ACTIVATION OF CELLULAR GENES FOLLOWING TRANSFORMATION BY SIMIAN VIRUS 40

Michael R.D. Scott,

Department of Biochemistry,

Imperial College of Science and Technology,

Submitted for the degree of Ph.D., 1982.

ACKNOWLEDGEMENTS

I would like to thank Peter Rigby for his support, advice and encouragement during the course of this project.

I would also like to thank my collaborators Stephanie Campbell-Smith and David Murphy, together with Guil Winchester and David Lane, all of whom allowed me access to their unpublished data.

I would also like to thank my colleagues in the C.R.C. group for advice and gifts of materials used in this project, Doug Hanahan for help with plasmid transfection, Glyn Millhouse for help with photography, Angela Barnwell, Sue Hayman and Cathy Headhouse-Benson for typing and the entire C.R.C. group for many happy memories.

ABSTRACT

The expression of cellular genes in mouse cells transformed by Simian Virus 40 has been investigated. Methods were developed which allowed the cloning of expressed mRNA sequences at high efficiency. When such cloned sequences were studied by conventional differential colony hybridization techniques, two conclusions were reached. Firstly, the technique was found to be extremely insensitive, being incapable of monitoring the expression of the majority of mRNAs which are of low abundance in cells. Secondly, no changes in the expression of abundant mRNAs could be detected using this approach within the system under study.

The detection threshold of differential colony hybridization was found to be limited by signal-to-noise ratio in such experiments. To improve the signal-to-noise ratio and hence increase the sensitivity of colony hybridization, a novel approach was adopted. A cDNA library derived from normal Balb/c 3T3 cell mRNA was amplified and the extracted DNA was attached to cellulose. The immobilized DNA was then used to remove homologous sequences from a radioactive cDNA probe derived from SV40-transformed cellular mRNA. This allowed the identification and purification of a probe which was specific for transformed cellular mRNA sequences which were represented at much lower levels, if at all,in normal cells.

When this purified probe was used to screen a library of cDNA clones derived from the mRNA of the transformed cell line, a number of positive clones were identified. Most of these were found to belong to one of four "Sets" when categorized on the basis of cross-hybridization. Two of these Sets were found to be related to a clone with an unusually large insert and further experiments established that all of these clones were of mitochondrial origin

and that the two clone Sets were related to the genes that encode two subunits of cytochrome oxidase.

The remaining two clone Sets were found to be of cellular origin, and members of each Set hybridized to several mRNAs in transformed cells which were either absent from or present at low levels in normal cells. The expression of mRNAs related to one of these Sets was found to show evidence of a connection with the expression of the viral transforming protein and the expression of mRNAs related to both Sets was found to correlate with the oncogenic potential of mouse cell lines transformed by SV40.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

bp	one nucleotide of RNA or DNA
BSA	bovine serum albumin
cDNA	complementary DNA
cfu	colony forming unit
cpm	counts per minute
ddCTP	dideoxycytidine triphosphate
ddTTP	dideoxythymidine triphosphate
ddH_2O	double distilled water
DNase	deoxyribonuclease
DMSO	dimethyl sulphoxide
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetic acid
h	hour
Kb	one thousand nucleotides of DNA or RNA
ND	experiment not performed
NP40	nonidet-P40
OD	optical density
oligo dT	oligomeric dTMP
PIPES	piperazine-NN'-bis-2-ethane-sulphonic acid
poly A or poly rA	polymeric AMP
poly D or poly rU	polymeric UMP
PM	preincubation medium
min	minute(s)
SDS	sodium dodecyl sulphate
sec	second(s)
sp. act.	specific radioactivity

TEMED	NNN'N'-tetramethylethylenediamine
Tris	tris(hydroxymethyl)methylamine
tRNA	transfer ribonucleic acid

Symbols-

<	less than
>	greater than
~	approximately

CHAPTER I

GENERAL INTRODUCTION

1.1 Transformation

Transformation has been defined by Martin (1981) as "the acquisition of growth characteristics not exhibited by the parental cells". This definition will be adhered to throughout this thesis. In this work, however, attention will be focussed upon one particular subset of transformation events, the aquisition of novel growth properties characteristic of cultured tumour cells, and hence relevant to the phenomenon of oncogenesis.

When the growth behaviours of normal primary cells and of cultured cells from malignant tumours are compared, many differences may be observed, including growth following serum depletion, capability for anchorage-independent growth, and ability to colonise and form tumours following injection into syngeneic or immunosuppressed animals. For an exhaustive list, the reader is referred to Tooze (1980) and Martin (1981). Many of the changes in growth behaviour may also be observed following treatment of cells in tissue culture with SV40 or other oncogenic agents. For this reason it may be surmised that these *in vitro* systems largely mimic the *in vivo* situation and that an understanding of the transformation of cells in culture will relate to the phenomenon of spontaneous tumorigenesis as it occurs *in vivo*.

The definition of transformation as quoted implies that it should be possible to transform a cell line already adjudged to be transformed by other criteria. This observation stresses the importance of two factors worthy of consideration;

(a) The growth parameters of the parental or "normal" cells.

(b) The criteria used to define the transformed state.

Thus many "normal" cells may be described as already transformed for certain growth characteristics. For example, the Balb/c 3T3

cells used in this study are aneuploid and immortal in tissue culture. Both of these properties are uncharacteristic of normal tissue cells or primary cells in culture but are common in tumour cells. Nevertheless, Balb/c 3T3 cells will not grow following serum deprivation, exhibit anchorage-independent growth or form tumours in syngeneic or immunosupressed animals. These properties are frequently displayed by tumour cells and very often following exposure of Balb/c 3T3 cells to SV40 virus or other oncogenic agents.

There sometimes appears to be a hierarchy of transformation parameters amongst cells selected following infection with SV40. Thus cells which have gained focus-forming activity do not necessarily exhibit anchorage-independent growth. In contrast, however, transformants distinguished by their ability to grow in agar invariably form dense foci on plastic (Risser and Pollack 1974). Transformation parameters may therefore be ordered to a certain extent on the basis of "stringency", with the least stringent characteristics being focus-formation in low serum, extending upward ultimately to tumorigenicity (Risser and Pollack 1974; Shin et al., 1975). This heirarchy may be further extended in both directions; hence immortality of cells in culture may be regarded as a still less stringent criterion than focus-formation, and the ability to metastasize would be a more stringent criterion than tumour formation.

Table 1.1 shows the results of an investigation of the biological properties and growth behaviour of six independently isolated lines of SV40-transformed Balb/c 3T3 mouse cells (Rigby et al.,1980; P.W.J. Rigby, M.Lovett, P.H.Gallimore and C.N. Cole, manuscript in preparation). These lines were made simultaneously using the same cells and the same preparation of SV40 virus, and were all selected by the low stringency criterion of focus formation in low concentrations

TABLE 1.1

Table 1.1 Biological Properties of a Number of SV40-transformed Mouse Cell Lines Contrasted with those of Normal Balb/c 3T3 Cells

The data for this table are taken from Rigby et al. (1980).

Cell Line	Moi. of Transformation	% Efficiency of Plating on 3T3 Monolayers	% Efficiency of Plating in Methocell	Fibronectin % of 3T3	Tumorigenicity in Nude Mice Dose: 2×10^6 Cells Tumor positive / total	Latent Period
Balb/c 3T3	-	0.1	<0.0005	100	$^0/_4$	>150
SV3T3 C120	100	80	43	8.5	$^1/_6$	63
SV3T3 C126	10	6	11	15	$^3/_3$	20
SV3T3 C138	1	2	4.35	35	$^3/_3$	19
SV3T3 C149	1	8	7.2	18	$^0/_3$	>150
SV3T3 C1H	0.5	41	8.3	7	$^3/_3$	18
SV3T3 C1M	0.5	7	30	10	$^2/_3$	14

of serum. The data of Table 1.1 illustrate that the biological properties of these cell lines are all quite different and that they do not correlate with the one intentional variable in the experiment, the multiplicity of infection during transformation. Although it is clear that all cell lines able to form tumours when injected into nude mice also show a significantly higher efficiency of plating in methocell and on monolayers of normal cells (two "stringent" criteria of transformation) than do untransformed Balb/c 3T3 cells, there is no correlation between the relative degres of expression of any two parameters of transformation. Thus lines which plate most efficiently in methocell (C120 and C1M) are not correspondingly the most tumorigenic in animals.

This treatise will concentrate upon the molecular changes that underlie transformation and tumorigenesis and the experimental system under scrutiny, transformation of Balb/c 3T3 cells with SV40 virus and selection for focus-formation in low serum (arguably) represents one of the simplest available for examining the effect of a transforming agent upon normal growth control. Consequently, the molecular changes that presumably relate to the transition from primary cell to secondary cell in culture (Todaro and Green, 1963) will not be investigated, although this in no way denies that this phenomenon is equally relevant to the oncogenic process.

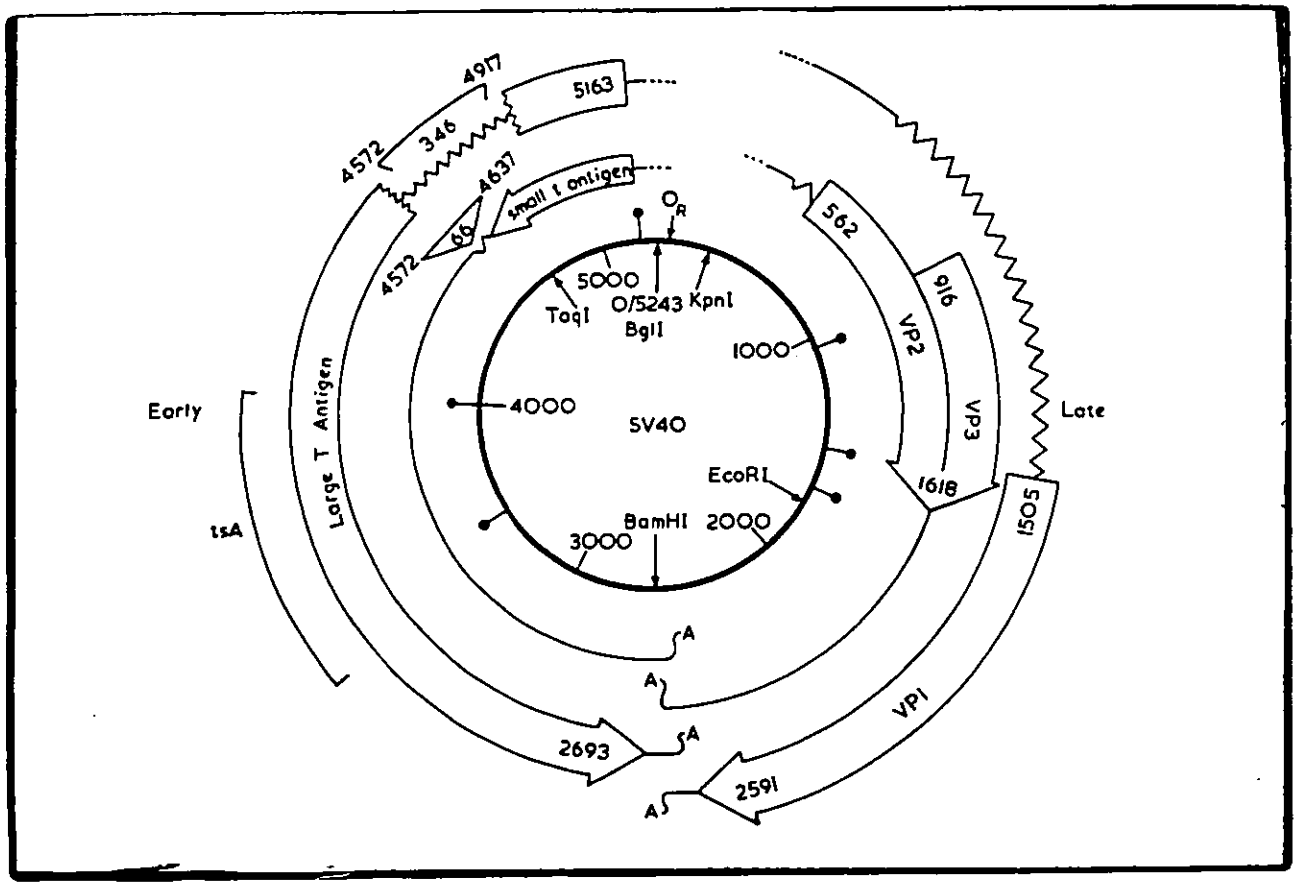
1.2 SV40 functions involved in transformation.

The life cycle and molecular biology of SV40 have been recently reviewed by Tooze (1980) and a detailed discussion of the data is beyond the scope of this study. The interrelationships of the SV40 mRNAs, the proteins they encode and the structure of the viral genome are summarised in Fig.1.1. It has been shown that only part of the SV40 genome is expressed early following infection of cells and that

FIGURE 1.1

Figure 1.1 A Schematic Representation of the SV40 Genome Indicating the Locations of some Principal Features

The SV40 genome is represented as a circle with the origin of DNA replication at the top. Numbers within the circle refer to the SV40 DNA sequence (Fiers et al., 1978). The symbol (↑) denotes Hind III cleavage sites. Arcs with arrowheads denote coding positions for mRNAs, with arrowheads pointing in the 5' to 3' direction. Boxed areas represent coding segments within the viral mRNA. Numbers within these boxed areas denote translational initiation and termination sites. Unboxed numbers denote the end points and extent of the early splices.


4917
5163
4572
346
4637
small t antigen
66
4572
O_R
5000
TaqI
O/5243
KpnI
BglI
1000
4000
SV40
EcoRI
BamHI
2000
3000
562
VP2
916
VP3
Late
1618
1505
Early
tsA
Large T Antigen
A
A
A
2693
VP1
2591

this early region is sufficient to induce transformation (Tooze, 1980). Two early viral gene products have been implicated in transformation by SV40 (see Fig.1.1). The first, large T-antigen, is a protein of molecular weight 94,000, originally identified by immunoprecipitation with sera taken from animals bearing virus-induced tumours. Subsequently it was shown (Prives et al.,1977) that viral RNA from infected cells can be translated into another protein of molecular weight 17,000, now known as small t-antigen. The two proteins share the same 82 residues at their amino termini (Paucha et al., 1978a). Subsequently, Berk and Sharp (1978) showed that there exist two distinct early SV40 mRNAs, generated by differential splicing of the same precursor; these encode the large T and small t proteins (Paucha et al.,1978b).

Large T-antigen.

That large T-antigen is required both for the establishment and maintenance of transformation seems beyond doubt. The evidence has been extensively reviewed by Tooze (1980) and Martin (1981) and is briefly summarized as follows:

(a) Temperature-sensitive point mutations (tsA) affect the synthesis, stability and function of large-T antigen (Tegtmeyer, 1972; Tegtmeyer et al., 1977; Tjian, 1978). Infection of cells at the non-permissive temperature not only prevents viral DNA replication in permissive monkey cells but also fails to induce transformation of non-permissive rodent cells (Tegtmeyer, 1972; Chou et al., 1974; Brugge and Butel, 1975; Kimura and Itagaki, 1975; Fluck and Benjamin, 1979).

(b) Transfection experiments using purified SV40 DNA fragments have shown that the early region of the genome (see Fig. 1.1) is both

necessary and sufficient for transformation (Graham et al., 1974; Abrahams et al., 1975; Martin, 1981).

(c) All of the SV40-transformed cell lines examined to date express the entire large T-antigen coding region (Martin, 1981).

(d) Under certain conditions (see below), deletion mutants unable to make small t-antigen are capable both of transformation in vitro and of tumour induction in vivo (Lewis and Martin, 1979).

Host DNA synthesis is stimulated following infection of both monkey and rodent cells with SV40 (Henry et al., 1966; Kit et al., 1967; Ritzi and Levine, 1969; Seif and Martin 1979b; Setlow et al., 1980). Moreover, infection of rodent cells causes transient cell division resulting in the generation, at high frequency, of abortive transformants, indicated by the formation of microcolonies in agar or microfoci on plastic (Sauer and Defendi, 1966; Scher, 1971; Smith et al., 1971). Only a small proportion of these events leads to stable transformation.

SV40(tsA) mutants at 40°C (the non-permissive temperature for viral DNA replication and establishment of transformation) will induce abortive transformation and host DNA synthesis (Chou and Martin, 1975; Fluck and Benjamin, 1979; Hiscott and Defendi, 1979). T-antigen does however appear to be involved in, at least, the induction of host DNA synthesis since (tsA) mutants at the more stringent temperature of 41.5°C failed to induce host replication although they succeeded at 40°C (Chou and Martin, 1975). Thus the tsA mutations have allowed dissection of the DNA synthesis stimulating activity of large T-antigen from the establishment of transformation function.

Small t-antigen.

The $\Delta_{.54-.59}$ mutations of SV40 are defective for small t-antigen production although synthesis of large T-antigen is unaffected. These mutants do not induce abortive transformation (Fluck and Benjamin, 1979) and several lines of evidence argue that small t-antigen is not required for, at the very least, the first round of host DNA synthesis (Martin, 1981). Furthermore, small t-deletion mutants are unable to effect transformation of resting cells, but they efficiently transform growing cells (Bouck et al., 1978; Martin et al., 1979b, c). Small t $\Delta_{.54-.59}$ mutants fail to transform cells in medium supplemented with human plasma but do so if the growth factor PDGF is added (Martin et al., 1979a). Moreover, it now seems that transformants generated by small t-deletion mutants fail to exhibit efficient anchorage-independent growth when grown in certain serum conditions. Supplementing the medium with other sera may compensate the defect (Steinberg and Pollack, 1979). These data, taken together, suggest that small t-antigen may possess a growth-factor like activity, thus serving to enhance transformation perhaps in a manner akin to the tumour promoters described in Section 1.5.

1.3 Interaction of SV40 with its host

Proviral integration sites

All lines stably transformed by SV40 contain viral DNA covalently linked to cellular DNA. Numerous experiments have demonstrated that different SV40-transformed lines contain different amounts of viral DNA inserted at a variety of sites in the cellular DNA and that the position of the junction is not specific (Sambrook et al., 1968; Gelb et al., 1971; Botchan et al., 1976; Ketner and Kelly,1976; Rigby et al., 1980; Clayton and Rigby, 1981).

Whereas in general a major conclusion of these studies is that integration of viral DNA is not specific with respect to either the viral or cellular DNA, there has been a suggestion of a preferential chromosomal integration site. Mougneau et al., (1980) reported the presence in three SV40-transformed rat cell lines of integrated structures indistinguishable by DNA-transfer hybridization experiments. In light of recent observations (Hayward et al., 1981; Neel et al., 1981; Noori-Daloii et al., 1981; Payne et al., 1981, 1982) on the mechanism of oncogenesis by certain retroviruses (see Section 1.4) the possibility that SV40 proviruses affect expression of adjacent cellular genes may be entertained. For example, gene activation could be accomplished by fusion to the SV40 early promoter. In addition, a 72 base-pair repeat sequence near the SV40 origin of DNA replication appears to possess a cis-acting activity which enhances the expression of adjacent genes (Banerji et al., 1981). Although most available evidence argues that specific integration is not mandatory for transformation by SV40, the apparent conflict between the results of Mougneau et al. (1980) and those of other investigators could be a consequence of differences in growth conditions or selection criteria used; hence the possibility that the provirus could exert a selectable effect upon cell growth behaviour by integrating next to and influencing the expression of particular cellular genes can not yet be dismissed.

Interaction of viral and cellular gene products

Cells transformed or infected by SV40 contain a protein of molecular wieght 53,000 (p53) which exists as a complex with large T-antigen (Lane and Crawford, 1979; McCormick and Harlow, 1980; Fanning et al., 1981). This protein is found at greatly reduced levels in normal 3T3 cells and a protein of this size has been shown

to exist at high levels in many cells transformed by other agents (Jay et al., 1979). The protein has also been observed at high levels in normal thymus cells, indicating that its presence may be related to rapid cell proliferation.

