MOLECULAR ANALYSIS OF HUMAN AND MOUSE INTERFERON α GENE STRUCTURE AND FUNCTION

By C. Bartholomew

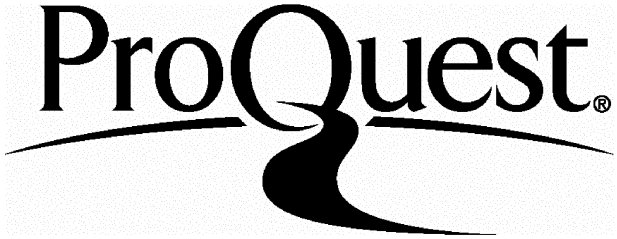
A thesis submitted for the degree of Doctor of Philosophy at the University of Leicester.

December 1986

Department of Biochemistry,

University of Leicester,

UK.



ACKNOWLEDGEMENTS

Thanks are due to many people, in particular my supervisor Dr.J.D. Windass for his advice, encouragement, patience and constructive criticism throughout. John's meticulous planning of experiments, dedication and enthusiasm is an example to us all. I would also like to thank Prof.W.J. Brammar for his continued support, guidance and many hours of discussion.

I am indepted to all the members of the I.C.I Corporate Bioscience Group. Specifically I would like to thank those members of labs p21 and p22 for their friendship, encouragement, advice and assistance. I am also very grateful to Dr.M. Edge of I.C.I Pharmaceuticals Division, Alderley Park, who supplied synthetic oligonucleotides for this work and Dr.A. Morris (University of Warwick) for performing IFN assays. In addition I would like to thank Prof.C. Weissmann, Drs.B. Howard, J. Lang, A. Jeffreys and T. Kunkel for their generous gifts of bacterial strains, plasmids and genomic DNAs.

My Mother, Father and family have always expressed interest and given their unrelented support. To them I owe a great deal. Last, but by no means least, I would like to thank Sian for her patience, encouragement and dedication to many hours spent typing and reading the manuscript.

DEDICATION

This thesis is dedicated to my brother Peter Robert Bartholomew

"Sweet is the Memory of Those That Loved Him"

MOLECULAR ANALYSIS OF HUMAN AND MOUSE INTERFERON α GENE STRUCTURE AND FUNCTION

C. Bartholomew

SUMMARY:-

Four human IFNα chromosomal genes have been isolated from a newly constructed placental DNA library in λL47. Restriction and sequencing analysis revealed that each gene had been described previously. However, one gene, $SMT111.1_A$, which encodes a full length IFN, is an allelic varient of a previously characterised pseudogene, thus indicating some degeneracy of the IFNα gene family.

A chimaeric gene comprising the MuIFNα$_1$ promoter (−188 to +52) and *cat* gene coding sequences has been constructed *in vitro*, enabling promoter function to be examined in mouse cells. Reproducible polyrI.rC mediated induction of CAT expression from the MuIFNα$_1$ promoter has been demonstrated in pools of stably transfected, but not transiently transfected, L929 cells. Monitoring mRNA production revealed the transient accumulation of correctly initiated hybrid gene transcripts which precede optimum CAT production.

Aspects of the structure/function relationship of the MuIFNα$_1$ promoter have been investigated by oligonucleotide site directed mutagenesis. Comparative studies of IFNα promoter sequences identified prospective regulatory regions for mutagenesis. Quantitative CAT assays have been employed for promoter assessment. Additionally, the construction of a pseudogene comprising the wildtype MuIFNα$_1$ promoter linked to an internally deleted *cat* gene has enabled both mutant and wildtype promoters functioning simultaneously in the same cell population to be assessed by S−1 nuclease protection studies using a common probe.

Such studies have revealed three distinct *cis*−acting regions implicated in MuIFNα$_1$ promoter function. Two are located upstream of the TATA box, defined by mutations at −87 and between −66 and −33 respectively. These reduce promoter activity 2 to 5 fold. The third, is defined by a mutation at +14, within the untranslated leader sequence. This enhances activity 2 to 3 fold.

A deletion derivative of the MuIFNα$_1$ promoter containing only 94bp of upstream sequence is inactive. Cis−activation by the Mo−MuSV enhancer restores inducibility to this promoter whereas the intact MuIFNα$_1$ promoter is refractory to this element. This suggests that distinct *cis*−acting sequences dictate the efficiency and regulation of MuIFNα$_1$ gene transcription.

ABBREVIATIONS

Ap - ampicillin

ATP - adenosinetriphosphate

bp - base pairs

BSA - bovine serum albumin

CAT - chloramphenicol acetyltransferase

CCC - covalently closed circular

conc. - concentration

cpm - counts per minute

dATP - deoxyadenosinetriphosphate

dCTP - deoxycytidinetriphosphate

DDW - double distilled water

DEAE - diethylaminoethyl

dGTP - deoxyguanosinetriphosphate

DMEM - Dulbecco's minimum essential medium

DMSO - dimethylsulphoxide

DNA - deoxyribonucleic acid

DNA pol I - DNA polymerase I

DNase - deoxyribonuclease

DTT - dithiothreitol

dTTP - deoxythymidinetriphosphate

ds - double stranded

EDTA - ethylenediaminetetra-acetic acid

F.C.S - foetal calf serum

F.G.M - full growth medium

F.T.L - freeze thaw lysate

G.R.E - glucocorticoid regulatory element

Hepes - N-2-Hydroxyethylpiperazine-N'-2-ethanesulphonic acid

H.R.E - heat regulatory element

hrs - hours

HSV - herpes simplex virus

IFN - interferon

I.R.E - interferon gene regulatory element

ITF - interferon gene transcription factor

kb - kilobases

L.M.P. - low melting point

mins - minutes

M.O.I - multiplicity of infection

Mo-MuSV - Moloney murine sarcoma virus

M.R.E - metal regulatory element

mRNA - messenger ribonucleic acid

mw - molecular weight

nt - nucleotides

OD - optical density

o/n - overnight

P.A.G.E - polyacrylamide gel electrophoresis

P.E.G - polyethyleneglycol

Pipes - piperazine-N-N-bis[2-ethane Sulphonic acid]

polyrI.rC - polyinosinic acid.polycytidylic acid

R.F. - replicative form

RNA - ribonucleic acid

RNase - ribonuclease

rpm - revolutions per minute

r.t - room temperature

sat. - saturated

S.D.S - sodium dodecyl sulphate

S.E - sonicated extract

secs - seconds

soln. - solution

ss - single stranded

SV40 - simian virus 40

Tc - tetracycline

TEMED - N, N, N', N' tetramethylethylenediamine

T.L.C. - thin layer chromatography

Tris - tris (Hydroxymethyl) aminomethane

V - volts

(v/v) - (volume/volume)

(w/v) - (weight/volume)

X-gal - 5-bromo-4-chloro-inolyl β galactosidase

CONTENTS

CHAPTER 1

INTRODUCTION

The interferons (IFN) are classically defined as antiviral proteins which are secreted by the host for protection against viral infection. They have been identified in all vertebrates examined (Wilson *et al*, 1983) but are most studied in man and mouse.

The genes encoding IFN activity are normally silent in most tissues, but can be activated by virus, bacteria or double stranded polyribonucleotides (Stewart *et al*, 1979). *In vitro* studies have identified major sources of IFN from leukocytes, lymphocytes and continuous fibroblast and lymphoblastoid cell lines (Pestka, 1981).

Polyclonal antisera have enabled the classification of IFNs into three subtypes: α, β and γ in both human and murine systems (Stewart *et al*, 1980a). IFNα, IFNβ, and IFNγ are also distinguishable physically, biologically and molecularly, according to their acid stability, characteristic profiles of target cell specificity and nucleotide sequence.

A wide range of cellular activities are associated with IFNs, including anti-viral, anti-proliferative and immunoregulatory roles. These tend to be mediated in a species specific manner (Stewart *et al*, 1979). Their anti-viral and anti-proliferative behaviour has stimulated substantial interest in their development as therapeutic agents.

In addition to their potential commercial significance the IFNs provide an important biological system for investigation. Understanding mechanisms of IFN gene regulation, species specificity and the pleiotropic response of target cells, are the subjects of intense research. Investigation of a family of virally induced proteins, which in turn regulate expression of diverse cellular genes, generating characteristic biochemical and phenotypic changes, provides an attractive model system for gene regulation studies.

1.1 Genetic Organisation and Physical Structure of the IFNs

IFNα

The isolation and characterisation of human (Nagata *et al*, 1980a; Lawn *et al*, 1981a) and mouse (Shaw *et al*, 1983) IFNα chromosomal genes demonstrated that they

comprised a family of genes (Brack *et al*, 1981; Henco *et al*, 1985; Zwarthoff *et al*, 1985; Kelley and Pitha, 1985a) exhibiting 80−90% and 60−70% nucleotide homology at the intra− and inter−species level respectively. Similarly bovine (Capon *et al*, 1985) and rat (Dijkema *et al*, 1984) IFNα genes have been isolated. These also comprise a gene family. The IFNαs are encoded by uninterupted genes (Nagata *et al*, 1980a; Lawn *et al*, 1981a; Shaw *et al*, 1983) which mostly generate a 189 amino acid precursor polypeptide, which is modified and exported from the cell by cleavage of a 23 amino acid signal peptide, producing the mature 166 amino acid protein (Mantei *et al*, 1980; Shaw *et al*, 1983). Exceptions include the 165 amino acid protein encoded by the LeIFNA or HuIFNα$_2$ genes (Goeddel *et al*, 1980a; Streuli *et al*, 1980) and the 167 and 151 amino acid polypeptide products of the MuIFNα$_2$ (Shaw *et al*, 1983) and MuIFNα$_4$ genes respectively (Zwarthoff *et al*, 1985; Kelley and Pitha, 1985a).

Currently twenty three loci for the human (Henco *et al*, 1985) and seven members of the mouse (Zwarthoff *et al*, 1985; Kelley and Pitha, 1985a) IFNα gene families have been characterised. The chromosomal location of these genes has been determined by hybridisation analysis of panels of somatic cell hybrids. This indicates the IFNα genes are clustered on human chromosome 9 (Owerbach *et al*, 1981) and mouse chromosome 4 (Kelley *et al*, 1983; Lovett *et al* 1984)

Many of these IFNα genes are expressed *in vitro*. Gene expression and cDNA cloning studies indicate a minimum of eight are transcribed in Sendai virus induced human leukocytes (Hiscott *et al*, 1984) and KG−1 cells (Goeddel *et al*, 1981). Similarly a minimum of five mouse genes are expressed in NDV virus induced L−929 cells (Zwarthoff *et al*, 1985; Kelley and Pitha, 1985b).

Purification and partial amino acid sequencing of IFNs has revealed the major anti−viral translation products of virus induced human leukocytes (Rubinstein *et al*, 1979) and lymphoblastoid Namalwa cells (Zoon *et al*, 1979; 1980) to be IFNα, comprising a minimum of eight distinct species (Levy *et al*, 1981; Allen, 1982). At least three species are translated in virus induced mouse C243 cells (De Maeyer−Guignard *et al*, 1978), N−terminal amino acid sequencing revealing a mixture of both α and β species (Taira *et al*, 1980).

Human IFNα, in contrast to mouse (Fujisawa, 1978), is not glycosylated (Allen and Fantes, 1980; Allen, 1982). However, a novel 27kd glycosylated species identified in Sendai virus induced Namalwa cells (Allen, 1982), may be the 172 amino acid polypeptide translation products of a closely related family of IFNα genes, designated $IFN\alpha_{11}$ (Capon *et al*, 1985; Hauptman *et al*, 1985). Nevertheless non−glycosylated recombinant mouse IFNα or human $IFN\alpha_{11}$, produced in *E.coli*, maintain anti−viral activity (Shaw *et al*, 1983; Hauptman *et al*, 1985), suggesting this post−transcriptional modification is unneccessary for biological activity.

The predicted amino acid sequence for IFNα contains 4 (human) or 5 (mouse) cysteine residues. Tryptic digests of purified recombinant IFNαA indicated the formation of disulphide bridges between the cysteine residues found at position 1 and 98, and also 29 and 138 (Wetzel, 1981). Intra− and inter−species conservation of these residues suggests that disulphide bridge formation may be a common structural feature of IFNα's.

IFNβ

Unique structural genes encoding human (Degrave *et al*, 1981; Gross *et al*, 1981; Houghton *et al*, 1981; Lawn *et al*, 1981b; Mory *et al*, 1981; Ohno and Taniguchi, 1981; Tavernier *et al*, 1981) and mouse (Fujita *et al*, 1985) IFNβ have been isolated and characterised. Inspection of human IFNα and IFNβ gene sequences demonstrate 45% and 29% homology at the nucleotide and amino acid level respectively (Taniguchi *et al*, 1980a). This strongly suggests they are derived from a common ancestral gene (Taniguchi *et al*, 1980b). Interspecies homology of IFNβ shows 63% nucleotide sequence conservation (Higashi *et al*, 1983).

The uninterrupted IFNβ genes encode 187 amino acid (human) and 181 amino acid (mouse) precursor polypeptides which are modified and exported from the cell as 166 and 161 amino acid mature proteins respectively, upon cleavage of a 21 amino acid signal peptide (Derynck *et al*, 1980; Higashi *et al*, 1983). The same procedures adopted to identify the chromosomal location of IFNα genes have allocated human IFNβ to chromosome 9 (Owerbach *et al*, 1981). The murine IFNβ gene is probably located on chromosome 4 (Lovett *et al*, 1984).

Although a multi-gene family of IFNα genes have been observed in man and mouse, only single species encoding IFNβ activity have been identified. In contrast, multiple IFNβ genes have been identified in cows and rabbits (Wilson *et al*, 1983; Capon *et al*, 1985). The presence of a single IFNβ gene in man and mouse has been contested by the identification of novel partial cDNAs which encode prospective IFNs, including HuIFNβ_2 (Weissenbach *et al*, 1980) and mouse pMIF3/10 and pMIF20/11 (Skup *et al*, 1982). These show no homology with previously characterised IFN genes. Moreover pMIF3/10 has recently been identified as the haemopoietic growth factor, erythrocyte potentiating factor (Gasson *et al*, 1985; Docherty *et al*, 1985). The origin of the IFNβ activity associated with pMIF3/10 and pMIF20/11 remains unexplained. The anti-viral activity associated with HuIFNβ_2 has also been disputed (Content *et al*, 1982).

The isolation of cDNAs encoding IFNβ demonstrated that these genes were transcribed in virus induced human and mouse fibroblasts (Taniguchi *et al*, 1980a; Derynck *et al*, 1980; Goeddel *et al*, 1980b; Higashi *et al*, 1983). Purification of IFNβ and partial N-terminal amino acid sequencing show it is the major anti-viral translation product of polyrI.rC induced human fibroblasts (Knight, 1976; Knight *et al*, 1980). In contrast polyrI.rC induced mouse fibroblasts generate a mixture of IFNα and IFNβ (DeMaeyer-Guignard *et al*, 1978; Taira *et al*, 1980). The apparent differences of expression of human and mouse IFNβ in fibroblasts may indicate differences in the cells examined, or a fundamental difference in the regulation of these genes.

Both human and mouse IFNβs are glycoproteins (Fujisawa *et al*, 1978; Tan *et al*, 1979). Nevertheless, expression of the mature human IFNβ polypeptide in *E.coli* demonstrated that this post-transcriptional modification is unnecessary for anti-viral activity (Goeddel *et al*, 1980b). In addition non-glycosylated mouse IFNβ produced in *E.coli* maintains its anti-viral properties (Windass, J.D and De-Maeyer, E., personal communication)

Mature human IFNβ contains three cysteine residues, one of which is located at position 17 and is conserved in mouse IFNβ, and two additional residues at position 31 and 141, analagous to the human and mouse IFNα cysteine residues shown to be cross-linked by a disulphide bridge. The isolation and expression in *E.coli* of a

non−antiviral IFNβ variant no longer recognised by anti−IFNβ antibody, which possesses a Cys to Tyr substitution at position 141, demonstrated the functional importance of this residue in IFN protein structure (Shepard *et al*, 1981). Site directed mutagenesis indicated however that Cys 17, can be substituted for serine (Mark *et al*, 1984). Recombinant IFNβ_{ser17} expressed and purified from *E.coli* has the same specific activity as natural IFNβ (Mark *et al*, 1984), indicating that this cysteine is not required for either anti−viral or anti−proliferative activities.

IFNγ

A unique gene encoding IFNγ has been isolated and characterised from the genome of human (Gray and Goeddel, 1982; Taya *et al*, 1982) and mouse cells (Gray and Goeddel, 1983). In contrast to α and β IFNs, the coding sequence is interrupted by three intervening sequences (Gray and Goeddel, 1982; Gray and Goeddel, 1983). The IFNγ gene shows no significant homology with the IFNα and IFNβ genes.

The human IFNγ gene, located on chromosome 12 (Trent *et al*, 1982), encodes a predicted 166 amino acid precursor polypeptide which is modified to a mature 146 amino acid glycoprotein (Rindernecht *et al*, 1984) upon removal of a 20 amino acid signal peptide. The mouse gene, located on chromosome 10 (Naylor *et al*, 1984) encodes a predicted 155 amino acid precursor polypeptide which is cleaved to generate a 136 amino acid mature glycoprotein, shows 64% homology at the nucleotide level with the human gene (Gray and Goeddel *et al*, 1983).

Natural IFNγ has been purified from antigen induced human lymphocyte cultures (Yip *et al*, 1982a; 1982b) and mitogen induced murine T−lymphocytes (Prat *et al*, 1984). Natural preparations of IFNγ consist of two glycoproteins (Kelker *et al*, 1983; Gribaudo *et al*, 1985) derived from a single precursor which apparently undergoes cleavage at the carboxyl terminus (Rindernecht *et al*, 1984; Gribaudo *et al*, 1985). Nevertheless, expression of biologically active IFNγ in *E.coli* demonstrates that neither glycosylation nor the removal of amino acids from the carboxy−terminus are essential for the anti−viral activity (Gray *et al*, 1982; Gray and Goeddel, 1983).

The mature polypeptide for human IFNγ has two cysteine residues at position 1 and

3 (Gray and Goeddel, 1982) in common with the mouse protein, which also has an additional cysteine at position 36 (Gray and Goeddel, 1983). The murine IFNγ protein may potentially possess a disulphide bridge, perhaps accounting for the observed greater stability of the purified protein, relative to the human species (Gray and Goeddel, 1983).

1.2 Regulation of Gene Expression by IFN

IFN Receptors

In the majority of cases the pleiotropic response initiated by the interaction of IFNs with target cells is effected by changes in cellular gene expression (see below). To effect these responses IFNs interact with high affinity cell surface receptors (Kd $1-2x10^{-11}$M). Radiolabelled IFNs (Aguet *et al*, 1980) used in combination with competition binding assays (Branca and Baglioni *et al*, 1981; Faltynek *et al*, 1983) have demonstrated the existence of distinct receptors for IFNα/β (Branca *et al*, 1982) and IFNγ (Anderson *et al*, 1982). These receptors have been identified by chemical cross-linking of IFNs bound to the cell surface, revealing 120kd and 117kd proteins which are thought to be human IFNα/β (Joshi *et al*, 1982; Faltynek *et al*, 1983) and IFNγ (Rashidibaigi *et al*, 1986) receptors, respectively.

Evidence for distinct IFNα/β and IFNγ receptors is strongly supported by the chromosomal locations of the receptor gene(s). Human/mouse somatic cell hybrids indicate that the human IFNα/β receptor gene(s) is present on chromosome 21 (Tan, 1976; Revel *et al*, 1976; Shulman *et al*, 1984), whereas the human IFNγ receptor gene is located on chromosome 6 (Rashidibaigi *et al*, 1986).

The events following receptor occupancy are poorly understood. Internalisation and degradation of receptor/IFN complexes have been observed (Branca *et al*, 1982) but inhibitors of this process (methylamine, chloroquine) indicate that receptor occupancy is sufficient to regulate the transcription of several interferon responsive genes in Daudi and T989 cells (Branca *et al*, 1982; Hannigan and Williams, 1986).

As suggested below all three IFN subtypes regulate the expression of common genes, but there are also some differences which may reflect binding to distinct receptors.

Moreover, receptors may dictate the species specificity commonly observed in the IFN sytem. Recent attempts to clone the human IFNα/β receptor gene(s), by transfection of heterologous cells with genomic DNA have generated recombinant mouse cells binding and responding to human IFNα (Fellous and Rosa, 1985). Thus, this suggests that IFNs are unable to interact productively with heterologous cell receptors. In addition, events following receptor occupancy must be highly conserved between the two species for receptors to function in a heterologous environment.

Interferon Induced mRNA and Protein Synthesis.

IFN receptor occupancy initiates the accumulation of specific mRNAs (Colonno and Pang, 1982) and proteins (Weil *et al*, 1983; Cheng *et al*, 1983) which are believed to elicit the anti-viral, anti-proliferative and immunoregulatory responses. Both IFNα and IFNβ induce a similar profile of proteins in human fibroblasts (Weil *et al*, 1983) while IFNγ induces a minimum of 12 additional polypeptides (Weil *et al*, 1983).

IFN induced enzymes, 2'-5' oligoadenylate synthetase and a ds RNA dependent protein kinase, have been purified (Sen *et al*, 1978; Hovanessian and Kerr, 1979; Galubru and Hovanessian, 1985) and their role as protein synthesis inhibitors in viral infection well established (Review: Lengyel *et al*, 1982). The isolation and characterisation of cDNAs for 2'-5'oligoadenylate synthetase (Merlin *et al*, 1983; Benech *et al*, 1985a; Saunders *et al*, 1985) has led to the discovery that IFN induced *de novo* synthesis of at least two mRNAs, which are derived by cell specific differential splicing of a single precursor transcript (Benech *et al*, 1985b; Saunders *et al*, 1985).

It is well established that IFNs enhance expression of histocompatability antigen (HLA) class I (Basham *et al*, 1982; Burrone and Milstein, 1982) and class II (Rosa *et al*, 1983; Koeffler *et al*, 1984; Collins *et al*; 1984) cell surface antigens in a variety of cell types. IFNs normally enhance expression of HLA mRNA from genes already being expressed (Fellous *et al*, 1982; Rosa *et al*, 1983), although stimulation of *de novo* transcription has been reported for class II genes with IFNγ (Collins *et al*, 1984).

To identify genes implicated in the IFN response, to assign precise physiological roles to the individual gene products and to investigate the mechanism by which genes

are regulated by IFN, many cDNAs encoding cryptic proteins derived from IFN regulated genes have been isolated and characterised. These include the IFNβ treated SV80 cell clone c56, encoding a 56kd protein (Chebath *et al*, 1983), IFNβ induced human fibroblast clones pIFN−IND1 encoding 42kd and 58kd proteins, and the c56 homologue pIFN−IND2 (Larner *et al*, 1984), seven clones from IFNα induced T989 cells, including metallothionein II and a HLA class I gene (Freidman *et al*,1984), and IFNγ induced U937 cell clone pIFN γ−31, which shows homology to platelet factor−4 and β−thromboglobulin derived proteins (Luster *et al*, 1985).

The examination of IFN inducible gene expression has revealed several mechanisms of gene regulation. Two classes of IFN induced gene have been identified, showing (a) transient, or (b) maintained gene expression (Kelly *et al*, 1985). The kinetics of induction can be immediate, increased transcription being apparent within 10 minutes for IFN−IND1 and IFN−IND2 (Larner *et al*, 1984), or delayed 4 hours or more, like the 2'−5'oligoadenylate synthetase gene (Chebath *et al*, 1983). Examination of IFN induced gene expression employing nuclear run off transcription assays has illustrated examples of both transcriptional (Larner *et al*, 1984; Friedman *et al*, 1984) and post−transcriptional (Friedman *et al*, 1984; Friedman and Stark, 1985) control.

The use of metabolic inhibitors suggest that in many cases induction of gene expression by IFN is a primary response (Friedman *et al*, 1984). However, in some cases other factors may contribute to gene expression. The regulation of two transcriptionally activated, transiently expressed IFN inducible genes has been investigated in the presence of the inhibitor of protein synthesis, cycloheximide (CHX). Cytoplasmic mRNA levels of c56 (Chebath *et al*, 1983), IFN−IND1 and IFN−IND2 (Larner *et al*, 1984) genes monitored by hybridisation analysis show that mRNA stability is increased by CHX when administered after IFN. This suggests a labile factor mediates rapid mRNA turnover in these cases. Further studies show that high levels of c56 mRNA are induced by IFNα but not IFNγ. However, pre−treatment of cells with IFNγ prior to induction restores IFNα mediated gene activation even in the presence of CHX. The interpretation of this data is that a labile IFN inducible protein is required for induction of this gene (Kusari and Sen, 1986).

As expected for the different properties of IFNα/β and IFNγ, differential regulation of interferon induced mRNAs have been observed (Kelly *et al*, 1985). Some genes are unique to IFNα or β, including pIFN−IND−1 and IND−2 (Larner *et al*, 1984), 6−16 (Kelly *et al*, 1985) and HLA class I (Fellous *et al*, 1982; Rosa *et al*, 1983), while others respond to IFNγ only, including HLA class II (Rosa *et al*, 1983; Collins *et al*, 1984), and pIFN γ−31 (Luster *et al*, 1985). Therefore, it is clear that although IFNα, IFNβ and IFNγ regulate the expression of common pathways, some genes are also uniquely regulated by these subtypes.

Inspection of the nucleotide sequences of HLA class I, class II, and metallothionein II genes, has revealed a homologous 28bp consensus sequence upstream of the TATA box, common to many IFN induced genes (Friedman and Stark, 1985). To experimentally investigate the location of regulatory sequences, *in vitro* mutagenesis together with gene transfer studies have been employed. Regulated expression of an IFN inducible human HLA class I gene (Yoshie *et al*, 1982) and the human 6−16 gene (Kelly *et al,* 1986) in heterologous mouse cells has been demonstrated, suggesting that the mechanism of gene activation is conserved in the two species. Surprisingly, deletion mutagenesis of an HLA class I gene revealed 5' flanking gene sequences, which include the IFN inducible consensus sequence, were not required for IFN induction. Nuclear run off transcription assays have demonstrated that IFN induces transcription of class I HLA genes (Freidman and Stark, 1985). However, as described previously, IFN can act post−transcriptionally (Freidman *et al*, 1984). Thus, this suggests that the results of Yoshie *et al* (1984) also probably result from post−transcriptional events. Further studies of this kind will be useful to decipher the precise structural features of IFN inducible genes.

IFNs also have negative effects on gene expression. This is illustrated by genes implicated in cell proliferation. Selective reduction of mRNA for the proto−oncogene c−*myc* (Battey *et al*, 1983) by IFNβ in Daudi cells (Jonak and Knight, 1984) correlates with growth arrest in the quiescent G_0/G_1 phase of the cell cycle (Einat *et al*, 1985a). IFNβ inhibits expression of genes normally induced by the competence factor, platelet derived growth factor (PDGF), in Balb/c3T3 cells (Einat *et al*, 1985b), including ornithine decarboxylase (McConologue *et al*, 1984), c−fos (Van Beveren *et al*, 1983) and

c−myc (Battey *et al*, 1983). Several reports of endogenously derived IFNs, expressed in response to growth factors, including IFNα by colony stimulating factor I (CSF−I) treated macrophages (Moore *et al*, 1984), IFNβ by PDGF treated fibroblasts (Zullo *et al*, 1985) and IFNβ_2 in differentiating U937 cells (Resnitsky *et al*, 1986), suggests a mechanism of autocrine growth arrest induced by IFNs, which is possibly mediated through suppression of those genes implicated in cell proliferation.

1.3 Regulation of IFNα and IFNβ Gene Expression.

Human and mouse IFNα and IFNβ are transiently expressed in the presence of virus or the double stranded RNA, polyrI.rC. The kinetics of induction depend on the cell type and inducer (Pestka, 1981), but the transient nature is apparently universal. Metabolic inhibitors have been employed to dissect this process. These have demonstrated that IFN genes are regulated both by transcriptional and post−transcriptional events.

Transcriptional regulation has been revealed by the inhibition of induced IFN production in cells treated with inhibitors of mRNA synthesis (Vilcek and Havell, 1973). More specifically, inhibitors of interferon induction by α−aminitin, which selectively inhibits RNA polymerase II activity, established a role for RNA polymerase II in the transcription of these genes (Raj *et al*, 1979).

Later studies revealed that IFN gene induction results in a transient accumulation of IFN mRNA. The kinetics of IFN production correlates with the appearance of translatable mRNA (Leblau *et al*, 1978; Raj and Pitha, 1977; Cavalieri *et al* 1977). Monitoring of cytoplasmic mRNA levels by hybridisation analysis using cloned IFN gene probes have confirmed these observations (Raj and Pitha, 1981; Hiscott *et al*, 1984; Higashi *et al* 1984; Zwarthoff *et al*, 1985; Kelley and Pitha, 1985b). Therefore, together these data strongly suggest IFN production is initiated by the induction of gene transcription.

The transcriptional regulation of these genes appears to be under both positive and negative control. It is well documented that IFN production can be enhanced by pre−treating cells with IFN (priming) prior to induction. This positive regulation results in either an accelerated release of IFN production (Content *et al*, 1980) or the increased

accumulation of IFN mRNA (Raj and Pitha, 1981), depending on the cells examined. In both cases IFN production is elevated. Further investigation of human fibroblasts revealed that priming is blocked in the presence of CHX. This suggests that the increased mRNA accumulation in primed cells is effected by a labile protein induced by IFN (Raj and Pitha, 1981). In this respect it is interesting that an IFN inducible *trans*−acting factor has recently been identified in IFN inducible cell lines which is absent from non−inducible Hela cells. This factor can complement these normally refractile cells to enable them to produce IFNβ in response to polyrI.rC (Enock *et al*, 1986).

In addition to the positive regulation additional studies with CHX have suggested that IFN gene transcription is also negatively regulated. Inhibitors of protein synthesis alone can activate IFNβ gene transcription in certain circumstances. Reports of CHX induction of the human IFNβ gene linked to an episomal BPV vector in mouse C127 cells (Maroteux *et al*, 1983) and an amplified human IFNβ gene in Chinese hamster ovary cells (Ringold *et* al, 1984), imply the existence of a labile repressor which normally prevents IFN gene transcription. Thus, the induction may result from either the inactivation or inhibition of synthesis of such a repressor in this case.

In addition to transcriptional control, a role for post−transcriptional events in IFN gene expression was revealed by the observation that metabolic inhibitors could also be used to increase the yield of IFN production. A combination of polyrI.rC, CHX and actinomycin D (ActD), administered sequentially, results in a significant increase in IFN production (superinduction) by human fibroblasts with a corresponding increase in the accumulation of translatable mRNA (Raj and Pitha, 1977; Cavalieri *et al*, 1977; Sehgal *et al*, 1977). Monitoring of mRNA levels by hybridisation analysis revealed that superinduction increased the stability of IFN mRNA (Raj and Pitha, 1981). This was confirmed by nuclear run off transcription assays, which suggested that in addition to the normal transcriptional activation of human IFNβ gene expression, superinduction interferes with post−transcriptional events normally associated with the rapid turnover of IFN mRNA.

Thus, IFN gene expression is regulated by transcriptional and post−transcriptional events. Normally transcription from these genes is repressed, probably by negative factors.

Upon induction the repression is relieved and in the presence of the appropriate positive factors, transcription is activated. Transcription is then rapidly shut off, resulting in only the transient accumulation of IFN mRNA which is then degraded by factors acting post−transcriptionally.

1.4 Transcriptional Activation in Eukaryotic Gene Expression

To fully understand the mechanisms of gene regulation it is essential to define both the precise nucleotide sequence of *cis*−acting sequences and their interaction with *trans*−acting factors. Initially, inspection of eukaryotic gene sequences identified the TATA (Corden *et al*, 1980; Table 1.1) and CCAAT (Efstratiadis *et al*, 1980; Table 1.1) box consensus sequences located approximately 30nt and 80nt respectively, upstream of the transcription initiation site. The construction and *in vitro* transcription (Weil *et al*, 1979) of promoter mutants soon established the TATA box as the functional promoter (Corden *et al*, 1980; Wasylyk *et al*, 1980). However, the identification of regulatory elements by comparative structural studies is limited by the heterogeneity of sequences flanking eukaryotic genes, which has prevented the identification of further prospective functional sites in this manner.

Ideally regulatory sequences should be identified on a functional basis. The development of "reverse genetics", whereby cloned genes are mutated *in vitro*, and assayed by reintroduction into amphibian (McKnight and Gavis, 1980) or mammalian cells (Wigler *et al*, 1979; Pellicer *et al* , 1980) has made it possible to investigate regions important in the control of a) constitutive (McKnight and Gavis, 1980; Gruss *et al*, 1981), b) tissue specific and developmentally regulated (Charnay *et al*, 1983; Chao *et al*, 1983; Wright *et al*, 1983; Wright *et al*, 1984) and c) induced (review, Serfling *et al*, 1985) gene expression.

Application of this methodology to the HSV−I *tk* (McKnight *et al*, 1981; McKnight and Kingsbury, 1982), human α globin (Mellon *et al*, 1981) and rabbit β globin (Dierks *et al*, 1983) genes has confirmed the importance of the TATA box as a probable recognition sequence for RNA polymerase II, which determined the accuracy of transcription initiation. Nevertheless a class of gene promoters, transcribed by RNA

TABLE 1.1 Some Functional Sequences of Eukaryotic Promoters

TATA box: TATAAA (Corden *et al*, 1980)

Constitutive Promoters

CAT box: CCAATT (Efstratiadis *et al*, 1980)

SP1 binding sites: GGGCGG (Gidoni *et al*, 1984)

Core enhancer: GTGGAAAG (Weiher *et al*, 1983)

Regulated Promoters

H.R.E: GTNGAANNTTCNAG (Pelham and Bienz, 1982)

M.R.E: TG T
CGCCCGGC C (Karin *et al*, 1984a)
GT C

G.R.E: TGGTACAAAATGTTCT (Karin *et al*, 1984a)

I.R.E: α GAGTGCATGAAGGAAAGCAAAAACAGAAATGGAAAGTGGCCCAGAA (Ryals *et al*, 1985)

β GAGAAGTGAAAGTGGGAAATTCCTCTGAATAGAGAGGAC (Goodbourn *et al*, 1985)

IFN regulated promoters: TTCN NACCTCNGCAGTTTCTC TCTCT (Friedman and Stark, 1985)
G C
C T

N = any nucleotide, G.R.E = glucocoricoid regulatory element, H.R.E = heat regulatory element, M.R.E = metal regulatory element and I.R.E = interferon gene regulatory element (minimum sequence of human IFNα_1 and IFNβ genes required to confer inducibility on a heterologous promoter)

polymerase II, lacking the TATA box has recently been defined. These include genes encoding "house keeping" functions (for review see Dynan, 1986), and others including the *ras* gene family (Ishii *et al,* 1985; Hall and Brown, 1985).

"Reverse genetics" has also allowed the classification of a new class of regulatory sequences which determine the efficiency of transcription initiation of a gene promoter, irrespective of the presence of a TATA box. These sequences, usually found upstream of the gene, are subdivided into two broad groups, distinguishable by the distance (relative to the promoter) overwhich they operate.

Upstream activators represent the first class of *cis*−acting transcriptional activators. The efficient transcription of the HSV−I *tk* gene (McKnight *et al*, 1980) is dependent upon two upstream, spatially separated inverted repeats (McKnight and Kingsbury, 1982), both of which contain the six nucleotides, GGGCGG (Table 1.1). Homologous sequences have been identified in a variety of viral and cellular gene promoters. These include the twenty one base pair repeats of the SV40 early promoter (Benoist and Chambon, 1981), the human Ha−c−*ras* gene promoter (Ishii *et al*, 1985), the human N−*ras* promoter (Hall and Brown, 1985), the human and mouse metallothionein gene promoters (Glanville *et al*, 1981) and the promoters of genes encoding "house keeping" functions (Dynan, 1986). The most upstream of the HSV−I *tk* promoter inverted repeats is also associated with an inverted CCAAT box (Graves *et al,* 1986). Similarly, expression of the human α_1 globin (Mellon *et al*, 1981) and rabbit β globin genes (Dierks *et al*, 1983) also require two upstream regions, including a CCAAT box, for maximum promoter activity. Although the upstream sequences of HSV−I *tk* and globin gene promoters are not closely related, they can functionally replace one another (Cochran and Weissmann, 1984). Thus, these sequences are able to regulate the expression from heterologous promoters. In addition, the location of the two upstream elements, with respect to one another or the TATA box, can be altered by a limited distance, or the orientation of the elements reversed, while still maintaining promoter activity (McKnight, 1982; Cochran and Weissmann, 1984).

Enhancers represent a second class of *cis*−acting transcriptional activator sequences (Moreau *et al*, 1981; Banerji *et al*, 1981; Gruss *et al*, 1981). Originally recognised in the

72bp repeat of SV40 (Moreau *et al*, 1981), they activate transcription (Treisman and Maniatis, 1985; Weber and Schaffner, 1985) of heterologous promoters (Banerji *et al*, 1981) in an orientation independent manner, over distances of several hundred nucleotides, whether they lie 5' or 3' to the gene (Khouri and Gruss, 1983). In addition enhancers can show tissue specificity (deVilliers and Schaffner, 1981; Laimins *et al*, 1982; Kriegler and Botchan, 1983), and can therefore dictate the viral host range (Bosze *et al*, 1986).

Enhancers have been identified in the genome of a variety of viruses. Although regions of extensive sequence homology have not been found in these functionally related elements, short functionally important consensus core sequences have been identified (Hen *et al* 1983; Lusky *et al*, 1983; Weiher *et al*, 1983). However, functional enhancers lacking these consensus core sequences have also been observed (Weber *et al*, 1984; Swimmer and Schenk, 1984). It appears therefore, that enhancers are comprised of multiple motifs. These probably act in concert to exert the transcriptional effect, although there is also, possibly, some degeneracy of information (Zenke *et al*, 1986).

Characterisation of the first cellular enhancers, identified in the introns of immunoglobulin heavy chain genes (Banerji *et al*, 1983; Gillies *et al*, 1983; Queen and Baltimore, 1983) and κ light chain genes (Picard and Schaffner, 1984) revealed that these elements were not peculiar to viruses. This highlighted their potential for regulating tissue specific expression in cellular differentiation. Subsequently, cellular enhancers have been implicated in the tissue specific expression for insulin (Walker *et al*, 1983; Edlund *et al*, 1985) and other pancreatic genes (Boulet *et al*, 1986). Enhancer−like sequences have also been identified in HLA genes (Kimura *et al*, 1986).

The catalogue of cellular enhancers has been extended by the inclusion of inducible gene promoters in this class of *cis*−acting sequences. Regulated transcription of a) the drosophila heat shock inducible protein 70 gene, *hsp*70, (Corces *et al*, 1981; Burke and Ish−Horowicz, 1982), b) the heavy metal inducible mouse metallothionein 1 gene, *MT*−1 (Mayo *et al*, 1982), c) the heavy metal and glucocorticoid inducible human metallothionein II gene, h*MT*−11_A, (Karin *et al*, 1983) and d) the glucocorticoid inducible mouse mammary tumour virus LTR, MMTV, (Hynes *et al*, 1981) and human

growth hormone (Robins *et al*, 1982) genes are well documented. Deletion mutagenesis and *in vivo* expression studies have identified prospective consensus sequences (Table 1.1) responsible for the induction by heat shock (Pelham, 1982; Heat Shock Regulatory Element, H.R.E.), heavy metals (Brinster *et al*, 1982; Richards *et al*, 1984; Metal Regulatory Element, M.R.E) and glucocorticoids (Karin *et al*, 1984a; Majors and Varmus, 1983; Glucocorticoid Regulatory Element, G.R.E). Such sequences of *hsp*70 (Pelham and Bienz, 1982), *MT*−1 (Brinster *et al*, 1982), h*MT*−II_A (Karin *et al*, 1984b) and MMTV (Chandler *et al*, 1983, Hynes, *et al*, 1983) confer inducibility upon heterologous promoters. In addition the H.R.E. of *hsp*70 and the G.R.E.s of MMTV and hMT−II_A genes, behave in a location and orientation independent fashion (Chandler *et al*, 1983; Karin *et al*, 1984b, Bienz and Pelham, 1986). Thus, both H.R.Es and G.R.Es are considered inducible enhancers (Bienz and Pelham, 1986; Chandler *et al*, 1984), while M.R.Es show analogous properties.

It is highly likely that *cis*−acting elements act by binding *trans*−acting factors. However very few *trans*−acting factors have been identified todate. Prospective transcription factors, SP1 (Dynan and Tjian, 1983) and CAT binding protein (CBP; Graves *et al*, 1986) or CAT transcription factor (CTF; Jones *et al*, 1985) have been purified. Originally they were shown to interact *in vitro* with the reiterated GGGCGG (SP1 binding sites) sequences of the 21bp repeat of the SV40 promoter (Gidoni *et al*, 1984) and the CCAAT box of the HSV−I *tk* promoter (Jones *et al*, 1985; Graves *et al*, 1986) respectively (review, McKnight and Tjian, 1986). Subsequently SP1 has been demonstrated to bind *in vitro* to other gene promoters containing homologous hexanucleotide sequences, including the HSV−I *tk* inverted repeats (Jones *et al*, 1985), the h*MT*I_A and h*MT*II_A promoters (Kadonaga *et al*, 1986), the mouse dihydrofolate reductase promoter (Kadonaga *et al*, 1986), two SV40 related monkey gene sequences (Dynan *et al*, 1985) and the HTLVIII/LAV LTR sequence (Kadonaga *et al*, 1986). Similarly CBP binds to the CCAAT box of MSV (Graves *et al*, 1986). Thus, both SP1 and CBP (or CTF) may be universal transcription factors which bind to promoter sequences of viral and cellular origin.

Both *in vivo* (Scholler and Gruss, 1984; Mercola *et al*, 1985) and *in vitro*

(Sassore−Corsi *et al*, 1984; Sassone−Corsi *et al*, 1985; Wilderman *et al*, 1984) transcription studies and *in vivo* binding studies (Church *et al*, 1985) strongly suggest that enhancers also function by interacting with trans−acting factors. Prospective transcription factors that interact *in vitro* with the inducible enhancers of the *hsp*70 and MMTV gene promoters have been purified. Thus, the heat shock transcription factor (HSTF; Parker and Topol, 1984) and glucocorticoid receptors (Scheiddereit *et al*, 1983) interact with induced *hsp*70 (Wu, 1984a) and MMTV promoters respectively.

Some inducible gene promoters contain prospective SP1 (mouse MT−1) and CBP (*xenopus hsp*70; Bienz, 1986) binding sites. Nevertheless, it is highly likely that different inducible promoters also require unique trans−acting factors to activate gene transcription. This is clearly evident by the heterogeneity of the functional, *cis*−acting, G.R.E.'s H.R.E.'s and M.R.E.'s, observed in promoters induced by different agents.

1.5 *Cis*−regulation of Human IFN Gene Expression.

As discussed previously, IFN gene expression is tightly regulated (Section 1.3). Inspection of both human IFNα_1 and IFNβ gene sequences has revealed a TATA box homology, 28bp (Nagata *et al*, 1980) and 24bp respectively (Ohno and Taniguchi, 1981), upstream of the transcription initiation sites. The TATA box sequence of the IFNα genes examined from human, mouse, rat and bovine sources, is a highly conserved variant, TATTTAA. The IFNβ genes examined possess a more conventional sequence. The relevance of this difference between IFNα and IFNβ genes is is unknown.

IFN gene promoters lack further consensus sequences commonly found in many other gene promoters. Consequently "reverse genetics" has been extensively employed to identify *cis*−acting sequences which contribute to the mechanism of virus and polyrI.rC regulated gene expression. As both human and mouse IFNα and IFNβ are distinguishable both by their different target cell specificities (Stewart, 1979) and by differential hybridisation, human IFN gene expression has tended to be examined in heterologous mouse cells.

The original demonstration of human IFNα_1 mRNA production by NDV induced mouse L cell lines stably transfected with integrated copies of the corresponding gene

(extending 5.4kb upstream and 1.2kb downstream of the coding sequence), with the same kinetics of endogenous mouse IFN mRNA production, indicated that the regulatory mechanism was conserved between the two species (Mantei and Weissmann, 1982). Similar studies soon followed with the virus or polyrI.rC inducible human IFNβ gene either integrated in the host genome (Canaani and Berg, 1982; Hauser *et al,* 1982; Ohno and Taniguchi, 1982) or on an extrachromosomally replicating bovine papilloma virus (BPV) vector (Zinn *et al*, 1982).

Both human IFNα and IFNβ genes behaved as expected in heterologous mouse cells. The kinetics of human IFN mRNA accumulation, monitored both by Northern blot analysis (Ohno and Taniguchi, 1982; Hauser *et al*, 1982; Zinn *et al*, 1982) and S−1 nuclease protection (Mantei and Weissmann 1982; Hauser *et al*,1982; Canaani and Berg, 1982; Zinn *et al,* 1982), confirmed the production of correctly initiated gene transcripts in induced cells only. As these genes were regulated in the same manner as endogenous IFN genes, described previously (Section 1.3), biologically relevant studies could be performed using gene transfer procedures.

Although the majority of cell lines investigated in such studies contained inducible integrated IFN genes, significant quantitative differences were observed between independent clones, which probably arose from the random positioning of genes within the host genome (Canaani and Berg, 1982; Hauser *et al*, 1982; Mantei and Weissmann, 1982; Ohno and Taniguchi, 1982). The variation has been successfully reduced in stable transfection systems by pooling independent colonies transfected with human IFNα_1 genes to give an average estimate of expression (Weidle and Weissmann, 1983) or by employing BPV vectors (Zinn *et al,* 1982).

Although both these solutions have emerged as the major procedures for examining IFN gene expression, a major criticism is that they are time consuming. Additionally, the disadvantage of pooling is that many transfected colonies are required to generate a representative population. BPV vectors on the other hand are severely limited by the host range, but more importantly they contain enhancers (Lusky *et al*, 1982) which may influence gene expression.

Location of Human IFNα and IFNβ *Cis*−acting Sequences

In attempts to locate the regulatory sequences of the human IFNα and IFNβ genes, attention has focussed upon the 5' flanking sequences. This is justified because of the results obtained by linking together the human IFNα_1 5' flanking sequence from −675 to −5 (relative to the transcription initiation site) and the rabbit β globin gene coding sequence (Weidle and Weissmann, 1983) or the human IFNβ gene 5' sequences from −284 to +20 and the HSV−I *tk* gene coding sequence (Ohno and Taniguchi, 1983). In each case viral regulated activation of heterologous gene transcription was seen in transfected mouse cells. Thus, the IFN promoters activate gene transcription in induced cells.

The 5' boundary of *cis*−acting sequences responsible for induction of the human IFNα_1 and IFNβ genes has been investigated by deletion mutagenesis together with *in vivo* expression studies. This revealed that 117bp (relative to the transcription initiation site) upstream of the IFNα_1 (Ragg and Weissmann, 1983) or IFNβ (Fujita *et al*, 1985) genes are necessary for full transcriptional activation by virus in mouse L cells. Surprisingly, only 77bp (relative to the transcription initiation site) upstream of the IFNβ gene are required for full polyrI.rC induced (Zinn *et al*, 1983) or virus induced (Goodbourn et al, 1985) activation of transcription of extrachromosomal (Zinn *et al*, 1983) or integrated (Goodbourn *et al*, 1985) genes in C127 cells. The basis for this descrepency has not been established but may be a result of different mechanisms of regulation in these cells (Goodbourn *et al*, 1985). Nevertheless, the BPV vectors used by Zinn *et al* (1983) contain enhancer sequences, as do the vectors employed to generate the cell populations containing integrated human IFNβ gene deletion mutants (Goodbourn *et al*, 1985). Thus, it is possible that the BPV vectors used may supplement sequences of the HuIFNβ gene promoter deleted between −117 and −77.

Subsequently, it has been demonstrated that IFN gene upstream sequences can dictate the efficiency of transcription initiation directed by heterologous gene promoters. A minimum of 46 nucleotides of the human IFNα_1 gene extending from −109 to −64 (Table 1.1), linked to a truncated rabbit β globin transcription unit (−56), is sufficient to produce correct and efficient transcription initiation from the globin TATA box

promoter in virus induced mouse L cells (Ryals *et al*, 1985). Likewise, linking the segment of the human IFNβ promoter extending from −77 to −37 (Table 1.1) to a hybrid transcription unit comprising the HSV−I *tk* gene promoter (−39 to +57) and the IFNβ gene coding sequences confers polyrI.rC inducibility upon the heterologous promoter (Goodbourn *et al*, 1985). The −77 to −37 sequence has been designated the interferon gene regulatory element (I.R.E., Goodbourn *et al*, 1985)

These properties suggested that the IFN gene upstream sequences may contain inducible enhancers. This has been confirmed by demonstrating that the I.R.E. conferred orientation independent polyrI.rC inducibility when located 900bp upstream or 350bp downstream of the poorly inducible −73 truncated IFNβ gene transcription unit (Goodbourn *et al*, 1985). This functional evidence is supported by the partial sequence homology of the I.R.E. with the consensus viral enhancer core sequence (Goodbourn *et al*, 1986).

Studies of HuIFNβ gene expression in virus induced stably transfected mouse L cells have failed to confirm these observations. Functional orientation independence of the −125 to −39 human IFNβ gene promoter fragment has been demonstrated when juxtaposed to the IFNβ gene containing only 38bp (relative to the transcription initiation site) of 5' flanking sequence (Fujita *et al*, 1985). However, when the −125 to −39 fragment was located 3' to this gene, no transcription was observed in the presence or the absence of virus (Fujita *et al*, 1985).

These apparently contradictory results were obtained with different inducers with extrachromosomal (Goodbourn *et al*, 1985) or integrated genes (Fujita *et al*, 1985). Thus, some of the observed differences may be attributed to this. However, no firm conclusions can be made concerning the behaviour of the IFNβ gene upstream sequences from the study of Fujita *et al* (1985) in the absence of any data demonstrating location independent activation of the 5' truncated IFNβ gene by a characterised enhancer.

Thus, strong evidence suggests that both human IFNα and IFNβ gene promoters possess upstream sequences which dictate the efficiency of transcription from the TATA box promoter element. They behave in a positive manner, significantly increasing the frequency of transcription initiation in the presence of virus or polyrI.rC. Hence

functional *cis*−acting sequences of IFN promoters share many properties of upstream elements of other eukaryotic genes described previously (Section 1.3). Futhermore, the human IFNβ gene I.R.E. at least, may be a member of the growing list of eukaryotic inducible cellular enhancers.

Although the human IFNβ gene therefore seems likely to contain an enhancer, it remains to be established if enhancer−like properties are common to other IFN promoter sequences. Prior to this study, the only other IFN gene promoter to be examined was the human IFNα_1 gene promoter. There are a minimum of 23 human IFNα genes, at least nine of which are expressed *in vivo* (Section 1.1). However, it is not clear if the remainder are also expressed in *in vivo*. Furthermore, differential regulation of human IFNα gene transcription has been described (Hiscott *et al*, 1984). Transcripts from the human IFNα_{14} gene are a minor component of virus induced leukocytes but a major component of induced leukemic myeloblast mRNA (Hiscott *et al*, 1984).

The functional basis for these observations are not known. "Reverse genetics" provides a means to compare the relative strengths of IFNα promoters. This should reveal if those genes not identified to be expressed *in vivo* actually contain a functional promoter. Futhermore, such studies may identify IFNα promoter sequences responsible for differential gene expression (Hiscott *et al*, 1984), perhaps determining tissue specificity. Studies of this nature should yield valuable information concerning the regulation of IFN gene expression. Thus, the first part of this thesis describes the isolation and characterisation of human IFNα chromosomal genes as a preliminary step toward investigating the functional activity of human IFNα gene promoters.

Negative *Cis*−acting Sequences in IFN Gene Expression

Despite the accumulating evidence for the positive regulation of IFN gene expression, IFN genes might also be under negative control. Deletion of the human IFNβ gene between −210 and −107 produces a two fold increase in polyrI.rC inducibility (Zinn *et al*, 1983). An analogous observation has also been reported for the human IFNα_1 gene (Ragg and Weissmann, 1983). Further investigation revealed the deletion mutant produced an acceleration of polyrI.rC induced mRNA accumulation (Zinn *et al*,

1983), reminiscent of the priming phenomenon described previously (Section 1.3). Therefore, it has been speculated that priming with IFN de−represses a control region located between −210 and −107 (Zinn *et al*, 1983). Thus, this defines a second regulatory region of IFNβ genes, upstream of the I.R.E. described previously.

Even more compelling evidence for negative regulation has been revealed by the ability of the I.R.E. to repress transcription from functional heterologous promoters in non−induced cells (Goodbourn *et al*, 1986). Thus, constitutive expression from the intact HSV−I *tk* promoter extending from −105 to +57, is significantly reduced by the orientation independent juxtaposition of the I.R.E. (Goodbourn *et al*, 1986). Other gene promoter sequences have been identified which also inhibit the activity of heterologous promoter activity. These sequences, identified in the rat 1 insulin gene (Laimins *et al*, 1986), the mouse c−*myc* gene (Remmers *et al*, 1986) and the virus HTLV−III/LAV LTR (Rosen *et al*, 1985) are able to inhibit heterologous promoter activity in an orientation and location independent fashion. Thus, they display the opposite properties to those of enhancers.

Thus, the human IFNβ gene promoters have both positive and negative regulatory properties (Zinn *et al*, 1983; Goodbourn *et al*, 1985; 1986; Fujita *et al*, 1985). Similar observations have been made with the human IFNα_1 gene promoter (Ragg and Weissmann, 1983, Ryals *et al*, 1985).

Cis−acting sequences probably interact with *trans*−acting factors. The interaction of proteins involved in the regulation of specific genes frequently correlates with changes in chromatin structure. Initially, selective DNase I digestion of globin (Weintraub and Groudine, 1976) or ovalbumin (Garel and Axel, 1976) genes revealed chromatin conformation changes in genes expressed in a tissue specific fashion. Transcriptionally active genes possess regions preferentially sensitive to DNase I (Stadler *et al*, 1980). The appearance of DNase I hypersensitive sites upon induction of drosophila heat shock protein genes (*hsp*; Wu *et al*, 1980) revealed protein binding sites (Wu, 1984b) and as a result a heat shock transcription factor (HSTF) has been purified (Wu, 1984a; Parker and Topol, 1984).

Similar techniques applied to human and mouse IFNβ genes indicates increased

sensitivity to DNase I in the induced state only (Coveney *et al*, 1984). Several distinct DNase I hypersensitive sites have been mapped around the mouse IFNβ gene (Higashi *et al*, 1984). One of these, located approximately 100bp upstream of the transcription initiation site, is absolutely dependent upon induction. This suggested the region corresponding to the *cis*−acting sequences identified by deletion mutagenesis interacts with *trans*−acting factors upon induction. The complexity of the IFNα gene family has prevented a similar approach.

These studies have been extended by the development of an *in vivo* genomic footprinting procedure to identify the interaction of *trans*−acting factors with DNA sequences (Zinn and Maniatis, 1986). By this means two regions of the IFNβ gene promoter have been defined which interact with *trans*−acting factors, located between −167 to −94 and −68 to −38 (relative to the transcription initiation site). After induction these factors dissociate and a new factor binds to a region located between −77 to −64. Based on these observations a model has been proposed for the regulation of IFN gene expression (Zinn and Maniatis, 1986).

Figure 1.1 summarises this model. It is suggested that transcription of the IFN promoter is repressed by *trans*−acting factors bound to the gene 5' flanking sequences at two distinct locations. The factor bound at the −167 to −94 location may explain the observation that deletion of this region results in increased promoter activity. Upon induction the repressors dissociate, by an unknown mechanism, from the promoter region enabling a prospective interferon transcription factor (ITF) to bind to the constitutive enhancer sequence located between −77 and −64 (Goodbourn *et al*, 1986). This factor behaves in a positive fashion, dictating efficient transcription initiation from the human IFNβ gene promoter.

Thus, this model provides for both the positive and negative aspect of interferon gene regulation discussed previously. The location of the positive regulatory sequences downstream of −100 (relative to the transcription initiation site) is consistent with the deletion mutagenesis studies of both human IFNα_1 and IFNβ genes. The negative control of gene activation is also consistent with the observations demonstrating the induction of IFN gene expression by inhibitors of protein synthesis (Section 1.4).

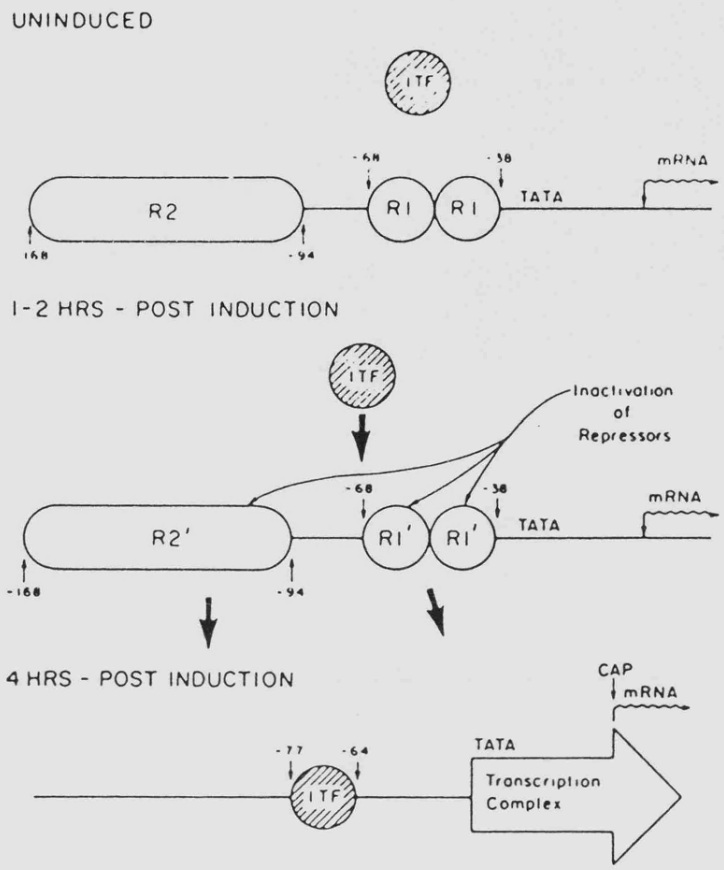


FIGURE 1.1 Molecular Mechanism of IFNβ Gene Regulation (copied from Zinn and Maniatis 1986)

R1 and R2 are repressor molecules. ITF is a prospective interferon gene transcription factor which interacts with the IFNβ gene promoter enhancer sequences activating gene transcription.

Cis–activation of Mouse IFNα Gene Expression

As discussed previously, human and mouse IFNα gene sequences are highly conserved (Section 1.1). Inspection of the primary nucleotide sequence of the mouse IFNα_1 gene reveals a TATA box 28bp upstream of the transcription initiation site (Shaw *et al*, 1983), analogous to the human IFNα_1 gene (Nagata *et al*, 1980). Further analysis reveals elements of homology between the human and mouse IFNα_1 5' flanking sequences (This work).

The expression of a cloned mouse IFNα gene after the introduction into mammalian cells has not been investigated previously. Such studies provide the first opportunity to examine polyrI.rC mediated activation of an IFNα gene, since human IFNα genes are not normally efficiently induced by polyrI.rC (Section 1.1). Therefore polyrI.rC mediated activation of the mouse IFNα_1 gene promoter has been investigated in the work described here.

Until now, IFN gene expression studies have been undertaken in heterologous cells (Section 1.3). Although satisfactory, the possibility of some incompatability between the human and mouse regulatory systems remains (Mantei and Weissmann *et al*, 1982). Thus, mouse IFNα_1 promoter function has been investigated in mouse cells.

To conduct expression studies in cells which normally produce mouse IFNα the gene promoter has been used to direct transcription of the bacterial gene, chloramphenicol acetyl transferase (*cat*). The CAT expression assay provides a sensitive and quantitative means of examining promoter activity which has been employed extensively in eukaryotic gene regulatory studies (Gorman *et al*, 1982). Normally mammalian cells do not contain the *cat* gene, therefore CAT can be assayed in the absence of background activity (Gorman *et al*, 1982). Thus, it provides an ideal system for examining promoter function in mammalian cells.

Eukaryotic promoter function has been examined in both transiently and stably transfected cells using the CAT system (Gorman *et al*, 1982; 1983). The problem of clonal variation with stable transfection studies has been discussed previously (Section 1.5). Transient expression assays have the advantage in quantitative studies of both the

extrachromosomal location of the exogenous DNA and speed. Thus, both transient and stable transfection have been examined here as means to investigate polyrI.rC induced mouse IFNα_1 promoter function in mouse L−929 cells.

Differential expression of endogenous mouse IFNα genes have been observed in virus induced mouse L−929 cells (Kelley and Pitha, 1985b; Zwarthoff *et al,* 1985). All IFNα genes examined are apparently transcribed, except the mouse IFNα_1 gene. Furthermore, S−1 nuclease protection studies have demonstrated an estimated 15 fold elevation of mouse IFNα_4 gene mRNA accumulation, relative to other IFNα genes examined, in virus induced mouse L−929 cells. The basis of these observations are unknown but may be a consequence of the respective gene promoters. Therefore, the CAT expression system has been employed to investigate the functional activity of the mouse IFNα_1 promoter in L−929 cells. Thus, this has provided an opportunity of assessing IFNα promoter function in cells in which promoter activity may be differentially regulated.

As described previously (Section 1.5), IFN promoters possess inducible enhancer like properties which stimulate efficient transcription initiation of IFN genes in the presence of polyrI.rC (Goodbourn *et al*, 1985). In view of the prospective weak mouse IFNα_1 promoter activity, the effect of a constitutive viral enhancer on both the regulation and efficiency of polyrI.rC induced *cat* gene transcription from the MuIFNα_1 gene promoter has also been investigated in L−929 cells. The Mo−MuSV enhancer was chosen for this purpose because it functions very efficiently in mouse fibroblast cells (Laimins *et al*, 1982).

For the first time, site directed mutagenesis has been employed to investigate the regulation of IFN gene expression. Previously, *in vitro* deletion mutagenesis has determined that a minimum of 46 nucleotides (Ryals *et al*, 1985) and 40 nucleotides (Goodbourn *et al*, 1985) of 5' flanking sequence respectively are sufficient for virus (human IFNα_1) and polyrI.rC (human IFNβ) regulation of gene expression. Deletion mutagenesis can only locate the prospective boundaries of *cis*−acting sequences and moreover, generates gross structural rearrangements of the surrounding genetic material. Site directed mutagenesis however, enables predetermined base substitutions to be made (Zoller and Smith, 1983) to determine the precise nucleotide sequence of *cis*−acting

regulatory sequences. Moreover, the contribution of these sites to promoter function can be examined with the surrounding DNA sequences intact.

Site directed mutagenesis has been employed succesfully to confirm the functional significance of a variety of *cis*−acting regulatory elements. These include the rabbit β globin TATA and CCAAT boxes (Dierks *et al*, 1983), the HSV−I *tk* inverted repeats (McKnight *et al*, 1984) and inverted CCAAT box (Graves *et al*, 1986) and the SV40 consensus core enhancer sequence (Weiher *et al*, 1983; Zenke *et al*, 1986).

Thus, a detailed comparative study of human and mouse IFNα promoter sequences has been undertaken to identify prospective regulatory sequences. Based on this analysis, specific nucleotide substitutions of the mouse IFNα_1 promoter have been generated by the oligonucleotide site directed mutagenesis procedure of Zoller and Smith (1983). Their functional activity has been evaluated by employing the CAT expression assay in transiently and stably transfected cells. These studies demonstrate the relative success of the stable transfection system to deduce useful information on the activity of upstream regulatory sequences of the MuIFNα_1 gene promoter.

CHAPTER 2

MATERIALS AND METHODS

2.1 General Methods

Siliconisation of Glassware

Glassware was siliconised by either coating directly with 2,4 dichlorodimethylsilane (5% (v/v) solution in $CHCl_3$) or by allowing the siliconising agent to vapourise under vacuum in a dessicator containing the glassware (for small items such as micropipettes). This procedure was performed in a fume cupboard.

Maintenance of Bacterial and Phage Stocks

Stocks of bacterial strains were stored at -70°C as frozen cultures, which were prepared by mixing 1ml of a fresh o/n culture with 1ml 80% glycerol. Stocks of λ recombinants were stored in either L.broth or λ buffer over 1% (v/v) $CHCl_3$

2.2 Sources of Chemicals and Biochemicals

Restriction endonucleases and DNA modifying enzymes were obtained from Amersham International, Amersham, UK, with the following exceptions: *Bst*NI, New England Biolabs; T_4 polynucleotide kinase, New England Biolabs; DNA polymerase I (Klenow), Boehringer Mannheim.

Radiochemicals were supplied by Amersham International with the exception of ^{14}C-chloramphenicol which was obtained from New England Nuclear.

Ultra-pure deoxy- and di-deoxy sequencing nucleotides, polyrI.rC, DEAE dextran, Sephadex G-50 and DE-52 cellulose were supplied by Pharmacia.

Tissue culture petri-dishes, flasks, media, serum, trypsin, additives including G418 and L-929 cells (ATCC No CCL 1) were supplied by BRL.

Nitrocellulose paper and filters were from Schleicher and Schull. Two types of X-ray film were used: Fuji RX and Kodak XAR-5. Ethidium bromide stained agarose and acrylamide gels were photographed using a Polaroid Land camera and Type 55 Polaroid film.

2.3 Bacterial and Phage Strains and Plasmids

Bacterial Strains			Source
C600	-	F^-, *sup*E44, *thi*1, *thr*1, *leu*1B6, *lac*Y1, *ton*A 21, λ^-	J. Windass
JA221	-	F^-, *hsd*R, *lac*Y1, *leu*B6, *rec*A, *trp*E5	J. Windass
JM101	-	*sup*E, *thi*Δ (*lac-pro*AB), [F'*tra*D36, *pro*AB, *lac* $I^qZ\Delta M15$]	Amersham
BHB 2688	-	*rec*A, *sup*°, (λ*imm*434, *cI*ts, *b*2, *red*3, Dam15, Sam7)	D. Burt
BHB 2690	-	*rec*A, *sup*°, (λ*imm*434, *cI*ts, *b*2, *red*3, Eam4, Sam7)	D. Burt
BW313	-	HfrKL16, Po/45, [*Lys*A(61-62)], *thi*-1, *rel*A1, *spo*T1, *dut*-1, *ung*-1	T. Kunkel
NM538	-	*met*B, *sup*E, *sup*F, $hsdR_K$, *ton*A, *trp*R, λ^-	D. Burt
NM539	-	*met*B, *sup*E, *sup*F, $hsdR_K$, *ton*A, *trp*R, λ^-, P2	D. Burt

Phage Strains			
M13mp8	-	Messing, 1983	Amersham
M13mp9	-	"	Amersham
M13mp10	-	"	Amersham
M13mp11	-	"	Amersham
λL47	-	λ*sbh*I°, *chi*A131, Δ(*s*RIλ1-2), *imm*434, *cI*, *s*RIλ4°, *nin*5, *s*RIλ5°, *shn*dIIIλ6°	W.J. Brammar

Plasmids

pAT153	- ApTc (Twigg and Sheratt, 1980)	J. Windass
pUC13	- Ap (Messing, 1983)	Pharmacia
pTCF	- Ap G418 (Grosveld *et al*, 1982)	J. Windass
pSV2*cat*	- Ap (Gorman *et al*, 1982)	B. Howard

2.4 Preparation of Buffers and Stock Solutions

Buffers and stock solutions mentioned but not detailed in the methods described below were prepared as follows:

T.E - 10mM Tris-HCl, pH7.5
1mM EDTA

λ Buffer - 6mM Tris-HCl, pH8
10mM $MgCl_2$
100mM NaCl
0.5mg/ml gelatine

Bacterial Buffer - 50mM Na_2HPO_4
20mM KH_2PO_4
70mM NaCl
1mM $MgCl_2$

PBS - 140mM NaCl
3mM KCl
8mM Na_2HPO_4
1.5mM KH_2PO_4
pH7.3

Proteinase K Buffer - 20mM Tris-HCl, pH8
1mM EDTA, pH8
0.1M NaCl
0.01% Triton X100

10× T_4 Ligation Buffer - 700mM Tris-HCl, pH7.5
70mM $MgCl_2$
0.7mM ATP

TNE - 50mM Tris-HCl, pH7.5
0.1M NaCl
5mM EDTA

SSC - 0.15M NaCl
0.015M Sodium citrate, pH7

SSPE - 0.18M NaCl
0.01M NaH_2PO_4, pH6.8
0.001M EDTA

T.A - 40mM Tris
20mM NaOAc, pH7.8
0.2mM EDTA

T.B.E - 0.9M Tris
0.9M Boric Acid, pH8.3
0.025M EDTA

L.broth - Difco Bacto Tryptone 10g
Difco Yeast extract 5g
NaCL 5g
Glucose 10g
H_2O to 1L, pH7

L.Agar - As for L-broth without glucose, plus 17g/L difco agar.

BBL.Agar - Baltimore Biol.Lab. Trypticase 10g
NaCl 5g
Difco Agar 15g*
*(6.5g top agar)
to 1L with H_2O

YT Medium - Bacto tryptone 8g
Yeast extract 5g
NaCl 2.5g
to 1L with H_2O

H Plates - Bacto tryptone 10g
NaCl 8g
Agar 12g (8g for top)
to 1L with H_2O

Sample Buffer - (for agarose and acrylamide gels)
20% (w/v) Ficoll
0.002% (w/v) Bromophenolblue

Agarose Beads	- Agarose	0.2g
	10mM Tris-HCl (pH7.5)	0.1ml
	20mM EDTA	2ml
	Glycerol	10ml
	0.01% (w/v) Bromophenolblue	10mg
	DDW	85ml

Formamide Dye Mix - 100ml of re-crystallised Formamide
0.1% (w/v) xylene cyanol FF
0.1% (w/v) bromophenol blue
4ml 500mM EDTA

5× Denhardt's	- Ficoll	0.5g
	polyvinylpyrrolidone	0.5g
	BSA (Boehringer Fraction V)	0.5g
	H_2O to 500ml	

2.5 General Molecular Biology Methods

Agarose Gel Electrophoresis

DNA samples were analysed by electrophoresis on various concentrations of agarose gel, depending on the size of the fragments anticipated. Generally, 0.6% to 1% gels were used for restriction digests with the exception of *Sau*3A partial digests of human genomic DNA which were analysed on 0.2% gels underlaid with a 1% gel to give support. Preparative gels for restriction fragment isolation were run on 1% LMP agarose (BRL). The buffer used was either T.A or T.B.E., both of which were used at various concs. from 0.3× to 1×. Conditions of electrophoresis varied from 25V o/n to 200V for 2hrs.

Polyacrylamide Gel Electrophoresis

Polyacrylamide gels were used both analytically and preparatively. In general, preparative gels were made and used in the same way as analytical gels except that they had larger wells. A stock solution 29:1 acrylamide:bis-acrylamide was stored at rt and diluted as required. Generally gels of 5% (w/v) were used and the running buffer was 1× T.B.E. Prior to pouring the gel solution is degassed then polymerisation is initiated by the addition of TEMED and ammonium persulphate.

Polyacrylamide/Urea denaturing gels and gradient gels for sequencing were prepared as described in the Amersham M13 Cloning and Sequencing handbook. The same method was used to prepare gels for the preparation of single stranded DNA probes for S-1 nuclease mapping studies, and the preparation of 20% acrylamide denaturing gels for analysis of end-labelled oligonucleotides.

Isolation of DNA Restriction Fragments

Restriction fragments were isolated from polyacrylamide or LMP agarose gels. Fragments were first visualised by ethidium bromide staining and subsequently excised by cutting the corresponding gel slice with a sterile scapel blade. The gel slice was placed into a dialysis bag containing 0.5mls 0.1× T.B.E. and electroeluted at 200V for 1hr. The DNA was released from the wall of the dialysis tubing by reversing the polarity for 30secs. The DNA solution was then transferred to a 1.5ml Eppendorf and any debris pelleted by centrifugation at 12,000rpm (M.S.E. microcentaur) for 10mins. The supernatant was transferred to a sterile Eppendorf and the DNA ethanol precipitated.

When DNA was recovered from normal agarose gels it was passed through a column of DE52 cellulose (equilibrated with 0.1× T.B.E.). The column was washed with 0.1× T.B.E. and the DNA solution passed through four times. The column was again washed with 0.1× T.B.E. and the DNA eluted

with 500μl elution buffer (1M NaCl, 50mM Tris-HCl, pH7.5, 1mM EDTA). DNA was recovered by ethanol precipitation.

Preparation of Competent Cells (JA221, C600 and JM101)

Cultures were grown in L.broth at 37°C until the OD_{650} reached 0.4 (JM101) to 0.6 (JA221 and C600). They were harvested in round bottom universal tubes (Nunc) by centrifugation at 4000rpm, 4°C for 10mins, the supernantant discarded and the cells resuspended in 0.1M $MgCl_2$ (JA221 only) to the original volume of culture and then repelleted. Cells were resuspended in 0.1M $CaCl_2$ to half the original volume and incubated on ice for 20mins. They were harvested as before and resuspended to 1/20 of the original volume of culture and stored on ice.

Transformation of Competent Cells (JA221 and C600)

The DNA to be used in transformation was diluted to 100μl in 1× SSC, and two subsequent 1/10 serial dilutions performed. 100μl DNA was mixed with 200μl competent cells and incubated on ice for 30mins, heat shocked at 42°C for 2mins, and returned to ice for 20mins. 1ml L.broth was added and the cells incubated with gentle shaking for 45mins at 37°C. Cells were harvested at 4000rpm, 4°C for 5mins, resuspended in 100μl bacterial buffer and plated onto an L.agar plate containing the appropriate selection.

Nick Translation of DNA (Rigby *et al*, 1977)

DNA was labelled using a nick translation kit (Amersham) according to the manufacturers instructions. Unincorporated nucleotides were removed by column chromatography through Sephadex G-50 pre-equilibrated in 1× T.E. 250μl fractions were collected and those containing labelled DNA were identified by Cerenkov counting. This gave an estimate of the specific activity of the labelled DNA probe.

Oligo-labelling of DNA Fragments by Hexanucleotide Primers (Feinberg and Vogelstein, 1984)

A known quantity of plasmid or phage DNA was cleaved with the appropriate restriction enzyme(s), to give the required fragment for labelling. The digest was run out on a 1% LMP agarose gel (BRL 5517). The gel was stained in ethidium bromide and the appropriate band excised carefully with the minimum amount of extraneous agarose. Each band was placed into a preweighed 1.5ml Eppendorf tube and DDW added at a ratio of 1.5mls H_2O/g agarose. The DNA conc. was estimated from the amount of DNA initially digested (50ng of DNA per labelling). The tube was then placed in a boiling water bath for 7mins to melt the agarose and denature the DNA (incubate at 65°C for 5mins). It was then kept at 37°C for 5mins prior to initiating the labelling reaction. The labelling reaction was performed using an oligo-labelling kit (Pharmacia) according to the manufacturers instructions. Unincorporated nucleotides were removed by column chromatography as decribed for nick translation.

5' End-labelling Oligonucleotides (Zoller and Smith, 1983)

200pmol oligonucleotide was lyophilised in a silanised Eppendorf then resuspended in 20μl DDW. Subsequently 3μl 1M Tris-HCl (pH8), 1.5μl 0.2M $MgCl_2$, 1.5μl 0.1M DTT, 3μl 1mM ATP (or DDW) and 1μl [γ-^{32}P]ATP were added and the contents mixed. The reaction was initiated by the addition of 4.5 units T_4 polynucleotide kinase and the mixture incubated for 45mins at 37°C. The tube was transferred to a 65°C waterbath to stop the reaction. Unincorporated nucleotides were removed by column chromatography as described for nick translation.

In Situ Plaque Hybridisation (Benton and Davis, 1977)

Nitrocellulose discs were carefully placed over cold agar plates containing phage plaques. The filters were allowed to wet, keyed to the

plate using a syringe needle and then lifted from the surface being careful not to remove any top agar. The phage were lysed and the DNA denatured by layering filters with the plaque side uppermost on a pad of Whatmann 3MM paper saturated in 0.5M NaOH for 5mins. The filters were then neutralised by incubating sequentially for 20secs each in 0.1M NaOH, 1.5M NaCl, twice in 0.5M Tris-HCl (pH7.5), 1.5M NaCl and finally in 2× SSC. They were then air dried and finally baked at 80°C for 2hrs. Hybridisation was performed as described for Southern blotting.

Colony Hybridisation (Grunstein and Hogness, 1975)

Prospective recombinant colonies were picked onto nitrocellulose discs layered on L.agar plates, containing an appropriate antibiotic selection, and incubated o/n at 37°C. Cells were lysed by layering the filters with colonies facing upward on a pad of Whatman 3MM paper soaked in 10% (w/v) S.D.S. for 3mins. The filter was then transferred to a similar pad saturated with 0.5M NaOH, 1.5M NaCl and incubated for 5mins to denature the DNA. Finally the filters were neutralised by transferring to a pad soaked in 1M Tris-HCl, 3M NaCl soln. for a further 5mins. The final step was repeated. The filters were then air dried and baked at 80°C for 2hrs. Hybridisation was performed as described for Southern Blotting.

Southern Blotting of Restriction Digests of DNA (Southern, 1975)

DNA samples were electrophoresed o/n (genomic DNA) at 25V or for 2hrs at 200V. The DNA was denatured by soaking the gel for 1hr in 1.5M NaCl, 0.5M NaOH and then neutralised for 30-60mins in 1M Tris-HCl (pH6.5), 1.5M NaCl. The gel was next soaked for 20mins in 20× SSC and placed on the transfer apparatus (Maniatis *et al*, 1980), ensuring no air bubbles were trapped between the gel and the filter paper of the wick. A sheet of nitrocellulose paper (soaked in 2× SSC) was carefully placed on top of

the gel and covered with a sheet of Whatman 3MM paper pre-soaked in 2× SSC. Six additional sheets of dry Whatman 3MM filter paper were then placed on top, followed by 4-6" of absorbant tissues and a 1kg weight. The transfer was carried out o/n using 20× SSC as the transfer buffer. The filter was baked at 80°C for 2hrs and was then ready for hybridisation.

Alternatively, the DNA was denatured for 30mins in 0.4M NaOH, 0.6M NaCl with gentle agitation and then placed inverted onto the wick which had been pre-wetted with the same buffer. GeneScreenPlus^{TM} (NEN) transfer membrane (Soaked in 0.4M NaOH, 0.6M NaCl) was placed over the gel followed by two sheets of wet (0.4M NaOH, 0.6M NaCl) and six sheets of dry Whatman 3MM paper respectively. Finally 2"-3" of absorbant tissues and a 0.5kg weight were placed on top and the transfer performed in a buffer comprising 0.4M NaOH and 0.6M NaCl. The transfer was performed for 16hrs after which the filter was neutralised for 15mins with agitation in 0.5M Tris-HCl (pH7), 1M NaCl. The filter was finally dryed at 55°C for 20mins.

Hybridisation of Southern Blots

Hybridisation was carried out in plastics bags. For nitrocellulose, the filters were pre-hybridised for 4hrs in 2× SSC, 0.5% (w/v) S.D.S., 5× Denhardt's and denatured Salmon sperm DNA (100μg/ml) in a final volume of 10ml at 65°C. Hybridisation was performed in a fresh but identical solution which also contained the denatured ^{32}P labelled DNA probe, at 65°C for 15 hrs. Filters were washed at 65°C in 0.5% (w/v) S.D.S. and 2× to 0.1× SSC with agitation prior to autoradiography

GeneScreenPlus^{TM} (NEN) was pre-hybridised for 4hrs in 5× SSPE, 0.25% (w/v) dried skimmed milk (Marvel^{RTM} Cadbury's), and denatured salmon sperm DNA (50μg/ml) for 6hrs in a final volume of 100ml at 65°C. Hybridisation was performed in 5× SSC, 1% (w/v) S.D.S., 10% (w/v) dextran

sulphate, denatured salmon sperm DNA (50μg/ml) and denatured ^{32}P labelled DNA probe in 10ml for 15hrs at 65°C. Filters were washed with agitation in 1-2% (w/v) S.D.S., 2× to 0.1× SSC at 65°C.

DNA Sequencing

DNA sequence determinations were carried out by the dideoxynucleotide chain-termination method of Sanger *et al* (1977).

The transformation of *E.coli* JM101 with the desired M13 phage DNA, the selection of recombinants, the preparation of ssDNA template DNA and the sequencing reactions with ^{32}P or ^{35}S labelled nucleotides were all performed as described in the Amersham M13 Cloning and Sequencing Handbook. However, Blou-gal (BRL) was substituted for X-gal for the selection of recombinant plaques.

2.6 Oligonucleotide Site Directed Mutagenesis

Reagents:- Solution A 0.2M Tris-HCl, pH7.5

0.1M $MgCl_2$

0.1M NaCl

0.01M DTT

Solution B 0.2M Tris-HCL(pH7.5)

0.1M $MgCl_2$

0.1M DTT

Solution C	Solution B	1μl
	10mM dCTP	1μl
	10mM dTTP	1μl
	10mM dGTP	1μl
	0.1mM dATP	0.5μl
	10mM ATP	1μl
	(α-^{32}P)dATP(410Ci/mmol)	1.5μl
	T_4 DNA ligase 2u/μl	1.5μl
	H_2O	2μl

Alkaline sucrose gradient stock solutions :-

X % (w/v) Sucrose

1M NaCl

0.2M NaOH

2mM EDTA

X = 20, 17.5, 15, 10, and 5.

autoclave and store at 4 °C.

Procedure A (Zoller and Smith 1983)

Alkaline sucrose gradient centrifugation

A 5% to 20% sucrose gradient was prepared for each sample in 5.5ml polycarbonate tubes (M.S.E. 1085). Step gradients were prepared by adding 1ml of each sucrose stock solution, starting with the 5% solution, and underlaying with increasing density solutions. Tubes were left at 4°C for 12hrs to linearise the gradients.

Mutagenesis

In a 1.5 ml Eppendorf tube 20pmol of 5' phosphorylated mismatched primer was annealed in 40-80 fold molar excess to M13 ssDNA template, with 1μl soln.A for 1.5 to 2 hrs at 55-60°C in a final volume of 10μl. Primer extension was initiated by the addition of 10.5μl soln.C and 2.5 units DNA polI (Klenow) to the annealing mix and incubated for 5mins at 23°C. 1μl cold 10mM dATP was added and incubation continued at 15°C for 20-24 hrs.

Unincorporated nucleotides were removed by the addition of 30μl H_2O and 50μl 1.6M NaCl/13% (w/v) $P.E.G._{6000}$. After incubation on ice for 15mins, precipitated DNA was pelleted for 5mins at 10,000rpm. The aqueous phase was withdrawn by pipetting, the pellet washed in 100μl 0.8M NaCl/6.5% (w/v) $P.E.G._{6000}$, re-pelleted for 30 secs and resuspended in 180μl T.E.

To the primer extended DNA, 20μl 2N NaOH was added and incubated for 5mins at 23°C. Samples were chilled on ice for 1min then layered gently onto sucrose gradients. Sucrose gradients were centrifuged at 37,000 rpm, 4°C, for 2hrs in a 6×5.5ml swing out rotor (M.S.E. 43127 126). Approximately 30 fractions per sample (5 drops/fraction) were collected, using a Gilson fraction collector (Microcol TDC80) and peristaltic pump (LKB Microperpex) and counted by Cerenkov counting (LKB 1215 Rackbeta). Fractions containing ds (CCC) DNA (lower half of gradient) were neutralised by the addition of approximately 50μl per fraction of 1M Tris-citrate (pH5.0) and stored at -20°C. 1μl, 5μl and 10μl aliquots were transfected into JM101.

Screening Mutants by Oligonucleotide Hybridisation

100 plaques were picked onto nitrocellulose discs (405316) on H plates then incubated overnight at 37°C. Colonies were denatured and neutralised as for colony hybridisation. Filters were pre-hybridised in

10ml 5× Denhardts, 0.5% (w/v) S.D.S., 6× SSC, for 1hr at 67°C. Filters were then washed once in 50ml 6× SSC, and then 4ml 5× Denhardt's, 6× SSC were added. 2-5×10^6 cpm [γ-^{32}P] ATP 5'-labelled oligonucleotide was added and incubation performed at 23°C for 1hr. Filters were washed 3× in 50ml 6× SSC at 23°C for 10mins. The stringency of hybridisation was increased by washing filters in 6× SSC for 1-10mins at progressively higher temperatures. Increments of 5-10°C were found to be most suitable. The melting temperature for each oligonucleotide was estimated by allowing 2°C and 4°C for each A/T and G/C base pairing respectively (Wallace *et al*, 1979). Filters were autoradiographed for 1-12 hrs after each wash. Positive colonies were picked from the master plates and ss and R.F DNA prepared for sequencing and cloning.

Procedure B (Kunkel, 1985)

Preparation of M13 Phage

One fresh plaque of the desired M13 recombinant phage was picked with a sterile 100μl micropipette and dispensed into a 1.5ml Eppendorf containing 1ml of Y.T. medium. The sample was vortexed until fully resuspended, then incubated at 70°C for 5mins to kill JM101 cells. The supernatant was then cleared by spinning for 5mins at 10,000g (M.S.E. Microcentaur) and transferred to a fresh 1.5ml Eppendorf tube. Phage stocks were used immediately to inoculate exponentially growing cells.

Preparation of Uracil Containing ssDNA Template

BW313 cells were grown at 37°C overnight in 2× Y.T. A 50ml culture was inoculated with 0.5ml o/n and incubated at 37°C for 1.5 to 2 hrs. 100ml 2× Y.T. was then inoculated with 5ml BW313 exponentially growing cells. Uridine (Sigma) was added to a final concentration of 0.25μg/ml and 100μl of the M13 phage preparation then added. Incubation was continued with vigorous shaking at 37°C for a further 6-8 hrs. The phage

were harvested and ssDNA prepared. The DNA was resuspended in 50-100μl T.E.

Mutagenesis

Primer annealing and extension reactions were carried out as described before with the following modifications. M13 ssDNA template was replaced by a uracil containing ssDNA template, and [α-^{32}P]dATP in soln.C was replaced by 1μl 10mM cold dATP. Primer extension and ligation was stopped by the addition of 20μl H_2O, and samples stored at -20°C. 5μl per sample was transfected into JM101. Plaques were picked and ss and R.F DNA prepared for sequencing and/or restriction analysis.

2.7 Preparation of Plasmid DNA

Cleared Lysate Method

Using a suitable selection, 200ml o/n cultures were grown and the cells harvested by centrifugation at 6000rpm, 4°C, for 5mins (M.S.E. 6x300ml rotor 43115-112). Cell pellets were resuspended in 3ml 25% (w/v) sucrose, 0.05M Tris-HCl (pH8.0), and treated with 0.5ml 10mg/ml lysozyme at 4°C for 15mins. After the addition of 1ml 0.25M EDTA, and a further 15mins at 4°C, the cells were lysed with 4ml Triton lysis buffer (2% (v/v) Triton X-100, 0.05M Tris-HCl, pH8.0, 0.0025M EDTA) and the lysate cleared by centrifugation at 18000rpm, 4°C, for 30mins (M.S.E. 8x50ml angle rotor 43114-143). To 7.4ml cleared lysate in 11.5ml polyallomer centrifuge tubes (Sorvall 03987), 7.1g CsCl was added and dissolved. Ethidium bromide was added (0.2ml 10mg/ml solution), and the tubes balanced with paraffin oil and sealed. Samples were centrifuged at 40,000rpm, 20°C, for 48hrs (Sorvall T-1270). Plasmid bands were harvested, after the removal of the upper chromosomal DNA band, using a syringe and needle from the side of the tube. Plasmid DNA was dialysed

o/n against 1× T.E., then extracted with an equal volume phenol (0.5M Tris-HCl pH8.0). The phenol phase was re-extracted, the aqueous phases pooled, extracted twice with an equal volume butan-1-ol (sat. H_2O), and ethanol precipitated in the presence 0.3M NaOAc at -70°C for 15mins. The DNA was resuspended in 1× T.E., the OD_{260nm} measured to estimate the conc., and then stored at -20°C.

Preparation of Plasmid DNA by Alkali Lysis (Birnboim and Doly, 1979)

Using a suitable antibiotic selection, 50ml cultures were grown o/n at 37°C. Cells were pelleted in round bottom universals (Nunc) at 4,000rpm, 4°C, for 15mins, resuspended in 2ml solution 1 (50mM glucose, 25mM Tris-HCl pH8.0, 10mM EDTA, 4mg/ml lysozyme), transferred to 30ml Oakridge tubes and left on ice for 10mins. Cells were lysed by the addition of 4ml cold solution 2 (0.2M NaOH, 1% (w/v) S.D.S.) and incubation on ice continued for 5mins prior to the addition of 3ml cold solution 3 (29ml glacial acetic acid, 61ml H_2O, adjusted to pH4.8 with 10M KOH). After a further 10mins on ice the cell debris was pelleted at 15,000rpm, 4°C, for 15mins (M.S.E. 8×50ml angle rotor 43114-143). Approximately 9ml supernatant was transferred to a round bottom universal and the DNA precipitated in 20ml ethanol. The DNA was then pelleted at 4,000rpm, 4°C, for 15mins (M.S.E. Chilspin), then washed in 1ml 80% (w/v) ethanol and repelleted. The DNA was resuspended in 500μl 10× T.E. and 50μl 3M NaOAc and transferred to a 1.5ml Eppendorf tube. It was then phenol extracted and chloroform extracted twice and ethanol precipitated. The DNA was resuspended in 200μl 1× T.E. containing 10μg/ml pancreatic RNase (Sigma) and stored at -20°C.

2.8 Preparation of Bacteriophage λ DNA

Solutions for step gradients

CsCl stock: 195g CsCl in 105 ml phage buffer

Density (gcm^3)	CsCl stock (ml)	Phage buffer (ml)
1.7	23.3	10
1.5	16.65	13.35
1.3	12	19.03

Plating cells

Plating cells were prepared from an o/n culture of *E.coli*. A fresh culture was inoculated by the addition of 200μl o/n to 20ml L.broth and incubated at 37°C for 5hrs. Cells were pelleted at 4000rpm (M.S.E. Chilspin), 4°C, for 10mins, and resuspended in 20ml 0.01M $MgSO_4$. These could be stored for one week at 4°C.

Phage Titres

Serial dilutions of phage were performed in phage buffer. In a phage tube, 100μl phage dilution were adsorbed to 200μl plating cells for 15mins at 4°C. 3ml BBL top agar (45-50°C) was added to the phage tube, mixed, and poured onto BBL agar plates on a level surface. Plates were left to set for 15mins, then incubated o/n at 37°C. Alternatively plates were poured with BBL top agar containing plating cells only. Aliquots of 10μl from phage dilutions were spotted onto solidified top agar and left to dry adjacent to a bunsen burner, prior to o/n incubation at 37°C.

Preparation of Phage - Plate Lysate

One to three fresh plaques were picked into 1ml phage buffer, 50μl $CHCl_3$ added and vortexed thoroughly. Three plate lysates and a minus phage control were prepared, by adding 500μl, 50μl and 5μl phage to 200μl plating cells. After 15mins, 4°C, 3ml BBL top agar was added. This was poured onto fresh, thick, moist L agar plates on a level surface, and subsequently incubated upright at 37°C for 8-12hrs. To harvest phage, 3ml L.broth was added to the plates, and the top agar scraped into a glass universal. After thorough vortexing, lysates were cleared by centrifugation at 4000rpm (M.S.E. chilspin), 4°C, for 10mins. Supernatants were transferred to a sterile glass bijou bottle, phage were titred, and stored over 50μl $CHCl_3$, 4°C.

Preparation of Phage - Liquid Lysate (Blattner *et al*, 1977)

A fresh culture of *E.coli* cells was inoculated by the addition of 1ml o/n to 200ml L. broth supplemented with 0.001M $MgCl_2$ in a 2L flask. The culture was incubated at 37°C with vigorous shaking to $O.D_{650nm}$ 0.4 (approx. 2×10^8 cells/ml). Phage prepared from the plate lysates were added to give an M.O.I. of approximately 2. The $O.D_{650nm}$ was followed for 3-6 hrs until cells lysed whereupon 1ml $CHCl_3$ was added and shaking continued for a further 10mins. Lysates were clarified at 6,000rpm (M.S.E. 6x300ml rotor 43115-112), 4°C, for 10mins ,and phage titred.

CsCl Step Gradients

From a stock solution of CsCl, three solutions of the concs. shown above were prepared. Step gradients were prepared in 14ml polycarbonate thin wall tubes (M.S.E. 1284) by adding 3ml $1.3gcm^3$ solution, and subsequently underlaying with 3ml each, of the two higher conc. solutions.

