants in both the crude DNA extract and the enzyme used. Thus, the reaction mixture was incubated for 15 minutes after which the tube was placed on ice to inhibit further enzymatic action.

In a fresh tube, 8 µl of linearized DNA was mixed with 5 µl of end-labeled primer (prepared as in **section C4a**) then placed in a boiling water bath for 10 min to denature the double-stranded template. The tube was then immediately placed on ice for at least 30 min to allow the primer to anneal with the template. (In the meantime, the deoxy/dideoxynucleotide mixes were prepared.) After collecting the solution to the bottom of the tube by quick centrifugation, 1 µl of Klenow polymerase (1 unit) and 1 µl of 0.2 M DTT were added, mixed thoroughly by vortexing, followed by a quick centrifugation.

The concentration of the deoxy- and dideoxy-nucleotides as suggested by Messing (1982) was slightly altered to enhance the sequencing ladder in the area immediately after the primer. The deoxy-mixes were prepared as shown in **Table 2-3** and dideoxy-mixes in **Table 2-4**. Equal volumes of a deoxy-mix and its corresponding dideoxy-mix (eg., 70 µl of ddGTP + 70 µl of G') were combined for each of the four mixes. (This could be stored at -20°C for up to 4 months without appreciable loss of activity. Four tubes marked G, A, T, and C, respectively, were

Table 2-3. Preparation of Deoxy-nucleotide Mixes

(All volumes in µl)

	G'	A'	T'	C'
Add 0.5 mM dGTP	2	20	20	20
Add 0.5 mM dATP	20	2	20	20
Add 0.5 mM dTTP	20	20	2	20
Add 0.5 mM dCTP	20	20	20	2
Hybridization buffer*	15	15	15	15
Total	77	77	77	77

*The hybridization buffer was a 10X concentrate containing 100 mM Tris.HCl pH 7.9, 600 mM NaCl, and 66 mM $MgCl_2$; stored in small aliquots at -20°C.

prepared for the 4 individual reactions. Using a micropipet, 2 µl of each of the four deoxy/ dideoxy-mixes was transferred into the corresponding tube that had been kept on ice. This was added with 3 µl of the template/primer plus polymerase mix that until now had been kept on ice. The two were mixed together by centrifugation and incubated at room temperature for 40 minutes. The reaction was stopped by adding 9 µl of stop/loading solution (0.1% (w/v) bromophenol blue, 0.1% xylene cyanol (w/v) in deionized, recrystallized formamide, and 1.0 mM Na_2EDTA), mixed by vortexing followed by quick centrifugation. The tubes were placed in a 90°C water bath for 3 min and the samples were either loaded immediately on a sequencing gel or kept at -20°C for storage. When a previously stored sample was used, reheating to 90°C was not necessary.

2. Sequencing Gel Electrophoresis

The sequencing gels were run on an IBI sequencing apparatus with an aluminum thermal plate to evenly distribute heat generated during electrophoresis. Two glass plates (46 cm X 38 cm and 43 cm X 38 cm, respectively; 0.5 cm thick) were matched so that their surfaces would give rise to gels with even thickness. The pairing of matched plates was done by checking for Newton's rings that appeared in areas where the distance between the two surfaces varied. Use of unmatched plates gave rise to anomalous migration of fragments in uneven areas.

Table 2-4. Preparation of Dideoxy-nucleotide Mixes

Mix 3.5 µl of 10 mM ddGTP +	66.5 µl H_2O	=	70 µl
Mix 5.6 µl of 10 mM ddATP +	64.4 µl H_2O	=	70 µl
Mix 8.4 µl of 10 mM ddTTP +	61.6 µl H_2O	=	70 µl
Mix 3.0 µl of 10 mM ddCTP +	66.5 µl H_2O	=	70 µl

The 10 mM dideoxy-nucleotide stocks were prepared form crystals in 10 mM Tris.Cl pH 8.0; stored in small aliquots at -20°C.

The procedure to bind the polyacrylamide gel on one plate to facilitate drying and auto-radiography was adapted. The shorter plate was treated with 5 ml of bind-silane solution (30 µl of 3-(trimethoxysilyl) propyl methacrylate [Aldrich] mixed with 10 ml absolute EtOH and 30 µl glacial acetic acid, prepared fresh) spread evenly with a Kimwipe® onto the surface, followed by rinsing with 5 ml of 95% EtOH then drying with Kimwipe® until the surface was shiny. The longer plate was treated in much the same way except that repel-silane (3% [v/v] dichloro-dimethyl silane [Kodak] in 1,1,1-trichloro-ethane [Caledon]) was used; this inhibits the adhesion of polyacrylamide on glass surfaces and, therefore, facilitates the separation after electrophoresis. The plates were taped together to sandwich a plastic strip (0.4 mm thick) on each side.

The acrylamide mix was prepared by dissolving 42 g urea, 8.5 g of combined acrylamide (BioRad) and BIS acrylamide in 20:1 ratio, in a 100 ml solution containing 1X TBE buffer. This was heated slightly to facilitate solution of crystals and then allowed to cool to room temperature. The solution was filtered through 0.45 µm Nalgene filter and transferred into another glass container. Immediately before use, 1 ml of 10% ammonium persulfate (stored at 4°C for no more than 4 months) and 30 µl of TEMED were added. The gel was cast with this solution using a pipet usually within a period of 3 minutes. The gel usually polymerized shortly after but complete polymerization was allowed to occur for a period of 1 hour.

The gel was pre-run at a constant 1800V for at least 1 hour before samples were loaded. Depending on the width of the sample well, usually 3 mm or 5 mm, 2-3 µl of sample were loaded using a sequencing pipettor. The actual electrophoresis of samples was carried out at a constant 2000V.

After electrophoresis, the plates were dismantled and the shorter plate, to which the gel was bonded, was soaked in 10% (v/v) glacial acetic acid and 5% (v/v) methanol to fix the gel and at the same time remove the urea. After 10 minutes of constant agitation, which aided in the removal of urea, the plate was further soaked in cold water to eliminate the acetic acid in the gel. The gel was dried at 37°C for 16 hours, or longer as required, and the radioactive bands were detected by auto-radiography on a Kodak X-ray film (XAR-5, 35 x 43 cm). A 6-hour exposure was usually sufficient when fresh isotope was used and proportionally longer with older stocks of isotope.

H. GEL PURIFICATION OF DNA

1. Purification from Agarose Gels

Restriction fragments greater than 300 bp were fractionated by agarose gel electrophoresis and purified using a) the method described by Girvitz *et al.* (1980) with modifications or b) using the commercially supplied GeneClean® kit. Gels were run protected from light to avoid breakage of DNA since ethidium bromide was customarily added to the electrophoresis buffer.

a. Trapping of DNA with dialysis membrane

In the Girvitz *et al.* method, the desired DNA fragment was trapped in chromatography paper with a dialysis membrane support placed directly in front of the migrating fragment during electrophoresis. This procedure was modified by

using a dialysis membrane alone. When TBE electrophoresis buffer was used, a relatively high voltage could be achieved without exceeding the output limits of ordinary power supplies. Thus DNA fragments could be trapped with a dialysis membrane under an electric field of between 150V-200V. The dialysis membrane was lifted in one smooth motion so as not to disturb the trapped DNA. The DNA, which was detectable when illuminated with a UV lamp, was washed off with about 200 µl TE buffer with 1% SDS, and then placed in a small tube. This was extracted twice with buffered phenol/chloroform and precipitated. Following drying in SpeedVac and dissolving in TE buffer, the DNA solution was ready for other applications without any subsequent purification.

b. Purification Using GENECLEAN®

The procedure suggested by the supplier was followed. This method was rapid and reliable and was the method of choice for purifying DNA from agarose gels when the quantity was small.

2. Purification from Polyacrylamide Gels

The procedure described by Maxam and Gilbert (1980) was used to purify restriction fragments fractionated with polyacrylamide gels.

I. DETERMINATION OF DNA CONCENTRATIONS

For many applications, the precise concentration of DNA stocks was not absolutely required. For example, standard cloning techniques allow a wide range of near-optimal conditions that quite often guarantee the success of obtaining the proper clones without the need to know the precise DNA concentrations of reactant fragments. However, in transfection assays, knowledge of the concentration of plasmid DNAs used was necessary for quantitative calculations. The following

combination of various methods to determine DNA concentration was applied for the mutants.

First, the concentrations of DNA stocks purified by cesium chloride and ethidium bromide isopycnic centrifugation technique was approximated by obtaining the optical density of a diluted solution at 260 nm. In the second step, the determined OD_{260} was used to prepare a diluted DNA stock at 100 ng/µl. The absolute concentration of this preparation was then determined by the diphenylamine method using a standard DNA solution of salmon sperm DNA whose concentration was determined from its optical density at 260nm wavelength.

1. Diphenylamine Method

A modified diphenylamine assay (Giles and Myers, 1965) was used to determine the absolute concentration of DNA stocks. Since this is a colorimetric reaction assay, the concentration is expressed as an interpolated value from a standard curve established by a simultaneous assay with known amounts of salmon sperm DNA. Thus, the DNA sample was diluted in doubly distilled water to a final volume of 250 µl in a glass disposable test tube. Using a pipetman, 250 µl of 20% perchloric acid was added and mixed by shaking the tube slightly. Immediately after, 500 µl of 4% diphenylamine (w/v in glacial acetic acid) was added and mixed. To this 50 µl of 0.15% acetaldehyde was added. The samples were incubated overnight at room temperature to allow a complete reaction to occur. The absorbance of the samples was determined at wavelength 595 nm using a Beckman® DU7 Spectrophotometer. Although a single wavelength reading at 595 nm was normally sufficient, the results were invariably improved if a second reading at 700 nm was subtracted from the first value.

The standard curve which was used to determine the unknown values was established using a software written for the Apple IIe computer. Thus the paired data (amount of DNA vs. OD_{595}-OD_{700}) were entered and the regression line was calculated. The reliability of the assay was assessed by the correlation coefficient, in which case 0.95 was considered acceptable; otherwise, the assay was repeated.

J. MAMMALIAN CELL SYSTEM

Two mammalian cell lines were used in transient expression assays: HeLa and 293 cells (Graham *et al.*, 1977). HeLa cells were grown in α-minimal essential medium (α-MEM) supplemented with 10% newborn calf serum. For prophylactic purposes, penicillin (100 U/ml), streptomycin (100 µg/ml), and fungizone (2.5 µg/ml) were added in the medium. 293 cells were grown in Joklik's medium with similar supplements except that 10% horse serum was used.

1. Growth in Tissue Culture Dishes

Both HeLa and 293 cells were propagated as monolayers in 150 mm petri dishes and maintained in incubators at 37°C with 96% humidity and 5% CO_2. Nearly confluent monolayers were passaged as follows. The spent medium was removed by vacuum aspiration and the cells were washed with phosphate buffered saline once (PBS^{2-}, prepared as 10X concentrate containing 80 g NaCl, 2.0 g KCl, 11.5 g Na_2HPO_4, 2.0 g KH_2PO_4 for every liter in distilled water; autoclaved and stored at 4°C). Trypsin (Gibco) in PBS^{2-} was added at 2 ml/dish and incubated at room temperature. After the cells in culture have partially detached (within approximately 5 min) from the bottom of the dish, the cells were completely dislodged by slightly tapping the side of the dish. Fresh medium was added and mixed, then proper dilutions were made to be distributed to fresh petri dishes.

2. Maintenance of Cells Used in Transfection Assays

The maintenance of HeLa cells used in transfection assays required some care to minimize variability that is often associated with this type of assay. First, cells were grown as monolayers until just before full confluence was reached. Cultures passaged at 90% confluence and diluted 1:5 usually reached the same cell density within 2 days, and sometimes 3 days. Second, at least two consecutive passages (including a change of medium) were performed within 4 days before the cells were used in transfection assays. This ensured that all the cells were in log-phase at the time of transfection. Third, cells were trypsinized at room temperature so that the possibility of cell damage during incubation with the enzyme was minimal. These precautions appeared to maintain the consistency of the behavior of the cells, at least as far as transfection assays were involved.

K. TRANSFECTION ASSAYS ON HELA AND 293 CELLS

1. Preparation of Carrier DNA

High molecular weight DNA (genomic) used in transfection studies was extracted from cell monolayers. Cells close to confluence were washed with sterile distilled water or PBS^- and to each 150-mM petri dishes was added 2 ml of pronase (0.5 mg/ml) in a buffer containing 10 mM Tris.Cl (pH 7.5), 5 mM Na_2EDTA, and 0.1% SDS. The solution was spread evenly by tilting the dish several times and then placed in a cell culture incubator for about 15 minutes. The mixture from four dishes was scraped using a silicon-rubber policeman and transferred into a 50-ml plastic centrifuge tube for further incubation (3-16 hours) at 37°C. The mixture was extracted twice with one volume of buffer-saturated redistilled phenol (with 0.1% 8-hydroxyquinoline) and once with one volume of

water-saturated diethyl-ether to remove residual phenol. The aqueous phase was recovered and placed in a beaker to which sodium acetate (pH 8.0) was added to a final concentration of 0.2 M. To precipitate the DNA, cold ethanol (-20°C) was poured slowly into the beaker while the solution was stirred with a pasteur pipette. The DNA was spooled around the pipette and excess ethanol was squeezed out. Under sterile conditions, the DNA was dissolved in 4 ml TE buffer (pH 8) and stored at 4°C.

2. Transfection Using the Calcium Technique

Transfection assays using the calcium technique (Graham and van der Eb, 1973a; Graham *et al.*, 1980) were carried out using HeLa or 293 cells. Actively growing cells from at least two previous passages were used. When HeLa cells were used, 2 x 10^6 cells were seeded in 100 mm petri dishes, while 293 cells from a monolayer (90% confluent) in 150 mm dishes were divided into five 100 mm dishes. After 16 hours, the cells were fed with a fresh 10-ml medium.

The transfection cocktail (for 4 dishes) was prepared as follows. Into one 15-ml centrifuge tube 2 ml of 2X HEPES-buffered saline (2X HeBS, prepared by dissolving 8.0 g NaCl, 0.37 g KCl, 0.1 g Na_2HPO_4, 5.0 HEPES and 1.0 g glucose in 500 ml distilled water; adjusted to pH 7.1; autoclaved, stored at 4°C) was dispensed at room temperature. In a separate tube, the plasmid and carrier DNAs were mixed together and the final volume was adjusted to 1.5 ml with TE (pH 8.0) buffer. The amount of carrier DNA that was used was empirically determined according to the method of Graham *et al.* (1980). To this tube, 0.5 ml of 1 M $CaCl_2$ was added, mixing the solution well. After a brief low-speed centrifugation to collect the solution at the bottom of the tube, the solution was added dropwise to the other tube containing 2X HeBS using a sterile pasteur pipette as air bubbles

were introduced, again using a pipette, through the HeBS solution. Almost immediately a fine precipitate was formed; the cocktail was allowed to stand at room temperature for at least 30 min prior to addition to the cells (1 ml per 100 mm dish).

Five hours after the addition of DNA, the cells were glycerol-shocked treating four dishes at a time, each dish undergoing the same treatment at intervals of ten seconds. The medium was aspirated and 2 ml of 10% glycerol (in complete medium) was added and spread over the cells by tilting the dish several times. The glycerol solution was left for exactly 1 min after which 8 ml of fresh medium was added. After 40 seconds the medium was aspirated and 10 ml of fresh medium was added. Exactly 24 hours after the initial addition of DNA the cells were harvested.

3. Preparation of Cell Extracts

a. Freeze-and-thaw technique

Transfected cells were washed once with ice-cold PBS^- and cells from 1 dish were scraped with a silicone rubber policeman in 1 ml PBS^- and transferred into a 1.5 ml plastic microcentrifuge tube. Cells were collected to the bottom by a 10-sec centrifugation and the cell pellets were either immediately used or kept at -70°C until needed. To prepare cell extracts, the cell pellet from one dish was resuspended in 200 µl of freeze/thaw buffer (FT; 10 mM Tris.Cl pH 7.4, 10 mM Na_2EDTA, 25 mM sucrose) by vortexing for several seconds. Cells were broken by four cycles of freezing in liquid nitrogen (30 sec) and thawing in a 37°C water bath (5 min). Cellular debris was separated from the cytoplasmic extract by centrifugation for 5 minutes at 4°C using a bench-top microcentrifuge. The clear supernatant was transferred into a fresh tube and was used immediately or kept at

-70°C until needed.

b. Ultrasonication technique

An alternative to the freeze/thaw technique which was equally effective but many times more laborious was cell sonication. Cell pellets were resuspended in FT buffer by vortexing then subjected to sonication with 10 one-second bursts. To minimize the effect of thermal heat generated during the process, the tube was contained in ice. After sonication, the lysate was centrifuged to remove cellular debris as noted above.

c. Chloroform extraction technique

Total cell extract was also prepared using a procedure that is analogous to one used in bacterial cells (Miller, 1972). The procedure is relatively fast and simple and could be used where chloroform has been previously shown not to affect subsequent enzymatic assays. Briefly, 300 µl of FT-buffered chloroform was added to cell pellet and vortexed vigorously to resuspend the cells. The cells are disrupted by the action of chloroform on the cell membrane. To the lysed cells, 200 µl of FT buffer was added and briefly mixed by vortexing. The tube was centrifuged for 30 seconds (Eppendorf) and the aqueous phase was recovered using a pipetman then transferred into another tube.

Although the procedure produced consistent results, the lysate that was obtained contained traces of chloroform which interfered with the BioRad protein assay. This procedure was therefore used to prepare lysates that were utilized only in preliminary assays.

4. Quantitation of Protein Concentration

The BioRad protein assay was used to determine the concentration of proteins in lysates. In this assay, the standard curve was established using known amounts of bovine plasma γ globulin that was provided in the kit. Although the procedure was often accurate, as determined by the correlation coefficient of the standard curve, there were certain situations when erroneous results were obtained as when samples were contaminated by chloroform.

L. ENZYMATIC ASSAYS

1. ß-Galactosidase Assay

A popular and inexpensive method for assaying *lac* Z activity in cell extracts is a colorimetric one based on the hydrolysis of o-nitrophenyl-ß-D-galactopyranoside (colorless) into a yellow product o-nitrophenolate (Miller, 1972). Since the assay is fast, simple and inexpensive, many samples can be tested routinely. This is perhaps the greatest advantage offered by this type of assay when the activity of various mutants is to be determined as accurately as possible.

To assay for ß-galactosidase activity, 50 µl (approximately 1/4) of the total cell extract was added to a reaction mixture containing 150 µl of Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$ and 50 mM ß-mercapto-EtOH, adjusted to pH 7.0) and 40 µl of 4 mg/ml o-nitrophenyl-ß-galactopyranoside (ONPG from Sigma, in 60 mM Na_2HPO_4 and 40 mM NaH_2PO_4 in distilled water; stored in small aliquots at -20°C), and incubated at 37°C. The reaction was stopped by adding 100 µl of 1 M Na_2CO_3 normally after 30-60 min incubation. The extent of reaction was quantitated by determining the spectroscopic absorbance at 420 nm and the specific activity was calculated using the following formula

(adapted from Miller, 1972):

$$\text{Sp. activity (Units/mg)} = \frac{A_{420}/.0045}{\text{incubation time (min) X mg of protein}}$$

(Equation 1)

M. COMPUTER AND STATISTICAL ANALYSES

Most computer analyses were carried out using the MicroGenie® software (BRL) for personal computers, or the University of Wisconsin Genetics Computer Group (UWGCG) Sequence Analysis Package Version 5.3 for the VAX computer. Spreadsheet calculations were carried out using the AppleWorks® software. A program written in BASIC and 6502 microprocessor assembly language was developed for the purpose of determining molecular weights in gels using a HiPad® digitizer that was interfaced with an Apple IIe microcomputer. All statistical analyses were performed as described by Spiegel (1975). A hand-held programmable calculator with an auxiliary statistical package (Texas Instrument TI59®) was used for statistical calculations.

CHAPTER III

INSERTION OF OLIGOMERS VIA A LOP-LABELED CASSETTE

A. Technical Difficulties in Insertion of Oligonucleotides

We wished to construct and characterize a large number of insertion mutants in the E1A region of Ad5. This required screening of large numbers of candidate plasmids. The standard methods for inserting oligonucleotides into plasmids are frequently characterized by their inefficiency in the identification and characterization of clones with inserts. An ideal technique would permit identification of desired mutants at the primary level of screening when bacterial transformants are grown on selector plates. This would be very convenient since subsequent steps could be devoted to full characterization of the candidates as in establishing the location of the insert using standard sequencing techniques. Although the brute-force approach to analyze individual clones using restriction enzymes and gel electrophoresis can often be substituted by *in situ* hybridization technique (cells fixed on nitrocellulose membranes and probed with radio-labeled oligonucleotides), the investment in time and labour can be substantial as the number of candidates to be analyzed increases. Such a situation often arises when for some reasons the target DNA does not readily ligate to the oligomers. This problem, in addition to some difficulty in establishing the structure of mutant at the sequence level, has not made this approach popular for studying protein structure and function. A simple method to facilitate these processes might make insertional

mutagenesis a more attractive alternative to other methods of genetic analysis.

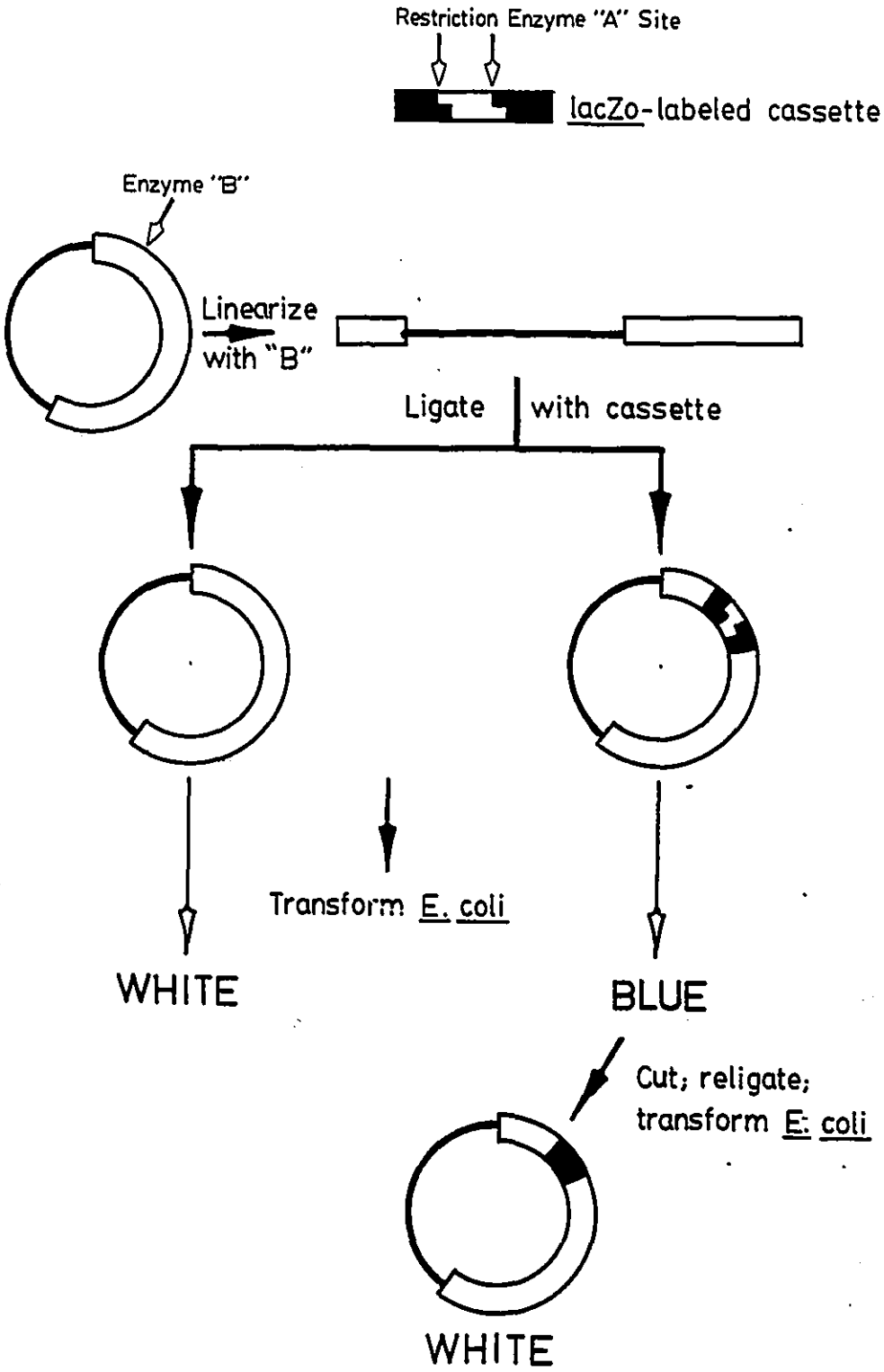
B. Development of a Novel Method for Inserting Oligonucleotides

The development of a simple and efficient method for inserting oligonucleotides into bacterial plasmids was undertaken. While it was originally intended solely for constructing E1A insertional mutants, the method is generally applicable for inserting oligomers of any sequence into plasmids. The main feature of the technique is the incorporation of the *lac* operator (*lacZo*) sequence within the oligonucleotides which one desires to insert. The biological property of *lacZo* to bind tightly to *lac* repressor (*lac* i) molecules serves as a phenotypic label since when present in multi-copy bacterial plasmids it is able to titrate out *lac* i molecules in the host, resulting in the constitutive expression of the *lac* operon (Marians *et al.*, 1976; Heynecker *et al.*, 1976; Sadler *et al.*, 1977). **Figure 3-1** illustrates how this method can be used to phenotypically label specific oligomers and how the ultimate insertion into a target plasmid is monitored by visual screening of bacterial transformants. As far as this technique is concerned, the most important *lac* operon product to be expressed is *lacZ* for which several colourimetric methods for detecting its expression have been established. For example, bacterial colonies harboring plasmids with the *lacZo*-labeled oligonucleotides are blue in colour when grown on agar containing the chromogenic substrate Xgal, while non-expressors are colourless. A similar idea for labeling DNA fragments with *lacZo* for cloning has been described previously (Sadler *et al.*, 1977; Danner, 1986).

Figure 3-1. **Phenotypic Labeling with Bacterial *lacZo* Sequence**

Oligomers containing the *lacZo* sequence (cross-hatched) and the sequence which one desires to insert (solid) into a target plasmid are combined to make a cassette via complementary sequences. The target DNA is linearized by a restriction enzyme "B" which generates ends similar to those of the cassette. Upon mixing and ligating, circular plasmids which had incorporated the cassette would confer on transformed bacterial cells both antibiotic resistance and constitutive expression of the *lac* operon by virtue of the *lacZo* sequence present in several copies that could titrate out lac repressor molecules. The phenotype is easily detectable in selector plates containing Xgal, a substrate which makes colonies appear blue when hydrolysed by *lac* Z. Their counterpart, those colonies transformed by a plasmid without an insert, remain colorless. The final step in obtaining the desired clone which contains the sequence of interest involves cleaving out the *lacZo* sequence using restriction enzyme "A" followed by religation. The plasmid containing the "collapsed" insert produces white colonies which are again distinguishable from their blue counterpart.

Figure 3-1. **Phenotypic Labeling with Bacterial *lacZo* Sequence**



A cassette designed to deliver sequences at random or into pre-determined restriction sites on a plasmid was developed. A similar cassette system for inserting oligonucleotides was earlier described in which the screening methodology differed in that the cassette was biologically marked with an antibiotic resistance gene (Barany, 1985) which could be purified as a restriction fragment from an especially constructed plasmid. In designing the present system, three features were incorporated. First, phenotypic labeling of the cassette by the *lacZo* sequence was used for easy identification of clones with inserts. Second, the cassette was designed for sufficient flexibility to introduce any desired sequence, limited only in size by the maximum length economically feasible with current automated synthesis techniques. And third, it allowed for fast and efficient sequencing of the plasmid with insert, primarily to establish the site and orientation of insert and to determine whether any modifications around the insert have occurred.

Figure 3-2a shows the general form of the cassette composed of up to 6 synthetic oligomers. By strand complementarity, the oligomers anneal to form the DNA duplex to be inserted. The two largest oligomers are 18 nucleotides long (A and B in **Figure 3-2b)** and make up the central core containing part of the *lacZo* sequence (Gilbert and Maxam, 1977) with 6 nucleotide-long 5'-overhangs on both sides. The cassette is completed by annealing with four other oligomers whose length and sequence can be varied depending on specific needs. In the example shown in **Figure 3-2c,** oligomers C, D and E were designed to anneal with each other and with A and B to generate a complete cassette having *Bam*HI recognition sites at each end. This particular cassette has blunt ends for insertion into sites generated by enzymes that are "blunt-end cutters" or into sites rendered blunt using

Figure 3-2. **Cassette Oligonucleotides**

Up to 6 oligonucleotides may form the complete cassette to be inserted (a). Two of these (b) comprise the core that contains the *lacZo* sequence. The flanking sequences (NNNNNN) are variable except for the 6-base overhang on both ends complementary to the core. Two examples of cassettes that could be formed by providing the necessary flanking oligonucleotides are shown: 1) a cassette with blunt termini flanked by *Bam*HI sites (using oligomers C, D and E), and 2) a cassette with 5'CG-overhang flanked by *Bam*HI sites (using oligomers F, G and H).

Figure 3-2. Cassette Oligonucleotides

a. General form of cassette

5'-NNNNNNCCCAATTGTGAGCGGATAACAATTNNNNNN-3'
3'-NNNNNNGGGTTAACACTCGCCTATTGTTAANNNNNN-5'

b. Oligos forming core of cassette

(A) 5'-CCCAATTGTGAGCGGATA-3'
(B) 3'-ACACTCGCCTATTGTTAA-5'

c. Flanking oligos with BamHI site and blunt ends

(C) 5'-GGATCC-3' 5'-ACAATTGGATCC-3' (E)
(D) 3'-CCTAGGGGGTTA-5' 3'-CCTAGG-5' (C)

d. Flanking oligos with BamHI site and 5'-CG overhang

(F) 5'-CGGATCC-3' 5'-ACAATTGGATC-3' (H)
(G) 3'-CTAGGGGGTTA-5' 3'-CCTAGGC-5' (F)

other means. In certain situations where insertion into sites with staggered ends is involved, designing the terminal with 'sticky' ends to match those of the target DNA is necessary. **Figure 3-2d** shows the type of oligomers (F, G, and H) that are necessary to complete a cassette having *Bam*HI sites and 5'-CG overhangs at each end. This particular cassette can ligate with DNA linearized with restriction enzymes *Hpa*II, *Hin*fI, *Taq*I and *Hin*PI as well as others.

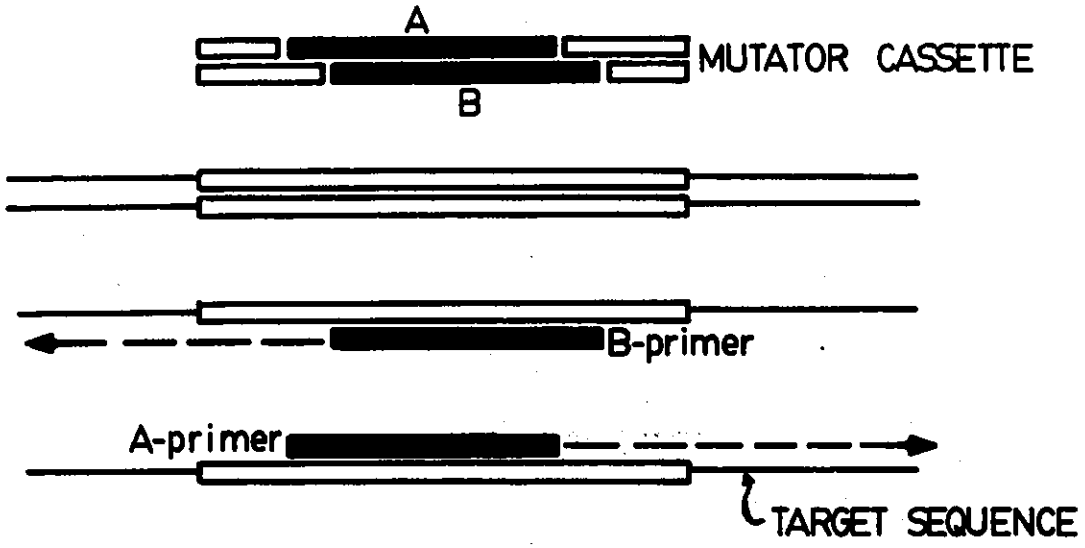
Sequence analysis of plasmids with the desired insert can be facilitated by making use of the core oligomers (**Figure 3-2b**) as primers for dideoxy-sequencing. These allow bi-directional sequencing (**Figure 3-3**) from the insert, and as shown below, the structure at the junction can be determined to establish the exact position and orientation of the insert in the target DNA provided the sequence of the latter is known.

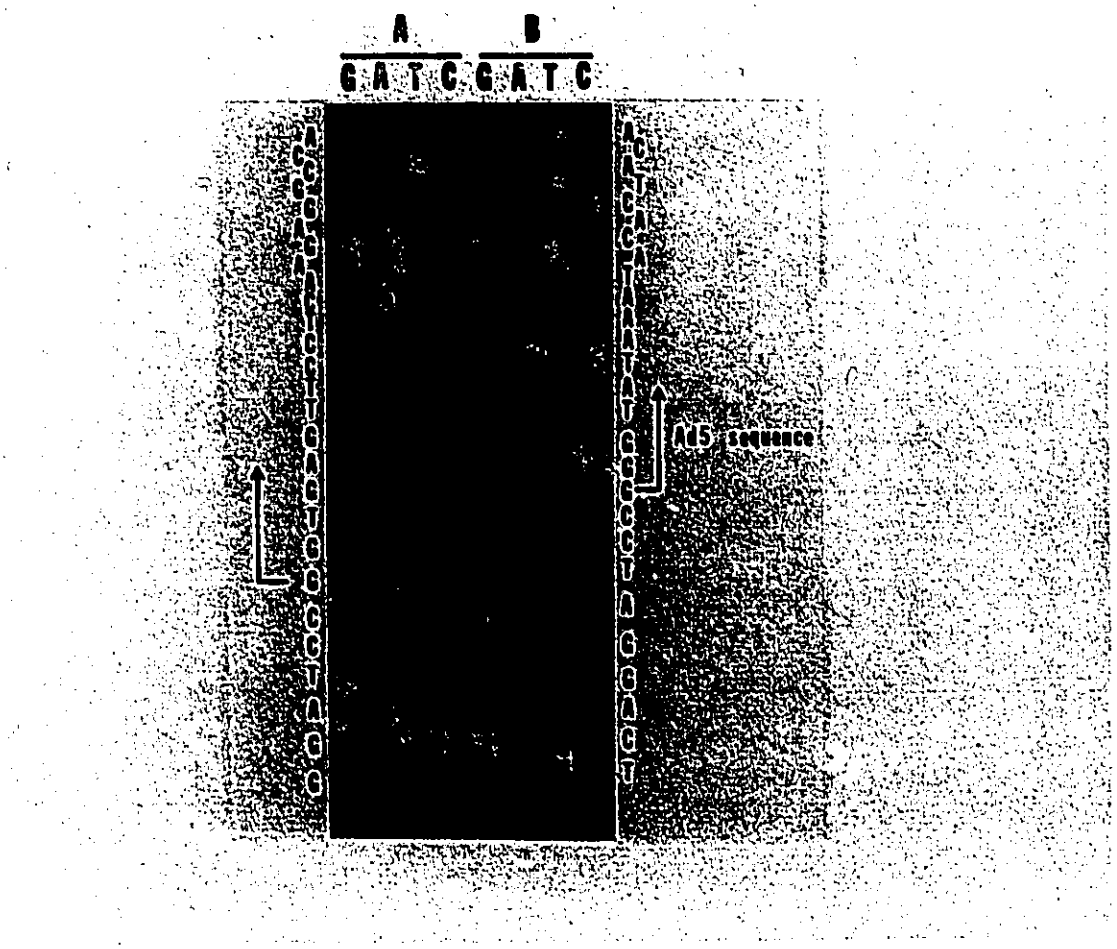
C. Description of the Mutagenesis Approach

The approach that was taken to mutagenize the E1A region involved the insertion of a 39-bp mutator DNA at each of a number of restriction sites in the E1A coding sequence. The mutator DNA was composed of six synthetic oligo-nucleotides that were designed to form a single duplex following annealing of complementary sequences. The termini of the mutator DNA were either blunt-ended to ligate with acceptor DNA molecules with similar termini, or contained a 5'-CG-overhang to ligate efficiently with those that have been linearized with enzymes such as *Hpa*II. The choice of using two mutators that differ only at the termini was dictated in part by the fact that many multi-cut enzymes that are commercially available generate termini of either type. Both mutator molecules

Figure 3-3. **Bi-Directional Sequencing**

The core oligonucleotides (shown here as A and B and also in **Figure 3-2**) may be used as primers for dideoxy-sequencing for the complete structural analysis of plasmids with the desired insert. Using double-stranded DNA templates, each of the core oligonucleotides can bind to its complementary sequence from which the primer extend during polymerase reaction. Reaction mixtures optimized to allow sequencing close to the primer (see **Materials and Methods**) gave results that allowed the unequivocal identification of the insertion site. In the example shown, a cassette with *Bam*HI flanking sequences was inserted into the Ad5 sequence defined by the enzyme *Hpa*II. Bi-directional sequencing from the insert identifies the *Bam*HI sites at the junctions between the insert and the target sequence. A sample gel showing results of sequencing using primers A and B is shown. Since the target sequence is known, the orientation of the insert can be established (see Chapter IV-D for more examples).


A
B
MUTATOR CASSETTE
B-primer
A-primer
TARGET SEQUENCE


A
B
G A T C G A T C
Ad5 sequence

could assume either orientation upon ligation. In one, referred to as "plus" or "+" orientation, the insert was designed to code for a stretch of 13 amino acid residues whose sequence depended on the translational reading frame. In the other, designated "minus" or "-" orientation, the insert contained a termination codon in each of the three reading frames so that wherever this orientation has been assumed by the mutator DNA in the coding region, translation would be terminated to generate a truncated protein. Since the termination codons were located within the insert, the truncated products would have varying C-termini composed of foreign amino acid residues whose length and sequence also depended on the reading frame.

Although there were actually two mutator molecules used, their coding potentials were identical. The predicted amino acid residues coded by the insert in the three reading frames are shown in **Figure 3-4**. One important property of the mutator DNA that made it a particularly attractive method to study E1A genetically was the potential to generate a third set of mutants from each mutant having either LOP insert with either orientation. This was made possible by "collapsing" the insert by joining the two *Bam*HI sites flanking the insert to generate a mutant with a 6-bp insert coding for 2 amino acid residues. It was thus possible to compare the effect of a 13aa residue insert with that of a shorter one of only two aa residues for each mutant generated.

D. Selection of a Bacterial Host for Optimal Expression of *LacZ*

A suitable host for cloning bacterial plasmids containing the *lacZo* sequence was selected from several candidates. Initially, three candidates, namely LE392,

Figure 3-4. **Mutator Cassettes**

The full sequence of the two mutator cassettes used to construct the E1A insertional mutants are shown. The predicted amino acid sequence in the three translational reading frames (I, II, and III) in two different orientations of the mutator cassettes are also shown. The stop codons in the closed orientation are underlined. Note that the order of reading frame was arbitrarily chosen.

Figure 3-4. **Mutator Cassettes**

1. CASSETTE WITH BLUNT ENDS

OPEN ORIENTATION:

```
          5'G-G-A-T-C-C-A-A-T-T-G-T-T-A-T-C-C-G-C-T-C-A-C-A-A-T-T-G-G-G-T-C-A-G-G-A-T-C-C3'
          3'CC-T-A-G-G-T-T-A-A-G-A-A-T-A-G-G-C-G-A-G-T-G-T-T-A-A-C-C-C-A-G-T-C-C-T-A-G-G5'
I           gly   ser   asn   cys   tyr   pro   leu   thr   ile   gly   ser   gly   ser
II        Gl asp   pro   ile   val   ile   arg   ser   gln   leu   gly   gln   asp   pro
III       G-Gl ile   gln   leu   leu   ser   ala   his   asn   trp   val   arg   ile  lC
```

CLOSED ORIENTATION:

```
          5'G-G-A-T-C-C-T-G-A-C-C-C-A-A-T-T-G-T-G-A-G-C-G-G-A-T-A-A-C-A-A-T-T-G-G-A-T-C-C-3'
          3'CC-T-A-G-G-A-C-T-G-G-G-T-T-A-A-C-A-C-T-C-G-C-C-T-A-T-T-G-T-T-A-A-C-C-T-A-G-G-5'
I           gly   ser   end
II        Gl asp   pro   asp   pro   ile   val   ser   gly   end
III       G-Gl ile   leu   thr   gln   leu   end
```

2. CASSETTE WITH 5'CG-OVERHANG

OPEN ORIENTATION:

```
          5'C-G-G-A-T-C-C-A-A-T-T-G-T-T-A-T-C-C-G-C-T-C-A-C-A-A-T-T-G-G-G-T-C-A-G-G-A-T-C
                C-T-A-G-G-T-T-A-A-C-A-A-T-A-G-G-C-G-A-G-T-G-T-T-A-A-C-C-C-A-G-T-C-C-T-A-G-G-C-5'
I           Cl gly   ser   asn   cys   tyr   pro   leu   thr   ile   gly   ser   gly   ser
II        C-Gl asp   pro   ile   val   ile   arg   ser   gln   leu   gly   gln   asp   pro
III         arg   ile   gln   leu   leu   ser   ala   his   asn   trp   val   arg   ile   arg
```

CLOSED ORIENTATION:

```
          5'C-G-G-A-T-C-C-T-G-A-C-C-C-A-A-T-T-G-T-G-A-G-C-G-G-A-T-A-A-C-A-A-T-T-G-G-A-T-C
                C-T-A-G-G-A-C-T-G-G-G-T-T-A-A-C-A-C-T-C-G-C-C-T-A-T-T-G-T-T-A-A-C-C-T-A-G-G-C-5'
I           Cl gly   ser   end
II        C-Gl asp   pro   asp   pro   ile   val   ser   gly   end
III         arg   ile   leu   thr   gln   leu   end
```

Table 3-1. Expression of *lacZ* in B-Broth Medium

HOST	(-)IPTG	(+)IPTG
LE392	++	+++
E5014	++	+++
HB101	-	+

End-point dilutions of bacterial culture in B broth were prepared so that when 50 ul was plated about 50 colonies could be detected in a single plate. IPTG was added at 0.01 ml of a 100 mM stock for every 0.3 ml diluted culture. Xgal was added to molten B-agar at 44°C at 1 ml (2% in DMF) for every 500 ml agar. "+" indicates detectable blue colonies, and "-" indicates no detectable colour.

E5014, and HB101, were compared for their ability to constitutively express *lacZ*, as detected in X-gal plates. Thus B-agar plates containing Xgal were used to plate out fresh cultures of the three hosts after the addition of IPTG in the culture for the induction of the *lac* operon. Results (**Table 3-1**) indicated that both LE392 and E5014 strains produced colonies that were deep blue in colour, and HB101 produced white colonies that contained blue patches in the middle. In subsequent studies, both LE392 and E5014 were used since the available strain of HB101 was not suitable.

The previous experiment indicated that both LE392 and E5014 have wild-type expression of *lacZ*. It also showed that they were capable of expressing in the absence of IPTG, indicating that there was a significant basal level of *lacZ* expression in B-agar, a relatively rich growth medium. This suggested that using a rich medium might not be appropriate for purposes of this study since there was no clear-cut difference in the colour of induced and uninduced colonies. Presumably, B-broth contains nonspecific inducers of the *lac* operon, resulting in blue colonies in the absence of IPTG.

The use of a minimal medium to discriminate between constitutive and non-constitutive expressors was investigated by examining the behaviour of LE392 and

Table 3-2. Expression of *lacZ* in M9 Minimal Medium

CARBON SOURCE	IPTG	Bacterial Host LE392	E5014
Glycerol	Add.	+++	+++
	Del.	+++	+++
Glucose	Add.	+++	+++
	Del.	-	-

Both LE392 and E5014 were grown in LB broth. Cells from small aliquots were sedimented by centrifugation and resuspended in sterile distilled water and plated on M9-based minimal medium agar with Xgal and with appropriate amino acids supplement. The inducer IPTG was either added or deleted as indicated. Colonies were detected 40 hours after plating. "+++" indicates deep blue colonies and "-" white colonies.

E5014 cells when grown under well-defined conditions. Two carbon sources were also evaluated, namely glucose and glycerol. The results (**Table 3-2**) indicated that when glycerol was used as the sole carbon source in M9 minimal medium, all colonies that appeared after plating suspensions at limiting dilutions were blue, in the presence or absence of the inducer IPTG. This result was in sharp contrast to the situation where glucose was the sole carbon source in that colonies only appeared blue when IPTG was added. The reason glycerol in the medium resulted in apparent constitutive expression of *lacZ* under uninduced conditions is likely due to the inherent ability of glycerol to elevate cAMP levels in *E. coli* (reviewed in Miller and Reznicoff, 1978). Under such conditions, the *lac* operon is largely under positive regulation by the CAP-cAMP complex which promotes transcription of *lac* operon genes, presumably by-passing the repressor level of control in the absence of an inducer (reviewed in Miller and Reznicoff, 1978).

