https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

AN ANALYSIS OF THE TRANSCRIPTIONAL CONTROL DOMAINS OF THE HUMAN c-*myc* PROTO-ONCOGENE

by

Christopher Bruce Alexander Whitelaw

A Thesis Presented for the Degree of
Doctor of Philosophy

in

The Faculty of Medicine
at the University of Glasgow

Beatson Institute for Cancer Research
Garscube Estate
Glasgow
G61 1BD

February 1987

ProQuest Number: 10995548

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10995548

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TO MY MOTHER, FATHER AND MY WEE FLO

CONTENTS

ACKNOWLEDGEMENTS

I wish to thank the Cancer Research Campaign for the funding of this work, through a Scholarship.

Dr. Neil M. Wilkie, my supervisor, for his invaluable support, advice and encouragement.

Dr. Jas C. Lang for his encouragement, discussion of results, and especially his initial assistance in the DNA transfection technique.

Dr. George D. Birnie for his helpful discussion.

Drs. T. Papas, B. Clements and J. Lang for supplying recombinant plasmids, and Dr. R. Tjian for supplying the purified protein fractions.

The friendship of the staff of the Beatson Institute.

Alison MacPherson for her typing excellence and patience.

The patience and support of Flora T. MacPherson and my family.

The understanding of Dr. A. John Clark.

The work included in this thesis was performed by the author.

ABREVIATIONS

A	adenine
bp	base pair
C	cytosine
Ci	curie
cm	centimeter
cpm	counts per minute
DTT	dithiothreitol
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
g	gram
G	guanine
hr	hour
HSV	herpes simplex virus
IE	immediate early
Kb	kilobase pair
L	litre
M	molar
min	minute
mg	milligram
ml	millilitre
mm	millimeter
mM	millimolar
mRNA	messenger ribonucleic acid
nm	nanometer
RNA	ribonucleic acid
sec	second
SDS	sodium dodecyl sulphate
T	thymidine
tk	thymidine kinase
TK	thymidine kinase negative
V	volts
v/v	volume per volume
W	watt
°C	degrees centigade
µg	microgram
µl	microlitre
%	percentage

SUMMARY

Transfection of mammalian cells with recombinant plasmid DNAs containing the bacterial chloramphenicol acetyl transferase (CAT) gene as a reporter has been used to analyse genomic sequences regulating the transcription of the human c-myc proto-oncogene. Several regulatory domains 5' to the c-myc coding region have been identified, and their locations defined by deletion analysis. Each of these sites encompass previously identified DNase I in vivo hypersensitive sites.

Published data suggested that the c-myc gene may be regulated in vivo by a repressor. The mapping of translocation breakpoints for Burkitt's lymphoma and murine plasmacytoma which involve the c-myc gene, suggests that the cis-acting recognition sequence for this putative repressor is located within the 5' flanking region. I have identified a negative regulatory element (NRE-2) in the 5' flanking region of the gene and localised it to a region between -1052 and -607bp 5' to the P1 start site of the c-myc mRNA, by deletion analysis. Subsequent competition experiments showed a 270bp sub-fragment to contain an essential component of the negative regulatory element. This element can function in an orientation independent manner, and has the ability to repress heterologous promoters (both viral and eukaryotic), but to a lesser degree than when acting in cis upon its homologous promoter. My data from both DNA titration and competition transfection analysis indicates that this repression is mediated by at least one trans-acting factor. Since the repression was observed in every cell line used as the transfection recipients, a certain promiscuity in the tissue- and species- specificity of the trans-acting repressor(s) is implied. In vivo footprint analysis tentatively identified two sequence-specific DNA-binding proteins which interact with this domain. Both the CCAAT-binding Transcription Factor (CTF) and Sp1 bind within the NRE-2 domain. This is the

first time either of these DNA-binding factors have been implicated in the transcriptional repression of a gene.

In addition, deletion analyses identified an Upstream Promoter Element (UPE), located between the NRE-2 and the c-myc mRNA major cap sites, which is responsible for activation of the high levels of CAT expression observed in cells transfected with the recombinant plasmids. This UPE appears to be a highly complex domain which was shown, by DNase I in vivo footprint analysis, to bind several Sp1-like factors. In addition, the UPE is somehow involved in the control of the repression function, although it is not required for the repression of heterologous promoters by the NRE-2.

Other data (assayed preliminary) suggested also that two other distal regulatory domains are involved in the control of c-myc expression. The more distal element (PRE) has an activating activity, and was localised to a region which showed sequence homology to enhancer elements. A second element (NRE-1) was tentatively identified which had a negative effect on CAT activity.

I conclude that the removal, and/or the rearrangement, of these transcription regulatory domains may play a crucial role in the deregulation of the expression of c-myc that is observed in some neoplastic cells.

CHAPTER ONE

INTRODUCTION

1.1. Introduction: Cancer and Oncogenes

Cancer can best be described as a cruel subversion from the normal growth parameters of a cell. Such a subversion, if not coded for by the cell, must be induced by external factors. In either case, since the majority of cells for the whole of their life-span pursue a normal existence, cancer must be due to the aberrant regulation of normal cellular processes. This aberrant control could manifest itself in many ways - temporally, qualitatively, quantitatively - but must act through disruption of the individual genes coding for the effectors of these cellular processes. The compelling questions are obvious: which genes, and how are they altered to give rise to the transformed phenotype?

Our first insight into the answers to these questions came from the study of a subset of animal viruses. The retroviruses are an heterogeneous group of RNA viruses (for review see Bishop, 1978), which can be further subdivided into two groups. The acutely transforming retroviruses contain a transforming gene (the viral oncogene), which we now know was derived by transduction of a cellular gene (for reviews see Bishop, 1983; Bishop and Varmus, 1985), while the slowly transforming retroviruses do not harbour a transforming gene but appear to activate cellular genes by promoter insertion (Hayward, Neel and Astrin, 1981). The cellular genes activated in either of these cases are termed proto-oncogenes, and we can now from this line of investigation account for about 25 of then. Upon further analysis, using several in vivo systems, these genes have been shown to be intimatly associated with the genesis of the neoplastic cell.

A separate line of research has led to the same conclusion. DNA-mediated gene transfer, which allows the identification of dominantly active oncogenes from tumour cells (for review see Marshall, 1985), has brought the total

number of proto-oncogenes to nearly 40. By their nature, these studies can only identify dominantly acting oncogenes; however, recent evidence from somatic cell hybridisation experiments has indicated the likelihood that recessive oncogenes (suppressor genes) may be involved in generating the transformed phenotype (for reviews see Stanbridge, 1985; Sager, 1986).

Thus, there appears to be a fairly large, very heterogeneous group of cellular genes which have the potential, when their actions are disrupted in some fashion, to cause cellular transformation. This concept has been supported by recent, elegant work using transgenic mice (Adams et al, 1985).

The members of the proto-oncogene family have been found to be highly conserved through evolution, being present in every mammalian and avian species examined to date, with *myb* and the *ras* family also represented in *Saccharomyces cerevisiae* (Katzen, Kornberg and Bishop, 1985). Such conservation implies essential and fundamental roles for these genes in normal cellular metabolism. These roles are only now beginning to be elucidated (for review see Bishop and Varmus, 1985). The gene products of a few proto-oncogenes have been identified. Not surprisingly these include a growth factor (*sis* and Platelet Derived Growth Factor; Waterfield et al, 1983) and growth factor receptors (*erb*-B and Epidermal Growth Factor receptor, Downward et al, 1984; *fms* and Colony Stimulating Factor receptor, Sacca et al, 1986). The identity of the other proto-oncogene products we can, as yet, only speculate about. The *ras*-gene family possess GTPase activity and may be related to the G-proteins (for reveiw see Balmain,1985). Some are found in the nucleus, (e.g. *myc*, *myb*, *ski*), while there are numerous examples of proto-oncogene products having protein kinase activity (for review see Bishop and Varmus, 1985). The proto-oncogenes appear to code for proteins which are essential for the sequential steps

required for the correct growth response of a cell. Presumably, in such a role, they act synergistically to bring about the normal growth of a cell.

The consensus of opinion, gleaned from all aspects of oncogenic research, points to cancer being a multistage process. However, when proto-oncogenes were first identified, it was considered possible that one genetic change could bring about the transformed state by the disruption of a single proto-oncogene. As we have learned more about these genes it is now thought that more than one activated proto-oncogene is required for the initiation and maintainance of the neoplastic state (Land, Parada and Weinberg, 1983; Adams et al, 1985; Oshimura, Gilmer and Barrett, 1985). Indeed, studies using the NIH-3T3 cell line have shown that fibroblast immortality is a prerequisite for transformation by the EJ c-Ha-ras oncogene when regulated by its own promoter domain (Newbold and Overell, 1983). However, in some cases the activation of one proto-oncogene may be enough for at least the initial stages (Spandidos and Wilkie, 1984b; Balmain, 1985).

At the basis of this concept is the disruption of the normal cellular role of the proto-oncogene. There is now evidence for both the alteration of coding sequences (for reveiw see Balmain, 1985; Bargmann, Hung and Weinberg, 1986) and the aberrant regulation of expression (for reveiw see Leder et al, 1983; Stewart, Pattengale and Leder, 1984; Adams et al, 1985). Such changes could be brought about by several mechanisms: amplification, deletion, translocation, point mutation, methylation. How the functions of these proto-oncogenes are subverted is of central importance in the quest to understand cancer. The work described in this thesis will hopefully help to shed light on how the expression of one of these proto-oncogenes, c-myc, is regulated, thus helping in understanding how regulation of this proto-oncogene is corrupted in the neoplastic cell.

1.2. Avian Myelocytomatosis Virus

On the basis of their oncogenicity avian retroviruses have been assigned to either of two classes; those that are weakly oncogenic and those that are highly oncogenic. The avian myelocytomatosis virus (AMV) in which the myc gene was first identified, falls into the latter category. These highly oncogenic retroviruses contain a transforming gene in their genome and cause disease after a short latency period of days to weeks. They are themselves split into the replication-competent sarcoma viruses and the replication-defective leukaemia viruses (DLV), which require a helper virus (a non-defective leukosis virus) for their replication (for review see Graf and Beug, 1978).

The oncogenicity of these viruses was found to be due to the presence of a transforming gene, a viral oncogene (v-onc), in their respective genomes. All the acutely transforming retroviruses have at least one, sometimes two viral oncogenes in their respective genomes (for review see Bishop and Vamus, 1985). These viral oncogenes have turned out to be the transduced forms, sometimes truncated, of normal cellular genes, the proto-oncogenes. Work done by Roussel et al (1979) showed that the DLV group represents recombinants between a vector related to a chicken endogenous virus and one of three cellular genes: erb, myb and in the case of AMV, myc. Other proto-oncogenes have since been identified through their transduction into avian retroviruses (Bishop and Varmus, 1983).

The AMV-type viruses (MC29, CMII, OK10 and MH2) belong to the DLV class of avian retrovirus. They transform macrophage-like cells in vitro and some strains induce myelocytomatosis in vivo. In addition, MC29 can transform fibroblasts and epithelial cells in vitro and induce carcinomas and occasionally sarcomas in vivo (for review see Hayward, 1983). The activation of c-myc by retroviral transduction is only one of the possible activation

mechanisms by which activation can occur (see section 1.5.3.)

1.3. v-myc

The myc gene was first defined as the unique transforming sequence of the MC29 family of retroviruses by Rousell et al, (1979). Since then, the topography of the v-myc gene has been found to differ from one AMV isolate to another. In the prototype, MC29, it is fused to a portion of the viral gag gene (Alitalo et al, 1983; Reddy et al, 1983) which encodes the internal structural protein of the virus. The hybrid gene gives rise to a polyprotein with a molecular weight of 110,000 (p110gag-myc; Bunte et al, 1984). While the closely related virus, CMII, encodes a similar protein (p90gag-myc; Hayman, Kitchner and Graf, 1979), the OK10 and MH-2 strains contain v-myc as an independent genetic unit that is expressed by means of a spliced subgenomic mRNA (Chiswell, Ramswey and Hayman, 1981; Pachl, Beigalke and Linial, 1982).

Recent hybridisation studies with the MH2 virus have detected a second v-onc (v-mil/mht) which again has a cellular counterpart (Coll et al, 1983; Kan et al, 1983). This gene is transformation specific, but as yet no direct function has been designed to it (Zhou et al, 1985). The MH2 virus is not the only avian retrovirus which harbours two viral oncogenes (Symonds et al, 1984; Watson et al, 1985).

The nucleotide sequence of v-myc, from the MC29 virus, was determined simultaneously by two independent research groups (Alitalo et al, 1983; Reddy et al, 1983) and revealed a single open reading frame that could encode a protein of 875 amino acids, with the v-myc sequences coding for 425 of them. The v-myc gene itself is capable of encoding an hydrophilic polypeptide which in its highly conserved (between viral isolates) carboxy-terminal region is rich in

basic residues, which may account for its DNA binding properties (Alitalo et al, 1983). More recent work has shown that the v-myc gene products from MC29- and MH2- type viruses have DNA binding properties, and that the MH2 product has a further RNA-binding property (Alitalo et al, 1983). In addition, the majority of the MC29-type virus gene product is found in the nucleus (Alitalo et al, 1983) while the majority of MH2 gene product is cytoplasmic (Bunte, Greiser-Wilke and Moelling, 1983; Bunte et al, 1984). The significance of this is not fully understood, as each virus causes only a slightly different spectrum of diseases (Roussel et al, 1979). The reported molecular weight of the v-myc gene product varies between 47000 and 58000 daltons (Alitalo et al, 1983; Bunte et al, 1984) possibly due to post-translational processing, degradation or the inherent lack of resolution in the assay techniques used (analogous migration variations in SDS-polyacrylamide gels is seen for other proteins, e.g. adenovirus E1a: Perrson et al, 1984).

It had been thought that the myc gene had only been transduced by avian retroviruses, until recent work from two separate research groups identified v-myc as the transforming gene of the acutely transforming variants of feline leukaemia virus (Neil et al, 1984; Levy, Gardner and Casey, 1984). It is tempting to speculate that as the transforming genes of more and more retroviral isolates are analysed, the transduction of the myc oncogene may be found to be a common occurrence. As a corollary of this, there may only be a subset of the cellular oncogenes which retroviruses have picked up. As this appears to be the case (new isolates are revealing previously identified v-onc), in order to identify all the cellular oncogenes we will have to utilise different avenues of research (see section 1.1.).

1.4. c-myc

Since being identified as the transforming gene of AMV

the cellular homologue, c-myc, has been detected in the normal DNA of species as diverse as chicken and man, with myc hybridising sequences being found in Drosophila (Shilo and Weinberg, 1981). Therefore this gene appears to have been conserved through evolution, with specific internal domains (e.g. the carboxy terminus) showing strong conservation, and as a correlation presumably has an important if not vital function in the normal cell. It is becoming apparent that, similar to the ras oncogene (Hall et al, 1983), there is a myc-like family of proto-oncogenes (Nau et al, 1985; Kohl et al, 1986). Gene mapping studies have assigned the three known human myc-like proto-oncogenes to separate chromosomes; c-myc to chromosome 8 (Dalla-Favera et al, 1982a; Neel et al, 1982; Taub et al, 1982), N-myc to chromosome 2 (Schwab et al, 1984) and L-myc to chromosome 1 (Nau et al, 1985). Further investigation may reveal the presence of additional members of this family.

1.4.1. Gene Structure

Heteroduplex analysis of the c-myc gene has shown that the v-myc hybridising regions are separated by an intron (Dalla-Favera et al 1982b). Since then, by comparative analysis of the RNA transcript and the cellular locus from the Raji Burkitt Lymphoma cell line a further 5'exon has been identified (Hamlyn and Rabbits, 1983), indicating the v-myc gene to be a truncated form of its cellular counterpart c-myc. Subsequently, this 5' exon has been identified in the murine c-myc (Bernard et al, 1983), chicken c-myc (Nottenburg and Varmus, 1986) and feline c-myc (Stewart et al, 1986), as well as in the human N-myc (Kohl et al, 1986; see figure 1.).

The sequences of the c-myc gene from man, mouse, cat and chicken (and N-myc) have been determined (Bernard et al, 1983; Battey et al, 1983; Stewart et al, 1986; Nottenburg and Varmus, 1986; Kohl et al, 1986). Although among the members of the family the coding sequences are reasonably

conserved, there are specific domains dispersed throughout the coding region which show strong homology (Kohl et al,1986; Stewart et al,1986). The best conservation at the amino acid level (70% for v- and c-myc, 50% for c- and N-myc) is seen in exon 3, in the form of small stretches of amino acids (Kohl et al,1986). It is within this region that the DNA-binding property of c-myc is thought to reside (Persson and Leder, 1984). Upon comparing this region to other known protein sequences, the only one with significant homology was c-fos, however the region of homology is relatively small (Kohl et al, 1986). A stuctural homology with the adenovirus E1a gene product has also been suggested (Ralston and Bishop, 1983; see section 1.6.4.).

Transcription initiation for the human c-myc can be from either of the two promoter sites, identified by the presence of the characteristic TATA-consensus sequence and S1 nuclease studies of mRNA transcripts (Watt et al, 1983; Hamlyn and Rabbits, 1983). In normal lymphoid cells, the relative ratios of the two 5' ends of stable mRNA imply that the downstream promoter is preferentially utilised (Taub et al, 1984 and see section 1.5.3.). A similar situation has been reported for the murine c-myc (Bernard et al, 1983). Recent evidence indicates the presence of an upstream start site (P0), approximatly 600bp 5' of P1 in the human c-myc (Bentley and Groudine, 1986), but its relevance to c-myc expression is unknown. The level of stable transcripts initiating from this site is very low, being less than that of P1. The transcription start sites for N-myc and chicken c-myc differ, in that there is no obvious main site but a heterogeneous cluster of starts, possibly reflecting a lack of a good TATA consensus in each case (Kohl et al,1986; Nottenburg and Varmus, 1986). All normal myc messages initiate translation at the 5' end of exon 2 (16bp downstream from the splice acceptor site of exon 1 for human c-myc :Saito et al, 1983). Thereby exon 1 is probably an untranslated leader sequence (containing several termination codons) in all the sequenced myc genes.

The untranslated region (including the dual promoter domain) has been proposed to have some functional value as all the myc genes sequenced possess such a domain (see section 1.5.3.). The human c-myc exon 1 encodes an RNA sequence with the potential to form four hairpin loops (Battey et al, 1983). Even though there is only slight homology between the exon 1 sequences from the various myc genes and between human c- and N-myc, the calculated secondary structure including potential loops shows some similarities (Kohl et al,1986; Nottenburg and Varmus, 1986; Stewart et al, 1986). This untranslated leader, even though its sequence is not conserved,, may therefore perform a similar funtion for the various members of the myc gene family. There is one report of an open reading frame, capable of encoding an 20 K protein, within exon 1 (Gazin et al, 1984). However, as no other sequence data for other cloned c-myc genes identified such a region, its existance is unlikely.

Presumably the 5' flanking sequences of the c-myc gene contain the important elements which regulate at least some aspects of its transcription. By looking at the DNase I sensitivity of chromatin, which is thought to indicate the presence of transcription regulatory elements (for review see Weintraub and Groudine, 1976), five hypersensitive sites have been identified in the human (Siebenlist et al, 1984) and murine c-myc genes (Fahrlander, Piechaczyk and Marcu, 1985). Two closely associated sites are localised around the dual promoter region, where there is very good interspecies sequence conservation including a repeated consensus (Corcoran, Cory and Adams, 1985; Stewart et al, 1986). Another two sites are associated with two similar nuclear protein consensus elements (i.e. NF-1: see section 1.9.4.), one of which is localised near the P0 upstream transcription initiation site. The fifth site is within a sequence strongly conserved between mouse and human about 2.0Kb upstream from exon 1, which contains two enhancer core consensus elements (see section 1.9.4.). Only sites I and

II-2 are strongly hypersensitive, with site III-2 being intermediate in strenth (see figure 2.). However, these are only the possible locations of regulatory elements as no functional analysis has been done. The work presented in this thesis concentrates on this topic.

1.4.2. Gene Product

The human c-myc gene codes for a 62K protein (Evan and Hancock, 1985), with several possible precursor or processed molecules (Ramsey, Evan and Bishop, 1984; Persson et al, 1986). Apparently this protein has a very short half life, about 30 minutes (Ramsey, Evan and Bishop, 1984; Rabbits et al, 1985), although recent data may indicate a product which is more stable, being generated post-translationally from the short lived precursors (Persson et al, 1986). The majority of p62 is found in the nucleus (Persson and Leder, 1984) but its exact location is still in question. Original studies indicated p62 to be associated with the nuclear matrix (Winqvist, Saksela and Alitalo, 1984), however, more recent analysis indicates that this association may be an artifact due to the isolation procedure, and therefore does not indicate what happens in vivo (Evan and Hancock, 1985).

As to what function the c-myc protein plays in the nucleus we have as yet very few clues. It does exhibit a high (possibly non-specific) affinity for double-stranded DNA (Perrson and Leder, 1984; Watt, Shatzman and Rosenberg, 1985). It is a generally hoped concept that the c-myc protein may play some role in the governing of gene expression (transcriptionally or post-transcriptionally). There are several examples of DNA-binding regulatory proteins which apparently bind non-specifically to DNA with high affinity, e.g. adenovirus Ela protein (for review see Kingston, Baldwin and Sharp, 1985), chick oviduct progesterone receptor (Hughes et al, 1981). Perhaps the c-myc protein, if it does indeed govern gene expression, may play a central role in the nuclear response to growth

signals (see sections 1.5.3., 1.6.3. and 1.6.4.). The nucleus is also the location of other oncogene products, e.g. c-myb which also has a DNA-binding affinity (Moelling et al, 1985; Klempnauer and Sippel, 1986). Whether these oncogene products interact with each other in some way, or have a similar function, is not known. Hopefully future analysis of the c-myc gene product, and the other nuclear oncogene products, will answer these questions. These answers would not only help in our understanding of the roles oncogenes play in the transformed cell, but also greatly expand our knowledge of gene regulation.

1.5. Cellular Role of myc

Since both the N- and c-myc (as yet very little is known about L-myc) have a similar gene layout, and show significant homology between specific domains of their protein products, especially at the carboxy terminus (Michitsch and Melera, 1985; Kohl et al, 1986), they may have similar functions in the cell. Why should the cell have several copies of a similar gene? Expression of the three genes, despite these apparent structural similarities, appears to be dramatically different with respect to tissue and developmental stage specificity, as well as tumour specificity (Zimmerman et al, 1985; Kohl et al, 1986). Therefore each gene may play the same (or similar) role in the different periods of the cell's life span (both N- and L-myc being expressed earlier during differention than c-myc: Zimmerman et al, 1985), or within the different cell types present in eukaryotic organisms (c-myc showing a more widespread expression range than either N- or L-myc: Zimmerman et al, 1985). It is conceivable that the different members of the human myc gene family may cross-regulate each other (see section 1.6.3.).

By analysing the expression patterns of the endogenous c-myc, and of c-myc exogenously introduced into various recipient cells, speculation as to its functional role in

the normal and neoplastic cell can be made.

1.5.1. Involvement in Proliferation

Among the various intracellular events which occur during the transition from the quiescent to the proliferative state (Chambard and Pouyssegur, 1986), c-myc expression is now known to be altered (see figure 3.). When quiescent cells (fibroblasts or lymphocytes) are stimulated by their respective mitogens (PDGF or Conconavalin A) to enter the cell cycle, c-myc expression is transiently, but dramatically elevated (Kelly et al, 1983). These studies have been expanded to show that c-fos expression is also transiently increased, but preceeds c-myc (Muller et al, 1984), and that both these oncogenes are induced by other growth factors (e.g. IL-2 and IL-3 on lymphocytes: Reed et al 1985; Conscience, Verrier and Martin, 1986). Elevation of c-myc expression (both the mRNA and protein levels are stimulated: Persson et al, 1984; Persson, Gray and Godeau, 1985; Rabbits et al, 1985) although an essential part of the induction of competence, appears to require additional synergistically acting growth mediator induced events for progression through the growth cycle (Armelin et al, 1984; Smeland et al, 1985; Bravo et al, 1985; Kaczmarek, Oren and Baserga, 1986) However, a v-myc containing retrovirus can (without structurally altering the endogenous c-myc locus), by itself abrogate the requirement for exogenously added growth factors in stimulating cell growth (Armelin et al, 1984; Rapp et al, 1985). Taken together, the above results indicate that c-myc may be an important component (for at least a subset of growth factors) in the cascade of events leading to cellular proliferation.

There is growing evidence that this induction of c-myc by growth factors (e.g. PDGF, one chain of which is encoded for by c-sis: Waterfeild et al, 1983) proceeds through the activation of protein kinase C (Reed et al, 1985; Bravo et al, 1985; Rodriguez-Pena and Rozengurt, 1985; Falletto,

Arrow and Macara, 1985; Coughlin et al, 1985). Numerous proto-oncogenes encode a product with protein kinase activity (for review see Bishop and Varmus, 1985), while the ras-gene family possess GTPase activity and may be related to the G proteins (for review see Balmain, 1985). In the light of these, and other findings (for review see Marshall, 1985; Goustin et al, 1985; Kelly et al, 1983), it is reasonable to propose a regulatory linkage between several different oncogene products. As our knowledge increases more and more of the known oncogenes will probably be found to be associated with each other in specific growth pathways (which are subverted during neoplastic growth).

Persistance of the competence state within the cell appears to be independent of c-myc expression (Bravo, Burchardt and Muller, 1985) even though it is apparently essential for its induction. Both messenger RNA and protein levels are constant, but low, throughout the cell cycle in exponentially growing cells (Rabbits et al, 1985; Persson, Gray and Godeau, 1985). Thus c-myc may be involved in the induction of a quiescent cell to enter the cell cycle, but not for the cell's transverse of the cycle (epidermal growth factor has been implicated in progression of the cell-cycle: Armelin et al, 1984).

Even though many studies have been done, how the expression of c-myc is regulated is still not understood. In the quiescent cell c-myc transcripts are undetectable (Campisi et al, 1984; Rabbits et al, 1985), but the level of transcription appears to be the same as that of a cycling cell (Blanchard et al, 1985; Thompson et al, 1985). This implies that the myc transcript in quiesent cells is extremely unstable. However, due to different cell lines being used (HL-60 and Chinese hampster lung fibroblasts respectively), and different methods utilised to arrest cell growth (contact inhibition and serum deprivation respectively), this direct correlation is not conclusive. In addition, the probe used for the in vitro run-on

transcription studies, since it relies upon cross-species hybridisation (Blanchard et al, 1985), does not preclude the presence of a transcription elongation block. Such a phenomenum does play some role in the decrease in c-myc message levels observed when HL-60 cells are induced to differentiate (Bentley and Groudine, 1986). Indeed, in a recent report (unfortunatly a different cell line again and the method of growth arrest is not reported), serum stimulation was shown to increase the transcription rate of c-myc (Endo and Nadal-Ginard,1986). Unfortunately therefore, because different systems have been utilised in these studies no clear picture emerges.

All the serum stimulation studies have used protein synthesis inhibitors, which when added in conjunction with serum, allow superinduction of c-myc message (Kelly et al, 1983; Endo and Nadal-Ginard, 1986). Once again conflicting results were obtained depending on the specific cell line and inhibitor employed, although the assay system was the same (nuclear in vitro run-on transcription assays) in the respective studies. Anisomysin treatment of quiescent NIH-3T3 cells prior to serum stimulation potentiated c-myc (c-fos and actin) expression at the transcriptional level (Greenberg, Hermanowski and Ziff, 1986). However the opposite, a decrease in transcription, was seen in the pheochromocytoma cell line PC12 (Greenberg, Hermanowski and Ziff, 1986). In similar studies, anisomysin or cycloheximide treatment resulted both in a marked increase in c-myc message stability and a slight decrease in transcription initiation upon serum stimulation of quiescent myoblasts (Endo and Nadal-Ginard, 1986). In another study, where an actinomysin chase was used, a marked increase in c-myc message stability was also observed for HeLa and HL-60 cells, while no change in stability was seen in Daudi Burkitt's lymphoma cells (Dani et al, 1984). Again no clear picture emerges from these studies.

Thus the actual mechanisms (transcriptional or post-

transcriptional) utilised by different cell lines in regulating c-myc expression within the various phases of the cell cycle (including G0) are still debatable. The main arguement (although not the only one) for an increase in the stability of the normally very unstable c-myc message (T1/2= 10-20 minutes: Dani et al, 1984; S.Graham personal communication) is based on data gleaned using protein synthesis inhibitors, and implies that a protein factor is selectively degrading the c-myc message. However this system revolves around a non-physiological state, in that the protein synthesis inhibitors used do not normally exist in the cell. Apart from this, the use of protein synthesis inhibitors may indeed give rise to artifacts, e.g. the resultant stabilisation seen after cycloheximide treatment, as this inhibitor functions to prevent termination, could be due to the accumulation of RNA polymerase II molecules on the transcript. The superinduction of c-myc expression seen upon addition of protein synthesis inhibitors may identify the presence of a protein factor (e.g. a double-strand-specific RNase or a transciption repressor) which functions to selectively degrade the c-myc message. Published data would suggest that this protein is present in normal and transformed cells (Dani et al, 1984; Endo and Nadal-Ginard, 1986). Thus, regulation at the post-transcriptional level may occur during serum stimulation, but there is no evidence which completely rules out a role for control at the transcriptional level.

To further elucidate the involvement of the c-myc gene in proliferation, studies on interferon (IFN) -induced growth arrest (as opposed to serum inducton of growth) of Daudi Burkitt's lymphoma cells have been reported. Upon IFN-β treatment cells are induced into a quiescent state with a corresponding decrease in the level of c-myc mRNA (Jonak and Knight, 1984). However not all BL cell lines respond equally to IFN, and indeed cell lines of different origins (e.g. HL-60 and Friend cells), although showing a reduced cell growth, do not arrest in G0 after IFN treatment

(Einat, Resnitzky and Kimchi, 1985). Thus the exact relevance of using the Daudi cell line as a model may be questioned. Nuclear run-on *in vitro* transcripton assays were undertaken in these studies, and the decrease in mRNA abundance attributed to a decrease in c-*myc* message stability (Knight et al, 1985; Dani et al, 1985). IFN-α has a similar affect on Daudi cells in that it induces growth arrest with a corresponding decrease in c-*myc* expression, however a decrease in c-*myc* transcription was seen (Einat, Resnitzky and Kimchi, 1985). This may mean that the forms of interferon have different affects on the control of c-*myc* expression.

The overall picture of the role of c-*myc* in cellular proliferation is still not clear. The fact that serum stimulation of quiescent cells induces a transient elevation of c-*myc* expression (with a presumed important function) is now generally accepted, but whether this is due to an increase in message stability or an increase in transcription is still under debate. Perhaps both mechanisms are differentially utilised by different cells or even by the same cell type at different stages of development or differentiation. Of note here, is that c-*fos* has a 5' *cis*-acting regulatory element which responds to serum induction in addition to a 3' element possibly involved in message stability (Treisman, 1985 and 1986).

1.5.2. Involvement in Differentiation

The occurrence of c-*myc* expression appears to be dependent on the differentiation state of the cell, being turned off in mature cells (Reitsma et al, 1983; Gonda and Metcalf, 1984; Lachman and Skoultchi, 1984; Dony, Kessel and Gruss, 1985; Brelvi and Studzinski, 1986). Several studies have shown that constitutive c-*myc* expression is incompatable (preventing) with the induction of terminal differentiation in HL-60 and MEL cells (Falcone, Tato and Alema, 1985; Symonds et al, 1986; Coppola and Cole, 1986;

Dmitrovsky et al, 1986; Prodownik and Kukowsky, 1986).

Once again whether this shut-off is transcriptionally or post-transcriptionally regulated is still in question. When HL-60 cells are treated with retinoic acid differentiation down the granulocytic pathway is induced, with an associated decrease in c-myc mRNA abundance. Nuclear in vitro run-on transcription assays have shown that a transcription elongation block (between exon 1 and 2) can account for most of this shut-off (at 6 days post induction: Bentley and Groudine, 1986). Thus neither regulation of the level of transcription induction nor post-transcriptional mechanisms need to be evoked in this system. In retinoic acid induced differentiation of F9 embryonic carcinoma cells c-myc is actively transcribed, even though there is a decrease in mRNA levels (Dony, Kessel and Gruss,1985), while the decrease in c-myc expression upon Friend cell differentiation is due to a decrease in the level of transcription (Grosso and Pitot, 1985). Thus the level at which regulation of c-myc expression occurs during differentiation may depend on the cell type.

Two indirectly relevant results come from different experimental approaches. On comparing the methylation patterns of the two c-myc alleles in a murine plasmacytoma cell line, the silent allele (that responsive to normal differentiation-specific regulation) is hypermethylated compared to the translocated, active allele (Dunnick et al, 1985). Secondly, upon HL-60 differentiation the DNase 1 hypersensitive sites 5' to the c-myc gene alter (see 1.4.1.). Both these results indicate that upon differentiation the c-myc gene is physically altered, a phenomenon which is commonly associated with a change in regulation of transcription. Obviously these results do not exclude post-transcriptional mechanisms from being involved.

However c-myc expression is controlled during terminal differentiation, c-myc shut-off may be a prerequisite for

normal differentiation to occur. Since c-myc has some role, in all likelihood an essential one, in inducing the growth factor stimulated competence state, the c-myc gene product is probably an essential component of the growth response of the cell. The recent studies on transgenic mice harbouring c-myc derived constructs, gives further credence to the c-myc gene having a role in cellular proliferation.

1.5.3. Involvement in Neoplasia

The c-myc gene has been implicated in the genesis of several forms of neoplasm. Classically, v-myc was shown to be the transforming gene present in specific avian retroviruses, while the cellular homologue (via promoter insertion) was implicated in the neoplasia by studies on certain slowly transforming retroviruses. More recently the activation of c-myc has been implicated in several human neoplasias by the analysis of tumour cell lines and primary biopsy samples.

(a). Retroviral Transduction

It was through the analysis of the MC-29 virus that the myc gene was first identified (see section 1.2.). With the recent identification of c-myc being transduced by a feline leukaemia virus (Neil et al, 1984; Levy, Gardner and Casey, 1984), the question of whether c-myc may be transduced by other viruses is raised. Perhaps other virus-associated leukaemias (which await analysis) may be due to a virally transduced c-myc.

Transduction presumably activates c-myc by placing it under the strong transcriptional control of the proviral enhancer within the LTR. The v-myc gene is a truncated version of c-myc, but still retains all the coding sequence. However important regulatory sequences (within exon 1) may have been lost during the truncation. Therefore it is conceivable that the loss of these sequences confers the

transforming ability upon the genes (see section 1.6.2.). Although the consequence of transduction is an over-expression of c-myc, it is also possible that the activating step is the removal of normal regulation of expression (i.e. deregulation). Although there are cases where a common mutation location has been observed (Papas and Lautenberger, 1985), not all v-myc isolates harbour this mutation (Stewart et al, 1986). Thus, an abnormal mutated gene product is not an absolute requirement for transforming potential (unless different mutations have the same effect: see section 1.5.3.c), but may conceivably have some role in directing the spectrum of disease.

(b).Promoter Insertional Activation

Although first demonstrated for the c-myc gene (Hayward, Neel and Astrin, 1981; Fung, Crittenden and Kung, 1982; Corcoran et al, 1984) proviral insertion adjacent to, and thereby activating, cellular oncogenes has now been documented for several viral-oncogene systems (Peters, Lee and Dickson, 1982; Fung et al, 1983; Neil et al, 1984; Nusse et al, 1984; Swift et al, 1985; Molders et al, 1985). After integration of the provirus the original DNase 1 hypersensitive sites of the endogenous oncogene are no longer detectable, with a new strong site being present in the proviral LTR (Schubach and Groudine, 1984). This has given credence to the promoter insertion theory for the activation of cellular oncogenes (for review see Temin, 1982), where expression of the cellular gene is driven by the strong proviral LTR.

Most proviral insertions are either 5' or 3' to the coding sequences in question. Therefore it is probably constitutive, deregulated expression of the cellular oncogene, rather than an altered coding potential which results in the transformed phenotype. As an aside to this, these situations probably give us an insight into how a retrovirus integrates into cellular DNA. In ALV-induced

lymphoid leukaemia (Hayward, Neel and Astrin, 1981) in the majority of cases the provirus integrates in the same orientation as the cellular oncogene, while in other systems proviral integration is in the opposite orientation (Corcoran et al, 1984). Thus the integration specificity of each type of provirus differs and through the study of these systems we may gain an insight into how.

In addition to provirus insertional activation, a movable repetitive DNA element ("retroposon") has been shown to activate c-myc (Katzir et al, 1985). Again the insertion was upstream of c-myc and caused deregulated expression. It is possible that the "retroposon" element directly increases c-myc expression but the disruption of the 5' control domain of c-myc is also a possibility.

(c). Chromosomal Aberrations Affecting c-myc

Slamon et al (1984) have shown that c-myc is expressed in several tumour types. Since then expression of c-myc has been associated with an ever increasing range of tumour types (Peschle et al, 1984; Erisman et al, 1985; Sikora et al, 1985). However, upon transfection of B-cell lymphoma DNA, the transformation-specific gene which was identified (B-lym) was not c-myc, although c-myc activation was inferred from analysis of the translocation break point (Goubin et al, 1983; Diamond et al, 1983; Devine, 1986). At face value this would give evidence for cancer being a multistep phenomenon, with the co-ordinate activation of more than one cellular oncogene. However, there are at present questions regarding the cloning strategy used to isolate the B-lym gene and its subsequent transforming activity, which are as yet unresolved.

Most information on c-myc involvement in neoplasia has come from the analysis of c-myc rearrangement in B-cell tumours (see below and Neil et al, 1984; Corcoran et al, 1984).

(i).Amplification

Amplification of certain oncogenes is common in some tumours (for review see Alitalo, 1985), and the increased dosage of an oncogene as a result of such a rearrangement may contribute to the genesis and/or progression of at least some forms of cancer. All three human myc genes have been found amplified in specific tumour types, both in tumour derived cell lines and clinical biopsy samples (Dalla-Favera et al, 1982a; Schwab, 1985; Nepveu et al, 1985; Nau et al 1985; Alitalo, 1985; Yokota et al, 1986).

How amplification is involved in the transformation process (N-myc amplification number correlates with malignant progression: Schwab et al, 1984), or if it is only a consequence of transformation (no myc gene has been found amplified in a normal cell) is not known. Indeed, whether all copies within the amplified domain are active is not known. However, correlating message levels with copy number (Graham, Tindle and Birnie, 1985), and the recent DNase 1 sensitivity studies on HL-60 cells (Dyson et al, 1985) and an A-MuLV-transformed fibroblast cell line (Nepveau et al, 1985), imply that at least most copies are active.

(ii).Chromosome translocation

Most of our insight into the ways in which chromosomal translocations can involve cellular oncogenes has come from the study of two types of lymphoid tumour: human Burkitts lymphoma (BL) and mouse plasmacytoma (MPC). Both of these tumours are characterised by a chromosomal translocation, in each case involving an immunoglobulin locus and the c-myc oncogene (see figure 4 and 5). Even though the actual breakpoint in the c-myc locus varies quite considerably in both tumour types, the end result of most translocations is constitutive (i.e. deregulated) c-myc expression, with an obvious over-expression occuring only infrequently (for reviews see Klein, 1983; Leder et al, 1983; Perry, 1983;

Mushinski et al, 1983; Rabbits et al, 1984; Sun, Showe and Croce, 1986).

Whenever the dual promoter region (Battey et al, 1983) is preserved during translocation, the relative usage of the two promoters changes, i.e. the more 5' promoter (P1) is preferentially activated (Leder et al, 1983; Taub et al, 1984; Fahrlander et al, 1985; Denny et al 1985; Yang et al, 1985). This is similar to the situation where c-myc is activated by proviral insertion in certain Feline leukaemias (D.Forrest, personal communication). However when the breakpoints occur within the exon/intron 1 region then truncated mRNA transcripts arising from cryptic promoters within intron 1, are generated (Saito et al, 1983; Keath, Kelekar and Cole, 1984). These transcripts originate in a heterogenous, bipolar fashion (Prehn, Mercola and Calame, 1984; Calabi and Neuberger, 1985) presumably reflecting a lack of a suitable TATA-consensus element. These truncated transcripts code for the same gene product as the full length transcripts generated from the normal dual promoter regions (Piechaczyk et al, 1985).

Since a stable stem-loop structure could form between a region in exon 1 which is highly complementary to a region in exon 2 (Saito et al, 1983), it has been proposed that lack of this structure in the truncated mRNA could result in their increased stability. Saito et al (1983) proposed that removal of the potential to form this stem-loop structure may, in addition, increase the translational efficiency of the mRNA. A similar secondary structure is potentially possible for all the myc-like genes sequenced to-date (see 1.4.1.).

Upon comparison of the relative stabilities of the various c-myc transcripts from the different MPC cell lines (where the breakpoints are either 5' or within exon 1) the truncated transcripts were found to have a substantially greater half life (Piechaczyk et al, 1985). To further

investigate the role of exon 1, truncated c-myc genes were constructed in vitro. Depending on the actual construct used either no variation in translation efficiency but an increase in message stablility (Butnick et al, 1985), or an increase in translation efficiency (Darveau et al, 1985) was observed for the truncated transcripts. These differing results are probably due to small differences in the construction strategies employed by each group. Since the constructs of Darveau et al (1985), included a 3' deletion this may account for the lack of variation in message stability seen in the other study. Taken together with the in vivo data, the untranslated 5' exon would appear to be involved to some extent in post-transcriptionally regulating the c-myc gene.

The relative stabilities of these different transcripts may be partially a red herring in the understanding of c-myc expression deregulation. It is probably the fact that c-myc is expressed at all that is important, remembering that c-myc is not expressed in the mature differentiated cell (see section 1.5.2.). The level of c-myc expression in these tumours varies quite considerably, from that of a normal lymphoblastoid cell to being highly elevated (Taub et al, 1984; Keath, Kelekar and Cole, 1984; Sun, Showe and Croce, 1986). The level of the protein also varies between different BL cell lines (Ramsey, Evan and Bishop, 1984; Hann and Eisenman, 1984).

Although there are cases when mutations within the structural gene have been detected (Rabbits et al, 1984; Denny et al, 1985, Murphy et al, 1986) and noted within the some v-myc isolates (Papas and Lautenberger, 1985), but not in the FeLV-transduced gene (Stewart et al, 1986) there is no unifying mutation found in all activated myc situations (unlike the ras oncogene activation: for review see Balmain, 1985). Recently, Martin et al (1986) transformed cells with an avian retrovirus containing and expressing a non-mutated human c-myc gene. Upon recloning of this gene from a

transformed cell it was still found not to contain any mutations. In addition, augmented transcription of the normal c-myc is sufficient for cotransforming activity with c-Ha-ras (Lee et al, 1985), makes primary fibroblasts tumorigeneic (Pragnell, Spandidos and Wilkie, 1985) and can transform cells alone in a rat embryo fibroblast focus formation assay (Kelekar and Cole, 1986). However, a mutated c-myc gene exhibited a reduced transforming ability (Murphy et al, 1986). These results do not exclude coding sequence mutations from affecting the disease spectrum of the activated, gene but do imply that a coding sequence mutation is not a prerequisite for activation.

(d). Cellular transformation

Recent evidence from transfection studies has further implicated the c-myc gene in the neoplastic process. The transfection of c-myc, under the control of various promoters, induces the recipient cells into a state of continuous growth with an associated reduced growth factor requirement (Palmieri, Kahn and Graf, 1983; Mouganeau et al, 1984; Pragnell, Spandidos and Wilkie, 1985; Connan, Rassoulzadegan and Cuzin, 1985). Recent experiments concerning the multistage nature of transformation (Land, Parada and Weinberg, 1983; Kelekar and Cole, 1986) have shown that two active oncogenes can bring about transformation of a recipient primary cell. In these experiments c-myc was found to confer extended growth characteristics (immortalisation) while c-Ha-ras induced morphological transformation upon the cells. However since these studies, c-myc has been shown to be capable of morphologically transforming cells and c-Ha-ras has the ability to extend the growth capacity of a cell (Pragnell, Spandidos and Wilkie, 1985). Thus to say that one oncogene has a specific role in the initiation or progression of the transformed phenotype is probably being simplistic. The actual role played by individual oncogenes will probably depend on the situation in which it finds itself.

It is conceivable that in some circumstances a single activated oncogene will be enough to trigger full transformation. Indeed, the c-Ha-ras oncogene alone, driven by two enhancer elements is capable of transforming primary cells (Spandidos and Wilkie, 1984b; see section 1.6.5.).

1.6. How is c-myc deregulated ?

Therefore, neither an overt over-production of normal c-myc nor the mutated form of c-myc appear to be central to its activation, although both have been observed. What appears to be of prime importance is the deregulated, constitutive expression of c-myc. The level of transcription may be less important than the potential for transcription at an inappropriate time during the cell-cycle or during differentiation. Understanding how c-myc is deregulated is of crucial importance if we are to comprehend the role c-myc plays in the transformation process.

1.6.1. Involvement of a Repressor.

Once again it has been studies on BL and MPC cell lines which, have given us an insight into how the c-myc gene is deregulated. The translocated c-myc gene is expressed constitutively, while the normal allele on the intact chromosome 8 is transcriptionally silent (or at least no transcripts can be detected). This situation is found both in BL and MPC cells (Bernard et al, 1983; Nishikura et al, 1984; Taub et al, 1984) and interestingly, only the active allele is hypomethylated (Dunnick et al, 1985).

Studies on somatic cell hybrids have shown that c-myc expression depends on the differentiated state of the cell (Nishukura et al, 1985). When a hybrid between a BL cell and a human lymphoblastoid cell is made, which retains the lymphoblastoid (immature) phenotype, only the normal allele is expressed (Croce et al, 1985; Nishikura, et al, 1985). However, in a BL/MPC hybrid (where the human chromosomes are

differentially lost), only the the translocated human allele is expressed (Nishikura et al, 1983; Croce et al, 1985). In these hybrids, which have a more mature phenotype than the lymphoblastoid hybrids previously described, the normal non-translocated allele appears to be repressed. This apparent repression of a normal c-myc gene is also seen in MPC/human fibroblast hybrids (Nishikura et al, 1984), but whether this is a differentiation- or tissue- specific event is unknown.

Thus, c-myc being expressed in the BL/MPC cell is abnormal in that its expression should be shut off in these mature cells. Either translocation allows one allele to escape the normal regulation (to which the untranslocated allele remains subject) or, the translocation event positively activates that allele (by the relocation of positive transcription elements). Since, in the somatic cell hybrid studies, the translocated allele is not expressed in the phenotypically lymphoblastoid hybrids (Nishikura et al, 1985) it is unlikely that deregulation is under the control of a cis-acting positive element, unlesss by a differentiation-speciic element. Differentiation-specific elements (e.g. the immunoglobulin enhancer) may in some cases be involved in this deregualtion (Croce et al, 1985). However, at least in some cases, translocation allows escape from the normal control of expression by the removal of (some of) the normal c-myc transcriptional regulatory elements.

The silence of the non-translocated allele has led Leder et al (1983) to propose that c-myc could be regulated by a trans-acting repressor factor that might use as its target exon 1 or adjacent 5' flanking sequences. Such a trans-acting factor could be a transcriptional repressor but complex models (as some breakpoints are outwith the transcribed unit), invoking an element that accelerates degradation of the c-myc RNA, can also be imagined. The translocated c-myc gene would not be affected by this

repression because of loss or damage to the particular control domain through which the repressor interacts.

1.6.2. Possible Sites for a Negative Transcription Element

A destabilisation of the already very unstable c-myc message (Dani et al, 1984) could come about in several ways. There could be an important sequence element within the transcribed sequences which confers instability to the message, or as discussed before (see section 1.4.1.) the potential secondary structure of the transcripts could be of relevance. If a trans-acting transcriptional repressor is envisaged a cis-acting responsive element (negative regulatory element, NRE) has to be hypothesised. However the non-specific binding for adenovirus Ela protein (for review see Kingston, Baldwin and Sharp, 1985) gives a precedent for transcription factors not requiring specific recognition signals (see section 1.6.4.). The circumstantial evidence for the presence of transcriptional control elements for the human c-myc gene is outlined below.

Somatic alterations in the structure of the exon 1, which range from its complete removal to the occurrance of scattered mutations within its sequence, have been associated with deregulation of c-myc expression (Taub et al, 1984; Rabbits et al, 1984). However, this does not always occur. Better evidence for the localisation of a negative regulatory element comes from the several studies which identified the DNase 1 hypersensitive sites around the 5' region of the human and murine c-myc genes (Siebenlist et al, 1984; Fahrlander, Piechaczyk and Marcu, 1985; Dyson et al, 1985). Although 5 hypersensitive sites have been localised within the 5' flanking sequences and dual promoter region (see section 1.4.1.), results from analysis of the two allelic forms in BL and MPC cells (Siebenlist et al, 1984; Fahrlander, Piechaczyk and Marcu, 1985) and the changes in the hypersensitivity pattern seen when HL-60 cells are induced to differentiate (Dyson et al, 1985) has

put site 1 into the limelight as a candidate for the differentiation-specific negative regulatory element. This region does show good homology between the human and murine c-myc genes (Fahrlander, Piechaczyk and Marcu, 1985).

However, pooling the known data on translocation breakpoints shows a clustering around the -450bp to P1 region, infering that a negative regulatory element may be closer to the dual promoter region than implied by the DNAase 1 data (Wiman et al, 1984; Yang, et al 1985; Corcoran, Cory and Adams, 1985; Pelicci et al, 1986). Additionally, the upstream transcriptional start (PO) has been localised to near DNase 1 hypersensitive site II-2 (Bentley and Groudine, 1986).

Both the studies on differentiation-specific and cell cycle regulated c-myc expression indicate the possible involvement of a repressor factor. Whether the repressor involved in each of these situations is the same factor or not is unknown. If different factors were involved then two cis-acting recognition elements would require to be postulated, although they could overlap. However it is conceivable that the same factor could function in both a cell-cycle and differentiation specific manner, in which case only one cis-acting recognition element requires to be postulated. The work presented in this thesis identifies several putative transcriptional control elements within the c-myc 5' flanking sequences.

1.6.3. Postulated Repressor Factor

Leder et al, (1983), also postulated that the repression of c-myc expression may be mediated directly or indirectly by the cellular level of the c-myc protein. Although no direct experiments have been done to test this, results from transfection of v-myc containing viruses (Rapp et al, 1985) or c-myc genes linked to viral promoters (Keath, Caimi and Cole, 1984), infer that the c-myc gene product may be

involved in its own regulation because in each case the expression of the endogenous c-myc is diminished.

More suggestive data came from analysing the sequence of the Raji BL cell line (Rabbits et al, 1984), where both c-myc alleles are expressed. Within the coding sequence numerous mutations were found, thus perhaps the translocated gene product is so altered that it can not function normally. If its normal function was to repress the normal allele then these mutations would allow that allele to escape repression and thus be expressed, as is indeed the case. However, the situation is slightly more complex than this. Since both alleles in the Raji cell line are expressing a gene product (N.B. there is no protein data to confirm this, it is the current thinking),

the translocated gene product (due to the mutations present in its coding domains) can not function as normal, i.e. to repress its own transcription. However, the normal gene product could. Thus, because of mutations within the promoter/first exon region of the translocated allele, the normal gene product could only interact (repress) with the normal allele. As a consequence, expression from the two alleles would be biased towards the translocated allele. This is indeed what is observed (Rabbits et al, 1984), with c-myc mRNA in the Raji cell line comprising of 65% translocated and 35% normal transcripts (approximately).

Using protein inhibitors in the serum induction (Kelly et atl, 1983) and liver regeneration experiment (Makino, Hayashi and Sugimura, 1984), results consistent with a labile protein regulating c-myc expression were obtained. As the c-myc protein has a short half-life of 20-30 minutes (Ramsey, Evan and Bishop, 1984), is located in the nucleus (Persson and Leder, 1984), and has a DNA-binding affinity (Alitalo et al, 1983), its physical properties are not unlike those expected of a transcriptional regulator. However, recent studies where a constitutvely expressed exogenous c-myc gene prevents induction of MEL cell

differentiation, both the exogenous and the endogenous c-myc genes are expressed (Dmitrovsky et al, 1986; Prochownik and Kukoska, 1986; Coppola and Cole, 1986).

One major problem in analysing the results descir bed above (apart from the use of different cell lines) is the lack of protein (p62) data in every report. Although c-myc is being transcribed (mRNA is detected), whether efficient or correct translation is occurring has not yet been shown. Thus whether c-myc does indeed regulate its own expression (directly or indirectly) is still debatable. Since the location of several cis-acting transcriptional elements have been identified (see section 4.) direct experiments to test whether c-myc regulates itself can be easily devised.

Two other putative proteins have been suggested as the c-myc repressor; one of which has since been discarded. Gazin et al (1984) reported that exon 1 of c-myc had the potential to encode a 20K protein. However, no other sequence data showed an open reading frame within the c-myc first exon (Bernard et al, 1983; Stewart et al, 1986; Kohl et al, 1986; Nottenburg et al, 1986). Working on variant t(6;15) translocations in several murine plasmacytomas, a breakpoints cluster in a 4.0Kb region has been identified, at least 72Kb distant from the c-myc dual promoter domain (Cory et al, 1985). Since c-myc is expressed at an elevated level in these cells, this variant translocation apparently activates the c-myc gene. This could occur directly via a long range cis-function (perturbation of chromatin suprastructure) or via the putative pvt-1 (break point locus) gene product (Cory et al, 1985; Graham, Adams and Cory, 1985). As yet no such gene product (or specific mRNA) has been identified so this is still speculation, and therefore does not remove the possibility that the c-myc gene product may regulate itself somehow. It is also conceivable that the different members of the myc gene family may cross regulate each other.

1.6.4. Similarity Between c-myc and Adenovirus E1a Gene Products

The myc and adenovirus E1a gene products show some degree of homology (Ralston and Bishop,1983), although the exact level of significance has been questioned (McLachlan and Boswell, 1985). Perhaps this structural similarity reflects a common function. Both these genes can complement Ha-ras in the co-transfection assay (Land, Parada and Weinberg, 1982). The E1a gene products can transcriptionally activate the human hsp70 gene while this gene has been shown to be normally regulated by an endogenous cellular E1a-like activity (Kao et al, 1985). Although no direct comparison with c-myc can be made, the c-myc gene product does stimulate expression of the Drosophila hsp70 gene (Kingston, Baldwin and Sharp,1985).

On a similar theme, certain exogenously introduced enhancer elements lose their activity when F9-EC cells are induced to differentiate, and that enhancer activity is associated with an E1a-like activity, since undifferentiated cells can support adenovirus mutants deficient in E1a activity (Imperiale et al, 1984). As a correlation, the induction of the differentiated state in these cells is associated with a decrease in c-myc expression (Sassone-Corsi and Borrelli, 1986).

Seperate studies have shown E1a to regulate enhancer activity in both a positive and a negative manner (for review see Jones, 1986). Although these are highly circumstantial connections, it is tempting to speculate that the c-myc gene, like E1a, may code for a trans-acting factor. In addition, through searching for protein sequences which show homology to c-myc, the only significant homology (albeit, only to a small region of the third exon, where the DNA binding domain is thought to reside, but not within a region showing homology between c-myc and E1a) was with the c-fos gene product (Kohl et al, 1986). Recently, c-fos has

been shown to encode a trans-acting transcription factor (Setoyama et al, 1986).

Since microinjection of E1a containing plasmids releases cells from growth arrest (Stabel, Argos and Philipson, 1985), and c-myc expression is induced during the competence stage of cell growth (Kelly et al, 1983), this speculation can be extended to implicate the c-myc gene having a role in governing the expression of the specific genes involved in the cells response to growth stimulation.

One apparent difference between the two gene products comes from data on the induction of heat-shock genes. Unlike E1a which has been shown to exhibit a non-sequence specific binding activity (for review see Kingston, Baldwin and Sharp, 1985), c-myc requires certain (as yet unidentified sequences) to regulate hsp70 expression (Kingston, Baldwin and Sharp, 1984). Therefore, even if these gene products do have a similar activity they are obviously not identical.

1.6.5. Summary: c-myc

We do not know if expression of c-myc in tumours is a consequence of that tumour's maturation arrest, or if c-myc expression prevents the neoplastic cells progression into a differentiated state, thus causing maturation arrest. What we do know, from the recent elegant transgenic experiments with c-myc (Adams et al, 1985; Leder et al, 1986), is that a deregulated c-myc gene appears to act as a heritable, predisposing factor which favours the accelerated development of tissue-specific tumours. However, for the development of full neoplasia other additional factors are required.

Constitutive expression is the result of c-myc activation in a chemically transformed cell (Campisi et al, 1984). How this deregulation of expression comes about is of central importance if we are to understand how normal

cellular genes can when activated induce the transformed phenotype. There is evidence for both transcriptional and post-transcriptional regulatory mechanisms. The work presented in this thesis is centered around the analysis of the transcriptional regulation of c-myc. The second part of this introduction therefore outlines our present knowledge of transcription regulation with a brief discussion of gene regulation in general.

1.7. Chromatin Structure of Active Genes

The nuclear DNA of eukaryotes is highly organised at several structural levels, from the protein coding gene unit to the ordering of solenoids into discrete loops (for reviews see North, 1984; Zehnbauer and Vogelstein, 1985). Indeed gene expression appears to be controlled at many stages (Efstratiadis et al, 1980) with transcriptional regulation the most frequent level (for review see Darnell, 1982). It is this level of control on which the work described in this thesis is based, and as such I will specifically concentrate this part of the introduction on eukaryotic transcriptional control mechanisms, with only a brief outline of the other levels of expression control.

Eukaryotic chromatin has a dynamic, complex hierarchical structure, with active gene transcription taking place only on a small proportion of it at any one time. Active genes are apparently packaged in an altered nucleosome structure probably established and maintained by histone H1 and its variants together with HMG proteins (for review see Weintraub, 1985). This altered nucleosome structure is described as being less condensed or more open than that associated with inactive domains. Obviously there are clear changes in chromatin structure between open-active and inactive domains and these can be detected by their sensitivity to DNase digestion (Kerem et al, 1984; Jackson, 1986). Presumably this open state allows the association of the transcriptional factors including the RNA polymerase

with the DNA, an association prohibited by the condensed state.

These active chromatin regions appear to be compartmentalised to certain regions within the nucleus. Many lines of evidence indicate that active regions are associated with the nuclear substructure (for review ses Jackson, 1986), and that transcriptionally active genes are apparently specifically anchored to the nuclear matrix/scaffold (for review see Zehnbauer and Vogelstein, 1985). Whether specific sequences are involved in this attachment is still under debate (for review see Jackson, 1986). If this is indeed the case, and transcription occurs in a complex tightly associated to the nuclear matrix/scaffold, then gene units could be activated by selective attachment, perhaps through specific sequences. Recent evidence has also implied that these transcriptionally active regions are under local torsional strain (for reviews see Weintraub, 1985; Lucknick, 1986; Jackson, 1986). From the data amassed to date, it seems plausible that it is the interaction of specific proteins to specific sequences located near to the gene unit, which induces this change in the level of supercoiling (for review see Weintraub, 1985). Indeed, Gasser and Laemmli, (1986), have identified specific DNA fragments which are selectivly attached to the nuclear scaffold (SARs), and that these are located 5' and 3' to three Drosophila genes. Within these regions there are both enhancer core consensus sequences and sequences closely related to the cleavage consensus of topoisomerase II.

In summary, active genes seem to be located in less condensed DNA chromatin domains associated with the nuclear matrix/scaffold. This association is probably mediated through specific protein-DNA interactions. In addition active regions of DNA are probably under local torsional stress (Cocherill and Gerrard, 1986). Regulation of this association is more than likely involved in the

developmental- and differentiation- specific control of expression.

1.7.1. DNase Sensitivity

Alterations in DNA chromatin structure may be identified by the presence or absence of DNase 1 sensitivity (Larsen and Weintraub, 1982; for a review see Elgin, 1984) and as a correlation, active genes can be functionally defined by their increased sensitivity to these nucleases. In addition, there have been many reports concerning DNase 1 hypersensitive sites which are associated with specific gene units. Such sites are generally found within the 5'and 3' flanking regions of active genes (McGuinnis et al, 1983; Peterson, 1985; for a review see Weintraub and Groudine, 1986) and their presence precedes transcription (for reviews see Weintraub et al, 1982; Elgin, 1984). As such they appear to represent a potential for transcription (Fritton et al, 1984; for a review see Reeves, 1984). Such sites are correspondingly lacking in 'silent' genes (for a review see Reudelhuber, 1984).

This sensitivity is considered to reflect the positions of potential interactions between specific effector proteins and DNA. In this respect specific proteins have been shown to bind to the upstream sequences of several genes, thereby inducing their DNase sensitive sites (Weisbrod and Weintraub, 1979; Emerson and Falsenfeld, 1984; Wu, 1984).

DNA-protein interaction may not be the only way in which DNAase 1 hypersensitive sites can be generated. Actively transcribed DNA is thought to be in the B-configuration, but DNA sequences of an alternating purine-pyrimidine nature have the potential ability to form the Z-DNA configuration (for review see Neidle, 1983; Ellison et al, 1985). Such a configuration causes a perturbation in the supercoiling of DNA and sequences with this potential have been noted in some DNase 1 hypersensitive sites (Benoist et al, 1980). The

switch from the B- to Z-configuration may be brought about by sequence-specific Z-DNA binding proteins (for review see Rich, 1982). There is, however, relatively little actual evidence for the in vivo presence of the alternative forms of DNA. Like Z-DNA, any structure mediated by the energy of DNA supercoiling, e.g. the ability of inverted repeats to form a cruciform formation (Lilley, 1983), may be of relevance in the regulation of gene expression.

1.7.2. DNA Methylation

DNA methylation patterns appear to correlate with nuclease sensitivity (Reeves, 1984; Keshet, Lieman-Hurwitz and Leder, 1986), and correspondingly active regions of chromatin are selectively hypomethylated (Davis et al, 1986). Whether methylation is a cause or consequence of gene activation is still open to debate.

1.8. Post-Transcriptional Control of Gene Expression

Gene expression can be regulated post-transcriptionally by a variety of mechanisms: formation and processing of the primary transcript, transport and stability of the mRNA and translational efficiency (for review see Nevins, 1983). Regulation at the level of processing of the gene products, although of considerable importance, is outwith the scope of this thesis and as such will not be discussed further.

(i). Formation and Processing of the Primary Transcript

The 5' mRNA terminus (transcription initiation site) is comprised of a terminal 7-methylguanosine residue (added post-transcriptionally) linked to a triphosphorylated purine (m7GpppPu), and is termed the cap site. Translation is mediated through cap-binding proteins and usually initiates from the first AUG codon encountered downstream from the transcription initiation site (for review see Shatkin, 1985; Kosak, 1986). The precise sequence surrounding the AUG

appears to exert some control over the site of translation initiation, with the optimal sequence being 5'-ACCAUG-3' (Kosak, 1986). There is evidence that false translational starts may be involved in control of gene expression (for review see Hunt, 1985).

The 3' terminus of the mRNA, unlike the 5' terminus does not correspond to the transcription termination site and in eukaryotes a shift in emphasis from transcription termination towards RNA processing can be perceived (for review see Nevins 1983). The polyadenylation signal has been identified as AATAAA (Nevins, 1983, Mason et al, 1985). While additional elements (located 3' to the polyadenylation signal) have been identified as essential for the correct formation of 3' termini in several genes (Georgiev and Birnstiel 1985; McLauchlan et al, 1985). Polyadenylation may require the presence of certain small nuclear ribonucleoproteins (Birnstiel, Busslinger and Straub, 1985; Hashimoto and Steilz, 1986).

The primary transcript can be further processed by splicing to remove non-coding sequences. In contrast to the rapid polyadenylation addition process the subsequent splicing of a precursor RNA to the mature mRNA is a relatively slow process (for review see Nevins, 1983), and probably occurs through consensus sequences by lariat formation (Reed and Mannatis, 1985). Thus, the ability to achieve alternative splicing obviously allows the cell to alter the output of a gene without affecting the transcriptional control.

(ii). Transport and Stability of RNA

Once the fully processed, mature, eukaryotic mRNA is formed, transport to the cytoplasm must take place to allow function (translation). Control of this transport and the subsequent stability of the message are both important levels of regulation. The poly(A) segment plays a role in

determining mRNA stability and transport efficiency (for review see Nevins, 1983). However, there is some data to suggest that other elements within the message are involved in stability (Garshi et al, 1985; Piechaczyk et al, 1985) and the mRNA cap may also play some role in protecting the message from enzymatic degradation (for review see Shatkin, 1985). Recent studies have started to implicate specific transcribed sequences which can confer instability to a message (Shaw and Kamen, 1896).

(iii). Efficiency of Translation

Even after the myriad of steps, each with its own level of control, involved in getting the mature mRNA to the polysomes, the level of translation efficiency may be regulated. Although no specific consensus sequences have been identified which modulate the efficiency of translation, 5' untranslated leader sequences have been implicated (McGarry and Lindquist, 1985; Pelletier and Sonenberg, 1985; Shpaer, 1985).

1.9. Promoter Signals for Genes Transcribed by RNA Polymerase II

The molecular mechanisms which underlie the regulation of transcription in eukaryotes are still a mystery. However, it has been established that eukaryotic cells contain three distinct classes of RNA polymerase (for review see Chambon, 1975): Class 1, which catalyses synthesis of ribosomal RNA; Class II, that of mRNA and Class III, that of tRNA, 5sRNA and some highly repeated sequences, e.g. B2 family. The efficiency of transcription of eukaryotic protein coding genes is thus dependent on the interaction of RNA polymerase II with the transcription cofactors and their regulatory sequences within the genes' promoters. Thus the mechanisms of this process are dependent on the precise DNA sequence of the regulatory regions.

Therefore defining transcription regulatory signals represents an essential step in our understanding of eukaryotic gene expression. Analysis of genes transcribed by RNA polymerase II has revealed a complex and interrelated set of promoter elements. By the comparison of the 5' flanking DNA sequences of a large number of eukaryotic genes, the concept of a functional promoter domain has been formed on the basis of the conserved sequences identified. Not all sequences identified are homologous between groups, or even present at all in some genes, and as such individual promoters are neither structurally nor functionally identical. In the following chapters the regions which make up the promoter domain and their interactions are discussed.

1.9.1. Transcription Initiation Site

The transcriptional initiation site may, although it has not been formally proven, correspond to the 5' terminus of the primary transcript (for review see Nevins, 1983). The location of transcription initiation is defined by the upstream control elements (Mathis and Chambon, 1981). Even though certain non-coding conserved sequences surrounding the initiation site have been identified (Efstratiadis et al, 1980; Sures, Levy and Kedes, 1980), and although not directly comparable to each other, they do imply common but as yet unknown functions. Several deletion studies have shown this short sequence of DNA around the cap to be non-essential for efficient gene expression (Benosit and Chambon, 1981; McKnight and Kingsbury, 1982; Grosveld et al, 1982), however recent contrasting data implies that the relative positioning with respect to helical periodicity may be important (Kovacs and Butterworth, 1986). In addition to this major start site, several minor upstream starts have been identified in the globin gene family (Zhu, Allan and Paul, 1984) and human dihydrofolate reductase gene (Masters and Attardi, 1985). The functional significance of such transcripts is not understood, although they appear to initiate from within putative upstream regulatory sequences

(Zhu, Allan and Paul, 1984).

Although the presence of only one major transcription start site is common, some genes possess two major starts, e.g. human c-myc. The best example analysed is that of the α-amylase gene where differential usage of the two start sites is regulated in a tissue-specific manner (Schibler et al, 1983). Additionally, the differential usage of start sites has been shown to vary within the adenovirus replication phase (Osborne and Berk, 1983). Perhaps the differential usage of start sites will be found to be more common than considered at the moment.

1.9.2. The TATA Consensus Sequence.

By studying the putative promoter regions of the Drosophila histone genes, Hogness (1979) first identified a conserved AT-rich sequence 70bp upstream from the ATG initiation codon. Then by comparing the upstream sequences of a number of viral and eukaryotic genes similar AT-rich consensus sequences were shown to be located 20-30 bp upstream from the transcription initiation sites in these genes (Gannon et al, 1979). This, the first RNA polymerase II promoter element to be identified, is usually termed the TATA consensus sequence (or sometimes the Goldberg-Hogness box), and has been regarded as analogous to the Pribnow sequence of prokaryotic organisms (Pribnow, 1975).

By comparison of the promoter sequences of 60 eukaryotic protein coding genes a consensus sequence of the form 5'- TATA$^{T}_{A}$A -3' has been compiled (Breathnach and Chambon, 1981). The exact position of the TATA-consensus varies slightly between genes. However, there are some eukaryotic coding genes and some viral genes which do not have an analogous sequence (Baker et al, 1979; Flavell et al, 1979; Brady ey al, 1982; Dynan, 1986).

There has been extensive in vitro (Grosscheld and

Birnstiel 1982) and in vivo analysis (Benoist and Chambon, 1981; Dierks et al, 1983), by deletion and point-mutation, with the general consensus that this element functions to fix spatially the initiation of transcription. This can be correlated with the naturally occuring 5' heterogeneity found within certain virus mRNA species, where no TATA-consensus is present (Brady et al, 1982). In addition, studies where the TATA-consensus has been deleted have implicated that it may also have a role in controlling transcription efficiency (Dierks et al, 1983). Recently, naturally occuring mutations in the TATA-consensus of the β-thalassaemia globin gene with a corresponding affect on transcription have been documented (Antonarakis et al, 1984) which again argues for the TATA-consensus having a role in governing efficiently the correct transcription initiation site.

How the TATA-consensus element brings about this regulation is not known. However, progress in understanding the mechanism should come soon with the isolation of a TATA-consensus binding protein (Wu, 1985).

1.9.3. The CCAAT Consensus Sequence

The next RNA polymerase II transcription control element was identified by comparison of homologous sequences within the 5' flanking region of a number of eukaryotic genes, and is termed the CCAAT consensus sequence (Efstratiadis et al, 1980; Benoist, et al, 1980). This sequence, 5'-GGTCAATCT-3', is found 50 to 90 bp upstream from the transcription initiation site (Benoist et al, 1980) and in addition to being highly conserved amongst eukaryotic genes (Dierks et al, 1983) is also found in some viral genes (Han et al, 1982). In addition, this region shows similarity to the -35 region of some Escherichia coli promoters (Benoist et al, 1980).

In general, the CCAAT sequence appears to be essential

for efficient gene transcription in vivo, but is not ubiquitous in nature (Anagnou et al, 1986). Again, how this element functions is only now beginning to be analysed with the recent isolation of a CCAAT-binding transcription factor (CTF: Kadonaga, Jones and Tijan, 1986; Cohen, Sheffery and Kim, 1986; Mcknight and Tjian, 1986). However due to its absence from certain genes and its non-essential role in certain other genes (Grosscheld and Birnstiel, 1982), the presence of additional upstream sequence elements is implied.

As data have become available on a large number of eukaryotic genes the 5' extent of the promotor sequences controlling transcription has become known. In general, sequences further upstream than the CCAAT element are absolutely required for efficient in vivo gene expression. Their identification has been a problem, as unlike the TATA-consensus and CCAAT element, there appears to be a lack of sequence homology between most of these upstream control regions in unrelated promoters. Thus it has not been possible to predict the precise location of these upstream regulatory elements by homology alone. In most cases, their initial identification has been achieved by determining the effect of DNA sequence alterations on the level of gene transcription. However, some of these regions do show a consensus sequence or at least a distinct repeated sequence.

1.9.4. The Upstream Promoter Elements

During the last few years the amount of data on upstream promoter elements has increased quite amazingly. As so many elements have been identified, a sub-classification can now begin to be made. I have grouped these elements into five tentative groups, which are themselves diverse in nature, and in some cases may indeed overlap each other. All the upstream promoter elements appear to have a common requirement for a trans-acting factor(s) to activate them. Most of these elements are 5' to the TATA-box, however some

have been located 3' to the coding region or within intronic sequences. The first group is the CCAAT-box which has been already discussed. It was discussed separately for historical reasons only (being the first identified). The last group I will mention will be the enhancers, as they have given the best insight into how such elements may function. Much of the data amassed for enhancers may also be considered of relevance to the other upstream elements.

(a). The GC Rich/SP1 Binding Element

Located at approximately the same distance from the transcription start site as the CCAAT element, and first identified in the SV40 promoter domain, is a GC-rich element with the consensus sequences 5'- $\overset{GG}{TA}$GGCG$\overset{GGGC}{TAAT}$-3' (Vigneran et al, 1984; Dynan and Tjian, 1985). This sequence plays a critical role in directing efficient transcription not only for certain viruses (for reviews see Kadonaga, Jones and Tjian, 1986; McKnight and Tjian, 1986) but also for a subset of mammalian promoters (Das and Piatigosky, 1986; Dynan et al, 1986; Kadonaga, Jones and Tjian, 1986), probably only a few of which have been identified (see section 4.). Many of the so- called housekeeping genes are in this group, possessing numerous GC-rich elements but no obvious TATA-consensus (Dynan, 1986).

This upstream promoter element appears to be functional in either orientation (possibly bipolar), and is usually present in multiple copies, with the most important site(s) usually being those proximal to the TATA-consensus (Dynan and Tjian, 1985; Kadonaga, Jones and Tjian, 1986; Ishii et al, 1986). The *trans*-acting factor which recognises these GC-rich motifs was first identified in SV-40 and is designated Spl. This protein has been isolated and purified, and shown to exhibit a sequence-specific DNA binding activity. Binding of Spl enhances transcription *in vitro* by RNA polymerase II 10-50 fold. Where additional promoter elements are present *in vivo* (e.g. the CCAAT element), Spl

appears to act in conjunction with other transcription factors (e.g. possibly CTF), to stimulate efficient transcription (for review see Kadonaga, Jones and Tjian 1986). Although possible, Sp1 has not been shown to be associated with a negative regulator of transcription.

How Sp1 might function is at the moment unknown. Since genes with no TATA-consensus but several GC-rich elements show heterogenous transcription start sites (for review see Dynan, 1986), perhaps Sp1 differs in its function to that of the TATA-consensus binding factor. The GC-rich domains of many promoters contain an unusually high number of GC nucleotides, and these dinucleotides are undermethylated when compared to the rest of the genome (Bird, 1986). Conceivably methylation of these sites may prevent Sp1 interaction. This is only a correlation of observations, but perhaps it may be of importance in control of expression. If demethylation of the GC dinucleotides is indeed an active process then a factor, such as Sp1, may play an role in their generation.

(b). Receptor-Induced Elements

This, by definition, is a diverse group of upstream promoter elements which is basically a subset of enhancer elements. More than likely this group will expand as more investigation is undertaken. These elements are activated by the binding of a trans-acting factor (e.g. a hormone activated receptor), which can have the ability to induce an individual or a group of genes thus allowing co-ordinate control of expression.

The best example is that of the glucocorticoid receptor which can activate both the metallothionein gene and Mouse Mammary Tumour Virus enhancer (Lato et al, 1984;Karen et al, 1984; von de Ahe et al, 1985; Moore et al, 1985). Using these two systems activation has been shown to be due to an active glucocorticoid hormone-recptor complex directly

binding to the respective enhancer elements (Miksicek et al, 1986). This binding is through the specific recognition sequence, 5'- GTGTGGAAAG -3' (Sassone-Corse and Borrelli, 1986). Since this enhancer can function without receptor activation, other factors must be able to interact with at least some part of the enhancer element (Sassone-Corse and Borrelli, 1986), but it is the receptor complex which confers the specificity to the system. This requirement for several factors is clearly seen for the metallothionein element where induction can be from heavy metals or glucocorticoids, thus impling the presence of several, possibly overlapping recognition elements (Sassone-Corse and Borrelli, 1986).

Other members of this family include the oestradiol-receptor complex (Jost, Geiser and Seldran, 1985)and the growth factor induction of the c-fos (Treisman, 1985) and probably c-myc genes (Kelly et al, 1983). These elements (and probably other upstream promoter elements) have a modular nature, in that the interaction of several factors are required (thus several cis-acting recognition sequences) for activity. In each case the cis-acting response elements will probably reside close to, although not necessarily adjacent to the promoters they regulate. Recently, the serum-responsive sequence required for transcriptional activation of the c-fos gene has been identified as an inverted repeat of 5'-GATGTCCATA-3' (Treisman, 1986). It will be interesting to see if any of the other serum inducible genes have this (or a similar) element adjacent to their promoters.

(d). Enhancer Elements

It is now becoming very hard to define an enhancer element because the characteristics originally thought to be specific for them are being found to be true for other upstream elements. Therefore some of the data already amassed for enhancer elements will probably be relevant in

the understanding of how upstream promoter elements function in general. However, enhancer elements can be found in addition to other upstream promoter elements in some systems.