Preparation of Phage DNA

Phage from the liquid lysate were precipitated o/n by the addition of 4g NaCl and 20g P.E.G.$_{6000}$ per 200ml supernatant. Phage were then pelleted at 5000rpm (M.S.E. 6×300ml rotor 43115-112) for 5mins, 4°C, and the pellet resuspended in 1ml phage buffer. Bacterial DNA was degraded by the addition of DNase (10μg/ml) and incubation for 1-2hrs at rt. Phage were pooled from 200ml lysates, and 3-5ml per 14ml tube of phage sample carefully layered onto the step gradients. Gradients were centrifuged at 24,000rpm, 20°C, for 4hrs (M.S.E. 6×14ml swing out rotor 43127-111). Phage bands were collected through the side of the tubes with a syringe needle and dialysed against 1× T.E. for 1hr at 4°C. Residual RNA was digested with pancreatic RNase (20mg/ml) for 1hr at rt, during dialysis. Dialysis was further continued at 37°C for 1hr in the presence of 1mg/ml proteinase k (Sigma) against proteinase k buffer. The dialysate was extracted twice with an equal volume 0.5M Tris-HCl (pH8.0) saturated phenol. Phage DNA was dialysed at 4°C for 4hrs with three changes T.E. buffer, the O.D.$_{260nm/280nm}$ ratio measured to estimate the conc. of DNA and λ DNA, concentrated by ethanol precipitation if necessary. DNA was stored at -20°C.

2.9 Preparation of a Genomic Library

Solutions for NaCl gradients:

Stock solution	5M NaCl	in 1xT.E.	
	1.25M NaCl	"	2.5ml
	2.5M NaCl	"	4.0ml
	3.75M NaCl	"	4.0ml
	5.0M NaCl	"	2.5ml

Packaging buffers:

Buffer A	1M Tris-HCl(pH8.0)	20μl
	1M $MgCl_2$	3μl
	β-Mercaptoethanol	0.5μl
	0.1M EDTA,pH7.0	10μl
	H_2O	966.5μl
Buffer B	1M Tris-HCl(pH7.5)	6μl
	0.1M Spermidine 3HCl(pH7.0)	300μl
	0.2M Putrescine 2HCl(pH7.0)	300μl
	0.1M ATP(pH7.0)	150μl
	β-Mercaptoethanol	2μl
	H_2O	224μl
Tris-sucrose	10% (w/v) sucrose	
	50mM Tris-HCl, pH7.5	
Lysozyme	2mg/ml in 0.25M Tris-HCl, pH7.5	

Preparation of λ Packaging Extracts (Hohn *et al*, 1979)

A freeze thaw lysate (F.T.L) was prepared from BHB2688 cells in the following manner. An o/n of the λ lysogen BHB2688 was grown at 28°C in L.broth. Three 2L flasks were inoculated by the addition of 5ml o/n to 250ml L.broth, and incubated at 28°C. At $O.D._{630nm}$ 0.6, a 1ml aliquot of cells were removed, and a drop of $CHCl_3$ added to test for induction of the λ lysogen. If the cells lyse, determined by rapid clearing of the culture medium, then the lysogenic phage has been induced. If the culture remains cloudy, as expected, then proceed as descibed below. Flasks were transferred to 37°C, 250ml L.broth prewarmed to 62°C added, and

incubation continued for 1hr with vigorous shaking. Cultures were cooled on ice and a 1ml aliquot removed for the induction test. Cells were harvested at 9,000rpm, 4°C, for 10mins (6×500ml rotor M.S.E. 43115-113). Pellets were resuspended gently in 0.5ml cold Tris-sucrose, and pooled into 10ml polyallomer centrifuge tubes. 75μl fresh lysozyme (2mg/ml; Calbiochem) was added, mixed gently, and cells frozen in liquid N_2. Extracts were thawed partially at room temperature, then completely on ice. 75μl buffer B was added, mixed thoroughly, and extracts centrifuged at 35,000rpm, 4°C for 35mins (8×35ml angle rotor) to clarify. The supernatant was then dispensed in 100μl aliquots into screw cap ampoules and frozen rapidly in liquid N_2. Extracts were stored at -70°C for several months.

On a separate day 3×250ml cultures of the lysogenic strain BHB2690 were incubated in L.broth at 28°C and induced as described for BHB2688. Cells were harvested as described above, resuspended in 0.5ml buffer A, pooled and transferred to a polypropylene tube. The suspension was diluted with 2.6ml buffer A, and sonicated on ice without foaming with 15 bursts of 5secs duration with 30secs interval. Cell debris was pelleted at 6,000rpm, 4°C, 10mins (M.S.E. chilspin), 50μl aliquots dispensed into screw cap ampoules, and frozen rapidly in liquid N_2. Sonicated extracts (S.E) were stored for several months at -70°C.

The packaging reaction was initiated by the successive addition of the following components to a 0.5ml Eppendorf tube :

Buffer A	7μl
DNA(0.1μg)	1μl
Buffer B	1μl
S.E	3.5μl
F.T.L.	5μl

Eppendorfs were centrifuged for 3secs, and incubated at rt for 90mins. Packing reactions were diluted with 230μl phage buffer and the phage plated. Packaged phage not plated were stored at 4°C over $CHCl_3$.

Continuous NaCl Gradients

Gradients were prepared by the addition of NaCl stock solutions to polycarbonate tubes (M.S.E. 1284). The 1.25M NaCl solution was added first and other stock solutions added by underlaying with the volumes indicated above. Gradients were left to diffuse o/n at rt. A control tube was prepared in parallel and a refractometer used to verify the linearity of the gradient.

Preparation of λL47 vector arms

100μg λL47 DNA was incubated o/n with *Bam*HI$_{(50units)}$, *Sal*I$_{(50units)}$, in 150μl final volume. Complete digestion was monitored by analysing an aliquot of DNA by agarose gel electrophoresis. Digested DNA was ethanol precipitated and resuspended in 100μl 1× T.E. λL47 vector arms were fractionated by applying the digested 100μg DNA to a continuous NaCl gradient, and centrifuging at 40,000rpm, 20°C, for 3hrs (6×14ml titanium swing out rotor M.S.E. 43127-111). Gradients were collected through a silanised capillary tube, using a peristaltic pump and Gilson fraction collector. 250μl fractions (14 drops) were collected, 250μl 1× T.E. added, and samples ethanol precipitated at -20°C o/n. Fractions were resuspended in 40μl 1× T.E., 4μl aliquots analysed by agarose gel electrophoresis, and samples containing vector arms were pooled and stored at -20°C.

*Sau*3A Partial Digests of Human Genomic DNA

i) Pilot digests

10μg genomic DNA was added to 10μl 10× *Sau*3A buffer in an Eppendorf tube, the volume adjusted to 100μl with H_2O, and incubated at 37°C for 15mins. 1 unit *Sau*3A was added and incubation at 37°C continued. Aliquots of 5μl were removed at 0, 5, 10, 15, 20, 25, and 30 min intervals into 20μl ice cold 10mM EDTA in 1× T.E. buffer, and incubated at 65°C for 5mins. A 5μl aliquot was removed from each sample, 10μl agarose beads and 5μl H_2O added and samples electrophoresed at 40V in 1× T.B.E. through a 0.2% agarose gel.

ii) Full Scale Digests

Conditions determined for partial digestion of human DNA with *Sau*3A were scaled up to digestion of 100μg DNA. After digestion, aliquots were analysed on a 0.2% agarose gel. The DNA was ethanol precipitated, resuspended in 100μl 1× T.E. and applied to a NaCl gradient. Gradients were centrifuged and collected as described previously. Aliquots of fractions were analysed on a 0.2% agarose gel to assess average DNA size.

Ligation of Insert DNA to Vector Arms

The ability of isolated vector arms to re-ligate was analysed. A ligation mix and controls were prepared as described below:

Vector arms	1μl	1μl	1μl
10× lig. buffer	2μl	2μl	2μl
T-4 ligase	2μl	2μl	-
Human DNA	1μl	-	-
H_2O	14μl	15μl	17μl

Ligations were incubated at 15°C for 3hrs, and the DNA analysed by agarose gel electrophoresis. Full scale ligations were performed using 5-10μg vector arms and 1-2μg *Sau*3A digested human DNA in 70μl final volume. Ligations were incubated o/n at 15°C, and a 5μl aliquot of the ligation mix analysed by agarose gel electrophoresis. Ligation mixes were stored at -20°C.

Plaque Purification

Plaque purification was performed by several rounds of plaque hybridisation. Subsequently four hybridisation positive plaques were picked and titred for each clone. Ten randomly chosen plaques were toothpicked from each of the four plates and plaque hybridisation performed to confirm that they were all recombinants.

2.10 Tissue Culture

Cell Culture

L-929 cells (Gibco) were grown in monolayer in F.G.M. consisting of DMEM supplemented with 10% foetal calf serum, glutamine(2mM), sodium pyruvate(2mM), and penicillin (5units/ml)/streptomycin (5μg/ml) at 37°C in a 5% CO_2 atmosphere. When confluent (twice weekly) cells were sub-cultured, washing twice in versene prior to trypsinising. Cells were

resuspended in F.G.M. and reseeded at 0.7-1x10^6 cells per 80cm^2 flask.

Freezing Down

Cells were trypsinised, counted, and pelleted at 750rpm (M.S.E. Mistral, 4x750ml rotor 43124-129), 4°C, for 5mins. Cell pellets were resuspended in F.C.S. containing 10% DMSO, at 4-5x10^6 cells/ml. 1ml aliquots were dispensed into screw cap ampoules (Nunc) and left o/n at -70°C. Ampoules were subsequently stored in liquid N_2.

Transfection of Mammalian Cells (Wigler et al, 1979)

One day prior to transfection, L-929 cells were seeded at 1x10^6 cells per 9cm tissue culture petri-dish (Nunc). On the following day the medium was changed for 9ml of freshly prepared F.G.M. 4hrs prior to transfection. A DNA/$CaCl_2$ solution was prepared from 1μg sterile pTCF (Grosveld *et al*, 1982) dissolved in 1mM Tris-HCl [pH7.9]/0.1mM EDTA, and 250mM $CaCl_2$ (BDH Analar). The DNA conc. was adjusted to 40μg/ml with salmon sperm DNA (Sigma). 1ml DNA /$CaCl_2$ solution was then added dropwise to an equal volume 2x HBS (280mM NaCl, 50mM Hepes, 1.5mM sodium phosphate [pH7.1]). As the DNA was added, bubbles were introduced through a sterile 1ml pipette. The $CaPO_4$/ DNA precipitate was left to form for 30-40mins at rt. The precipitate was then mixed by pipetting, and 1ml added per petri-dish. This was performed rapidly to avoid the pH of the medium changing. Cells were incubated for 24hrs, then washed with 2x10ml DMEM, and re-fed with 10ml F.G.M. Selection with 400μg/ml G418 (Geneticin BRL) was started 24hrs later. The medium was changed twice weekly. After 3 weeks colonies were counted, trypsinised and pooled for subsequent studies. Cells were maintained under continuous selection pressure.

Transient Transfection Assays

The same procedure described above was employed to transfect L-929 cells with 10μg plasmid DNA. Normally cells were harvested 48hrs post-transfection and extracts prepared for CAT assays.

PolyrI.rC Induction of L-929 Cells (Trapman, 1979)

To induce interferon synthesis, L-929 cells were seeded at 1×10^6 cells per 9cm tissue culture petri-dish (Nunc), and grown to confluence. Cells were re-fed and incubated for a further 3 days prior to induction. Confluent monolayers were washed with 2×10ml DMEM and 3ml DMEM added. 800μg/ml DEAE dextran (mw 5×10^5, Pharmacia), was added to the cells and mixed thoroughly prior to addition of 20μg/ml polyrI.rC (Pharmacia). L-929 cells, induced and non-induced were incubated for 8hrs then the monolayers were washed in 2×10ml DMEM to remove residual DEAE dextran/polyrI.rC, re-fed with F.G.M. supplemented with 2.5% F.C.S. and incubated for a further 16hrs.

Extraction of High mw. DNA from Tissue Culture Cells (Pellicer *et al*, 1978; Maniatis *et al*, 1982)

Cells were grown to confluence, harvested with a rubber policeman, and resuspended in 10ml (10^7-10^8 cells) ice cold PBS. Cells were then pelleted at 750rpm, 4°C, 5mins (M.S.E. Mistral 4×750ml rotor 43124-129), resuspended in 1ml hypotonic buffer (10mM Tris-HCl pH7.5, 10mM NaCl, 3mM $MgCl_2$), and left on ice for 5mins. The cell suspension was lysed by Dounce homogenisation (Jencons), and the nuclei pelleted at 7000rpm, 6°C, for 5mins (M.S.E. 4×50ml swing out rotor 43124-705) in a 15ml corex tube. The nuclei were resuspended in 1ml hypotonic buffer/Triton-×100 (0.2% [w/v]), and Dounce homogenisation and centrifugation steps repeated. Nuclei were resuspended in 1ml T.N.E. buffer (10mM Tris-HCl [pH7.9], 0.4M NaCl, 2mM EDTA), and lysed by the addition of S.D.S. to 0.5% (w/v).

Extracts were incubated on ice for 10mins, proteinase k added (200μg/ml), and incubation continued o/n at 37°C with gentle mixing. DNA was extracted with an equal volume of phenol (0.5M Tris-HCl, [pH8.0]):chloroform:isoamyl-alcohol (25:24:1), the organic phase removed and re-extracted with an equal volume of T.N.E. buffer and the aqueous phases pooled and re-extracted with an equal volume chloroform:isoamyl-alcohol (24:1). DNA was precipitated by the addition of 2× volume ethanol slowly to the aqueous phase. DNA was spooled out quickly with a looped glass rod into 1ml T.E. buffer. The DNA was dissolved and left to dialyse o/n against T.E. The OD_{260nm} was measured to estimate the DNA conc. and the DNA stored at -20°C.

Total RNA Extraction From Tissue Culture Cells (Auffray and Rougeon, 1980)

Precautions were taken to ensure that all glassware, solutions, and plastic ware were RNase free. All solutions were prepared using baked spatulas, in baked glassware (220°C for 2hrs) and then autoclaved. Glass Corex tubes, and Eppendorf tubes were siliconised, the glassware was subsequently baked, and Eppendorf tubes autoclaved, prior to use. Cells were grown to confluence, harvested and washed in PBS as described for DNA preparation above. The cell pellet was resuspended in 7.5ml 6M urea, 3M LiCl, transferred to 13ml Sarstedt tubes (60.541) and homogenised in a Polytron (Kinematica) at full speed for 2mins. The RNA was precipitated o/n on ice, and pelleted at 10,000g, 4°C, for 20mins (M.S.E. 4×50ml swing out rotor 43124-705). The RNA pellet was resuspended in 500μl buffer III (10mM Tris-HCl pH7.6, 0.5% (w/v) S.D.S.) and extracted with an equal volume chloroform:isoamyl-alcohol (24:1). The organic phase was re-extracted with an equal volume of buffer III, the aqueous phases mixed, LiOAc added to 0.2M, and RNA precipitated by the addition of 2 volumes of cold ethanol, o/n at -20°C. RNA was pelleted at 10,000g, 4°C,

20mins (M.S.E. 43124-705), washed in 70% cold ethanol, re-pelleted, and resuspended in buffer III. The $OD_{260/280nm}$ was measured and the RNA stored at -70°C.

Protein Determination (Lowry *et al*, 1957)

A : 2% (w/v) Na_2CO_3 in 0.1N NaOH

B_1: 1% (w/v) $CuSO_4$

B_2: 2% (w/v) Sodium or Potassium Tartrate

C : Mix 50ml A with 0.5ml B_1 and 0.5ml B_2 (Fresh)

D : Folin-Ciocateu Phenol Reagent 143ml + 250ml DDW

An aliquot of the desired cell extract (usually 5-10μl) was diluted to 500μl in DDW, 2ml reagent C added, mixed and allowed to stand at rt for at least 10mins. 0.2ml reagent D was then added and mixed immediatley. Incubation at rt was continued 30min and then the $O.D_{750nm}$ read.

Prepare standards with various concs. BSA, usually 10μg to 160μg.

Chloramphenicol Acetyl Transferase (CAT) Assays (Shaw, 1979; Gorman *et al*, 1982)

Preparation of Cell Extracts

Transiently or stably transfected L-929 cells were grown to confluence in 9cm vented tissue culture petri-dishes (Nunc) in F.G.M. with suitable selection (G418 400μg/ml for stably transfected cell lines). The F.G.M. was discarded, the cells washed twice in 10ml ice cold P.B.S and the monolayer harvested using a rubber policeman in 1ml ice cold P.B.S. The cell suspension was then transferred to a 1.5ml Eppendorf and pelleted for 30secs at 1000rpm in a microfuge (M.S.E. Microcentaur). The supernantant was discarded and the cells resuspended in 100μl 0.25M Tris-HCl (pH7.8). Cells were lysed by two cycles of rapid freeze thawing

and subsequent sonication (M.S.E. Soniprep) for 10 seconds at full power using a titanium exponential microprobe (end diameter 3mm; M.S.E. 38121-114). The cell debris was pelleted for 15mins at 12000 rpm in a pre-cooled microfuge (in cold room), the supernatant transferred to a clean 1.5ml Eppendorf and stored at -20°C.

CAT Assays

CAT assays were performed in 1.5ml Eppendorfs. Each reaction contained 100μl 0.25M Tris-HCl [pH7.5], 5μl ^{14}C-chloramphenicol (0.2μCi; N.E.N), 5μl to 55 μl cell extract and DDW to a final volume 160μl. Each was pre-incubated at 37°C for 5mins prior to starting the reaction by the addition of 20μl freshly prepared 4mM acetyl-CoA (Pharmacia). Assays performed up to eight hours were supplemented by the addition of 1μl freshly prepared 80mM acetyl-CoA every hour since the coenzyme is unstable.

Reactions were stopped by the addition of 1ml ice cold ethyl acetate (B.D.H. Analar) and tubes mixed vigourously to extract ^{14}C-chloramphenicol from the aqueous phase. The two phases were separated by centrifugation for 1min at 12,000rpm (M.S.E. Microcentaur) and the organic phase carefully transferred to 1.5ml topless Sarstedt (72/696) tubes. Samples were vacuum dried in a Savant Speedvac for 30mins to 1hr and subsequently resuspended in 10μl ice cold ethyl acetate.

Ascending Thin Layer Chromatography (T.L.C.)

All 10μl of the samples from CAT assays were carefully spotted onto activated (pre-baked at 100°C for 15 to 30mins to remove hydroxyl groups) silica gel T.L.C. plates (Kodak Chromagram). The ^{14}C subtrate, mono- and di-acetylated products of the CAT reaction were separated by ascending T.L.C. in 95:5 chloroform:methanol in a pre-saturated T.L.C. tank. Activation of the silica gel T.L.C. plates and pre-saturation of tanks

ensured the even running of samples. The solvent front was allowed to migrate 10cms after which the T.L.C. plates were air dried and the ^{14}C-chloramphenicol substrate and derivatives visualized by autoradiography.

Quantitative CAT Assays

Scintillation Fluid Cocktail:

Beckman CP 2L

Beckman MP 2L

Propan-1-ol 500ml

Glacial Acetic Acid (B.D.H. Aristar) 50ml

Assays were performed as described above except 40μl aliquots were removed from the reaction mix at specified intervals into 1ml ice cold ethyl acetate. T.L.C. and autoradiography were performed as described above. ^{14}C-chloramphenicol substrate (S) is acetylated by CAT to give ^{14}C-chloramphenicol-3-acetate (P). The spots corresponding to S and P were located by realigning the autoradiograph obtained with the T.L.C. plate. Spots were cut out of the T.L.C. plates, the silica gel dissolved in the scintillation cocktail and the ^{14}C present quantitated by scintillation counting (Beckmann LS 3800). The percentage acetylation of substrate to product by CAT was calculated using the formula:

$$\% \text{ acetylation} = \frac{Pcpm}{Scpm + Pcpm} \times 100$$

The percentage acetylation at different time points was plotted against time. The linear gradient was calculated, giving an estimate of

the CAT enzyme reaction rate which when corrected for total cell protein allowed a relative estimate of CAT production in independent cell extracts to be made, expressed as % acetylation/min/μg of total cell protein.

Statistical Analysis

To compare the means of two populations:

The sum of the squares of population 1 is: $\Sigma(x_1 - \bar{x}_1)^2 = \Sigma x_1^2 - \frac{(\Sigma x_1)^2}{n_1}$

The sum of the squares of population 2 is: $\Sigma(x_2 - \bar{x}_2)^2 = \Sigma x_2^2 - \frac{(\Sigma x_2)^2}{n_2}$

The pooled sample varience is: $S_p^2 = \frac{\Sigma (x_1 - \bar{x}_1)^2 + \Sigma (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}$

The random variable is: $t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{S_p^2 (1/n_1 + 1/n_2)}}$

when $t \geq t_{0.05/2}$ for $n_1 + n_2 - 2$ degrees of freedom*

or $t \leq -t_{0.05/2}$ for $n_1 + n_2 - 2$ degrees of freedom*

then it is 95% certain that the two populations have the same mean.

where: x_1 and x_2 are values in population one and two respectively

$\bar{x}_1$ and $\bar{x}_2$ are the mean of population one and two respectively

n_1 and n_2 are the sample size of population one and two respectively

* determined from t-distribution tables

Preparation of Single Stranded Probes for S-1 Nuclease Mapping

M13 ssDNA template was annealed to the universal primer as for sequencing studies. Primer extension was initiated by the addition of 2 U DNA PolI (Klenow) to the reaction mix prepared as shown below:

Primer annealed template DNA	10μl
0.5mM dCTP	1μl
0.5mM dGTP	1μl
0.5mM dTTP	1μl
20μCi dATP (450Ci/mmol	2μl

The reaction was incubated for 30mins at rt, 2μl chase was added and incubated for a further 15mins at rt. DNA was digested, usually with *Eco*RI, for one hr at 37°C after which 10μl formamide dye mix was added, the sample denatured for 5mins in a boiling water bath and immediately loaded onto a 5% polyacrylamide/urea denaturing gel. The newly synthesised single stranded DNA fragment was located by autoradiography (≈ 5mins), excised and electoeluted from the gel. The probe was ethanol precipitated, resuspended in S-1 hybridisation buffer and counted (≈ 1×10^8) prior to immediate use.

S-1 Nuclease Mapping (Weaver and Weissmann, 1979; Favaloro *et al* 1980)

5× hybridisation buffer: 2M NaCl
0.2M Pipes, pH6.4
5mM EDTA

S-1 digestion buffer: 0.25M NaCl

30mM NaOAc pH4.6

1mM $ZnSO_4$

0.2μg/ml heat denatured Salmon sperm DNA added to ice cold S-1 buffer prior to use.

S-1 mapping was carried out typically with 25μg to 50μg total RNA. The RNA was lyophilised in a sterile siliconised 1.5ml Eppendorf tube. RNA was thoroughly resuspended in 20μl 1× hybridisation buffer / 50% (v/v) recrystalised formamide (Maniatis *et al* 1980) containing 20,000dpm of probe. A drop of paraffin oil was overlayed and the samples were subsequently incubated at 85°C for 15mins prior to annealing o/n at 52°C. 400μl ice cold S-1 digestion buffer was added to the samples which were immediately placed on ice. S-1 digestion was performed by the addition 200 units S-1 nuclease (Amersham) for 40mins at 30°C. Samples were phenol extracted, chloroform extracted and ethanol precipitated. Samples were then resuspended in 4μl formamide dye mix and 1μl DDW, denatured for 3mins in a boiling water bath and immediately loaded onto a sequencing gel. S-1 nuclease protected fragments were visualised by autoradiography.

CHAPTER 3

CONSTRUCTION OF A HUMAN GENOMIC LIBRARY: THE IDENTIFICATION AND ISOLATION OF HUMAN IFNα CHROMOSOMAL GENES

3.1 Introduction

The initial objective of these studies was to investigate the structure/function relationship of IFNα gene promoters. Thus, as a preliminary step, the first part of this work was primarily concerned with obtaining promoters which could eventually be used for this purpose. Although human IFNα chromosomal genes had been isolated previously, there were several potential advantages in isolating IFNα gene sequences from a new library to obtain the promoter sequences for regulatory studies. Firstly, the human IFNαs are encoded by a large gene family and so it was not certain that all the members had been identified. Secondly, at the outset of this work these genes had only been isolated from a single genomic library and thus there was the prospect of isolating allelic variants. Finally, the limited number of probes which had been used in screening genomic libraries meant that distantly related members of the family may exist which would possibly be identified using an alternative probe.

Genomic libraries prepared in λ*charon*-4A from foetal DNA (Maniatis *et al*, 1978), cosmid pHC79-2cos/tk from placental DNA (Lund *et al*, 1984) and cosmid pJB8 from placental and buffy coat DNAs (Todokoro *et al*, 1984) have been used for the isolation of human IFNα genes. The libraries were screened by *in situ* plaque hybridisation using cDNAs containing the human IFNα_1 (Nagata *et al*, 1980a; Todokoro *et al*, 1984) and LeIF-A (Lawn *et al*, 1981a) gene sequences or 5' end-labelled synthetic 17bp oligonucleotide probes (Torozynski *et al*, 1984). From these studies 23 distinct loci, containing IFNα related sequences have been isolated and characterised (Henco *et al*, 1985). Furthermore, two novel IFNα cDNA clones, GX-1 (Sloma *et al*, 1983) and LeIF-F (Goeddel *et al*, 1981) have been characterised, but their genomic equivalents not described (Henco *et al*, 1985).

At the outset of this work, many of the above genes had not been characterised. This Chapter describes the construction of a genomic library, prepared in λL47 (Loenen and Brammar, 1980) from human placental DNA (kindly provided by Dr.A. Jeffreys), constituting 300,000 recombinant phages, from which four IFNα gene related clones were identified.

3.2 Construction of a Human Genomic Library in λL47

Coliphage λL47 was chosen for the construction of a human genomic library because it is a versatile replacement vector with a relatively large cloning capacity for foreign DNA fragments ranging from 5kb to 22kb with cohesive termini complementary to *Eco*RI, *Hin*dIII or *Bam*HI restriction sites (Fig.3.1A). Recombinant derivatives of λL47 lack phage specific recombination genes, *red* and *gam*, and hence are viable on *rec*$^+$ hosts only (Stahl *et al*, 1975). The presence of a *chi* site (Murray *et al*, 1977) within the phage genome permits efficient utilisation of the host *rec*A system, thus increasing recombinant phage viability. A valuable property of recombinant λL47 phage is that because they lack *red* and *gam* genes they will grow on P2 lysogenic strains of *E.coli* - the so called *spi*$^-$ phenotype. Because of this property, parental *red*$^+$ *gam*$^+$ λL47 phage can be readily eliminated from recombinant libraries. The viability of only recombinant λL47 phage on bacteriophage P2 lysogenic strains of *E.coli* (Loenen and Brammar, 1980) provides a convenient selection for the identification of recombinant phage.

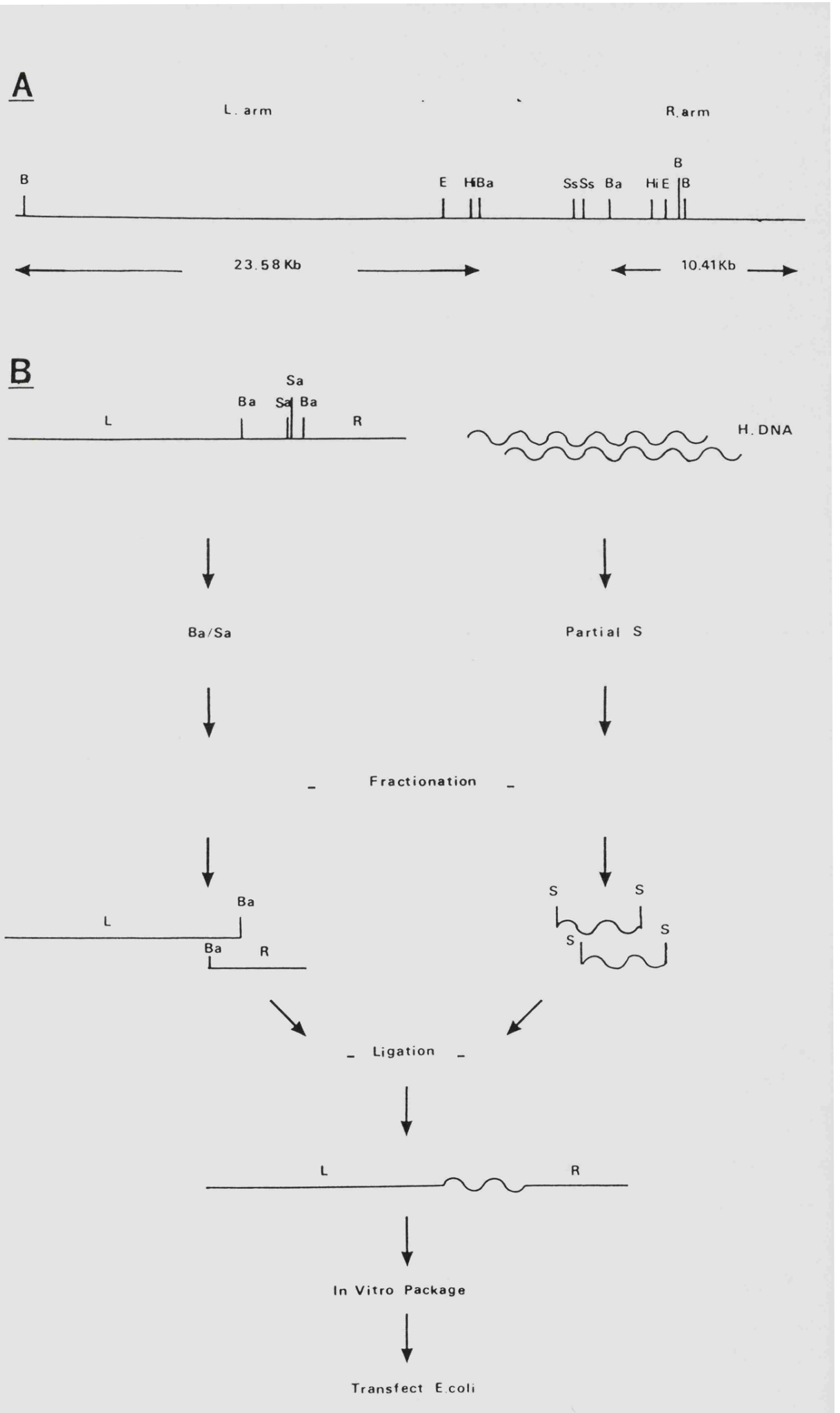
The use of λL47 for the preparation of a mammalian genomic library has been described previously (Jeffreys *et al*, 1982). The procedure employed here to construct a human genomic library is illustrated in Figure 3.1B. Random fragments of DNA were generated by partial digestion with *Sau*3A. This enzyme has a tetranucleotide recognition sequence which occurs frequently within the genome (theoretically every 256bp). Digestion of λL47 DNA with *Bam*HI produces complementary cohesive termini to those generated by *Sau*3A. Thus, size-fractionated, *Sau*3A partially digested DNA can be cloned into the *Bam*HI site of λL47.

FIGURE 3.1 A:- Partial restriction map of λL47 (Loenen and Brammar, 1980)

B:- Strategy for the construction of a human genomic library

B: *Bgl*II, Ba: *Bam*HI, E: *Eco*RI, Hi: *Hin*dIII, S: *Sau*3A, Sa: *Sal*I, Ss: *Sst*I.

L: left arm, R: right arm.


A
L. arm
R. arm
B
E
HiBa
SsSs
Ba
HiE
B
23.58Kb
10.41Kb
B
Sa
Ba
Sa
Ba
L
R
H. DNA
Ba/Sa
Partial S
Fractionation
Ba
L
Ba
R
S
S
S
S
Ligation
L
R
In Vitro Package
Transfect E.coli

Preparation of λL47 vector arms

In order to increase the proportion of recombinants during the library construction, the left and right arms of *Bam*HI digested λL47 DNA were purified by fractionation on a continuous NaCl gradient (Chapter 2). This procedure minimises the likelihood of the λL47 "stuffer" fragment competing with the genomic DNA during the ligation reaction used to generate recombinant phage.

λL47 DNA was prepared as described in Chapter 2 and its structure (Fig.3.1A) verified by agarose gel electrophoresis following digestion with *Eco*RI, *Bam*HI and *Hin*dIII, each of which generated the anticipated DNA fragments. However, salt gradient fractionation of *Bam*HI, which generated λL47 vector arms, was found to give only poor resolution of the 10.41kb right arm and 6.64kb "stuffer" fragments. Thus, prior to fractionation the vector DNA was also digested with an additional enzyme, *Sal*I, which cleaves the 6.64kb fragment generating 4.45kb, 1.79kb and 0.35kb fragments. Aliquots of the fractionated DNA were analysed by agarose gel electrophoresis revealing that satisfactory separation of discrete phage arms and "stuffer" fragments had been achieved (Fig.3.2).

The fractions identified as containing the 23.58kb and 10.41kb left and right vector arms and lacking detectable internal DNA fragments, were pooled and stored at -20°C. Their suitability for cloning purposes was established by demonstrating their ability to religate. Ligation was assessed by agarose gel electrophoresis.

Preparation of Human DNA

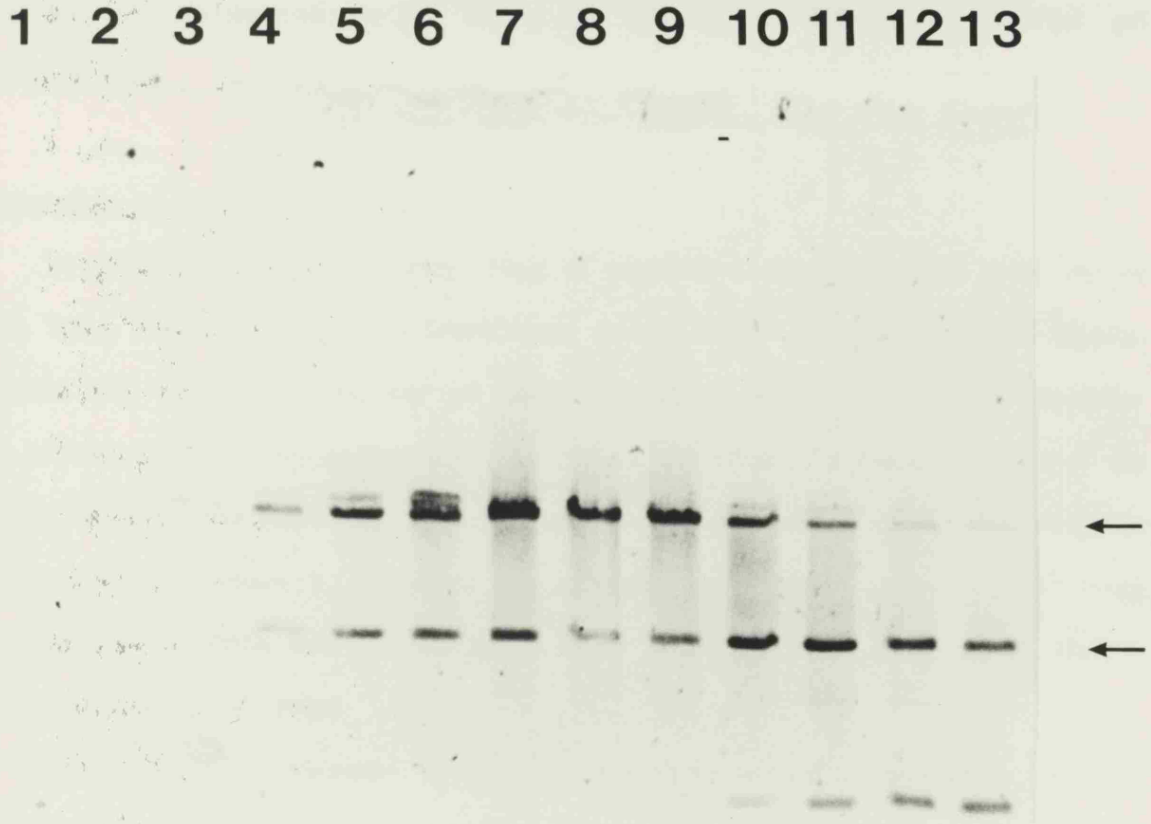
Preparation of the insert DNA is important to the final integrity of the genomic library. To generate a library representative of the genome it is important to produce a series of random, overlapping DNA fragments. For this purpose partial DNA digestion with *Sau*3A was employed (Chapter 2). The extent of enzyme digestion was assessed by agarose gel electrophoresis and the most suitable digestion conditions determined were scaled up for the preparation of insert DNA for cloning purposes.

To reduce the number of recombinants required for a representative library, *Sau*3A partially digested DNA can be size fractionated (as described for the vector

FIGURE 3.2 Agarose Gel Electrophoresis of Fractionated *Bam*HI/*Sal*I Digested λL47 DNA

250μl fractions were collected from the NaCl gradients and ethanol precipitated as described in Chapter 2. Fractions were resuspended in 40μl of 1× T.E. and 4μl aliquots analysed by agarose gel electrophoresis on a 0.8% gel. The arrows, upper and lower, indicate the 23.58kb and 10.41kb *Bam*HI vector arms respectively.

1. fraction 6
2. " 7
3. " 8
4. " 9
5. " 11
6 " 12
8. fraction 13
9. " 14
10. " 16
11. " 17
12. " 18
13. " 19


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

arms), to enrich for the larger fragments prior to cloning (Chapter 2). To assess the enrichment for specifically sized DNA species, aliquots of fractions were analysed by agarose gel electrophoresis and sizes estimated according to parallel molecular weight markers (Fig.3.3). Fractions enriched for DNA fragments between 11kb and 17kb were pooled and stored at -20°C. Again the suitability of this DNA for cloning purposes was determined by its ability to religate, as assessed by agarose gel electrophoresis.

Preparation of a Genomic Library

To generate a genomic library, 10μg of purified λL47 vector arms were ligated with 2.5μg *Sau*3A digested, size-fractionated human DNA. An aliquot of the ligation mixture was analysed by agarose gel electrophoresis to assess ligation. To determine the efficiency of ligation and estimate the size of the library, a further aliquot of the ligation mixture was taken and packaged using λ *in vitro* packaging extracts, prepared as described in Chapter 2. These extracts routinely gave approximately 10^8-10^9pfu/μg of wild type λ DNA and 10^7pfu/μg of λL47 DNA when titred on *E.coli* NM538 cells (Frischauf *et al*, 1983).

The P2 lysogenic derivative of NM538, NM539 (Frischauf *et al*, 1983), was used to estimate the frequency of recombinant phage. 200 prospective recombinant phage and 10 λL47 phage were transferred to duplicate lawns of NM538 and NM539 cells. All 200 phage were viable on both strains, whereas λL47 phage grew on *E.coli* NM538 only, indicating a high percentage of phage were genuine recombinants.

Phages derived from the library were examined further to confirm that they contained human DNA inserts. The presence of interspersed highly repetitive sequences throughout the human genome (Deininger *et al*, 1981) allows a simple evaluation of this. Thus, hybridisation to a total human DNA probe indicates the presence of these sequences. Approximately 500 recombinant phage were screened by plaque hybridisation using a nick translated total human DNA probe: >95% hybridised proving that they contained human DNA sequences.

After these pilot studies the remaining ligation mixture was packaged and the

FIGURE 3.3 Agarose Gel Electrophoresis of Size Fractionated *Sau*3A Partial Digests of Human Placental DNA

100μg aliquots of high molecular weight human placental DNA was partially digested for 15mins (A) or 20mins (B) with 1 unit of *Sau*3A (Chapter 2) and fractionated on a NaCl gradient (See legend to Fig.3.2). Aliquots of selected DNA fractions were analysed by agarose gel electrophoresis on a 0.4% gel.

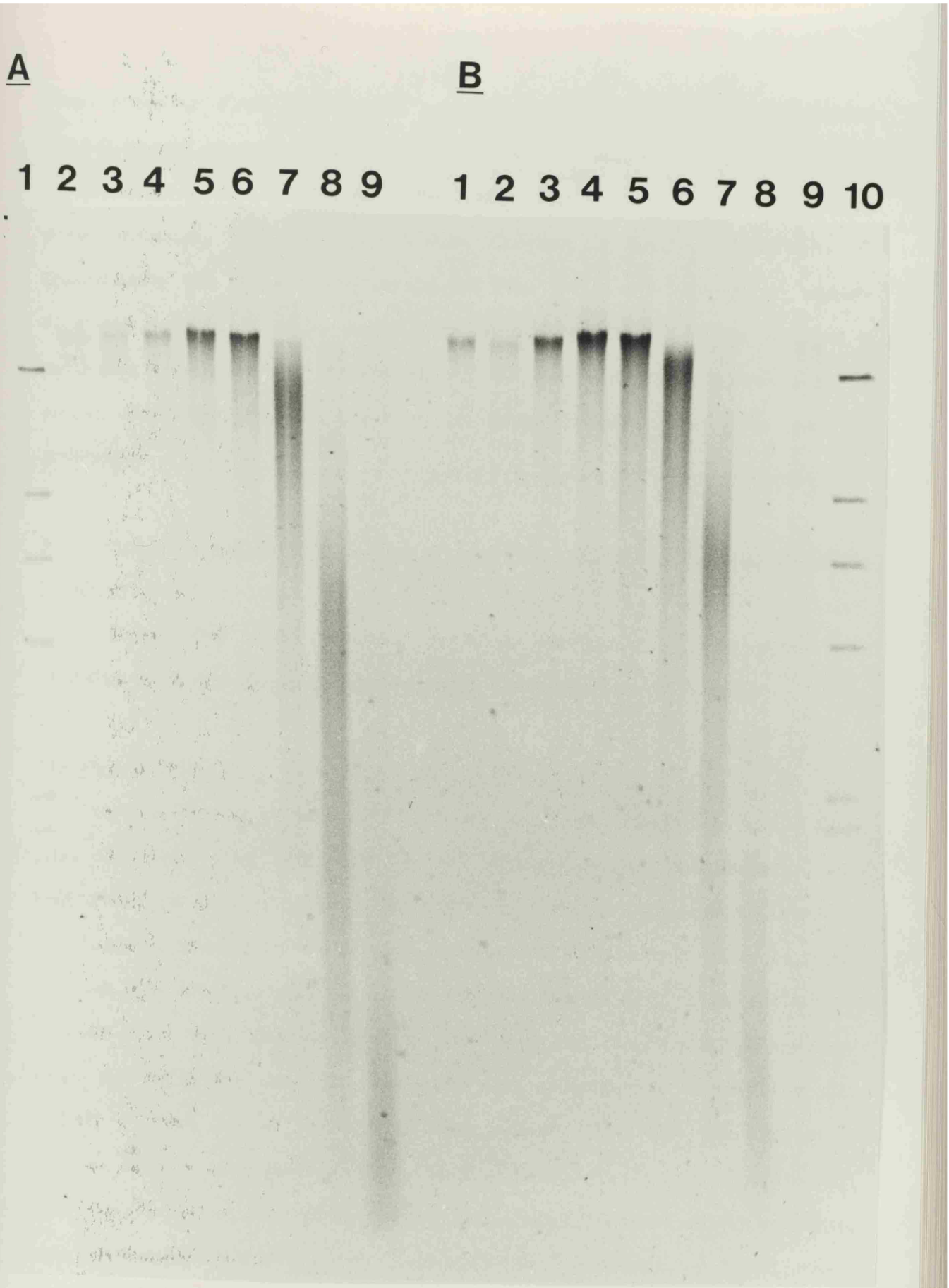
A)

1. λ *Hin*dIII Size Marker: 23.75kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb, 1.96kb, 0.59kb and 0.1kb.

2. fraction 1
3. " 5
4. " 10
5. " 15
6. " 20
7. " 25
8. " 30
9. " 35

B)

1. fraction 1
2. " 5
3. " 10
4. " 15
5. " 20
6. " 25
7. " 30
8. " 35
9. " 40
10. λ *Hin*dIII Size Marker


A
1 2 3 4 5 6 7 8 9
B
1 2 3 4 5 6 7 8 9 10

phage plated on *E.coli* NM538 cells generating a library constituting approximately 300,000 plaques, as determined by the phage titre. A complete library, representative of the human genome, constructed in this manner requires about 0.5×10^6 recombinant phage (Brammar, 1982). By this criteria, therefore, a library representative of approximately 50% of the human genome had been constructed. Although the number of clones is sufficient to be screened for IFNα gene sequences which comprise a multi-gene family, the full complement of IFNα gene sequences are unlikely to be present. Realistically perhaps only 5 to 10 clones containing IFNα genes could be anticipated.

3.3 Identification and Isolation of Recombinant Phage Containing Human IFNα Gene Sequences

Recombinant phage containing sequences homologous to IFNα genes were identified by *in situ* plaque hybridisation using a human IFNα cDNA probe.

The Human IFNα cDNA probe pHIF24.40

For the purpose of screening the newly constructed library for human IFNα genes the cDNA clone pHIF24.40 (Dr. J.D. Windass, personal communication) was used. pHIF24.40 contains a 994bp full length IFNα cDNA which was derived from NDV induced Namalva cells. The coding sequence of pHIF24.40 is shown in Figure 3.4 were it is compared with the cDNAs HuIFNα_1 (Nagata *et al*, 1980) and LeIF-E (Goeddel *et al*, 1981). Comparison of pHIF24.40 with HuIFNα_1 reveals 85% homology within the coding sequence at the nucleotide level. A comparison of pHIF24.40 with LeIF-E, a cDNA derived from Sendai virus induced KG–1 cells and believed to originate from a transcribed IFN pseudogene (Goeddel *et al*, 1981), indicates that these IFNα genes are substantially similar. Their coding sequences display only three base substitutions. Perfect homology extends to the 3' non-translated region, except for one additional base in pHIF24.40. It is highly likely, therefore, that these two clones represent alleles of the same pseudogene.

The suitability of pHIF24.40 to detect members of the IFNα gene family was

```
                1         10         20         30         40         50
pHIF24.40      :ATGGACTTGCCCTTTACTTTACTGATGGCTATGATGGTGCTCAGCTGCAAGTCCATCTGCC
LeIF-E         :                                                     -----
HuIFNα1        :----C--C-------G------------TCC--G------------------A-GTGC--

          70         80         90         100        110        120        130
TCTGGGCTGTGATCTGCCTCAGGCCCACAGCGTGGGTAACAGGAGGGCCTTCATACTCCTGACACAAATGAGGAG
------------------------------------------------------------------------------
--------------C---G--A--------C---A----------A----G--G----------------C--

    140        150        160        170        180        190        200
AATCTCTCCTTTTTCTTACCTGAAGGACAGACATGACTTTGATTTTCCATCATCAGGTGTTTCATGGCAACCACT
------------------------------------------------------------------------------
-----------CC--C-GT----T----------------G------C --GG---A----G----------G-

         220        230        240        250        260        270
TCCAGAAGGTTCAAGCTATCTTCCTTTTCCATGAGATGATGCAGCAGACCTTCAACCTCTTCAGCACAAAGGACT
------------------------------------------------------------------------------
---------C--C---C----CTG-CC--------C----C-------T------------T-C------A--T-

    290        300        310        320        330        340        350
CATCTGATACTTTGGATGCGACCCTTTTAGACAAATTCTACACTGAACTTTACCAGCAGCTGAATGACCTGGAAG
------------G-----A----------------C-------------------------------------
------C-G-----------GA---CC------------G---C-----C------------------T------

         370        380        390        400        410        420        430
CCTGTGTGATGTAGAAGGTTGGAGTGGAAGAGACTCCCCTGAGGAATGTGGACTCCATCCTGGCTGTGAGAAAAT
------------------------------------------------------------------------------
-----------C--G---AGA-G----G---A----------T----------------T---------AG----

    440        450        460        470        480        490        500
ACTTTCAAAGAATCACTCTTTATCTGACAAAGAAGAAGTATAGCCCTTGTTCCTGGGAGGCTGTCAGAGCAGAAA
------------------------------------------------------------------------------
----C-G-----------C---------A-------A--C---------G---------T--------------

         520        530        540        550        560
TCATGAGATCCTTCTCTTTATGAACGAACTTGCAGGAAAGATTAAGGAGGAAGGAATGA
------------------------------------------------------------
-----------C---------C---A--------A---------------------A-
```

FIGURE 3.4 Comparison of the Nucleotide Sequence of pHIF24.40 with HuIFNα1 and LeIF-E

The nucleotide sequence of the cDNA clone pHIF24.40 is illustrated. Below are aligned the sequences of the human cDNA clones $IFN\alpha_1$ (Nagata *et al*, 1980a) and LeIF-E (Goeddel *et al*, 1981). Regions of homology are depicted by dashes (-). Specific nucleotide differences are shown using the conventional symbols.

investigated by Southern blot analysis of *Eco*RI digested human DNA. For this purpose the 994bp *Pst*I insert isolated from pHIF24.40 by polyacrylamide gel electrophoresis (P.A.G.E.) was used as probe. The stringency of hybridisation employed (2xSSC, 65°C) should allow the detection of sequences showing >85% homology to pHIF24.40 (Meinkoth and Wahl, 1984). Figure 3.5 shows that pHIF24.40 hybridises to at least 11 distinct *Eco*RI fragments in human DNA. Thus pHIF24.40 hybridises to multiple species in human DNA, in keeping with a probe homologous to a family of approximately 20 genes. Therefore, it was used to screen the newly constructed human genomic DNA library for IFNα gene sequences.

Screening the Genomic Library with pHIF24.40

Recombinant phage within the genomic library containing IFNα genes were identified by *in situ* plaque hybridisation using the *Pst*I insert of pHIF24.40 as probe. This procedure was performed prior to amplification of the library to avoid the potential loss of less vigorous recombinant phage and to minimise possible rearrangements (Fritsch *et al*, 1980; Grosveld *et al*, 1981).

A total of 300,000 plaques were screened using similar conditions to those described above for Southern blot analysis. Recombinant phage, containing sequences homologous to the pHIF24.40 insert, were identified by autoradiography (Fig.3.6). Hybridisation-positive plaques were picked for subsequent rounds of hybridisation and plaque purification. A total of 22 pHIF24.40 related primary clones (I to XXII) identified were rescreened. Of these, four phages λSMTI, λSMTIII, λSMTVI and λSMTXIV were positive on subsequent rounds of screening. λSMTI.2, λSMTIII.1 and λSMTVI.1 were purified as described in Chapter 2 (Fig.3.7). λSMTXIV has not been characterised further.

3.4 Characterisation of pHIF24.40 Related Clones by Southern Hybridisation

To confirm that the purified phage λSMTI.2, λSMTIII.1 and λSMTVI.1 contained sequences complementary to IFNα genes, DNA was prepared for Southern blot analysis. DNA cloned into the *Bam*HI site of λL47 is flanked by *Hin*dIII and *Eco*RI

FIGURE 3.5 Southern Blot Analysis of Human DNA with pHIF24.40

10μg of *Eco*RI digested human placental DNA was loaded onto a 0.8% agarose gel and run o/n at 25V. The DNA was transferred to nitrocellulose (Chapter 2) and hybridisation performed in 5× Denhardt's (2×SSC, 65°C o/n) using the nick translated 994bp *Pst*I insert from pHIF24.40 as probe (1×10^8 cpm/μg DNA). The filter was washed in 2×SSC at 65°C and autoradiography performed at -70°C for 48hrs using Fuji RX film with an intensifying screen (Ilford, fast tungstate). Arrows show *Hind*III digested λ DNA size markers, 23.72kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb and 1.96kb.

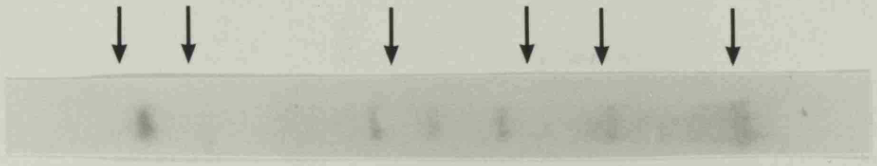


FIGURE 3.6 Identification of IFNα Genomic Clones by Plaque Hybridisation

Approximately 20,000 plaques per 9cm nitrocellulose filter were screened as described in Chapter 2. The conditions of hybridisation and the probe used were as described in the legend to Fig.3.5. Autoradiography was performed o/n. Arrows indicate positively hybridising plaques.

FIGURE 3.7 Plaque Purification

Purified plaques were transferred onto a lawn of C600 plating cells (Chapter 2). Hybridisation was performed as described in the legend to Fig.3.5. Autoradiography was performed o/n.

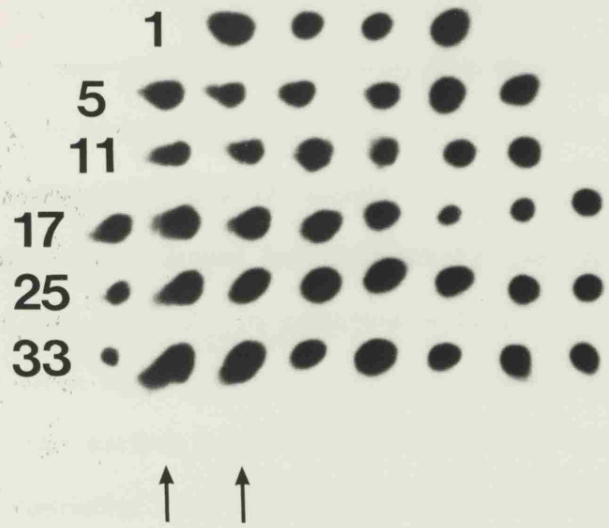
A) 1 to 40 λSMTI.2

B) 1 to 40 λSMTIII.1

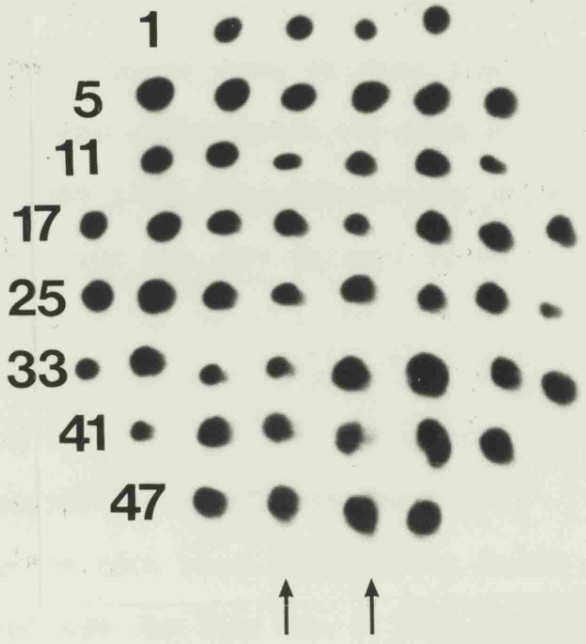
41 to 50 λSMTVI.1

Arrows indicate λL47 phage controls.

A



B



restriction sites (Fig.3.1A). In the construction of the genomic library the *Bam*HI sites of λL47 may be recreated, however in most cases the ligation of *Bam*HI and *Sau*3A digested DNAs will destroy the vector cloning sites. Thus, to resolve cloned DNA inserts from the phage vector arms by agarose gel electrophoresis, *Eco*RI or *Hin*dIII digests can be used, but there may be internal sites and thus excision of a single fragment was not necessarily anticipated. This procedure can be used to estimate the DNA insert size by totalling the size of specific DNA fragments generated by restriction enzyme digestion. The distance between the λL47 cloning sites and *Eco*RI or *Hin*dIII sites is known and hence can be compensated for.

Recombinant phage DNA from all three clones was digested with *Eco*RI, *Hin*dIII or *Bam*HI, and Southern blot analysis performed. Again the pHIF24.40 *Pst*I insert was used as a probe. The restriction and hybridisation patterns indicate that all three clones contain IFNα gene related sequences (Figs.3.8, 3.9 and 3.10).

The hybridisation and restriction digest patterns for each of the recombinant clones were different. The number and size of fragments generated by *Bam*HI, *Eco*RI and *Hin*dIII digests of DNA derived from all three genomic clones, and those fragments containing IFNα related gene sequences are shown in Table 3.1.

λSMTI.2 contains a 9kb DNA insert. The absence of any *Bam*HI sites in λSMTI.2 DNA indicates that these sites were disrupted by the cloning procedure and also that no sites are present within the DNA insert (Fig.3.8A lane 7). Neither are there any internal *Eco*RI sites in this clone (Fig.3.8A lane 8). Digestion with *Hin*dIII released two clone specific fragments, including a 3.5kb fragment (Fig.3.8A lane 6) which hybridised to the pHIF24.40 insert (Fig.3.8B lane 6).

λSMTIII.1 contains an 8kb DNA insert. Digestion with *Bam*HI generated a 10.1kb fragment which comigrated with the right arm of λL47 (Fig.3.9A lane 4 and 7) suggesting that the right hand *Bam*HI site had been recreated during the cloning procedure. However, no *Bam*HI sites are present within the human DNA insert itself. Digests of λSMTIII.1 DNA with *Eco*RI produced six clone specific fragments (Fig.3.9A lane 8), including both 2kb and 1kb fragments which hybridised to the pHIF24.40 insert (Fig3.9B lane 8). The 1kb fragment is actually a doublet thus either

FIGURE 3.8 Southern Blot Analysis of λSMTI.2 DNA

DNA digests were performed according to the manufacturers recommendation. 1μg aliquots of digested DNAs were analysed by agarose gel electrophoresis on a 0.8% gel and bands visualised by ethidium bromide staining (A). Southern blot analysis was performed as described in the legend to Fig.3.5. Autoradiography was for 5hrs at -70°C using Fuji RX film with an intensifying screen (B).

1. λ *Eco*RI Size Marker*	6. λSMTI.2 *Hin*dIII
2. " *Hin*dIII Size Marker†	7. " *Bam*HI
3. λL47 *Hin*dIII	8. " *Eco*RI
4. " *Bam*HI	9. λ *Hin*dIII
5. " *Eco*RI	10. " *Eco*RI

Arrows indicate λ *Eco*RI size markers (B)

* 21.8kb, 7.55kb, 5.93kb, 5.54kb, 4.8kb and 3.38kb

† 23.72kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb, 1.96kb, 0.59kb and 0.1kb.

A

B

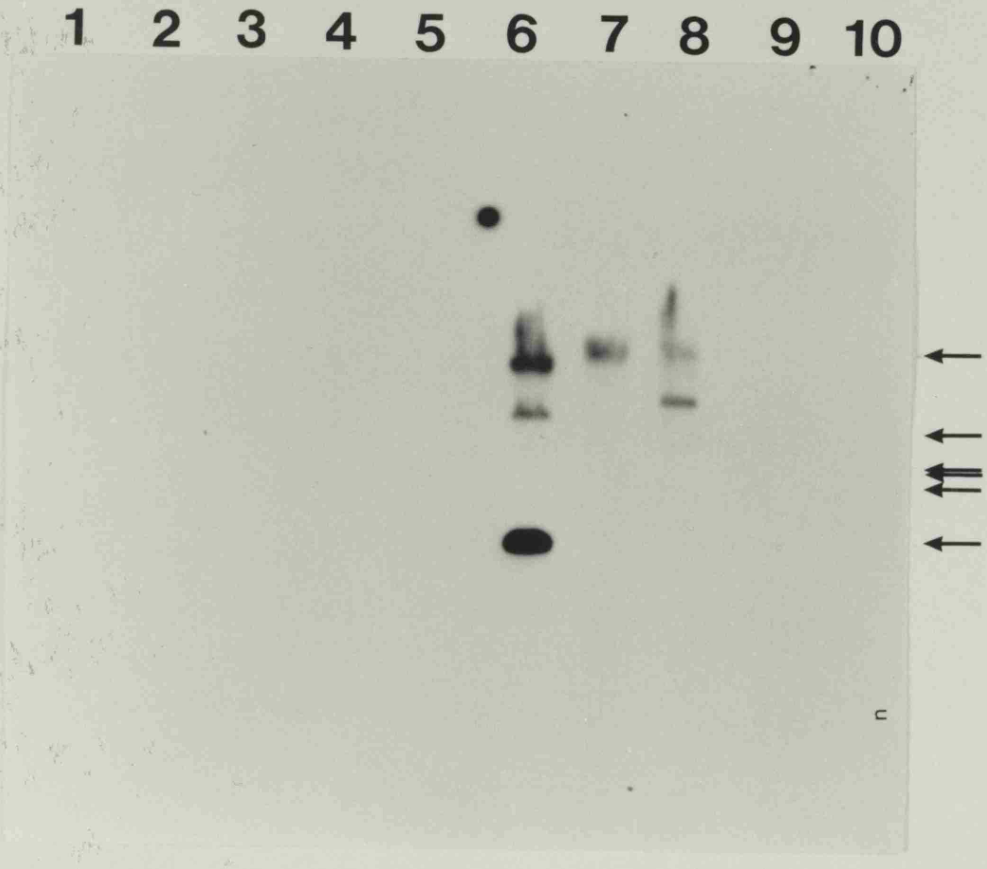


FIGURE.3.9 Southern Analysis of λSMTIII.1 DNA

See legend to Fig.3.8.

1. λ *Eco*RI Size Marker
2. " *Hin*dIII Size Marker
3. λL47 *Hin*dIII
4. " *Bam*HI
5. " *Eco*RI
6. λSMTIII.1 *Hin*dIII
7. " *Bam*HI
8. " *Eco*RI
9. λ *Hin*dIII
10. " *Eco*RI

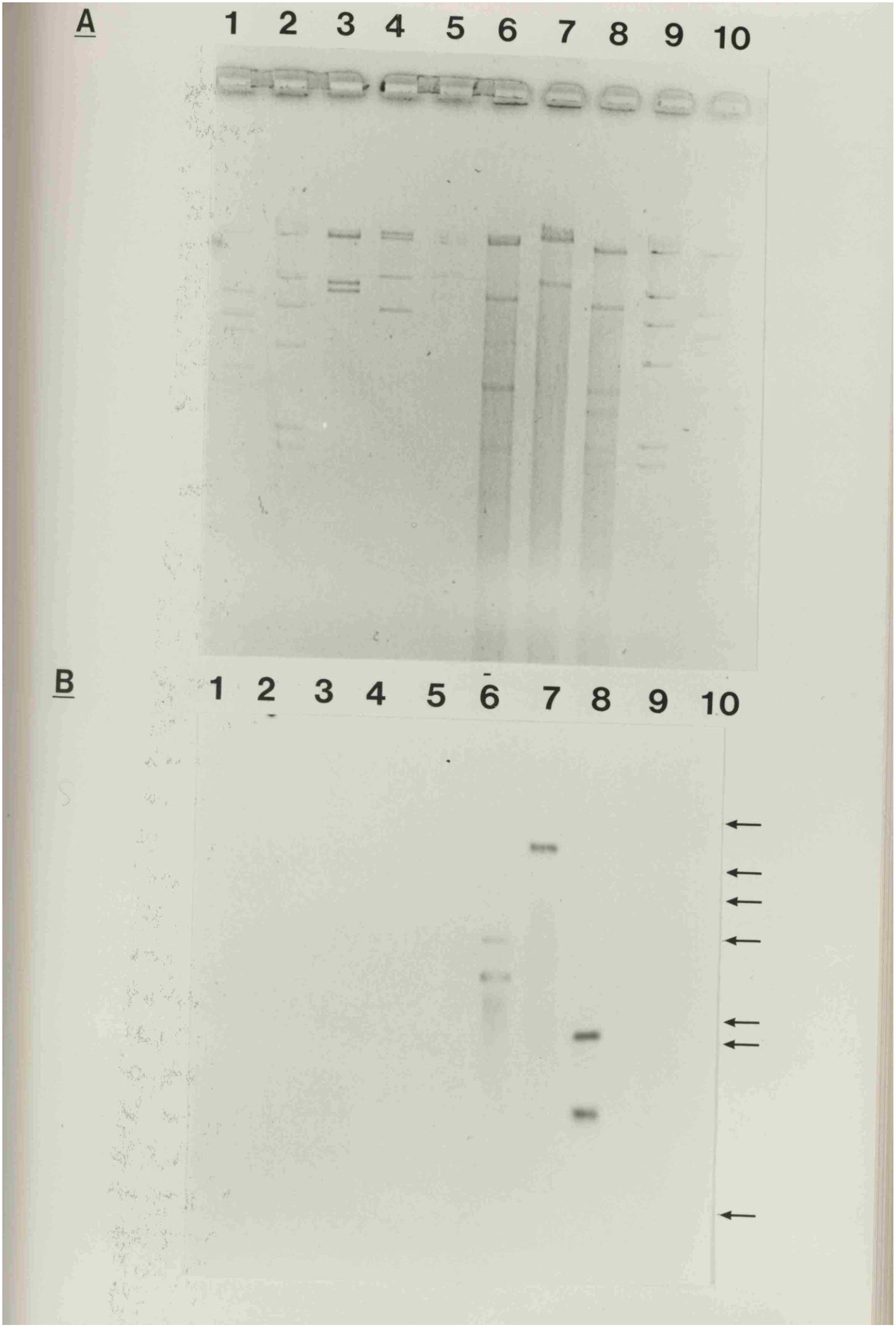
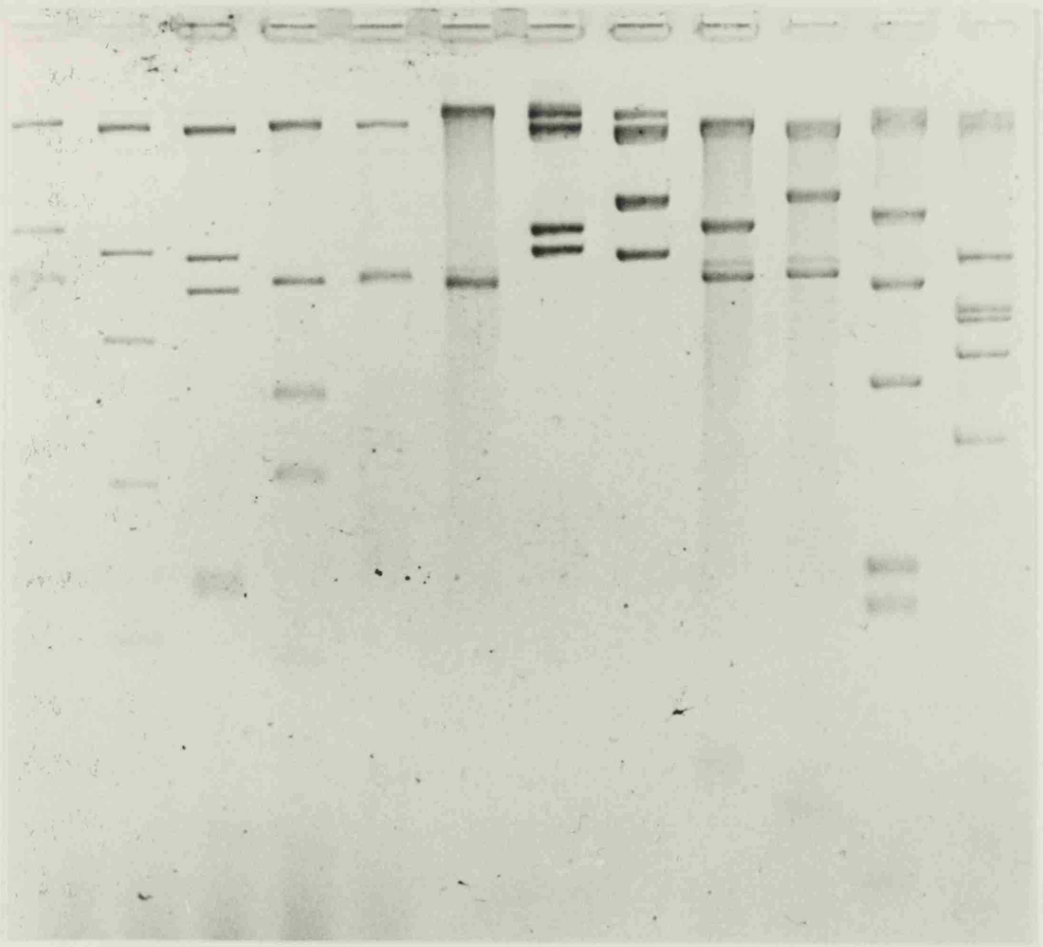

A
1 2 3 4 5 6 7 8 9 10
B
1 2 3 4 5 6 7 8 9 10

FIGURE 3.10 Southern Analysis of λSMTV.1 DNA

See legend to Fig.3.8 (*Bam*HI digests not shown)

1. λSMTVI.1 *Bgl*II
2. " *Hin*dIII
3. " *Eco*RI
4. " *Bgl*II/*Hin*dIII
5. " *Bgl*II/*Eco*RI
6. λL47 *Bgl*II
7. " *Hin*dIII
8. " *Eco*RI
9. " *Bgl*II/*Hin*dIII
10. " *Bgl*II/*Eco*RI
11. λ *Hin*dIII Size Marker
12. " *Eco*RI Size Marker

A



B

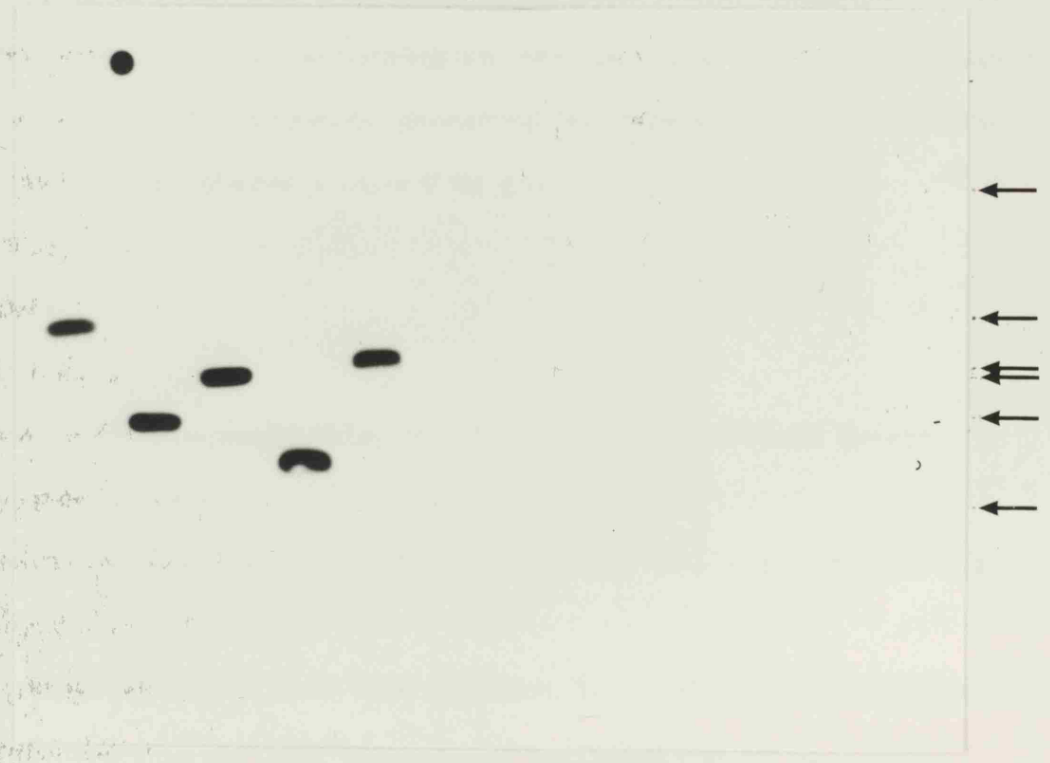
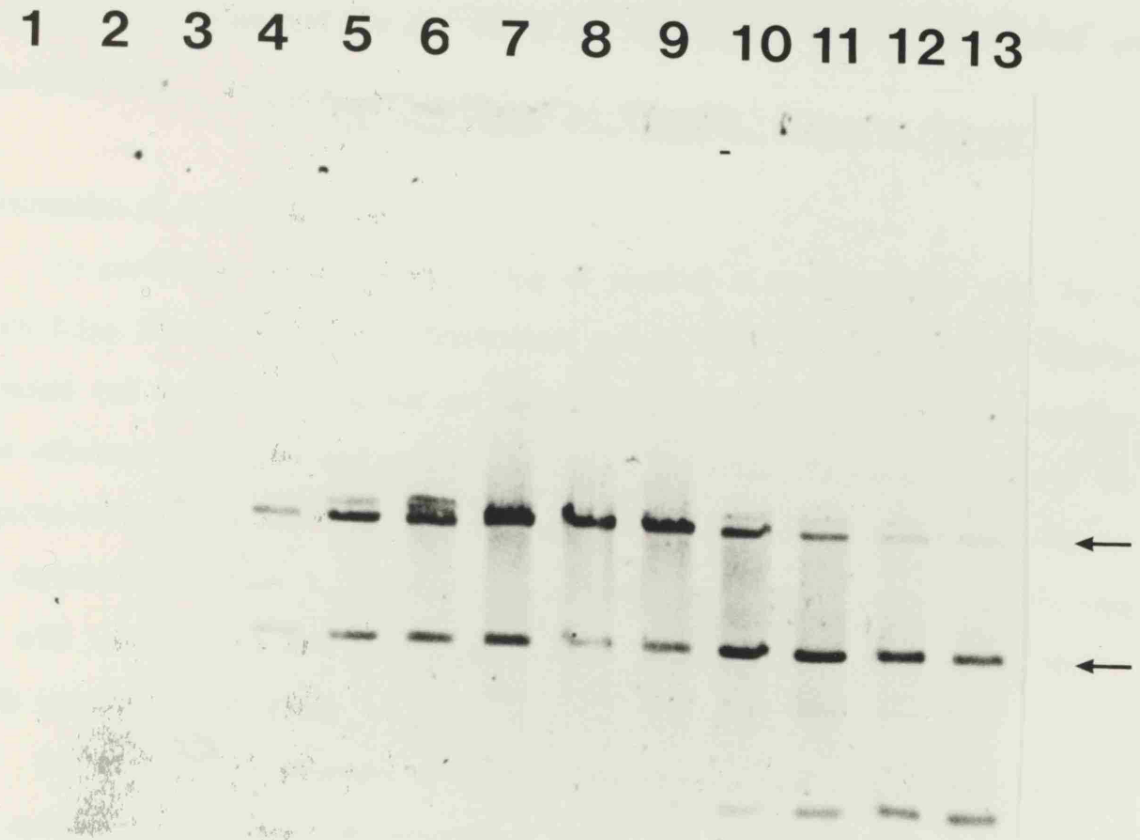


FIGURE 3.2 Agarose Gel Electrophoresis of Fractionated *Bam*HI/*Sal*I Digested λL47 DNA

250μl fractions were collected from the NaCl gradients and ethanol precipitated as described in Chapter 2. Fractions were resuspended in 40μl of 1× T.E. and 4μl aliquots analysed by agarose gel electrophoresis on a 0.8% gel. The arrows, upper and lower, indicate the 23.58kb and 10.41kb *Bam*HI vector arms respectively.

1. fraction 6
2. " 7
3. " 8
4. " 9
5. " 11
6 " 12
8. fraction 13
9. " 14
10. " 16
11. " 17
12. " 18
13. " 19


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

arms), to enrich for the larger fragments prior to cloning (Chapter 2). To assess the enrichment for specifically sized DNA species, aliquots of fractions were analysed by agarose gel electrophoresis and sizes estimated according to parallel molecular weight markers (Fig.3.3). Fractions enriched for DNA fragments between 11kb and 17kb were pooled and stored at -20°C. Again the suitability of this DNA for cloning purposes was determined by its ability to religate, as assessed by agarose gel electrophoresis.

Preparation of a Genomic Library

To generate a genomic library, 10μg of purified λL47 vector arms were ligated with 2.5μg *Sau*3A digested, size-fractionated human DNA. An aliquot of the ligation mixture was analysed by agarose gel electrophoresis to assess ligation. To determine the efficiency of ligation and estimate the size of the library, a further aliquot of the ligation mixture was taken and packaged using λ *in vitro* packaging extracts, prepared as described in Chapter 2. These extracts routinely gave approximately 10^8-10^9pfu/μg of wild type λ DNA and 10^7pfu/μg of λL47 DNA when titred on *E.coli* NM538 cells (Frischauf *et al*, 1983).

The P2 lysogenic derivative of NM538, NM539 (Frischauf *et al*, 1983), was used to estimate the frequency of recombinant phage. 200 prospective recombinant phage and 10 λL47 phage were transferred to duplicate lawns of NM538 and NM539 cells. All 200 phage were viable on both strains, whereas λL47 phage grew on *E.coli* NM538 only, indicating a high percentage of phage were genuine recombinants.

Phages derived from the library were examined further to confirm that they contained human DNA inserts. The presence of interspersed highly repetitive sequences throughout the human genome (Deininger *et al*, 1981) allows a simple evaluation of this. Thus, hybridisation to a total human DNA probe indicates the presence of these sequences. Approximately 500 recombinant phage were screened by plaque hybridisation using a nick translated total human DNA probe: >95% hybridised proving that they contained human DNA sequences.

After these pilot studies the remaining ligation mixture was packaged and the

FIGURE 3.3 Agarose Gel Electrophoresis of Size Fractionated *Sau*3A Partial Digests of Human Placental DNA

100μg aliquots of high molecular weight human placental DNA was partially digested for 15mins (A) or 20mins (B) with 1 unit of *Sau*3A (Chapter 2) and fractionated on a NaCl gradient (See legend to Fig.3.2). Aliquots of selected DNA fractions were analysed by agarose gel electrophoresis on a 0.4% gel.

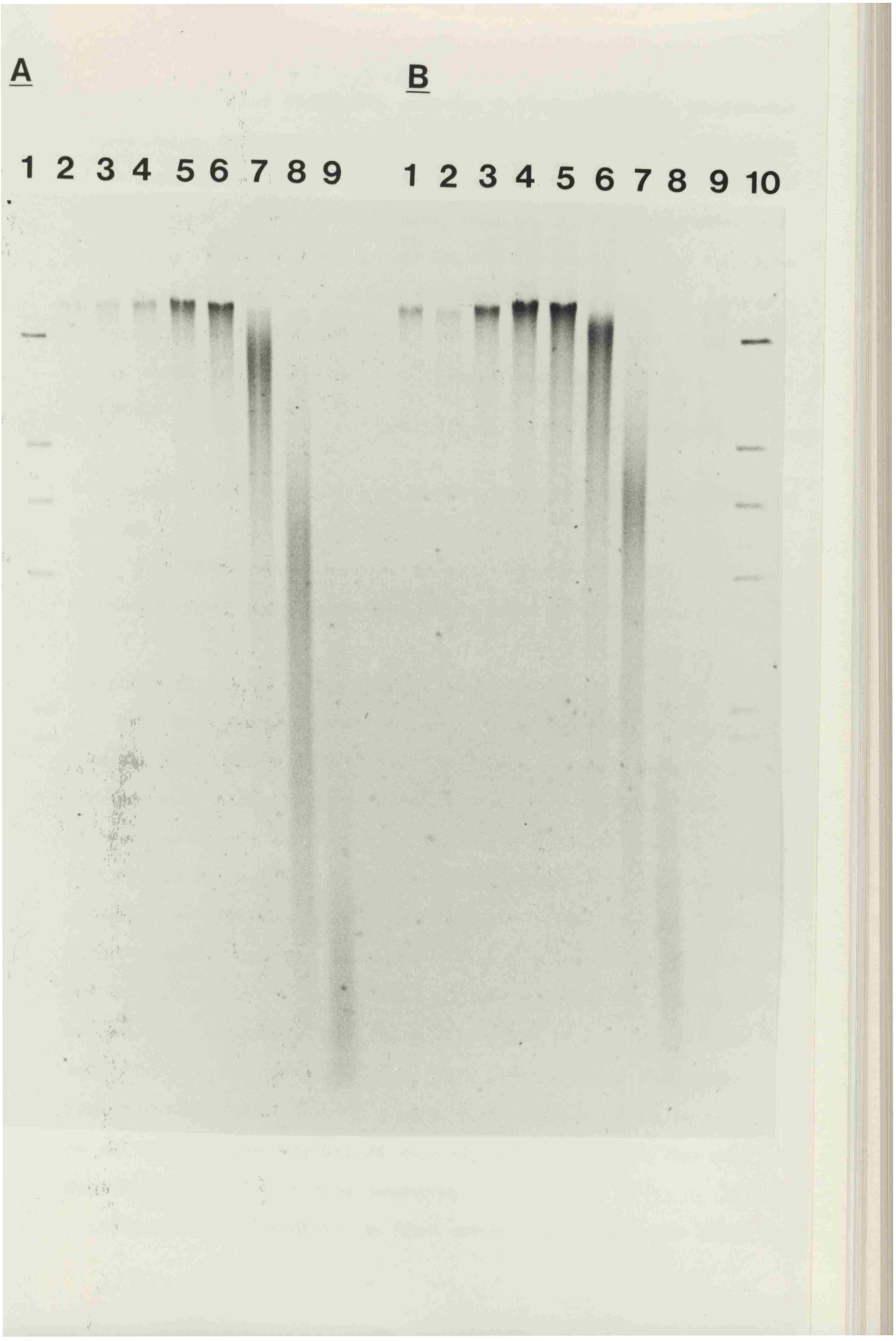
A)

1. λ *Hin*dIII Size Marker: 23.75kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb, 1.96kb, 0.59kb and 0.1kb.

2. fraction 1
3. " 5
4. " 10
5. " 15
6. " 20
7. " 25
8. " 30
9. " 35

B)

1. fraction 1
2. " 5
3. " 10
4. " 15
5. " 20
6. " 25
7. " 30
8. " 35
9. " 40
10. λ *Hin*dIII Size Marker


A
1 2 3 4 5 6 7 8 9
B
1 2 3 4 5 6 7 8 9 10

phage plated on *E.coli* NM538 cells generating a library constituting approximately 300,000 plaques, as determined by the phage titre. A complete library, representative of the human genome, constructed in this manner requires about 0.5×10^6 recombinant phage (Brammar, 1982). By this criteria, therefore, a library representative of approximately 50% of the human genome had been constructed. Although the number of clones is sufficient to be screened for IFNα gene sequences which comprise a multi-gene family, the full complement of IFNα gene sequences are unlikely to be present. Realistically perhaps only 5 to 10 clones containing IFNα genes could be anticipated.

3.3 Identification and Isolation of Recombinant Phage Containing Human IFNα Gene Sequences

Recombinant phage containing sequences homologous to IFNα genes were identified by *in situ* plaque hybridisation using a human IFNα cDNA probe.

The Human IFNα cDNA probe pHIF24.40

For the purpose of screening the newly constructed library for human IFNα genes the cDNA clone pHIF24.40 (Dr. J.D. Windass, personal communication) was used. pHIF24.40 contains a 994bp full length IFNα cDNA which was derived from NDV induced Namalva cells. The coding sequence of pHIF24.40 is shown in Figure 3.4 were it is compared with the cDNAs HuIFNα$_1$ (Nagata *et al*, 1980) and LeIF-E (Goeddel *et al*, 1981). Comparison of pHIF24.40 with HuIFNα$_1$ reveals 85% homology within the coding sequence at the nucleotide level. A comparison of pHIF24.40 with LeIF-E, a cDNA derived from Sendai virus induced KG–1 cells and believed to originate from a transcribed IFN pseudogene (Goeddel *et al*, 1981), indicates that these IFNα genes are substantially similar. Their coding sequences display only three base substitutions. Perfect homology extends to the 3' non-translated region, except for one additional base in pHIF24.40. It is highly likely, therefore, that these two clones represent alleles of the same pseudogene.

The suitability of pHIF24.40 to detect members of the IFNα gene family was

```
                1         10         20         30         40         50
pHIF24.40      :ATGGACTTGCCCTTTACTTTACTGATGGCTATGATGGTGCTCAGCTGCAAGTCCATCTGCC
LeIF-E         :                                                      -----
HuIFNα1        :----C--C-------G------------TCC--G------------------A-GTGC--

          70         80         90         100        110        120        130
TCTGGGCTGTGATCTGCCTCAGGCCCACAGCGTGGGTAACAGGAGGGCCTTCATACTCCTGACACAAATGAGGAG
------------------------------------------------------------------------------
--------------C---G--A--------C---A----------A----G--G----------------C--

    140        150        160        170        180        190        200
AATCTCTCCTTTTTCTTACCTGAAGGACAGACATGACTTTGATTTTCCATCATCAGGTGTTTCATGGCAACCACT
------------------------------------------------------------------------------
-----------CC--C-GT----T----------------G------C --GG---A----G----------G-

        220        230        240        250        260        270
TCCAGAAGGTTCAAGCTATCTTCCTTTTCCATGAGATGATGCAGCAGACCTTCAACCTCTTCAGCACAAAGGACT
------------------------------------------------------------------------------
---------C--C---C----CTG-CC--------C----C-------T------------T-C------A--T-

    290        300        310        320        330        340        350
CATCTGATACTTTGGATGCGACCCTTTTAGACAAATTCTACACTGAACTTTACCAGCAGCTGAATGACCTGGAAG
------------G-----A----------------C-------------------------------------
------C-G-----------GA---CC------------G---C-----C-----------------T------

          370        380        390        400        410        420        430
CCTGTGTGATGTAGAAGGTTGGAGTGGAAGAGACTCCCCTGAGGAATGTGGACTCCATCCTGGCTGTGAGAAAAT
------------------------------------------------------------------------------
-----------C--G---AGA-G----G---A----------T----------------T---------AG----

    440        450        460        470        480        490        500
ACTTTCAAAGAATCACTCTTTATCTGACAAAGAAGAAGTATAGCCCTTGTTCCTGGGAGGCTGTCAGAGCAGAAA
------------------------------------------------------------------------------
----C-G-----------C---------A-------A--C---------G---------T--------------

          520        530        540        550        560
TCATGAGATCCTTCTCTTTATGAACGAACTTGCAGGAAAGATTAAGGAGGAAGGAATGA
------------------------------------------------------------
-----------C---------C---A--------A---------------------A-
```

FIGURE 3.4 Comparison of the Nucleotide Sequence of pHIF24.40 with HuIFNα1 and LeIF-E

The nucleotide sequence of the cDNA clone pHIF24.40 is illustrated. Below are aligned the sequences of the human cDNA clones $IFN\alpha_1$ (Nagata *et al*, 1980a) and LeIF-E (Goeddel *et al*, 1981). Regions of homology are depicted by dashes (-). Specific nucleotide differences are shown using the conventional symbols.

investigated by Southern blot analysis of *Eco*RI digested human DNA. For this purpose the 994bp *Pst*I insert isolated from pHIF24.40 by polyacrylamide gel electrophoresis (P.A.G.E.) was used as probe. The stringency of hybridisation employed (2xSSC, 65°C) should allow the detection of sequences showing >85% homology to pHIF24.40 (Meinkoth and Wahl, 1984). Figure 3.5 shows that pHIF24.40 hybridises to at least 11 distinct *Eco*RI fragments in human DNA. Thus pHIF24.40 hybridises to multiple species in human DNA, in keeping with a probe homologous to a family of approximately 20 genes. Therefore, it was used to screen the newly constructed human genomic DNA library for IFNα gene sequences.

Screening the Genomic Library with pHIF24.40

Recombinant phage within the genomic library containing IFNα genes were identified by *in situ* plaque hybridisation using the *Pst*I insert of pHIF24.40 as probe. This procedure was performed prior to amplification of the library to avoid the potential loss of less vigorous recombinant phage and to minimise possible rearrangements (Fritsch *et al*, 1980; Grosveld *et al*, 1981).

A total of 300,000 plaques were screened using similar conditions to those described above for Southern blot analysis. Recombinant phage, containing sequences homologous to the pHIF24.40 insert, were identified by autoradiography (Fig.3.6). Hybridisation-positive plaques were picked for subsequent rounds of hybridisation and plaque purification. A total of 22 pHIF24.40 related primary clones (I to XXII) identified were rescreened. Of these, four phages λSMTI, λSMTIII, λSMTVI and λSMTXIV were positive on subsequent rounds of screening. λSMTI.2, λSMTIII.1 and λSMTVI.1 were purified as described in Chapter 2 (Fig.3.7). λSMTXIV has not been characterised further.

3.4 Characterisation of pHIF24.40 Related Clones by Southern Hybridisation

To confirm that the purified phage λSMTI.2, λSMTIII.1 and λSMTVI.1 contained sequences complementary to IFNα genes, DNA was prepared for Southern blot analysis. DNA cloned into the *Bam*HI site of λL47 is flanked by *Hin*dIII and *Eco*RI

FIGURE 3.5 Southern Blot Analysis of Human DNA with pHIF24.40

10μg of *Eco*RI digested human placental DNA was loaded onto a 0.8% agarose gel and run o/n at 25V. The DNA was transferred to nitrocellulose (Chapter 2) and hybridisation performed in 5× Denhardt's (2×SSC, 65°C o/n) using the nick translated 994bp *Pst*I insert from pHIF24.40 as probe (1×10^8 cpm/μg DNA). The filter was washed in 2×SSC at 65°C and autoradiography performed at -70°C for 48hrs using Fuji RX film with an intensifying screen (Ilford, fast tungstate). Arrows show *Hind*III digested λ DNA size markers, 23.72kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb and 1.96kb.

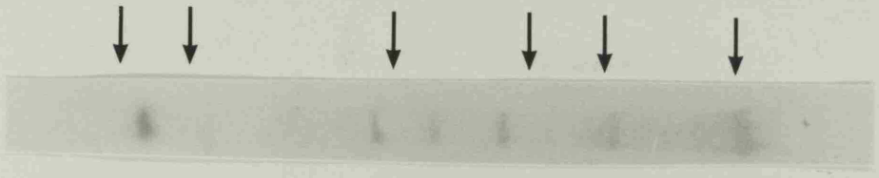


FIGURE 3.6 Identification of IFNα Genomic Clones by Plaque Hybridisation

Approximately 20,000 plaques per 9cm nitrocellulose filter were screened as described in Chapter 2. The conditions of hybridisation and the probe used were as described in the legend to Fig.3.5. Autoradiography was performed o/n. Arrows indicate positively hybridising plaques.

FIGURE 3.7 Plaque Purification

Purified plaques were transferred onto a lawn of C600 plating cells (Chapter 2). Hybridisation was performed as described in the legend to Fig.3.5. Autoradiography was performed o/n.

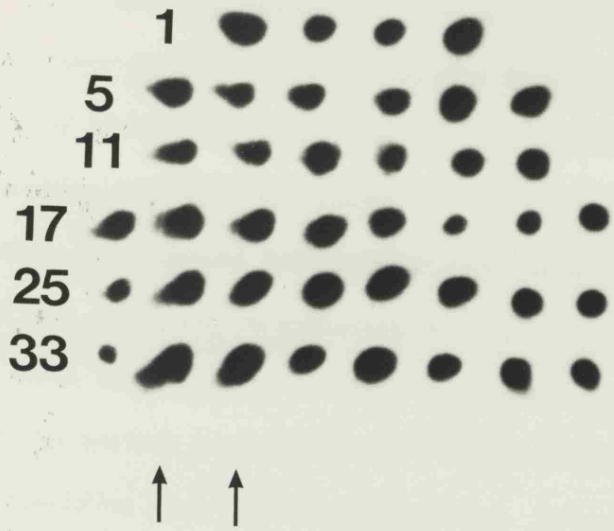
A) 1 to 40 λSMTI.2

B) 1 to 40 λSMTIII.1

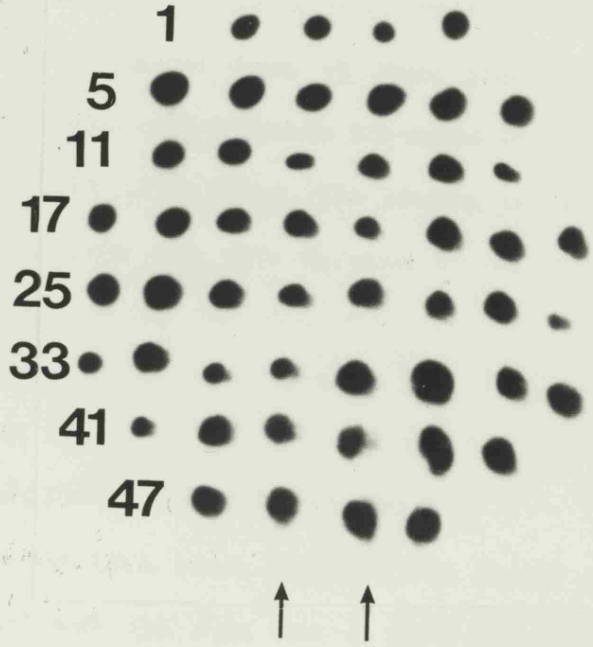
41 to 50 λSMTVI.1

Arrows indicate λL47 phage controls.