The results in **Tables 3-1** and **3-2** when taken together indicated that the medium used for identifying constitutive expressors was critical when one desires to make a clear distinction between colonies carrying plasmids with *lacZo* inserts and those without. The use of a minimal medium gave the best results but only

when the appropriate carbon source was used. The conditions for clear discrimination between the two types of colonies were therefore M9 minimal medium + Xgal with glucose as carbon source.

The previous experiments did not establish whether plasmid-borne *lacZo* could result in constitutive expression of *lac*Z. Thus, in a subsequent experiment, the inducibility of *lac*Z was tested by using plasmids with or without the *lacZo* sequence. Three plasmids were used: pUCHI (a pUC8-based plasmid in which the plasmid-borne *lac*Z for α-complementation was inactivated by inserting the 1.0-kbp Ad5 *Hind*III-I fragment at the multiple cloning site; this plasmid contained the natural *lacZo*), pEST (a derivative of pUC8 in which a small *Pvu*II fragment was removed resulting in the deletion of *lacZo* sequences from pUC), and pLOP (a derivative of pBR322 in which a 33bp synthetic *lacZo* flanked by *Bam*HI sites was inserted at the *Pvu*II site; F.L. Graham, pers. comm.). These plasmids were used to transform to ampicillin resistance either LE392 or E5014 and transformants were selected using M9-agar with glucose as the sole carbon source. Results showed (**Table 3-3**) that on Xgal M9-glucose selector plates, transformants from both pUCHI and pLOP were blue in colour and those from pEST were white. Since pEST was the only plasmid among the three that did not have any *lacZo* sequence it was concluded that the blue colour of the colonies transformed by pUCHI and pLOP was due to the *lacZo* sequence. Thus, the conditions needed to clearly differentiate expressors from non-expressors were established.

It was also noted that although colonies transformed by either pUCHI and pLOP were distinctively blue, they differed qualitatively in intensity of colour. Those transformed by pUCHI were generally a deeper blue than those carrying

Table 3-3. Induction of *lacZ* with Plasmids Containing *lacZo* Sequences

PLASMID	BACTERIAL HOST LE392	E5014
pUCHI	blue eff=0.4	blue eff=65.0
pEST	white eff=0.5	white eff=78.0
pLOP	blue eff=0.2	blue eff=30.0

Both LE392 and E5014 were transformed as described in Materials and Methods. Transformation mixtures were plated on M9-based minimal medium agar containing glucose, Xgal, ampicillin, L-trp and L-met. Under these conditions E5014 colonies appeared after 20 hours while LE392 colonies required longer incubation up to about 40 hours total. The values for eff (efficiency of transformation) equal the no. of transformants (in 10^{-5}) per microgram supercoiled DNA.

pLOP. This difference was presumably due to the plasmid copy number, a factor of particular significance considering that the screening method depended on the ability to titrate out *lac* i molecules. It is known that plasmids derived from the pUC vectors (such as pUCHI) have a higher copy number than those derived from pBR322 (such as pLOP). Hence the degree of discrimination was best when high copy number plasmids were used.

E. Conditions for Optimal Insertion of Linkers

Since the method described above permitted the identification of clones with the desired insert phenotypically marked by *lacZo*, it was now possible to establish optimal conditions for ligating linkers into plasmids. This was done by inserting the mutator DNA with blunt termini into the single *Pvu*II site of pEST and determining the ratio of blue to white colonies. Varying amounts of mutator duplex were added to a fixed amount (0.5 μg) of linearized pEST in a fixed volume, and after ligation were used to transform E5014. The results (**Table 3-4**) show that the

Table 3-4. **Optimization of Linker Ligation**

	AMOUNT OF LINKER (pmole)			
	0.5	1.0	3.5	7.0
Total # blue colonies	56	253	1137	330
Total # transformants	1038	903	1312	419
% blue colonies	5.4%	28%	87%	79%

The plasmid pEST (2364 bp) was linearized at the single *Pvu*II site followed by heat-inactivation of enzyme prior to ligation. In a ligation reaction volume of 20 *u*l, 0.3 pmole of linearized pEST (about 0.5 microgram) was ligated with increasing amounts of linker DNA. All blue colonies were assumed to harbor plasmids with inserts. For each ligation, 5 units of T4 DNA ligase was added and incubated overnight at 15°C.

number of blue colonies per ligation was at a maximum when the molar ratio of linker to vector was about 12:1 (3.5 pmoles of linker to 0.3 pmoles of vector), assuming that all oligonucleotides participated in forming linker DNA. The efficiency of transformation was slightly reduced when the ratio was doubled to 24:1 while the percentage of blue colonies in the total population remained approximately constant. These results suggest that the efficiency of inserting linkers increases with the amount of linker used until a maximum is reached which is presumably limited by the efficiency of circularization of the targeted vector.

CHAPTER IV

INSERTIONAL MUTAGENESIS OF AD5 E1A

A. Target Plasmid for Mutagenesis: Construction of pKH101

The mutagenesis of E1A using the insertional mutator approach described in Chapter III was designed to permit the phenotypic characterization of the mutants from the plasmid level up to virus level. As an initial step, the entire E1A coding region was cloned to facilitate mutagenesis. There were important strategic considerations that determined the approach used in cloning. Firstly, it was shown in Chapter III that targeting of plasmids with a high copy number by the LOP-mutagenesis approach was more efficient in discriminating between blue and colourless colonies than when plasmids of low copy number were used. Secondly, random insertion of the mutator DNA required structural characterization at the level of DNA sequencing, hence the mutants should be readily amenable to such analysis. With this in mind, the logical approach was to sequence by the dideoxy-method using double-stranded DNA which would support bidirectional sequencing from the insert. Since this required relatively high concentrations of template DNA, it was decided that the E1A coding sequence be put in a pUC vector which consistently replicates to high copy number in bacterial hosts. The goal was to use a single crude preparation of plasmid DNA from a particular candidate for complete characterization at the structural level. Thirdly, the mutator DNA contained 2 flanking *Bam*HI sites to facilitate restriction enzyme analysis and to allow generation of shorter insertion mutants by cleavage with *Bam*HI and

religation. This feature was designed for target plasmids lacking *Bam*HI sites but the mutator could, of course, be modified to have alternative restriction sites if desired. Fourthly, the pKH101 vector was designed to permit transfer of the mutated E1A sequence into a larger plasmid containing the E1B region (for transformation assays) and eventually into an intact viral genome (for studies at the viral level). The strategy for constructing a plasmid, called pKH101, that satisfied these requirements is illustrated in **Figure 4-1.** Characterization of the E1A-containing plasmid pKH101 and resulting insertional mutants at the plasmid level included *trans*-activation and enhancer *trans*-repression activities. The mutations could be easily reconstituted with E1B to generate a complete Ad5 E1 region. The resulting plasmids could then be used in a stable cell-focus transformation assays. The construction and screening of the reconstituted E1-containing plasmids was facilitated by restriction enzymes *Eco*RI and *Kpn*I (see McGrory, 1988). Finally, the mutation could be rescued efficiently into intact viruses via homologous recombination (McGrory *et al.*, 1988). These mutant viruses could then be used in various studies to determine their effect on several aspects of viral life cycle.

B. Mutagenesis of pKH101 - Construction of 39-bp Insertion Mutants

Analysis of the E1A coding sequence with the aid of a computer indicated that a relatively large number of mutants could be generated using several multi-cut restriction enzymes using the approach described in **Chapter 3.** Among these were those which generate termini that are either blunt or with 5'CG-overhang. These enzymes are listed in **Table 4-1** along with their recognition specificity, the

Figure 4-1. **Construction of pKH101**

The *Eco*RI-*Kpn*I fragment of pXC38 (see **Appendix H**) that contains the E1A coding sequence was gel-purified and cloned into the corresponding sites of pUC19, utilizing the α-complementation feature of pUC to detect clones with inserts. The resulting plasmid was designated pKH19 (4730 bp) to indicate that the *Kpn*I H-fragment of Ad5 was contained in the plasmid. This plasmid contained the natural *lacZo* sequence, hence subsequent steps were carried out to delete it. pKH19 was linearized with *Bam*HI and the ends were repaired by a Klenow polymerase fill-in reaction. After heat-inactivating the enzymes, a second digest with *Aat*II was carried out to cleave out the Ad5 sequence. This fragment was cloned into a plasmid called pESTLOP (see **APPENDIX D**) which contained a *Bam*HI site replacing the *Pvu*II site in pEST. The vector was first linearized with *Bam*HI and the ends were repaired by Klenow treatment. Then a second cleavage by *Aat*II was carried out, followed by gel purification of the larger fragment. Following ligation of the two fragments, the resulting plasmid called pKH101 with a predicted size of 4574bp was isolated.

Figure 4-1. **Construction of pKH101**

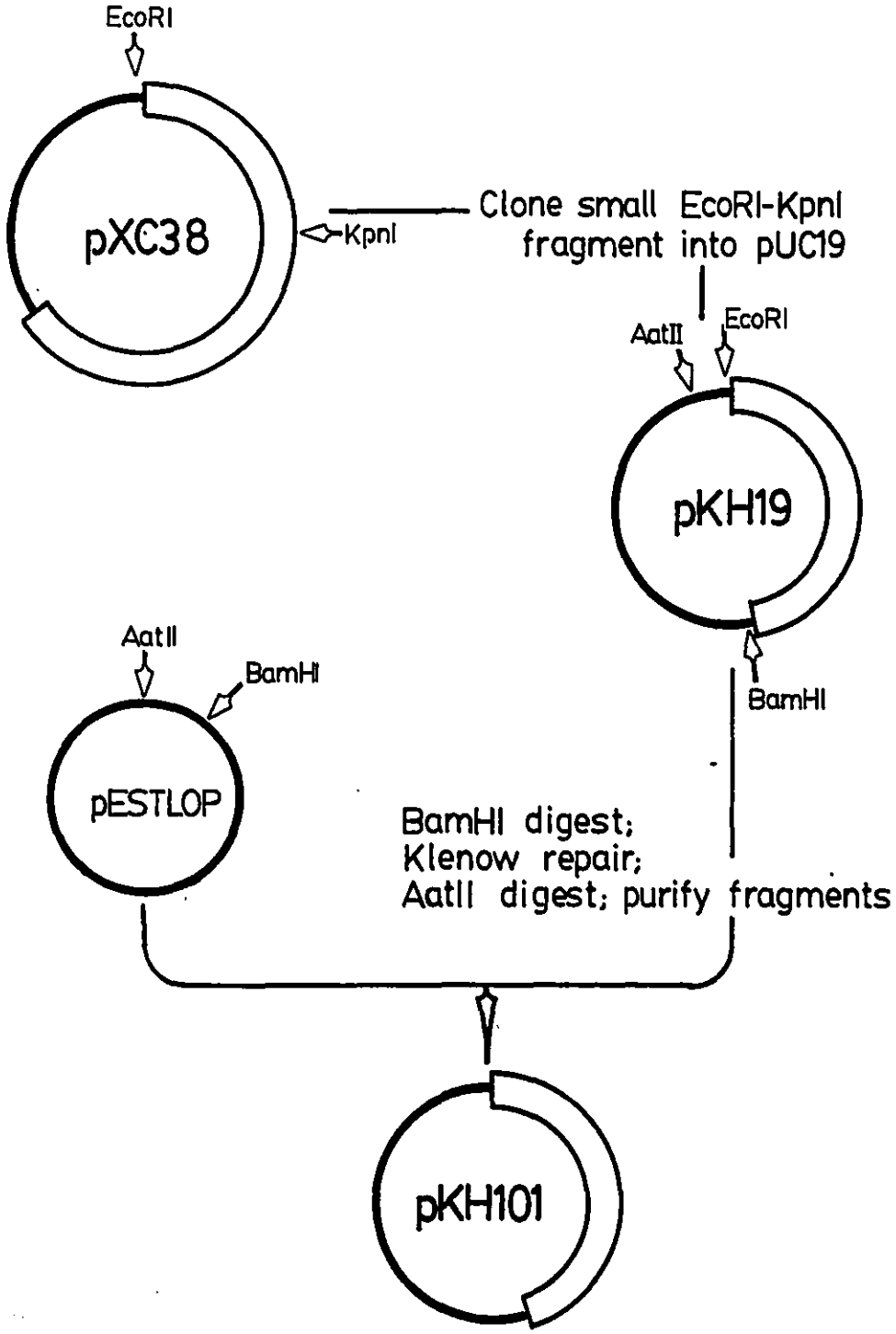


Table 4-1. Restriction Enzymes Used in the Mutagenesis of Ad5 E1A

ENZYME	# SITES IN pKH101	# SITES IN Ad5 SEQ.	# SITES IN 13S
*Alu*I (AG'CT)	21	7	3
*Dpn*I (GA'TC)	21	5	4
*Dra*I (TTT'AAA)	6	2	2
*Hae*III (GG'CC)	18	6	3
*Hin*PI (G'CGC)	24	10	2
*Hpa*II (C'CGG)	30	18	12
*Mnl*I (CCTCnnnnnnn')	38	25	15
*Nla*IV (GGn'nCC)	20	10	9
*Rsa*I (GT'AC)	8	6	3
	186	89	53

These enzymes, as a group, were expected to generate a large number of mutants using 2 mutator cassettes that differ only at the termini; i.e., either with blunt ends or having 5'CG-overhang. Other multi-cut enzymes were also useful except that mutator cassettes with different termini from those mentioned would have been required.

number of times they were expected to cut in the plasmid pKH101 (186), in the Ad5 sequence alone (89), or in the coding sequence for the 289aa-residue E1A product (53). These restriction enzymes were individually used to generate linear pKH101 DNA by partial digestion using the method of Parker *et al.* (1977). For each enzyme, the required incubation time to result in approximately 10% linear form of the total input DNA was established empirically. **Figure 4-2** shows a typical time course using *Hpa*II to linearize pKH101. The result shows that the appearance of the linear form depended on incubation time when concentrations of ethidium bromide and restriction enzyme and when incubation temperature were held constant. After establishing the desired incubation time, a batch-digest was carried out using identical conditions. The digest, which was done separately for each enzyme, was fractionated and the linear form purified using agarose gel electrophoresis as described. The purified linear form was ligated with reconstituted mutator DNA and used to transform the bacterial host E5014. Typically, between 1000-10,000 ampicillin-resistant colonies were obtained from transformation with

Figure 4-2. **Mutagenesis of pKH101**

Mutagenesis of the E1A-containing pKH101 was accomplished by insertion either one of the two mutator cassettes into restriction sites recognized by various multi-cut enzymes. In the presence of ethidium bromide, supercoiled plasmid DNA was linearized using conditions empirically determined for each enzyme. The linear form in the partially digested DNA was initially identified in a separate experiment as the band that had a mobility identical to that of a fully linearized pKH101 in 0.8% agarose gel. In the example shown, 50 μg of supercoiled pKH101 was digested with *Hpa*II in the presence of ethidium bromide (see **Materials and Methods**). Aliquots were removed every 5 minutes, the reaction was stopped by adding a gel loading buffer containing 1% SDS, and run on an agarose gel to determine the extent of reaction. The incubation time that resulted in the generation of approximately 10% linear form was selected, in the example shown 15 minutes, to ensure that molecules with more than one cleavage would comprise no more than approximately 1% of the population. The conditions determined for partial digestion with a specific enzyme were then used to generate linear DNA from a different preparation and purified from agarose gels. Linear DNA generated by various enzymes was ligated with the appropriate mutator cassette (blunt termini or with 5'CG-overhang, **Figure 3-4**) separately. After transformation of *E. coli* E5014 to ampicillin resistance, blue colonies were screened.

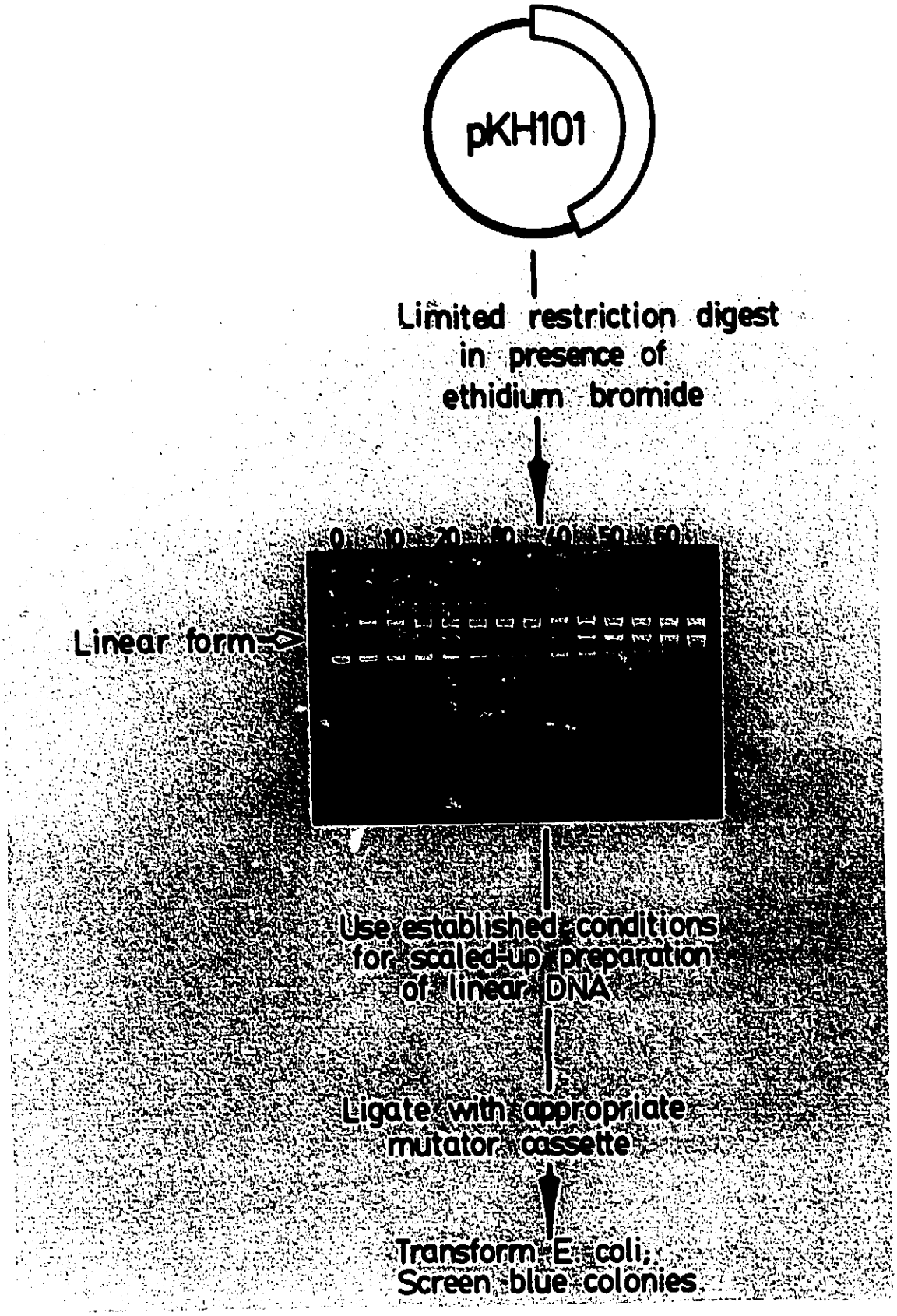

pKH101
Limited restriction digest
in presence of
ethidium bromide
0 10 20 30 40 50 60
Linear form
Use established conditions
for scaled-up preparation
of linear DNA
Ligate with appropriate
mutator cassette
Transform E. coli;
Screen blue colonies

Table 4-2. Efficiency of Insertion of LOP Cassettes into Plasmid DNA

Enzyme	Terminus	% Blue colonies
*Alu*I	blunt	80
*Dpn*I	blunt	3
*Dra*I	blunt	1
*Hae*III	blunt	48
*Hin*PI	5'CG	42
*Hpa*II	5'CG	55
*Mnl*I	blunt?[a]	0.001
*Mnl*I	blunt[b]	85
*Nla*IV	blunt	22
*Pvu*II	blunt[c]	87
*Rsa*I	blunt	8
*Sma*I	blunt	5
DNaseI	variable	0.001
DNaseI	blunt[b]	85

[a]MnlI is reported to generate blunt ends.
[b]Linear DNA was purified and treated with Klenow polymerase and dNTP's.
[c]This enzyme was tested using the plasmid pEST as described in Chapter 3.

ligation mixtures containing 1 μg of pKH101 and the appropriate linker DNA. The relatively low transformation efficiency was probably due to reduced circularization efficiency of agarose-purified vector DNA. Nonetheless, generally, more than adequate numbers of transformants were obtained for further analysis.

C. Insertion of LOP-Cassettes Into Various Restriction Sites

In **Chapter III**, it was shown that the LOP-cassette with blunt termini could be inserted into the *Pvu*II site of the test plasmid pEST with an efficiency of as high as 87%. Those results (**Table 3-3**) suggested that once the conditions for inserting oligonuclotides into plasmids were optimized, the efficiency of insertion could be very high so that the need for a method to easily detect plasmids with inserts might not be pressing. However, this high efficiency was not consistently seen when different enzymes were used. As summarized in **Table 4-2**, the efficiency of insertion of the LOP cassette into pKH101 (measured as the

frequency of occurrence of blue colonies among amp-resistant transformants) that was linearized by various enzymes varied from less than 0.001% to near 100%. Insertion of the LOP cassette into sites specified by enzymes such as *Dpn*I, *Dra*I, *Rsa*I and *Sma*I was consistently inefficient, suggesting that these enzyme preparations contained nucleases. That is, the termini of the restricted DNA could have been altered by contaminating nucleases so that they no longer matched those of the linkers, resulting in inefficient ligation. In the case of *Mnl*I, insertion of the linkers with blunt termini into sites specified by this enzyme (reported to generate blunt termini, Kessler and Holtke, 1986) was extremely inefficient (over 1000-fold lower than *Dpn*I, *Dra*I, *Rsa*I or *Sma*I). After 'repairing' *Mnl*I-digested DNA with Klenow and dNTP's, the efficiency of insertion was raised to near 100%. Recently, Brinkley and Graham (pers. comm.) have determined that *Mnl*I produces termini with a one-base 3'-overhang, which explains the failure to ligate linkers with blunt termini.

Insertion of the LOP cassette whose termini contain 5'-CG overhang into sites specified by *Hpa*II and *Hin*PI were quite efficient and comparable to either *Alu*I or *Pvu*II, both of which produce blunt ends. Since DNA fragments with staggered termini could ligate more efficiently than those that are blunt, the failure to observe a significant increase in inserting the LOP cassettes in DNA fragments generated HinPI or *Hpa*II relative to either *Alu*I or *Pvu*II could be due to the fact that the termini generated by either *Hin*PI or *Hpa*II only contain a 2-base overhang which might not be sufficient to improve the insertion efficiency over that when blunt termini were used. It is interesting then to note that high frequency of insertion did not necessarily depend on whether the termini have blunt or 'sticky'

termini, but more perhaps on whether the DNA fragments to be ligated contained matching termini.

D. Screening of Candidate Mutants

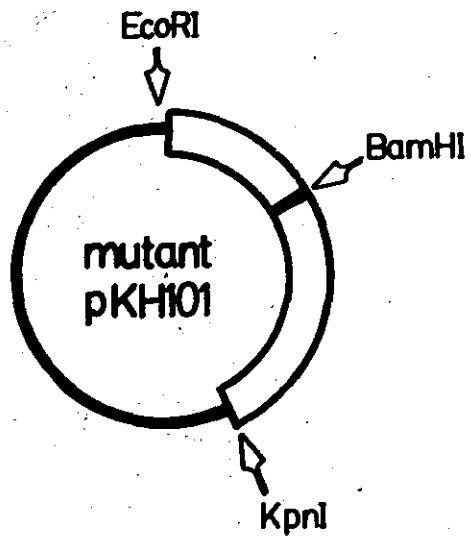
The plasmid DNA preparations from blue colonies were initially analyzed by restriction and gel electrophoresis to establish the location of the insert in the target plasmid. This was relatively straightforward since the insert contained *Bam*HI sites. Thus the disappearance from a gel of a restriction fragment following *Bam*HI digest maps the location of the insert to that fragment. When pKH101 was simultaneously digested with *Eco*RI and *Kpn*I, two fragments appeared on agarose gels corresponding to the Ad5 (2.1kbp) and vector (2.5kbp) sequences, respectively **(Figure 4-3)**. A clone with an insert in the Ad5 fragment was identified when the 2.1kbp fragment disappeared following simultaneous digestion with *Eco*RI, *Kpn*I and *Bam*HI. Thus candidate clones in which the smaller fragment disappeared following *Bam*HI digest were grouped according to restriction pattern, with the assumption that each group represented a unique position of insertion. The grouping was done to avoid the sequencing of more duplicate clones than was necessary to identify all possible insertions.

E. Sequence Analysis of Candidate Mutants

The sequencing of candidate clones was carried out to determine the exact location and orientation of the insert and the structure of the target DNA around it. Clones that were found to have inserts in the Ad5 sequence by restriction mapping were initially sequenced using AB10 as primer **(Figure 4-4)**. The sequencing

Figure 4-3. **Secondary Screening of Candidate Mutants by Restriction Mapping**

Candidate insertion mutants were first identified by their blue color in Xgal selector plates. Crude DNA extracts from 1.5 ml cultures were prepared and subjected to analyses with restriction enzymes and gel electrophoresis. In the example shown, 5 candidate mutants (numbered 1 to 5) were analyzed by double-enzyme digestion with *Eco*RI and *Kpn*I (lane A) and by triple-enzyme digestion with *Eco*RI, *Kpn*I and *Bam*HI (lane B). The first digest generated two fragments: 2.5kbp for the vector and 2.1kbp for the Ad5-containing fragment (Lane A). A detectable change in the mobility of either fragment would indicate that the fragment contained a deletion, hence that particular clone was discarded. Of 400-plus clones analyzed this way, about 5% were found with deletions. The second digest was used to determine if the insert was in Ad5 sequence. Since the insert contained *Bam*HI sites, enzyme digestion with *Eco*RI, *Kpn*I and *Bam*HI resulted in the disappearance of either the 2.5kbp or 2.1kbp fragments depending on the location of the insert. Only those with inserts in the 2.1kbp fragment (Ad5 sequence) were retained for further analysis. In the example shown only clone number 1 contained an insert in Ad5.



Digest A: EcoRI + KpnI
B: BamHI + EcoRI + KpnI

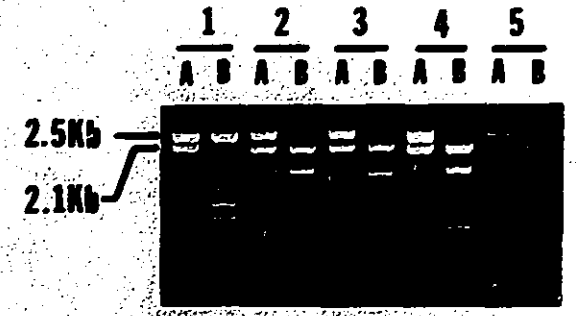
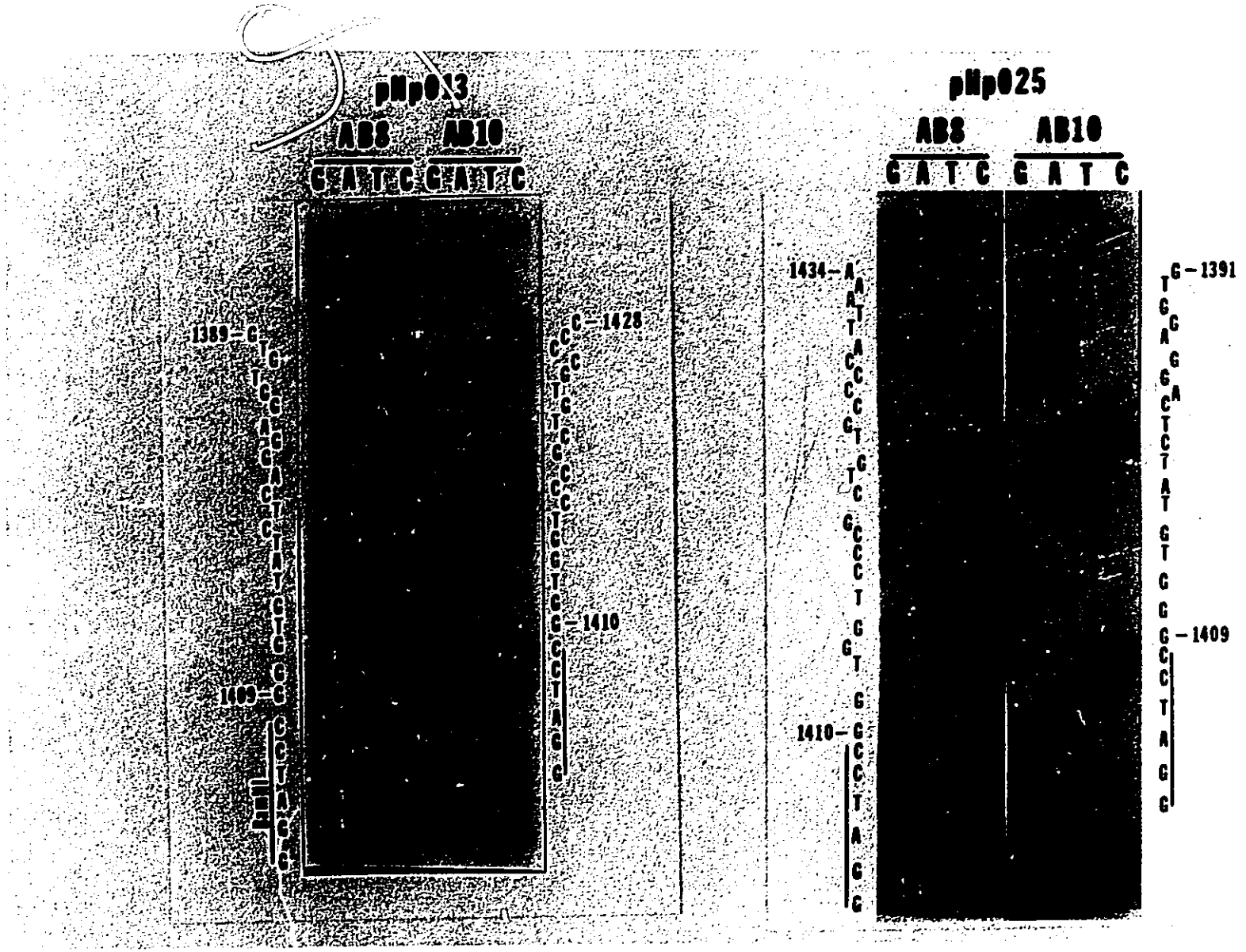


Figure 4-4. **Examples of Sequence Analysis of Candidate Mutants**

Candidate mutants with inserts in the 2.1 kbp *Eco*RI-*Kpn*I fragment were further characterized by sequencing using both AB8 and AB10 as primers. A typical sequencing gel showing the junction between the target and the inserted cassette sequences was used to identify the exact position of the insert. The *Bam*HI sites in the cassette could both be identified when both primers were used. By doing a computer search to determine location of matching sequence in Ad5, the location as well as the orientation of the insert was established unequivocally. In the examples shown, mutants at nt1408 in both orientations were identified: the insert in pHp013 is in the "+" orientation and pHp025 in "-" orientation.


pHp013
AB8
AB10
G A T C G A T C
1389
1409
1428
1410
pHp025
AB8
AB10
G A T C G A T C
1434
1410
G – 1391
G – 1409

ladder produced when this primer was used corresponded to the mRNA sense if the insert was in the "+" orientation. The initial results were confirmed by sequencing in the opposite direction with AB8 as primer. By comparing the sequence data with published Ad5 sequence or with the aid of a microcomputer, the precise structure of a particular mutant was established. Candidate clones in the same restriction group were all sequenced until at least one mutant for each of the two orientations was identified. When only one orientation was isolated, the other was obtained using the procedure described below.

Up to about 120 bases were readable from the sequencing gels. The relatively short readable portion was the result of using higher concentrations of dideoxy-NTP's in the sequencing reaction mixtures to favor the sequencing close to the primer. Thus, when sequenced in both directions from the insert using both primers, greater than 200 bases have been sequenced for a particular mutant. This represents about 10% of the entire Ad5 sequence present in the plasmid. No discrepancies were detected in Ad5 sequences obtained from sequencing of mutants relative to the published sequence suggesting that no alteration has occurred in the viral sequence, except at the site of insertion.

F. Construction of 6-bp Insertion Mutants

For each mutant the 39 bp insert could be collapsed to generate a 6bp insert by *Bam*HI digestion followed by religation. The screening for the smaller inserts on Xgal selector plates was relatively easy since they produce colourless colonies as opposed to blue when the insert remained intact.

G. Construction of Mutants with the Other Orientation

Of 18 the positions in E1A that were mutated (see below), there were a total of 4 sites in which either only a single candidate was found or all the mutants for a particular site were in the same orientation. In such cases, the mutant with the other orientation was isolated by re-inserting a new set of oligonucleotides corresponding to full length mutator DNA. Thus after "collapsing" the insert in the original isolates, a full-length mutator DNA was re-inserted via *Bam*HI complementary ends (**Figure 4-5**). Synthesizing a new set of oligomers as opposed to purifying the insert following restriction digest from one mutant appeared to be a better choice because the latter required more intensive labour than the former to achieve the same results. To preclude multiple insertions, the oligomers were not phosphorylated prior to ligation. Those with inserts produced blue colonies and were picked for sequence analysis as before. Approximately 1/2 of the blue colonies had plasmid with inserts in the desired orientation.

H. Sequencing Across the Inserts

Sequencing with the two core primers did not directly determine the internal structure of the insert *per se* since the primers are themselves components of the insert. To sequence across a selection of inserts, a 21-base long primer was synthesized corresponding to nucleotides 779 to 799 of the Ad5 sequence. Because of their proximity to this primer, several mutants (at nt812, nt819 and nt827) were sequenced across the entire insertions (**Figure 4-6**). Of a total of 14 mutants in these three positions sequenced, all had inserts with predicted structures.

Figure 4-5. **Construction of Mutants with the Other Orientation**

When after the initial screening of candidate mutants only one of the two possible orientations was found for a particular site, the other mutant was obtained by re-inserting a full length DNA duplex after cleaving the original insert by *Bam*HI digestion. Removal of the insert by *Bam*HI cleavage "collapses" the insert from 39bp to 6bp which could be identified following bacterial transformation since the colonies were colorless. The blue colonies present from the same transformation contained plasmids in which the cleaved insert religated, assuming *Bam*HI digestion was complete. However, in such a case one cannot be absolutely certain that several copies have not been inserted. This is not unlikely because the cleaved insert can ligate together to form concatemers. Thus a full length insert with corresponding length and sequence was separately synthesized and was reinserted into the *Bam*HI site of a collapsed mutant without phosphorylating the oligonucleotides to preclude multiple insertion. The two oligonucleotides that made up the full-length cassette were AB238 and AB239 (**Table 2-2**). Those with inserts were again selected on the basis of phenotypic labeling with *lacZo* and sequenced directly to screen for those with the desired orientation. The sequence of the re-inserted full-length cassette was determined for pOT0812 which was derived from pKP0812, the collapsed form of pHp083. An example of this is shown below. The procedure for sequencing across the insert is described in the following section.

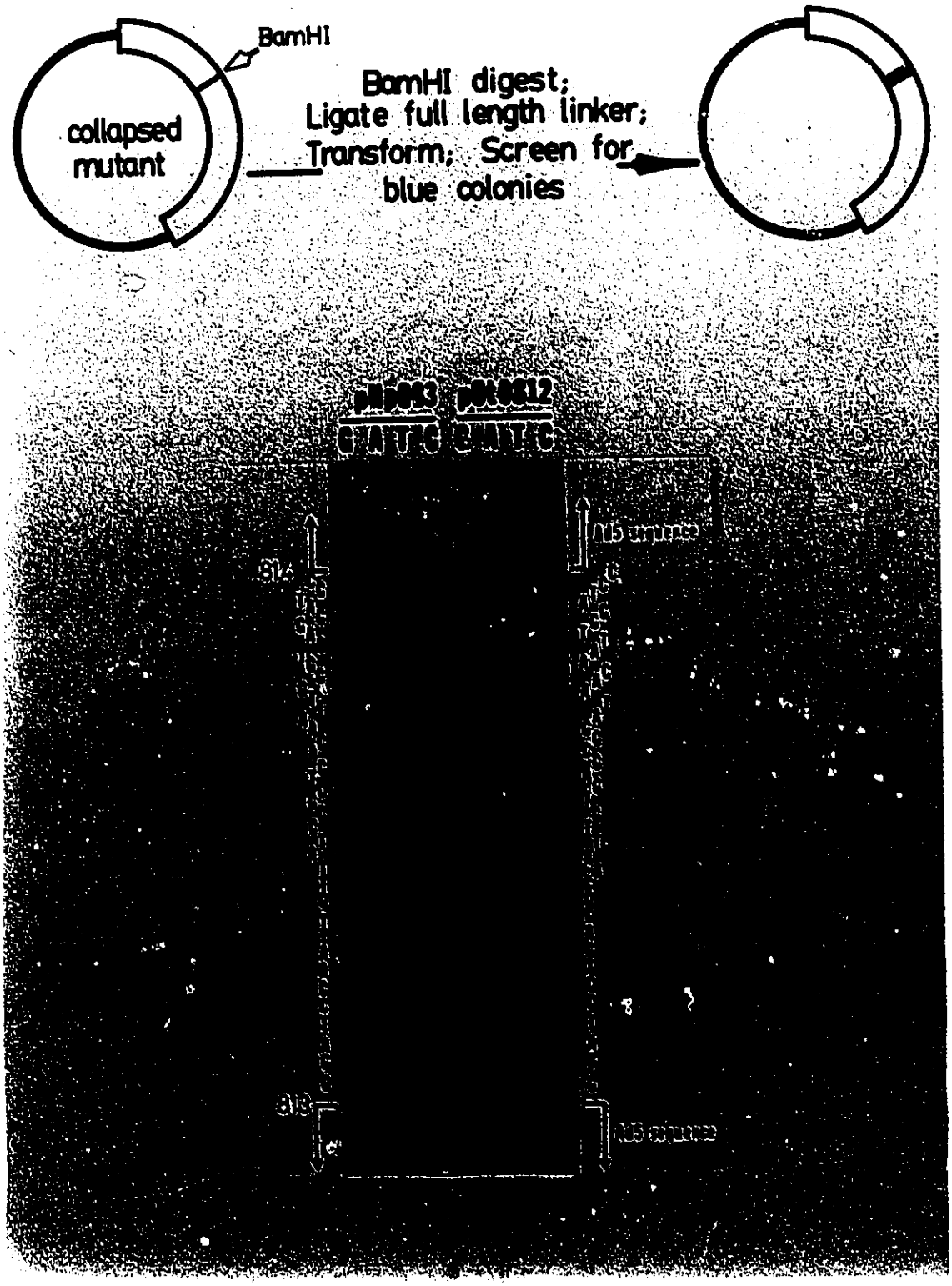
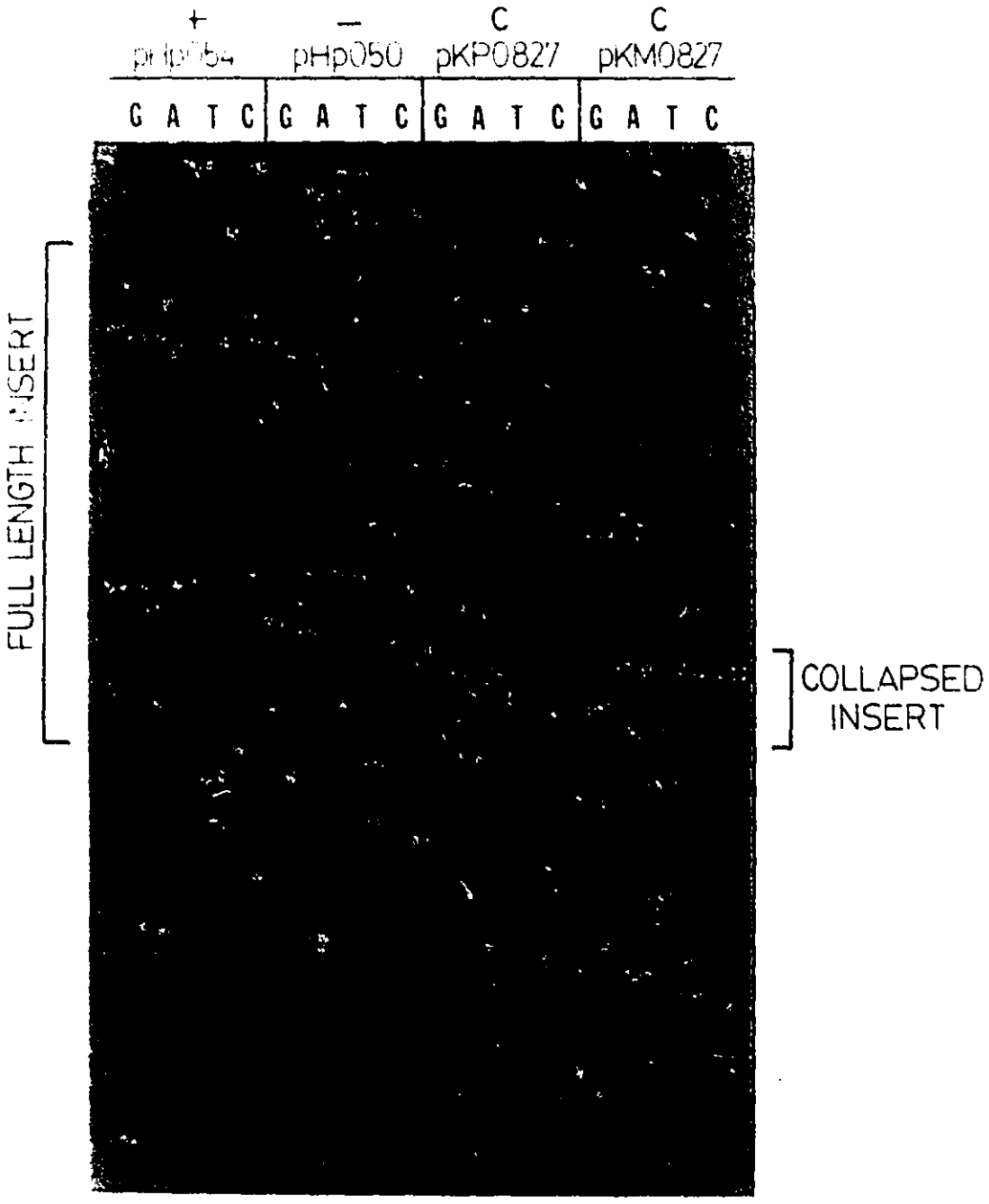

BamHI
collapsed mutant
BamHI digest;
Ligate full length linker;
Transform; Screen for
blue colonies

Figure 4-6. **Examples of Sequence Analysis of Mutants Across Insertions**

The primer AB289 (**Table 2-2**) corresponds to nt779 to nt799 of the r-strand of Ad5 so that sequencing with it would proceed in the rightward direction. Mutants at positions nt812, nt819, nt827 and nt863, a total of 14, were resequenced using this primer to determine the structure of the individual insert. None of these deviated from what was expected. The sequence of mutants at nt827 are illustrated with their inserts as marked (full length or collapsed).


+
pHp054
–
pHp050
C
pKP0827
C
pKM0827
G A T C G A T C G A T C G A T C
FULL LENGTH INSERT
COLLAPSED INSERT

I. Analysis of Mutants with *Hpa*II Restriction Enzyme and PAGE

A fine-structure analysis using the *Hpa*II restriction enzyme and 6% PAGE was carried out on the mutants. This analysis was done to determine whether the insertion mutants have identical structure as wild-type pKH101 except those changes that could be attributed to the insert. Structural alterations that could be detected by this type of analysis include deletions that might have been acquired during cloning, more importantly in regions outside those that have been sequenced. With the approach used in constructing and screening the insertion mutants, it would have been highly unlikely to both generate and not detect deletions at the same time. Nonetheless, the mutants were subjected to this structural analysis as a precautionary measure.

Figure 4-7 shows the *Hpa*II restriction pattern of both wild-type pKH101 and mutants in 6% PAGE. All mutants generated by *Hpa*II (lanes B to L,N,O) appeared identical to pKH101 (lanes A and M) since their inserts could be cleaved out by this enzyme. All other mutants, however, have inserts that could not be cleaved out by *Hpa*II digestion, hence their restriction patterns would differ from that of wild-type pKH101 by a shift in mobility of one of the bands in the mutants due to the insert. Thus insertions at nt1415 (lane P) and nt1523 (lane Q) are found in a 561-bp fragment and the gel showed the appearance, in both cases, of a band migrating slower than the 561-bp band (which comigrated with a 501-bp fragment), which is consistent with the notion that the fragment contained an insert. The insert at nt1304 (lane R) was expected to affect the 109-bp fragment (which comigrated with a 110-bp fragment) increasing its size to 148bp and therefore shifting its location slightly above the 147-bp band on the gel. Lane S shows the pattern for a

Figure 4-7. **PAGE Analysis of Mutants With *Hpa*II Enzyme**

Representative mutants at various positions in E1A were digested (600 ng) with *Hpa*II and analyzed in 6% PAGE. The restriction pattern of the wild-type plasmid, pKH101, is shown in lane A on the top gel and M on the bottom gel. The diagram on the right of the top gel shows the *Hpa*II restriction pattern of pKH101 as predicted by the linear relationship between log(MW) and R_f. Note that the restriction pattern on the gel does not follow the predicted pattern. Assignment of sizes of the bands are therefore not absolute.