Oren et al. (1981) have demonstrated that poly A^+ RNA taken from normal cells can be translated into p53 as efficiently as RNA from an SV40-transformed derivative. In addition, the protein appears to be synthesized in normal cells but is rapidly degraded. Thus, the apparently elevated levels of p53 in transformed cells may be a consequence of stabilization of the protein by virtue of its association with T-antigen. It is perhaps germane to point out that the instability of the protein in normal cells demands caution in interpreting much of the data on the relative levels of the protein in cells since all these studies were performed by in vivo labelling and the conditions of labelling used might distort absolute measurements of protein levels. Nevertheless it is tempting to speculate that synthesis and degradation of p53 may be normally involved in the control of cell division and that the observed association with T-antigen may pertain to the mechanism by which SV40 over-rides these normal controls.

It is of interest to note that when several SV40-transformed cell lines were ordered with respect to the total amount of T-antigen per cell, the order obtained corresponded closely with the relative growth rates, in subconfluent culture, of the same cell lines (G. Winchester and D.P. Lane, personal communication). Thus, cells expressing large amounts of T-antigen were found to divide more rapidly than those containing low levels of the protein. It may be speculated that the rate of cell division is in reality proportional to the amount of p53 per cell. Complexing with, and consequently stabilizing p53,

T-antigen might effect rapid cell proliferation by increasing the net amount of p53 available. The observation that the overall amounts of large T-antigen and p53 in transformed cells appear to vary in parallel (Lane and Crawford, 1979; Linzer and Levine, 1979; Crawford et al., 1980) lends credence to this hypothesis. Since T-antigen is known to possess a DNA binding activity and clearly plays a central role in viral DNA replication and SV40-mediated induction of cellular DNA synthesis (this is discussed later in this Section and in Section 1.2), p53 could be a specificity factor enabling T-antigen to recognize cellular replication sites.

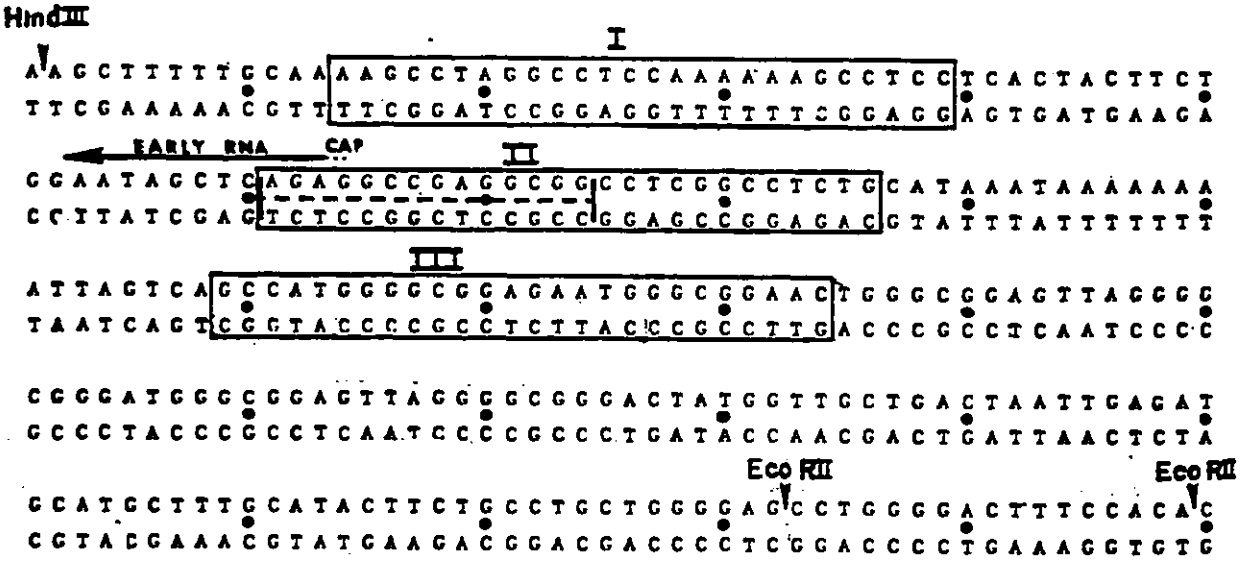
Interaction between T-antigen and DNA

There is also evidence for direct interaction of large T-antigen with DNA. Nuclease protection experiments have demonstrated that the large T-related D2-T protein encoded by the adenovirus 2-SV40 hybrid virus $Ad2^{+}D2$ (Hassel et al., 1978), which contains all of the large T sequences except for the region shared with small t-antigen, specifically binds to the SV40 replication origin. Both D2-T (Myers et al., 1981a) and large T polypeptide (Shalloway et al., 1980; Myers et al., 1981b; Thummel et al., 1981) bind tightly in a sequential and cooperative manner to three tandem recognition sites, illustrated in Fig.1.2. This situation is strikingly reminiscent of the interaction between the bacteriophage λ repressor protein and its operator (Maniatis and Ptashne, 1973; Meyer et al., 1975) and raises the question whether direct interaction between large-T antigen and SV40 DNA is involved in the regulation of viral transcription. Using an in vitro transcription assy, Rio et al. (1980) have shown that D2-T will selectively inhibit early SV40 transcription without affecting transcription of either SV40 or adenovirus 2 late regions. Purified large T-polypeptide exhibits a similar activity. In addition, Conrad

FIGURE 1.2

Figure 1.2 The Binding Sites on SV40 DNA for T-antigen

This figure was adapted from Myers et al., (1981b). The DNA sequence in the region of the origin of replication is shown. The three T-antigen binding sites are enclosed by boxes. Dots denote 10 bp intervals. The top sequence line is the early strand sequence and the 14 bp region of homology with the human "Alu" family of sequences is designated by a dotted line.


HindIII
I
EARLY RNA
CAP
II
III
Eco RII
Eco RII

and Botchan (1981) have isolated human DNA segments which have homology to the SV40 origin, bind T-antigen and which can enhance the expression of adjacent genes in a manner analogous to the SV40 72 base -pair repeats (Banerji et al., 1981). Sequences showing homology with the SV40 replication origin have also been cloned from the genomic DNA of uninfected monkey cells (McCutchan and Singer, 1981), and a major class of repeated DNA sequences (the "Alu" family), present in human DNA, contains a 14 nucleotide homology to part of the SV40 origin of replication (Jelinek et al., 1980; see also Fig.1.2). This 14 base-pair sequence represents part of one of the three putative binding sites for T-antigen (Fig.1.2) and is located within a region that is essential for origin function (Gutai and Nathans, 1978; Subramanian and Shenk, 1978). Thus several lines of evidence imply that direct interaction between large T-antigen and SV40 DNA may be involved both in viral replication and in autoregulation of viral transcription. Furthermore, there exist potential cellular targets for the protein suggesting that large T polypeptide could exert an effect upon the expression and/or replication of the host genome, correlating with the fact that T-antigen is an absolute requirement for the induction of cellular DNA synthesis and the establishment and maintenance of transformation by SV40.

1.4 Activation of cellular genes in transformation.

Viral oncogenes

Most retroviruses, including the acute leukemia and sarcoma viruses, transform cells by virtue of "onc" genes carried on the viral genome which confer the ability to transform cells in vitro and induce tumours in animals (Frankel and Fischinger, 1976; Stehelin et al., 1976; Rousell et al., 1979).

In all cases so far tested, a cellular counterpart of the transforming region ("proto-oncogene") has been identified, and it has been demonstrated that normal uninfected cells contain, at much lower levels, a protein either identical to or closely related to the putative proto-oncogene product (Opperman et al., 1979; Goff et al., 1980; Hayman, 1981). It seems likely that transformation by retroviruses is accomplished by overexpression of cellular gene products involved normally in the control of cell growth (Opperman et al., 1979; Roussel et al., 1979). In support of this theory, Oskarsson et al. (1980) have shown, by DNA transfection of molecularly cloned cellular sequences related to the Moloney murine sarcoma virus transforming region, that the normal homologue can indeed induce transformation when expression is forced by fusion to a viral promoter.

However, not all tumorigenic retroviruses contain transforming genes. Until recently the mechanism of transformation of one group, the lymphoid leukosis viruses (LLV), remained a mystery. The LLV are replication competent viruses, unlike the majority of acute transforming and sarcoma viruses (Graf and Beug, 1978; Hayman, 1981), and are further distinguished by the relatively long latency periods, usually 4-12 months, required before tumours are detected in infected animals (Graf and Beug, 1978). Acute viruses induce neoplasms in 2-4 weeks.

Neel et al. (1981) and Payne et al. (1981) examined the structure and expression of the integrated proviral genome in several LLV-induced tumours. Most tumours examined, both primary and metastatic, appeared clonal but, strikingly, tumours from different animals were found to have integrated viral sequences at the same sites. Furthermore, hybrid poly A^+ RNAs comprising covalently linked viral and cellular sequences were observed, suggesting that

expression of the same cellular gene(s) flanking the viral integration site might be accomplished from a promoter located within the proviral DNA. In light of the fact that many of the chromosomal sites seemed similar, this suggests that the LLV may transform by insertion beside a cellular gene which when expressed at elevated levels from a viral promoter causes the switch to the transformed state. Support for this view comes from the observation (Hayward et al., 1981; Neel et al., 1981; Payne et al., 1981, 1982) that many of the tumours studied showed elevated levels of RNA homologous to the transforming region of the acute leukemia virus MC29, and that the proviral insertion was indeed adjacent to the MC29 proto-oncogene. Most of the data were consistent with the "promoter insertion" model outlined above, although, surprisingly, Payne et al. (1982) reported that in some cases the viral sequences were located downstream of the MC29 specific region. This might imply that integration of the viral genome can activate expression of sequences on either side of the insert, without directly promoting transcription, perhaps in a manner analogous to that of the SV40 72 base-pair repeat described by Banerji et al. (1981) and discussed in Section 1.3.

Cellular transforming genes.

The advent of gene transfer techniques (Wigler et al., 1978) has allowed direct analysis of the transforming potential of normal cellular genes when taken out of their habitual context (Shih et al., 1979; Cooper et al., 1980; Krontiris and Cooper, 1981; Shih et al., 1981). Genomic DNA from chemically-transformed mouse cells (Shih et al., 1979; Cooper et al., 1980) and certain human carcinomas (Krontiris and Cooper, 1981; Shih et al., 1981) have been shown to induce transformation of NIH3T3 fibroblasts where DNA from

untransformed cells did not. Subsequently Murray et al. (1981) showed that when DNAs from human colon and bladder carcinoma cell lines and one promyelocytic leukemia line were used, the recipient mouse cells acquired a large number of human "Alu" type dispersed repetitive sequences. When cell DNA from primary foci was used for a subsequent cycle of transfection, secondary foci were obtained which now contained very little human DNA, as evidenced by the presence of Alu sequences. DNAs from several different primary foci were used and when a number of secondaries were examined, all were found to share a subset of human sequences in addition to a number of uncommon fragments. Murray et al. (1981) postulated that this indicated that all the transformants obtained had acquired the same human DNA sequences and that the ubiquitous fragments observed on their blots were those responsible for the transforming potential of the human tumour DNA. Furthermore, comparison of the DNA from secondary transformants derived from transfection with tumour DNA of different origins indicated that a different transforming sequence has been inherited in each case.

One question of fundamental importance regarding all these transfection analyses has yet to be answered. That is whether the donor sequences which induce transformation to focus formation and colony formation in soft agar are truly representative of the primary lesion responsible for the original transformation. In light of evidence to be discussed in the following section, it might be argued that activation of cellular genes may be necessary for the progression of a "primary" transformed cell to a more malignant state. Within this context, an activated gene which served merely to enhance transformation at some stage in the evolution of a tumour might induce transformation in a heterologous system, perhaps even where the

original lesion could not. Further to this point, it should be noted that whereas Hayward et al. (1981) and Payne et al.,(1982) presented evidence that transformation by lymphoid leukosis viruses involves proviral insertion close to the putative transforming gene, Cooper and Neiman (1980, 1981) demonstrated that transformed NIH3T3 clones obtained following transfection with DNA derived from LLV-induced bursal lymphomas did not contain MC29 virus-related sequences. Perhaps the "transforming gene" observed by Cooper and Neiman (1980, 1981) was in reality generated as a side effect of tumour development, being involved more in the progression towards malignancy than in the primary neoplastic event.

Evidence of gene activation in transformed cells.

Many changes have been reported in the behaviour of cells following transformation by SV40 and other agents. For reviews, the reader is referred to Martin (1981) and Tooze (1980). In most cases it appears that these changes, usually the apparent induction of host functions, are secondary effects on cell metabolism caused by uncontrolled growth of the transformed cells, there being little evidence for gene activation as a mediator of these changes (Martin, 1981).

There exist, however, limited data pertaining to the activation of particular cellular genes. Infection of mouse/ human cell hybrids with SV40 causes induction of the normally quiescent ribosomal complement, irrespective of which of the two sets are normally expressed in the hybrid (Soprano et al., 1980a,b). This effect is dependent upon functional large T-antigen but occurs independently of small t-antigen. In this case it may be argued that large T protein, whether free or in complex with other cellular proteins, may interact directly with, at least, ribosomal genes to directly influence their

expression. An effect on the expression of cellular protein-coding mRNA species has also been established. Williams et al. (1977) showed, by RNA/cDNA cross hybridization, that activation of a small number of poly A^+ RNAs (of the order of 30 average sized messages) had occurred in SV40-transformed human fibroblasts, whereas expression of the vast majority of cellular poly A^+ RNAs remained unaltered, within the limits of detection of the technique, which is insensitive to small fluctuations in relative abundance of mRNAs.

Limited gene activation may not be a characteristic unique to SV40 transformation. Groudine and Weintraub (1980) have shown that transformation of primary chick embryo fibroblasts by Rous sarcoma virus is accompanied by activation of about 1,000 nuclear transcription units of average size. Unfortunately they did not analyze what proportion, if any, of these induced transcripts became transported to the cytoplasm as mature poly A^+ RNA. The relevance of these observations to those of Williams et al. (1977) is therefore in doubt.

It should also be mentioned that gene activation has long been suspected to occur in tumours which often evidence ectopic production of hormones not normally present in the tissue of origin. There is yet, however, no evidence that the hormone accumulation observed in these tumours results from a mechanism of gene activation (for review see Baylin and Mendelsohn, 1980).

1.5 Tumour promoters and neoplastic progression.

Tumour promoters, or cocarcinogens, are compounds which show negligible carcinogenic activity but which, when administered in the presence of a known carcinogen, greatly lower the dosage of carcinogen required for the induction of tumours in animals. Their role appears to be one of accelerating the development of the neoplastic process

following a primary carcinogenic event (Diamond et al., 1980; Cairns, 1981). The promotion of neoplastic transformation by these compounds shows many similarities to the development of natural tumours and the progression of the neoplastic process (Pitot and Sirica, 1980; Ponder, 1980).

A large number of biochemical and phenotypic effects have been ascribed to the tumour promoters, particularly the phorbol esters such as 12-O-tetradecanoyl phorbol-13-acetate (TPA). These include stimulation of DNA, RNA and protein synthesis (Baird et al., 1971; Yuspa et al., 1976a,b; Dicker and Rozengurt, 1978), increased plasminogen activator secretion (Wigler and Weinstein, 1976) and acquisition of anchorage independent growth (Colburn et al., 1979). Many of these changes are observed in cells transformed or even infected by oncogenic viruses such as SV40 (Tooze, 1980; Martin,1981) suggesting that SV40 may exhibit promoter-like activity, perhaps triggering in cells expressing SV40 products the same general response as occurs following exposure to a tumour promoter. The end result in both cases would be to allow activation of particular cellular genes which could enhance expression or bring about de novo expression of transformation parameters that confer a selective advantage upon the cell.

In support of this view, changes in the expression of particular cellular genes have been observed following treatment of cells with tumour promoters. TPA has been shown to elicit specific changes in polypeptide synthesis (Laszlo et al., 1981; Laskin et al., 1981) and in protein phosphorylation (Laszlo et al., 1981). However, most, if not all of these changes were modulations in relative amount of polypeptide rather than representations of de novo synthesis. The effect of TPA-treatment on the synthesis and excretion of the low

molecular weight major excreted protein (MEP) has also been studied (Gottesman and Sobel, 1980). This protein is barely secreted by non-transformed cells, yet may comprise up to 1% of the total protein synthesized by cells transformed by a number of agents (Gottesman, 1978). Treatment of NIH3T3 cells with TPA resulted in a 5-10 fold increase in secretion of MEP. Pulse-labelling experiments showed that this phenomenon was accompanied by an increase in the rate of MEP synthesis and, furthermore, was reproduced when other promoters were substituted but was not observed following treatment with non-promoting analogues. Significantly, an increase in the level of translatable mRNA was indicated, implying an effect at the level of mRNA synthesis. Thus tumour promoters may indeed serve to activate particular cellular genes by elevating the level of expression of cytoplasmic mRNA. It should be noted that although Gottesman and Sobel failed to notice any other changes in protein biosynthesis after TPA-treatment, the technique they employed was very insensitive and can in no way exclude the possibility that induction of other genes might occur.

Varshavsky (1981a) suggested that tumour promoters might act in cocarcinogenesis by increasing the probability of activation of specific cellular genes which when over expressed would alter the phenotype of the cell in a selectable manner. Varshavsky proposed that this could be accomplished by the anachronistic "misfiring" of origins of replication, leading to local DNA amplification. Eukaryotic genes often show a dosage effect upon expression; thus overproduction of dihydrofolate reductase (DHFR) following treatment with the cytotoxic drug methotrexate is usually accompanied by a parallel increase in the number of copies of DHFR genes in the cell (Alt et al., 1978). In a similar fashion, genes amplified by

TPA-promotion would exhibit elevated expression which by consideration of the data discussed in Section 1.4, might be sufficient to induce a switch to a more advanced stage of neoplastic development.

That gene amplification can indeed be observed following growth of cells in the presence of tumour promoters has subsequently been demonstrated (Varshavsky, 1981b). As mentioned above, when cells in culture are exposed to methotrexate, a cytotoxic inhibitor of DHFR, spontaneously arising drug resistant variants may be selected which are frequently found to have sustained an amplification of the DHFR gene. Varshavsky observed that treatment of 3T6 cells with TPA increased by up to 100-fold the incidence of methotrexate resistant colonies following drug treatment. Application of the dot hybridization technique (Kafatos et al., 1979) showed that the drug resistant variants contained amplified DHFR coding sequences.

Thus tumour promoters may accomplish their objective, at least in part, by limited gene activation. As described earlier in this Section, there exist data which indicate a tumour-promotional activity for SV40, largely inferred by the similarity of biochemical and phenotypic changes that these different agents are able to induce. Perhaps most strikingly, TPA has been shown to cause reactivation of quiescent ribosomal genes in human/mouse hybrid cells (Soprano and Baserga, 1980) in exactly the same manner as has been described for SV40 (Soprano et al., 1980a,b). It should be noted also that the time course of induction of ribosomal genes and of MEP synthesis (Gottesman and Sobel, 1980) as far as can be ascertained from the published data, is probably too rapid to be explained by a gene amplification model. It seems reasonable that tumour promoters may effect a general response, perhaps one also induced by viruses such as SV40, and that this leads both to direct gene activation and indirect activation by

amplification. Implicit in this scheme is the very likely possibility that part of this response may occur independently of any effect upon mRNA expression, for example by influencing translation or processing of proteins, which could in part explain the changes observed by Laszlo et al.(1981) and Laskin et al.(1981).