A



B



restriction sites (Fig.3.1A). In the construction of the genomic library the *Bam*HI sites of λL47 may be recreated, however in most cases the ligation of *Bam*HI and *Sau*3A digested DNAs will destroy the vector cloning sites. Thus, to resolve cloned DNA inserts from the phage vector arms by agarose gel electrophoresis, *Eco*RI or *Hin*dIII digests can be used, but there may be internal sites and thus excision of a single fragment was not necessarily anticipated. This procedure can be used to estimate the DNA insert size by totalling the size of specific DNA fragments generated by restriction enzyme digestion. The distance between the λL47 cloning sites and *Eco*RI or *Hin*dIII sites is known and hence can be compensated for.

Recombinant phage DNA from all three clones was digested with *Eco*RI, *Hin*dIII or *Bam*HI, and Southern blot analysis performed. Again the pHIF24.40 *Pst*I insert was used as a probe. The restriction and hybridisation patterns indicate that all three clones contain IFNα gene related sequences (Figs.3.8, 3.9 and 3.10).

The hybridisation and restriction digest patterns for each of the recombinant clones were different. The number and size of fragments generated by *Bam*HI, *Eco*RI and *Hin*dIII digests of DNA derived from all three genomic clones, and those fragments containing IFNα related gene sequences are shown in Table 3.1.

λSMTI.2 contains a 9kb DNA insert. The absence of any *Bam*HI sites in λSMTI.2 DNA indicates that these sites were disrupted by the cloning procedure and also that no sites are present within the DNA insert (Fig.3.8A lane 7). Neither are there any internal *Eco*RI sites in this clone (Fig.3.8A lane 8). Digestion with *Hin*dIII released two clone specific fragments, including a 3.5kb fragment (Fig.3.8A lane 6) which hybridised to the pHIF24.40 insert (Fig.3.8B lane 6).

λSMTIII.1 contains an 8kb DNA insert. Digestion with *Bam*HI generated a 10.1kb fragment which comigrated with the right arm of λL47 (Fig.3.9A lane 4 and 7) suggesting that the right hand *Bam*HI site had been recreated during the cloning procedure. However, no *Bam*HI sites are present within the human DNA insert itself. Digests of λSMTIII.1 DNA with *Eco*RI produced six clone specific fragments (Fig.3.9A lane 8), including both 2kb and 1kb fragments which hybridised to the pHIF24.40 insert (Fig3.9B lane 8). The 1kb fragment is actually a doublet thus either

FIGURE 3.8 Southern Blot Analysis of λSMTI.2 DNA

DNA digests were performed according to the manufacturers recommendation. 1μg aliquots of digested DNAs were analysed by agarose gel electrophoresis on a 0.8% gel and bands visualised by ethidium bromide staining (A). Southern blot analysis was performed as described in the legend to Fig.3.5. Autoradiography was for 5hrs at -70°C using Fuji RX film with an intensifying screen (B).

1. λ *Eco*RI Size Marker*
2. " *Hin*dIII Size Marker†
3. λL47 *Hin*dIII
4. " *Bam*HI
5. " *Eco*RI
6. λSMTI.2 *Hin*dIII
7. " *Bam*HI
8. " *Eco*RI
9. λ *Hin*dIII
10. " *Eco*RI

Arrows indicate λ *Eco*RI size markers (B)

* 21.8kb, 7.55kb, 5.93kb, 5.54kb, 4.8kb and 3.38kb

† 23.72kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb, 1.96kb, 0.59kb and 0.1kb.

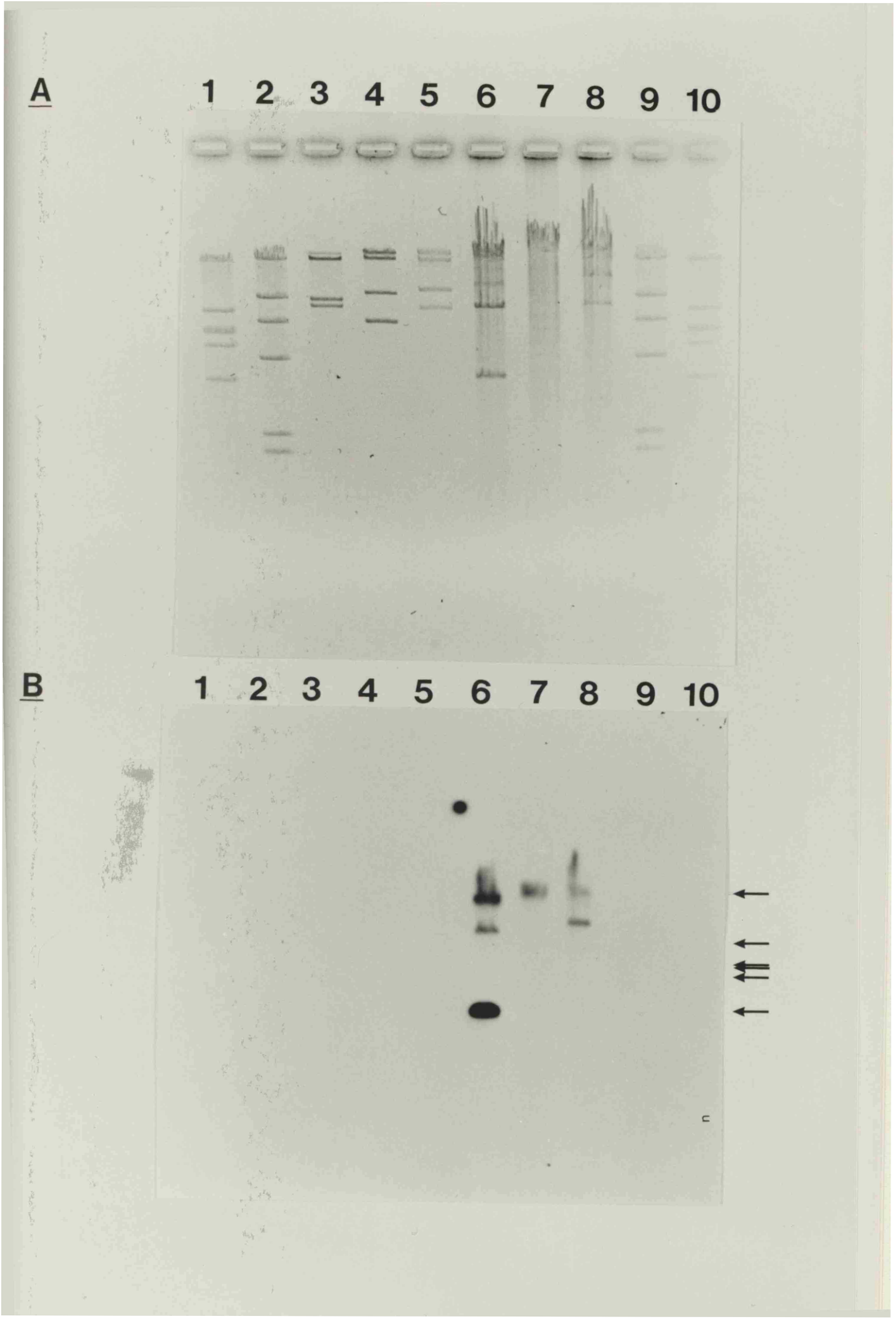
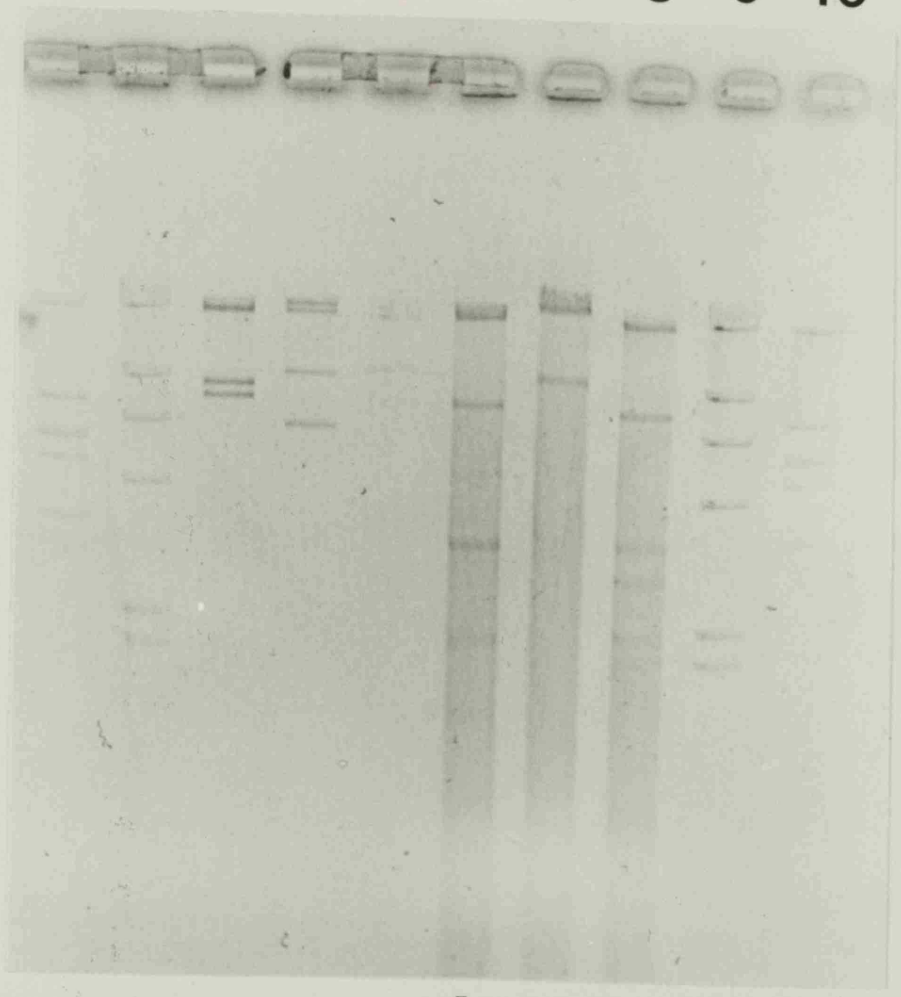

A
1 2 3 4 5 6 7 8 9 10
B
1 2 3 4 5 6 7 8 9 10

FIGURE.3.9 Southern Analysis of λSMTIII.1 DNA

See legend to Fig.3.8.

1. λ *Eco*RI Size Marker
2. " *Hin*dIII Size Marker
3. λL47 *Hin*dIII
4. " *Bam*HI
5. " *Eco*RI
6. λSMTIII.1 *Hin*dIII
7. " *Bam*HI
8. " *Eco*RI
9. λ *Hin*dIII
10. " *Eco*RI

A



B

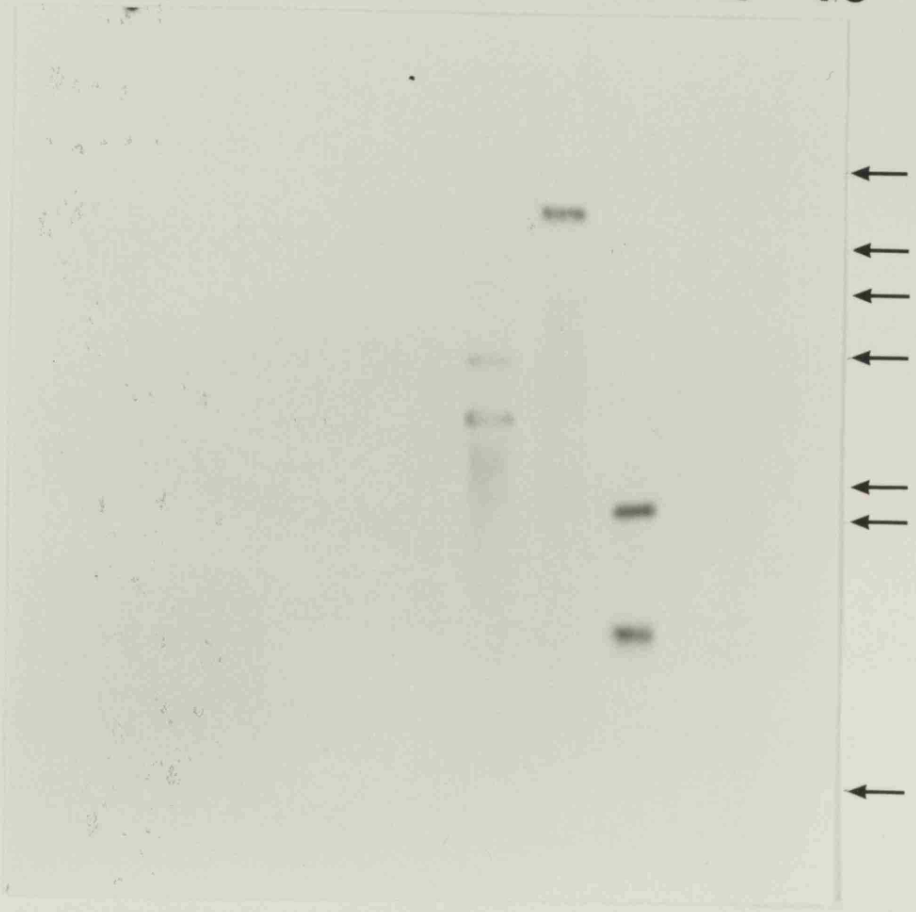
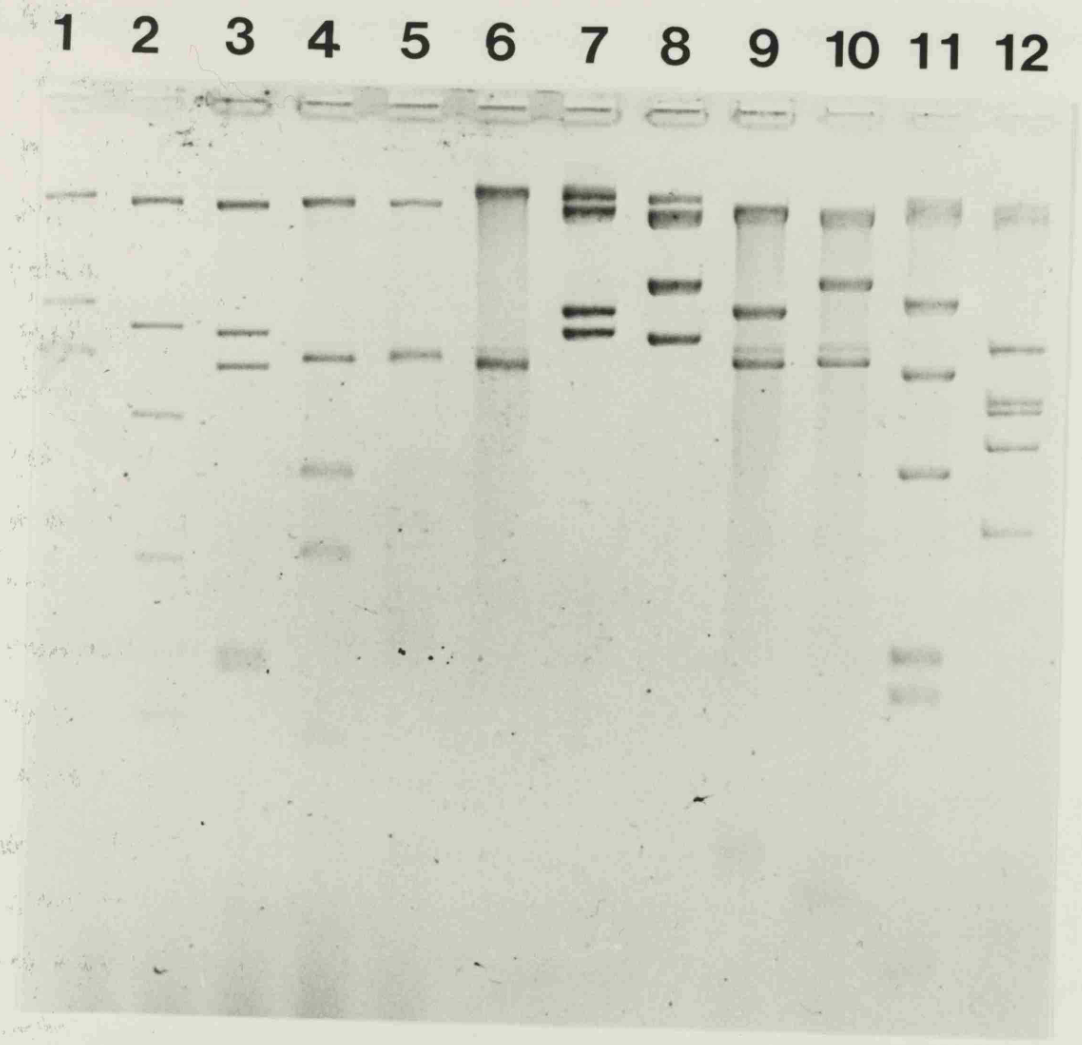


FIGURE 3.10 Southern Analysis of λSMTV.1 DNA

See legend to Fig.3.8 (*Bam*HI digests not shown)

1. λSMTVI.1 *Bgl*II
2. " *Hin*dIII
3. " *Eco*RI
4. " *Bgl*II/*Hin*dIII
5. " *Bgl*II/*Eco*RI
6. λL47 *Bgl*II
7. " *Hin*dIII
8. " *Eco*RI
9. " *Bgl*II/*Hin*dIII
10. " *Bgl*II/*Eco*RI
11. λ *Hin*dIII Size Marker
12. " *Eco*RI Size Marker

A



B

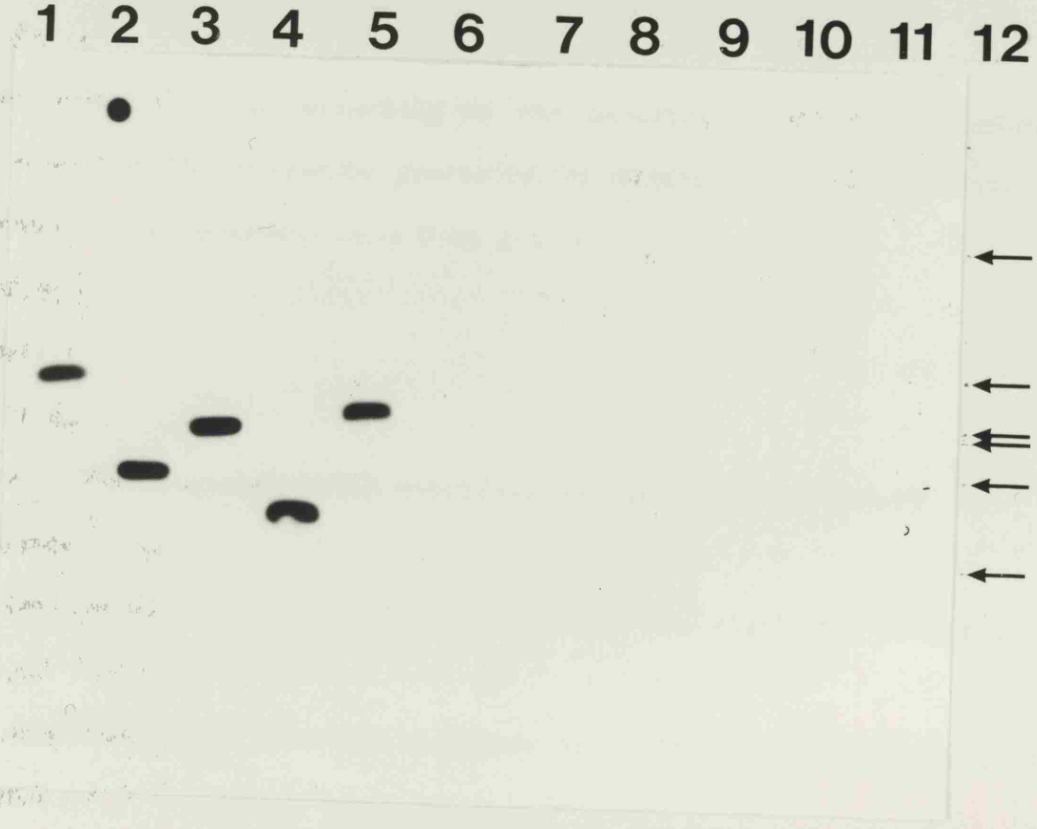


TABLE 3.1 Restriction Fragments Generated by *Bam*HI, *Eco*RI and *Hind*III Digests of λSMTI.2, λSMTIII.1 and λSMTVI.1 DNAs

*Bam*HI			*Eco*RI			*Hind*III		
1	2	3	1	2	3	1	2	3
43kb*	-	-	21.7kb	21.7kb	21.7kb	23kb	23kb	23kb
-	32kb*	-	13.5kb*	-	-	8.1kb(D)	8.1kb	8.1kb
-	-	40.5kb*	7.6kb	7.6kb	7.6kb	-	-	5.2kb*
-	10.1kb	-	-	-	6.4kb*	3.5kb*	3.5kb*(D)	-
			-	3.5kb	-	-	-	3.1kb
			-	3kb	-	-	2.2kb	-
			-	2.4kb	-	-	1.6kb	1.5kb
			-	-	2.1kb	-	0.6kb	-
			-	2kb*	2kb			
			-	1kb*(D)				

DNAs were digested according to the manufacturers recommendations. The size of the DNA fragments generated by digestion were determined by agarose gel electrophoresis on a 0.8% gel.

1= λSMTI.2

2= λSMTIII.1

3= λSMTVI.1

* Indicates DNA fragments which hybridise to the pHIF24.40 insert.

(D) Indicates doublet

*Eco*RI digested λL47, left arm fragment is 21.7kb, right arm fragment is 7.6kb

*Hind*III digested λL47, left arm fragment is 23kb, right arm fragment is 8.1kb.

one or both hybridise to pHIF24.40. This analysis suggested that there may be two IFNα related genes within this clone or an intragenic *Eco*RI site. The former conclusion has since been verified (Chapter 4). Digestion of λSMTIII.1 DNA with *Hin*dIII also produced five clone specific fragments (Fig.3.9A lane 6), including a 3.5kb doublet of which at least one hybridise to pHIF24.40 (Fig.3.9B lane 6).

λSMTVI.1 contains a 6.5kb DNA insert. Again this clone is devoid of any *Bam*HI sites, indicating both sites were disrupted by the cloning procedure and that no internal *Bam*HI sites are present. Digestion of λSMTVI.1 with *Eco*RI produced three fragments (Fig.3.10A lane 3), including a 6.4kb fragment which hybridised to the pHIF24.40 insert (Fig.3.10B lane 3). *Hin*dIII digestion of λSMTVI.1 DNA also produced 3 fragments (Fig3.10A lane 2), including a 5.2kb fragment which hybridised to pHIF24.40 (Fig.3.10B lane 2).

Therefore, this analysis demonstrated that three distinct λ DNA clones had been identified which contained segments of human DNA sequences complementary to the human IFNα pseudogene pHIF24.40 (≡ LeIF-E). Thus, this strongly suggested that at least three human IFNα structural genes had been isolated.

3.5 Conclusions

Four λL47 recombinant phage potentially containing human IFNα gene sequences have been isolated from screening approximately 300,000 recombinant phages of the newly constructed placental DNA genomic library. Three of these, λSMTI.2, λSMTIII.1 and λSMTVI.1 have been purified, DNA prepared and the presence of inserts complementary to an IFNα gene confirmed by preliminary restriction and Southern blot analysis.

Other workers have previously isolated 16 recombinant phage which hybridised to a human IFNα_1 cDNA probe (Brack *et al*, 1981) in a foetal genomic library of only 240,000 clones. The human IFNα gene family consists of at least 23 distinct loci (Henco *et al*, 1985). Therefore, five to ten distinct clones hybridising to pHIF24.40 in a library representing 50% (300,000 clones) of the human genome would be anticipated.

There may be several explanations accounting for the relatively small number of IFNα genes identified in this study. Firstly, the genomic clones characterised have small inserts despite the precautions taken to enrich for 11kb to 17kb fragments and thus, the size of the library may be overestimated. Secondly, the conditions used for *in situ* plaque hybridisation, suitable for the detection of sequences showing >85% homology to pHIF24.40, may have been too stringent. In this respect, some of the clones isolated using the human IFNα_1 gene cDNA probe hybridised only very weakly (Brack *et al*, 1981). Thirdly, the density of plaques screened in this study may also have contributed to the relatively small number of recombinant phage containing IFNα gene sequences identified. Plaque hybridisation was performed on 20,000 plaques per 9cm nitrocellulose disc (Maniatis *et al*, 1981) compared with only 3,000 plaques per 9cm nitrocellulose disc screened in other studies (Brack *et al*, 1981). Screening at a lower plaque density may give a stronger hybridisation signal allowing the identification and isolation of specific clones to be performed more readily. Finally, twenty two prospective recombinant phage were identified from the primary screen to hybridise to pHIF24.40, eighteen of which failed to hybridise on subsequent rounds of screening. It is unlikely that all were artifacts. Probably some genuine recombinant phage were lost by the failure to correctly realign the autoradiographs to the original stockplates containing the phage library. This problem is enhanced by the high density of plaques screened in this study and the tendency of nitrocellulose filters to shrink during these procedures.

CHAPTER 4

CHARACTERISATION OF HUMAN IFNα GENOMIC CLONES

Three genomic clones isolated from the human placental DNA library described in Chapter 3 were characterised in detail to verify that they contained authentic IFNα genes. For this purpose, restriction digest analysis and sub−cloning, in combination with Southern blot analysis, were undertaken to locate and map IFNα gene sequences within the genomic DNA. Confirmation that genuine IFNα genes had been isolated was obtained by sequencing studies. Through this analysis four IFNα genes have been identified and characterised in some detail.

4.1 Characterisation of an IFNα Gene in λSMTI.2

Restriction Mapping of λSMTI.2

λSMTI.2 contains a 9kb human genomic DNA insert devoid of *Eco*RI and *Bam*HI recognition sequences (Chapter 3). Restriction with *Hin*dIII generated 8.1kb and 3.5kb fragments, the latter containing IFNα gene related sequences (Chapter 3). In addition *Hin*dIII digests produced a 23kb and 8.1kb species derived from the left and right arm of λSMTI.2 respectively (Table 3.1).

Phage λSMTI.2 DNA, as with all λ recombinant phage DNAs described in this study, was initially digested with a range of restriction endonucleases and the products analysed by agarose gel electrophoresis. Digestion of DNA with *Bgl*II revealed a single site located within the insert. Therefore, *Bgl*II/*Hin*dIII double digests were performed to map the unique *Hin*dIII and *Bgl*II sites. Since *Bgl*II cleaves within the right arm of λSMTI.2 this analysis also resolved the 8.1kb *Hin*dIII doublet. These analyses were undertaken in the same manner as those shown in Figure 3.8. Table 4.1 summarises the size of restriction fragments generated by these digests. Southern blot analysis with the probe pHIF24.40 (994bp *Pst*I insert) identified fragments containing IFNα gene

related sequences. *Bgl*II digestion produced a 6.8kb DNA fragment from the right arm of λSMTI.2, common to wild type λL47 and recombinant phage. The 33kb *Bgl*II fragment from the left arm of λL47 is reduced to 27.5kb and 8.1kb fragments in λSMTI.2. Hence a single *Bgl*II site is present 8.1kb from the *Bgl*II site within the right arm of λSMTI.2 (Fig.4.1A).

Table 4.1 also shows that the λSMTI.2 27.5kb *Bgl*II fragment hybridises to pHIF24.40. Thus, IFNα gene sequences are located to the left of the insert specific *Bgl*II site (Fig.4.1A). The 3.5kb *Hin*dIII hybridising fragment must also, therefore, be flanked on one side by the *Hin*dIII site of the left arm of λSMTI.2 (Fig.4.1A).

Double digests of λSMTI.2 with *Bgl*II and *Hin*dIII confirm the map shown in Figure 4.1A. Figure 4.2 shows agarose gel electrophoresis of *Bgl*II and *Hin*dIII digested λSMTI.2 and λL47 DNA. Comparison of *Hin*dIII digested DNA (lane 3) and *Bgl*II/*Hin*dIII digested DNA (lane 4), demonstrated that the 8.1kb *Hin*dIII fragment is a doublet. *Bgl*II cleaved the 8.1kb *Hin*dIII fragment generating 6.8kb and 1.3kb fragments in both phage (Table 4.1). In addition λSMTI.2 DNA generated 7.6kb and 0.5kb fragments confirming the presence of an 8.1kb *Hin*dIII doublet, locating the *Hin*dIII site 0.5kb from the *Bgl*II site in λSMTI.2 (Fig.4.1A).

*Pvu*II recognition sequences are commonly found within the nucleotide sequences encoding the signal peptide and mature polypeptide of human IFNα genes (Goeddel *et al*, 1981). Therefore, the presence of *Pvu*II sites within λSMTI.2 was investigated. λSMTI.2 DNA was digested with *Pvu*II and *Hin*dIII/*Pvu*II and Southern blot analysis performed using the pHIF24.40 probe (*Pst*I insert). A single *Pvu*II recognition sequence was mapped within the *Hin*dIII 3.5kb fragment (Fig.4.1A). From these studies the restriction map in Figure 4.1A was derived.

Subcloning the 3.5 kb *Hin*dIII Fragment

Having demonstrated that the IFNα gene complementary sequences of λSMTI.2 lie within a 3.5kb *Hin*dIII fragment, it was subcloned into the *Hin*dIII site of pAT153 (Twigg and Sheratt, 1980) for further characterisation. Plasmid pAT153 was chosen because of its well defined restriction map. *Hin*dIII digested λSMTI.2 and pAT153 DNAs

FIGURE 4.1 Restriction Map of λSMTI.2 and pSMTI.2

A: *Alu* I, B: *Bgl*II, Ba: *Bam*HI, E: *Eco*RI, H: *Hae*III, Hi: *Hin*dIII, P: *Pvu*II, Ps: *Pst*I, R: *Rsa*I, S: *Sau*3A, Sm: *Sma*I, Ss: *Sst*I.

A) shows the restriction map of λSMTI.2 derived by restriction and Southern blot analysis. The shaded region indicates sequences hybridising to the *Pst*I insert of pHIF24.40. The arrow indicates the orientation of the IFNα gene. B) shows the restriction map of pSMTI.2 derived by restriction, Southern blot and sequencing analysis. The arrow above the map indicates the orientation of the IFNα gene, whereas those below correspond to the restriction fragments sequenced. The shaded region indicates the coding sequence of SMTI.2.

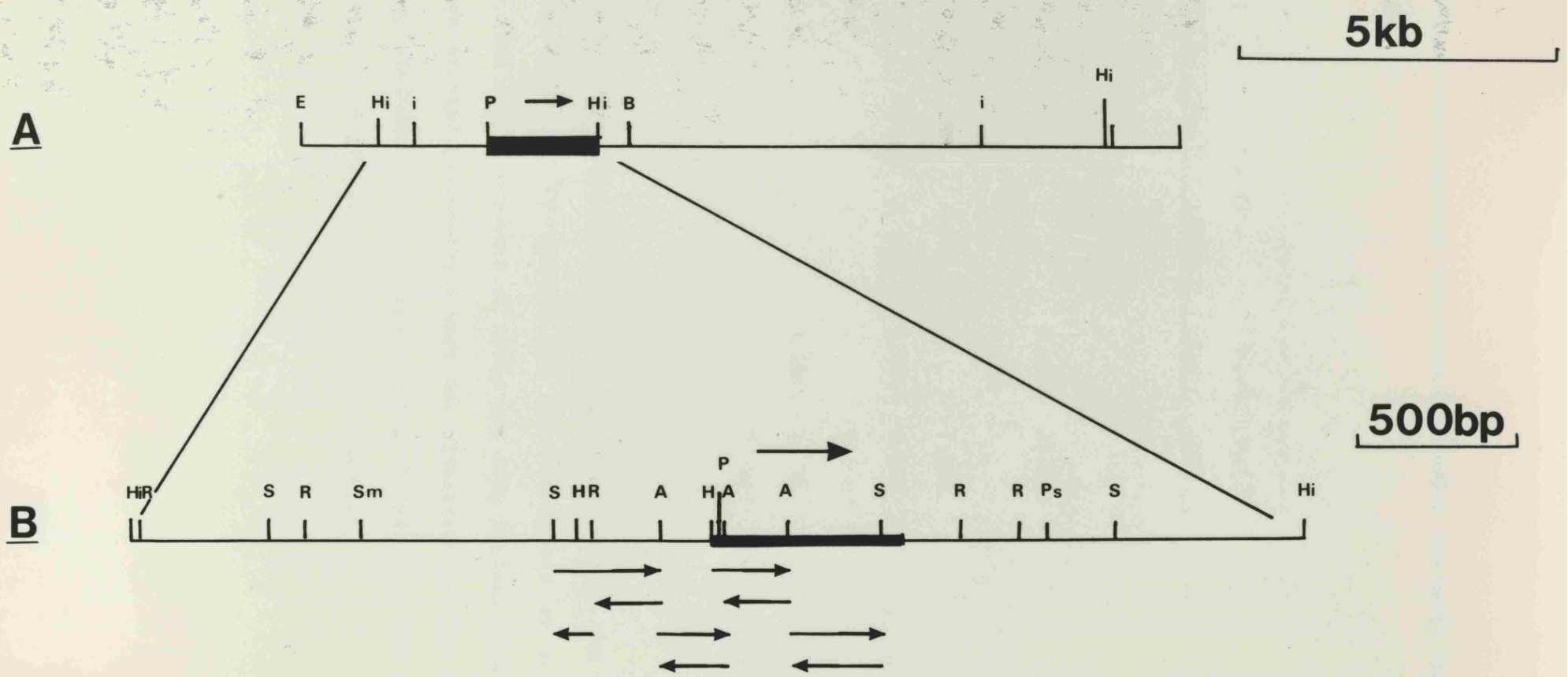

5kb
A
E
Hi
i
P
Hi
B
i
Hi
500bp
B
HiR
S
R
Sm
S
HR
A
H
P
A
A
S
R
R
Ps
S
Hi

TABLE 4.1 Restriction Fragments Generated by *Hin*dIII and *Bgl*II Digests of λSMTI.2

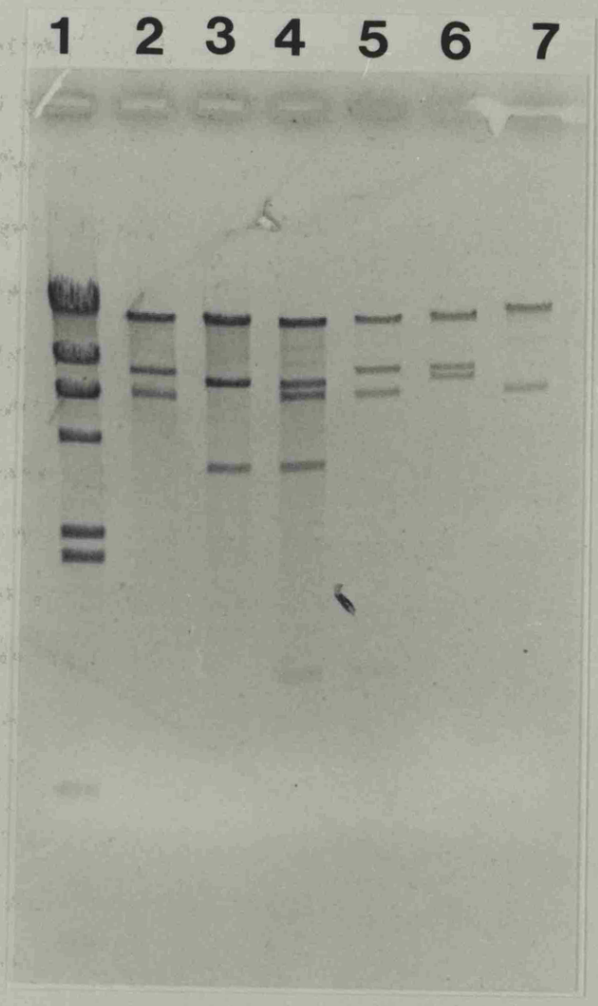
*Hin*dIII		*Bgl*II		*Hin*dIII/*Bgl*II	
λL47	λSMTI.2	λL47	λSMTI.2	λL47	λSMTI.2
-	-	33kb	27.5kb*	-	-
23kb	23kb	-	-	22.5kb	22.5kb
9.3kb	-	-	-	9.3kb	-
8.1kb	8.1kb(D)	-	8.1kb	-	7.6kb
-	-	6.8kb	6.8kb	6.8kb	6.8kb
-	3.5kb*	-	-	-	3.5kb*
-	-	-	-	1.3kb	1.3kb
-	-	-	-	-	0.5kb

The size of DNA fragments generated by digestion with the indicated enzymes were determined by agarose gel electrophoresis on a 0.8% gel (Chapter 2). Two *Bgl*II fragments of 470bp and 60bp derived from the left and right arm of λL47 respectively were not visualized in this study and have not been included in the analysis above. * indicates DNA fragments which hybridise to pHIF24.40. (D) indicates a doublet.

FIGURE 4.2 Restriction analysis of λSMTI.2 DNA with *Bgl*II and *Hin*dIII

1μg of each DNA sample was digested with *Bgl*II and/or *Hin*dIII according to the manufacturers recommendations and analysed by agarose gel electrophoresis on a 0.8% gel. The gel was stained with ethidium bromide (1μg/ml) to reveal the DNA.

1. λ *Hin*dIII Marker: 23.75kb, 9.46kb, 6.67kb, 4.26kb, 2.25kb, 1.96kb, 0.59kb and 0.1kb.
2. λSMTI.2 *Bgl*II
3. " *Hin*dIII
4. " *Bgl*II/*Hin*dIII
5. λL47 *Bgl*II
6. " *Hin*dIII
7. " *Bgl*II/*Hin*dIII


1 2 3 4 5 6 7

were ligated and *E.coli* JA221 was transformed, selecting colonies for ampicillin resistance. Prospective recombinant clones were recognised by their inability to grow on L.plates containing 10μg/ml tetracycline. Five such clones were identified. Plasmid DNA prepared from these colonies was digested with *Hin*dIII and in combination with Southern blot analysis using the pHIF24.40 (*Pst*I insert) probe, one recombinant containing the 3.5kb *Hin*dIII hybridising fragment from λSMTI.2 was identified. This plasmid was designated pSMTI.2.

Restriction Mapping pSMTI.2

Plasmid pSMTI.2 was subjected to detailed restriction analysis. Digests with *Pst*I and *Sma*I identified single sites for these enzymes within the insert. These sites were located by double digests with *Hin*dIII or *Eco*RI (Fig.4.1B).

A detailed map of the *Hin*dIII fragment was also produced for the tetranucleotide restriction endonucleases *Alu*I, *Rsa*I and *Sau*3A (Fig.4.1B). Fragments were orientated by double digests with *Hin*dIII, *Eco*RI, *Pst*I, or *Sma*I. Ambiguities were subsequently resolved by sequence analysis. Figure 4.1B shows the detailed restriction map of pSMTI.2 derived from these studies.

Primary Nucleotide Sequence of the IFNα Gene SMTI.2

The objective of these studies was to isolate and characterise human genomic IFNα genes. Therefore, to verify that pSMTI.2 contained an authentic IFNα gene, sequencing studies were undertaken.

Thus, the IFNα gene in pSMTI.2 was sequenced by the dideoxy chain termination method (Sanger *et al*, 1977). For this purpose restriction fragments generated by *Sau*3A digestion of pSMTI.2 DNA were "shot gun cloned" into the *Bam*HI site of M13mp10. *E.coli* JM101 was transfected with ligated DNA generating greater than 90% recombinants as determined by the frequency of colourless plaques identified on plates containing Bluo−gal (Chapter 2).

Recombinant M13mp10 phage containing IFNα gene sequences were identified by plaque hybridisation (Chapter 2) again using the pHIF24.40 insert as probe. Four positive

plaques were picked for further analysis. To confirm that these recombinant phage contained DNA sequences specific to pSMTI.2, R.F. DNA was prepared, digested with *Sau*3A and analysed by P.A.G.E. Two of the four plaques isolated contained a 1000bp *Sau*3A fragment, comigrating with a pSMTI.2 *Sau*3A fragment. These two phages were employed in sequencing studies. A total of 365bp of sequencing data obtained from one end of one of these was used to search the EMBL data base using the computer programme Seqfit (Staden, 1977) for IFNα related gene sequence homology. Perfect homology was established with 310bp of 5' flanking sequence from the IFNα gene λ2h (Lawn *et al*, 1981a). The remaining 53bp are 5' to the available sequence of λ2h. DNA from the other recombinant phage proved difficult to sequence and thus was not investigated further. The remaining 5' region and most of the structural gene was sequenced using the strategy indicated in Figure 4.1B. The 1000bp *Sau*3A fragment, isolated as an *Eco*RI/*Hin*dIII fragment from the recombinant M13 phage by P.A.G.E, was further digested with either *Hae*III, *Alu*I, or *Rsa*I, ligated to the *Sma*I digested M13 phage DNA and transfected into *E.coli* JM101. The proportion of recombinants varied from 50% to 100% using this procedure. Four to eight plaques were picked from each cloning and sequenced.

Thus, λSMTI.2 contains an authentic IFNα gene. Available sequence information for the gene SMTI.2 is shown in Figure 4.3, where it is compared with λ2h. A total of 974bp was determined including 393bp of 5' flanking sequence. The remaining 50bp of anticipated 3' translated and additional non−translated gene sequence was not determined. In Figure 4.3 the prospective TATTTAA promoter element, the putative CAP site and translation initiation codon are each highlighted. Clearly λ2h and SMTI.2 genes are identical except for two additional bases seen in the 5' flanking sequence of SMTI.2. SMTI.2 extends the sequence of the λ2h clone an additional 52bp upstream of the IFNα gene. Furthermore, two bases not determined in λ2h were identified in λSMTI.2 at −171 and −172. Comparison of the coding sequence of SMTI.2 with other IFNα genes shows the next closest homology is IFNα_5 which exhibits 48 base substitutions, encoding a full length IFNα polypeptide with 24 amino acid changes relative to SMTI.2 and λ2h. From this analysis it can be concluded that λSMTI.2 and λ2h (Lawn *et al*, 1982) are

FIGURE 4.3 Comparison of the Primary Nucleotide Sequence of SMTI.2 with λ2h

The nucleotide sequence determined for SMTI.2 is shown with the putative transcription initiation site, translation initiation codon and TATA box sequences underlined. The complete coding region for this gene was not determined. The SMTI.2 sequence is aligned with the available complementary sequence data for λ2h. Dashes (-) indicate homology. Spaces at -171 and -172 indicate the precise nucleotides at this location, which were not determined for λ2h. Two additional nucleotides were also revealed at -333 and -338 within SMTI.2 which are not found in λ2h.

```
  -390        -380        -370        -360        -350        -340
GATCCATCTG CCTTGGTCTC CCAAAGTGCT GGCCTTACAA GTGTGAGCCA CCGGGTCCCAG
                                                        --- -----

  -330        -320        -310        -300        -290        -280
TCCCATTTGG TCATTCTTAA CACTATGTTA AATAAAAAGA TTAAACTTTA CACTCCTTATA
 --------- ---------- ---------- ---------- ---------- ----------

  -270        -260        -250        -240        -230        -220
AATAGATATG TACAGTATAT CAACAAATAT ATGGTATGTC TGTGTTATTA AAATTTAATGG
---------- ---------- ---------- ---------- ---------- ----------

  -210        -200        -190        -180        -170        -160
GACTTTGAAT TAGAAAGAAA TTTCTAAAAA GCCCATGGGG CAGGGAAAGA TGAGGTAATAT
---------- ---------- ---------- ----------   --------- ----------

  -150        -140        -130        -120        -110        -100
ACTGAAAATA AAAGTGGTTG AGAAACTGCT CTACACCCAT GTAGACAGGA CATAAAGGAAA
---------- ---------- ---------- ---------- ---------- ----------

  -90         -80         -70         -60         -50         -40
GCCAAAAGAG AAGTAGAAAA AAACATGAAG ACGCTTCAGA AAATGGAAGC TAGTATGTTCC
---------- ---------- ---------- ---------- ---------- ----------

  -30         -20         -10                     +10         +20
TTATTTAAGA CCTATGCACA GAGCAAGGTC TTCAGAAAAC CTACAACCCA AGGTTCAGTGT
---------- ---------- ---------- ---------- ---------- ----------

  +30         +40         +50         +60         +70         +80
TTACCCCTCA TCAACCAGCC CAGCAGCATC TTCGGGATTC CCAATCGCAT TGCCCTTTGCT
---------- ---------- ---------- ------ --- ---------- ----------

  +90         +100        +110        +120        +130        +140
TTAATGATGG CCCTGGTGGT GCTCAGCTGC AAGTCAAGCT GCTCTCTGGG CTGTAATCTGT
---------- ---------- ---------- ---------- ---------- ----------

  +150        +160        +170        +180        +190        +200
CTCAAACCCA CAGCCTGAAT AACAGGAGGA CTTTGATGCT CATGGCACAA ATGAGGAGAAT
---------- ---------- ---------- ---------- ---------- ----------

  +210        +220        +230        +240        +250        +260
CTCTCCTTTC TCCTGCCTGA AGGACAGACA TGACTTTGAA TTTCCCCAGG AGGAATTTGAT
---------- ---------- ---------- ---------- ---------- ----------

  +270        +280        +290        +300        +310        +320
GGCAACCAGT TCAGAAAGCT CAAGCCATCT CTGTCCTCCA TGAGATGATG CAGCAGACCTT
---------- ---------- ---------- ---------- ---------- ----------

  +330        +340        +350        +360        +370        +380
CAATCTCTTC AGCACAAAGA ACTCATCTGC TGCTTGGGAT GAGACCCTCC TAGAAAAATTC
---------- ---------- ---------- ---------- ---------- ----------

  +390        +400        +410        +420        +430        +440
TACATTGAAC TTTTCCAGCA AATGAATGAC CTGGAAGCCT GTGTGATACA GGAGGTTGGGG
---------- ---------- ---------- ---------- ---------- ----------

  +450        +460        +470        +480        +490        +500
TGGAAGAGAC TCCCCTGATG AATGAGGACT CCATCCTGGC TGTGAAGAAA TACTTCCAAAG
---------- ---------- ---------- ---------- ---------- ----------

  +510        +520        +530        +540        +550        +560
AATCACTCTT TATCTGATGG AGAAGAAATA CAGCCCTTGT GCCTGGGAGG TTGTCAGAGCA
---------- ---------- ---------- ---------- ---------- ----------

  +570        +580
GAAATCATGA GATC
---------- ----
```

almost certainly alleles.

4.2 Identification of Two IFNα Genes in λSMTIII.1

Restriction Mapping of λSMTIII.1

The genomic clone λSMTIII.1 contains an 8kb DNA insert in λL47 (Chapter 3). Preliminary restriction digests indicated a complex pattern of *Hin*dIII and *Eco*RI sites within the genomic DNA (Chapter 3). Furthermore the cloning procedure recreated the *Bam*HI site at the right arm of λSMTIII.1 (Chapter 3). This feature was used to orientate *Eco*RI and *Hin*dIII fragments derived from the right arm.

Restriction digestion of λSMTIII.1 performed with *Bgl*II again defined a single site within the insert. This site was 7kb from the λSMTIII.1 right arm *Bgl*II site (Fig.4.4A). Southern blot analysis revealed that both the 7kb and 28kb *Bgl*II fragments derived from the right and left arms of λSMTIII.1, respectively, hybridised to pHIF24.40 (*Pst*I insert) and hence contained IFNα gene related sequences. Southern blot analysis of *Eco*RI digested λSMTIII.1 DNA also identified two fragments of 2kb and 1kb hybridising to pHIF24.40 (Chapter 3). Therefore, either *Bgl*II and *Eco*RI sites occur within a single IFNα gene present in λSMTIII.1 or two genes are present.

To determine the number of IFNα genes, Southern blot analysis of *Eco*RI/ *Bgl*II digested DNA was performed. Figure 4.5 shows the banding pattern generated by *Bgl*II and *Eco*RI digested λL47 and λSMTIII.1 DNAs after agarose gel electrophoresis. Table 4.2 shows the size of the specific DNA fragments identified by this analysis.

This analysis demonstrated that two separate regions within λSMTIII.1 contain IFNα gene sequences. *Bgl*II cleaved a 3kb *Eco*RI fragment which does not hybridise to pHIF24.40. Therefore, the *Bgl*II site cannot lie within the IFNα gene. The two *Eco*RI generated hybridising fragments, 2kb and 1kb, are not digested by *Bgl*II. As fragments derived from both sides of the *Bgl*II site hybridised to pHIF24.40 the two *Eco*RI fragments are separated, located on either side of the *Bgl*II site.

Southern blot analysis of *Hin*dIII digested λSMTIII.I suggested that a 3.5kb fragment hybridised to the pHIF24.40 insert (Chapter 3). However, this fragment is a doublet

FIGURE 4.4 Restriction Maps for λSMTIII.1 and pSMTIII.1

See legend to Fig.4.1. A) shows the restriction map of λSMTIII.1 B) shows the restriction map of pSMTIII.1_A. * Not all *Alu*I sites are shown. The *Hind*III fragment is shown in the opposite orientation to that displayed in A.

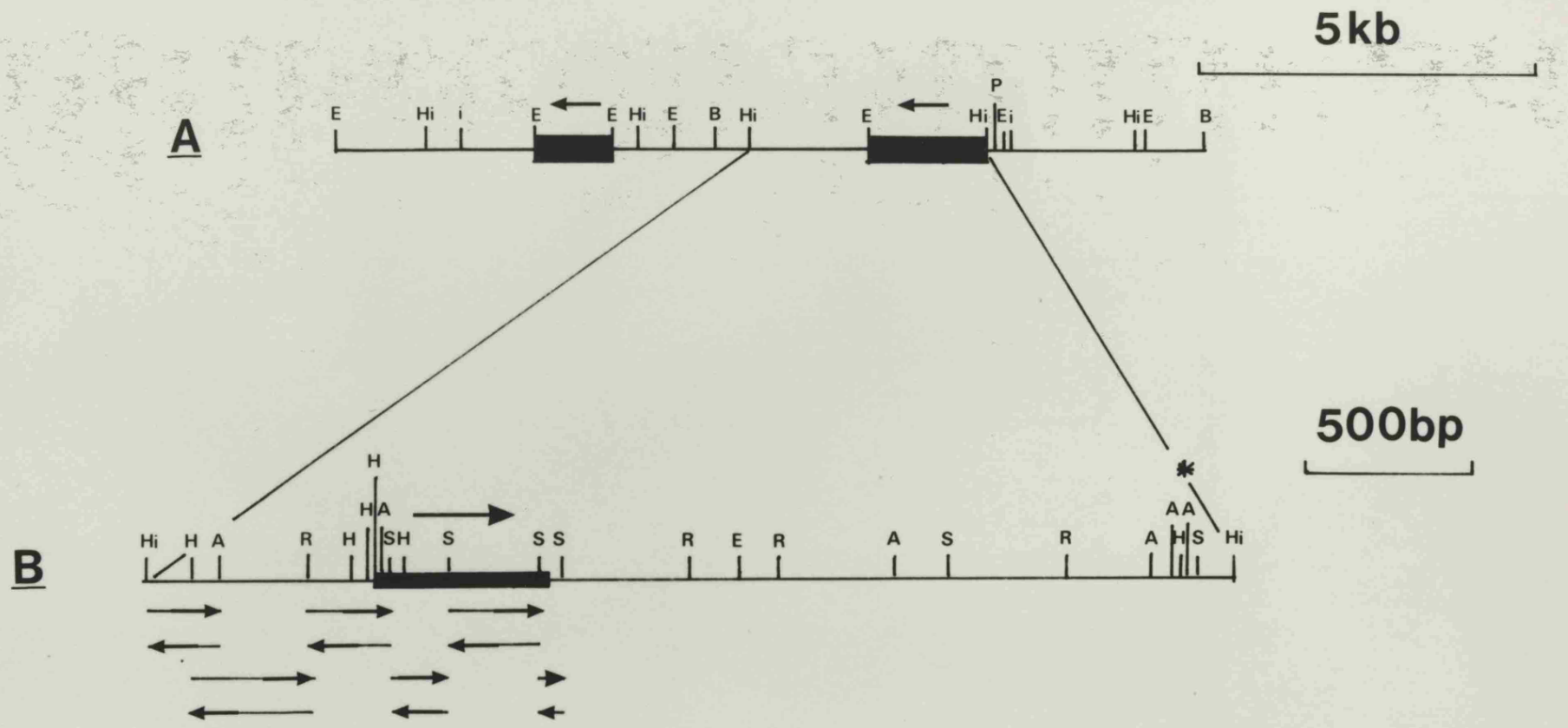

5kb
A
E
Hi
i
E
E
Hi
E
B
Hi
E
Hi
P
Ei
HiE
B
500bp
B
H
HA
Hi
H
A
R
H
SH
S
S
S
R
E
R
A
S
R
A
A
H
S
Hi

FIGURE 4.5 Southern Blot Analysis of *Bgl*II and *Eco*RI Digested λSMTIII.1 DNA with pHIF24.40 (*Pst*I insert)

1μg of each DNA sample was digested with *Bgl*II and/or *Eco*RI and analysed by agarose gel electrophoresis on a 0.8% gel. The gel was stained with ethidium bromide to visualise the DNA (A). DNA was transferred to a nitrocellulose membrane as described in Chapter 2 and then hybridisation performed using the nick translated *Pst*I insert of pHIF24.40 (1×10^8cpm/μg DNA) as probe. Autoradiography was for 5hrs using Fuji RX film at -70°C with an intensifying screen (B). Arrows indicate the size of specific hybridising bands.

1. λ *Hin*dIII Marker (see legend to Fig.4.2)
2. λSMTIII.1 *Bgl*II
3. " *Eco*RI
4. " *Bgl*II/*Eco*RI
5. λL47 *Bgl*II
6. " *Eco*RI
7. " *Bgl*II/*Eco*RI

A

B

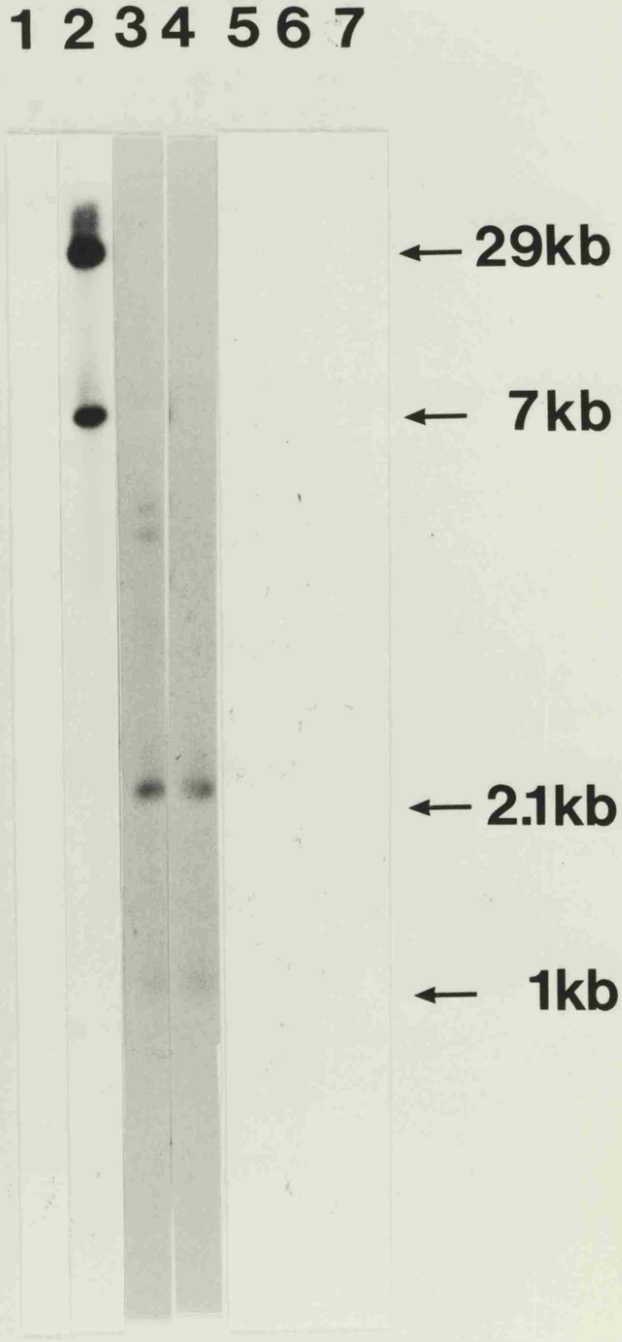


TABLE 4.2 Restriction Fragments Generated by *Eco*RI and *Bgl*II Digests of λSMTIII.1

*Eco*RI		*Bgl*II		*Eco*RI/*Bgl*II	
λL47	λSMTIII.1	λL47	λSMTIII.1	λL47	λSMTIII.1
-	-	33kb	28kb*	-	-
21.7kb	21.7kb	-	-	21.7kb	21.7kb
11.3kb	-	-	-	11.3kb	-
7.6kb	7.6kb	-	-	-	-
-	-	6.8kb	6.8kb(D)*	6.8kb	6.8kb
-	3.5kb	-	-	-	3.5kb
-	3kb	-	-	-	-
-	2.4kb	-	-	-	2.1(D)*
-	2kb*	-	-	-	2kb
-	1kb(D)*	-	-	-	1kb(D)*
-	-	-	-	-	0.9kb
-	-	-	-	0.7kb	0.7kb

See legend to Table 4.1.

(Chapter 3). As shown in Figure 4.4A the two 3.5kb *Hin*dIII fragments are derived from either side of the *Bgl*II site and thus the IFNα genes must be located on two separate 3.5kb *Hin*dIII fragments (Fig.4.4A).

As described before, *Pvu*II sites are common within IFNα genes. Therefore, λSMTIII.1 DNA was analysed for *Pvu*II recognition sequences. Southern blot analysis of *Pvu*II/*Eco*RI digested λSMTIII.1, with the pHIF24.40 insert, revealed a *Pvu*II site 100bp from one end of the 2kb *Eco*RI hybridising fragment.

Further double digests were performed to orientate the *Hin*dIII and *Eco*RI sites. By these means the restriction map in Figure 4.4A was generated. The location of two IFNα gene related sequences within the genomic DNA of λSMTIII.1, designated SMTIII.1_A and SMTIII.1_B, are shown (Fig.4.4A).

Subcloning the 3.5kb *Hin*dIII Fragment Containing the SMTIII.1_A Gene

Having established that the IFNα complementary sequences in λSMTIII.1 both lay within 3.5kb *Hin*dIII, steps were taken to subclone them into pAT153 to allow more detailed characterisation. Thus, *Hin*dIII digested λSMTIII.1 and pAT153 DNAs were ligated and transformed into *E.coli* JA221. Transformants were selected for ampicillin resistance and prospective recombinants identified as described previously. Four ampicillin resistant, tetracycline sensitive colonies were obtained. The plasmid DNA isolated from one of these colonies contained a 3.5kb *Hin*dIII DNA insert with homology for the pHIF24.40 *Pst*I DNA insert. Thus, one of the IFNα complementary regions from λSMTIII.1 had been isolated. This plasmid was designated pSMTIII.1_A.

Restriction Mapping pSMTIII.I_A

Plasmid pSMTIII.1_A was subjected to detailed restriction analysis. Digests with *Eco*RI and *Ava*II revealed a single site for *Eco*RI and two sites for *Ava*II respectively (Fig.4.4B). As only one of the two possible 3.5kb *Hin*dIII fragments contained an *Eco*RI site its presence established that the SMTIII.1_A gene sequences had been isolated.

A detailed restriction map of pSMTIII.1_A was contructed for *Sau*3A, *Rsa*I, *Alu*I, and *Hae*III restriction endonucleases (Fig.4.4B). Ambiguities were ultimately resolved by

subsequent sequence studies (next section).

From the restriction studies performed with λSMTIII.1 DNA described above, it was not possible to precisely localise the *Pvu*II site close to gene SMTIII.1_A however as *Pvu*II does not digest pSMTIII.1_A DNA the *Pvu*II site must lie just outside the 3.5kb *Hin*dIII fragment, in the position shown in Figure 4.4A.

Thus, restriction mapping and Southern blot analysis have located two prospective IFNα genes in the recombinant phage DNA λSMTIII.1 (Fig.4.4A). The 3.5kb *Hin*dIII fragment encompassing the SMTIII.1_A gene has been subcloned and mapped in more detail (Fig.4.4B).

Primary Nucleotide Sequence of the IFNα Gene SMTIII.1_A

Initially the same approach which had been employed to obtain preliminary sequence data for SMTI.2 was adopted. Thus, fragments from *Sau*3A digested pSMTIII.1_A DNA were "shot gun cloned" into the *Bam*HI site of M13mp10 DNA. *E.coli* JM101 was transfected with ligated DNA generating approximately 80% recombinants.

Plaques were screened to identify recombinants containing IFNα gene complementary sequences by plaque hybridisation (see λSMTI.2). Initially five plaques hybridising to pIF24.40 (*Pst*I insert) were isolated. R.F. DNA was prepared, digested with *Sau*3A and analysed by P.A.G.E. All five clones contained *Sau*3A fragments, two of which comigrated with a 271bp fragment, two with a 178bp fragment and one with a 70bp fragment of *Sau*3A digested pSMTIII.1_A DNA.

Recombinant phage were sequenced to determine if they contained segments of an IFNα gene. This procedure revealed that the 178bp fragment had been isolated in both orientations, while the 271bp fragments were both in the same orientation. The sequence data obtained was used to search the EMBL data base for homology to human IFNα sequences using the Seqfit programme (Staden, 1977). This demonstrated that the *Sau*3A fragments were probably co−linear in that together they exhibit a continuous homology of 523bp with the coding sequence of several human IFNα genes.

The strategy employed to sequence the remainder of the SMTIII.1_A gene is shown in Figure 4.4B. The 5' clones were isolated initially by sub−cloning the *Hin*dIII/*Sau*3A

700bp fragment from pSMTIII.1_A DNA (Fig.4.4B) into the *Hin*dIII/*Bam*HI site of M13mp8. Recombinant phage with DNA inserts corresponding to the 5' region of the SMTIII.1_A gene were identified by restriction analysis. R.F DNA was prepared from eight prospective recombinant phage, digested with *Hin*dIII/*Sau*3A, and analysed by P.A.G.E. Three recombinants containing a 700bp fragment comigrating with the *Hin*dIII/*Sau*3A fragment from pSMTIII.1_A were identified.

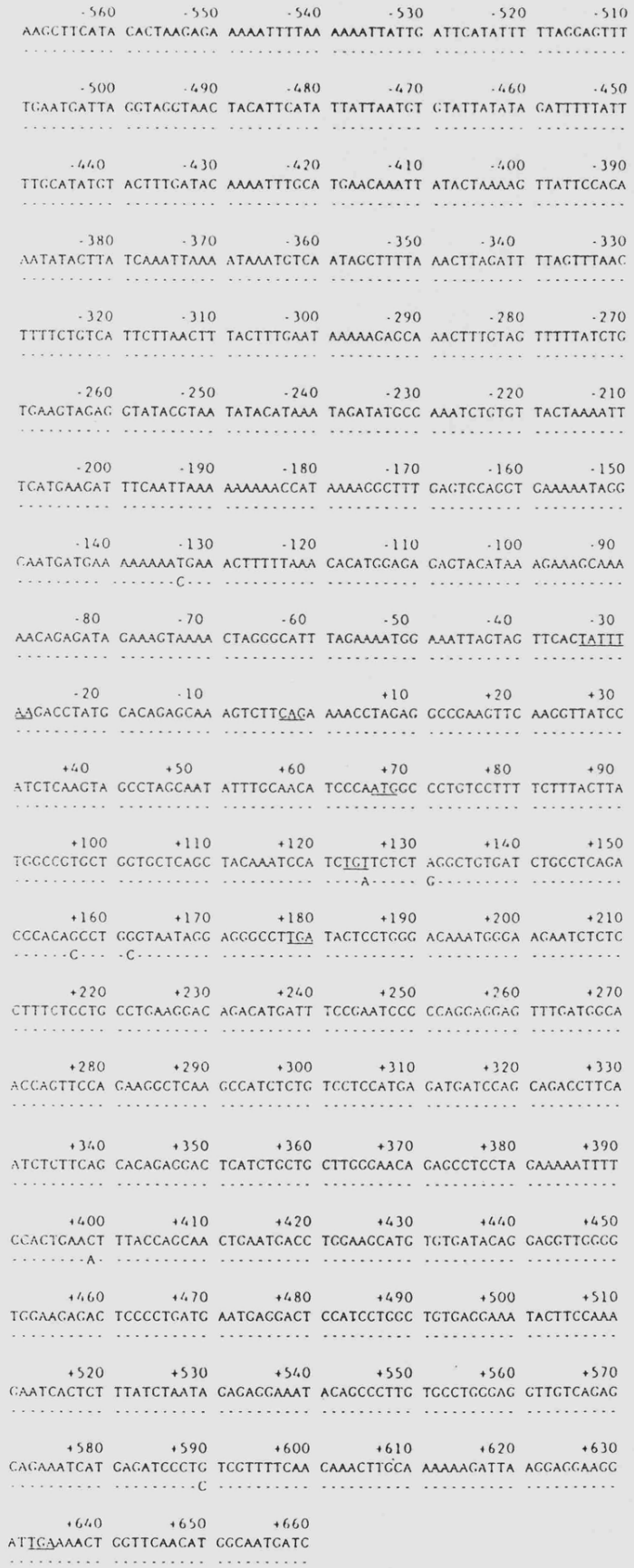
Sequencing information covering 214bp from the *Hin*dIII site of this clone was obtained to the *Alu*I site shown in Figure 4.4B. The data obtained was used to search the EMBL database using Seqfit (Staden, 1977). Perfect homology to the pseudogene ψLeIF−L (Ullrich *et al*, 1982) was observed.

The remainder of the 5' region of the λSMTIII.1 gene was determined to confirm the homology observed with the previously characterised gene ψLeIF−L. Thus, the 700bp *Hin*dIII/*Sau*3A fragment was isolated from *Eco*RI/*Hin*dIII digested M13mp8 recombinant DNA by P.A.G.E, digested with either *Alu*I, *Hae*III, or *Rsa*I and ligated to *Sma*I digested M13 DNA vectors. *E.coli* JM101 was transfected with ligated DNA generating 30% to 100% recombinants. DNA was prepared from four prospective recombinant phage for each cloning and sequenced. An additonal 707bp of sequence was obtained by this procedure. In total 680bp of 5' flanking DNA sequence and the entire coding sequence of the SMTIII.1_A IFN gene was determined.

The total available sequence information on SMTIII.1_A is shown in Figure 4.6, where it is compared with the complementary region of ψLeIF−L (Ullrich *et al*, 1982). The considerable homology shared throughout the coding sequence and 5' flanking sequence extending to −567 confirms that these genes are almost certainly alleles. The promoter element TATTTAA is underlined, along with the putative CAP site, and the IFNα gene initiation and termination codons (Fig.4.6). Comparison of SMTIII.1_A with other IFNα genes demonstrates that only 11 and 12 base substitutions are found in the coding sequences of λ2h and IFNα_{4b} respectively, each encoding potential full length IFNα polypeptides with 8 amino acid substitutions relative to SMTIII.1_A. Clearly these genes are closely related but within 628bp of 5' flanking sequence 53 and 24 nucleotide changes are apparent in λ2h and IFNα_{4b} respectively, relative to SMTIII.1_A. Thus, this

FIGURE 4.6 Comparison of the Primary Nucleotide Sequence of SMTIII.1_A with LeIF-L (Ullrich *et al*, 1982)

See legend to Fig.4.3 (also underlined is the translation termination codon). Extensive 5' flanking sequence and the entire coding sequence determined for SMTIII.1_A is shown. LeIF-L sequences are compared below. Dashes (-) indicate sequence homology. Nucleotide substitutions are shown by conventional symbols. Codon 20 of the signal peptide is underlined illustrating the single base transition found in SMTIII.1_A which changes the nonsense codon of LeIF-L to a triplet encoding cysteine.


-560 -550 -540 -530 -520 -510
AAGCTTCATA CACTAAGAGA AAAATTTTAA AAAATTATTG ATTCATATTT TTAGGAGTTT
-500 -490 -480 -470 -460 -450
TGAATGATTA GGTAGGTAAC TACATTCATA TTATTAATGT GTATTATATA GATTTTTATT
-440 -430 -420 -410 -400 -390
TTGCATATGT ACTTTGATAC AAAATTTGCA TGAACAAATT ATACTAAAAG TTATTCCACA
-380 -370 -360 -350 -340 -330
AATATACTTA TCAAATTAAA ATAAATGTCA ATAGCTTTTA AACTTAGATT TTAGTTTAAC
-320 -310 -300 -290 -280 -270
TTTTCTGTCA TTCTTAACTT TACTTTGAAT AAAAAGAGCA AACTTTGTAG TTTTTATCTG
-260 -250 -240 -230 -220 -210
TGAAGTAGAG GTATACGTAA TATACATAAA TAGATATGCC AAATCTGTGT TACTAAAATT
-200 -190 -180 -170 -160 -150
TCATGAAGAT TTCAATTAAA AAAAAACCAT AAAAGGCTTT GAGTGCAGGT GAAAAATAGG
-140 -130 -120 -110 -100 -90
CAATGATGAA AAAAAATGAA ACTTTTTAAA CACATGGAGA GAGTACATAA AGAAAGCAAA
-------C---
-80 -70 -60 -50 -40 -30
AACAGAGATA GAAAGTAAAA CTAGGGCATT TAGAAAATGG AAATTAGTAG TTCACTATTT
-20 -10 +10 +20 +30
AAGACCTATG CACAGAGCAA AGTCTTCAGA AAACCTAGAG GCCGAAGTTC AAGGTTATCC
+40 +50 +60 +70 +80 +90
ATCTCAAGTA GCCTAGCAAT ATTTGCAACA TCCCAATGGC CCTGTCCTTT TCTTTAGTTA
+100 +110 +120 +130 +140 +150
TGGCCGTGCT GGTGCTCAGC TACAAATCCA TCTGTTCTCT AGGCTGTGAT CTGCCTCAGA
----A----- G---------
+160 +170 +180 +190 +200 +210
CCCACAGCCT GGGTAATAGG AGGGCCTTCA TAGTCCTGGG ACAAATGGGA AGAATCTCTC
------C--- -C-------- ---------- ---------- ---------- ----------
+220 +230 +240 +250 +260 +270
CTTTCTCCTG CCTGAAGGAC AGACATGATT TCCGAATCCC CCAGCAGGAG TTTGATGGCA
+280 +290 +300 +310 +320 +330
ACCAGTTCCA GAAGGCTCAA GCCATCTCTG TCCTCCATGA GATGATCCAG CAGACCTTCA
+340 +350 +360 +370 +380 +390
ATCTCTTCAG CACAGACGAC TCATCTGCTG CTTGCGAACA GAGCCTCCTA GAAAAATTTT
+400 +410 +420 +430 +440 +450
CCACTGAACT TTACCAGCAA CTGAATGACC TGGAAGCATG TGTGATACAG GAGGTTGGGG
--------A- ---------- ---------- ---------- ---------- ----------
+460 +470 +480 +490 +500 +510
TGGAAGAGAC TCCCCTGATG AATGAGGACT CCATCCTGGC TGTGAGGAAA TACTTCCAAA
+520 +530 +540 +550 +560 +570
GAATCACTCT TTATCTAATA GAGAGGAAAT ACAGCCCTTG TGCCTGCGAG GTTGTCAGAG
+580 +590 +600 +610 +620 +630
CAGAAATCAT GAGATCCCTG TCGTTTTCAA CAAACTTGCA AAAAAGATTA AGGAGGAAGG
---------C
+640 +650 +660
ATTGAAAACT GGTTCAACAT GGCAATGATC

strongly supports the assignment of SMTIII.1 as an allele of ψLeIF−L rather than of these other closely related IFN genes.

Inspection of SMTIII.1_A and ψLeIF−L sequences in Figure 4.6 reveals six base substitutions within the coding sequence. The base substitutions occur at codon 20 and 22 (coding for the signal peptide), codon 8, codon 10, codon 89 and codon 155 (coding for the mature IFNα polypeptide). Codon 20 in the signal peptide of ψLeIF−L is a translation termination codon. Hence ψLeIF−L has been designated a pseudogene. However, SMTIII.1_A contains a single base substitution in this region converting the UGA termination codon (ψLeIF−L) to UGU coding for the amino acid cysteine. This amino acid is highly conserved in the other human IFNα genes. The base substitution at codon 22 is a conservative change from CUG (ψLeIF−L) to CUA (SMTIII.1_A) coding for the amino acid leucine. Further base substitutions at codon 8, 10 and 89 produce amino acid substitutions of threonine to serine, arginine to glycine, and isoleucine to leucine in ψLeIF−L and SMTIII.1_A respectively. These amino acids are also commonly found in other IFNα genes at these locations. Finally the base substitution in codon 155 is a conservative change of CUC (ψLeIF−L) to CUG (SMTIII.1_A) specific for the amino acid leucine.

Thus, the clone λSMTIII.1 contains an authentic IFNα gene, SMTIII.1_A, showing considerable homology with the pseudogene ψLeIF−L. However, unlike ψLeIF−L, the SMTIII.1_A gene potentially encodes a full length IFNα polypeptide and, therefore, is not a pseudogene.

Partial Sequence of the SMTIII.I_B Gene

The data presented above (Section 4.2) demonstrates that the λSMTIII.1_A gene is equivalent to ψLeIF−L. ψLeIF−L has been characterised previously in a 9kb region of the human genome within the clone λHLeIF1, which encompasses one partial and two complete IFNα genes, ψLeIF−M, ψLeIF−L, and LeIF−J organised in tandem (Ullrich *et al*, 1982). Inspection of the restriction map for λHLeIF1 indicated the LeIF−J gene is downstream of ψLeIF−L. By analogy the gene downstream of SMTIII.1_A, SMTIII.1_B, seemed likely to correspond to LeIF−J.

To determine whether SMTIII.1_B and LeIF−J were indeed alleles, preliminary studies were performed to characterise SMTIII.1_B. Inspection of the restriction map for λSMTIII.1 shows that SMTIII.1_B gene sequences are partly located on a 1kb *Eco*RI fragment (Fig.4.4A). Therefore, λSMTIII.1 DNA was digested with *Eco*RI and cloned into similarly digested M13mp10 and mp11 DNA.

A total of eight prospective recombinant phage were sequenced to see if they contained IFNα gene sequences. Sequence data generated was compared to LeIF−J using the Seqfit program (Staden, 1977). A total of 73bp of DNA sequence from one clone was identical to part of the coding sequence of LeIF−J. The sequence homology extended from an *Eco*RI site at +238 (relative to the putative CAP site) 5' toward the promoter region. This *Eco*RI site is unique to LeIF−J amongst currently characterised IFNα genes as is the amino acid glutamic acid at position 35 (mature polypeptide) within the sequenced region. The closest homology other than LeIF−J is observed with λ2h, LeIF−L, IFNα_{4b} and LeIF−F, all of which show 3 base substitutions relative to the SMTIII.1_B primary sequence within the 73bp region characterised.

Thus, preliminary evidence supports the conclusion that SMTIII.1_B is complementary to LeIF−J. It therefore seems probable that the segment of the human genome cloned in λSMTIII.1 is homologous to that in λHLeIF1 (Ullrich *et al*, 1982). This conclusion is borne out by the close similarity observed for the restriction map of λSMTIII.1 with a computer generated restriction map (Analyseq; Staden, 1984) of λHLeIF1 (Fig.4.7).

4.3 Characterisation of an IFNα Gene in λSMTVI.1

Restriction Mapping of λSMTVI.1

The genomic clone λSMTVI.1 contains a 7.1kb DNA insert which has several restriction sites for *Eco*RI, *Hin*dIII, and *Bgl*II (Chapter 3). The size of DNA fragments generated with these restriction endonucleases from the analysis shown in Figure 3.10 are indicated in Table 4.3. Also shown are fragments hybridising to the pHIF24.40 insert as determined by Southern blot analysis.

Table 4.3 shows that *Bgl*II digestion of λSMTVI.1 DNA generated two clone specific

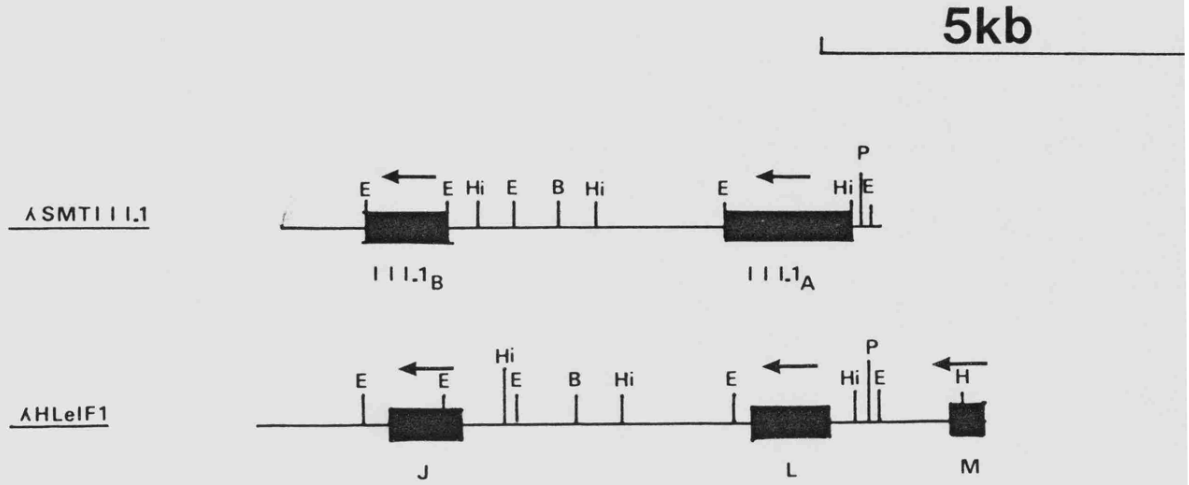


FIGURE 4.7 Comparison of the Restriction Maps of λSMTIII.1 and λHLeIF1 (Ullrich *et al*, 1982)

See legend to Fig.4.1. The restriction map generated for λSMTIII.1 is shown compared with a computer generated restriction map of λHLeIF1 derived from the EMBL data base. The location (block) and orientation (arrow) of the SMTIII.1_A and III.1_B genes are displayed as are the transcription units of the LeIF-M, LeIF-L and LeIF-J genes of λHLeIF1.

fragments of 8.5kb and 1.2kb. Table 4.3 also indicates that the 8.5kb *Bgl*II fragment derived from the right arm of λSMTVI.1 hybridised to pHIF24.40, demonstrating that it contains IFNα gene related sequences. The second *Bgl*II site is 1.2kb from this site (Fig.4.8A).

*Bgl*II/*Hin*dIII double digests were performed to orientate the *Hin*dIII sites in λSMTVI.1 (Table 4.4). *Bgl*II cleaved two clone specific *Hin*dIII fragments. One of these, the 5.2kb *Hin*dIII pHIF24.40 complementary fragment, generated a 4.2kb fragment hybridising to the pHIF24.40 insert and a 0.9kb fragment which did not hybridise (Table 4.4). This located a *Hin*dIII site 4.2kb from the *Bgl*II site, within the 8.5kb *Bgl*II hybridising fragment (Fig.4.8A). Table 4.4 also shows that *Bgl*II cleaved the 1.5kb *Hin*dIII fragment, generating a 1.4kb fragment. This located the 1.5 kb *Hin*dIII fragment proximal to the left arm of λSMTIII.1, 100bp from the second *Bgl*II site (Fig.4.8A). Therefore, the remaining 3.1kb *Hin*dIII fragment must be proximal to the right arm of λSMTVI.1 (Fig.4.8A).

A similar analysis was undertaken to locate the *Eco*RI sites within λSMTVI.1. Both *Bgl*II/*Eco*RI and *Hin*dIII/*Eco*RI double digests were performed. In addition advantage was taken of a single *Sst*I site recognised and located by *Hin*dIII/*Sst*I digests, as shown in Figure 4.8A. *Sst*I cleaved the 6.4 kb *Eco*RI hybridising fragment, mapping this fragment to the right arm of λSMTIII.1 (Fig.4.8A).

Figure 4.9 shows the result of Southern blot analysis of *Eco*RI/*Hin*dIII double digested λSMTVI.1 DNA, using the pHIF24.40 insert probe. These data show that *Eco*RI cleaves the 5.2kb *Hin*dIII hybridising fragment, generating a 3.0kb hybridising fragment containing the IFNα gene complementary sequences of λSMTVI.1 DNA. Restriction and Southern blot analysis of λSMTVI.1 has, therefore, generated the restriction map shown in Figure 4.8A.

Subcloning the 3 kb *Eco*RI/ *Hin*dIII λSMTVI.1 Fragment

To obtain a detailed map of the physical organisation of the DNA encompassing the IFNα gene complementary sequences in λSMTVI.1, the 3kb *Eco*RI/*Hin*dIII fragment with homology for the pHIF24.40 insert was subcloned into pAT153. *Eco*RI/ *Hin*dIII digested

TABLE 4.3 Restriction Fragments Generated by *Eco*RI, *Hin*dIII *and Bgl*II digests of λSMTVI.1

*Eco*RI		*Hin*dIII		*Bgl*II	
λL47	λSMTVI.1	λL47	λSMTVI.1	λL47	λSMTVI.1
21.7kb	21.7kb	23kb	23kb	33kb	24.5kb
11.3kb	-	9.3kb	-		8.5kb*
7.6kb	7.6kb	8.1kb	8.1kb	6.8kb	6.8kb
	6.4kb*	-	5.2kb*	-	1.2kb
-	2.1kb	-	3.1kb		
-	2kb	-	1.5kb		

See legend to Table 4.1

FIGURE 4.8 Restriction Map of λSMTVI.1 and pSMTVI.1

See legend to Fig.4.1. A) shows the restriction map for λSMTVI.1. B) shows the restriction map for pSMTVI.1. Not all *Rsa*I sites are shown.

* The relative orientation of these *Hae*III and *Rsa*I sites has not been resolved.

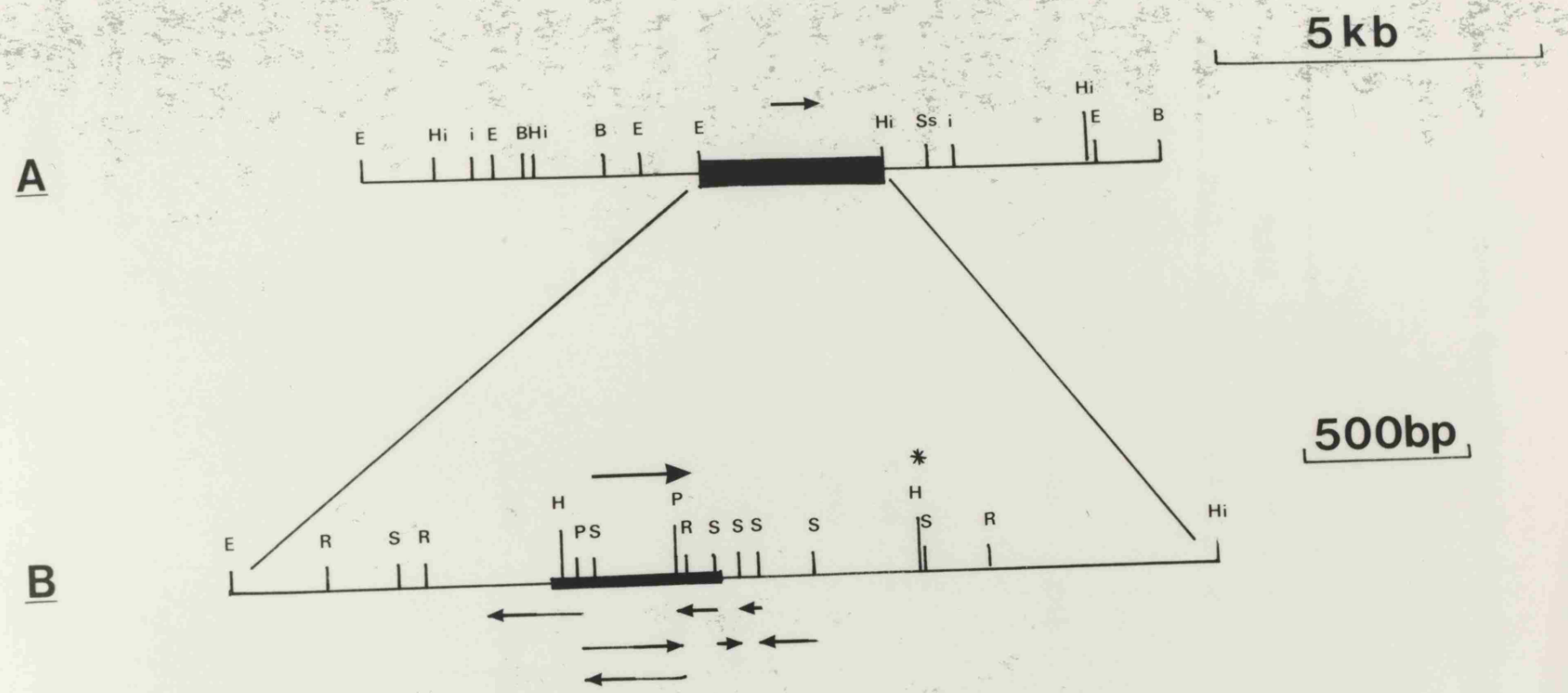

5kb
A
E
Hi
i
E
BHi
B
E
E
Hi
Ss
i
Hi
E
B
500bp
B
E
R
S
R
H
P
S
R
S
S
S
S
*
H
S
R
Hi

TABLE 4.4 Restriction Fragments Generated by *Hin*dIII and *Bgl*II Digestion of λSMTVI.1 DNA

*Hin*dIII		*Bgl*II		*Hin*dIII/*Bgl*II	
λL47	λSMTVI.1	λL47	λSMTVI.1	λL47	λSMTVI.1
-	-	33kb	24.5kb	-	-
23kb	23kb	-	-	23kb	23kb
9.3kb	-	-	-	-	-
-	-	-	8.5kb*	-	-
8.1kb	8.1kb	-	-	8.1kb	-
-	-	6.8kb	6.8kb	6.8kb	6.8kb
-	5.2kb*	-	-	-	-
-	-	-	-	-	4.2kb*
-	3.1kb	-	-	-	3.1kb
-	1.5kb	-	-	-	-
-	-	-	-	-	1.4kb
-	-	-	-	0.9kb	0.9kb(D)

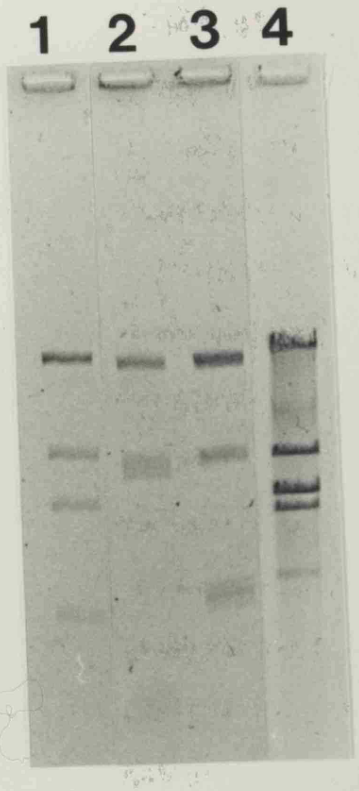
See legend to Table 4.1

FIGURE 4.9 Southern Blot Analysis of *Eco*RI and *Hin*dIII Digested λSMTVI.1 DNA with pHIF24.40 (*Pst*I insert)

See legend to Fig.4.5. A) shows an ethidium bromide stained 0.8% agarose gel containing the samples indicated below. Autoradiography was for 5hrs using Fuji RX film at -70°C with an intensifying screen (B).

1. *Hin*dIII λSMTVI.1
2. *Eco*RI "
3. *Hin*dIII/*Eco*RI λSMTVI.1

A



B

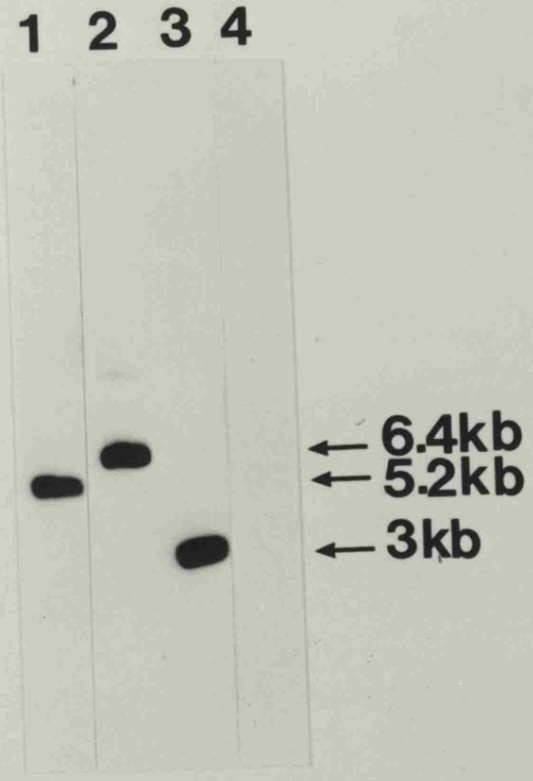


FIGURE 4.10 Comparison of the Primary Nucleotide Sequence of SMTVI.1 with IFNα_6 (Henco *et al*, 1985)

See legend to Fig.4.3 (also underlined is the translation termination codon). Extensive 5' and 3' flanking sequence and the entire coding sequence determined for SMTVI.1 is shown. IFNα_6 sequences are compared below. Dashes (–) indicate sequence homology. Nucleotide substitutions are shown by conventional symbols.

```
      -120        -110        -100        -90         -80         -70
TGAAAAACTA CTCTATACCC ATGTAGAGAG TAAATAAATG AAAGCAAAAT CAGACGTAGA
---------- ---------- ---------- ---------- ---------- ----------

       -60         -50         -40         -30         -20         -10
AAGTAAATTC TGAAAATGGA AACTAGTATG TTCCCTATTT AAGACCTACA CATAAAGCAA
---------- ---------- ---------- ---------- ---------- ----------

                   +10         +20         +30         +40         +50
GGTCTTCAGA GAACCTAGAG CTGAAGGTTC AGAGTCACCC ATCTCAACAA GTCCAACAGC
---------- ---------- ---------- ---------- ---------- ----------

       +60         +70         +80         +90        +100        +110
ATCTGCAACA TCTACAATGG CTTTGCCTTT TGCTTTACTG ATGGCCCTGG TGGTGCTCAG
---------- ---------- ---------- ---------- ---------- ----------

      +120        +130        +140        +150        +160        +170
CTGCAAGTCA AGCTGCTCTC TGGACTGTGA TCTGCCTCAG ACCCACAGCC TGGGTCACAG
---------- ---------- ---------- ---------- ---------- ----------

      +180        +190        +200        +210        +220        +230
GAGGACCATG ATGCTCCTGG CACAAATGAG GAGAATCTCT CTTTTCTCCT GTCTGAAGGA
---------- ---------- ---------- ---------- ---------- ----------

      +240        +250        +260        +270        +280        +290
CAGACATGAC TTCAGATTTC CCCAGGAGGA GTTTGATGGC AACCAGTTCC AGAAGGCTGA
---------- ---------- ---------- ---------- ---------- ----------

      +300        +310        +320        +330        +340        +350
AGCCATCTCT GTCCTCCATG AGGTGATTCA GCAGACCTTC AACCTCTTCA GCACAAAGGA
---------- ---------- ---------- ---------- --T------- ----------

      +360        +370        +380        +390        +400        +410
CTCATCTGTT GCTTGGGATG AGAGGCTTCT AGACAAACTC TATACTGAAC TTTACCAGCA
---------- ---------- ---------- ---------- ---------- ----------

      +420        +430        +440        +450        +460        +470
GCTGAATGAC CTGGAAGCCT GTGTGATGCA GGAGGTGTGG GTGGGAGGGA CTCCCCTGAT
---------- ---------- ---------- ---------- ---------- ----------

      +480        +490        +500        +510        +520        +530
GAATGAGGAC TCCATCCTGG CTGTGAGAAA ATACTTCCAA AGAATCACTC TCTACCTGAC
---------- ---------- ---------- ---------- ---------- ----------

      +540        +550        +560        +570        +580        +590
AGAGAAAAAG TACAGCCCTT GTGCCTGGGA GGTTGTCAGA GCAGAAATCA TGAGATCCTT
---------- ---------- ---------- ---------- ---------- ----------

      +600        +610        +620        +630        +640        +650
CTCTTCATCA AGAAACTTGC AAGAAAGGTT AAGGAGGAAG GAATAAGACC TGATCCAACA
---------- ---------- ---------- ---------- ---------- ----------

      +660        +670        +680        +690        +700        +710
CAGAAACGAC TCGCATTGAC GACTACACCA GCTTGCACTT TCATGATCTG CCATTTTAAA
---------- ---------- ---------- ---------- ---------- ----------

      +720        +730        +740        +750        +760        +770
GACTCTTGTT TCTGCTATAA GCATACCATG AGTTGAATCA AACGCGTCAA GTATTTTCAA
---------- ---------- ---------- ---------- ---------- ----------

      +780        +790        +800        +810        +820        +830
GTGTGTTAAG CAACATCGTG TTCAGTTGCA CAGGAACTAG TCCCTTACAG ATGACTAAGC
---------- ---------- ---------- ---------- ---------- ----------

      +840
TGATGGATC
---------
```

λSMTVI.1 and pAT153 DNAs were ligated and transformed into *E.coli* JA221. Restriction analysis of plasmid DNA prepared from thirty two ampicillin resistant colonies demonstrated nine were recombinant, one containing a 3kb *Eco*RI/ *Hin*dIII fragment which hybridised to pHIF24.40. This plasmid was designated pSMTVI.1.