The list of such elements has expanded rapidly since the prototype (the 72 bp repeat element of SV-40) was originally identified (Banergi, Rusconi and Schaffner, 1981) and includes both viral and eukaryotic examples (for review see Khoury and Gruss, 1983).

Although little homology exists among enhancer elements from different species, a core sequence has been identified as 5'-GTGG $\overset{TTT}{AAA}$ G-3', but even this is not universal (for review see Khoury and Gruss, 1983). By itself this sequence is not sufficient for full function, suggesting that neighbouring sequences contribute to enhance function (for review see Yaniv, 1984). These elements function optimally as an intact element, the regions of importance apart from the core element, are still little characterised (Cochran and Weissman,1984; Zenke et al, 1986; Hearing and Shenk, 1986). Thus enhancers appear to have a modular nature, possibly with each sub-element having its own specific individual role in allowing the enhancement of transcription (e.g. hormone receptor responsive).

In addition to their lack of sequence homology, there is little structural similarity between enhancers. Duplicated sequences are common (Laimins et al, 1984) but are not an absolute requirement, as related viruses without duplicated enhancer sequences remain viable (Khoury and Gruss, 1983). Of note, there appears to be an homologous duplicated sequence upstream from the Xenopus laevis ribosomal gene spacer (transcribed by RNA polymerase I: Labhart and Reeder, 1984). However, enhancer elements do share local homologous twist-angle variations with a helical periodicity. This recurring spatial geometry may account in part for their similar characteristics (Nussinov et al, 1984).

A comparison of the available data has allowed the following *in vivo* enhancer characteristics to be identified (for reviews see Khoury and Gruss, 1983; Serfling, Jasin and Schaffner, 1985).

(i).They are able to activate heterologous promoters. Due to this property, enhancers have been used to construct high level expression vectors for mammalian cells and thus allow genes with weak promoters to be analysed.

(ii).They function in an orientation independent manner.

(iii).Enhancers have the property of being able to exert their effect over relatively long distances (in some cases over several kilobases) although the presence of certain sequences may interfere with this property. When adjacent to more than one promoter, the enhancement was found to be preferentially for the most proximal.

(iv).They can initiate transcription in the absence of an apparent endogenous promoter.

(v).They can be located 5'or 3' to, or within an intron of the gene.

(vi).There is now quite considerable evidence suggesting that there are subtle differences between enhancer elements and as a consequence they can function in a tissue- and/or species-specific manner.

These properties had set them apart from the upstream promoter elements. More recent investigations however, suggest that enhancer and promoter elements may overlap both physically and functionally (Serfling, Jasin and Schaffner, 1985; Foster, Stafford and Queen, 1985). How an enhancer may function is only now beginning to be understood. Several studies have shown enhancer activity to be due to the interaction of the DNA sequence with a *trans*-acting factor

(Sassone-Corsi, Wilderman and Chambon, 1985; Scholer and Gruss, 1985; Gaffeny et al, 1985; Treisman, 1986; Ohlsson and Edhund, 1986). In addition, the SV-40 enhancer has been shown to increase directly the polymerase density within a linked gene, therefore, apparently increasing the rate of transcription initiation without influencing the specificity of initiation (Weber and Schaffner, 1985; Treisman and Maniatis, 1985). Also, DNase sensitive sites sometimes correlate with the presence of enhancer sequences in chromatin (for review see Khoury and Gruss, 1983).

From this data a speculative model of how enhancers (and probably other upstream promoters) may function can be proposed. The enhancer element could function in association with nuclear proteins to present an "open-window" thereby allowing entry of the RNA polymerase II. Linear diffusion of the polymerase to the control promoter elements could then occur, (e.g. CCAAT, Sp1, TATA), which then spacially fix transcription initiation, presumably through their respective trans-acting factors. The modular nature of enhancers allows for the interaction of several trans-acting factors, possibly each conferring its own degree of specificity upon the enhancer.

A recent report on transcription control of the SV-40 early promoter has provided more information on this aspect. This promoter domain contains an enhancer element, a Sp1/GC-rich domain and a classical TATA-consensus. Upon inserting odd multiples of half a DNA turn between each of these respective elements (in various combinations), a marked decrease in in vivo activity was observed (Takahashi et al, 1986). This indicates that the helical periodicity of the respective cis-acting elements is crucial, implying that the correct ster eospecific alignments between the trans-acting factors binding to these elements are essential for efficient transcription regulation.

The interaction of upstream elements was shown by recent

analysis of the HSV-1 IE 4/5 gene (Gaffney et al, 1985). The HSV enhancer, the 5'-TAATGARAT-3' consensus sequence, responds to a *trans*-acting viral component, presumably opening up the chromatin adjacent to the gene, thereby allowing entry of the polymerase. However the GC-rich sequences (possibly Sp1) near the transcription initiation site can affect the magnitude of transcription. The enhancer is conferring the specificity to the system.

In addition, several studies have indicated that a potential to form Z-DNA may participate in the formation of such an "open-window" (Nordheim and Rich, 1983, Azorin and Rich, 1985), but there is as yet little direct evidence that Z-DNA actually exists. Recent studies have indicated that it is the helical periodicity within and between adjacent consensus elements which is of importance (Das and Sulzman, 1985; Takahashi et al, 1986; Rhodes and Klug, 1986).

Since it has become apparent that full enhancer activity can not be achieved *in vitro* with soluble transcription extracts, it has been proposed that higher-order nuclear structures are also required. Recent reports give evidence for this by identifying nuclear matrix/scaffold-associated regions (SARs) associated with both enhancer consensus sequences and topoisomerase II recognition elements (murine immunoglobulin gene: Cockerill and Garrard, 1986; *Drosophila*: Gasser and Laemmli, 1986).

Enhancers function as their name implies, to stimulate a positive enhancement of transcription. However, given the correct *trans*-acting factor these *cis*-acting elements can cause the repression of a gene (Gorman, Rigby and Lane, 1985; Jones, 1986). How this occurs is unknown, but presumably it is a function of the respective *trans*-acting factors. To complicate matters the E1a gene product can both activate and repress enhancer action (for review see Jones, 1986).

(d). Other Upsteam Promoter elements

This is a diverse category containing all the elements known to-date which do not fall into one of the other groups. As these elements are further analysed they will probably be found to be members of the other categories. To describe each here would be nonproductive, but there are two factors of interest with respect to the work presented in this thesis.

Firstly, the adenovirus EIa gene product (see section 1.6.4.) has been shown to regulate transcription from its own viral genes and from cellular genes. Surprisingly perhaps, this trans-acting factor functions in a sequence-independent manner (Hearing and Shenk, 1985; Kingston, Kaufman and Sharp, 1984). Here we have a trans-acting factor which regulates transcription without the requirement of an obvious promoter sequence! However, another possibility is that the E1a gene product functions in association with other sequence-specific DNA-binding proteins and is therefore not truly sequence non-specific (Nevins, personal communication).

As already mentioned (see section 1.4.1.), the c-myc gene may have at least one Nuclear Factor-1 (NF-1) binding site in its 5' flanking sequences (Siebenlist et al, 1984). There are now several genes (Hennighausen and Fleckenstein, 1986) which have the consensus sequence, 5'-TGGCANNATGCCA-3' and which bind NF-1 in vivo (for review see Dynan, 1985). Until recently no role in transcription regulation had been demonstrated for this element, although it had been associated with sites of DNase 1 hypersensitivity (for review see Dynan, 1985). Upon removal of the domain containing the NF-1 recognition sequence from the Hepatitis B virus surface antigen promoter there is a 10-fold decrease in gene expression observed by transient transfection assay. In this system this region has been shown to bind NF-1 factor in vitro (Shaul, Ben-Levy and De-Medina, 1986).

Although this is not proof of an in vivo functional interaction, it implies that it may occur. NF-1 has been shown to stimulate adenovirus DNA replication (de Vries et al, 1985). If a role in transcription is positively assigned to this factor, it is currently thought likely that it would function similarly to Sp1 and CTF in that it would augment expression control, and not confer specificity to the system directly (see section 4.). Since these studies another factor, NF-III which, has similar characteristics, and binds to a sequence which overlaps the NF-I consensus has been identified (Pruijin, van Driel and van der Vliet, 1986).

1.9.5. Summary of Upstream Promoter Elements

Both upstream promoters and enhancer elements may have the ability to confer specificity to their individual genes, through interaction with specific trans-acting factors. They both probably act, enhancers almost certainly do, by altering the chromatin state around the gene to be transcribed into a more open conformation, thus presumably allowing easier access of RNA polymerase II and any attendant transcription cofactors. In addition, there would appear to be a subclass of upstream promoter elements, with their respective trans-acting factors (Sp1, CTF and possibly NF-1) which somehow augmentt the level of transcription.

Although in the above discussion these elements have been termed upstream promoter elements, it must be remembered that transcription regulatory elements can be located 3' to or within the transcribed unit.

1.10. Negative Regulatory Elements

Specific sequences which promote transcription may only be one side of cellular transcription control mechanisms. Recent data from several systems has shown that repression of transcription may be just as an important mechanism with

which cellular integrity is maintained. Two viral gene products are known to repress their own synthesis by interacting directly with their own promoters: HSV-VP175 (Dixon and Schaffner, 1980) and the SV-40 large T-antigen (Rio and Tjian, 1983).

Recently, several genes have been shown to be regulated, at least in part, by some form of repressor : the Beadex locus of Drosphila (Mattox and Davidson, 1984), the H-1 region of β-globin (Gilmour et al, 1984), p53 (Bienz-Tadmor et al, 1985), rat insulin-1 gene (Laimins, Holmgren-Konig and Khoury, 1986), human immunoglobin heavy chain locus (Gough, 1985), HMG coenzyme A reductase gene (Gaub et al, 1985), ovalbumin gene (Osborne, Goldstein and Brown, 1985) and the yeast mating type locus (Brand et al, 1985). The name given to these regions varies (negative regulatory element, dehancer, blocker, silencer), however, in each case they function similarly to repress transcription (albeit to various degrees). The cases which have been studied in greater detail show that these negative regulatory elements are cis-acting, orientation independent and relatively lax about their positioning to homologous or heterologous promoters (Gilmour et al, 1984; Brand et al, 1985; Laimins; Holmgren-Konig and Khoury, 1986).

At the present moment we can only speculate at how they function, although several candidate properties have been postulated to be of possible importance. The Hl locus (Gilmour et al, 1984) has the potential to form Z-DNA while the rat insulin locus (Laimins, Holmgren-Konig and Khoury, 1986) was identified as a member of a large interspersed repetitive family. In addition, transcription of the HMG CoA reductase gene (Reynolds et al, 1984) appears to be inhibited by sequences present in the untranslated leader sequences. Whether these properties are of any relevance to the repression function awaits analysis. Although the individual DNA sequences will be essential, the effector must be a trans-acting factor. This has been clearly

illustrated by the extinction experiments in cell hybrids (Killary and Fournier, 1984) where the extinction (repression) of a given tissue-specific trait was correlated to a specific locus.

Transcription repression requires cis-acting sequences (negative regulatory elements: NRE) to be postulated (Renan, 1985). More than likely these sequences will only function when the correct trans-acting factor (repressor) interacts with them. In possibly analogous bacterial systems, more is known about this repressor-NRE interaction (for review see Papo and Sauer, 1984). The bacterial repressor is thought to be a symmetrical protein which binds to one side of the double helix (i.e. the recognition sites are 10bp apart). This recognition sequence (to which the repressor interacts) is thought to reside in the major groove of the B-DNA double helix (Pabo and Sauer, 1984). The periodicity and thus the relative positioning of the recognition sequence may be crucial for repressor binding, similar to what is currently thought for positive regulators. Activation of the gene would then require removal or inactivation of the repressor, or the presence of an activator.

The work described in this thesis identifies a NRE in the 5' flanking sequences of the human c-myc gene and as such this topic will be expanded in section 4 of this thesis.

1.11. The Transcription Termination Signal

Transcription 3' end processing in eukaryotes is heterogeneous, but regulated, and depends on specific DNA sequences downstream of the 3' processing site (s) (McLauchlan et al, 1985; Birnsteil, Busslinger and Strub, 1985). The factors which regulate termination of transcription and subsequent 3' end processing are as yet only vaguely understood.

Since alternative 3' ends can be generated, and the actual 3' end achieved may in turn control processing events or affect the stability of the transcript, this is, an obvious point at which expression could be controlled. However, this is not at the level of control of transcription. There are examples where 3' sequences affect the level of expression of a gene but this is presumably due to post-transcriptional effects (Triesman, 1985). As of yet there are no examples where 3' sequences regulate the level of transcription initiation.

1.12. The Promoter Domain : A Summary

Although I have described each transcriptional element individually it cannot be over-stressed that they function together as a whole to form the full functional promoter domain of a gene. Although individual elements have been identified how they function is still unknown, but speculations can be made.

One interpretation of what happens in the promoter domain could be as follows. Upon activation of the upstream promoter/enhancer elements, mediated by a *trans*-acting factor, an open chromatin state is induced, which we can detect due to its increased sensitivity to DNase 1 digestion. Somehow these active regions are associated with the nuclear substructure and there is now some evidence that this may involve enhancer-like elements. This state then allows entry of the RNA polymerase II, and any required cofactors, onto the DNA. The polymerase then diffuses along the DNA until recognising the transcription initiation control elements; possibly the TATA- , SP1- or CCAAT-elements as they are in nature conserved sequences. These elements (and thus the *trans*-acting factors which recognise them) may help to regulate the magnitude of transcription. The specific sequences of the individual upstream promoter/enhancer elements probably confer the species, tissue and temporal regulation upon specific genes.

Negative regulatory elements could function to prevent the binding or movement of the polymerase, presumably by the binding of a specific factor (repressor) to this site. At a higher level nucleosome spacing could form the basis of the selective gene expression during development and differentiation (Stein and Kunzler, 1983).

Co-ordinate control of an individual or group of genes could be mediated by trans-acting factors binding to gene specific upstream control sequences. Specificity in these systems could be conferred by the specific cis-acting element, the trans-acting factor or the presence of several trans-acting factors (repression or activation).

1.13. RNA Polymerase II Associated Factors

These are general factors unlike the specific proteins previously discussed, which are involved in generating a transcription complex and not transcription activation. By definition, RNA polymerase II must be capable of binding to DNA to function, but it lacks any inherent ability to recognise specific sequences and as such will initiate transcription totally at random (for review see Lewis and Burgess, 1982). Initiation only occurs at specific sites upon the presence of the previously discussed promoter-associated factors. It is now thought that there are other factors which interact with the polymerase to form preinitiation and initiation complexes, and only then transcription occurs (Culotta, Wides and Sollner-Webb, 1985; Safer et al, 1985; Sawadogo and Roeder, 1985).

1.14. Rational for Work Presented in this Thesis

Since aberrant expression of the human c-myc oncogene is observed in transformed cells, and expression of c-myc in conjunction with the c-ras oncogene induces transformation, the regulation of its expression can be expected to be tightly regulated in the normal cell. From serum induction

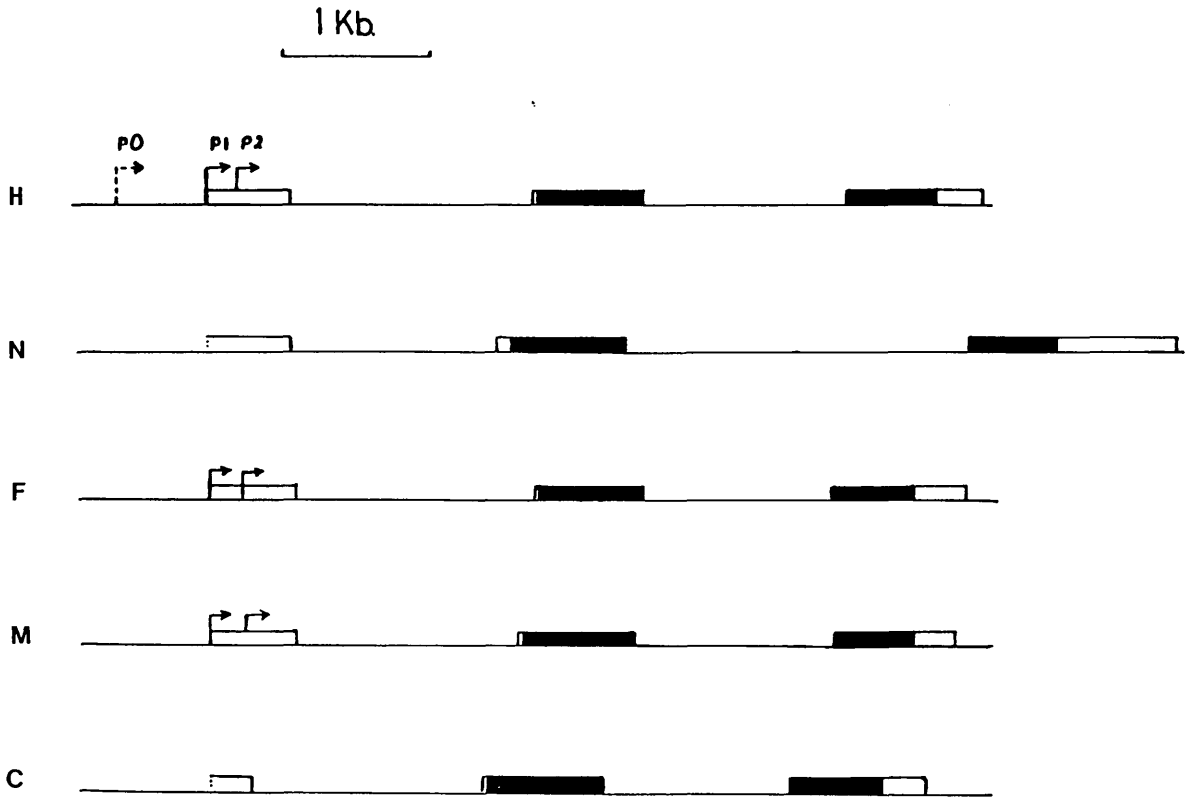
studies (Kelly et al, 1983) and the expressional silence of the normal allele in Burkitt's lymphoma cells (for review see Leder et al, 1983), a repressor has been postulated to control, at least in part, c-myc transcription. As previously discussed, most transcriptional regulatory elements are located 5' to a gene, but elements 3' or even within introns have been identified (but current data would implicate the 5' region in c-myc: for review see Leder et al, 1983). DNase I sensitivity analysis (Siebenlist et al, 1984) correlates with the location of chromosome translocation breakpoints (for review seee Leder et al, 1983) to locate the hypothetically required cis-acting negative regulatory element to the 5' flanking sequences of the c-myc oncogene. However it is possible that a non-specific interaction could occur thus negating any requirement for a cis-responsive sequence element.

The aim of the work described in this thesis was to identify and localise such a negative regulatory element (NRE), if one was present, by analysis of the 5' flanking sequences of the human c-myc oncogene with functional expression assay, in cell cultures growing in log phase. As a consequence of such a search other cis-acting elements present in this region could be identified (as described c-myc expression is regulated by several factors: serum, interferon, the differentiation state). In addition, although the dual promoter region contains two TATA-boxes (Battey et al, 1983) there is no easily identifiable upstream promoter, although there are candidate domains. The presence of such a promoter (again if one was present), or any additional transcriptional regulatory element, would also be identified in the line of this research.

By identifying the transcriptional regulatory elements involved in c-myc expression a greater insight into how c-myc is deregulated in the transformed cell will hopefully be achieved.

Figure 1.

Comparison of the Sequenced Genomic c-myc Genes.



Depicted in this diagram is the genomic structure, to scale, of the various cellular myc genes sequenced. H = human c-myc; N = human N-myc; F = feline c-myc; M = murine c-myc; C = chicken c-myc (Bernard et al, 1983; Nottenburg and Varmus, 1986; Stewart et al, 1986; Kohl et al, 1986). Boxes represent exons; open being untranslated, shaded being translated. The known transcription initiation sites are indicated by arrows (with the human major start sites designated P0, P1 and P2); where heterogeneous starts occur these have been shown by an open end (dots) to the first exon.

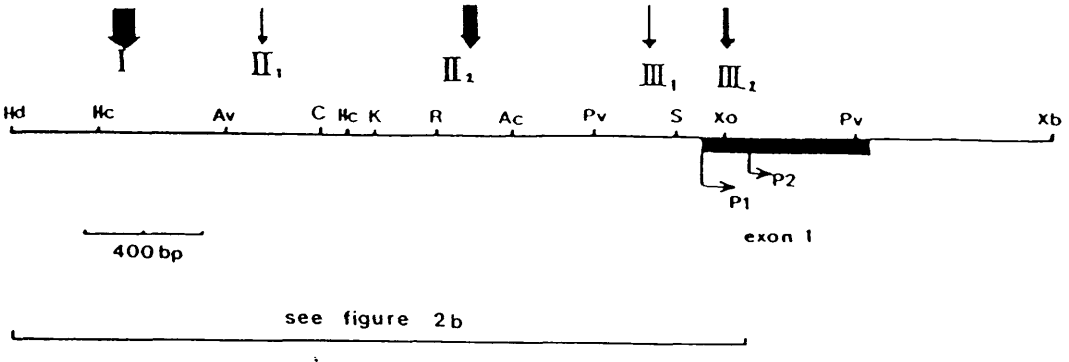
Figure 2.

Human c-myc 5' Flanking Region.

Shown below (A.) are the in vivo DNase I hypersensitve sites (indicated by Roman numerals and solid arrows, Siebenlist etal, 1984) within the human c-myc region, and relevant restriction sites (see figure 8 for key). The relative intensities of the DNase I hypersensitve sites are indicated by the width of the arrow. The sequence corresponding to the indicated region (sequence) is shown on the next page (B.). P1 and P2 refer to the two major transcription initiation sites (P0 has not been localised accuratly; Bentley and Groudine, 1986).

Key:

core enhancer sequence
Sp1 consensus (showing orientation)
CTF consensus
NF-1 consensus



B.

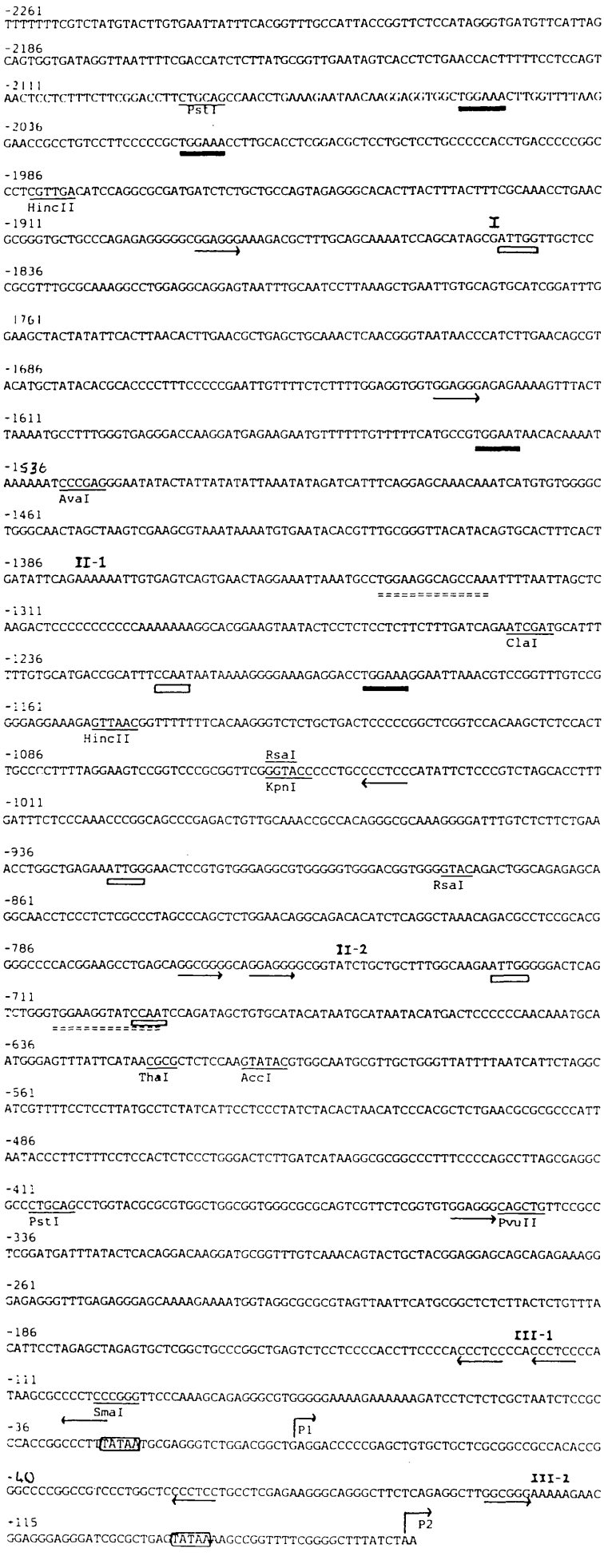

-2261
TTTTTTTCGTCTATGTACTTGTGAATTATTTCACGGTTTGCCATTACCGGTTCTCCATAGGGTGATGTTCATTAG
-2186
CAGTGGTGATAGGTTAATTTTCGACCATCTCTTATGCGGTTGAATAGTCACCTCTGAACCACTTTTTCCTCCAGT
-2111
AACTCCTCTTTCTTCGGACCTTCTGCAGCCAACCTGAAAGAATAACAAGGAGGTGGCTGGAAACTTGGTTTTAAG
PstI
-2036
GAACCGCCTGTCCTTCCCCCGCTGGAAACCTTGCACCTCGGACGCTCCTGCTCCTGCCCCCACCTGACCCCCGGC
-1986
CCTCGTTGACATCCAGGCGCGATGATCTCTGCTGCCAGTAGAGGGCACACTTACTTTACTTTCGCAAACCTGAAC
HincII
-1911
I
GCGGGTGCTGCCCAGAGAGGGGGCGGAGGGAAAGACGCTTTGCAGCAAAATCCAGCATAGCGATTGGTTGCTCC
-1836
CGCGTTTGCGCAAAGGCCTGGAGGCAGGAGTAATTTGCAATCCTTAAAGCTGAATTGTGCAGTGCATCGGATTTG
1761
GAAGCTACTATATTCACTTAACACTTGAACGCTGAGCTGCAAACTCAACGGGTAATAACCCATCTTGAACAGCGT
-1686
ACATGCTATACACGCACCCCTTTCCCCCGAATTGTTTTCTCTTTTGGAGGTGGTGGAGGGAGAGAAAAGTTTACT
-1611
TAAAATGCCTTTGGGTGAGGGACCAAGGATGAGAAGAATGTTTTTTGTTTTTCATGCCGTGGAATAACACAAAAT
-1536
AAAAAATCCCGAGGGAATATACTATTATATATTAAATATAGATCATTTCAGGAGCAAACAAATCATGTGTGGGGC
AvaI
-1461
TGGGCAACTAGCTAAGTCGAAGCGTAAATAAAATGTGAATACACGTTTGCGGGTTACATACAGTGCACTTTCACT
-1386
II-1
GATATTCAGAAAAAATTGTGAGTCAGTGAACTAGGAAATTAAATGCCTGGAAGGCAGCCAAATTTTAATTAGCTC
==========
-1311
AAGACTCCCCCCCCCCCAAAAAAAGGCACGGAAGTAATACTCCTCTCCTCTTCTTTGATCAGAATCGATGCATTT
ClaI
-1236
TTTGTGCATGACCGCATTTCCAATAATAAAAGGGGAAAGAGGACCTGGAAAGGAATTAAACGTCCGGTTTGTCCG
-1161
GGGAGGAAAGAGTTAACGGTTTTTTTCACAAGGGTCTCTGCTGACTCCCCCGGCTCGGTCCACAAGCTCTCCACT
HincII
-1086
RsaI
TGCCCCTTTTAGGAAGTCCGGTCCCGCGGTTCGGGTACCCCCTGCCCCTCCCATATTCTCCCGTCTAGCACCTTT
KpnI
-1011
GATTTCTCCCAAACCCGGCAGCCCGAGACTGTTGCAAACCGCCACAGGGCGCAAAGGGGATTTGTCTCTTCTGAA
-936
ACCTGGCTGAGAAATTGGGAACTCCGTGTGGGAGGCGTGGGGGTGGGACGGTGGGGTACAGACTGGCAGAGAGCA
RsaI
-861
GGCAACCTCCCTCTCGCCCTAGCCCAGCTCTGGAACAGGCAGACACATCTCAGGCTAAACAGACGCCTCCGCACG
-786
II-2
GGGCCCCACGGAAGCCTGAGCAGGCGGGGCAGGAGGGGCGGTATCTGCTGCTTTGGCAAGAATTGGGGGACTCAG
-711
TCTGGGTGGAAGGTATCCAATCCAGATAGCTGTGCATACATAATGCATAATACATGACTCCCCCCAACAAATGCA
==========
-636
ATGGGAGTTTATTCATAACGCGCTCTCCAAGTATACGTGGCAATGCGTTGCTGGGTTATTTTAATCATTCTAGGC
ThaI
AccI
-561
ATCGTTTTCCTCCTTATGCCTCTATCATTCCTCCCTATCTACACTAACATCCCACGCTCTGAACGCGCGCCCCATT
-486
AATACCCTTCTTTCCTCCACTCTCCCTGGGACTCTTGATCATAAGGCGCGGCCCTTTCCCCAGCCTTAGCGAGGC
-411
GCCCTGCAGCCTGGTACGCGCGTGGCTGGCGGTGGGCGCGCAGTCGTTCTCGGTGTGGAGGGCAGCTGTTCCGCC
PstI
PvuII
-336
TCGGATGATTTATACTCACAGGACAAGGATGCGGTTTGTCAAACAGTACTGCTACGGAGGAGCAGCAGAGAAAGG
-261
GAGAGGGTTTGAGAGGGAGCAAAAGAAAATGGTAGGCGCGCGTAGTTAATTCATGCGGCTCTCTTACTCTGTTTA
-186
III-1
CATTCCTAGAGCTAGAGTGCTCGGCTGCCCGGCTGAGTCTCCTCCCCACCTTCCCCACCCTCCCCACCCTCCCCA
-111
TAAGCGCCCCTCCCGGGTTCCCAAAGCAGAGGGCGTGGGGGAAAAGAAAAAAGATCCTCTCTCGCTAATCTCCGC
SmaI
-36
P1
CCACCGGCCCTTTATAATGCGAGGGTCTGGACGGCTGAGGACCCCCGAGCTGTGCTGCTCGCGGCCGCCACACCG
-40
III-1
GGCCCCGGCCGTCCCTGGCTCCCCTCCTGCCTCGAGAAGGGCAGGGCTTCTCAGAGGCTTGGCGGGAAAAAGAAC
+115
P2
GGAGGGAGGGATCGCGCTGACTATAAAAAGCCGGTTTTCGGGGCTTTATCTAA

Figure 3.

The Cell Cycle.

There are three phases to within the G0-G1 period of the cell cycle. There is G0, where the cells are quiscent. Certain growth factors can stimulate cells in G0 to enter G1. They are termed Competence Factors (e.g. PDGF, platelet-derived growth factor; EGF, epidermal growth factor; FGF, fibrobast growth factor; IL-2 and IL-3, interlukin 2 and 3). The Progression Factors are then required to allow traverse of the G1 phase (through a restriction point, RP - still ill defined). In the final phase the cells are commited to finish the cell cycle (although there may be another restriction point within G2). Where the expression of an oncogene is transiently effected within the cycle, it has been indicated. A 68K protein has been proposed to be involved in the overcoming of the G1 restriction point. (Stiles et al, 1979; Kelly et al, 1983; Reed et al, 1985; Conscience, Verrier and Martin, 1986; Treisman, 1986; and data presented at a UCLA Symposium, 1986).

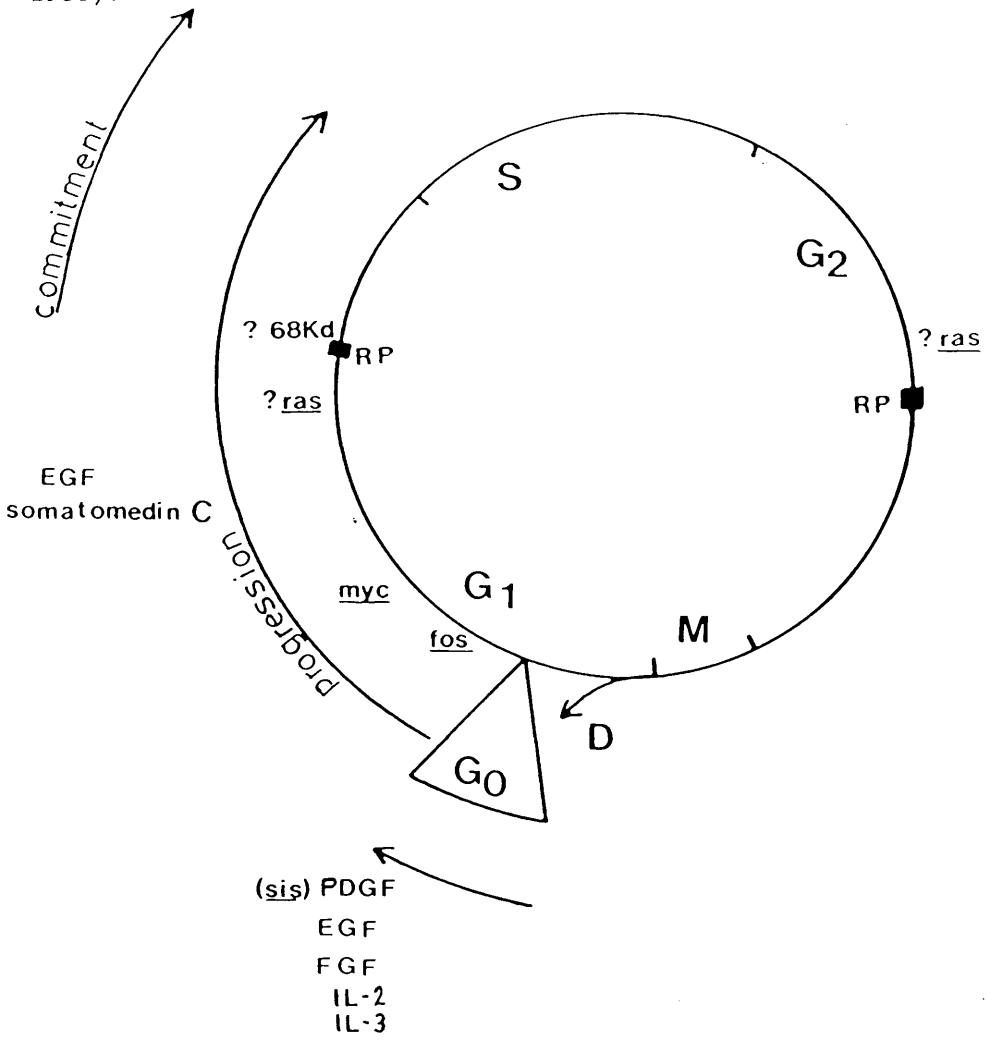


Figure 4.

Selected Burkitt's Lymphoma Breakpoints.

Depicted in the diagram are some (not all) of the known Burkitt's lymphoma (BL) breakpoint locations. The Roman numerals indicate the appropriate DNase 1 hypersensitive sites (Siebenlist et al, 1984). The arrows indicate the two major human transcription start sites (P1 and P2). The individual BL cell lines are as described in Bernard et al, (1983); Saito et al, (1983); Taub et al, (1984a and 1984b); Rabbits et al, (1984); Denny et al, (1985); Show et al, (1985).

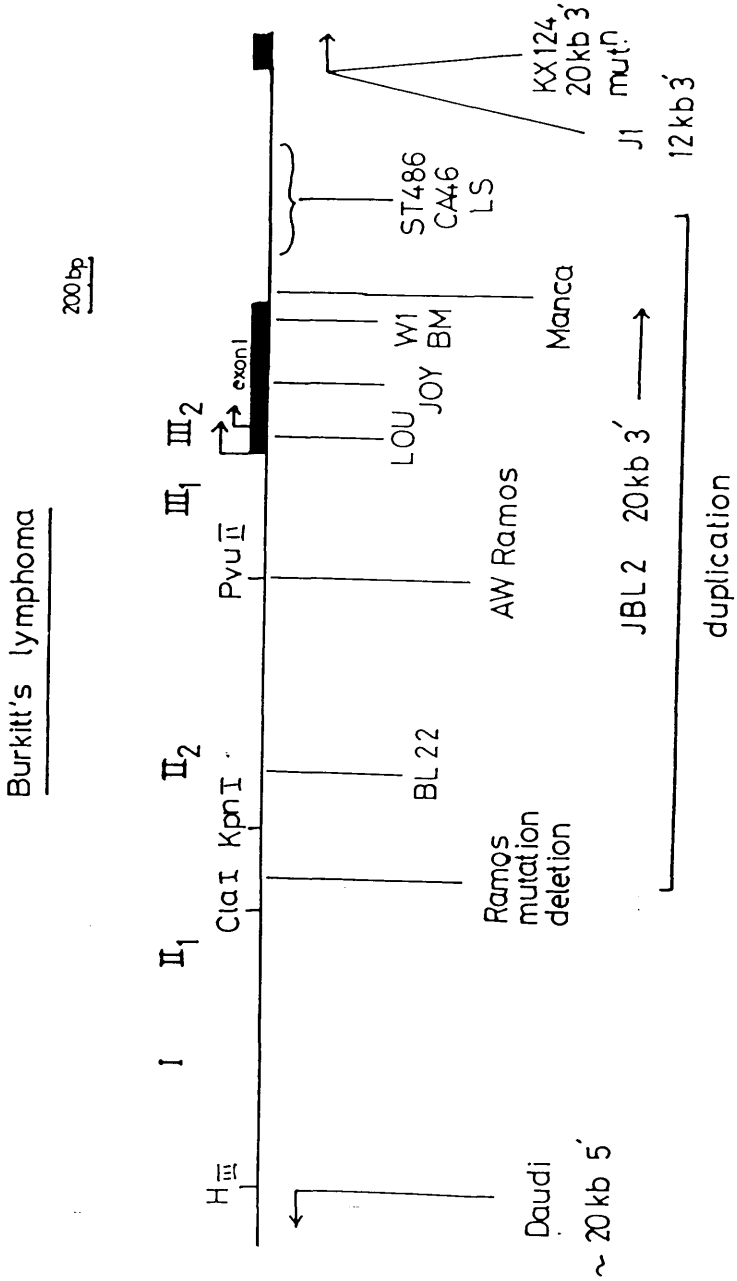
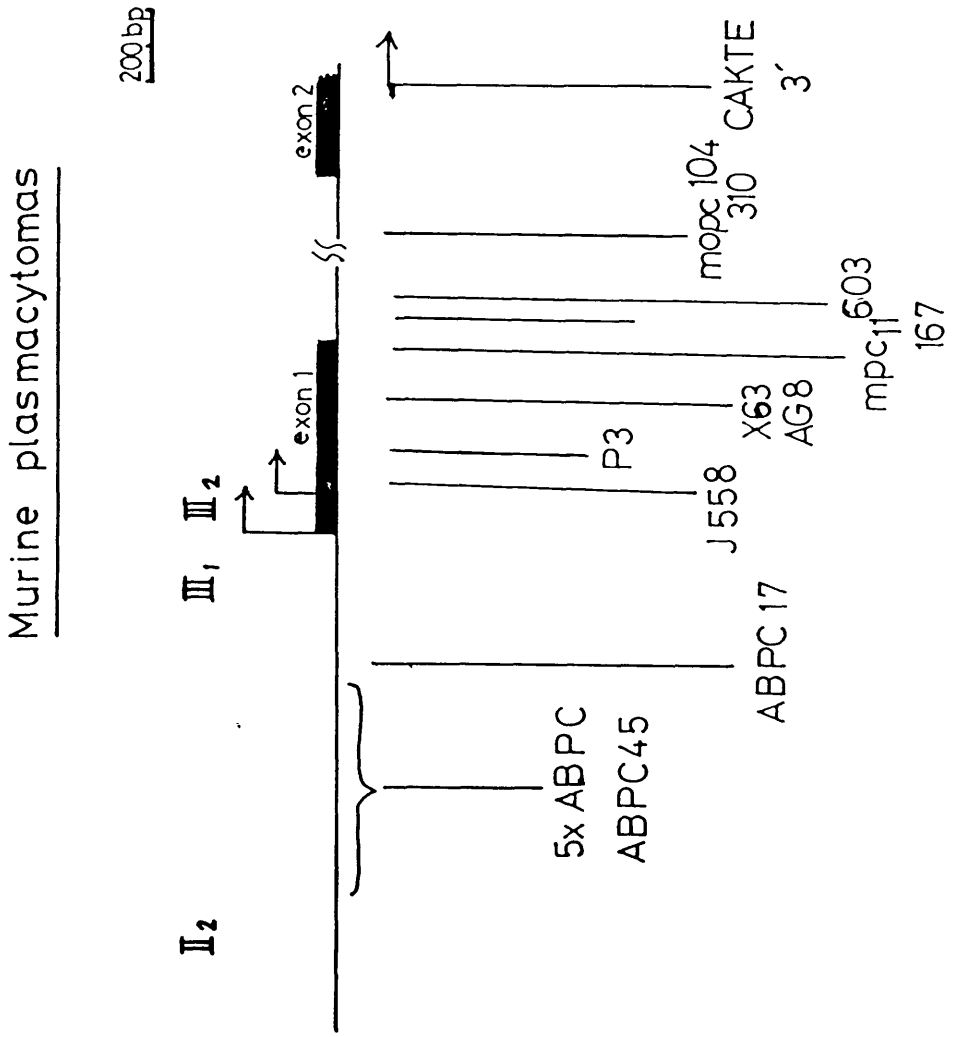


Figure 5.

Selected Murine Plasmacytoma Breakpoints.

Depicted in the diagram are some (not all) of the known murine plasmacytoma (MPC) breakpoint locations. The Roman numerals indicate the appropriate DNase 1 hypersensitive sites (Fahrlander, Piechaczyk and Marcu, 1985). The arrows indicate the two major murine transcription start sites (P1 and P2). The individual MPC cell lines are as described in Bernard et al, (1983); Cory, Gerondakis and Adams, (1983); Yang et al, (1985); Fahrlander, Piechazyk and Marcu, (1985); Calibi and Neuberger, (1985).



CHAPTER TWO

MATERIALS AND METHODS

2.1. Materials

Unless othrewise stated, all chemicals were obtained from BDH Chemicals, Poole, Dorset, UK, or Sigma Chemical Co, St Louis, Missourri, USA.

Biorad Laboratories, Richmond California, USA:
Ammonium persuphate, TEMED, protein assay kit II.

Pharmacia Fine Chemicals, Uppsala, Sweden:
Sephadex G50, polynucleotide kinase.

Vestric, Glasgow, Scotland, UK:
Ampicillin.

Fluka, Buchs, Switzerland:
Piperidine.

Aldridge Chemical CO Ltd, Gillingham, Dorset, UK:
Dimethylsulphate.

Schwarz Mann Biochemicals, Wembley, Middlesex, UK:
Urea.

BRL, Rockville, Maryland, USA:
Agarose, low melting point agarose, all restriction endonucleases, bacterial alkaline phosphatase, Klenow large fragment, T4 DNA polymerase.

Amersham International, Ailesbury, UK
^{14}C chloramphenicol.

Collaborative Research Inc, Massachusetts, USA:
Molecular linkers.

Boehringer Mannheim, Lewes, East Sussex, UK:
DNase I.

Gibco Ltd, Paisley, Scotland, UK:
Special liquid medium.

Fisons PLC, Loughborough, UK:
Ethyl acetate.

Difco Laboratories, Detroit, Mitchigan, USA:
Bactopeptone, yeast extract.

Northumbrian Biologicals Ltd, Cramlington, UK:
Foetal bovine serum, glutamine.

Camlab, Cambridge, UK:
Chromatography plates.

2.2. Construction of Recombinant Plasmids

All recombinants used in this thesis were constructed by manipulation of one of three existing plasmids; pMC41 (kindly given by T.Papas), pLW2 (kindly given by B.Clements) and pε0-1 (kindly given by J.Lang). The vector plasmids used were pUC12, pUC18 (Messing and Vieira, 1982) and pAT153 (Twigg and Sherrat, 1980).

2.2.1. Growth and Storage of Parental Plasmids

10 ml of L-broth (1% tryptone, 0.5% yeast extract, 1% NaCl) containing 100 μg/ml ampicillin was inoculated with host bacterium Esherichia coli K12 strain JM83 containing parental plasmiids (containg the amp^R gene) and incubated overnightat 37°C in an orbital shaker. 9 ml of the resulting culture were centrifuged at 2 G for 15 min and resuspended in 1 ml of 80% glycerol, 2% bactopeptone, therefore allowing storage at -20°C. The remaining 1 ml was then used to inoculate 1-2 volumes 500 ml L-broth, and incubated for 20 hr at 37°C in an orbital shaker. Bacterial cultures were subsequently harvested by centrifugation for 1 min at 8000 G and 4°C.

2.2.2. Extraction of Plasmid DNA

Plasmid DNA was isolated by modification of a previously described method (Clewell and Helinski, 1970). The pellets were resuspended in 10 ml of 25% (w/v) sucrose in 50 mM Tris-HCl pH 8.0, 5 ml 5 mg/ml (w/v) lysozyme and 5 ml 0.2 M EDTA pH 8.0 added respectively, and the bacterial lysate left for 30 min at room temperature. 600 μl of 10% Nonidet P-40 in 50 mM Tris-HCL pH 8.0, was then added and the tubes inverted until the lysate became viscous. The lysate was spun at 15000 G and 4°C for 30 min. The 20 ml supernatant was removed and extracted once with an equal volume of phenol, equilibrated in 50 mM Tris-HCl pH 8.0, and once with an equal volume of chloroform : isoamyl alcohol

(24:1 v/v). The supernatant was removed, 50 µl of pre-boiled (10 min) RNase A (1 mg/ml): RNase T1 (10000 unit/ml) added (Lonsdale et al, 1979), and incubated at 37°C for 30 min. The plasmid preparation was then extracted once with an equal volume of equilibrated phenol, and once with an equal volume chloroform : isoamyl alcohol (24:1 v/v). The purified plasmid was then precipitated for 90 min by the addition of 1/10 volume 7.5 M ammonium acetate pH 7.5, and an equal volume of isopropanol (Marmur, 1961). The DNA pellet was quickly washed with 100% ethanol, lyophilised and resuspended in 500-1000 µl of 0.1 M TrisHCl pH 8.0, 0.1 M EDTA. The resulting DNA preparation contains mostly supercoil plasmid, some relaxed circle plasmid and very slight traces of contaminating E.coli DNA.

2.2.3. Identification of Plasmids by Restriction Endonuclease Digestion and Gel Electrophoresis

All recombinant plasmids constructed were confirmed by extensive restriction enzyme digestion, and subsequent identification by either agarose or polyacrylamide gel electrophoresis.

(i) Restriction Endonuclease Digestion

An appropriate amount (1-3 µg) of plasmid DNA was digested by incubation for 3 hr at 37°C in the suppliers recommended buffers with spermidine added to 1 mM. Digestion was terminated by the additon of 1/10 volume gel loading dye containing, 10% w/v Ficoll 400, 0.1 M EDTA and 0.05 w/v bromophenol blue.

(ii) Agarose Gel Electrophoresis

DNA samples were electrophoresed on 0.5-2.0% horizontal gels (depending on expected molecular weights of DNA fragments generated) containing 0.5 µg/ml ethidium bromide at 2-5 V/cm for 1 hr-overnight, at room temperature

in 1 x TBE buffer (0.089 M Tris-borate, 0.089 M boric acid, 10 mM EDTA pH 8.0).

(iii) Polyacrylamide Gel Electrophoresis

Gels were formed by the addition of an appropriate volume of 29% w/v acrylamide : 1% w/v N,N'-methylene-bisacrylamide stock solution diluted as required by the addition of water, 1 x TBE buffer, and 500 µl 10% ammonium persulphate, and gel formation catalysed by the addition of 50 µl N,N,N',N'-tetramethylethylenediamine (TEMED) / 100 ml. DNA samples were electrophoresed on 4-8% gels in 1 x TBE buffer at 20 V/cm and room temperature. Gels were stained in 1 x TBE buffer containing 1.0 µg/ml ethidium bromide.

(iv) Gel Photography

Gels were photographed under UV irradiation (365 nm) using Polaroid type 55 or type 57 film.

2.2.4. Isolation of Purified Restriction Fragments

Restriction endonuclease digested DNA was electrophoresed on low melting point agarose (LMA) gels in 1 x TBE buffer in the presence of 0.3 µg/ml ethidium bromide. The appropriate DNA fragments were then excised and extracted by a modification of the method of Weislander, (1979). The excised LMA fragment was heated to 70°C for 5-10 min in 1 volume distilled water and then allowed to cool to 40°C. The DNA solution was extracted once with ice cold phenol, then twice with room temperature phenol, and the volume reduced to 200 µl with butan-2-ol. The DNA was then precipitated with 1/10 volume 7.5 M ammonium aetate pH 7.5, and 1 volume isopropanol at -70°C for 60 min. The purity of the DNA fragment was checked on an agarose gel (see section 2.2.3.).

2.2.5. Preparation of Vector DNA

Vector DNA was subjected to appropriate restriction enzyme digestion and electrophoresed on a low melting point agarose gel at 3 V/cm for 1 hr at 4°C, and purified as previously described. Because the colour selection technique (pUC vectors: Messing and Vieira, 1982) was used, there was no requirement to phosphatase the vector DNA. However, when very small fragments were to be cloned and the colour selection technique was not available, the vector was phosphatased. Rather than putting the digested vector DNA through a LMA gel, the digestion solution was diluted with one volume 50 mM TrisHCl pH 7.5, and incubated at 65°C for a further 2 hr with bacterial alkaline phosphatase (BAP) as recommended by the suppliers. The phosphatased DNA was extracted twice with phenol, once with chloroform and ethanol precpitated. Subsequently, the prepared DNA was ligated as described in section 2.2.7.

2.2.6. Use of Molecular Linkers

The presence of fortuitous restriction enzyme sites correctly positioned to allow constuction of particular recombinant plasmids is a rare event, and more often it is necessary to convert undesirable restriction sites to those required by the addition of synthetic molecular linkers (Heyneker et al, 1976). For this thesis only BamHI and HindIII linkers were used.

(i) Molecular Linker Addition to Blunt Ended DNA

0.5-2 µg purified DNA previously cut with restriction enzymes which generate blunt-ended molecules (HincII, RsaI, SmaI, ThaI and PvuII for this thesis) was ligated with 50 x excess of phosphorylated synthetic linker molecules under conditions descibed in section 2.1.7. Ligation was terminated by heating at 65°C for 10 min, and the DNA cleaved with the appropriate restriction endonuclease to

remove excess copies of the linker molecules. DNA was then separated from unused linkers by low melting point agarose gel electrophoresis, or by precipitation with 1/2 volume isopropanol at room temperature (a technique which allows selective precipitation of DNA while the molecular linkers remain in solution : J.Lang, personal communication).

(ii). Molecular Linker Addition to 'Sticky End' DNA (5'Overhang)

On DNA cleavage certain restriction enzymes produce a 5' overhang which must be removed before molecular linker addition is possible. This was done by the addition of the nucleotides complementary to those of the 5' overhang, catalysed by Klenow large fragment DNA polymerase. 2-3 μg of cleaved DNA was incubated for 1hr at 15°C in 50 μl reaction volume and 50 mM TrisHCl pH 7.8, 5 mM $MgCl_2$, 1 mM DTT, with complementary nucleotides to 20 μM and 5 units Klenow. The reaction was terminated by phenol extraction and the DNA was selectively precipitated with isopropanol at room temperature.

(iii) Molecular Linker Addition to Vector DNA

When molecular linkers were ligated to a restriction site within vector DNA molecules, removal of excess linker molecules by restriction enzyme recleavage was considered unnecessary. Under these circumstances, the ligated product was transfected directly into the host bacteria. Routinely, a high proportion (up to 80%) of the recombinant clones contained the linker insert, either produced by circularisation of the vector molecules *in vitro* after linker addition, or by homologous recombination across a ladder of linker inserts *in vivo*. This method has allowed the generation of recombinant clones containing analogous linker inserts at a number of sites.

2.2.7. Ligation of 'Target' and Vector DNA

Ligation was performed by a modification of a previously described method (Tanaka and Weisblum, 1975). 50-100 μg of vector DNA was mixed with a 3 molar excess of 'target' DNA (fragments to be recloned) and ligated in the presence of 10 mM TrisHCl pH 7.6, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP and 1-2 units T4 DNA ligase in a 10 μl reaction volume, for 5-16 hr at room temperature. For all such ligations the plasmid vector was pAT153 (a deletion mutant of pBR322; Twigg and Sherratt, 1980), or more often one of the pUC family of plasmids (Messing and Vieira, 1982).

2.2.8. Transformation of Competent Bacterial Host Cells

Host JM83 bacteria were transformed with ligated plasmid as previously described (Cohen et al, 1972). 100 ml of E.coli K12 strain JM83 bacterial cells were grown to an optical density of 0.4 at 600 nm, harvested by centrifugation for 1 min at 8000 G and 4°C, washed by resuspending in 1 volume of ice cold 10 mM TrisHCl pH 7.5, 100 mM $MgCl_2$, and swirled gently for 5 min, before harvesting as before. The bacterial pellet was resuspended in 1 volume 10 mM TrisHCl pH 7.5, 100 mM $CaCl_2$, and left on ice for 1 hr. The cells were harvested as before. The cells were then resuspended in 1/50 volume 10 mM TrisHCl pH 7.5, 100 mM $CaCl_2$, and if kept at 4°C were found to be viable for up to 5 days (with the highest effiency after 12-24 hr). Ligated plasmids were then added to 200 μl of competent cells and left on ice for 30 min, to allow the DNA to come into contact with the bacterial cells. To facilitate entry of the DNA into the bacteria, the cell-DNA mixture was heat-shocked at 37°C for 5 min. The cell suspension was incubated with 800 μl of L-broth for 90 min at 37°C. 300 μl aliquots were then spread over 90 mm L-broth 1.5% agar plates containg 100 μg/ml ampicilin, 25μg/ml X-gal (5-bromo-4-chloro-3-indolyl-β-galactoside : chromogenic substrate for β-galactosidase) and incubated overnight at

37°C.

2.2.9. Screening Bacterial Colonies for Specific Recombinants

Ampicillin resistant colonies were screened for the desired recombinant plasmids by 'mini-preps' analysis. This method was thought to be more efficient than filter hybridisation.

(i) Mini Preps

12-48 single bacterial colonies were picked from the original plates with sterile cocktail-sticks, and shaken into 1ml of L-broth containing 100 μg/ml ampicillin, and grown overnight at 37°C in an orbital shaker. The cells were harvested by spinning at 16000 G for 20 sec. The pellets were resuspended in 75 μl STET buffer (8% w/v sucrose, 50 mM TrisHCl pH 8.0, 50 mM EDTA, 5% v/v Triton X-100) with fresh lysozyme added to 1 mg/ml. This solution was placed in a boiling water bath for 50 sec. The bacterial lysate was then centrifuged at 16000 G for 8 min. The supernatant was recovered, and plasmid DNA precipitated by addition of 1 volume isopropanol on dry ice for 5 min. The DNA was recovered by centrifugation at 16000 G for 10 min. The pellets were then lyophilised for 1 hr, and resuspended in 50 μl of water. This method yields about 6 μg of plasmid DNA.

(ii). Identification of Correct Recombinant by Gel Electrophoresis

Recombinant clones identified by the 'mini-prep' method were then digested with appropriate restriction enzymes and electrophoresed on agarose gels at 4.5 V/cm for 1-3 hr, or after a 30 min treatment with RNase A/T1 (see section 2.2.2.), followed by phenol extraction on polyacrylamide gels at 20 V/cm for 1-2 hr.

2.3. Transient Expression Assays

In the studies presented within this thesis various DNA domains were used to drive the expression of a reporter gene (CAT) in transient transfection studies.

2.3.1. Recipient Cells

All the cell lines described were obtained from the Beatson Institute reference stocks. During these studies the following cells were used.

(a). LaTK$^-$ cells, (Kit et al,1963); a thymidine kinase negative varient of the mouse fibroblast derived cell line (L cells).

(b). HeLa cells, (Scherer, Syverton and Gey, 1953); derived from a human epithelial carcinoma.

(c). K-562 cells, (Lozzio and Lozzio, 1975); derived from a human erythroid leukaemia, the actual varient used was selected for its increased adhesive property (D.Spandidos).

(d). NIH-3T3 cells (Jainchill, Aaronson, and Todaro, 1969); a fibroblastoid contact-inhibited varient derived from a NIH Swiss mouse embryo.

(e). GM468 cells (NIGMS, Human Genetic Mutant Cell Repository); derived from normal human fibroblasts.

(f). MRC-5 cells (Jacobs, Jones and Baille, 1970); a human primary fibroblast culture, capable of attaining 42-46 population doublings before onset of decline in proliferation.

2.3.2. Cell Culture

Cells were grown in Special Liquid Medium (SLM) supplemented with glutamine to 4 mM and foetal calf serum to 13% at 37°C, in 10% CO_2/air. Cells were refed every 2 days to ensure optimal health and were passaged by sequential washes in phosphate buffered saline (PBS), 10% trypsin/ citruline and finally resuspended in an appropriate volume of medium. Cell numbers were measured in a coulter counter.

Cell stocks were made by resuspending 10^6 cells in medium plus 10% glycerol, and immediatly placing them at -70°C, with final storage being under liquid nitrogen.

2.3.3. DNA Mediated Gene Transfer

The introduction of CAT recombinant plasmids into mammalian cells in culture (DNA mediated gene transfer) was performed by the calcium phosphate co-precipitation technique (Graham and Van Der Eb, 1973), using a modification of the method described by Wigler et al, (1977). The levels of CAT expressed was then measured by enzyme assay shortly after transfection (i.e. transient expression).

(i). Estimation of DNA Concentration

The DNA concentration was measured by obtaining an optical density reading. This value was confirmed by a visual estimation of DNA concentration. Uncut plasmid was electrophoresed on 0.6% agarose gels and the DNA 'eyeballed'. By using a range of sample amounts and 1 μg of commercial pUC12 DNA an accurate value was obtained. This also allowed the DNA purity to be observed. Such a technique ensures molar equivalents of each recombinant plasmid are used when introducing DNA into recipient cells.

(ii). Calcium Phosphate Co-Precipitation of DNA

Recombinant plasmid DNA and commercial salmon sperm DNA as carrier (except in the competition experiments, where pUC12 plasmid DNA was used), were diluted to a final concentration of 20 μg/ml by the addition of an appropriate volume of water HEPES bufered saline (5 g/L HEPES [N- 2-hydroxyethyl-piperazine- N'- 2- ethenesulfonic acid], 8 g/L NaCL, 0.1 g/L Na_2HPO_4 pH 6.95, and subsequently $CaCL_2$ to 125 mM. The solution was then left to stand for 30 min at room temperature, to allow a calcium-phosphate precipitate to

form and concomitantly co-precipitate the DNA. 1 ml of the precipitate (in suspension) was then added, by dropping, to approximately 10^6 recipient cells (5 x 10^6 cells when using K-562 cells), exponentially growing at 37°C in tissue culture flasks of surface area 25 cm^2, in 9 ml of culture medium.

(iii). Harvesting of Cells

Cells were routinely harvested 44-48 hr after transfection, basically by the method of Gorman, Moffat and Howard, (1982). The cells were washed with PBS carefully, and then allowed to sit at room temperature in 1 ml TEN buffer (0.04 M TrisHCl pH 7.4, 1 mM EDTA, 0.15 M NaCl). The cells were scraped off and pelleted (2 min in microfuge at room temperature). Pellets were dispersed by vortexing in 100 µl 0.25 M TrisHCl pH 7.8, and cells disrupted by 3 cycles of freeze-thaw. The debris were spun down (8 min in microfuge) to give the cell extract (supernatant) ready for assaying. The cell extract can be stored at -20°C at this stage if required.

2.3.4. Chloramphenicol Acetyltransferase (CAT) Assay

Obviously, the amount of cell extract and the time of enzyme assay depends on the promoter domain to be assayed, the amount of DNA transfected and the cell type used as recipient. However these variables were standardised as discused in the result section of this thesis.

(i). Assay of Transient Expression of CAT Enzyme Levels

All reactions were done in 90 µl final volume containing 1 µl ^{14}C Chloramphenicol, 5 µl 50 mM acetyl coenzyme A, 20-84 µl cell extract (made up to volume as required with 0.25 M TrisHCl, pH 7.8). Typical reaction times were 30-90 min at 37°C. The chloramphenicol and its acetylated forms were extracted with 600 µl ethyl acetate by

shaking for 30 sec, and separated from the cell extract by spining in a microfuge for 30 sec. The top organic phase was removed (avoiding interface) and dried down using a vacuum aspirator (about 1hr).

(ii). Thin Layer Chromatography of Labelled Chloramphenicol

The dried down chloramphenicol samples were resuspended in 20 µl ethyl acetate, and spotted onto silica gel thin layer chromatography (TLC) plates, trying to keep the spot size as small as possible. The TLC plates were developed in chloroform : methanol (95:5) ascending.

(iii). Representation and Quantification of Results

The developed TLC plates were exposed to XAR-5 Kodak film overnight at -70°C. These autoradiographs gave a good visual qualitative representation of the result (see figure 6). Quantitative results were obtained by cutting out and subsequently, counting the respective acetylated chloramphenicol spots in a LKB scintillation counter.

2.4. In vitro 'Footprint' Analysis

This technique allows the 'footprint' of a protien to be localised on a fragment of DNA in vitro. By using purified protein extracts 'footprints' of specific proteins can be identified. The purified protein fractions used in these studies were a kind gift from R.Tjian (Kadonaga, Jones and Tjian, 1986). Basically, nuclei are isolated and a nuclear extract made. This extract is subsequently run over a series of differing affinity columms (the final enrichment step being a DNA affinity column). By this method a partially purified protein extract, and two highly enriched extracts (for Sp1 and CTF) were obtained (see figure 7).

2.4.1. ^{32}P-Endlabelling of DNA 5' Termini

Purified restriction fragments were labelled at their 5' termini as described by Maxam and Gilbert, (1980). Approximately 4 µg of purified plasmid was linearlised with a selected restriction enzyme and treated with BAP (see section 2.2.5.). The DNA was ethanol precipitated and washed with 95% v/v ethanol, lyophilised and dissolved in 50 mM TrisHCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM EDTA, 0.1 mM spermidine. 5 µl α-^{32}P-ATP (3000Ci/mmol) was added and the reaction catalysed with 10-20 units of T4 polynucleotide kinase in a 12.5 µl reaction volume for 30 min at 37°C. The reaction was then terminated by phenol extraction, the DNA precipitated by ethanol, and subsequently digested with a further restriction enzyme to generate restriction fragments 5' labelled at one end. The fragments were then separated by electrophoresis on a polyacrylamide gel at 20 V/cm. The gels were then autoradiographed while wet using radioactive ink as a marker and the bands removed to Eppendorf tubes.

2.4.2. Purification of Labelled Fragment

The extraction of labelled DNA was by isotachophoresis (Ofverstedt et al, 1984). The polyacrylamide gel fragments were stained in 5% phenol red and placed on top of a sephadex G50 column (sealed with dialysis membrane), equilibrated with 40 mM TrisHCl pH 7.5, and overlayed with 0.1 M 6-aminocaproic acid. A current of 2-6 mA was applied until sharp focusing of the indicator dye is achieved (the DNA runs just ahead of the dye). The DNA was then collected in fractions by removal of the dialysis membrane, pooled and ethanol precipitated on dry ice for 15 min. After lyophilising the labelled, purified DNA was resuspended in an appropriate volume of water.

2.4.3. G Residue Sequencing

G-track markers were made by using the DNA sequencing

protocol described by Maxam and Gilbert, (1980). 200 µl of DMS buffer (50 mM sodium cacodylate pH8.0, 10 mM $MgCl_2$, 1 mM EDTA) and 1 µl calf thymus carrier DNA (1 mg/ml) were added to 5 µl ^{32}P-DNA chilled on ice, 1 µl dimethylsulphate (DMS) added, and incubated at 20°C for 10 min. The reaction was then chilled, 50 µl DMS stop (1.5 M sodium acetate pH 7.0, 1 M 2-mercaptoethanol, 100 µg/ml tRNA) added, and the DNA precipitated by the addition of 750 µl ethanol, and transfer to a dry-ice-ethanol bath for 75 min. The DNA was recovered by centrifugation at 4°C for 5 min, in an Eppendorf centrifuge. The pellet was redissolved in 250 µl of 0.3 M sodium acetate pH 6.5, and ethanol precipitated as before, washed with 750 µl 95% ethanol and lyophilised. 100 µl 1 M piperidine was then added, and the tubes heated to 90°C for 30 min, and the solution then frozen on dry ice and lyophilised (5 hr-overnight). 10 µl water was added, lyophilised as before, this step repeated, and the pellet dissolved in 6-20 µl formamide dyes (80% v/v formamide, 1 x TBE, 0.1% w/v xylene cyanol FF, 0.1% w/v bromophenol blue), heated for 1 min, and quick chilled on ice.

2.4.4. DNase I Digestion

End-labelled DNA restriction fragments (10-30 cpm), 10 µl 10% polyvinyl alcohol, 1 µl 1 mg/ml calf thymus DNA (with $10A_{260nm}$ unit/ml poly d(I-C) unless purified protein extracts were used) made up to 25 µl with water, were incubated with protein extract (0-25 µg), and adjusted to 25 µl volume with Z'Buffer (25 mM HEPES (K+) pH 7.8, 12.5 mM $MgCl_2$, 1 mM DTT, 20% v/v glycerol, 0.1% v/v Nonidet P-40 plus 0.1 M KCl) for 15 min, on ice. The sample was left at room temperature for 1 min, then 50 µl (10 mM $MgCl_2$, 5 mM $CaCl_2$) was added, and mixed at room temperature for 1 min. DNase I (0.5 µl; 2.5 mg/ml) was added to 1 ml ice cold water, and dilutions (1:1000 or 1:2000) were added to the DNA-protein solution, mixed and left at room temperature for 1 min. The reaction was terminated with 90 µl Stop Solution (20 mM EDTA pH 8.0, 1% w/v SDS, 0.2 M NaCl, 250 µg/ml tRNA).

The samples were then extracted first with phenol, then chloroform : isoamyl alcohol, and then ethanol precipitated on dry ice for 10 min. After a 75% ethanol wash and lyophilising, the pellet was dissolved in 7.5 µl formamide dye, boiled for 3.5 min, and chilled on ice ready to be loaded onto the gel.

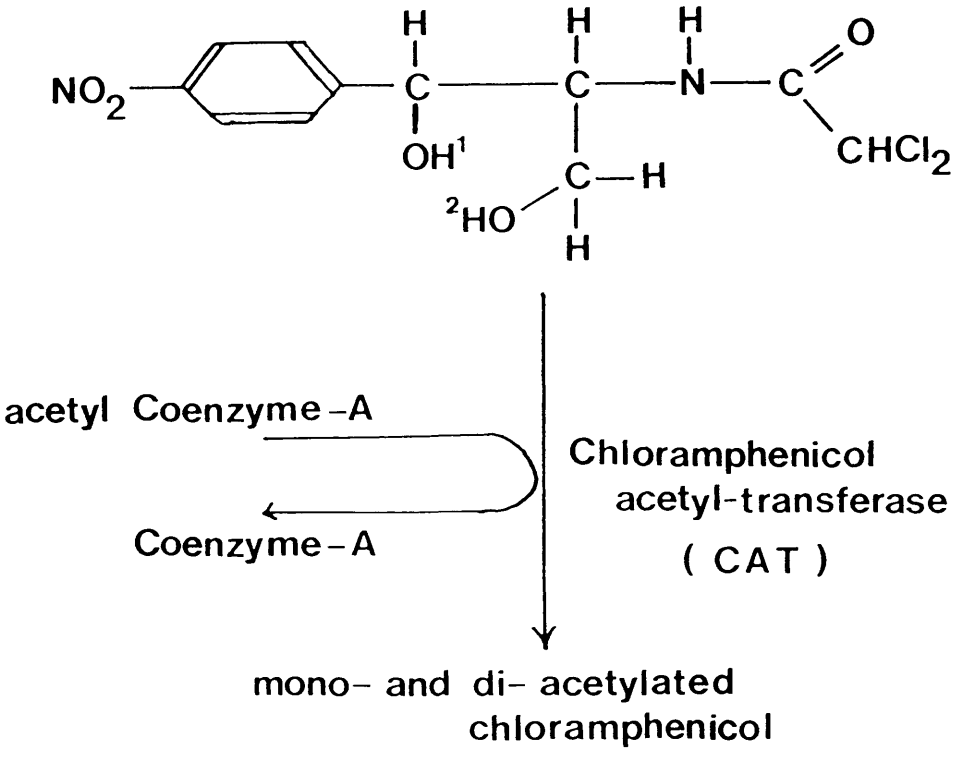
2.4.5. Polyacrylamide Gel Electrophoresis

The footprinted samples, plus G-track marker, were then loaded onto a 6% or 12% polyacrylamide sequencing gel, made by dilution of stock deionised 20% acrylamide solution (29:1 w/w acrylamide: NN'-methylenebisacrylamide) with 10 M urea in 1 x TBE and 0.07% w/v ammonium persulphate, and polymerisation was catalysed by the addition of 27 µl TEMED. The gels were pre-electrophoresed for 1-2 hr, at 35 W, and subsequently electrophoresed at 35 W until optimal positioning was acheived. These gels were autoradiographed by exposure overnight to XAR-5 Kodak film at -70°C.

Figure 6.

The CAT Assay

I.



II. Autoradiograph of CAT Result.

Key.

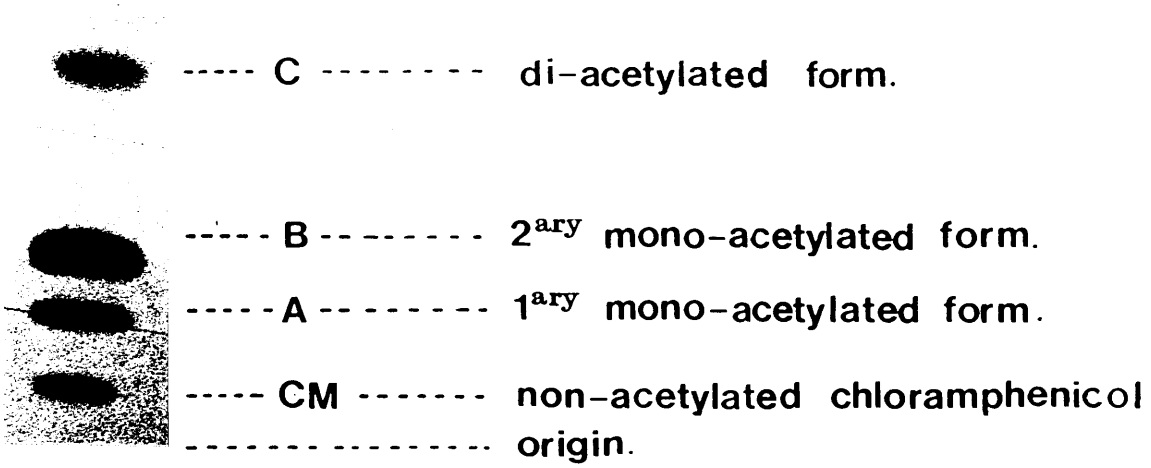
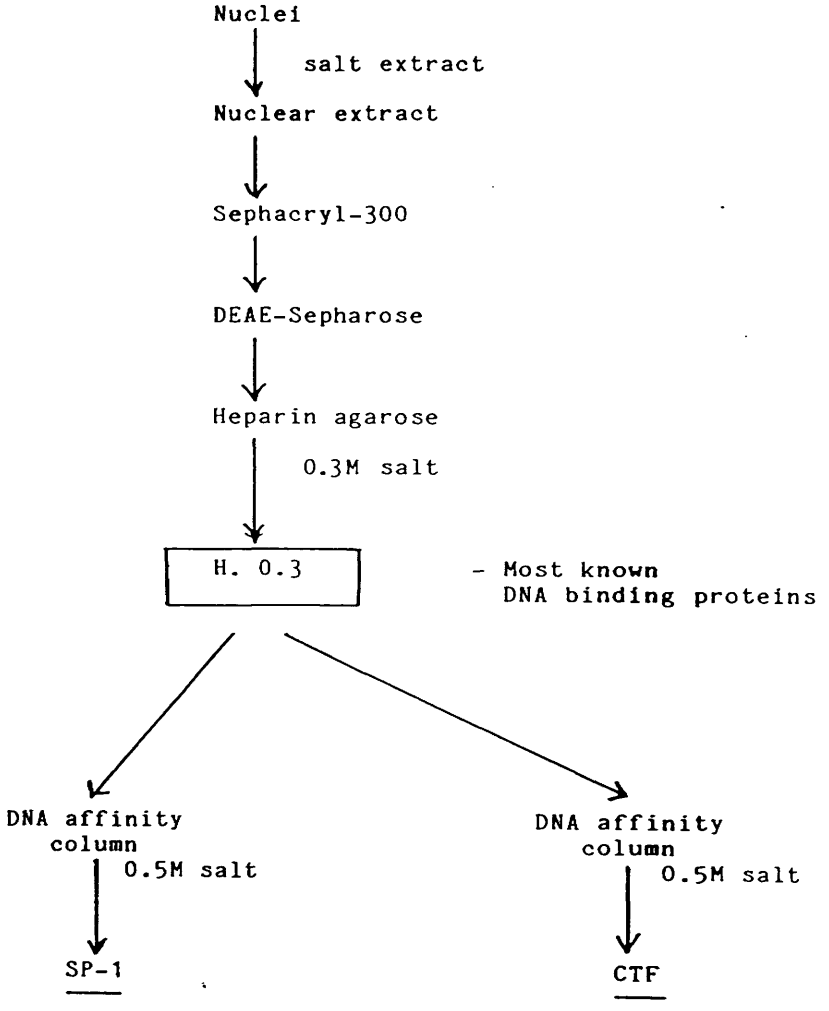


Figure 7.

Basic Protocol used to Purify Protein Extracts for *in vitro* DNase 1 Footprinting.

The respective sequences of the two consensus elements are shown (for a more detailed description see sections 2.4, 3.4.6., and Kadonaga, Jones and Tjian, 1986).



G/T G G G C G G G/A A G/C T . . . C C A A T . . . .

(A)(A) (A) (T) (G)

(C)

CHAPTER THREE

RESULTS

3.1. Optimisation of CAT Assay

3.1.1. The Chloramphenicol Acetyltransferase Assay

Although RNA levels represent the most definitive gauge of promoter activity, quantification of RNA is tedious and it is often difficult to obtain accurate measurements unless the promoter being analysed is particularly strong or amplified. It is often preferable to determine the function of a promoter by joining it to a second gene segment (which lacks its own promoter) coding for a readily assayable protein, the reporter gene. In the CAT assay (Gorman, Moffat and Howard, 1982) the bacterial chloramphenicl acetyltransferase (CAT) gene is used, with the enzymatic function of CAT being both quickly and precisely assayable.

Two assays for CAT activity have been developed. The first, a spectrophotometric assay, unfortunately depends on the cell type used and as such requires a control to be run for each cell type used. A better assay (Gorman, Moffat and Howard, 1982) involves the measurement of the conversion of ^{14}C-chloramphenicol to its acetylated forms by thin layer chromatography (see figure 6). The great advantage of this assay is that any cell type can be used, the only restraints being those imposed by the specific DNA sequences being analysed. It is now the most common transient functional assay used to analyse gene expression.

This assay shows a high level of specificity due to the CAT gene being absent from all mammalian cells looked at, (metabolism of chloramphenicol being via glucosyl transferase). Thus there is no endogenous background CAT activity. In addition, there is no interference from other endogenous enzymatic activities which could compete for the substrate or cofactors. The assay is rapid, sensitive and reproducible, as well as requiring only a low grade radioisotope. Finally, quantification of the result can be easily and accurately carried out in a scintillation

counter. These advantages make the CAT gene marker, at the moment, the most convenient developed assay for transient expression studies. There are other transient assay systems available, eg.β-galactosidase, but they are either more cumbersome or less sensitive than the CAT assay.

3.1.2. Standard Positive CAT Construct

Figure 8 shows the Key for all the figures in this section.

Before undertaking any functional experiments on the c-myc-CAT plasmid recombinants, characterisation of the obtainable expression levels is required. For this, and as a positive control in subsequent experiments, a strong promoter-driven CAT recombinant was kindly obtained from B. Clements, pLW2 (Gaffney et al, 1985). The 210bp BamHI fragment of pLW2 (see figure 10) contains the promoter domain (TATA-consensus and two GC-elements) plus a small part of the 5' mRNA leader sequence from the HSV-2 immediate-early 4/5 gene (HSV-IE). This promoter directs a high level of activity when compared to other promoters, e.g. HSV-tk (J. Lang, personal communication) and ε-globin (this thesis). Although an enhancer element has been described which controls this gene it is not present in pLW2 (Gaffney et al, 1985). The first ATG initiation codon lies within the CAT gene.