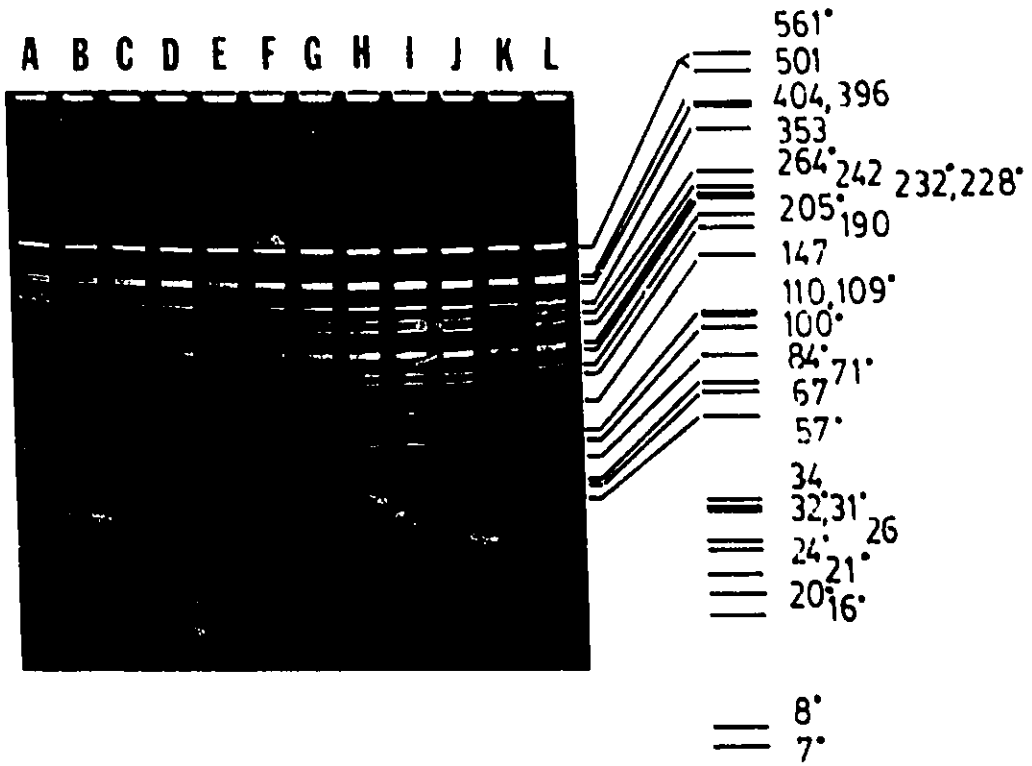
Legend:

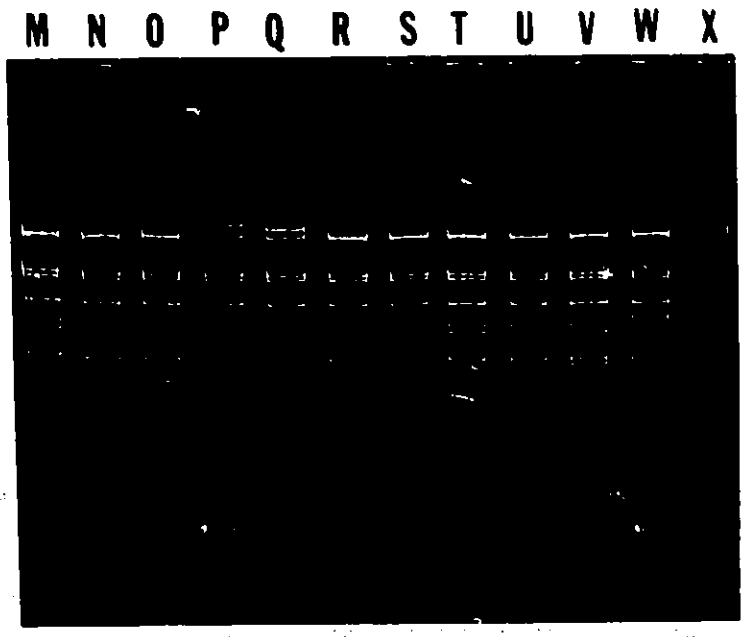
Top Gel:

A. pKH101	wild type
B. pHp083	812+
C. pOt0812	812-
D. pKP0812	812c
E. pHp134	819+
F. pHp144	819-
G. pKP0819	819c
H. pKM0819	819c
I. pHp054	827+
J. pHp050	827-
K. pKP0827	827c
L. pKM0827	827c

Bottom Gel:

M. pKH101	wild type
N. pKM1039	1039c
O. pKM1008	1008c
P. pAlu009	1523+
Q. pNl025	1415-
R. pNla140	1304+
S. pNla142	1056+
T. pRs008	906+
U. pNla136	882+
V. pNl033	882-
W. pNl032	717-
X. pNl020	717+


A B C D E F G H I J K L
561*
501
404, 396
353
264*
242
232*, 228*
205*
190
147
110, 109*
100*
84*
71*
67
57*
34
32*, 31*
26
24*
21*
20*
16*
8*
7*


M N O P Q R S T U V W X

mutant at nt1056 which was expected to (and did) alter the mobility of the 228-bp fragment. Insertions at nt906 (lane T) and at nt882 (lanes U and V) were expected in fragments that are 24- and 21-bp, respectively, neither of which could be resolved in the gel. Finally, insertions at nt717 (lanes W and X) were expected to affect a 264-bp fragment and change its location. However, the results showed that the affected band had an assigned size of 242bp. This inconsistency is probably due to anomalous migration of the 264-bp fragment (migrating faster than it should) and not to structural alterations since the two independent isolates, pNl032 (lane W) and pNl020 (lane X), showed identical restriction pattern. All fragments that migrated differently from wild-type pattern were therefore accounted for. Taken together, these results provide further evidence that the mutants have no sequence alterations anywhere else in the plasmid aside from those that could be explained by the inserts.

J. Structural Features of Mutants

Table 4-3 lists all the mutants that were isolated and mapped on the Ad5 sequence in the plasmid pKH101. The position of insertion is the nucleotide number that precedes the actual inserted piece of DNA. References to these mutants is according to insertion site (e.g., nt812 means insertion between nucleotides 812 and 813), orientation (+ or -) or collapsed (c). **Figure 4-8** shows the position of the mutated sites in the *Kpn*I-H fragment of Ad5 relative to the products of both E1A and E1B. Mutants in nt188, nt420 and nt477, all of which were generated by *Hpa*II, are located in the regulatory sequences of E1A. Position nt548, also generated by *Hpa*II, is in the untranslated leader sequence of the E1A mRNAs. Position nt717 which was generated by *Nla*IV and which produced the

Table 4-3. **List of Insertion Mutants in the Left 2kbp of Ad5**

The 25 mutated sites in the Ad5 KpnI H-fragment are listed according to nucleotide position in the Ad5 sequence. These mutants were classified as to the type of insertion: "+" for the orientation in which all translational reading frames are open; "-" for closed reading frame orientation; and "c" for the collapsed version. When both orientations for a particular site were independently isolated, both were used to generate the collapsed mutant for that site. The underlined sequences are those of the collapsed mutant for that site; i.e., either blunt-ended or with a 5'CG-overhang. The translational reading frame (R.F.) of the inserts are as designated in **Figure 3-4**, and the predicted modification caused by the collapsed insertion in both DNA and amino acid sequences are shown under "MODIFICATION"; the exact sequence of the larger insert can be easily derived by substituting the smaller insert as shown in the Table with the larger insert in the appropriate reading frame. Note that the insertions at nt1056 and nt1523 resulted in amino acid substitutions: D_{166} to E_{166} for mutants at nt1056, and S_{283} to R_{283} for mutants at nt1523.

Table 4-3. List of Insertion Mutants in the Left 2kbp of Ad5

POSITION	MUTANT	TYPE	R.F.	MODIFICATION
188	pHp042 pKM0188	- c		GTGCGCCGGATCCGGTGT CACGCGGCCTAGGCCACA
420	pHp058 pKP0420	+ c		GCGTTCCGGATCCGGGTC CGCAAGGCCTAGGCCCAG
477	pHp088 pKM0477	- c		TATACCCGGATCCGGTGA ATATGGGCCTAGGCCACT
548	pHp070 pOt0548 pKM0548	+ - c		CGACACCGGATCCGGGAC GCTGTGGCCTAGGCCCTG
717	pNL020 pNL032 pKP0717 pKM0717	+ - c c	II	52 53 54 T A D P P ACG GCG GAT CCC CCC TGC CGC CTA GGG GGG
812	pHp083 pOt0812 pKP0812	+ - c	II	84 85 86 P P D P A CCG CCG GAT CCG GCG GGC GGC CTA GGC CGC
819	pHp134 pHp144 pKP0819 pKM0819	+ - c c	I	86 87 88 A P G S G GCG CCC GGA TCC GGT CGC GGG CCT AGG CCA
827	pHp054 pHp050 pKP0827 pKM0827	+ - c c	II	89 90 91 S P D P E TCT CCG GAT CCG GAG AGA GGC CTA GGC CTC
863	pOt0863 pHp043 pKM0863	+ - c	II	101 102 103 Q P D P E CAG CCG GAT CCG GAG GTC GGC CTA GGC CTC

Table 4-3. List of Insertion Mutants in the Left 2kbp of Ad5

POSITION	MUTANT	TYPE	R.F.	MODIFICATION
882	pNla136	+	II	107 108 109
	pNl033	-		L G D P P
	pKP0882	c		TTG GGG GAT CCT CCG
	pKM0882	c		AAC CCC CTA GGA GGC
884	pHp057	+	II	108 109 110
	pHp009	-		G P D P V
	pKP0884	c		GGT CCG GAT CCG GTT
	pKM0884	c		CCA GGC CTA GGC CAA
906	pRs008	+	II	115 116 117
	pOt0906	-		L V D P P
	pKP0906	c		CTT GTG GAT CCA CCG
				GAA CAC CTA GGT GGC
908	pHp038	+	II	116 117 118
	pHp018	-		V P D P E
	pKP0908	c		GTA CCG GAT CCG GAG
	pKM0908	c		CAT GGC CTA GGC CTC
1008	pOt1008	+	I	149 150 151
	pHp084	-		H P G S G
	pKM1008	c		CAC CCC GGA TCC GGG
				GTG GGG CCT AGG CCC
1039	pHp104	+	III	159 160 161
	pHp035	-		Y H R I R
	pKP1039	c		TAT CAC CGG ATC CGG
	pKM1039	c		ATA GTG GCC TAG GCC
1056	pNla142	+	II	165 166 167
	pNl050	-		G E D P P
	pKP1056	c		GGG GAG GAT CCC CCA
	pKM1056	c		CCC CTC CTA GGG GGT
1267	pHp103	+	II	197 198 199
	pHp094	-		E P D P E
	pKP1267	c		GAA CCG GAT CCG GAG
	pKM1267	c		CTT GGC CTA GGC CTC

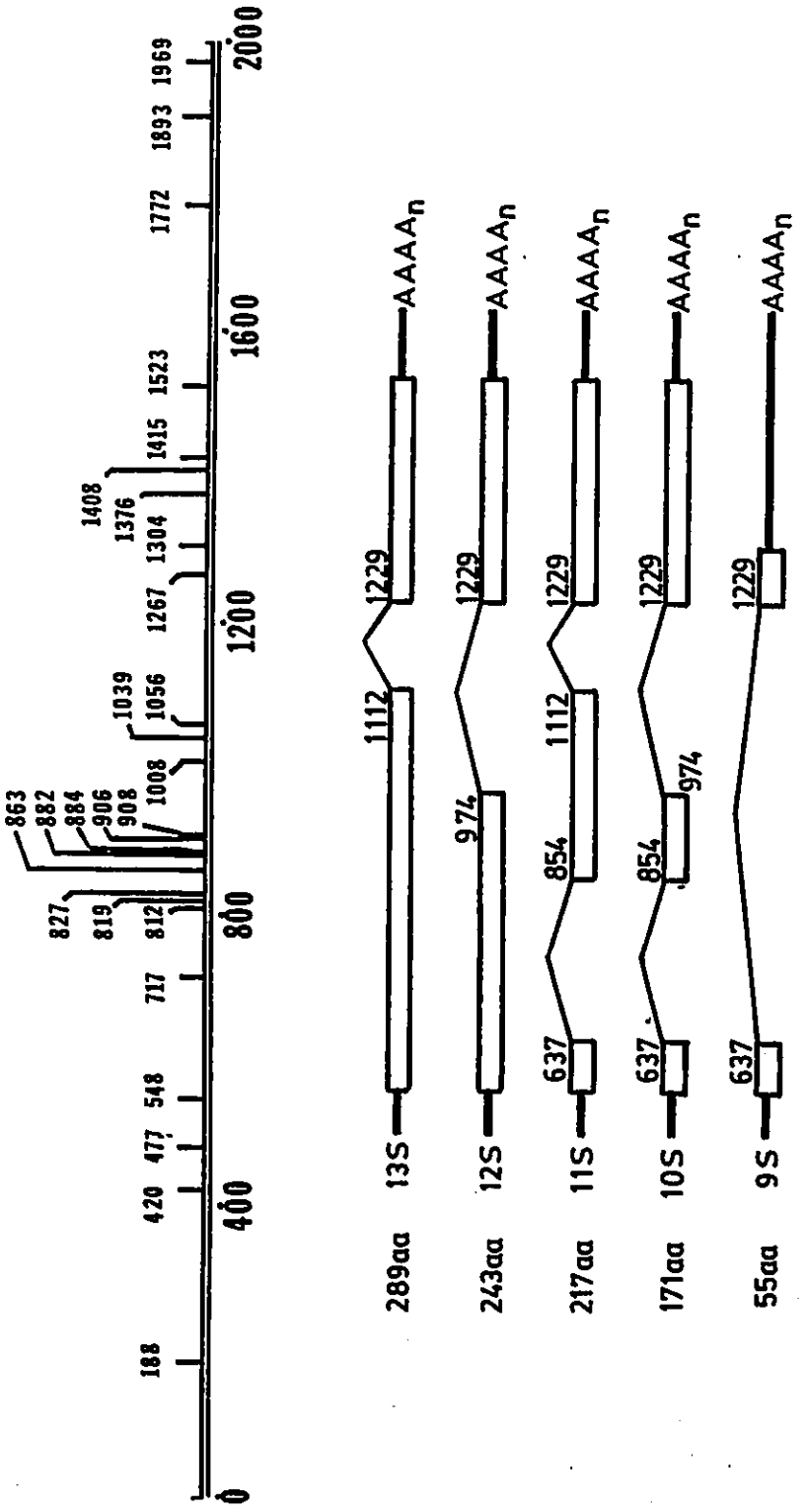
Table 4-3. List of Insertion Mutants in the Left 2kbp of Ad5

POSITION	MUTANT	TYPE	R.F.	MODIFICATION
1304	pN1a140 pOt1304 pKP1304	+ - c	II	209 210 211 M A D P P ATG GCG GAT CCG CCT TAC CGC CTA GGC GGA
1376	pHp121 pHp112 pKP1376 pKM1376	+ - c c	I	233 234 235 D S G S G GAC TCC GGA TCC GGT CTG AGG CCT AGG CCA
1408	pHp013 pHp025 pKP1408 pKM1408	+ - c c	II	244 245 246 H P D P V CAC CCG GAT CCG GTG GTG GGC CTA GGC CAC
1415	pOt1415 pN1025 pKM1415	+ - c	II	246 247 248 V V D P P GTG GTG GAT CCC CCG CAC CAC CTA GGG GGC
1523	pAlu009 pOt1523 pKP1523	+ - c	II	282 283 284 L R D P C TTG AGG GAT CCC TGT AAC TCC CTA GGG ACA
1772	pA1001 pKP1772	- c	II	19 20 21 Q R D P S CAG AGG GAT CCC TCT GTC TCC CTA GGG AGA
1893	pA1008 pKM1893	- c	I	59 60 61 G E G S L GGT GAG GGA TCC CTG CCA CTC CCT AGG GAC
1969	pHp003 pKM1969	- c	II	85 86 87 T P D P G ACA CCG GAT CCG GGG TGT GGC CTA GGC CCC

Figure 4-8. **Location of Insertion Mutants in the Left 2kbp of Ad5**

The mutated sites in the KpnI-H fragment of Ad5 are shown in the diagram. The E1A products are illustrated below the linear map and the relative positions of their coding sequence are as indicated.

Figure 4-8. Location of Insertion Mutants in the Left 2kbp of Ad5



insertions between amino acid residues 53 and 54, maps in conserved region 1 (residues 40 to 80) of the 13S and 12S products. Mutants in positions nt812, nt819, nt827 and nt863, all of which were generated by *Hpa*II, map between conserved regions 1 and 2. Four positions in conserved region 2 (residues 108 to 136) were mutated, namely, nt882, nt884, nt906 and nt908. These were generated by *Nla*VI (nt882), *Rsa*I (nt906), and *Hpa*II (nt884 and nt908). Three positions (nt1008,nt1039 and nt1056) in the unique region (residues 140 to 184) were also mutated, all of which were generated by *Hpa*II. Six positions (nt1267, nt1304, nt1376, nt1408, nt1415 and nt1523) in exon 2, which is found in all E1A mRNAs except 9S, have been mutated. Such mutants were generated by *Hpa*II (nt1267, nt1376, nt1408), *Nla*IV (nt1415), and *Alu*I (nt1523), the latter being in a short stretch of highly conserved residues (279 to 283) near the COOH-terminus.

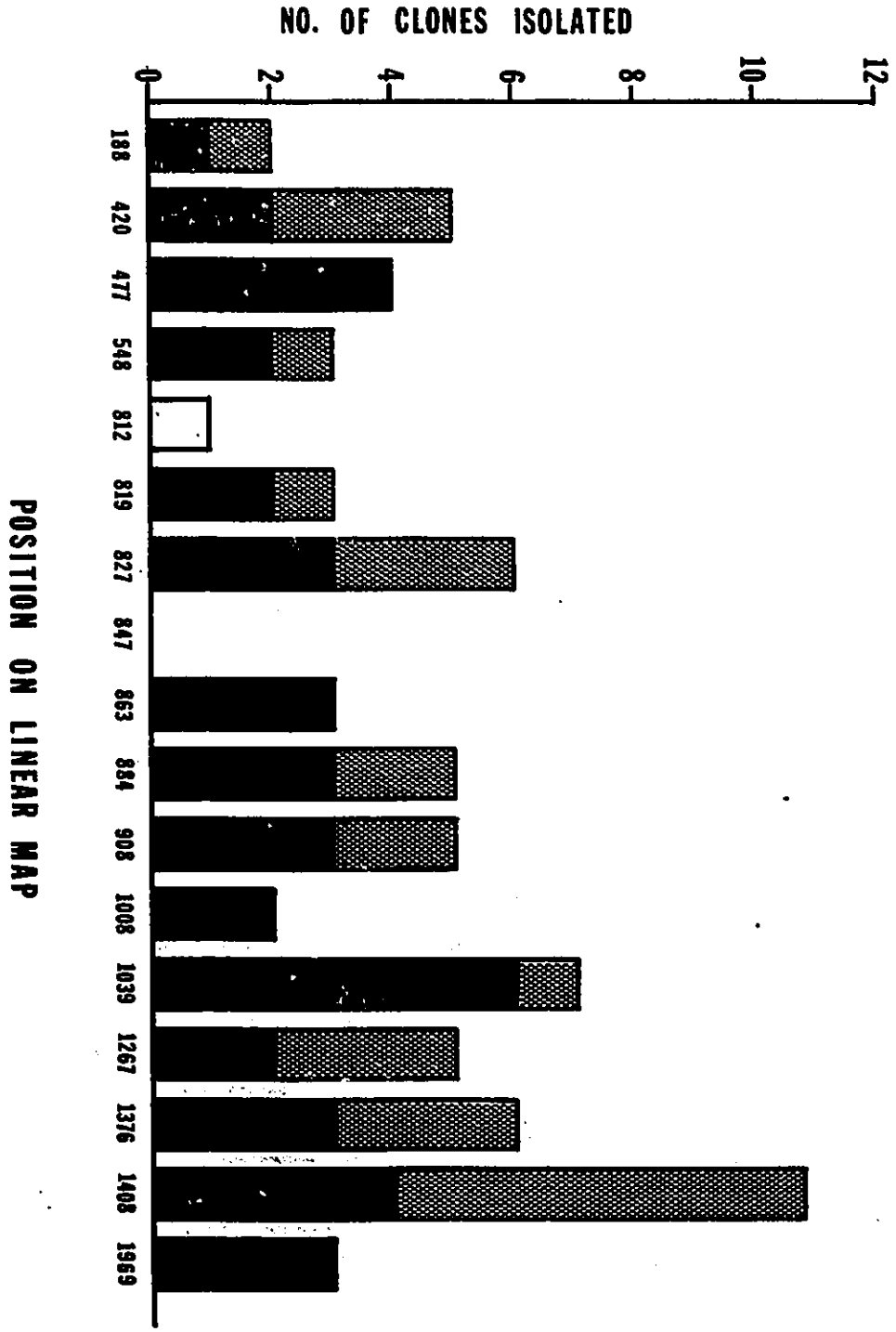
K. Frequency of Insertion in *Hpa*II Sites

A total of 71 mutants generated by *Hpa*II were identified and sequenced fully to establish the precise insertion site and orientation of the LOP-cassette in the 2.1kb Ad5 fragment of pKH101. This was a considerable sample size from which one can determine the frequency at which a multi-cut enzyme cleaves a target DNA segment having several available sites. It was assumed that the frequency of insertion at a particular site is equivalent to the frequency of cleavage by the enzyme at that particular site. The results (**Figure 4-9**) indicate that the distribution of insertion mutations was relatively uniform except for two sites, one (nt847) where insertion was never observed, and another (nt1408) where the frequency of insertion was apparently higher than average. The absence of

Figure 4-9. **Distribution of Mutants at *Hpa*II sites in Ad5 DNA of pKH101**

The numbers on the x-axis refer to the position of the *Hpa*II sites in the left 2.1kb of Ad5 DNA. The numbers of mutants isolated at each available *Hpa*II site are indicated as stippled bars for inserts in the open orientation and as solid bars for closed orientation mutants. The statistic X^2 for the distribution of the mutants (with Yate's correction) was 18.6, which was less than the critical value at 95% significance level with 17 degrees of freedom (26.3), thus indicating that the observed pattern of insertion was not significantly different from a uniform distribution. But since the 18.6 was greater than the critical value at the 5% significance level with 17 degrees of freedom (7.96) then the agreement was not exceptionally good either.

Figure 4-9. **Distribution of Mutants at HpaII sites in Ad5 DNA of pKH101**



insertions at nt847 could not be explained by a mutational change in the target DNA, since sequence analysis confirmed the existence of this particular *Hpa*II site. It is possible, however, that further screening of candidate mutants would have allowed us to isolate one at nt847, since X^2 analysis indicated that the pattern of insertions seen in **Figure 4-9** did not differ significantly from a uniform distribution.

CHAPTER V

DEVELOPMENT OF AN ASSAY FOR E1A REGULATORY FUNCTIONS

A. Transient Expression Assays for E1A Regulatory Functions

Transient expression in established human cell lines of transfected DNA is commonly used to study the transcriptional regulatory activity of E1A proteins. As a regulator of transcription, E1A activates both viral and several cellular promoters but also represses some through its ability to reduce the activity of various enhancers. In transient expression assays, the E1A region or its mutant variant is typically present in a bacterial plasmid, and its regulatory roles in expression are assessed by the ability of the plasmid to activate or repress the expression of a reporter gene which is contained in a separate plasmid. The reporter gene is often a chimera between a potentially E1A-responsive promoter and the coding region for an enzyme for which an enzymatic assay is available.

The most commonly used reporter gene to assay for E1A *trans*-activation function consists of a fusion between the E3 promoter and the bacterial gene coding for chloramphenicol acetyl transferase (CAT) (Weeks and Jones, 1983). This assay provided an easier alternative for testing this particular E1A function compared to the more tedious, but perhaps more appropriate, direct measurement of mRNA from E1A-responsive genes by Northern blot, S1 or primer-extension analyses. In the so-called CAT assay (Gorman *et al.*, 1982), lysates from

transfected cells are prepared and directly assayed for CAT activity without purifying the enzyme. The enzymatic assay consists of the catalytic acetylation of ^{14}C-chloramphenicol. The reaction mixture is fractionated by thin layer chromatography (TLC) which separates the acetylated and unacetylated forms. Auto-radiography of the chromatogram allows visualization of the spots which can then be scraped and quantitated by a liquid scintillation counter. The level of CAT activity has been shown to directly correspond to the level of CAT-specific mRNA and, therefore, the enzymatic activity is a good indicator of the level of transcription from the E3 promoter, which is responsive to E1A stimulation.

The CAT assay in this form requires several steps to quantitate the level of enzyme activity, and the process includes many steps in which errors may be incurred. A significant modification of the technique (Nordeen *et al.*, 1984) has led to less experimental manipulation of the material so that the possibility of introducing variation is reduced. The modified version does not involve running TLC, locating the spots by chromatography and scraping them to be quantitated. Rather, ^{3}H-labeled acetyl coenzyme A is used as the donor of the acetyl group to cold chloramphenicol. Acetylated chloramphenicol (which is now ^{3}H-labeled) can be separated from the other labeled compound by virtue of its solubility in benzene. Recovery of the labeled reaction product is by drying in vacuum after which the radioactivity is determined by a liquid scintillation counter, thus by-passing the fractionation step involved in the original version. Despite the improvement, drying the material before quantitation is usually long (12 to 16 hours) depending on the number of samples and performance quality of the desiccator.

In this project, it was anticipated that the amount of work required to functionally characterize the relatively large number of mutants which have been generated could be considerable as repetitions were required in order to obtain accurate values. This would involve an assay that is at least as accurate and reliable but considerably faster than the CAT assay was desirable. Hence, a colourimetric assay based on the bacterial *lacZ* gene which has been extensively used by investigators of prokaryotic gene expression and recently of eukaryotes, was developed. This assay has several features that make it attractive for use in transient expression assays. One, it is simple in that allows direct quantitation of the reaction product without subsequent manipulation, resulting in a significant reduction in time and effort. Two, it allows the user to visually monitor the reaction as it progresses, hence incubation time may be shortened or extended as the need arises. Furthermore, fusion protein products with *lacZ* can be made without apparent change in enzyme activity, hence construction of reporter plasmids could be facilitated in cases where available restriction sites do not allow simple promoter/reporter gene fusion. These features and the fact that the assay requires fewer and uncomplicated steps in the absence of radio-labeled reagents provided the impetus for the development of a colourimetric assay for E1A regulatory functions based on the bacterial *lacZ* gene.

One objective of this study was to assess the regulatory functions of E1A on several Ad5 early promoters and to determine whether a common mechanism is involved in their regulation in spite of sequence diversity of such promoters. As noted elsewhere, the reported studies on E1A regulatory functions were not extended beyond testing one or two viral promoters so that there was no clear

evidence to suggest that E1A responsive promoters require the same domain(s). The establishment of this fact is crucial to the understanding of the mechanism for *trans*-activation. Thus, an objective of the studies described in the next chapter on E1A transcriptional regulatory functions was to include assays on more than one promoter.

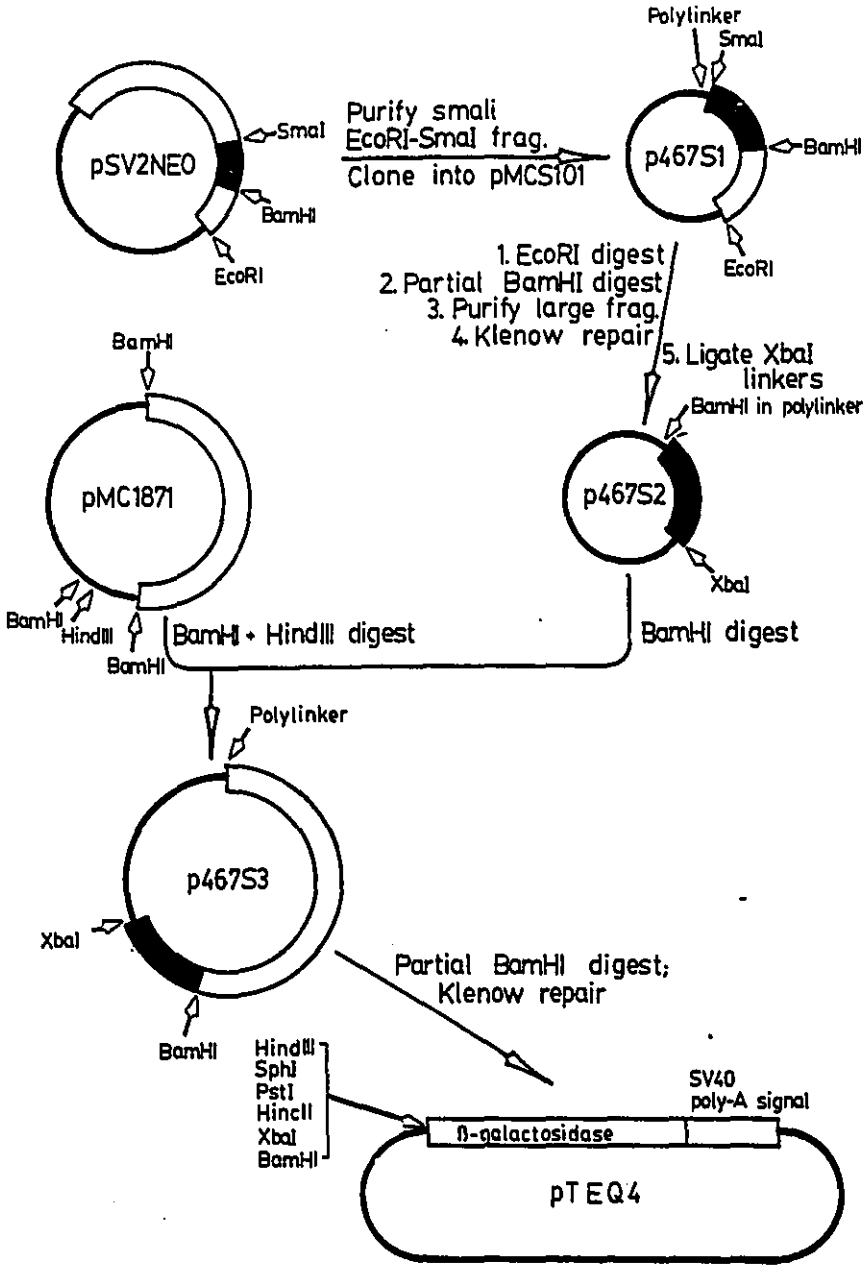
B. Construction of pTEQ4 for Expression of *lac*Z in Mammalian Systems

An expression capsule containing the coding region of *lac*Z was constructed (**Figure 5-1**) for the purpose of assessing promoter activity in transient expression assays. This expression capsule, called pTEQ4 (6259bp), was intended to accept various promoters for expression in transfection assays. It contained 4 unique restriction sites (*Hin*dIII, *Sph*I, *Sal*I and *Bam*HI) in front of the 10th amino-acid residue of *lacZ*. It does not, however, contain a translation initiation codon so that must be provided either as a component of the heterologous promoter or be added by introduction of a synthetic oligonucleotide. Several properties of the plasmid make it a desirable test gene for transfection studies of various promoters. One, by virtue of several restriction sites in front of the *lac*Z coding sequence, the insertion, with proper orientation of restriction fragments containing promoters could be achieved relatively easily, usually through simple cutting and ligating of restriction fragments. Second, it contains the polyadenylation site of SV40 for proper expression and transport into the cytoplasm for translation. Third, the plasmid is over-amplified in bacteria so that a relatively good yield of DNA can be obtained from a single extraction. The complete restriction map and the assembled sequence of the plasmid is shown in the **APPENDIX.**

Figure 5-1. **Construction of pTEQ4**

The first step involved the cloning of the SV40 poly-A signal from the plasmid pSV2NEO into a plasmid named pMCS101 which was essentially pUC19 except that the *lacZo* sequence was removed (see the **Appendix**). The SV40 poly-A signal is contained in a restriction fragment bounded by *Eco*RI and *Sma*I sites, the latter being at the coding sequence for the C-terminus of the neo gene. This fragment was gel-purified and cloned into pMCS101 at the corresponding *Eco*RI and *Sma*I sites to produce a plasmid called p467S1. The plasmid p467S1 was further reduced in size by removing the small *Bam*HI-*Eco*RI fragment downstream from the poly-A signal. This was achieved by first linearizing the plasmid with *Eco*RI then partially digesting with *Bam*HI, which resulted in 2 fragments, the larger of which was gel-purified. This fragment was blunt-ended with Klenow polymerase and dNTP's then *Xba*I linkers were introduced using the *lacZo*-labeled cassette method for inserting oligomers. Oligomers AB6, AB7, AB9, in addition to core oligomers AB8 and AB10, were used to make the *Xba*I-cassette. The resulting plasmid, p467S2, contains the SV40 poly-A signal bounded by an *Xba*I site at its 3' end and multiple cloning sites at the 5' end. The next step was the fusion of the coding sequence of lac Z in front of the SV40 poly-A signal and in the proper orientation. The gene was contained in the plasmid pMC1871 (Casadaban *et al.*, 1978) which was doubly digested with *Bam*HI to yield the *lac* Z gene and the vector DNA and *Hin*dIII to further cleave the latter to prevent its religation. This was mixed with *Bam*HI-linearized and phosphatase-treated p467S2 and ligated to produce a plasmid called p467S3A. Finally, the second *Bam*HI site at the end of the lac Z coding sequence was removed to produce the plasmid called pTEQ4. The sequence of the cloning sites in front of *lac* Z is shown below the diagram.

Figure 5-1. **Construction of pTEQ4**



1 10 20 30 40

A AGC TTG CAT GCC TGC AGG TCG ACT CTA GAG GAT CCC GTC LAC Z

HindIII SphI PstI HincII XbaI BamHI Val

C. Expression of *lac*Z Under Control by Ad5 Early Promoters

1. Construction of pE3*lac*Z

To assess whether the *lac*Z gene can be used as a reporter gene in a mammalian system, it was decided that the Ad5 E3 promoter be used as a test promoter since it appeared to be most responsive to E1A stimulation amongst the early promoters. The actual construction of the E3 and *lac*Z fusion plasmid, called pE3*lac*Z, was relatively simple as revealed by inspecting the DNA sequence for several convenient cloning sites around the E3 promoter site. In this construct, the translation initiation codon for *lac*Z was introduced using synthetic oligomers using the LOP-labeled cassette method. The site that was chosen for initiating translation resulted in a good match for the consensus sequence surrounding the initiation codon (Kozak, 1986). The strategy for cloning that was adopted is shown in **Figure 5-2.** In experiments similar to those described in the following section, the pE3*lac*Z reporter plasmid was tested for its ability to respond to E1A stimulation in co-transfection assays, and was found to be a sensitive assay for E1A *trans*-activation function with very reproducible results. This encouraged the construction of two other early promoter-*lac*Z fusions as described below.

2. Construction of pE1B*lac*Z

The strategy for expressing *lac*Z gene under the control of the E1B promoter involved making a fusion product between the first 19 amino acid residues of E1B 19K protein and *lac*Z. This scheme was adopted solely for convenience since it did not involve introducing a translation codon as was required for pE3*lac*Z. Hence the structure of the construct would predict that both transcriptional and translational initiation sites would be identical to those of the

Figure 5-2. **Construction of pE3*lacZ***

The source of the E3 promoter was the plasmid pHB1 which contains the *Hin*dIII B-fragment of the Ad5 genome (72.8 to 89.1 mu) cloned into the vector pBR322 (F.L. Graham, pers. comm.). At 81.0 mu a single *Kpn*I site exists, thus serving as a convenient cloning site for the E3 promoter which is located at 76.6 mu. The *Kpn*I-*Hin*dIII fragment (81.0 to 72.8 mu) of pHB1 was cloned into the vector pMCS101 using the corresponding restriction sites to produce a plasmid called p467S4B, which contains the E3 promoter at position 76.6 mu. (Note that the cloned fragment also contains the E2A(e) promoter at 75.1 mu in the opposite strand.) The plasmid was cleaved with *Xma*III at the +77 position with respect to the E3 cap site then filled with Klenow and dNTP's. A *lacZo*-labeled cassette with flanking sequences comprised of oligomers AB42, AB43 and AB44 (see **Table 2-2**) was then ligated to produce the plasmid p467S5X. This cassette contained the translation initiation codon and a *Bam*HI site immediately after to provide a convenient fusion site with the *lacZ* gene in pTEQ4. The complete E3 promoter (black thick line) is now bounded by a 5' *Hin*dIII site and a 3' *Bam*HI site. Purification of this fragment and ligation into the *Hin*dIII-*Bam*HI sites of pTEQ4 produced the plasmid pE3*lacZ*. The structure of the fusion gene is shown at the bottom diagram. The cloned E3 promoter also contains the E2A(e) enhancer, the normal TATAA box and transcription initiation site. Note that the modification of the *Xma*III site to accept the ATG/*Bam*HI-containing cassette resulted in the creation of an *Nco*I site around the +81 position (confirmed by sequencing of p467S5X, data not shown). This site is commonly found in eukaryotic initiation codons and has been postulated to provide efficient initiation of translation (reviewed in Kozak, 1986).

Figure 5-2. **Construction of pE3*lacZ***

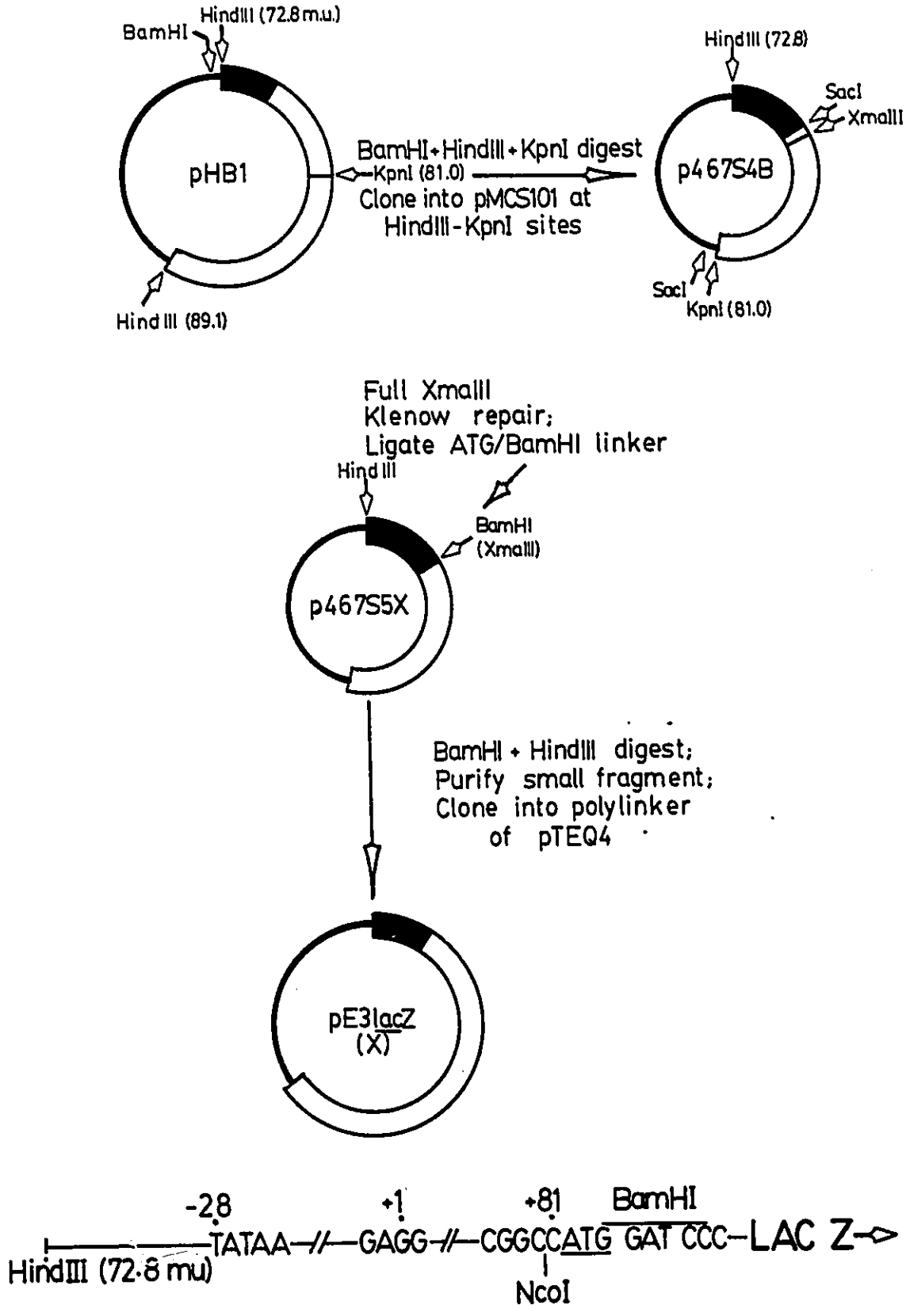
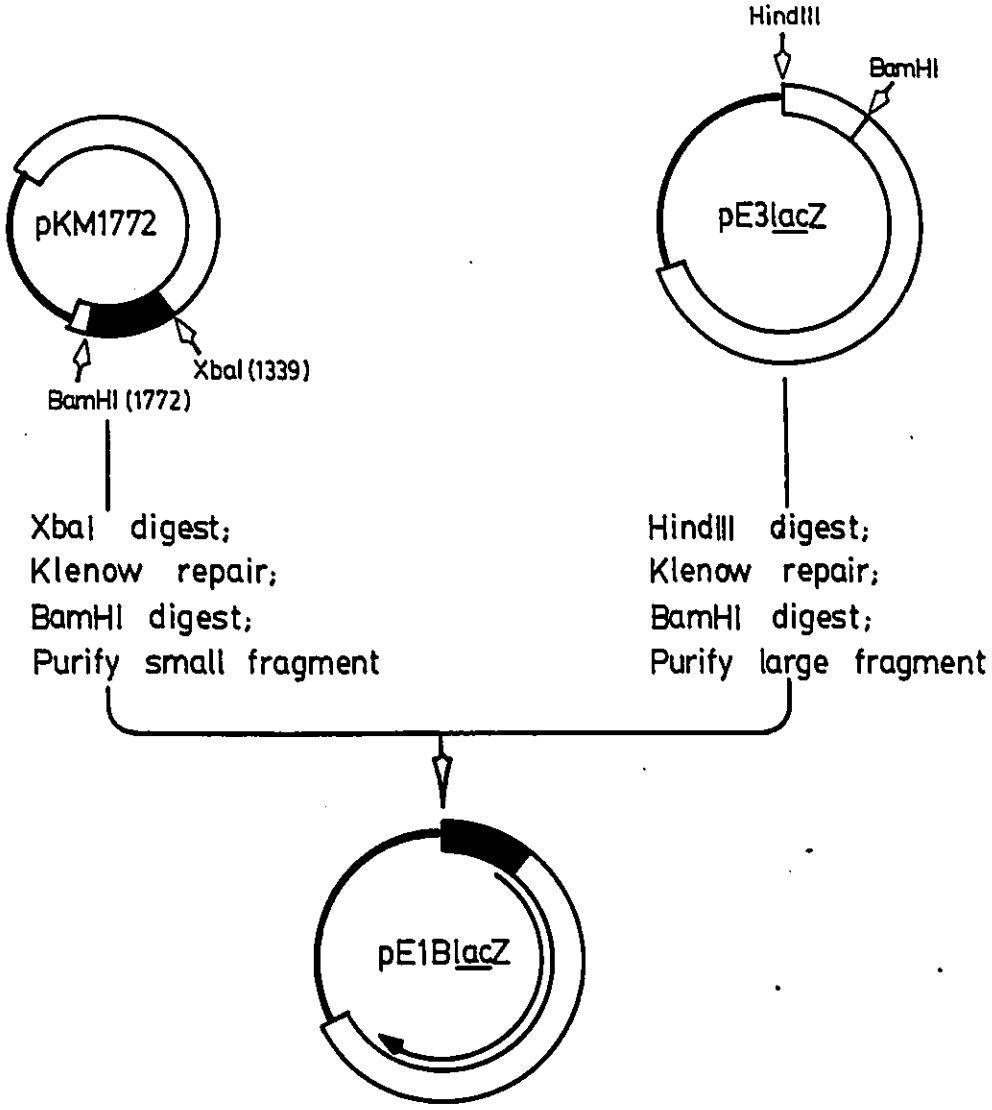
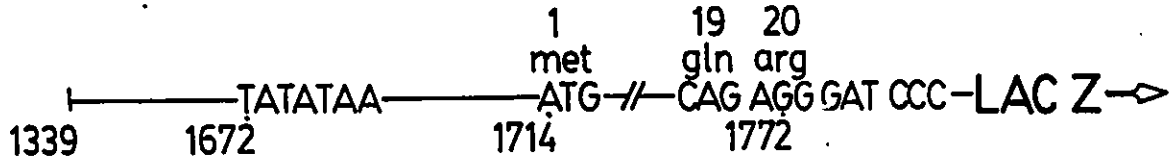


Figure 5-3. **Construction of pE1B*lacZ***

The source of the E1B promoter was the plasmid pAlu010, one of the mutant plasmids with a LOP-cassette insert at position 1772 (in the E1B region). Fortuitously, the position of the *Bam*HI site enables one to make a direct fusion with the lacZ gene in the pTEQ4 construct and still maintain an in-phase translational reading frame. The cloning thus was greatly simplified since it did not involve introducing an ATG in front of lacZ as was done in constructing pE3*lacZ*. The plasmid pAluG10 was digested with *Bam*HI then religated to collapse the linker resulting in the plasmid was called pKM1772. This plasmid was cleaved at the *Xba*I site (position 1339 on the Ad5 sequence) and the ends were filled with Klenow and dNTP's. Next, the desired E1B promoter-containing fragment was cleaved with *Bam*HI and the fragment purified from polyacrylamide gel. The source of lacZ was pE3*lacZ*, which was cut with *Hin*dIII, blunt-ended then digested with *Bam*HI. The two fragments were ligated and the resulting plasmid was called pE1B*lacZ*. The structure of the chimeric gene is shown below the diagram. Note that the chimeric lacZ product includes the first 20 aa residues of the 55KDa product of E1B.