Restriction Mapping of pSMTVI.1

Plasmid pSMTVI.1 was characterised by restriction mapping. As described previously cleavage sites for *Pvu*II are commonly found in human IFNα gene sequences. Therefore, *Pvu*II digests were performed in combination with Southern blot analysis (pHIF24.40 *Pst*I insert). A 300bp fragment hybridising to pHIF24.40 was revealed. As pAT153 does not contain a *Pvu*II site this demonstrates pSMTVI.1 contains two *Pvu*II recognition sequences which are probably located within the coding sequence of an IFNα gene. The *Pvu*II sites were mapped by *Eco*RI/*Pvu*II and *Hin*dIII/*Pvu*II double digests (Fig.4.8B).

A more detailed restriction map of pSMTVI.1 was generated for *Rsa*I, *Sau*3A, and *Hae*III restriction endonucleases. Specific fragments were orientated by performing double digests with either *Eco*RI, *Hin*dIII, or *Pvu*II (Fig.4.8B). Any ambiguities were ultimately resolved by seqencing studies.

Primary Nucleotide Structure of the IFNα gene SMTVI.1

Sequencing studies were undertaken to confirm that pSMTVI.1 DNA contained an authentic IFNα gene. The strategy adopted is shown in Figure 4.8B. The 300bp *Pvu*II fragment was subcloned into the *Sma*I site of M13mp10. *E.coli* JM101 was transfected with ligated DNA generating only colourless plaques on Bluo−gal containing plates. D.I.G.E analysis (Chapter 2) of four phage DNAs verified that they were all recombinants.

Upon sequence analysis three of the four recombinant phage gave 301bp of identical sequence, which when compared to characterised human IFNα genes using the Seqfit program (Staden, 1977) revealed significant homology within the coding sequence, thus verifying that λSMTVI.1 contained authentic IFNα gene sequences.

The sequence data obtained was compared visually with 19 previously characterised

human IFNα genes (Henco *et al*, 1985). One gene, IFNα_6, contained only a single base substitution relative to this sequence within the 301bp region. Sequence data obtained from the fourth recombinant was in the opposite orientation to that described above. This verified the 301bp *Pvu*II fragment had a single base change relative to the comparable region of the IFNα_6 gene.

To assess whether the two IFNα genes SMTVI.1, and IFNα_6 were alleles, sequence analyses were extended to the 5' and 3' regions of the gene. An M13 recombinant phage containing the 5' region of pSMTVI.1 was isolated by cloning the pSMTVI.1 *Pvu*II/*Eco*RI fragment (Fig.4.8B) into *Sma*I/*Eco*RI digested M13mp10. 240bp of further sequence was obtained from these clones, extending to −127 (relative to the transcription intitiation site). Similarly 3' fragments were isolated and sequenced. Initially the pSMTVI.1 *Pvu*II/*Hin*dIII fragment (Fig.4.8B) was subcloned into *Sma*I/*Hin*dIII digested M13mp10. However, the nucleotide sequence obtained from the *Hin*dIII site did not match any available human IFNα sequences associated with the EMBL database, indicating that this region is probably downstream of previously characterised IFNα gene sequences. Insert DNA from this recombinant was isolated as an *Eco*RI/*Hin*dIII fragment by P.A.G.E, digested with *Sau*3A, and subcloned into *Bam*HI digested M13 mp10 DNA. Eight prospective recombinants were sequenced generating a total of 429bp of additional nucleotide sequence, including the remaining 3' IFNα gene coding sequence, and 204 bp of 3' flanking sequence.

The complete sequence derived for the gene SMTVI.1 is shown in Figure 4.10, where it is compared to the HuIFNα_6 gene. The promoter element TATTTAA is highlighted together with the putative CAP site and translation initiation and termination codons. A single conservative base substitution within the coding sequence, changing codon 65 (mature IFNα polypeptide) from AAC (IFNα_6) to AAT (SMTVI.1), is apparent. Both genes, therefore, encode identical full length IFNα polypeptides. Comparison of the coding sequence of SMTVI.1 with other available IFNα gene sequences demonstrated the next most closely related gene to be IFNα_5 showing 49 base substitutions which overall result in 26 amino acid changes. It is, therefore, highly likely that SMTVI.1 and IFNα_6 represent alleles of the the same gene.

4.4 Comparison of SMTI.2, SMTIII.1_A and SMTVI.1 Gene Promoters with the HuIFNα_1 Gene Promoter

To identify potential IFNα regulatory sequences a comparison of the primary sequence of the promoter region of these genes was made with that of the HuIFNα_1 gene. When the nucleotide sequences were aligned they display a high degree of homology (Fig.4.11). In particular the regions around the TATA box and putative transcription initiation sites are highly conserved. All currently characterised human IFNα genes including those identified in this study possess the TATA box variant TATTTAA, with the exception of IFNα_5 (TTTTTAA; Henco *et al*, 1985). Some divergence of promoter sequences, evident between −61 and −72 are common between HuIFNα_1, SMTI.2, and SMTIII.1_A. Potential deletions or insertions of 8bp and 11bp are also seen between HuIFNα_1 and λSMTIII.2 or λSMTVI.1 genes, respectively, in this region (Fig.4.11).

Figure 4.11 also indicates repeat structures identified in HuIFNα_1 as being important for NDV induced expression in heterologous mouse L cells (Ryals *et al*, 1985). These sequences are partially deleted in the SMTIII.1 and SMTVI.1 gene promoters. It is interesting in this respect that transcripts derived from the SMTVI.1 allele IFNα_6 have not been observed previously (Hiscott *et al*, 1984). In view of these observations functional analysis of these IFNα gene promoters would clearly be worthwhile.

4.5 Conclusions

Four IFNα genes have been identified within three isolates from a newly constructed human genomic library (Chapter 3). The primary sequence of three of these genes was determined completely. Alleles of these genes have been described previously. SMTIII.1_A is clearly an allele of the pseudogene ψLeIF−L. However, SMTIII.1_A contains six base substitutions within the coding region, including codon 20 which changes the termination codon UGA (ψLeIF−L) to UGU. Therefore, SMTIII.1 is not a pseudogene.

Interestingly, ψLeIF−L has been isolated on three previous occassions (Brack *et al*, 1981, Pestka *et al*, 1983 Torcynski *et al*, 1984, Henco *et al*, 1985). Two alleles, both pseudogenes, have been identified by these studies. However, on each occasion the same

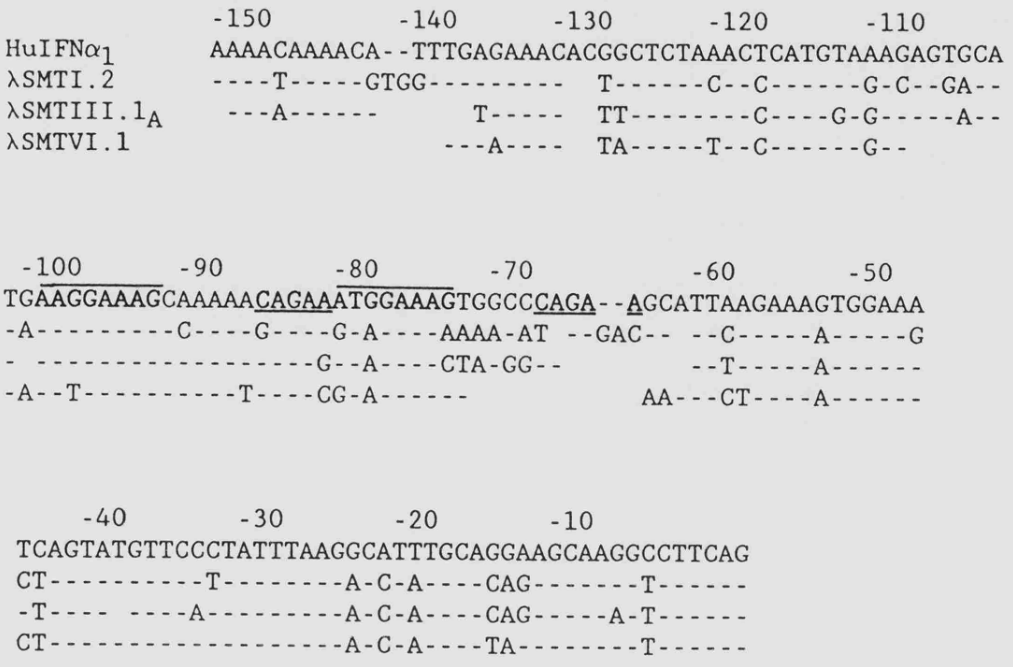


FIGURE 4.11 Comparison of the 5' Flanking Sequences of SMTI.2, SMTIII.1_A and SMTVI.1 with the Human IFNα_1 Gene Promoter

The nucleotide sequence of the human IFNα_1 gene promoter region is shown. Homologous sequences from the human IFNα genes SMTI.2, SMTIII.1_A and SMTVI.1 are aligned below. Dashes (-) indicate regions of homology. Specific base substitutions are shown using conventional symbols. Spaces have been included to achieve the best alignment. The repeat sequences of the human IFNα_1 gene which are believed to be important for viral induction (Ryals *et al*, 1985) have been underlined and overlined.

genomic library was employed as the primary source of the genes characterised. The coding sequence of the cDNA, LeIF−C (Goeddel *et al*, 1981), derived from Sendai virus induced KG−1 cells is closely related to that of ψLeIF−L suggesting they are allelic (Ullrich *et al*, 1982; Henco *et al*, 1985). The LeIF−C gene sequence suggests it encodes a functional IFNα protein.

Inspection of the coding sequences for SMTIII.1_A and LeIF−C reveals only two conservative base substitutions at codon 22 from CUA (SMTIII.1_A) to CUG (LeIFC) and codon 32 from CUG (SMTIII.1_A) to CUC (LeIFC), encoding the amino acid leucine. Therefore, the genomic clone SMTIII.1_A and the previously described cDNA LeIF−C encode identical proteins. Thus, together with the relatedness of SMTIII.1_A to ψLeIF−L, this strongly supports the assignment of ψLeIF−L and LeIF−C as alleles. Therefore, SMTIII.1_A is the the first genomic clone described which corresponds precisely to the IFNα mRNA from which the LeIF−C cDNA was derived.

Functional activity for a recombinant polypeptide derived from ψLeIF−L has been reported. The ψLeIF−L gene containing a modified signal sequence has been expressed as a pre−interferon polypeptide from the M13mp11 *lac* promoter, producing 10^6 − 10^7 units/L of IFN activity (Fuke *et al*, 1984). Thus, evidence from both cDNA cloning and bacterial expression studies indicate that this gene can be transcribed and encodes a protein with anti−viral properties.

The alleles ψLeIF−L and SMTIII.1_A have been isolated from two independently derived human genomic libraries. It would be of interest to investigate the relative frequencies of these alleles in the population. This could be achieved by Southern blot analysis of DNA derived from a representative sample of the population using oligonucleotide hybridisation probes specific for the region surrounding codon 20 of the LeIF−L gene. Inspection of this region suggests oligonucleotides specific for each allele, distinguishing them from other currently characterised IFNα genes, could be designed.

Interferon gene SMTI.2 is an allele of λ2h (Lawn *et al*, 1981a), which is also an allele of IFNα_{14} (Brack *et al*, 1981) and the cDNA LeIF−H (Goeddel *et al*, 1981) as defined by Henco *et al* (1985). If this assignment is correct, then clearly this allele is expressed. Expression studies with IFNα_{14} demonstrate it is a minor component of virus

induced leukocyte mRNA but a major constituent of leukaemic myeloblast mRNA (Hiscott *et al*, 1984). Thus, this suggests that IFNαs may be produced in a tissue specific fashion. Expression studies upon reintroduction of the SMTI.2 gene into various mammalian cell lines therefore, may be of considerable interest.

Interferon gene SMTVI.1 is an allele of IFNα_6 (Henco *et al*, 1985). Another allele, LeIF−K, identical to IFNα_6 has been described (Henco *et al*, 1985). This gene has not been isolated previously as a cDNA nor has its expression been detected in leukocytes (Hiscott *et al*, 1984). As described previously (Section 4.4), this promoter lacks 11bp of promoter sequence found in the HuIFNα_1 gene, which is essential to viral induction of gene expression in mouse fibroblasts (Ryals *et al*, 1985). Again it would be useful to investigate expression of this gene upon re−introduction into mammalian cells to determine if it contains a functional promoter.

The IFNα gene family consists of 23 currently characterised loci, corresponding to 15 potentially functional genes, 6 pseudogenes, and 2 uncharacterised (Henco *et al*, 1985). The data presented above suggest that at least four alleles for the LeIF−L loci exist. Furthermore at least three alleles for λ2h and two for IFNα_6 have now been described. The fact that previously identified alleles of ψLeIF−L are pseudogenes and SMTIII.1 (and cDNA LeIF−C) is not, would be consistent with some degeneracy of the IFNα gene system.

CHAPTER 5

THE CONSTRUCTION AND FUNCTIONAL ANALYSIS OF A MOUSE IFNα1 PROMOTER/CAT GENE CHIMAERA

The principal objective of these studies was to investigate transcriptional control of interferon (IFN) genes after the introduction of cloned genes into mammalian cells. While characterisation of the human IFNα chromosomal genes was in progress (Chapter 3 and 4), the first mouse IFNα (MuIFNα) genomic clone was isolated (Shaw *et al*, 1983). Plasmid pGS1, which contains this MuIFNα1 gene extending 950bp 5' and 100bp 3' to the coding sequence, was obtained from Prof. C. Weissmann. This allowed a functional analysis of an IFN promoter to begin.

Functional analyses of human IFNα1 (HuIFNα1) and human IFNβ (HuIFNβ) genes in heterologous cells has been documented but the transcriptional control of a cloned MuIFNα gene has not been studied previously in this manner (Chapter 1). To monitor MuIFNα1 gene promoter activity *in vivo* it has been used to direct the expression of the bacterial gene encoding chloramphenicol acetyltransferase (CAT). PolyrI.rC mediated CAT production has been assessed in transiently and stably transfected mouse L−929 cells to characterise and validate this system as a means of studying IFNα gene expression.

5.1 The Construction of an Expression Vector to Analyse MuIFNα1 Promoter Activity in Homologous Cells

A cloning strategy was devised to a) construct a suitable expression vector for the routine monitoring of MuIFNα1 gene promoter activity in mouse L−929 cells by linking it to a reporter gene (*cat*), and b) afford intermediates which could serve the dual purpose of being both the source of a promoter module and the substrate for planned *in vitro* site directed mutagenesis studies (Chapter 6).

A convenient restriction fragment encompassing the prospective polyrI.rC regulatory elements of the MuIFNα1 gene promoter was identified. To facilitate manipulation of the promoter, flanking heterologous restriction sites were introduced which allowed shuttling

of the promoter module between M13 (for site directed mutagenesis and sequencing) and a promoterless CAT gene plasmid.

Identification of a MuIFNα_1 Promoter Containing DNA Fragment

To help identify the MuIFNα_1 gene promoter, the primary nucleotide sequence upstream of the MuIFNα_1 gene was aligned and compared with the well characterised HuIFNα_1 gene promoter. Inspection of the restriction map for pGS1 (Shaw *et al*,1983) revealed a *Hin*dIII fragment encompassing MuIFNα_1 promoter sequences. Published sequencing data (Shaw *et al*, 1983) showed that this extended 52bp into the untranslated gene leader sequence and 188bp 5' to the transcription initiation site.

Figure 5.1 shows a comparison of the nucleotide sequence of this 241bp *Hin*dIII fragment with the 204bp of the HuIFNα_1 gene which contains a virus inducible promoter (Ragg and Weissmann, 1983). Identical domains of 17bp (−4 to +12) and 11bp (−35 to −25) are apparent, encompassing both the transcription initiation site and a variant of the common RNA polymerase II promoter TATA box element (Corden *et al*, 1980) respectively. Overall 71bp substitutions occur within the 204bp region compared, demonstrating 65% homology between these sequences.

Thus, the extent of homology strongly suggested the *Hin*dIII fragment included sequences required for virus induction of the MuIFNα_1 gene promoter. An early objective of this study, therefore, was to determine by direct experimental demonstration that this sequence also contained the information required for polyrI.rC regulation of the MuIFNα_1 gene promoter.

Construction of a MuIFNα_1 Promoter *cat* Hybrid Gene

The strategy adopted to construct a MuIFNα_1 promoter/*cat* gene hybrid for IFN gene regulation studies, illustrated in Figure 5.2, utilised the plasmid vector p22 (gift from Dr. J. Lang). This plasmid possesses unique *Sst*I and *Bam*HI sites allowing insertion of promoter fragments 5' to a modified *cat* gene, flanked by a 3' HSV−2 polyA signal sequence, AATAAA (Proudfoot and Brownlee, 1976), suitable for expression studies in mammalian cells (Fig.5.3B).

FIGURE 5.1 Comparison of the 5' Flanking Sequences of the MuIFNα_1 Gene with the HuIFNα_1 Gene Promoter

The primary nucleotide sequence of the MuIFNα_1 gene *Hin*dIII fragment (-188 to +52) is illustrated. Below, the corresponding nucleotides of available HuIFNα_1 gene promoter sequence are compared. A single space has been incorporated within the MuIFNα_1 gene sequence to achieve the best alignment. Conserved nucleotides between the two promoters are indicated by a dash (-) in the HuIFNα_1 gene sequence. Nucleotide substitutions are shown using the conventional symbols. The TATA box and transcription initiation sites (CAP) are underlined. Repeat sequences implicated in viral induction of the HuIFNα_1 gene promoter are shown underlined and overlined (R1, R1', R2 and R2'; Ryals *et al*, 1985), the corresponding MuIFNα_1 gene sequences are also underlined and overlined. † C to T transversion identified from sequencing analysis (see sequence of Shaw *et al*, 1983).

```
          HindIII     †
                    -180       -170       -160       -150       -140
MuIFNα1:   AAGCTTT TGATGAGGACCAGTGAAAGAGGAAGCAATAATGAAAACCACAATGGTTTA
HuIFNα1:                                                AA--C-AA---T---AGA-
```

```
 -130       -120       -110       -100       -90        -80        -70
GAAAACACCCAGACGCAAGCAGAGAATGAGTTAAAGAAAGTGAAAAGACAAGTGGAAAGTGATGGAA
AC-CGGCT-T-A--T--T-T-A---G--CA-G--G-----CA----C-G--A---------GCCC-G
                                  AAGGAAAGCAAAAACAGAAATGGAAAGTGGCCCAG
                                     R1           R2    R1'          R2'
```

```
                                    TATA                            CAP
    -60        -50        -40        -30        -20        -10
GGGCATTCAGAAAGTAAAAACCAGTGTTTGCCCTATTTAAGACACATTCACCCAGGATGGTCTTCAG
AA-----A--------G---T----A-G-T----------G--TT-G--GGA--C-A--C------
AA
```

```
                                                    HindIII
       +10        +20        +30        +40        +50
AGAACCTAGAGGGGAAGGATCAGGACCAAACAGTCCAGAAGACCA GAAGCTT
-----------CCC----T----AGT--CC--TCT------C---C----TA-
```

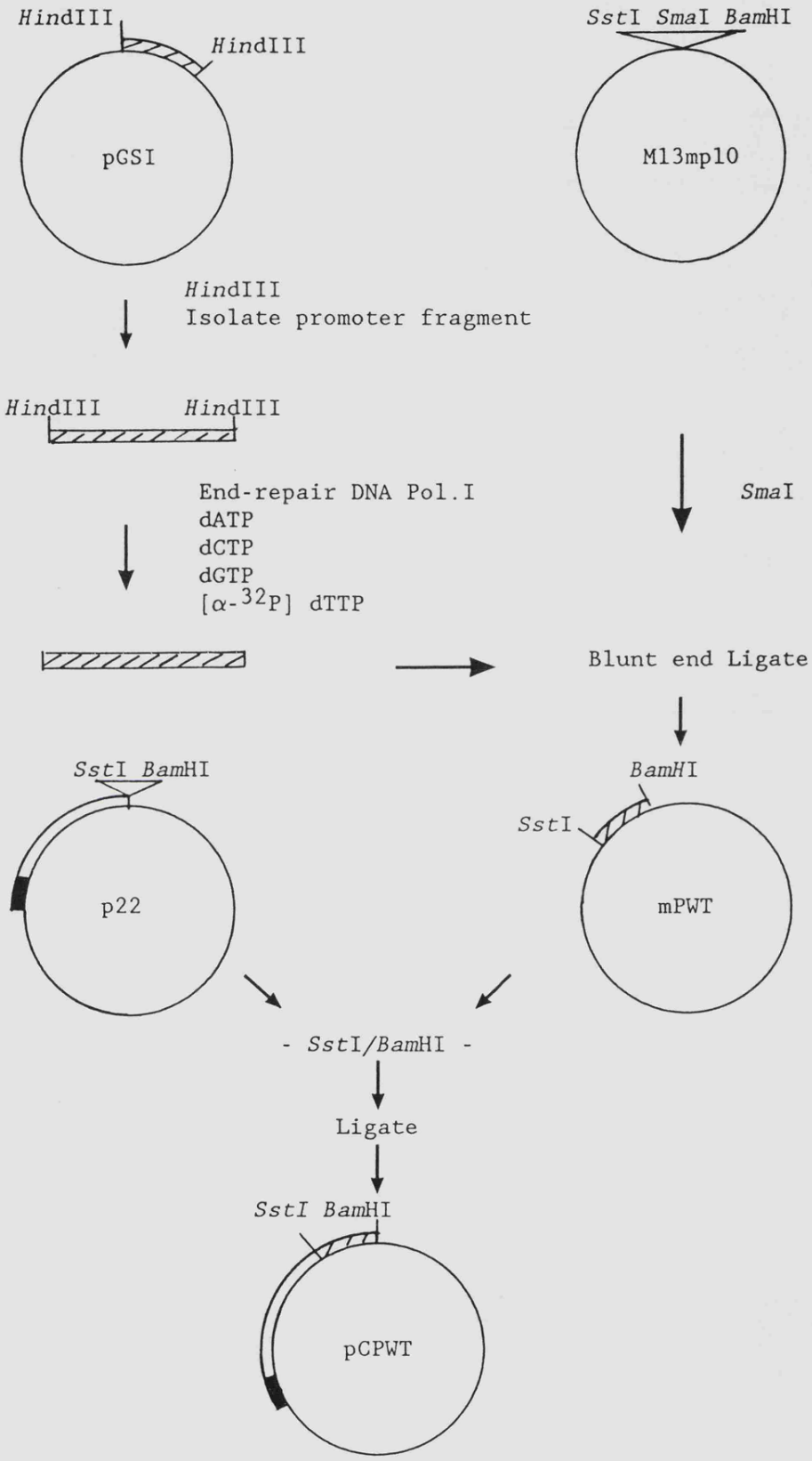


FIGURE 5.2 STRATEGY FOR THE CONSTRUCTION OF A MuIFNα1 PROMOTER *cat* GENE.
Striped box = MuIFNα_1 promoter, open box = *cat* gene and filled box = HSV-2 PolyA. Only essential restriction sites are shown

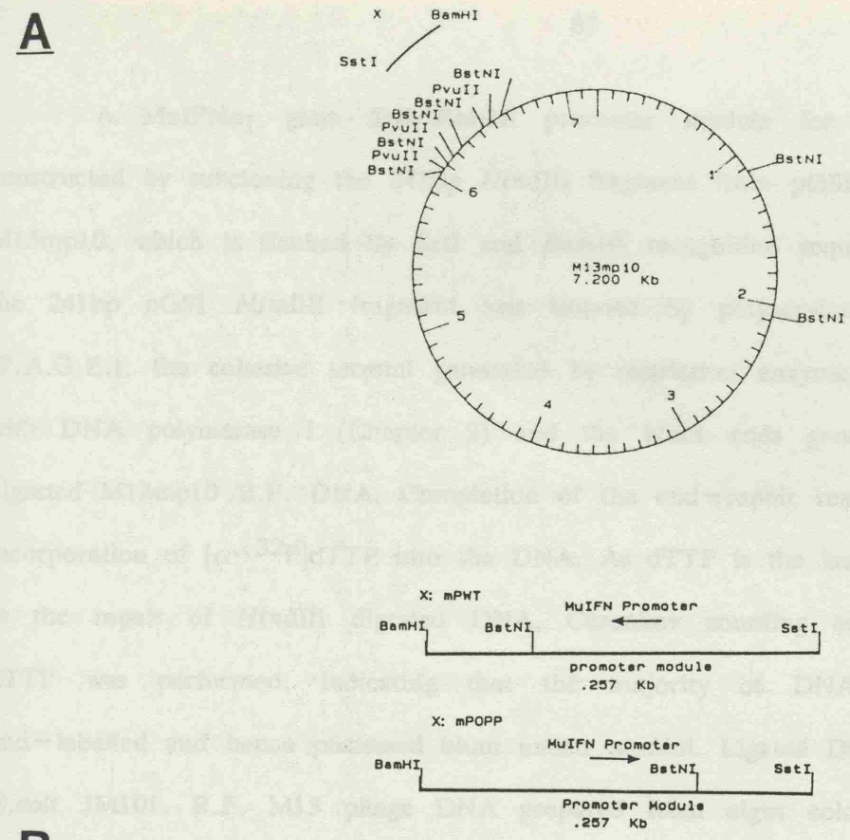
FIGURE 5.3 Restriction Maps of M13mp10, mPWT, mPOPP, p22, pCPWT, pCOPP, p41 and p43

Partial restriction maps of M13mp10 and p22 for *Bam*HI, *Bst*NI, *Eco*RI, *Pvu*II and *Sst*I restriction endonucleases are illustrated. Derivatives of these contain insertions at position X:. The insertion X, and the corresponding constructs created are shown. These are not drawn to the same scale.

A) Shows the restriction maps of M13mp10, mPWT and mPOPP, indicating the *Bst*NI and *Eco*RI sites used to orientate the *Sst*I/*Bam*HI MuIFNα_1 gene promoter module.

B) Shows the restriction maps of p22, pCPWT and pCPOPP. The *cat* gene coding and polyA regions of p22 are shown along with the *Bst*NI and *Eco*RI sites used to orientate the *Sst*I/*Bam*HI MuIFNα_1 gene promoter module.

C) Shows the HSV-1 *tk* and Mo-MuSV LTR promoter fragments present in the p22 derivatives p41 and p43.


A
X
BamHI
SstI
BstNI
PvuII
BstNI
BstNI
PvuII
BstNI
PvuII
BstNI
BstNI
M13mp10
7.200 Kb
1
2
3
4
5
6
7
X: mPWT
MuIFN Promoter
BamHI
BstNI
SstI
promoter module
.257 Kb
X: mPOPP
MuIFN Promoter
BamHI
BstNI
SstI
Promoter Module
.257 Kb
B
X
BamHI
SstI
EcoRI
PvuII
BstNI
PvuII
cat
EcoRI
BstNI
BstNI
BstNI
HSV-2 PolyA
HindIII
BstNI
PvuII
BstNI
BstNI
BstNI
p22
3.600 Kb
1
2
3
X: pCPWT
MuIFN Promoter
BamHI
BstNI
SstI
promoter module
.257 Kb
X: pCPOPP
MuIFN Promoter
BamHI
BstNI
SstI
Promoter Module
.257 Kb
C
X: p41
HSV-1 Tk Promoter
BamHI
BglII
.250 Kb
X: p43
Mo-MSV LTR
BamHI
BamHI
.730 Kb

A MuIFNα_1 gene *Sst*I/ *Bam*HI promoter module for cloning into p22 was constructed by subcloning the 241bp *Hin*dIII fragment from pGS1 into the *Sma*I site of M13mp10, which is flanked by *Sst*I and *Bam*HI recognition sequences. For this purpose the 241bp pGS1 *Hin*dIII fragment was isolated by polyacrylamide gel electrophoresis (P.A.G.E.), the cohesive termini generated by restriction enzyme digestion end−repaired with DNA polymerase I (Chapter 2) and the blunt ends generated ligated to *Sma*I digested M13mp10 R.F. DNA. Completion of the end−repair reaction was monitored by incorporation of [$\alpha-^{32}$P]dTTP into the DNA. As dTTP is the last base to be introduced in the repair of *Hin*dIII digested DNA, Cerenkov counting of incorporated [$\alpha-^{32}$P] dTTP was performed, indicating that the majority of DNA molecules had been end−labelled and hence possessed blunt ended termini. Ligated DNA was transfected into *E.coli* JM101. R.F. M13 phage DNA prepared from eight colourless plaques, digested with *Sst*I/ *Bam*HI, and analysed by P.A.G.E., released a 257bp fragment in each case. This suggested the 241bp *Hin*dIII fragment had been inserted in M13. The orientation of the promoter fragment was determined by *Eco*RI/ *Bst*NI double digests and P.A.G.E. This demonstrated that four recombinant phage were likely to contain the *Hin*dIII promoter fragment in each of the two possible orientations (corresponding to mPWT and mPOPP; Fig.5.3A).

The structure and orientation of the promoter inserts were confirmed by sequencing a recombinant for both mPWT and mPOPP. This unequivocally demonstrated that the desired recombinant phage had been isolated, although a single C to T transversion at position −179 (Fig.5.1), relative to the published sequence (Shaw *et al*, 1983), was identified. This has subsequently been confirmed in the primary nucleotide sequence of the MuIFNα_1 gene described by Kelley and Pitha (1985). Phage mPWT satisfies the criteria discussed previously, being suitable for site directed mutagenesis (Zoller and Smith, 1983), sequencing and routine orientation dependent cloning of the promoter module into p22.

To complete the construction of vectors carrying MuIFNα_1 promoters coupled to the *cat* gene the *Sst*I/*Bam*HI promoter modules were excised from mPWT and mPOPP, ligated to *Sst*I/*Bam*HI digested p22, and transformed into *E.coli* JA221. Plasmid DNA

prepared from eight ampicillin resistant colonies for each ligation were digested with *Sst*I/ *Bam*HI and analysed by P.A.G.E. This revealed that approximately 75% and 25% of the colonies examined from each transformation, contained recombinant plasmids with the promoter module inserted. Thus, plasmids designated pCPWT and pCPOPP (Fig.5.2) had been created. Again the restriction maps for these plasmids (Fig.5.3B) were confirmed by *Eco*RI/ *Bst*NI double digests and P.A.G.E.

Through this work two vectors have been produced which are suitable for the investigation of MuIFNα_1 gene promoter transcriptional control of a heterologous gene, the product of expression, CAT, being readily assayed in both transiently and stably transfected mammalian cells (Gorman *et al,* 1982; 1983). The intermediate construct, mPWT, is also an appropriate species for site directed mutagenesis (Chapter 6).

5.2 Expression of the Heterologous CAT Gene from the MuIFNα_1 Promoter in Mouse L−929 Cells

To monitor *in vivo* regulation of CAT production from the MuIFNα_1 gene promoter it was necessary to introduce pCPWT and pCPOPP into mammalian cells. Two alternative procedures, involving transient and stable transfection were adopted. L−929 cells were chosen for these studies as they have been employed extensively in DNA transfection studies (Fujita *et al*, 1985) and their endogenous IFN genes are readily induced by polyrI.rC (Trapman *et al*, 1979). Therefore, prior to analysing expression of the transfected MuIFNα_1 gene promoter/*cat* genes it was necessary to establish conditions for DNA transfection and polyrI.rC induction in L−929 cells.

Transient Expression of CAT in Mouse L−929 Cells

Transient expression of CAT was adopted to optimise the conditions for DNA uptake by L−929 cells for both transient and stable transfection protocols. To investigate the transient expression of CAT in L−929 cells, the p22 derivatives p41 and p43 (Fig.5.3) were used (gifts of Dr. J. Lang). These plasmids should direct constitutive CAT gene expression from the HSV−I *tk* (McKnight and Gavis, 1980) and Mo−MSV LTR (Kreigler and Botchan, 1983) promoters respectively.

Firstly, p41 was introduced into L−929 cells by incubating them with a DNA/$CaPO_4$ co−precipitate (Wigler *et al*, 1979) for 4hrs or 24hrs and CAT activity assayed 48hrs post−transfection. Cells exposed to p41 for 24hrs produced higher CAT activity than those exposed for 4hrs (Fig.5.4 A;B), suggesting increased DNA incorporation. It has been reported that the addition of a 2min glycerol treatment (20%) immediately after exposure of cells to the DNA/$CaPO_4$ co−precipitate increases uptake. This was investigated in a parallel study, which suggested CAT activity and hence DNA uptake is enhanced after 4hrs only (Fig.5.5). Transient expression can also increase up to 72hrs post transfection (Lopata *et al*, 1984), but no significant improvement was observed when this was attempted (data not shown).

Prior to this study CAT expression from neither p41 nor p43 had been investigated (Dr J. Lang, personal communication). Detection of CAT production from p41 in L−929 cells required long exposures (legend to Figs. 5.4 and 5.5), suggesting either poor DNA uptake or poor promoter activity. Moreover, when CAT activity was seen it was not substantially elevated over that detected for the promoterless plasmid p22. Subsequently, many transient expression assays performed with p43 in L−929 cells, produced reproducibly high CAT activity which unlike p41 was observed in the absence of detectable promoterless CAT production (Fig.5.6). Thus, the low CAT production observed with p41 is probably attributable to weak promoter activity in these cells.

From this pilot study it was concluded that suitable conditions for transient assay studies in L−929 cells had been established. The most suitable conditions for DNA uptake required the incubation of L−929 cells with the DNA/$CaPO_4$ co−precipitate for 24hrs. These conditions have, therefore, been used in all subsequent transient and stable transfection studies.

Expression of CAT in Stably Transfected L929 Cells

Stably transfected cells provided an alternative to transient expression for analysing gene activity. Preliminary experiments were performed to demonstrate both stable transfection and CAT expression in L−929 cells. Constitutive expression of CAT in mammalian cells stably transfected with pSV2*cat* had been reported (Gorman *et al*,

FIGURE 5.4 Transient CAT Assays with p41 in L-929 Cells

Transient CAT assays were performed as described in Chapter 2. Cells were incubated with the DNA/$CaPO_4$ co-precipitate for 4hrs (A) or 24hrs (B), after which the medium was replaced with F.G.M. Cells were harvested 48hrs post-transfection and CAT assays performed with 50μl of extract using 1μCi of ^{14}C-chloramphenicol/assay for 2hrs (Chapter 2). T.L.C was performed in 95:5 Chloroform: Methanol. The solvent front was allowed to migrate 10cm from the origin. Autoradiography was performed for 6 days using Fuji RX film.

a) 20μg salmon sperm DNA

b) 10μg p22 +10μg salmon sperm DNA

c) 10μg p41 +10μg " " "

d) Pure CAT (Pharmacia)

e) ^{14}C-chloramphenicol only

O: Origin

C: ^{14}C-chloramphenicol

1'A: chloramphenicol-1-acetate

3'A: chloramphenicol-3-acetate

1'3'DiA: chloramphenicol-1,3-diacetate

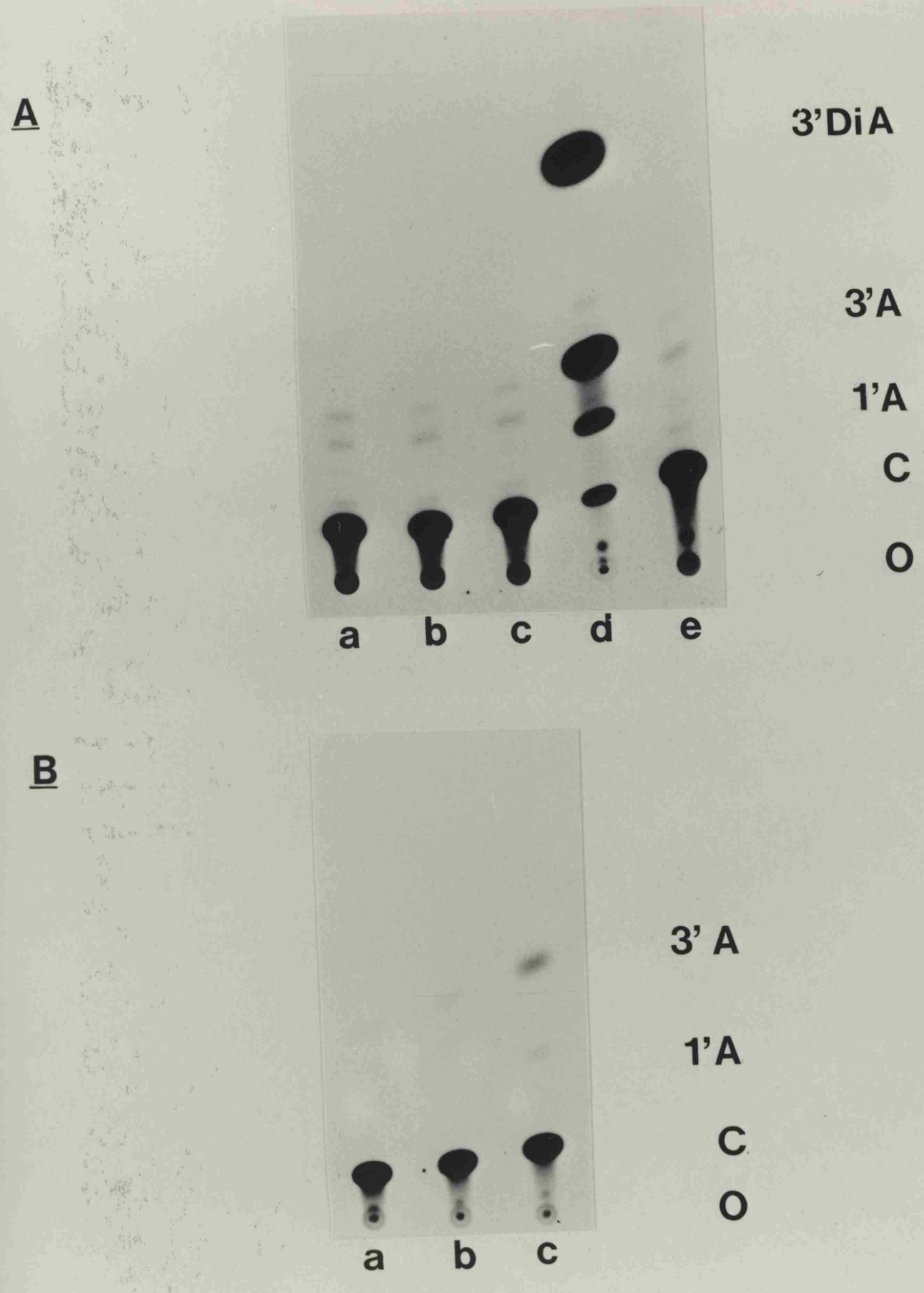

A
3'DiA
3'A
1'A
C
O
a
b
c
d
e
B
3' A
1'A
C
O
a
b
c

FIGURE 5.5 Transient CAT assays with p41 in L-929 Cells

See legend to Fig.5.4. After the cells had been incubated with the DNA/$CaPO_4$ co-precipitate for 4hrs the cells were incubated for 2mins in 20% glycerol, then washed and the medium replaced with F.G.M.

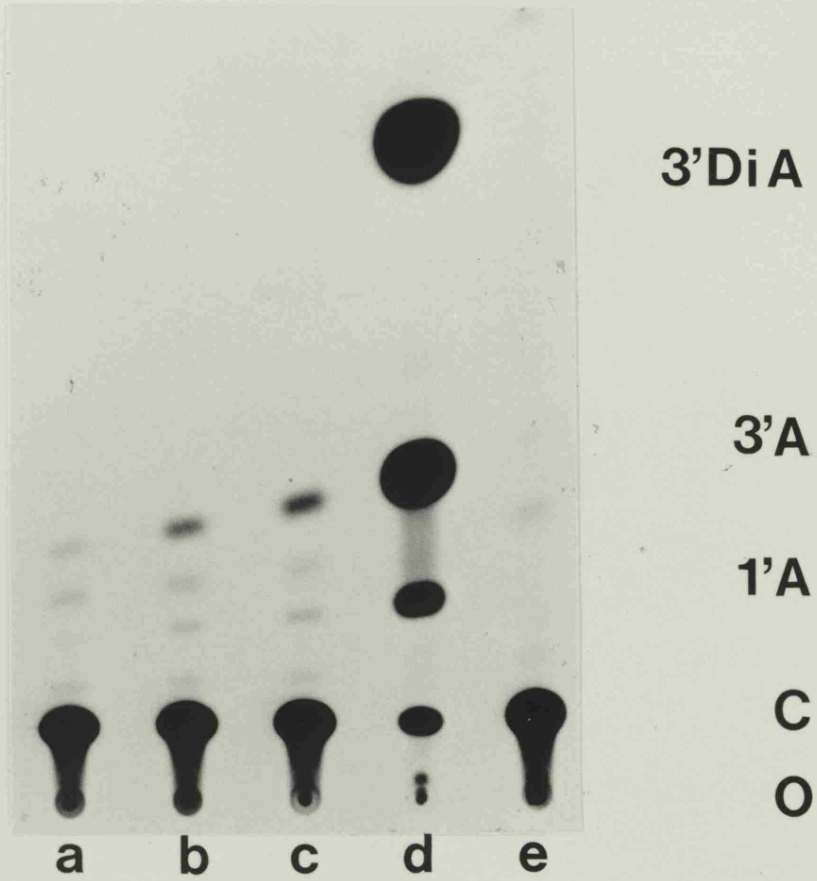

3'DiA
3'A
1'A
C
O
a
b
c
d
e

FIGURE 5.6 Transient CAT Assays with p43 in L-929 Cells

See legend to Fig.5.4 except 0.2μCi of ^{14}C-chloramphenicol was used per assay. Four independent transfections with p43 were performed in parallel. L-929 cells were incubated with the DNA/$CaPO_4$ co-precipitate for 24hrs and the medium subsequently replaced with F.G.M. Cells were harvested and CAT assays performed with 10μl of cell extract for 1hr. Autoradiography was performed for 12hrs using Fuji RX film (A). The structure of the *cat* gene and flanking sequences of p22 and p43 are also shown (B).

a) Pure CAT

b) 10μg p22 + 10μg salmon sperm DNA

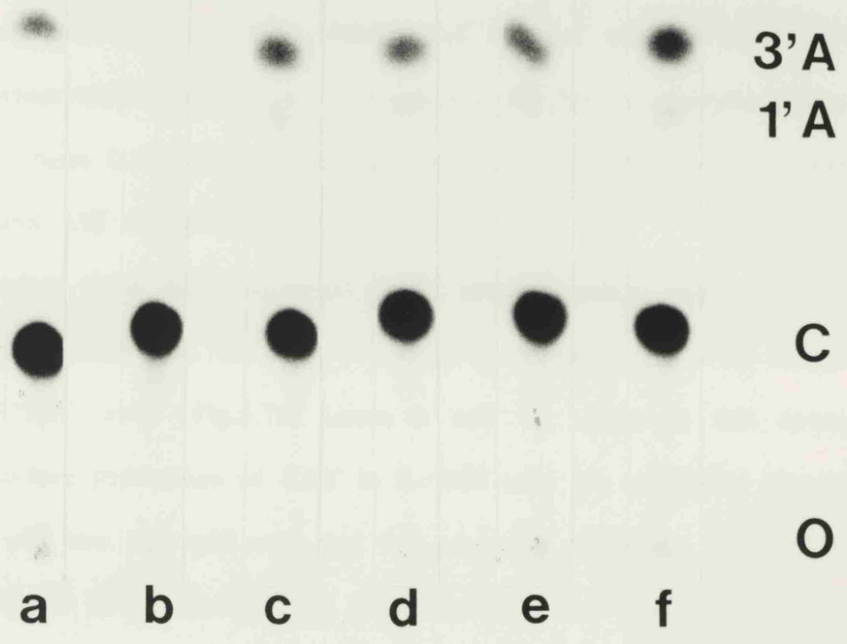
c) " p43 " " plate 1

d) " " " " plate 2

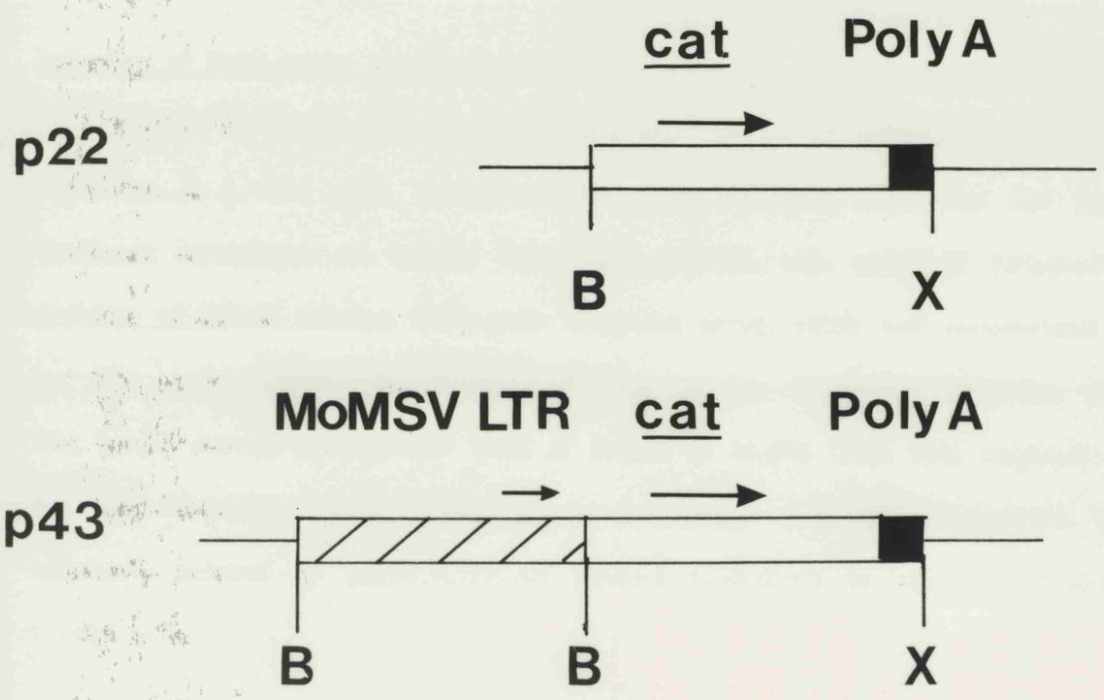
e) " " " " plate 3

f) " " " " plate 4

A



B



1983). Therfore, this plasmid (gift of Dr. B. Howard) was introduced in parallel with the promoterless plasmid p22 into L−929 cells by co−transfection with pTCF (Grosveld *et al*, 1982), to allow for selection of stably transfected colonies in G418 (400μg/ml). At 8 to 14 days G418 resistant colonies were first visible, and counted after 18 days (Table 5.1). The colonies of independent transfections were pooled for further analysis (Chapter 1), generating cell populations L−929p22 (1 and 2) and L929pSV2*cat* (1 and 2) which should contain the plasmids p22 or pSV2*cat* respectively.

To determine whether these cells expressed CAT, enzyme assays were performed. These demonstrated readily detectable activity in L929pSV2*cat* cells (Fig.5.7A Lanes c and d) but not in L929p22 cells (Fig.5.7A Lanes a and b). Although this strongly suggested promoter dependent production of CAT in L−929 cells the possibility remained that co−transfection of p22 into L929p22 cells had not occurred. However, Southern blot analysis verified that L929p22 cells contained p22 sequences (data not shown).

Thus, this pilot study showed that *cat* genes could be stably co−transfected with pTCF to generate cell populations producing readily detectable levels of CAT in the expected constitutive fashion in L−929 cells. CAT production was completely dependent upon the presence of the SV40 early promoter upstream of the *cat* gene.

Induction of Endogenous IFN Gene Expression by PolyrI.rC in Mouse L929 Cells

Finally, to characterise the production and kinetics of polyrI.rC induced IFN gene expression in L−929 cells, endogenous host cell IFN gene expression was investigated. Confluent monolayers of L−929 cells were induced with polyrI.rC (20μg/ml) in the presence of DEAE dextran (800μg/ml; Trapman *et al*, 1979) and supernatants removed for IFN assays. These were kindly performed by Dr. A. Morris (Warwick University). The results showed optimal IFN titres of 30,000 to 40,000 U/ml were obtained 15 to 24 hrs post−induction. Thus, these results established that IFN production could be efficiently induced by polyrI.rC/DEAE dextran (polyrI.rC) in confluent monolayers of L−929 cells.

TABLE 5.1 Number of Colonies Obtained by Co-transfection of p22 or pSV2*cat* with pTCF in L-929 Cells

Plasmid (10μg)	pTCF (1μg)	Cell Population	Colonies/Dish
1) p22	+	$L929p22_1$	150
2) p22	+	$L929p22_2$	110
1) pSV2CAT	+	$L929pSV2CAT_1$	100
2) pSV2CAT	+	$L929pSV2CAT_2$	150
1) SS DNA	-	-	0
2) SS DNA	-	-	0

A total of 20μg of DNA, was added to 10^6 L-929 cells as a DNA/$CaPO_4$ co-precipitate (Chapter 2) and incubated for 24 hrs before replacing with fresh medium. G418 (400μg/ml) selection was initiated 24hrs later and the medium replaced every 3 days thereafter. G418 resistant colonies were counted 18 days post transfection.

1) Transfection 1

2) " 2

FIGURE 5.7 CAT Expression in L-929 Cells Stably Transfected with pSV2*cat*

See legend to Fig.5.6, except ^{14}C-chloramphenicol was from N.E.N. 1hr CAT assays were performed on 10μl aliquots of cell extracts prepared from confluent monolayers of L-929 cell populations stably transfected with p22 or pSV2*cat* (see legend to Table 5.1). Autoradiography was performed for 12hrs with Fuji RX film (A). The structure of the *cat* gene and 5' flanking sequences of p22 and pSV2*cat*, indicating the SV40 early promoter, splice and polyA sequences, are also shown (B).

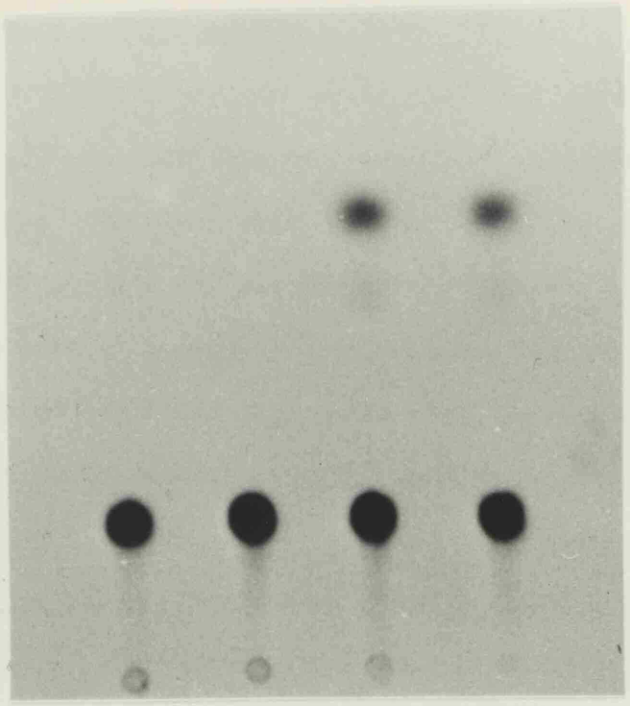
a) L929p22$_1$

b) L929p22$_2$

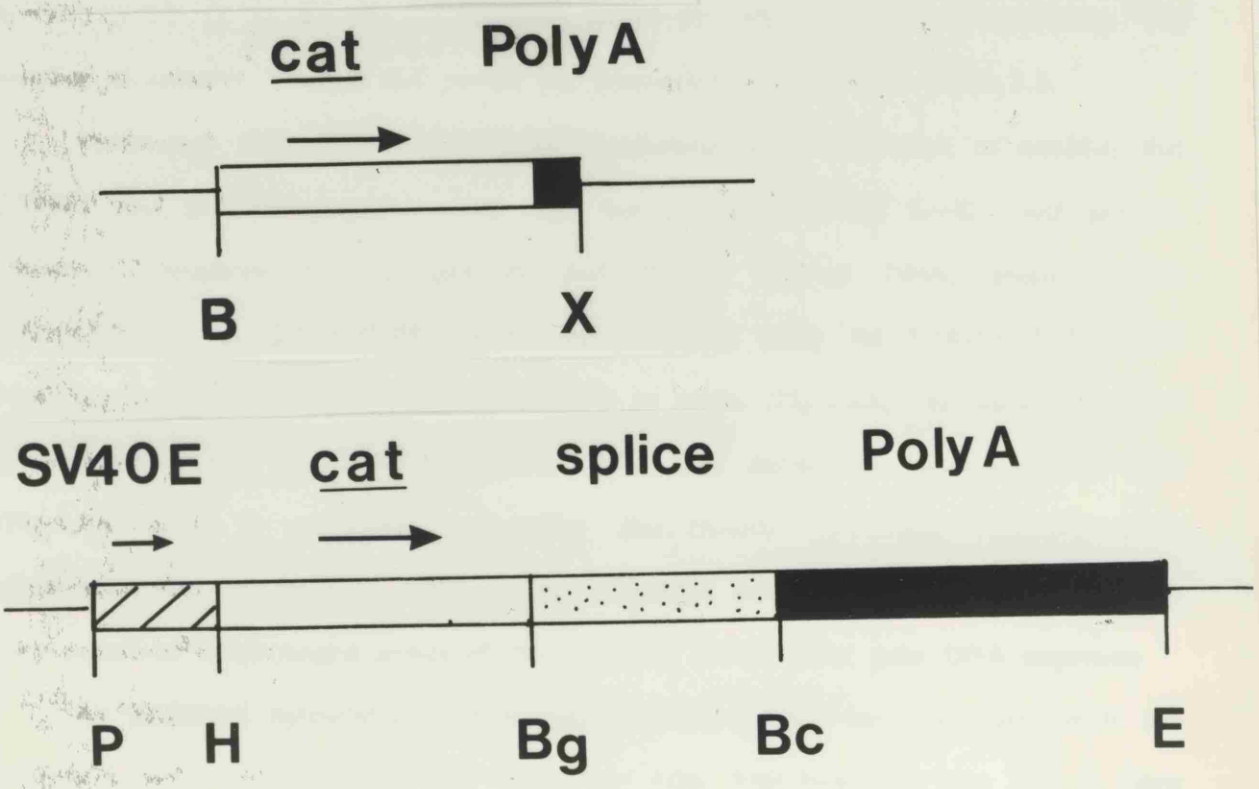
c) L929pSV2*cat*$_1$

d) L929pSV2*cat*$_2$

A



B



Transient Expression of CAT from the MuIFNα_1 Promoter

The MuIFNα_1 gene promoter activity was then examined, initially by transient transfection of pCPWT. L929 cells were transfected as described previously. Transfected cells were induced 24hrs post−transfection with polyrI.rC and CAT assays performed after a total of 48hrs. However, this assay proved unsatisfactory. Very low levels of CAT production were observed in polyrI.rC induced cells. The major conclusion, derived from a series of transient transfection experiments, was that this assay was incompatible with efficient polyrI.rC mediated activation of IFN gene expression.

Expression of CAT from the MuIFNα_1 Promoter in Stably Transfected L929 Cells

In view of the problems experienced studying polyrI.rC regulated expression from the MuIFNα_1 gene promoter in transiently transfected cells, stable transfection studies were undertaken. Six independently derived cell populations were generated by co−transfection of pCPWT or pCPOPP with pTCF, as described previously, creating L−929pCPWT (1 to 3) and L929CPOPP (1 to 3) cells populations respectively. The number of colonies obtained and pooled per population are shown in Table 5.2.

Preliminary characterisation of these populations were undertaken to establish that pCPWT and pCPOPP sequences had been incorporated into the L−929 cell genome. Therefore, Southern blot analysis of *Sst*I/*Hin*dIII digested DNA, prepared from L929pCPWT$_1$ and L929pCPOPP$_1$ cells, was performed using the 1,163bp *Sst*I/*Hin*dIII MuIFNα_1 promoter/*cat* gene insert of pCPWT as probe (Fig.5.8B). As shown in Figure 5.8A (Lanes 1 and 2), DNA derived from these cell populations but not L−929 cells (Fig.5.8A Lane 3) contained a 1,163bp *Sst*I/*Hin*dIII hybridising fragment which comigrated with the pCPWT DNA marker (Fig.5.8A Lane 4), thus demonstrating that they contained unrearranged copies of the MuIFNα_1 promoter/*cat* gene DNA sequences.

An additional hybridising band of approximately 1050bp was also observed in this analysis in both transfected and non−transfected cells. This band has been seen in other Southern blot analyses described in these studies using the *cat* gene probes, including those lacking MuIFNα_1 gene promoter sequences. Thus, it is likely that either sequences within the *cat* gene or 3' HSV−2 polyA region are complementary to sequences within

TABLE 5.2 Number of Colonies Obtained by Co-transfection of pCPWT or pCPOPP with pTCF, in L-929 Cells

Plasmid (10μg)	pTCF(1μg)	Cell Population	Colonies/dish
1) pCPWT	+	L929pCPWT$_1$	400
2) "	+	L929pCPWT$_2$	40
3) "	+	L929pCPWT$_3$	200
1) pCPOPP	+	L929pCPOPP$_1$	140
2) "	+	L929pCPOPP$_2$	100
3) "	+	L929pCPOPP$_3$	50
1) SSDNA	-	-	0
2) "	-	–	0
3) "	-	-	0

See legend to Table 5.1

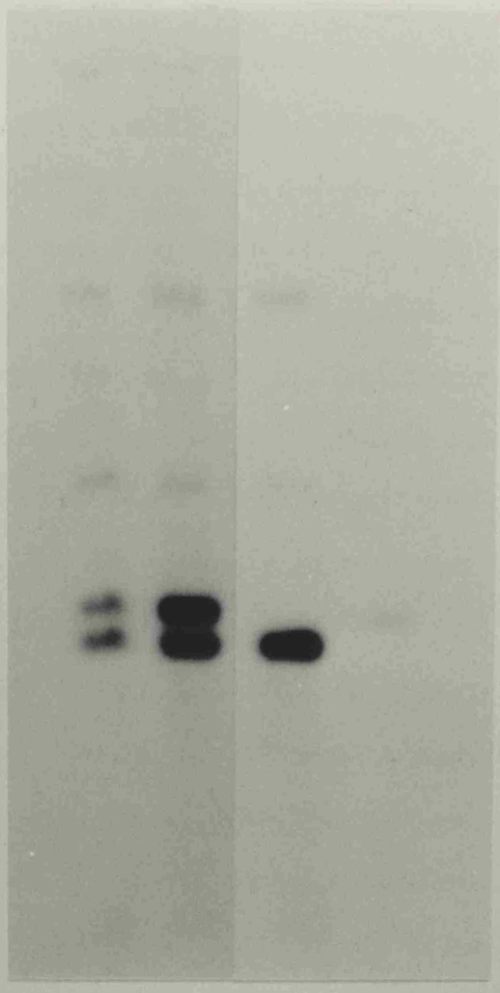
1) Transfection 1, 2) Transfection 2 and 3) Transfection 3.

FIGURE 5.8 Southern Blot Analysis of L929pCPWT$_1$ and L929pCPOPP$_1$ Cells

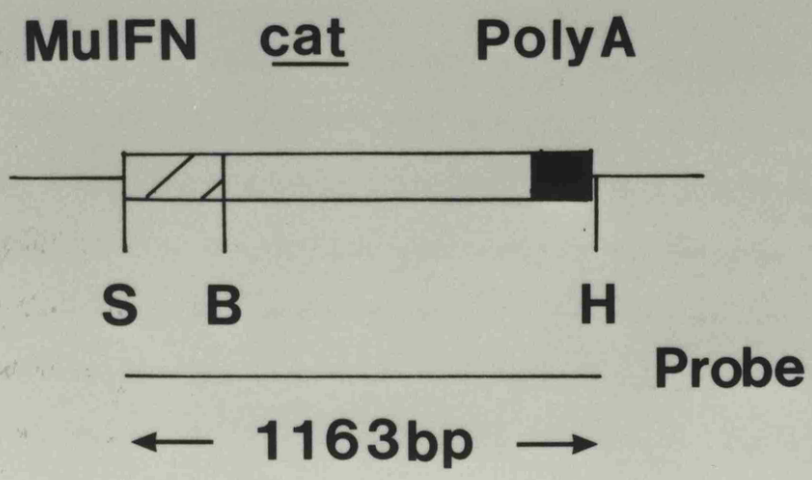
10μg samples of *Sst*I/*Hin*dIII digested genomic DNA prepared (Chapter 2) from L929pCPWT$_1$, L929pCPOPP$_1$ (see Table 5.2) and L-929 cells were loaded onto a 0.8% agarose gel and run o/n at 25V. The DNA was transferred to GeneScreenPlus^{TM} (NEN) and hybridisation performed in 5×SSPE, 1% SDS and 10% dextran sulphate at 65°C (Chapter 2) using the oligo-labelled 1163bp *Sst*I/*Hin*dIII MuIFNα_1 promoter/*cat* gene insert of pCPWT (B) as probe (5×10^{8}cpm/μg of DNA). The filter was washed in 0.1×SSC/2%SDS at 65°C and autoradiography performed at -70°C for 12hrs using Fuji RX film with an intensifying screen (A). The 1063bp hybridising fragment is indicated by an arrow.

1) 10μg L929pCPOPP$_1$ *Sst*I/*Hin*dIII Digested DNA
2) " L929pCPWT$_1$ " " "
3) " L-929 " " "
4) 1pg pCPWT " " "

A



B



L−929 cell DNA.

To investigate the functional activity of the MuIFNα_1 gene promoter in stably transfected cells, *cat* gene expression was monitored. Confluent monolayers of L929pCPWT (1) and L929pCPOPP (1 to 3) cell populations were induced with polyrI.rC under optimum conditions defined previously. As expected, CAT production was readily detectable in L929pCPWT (1) cells 24hrs post−induction (Fig.5.9A Lane c) but not in the absence of the inducer (Fig.5.9A Lane b). Furthermore, all three L929pCPWT cell populations examined responded reproducibly to induction (see below). However, with the MuIFNα_1 gene promoter in the reverse orientation (pCPOPP) or in its absence (p22; not shown), CAT production was not detectable either in the presence or absence of the inducer (Fig.5.9 Lanes d,e,f,g,h and i).

Thus, the production of CAT in L−929 cells stably transfected with pCPWT and pCPOPP has been demonstrated to be completely dependent both upon the MuIFNα_1 gene promoter being in the correct orientation and upon induction with polyrI.rC. Therefore, the 241bp *Hin*dIII fragment extending from −188 to +52 (relative to the transcription initiation site) appears to contain all the necessary information to enable it to respond in the same manner to induction as endogenous IFN genes. Physiologically relevant studies of the MuIFNα_1 gene promoter function can, therefore, be made with pCPWT related plasmids stably introduced into L−929 cells.

Quantitation of PolyrI.rC Induced CAT Expression in Stably Transfected Cell Populations

To investigate the inducibility and reproducibility of CAT expression in the independently derived cell populations L929pCPWT 1 to 3, polyrI.rC regulated activation of CAT production from the MuIFNα_1 gene promoter was quantified. For this purpose time course CAT assays were performed on non−induced and induced cell extracts. The rate of acetylation was then quantitated by scintillation counting of the acetylated and non−acetylated species of ^{14}C−chloramphenicol (Chapter 2). Figure 5.10A shows a typical time course of CAT enzyme activity and in Figure 5.10B the rate of accumulation of acetylated ^{14}C−chloramphenicol with incubation time, derived from this assay, is plotted. This plot is clearly linear. It is, therefore, possible to obtain a

FIGURE 5.9 PolyrI.rC Induced CAT Production in L-929 Cells Stably Transfected with pCPWT or pCPOPP

Confluent monolayers of cell populations (One 9cm petri-dish/population) were induced with polyrI.rC as described in Chapter 2. Cells were harvested 24hrs post-induction and CAT assays performed with 50μl of extract for 7hrs. Autoradiography was performed for 12hrs (A). Non-induced (-) cells were treated with DMEM + 2% F.C.S only (maitenance medium). Induced cells (+) were treated with 20mg/ml polyrI.rC and 800μg/ml DEAE dextran for 8hrs, washed and incubated for a further 16hrs in maintenance medium. The structure of the *cat* gene and flanking sequences of pCPWT and pCPOPP are also illustrated (B).

a) ^{14}C-chloramphenicol

b) L929pCPWT$_1$ -

c) " +

d) L929pCPOPP$_1$ -

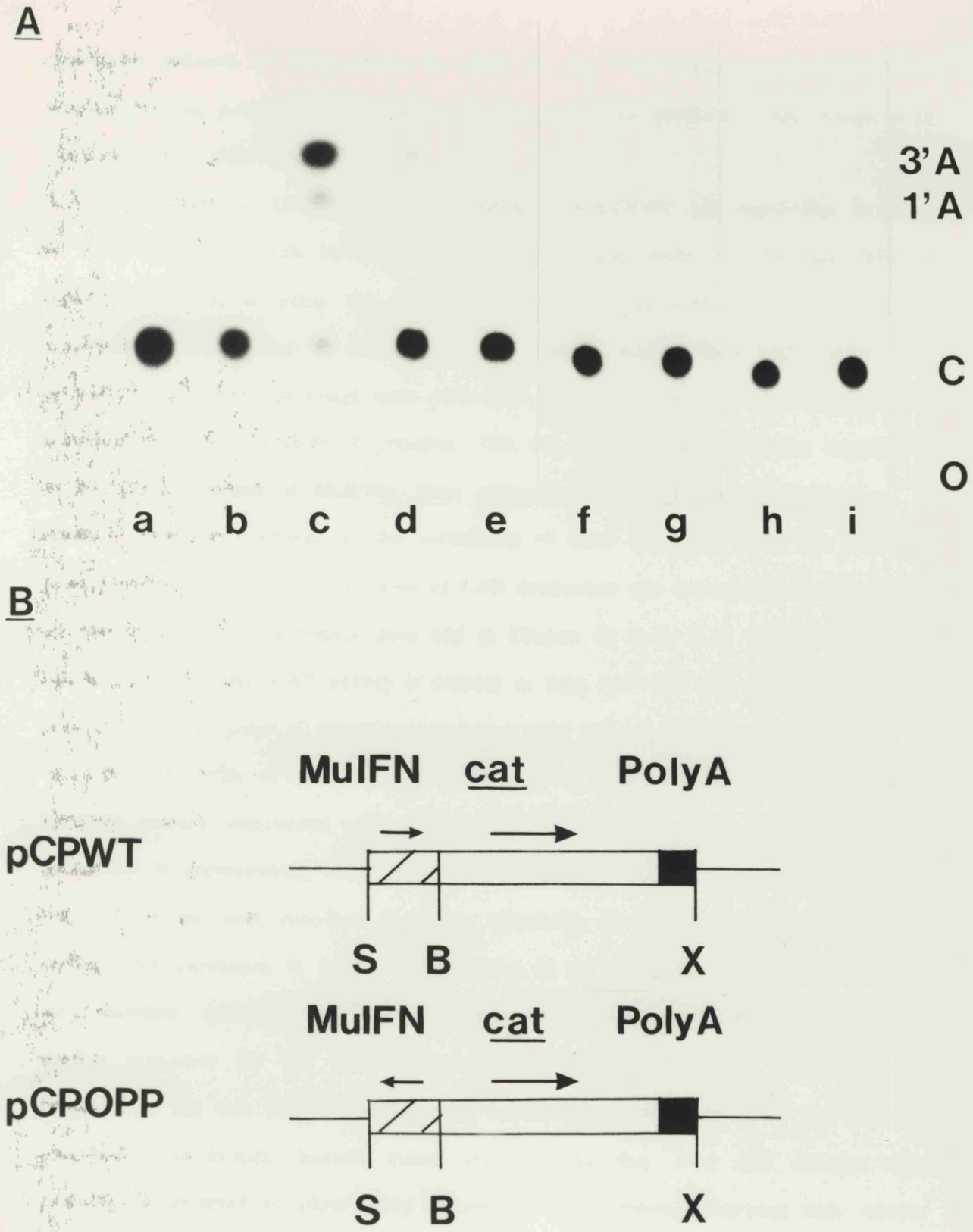
e) " +

f) L929pCPOPP$_2$ -

g) " +

h) L929pCPOPP$_3$ -

i) " +


A
3'A
1'A
C
O
a b c d e f g h i
B
MuIFN cat PolyA
pCPWT
S B X
MuIFN cat PolyA
pCPOPP
S B X

quantitative estimate of CAT activity in terms of cpm ^{14}C-chloramphenicol-3-acetate produced/min/mg protein in cell extracts. In this way CAT production was compared in independent cell populations.

A comparison of CAT production in induced L929pCPWT cell populations is shown in Table 5.3. These data indicated that CAT production could vary by upto 50% in separate inductions of either the same or different cell populations. A similar analysis undertaken in an attempt to assess CAT production in non-induced L929pCPWT cells proved difficult. Although assays were performed over eight hours, in some cases activity remained below the threshold of detection. This was a frequent difficulty also experienced in the characterisation of MuIFNα_1 gene promoter mutants (Chapter 6). Therefore, to obtain a minimum estimate for the inducibility of CAT expression from the MuIFNα_1 gene promoter, an average basal level of CAT production was assumed (determined from all non-induced extracts assayed here and in Chapter 6; Table 5.3). Using this value it can be concluded that CAT activity is induced at least 300-750 fold from the MuIFNα_1 gene promoter by polyrI.rC in stably transfected L929 cell populations.

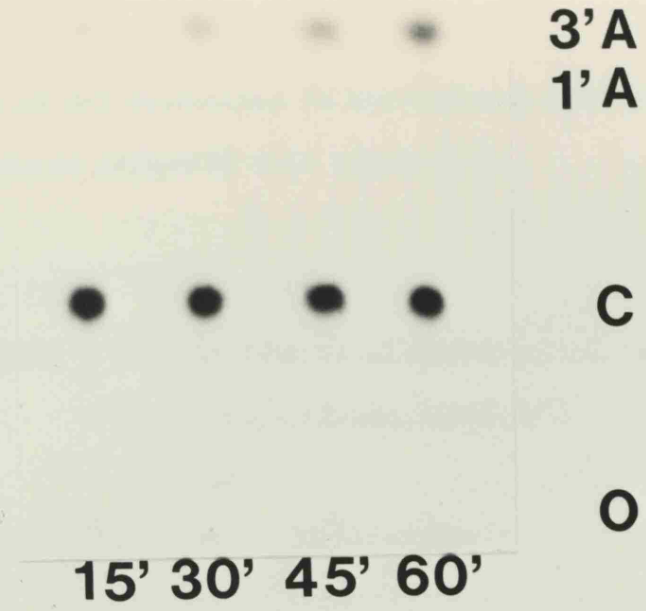
The inducibility of HuIFNα_1 and HuIFNβ genes in heterologous mouse cells quoted by other workers, determined using an S-1 nuclease protection assay to quantify mRNA production, is approximately 100 and 400 fold respectively (Ryals *et al*, 1985; Zinn *et al*, 1983). From the work described here, the inducibility of the MuIFNα_1 gene promoter driving CAT expression in L-929 cells appears to be of a very similar magnitude. It was, therefore, concluded that the MuIFNα_1 gene promoter/*cat* system represented a suitable procedure for the evaluation of inducer mediated promoter mutant activity. However, it was felt that the variation in inducibility of both the same and different transfectant populations, induced independently, mean that $\geq$ 5 fold changes must probably be achieved to allow strong conclusions to be drawn concerning their relative transcription efficiencies.

FIGURE 5.10 Graph Showing Quantitation of Cellular CAT Activity

A time course assay of CAT activity (Chapter 2) performed with 10μl of polyrI.rC induced L929pCPWT cell extract over a period of 60mins is shown (A). The percentage chloramphenicol acetylation was determined at each time point by scintillation counting (Chapter 2). The corresponding linear plot derived shows (B) the accumulation of acetylated ^{14}C-chloramphenicol with incubation time.

The vertical axis A= the % acetylation/μg of cell protein.

A


3'A
1'A
C
O
15' 30' 45' 60'

B

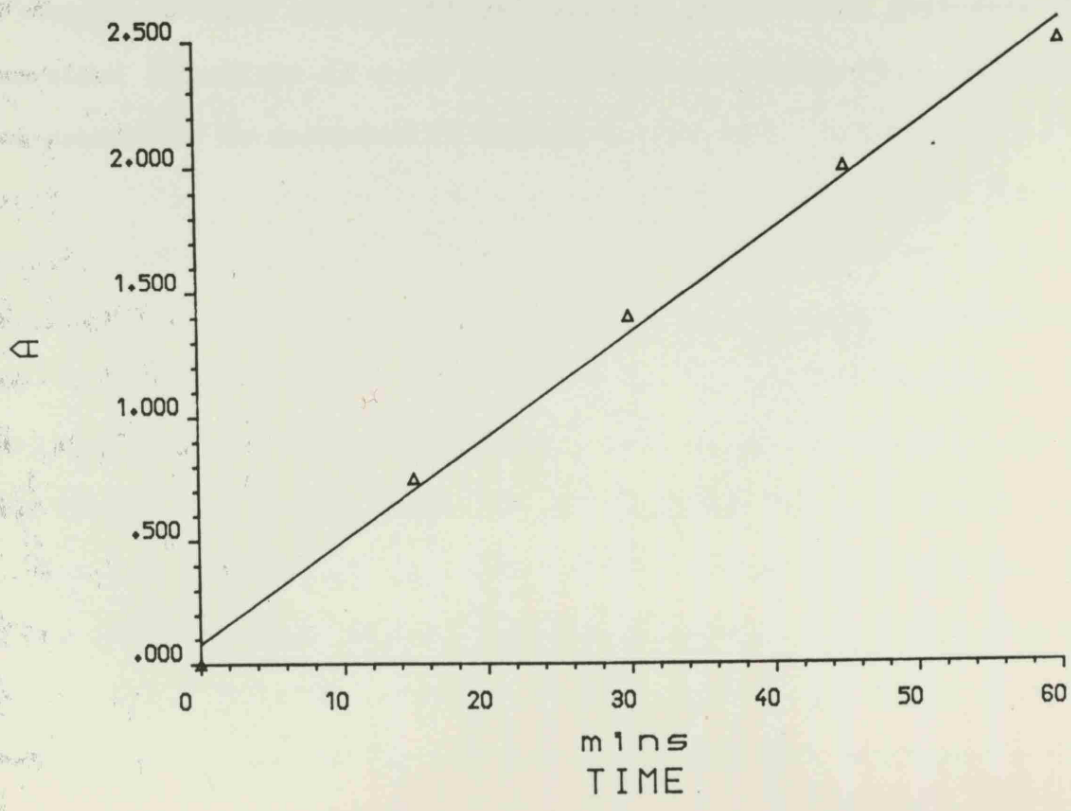

2.500
2.000
1.500
1.000
.500
.000
A
0
10
20
30
40
50
60
mins
TIME

TABLE 5.3 Comparison of CAT Production in Non-induced and PolyrI.rC/DEAE Dextran Induced L929pCPWT Cell Populations

Cell Population	CAT Activity (% Chloramphenicol Acetylation/min/μg Total Cell Protein)	Inducibility
$L929pCPWT_1$	0.03-0.038	300-380
$L929pCPWT_2$	0.038-0.075	380-750
$L929pCPWT_3$*	0.05	500
Uninduced (Average)	1×10^{-4}	

Duplicate time course CAT and protein assays were performed from independent inductions of each cell population (except *), and reaction rates determined as described in Chapter 2.

5.3 The Influence of an Enhancer on CAT Expression from the MuIFNα1 Promoter in L−929 cells

Previously it has been reported that mRNA production from the endogenous MuIFNα1 gene is undetectable in L−929 cells (Kelley and Pitha, 1985b; Zwarthoff *et al*, 1985). Several reasons may explain the discrepency between this endogenous expression data and the exogenous gene expression results described here (Section 5.2). These include a) the different inducers used, b) the chromosomal locations, c) different L−929 cells d) genetic lesions or e) that despite the readily detectable inducibility of CAT production, the MuIFNα1 gene promoter is in fact still weak relative to other MuIFNα gene promoters. To address the latter possibility further, the impact of a transcriptional activator (enhancer) upon MuIFNα1 gene promoter activity has been investigated.

Insertion of the Mo−MuSV 73bp/72bp Repeat Enhancer Proximal to the MuIFNα1 Promoter CAT Gene Hybrid

The Mo−MuSV 73bp/72bp repeat enhancer was chosen to assess the impact of such an element on polyrI.rC regulated CAT production from the MuIFNα1 promoter in mouse L−929 cells. This choice was based upon its high efficiency in mouse fibroblasts (Laimins *et al*, 1982).

Thus, vectors containing the Mo−MuSV virus enhancer in both orientations upstream of the MuIFNα1 promoter/*cat* gene were constructed. The procedure employed to create plasmids containing the *cat* gene of p22 and the 73bp/72bp repeat Mo−MuSV enhancer, in pUC13, while maintaining the *Sst*I/ *Bam*HI restriction sites for the insertion of the MuIFNα1 gene promoter module (section 5.1), is illustrated schematically in Figure 5.11. Inspection of the DNA sequence for this enhancer revealed a 204bp *Sau*3A/ *Xba*I restriction fragment (Fig.5.12A), which should be devoid of promoter activity (Kreigler and Botchan, 1983). This 204bp fragment was isolated from plasmid p43 (Fig.5.3) by P.A.G.E. and the cohesive termini were end−repaired (Chapter 2). It was then initially introduced into the *Sma*I site of M13mp10 to verify its identity by sequence analysis (Fig.5.12A). Subsequently it was ligated to *Eco*RI digested pUC13, the cohesive termini of which had also been end−repaired with DNA polymerase I treatment

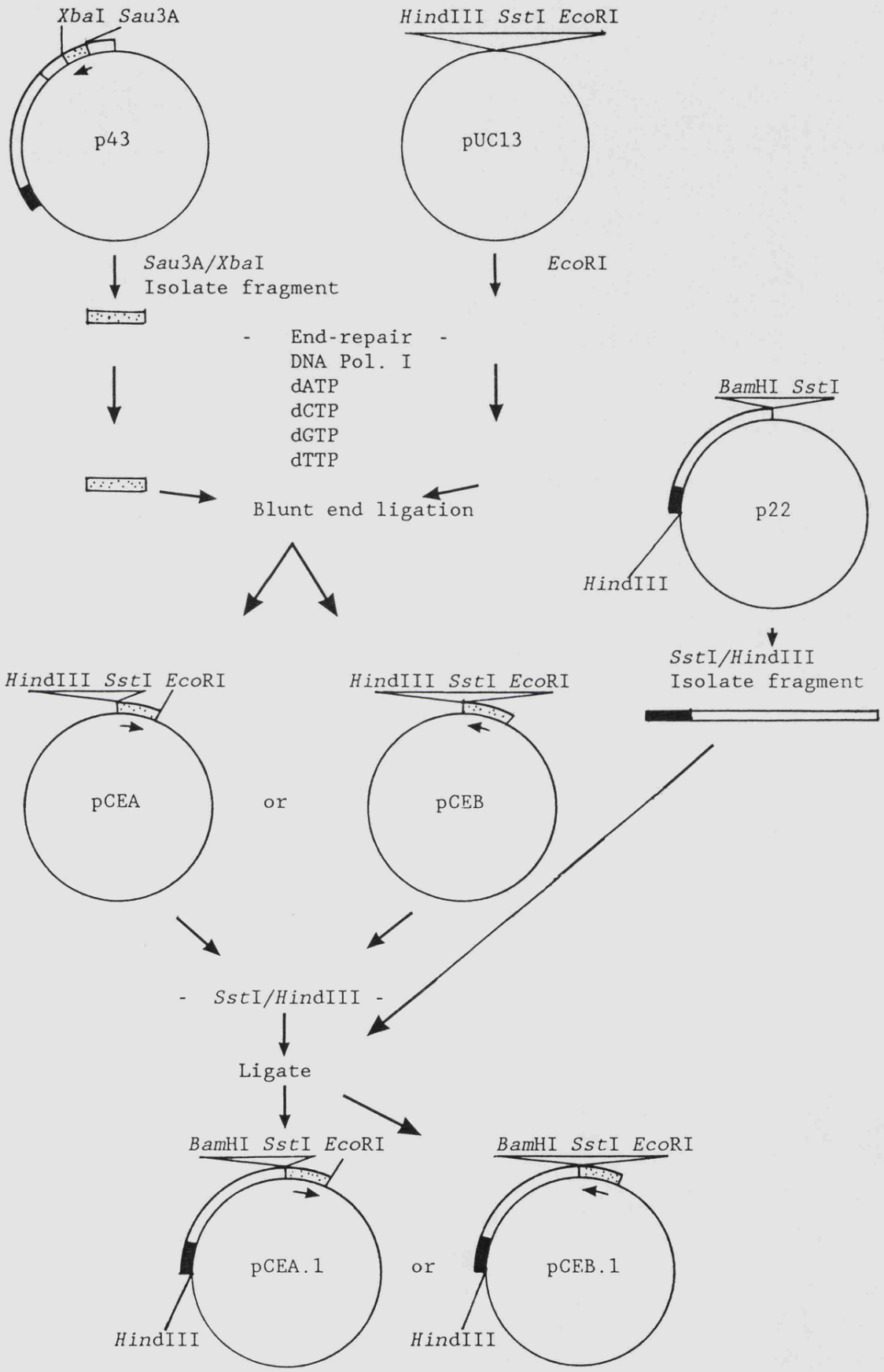


FIGURE 5.11 STRATEGY FOR THE CONSTRUCTION OF pCEA.1 and pCEB.1

Stippled box = Mo-MuSV enhancer *Sau*3A/*Xba*I fragment, Open box = *cat* gene and filled box = HSV-2 PolyA. Only restriction sites essential to construction are shown.

A

```
                                                    Sau3A
                                                          10        20
                                                   GATCAAGGTCAGGAACAGAGAG

 PvuII
       30        40        50        60        70        80        90
ACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGCTCAGGGCCAAGAACAGATGGA

 PvuII
   100       110       120       130       140       150       160
ACAGCTGAAT TGGGGCAAACAGGATATCTGCGGTAAGCAGTTCCTGCCCCGCTCAGGGCCAAGAACAGATGG

                                 XbaI
170       180       190       200
TCCCCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGA
```

B

```
                EcoRI
pUC13:      ...GAATTC...  EcoRI  AATTC...   End Repair  AATTC...
            ...CTTAAG...              G...                TTAAG...

                Sau3A
Enhancer:   ...GATC...    Sau3A   GATC...   End Repair   GATC...
            ...GATC...                ...                GATC...

                XbaI
Enhancer:   ...TCTAGA...  XbaI    CTAGA...  End Repair   CTAGA...
            ...AGATCT...              T...                GATCT...

                Sau3A/EcoRI                      XbaI/EcoRI *
Ligation:   ...GATC AATTC...                 ...TCTAG AATTC...
            ...CTAG TTAAC...                 ...AGATC TTAAG...
```

FIGURE 5.12 Primary Nucleotide Sequence of the 204bp *Sau*3A/*Xba*I Mo-MuSV Enhancer Fragment (Laimins *et al*, 1982)

The primary nucleotide sequence of the Mo-MuSV virus LTR enhancer fragment used in these studies is illustrated (A). The 73bp/72bp enhancer repeat sequence is shown underlined. Restriction sites for *Pvu*II, *Sau*3A and *Xba*I are underlined. Below (B) are illustrated the sequences created by end-repair and blunt-end ligation of *Eco*RI, *Sau*3A or *Xba*I generated cohesive termini. Blunt-end ligation of end-repaired *Eco*RI/*Xba*I sites recreates the *Eco*RI recognition sequence *.

and then transformed into *E.coli* JA221. One hundred ampicillin resistant colonies were screened by colony hybridisation (Chapter 2) using the 204bp *Xba*I/ *Sau*3A enhancer fragment as probe. This identified 27 positively hybridising clones. Because the enhancer fragment contains two characteristic *Pvu*II sites (Fig.5.12A) and the cloning procedure should recreate an *Eco*RI site at the *Eco*RI/ *Xba*I junction (Fig.5.12B), *Eco*RI/ *Pvu*II digests were performed to confirm that the plasmids designated pCEA and pCEB (Fig.5.11) had been created. Eight prospective recombinant plasmid DNAs examined, revealed six with the expected restriction pattern for pCEA and two for pCEB (Fig.5.13).

To complete these manipulations a 922bp *Sst*I/ *Hin*dIII fragment containing the *cat* gene and HSV−2 polyA sequences was excised from plasmid p22 (Fig.5.3), isolated by P.A.G.E., ligated to *Sst*I/ *Hin*dIII digested pCEA or pCEB and transformed into *E.coli* JA221. Plasmid DNAs prepared from four ampicillin resistant colonies for both ligations were digested with *Sst*I/ *Hin*dIII and analysed by P.A.G.E. This revealed, in both cases, two out of four recombinants examined to be the plasmids designated pCEA.1 and pCEB.1 (Fig.5.11). Their restriction maps (Fig.5.13) were confirmed by *Eco*RI/ *Pvu*II and *Eco*RI/ *Bst*NI double digests, analysed by P.A.G.E. Plasmids pCEAPWT and pCEBPWT containing the MuIFNα_1 gene promoter were then created by inserting the *Sst*I/ *Bam*HI promoter module (from mPWT) into the *Sst*I/ *Bam*HI site of pCEA.1 and pCEB.1 respectively. Again the restriction maps (Fig.5.13) were verified as described above.

Thus, these manipulations generated two expression vectors identical to pCPWT except for the presence of the Mo−MuSV enhancer in both orientations proximal to the MuIFNα_1 gene promoter.

Production of CAT from the MuIFNα_1 Promoter in the Presence of an Enhancer, in Stably Transfected L929 Cells

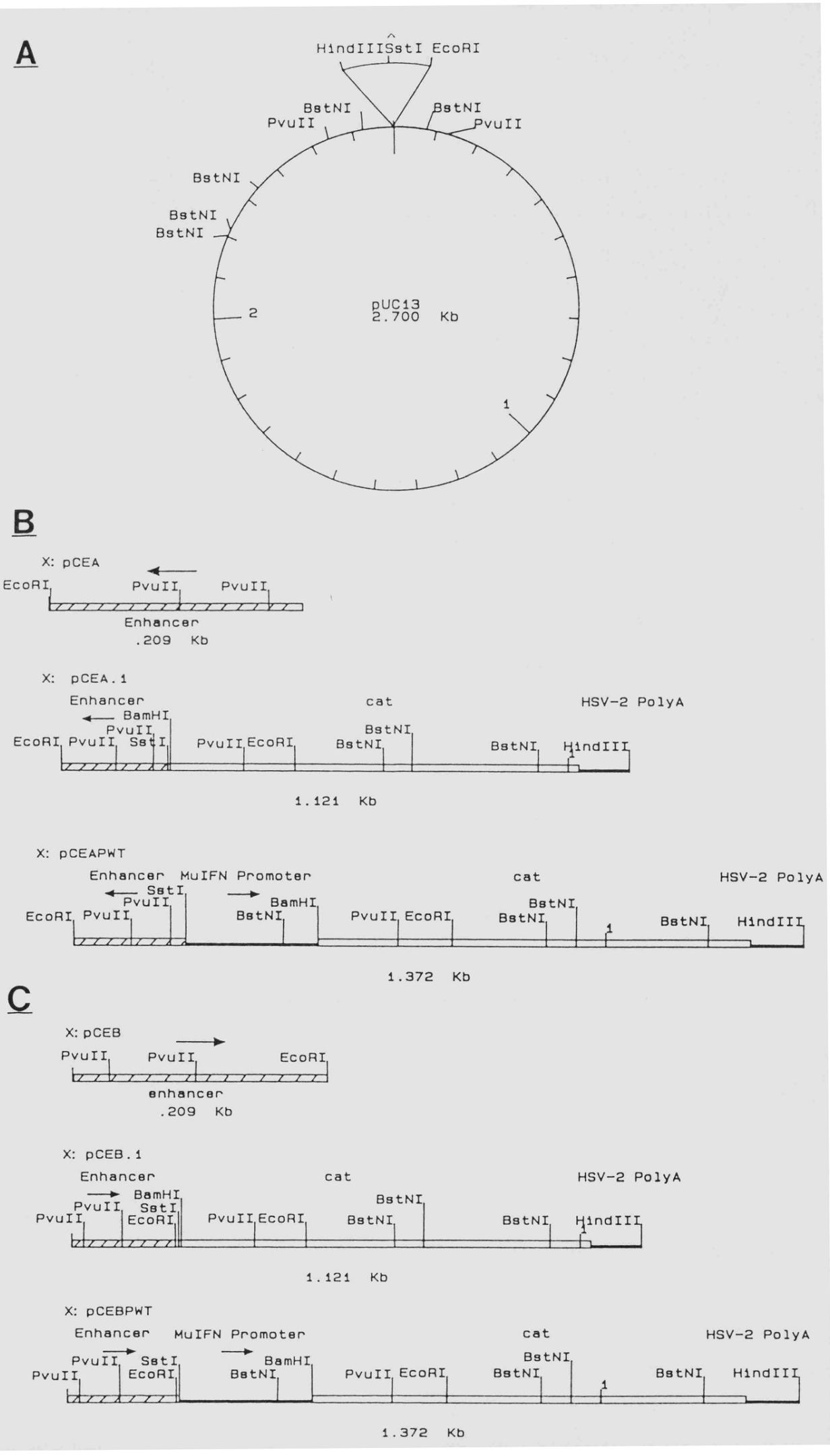
The plasmids pCEAPWT and pCEBPWT were used to investigate the impact of the Mo−MuSV enhancer on CAT expression from the MuIFNα_1 gene promoter. Initially their expression was examined by a transient transfection assay. However, this analysis, as before (Section 5.2), failed to demonstrate efficient polyrI.rC mediated production of

FIGURE 5.13 Restriction Maps of Mo-MuSV Enhancer Constructs

A) Shows a partial restriction map of pUC13 for *Bst*NI, *Eco*RI, *Hin*dIII, *Pvu*II and *Sst*I, illustrating sites vital both for the construction of the enhancer constructs and those important for determining the orientation of specific fragments and verifying the structure of recombinant plasmids. Derivatives contain fragments which have been inserted at region X.

B) Shows vital restriction sites of fragments introduced at region X into pUC13 to create the derivatives pCEA, pCEA.1 and pCEAPWT. The *Sau*3A/*Xba*I 204bp enhancer fragment was introduced into the *Eco*RI site of pUC13, recreating the *Eco*RI site at the *Xba*I/*Eco*RI junction (Fig. 5.13), to produce pCEA. Subsequently the *Sst*I/*Hin*dIII *cat* gene fragment from p22 was introduced to create pCEA.1. Finally the *Sst*I/*Bam*HI MuIFNα_1 gene promoter module from mPWT was introduced to produce pCEAPWT.

C) As for B) but with the enhancer in the opposite orientation. Restriction maps in B) and C) are not drawn to the same scale as pUC13 in A).


A
HindIII SstI EcoRI
BstNI
PvuII
BstNI
PvuII
BstNI
BstNI
BstNI
pUC13
2.700 Kb
2
1
B
X: pCEA
EcoRI
PvuII
PvuII
Enhancer
.209 Kb
X: pCEA.1
Enhancer
cat
HSV-2 PolyA
BamHI
PvuII
BstNI
EcoRI PvuII SstI
PvuII EcoRI
BstNI
BstNI
HindIII
1.121 Kb
X: pCEAPWT
Enhancer MuIFN Promoter
cat
HSV-2 PolyA
SstI
PvuII
BamHI
BstNI
EcoRI PvuII
PvuII EcoRI
BstNI
BstNI
BstNI
HindIII
1.372 Kb
C
X: pCEB
PvuII
PvuII
EcoRI
enhancer
.209 Kb
X: pCEB.1
Enhancer
cat
HSV-2 PolyA
BamHI
PvuII
SstI
BstNI
PvuII
EcoRI
PvuII EcoRI
BstNI
BstNI
HindIII
1.121 Kb
X: pCEBPWT
Enhancer MuIFN Promoter
cat
HSV-2 PolyA
PvuII
SstI
BamHI
BstNI
PvuII
EcoRI
BstNI
PvuII EcoRI
BstNI
BstNI
HindIII
1.372 Kb

CAT in L−929 cells.