Transcription 3' end processing is regulated by a 100bp SmaI/XbaI fragment derived from the 3' terminus of the HSV-2 IE gene-5 (Gaffeny et al, 1985), which makes this construct differ from the more commonly used prototype CAT vector, pSV2-CAT (Gorman, Moffat and Howard, 1982). The HSV-derived termination sequence appears to allow a slighly higher level of activity, in mouse fibroblasts, when compared to the SV-40 derived recombinant (unpublished results).

3.1.3. Generation of CAT Vector

By deleting the 210bp BamHI fragment from pLW2 (Gaffney et al, 1985), the CAT vector pB9 (containing no obvious promoter or cap site), was generated (see figure 11). This plasmid was used as the basic vector for all the recombinants described in this thesis, and as a negative control in all the experiments.

3.1.4. Restriction Analysis of Basic CAT Recombinant

HindIII or EcoRI digestion of both pB9 and plW2 results in a linear plasmid, of respectively 3.6 and 3.8Kb as seen in figures 12. To confirm that the 210bp HSV-IE promoter domain had been deleted in the CAT vector, pB9 was digested with HinfI (see figure 12). As expected the pUC-derived fragments of 1413, 517, 396, 75 and 65bps are seen. In addition, the 214bp pUC fragment containing the polylinker domain is replaced by a 301bp (HSV-IE termination domain) and a 793bp (most of CAT structural gene), due to the presence of a HinfI site being present within the CAT sequence. PstI digestion of pLW2 gives a 1000bp fragment, containing the 780bp CAT gene and the 210bp HSV-IE promoter domain (see figure 12).

3.1.5. Optimisation of Expression Time

The activity time course observed when cells were harvested at intervals of 8 hours from when the DNA-calcium-phosphate precipitate was washed off is shown in figure 13. In each experiment the positive control pLW2 (5μg) was used. For LaTK$^-$cells CAT enzyme levels start to decrease after 48 hours, while a plateau of activity is observed for NIH-3T3 cells with a slight maximum occuring at 40 hours. This time course is in agreement with results using the *tk* enzyme as a marker (Spandidos and Wilkie, 1984a; Lang, 1985). Hela cells (result not shown) show maximal activity at 40 hours in this assay. Since the

absolute activity levels observed in LaTK$^-$ cells was far greater than those observed in NIH-3T3 cells, the former were the cells of choice for the majority of experiments described in this thesis. NIH-3T3 cells have been found to be poor transient transfection recipients by other workers (Alwine, 1985; see section 3.1.8.).

3.1.6. Tranfected DNA Titration

To investigate the effect of varying the amount of transfected DNA, increasing amounts of pLW2 DNA (the positive control) were transfected into LaTK$^-$cells. As shown in figure 6, the acetylated chloramphenicol has 2 forms, the mono-acetylated (primary and secondary) and di-acetylated, both of which are plotted along with total activity in figure 14. Clearly, although at higher DNA transfected amounts the production rate of the diacetylated (C) form is in the linear part of the curve with respect to increasing DNA amounts, at the lower DNA amounts it is not an accurate representation of CAT activity. Since in the studies described in this thesis repression of an active state is analysed, many results would presumably involve relatively low CAT levels, the measurement of the diacetylated form alone would be an inappropiate representation of enzyme activity (expression). Summation of the two mono-acetylated forms (A + B) seems unreliable, since at high rates of conversion the di-acetylated form was produced from the mono-acetylated forms in preference to mono-acetylation. However, total counts (A + B + C) gave a reproducible representation of CAT activity, above 1 ug of transfected DNA. In addition, total counts gave at the lower DNA amounts a high enough value to allow easy and reproducible values to be obtained.

Since the absolute value between experiments using the same recipient cell line vary, results are expressed as averages of total counts or as relative ratios (from a standard positive). Experiments with less than 15% variation

in duplicate samples from control plasmids were deemed representative and tabulated. In this way, the different cell lines can be qualitatively compared for relative differences between test plasmids. The variation observed between experiments is probably due to a complex interaction of many factors, (e.g. subtle changes in temperature or pH, slight differences between various medium preparations) which compoundly affect the growth of the recipient cell. Thus this variation probably reflects the growth state of the recipient cell at time of transfection.

In the experiments described here, the minimum amount of DNA to be transfected was taken to be 1ug, since below this amount non-reproducible results were obtained. In addition, when more than 10μg were transfected only very slight increases (if any) in activity were observed, probably indicating an "assay-saturation" (Weintraub, Cheng and Conrad, 1986; and see figure 30). Whether this saturation is relative to DNA uptake by the cells, or to the cells ability to express (titration of cellular pool of transcription factors) the exogenously introduced DNA is not known. Although not seen in this experiment, when DNA amounts of greater than 10μg were used a near plateau of activity was observed (see section 3.2.6.). Thus the linear part of the expression curve for LaTK$^-$cells in these experiments was, taken to be, between 1μg and 10μg. To ensure equal numbers of available plasmid molecules equimolar-amounts (standardising all plasmids to the 3.3Kb pB9,i.e. pB9=1) were used in all experiments.

The amount of DNA transfected could play a crucial role in the outcome of an experiment. Therefore experiments were carried out using a DNA titration wherever possible, thus hopefully allowing any biologically relevant variation in activity due to DNA concentration to manifest itself. In experiments where only a single DNA amount was used, a low amount was considered the most appropriate to avoid the possible saturation effects in the assay.

3.1.7. Optimisation of CAT Assay Parameters

To allow optimisation of the CAT enzyme activity assay LaTK$^-$ and NIH-3T3 cells were transfected with 5μg of pLW2 (the positive control), and the cell extract aliquoted accordingly such that the effect of varying the ammount of extract, or the time of incubation in the enzyme assay could be examined (see figure 15). The cell extract containing CAT enzyme activity was stable at -20°C for several months. Again the graphs obtained for the various acetylated forms were compared (results not shown) and the previous conclusions were confirmed (see section 3.1.6.).

An assay time course of 20 to 120 minutes for cell extracts from both LaTK$^-$ and NIH-3T3 cells and the resultant CAT activity plotted in figure 15. The assay was linear between 60 and 120 minutes for both LaTK$^-$ and NIH-3T3 cells, and longer if additional coenzyme A was added at 100 minutes (data not shown). The tail-off after 120 minutes is probably due to substrate exhaustion, but the presence or accumulation of inhibitors cannot be excluded. It was noted that reproducible results were not obtained if a time point of less than 30 minutes was taken.

Even though the time of assay incubation was doubled (i.e. betwwen 40 and 80 minutes) the resultant CAT activity did not. This may reflect the effect of progressive substrate exhaustion. To exclude any possible discrepancies between results due to this effect, for the studies presented within this thesis the CAT assay was routinely run for 60 minutes.

A linear graph (see figure 15) was obtained when the amount of extract was varied for LaTK$^-$ cells. In these experiments the observed CAT activity was always substantially lower for NIH-3T3 cells than for LaTK$^-$ cells. In addition, the curve seen for NIH-3T3 cells was linear but even though the amount of cell extract used was quadrupled,

the corresponding CAT enzyme activity was only doubled. This could be due to an endogenous enzyme (which is present at high levels in NIH-3T3 cells), which is utilising the substrate. However, since no spurious chloramphenicol spots were observed on the chromatogram (although this does not exclude the possibility), it is unlikely that such an enzyme exists. Chloramphenicol is broken down in mammalian cells by glucosyl transferase, but even if no spurious spots were produced during this reaction a decrease in substrate amount would (but is not) be seen. Possibly the most likely explanation for the observed NIH-3T3 cells activity curve may be the presence of an inhibitor. A preliminary experiment involving the mixing of NIH-3T3 and LaTK$^-$ cell extracts indicates that such an inhibitor may exist within the NIH-3T3 cell extract (unpublished results). Mixing cell extracts has indicated that Friend erythroleukaemia cells may also contain an inhibitor (G.Sibbet, personal communication). Due to the low transfection efficiency and possible assay inhibitor, NIH-3T3 cells were only rarely used as transfection recipients in the work presented in this thesis.

3.1.8. Reproducibility of CAT Assay

The fact that with at least 8 separate transfections of pB12 and pB14 during these studies (see section 3.2.3), using several plasmid preparations, similar results (comparable ratios) were obtained for LaTK$^-$ cells, implies reproducibility within this assay. Several additional experiments confirmed this reproducibility. In the early experiments, quadruplet samples were assayed, with only a 15% variation between samples being observed. To ensure that this was not due to cell harvesting variations (i.e. total cell number harvested), the protein extract amount was quantified with only a 11% variation being present (see table 1). However, this variation was not directly accountable for the 15% variation in activity, as assays performed after protein equalisation still showed variation.

When duplicate assays were performed on the respective halves of the same cell extract there was still a slight variation observed (12%). Thus an approximately 10-15% error would appear to be inherent in this assay which may be due to slight variations in the individual assay sample handling (i.e. human error, loss during ethyl acetate extraction).

Figure 8.

Restriction Enzyme Maps for Recombinant Plasmids.

Maps represent recombinant plasmids characterised by restriction endonuclease digestion in the relevant sections of this thesis. The specific gel photograph corresponding to each restriction map is indicated by reference to the appropriate figure. The type of gel (and percentage) is shown along side each photograph. Only the relevant restriction sites are shown (see sections 3.2.2.,3.3.2.,3.4.2.,3.5.2. and 3.6.2.). HindIII and Hinf I pAT153 were the size markers used. Fragment size is shown to the nearest 10bp only, since due to molecular linker additions, accurate size can only be determined by DNA sequencing. (The numbers inside the plasmid circle refer to the HinfI fragment sizes only). P1 and P2 refer to the two major human c-myc transcription start sites. The orientation of the respective fragments of the recombinants is indicated by an internal arrow. Roman numerals are used to represent the location of the in vivo DNase 1 hypersensitive sites (Siebenlist et al, 1984). The origin of each fragment is illustrated as described in the key below:

Key to Plasmid Fragment Source.

Vector sequences (pUC12,pUC18 or pAT153 as indicated in the text),

human c-myc sequences,

HSV-IE derived sequences,

human ε-globin derived sequences,

bacterial CAT gene sequences,

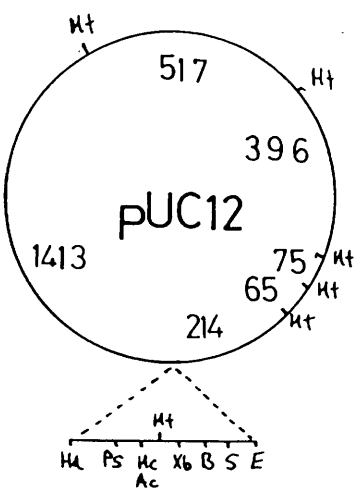
Restriction Enzyme Key:

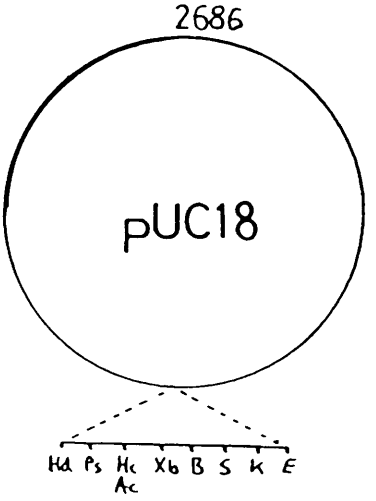
Ac	Acc I	Ps	Pst I
Av	Ava I	Pv	Pvu II
B	BamH I	R	Rsa I
C	Cla I	S	Sma I
E	EcoR I	Sa	Sau 3A
Hc	Hinc II	T	ThaI
Hd	Hind III	Xo	Xho I
Hf	Hinf I	Xb	Xba I
K	Kpn I		

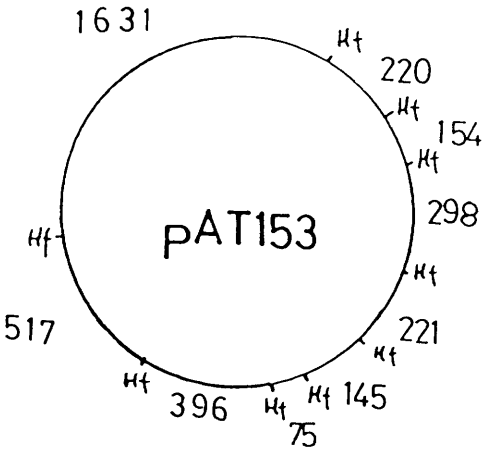
Linker addition. CCAAGCTTGG = HindIII
CCGGATCCGG = BamHI

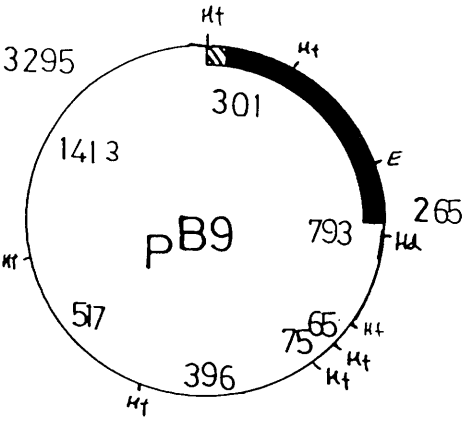
An unbound copy of this figure can be found in the sleeve of this thesis.

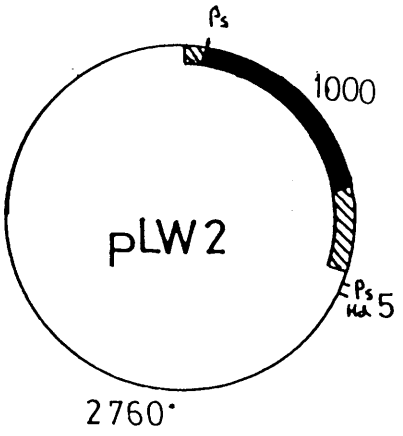
Figure 8a


Ht
517
Ht
396
pUC12
1413
75
Ht
65
Ht
214
Ht
Ht
Hd Ps Hc Xb B S E
Ac


2686
pUC18
Hd Ps Hc Xb B S K E
Ac


1631
Ht
220
Ht
154
Ht
298
pAT153
Ht
Ht
517
221
Ht
Ht
396
Ht
Ht
145
75


Ht
Ht
3295
301
E
1413
265
Hd
pB9
793
Ht
517
Ht
65
75
Ht
Ht
396
Ht


Ps
1000
pLW2
Ps
Hd
5
2760

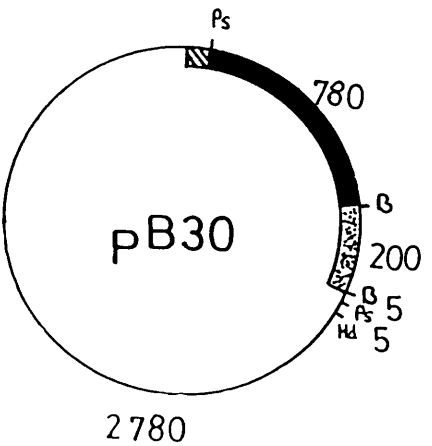
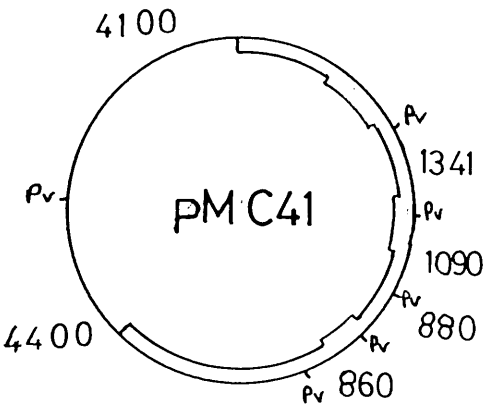
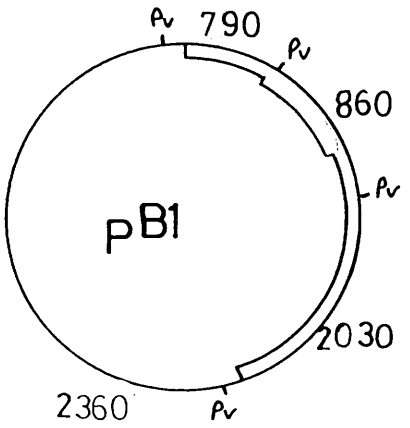
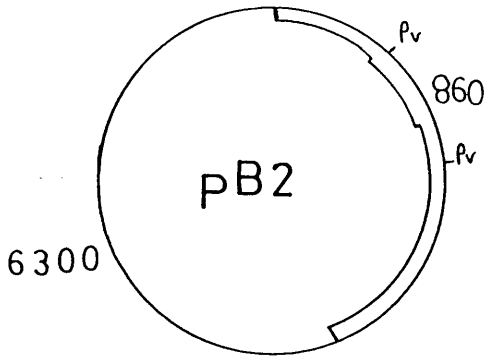
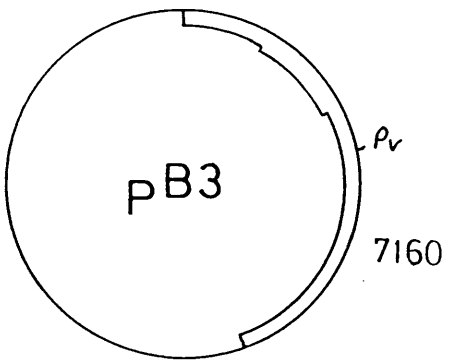

Ps
780
B
pB30
200
B
Ps
5
Hd
5
2780

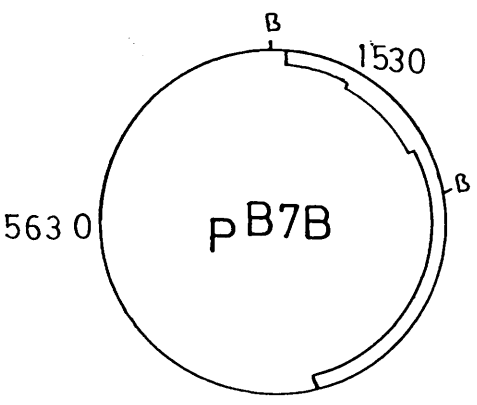
Figure 8b


4100
Pv
1341
Pv
1090
880
Pv
860
Pv
4400
Pv
pMC41


Pv
790
Pv
860
Pv
2030
Pv
2360
pB1


Pv
860
Pv
6300
pB2


Pv
7160
pB3


B
1530
B
5630
pB7B

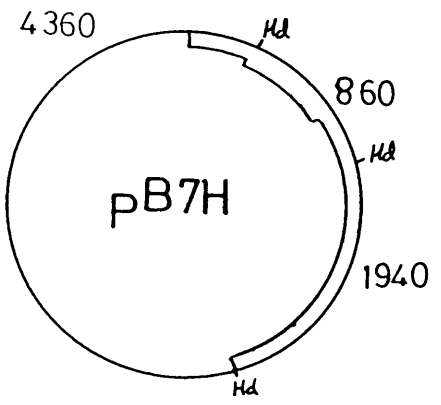

4360
Hd
860
Hd
1940
Hd
pB7H

Figure 8c

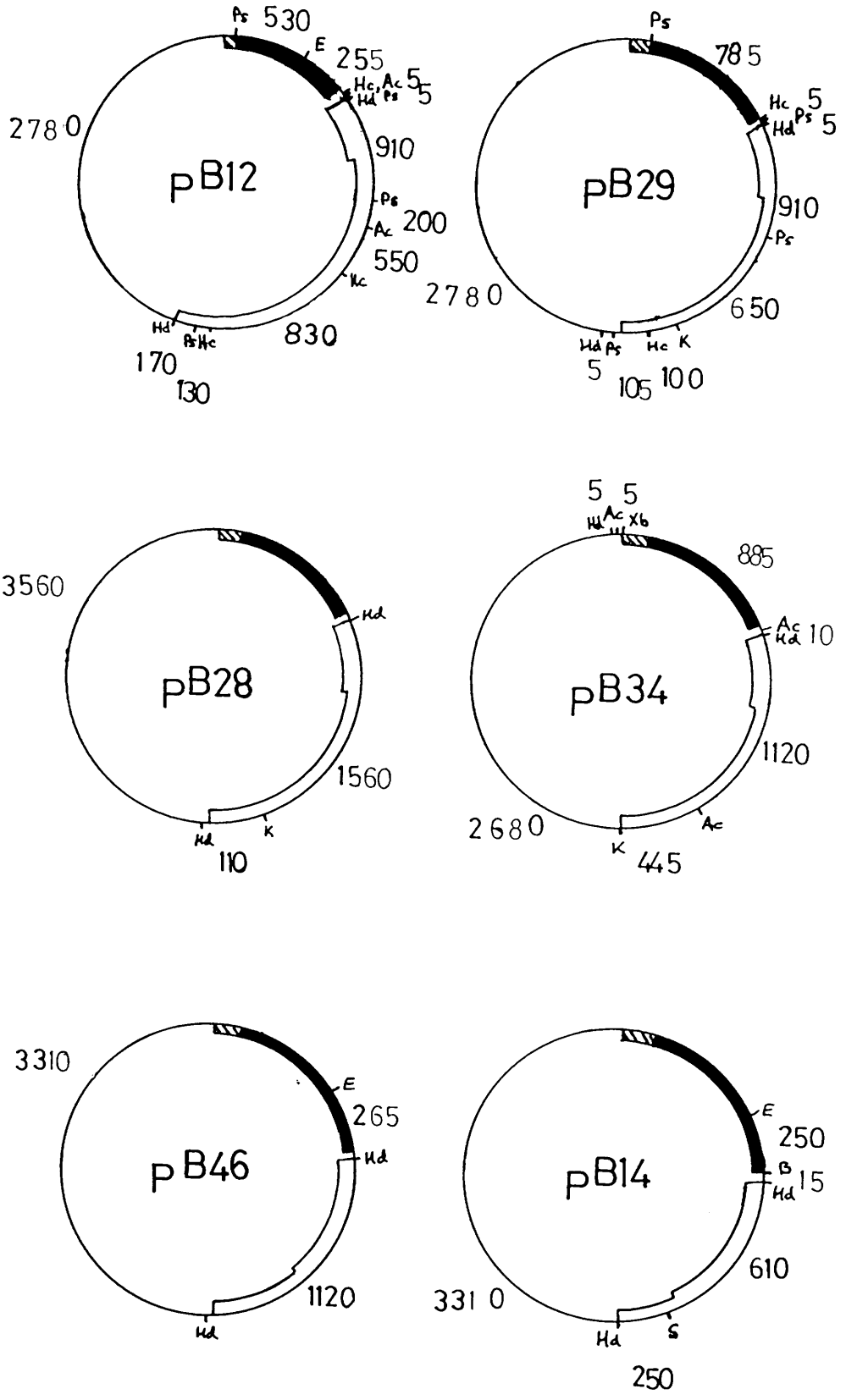

pB12
Ps 530
E 255
Hc, Ac 5
Hd Ps 5
910
Ps
Ac 200
550
Hc
830
Hd
Ps Hc
170
130
2780
pB29
Ps
785
Hc 5
Hd Ps 5
910
Ps
650
Hd Ps
Hc K
5 105 100
2780
pB28
Hd
1560
Hd K
110
3560
pB34
5 5
Hd Ac Xb
885
Ac
Hd 10
1120
Ac
K 445
2680
pB46
E
265
Hd
1120
Hd
3310
pB14
E
250
B
Hd 15
610
Hd S
250
3310

Figure 8d

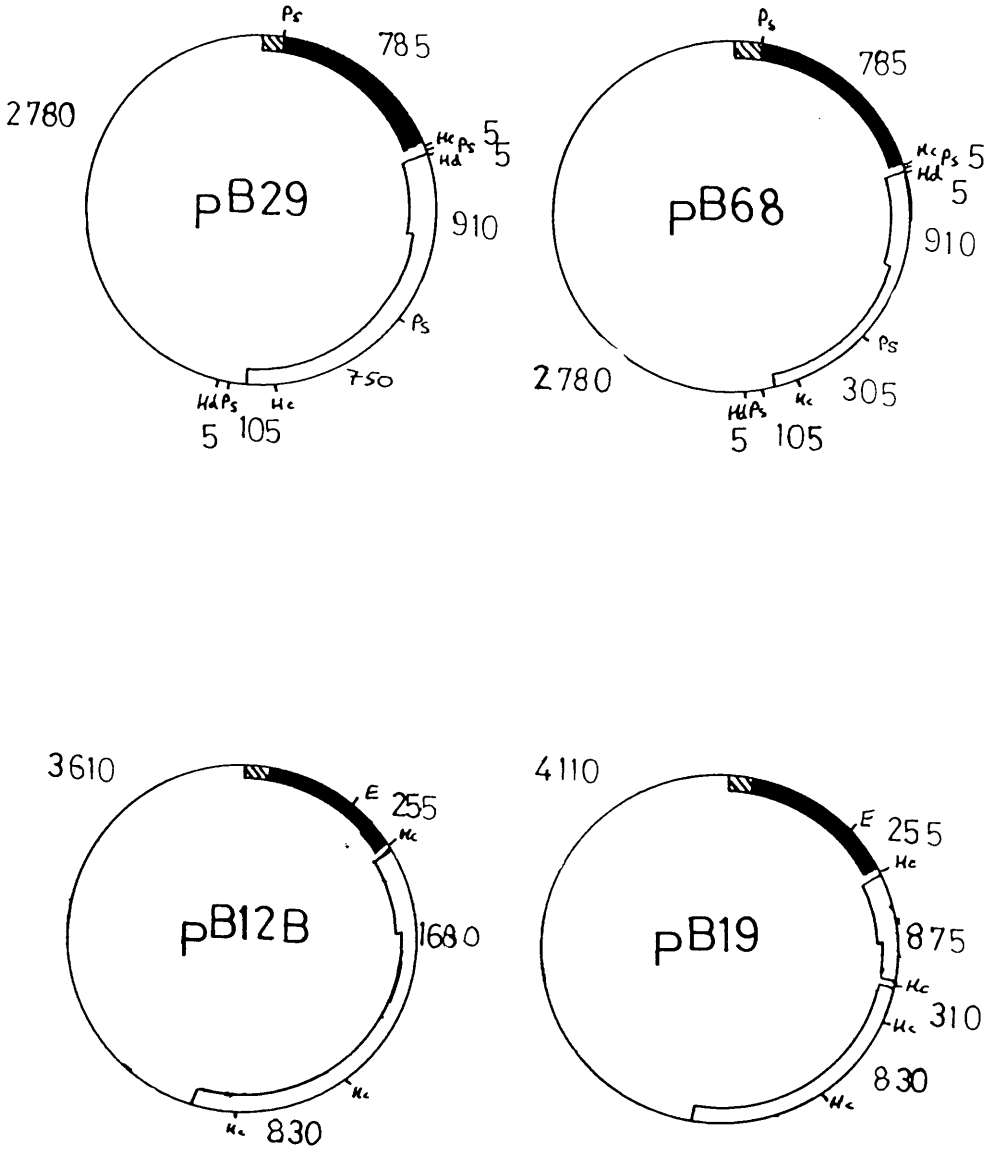

pB29
2780
785
910
750
105
5
5
5
pB68
785
910
2780
305
105
5
5
pB12B
3610
255
1680
830
pB19
4110
255
875
310
830

Figure 8e

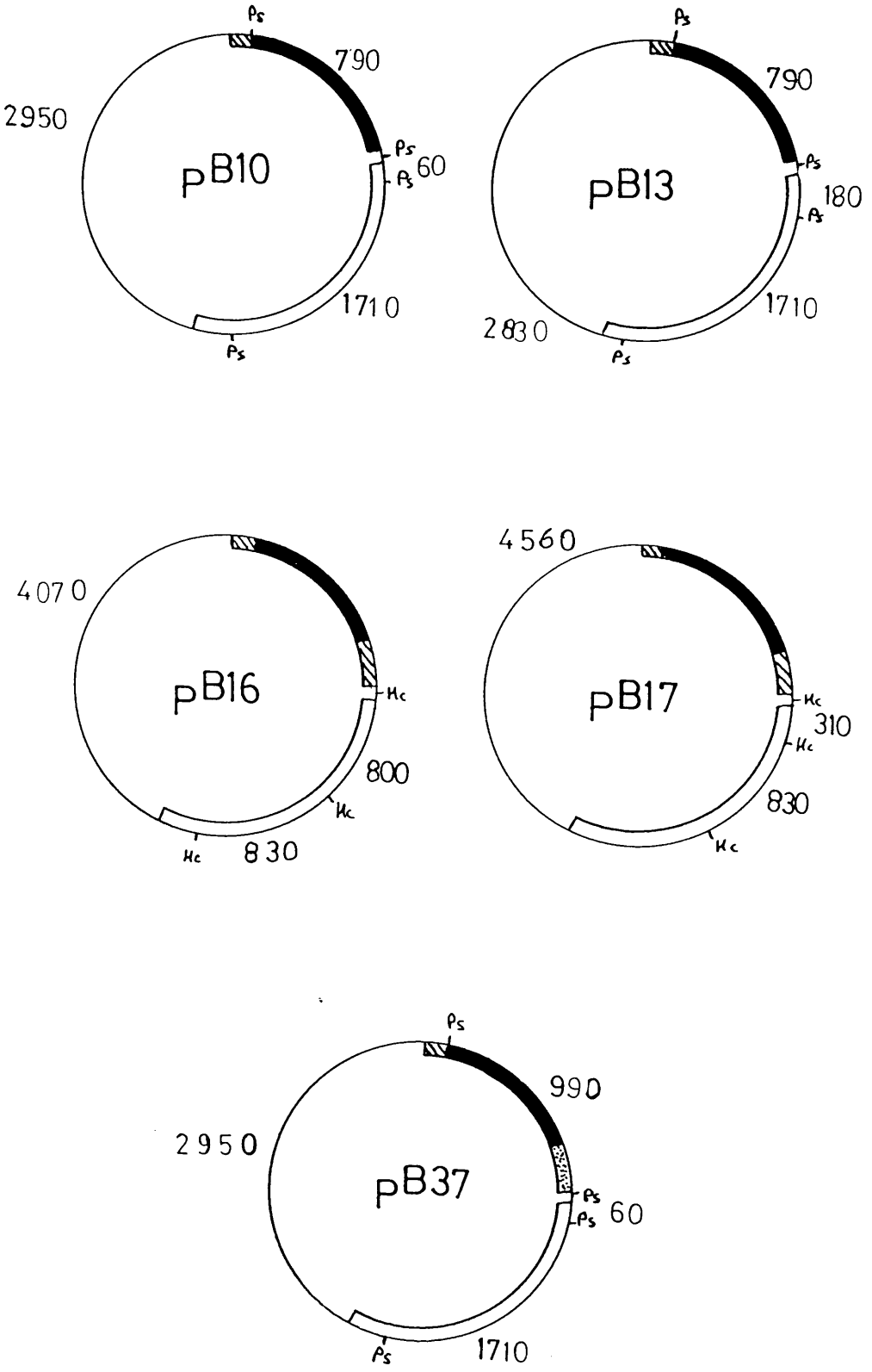
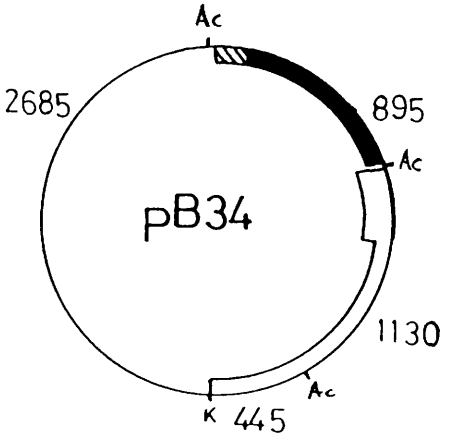
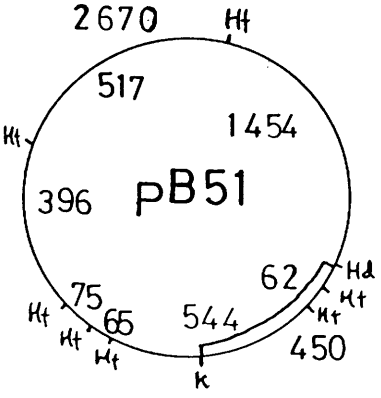
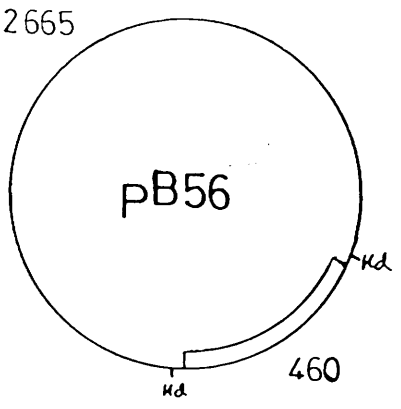
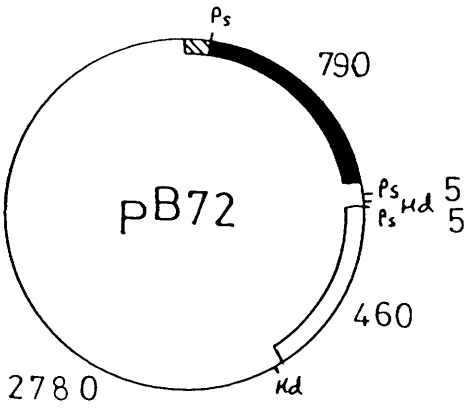

Ps
790
2950
pB10
Ps
Ps
60
1710
Ps
Ps
790
pB13
Ps
180
Ps
1710
2830
Ps
4070
pB16
Hc
800
Hc
Hc
830
4560
pB17
Hc
310
Hc
830
Hc
Ps
990
2950
pB37
Ps
Ps
60
Ps
1710

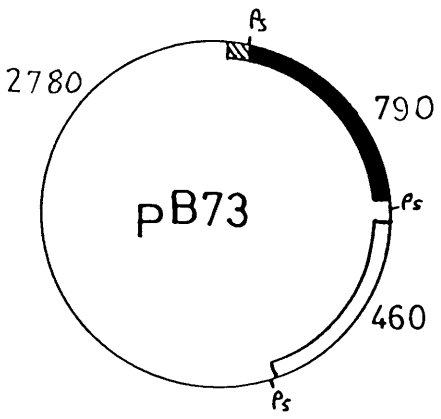
Figure 8f


Ac
2685
895
Ac
pB34
1130
Ac
K
445


2670
Ht
517
1454
Ht
396
pB51
62
Hd
Ht
Ht
75
Ht
65
544
Ht
Ht
450
K


2665
pB56
Hd
460
Hd


Ps
790
Ps
Ps
Hd
5
5
pB72
460
2780
Hd


Ps
2780
790
pB73
Ps
460
Ps

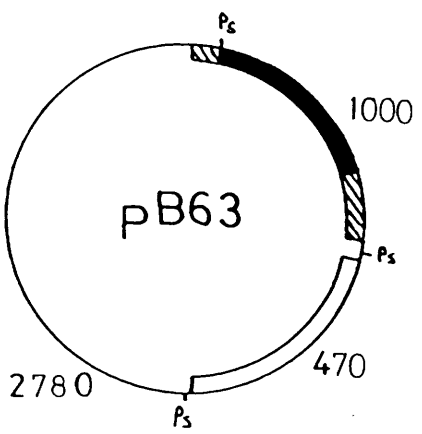

Ps
1000
pB63
Ps
470
2780
Ps

Figure 8g

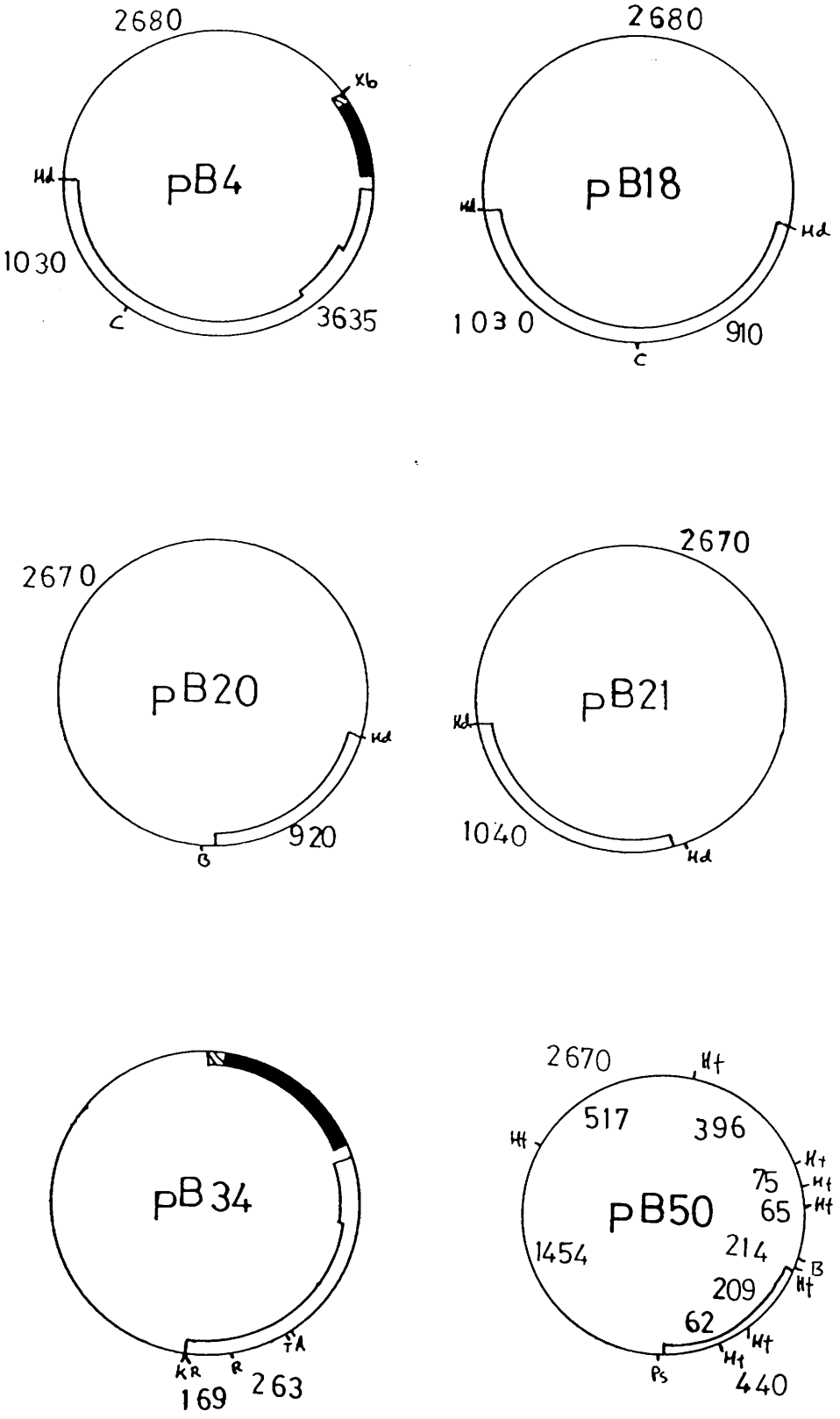

2680
Xb
Hd
pB4
1030
C
3635
2680
pB18
Hd
Hd
1030
C
910
2670
pB20
Hd
B
920
2670
pB21
Hd
1040
Hd
pB34
K R
R
TA
169
263
2670
Ht
517
396
Ht
Ht
75
Ht
Ht
pB50
65
1454
214
B
209
Ht
62
Ht
Ht
Ps
440

Figure 8h

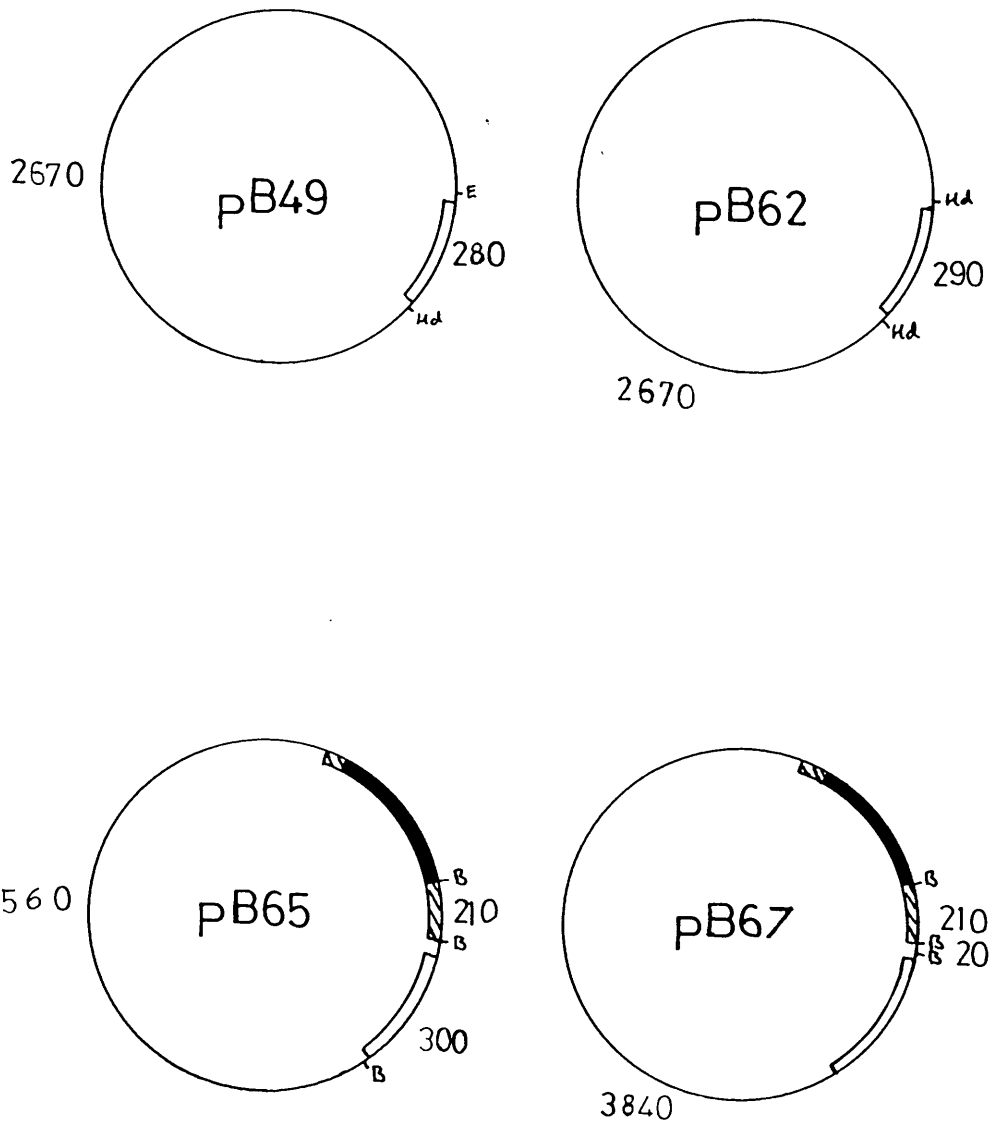
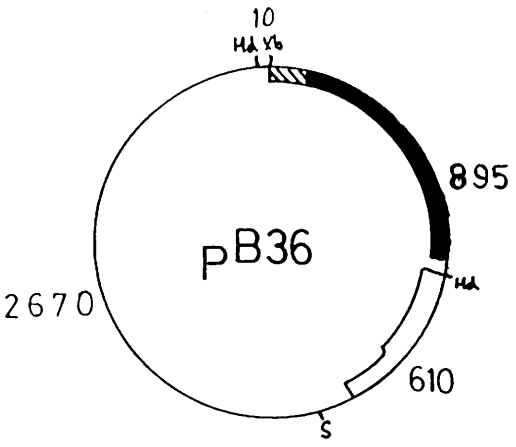
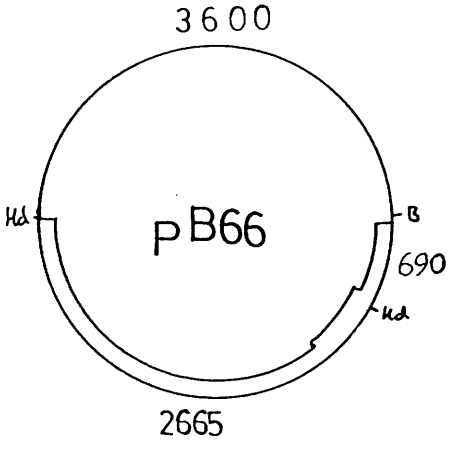
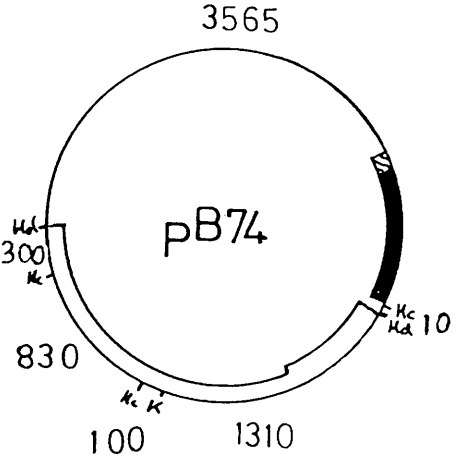

2670
pB49
E
280
Hd
pB62
Hd
290
Hd
2670
3560
pB65
B
210
B
300
B
pB67
B
210
B
B 20
3840

Figure 8i


10
895
pB36
2670
610
S


3600
pB66
690
2665


3565
pB74
300
830
100
1310
10

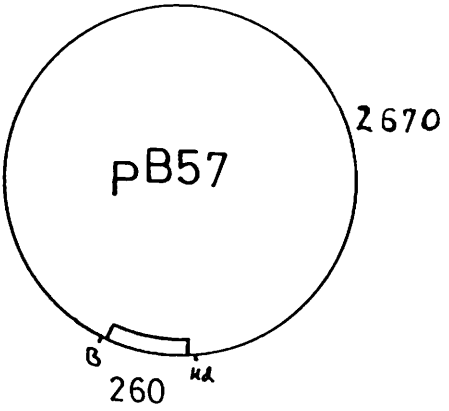

2670
pB57
260

Figure 8j

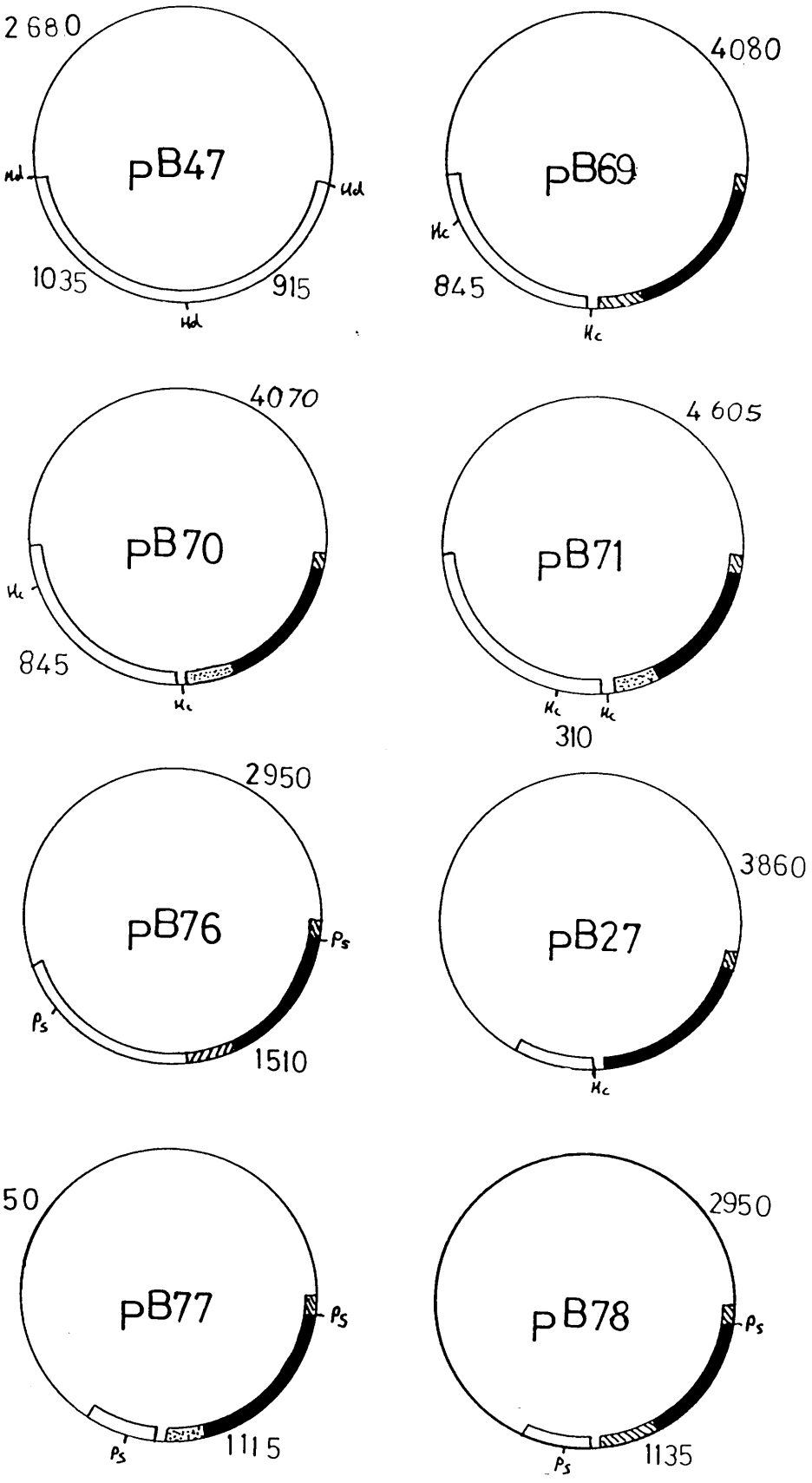
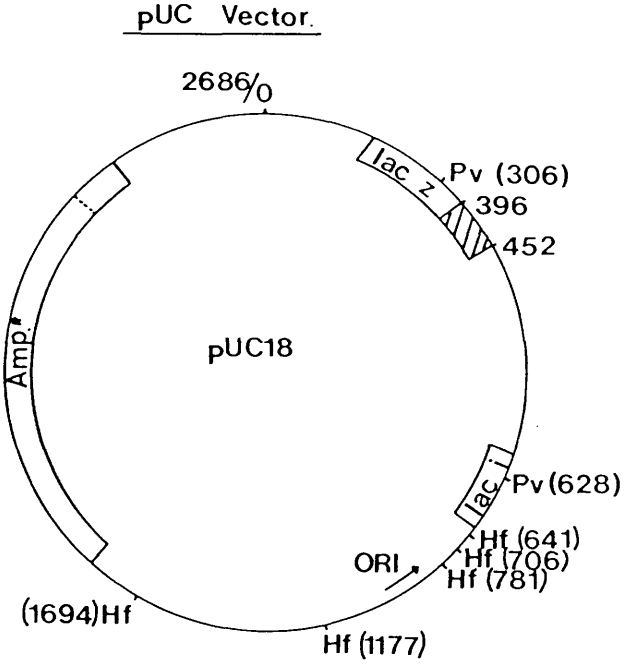

2680
pB47
Hd
Hd
1035
915
Hd
4080
pB69
Hc
845
Hc
4070
pB70
Hc
845
Hc
4605
pB71
Hc
Hc
310
2950
pB76
Ps
Ps
1510
3860
pB27
Hc
2950
pB77
Ps
Ps
1115
2950
pB78
Ps
Ps
1135

Figure 9.

Structure of Basic Cloning Vector.

The structure of pUC18 (Messing and Vieira, 1982) is depicted in the diagram. The hatched bar represents the polylinker cloning region and its sequence is presented below the plasmid map, showing the single cutting restriction cloning sites. Also used as a cloning vector, pUC12 is essentially the same except that it does not contain the KpnI and SphI cloning sites. For one ligation pAT153 was used as a cloning vector (Twigg and Sherrat, 1980).



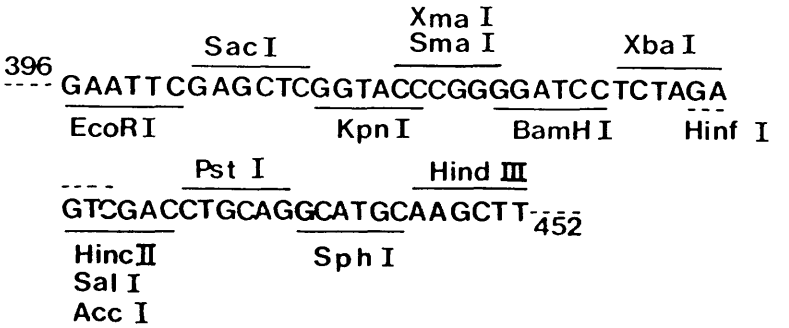


Figure 10.

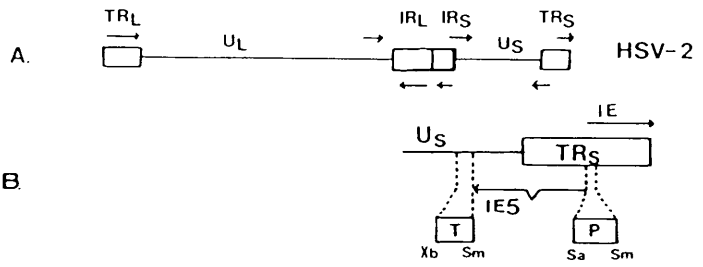
Structure of HSV-2 Genome and Derivation of HSV-CAT Recombinant.

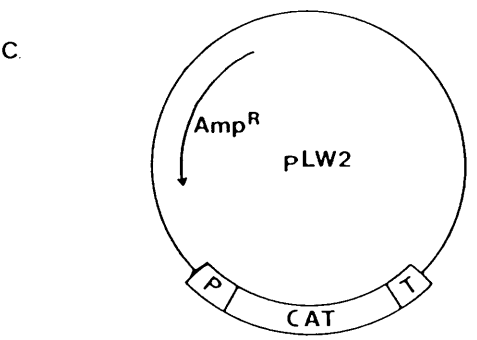
A. The stucture of the HSV-2 genome. Arrows indicate the site and direction of immediate early (IE) transcription. TR = terminal repeat; IR = inverted repeat; U = unique region; L = long region; S = short region.

B. Expanded region of HSV-2 genome showing the IE-5 gene transcription unit. T = 3' end processing signal; P = promoter domain.

C. Map of pLW2. The 210bp SmaI Sau3a promoter domain was inserted 5' to the bacterial CAT gene, with transcription 3' processing being regulated by the 100bp IE-5 region (Gaffeny et al, 1985).

D. The sequence of the 210bp promoter domain showing Sp1 consensus sequences (underlined), the TATA-consensus (boxed) and the cap site (arrow).





D.

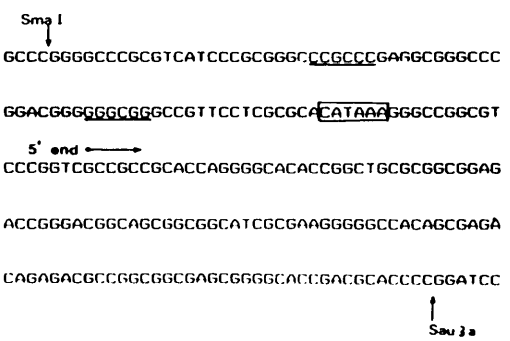


Figure 11.

Derivation of the Basic CAT Vector.

By digesting pLW2 with BamHI and subsequent self-ligation of the resultant linear (after removal of the 210bp promoter fragment by low melting agarose electrophoresis), the CAT vector pB9 was generated. This CAT vector does not contain an obvious promoter domain, nor a cap site. Transcription 3' processing is regulated by the HSV-2-IE termination signal (Gaffeny et al, 1985). The arrow indicates the direction and extent of transcription. See figures 8a and 12.

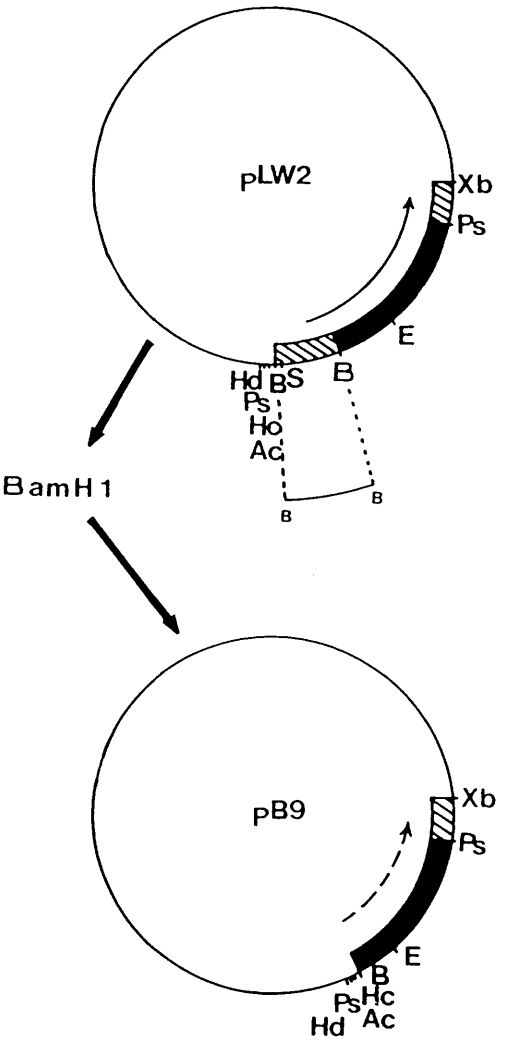


Figure 12.
1% agarose gel.

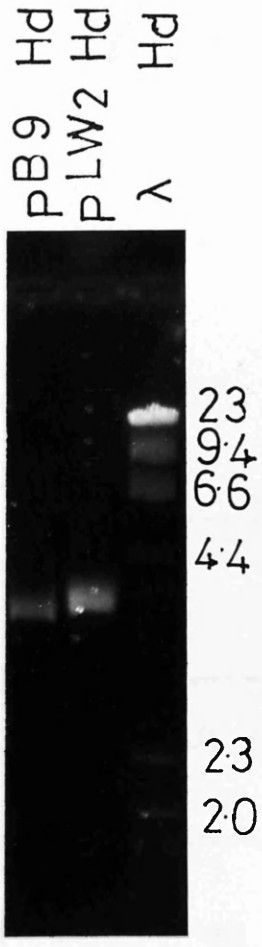
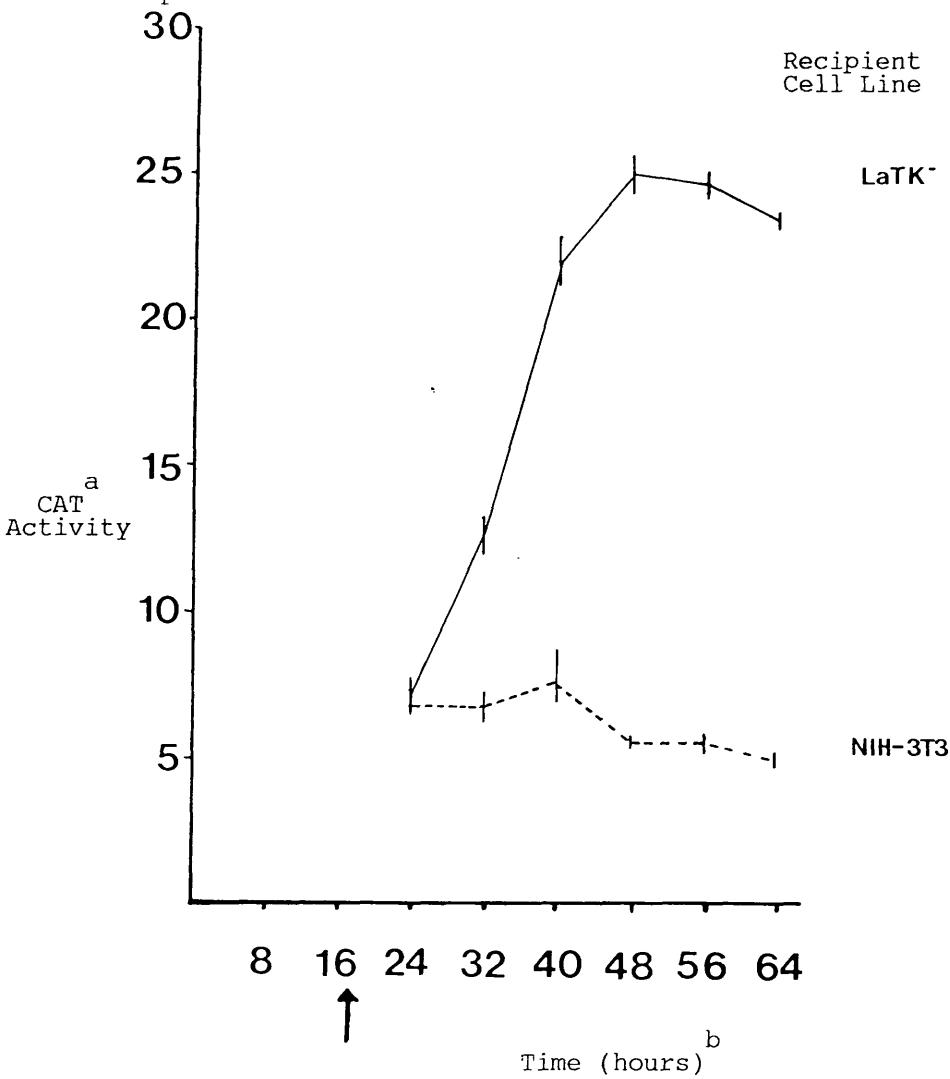


Figure 13.

Standardisation of Transient CAT Assays:
Expression Time.



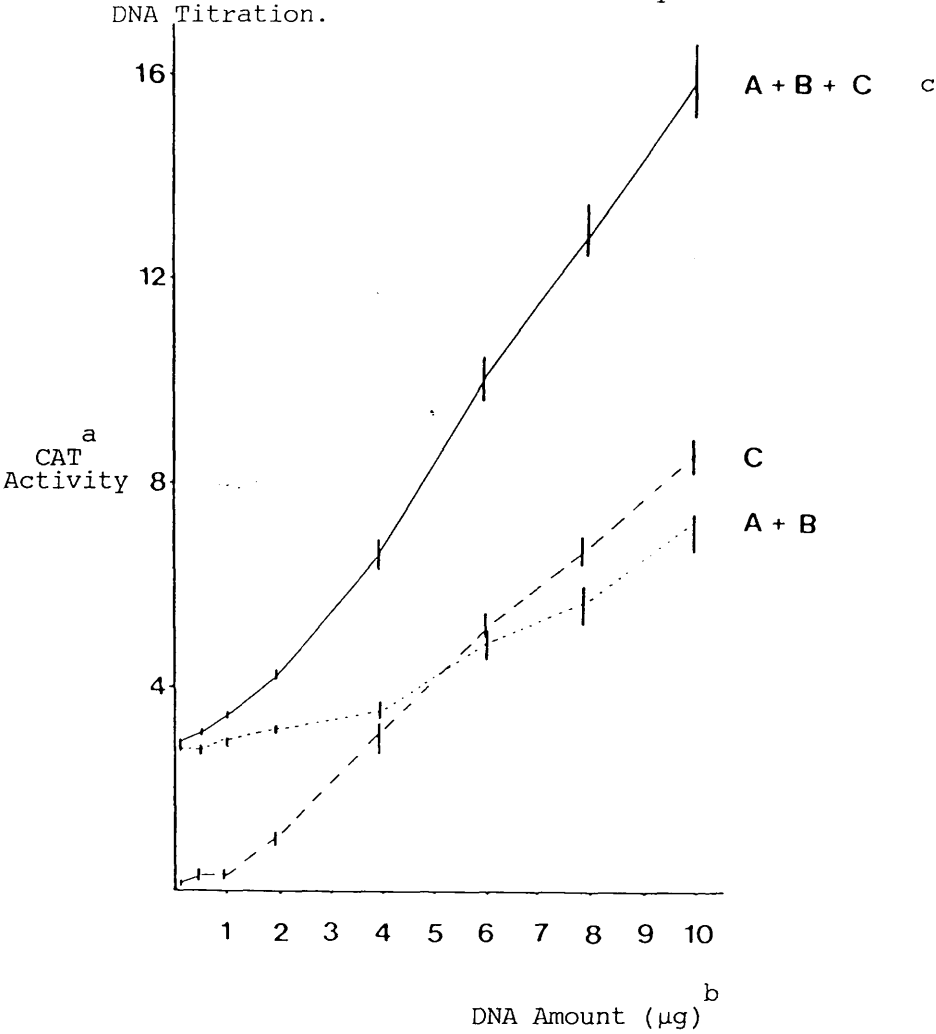
a. Average total counts (10^3cpm) from two experiments of duplicate flasks. Vertical bars indicate the range of values observed.

b. Hours post transfection (i.e. post addition of precipitate). The arrow indicates the time point of removal of precipitate and refeeding of cells with medium.

c. 5µg of pLW2 was transfected into LaTK and NIH-3T3 cells as described in Materials and Methods. The resultant CAT activity was assayed as described in Materials and Methods.

Figure 14.

Standardisation of Transient CAT Assay:
DNA Titration.



a. Average total counts (10^3cpm) from two experiments of duplicate flasks. Vertical bars indicate the range of values observed.

b. 1-10μg of pLW2 was transfected into LaTK$^-$ cells as described in Materials and Methods. The resultant CAT activity was assayed as described in MAterials and Methods.

c. The various acetylated forms of chloramphenicol (see figure 6.).
A and B = monoacetylated chloramphenicol.
C = diacetylated chloramphenicol.

Figure 15.

Standardisation of Transient CAT Assay.

A. Cell Extract Titration.

B. Assay Time Course.

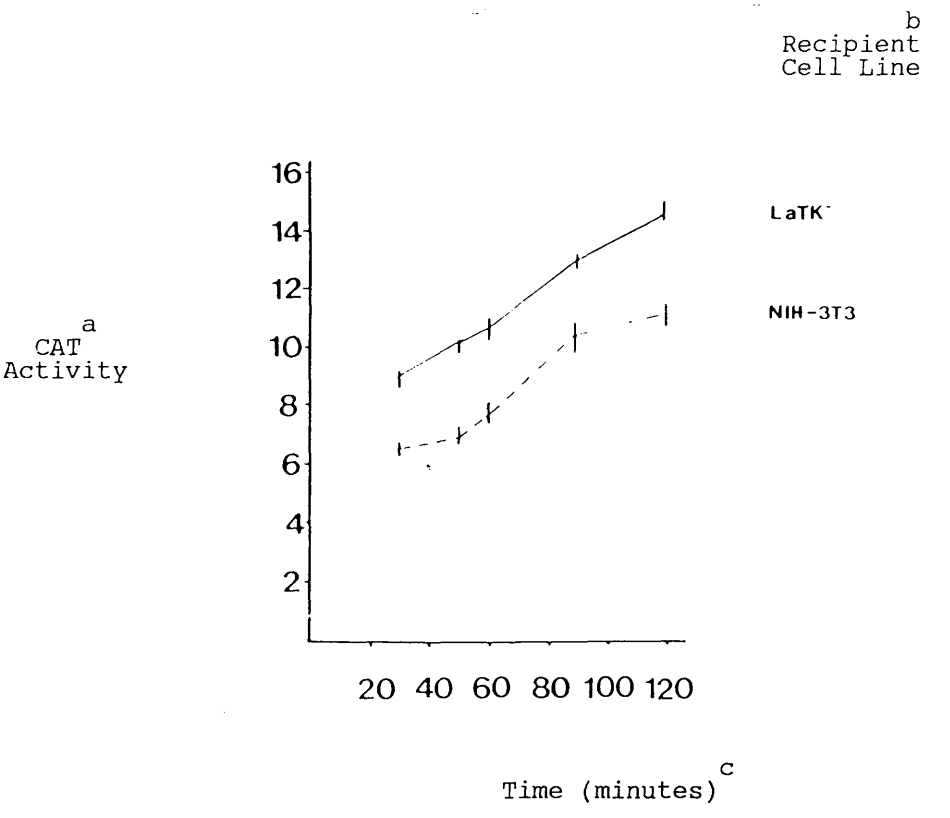
a. Average total counts (10^3cpm) from two experiments of duplicate flasks. Vertical bars indicate the range of values observed.

b. 10μg of pLW2 was transfected into LaTK$^-$ and NIH-3T3 cells as described in Materials and Methods. The resultant CAT activity was assayed as described in Materials and Methods.

c. Incubation time of CAT assay. 20μl of cell extract was assayed for CAT activity.

d. Volume of cell extract assayed. The CAT assay was incubated for 60 minutes.

A. CAT Assay Time Course.



B. Cell Extract Titration.

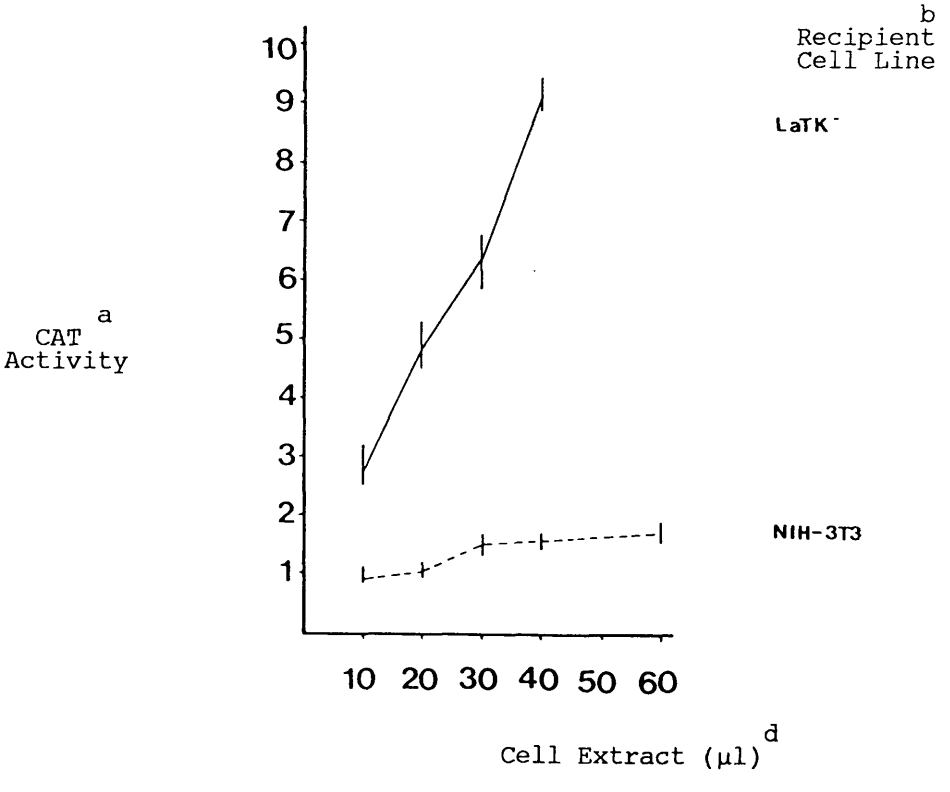


Table 1.

Comparison of Cell Extracts after Protein Estimation

	[a] Protein Concentration	[b] CAT Activity
A.	130μg	2367
	130μg	2409
	130μg	2278
	130μg	2331
B.	555μg	9067
	560μg	9409
	530μg	8799
	515μg	9110

A. The cell extract from one flask of LaTK$^-$ cells was aliquoted into 4 samples and assayed for CAT activity as described in Materials and Methods. The protein concentration (Biorad protein assay, gave values to the nearest 5μg) was measured.

B. The protein content of 4 duplicate flasks was determined (Biorad assay), and the CAT activity of the total cell extract was assayed for as desrcibed in Materials and Methods.

a. Protein concentration as measured by the Biorad assay, as recomended by the suppliers.

b. CAT activity as total counts (cpm) of acetylated forms. (see figure 6.).

c. In each case 5μg of pLW2 was transfected into LaTK$^-$cells as described in Materials and Methods. The CAT activity was assayed as described in Materials and Methods.

3.2. Identification and Localisation of a Negative Regulatory Element Within the 5' Flanking Sequences of the Human c-myc Gene

If a cis-acting negative regulatory element (NRE) exists in the 5' flanking region of the human c-myc, as postulated, then progressive 5' deletions of this region might allow its identification. CAT activity driven by various c-myc 5' deletion fragments was analysed in the various cell lines previously described (2.3.1.), and such a domain localised within the human c-myc 5' flanking region.

3.2.1. Recombinant Plasmid Constructions

All recombinant plasmids in this series, were constructed using the CAT vector (pB9) and various regions of the c-myc 5' flanking sequences plus exon 1 derived from pMC41 (see figure 16), kindly given by T.Papas (Dalla-Favera et al, 1982b).

(i). Isolation of c-myc 5' fragment

The 3.6Kb HindIII-XbaI fragment from pMC41 was inserted into HindIII-XbaI digested pUC12 (pB1; see figure 17) and subsequently the resulting 3.6Kb HindIII-BamHI fragment (the BamHI site derived from the pUC polylinker) was cloned into HindIII-BamHI digested pAT153 (pB2; see figure 18). These steps were required to reduce the number of PvuII sites present to two (pB2) thereby allowing easy partial digestion with PvuII and subsequent HindIII molecular linker modification to give pB3 (see figure 18). The dual c-myc promoter region, plus the majority of exon 1 and 2.3Kb of 5' flanking sequences (originally encompassed within the 2.8Kb HindIII-PvuII fragment of pMC41) is now present in pB3 as a 2.8Kb HindIII fragment. This conversion of a PvuII site to a HindIII site was done to enable easy construction of the 5' deletion recombinants. HindIII linker or BamHI linker modification after PvuII linearisation of pB3, gave

respectively, pB7H and pB7B (see figure 18).

(ii). Construction of the 5' Deletion Recombinants

The c-myc recombinant (pBl2; see figure 19), i.e. containing all 2.3Kb of flanking sequences, was constructed by inserting the 2.8Kb HindIII fragment from pB3 into HindIII digested pB9. All further c-myc 5' deletion recombinants are variations on this plasmid, and consequently they all contain 510 bp of exon 1 (i.e. except the distal 40bp), and both TATA-consensus elements.

The strategy used to construct the 5' deletion recombinants is described below, in order of increasing deleted fragment size. The ClaI-XbaI fragment from pBl2 was inserted into AccI-XbaI digested pUCl2 to give pB29 (both the ClaI and AccI sites being destroyed on ligation; see figure 20), which contains 1.25Kb of 5' flanking sequence. The 1.7Kb HincII fragment from pBl2 was inserted into HincII digested pB9 to give pB28 (see figure 20), which contains 1.15Kb of 5' flanking sequence. By inserting the KpnI-XbaI fragment of pBl2 into KpnI-XbaI digested pUCl8, pB34 was generated containing 1.05Kb of 5' flanking sequence (see figure 20). AccI digestion of pB28 released a 1.1Kb fragment which was inserted into AccI digested pUCl2 to give pB45 (the 5' AccI site being destroyed as this ligation involved the two isoschizomeres of AccI; see figure 21). This allowed a 1.1Kb HindIII fragment (the 5' HindIII site was derived from the pUC polylinker), containing 610bp of 5' flanking sequence, to be inserted into HindIII digested pB9 generating pB46 (see figure 21). The final recombinant in this series pB14 (see figure 22), containing only 350bp of 5' flanking sequence was generated by insertion of the 860bp HindIII fragment from pB7H into HindIII digested pB9. Figure 28 shows these 5' deletion constructions and the location of relevant in vivo DNase 1 hypersensitive sites (HSS).

(iii). Construction of the Upstream Internal Deletion

AccI partial digestion plus KpnI limit digestion of pB29, followed by Klenow flush-ending and subsequent religation gave pB68 (see figure 23). This internally deleted recombinant (pB68) is identical to pB29 (the c-myc recombinant with 1.25Kb of 5' flanking sequences), except that the 445bp KpnI AccI fragment containing the in vivo DNase 1 HSS II-2 has been deleted.

3.2.2. Restriction Enzyme Analysis of Recombinant Plasmids

Extensive restriction enzyme analysis has been used to identify each of the recombinant plasmids described. Representative digests are shown in figures 24, 25, 26,27 and relevant restriction maps in figures 8a, 8b, 8c, 8d.

(i). Plasmid with the Full 5' Flanking Region and Parental Plasmid pMC41.

PvuII digestion of pMC41 reveals the 5 internal human c-myc restriction sites (1341, 1091, 880 and 860bp) and the plasmid vector site (4400 and 4100bp). The recombinants pB1, pB2 and pB3 were made to allow the number of PvuII sites present to be decreased as shown by PvuII digestion of the respective recombinants; pB1 (2360, 2030, 860 and 790bp), pB2 (6300 and 860bp) and pB3 (7160bp). The 860bp fragment of pB1 and pB2 contains the majority of the c-myc exon 1 sequences. This remaining PvuII site was converted to a BamHI site by linker modification to give pB7B, which gives upon BamHI digestion fragments of 5630 and 1530bp in size. In pB7H the remaining PvuII site (of pB3) was converted to HindII by linker modification, which then gave upon HindIII digestion fragments of 4360, 1940 and 860bp.