Figure 5-3. **Construction of pE1B*lacZ***





native E1B gene. **Figure 5-3** illustrates the cloning strategy.

3. Construction of pE1A*lacZ*

The approach used to clone the E1A promoter in front of the lac Z gene involved fusing a fragment that also contained the E1A enhancer. Since E1A is modulated by its own products, this construct could also serve as the reporter plasmid for the enhancer repression assay. The construct contains the native E1A cap site but a synthetic translation initiation codon was introduced as was done for pE3*lac*Z. The strategy is illustrated in **Figure 5-4.**

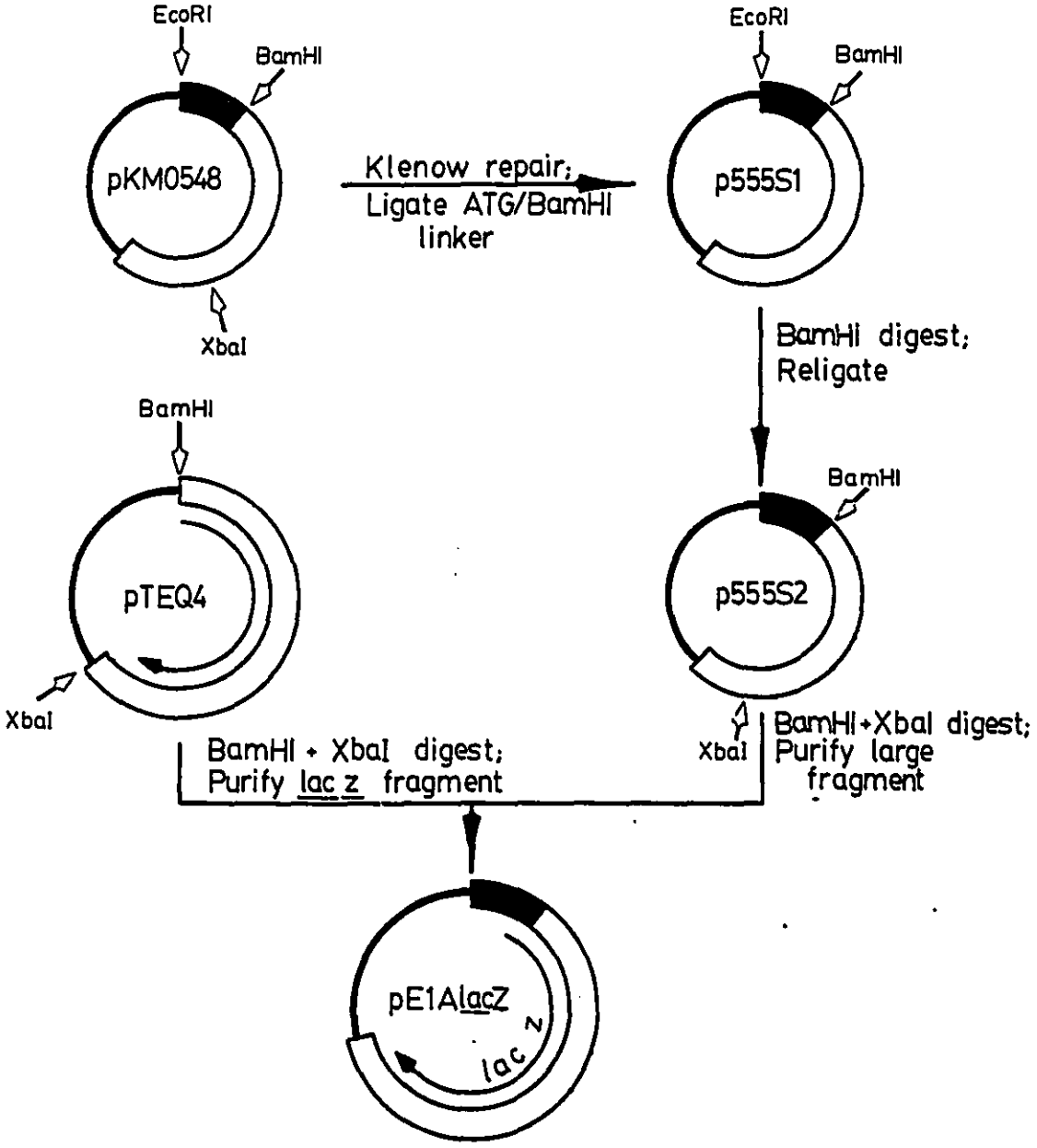
D. EXPRESSION OF VARIOUS *Lac*Z PLASMIDS IN HeLa CELLS

Preliminary experiments were carried out with wild-type E1A to establish the dose response for *trans*-activation and to identify linear region of the response curve for subsequent assays with E1A mutants. **Figure 5-5** shows the results of these experiments. The three reporter plasmids were individually used in co-transfection assays with pKH101, the plasmid which contains the wild-type E1A gene. The data indicate that the three *lac*Z plasmids were activated in a dose-dependent manner by pKH101 and induction was essentially identical at low doses of pKH101 (0 to 300 ng/dish). A reproducible difference can be identified at higher amounts since for pE1A*lac*Z linearity continued up to 500 ng/dish where maximum induction was observed even when 1.0 μg was used. In contrast, when pE3*lac*Z and pE1B*lac*Z were used, response to pKH101 declined beyond 300 ng/dish and continued to do so at higher amounts. These results show that kinetics of induction are essentially identical for pE1B*lac*Z and pE3*lac*Z but differ only

Figure 5-4. **Construction of pE1A*lacZ***

The source of the E1A enhancer/promoter was pKM0548, one of the insertional mutants with a 6 bp insert at position 548 produced by 'collapsing' its original LOP-cassette insert. This 6 bp insert was for a *Bam*HI recognition site which was deleted by filling the ends with Klenow polymerase and dNTP's. A *lacZo*-labeled cassette containing an ATG and a *Bam*HI site (similar to that used in constructing pE3*lacZ*) was ligated at this site. The resulting plasmid was called p555S1. The plasmid was cleaved with *Bam*HI to collapse the cassette core, resulting in a plasmid called p555S2. Then the lac Z coding sequence from pTEQ4 was cleaved using *Bam*HI and *Xba*I and purified. This fragment also contains the polyadenylation signal. To position this fragment behind the E1A promoter, p555S2 was digested with *Bam*HI and *Xba*I (pos. 1339) and the large purified fragment was ligated with the *lacZ* fragment from pTEQ4. The resulting plasmid, called pE1A*lacZ*, contained the E1A promoter in front of the lacZ sequence. The structure of the chimeric *lacZ* is shown below the diagram. The fusion was predicted to initiate transcription in the natural E1A cap site and translation in the synthetic start site near the *Bam*HI site.

Figure 5-4. **Construction of pE1A*lac*Z**



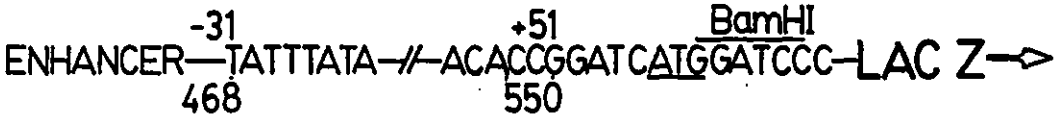
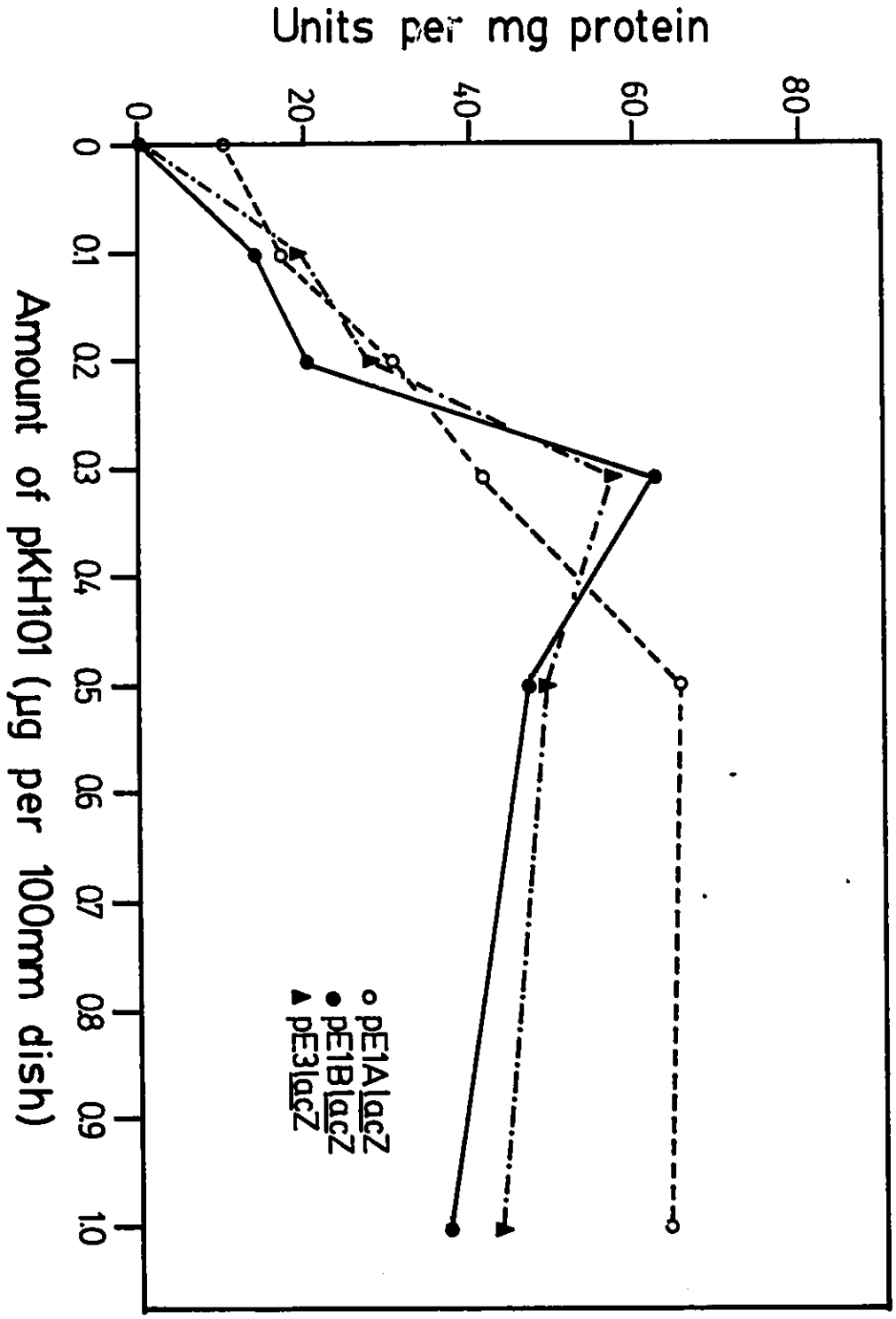


***Figure 5-5. Trans*-Activation Assays Using Various lacZ Plasmids**

Experiments which were initially carried out with the reporter plasmid pE3*lac*Z to determine optimal conditions for the *trans*-activation assay such as the amounts of reporter plasmid and the activator E1A, have indicated that 10 ug of reporter plasmid/100 mm dish was sufficient for induction in HeLa cells by the wild-type E1A to a level that could be easily measured using the standard colourimetric assay for *lac*Z (data not shown). Therefore, appropriate amounts of pKH101 were co-transfected as described in **Materials and Methods** to test the kinetics of induction of the three reporter plasmids namely pE1A*lac*Z, pE1B*lac*Z and pE3*lac*Z. Each point represents the average from 4 dishes and the error bar (deleted for clarity) represents one standard deviation on both sides with a value that range between 7 to 23% of the mean.

Figure 5-5. ***Trans*-Activation Assays Using Various *Lac*Z Plasmids**



slightly for pE1A*lac*Z. It was also established from these that relatively low amounts of E1A were sufficient to activate the reporter genes. In subsequent experiments, a standard amount of 100 ng/dish of E1A plasmids was used to assay for the *trans*-activation function.

CHAPTER VI

TRANSCRIPTIONAL REGULATORY ACTIVITY OF E1A INSERTIONAL MUTANTS

A. TRANSCRIPTIONAL ACTIVATION BY E1A INSERTIONAL MUTANTS

The *trans*-activation phenotypes of the insertion mutants were determined using the *lac*Z-based assay described in the previous chapter. Co-transfection experiments with individual mutants and each of the three reporter genes (pE1A*lac*Z, pE1B*lac*Z and pE3*lac*Z) were carried out using HeLa cells as recipients. The *trans*-activation activities of the mutants are expressed as percentage of wild-type control after subtracting the background (level of lac Z activity in the absence of an E1A-containing plasmid). The data for the *trans*-activation of pE3*lac*Z are represented in histograms as shown in **Figure 6-1.** The same results are also shown in **Table 6-1** which contains in addition the data for the *trans*-activation of pE1A*lac*Z and pE1B*lac*Z. The following discussion pertains to the data obtained by co-transfection with pE3*lac*Z but applies equally well to results obtained with the other two plasmids, namely, pE1A*lac*Z and pE1B*lac*Z.

1. Mutants in Exon I

Nucleotide positions 717, 812, 819, 827, 863, 884, 906, and 908 are located in exon 1 of the E1A 12S product, a region which contains the two conserved regions postulated to be associated with the transformation function. This region is rich in proline residues, which are distributed in clusters most notably between conserved regions 1 and 2. The conserved regions in this exon do not appear to be

***Figure 6-1. Trans*-Activation of pE3*lacZ* by the E1A Insertion Mutants**

The *trans*-activation phenotypes of the E1A insertion mutants were initially determined using the pE3*lacZ* reporter plasmid. The results are virtually identical when the other 2 reporter plasmids were used (see Table 6-1 which contains the complete results using the three reporter plasmids), hence for simplicity only the data for *trans*-activation of pE3*lacZ* are shown in the histogram. Each value represents the average of at least 2 experiments, in which four dishes were used per experiment, and was calculated as the percentage of wild-type activity after subtracting uninduced activity (*lac* Z activity when the reporter plasmid was transfected alone).

Legend:

empty bar	-39bp insert in "plus" orientation
solid bar	-39bp insert in "minus" orientation
light stippled bar	-6bp insert derived from the "plus" orientation
dark stippled bar	-6bp insert derived from the "minus" orientation

Figure 6-1. ***Trans*-Activation of pE3*lac*Z by the E1A Insertion Mutants**

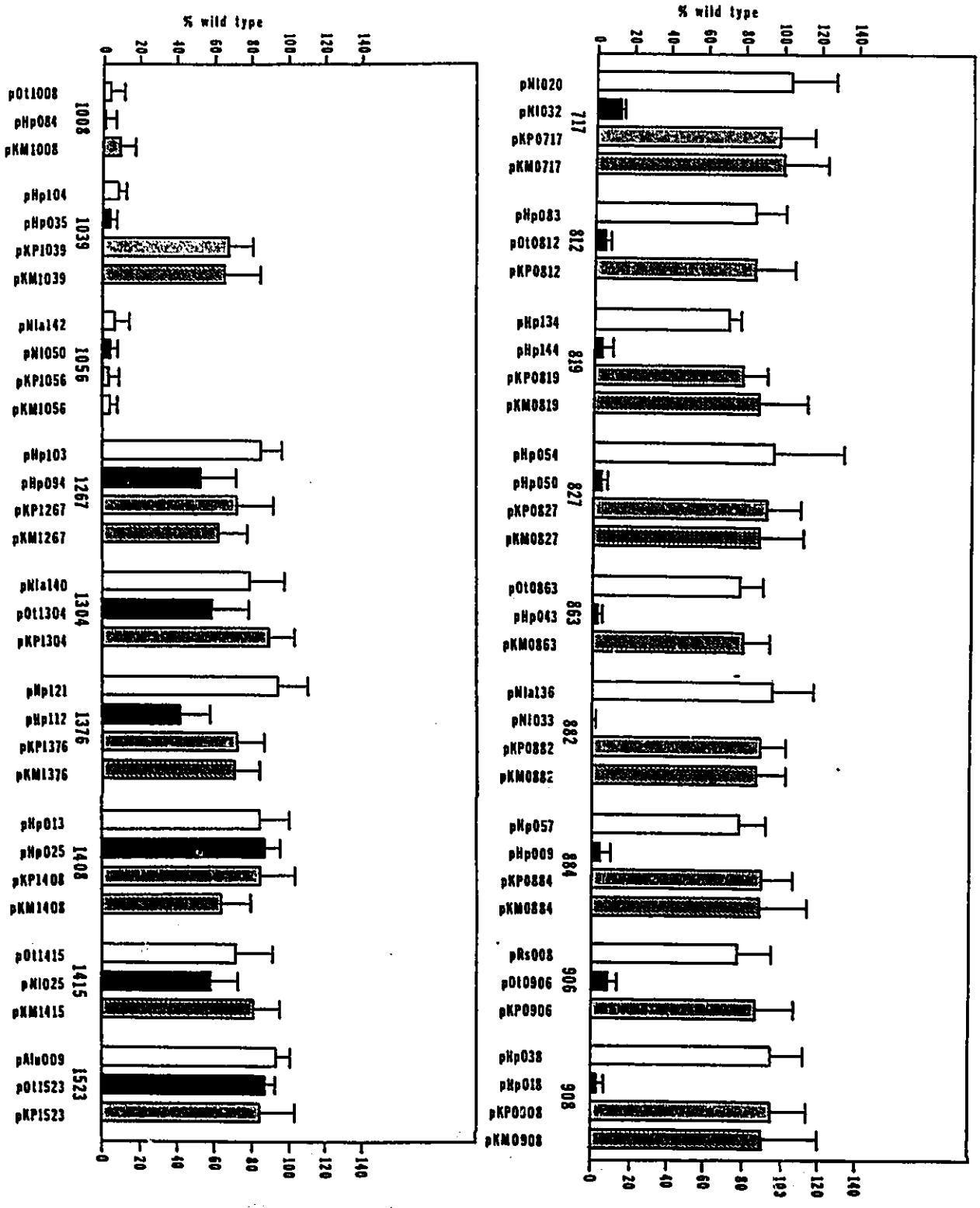


Table 6-1. Data for *Trans*-Activation by the E1A Insertion Mutants

MUTANT	POSITION	pE1Alacz	pE1Blacz	pE3lacz
pNl020	717+	98.6 ± 20.2	115.5 ± 11.0	103.9 ± 23.9
pNl032	717–	nd	6.1 ± 4.1	12.3 ± 2.2
pKP0717	717c	87.9 ± 10.3	146.0 ± 28.0	97.3 ± 18.7
pKM0717	717c	nd	91.8 ± 14.4	100.2 ± 23.4
pHp083	812+	84.1 ± 10.1	54.3 ± 5.8	84.3 ± 17.1
pOt0812	812–	nd	4.8 ± 5.0	5.0 ± 3.2
pKP0812	812c	86.3 ± 20.8	75.3 ± 13.5	85.6 ± 20.3
pHp134	819+	84.1 ± 16.7	91.1 ± 19.3	70.8 ± 7.5
pHp144	819–	nd	1.8 ± 0.7	3.3 ± 6.2
pKP0819	819c	86.5 ± 17.6	120.1 ± 12.1	79.3 ± 12.5
pKM0819	819c	nd	102.6 ± 18.6	86.9 ± 26.6
pHp054	827+	110.4 ± 14.4	106.9 ± 29.9	95.0 ± 38.0
pHp050	827–	nd	9.3 ± 1.3	4.0 ± 2.7
pKP0827	827c	103.5 ± 17.6	96.8 ± 9.2	91.3 ± 18.5
pKM0827	827c	nd	76.6 ± 15.8	88.3 ± 22.5
pOt0863	863+	83.2 ± 17.8	78.4 ± 11.2	77.9 ± 12.0
pHp043	863–	nd	3.3 ± 1.6	1.8 ± 1.8
pKM0863	863c	88.8 ± 19.6	90.1 ± 12.5	79.3 ± 13.8
pNla136	882+	109.1 ± 16.7	95.3 ± 20.5	95.4 ± 22.1
pNl033	882–	nd	0.0 ± 3.2	-0.6 ± 1.8
pKP0882	882c	101.9 ± 22.8	87.4 ± 15.8	89.5 ± 13.6
pKM0882	882c	nd	89.5 ± 31.4	87.3 ± 15.7
pHp057	884+	75.4 ± 14.0	96.3 ± 13.1	79.2 ± 12.6
pHp009	884–	nd	16.4 ± 5.0	4.6 ± 5.7
pKP0884	884c	98.0 ± 11.7	126.4 ± 9.9	90.7 ± 16.2
pKM0884	884c	nd	102.6 ± 6.0	89.6 ± 24.3
pRs008	906+	103.6 ± 14.5	110.3 ± 23.8	78.7 ± 16.7
pOt0906	906–	nd	nd	8.6 ± 4.9
pKP0906	906c	90.1 ± 27.4	95.1 ± 19.7	87.0 ± 20.7
pHp037	908+	93.2 ± 10.1	84.0 ± 10.5	95.4 ± 16.6
pHp018	908–	nd	0.0 ± 1.3	2.8 ± 3.6
pKP0908	908c	72.9 ± 8.9	97.8 ± 18.7	95.3 ± 19.6
pKM0908	908c	nd	71.6 ± 12.6	90.9 ± 29.4

MUTANT	POSITION	pE1Alacz	pE1Blacz	pE3lacz
pOt1008	1008+	6.8 ± 5.2	13.3 ± 6.5	3.6 ± 7.9
pHp084	1008–	nd	0.3 ± 2.1	1.1 ± 6.1
pKM1008	1008c	1.6 ± 6.0	18.9 ± 5.5	9.5 ± 7.9
pHp104	1039+	-0.3 ± 5.8	7.1 ± 3.2	8.5 ± 4.4
pHp035	1039–	nd	4.7 ± 2.7	3.9 ± 3.7
pKP1039	1039c	57.7 ± 7.4	62.9 ± 9.8	68.3 ± 12.6
pKM1039	1039c	nd	73.0 ± 9.6	71.4 ± 13.3
pNla142	1056+	14.6 ± 4.4	5.4 ± 3.0	5.8 ± 3.5
pNl050	1056–	nd	9.4 ± 3.6	4.6 ± 3.3
pKP1056	1056c	0.9 ± 3.5	4.0 ± 4.0	3.6 ± 5.5
pKM1056	1056c	nd	5.3 ± 4.4	3.5 ± 4.6
pHp103	1267+	89.0 ± 9.7	70.5 ± 25.3	85.2 ± 11.4
pHp094	1267–	nd	65.7 ± 11.6	53.2 ± 18.5
pKP1267	1267c	87.4 ± 19.2	74.5 ± 8.3	71.4 ± 20.0
pKM1267	1267c	nd	79.8 ± 11.0	61.9 ± 15.7
pNla140	1304+	106.4 ± 9.9	78.9 ± 26.8	78.4 ± 18.6
pOt1304	1304–	nd	67.1 ± 18.9	59.1 ± 19.6
pKP1304	1304c	nd	107.4 ± 25.3	89.2 ± 13.4
pHp121	1376+	81.3 ± 7.9	111.5 ± 18.9	93.7 ± 16.0
pHp112	1376–	nd	74.0 ± 20.8	41.7 ± 15.7
pKP1376	1376c	86.9 ± 19.8	87.8 ± 19.9	71.5 ± 14.8
pKM1376	1376c	nd	91.7 ± 14.6	70.3 ± 13.6
pHp013	1408+	81.4 ± 19.3	89.1 ± 15.6	85.0 ± 15.5
pHp025	1408–	nd	25.7 ± 1.4	87.4 ± 8.0
pKP1408	1408c	90.8 ± 10.0	84.7 ± 26.7	85.0 ± 18.3
pKM1408	1408c	nd	95.8 ± 21.7	63.9 ± 16.5
pOt1415	1415+	90.7 ± 20.1	85.1 ± 5.9	72.1 ± 19.3
pNl025	1415–	nd	102.3 ± 12.5	58.4 ± 14.5
pKP1415	1415c	82.1 ± 14.3	85.6 ± 30.8	81.3 ± 13.9
pAlu009	1523+	86.5 ± 17.8	89.0 ± 16.3	93.7 ± 7.1
pOt1523	1523–	nd	86.7 ± 20.5	87.7 ± 5.2
pKP1523	1523c	93.5 ± 15.3	122.8 ± 24.2	85.0 ± 18.6

particularly rich in proline residues. Only mutants in this region whose inserts were in the closed reading frame orientation were defective in ability to *trans*-activate pE3*lac*Z, whereas neither mutants with large inserts in the open orientation nor those which had been collapsed showed significant decrease in activity. That mutants in this region in the "closed orientation (those whose products are predicted to be truncated at the insertion site) did not show activity that was significantly higher than background levels suggest that either (a) the truncated products did not contain the *trans*-activation domain, or (b) that the mutant protein products were sufficiently unstable that they failed to carry out the *trans*-activation function. Similarly, the wild-type activity of the "+" and "c" mutants in exon I might indicate (a) that there was no *trans*-activation domain(s) in this exon, or (b) that a *trans*-activation domain was present but was insensitive to insertional mutagenesis. No attempt was made to distinguish between these possibilities. It would appear that these results are in agreement with those found by others (see **INTRODUCTION**) who have concluded on the basis of other E1A mutants that exon I does not contain domains for *trans*-activation. Since insertion of segments as long as 13aa residues in this exon did not affect activity, it is suggestive that the higher order structure contributed by this region does not contribute to function.

The protein product of the E1A 11S mRNA is similar in sequence to the 13S product except for the absence of amino acid residues 26 to 99 that map in exon I (Ulfendahl *et al.*, 1987; Stephens and Harlow, 1987). Mutants at nucleotide positions 717, 812, 819 and 827 map in this intron, hence these mutants (regardless of type) would be expected to produce a normal 11S product, which has been shown able to *trans*-activate (Ulfendahl *et al.*, 1987). The results of the *trans*-

activation assays shown here, however, do not appear to be in agreement with this conclusion since mutants in the "closed" orientation at those 4 positions failed to *trans*-activate. It is possible that in the assays used here the 11S product is expressed at very low levels or not expressed at all so that the putative *trans*-activating function of the 11S product was not observed.

2. Mutants in the Unique Region

There were three mutated sites in the unique region: nt1008, nt1039, and nt1056. All mutants with large inserts failed to *trans*-activate pE3*lac*Z regardless of orientation. This failure to *trans*-activate strongly suggested that these sites are within an important region of the *trans*-activation domain. Analysis of the collapsed form of the mutants in these three sites, however, not only supported the conclusion that these positions are in a highly sensitive domain of the *trans*-activation function, but might also refine our understanding of the functions of the unique region since only two, out of three mutant sites in the unique region, namely nt1008 and nt1056, appeared vulnerable to 2aa residue insertion. These failed to *trans*-activate pE3*lac*Z, but the 2aa insert at nt1039, in contrast, had wild-type *trans*-activating ability. Again, these results are in agreement with previous works which showed that the unique region is an important domain of the *trans*-activation function.

3. Mutants in Exon II

Mutants in positions nt1267, nt1304, nt1376, nt1415, and nt1523 reside in exon II, a region which is common to the 13S, 12S, 11S and 10S mRNAs. *Trans*-activation using pE3*lac*Z indicated that no particular site in this region was highly vulnerable to the three types of mutations. It did not appear important whether the

insert was large or small, or whether the products were truncated due to the closed orientation of the insert. The wild-type activity of the "-" mutants suggested that the C-terminus did not contribute to *trans*-activation which suggests that the rapid nuclear localization signal found at the C-terminus of this exon is not required for full activity. Either rapid localization to the nucleus is unimportant or another nuclear localization signal is present upstream of nt1267.

4. *Trans*-Activation Assay Using pE1A*lacZ* and pE1B*lacZ* Reporter Plasmids

One interesting question regarding E1A *trans*-activation functions is concerned with the fact that E1A-responsive promoters do not appear to have any sequence homology which can be specifically associated with E1A-dependence. In a related question, neither is it known whether there exists a single domain or set of domains in E1A proteins that is invariably required by all E1A-activated promoters, or whether these different promoters have different domain requirements. It was therefore of particular interest to determine if structural differences in E1A-responsive promoters could be correlated with different domains for *trans*-activation by the E1A proteins. In an effort to address this question, the *trans*-activation activity of the various mutants was assayed using two other *lacZ* reporter plasmids, namely pE1A*lacZ* and pE1B*lacZ*, to determine if these promoters require the same E1A domains.

Mutants were assayed using pE1A*lacZ* and pE1B*lacZ* as reporter plasmids essentially as was carried out with pE3*lacZ*. **Table 6-1** shows the average values obtained from several experiments for the *trans*-activation of pE1A*lacZ* and pE1B*lacZ* in comparison with pE3*lacZ*. Statistical analysis using Student's t-test method indicated that the *trans*-activation of both pE1B*lacZ* and pE1A*lacZ* by the

various mutants was similar in all respects to activation of pE3*lac*Z; that is, mutants that induced pE3*lac*Z also induced the other two promoters and those that did not induce pE3*lac*Z were likewise incapable of inducing pE1B*lac*Z and pE1A*lac*Z. From these results, it appears that only the unique region is required by the three different promoters for E1A *trans*-activation when analyzed by insertional mutagenesis.

B. ENHANCER REPRESSION BY E1A INSERTIONAL MUTANTS

1. *Trans*-Repression Assays

One E1A function related to gene regulation that is well-documented is *trans*-repression of enhancers. It is widely believed that the repression function is not coupled to the *trans*-activation domain since mutants that fail to activate newly transferred genes may not necessarily be *trans*-repression negative (see INTRODUCTION).

The *trans*-repression function of E1A is frequently assayed on reporter genes driven by either SV40 or polyoma enhancer. In this study, this E1A function was determined for the various mutants using a system in which the reporter gene was driven by the E1A enhancer, since it is well-known that E1A negatively regulates its own gene through its enhancer (Borrelli *et al.*, 1984). The same plasmid, pE1A*lac*Z, as was used in the *trans*-activation assay described in the previous section was used for the repression assay.

2. Enhancer Repression Assay Using pE1A*lac*Z

An initial study to determine the feasibility of using pE1A*lac*Z as the reporter gene in the *trans*-repression assay was carried out (**Figure 6-2**). When 25

μg of pE1A*lac*Z/dish was used, the level of constitutive expression in HeLa cells was 23 units/mg protein ±2.6 with one standard deviation of the mean of three dishes. The standard deviation of the mean from one experiment to another ranged from 2.6 to 39.5. Reduction from this basal level of *lac*Z to 20-30% was achieved when 1 μg of pHp084/dish was co-transfected with the reporter plasmid. The plasmid pHp084 contained an insertion of the mutator DNA in the closed reading frame orientation at nt1008, which was predicted to produce a truncated 289aa residue protein (product of 13S) but a full length 243aa residue protein (product of 12S). The observed reduction of expression of pE1A*lac*Z was dependent on the amount of co-transfected pHp084 as shown in **Figure 6-2.** Thus, in the repression assay 25 μg pE1A*lac*Z and 1 μg of E1A-containing plasmid were used as the standard amounts per dish.

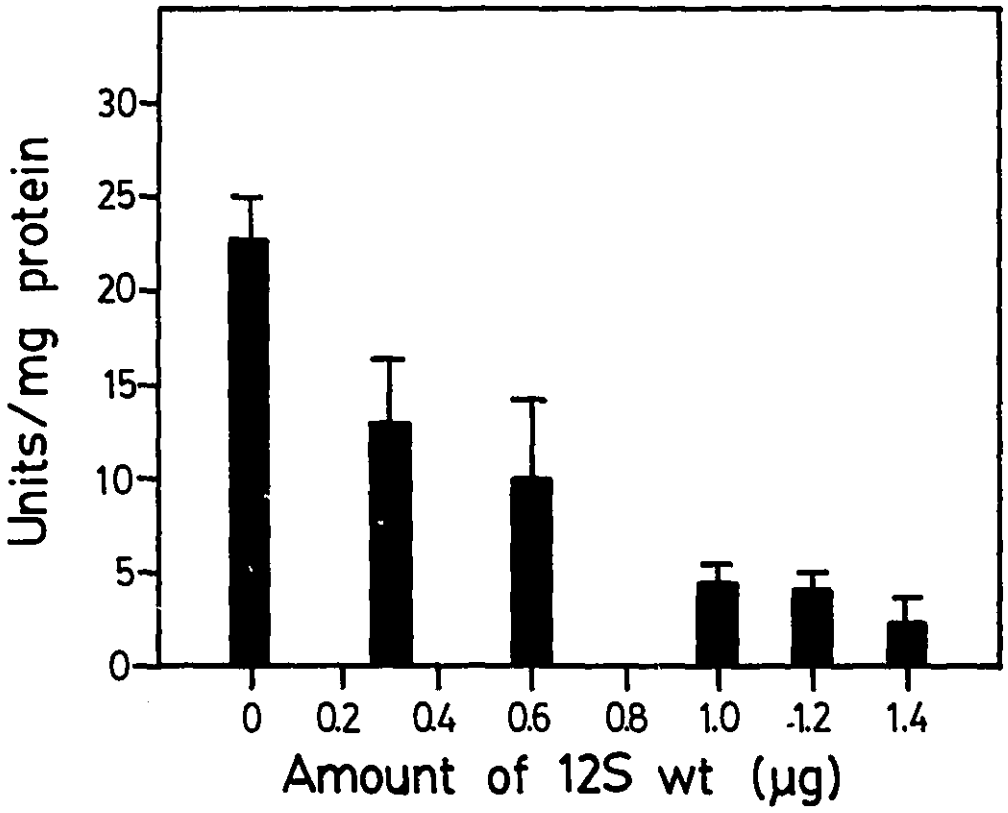
3. Conversion of Mutants into 12S Expressors

Experiments carried out using the pE1A*lac*Z reporter plasmid showed that wild-type E1A-containing plasmid pKH101 increased enzyme activity even at high amounts of input DNA. This indicated that the *trans*-repression function was effectively masked by the *trans*-activation function when both were allowed to be simultaneously expressed in these transfection studies. As noted in the **INTRODUCTION,** the 12S E1A mRNA is mainly responsible for the *trans*-repression function, so that it seemed appropriate to assay for this activity in the absence of 13S mRNA expression. Therefore it was necessary to modify all mutants so that they expressed the 12S product alone. This required insertion of an in-phase stop codon within the unique region so that the protein product of the 13S mRNA was truncated at the SmaI site, thus expressing only 12 amino acid residues

Figure 6-2. ***Trans*-Repression Assay Using pE1A*lacZ***

The amount of pE1A*lacZ* used for the repression assay was 25 ug/100 mm dish. Identical transfection conditions as in *trans*-activation assays were used. Each point is the average from 4 dishes using appropriate amounts of pHp084, the 12S wild type control.

Figure 6-2. ***Trans*-Repression Assay Using pE1A*lacZ***



(139-150) of the unique region. **Figure 6-3** illustrates the strategy used and contains the details of construction. The structure of one such mutant is shown in **Figure 6-4** which shows the sequence across the insert containing a HindIII site and a stop codon at the *Xma*I site. The reconstructed mutants were used in an enhancer repression assay as described below under the assumption that the truncated product did not interfere.

Except for positions nt819, nt1008 and nt1415, all mutants have been reconstructed as 12S expressors. These mutants are by definition double mutants since they express a truncated 13S product and 12S products with various mutations due to linker insertion. Mutants at nt1008 which have disrupted the *Xma*I site due to insertion of the LOP cassette were not reconstructed. In fact, mutants in the unique region (nt1008, nt1039 and nt1056) would have been irrelevant in this assay since they would have no effect on the 12S product.

4. *Trans*-Repression Activity of E1A Mutants

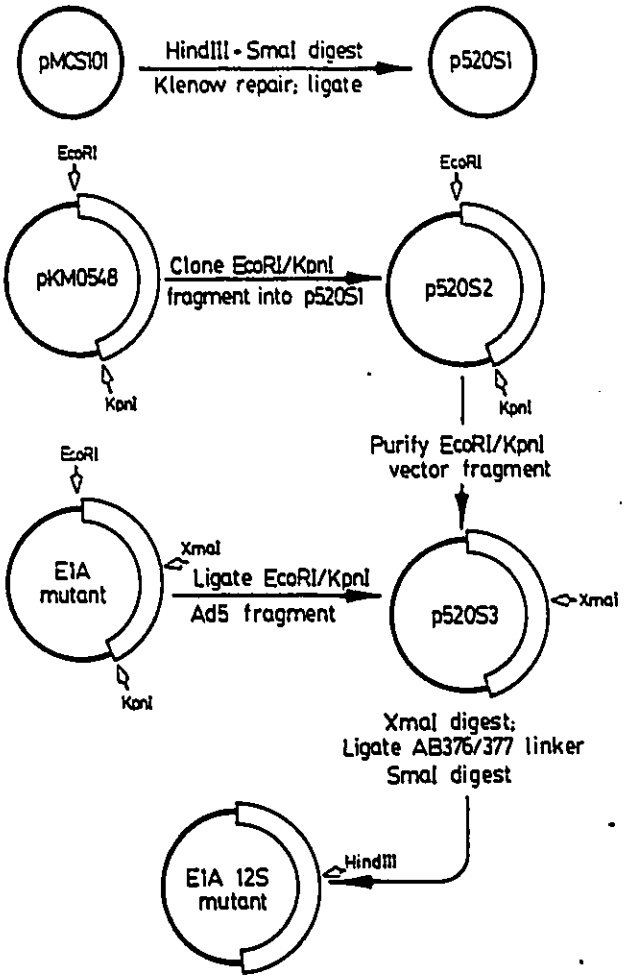
1. Mutants in Exon I

Eight mutated sites in exon I (nt717, nt812, nt827, nt863, nt882, nt884, nt906, and nt908) were assayed for the *trans*-repression function as shown in **Figure 6-5.** Mutants in the closed orientation all failed to repress the expression of pE1A*lac*Z while those in the open orientation and those in the collapsed form had wild-type or near wild-type levels of activity except for position nt908. For this particular position, only the collapsed form repressed while both open and closed orientations did not. The inability of the closed orientation mutants in exon I to repress indicated that exon I contains sequences which are important, and the failure of p12S908+ to repress identified a region in exon I that was sensitive to insertion

Figure 6-3. **Construction of Mutants Expressing the 12S Product Alone**

The strategy involved inserting synthetic linkers containing an in-phase stop codon and a *Hin*dIII site at the *Xma*I site in the unique region of E1A. The oligonucleotides AB376 and AB377 were used (see **Table 2-2**). Since the parent plasmid pKH101 contained 2 *Xma*I sites, the insertion of the linker was not completely straightforward. First, the vector pMCS101 (see the **APPENDIX**) was modified by deleting both the single *Xma*I and the single *Hin*dIII sites using Klenow and dNTP's. The resulting plasmid, p520S1, was doubly digested with *Eco*RI and *Kpn*I and the E1A-containing *Eco*RI/*Kpn*I fragment of pKM0548 was initially inserted into the vector to produce p520S2. A batch preparation of the *Eco*RI/*Kpn*I vector fragment was made from this plasmid, the idea being that it was easier to monitor on gels the complete double-enzyme digest of this plasmid than the modified pMCS101 where the *Kpn*I and *Eco*RI sites are very close to each other. Then individual E1A-containing *Eco*RI/*Kpn*I fragments of each of the mutant were purified and cloned to the purified vector. The next step involved digestion of these intermediate plasmids with *Xma*I with which the reconstituted linker, which was not phosphorylated, was ligated. Since the linker was designed to eliminate the *Sma*I recognition site, the ligation mixture was incubated with *Sma*I (the isoschizomer of *Xma*I) overnight to enrich for the population of recircularized plasmids with the desired insert. The desired double mutant was identified on gels by the presence of a single *Hin*dIII site found in the inserted linker. The structure of the linker (underlined) when inserted into the nt1007 XmaI site is shown in its two orientations. Note that the linker is not perfectly symmetrical so that the sequence of the two orientations differs; however, termination of translation in each orientation produces identical results truncated products.

Figure 6-3. **Construction of Mutants Expressing the 12S Product Alone**



his pro gly his
—CAC CCC GGG CAC—
—GTG GGG CCC GTG—

1007 end 1008
CAC CCC GGT TAA GCT TTA ACC GGG CAC
GTG GGG CCA ATT CGA AAT TGG CCC GTG
HindIII

end
CAC CCC GGT TAA AGC TTA ACC GGG CAC
GTG GGG CCA ATT TCG AAT TGG CCC GTG
HindIII

Figure 6-4. **Sequence Analysis across the 12S Mutation**

A single double mutant, p12S1056(-), was sequenced to determine the structure of the inserted linker at the nt1007 *Xma*I site. This mutant contained the full-length mutator cassette at nt1056 in the "closed" orientation (shown as stippled bar). This allowed sequencing using the core oligomer AB10 as primer (solid bar) from that position in a leftward direction, and therefore across nt1007 which contained the linker with termination codon. The sequencing result clearly showed the correct structure of the insert as predicted in **Figure 6-3.** Similar analysis for other double mutants was not carried out because the presence of the *Hin*dIII site was considered to be adequate evidence for the correct insertion of the terminator.