Subsequently, their expression was examined in stably transfected L−929 cells. The stably transfected cell populations listed in Table 5.4 were, therefore, generated by co−transfection of the corresponding plasmids with pTCF as described previously.

CAT production in these cell populations was analysed using the polyrI.rC induction conditions defined previously. Surprisingly, a comparison of induced CAT production in cell populations containing pCPWT or pCEBPWT revealed approximately equal levels of expression in both cases (Fig.5.14 Lanes c, and e). Similar results were obtained with the enhancer in the opposite orientation (see below). Interestingly, non−induced CAT production was not readily detectable in these cell populations or in induced or non−induced cell populations containing the promoterless constructs, pCEA.1 and pCEB.1 (Fig.5.14 Lanes b,d,f,g,h and i).

Thus, the MuIFNα_1 gene promoter is still activated in a regulated manner even in the presence of the Mo−MuSV enhancer in either orientation. The enhancer, therefore, has had no detectable influence on polyrI.rC mediated MuIFNα_1 gene promoter activity.

To evaluate the impact of this enhancer on the MuIFNα_1 gene promoter activity quantitative evaluation of CAT production in non−induced and induced cells containing pCEAPWT (L929pCEAPWT 1 and 2) and pCEBPWT (L929pCEBPWT 1 and 2) were made in parallel to the studies described with the enhancerless MuIFNα_1 gene promoter (L929pCPWT 1 to 3). A plot of typical quantitative CAT assays performed are illustrated in Figure 5.15, which shows a comparison of CAT production from the MuIFNα_1 gene promoter in the presence and absence of enhancer sequences. The results are summarised in Table 5.5. Thus, perhaps rather surprisingly, neither the presence nor absence of the Mo−MuSV enhancer appeared to have an effect on the maximal levels of polyrI.rC mediated CAT production in L−929 cells, 24hrs after induction. This enhancer, therefore, does not appear to potentiate the strength of the MuIFNα_1 gene promoter.

However, when the levels of non−inducible expression from the MuIFNα_1 gene promoter in the presence of this enhancer were examined, a reproducible elevation of CAT production was observed (Figure 5.15B: Table 5.5). A minimum estimate of a 2 to 5 fold increase in basal CAT production was obtained. The elevation of basal expression

TABLE 5.4 Number of Colonies Obtained by Co-transfection of pCEA.1, pCEAPWT, pCEB.1, or pCEBPWT with pTCF, in L-929 Cells

Plasmid(10μg)	pTCF(1μg)	Cell Population	Colonies/dish
1) pCEA.1	+	L929pCEA.1_1	110
2) "	+	L929pCEA.1_2	440
1) pCEAPWT	+	L929pCEAPWT$_1$	120
2) "	+	L929pCEAPWT$_2$	100
1) pCEB.1	+	L929pCEB.1_1	110
2) "	+	L929pCEB.1_2	160
1) pCEBPWT	+	L929pCEBPWT$_1$	270
2) "	+	L929pCEBPWT$_2$	80
1) SS DNA	-	-	0
2) "	-	-	0

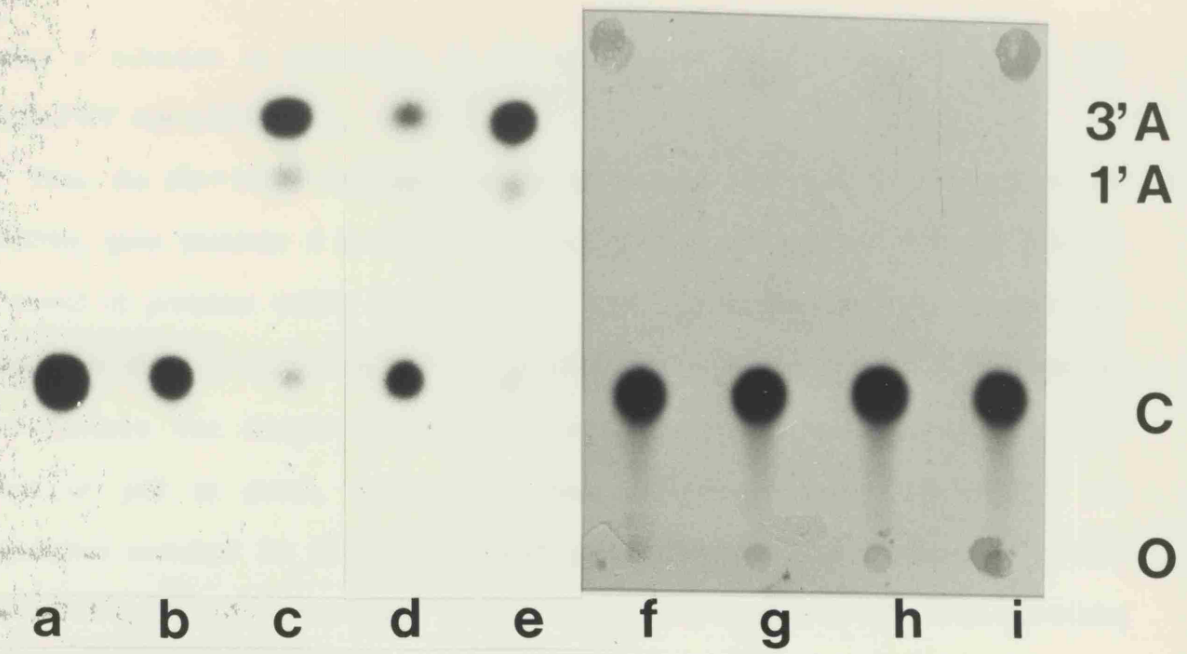
See legend to Table 5.1

FIGURE 5.14 PolyrI.rC Induced CAT Production in L-929 Cells Stably Transfected with pCPWT, pCEA.1, pCEB.1 and pCEBPWT

See legend to Fig.5.9. Autoradiography was performed for 12hrs using Fuji RX film (A). The structure of the *cat* gene and flanking sequences of pCPWT, pCEA.1, pCEB.1 and pCEBPWT are also illustrated (B). PolyrI.rC induced (+), non-induced (-).

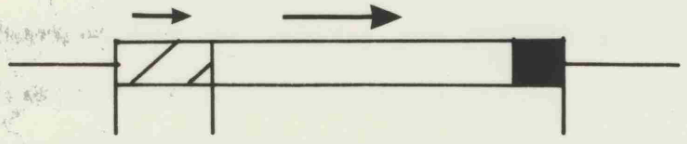
a) ^{14}C-chloramphenicol
b) L929pCPWT$_1$ -
c) " +
d) L929pCEBPWT -
e) L929pCEBPWT +
f) L929pCEA.1 -
g) " +
h) L929pCEB.1 -
i) " +

A



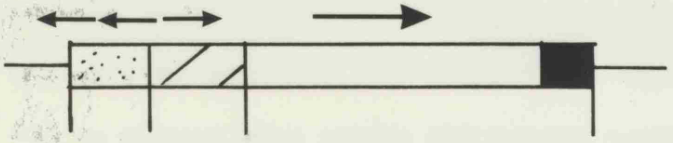
MuIFN cat PolyA

B



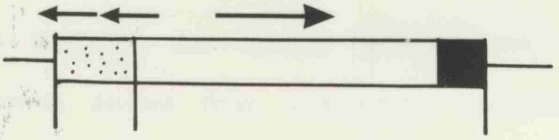
S B X

En MuIFN cat PolyA



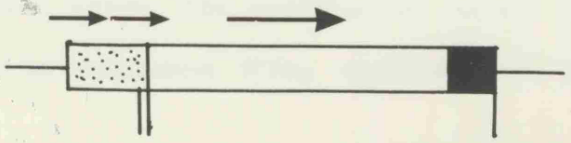
E S B X

En cat Poly A



E B X

En cat PolyA



E B X

caused a reduction in inducibility of 300 to 750 fold (pCPWT) to 80 to 200 fold (pCEAPWT and pCEBPWT).

Thus, the Mo−MuSV enhancer appeared to increase basal gene expression from the MuIFNα_1 gene promoter 2 to 5 fold. Although CAT assays suggested that the enhancer is devoid of promoter activity (Fig.5.14 Lanes f,g,h and i), the possibility remained that these cells did not contain pCEA.1 or pCEB.1 sequences. To prove this was not the case Southern blot analysis was performed using the 922bp *Bam*HI/ *Hin*dIII *cat* gene insert of p22 as probe, demonstrating that L929pCEA.1 and L929pCEB.1 cell populations contained the 922bp *Sst*I/ *Hin*dIII promoterless *cat* gene sequences of pCEA.1 and pCEB.1 (Fig.5.16 Lanes 1 and 2). The 1050bp *Sst*I/ *Hin*dIII L−929 DNA hybridising fragment, described previously, was also observed.

Thus, in the presence of the Mo−MuSV enhancer regulated activation of CAT expression from the MuIFNα_1 gene promoter is maintained, although basal expression is elevated. These observations may be interpreted in several ways. The inability to increase CAT production suggests that the promoter is fully activated by polyrI.rC, and enhancer sequences are unable to increase expression further. Alternatively, the enhancer may not be functioning as expected, although elevated basal gene expression and data to be presented in Chapter 7 strongly suggest the 204bp *Sau*3A/ *Xba*I fragment does retain enhancer−like properties. It is intriguing that regulated promoter function is maintained with an enhancer. One possible interpretation would be that the MuIFNα_1 gene promoter is suppressed in the non−induced state, preventing any upstream transcriptional activator sequences from functioning. Such possibilities will be discussed further in Chapter 8.

The inability of the enhancer to further increase polyrI.rC regulated promoter activity suggested the induced promoter may be fully activated, contrary to the expectations derived from undetectable endogenous L−929 MuIFNα_1 gene expression observed by others (Kelley and Pitha, 1985; Zwarthoff *et al*, 1985). Obviously firm conclusions regarding the strength of the MuIFNα_1 gene promoter relative to other MuIFNα gene promoters would require a direct comparison, perhaps employing this CAT expression system. The possibility still remains that the discrepency between published results on endogenous IFNα_1 expression and these reports on the behaviour of an

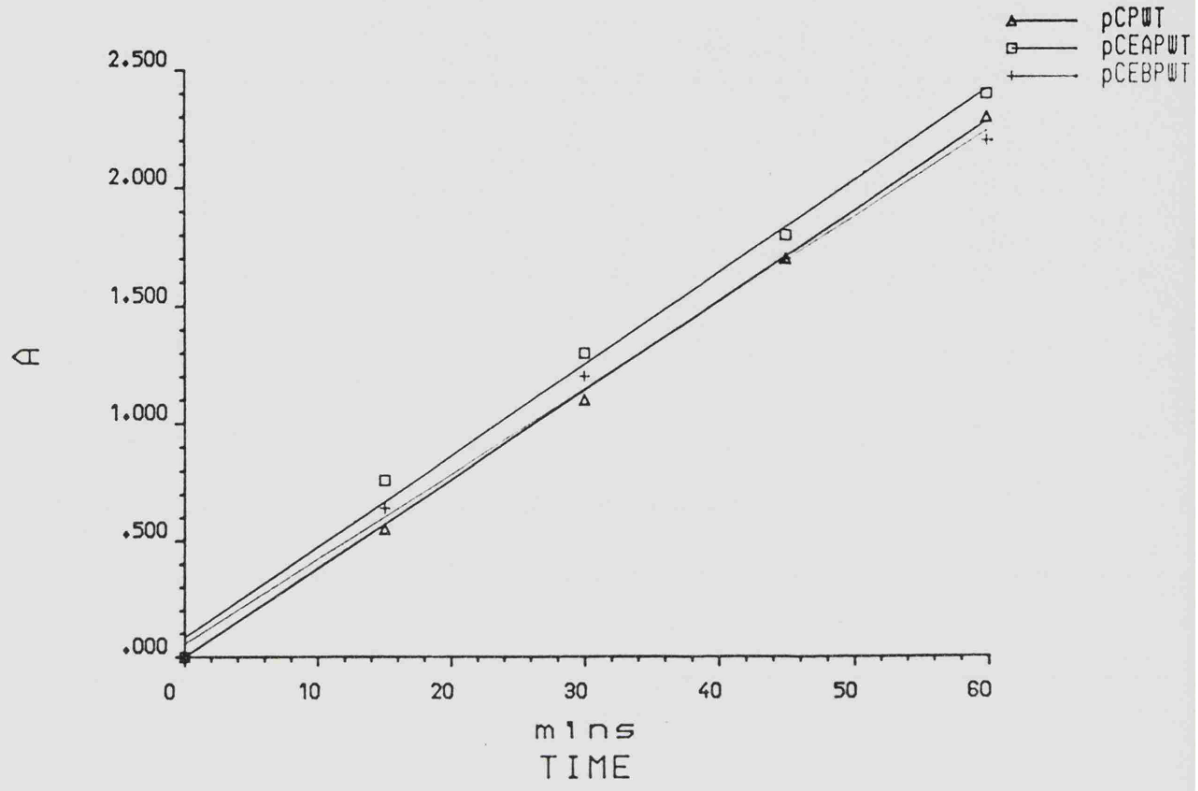
FIGURE 5.15 Graph Showing a Comparison of CAT Production from the MuIFNα_1 Gene Promoter in the Presence and Absence of the Mo-MuSV Enhancer

See legend to Fig.5.10.

A) Parallel time course CAT assays of 10μl of extract from polyrI.rC induced L-929 cells stably transfected with the indicated plasmid, performed over 60mins. A= % acetylated ^{14}C-chloramphenicol per μg of cell extract.

B) Parallel time course CAT assays of 50μl of extract from non-induced L-929 cells stably transfected with the indicated plasmid, performed over 8hrs (Chapter 2). A= % acetylated ^{14}C-chloramphenicol per μg of cell protein.

A


pCPWT
pCEAPWT
pCEBPWT
2.500
2.000
1.500
1.000
.500
.000
A
0
10
20
30
40
50
60
mins
TIME

B

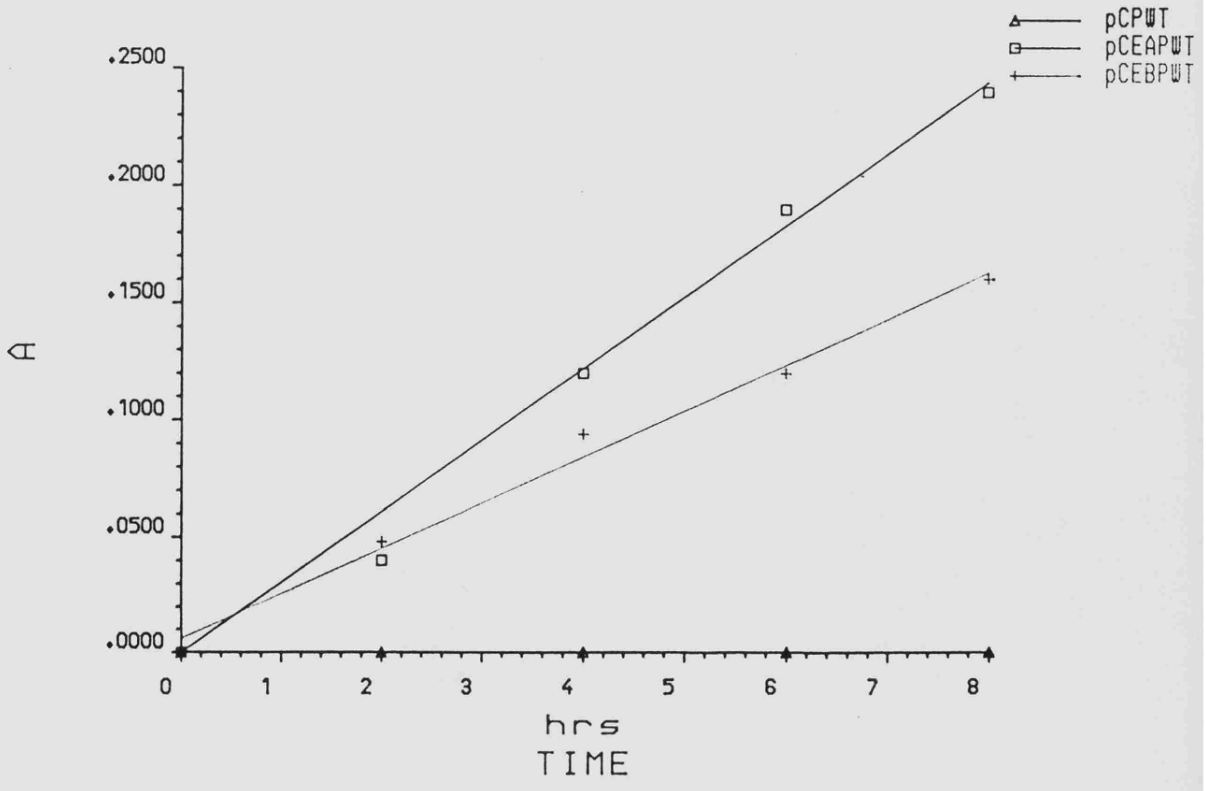

pCPWT
pCEAPWT
pCEBPWT
.2500
.2000
.1500
.1000
.0500
.0000
A
0
1
2
3
4
5
6
7
8
hrs
TIME

TABLE 5.5 Comparison of CAT Production in Non-induced and PolyrI.rC/DEAE Dextran induced L929pCEAPWT and L929pCEBPWT

Cell Population	CAT Activity (% Chloamphenicol Acetylation/min/μg Total Cell Protein		Inducibility
	Induced	Uninduced ($\times 10^4$)	
L929pCEAPWT$_2$	0.054	5.3	100
L929pCEAPWT$_3$	0.04	4.7	90
L929pCEBPWT$_1$*	0.04-0.059	3.3-5.4	70-180
L929pCEBPWT$_3$	0.025	1.3	190

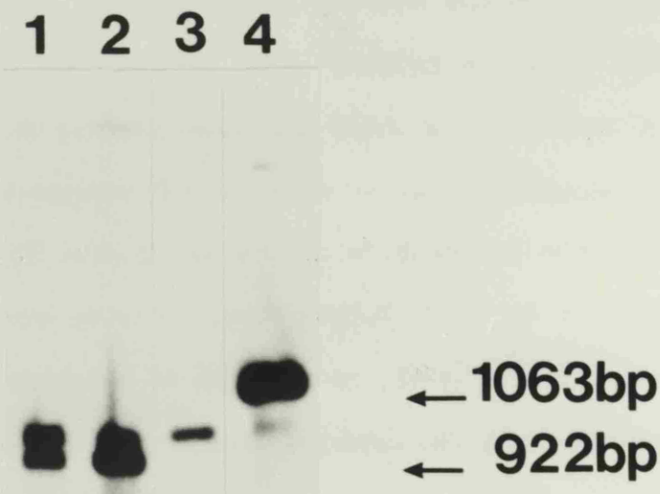
Time course CAT and protein assays were performed on extracts from induced and non-induced cell populations and the reaction rates determined as described in Chapter 2.

* Duplicate assays performed from independent inductions

FIGURE 5.16 Southern Blot Analysis of L929pCEA.1_2 and L929pCEB.1_2 Cells

See legend to Fig.5.8. An oligolabelled 922bp *Bam*HI/*Hin*dIII *cat* gene fragment of p22 (B) was used as probe (5×10^8cpm/μg DNA). Autoradiography was performed at -70°C for 15hrs using Fuji RX film with an intensifying screen (A). The upper arrow indicates the 1163bp *Sst*I/*Hin*dIII pCPWT marker fragment. The lower arrow indicates the 922bp hybridising fragment is indicated by an arrow.

1) 10μg of L929pCEA.1_2 *Sst*I/*Hin*dIII digested DNA
2) " L929pCEB.1_2 " " "
3) " L-929 " " "
4) 100pg pCPWT " " "


1 2 3 4
1063bp
922bp

exogenous MuIFNα_1/*cat* gene may result from either a positional effect or a genetic lesion of the endogenous L−929 MuIFNα_1 gene.

5.4 Conclusions

This study reveals that based on both comparative structural and functional studies, the 241bp *Hin*dIII fragment from the MuIFNα_1 gene, identified as having significant homology with human IFNα_1 gene promoter sequences which are implicated in viral regulation of gene expression, can regulate the expression of the heterologous bacterial gene *cat* in stably transfected L−929 cells. In the absence of an inducer very low levels of CAT are produced by the hybrid gene but upon polyrI.rC treatment the MuIFNα_1 gene promoter directs efficient transcription of the *cat* gene. Thus, the 241bp promoter fragment contains sequences required for polyrI.rC activation of gene expression in L−929 cells.

In contrast to studies employing stably transfected L−929 cell populations, transient assays failed to reveal significant MuIFNα_1 gene promoter activity in induced L−929 cells, even in the presence of an enhancer. In view of the demonstrable MuIFNα_1 promoter activity *in vivo*, the most likely explanation for this observation is that optimal transient expression and polyrI.rC induction protocols are probably incompatible.

Therefore, the stable transfection system has been employed to compare the functional activity of chimaeric constructs containing the bacterial *cat* gene and the MuIFNα_1 gene promoter in the presence and absence of the Mo−MuSV enhancer, in independent cell populations induced with polyrI.rC. Quantitative CAT assays demonstrated that induced CAT production can vary by up to 50% in the same or independently derived cell populations. Surprisingly they also revealed that regulated promoter function is maintained in the presence of an enhancer, although a reproducible 2 to 5 fold enhancement of basal gene expression was observed. The refractory behaviour of the MuIFNα_1 promoter to the enhancer is interesting and will be discussed further in Chapter 8.

Thus, these studies have confirmed the suitability of this system for examining the behaviour of closely related MuIFNα_1 promoter mutants present in independent cell

populations. Firm conclusions regarding the relative strength of promoters, however, would require $\geq$5 fold changes in activity. The major sources of the observed variation are likely to result from a) differences in independent cell populations and b) the efficiency of polyrI.rC induction. Pooling colonies significantly reduces the variation of independently transfected cell lines which result from random chromosomal integrations and variable copy number of the newly introduced genes (Chapter 1). However, variation in the efficiency of polyrI.rC induction, inevitably, cannot be avoided in this manner.

The CAT assay is clearly very useful for analysing IFN promoter function in homologous cells. The ease and relative reproducibility of CAT assays combined with a simple means of quantitation has demonstrated the suitability of this system for MuIFNα_1 promoter functional studies in L$-$929 cells.

Thus, a suitable MuIFNα_1 gene promoter module has been constructed which can be used both as a substrate for site directed mutagenesis and for expression studies when linked to the plasmid p22. Therefore, a system has been devised for the functional analysis of the MuIFNα_1 gene promoter, which can be employed to compare the relative activity of mutant derivatives present in independent cell populations. Such studies are described in detail in Chapter 6.

CHAPTER 6

STRUCTURE/FUNCTION STUDIES OF THE MuIFNα_1 PROMOTER PART I: ANALYSIS OF PROMOTER MUTANTS USING THE CAT EXPRESSION ASSAY

The generation of mutants of cloned genes *in vitro* together with *in vitro* or *in vivo* expression studies (Chapter 1) has proved invaluable for investigating promoter function. Deletion mutagenesis, commonly exploited to define both the location and boundaries of functionally important elements of gene promoters, including HuIFNα_1 and HuIFNβ genes (Chapter1), has the prospective disadvantage of producing gross structural changes in genetic organisation. Ultimately, it will be important to identify the precise nucleotides contributing to promoter function. Site directed mutagenesis, a procedure for introducing predetermined point mutations into a cloned gene (Razin *et al*, 1978), together with *in vivo* expression studies, provides a means to this end.

Chapter 5 described polyrI.rC mediated activation of CAT expression from the 241bp *Hin*dIII MuIFNα_1 promoter fragment (-188 to $+52$) in stably transfected L$-$929 cells. To investigate those sequences which might contribute to the regulation and strength of polyrI.rC induced IFN gene expression, mutant promoters have been constructed by site directed mutagenesis and their behaviour examined using the CAT expression assay in stably transfected cell lines.

6.1 Location of Prospective Functional Sequences in Mouse IFNα Gene Promoters

Prospective functional sites in the MuIFNα_1 promoter were identified as potential targets for site directed mutagenesis. Initially to establish the probable locations of regulatory sequences, the inter$-$ and intra$-$species homology of IFNα promoter sequences was examined.

Primary Nucleotide Structure of MuIFNα_1 Gene Promoters

Mouse IFNα is encoded by a family of genes, analogous to the HuIFNα gene family (Chapter 3 and 4). Both cDNA clones (Shaw *et al*, 1983, Kelley *et al*, 1983) and

genomic clones (Shaw *et al*, 1983; Daugherty *et al*, 1984; Zwarthoff *et al*, 1985; Kelley and Pitha, 1985a) have been isolated and seven members of the mouse IFNα gene family characterised in some detail. Southern blot analysis of *Eco*RI digested spleen DNA from AKR and C57 mice and NIH/3T3 cell DNA under low stringency conditions (2×SSC, 65°C) using the MuIFNα cDNA probe, pMIF1204 (Kelley *et al*, 1983), revealed at least 10 distinct hybridising fragments (Fig.6.1), strongly suggesting the existence of more than seven MuIFNα genes.

To identify structural features common to the IFNα gene promoters, the 5' sequences of six of the currently characterised mouse genes and HuIFNα_1 were compared. For this purpose the promoters were aligned with the 241bp promoter sequence of the MuIFNα_1 gene (−188 to +52) as described in Chapter 5 (Fig.6.2). The striking feature of this comparison is the degree of homology shared by these sequences. Within a region of 166 nucleotides from −113 to +52 (relative to the transcription initiation site) the intra− and inter−species homology is 75% to 93% and 63% to 70% respectively. Through this analysis, sequence motifs were recognised and designated regions I to IV (Fig.6.2). These are discussed separately below.

Region I

This highly conserved region contains a TATTTAA sequence (Fig.6.2) variant of the TATA box, common to many genes transcribed by RNA polymerase II (Corden *et al*, 1980). This sequence, found in all currently characterised human (except IFNα_5, TTTTTAA), mouse, rat and bovine IFNα genes (Chapter1), is strongly implicated in promoter function. Human and mouse IFNβ genes have the better consensus sequence TATAA. Since 80% of polyrI.rC induced IFN production by L−929 cells is IFNβ (Trapman *et al*, 1980), the effect upon MuIFNα_1 promoter activity of changing this region to more closely resemble the consensus, like IFNβ genes, was investigated. For this purpose the mutant promoter P1 was constructed by site directed mutagenesis (Fig.6.3).

FIGURE 6.1 Southern Blot Analysis of Mouse DNAs with pMIF1204

10μg of *Bam*HI digested mouse DNAs were loaded onto a 0.8% agarose gel and run o/n at 25V. The DNA was transferred to nitrocellulose (Chapter 2) and hybridisation performed in 5× Denhardt's (2×SSC, 65°C o/n) using the nick translated 820bp *Pst*I insert from pMIF1204 as probe (1×10^8 cpm/μg DNA). The filter was washed in 2×SSC at 65°C and autoradiography performed at -70°C for 14 days using Fuji RX film with an intensifying screen.

1) AKR mouse spleen DNA

2) C57 " " "

3), 4) and 5) NIH/3T3 cell DNA

6) Human Placental DNA

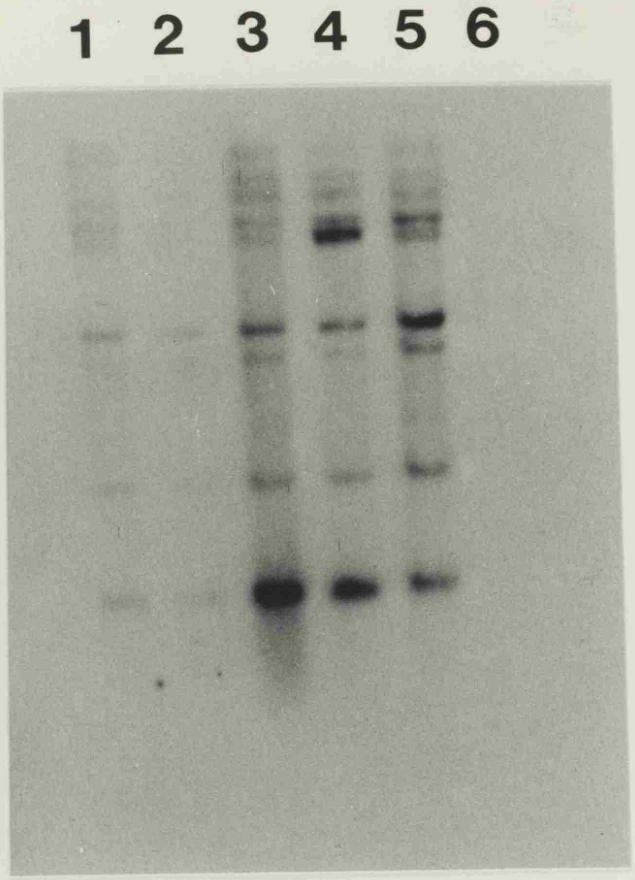

1 2 3 4 5 6

FIGURE 6.2 Comparison of Mouse and Human IFNα Gene Promoter Sequences

The MuIFNα_1 promoter primary nucleotide sequence is illustrated and below is aligned other MuIFNα and the HuIFNα_1 gene promoter sequences. Gaps have been incorporated to achieve optimum homology. The numbering of nucleotides is based on the MuIFNα_1 sequence. Conserved nucleotides are denoted by a dash (–). Specific nucleotide substitutions are shown using the conventional symbols. Sequences of prospective functional significance (described in the text) are underlined (bold type) and labelled regions I to IV. The TATA box (region I) and prospective transcription initiation sites (CAP) are also indicated.

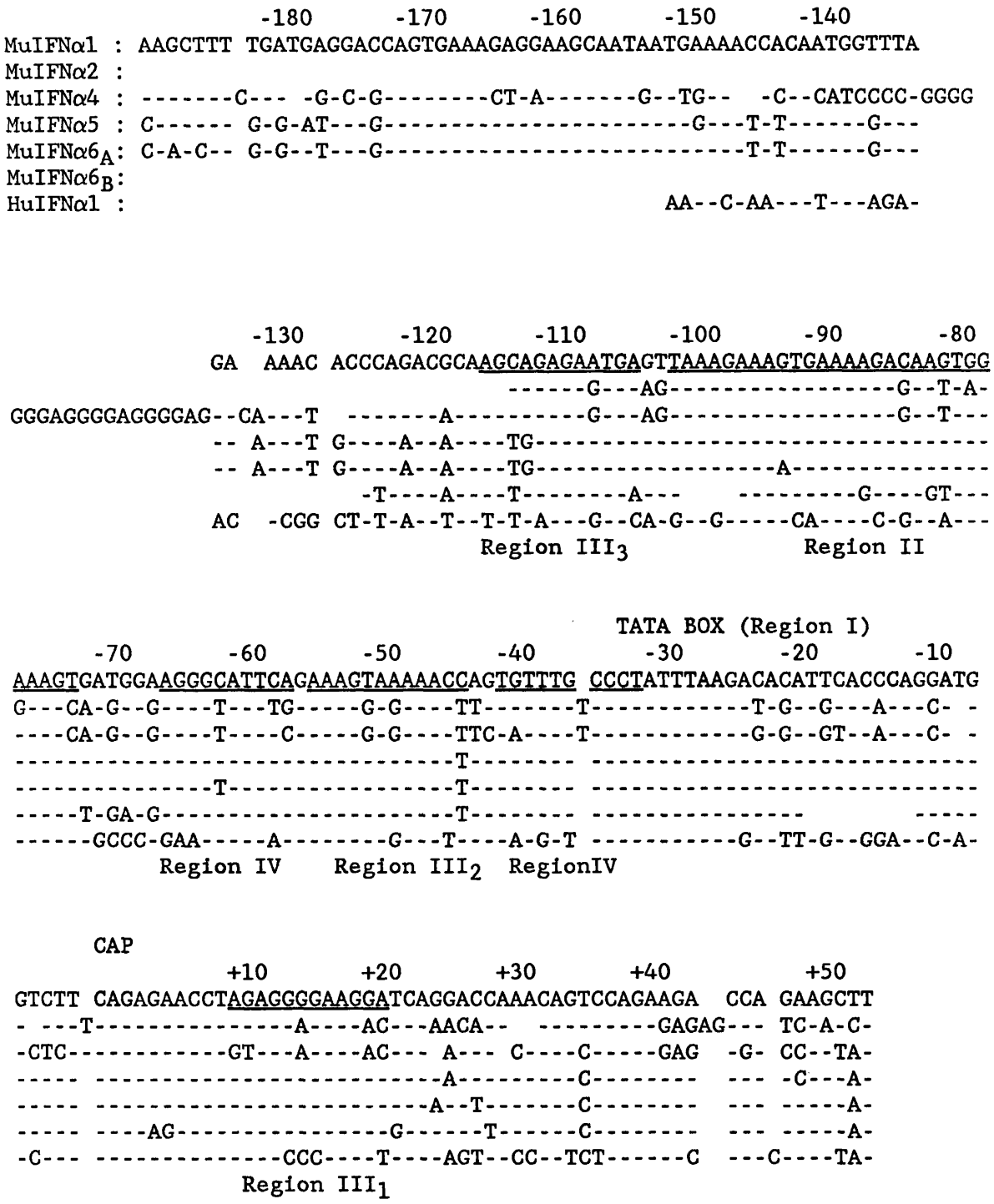

-180 -170 -160 -150 -140
MuIFNα1 : AAGCTTT TGATGAGGACCAGTGAAAGAGGAAGCAATAATGAAAACCACAATGGTTTA
MuIFNα2 :
MuIFNα4 : -------C--- -G-C-G--------CT-A-------G--TG-- -C--CATCCCC-GGGG
MuIFNα5 : C------ G-G-AT---G----------G---T-T------G---
MuIFNα6A: C-A-C-- G-G--T---G----------T-T------G---
MuIFNα6B:
HuIFNα1 : AA--C-AA---T---AGA-
-130 -120 -110 -100 -90 -80
GA AAAC ACCCAGACGCAAGCAGAGAATGAGTTAAAGAAAGTGAAAAGACAAGTGG
------G---AG----------G--T-A-
GGGAGGGGAGGGGAG--CA---T -------A----------G---AG----------G--T---
-- A---T G----A--A----TG----------
-- A---T G----A--A----TG----------A----------
-T----A----T--------A--- ---------G----GT---
AC -CGG CT-T-A--T--T-T-A---G--CA-G--G-----CA----C-G--A---
Region III3 Region II
TATA BOX (Region I)
-70 -60 -50 -40 -30 -20 -10
AAAGTGATGGAAGGGCATTCAGAAAGTAAAAACCAGTGTTTG CCCTATTTAAGACACATTCACCCAGGATG
G---CA-G--G----T---TG-----G-G----TT-------T----------T-G--G---A---C- -
----CA-G--G----T----C-----G-G----TTC-A----T----------G-G--GT--A---C- -
----------T----------T-------- ----------
-----T-GA-G----------T-------- ---------- -----
------GCCC-GAA-----A--------G---T----A-G-T ----------G--TT-G--GGA--C-A-
Region IV Region III2 RegionIV
CAP
+10 +20 +30 +40 +50
GTCTT CAGAGAACCTAGAGGGGAAGGATCAGGACCAAACAGTCCAGAAGA CCA GAAGCTT
- ---T----------A----AC---AACA-- ---------GAGAG--- TC-A-C-
-CTC----------GT---A----AC--- A--- C----C-----GAG -G- CC--TA-
----- ----------A---------C-------- --- -C---A-
----- ----------A--T-------C-------- --- -----A-
----- ----AG----------G------T------C-------- --- -----A-
-C--- ----------CCC----T----AGT--CC--TCT------C ---C----TA-
Region III1

FIGURE 6.3 Primary Nucleotide Sequence of MuIFNα_1 and Mutant Promoters

The primary nucleotide sequence of the MuIFNα_1 241bp *Hin*dIII promoter fragment (-188 to +52) is illustrated. Below are aligned mutant MuIFNα_1 gene promoter sequences derived by site directed mutagenesis, which have been verified by sequencing studies. The specific nucleotide substitutions created are indicated for each promoter sequence. Unchanged nucleotides are denoted by a dash (-). The 77bp deletion of P3Δ77 is indicated by a space.

```
                    -180       -170       -160       -150       -140       -130
MUIFNα1 :    AGCTTTTGATGAGGACCAGTGAAAGAGGAAGCAATAATGAAAACCACAATGGTTTAGAAAACA
P1           ------------------------------------------------------------
P2           ------------------------------------------------------------
P3           ------------------------------------------------------------
P3Δ77        ----------------
P4           ------------------------------------------------------------
P5           ------------------------------------------------------------
P6           ------------------------------------------------------------
P6t          ------------------------------------------------------------
P7           ------------------------------------------------------------
P7[8]        ------------------------------------------------------------
P8           ------------------------------------------------------------

    -120       -110       -100       -90        -80        -70        -60
CCCAGACGCAAGCAGAGAATGAGTTAAAGAAAGTGAAAAGACAAGTGGAAAGTGATGGAAGGGCATTCAGAAAGT
------------------------------------------------------------------------
-----------------------------------GGACAA----------------------------------
-----------------------------------GGAGAA----------------------------------
                                  -------GGAGAA----------------------------------
------------------------------------------------------------------------GG
------------------------------------------------------------------------GG
------------------------------------------------------------------------
------------------------------------------------------------------------
------------------------------------------------------------------------
------------------------------------------------------------GCGG----------
------------------------------------------------------------GCGG----------

                  TATA BOX                          CAP
50        -40        -30        -20        -10                  +10        +20
AAAAACCAGTGTTTGCCCTATTTAAGACACATTCACCCAGGATGGTCTTCAGAGAACCTAGAGGGGAAGGATCA
------------------TATAAAA-----------------------------------------------
------------------------------------------------------------------------
------------------------------------------------------------------------
------------------------------------------------------------------------
AAAA--------------------------------------------------------------------
AGAA--------------------------------------------------------------------
------------------------------------------------------------GGAGAA------
-----------------------------------C------------------GGAGAA------
--------------CCGC----------------------------------------------------
--------------CCGC----------------------------------------------------
------------------------------------------------------------------------

     +30        +40        +50
GGACCAAACAGTCCAGAAGACCAGAAGCT
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
-----------------------------
```

Region II

A second region of substantial inter− and intra− species homology was identified within the IFNα gene promoters, located between −102 and −73. This was designated region II (Fig.6.2). This region contained a 14bp sequence (−94 to −81; Fig.6.4A) with an imperfect repeat (−171 to −158) located within another relatively conserved region at −172 to −147.

The potential functional importance of region II, defined here by its conservation in distinct IFNα genes is reinforced by the fact that deletion mapping studies (Ryals *et al*, 1985; Goodbourn *et al*, 1985) have also focussed attention on corresponding sequences in the HuIFNα_1 and HuIFNβ gene promoters. Figure 6.5 shows the MuIFNα_1 promoter sequence with regions I to IV indicated. Aligned below, at homologous locations, are repeat sequences of the virus inducible HuIFNα_1 and polyrI.rC inducible HuIFNβ gene promoters which have been functionally identified by deletion mutagenesis (Ryals *et al*, 1985; Goodbourn *et al*, 1985). Such comparisons, therefore, reveal that region II corresponds to both the HuIFNα_1 and two 13bp repeat HuIFNβ sequences, strongly suggesting a common function in human and mouse IFN promoter regulation. The MuIFNα_1 gene promoter also possesses an additional motif which is homologous to part of region II, located between −171 to −158, which suggests either other regulatory sequences or some degeneracy within this promoter.

Clearly region II is likely to be implicated both in the mechanism of virus and polyrI.rC induced gene regulation. As described below (region III), a possible relationship was identified between region II and region III sequences. A sequence related to the hexanucleotide consensus sequence (region III; see below) was identified between −87 to −81 which is located within region II. Therefore, the prospective contribution of this sequence to MuIFNα_1 gene promoter function was investigated by creating two mutant promoters, P2 and P3 (Fig.6.3), which contained improved and perfect consensus sequences, respectively, at this location.

FIGURE 6.4 Repeat Sequences of MuIFNα Gene Promoters

Repeat sequences identified within MuIFNα gene promoters are aligned for region II (A), region III (B) and region IV (C) respectively. Their location within the MuIFNα promoters are indicated, numbered relative to the transcription initiation site. For region II (A) and IV (C), sequences from the MuIFNα_1 promoter only are shown. The inverted repeat sequences are aligned in opposite orientations, depicting complementarity of the two regions (C, mismatches are shown as *). Region III (B) sequences are shown for all six MuIFNα genes analysed in this study. Below the three region III sequences is a consensus sequence derived from the most frequently occuring nucleotides at each position. The hexanucleotide sequence GGAGAA is underlined. X= A, G or T.

A) Region II: 14bp repeat (α_1)

	−94	AGTGAAAAGACAAG -87
	-171	AGTGAAAGAGGAAG -158

B) Region III: 12bp repeats

		+9	+20
Region III$_1$	α_1	AGAGGGGAAGGA	
	α_2	AGAGGAGAAGAC	
	α_4	GTAGGAGAAGAC	
	α_5	AGAGGGGAAGGA	
	α_{6A}	AGAGGGGAAGGA	
	α_{6B}	AGAGGGGAAGGA	

		-55	-44
Region III$_2$	α_1	AAAGTAAAAACC	
	α_2	AAAGGAGAAACT	
	α_4	AAAGGAGAAACT	
	α_5	AAAGTAAAAACT	
	α_{6A}	AAAGTAAAAACT	
	α_{6B}	AAAGTAAAAACT	

		-115	-104
Region III$_3$	α_1	AGCAGAGAATGA	
	α_2	CAGAGAGTGA	
	α_4	AGCAGAGAGTGA	
	α_5	AGTGGAGAATGA	
	α_{6A}	AGTGGAGAATGA	
	α_{6B}	AGTAGAGAATGA	

Consensus:		AGAGGAGAAXGA

C)
Region IV: 10bp inverted repeat (α_1)

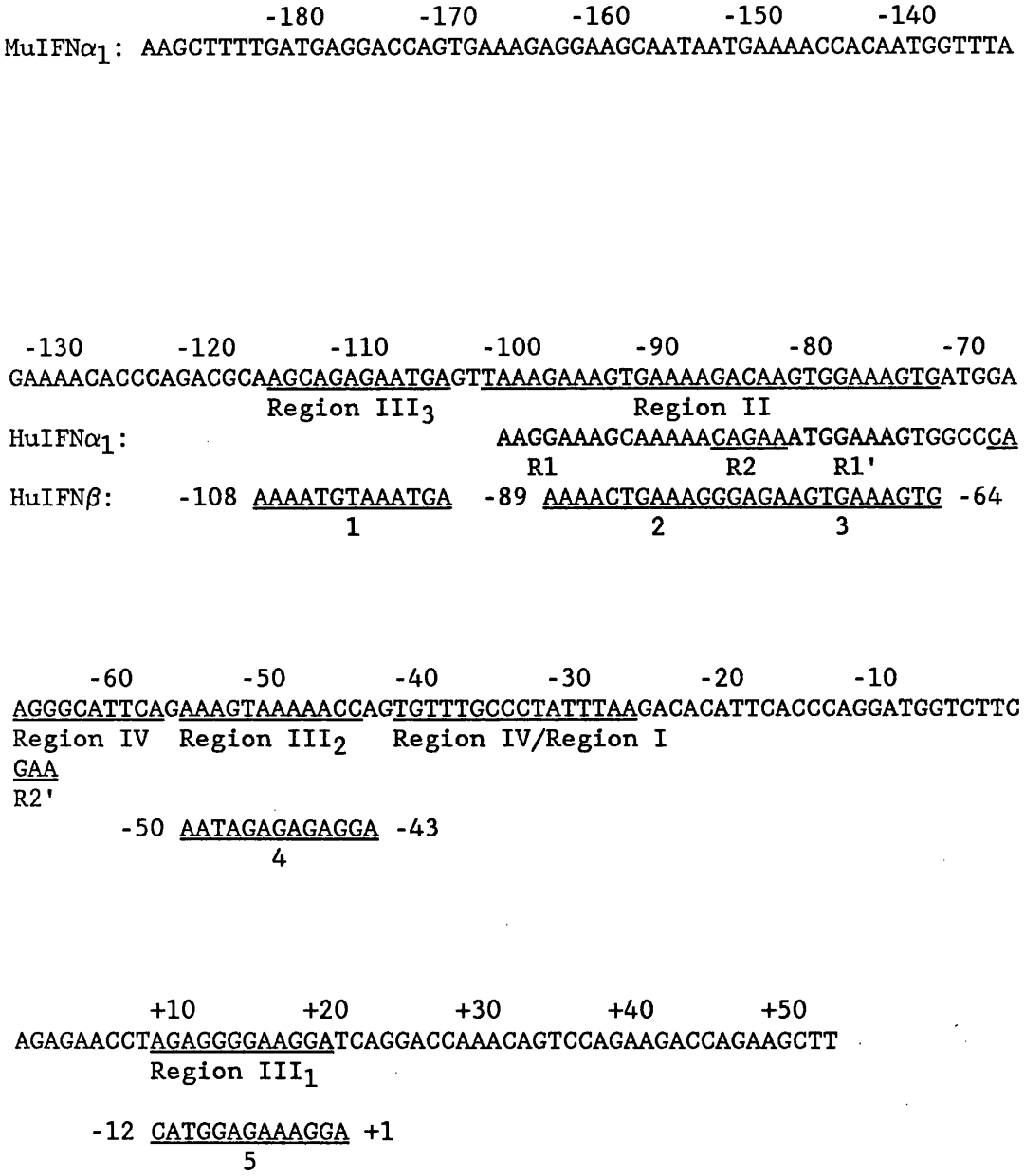
5'	-66	AGGGCATTCA	-57 3'
3'	-32	TCCCGT**GT	-41 5'
		TT	

FIGURE 6.5 Comparison of MuIFNα_1 Sequences with HuIFNα_1 (Ryals *et al* 1985) and HuIFNβ (Goodbourn *et al*, 1985) Repeats

The primary nucleotide sequence of the 241bp *Hin*dIII fragment (-188 to +52) of the MuIFNα_1 gene promoter is shown. The nucleotide sequences of prospective regulatory regions I to IV are underlined. Aligned below the MuIFNα_1 sequence are repeat sequences identified previously by deletion mutagenesis of the HuIFNα_1 and HuIFNβ genes (Ryals *et al*, 1985; Goodbourn *et al*, 1985).

The HuIFNα_1 repeats: R1, R1', R2 and R2'

The HuIFNβ repeats: 1, 2, 3, 4 and 5


-180 -170 -160 -150 -140
MuIFNα1: AAGCTTTTGATGAGGACCAGTGAAAGAGGAAGCAATAATGAAAACCACAATGGTTTA
-130 -120 -110 -100 -90 -80 -70
GAAAACACCCAGACGCAAGCAGAGAATGAGTTAAAGAAAGTGAAAAGACAAGTGGAAAGTGATGGA
Region III3
Region II
HuIFNα1: AAGGAAAGCAAAAACAGAAATGGAAAGTGGCCCA
R1 R2 R1'
HuIFNβ: -108 AAAATGTAAATGA -89 AAAACTGAAAGGGAGAAGTGAAAGTG -64
1 2 3
-60 -50 -40 -30 -20 -10
AGGGCATTCAGAAAGTAAAAACCAGTGTTTGCCCTATTTAAGACACATTCACCCAGGATGGTCTTC
Region IV Region III2 Region IV/Region I
GAA
R2'
-50 AATAGAGAGAGGA -43
4
+10 +20 +30 +40 +50
AGAGAACCTAGAGGGGAAGGATCAGGACCAAACAGTCCAGAAGACCAGAAGCTT
Region III1
-12 CATGGAGAAAGGA +1
5

Region III

Three potential 12bp imperfect repeat sequences (III_1, +9 to +20; III_2, −55 to −44; III_3, −115 to −104), designated region III 1 to 3 respectively, were recognised within conserved regions of the MuIFNα promoters (Fig.6.2). These repeats (Fig.6.4B) contain a consensus hexameric core sequence, GGAGAA, which is present in the MuIFNα_2, α_4 (region III_1 and III_2), α_5 and α_{6A} (region III_3) promoters. The HuIFNβ gene promoter also possesses two such sequences (Hauser *et al*, 1984). The basic region III motifs are also conserved in MuIFNα_1, HuIFNα_1 and other HuIFNα promoters, but the perfect hexameric sequence is absent in these genes. Thus, all MuIFNα gene promoters examined, the endogenous genes of which are efficiently expressed in L−929 cells (Chapter 1), retain one or more copies of the perfect consensus sequence, but the MuIFNα_1 gene has none.

Close inspection of region III sequences indicated that they may be related to the HuIFNβ 13bp repeats (Fig.6.5). Thus, this suggested a relationship between region III and region II, described previously, which also show homology to the HuIFNβ repeats. This raised the possibility that the GGAGAA motif is the consensus core sequence of these repeats.

Direct evidence for the potential functional significance of these sequences in the HuIFNβ gene has been provided by deletion analysis (Zinn *et al*, 1983). As described in Chapter 1, such studies indicated that 77bp of 5' flanking sequence was required for full polyrI.rC induction in C127 cells (Zinn *et al*, 1983). Deletion to position −73 resulted in a large reduction in induced transcription levels, although low level inducibility was retained (Zinn *et al*, 1983). This deletion corresponded to the loss of a copy of the hexameric sequence GGAGAA (observation from this analysis). Thus, this suggested that region III sequences may be functionally relevant, possibly contributing to the efficiency of transcription initiation of IFN gene promoters.

Human and mouse fibroblasts induced by polyrI.rC produce mainly IFNβ, or a mixture of IFNα and IFNβ respectively (Chapter 1). The different cell specificity observed with HuIFNα and HuIFNβ gene expression may reflect differences in the structure of the gene promoters. In this respect it is interesting that the MuIFNα

promoters appear to share structural features common to HuIFNα (region II) and HuIFNβ (region III) promoters. Endogenous expression of virus induced MuIFNα_2, α_4, α_5 and α_{6A} genes has been demonstrated in L−929 cells, but not MuIFNα_1 (Chapters 1 and 5). Therefore, the absence of a perfect hexameric repeat in the MuIFNα_1 gene promoter may contribute to the undetectable endogenous gene expression observed in NDV induced L−929 fibroblast cells.

To investigate the potential role of GGAGAA sequences in polyrI.rC induced MuIFNα_1 promoter function in L−929 cells, the promoters P2, P3, P4, P5 and P6 (Fig.6.3) were constructed by site directed mutagenesis. The promoter mutants, P6 and P4/P5 created improved or perfect hexameric repeat sequences corresponding to regions III_1 and III_2, respectively. In addition, a third prospective region III core sequence was identified within region II. The perfect consensus sequence has not been identified in this region in any native MuIFNα promoters examined. Therefore, the promoters P2 and P3, which contain this sequence, were constructed.

Region IV

Computer analysis was undertaken to further extend characterisation of the MuIFNα_1 gene promoter structure. Analysis of the promoter sequence, employing Analyseq (Staden, 1977), revealed an imperfect 10bp inverted repeat (Fig.6.4C) located at −41 to −32 and −66 to −57 which has been designated region IV (Fig.6.2). Comparison of MuIFNα and HuIFNα_1 promoters shows that sequences both within and between the inverted repeats are conserved.

To investigate the possible relevance of these sequences in polyrI.rC regulated activation of the MuIFNα_1 gene promoter, mutations were designed to examine the effect of disrupting and recreating homology between these sites. Thus, the mutant promoters P7, $P7^8$, and P8 (Fig.6.3) were constructed by site directed mutagenesis.

The identification of significantly conserved mouse and human IFNα gene promoter elements, although strongly suggesting a major functional importance, provides no direct proof of any precise role in promoter activity. However, the comparison of the regions identified by deletion analysis as being functionally important to HuIFNα_1 and HuIFNβ

gene regulation, with MuIFNα promoter organisation, reinforces the conclusions drawn from inter− and intra−specific sequence comparisons that regions I, II, III and IV may be of major importance. Thus, as a preliminary step to assess the possible relevance of these sequences in polyrI.rC mediated induction of MuIFNα_1 gene promoter activity, the mutants illustrated in Figure 6.3 were created.

6.2 Strategy for Site Directed Mutagenesis of the MuIFNα_1 Promoter

To generate the promoter mutants described in Section 6.1, oligonucleotide mediated site directed mutagenesis (Razin *et al*, 1978) of the MuIFNα_1 promoter was undertaken. The procedure adopted (Zoller and Smith, 1983), shown schematically in Figure 6.6, utilised the M13 phage derivative mPWT (Chapter 5) which contains the MuIFNα_1 gene promoter module. Synthetic, mismatched oligonucleotides directed to specific targets within the promoter fragment were hybridised to the ssDNA template to direct *in vitro* synthesis of the complementary DNA strand, generating a heteroduplex (Fig.6.6). Propagation *in vivo* of the *in vitro* synthesised DNA heteroduplex produced wildtype and mutant phage derivatives which were distinguishable by differential hybridisation, employing the oligonucleotide which had been used to generate the mutation as probe.

Design of Oligonucleotides for Mutagenesis

Oligonucleotides were designed and synthesised for use as mismatched primers to construct the MuIFNα_1 promoter mutants. Computer assisted analysis, using the Seqfit programme (Staden, 1977), was employed to help design oligonucleotides which preferentially hybridised to the predetermined target sites, with minimum complementarity elsewhere in mPWT template DNA.

The minimum size of the oligonucleotides to readily achieve the best target specificity was 19 nucleotides. In each case no more than thirteen of nineteen bases were complementary, other than the target site, to secondary sites within the M13 phage mPWT genome. The primary nucleotide sequence of the oligonucleotides used for site directed mutagenesis are shown in Table 6.1. All oligonucleotides for this type of study were kindly synthesised by Dr. M.D. Edge.

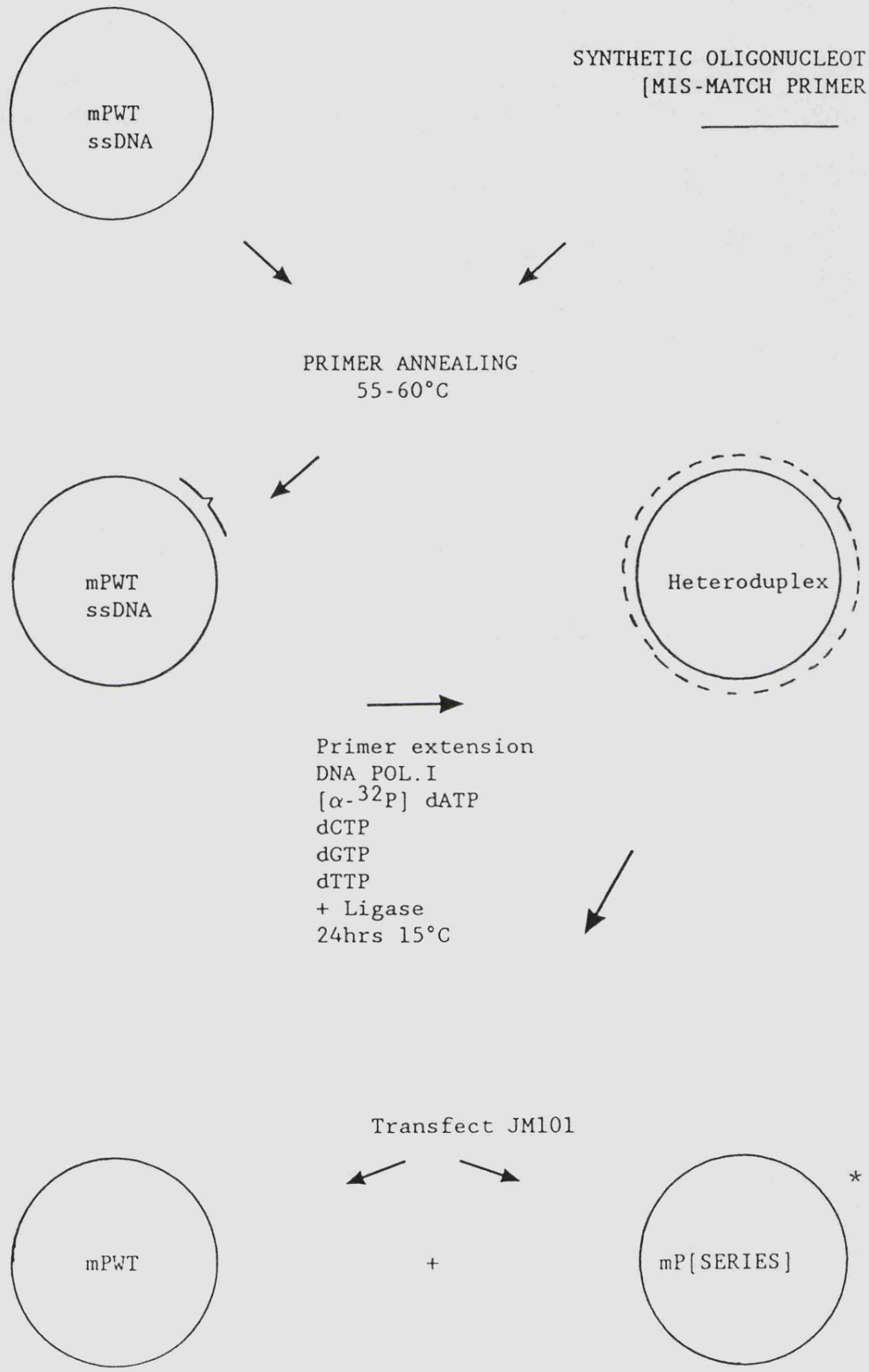


FIGURE 6.6 Generation of Mutant MuIFNα_1 Promoters by Oligonucleotide Site Directed Mutagenesis

* new mutation

Target Specificity of Oligonucleotides

Prior to being employed for site directed mutagenesis, the quality and specificity of each oligonucleotide was experimentally evaluated by determining both the efficiency of 5' phosphorylation with T4 polynucleotide kinase and the site of initiation in sequencing reactions (Chapter 2) performed with mPWT template DNA.

The efficiency of 5' phosphorylation in addition to indicating the quality of an oligonucleotide preparation, is an important parameter with respect to site directed mutagenesis, since it is necessary both for DNA ligation, and hence heteroduplex formation, and for oligonucleotide hybridisation. Oligonucleotides 1 to 8 were all efficiently phosphorylated by T4 polynucleotide kinase ($2-5\times10^6$cpm/μg DNA). The products of these reactions were also analysed by denaturing P.A.G.E., confirming in every case the presence of major species of the expected size (19 nucleotides).

To demonstrate that the oligonucleotides were priming at the target site within the promoter of mPWT DNA, they were each used as primers for dideoxy sequencing (Chapter 2). Figure 6.7 shows the results achieved with four oligonucleotides. Comparison with the sequence of the MuIFNα_1 promoter demonstrated that they initiated DNA synthesis as expected. However oligonucleotides 2, 4, 6 and 7 gave poor sequencing data and thus, it was not possible to determine if they primed correctly. Nevertheless, subsequent studies show these oligonucleotides could still be successfully employed to create the desired mutant promoters, although generally with reduced efficiency.

Oligonucleotide Site Directed Mutagenesis

The oligonucleotides were then used to create the mutant promoters shown in Figure 6.3 (Chapter 2). To increase the percentage of mutants obtained an enrichment for *in vitro* synthesised covalently closed circular (CCC) heteroduplex DNA was undertaken by including an alkaline sucrose gradient centrifugation step (Kudo *et al*, 1981). Fractions collected from the sucrose gradients were analysed by Cerenkov counting to identify those containing CCC heteroduplex DNA. A typical profile of the DNA fractions recovered is shown in Figure 6.8.

Heteroduplex DNA recovered by this procedure was transfected into *E.coli* JM101.

TABLE 6.1 Structure of Oligonucleotides for Site Directed Mutagenesis

Oligonucleotide	Oligonucleotide Structure 5' 3'	Mutant Promoter Created
	-19 ** -38	
1	ATGTGTCTTTTATAGGGCA	P1 (region I)
	-77 * -95	
2	TTCCACTTGTCCTTTCACT	P2 (region II)
	-77 * * -95	
3	TTCCACTTCTCCTTTCACT	P3 (region II)
	-42 * -60	
4	ACTGGTTTTTCCTTTCTGA	P4 (region III_2)
	-42 * * -60	
5	ACTGGTTTCTCCTTTCTGA	P5 (region III_2)
	+24 * +6	
6	CTGATCCTTCTCCTCTAGG	P6 (region III_1)
	-26 * * -44	
7	CTTAAATAGCGGAAACACT	P7 $+P7^8$ (region IV)
	-54 * * -72	
8	TTTCTGAATCCGCTTCCAT	P8 $+P7^8$ (region IV)

* shows site of mutation

FIGURE 6.7 Sequencing with Mismatched Oligonucleotide Primers

Sequencing reactions were performed using oligonucleotide primers and mPWT single stranded template DNA using standard conditions described in Chapter 2. The products were analysed by denaturing P.A.G.E on an 8% sequencing gel.

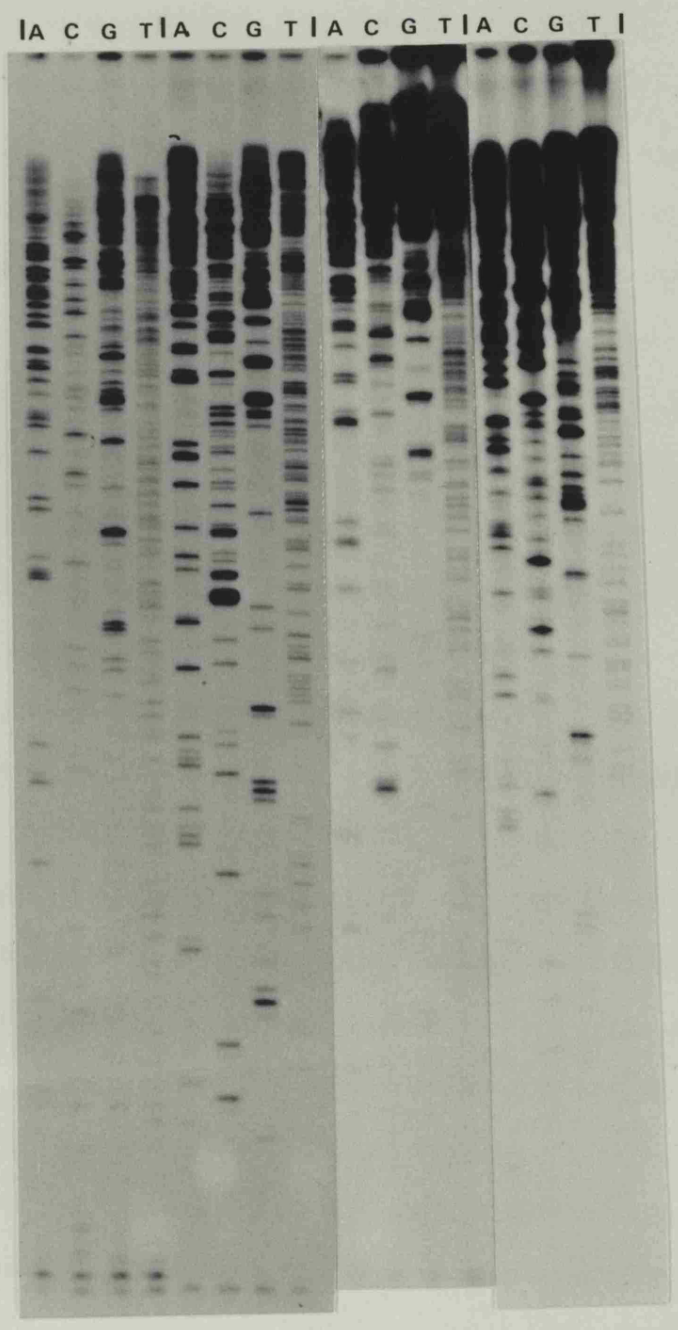
1) Oligonucleotide 5 (Table 6.1)

2) " 3

3) " 1

4) " 8

1 2 3 4



The frequency of transfection varied according to the oligonucleotide used, being particularly low with oligonucleotides 7 and 8 (Table 6.2).

To identify prospective mutants, one hundred phage obtained from each mutagenesis were screened by colony hybridisation (Chapter 2) with the corresponding 5' phosphorylated oligonucleotide as probe ($1-2\times10^6$cpm/μg), under low stringency conditions (6×SSC, 20°C). The anticipated melting temperature for each oligonucleotide probe, calculated as described in Chapter 2, was approached by sequentially increasing the hybridisation temperature by 5°, 10° or 15°C steps, allowing the distinction of lysogens carrying mutant and wildtype phage DNA by differential hybridisation. An example of the results obtained using this procedure is shown in Figure 6.9. Prospective mutants were identified at a frequency between 2% and 15%, depending upon the oligonucleotide used (Table 6.2). The highest mutation frequency was obtained with oligonucleotides 1 (15%) and 5 (12%), both previously identified as good sequencing primers (Fig.6.7).

Sequencing studies were undertaken to prove that the desired mutants had been isolated. For this purpose the universal and oligonucleotide 1 primers were separately employed to sequence the entire promoter region. Figures 6.10, 6.11 and 6.12 show typical results of sequencing studies, performed to compare the primary nucleotide sequence of the mutant and wildtype promoters. In some cases several prospective mutant phage DNAs were sequenced to obtain the desired mutants (Table 6.2).

Surprisingly, in addition to the expected promoter mutants (Fig.6.3), two others were detected. Site directed mutagenesis with oligonucleotide 3 generated an identical 77bp deletion, in four of five prospective mutants examined, in addition to the desired mutation (Fig.6.3). A likely explanation for the origin of the 77bp deletion is depicted in Figure 6.13, suggesting that the oligonucleotide hybridised simultaneously to two distinct regions of the MuIFNα_1 gene promoter. This likelihood highlights the observed homology between the nucleotides at −94 to −81 and −171 to −158, discussed previously in Section 6.1 (region II). An additional mutation was also obtained using oligonucleotide 6, where both the desired mutation and a spontaneous G to C transition (−8) was identified in one prospective mutant analysed (Fig.6.3).

Thus, the primary nucleotide sequence of the intended mutant promoters and the

FIGURE 6.8 Histogram Showing Enrichment for CCC DNA After Sucrose Gradient Centrifugation

Centrifugation was performed as described in Chapter 2. Fractions were analysed by Cerenkov counting to identify those containing CCC DNA (lower half of gradient).

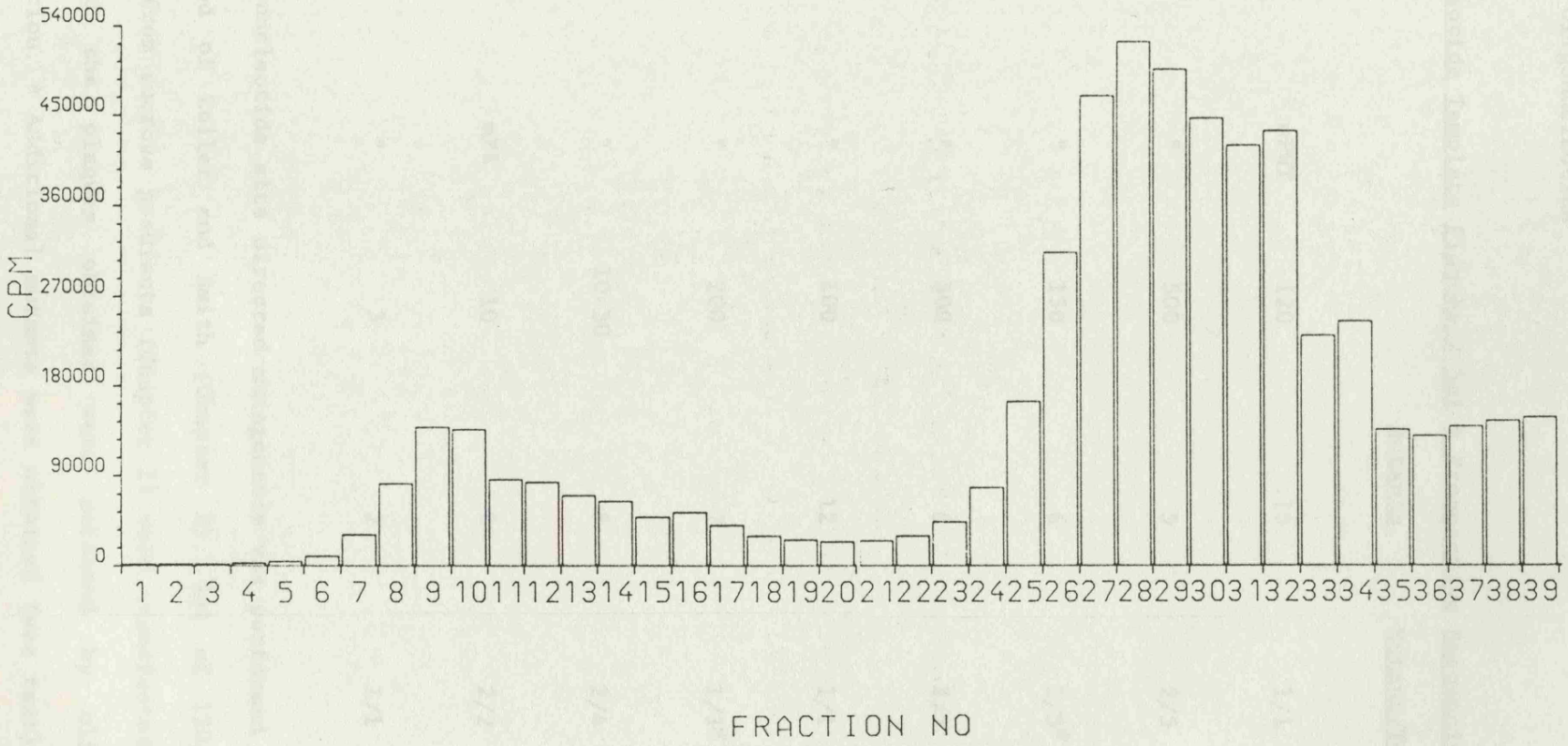

540000
450000
360000
270000
180000
90000
0
CPM
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
FRACTION NO

TABLE 6.2 Efficiency of Site Directed Mutagenesis of the MuIFNα_1 Promoter with Each Oligonucleotide

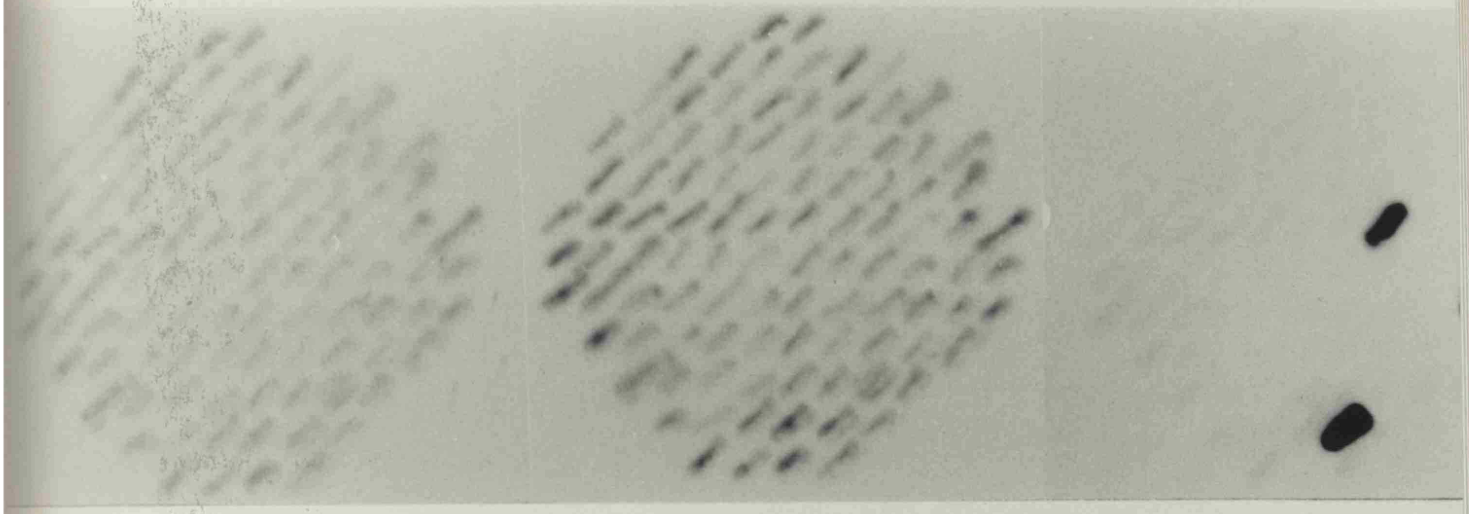
Oligonucleotide	Template	Plaques / 5μl	% Prospective Mutants	Sequencing Mutant/Total
1	mPWT	120	15	1/1
2	"	500	5	2/5
3	"	130	6	1/5*
4	"	500	8	3/5
5	"	100	12	1/1
6	"	200	7	1/3*
7†	"	10-30	5	2/4
7	mP8	10	2	2/2
8	"	3	2	1/1

Oligonucleotide site directed mutagenesis was performed according to the method of Zoller and Smith (Chapter 2). 5μl of 120μl fractions obtained from sucrose gradients (Chapter 2) were transfected into *E.coli* JM101 and the plaques obtained were screened by oligonucleotide hybridisation. * Additional mutants were obtained (see text). † Repeated three times to obtain desired mutant.

FIGURE 6.9 Identification of Prospective Mutant Phage by Colony Hybridisation

100 plaques derived from site directed mutagenesis with oligonucleotide 8 (Chapter 2), used to create the P8 promoter (Fig.6.3), were screened by colony hybridisation using this oligonucleotide labelled at the 5' end as probe (2×10^6 cpm). A) Shows autoradiographs of filters after hybridisation, washed sequentially in 6×SSC at room temperature (rt), 37°C and 50°C. The exposures are for 1hr, 2 hrs and 12hrs respectively, at -70°C using Fuji RX film with an intensifying screen. B) Shows the primary nucleotide structure of oligonucleotide 8 and the anticipated complementarity with wildtype and mutant MuIFNα_1 gene promoter sequences at the target site, indicating two mismatched bases with the former but a perfect match with the latter.

A



rt 37°C 50°C

B

```
                    G   G
WT   5' ATGGAA  G   C  CATTCAGAAA 3'
OI8  3' TACCTTG CGGTAAGTCTTT 5'
P8   5' ATGGAACGCCATTCAGAAA 3'
```

FIGURE 6.10 Sequence Confirmation of the Primary Nucleotide Structure of the MuIFNα_1 Mutant Promoter P1

Sequencing reactions were performed with the universal primer (Amersham 17bp) and the wildtype mPWT or P1 mutant derivative single stranded template DNAs, using the standard procedure described in Chapter 2. The products were analysed in parallel by denaturing gel electrophoresis on an 8% sequencing gel. Autoradiography was performed at -70°C o/n with Fuji RX film. A region of the sequencing ladders indicating two anticipated base substitutions (arrows) in the P1 promoter sequence relative to wildtype (WT) is shown. Inspection of the remainder of the sequence indicated they were identical. The complement of part of the sequence (the 19 nucleotides corresponding to oligonucleotide 1) derived from this gel is shown at the side, indicating two nucleotide changes in the mutant promoter relative to the wildtype.

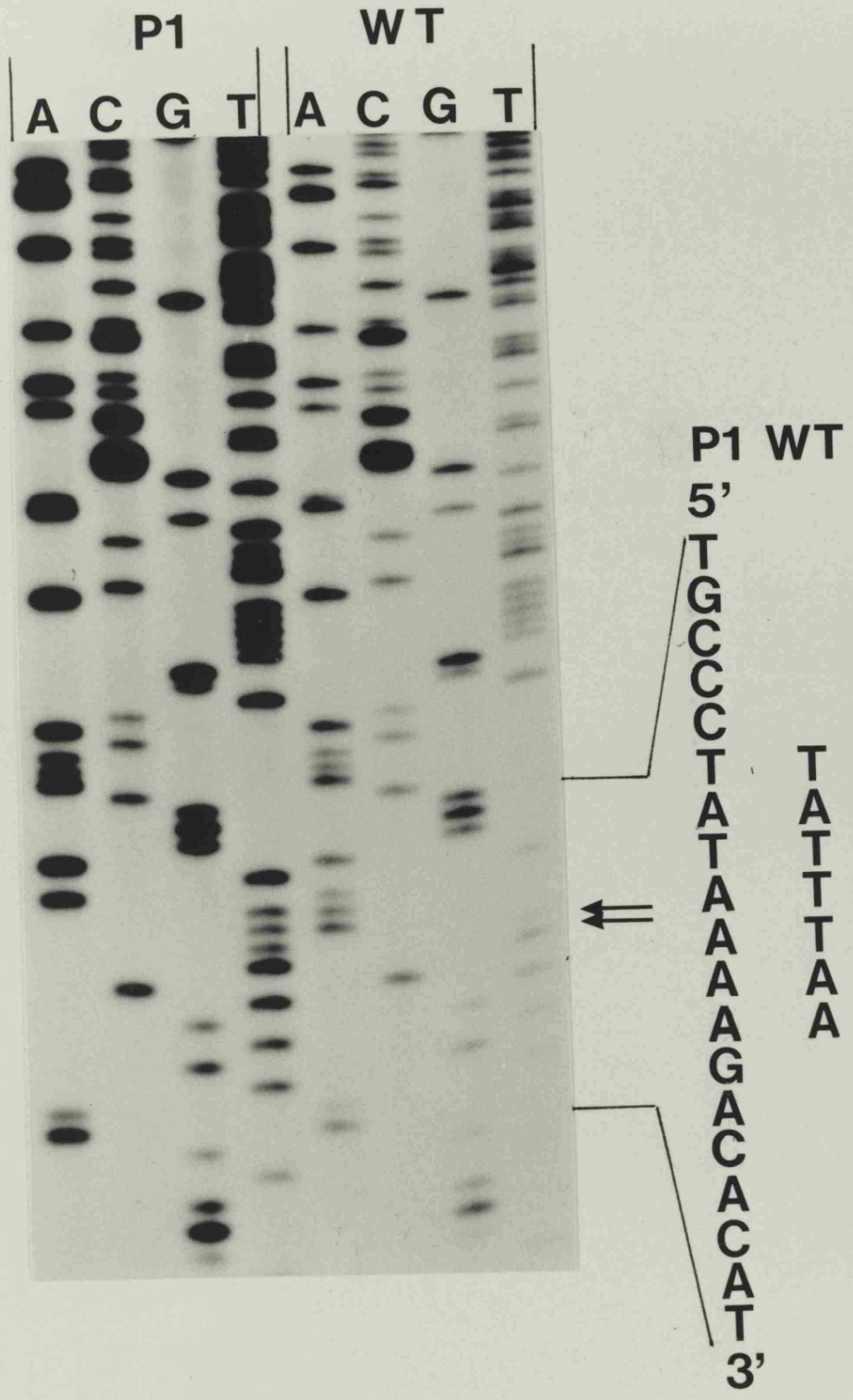

P1
WT
A C G T
A C G T
P1 WT
5' TGCCCTATAAAGACACAT 3'
TATTTAA

FIGURE 6.11 Sequence Confirmation of the Primary Nucleotide Structure of the MuIFNα_1 Mutant Promoter P5

See legend to Fig.6.10

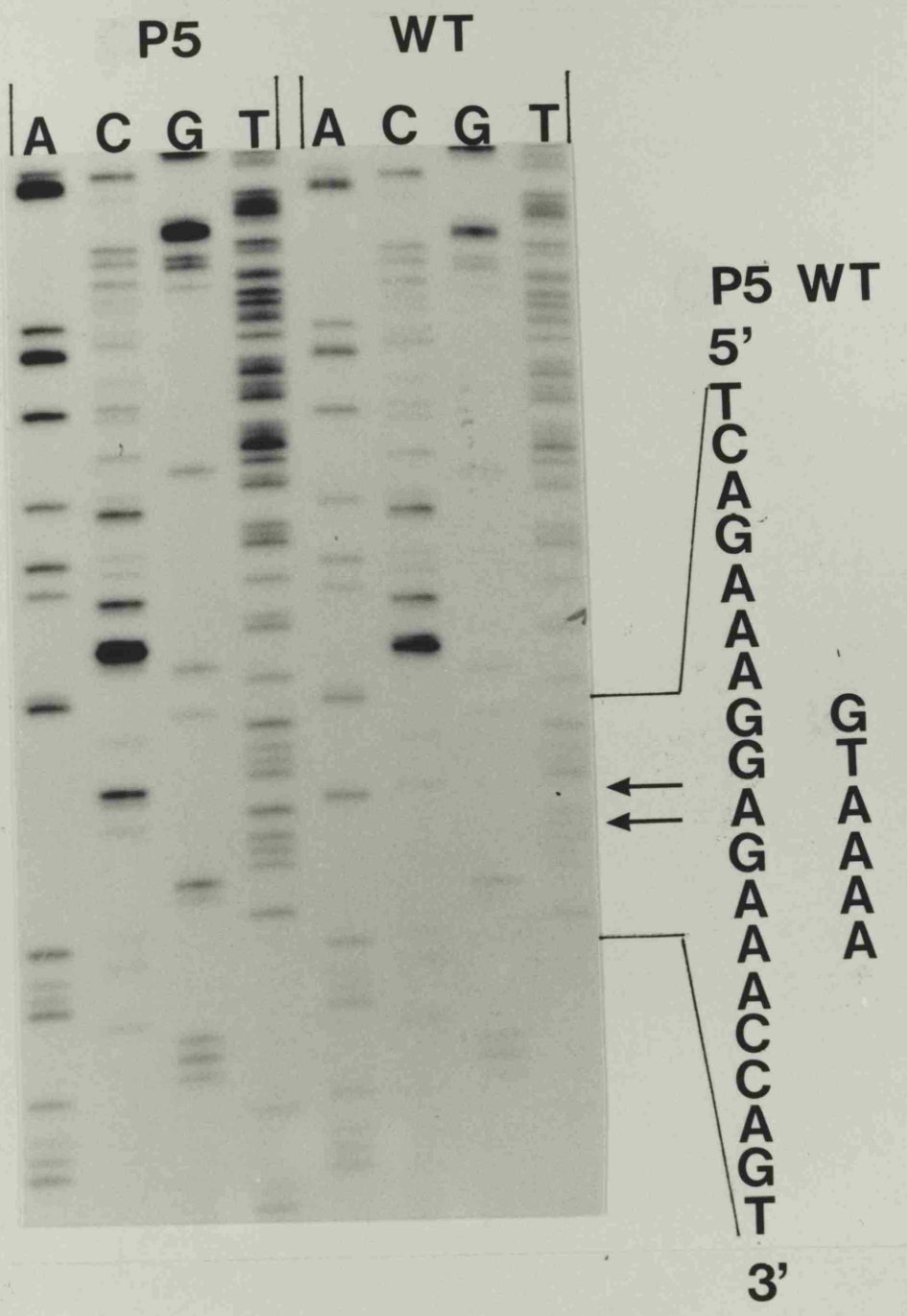
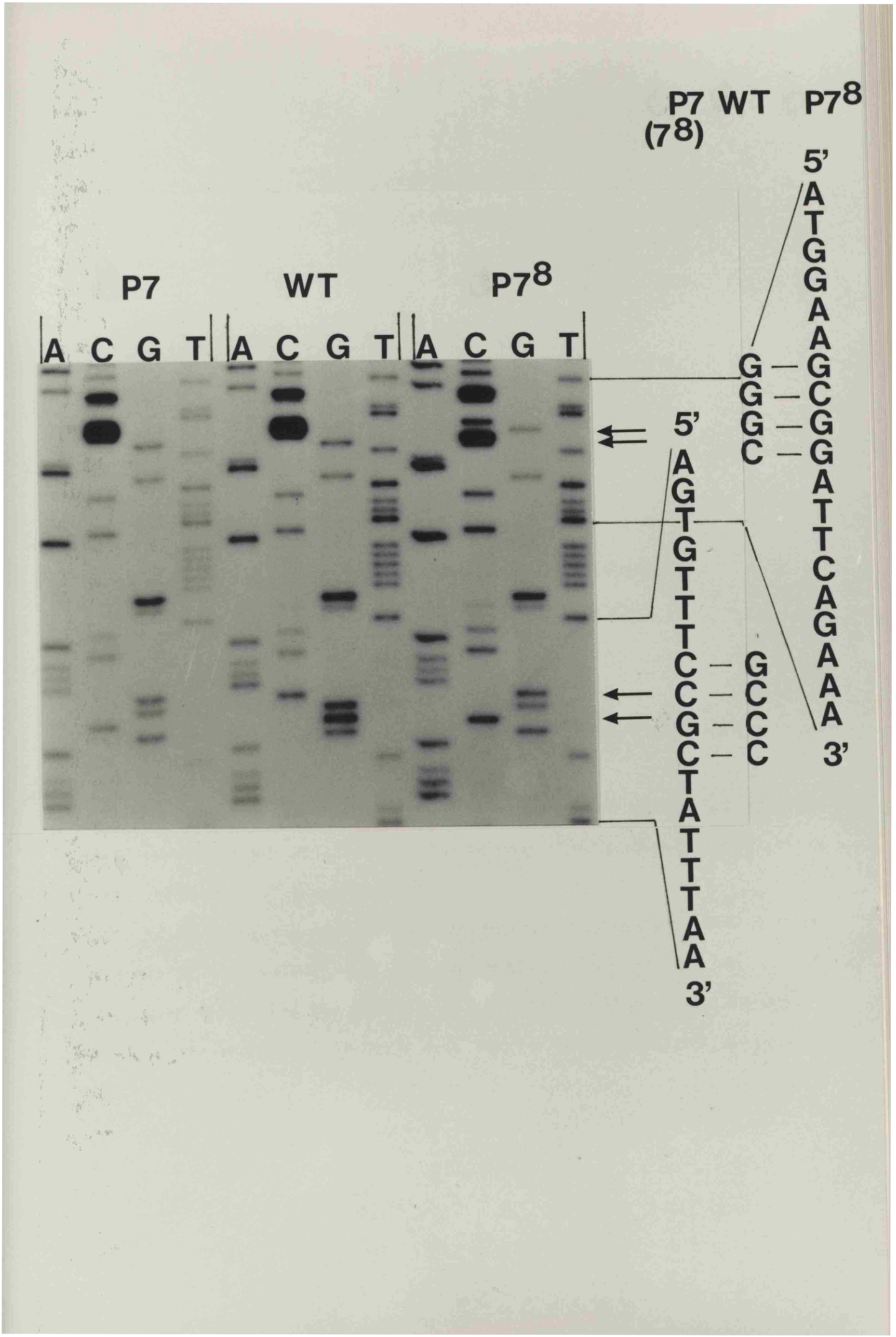

P5
WT
A C G T A C G T
P5 WT
5'
TCAGAAAGGAGAAACCAGT
GTAAAA
3'

FIGURE 6.12 Sequence Confirmation of the Primary Nucleotide Structure of the MuIFNα_1 Mutant Promoters P7 and P7^8

See legend to Fig.6.10


P7 WT P7⁸
(7⁸)
5'
ATGGAAGCGGATTCAGAAA
3'
P7
WT
P7⁸
A C G T
A C G T
A C G T
GGGC
5'
AGTGTTTCCGCTATTTAA
3'
GCCC

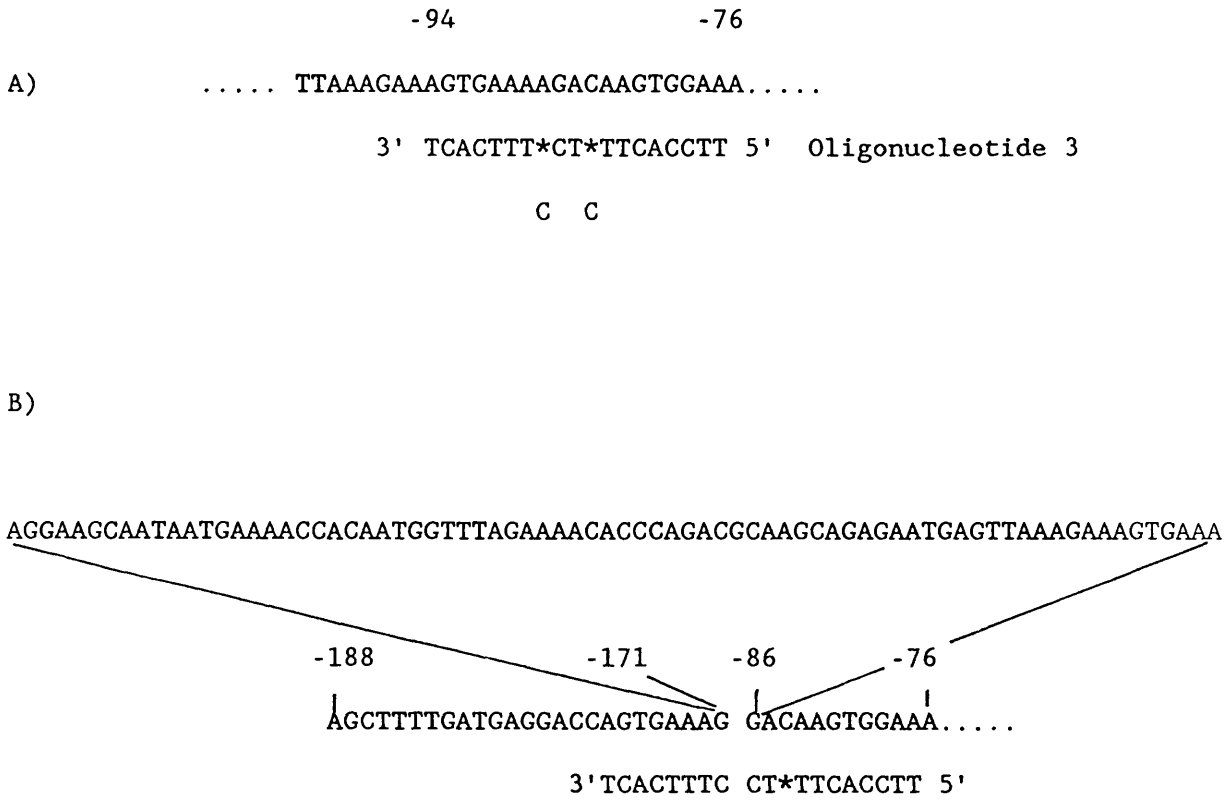


FIGURE 6.13 Prospective Mechanism for Generating the 77bp Deletion Derivative of the MuIFNα_1 Gene Promoter

A) Shows the base pairing of oligonucleotide 3 (Table 6.1) to the anticipated target site within the MuIFNα_1 gene promoter, from -94 to -76, indicating two mismatches (*).

B) Shows an alternative base pairing of oligonucleotide 3 with two regions of the MuIFNα_1 gene promoter from -171 to -164 and -86 to -76, resulting in looping out of 77bp of DNA sequence which would be deleted upon primer extension. In this case only a single mismatch is expected (*).

two additional ones were verified (Fig.6.3). Preliminary studies performed to assess the quality of the oligonucleotides were good indicators of their performance in mutagenesis. In general, those oligonucleotides which were good sequencing primers were most efficient for mutagenesis, whereas poor primers were either inefficient, or generated a high frequency of false mutants.

6.3 Construction of MuIFNα_1 Promoter Mutant/*cat* Gene Hybrids

It was intended to analyse the regulation of MuIFNα_1 mutant promoter activity using the CAT expression assay described in Chapter 5. To this end, each of the mutant promoters was excised from the recombinant M13 phage genomes in which they had been generated, as *Sst*I/*Bam*HI fragments, and inserted into the *Sst*I/*B*amHI site of the promoterless *cat* gene plasmid p22. This location was shown previously (Chapter 5) to be suitable for monitoring wildtype promoter activity.

To confirm the successful cloning of the promoter modules, plasmid DNA was prepared from ampicillin resistant colonies obtained after transformation of *E.coli* JA221 with the ligated DNAs, digested with *Sst*I/*Bam*HI and analysed by P.A.G.E. The number of colonies screened and those recognised in each case to be containing the MuIFNα_1 mutant promoters are summarised in Table 6.3. The expected restriction map for each recombinant plasmid was confirmed by *Bst*NI/ *Eco*RI digestion together with P.A.G.E. (Chapter 5; Fig.5.3B).

6.4 Expression of MuIFNα_1 Promoter Mutants in L929 Cells

To assess *in vivo* transcription from the mutant promoters, the plasmids pCP1 to pCP8 were introduced into L−929 cells, employing the stable transfection procedure described in Chapter 5 (Section 5.2).