(ii). Plasmids with 5' Deletions of 5' c-myc Flanking Region

Since pB12 was used as the parent of all the deleted

c-myc recombinants, its restriction map is shown in detail, and is a reference for the recombinants in this series. HindII digestion of pB12 gives fragments containing the c-myc derived sequences (2790bp) and the pUC vector and CAT gene sequences (3575bp, basically pB9. Upon HincII digestion, fragments of 3865, 1670 and 830bp are produced. AccI digestion gives two fragments of 5250 and 1115bp (this digest went only to partial completion, thus the upper band corresponds to the linear plasmid, while PstI digestion reveals 4 fragments of 2950bp (pUC and HSV-IE terminator sequences), 1710bp (c-myc 5' flanking sequences), 915 (c-myc 5' adjacent sequences and the majority of exon 1) and 890bp (CAT structural gene).

The recombinants pB9 (CAT vector), pB14, pB46 and pB12 were digested with EcoRI giving respectively 3.6, 4.4, 4.7 and 6.4Kb fragments, while pB28 and pB29 digested with KpnI gave 5.2 and 5.3Kb fragments, and pB34 digested with XbaI gave a 5.1Kb fragment. The c-myc derived fragments were identified by HindIII digestion of pB29 (1770bp), pB28 (1670bp), pB46 (1120bp) and pB14 (860bp), while HindIII KpnI digestion of pB34 gave the c-myc derived fragment (1575bp) and the CAT-HSV terminator domain fragment (895bp). HindIII digested pB9 is shown to indicate the basic vector size (3.6Kb). BamHI SmaI digestion of pB14 reveals fragments of 625 and 3795bp.

(iii). Plasmid with Upstream Internal Deletion

Since pB68 is derived from pB29 by deletion of a 445bp region, the restriction digests of these two recombinants are shown beside each other. PstI digestion of pB29 reveals identical fragments except for the 855bp in pB29 which is 410bp in pB68. HincII digestion also confirms the deletion (1670bp in pB29 and 1225bp in pB68), as does HindIII digestion (1770 and 1325bp respectively).

3.2.3. Identification of a cis-acting Negative Regulatory Element

Transient transfection studies using an equimolar ratio of 10μg (pB9) for pB12 (containing 2.3Kb 5' flanking sequences), pB14 (containing 350bp 5' flanking sequences), pLW2 (positive control) and pB9 (negative control) were undertaken in an attempt to identify a dominant cis-acting transcription regulatory domain. The resultant CAT activities are presented in table 2 (and see figure 45). The 6 recipient cell lines have been described previously (see section 2.3.1.). The c-myc recombinant containing 2.3Kb of upstream sequences (pB12) was observed to have a relatively weak promoting activity when compared to pLW2. However, 5' deletion to the PvuII (pB14) dramatically elevated CAT levels by 3-to 12-fold. Therefore the 2.0Kb HindIII-PvuII fragment 5' to the c-myc dual promoter contains a cis-acting negative regulatory element (NRE) and deletion of this element relieves the promoter domain from repression. Due to subsequent experimentation this domain has been designated NRE-2. The relatively low repression ratio observed (still significant) is due to the amount of DNA (equimolar ratio of 10ug) transfected. With lower DNA amounts a much greater repression ratio is observed (see section 3.4.4.).

There appears to be no obvious specificity to this element as it functions both in human and murine cells, of an erythroid, epithelial or fibroblastic origin. Slight variations in the repression ratio (2-4 fold) are seen among LaTK⁻, HeLa and NIH-3T3 cells. Such small variations are unlikely to reflect a difference in the tissue- or species-specificity. As described before (see section 3.1.6.), the observed ratios can only be qualitatively compared between cell lines. The repression ratio observed in K-562 cells is at least 2-3 times that seen in any of the other cell lines used as recipients in these studies. The presence of repression activity in K-562 cells was unexpected. K-562 exhibit elevated levels of c-myc RNA

without any obvious rearrangement of the c-myc locus (Wong-Staal et al, 1983), and it was therefore hoped that they would not contain the postulated repressor and thus be permissive for c-myc expression. This has implications for the regulation of c-myc expression in K562 cells (see section 3.4.4.).

However, the c-myc recombinants, even though exhibiting the repression function (see pB14 and pB12), have only relatively weak activity in Hela and NIH-3T3 cells when compared to pLW2. By comparing absolute counts (pLW2 control not done) the c-myc-CAT recombinants are probably also only weakly expressed in MRC-5 and GM468 cells. This may imply a certain level of specificity between cell types, or may just represent differing transfection efficiencies. Other workers find GM468 and MRC-5 cells difficult to transfect into (D. Spandidos, personal communication). Irrespective of these differences, the repression of pBl2 compared to pBl4 levels is constantly observed in all these cell lines.

3.2.4. Localisation of the Negative Regulatory Element

To localise this element within the 2.0Kb HindIII-PvuII 5' region, equimolar amounts (of 5μg) of the various c-myc 5' deletion recombinants were transfected into LaTK and HeLa cells and the resultant CAT activities are presented in figure 28. Deletion of the distal 1225bp (pB29, pB28 and pB34) which contains DNase I HSS I and II-1 had only a small effect on the activity observed when compared to pB14 (see section 3.6.). A further 445bp deletion (pB46) to the AccI site allowed a release from repression similar to that of pBl4. Analysis of pBl2, pB34 and pBl4 in Hela cells gave similar results, however, in HeLa cells the repression ratio was less than that observed in LaTK$^-$ cells. Whether such a slight difference represents cell specificity is questionable. It may just reflect the state of these in vitro cultured cells and their differing abilities to take up and express exogenously introduced DNA (see section

3.2.5.).

One anomaly which will be discussed later is the apparent slight release from repression observed with pB34. Perhaps this deletion has in some way affected the NRE-2 element structure or conformation (see 3.4.4.).

3.2.5. Internal Deletion of the Negative Regulatory Element (NRE-2)

To confirm these results an internal 445bp deletion (KpnI to AccI, which is the region deleted between pB34 and pB46) from pB29 was constructed (pB68). Equimolar amounts of pB29 (containing 1.25Kb of upstream sequences), pB68 (identical to PB29 except that the 445bp fragment is internally deleted), pB14 (containig 350bp upstream sequences) and pB9 (negative control) were transfected into LaTK$^-$ cells. The resultant CAT activities are plotted in figure 29. Although this internal deletion does not restore full activity (i.e. to that of pB14) its CAT levels are substantially greater than that of its parent, pB29. This indicates that at least part of the negative regulatory element resides within this 445bp region as implied by the 5' deletion studies, however, flanking sequences (or indeed distal sequences) may be additionally required for full repression. Therefore this 445bp domain does define an essential component of the NRE-2. It is, however, possible that this 445bp region contains all the sequences requiired for repression (i.e. contains the whole NRE-2), and the difference between pB14 and pB68 CAT levels could be due to some structural alteration in the general chromatin structure of the c-myc 5' control regions in the two recombinants (in addition to the NRE-2 deletion) brought about by the very nature of the internal deletion. Due to subsequent experimentation, the former proposal (i.e. the 445bp region only deliniates an essential part but not all of the NRE-2) is favoured.

A point to note, one which will be extensively described later (see section 3.4.4.), is that the magnitude of repression (compare pBl4 and pB29; see figure 29) decreases with increasing amount of DNA transfected.

From these studies, at least an essential component of the NRE-2 is located between the KpnI site (-1055bp) and the AccI site (-610bp) upstream from the human c-myc transcription start site, when assayed transiently in LaTK cells. In addition this transcription regulatory element shows an apparent lack of tissue- or species-specificity. In subsequent experimentation 3 different fragments containing the NRE-2 domain were used; a 2.0 Kb HindIII-PvuII fragment, a 445 bp KpnI-AccI fragment and a 270 bp RsaI-ThaI fragment (see figure 16).

3.2.6. Characterisation of c-myc Promoter

The relative activity of pBl4, the 5' deletion which leaves only 350bp upstream of P1 but contains all the 5' flanking sequences required for promoter activity in these studies, was compared to that of pLW2, by DNA titration transfection experiments, with equimolar amounts of the test recombinants being transfected into LaTK$^-$ cells. The resultant CAT activities are plotted in figure 30. In this experiment pBl4 was nearly as active as pLW2. Since pLW2 has already been characterised as a strong promoter (Gaffney et al, 1985; J. Lang, personal communication ;and see section 3.1.6.) this would indicate that pBl4, also contained a highly active promoter, when assayed in these cells.

The CAT vector alone (pB9) showed only very slight activity (1-2% of the full promoter) in LaTK$^-$and HeLa cells. However, in NIH-3T3 and K-562 cells the CAT vector alone had an activity value of about 10% of the c-myc promoter. Since no activity was observed when pMC41 was assayed, a non-CAT containing plasmid (data not shown), the low level activity

seen for pB9 is not an assay artifact. This extremely low level of activity may be due to the presence of cryptic promoters within the pUC sequences, or may somehow be an innate property of a covalently closed plasmid DNA containing an ATG codon. This low level expression from a promoter-less / capless CAT recombinant differs from previously reported results. In experiments using pSV0-CAT, a promoter-less / capless CAT recombinant which contains a SV-40 polyadenylation signal, no expression (CAT activity) was seen (Alwine, 1985). This may reflect a function of the different polyadenylation signals (pB9 contains a HSV derived polyadenlyation signal: see section 3.1.3. and 3.1.4), or may just reflect a difference in the sensitivity between the two series of experiments.

However, pB14 is only 10% and 16% as active as pLW2 in Hela and NIH-3T3 cells respectively, while it is 2-3 times more active than pLW2 in K-562 cells.These differences in absolute values may represent tissue- and/or species-specificity for either one of the two promoters studied. There is no detailed data on the expression specificity of the HSV-IE promoters. Gross expressional variation (i.e. non-expression) has been taken to indicate specificity in transfection studies. However, where a gene is expressed but to differing levels depending on the cell type, absolute quantitation of that variation would be difficult to achieve. It is quite hard to conceive how to unambiguously standardise any experimentation involving such a gene, as the standard control, in effect, could not itself be conclusively defined to have an unvarying expression irrespective of recipient cell line. Thus expression patterns at the moment can only be described in general terms: no expression, weak expression or high expression. Applying this arguement to these studies, the differences seen between the relative activities of the c-myc and HSV-2 IE promoters in the various recipient cell lines can not be attributed to, specifically, one or the other promoter. Of note, is that in all cell lines analysed the c-myc promoter

(and repression) was active.

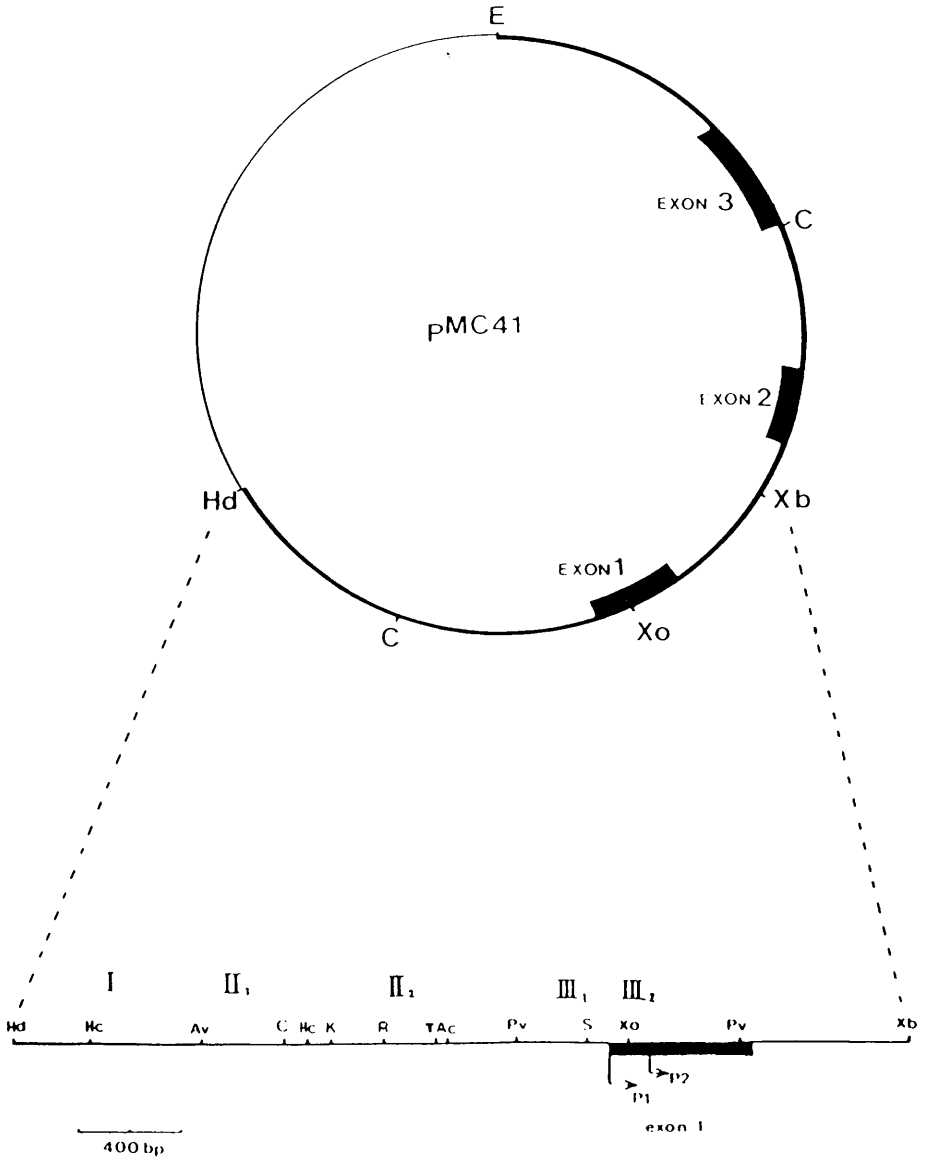
In these experiments, an assay saturation effect was observed: activity for both pLW2 and pB14 start to plateau of when more than 10μg of DNA are used for the transfection experiments (see figure 30; section 3.1.6.; Weintraub, Cheng and Conrad, 1986).

These results identify a negative regulatory element within the 5' flanking sequences of the human c-myc oncogene. This element is active in several cell types of both human and murine origins. It has been localised to a 445bp region, approximately 700bp 5' to P1 (-1052 to -607bp), when LaTK$^-$ cells were used as the transfection recipient. Since the transfected gene is of human origin, and the recipient cell of murine origin, a cross-species interaction is implied. In addition, the repression function was observed in fibroblastic, epithelial, and erythroid derived cells, thus implying a lack of tissue-specificity as well. This region contains the previously identified DNase I hypersensitive site II-2 (Siebenlist et al, 1984). Due to subsequent experimentation this region is designated NRE-2.

Figure 16

Parental c-myc Recombinant and Restriction Map of 5' Flanking Sequences and Exon 1.

The heavy lines represent human c-myc sequences (HindIII-EcoRI). The 5' flanking region, exon 1 and part of intron 1 are shown at the bottom of the page, with the relevant restriction enzyme sites indicated. P1 and P2 correspond to the two major c-myc transcription initiation sites. The position of the in vivo DNase 1 hypersensitive sites (Siebenlist et al, 1984) are designated by Roman numerals.



An unbound copy of this figure can be found in the sleeve of this thesis.

Figure 17.

Structure and Cloning Strategy of c-myc Recombinant pB1.

To enable 5' deletion recombinants to be generated the 5' region (2.3Kb of 5' flanking sequences, the whole 550bp of exon 1 and part of intron 1) were cloned into pUC12 as a HindIII-XbaI fragment. For a more detailed restriction map see figure 16. For a detailed description of the cloning strategy see section 3.2.1. See figure 8b and 24.

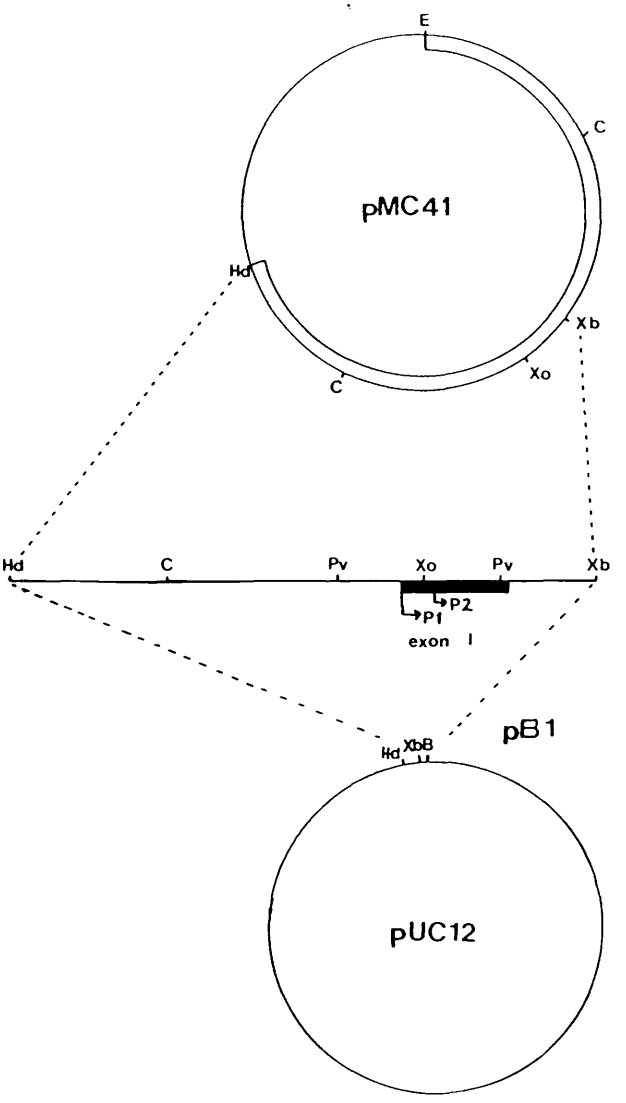


Figure 18.

Structure and Cloning Strategy of c-myc Recombinants pB2, pB3, pB7B and pB7H.

For a detailed description of the cloning strategy see section 3.2.1. See figures 8b and 24.

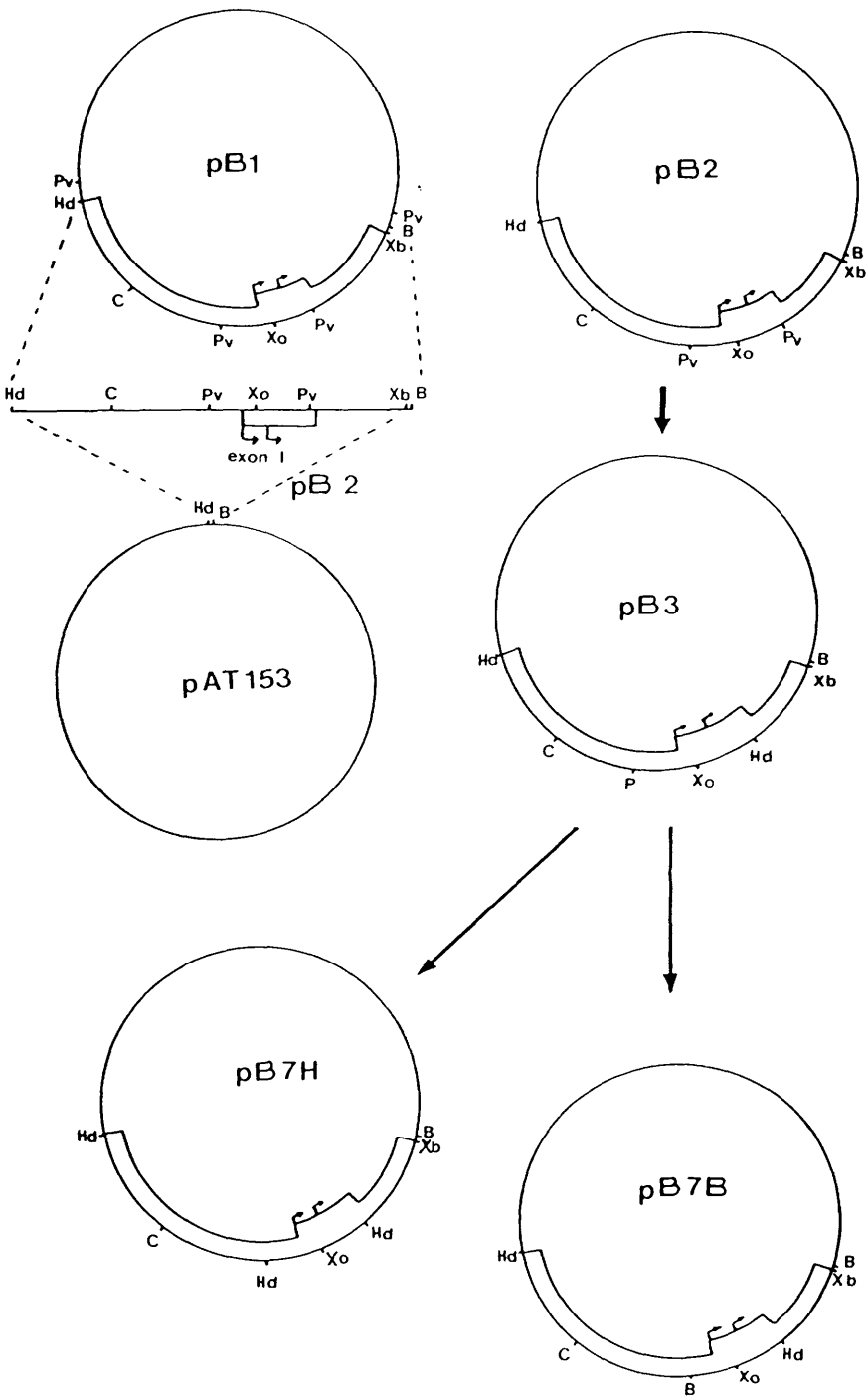


Figure 19.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB12.

For a detailed description of the cloning strategy see section 3.2.1. See figure 8b, 8c, 25 and 26.

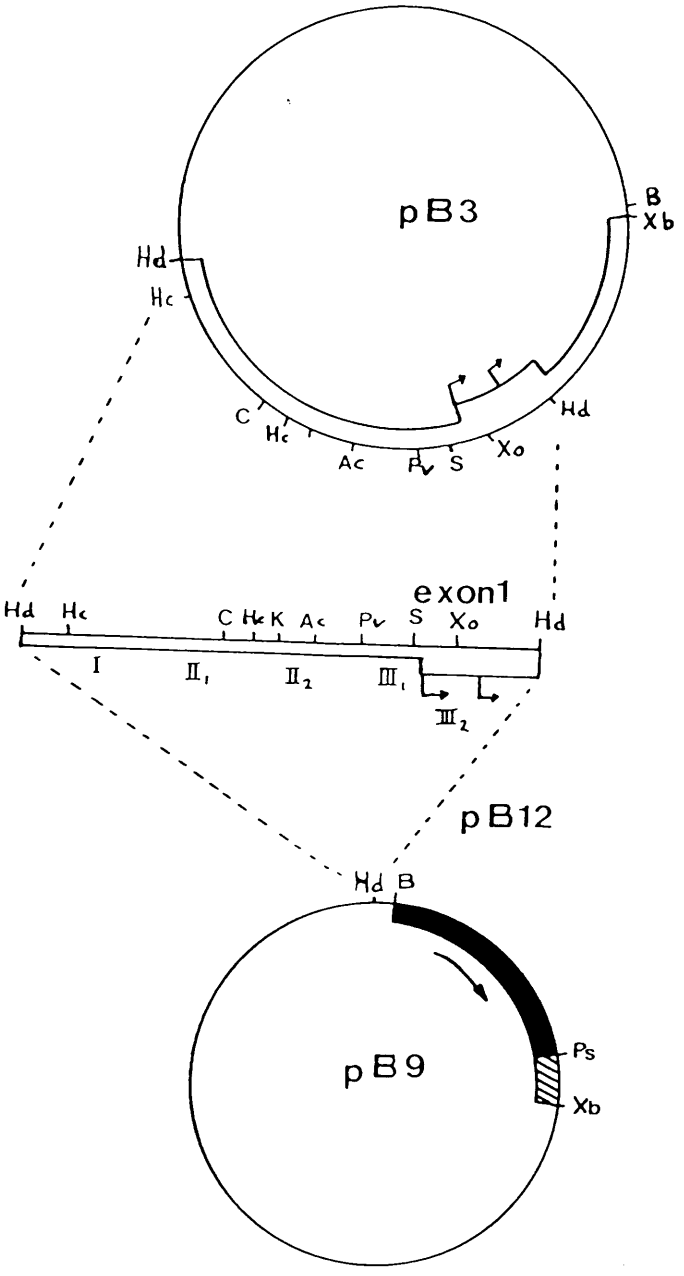


Figure 20.

Structure and Cloning Strategy of 5' Deleted c-myc-CAT Recombinants pB29, pB28 and pB34.

For a detailed description of the cloning strategy see section 3.2.1. See figure 8c, 25, 26 and 27.

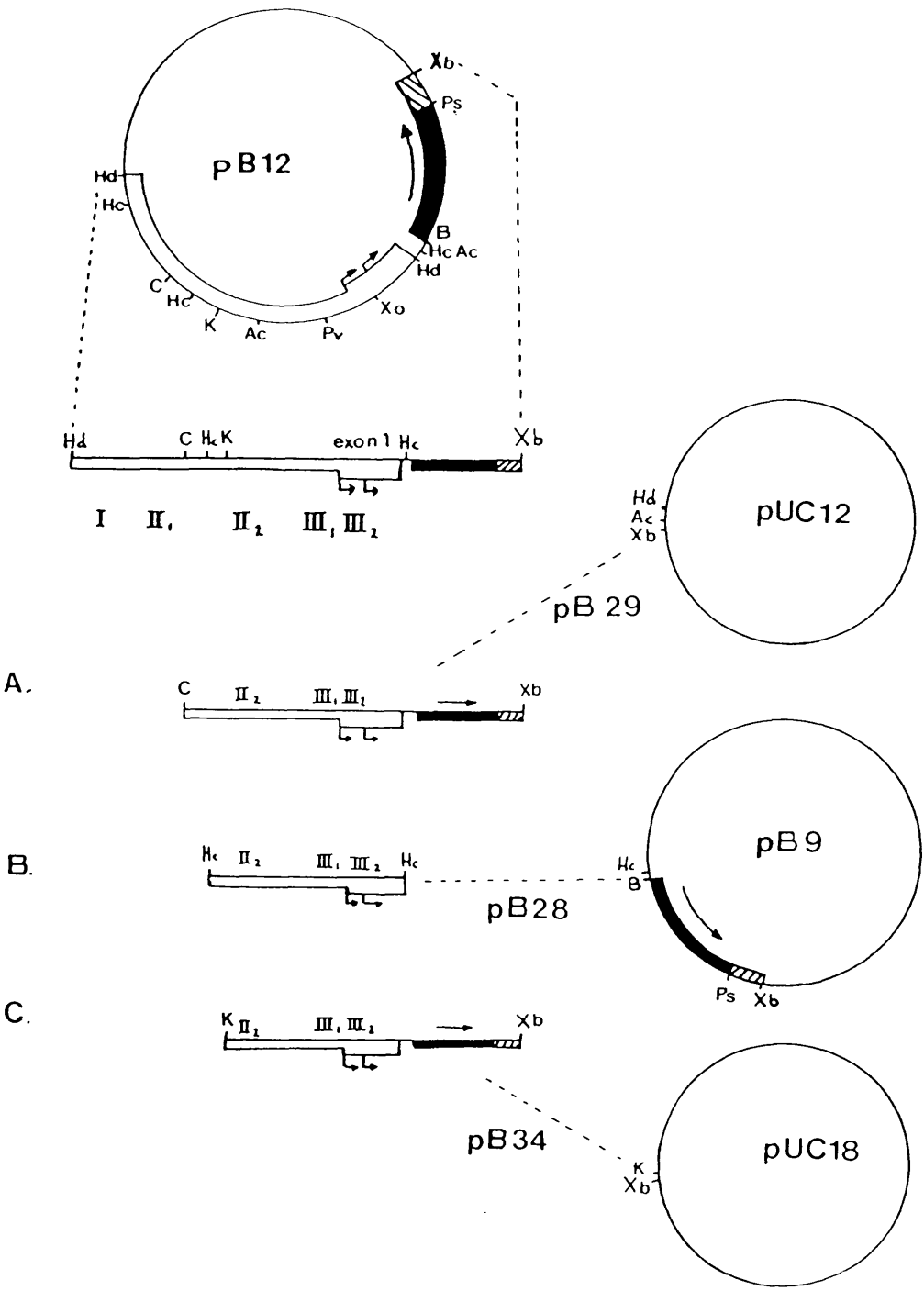


Figure 21.

Structure and Cloning Strategy of 5' Deleted c-myc-CAT Recombinant pB46, and Intermediate Recombinant pB45.

For a detailed description of the cloning strategy see section 3.2.1. See figure 8c and 26.

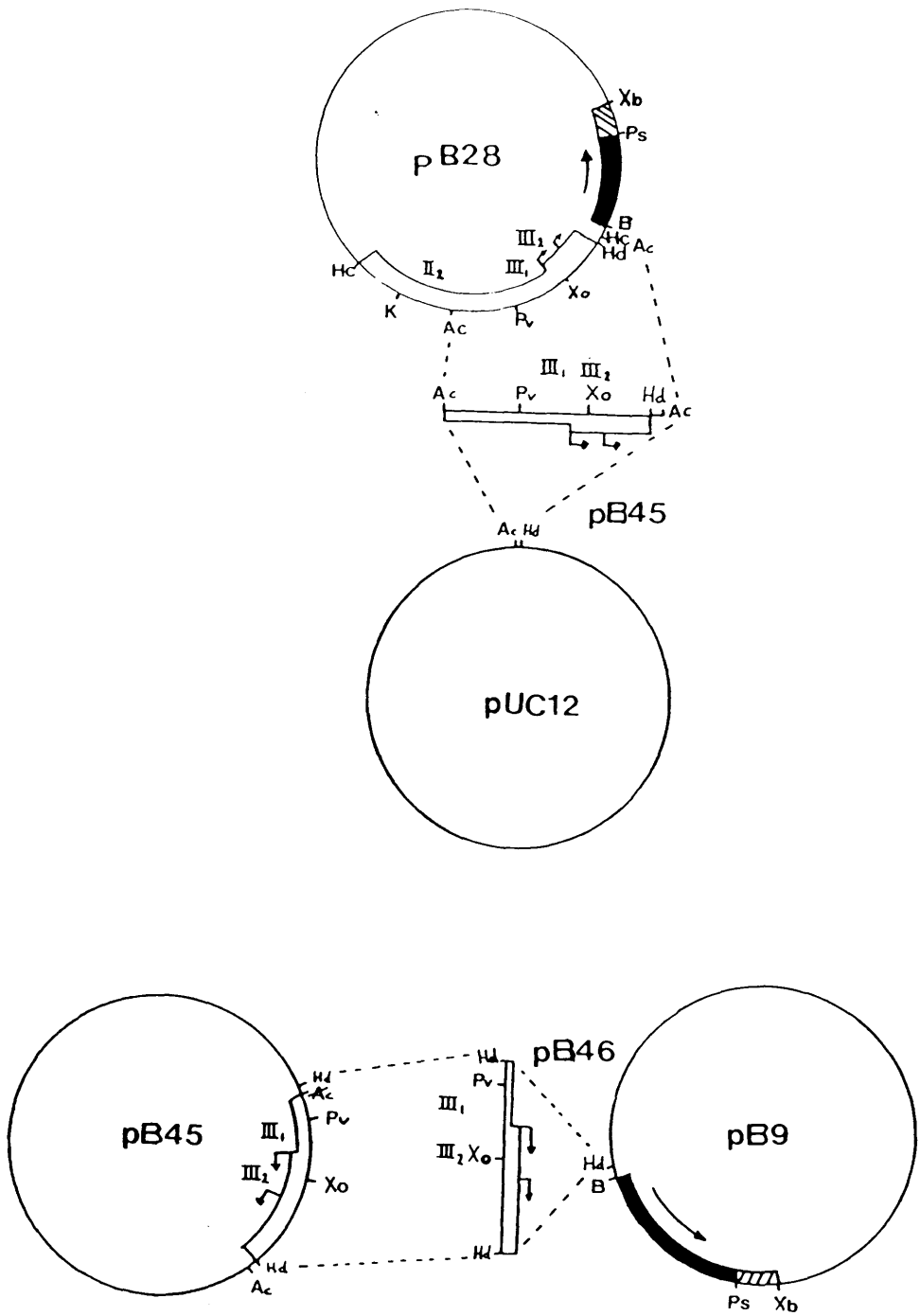


Figure 22.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB14.

For a detailed description of the cloning strategy see section 3.2.1. See figure 8c, 25 and 26.

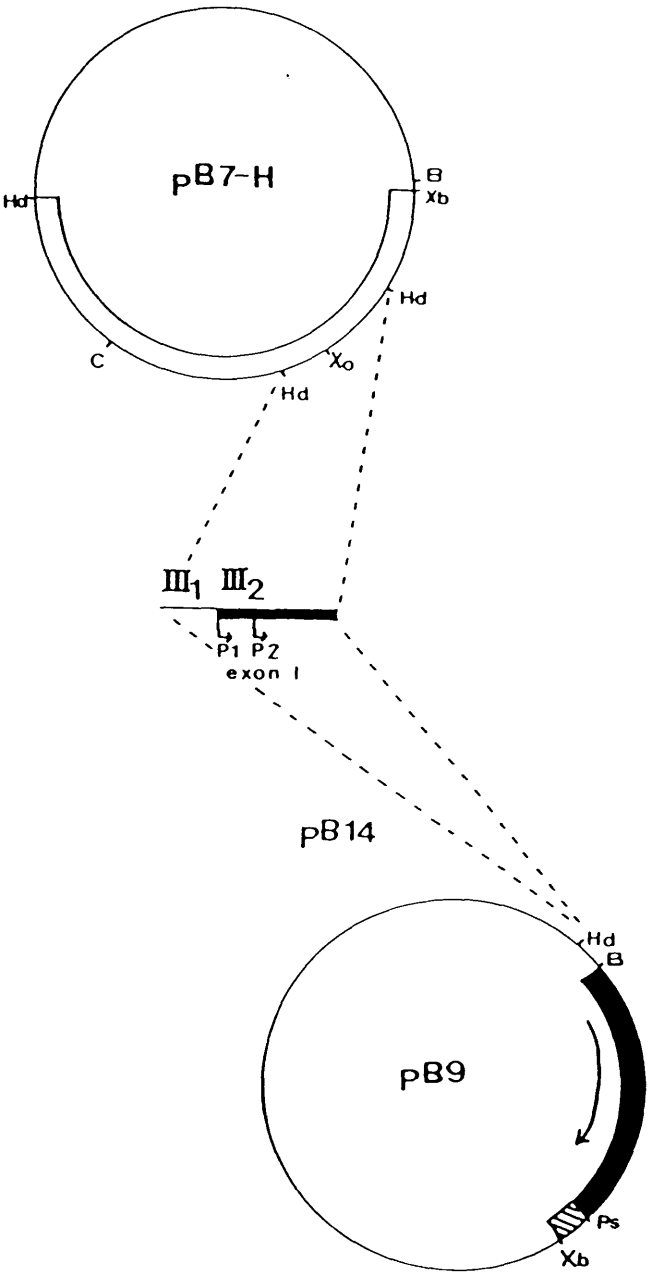
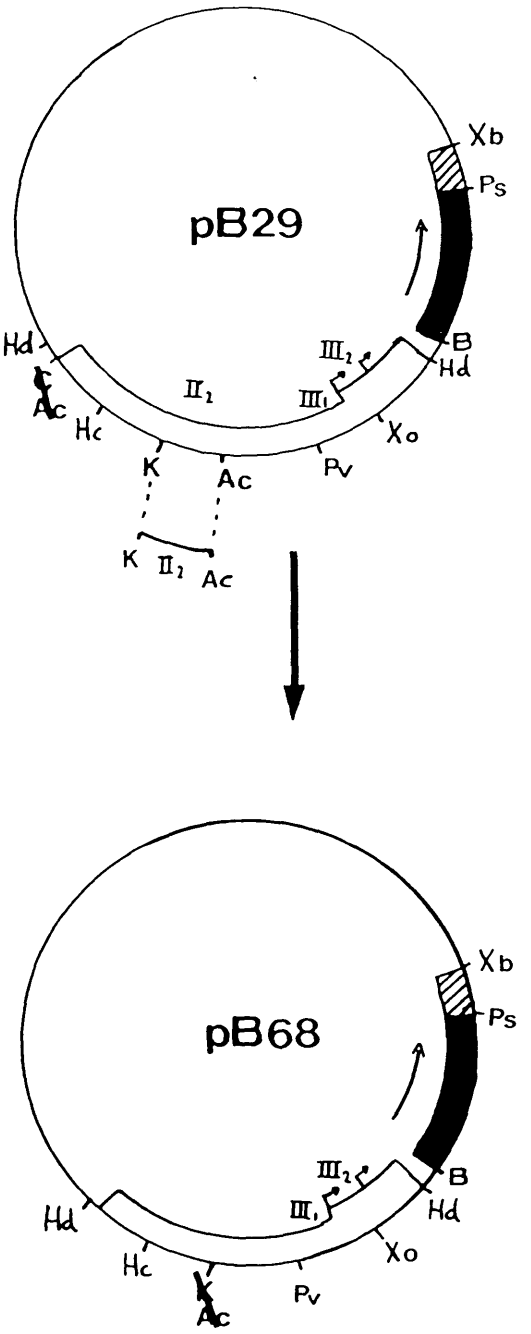
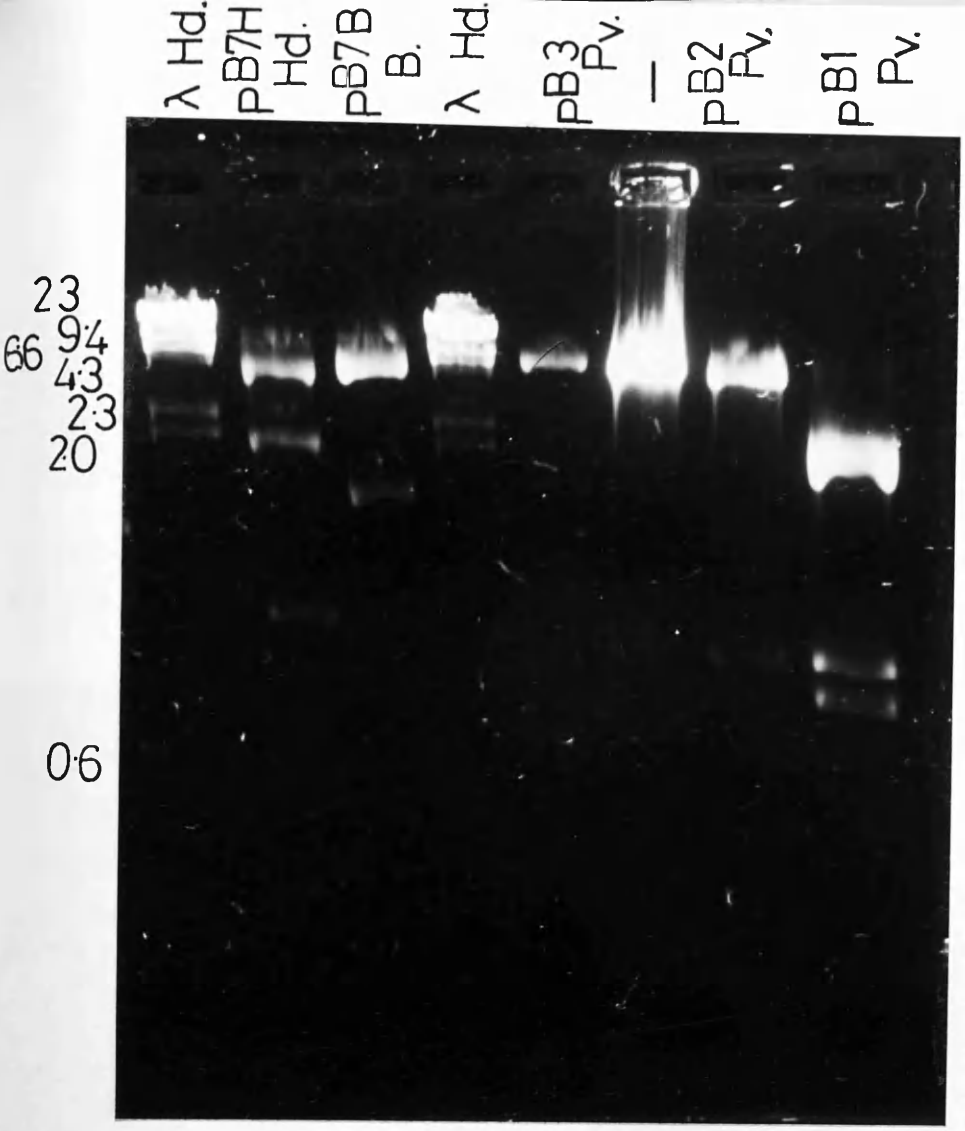


Figure 23.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB68.

Restriction sites which are destroyed upon ligation are marked with a diagonal line across them. For a detailed description of the cloning strategy see section 3.2.1. See figure 8d and 27.

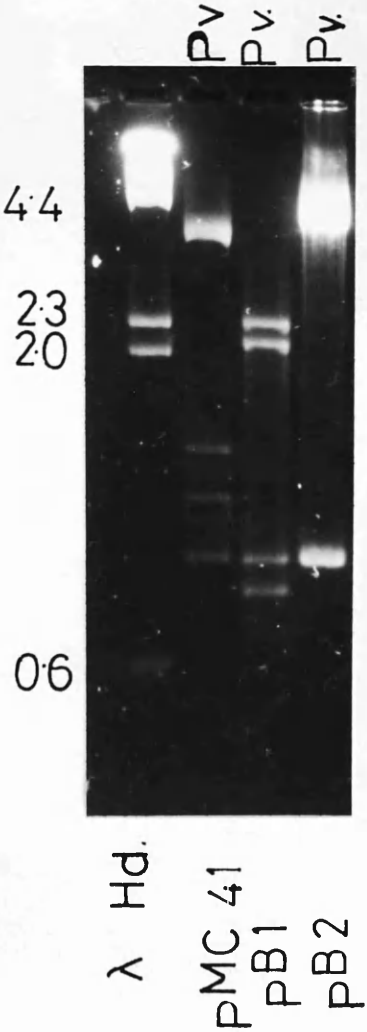




(a).

1%agarose gel.

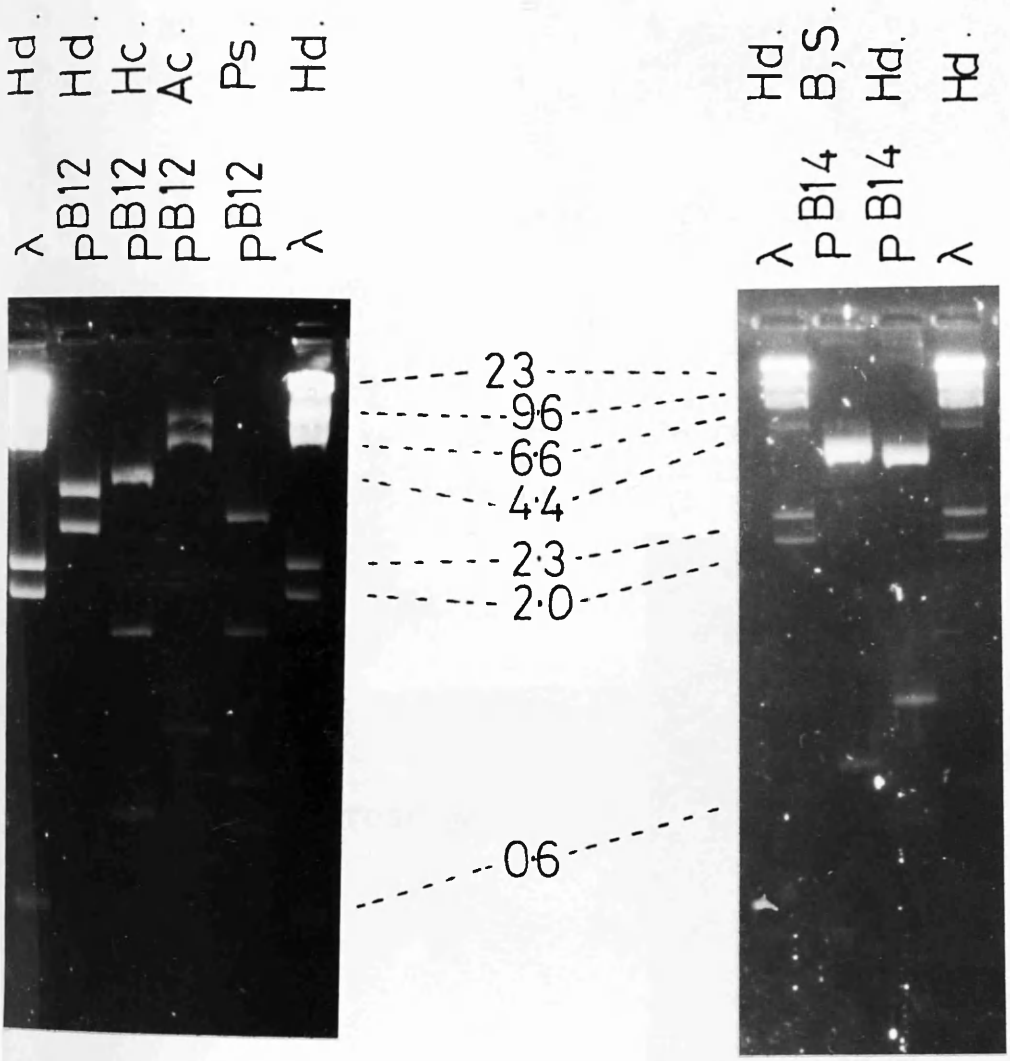
Figure 24.



(b).

1%agarose gel.

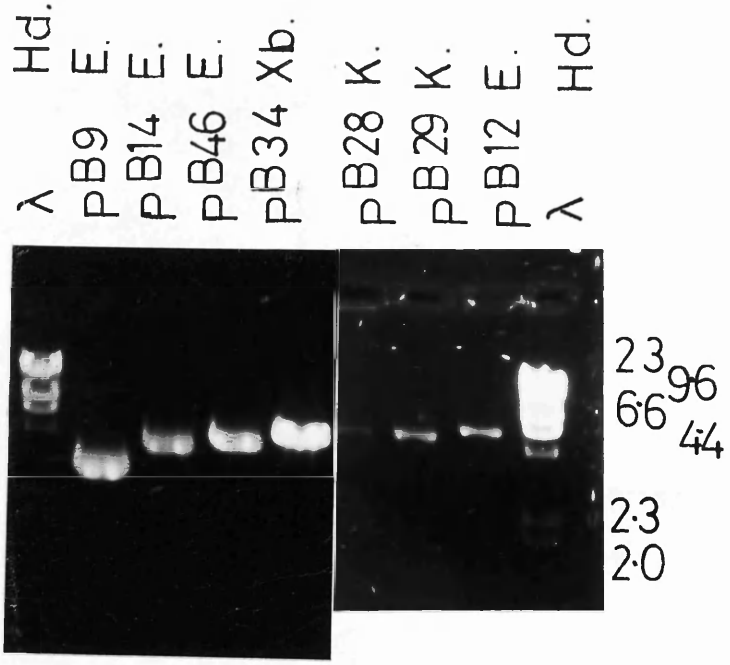
Figure 25.



1% agarose gel.

Figure 26.

(a).
0·8 % agarose gel.



(b).
1% agarose gel.

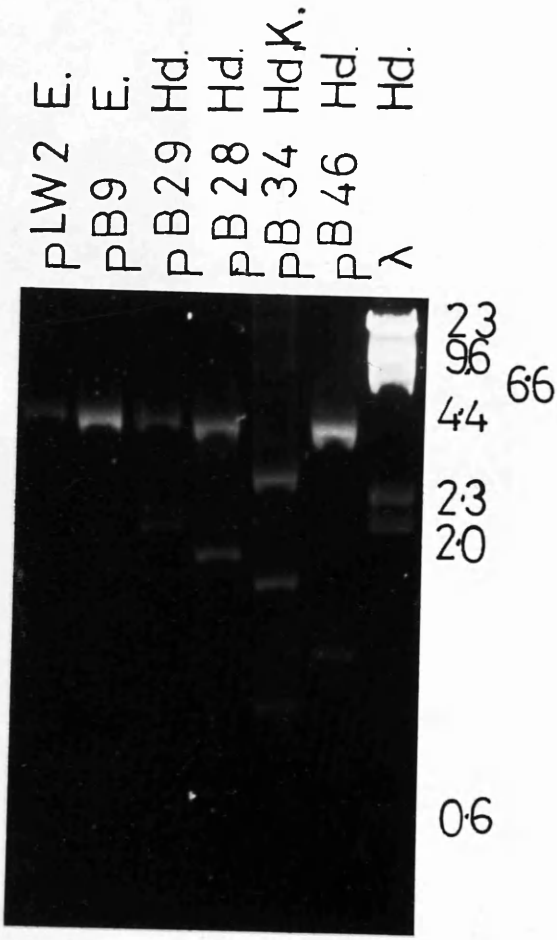
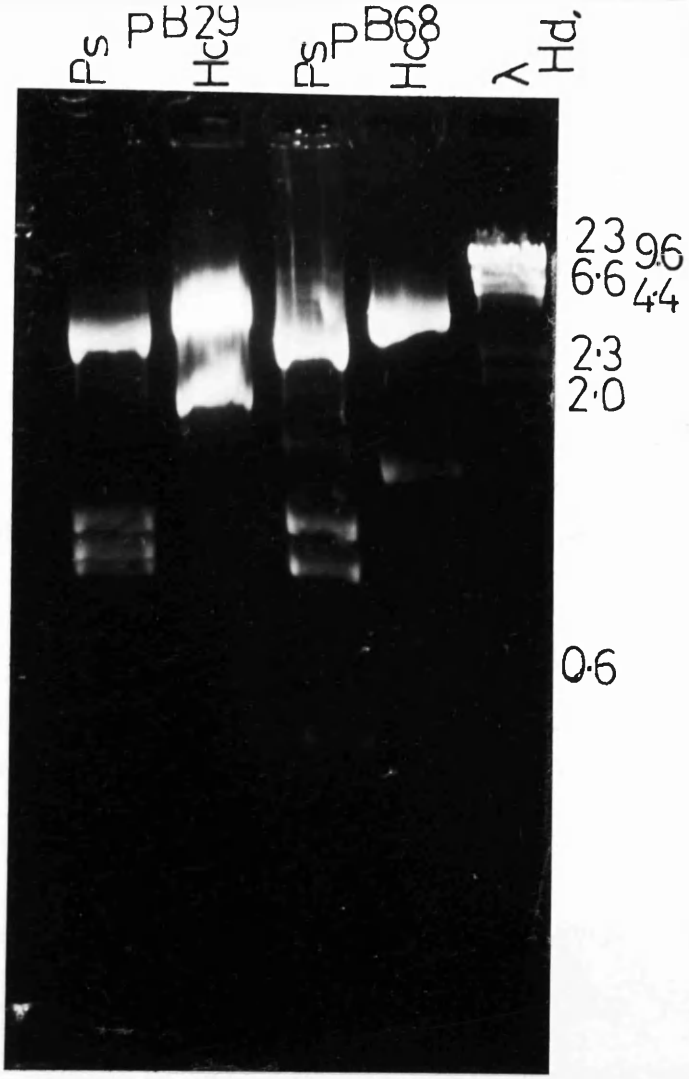
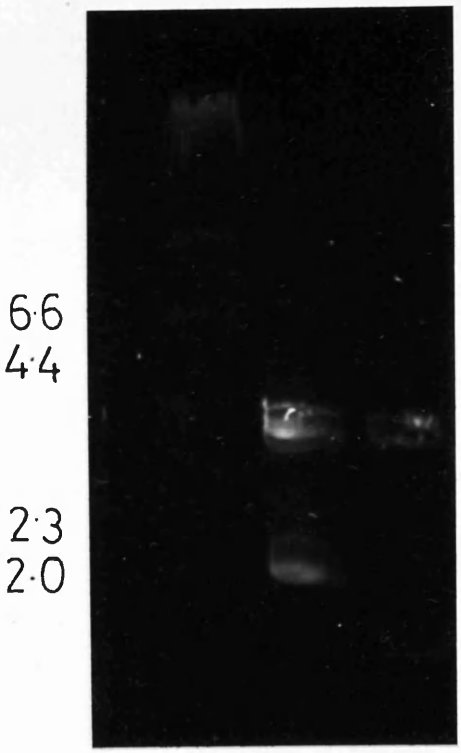


Figure 27.



(a).
1·2% agarose gel.



(b).
1·2% agarose gel.

Table 2.

Comparison of Basic c-myc and HSV-IE Promoter Recombinants Expression in Different Cell Lines.

Recipient Cell Lines.	Donor DNA pB9 [a]	[b]	pB14		pB12		pLW2		[c] expts.
LaTK⁻	1372	(1)	169523	(87)	50924	(26)	194022	(100)	8
HeLa	701	(2)	3100	(10)	805	(3)	30333	(100)	3
NIH-3T3	491	(10)	857	(17)	518	(10)	5149	(100)	3
K-562	1009	(8)	30053	(229)	3146	(24)	13117	(100)	3
MRC-5	491	(45)	1088	(100)	568	(52)	ND		2
GM468	496	(19)	2661	(100)	661	(25)	ND		2

a. Average of total counts (cpm) per flask (with no more than 15% variation).

b. Average counts (cpm) represented as a percentage of pLW2 (pB14 for MRC-5 and GM468).

c. Number of separate experiments of duplicate flasks.

d. Donor DNA (equimolar ratio of 10μg) was mixed with carrier salmon sperm DNA to a final concentration of 20μg/ml in the calcium-phosphate precipitate and transfected as described in Material and Methods.

e. CAT activity was assayed as described in Material and Methods using only 50μl of cell extract.

Figure 28.

Analysis of 5' Deleted c-*myc*-CAT Recombinants.

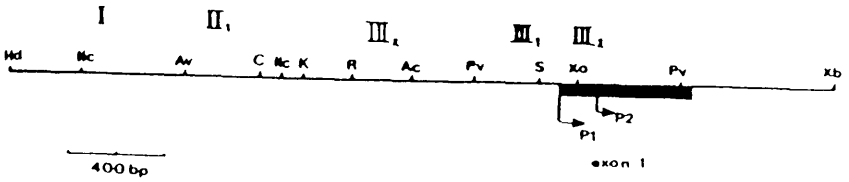
A. Diagrammatic representation of 5' deletions and relevant restriction map of c-*myc* 5' flanking region. The location of the DNase 1 *in vivo* hypersensitive sites of Siebenlist et al, (1984), are represented by Roman numerals. pB9 is the promoterless CAT recombinant (the negative control).

B. Histogram of the results shown in A.

a. For each plasmid an equimolar ratio of 5μg was transfected into either LaTK⁻ or HeLa recipient cells as described in Materialls and Methods.

b. Average total counts (cpm) from two to nine separate experiments of duplicate flasks, represented as a percentage of pB14.

A.



Donor DNA[a]		CAT Activity[b]	
		LaTK⁻	HeLa
pB12	I II1 II2 III1 III2	16·4	26·0
pB29	II2 III1 III2	23·6	
pB28	II2 III1 III2	14·6	
pB34	II2 III1 III2	36·8	40·3
pB46	III1 III2	109·3	
pB14	III1 III2	100	100
pB9		1·5	19·2

B

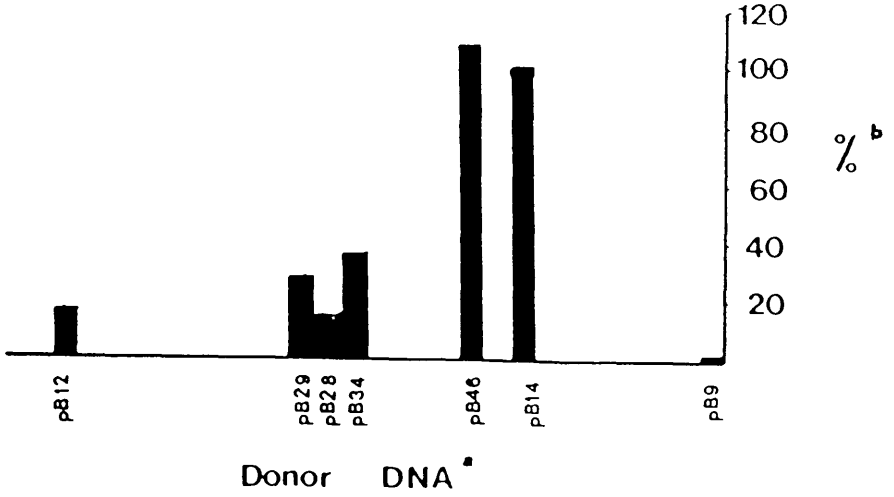
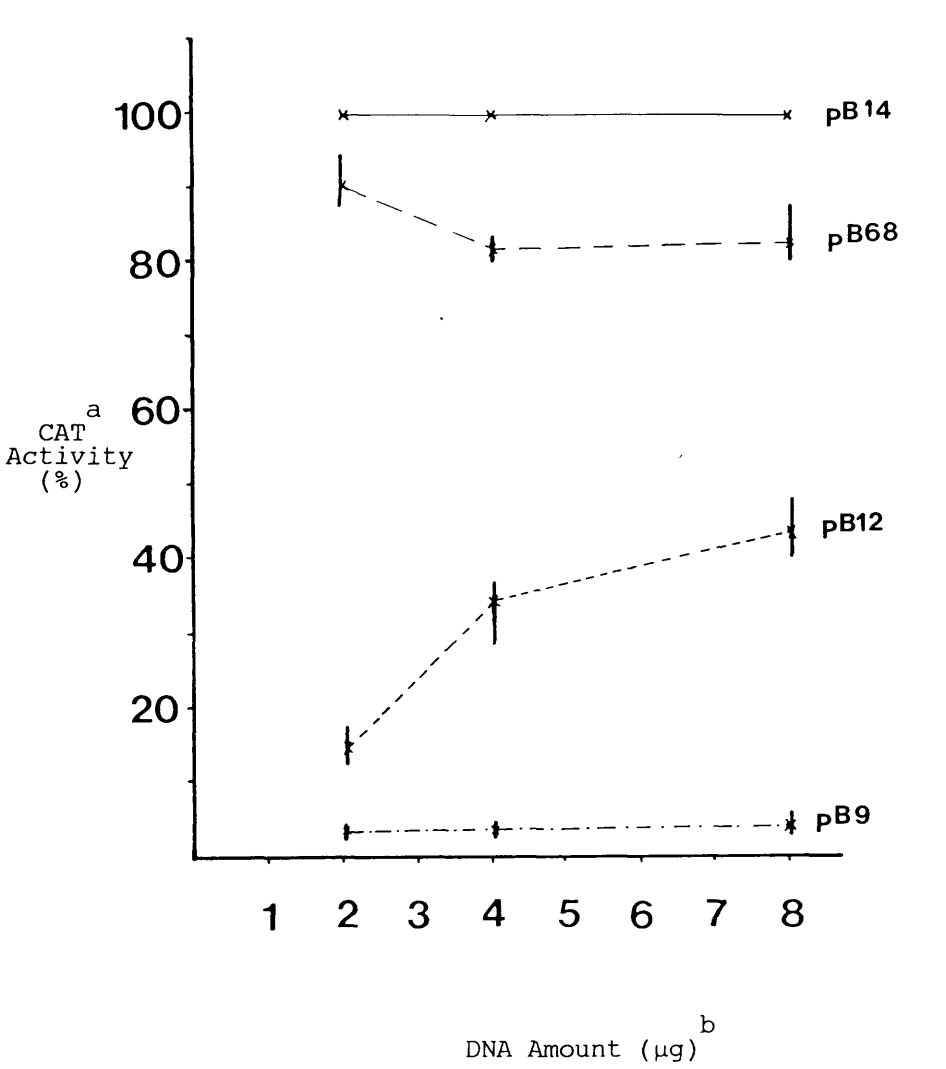


Figure 29.

Activity of Recombinants Containing an Internal Deletion of the NRE-2 Domain.

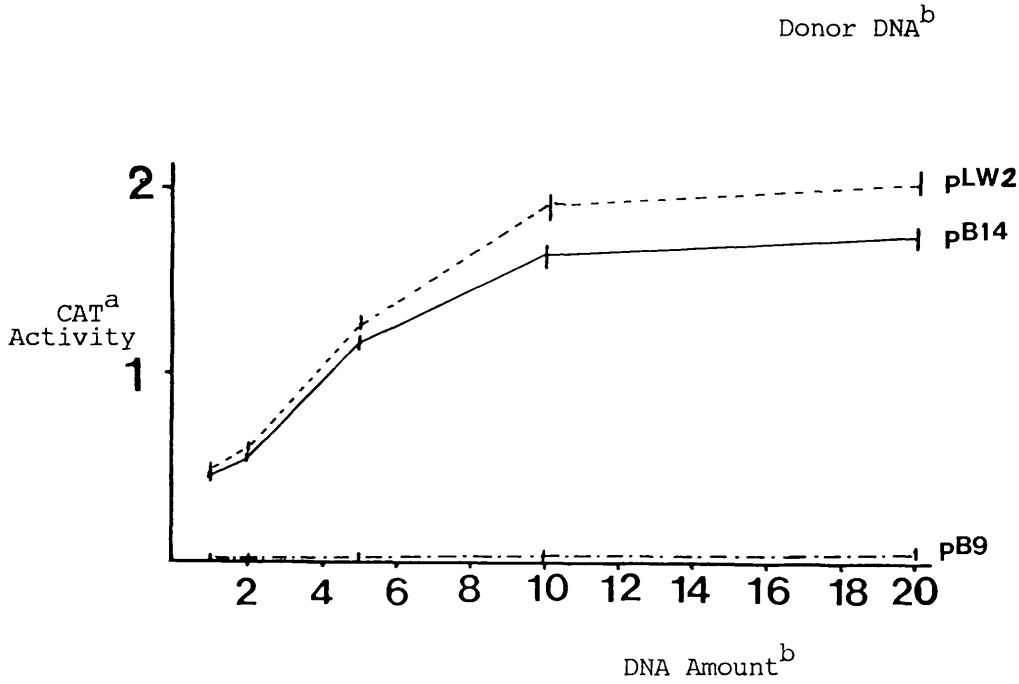


a. Average total counts (cpm) from two experiments of duplicate flasks as a percentage of pB14. Vertical bars indicate the range of values observed.

b. Equimolar amounts of each donor DNA (1-8μg) were transfected into LaTK⁻cells as described in Materials and Methods. The resultant CAT activity was assayed as described in Materials and Methods.

Figure 30.

DNA Titration of c-myc and HSV-IE Promoter Driven CAT Recombinants in LaTK⁻ Cells.



a. Average of total counts (10^5 cpm) from four separate experiments of duplicate flasks. Vertical bars indicate the range of values observed.

b. Equimolar amounts (1-10μg) of donor DNA were transfected into LaTK⁻ cells as described in Materials and Methods. CAT activity was assayed as described in Materials and Methods.

3.3. Characterisation of the Negative Regulatory Element

Transcriptional control elements have been classified by their specific characteristics with respect to magnitude of effect, activity on heterologous promoters, orientation and distance dependence from the transcription start site. The NRE-2 was analysed with respect to these functional parameters to give preliminary clues to how this element might function.

3.3.1. Recombinant Plasmid Constructions

The basic parental recombinants have already been described (see 3.1.2. and 3.2.1.). In addition, pB30 (see figure 31) was constructed by adding BamHI linkers to the PvuII site of pε01 (kindly given by J. Lang, Beatson Institute), and subsequent insertion of the 200bp BamHI fragment into BamHI digested pB9. This fragment contains the human epsilon globin gene promoter domain comprising of a CCAAT- and a TATA- element and the first 20bp of the mRNA 5' leader sequence (see figure 32; Efstratiadis et al, 1981). This construct, while containing a different promoter sequence, has essentially the same structural layout as pLW2 (see section 3.1.3.).

(i). Recombinants Containing c-myc 5' Region in Both Orientations

The 2.0Kb HindIII fragmennt of pB7H, containing DNase I HSS I, II-1, and II-2 (the NRE-2) was inserted, in both orientations into pB9 (pB10 and pB13), pLW2 (pB16 and pB17) and pB30 (pB37, only 5' to 3' orientation) all cut with HindIII as shown in figure 33. To allow the "flipping" of the upstream region in front of its homologous promoter, the 2.8Kb HindIII fragment from pB7B was first inserted into HindII digested pB9, giving pB12B (see figure 34). Then by inserting the 900bp BamHI fragment from pB12B into BamHI digested pB13, the same 5' region is "flipped" in pB19

compared to pB12 (see figure 35).

(ii). Recombinants Containing 445bp and 270bp NRE-2 fragments

By AccI digestion of pB34 and self-ligation of the resulting linear, pB51 (see figure 36) was constructed which contains the 445bp NRE-2 fragment (containing DNase I HSS II-2). EcoRI digestion of pB51 followed by Klenow treatment to produce flush-ends, and subsequent HindIII linker modification, allows this 445bp NRE-2 fragment to reside as a HindIII fragment (pB56; see figure 36). This fragment was inserted in both orientations into HindIII digested pB9 (pB72 and pB73; see figure 36) and only in the 3'-5' orientation into HindIII digested pLW2 (pB63; see figure 36). The DNase I HSS II-2 is located within the 3' portion of the 445bp fragment, and from analysis of the sequence, the 270bp sub-fragment (RsaI-ThaI) may define the extent of the NRE-2. Thus the RsaI-ThaI 270bp fragment from pB51 was inserted into the SmaI site of pUC12 destroying each respective restriction site upon ligation (pB49; see figure 37). EcoRI digestion followed by Klenow flush-ending, and HindIII linker modification, allowed this region to reside as a 300bp HindIII fragment (pB62; see figure 37). This HindIII fragment was then inserted in either orientation (pB65 and pB67; see figure 37) into HindIII digested pLW2.

(iii). Recombinant Containing Splice Donor Site

The 3.6Kb HindIII-BamHI fragment from pB2 containing all the 5' region, all of exon 1 and 600bp of intron 1 (to the XbaI site), was inserted into HindIII-BamHI digested pB9 to give pB4 (see figure 38).

3.3.2. Restriction Enzyme Analysis of Recombinant Plasmids

Extensive restriction enzyme analysis has been used to identify each of the recombinant plasmids described.

Representative digests are shown in figures 39, 40, 41, 42, 43, 44 and relevant restriction maps in figures 8a, 8d, 8e, 8f, 8g, 8h.

(i). Plasmids Containing 5' Region and the ε-Globin Promoter.

The basic ε-globin promoter-CAT recombinant (pB30) is defined by HindIII linearalisation (3.8Kb), BamHI digestion to reveal the 200bp promoter domain, and PstI digestion to show the promoter-CAT fragment (985bp).

The recombinants containing the 1940bp c-myc upstream region 5' to heterologous promoter domains driving the CAT gene are defined by restriction enzymes which confirm the orientation of the c-myc derived sequences. PstI digestion of pB10 and pB37 give the diagnostic fragment of 60bp, while a 180bp fragment is seen for pB13. The 1710bp fragment is derived from the inserted c-myc sequences. Because pB37 contains the 200bp ε-globin promoter domain the CAT fragment is larger (990bp) than that seen for pB10 or pB13 (790bp). HincII digestion of pB16 gives a diagnostic fragment of 800bp, while pB17 gives a 310bp fragment. The 830bp fragment is derived from the inserted c-myc sequences.

EcoRI linearalisation of pB19 and pB12 confirms their near identical (only 10bp difference) size of 6.4Kb. To confirm the c-myc upstream region is in the "flipped" oreintation in pB19 compared to pB12, both recombinants were digested with HincII to give fragments of 4365, 875, 830 and 310bp for pB19, and 3865, 1670 and 830bp for pB12. The 875bp and 310bp fragments of pB19 confirm the orientation of the upstream sequences.

(ii). Plasmids Containing NRE-2

The 445bp NRE-2 fragment is shown in pB51 by HindIII KpnI digestion, in pB56 by HindIII digestion (apparent

disparity in size is due to linker modification), and in pB72 by HindIII digestion. AccI-KpnI digestion of pB34 indicates the size of the original NRE-2 fragment (445bp). The orientation of the NRE-2 fragment within pB51 was confirmed by HinfI digestion. The 62bp fragment derived solely from c-myc sequences, while the 1454 and 544bp fragments indicating the insert oreintation. PstI digestion of pB72 giving fragments of 3240, 790 and 10bp, and of pB73 giving fragments of 2780, 790 and 460bp confirms the NRE-2 domain orientation in these recombinants. Since PstI digestion of pB63 reveals a 470bp fragment, similar to pB73, this confirms that the NRE-2 domain in these two recombinants is the same.

The 270bp sub-fragment is shown in pB49 by EcoRI HindIII digestion and in pB62 by HindII digestion (apparent disparity in size is due to linker modification). This fragment was inserted into pLW2, where its orientation was confirmed by BamHI digestion: giving 3560, 300 and 210bp (HSV-IE promoter domain) in pB65, and 3840, 210 and 20bp in pB67.

(iii). Plasmid Containing the Splice Donor Site.

The recombinant containing the splice donor site pB4 is defined by ClaI linearalisation (7.3Kb), and XbaI-HindIII digestion which reveals the pUC fragment (2680bp) and the c-myc-CAT fragment (4665bp, compare to HindIII digested pB12; see figure 25).

3.3.3. Effect on Heterologous Promoters

The ability of the NRE-2 to regulate (repress) heterologous promoters was investigated. Since the greatest repression was observed when the HindIII-PvuII 2.0Kb region was deleted, this region was used in these experiments. Two heterologous prmoters were used in these experiments; the HSV-IE gene promoter (viral) and the human ε-globin gene

promoter (eukaryotic). Equimolar amounts of pLW2 (HSV-IE promoter), pB30 (ε-globin promoter), pB14 (c-myc promoter) and the respective NRE-2 driven promoter-CAT recombinants (pB16, pB37, pB12) plus the negative control pB9, were transfected into LaTK$^-$, HeLa, NIH-3T3 and K-562 recipient cells. The resultant CAT values are shown in table 3 (see figure 45).

The 2.0Kb HindIII-PvuII 5' region (containing DNase I HSS I, II-1 and II-2) confers a 6-fold repression upon its own homologous promoter when orientated in the same direction (5'-3') to it, in LaTK$^-$cells. When placed in the same orientation in front of either the HSV-IE (pB16) or ε-globin (pB37) gene promoters their respective CAT activities are only halved in LaTK$^-$cells. However, in Hela cells the heterologous HSV-IE promoter (pB16) is repressed 9-fold, while repression of the homologous promoter (pB12) is only 4-fold. These differences in repression ratios have no obvious explanation, except that they may reflect the different cell lines ability to express the exogenously introduced viral promoter, or reflect a titration effect (see 3.4.4.). The repression seen in NIH-3T3 cells is approximately the same for the homologous (pB12) and heterologous (pB16) promoters. Therefore the 2.0Kb HindIII-PvuII NRE-2 region can repress heterologous promoters, although not as efficiently as the homologous promoter, in this assay system.

3.3.4. Orientation Dependence of Repression

To investigate the orientation dependence of the NRE-2, the 2.0Kb HindIII-PvuII 5' region (containing DNase I HSS I, II-1 and II-2) was inserted in either orientation (pB12 and pB19) 5' to its homologous promoter (pB14) and 5' to the heterologous HSV-IE promoter (pB16,pB17; pLW2). Equimolar ratios of these recombinants were transfected into LaTK$^-$, HeLa, NIH-3T3 and K-562 cells and the resultant CAT activities are shown in table 3. When the 2.0Kb NRE-2 region

was placed 5' to its homologous promoter, the 3'-5' orientation (i.e. the "flipped" orientation) consistently gave the greater magnitude of repression in LaTK$^-$cells, the same level of repression in K-562 and NIH-3T3 cells, and slightly less repression in Hela cells. When regulating the HSV-IE gene promoter (pBl6 and pBl7) in both LaTK$^-$ and NIH-3T3 cells, the 3'-5' orientation again showed the greater level of repression.

These orientation experiments were done on covalently closed plasmids, so the activity seen when the c-myc fragment containing the NRE-2 is in the "flipped" orientation could be due to its function operating the "long-way-round" the plasmid. However, the mere transfection of linear plasmids does not exclude this either, due to the large probability that there is immediate recombination and recirculisation of transfected DNA on entering the cell (Weber and Shaffner, 1985), or possibly even before while in the co-precipitate stage.

Why the "flipped" orientation in some instances causes a greater repression is not obviously clear. This can not be due directly to the positioning of the NRE-2 element as it is in the 3' region of this fragment, i.e. the "flipped" orientation has it distal to the promoter. The presence of an additional negative element, although not seen in the 5' deletion studies, can not be excluded (see 3.6.4.).

Thus, using the 2.0Kb HindIII-PvuII NRE-2 fragment, repression can be conferred upon heterologous promoters, both of viral and eukaryotic origin, in an apparently orientation independent manner. Repression ratios vary between 2- and 10- fold depending on both the orientation and cell line being analysed. These results imply that there is no essential interaction with c-myc sequences located outwith this region required for this repression, however, these studies do not provide evidence to say such an interaction definitely does not occur.

3.3.5. Effect of Varying Distance

These results came from the strategy used to construct previously described plasmids, and the various experiments in which they were analysed (see table 3). As a consequence of the cloning strategy, the negative element is in a different location with respect to P1 in each recombinant, from -530bp (pBl6) and -700bp (pBl2) to -1670bp (pBl7) and -1840bp (pBl9). This would imply that the exact positioning of the NRE-2 with respect to the transcription start site is relatively unimportant at these distances. However, since in pBl7 and pBl9 the c-myc sequences have been "flipped", there are other explanations for this (see 3.6.4.), and this conclusion has to be treated with some reserve.

3.3.6. Analysis of NRE-2 Fragments

To further analyse the repression function the negative element (NRE-2), as defined by the deletion analysis, was inserted as a 445bp KpnI-AccI or 270bp RsaI-ThaI fragment into pLW2 and transient transfection experiments using equimolar amounts of the respective recombinants were performed. This set of experiments was done in LaTK$^-$ cells and the results are presented in table 4.

When placed as a 445bp fragment (pB63) or a 270bp fragment in either orientation (pB65 and pB67) upstream of the HSV-IE gene promoter no repression of CAT activity compared to the parental recombinant (pLW2) was observed. This may indicate, that although this region is essential for the repression activity, it does not contain the whole negative regulatory element. As a corollary, sequences outwith the KpnI-AccI fragment are required for it to function. Alternatively the relative positioning of the NRE-2 may be of crucial importance when proximal to the promoter, being in these constructs, too close to the promoter domain, or in the wrong local helical twist to function. Analysis of this anomaly requires detailed

examination in the future.

3.3.7. Effect of the NRE-2 on a Promoter-less CAT Gene Alone

The recombinant pB9 contains no obvious promoter domain, nor a cap site, However, there is a very weak level of background activity (see 3.2.3.). When enhancer elements are placed upstream of a promoterless gene, transcriptional activation of that gene occurs (Lang, 1985). The NRE-2 was analysed for such activity. The 2.0Kb NRE-2 fragment and the 445bp NRE-2 fragment were inserted, in both oreintations, 5' to the CAT structural gene and transfected into LaTK$^-$ and K-562 cells as before. The resultant CAT activity is presented in table 5.

When the 2.0Kb HindIII-PvuII upstream c-myc region (containing DNase HSS I, II-1 and II-2) was inserted upstream of pB9 in either orientation (pB10 and pB13) no affect on CAT enzyme activity was observed. However, when the 445bp fragment (KpnI-AccI), which contains at least an essential domain of the negative element, is inserted in either orientation 5' to the promoterless CAT gene (pB72 and pB73) there is an approximately 15-fold increase in CAT activity when compared to pB9. This level of activity is nearly half that of the c-myc promoter (pB14). The NRE-2 (the 445bp KpnI-AccI fragment) in these experiments is acting to promote expression.

These results may indicate that there are additional sequences outwith the 445bp fragment required to bring about repression. Whether pB10 and pB13 give no activity due to an intact functioning negative regulatory element or that the element(s) conferring promotion are too distant from the promoter to have any activity is not known. When viewed as a whole the results presented (5' deletions, internal deletion and heterologous promoters) indicate that this 445bp fragment contains an essential region but perhaps not the whole domain of the negative regulatory element, NRE-2,

which can in specific situations act to promote expression.