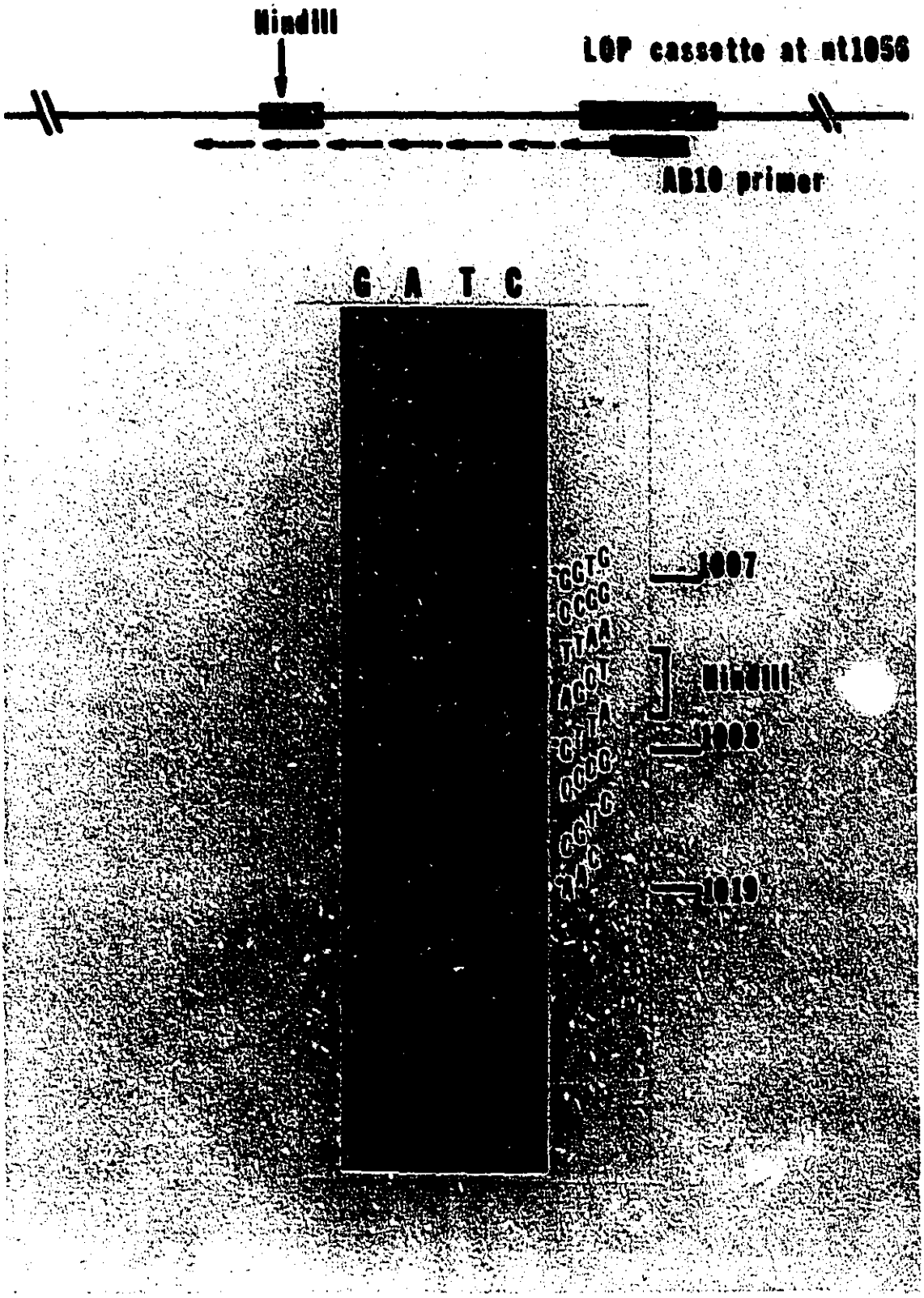

HindIII
LOP cassette at nt1056
AB10 primer
G A T C
1007
HindIII
1008
1019

Figure 6-5. ***Trans*-Repression Activity of Insertional Mutants**

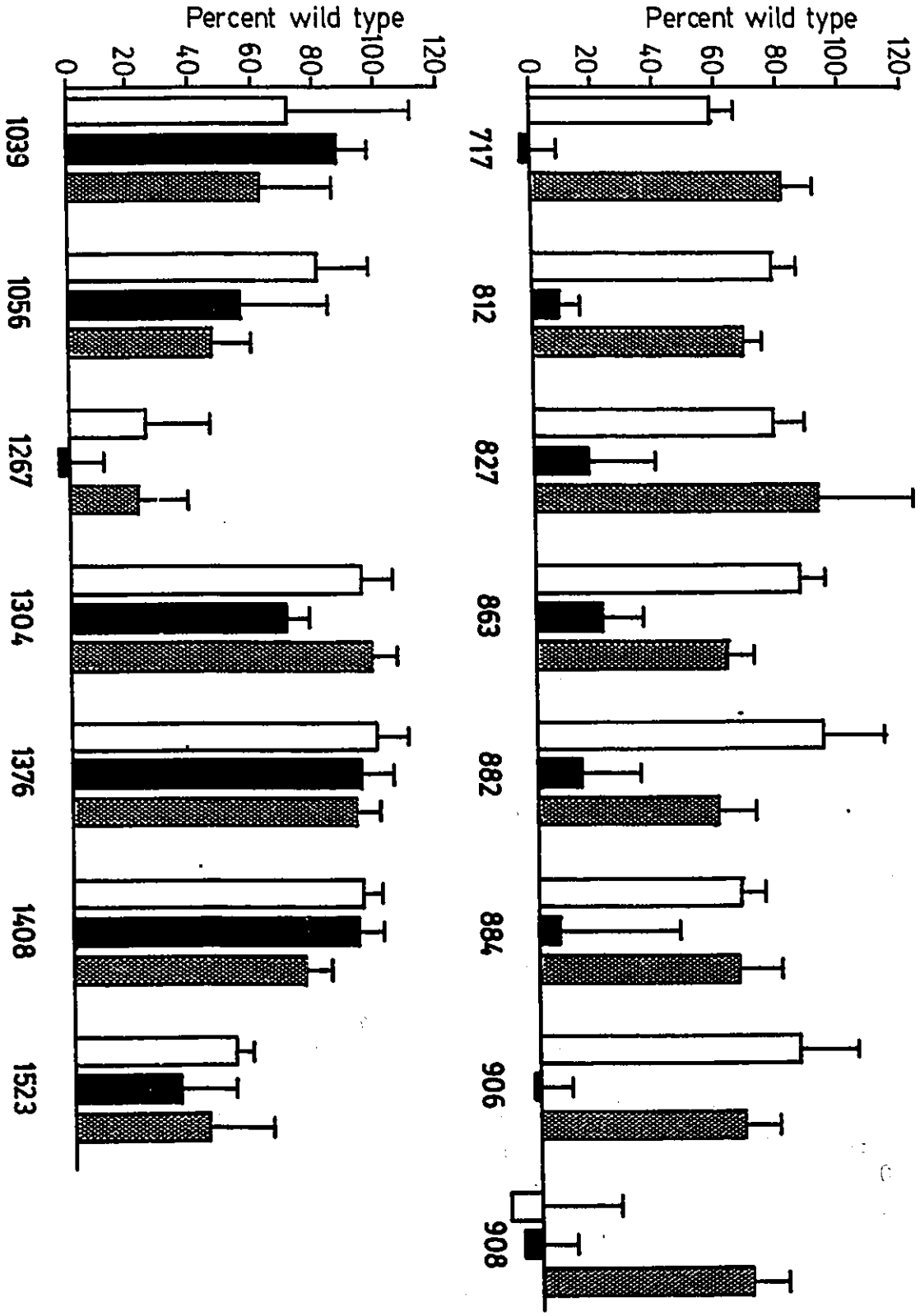
The histogram represents the average activity of the mutant as percentage of the wild type control (pHp084) from at least two experiments using two dishes per transfection. 25ug/100 mm dish of pE1A*lac*Z was used under identical transfection conditions as for *trans*-activation. 1ug/100 mm dish of mutant and wild-type 12S E1A DNA was added. The error bars represent one standard deviation on both sides of the mean. For calculations, the wild-type activity was the specific lac Z activity when pE1A*lac*Z was used alone (A) minus that when wild-type pHp084 was co-transfected (B). Given B' as the *lac*Z activity when a mutant was used, the *trans*-repression activity of a mutant is calculated as:

$$\frac{A - B'}{A - B} \times 100\%$$

Legend:

empty bar	-39bp insert in "plus" orientation
solid bar	-39bp insert in "minus" orientation
stippled bar	-6bp insert derived from either the "plus" or "minus" orientation

Figure 6-5. Trans-Repression Activity of the Insertional Mutants



of 13aa but not to 2aa residues.

2. Mutants in Exon II

Mutants at nt1267, nt1304, nt1376, nt1408, and nt1523 were also assayed for their repression activity. For nt1267, it appeared that all three types (open, closed and collapsed) have either reduced or zero levels of repression activity, indicating that this position is quite important for the repression function. The mutants in other positions in this region showed wild-type or near wild-type levels.

Thus, the repression assay indicated that the repression function was contained mainly in exon I and extended a small distance into exon II demarcated by and including nt1267. There were two positions that showed sensitivity to 13aa residue insertions (nt908 and nt1267) although these two positions differed in their sensitivity to 2aa residue insertion.

CHAPTER VII

DISCUSSION

Modern genetic approaches have been largely successful in unravelling roles played by key gene products in fundamental cellular processes. With the development of techniques that allow isolation of genes and introduction of any desired changes into such genes, it has been possible to study more closely the relationship between structure and function of gene products, the knowledge of which might provide clues to the mechanisms of cellular processes in which such gene products are involved. To this end various mutagenesis techniques have been developed, each involving different approaches for introducing changes into the coding sequence. These techniques produce changes in the final mutated polypeptide chain that in general may be classified as *deletion*, *substitution*, or *insertion* of amino acid residues. One approach is to use site directed mutagenesis so that specific well-defined alterations are made in regions of the polypeptide chain which are suspected of comprising functional domains. On the other hand, random mutations may be desirable in cases where no prior knowledge of structure-function relationship has been established.

The different types of mutations can be achieved through different routes. For example, deletion mutagenesis may be achieved simply by cutting out a region in the coding sequence defined by restriction enzymes, or by the more sophisticated deletion loop technique through the use of synthetic oligonucleotides. Amino acid substitutions are primarily generated through synthetic oligonucleotides

that bear the desired change in the coding sequence, but can also be achieved in a less defined manner by means of chemicals such as sodium bisulphite. Finally, insertion mutations may be produced by introducing synthetic oligomers into sites defined by restriction enzymes or randomly cleaved DNA target by DNAseI in the presence of manganese ions. Gross definition of gene functions has also been achieved using transposable elements as mutagens but the technique is generally of less value when defined genes are under study.

The new insertional mutagenesis technique described in this report is characterized by its simplicity and efficiency in generation of a set of mutants with random insertions throughout a particular DNA segment. The advantages of the technique are several: First, the oligonucleotides that comprise the inserted DNA can be readily obtained from commercial companies or can be made 'in house'. Second, by virtue of phenotypic labeling with the bacterial *lac* operator sequence, the often labourious task of isolating clones with inserts is reduced to merely identifying coloured colonies at the primary level of selection for antibiotic resistance. Third, full characterization of individual mutants at the DNA level could be accomplished by utilizing oligomers that are components of the cassette as primers for sequencing, allowing the identification of the precise insertion site as well as the structure of the insert. This technique is suitable as a general method for modifying DNA sequences in a plasmid. For example, since there is a common core, only the oligonucleotides at the ends need be tailored for different insertion sites and, hence, virtually any sequence of interest may be introduced into a plasmid. However, the utility of the technique is best realized when insertions at various locations, as opposed to a single one, on the plasmid are desired, as in isolating a set of mutants of a particular gene.

The approach used to generate insertion mutants in this report was to employ various multi-cut restriction enzymes to linearize the target plasmid at a one of a number of possible sites where the mutator DNA could be inserted. Thus, the set of mutants that can be generated is limited in size by the number of recognition sites of these enzymes on the target DNA. There are many such enzymes available commercially so that virtually any sequence can be mutated in this manner. The distribution of mutable sites can be influenced by base composition of target DNA; that is, a G+C rich sequence will be expected to have a high frequency of cuts with enzymes like *Hpa*II (which recognizes 5'-CCGG-3'), whereas DNA segments with equal composition of the 4 bases would be recognized more frequently by enzymes like *Alu*I (5'-AGTC-3'). In certain instances, however, the number of mutable sites may be greatly reduced if, for example, the target sequence is A+T-rich since no presently known restriction recognition site is composed only of A and T bases except *Dra*I and *Ssp*I, both of whom have 6bp recognition sites and, therefore, are expected to cut infrequently.

Mutagenesis using the linker scanning approach (McKnight and Kingsbury, 1982) is another application in which the LOP-cassette method may be suitable. Since this method involves the systematic replacement of sequences with usually those of a restriction enzyme site obtained in synthetic oligomers, the initial isolation and characterization of plasmids with insertions from which matching fragments are later obtained to make the linker scan mutants would be greatly facilitated.

*Lac*Z-based Reporter Genes

Gene transfer experiments have been moderately successful in unravelling basic regulatory controls of eukaryotic gene expression. Although gene expression

is formally measured by the levels of transcript made, it has become common to measure expression indirectly by assaying the biological or biochemical activity of the protein product of the gene under investigation, whether in terms of quantifiable binding to antibodies or measurable enzymatic activities, to name the two most popular ones. Part of the reason is that assaying for protein product activity is often simpler to execute than quantitation of transcripts. The innate assumption in these studies that the level of biological activity of the protein product is directly proportional to transcript level is usually substantiated by an initial comparison between the two; that is, the relative change in the RNA transcripts should match that of the protein product. There are certain situations, however, where the proportionality could fail so that direct measurement of gene product biological activity would not represent the actual rate of transcription. For example, the stability of the protein might be disproportionate to that of the transcript, or when the activity of the protein is controlled at some other level such as transcript maturation (splicing, polyadenylation, transport from nucleus), in which case measurement of the initial rate of transcription may be absolutely necessary. There are also situations in which measurement of the transcript of the native gene cannot reliably indicate the promoter strength, such as when the protein product auto-regulates its promoter. In this scenario, it would be difficult to assess promoter strength without first preventing translation, which could lead to an even more complex situation when the promoter under investigation requires some other protein for expression.

Some of these problems have been circumvented by assessing transcription of a promoter that is fused to a bacterial gene as a reporter gene. In most cases the initial purpose of the chimeric gene is to provide an easy alternative to measuring

gene expression through enzymatic assays available for the bacterial gene, but it might very well be the most appropriate method for assessing eukaryotic promoter activity when the native product plays a role in its own regulation. At present, both the CAT and *lac*Z genes have been commonly used as reporter genes when fused with several promoters, largely because of the availability of inexpensive and simple enzymatic assays for these bacterial enzymes.

Adenovirus E1A regulatory functions are customarily assayed in DNA transfer experiments using the CAT reporter gene (Weeks and Jones,1983). An alternative assay based on the *lac*Z gene was developed (this report) mainly for the purpose of utilizing the benefits of a colourimetric assay which is currently not available for the CAT enzyme. Since a colourimetric assay is available for *lac*Z, the products of reaction are readily quantifiable by spectroscopy without any further manipulation, thereby reducing experimental error. During the enzymatic assay, the extent of reaction can be monitored visually so that incubation time can be terminated or extended as desired thereby allowing greater control of the whole process. The absence of radio-labeled substrate in the assay adds to its simplicity and cost-effectiveness. The benefits derived from the ease and simplicity of this assay encourages repetition and large data sampling which are both necessary and desirable for the accurate representation of a measured biological phenomenon.

Trans-Activation Phenotype of E1A Insertion Mutants

The *trans*-activation phenotype of the E1A insertion mutants constructed in this study pointed to the complex nature of the *trans*-activation function. Many studies have shown that the major *trans*-activation domain resides in the unique region found in the 13S E1A product (reviewed in Moran and Mathews, 1987). The same conclusions were reached by this present study which can be

summarized as follows: 1) mutants with full-length insertion in the closed orientation inactivated E1A *trans*-activation function if they were any where in exon I but had no effect in exon II; 2) mutants with full-length insertions in the open orientation exhibited wild-type or near wild-type activity except when they were located in the unique region, all three of which (nt1008, nt1039, and nt1056) failed to *trans*-activate; 3) mutants with 2aa residue insertion showed wild-type activity except two (at nt1008 and nt1056, again in the unique region) which failed to *trans*-activate; and 4) *trans*-activation of three different Ad promoters (E1A, E1B and E3) by the various mutants showed identical patterns, indicating that the same E1A functional domain is likely important for *trans*-activation of all three promoters, and possibly suggesting that the activation of these promoters may be mediated by a similar mechanism. These findings are in agreement with previous results by other investigators who have identified the unique region, using other mutational approaches, as the only essential domain for the *trans*-activation function.

The approach that was used in this study to mutate E1A allowed the opportunity to assess the sensitivity of a particular position to either 13aa or 2aa residue insertions. It was not possible to establish *a priori* the effect of these inserts on the protein structure, aside from making predictions with respect to various chemical properties such as acidity and hydrophilicity. The underlying assumption for any insertional mutagenesis is that the introduction of a localized change in the structure of the native protein would reveal the sensitivity of that particular position to alteration. In designing a collapsible mutator DNA, a basic assumption was made that the degree of sensitivity to mutation might depend on the size of the insert. Those sites that are only sensitive to a large insert could be

referred to as "slightly" sensitive, as opposed to the "highly" sensitive ones in which either size of insert would produce phenotypes that are indistinguishable from background levels. Such an assumption rests on the possibility, for example, that some higher order of protein structure contributes to function although the contribution may be minor. Such higher order structures (e.g., dimerization, complex formation) might be expected to be particularly sensitive to large inserts and that smaller inserts may not have noticeable effects. However, sites that are sensitive even to small inserts might be expected to be major functional domain(s). In reality, however, these assumptions may not apply, since the overall activity of the protein might depend almost exclusively on its folding. Therefore, depending on the residue(s) inserted, the structure of the protein may be affected greatly even by a single residue or not at all by a longer insert.

The observation that 13aa residue insertions only affected the *trans*-activation function when located in the unique region, and nowhere else, is intriguing since it suggests a) that the size of insert might be insufficient to affect the overall structure of the protein, assuming that folding is important for function, or b) that the higher order of E1A protein structure might either be absent or unimportant for *trans*-activation. One can only speculate as to the effect of an insert 13aa residues long coded by the LOP-mutator in the structure of the E1A proteins. What was observed for E1A seems to indicate then that *trans*-activation was highly resistant to the effect of 13aa residue insertions, except in the unique region. This might suggest that as long as the structure of the unique region is preserved, other parts of the protein can assume other structures without affecting function. This is consistent with the report that a synthetic oligopeptide whose primary structure corresponds to the unique region was demonstrated capable of

trans-activating the E2A promoter in microinjection studies (Green *et al.*, 1988). What these observations might predict is that the unique region may be placed in the context of a totally foreign protein and still retain its function in most cases. It has been reported, however, that an E1A gene construct in which exons I and II have been deleted failed to *trans*-activate (Jelsma *et al.*, 1988), leading to speculations that these domains, while not having *trans*-activation functions, are nevertheless important for maintaining stability of the E1A proteins when expressed *in vivo* as the relative concentration may be critical for function.

The results of the *trans*-activation assays was suggestive of an apparent division of the *trans*-activation domain into subdomains. This arises from the fact that the three mutated sites in the unique region were equally sensitive to 13aa residue inserts but not to 2aa residue inserts. One reason for this could be that these mutants differ in activity simply because they all have different sequence of inserts. The mutation at nt1008 caused an insertion of glycine-serine between residues proline$_{150}$ and glycine$_{151}$ (see **Figure 7-1**). The R-groups of glycine and serine are polar but uncharged. For position nt1039, histidine$_{160}$ and arginine$_{161}$ are separated by arginine-isoleucine: while the R-group of isoleucine is nonpolar (hydrophobic), that of arginine is positively polar with pK=12.48. The third mutation in the unique region at nt1056 causes aspartic acid$_{166}$ to be replaced by glutamic acid, a change that may not affect the overall acidity because both have negatively charged R-groups ($pK'R_{asp}$=3.86 and $pK'R_{glu}$=4.25). The insertion sequence at this position is aspartic acid-proline. In terms of changes in acidity, the mutation that is predicted to have a most dramatic effect would be at nt1039 because of the insertion of a highly basic arginine, but the *trans*-activation phenotype of this mutant is near wild-type levels, while the other two were close

Figure 7-1. **Location and Structure of 'Collapsed' Mutants in Unique Region**

The top diagram shows the structures of the 3 mutants in the unique region with 2aa residue inserts. The lower diagram shows the location of these inserts in the alignment of E1A proteins.

Figure 7-1. Location and Structure of 'Collapsed' Mutants in Unique Region

POSITION	MODIFICATION				
1008	149	150			151
	H	P	G	S	G
	CAC	CCC	GGA	TCC	GGG
	GTG	GGG	CCT	AGG	CCC
1039	159	160			161
	Y	H	R	I	R
	TAT	CAC	CGG	ATC	CGG
	ATA	GTG	GCC	TAG	GCC
1056	165	166			167
	G	E	D	P	P
	GGG	GAG	GAT	CCC	CCA
	CCC	CTC	CTA	GGG	GGT

	1008	1039	1056
Ad2	EEFVLDYVEHP	GHGCRSCHYH	RRNTGD	PDIMCSLCYMRTCGMFVYS
Ad4	ESFALDCPPLP	GHGCKSCEFH	RINTGD	KAVLCALCYMRAYNHCVYS
Ad5	EEFVLDYVEHP	GHGCRSCHYH	RRNTGD	PDIMCSLCYMRTCGMFVYS
Ad7	DVFKLDCPELP	GHGCKSCEFH	RNNTGM	KELLCSLCYMRMHCHFIYS
Ad12	EEFQLDHPELP	GHNCKSCEHH	RNSTGN	TDLMCSLCYLRAYNMFIYS
Ad40	NELVLDCPENP	GRGCRACDFH	RGTSGN	PEAMCALCYMRLTGHCIYS
Ad41	SQLVLDCPENP	GRGCRACDFH	RGSSGN	PEAMCALCYMRLTGHCIYS
AdTS	. . VFPDCPERP	GQECRSCKQH	REMSGD	PSILCSLCYMRLTACFVYS
SA7	DDFRLDCPSVP	GHGCSSCDYH	RKTSGC	PEILCSLCYLRANSMFIYS
MAV1	EVF.	. PPCSTCGGH	. . . . .	E VNGFCSLCYLRGLTGKVF
Consensus	EEF-LDCPE-P	GHGCRSC-FH	R- - TG-	PEIMCSLCYMR- - - - FIYS

UNIQUE REGION

to background levels. Therefore, the explanation for the phenotypic differences between the unique region "collapsed" mutants cannot be accounted for solely by the charge differences of the insertions.

Figure 1-2 shows the alignment of several adenovirus E1A proteins in the unique region for Ad5. As previously mentioned, the E1A phosphoproteins are proline-rich except in the unique region. For Ad5 E1A, the unique region contains only 2 proline residues (pro_{150} and pro_{167}). Both prolines are highly conserved being present in the large E1A protein of various serotypes. More importantly, the position of these prolines in the unique region is an area of the highest conservation in the protein so that it is highly probable that they contribute significantly to the structure required for the *trans*-activation function. Hence, alteration of the spatial arrangement of the components (both conformation and configuration) of this region could be expected to result in drastic effects on function. How the effect might come about is revealed by comparing the structural changes in the three mutants with small inserts in the unique region. For the mutant at nt1008, insertion of glycine-serine between the highly conserved pro_{150}-gly_{151} residues might not grossly affect the structure because the insert juxtaposed a glycine residue immediately after the proline residue, which is the natural sequence. Rather, the addition of a serine residue immediately after the glycine residue might be the ultimate cause for the abrogation of activity for this mutant. In the example of the 2aa insert at nt1039, insertion of a highly positively charged arginine residue might be expected to affect function but this was not the case as this mutant *trans*-activated at wild-type levels. This might be explained by the fact that the mutation caused an insertion of arg-ile residues between his_{160}-arg_{161}, which, although highly conserved, may not have as great an effect because the insertion is

rather similar to the natural sequence. For the mutation at nt1056, the abrogation of the *trans*-activation function by 2aa residues could be explained by the fact that it caused the change from highly conserved aspartic $acid_{166}$ to glutamic $acid_{166}$. It is highly possible then that asp_{166} is involved directly in the function. It also caused the duplication of the proline residue in tandem by the insertion of asp-pro, in which case the extra proline residue might have considerable effect on structure. Thus, the effect on function by the 13aa or 2aa insertion can be summarized as follows. Insertions of 13 aa residues in the three sites in the unique region may have resulted in the *trans*-activation-negative phenotype probably because the structure of the unique region that is required for the activity was severely disrupted. When the insert was reduced in size to 2 residues, the elimination of activity in two out three cases could be attributed to the disruption of the structure normally maintained by highly conserved residues.

The distribution of the mutants, being mostly generated by insertion into *Hpa*II sites, is biased towards C+G rich regions and in particular to proline-rich regions of E1A coding sequence as a consequence of the CCGG recognition site and the fact that proline is encoded by CCN (N being any base). Proline, being an imino acid in which the N atom is part of a rigid ring, distorts the regular conformation of the polypeptide chain by reducing the allowable angle of rotation of the ring N-C bond. Thus, the presence of prolines in many parts of the E1A molecule likely makes it a relatively rigid molecule particularly in proline-rich regions. Thus, one may imagine the structure of the large E1A protein as consisting of a highly acidic unique region flanked by structures that are relatively rigid by virtue of proline residues. Insertions into the flanking regions have little effect on the overall structure since the insert may prove insufficient to affect a

rigid structure, especially so when most insertions maintain the proline codon. Previous studies have indicated that these flanking domains do not contribute significantly to the *trans*-activation function, hence it is perhaps not surprising in retrospect that distortion of their structures by insertions of as long as 13 residues has little effect on *trans*-activation.

TRANS-ACTIVATION OF E1A, E1B AND E3 PROMOTERS

Trans-activation of the three viral promoters, E1A, E1B and E3, as assayed with the bacterial *lac*Z gene was similar qualitatively but not quantitatively. It was shown earlier that the level of induction by the wild type E1A gene (expressing both 13S and 12S) was 9- to 13-fold for both E1B and E3 promoters and only 4-6 fold for E1A. The quantitative difference may be attributed to the assay used as one possible explanation, since *trans*-activation was measured with plasmids in which no attempt was made to allow the exclusive expression of 13S product. That is, because the 12S product exerts a *trans*-repression effect on the E1A enhancer, the full effect of the *trans*-activating 13S product may not be revealed when the assay is carried out on the E1A promoter. If this hypothesis is correct then one might expect that a similar level (9-13 fold) of induction of the E1A promoter would be observed when the gene expressing only the 13S product was used instead. However, this hypothesis fails to reconcile the observation that mutants at positions nt908 and nt1267, which were *trans*-repression defective, did not show over-induction of the pE1A*lac*Z reporter gene. The difference was in the basal level where E1A promoter activity was consistently 2 to 3 fold higher than that of E1B or E3. It appears, therefore, that the *trans*-activation function of E1A was not to multiply the activity of an inducible promoter by a certain fixed factor but rather

to maximize it to a certain fixed level. This explanation seems reasonable because it would mean that the limit is determined by the capacity of the recipient cells to support transcription rather than by the activity of the inducer.

The observation that all *lac*Z reporter plasmids placed under the individual control of the E1A, E1B and E3 promoters responded in a virtually identical manner to the mutations in E1A suggests that a single mechanism may be operational for the E1A-mediated activation of these promoters, at least in gene transfer experiments. As noted in the **INTRODUCTION**, these promoters are similar as far as the presence of a typical TATA box is concerned but differ in terms of sequence and composition of other cellular factor binding sites. The simplest of these three promoters appears to be that of E1B since its basal activity and E1A-inducibility can be localized within a region that includes only the binding sites for Sp1 and the TATA-box factor (TFIID), and interestingly mutations that affect binding of factors to the TATA box are the only ones that affect inducibility of this promoter by E1A (Wu *et al.*, 1987). Hence it is likely that E1A-inducibility of the three promoters used in the *trans*-activation assays described here is mediated by a common mechanism that involves only the TATA box, in which case, transient expression assays may not be a true measure of *in vivo* activity. This is entirely possible since class III genes such as the *Drosophila* $tRNA^{arg}$ are stimulated only when newly introduced but not when within a chromosomal context (Gaynor *et al.*, 1985) suggesting that the chromosomal context may greatly influence responsiveness to E1A.

ENHANCER *TRANS*-REPRESSION

It appeared that the only way to assay the enhancer repression function was

to prevent translation of the 13S message, one consideration that is important in the interpretation of this particular E1A function (see below). The enhancer repression function of the E1A mutants was assayed by first constructing a second mutation in each mutant so that only the 12S product was expressed. This was done by inserting a stop codon in the unique region that results in a truncated 289R polypeptide product.

When such E1A double mutants were assayed for their ability to repress the expression of the *lac*Z reporter gene under the control of the E1A enhancer/promoter, two sites, nt908 and nt1267, were found to be particularly sensitive to a 13aa residue insertion (both insertions are in translational reading frame II; see **Table 4-3**). The large insert at nt908 caused total elimination of ability to repress but activity was restored when the insert was reduced in size to 2aa residues. This is in contrast to the mutations at nt1267 in which both large and small insertions reduced the activity to about 20% wild type. These observations are not easily interpretable in the manner that was proposed for the effect seen on *trans*-activation with the unique region mutants. For example, it is not readily obvious why nt908, but not nt906, was found to be sensitive to the large insert, especially since both contain identical inserts (i.e., in identical reading frame) and that both insertions occur in an area which is highly conserved. Previous studies have indicated that the conserved regions 5' of the unique region are important for the *trans*-repression function. In this study, it was shown that at least a small region after the unique region showed a measurable degree of sensitivity at the nt1267 position. This region is proline-rich, being found in most adenoviruses except Ad7 and simian Ad7, among those whose sequences are known. These results agree with the findings of Subramanian *et al.* (1988) and Velcich and Ziff

(1988) who have identified a region downstream from the unique region which showed sensitivity to a small deletion when assayed for the *trans*-repression function. In the 243R polypeptide, which is the major trans-repressor, the two sites at nt908 and nt1267 are immediately adjacent to each other since the unique region is not present. Thus it is most likely that at least a small region in exon II contributes to the *trans*-repression domain.

It has been implicit in the assay used here that only the 12S product has significant contribution to the *trans*-repression function, the 13S being largely ignored since its *trans*-activation function either masks or lacks the repression activity. Removal of the relatively proline-free unique region brings together the flanking regions which are rich in prolines. The 243 aa residue protein is very high in proline residues (18.1% for Ad5) which would be expected to provide a very rigid molecular structure. Furthermore, proline residues, aside from being powerful eliminators of secondary structures (both alpha-helix and beta-sheet) (Chou and Fasman, 1978), also undergo isomerization about the *cis-trans* conformation, making it likely that the conformation of the protein is not limited to one or a few, but many structures. The complexity of E1A structure is further increased by various post-translational modifications most notably phosphorylation, the effect of which on the structure of an unrelated protein was recently established (Sprang *et al.*, 1988). Perhaps it is more appropriate to describe the structure of these proteins as variable.

TRANS-ACTIVATION, ENHANCER REPRESSION AND VIRAL INFECTIVITY

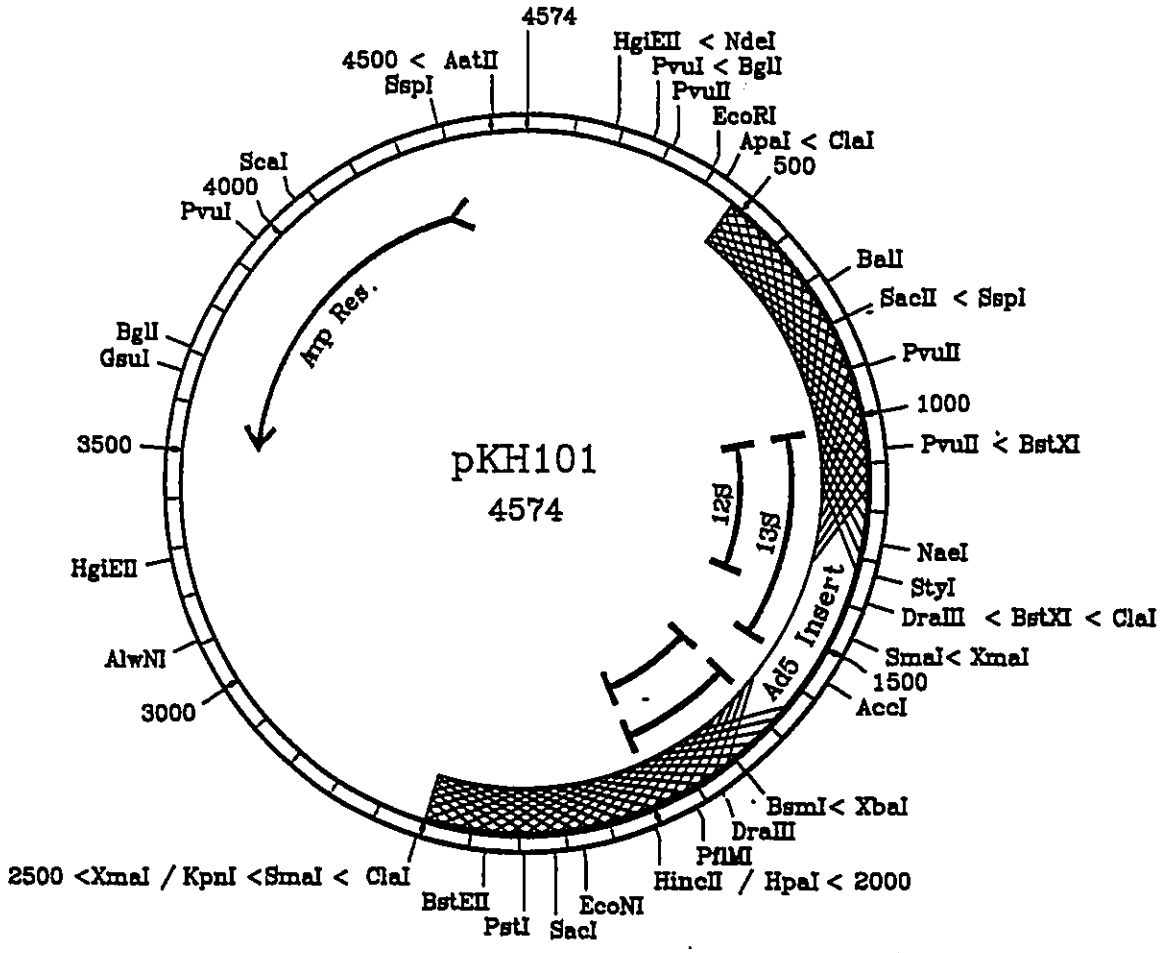
The E1A insertion mutants have been reconstructed into intact viruses (McGrory, 1988) using a homologous recombination approach (McGrory *et al.*, 1988). In infectivity assays using either HeLa or 293 cells, it was shown that the pattern of plaquing efficiency in HeLa cells of the mutant viruses was in perfect correlation with the ability of the corresponding E1A to *trans*-activate in gene transfer experiments, leaving no doubt that the *trans*-activation function plays a major role in the infectivity of the virus in cell culture. When a similar comparison was made between infectivity and ability to *trans*-repress the E1A enhancer, it was found that no lesion in E1A affected infectivity and at the same time *trans*-repression function. These results may not necessarily indicate that the enhancer repression function is not important in the life cycle of the virus since plaquing efficiency and not the kinetics of viral growth was the basis of comparison. That is, if the repression function plays only a minor role in viral infectivity, it might only affect time-dependent events in the viral cycle in subtle ways (e.g., rate of viral growth) and not necessarily burst-size. There also exists the possibility that this function is manifested only in certain situations. For example, in order to be properly measured in transfection experiments using HeLa cells, it was obligatory to prevent the expression of the 13S product, one situation which undoubtedly would severely impair viral growth and infectivity if the assay were done in intact viruses. Similarly, in intact viruses the *trans*-repression function is manifested only when 12S alone is expressed, making it necessary to grow such viruses in cells that complement E1A defects. These might suggest that *trans*-repression is a relatively unimportant function. However, it is possible that there is a cell-cycle

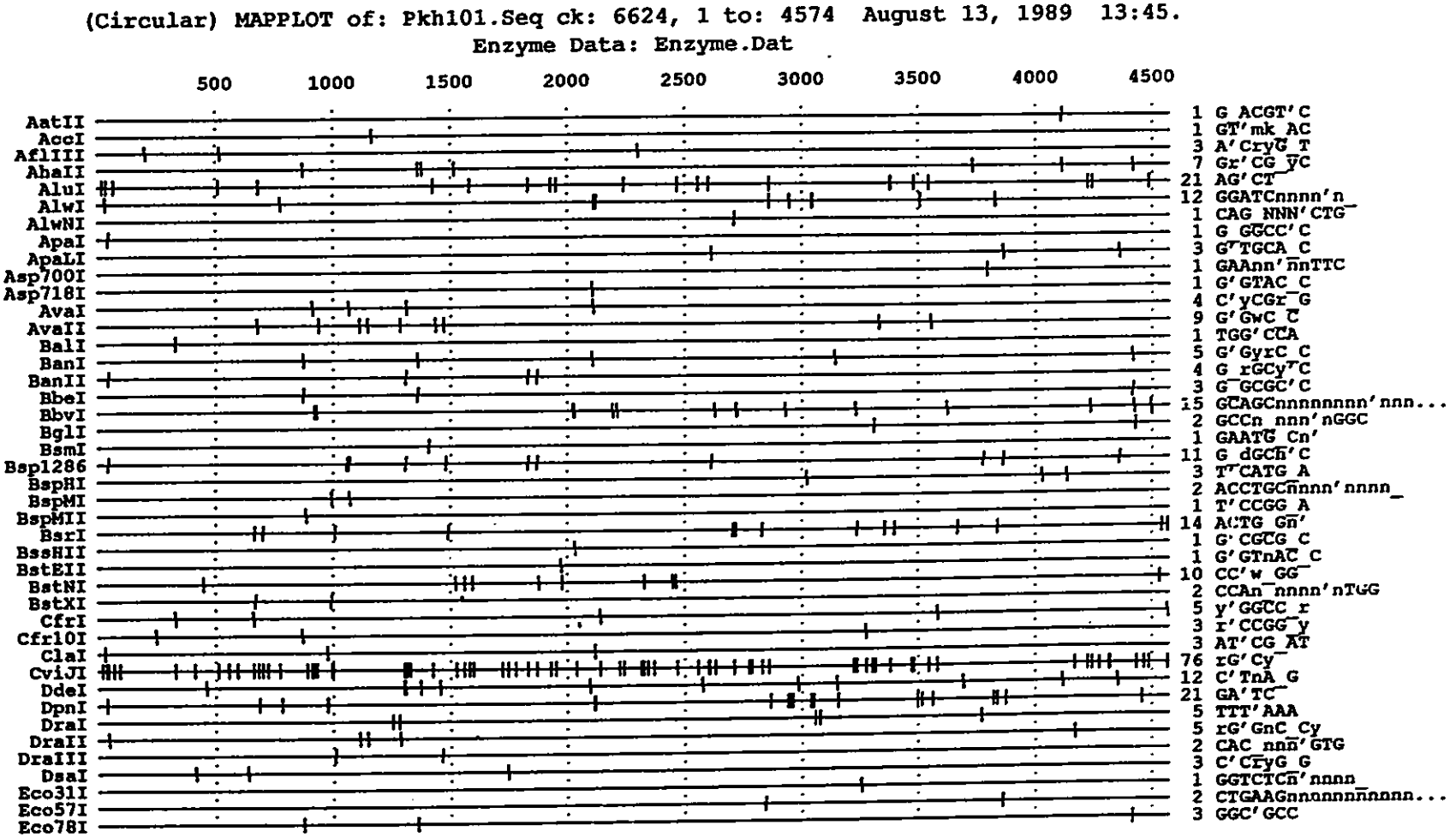
dependence for the function to be manifested in the presence of the 13S product such as in the course of natural infection where terminally-differentiated, G_0-arrested cells are the natural hosts to the virus. Such function might be needed for viral replication or in establishing latency under those conditions. As noted in the **INTRODUCTION**, induction of DNA synthesis and cell cycle progression are two other functions ascribed to the E1A region, in particular the 12S product (reviewed in Moran and Mathews, 1987). The requirement of these two functions, and possibly *trans*-repression by virtue of its linkage to the 12S product, in proliferating cells in culture may be reduced or nil and, during infection of such cells, the *trans*-repression function may be concealed by the antagonistic *trans*-activation function. In addition, that the *trans*-repression function is required to regulate transcription, specifically that of E1A, is equally possible. Auto-regulation at the level of transcription is perhaps an important function for the maintenance of a balanced expression of transcription units.