CAT Expression from MuIFNα_1 Mutant Promoters in Stably Transfected L929 Cells

To investigate and compare non−induced and polyrI.rC mediated CAT expression from the MuIFNα_1 mutant and wildtype promoters, three transfections for each plasmid were performed to generate independent cell populations (Chapter 5) of 50 to 200 pooled

TABLE 6.3 Cloning MuIFNα_1 Promoter Mutants into p22

Promoter Module	Colonies Screened		Expression Vector Created
	Total	+Promoter	
P1	8	7	pCP1
P2	2	1	pCP2
P3	4	3	pCP3
P3Δ77	8	7	pCP3Δ77
P4	4	2	pCP4
P5	8	7	pCP5
P6	3	1	pCP6
P6t	8	6	pCP6t
P7	4	2	pCP7
$P7^8$	4	2	$pCP7^8$
P8	12	7	pCP8

G418 (400μg/ml) resistant colonies (Table 6.4) for further analysis.

Preliminary characterisation of each of the cell populations was undertaken to prove the presence of MuIFNα_1/*cat* gene sequences in the L−929 cell genome, as described previously (Chapter 5). Southern blot analysis of DNA from cells derived from transfection 1 (Table 6.4) confirmed that the cell populations contained unrearranged copies of the 1163bp (1093bp for L929pCP3$\Delta77_1$ cells) *Sst*I/*Hin*dIII fragment, corresponding to the MuIFNα_1/*cat* gene (Fig.6.14, Lanes 1 to 8). The L−929 *Sst*I/*Hin*dIII 1050bp fragment described in Chapter 5 was again observed here (Fig.6.14, lane 9).

Having established that L−929 cells had been stably co−transfected with the MuIFNα_1/*cat* gene mutants, polyrI.rC regulated CAT production was examined. CAT assays performed on extracts of cells derived from transfection 1, revealed readily detectable CAT production in most induced cell populations (Fig.6.15 Lanes b,d,h,l,n,p and r) except those containing the mutant promoters P3Δ77 or P6t (Fig.6.13 Lanes f and j), which possess a 77bp deletion and single G to C transition (−8) respectively. In contrast only low levels of CAT production were observed in non−induced cells (Fig.6.15 Lanes a,c,e,g,i,k,m,o and q).

Thus, the control of the MuIFNα_1 gene promoter is maintained with most of these mutations. In non−induced cells, inefficient transcription initiation produced low levels of CAT. PolyrI.rC mediated the activation of efficient transcription initiation, producing high levels of readily detectable CAT activity. No induced CAT production was observed in any of the L929pCP3Δ77 or L929pCP6t cell populations examined, suggesting the 77bp deletion and single base transition (−8) might both result in a reduction of either the transcription efficiency or polyrI.rC activation of the MuIFNα_1 promoter. Therefore, 94bp of 5' flanking sequence are insufficient for polyrI.rC mediated activation of the MuIFNα_1 gene promoter. However, it remained possible that the other delibrately introduced base substitutions in these promoters also contributed to the observed reduction of activity (Fig.6.3).

A deletion of the HuIFNα_1 (−93) gene closely corresponding to that described for the MuIFNα_1 gene promoter above, maintains only 3 fold induction by virus in mouse

TABLE 6.4 Number of G418 Resistant Colonies Obtained by Co-transfection of Promoter Mutant *cat* Constructs with pTCF in L-929 Cells

Plasmid (10μg)	pTCF (1μg)	Cell Population	Colonies/Dish
1) pCP1	+	L929pCP1$_1$	210
2) "	+	L929pCP1$_1$	65
3) "	+	L929pCP1$_3$	50
1) N/D			
2) "	+	L929pCP2$_1$	100
3) "	+	L929pCP2$_2$	130
1) pCP3	+	L929pCP3$_1$	180
2) N/D			
3) N/D			
1) pCP3Δ77	+	L929pCP3Δ77$_1$	200
2) "	+	L929pCP3Δ77$_2$	120
3) "	+	L929pCP3Δ77$_3$	100
1) N/D			
2) "	+	L929pCP4$_1$	100
3) "	+	L929pCP4$_2$	127
1) pCP5	+	L929pCP5$_1$	220
2) "	+	L929pCP5$_2$	84
3) "	+	L929pCP5$_3$	80
1) pCP6	+	L929pCP6$_1$	224
2) "	+	L929pCP6$_2$	100
3) "	+	L929pCP6$_3$	80
1) pCP6t	+	L929pCP6t$_1$	224
2) "	+	L929pCP6t$_2$	200
3 "	+	L929pCP6t$_3$	60
1) pCP7	+	L929pCP7$_1$	142
2) "	+	L929pCP7$_2$	87
3) "	+	L929pCP7$_3$	100
1) pCP7^8	+	L929pCP7$^8_1$	98
2) "	+	L929pCP7$^8_2$	90
3) "	+	L929pCP7$^8_3$	90
1) pCP8	+	L929pCP8$_1$	84
2) "	+	L929pCP8$_2$	60
3) "	+	L929pCP8$_3$	100
1) SS DNA	−	−	0
2) "	−	−	0
3) "	−	−	0

1) Transfection 1

2) Transfection 2

3) Transfection 3

FIGURE 6.14 Southern Blot Analysis of Transfected L-929 Cells

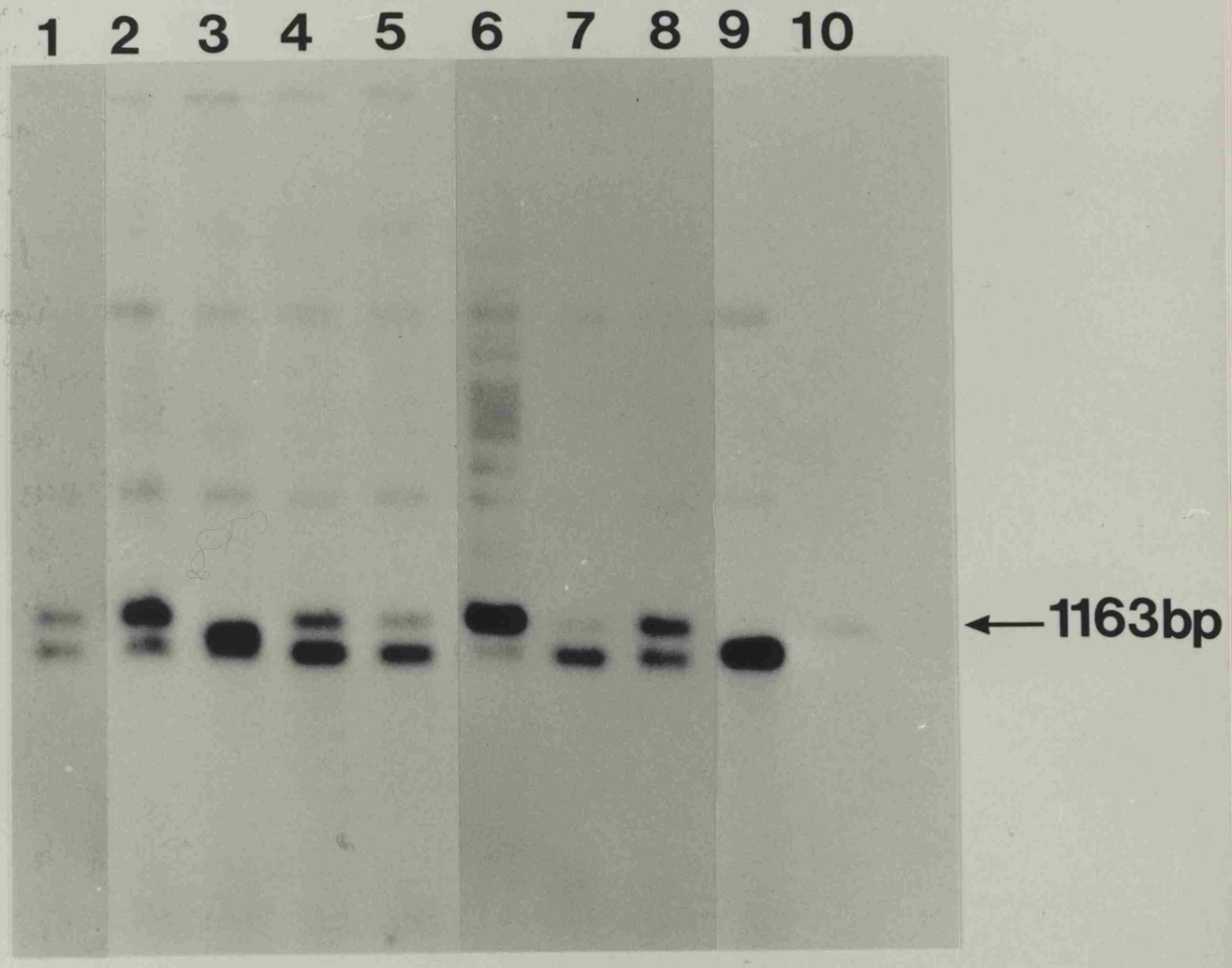
See legend to Fig.5.8. Autoradiography was performed for 12hrs at -70°C using Fuji RX film with an intensifying screen.

A)

1) 10μg L929pCP1$_1$ *Sst*I/*Hin*dIII Digested DNA
2) " L929pCP3$_1$ " " "
3) " L929pCP3Δ77$_1$ " " "
4) " L929pCP6t$_1$ " " "
5) " L929pCP6$_1$ " " "
6) " L929pCP7$_1$ " " "
7) " L929pCP7$^8_1$ " " "
8) " L929pCP8$_1$ " " "
9) " L929 " " "
10) 1pg pCPWT Marker " " "

B) Probe

A



B

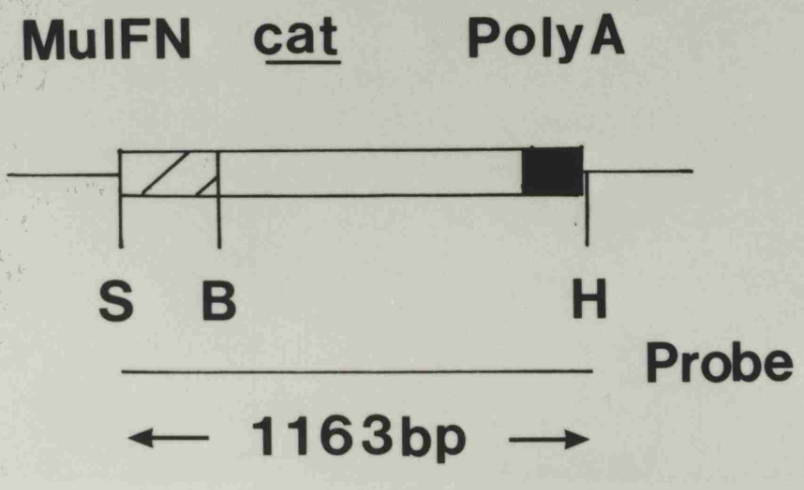


FIGURE 6.15 PolyrI.rC Induced CAT Production in Stably Transfected L-929 Cells

See legend to Fig.5.10. Non-induced cells (-) were incubated in DMEM +2% F.C.S (maintenance medium). Induced cells (+) were treated with 20μg/ml polyrI.rC and 800μg/ml DEAE dextran for 8hrs, washed and incubated for an additional 16hrs in maintenance medium. Autoradiography was performed for 12hrs with Fuji RX film.

A)

a) L929pCP1_1 cells -

b) " " +

c) L929pCP3_1 cells -

d) " " +

e) L929pCP3$\Delta 77_1$" -

f) " " +

g) L929pCP4_1 " -

h) " " +

i) L929pCP$6t_1$ " -

j) " " +

k) L929pCP6_1 " -

l) " " +

m) L929pCP7_1 " -

n) " " +

o) L929pCP7^8_1 " -

p) " " +

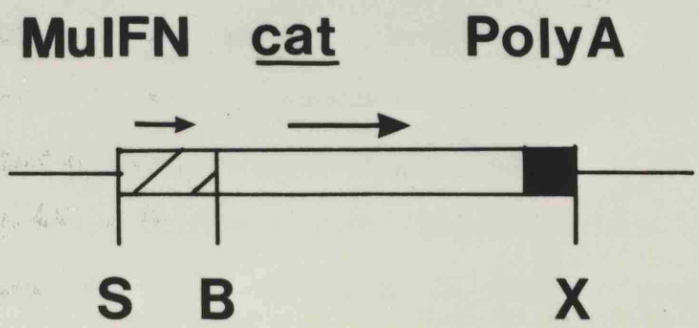
q) L929pCP8_1 " -

r) " " +

B) Shows the structure of the MuIFNα_1 promoter/*cat* genes used in these studies.

A

B



cells (Ryals *et al*, 1985). This result therefore, confirmed the structural and functional similarity observed between the mouse and human gene promoters. Moreover it demonstrated that sequences 5' to −94 are required for full transcriptional activity of the MuIFNα_1 gene promoter.

This analysis has established that most promoter mutants described above have remained capable of driving transcription and retained polyrI.rC mediated inducibility of the heterologous *cat* gene. To compare the strength of transcription and polyrI.rC inducibility of the mutant and wildtype promoters, CAT production in all the cell populations isolated (Table 6.4) was quantitated as described previously (Chapter 5). The data from these analyses has been obtained in parallel to that described for L929pCPWT 1 to 3 cells (Chapter 5). The behaviour of each class of promoter mutants, defined by the "region" of the MuIFNα_1 promoter under investigation is described separately below.

CAT Production from the Region I Mutant Promoter

Region I incorporated the TATA box variant sequences, TATTTAA, common to all IFNα genes. As discussed above (Section 6.1) a mutant promoter was constructed to investigate the possible relevance of the different TATA box sequences found in IFNα and IFNβ genes. Figure 6.3 shows the primary sequence of this promoter (P1), containing a closer consensus TATA box.

CAT production in polyrI.rC induced L−929 cells containing pCP1 was quantitated. A time course of CAT enzyme activity from induced L929pCP1 cells again demonstrated the linear accumulation of acetylated ^{14}C−chloramphenicol with time (Fig.6.16). Table 6.5 summarises the results of several studies, comparing CAT production in polyrI.rC induced L−929 cells containing the mutant P1 and wildtype (Chapter 5) MuIFNα_1 promoters. Student t analysis (Chapter 2) performed to compare the means of the distribution of CAT production, shows that there is no significant difference in *cat* gene expression in these cell populations (Table 6.5). A minimum estimate for the inducibility of CAT production of 250 to 400 fold compared to 300 to 750 fold from the mutant P1 and wildtype MuIFNα_1 promoters respectively (Table 6.5) suggested, therefore, that this mutation has no detectable effect on either the strength of transcription or polyrI.rC

FIGURE 6.16 Graph Showing a Quantitative Comparison of PolyrI.rC Induced CAT Production from the MuIFNα_1 Wildtype and P1 Promoters

See legend to Fig.5.15A.

pCPWT = CAT production in L929pCPWT$_1$ Cells

pCP1 = " " " L929pCP1 Cells

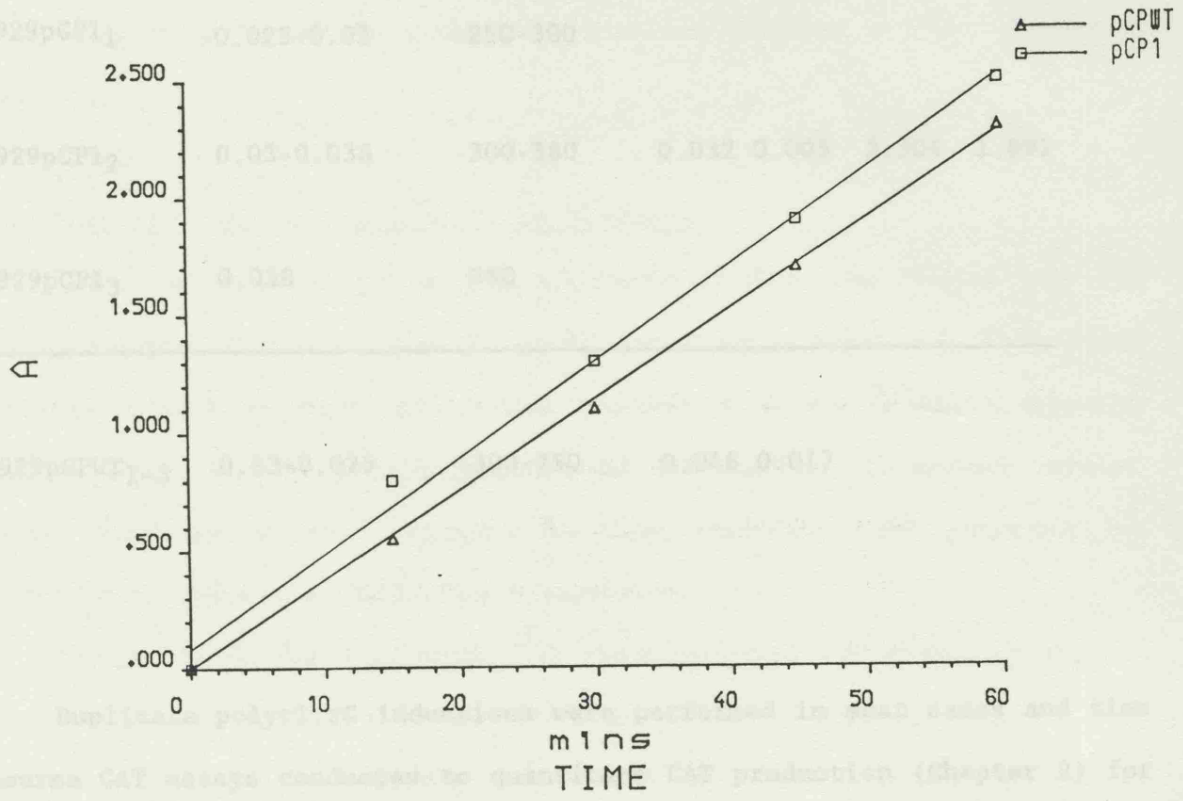

pCPWT
pCP1
2.500
2.000
1.500
1.000
.500
.000
0
10
20
30
40
50
60
mins
TIME

TABLE 6.5 Comparison of CAT Production in PolyrI.rC Induced L929pCP1 and L929pCPWT Cell Populations.

Populations	CAT Activity	Inducibility	Mean	σ	$t_{(0.05)}$	t_c
$L929pCP1_1$	0.025-0.03	250-300				
$L929pCP1_2$	0.03-0.038	300-380	0.032	0.005	2.306	1.891
$L929pCP1_3$	0.038	380				
$L929pCPWT_{1-3}$	0.03-0.075	300-750	0.046	0.017		

Duplicate polyrI.rC inductions were performed in most cases and time course CAT assays conducted to quantitate CAT production (Chapter 2) for independently derived cell populations. CAT activity (% Acetylation of chloramphenicol/min/mg total cell protein) was compared in cell lines containing mutant promoters with those containing the wildtype MuIFNα_1 promoter. The mean of the distribution of CAT production in the sets of populations (mutant and wildtype cells) are compared by a student t analysis (Chapter 2).

σ standard deviation

$t_c > t_{(0.05)}$ difference in CAT production is significant

$-t_c < -t_{(0.05)}$ " " " " " "

regulated inducibility of the MuIFNα_1 promoter in L−929 cells.

As described previously (Section 6.1) the TATA box variant TATTTAA is conserved in IFNα gene promoters. The introduction of a better consensus sequence, similar to that present in IFNβ gene promoters has no effect on the regulation of the MuIFNα_1 promoter in stably transfected L−929 cells. This suggests that the highly conserved IFNα sequence has no relevance with respect to either the mechanism of polyrI.rC regulated gene expression or the superior mouse IFNβ production compared to IFNα, observed in these cells.

CAT Production from the Region III Mutant Promoters

Region III promoter mutants were constructed at three sites (Section 6.1). The primary sequence of each of these (P2 to P6) are shown in Figure 6.3. They contain nucleotide substitutions which have created improved or perfect GGAGAA sequences, normally present in other MuIFNα promoters but not MuIFNα_1. To compare polyrI.rC induced regulation of gene expression by these promoters, CAT production was quantitated in L929pCP2 to L929pCP6 cell populations.

The plot of the first time course CAT assays performed with extracts of polyrI.rC induced L929pCP2 to L929pCP6 cell populations (Fig.6.17), suggested improved CAT production from the P6 promoter and reduced production from the P2 and P3 promoters, relative to the wildtype MuIFNα_1 promoter. However, CAT production by P4, P5 and wildtype MuIFNα_1 mutant promoters were of a similar magnitude.

Table 6.6 summarises the results obtained from a series of independent experiments performed with these cell populations. Inspection of these data reveals that CAT production in L−929 cells containing the P6 mutant promoter (containing a G to C transversion at +14) is highly variable, with inducibilty ranging from 150 to 1000 fold compared to the 300 to 750 fold observed in L929pCPWT cells. Statistical analysis (student t) suggested, therefore, that there was no genuinely significant difference in CAT expression from the P6 and wildtype MuIFNα_1 gene promoters. Nevertheless because of the variable expression observed (Table 6.6) it was concluded that further studies would be appropriate to determine the relative strength of polyrI.rC induced transcription from

FIGURE 6.17 Graph Showing a Quantitative Comparison of PolyrI.rC Induced CAT Production from the MuIFNα_1 Wildtype, P2, P3, P4, P5 and P6 Promoters

See legend to Fig.5.15A.

pCPWT = CAT production in L929pCPWT$_1$ Cells

pCP2 = " " " L929pCP2$_2$ Cells

pCP3 = " " " L929pCP3$_1$ "

pCP4 = " " " L929pCP4$_2$ "

pCP5 = " " " L929pCP5$_1$ "

pCP6 = " " " L929pCP6$_1$ "

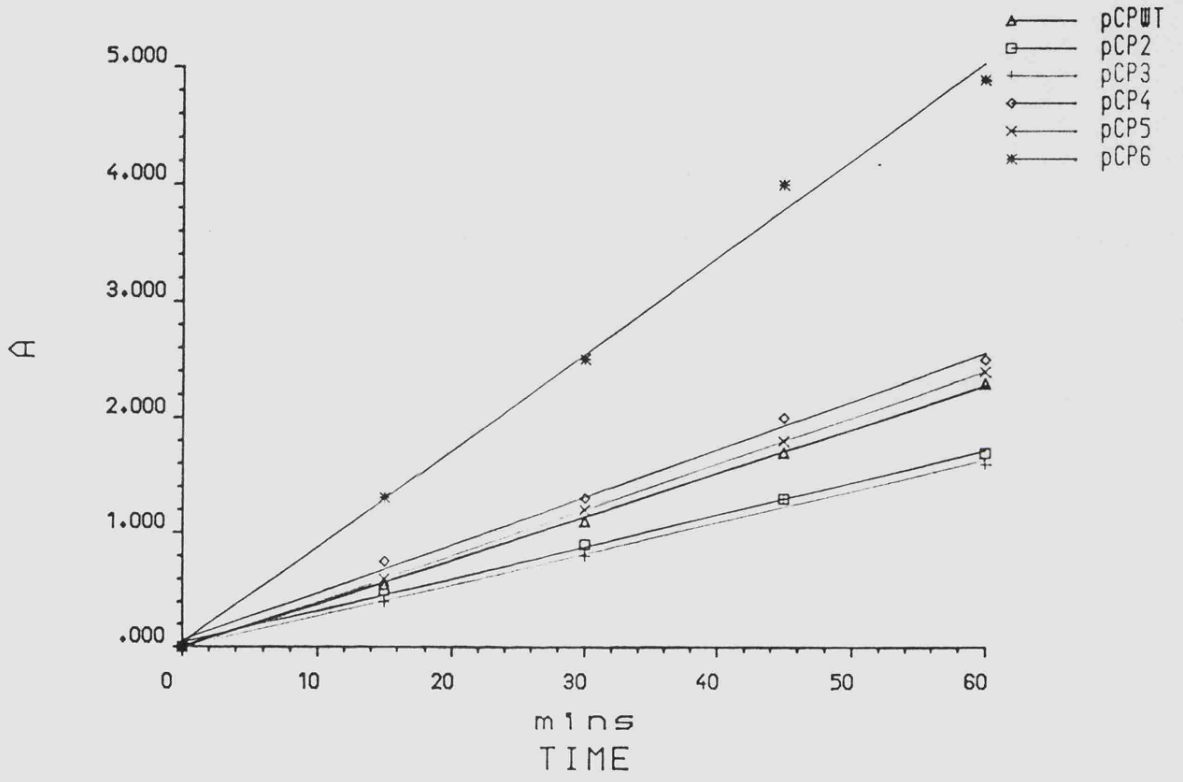

pCPWT
pCP2
pCP3
pCP4
pCP5
pCP6
5.000
4.000
3.000
2.000
1.000
.000
A
0
10
20
30
40
50
60
mins
TIME

TABLE 6.6 Comparison of CAT Production in PolyrI.rC Induced L929pCP2, L929pCP3, L929pCP4, L929pCP5, L929pCP6 and L929pCPWT Cell Populations.

Population	CAT Activity	Inducibility	Mean	σ	$t_{(0.05)}$	t_c
$L929pCP2_2$	0.03	300				
$L929pCP2_3$	0.014	140	0.023	0.007	2.365	2.718
$L929pCP3_1$	0.018-0.029	180-290				
$L929pCP4_2$	0.044-0.075	440-750				
$L929pCP4_3$	0.03-0.075	300-750	0.049	0.019	-2.447	-0.297
$L929pCP5_1$	0.04-0.05	400-500				
$L929pCP5_2$	0.053-0.06	530-600	0.050	0.007	-2.306	-1.920
$L929pCP5_3$	0.045	450				
$L929pCP6_1$	0.015-0.086	150-860				
$L929pCP6_2$	0.02-0.1	200-1000	0.054	0.034	2.306	-0.476
$L929pCP6_3$	0.05	500				
$L929pCPWT_{1-3}$	0.03-0.075	300-750	0.046	0.017		

See Table 6.5. σ:standard deviation,

$t_c > t_{(0.05)}$ difference in CAT production is significant

$-t_c < -t_{(0.05)}$ " " " " " "

this promoter (Chapter 7).

The introduction of a perfect region III motif into the MuIFNα_1 promoter, at the novel site (−87 to −82, Fig.6.3, P3) appeared to reduce polyrI.rC induced CAT production (L929pCP3 1) from the MuIFNα_1 promoter (Table 6.6). CAT production was also reduced in L929pCP2 1 and L929pCP2 2 cells, where the mutant promoter P2 responsible, contained the same A to G transversion at −87 as the P3 promoter (Fig.6.3). The additional transition at −84 found in P3 is probably not implicated in the observed reduction of CAT expression as it is shared by the active MuIFNα_2, MuIFNα_4 and HuIFNα_1 gene promoters. Student t analysis confirmed that the apparent reduction of CAT production observed in these cell populations is probably statistically relevant (Table 6.6).

Thus, the gene transfer studies described here, suggest that a transversion at −87 is probably responsible for a reduction in CAT expression from the MuIFNα_1 gene promoter. Inspection of this region (Fig.6.2) revealed a GTGAAAAG sequence (−94 to −86), which is homologous to the functionally important enhancer core sequence, GTGGAAAG (Weiher *et al*, 1983), first identified in the SV40 enhancer. Just downstream of this, there is a perfect enhancer core sequence (−81 to −74). Therefore, the reduction in promoter activity observed here may be the result of changing upstream activator sequences which share significant structural homology with enhancers.

The mutation at −87 also occurs within a region homologous to a repeat structure identified in the HuIFNα_1 gene and believed to be essential for rendering heterologous promoters inducible by virus (Ryals *et al*, 1985; Fig.6.3). Therefore, based on both functional and structural studies this sequence is strongly implicated in IFN gene regulation.

It is, therefore, likely that the reduced inducibility or strength of transcription from the P3 promoter is a consequence of altering the region II sequences, rather than the introduction of a region III motif. Additionally the −87 mutation occurs within the 14bp motif which is repeated at −171 to −159 in MuIFNα gene promoters (Section 6.1), raising the possibility that partial loss of function within region II can be compensated by upstream sequences. These observations are consistent with a common

role for this region in both virus and polyrI.rC induction, perhaps indicating a similar mechanism of gene activation or that the same sequences dictate the strength of transcription from activated promoters.

The two promoters, P4 and P5, contain one and two base substitutions respectively, creating a partial or complete region III_2 motif (Fig.6.3). Table 6.6 summarises the results comparing CAT production in L929pCP4 and L929pCP5 cells with L929pCPWT cells. Inducibilties of 300−750 and 400−700 fold from the P4 and P5 promoters respectively were determined (Table 6.6) demonstrating that these mutations have little influence on the inducibility or strength of transcription from the MuIFNα_1 promoter in polyrI.rC induced L−929 cells.

Thus, region III mutations, frequently found within other MuIFNα gene promoters, have been introduced into the MuIFNα_1 gene promoter and polyrI.rC mediated *cat* gene activation quantitated in L−929 cells. The apparent failure to significantly elevate CAT production in L−929 cells with these promoters may reflect the conclusions derived previously (Chapter 5), that the MuIFNα_1 promoter is fully active in this system. Thus, these analyses do not provide any evidence to support the hypothesis that the GGAGAA sequences common to HuIFNβ and MuIFNα but not HuIFNα genes might contribute to polyrI.rC mediated IFN production in human or mouse fibroblasts.

Despite the demonstration from this work that improving the level of homology of the MuIFNα_1 promoter region III sequences to the consensus GGAGAA motif does not significantly improve the function of this promoter, the possibility still remains that these relatively conserved regions may be important in IFN expression. This issue might best be addressed by disruption of these regions by site directed mutagenesis together with gene transfer studies.

CAT Production from the Region IV Mutant Promoters

Figure 6.3 shows the primary nucleotide structure of the region IV promoter mutants, P7, $P7^8$ and P8, containing mutations within an inverted repeat region, 5' to the TATA box. Figure 6.18 shows the first time course CAT assays performed on polyrI.rC induced cell extracts, which suggested that there might be a small reduction in

CAT production in L929pCP7 and L929pCP7^8 cells compared to L929pCPWT, whereas L929pCP8 produced elevated levels of CAT activity.

PolyrI.rC induced CAT production in these cell populations was quantitated as described previously and the data obtained is summarised in Table 6.7. These results show that in a series of assays these mutant promoters, P7, P7^8 and P8, produced polyrI.rC mediated CAT induction over a range of 180 to 500, 130 to 2000 and 330 to 1000 fold each, compared with 300 to 750 fold from the wildtype MuIFNα_1 promoter. The apparent reduction in CAT production originally observed from the P7 and P7^8 promoters, and increased expression from the P8 promoter is thus, not statistically significant (Table 6.7). Nevertheless, again because of the variable expression observed, these promoters were investigated further by an alternative procedure which is described in Chapter 7.

Thus, these data directly demonstrated that the complementarity observed between the inverted repeat sequences is not required for polyrI.rC induced gene regulation because disruption of homology at either site (promoters P7 and P8) does not significantly affect inducibility (Table 6.7).

As discussed previously basal gene expression from the wildtype MuIFNα_1 promoter was difficult to quantitate due to the low levels of expression observed (Chapter 5). However, the cell populations containing plasmid pCP7 presented one of the few opportunities where basal CAT expression could be consistently measured. Table 6.8 shows CAT production derived from several experiments performed with these cells to quantitate basal gene expression. From this analysis an estimate of a 2−5 fold increase relative to wildtype was determined. Therefore, it is likely that these mutations result in an increase of basal expression from the MuIFNα_1 promoter, possibly suggesting a negative regulatory role in the control of IFN gene expression. However these data are not supported by non−induced CAT expression in cells transfected with pCP7^8. This promoter contained the same mutation as pCP7, plus two additional base substitutions at the distal site (Fig.6.3).

FIGURE 6.18 Graph Showing a Quantitative Comparison of PolyrI.rC Induced CAT Production from the MuIFNα_1 Wildtype, P7, P7^8 and P8 Promoters

See legend to Fig.5.15A.

pCPWT = CAT production in L929pCPWT$_1$ Cells

pCP7 = " " " L929pCP7$_1$ Cells

pCP7^8 = " " " L929pCP7$^8_1$ "

pCP8 = " " " L929pCP8$_1$ "

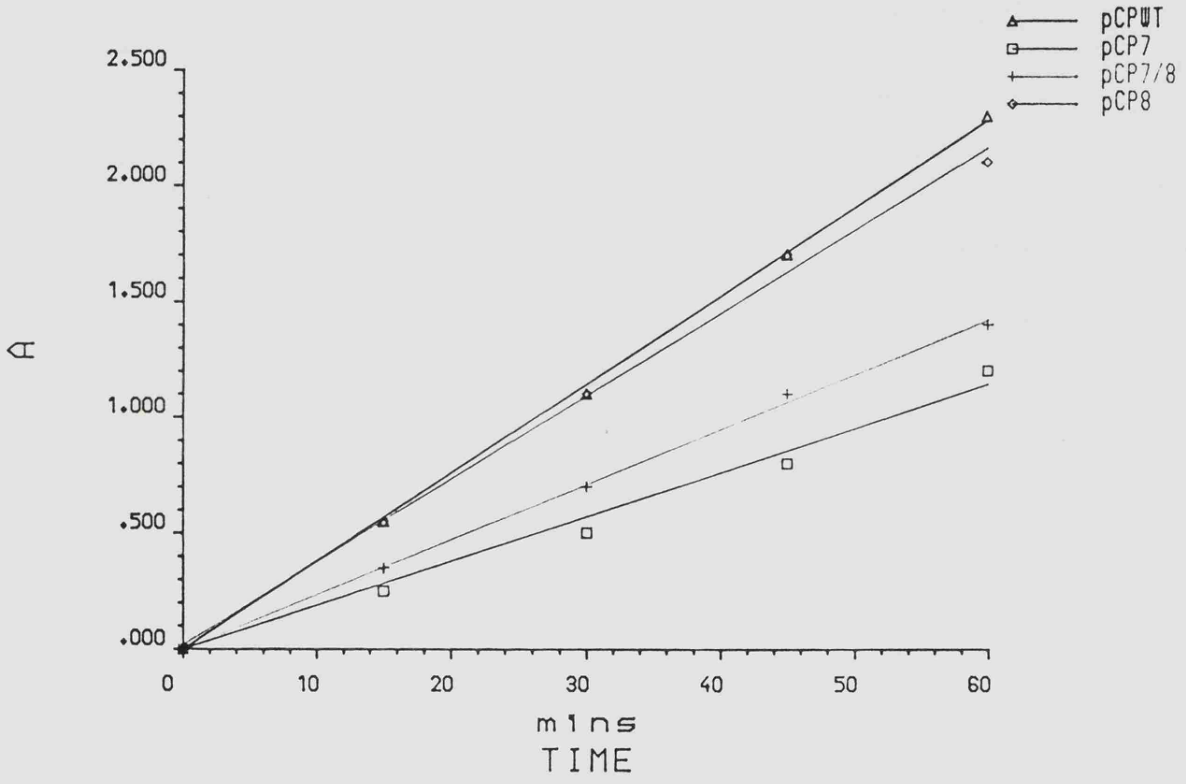

pCPWT
pCP7
pCP7/8
pCP8
2.500
2.000
1.500
1.000
.500
.000
A
0
10
20
30
40
50
60
mins
TIME

TABLE 6.7 Comparison of CAT Production in PolyrI.rC Induced L929pCP7, L929pCP7[8], L929pCP8 and L929pCPWT Cell Populations

Population	CAT activity	Inducibilty	Mean	σ	$t_{(0.05)}$	t_c
$L929pCP7_1$	0.018-0.025	180-250				
$L929pCP7_2$	0.02-0.04	200-400	0.03	0.012	2.306	1.116
$L929pCP7_3$	0.05	500				
$L929pCP7^8{}_1$	0.013-0.025	130-250				
$L929pCP7^8{}_2$	0.017-0.033	170-330	0.058	0.07	-2.306	-0.829
$L929pCP7^8{}_3$	0.2	2000				
$L929pCP8_1$	0.033-0.038	330-380				
$L929pCP8_2$	0.033-0.1	330-1000	0.061	0.03	-2.306	-0.939
$L929pCP8_3$	0.1	1000				
$L929pCPWT_{1-3}$	0.03-0.075	300-750	0.046	0.017		

See Table 6.5. σ: standard deviation

$t_c > t_{(0.05)}$ difference in CAT production is significant

$-t_c < t_{(0.05)}$ " " " " " "

TABLE 6.8 Comparison of CAT Production in Non-induced L929pCP7 and L929pCPWT Cell Populations

Population	CAT Activity ($\times 10^{-4}$)
$L929pCP7_1$	1.1-1.9
$L929pCP7_2$	5.3-7.3
$L929pCP7_3$	1.6
L929pCPWT	4.5-5
$L929pCPWT_2$	und-0.28
$L929pCPWT_3$	0.42

CAT time course assays were performed over a period of 8 hrs (Chapter 2). CAT activity is expressed as the % acetylation of chloramphenicol/min/mg total cell protein.

Conclusions

Several structural features common to IFNα gene promoters have been identified and experiments conducted to begin dissection of the structure/function relationships of the IFN promoter. Both conserved sequence homology and expression data of deletion mutants performed here and by others has confirmed the structural and functional similarities of HuIFNα$_1$ and MuIFNα$_1$ gene promoters.

The ability of the MuIFNα$_1$ gene promoter to confer polyrI.rC inducibility upon the heteologous *cat* gene (Chapter 5) is lost upon deletion of 77bp of 5' flanking DNA sequence. Thus, 94bp of promoter sequence are insufficient to maintain polyrI.rC mediated activation of gene expression. This deletion removes sequences homologous to those required for full activation of the HuIFNα$_1$ gene by virus, suggesting a common function with both inducers. In addition, a single G to C transition at −8 apparently inactivates MuIFNα$_1$ gene promoter activity, but the reasons for this are unclear (Chapter 7).

Site directed mutagenesis has been employed successfully to create a series of predetermined promoter mutants which have focused attention on the functional relevance of specific structural motifs within the MuIFNα$_1$ gene promoter. Some mutations appear to have a minor effect on promoter function *in vivo*. For example the nucleotide transitions at −33 and −35 (P7) within the proximal inverted repeat (Fig.6.3) possibly elevate basal gene transcription, thus, implying a negative regulatory role for the inverted repeat sequences (region IV) in the control of IFN gene expression.

Mutation of the MuIFNα$_1$ gene promoter at −87, from A to G (P2 and P3), within a sequence which shows significant homology to enhancer sequences, produces a statistically significant 2 fold reduction of polyrI.rC mediated promoter activity. The behaviour of these promoters, together with the deletion derivative, are consistent with the requirement for upstream regulatory elements with structural features of enhancers. These elements make a positive contribution to the efficiency of transcription initiation from the induced MuIFNα$_1$ gene promoter.

Thus, examples of both positive and negative effects upon the MuIFNα$_1$ gene promoter have been observed. Other regions possibly influencing the transcription

efficiency of the activated MuIFNα_1 gene promoter were also identified but firm conclusions concerning their behaviour was prevented by the variation in expression observed. These again included the nucleotide transitions at −33 and −35 (P7) and the single base transversion at +14 (P6) creating a perfect GGAGAA motif, which possibly depress and elevate induced CAT production from this promoter respectively.

Therefore, this analysis has begun to highlight regions of functional relevance within the MuIFNα_1 gene promoter. Clearly the ability to make predetermined promoter mutations by site directed mutagenesis is a very valuable approach to identifying sequences involved in the regulation of gene expression. These studies now need to be extended to concentrate mutations to specific sites for a more comprehensive analyses to define both the important nucleotides and their precise role in the control of IFNα gene expression.

CHAPTER 7

STRUCTURE/FUNCTION ANALYSIS OF THE MuIFNα1 PROMOTER PART II: ADDITIONAL MUTANTS AND S−I NUCLEASE ANALYSIS OF MuIFNα1 PROMOTER FUNCTION

The double stranded RNA polyrI.rC characteristically induces transient expression of IFNα and IFNβ mRNAs and proteins in a variety of mammalian cells including those of mouse (Chapter 1). Work described in both Chapters 5 and 6 demonstrated polyrI.rC/DEAE dextran (polyrI.rC) mediated production of CAT by the MuIFNα1/*cat* gene in L−929 cells. However, the kinetic aspects of expression were not addressed. Additionally, the transcription initiation site, an important aspect of promoter function was not determined. To address these issues, time course CAT assays and S−1 nuclease protection studies (Weaver and Weissmann, 1979) have been undertaken.

To provide additional information on the behaviour of the mutant IFN promoters described in Chapter 6 further mutants have been created. In addition, although the CAT assay system used to examine IFN promoter function previously (Chapter 5 and 6) had proved to be adequate for demonstrating major changes in promoter activity, the experimental variation observed by the direct comparison of CAT production in essentially identical cell populations had made it difficult to detect subtle changes in promoter activity. An S−1 nuclease protection assay which allows the direct comparison of transcription from both mutant and wildtype promoter/*cat* gene constructs within the same cell population has, therefore, been developed in an attempt to provide a more sensitive means of comparing promoter activity.

7.1 Kinetics of PolyrI.rC Induced CAT Production

The production of CAT by L−929 cells stably transfected with the wildtype MuIFNα1 promoter/*cat* gene (L929pCPWT cells) had previously been monitored 24hrs after polyrI.rC treatment (Chapter 5). However to ensure that the kinetics of MuIFNα1 promoter/*cat* gene expression were no different to the endogenous IFN genes, CAT

production was monitored over a 24hr period in induced L929pCPWT cells.

As shown in Figure 7.1 CAT activity is detectable at 9hrs, is maximal between 12−16hrs and declines to approximately 30% to 50% of this level by 24hrs post−induction. This suggests that the CAT assays described previously (Chapters 5 and 6) were not performed under optimal conditions, although a significant proportion of enzyme activity still remains after 24hrs. Nevertheless, a transient profile of CAT production was observed, corresponding to mouse IFN gene expression. The mixture of IFNs induced by polyrI.rC in L−929 cells are the products of a family of genes (Chapter 1). The apparent differences in the kinetics of CAT (12−16hrs) and IFN (16−24hrs) production may suggest subtle differences in either the rate of the response of individual MuIFNα genes to the inducer or the relative stability of *cat*/MuIFN gene mRNA or their proteins. The question of relative protein stability is difficult to assess as normally IFN is exported from the cell. The cellular location of the CAT protein has not been examined here, but it is probably intra−cellular.

7.2 Kinetics and Transcription Initiation Site of Wildtype MuIFNα$_1$ Promoter/*cat* Gene mRNA Transcription

To investigate the MuIFNα$_1$ promoter/*cat* gene system further, the kinetics of polyrI.rC induced mRNA accumulation has been examined. For this purpose an S−1 nuclease protection assay was employed to allow a simultaneous verification that the anticipated transcription intiation site was utilised. The principle of this assay is illustrated schematically in Figure 7.2B.

A suitable recombinant phage for generating a uniformly labelled single stranded DNA (ssDNA) probe (Chapter 2), complementary to chimaeric MuIFNα$_1$/*cat* gene transcripts, was created for S−1 nuclease protection studies. This was constructed by inserting a 410bp *Eco*RI/*Pvu*II fragment from pCPWT, into the *Eco*RI/*Sma*I site of M13mp10, producing mS−1 (Fig.7.2).

The sequence of the entire 410bp insert of mS−1 has been verified (Fig.7.3B), facilitated by the use of the oligonucleotide 5 mismatch sequencing primer (Chapter 6) which primes DNA synthesis from a midway point in this sequence. Preparation of a

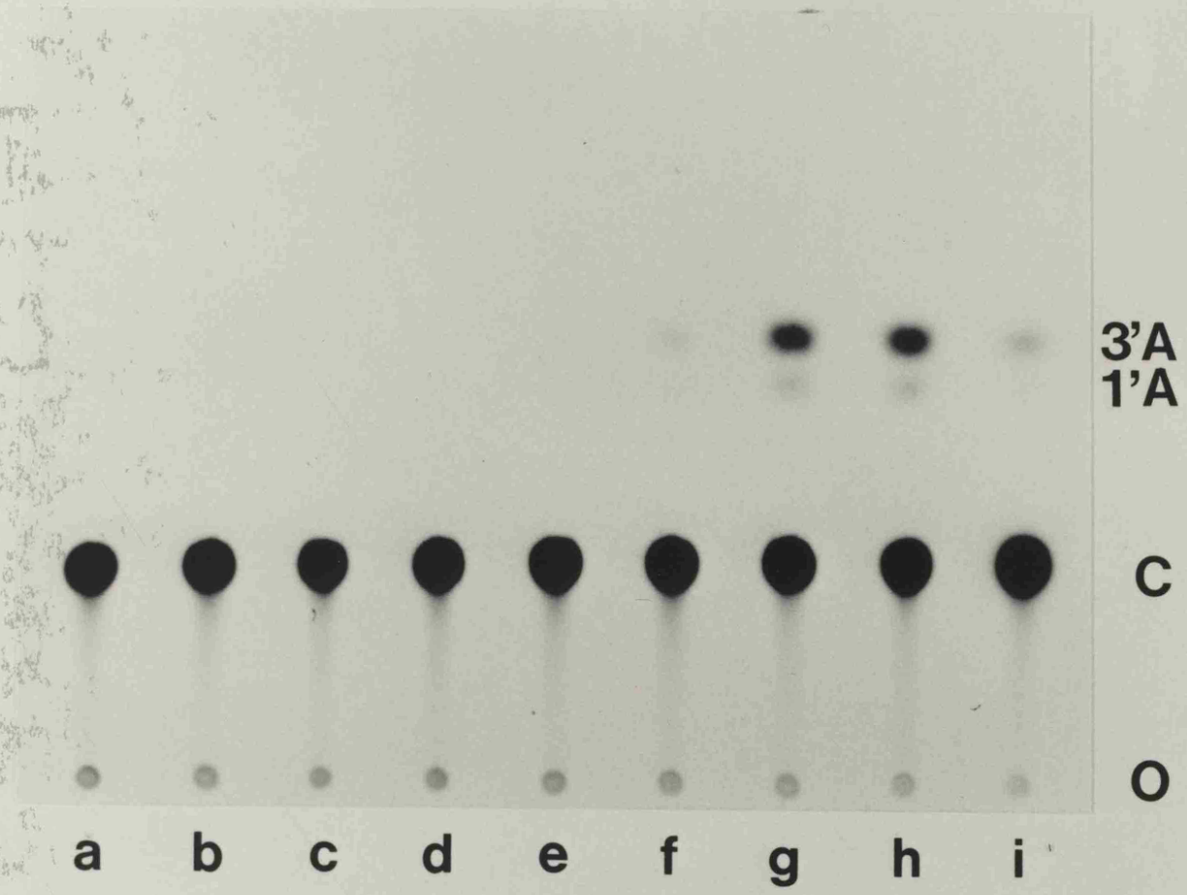
FIGURE 7.1 Kinetics of CAT Production in PolyrI.rC Induced L929pCPWT Cells

Confluent monolayers of L929pCPWT$_3$ cells (in 9cm Nunc tissue culture petri-dishes) were induced with polyrI.rC and extracts prepared at the specified time points. CAT assays were performed with 10μl of extract using 0.2μCi of ^{14}C-chloramphenicol (NEN) for 1hr. T.L.C. was performed in 95:5 chloroform:methanol (Chapter 2). Autoradiography was performed for 15hrs using Fuji RX film (A).

a) non-induced (-)
b) 1hr polyrI.rC
c) 3hrs "
d) 5hrs "
e) 7hrs "
f) 9hrs polyrI.rC
g) 12hrs "
h) 16hrs "
i) 24hrs "

Below the structure of the MuIFNα_1/*cat* gene in L929pCPWT cells is shown (B).

A



B

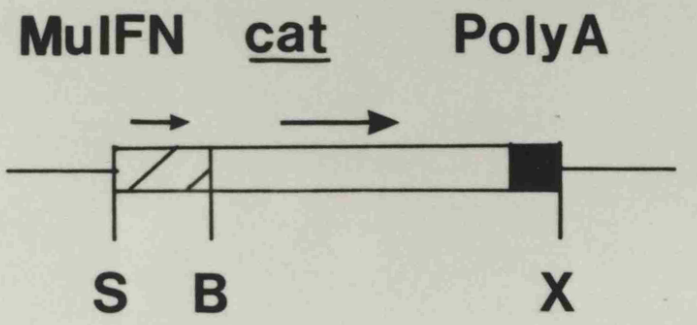


FIGURE 7.2 The Transcription Initiation Site of the Chimaeric Gene in L929pCPWT Cells

25μg samples of total RNA, prepared from polyrI.rC induced L929pCPWT$_3$ cells at specified time points were used for S-1 nuclease protection studies (Chapter 2). After S-1 treatment samples were analysed on a 5% polyacrylamide/urea denaturing gel and then autoradiography was performed for 48hrs at -70°C using Fuji RX film with an intensifying screen (A).

1) Probe only - S-1

2) " only + S-1

3) Probe +RNA, 5hrs post-induction + S-1

4) " " 7hrs " "

5) " " 9hrs " "

6) " " 12hrs " "

7) " " 16hrs " "

8) Probe only - S-1

A,C,G and T = sequencing ladder marker prepared using the oligonucleotide 12 primer and mS-1 template DNA (Fig.7.3). P = Probe, A = Probe protected by readthrough transcription and T = Probe protected by correctly initiated transcripts.

The complement of partial sequence data derived from the marker is shown indicating the expected transcription initiation site of the MuIFNα_1 gene promoter (arrow) which comigrates with the protected fragment T.

Below is illustrated schematically the principle of the S-1 nuclease protection assay (B). The primary transcript (pt) derived from the transcription initiation site (Tis) forms a double stranded RNA/DNA hybrid with complementary sequences within the probe (P), protecting it from S-1 nuclease digestion. This generates a 206nt protected fragment (T). E = *Eco*RI, Pv = *Pvu*II, the striped box = the MuIFNα_1 promoter, the open box = the *cat* gene and the shaded box = the HSV polyA sequences.

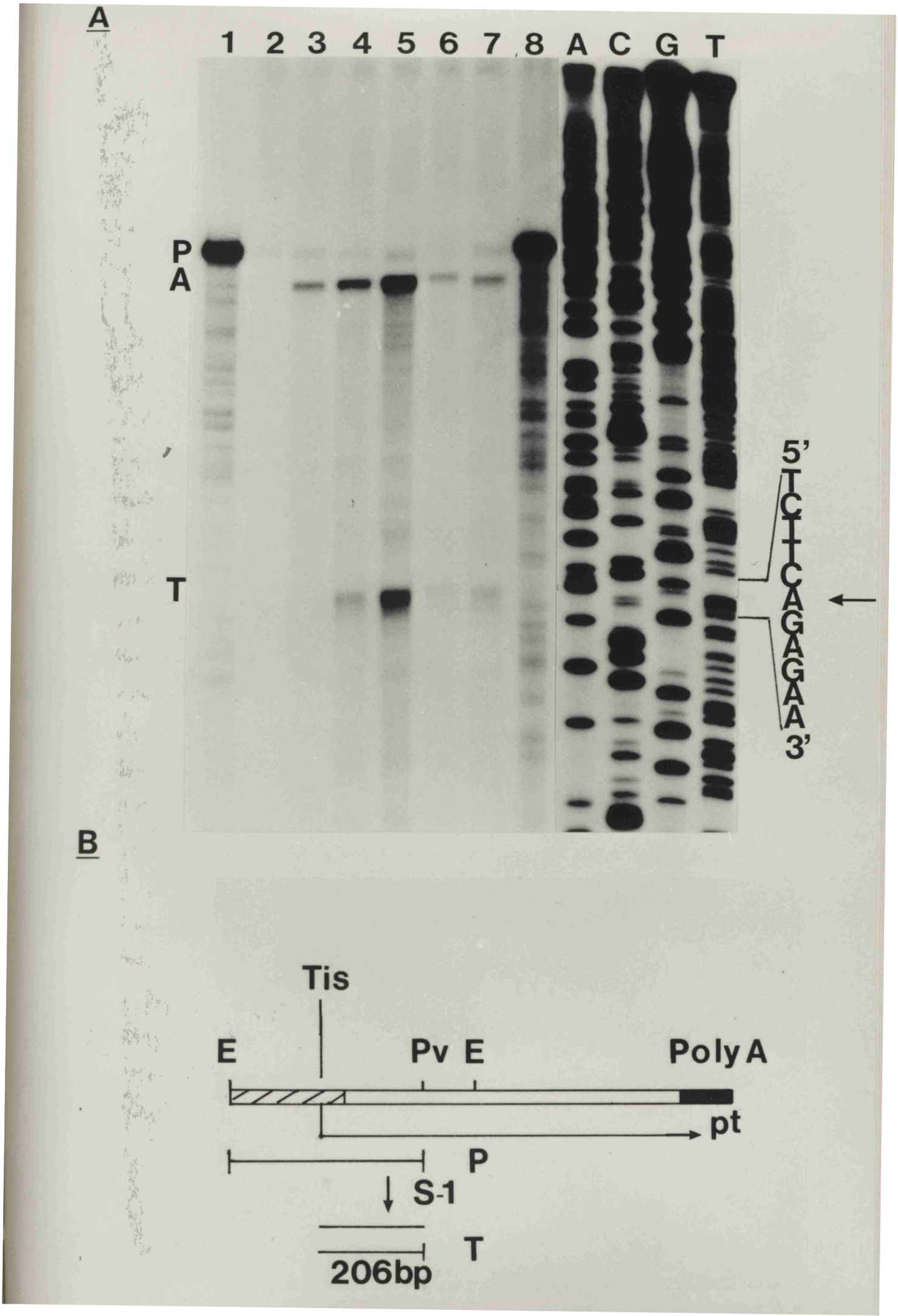
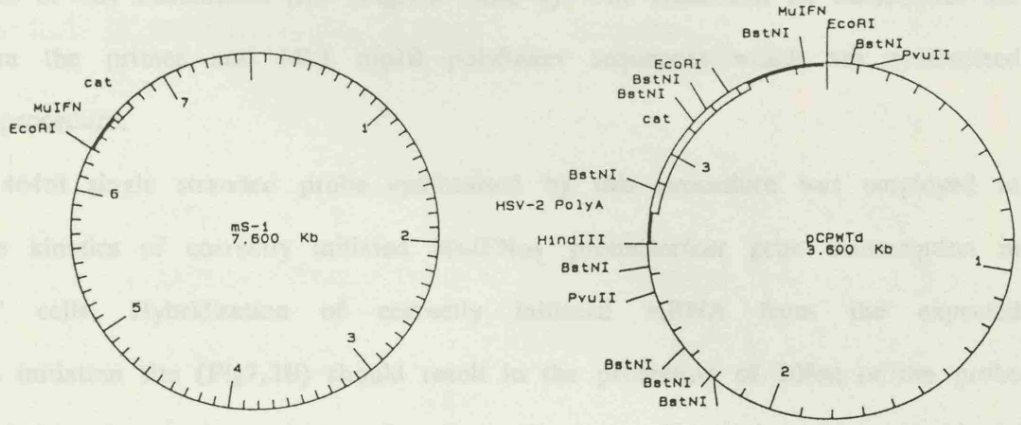

A
1 2 3 4 5 6 7 8 A C G T
P
A
T
5' TCTTCAGAGAA 3'
B
Tis
E
Pv E
Poly A
pt
P
S-1
T
206bp

FIGURE 7.3 Restriction Maps of mS-1 and pCPWTΔ (pCPWTd)

The restriction map of pCPWTΔ for *Bst*NI, *Eco*RI, *Hin*dIII and *Pvu*II restriction enzymes is shown (A). Below (B) is illustrated the primary nucleotide sequence of the 410bp *Eco*RI/*Pvu*II (underlined) fragment of pCPWT which has been cloned into M13mp10 to create mS-1 (A). The MuIFNα_1 gene promoter TATA box, transcription initiation site (arrow) and *cat* gene initiation codon (ATG) have been underlined (B). Futhermore, the oligonucleotide 12 sequencing and oligonucleotide 13 mismatched primers are shown aligned to their respective complementary sequences of the *cat* gene coding region (B).

A)



B)

*Eco*RI 10 20 30 40 50 60

GAATTCGAGC TCGCCCAGCT TTTGATGAGG ACCAGTGAAA GAGGAAGCAA TAATGAAAAC

70 80 90 100 110 120

CACAATGGTT TAGAAAACAC CCAGACGCAA GCAGAGAATG AGTTAAAGAA AGTGAAAAGA

130 140 150 160 170 180

CAAGTGGAAA GTGATGGAAG GGCATTCAGA AAGTAAAAAC CAGTGTTTGC CCTATTTAAG

190 200 ↓ 210 220 230 240

ACACATTCAC CCAGGATGGT CTTCAGAGAA CCTAGAGGGG AAGGATCAGG ACCAAACAGT

250 260 270 280 290 300

CCAGAAGACC AGAAGCTGGG GATCCGTCGA GATTTTCAGG AGCTAAGGAA GCTAAAATGG

310 320 330 340 350 360

AGAAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCCA ATGGCATCGT AAAGAACATT

3' A

370 380 390 400 *Pvu*II

TTGAGGCATT TCAGTCAGTT GCTCAATGTA CCTATAACCA GACCGTTCAG

AACTCC*TAA *GTCAGTC 5' 3' GATATTGGT CTGGCAAGTC 5'

T G

Oligonucleotide 13 Oligonucleotide 12

uniformly labelled single stranded DNA probe from the mS−1 template using the universal primer (Amersham 17mer) as described in Chapter 2 produced an anticipated major species of 464 nucleotides (nt) (Fig.7.2 Lane 1). The additional 54 nucleotides are derived from the primer and M13 mp10 polylinker sequences which are synthesised during this procedure.

The 464nt single stranded probe synthesised by this procedure was employed to examine the kinetics of correctly initiated MuIFNα_1 promoter/*cat* gene transcription in L929pCPWT cells. Hybridisation of correctly initiated mRNA from the expected transcription initiation site (Fig7.3B) should result in the protection of 206nt of the probe from degradation by S−1 nuclease digestion. Thus, a 206nt fragment should be specifically seen from induced cells only. Figure 7.2 shows the appearance of such a probe fragment (T), protected by the mRNA of polyrI.rC induced cells (Fig.7.2 Lanes 3,4,5,6 and 7).

The exact size of fragment T was determined by comparison with the sequencing ladder marker (Fig.7.2), prepared using the oligonucleotide 12 sequencing primer (Table 7.1) and mS−1 template DNA (Fig.7.3 A and B). Both the protected fragment and derivatives of the sequencing reaction should share common 5' termini, thus, fragments of the same size within the S–1 and sequencing lanes should indicate transcription initiation and chain termination at the same nucleotide respectively. This procedure should, therefore, reveal the precise nucleotide sequence from which MuIFNα_1/*cat* gene transcripts originate. Inspection of the nucleotide sequence displayed in Figure 7.2 revealed that the inducer dependent protected fragment (T) comigrated with the sequence around CAG. The most likely conclusion, therefore, was that 206nt of the probe had been protected by MuIFNα_1/*cat* gene transcripts derived from the anticipated transcription initiation site of the MuIFNα_1 gene promoter (Fig.7.3B). The appearance of the S−1 nuclease protected probe fragment is, therefore, indicative of correct transcription from the MuIFNα_1/*cat* gene.

The kinetics of induced MuIFNα_1 gene transcription was also examined by monitoring the accumulation of the 206nt protected probe fragment. Inspection of Figure 7.2 shows this fragment is undetectable at time 0, appears by 7hrs, reaches a maximum

TABLE 7.1 Structure of Oligonucleotides for Site Directed Mutagenesis and Sequencing.

Oligonucleotide No	Oligonucleotide Structure 5' 3'	Purpose
	-102 * -120	
9	AACTCATTCTCCGCTTGCG	Mutagenesis: P9(region III_3)
	-77 * * -95	
10	TTCCACTTGTCTTTTCACT	Mutagenesis: P3Δ77R
	+24 * +6	
11	CTGATCCTTTTCCTCTAGG	Mutagenesis: P6tR
	+205 +187	
12	CAGAACGGTCTGGTTATAG	Sequencing Primer
	+173 * * +155	
13	CTGACTGGAATTCCTCAAA	Mutagenesis: mS-1RI

* site of mutation

by 9hrs and declines partially by 12−16hrs (Fig.7.3 Lanes 3,4,5,6 and 7). Thus, polyrI.rC mediates transient activation of transcription from the MuIFNα_1 promoter in L−929 cells. Together with the data described previously (Section 7.1) it was concluded that *cat* gene transcription precedes CAT production in polyrI.rC induced L929pCPWT cells. The abundance of *cat* mRNA being maximal approximately 9hrs after induction.

The presence of a separate transcript (A) protecting approximately 410nt of the 464nt probe was also suggested by this study (Fig.7.2). The most likely explanation for this is readthrough transcription from upstream cryptic promoters located within plasmid DNA (Fig.7.4). The approximate size of this fragment suggested these transcripts protect all 410nt of the probe *Eco*RI/*Pvu*II fragment, supporting this interpretation.

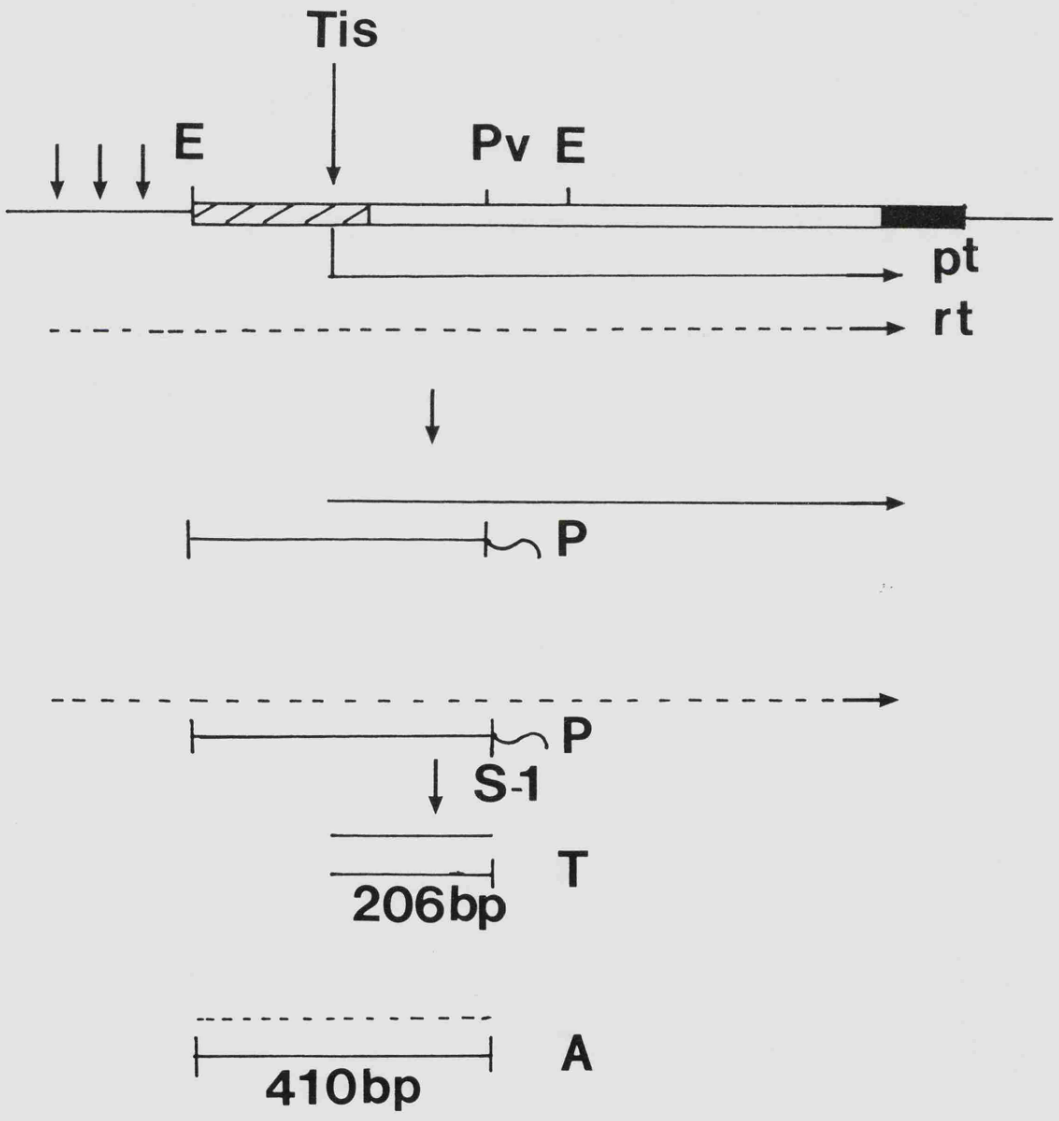
Despite the above conclusion it is unlikely that readthrough transcription contributes to CAT activity in these cells. Computer assisted analysis of the MuIFNα_1 promoter/*cat* gene employing Analyseq (Staden, 1984) has revealed in frame nonsense mutations within the promoter sequence. Thus, the formation of a functional CAT fusion protein is unlikely. This is supported by the fact that this transcript is abundant in the absence of polyrI.rC treatment when CAT production is undetectable in L929pCPWT cells.

Although the probable readthrough transcripts are expressed at high basal levels scanning densitometry of autoradiographs suggested they are induced approximately 7 fold in polyrI.rC treated cells, following the same profile of expression observed with the correctly initiated transcript (Fig.7.2 Lanes 3,4,5,6, and 7). This suggested either, that the cryptic promoter(s) might also respond directly to the inducer or, more likely, that the efficiency of transcription initiation is potentiated by the activation of the MuIFNα_1 promoter.

Transcriptional activation by elements 3' to a heterologous promoter is characteristic of enhancers (Chapter 1). This ability to activate transcription of heterologous promoters in induced cells, therefore, would be consistent with the MuIFNα_1 promoter containing an inducible cellular enhancer element. This observation might be worthy of further investigation to confirm, locate and define the sequences responsible using heterologous promoters which have been employed extensively in the characterisation of enhancer elements (Chapter 1).

FIGURE 7.4 Schematic Representation of the Likely Origin of the 410nt and 206nt S-1 Protected Fragments

See legend to Fig.7.3. The primary (pt) and readthrough (rt) transcripts derived from the Tis and cryptic promoters (arrows) upstream of the MuIFNα_1 gene promoter (striped box) respectively, protect 206nt (T) and 410nt (A) of the 464nt ssDNA probe from S-1 nuclease digestion. The additional 54nt of the probe (~) are not complementary, being derived from the universal primer and M13 sequences during the synthesis of the probe from mS-1.


Tis
E
Pv
E
pt
rt
P
P
S-1
T
206bp
A
410bp

Thus, S−1 analysis has identified two inducible transcripts, one of which is derived from the MuIFNα1 promoter/*cat* gene in polyrI.rC induced L929pCPWT cells. This transcript originates exactly from the anticipated transcription initiation site of the hybrid gene. Its maximal abundance precedes that of CAT expression. The second transcript is probably derived by readthrough transcription of cryptic promoters located upstream of the chimaeric MuIFNα1/*cat* gene. Its expression does not correspond with CAT activity but is enhanced by polyrI.rC.

7.3 Construction of a MuIFNα1 Promoter/*cat* Pseudogene

As discussed in Chapter 6, it was possible to discern certain trends in the relative strength of MuIFNα1 and mutant promoter constructs by monitoring polyrI.rC induced CAT production in stably transfected L−929 cells. However, experimental variation between duplicate assays performed on both the same and independently derived cell populations containing identical promoter/*cat* constructs made it extremely difficult to demonstrate subtle changes in promoter activity. In an attempt to address this difficulty an S−1 nuclease protection assay has been developed which allows a direct comparison of transcription of the *cat* gene from two different promoters functioning within a single cell population. This assay is dependent upon the creation of a pseudogene comprising the wildtype MuIFNα1 promoter and an internally deleted *cat* gene. This enabled the distinction of correctly initiated transcripts of two distinctly sized mRNAs, using in both cases the same 464nt S−1 probe described in Section 7.2.

A suitable pseudogene for this purpose, pCPWTΔ, was created for this purpose by introducing a 366bp deletion into the *cat* gene of pCPWT (Fig.7.5). The first stage of this process required the excision of a 273bp *Eco*RI fragment from the plasmid p22, performed by *Eco*RI digestion, religation and transformation of plasmid DNA into *E.coli* JA221 selecting for ampicillin resistant colonies. Deletion derivatives lacking the 273bp fragment were identified in three of four plasmid DNAs examined by *Eco*RI restriction analysis (p22ΔRI).

The wildtype MuIFNα1 promoter and 5' *cat* gene sequences were introduced into p22ΔRI as an *Eco*RI module to create pCPWTΔ (Fig.7.2). This was achieved in two

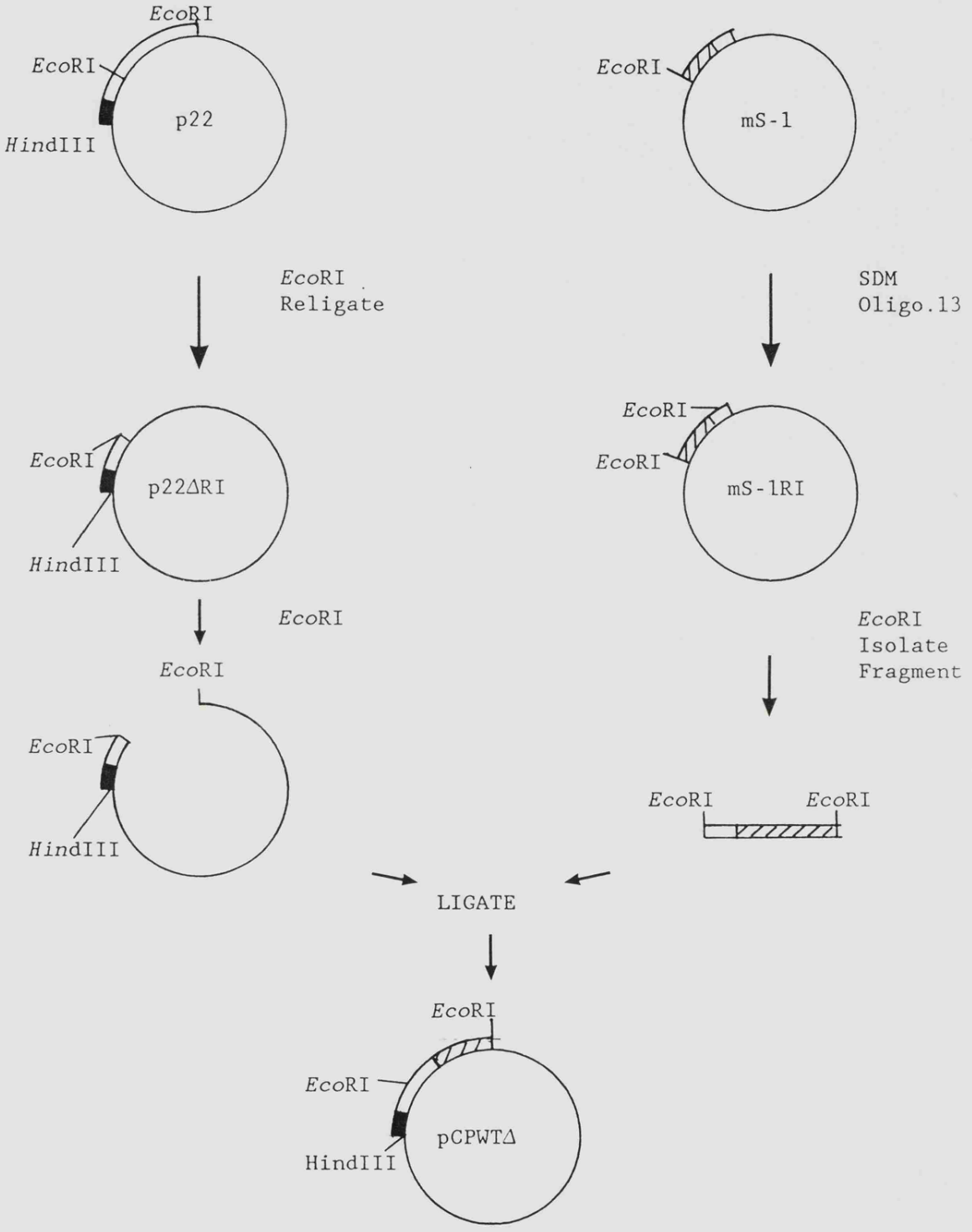


FIGURE 7.5 Strategy for the Construction of pCPWTΔ

Striped box = MuIFNα_1 promoter, open box = *cat* gene and filled box = HSV-2 PolyA.

stages a) by creating a novel *Eco*RI site within the *cat* gene sequence of the M13 recombinant phage mS−1, and b) by cloning the newly created 365bp *Eco*RI promoter/5' *cat* module into p22ΔRI (Fig.7.4).

A development of the site directed mutagenesis procedure used previously (Chapter 6) which increases the efficiency and ease of mutant isolation was adopted to generate the *Eco*RI site in mS−1 (Kunkel, 1985). Thus, the mismatch oligonucleotide 13 (Table 7.1) was designed to create an *Eco*RI site within mS−1 DNA at the location illustrated in Figure 7.3. Only two base changes were necessary for this purpose.

To create the mutation, oligonucleotide 13 was used to prime *in vitro* synthesis of heteroduplex DNA from a uracil containing M13ssDNA (U−DNA) template derived from mS−1 (Chapter 2). The heteroduplex was transfected into *E.coli* JM101. As shown in Table 7.2, the heteroduplex is approximately 1000 fold more efficient at transfecting JM101 cells than the U−DNA template, consistent with the expected positive selection for the *in vitro* synthesised DNA or mutant strand (Kunkel, 1985). *Eco*RI restriction analysis of twelve prospective mutant phage R.F DNAs examined demonstrated that eight contained an additional *Eco*RI site, as revealed by the release of a 365bp fragment detected by P.A.G.E. Sequencing analysis of single stranded DNA for one of these phage confirmed that an *Eco*RI site had been created at the intended location and that the remaining sequence was intact. Thus, mS−1RI had been generated (Fig7.5).

Construction of the pseudogene pCPWTΔ was completed by cloning the 365bp *Eco*RI fragment from mS−1RI into the *Eco*RI site of p22ΔRI and transforming into *E.coli* JA221 selecting for ampicillin resistance. Prospective recombinants were identified by colony hybridisation using a 5' end−labelled oligonucleotide 10 (Table 7.1) probe which recognises MuIFNα_1 promoter sequences (Table 7.1). Plasmid DNA prepared from ten positives identified from one hundred colonies screened, were examined by restriction analysis with *Sst*I/*Hin*dIII, identifying two possessing the expected restriction map for pCPWTΔ (Fig.7.2). The remaining eight clones contained the 365bp *Eco*RI fragment in the opposite orientation.

Thus, a plasmid (pCPWTΔ) containing the wildtype MuIFNα_1 promoter upstream of an internally deleted *cat* gene had been generated. Correctly initiated transcripts from

TABLE 7.2 Relative Transfection Efficiency of M13 Uracil Containing ssDNA and *in Vitro* Synthesised Hetroduplex DNA Templates in *E.coli* JM101 and BW313

M13 Template DNA	BW313	JM101
mS-1 U.ssDNA	1000	3
mS-1 U.H.DNA	200	200

0.2μg of uracil-containing ssDNA (U) or *in vitro* synthesised heteroduplex DNA (Chapter 2) was transfected into competent *E.coli* JM101 or BW313 cells (Chapter 2). Plaques were counted after an o/n incubation at 37°C.
ssDNA: Single stranded DNA

H. DNA: Heteroduplex DNA

this gene should protect an expected 162nt of the 464nt probe from S−1 nuclease digestion. This should be readily distinguishable from the 206nt probe fragment protected by transcripts from the MuIFNα_1 promoter upstream of the intact *cat* gene.

7.4 Analysis of MuIFNα_1 Mutant Promoter and Pseudogene Function by S−I Nuclease Mapping

To demonstrate the practicality of analysis of IFN gene regulation using the S−1 nuclease protection assay, L−929 cells were co−transfected with pairs of genes using the selectable plasmid pTCF (Chapters 5 and 6). These comprised a mutant MuIFNα_1 promoter/wildtype *cat* test gene and the wildtype MuIFNα_1 promoter/mutant *cat* pseudogene. The cell populations generated are indicated in Table 7.3.

The cell populations were examined to confirm both the presence and relative abundance of the test gene and pseudogene sequences. For these purposes Southern blot analysis of *Pvu*II digested DNA, with the 1,163bp *Sst*I/*Hin*dIII fragment MuIFNα_1/*cat* gene probe (Chapters 5 and 6) was performed. As both chimaeric genes are flanked by *Pvu*II sites within pUC13 sequences, but the pseudogene has lost the internal *cat* gene *Pvu*II site (Figures 5.2 and 7.2), distinct *Pvu*II hybridising fragments of 1299bp (pseudogene) and 945bp plus 497bp (test gene) were expected. Some results obtained from this analysis are shown in Figure 7.6. Subsequent analysis demonstrated, unfortunately, a *Pvu*II hybridising fragment within L−929 DNA (see also Chapters 5 and 6) which comigrated with the expected 1299bp *Pvu*II fragment of the pseudogenes. Thus, although this analysis confirmed the presence of the test gene, neither the presence nor relative abundance of the pseudogenes could be determined.

The activity of the test genes were also examined in these cells prior to S−1 nuclease protection studies. It was assumed and later confirmed that the pseudogene would not produce functional CAT protein due to the internal deletion. Thus, CAT assays were routinely performed to monitor test gene promoter activity.

TABLE 7.3 Cell Populations Generated for S-1 Nuclease Protection Studies

Test Gene (10μg)	Pseudogene (10μg)	pTCF (1μg)	Cell Population
pCP4	pCPWTΔ	+	L929ψpCP4
pCP5	"	"	L929ψpCP5
$pCP7^8$	"	"	$L929\psi pCP7^8$
pCP8	"	"	L929ψpCP8
pCP1	"	"	L929ψpCP1
pCP6	"	"	L929ψpCP6
pCP6t	"	"	L929ψpCP6t
pCP6tR	"	"	L929ψpCP6tR
$pCP8^5$	"	"	$L929\psi pCP8^5$
$pCP8^6$	"	"	$L929\psi pCP8^6$
$pCP8^{5/9}$	"	"	$L929\psi pCP8^{5/9}$
pCEAPWT	"	"	L929ψpCEAPWT
pCEBPWT	"	"	L929ψpCEBPWT
pCEAP3Δ77	"	"	L929ψpCEAP3Δ77
pCEBP3Δ77	"	"	L929ψpCEBP3Δ77

An average of 10 to 20 colonies were obtained by co-transfection of L-929 cells with pTCF (Chapter 2). After 18 to 24 days selection in G418 (400μg/ml) the colonies of independent transfections were pooled and the cells grown as populations for further analysis.

FIGURE 7.6 Southern Blot Analysis of Stably Transfected L929 Cells Containing both Pseudogene and Test Gene Sequences

See the legend to Fig.5.9. Southern blot analyses of *Pvu*II digested genomic DNAs were performed with the same probe described in Fig.5.9. Autoradiography was performed for 24hrs (a) or 48hrs (b) at -70°C using Fuji RX film with an intensifying screen. The arrows indicate the 1299bp (pseudogene) and 945bp (test gene) fragments. Later studies revealed a 1300bp *Pvu*II hybridising fragment in L-929 cells which comigrated with the 1299bp fragment. The additional bands seen are probably endogenous MuIFNα gene sequences which hybridise to the MuIFNα_1 promoter sequences within the probe.

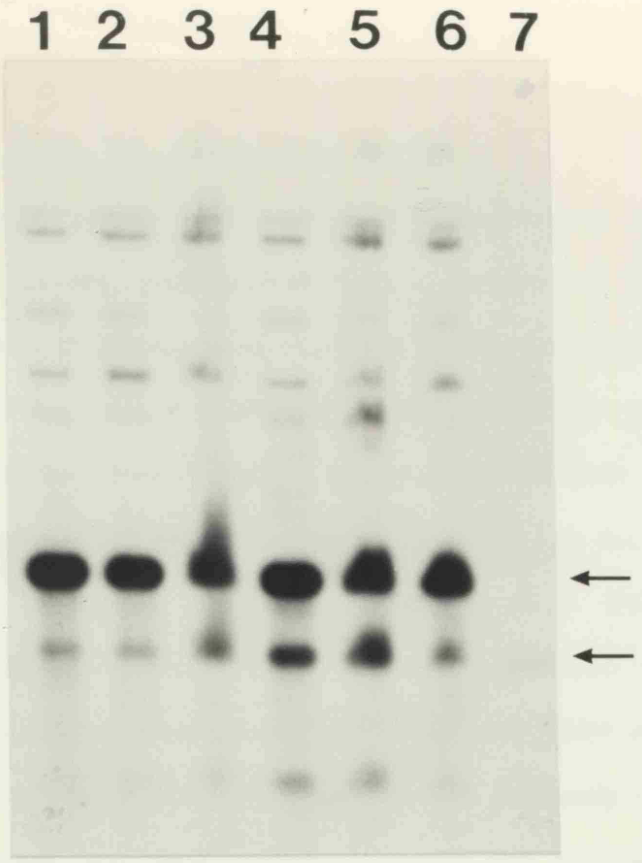
a:

1) 10μg of *Pvu*II digested L929ψpCP5 DNA
2) " " " L929ψpCP6 "
3) " " " L929ψpCP7[8] "
4) " " " L929ψpCP8 "
5) " " " L929ψpCP4 "
6) " " " L929ψpCP2 "
7) 10pg " " pCPWT DNA

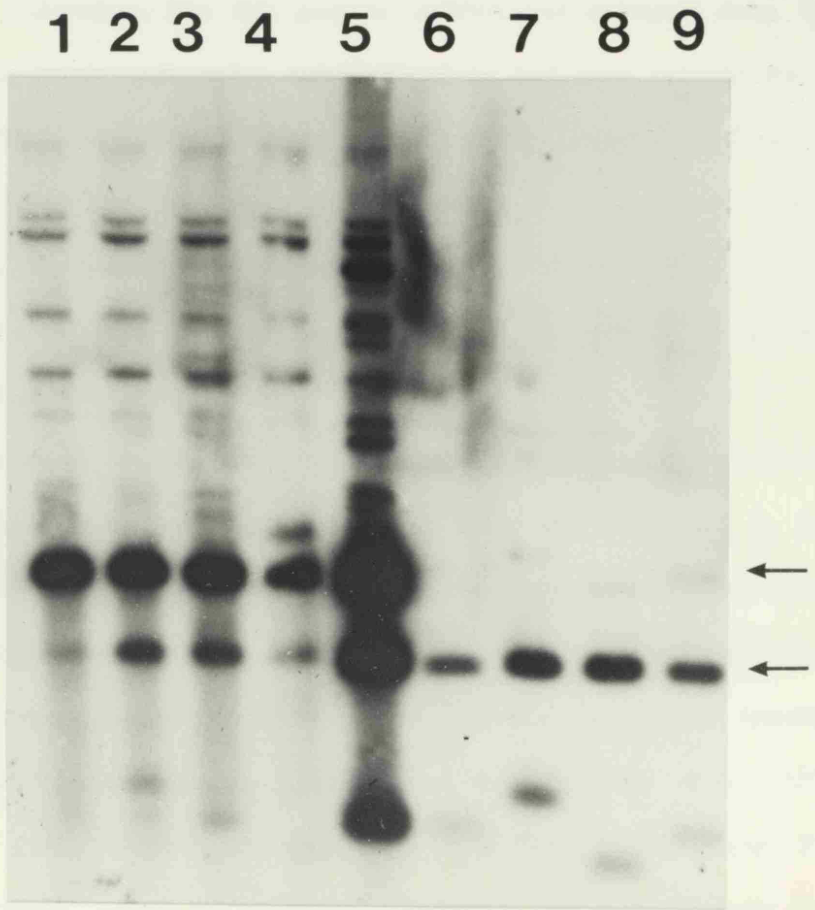
b:

1) 10μg of *Pvu*II digested L929ψpCEAPWT DNA
2) " " " L929ψpCEBPWT "
3) " " " L929ψpCP3Δ77 "
4) " " " L929ψpCEAP3Δ77"
5) " " " L929ψpCEBP3Δ77"
6) 10pg " " pCEAPWT "
7) " " " pCEBPWT "
8) " " " pCEAP3Δ77 "
9) " " " pCEBP3Δ77 "

a



b



Functional Analysis of the Region III_2 Promoters P4 and P5

Having established these criteria S−1 nuclease protection studies were undertaken. The behaviour of some mutant promoters previously characterised by CAT assays (Chapter 6) and some additional mutants were examined by this procedure.

The region III_2 MuIFNα_1 mutant promoters P4 and P5 contain mutations creating either a partial or complete hexameric GGAGAA sequence motif located at −52 to −47 (Chapter 6; Fig.6.3). Although CAT production studies suggested that the function of the MuIFNα_1 promoter was unaffected by these mutations their activity has been reinvestigated using an alternative assay. Thus, the cell populations L929ψpCP4 and L929ψpCP5 (Table 7.3), containing transfected pCPWTΔ (ψ) and pCP4 or pCP5 plasmid DNAs, were generated. CAT assays, performed as described previously (Chapters 5 and 6), demonstrated as expected that these mutant promoters are regulated by polyrI.rC (Fig.7.7 Lanes a,b,c and d).

To investigate gene transcription S−1 protection studies with the 464nt probe (Section 7.2) were undertaken. For this purpose, mRNA was extracted from L929ψpCP4 and L929ψpCP5 cells which were either non−induced or had been induced for 9hrs with polyrI.rC as described previously (Section 7.2). Correctly initiated inducer dependent transcripts derived from the pseudogene were observed in both cases, protecting the anticipated 162nt of the probe from degradation by S−1 digestion (Fig.7.8). Thus, this demonstrated that the wildtype promoter coupled to the CAT pseudogene had retained its previously characterised functional activity.

The identification of an additional fragment of 206nt protected from S−1 nuclease digestion by correctly initiated transcripts, initiated from the test genes in these cells, confirmed that regulated transcription of two independent genes functioning simultaneously within a single population can be demonstrated by this analysis (Fig.7.8). Scanning densitometry of autoradiographs was performed to examine the relative intensity of the 206nt and 162nt protected fragments. This enabled a comparison of the efficiency of polyrI.rC induced gene transcription from the MuIFNα_1 wildtype and P4 or P5 promoters. These results, shown in Table 7.4, confirmed that induced gene transcription from these promoters is indistinguishable in L−929 cells.

FIGURE 7.7 CAT Production in L929ψpCP4 and L929ψpCP5 Cells

See legend to Fig.5.10, the only difference being that 10μl of cell extracts were assayed for 1hr. Autoradiography was performed for 12hrs using Fuji RX film (a). The structure of the MuIFNα_1/*cat* gene is shown below (b).

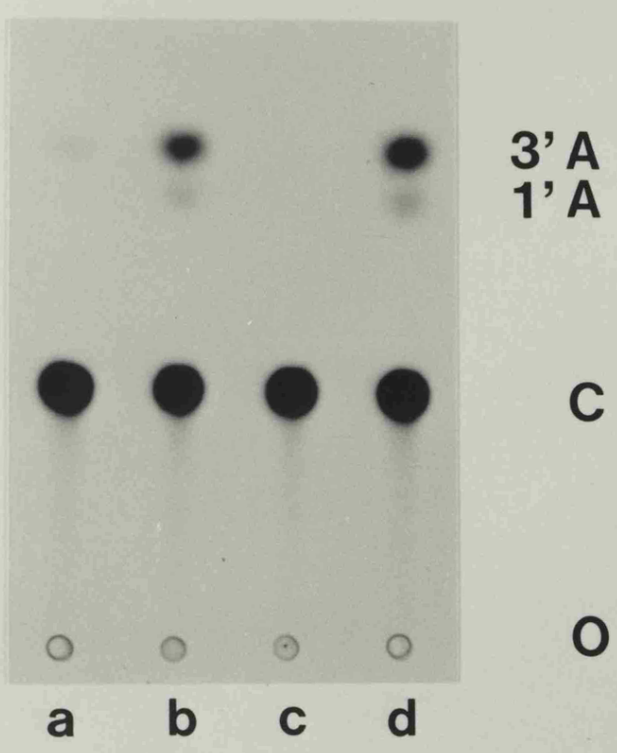
a) L929ψpCP4 cells, non-induced (-)

b) " " induced (+)

c) L929ψpCP5 " -

d) " " +

A



B

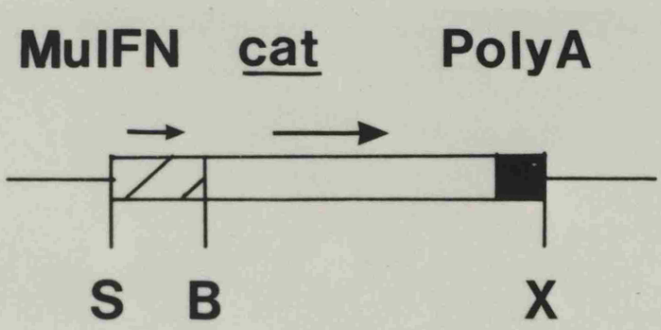
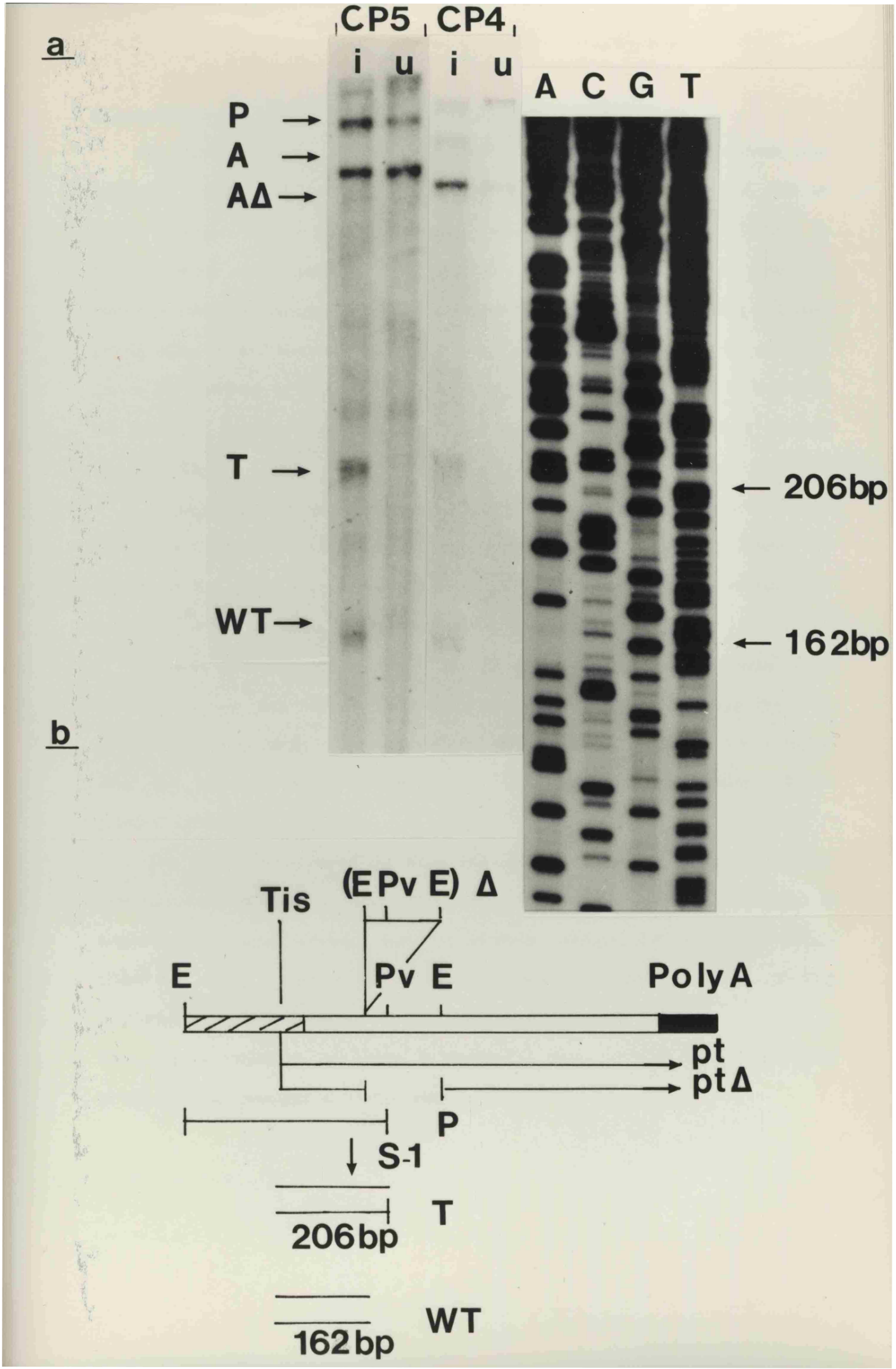


FIGURE 7.8 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCP4 and Lp29ψpCP5 Cells

See legend to Fig.7.3, except that 50μg of total RNA from induced (i) or non-induced (u) cells were used per assay. Autoradiography was performed for 2 weeks at -70°C using Kodak XAR-5 film with an intensifying screen (a).