3.3.8. Presence of Splice Donor Site

The recombinant (pB4) which contains all the c-myc 5' sequences, all of exon 1 and 600bp of intron 1 was originally made as a negative control. To test this recombinant (pB4) for activity, equimolar amounts (of 10μg) of it and pB9 were transfected into LaTK⁻and K-562 cells, and the resultant CAT activity shown in table 5. The CAT activity for pB4 is zero when compared to that of pB9. This was expected as the presumed splice donor site is retained in the construct while no obvious splice acceptor site is present. Presumably the CAT structural gene is spliced out, thus not transcribed (no data). Several other studies (D. Spandidos and N. Lowdnes, personal communication) and elsewhere (Remmers, Yang and Marcu, 1986) also retain splice donor sites in their constructs (c-Ha-ras and c-myc respectively) with no obvious effect on expression. This may be explained for the c-Ha-ras recombinants in that it is another system, but the c-myc recombinants (Remmers, Yang and Marcu, 1986) present a definite anomaly. A much smaller intronic region (approximately 20bp) is retained in their recombinant (compared to 600bp in pB4) and it is conceivable that the correct mRNA structure required for splicing can not therefore be attained.

The other obvious difference between pB4 and pBl2, is the inclusion of the 3' 40bp of exon 1. This full length untranslated exon 1 may be involved in messenger stability (Piechaczyk et al, 1985), and as such its inclusion may confer an extreme instability upon the CAT induced message. If this was the case again Remmer's recombinants would (but do not; Remmer, Yang and Marcu, 1986) show the same lack of CAT enzyme activity. Thus this anomaly remains unresolved, but could be a point for future investigation.

The relatively high c-myc-driven CAT activity in these

studies (compare recombinants pB14 and pLW2; see table 2) implies the presence of a fairly abundant CAT protein. Thus the inclusion of 500bp (the majority) of the c-myc exon 1 sequences probably does not confer an extreme instability upon the resulting hybrid message.

Figure 31.

Structure and Cloning Strategy of ε- Globin Promoter-CAT Recombinant pB30.

For a detailed description of the cloning strategy see section 3.3.1. See figure 8a, 32, 39 and 40.

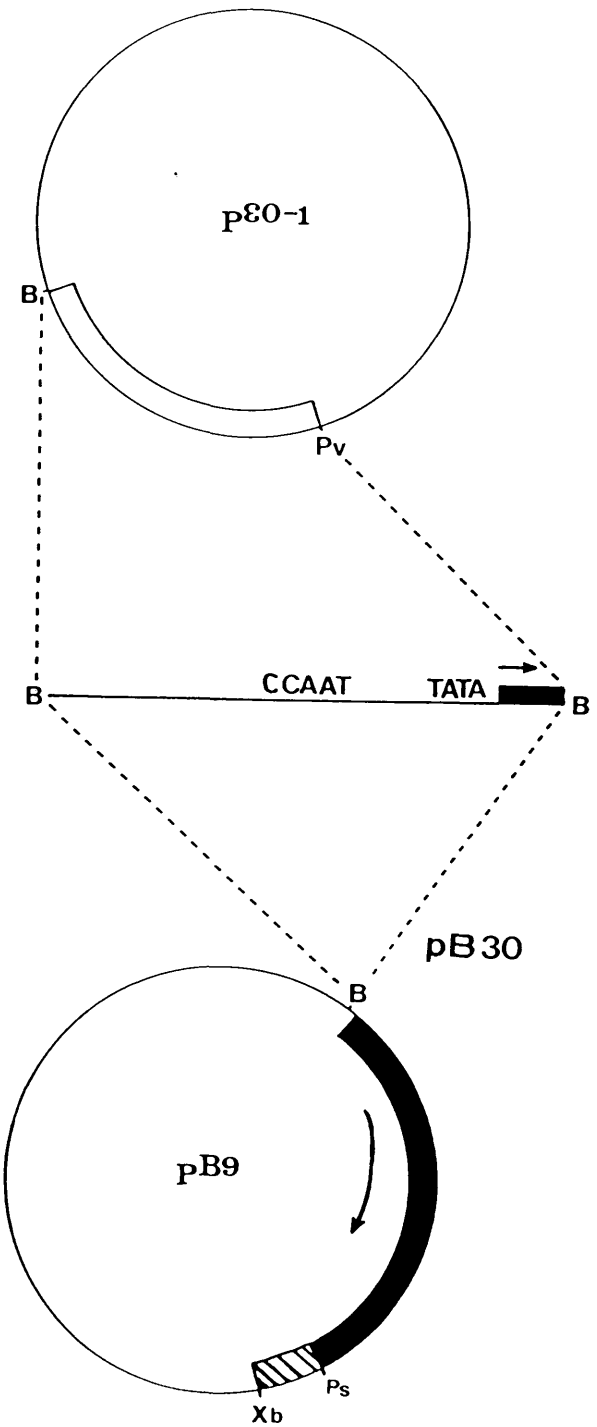
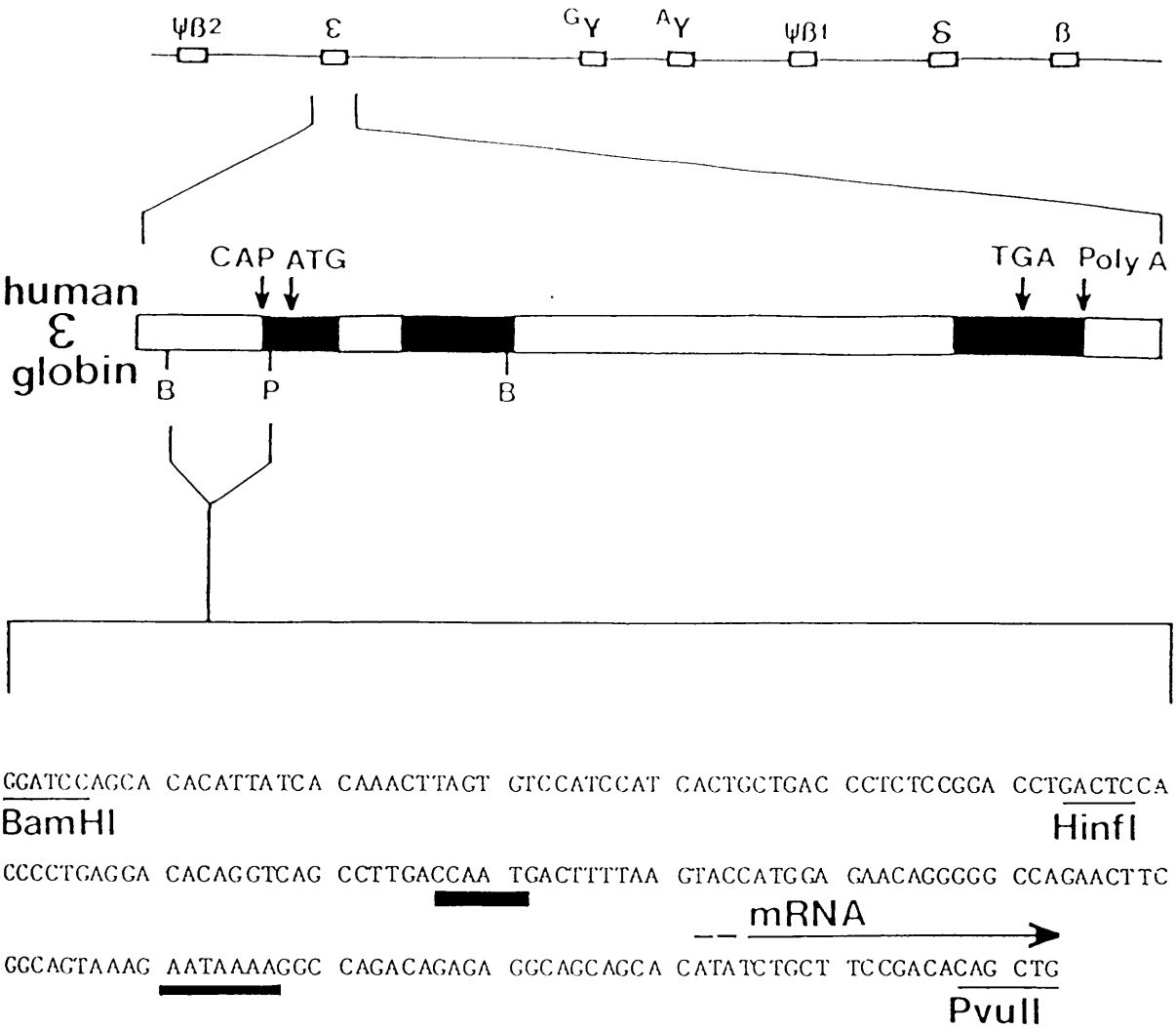


Figure 32

Structure of the Human Globin Gene Family and Derivation of the ε-Globin Gene Promoter Domain.



The position of the human ε globin gene is depicted within the β-like globin gene family, and the DNA sequence of the 197 bp Bam HI-Pvu II promoter region shown. Shaded areas represent the coding regions of the gene separated by intervening introns. Underscored black bars represent the CAAT and TATA consensus sequences.

See figure 31.

Figure 33.

Structure and Cloning Strategy of Recombinants pB10, pB13, pB16, pB17 and pB37.

For a detailed description of the cloning strategy see section 3.3.1. See figure 8e and 40.

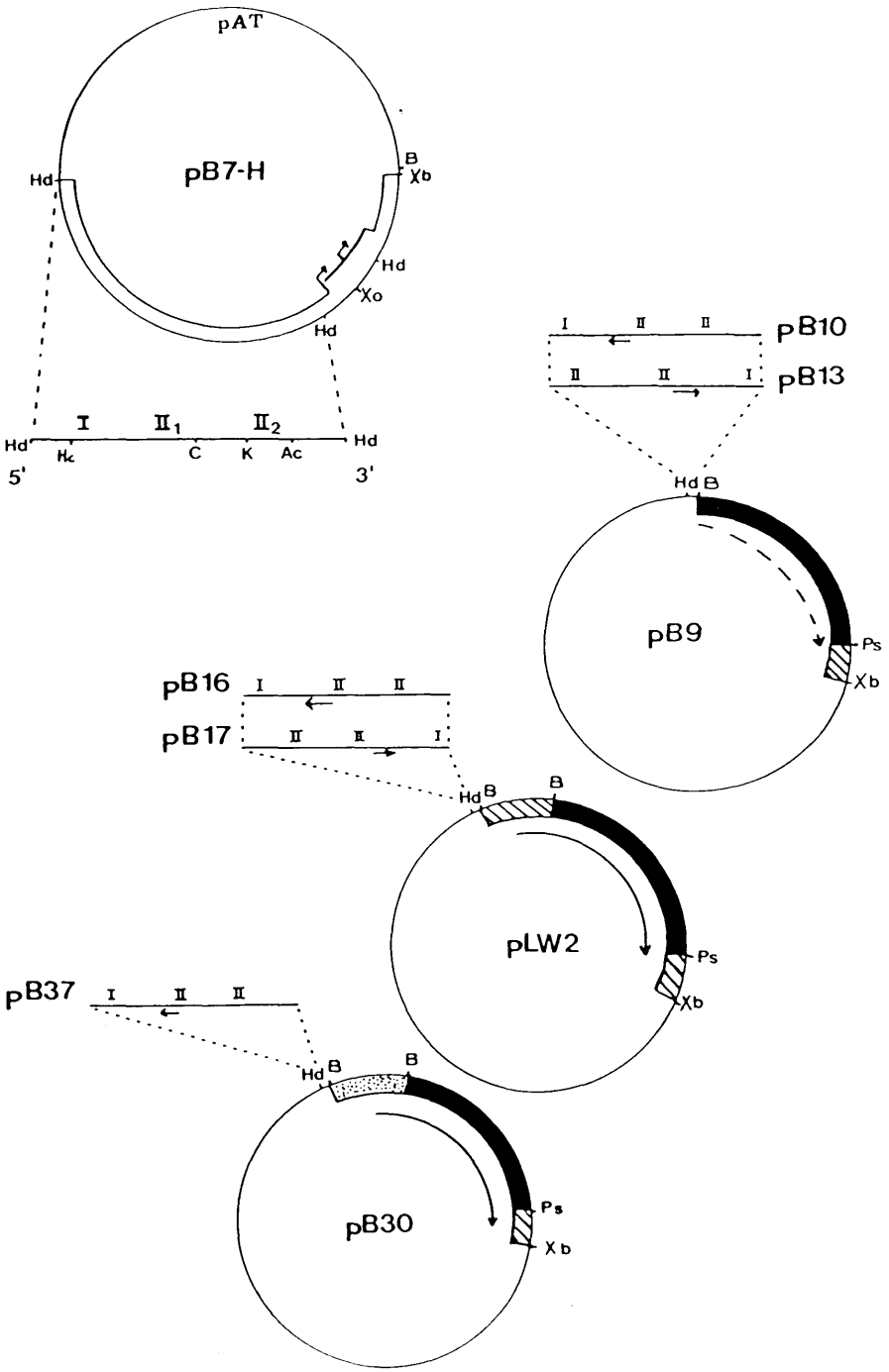


Figure 34.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB12B.

For a detailed description of the cloning strategy see section 3.3.1. See figure 8d.

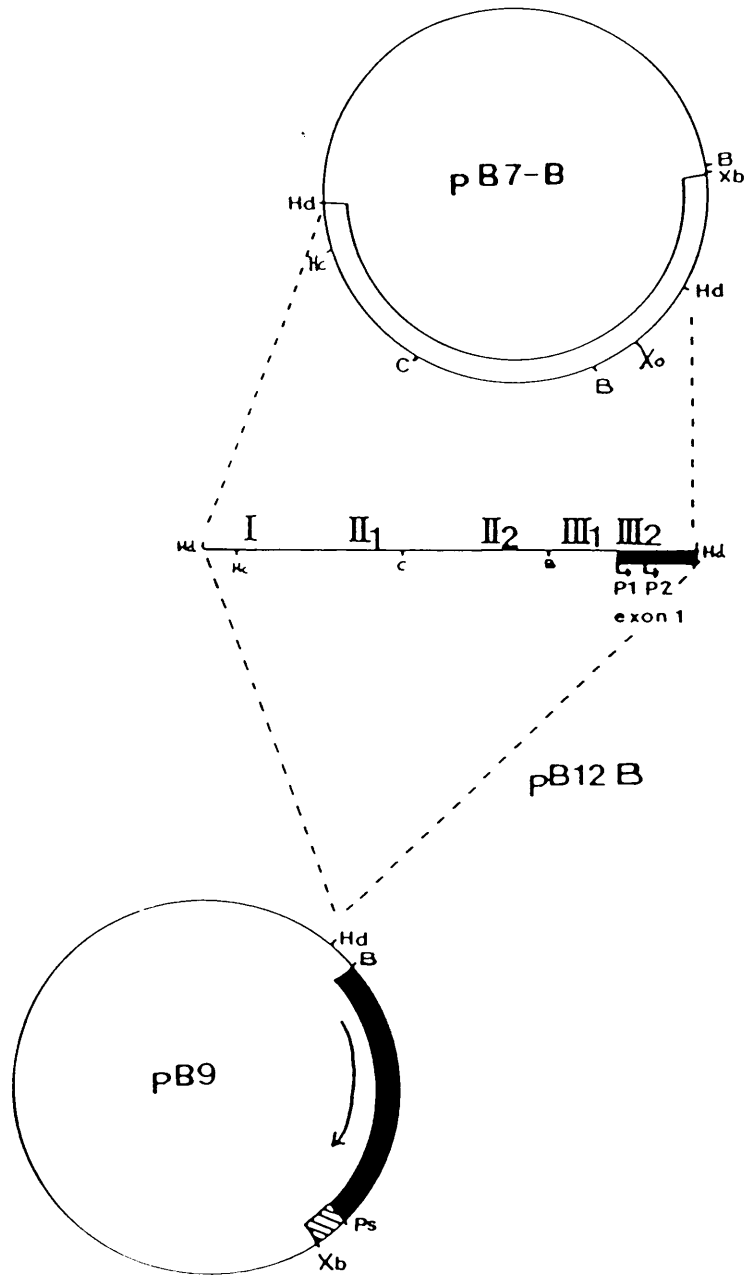
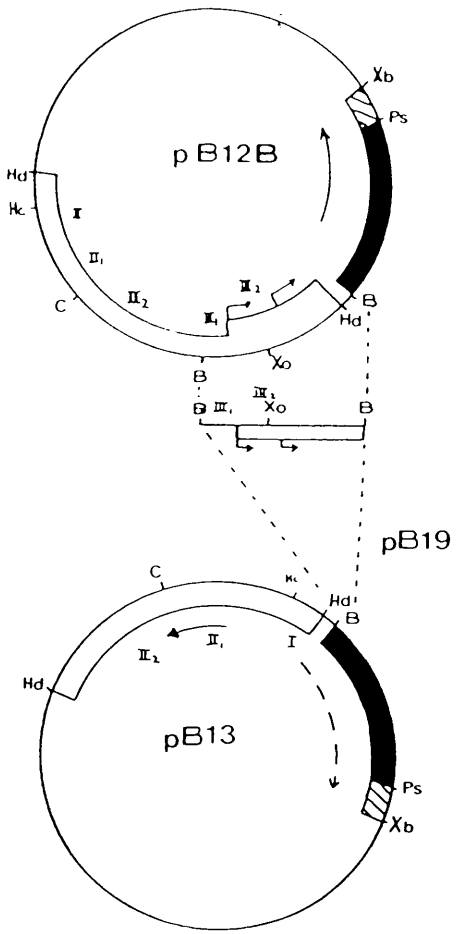


Figure 35.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB19.

The comparable recombinants pB12 and pB19 are shown diagrammatically at the bottom of the page. In effect the major difference between them is the "flipping" of the 2.0Kb HindIII-PvuII upstream region. For a detailed description of the cloning strategy see section 3.3.1. See figure 8d, 19, 33 and 41.



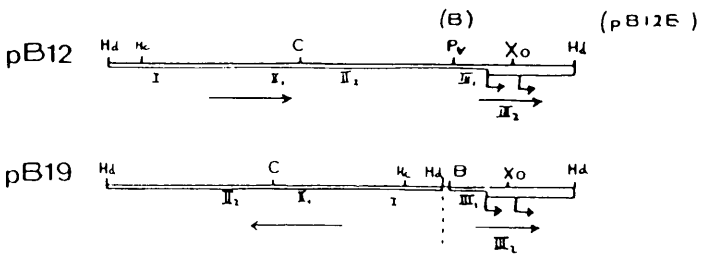


Figure 36.

Structure and Cloning Strategy of Recombinants pB51, pB56, pB62, pB72 and pB73 Containing the 445bp NRE-2 Region.

Restriction sites which are destroyed upon ligation are marked with a diagonal line across them. For a detailed description of the cloning strategy see section 3.3.1. See figure 8f, 42, 43 and 47.

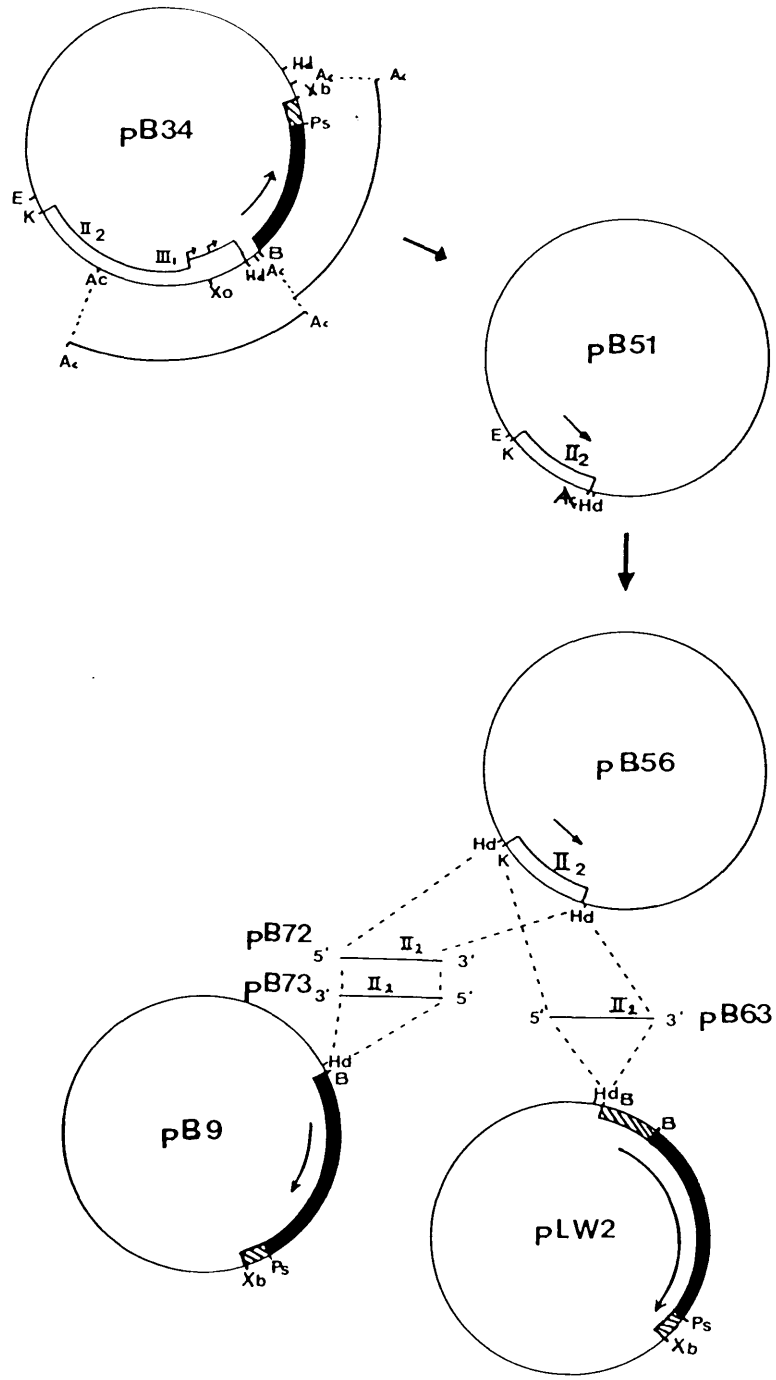


Figure 37.

Structure and Cloning Strategy of Recombinants pB49, pB62, pB65 and pB67 Containing the 270bp NRE-2 Sub-Region.

Restriction sites which are destroyed upon ligation are marked with a diagonal line across them. For a detailed description of the cloning strategy see section 3.3.1. See figure 8h, 36 and 44.

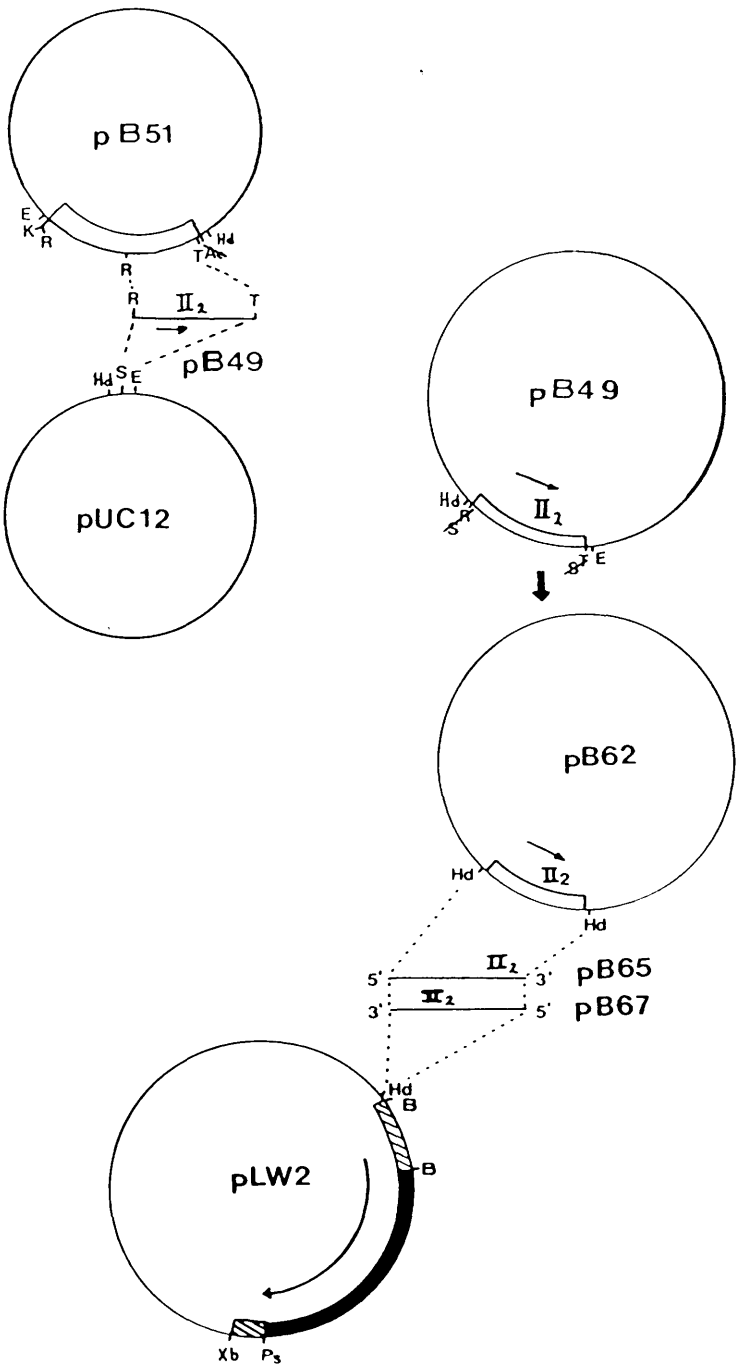
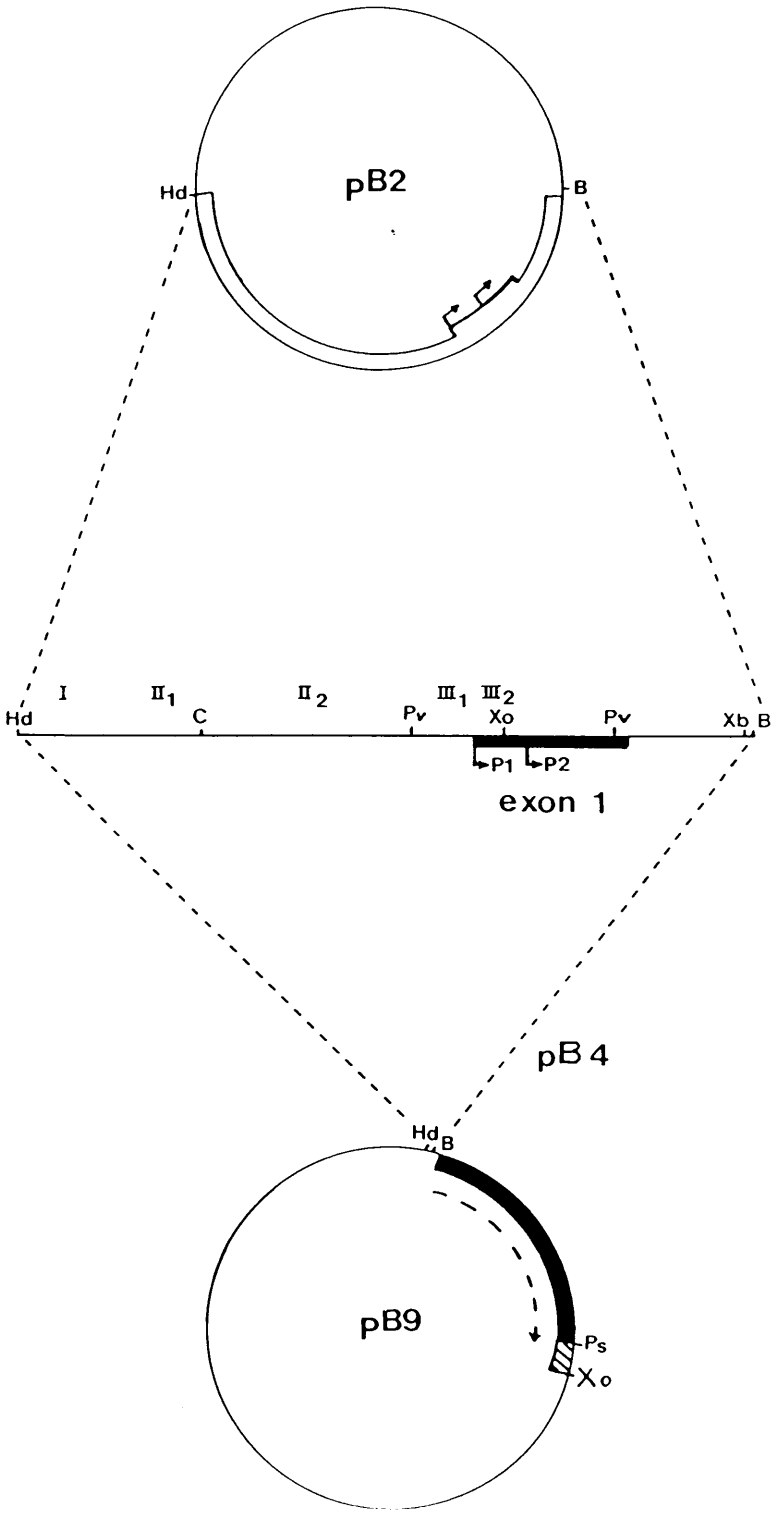


Figure 38.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB4.

For a detailed description of the cloning strategy see section 3.3.1. See figures 8g and 44.



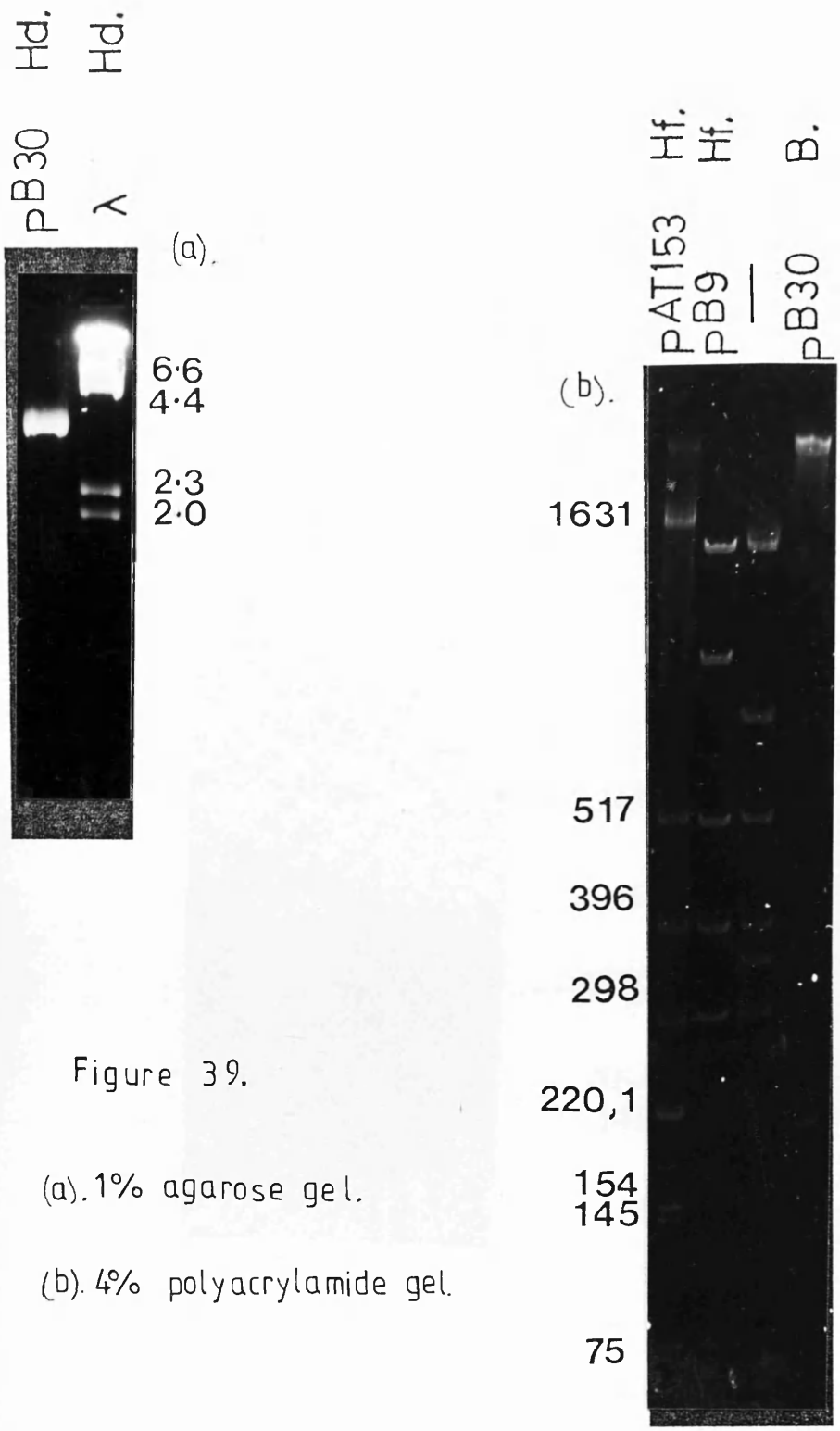
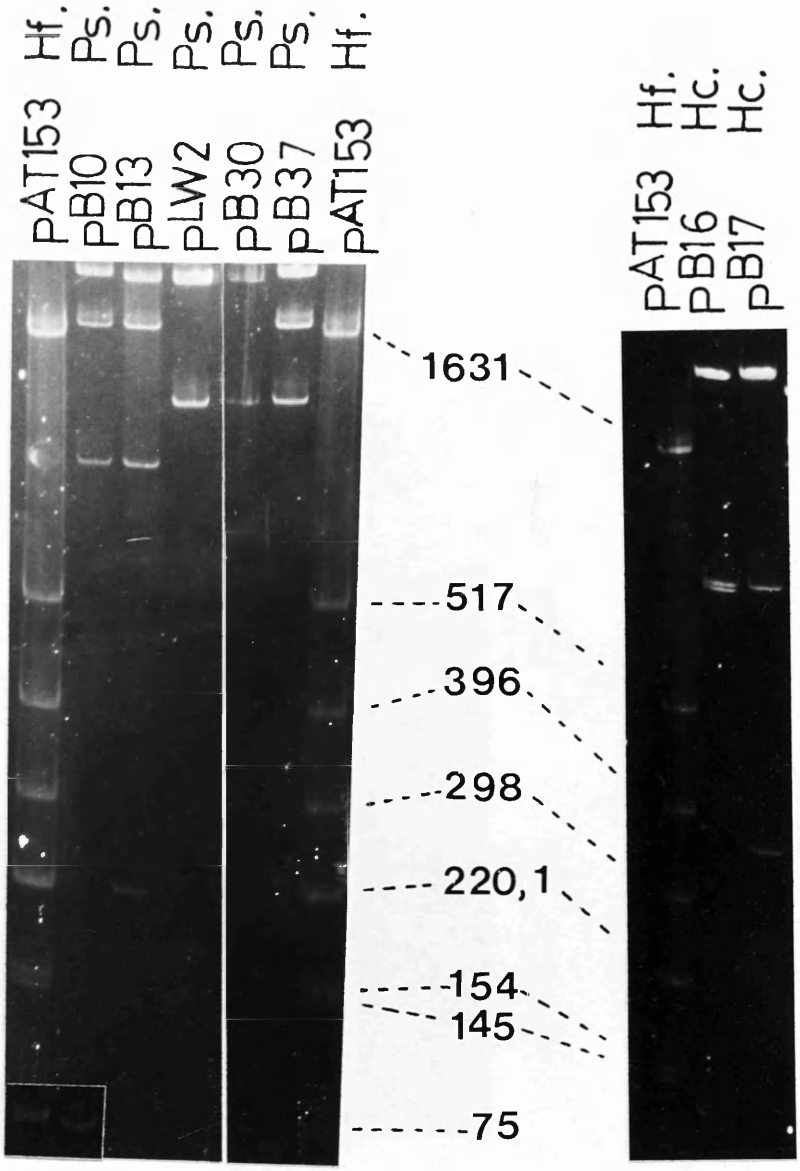


Figure 39.

(a). 1% agarose gel.

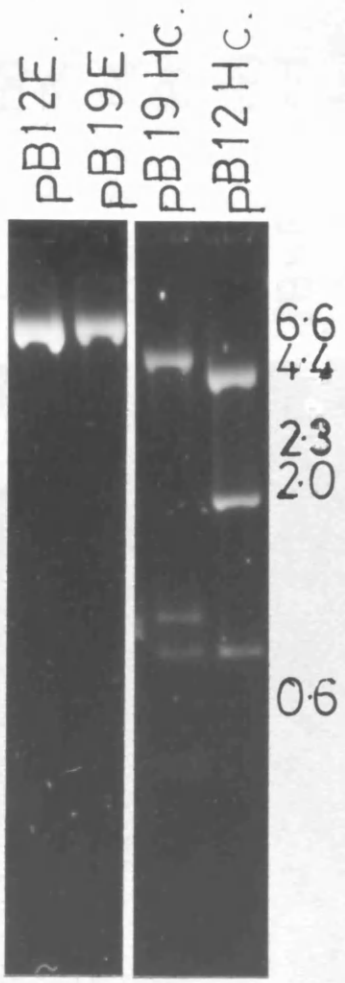
(b). 4% polyacrylamide gel.

Figure 40.



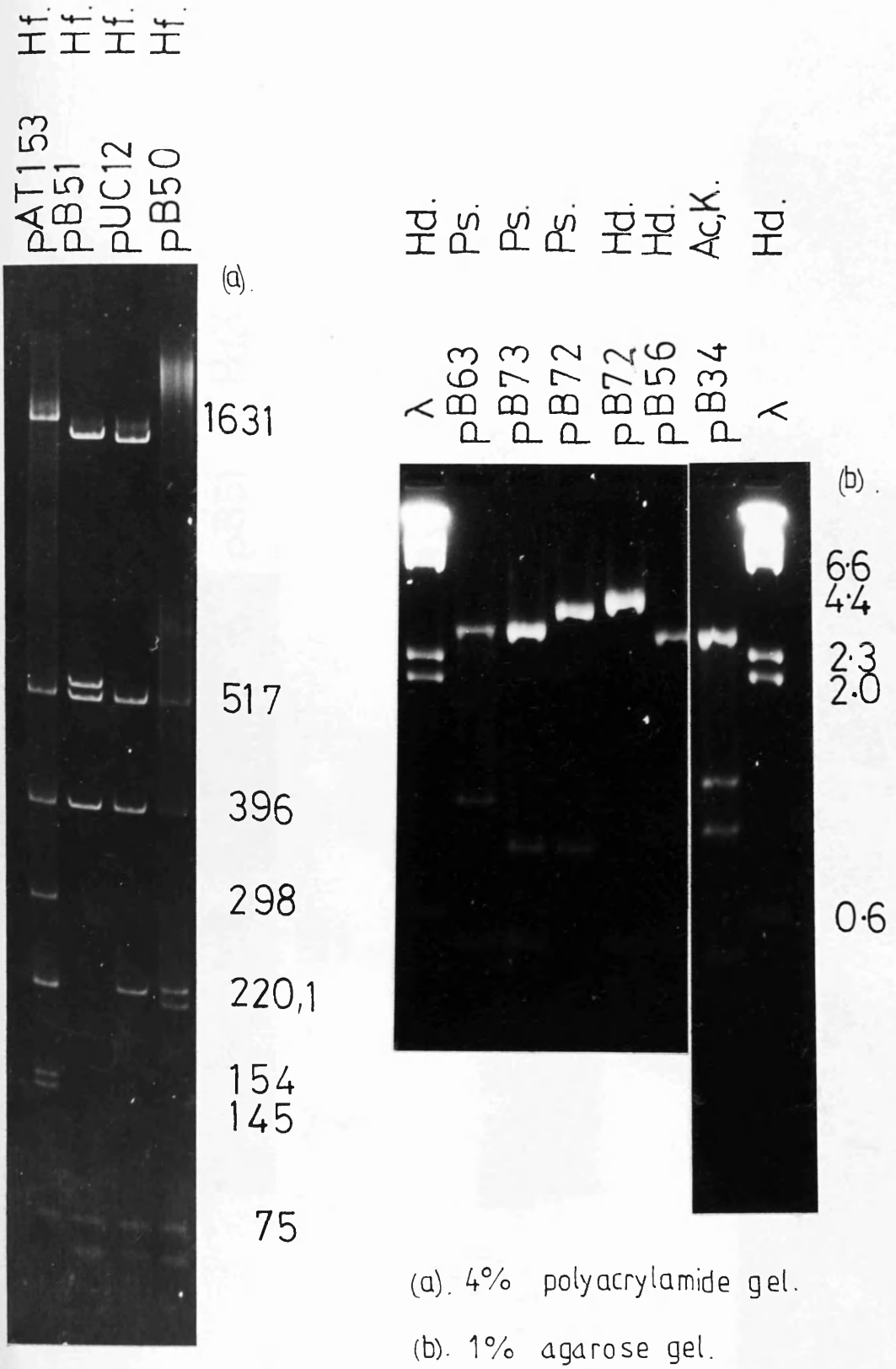
4% polyacrylamide gel.

Figure 41.



1% agarose gel.

Figure 42.



(a). 4% polyacrylamide gel.

(b). 1% agarose gel.

Figure 43.

(a), (c). 4% polyacrylamide gel

(b). 1% agarose gel.

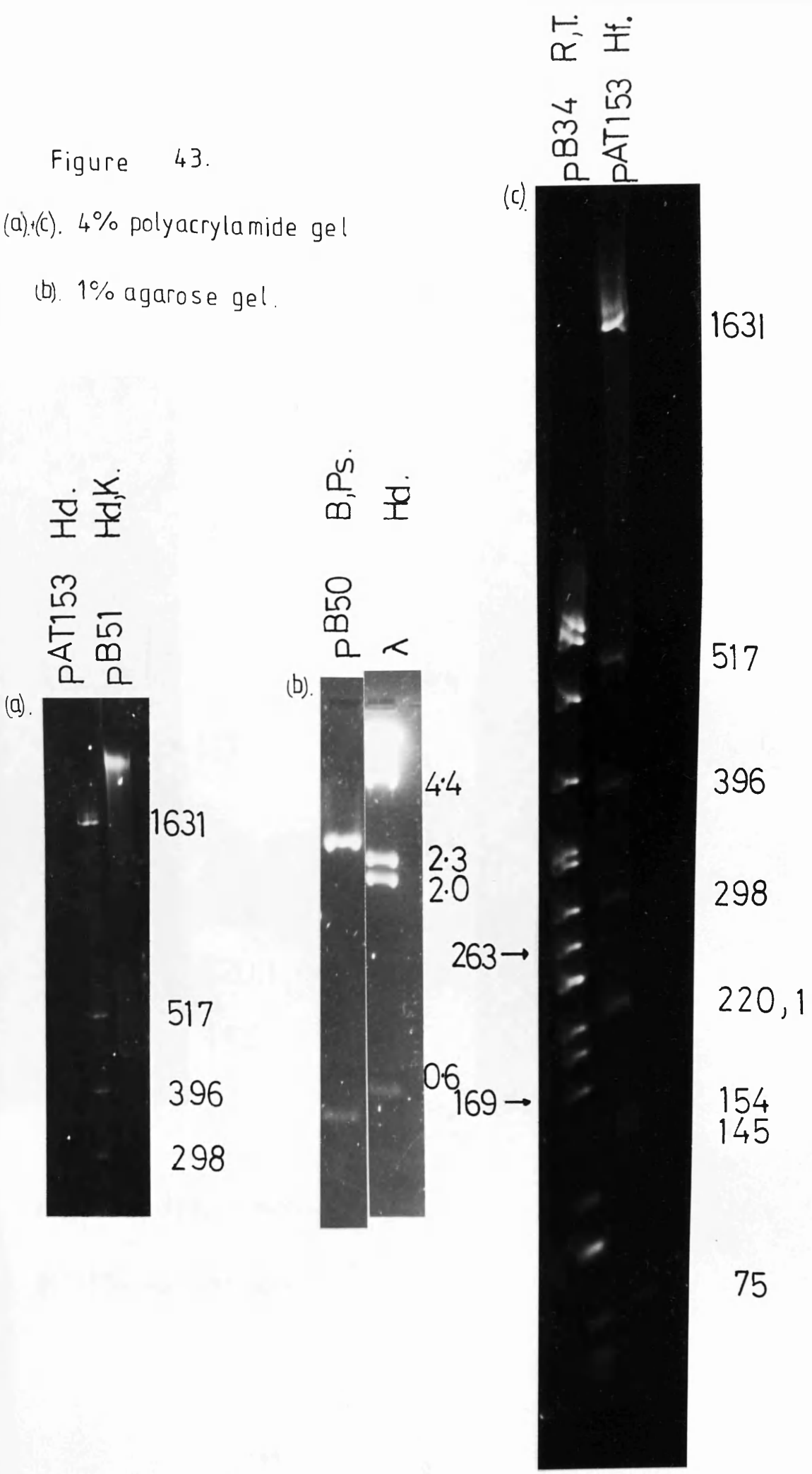
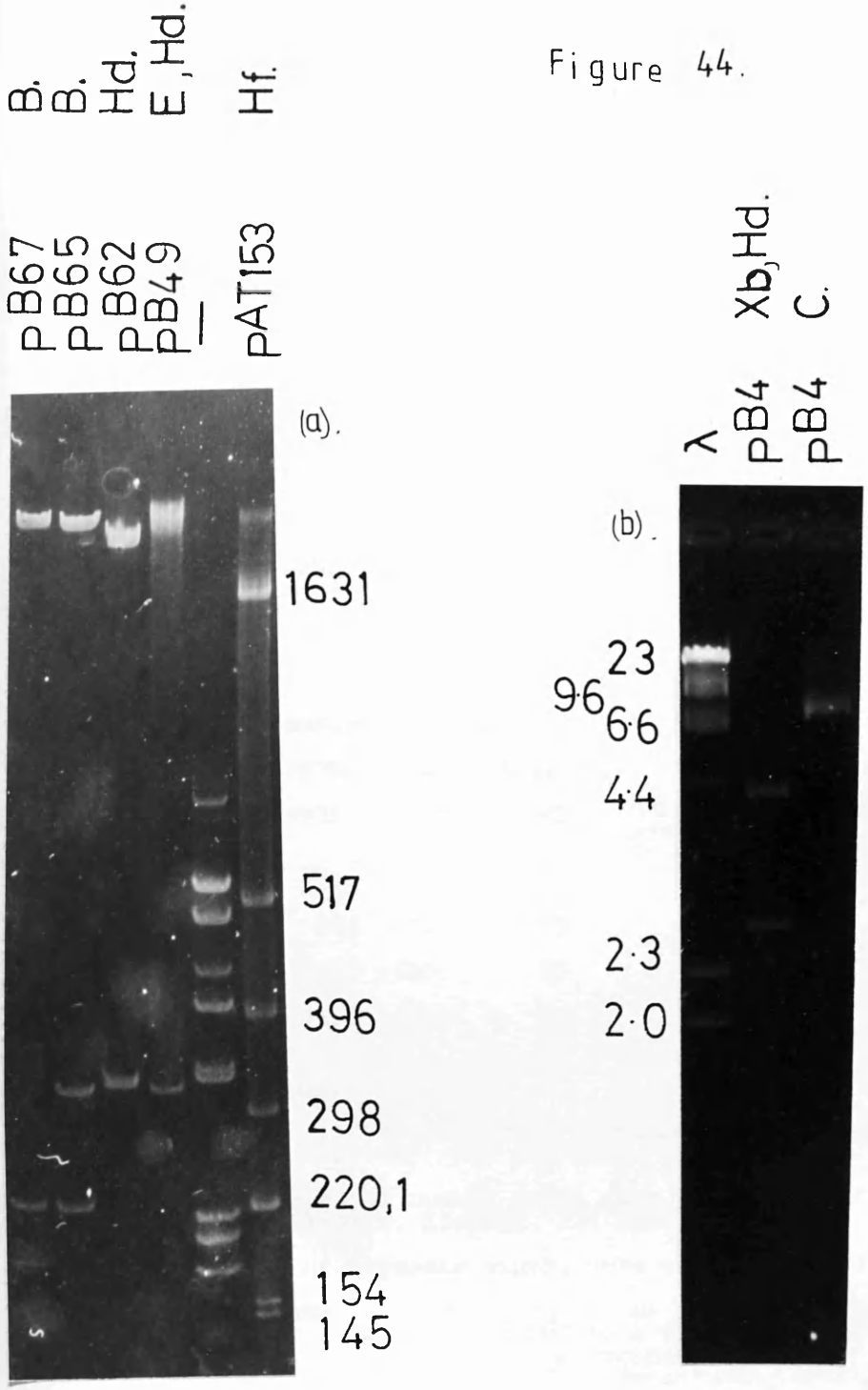


Figure 44.



(a). 4% polyacylamide gel.

(b). 1% agarose gel.

Table 3.

Activity of HindIII-PvuII c-myc Upstream Sequences upon Homologous and Heterologous Promoters.

Donor DNA	Recipient Cell Lines: LaTK⁻		HeLa		NIH-3T3		K-562	
	[a]	[b]						
I. pB9	462	(1)	436	(20)	491	(57)	803	(5)
pB14	109186	(104)	2185	(100)	857	(100)	16526	(100)
pB12	17909	(16)	569	(25)	514	(60)	1423	(8)
pB19	5052	(5)	881	(40)	483	(56)	1308	(8)
expts [c]	3		3		4		2	
II. pB9	462	(1)	436	(2)	491	(8)	ND	
pLW2	139818	(100)	19314	(100)	6567	(100)	ND	
pB16	72702	(52)	2188	(11)	2646	(40)	ND	
pB17	29955	(21)	ND		850	(13)	ND	
expts	3		2		4			
III. pB9	462	(16)	ND		ND		ND	
pB30	2981	(100)	ND		ND		ND	
pB37	1429	(48)	ND		ND		ND	
expts	3		-		-		-	

a. Average of total counts (cpm) per flask (no more than 15% variartion). ND= not done.

b. Average total counts (cpm) as a precentage of promoter alone (I:pB14, II:pLW2, III:pB30).

c. Number of separate experiments of duplicate flasks.

d. Donor DNA (eqimolar ratio of 5μg) was mixed with carrier salmon sperm DNA to a final concentration of 20μg/ml in the calcium-phosphate precipitate and transfected as described in Materials and Methods.

e. CAT activity was assayed as described in Materials and Methods using total cell extract.

(see figure 45.).

Figure 45.

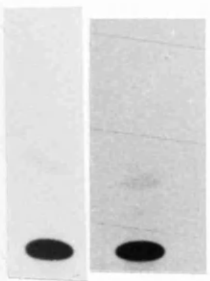
Examples of CAT Assay Autoradiographs.

A. The effect of the 2.0Kb HindIII-PvuII upstream sequence (pB16) on the activity of the HSV-IE gene promoter (pLW2) in LaTK$^-$ cells. (see table 3-II.)

B. The activity of the 5' deleted c-myc-CAT recombinants pB12 and pB29 compared to pB14 in LaTK$^-$ cells. The promoterless CAT vector pB9 is also shown. (see figure 28.)

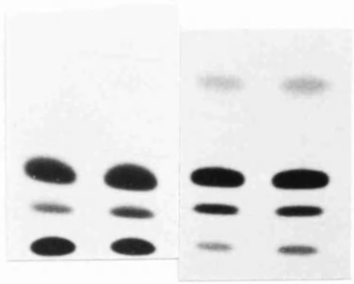
The equimolar amount transfected, and the plasmid name, is shown below each autoradiograph.

A.



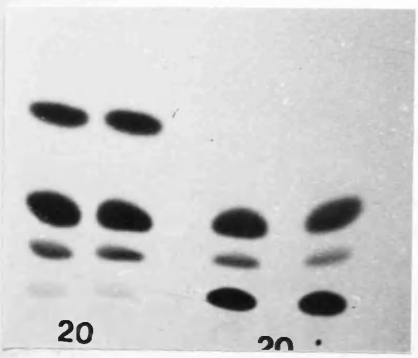
1 10

pB9



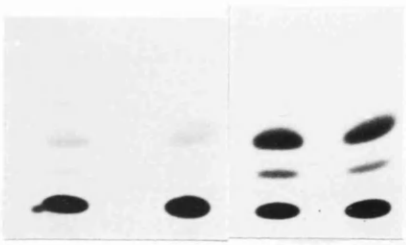
1 10

pB14



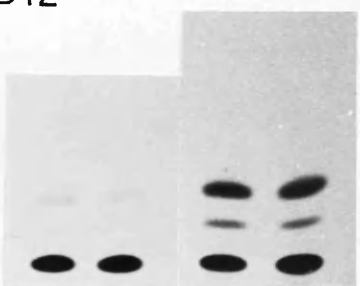
20 20

pB14 pB12



1 10

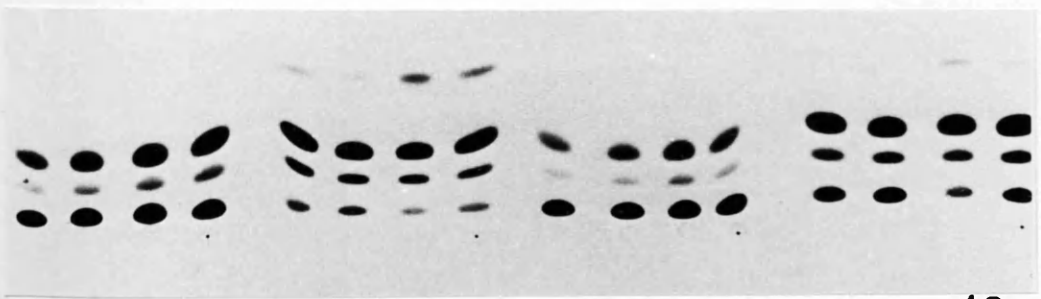
pB12



1 10

pB 28

B.



1 2 5 10 1 2 5 10

pLW2 pB16

Table 4.

Activity of the NRE-2 Sub-Regions upon the HSV-IE Gene Promoter in Transiently Transfected LaTK⁻Cells.

Donor DNA	CAT Activity [a] Counts	[b] %	[c] expts
pB9	5685 (11%)	2.1	3
pLW2	270952 (16%)	100	3
pB63	350850 (18%)	129	3
pB65	293293 (12%)	108	2
pB67	277120 (8%)	102	2

a. Average total counts (cpm) per flask with range of values expressed as a precentage in brackets.

b. Average total counts represented as a percentage of pLW2.

c. Number of separate experiments of duplicate flasks.

d. Donor DNA (equimolar ratio of 4μg) was mixed with carrier salmon sperm DNA to a final concentration of 20μg/ml in the calcium-phosphate precipitate and transfected as described in Material and Methods.

e. CAT activity was assayed as described in Materials and Methods using total cell extract.

Table 5.

Transient Expression of Promoterless/Capless CAT Recombinants.

Donor DNA	Recipient Cell Lines: LaTK⁻		K-562	
pB9	1.0[a]		1.0	
pB14	36.2	(±3.6)[b]	20.6	(±2.0)
pB10	1.2	(±0.1)	0.8	(±0.1)
pB13	1.0	(±0.1)	1.1	(±0.1)
pB72	12.7	(±3.3)	ND	
pB73	17.3	(±2.3)	ND	
pB4	1.0	(±0.1)	0.9	(±0.1)

a. Values represent the average total counts (cpm) from two separate experiments of duplicate flasks, normalised to the CAT activity of pB9. i.e. pB9=1. ND= not done.

b. Values in brackets indicate the range of activity between flasks observed.

c. Donor DNA (equimolar ratio of 10μg) was mixed with carrier salmon sperm DNA to a final concentration of 20μg/ml in the calcium-phosphate precipitate and transfected as described in Material and Methods.

d. CAT activity was assayed as described in Materials and Methods using total cell extract.

3.4. Evidence for Repression Involving a Trans-Acting Factor

Since the studies on the location of the negative regulatory element by 5' deletion, showed that with increasing amount of transfected DNA the level of repression decreased, the presence of a trans-acting factor which was being titrated out was implied. To investigate this further, additional DNA-titration and competition experiments were undertaken.

3.4.1. Recombinant Plasmid Constructions

Many of the recombinants used in this series of experiments have been described previously (see sections 3.2.1. and 3.3.1.).

(i). Recombinants for Competition Experiments

The 2.0Kb HindIII fragment from pB7H containing DNase I HSS I, II-1 and II-2, was inserted into HindIII digested pUC12 to give pB18 (see figure 46). Subsequent AccI partial digest to linearlise pB18, followed by ClaI digestion allowed pB20 (comtaining DNase I HSS II-2; see figure 46) to be generated by self-ligation and the insertion of the 1.1Kb AccI-ClaI fragment into AccI digested pUC12 gave pB21 (containing DNase I HSS I and II-1; see figure 46).

(ii). Recombinants for *in vitro* Footprint Experiments

The following plasmid was made unintentionally but it proved to be a very fortuitous ligation! The 445bp NRE-2 fragment from pB18, by KpnI-AccI digestion and subsequent low melting agarose gel isolation, was further digested with RsaI and ThaI. The resulting fragment mixture (containing both the 169bp RsaI fragment and the 263bp RsaI-ThaI fragment) was ligated into HincII digested pUC12. The 445bp NRE-2 fragment is contained in pB50 in a mixed orientation, i.e. 5' - 270bp - 170bp - 3' (see figure 47), by ligation of

the flush RsaI and ThaI ends of the internal fragments (see section 3.4.2.). The whole of the 445bp fragment is present, but the blunt ended fragments are in each others position, i.e. the originally 5' 170bp fragment is now located 3' to the 270bp fragment. The orientation of both fragments remains the same.

3.4.2. Restriction Enzyme Analysis of Recombinant Plasmids

Extensive restriction enzyme analysis has been used to identify each of the recombinant plasmids described. Representative digests are shown in figure 42, 43, 48 and relevant restriction maps in figures 8f, 8g.

(i). Plasmids for Competition Experiments

The 1940bp c-myc fragment can be dissected from pB18 by HindIII-ClaI digestion to reveal the 5' (1030bp) and 3' (910bp) fragments. These fragments can be identified in pB21 by HindIII digestion and pB20 by BamHI-HindIII digestion respectively (apparent disparity in sizes is due to linker modification or the presence of pUC derived polylinker sequences). Both pB56 and pB62 have been described previously (see 3.3.2.).

(ii). Plasmids for in vitro Footprint Experiments

To define pB50 the insert was identified by BamHI-PstI digestion (440bp), and its orientation (which is mixed) was confirmed by HinfI digestion. The apparent disparity in size of the 1454bp fragment was probably due to the proximity of this lane to the edge of the gel. The c-myc derived 62bp fragment is diagnostic of the NRE-2 domain, while the 209bp and 1454bp fragments confirm the relative positions of the two internal fragments in this recombinant. RsaI-ThaI digestion of pB34 shows the parental RsaI and RsaI-ThaI fragments.

In addition to restriction enzyme analysis, the configuration of the insert in pB50 was confirmed by the 'G' sequence tract used in the footprint analysis.

3.4.3. Expression Assay Results

By increasing the amount of DNA transfected we assume the amount of DNA taken up by the cells is increased, which would (and does) manifest itself as an increase in CAT activity (see figure 14; Weintraub, Cheng and Conrad, 1986). Thus any cellular factor which functions in *trans* on the DNA might be the limiting parameter and could in effect be titrated out, as the amount of DNA transfected increases. In addition, if a competitor plasmid (containing only the *cis*-acting element sequences and no coding sequences) is co-transfected in increasing amounts, with the indicator recombinant (containing the transcription regulatory element in *cis* to the reporter gene) again the *trans*-acting factor(s) might be titrated out by interaction with the competitor plasmid. Both these strategies have been used to positively identify a role for a *trans*-acting factor(s), interacting with the *cis*-acting NRE-2, to bring about the repression activity seen in these studies.

3.4.4. DNA Titration Studies

These experiments involved the transfection into LaTK$^-$ cells of increasing equimolar amounts of plasmid DNA. The resultant CAT activites were then represented as a ratio to pB14 (i.e. at each DNA amount pB14 gives a value of 1), as shown in figure 49. The test plasmids used were pB12, pB28, pB34, pB46, pB14 and pB9 (see section 3.2.4.). When pB46 and pB14 are compared by DNA titration (pB46 gave the same activity as pB14) the resulting CAT activity implied that they have an equal potential for transcription, as was shown for previous experiments (see section 3.2.4.). This infers that only the proximal 350bp upstream of the c-*myc* promoter P1 are required for full promotion of the c-*myc* gene, and

that the -607 to -350bp region has no obvious functional role in control of expression in this system.

The CAT control (pB9) also increases linearly with pBl4 but at only a fraction of its activity level. As already mentioned this is probably not due to an artifact of the CAT assay but an intrinsic property of pB9 (see section 3.2.6.). However, when pBl2 is compared with pBl4 (or pB46) the repression ratio at low equimolar amounts (equimolar amounts are used to ensure that an equal plasmid copy number is available to the cell) is 10-fold greater than that at high equimolar amounts (20µg) of transfected DNA. This implies the presence of a trans-acting factor (repressor), which is being titrated out as the number of exogenously introduced plasmid copies getting into the cell increases (i.e. self titration).

The titration effect can be seen for pBl2, pBl9 and pB28. However, pB34 is repressed when compared to pBl4 but not to the same extent as pBl2 (see 3.2.1. and 3.3.3.). This indicates that part of the NRE-2 domain may be 5' to the KpnI site. This region (around the KpnI site) may be involved in conferring a structural alteration upon the NRE-2 domain induced by trans-acting factors binding to the NRE-2 domain directly. Why such sequences are 5' to the NRE-2 is not immediatly obvious, however these studies would not identify 3' sequences with a similar property. Since the competition studies (see 3.4.5.) show that the 445bp KpnI-AccI fragment can greatly relieve repression, it is unlikely that sequences 5' to the KpnI site act as recognition elements for trans-acting factors. However it remains a possibility that the 445bp fragment could conceivably be titering out only an essential factor(s), but not all the factors required for repression. It is also posssible that the deletion from the KpnI restriction site could truncate part of the negative regulatory element trans-acting factor recognition domain.

However, the differences in repression ratio between pB12, pB28 and pB34 are relatively small when compared to the actual repression ratio observed. Therefore this may reflect the assay variation rather than functional domains (see section 3.6.). Interestingly, repression appears to be greater in the larger recombinants (pB12 > pB28 > pB34). Therefore, it is possible that the difference in repression ratio may just reflect a more efficient repressor-NRE-2 interaction in the larger recombinants, possibly due to some form of conformational restraints.

Titration experiments for pB12, pB14 and pB9, were also done with K-562 cells acting as the transfection recipient. No titration-effect is seen in K-562 cells, which show the greatest repression activity. K-562 cells exhibit a very poor transfection efficiency (D.Spandidos and P. Montague, personal communication) and this result may reflect this. If the assumption is made that this poor efficiency is due to a poor DNA uptake, then even though 10μg of DNA are transfected the cell may only be taking up a small amount of the available co-precipitate. (N.B. The poor efficiency could also be due to poor DNA transport to the nucleus or inefficient DNA expression.) If this is so, then these results may only be at the start of a DNA-titration-curve and explain why no titration effect is observed. Experiments using larger amounts of DNA (20μg and 40μg) were attempted but this heavy precipitate resulted in a dramatic change from adherent cells to suspension cells (the K-562 cell line was selected for its adherence properties: see section 2.3.1.) and although results were obtained (not shown) they were not reproducible.

If this cell line is indeed only taking up a small proportion of the available DNA this could explain the results of section 3.2.3. The 5' deletion results infered that K-562 cells contained either a very active repressor(s), or that the repressor factor(s) was present in this cell line at high concentrations. The titration results

indicated that the lower the amount of DNA transfected into the cell, the greater the repression ratio is (i.e. less self titration of repressor). Taking these two results together suggests that the high level of repression seen in K-562 cells may be a consequence of the low amount of DNA taken up by these cells (i.e. low nuclear levels of exogenous plasmid).

To investigate whether a titration effect could be observed in other cells, pB16 (containing the 2.0Kb NRE-2 regulating transcription from the HSV-IE promoter) was transfected in increasing amounts into LaTK$^-$and HeLa cells. The resultant CAT activities are plotted in figure 51. The titration-effect was also observed for the 5' c-myc region driven repression of the HSV-IE gene promoter (pB16) in LaTK$^-$ and Hela cells. To confirm earlier results with this recombinant the repression appears to be greater in Hela than LaTK$^-$ cells. Whether this is due to an increased presence of c-myc-interacting trans-acting factor(s) in these cells or the respective ability of each cell to express the viral promoter is unknown. Because LaTK$^-$ cells function as very efficient transfection recipients it is possible that at equal DNA amounts (i.e. at 1μg) the number of plasmid copies taken up and expressed by the nucleus is far greater for the LaTK cells than for the HeLa cells. This would make it technically difficult to see the lower end of the titration curve (see section 3.1.6.). The limited titration done for pB30 and pB37 (containing the human ε-globin gene promoter) also showed the titration-effect (unpublished result).

Although these results imply the presence of a trans-acting repressor which is being titrated out, the titration out of positive factors can not be dismissed. Indeed the expression curve seen for the c-myc-promoter (pB14) and the HSV-IE promoter (pLW2) do show a slight leveling off before the saturation plateau is reached. This may indicate a titration-effect on a positive factor. Therefore, to further

indicate that it is the repressor which is being titrated out competition experiments were undertaken.

3.4.5. Competition Experiments

In these experiments the indicator recombinant containing the reporter gene (CAT) driven by the desired cis-acting regulatory elements was transfected (constant equimolar amount) in the presence of increasing equimolar amounts of the competitor recombinant (containing only the relevant cis-acting regulatory element). LaTK$^-$ cells were used as the transfection recipients for these competition experiments. The carrier DNA used for these competition experiments was pUC plasmid DNA, thus ruling out the titrating out of trans-acting factors by the plasmid vector sequences.

The 2.0Kb HindIII-PvuII NRE-2 fragment has been shown to represses heterologous promoters, therefore it was used as the first competitor (see figure 52). Increasing amounts of this fragment inserted in pUC12 (pBl8) were used as the competitor to the indicator plasmid pBl2. At a 5 molar excess of competitor to indicator, the titration-out of repression was observed. No CAT activity was observed for pBl8 alone (data not shown).

Since the repression activity had been localised to a fragment containing DNase I HSS II-2 located in the the 3' half of the c-myc sequences in pBl8, this region was used as a competitor (pB20) against pB28 (c-myc 5' deletion from HincII retaining NRE-2). Again, release from repression was observed at 5 molar excess of competitor to indicator (see figure 52). When the 5' half of pBl8 is used as a competitor (pB2l) against the indicator pB28 there is no obvious release from repression (there is a very slight rise in activity; see figure 52). This would apparently exclude any sequences 5' to the ClaI site from acting as an additional "sink" (see 3.6.4.).

Previous results indicated that the 445bp KpnI-AccI fragment although containing an essential component of the negative element may not include the whole domain. Therefore, competition experiments similar to those described above were done using the 445bp NRE-2 region (pB56) or the 270bp NRE-2 sub-region (pB62) to see if these fragments alone could compete out the repression activity (see figure 53). Both the 445bp fragment and the 270bp sub-fragment compete out repression at approximately the same molar excess of competitor to indicator as pB20 (about 5-10 fold). Unfortunately, during these experiments the growth state of the recipient cells was somehow impaired (slower growth), as reflected by the low absolute values obtained. However , the result was obtained from two separate experiments thus indicating a competition effect. This implies that the 445bp and its 270bp sub-fragment contain, at least, an essential component of the cis-acting negative element (NRE-2), and that repression is mediated through the interaction of at least one trans-acting factor (the repressor).

3.4.6. Analysis of in vitro Footprint Studies

Since titration and competition results imply the presence of a trans-acting repressor factor(s) the obvious question is whether and which specific DNA binding protein(s) interact with the NRE-2 domain. One method for the identification of such an interaction is by DNase I in vitro footprinting (Galas and Schmitz, 1978), which is now used quite extensively to answer this type of question (e.g. analysis of Spl : Kadonaga, Jones and Tjian, 1986).

In the basic technique a ^{32}P-end-labelled DNA probe containing at least one putative protein binding site is incubated with a protein extract, lightly treated with DNase I and then electrophoresed on a denaturing polyacrylamide gel. Regions of DNA that are bound by a trans-acting DNA-binding protein will appear as protected areas (the

"footprint"). DNA footprinting was carried out on the NRE-2 domain (the 445bp KpnI-AccI fragment: Whitelaw et al, in press) using a partially purified nuclear DNA-binding protein preparation (HO.3), and two highly purified preparations prepared by specific DNA affinity chromatography from Hela cells, highly enriched for transcription factors Spl and CAATT Binding Transcription Factor (CTF) repectively (Kadonaga and Tjian, in press; see figure 7). As a partially purified protein extract (HO.3) is used it is possible that not all the protein factors present in the nucleus were included.

Figures 54 and 56 show the DNA footprints, and where they map, obtained from the 445bp NRE-2 (pB50). Four distinct footprints were obtained, one of which can be accounted for by Spl binding, one by a Spl-like factor and two by CTF binding. In addition, several *in vitro* DNase I hypersensitive sites were observed. The limits of the footprints and the location of the *in vitro* DNase I hypersensitive sites are shown relative to the DNA sequence in figure 56. These results indicate (from 5'-3') one Spl site (5'-GCCCCTCCCA-3'), one Spl-like site (5'-CAGGAGGGGC-3') and two CTF sites in opposite orientations respectively (5'-TTTGG-3') and (5'-CCAAT-3') are located within this fragment, as defined by *in vitro* analysis.

Weak Spl binding sites have been observed for other gene promoter regions (Kadonaga, Jones and Tjian, 1986), and therefore what is termed a Spl-like factor binding site may just reflect a weak Spl site. The 5' CTF consensus is not as good an homology as would be expected. There is another consensus (3'-ATTGG-5') nearer the 3' end of this footprint, but this would then make the symmetry of the two CTF footprints different. As yet it is not known whether symmetry is essential or not. Thus whether both footprints indicate a similar binding activity (or even an identical protein) is still debatable. There is also a relatively

strong hypersensitive site within the 5' CTF footprint.

These results imply the possible involvement of either, or both, Sp1 and CTF in the *trans*-mediation of repression observed in these studies, but in no way excludes the presence of additional protein interactions.

Figure 46.

Structure and Cloning Strategy of c-myc-CAT Recombinants pB18, pB20 and pB21.

For a detailed description of the cloning strategy see section 3.4.1. See figure 8g and 48.

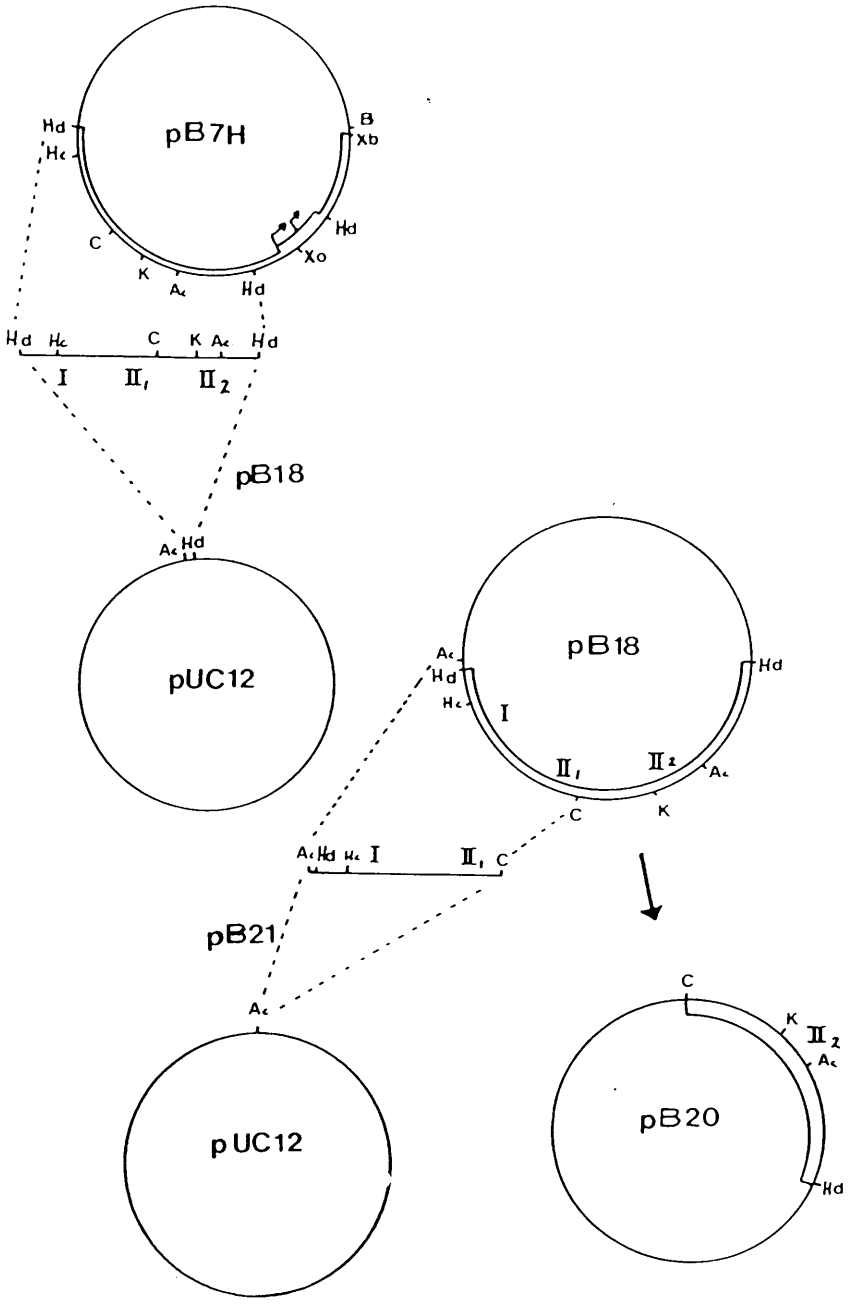


Figure 47.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB50.

Restriction sites which are destroyed upon ligation are marked with a diagonal line across them. For a detailed description of the cloning strategy see section 3.4.1. See figure 8g, 42, 43 and 46.