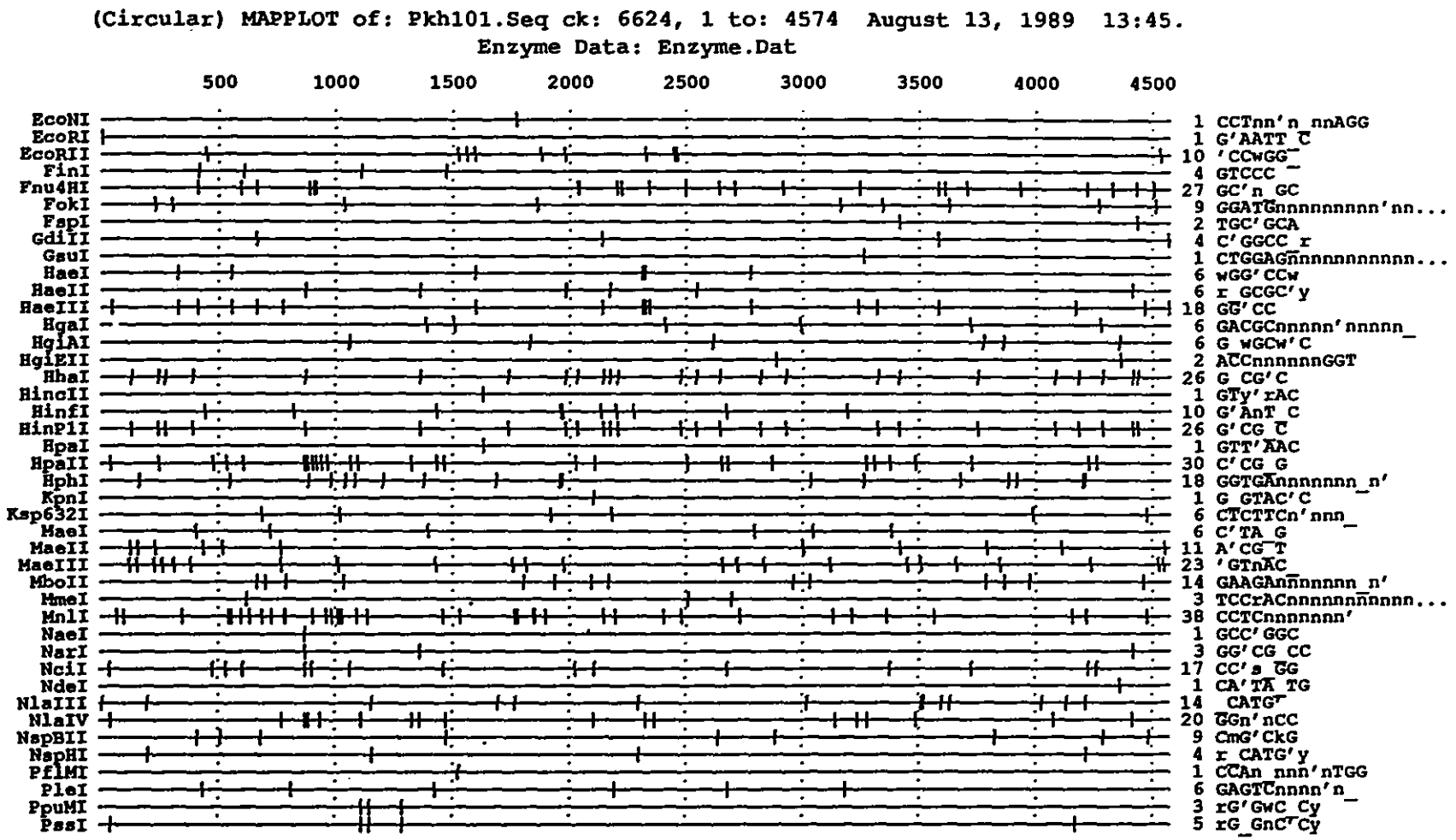
APPENDIX A. SEQUENCE AND RESTRICTION MAP OF pKH101

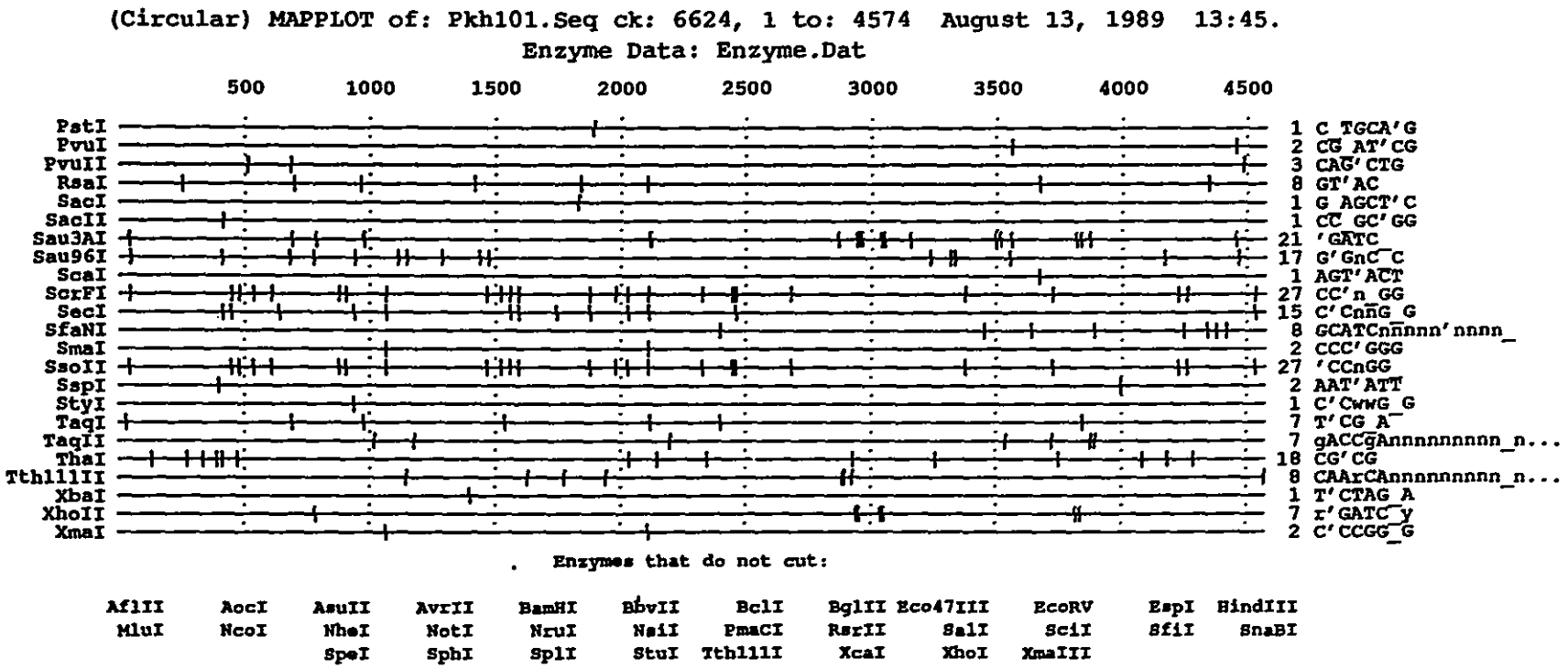
The assembled sequence of pKH101 starts at the unique *Eco*RI site. The complete sequence contains:

1 - 41	-*Eco*RI-*Bam*HI of pBR322; *Hin*dIII-*Bam*HI deleted by Klenow treatment
42 - 78	-unknown sequence introduced through cloning of Ad5 DNA (McKinnon and Graham, pers. comm., and this study)
79 - 2110	-Ad5 sequence from nt22 to nt2048 (*Kpn*I site)
2111 - 4574	-pUC18 sequences




(Circular) MAPPLOT of: Pkh101.Seq ck: 6624, 1 to: 4574 August 13, 1989 13:45.
Enzyme Data: Enzyme.Dat
500
1000
1500
2000
2500
3000
3500
4000
4500
AatII 1 G ACGT'C
AccI 1 GT'mk AC
AflIII 3 A'CryG T
AhaII 7 Gr'CG yC
AluI 21 AG'CT
AlwI 12 GGATCnnnn'n
AlwNI 1 CAG NNN'CTG
ApaI 1 G GGCC'C
ApaLI 3 G'TGCA C
Asp700I 1 GAAnn'nnTTC
Asp718I 1 G'GTAC C
AvaI 4 C'yCGr G
AvaII 9 G'GwC C
BalI 1 TGG'CCA
BanI 5 G'GyrC C
BanII 4 G rGCy'C
BbeI 3 G GCGC'C
BbvI 15 GCAGCnnnnnnnn'nnn...
BglI 2 GCCn nnn'nGGC
BsmI 1 GAATG Cn'
Bsp1286 11 G dGCh'C
BspHI 3 T'CATG A
BspMI 2 ACCTGCnnnn'nnnn
BspMII 1 T'CCGG A
BsrI 14 ACTG Gn'
BssHII 1 G'CGCG C
BstEII 1 G'GTnAC C
BstNI 10 CC'w GG
BstXI 2 CCAn nnnn'nTGG
CfrI 5 y'GGCC r
Cfr10I 3 r'CCGG y
ClaI 3 AT'CG AT
CviJI 76 rG'Cy
DdeI 12 C'TnA G
DpnI 21 GA'TC
DraI 5 TTT'AAA
DraII 5 rG'GnC Cy
DraIII 2 CAC nnn'GTG
DsaI 3 C'CryG G
Eco31I 1 GGTCTCn'nnnn
Eco57I 2 CTGAAGnnnnnnnnnnnn...
Eco78I 3 GGC'GCC


(Circular) MAPPLOT of: Pkh101.Seq ck: 6624, 1 to: 4574 August 13, 1989 13:45.
Enzyme Data: Enzyme.Dat
500
1000
1500
2000
2500
3000
3500
4000
4500
EcoNI 1 CCTnn'n_nnAGG
EcoRI 1 G'AATT_C
EcoRII 10 'CCwGG_
FinI 4 GTCCC
Fnu4HI 27 GC'n_GC
FokI 9 GGATGnnnnnnnnn'nn...
FspI 2 TGC'GCA
GdiII 4 C'GGCC_r
GsuI 1 CTGGAGnnnnnnnnnnn...
HaeI 6 wGG'CCw
HaeII 6 r_GCGC'y
HaeIII 18 GG'CC
HgaI 6 GACGCnnnnn'nnnnn_
HgiAI 6 G_wGCw'C
HgiEII 2 ACCnnnnnnGGT
HhaI 26 G_CG'C
HincII 1 GTy'rAC
HinfI 10 G'AnT_C
HinPII 26 G'CG_C
HpaI 1 GTT'AAC
HpaII 30 C'CG_G
HphI 18 GGTGAnnnnnnn_n'
KpnI 1 G_GTAC'C
Ksp632I 6 CTCTTCn'nnn_
MaeI 6 C'TA_G
MaeII 11 A'CG_T
MaeIII 23 'GTnAC
MboII 14 GAAGAnnnnnnn_n'
MmeI 3 TCCrACnnnnnnnnnnn...
MnlI 38 CCTCnnnnnnn'
NaeI 1 GCC'GGC
NarI 3 GG'CG_CC
NciI 17 CC's_GG
NdeI 1 CA'TA_TG
NlaIII 14 _CATG'
NlaIV 20 GGn'nCC
NspBII 9 CmG'CkG
NspHI 4 r_CATG'y
PflMI 1 CCAn_nnn'nTGG
PleI 6 GAGTCnnnn'n_
PpuMI 3 rG'GwC_Cy
PssI 5 rG_GnC'Cy


(Circular) MAPPLOT of: Pkh101.Seq ck: 6624, 1 to: 4574 August 13, 1989 13:45.
Enzyme Data: Enzyme.Dat
500
1000
1500
2000
2500
3000
3500
4000
4500
PstI 1 C TGCA'G
PvuI 2 CG AT'CG
PvuII 3 CAG'CTG
RsaI 8 GT'AC
SacI 1 G AGCT'C
SacII 1 CC GC'GG
Sau3AI 21 'GATC
Sau96I 17 G'GnC C
ScaI 1 AGT'ACT
ScrFI 27 CC'n GG
SecI 15 C'CnnG G
SfaNI 8 GCATCnnnnn'nnnn
SmaI 2 CCC'GGG
SsoII 27 'CCnGG
SspI 2 AAT'ATT
StyI 1 C'CwwG G
TaqI 7 T'CG A
TaqII 7 gACCGAnnnnnnnnn n...
ThaI 18 CG'CG
Tth111II 8 CAArCAnnnnnnnnn n...
XbaI 1 T'CTAG A
XhoII 7 r'GATC y
XmaI 2 C'CCGG G
Enzymes that do not cut:
AflII MluI
AocI NcoI
AsuII NheI SpeI
AvrII NotI SphI
BamHI NruI SplI
BbvII NsiI StuI
BclI PmaCI Tth111I
BglII RsrII XcaI
Eco47III SalI XhoI
EcoRV SciI XmaIII
EspI SfiI
HindIII SnaBI

pKH101.seq Length: 4574 November 16, 1988 19:49 Check: 6624 ..

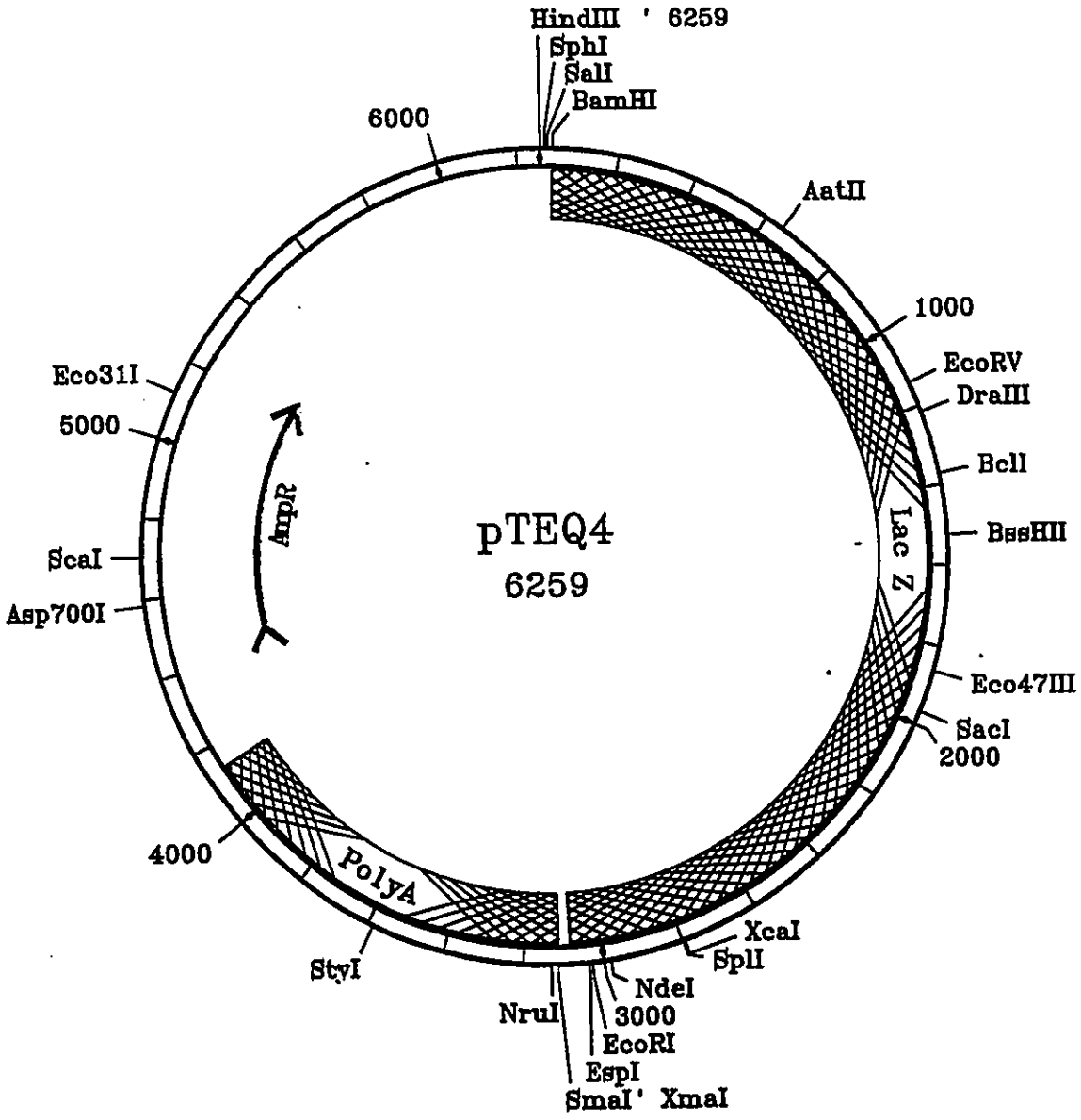
```
   1  GAATTCTCAT GTTTGACAGC TTATCATCGA TAAGCTGATC CGGGCCCCCA
  51  TTTCCCCTCC CTTCCAGCTC TCTGCCCCTT TTGGATTGAA GCCAATATGA
 101  TAATGAGGGG GTGGAGTTTG TGACGTGGCG CGGGGCGTGG GAACGGGGCG
 151  GGTGACGTAG TAGTGTGGCG GAAGTGTGAT GTTGCAAGTG TGGCGGAACA
 201  CATGTAAGCG ACGGATGTGG CAAAAGTGAC GTTTTTGGTG TGCGCCGGTG
 251  TACACAGGAA GTGACAATTT TCGCGCGGTT TTAGGCGGAT GTTGTAGTAA
 301  ATTTGGGCGT AACCGAGTAA GATTTGGCCA TTTTCGCGGG AAAACTGAAT
 351  AAGAGGAAGT GAAATCTGAA TAATTTTGTG TTACTCATAG CGCGTAATAT
 401  TTGTCTAGGG CCGCGGGGAC TTTGACCGTT TACGTGGAGA CTCGCCCAGG
 451  TGTTTTTCTC AGGTGTTTTC CGCGTTCCGG GTCAAAGTTG GCGTTTTATT
 501  ATTATAGTCA GCTGACGTGT AGTGTATTTA TACCCGGTGA GTTCCTCAAG
 551  AGGCCACTCT TGAGTGCCAG CGAGTAGAGT TTTCTCCTCC GAGCCGCTCC
 601  GACACCGGGA CTGAAAATGA GACATATTAT CTGCCACGGA GGTGTTATTA
 651  CCGAAGAAAT GGCCGCCAGT CTTTTGGACC AGCTGATCGA AGAGGTACTG
 701  GCTGATAATC TTCCACCTCC TAGCCATTTT GAACCACCTA CCCTTCACGA
 751  ACTGTATGAT TTAGACGTGA CGGCCCCCGA AGATCCCAAC GAGGAGGCGG
 801  TTTCGCAGAT TTTTCCCGAC TCTGTAATGT TGGCGGTGCA GGAAGGGATT
 851  GACTTACTCA CTTTTCCGCC GGCGCCCGGT TCTCCGGAGC CGCCTCACCT
 901  TTCCCGGCAG CCCGAGCAGC CGGAGCAGAG AGCCTTGGGT CCGGTTTCTA
 951  TGCCAAACCT TGTACCGGAG GTGATCGATC TTACCTGCCA CGAGGCTGGC
1001  TTTCCACCCA GTGACGACGA GGATGAAGAG GGTGAGGAGT TTGTGTTAGA
1051  TTATGTGGAG CACCCCGGGC ACGGTTGCAG GTCTTGTCAT TATCACCGGA
1101  GGAATACGGG GGACCCAGAT ATTATGTGTT CGCTTTGCTA TATGAGGACC
1151  TGTGGCATGT TTGTCTACAG TAAGTGAAAA TTATGGGCAG TGGGTGATAG
1201  AGTGGTGGGT TTGGTGTGGT AATTTTTTTT TTAATTTTTA CAGTTTTGTG
1251  GTTTAAAGAA TTTTGTATTG TGATTTTTTT AAAAGGTCCT GTGTCTGAAC
1301  CTGAGCCTGA GCCCGAGCCA GAACCGGAGC CTGCAAGACC TACCCGCCGT
1351  CCTAAAATGG CGCCTGCTAT CCTGAGACGC CCGACATCAC CTGTGTCTAG
1401  AGAATGCAAT AGTAGTACGG ATAGCTGTGA CTCCGGTCCT TCTAACACAC
1451  CTCCTGAGAT ACACCCGGTG GTCCCGCTGT GCCCCATTAA ACCAGTTGCC
1501  GTGAGAGTTG GTGGGCGTCG CCAGGCTGTG GAATGTATCG AGGACTTGCT
1551  TAACGAGCCT GGGCAACCTT TGGACTTGAG CTGTAAACGC CCCAGGCCAT
1601  AAGGTGTAAA CCTGTGATTG CGTGTGTGGT TAACGCCTTT GTTTGCTGAA
```

1651 TGAGTTGATG TAAGTTTAAT AAAGGGTGAG ATAATGTTTA ACTTGCATGG
1701 CGTGTTAAAT GGGGCGGGGC TTAAAGGGTA TATAATGCGC CGTGGGCTAA
1751 TCTTGGTTAC ATCTGACCTC ATGGAGGCTT GGGAGTGTTT GGAAGATTTT
1801 TCTGCTGTGC GTAACTTGCT GGAACAGAGC TCTAACAGTA CCTCTTGGTT
1851 TTGGAGGTTT CTGTGGGGCT CATCCCAGGC AAAGTTAGTC TGCAGAATTA
1901 AGGAGGATTA CAAGTGGGAA TTTGAAGAGC TTTTGAAATC CTGTGGTGAG
1951 CTGTTTGATT CTTTGAATCT GGGTCACCAG GCGCTTTTCC AAGAGAAGGT
2001 CATCAAGACT TTGGATTTTT CCACACCGGG GCGCGCTGCG GCTGCTGTTG
2051 CTTTTTTGAG TTTTATAAAG GATAAATGGA GCGAAGAAAC CCATCTGAGC
2101 GGGGGGTACC CGGGGATCGA TCCCTGCATT AATGAATCGG CCAACGCGCG
2151 GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA
2201 CTCGCTGCGC TCGGTCGTTC GGCTGCGGCG AGCGGTATCA GCTCACTCAA
2251 AGGCGGTAAT ACGGTTATCC ACAGAATCAG GGGATAACGC AGGAAAGAAC
2301 ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA AGGCCGCGTT
2351 GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC
2401 GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG
2451 GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC
2501 GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT
2551 CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC
2601 AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC GCTGCGCCTT
2651 ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC
2701 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC
2751 GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG
2801 GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA
2851 GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT
2901 TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG GATCTCAAGA
2951 AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
3001 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG
3051 ATCCTTTTAA ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA
3101 GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT
3151 CAGCGATCTG TCTATTTCGT TCATCCATAG TTGCCTGACT CCCCGTCGTG
3201 TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT
3251 GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
3301 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC

3351 TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC
3401 AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT
3451 CACGCTCGTC GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA
3501 AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG TTAGCTCCTT
3551 CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
3601 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA
3651 TGCTTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG
3701 TATGCGGCGA CCGAGTTGCT CTTGCCCGGC GTCAATACGG GATAATACCG
3751 CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG
3801 GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA
3851 ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
3901 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA
3951 AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA
4001 TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT
4051 GTATTTAGAA AAATAAACAA ATAGGGGTTC CGCGCACATT TCCCCGAAAA
4101 GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT TAACCTATAA
4151 AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
4201 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG
4251 TAAGCGGATG CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT
4301 TGGCGGGTGT CGGGGCTGGC TTAACTATGC GGCATCAGAG CAGATTGTAC
4351 TGAGAGTGCA CCATATGCGG TGTGAAATAC CGCACAGATG CGTAAGGAGA
4401 AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA ACTGTTGGGA
4451 AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG
4501 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCCAGTCA
4551 CGACGTTGTA AAACGACGGC CAGT

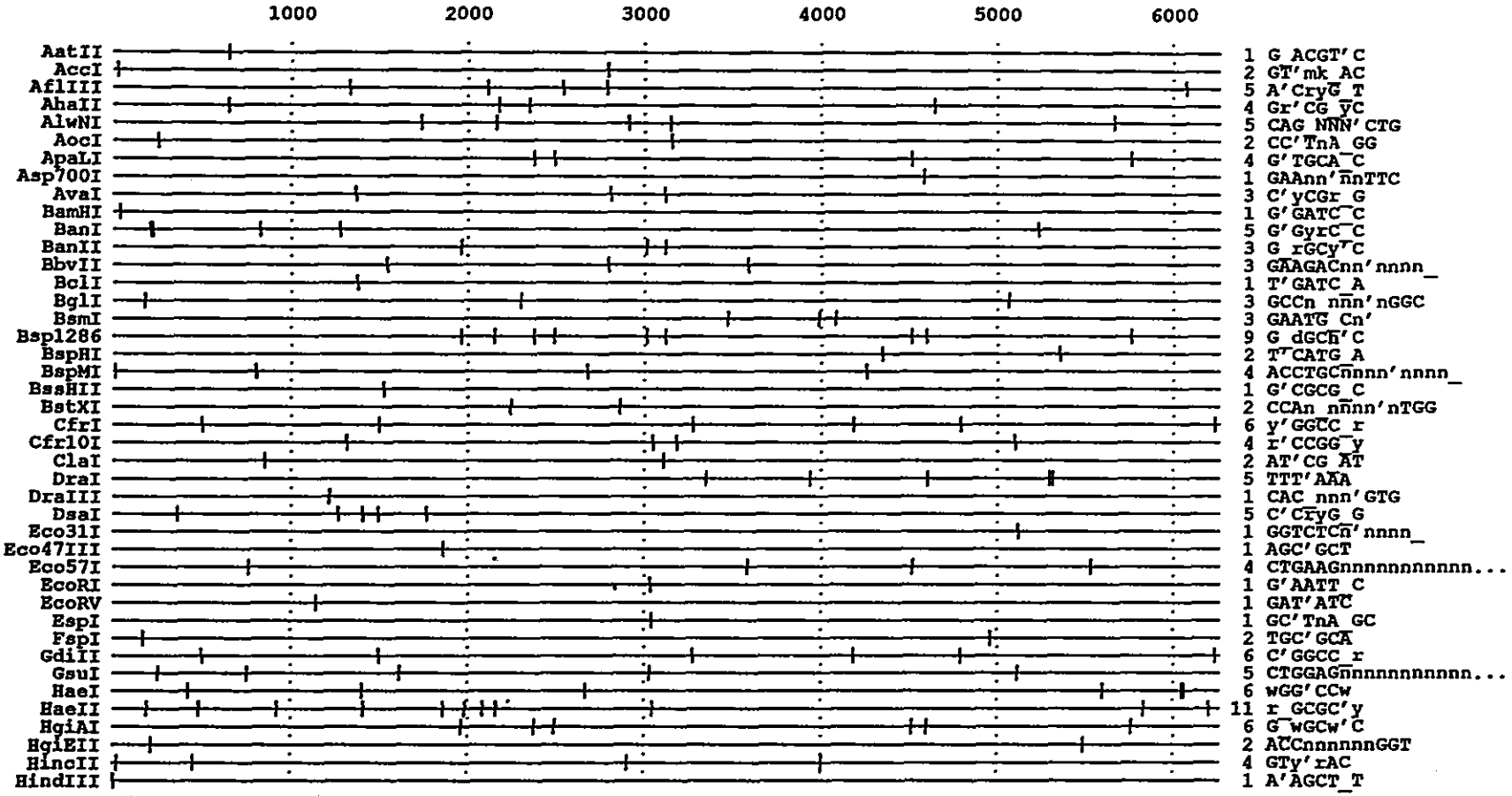
APPENDIX B. RESTRICTION MAP OF pTEQ4

The circular map of pTEQ4 is shown in the diagram. Unique restriction sites are indicated. The complete restriction map using enzymes with a recognition sequence that is 6-bases or more in length are shown in the following pages.



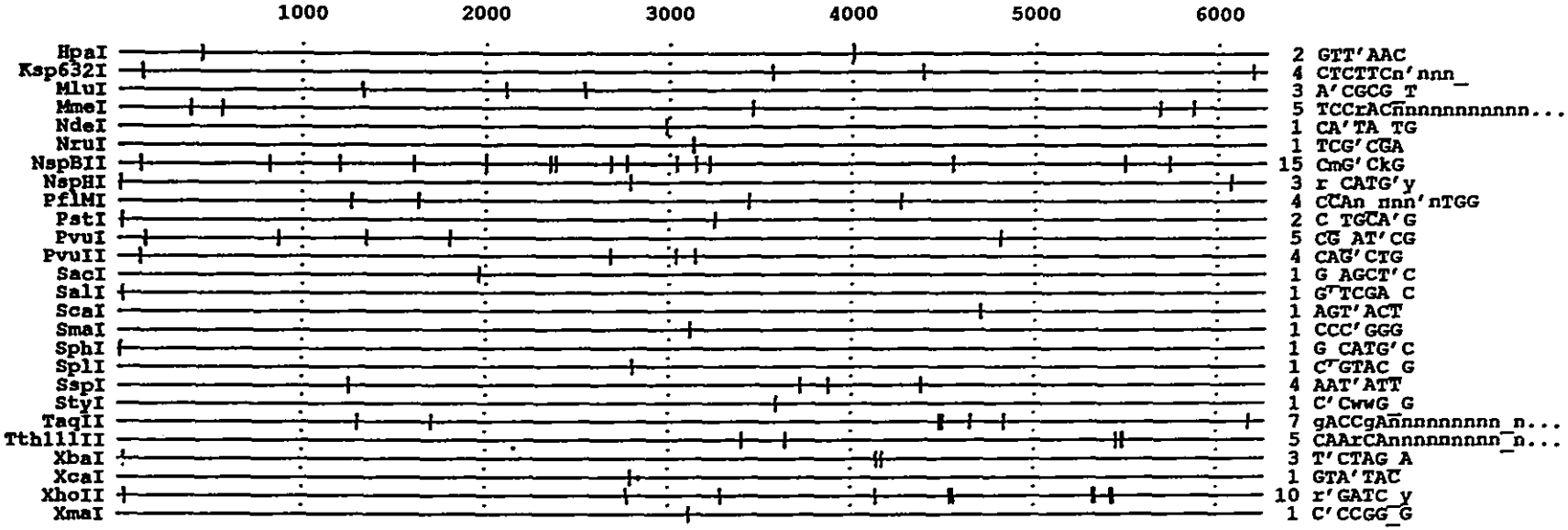
(Circular) (Six-Base) MAPPLOT of: Pteq4.Seq ck: 1767, 1 to: 6259 August 8, 1989 16:25.

Enzyme Data: Enzyme.Dat



(Circular) (Six-Base) MAPPLOT of: Pteq4.Seq ck: 1767, 1 to: 6259 August 8, 1989 16:25.

Enzyme Data: Enzyme.Dat



Enzymes that do not cut:

AflII	ApaI	Asp718I	AsuII	AvrII	BalI	BbeI	BglII	BspMII	BstEII	DraII	Eco78I
EcoNI	KpnI	NaeI	NarI	NcoI	NheI	NotI	NsiI	PmaCI	PpuMI	PssI	RsrII
	SacII	SciI	SfiI	SnaBI	SpeI	StuI	Tth111I	XhoI	XmaIII		

APPENDIX C. CONSTRUCTION OF OTHER PLASMIDS

The construction of the following plasmids are shown below:

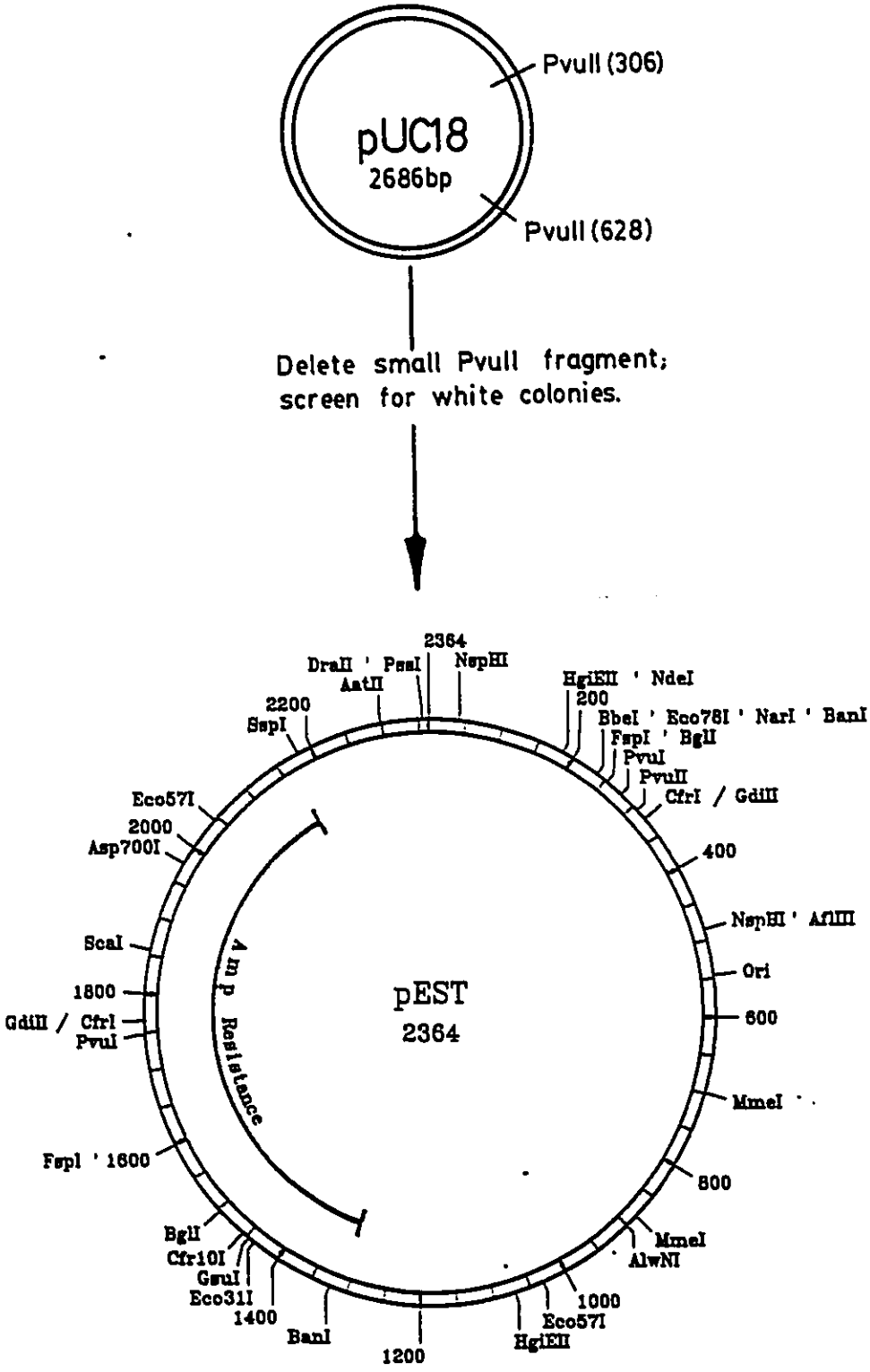
1. **pEST**
2. **pESTLOP**
3. **pSEQ18**
4. **pMCS101**
5. **pHE38** and **pXC38**

The assembled sequences of these and other plasmids whose constructions have been described elsewhere in the thesis are now stored in a floppy disk for the IBM PC and compatibles in both the University of Wisconsin GCG package and ASCII formats. These could be obtained from the author or from the supervisor, Dr. F. L. Graham.

1. CONSTRUCTION OF pEST

A *Pvu*II fragment containing the multiple cloning site of pUC18 was deleted. The resulting plasmid, pEST, produced colorless colonies in Xgal plates (in LB or B medium) using hosts such as JM83 since its α-complementation function has been deleted, as well as in Xgal-glucose plates (in minimal medium) using hosts such as E5014 and LE392 since the operator has been removed. It was used primarily in studies to optimize the conditions for inserting the LOP-cassette into multicopy plasmids (see **Chapter III**). The exact sequence may be assembled from GENBANK's file VEC:PUC18C by deleting nt307-nt628 inclusive. The plasmid is exactly 2364bp long.

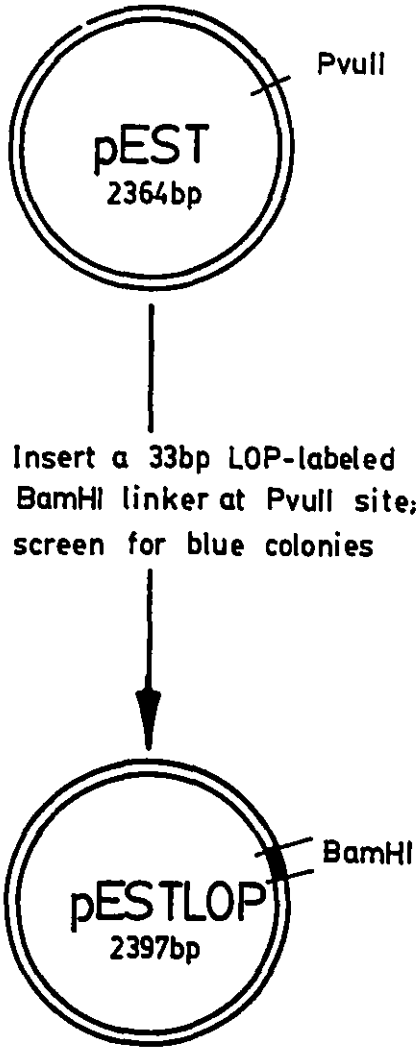
1. CONSTRUCTION OF pEST



2. CONSTRUCTION OF pESTLOP

The plasmid pESTLOP was constructed by inserting a 33-bp linker at the *Pvu*II site of pEST. The linker consisted of three oligomers (SAM67, SAM68 and SAM69, see *Table 2-2*, page 57) with a core containing the *lacZo* sequence flanked by *Bam*HI sites, and its insertion was screened using the LOP-labeling method described in **Chapter III.** The insert could assume two orientations; however, this was not determined for pESTLOP. The plasmid was used subsequently for constructing other plasmids as noted elsewhere in the text.

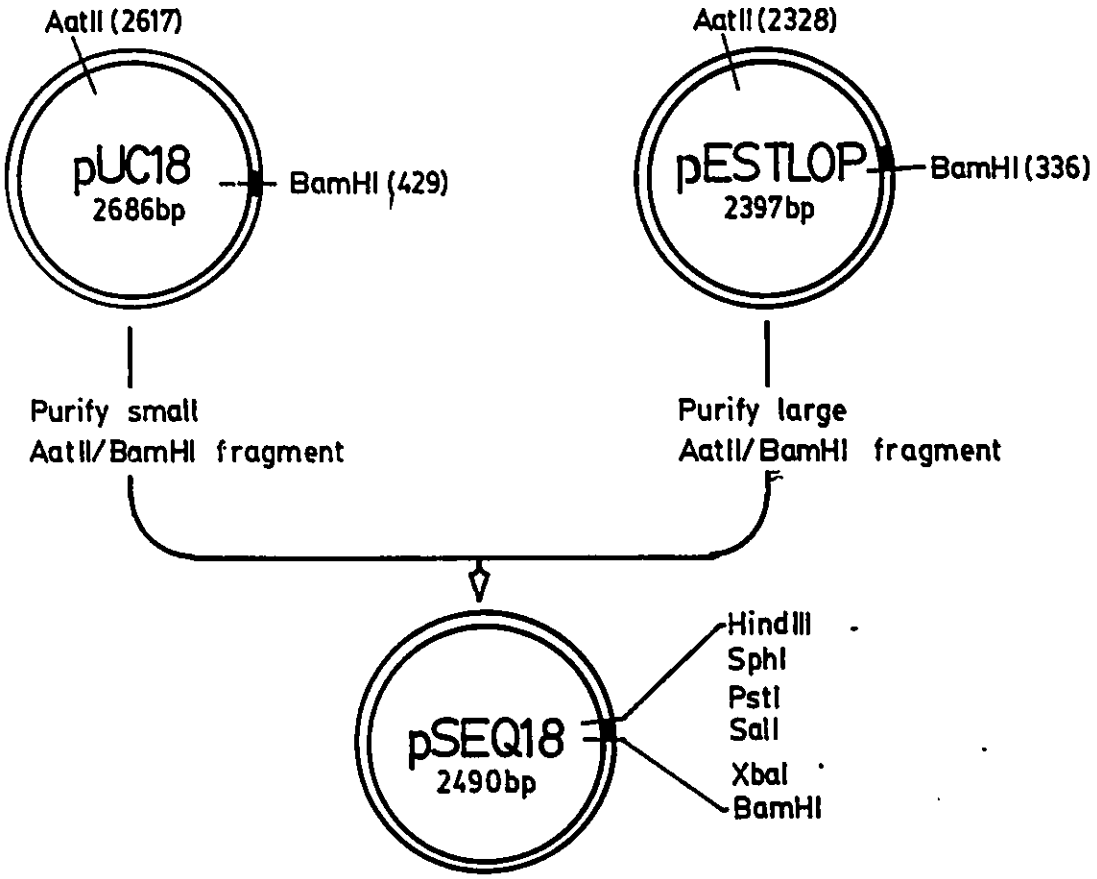
2. CONSTRUCTION OF pESTLOP



3. CONSTRUCTION OF pSEQ18

The plasmid pSEQ18 was constructed by fusing the small *Aat*II/*Bam*HI fragment of pUC18 and the large *Aat*II/*Bam*HI fragment of pESTLOP to produce a plasmid that was 2490bp in size. The construct contained a partial multiple cloning site of pUC18 (from *Bam*HI to *Hin*dIII) and was originally intended for sequencing LOP-labeled fragments as follows: First, the small LOP-labeled fragment generated by digestion of the source plasmid with one or a combination of multicut enzymes that generate blunt termini is shotgun-cloned into the unique *Hin*cII site of pSEQ18. The desired clone would appear as a blue colony (due to the LOP sequence) in Xgal-glucose minimal medium plates. Sequencing is done on double-stranded templates using primers on either side of the multiple cloning site. Although in a few occasions in this study this procedure was successfully used to determine the structure of some mutants, it proved to have limited usefulness since many times the region of interest to be sequenced was at a distance from the primer where resolution on the sequencing gel was poor. It was later abandoned for a method in which the candidate mutant was sequenced from within the insert as described in **Chapter III.** Since pSEQ18 does not have the *lacZo* sequence, it could be used for cloning fragments for mutagenesis using the LOP-cassette method.

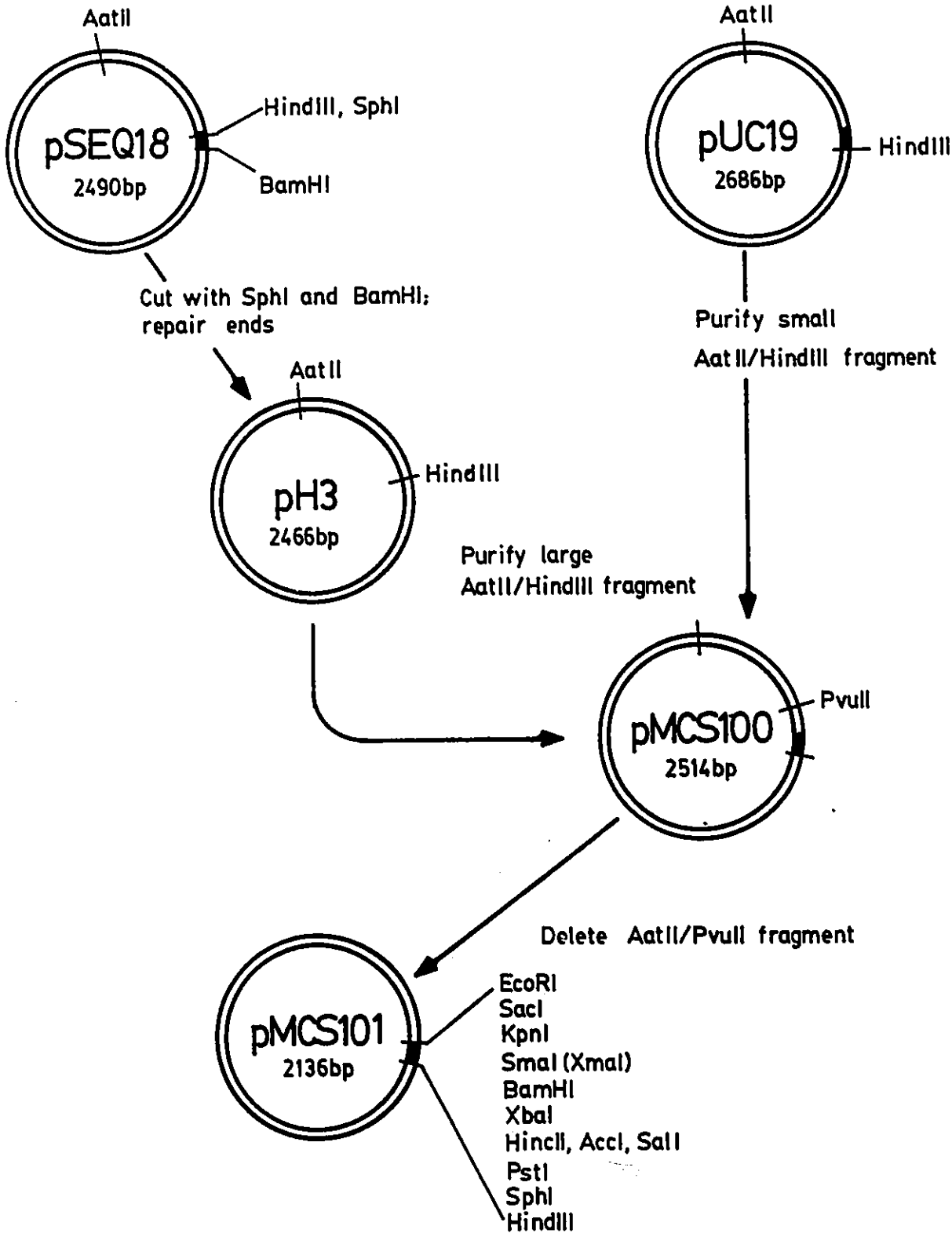
3. CONSTRUCTION OF pSEQ18



4. CONSTRUCTION OF pMCS101

A cloning vector containing the entire multiple cloning site of the pUC18/19 vectors but without the *lacZo* sequence was constructed as an alternative to pSEQ18. First, both SphI and BamHI were used to cleave pSEQ18 and then the termini were blunt-ended with Klenow and dNTP's. The resulting plasmid, pH3 (2466bp) was then cleaved with AatII and HindIII and the large fragment was purified. This was ligated with a small AatII/HindIII fragment from pUC19 resulting in pMCS100 (2514bp). To reduce the size, an AatII/PvuII fragment was deleted to produce the plasmid pMCS101 (2136). This vector was successfully used in the construction of pTEQ4 (see **Chapter V**, page 132). The construction of pMCS101 was done in collaboration with J. McGrory.

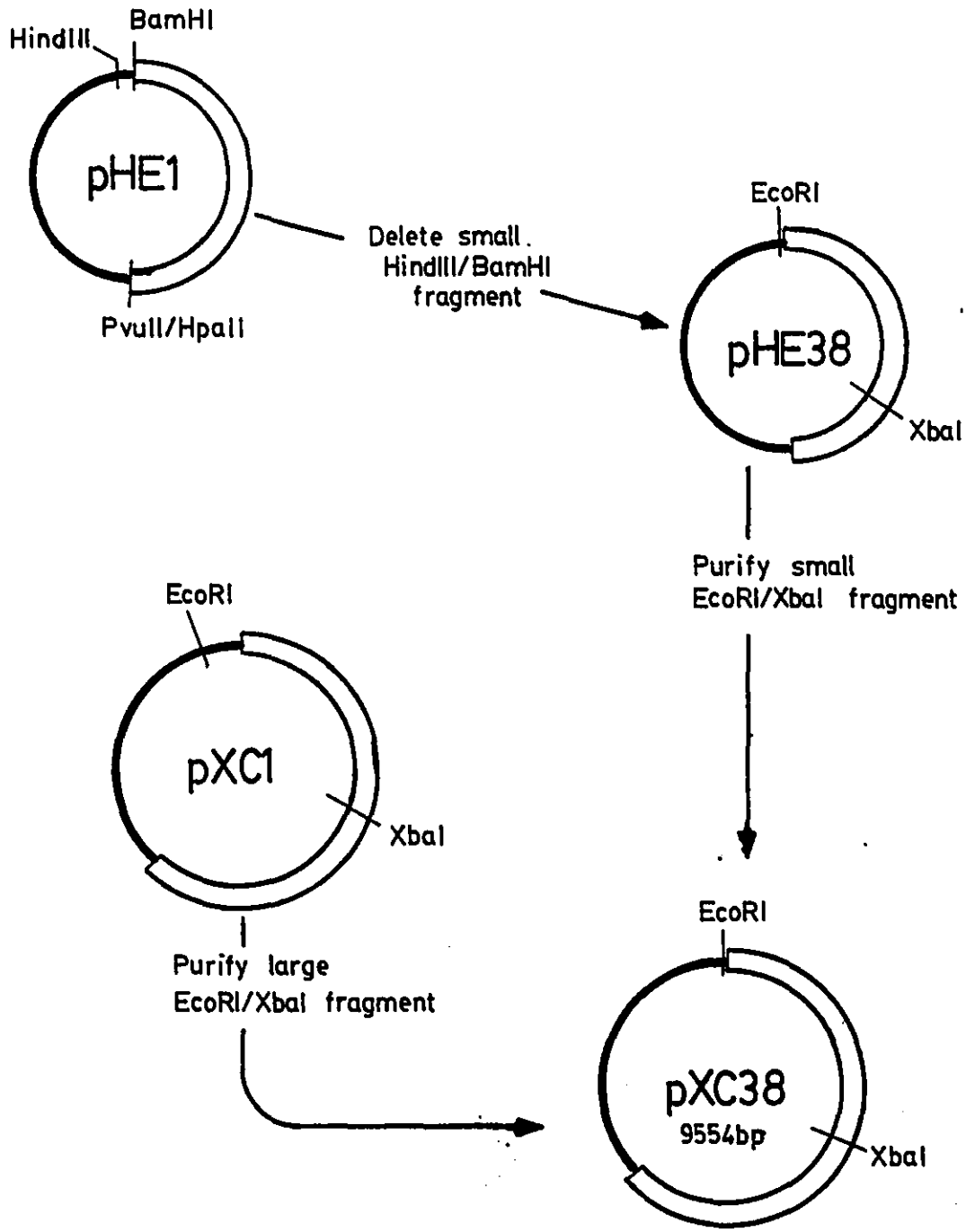
4. CONSTRUCTION OF pMCS101

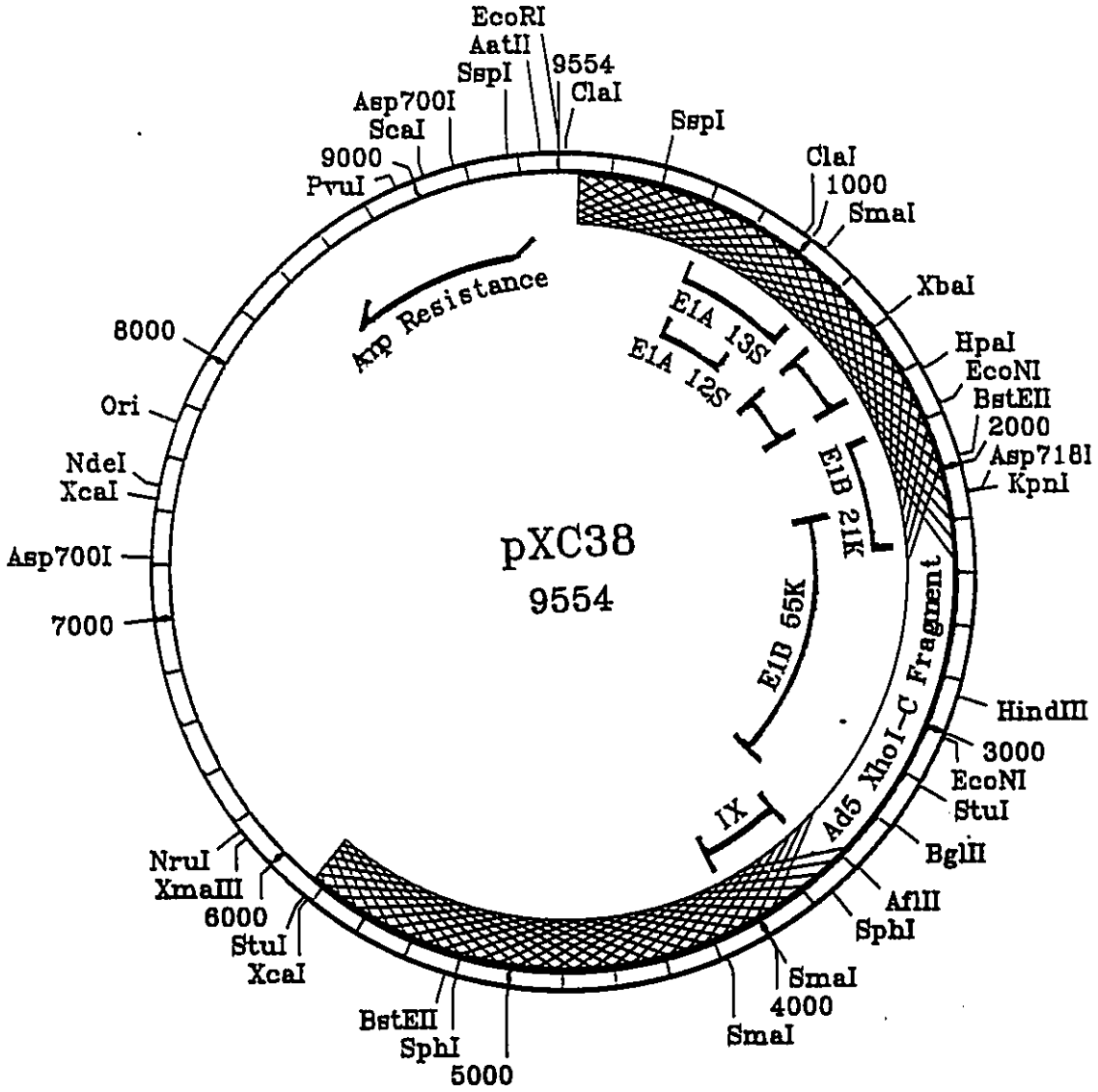


5. Construction of pHE38 and pXC38

The plasmids pHE38 and pXC38 were constructed as derivatives of pHE1 and pXC1, respectively, in which both the HindIII and BamHI sites in the pBR sequence were deleted. This was intended for the rescue of the E1A insertion mutants described in Chapter IV, all of which contained the recognition sequence for BamHI in the inserts, to simplify structural analysis. Thus, a small HindIII/BamHI fragment in the pBR322 sequence of pHE1 was deleted to produce pHE38. The small EcoRI/XbaI fragment containing the deletion was used to replace a corresponding fragment in pXC1 to produce pXC38 (9554bp). The reconstitution of the E1A insertion mutants into intact E1 using pXC38 was described previously (McGrory, 1988).