1.6 Conclusion

There exists evidence that SV40 viral gene products may interact directly with cellular proteins and with sites on the genomic DNA. Part of the effect of transformation upon cellular behaviour may be explained by these interactions. One such effect may involve the activation of a limited subset of cellular genes which could either be involved directly in the expression of transformation or might serve to enhance transformation by encouraging the expression of growth properties not essential for basic transformation but representing an advancement of the neoplastic potential of the cell. The experiments to be described in Chapters 3, 4 and 5 will be addressed to ways of screening for changes in cellular gene expression, the application of these techniques to the specific problem of whether gene activation does, as has been proposed, accompany transformation by SV40 and finally to the task of characterisation and analysis of cloned segments of DNA obtained in the investigation.

CHAPTER II

MATERIALS AND METHODS

2.1 Mammalian Cells

Growth of Mammalian Cells in Culture

Cells were generally grown at 37°C in 9 cm petri dishes (Nunclon) in an incubator gassed with 5% (v/v) CO_2 in air. For the experiment described in Section 5.7 cells were grown at 32°C or 39.5°C as appropriate.

Mouse Balb/c 3T3 clone A31 cells were obtained from Dr. W.A. Nelson - Rees Naval Biomedical Research Laboratory, Oakland, California, and were grown in Dulbecco's modification of Eagle's medium (Flow Laboratories) supplemented with 10% (v/v) foetal calf serum (Gibco), 500 units ml^{-1} penicillin (Sigma) and 100 µg ml^{-1} streptomycin sulphate (Sigma). The SV40-transformed Balb/c 3T3 lines were grown as above with 2% (v/v) foetal calf serum, except in the experiments of Section 5.6 and Section 5.7 where the SV40 and SV40 (tsA) transformed lines were maintained on 10% (v/v) FCS.

Origins of SV40-transformed Cell Lines

The SV40-transformed mouse cell lines, SV3T3 C120, SV3T3 C126, SV3T3 C138, SV3T3 C149, SV3T3 C1H and SV3T3 C1M were derived following infection of confluent monolayers of Balb/c 3T3 clone A31 cells with virions of the wt 830 strain of SV40. Transformants were selected as dense foci in medium containing 2% (v/v) foetal calf serum; each cell line was cloned at least three times prior to further characterization (Rigby *et al.*, 1980; P.W.J. Rigby, M. Lovett , P.H. Gallimore and C.N. Cole, manuscript in preparation). Subclones of the SV3T3 C138 line were

provided by Mrs. V. Lewis, Imperial College, and were derived by plating in semi-solid medium (Methocel) and subsequently picking colonies. The SV3T3 A255B1b, SV3T3 A7B4b and SV3T3 A209B4a lines were supplied by Dr. W. Brockman, University of Michigan, Ann Arbor and are described in Brockman (1978).

2.2 Bacterial Strains

E. coli HB101 ($\underline{hsd}_K$ R^-, M^-, $\underline{Leu}^-$, $\underline{pro}^-$, $\underline{recA}^-$) was obtained from Dr. D. Glover, Imperial College. *E. coli* DH1 (recA1, nalA, $\underline{hsd}_K$ R^-, M^+, $\underline{endoI}^-$, $\underline{thi}^-$, relA1) was obtained from D. Hanahan, Harvard University.

2.3 Bacterial Plasmids

pAT153 (Twigg and Sherratt, 1980) was obtained from Dr. Sherratt, University of Sussex. pXf3 (D. Hanahan, personal communication) was obtained from D. Hanahan, Harvard University. pCD5 was constructed in this laboratory and is derived directly from pXf3 (M.R.D. Scott and P.W.J. Rigby, manuscript in preparation). pAT153 and clones based on this vector were propagated in *E. coli* HB101. pXf3, pCD5 and clones based on these vectors were propagated in *E. coli* DH1.

2.4 DNA Preparation

Mammalian Genomic DNA (Gross-Bellard *et al.*, 1973)

Tissue culture monolayers were lysed on the plate and incubated at 37°C overnight in a solution containing 50 mM Tris.HCl, pH 8.1, 1 mM EDTA, 0.5% (w/v) SDS and 100 μg ml^{-1} Proteinase K. The lysate was extracted with

TABLE 2.1

Media and solutions used for *E. coli* strains

L-broth	1% (w/v) Difco tryptone, 0.5% (w/v) Difco yeast extract, 0.5% (w/v) NaCl
L-agar	L-broth containing 1.5% (w/v) Difco agar
F-broth	L-broth containing 5% (w/v) glycerol
F-agar	L-agar containing 5% (w/v) glycerol
ψ-broth	2% (w/v) Difco tryptone, 0.5% (w/v) Difco yeast extract, 10 mM NaCl, 20 mM $MgSO_4$ adjusted to pH 7.6 with NaOH
ψ-agar	ψ-broth containing 1.5% (w/v) Difco agar
Tfb I	30 mM potassium acetate, 50 mM $MnCl_2$, 100 mM RbCl, 10 mM $CaCl_2$, 15% (w/v) glycerol adjusted to pH 5.8 with glacial acetic acid
Tfb II	10 mM NaPIPES buffer pH 6.5, 10 mM RbCl, 75 mM $CaCl_2$, 15% (w/v) glycerol

TABLE 2.2

Standard buffers and solutions used

Fragmented salmon sperm DNA	Dissolved to 10 mg ml^{-1} in 0.1M NaOH, 1 mM EDTA and boiled for 20 min.
Gel loading buffer	60% (w/v) sucrose, 1.5% (w/v) Ficoll, 25 mM EDTA (neutralized with NaOH), 0.2% (w/v) bromocresol purple, xylene cyanol or orange G
TD	25 mM Tris.HCl, pH 7.4, 140 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 1mM EDTA
TE	10 mM Tris.HCl, pH 8.1, 1 mM EDTA
TSE	10 mM Tris.HCl, pH 8.1, 100 mM NaCl, 1 mM EDTA
TAE	40 mM Tris, 5 mM sodium acetate, 1 mM EDTA adjusted to pH 7.8 with acetic acid
TBE	89 mM Tris, 89 mM boric acid, 2.5 mM EDTA
MAE	20 mM MOPS, 5 mM sodium acetate, 1 mM EDTA, adjusted to pH 7.2 with NaOH
PM (x1)	0 02% (w/v) polyvinylpyrrolidone, 0.02% (w/v) bovine serum albumin, 0.02% (w/v) Ficoll
SSC (x1)	150 mM NaCl, 15 mM trisodium citrate

neutral phenol (equilibrated with 0.5 M Tris.HCl, pH 8.9) and dialysed against TE. Pancreatic RNase was added to a final concentration of 150 μg ml^{-1} and the solution was incubated at 37°C for 3 to 6 hours. The solution was made 0.5% (w/v) in SDS and 100 μg ml^{-1} in Proteinase K and incubation was continued at 37°C overnight. Following two extractions with neutral phenol, the DNA solution was dialysed extensively against TE.

Bulk Preparation of Plasmid DNA

500 ml of late log-phase (OD_{600} = 0.6-1.0) cultures of *E. coli* strains carrying plasmids were grown at 37°C in L-broth containing 100 μg ml^{-1} ampicillin (Sigma). Following amplification overnight at 37°C with 200 μg ml^{-1} chloramphenicol (Sigma), the cells were harvested by centrifugation at 6,000 rpm for 10 min. at 4°C in a Sorvall GS-3 rotor. The cellular pellet was resuspended in 5 ml of 50 mM glucose, 10 mM EDTA, 25 mM Tris.HCl, pH 8.1. 1 ml of the above buffer containing 10 mg ml^{-1} lysozyme (Sigma) was added and the mixture was left at room temperature for 2 min., then transferred to 4°C. 12 ml of freshly prepared 0.2 M NaOH, 1% (w/v) SDS were added with gentle mixing. 6 ml of ice-cold 5 M potassium acetate, pH 4.8 was then added with gentle mixing. After 15 min. on ice, the clot of denatured chromosomal DNA and cell debris was removed by centrifugation for 20 min. at 15,000 rpm in a Sorvall SS-34 rotor, at 4°C. 6 ml of 50% (w/v) polyethylene glycol (average molecular weight 6,000) were added to the supernatant and the mixture was left on ice for 30 min.. The plasmid DNA precipitate was collected

at 8,000 rpm for 10 min. at 4°C in a Sorvall HB-4 rotor and resuspended in 20 ml of TE by occasional mixing during a period of 1 hour on ice. In this solution were dissolved 20 g of CsCl and 0.4 ml of 10 mg ml^{-1} ethidium bromide (Sigma) and the solution was transferred to a 40 ml Beckman VTi tube and centrifuged at 40,000 rpm for 16 hours at 15°C under mineral oil in a Beckman VTi 50 rotor. The plasmid DNA was collected in a 5 ml syringe by side puncture under long-wavelength ultra-violet illumination. CsCl and ethidium bromide were removed by dialysis against several changes of TE followed by one extraction with neutral phenol. The DNA was precipitated by addition of 0.1 volumes sodium acetate (pH 5.5) and 2.5 volumes ethanol and was collected by centrifugation at 10,000 rpm for 10 min. at 4°C in a Sorvall HB-4 rotor after at least 1 hour at -20°C. Following drainage and brief drying under high vacuum, the pellet was resuspended in 0.5 - 1.0 ml of TE.

Small Scale ("Miniprep") Preparation of Plasmid DNA

1 ml of late log-phase cultures of *E. coli* strains bearing recombinant plasmids was added to 1 ml of L-broth containing 400 µg ml^{-1} of chloramphenicol (Sigma) and incubated overnight at 37°C with gentle (150 rpm) agitation. 1.5 ml of each culture were centrifuged in an Eppendorf microfuge for 2 min. at 4°C. The cells were resuspended in 90 µl of 50 mM glucose, 10 mM EDTA, 25 mM Tris.HCl, pH 8.1. 10 µl of 10 mg ml^{-1} lysozyme in the same buffer were added; the solution was left at room temperature for 2 min., then transferred to 4°C. After 5 min. on ice, 150 µl of ice-cold potassium acetate, pH 4.8, were added

and the mixture left for 15 min. on ice. The precipitate was removed by centrifugation for 5 min. in an Eppendorf microfuge at 4°C and the supernatant was transferred to a fresh tube containing 720 µl ethanol. This was left on ice for 2 min. and the DNA was collected by centrifugation for 5 min. at 4°C in an Eppendorf microfuge. The plasmid DNA pellet was drained, dried briefly under high vacuum and reprecipitated from 100 µl of 0.2 M sodium acetate, pH 5.5 with 200 µl ethanol. After collection as above, the DNA was resuspended in 50 µl of TE.

2.5 Preparation of Poly A^+ Cytoplasmic RNA

Total Cytoplasmic RNA

Cells were harvested at 25% to 50% confluence. The monolayer was washed twice in TD at room temperature and incubated for 10 min. at room temperature in TD containing trypsin (0.025%) and 25 mM EDTA. Cells were collected by centrifugation at 2,000 rpm for 5 min. at room temperature in a Sorvall HB-4 rotor and resuspended in TSE at 4°C (0.5 ml per 90 cm dish). $MgCl_2$ was added to 5 mM then NP40 to 0.5% (v/v). Following vortexing for 10-20 sec., the nuclei were removed by centrifugation at 8,000 rpm for 2 min.. The supernatant was extracted at 65°C for 10 min. with an equal volume of water saturated phenol, in the presence of 0.5% (w/v) SDS and 25 mM EDTA. The aqueous phase was re-extracted once with 2 volumes of 1:1 phenol/$CHCl_3$ mixture and the RNA was precipitated by addition of sodium acetate, pH 5.5 to 0.2 M and 2.5 volumes of ethanol; this was stored at -20°C until required.

The above procedure (Method I) was used to prepare

the Balb/c 3T3 and SV3T3 C138 RNA used for all of the experiments of Chapters 3 and 4 and the initial comparisons of Section 5.3. In later experiments a simplified procedure was used (Method II). In this, the NP40 lysate was added directly to the phenol/SDS/EDTA misture. In the subsequent extraction with phenol at 65°C the nuclei precipitate largely intact onto the interface.

Selection of Poly A^+ RNA on Poly U-Sepharose

Preparation of total cytoplasmic RNA stored in 70% (v/v) ethanol were collected by centrifugation at 10,000 rpm for 10 min. at 4°C in a Sorvall HB-4 rotor, drained, dried briefly under high vacuum and resuspended in a small volume of ddH_2O. The RNA solution was then adjusted to be 10 mM Na(PIPES) buffer, pH 7.0, 10 mM EDTA, 25% (v/v) formamide, 0.7 M NaCl, 0.1% (w/v) SDS and passed three times over poly-U sepharose (Miles Laboratories) at room temperature. Generally the RNA solution was 1-2 mg ml^{-1} and the columns (poured in disposable syringes) contained 0.5-1.0 ml of swollen gel per mg of RNA. The gel was washed with 5 bed volumes of 10 mM Na(PIPES) buffer, pH 7.0, 10 mM EDTA, 25% (v/v) formamide, 0.7 M NaCl, 0.1% (w/v) SDS, followed by 10 bed volumes of 10 mM Na(PIPES) buffer, pH 7.0, 10 mM EDTA, 50% (v/v) formamide, 0.5 M NaCl, 0.1% (w/v) SDS. The bound poly A-containing fraction was eluted with 10 mM Na(PIPES) buffer, pH 7.0, 10 mM EDTA, 95% (v/v) formamide, 0.1% (w/v) SDS. Fractions (0.2 bed volume) were collected and 1 µl aliquots were dried onto a 1% (w/v) agarose gel containing 0.5 µg ml^{-1} of ethidium bromide. The peak fractions were located under long-wave ultra-violet

illumination and pooled. Ammonium acetate was added to 0.5 M together with 2.5 volumes of ethanol and the poly A^+ RNA was left to precipitate at -70°C for 10-30 min. Following centrifugation, either for 5 min. at 4°C in an Eppendorf microfuge or at 10,000 rpm for 10 min. at 4°C in a 15 ml Corex tube using a Sorvall HB-4 rotor, the RNA was drained, dried briefly under high vacuum and resuspended in 0.2 M sodium acetate, pH 5.5. 2.5 volumes of ethanol were added and the RNA stored as a precipitate at -20°C until required.

2.6 Enzyme Reactions

Restriction Enzyme Digestion Conditions (Roberts, 1978)

These are listed in Table 2.3. In general, DNA was included at 50-200 μg ml^{-1}. Double digests were performed sequentially, starting with the enzyme requiring the lowest ionic strength. After completion of the first reaction, the buffer was adjusted to the appropriate conditions for the second digest. Reactions were terminated by addition of EDTA to a final concentration of 25 mM and, sometimes, extraction with neutral phenol.

All restriction endonucleases used were assayed and supplied by C. Headhouse-Benson, C. Watson and L. Woods of this department and digests were performed using a two-fold excess of enzyme.

Nick Translation (Rigby *et al.*, 1977)

Final reaction conditions were 50 mM Tris.HCl, pH 7.4, 5 mM $MgCl_2$, 15 μM unlabelled dNTPs (usually dATP and dGTP), 7.5 μM α-^{32}P-labelled dCTP and dTTP (Amersham; specific activity = 200-400 Ci $mmol^{-1}$), 15 μg ml^{-1} of DNA, 10 ng ml^{-1}

TABLE 2.3

Restriction Enzyme Digestion Conditions

Enzymes	Buffer	Temperature
Bst I, Taq I	6 mM Tris.HCl, pH 7.6; 6 mM $MgCl_2$; 6 mM 2-mercapto-ethanol	65°C
EcoRI, HindIII, Hinf I	10 mM Tris.HCl, pH 7.6; 10 mM $MgCl_2$; 50 mM NaCl; 1 mM DTT.	37°C
Sal I	6 mM Tris.HCl, pH 7.6; 6 mM $MgCl_2$; 100 mM NaCl; 6 mM 2-mercaptoethanol	37°C
Cla I	10 mM Tris.HCl, pH 8.1; 10 mM $MgCl_2$	37°C

DNase I and an optimal amount of pure *E. coli* DNA polymerase I provided by Professor A. Fersht, Imperial College. Incubation was at 14°C for 45-90 min. and the labelled DNA was purified by addition of EDTA to a final concentration of 25 mM followed by passage over a 1 ml column of Sephadex G50M (Pharmacia) in TE. The excluded peak of radioactivity was pooled, denatured by addition of NaOH to a final concentration of 0.1 M and diluted as required into hybridization reactions. For nick translations of the "mini-prep" DNAs used in Section 5.2, 5 µl of the preparation was used in a total reaction volume of 10 µl.

Routinely 50% of labelled dNTP was incorporated into the G50M-excluded fraction, giving a final sp. act. of $5 \times 10^7 - 2 \times 10^8$ cpm μg^{-1} DNA.

Labelling of RNA Using Polynucleotide Kinase

RNA was partially hydrolysed by incubation for 1 hour at 4°C in 0.1 M NaOH. 0.1 volumes of 2 M acetic acid were then added and the RNA precipitated at -70°C for 10-30 min. following addition of 2.5 volumes of ethanol. The RNA was resuspended in a buffer containing 50 mM Tris.HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM spermidine (Sigma), 15 mM DTT (Sigma) and 2 mCi ml^{-1} of γ-^{32}P-ATP (sp. act. = >1000 Ci $mmol^{-1}$, Amersham). T4 polynucleotide kinase (donated by L. Woods, C. Headhouse-Benson and C. Watson) was added to 25-50 units ml^{-1} and the reaction incubated at 37°C for 1 hour. EDTA was added to a final concentration of 25 mM and the labelled RNA purified by chromatography over Sephadex G50M in TSE. RNA was included in most reactions at 50 µg ml^{-1} and specific activities of 2-3 x 10^7 cpm μg^{-1} were routinely obtained.

Synthesis of cDNA

This was performed in a buffer containing 50 mM Tris.HCl pH 8.3, 6 mM magnesium acetate, 5 mM DTT, 1 mM dATP, 1 mM dGTP, 1 mM dTTP, 20-200 μM dCTP (see below), 100 μg ml^{-1} actinomycin D (Sigma), 5 μg ml^{-1} oligo dT (Collaborative Research), 10-25 μg ml^{-1} poly A^+ RNA, and 25 units ml^{-1} of rat liver ribonuclease inhibitor, (obtained from Dr. S. Emtage, G.D. Searle, High Wycombe.) Reverse transcriptase (AMV), provided by Dr. J. Beard, Life Sciences Inc., St. Petersburg, Florida, was added to 250 units ml^{-1} and reactions were incubated at 42°C for 30-60 min.. Where these conditions were varied, they are specified in the text. For cloning experiments, the concentration of dCTP was 200 μM, using a proportion of labelled nucleotide sufficient to give a final cDNA specific activity of 10^5 - 10^6 cpm μg^{-1} of cDNA (assuming an average dCMP composition of 25%). For analytical reactions, where higher specific activities were desired, ^{32}P-labelled dCTP was included to give a final sp. act. of 10^6 - 10^7 cpm μg^{-1} and the concentration of dCTP was reduced to 50 μM. For high specific activity labelling (Section 4.3), the dCTP was used at a specific activity of 50-100 Ci $mmol^{-1}$ at a final concentration of 20 μM. The RNA concentration was increased to 50 μg ml^{-1} to ensure maximal incorporation of labelled nucleotide. Routinely, 20-50% of ^{32}P-labelled dCTP was incorporated under these conditions, with the calculated specific activity usually $>10^8$ cpm μg^{-1} of cDNA.