CP4 = L929ψpCP4 cells, CP5 = L929ψpCP5 cells, P = probe, A = probe protected by readthrough transcripts of the test gene, AΔ = probe protected by readthrough transcripts of the pseudogene, T = 206nt of the probe protected by correctly initiated transcripts of the test gene and WT = 162nt of probe protected by correctly initiated transcripts of the pseudogene. Lanes are from independent gels and hence arrows do not align perfectly to the corresponding probe fragments in all cases.

Below (b) the principle of this assay is illustrated schematically. This shows two distinct transcripts derived from the test gene (pt) and pseudogene (ptΔ). The latter contains the deletion Δ shown. Thus, they are both complementary to the same probe but protect distinct sized species from S-1 digestion.


a
CP5
CP4
i
u
i
u
A C G T
P
A
AΔ
T
WT
206bp
162bp
b
Tis
(E Pv E) Δ
E
Pv E
PolyA
pt
pt Δ
P
S-1
T
206bp
WT
162bp

Re−examination of $P7^8$ and P8 Region IV Promoter Mutant Activity

The mutant promoters $P7^8$ and P8, previously described in Chapter 6, contain base substitutions within 10bp inverted repeat sequences located −41 to −32 and −66 to −57nt upstream of the MuIFNα_1 promoter transcription initiation site (Chapter 6; Fig.6.3). Initial examination of CAT production had suggested that these promoter mutations caused an apparent reduction and potentiation of induced MuIFNα_1 promoter activity (Chapter 6). However, these preliminary results were not statistically significant due to the variability of CAT production observed. It was therefore decided to investigate the behaviour of these promoters using the prospectively more sensitive S−1 nuclease protection assay.

To investigate $P7^8$ and P8 promoter activity of these mutant promoters the cell populations L929ψpCP7^8 and L929ψpCP8 were generated (Table 7.3). These contain the pseudogene and either the plasmid pCP7^8 or pCP8. CAT assays again confirmed that functionally active test genes had been introduced into these cells (Fig.7.9).

A direct comparison of the relative strength of polyrI.rC induced gene transcription from the mutant and wildtype MuIFNα_1 promoters was undertaken using the S−1 nuclease protection assay. Correctly initiated gene transcripts protecting the expected 206nt and 162nt of the probe from S−1 nuclease digestion were identified only in polyrI.rC induced cells (Fig.7.10).

The relative intensities of the 206nt and 162nt S−1 nuclease protected fragments were quantitated by scanning densitometry. This analysis established that the *cat* gene is transcribed with equal efficiency from the MuIFNα_1 wildtype, $P7^8$ and P8 promoters (Table 7.4) adding support to the conclusions derived by the statistical analysis of CAT production in polyrI.rC induced L−929 cells containing these constructs (Chapter 6). Thus, these mutations do not appear to significantly affect polyrI.rC mediated activation of the MuIFNα_1 promoter in L−929 cells.

TABLE 7.4 Relative Abundance of PolyrI.rC Induced MuIFNα_1 Wildtype and Mutant Promoter Transcripts in L-929 Cells.

Population	% T	% WT	Corrected %WT	Ratio T/WT
L929ψpCP4	60	40	55	1
L929ψpCP5	60	40	55	1
L929ψpCP7[8]	60	40	55	1
L929ψpCP8	59	41	57	1
L929ψpCP1	38	62	87	0.4
L929ψpCP6	79	21	29	2.7
L929ψpCP6t	83	17	24	3.0
L929ψpCP6tR	70	30	42	1.6
L929ψpCP8[5]	20	80	112	0.2
L929ψpCP8[6]	55	45	63	1
L929ψpCP8[5/9]	34	66	92	0.4
L929ψpCEAPWT	25	75	105	0.2
L929ψpCEBPWT	48	34	59	1.0
L929ψp3Δ77	-	100	-	-
L929ψpCP3Δ77R	-	100	-	-
L929ψpCEAP3Δ77	57	43	60	1.0

The relative abundance of the 206nt and 162nt protected probe fragments were determined by scanning densitometry of autoradiographs. Values have been corrected to account for the higher radioactivity associated with the larger 206nt protected fragment by multiplying %WT by 1.4.

FIGURE 7.9 CAT Production in L929ψpCP7[8] and L929ψpCP8 Cells

See legend to Fig.7.7.

Autoradiograph (A)

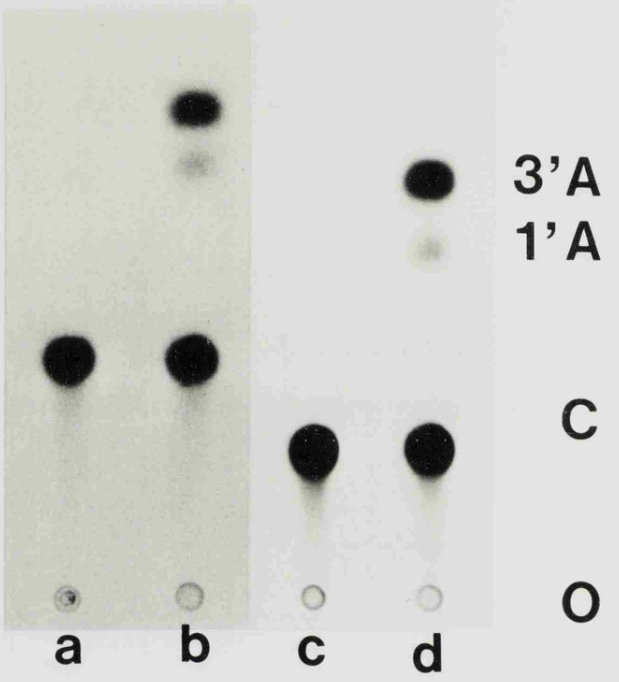
a) L929ψpCP7[8] -

b) " +

c) L929ψpCP8 -

d) " +

a



b

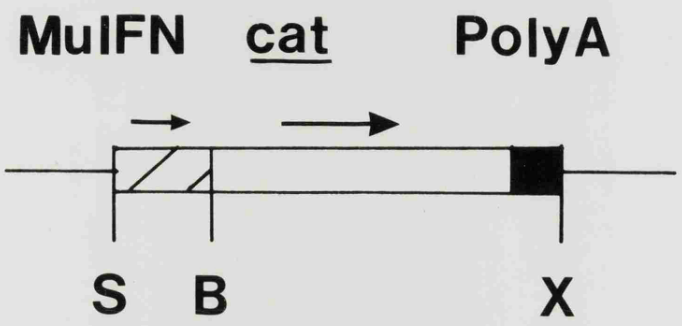


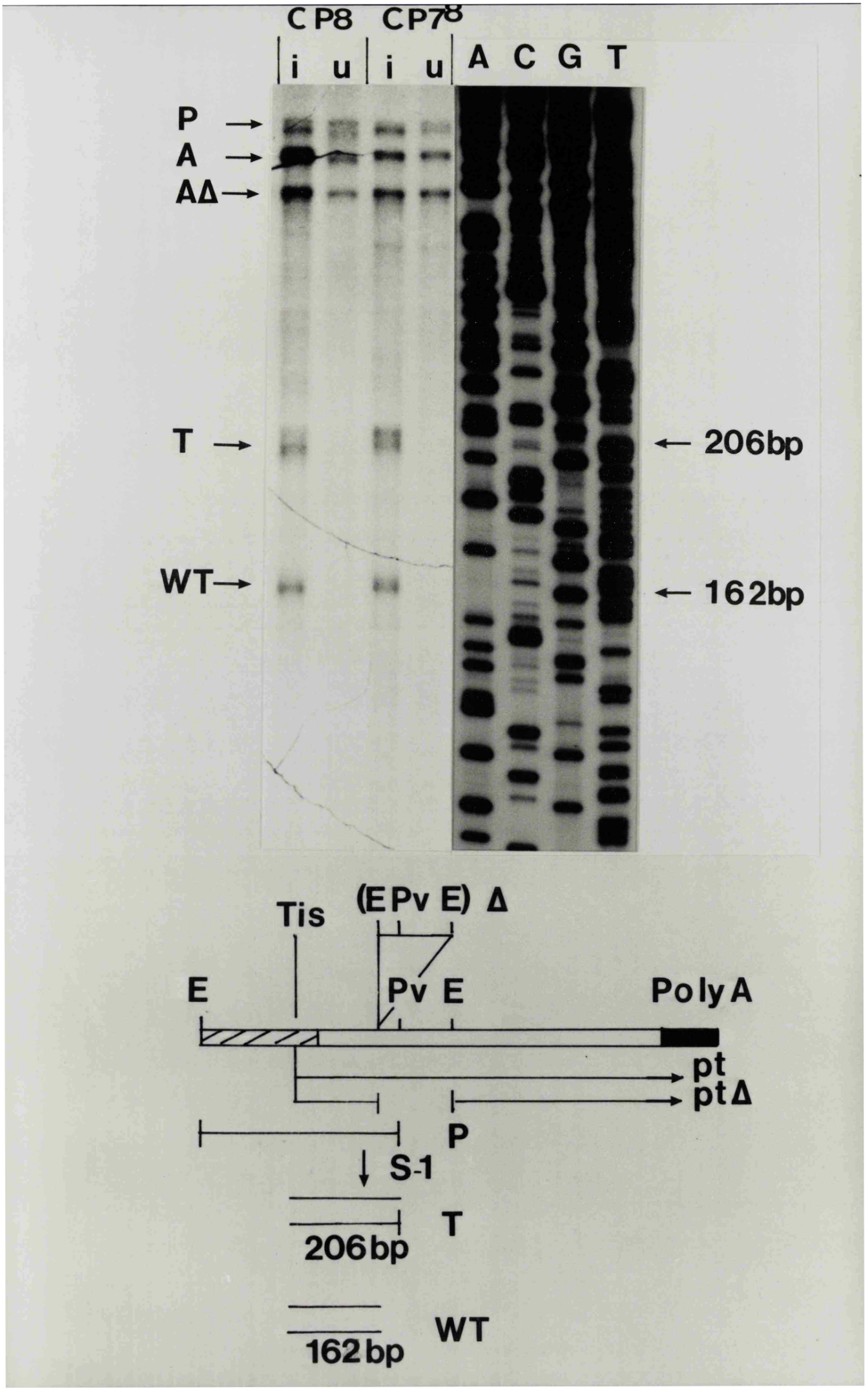
FIGURE 7.10 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCP7^{8} and L929ψpCP8 Cells

See legend to Fig.7.8.

Autoradiograph

CP7^{8} = L929ψpCP7^{8}

CP8 = L929ψpCP8


CP8
CP78
i
u
i
u
A C G T
P
A
AΔ
T
WT
206bp
162bp
Tis
(EPv E) Δ
E
Pv E
PolyA
pt
ptΔ
P
S-1
T
206bp
WT
162bp

An A to G Transition at −87 Reduces PolyrI.rC Induction of the MuIFNα_1 Promoter by 2 to 3 Fold

The MuIFNα_1 promoter mutants P2 and P3 contain a common A to G transition located at −87, created by introducing a region III hexanucleotide GGAGAA sequence motif by site directed mutagenesis, at a new site within region II (Chapter 6; Fig.6.3). CAT assays had previously indicated a statistically significant reduction of inducible CAT production from these promoters (Chapter 6). It was concluded that the −87 transition was responsible. To confirm these observations a corresponding reduction of correctly initiated gene transcripts should be seen. Therefore, the cell population L929ψpCP2 was generated (Table 7.3). These cells contained both the pseudogene and the plasmid pCP2, which carries the P2 promoter.

Initially CAT assays were performed to confirm the expected functional activity of the P2 promoter (Fig.7.11). The relative strength of the P2 and wildtype promoters was then investigated by employing the S−1 nuclease assay. Correctly initiated transcripts derived from the mutant and wildtype promoters were observed (Fig.7.12) in polyrI.rC induced cells only. However, direct comparison of the intensity of the protected fragments by scanning densitometry confirmed the previous conclusion that the −87 mutation reduced transcription of correctly initiated transcripts from the MuIFNα_1 promoter 2−3 fold in polyrI.rC induced L−929 cells (Table 7.4).

Thus a single base substitution has revealed *cis* regulatory sequences within the MuIFNα_1 promoter which are worthy of further investigation to determine both the extent and contribution of this region to IFN gene regulation. These data in conjunction with previous observations (Chapter 6) strongly implicate the interspecies conserved sequences identified as region II (Chapter 6; Fig.6.3) as being satisfactory in polyrI.rC mediated gene activation. As discussed previously a homologous region has been implicated in the virus induced activation of HuIFNα_1 promoter function, suggesting a common mechanism for the two inducers or at least an interaction with common parts of the gene.

FIGURE 7.11 CAT Production in L929ψpCP2 Cells

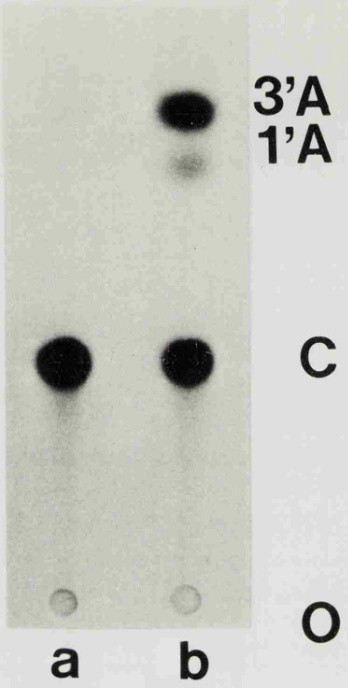
See legend to Fig.7.7.

Autoradiograph (a)

a) L929ψpCP2 -

b) " +

a



b

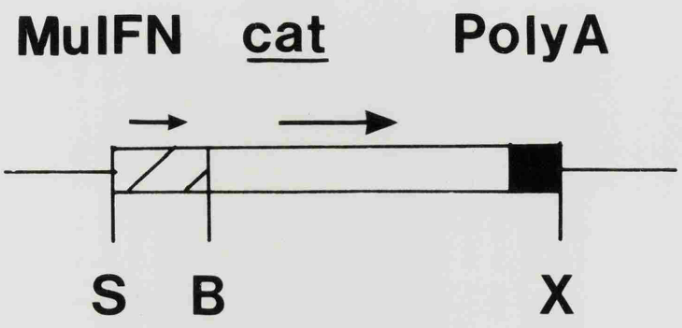
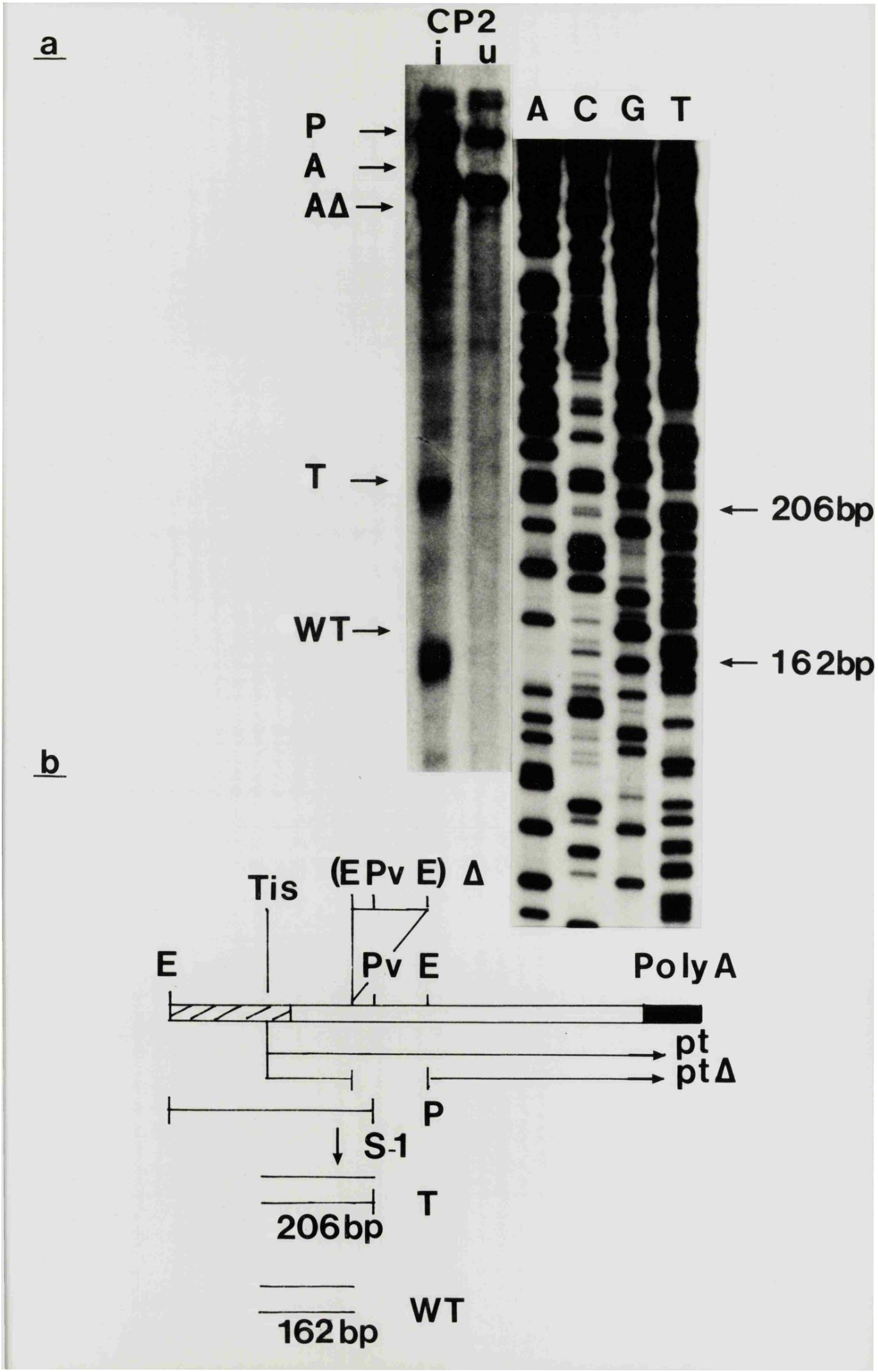


FIGURE 7.12 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCP2 Cells

See legend to Fig.7.8.

Autoradiograph (a)

CP2 = L929ψpCP2 Cells


a
CP2
i u
A C G T
P
A
AΔ
T
WT
206bp
162bp
b
Tis
(E Pv E) Δ
E
Pv E
Poly A
pt
pt Δ
P
S-1
T
206bp
WT
162bp

Transcription is Elevated from MuIFNα_1 Promoters Containing the Region III_1 Sequence Motif GGAGAA

In Chapter 6 the construction and functional analysis of two MuIFNα_1 promoters, P6 and P6t, containing a G to A transition at +14 to create the hexanucleotide motif GGAGAA at region III_1, was described (Fig.6.3). One of these promoters, P6t, contained an additional base substitution at −8, identified during the isolation of P6, which apparently inactivated the MuIFNα_1 promoter. To ascertain if the −8 mutation was solely responsible for the loss of promoter function a revertant containing only this mutation was isolated by site directed mutagenesis and its activity analysed in L−929 cells. In contrast to P6t the promoter P6 appeared to elevate polyrI.rC induced CAT production relative to the wildtype promoter in L−929 cells. The location of the mutation within the untranslated leader sequence of the MuIFNα_1 promoter/*cat* gene suggests this region could function at the transcriptional or post−transcriptional level. To investigate these possibilities S−1 nuclease mapping studies were undertaken.

The apparent loss of promoter function observed with P6t was investigated further by isolating a revertant promoter containing only the −8 mutation (Fig.7.13). This was achieved by the use of oligonucleotide 11 (Table 7.1), designed to recreate the wildtype MuIFNα_1 promoter sequence at +14, whilst leaving the −8 mutation of P6t intact. Site directed mutagenesis was performed according to the method of Kunkel (Chapter 2) using a uracil containing DNA template prepared from the M13 recombinant phage containing the corresponding promoter (Chapter 6). The mutant phage were selected on *E.coli* JM101 (Table 7.5) and two prospective mutants analysed by sequencing using the universal and oligonucleotide 5 primers (Chapter 6). This confirmed that both contained the desired reversion while the remaining sequences were unchanged (Fig.7.13). Subsequently this promoter was subcloned into p22 as described previously (Chapters 5 and 6) to create pCP6$_t$R.

The expression of the region III_1 mutant promoters P6, P6t and P6tR were compared directly with the wildtype MuIFNα_1 promoter by generating the cell populations L−929ψpCP6, L929ψpCP6t and L929ψpCP6tR (Table 7.3). In addition to the corresponding P6 promoters, these cell populations contained the pseudogene.

```
                          -180       -170       -160       -150       -140
MuIFNα1          :    AGCTTTTGATGAGGACCAGTGAAAGAGGAAGCAATAATGAAAACCACAATGC
P6tR                  ----------------------------------------------------
P8^5                  ----------------------------------------------------
P8^6                  ----------------------------------------------------
P8^5/9                ----------------------------------------------------
P3Δ77R                ----------------

     -130       -120       -110       -100       -90        -80        -70
TTTAGAAAACACCCAGACGCAAGCAGAGAATGAGTTAAAGAAAGTGAAAAGACAAGTGGAAAGTGATGGAA
------------------------------------------------------------------------
------------------------------------------------------------------------
------------------------------------------------------------------------
------------------------GGAGAA------------------------------------------
                                  ----------------------------

                                  TATA BOX                          CAP
    -60        -50        -40        -30        -20        -10
GGGCATTCAGAAAGTAAAAACCAGTGTTTGCCCTATTTAAGACACATTCACCCAGGATGGTCTTCAGAGAA
-----------------------------------------------------------C------------
GCGG---------GGAGAA-----------------------------------------------------
GCGG-----------------------------------------------------------------
GCGG---------GGAGAA-----------------------------------------------------
------------------------------------------------------------------------

   +10        +20        +30        +40        +50
CCTAGAGGGGAAGGATCAGGACCAAACAGTCCAGAAGACCAGAAGCT
-----------------------------------------------
------GGAGAA----------------------------------
-----------------------------------------------
-----------------------------------------------
```

FIGURE 7.13 Primary Nucleotide Sequence of MuIFNα_1 and Mutant Promoters

See legend to Fig.6.3.

TABLE 7.5 Site Directed Mutagenesis by the Method of Kunkel (1985)

Promoter	Template	Oligonucleotide	BW313	JM101	Sequencing	Cloned
P6tR	U-mP6t	-	>1000	1		
	H-mP6t	11	N/D	890	2/2	2/5
$P6^{5/9}$	U-mP8[5]	-	>1000	1		
	H-mP8[5]	9	N/D	700	2/2	1/5
P3Δ77R	U-mP3Δ77	-	>1000	1		
	H-mP3Δ77	10	N/D	23	2/2	2/4

0.2 μg of uracil containing ssDNA (U) or *in vitro* synthesised heteroduplex DNA (Chapter 2) was transfected into *E.coli* JM101 or BW313. *E.coli* JM101 selects against the parental U-DNA strand of the heteroduplex, resulting in an enrichment for the mutant phage. The number of plaques obtained by transfection of BW313 and JM101 cells is shown. Random plaques selected on *E.coli* JM101 were analysed by sequencing and those containing the desired mutation cloned into p22. The ratio of mutants identified to the total analysed by sequencing is shown. Similarly the ratio of recombinant plasmids identified by restriction analysis to those screened is also indicated.

H-: Heteroduplex DNA synthesised *in vitro* from the U-DNA template and oligonucleotide indicated

Preliminary CAT assays (Fig.7.14) established that all except L929ψpCP6t cells contained functional test genes. As observed previously (Chapter 6) no detectable CAT production was observed in L−929 cells containing the P6t promoter coupled to a *cat* gene (Fig.7.14 Lanes c and d). Surprisingly however, the removal of the +14 mutation of $P6_t$ to create $P6_tR$ (Fig.7.13) apparently restored MuIFNα_1 promoter activity (Fig.7.14 Lanes e and f). Thus, these results failed to confirm the previous conclusion (Chapter 6) that a G to C transition at −8 inactivated the MuIFNα_1 promoter.

To investigate this observation further, the relative strength of transcription from the wildtype, P6, P6t and P6tR promoters has been compared by S−1 nuclease protection assays. Surprisingly, correctly initiated transcripts derived from all three P6 test promoters were observed in polyrI.rC induced cells (Fig.7.15).

These data demonstrated that the MuIFNα_1 promoter mutant P6t containing both the −8 and +14 mutations (Fig.6.3) produces correctly initiated gene transcripts but no CAT production in induced L−929 cells. Both the P6 and P6tR promoters containing the +12 or −8 mutations, respectively (Fig.6.3; Fig.7.13), transcribe and translate *cat* gene sequences producing a functional CAT protein. The most likely explanation for these observations therefore, is that the P6t promoter/*cat* gene transcripts are not translated to give a functional protein.

To investigate the effect of the +14 mutation of promoters P6 and P6t upon MuIFNα_1 promoter activity, the relative intensity of the 206nt and 162nt S−1 nuclease protected fragments (Fig.7.15) were compared by scanning densitometry. This reveals an approximate three fold elevation of induced transcription from the MuIFNα_1 promoter containing the GGAGAA motif at +12 to +17 (Table 7.4). Thus, this reinforces the previous conclusions from CAT production studies (Chapter 6) that this mutation enhances polyrI.rC mediated activation of the MuIFNα_1 promoter.

Therefore, the transition at +14, located within the transcribed region, defines a second site distinct from the region II −87 mutation described previously, that influences regulated MuIFNα_1 promoter activity. Its location indicates it could either increase the efficiency of transcription or mRNA stability of the chimaeric MuIFNα_1/*cat* gene. Further investigation by nuclear run off transcription assays could resolve this issue.

FIGURE 7.14 CAT Production in L929ψpCP6, L929ψpCP6$_t$ and L929ψpCP6$_t$R Cells

See legend to Fig.7.7.

Autoradiograph (a)

a) L929ψpCP6 Cells -

b) " +

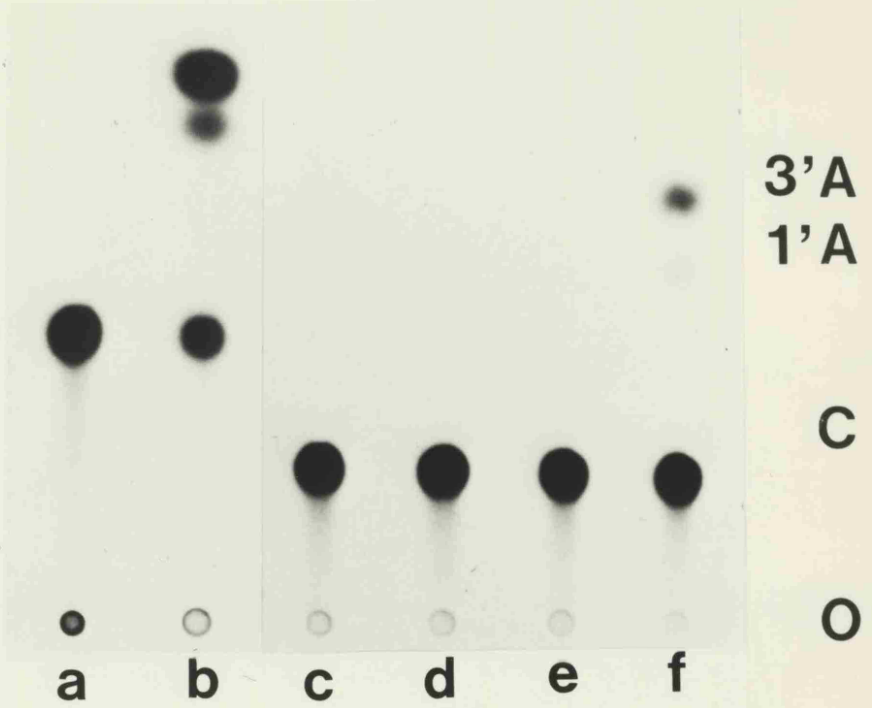
c) L929ψpCP6$_t$ Cells -

d) " +

e) L929ψpCP6$_t$R Cells -

f) " +

a



b

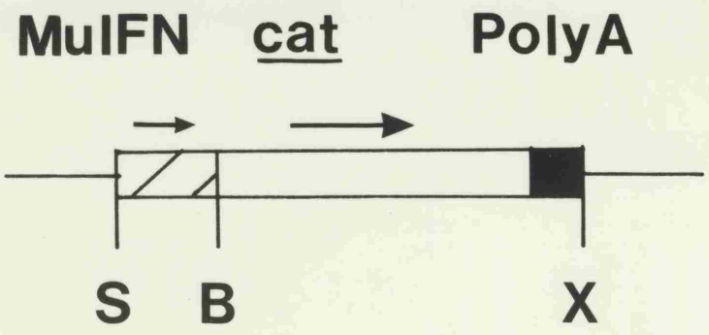


FIGURE 7.15 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCP6, L929ψpCP6$_t$ and L929ψpCP6$_t$R Cells

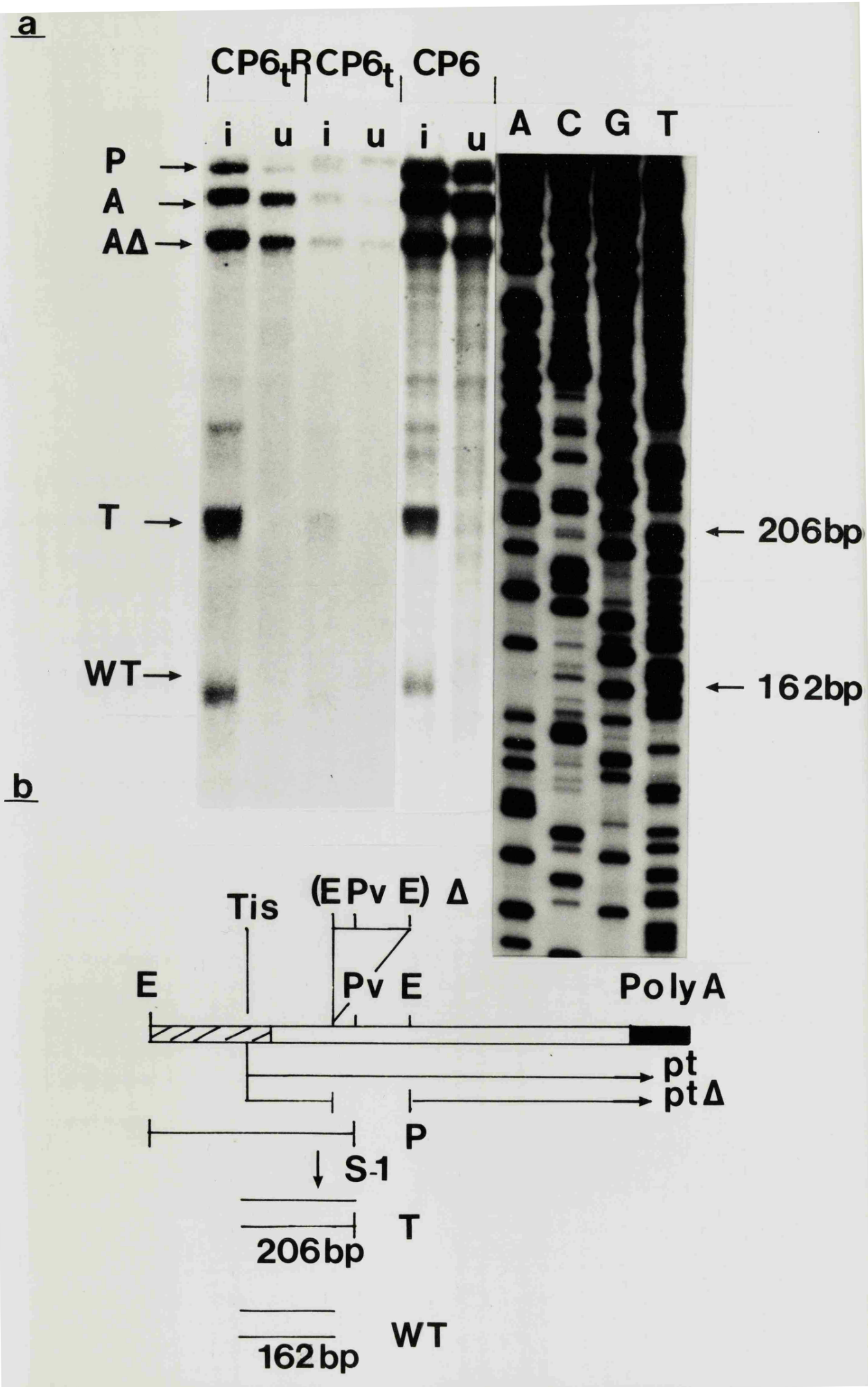
See legend to Fig.7.8.

Autoradiograph (a)

CP6 = L929ψpCP6 Cells

CP6$_t$ = L929ψpCP6$_t$ Cells

CP6$_t$R = L929ψpCP6$_t$R Cells


a
CP6t
R
CP6t
CP6
i u i u i u
A C G T
P
A
AΔ
T
WT
206bp
162bp
b
Tis
(E Pv E) Δ
E
Pv E
PolyA
pt
pt Δ
P
S-1
T
206bp
WT
162bp

7.5 Functional Analysis of MuIFNα_1 Promoters Containing Multiple Mutations

As described here and in Chapter 6, subtle changes in MuIFNα_1 promoter function have been observed with some of the first generation mutant promoters constructed on the basis of IFNα promoter comparative structural studies (Chapter 6). These results suggest that sequences contributing to promoter function may not be confined to a single location. A preliminary study has therefore been made to examine the behaviour of MuIFNα_1 promoters containing mutations at more than one site, by creating promoter mutants possessing a combination of selected mutations.

For this purpose the region III_1, region III_2 and region III_3 GGAGAA sequence motifs, described in Chapter 6, were introduced into the promoter mutant P8, which contains two substitutions within the inverted repeat located at -70 to -61 in the MuIFNα_1 promoter (Fig.6.3). This choice was based upon early indications that the promoter mutant P8 appeared to increase promoter activity. It should however be noted that subsequent studies suggested this promoter is indistinguishable from wildtype. The introduction of a GGAGAA sequence motif at region III_1 however has been demonstrated to increase MuIFNα_1 promoter activity 2 to 3 fold in polyrI.rC induced L$-$929 cells. To determine if the homologous sequence motifs of region III_2 and region III_3 could also influence expression the mutant promoters $P8^6$, $P8^5$, and $P8^{5/9}$ (Fig.7.13) were created by site directed mutagenesis.

The mutant promoters $P8^6$ and $P8^5$ were constructed using the M13 recombinant phage mP8 containing the P8 promoter (Chapter 6) and the oligonucleotides 5 and 6 (Chapter 6), by site directed mutagenesis as described by Zoller and Smith (Chapter 2). A mutation efficiency of 5% and 8% was obtained by this procedure for oligonucleotides 3 and 5 respectively (Table 7.6). Sequencing analysis of prospective mutants identified by oligonucleotide hybridisation confirmed the anticipated sequence of the desired mutant promoters $P8^6$ and $P8^5$ (Fig.7.13).

Mutant $P8^{5/9}$ was constructed by site directed mutagenesis using the method of Kunkel (1985) with oligonucleotide 9 (Table 7.1), designed to create an A to G transversion at -112, and a uracil containing M13 DNA template prepared from $mP8^5$, containing the corresponding mutant promoter. Mutant phage were selected on JM101

(Table 7.5) and sequence analysis identified two phage containing the desired mutations, confirming the structure shown in Figure 7.13.

The mutant *SstI/Bam*HI promoter modules were excised from their respective M13 recombinant phage and cloned into p22 as described previously (Chapters 5 and 6; Tables 7.5 and 7.6). Having obtained the mutant promoter/*cat* gene plasmids pCP8^{6}, pCP8^{5} and pCP8$^{5/9}$, their behaviour was investigated by generating the cell populations L929ψpCP8^{6}, L929ψpCP8^{5} and L929ψpCP8$^{5/9}$ (Table 7.1). These cells contain the pseudogene in addition to their respective test gene plasmids. Once again preliminary CAT assays confirmed they each contained functional test genes (Fig.7.16).

The relative transcription efficiencies of these promoters were examined by employing the S−1 nuclease transcription assay. In each case correctly initiated gene transcripts were identified in polyrI.rC induced cells only (Fig.7.17). Data obtained by scanning densitometry of this autoradiograph are summarised in Table 7.4. Surprisingly, this suggests P8^{5} and P8$^{5/9}$ result in a 3−5 fold reduction in gene expression from the MuIFNα_1 promoter. However the P8^{6} promoter behaves as wildtype.

Independently, neither the introduction of a GGAGAA sequence at −52 to −47 (region III$_2$) nor the substitutions within the inverted repeat located at −70 to −61 (region IV) affect MuIFNα_1 promoter activity (Chapter 6). The activity of a MuIFNα_1 promoter containing a GGAGAA sequence at −112 to −107 (region III$_3$) has not been investigated previously, however this sequence is present in the functionally active MuIFNα_5 and MuIFNα_{6A} promoters which suggests it is unlikely to contribute to the observed reduction in transcription of the P8$^{5/9}$ promoter. Thus the reduction of transcription from both the P8^{5} and P8$^{5/9}$ promoters is consistent with the combined region III$_2$ and region IV mutations being responsible. In this respect the inverted repeat sequences defined as region IV encompass the region III$_2$ sequence (Chapter 6; Fig.6.2).

7.6 Characterisation of the Transcription Initiation Site for the MuIFNα_1 Promoter in the Presence of an Upstream Enhancer Sequence

Previous studies have been described (Chapter 5) which demonstrated regulated MuIFNα_1 promoter function in the presence of the Mo−MuSV enhancer. Quantitative

TABLE 7.6 Site Directed Mutagenesis by the Method of Zoller and Smith (1983)

Promoter	Template	Oligonucleotide	% Mutants	Sequenced	Cloned
$P8^5$	mP8	5	8	2/3	1/4
$P8^6$	mP8	6	3	2/3	4/4

Templates were derived from recombinant M13 phages described in Chapter 6. 100 colonies for each mutagenesis were screened using 5' end-labelled oligonucleotide probes. The number of prospective positives identified by this screen are indicated above. See legend to Table 7.5.

FIGURE 7.16 CAT Production in L929ψpCP8^{5}, L929ψpCP8^{6} and L929ψpCP8$^{5/9}$ Cells

See legend to Fig.7.7.

Autoradiograph (A)

a) L929ψpCP8^{5} Cells -

b) " +

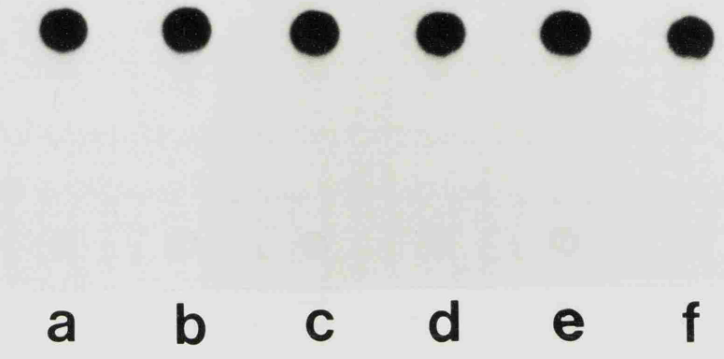
c) L929ψpCP8^{6} Cells -

d) " +

e) L929ψpCP8$^{5/9}$ Cells -

d) " +

A



B

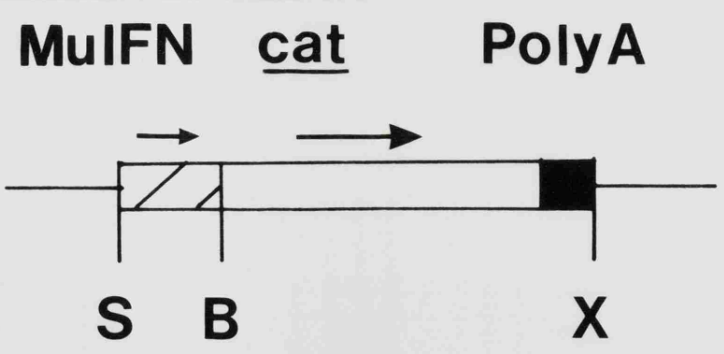


FIGURE 7.17 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCP8^5, L929ψpCP8^6 and L929ψpCP8$^{5/9}$ Cells

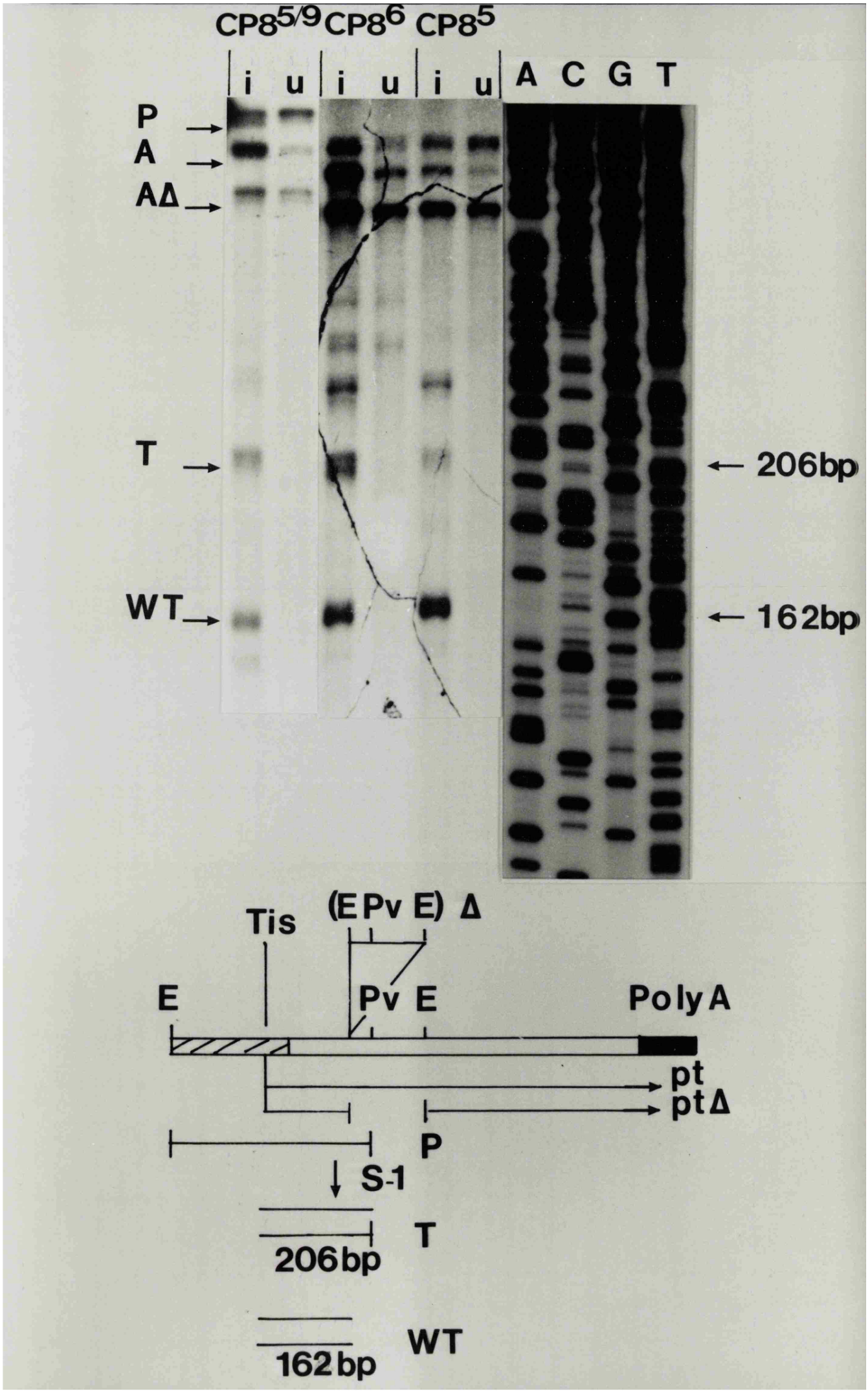
See legend to Fig.7.8.

Autoradiograph (A)

CP8^5 = L929ψpCP8^5 Cells

CP8^6 = L929ψpCP8^6 Cells

CP8$^{5/9}$ = L929ψpCP8$^{5/9}$ Cells


CP8^5/9
CP8^6
CP8^5
i u i u i u
A C G T
P
A
AΔ
T
WT
← 206bp
← 162bp
Tis
(E Pv E) Δ
E
Pv E
PolyA
pt
ptΔ
P
S-1
T
206bp
WT
162bp

CAT assays performed to compare the efficiency of non−induced and polyrI.rC induced gene transcription from the MuIFNα_1 promoter suggested that the enhancer had a 2 to 5 fold increase in basal gene expression, but no effect on induced promoter activity.

To establish whether the regulated expression observed arose as a result of transcripts from the MuIFNα_1 promoter, S−1 nuclease protection studies were performed. This enabled a simultaneous reinvestigation of the effect of the Mo−MuSV enhancer on MuIFNα_1 promoter function.

Thus, the cell populations L929ψpCEAPWT and L929ψpCEBPWT, containing the pseudogene and plasmids pCEAPWT or pCEBPWT respectively (Fig.5.14), were generated (Table 7.3). Preliminary CAT assays established, as expected (Chapter 5), that the test gene promoters were functional (Fig.7.18).

Regulated gene transcription was investigated, employing the 464nt probe used previously for S−1 nuclease protection assays. Correctly initiated gene transcripts of 206nt in polyrI.rC induced L929ψpCEAPWT and L929ψpCEBPWT cells demonstrated that the transcription initiation site of the MuIFNα_1 promoter is maintained in the presence of the Mo−MuSV enhancer (Fig.7.19). Correctly initiated gene transcripts were not identified in non−induced cells preventing further characterisation of the elevated basal gene expression previously observed.

A comparison of the intensity of the 206nt and 162nt protected fragments derived from the enhancer containing and wildtype MuIFNα_1 promoter transcripts respectively, by scanning densitometry, revealed a reduction in induced promoter activity in L929ψpCEAPWT cells, but comparable activity in L929ψpCEBPWT cells (Table 7.4). This is inconsistent with previous conclusions that polyrI.rC induced CAT production from the MuIFNα_1 promoter is unaffected by the Mo−MuSV enhancer independently of its orientation (Chapter 5). Therefore, the observation here probably reflects the cell population.

The MoMuSV Enhancer Restores MuIFNα_1 Promoter P3Δ77 Function

During the isolation of the MuIFNα_1 promoter mutants by site directed mutagenesis a 77bp deletion derivative, P3Δ77, arose (Chapter 6; Fig.6.3). CAT production from this

FIGURE 7.18 CAT Production in L929ψpCP3Δ77, L929ψpCEAPWT, L929ψpCEBPWT, L929ψpCEAP3Δ77 and L929ψpCEBP3Δ77 Cells

See legend to Fig.7.7.

Autoradiograph (A)

a) L929ψpCP3Δ77 -

b) " +

c) L929ψpCEAPWT -

d) " +

e) L929ψpCEBPWT -

f) " +

g) L929ψpCEAP3Δ77 -

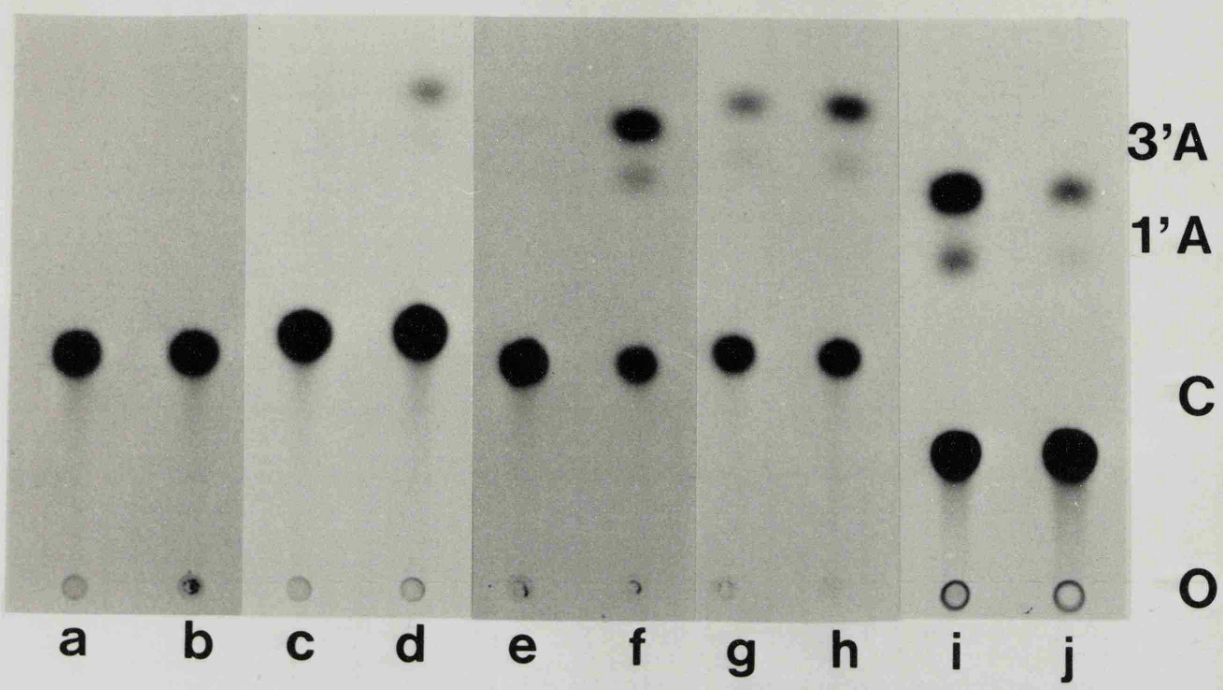
h) " +

i) L929ψpCEBP3Δ77 -

j) " +

Below (B) is illustrated the the structure of the MuIFNα_1/*cat* genes in these cells. The striped box = the MuIFNα_1 gene promoter, the open box = the *cat* gene, the blocked box = the HSV-2 polyA sequence and the stippled box = the Mo-MuSV enhancer shown in both orientations.

A


3'A
1'A
C
O
a
b
c
d
e
f
g
h
i
j

B

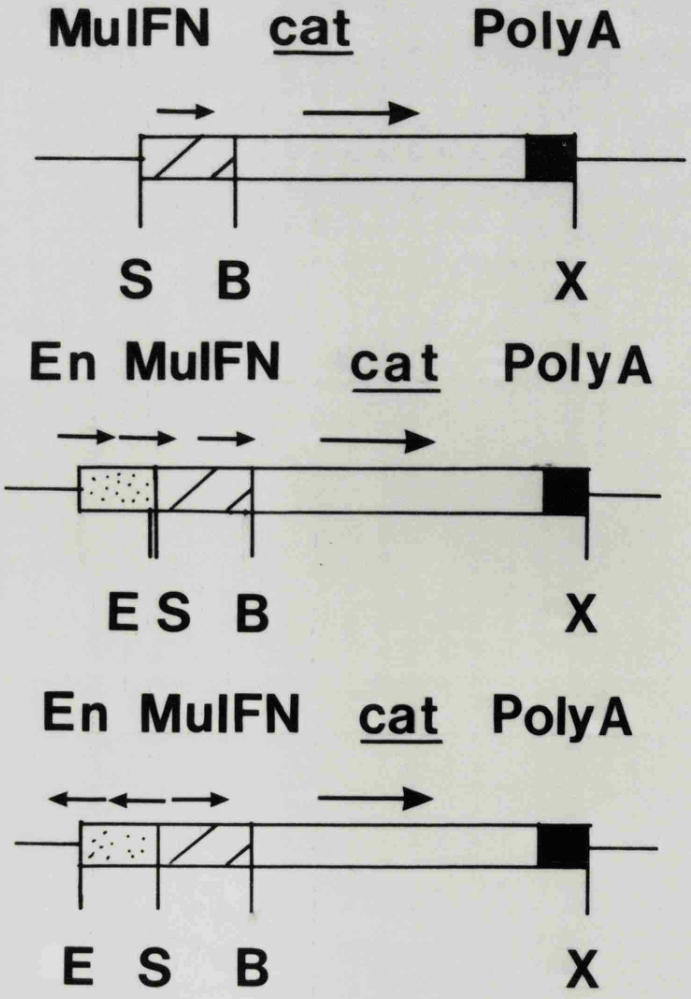

MuIFN cat PolyA
S B X
En MuIFN cat PolyA
E S B X
En MuIFN cat PolyA
E S B X

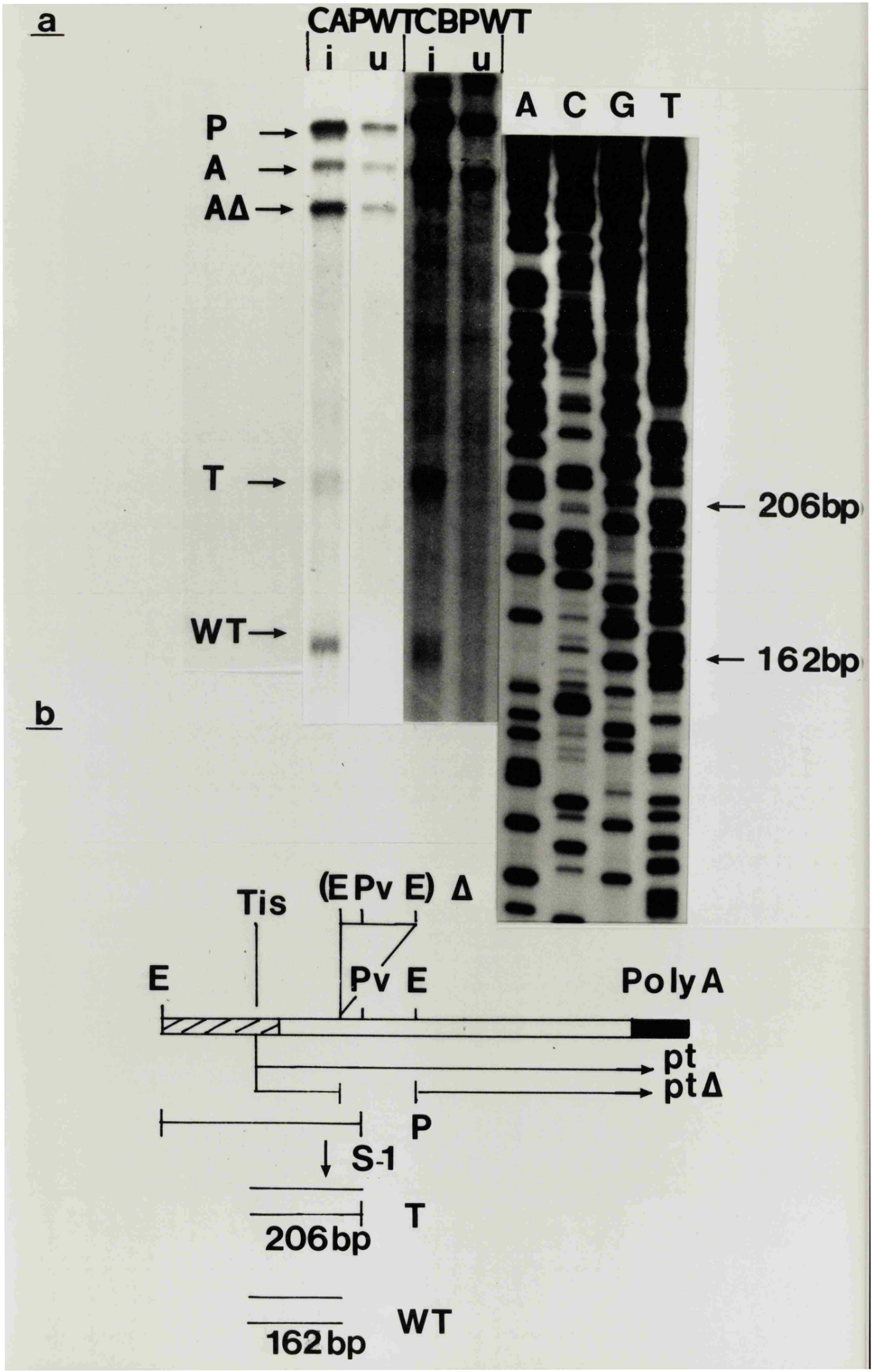
FIGURE 7.19 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCEAPWT and L929ψpCEBPWT Cells

See legend to Fig.7.8.

Autoradiograph (a).

CAPWT = L929ψpCEAPWT Cells

CBPWT = L929ψpCEBPWT Cells


a
CAPWTCBPWT
i u i u
A C G T
P
A
AΔ
T
WT
206bp
162bp
b
Tis
(E Pv E) Δ
E
Pv E
PolyA
pt
pt Δ
P
S-1
T
206bp
WT
162bp

promoter, linked to the *cat* gene, was not detectable in L−929 cells. To investigate the basis of the apparent loss of promoter function, further analysis of P3Δ77 was undertaken. P3Δ77 in addition to the 77bp deletion has base substitutions at −85 and −87, the latter having been shown previously to reduce induced MuIFNα_1 promoter activity 2−3 fold in L−929 cells. To confirm that the deletion was solely responsible for the behaviour of P3Δ77, a revertant was first created with wildtype MuIFNα_1 promoter sequences at −85 and −87 (Fig 7.13).

For this purpose site directed mutagenesis was performed by the method of Kunkel (1985) using the oligonucleotide 10 mismatch primer (Table 7.1) and a uracil containing M13 DNA template derived from mP3Δ77 (Chapter 6), containing the corresponding promoter. Using this procedure the promoter P3Δ77R was isolated (Table 7.5) possessing the expected primary sequence illustrated in Figure 7.13. This was cloned into p22 creating pCP3Δ77 as described previously (Chapters 5 and 6)

During these studies the opportunity arose to compare the relative efficiency of site directed mutagenesis by the method of Zoller and Smith (1983) with the modified procedure of Kunkel (1985), both adopted here. Comparison of the data shown in Tables 7.4 and 7.5 summarising the results obtained creating mutants by both procedures reveals the superior mutation efficiency obtained with the modified procedure. In combination with the site directed mutagenesis creating the *Eco*RI site in the construction of the pseudogene (Section 7.3), the average mutation efficiency observed with the modified procedure is 66%, relative to the 2−8% obtained with the original method employed earlier in these studies.

In addition to the creation of P3Δ77R two further derivatives were made. These were designed to establish whether upstream enhancer sequences could restore functional activity to the deleted promoter. For this purpose the P3Δ77 *Sst*I/*Bam*HI promoter module was excised from the corresponding M13 phage derivative (Chapter 6) and cloned into the *Sst*I/*Bam*HI site of pCEA.I and pCEB.I (Chapter 5) to create pCEAP3Δ77 and pCEBPΔ77. Thus these plasmids contain the P3Δ77 promoter with an upstream Mo−MuSV enhancer in both orientations, 5' to the *cat* gene.

These constructs were introduced into L−929 cells to create L−929ψpCP3Δ77,

L929ψpCP3Δ77R, L929ψpCEAP3Δ77 and L929ψpCEBP3Δ77 cell populations (Table 7.3) to assess promoter function. These cells contain the respective plasmids indicated in addition to the pseudogene.

An initial examination of promoter behaviour was undertaken by assessing CAT production in these cells. CAT production was not detected from either P3Δ77 or P3Δ77R promoters in L−929 cells (Fig.7.18 Lanes a and b). This confirmed that the loss of promoter activity, detected previously (Chapter 6), could indeed be attributed to the 77bp deletion which had originally arisen during site directed mutagenesis of the MuIFNα_1 promoter (Chapter 6).

Surprisingly, in the presence of an enhancer CAT production is restored from the 77bp deletion derivative of the MuIFNα_1 gene promoter in L929ψpCEAP3Δ77 and L929ψpCEBP3Δ77 cells (Fig.7.19 Lanes g,h,i and j). Futhermore, polyrI.rC regulated promoter function was revealed in L929ψpCPEAP3Δ77 cells (Fig.7.20 Lanes g and h). Therefore, this indicated that a transcriptional enhancer is able to reactivate an inactive MuIFNα_1 gene promoter containing only 94bp of DNA sequence upstream of its transcription initiation site.

The possibility remained that transcription was initiated from within enhancer sequences. Thus, to extend this analysis gene transcription from the mutant MuIFNα_1 promoters in these cells were investigated using the S−1 protection assay. In the absence of enhancer sequences, as expected, no P3Δ77 promoter specific transcripts could be detected in either non−induced or induced L929ψpCP3Δ77 cells. Correctly initiated transcripts were seen from the pseudogene in these cells, which protected 162nt of the probe from S−1 digestion (Fig.7.20). This confirmed that the intact MuIFNα_1 promoter of the pseudogene was functional in these cells, whereas the 77bp deletion derivative was not. This reinforced the previous conclusions (Chapter 6) that the transcriptional activation of the MuIFNα_1 gene promoter was lost upon deletion of the 77bp of sequence upstream of −94.

Next L929ψpCEAP3Δ77 and L929pCEBP3Δ77 cells were analysed using the transcription assay. In one orientation (pCEAP3Δ77), the presence of the Mo−MuSV enhancer immediately upstream of the P3Δ77 promoter produced inducer dependent

correctly initiated gene transcripts, which protected 206nt of the probe (Fig.7.20 Lanes c and d). This demonstrated that the enhancer had restored regulated promoter function to the deleted MuIFNα_1 promoter. A 162nt protected fragment was also observed in these cells, derived from the pseudogene (Fig.7.20 Lanes c and d). The intensity of the protected fragments were compared by scanning densitometry, indicating that both promoters transcribed *cat* with equal efficiencies in polyrI.rC induced L−929 cells (Table 7.4).

With the enhancer in the reverse orientation (pCEBP3Δ77) no correctly initiated transcripts were detected from either the pseudogene or test gene in L929ψpCEBP3Δ77 cells, but a highly abundant larger transcript was seen (Fig.7.20 A). This has been observed in all cell populations containing the P3Δ77 promoter and therefore is probably the readthrough transcript described previously (Section 7.2).

The inability to detect inducible promoter activity in L929ψpCEBP3Δ77 cells may reflect the cell population. Indeed, Southern blot analysis indicated a significantly higher copy number of transfected *cat* genes in L929ψpCEBP3Δ77 cells than other cells examined. Therefore, this promoter should be re−examined to determine if it restores either constitutive or inducible activity to the inactivated MuIFNα_1 promoter containing only 94bp of 5' flanking sequence.

Thus, this analysis demonstrates restoration of function to an apparently inactive deletion derivative of the MuIFNα_1 promoter by the Mo−MuSV enhancer. This suggests the basis for the behaviour of the deleted promoter is that upstream transcription activating functions have been lost. The ability to re−establish polyrI.rC inducibilty upon the deleted promoter shows that the sequences required for induction are probably still intact. Thus, this is consistent with the MuIFNα_1 promoter possessing at least two components, one controlling gene regulation and a second being responsible for efficient transcription from the activated promoter.

FIGURE 7.20 S-1 Nuclease Analysis of Chimaeric Gene Transcripts in L929ψpCP3Δ77, L929ψpCEAP3Δ77 and L929ψpCEBP3Δ77 Cells

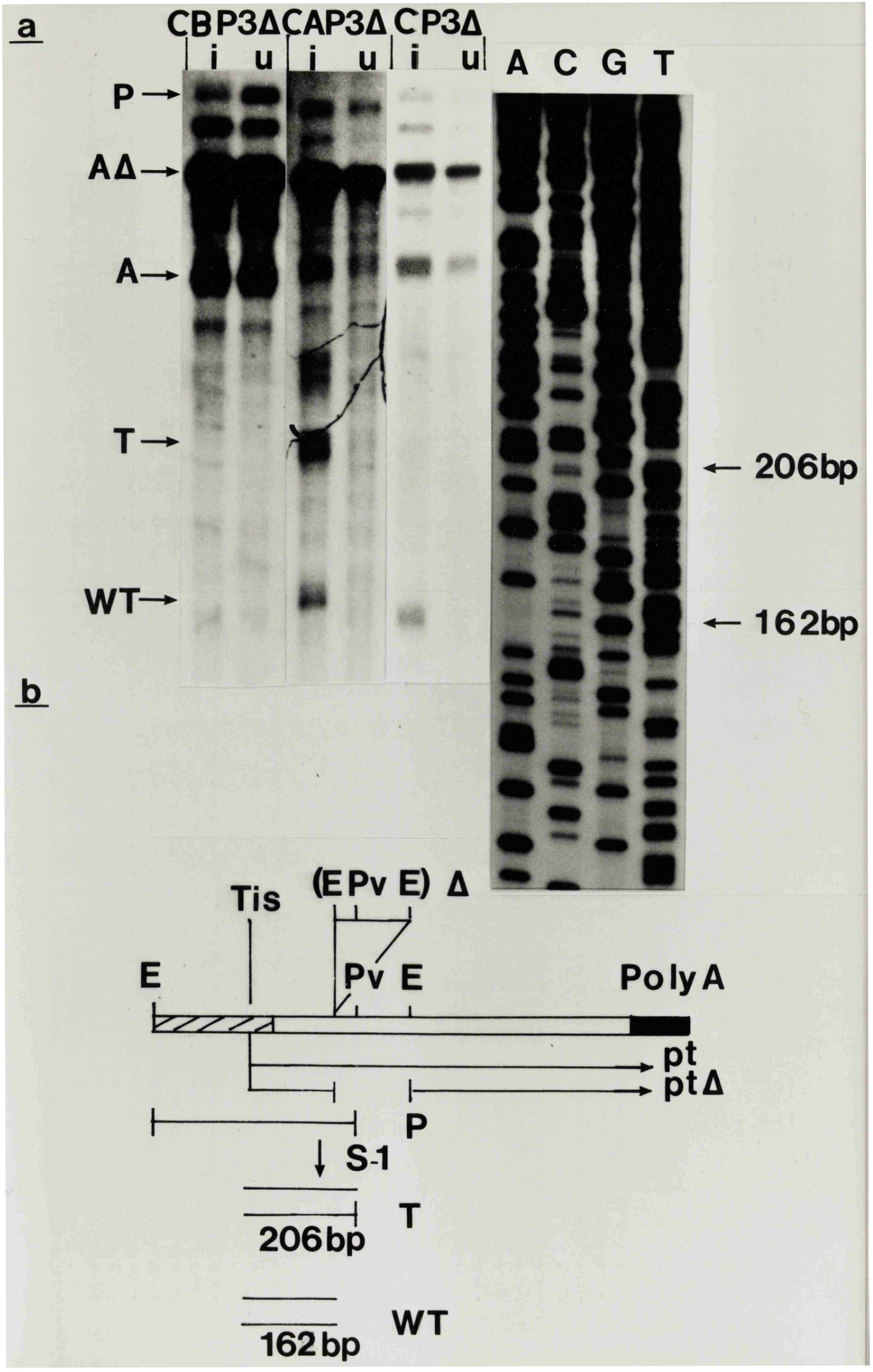
See legend to Fig.7.8.

Autoradiograph (a).

CP3Δ = L929ψpCP3Δ77 Cells

CAP3Δ = L929ψpCEAP3Δ77 Cells

CBP3Δ = L929ψpCEBP3Δ77 Cells


a
CBP3Δ
CAP3Δ
CP3Δ
i
u
i
u
i
u
A C G T
P
AΔ
A
T
WT
206bp
162bp
b
Tis
(E Pv E) Δ
E
Pv E
Poly A
pt
ptΔ
P
S-1
T
206bp
WT
162bp

7.7 Conclusions

These data demonstrate the use of an S-1 nuclease protection assay, designed to examine MuIFNα_1 promoter function. Preliminary investigations revealed, as expected, that polyrI.rC mediated the transient accumulation of correctly initiated chimaeric MuIFNα_1/*cat* gene transcripts in stably transfected L-929 cells. Maximum mRNA levels were identified approximately nine hours after induction, preceeding optimum CAT production by three to seven hours. Thus, this confirmed the conclusions derived in Chapter 6 that biologically relevent studies relating to MuIFNα_1 promoter function could be addressed using this hybrid gene system.

The S-1 nuclease protection assay has also been used to quantitate the relative abundance of chimaeric gene transcripts derived from MuIFNα_1 promoters. The construction of a functional pseudogene enabled the direct comparison of wildtype and mutant MuIFNα_1 gene promoters which are expressed simultaneously within the same cell populations. This analysis has confirmed and extended conclusions concerning MuIFNα_1 promoter function derived from CAT expression assays described in Chapter 6.

Thus site directed mutagenesis has revealed at least three distinct regions which may contribute to MuIFNα_1 promoter function. Two of these are defined here by single nucleotide substitutions at -87 and +14 which reduce and enhance respectively, polyrI.rC mediated gene transcription. The -87 mutation, located within a region highly conserved in IFN promoters, has been discussed previously (Chapter 6). The +14 mutation is located downstream of the transcription initiation site within the non-translated leader sequence of the MuIFNα_1 gene. It creates a GGAGAA sequence between +12 and +17 which is normally found in the MuIFNα_2 and MuIFNα_4 gene promoters. In this location the mutation could enhance either the efficiency of transcription or act post-transcriptionally.

A third region has been revealed by the reduced transcription of a MuIFNα_1 promoter containing two nucleotide transitions at -64 and -62 and two transversions at -51 and -49 located within part of an inverted repeat and a prospective GGAGAA motif resepectively. Neither mutation affects promoter activity independently (Chapter 6)

but combined a 3 to 5 fold reduction in transcription is observed. Interestingly, close inspection of autoradiographs obtained from S-1 protection of transcripts derived from mutant promoters containing the -64 and -62 mutations within the inverted repeat sequence (region IV, Chapter 6) all produce dual transcripts. Therefore, these sequences may also influence the accuracy of transcription initiation from this promoter.

Perhaps the most interesting observation of all in the work described here is the ability of the Mo-MuSV enhancer to restore polyrI.rC regulated transcription to an inactive MuIFNα_1 promoter containing only 94bp of promoter sequence upstream of the transcription initiation site. With the enhancer in one orientation, CAT assays revealed elevated basal gene transcription, the efficiency of which was significantly enhanced by polyrI.rC. PolyrI.rC mediated activation of correctly initiated gene transcription confirmed that the authentic MuIFNα_1 gene promoter had been reactivated. Thus, the ability of a constitutively active viral enhancer to functionally replace sequences normally required for efficient MuIFNα_1 gene transcription in induced L-929 cells is consistent with this promoter normally containing an enhancer.

CHAPTER 8

DISCUSSION

8.1 Human IFNα Chromosomal Genes

The first part of this work involved the isolation of human IFNα_1 genes from a newly constructed genomic library in λL47 (Chapter 3). Four IFNα genes have been described. A detailed analysis has revealed that each had been isolated previously (Chapter 4). In the past, at least 23 distinct human IFNα gene loci have been identified (Henco *et al,* 1985). Thus, because so many IFNα genes are now known and because each of the genes isolated here have been described before, it is highly likely that there are few more distinct members of this family which remain to be discovered.

Whilst this may be true for each IFN gene locus, some of the results described here do however indicate that there may be more alleles for particular loci awaiting discovery. In some instances functional alleles of loci previously designated as pseudogenes may exist. This possibility is exemplified by one of the genes described here, SMTIII.1_A (Section 4.2), which was isolated from a λ clone containing two IFNα genes. Both restriction mapping and sequencing studies confirmed SMTIII.I_A to be an allele of the ψLeIF−L gene (Ullrich *et al,* 1982). However, unlike the pseudogene ψLeIF−L gene, which contains a single nonsense mutation within the nucleotide sequence encoding the signal peptide, SMTIII.1_A encodes a full length IFN (Section 4.2). Futhermore, the primary amino acid sequence of this IFN corresponds exactly with that of the cDNA, LeIF−C (Chapter 4). Thus, the cDNA LeIF−C (Goeddel *et al,* 1981) has probably been derived from a transcript originating from a gene similar to SMTIII.1_A. Therefore, this is the first description of a chromosomal gene equivalent to LeIF−C.

Thus, although most of the human IFNα genes have now been described, the possibility remains that other allelic variants exist. As described here, these may encode full length interferons from loci previously thought to be non−functional. Similarly, there may well be pseudogene equivalents of genes previously regarded as functional. The frequency of particular alleles within the human population is of course unknown.

Clearly, therefore, caution should be exercised concerning the conclusions about genes which have been isolated from a very limited source of DNA. A strong implication of the existence of large numbers of IFNα genes, several members of which may be functional or non−functional, is that significant degeneracy can be tolerated in the human IFN system. Thus, whatever the selection pressures which have encouraged the development of such a relatively large gene family (Golding and Glickman, 1985; Henco *et al*, 1985), there still remains little selective disadvantage arising from having a few members of the family inactive.

8.2 The MuIFNα$_1$ Promoter/*cat* Gene System

Data are presented describing the identification of a MuIFNα$_1$ promoter by both comparative structural and functional studies. A 241bp MuIFNα$_1$ promoter fragment (−188 to +52; Section 5.1) has been used to drive polyrI.rC regulated transcription of the bacterial gene chloramphenicol acetyltransferase (*cat*) in pooled populations of stably transfected L−929 cells. The production of CAT is reproducible in independent cell populations providing a valid estimate of MuIFNα$_1$ promoter activity (Section 5.2).

The hybrid gene is regulated in a manner typical of IFN genes. The production of CAT induced by polyrI.rC follows a transient profile of expression (Section 7.1) which is preceded by the accumulation of correctly initiated transcripts derived from the hybrid gene in stably transfected L−929 cells (Section 7.2). Thus, these observations have demonstrated that biologically relevant studies, concerning MuIFNα$_1$ promoter activity, could be performed with this hybrid gene transcription system.

Through this work a preliminary characterisation of the MuIFNα$_1$ promoter, using *in vitro* oligonucleotide mediated site directed mutagenesis together with *in vivo* expression studies using the hybrid gene transcription system, has defined several sequences which appear to contribute to MuIFNα$_1$ promoter activity.

Evidence has been presented which suggests that an inverted repeat structure (region IV), a reiterated hexanucleotide sequence, GGAGAA (region III), and a region highly conserved in HuIFNα, HuIFNβ and MuIFNα genes (region II) each affect MuIFNα$_1$ promoter function with respect to polyrI.rC regulation in stably transfected L−929 cells

(Chapter 6). However, their limited impact upon gene expression suggests that a) those nucleotides vital for promoter activity have not been mutated or b) that the wrong transitions or transversions have been made. To assess these small changes in gene expression more accurately an alternative procedure was designed to examine the relative transcription efficiencies of MuIFNα_1 promoter mutants. This relied upon an S−I nuclease transcription assay which allowed the direct comparison of mutant and wild type MuIFNα_1 promoter activity simultaneously in the same cell population (Chapter 7).

The ability to identify and mutate specific regions of the MuIFNα_1 promoter in a predetermined fashion clearly provides a valuable approach to dissecting functional activity. However, certain aspects of IFNα gene regulation which may otherwise be difficult to investigate may also be approached using the MuIFNα_1 promoter/*cat* gene system. In addition to defining *cis*−acting sequences in promoter function, the identification, location and purification of *trans*−acting factors which interact with specific nucleotides will be essential to a full understanding of the mechanisms of gene regulation. It is well established that the chromatin of transcriptionally active genes exhibit nuclease hypersensitive sites at the 5' region (Stadler *et al*, 1980). However, such methodology is impractical when an analysis of a closely related gene family is being conducted (Higashi, 1984). The use of well characterised cell lines containing single copies of the hybrid gene, which are expressed in a regulated fashion, may be an attractive alternative for investigating changes in chromatin structure during the activation of a single IFNα gene promoter. The identification and localisation of *in vivo* DNA/ protein interactions through DNA footprinting procedures (Zinn and Maniatis, 1986) may also be more readily achieved in the IFNα gene system using such cell lines.

Thus, this study has demonstrated the suitability of the chimaeric MuIFNα_1/*cat* gene system for investigating MuIFNα_1 gene promoter activity in mouse cells. Several promoter mutants have been constructed by site directed mutagenesis and demonstrated to affect polyrI.rC mediated MuIFNα_1 promoter function, as assessed both by quantitative CAT assays and S−1 nuclease protection studies.

8.3 Activity of the MuIFNα_1 Gene

At the outset of this study the endogenous MuIFNα_1 gene had only been observed to be expressed at low levels relative to the other available IFNα genes. Thus, in NDV induced Ehrlich ascites tumour cells transcription levels were 6 to 15 fold less than from the MuIFNα_2 gene in the same cells (Shaw *et al*, 1983). The failure to detect MuIFNα_1 mRNA production in induced L−929 cells (Kelley and Pitha, 1985; Zwartoff *et al*, 1985) tended to confirm the low activity of this gene relative to other MuIFNα genes examined. Such observations are perhaps most consistent with the weak promoter activity associated with this gene.

Results presented here bear directly upon the origin of this observed low expression. The demonstration of MuIFNα_1 promoter activity following the introduction of the chimaeric gene into L−929 cells shows the MuIFNα_1 gene possesses a functional promoter (Section 5.2). Thus, at least this component of the MuIFNα_1 gene is not inherently inactive. Without undertaking a direct comparison of MuIFNα promoters employing the experimental system used in the work described here, it is not possible to assess their relative efficiencies. However, two lines of evidence are consistent with the MuIFNα_1 gene containing a fully functional promoter. Firstly, the creation and functional evaluation of MuIFNα_1 promoter mutants containing base substitutions commonly found in other endogenously active MuIFNα gene promoters (Region III_1, III_2 and III_3) has failed to significantly elevate promoter activity. Secondly, the failure of the powerful Mo−MuSV enhancer (Laimins *et al*, 1982) to potentiate induced transcription from the intact MuIFNα_1 promoter (Section 5.3) supports this conclusion.

Although these results suggest the MuIFNα_1 gene promoter may be functioning at its maximum capacity in polyrI.rC induced cells, it is possible that suppressor sequences are normally located either upstream or downstream of the promoter fragment (−188 to +52) used in this analysis. Such sequences have been identified in several mammalian cellular and viral genes including the rat insulin I gene (Laimins *et al*, 1986), the mouse c−*myc* gene (Remmers *et al*, 1986) and the HTLV−III/LAV LTR (Rosen *et al*, 1985). Futhermore, it is still possible that mutations other than those desribed here would increase gene expression.

The work described here would nevertheless be most consistent with the conclusion that the MuIFNα_1 gene has an active, fully functional promoter. Inherent promoter weakness is, therefore, unlikely to account for the differential gene expression observed with endogenous genes. An alternative explanation for low level expression of the endogenous MuIFNα_1 gene ought, therefore, to be sought.

Firstly, it is possible that the MuIFNα_1 promoter responds more efficiently to polyrI.rC mediated induction rather than NDV, although the inactivation of promoter function by deletion to −94 (Sections 6.4 and 7.6) indicates both inducers require common *cis*−acting sequences for induction. Secondly, the impact of the chromosomal location on expression of the endogenous gene is also a distinct possibility, although the MuIFNα_1 gene was isolated 7.6kb from the very active MuIFNα_4 gene on the same λ clone (Kelley and Pitha, 1985b; Zwarthoff *et al*, 1985).

Thirdly, differential gene expression could reflect different tissue specificities of the individual MuIFNα genes. There are precedents for this in IFN gene expression. For example IFNβ is the major anti−viral product of polyrI.rC induced human fibroblasts whereas human leukocytes produce mainly IFNα (Chapter 1). Tissue specific differential gene expression of the human IFNα genes has also been observed (Hiscott *et al*, 1983). Transcripts from the human IFNα_{14} gene are undetectable in normal blood leukocytes but are highly expressed in leukaemic cells. For the MuIFNα_1 gene it would be necessary to assume that the information dictating tissue specificity is either located upstream of −188, downstream of +52 or lost upon introduction of the MuIFNα_1/*cat* gene into L929 cells.

Thus, these results demonstrate that the MuIFNα_1 gene contains a functional promoter. Therefore, it is most likely that some other factor(s) are responsible for the low expression of the endogenous gene observed by other workers.

8.4 Properties of the MuIFNα_1 Promoter

Significant progress has been made by other workers in defining the functional boundaries of human IFNα_1 and IFNβ gene promoters by deletion mutagenesis (Chapter 1). However, apart from their ability to confer polyrI.rC or virus inducibility upon

heterologous genes, little else was known about the nature of these regulatory sequences at the outset of this work. Studies described here provide valuable clues which bear directly upon this issue, providing some further insight into the organisation of IFN gene promoters.

Firstly, S−1 nuclease protection studies have revealed two distinct transcripts, specific to chimaeric gene sequences, in L−929 cells stably transfected with the MuIFNα_1/*cat* gene (Section 7.2). One of these is derived from the expected MuIFNα_1 gene promoter and is completely dependent upon polyrI.rC for induction. The second transcript probably arises as a readthrough transcript from cryptic promoters located upstream of the MuIFNα_1/*cat* gene within plasmid DNA (Section 7.2). The readthrough transcript is expressed at high basal levels but is significantly enhanced by polyrI.rC (Section 7.2). It is unlikely that the cryptic promoter(s) responds to the inducer directly, raising the possibility that the MuIFNα_1 gene promoter sequences are responsible. The ability to activate transcription of upstream heterologous gene promoters, perhaps over long distances, is characteristic of enhancers (Chapter 1). Thus, this observation is consistent with the MuIFNα_1 gene promoter sequences exhibiting inducible enhancer−like properties.

This conclusion is considerably strengthened by results obtained with viral enhancer sequences and an inactive deletion derivative of the MuIFNα_1 gene promoter. A MuIFNα_1 gene promoter was fortuitously obtained, which contained a 77bp deletion, from site directed mutagenesis studies (Section 6.2). Investigation of the functional activity of this promoter indicated that it was no longer able to confer polyrI.rC inducibility upon the heterologous *cat* gene (Sections 6.4). Subsequent studies revealed that the Mo−MuSV enhancer could complement for the loss of sequences required for MuIFNα_1 gene promoter activity (Section 7.6). Thus, in the presence of this enhancer, polyrI.rC induction is restored to the deleted MuIFNα_1 gene promoter.

The Mo−MuSV enhancer is normally constitutively active in mouse L929 cells (Kreigler and Botchan, 1983). Therefore, it is most likely that this enhancer functionally replaces MuIFNα_1 upstream sequences, that are absent from the deletion derivative, which are responsible for efficient gene transcription from this promoter in the induced

state. Thus, again this is consistent with this promoter containing a cellular enhancer. Furthermore, whatever keeps the MuIFNα_1 gene promoter from being expressed in non−induced cells also works on this enhancer.

A precedent for the existence of an IFN gene enhancer has been indicated in the HuIFNβ gene promoter (Chapter 1). Functionally important sequences of this promoter (Goodbourn *et al*, 1986) display significant homology to two regions of the MuIFNα_1 promoter between the bases at −171 to −155 and −99 and −76 (Table 8.1). Thus, these comparative sequence studies also support the functional evidence attributing enhancer properties to the MuIFNα_1 promoter.

In addition to the direct experimental evidence for the existence of enhancer like sequences there is also substantial structural data consistent with this conclusion. Although enhancers do not share regions of extensive sequence homology a short functionally important core sequence, first identified in the SV40 72bp repeat, is common to many viral and cellular enhancers (Chapter 1). This precise consensus core sequence is also found within the promoter region of the MuIFNα_1 gene between the bases −81 to −74. A detailed comparison of the consensus core enhancer sequence with regions of the MuIFNα gene promoters are shown (Table 8.1). This reveals additional elements of the MuIFNα promoters showing homology to the viral enhancer core sequence. Therefore, both structural and functional evidence strongly suggests the presence of a cellular enhancer component in the MuIFNα_1 gene promoter, which contributes to the promoter activity of IFN genes in L−929 cells.

An interesting additional property of the intact MuIFNα_1 promoter is that it is refractory to the influence of the Mo−MuSV enhancer in the non−induced state (Section 5.3). Moreover, this property is not merely confined to the combination of the MuIFNα_1 gene promoter and Mo−MuSV enhancer. For example, examination of data obtained by other workers reveal that the regulation of the HuIFNβ gene promoter is maintained in replicating BPV (Zinn *et al*, 1982) and SV40$_{ori}$ based vectors (Tavernier *et al*, 1983), each of which contain enhancer sequences.

However, the 77bp deletion derivative of the MuIFNα_1 gene promoter, studied in this work, responds to the Mo−MuSV enhancer (Section 7.6). Thus, upon the deletion

TABLE 8.1 Location of Sequence Motifs Showing Homology to Enhancer Sequences

A) MuIFNα_1 Promoter

-104 -52

AGTTAAAGAAAGTGAAAAGACAAGTGGAAAGTGATGGAAGGGCATTCAGAAAG

AG G T A G TG/C

B)

-170	-164		-104	-97		-94*	-86	
AGTGAAAG	$\alpha_{1,2,5,6A/B}$		AGTTAAAG	$\alpha_{1,5,6A}$		GTGAAAAG	$\alpha_{1,2,4,5,6A}$	
AGTGAAAC	α_4		AAGTAAAG	$\alpha_{2,4}$		GTGAAAGG	α_{6B}	

-81	-74		-59†	-52	
GTGGAAAG	$\alpha_{1,5,6A}$		TCAGAAAG	$\alpha_{1,5,6A/B}$	
TTGGAAAG	$\alpha_{4,6B}$		TTGGAAAG	α_2	
TTAGGAAG	α_2		TTCGAAAG	α_4	

Core Enhancer Sequences:
SV40* GTGGAAAG (Weiher *et al*, 1983)
CMV† TTGGAAAG (Boshart *et al*, 1985)

C)
Comparison of MuIFNα_1 Gene promoter Sequences with the HuIFNβ Enhancer

HuIFNβ Enhancer: -91 GGAAAACTGAAAGGGAGAAGTGAAA -67 (Goodbourn *et al*, 1986)
MuIFNα_1 : -99 AAG---G-----A --C-----G-- -76
MuIFNα_1 : -171 -G------A- ---- -C-- -155

The primary nucleotide sequence of part of the MuIFNα_1 gene promoter (A) is shown from -104 to -52 (relative to the transcription initiation site). Regions showing homology to the consensus core enhancer sequence are shown underlined. Base substitutions within other MuIFNα promoters in these regions are shown. B) shows the nucleotide sequence of these regions, indicating the location of these sequences within the MuIFNα_1 gene promoter. Homologous sequences within these regions from other MuIFNα gene promoters examined in this study (Fig.6.2) are also shown. Below is shown the core enhancer sequences of SV40 and the cytomegalovirus LTR (CMV). * and † indicates sequences of MuIFNα promoters revealed to be identical to these viral enhancer motifs. C) shows a comparison of the HuIFNβ gene promoter enhancer sequences with two regions of the MuIFNα_1 gene promoter. The nucleotides for specific base differences are shown. Homology is indicated by a dash (-).

of this region of the promoter, the efficiency of basal gene transcription is significantly increased in stably transfected L−929 cells in the presence of this enhancer. With the Mo−MuSV enhancer in one orientation, high constitutive transcription is observed from the MuIFNα_1 gene promoter (Section 7.6). However, in the opposite orientation, inducibility is maintained despite elevated basal gene expression. Thus, it is clear from this that at least some of the sequences responsible for the refractory behaviour of the MuIFNα_1 gene promoter to the Mo−MuSV enhancer sequences are maintained in the deleted promoter.

These results are consistent with the MuIFNα_1 gene promoter being under negative control in addition to the positive regulation described in this work. Two possible interpretations of the repressed state, with respect to the refractory behaviour are that promoter selection by the RNA polymerase II is prevented or that sequences within the MuIFNα_1 promoter region can interfere with positive regulation, preventing gene transcription in the non−induced state. DNA fragments with "silencer" properties similar to this which can inhibit enhancer function have been identified in the rat insulin I gene (Laimins *et al*, 1986), the mouse c−*myc* gene (Remmers *et al*, 1986) and the virus HTLV−III/LAV LTR (Rosen *et al*, 1985). However, further studies are required to determine if the MuIFNα_1 gene promoter contains sequences which fall into this category of potential regulatory sequences.

Thus, the possibility that the MuIFNα_1 promoter contains both enhancer and "silencer" properties has important implications with respect to the regulation of gene expression. For example it is possible that the strength of gene transcription is dependent upon the enhancer, and the inducible component of the promoter is determined by negative regulation which in the repressed state inhibits enhancer activity. Induction would relieve the repression allowing the enhancer to activate full transcription from the IFN promoter.

8.5 Modular Structure of the MuIFNα_1 Promoter

Interferon gene expression is complex. The IFN genes comprise a family of tightly regulated genes which are inducible by a variety of agents including polyrI.rC, virus and

growth factors including PDGF and CSF−1 (Zullo *et al*, 1985; Moore *et al*, 1984). Interferon also regulates its own production, a phenomonen exploited in IFN production (priming; Chapter 1). In certain circumstances IFN promoters are refractory to induction. For example in early stages of development there is a latent period. This is also seen in undifferentiated embryonic stem cells (Coveney *et al*, 1984).

A common feature emerging from the analysis of the structure of the promoters of higher eukaryotes and their viruses is that they are composed of multiple functional elements (reviews: Serfling *et al*, 1985; McKnight and Tjian, 1986). Generally, promoters recognised by RNA polymerase II possess a TATA box determining the accuracy of transcription initiation. However, as described previously not all genes show TATA box homology (Chapter 1). Nevertheless, whether or not the TATA box is present, upstream elements dictate the frequency of initiation of gene transcription (review: Sassone−Corsi and Borrelli, 1986).

Data obtained here from site directed mutagenesis and gene expression studies (Chapters 6 and 7) suggest that the MuIFNα_1 gene promoter is no exception to this trend of modular organisation. The first promoter element to be identified within these genes was the TATA box (Shaw *et al*, 1983). All IFNα gene promoters examined possess TATA box homology, which is located approximately 30nt upstream of the transcription initiation site (Henco *et al*, 1985). However, this homology is imprecise, IFNα genes commonly having the sequence TATTTAA (Henco *et al*, 1985).

The prospective relevance of the variant sequence, TATTTAA, has been investigated directly in this study by site directed mutagenesis (Chapter 6). The construction of a MuIFNα_1 gene promoter mutant containing a more conventional TATA box (Section 6.2) was undertaken to address this very issue. However, functional studies demonstrated that there was no obvious impact of the presence of this variant sequence with respect to polyrI.rC mediated activation of the MuIFNα_1 gene promoter in L−929 cells (Section 6.4).

These observations are, therefore, consistent with the MuIFNα_1 gene TATA box sequence, having no major role either in the efficiency or mechanism of activation of gene transcription from this promoter. This is consistent with observations by other

workers investigating the role of the TATA box in *in vivo* gene expression. An absolute requirement for the TATA box sequence in gene expression has only been demonstrated *in vitro* (Grosveld *et al*, 1981). However, deletion of the SV40 early promoter TATA box does result in heterogeneity of *in vivo* transcription initiation (Benoist and Chambon, 1981). It is now accepted that the TATA box determines the accuracy of transcription initiation *in vivo* (Benoist and Chambon, 1981; Grosveld *et al*, 1982). Thus, the TATA box sequence peculiar to IFNα genes probably have a similar function in determining the location of the transcription initiation site.

Immediately upstream of the TATA box of the MuIFNα_1 promoter is a sequence of ten nucleotides (−42 to −31) which forms part of an imperfect inverted repeat (−66 to −57) sequence (region IV; Section 6.1). These sequences are separated by fifteen bases of intervening DNA sequence, thus spanning 35nt overall. Mutant promoters were created to investigate the relevance of these inverted repeat structures (Section 6.1). *In vivo* expression studies revealed that the complementarity of these sequences was not important for MuIFNα_1 gene promoter function (Sections 6.4 and 7.4). However, this region is implicated both in positive and negative regulation of gene expression. Site directed mutagenesis of the sequences located between −42 to −31, designed to disrupt the complementarity of the inverted repeats (Section 6.1), produced a reproducible increase in basal gene expression from the MuIFNα_1 promoter *in vivo* (Section 6.4). This suggests a negative regulatory function for this region in MuIFNα_1 gene promoter regulation. Additionally, the combined mutagenesis at two locations, one between −66 to −57 and the second between −52 and −47, resulted in a significant reduction of polyrI.rC activation of the MuIFNα_1 promoter (Section 7.5). Thus, this is consistent with this region also having a positive role in gene expression.