5. Construction of pHE38 and pXC38

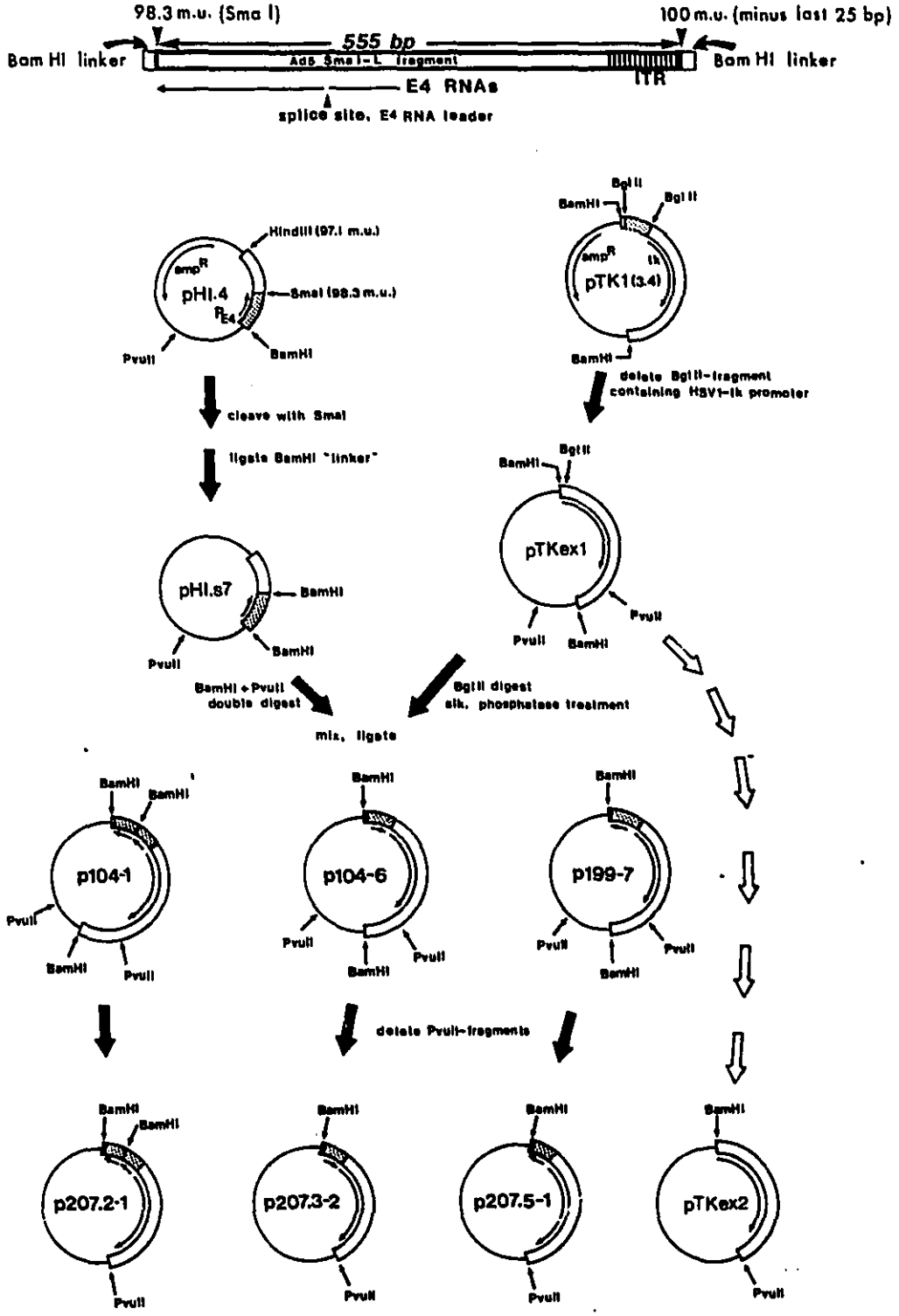



pXC38
9554
EcoRI
AatII
SspI
9554
ClaI
SspI
ClaI
1000
SmaI
XbaI
HpaI
EcoNI
BstEII
2000
Asp718I
KpnI
HindIII
3000
EcoNI
StuI
BglII
AflII
SphI
SmaI
4000
SmaI
BstEII
SphI
5000
NruI
XmaIII
6000
StuI
XcaI
Asp700I
7000
NdeI
XcaI
Ori
8000
Asp700I
ScaI
9000
PvuI
Amp Resistance
E1A 13S
E1A 12S
E1B 21K
E1B 55K
IX
Ad5 XhoI-C Fragment

APPENDIX D. CHARACTERIZATION OF A RIGHTWARD READING PROMOTER AT THE END OF THE AD5 GENOME

This study was initiated to characterize the early region 4 (E4) promoter found at the extreme right-end of the viral genome (at 99.1 m.u.) and normally transcribed in a leftward direction (reviewed by Berk, 1986b). Such a study involved the expression of the herpes simplex type 1 (HSV-1) thymidine kinase (tk) gene under the control of the E4 promoter found in the small *Sma*I-L fragment (98.3 to 100 m.u.). The ability of the E4 promoter to drive the expression of the tk gene was then assessed in a stable cell focus assay involving the biochemical transformation of tk-deficient mouse LTA cells to the tk^+ phenotype. **Figure 1** shows the strategy used in the construction of Ad5-tk fusion plasmids. The top diagram shows the structure of the *Sma*I-L fragment containing the E4 promoter and part of the region's coding sequences. The fragment was obtained from the plasmid pHI4 (F.L. Graham, pers. comm.) which contained the *Hin*dIII-I fragment (97.1-100 m.u.) of Ad5 and which was modified by inserting a *Bam*HI linker at the *Sma*I site (98.3 m.u.) to produce a plasmid called pHIs7. Sequencing of this fragment has shown that the terminal 25bp have been deleted and replaced by linker sequences used in its cloning (data not shown). The E4-promoter-containing fragment could now be cleaved out as a small *Bam*HI fragment. The promoterless tk gene was obtained from pTKex1 (J. Smiley, pers. comm.) which was derived

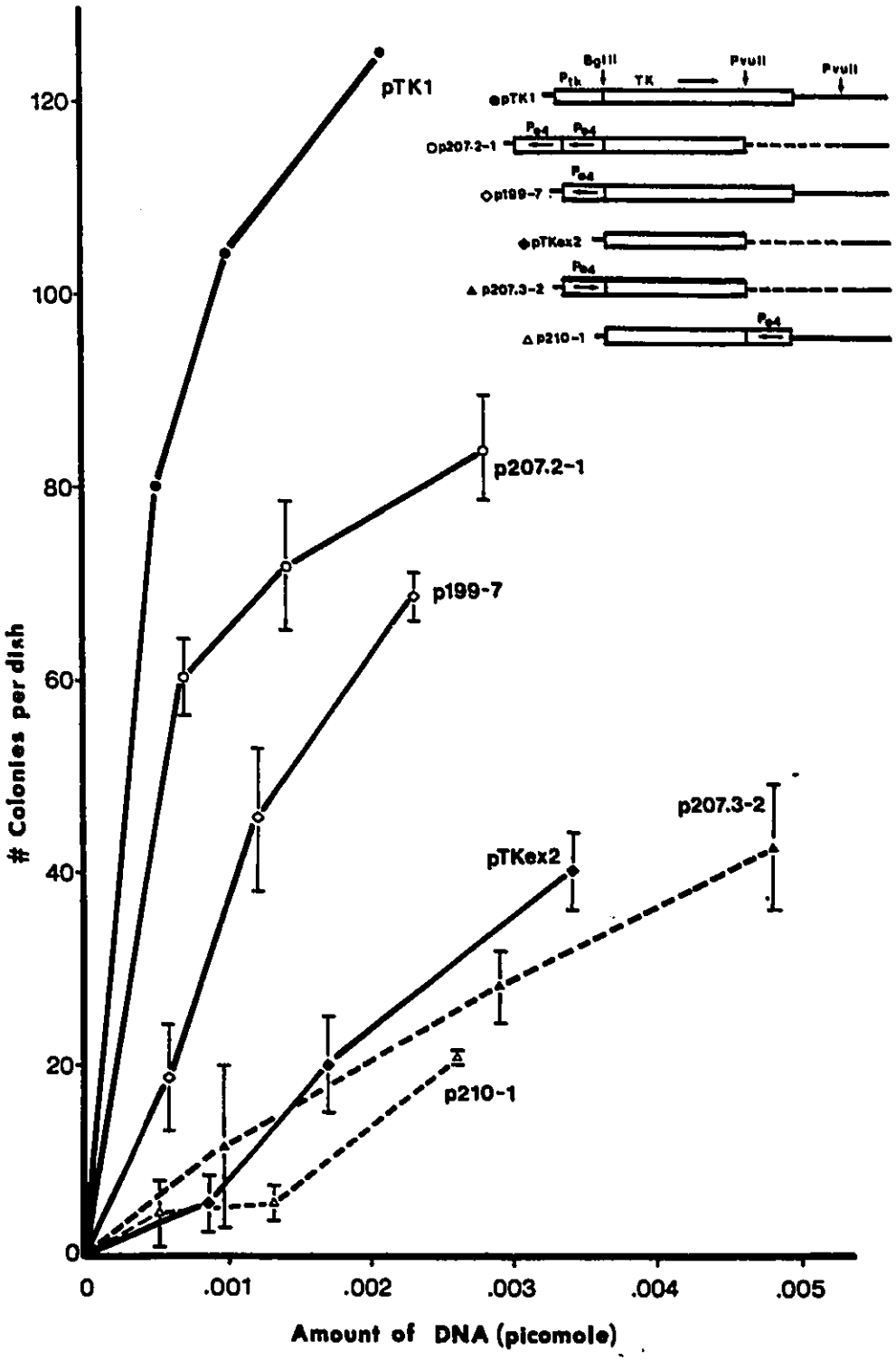
Figure 1. **Construction of Ad5-TK Fusion Genes**



from pTK1(3.4kb) by deleting a small *Bgl*II fragment containing the major transcriptional regulatory sequences. The *Bam*HI fragment containing the E4 promoter could ligate in the single *Bgl*II site of pTKex1 and assume the two possible orientations. Three independent isolates were initially isolated and by sequencing using the Maxam and Gilbert (1980) protocol (data not shown), the structures of these clones were established: p104-1 contained 2 inserts arranged in tandem in which the orientation of the E4 promoter was reversed with respect to that of the tk gene; p104-6 contained a single insert in which the E4 promoter had the same orientation as the tk gene; and p199-7 contained a single insert in the reversed orientation. Derivatives of these clones were made by deleting a small *Pvu*II fragment that mapped downstream from the tk gene and into the pBR sequences.

The activity of these constructs to biochemically transform tk-deficient LTA cells to the tk^+ phenotype was determined. **Figure 2** shows the results of one such assay using constructs whose structures are shown above the graph. Using the plasmids pTK1 (positive) and pTKex2 (negative) as controls, it was found that the cloned *Sma*I-L fragment was able to significantly increase the transforming activity of the tk gene when present at the 5'-end but only in the reversed orientation (i.e., the direction of the E4 promoter was reversed with respect to tk) as shown for p207.2-1 and p199-7. When the orientation of the E4 promoter was correct (p207.3-2), the transforming activity did not differ from the negative control. The failure of the E4 promoter to increase the transforming activity of tk (which was assumed to be a function of the levels of tk expression) was attributed to the fact that its optimal expression requires E1A functions, in which case the basal level of E4 promoter activity was undetectable by this type of assay. The results, however,

Figure 2. Transformation Assays With Ad5-TK Fusion Plasmids

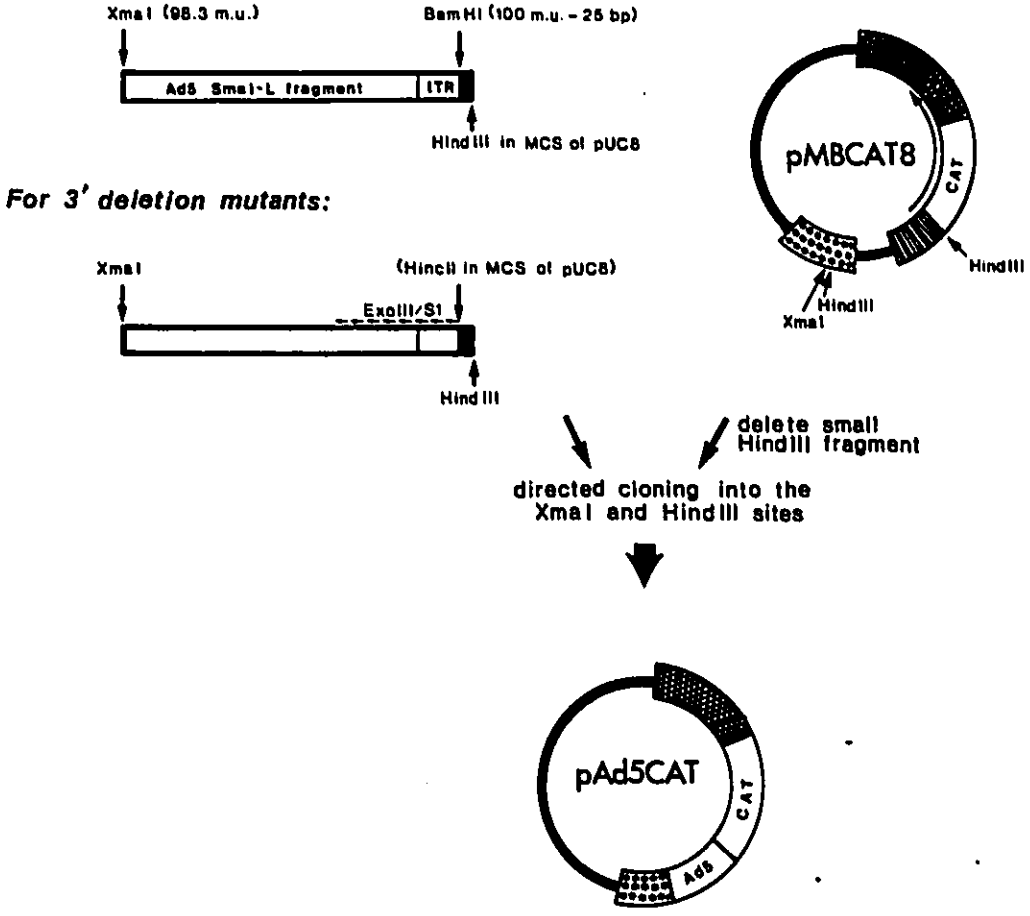


indicated that the fragment contained another transcriptional element with an activity directed away from the E4 region. This element could be a promoter, an enhancer, or a combination of the two. To test whether the fragment contained an enhancer, the plasmid p210-1 was constructed and its activity was determined in a similar fashion. This plasmid contained the *Sma*I-L fragment at the 3' end of tk and was intended to test the position-independence criterion for an enhancer. The transforming activity of this construct appears in **Figure 2** which shows that it did not differ from that of the negative control. Therefore, the putative transcriptional element was most likely a simple promoter as further evidenced by a strict orientation-dependency in its ability to drive the expression of tk.

To further characterize the putative promoter at the end of the viral genome, a transient expression assay based on the bacterial acetyl transferase gene (CAT) was used, first to demonstrate that it could drive the expression of another reporter gene, and second to determine the 3'-boundary of the element by making a series of deletion mutants from the 3' terminus. **Figure 3** shows the strategy for making the constructs. First, a plasmid called pMBCAT8 was constructed. This plasmid was derived from pSV2CAT (Gorman *et al.*, 1982) by inserting a *Hae*II fragment from pUC8 containing the *lac* z gene for α-complementation. pMBCAT8 therefore contained the colony color-indication as well as the multiple cloning site features of the pUC plasmids.

The *Sma*I-L fragment was taken from pHI4 as a *Xma*I-*Bam*HI fragment and cloned into pUC8 into corresponding sites. The fragment could now be cleaved as an *Xma*I-*Hin*dIII fragment due to a *Hin*dIII site in the multiple cloning site of pUC8 adjacent to the *Bam*HI site. This fragment was used to replace a corresponding fragment in pMBCAT8 to position the putative promoter in front of

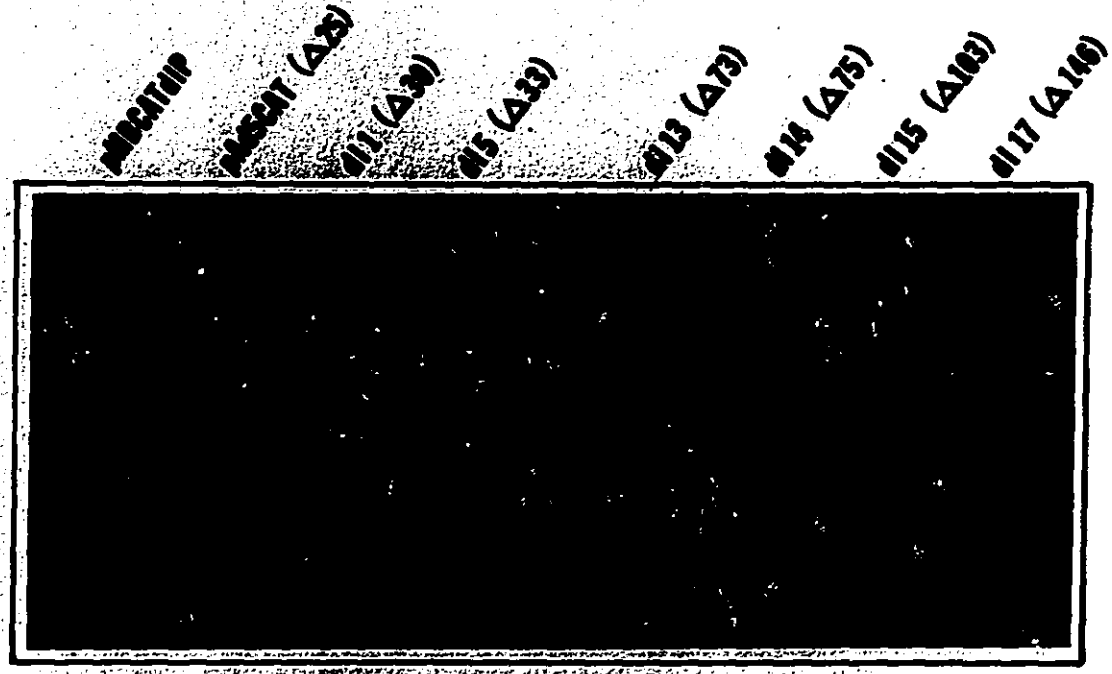
Figure 3. Construction of Ad5-CAT Fusion Genes



the CAT gene in the proper orientation. The resulting plasmid was called pAd5CAT (**Figure 3**). The series of 3' deletion mutants was constructed starting from the *Bam*HI site of pHI4 by stepwise digestion with ExoIII and S1 nucleases. By cutting with *Xma*I at the 5'-end, fragments with deletions were cloned into pUC8 at the *Hin*cII (blunt) and *Xma*I sites. Candidate clones were further characterized by sequencing using the M13/dideoxy technique (Messing, 1983). The deletion mutants in pUC were then cloned into pMBCAT8 using the *Xma*I and *Hin*dIII sites as described for pAd5CAT. Six mutants whose deletions ranged from 33 to 146bp were isolated for further characterization in transient expression assays.

The transient expression assays were carried out using 293 cells, a human cell line that constitutively express adenovirus E1 proteins (Graham *et al.*, 1977). The amount of transfected plasmid DNA using thecalcium technique (Graham *et al.*, 1980) was 12 µg per 100mm-petri dish which was incubated with the cells for 24 hours. For the enzymatic reaction, one-half of the total cell extract was used. **Figure 4** shows the results of the assay using the 6 deletion mutants in addition to a negative control pMBCAT*dl*P, which was derived from pMBCAT8 by deleting a *Hin*dIII fragment containing the SV40 enhancer-promoter, and the wild-type fusion pAd5CAT. The expression of pAd5CAT was low and deletions of up to nt73 (counting from the right terminus) caused further decrease. However, deletions between nt75 to nt103 (the latter being the complete deletion of the ITR) resulted in increased expression and seemed to decrease when the last 146bp were deleted. Since the tk transformation assays have indicated that the level of expression from this promoter in mouse cells was high relative to the negative control and the CAT assays have indicated low expression until the ITR was removed, then the promoter

QUANTITATIVE CAT ASSAY: DETERMINATION OF THE 3′ BOUNDARY OF THE TRANSCRIPTIONAL REGULATORY ELEMENT



must be under a negative regulatory control when expressed in 293 cells and the ITR must be a target of such control.

It is difficult to determine from the evidence presented above as to the nature of the regulatory control proposed for the rightward promoter since the assay was carried out in 293 cells. Such cells express proteins from E1A and E1B and since both transcriptional units make products which can activate or repress transcription, it is not possible to determine which unit actually represses the putative promoter's activity in 293 cells. However, one can invoke a mechanism whereby the E1 proteins are not involved. That is, since the ITR contains binding sites for CTF/NF1 (see **Chapter I**, p.38), it is possible that binding of the factor is a physical interference of transcription from the rightward promoter; its removal. therefore, increases the level of transcription, which is consistent with these results. Alternatively, the E1 proteins might promote the binding of CTF/NF1 resulting in reduction of expression.

The existence of a rightward reading promoter is interesting since there is no apparent protein-coding sequence in its direction. One attractive hypothesis to explain its presence is that it is involved in the modulation of the E4 promoter. There are at least two other cases in adenovirus where one can find back-to-back arrangements of two promoters, one being the E2A(e) (75.0 m.u.) and E3 (76.6 m.u.) promoters (see **Chapter 1**, page 37) as well as the adenovirus major late 16.4 m.u.) and IVa2 (15.9 m.u.) promoters. How these arrangements contribute to the regulation of these promoters is still not clear although one assumes that some kind of decision must be made by the transcriptional machinery which of the two is preferred at specific times during the viral cycle. In the case of the rightward promoter and that of E4, it is possible that the E4 promoter is shut off during

transcription of the rightward promoter and in the presence of *trans*-activating E1A and E1B proteins, the E4 promoter is turned on by shutting off the rightward promoter. This model, which is at this stage is speculative, needs further investigation following a full characterization of the rightward promoter as to its precise initiation site and modes of control.

BIBLIOGRAPHY

Abmayr, S.M., J.L. Workman, and R.G. Roeder (1988). The pseudorabies immediate early protein stimulates in vitro transcription by facilitating TFIID:promoter interactions. *Genes Dev.* 2:542-553.

Allan, M., J.-de Zhu, P. Montague, and J. Paul (1984). Differential response of multiple ε-globin cap sites to *cis*- and *trans*-acting controls. *Cell* 38:399-407.

Andersson, M., A. McMichael, and P.A. Peterson (1987). Reduced allorecognition of adenovirus-2 infected cells. *J. Immunol.* 138:3960-6.

Andersson, M., S. Paabo, T. Nilsson, and P.A. Peterson (1985). Impaired intracellular transport of class I MHC antigens as a possible means for adenoviruses to evade immune surveillance. *Cell* 43:215-222.

Angel, P.E.A., Allegretto, S. Okino, K. Hattori, W.J. Boyle, T. Hunter, and M. Karin (1988a). Oncogene *jun* encodes a sequence specific *trans*-activator similar to AP-1. *Nature* 332:166-71.

Angel, P., K. Hattori, T. Smeal, and M. Karin (1988b). The *jun* proto-oncogene is positively autoregulated by its product, Jun/Ap-1. *Cell* 55:875-885.

Angel, P., M. Imagawa, R. Chiu, B. Stein, R.J. Imbra, H.J. Rahmsdorf, Cc. Jonat, P. Herrlich, and M. Karin (1987). Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell* 49:729-39.

Babiss, L.E. and Ginsberg, H.S. (1984a). Adenovirus type 5 early region 1b gene product is required for efficient shutoff of host protein synthesis. *J. Virol.* 50:202-212.

Babiss, L.E., Ginsberg, H.S. and Darnell, J.E., Jr. (1985). Adenovirus E1B proteins are required for accumulation of late viral mRNA and for effects on cellular mRNA translation and transport. *Mol. Cell. Biol.* 5:2552-2558.

Baker, C.C. and Ziff, E.B. (1981). Promoters and heterogeneous 5' termini of the messenger RNAs of adenovirus serotype 2. *J. Mol. Biol.* 149:189-221.

Ball, A.B., M.E. Williams, and K.R. Spindler (1988). Identification of mouse adenovirus type 1 early region 1: DNA sequence and a conserved transactivating function. *J. Virol.* 62:3947-57.

Barany, F. (1985). Two-codon insertion mutagenesis of plasmid genes by using single-stranded hexameric oligonucleotides. *Proc. Natl. Acad. Sci. USA* **82**:4202-4206.

Barrett, P., L. Clark, and R.T. Hay (1987). A cellular protein binds to a conserved sequence in the adenovirus type 2 enhancer. *Nucl. Acids Res.* **15**:2719-35.

Bellgrau, D., W.A. Walker and J.L. Cook (1988). Recognition of adenovirus E1a gene products on immortalized cell surfaces by cytotoxic T lymphocytes. *J. Virol.* **62**:1513-1519.

Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon (1980). The ovalbumin gene; sequence of putative control regions. *Nucl. Acids Res.* **8**:127-142.

Berger, S.L. and W.R. Folk (1985). Differential activation of RNA polymerase III-transcribed genes by the polyomavirus enhancer and the adenovirus E1A gene products. *Nucl. Acids Res.* **13**:1413-28.

Bergman, Y. and D. Shavit (1988). Regulation of the Ig kappa-chain enhancer by the adenovirus E1A gene products: repression in lymphoid cells, activation in fibroblasts. *J. Immunol.* **140**:2073-80.

Berk, A.J. (1986a). Functions of adenovirus E1A. *Cancer Surveys* **5**:367-387.

Berk, A.J. (1986b). Adenovirus promoters and E1A transactivation. *Ann. Rev. Genet.* **20**:45-79.

Berk, A.J. and Sharp, P.A. (1978). Structure of the adenovirus 2 early mRNAs. *Cell* **14**:695-711.

Bernards, R., P.I. Schrier, A. Houweling, J.L. Bos, A.J. van der Eb, M. Zielstra, and C.J.M. Melief (1983). Tumorigenicity of cells transformed by adenovirus type 12 by evasion of T-cell immunity. *Nature* **305**:776-79.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* **7**:1513-1523.

Blanton, R.A and T. Carter (1979). Autoregulation of adenovirus type 5 early gene expression. III. Transcription studies in isolated nuclei. *J. Virol.* **29**:458-465.

Boeuf, H., D.A. Zajchowski, T. Tamura, C. Hauss and C. Kodinger (1987). Specific cellular proteins bind to critical sequences of the adenovirus early E2A promoter. *Nucl. Acids Res.* **15**:509-27.

Bohmann, D., T.J. Bos, A. Admon, T. Nishimura, P.K. Vogt, and R. Tjian (1987). Human proto-oncogene c-jun encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* **238**:1386-92.

Borrelli, E., Hen, R. and Chambon, P. (1984). Adenovirus-2 E1a products repress enhancer-induced stimulation of transcription. *Nature* **312**:608-612.

Bos, J.L., H.C. ten Wolde Kraamwinkel (1983). The E1B promoter of Ad12 in mouse L-cells is activated by adenovirus region E1A. *EMBO J.* **2**:73-76.

Boyer, H.W. and D. Roulland-Dussoix (1969). A complementation analysis of the restriction and modification of DNA in Escherichia coli. *J. Mol. Biol.* **41**:459-72.

Branton, P.E., S.T. Bayley, and F.L. Graham (1985). Transformation by human adenoviruses. *Biochim. Biophys. Acta* **780**:67-94.

Breathnatch, R. and P. Chambon (1981). Organization and expression of eukaryotic split genes coding for proteins. *Ann. Rev. Bochem.* **50**:349-83.

Briggs, M.R., J.T. Kadonaga, S.P. Bell, and R. Tjian (1986). Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* **234**:47-52.

Brown, RS., C. Sander and P. Argos (1985). The primary structure of transcription factor TFIIIA has 12 consecutive repeats. *FEBS Lett.* **186**:271-4.

Bruner, M., B. Dalie, R., Spangler, M.L. Harter (1988). Purification and biological characterization of an adenovirus type 2 E1A protein expressed in E. coli. J. Biol. Chem. 263:3984-9.

Burgert, H.G. and S. Kvist (1987). The E3/19k protein of adenovirus type 2 binds to the domains of histocompatibility antigens required for CTL recognition. *EMBO J.* **6**:2019-2026.

Burgert, H.G., J.L. Maryanski, and S. Kvist (1987). "E3/19K" protein of adenovirus type 2 inhibits lysis of cytolytic T lymphocytes by blocking cell-surface expression of histocompatibility class I antigens. *Proc. Natl. Acad. Sci. USA* **84**:1356-60.

Carlock, L.R. and Jones, N.C. (1981). Transformation-defective mutant of adenovirus type 5 containing a single altered E1A mRNA species. *J. Virol.* **40**:657-664.

Casadaban, M.J., A. Martinez-Arias, S.K. Shapira, and J. Chou (1983). Beta-galactosidase gene fusions for analyzing gene expression in Escherichia coli and yeast. *Methods Enzymol.* **100**:293-309.

Chaterjee, P.K., M. Bruner, S.J. Flint, and M.L. Harter (1988). DNA-binding properties of an adenovirus 289R E1A protein. *EMBO J.* **7**:835-841.

Chen, J.-M. and W.-T. Chen (1987). Fibronectin-degrading proteases from the membranes of transformed cells. *Cell* **48**:193-203.

Chen, M.J., B. Holskin, J. Strickler, J. Gorniak, M.A. Clark, P.J. Johnson, M. Mitcho, and D. Shalloway (1987). Induction by E1A oncogene expression of cellular susceptibility to lysis by TNK. *Nature* **330**:581-3.

Chinnadurai, G. (1983). Adenovirus 2 lp+ locus codes for a 19kd tumor antigen that plays an essential role in cell transformation. *Cell* **33**:759-766.

Chiu, R., M. Imagawa, R.J. Imbra, J.R. Bockoven, and M. Karin (1987). Multiple cis- and trans-acting elements mediate the transcriptional response to phorbol esters. *Nature* **329**:648-51.

Chiu, R., W.J. Boyle, J. Meek, T. Smeal, T. Hunter, and M. Karin (1988). The c-fos protein interacts with c-jun/AP-1 to stimulate transcription of AP-1 responsive genes. *Cell* **54**:541-52.

Chodosh, L.A., A.S. Baldwin, R.W. Carthew, and P.A. Sharp (1988a). Human CCAAT-binding proteins have heterologous subunits. *Cell* **53**:11-24.

Chodosh, L.A., J. Olesen, S. Hahn, A.S. Baldwin, L. Guarente, and P.A. Sharp (1988b). A yeast and a human CCAATT-binding protein have heterologous subunits that are functionally interchangeable. *Cell* **53**:25-35.

Chow, K.-C., and G.D. Pearson (1985). Adenovirus infection elevates levels of cellular topoisomerase I. *Proc. Natl. Acad. Sci. USA* **82**:2247-2251.

Chow, L.T., T.R. Broker, and J.B. Lewis (1979a). Complex splicing patterns of RNAs from the early regions of adenovirus-2. *J. Mol. Biol.* **134**:265-303.

Cladaras, C. and W.S.M. Wold (1985a). DNA sequence of the early E3 transcription unit of adenovirus 5. *Virology* **140**:28-43.

Cladaras, C., B. Bhat, and W.S.M. Wold (1985b). Mapping the 5' ends, and splice sites of mRNAs from the early E3 transcription unit of adenovirus 5. *Virology* **140**:44-54.

Coffino, P., H.R. Bourne, U. Friedrich, J. Hochman, P.A. Insel,I. Lemaire, K.L. Melman, and G.M. Tomkins (1976). Molecular mechanism of cyclic AMP action: a genetic approach. *Recent Prog. Horm. Res.* **32**:669-84.

Comb, M., N.C. Birnberg, A. Seasholtz, E. Herbert, and H.W. Goodman (1986). A cyclic AMP- and phorbol ester-inducible DNA element. *Nature* **323**:353-6.

Cook, J.L., and A.M. Lewis, Jr. (1984). Differential NK cells and macrophage

killing of hamster cells infected with nononcogenic or oncogenic adenovirus. *Science* 224:612:5.

Cook, J.L., D.L. May, A.M. Lewis, Jr., and T.A. Walker (1987). Adenovirus E1A gene induction of susceptibility to lysis by natural killer cells and activated macrophages in infected rodent cells. *J. Virol.* 61:3510-3520.

Cook, J.L., T.A. Walker, A.M. Lewis, Jr., H.E. Ruley, F.L. Graham, and S.H. Pilder (1986). Expression of the adenovirus E1A oncogene during cell transformation is sufficient to induce susceptibility to lysis by host inflammatory cells. *Proc. Natl. Acad. Sci. USA* 83:6965-9.

Courey, A.J. and R. Tjian (1988). Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* 55:887-98.

Culp, J.S., L.C. Webster, D.J. Friedman, C.L. Smith, W.-J. Huan, F.Y.-H. Wu, M. Rosenberg, and R.P. Ricciardi (1988). The 289-amino acid E1A protein of adenovirus binds zinc in a region that is important for trans-activation. *Proc. Nat. Acad. Sci. USA* 85:6450-54.

Curran, T. and B.R. Franza, Jr. (1988). Fos and Jun: the AP1 connection. *Cell* 55:395-97.

Danner, D.B. (1986). The lac operator as a phenotypic label for DNA fragments cloned in Escherichia coli. *Gene* 44:193-199.

DePamphilis, M.L. (1988). Transcriptional elements as components of eukaryotic origins of DNA replication. *Cell* 52:635-638.

Deutsch, P.J., J.L. Jameson, and J.F. Habener (1987). cAMP responsiveness of human gonadotropin-alpha gene transcription is directed by a related 18 bp enhancer. *J. Biol. Chem.* 262:12169-12174.

D'Halluin, J.C., C. Allart, C. Cousin, P.A Boulanger, and G. Martin (1979). Adenovirus early function required for the protection of viral and cellular DNA. *J. Virol.* 32:61-71.

Doherty, P.C., B.B. Knowles, and P.J. Wettstein (1984). Immunological surveillance of tumours in the context of major histocompatibility restriction of T cell function. *Adv. Cancer Res.* 42:1-65.

Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis (1987). A multiplicity of CCAAT box-binding proteins. *Cell* 50:863-872.

Dynan, W.S. and R. Tjian (1983a). Isolation of transcription factors that discriminate between different promoters recognized by RNA pol II. *Cell*

32:669-80.

Dynan, W.S. and R. Tjian (1983b). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* 35:79

Eddy, B.E., G.S. Borman, G.E. Grubbs, and R.D. Young (1962). Identification of the oncogenic substance in rhesus monkey kidney cell cultures as simian virus 40. *Virology* 17:65.

Efstratiadis, A., J.W. Posakony, T. Maniatis, R.M. Lawn, C. O'Connell, R.A. Spritz, J.K. DeRiel, B.G.Forget, S.M. Weissman, J.L. Slighton, A.E. Blechl, O. Smithies, F.E. Baralle, C.C. Shoulders, and N.J. Proudfoot (1980). The structure and evolution of the human beta-globin gene family. *Cell* 21:653-68.

Eliyahu, D., A. Raz, P. Gruss, D. Givol, and M. Oren (1984). Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. *Nature* 312:646-49.

Engel, D.A., S. Hardy, and T. Shenk (1988). cAMP acts in synergy with E1A protein to activate transcription of the adenovirus early genes E4 and E1A. Genes Dev. 2:1517-28.

Enver, T. A.C. Brewer, and R.K. Patient (1988). Role for DNA replication in Beta-globin gene activation. *Mol. Cell. Biol.* 8:1301-1308.

Enzo, H., Lai Fatt, R.B. and Mak, S. (1981). Degradation of intracellular DNA in KB cells infected with cyt mutants of human adenovirus type 12. *J. Virol.* 40:20-27.

Epstein, M.A., B.G. Achong, and Y.M. Barr (1964). Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet* ii:702.

Esche, H., Mathews, M.B. and Lewis, J.B. (1980). Proteins and messenger RNAs of the transforming region of wild-type and mutant adenoviruses. *J. Mol. Biol.* 142:399-417.

Evans, R.M. and S.M. Hollenberg (1988). Zinc fingers: gilt by association. *Cell* 52:1-3. (see also published erratum in Cell 52:783.)

Everett, R.D. and M. Dunlop (1984). Trans activation of plasmid-borne promoters by adenovirus and several herpes group viruses. *Nucl. Acids Res.* 12:5969-78.

Falgout, B., and G. Ketner (1987). Adenovirus early region 4 is required for efficient virus particle assembly. *J. Virol.* 61:3759-68.

Ferguson, B., Krippl, B., Andrisani, O., Jones, N., Westphal, H. and Rosenberg, M. (1985). E1A 13S and 12S mRNA products made in Escherichia coli both function as nucleus-localized transcription activators but do not directly bind DNA. *Mol. Cell. Biol.* **5**:2653-2661.

Figge, J., T. Webster, T.F. Smith, and E. Paucha (1988). Prediction of similar transforming regions in simian virus 40 large T, adenovirus E1A, and myc oncoproteins. *J. Virol.* **62**:1814-1818.

Flint, S.J. (1981). Structure and genomic organization of adenoviruses. In *DNA TUMOR VIRUSES: Molecular Biology of Tumor Viruses* (Tooze, J, ed.), pp. 383-441, Cold Spring Harbor Laboratory, NY.

Flomemberg, P.R., M. Chen, and M.S. Horwitz (1987). Characterization of a major histocompatibility complex class I antigen-binding glycoprotein from adenovirus type 35, a type associated with immunocompromised hosts. *J. Virol.* **61**:3665-71.

Gallimore, P.H., Sharp, P.A. and Sambrook, J. (1974). Viral DNA in transformed cells. II. A study of the sequences of adenovirus 2 DNA in nine lines of transformed rat cells using specific fragments of the viral genome. *J. Mol. Biol.* **89**:49-72.

Garcia, J., F. Wu, and R. Gaynor (1987). Upstream regulatory regions required to stabilize binding to the TATA sequence in an adenovirus early promoter. Nucleic Acids Res. 15:8367-85.

Gaynor, R.B. and A.J. Berk (1983). Cis-acting induction of adenovirus transcription. *Cell* **33**:683-93.

Gaynor, R.B., D. Hillman, and A.J. Berk (1984). Adenovirus early region 1A protein activates transcription of a nonviral gene introduced into mammalian cells by infection or transfection. *Proc. Natl. Acad. Sci. USA* **81**:1193-7.

Gaynor R.B., L.T. Feldman, and A.J. Berk (1985). Transcription of class III genes activated by viral immediate early proteins. *Science* **2230**:447-450.

Gilardi, P. and M. Perricaudet (1986). The E4 promoter of adenovirus type 2 contains an E1A-dependent cis-acting element. *Nucl. Acids Res.* **14**:9035-49.

Gilbert, W. and A. Maxam (1973). The nucleotide sequence of the lac operator. *Proc. Natl. Acad. Sci. USA* **70**:3581-3584.

Giles, K.W. and Myers, A. (1965). An improved diphenylamine method for the estimation of deoxyriboucleic acid. *Nature* **206**:93.

Gilman, A.G. (1984). G-proteins and dual control of adenylate cyclase. *Cell*

36:577-79.

Girardi, A.J., B.H. Sweet, V.B. Slotnick, and M.R. Hilleman (1962). Development of tumors in hamsters inoculated in the neo-natal period with vacuolating virus, SV40. *Proc. Soc. Exp. Biol. Med.* **109**:649.

Girvitz, S.C., S. Bacchetti, A.J. Rainbow, and F.L. Graham (1980). A rapid and efficient procedure for the purification of DNA from agarose gels. *Anal. Biochem.* **106**:492.

Goding, C., P. Jalinot, D. Zajchowski, H. Boef, and C. Kedinger (1985). Sequence-specific trans-activation of the adenovirus EIIa early promoter by the viral EIV transcription unit. *EMBO J.* **4**:1523-8.

Gooding, L.R., L.W. Elmore, A.E. Tolefson, H.A. Brady, and W.S.M. Wold (1988). A 14,700 mw protein from the E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. *Cell* **53**:341-346.

Goodman, H.M and R.J. McDonald (1979). Cloning of hormone genes from a mixture of cDNA molecules. *Methods Enzymol.* **68**:75-90.

Gorman, C.M., B.H. Howard, R. Reeves (1983). Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. *Nucl. Acids Res.* **11**:7631-7648.

Gorman, C.M., L.F. Moffat, and B.H. Howard (1982). Recombinant genomes which express chloramphenicol acetyl transferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044.

Graham, F.L. (1984). Transformation by adenovirus. In *The Adenoviruses* (Ginsberg, H.S., ed.), pp. 339-398, Plenum Press, New York.

Graham, F.L. and van der Eb, A.J. (1973a). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**:456-467.

Graham, F.L., S. Bacchetti, R. McKinnon, C. Stanners, B. Cordell, H. Goodman (1980). Transformation of mammalian cells with DNA using the calcium technique. In *Introduction of Macromolecules into Viable Mammalian Cells*, Alan R. Liss, Inc., New York.

Graham, F.L. and van der Eb, A.J. (1973b). Transformation of rat cells by DNA of human adenovirus 5. *Virology* **54**:536-9.