Double Stranded cDNA Synthesis

The first strand reaction mixture was adjusted to be 10 mM in EDTA and 3 M NaOH was added to give a final concentration of 0.3 M, followed by incubation at 37°C for 2 hours to hydrolyse the RNA template. The mixture was neutralized by addition of 0.5 volumes of 1 M acetic acid and passed over a 1 ml column of Sephadex G200M (Sigma or Pharmacia) in TSE buffer. The excluded peak was pooled and precipitated by addition of 0.1 volumes of 2 M sodium acetate, pH 5.5, and 2.5 volumes of ethanol. Following incubation for 4-16 hours at -20°C the purified cDNA was collected by centrifugation at 10,000 rpm for 30 min. at 4°C in a 30 ml siliconized Corex tube, using a Sorvall HB-4 rotor. The pellet was drained, dried briefly under high vacuum and resuspended carefully in a small volume of ddH_2O. The second-strand reaction contained 50 mM Tris.HCl, pH 8.3, 9 mM magnesium acetate, 20 mM DTT, 1 mM each of all four dNTPs (unlabelled) and 2-10 μg ml^{-1} of cDNA. Reverse transcriptase was added to 500 units ml^{-1} and the reaction was incubated at 46°C. The reactions, monitored by endonuclease S1 assay (see Section 2.12), routinely plateaued at 40-70% resistance to S1 after 3-4 hours of incubation.

The mixture was then adjusted to endonuclease S1 cleavage conditions by addition of 0.1 volumes of 300 mM sodium acetate, pH 3.75 (this adjusts the solution to a final pH of 4.5-4.6), 2.5 M NaCl, 45 mM $ZnSO_4$ and diluted by addition of 5 volumes of 30 mM sodium acetate, pH 4.55 250 mM NaCl, 4.5 mM $ZnSO_4$. Dilution is necessary since the high nucleotide concentrations used in the second-strand synthesis were found to severely inhibit endonuclease S1.

An optimal amount (2 units ng^{-1} of DNA; see Section 3.3) of endonuclease S1 (Sigma) was added and the reaction incubated at 20°C for 30 min.. The double-stranded cDNA was then purified by addition of nuclease-free BSA (a gift of Dr. D. Woods, St. Mary's Hospital Medical School) to 10 $\mu g\ ml^{-1}$, 0.1 volumes of sodium acetate, pH 5.5, EDTA to a final concentration of 10 mM and 2.5 volumes of ethanol. After 4-16 hours at -20°C the double-stranded cDNA was collected by centrifugation as described for the cDNA first-strand. Following resuspension in a small volume of TE, the cDNA solution was adjusted to be 0.1 M in NaCl and passed over a 1.5 ml column of Biogel A150M in TSE. Fractions (0.2 bed volumes) were collected and the position of the eluted cDNA established by Cerenkov counting. Usually only the leading two-thirds of the peak were retained for cloning (see Section 3.3).

Purified double-stranded cDNA was pooled and $MgSO_4$ (final concentration 10 mM) and isopropanol (2.5 volumes) were added. After freezing at -70°C for 15 min., the cDNA was collected by centrifugation by 5 min. at 4°C in an Eppendorf microfuge. The pellet was drained, dried briefly under high vacuum, resuspended at approximately 100 $\mu g\ ml^{-1}$ in TE and stored at -70°C till required.

Tailing of DNA Using Terminal Transferase

This was performed in a buffer containing 30 mM Tris, 140 mM cacodylic acid, 0.1 mM DTT, adjusted to pH 7.6 with NaOH. $CoCl_2$ (10 mM) was added to 1 mM immediately before addition of enzyme (1000 units ml^{-1}, obtained from Ratliff Biochemicals, Los Alamos, New Mexico). DNA to be tailed was included at 100-500 $\mu g\ ml^{-1}$ (vector DNA) or 20-100 $\mu g\ ml^{-1}$

(cDNA or λ fragments), thus giving a concentration of termini of 0.1-0.5 μM.

When tailing cDNA which had been previously quantitated using the dideoxynucleotide assay (see Sections 2.12 and 3.4), unlabelled dGTP was included at a 20-fold excess of nucleotide over termini (i.e. usually at 4-20 μM) and the reaction was incubated at 37°C for 1 hour. This procedure yields tailed cDNA which consistently gives acceptable efficiencies of cloning. In experiments where trace label was included, 40-60% of the nucleotide became incorporated giving an average tail length of optimal size (8-12 nucleotides per end).

When tailing vector DNA and insert fragments which had not been previously assayed, ^{3}H-dCTP (or other dNTP) was included at 50 μM, usually at a specific activity of 1-10 Ci mmol^{-1}. Pilot reactions were performed using one tenth or one twentieth of the reaction mix, monitored by incorporation of ^{3}H-labelled nucleotide with time (as described in Section 2.12) and the number of residues added per end calculated from the known concentration of nucleotide (50 μM) and the estimated concentration of termini. The remainder of the reaction was tailed for the time required to add the appropriate number of residues and was then diluted with several volumes of TSE, extracted once with neutral phenol and three times with five volumes of water-saturated diethyl ether. Residual ether was removed with a gentle stream of N_2 gas.

Preparations of cDNA were stored at -70°C without further treatment but preparations of tailed vector DNA were routinely passed over a 1 ml column of Sephadex G-200M in TSE

buffer and stored at 4°C till required.

End Labelling With Reverse Transcriptase

This was performed in a buffer containing 50 mM Tris.HCl pH 8.3, 10 mM $MgCl_2$, 5 mM DTT, 1 mM unlabelled dATP and dGTP and 100-200 µCi ml^{-1} of α-^{32}P-dCTP and dTTP (Amersham), sp. act. = 200-400 Ci $mmol^{-1}$. Restriction endonuclease digests were adjusted to these conditions and incubated with reverse transcriptase (1000 units ml^{-1}) at 42°C for 30 min. Samples were loaded directly onto 8% (w/v) polyacrylamide slab gels. End-labelled plasmid or bacteriophage λ DNA marker fragments were prepared using 0.1 µg of digested DNA in a final volume of 20 µl. Fingerprinting analyses were performed as described in Chapter 5.2.

2.7 Electrophoretic Techniques

Neutral Agarose Gels (Favoloro *et al.*, 1980)

Gels containing 0.8-2.0% (w/v) agarose were cast and electrophoresed in TAE buffer (Table 2.2). Gels usually contained 0.5 µg ml^{-1} ethidium bromide and were electrophoresed submerged at 1.5 V cm^{-1} for 12-18 hours. 0.05 volumes of gel sample buffer were added to all samples prior to loading (Table 2.2).

Alkaline Agarose Gels (Favoloro *et al.*, 1980)

These contained 2.0% (w/v) agarose and were cast in 50 mM NaCl, 1 mM EDTA and were electrophoresed submerged in 30 mM NaOH, 1 mM EDTA, usually at 1.5 V cm^{-1} for 12-18 hours. If staining was required (to visualize DNA marker

fragments), the gel was soaked in 0.2 M ammonium acetate containing ethidium bromide (0.5 μg ml^{-1}). Gels to be autoradiographed were fixed in 10% (v/v) acetic acid for 30 min. and dried down on a BioGrad gel drier. 0.05 volumes of gel sample buffer (Table 2.2) were added immediately prior to loading.

Formaldehyde Agarose Gels (Lehrach *et al.*, 1977)

These contained 1% (w/v) agarose, and were prepared by boiling the agarose in ddH_2O, cooling it to 55°C and adjusting it to be 3% (w/v) in formaldehyde in MAE buffer (Table 2.2) Samples to be electrophoresed were adjusted to 50% (v/v) formamide, 6% (w/v) formaldehyde in MAE buffer, heated to 65°C for 5 min. and cooled to room temperature. 0.05 volumes of gel loading buffer were added immediately prior to loading. Electrophoresis was for 5 hours at 7.5 V cm^{-1} or 15-20 hours at 1.5 V cm^{-1}. Gels to be stained (for visualization of RNA markers) were soaked twice for 30 min. each in 5 gel volumes of 0.2 M acetic acid. Staining was then achieved by soaking the gels in 0.2 M ammonium acetate, containing 1.0 μg ml^{-1} ethidium bromide.

Polyacrylamide Thin Gels (Sanger and Coulson, 1978)

These contained 7.6% (w/v) acrylamide and 0.4% (w/v) bis-acrylamide cast in TBE buffer (Table 2.2). Polymerization was effected by addition of 0.05% (v/v) TEMED and 0.05% (w/v) ammonium persulphate immediately prior to pouring. Samples to be loaded were mixed with 0.05 volumes of gel loading buffer before loading.

Electrophoresis was performed in TBE buffer at 5 V cm^{-1} for 8 hours.

Two Dimensional Agarose Gels (Favoloro *et al.*, 1980)

Gels containing 1.5-2.0% (w/v) agarose were cast and run in TAE buffer as described for neutral agarose gels, using a single 2 mm diameter cylindrical sample gel slot. After running the first dimension, the gels were equilibrated in 5-10 gel volumes of alkaline electrophoresis buffer (30 mM NaOH, 1 mM EDTA) for >1 hour. The second dimension was electrophoresed and the gel dried as described for alkaline agarose gels.

Size Markers for Gel Electrophoresis

In general restriction digests of plasmid or bacteriophage λ DNA were used. Marker DNA fragments for formaldehyde agarose gels were calibrated if necessary by comparison with RNA standards of known length.

Preparation of Filters for DNA Gel Transfer Hybridization

Transfer of DNA from agarose gels in TAE buffer to nitrocellulose was performed essentially as described by Southern (1975) incorporating some of the modifications of Wahl *et al.* (1979). Gels were treated twice for 15 min. with 5 gel volumes of 0.25 M HCl. They were then rinsed briefly with water, treated twice for 30 min. with 5 volumes of 0.5 M NaOH, 1.5 M NaCl and neutralised in 5 volumes of 0.5 M Tris.HCl, pH 5.5, 3 M NaCl (2 changes over a total of 1-2 hours).

Nitrocellulose sheets (Schliecher and Schüll)

were wetted in water and rinsed in 2x SSC for 30 min. before use. Transfer was performed overnight using 20x SSC. Following transfer the sheets were rinsed in 2X SSC, blotted dry and baked for two hours at 80°C.

Preparation of Filters for RNA Gel Transfer Hybridization

Transfer of nucleic acids from formaldehyde agarose gels was performed by a modification of the procedure of Thomas (1980). The gel was placed untreated onto a conventional transfer apparatus and a dry nitrocellulose sheet laid on top of it. Transfer was started immediately by placing on top of the filter a pile of filter paper sheets. 20x SSC was used as transfer buffer. After 4-20 hours, the filter was removed, rinsed in 10x SSC, blotted dry and baked at 80°C for 2 hours.

2.8 Preparation of Filters for Colony Hybridization

82 mm diameter Millipore HAWP filters were used. They were boiled for 30 min. in ddH_2O and sandwiched, still moist, between two layers of Whatman 3 MM paper. Stacks of 5-10 filters were wrapped in aluminium foil and pressure cooked for 10 min. at 15 lb in^{-2} and stored at 4°C till required.

Ordered Arrays (Grunstein and Hogness, 1975)

Colonies to be tested were stabbed in L-agar plates containing 100 μg ml^{-1} ampicillin and grown overnight at 37°C. A sterile 82 mm filter was laid onto each plate, removed and inverted before being placed on a fresh L-agar plate containing 500 μg ml^{-1} chloramphenicol. The colonies

largely adhere to the filter leaving some cells stabbed into the agar; these can be retrieved as required. Filters were incubated overnight at 37°C and treated exactly as described for those bearing unordered arrays of colonies.

Unordered Arrays of Colonies (Hanahan and Meselson, 1980)

Cells were plated onto 82 mm filters either directly from transfection experiments as described in Section 2.11 or by spreading from previously titred frozen stocks. The filters were incubated at 37°C on L-agar plates containing 100 μg ml^{-1} ampicillin till a colony size of 0.5 mm was attained. Several replicas of each filter were made as follows. The template filter was removed and placed, colony side upwards, onto several sheets of sterile blotting paper. A sterile 82 mm filter was moistened on L-agar containing 100 μg ml^{-1} ampicillin and laid, moistened side downward, on top of the master filter. The two were then covered with blotting paper and firmly pressed together. The two filters were then keyed together using a scalpel blade, peeled apart and replaced on their original plates. The master filter was then used for a second cycle of replication, then grown for 1 hour at 37°C before being used to prepare a third and fourth replica exactly as described above.

The original master filter was stored at 4°C on its agar plate and the replicas were incubated at 37°C till a colony size of 1-1.5 mm was reached. They were then transferred to L-agar plates containing chloramphenicol (500 μg ml^{-1}) and incubated overnight at 37°C.

For long term storage of unordered colony arrays, the

master filters were grown in F-agar containing antibiotic. A second filter, prewetted on a similar plate, was pressed against it and the "sandwich" placed between sheets of sterile blotting paper, sealed in aluminium foil and stored at -70°C until required. They were then put at room temperature for several minutes and the filters were peeled apart and laid on fresh plates.

Treatment of Filters for Hybridization

Filters bearing ordered or unordered arrays of colonies were placed colony side upward onto blotting paper saturated with 0.5 M NaOH for 5 min., transferred to fresh paper saturated with 0.5 M NaOH for 5 min. then treated in the same way with 1 M Tris.HCl, pH 7.4 (twice for 5 min. each) and with 0.5 M Tris.HCl, pH 7.4, 1.5 M NaCl (once for 15 min.). They were then rinsed in 2x SSC and left colony side up on blotting paper till dry. The filters were then baked at 80°C for 2 hours and hybridized as described in Section 2.10.

2.9 Preparation of DNA-Cellulose (Noyes and Stark, 1977)

4 mg of plasmid DNA were diluted into 3 ml of TE. 0.75 ml of 1 M HCl were added and the DNA partially depurinated for 5 min. at room temperature. 5.25 ml of 1 M NaOH were added and the mixture left at room temperature for 1 hour to complete cleavage at depurination sites and to denature the DNA. The DNA becomes fragmented to an average size of 1.5 kb under these conditions. 1 ml of 5 M acetic acid was then added, together with 20 ml of ethanol. The DNA was left at -20°C overnight and the

precipitate collected by centrifugation at 10,000 rpm for 15 min. in a Sorvall HB-4 rotor at 4°C. The DNA was resuspended in a small volume of 25 mM sodium phosphate buffer, pH 6.5 and four volumes of DMSO were added.

72 mg (dry weight) of m-amino benzyloxymethyl cellulose (Miles Laboratories) were dissolved in 7.2 ml of ammoniacal $Cu(OH)_2$ (0.45 g $Cu(OH)_2$, 0.1 g sucrose and 6 ml ammonia solution per 10 ml solution). A further 3.6 ml of ammonia solution were added and undissolved material removed by centrifugation (5 min., 10,000 rpm at 20°C in a Sorvall HB-4 rotor). 6 volumes of ddH_2O (32 ml) were added followed by heating to 70°C. The cellulose was then precipitated by addition of 10% (v/v) H_2SO_4 at 70°C dropwise to pH = 6.0, centrifuged as above and washed 3 times with ddH_2O at 4°C. The cellulose was then resuspended in 8 ml of ddH_2O at 4°C and 16 ml of 1.8 M HCl and 0.64 ml of a freshly prepared solution of $NaNO_2$ (10 mg ml^{-1} in ddH_2O) were added. The reaction mixture was stirred at 0°C for 30 min. then excess HNO_2 was destroyed by adding solid urea until the reaction with starch-iodide paper was almost negative. The diazotized cellulose was collected by centrifugation, washed once each with ddH_2O, cold 25 mM sodium phosphate, pH 6.5, and 80% (v/v) DMSO containing 5 mM sodium phosphate, pH 6.5. The cellulose was then resuspended directly in the DNA solution at a cellulose concentration of >10 mg ml^{-1} (dry weight). Approximately 10^5 cpm (1 ng) of ^{32}P-labelled bacteriophage λ DNA were included to monitor the binding reaction which was incubated for 3 days at 4°C with gentle shaking. The DNA-cellulose was collected by centrifugation and washed twice in 80% (v/v) DMSO at 50°C

4 times in 0.1x SSC at room temperature, twice in 99% (v/v) formamide, 0.1% (w/v) SDS at 80°C and finally twice in hybridization buffer (see Section 2.8), in which it was stored at 4°C until required. The DNA-cellulose was washed with fresh hybridization buffer before use in hybridization experiments.

2.10 Hybridization Reactions

Filter Hybridization

All filters were hybridized as follows. The filter was sealed in a polythene bag containing hybridization buffer (50% (v/v) formamide, 50 mM sodium phosphate buffer, pH 6.5, 200-1,000 μg ml^{-1} denatured fragmented salmon sperm DNA, 5x SSC, 5x PM and 0.1% (w/v) SDS. 0.5 ml of buffer were used per cm^2 of filter. Prehybridization was for 4-16 hours at 42°C. The bag was then opened, the radioactive probe was added and the bag resealed. Hybridizations with ^{32}P-labelled DNA labelled by nick translation (Section 2.4) used 10-25 ng ml^{-1} of DNA of a specific activity of 0.5 - 2.0 x 10^8 cpm μg^{-1}. Hybridizations with kinase-labelled RNA (Section 2.4) used 10^7 cpm ml^{-1} (350 ng ml^{-1}) of probe. Filter hybridizations with RNA probes were further supplemented with 20 μg ml^{-1} poly rA (Sigma) and ones with cDNA probes used 20 μg ml^{-1} poly rU (Sigma).

Hybridization was for 12-16 hours at 42°C. The filters were then washed 3 times for 10 min. each in 2x SSC, 0.1% (w/v) SDS at room temperature, 4 times for 30 min. each at 60°C in 2x SSC 0.1% (w/v) SDS and 6 times for 30 min. each in 0.1x SSC, 0.1% (w/v) SDS at 55-60°C. Colony hybridization filters were further rinsed at room temperature in 3 mM Tris

(twice for 15 min. each). The filters were then blotted dry and autoradiographed at -70°C using preflashed Fuji X-ray film and Fuji Mach II intensifying screens as described by Laskey and Mills (1977).

Hybridization to DNA-Cellulose

This was performed in cellulose hybridization buffer, containing 40 mM NaPIPES, pH 6.5, 1 mM EDTA, 0.6 M NaCl, 0.1% (w/v) SDS, 0.25 mg ml^{-1} poly rU, 0.5 mg ml^{-1} yeast tRNA. The reaction were incubated at 37°C, usually for 24 hours, with gentle shaking. Hybridizations contained ~180 mg ml^{-1} of cellulose and ~8 mg ml^{-1} plasmid DNA, corresponding to an estimated concentration of cDNA equivalents of ~1 mg ml^{-1}. ^{32}P-Labelled probe was used at ~1 µg ml^{-1}.