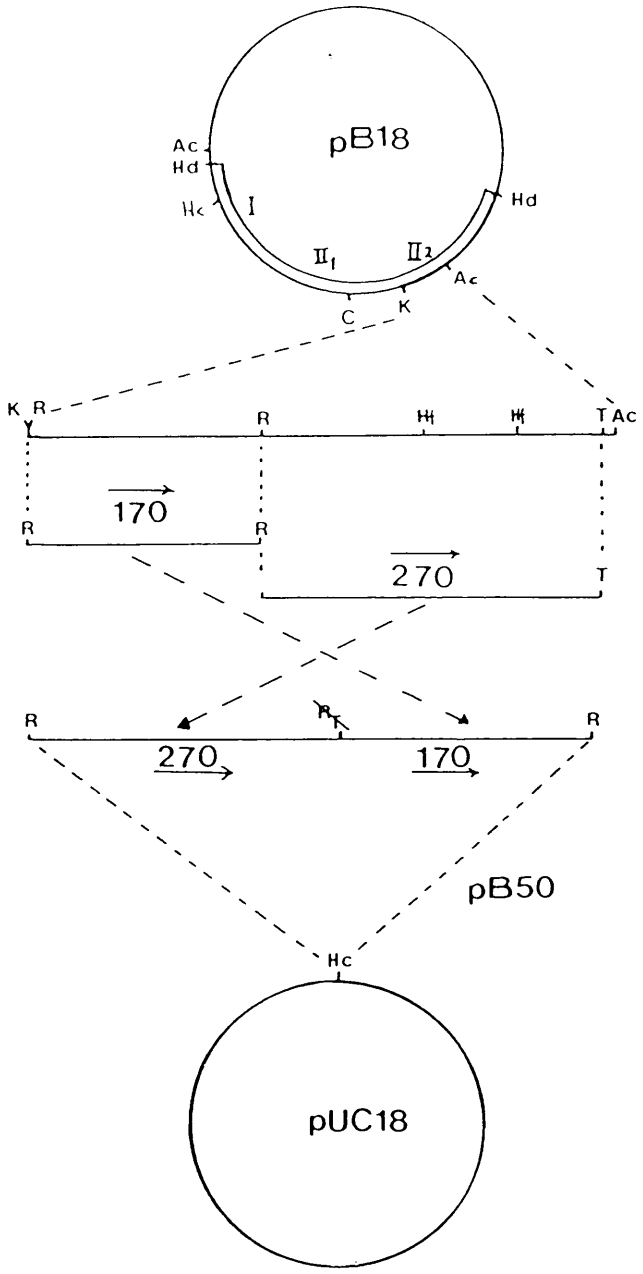
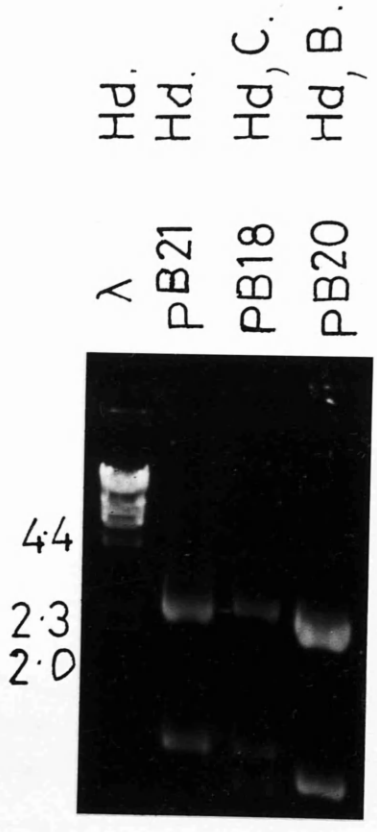


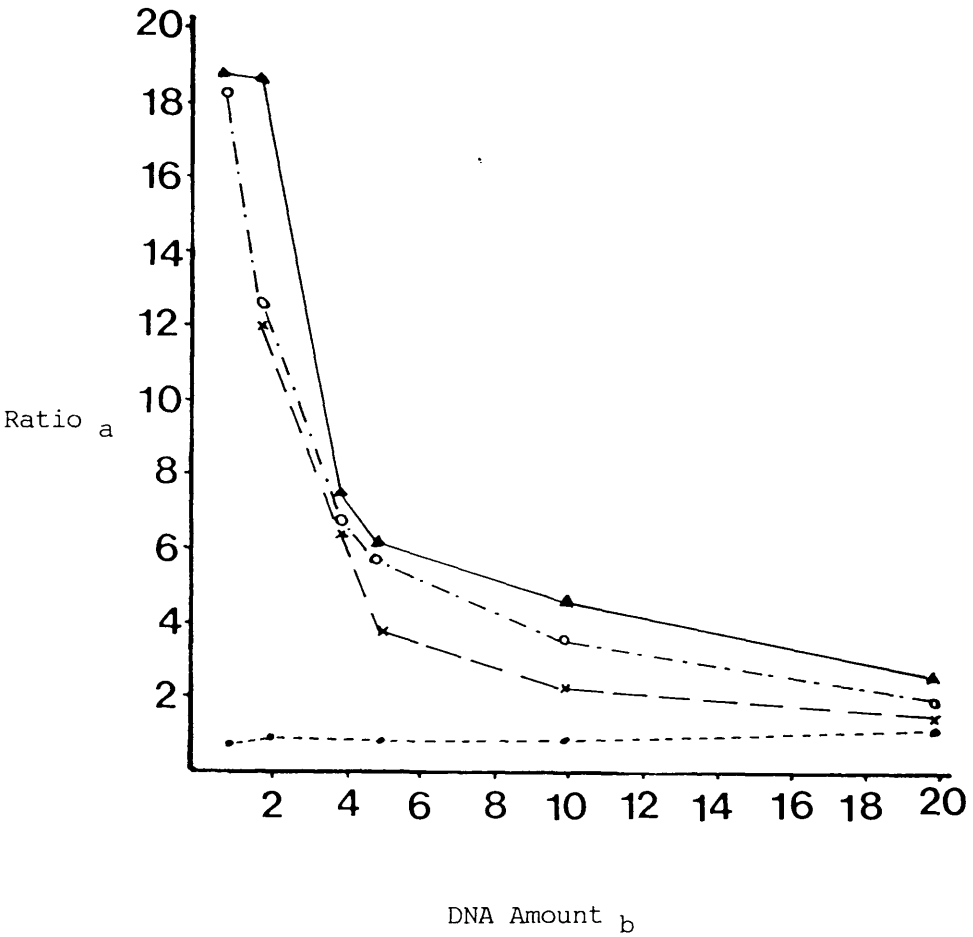
Figure 48.



1% agarose gel.

Figure 49.

DNA Titration Analysis of 5' Deleted c-myc-CAT Recombinants in LaTK⁻cells.



a.Ratio represents the fold difference between pB14 and the various c-myc-CAT recombinants. Values represent (as a ratio) the average total counts (cpm) of two to ten separate experiments of duplicate flasks.

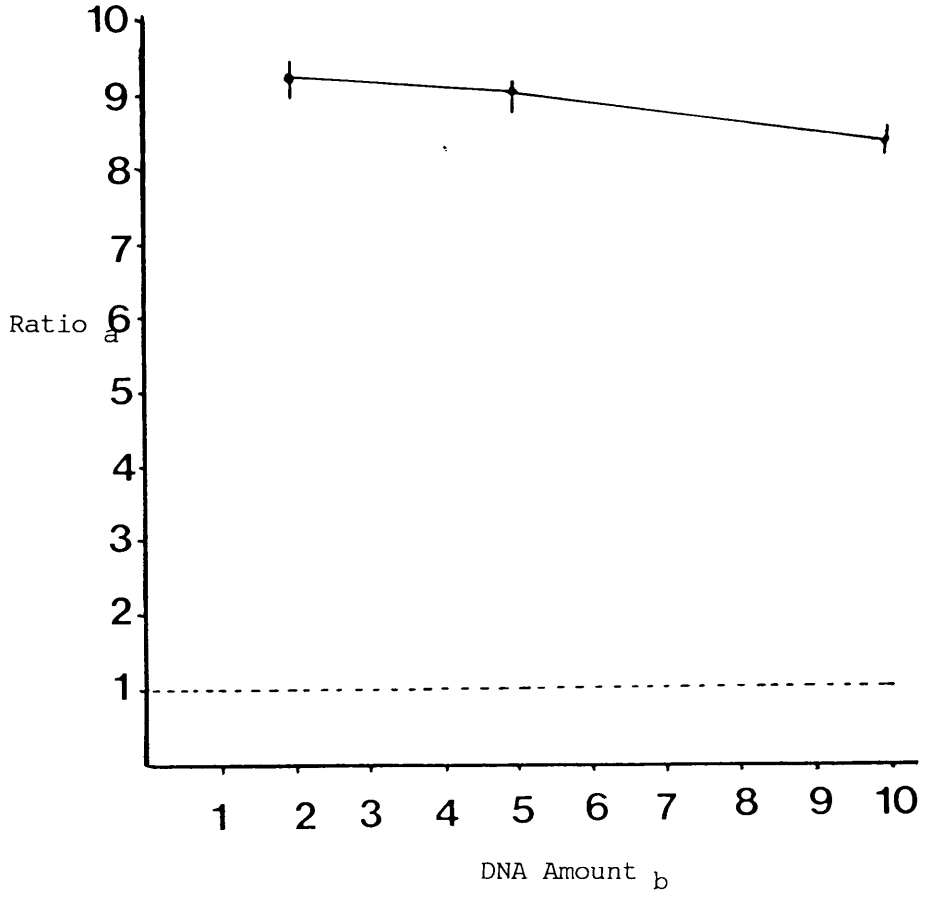
b. Equimolar ratios of donor DNA were transfected into LaTK⁻ cells as described in Materials and Methods. CAT activity was assayed as described in Materials and Methods.

c. Key: ——— =pB12; ·-·-·-·- =pB46; — — — =pB34.

(see figure 45).

Figure 50.

Expression of 5' Deleted c-myc-CAT Recombinants in K-562 Cells.



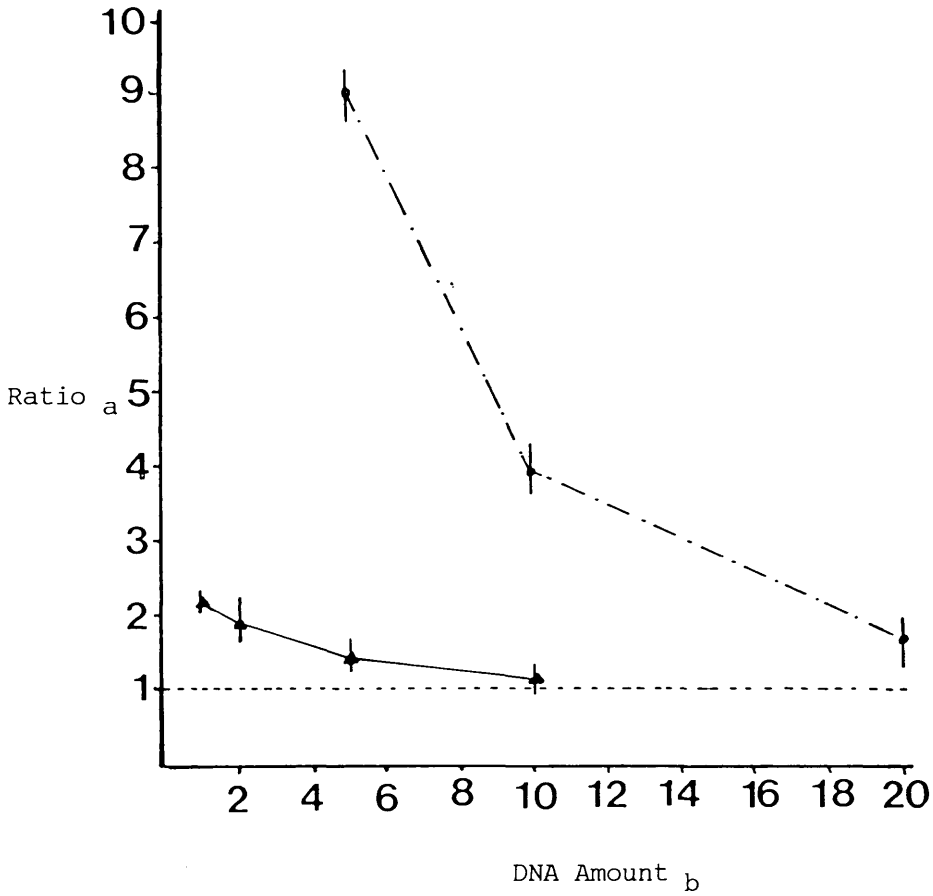
a. Ratio represents the fold difference between pB14 and pB12. Values represent the average of total counts (cpm) from three separate experiments of duplicate flasks. Vertical bars represent range of values obtained.

b. Equimolar ratios of donor DNA were transfected into K-562 cells as described in Materials and Methods. CAT activity was assayed as described in Materials and Methods.

c. Key: ———— =pB12; ------- =pB14.

Figure 51.

DNA Titration Analysis of c-myc-HSV-CAT Recombinants in LaTK$^-$ and HeLa Cells.



a. Ratio represents the fold difference between pLW2 and pB16. Values represent the average of total counts (cpm) from two separate experiments of duplicate flasks. Vertical bars represent range of values obtained.

b. Equimolar ratios of donor DNA were transfected into LaTK$^-$ and HeLa cells as described in Materials and Methods. CAT activity was assayed as described in Materials and Methods.

c. Key: ..... =pLW2; _____ =pB16 in LaTK$^-$ cells; —·—·— =pB16 in HeLa cells.

(see figure 45).

Figure 52.

Competition Assays with the c-myc Upstream Sequences.

A. Competition experiment with pB12 (indicator) and pB18 (competitor).

B. Competition experiment with pB12 (indicator) and either pB20 or pB21 as the competitor.

a. Average total counts (cpm) from two experiments of duplicate flasks represented as a percentage of pB14. Vertical bars indicate the range of values obtained.

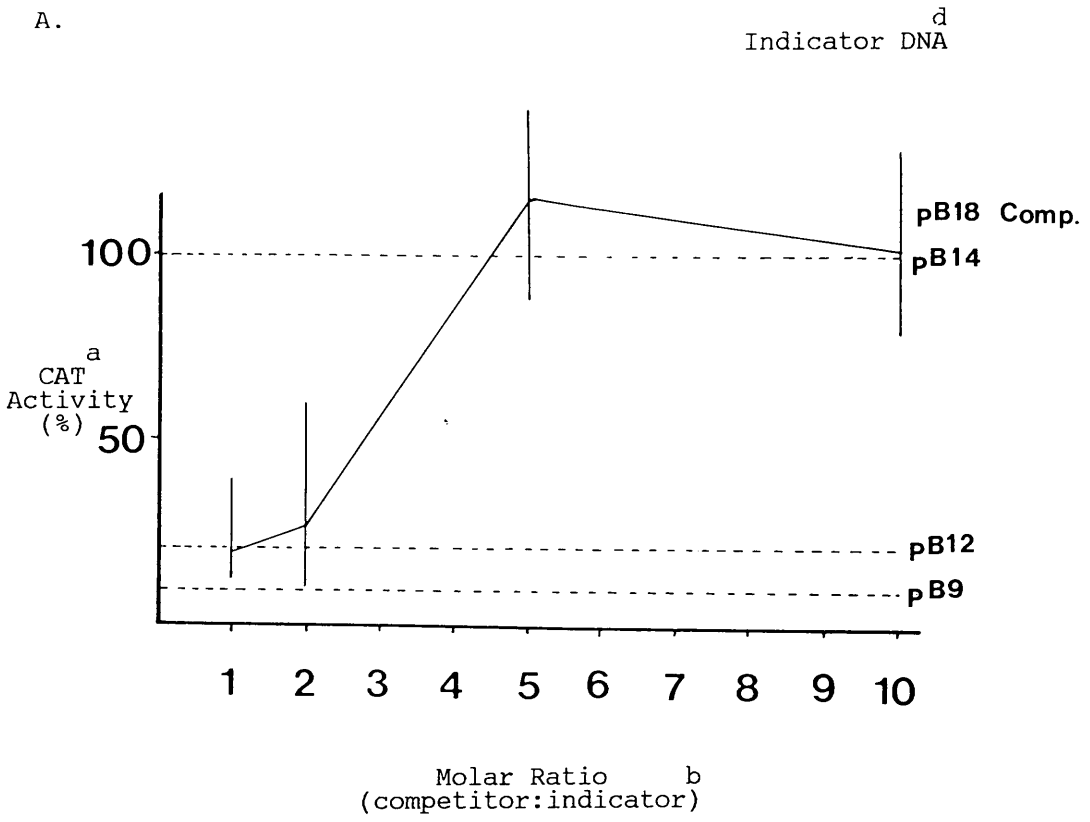
b. Donor indicator DNA (equimolar amount of 2μg) was mixed with increasing amounts (represented as a molar ratio) of competitor DNA (equimolar amounts of 1-10μg) , and adjusted to 20μg/ml with pUC12 carrier DNA.

c. DNA was transfected into LaTK$^-$ cells as described in Materials and Methods. The resultant CAT activity was assayed as described in Materials and Methods.

d. The hatched lines represent the activity of the indicator DNA alone, while the continuous lines represent the activity of the indicator in the presence of competitor DNA.

Competition Assays with c-myc Upstream Sequences.

A.



B.

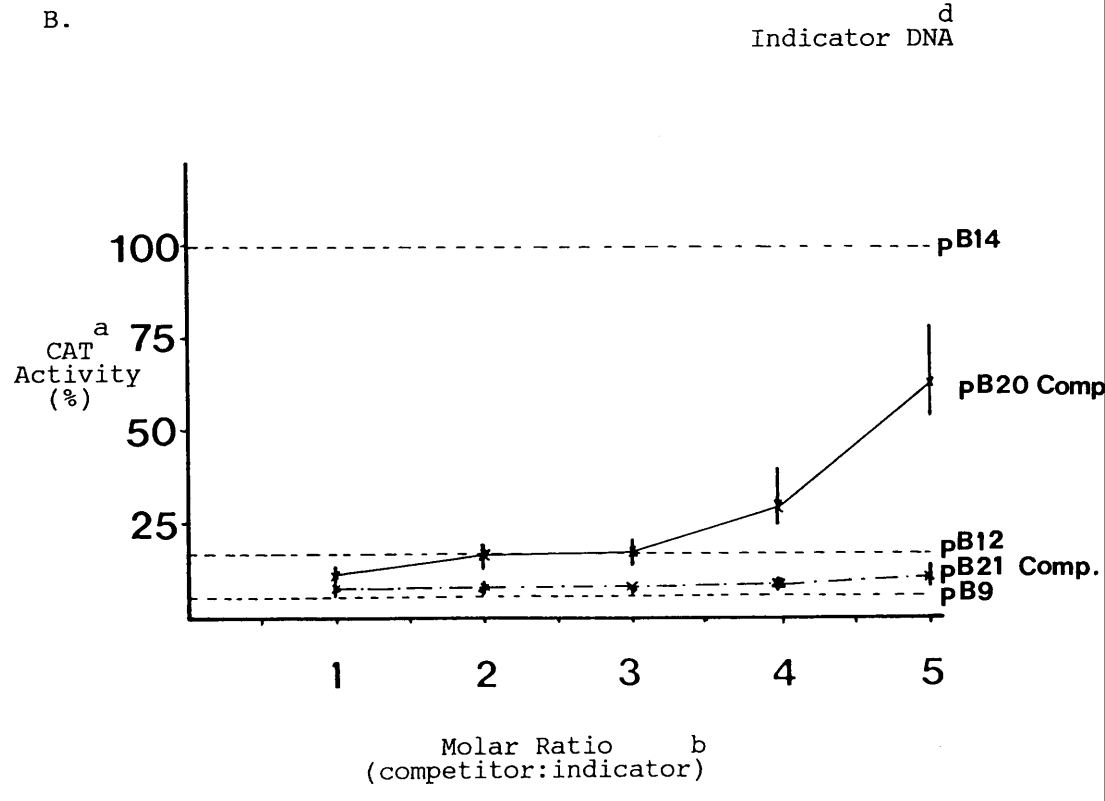
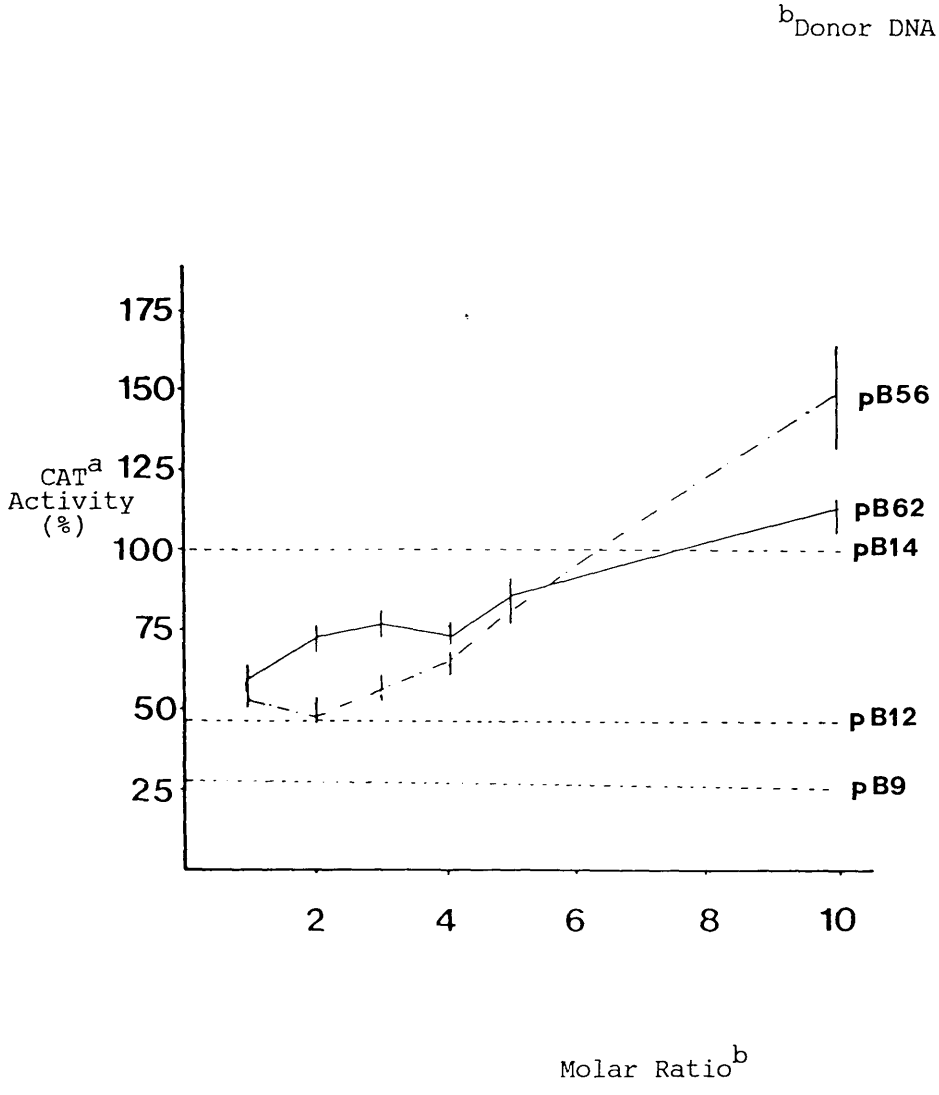


Figure 53.

Competition Assay With the NRE-2 Sub-fragments as Competitors.



a. Average total counts (cpm) from two separate experiments of duplicate flasks as a percentage of pB14. Vertical bars represent the range of values obtained.

b. Donor indicator DNA (equimolar amounts of 2µg) was mixed with increasing equimolar ratios (0-10) of competitor DNA (pB56 or pB62) and adjusted to 20µg DNA/ml with pUC12 carrier DNA.

c. The calcium-phosphate transfection of LaTK$^-$cells was done as described in Materials and Methods. CAT activity was assayed as described in Materials and Methods.

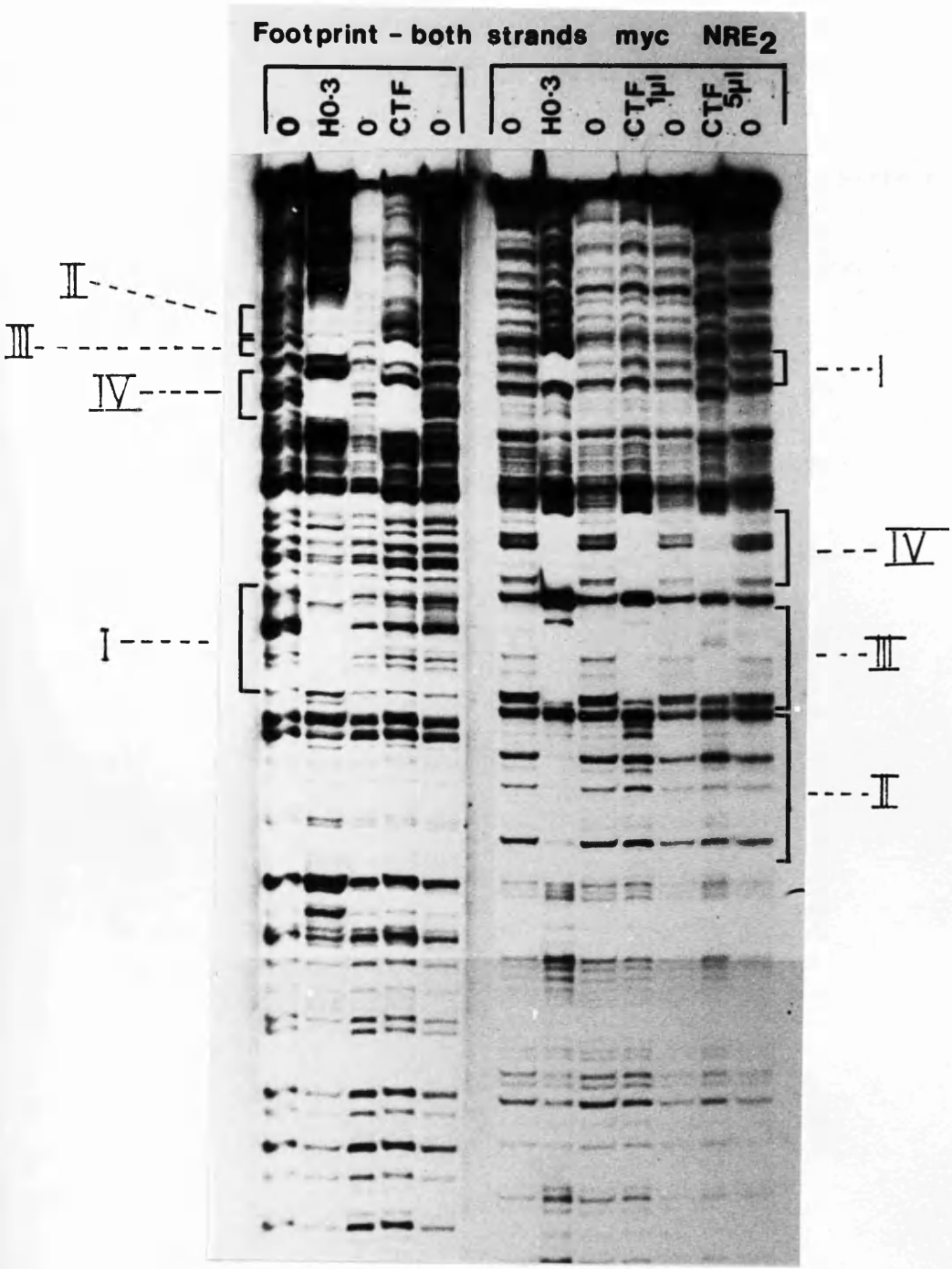

Footprint – both strands myc NRE2
0
HO-3
0
CTF
0
0
HO-3
0
CTF 1µl
0
CTF 5µl
0
II
III
IV
I
I
IV
III
II

Figure 55.

Photoraph of the *in vitro* DNase I Hypersensitivity Pattern of the NRE-2 Domain showing Sp1 Interaction.

The H.03, CTf-enriched and Sp1-enriched cell extracts were used as indicated (see Materials and Methods for detailed description of the experimental technique). The fragment are labelled at the 5' end. The Roman numerals refer to the footprint (DNase I protected domains) observed. From this gel sites III and IV are again shown to bind CTF, while sites I and II appear to bind Sp1. In addition, a strong *in vitro* hypersensitive site was observed between the two CTF protected domains. These results have been confirmed by gaining the same results from two further experiemnts.

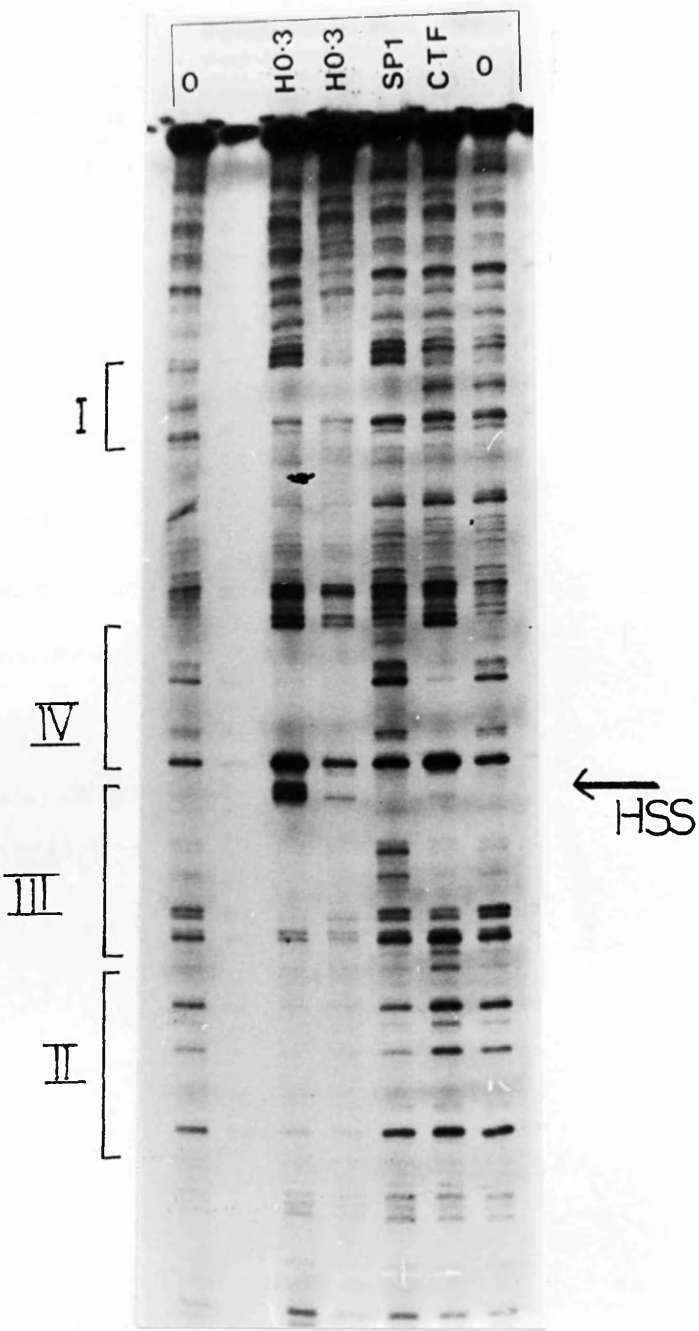
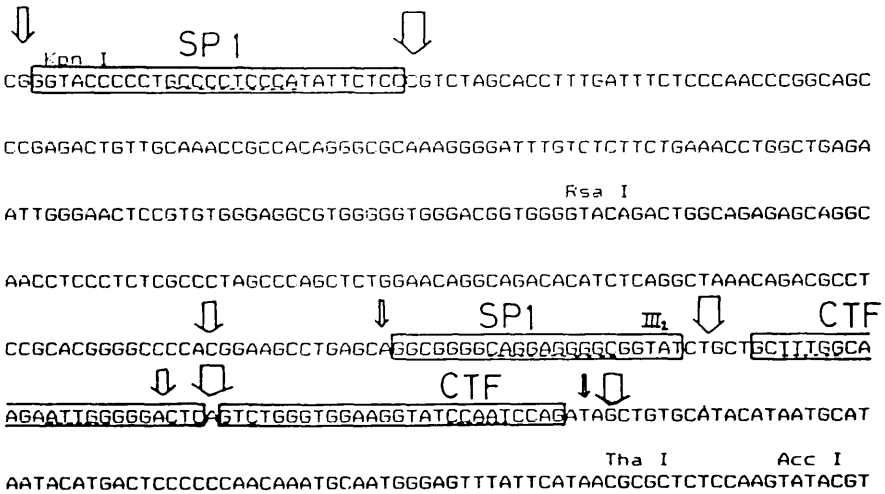

O
H0·3
H0·3
SP1
CTF
O
I
IV
III
II
HSS

Figure 56.

Sequence of the NRE-2 Region Indicating the in vitro DNase I Footprints (Protected Domains).

A. The footprints indicated were obtained using Sp1 and CTF enriched protein extracts (see figures 54 and 55). The relevant restriction sites are shown. The in vitro DNase I protected domains are boxed. The open arrows indicate in vitro DNase I hypersensitve sites (width of arrow correspomnding to the intensity of the site).The consensus sequences (see figure 74) for Sp1 and CTF are shown by dotted lines. The in vivo DNase I hypersensitive site (Siebenlist et al, 1984) is shown by the Roman numeral II-2.



3.5. Identification of an Upstream Promoter Element

A repeated, fairly well conserved sequence occurs approximately 110bp 5' to P1 in the c-myc gene (human: Siebenlist et al, 1984; feline: Stewart et al, 1986; murine: Bernard et al, 1983). As it is by nature a conserved non-coding sequence (this is the argument first used to identify transcription control elements) it is thought likely to have same role in the regulation of transcription. In addition, this region contains the in vivo DNase I HSS III-1 (Siebenlist et al, 1984). Transient functional analysis of this region identified it to pocess weak promoter activity, again with an apparent lack of tissue-specificity.

3.5.1. Recombinant Plasmid Constructions

Again some of the recombinants used in this series of experiments have been described previously (see 3.2.1.).

(i). Recombinants for Upstream Promoter Element Analysis

By inserting the SmaI-XbaI fragment from pBl2 into SmaI-XbaI digested pUCl2, the c-myc driven CAT recombinant pB36 was generated (see figure 57), which contained only 97bp of upstream sequences.

PvuII SmaI double digestion of pB3 and subsequent self-ligation gave pB66 (see figure 57), which contains the whole c-myc 5' flanking region except DNase I HSS III-1 region (a 250bp fragment). Subsequent insertions of the 2.6Kb HindIII fragment into HindIII digested pB9 gave pB74, (see figure 57), the c-myc driven CAT recombinant which had all the 5' flanking region except DNase I HSS III-1.

(ii). Recombinant for in vitro Footprint Analysis

PvuII SmaI double digest of pB2 releases a 250bp

fragment, containing DNase I HSS III-1, which when inserted into HincII digested pUC12 gave pB57 (see figure 58).

3.5.2. Restriction Enzyme Analysis of Recombinant Plasmids

Extensive restriction enzyme analysis has been used to identify each of the recombinant plasmids described. Representative digests are shown in figure 59 and relevant restriction maps in figures 8b, 8c and 8i.

(i). Recombinants Containing DNase I HSS III-1

The c-myc deleted recombinant pB36 upon XbaI digestion gives a linear fragment of 4.2Kb (compare with pB14 digested with EcoRI; see figure 26). SmaI-HindIII digestion of pB36 reveals the c-myc fragment of 610bp and the CAT-terminator domain fragment of 895bp. The 250bp 5' adjacent domain of c-myc was deleted from pB12 to give pB74. HincII digestion of pB12 and pB74 gives respectively fragments of 3865, 1670, 830bp and 3865, 1420, 830bp, with the deleted fragment easily identified. To confirm this, pB74 was digested with HindIII to give fragments of 3575bp and 2540bp (pB12 has 2375bp and 2790bp fragments), and linearalised with KpnI to give a 6115bp fragment (250bp shorter than linear pB12 of 6365bp).

(ii). Recombinant used for *in vitro* Footprinting Studies

The c-myc derived 250bp fragment is shown in pB57 by BamHI HindIII digestion (260bp).

3.5.3. Identification and Localisation of the Upstream Promoter Element

A DNA titration experiment with pB36, pB14 and pB9 using LaTK and HeLa cells as the recipients was done in an attempt to functionally study the -350 to-100bp region 5' to P1 (see figure 60). As described before (see section

3.2.6.), 5' deletion to the PvuII site 353bp 5' to P1 (pB14) retains all the sequences required for promoter activity. Further 5' deletion to the SmaI site, leaving only 97bp 5' to P1 (pB36), results in a halving of activity. This result was reproducible and observed in LaTK$^-$ and Hela cells (also MRC-5 cells, data not shown). Such a decrease in activity implies the presence of an element weakly involved in the promoter domain of c-myc, although this deletion may be just disecting an element. In addition, in both LaTK$^-$ and Hela cells a weak titration-effect was observed, possibly indicating the involvement of a trans-acting factor. It must be stressed that the activity seen for this region is very weak (2 fold), but since it is both reproducible and present in two different cell lines, it probably reflects some in vivo involvement in promoter activity. Again a lack of specificity (tissue and species) between the cell lines analysed was observed.

3.5.4. Analysis of in vitro Footprint Studies

The same protein preparations as previously described (see sections 3.4.6.) were used to in vitro footprint the UPE domain (pB57; see Whitelaw et al, in press.)

The footprints observed for pB57 comprised a long, partially overlapping region of protection (see figures 61 and 62). This footprint cluster coincides with several (well conserved in mouse and feline c-myc genes: see figure 70) Sp1 consensus sequences, implying the presence of several putative Sp1 in vitro binding sites located between -160 and -100. No footprints were observed between -350 and -160bp 5' to P1. These results imply an upstream region (UPE) of partially overlapping weak Sp1 (or Sp1-like sites).

DNase 1 in vitro footprint analysis of the region spanning from -100 to +500 (which includes P1 and P2: N.Wilkie and J.Lang, personal communication) has identified several other GC-rich elements, which bind protein factors

(the HO.3 extract was used), so whether they bind Sp1 is not known). No DNase 1 footprints were observed adjacent to P2, although hypersensitive sites were seen after protein binding. On the basis of HO.3 titration experiments, the upstream sites appear to be weaker targets for Sp1 than the proximal site, or targets for a different DNA binding protein(s).

These results might imply an involvement of Sp1 factors in the trans-mediation of the promoter domain. Obviously, other factors are likely to be involved.

3.5.5. Additional Studies on the Upstream Promoter Element

The evidence described in this thesis makes it probable (see section 5.1.), that during a competition co-transfection experiment cis-recombination may occur. However, the additionally transfected competitor sequences will function as trans-acting factor sinks and thus the experiment is still a valid indicator for the presence of such a factor.

Competition experiments were done as described in section 3.4.5., to preliminarily analyse the UPE domain. The resultant CAT activity is represented as a percentage of pB14, as shown in figure 63. When the 250bp UPE domain in a pUC vector (pB57) were used as a competitor against the indicator plasmid pBl2 (the myc recombinant containing all the 5' sequences) the observed CAT activity was slightly greater than that expected for pBl2 alone (see figure 63). This elevation of expression was considerably less than that observed when the NRE-2 competitor (see figure 52) was used. It is possible that the UPE region (see section 3.5.4.) may be acting as a "sink" for the Sp1 factor required for the NRE-2 (in vitro footprint requirement), thus marginally reduing the functional efficiency of the NRE-2 domain.

However, when pB57 (containing the 250bp UPE domain) is

used as a competitor against the indicator plasmid pBl4 (the c-myc 5' deletion recombinant containing 350bp of 5' flanking sequences) a large increase in CAT activity, compared to pBl4 alone, was observed with a titration-effect occuring at the higher molar ratios. Perhaps the initial increase in activity is due to cis-activation due to recombination between the plasmids after transfection (which could also explain the pB57/pBl2 results above; see section 5.1.) while at the higher molar ratios the competitor sequences are acting as a Spl "sinks". However, if this result were taken in isolation, the competing out of a repression function would be the obvious conclusion. Thus, the UPE domain may posses different functions in different situations.

However, when the 250bp PvuII SmaI fragment containing the UPE is deleted from pBl2 (pB74) the resulting CAT activity upon transfection into LaTK⁻cells was greater than that of pBl2, but less than that of pB36 (and thus pBl4, i.e. the c-myc promoter driven CAT recombinant). This result (see figure 64), in agreement with the pB57/pBl4 competition result, indicates that this region could function as a negative regulator of expression. This presents an anomaly in this series of experiments which as yet can not be explained fully. Perhaps the internal deletion affects the chromatin structure of the entire c-myc sequences present, thus affecting the functional efficiency of the NRE-2 domain. Alternatively the UPE domain may function to mediate the repression instigated by the NRE-2 domain, i.e. the UPE may function as a positive regulator of transcription until there is some form of interaction between it and the NRE-2. The internal UPE deletion result implies that this interaction would be essential for maximal repression to occur. However previous results imply that the NRE-2 can function independently from the UPE (repression of heterologous promoters, see 3.3.3.), thus the postulated NRE-2/UPE interaction is not an absolute requirement for the NRE-2 repression function.

Figure 57.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB36,pB66 and pB74.

Restriction sites which are destroyed upon ligation are marked with a diagonal line across them. For a detailed description of the cloning strategy see section 3.5.1. See figure 8i, 19 and 62.

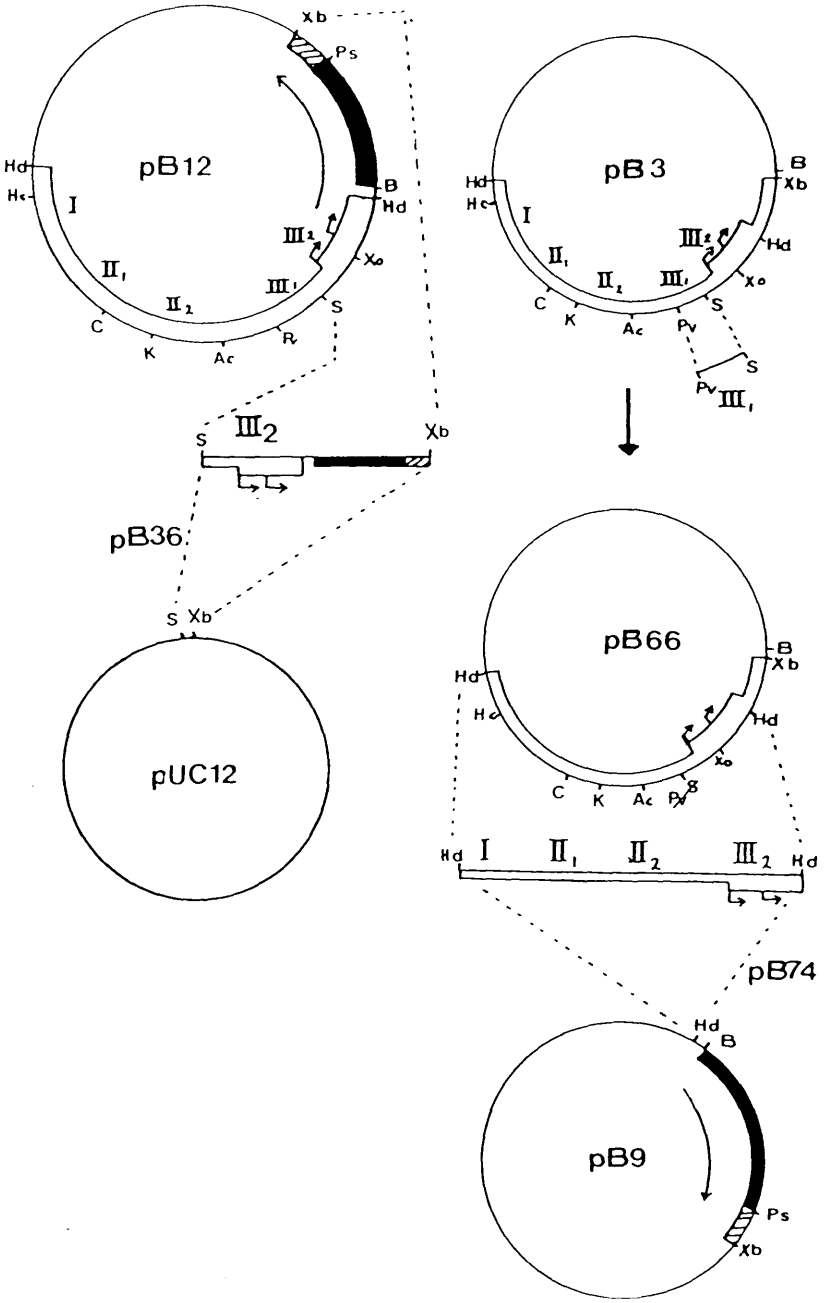


Figure 58.

Structure and Cloning Strategy of c-myc Recombinant pB57.

For a detailed description of the cloning strategy see section 3.5.1. See figure 8i, 18 and 67.

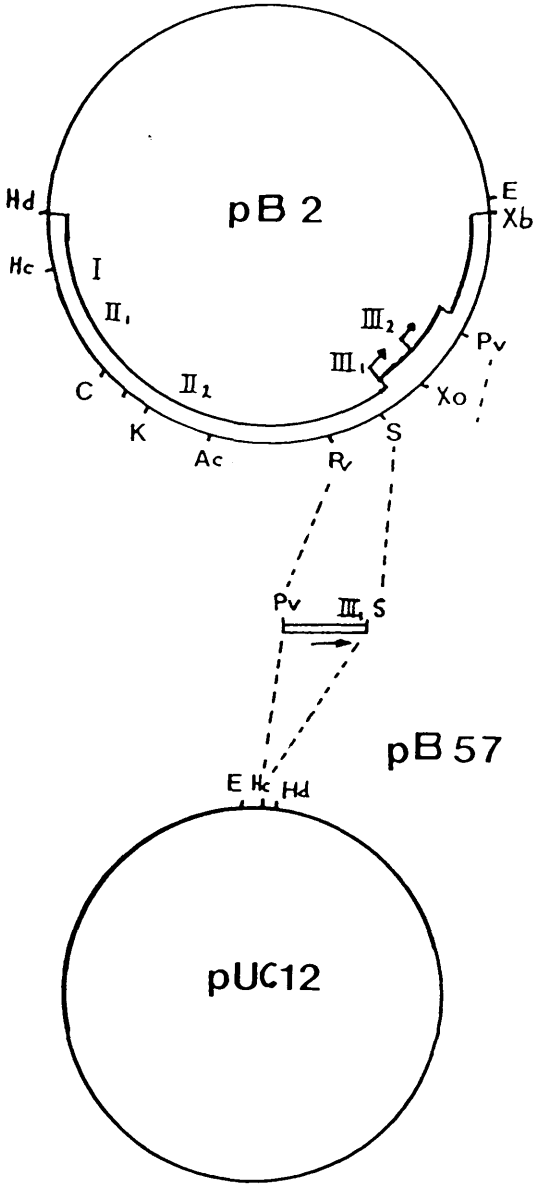
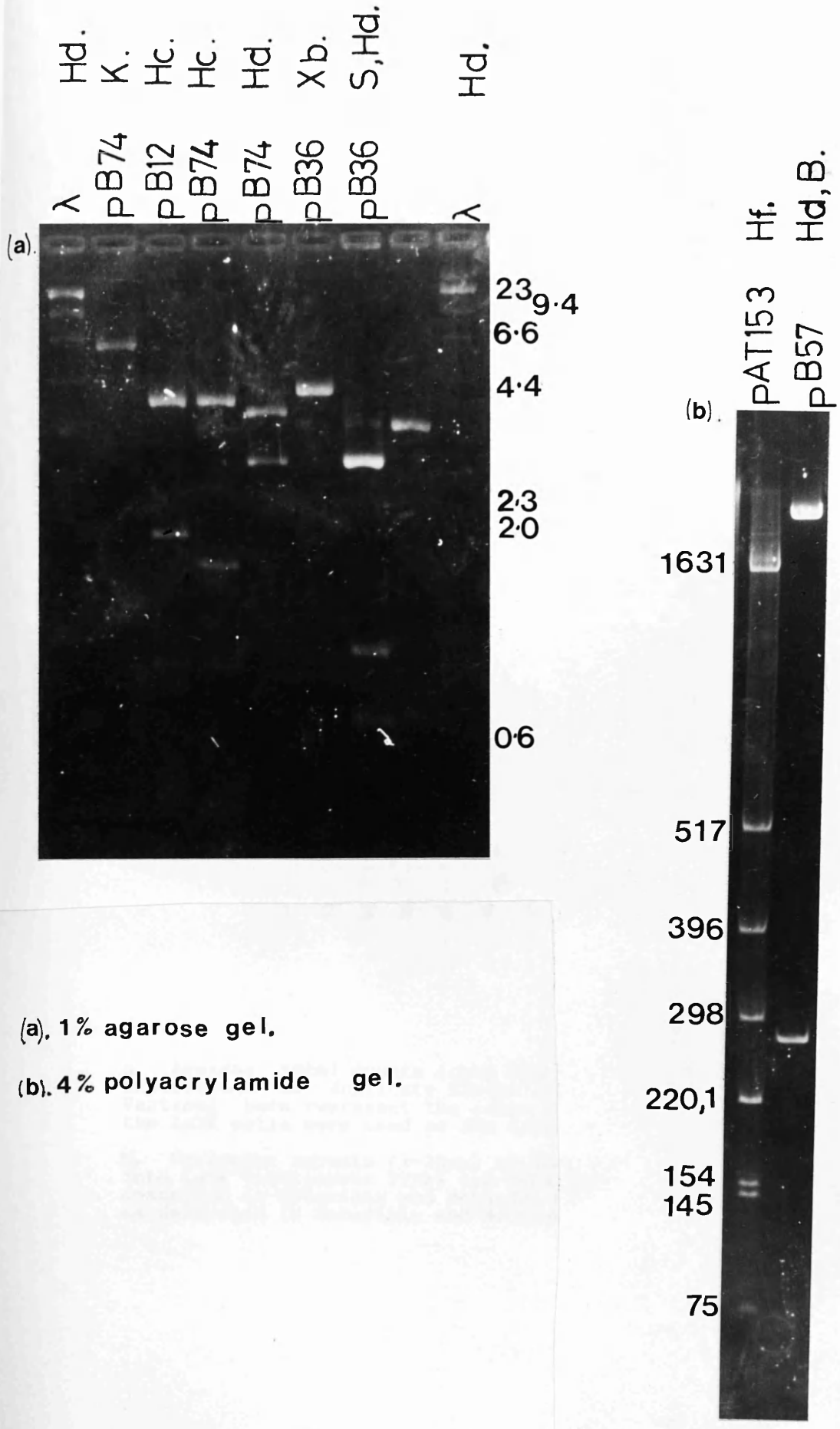


Figure 59.

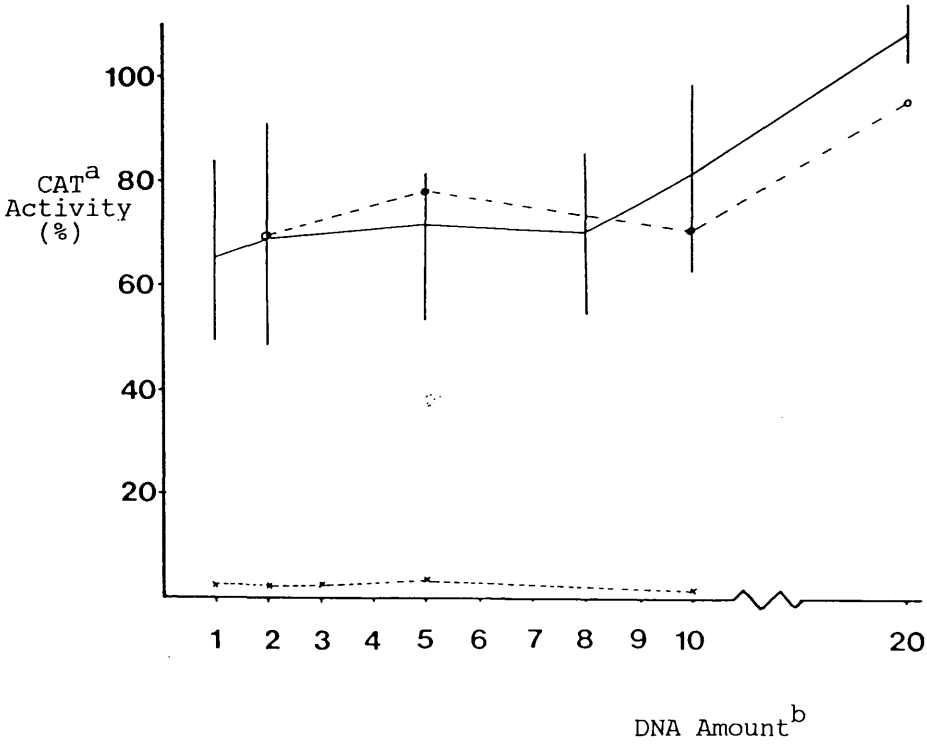


(a). 1% agarose gel.

(b). 4% polyacrylamide gel.

Figure 60.

CAT Activity of c-myc-CAT Recombinants with the Upstream Promoter Element 5' Deleted.



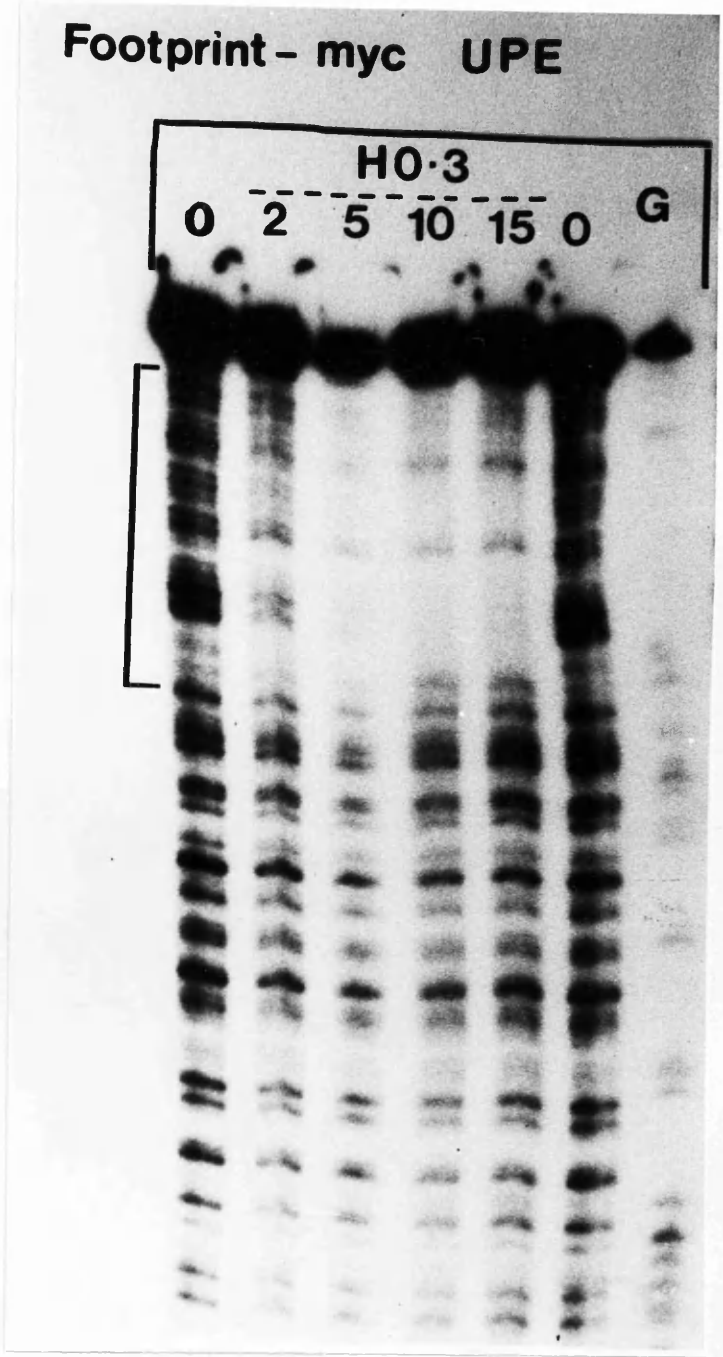
a. Average total counts (cpm) from three to six separate experiments of duplicate flasks as a percentage of pB14. Vertical bars represent the range of values obtained when the LaTK cella were used as the transfection recipient.

b. Equimolar amounts (1-20μg) of donor DNA were transfected into LaTK$^-$ (continuous line) and HeLa cells (hatched line) as described in Materials and Methods. CAT activity was assayed as described in Materials and Methods.

Figure 61.

Photoraph of the *in vitro* DNase I Hypersensitivity Pattern of the UPE Domain.

The H.O3 cell extract was used (see Materials and Methods for detailed description of the experimental technique). The Sp1-enriched cell extract has shown these sites to be at least Sp1-like (J.lang and N.Wilkie, personel communication). The samples shown are labelled at the 5' end (the experiment was repeated twice). The Roman numeral refer to the footprint (DNase I protected domain) observed. It is possible that this one relatively large protected domain comprises of three smaller domains (unfortunatly no gel could unambiguously clarify this).



V

Figure 62.

Sequence of the UPE Region Indicating the *in vitro* DNase I Footprints (Protected Domains).

The footprints indicated were obtained using the Sp1-enriched protein extract (see figures 64). The relevant restriction sites are shown. The *in vitro* DNase I protected domain is boxed. The arrows indicate *in vitro* DNase I hypersensitve sites.The consensus sequences (see figure 74) for Sp1 is shown by dotted lines. The *in vivo* DNase I hypersensitive site (Siebenlist et al, 1984) is shown by the Roman numeral III-2.

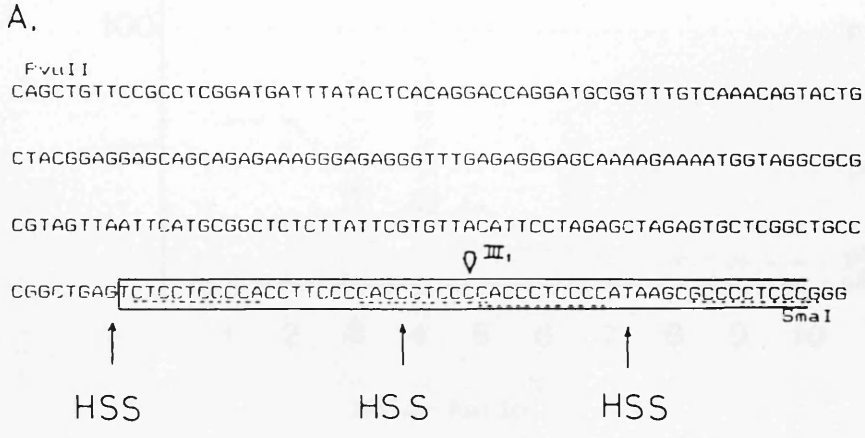
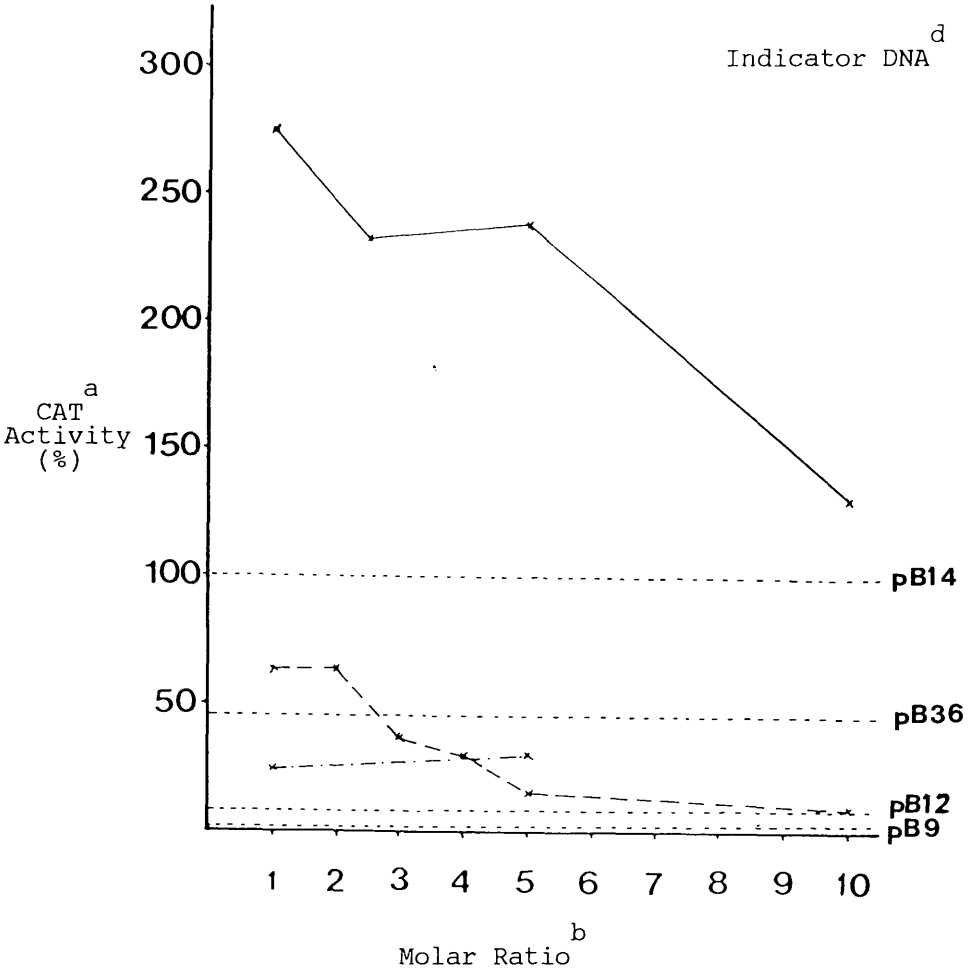


Figure 63.

Competition Assay using the Upstream Promoter Element (UPE) in LaTK Cells.



a. Average total counts (cpm) from one experiment of duplicate flasks represented as a percentage of pB14.

b. Donor indicator DNA (equimolar amounts of 2μg) was mixed with increasing equimolar ratios (0-10) of competitor DNA (pB57), and adjusted to 20μg/ml with pUC12 carrier DNA.

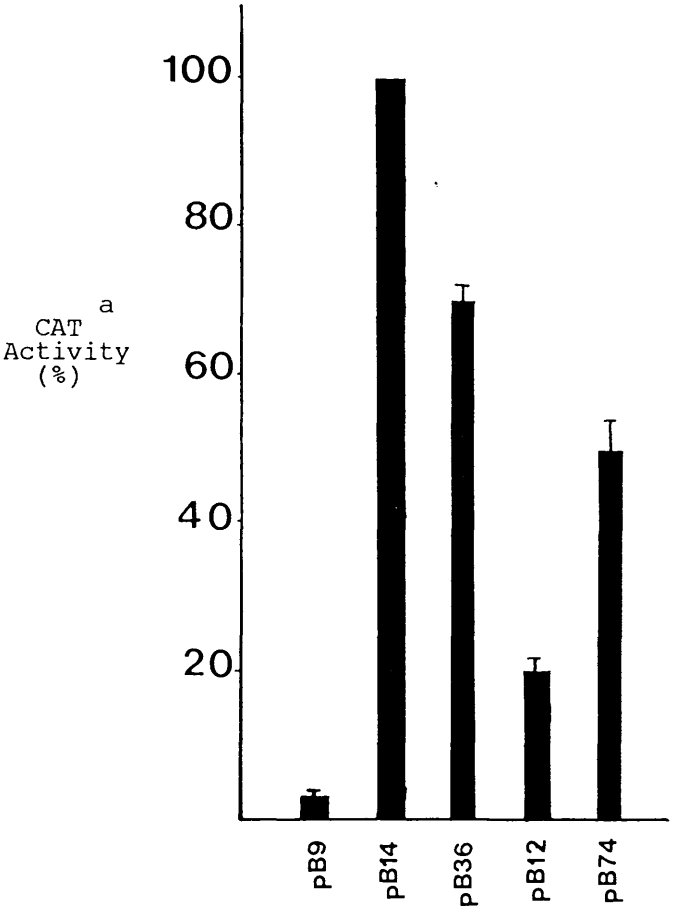
c. The calcium-phosphate transfection of LaTK$^-$ cells was done as described in Materials and Methods. The resultant CAT activity was assayed as described in Materials and Methods.

d. Key : Indicator : Competitor
continuous line = pB14 : pB57
hatched line = pB36 : pB57
hatched/dotted line = pB12 : pB57
dotted lines = indicator plasmids alone (pB14, pb36, pB12, pB9)

e. Schematic representation of relevant plasmids.

Figure 64.

Deletion Analysis of the UPE in LaTK Cells.



a. Average total counts (cpm) from two separate experiments of duplicate flasks as a precentage of pB14. The bars indicate the percentage variation in values observed.

b. Equimolar amounts (of 8µg) of each donor DNA was transfected into Latk$^-$ cells as described in Materials and Methods. The resultant CAT activity was assayed as described in Materials and Methods.

3.6. Preliminary Analysis of Far Upstream Region

Although the 5' deletion studies did not indicate any obvious functional significance in the far upstream region (-2290 to -1248bp) there are two *in vivo* DNase I hypersensitive sites (I and II-1) located within this region. One of these sites (II-1) is located near an Nuclear Factor-1 (NF-1) consensus sequence, and does indeed compete for NF-1 binding in an *in vitro* assay (Siebenlist et al, 1984). In addition, sequence analysis shows the presence of several interesting putative consensus elements.

Due to the restricted range of recombinants analysed, these results are only preliminary, however, they do indicate the presence of two putative domains (one positive, the other negative) within the far upstream region (-2290 to -1248) which deserve further analysis.

3.6.1. Recombinants Plasmid Constructions

All constructions used pB9, pB12, pLW2, pB30 and pB18 as the parental plasmids (see sections 3.2.1.and 3.3.1.).

(i). Recombinants Containing the Far Upstream Region

ClaI digestion and Klenow flush-ending of pB18 followed by HindIII linker modification and self-ligation gave pB47 (see figure 65). HindIII digestion of pB47 releases a 1.1Kb HindIII fragment (containing DNase I HSS I and II-1) which was inserted in both orientations (pB70 and pB71; see figure 65) into HindIII digested pB30 (containing the ε-globin promoter), and in a 5' to 3' orientation (pB69; see figure 65) into HindIII digested pLW2 (containing the HSV-IE promoter).

(ii). Recombinants Containing Sub-Regions of the Far Upstream Region

By digesting pB69 with AccI and SmaI, plus Klenow flush-ending and subsequent self-ligation, pB76 was generated (see figure 66). This recombinant contains only the c-myc sequences encompasing DNase I HSS I (-2290 to -1530) upstream of the HSV-IE gene promoter (pLW2). HincII digestion and subsequent self-ligation of the resulting linears of pB12, pB69 and pB70, recombinants with only the very far 5' region (307bp) of c-myc upstream sequences (-2290 to -1983bp) located 5' to the promoterless CAT gene (pB27), the HSV-IE gene promoter (pB78) and the human ε-globin gene promoter (pB77) were constructed (see figure 67). This 307bp fragment is designated E in these experiments.

3.6.2. Restriction Enzyme Analysis of Recombinant Plasmids

Extensive restriction enzyme analysis has been used to identify each of the recombinant plasmids described. Representative digests are shown in figures 68, 69 and relevant restriction maps in figures 8j.

(i). Plasmids Containing the Whole Far Upstream Region

The ClaI site of Pb18 aws converted by linker modification to HindIII in pB47, and upon HindIII digestion of pB47 the two c-myc fragments (1035bp and 915bp) identified. The 5' region (1035bp fragment) was inserted into pLW2 (pB69) and pB30 (pB70 and pB71). HincII digestion of pB69 (4080bp and 845bp), pB70 (4070bp and 845bp) and of pB71 (4605bp and 310bp) confirmed the oreintation (the size of the smaller fragment was diagnostic) of the c-myc derived sequences.

(ii). Plasmids Containing Sub-Regions of the Far Upstream Region

PstI digestion of pB76, pB77 and pB78 gave respectively diagnostic fragments of 1510bp, 1115bp and 1135bp, confirming the size of the c-myc derived fragment. HincII digestion of pB27 gave a 3.8Kb linear fragment, again identifying the size of the c-myc derived fragment.

3.6.3. Activity Within the Far Upstream Region

Several transient transfection experiments with thse far upstream recombinants were done using LaTK cells as recipients. Depending on the test promoter different equimolar DNA amounts were transfected: equimolar ratios of 2μg for the HSV-IE promoter and promoterless CAT recombinants and an equimolar ratio of 10μg for the ε-globin promoter recombinants. The greater DNA amounts were required for the ε-globin promoter recombinants due to the very weak activity observed when using this promoter.

In this set of experiments pB69, pB70 and pB71, with the controls pB9, pB30 and pLW2 were transfected into LaTK cells (see table 6). 5' deletion through this region had no affect on CAT activity (see section 3.2.4.), however, when inserted 5' of the HSV-IE promoter (pB69) a slight reduction in CAT activity was observed. A greater reduction (4- to 5-fold) was observed when this region was inserted, in either orientation, 5' to the ε-globin gene promoter (pB70 and pB71). Thus, when placed 5' to a promoter this region represses expression (although to a lesser degree than NRE-2), and these results therefore imply the presence of a putative weak second negative regulatory element (NRE-1) within this region.

The two obvious candidate locations for such a region are the two in vivo DNase I hypersensitive sites (I and II-1), although other regions can not be excluded. In a

preliminary attempt to localise the NRE-1, recombinants containing either the DNase I HSS I or II-1 were constructed and analysed by transient transfection into LaTK$^-$ cells as descibed below.

3.6.4. Activity upon Deletion of Site II-1 from the Far Upstream Region

Deletion of site II-1 results in only the far 5' 760bp of c-myc upstream flanking sequences remaining. This region not only contains the *in vivo* DNase HSS I which has been shown to be conserved in the mouse c-myc (Fahrlander, Piechaczyk and Marcu, 1985), but 3 closely-associated enhancer core sequences (for review see Khoury and Gruss, 1983). Obviously sequence inspection does not imply function but it does allow preliminary experimentation to be directed.

When this region was inserted, 5'-3' (pB76), upstream of the HSV-IE gene promoter (pLW2) no alteration in CAT activity was observed (see table 6). This result implies that the region deleted from pB69 to give pB76 (-1530 to -1248) would contain the repression activity seen for the whole far upstream region (see section 3.3.4.). This 282bp region contains the DNase I HSS II-1 which has an associated NF-1 concensus sequence, but does not bind NF-1 in an *in vitro* assay (Siebenlist et al, 1984). A putative weak second negative regulatory element, NRE-1, is tentatively located within this region approximately 1300bp upstream from P1.

This region contains homologous sequence strings to the NRE-2 domain (e.g. the NF-2 element), and to investigate if there were any functional similarities between the two, a competition experiment was done (see section 3.4.5). The competitor pB51 (containing the 445bp NRE-2 fragment) was transfected at a 5 molar excess with the indicator pB69 (see table 6). The 445bp KpnI-AccI NRE-2 competitor fragment relieved the slight repression, observed for pB69 when

compared to pLW2. Previously pB21 was shown not to (or only very slightly: see section 3.4.5.) compete out repression, i.e. NRE-1 does not compete out NRE-2 but the NRE-2 can relieve the slight NRE-1 mediated repression. This implies that at least one of the trans-acting factors (probably an accessory factor) involved in the repression mediated through NRE-2 may be required for NRE-1 activity.

3.6.5. Preliminary Analyis of Enhancer Core Sequences

Upon sequence inspection of this far-upstream region, 3 enhancer core sequences were found near to the DNase I HSS I. Two of these consensus sequences were exactly 40bp apart, and are exact copies of each other. Several possible Sp1 consensus sequences are also present in this region. It must be stressed that these enhancer core sequences have been identified by sequence homology alone. They are located in the distal 320bp region of the c-myc 5' flanking sequences of pMC41, just upstream of the in vivo DNase I HSS I (Siebenlist et al, 1984). The 320bp HindIII HincII fragment (designated E) used in this series of experiments contains the 3 enhancer core sequences, but does not contain the centre of the in vivo DNase I HSS I.

LaTK$^-$ cells were used as the recipients in these experiments (see section 3.6.4. and table 6). Earlier results from analysing pB76 expression levels implied that these sequences do not function as an enhancer, and pB78 CAT activity (being the equal to that of pLW2) agrees with this conclusion. Both these constructs use the HSV-IE gene promoter to drive CAT expression. However, when this E region is placed 5' to a promoterless CAT gene (pB27) a dramatic induction in CAT activity is observed. This level of activity although substantial (12 times that of pB9) it is only 1/6th that of pLW2. This level of activity is comparable to results obtained on enhancer activation of a promoterless gene (Lang, 1984). In addition, when this region is placed 5' to the human ε-globin gene promoter

(pB77) the resulting CAT activity is double that of pB30.

These results tentatively imply the presence of a positive regulatory element (PRE) located in this distal 320bp fragment.

Figure 65.

Structure and Cloning Strategy of Far Upstream Region c-myc-CAT Recombinants pB47, pB69, pB70 and pB71.

For a detailed description of the cloning strategy see section 3.6.1. See figure 8j, 71 and 72.

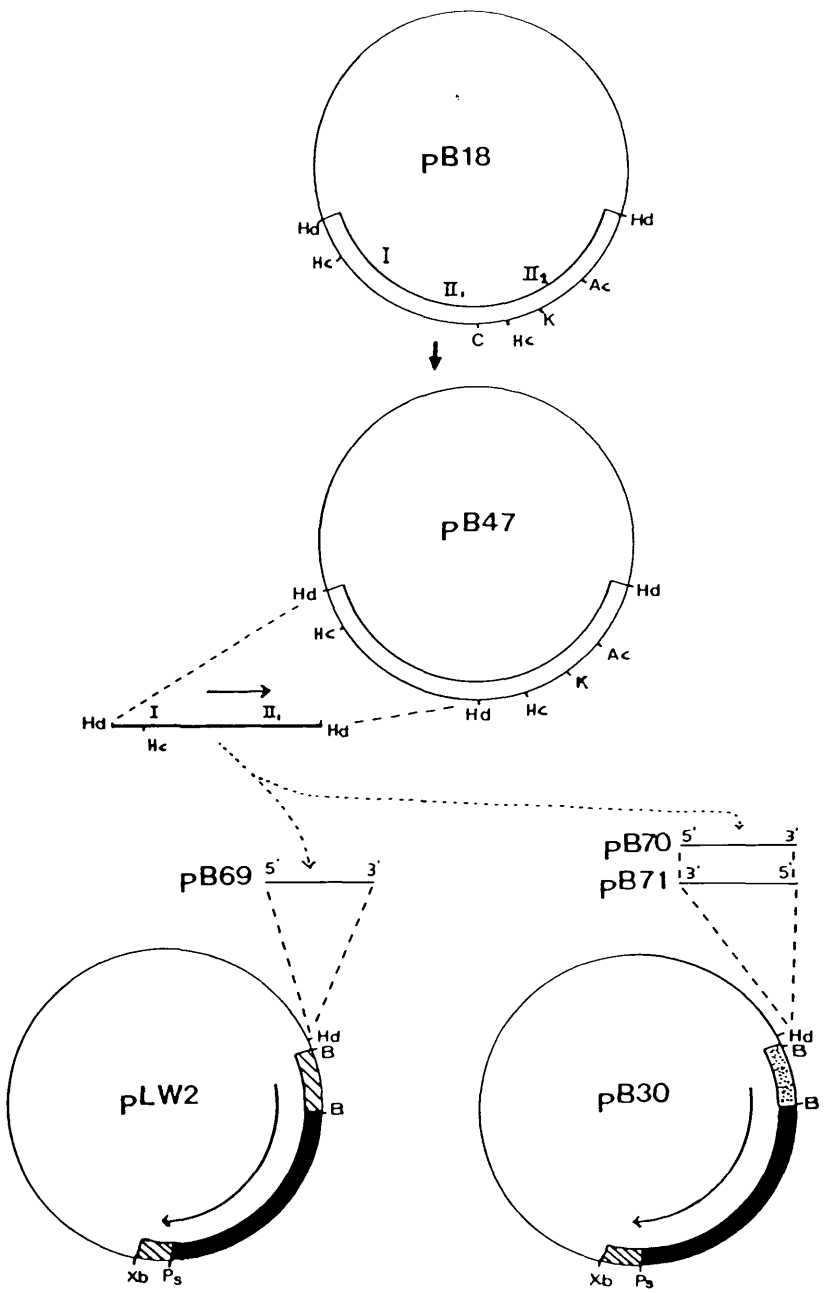


Figure 66.

Structure and Cloning Strategy of c-myc-CAT Recombinant pB76.

Restriction sites which are destroyed upon ligation are marked with a diagonal line across them. For a detailed description of the cloning strategy see section 3.6.1. See figure 8j and 71.

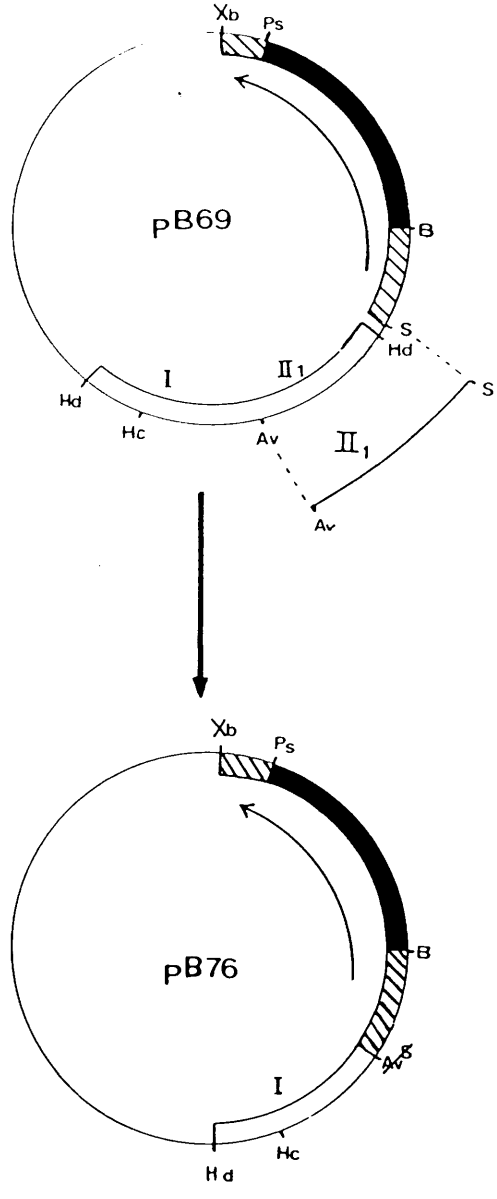


Figure 67.

Structure and Cloning Strategy of "E" Domain c-myc-CAT Recombinants pB27, pB77 and pB78.

For a detailed description of the cloning strategy see section 3.6.1. See figure 8j and 71.