Graham, F.L., Smiley, J., Russell, W.C. and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**:59-72.

Graham, F.L., van der Eb, A.J. and Heijneker, H.L. (1974). Size and location of the transforming region in human adenovirus type 5 DNA. *Nature* 251:687-691.

Graves, B.J., P.F. Johnson, and S.L. McKnight (1986). Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* 44:565-576.

Green, M., W.S.M. Wold, J.K. Mackey, and P. Rigden (1979). Analysis of human tonsil and cancer DNAs and RNAs for DNA sequences of group C (serotypes 1, 2, 5, and 6) human adenoviruses. *Proc. Natl. Acad. Sci. USA* 76:6606.

Green, M., P.M. Loewenstein, R. Pusztai, and J.S. Symington (1988). An adenovirus E1A protein domain activates transcription in vivo and in vitro in the absence of protein synthesis. *Cell* 53:921-26.

Green, M.R., Treisman, R. and Maniatis, T. (1983). Transcriptional activation of cloned human beta-globin genes by viral immediate-early gene products. *Cell* 35:137-148.

Green, M., W.S.M. Wold, J.K. Mackey, and P. Rigden (1979). *Proc. Natl. Acad. Sci. USA* 76:6606-6610.

Grodzicker, T. and N. Hopkins (1981). Origins of contemporary DNA tumor virus research. In *J. DNA TUMOR VIRUSES: Molecular Biology of Tumor Viruses* (Tooze, J, ed.), pp. 1-59, Cold Spring Harbor Laboratory, NY.

Guilfoyle, R.A., Osheroff, W.P. and Rossini, M. (1985). Two functions encoded by adenovirus early region 1A are responsible for the activation and repression of the DNA-binding protein gene. *EMBO J.* 4:707-713.

Halazonetis, T.D., K. Georgopoulos, M.E., Greenberg, and P. Leder (1988). c-Jun dimerizes with itself and with c-Fos, forming complexes of different DNA binding affinities. *Cell* 55:917-24.

Halbert, D.N., J.R. Cutt, and T. Shenk (1985). Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *J. Virol.* 56:250-7.

Hanahan, D. (1983). Studies on Transformation of Escherichia coli with plasmids. *J. Mol. Biol.* 166:267-274.

Hanahan, D. and Y. Gluzman (1984). Rescue of functional replication origins from embedded configurations in a plasmid carrying the adenovirus genome. *Mol. Cell. Biol.* 4:302-309.

Hanaka, S., T. Nishigaki, P.A. Sharp, and H. Handa (1987). Regulation of in vitro and in vivo transcription of early-region IV of adenovirus type 5 by multiple cis-acting elements. *Mol. Cell. Biol.* 7:2578-87.

Handa, H., R.E. Kingston, and P.A. Sharp (1983). Inhibition of adenovirus early region IV transcription unit in vitro by purified viral DNA binding protein. *Nature* 302:545-47.

Hardy, S. and T. Shenk (1988). Adenoviral control regions activated by E1A and the cAMP response element bind the same factor. *Proc. Natl. Acad. Sci. USA* 85:4171-75.

Hasimoto, S., W. Schmid, and G. Schutz (1984). Transcriptional activation of the rat liver tyrosine aminotransferase gene by cAMP. *Proc. Natl. Acad. Sci. USA* 81:6637-41.

Hatamochi, A., B. Paterson, and B. de Crombrugghe (1986). Differential binding of a CCAAT DNA binding factor to the promoters of the mouse alpha2(I). and alpha1(III). collagen genes. *J. Biol. Chem.* 261:11310-4.

Hearing, P. and T. Shenk (1983a). functional analysis of the nucleotide sequence surrounding the cap site for adenovirus type 5 region E1A messenger RNAs. *J. Mol. Biol.* 167:809-22.

Hearing, P. and T. Shenk (1985). Sequence-independent auto-regulation of the adenovirus type 5 E1A transcription unit. *Mol. Cell. Biol.* 5:3214-21.

Hearing, P. and T. Shenk (1983b). The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell* 33:695-703.

Heberman, R.B. and J.R. Ortaldo (1981). Natural killer cells: their role in defences against disease. *Science* 214:305-377.

Hen, R., E. Borrelli, P. Sassone-Corsi, and P. Chambon (1983). An enhancer element is located 340 base pairs upstream from the adenovirus-2 E1A cap site. *Nucl. Acids Res.* 11:8747-61.

Hen, R., E. Borrelli, and P. Chambon (1985). Repression of the immunoglobulin heavy chain enhancer by the adenovirus-2 E1A products. *Science* 230:1391-94.

Herbst, R.S., H. Hermo, Jr., P.B. fisher, and L.E. Babiss (1988). Regulation of adenovirus and cellular gene expression and of cellular transformation by the E1B-encoded 175-amino-acid protein. *J. Virol.* 62:4634-43.

Hermann, C.H., C.V. Dery, and M.B. Mathews (1987). Transactivation of host and viral genes by the adenovirus E1B 19k tumor antigen. *Oncogene* 2:25-35.

Heynecker, H.L., J. Shine, H.M. Goodman, H.W. Boyer, J. Rosenberg, R.E. Dickerson, S.A. Narang, K. Itakura, S. Lin, and A.D. Riggs (1976). Synthetic lac operator DNA is functional *in vivo*. *Nature* (London) **263**:748-52.

Hilleman, M.R., R.A. Stallows, R.L. Gould, M.S. Waterfield, and S.A. Anderson (1956). Prevention of acute respiratory illness in recruits by adenovirus. *Proc. Soc. Exp. Biol. Med.* **92**:377.

Hoeffler, W.K., and R.G. Roeder (1985). Enhancement of RNA polymerase III transcription by the E1A gene product of adenovirus. *Cell* **41**:955-63.

Hoeffler, W.K., R. Kovelman, and R.G. Roeder (1988). Activation of transcription factor IIIC by the adenovirus E1A protein. *Cell* **53**:907-20.

Horikoshi, M., M.F. Carey, H. Kakidani, and R.G. Roeder (1988b). Mechanism of action of a yeast activator: direct effect of GAL4 derivatives on mammalian TFIID-promoter interactions. *Cell* **54**:665-669.

Horikoshi, M., T. Hai, Y.-S. Lin, M.R. Green, and R.G. Roeder (1988a). Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* **54**:1033-42.

Houweling, A., van den Elsen, P.J. and van der Eb, A.J. (1980). Partial transformation of primary rat cells by the left-most 4.5% fragment of adenovirus 5 DNA. *Virology* **105**:537-550.

Huebner, R.J., W.P. Rowe, and W.T. Lane (1962). Oncogenic effects in hamsters of human adenovirus types 12 and 18. *Proc. Natl. Acad. Sci. USA* **48**:2051.

Hurst, H.C., and N.C. Jones (1987). Identification of factors that interact with the E1A-inducible adenovirus E3 promoter. *Genes Dev.* **1**:1132-46.

Imler, J.L., C. Schatz,C. Wasylyk, B. Chatton, and B. Wasylyk (1988). A Harvey-*ras* responsive transcription element is also responsive to a tumor-promoter and to serum. *Nature* **332**:275-78.

Imperiale, M.J. and J.R. Nevins (1984). Adenovirus 5 E2 transcription unit: an E1A-inducible promoter with an essential element that functions independently of position and orientation. *Mol. Cell. Biol.* **4**:875-82.

Imperiale, M.J., L.T. Feldman, and J.R. Nevins (1983). Activation of gene expression by adenovirus and herpesvirus regulatory genes acting in trans and by a cis-acting adenovirus enhancer element. *Cell* **35**:127-36.

Imperiale, M.J., R.P. Hart, and J.R. Nevins (1985). An enhancer-like element in the

adenovirus E2 promoter contains sequences essential for uninduced and E1A-induced transcription. *Proc. Natl. Acad. Sci. USA* **82**:381-85.

Jacobsen, H., H. Klenow, and K. Ovargaard-Hansen (1974). The N-terminal amino-acid sequences of DNA polymerase I from *Escherichia coli* and of the large and the small fragments obtained by a limited proteolysis. *Eur. J. Biochem.* 45:623.

Jalinot, P. and C. Kedinger (1986). Negative regulatory sequences in the E1A-inducible enhancer of the adenovirus-2 early EIIa promoter. *Nucl. Acids Res.* **14**:2651-69.

Jelsma, T.N., Howe, J.A., Evelegh, C.M., Cunniff, N.F., Skiadopoulos, M.H., Floroff, M.R., Denman, J.E. and Bayley, S.T. (1988). Use of deletion and point mutants dpanning the coding region of the adenovirus 5 E1A gene to define a domain that is essential for transcriptional activation. *Virology* **163**:494-502.

Jenkins, J.R., K. Rudge, and G.A. Currie (1984). Cellular immortalization by a cDNA clone encoding the transformation-associated phosphoprotein p53. *Nature* **312**:651-654.

Jochemsen, A.G., L.T.C. Peltenburg, M.F.W. te Pas, C.M. de Wit, J.L. Bos, and A.J. van der Eb (1987). Activation of adenovirus 5 E1A transcription by region E1B in transformed primary rat cells. *EMBO J.* **6**:3399-3405.

Johnson, P.F., W.H. Landshulz, B.J. Graves, and S.L. McKnight (1987). Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. *Genes Dev.* **1**:133-146.

Jones, K.A, K.R. Yamamoto, and R. Tjian (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. *Cell* **42**:559-572.

Jones, K.A., J.T. Kadonaga, P.J. Rosenfeld, T.J. Kelly, and R. Tjian (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48**:79-89.

Kaczmarek, L., Ferguson, B., Rosenberg, M. and Baserga, R. (1986). Induction of cellular DNA synthesis by purified adenovirus E1A proteins. *Virology* **152**:1-10.

Kampe, O., D. Bellgrau, U. Hammerling, P. Lind, S. Paabo, L. Severinson, and P.A. Peterson (1983). Complex formation of class I transplantation antigens and a viral glycoprotein. *J. Biol. Chem.* **258**:10594-8.

Kalderon, D. and A.E. Smith (1984). In vitro mutagenesis of a putative DNA

binding domain of SV40 large-T. *Virology* **139**:109-137.

Kessler, K. and H.-J. Holtke (1986). Specificity of restriction endonucleases and methylases--a review (Edition 2). *Gene* **47**:1-153.

Kimmelman, D., Miller, J.S., Porter, D. and Roberts, B.E. (1985). E1A regions of the human adenoviruses and of the highly oncogenic simian adenovirus 7 are closely related. *J. Virol.* **53**:399-409.

Klug, A. and D. Rhodes (1987). 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* **12**:464-9.

Korneluk, R.G., G. Quan, and R.A. Gravel (1985). Rapid and reliable dideoxy sequencing of double-stranded DNA. *Gene* **40**:317-323.

Kovesdi, I., R. Reichel, and J.R. Nevins (1986a). Identification of a cellular factor involved in E1A trans-activation. *Cell* **45**:219-228.

Kovesdi, I., R. Reichel, and J.R. Nevins (1986b). E1A transcription induction: enhanced binding of a factor to upstream promoter sequences. *Science* **231**:719-22.

Kovesdi, I., R. Reichel, and J.R. Nevins (1987). Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc. Natl. Acad. Sci. USA* **84**:2180-84.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiation codon that modulates translation by eukaryotic ribosomes. *Cell* **44**:283-292.

Krebs, E.G. and J.A. Beavo (1979). Phosphorylation-dephosporylation of enzymes. *Ann. Rev. Biochem.* **48**:923-59.

Kuo, J.F. and P. Grengard (1969). Cyclic nucleotide-dependent protein kinases IV. Widespread occurrence of cAMP-dependent protein kinase in various tissues and phyla of the animal kingdom. *Proc. Natl. Acad. Sci. USA* **64**:1349-55.

Kvist, S., L. Ostberg, H. Persson, L. Phillipson, and P.A. Peterson (1978). Molecular association between transplantation antigens and surface antigen in adenovirus-transformed cell line. *Proc. Natl. Acad. Sci. USA* **75**:5674-8.

Lai Fatt, R.B. and Mak, S. (1982). Mapping of an adenovirus function involved in the inhibition of DNA degradation. *J. Virol.* **42**:969-977.

Lamers, W.H., R.W. Hanson, and H.W. Miesner (1982). cAMP stimulates transcription of the gene for cytosol phosphoenolpyruvate carboxykinase in rat liver nuclei. *Proc. Natl. Acad. Sci. USA* **79**:5737-41.

Land, H., Parada, L.F. and Weinberg, R.A. (1983). Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature* **304**:596-602.

Landshulz, W.H., P.F. Johnson, E.Y. Adashi, B.J. Graves, and S.L. McKnight (1988). Isolation of a recombinant copy of the gene encoding C/EBP. *Genes Dev.* **2**:786-800.

Lee, K.A., T.Y. Hai, L. SivaRaman, B. Thimmappaya, H.C Hurst, N.C. Jones and M.R. Green (1987). A cellular protein, activating transcription factor, activates transcription of multiple E1A-inducible adneovirus early promoters. *Proc. Natl. Acad. Sci. USA* **84**:8355-9.

Lee, W., A. Haslinger, M. Karin, and R. Tjian (1987a). Activation of transcription by two factors that bind promoter and enhancer sequences of the human matallothionein gene and SV40. *Nature* **325**:368-72.

Lee, W., P. Mitchell, and R. Tjian (1987b). Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* **49**:741-52.

Lee., K.A.W. and M.R. Green (1987). A cellular transcription factor E4F1 interacts with an E1A-inducible enhancer and mediates constitutive function in vitro. *EMBO J.* **6**:1345-1353.

Leff, T. and P. Chambon (1986). Sequence-specific activation of transcription by adenovirus E1A products is observed in HeLa cells but not in 293 cells. *Mol. Cell. Biol.* **6**:201-208.

Leff, T., Elkaim, R., Goding, C.R., Jalinot, P., Sassone-Corsi, P., Perricaudet, M., Kedinger, C. and Chambon, P. (1984). Individual products of the adenovirus 12S and 13S E1A mRNAs stimulate viral EIIa and EIII expression at the transcriptional level. *Proc. Natl. Acad. Sci. USA* **81**:4381-4385.

Leff, T., J. Corden, R. Elkaim, and P. Sassone-Corsi (1985). Transcriptional analysis of the adenovirus-5 EIII promoter: absence of sequence specificity for stimulation by E1A gene products. *Nucl. Acids Res.* **13**:1209-21.

Lewis, E., Diann, J.L. Manley (1985). Control of adenovirus late promoter expression in two human cell lines. *Mol. Cell. Biol.* **5**:2433-42.

Lewis, E.J., C.A. Harrington, D.M. Chikaraishi (1987). Transcriptional regulation of the tyrosine hydroxylase gene by glucocorticoid and cyclic AMP. *Proc. Natl. Acad. Sci. USA* **84**:3550-54.

Leza, M.A., and P. Hearing (1988). Cellular transcription factor binds to adenovirus early region promoters and to a cyclic AMP response element. *J. Virol.* **62**:3003-13.

Lillie, J.W., Loewenstein, P.M., Green, M.R. and Green, M. (1987). Functional domains of adenovirus type 5 E1a proteins. *Cell* **50**:1091-1100.

Lillie, J.W., M. Green, and M.R. Green (1986). An adenovirus E1A protein region required for transformation and transcriptional repression. *Cell* **46**:1043-1051.

Lin, Y-S., M.F. Carey, M. Ptashne, and M.R. Green (1988). GAL4 derivatives function alone and synergistically with mammalian activators in vitro. *Cell* **54**:659-64.

Lin, Y.-S., and M.R. Green (1988). Interaction of a common cellular transcription factor, ATF, with regulatory elements of both E1A and cyclic AMP inducible promoters. *Proc. Natl. Acad. Sci. USA* **85**:3396-3400.

Lyons, R.H., B.Q. Ferguson, and M. Rosenberg (1987). Pentapeptide nuclear localization signal in adenovirus E1A. *Mol. Cell. Biol.* **7**:2451-2456.

Mackey, J.K., P.M. Rigden, and M. Green (1976). Do highly oncogenic group A human adenoviruses cause human cancer? Analysis of human tumors for adenovirus 12 transforming DNA sequences. *Proc. Natl. Acad. Sci. USA* **73**:4657-61.

Maity, S.N., P.T. Golumbek, G. Karsenty, B. de Crombrugghe (1988). Selective activation of transcription by a novel CCAAT binding factor. *Science* **241**:582-585.

Maki, Y., T.J. Bos, C. Davis, M. Starbuck, and P.K. Vogt (1987). Avian sarcoma virus 17 carries the jun oncogene. *Proc. Natl. Acad. Sci. USA* **84**:2848-52.

Mandel, M. and Higa, A. (1970). Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**:159-162.

Maniatis, T., E.F. Fritsch, and J. Sambrook (1982). In *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Marians, K.J., R.Wu, J. Stawinski, T. Hozumi, and S.A. Narang (1976). Cloned synthetic *lac* operator DNA is functional *in vivo*. *Nature* (London) **263**:744-8.

Martin, G.R. (1980). Teratocarcinomas and mammalian embryogenesis. *Science* **209**:768-776.

Mathis, D.J., R. Elkaim, C. Kedinger, P. Sassone-Corsi, and P. Chambon (1981). Specific in vitro initiation of transcription on the adenovirus type 2 early

and late EII transcription units. *Proc. Natl. Acad. Sci. USA* **78**:7383-87.

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

McGlade, C.J., M.L. Tremblay, S.-P. Yee, R. Ross, and P.E. Branton (1987). Acylation of the 176R (19 kilodalton) early region 1B protein of human adenovirus 5. *J. Virol.* **61**:3227-3234.

McGrory, W.J. (1988). M. Sc. Thesis. McMaster University.

McGrory, W.J., Bautista, D.S. and Graham, F.L. (1988). A simple technique for the rescue of early region 1 mutations into infectious human adenovirus type 5. *Virology* **163**:614-617.

McKinnon, R.D., J.S. Waye, D.S. Bautista, and F.L. Graham (1985). Nonrandom insertion of Tn5 into cloned human adenovirus DNA. *Gene* **40**:31-38.

McKnight, S. and R. Tjian (1986). Transcriptional selectivity of viral genes in mammalian cells. *Cell* **46**:795-806.

McKnight, S.L. (1982). Functional relationships between transcriptional control signals of the thymidine kinase gene of herpes simplex virus. *Cell* **31**:355-65.

Messing, J. (1983). New M13 vectors for cloning. *Methods Enzymol.* **101**:20-78.

Messing, J. and J. Vieira (1982). A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* **19**:269-276.

Miller, J., A.D. McLachlan, and A. Klug (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**:1609-14.

Miller, J.H. (1972). *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold spring Harbor, New York.

Miller, J.H. and W.S. Reznicoff (1978). *The Operon.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Montell, C., E.F. Fisher, M.H. Caruthers, and Berk, A.J. (1982). Resolving the functions of overlapping viral genes by site-specific mutagenesis at a mRNA splice site. *Nature* **295**:380-384.

Montell, C., G. Courtois, C. Eng, and A.J. Berk (1984). Complete transformation by adenovirus 2 requires both E1A proteins. *Cell* **36**:951-61.

Montminy, M.R. and L.M. Bilezikjian (1987). Binding of a nuclear protein to the cAMP response element of the somatostatin gene. *Nature* **328**:175-8.

Montminy, M.R., K.A. Sevarino, J.A. Wagner, G. Mandel, and R.H. Gooman (1986). Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. *Proc. Natl. Acad. Sci. USA* **83**:6682-86.

Moran (1988). A region of SV40 large T antigen can substitute for a transforming domain of the adenovirus E1A products. *Nature* **334**:168-170.

Moran, E. and M.B. Mathews (1987). Multiple functional domains in the adenovirus E1A gene. *Cell* **48**:177-178.

Moran, E., Grodzicker, T., Roberts, R.J., Mathews, M.B. and Zerler, B. (1986b). Lytic and transforming functions of individual products of the adenovirus E1A gene. *J. Virol.* **57**:765-775.

Murray, J.D., A.J.D. Bellet, A.W. Braithwaite, L.K. Waldron, and I.W. Taylor (1982). Altered cell cycle progression and aberrant mitosis in adenovirus-infected rodent cells. *J. Cell. Physiol.* **111**:89-96.

Murray, N.E. et al. (1977). Lambdoid phages that simplify the recovery of in vitro recombinants. *Mol. Gen. Genet.* **150**:53-61.

Murthy, S.C.S, G.P Bhat, and B. Thimmappaya (1985). Adenovirus EIIA early promoter: Transcriptional control elements and induction by the viral pre-early E1A gene, which appears to be sequence independent. *Proc. Natl. Acad. Sci. USA* **82**:2230-2234.

Nagata, K., R.A. Guggenheimer, and J. Hurwitz (1983a). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. *Proc. Natl. Acad. Sci. USA* **80**:6177-6181.

Nakabeppu, Y., K. Ryder,and D. Nathans (1988). DNA binding activities of three murine jun proteins: stimulation by Fos. *Cell* **55**:907-15.

Nakagawa, J., D. von der Ahe, D. Perason, B.A. Hemmings, S. Shibahara, and Y. Nagamine (1988). Transcriptional regulation of a plasminogen activator gene by cAMP in a homologous cell-free system. *J. Biol. Chem.* **263**:2460-8.

Nakajima, T., M. Masuda-Murata, E. Hara, and K. Oda (1987). Induction of cell cycle progression by adenovirus E1A gene 13S- and 12S-mRNA products in quiescent rat cells. *Mol. Cell. Biol.* **7**:3846-52.

Natarajan, V. and N.P. Salzman (1985). Cis and trans-activation of adenovirus IVa2 gene transcription. *Nucl. Acids Res.* **13**:4067-83.

Nevins, J. and J. Winkler (1980). Regulation of early adenovirus transcription: a protein product of early region 2 specifically represses region 4 transcription. *Proc. Natl. Acad. Sci. USA* 77:1893-97.

Nevins, J.R. (1981). Mechanism of activation of early viral transcription by the adenovirus E1A gene product. *Cell* 26:213-220.

Nevins, J.R. (1982). Induction of the synthesis of a 70,000 dalton mammalian heat shock protein by the adenovirus E1a gene product. *Cell* 29:913-919.

Nimmo, H.G. and P. Cohen (1977). Hormonal control of protein phosphorylation. *Adv. Cyclic Nucleotide Res.* 8:146-266.

Nishigaki, T., S. Hanaka, R.E. Kingston, and H. Handa (1988). A specific domain of the adenovirus EIV promoter is necessary to maintain susceptibility of the integrated promoter to E1A transactivation. *Mol. Cell. Biol.* 8:353-360.

Nishizuka, Y. (1986). Studies and perspectives of protein kinase C. *Science* 233:305-310.

Niven, J.S.F, J.A. Armstrong,C.H. Andrewes, H.G. Pereira, and R.C. Valentine (1961). Subcutaneous "growths" in monkeys produced by poxvirus. *J. Pathol. Bacteriol.* 81:1.

Nordeen, S.K., P.P Green and D.M. Fowlkes (submitted to DNA). A rapid, sensitive, and inexpensive assay for chloramphenicol acetyltransferase. *DNA* 6:173-8.

Oikarinen, J., A. Hatamochi, and B. de Crombrugghe (1987). Separate binding sites for nuclear factor 1 and a CCAAT DNA binding factor in the mouse alpha2(I). collagen promoter. *J. Biol. Chem.* 262:11064-11070.

Osborne, T.F. and A.J. Berk (1983). Far upstream initiation sites for adenovirus early region 1A transcription are utilized after the onset of viral DNA replication. *J. Virol.* 45:594-99.

Osborne, T.F., D.N. Arvidson, E.-S. Tyau, m. Dunsworth-Browne, and A.J. Berk (1984). Transcription control region within the protein-coding portion of adenovirus E1A genes. *Mol. Cell. Biol.* 4:1293-305.

Osborne, T.F., R.B. Gaynor, and A.J. Berk (1982). The TATA homology and the mRNA 5' untranslated sequence are not required for expression of essential adenovirus E1A functions. *Cell* 29:139-48.

Paabo, S., F. Weber, T. Nilsson, W. Schaffner, and P.A. Peterson (1986). Structural and functional dissection of an MHC class I antigen-binding adenovirus glycoprotein (1986). *EMBO J.* 5:1921-7.

Paabo, S., T. Nilsson, and P.A. Andersson (1986). Adenoviruses of subgenera B, C, D, and E modulate cell-surface expression of major histocompatibility complex class I antigens. *Proc. Natl. Acad. Sci. USA* 83:9665-9.

Paabo, S., F. Weber, O. Kampe, W. Shaffner, and P.A. Peterson (1983). Association between transplantation antigens and a viral membrane protein synthesized form a mammalian expression vector. *Cell* 33:445-53.

Parada, L.F., H. Land, R.A. Weinberg, D. Wolf, and V. Rotter (1984). Cooperation between gene encoding p53 tumour antigen and ras in cellular transformation. *Nature* 312:649-51.

Parker, C.S., and J. Topol (1984). A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell* 36:357-369.

Parker, R.C., R.M. Watson, and M. Vinograd (1977). Mapping of closed circular DNAs by cleavage with restriction endonucleases and calibration by agarose gel electrophoresis. *Proc. Natl. Acad. Sci. USA* 74:851-55.

Perricaudet, M., Akusjarvi, G., Virtanen, A. and Pettersson, U. (1979). Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses. *Nature* 281:694-696.

Phelps, W.C., C.L. Yee, K. Munger, and P.M. Howley (1988). The human papillomavirus type 16 E7 gene encodes transactivation and transformation function similar to those of adenovirus E1A. *Cell* 53:539-547.

Pilder, S., Logan, J. and Shenk, T. (1984). Deletion of the gene encoding the adenovirus 5 early region 1B 21,000-molecular-weight polypeptide leads to degradation of viral and host cell DNA. *J. Virol.* 52:664-671.

Pilder, S., Moore, M., Logan, J. and Shenk, T. (1986). The adenovirus E1B-55K transforming polypeptide modulates transport or cytoplasmic stabilization of viral and host cell mRNAs. *Mol. Cell. Biol.* 6:470-476.

Quinlan, M.P. and T. Grodzicker (1987). Adenovirus E1A 12S protein induces DNA synthesis and proliferation in primary epithelial cells in both the presence and absence of serum. *J. Virol.* 61:673-682.

Quinlan, M.P., N. Sullivan, and T. Grodzicker (1987). Growth factor(s) produced during infection with an adenovirus variant stimulates proliferation of non-established epithelial cells. *Proc. Natl. Acad. Sci. USA* 84:3283-87.

Ran, W., M. Dean, R.A. Levine, C. Henkle and J. Campisi (1986). Induction of c-*fos* and c-*myc* mRNA by epidermal growth factor or calcium ionosphere is

cAMP dependent. *Proc. Natl. Acad. Sci. USA* **83**:8216-20.

Raymondjean, M., S. Cereghini, and M. Yaniv (1988). Several distinct "CCAATT" box binding proteins coexist in eukaryotic cells. *Proc. Natl. Acad. Sci. USA* **85**:757-761.

Rossini, M. (1983). The role of adenovirus early region 1A in the regulation of early regions 2A and 1B expression. *Virology* **131**:49-58.

Rowe, W.P., R.. Huebner, L.K. Gillmore, R.H. Parrot, and T.G. Ward (1953). Isolation of a cytogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Biol. Med.* **84**:570.

Ruley, H.E. (1983). Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. *Nature* **304**:602-606.

Sachs, L. and E. Winocour (1959). Formation of different cell-virus relationships in tumour cells induced by polyoma. *Nature* **184**:1702.

Sadler, J.R., M. Tecklenburg, J.L. Betz, D.V. Goeddel, D.G. Yansura, and M.H. Caruthers (1977). Cloning of chemically synthesized lactose operators. *Gene* **1**:305-321.

Samulski, R.J. and T. Shenk (1988). Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J. Virol.* **62**:206-10.

Sassone-Corsi, P. (1988). Cyclic-AMP induction of early adenovirus promoters involves sequences required for E1A trans-activation. *Proc. Natl. Acad. Sci. USA* **85**:7192-96.

Sassone-Corsi, P., R. Hen, E. Borrelli, T. Leff, and P. Chambon (1983). Far upstream sequences are required for efficient transcription from the adenovirus 2-E1A transcription unit. *Nucl. Acids Res.* **11**:8735-45.

Sassone-Corsi, P., and E. Borreli (1987). Promoter transactivation of proto-oncogenes c-fos and c-myc, but not c-Ha-ras, by products of adenovirus early region 1A. *Proc. Natl. Acad. Sci. USA* **84**:6430-6433.

Sawada, Y., B. Fohring, T.E. Shenk, and K. Raska (1985). Tumorigenicity of adenovirus-transformed cells: region E1A of adenovirus 12 confers resistance to natural killer cells. *Virology* **147**:413-21.

Sawadogo, M. and R.G. Roeder (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA-box region. *Cell* **43**:165-75.

Schonthal, A., P. Herrlich, H.J. Rahmsdorf, and H. Ponta (1988). Requirement for fos gene expression in the transcriptional activation of collagenase by other oncogenes and phorbol esters. *Cell* **54**:325-34.

Schrier, P.I., R.T. Bernards, M.J. Vaessen, A. Houweling, and A.J. van der Eb (1983). Expression of class I major histocompatibility antigens switched off by highly oncogenic adenovirus 12 in transformed rat cells. *Nature* **305**:771-75.

Schwab, M., Varmus, H.E., and J.M. Bishop (1985). Human N-myc gene contributes to neoplastic transformation of mammalian cell in culture. *Nature* **316**:160-62.

Severinsson and Peterson (1985). Abrogation of cell surface expression of human class I transplantation antigens by an adenovirus protein in *Xenopus laevis* oocytes. *J. Cell Biol.* **101**:540-7.

Severinsson, L. I. Martens, and P.A. Peterson (1986). Differential association between two human MHC class I antigens and an adenoviral glycoprotein. *J. Immunol.* **137**:1003-9.

Sharp, P.A. (1984). Adenovirus Transcription. In *The Adenoviruses* (Ginsberg, H.S., ed.). Plenum Press, New York, pp.173-204.

Shenk, T., N. Jones, W. Colby, and D. Fowlkes (1979). Functional analysis of adenovirus-5 host-range deletion mutants defective for transformation of rat embryo cells. *Cold Spring Harbor Symp. Quant. Biol.* **44**:367-75.

Shimojo, H. and Yamashita, T. (1968). Induction of DNA synthesis by adenoviruses in contact-inhibited hamster cells. *Virology* **36**:422-433.

Shiroki, K., and M. Toth (1988). Activation of the human beta interferon gene by the adenovirus type 12 E1B gene. *J. Virol.* **62**:325-30.

Short, J.M., A. Wynshaw-Boris, H.P. Short, and R.W. Hanson (1986). Characterization of the phosphoenolpyruvate carboxykinase (GTP). promoter regulatory region. *J. Biol. Chem.* **261**:9721-26.

Signas, C., M.G. Katze, H. Persson, and L. Phillipson (1982). An adenovirus glycoprotein binds heavy chains of class I transplantation antigens from man and mouse. *Nature* **299**:175-8.

Silver, J.J., J.A. Bokar, J.B. Virgin, E.A. Vallen, A. Milsted, and J.M. Nilson (1987). Cyclic AP regulation of the human glycoprotein hormone alpha-subunit gene is mediated by an 18-base-pair element. *Proc. Natl. Acad. Sci. USA* **84**:2198-2202.

Simon, M.C., K. Kitchener, H.T. Kao, E. Hickey, L. Weber, R. Voellmy, N. Heintz, and J.R. Nevins (1987). Selective induction of human heat shock gene transcription by the adenovirus E1A gene products including the 12S E1A product. *Mol. Cell. Biol.* 7:2884-90.

Simon, M.C., T.M. Fisch, B.J. Benecke, J.R. Nevins, and N. Heintz (1988). Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation. *Cell* 52:723-729.

SivaRaman, L., and B. Thimmappaya (1987). Two promoter-specific host factors interact with adjacent sequences in an E1a-inducible adenovirus promoter. *Proc. Natl. Acad. Sci. USA* 84:6112-6.

SivaRaman, L., S. Subramanian, and B. Thimmappaya (1986). Identification of a factor in HeLa cells specific for an upstream transcriptional control sequence of an E1A-inducible adenovirus promoter and its relative abundance in infected and uninfected cells. *Proc. Natl. Acad. Sci. USA* 83:5914-18.

Smith, D.H., Kegler, D.M. and Ziff, E.B. (1985). Vector expression of adenovirus type 5 E1A proteins: Evidence for E1a autoregulation. *Mol. Cell. Biol.* 5:2684-2696.

Southern, P.J. and P. Berg (1982). [pSV2neo ref] *J. Mol. Appl. Genet.* 1:327-41.

Spector, D.J., Crossland, L.D., Halbert, D.N. and Raskas, H.J. (1980). A 28K polypeptide is the translation product of 9S RNA encoded by region 1A of adenovirus 2. *Virology* 102:218-221.

Spiegel, M.R. (1975). *Theory and Problems of Probability and Statistics* in Schaum's Outline Series, McGraw-Hill Book, Inc., N.Y.

Sprang, S.R., K.R. Acharya, E.J. Goldsmith, D.I. Stuart, K. Varvill, R.J. Fletterick, N.B. Madsen and L.N. Johnson (1988). Structural changes in glycogen phosphorylase induced by phosphorylation. *Nature* 336:215-221.

Stein, R.W. and Ziff, E.B. (1987). Repression of insulin gene expression by adenovirus type 5 E1A proteins. *Mol. Cell. Biol.* 7:1164-1170.

Stein, R.W. and Ziff, E.B. (1984). HeLa cell beta-tubulin gene transcription is stimulated by adenovirus 5 in parallel with viral early genes by an E1a-dependent mechanism. *Mol. Cell. Biol.* 4:2792-2801.

Stephens, C. and Harlow, E. (1987). Differential splicing yields novel adenovirus 5 E1A mRNAs that encode 30 kd and 35 kd proteins. *EMBO J.* 6:2027-2035.

Stewart, S.E., B.E. Eddy, and N. Borgese (1958). Neoplasms in mice inoculated with a tumor agent carried in tissue culture. *J. Natl. Cancer Inst.* **20**:1223.

Stillman, B.W., J.B. Lewis, L.T. Chow, M.B. Mathews, and J.E. Smart (1981). Identification of the gene and mRNA for the adenovirus terminal protein precursor. *Cell* **23**:497-508.

Subramanian, T., M. Kappuswamy, J. Gysbers, S. Mak, and G. Chinnadurai (1984a). 19 kDa tumor antigen coded by early region E1b of adenovirus 2 is required for efficient synthesis and for protection of viral DNA. *J. Biol. Chem.* **259**:11777-83.

Subramanian, T., M. Kappuswamy, S. Mak, G. Chinnadurai (1984b). Adenovirus cyt$^+$ locus, which controls cells transformation and tumorigenicity, is an allele of lp$^+$ locus, which codes for a 19-kilodalton tumor antigen. *J. Virol.* **52**:336-343.

Subramanian, T., M. Kappuswamy, R.J. Nasr, and G. Chinnadurai (1988). An N-terminal region of adenovirus E1A essential for cell transformation and induction of an epithelial growth factor. *Oncogene* **2**:105-112.

Sutherland, E.W. (1972). Studies on the mechanism of hormone action. *Science* **177**:401-408.

Svensson, C. and Akusjarvi, G. (1984). Adenovirus 2 early region 1A stimulates expression of both viral and cellular genes. *EMBO J.* **3**:789-791.

Takemori, N., Cladaras, C., Bhat, B., Conley, A.J. and Wold, W.S.M. (1984). cyt gene of adenoviruses 2 and 5 is an oncogene for transforming function in early region E1B and encodes the E1B 19,000-molecular-weight polypeptide. *J. Virol.* **52**:793-805.

Takemori, N., J.L. Riggs, and C. Aldrich (1968). Genetic studies with tumorigenic adenoviruses. I. Isolation of cytocidal (cyt) mutants of adenovirus type 12. *Virology* **36**:575-86.

Tanaka, Y. and S.S. Tevethia (1988). Differential effect of adenovirus 2 E3/19K glycoprotein on the expression of H-2K^b and H-2D^b class I antigens and H-2K^b- and H-2D^b-restricted SV40-specific CTL-mediated lysis. *Virology* **165**:357-66.

Toda, T., S. Cameron, P. Sass, M. Zoller, and M. Wigler (1987). Three different genes in S. cerevisiae encode the catalytic subunits of the cAMP-dependent protein kinase. *Cell* **50**:277-287.

Tooze, J. (1981). *The Molecular Biology of Tumor Viruses*, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Toth, M.I., B. Arya, R. Pusztai, K. Shiroki, and I. Beladi (1987). Interferon induction by adenovirus type 12: stimulatory function of early region 1A. *J. Virol.* **61**:2326-30.

Trentin, J.J., Yabe, Y. and Taylor, G. (1962). The quest for human cancer viruses. *Science* **137**:835-841.

Tsukamoto, A.S., A. Ponticelli, A.J. Berk, and R.B. Gaynor (1986). Genetic mapping of a major site of phosphorylation in adenovirus type 2 E1A proteins. *J. Virol.* **59**:14-22.

Ulfendahl, P.J., S. Linder, J.-P. Kreivi, K. Nordqvist, C. Sevensson, H. Hultberg, and G. Akusjarvi (1987). A novel adenovirus-2 E1A mRNA encoding a protein with transcription activation properties. *EMBO J.* **6**:2037-2044.

Vaessen, R.T.M.J., A. Houweling, and A.J. van der Eb (1987a). Post-translational control of class I MHC mRNA expression in adenovirus 12-transformed cells. *Science* **235**:1486-8.

Van Buskirk, R., T. Corcoran, and J.A. Wagner (1985). Clonal variants of PC12 pheochromocycloma cells with defects in cAMP-dependent protein kinases induce ornithine decarboxylase in response to nerve growth factor but not to adenosine agonists. *Mol. Cell. Biol.* **19**:84-92.

van den Elsen, P., Houweling, A. and van der Eb, A. (1983c). Expression of region E1b of human adenoviruses in the absence of region E1a is not sufficient for complete transformation. *Virology* **128**:377-390.

Velcich, A. and E. Ziff (1988). Adenovirus E1A ras cooperation activity is separate from its positive and negative transcription regulatory functions. *Mol. Cell. Biol.* **8**:2177-2183.

Velcich, A. and Ziff, E. (1985). Adenovirus E1a proteins repress transcription from the SV40 early promoter. *Cell* **40**:705-716.

Velcich, A., F.G. Kern, C. Basilico and E.B. Ziff (1986). Adenovirus E1A proteins repress expression from polyomavirus early and late promoters. *Mol. Cell. Biol.* **6**:4019-25.

Vogt, M. and R. Dulbecco (1960). Virus-cell interaction with a tumor-producing virus. *Proc. Natl. Acad. Sci. USA* **46**:365.

Vogt, P.K., T.J. Bos,and R.F. Doolittle (1987). Homology between the DNA binding domain of the GCN4 regulatory protein of yeast and the carboxy-terminal region of a protein coded for by the oncogene jun. *Proc. Natl. Acad. Sci. USA* **84**:3316-19.

Walter, U., J. Uno, A.Y.C. Liu, and P. Greengard (1977). Identification, characterization, and quantitative measurement of cyclic AMP receptor proteins in cytosol of various tissues using a photo-affinity ligand. *J. Biol. Chem.* **252**:6494-6500.

Wasylyk, C., J.L. Imler, and B. Wasylyk (1988). Transforming but not immortalizing oncogenes activate the transcription factor PEA1. *EMBO J.* **7**:2475-83.

Watanabe, H., T. Imai, P.A. Sharp, and H. Handa (1988). Identification of two transcription factors that bind to specific elements in the promoter of the adenovirus early-region 4. *Mol. Cell. Biol.* **8**:1290-1300.

Weeks, D.L. and N.C. Jones (1983). E1A control of gene expression is mediated by sequences 5' to the transcriptional starts of the early viral genes. *Mol. Cell. Biol.* **3**:1222-1234.

Weeks, D.L. and N.C. Jones (1985). Adenovirus E3-early promoter sequences required for activation by E1A. *Nucl. Acids Res.* **13**:5389-402.

Weiher, H., M. Konig, and P. Gruss (1983). Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219**:626-31.

Weinberg, D.H., and G. Ketner (1986). Adenoviral early region 4 is required for efficient viral DNA replication and for late gene expression. *J. Virol.* **57**:833-8.

Weiss, R., N. Teich, H. Varmus, and J. Coffin (1982). *RNA TUMOR VIRUSES: Molecular Biology of Tumor Viruses*, pp. 1-24, Cold Spring Harbor Laboratory, New York.

White, E., A. Denton, and B. Stillman (1988). Role of the adenovirus E1B 19,000-Dalton tumor antigen in regulating early gene expression. *J. Virol.* **62**:3445-54.

White, E. and B. Stillman (1987). Expression of adenovirus E1B mutant phenotypes is dependent on the host cell and on synthesis of E1A proteins. *J. Virol.* **61**:426-35.

White, E., Grodzicker, T. and Stillman, B.W. (1984). Mutations in the gene encoding the adenovirus early region 1B 19,000-molecular-weight tumor antigen cause the degradation of chromosomal DNA. *J. Virol.* **52**:410-419.

White, E., B. Faha, and B. Stillman (1986). Regulation of adenovirus gene expression in human W138 cells by an E1B-encoded tumor antigen. *Mol. Cell. Biol.* **6**:3763-73.

Wilson, J.S., P.J. Grant, D.L. Miller, C.E. Taylor, and J.C. MacDonald (1960). Trial of adenovirus vaccine on Royal Air Force recruits. *Brit. J. Med.* 1:1081.

Winberg, C. and T. Shenk (1984). Dissection of overlapping functions within the adenovirus type 5 E1A gene. *EMBO J.* 3:1907-12.

Wold, W.S.M., C. Cladaras, S.L. Deutscher, and Q.S. Kapoor (1985). The 19kDa glycoprotein by region E3 of adenovirus. *J. Biol. Chem.* 260:2424-31.

Wu, L. and A.J. Berk (1988a). Constraints on spacing between transcription factor binding sites in a simple adenovirus promoter. *Genes Develop.* 2:403-411.

Wu, L. and A.J. Berk (1988b). Transcriptional activation by the pseudorabies virus immediate protein requires the TATA box element in the adenovirus E1B promoter. *Virology* 167:318-322.

Wu, L., D.S.E. Rosser, M.C. Schmidt, and A. Berk (1987). A TATA box implicated in E1A transcriptional activation of a simple adenovirus 2 promoter. *Nature* 326:512-515.

Yancopoulos, G.D., P.D. Nisen, A. Tesfaye, N.E. Kohl, M.P. Goldfarb, and F.W. Alt (1985). N-myc can cooperate with ras to transform normal cells in culture. *Proc. Natl. Acad. Sci. USA* 82:5455-5459.

Yanish-Peron et al. (1985). Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vecotrs. *Gene* 33:103.

Yee, A.S., R. Reichel, I. Kovesdi and J.R. Nevins (1987). Promoter interaction of the E1a-inducible factor E2F and its potential role in the formation of a multi-component complex. *EMBO J.* 6:2061-2068.

Yee, S.-P. and P.E. Branton (1985a). Analysis of multiple forms of human adenovirus type 5 E1A polypeptides using an anti-peptide antiserum specific for the amino terminus. *Virology* 146:315-322.

Yee, S.-P. and P.E. Branton (1985b). Detection of cellular proteins associated with human adenovirus type 5 early region 1A polypeptides. *Virology* 147:142-153.

Yoder, S.S., and S.M. Bergert (1986). Role of adenovirus type 2 early region 4 in the early-to-late switch during productive infection. *J. Virol.* 60:779-81.

Yoshida, K., Venkatesh, L., Kuppuswamy, M. and Chinnadurai, G. (1987). Adenovirus transforming 19-kD T antigen has an enhancer-dependent trans-activation function and relieves enhancer repression mediated by viral and

cellular genes. *Genes and Dev.* 1:645-658.

Yoshinaga, S., N. Dean, M. Han, and A.J. Berk (1986). Adenovirus stimulation of transcription by RNA polymerase III: evidence for an E1A-dependent increase in transcription factor IIIC concentration. *EMBO J.* 3:43-54.

Zajchowski, D.A., H. Boeuf, and C. Kedinger (1985). The adenovirus-2 early EIIA transcription unit possesses two overlapping promoters with different sequence requirements for E1A dependent stimulation. *EMBO J.* 4:1293-1300.

Zajchowski, D.A., P. Jalinot, and C. Kedinger (1988). E1A-mediated stimulation of the adenovirus EIII promoter involves an enhancer element within the nearby EIIa promoter. *J. Virol.* 62:176-67.

Zerler, B, R.J. Roberts, M.B. Mathews, and E. Moran (1987). Different functional domains of the adenovirus E1A gene are involved in regulation of host cell cycle products. *Mol. Cell. Biol.* 7:821-829.

Zerler, B., Moran, B., Maruyama, K., Moomaw, J., Grodzicker, T. and Ruley, H.E. (1986). Adenovirus E1a Coding Sequences that enable ras and pmt Oncogenes to Transform Cultured Primary Cells. *Mol. Cell. Biol.* 6:887-899.