Following hybridization the cellulose was washed three times at room temperature with hybridization buffer lacking poly rU and tRNA. In these experiments, the unhybridized fraction and the washes were pooled. The hybridized material was eluted by resuspending the cellulose in 99% (v/v) formamide, 0.1 % (w/v) SDS and heating to 70°C for 2 min.. The cellulose was then pelleted (30 sec. in an Eppendorf microfuge at 20°C) and the elution repeated until no radioactive material remained. The pooled unhybridized fraction was precipitated by the addition of 2 volumes of ethanol, after 20 min. at -20°C, and was centrifuged at 10,000 rpm for 10 min. at 4°C in a Sorvall HB-4 rotor. The pellet was drained, dried briefly under high vacuum, resuspended in a small volume of ddH_2O and readjusted to hybridization conditions prior to a subsequent cycle of hybridization or

stored at -20°C until required.

mRNA/cDNA Hybridization In Solution (Bishop *et al.*, 1974b; Williams *et al.*, 1977)

Hybridization was performed in 10 mM NaPIPES buffer, pH 6.5, 1 mM EDTA, 0.2% (w/v) SDS, 2x SSC at 65°C. Analytical reactions (10-50 µl) contained 50-200 µg ml^{-1} of poly A^+ RNA and >10^7 cpm ml^{-1} of ^{32}P-labelled cDNA. Aliquots (1-2 µl) of the mixture were sealed in 10 µl siliconized glass microcapillaries, boiled for 3 min. and quenched in ice/water; the samples were then incubated together at 65°C. At appropriate times of incubation samples were removed, quenched in ice/water and expelled into 1 ml of cold S1 assay buffer prior to determination of the percentage nuclease resistance (see Section 2.12). Preparative reactions were performed in siliconized 1.5 ml Eppendorf microfuge tubes in a total volume of 50 µl. Poly A^+ RNA was included to 100 µg ml^{-1} and ^{32}P-labelled cDNA at approximately 2 µg (2 x 10^8 cpm) ml^{-1}. Hybridization was commenced by boiling for 3 min. in a water bath, quenching in ice/water and the reaction was incubated at 65°C to the appropriate R_0t value (calculated assuming that an R_0t = 1 corresponds to hybridization for 1 hour at a concentration of poly A^+ RNA of 83 µg ml^{-1}). The reaction was then quenched in ice/water, 5 ml of cold S1 assay buffer (see Section 2.12) was added together with a previously determined quantity of endonuclease S1 sufficient to produce a limit digest of single stranded cDNA under the same conditions. Following incubation at 20°C for 1 hour the reaction was terminated by addition of tRNA to 10 µg ml^{-1}, 0.1 volumes of sodium acetate, pH 5.5, EDTA to a final

concentration of 10 mM and 2.5 volumes of ethanol. After 12-16 hours at -20°C the cDNA was collected by centrifugation at 10,000 rpm for 30 min. in a Sorvall HB-4 rotor at 4°C. The pellet was drained, dried briefly under high vacuum and resuspended in 50 µl of TE. RNA was removed by addition of 3 M NaOH to a final concentration of 0.3 M followed by incubation at 65°C for 15 min.. The mixture was then neutralised by the addition of 0.5 volumes of 1 M acetic acid and the cDNA was passed over a 1.5 ml column of Sephadex G50M (Pharmacia) in TSE buffer. The excluded peak was pooled and stored at -20°C until required.

2.11 Molecular Cloning Techniques

Annealing of Homopolymer-Tailed DNA Molecules

Tailed vector and cDNA or other insert fragments, usually at a final DNA concentration of 0.2-1.0 µg ml^{-1}, were mixed in TSE at a molar vector:insert ratio of 2:1 (except where otherwise specified in the text). Annealing was achieved by successive incubations at 65°C for 2 min., 42°C for 2 hours followed by gradual cooling to room temperature over a period of 2-3 hours. The annealed DNA was then used directly for transfection.

Subcloning

Vector restriction fragments (0.1-1.0 µg) were prepared as described in Section 2.6. Target DNA fragments were purified from agarose gels. The gel slice was finely minced in an equal volume of TSE and dissolved in one volume of neutral phenol. The mixture was vortexed frequently over a period of 30 min. at 65°C and placed on ice for 1 hour. The phases were separated by centrifugation at 4°C for 10 min. in

an Eppendorf microfuge. The aqueous phase was removed, extracted once with 5 volumes of water-saturated diethyl ether, and DNA was precipitated by addition of $MgSO_4$ (to a final concentration of 10 mM) and 2.5 volumes of isopropanol. After being placed at -70°C for 30 min. the DNA was collected by centrifugation as above, drained, dried briefly under high vacuum and resuspended in TE.

Vector and target DNA fragments were then mixed at an approximately 1:3 ratio of vector DNA:target DNA in a buffer containing 50 mM Tris.HCl, pH 7.6, 10 mM $MgCl_2$, 20 mM DTT, 0.1 mM ATP at a final DNA concentration of 10-50 $\mu g\ ml^{-1}$. An optimal amount of T4 DNA Ligase (supplied and assayed by C. Watson, C. Headhouse-Benson and L. Woods) was added and the mixture incubated at 14°C overnight, hybrid molecules being recovered by transfection the following day. Subclones were identified by restriction endonuclease digestion and gel electrophoresis (see Sections 2.6 and 2.7) of plasmid "miniprep" DNA preparations made from 1 ml cultures (see Section 2.4).

Transfection

The process of introducing hybrid plasmid molecules into *E. coli* strains is referred to as transfection rather than transformation throughout this work to avoid confusion with oncogenic transformation defined at the outset of Chapter 1.

Competent *E. coli* HB101 cells were prepared as follows. A mid-log-phase culture of *E. coli* HB101 was adjusted to 5% (w/v) glycerol and stored frozen at -70°C. The day before competent cells were to be prepared, cells were

streaked from this stock onto an L-agar plate and incubated at 37°C overnight. A single large colony was picked, dispersed in 5 ml of L-broth and incubated with agitation (150 rpm) at 37°C until faintly turbid. This culture was then used to inoculate 100-500 ml of prewarmed L-broth, grown to an OD_{600} = 0.6 with agitation (150 rpm) at 37°C and harvested by centrifugation at 6000 rpm for 10 min. at 4°C in a Sorvall GS-3 rotor. The cells were gently resuspended in 0.5 original volumes of cold 10 mM NaPIPES buffer, pH 6.5, 10 mM RbCl. Following a second centrifugation, the cells were resuspended gently in 0.4 original volumes of 10 mM NaPIPES buffer, pH 6.5, 10 mM RbCl, 75 mM $CaCl_2$. The suspension was left on ice for 30 min. and recentrifuged. The cells were finally resuspended in 0.1 original volumes of 10 mM NaPIPES buffer, pH 6.5, 10 mM RbCl, 75 mM $CaCl_2$ containing 15% (w/v) glycerol. Aliquots (0.1-0.5 ml) were dispensed, using a pre-chilled pipette, into 1.5 ml Eppendorf microfuge tubes, snap frozen in liquid N_2, and stored at -70°C until required, whereupon aliquots were thawed on ice for >1 hour before addition of DNA.

Competent *E. coli* DH1 cells were prepared as follows. A mid-log-phase culture of DH1 cells in ψ-broth was frozen at -70°C following addition of glycerol to 5% (w/v). Cells were streaked from this stock onto ψ-agar plates and incubated overnight at 37°C. A single large colony was dispersed in 5 ml warm ψ-broth and incubated at 37°C with vigorous agitation (>250 rpm) to an OD_{550} of approximately 0.3. This culture was diluted into 100 ml of warm ψ-broth and incubated at 37°C at 350 rpm until an OD_{550} = 4.8 was reached. Cells were harvested by centrifugation at 6000 rpm for 5 min. at 4°C in

30 ml Corex tubes in a Sorvall HB-4 rotor, washed once in 0.4 original volumes of Tfb I (see Table 2.1) left on ice for 5 min. and recentrifuged as described above. The cells were then resuspended in 0.04 original volumes of cold Tfb II (see Table 2.1), aliquotted with a chilled pipette into 1.5 ml Eppendorf microfuge tubes and snap-frozen in liquid N_2. Aliquots were stored at -70°C, removed as required, thawed rapidly until just liquid then placed on ice for 10 min. before addition of DNA.

DNA (either in TSE or in ligation buffer) at a concentration of <1.0 μg ml^{-1} was added to at least four volumes of competent *E. coli* (prepared as described above). The mixture was left on ice for 20 min.. HB101 cells were then transferred to a water bath at 25°C for 10 min.. DH1 cells were heat-shocked at 42°C for 90 min.. The cells were then placed on ice for 1-2 min. then 4 volumes of broth (L-broth for HB101, ψ-broth for DH1) were added and the tubes incubated with gentle shaking (150 rpm) at 37°C for 50 min.. Samples were then either spread directly onto L-agar containing antibiotic (usually 100 μg ml^{-1} ampicillin) or diluted to 10 ml with further broth and filtered directly, with gentle suction, onto 82 mm diameter Millipore HAWP filters (see Section 2.8). The filters were then transferred, colony side up, onto L-agar containing antibiotic and incubated at 37°C overnight.

2.12 Quantitation of Nucleic Acids

Spectrophotometric Assays

Samples to be quantitated were diluted in TE and the absorbance at 260 nm was measured. Concentrations were calculated assuming that an OD_{260} = 1 corresponds to 50 µg ml^{-1} of double-stranded DNA or 40 µg ml^{-1} of RNA.

Radiometric Assays

In general, DNA synthetic reactions (such as reverse transcription, nick-translation or tailing with terminal transferase) were monitored by percentage incorporation of either ^{3}H- or ^{32}P-labelled nucleotide into acid insoluble material. An aliquot (0.01-0.1 reaction volumes) was dispersed in 0.1-1.0 ml TSE. An aliquot of this was dried onto a 2.0 mm Whatman GFC filter. A further aliquot was added to 20 µg of denatured fragmented salmon sperm DNA (see Table 2.2), 5 ml of 1 M HCl containing 1% (w/v) NaH_2PO_4, 1% (w/v) $Na_4P_2O_7$, was added and the mixture left at 4°C for 2-5 min.. The precipitated mucleic acids were collected by filtration onto 2.0 mm GFC filters, washed extensively in acid solution and ethanol and dried. All filters were counted in toluene-based scintillation fluid. The yield of DNA synthesised was calculated from the percentage incorporation and from the known concentration of nucleotides. Specific activities were calculated from the known specific activity of the labelled nucleotide, assuming a 25% representation of each nucleotide in cDNA.

Endonuclease S1 Assays

These were performed in S1 assay buffer, containing 30 mM sodium acetate, pH 4.55, 250 mM NaCl, 4.5 mM $ZnSO_4$, 10 μg ml^{-1} denatured, fragmented salmon sperm DNA (see Table 2.2). Radioactive samples were diluted into 1 ml of this buffer. Duplicate 0.4 ml aliquots were removed, one being incubated with an empirically determined amount of endonuclease S1, sufficient to produce a limit digest of ^{32}P-labelled cDNA. Both aliquots were then precipitated by addition of 20 μg of denatured, fragmented salmon sperm DNA (Table 2.2) and 5 ml of cold 1 M HCl containing 1% (w/v) NaH_2PO_4 and 1% (w/v) $Na_4P_2O_7$. After 2-5 min. on ice the nucleic acids were collected by filtration onto Whatman GFC filters, washed with acid solution and ethanol and dried. Percentage resistance to endonuclease S1 was then calculated.

^3H-Poly-U Assay (Bishop *et al.*, 1974a)

An aliquot of each RNA to be assayed (usually 0.1-1.0 μg) was diluted into 0.2 ml of 2x SSC. A saturating amount (0.1 μg) of ^3H-poly-U (sp. act. = 2 x 10^5 cpm g^{-1}, a gift from M. Lovett) was added followed by incubation at 42°C for 30 min.. The samples were then chilled in ice/water, 50 μl of 0.5 mg ml^{-1} ribonuclease A (Worthington) were added and the reactions were left in ice/water for 30 min.. 20 μg of denatured, sonicated salmon sperm DNA were then added followed by 5 ml of cold 1 M HCl containing 1% (w/v) NaH_2PO_4 and 1% (w/v) $Na_4P_2O_7$. After 2-5 min. on ice the nucleic acids were collected by filtration onto 2.0 mm Whatman GFC filters, washed in acid solution and ethanol, dried and counted in toluene based scintillation fluid. RNA concentrations

were calculated assuming a 1:1 correspondence between poly A and poly U and a 7.5% representation of poly A in mRNA (Bishop *et al.*, 1974b).

Dideoxynucleotide Tailing Assay (Section 3.4)

The reaction conditions were the Tris-cacodylate buffer described in Section 2.6. A standard mixture of 100 μM ^{3}H-dTTP and 2.5 μM ddTTP was diluted 10-fold into the final assay mix of 5-10 μl. Reactions used serially diluted samples of DNA to be titrated to ensure nucleotide excess. Incubation was for 15 min. at 37^{o}C with 1000 units ml^{-1} of terminal deoxynucleotidyl transferase (Ratliff Biochemicals, Los Alamos, New Mexico). Concentrations of termini were calculated by measuring the incorporation of label into acid insoluble material (as described above), the nucleotide mixture being previously calibrated using a DNA standard of known molar concentration (see Section 3.4).

CHAPTER III

SYNTHESIS AND CLONING OF cDNA

3.1 Introduction.

Several DNA polymerases, including E.coli DNA polymerase I, are able to transcribe a DNA copy of an RNA template under appropriate conditions, using a suitable primer. The most efficient and widely used are the RNA-directed DNA polymerases or reverse transcriptases of RNA tumour viruses, such as avian myeloblastosis virus (AMV).

The first non-viral template mRNA to be successfully copied into complementary DNA (cDNA) was rabbit globin mRNA (Kacian et al. 1972; Ross et al., 1972; Verma et al., 1972). A short time later cDNA transcripts of several mRNA species were being employed in various studies, particularly as hybridization probes where the increased specific activities that they provided afforded a significant improvement in sensitivity. One very important application of cDNA probes was in the technique of cDNA/mRNA hybridization, which allowed the constitution of complex populations of mRNA to be examined (Bishop et al., 1974b).

The construction and subsequent cloning in E.coli of a double-stranded DNA copy of the rabbit β-globin messenger RNA was an important landmark in molecular biology (Efstratiadis et al., 1976; Maniatis et al., 1976). The basic elements of the approach used have been successfully employed to clone cDNA copies of a number of mRNAs, although methodological details have varied (Efstratiadis and Villa-Komaroff, 1978). A generalised scheme for the cloning of double-stranded cDNA is presented in Fig. 3.1. The procedure may be divided into three basic stages:

(a) Synthesis of a cDNA copy of the template RNA and subsequent removal of the RNA.

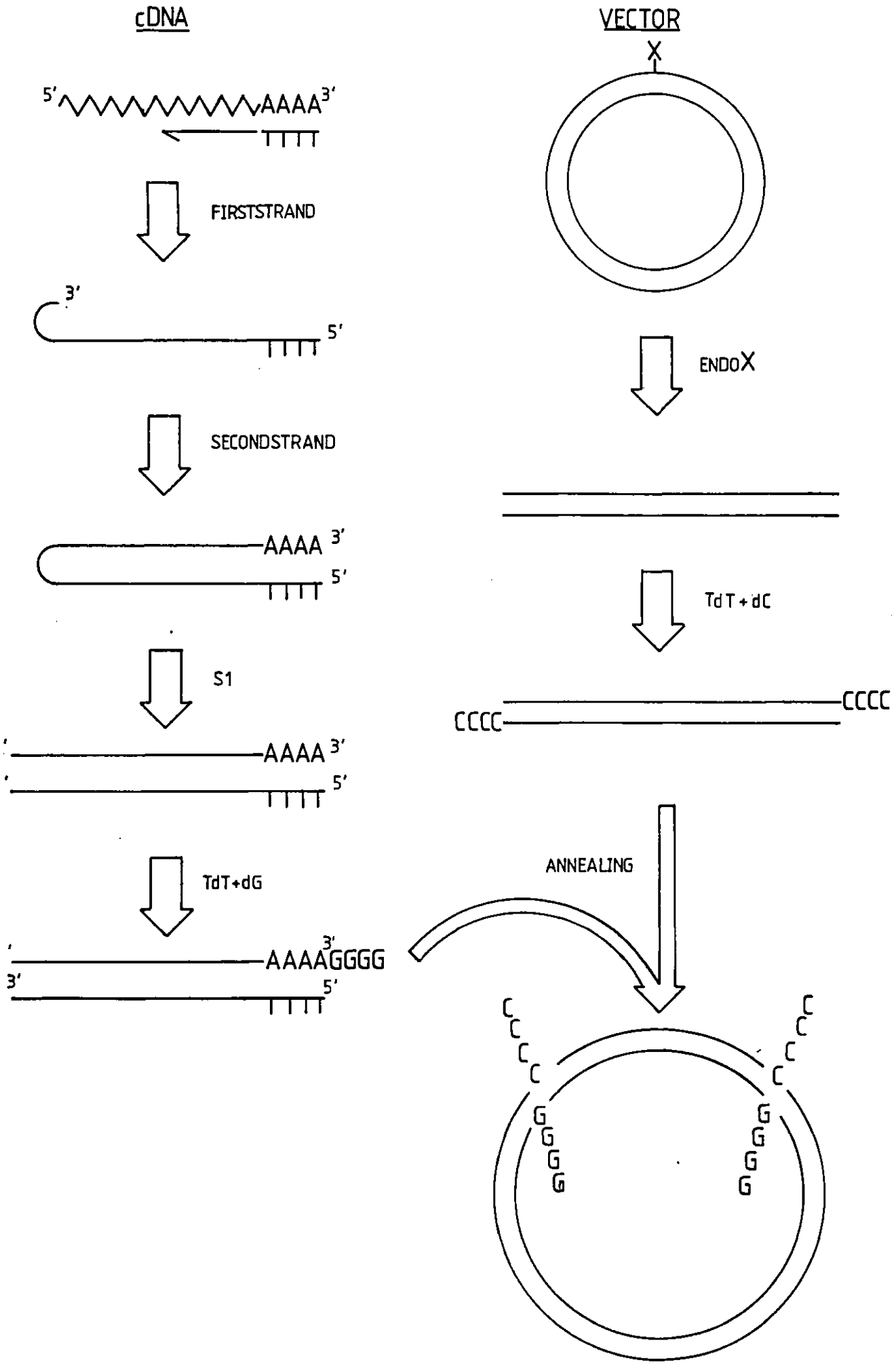
(b) Conversion to a double-stranded, blunt-ended DNA molecule.

(c) Insertion into a suitable vector molecule (usually a plasmid) and recovery of hybrid molecules in E.coli.

FIGURE 3.1

Figure 3.1 A Generalized Scheme for the Synthesis and Cloning of Double-Stranded cDNA.

The vector molecule is cleaved with one (or more) site-specific endonuclease(s) (endoX) and a homopolymer tail is added using terminal deoxynucleotide transferase (TdT). Single stranded cDNA is synthesized from a polyA$^+$ RNA template, purified free of template RNA and converted to a double-stranded hairpin structure. Following cleavage of the hairpin loop with endonuclease S1 a complementary homopolymer tail is added. The two molecules are then annealed and the hybrid molecules recovered by transfection into E. coli.


cDNA
5'
AAAA
3'
FIRSTSTRAND
3'
5'
SECONDSTRAND
AAAA
3'
5'
S1
AAAA
3'
5'
TdT+dG
3'
AAAAGGGG
3'
5'
VECTOR
X
ENDOX
TdT+dC
CCCC
CCCC
ANNEALING
C
G

The two most important parameters to optimize in any cDNA cloning exercise are efficiency and fidelity of copying. Thus a compromise must be reached between the yield of product at each step of the procedure and the quality of the product in terms of its length and, if a heterogeneous template is used, the degree of representation of the total sequences present in the starting mRNA population. In this Chapter the three stages of first-strand synthesis, conversion to a double-stranded form and recovery in E.coli will be considered in turn and the construction of two "libraries" of cDNA clones derived from mouse tissue culture cell total mRNA will be described.