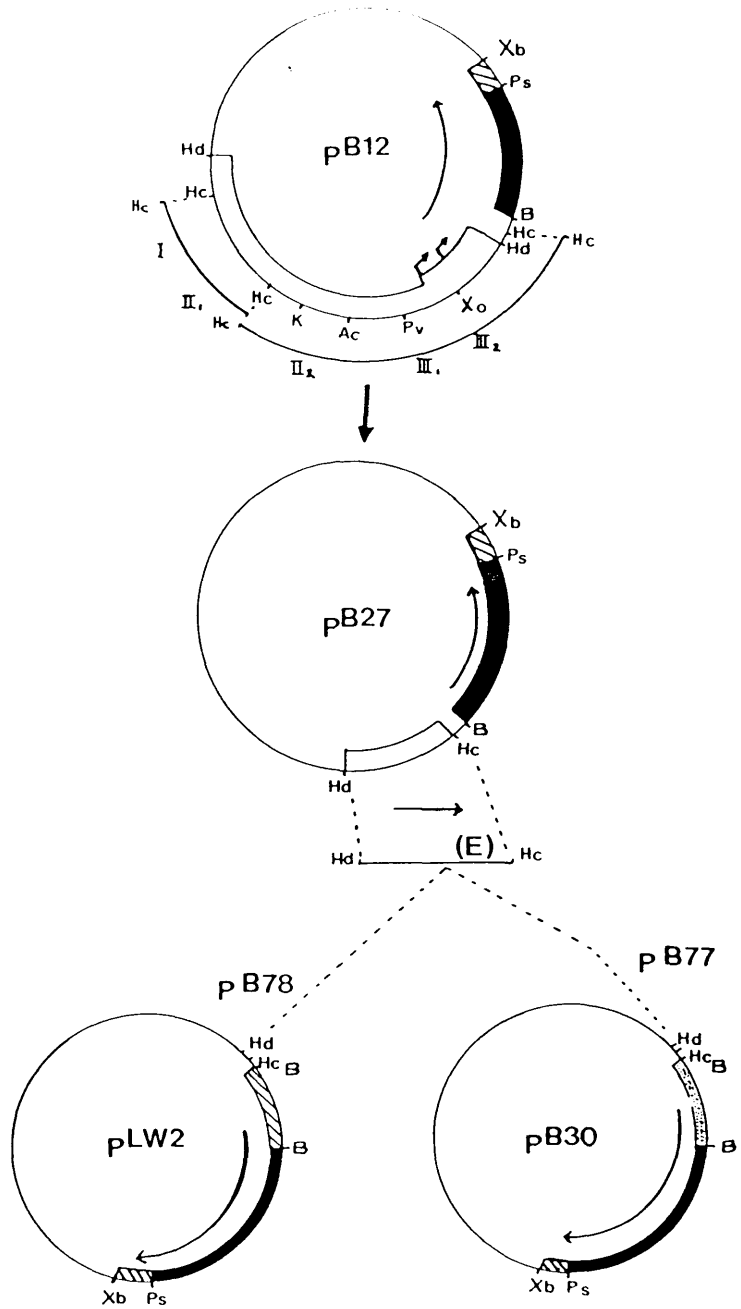
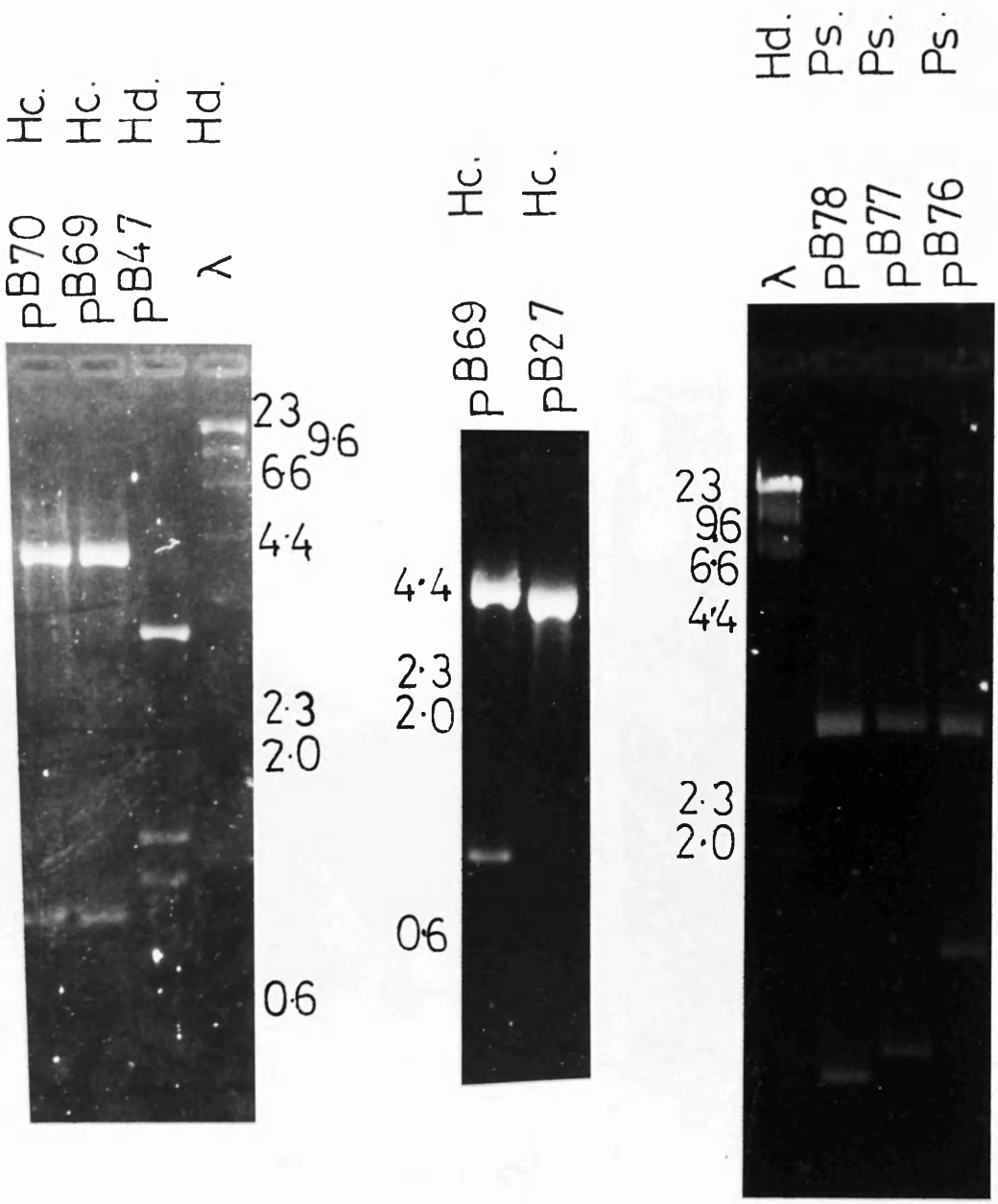


Figure 68 .



1% agarose gel.

Figure 69.

1% agarose gel.

Table 6.

Transient Expression of Far Upsteam Region Recombinants in LaTK⁻Cells.

Donor DNA	DNA Amount[a] 2	10	expts.[b]
pB9	1.4[c] ±0.4[d]	18.9 ±2.1	4
pB27	16.4 ±0.9	ND	4
pLW2	100.0	ND	4
pB69	74.2 ±8.9	ND	3
Comp.pB51[e] pB69	116.0 ±11.0	ND	3
pB76	102.9 ±6.3	ND	2
pB78	95.3 ±8.9	ND	3
pB30	ND	100.0	3
pB70	ND	27.2 ±1.1	3
pB71	ND	19.9 ±3.2	3
pB77	ND	209.0 ±15.2	2
Comp.pB48 pB77	ND	53.8 ±1.1	2

a. Equimolar amounts of 2 or 10µg Donor DNA were transfected.

b. Number of separate experiments of duplicate flasks.

c. Average total counts (cpm) per flask as a percentage of pLW2 and pB30 respectively. ND= not done.

d. Range of variation of values.

e. Competitor plasmids (prefix = Comp.) were co-transfected at a 5 molar excess over the indicator plasmid.

f. Donor DNA was mixed with carrier salmon sperm DNA to a final concentration of 20µg/ml in the calcium-phosphate precipitate and transfected as described in Material and Methods.

g. CAT activity was assayed as described in Materials and Methods using only 50µl of cell extract.

CHAPTER FOUR

DISCUSSION

Before discussing the significance of the results of this work, and their relevance to observed biological characteristics, the usefulness of tissue culture experimentation, and any drawbacks with the respective assay systems utilised, need to be mentioned.

4.1. Gene Expression Assay Systems

Expression can be defined in several ways: the level of transcription, mRNA abundance or the amount of translation product. Quantification of expression can be by assay of the structural gene product, either for the RNA or protein, or by measuring its effect in a biological system. Obviously the level of expression can be thought of as a direct consequence of the activity of the gene's regulatory domains (transcriptional or post-transcriptional). Thus any expression assay concommitantly measures that gene's transcriptional regulatory regions (within the constraints of the assay system used). Recombinant DNA techniques can be used to "remodel" the cellular chromatin *in vivo* in defined ways, and thus precise characterisation of the regulatory domains can be carried out. A large number of assay systems have been developed which allow measurement of gene expression within cell-free lysates *in vitro* and within live animal cells *in vivo*.

In vitro assay systems (Weil et al, 1979, Manley et al, 1980) have allowed the characterisation of a number of eukaryotic gene promoter sequences. Unfortunately such assays do not always reflect the true *in vivo* situation for two reasons. Firstly they rely on fully functional transcriptional machinery within the cell-free lysates and secondly, the template is added naked DNA and as such the presumably important role of secondary structure is not taken into account.

The *in vivo* assay systems allow the quantification of the level of expression of exogenous genes introduced into

cultured mammalian cells, thus bypassing some of the drawbacks of the in vitro systems. Such a system must incorporate within its methodology a gene transfer mechanism for the introduction of genes into the cell nucleus. There are now many such techniques, e.g. calcium phosphate precipitation,DEAE-dextran, microinjection, electroporation, liposome-mediated transfer, protoplast fusion, chromosome-mediated transfer. Each of these methods have their own specific preferred application, e.g. DEAE-dextran transfer for assaying the infectivity of viral DNA, microinjection when cell numbers are limiting (Spandidos and Wilkie, 1984).

However, the most widely used technique due to its ease, and versatility is that involving cellular uptake of the exogenous DNA via a calcium phosphate precipitation (Graham and van der Eb, 1973). Although this method is extremely useful for introducing exogenous DNA into cells, there are some cell lines in which the take up of DNA is very inefficient, e.g. HL-60 and Friend cells (G.Birnie and J.Frampton, personal communication). Perhaps electroporation will allow a more efficient DNA transfer into these cells. Recenty, vectors based on viruses, e.g. vaccina virus, adenovirus and retrovirus, have been developed and early results imply that they may prove very useful in the future. Each vector system has its own points of merit, e.g. retroviruses allow single copy transformants to be generated. However, these systems also have drawbacks (see section 4.1.2.). Recently techniques have been developed which allow the entry of DNA (naked cloned DNA or retrovirus vectors) into fertilised oocytes thus producing transgenic animals (Adams et al, 1986; Palmiter and Brinster, 1986). Perhaps this will allow unambiguous analysis of gene regulation in vivo.

After introduction of the exogenous DNA into the cell, expression can be measured either after a few days (transient assay; Benoist and Chambon, 1981; Gorman et al, 1982) or after a few weeks (a long-term expression assay;

(Spandidos and Wilkie, 1984). Generally transient expression assays are convenient monitors of promoter activity in that they are both rapid and measure gene expression before integration into the host cell genome.

Two basic methods have been employed for transient assays. Firstly, by the direct ligation of a promoter sequence to the coding sequence of a reporter gene (e.g. CAT, tk), the promoter's activity can be measured (Gorman, Moffat and Howard, 1982a; Lang, Spandidos and Wilkie, 1984). This technique relies on the reporter product being either easily detected above the endogenous gene product, or there being no endogenous gene product, e.g. the promoter product is of prokaryotic origin. Secondly, the gene of interest (control and coding sequences) can be transfected into the relevant recipient cell and the resulting mRNA levels can be quantified (e.g. by Northern blot, S1 analysis, in vitro run-on assay). This technique is not as easy, quick or versatile as that of the reporter system, e.g. where a eukaryotic gene is being analysed in a mammalian cell line the expression of an exogenously introduced gene has to be easily identified above that of the endogenous gene's mRNA.

Long-term transformation assays have the advantages that stable cell lines can be developed and that induced phenotypic changes can be monitored (Spandidos and Wilkie, 1984). To enable the generation and identification of stable transformants a selectable marker is commonly used, e.g. tk, APH, DHFR. However, some of these markers require the use of mutant cell lines as recipients, e.g. tk, and thereby limit their versatility. Although this method allows the stable host cell regulation of the introduced gene, the actual site of integration into the host cell's genome could have a drastic effect on the specificity of expression. In addition, there is the possibility that the introduced DNA may become rearranged during either its intergration into the cell or sometime after (Sussman and Milman, 1984).

There is a good correlation between the expression levels obtained during the transient phase of expression and the subsequent biochemical transformation frequences, but biochemical transformation can be the more sensitive technique (Sussman and Milman, 1984; Spandidos and Wilkie, 1984; Lang, 1985). In the studies described in this thesis, transient assays were undertaken in preference to long-term transformation to allow a greater range of recombinant plasmids to be analysed, therefby hopefully answering more questions. To investigate the same range of recombinants by the establishment, and the subsequent characterisation, of stable transformants would take considerably longer than if the same experiments were done in a transient expression system. The actual assay system, using the chloramphenicol acetyltransferase (CAT) gene as the reporter gene, is described below (see section 4.1.1.).

4.1.1. A Critique of The CAT Assay

Although RNA levels represent a good gauge of promoter activity, quantification of RNA is tedious and it is often difficult to obtain accurate measurements unless the promoter being analysed is particularly strong or amplified. It is often preferable to determine the function of a promoter by joining it to a second gene segment (which lacks its own promoter) coding for a readily assayable protein, the reporter gene. In the CAT assay (Gorman, Moffat and Howard, 1982) the bacterial chloramphenicol acetyltransferase (CAT) gene is used, with the enzymatic function of CAT being both quickly and precisely assayable.

The transient assay system based on the control of the bacterial chloramphenicol acetyltransferase (CAT) gene, the reporter gene, is a convenient technique for the characterisation of eukaryotic transcriptional control sequences in a cell culture situation. Importantly, it is sensitive enough to detect the variations in expression levels obtained with different promoters. Primer extension and S1 nuclease studies have been used to show that the

steady state mRNA levels correlate with the level of CAT activity (Gorman et al, 1982), and that RNA initiation occurs at the correct point (Gorman et al, 1982; Alwine, 1985; Remmers, Yang and Marcu, 1986). A major advantage of the CAT assay over some other transient assay systems (e.g. tk), is that any cell line could probably be used as the transfection recipient. Thus the CAT assay presents itself as a convenient system in which to monitor the expression of specific promoter DNA sequences, which correlates very well with the actual mRNA levels. Nevertheless, this does not imply that all promoter sequences can be identified by this system (e.g. control elements may be located within transcribed sequences).

This assay shows a high level of specificity due to the CAT gene being absent from all mammalian cells looked at, (metabolism of chloramphenicol being via glucosyl transferase). Thus there is no endogenous background CAT activity. In addition, there is no interference from other endogenous enzymatic activities which could compete for the substrate or cofactors (Gorman, Moffat and Howard, 1982). The assay is rapid, sensitive and reproducible, as well as only requiring the low grade radioisotope ^{14}C as its labelled marker. Finally, quantification of the result can be easily and accurately carried out in a scintillation counter. These advantages make the CAT gene marker, at the moment, the most convient developed assay for transient expression studies. There are other transient assay systems available, eg.β-galactosidase, but they are either more cumbersome or less sensitive than the CAT assay.

4.1.2. A Critique of Cell-Culture Experimentation

Firstly, the cells used as recipients, although considered to represent the cell-type they were originally derived from, require their phenotype to be questioned. After long-term tissue culture propagation their actual relatedness to their parental phenotype is not always clear.

They are abnormal, immortal and some being polyploid (Freshney, 1983). Therefore the molecular events which occur within them need not necessarily mimic what occurs in the parental cell type in its normal environment. In the studies presented within this thesis HeLa cells were used at two stages: once for transfection recipients, and again as the source of the nuclear protein extracts. In the first instance, they are an adherent cell-line, while the nuclear extracts were obtained from a suspension cell line. Both cell-lines go under the name HeLa, but they have been sub-cultured for their respective properties. It is hard to imagine they are identical. Another example concerns mammary epithelial cells. No mammary cell line available responds correctly to the normal cocktail of lactation hormones to which the mammary gland is responsive in vivo, e.g. expression of the milk genes.

Secondly, any exogenously introduced DNA may not form the same overall conformation with respect to the protein: DNA association as found in vivo. In addition, for transient expression assays, the actual plasmid copy number per cell may be as many as 1,000 (Perucho et al, 1980). Again this is very different from the in vivo environment. Obviously stable transformation overcomes part of this problem, but as a consequence other problems (e.g. the chromatin state at the site of integration can confer position effects) are raised. However, as long as these caveats are put upon any interpretation of transfection results (or any cell-culture analysis) then meaningful studies will be obtained by such experimentation. Obviously, such in vivo analysis will be superior to the in vitro techniques we have available, although the in vitro assays could be used to complement the in vivo results. Perhaps, the recently developed transgenic techniques will allow the accurate in vivo analysis of these types of questions. Having just criticised both the present day in vitro and in vivo techniques, their already proven usefulness in quickly giving an insight (but it is only that)

into the answers, has to be stated

4.1.3. Cell-Specific Expression and Transfection Efficiency

Expression of the c-myc recombinants was observed in each cell line analysed, i.e. fibroblastoid, epithelial, erythroid - both of human and murine origin. In addition, the repression function was observed in all the cell lines transfected in these studies. This would imply both a lack of tissue- and species- specificity for the c-myc transcriptional control domains identified in these studies.

The repression ratio observed for the c-myc recombinants varied between the different cell lines (see section 3.2.3.). The repression ratio represents the fold difference in activity (cpm) between the repressed (NRE-2 containing) and the non-repressed (NRE-2 deleted) c-myc-CAT recombinants. There are two obvious candidate explanations for this. Firstly, the variations in repression ratio could reflect the tissue and species from which these cell lines were derived. However, there is no obvious correlation between the variations in repression ratios and the various cell lines with respect to their origins. In addition, the variations in repression ratios are small, while the variation in absolute expression levels are relatively large. Thus the second explanation, that the variation in expression is due to the different transfection efficiencies (DNA uptake and subsequent expression) of the different cell lines, is more plausible, i.e. different numbers of plasmid are being expressed in the different cell lines. LaTK$^-$ and HeLa cells are routinely used for transient expression studies in this and other laboratories, while NIH-3T3 cells have been the recipient cells used in the now historical identification of the ras gene family. Therefore, these cells have an assumed high transfection (DNA uptake and expression) efficiency, and in correlation, it is in these cell lines in which c-myc recombinants are fairly active. In addition GM468 and MRC-5 have been found in our

laboratory to be poor recipients (D.Spandidos and J.Lang, personal communication), and again correlating with this, it is in these cell lines in which a low level of activity was observed. However, K-562 cells in this laboratory are considered poor recipients (D.Spandidos and P.Montague, personal communication) but give relatively high c-myc expression levels. This may in part be due to the uptake and expression efficency of K-562 cells. DNA titration experiments with K-562 cells (see section 3.4.4.), even at 10-20μg of transfected DNA did not show a titration effect. Indeed, the effect was comparable to that seen at 1-2μg for LaTK$^-$cells. This may indicate that only a small proportion of the DNA present in the transfected precipitate is actually expressed within K-562 cells.

In an attempt to identify any variation in transfection efficiency present between the respective cell lines, a positive control was used. The positive control in these studies was the CAT gene driven by an HSV-IE gene promoter (see section 3.2.3.). The proportional levels of activity for the two c-myc recombinants (pBl2 and pBl4) varied quite drastically from the two controls (pB9 and pLW2) in the various cell lines. This may reflect differential basal levels for c-myc expression in these cell types, however, there is no correlation between the cell types used and the activity of the c-myc recombinants in these studies. Another possibility, is that the basal level of expression of the c-myc recombinants may be constant, and it is the positive control (pLW2) which is showing a variation in its specificity of expression. No studies have been done to elucidate this (with reference to pLW2) as far as I know. The tissue specificity of the immediate early enhancer domain of HSV-1 has been analysed (Spandidos, Lang and Wilkie, 1984), but the domain used in these studies is not directly comparable to the promoter domain of pLW2.

In conclusion, without a positve control, which has its tissue and species specificity absolutely defined

(impossible with to-day's methodology), any variation in expression levels observed in this type of assay could be due to either specificity of the promoter in question, or differences in transfection efficiency between the recipient cell lines. However, the c-myc-CAT recombinants were expressed, and the repression function was observed (when the NRE-2 domain was included in the recombinant), in each cell line analysed. This implies a certain promiscuity with respect to tissue- and species- specificity for the regulation of c-myc expression.

The erythroid cell line K-562 (a human erythroleukaemia derived cell line) was included in these studies, as its endogenous c-myc expression levels are greater than that found in other erythroid cells (Wong-Staal et al, 1983). It was therefore hoped that this cell line would be permissive for expression of the exogenously introduced c-myc recombinants. If this was indeed the case, it could represent a lack of repressor (or functional repressor), as no gross rearrangement of the endogenous c-myc locus has been found. However, the results presented show clearly that a functional repressor is present and active in the K-562 cell line, as it represses exogenously introduced c-myc-CAT recombinants. As a correlation to this, the elevated endogenous c-myc mRNA levels may not be due to a lack of repressor, but may indicate that either a rearrangement of the NRE-2 domain or an alteration in stability of the c-myc message has occurred. Another possibility is that of some functional alteration in the positive regulation of the endogenous c-myc in this cell line. Sequencing of the endogenous K-562 c-myc 5' flanking sequences would help in our understanding of the mechanism by which the c-myc gene is deregulated in these cells.

4.1.4. A Critique of in vitro DNase 1 Footprinting

Once having identified a requirement for a trans-acting factor, identification of this factor is obviously an

important consideration. To start these studies in vitro DNase 1 footprint analysis was carried out. This is now a widely used technique, but there are obvious drawbacks associated with it.

This is an in vitro technique and therefore cannot be unambiguously taken to indicate what happens in vivo. At best, it will tell us the likely involvement and location of DNA-binding proteins, allowing if nothing else, a strategy for further analysis to be proposed. To give an example of the possible problems: the sequence 5'-AGCCCTCCCG-3' (located in the middle of exon 1) which fits the Sp1 consensus except for the 5' G (should be C or T), gave in only one experiment (result not shown) a very faint footprint when using the Sp1 enriched extract. Possibly this footprint was an in vitro assay antifact, caused by the presence (in an inappropriate position) of a Sp1-like consensus sequence. However an in vivo role for a protein interaction at this site can not be ruled out. There are several other Sp1-like consensus sequences which do not bind a protein, by in vitro footprint analysis (personal communication R.Tjian: this thesis). Therefore the presence of a consensus sequence does not always result in a protein interaction. Presumably the actual recognition sequence requirements for such a protein have yet to be stringently defined.

When in vitro footprinting is combined with in vitro transcription analysis, as has been done for Sp1 (Kadagowa, Jones and Tjian, 1986), a direct in vitro correlation between protein binding and transcriptional activation can be demonstrated. This, although not by in vivo methods, still forms a good argument for a functional role for the trans-acting factor in transcription regulation. Since some of these factors have been purified, the near future hopefully holds some exciting developments regarding the trans-acting factors which regulate gene transcription.

The protein extracts used in the work described in this thesis were kindly given by B.Tjian. They represent various degrees of purification. THe Sp1 and CTF preparations are essentially pure, while the HO.3 extract is basically a crude salt preparation which is thought to contain the majority of cellular trans-acting factors (B.Tjian, personal communication). These preparations are all from HeLa cells. Since the HO.3 extract has had some enrichment it may not contain all trans-acting factors (especially since we can only identify a few).

A technical point of relevance, regarding the actual footprint observed, comes from a report on the 3-dimensional structure of the DNase 1 molecule (Suck and Oefner, 1986). Due to the size of the enzyme, steric hindrance between it and the trans-acting factor already bound to the DNA fragment in question will cause the footprint (DNase 1 sensitive region) to be larger than the actual DNA region bound by protein (by 6 bp at the 5' end and 4 bp at the 3' end). This should help in identifying the actual recognition sequence used by the factor, by reducing the observed size of the protected domain.

One avenue of investigation not attempted during this study, which could utilise the DNase 1 footprinting assay, is to attempt to analyse the presumably different trans-acting factors present in proliferating versus quiescent, or immature versus mature cells. Since several transcriptional control domains have been identified during these studies, there are now localised candidate sites for the interaction of these factors, and experiments could be devised to investigate these questions.

4.2. Identification and Analysis of the Transcriptional Control Elements of the c-myc Gene

The CAT assay allows quantification of expression level, and as such, can not directly distinguish between

transcriptional and post-transcriptional (e.g. stability, transport, translation) regulation. However, presumably cis-acting elements, e.g. the NRE-2 and UPE domains discussed in this thesis, function to regulate the level of transcription, as it is hard to imagine how sequences located outwith the trancribed unit can affect the expression level of that gene by regulating post-transcriptional activities. In the following discussion the accepted nomenclature for nucleotide numbering is used, i.e. increasing negative values to all nucleotides 5' to be mRNA cap site (taken to be P1). The results described in this thesis have identified two transcriptional control regions 5' to the human c-myc and hinted at two additional regulatory elements (see figure 73). Evidence is also presented which indicates a role for at least one trans-acting factor in the transcriptional regulation of the human c-myc gene.

In the work presented in this thesis, the CAT assay has allowed the identification of both positive and negative regulatory elements, located within the human c-myc 5' control region. Due to the versatility of the CAT assay regarding cell line usage, the relative specificity of these elements has been tested. During the course of these studies one important aspect of DNA transfection (probably applicable irrespective of marker system) has emerged, which when utilised has allowed more information on the respective elements to be analysed. Upon increasing the amount of transfected DNA a titration effect was observed irrespective of the test recombinant plasmid. This is considered to represent a titration out of the trans-acting factors required in the regulation of the individual transcriptional control elements due to increasing input of test plasmid molecules. This titration effect, when used in combination with competition experiments, has allowed the identification of a requirement for a trans-acting factor in the repression function seen in the c-myc transcriptional control domain.

A point raised by this titration effect, but not investigated, is what effect do large numbers of exogenously introduced plasmid molecules have on the expression of the endogenous c-myc gene. It is conceivable that the presence of numerous plasmids containing the negative element would allow release from any repression the endogenous gene may be under. Perhaps a more exciting question is, do vast amounts of exogenously introduced DNA fragments containing a positive regulatory element (e.g. Sp1) have an effect on the endogenous gene expression of the recipient cell? To my knowledge, although many transcriptional control domains have been identified by transient expression assays the effect on the respective endogenous genes has not been analysed. This is an interesting question in itself.

Evidence has been presented (see section 5.1.) to indicate that plasmids co-transfected recombine in cis at some stage during the transfection procedure. The presence of very large concatemers post-transfection has been previously reported (Wigler et al, 1979; Weber and Shaffner, 1985). However, not all studies show concatemerisation (Weinberg, Cheng and Conrad, 1986). As discussed, this cis-involvement of co-transfected plasmids must be borne in mind when interpretating competition experiment results. How and when cis-recombination occurs, indeed whether it involves many sites or a specific site, is unknown. Southern analysis of transiently transfected cells may give an answer, but could equally just give a mass of indecipherable bands. Initially, as a negative control, to confirm that insertion of a DNA fragment upstream of the CAT transcription unit did not induce expression if no obvious promoter was present, pB4 was made. This construct although containing the c-myc dual promoter region also contained the intron 1 splice donor site and the first 620 bp of intron 1. As expected no CAT activity was observed. Since the level of activity was slightly less than that of pB9 (promoterless CAT recombinant), in all likelihood the CAT

coding sequences have been spliced out. A dramatic decrease in mRNA stability (or any other post-transcriptional mechanism) brought on by the inclusion of the 3' 40bp of exon 1 can not be excluded. The untranslated first exon of c-myc may confer instability upon the message (see section 1.5.3.). For the low level of CAT activity found for pB4 to be due to a post-transcriptional mechanism, a very dramatic change in the mRNA processing requires to be postulated. Without analysis of the resultant RNA this question can not be conclusively answered. Analysis of the stability of the c-myc message was considered to be outwith the scope of this thesis, but it would merit future investigation.

4.2.1. The Promoter Domain

The promoter domain of the human c-myc gene appears to be a complex region spanning from -350 to +164 (i.e.P2): the first transcribed base of the P1 transcript is taken as +1. Both mRNA start sites have associated, and highly conserved, TATA consensus sequences (murine, feline and human; see figure 72) which are positioned approximatly 30 bp 5' to P1 and P2. These regions have been footprinted and neither have been shown to be protected, although several hypersensitive sites were identified implying some protein:DNA interaction (J.Lang and N.Wilkie, personal communication). TATA-binding proteins have been identified (Wu, 1984). Therefore the lack of a TATA footprint may reflect the absence of the relevant binding protein(s) from the specific protein extract (HO.3) used in these experiments, or that the these protein(s) are present but fail to function unless in a highly enriched form. However the possibility that this sequence is not involved in c-myc transcription (although very unlikely) can not be ruled out.

By sequential 5' deletion of the c-myc upstream sequences, only the proximal 350bp (pB14) have been shown to be required for maximal promoter activity. Upon further 5' deletion, leaving only the proximal 100bp (pB36), a halving

of promoter activity was observed. These effects were shown for both LaTK⁻and HeLa cells, possibly reflecting a certain promiscuity with respect to tissue- and species-specificity. In addition, titration results tentatively imply a possible interaction of a trans-acting factor with the -350 to -100 domain (termed the Upstream Promoter Element, UPE).

The whole promoter region is GC rich (63%) with several (conserved) consensus sequences for the GC-rich/Sp1 binding element being present (see figures 70 and 71). By DNase 1 in vitro footprinting a diffuse footprint was identified (-152 to -100) with the H0.3 extract. This footprint is located on a very GC-rich region (73%) containing 4 putative Sp1-binding elements. The in vivo hypersensitive site III-1 (Siebenlist et al, 1984; Dynan and Tjian, 1985) is localised around -126, i.e. in the middle of this footprint. Since in the in vivo studies, this hypersensitive site is located in a region of relatively poor resolution on the gel, rather than defining a hypersensitive site, it only localises a hypersensitive domain or a close cluster of sites. Therefore the domain around -126 (probably about 50 bp) is hypersensitive to DNase I in vivo. Thus the observed in vitro footprints fall within the in vivo hypersensitive domain.

Subsequent experimentation has further identified footprints (with H0.3 extract) near to P1, located on Sp1 consensus sequences. No footprints were found in the region between P1 and P2 although several in vitro DNase 1 hypersensisitve sites were identified. These sites correlate well with the in vivo site III-2 (Siebenlist et al, 1984), and were in the same relative positions to the sites surrounding Sp1 footprints 5' to P1 (J.Lang and N.Wilkie, personal communication). Thus, there may be a cluster of Sp1 elements just 5' to the c-myc transcription start sites (see figure 70). Such regions within promoter domains have been found before, e.g. for the SV-40 immediate

early, HSV immediate early, mouse DHFR and Ha-ras genes (Kadonaga, Jones and Tjian, 1986; Ishii et el, 1986).

The c-myc promoter recombinant (pB14) contains all of this GC-rich region and additional sequences to -353, (although no additional footprints have been found between -353 and -152). Upon 5' deletion to -100 (pB36) the promoting activity of this region is halved. It is therefore likely that this region functions as an upstream promoter element, possibly by the trans-action of Sp1 (or Sp1-like) factors. Perhaps, as shown for other genes (e.g. Ha-ras: Ishii et al, 1986), for c-myc it is the proximal (putative) Sp1 elememts which have the major bearing on transcription, while the more distal have only a minor role in the regulation of transcription. This is in agreement with the observed results. Why there are (in several genes) clusters of Sp1 sites, when only the proximal few are required for optimal transcription is unknown. It is conceivable that under certain circumstances this region could have a greater function (e.g. respond to specific inducers) which has not been detected by these studies.

Preliminary DNase 1 in vitro footprint analysis on the chicken c-myc gene (which does not contain an obvious TATA consensus) has been done (M. Plumb, in press). Although no purified protein extracts were used, several 20-30bp footprints were found just 5' to the mRNA initiation sites (which cluster to a 7bp domain: Nottenburg and Varmus,1986). Some of these sites overlay Sp1 consensus sequences, of which again there are many in this region (see figure 70).

From these results, it is possible that Sp1 or similar factors may play a crucial role in the promoter domain function of c-myc. When a species comparison is done for this region a strong conservation of these Sp1 sites is found (see figure 70). Both orientations are found, but whether Sp1 functions in an orientation independent manner is now considered unlikely (Kadonaga, Jones and Tjian,

1986). The presence of weak transcripts originating from this region but from the anti-sense strand (Bentley and Groudine, 1986), may be due to the presence of these putative Sp1 elements, since Sp1 has been shown to induce bipolar transcription (for review see Dynan, 1986). Although current data implicates Sp1 as a positive regulator of transcription (Kadonaga, Jones and Tjian, 1986) results presented in this thesis may imply that it has a broader function (see section 4.2.2.).

There is slowly emerging a group of genes which, by lacking a TATA consensus, do not fit in with our now classical description of promoter/initiation sites. These new promoters have a very high GC content, and are nearly all limited to housekeeping genes (for review see Dynan, 1986). Perhaps the c-myc gene could be the classified among them, although it does possess a TATA consensus. Conceivably, genes containing both GC-rich elements and a TATA consensus may form a third distinct class.

However, the UPE domain appears to have a dual personality in respect to its transcription regulatory function. When the 250bp UPE domain (-350 to -100) is internally deleted, or is used as the competitor in competition experiments, the results imply a role for the UPE in regulating the repression of the c-myc gene. This involvement of the UPE in the repression function is not an absolute one, since a DNA fragment containing the NRE-2 domain can repress heterologous promoters (see section 3.3.3.). The interaction of the NRE-2 and the UPE is discussed in section 4.3.

This region of the human c-myc gene (-400 to -100, which contains the UPE defined in this thesis) has been shown to contain a negative regulatory element but no positive activity, by R. Dalla-Favera (Marcu et al, 1986). The 5'deletion results presented in this thesis indicate the UPE to have a role in the positive regulation of c-myc

transcription, although an involvement in the repression function is also seen. As far as I can tell, without seeing the actual data, basically the same constructs and cell lines were used in the work described in this thesis and the work presented by R.Dalla-Favera. At the moment this anomaly remains. In the murine system no negative function has been assigned to the colinear region

4.2.2. The Negative Regulatory Element

(i). Identification and Localisation of the NRE-2.

By 5' and internal deletion (this thesis) an essential component of a Negative Regulatory Element (NRE) has been identified and located to -1052 to -607 (a 445bp region), with respect to the human c-myc gene. The DNase 1 hypersensitive site II-2 (Siebenlist et al, 1984), is located within this region at -750. Recent work on the murine c-myc gene has localised a similar element (the dehancer), to a region between -1188 and -428bp 5' of the first murine c-myc promoter (Remmers, Yang and Marcu, 1986), which encompases the location of the corresponding negative regulatory element in the human c-myc gene. The location of this element correlates well with the proposal of Yang et al (1985), which was based on analysis of murine plasmacytoma translocation breakpoint clusters.

However, there are some surprising differences between the murine dehancer (Remmers, Yang and Marcu, 1986) and the human NRE-2 described in this thesis. They both function in an orientation-independent manner, but the dehancer has a greater repressing activity on an heterologous promoter than on the normal homologous c-myc promoter. The converse is found for the human NRE-2. Although competition experiments were attempted, no evidence for the involvement of a trans-acting factor was found for the murine dehancer, although such an involvement has been convincingly shown for the human NRE-2. Additionally, although a similarly

constructed recombinant to pB36 (5' deleted c-myc-CAT recombinant retaining omly 100bp of 5' flanking sequences) was analysed, no evidence for UPE activity (see section 4.2.1.) was found. Why there are such differences remains a matter for speculation.

Perhaps the same gene is regulated by subtly different mechanisms in the different species. However, this seems unlikely. In the studies presented in this thesis, the human NRE-2 domain responded to trans-acting factors from murine cells, implying conservation of the regulatory mechanism. In the experiments of Remmers, Yang and Marcu, (1986) on the murine c-myc dehancer, the competition was done at only a 2 molar excess of competitior over indicator recombinant. Therefore it is highly likely that the competition effect, rather than not being present, was not observed through inadequate experimentation (i.e. from the competition experiments described in this thesis, at least a 5 molar excess was required to see an effect). Different cell lines (although some are in common) and promoters (SV40 in the murine studies, HSV and ε-globin in the human studies) were used, and this may account for some of the differences. The answer to this question awaits further experimentation.

In addition, the basic constructs used in these two lines of work differ. The murine c-myc-CAT recombinants include about 10bp of intronic sequences (i.e. contains the splice donor site). As already described (see section 3.3.8.), when a similar recombinant was tested for the human c-myc (pB4), no activity was detected. Perhaps the small amount of murine intronic sequence is not enough to form the correct environment for the splicing event. However, these results do imply that the full first exon does not confer extreme instability upon the c-myc message directly (or at least not in the murine system).

(ii). NRE-2 Directed Repression is Mediated by at Least One *trans*-Acting Factor.

Both DNA titration and competition experiiments indicated that the human NRE-2 functioned by the interaction of at least one *trans*-acting factor. *In vitro* DNase 1 footprint analysis showed four protected areas to reside within the 445bp region. All the protected regions obtained with partially purified DNA binding proteins (HO.3) could be associated with footprints obtained with the purified extracts, and identified one strong 5' Sp1 site, one weak Sp1 site, and two strong CTF sites (see section 3.4.6.; figure 71). Transfection competition experiments indicates that at least an essential function of the NRE resided in the 270bp sub-fragment (-882 to -607). The two CTF, and the weak Sp1 footprints are located within this sub-fragment, as is the *in vivo* DNase 1 hypersensitive site II-2 (Siebenlist et al, 1984). As discusssed (see section 4.2.1.), the location of this *in vivo* hypersensitive site probably defines a domain of 50-80bp (or close cluster of sites) which is hypersensitive *in vivo*. Thus the two CTF sites and the weak Sp1-like site are all located within a region which is hypersensitive to DNase I digestion *in vivo*.

However, the two CTF footprints differ slightly, possibly implying a different binding activity, or even the presence of additional proteins. Indeed, there may be two distinct factors (CCAAT binding protein, CBP and CCAAT Transcription factor, CTF) capable of binding to the CCAAT consensus (McKnight and Tjian, 1986). Within the 3' CTF footprint is a weak homology to the NF-1 consensus sequence, which has previously been shown not to bind NF-1 *in vitro* (Siebenlist et al, 1984).

Recently however, NF-1 has been shown to be similar to CTF (if not actually the same; R.Tjian, personal communication). The data presented by Seibenlist et al (1984) constitutes only weak evidence for a NF-1-like factor

interacting with the in vivo DNase I hypersensitive site II-1, and no interaction occurring at site II-2. In this experiment a crude nuclear extract was used to identify the retention of cloned DNA fragments on nitrocellulose filters. The data presented for the identification of the binding domains by restriction digestion of the cloned c-myc DNA fragments does not clearly distinguish the hypersensitive domains II-1 and II-2. Indeed, when different temperatures were used for their protein-DNA binding reaction, a protein was selectively retained by both II-1 and II-2. In addition, they found no protein interaction with the UPE domain described in this thesis. Since the UPE domain was shown (see section 3.5.4.) to bind several Sp1-like factors, and Sp1 is thought to be a relatively common (and abundant) protein, the actual composition of their protein extracts is unclear. NF-1 has been shown to have a role in the transcriptional regulation of at least one viral gene (Shaul, Ben-Levy and De-Medina, 1986). Obviously, these points need to be clarified in the future. Possibly site-directed mutagenesis of each individual site, and in vitro transcription experiments using purified protein extracts would give a clearer picture. At the moment a synthetic oligonucleotide of this region, containing the two CTF sites and the weak Sp1 site is being analysed in the laboratory (N.Wilkie and J.Lang, personal communication).

This suggests that either or both CTF sites and the associated Sp1 site may have a role in modulating the negative regulatory effect. If this is true, it is of considerable interest, since previously both Sp1 and CTF have only been associated with up-regulation of promoter and/or enhancer-like domains (Dynan and Tjian, 1985). Both these elements may have only a supplementary or ancillary (albeit essential) role in regulating transcription. Thus their presence in a negative element would not be too surprising. They may function to hold specific DNA domains in a correct conformation for the interaction of other proteins or in some way increase the binding efficiency of

other proteins. The *in vitro* results presented do not exclude the involvement of other and/or additional *trans*-acting factors (see section 4.4.3.). Perhaps the interaction of multiple closely located sequence-specific transcription factors will emerge as a common property of transcriptional control sequences in general.

Recent data presented by Bentley and Groundine (1986) reported the existence of an upstream transcript (P0) which orginated from the -650 to -550 region. They also identified, at a low level, transcripts originating at this point but from the anti-sense strand. It is conceivable that the Sp1 and/or the CTF sites localised to the NRE-2 may *in vivo* induce this upstream transcript. Perhaps spurious transcription initiation is a consequence of the NRE-2 domain with its associated multiple protein interactions. Upstream start sites, which have been subsequently localised to transcriptional control elements have been found for other genes (Zhu, Allan and Paul, 1984). The open reading frames identified for this transcript from human cells (Bentley and Groudine, 1986) are noticeably not conserved in the colinear murine c-*myc* sequence. What role, if any, this P0 transcript has on the expression of the c-*myc* gene is unknown.

(iii). Characterisation of the NRE-2 Domain.

The results presented in section 3.3.6. and 3.3.7., show that the 445bp (and the 270bp sub-fragment) which contain at least an essential component of the NRE-2, do not confer repression upon an heterologous promoter (HSV-IE promoter). When the same fragments are inserted, in either orientation, upstream of a promoterless/capless CAT gene, transcription of the CAT gene is induced. It is possible that these complementary fragments contain the promoter responsible for the P0 transcript. Since Sp1 functions independently of orientation to promoter transcription (for review see Dynan, 1986), it is possible that it is the Sp1

site within the 270bp domain of the NRE-2 which activates the CAT gene, and in vivo may regulate the P0 transcription start site. This result (the 445bp domain does not repress an heterologous promoter) implies that the repression function may require at least one other trans-acting factor and additional cis-acting sequences not present within the 445bp domain, i.e. the 445bp fragment may have truncated the NRE-2 domain..

Two other potentially important structures of the NRE-2 were gleaned from sequence inspection . Just downstream of the 3' CTF in vitro site lies a small stretch of alternating purines and pyrimidines. Such sequences have the potential to form Z-DNA. Whether this region is required for functional repression is not known, but what makes this small stretch noticeable is the near complete lack of such sequences elsewhere 5' to P1. Secondly the whole 445bp region is GC-rich (51%) with localised sub-regions being 75% GC-rich, one of which contains the above-mentioned weak Spl site.

When a sequence comparison is made with the murine c-myc there are several large regions of homology 5' to exon 1 (Corcoran, Cory and Adams, 1985). However, the sequences about DNase 1 hypersensitive site II-2 do not fall into one of these highly conserved regions. Upon closer inspection small conserved stretches can be identified. The 5' Sp1 region has two possible homologous domains in the murine c-myc, while the weaker Spl footprint has no obvious homology except for a stretch of 15 Gs. Again the region just 5' to the CTF sites is very GC-rich as it is in the human c-myc. However, there is no appreciable alternating purine-pyrimidine stretch to the 3' end of the region as there is in the human sequence. Both CTF footprint regions are fairly well conserved (5' site at 67% and 3' site at 58%) but within this are small stretches of conserved sequences, including the putative CTF recognition sites. Perhaps more important than overall sequence homology is the

general spacing of these elements (see section 4.3.).

Footprint analysis of the murine region would help to clarify these questions. The in vivo DNase 1 hypersensitive site for both the human and murine sequences is colinear (Fahrlander, Piechaczyk and Marcu, 1985) thereby implying some interaction event occuring at this site. Although individual sequences may form the recognition sites for trans-acting factors perhaps the overall geometry of the domain is required for effective binding.

Several other negative regulatory elements have recently been identified in association with eukaryotic genes. There are no striking homologies between these elements (Remmers, Yang and Marcu, 1986), especially since most of these elements have only been given approximate locations. The H1 domain of the mouse β-major globin gene has been shown to repress expression in an orientation independent manner (Gilmour et al, 1984), as does the human c-myc NRE-2. However, unlike the H1 domain, the c-myc NRE-2 does not have a large potential Z-DNA domain. Other negative regulatory elements can repress independently of their orientation. Perhaps, in a similar fashion to enhancers (Sassone-Corsi and Borrelli, 1986), certain negative regulatory domains may share common factors.

Recent investigation into enhancer function has shown that in certain situations these elements can be repressed by some cellular or viral trans-acting factor (Borrelli and Chambon, 1984; Gorman, Rigby and Lane, 1985). The human c-myc NRE-2 domain may have the ability to act both positively and negatively with respect to transcription regulation. It is slowly being realised that the repression of genes may be a common phenomenon.

4.2.3. The Far Upstream Region

The far upstream region (-2280 to -1240) contains two

sequence and two clustered enhancer core consensus sequences (Siebenlist et al, 1984). Associated with the 3' CTF footprint of the NRE-2 is the sequence 5'-TGGAAGGTATCCAAT-3' which is highly homologous to the reported target sequence for the cellular protein nuclear factor -1 (NF-1), but which may not actually bind NF-1 as shown by preliminary *in vitro* competition experiment (Siebenlist et al, 1984; see section 4.2.2.). As well as the presence of this sequence in the NRE-2 domain a similar sequence 5'-TGGAAGGCAGCCAA-3' is located adjacent to the *in vivo* DNase 1 hypersensitive site II-1. Previous studies have suggested that this sequence is a target for NF-1 binding (Siebenlist et al, 1984). Comparing this sequence with that of the murine c-*myc* colinear (to site II-1) sequence, shows this region to be fairly well conserved with two completely homologous stretches of 11 bp and 16 bp. However, the NF-1 consensus is not conserved (Corcoran, Cory and Adams, 1985). In addition there are no CTF recognition sites in either sequence, and neither sequences show any obvious Spl consensus signals.

Of interest, a recent report using computer-assisted analysis of several genes (viral and eukaryotic) identified a consensus sequence, 5'- TTGnnnTTTTTT -3', which was postulated to be involved in negative control of transcription (Renan, 1985). No functional studies have been done to substantiate these observations. Homologues of the consensus sequence, as stated, can be found at many sites throughout the c-*myc* gene, but whether they bind a factor or have any functional significance awaits investigation.

Results described in this thesis implicate the -1525 to -1240 region in possessing another negative regulatory element, and it is therefore tentatively termed NRE-1. This negative function can be conferred upon an heterologous promoter (HSV-1E), and shows a certain promiscuity with respect to its tissue- and species- specificity. The repression conferred by this domain can be competed out by a cotransfected plasmid containing the NRE-2 domain. Thus

there is at least one factor common to both domains. Further investigation is required to characterise this domain. Whether this element has a role in controlling c-myc expression is not known.

Near to the conserved DNase 1 in vivo hypersensitive site I are two core enhancer consensus sequences (5'-GCTGGAAAC-3': Sassone-Corsi and Borrelli, 1986) which are 40 bp apart (see figure 72), the 3' consensus being conserved in the murine sequence (Fahrlander, Piechaczyk and Marcu, 1985). Evidence summarised by Fahrlander, Piechaczyk and Marcu (1985), correlated this region with transcriptionally active genes, however, previous arguments pointed to this region being the recognition site for a repressor (Siebenlist et al, 1984). This region is located approximately 1900bp 5' to P1. This region has the ability when covalently linked to either the CAT gene alone (promoterless) or to a heterologous promoter-CAT hybrid (ε-globin) to activate transcription. Therefore this region is tentatively termed a Positive Regulatory Element, PRE. In competition studies this activator/promoter activity could be competed out, hence the interaction of a trans-acting factor is implied. Again a certain promiscuity with respect to its tissue- and species- specificity was observed.

Sequence scrutinisation of the 5' flanking region of another oncogene, c-mos, which is thought to be down-regulated by a repressor (exhibits very strict tissue specific expression), shows various consensus sequence coincident to c-myc. There are several CTF and enhancer consensus elements, in approximatly the same positions (relative to the transcription start site) as c-myc, within its 5' flanking regiom (Wood et al, 1984). It is conceivable that both the c-myc and the c-mos genes may be regulated, at least in part, by similar mechanisms. However, recent data (from S1 nuclease and nuclear run-on experimentation) indicates that sequences within the c-mos 5' region may function as a transcription termination signal. This region can also, when inserted between the HSV-1 thymidine kinase

gene and its promoter, cause a 99% drop in tk transforming activity (McGeady et al, 1986). The 5' flanking sequences of the c-myc gene (as far as is known) do not contain a transcription terminator activity, thus presumably making the regulatory machinary for these two oncogenes different.

This difference in control does not exclude a role for CTF in the regulation of the c-mos gene. Seibenlist et al (1984) presented tentative evidence for a similar protein interaction with both the DNase I hypersensitive sites II-1 and II-2, and the transcription termination domain of the c-myc gene. If CTF does bind to the 5' flanking region of c-mos gene (and this awaits investigation) and this region does indeed possess a transription terminator, could CTF have a role in transcription termination?

4.3. How Might These Elements Function ?

Although the trans-acting factors involved in the repression of c-myc expression have not been conclusively identified, how the repression function may occur can be speculated on. From the work presented in this thesis, there is a cis-acting element and at least one trans-acting factor involved in this repression. In addition the deletion studies on the UPE imply that this region has some involvement as well as the NRE-2. However, because repression of heterologous (eukaryotic and viral) promoters occurs when only sequences containing the NRE-2 and not the UPE domain are inserted adjacent to them, the involvement of the UPE is probably not a prerequisite for repression (unless both the HSV-IE and human ε-globin promoter domains contain an analogous element). Although only speculation, several models of how the interaction of the repressor with the NRE-2 domain may occur, can be thought of.

4.3.1. Direct Steric Hinderance

According to this model, the repressor upon binding to

the NRE-2 physically blocks the interaction of the RNA polymerase II or a positive transcription factor to the DNA. This is the mechanism used by many bacterial repressors (Pabo and Sauer, 1984). This model evokes the requirement of a positive activator of transcription which would also use the NRE-2 as a recognition signal, or in some situations a lack of repressor. It is hard to envisage a role for the UPE in repression in this model, although it does fit with the enhancer-like properties of the NRE-2.

4.3.2. Indirect Steric Hindrance

According to this model, the repressor once bound to the NRE-2 domain interacts with auxillary proteins (e.g. a histone complex) such that the auxillary proteins block the accessibility of the essential promoter sequences to the polymerase or a transcription factor. This would involve a specific higher order interaction of proteins. The association of transcriptional control domains with scaffold binding regions has been reported recently (Grasser and Laemmli, 1986). A function for the UPE domain as a positive regulator of transcription fits into this model, as does the enhancer-like properties described for the NRE-2. However, again it is hard to conceive how the UPE may augment repression.

4.3.3. Alteration of DNA Topology

According to this model, the interaction of the repressor with the NRE-2 domain induces a structural change in the local topology of the DNA thereby preventing the interaction of the RNA polymerase or a transcription factor with this region. There is a substantial amount of evidence to implicate this method as a form of repressional control (Peck and Wang, 1985; Brahms et al, 1985; Weintraub, Cheng and Conrad, 1986). The easiest explanation for this would involve an alteration (e.g. supercoiling) of the DNA.

This can explain UPE involvement, if this change of DNA structure upon reaching the UPE domain causes a change in that region as well. Such a change could allow a different factor to interact with this region, thus causing repression rather than activation. This factor may then work by steric hindrance (see previous models). This model evokes two proteins, one a repressor, the other an activator, both of which are capable of specifically recognising and binding to the UPE domain. However, it must be noted that repression can occur without the involvement of the UPE domain, although in these cases other promoter elements (i.e. the HSV-IE and ε-globin promoters) may function analogously and somehow innteract with the NRE-2. However, the promoter domain of the HSV-IE gene contains GC-rich elements (possibly Sp1), while the ε-globin gene promoter is A-T rich and contains a CCAAT consensus. Thus domains of these two promoters are not obviously similar. The greatest repression is seen when the NRE-2 is regulating its homologous promoter. Recent evidence indicates, that positive factors also function optimally on their homologous promoter (Garcia et al, 1986). Perhaps the repression function of the UPE can be interchanged with other promoter domains irrespective of their individual elements, with only a quantitative change in the repression function.

4.3.4. Diffusion Blockage

According to this model, the binding of the repressor to the NRE-2 prevents movement of the polymerase or a transcription factor to the transcription start site. The whole upstream region, of the c-myc 5' sequences (HindIII to PvuII), can repress even when in a flipped orientation, this would insist (if this model was true) that the polymerase or co-factor first interacted with the DNA more than 2.3Kb distant from the transcription start site. Whether RNA polymerases can diffuse over such distances is not known. The looping out of the intervening sequences is a possible mechanism by which the potential problem of long distances

could be negotiated. Again it is hard to conceive how the UPE domain may play a role in this model.

4.3.5. Repressor Diffusion

According to this model, the repressor upon interaction with the NRE-2 diffuses to another location where it prevents the polymerase or a transcription factor from reaching the transcription start site. This model is evoked to explain the involvement of the UPE domain. The repressor after using the NRE-2 as an entry site diffuses to the UPE domain thereby shutting down transcription by preventing the normal protein-DNA interaction at this site. In the absence of the repressor the UPE domain would function as a positive regulator of transcription. The repressor need not displace the normal protein-UPE interaction, if upon becoming close to each other, an interaction between the two proteins can take place. Such co-operative binding has been reported for the λ repressor (Hochschild and Ptashne, 1986; Ptashne, 1986).

Whether one of these models, a combination of them or another model altogether accounts for the repression of c-myc can only be speculated on as yet. Hopefully future investigation will answer the question.

4.3.6. How the Far Upstream Elements May Function

Very little is known about these elements. If the PRE located about -1900 bp 5' to P1 has a role (unkown as yet) in the regulation of c-myc transcription, perhaps it is to allow the entry of the RNA polymerase II molecules onto the DNA, or in some way keeps the whole gene domain in an open chromatin configuration by some interaction with the nuclear superstructure. Recently, the association of enhancer domains with scaffold binding regions has been reported (Grasser and Laemmli, 1986). There are two enhancer consensus elements within this domain, one of which is

conserved in the murine c-myc sequence. It is tempting to speculate that the PRE may be involved in the association of the active c-myc gene with the nuclear matrix (see section 4.4.3.).

What role, if any, the NRE-1 domain has on the regulation of c-myc transcription again is unknown. The competition results imply the interaction of a factor common to both NRE-1 and NRE-2. It is conceivable that the NRE-1 domain could regulate repression in other cell types, or stages of cell growth and differentiation, in which the NRE-2 domain functions.

4.4 Involvement of These Elements in Cellular Biological Events

The exact level of involvement of the different transcriptional control domains located 5' to the human c-myc in the various biological situations where c-myc expression is known to be up or down regulated is not as yet known. However, speculation as to their roles can be made.

4.4.1. Cell Division

Although the function of c-myc is unknown, there is great expectation that it may be involved in regulating the normal division of a cell. If this is indeed the case then obviously a cell would want to keep such a gene under tight expressional control. The tight down-regulation (repression) of such a gene is a good candidate for this control. Thus the c-myc gene would be expressionally silent (repressed) until required, being turned on either by removal of the controlling repressor or the involvement of an activator. Both of these factors require a cis-acting recognition sequence, with both sequences conceviably lying adjacent to or even overlapping each other. The NRE-2 domain and/or the UPE domain are good candidates for this element.

Upon serum stimulation of quiescent cells c-myc expression is induced along with the growth response of the cell. Both transcriptional and messenger stability changes have been suggested as the cause of this increase in expression (see section 1.5.). Since in the results described in this thesis transcriptional control domains were identified, the following discussion will concentrate on the possible involvement of these transcription regulatory elements in the control of c-myc gene expression, without directly discussing the possible involvement of post-transcriptional control mechanisms. The c-fos gene is also serum responsive, and a cis-element has been identified which confers this upon the gene (Treisman, 1986). This serum-responsive element, like the NRE-2, exhibits enhancer-like properties. Could the NRE-2, the site from where repression is instigated, also act as the serum-responsible element? Serum stimulation could induce an activator or inactivate (or convert to an activator) the repressor. In either case a structural alteration requires to be envisaged, as protein inhibitors superinduce c-myc expression in this system (Kelly et al, 1983). However, there is no obvious sequence homology between the c-fos serum responsive element and the c-myc NRE-2. Of note is that when the NRE-2 domain is isolated from its normal position it can function to promote transcription of a heterologous promoter (see section 4.2.2.).

The search for the serum responsive domain of the c-myc gene would form a good line for future research. In addition, ingenious experiments could be devised utilising antisense mRNA (Weintraub, Izant and Harland, 1985) to investigate the role of c-myc in the growth response, as has been done for c-fos (Holt et al, 1986).

It is becoming apparent that the same upstream element can act both positively and negatively with respect to regulation of transcription (Borrelli, Hen and Chambon, 1984; Gorman, Rigby and Lane, 1985; for review see Jones,

1986). This change of personality must be due to the specific trans-acting factors involved in each situation. It is conceivable that the NRE-2 domain could have either a negative or positive function, depending on the specific trans-acting factor (or cocktail of factors) interacting with it.

4.4.2. Differentiation

In the mature cell c-myc transcription is repressed. Transcriptional and post-transcriptional mechanisms have been suggested, but the complete repression of transcription would appear most sensible for a cell. Although not always clear, the data amassed to-date implicates the in vivo DNase 1 hypersensitive site I region to be associated with this phenomenon (Siebenlist et al, 1984; Dyson et al, 1985; Fahrlander, Piechaczyk and Marcu, 1985). From the results presented in this thesis, this region may have characteristics of a positive regulator of expression, however, no direct role in c-myc expression has been shown. Perhaps this domain could function to keep the whole c-myc gene domain in an open chromatin-state, possibly by interacting with the nuclear matrix. This site becomes strongly hypersennsitive after induction of HL-60 cell differentiation by DMSO (Dyson et al, 1985), and is the only hypersensitive site seen for the silent c-myc allele in some Burkitt Lymphoma cell lines (Siebenlist et al, 1984; Fahrlander, Piechaczyk and Marcu, 1985). Could this site be the recognition signal for a differentiation-specific protein which somehow represses c-myc gene expression in the mature cell, possibly by preventing association with the nuclear matrix?

This site could conceivably form the key to allowing expression of the c-myc gene. Depending on the trans-acting factor(s) interacting with the PRE, the gene could be in a competent or repressed state (for reveiw see Weintraub, 1985). Once in a competent state, additional factors would

be required to induce (or repress) expression, perhaps with the NRE-2 domain.

4.4.3. Neoplasia

Although there are several instances where the c-myc gene is amplified (Graham, Tindle and Birnie, 1985, for review see Alitalo, 1985), no sequence data is available for these cases. However, in Burkitt's lymphoma or murine plasmacytoma cells the c-myc locus has been (in most cases) rearranged by reciprocal translocation with an immunoglobin locus. In some cases the Ig heavy chain enhancer is positioned adjacent to the c-myc coding exons, but this is rare (Marshall, 1985). However, in all cases the c-myc gene appears to be expressionally deregulated, as it is constitutively expressed.

By analysing the chromosome breakpoints of these cells it is clear that in the majority of cases the NRE-2 domain has been removed. Indeed some of the breakpoints found in murine plasmacytoma cells cluster between -500 and -350 bp 5' of the first c-myc exon (Yang et al, 1985). Translocation breakpoints within this region are also found in Burkitt's lymphoma cells (Dyson and Rabbits, 1985). Thus these translocatons would remove the NRE-2 domain from its normal location. Perhaps the observed deregulated c-myc expression is due to this translocation event which rearranges the positioning of the NRE-2 with respect to the promoter domain. Leder et al (1983) proposed a similar model, except that the promoter domain was implicated as the recognition site for the repressor.

The cases where the breakpoints leave the NRE-2 domain adjacent to the c-myc gene can not be explained as easily. In such cases the constitutive c-myc expression may be due to some interaction with the immunoglobulin locus as a consequence of the translocation which overrides the activity of the NRE-2, or there could be small rearrangements (e.g. point mutations) within the NRE-2

domain, thereby destroying its function. For the majority of known cases no sequence data is available. However, in a B-cell lymphocytic leukaemia, even though the translocation breakpoint is 1077 bp 5' to P1 (i.e. 5' to the NRE-2 domain), several mutations have been identified within the transcribed sequences (Care et al, 1986). Perhaps point mutations/deletions could have disrupted the NRE-2 domain rendering it inactive (sequencing of the 5' flanking region would give the answer). There are other possibilities to explain these translocations and the associated deregulated c-myc, (e.g. mutation of c-myc repressor gene; possibly the pvt-1 locus: Cory et al, 1985).

This raises the question of the identity of the repressor (see sections 1.6.3., 1.6.4.). One candidate is the c-myc gene product itself. No direct experimentation on this question was undertaken during these studies, but the K-562 cell results are of relevance to this point. These cells have elevated c-myc transcript levels (thus presumably elevated protein levels). They also showed the greatest repression ratio in these studies. The DNase I footprinting results described in this thesis implicate a role for CTF in the repression of the c-myc gene. Coincidentally the purified CTF extract of Cohen, Sheffery and Kim, (1986) contained as its major component a 64 Kd protein, which is the same size as the c-myc gene product. In addition, studies on the hsp70 gene showed that its expression could be enhanced by co-transfection of the c-myc gene (Kingston, Baldwin and Sharp, 1985). By in vitro DNase I footprint analysis, the only protected domains within the region required for optimal expression of the hsp70 gene were a HSE (heat shock element) and a CTF site (Bienz, 1986). Highly speculatively (and with a good imagination), it is conceivable that the c-myc gene family could encode a CTF-like family of transcriptional regulators.

At nearly every stage of the investigation into the c-myc oncogene, cross-reference to the adenovirus Ela gene

is possible (see section 1.6.4.). The E1a gene product regulates its own transcription (Borrelli, Hen and Chambon, 1984): could the c-myc gene product do likewise? Perhaps these two genes will, when fully understood, be categorised together. Recent evidence suggests that the E1a gene product may function by binding to sequence-specific DNA-binding proteins. Perhaps the c-myc gene transcriptional repressor (possibly the c-myc gene product itself) might function in an analogous manner: binding to either the CTF or Sp1 (or both) factors which have been shown in vivo to bind to the NRE-2 domain (this thesis).

The experiments of Borrelli, Hen and Chambon, (1984), which identified the adenovirus E1a gene product as having the ability to repress certain enhancer elements, and similar studies which implied that the c-fos gene product could trans-activate certain promoter domains (Setoyma et al, 1986), could be mimicked for the c-myc.

4.4. Conclusion

The work described in this thesis identifies, and partially characterises, several transcriptional control domains within the 5' flanking region of the human c-*myc* proto-oncogene (see figure 73). Both negative and positive regulatory domains have been localised to regions previously found to be hypersensitve to DNase I *in vivo* (Siebenlist et al, 1984). By comparing the sequence of the colinear domains from the murine (Bernard et al, 1983: Fahrlander, Piechaczyk and Marcu, 1985) and feline c-*myc* genes (Stewart et al, 1986), regions of homology are present. Therefore it is probable that these domains represent conserved transcriptional control elements.

The activity of these domains was identified when either human- or mouse- derived cells were used as the recipients in transient transfection studies. In addition, activity was observed in cells of erythroid, fibroblastic or epithelial origin. Presumably this indicates a certain promiscuity with respect to tissue- and species- specificity for the transcriptional control domains identified. However the whole story of how these elements function is far from known. For example, the NRE-2 domain can function to promote transcription, while the repression function although instigated by the NRE-2 domain, can somehow involve the UPE domain as well. Future investigation into these domains will hopefully clarify the picture.

Both *in vivo* transfection competition and *in vitro* DNase I footprinting studies implicate roles for *trans*-acting factors in repression and activation (promotion) of the human c-*myc* gene. The *in vitro* DNase I footprint data implicated that either, or both CTF and Sp1 factors were involved with the repression function, while Sp1-like factors were involved in the functioning of the promoter domain. This is the first time either CTF or Sp1 have been associated with the negative regulation of a gene. The data

described in this thesis suggest that the transcriptional regulation of the c-myc gene requires the interaction of several factors. Perhaps the interaction of multiple closely located sequence-specific transcription factors will emerge as a common property of transcriptional control sequences in general.

It is conceivable that one or more of these domains may be involved in either, or both, the serum-induced activation or the differentiation-associated repression of the c-myc gene. Since several transcriptional control domains for the c-myc have been identified (this thesis, Remmers et al, 1986), experiments to test this involvement can now be devised.

The c-myc gene is intimately associated with the genesis of the neoplastic state. In transgenic mice which contain an exogenously introduced c-myc gene under the transcriptional control of an inducible promoter, the resultant constitutive transcriptional deregulation of the gene (in cells which are permissive for the promoter) appears to act as a heritable, predisposing factor favouring the development of tissue-specific neoplasms (Stewart, Pattengale and Leder, 1984). This predisposition may in part be due to a maturation arrest of the sensitive cells (Langdon et al, 1986). Thus the deregulation of the c-myc gene appears to be a central step in the progression from a normal cell to the full neoplastic state.

With this in mind, deregulation of the c-myc gene is associated with Burkitt's lymphoma and murine plasmacytoma cells, where reciprocal chromosome translocation joins the c-myc gene with one of the immunoglobulin loci (see section 1.5.3.c). Comparison of the location of, some of, the translocation breakpoints shows a clustering around -500 to -350 from the transcription start site (Yang et al, 1985; Dyson and Rabbits, 1985). These translocations would rearrange the NRE-2 domain described in this thesis, with

respect to the c-myc promoter-gene unit. It is therefore not unreasonable to suggest that, in at least some cases, removal or rearrangement of the NRE-2 domain would result in deregulation of expression of the c-myc gene, and that this deregulation is an important step in the genesis of the observed cancer.

Figure 70.

Comparison of Species Conservation of GC-Rich Elements.

This diagram shows the relative positions (by sequence analysis) of the GC-rich elements (open circles) and their orientation (arrow), in the various c-myc genes sequenced; M= murine, H= human, F= feline, C= chicken. P1 and P2 are the two major transcription initiation sites (heterogenous starts are found for chicken c-myc). (Bernard et al, 1983; Stewart et al, 1986; Nottenburg and Varmus, 1986).
See figures 65, 74 and 76.

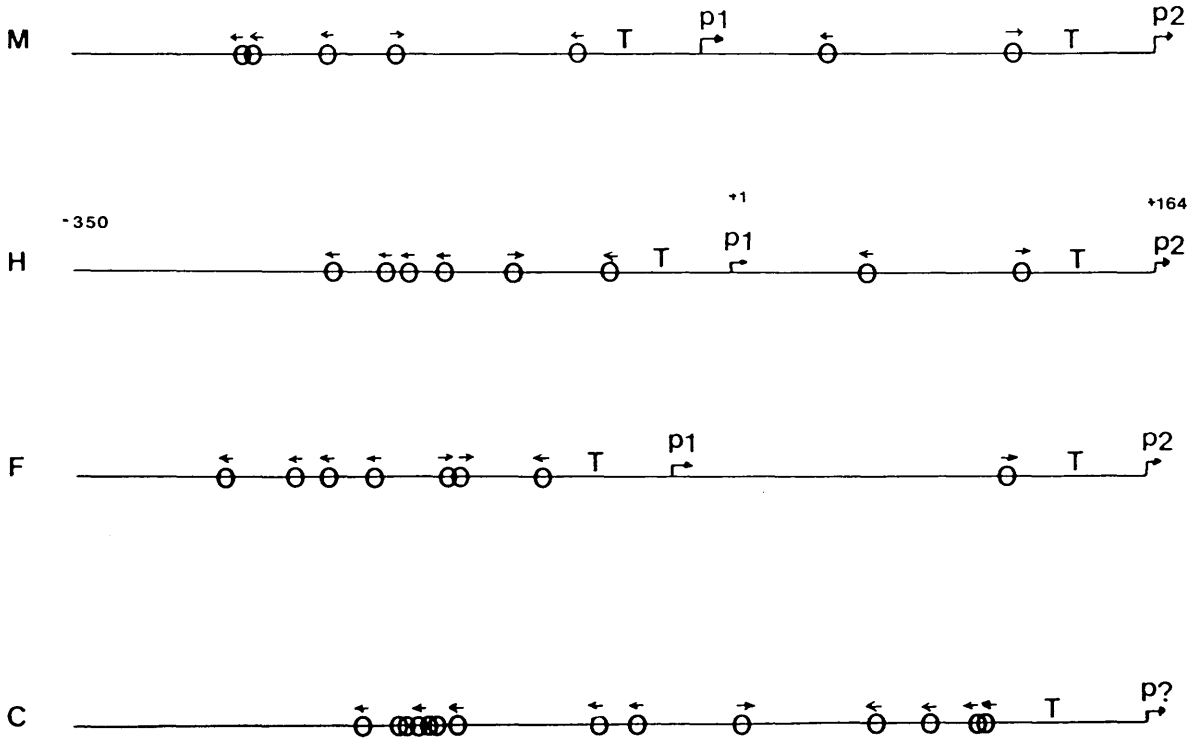


Figure 71.

Comparison of DNase I Protected (Footprint) Sequences.

A. h-c-myc Sp1-like sites (detected using Sp1 enriched protein extract). They are arranged 5' to 3' (i.e. the first is the most distant from P1 the 5' major transcription initiation site). The consensus sequence is shown, with small letters referring to the less frequently observed bases (Kadonaga, Jones and Tjian, 1986; R.Tjian, personel communication)

```
                T G G G A G G G G C
                C A G G A G G G G C
                G G G G A G G A G A
                G G G G A G G G T G
                C G G G A G G G G C
                A G G G C G T G G G
                T G G G C G G A G A

consensus       G G G G C G G G G C
                T       a   t A A T
                a               g
                g
```

Four of the sequences footprinted in these studies show complete homology with the consensus sequence, while three show a single base miss-match (underlined).

B. h-c-myc CTF sites (detected by using a CTF enriched protein extract).They are arranged 5' to 3', with both the possible recognition sites for the 5' footprint shown. The consensus sequence is shown (Benoist et al, 1980; R.Tjian personel communication).

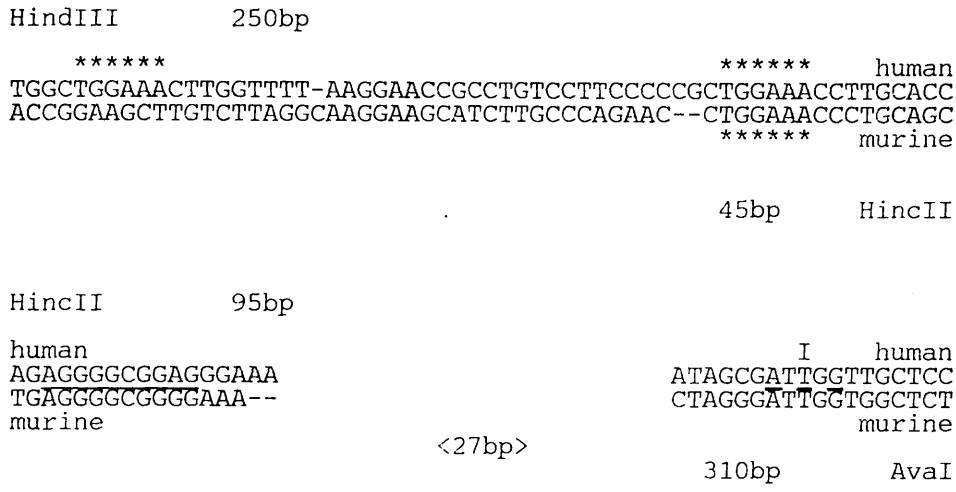
```
                c c C C A A T t c
                t G C C A A a g c
                a t C C A A T C c

consensus       G G C C A A T C A
                    T           T
```

The base miss-matches are shown as small letters. The core consensus is found in two of the possible recognition sites, with a one base miss-match (T to A) being present in the other sequence.

Figure 72.

A.Conservation of the Far Upstream Regulatory Domain.



The best alinement of the human and murine c-myc sequences for the colinear Far Upstream Doamin are shown above (Fahrlander, Piechacyzk and Marcu, 1985).The location of the in vivo Dnase I hypersensitive site is indicated by the Roman numeral I. Distances are given in base pairs (bp) from the indicated restriction sites to the relevant sequence, and between the putative Sp1 (continuous line) and CTF (dashed line) consensus elements (designated by sequence only). The putative enhancer core elements are indicated by asterisks.

B. Species Conservation of the TATA-Consensus Domain.

P1	human	T T T A T A A T G C
	murine	T T T A T A T T C C
	feline	T T T A T A T G C G
	chicken	G C T T T A A A G A
P2	human	A G T A T A A A A G
	murine	A G T A T A A A A G
	feline	T G T A T A A A A G

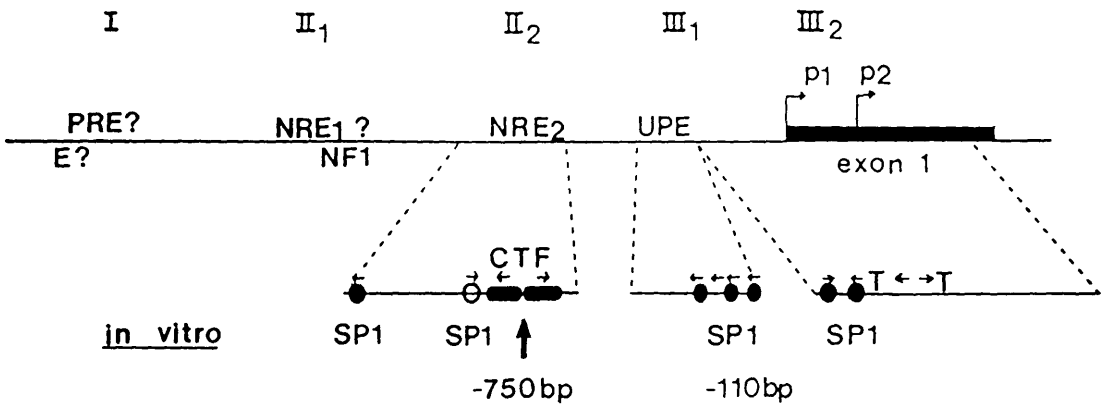
This diagram shows the high degree of conservation that exists between the sequenced c-myc genes within the TATA-consensus domain. P1 and P2 refer to the two major c-myc transcription initiation sites. The greatest species homology is found within the P2 domain. In addition, the chicken sequence, which contains only one putative TATA-conensus, shows better homology to the P2 domain than to the P1 domain. The above sequence data was taken from Bernard et al, 1983; Siebenlist et al, 1984; Nottenburg and Varmus, 1986; Stewart et al, 1986.

Figure 73.

Model for cis-Acting Transcriptional Control Domains.

The 5' flanking region and first exon of the human c-myc gene is shown. P1 and P2 refer to the two major transcription initiation sites. Roman numerals refer to the in vivo DNase I hypersensitive sties of Siebenlist et al, (1984). Horizontal arrows indicate Sp1 consensus elements and their orientation. Circles indicate Sp1-like protected domains (filled = strongly protected, open = weakly protected). The black bars represent CTF proteced domains, while the vertical arrow represents the location of a strong in vitro DNase I hypersensitive site. Distances are from P1.

Key: PRE = Positive Regulatory Element
E = putative Enhancer domain
NRE = Negative Regulatory Element
UPE = Upstream Promoter Element
NF1 = Nuclear Factor-1 consensus sequence
CTF = CCAAT-binding Transcription Factor
Sp1 = GC-rich sequence binding transcription factor



APPENDIX

5. Appendix

5.1. Evidence for Possible Cis-Recombination Between Co-Transfected Supercoiled Plasmids

Initial results (see table 7) showed that pB25 (MoLTR enhancer sequences in a pUC vector, not described) could compete out repression to a greater extent that either pB18 or pB20 (c-myc 5' flanking sequences) from the indicator plasmid pB12 (c-myc-CAT recombinant: see section 3.4.5.). This was initially taken to indicate that the trans-acting factor(s) involved in repression also recognised the enhancer element. However, as controls in latter experiments, pB25 was shown to elevate the expression level obtained from pB14 (c-myc-CAT recombinant which does not contain the negative element) by co-transfection, while pB18 had no affect on pB14. This may imply that the MoLTR enhancer element was functioning in cis to activiate the c-myc promoter. Thus, in the co-transfection experiment although the enhancer may be acting as a sink for the repressor factor(s), it could also be cis-activating (possibly by over-coming the NRE-2) the c-myc recombinant.