The presence of an inverted repeat structure within the HuIFNβ gene promoter has also recently been pointed out by Goodbourn *et al* (1986). This structure has also been functionally implicated in the negative regulation of this promoter (Goodbourn *et al*, 1986). The relatedness in structure and location of the inverted repeats of the MuIFNα_1 gene promoter described here and the HuIFNβ gene inverted repeat would again, therefore, be consistent with a common function

A third functionally important element of the MuIFNα_1 promoter, upstream of the inverted repeat sequence, is located between −102 and −77 (Section 6.1) within a region of extensive sequence conservation between human and mouse IFNα and IFNβ gene promoters (this work). The functional relevance of this region is indicated in this work by the mutation analysed at position −87. This results in a moderate decrease in MuIFNα_1 gene promoter activity. This mutation falls within the DNA sequence homologous to the interferon gene regulatory element (IRE) defined in the HuIFNβ gene (Goodbourn *et al*, 1985) and the IFNα repeats identified in the HuIFNα_1 gene promoter (Ryals *et al*, 1985).

As described in Section 8.3 (Table 8.1), an element within this region shows homology to the SV40 enhancer core element. Thus, this is consistent with these sequences, initially identified on a structural basis, having a functional relevence also. However, it is also very interesting to note that the very same region of the MuIFNα_4 gene contains a perfect CCAAT box sequence, but in an inverted orientation, between the nucleotides present at −79 to −83 (Table 8.2). Inspection of the MuIFNα_1 gene promoter indicates it also has partial CCAAT box homology in this region (Table 8.2). Furthermore, examination of other MuIFNα, HuIFNα_1 and HuIFNβ gene sequences reveals that they too contain CCAAT box homology (Table 8.2).

Therefore, the MuIFNα_1 gene promoter has an imperfect CCAAT box just downstream of the mutation at −87. Aside from the relatedness in general location (80bp upstream of the transcription initiation site) and DNA sequence to this element it is also intriguing that the perfect sequence homology of the MuIFNα_4 promoter (Table 8.2) correlates with a significant elevation of inducible expression of this endogenous gene, relative to other MuIFNα genes examined, in L−929 cells (Chapter 1).

Thus, both comparative structural and functional studies clearly indicate a role for this region in MuIFNα_1 gene promoter function. It has structural similarities with several DNA sequences which are believed to interact with transcription factors (Graves *et al*, 1986; Mercola *et al*, 1985). Therefore, this is consistent with this region being a DNA protein binding site. However, clearly this conclusion is based mainly upon structural observations and functional evidence must be sought to extend this possibility.

TABLE 8.2 CCAAT Box Homology in IFN Gene Promoters

MuIFNα_1	-72	CACTTTCCACTTGTCT	-87
MuIFNα_2	-73	TGCTTCCTAATTCTCT	-88
MuIFNα_4	-73	TGCTTTCCAATTCTCT	-88
MuIFNα_5	-72	CACTTTCCACTTGTCT	-87
MuIFNα_{6A}	-72	CACTTTCCACTTGTCT	-87
MuIFNα_{6B}	-64	AACTTTCCAACTGTCC	-79
HuIFNα_1	-72	AACTTTCCATTTCTGT	-87
HuIFNβ	-59	AATTTCCCACTTTCTC	-74
Hu β-GLOBIN		GGTTGGCCAATCTACT	
Mu β-GLOBIN^{maj}		TAAGGGCCAATCTGCT	
HSV-tk	-76	AATTCGCCAATGACAA	-91
MSV LTR	-97	AACTAACCAATCAGTT	-72

Partial sequence data of promoter sequences of several genes are shown. Underlined are the CCAAT box sequences identified in the human and mouse β-globin genes and similar sequences identified in the HSV *tk* promoter and Mo-MSV LTR sequences (Graves *et al*, 1986). Above these, are aligned sequences of part of several human and mouse IFNα and IFNβ gene promoters revealed to show CCAAT box homology. These sequences are shown in the opposite orientation. The CCAAT box homology is underlined. The location of these sequences are depicted by numbering relative to the transcription initiation sites.

During the construction of MuIFNα_1 promoter mutants by site directed mutagenesis a deletion derivative was obtained (Section 6.2). Sequencing revealed that 77bp had been deleted from this promoter, leaving only 94bp upstream of the transcription initiation site. Functional analysis has demonstrated that neither CAT assays (Section 6.4) nor S−1 protection studies (Section 7.6) could detect inducible transcription from this disabled promoter. Surprisingly, it has been demonstrated here that the Mo−MuSV enhancer is able to re−activate expression from this promoter (Section 7.6).

Complementation of the disabled MuIFNα_1 promoter deletion mutant by an enhancer highlights two features of IFN promoter structure and function. Firstly, the *cis* regulatory sequences required for polyrI.rC induction are located downstream of −94. This is in agreement with the observations from deletion mutagenesis of the HuIFNβ (−77) gene promoter using BPV vectors (Zinn *et al*, 1984). Secondly the frequency of initiation is dictated by sequences disrupted by, or upstream of, the −94 deletion. This observation is in agreement with deletion mutagenesis of the HuIFNα_1 and HuIFNβ gene promoters where the 5' boundary for maximum NDV induction in L929 cells is found approximately 109bp upstream of the respective transcription initiation sites (Ryals *et al*, 1985; Fujita *et al*, 1985).

Thus, the MuIFNα_1 gene promoter can be divided into two distinct upstream elements. Figure 8.1A illustrates this. One of these elements is responsible for the regulation of IFN gene expression. The second element dictates the strength of induced gene transcription. The functional substitution of the second element with a viral enhancer is consistent with it normally being a cellular enhancer itself.

These results only demonstrated restoration of the deleted MuIFNα_1 gene promoter function with the Mo−MuSV enhancer in one orientation (Section 8.3). Even in this case, the basal level of gene expression is markedly elevated (Section 7.6). Thus, additional properties of the intact MuIFNα_1 gene promoter must normally maintain the tight regulation of IFN production usually observed with these genes. Therefore, it is not sufficient merely to suggest that upstream sequences of this promoter have properties of cellular enhancers. Clearly, the MuIFNα_1 promoter sequences have additional properties to suppress basal expression levels. In this respect it is interesting that negative regulatory

sequences have been identified in human IFNα and IFNβ genes located approximately between 200bp and 110bp upstream of the transcription initiation site (Ryals *et al*, 1985; Zinn *et al*, 1983).

The observation that viral enhancer sequences can functionally substitute for upstream sequences of IFN gene promoters may have some implications with respect to the deletion mutagenesis studies performed by others. Deletion mutagenesis of the HuIFNβ gene promoter and expression studies have been performed with BPV vectors which contain enhancer sequences (Zinn *et al*, 1983). This analysis has demonstrated that the 5' boundary of polyrI.rC inducibility of this promoter is only 77bp upstream of the transcription initiation site (Zinn *et al*, 1983). However, other investigators have located the 5' boundary of this promoter between −119 and −110 by analysing the expression of integrated mutant HuIFNβ genes (Fujita *et al*, 1985). Based on the results described here, this obvious discrepency may perhaps be explained by the complementation of deleted upstream HuIFNβ promoter sequences by *cis*−acting enhancers within the the BPV vectors. Thus, this data may allow a possible rationalisation of the contrasting conclusions obtained by others.

The existence of enhancer sequences within IFN promoters is currently the subject of debate. Some investigators have strongly suggested that enhancers are integral components of IFN gene promoters (Goodbourn *et al*, 1985), whereas others have failed to demonstrate such properties (Fujita *et al*, 1985).

The apparent discrepency in these observations is difficult to reconcile. The distinction of upstream regulatory sequences and enhancers is becoming increasingly blurred. As described in the Chapter 1 the major distinction between these elements is that enhancers can function over considerable distances to activate efficient transcription of heterologous promoters. However, other properties are shared by enhancers and upstream regulatory sequences. Both are able to activate efficient transcription of heterologous promoters and appear to operate in an orientation independent fashion (Section 1.4).

Functional orientation and location independence for the IRE of HuIFNβ (Goodbourn *et al*, 1985) has been observed only under special circumstances, when

promoter sequences located between −73 to −60 are duplicated. It is possible that the duplication of upstream elements responsible for the frequency of transcription from IFN promoters, therefore, may actually create an inducible enhancer. Precedent for such an effect is provided by the observation that enhancer sequences were created by the duplication of SV40 sequences not normally associated with such activity (Swimmer and Shenk, 1984). Similarly duplication of the heat shock element (HSE) from the *hsp*70 gene which normally exibits properties of upstream regulatory sequences also generates an enhancer (Bienz and Pelham, 1986).

The potential enhancer activity within the MuIFNα_1 promoter indicated by this study may, therefore, also involve partial duplication of the sequence homologous to the HuIFNβ IRE located between the bases at −99 to −76 and −171 to −155 (Table 8.1). Thus, the MuIFNα_1 promoter may possess enhancer function as a result of this duplication. Clearly these observations are worthy of further investigation.

In addition to the motifs already described this study also focussed upon assessing the possible significance of a hexanucleotide repeat motif, GGAGAA, identified in the HuIFNβ and MuIFNα gene promoters (Section 6.1). The creation of the consensus region III_1 motif by site directed mutagenesis produced a 2 to 3 fold increase in mRNA and CAT production from the MuIFNα_1 promoter in polyrI.rC induced L−929 cells (Sections 6.4 and 7.4). The consensus region III_2 motif in combination with a mutation of part of an inverted repeat, which surprisingly resulted in decreased promoter function, has been discussed previously (Section 8.4).

The significance of these elements remains to be established. The moderate impact on gene expression observed may imply the consensus sequence is not essential to polyrI.rC regulation but critical to some other aspect of IFN gene control. Alternatively they may reflect some degeneracy of IFN gene promoters. Their significance, if any, may only be established by disrupting one or more of these region by site directed mutagenesis.

The major components of the MuIFNα_1 promoter, evidence for which has been obtained in this study, are summarised in Figure 8.1. The precise functional role of these sequences remains somewhat speculative although it is likely that these components

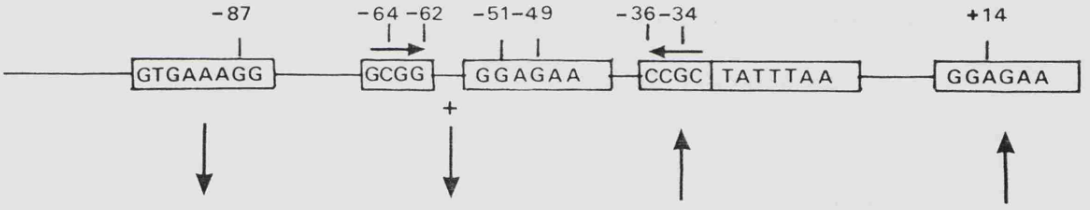
FIGURE 8.1 The Modular Structure of the MuIFNα_1 Gene Promoter

A) Shows that the MuIFNα_1 gene promoter is comprised of at least three elements. These include the TATA box (TATTTAA) and two distinct upstream regions which control the regulation of IFN gene expression (RE) and the strength of transcription (E) respectively. E: prospective enhancer, RE: Regulatory Element. B) Shows the location and nucleotide sequence of specific mutations created in the MuIFNα_1 promoter (-188 to +52 *Hin*dIII fragment), in this study, which affect its expression. The horizontal arrows depict inverted repeat structures. The vertical arrows indicate whether the corresponding promoter mutation increased or decreased polyrI.rC mediated gene expression. ↑: increase, ↓: decrease, +: the combination of mutations at -64, -62, -51 and -49.

A


E
RE
TATTTAA

B


-87
-64 -62
-51-49
-36-34
+14
GTGAAAGG
GCGG
GGAGAA
CCGC
TATTTAA
GGAGAA
+

act in concert to regulate IFN gene expression. More extensive site directed mutagenesis and gene expression studies using procedures of the type developed in these studies should help define more clearly the boundary, critical primary nucleotide sequence and role of such *cis*−regulatory sites.

8.6 Regulation of IFN Gene Expression

Thus, evidence has been presented which strongly supports the conclusion that there are both positive and negative factors influencing MuIFNα_1 gene expression. All the evidence provided by this work are consistent with the control of the MuIFNα_1 gene promoter function resulting from an inducible "silencer" element (Laimins *et al*, 1986) and an enhancer species located downstream and upstream of the position −94 respectively. Clearly the silencer is capable, at least in part, of overriding enhancer activity.

The best published model, incorporating both positive and negative regulatory elements has been provided by Zinn and Maniatis (1986). These workers proposed that the HuIFNβ enhancer is negatively regulated by the association of repressors with the IFN promoter preventing access of pre−existing regulatory factors in the non−induced state. Induction correlates with the concomittant dissociation of repressors and the association of positive factors which activate transcription (Section 1.5).

The majority of the data presented here is consistent with the control of the MuIFNα_1 gene working according to a similar model to that of Zinn and Maniatis (1986). The MuIFNα_1 gene promoter possesses enhancer−like properties that can be functionally substituted by the Mo−MuSV enhancer sequences. However, with an intact MuIFNα_1 gene promoter or a deletion derivative also containing Mo−MuSV enhancer sequences, negative factors can override enhancer activity. Repression is relieved upon induction with polyrI.rC, giving rise to efficient gene transcription. Thus, this suggests that the prospective MuIFNα_1 gene enhancer is under negative control. This is consistent with the model of Zinn and Maniatis (1986) which indicates that the HuIFNβ gene enhancer is also under negative control (Goodbourn *et al*, 1986).

However, whilst it is possible that the Zinn and Maniatis model fully explains the

regulation of both HuIFNβ and MuIFNα_1 gene expression, some consideration should be given to the precise consistency of all the available data. In the transient assays described here the expected IFN gene regulation was maintained in the non−induced state − no expression was observed. The transient assay procedure normally introduces a high copy number of exogenous gene sequences into host cells. Therefore, unless it were present in vast excess, the IFN gene repressors should be titrated by this influx of prospective binding sites, leaving other templates available to interact with positive regulatory factors. In such circumstances constitutive transcription from the MuIFNα_1 promoter would be anticipated according to the model of Zinn and Maniatis (1986).

Obviously, sufficiently extensive titration analysis has probably not been undertaken here for such a conclusive interpretation of the transient expression data. However, this argument is consistent with previous studies, performed by others, demonstrating regulated transcription from IFN promoters using replicating BPV (Zinn *et al*, 1982) and $SV40_{ori}$ (Tavernier *et al*, 1983) vectors which achieve approximately 50 and 30,000 copies per cell respectively.

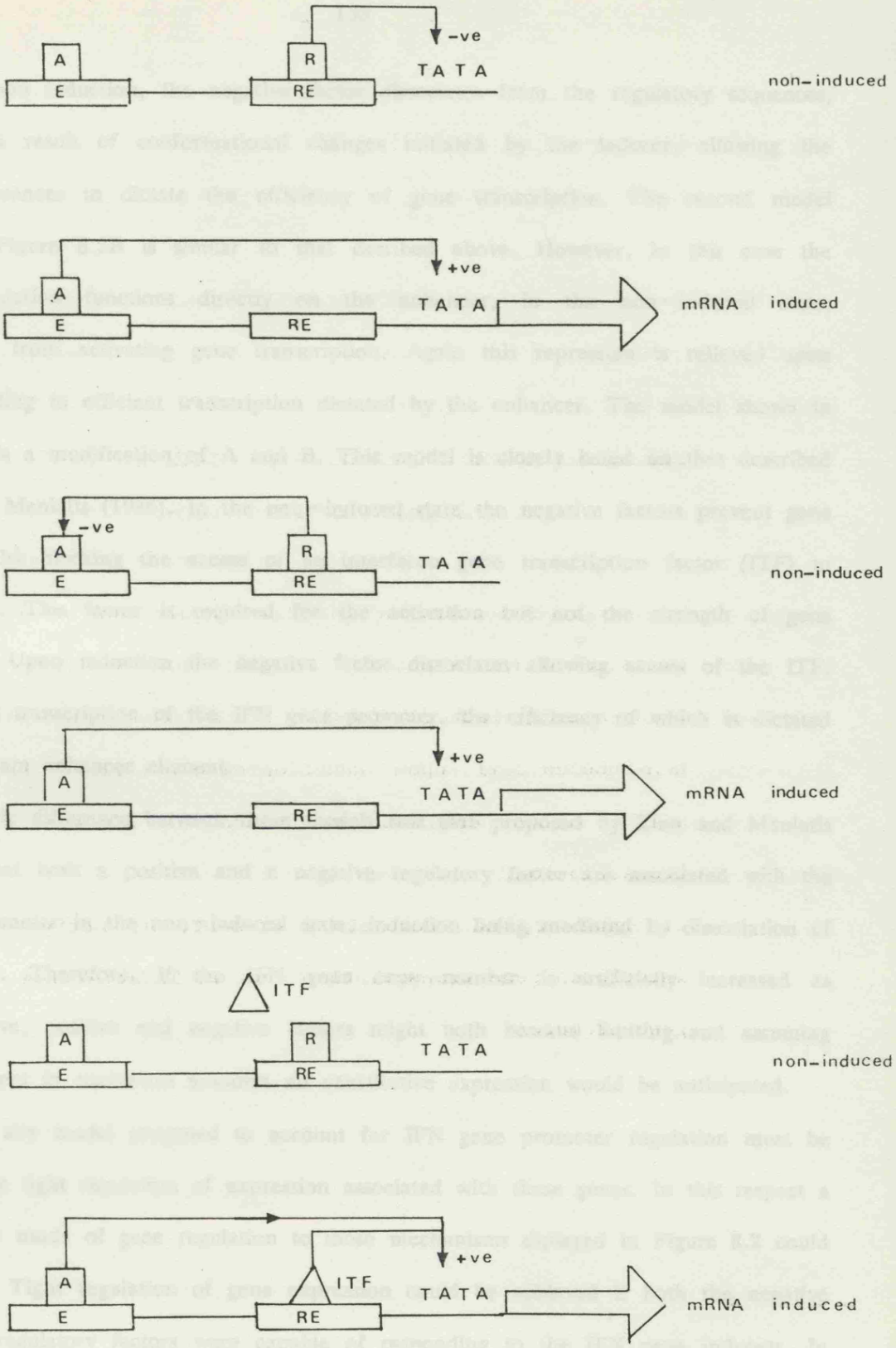
The original model of Zinn and Maniatis is based on the close physical proximity of "silencer" and enhancer sequences, such that the occupation of the "silencer" region prevents the access of positive factors to this enhancer. The data presented in this work indicates that the "silencer"/enhancer interaction is able to operate over more extensive distances. Thus, the MuIFNα_1 "silencer" can operate on the Mo−MuSV enhancer when located approximately 200bp upstream of the transcription initiation site. This may be accounted for by DNA looping, which would allow the interaction of negative factors at more than one site (Ptashne, 1986).

Therefore, it is possible that an alternative model may more precisely account for all these observations. Figure 8.2 shows three prospective models which are all consistent with the observations described in this study. In each case the upstream elements consist of two components. These are an enhancer (E) and a regulatory element (RE). In model A, both positive factors (A) and negative regulatory factors (R) are interacting with the promoter sequences in the non−induced state. The negative factors are assumed to be dominant to the positive factors by inhibiting transcription initiation from the IFN gene

FIGURE 8.2 Prospective Mechanisms for the Regulation of MuIFNα_1 Gene Expression by PolyrI.rC

Three mechanisms for the regulation of IFN gene expression are illustrated (A, B and C). In each case activator (A) and repressor (R) *trans*-acting factors interact with the gene promoter region in the non-induced cell. Upon induction with polyrI.rC the repressor dissociates, which allows enhancer sequences to dictate the strength of gene transcription.

E: Enhancer, RE: Regulatory element, A: Activator, R: Repressor, ITF: Interferon gene transcription factor, TATA:TATTTAA sequence of the MuIFNα_1 gene promoter


A
R
-ve
E
RE
TATA
non-induced
A)
+ve
A
E
RE
TATA
mRNA
induced
-ve
A
R
E
RE
TATA
non-induced
B)
+ve
A
E
RE
TATA
mRNA
induced
ITF
A
R
E
RE
TATA
non-induced
C)
+ve
A
ITF
E
RE
TATA
mRNA
induced

promoter. Upon induction, the negative factor dissociates from the regulatory sequences, possibly as a result of conformational changes initiated by the inducer, allowing the enhancer sequences to dictate the efficiency of gene transcription. The second model depicted in Figure 8.2B is similar to that desribed above. However, in this case the negative regulation functions directly on the enhancer, in the non−induced state, preventing it from activating gene transcription. Again this repression is relieved upon induction leading to efficient transcription dictated by the enhancer. The model shown in Figure 8.2C is a modification of A and B. This model is closely based on that described by Zinn and Maniatis (1986). In the non−induced state the negative factors prevent gene transcription by blocking the access of an interferon gene transcription factor (ITF) to the promoter. This factor is required for the activation but not the strength of gene transcription. Upon induction the negative factor dissociates allowing access of the ITF. This activates transcription of the IFN gene promoter, the efficiency of which is dictated by the upstream enhancer element.

The basic difference between these models and that proposed by Zinn and Maniatis (1986) are that both a positive and a negative regulatory factor are associated with the MuIFNα_1 promoter in the non−induced state, induction being mediated by dissociation of the repressor. Therefore, if the IFN gene copy number is artificially increased as discussed above, positive and negative factors might both become limiting and assuming they are present in equivalent amounts no constitutive expression would be anticipated.

Clearly any model proposed to account for IFN gene promoter regulation must be capable of the tight regulation of expression associated with these genes. In this respect a more complex mode of gene regulation to those mechanisms diplayed in Figure 8.2 could be envisaged. Tight regulation of gene expression could be achieved if both the negative and positive regulatory factors were capable of responding to the IFN gene inducers. In this case induction would inactivate repression, possible by dissociation from ther "silencer" region, while simultaneously activating positive factors which could then interact with enhancer sequences. Thus, this would activate transcription.

There is some data presented here which is consistent with the MuIFNα_1 gene containing such an inducible enhancer. Although the Mo−MuSV enhancer failed to

elevate transcription from the intact MuIFNα_1 gene promoter, it clearly elevated basal gene expression 2 to 5 fold. Thus, it is not sufficient to imply that the MuIFNα_1 gene contains an enhancer with entirely identical properties to the constitutively active Mo−MuSV enhancer.

The regulation of both negative and positive factors in the manner described above could lead to extremely tight coordinate control. IFNs are potent molecules which can block cell proliferation and destroy (at least transiently) protein synthesis. Thus, tight regulation may be vital in the IFN system where the consequences of inappropriate gene expression may be very severe for the cell physiology.

Clearly the results presented here are consistent with the MuIFNα_1 gene promoter being regulated both positively and negatively. However, neither the identity nor mode of action of either negative or positive regulatory sequences has been unequivocally determined although data described here suggest that the MuIFNα_1 promoter contains an enhancer and a "silencer". The identification, isolation and purification of *trans*−acting factors involved in IFN gene regulation is an exciting prospect for the future. Studies of this nature in combination with site directed mutagenesis should help clarify the location, precise primary nucleotide structure and role of the *cis*−regulatory sequences in the mechanism of polyrI.rC and virus induced gene expression.

REFERENCES

Aguet, M. (1980). High−Affinity Binding of ^{125}I−Labelled Mouse Interferon to a Specific Cell Surface Receptor. *Nature, Lond.* **284**, 459−461.

Allen, G. (1982). Structure and Properties of Human Interferon α from Namalwa Lymphoblastoid Cells. *Biochem. J.* **207**, 397−408.

Allen, G. and Fantes, K.H. (1980). A Family of Structural Genes for Human Lymphoblastoid (Leukocyte−Type) Interferon. *Nature, Lond.* **287**, 408−411.

Anderson, P., Yip, Y.K. and Vilcek, J. (1982). Specific Binding of ^{125}I−Human Interferon−γ to High Affinity Receptors on Human Fibroblasts. *J. Biol. Chem.* **257**, 11301−11303.

Auffray, C. and Rougeon, F. (1980). Purification of Mouse Immunoglobulin Heavy−chain Messenger RNAs from Total Myeloma Tumor RNA. *Eur. J. Biochem.* **107**, 303−314.

Banerji, J., Olson, S. and Schaffner, W. (1983). A Lymphocyte−Specific Cellular Enhancer is Located Downstream of the Joining Region in Immunoglubulin Heavy Chain Genes. *Cell* **33**, 729−740.

Banerji, J., Rusconi, S. and Schaffner, W. (1981). Expression of a β−Globin Gene is Enhanced by Remote SV40 DNA Sequences. *Cell* **27**, 299−308.

Basham, T.Y., Bourgeade, M.F., Creasey, A.A. and Merigan, T.C. (1982). Interferon Increases HLA Synthesis in Melanoma Cells: Interferon−Resistant and −Sensitive Cell Lines. *Proc. natn. Acad. Sci., U.S.A.* **79**, 3265−3269.

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 Burkitts Lymphoma. *Cell* **34**, 779−787.

Benech, P., Mory, Y., Revel, M. and Chebath, J. (1985a). Structure of Two Forms of the Interferon−Induced (2'−5') Oligo A Synthetase of Human Cells Based on cDNAs and Gene Sequences. *EMBO. J.* **4**, 2249−2256.

Benech, P., Merlin, G., Revel, M. and Chebath, J. (1985b). 3' End Structure of the Human (2'−5') Oligo A Synthetase Gene: Prediction of Two Distinct Proteins with Cell Type−Specific Expression. *Nucl. Acids Res.* **13**, 1267−1281.

Benoist, C. and Chambon, P. (1981). *In vivo* Sequence Requirements of the SV40 Early Promoter Region. *Nature Lond.* **290**, 304−310.

Benton, W.D. and Davis, R.W. (1977). Screening λgt Recombinant Clones by Hybridisation to Single Plaques *in situ. Science, N,y.* **196**, 180−182.

Bienz, M. (1986). A CCAAT Box Confers Cell−Type−Specific Regulation on the *Xenopus hsp*70 Gene in Oocytes. *Cell* **46**, 1037−1042.

Bienz, M. and Pelham, H.R.B. (1986) Heat Shock Regulatory Elements Function as an Inducible Enhancer in the *Xenopus hsp*70 Gene and When Linked to a Heterologous Promoter. *Cell* **45**, 753−760.

Birnboim, H.C. and Doly, J. (1979). A Rapid Alkaline Extraction Procedure for Screening Recombinant Plasmid DNA. *Nucl. Acids Res.* **7**, 1513−1523.

Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston−Thompson, K., Faber, H.E., Furlong, L.−A., Grunwald, D.J., Keifer, D.O., Moore, D.D., Sheldon, E.L. and

Smithies, O. (1977). Charon Phages: Safer Derivatives of Bacteriophage Lambda for DNA Clining. *Science N.y.* 196, 161−169.
Immediate Early Gene of Human Cytomegalovirus. *Cell* 41, 521−530.

Boshart, M., Weber, F., Jahn, G., Dorsch−Hasler, K.. Fleckenstein, B. and Schaffner, W. (1985). A Very Strong Enhancer is Located Upstream of an Immediate Early Gene of Human Cytomegalovirus. *Cell* 41, 521−530.

Boulet, A.M., Erwin, C.R. and Rutter, W.J. (1986). Cell−specific Enhancers in the Rat Exocrine Pancreas. *Proc. natn. Acad., Sci. U.S.A.* 83, 3599−3603.

Bosze, Z., Thiesen, H.J and Charnay, P. (1986). A Transcriptional Enhancer with Specificity for Erythroid Cells is Located in the Long Terminal Repeat of the Friend Murine Leukaemia Virus. *EMBO. J.* 5, 1615−1623.

Brack, C., Nagata, S., Mantei, N. and Weissmann, C. (1981). Molecular Analysis of the Human Interferon−α Gene Family. *Gene* 15, 379−394.

Brammar, W.J. (1982). Vectors Based on Bacteriophage Lambda. *Genetic Engineering* (Acad. Press) 3, 52−81.

Branca, A.A. and Baglioni, C. (1981). Evidence that Type I and II Interferons have Different Receptors. *Nature Lond.* 294, 768.
Branca, A.A., Faltynek, C.R., D'Alessandro, S.B. and Baglioni, C. (1982). Interaction of Interferon with Cellular Receptor: Internalisation and Degradation of Cell−bound Interferon. *J. Biol. Chem.* 257, 13291−13296.

Brinster, R.L., Chen, H.Y., Warren, R., Sarthy, A. and Palmiter, R.D. (1982). Regulation of Metallothionein−Thymidine Kinase Fusion Plasmids Injected into Mouse Eggs. *Nature, Lond.* 296, 39−42.

Burke, J.F. and Ish−Horowicz, D. (1982). Expression of *Drosphila* Heat Shock Genes is Regulated in Rat−1 Cells. *Nucl. Acids Res.* 10 3821−3830.

Burrone, O.R. and Milstein, C. (1982). Control of HLA−A,B,C Synthesis and Expression in Interferon−Treated Cells. *EMBO.J.* 1, 345−349.

Canaani, D. and Berg, P. (1982). Regulated Expression of the Human Interferon β_1 Gene After Transduction into Cultured Mouse and Rabbit Cells. *Proc. natn. Acad. Sci., U.S.A.* 79, 5166−5170.

Capon, D.J., Shepard, H.M. and Goeddel, D.V. (1985). Two Distinct Families of Human and Bovine Interferon−α Genes are Coordinately Expressed and Encode Functional Poypeptides. *Mol. Cell. Biol.* 5, 768−779.

Cavalieri, R.L., Havell, E.A., Vilcek, J. and Pestka, S. (1977). Induction and Decay of Human Fibroblast Interferon mRNA. *Proc. natn. Acad. Sci., U.S.A.* 74, 4415−4419.

Chandler, V.L., Maler, B.A. and Yamamoto, K.R. (1983). DNA Sequences Bound Specifically by Glucocorticoid Receptor *In Vitro* Render a Heterologous Promoter Responsive *In Vivo. Cell* 33, 489−499.

Chandler, V.L., Maler, B.A. and Yamamoto, K.R. (1984). Are Glucocorticoid Receptor Binding Domains Regulatable "Enhancer" Elements? In "Gene Transfer and Cancer" (Ed. M. C. Pearson and N. Sternberg), pp. 303.

Chao, M.V., Mellon, P., Charnay, P., Maniatis, T. and Axel, R. (1983). The Regulated Expression of *B*−globin Genes Introduced into Mouse Erythroleukemia Cells. *Cell* 32, 483−493.

Charnay, P., Treisman, R., Mellon, P., Chao, M., Axel, R. and Maniatis, T. (1984). Difference in Human α- and B-globin Gene Expression in Mouse Erythroleukemia Cells: The Role of Intragenic Sequences. *Cell* 38, 251-263.

Chebath, J., Merlin, G., Metz, R., Benech, P. and Revel, M. (1983). Interferon-induced 56,000 Mr Protein and Its mRNA in Human Cells: Molecular Cloning and Partial Sequence of the cDNA. *Nucl. Acids Res.* 11, 1213-1226.

Cheng, T-S. E., Colonno, R.J. and Yin, F.H. (1983). Interferon Induction of Fibroblast Proteins with Guanylate Binding Activity. *J. Biol. Chem.* 258, 7746-7750.
Proc. natn. Acad. Sci. U.S.A., 72, 4361-4365.

Cochran, M.D. and Weissmann, C. (1984). Modular Structure of the β-Globin and *Tk* Promoters. *EMBO. J.* 3, 2453-2459.

Collins, T., Korman, A.J., Wake, C.T., Boss, J.M., Kappes, D.J., Fiers, W., Ault, K.A., Gimbrone Jr, M.A., Strominger, J.L. and Pober, J.S. (1984). Immune Interferon Activates Multiple Class II Major Histocompatability Complex Genes and the Associated Invariant Gene in Human Endothelial Cells and Dermal Fibroblasts. *Proc. natn. Acad. Sci., U.S.A.* 81, 4917-4921.

Colonno, R.J. and Pang, R.H.L. (1982). Induction of Unique mRNAs by Human Interferons. *J. Biol. Chem.* 257 9234-9237.

Content, J.,De Wit, L., Pierard, D., Derynk, R., DeClerq, E. and Fiers, W. (1982). Secretory Proteins Induced in Human Fibroblasts Under Conditions used for the Production of IFNβ. *Proc. natn. Acad. Sci. U.S.A.* 79, 2768-2772.

Content, J., Johnston, M.I., De Wit, L., De Maeyer-Guignard, J. and De Clercq, E. (1980). Kinetics and Distribution of Interferon mRNA in Interferon Primed and Unprimed Mouse L-929 Cells. *Biochem and Biophysi Res. Comm.* 96, 415-424.

Corces, V., Pellicer, A., Axel, R. and Meselson, M. (1981). Integration, Transcription and Control of a *Drosphila* Heat Shock Gene in Mouse Cells. *Proc. natn. Acad. Sci., U.S.A.* 78, 7038-7042.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980). Promoter Sequences of Eukaryotic Protein-Coding Genes. *Science* 209, 1406-1414.

Coveney, J., Scott, G., King, R., Burke, D.C. and Skup, D. (1984). Changes in the Conformation of the Interferon β Gene During Differentiation and Induction. *Biochem. and Biophys. Res. Comm.* 121, 290-296.

Degrave, W., Derynck, R., Tavernier, J., Haegeman, G. and Fiers, W. (1981). Nucleotide Sequence of the Chromosomal Gene for Human Fibroblast (β_1) Interferon and of the Flanking Region. *Gene* 14, 137-143.

Deininger, P.l, Jolly, D.J., Rubin, C.M., Freidman, T. and Schmid, C.W. (1981). Base Sequence Studies of 300 Nucleotide Renatured Repeated Human DNA Clones. *J. Mol. Biol.* 151, 17-33.

De Maeyer-Guignard. (1975). Delayed-Type Hypersensitivity to Sheep Red Blood Cells: Inhibition of Sensitization by Interferon. *Nature, Lond.* 190, 574-576.

De Maeyer-Guignard, J., Tovey, M.G., Gresser, I. and De Maeyer, E. (1978). Purification of Mouse Interferon by Sequential Affinity Chromotagraphy on Poly(U)-and Antibody-Agarose Columns. *Nature, Lond.* 271, 622-625.

Derynck, R., Content, J., DeClercq, E., Volckaert, G., Tavernier, J., Devos, R. and

Fiers, W. (1980). Isolation and Structure of a Human Fibroblast Interferon Gene. *Nature, Lond.* **285**, 542–546.

De Villiers, J. and Schaffner. W. (1981). A Small Segment of Polyoma Virus DNA Enhances the Expression of a Cloned β–Globin Gene Over a Distance of 1400 Base Pairs. *Nucl. Acids Res.* **9**, 6251–6264.
its Expression in Eukaryotic Cells. *Nucl. Acids Res.* **10**, 2487–2501.

Devos, R., Cheroute, H., Taya, Y., Degrave, W., Van Heuverswyn, H. and Fiers, W. (1982). Molecular Cloning of Human Immune Interferon cDNA and its Expression in Eukaryotic cells. *Nucl. Acids Res.* **10**, 2487–2501.

Dierks, P., van Ooyen, A., Cochran, M.D., 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 3T6 Cells. *Cell* **32**, 695–706.

Dijkema, R., Pouwels, P., deReus, A. and Schellekens (1984). Structure and Expression in *Escherichia coli* of a Cloned Rat Interferon–α Gene. *Nucl. Acids Res.* **12**, 1227–1242.

Docherty, A.J.P., Lyons, A., Smith, B.J., Wright, E.M., Stephens, P.E. and Harris, T.J.R. (1985). Sequence of Human Tissue Inhibitor of Metalloproteinases and its Identity to Erythroid–potentiating Activity. *Nature, Lond.* **318**, 66–69.

Dynan, W.S. (1986). Promoters for Housekeeping Genes. *TIG.* **2**, 196–197.

Dynan, W.S., Saffer, J.D., Lee, W.S., and Tjian, R. (1985). Transcription Factor Sp1 Reconises Promoter Sequences from the Monkey Genome that are Similar to the Simian Virus 40 Promoter. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4915–4919.

Dynan, W.S. and Tjian, R. (1983). The Promoter–Specific Transcription Factor Sp1 Binds to Upstream Sequences in the SV40 Early Promoter. *Cell* **35**, 79–87.

Edlund, T., Walker, M.D., Barr, P.J. and Rutter, W.J. (1985). Cell–Specific Expression of the Rat Insulin Gene: Evidence for Role of Two Distinct 5'Flanking Elements. *Science, N.y.* **230**, 912–916.

Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Fotget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, 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. (1985a). Close Link Between Reduction of c–*myc* Expression by Interferon and G_0/G_1 Arrest. *Nature, Lond.* **313**, 597–600.

Einat, M., Resnitzky, D. and Kimchi, A. (1985b). Inhibitory Effects of Interferon on the Expression of Genes Regulated by Platelet–derived Growth Factor. *Proc. natn. Acad., Sci., U.S.A.* **82**, 7608–7612.

EMBL Nucleotide Data Library, European Molecular Biology Laboratory, Postfach 102209, D–6900 Heidelburg–FRG.

Enock, T., Zinn, K. and Maniatis, T. (1986). Activation of the Human β–Interferon Gene Requires an Interferon Inducible Factor. *Mol. Cell. Biol.* **6**, 801–810.

Ephrussi, A., Churgh, G.M., Tonegawa, S. and Gilbert, W. (1985). B Lineage–Specific Interactions of an Immunoglobulin Enhancer with Cellular Factors *in Vivo. Science, N.y.* **227**, 134–140.

Faltynek, C.N., Branca, A.A., McCandless, S. and Baglioni, C. (1983). Characterisation of an Interferon Receptor on Human Lymphoblastoid Cells. *Proc. natn. Acad. Sci., U.S.A.* 80, 3269–3273.

Favaloro, J., Treisman, R. and Kamen, R. (1980). Transcription Maps of Polyoma Virus–specific RNA: Analysis by Two–dimensional Nuclease S1 Gel Mapping. *Methods Enzymol.* 65, 718–749.

Feinberg, A. and Vogelstein, B. (1984). *Annl. Biochem.* 137, 266.

Fellous, M., Nir, U., Wallach, D., Merlin, G., Rubinstein, M. and Revel, M. (1982). Interferon–dependent Induction of mRNA for the Major Histocompatability Antigens in Human Fibrolasts and Lymphoblastoid Cells. *Proc. natn. Acad. Sci., U.S.A.* 79, 3082–3086.

Fellous, M. and Rosa, F. (1985). Interferons and HLA Expression. *Symposium on "Interferons as Regulators of Growth and Function".* BACR and IACR Joint Meeting.

Freidman, R.L., Manly, S.P., McMahon, M., Kerr, I.M. and Stark, G.R. (1984). Transcriptional and Posttranscriptional Regulation of Interferon–Induced Gene Expression in Human Cells. *Cell* 38, 745–755.

Freidman, R.L. and Stark, G.R. (1985). α–Interferon–Induced Transcription of HLA and Metallothionein Genes Containing Homologous Upstream Sequences. *Nature, Lond.* 314, 637–639.

Frischauf, A–M., Lehrach, H. Postka, A. and Murray, N. (1983) Lambda Replacement vectors carrying Polylinker Sequences. *J. Mol. Biol.* 170, 827–842.

Fritsch, E.F., Caron, R.M. and Maniatis, T. (1980). Molecular Cloning and Characterisation of the Human β–like Globin Gene Cluster. *Cell* 19, 959

Fujisawa, J., Iwakura, Y. and Kawade, Y. (1978). Nonglycosylated Mouse L Cell Interferon Produced by the Action of Tunicamycin. *J. Biol. Chem.* 253, 8677–8679.

Fujita, T., Ohno, S., Yasumitsu, H. and Taniguchi, T. (1985). Delimitation and Properties of DNA Sequences Required for the Regulated Expression of Human Interferon–β Gene. *Cell* 41, 489–496.

Fuke, M., Hendrix, L.C. and Bollon, A.P. (1984). Pseudogene IFN–αL: Removal of the Stop Codon in the Signal Sequence Permits Expression of Active Human Interferon. *Gene* 32, 135–140.

Galabru, J. and Hovanessian, A.G. (1985). Two Interferon–induced Proteins are Involved in the Protein Kinase Complex Dependent on Double Stranded RNA. *Cell* 43, 685–694.

Garel, A. and Axel, R. (1976). Selective Digestion of Transcriptionally Active Ovalbumin Genes from Oviduct Nuclei. *Proc. natn. Acad. Sci., U.S.A.* 73, 3966–3970.

Gasson, J.D., Golde, D.W., Kaufman, S.E., Westbrook, C.A., Hewick, R.M., Kaufman, R.J., Wong, G.G., Temple, P.A., Leary, A.C., Brown, E.L., Orr, E.C. and Clark, S.C. (1985). Molecular Characterisation and Expression of the Gene Encoding Human Erythroid–potentiating Activity. *Nature, Lond.* 351, 768–771.

Gidoni, D., Dynan, W.S. and Tjian, R. (1984) Multiple Specific Contacts Between a Mammalian Transcription Factor and its Cognate Promoters. *Nature, Lond.* 312, 409–413.

Gillies, S.D., Morrison, L.S., Oi, V.T. and Tonegawa, S. (1983) A Tissue–specific Transcription Enhancer is Located in the Major Intron of a Rearranged Immunoglobulin

Heavy Chain Enhancer Gene. *Cell* 33, 717−728.

Glanville, N., Durnham, D.M. and Palmiter, R.D. (1981). Structure of Mouse Metallothionein−1−gene and its mRNA. *Nature Lond.* 292, 267−269.

Goeddel, D.V., Shephard, H.M., Yelverton, E., Leung, D., Crea, R., Sloma, A. and Pestka, S. (1980b). Synthesis of Human Fibroblast Interferon by *E.coli*. *Nucl. acids Res.* 8, 4057−4074.

Goeddel, D.V., Leung, D.W., Dull, T.J., Gross, M., Lawn, R.M., McCandliss, R., Seeburg, P.H., Ullrich, A., Yelverton, E. and Gray, P.W. (1981). The Structure of Eight Distinct Cloned Human Leukocyte Interferon cDNAs. *Nature, Lond.* 290, 20−26.

Goeddel, D.V., Yelverton, E., Ullrich, A., Heyneker, H.L., Miozzari, G., Holmes, W., Seeburg, P.H., Dull, T., May, L., Stebbing, N., Crea, R., Maeda, S., McCandliss, R., Sloma, A., Tabor, J.M., Gross, M., Familletti, P.C. and Pestka, S. (1980a). Human Leukocyte Interferon Produced by *E.coli* is Biologically Active. *Nature, Lond.* 287, 411−416.

Golding, G.B. and Glickman, B.W. (1985). Sequence−directed Mutagenesis: Evidence from a Phlogenetic Hitory of Human α−Interferon Genes. *Proc. natn. Acad. Sci., U.S.A.* 1985, 8577−8581.

Goodbourn, S., Zinn, K. and Maniatis, T. (1985). Gene Expression is Regulated by an Inducible Enhancer Element. *Cell* 41, 509−520

Goodbourn, S., Burstein, H. and Maniatis, T. (1986). The Human β Interferon Gene Enhancer is Under Negative Control. *Cell* 45, 601−610.

Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982). Recombinant Genomes Which Express Chloramphenicol Acetyltransferase in Mammalian Cells. *Mol. Cell. Biol.* 2, 1044−1051.

Gorman, C., Padmanabhan, R. and Howard, B. (1983). High Efficiency DNA−mediated Transformation of Primate Cells. *Science N.y.* 221, 551−553.

Graves, B.J., Johnson, P.F. and McKnight, S.L. (1986). Homologous Recognition of a Promoter Domain Common to the MSV LTR and the HSV tk Gene. *Cell* 44, 565−576.

Gray, P.W. and Goeddel, D.V. (1982). Structure of the Human Immune Interferon Gene. *Nature, Lond.* 298, 859−863.

Gray, P.W. and Goeddel, D.V. (1983). Cloning and Expression of Murine Immune Interferon cDNA. *Proc. natn. Acad. Sci., U.S.A.* 80, 5842−5846.

Gray, P.W., Leung, D.W., Pennica, D., Yelverton, E., Najarian, R., Simonsen, C.C., Derynck, R., Sherwood, P.J., Wallace, D.M., Berger, S.L., Levinson, A.D. and Goeddel, D.V. (1982). Expression of Human Immune Interferon cDNA in *E.coli* and Monkey Cells. *Nature, Lond.* 295, 503−508.

Gresser, I., De Maeyer−Guignard, J., Tovey, M.G. and De Maeyer, E. (1979). Electrophoretically Pure Mouse Interferon Exerts Multiple Biologic Effects. *Proc. natn. Acad. Sci., U.S.A.* 76, 5308−5312.

Gribaudo, G., Cofano, F., Prat, M., Baiocchi, C., Cavallo, G. and Landolfo, S. (1985). Natural Murine IFN γ: Evidence for Post−translational Proteolytic Processing. *J. Biol. Chem.* 260, 9936−9940.

Gross, G., Ulrich, M., Bruns, W., Grosveld, F., Dahl, H−H.M. and Collins, J. (1981). The Structure of a Thirty Six Kilobase Region of the Human Chromosome including the

Fibroblast Interferon Gene IFN−β. *Nucl. Acids Res.* 9, 2415−2507.

Grosveld, G.C., Shewmaker, G.K., Jat, P. and Flavell, R.A. (1981). Localization of DNA Sequences Necessary for Transcription of the Rabbit β−Globin Gene *in Vitro*. *Cell* 25, 215−226.

Grosveld, F.G., Dahl, H.−H.M., deBoer, E. and Flavell, R.A. (1981). Isolation of β−globin−related Genes from a Human Cosmid Library. *Gene*, 13, 227−237.

Grosveld, F.,G., Lund, T., Murray, E.J., Mellor, A.L., Dahl, H.H.M. and Flavell, R.A. (1982). The Construction of Cosmid Libraries Which Can be Used to Transform Eukaryotic Cells. *Nucl. Acids Res.* 10, 6715−6732

Grunstein, M. and Hogness, D.G. (1975). Colony Hybridisation: A Method for the Isolation of Cloned DNAs that Contain a Specific Gene. *Proc. natn. Acad. Sci. U.S.A.* 72, 3961−3965.

Gruss, P., Dhar, R. and Khoury, G. (1981) Simian Virus 40 Tandem Repeated Sequences as an Element of the Early Promoter. *Proc. natn. Acad. Sci. U.S.A.* 78, 943−947.

Hall, A. and Brown, R. (1985). Human N−*ras*: cDNA Cloning and Gene Structure. *Nucl. Acids Res.* 13, 5255−5268.

Hannigan, G. and Williams, B.R.G. (1986). Transcriptional Regulation of Interferon−responsive Genes is Closely Linked to Interferon Receptor Occupancy. *EMBO.J.* 5, 1607−1613.

Hauptmann, R. and Swetly, P. (1985). A Novel Class of Human Type I Interferons. *Nucl. Acids Res.* 13, 4739−4749.

Hauser, H., Dinter, H. and Collins, J. (1984). Analysis of the Control Regions of the Human Interferon β Gene. *The Biology of the Interferon System* (Elsevier) pp21−26

Hauser, H., Gross, G., Bruns, W., Hochkeppel, H−K., Mayr, U. and Collins, J. (1982). Inducibility of the Human β−Interferon Gene in Mouse L−cell Clones. *Nature, Lond.* 297, 650−654.

Hen, R., Borrelli, E., Sassone−Corsi, P. and Chambon, P. (1983). An Enhancer is Located 340 base pairs Upstream from the Adenovirus−2 E1A Capsite. *Nucl. Acids Res.* 11, 8747−8760.

Henco, K., Brosius, J., Fujisawa, Fujisawa, A., Haynes, J.R., Hochstadt, J., Kovavic, T., Pasek, M., Shambrock, A., Schmid, J., Todokoro, K., Walchli, M., Nagata, S. and Weissmann, C. (1985). Structural Relationship of Human Interferon Alpha Genes and Pseudogenes. *J. Mol. Biol.* 185, 227−260.

Higashi, Y. (1984). Changes of Chromatin Conformation Around Mouse Interferon−β Gene Associated with Induction of Interferon Synthesis. *Nucl. Acids Res.* 13, 5157−5172.

Higashi, Y., Sokawa, Y., Watanabe, Y., Kawade, Y., Ohno, S., Takaoka, C. and Taniguchi, T. (1983). Structure and Expression of a Cloned cDNA for Mouse Interferon−β. *J. Biol. Chem.* 258, 9522−9529.

Hiscott, J., Cantell, K. and Weissmann, C. (1984). Differential Expression of Human Interferon Genes. *Nucl. Acids Res.* 12, 3727−3746.

Hohn, B. (1979). *In Vitro* Packaging of λ and Cosmid DNA. *Methods in Enzymol.* 68, 299−309.

Hohn, B. and Murray, K. (1977). Packaging Recombinant DNA Molecules into Bacteriophage Particles *in Vitro*. *Proc. natn. Acad. Sci., U.S.A.* **74**, 3259–3263.

Houghton, M., Jackson, I.J., Porter, A.G., Doel, S.M., Catlin, G.H., Barber, C. and Carey, N.H. (1981). The Absence of Introns Within a Human Fibroblast Interferon Gene. *Nucl. Acids. Res.* **9**, 247–266.

Hovanessian, A.G. and Kerr, I.M. (1979). The (2'–5') Oligoadenylate (pppA2'–5'A2'–5'A) Synthetase and Protein Kinase(s) from Interferon–treated Cells. *Eur. J. Biochem.* **93**, 515–526.

Hynes, N.E., Kennedy, N., Rahmsdorf, U. and Groner, B. (1981). Hormone–responsive Expression of an Endogenous Proviral Gene of Mouse Mammary Tumour Virus after Molecular Cloning and Gene Transfer into Cultured Cells. *Proc. natn. Acad. Sci., U.S.A.* **78**, 2038–2042.

Hynes, N., Van Ooyen, A.J.J., Kennedy, N., Herrlich, P., Ponta, H. and Groner, B. (1983). Subfragments of the Large Terminal Repeat cause Glucocorticoid–Responsive Expression of Mouse Mammary Tumour Virus and of an Adjacent Gene. *Proc. natn. Acad. Sci., U.S.A.* **80**, 3637–3641.

Ishii, S., Merlino, G.T. and Pastan, R. (1985). Promoter Region of the Human Harvey *reas* Proto–oncogene: Similarity to the EGF Receptor Proto–Oncogene Promoter. *Science* **230**, 1378–1381.

Jeffreys, A.J., Barrie, P.A., Harris, S., Fawcett, D.W. and Nugent, Z.J. (1982). Isolation and Sequence Analysis of a Hybrid δ–Globin Pseudogene from the Brown Lemur. *J. Mol. Biol.* **156**, 487–503.

Jonak, G.J. and Knight Jr., E. (1984). Selective Reduction of c–*myc* mRNA in Daudi Cells by Human β–Interferon. *Proc. natn. Acad. Sci., U.S.A.* **81**, 1747–1750.

Jones, K.A., Yamamoto, K.R. and Tjian, R. (1985). Two Distinct Transcription Factors Bind to the HSV Thymidine Kinase Promoter *In Vitro*. *Cell* **42** 559–572 1985.

Joshi, A.R., Sarkar, F.H. and Gupta, S.L. (1982). Interferon Receptors:– Cross Linking of Human Leukocyte IFNα–2 to its Receptor on Human Cells. *J. Biol. Chem.* **257**, 13884–13887.

Kadonaga, J.T, Jones, K.A. and Tjian, R. (1986). Promoter–specific Activation of RNA Polymerase II Transcription by Spl. *TIBS*. **11**, 20–23.

Karin, M., Cathala, G. Nguyen–Huu, M.C. (1983). Expression and Regulation of a Human Metallothionein Gene Carried on an Autonomously Replicating Shuttle Vector. *Proc. natn. Acad. Sci., U.S.A.* **80**, 4040–4044.

Karin, M., Haslinger, A., Holtgreve, H., Richards, R.I., Krauter, P., Westphal, H.M. and Beato, M. (1984a). Characterization of DNA Sequences Through Which Cadmium and Glucocorticoid Hormones Induce Human Metallothionein–II_A Gene. *Nature, Lond.* **308**, 513–519.

Karin, M., Haslinger, A., Holrgreve, H., Cathala, G., Slater, E. and Baxter, J.D. (1984b). Activation of a Heterologous Promoter in Response to Dexamethasone and Cadmium by Metallothionein Gene 5'–Flanking DNA. *Cell* **36**, 371–379.

Kelley, K.A., Kozak, C.A., Dandoy, F., Sor, F., Skup, D., Windass, J.D., De Maeyer–Guignard, J., Pitha, P.M. and De Maeyer, E. (1983). Mapping of Murine Interferon–α Genes to Chromosome 4. *Gene.* **26**, 181–188.

Kelley, K.A. and Pitha, P.M. (1985a). Characterisation of a Mouse Interferon Gene Locus I: Isolation of a Cluster of Four α Interferon Genes. *Nucl. Acids Res.* **13**, 805–823.

Kelley, K.A. and Pitha, P.M. (1985b). Characterisation of a Mouse Interferon Gene Locus II: Differential Gene Expression of α–Interferon Genes. *Nucl. Acids Res.* **13**, 825–839.

Kelly, J.M., Gilbert, C.S., Stark, G.R. and Kerr, I.M. (1985). Differential Regulation of Interferon Induced mRNAs and c–*myc* mRNA by α– and γ Interferons. *Eur. J. Biochem.* **153**, 367–371.

Kelly, J.M., Porter, A.C.G., Chernajovsky, Y., Gilbert, C.S., Stark, G.R. and Kerr, I.M. (1986). Characterisation of a Human Gene Inducible by α– and β– Interferons and its Expression in Mouse Cells. *EMBO. J.* **5**, 1601–1606.

Kelker, H.C., Yip, Y.K., Anderson, P. and Vilcek, J. (1983) Effects of Glycosidase Treatment on the Physicochemical Properties and Biological Activity of Human Interferon–γ. *J. Biol. Chem.* **258**, 8010–8013.

Khoury, G. and Gruss, P. (1983). Enhancer Elements. *Cell* **33**, 313–314.

Kimura, A., Israel, A., Le Bail, O. and Kourilsky, P. (1986). Detailed Analysis of the Mouse H–$2k^b$ Promoter: Enhancer–like Sequences and Their Role in the Regulation of Class I Gene Expression. *Cell* **44**, 261–272.

Knight Jr., E. (1976). Interferon: Purification and Initial Characterisation from Human Diploid Cell. *Proc. natn. Acad. Sci. U.S.A.* **73**, 520–523.

Knight Jr., E., Hunkapiller, M.W., Korant, B.D., Hardy, R.W.F. and Hood, L.E. (1980). Human Fibroblast Interferon: Amino Acid Analysis and Amino Terminal Amino Acid Sequence. *Science, N.y.* **207**, 525–526.

Koeffler, H.P., Ranyard, J., Yelton, L., Billing, R. and Bohman, R. (1984). γ–Interferon Induces Expression of the HLA–D Antigens on Normal and Leukemic Human Myeloid Cells. *Proc. natn. Acad. Sci., U.S.A.* **81**, 4080–4084.

Kriegler, M. and Botchan, M. (1983). Enhanced Transformation by a SV40 Recombinant Virus Containing a Harvey Murine Sarcoma Virus LTR. *Mol. Cell. Biol.* **3**, 325–339.

Kudo, I., Leineweber, M. and RajBhandary, U.L. Site–specific Mutagenesis on Cloned DNAs: Generation of a Mutant of *Escherichia coli* Tyrosine Suppressor tRNA in which the G–T–T–C corresponding to the Universal G–T–ψ–C Sequence of tRNAs is changed to G–A–T–C. *Proc. natn. Acad. Sci., U.S.A.* **78**, 4752–4757.

Kunkel, T. (1985). Rapid and Efficient Site–directed Mutagenesis Without Phenotypic Selection. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492.

Kusari, J. and Sen, G.C. (1986). Regulation of Synthesis and Turnover of an Interferon Inducible mRNA. *Mol. Cell. Biol.* **6**, 2062–2067.

Lawn, R.M., Adelman, J., Dull, T.J., Gross, M., Goeddel, D. and Ullrich, A. (1981). DNA Sequence of Two Closely Linked Human Leukocyte Interferon Genes. *Science, N.y.* **212**, 1159–1162.

Laimins, L.A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982). Host–specific Activation of Transcription by Tandem Repeats from Simian Virus 40 and Moloney murine Sarcoma Virus. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6453–6457.

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. natn. Acad. Sci. U.S.A.* **83**, 3151–3155.

Lawn, R.M., Adelman, J., Franke, A.E., Houck, C.M., Gross, C.M., Najarian, R. and Goeddell, D.V. (1981b). Human Fibroblast Interferon Gene Lacks Introns. *Nucl. Acids Res.* **9**, 1045–1052.

Lawn, R.M., Gross, M., Houck, C.M., Franke, A.E., Gray, P.V. and Goeddel, D.V. (1981a). DNA Sequence of a Major Human Leukocyte Interferon Gene. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5435–5439.

Larner, A.C., Jonak, G., Cheng, Y.S.E., Korant, B., Knight, E. and Darnel Jr, J.E. (1984). Transcriptional Induction of Two Genes in Human Cells by β–interferon. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6733–6737.

Lebleu, B., Hubert, E., Content, J., De Wit, L., Braude, I.A. and De Clercq, E. (1978). Translation of Mouse Interferon mRNA in *Xenopus Laevis* Oocytes and in Rabbit Reticulocyte Lysates. *Biochem. Biophys. Res. Com.* **82**, 665–673.

Lee, F., Mulligan, R., Berg, P. and Ringold, G. (1981). Glucocorticoids Regulate Expression of Dihydrofolate Reductase cDNA in Mouse Mammary Tumour Virus Chimaeric Plasmids. *Nature, Lond.* **294**, 228–232.

Lengyel, P. (1982). Biochemistry of Interferons and Their Actions. *Ann. Rev. Biochem.* **51**, 251–282

Levy, W.P., Rubinstein, M., Shively, J., Del Valle, U., Lai, C–Y., Moschera, J., Brink, L., Gerber, L., Stein, S. and Pestka, S. (1981). Amino Acid Sequence of a Human Leukocyte Interferon. *Proc. natn. Acad. Sci., U.S.A.* **78**, 6186–6190.

Loenen, W.A.M. and Brammar, W.J. (1980). A Bacteriophage Lambda Vector for Cloning Large DNA Fragments made with Several Restrction Enzymes. *Gene* **10**, 249–259.

Lopata, M.A., Cleveland, D.W and Sollner–Webb, B, B. (1984). High Level Expression of a Chlramphenicol Acetyltransferase Gene by DEAE–dextran mediated DNA Transfection Coupled with a Dimethyl Sulphoxide or Glycerol Shock Treatment. *Nucl. Acids Res.* **12**, 5707–5717.

Lovett, M., Cox, D.R., Yee, D., Boll, W., Weissmann, C., Epstein, C.J. and Epstein, L.B. (1984). The Chromosomal Location of Mouse Interferon α Genes. *EMBO. J.* **3**, 1643–1646.

Lowry, O.H., Roubrough, N.J *et al.* (1951). *J. Biol. Chem.* **193**, 265.

Lund, B., Edlund, T., Lindenmaier, W., Ny, T., Collins, J., Lundgren, E. and von Gabain, A. (1984). Novel Cluster of α–Interferon Gene Sequences in a Placental Cosmid DNA Library. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2435–2439.

Lusky, Berg, Weiher and Botchan. (1983). Bovine Papilloma Virus Contains an Activator of Gene Expression at the Distal End of the Early Transcription Unit. *Mol. Cell. Biol.* **3**, 1108–1122.

Luster, A.D., Unkeless, J.C. and Ravetch, J.V. (1985). γ–Interferon Transcriptionally Regulates an Early–Response Gene Containing Homology to Platelet Proteins. *Nature, Lond.* **315**, 672–676.

Maniatis, T., Hardison, R.C., Lacy, E., Lamer, J.,, O'Connel, C, Quon, D., Sim, D.K. and Efstratiadis, A. (1978). The Isolation of Structural Genes from Libraries of Euakaryotic DNA. *Cell* **15**, 687–701.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). *Molecular Cloning* (A Laboratory Manual), Cold Spring Harbour.

Majors, J. and Varmus, H.E. (1983). A Small Region of Mouse Mammary Tumor Virus Long Terminal Repeat confers Glucocorticoid Hormone Regulation on a Linked Heterologous Gene. *Proc. natn. Acad. Sci., U.S.A.* **80**, 5866–5870.

Mantei, N. and Weissmann, C. (1982). Controlled Transcription of a Human α–Interferon Gene Introduced into Mouse L Cells. *Nature, Lond.* **297**, 128–132.

Mantei, N., Schwarzstein, M., Streuli, M., Pasem, S., Nagata S. and Weissmann. C. (1980). The Nucleotide Sequence of a Cloned Human Leukocyte Interferon cDNA. *Gene* **10**, 1–10.

Mark, D.F., Lu, S.D., Creasey, A.A., Yamamoto, R. and Lin, L.S. (1984). Site–specific Mutagenesis of the Human Fibroblast Interferon Gene. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5662–5666.

Maroteaux, L., Kahana, C., Mory, Y., Groner, Y. and Revel, M. (1983). Sequences Involved in the Regulated Expression of the Human Interferon–β_1 Gene in Recombinant SV40 DNA Vectors Replicating in Monkey Cells. *EMBO. J.* **2**, 325–332.

Mayo, K.E., Warner, R. and Palmiter, R.D. (1982). The Mouse Metallothionein–I Gene is Transcriptionally Regulated by Cadmium Following Transfection into Human or Mouse Cells. *Cell,* **29**, 99–108.

McConologue, L., Gupta, M., Wu, L. and Coffino, P. (1984). Molecular Cloning and Expression of the Mouse Ornithine Decarboxylase Gene. *Proc. natn. Acad. Sci. U.S.A.* **81**, 540–544.

McCormick, FF., Trahey, M. Innis, Dieckmann, B. and Ringold, G.M. (1984). Inducible Expression of Amplified Human β Interferon Genes in CHO Cells. *Mol. Cell. Biol.* **4**, 166–172.

McKnight, S.L. (1982). Functional Relationships between Transcriptional Control Signals of the Thymidine Kinase Gene of Herpes Simplex Virus. *Cell* **31**, 355–365.

McKnight, S.L. and Gavis, E.R. (1980). Expression of the Herpes Thymidine Kinase Gene in *Xenopus laevis* Oocytes; an Assay for the Study of Deletion Mutants Constructed *in Vitro*. *Nucl. Acids Res.* **8**, 5931–5948.

McKnight, S.L., and Kingsbury, R. (1982). Transcriptional Control Signals of a Eukaryotic Protein–Coding Gene. *Science, N.y.* **217**, 316–324.

McKnight, S.L., Gavis, E.R., Kingsbury, R. and Axel, R. (1981). Analysis of Transcriptional Regulatory Signals of the HSV Thymidine Kinase Gene; Identification of an Upstream Control Region. *Cell* **25**, 385–398.

McKnight, S.L., Kingsbury, R.C., Spence, A. and Smitch, M. (1984). The Distal Transcription Signals of the Herpes Virus tk Gene Share a Common Hexanucleotide Control Sequence. *Cell* **37**, 253–262.

McKnight, S.L. and Tjian, R. (1986). Transcriptional Selectivity of Viral Genes in Mammalian Cells. *Cell* **46**, 795–805.

Meinkoth, J. and Wahl, G. (1984). Hbridisation of Nucleic Acids Immobilised on Solid Supports. *Anal. Biochem.* **138**, 267–284

Mellon, P., Parker, V., Gluzmann, Y. and Maniatis, T. (1981). Identification of DNA

Sequences Required for Transcription of the Human α1−Globin Gene in a New SV40 Host−Vector System. *Cell* **27**, 279−288.

Mercola, M., Goverman, J., Mirell, C. and Calame, K. (1985). Immunoglobulin Heavy−Chain Enhancer Requires One or More Tissue−Specific Factors. *Science, N.y.* 277, 266−270.

Merlin, G., Chebath, J., Benech, P., Metz, R. and Revel, M. (1983). Molecular Cloning and Sequence of Partial cDNA for Interferon−Induced (2'−5') Oligo(A) Synthetase mRNA from Human Cells. *Proc. natn. Acad. Sci., U.S.A.* **80**, 4904−4908.

Messing, J. (1983). New M13 Vectors for Cloning. *Methods Enzymol.* **101**, 20−89.

Moore, R.N., Larsen, H.S., Horokov, D.W. and Rouse, B.T. (1984). Endogenous Regulation of Macrophage Proliferative Expansion by Colony−Stimulating Factor Induced Interferon. *Science, N.y.* **232**, 178−180.

Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M.P. and Chambon, P. (1981). The SV40 72base pair Repeat has a Striking Effect on Gene Expression both in SV40 and Other Chimeric Recombinants. *Nucl. Acids Res.* **9**, 6047−6068.

Mory, Y., Chernajouvsky, Y., Feinstein, S.I., Chen, L., Nir, U., Weissenbach, J., Malpiece, Y., Tiollais, P., Marks, D., Ladner, M., Colby, C. and Revel, M. (1981). Synthesis of Human Interferon β_1 in *Escherichia coli* Infected by a Lambda Phage Recombinant Containing a Human Genomic Fragment. *Eur. J. Bioch.* **120**, 197−202.

Murray, N.E., Brammar, W.J. and Murray, K. (1977). Lambdoid Phages that Simplify the Recovery of *In Vitro* Recombinants. *Mol. Gen. Genet.* **150**, 53.

Nagata, S., Mantei, N. and Weissmann, C. (1980a). The Structure of One of the Eight or More Distinct Chromosomal Genes for Human Interferon α. *Nature, Lond.* **287**, 401−408.

Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsodi, J., Boll, W., Cantell, K. and Weissmann, C. (1980b). Synthesis in *E.coli* of a Polypeptide with Human Leukocyte Interferon Activity. *Nature, Lond.* **284**, 316−320.

Naylor, S.L., Lalley, P.A., Sakaguchi, A.Y. and Gray, P.W. (1984). *Cytogenet. Cell Genet.* **37**, 550

Ohno, S. and Taniguchi, T. (1981). Structure of a Chromosomal Gene for Human Interferon β. *Proc. natn. Acad. Sci., U.S.A.* **78**, 5305−5309.

Ohno, S. and Taniguchi, T. (1982). Inducer−Responsive Expression of the Cloned Human Interferonβ_1 Gene Introduced into Cultured Mouse Cells. *Nucl. Acids Res.* 10, 967−977.

Ohno, S. and Taniguchi, T. (1983). The 5' Flanking Sequence of Human Interferon−β_1 Gene is Responsible for Viral Induction of Transcription. *Nucl. Acids Res.* **11**, 5403−5412.

Owerbach, D., Rutter, W.J, Shows, T.B., Gray, P., Goeddel, D.V. and Lawn, R.M. (1981). Leukocyte and Fibroblast Interferon Genes are Located on Human Chromosome 9. *Proc. natn. Acad. Sci., U.S.A.* **78**, 3123−3127.

Parker, C.S. and Topol, J. (1984). A *Drosophila* RNA Polymerase II Transcription Factor Binds to the Regulatory Site of an *hsp*70 Gene. *Cell* **37**, 273−283.

Pelham, H.R.B. (1982). A Regulatory Upstream Promoter Element in the *Drosophila* Hsp 70 Heat−Shock Gene. *Cell* **30**, 517−528.

Pelham, H.R.B. and Bienz, M. (1982). A Synthetic Heat−shock Promoter Element Confers Heat−Inducibility on the *Herpes Simplex* Virus Thymidine Kinase Gene. *EMBO. J.* 1, 1473−1477.

Pellicer, A., Robins, D., Wold, B., Sweet, R., Jackson, J., Lowy, I., Roberts, J.M., Sim, G.K., Silverstein, S. and Axel, R. (1980). Altering Genotype and Phenotype by DNA−Mediated Gene Transfer. *Science, N.y.* **209**, 1414−1422.

Pellicer, A. Wigler, M. Axel, R. and Silverstein, S. (1978). The Transfer and Stable Integration of the HSV Thymidine Kinase Gene into Mouse Cells. *Cell* **14**, 133−141.

Pestka, S. (ed. 1981). Interferons. *Methods Enzymol.* **78**

Picard, D. and Schaffner, W. (1984). A Lymphocyte−specific Enhancer in the Mouse Immunoglobulin K Gene. *Nature, Lond.* **307**, 80−82.

Proudfoot, N.J. and Brownlee, G.G. (1976). 3' Non−coding Region Sequences in Eukaryotic Messenger RNA. *Nature, Lond.* **263**, 211−214.

Prat, M., Gribaudo, G., Comogeio, P.M., Cavallo, G. and Landalfo, S. (1984). Monoclonal Antibodies Against Murine γ Interferon. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4515−4519.

Ptashne, M. (1986). Gene Regulation by Proteins Acting Nearby and at a Distance. *Nature, Lond.* **322**, 697−201

Queen, C. and Baltimore, D. (1983). Immunoglobulin Gene Transcription is Activated by Downstream Sequence Elements. *Cell* **33**, 741−748.

Ragg, H. and Weissmann, C. (1983). Not More Than 117 base pairs of 5'−flanking Sequence are Required for Inducible Expression of a Human IFN−α Gene. *Nature, Lond.* **303**, 439−442.

Raj, N.B.K., Fernie, B.F. and Pitha, P.M. (1979). Correlation between the Induction on Mouse Interferon and the Amount of Its mRNA. *Eur. J. Biochem.* **98**, 215−221.

Raj, N.B.K. and Pitha, P.M. (1977). Relationship between Interferon Production and Interferon messenger RNA Synthesis in Human Fibroblasts. *Proc. natn. Acad. Sci., U.S.A.* **74**, 1483−1487.

Raj, N.B.K. and Pitha, P.M. (1981). Analysis of Interferon mRNA in Human Fibroblast Cells Induced to Produce Interferon. *Proc. natn. Acad. Sci., U.S.A.* **78**, 7426−7430.

Raj, N.B.K. and Pitha, P.M. (1983). Two Levels of Regulation of β−interferon Gene Expression in Human Cells. *Proc. natn. Acad. Sci., U.S.A.* **80**, 3923−3927.

Rashidbaigi, A., Langer, J.A., Jung, V., Jones, C., Morse, H.G., Tischfield, J.A., Trill, J.J., Kung, H−F. and Pestka, S. (1986). The Gene for the Human Immune Interferon Receptor is Located on Chromosome 6. *Proc. natn. Acad. Sci., U.S.A.* **83**, 384−388.

Razin, A., Hirose, T., Itakura, K. and Riggs, A.D. (1978). Efficient Correction of a Mutation by Use of Chemically Synthesised DNA. *Proc. natn. Acad. Sci., U.S.A.* **75**, 4268−4270.

Remmers, R.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.

Resnitzky, D., Yarden, A., Zipori, D. and Kimchi, A. (1986). Autocrine β−Related Interferon Controls c−*myc* Suppression and Growth Arrest during Hematopoietic Cell

Differentiation. *Cell* **46**, 31–40.

Revel, M., Bash, D. and Ruddle, F.H. (1976). Antibodies to a Cell–Surface Component Coded by Human Chromosome 21 Inhibit Action of Interferon. *Nature, Lond.* **260**, 139–141.

Richards, R.I., Heguy, A. and Karin, M. (1984). Structural and Functional Analysis of the Human Metallothionein–1_A–gene: Differential Induction by Metal Ions and Glucocorticoids. *Cell* **37**, 263–272.

Rigby, P.J.W., Dieckmann, M., Rhodes, C. and Berg, P. (1977). Labellind Deoxyribonucleic Acid to high Specific Activity *in Vitro* by Nick Translation with DNA Polymerase I. *J. Mol. Biol.* **113**, 237–251.

Rinderknecht, E., O'Conner, B.H. and Rodriguez, H. (1984). Natural Human Interferonγ Complete Amino Acid Sequence and Determination of Sites of Glycosylation. *J. Biol. Chem.* **259**, 6790–6797.

Ringold, G.M., Dieckmann, B., Vannice, J.L., Trahey, M. and McCormick, F. (1984). Inhibition of Protein Synthesis Stimulates the Transcription of Human β–Interferon Genes in Chinese Hamster Ovary Cells. *Proc. natn. Acad. Sci., U.S.A.* **81**, 3964–3968.

Robins, D.M., Paek, I., Seeburg, P.H. and Axel, R. (1982). Regulated Expression of Human Growth Hormone Genes in Mouse Cells. *Cell*, **29**, 623–631.

Rosa, F., Hatat, D., Abadie, A., Wallach, D., Revel, M. and Fellous, M. (1983). Differential Regulation of HLA–DR mRNAs and Cell Surface Antigens by Interferon. *EMBO. J.* **2**, 1585–1589.

Rosen, C.A., Sodroski, J.G. and Haseltine, W.A. (1985). The Location of *Cis*–Acting Regulatory Sequences in the Human T Cell Lymphotropic Virus Type III (HTLV–III/LAV) Long Terminal Repeat. *Cell* **41**, 813–823.

Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R.S., Waldman, A.A. and Pestka, S. (1979). Human Leukocyte Interferon: Production, Purification to Homogeneity and Initial Characterisation. *Proc. natn. Acad. Sci., U.S.A.* **76**, 640–644.

Ryals, J., Dierks, P., Ragg, H. and Weissmann, C. (1985). A 46 Nucleotide Promoter Fragment from an IFNα Gene Renders an Unrelated Promoter Inducible by Virus. *Cell* **41**, 497–507.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA Sequencing with Chain Terminating Inhibitors. *Proc. natn. Acad. Sci., U.S.A.* **74**, 5463–5467.

Sassone–Corsi, P. and Borrelli, E. (1986). Transcriptional Regulation by Trans–acting Factors. *T.I.G.* **2**, 215–219.

Sassone–Corsi, P., Dougherty, J.P., Wasylyk, B. and Chambon, P. (1984). Stimulation of *in Vitro* Transcription from Heterologous Promoters by the Simian Virus 40 Enhancer. *Proc. natn. Acad. Sci., U.S.A.* **81**, 308–312.

Sassone–Corsi, P., Wildeman, A. and Chambon, P. (1985). A Transacting Factor is Responsible for the Simian Virus 40 Enhancer Activity *in Vitro*. *Nature, Lond.* **313**, 458–463.

Saunders, M.E., Gewert, D.R., Tugwell, M.E., McMahon, M. and Williams, B.R.G. (1985). Human 2–5A Synthetase: Characterisation of a Novel cDNA and Corresponding Gene Strucure. *EMBO. J.* **4**, 1761–1768.

Scheidereit, C., Geisse, S., Westphal, H.M. and Beato, M. (1983). The Glucocorticoid

Receptor Binds to Defined Nucleotide Sequences near the Promoter of Mouse Mammary Tumour Virus. *Nature, Lond.* **304**, 749–752.

Scholer, H.R. and Gruss, P. (1984). Specific Interaction between Enhancer–Containing Molecules and Cellular Components. *Cell*, **36**, 403–411.

Sehgal, P.B., Dobberstein, B. and Tamm, I. (1977). Interferon messenger RNA Content of Human Fibroblasts during Induction, Shutoff and Superinduction of Interferon Production. *Proc. natn. Acad. Sci., U.S.A.* **74**, 3409–3413.

Sen, G.C., Taira, H. and Lengyel, P. (1978). Interferon, Double–stranded RNA and Protein Phosphorylation. *J. Biol. Chem.* **253**, 5915–5921.

Sergeant, A., Bohmann, D., Zentgraf, H., Weiher, H. and Keller, W. (1984). A Transcription Enhancer Acts *in Vitro* over Distances of Hundreds of Base–pairs on both Circular and Linear Templates, but not on Chromatin–reconstituted DNA. *J. Mol. Biol.* 180, 577–600.

Serfling, E. Jasin, M. and Schaffner, W. (1985). Enhancers and Eukaryotic Gene Transcription. *TIG*, **1**, 224–230

Shaw, G.D., Boll, W., Taira, H., Mantei, N., Lengyel, P. and Weissmann, C. (1983). Structure and Expression of Cloned Murine IFN–α Genes. *Nucl. Acids Res.* **11**, 555–573.

Shaw, W.V. (1975) Chloramphenicol Acetyltransferase from Chloramphenicol Resistant Bacteria. *Methods Enzymol.* 737– 755.

Shepard, H.M., Leung, D., Stebbing, N. and Goeddel, D.V. (1981). A Single Amino Acid Change in IFN–β_1 Abolishes its Antiviral Activity. *Nature, Lond.* **294**, 563–565.

Shulman, L.M., Kamark, M.E., Slate., Ruddle. F.H., Branca. A.W., Baglioni. C., Maxwell. B.L., Gutterman. J., Anderson. P., Nagler. C., Vilcek. J. (1984). Antibodies to Chromosome 21 Coded Cell Surface Components Block Binding of Human αIFN but not γIFN to Human Cells. *Virology*, 137, 422–427.

Skup, D., Windass, J.D., Sor, F., George, H., Williams, B.R.G., Fukuhara, H., De Maeyer–Guignard, J. and De Maeyer, E. (1982). Molecular Cloning of Partial cDNA Copies of Two Distinct Mouse IFN–β mRNAs. *Nucl. Acids Res.* **10**,

Sloma, A. (1983). European Patent Application. 83102893.1.

Staden, R. (1977). Sequence Data Handling by Computer. *Nucl. Acids Res.* **12**, 4037–4051

Staden, R. (1984). Graphic Methods to Determine the Function of Nucleic Acid Sequences. *Nucl. Acids Res.* **12**, 521–538.

Stadler, J., Larsen, A., Engel, J.D., Dolan, M., Groudine, M. and Weintraub, H. (1980). Tissue–Specific DNA Cleavages in the Globin Chromatin Domain Introduced by DNase I. *Cell* **20**, 451–460.

Stahl, F.W., Crasemann, J.W. and Stahl, M.M. (1975). Rec–mediated Recombinational Hot Spot Activity in Bacteriophage Lambda.III *Chi* Mutations are Site Mutations Stimulating *Rec*–mediated Recombination *J. Mol. Biol.* **94**, 203–212.

Stewart II, W.E. (1979). *The Interferon System.* New York: Springer.

Stewart II, W.E., Sarkar, F.H., Taira, H., Hall, A., Nagata, S. and Weissmann, C. (1986). Comparison of Several Biological and Physicochemical Properties of Human

Leukocyte Interferons Produced by Human Leukocytes and by *E.coli*. *Gene* **11**, 181–186.

Streuli, M., Nagata, S., and Weissmann, C. (1980). At Least Three Human Type α Interferons: Structure of α2. *Science, N.y.* **209**, 1343–1347.

Swimmer, C. and Shenk, T. (1984). A Viable Simian Virus 40 Variant that Carries a Newly Generated Sequence Reiteration in Place of the Normal Duplicated Enhancer Element. *Proc. natn. Acad. Sci., U.S.A.* **81**, 6652–6656.

Taira, H., Broeze, R.J., Jayaram, B.M., Lengyel, P., Hunkapiller, M.W. and Hood, L.E. (1980). Mouse Interferons: Amino Terminal Amino Acid Sequences of Various Species. *Science, N.y.* **207**, 528–529

Tan, Y.H. (1976). Chromosme 21 and the Cell Growth Inhibitory Effect of Human Interferon Preparations. *Nature, Lond.* **260**, 141–143.

Tan, Y.H., Barakat, F., Berthold, W., Smith–Johannsen, H. and Tan, C. (1979). The Isolation and Amino Acid/Sugar Composition of Human Fibroblastoid Interferon. *J. Biol. Chem.* **254**, 8067–8073.

Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Muramatsu, M. and Weissman, C. (1980b). Human Leukocyte and Fibroblast Interferons are Structurally Related. *Nature, Lond.* **285**, 547–549.

Taniguchi, T., Ohno, S., Fujii–Kuriyama, Y. and Muramatsu, M. (1980a). The Nucleotide Sequence of Human Fibroblast Interferon cDNA. *Gene* **10**, 11–15.

Tavernier, J., Derynck, R. and Fiers, W. (1981). Evidence for a Unique Human Fibroblast Interferon (IFN–β_1) Chromosomal Gene, Devoid of Intervening Sequences. *Nucl. Acids Res.* **9**, 461–472.

Taverier, J., Gheysen, D., Duerinck, F., Van der Heyden, J. and Fiers, W. (1983). Deletion Mapping of the Inducible Promoter of Human IFN–β Gene. *Nature, Lond.* **301**, 634–636.

Taya, Y., Devos, R., Tavernier, J., Cheroute, H., Engler, G. and Fiers, W. (1982). Cloning and Structure of the Human Immune Interferon –γ Chromosomal Gene. *EMBO. J.* **1**, 953–958.

Todokoro, K., Kioussis, D. and Weissmann, C. (1984). Two Non–allelic Human Interferon α Genes with Identical Coding regions. *EMBO. J.* **3**, 1809–1812.

Torczynski, R.M., Fuke, M. and Bollon, A.P. (1984). Human Genomic Library Screened with 17–base Oligonucleotide Probes Yields a Novel Interferon Gene. *Proc. natn. Acad. Sci., U.S.A.* **81**, 6451–6455.

Trapman, J. (1979). A Systematic Study of Interferon Production by Mouse L–929 Cells Induced with PolyrI.rC and DEAE–dextran. *FEBBS Lett.* **98**, 107–110.

Trapman, J. (1980). Distinct Antigenic Character of Two Components of Poly(I).Poly(C)–Induced Mouse L Cell Interferon. *FEBBS. Lett.* **109**, 137–139.

Treisman, R. and Maniatis, T. (1985). Simian Virus 40 Enhancer Increases Number of RNA Polymerase II Molecules on Linked DNA. *Nature Lond.* **315**, 72–75.

Trent, J.M., Olson, S. and Lawn, R.M. (1982). Chromosomal Localisation of Human Leukocyte, Fibroblast and Immune IFN Genes by Means of *in Situ* Hybridisation. *Proc. natn. Acad. Sci., U.S.A.* **79**, 7809–7813.

Twigg, A.J. and Sherratt, D. (1980). Trans–complementable Copy–number Mutants of

Plasmid ColE1. *Nature Lond.* 283, 216.

Ullrich, A., Gray, A., Goeddel, D.V. and Dull, T.J. (1982). Nucleotide Sequence of a Portion of Human Chromosome 9 Containing a Leukocyte Interferon Gene Cluster. *J. Mol. Biol.* 156, 467–486.

Van Beveren, C., van Straaten, F., Curran, T., Muller, R. and Verma, I.M. (1983). Analysis of FBJ–MuSV Provirus and c–*fos* (mouse) Gene Reveals That Viral and Cellular *fos* Gene Products Have Different Carboxy Termini. *Cell* 32, 1241–1255.

Van Damme, J., Opdenakker, G., Billiau, A., De Somer, P., De Wit, L., Poupart, P. and Content , J. (1985). Stimulation of Fibroblast Interferon Production by a 22K Protein from Human Leukocytes. *J. Gen. Virol.* 66, 693–700.

Vilcek, J. and Havell, E.A. (1973). Stabilisation of Messenger RNA Activity by Treatment with Metabolic Inhibitors and Lowering of the Incubation Temperature. *Proc. natn. Acad. Sci., U.S.A.* 70, 3909–3913.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T. and Itakura, K. (1979). Hybridisation of Synthetic Oligodeoxyribonucleotides to $\Phi\chi174$ DNA: the Effect of Single Base Pair Mismatch. *Nucl. Acids Res.* 6, 3543–3557.

Walker, M.D., Edlund, T., Boulet,. A.M. and Rutter, W.J. (1983). Cell–specific Expression Controlled by the 5'–flanking Region of Insulin and Chymotrypsin Genes. *Nature Lond.* 306, 555–561.

Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roger, A., Teoule, R., and Chambon, P. (1980). Specific *in Vitro* Transcription of Conalbumin Gene is Drastically Decreased by Single–point Mutations in T–A–T–A Box Homology Sequence. *Proc. natn. Acad. Sci., U.S.A.* 77, 7024–7028.

Weaver, R.R. and Weissmann, C. (1979). Mapping of the Berk–Sharp Procedure: The 5' Termini of 15S β–globin mRNA precursors and Mature 10S β–globin mRNA have Identical Map Coordinates. *Nucl. Acids Res.* 7, 1175–1193.

Weber, F., de Villiers, J. and Schaffner, W. (1984). An SV40 "Enhancer Trap" Incorporates Exogenous Enhancers or Generates Enhancers from Its Own Sequences. *Cell* 36, 983–992.

Weidle, U. and Weissmann, C. (1983). The 5' Flanking Region of a Human IFN–α Gene Mediates Viral Induction of transcription. *Nature, Lond.* 303, 442–446.

Weiher, H., Konig, M. and Gruss, P. (1983). Multiple Point Mutations Affecting the Simian Virus 40 Enhancer. *Science, N.y.* 219, 626–631.

Weil, P.A., Luse, D.S., Segall, J. and Roeder, R.G. (1979). selective and Accurate Initiation of Transcription at the Ad2 Major Late Promotor in a Soluble System Dependent on Purified RNA Polymerase II and DNA. *Cell* 18, 469–484.

Weil, J., Epstein, C.J., Epstein, L.B., Sedmak, J.J., Sabran, J.L. and Grossberg, S.E. (1983). A Unique set of Polypeptides is Induced by γ Interferon in addition to those Induced in Common with α and β Interferons. *Nature, Lond.* 301, 437–439.

Weintraub, H. and Groudine, M. (1976). Chromosomal subunits in Active Genes have an Altered Conformation. *Science, N.y.* 193, 848–856.

Weissenbach, J., Cherajvsky, Y., Zeevi, M., Schulman, L., Soreq, H., Nir, U., Wallach, D., Perricaudet, M., Tiollais, P. and Revel, M. (1980). Two Interferon mRNAs in Human Fibroblasts: *In Vitro* Translation and *E.coli*. Cloning Studies. *Proc. natn. Acad. Sci., U.S.A.* ?, 7152–7156.

Wetzel, R. (1981). Assignment of the Disulphide Bonds of Leukocyte Interferon. *Nature Lond.* 289, 606–607.

Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G. and Chasin, L., (1979). DNA–mediated Transfer of the Adenine Phosphoribosyl Transferase Locus into Mammalian Cells. *Proc. natn. Acad. Sci., U.S.A.* 76, 1373–1376.

Wildeman, A.G., Sassone–Corsi, P., Grundstrom, T., Zenke, M. and Chambon, P. (1984). Stimulation of *in Vitro* Transcription from the SV40 Early Promoter by the Enhancer Involves a Specific *Trans*–acting Factor. *EMBO. J.* 3, 3129–3133.

Wilson, V., Jeffreys, A.J., Barrie, P.A., Boseley, P.G., Slocombe, P.M, Easton, A. and Burke, D.C. (1983). A Comparison of Vertebrate Interferon Gene Families Detected by Hybridisation with Human DNA. *J. Mol. Biol.* 166, 457–475.

Wright, S., Rosenthal, A., Flavell, R. and Grosveld, F. (1984). DNA Sequences Required for Regulated Expression of β–globin Genes in Murine Erythroleukaemia Cells. *Cell* 38, 265–273.

Wright, S., deBoer, E., Grosveld, F. and Flavell, R. (1983). Regulated Expression of the Human β–globin Gene Family in Murine Erythroleukaemia Cells. *Nature Lond.* 305, 333–336.

Wu, C. (1980). The 5' Ends of *Drosophila* Heat Shock Genes in Chromatin are Hypersensitive to DNase I. *Nature, Lond.* 286, 854–860.

Wu, C. (1984b). Two Protein–binding sites in Chromatin Implicated in the Activation of Heat–shock Genes. *Nature, Lond.* 309, 229–234.

Wu, C. (1984a). Activating Protein Factor Binds *in Vitro* to Upstream Control Sequences in Heat Shock Gene Chromatin. *Nature, Lond.* 311, 81–84.

Yip, Y.K., Barrowclough, B.S., Urban, C. and Vilcek, J. (1982a). Purification of Two Subspecies of Human γ(immune) Interferon. *Proc. natn. Acad. Sci., U.S.A.* 79, 1820–1824.

Yip, Y.K., Barrowclough, B.S., Urban, C. and Vilcek, J. (1982b). Molecular Weight of Human Gamma Interferons is Similar to that of Other Human Interferons. *Science, N.y.* 215, 411–412.

Yoshie, O., Schmidt, H., Reddy, E.S.P., Weissman, S. and Lengyel, P. (1982). Mouse Interferons Enhance the Accumulation of a Human HLA RNA and Protein in Transfected Mouse and Hamster Cells. *J. Biol. Chem.* 257, 13169–13172.

Yoshie, O., Schmidt, H., Lengyel, P., Reddy, E.S.P., Morgan, W.R. and Weissman, S.M. (1984). Transcripts of Human HLA Gene Fragments Lacking the 5' Terminal Region in Transfected Mouse Cells. *Proc. natn. Acad. Sci., U.S.A.* 81, 649–653.

Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A. and Chambon, P. (1986). Multiple Sequence Motifs are Involved in SV40 Enhancer Function. *EMBO. J.* 5, 387–397.

Zinn, K., DiMaio, D. and Maniatis, T. (1983). Identification of Two Distinct Regulatory Regions Adjacent to the Human β–Interferon Gene. *Cell*, 34, 866–879.

Zinn, K. and Maniatis, T. (1986). Detection of Factors That Interact with the Human β–Interferon Regulatory Region *in Vivo* by DNase I Footprinting. *Cell* 45, 611–618.

Zinn, K., Mellon, P., Ptashne, M. and Maniatis, T. (1982). Regulated Expression of an

Extrachromosomal Human β−Interferon Gene in Mouse Cells. *Proc. natn. Acad. Sci., U.S.A.* 79, 4897−4901.

Zoller, M.J. and Smith, M. (1983). Oligonucleotide−directed Mutagenesis of DNA Fragments Cloned into M13 Vectors. *Methods Enzymol.* **100**, 468−500.

Zoon, K.C., Smith, M.E., Bridgen, P.J., Anfinsen, C.B., Hunkapíller, M.W. and Hood, L.E. (1980). Amino Terminal Sequence of the Major Component of Human Lymphoblastoid Interferon. *Science, N.y.* **207**, 527−528.

Zoon, K., Smith, M.E., Bridgen, P.J., zur Nedden, D. and Anfinsen, C.B. (1979). Purification and Partial Characterisation of Human Lymphoblastoid Interferon. *Proc. natn. Acad. Sci., U.S.A.* **76**, 5601−5605.

Zullo, J.N., Cochran, B.H., Huang, A.S. and Stiles, C.D. (1985). Platelet Derived Growth Factor and Double Stranded Ribonucleic Acids Stimulate Expression of the Same Genes in 3T3 Cells. *Cell*, **43**, 793−800.

Zwarthoff, E.C., Mooren, A.T.A. and Trapman, J. (1985). Organisation, Structure and Expression of Murine Interferon Alpha Genes. *Nucl. Acids Res.* **13**, 791−804.

MOLECULAR ANALYSIS OF HUMAN AND MOUSE INTERFERON α GENE STRUCTURE AND FUNCTION

C. Bartholomew

SUMMARY:-

Four human IFNα chromosomal genes have been isolated from a newly constructed placental DNA library in λL47. Restriction and sequencing analysis revealed that each gene had been described previously. However, one gene, $SMT111.1_A$, which encodes a full length IFN, is an allelic varient of a previously characterised pseudogene, thus indicating some degeneracy of the IFNα gene family.

A chimaeric gene comprising the $MuIFN\alpha_1$ promoter (-188 to +52) and *cat* gene coding sequences has been constructed *in vitro*, enabling promoter function to be examined in mouse cells. Reproducible polyrI.rC mediated induction of CAT expression from the $MuIFN\alpha_1$ promoter has been demonstrated in pools of stably transfected, but not transiently transfected, L929 cells. Monitoring mRNA production revealed the transient accumulation of correctly initiated hybrid gene transcripts which precede optimum CAT production.

Aspects of the structure/function relationship of the $MuIFN\alpha_1$ promoter have been investigated by oligonucleotide site directed mutagenesis. Comparative studies of IFNα promoter sequences identified prospective regulatory regions for mutagenesis. Quantitative CAT assays have been employed for promoter assessment. Additionally, the construction of a pseudogene comprising the wildtype $MuIFN\alpha_1$ promoter linked to an internally deleted *cat* gene has enabled both mutant and wildtype promoters functioning simultaneously in the same cell population to be assessed by S-1 nuclease protection studies using a common probe.

Such studies have revealed three distinct *cis*-acting regions implicated in $MuIFN\alpha_1$ promoter function. Two are located upstream of the TATA box, defined by mutations at -87 and between -66 and -33 respectively. These reduce promoter activity 2 to 5 fold. The third, is defined by a mutation at +14, within the untranslated leader sequence. This enhances activity 2 to 3 fold.

A deletion derivative of the $MuIFN\alpha_1$ promoter containing only 94bp of upstream sequence is inactive. Cis-activation by the Mo-MuSV enhancer restores inducibility to this promoter whereas the intact $MuIFN\alpha_1$ promoter is refractory to this element. This suggests that distinct *cis*-acting sequences dictate the efficiency and regulation of $MuIFN\alpha_1$ gene transcription.