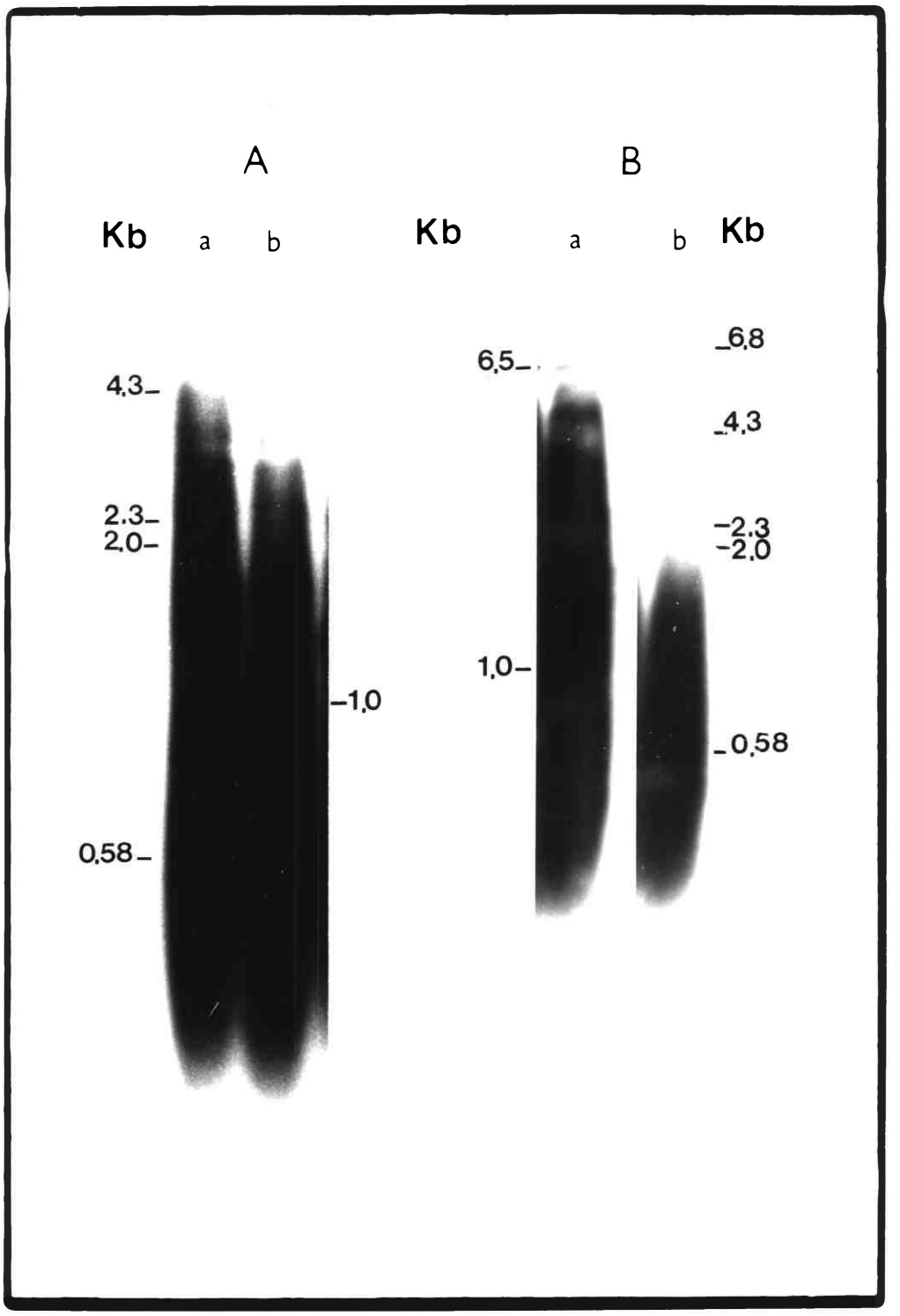
3.2 The first strand reaction.

Most protein-coding mRNA sequences are poly-adenylated at their 3' end. This allows synthesis of cDNA using an oligo-dT primer. Various authors have described conditions for optimizing both length and yield of first strand product (Efstratiadis and Villa-Komaroff, 1978). The most striking conclusion from these studies is that different templates work optimally under different conditions. In Fig. 3.2A the effect of two sets of reaction conditions on the length of product obtained with AMV reverse transcriptase is examined using Xenopus laevis liver poly A^+ RNA as template. Both sets of reaction conditions gave optimal results on ovalbumin mRNA (Monahan et al., 1976; Wickens et al., 1978) yet gave markedly different results using the Xenopus template. Rat liver ribonuclease inhibitor was found to have a pronounced effect under both reaction conditions and was included in all subsequent first-strand reactions to inhibit the low level of ribonuclease that presumably contaminates the enzyme preparation. The best results with this template were therefore obtained by including ribonuclease inhibitor in the reaction conditions of Fig. 3.2A track a. Transcription under these

FIGURE 3.2

Figure 3.2 Optimization of cDNA First-Strand Reaction Conditions

Analytical cDNA synthesis reactions (10µl) were performed. The conditions for Fig. 3.2A track a were essentially those of Monahan et al. (1976), containing 50mM Tris HCl pH8.3, 10mM $MgCl_2$, 20mM DTT, 400µM dATP, dGTP and dTTP, 50µM dCTP, 36µg ml^{-1} actinomycin D, 5µg ml^{-1} oligo dT_{12-18} and 50µg ml^{-1} of Xenopus laevis $polyA^+$ RNA (a gift of Dr. D.F. Smith, Imperial College). AMV reverse transcriptase was added to 250 units ml^{-1} and the reaction incubated at 46°C for 5 min. The conditions for Fig. 3.2A track b were essentially those of Wickens et al. (1978), containing 50mM Tris HCl pH8.3, 10mM $MgCl_2$, 30mM 2-mercaptoethanol, 140mM KCl, 500µM dATP, dGTP and dTTP, 50µM dCTP, 5µg ml^{-1} of oligo dT_{12-18} and 50µg ml^{-1} of Xenopus RNA. Reverse transcriptase was added to 250 units ml^{-1} and incubation was at 42°C for 60 min. The reaction conditions for Fig. 3.2B were as described for Fig. 3.2A track a except that rat-liver ribonuclease inhibitor was added to 25 units ml^{-1} in both cases and SV3T3 C138 $polyA^+$ RNA was used in track b. ^{32}P-labelled dCTP was included in all reactions to give a final cDNA sp. act. of 10^6 cpm/ug^{-1} and electrophoresis was performed in 2.0% agarose gels in alkaline buffer as described in Chapter 2. The equivalent conditions described in Chapter 2 were found to give identical results with both templates and were used in subsequent experiments.


A
Kb
a
b
4,3
2,3
2,0
1,0
0,58
B
Kb
a
b
Kb
6,5
6,8
4,3
2,3
2,0
1,0
0,58

conditions yielded a cDNA product which appeared to contain a high proportion of full-length strands (Fig. 3.2B). This is made particularly evident by the presence of a discrete band of approximately 6.5 Kb which probably represents a full-length transcript of the Xenopus vitellogenin mRNA (Smith et al., 1979).

When these conditions were applied to mouse SV3T3 C138 poly A^+ RNA, however, much poorer results were obtained (Fig. 3.2B). To improve the length of the product, different monovalent cations were included in the reaction. The results of one such experiment are shown in Fig.3.3; inclusion of NH_4^+ to 0.1M markedly improved the length of product obtained. The effect of varying the concentration is shown in Fig. 3.4, which indicates that a concentration of 0.2M is optimal for both yield and length. Reverse transcription under these conditions consistently yielded an acceptable product.

3.3 Conversion to double-stranded form.

Following removal of RNA by alkaline hydrolysis, and purification of the single-stranded cDNA by Sephadex G200M chromatography and ethanol precipitation as described in Chapter 2, the cDNA must be converted into a form suitable for cloning. Leis and Hurwitz (1972) demonstrated that cDNA freed from its template RNA could support second strand synthesis when new polymerase was added and postulated the existence of a partial 3' terminal hairpin structure on the cDNA. This property allows a complete double-stranded hairpin structure to be created by treatment of single-stranded cDNA with either E.coli DNA polymerase I, T4 DNA polymerase or reverse transcriptase (Efstratiadis and Villa- Komaroff, 1978). The second-strand reaction may be monitored in two ways:

(a) By making a radioactively-labelled first strand and assaying for double-strandedness using the single strand specific endonuclease S1.

FIGURE 3.3

Figure 3.3 The Effect of Varying Monovalent Cation Upon the cDNA First-Strand Reaction.

Analytical cDNA synthesis reactions (10μl) were performed using the conditions described in Chapter 2. (-) and the same supplemented with 0.1M ammonium acetate (NH_4^+), 0.1M KCl (K^+) and 0.1M NaCl (Na^+). Electrophoresis was as described in the legend to Fig. 3.2.

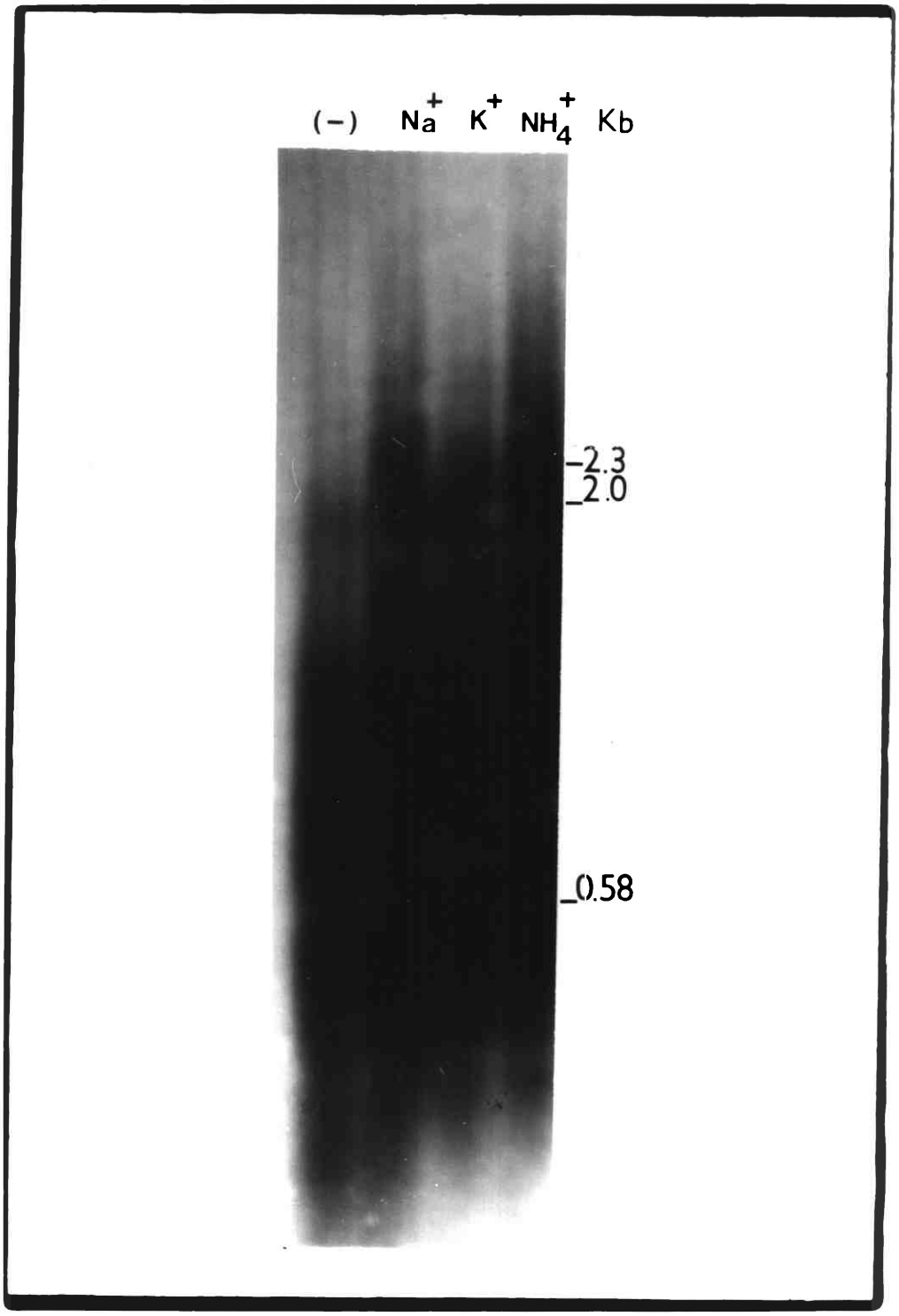
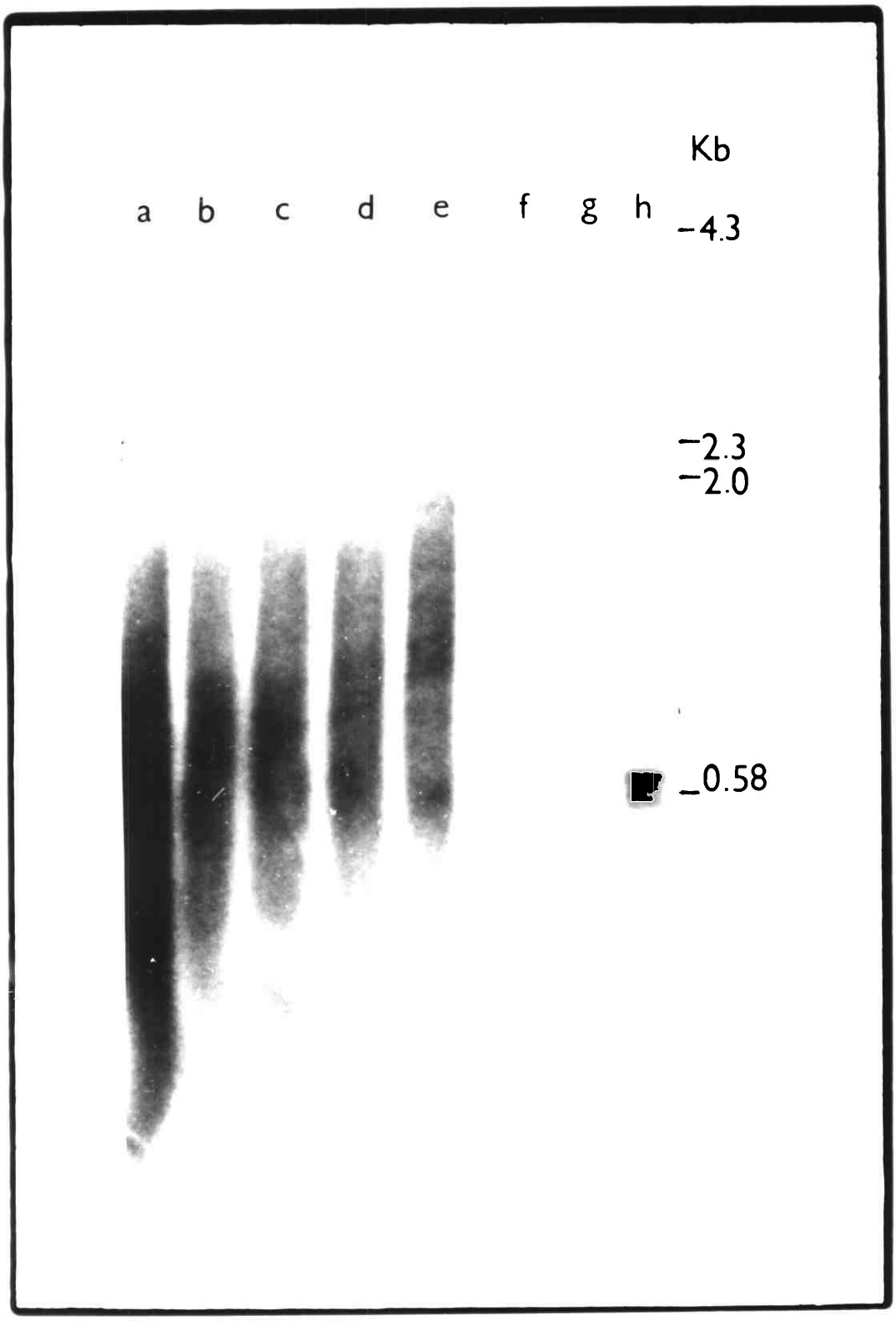

(−)
Na^+
K^+
NH_4^+
Kb
2.3
2.0
0.58

FIGURE 3.4

Figure 3.4 The Effect of Varying Concentrations of Ammonium Ion on the cDNA First-Strand Reaction

Analytical cDNA synthesis reactions were performed using the conditions described in Chapter 2. The inclusions are (a) 0 M, (b) 0.02 M, (c) 0.05 M, (d) 0.1 M, (e) 0.2 M, (f) 0.5 M, (g) 1.0 M ammonium acetate. Track (h) denotes single-stranded mouse globin cDNA made under the conditions of track (d).


Kb
a b c d e f g h
4.3
2.3
2.0
0.58

(b) By agarose gel electrophoresis. Since the product of the reaction is a double-stranded hairpin (Fig. 3.1) the reaction may be monitored by electrophoresis in a denaturing gel, conveniently an alkaline agarose system. The double-stranded product migrates as a molecule of twice the length of the first-strand product under these conditions. This is readily monitored if the template is a single species such as globin RNA. The second-strand reaction of a complex template can be monitored by two-dimensional electrophoresis. This is demonstrated in Fig. 3.5. Single-stranded SV3T3 C138 cDNA, labelled with ^{32}P, was incubated with reverse transcriptase as described in Chapter 2 and the reaction monitored using the endonuclease S1 assay. After incubation for 4 hours approximately 65% of the labelled first strand was protected from S1 cleavage. An aliquot was electrophoresed on a two-dimensional gel which was then dried down and autoradiographed. Two separate diagonals are observed on the gel, resolving the double-stranded hairpin form from the single-stranded unreacted species. The upper diagonal is the hairpin form, which migrated the more slowly in the denaturing dimension. The lower diagonal of single-stranded material is not colinear with the upper diagonal along the native dimension since single-stranded DNA migrates faster than double-stranded DNA of the same length. The relative intensities of the two diagonals fit acceptably with the result of the S1 assay.

The conversion to double-strand form is more readily assayed using a pure template. In Fig. 3.6, mouse globin RNA was used as a template for the synthesis of labelled cDNA. Denaturing gel electrophoresis resolved the cDNA into two bands, representing reverse transcripts of the α- and β-globin mRNAs. After second strand synthesis these bands were largely replaced by two which represent the

FIGURE 3.5

Figure 3.5 The Second Strand Reaction of a Complex Template.

A 50µl cDNA first strand synthesis reaction was performed using 0.5µg of SV3T3 C138 polyA$^+$ RNA. The purified cDNA was adjusted to double-stranded cDNA synthesis conditions (25µl reaction) and incubated with AMV reverse transcriptase 250 units ml^{-1} for 1h at 42^{o}C. An aliquot (2µl) was removed and electrophoresed through a 2.0% two dimensional agarose gel which was dried down and autoradiographed. All conditions were as described in Chapter 2. Another aliquot (2µl) was subjected to S1 assay (see Chapter 2) after second-strand synthesis to determine the degree of nuclease resistance (see Section 3.3).