Further evidence for cis-activation came from experiments with the UPE domain. When pB57 (containg the 250bp UPE domain) was co-transfected with pB36 (c-myc-CAT recombinant containing only 97bp of 5' flanking sequences), an elevation of expression was observed at the lower molar ratios, with a gradual titration-effect being seen as the molar excess of competitor over indicator was increased. This may imply that the UPE domain (from pB57) was acting in cis to activate expression. However, an alternative explanation could be that the UPE domain was competing out a repressor factor which interacted with sequences within pB36.

Although these results tentatively imply that cis-recombination between co-transfected supercoiled plasmids may occur at a relatively high rate. However, the

sequences present in excess could concurrently titrate out any trans-acting factor which bound to that sequence (as seen in the competition experiments, see section 3.4.5.).

Table 7.

Competition Experiments with Enhancer Elements in LaTK⁻ Cells.

Donor DNA[a]	CAT Activity[b]	[c]	%[d]
pB9	385	±29	1.2
pB14	31810	±1111	100
pB14 Comp.pB25	155687	±8011	489
pB14 Comp.pB18	32189	±2568	101
pB12	3533	±479	11
pB12 Comp.pB18	13281	±922	42
pB12 Comp.pB20	10823	±562	34
pB12 Comp.pB25	30869	±434	87

a. Equimolar ratios of 1μg of indicator plasmid was transfected, and where relevant (prefix = Comp.), a 5 molar excess of competitor plasmid was cotransfected with the indicator.

b. Activity was expressed as the average of total counts (cpm) per flask for two separate experiments of duplicate flasks.

c. Range of variation of values.

d. Average counts (cpm) represented as a percentage of pB14.

e. Donor DNA was mixed with carrier salmon sperm DNA to a final concentration of 20μg/ml in the calcium-phosphate precipitate and transfected as described in Material and Methods.

f. CAT activity was assayed as described in Materials and Methods using total cell extract.

References

Adams, J.M., Harris, A.W., Pinkert, C.A., Corcoran, L.M., Alexander, W.S., Cory, S., Palmiter, R.D., and Brinster, R.L., (1985). The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid maligancy in transgenic mice. Nature, 318, 533-538.

Alitalo, K., Bishop, J.M., Smith, D.H., Chen, E.Y., Colby, W.W., and Levinson, A.D., (1983). Nucleotide sequence of the v-myc oncogene of avian retrovirus MC29. Proc. Natl. Acad. Sci., 80, 100-104.

Alitalo, K., (1985). Amplification of cellular oncogenes in cancer cells. TIBS, 194-197.

Alwine, J.C., (1985). Transient gene expression control: effects of transfected DNA stability and trans-activation of viral early genes. Mol. Cel. Biol., 5, 1034-1042.

Anagnou, N.P., Karlson, S., Moulton, A.D., Keller, G., and Nienhius, A.W., (1986). Promoter sequence required for the function of the human globin gene in erythroid cells. EMBO J., 5, 121-126.

Antonarakis, S.E., Orkin, S.H., Cheng, T-C., Scott, A.F.,Sexton, J.P., Trusko, S.P., Charache, S., and Kazazian, H.H., (1984). β-Thalassaenia in American blacks: novel mutations in the TATA box and an acceptor splice site. Proc. Natl. Acad. Sci. USA., 81, 1154-1158.

Armelin, H.A., Armelin, M.C.S., Kelly, K., Stewart, T., Leder, P., Cochran, B.H., and Stiles, C.D., (1984). Functional role for c-myc in mitogenic resonse to platelet-derived growth factor. Nature, 310, 655-660.

Azorin, F., and Rich, A., (1985). Isolation of Z-DNA binding proteins from SV40 minichromosomes : evidence for binding of the viral control region. Cell, 41, 365-374.

Baker, C.C., Herisse, J., Courtois, G., Galibert, F., and Ziff, E., (1979). Messenger RNA for the AD2 DNA binding proteins DNA sequences encoding the first leader and heterogeneity at the mRNA 5' end. Cell, 18, 569-580.

Balmain, A. (1985). Transforming ras oncogenes and multistage carcinogenesis. Br. J. Cancer, 51, 1-7.

Banergi, J., Rusconi, S., and Schaffner, W., (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell, 27, 299-308.

Bargmann, C.I., Hung, M.C., and Weinberg, R.A., (1986). Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. Cell, 45, 649-657.

Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G., and Leder, P., (1983). The human c-myc oncogene : structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell, 34, 779-789.

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

Benoist, C., and Chambon, P., (1981). In vivo sequence requirements of the SV40 early promoter region. Nature, 290, 304-310.

Bentley, D.L., and Groudine, M., (1986). A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL-60 cells. Nature, 321, 702-706.

Bergman, Y., Rice, D., Grosschedl, G., and Baltimore, D., (1984). Two regulatory elements for immunoglobin K light chain gene expression. Proc. Natl. Acad. Sci. USA., 81, 7041-7045.

Bernard, O., Cory, S., Gerondakis, S., Webb, E., and Adam, J.M., (1983). Sequence of the murine and cellular myc oncogene and two modes of myc transcription resulting from chromosome translocation in B lymphoid tumours. EMBO J., 2, 2375-2383.

Bienz, M., (1986). A CCAAT box confers cell-type-specific regulation on the Xenopus hsp70 gene in oocytes. Cell, 46, 1037-1042.

Bienz-Tadmor, B., Zakut-Houri, R., Libresco, S., Givol, D., and Oren, M., (1985). The 5' region of the p53 gene : evolutionary conservation and evidence for a negative element. EMBO J., 4, 3209-3213.

Bird, A. (1986). CpG-rich islands and the function of methylation. Nature, 321, 209-213.

Birnstiel, M.L., Busslinger, M., and Strub, K., (1985). Transcription termination and 3' processing : the end is in site! Cell, 40, 349-359.

Bishop, J. M., (1978). Retroviruses. Ann. Rev. Biochem., 47, 35-88.

Bishop, J.M., (1983). Cellular oncogenes and retroviruses. Ann. Rev. Biochem., 52, 301-354.

Bishop, M.J., and Varmus, H., (1985). Functions and origins of retroviral transforming genes. In RNA Toumour Viruses: Weiss, R., Teich, N., Varmus, H., Coffin, J., eds., Cold Spring Harbour Laboratory, 249-356.

Blanchard, J.M., Piechaczyk, M., Dani, C., Chambard, J-C., Franchi, A., Pouyssegur, J., and Jeanteur, P., (1985). c-myc gene is transcribed at high rate in GO-arrested fibroblasts and is post-transcriptionally regulated in response to growth factors. Nature, 317, 443-445.

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

Brady, J., Radonovich, M., Vodkin, M., Natarajan, V., Thoren, M., Das, G., Janik, J., and Salzman, N.P., (1982). Site-specific base substitution and deletion mutations that enhance or suppress transcription of the SV-40 major late RNA. Cell, 31, 625-633.

Brahms, J.G., Dargouge, O., Brahms. S., Ohara, Y., and Vagner, V., (1985). Activation and inhibition of transcription by supercoiling. J. Mol. Biol., 181, 455-465.

Brand, A.H., Breeden, L., Abraham, J., Sternglanz, R., and Naysmith, K., (1985). Characterisation of a "silencer" in yeast : a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell, 41, 41-48.

Bravo, R., Burckhardt, J., Curran, T., and Muller, R., (1985). Stimulation and inhibition of growth by EGF in different A431 cell clones is accompanied by the rapid induction of c-fos and c-myc proto-oncogenes. EMBO J., 4, 1193-1197.

Bravo, R., Burckhardt, J., and Muller, R., (1985). Persistence of the competent state in mouse fibroblasts is independent of c-fos and c-myc expression. Exp. Cell. Res., 160, 54--543.

Breathnach, R., and Chambon, P., (1981). Organisation and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem., 50, 349-383.

Brelvi, Z.S., and Studzinski, G.P., (1986). Changes in the expression of oncogenes encoding nuclear phosphoproteins but not c-Ha-ras have a relationship to monocytic differentiation of HL-60 cells. J. Cell. Biol., 102, 2234-2243.

Bunte, T., Greiser-Wilke, I., and Moelling, K., (1983). The transforming protein of the MC29-related virus CMII is a nuclear DNA-binding protein whereas MH2 codes for a cytoplasmic RNA-DNA binding polyprotein. EMBO J., 2, 1087-1092.

Bunte, T., Donner, P., Pfaff, E., Reis, B., Greiser-Wilke, I., Schaller, H., and Moelling, K., (1984). Inhibition of DNA binding of purified p55 v-myc in vitro by antibodies against bacterially expressed myc protein and a synthetic peptide. EMBO J., 3, 1919-1924.

Butnick, N.A., Miyamoto, C., Chizzanite, R., Cullen, B.R., Ju, G., and Skalk, A.M., (1985). Regulation of the human c-myc gene: 5' noncoding sequences do not affect translation. Mol. Cell. Biol., 5, 3009-3016.

Calabi, F., and Neuberger, M.S. (1985). Chromosome translocation activates heterogeneously initiated, bipolar transcription of a mouse c-myc gene. EMBO J., 4, 667-674.

Campisi, J., Gray, H.E., Pardee, A.B., Dean, M., and Sanenshein, G.E., (1984). Cell cycle control of c-myc but not c-ras expression is lost following chemical transformation. Cell, 36, 241-247.

Care, A., Cianetti, L., Giampaolo, A., Sposi, N.M., Zapparigna, V., Mavillo, F., Alimena, G., Amadari, S., Mandelli, F., and Peschle, C., (1986). Translocation of c-myc into the immunoglobulin heavy-chain locus in human acute B-cell leukaemia. A molecular analysis. EMBO J., 5, 905-911.

Carthew, R.W., Chodosh, L.A., and Sharp, P.A., (1985). An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell, 43, 439-448.

Cato, A.C.B., Geisse, S., Wenz, M., Westphal, H.M., and Beato, M., (1984). The nucleotide sequences recognized by the gluococorticoid receptor in the rabbit uteroglobin gene region are located far upstream from the initiation of transcription. EMBO J., 3, 2771-2778.

Chambard, J.C., and Pouyssegur, J., (1986). Intracellular pH controls growth factor-induced ribosomal protein S6 phosphorylation and protein synthesis in the G0-G1 transition in fibroblasts. Exp. Cell. Res., 164, 282-294.

Chambon, P., (1975). Eukaryotic nuclear RNA polymerase. Ann. Rev. Biochem., 44, 613-638.

Chiswell, D.J., Ramsey, G., and Hayman, M.J., (1981). Two virus-specific RNA species are present in cells transformed defective leukaemia virus OK10. J. Virol., 40, 301-304.

Clewell, D.B., and Helsinki, D.R. (1970). Properties of a supercoiled deoxyribonucleic acid-protein relaxation complex and strand specificity of the relaxation event. Biochem., 9, 4428-4440.

Cochran, M.D., and Weissman, C., (1984). Modular structure of the β-globin and TK promoters. EMBO J., 3, 2453-2459.

Cockerill, P.N., and Gerrard, W.T., (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell, 44, 273-282.

Cohen, S.N., Chang, A.C.Y., and Hsu, L., (1972). Non-chromosomal antibiotic resistance in bacteria: genetic transformation by R-factor DNA. Proc. natl. acad. Sci. USA., 69, 2110-2114.

Cohen, R.B., Sheffery, M., and Kim, C.G., (1986). Partial purification of a nuclear protein that binds to the CCAAT box of the mouse α1-Globin gene. Mol. Cell. Biol., 6, 821-832.

Coll, J., Gighi, M., Taisne, C.D., Dissous, C., Gegonne, A., and Stehelin, D., (1983). Molecular cloning of the avian acute transforming retrovirus MH2 reveals a novel cell-derived sequence (v-mil) in addition to the myc oncogene. EMBO J., 2, 2189-2194.

Connan, G., Rassoulzadegan, M., and Cuzin, F., (1985). Focus formation in rat fibroblasts exposed to a tumour promoter after transfer of polyoma plt and myc oncogenes. Nature, 314, 277-280.

Conscience, J.F., Verrier, B., and Martin, G., (1986). Interleukin-3-dependent expression of the c-myc and c-fos proto-oncogenes in hemopoietic cell lines. EMBO J., 5, 317-323.

Coppola, J.A., and Cole, M.D., (1986). Constitutive c-myc oncogene expression blocks mouse erythroleukaemia cell differentiation but not commitment. Nature, 320, 760-763.

Corcoran, L.M., Adams, J.M., Dunn, A.R., and Cory, S., (1984). Murine T lymphomas in which the cellular myc oncogene has been impicated by retroviral insertion. Cell, 37, 113-122.

Cororan, L.M., Cory, S., and Adams, J.M., (1985). Transposition of the immunoglobulin heavy chain enhancer to the myc oncogene in a murine plasmocytoma. Cell, 40, 71-79.

Cory, S., Gerondakis, S., and Adams, J.M., (1983). Interchromosomal recombination of the cellular oncogene c-myc with the immunoglobulin heavy chain locus in murine plasmocytomas in a reciprocal exchange. EMBO J., 2, 697-703.

Cory, S., Graham, M., Webb, E., Corcoran, L., and Adams, J.M., (1985). Varient (6;15) translocations in murine plasmocytomas involve a chromosome 15 locus at least 72 Kb from the c-myc oncogene. EMBO J., 4, 675-681.

Couglin, S.R., Lee, W.M.F., Williams, P.W., Giels, G.M., and Williams, L.T., (1985). c-myc gene expression is stimulated by agents that activate protein kinase C and does not account for the mitogenic effect of PDGF. Cell, 43, 243-251.

Croce, C.M., Erikson, J., Huebner, K., and Nishikura, K., (1985). Co-expression of translocated and normal c-myc oncogenes in hybrids between Daudi and lymphoblastoid cells. Science, 227, 1235-1237.

Culotta, V.C., Wides, R.J., and Sollner-Webb, B., (1985). Eukaryotic transcription compexes are specifically associated in large sedimentable structures: rapid isolation of polymerase I, II and III transcription factors. Mol. Cell. Biol., 5, 1582-1590.

Dalla-Favera, R., Bregui, M., Erikson, J., Patterson, D., Gallo, R.C., and Croce, C., (1982a). Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA., 79, 7824-7827.

Dalla-Favera, R.,Gellman, E.P., Martinotti, S., Franchini, G., Papas, T.S., Gallo, R., and Wong-Staal, F., (1982b). Cloning and characterisation of different human sequences related to the onc gene (v-myc) of avian myelocytomatosis virus (MC29). Proc. Natl. Acad. Sci. USA., 79, 6497-6501.

Dani, C., Blanchard, M., Piechazyk, M., Sabouty, S.E., Marty, L., and Jeanteur, P., (1984). Extreme instability of myc mRNA in normal and transformed human cells. Proc. Natl. Acad. Sci. USA., 81, 7046-7050.

Dani, C., Mechti, N., Piechazyk, M., Lebleu, B., Jeanteur, P., and Blanchard, J.M., (1985). Increased rate of degradation of c-myc mRNA in interferon-treated Daudi cells. Proc. Natl, Acad. Sci. USA., 82, 4896-4899.

Darnell, Jr., J.E., (1982). Variety in the level of gene control in eukaryotic cells. Nature, 297, 365-371.

Darveau, A., Pelletier, J., and Sonenberg, N., (1985). Differential efficiencies of in vitro translation of mouse c-myc transcripts differing in the 5' untranslated region. Proc. Natl. Acad. Sci. USA., 82, 2315-2319.

Das, G.C., and Piatigorsky, J., (1986). The chicken δ-crystallin gene promoter: binding of transcription factor(s) to the upstream G+C-RICH region is necessary for promoter function in vitro. Proc. Natl. Acad. Sci. USA., 83, 3131-3135.

Das, G.C., and Salzman, N.P., (1985). Simian virus 40 early promoter mutations that affect promoter function and autoregulation by large T antigen. J.Mol. Biol., 182, 229-239.

Davis, T., Rinaldi, A., Clark, L., and Adams, R.L.P., (1986). Methylation of chromatin in vitro. Biochem. Biophys. Acta, 866, 233-241.

Denny, C.T., Hollis, G.F., Magrath, I.T., and Kirsch, I.R., (1985). Burkitt lymphoma cell line carrying a varient translocation creates new DNA at the breakpoint and violates the heirarchy of immunoglobulin rearrangement. Mol. Cell. Biol., 5, 3199-3207.

Devine, J.M., (1986). Mechanism of activation of HuBlym-1 gene unresolved. Nature, 321, 437-439.

de Vries, E.,Driel, W.V., Tronp, M., Boom, J.V., and Vliet, P.C., (1985). Adenovirus DNA replication in vitro : site-directed mutagenesis of the nuclear factor 1 binding site of the AD2 origin. Nucl. Acid. Res., 13, 4935-4952.

Diamond, A., Cooper, G.M., Fitz, J., and Lane, M.A., (1983). Identification and molecular cloning of the human B-lym transforming gene activated in Burkitt lymphoma. Nature, 305, 112-116.

Dierks, P., Ooyen, A.V., Dobkin, C., Reiser, J., and Weissmann, C., (1983). Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T3 cells. Cell, 32, 695-706.

Dixon, R.A.F., and Schaffer, P.A., (1980). Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. J. Virol, 36, 189-203.

Dmitrovsky, E., Kuehl, W.M., Hollis, G.F., Kirsch, I.R., Bender, T.P., and Segal, S., (1986). Expression of a transfected human c-myc oncogene inhibits differentiation of a mouse erythroleukaemic cell line. Nature, 322, 748-750.

Donner, P., Greiser-Wilke, I., and Moelling, K. (1982).Nuclear localisation and DNA binding of the transforming gene product of the avian myelocytomatosis virus. Nature, 296, 262-266.

Dony, C., Kessel, M., and Gruss, P., (1985). Post-transcriptional control of myc and p53 expression during differentiation of the embryonal cell line F9. Nature, 317, 636-639.

Downward, J., Yarden, Y., Mayes, E., Scarce, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J., and Waterfeild, M.D., (1984). Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. Nature, 307, 521-527.

Dunnick, W., Baumgartner, J., Fradkin, L., Schultz, C., and Szureki, P, (1985). Methylation of plasmocytoma c-myc genes. Gene, 39, 287-292.

Dynan, W.S., (1985). A single protein that binds to enhancers, promoters and replication origins? TIG, 269-270.

Dynan, W.S., (1986). Promoters for housekeeping genes. TIG, 196-197.

Dynan, W.S., and Tjian, R., (1985). Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature, 316, 774-778.

Dynan, W.S., Sazer, S., Tjian, R., and Schimke, R.T., (1986). Transcription factor Sp1 recognises a DNA sequence in the mouse dihydrofolate reductase promoter. Nature, 319, 246-248.

Dyson, P.J., and Rabbits, T.H., (1985). Chromatin structure around the c-myc gene in Burkitt lymphomas with upstream and downstream translocation points. Proc. Natl. Acad. Sci. USA., 82, 1984-1988.

Dyson, P.J., Littlewood, T.D., Forester, A., and Rabbits, T.H., (1985). Chromatin structure of transcriptionally active and inactive human c-myc alleles. EMBO J., 4, 2885-2891.

Efstratiadis, A., Posakany, J.W., Maniatis, T., Laum, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightam, J.L., Blechl, A.E., Smithies, O., Borulle, F.E., Shoulders, C.C., and Proudfoot, N.J., (1980). The structure and evolution of the human β-globin gene family. Cell, 21, 653-668.

Einat, M., Resnitzky, D., and Kimchi, A., (1985). Close link between reduction of c-myc expression by interferon and G0/G1 arrest. Nature, 313, 597-600.

Elgin, S.C.R., (1984). Anatomy of hypersensitive sites. Nature, 309, 213-214.

Ellison, M.J., Kelleher, R.T., Wong, A.H-J., Habener, J.F., and Rich, A., (1985). Sequence-dependent energetics of the B-Z transition in supercoiled DNA containing nonalternating purine-pyrimidine sequences. Proc. Natl. Acad. Sci. USA., 82, 8320-8324.

Emerson, B.M., and Felsenfeld, G., (1984). Specific factor conferring nuclease hypersensitivity atthe 5' end of chicken adult β-globin gene. Proc. Natl. Acad. Sci. USA., 81, 95-99.

Endo, T., and Nadal-Ginard, B., (1986). Transcriptional and postranscriptional control of c-myc during myogenesis : its mRNA remains inducible in differentiated cells and does not suppress the differentiated phenotype. Mol. Cell. Biol., 6, 1412-1421.

Erisman, M.D., Rothberg, P.G., Diehl, R.E., Morse, C.C., Spandorfer, J.M., and Astrin, S.M., (1985). Deregulation of c-myc gene expression in human colon carcinoma is not accompanied by amplification or rearrangement of the gene. Mol. Cell. Biol., 5, 1969-1976.

Evan, G.I., and Hancock, D.C., (1985). Studies on the interaction of the human c-myc protein with cell nuclei: p62 c-myc as a member of a discrete subset of nuclear proteins. Cell, 43, 253-261.

Fahrlander, P.D., Piechaczyk, M., and Marcu, K.B., (1985). Chromatin structure of the murine c-myc locus : implications for the regulation of normal and chromosmally translocated genes. EMBO J., 4, 3195-3202.

Falcone, G., Tato, F., and Alema, S., (1985). Distinctive effects of the viral oncogenes myc, erb, fps and src on the differentiation of program of quail myogenic cells. Proc. Natl. Acad. Sci. USA., 82, 426-430.

Faletto, D.L., Arrow, A.S., and Macara, I.G., (1985). An early decrease in phosphatidylinositol turnover occurs on induction of friend cell differentiation and precedes the decrease in c-myc expression. Cell, 43, 315-325.

Feo, S., ar-Rushdi, A., Huebner, K., Finan, J., Nowell, P.C., Clarkson, B., and Croce, C.M., (1985). Suppression of the normal mouse c-myc oncogene in human lymphoma cells. Nature, 313, 493-495.

Flavell, A.J., Cowie, A., Legon, S., and Kamen, R., (1979). Multiple 5' terminal cap structures in late polyoma virus DNA. Cell, 16, 357-371.

Foster, J., Stafford, J., and Queen, C., (1985). An immunoglobulin promoter displays cell-type specificity independent of the enhancer. Nature, 315, 423-425.

Freshney, R.I., (1983). Culture of animal cells. A manuel of basic technique. Alan R. Liss Inc., New York, USA.

Fritton, H.P., Igo-Kemenes, T., Nowock, J., Strech-Jurk, U., Theisen, M., and Sippel, A.E., (1984). Alternative sets of DNAase-1-hypersenitive sites characterise the various functional states of chicken lysozyme gene. Nature, 311, 163-165.

Fung, Y.K., Crittenden, L.B., and Kung, H.J., (1982). Orientation and position of avian leukosis virus DNA relative to the cellular oncogene c-myc in B-lymphoma tumours of highly susceptible 151 x 7 chickens. J. Virol., 44, 742-746.

Fung, Y.K.T., Lewis, W.G., Crittenden, L.B., and Kung, H-J., (1983). Activation of the cellular oncogene c-erbB by LTR insertion : molecular basis for induction of erythroblastosis by avian leukosis virus. Cell, 33, 357-368.

Gaffney, D.F., McLauchlan, J., Whitton, J.L., and Clements, J.B., (1985). A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucl. Acid. Res., 13, 7847-7863.

Galas, D.J., and Schmitz, A., (1978). DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucl. Acid. Res., 5, 3157-3170.

Garcia, J.V., Bich-Thuy, L.T., Stafford, J., and Queen, C., (1986). Synergism between immunoglobulin enhancers and promoters. Nature, 322, 383-385.

Gaub, M.P., et al, (1985). pp123-131. Gluzman Y., ed., Cold Spring Harbour Laboratory, USA.

Gannon, F., O'Hare, K., Perrin, F., LePennee, J.P., Benoist, C., Cechet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B., and Chambon, P., (1979). Organisation and sequences at the 5' end of a cloned complete ovalbumin gene. Nature, 278, 428-434.

Gazin, C., Dimechi, S.D., Hampe, A., Masson, J-M., Martin, P., Stehelin, D., and Galibert, F., (1984). Nucleotide sequence of the human c-myc locus : provocative open reading frame within first exon.EMBO J., 3, 383-387.

Georgiev, O., and Birnstiel, M.L., (1985). The conserved CAAGAAAGA spacer sequence is an essential element for the formation of 3' termini of the sea urchin H3 histone mRNA by RNA processing. EMBO J., 4, 481-489.

Gilmour, R.S., Spandidos, D., Vass, K., Gow, J., and Paul, J., (1984). A negative regulatory sequence near the mouse β-major globin gene associated with a region of potential Z-DNA. EMBO J., 3, 1263-1272.

Gonda, T.J., and Metcalf, D., (1984). Expression of myb, myc and fos proto-oncogenes during the differentiation of a murine myeloid leukaemia. Nature, 310, 249-251.

Gorman, C.M., Moffat, L.F., and Howard, B.H., (1982). Recombinant genomes which express chloramphenicol acetytransferase in mammalian cells. Mol. Cell. Biol. 2, 1004-1051.

Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I., and Howard. B.H., (1982). The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc. Natl. Acad. Sci. USA., 79, 6777-6781.

Gorman, C.M., Rigby, P.W.J., and Lane, D.P., (1985). Negative regulation of viral enhancers in undifferentiated embryonic stem cells. Cell, 42, 519-526.

Gorski, K., Roch, J-M., Prentki, P., and Kirsch, H.M., (1985). The stability of bacteriophage T4 gene 32 mRNA a 5' leader sequence that can stabalise mRNA transcripts. Cell, 43, 461-469.

Goubin, G., Goldman, D.S., Luce, J., Neiman, P.E., and Cooper, G.M., (1983). Moleclular cloning and nucleotide sequence of a transforming gene detected by transfection of chicken B-cell lymphoma DNA. Nature, 302, 114-119.

Gough, N., (1985). Tissue-specific expression of immunoglobulin genes : a reappraisal. TIG, 267-268.

Goustin, A.S., Betsholtz, C., Pfeifer-Ohlsson, S., Persson, H., Rychert, J., Bywater, M., Holmgreen, G., Heldin, C.H., Westermark, B., and Ohlsson, R., (1985). Co-expression of the sis and myc proto-oncogenes in developing human placenta suggests autocrine control of trophoblast growth. Cell, 41, 301-312.

Graf, T., and Beug, H., (1978). Avian leukaemia viruses interaction with their target cells in vivo and in vitro. Biochem. Biophys. Acta., 516, 269-299.

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

Graham, M., Adams, J.M., and Cory, S., (1985). Murine T lumphomas with retroviral inserts in the chromosomal 15 locus for plasmocytoma varient translocations. Nature, 314, 740-743.

Graham, S., Tindle, R., and Birnie, G., (1985). Variation in myc gene amplification and expression in sublines of HL60 cells. Leuk. Res., 9, 239-247.

Grasser, S.M., and Laemmli, U.K., (1986). The organisation of chromatin loops: characterisation of a scaffold attachment site. EMBO J., 5, 511-518.

Greenberg, M.E., Hermanowski, A.L., and Ziff, E.B., (1986). Effect of protein synthesis inhibitors on growth factor activation of c-fos, c-myc and actin gene expression. Mol. Cell. Biol., 6, 1050-1057.

Grinnell, B.W., and Wagner, R.R., (1985). Inhibition of DNA-dependent transcription by the leader RNA of vesicular stomatitis virus : role of specific nucleotide sequences and cell protein binding. Mol. Cell. Biol., 5, 2502-2513.

Grosschedl, R., and Birnsteil, M.L., (1982). Delimitation of far upstream sequences required for maximal in vitro transcription of an H2A histone gene. Proc. Natl. Acad. Sci. USA., 79, 297-301.

Grosso, L.E., and Pitot, H.C., (1985). Transcriptional regulation of c-myc during chemically induced differentiation of HL60 cultures. Cancer Res., 45, 847-850.

Grosveld, G.C., de Boer, E., Shewmaker, L.K., and Flavell, R.A., (1982). DNA sequences necessary for transcription of the rabbit β-globin gene in vivo. Nature, 295, 120-126.

Hall, A., Marshall, C.J., Spurr, N.K., and Weiss, R.A., (1983). Identification of transforming gene in two human sarcoma cell lines as a new member of the ras gene family located on chromosome 1. Nature, 303, 396-400.

Hamlyn, P.H., and Rabbits, T.H., (1983). Translocation joins c-myc and immunoglobulin genes in a Burkitt lymphoma revealing a third exon in the c-myc oncogene. Nature, 304, 135-139.

Hann, S.R., and Eisenman, R.N., (1984). Proteins encoded by the human c-myc oncogene : differential expression in neoplastic cells. Mol. Cell. Biol., 4, 2486-2497.

Hann, S.R., Thompson, C.B., and Eisenman, R.N., (1985). c-myc oncogene protein synthesis is independent of the cell cycle in human and avian cells. Nature, 314, 366-369.

Hashimoto, C., and Steitz, J.A., (1986). A small nuclear ribonucleoprotein assiociates with the AAUAAA polyadenylation signal in vitro. Cell, 45, 581-591.

Hayday, A.C., Gillies, S., Saito, H., Wood, C., Wiman, K., Howard, W., and Tonegawa, S., (1984). Activation of a translocated human c-myc gene by an enhancer in the immunoglobulin heavy-chain locus. Nature, 307, 334-340.

Hayman, M.J., Kitchner, G., and Graf, J., (1979). Cells transformed by avian myelocytomatosis virus strain CM2 contain a gag-related 90K protein. Virology, 98, 191-199.

Hayward, W.S., Neel, B.G., and Astrin, S.M., (1981). Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature, 290, 475-480.

Hayward, W.S., (1983). The v-myc and c-myc genes : roles in oncogenesis. In Genes and Proteins in Oncogenesis : eds., 225-232.

Hearing, P., and Shenk, T., (1985). Sequence-independent autoregulation of the adenovirus type 5 E1A transcription unit. Mol. Cell. Biol. 5, 3214-3221.

Hearing, P., and Shenk, T., (1986). The adenovirus type 5 E1A enhancer contains two functionally distinct domains : one is specific for E1A and the other modulates all early units in cis. Cell, 45, 229-236.

Hen, R., Sassone-Corsi, P., Corden, J., Gaub, M.P., and Chambon, P., (1982). Sequences upstream from the TATA box are required *in vivo* and *in vitro* for efficient transcription from the adenovirus serotype 2 major late promoter. Proc. Natl. Acad. Sci. USA., 79, 7132-7136.

Hennighausen, L., and Fleckenstein, B., (1986). Nuclear factor 1 interacts with five DNA elements in the promoter region of the human cytomegalovirus major immediate early gene. EMBO J., 5, 1367-1371.

Heyneker, H.L., Shine, J., Goodman, H.M., and Boyer, H.W., (1976). Synthetic lac operator is functional *in vivo*. Nature, 263, 748- 752.

Hochschild, A., and Ptashne, M., (1986). Cooperative binding of 1 repressor to sites seperated by intergtal turns of the DNA helix. Cell, 44, 681-687.

Hogness, D.A., (1979). PhD Thesis, University of Stanford, USA.

Holt, J.T., Gopal, T.V., Moulton, A.D., and Nienhuis, A.W., (1986). Inducible production of c-*fos* antisense RNA inhibits 3T3 cell proliferation. Proc. Natl. Acad. Sci. USA, 83, 4794-4798.

Hughes, M., Compton, J.G., Schaeler, W.T., and O'Mally, B.W., (1981). Interaction of the chick oviduct progesterone receptor with deoxyribonecleic acid. Biochem., 20, 2481-2491.

Hunt, T., (1985). False starts in translational control of gene expression. Nature, 316, 580-581.

Imperiale, M.J., Kao, H-J., Feldman, L.T., Nevins, J.R., and Strickland, S., (1984). Common control of the heat shock gene and early adenovirus genes: evidence for a cellular E1a-like activity. Mol. Cell. Biol., 4, 867-874.

Ishii, S., Kadonaga, J.T., Tjian, R., Brady, J.N., Merlino, G.T., and Pastan, I., (1986). Binding of Sp1 transcription factor by the human Harvey *ras* proto-oncogene promoter. Science, 232, 1401-1413.

Jackson, D.A., (1986). Organisation beyond the gene. TIBS, 11, 249-252.

Jainchill, J.L., Aaronson, S.A., and Todaro, G.J., (1967). Murine sarcoma and leukaemia viruses: assay using clonal lines of contact-inhibited mouse cells. J.Virol.,4, 549-553.

Jocobs, J.P., Jones, C.M., and Baille, J.P., (1970). Characterisation of a human diploid cell designated MRC-5. Nature, 227, 168-170.

Jonak, G.J., and Knight, E., (1984). Selective reduction of c-*myc* mRNA in Daudi cells by human β-interferon. Proc. Natl. Acad. Sci. USA., 81, 1747-1750.

Jones, N.C., (1986). Negative regulation of enhancers. Nature, 321, 202-203.

Jost, J.P., Geiser, M., and Seldran, M., (1985). Specific modulation of the transcription of cloned avian vitellogenin II gene by estradiol-receptor complex *in vitro*. Proc. Natl. Acad. Sci. USA., 82, 988-991.

Kaczmarek, L., Oren, M., and Baserga, R., (1986). Co-operation between the p53 protein tumour antigen and platelet-poor plasma in the induction of cellular DNA synthesis. Exp. Cell. Res., 162, 268-272.

Kadonaga, J.T., Jones, K.A., and Tjian, R., (1986). Promoter-specific activation of RNA polymerase II transcription by Spl. TIBS, 11, 20-23.

Kakkis, E., Prehn, J., and Calame, K., (1986). An active chromatin structure acquired by translocated c-myc gene. Mol. Cell. Biol., 6, 1357-1361.

Kan, N.C., Flordellis, C.S., Garon, C.F., Duesberg, P.H., and Papas, T.S., (1983). Avian caricinoma virus MH2 contains a transformation specific sequence, mht, and shares the myc sequence with MC29, CMII and OK10 viruses. Proc. Natl. Acad. Sci. USA., 80, 6566-6570.

Kao, H.T., Capasso, O., Heintz, N., and Nevins, J.P., (1985). Cell-cycle control of the human hsp70 gene : implications for the role of a cellular Ela-like function. Mol. Cell. Biol., 5, 628-633.

Karin, M., Haslinger, A., Holtgreve, H., Richards, R.I., Krauter, P., Westphal, H.M., and Beato, M., (1984). Characterisation of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein-IIa gene. Nature, 308, 513-519.

Katzen, A.L., Kornberg, T.B., and Bishop, J.M. (1985). Isolation of the proto-oncogene c-myb from D. Melanogaster. Cell, 41, 449-456.

Katzir, N., Rechair, G., Cohen, J.B., Unger, T., Simoni, F., Segal, S., Cohen, D., and Givol, D., (1985). "Retroposon" insertion into the cellular oncogene c-myc in canine transmissible veneral tumour. Proc. Natl. Acad. Sci. USA., 82, 1054-1058.

Keath, E.J., Caimi, P.G., and Cole, M.D., (1984). Fibroblast lines expressing activated c-myc oncogenes are tumourigenic in nude mice and syngenic animals. Cell, 39, 339-348.

Keath, E.J., Kelekar, A., and Cole, M.D., (1984). Transcriptional activation of the translocated c-myc oncogene in mouse plasmocytomas : similar RNA levels in tumour and proliferating normal cells. Cell, 37, 521-528.

Kelekar, A., and Cole., M.D., (1986). Tumorigenicity of fibroblast lines expressing the adenovirus Ela, cellular p53, or normal c-myc genes. Mol. Cell. Biol., 6, 7-14.

Kelly, K., Cochran, R., Stiles, C.D., and Leder, P., (1983). Cell-specific regulation of the c-myc gene by lymphocyte mitogens and platelet-derived growth factor. Cell, 35, 603-610.

Kerem, B.S., Goitein, R., Diamond, G., Cedar, H., and Marcus, M., (1984). Mapping of DNAase I sensitive regions on mitotic chromosomes. Cell, 38, 493-499.

Keshet, I., Lienan-Hurwitz, J., and Cedar, H., (1986). DNA methylation affects the formation of active chromatin. Cell, 44, 535-543.

Khoury, G., and Gruss, P., (1983). Enhancer elements. Cell, 33, 313-314.

Killary, A.M., and Fournier, R.E.K., (1984). A genetic analysis of extinction: trans-dominant loci regulate expression of liver-specific traits in hepatoma hybrid cells. Cell, 38, 523-534.

Kingston, R.E., Baldwin, A.S., and Sharp, P.A., (1985). Transcription control of oncogenes. Cell, 41, 3-5.

Kingston, R.E., Kaufman, R.J., and Sharp, P.A., (1984). Regulation of transcription of the adenovirus EII promoter by ElA gene products : absence of sequence specificity. Mol. Cell. Biol., 4, 1970-1977.

Kit, S., Dubbs, D.R., Pierkarski, L.J., and Hsu, T.C., (1963). Deletion of thymidine kinase activity from L cells resistant to bromodeoxyuridine. Exp.Cell Res., 31, 297-312.

Klein, G., (1983). Specific chromosomal translocations and the genesis of B-cell derived tumours in mice and men. Cell, 32, 311-315.

Klempnauer, K.H., and Sippel, A.E., (1986). Subnuclear localisation of proteins encoded by the oncogene v-myb and its cellular homolog c-myb. Mol. Cell. Biol., 6, 62-69.

Knight, E., Anton, E.D., Fahey, D., Friedland, B.K., and Jonak, G.L., (1985). Interferon regulates c-myc gene expression in Daudi cells at the post-transcriptional level. Proc. Natl. Acad-Sci, USA, 82, 1151-1154.

Kohl, N., Legany, E., DePinho, R.A., Nisen, P.D., Smith. P.K., Gee, C.E., and Alt, F.W., (1986). Human N-myc is closely related in organisation and nucleotide sequence to c-myc. Nature, 319, 73-77.

Kovacs, B.J., and Butterworth, P.H.W., (1986). The effect of changing the distance between the TATA-box and cap site by up to three base pairs on the selection of the transcriptional start site of a cloned eukaryotic gene in vitro and in vivo. Nucl. Acid. Res., 14, 2429.

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

Labhart, P., and Reeder, R.H., (1984). Enhancer-like properties of the 60/81 bp elements in the ribosomal gene spacer of Xenopus laevis. Cell, 37, 285-289.

Lachman, H.M., and Skoulitchi, A.I., (1984). Expression of c-myc changes during differentiation of mouse erythroleukaemia cells. Nature, 310, 592-594.

Laimins, L.A., Gruss, P., Pozzatti, R., and Khoury, G., (1984). Characterisation of enhancer elements in the long terminal repeat of moloney murine sarcoma virus. J. Virol, 49, 183-189.

Laimins, L., Holmgren-Konig, M., and Khoury, G., (1986). Transcriptional "silencer" element in rat repetitive sequences associated with the rat insulin-1 gene locus. Proc. Natl. Acad. Sci. USA., 83, 3151-3155.

Land, H., Parada, L.F., and Wienberg, R.A., (1983). Tumorigenic conversion of primary embryo fibroblasts requires at least two co-operating oncogenes. Nature, 304, 396-602.

Lang, J.C., (1985). PhD Thesis, University of Glasgow, U.K.

Lang, J.C., Spandidos, D.A., and Wilkie, N.M., (1984). Transcriptional regulation of a herpes simplex virus immediate early gene is mediated through an enhancer-type sequence. EMBO J., 3, 384-395.

Langdon, W.Y., Harris, A.W., Cory, S., and Adams, J.M., (1986). The c-myc oncogene perturbs B lymphocyte development in Eμ-myc transgenic mice. Cell, 47, 11-18.

Langer, K.D., Vardiman, L., Renz, D., and Doefler, W., (1984). DNA methylation of three 5'-CCGG-3' sites in the promoter and 5' region inactivates the E2a gene of adenovirus type 2. Proc. Natl. Acad. Sci. USA, 81,2950-2954.

Larsen, A., and Weingraub, H., (1982). An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell, 29, 609-622.

Leder, A., Pattengale, P.K., Kuo, A., Stewart, T.A., and Leder, P., (1986). Consequence of widespread deregulation of the c-myc gene in transgenic mice: multiplr neoplasms and normal development. Mol. Cell. Biol., 45, 485-495.

Leder, P., Battey, J., Lenoir, G., Moulding, C., Murphy, W., Potter, H., Stewart, T, and Taub, R., (1983). Translocations among antibody genes in human cancer. Science, 220, 765-771.

Lee, W.M.F., Schwab, M., Westaway, D., and Varmus, H.E., (1985). Augmented expression of normal c-myc is sufficient for cotransformation of rat embryo cells with a mutant ras gene. Mol. Cell. Biol., 5, 3345-3356.

Levy, L.S., Gardner, M.B., and Casey, J.W., (1984). Isolation of a feline leukaemia provirus containing the oncogene myc from a feline lymphosarcoma. Nature, 308, 853-858.

Lewis, M.K., and Burgess, R.R., (1982). Eukaryotic RNA polymerase. In The Enzymes, 3rd Edit., P.D. Boyer, Ed., New York : Academic Press, 15B, 109-153.

Lilley, D.M.J., (1983). Structural perturbation in supercoiled DNA : hypersensitivity to modification by a single-strand-selective chemical reagent conferred by inverted repeat sequences. Nucl. Acid. Res., 11, 3097-3113.

Lonsdale, D.M, Brown, S.M., Subak-Sharp, J.H., Warren, K.G., and Koprowski, H., (1979). The polypeptide and the DNA restriction enzyme profiles of spontaneous isolates of herpes simplex virus type 1 from explants of hu,an teigeminal, superior cervical and vagus ganglia. J.Gen. Virol., 43, 151-171.

Loyter, A., Scangos, G.A., and Ruddle, F.H., (1982). Mechanisms of DNA uptake by mammalian cells: fate of exogenously added DNA monitored by the use of fluorescent dyes. Proc. Natl. acad. Sci. USA., 79, 422-426.

Lozzio, C.B., and Lozzio, B.B., (1975). Human chromic myelogenous leukaemia cell line with positive Philadelphia chromosome. Blood, 45, 321-334.

Luchnik, A.N., (1986). Long-distance signal transfer in transcriptionally active chromatin - How does it occur? Bio Essays, 3, 249-252.

Makino, R., Hayashi, K., and Sugimura, T., (1984). c-myc transcript is induced in rat liver at a very early stage of regeneration or by cycloheximide treatment. Nature, 310, 697-698.

Manley, J., Fire, A., Cano, A., Sharp, P.A., and Gefter, M.L., (1980). DNA-dependent transcription of adenovirus genes in a soluable whole-cell extract. Proc. Natl. Acad. Sci. USA, 77, 3855-3859.

Marcu, K., Melchers, F., Morse, H.C., and Potter, M., (1986). Mechanisms in B-cell neoplasia. Immun. Today, 7, 249-253.

Marmur,J., (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J.Mol.Biol., 3, 208-218.

Martin, P., Henry, C., Ferre F.,et al, (1986). Transformation of quail embryo fibroblasts by a retrovirus carrying a normal human c-myc gene. EMBO J., 5, 1529-1533.

Marshall, C., (1985). Human Oncogenes. In RNA Tumour Viruses : Weiss, R., Teich, N., Varmus, H., Coffin, J., eds., Cold Spring Harbour Laboratory, 487-558.

Mason, P.J., Jones, M.B., Elkington, J.A., and Williams, J.G., (1985). Polyadenylation of the Xenopus Bl globin mRNA at a downstream minor site in the absence of the major site and utilisation of an AAUACA polyadenylation signal. EMBO J., 4, 205-211.

Masters, J.N., and Attardi, G., (1985). Discrete human dihyrofolate reductase gene transcripts present in polysomal RNA map with their 5' ends several hundred nucleotides upstream of the main mRNA start site. Mol. Cell. Biol., 5, 493-500.

Mathis, D.J., and Chambon, P., (1981). The SV40 early region TATA-box is required for accurate in vitro initiation of transcription. Nature, 290, 310-315.

Mattox, W.W., and Davidson, N., (1984). Isolation and characterisation of the beadex locus of Drosophila melanogaster : a putative cis-acting negative regulatory element for the heldup-a gene. Mol. Cell. Biol., 4, 1343-1353.

Maxam, A., and Gilbert, W., (1977). A new method for DNA sequencing. Proc. Natl. Acad. Sci. USA., 74, 560-564.

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

Michitsch, R.W., and Melera, P.W., (1985). Nucleotide sequence of the 3' exon of the human N-myc gene. Nucl. Acid. Res., 13, 2545.

Miksicek, R., Heber, A., Schmid, W., Danesch, U., Posseckert, G., Beato, M., and Schutz, G., (1986). Glucorcoticoid responsiveness of the transcriptional enhancer of Moloney murine sarcoma virus. Cell, 46, 283-290.

Miyamoto, N.G., Moncollin, V., Wintzerith, M., Hen, R., Egly, J.M., and Chambon, P., (1984). Stimulation of in vitro transcription by the upstream element of the adenovirus-2 major late promoter involves a specific factor. Nucl. Acid. Res., 12, 8779-8799.

Moelling, K., Pfaff, E., Beug, H., Beimling, P., Bunte, T., Schaller, H.E., and Graf, T., R., (1985). DNA-binding activity is associated with purified myb proteins from AMV and E26 viruses and is temperature-sensitive for E26 ts mutants. Cell, 40, 983-990.

Molders, H., Defesche, J., Muller, D., Bonner, Rapp. U.R., and Muller, R., (1985). Integration of transfected LTR sequences into the c-raf proto-oncogene : activation by promoter insertion. EMBO J., 4, 693-698.

Moore, D.D., Marks, A.L., Buckley, D.I., Kapler, G., Payvar, F., and Goodman, H.M., (1985). The first intron of the human growth hormone gene contains a binding site for glucocorticoid receptor. Proc. Natl. Acad. Sci. USA., 82, 699-702.

Moroy, T., Marchio, A., Etiemble, J., Trepo, C., Tiollais, P., and Buendia, M-A., (1986). Rearrangement an denhanced expression of c-myc in hepatocellular carcinoma of hepatitis virus infected woodchuck. Nature, 324, 276-279.

Mougneau, E., Lemieux, L., Rassoulzadegan, M., and Cuzin, F., (1984). Biological activities of v-myc and rearranged c-myc oncogenes in rat fibroblast cells in culture. Proc. Natl. Acad. Sci. USA, 81, 5758-5762.

Muller, R., Bravo, R., Buckhardt, J., and Curran, T., (1984). Induction of c-fos gene and protein by growth factors precedes activation of c-myc. Nature, 312, 714-720.

Murphy, W., Sariel, J., Taub, R., Vasicek, T., Battey, J., Lenoir, G., and Leder, P., (1986). A translocated human c-myc oncogene is altered in a conserved coding sequence. Proc. Natl. Acad. Sci. USA., 83, 2939-2943.

Mushinski, J.F., Bauer, S.R., Potter, M., and Reddy, E.P., (1983). Increased expression of myc-related oncogene mRNA characterises most BALB/c plasmacytomas induced by pristane or Abelson murine leukemia virus. Proc. natl. Acad. Sci. USA, 80, 1073-1077.

McGarry, T.J., and Lindquist, S., (1985). The preferential translation of Drosophila hsp70 mRNA requires sequences in the untranslated leader. Cell, 42, 903-911.

McGeady, M.L., Wood, T.G., Maizel, J.V., and Vande Wood, G.F., (1986). Sequences upstream from the mouse c-mos oncogene may function as a transcription termination signal. DNA, 5, 289-298.

McGuinnis, W., Shermoen, A.W., Hamskerk, J., and Beckendorf, S.K., (1983). DNA sequence change in an upstream DNAase-1 hypersensitive region are correlated with reduced gene expression. Proc. Natl. Acad. Sci. USA., 80, 1063-1067.

McKnight, S.L., and Kingsbury, R., (1982). Transcriptional control signals of a eukaryotic protein-coding gene. Science, 217, 316-324.

McKnight, S., and Tjian, R., (1986). Transcriptional selectivity of viral genes in mammalian cells. Cell, 46, 795-805.

McLachlan, A.D., and Boswell, D.R., (1985). Confidence limits for homology in protein or gene sequences. J. Mol. Biol., 185, 39-49.

McLauchlan, J., Gaffney, D., Whitton, J.L., Clements, J.B., (1985). The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. Nucl. Acid. Res., 13, 1347-1365.

Nau, M.M., Brooks, B.J., Battey, J., Sausville, E., Garzadar, A.F., Kirsch, I.R., McBride, O.W., Bertness, V., Holts, G.F., and Minna, J.D. (1985). L-myc, a new myc-related gene amplified and expressed in human small cell lung cancer. Nature, 318, 69-72.

Neel, B.G., Jhanwar, S.C., Chaganti, R.S.K., and Hayward, W.S., (1982). Two human c-onc genes are located on the long arm of chromosome 8. Proc. Natl. Acad. Sci. USA., 79, 7842-7846.

Neidle, S., (1983). New twists to left-handed DNA. Nature, 302, 574.

Neil, J.C. Hughes, D., McFarlane, R., Wilkie, N.M., Onions, D.E., Lees, G., and Jarrett, O., (1984). Transduction and rearrangement of the myc gene by feline leukaemia virus in naturally occuring T-cell leukaemia. Nature, 308, 814-820.

Nepveu, A., Fahrlander, P.D., Yang, J-Q., and Marcu, K.B., (1985). Amplification and altered expression of the c-myc oncogene in A-MuLV-transformed fibroblasts. Nature, 317, 440-443.

Nevins, J.R., (1983). The pathway of eukaryotic mRNA formation. Ann. Rev. Biochem., 52, 441-466.

Newbold, R.F., and Overell, R.W., (1983). Fibroblast immortality is a prerequisite for transformation by EJ c-Ha-ras oncogene. Nature, 304, 648-651.

Nishikura, K, ar-Rushdi, A., Erikson, J., DeJesus, E., Dugan, D., and Croce C., (1984). Repression of rearranged μ gene and translocated c-myc in mouse 3T3 cells X Burkitt lymphoma cell hybrids. Science, 224, 399-402.

Nishikura, K., Erikson, J. ar-Rushdi, A., Huebner, K., and Croce, C.M., (1985). The translocated c-myc oncogene of Raji Burkitt lymphoma cells is not expressed in human lymphoblastoid cells. Proc. Natl. Acad. Sci. USA., 82, 2900-2904.

Nordheim, A., and Rich, A., (1983). Negative supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature, 303, 674-679.

North, G., (1984). Multiple levels of gene control in eukaryotic cells. Nature, 312, 308-309.

Nottenburg, C., and Varmus, H.E., (1986). Features of the chiken c-myc gene that influences the stucture of c-myc RNA in normal cells and bursal lymphomas. Mol. Cell. Biol., 6, 2800-2806.

Nusse, R., van Ooyen, A., Cox, D., Fung, Y.K.T., and Varmus, H., (1984). Mode of proviral activation of a putative mammary oncogene (int-1) in mouse chromosome 15. Nature, 307, 131-136.

Nussinov, R., Shapiro, B., Lipkin, L.E., and Maizel, J.V., (1984). Enhancer elements share local homologous twist-angle variations with a helical periodicity. Biochem. Biophys. Acta, 783, 246-257.

Ofverstedt, et al, (1984). Rapid and quantitative recovery of DNA fragments from gels by displacement electrophoresis. Biochim. Biophys. Acta, 782, 120-126.

Osborne, T.F., Gaynor, R.B., and Berk, A.J., (1982). The TATA homology and the mRNA 5' untranslated sequence are not required for expression of essential adenovirus E1A function. Cell, 29, 139-148.

Osborne, T.F., Goldstein, J.L., and Brown, M.S., (1985).5' end of HMG CoA reductase gene contains sequences responsible for cholesterol-mediated inhibition of expression. Cell, 42, 302-212.

Ohlsson, H., and Edlund, T., (1986).
Sequence-specific interactions of nuclear factors with the insulin gene enhancer.
Cell, 45, 35-44.

Oshimura, M., Gilmer, T.M., Barrett, C., (1985). Nonrandom loss of chromosome 15 in syrian hamster tumour induced by v-Ha-ras plus v-myc oncogenes. Nature, 316, 636-639.

Pabo, C.O., and Sauer, R.T., (1984). Protein-DNA recognition. Ann. Rev. Biochem., 53, 293-321.

Pachl, C., Biegalke, B., and Linial, M., (1982). RNA and protein encoded by MH2 virus: evidence for subgenomic expression of v-myc. J. Virol., 45, 133-139.

Palmieri, S., Kahn, P., and Graf, T., (1983). Quail embryo fibroblasts transformed by four v-myc-containing virus isolates show enhanced proliferation but are non tumorigenic. EMBO J., 2, 2385-2389.

Palmiter, R.D., and Brinster, K.L., (1986). Germline transformation of mice. Ann. Rev. Gen.

Papas, T.S., and Lautenberger, J.A., (1985). Sequence curiosity in v-myc oncogene. Nature, 318, 237.

Peck, L.J., and Wang, J.C., (1985). Transcriptional block caused by a negative supercoiling induced structural change in an alternating CG sequence. Cell, 40, 129-137.

Pelicci, P.G., Knowles, D.M., Magarth, I., and Dalla-Favera, R., (1986). Chromosomal breakpoints and structural alterations of the c-myc locus differ in endemic and sporadic forms of Burkitt lymphoma. Proc. Natl. Acad. Sci. USA., 83, 2984-2988.

Pelletier, J., and Sonenberg, N., (1985). Insertion mutagenesis to increase secondary structure within the 5' noncoding regin of a eukaryotic mRNA reduces translational efficiency. Cell, 40, 515-526.

Perry, R.P., (1983). Consequences of myc invasion of immunoglobulin loci; facts and speculation. Cell,33,647-649.

Persson, H., and Leder, P., (1984). Nuclear localisation and DNA-binding properties of a protein expressed by human c-myc oncogene. Science, 225, 718-721.

Persson, H., Heminghausen, H., Taub, R., BeGrado, R., and Leder, P., (1984). Antibodies to human c-myc oncogene product : evidence of an evolutionary conserved protein induced during cell proliferation. Science, 225, 687-693.

Persson, H., Gray, H.E., and Godeau, F., (1985). Growth-dependent synthesis of c-myc-encoded proteins: early stimulation by serum factors in synchronised mouse 3T3 cells. Mol. Cell. Biol., 5, 2903-2913.

Persson, H., Gray, H.E., Godeau, F., Braunhat, S., and Bellve, A.R., (1986). Multiple growth-associated nuclear proteins immunoprecipitated by antisera raised against human c-myc peptide antigens. Mol. Cell. Biol., 6, 942-949.

Peruchio, M., Hanahan, D., and Wigler, M., (1980). Genetic and physical linkage of exogenous sequences in transformed cells. Cell, 22, 309-317.

Peschle, C., Mavilio, F., Sposi, N.M., Gianpaolo, A., Care, A., Bottero, L. Bruno, M., Masteradino, G., Gastaldi, R., Testa, M.G., Alimena, G., Amadari, S., and Mandelli, F., (1984). Translocation and rearrangement of c-myc into immunoglobulin and heavy chain locus in primary cells from acute lymphocytic leukaemia. Proc. Natl. Acad. Sci. USA., 81, 5514-5518.

Peter, G., Lee, A.E., Dickson, E., (1984). Activation of cellular gene by mouse mammary tumour virus may occur early in mammary tumour development. Nature, 309, 273-274.

Peterson, D.O., (1985). Alterations in chromatin structure associated with gluococorticoid-induced expression of endogenous mouse mammary tumour virus genes. Mol. Cell. Biol., 5, 1104-1110.

Piechaczyk, M., Yang, J.Q., Blanchard, J.M., Jeanteur, P., and Marcu, K.B., (1985). Post-transcriptional mechanisms are responsible for accumulation of truncated c-myc RNAs in murine plasma cell tumours. Cell, 42, 589-597.

Pragnell, I.B.., Spandidos, D.A., and Wilkie, N.M., (1985). Consequence of altered oncogene expression in rodent cells. Proc. R. Soc. Lond. B., 226, 107-119.

Prehn, J., Mercola, M., Calame, K., (1984). Translocation affects normal c-myc promoter usage and activates fifteen cryptic c-myc transcription starts in plasmocytoma M603. Nucl. Acid. Res., 12, 8987-9007.

Pribnow, D., (1975). Nucleotide sequence of a RNA polymerase binding site at an early T7 promoter. Proc. Natl. Acad. Sci. USA., 72, 784-788.

Prochownik, E.V., and Kukowska, J., (1986). Deregulated expression of c-myc by murine erythroleukaemia cells prevents differentiation. Nature, 322, 848-850.

Pruijn, G.J.M., van Driel, W., and van der Vliet, P.C., (1986). Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. Nature, 322, 656-659.

Ptashne, M., (1986). Gene regulation by proteins acting nearby and at a distance. Nature, 322, 697-701.

Rabbits, P.H., Watson, J.V., Lamand, A., Forester, A., Stinson, M.A., Evan, G., Fischer, W., Atherton, E., Sheppard, R., and Rabbits, T.H., (1985). Metabolism of c-myc gene products :c-myc mRNA and protein expression in the cell cycle. EMBO J., 4, 2009-2015.

Ralston, R., and Bishop, J.M., (1983). The protein products of the myc and myb oncogenes and adenovirus Ela are structurally related. Nature, 306, 803-805.

Ramsey, G., Evan, G.I, and Bishop, J.M., (1984).The protein encoded by the human proto-oncogene c-myc. Proc. Natl. Acad. Sci. USA, 81, 7742-7746.

Rapp, U.R., Cleveland, J.L., Brightman, K., Scott, A., and Ihle, J.N., (1985). Abrogation of IL-3 and IL-2 dependence by recombinant murine retroviruses expressing v-myc oncogenes. Nature, 317, 434-438.

Reddy, E.P., Reynolds, R.K., Watson, D.K., Schultz, R.A., Lautenberger, J., and Papas, T.S., (1983). Nucleotide sequence analysis of the proviral genome of avian myelocytomatosis virus (MC29). Proc. Natl. Acad. Sci. USA., 80, 2500-2504.

Reed, J.C., Sabath, D.E., Hoover, R.G., and Prystowsky, M.B., (1985). Recombinant interleukin-2 regulates levels of c-myc mRNA in a cloned murine T-lymphocyte. Mol. Cell. Biol. 5, 3361-3368.

Reed, R., and Maniatis, T., (1985). Intron sequences involved in lariat formation during pre-mRNA splicing. Cell, 41, 95-105.

Reeves, R., (1984). Transcriptionally active chromatin. Biochem. Biophys. Acta., 782, 343-393.

Reitsma, P.H., Rothberg, P.G., Astrin, S.M., Trial, J., Bar-Shavit, Z., Hall, A., Teitelbaum, S.L., and Kahn, A.J., (1983). Regulation of myc gene expression in HL-60 leukaemia cells by a vitamin D metabolite. Nature, 306, 492-494.

Remmers, E.F., Yang, J.Q., and Marcu, K.B., (1986). A negative transcriptional control element located upstream of the murine c-myc gene. EMBO J., 5, 899-904.

Renan, M.J., (1985). Putative repressor binding sites in the regions mediating transcriptional control of viral and cellular genes. Biosci. Rep., 5, 739-753.

Reudelhuber, J., (1984). A step closer to the principles of eukaryotic transcription control. Nature, 311, 301.

Reynolds, G.A., Basen, S.K., Osborne, T.F., Chin, D.J., Gil, G., Brown, M.S., Goldstein, J.L., and Lusky, K.L., (1984). HMG CoA reductase : a negatively regulated gene with unusual promoter and 5' untranslated regions. Cell, 38, 275-285.

Rhodes, D., and Klug, A., (1986). An underlying repeat in some transcriptional control sequences corresponding to half a double helical turn of DNA. Cell, 46, 123-132.

Rich, A., (1982). Right-handed and left-handed DNA : conformatinal information in genetic material. Cold Spring Harbour Symp. Quant-Biol., 47, 1-12.

Rio,D.C., and Tjian,R.,(1983). SV40 T antigen binding site mutations that affect autoregulation. Cell, 32, 1227-1240.

Rodriguez-Pena, A., and Rozengurt, E., (1985). Serum, like phorbol esters, rapidly activates protein kinase C in intact quiescent fibroblasts. EMBO J., 4, 71-76.

Roussel, M., Saule, S., Lagrau, C., Rommens, C., Beug, H., Graf, T., and Stehelin, D., (1979). Three new types of viral oncogene of cellular origin specific for heamotopoetic cell transformation. Nature, 281, 452-455.

Sacca, R., Stanley, E.R., Sherr, C.J., and Rettenmier, C.W., (1986). Specific binding of the mononuclear phagocyte colony-stimulating factor CSF-1 to the product of the v-fms oncogene. Proc. Natl. Acad. Sci. USA., 83, 3331-3335.

Safer, B., Yang, L., Tolunay, E., and Anderson, W.F., (1985). Isolation of stable preinitiation, initiation and elongation compexes from RNA polymerase II - directed transcription. Proc. Natl. Acad. Sci. USA., 82, 2632-2636.

Sager, R., (1986). Genetic supression of tumour formation : a new frontier in cancer research. Cancer Res.,46,1573-1580.

Saito, H., Hayday, A.C., Wiman, K., Hayward, W.S., and Tonegawa, S., (1983). Activation of the c-myc gene by translocation : a model for translational control. Proc. Natl. Acad. Sci. USA., 80, 7476-7480.

Sassone-Corsi, P., and Borrelli, E., (1986). Transcriptional regulation by trans-acting factors. TIG, 215-219.

Sassone-Corsi, P., Wilderman, A., and Chambon, P., (1985). A trans-acting factor is responsible for the simian virus 40 enhancer activity in vitro. Nature, 313, 458-463.

Sawadogo, M., and Roeder, R.G., (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA-box region. Cell, 43, 165-175.

Scherer, W.F., Syverton, J.T., and Gey, G.O., (1953). Studies on the propagation in vitro of poliomyelitis viruses. J. Exp. Med., 97, 695-710.

Schibler, U., Hagenbuchle, O., Wallauer, P.K., and Pittet, A.C., (1983). Two promoters of different strengths control the transcriptin of the mouse alpha-amylase gene Amy-1a in the parotid gland and the liver. Cell, 33, 501-508.

Scholer, H.R., and Gruss, P., (1985). Cell type-specific transcriptional enhancement in vitro requires the presence of trans-acting factors. EMBO J., 4, 3005-3013.

Schubach, W., and Groudine, M., (1984). Alteration of c-myc chromatin structure by avian leukosis virus integration. Nature, 307, 702-708.

Schwab, M., Varmus, H., Bishop, J.M., Grzeschik, K-H., Naylor, S.L., Sakaguchi, A.Y., Brodeuk, G., and Trent, J., (1984). Chromosome localisation in normal human cells and neuroblastoma of a gene related to c-myc. Nature, 308, 288-290.

Schwab, M., (1985). Amplification of N-myc in human neuroblastomas. TIG, 272-275.

Serfling, E., Jasin, M., and Schaffner, W., (1985). Enhancers and eukaryotic gene transcription. TIG, 224-230.

Setoyama, C., Franzio, R., Liau, G., Mudryj, M., and de Crambrugghe, B., (1986). Transcription activator encoded by the v-fos gene. Proc. Natl. Acad. Sci. USA., 83, 3213-3217.

Shatkin, A.J., (1985). mRNA cap binding proteins : essential factors for initiating translation. Cell, 40. 223-224.

Shaul, Y., Ben-levy, R., and De-Medina, T., (1986). High affinity binding site for nuclear factor 1 next to the hepatitis B virus S-gene promoter. EMBO J., 5, 1967-1971.

Shaw, G., and Kanen, R., (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA regradation. Cell, 46, 659-667.

Shih, C.K., Linial, M., Goodenoe, M.M., and Hayward, W.S., (1984). Nucleotide sequence 5' of the chicken c-myc coding region : localisation of a noncoding exon that is absent from myc transcripts in most avian leukosis virus-induced lymphomas. Proc. Natl. Acad. Sci. USA., 81, 4697-4701.

Shilo, B.Z., and Weinberg, R.A., (1981). DNA sequences homologous to vertegrate oncogenes are conserved in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA., 78, 6789-6792.

Shpaer, E.G., (1985). The secondary structure of mRNAs from Escherichia coli : its possible role in increasing the accuracy of translation. Nucl. Acid. Res., 13, 275-288.

Siebenlist, U., Henninghausen, L.,Battey, J., and Leder, P., (1984). Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. Cell, 37, 381-391.

Sikora, K., Evan, G., Stewart, J., and Watsom, J.V., (1985). Detection of the c-myc oncogene product in testicular cancer. Br. J. Cancer., 52, 171-176.

Slamon, D.J., deKerinon, J.B., Verma, I.M., and Cline, M.J., (1984). Expression of cellular oncogenes in human malignancies. Science, 224, 256-262.

Smeland, E., Godal, T., Ruud, E., Beiska, K., Funclerud, S., Clark, E.A., Pfeifer-Ohlsson, S., and Ohlsson, R., (1985). The specific induction of myc protooncogene expression in normal human B cells is not a sufficient event for acquisition of competence to proliferate. Proc. Natl. Acad. Sci. USA., 82, 6255-6259.

Spandidos, D., and Wilkie, N.M., (1984a). Expression of exogenous DNA in mammalian cells. In Transcription and Translation : Hames, B.D., and Higgins, S.J., eds., IRL Press, 1-48.

Spandidos, D., and Wilkie, N.M., (1984b). Malignant transformation of early passage rodent cells by a single mutated oncogene. Nature, 310, 469-475.

Stabel, S., Argos, P., and Philipson, L., (1985). The release of growth arrest by microinjection of adenovirus E1A DNA. EMBO J., 4, 2329-2336.

Stanbridge, E., (1985). A case for human tumour-suppressor genes. Bio Essays, 3, 252-255.

Stein, A., and Kunzler, P., (1983). Histone H5 can correctly align randomly arranged nucleosomes in a define in vitro system. Nature, 302, 548-551..

Stewart, M.A., Forrest, D., McFarlane, R., Onions, D., Wilkie, N., and Neil, J.C., (1986). Conservation of the c-myc coding sequence in transduced feline v-myc genes. Virology, 153, 121-134.

Stewart, T.A., Pattengale, P.K., and Leder, P., (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell, 38, 627-637.

Stiles, C.P., Capone, G.T., Scher, C.D., Antoniades, H.N., Van Wyke, J.J., and Pledger, W.J., (1979). Dual control of cell growth by somatomedins and PDGF. Proc. Natl. Acad. Sci USA., 76, 1279-1283.

Suck, D., and Oefner, C., (1986). Structure of DNase I at 2.0A resolution suggests a mechanism for binding to and cutting DNA. Nature, 321, 620-625.

Sun, L.K., Showe, L.C., and Croce, C.M., (1986). Analysis of the 3' flanking region of the human c-myc gene in lymphomas with the t(8;22) and t(2;8) chromosome translocation. Nucl. Acid. Res., 14, 4037-4050.

Sures, I., Levy, S., and Kedes, L.H., (1980). Leader sequences of Strongylocentrotus purpuratus histone mRNAs start at a unique heptanucleotide common to all five histone genes. Proc. Natl. Acad. Sci. USA., 77, 1265-1269.

Sussman, D.J., and Milman, G., (1984). Short-term, high-efficiency expression of transfected DNA. Mol. Cell. Biol., 4, 1641-1643.

Swift, R.A., Shaller, E., Witter, R.L., and Kung, H-J., (1985). Insertional activation of c-myc by reticuloendotheliosis virus in chicken B lymphoma : nonrandom distribution and orientation of the provirus. J. Virol., 54, 869-872.

Symonds, G., Stubblefeild, E., Guyaux, M., and Bishop, J.M., (1984). Cellular oncogenes (c-erb-A and c-erb-B) located on different chicken chromosomes can be transduced into the same retroviral genome. Mol. Cell. Biol., 4, 1627-1630.

Symonds, G., Klempnauer, K-H., Snuder, M., Mozovici, G., Mozovici, C., and Bishop, J.M., (1986). Co-ordinate regulation of myelomonocytic phenotype by v-myb and v-myc. Mol. Cell. Biol. 6, 1796-1802.

Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M., and Chambon, P., (1986). Requirement of stereospecific alignments for initiation from the simian virus early promoter. Nature, 319, 121-126.

Taub, R., Kirsch, I., Morton, C, Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P., (1982). Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmocytoma cells. Proc. Natl. Acad. Sci. USA., 79, 7837-7841.

Taub, R., Moulding, C., Battey, J., Murphy, W., Vasicek, T., Lenoir, G.M., and Leder, P., (1984). Activation and somatic mutation of the translocated c-myc in Burkitt lymphoma cells. Cell, 36, 339-348.

Temin, H., (1982). Function of the retrovirus long terminal repeat. Cell, 28, 3-5.

Thompson, C.B., Challoner, P.B., Neiman, P.E., and Groudine, M., (1985). Levels of c-myc oncogene mRNA are invarient throughout the cell cycle. Nature, 314, 363-369.

Treisman, R., (1985). Transient accumulation of c-fos RNA following serum stimulation requires a conserved 5' element and c-fos 3' sequences. Cell, 42, 889-902.

Treisman, R., (1986). Identification of a promoter-binding site that mediated transcriptional reponse of the c-fos gene to serum factors. Cell, 46, 567-574.

Treisman, R., and Maniatis, T., (1985). Simian virus 40 enhancer increases number of RNA enhancer increases number of RNA polymerase II molecules on linked DNA. Nature, 315, 72-75.

Tsuda, M., and Suzuki, Y., (1981). Faithful transcription initiation of fibroin gene in an homologous cell-free system reveals an enhancing effect of 5' flanking sequence far upstream. Cell, 27, 175-182.

Twigg, A., and Sherrat, D., (1980). Trans-complementable copy-number mutants of plasmid Col E1. Nature, 283, 216-218.

Tyndall, C., Mantia, G.L., Thacker, C.M., Favaloro, J., and Kamen, R., (1981). A region of the polyoma virus genome between the replication origin and late protein coding sequences is required in cis for both early gene expression and viral DNA replication. Nucl. Acid. Res., 9, 6231-6250.

Vigneron, M., Mantia, G.L., Thacker, C.M., Favalaro, J., and Kamen, R., (1984). Effect of the 21 bp repeat upstream element on in vitro transcription from the early and late SV40 promoters. EMBO J., 3, 2373-2383.

von de Ahe, D., Janich, S., Scheiclereit, C., Renkawitz, R., Schutz, G., and Beato, M., (1985). Glucocorticoid and progesterone receptors bind to the same sites in two hormonally regulated promoters. Nature, 313, 706-709.

Waterfield, M.D., Scarce, T., Whittle, N., Stroobant< p., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C.H., Huang, J.S., and Deud, J., (1983). Platelet-derived growth factor is structurally related to the putative transforming protein p28sis of simian sarcoma virus. Nature, 304, 35-39.

Watson, D.K., McWilliams-Smith, M.J., Nunn, M.F., Duesberg, P.H., O'Brien, S.J., and Papas, T.S., (1985). The ets sequence from the transforming gene of avian erythroblastosis virus, E26, has unique domains on human chromosome 11 and 21 : both loci are transcriptionally active. Proc. Natl. Acad. Sci, USA., 82, 7294-7298.

Watt, R., Nishikura, K., Sorrentino, J., ar-Rushdi, A., Croce, C.M., and Rovera, G., (1983). The structure and nucleotide sequence of the 5' end of the human c-myc oncogene. Proc. Natl. Acad. Sci. USA., 80, 6307-6311.

Watt, R.A., Shatzman, A.R., and Rosenberg, M., (1985). Expression and characterisation of the human c-myc DNA-binding protein. Mol. Cell. Biol. 5, 448-456.

Weber, F., and Schaffner, W., (1985). Simian virus 40 enhancer increases RNA polymerase density within the linked gene. Nature, 315, 75-78.

Weil, P., Luse, D.S., Segall, J., and Roeder, R.G., (1979). Selective and accurate initiation of transcription at the AD2 major late promoter in a soluable system dependent on purified RNA polymerase II and DNA. Cell, 18, 469-484.

Weintraub, H., Beug, H., Groudine, M., and Graf, T, (1982). Temperature-sensitive changes in the structure of globin chromatin in lines of red cell precursors transformed by ts-AEV. Cell, 28, 931-940.

Weintraub, H., and Groudine, M., (1986). Chromosomal subunits in active genes have an altered conformation. Science, 193, 848-858.

Weimtraub, H., Izant, J.G., and Harland, R.M., (1985). Anti-sense RNA as atool for genetic analysis. TIG, 22-25.

Weintraub, H., (1985). Assembly and propagation of repressed and derepressed chromosomal states. Cell, 42, 705-711.

Weintraub, H, Chang, P.F., and Conrad, K., (1986). Expression of transfected DNA depends on DNA topology. Cell, 46, 115-122.

Weisbrod, S., and Weintraub, H., (1979). Isolation of a subclass of nuclear proteins responsible for conferring a DNAase-1-sensitive structure on globin chromatin. Proc. Natl. Acad. Sci. USA., 76, 630-634.

Whitelaw, C.B.A., Wilkie, N.M., Jones, J.A., Kadonaga, J.T., Tjian, R., and Lang, J.C., (1986). Transcriptionally active domains in the 5'flanking sequence of human c-myc. In Press.

Wigler, M., Sweet, R., Sim, G.F., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R., (1979). Transformation of mammalian cells with genes from prokaryotes and eukaryotes. Cell, 16, 777-785.

Wiman, K.G., Clarkson, B., Hayday, A.C., Saito, H., Tonegawa, S., and Hayward, W.S., (1984). Activation of a translocated c-myc gene : role of structural alterations in the upstream region. Proc. Natl. Acad. Aci. USA., 81, 6798-6802.

Winquist, R., Saksela, K., and Alitalo, K., (1984). The myc proteins are not associated with chromatin in mitotic cells. EMBO J., 3, 2947-2950.

Wong-Staal, F., Josephs, S., Dalla Favera, R., Westin, E., Gelmann., Franchini, G., and Gallo, R.C., (1983). Cellular onc genes: their role as progenitors of viral onc genes and their expression in human cells. In Haematology and Blood Transfusion : Neth, Gallo, Greaves, Moore and Winkler, 28, 178-185.

Wood, T.G., Mcgeady, M.L., Baroudy, B.H., Blair, D.G., and Vande Woude, G.F., (1984). Mouse c-mos oncogene activation is prevented by upstream sequences. Proc. Natl. Acad. Sci. USA., 81, 7817-7821.

Wu, B.J. and Morimoto, R.I., (1985). Transcription of the human hsp70 gene is induced by serum stimulation. Proc. Natl. Acad. Sci. USA., 82, 6070-7074.

Wu, C., (1984). Activating protein factor binds in vitro to upstream control sequences in heat shock gene chormatin. Nature, 311, 81-84.

Wu, C., (1985). An exonuclease protection assay reveals heat-shock element and TATA-box DNA-binding proteins in crude nuclear extracts. Nature, 317, 84-87.

Yaffe, D., Nudel, U., Mayer, Y., and Neuman, S., (1985). Highly conserved sequences in the 3' untranslated region of mRNAs coding for homologous proteins in distantly related species. Nucl. Acid. Res., 13, 3723-3737.

Yang, J.Q., Bauer, S.R., Mushinski, J.F., and Marcu, K,B., (1985). Chromosome translocations clustered 5' of the murine c-myc gene qualitatively affect promoter usage : implications for the site of normal c-myc regulation. EMBO J., 4, 1441-1447.

Yaniv, M., (1984). Regulation of eukaryotic gene expression by trans-activating proteins and cis acting DNA elements. Biol. Cell., 50, 203-216.

Yokota, J., Yokota, Y.T., Buttifora, H., Fevre, C.L., and Cline, M.J., (1986). Alterations of myc, myb and Ha-ras proto-oncogenes in cancers are frequent and show clinical correlation. Science, 231, 261-265.

Zehnbauer, B.A., and Vogelstein, B., (1985). Supercoiled loops and the organisation of replication and transcription in eukaryotes. Bio. Essays, 2, 52-54.

Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wilderman, A., and Chambon, P., (1986). Multiple sequence motifs are involved in SV40 enhancer function. EMBO J., 5, 387-397.

Zhu, J.D., Allan, M., and Paul, J., (1984). The chromatin structure of the human ε-globin gene : nuclease hypersensitive sites correlate with multiple initiation sites of transcription. Nucl. Acid. Res., 12, 9191-9204.

Zhou, R.P., Kan, N., Papas, T., and Duesberg, P., (1985). Mutagenesis of avian carcinoma virus MH2: only one of two potential transforming genes (gag-myc) transforms fibroblasts. Proc. Natl. Acad. Sci. USA., 82, 6389-6393.

Zimmerman et al. (1986).
Differential expression of myc family genes during murine development. Nature, 319, 780.