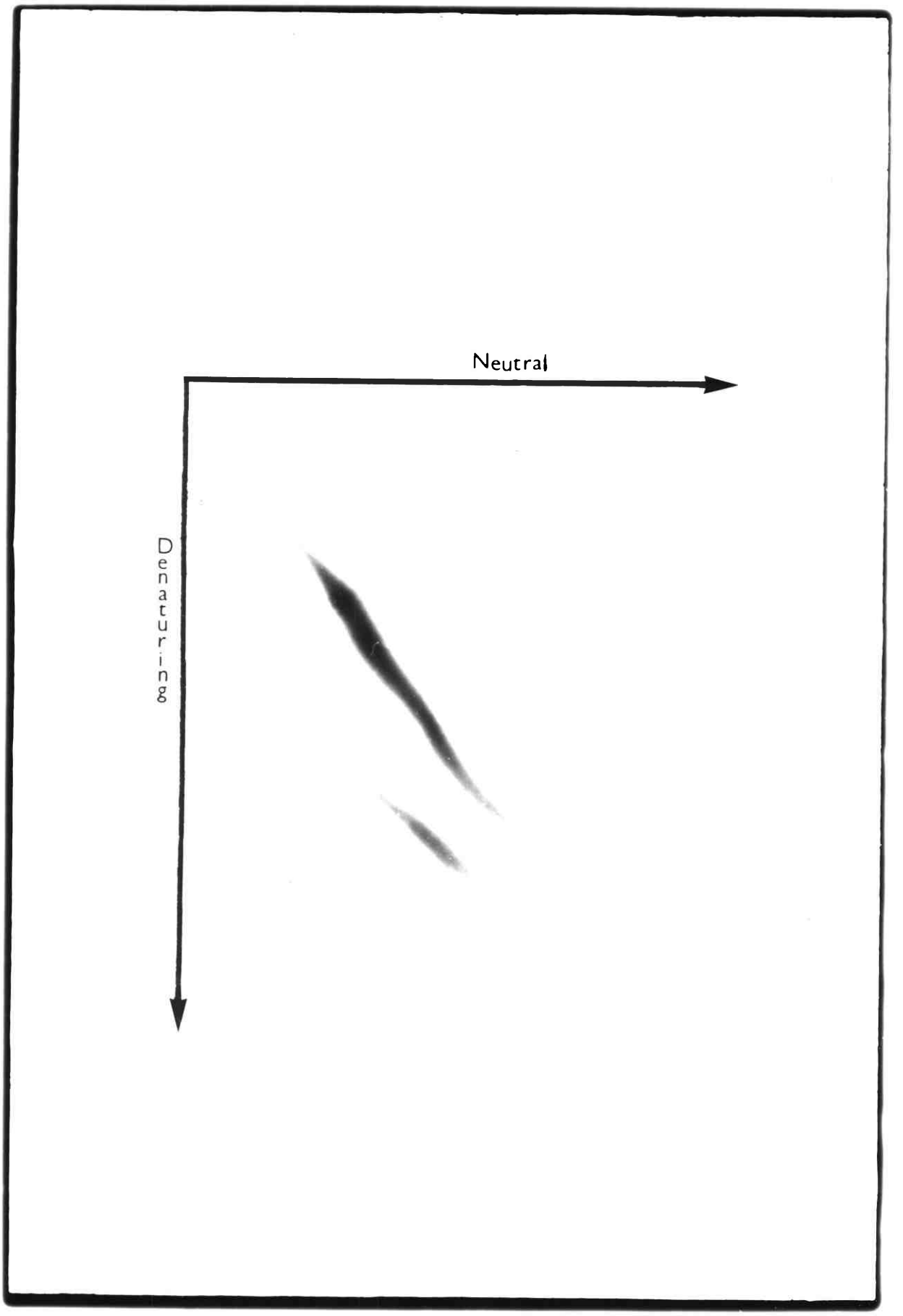
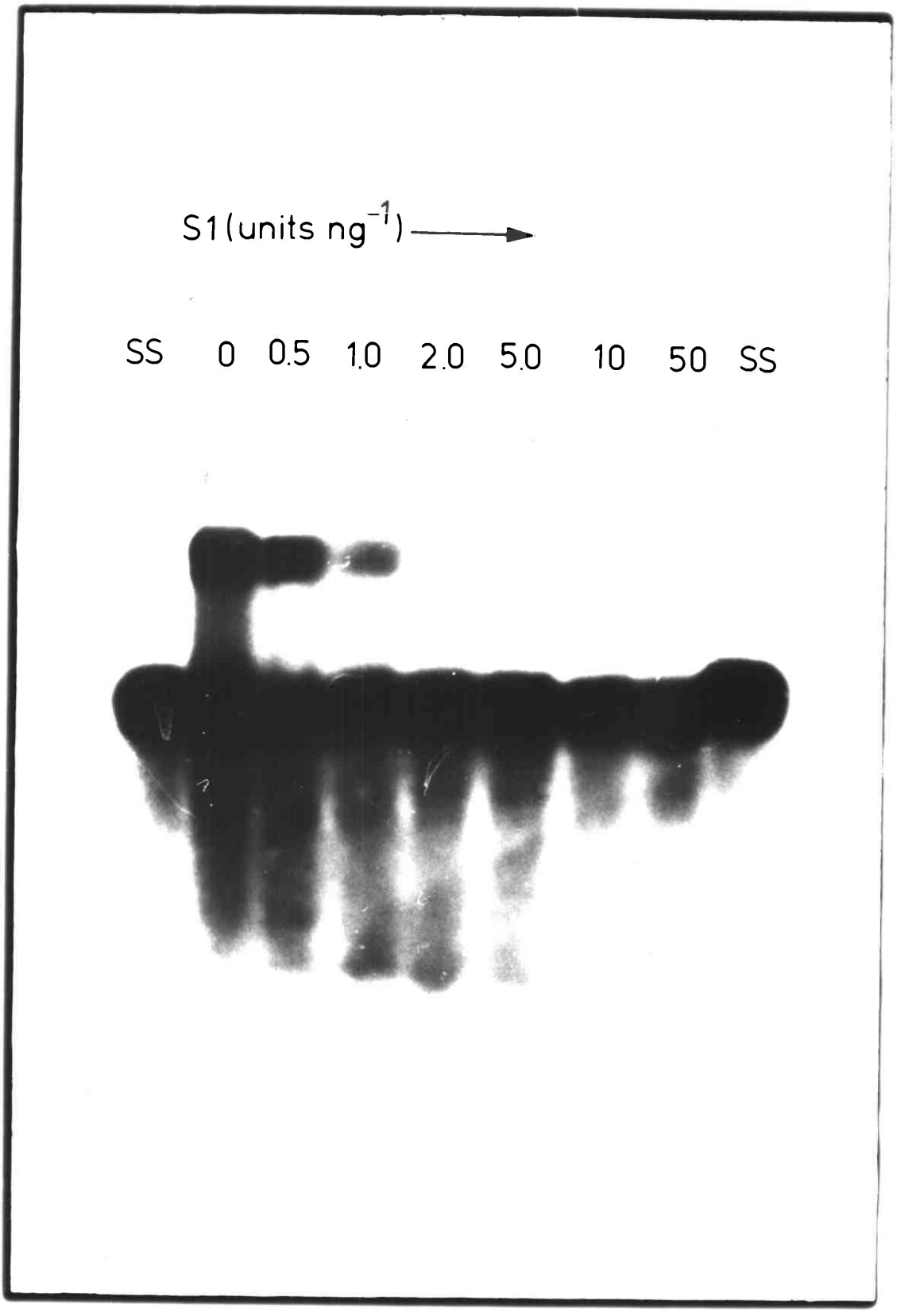

Neutral
Denaturing

FIGURE 3.6

Figure 3.6 Second-Strand Synthesis and Endonuclease S1 Assay using Mouse Globin RNA template

A 20 μl reaction containing 2 μg of mouse globin mRNA was converted to double-strand form as described in Section 2.6. Aliquots (1/10) were subjected to S1 treatment using varying concentrations of endonuclease S1 as denoted in the Figure. The control tracks, marked (SS) are single-stranded cDNA removed prior to the double-stranded reaction.


S1 (units ng^{-1}) ⟶
SS 0 0.5 1.0 2.0 5.0 10 50 SS

hairpin form. Under appropriate conditions, the hairpin loop may be cleaved using endonuclease S1. This reaction is also shown on Fig.3.6; as increasing concentrations of S1 were added to the double-stranded product, conversion to a form that co-migrated with first strand was observed, consistent with cleavage of the hairpin loop. In Fig 3.7, double-stranded SV3T3 C138 cDNA was incubated with 0.5 units ng^{-1} of S1, and an aliquot run on a two-dimensional gel. Digestion in this case was incomplete, as expected, allowing the conversion of hairpin to blunt-ended form to be observed. The upper diagonal is undigested hairpin DNA and the lower represents the double-stranded, blunt ended species.

Following S1 treatment, the cDNA is in a form suitable for cloning. The double-stranded cDNA should first be purified free of unincorporated deoxynucleoside triphosphates and short fragments of DNA. This may be accomplished by Biogel A15OM chromatography followed by ethanol precipitation. This procedure also fractionates on the basis of size, allowing cDNA fractions of different lengths to be cloned in pools. An example of such a purification is presented in Fig. 3.8. Here double-stranded SV3T3 C138 cDNA was cleaved with S1 and passed over Biogel as described in Chapter 2. Aliquots of each fraction were electrophoresed through neutral agarose and the gel dried and autoradiographed. The remainder of each fraction was counted by Cerenkov radiation and the column profile plotted. In all cloning experiments the leading two-thirds of the peak were used for cloning, and the remainder was discarded as being predominantly of low molecular weight and contaminated with short fragments.

3.4 Cloning of cDNA.

Double-stranded cDNA is most conveniently cloned using the homopolymer tailing method. Roychoudry et al. (1976) showed that

FIGURE 3.7

Figure 3.7 The Conversion of Hairpin cDNA to Blunt-Ended DNA.

The remainder of the preparation of double-stranded SV3T3 C138 cDNA described in the legend to Fig. 3.5 (0.15μg) was adjusted to endonuclease S1 reaction conditions and incubated with 75 units (0.5 units ng^{-1}) of endonuclease S1. An aliquot (4μl) was electrophoresed through a two dimensional gel. All conditions were as described in Chapter 2.

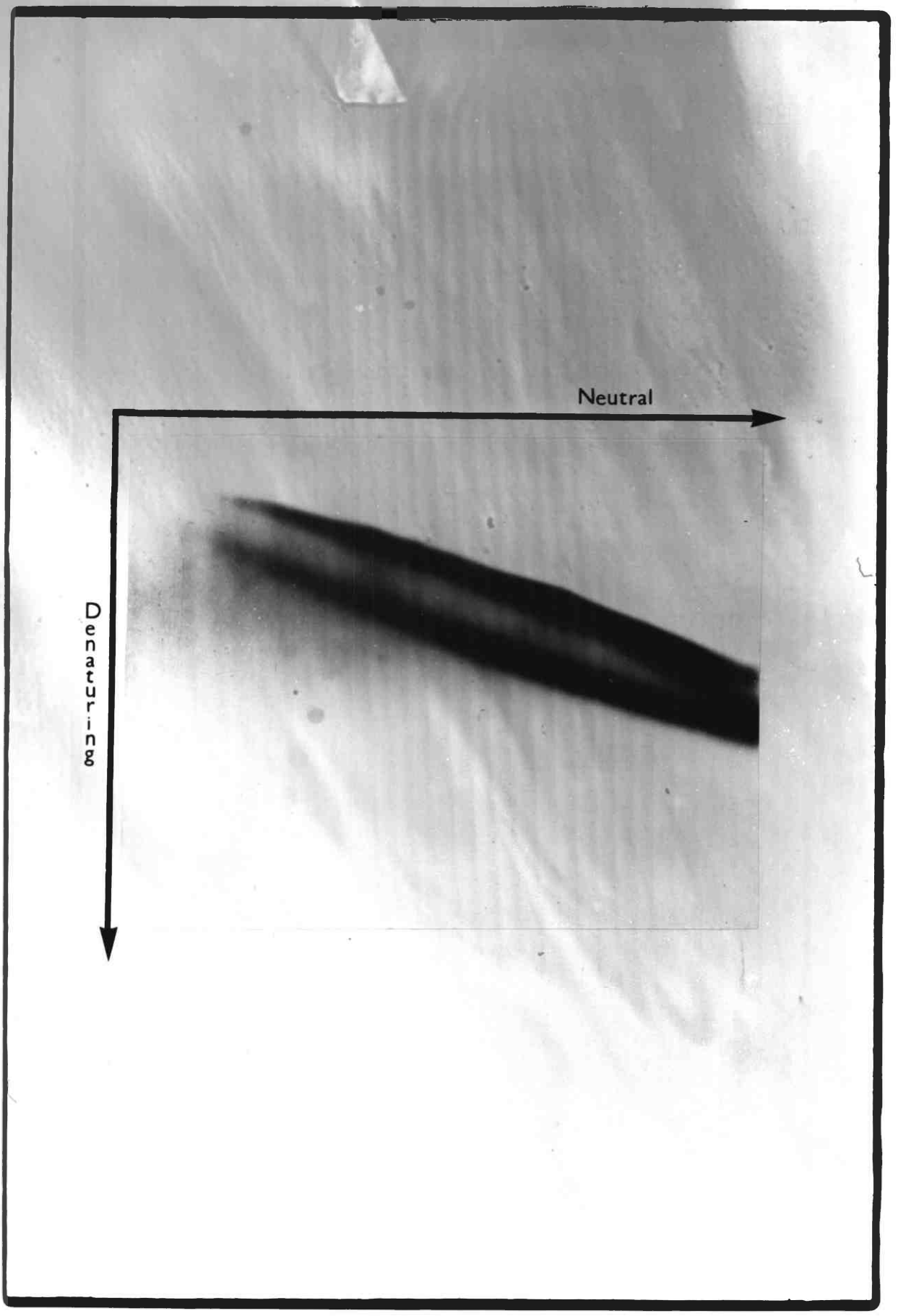
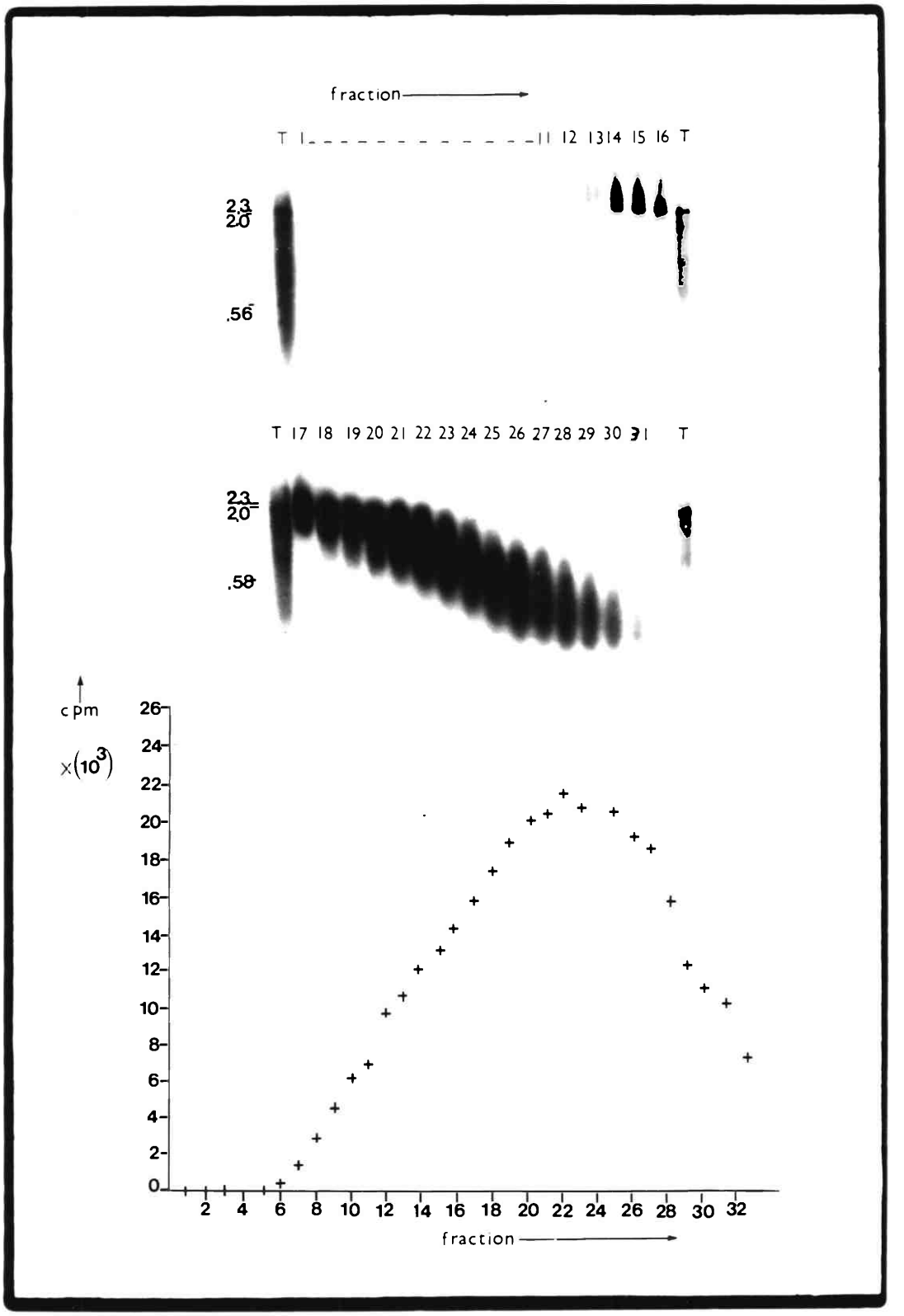

Neutral
Denaturing

FIGURE 3.8

Figure 3.8 Size Fractionation of Double-Stranded cDNA prior to Cloning

A preparation of double-stranded SV3T3 C138 cDNA was fractionated over a Biogel A.150M column in TSE as described. Aliquots were counted by Cerenkov radiation and further aliquots were electrophoresed on a 2% alkaline agarose gel (see Chapter 2).


fraction
T 1 11 12 13 14 15 16 T
2.3
2.0
.56
T 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 T
2.3
2.0
.58
cpm
x(10^3)
26
24
22
20
18
16
14
12
10
8
6
4
2
0
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32
fraction

terminal deoxynucleotidyl transferase could add a homopolymer tail to blunt-ended DNA if Co^{2+} was used instead of Mg^{2+} as cofactor. Thus cDNA may be cloned without further treatment such as repair of the ends.

Three factors must be considered when optimising tailing and annealing conditions:

(a) Tail composition.

(b) Tail length.

(c) End ratio during annealing.

Each of these will be considered in turn here.

Effect of varying the tail composition.

In Table 3.1 are presented the results of an experiment designed to test the effect of tail composition. The plasmid vector pAT153 (see Fig. 3.14) was cleaved with EcoRI and tailed with approximately 20-40 nucleotides using each of the four deoxynucleoside triphosphates. TaqI fragments of bacteriophage λ DNA (approximate mean size = 0.5 Kb) were similarly tailed with all four triphosphates. The rate of mononucleotide addition to each DNA, using each triphosphate, was determined by assaying the incorporation of ^{3}H-labelled triphosphate with time; a typical profile is shown in Fig. 3.9. From the known concentration of DNA termini, the number of residues added per end may be determined. The DNA is then tailed for the relevant length of time to allow the desired number of residues to be added. Tailing the vector with dCTP and the insert with dGTP yielded the highest frequencies following annealing and transfection into E.coli HB101, as described in Chapter 2.

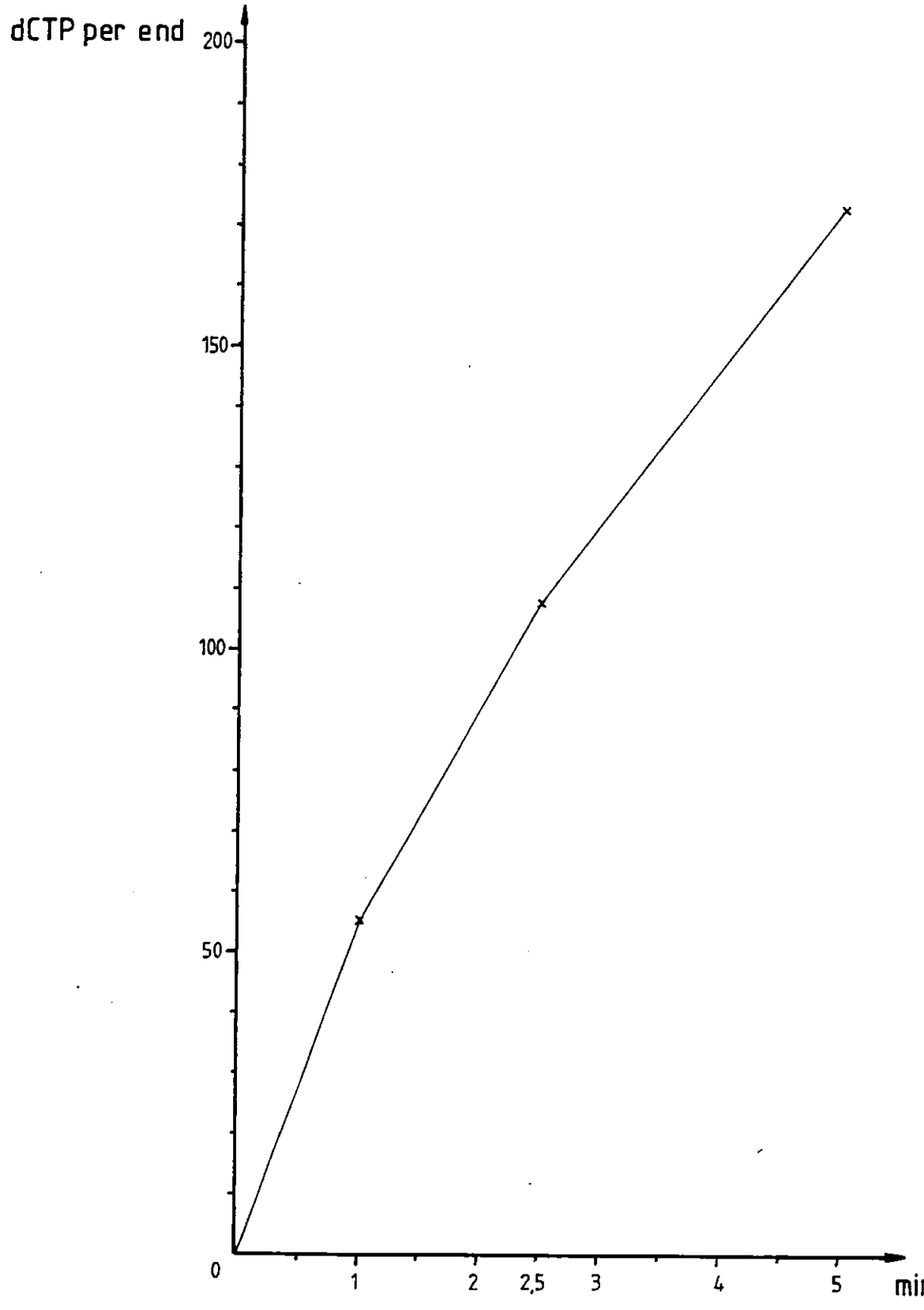
Effect of varying the tail length.

Table 3.2 shows the effect of different tail lengths using the optimum end composition described above. Best results were obtained

FIGURE 3.9

Figure 3.9 Kinetics of a Typical Tailing Reaction

A 100 µl tailing reaction was set up as described in Section 2.6, using dCTP as substrate. 5 µl aliquots were removed at intervals and assayed as described in Section 2.12.



ADDITION OF dCTP TO RI-pAT 153 DNA

TABLE 3.1

The effect of varying tail composition upon the efficiency of cloning of λ DNA fragments

average number of mononucleotides added		number of colonies obtained per μg of DNA	
to vector	to insert	cfu μg^{-1} vector	cfu μg^{-1} insert
42 dCMP	26 dGMP	7.8×10^4	5.7×10^5
31 dGMP	21 dCMP	1.1×10^4	8.0×10^4
33 dTMP	29 dAMP	4.4×10^3	3.2×10^4
35 dAMP	24 dTMP	3.8×10^3	2.8×10^4
42 dCMP	-	7.2×10^2	-
31 dGMP	-	9.0×10^2	-
37 dTMP	-	1.1×10^3	-
35 dAMP	-	8.5×10^2	-

Annealings (10 μl) were performed containing EcoRI cleaved pAT153 DNA (0.5 μg ml^{-1}) and Taq I cleaved λ DNA (0.06 μg ml^{-1}). Assuming an average size of 3.65 kb for the vector and 0.5 kb for the insert λ fragments, this corresponds to an approximately equimolar ratio of DNA termini. The number of residues of mononucleotide added per end is shown together with the number of colonies (cfu) obtained per μg of either vector or insert DNA following transfection into E. coli HB101 as described in Chapter 2.. The results obtained with each preparation of tailed vector DNA in the absence of insert DNA is also shown. The transfection frequency of uncleaved vector in this case was 6.0×10^6 cfu μg^{-1}.

TABLE 3.2

The effect of varying tail length upon the efficiency of cloning of λ DNA fragments

average number of mono-nucleotides added		number of colonies obtained per μg of DNA	
to vector	to insert	cfu μg^{-1} vector	cfu μg^{-1} insert
42 dCMP	26 dGMP	5.5×10^4	4.0×10^5
42 dCMP	18 dGMP	4.0×10^4	2.9×10^5
42 dCMP	10 dGMP	2.5×10^4	1.8×10^5
20 dCMP	26 dGMP	7.2×10^4	5.2×10^4
20 dCMP	18 dGMP	7.1×10^4	5.2×10^5
20 dCMP	10 dGMP	9.0×10^4	6.6×10^5
9 dCMP	26 dGMP	4.0×10^4	2.9×10^5
9 dCMP	18 dGMP	4.0×10^4	2.9×10^5
9 dCMP	10 dGMP	3.5×10^4	2.5×10^5

Annealings (10 μl) were performed containing EcoRI cleaved pAT153 DNA (0.5 μg ml^{-1}) and Taq I cleaved λ DNA (0.06 μg ml^{-1}). This corresponds to an equimolar ratio of DNA termini (see Table 3.1). The number of residues of mononucleotide added per end is shown together with the number of colonies (cfu) obtained per μg of vector and per μg of insert DNA, following incubation and transfection into E. coli HB101 as described in Chapter 2. The transfection frequency of uncleaved vector in this case was 3.0×10^6 cfu μg^{-1}.

using a vector tail length of 20 and an insert tail length of 10.

Effect of the vector:insert ratio in the annealing reaction

The results of an experiment designed to test the effect of varying the ratio of vector ends to insert ends are presented in Fig. 3.10. Approximately 33 residues per end of dCMP were added to a preparation of EcoRI-cleaved pAT153 DNA. The insert λ HinfI-fragments were tailed to a length of 11 residues. The DNAs were annealed at varying end ratios as described in the Figure legend, and transfected into E.coli HB101. The number of colonies obtained per μg of vector DNA and per ug of inserted DNA were determined for each ratio, the results being shown in Fig. 3.10; clearly the efficiency per μg of insert DNA (the more important parameter since cDNA is usually scarce) rose markedly as the ratio of vector to insert ends approached unity at which point it started to plateau. Little increase in the yield of transfectants was observed at higher vector:insert ratios. The result of a similar experiment, this time using double-stranded Balb/c 3T3 cDNA as inserted material, is shown in Fig. 3.11. In this experiment the overall efficiency of cloning was lower than in the previous experiment, introducing a complicating factor, that of vector background caused by low levels of contaminating uncleaved supercoiled DNA in the preparation of vector DNA. The uncorrected transfection frequency per μg of cDNA seemed to continue increasing beyond a ratio of 1; however, when the number of colonies obtained in the absence of insert DNA was subtracted for each sample, no increase in the net transfection frequency per μg of insert was observed, with the optimum being reached at a vector to insert ratio of between 1 and 2.

In the two library cloning experiments to be described at the end of this Chapter, the optimal parameters derived in this Section were strictly adhered to. Thus vector molecules were tailed with

FIGURE 3.10

Figure 3.10 The Effect of Varying Vector/Insert Ratios on the Efficiency of Cloning HinfI fragments.

Ratio Vector/Insert	cfu μg^{-1} Vector	cfu μg^{-1} Insert
0.19	8.1×10^4	1.5×10^5
0.32	6.8×10^4	1.7×10^5
0.48	6.6×10^4	3.0×10^5
0.96	3.6×10^4	3.2×10^5
1.1	4.1×10^4	4.2×10^5
2.4	2.2×10^4	5×10^5

Annealings (10µl) were performed containing a constant amount of dC-tailed EcoRI cut pAT153 (0.5µg ml^{-1}) and increasing amounts of dG-tailed λ HinfI fragments. Following incubation and transfection as described in the text (see also Chapter 2) the transfection frequencies obtained per µg of vector and per µg of insert were calculated. The molar ratio of vector/insert was calculated assuming a length of 3.65Kb for pAT153 and an average length determined by gel electrophoresis of 385 nucleotides for the λ HinfI fragments. All the transfection frequencies have been corrected for vector background which was never greater than 2.5% in this experiment.

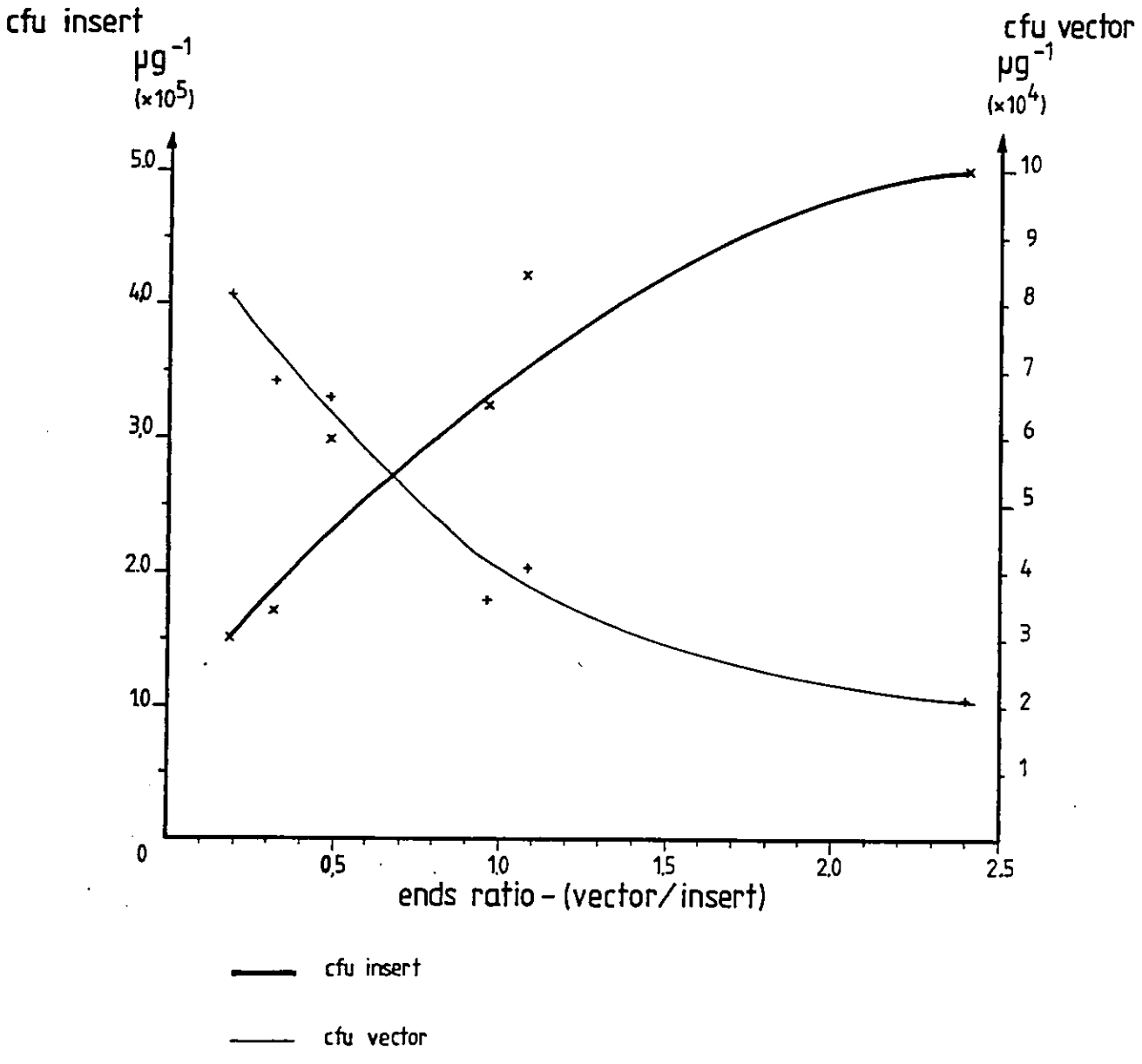
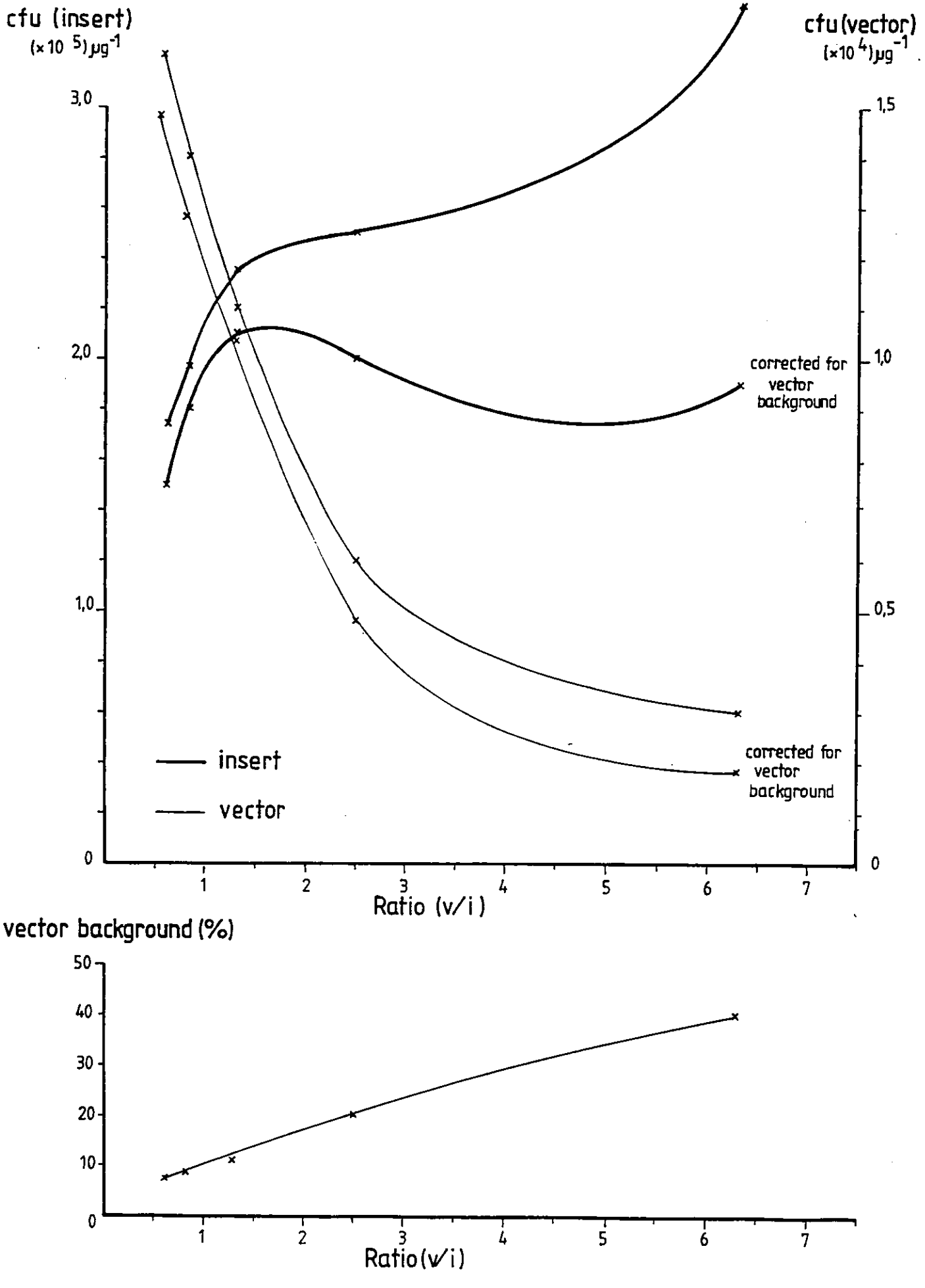

cfu insert
µg⁻¹
(×10⁵)
cfu vector
µg⁻¹
(×10⁴)
5.0
4.0
3.0
2.0
1.0
0
10
9
8
7
6
5
4
3
2
1
0,5
1.0
1.5
2.0
2.5
ends ratio – (vector/insert)
cfu insert
cfu vector

FIGURE 3.11

Figure 3.11 The Effect of Varying Vector/Insert Ratios on the Efficiency of Cloning cDNA.

Ratio	cfu μg^{-1} Vector		cfu μg^{-1} Insert		% Vector
Vector/Insert	Uncorrected	Corrected	Uncorrected	Corrected	Background
0.63	1.6×10^4	1.5×10^4	1.7×10^5	1.6×10^5	7.3
0.83	1.4×10^4	1.3×10^4	2.0×10^5	1.8×10^5	8.6
1.3	1.1×10^4	9.8×10^3	2.3×10^5	2.1×10^5	11
2.5	6.0×10^3	4.8×10^3	2.5×10^5	2.0×10^5	20
6.3	3.0×10^3	1.8×10^3	3.2×10^5	1.9×10^5	40

Annealings (10µl) were performed with a constant amount of dC-tailed EcoRI-cut pAT153 DNA (0.5µg ml^{-1}) and increasing amounts of dG-tailed double-stranded cDNA. Following incubation and transfection as described in the text (see also Chapter 2), the transfection frequencies obtained per µg of vector and per µg of insert were calculated before and after correction for vector background. The percentage vector background is also shown. The molar ratio of vector/insert was calculated assuming a length of 3.65Kb for pAT153 and an average length of 250 nucleotides for the cDNA, determined by gel electrophoresis.


cfu (insert)
(×10 5) µg-1
cfu (vector)
(×10 4) µg-1
3,0
2,0
1,0
0
1,5
1,0
0,5
0
corrected for vector background
corrected for vector background
insert
vector
1
2
3
4
5
6
7
Ratio (v/i)

vector background (%)
50
40
30
20
10
0
1
2
3
4
5
6
7
Ratio (v/i)

approximately 20 dCTP residues, insert cDNA was tailed with approximately 10 dGTP residues and the two were annealed with a two-fold excess of vector.

The use of dideoxynucleotides to control tail length.

One of the major problems encountered in a cDNA cloning exercise is that it is difficult to derive an accurate value for the concentration of cDNA termini, since that figure must be calculated using estimates of cDNA yield and size which are at best approximate. This inherent uncertainty results in poor reproducibility between results obtained with different preparations of cDNA. To overcome this, a new strategy was developed which not only allows the concentration of insert termini to be accurately determined, but also serves as a sensitive assay for contamination with short fragments which interfere in the cloning reactions.

The method exploits two observations:

(a) That terminal transferase will incorporate dideoxynucleotide triphosphates.

(b) That terminal transferase will incorporate 50% of available triphosphate when permitted.

The effect on tail length obtained, following tailing to saturation, of mixing increasing quantities of dideoxycytidine triphosphate (ddCTP) with a set amount of dCTP is presented in Table 3.3. EcoRI-digested pAT153 DNA was included in the reaction at a concentration of 14.3nM. ^{3}H-labelled dCTP was included at 10μM along with increasing concentrations of ddCTP. Terminal transferase was added as described in Chapter 2, and the samples incubated at 37°C for 10min. Reference to Fig. 3.9 shows that the rate of addition of dCTP to DNA is rapid, with in excess of 300 residues being added in this time. The result obtained in the absence of ddCTP (Table 3.3)

TABLE 3.3

The Effect of Varying dCTP/ddCTP Ratios Upon Tail Length Following Saturation Tailing

µM ddCTP	Ratio dCTP/ddCTP	No. Residues Added Per End
0	-	341
0.2	50	53
0.5	20	30
1.0	10	17
5.0	2	12
25.0	0.4	3
100.0	0.1	2

Several aliquots of a standard tailing reaction (see Chapter 2) containing 14.3nM (18µg ml^{-1}) of EcoRI cut pAT153 and 10µM ^{3}H-labelled dCTP were supplemented with increasing amounts of ddCTP. Following incubation with 1000 units ml^{-1} of terminal deoxynucleotidyl transferase for 10 min. at 37°C the number of residues added per end was calculated from the percentage incorporation of label into acid precipitable material from the known concentrations of nucleotide and DNA termini.

is in excellent agreement with this prediction. As ddCTP was added, however, the length of tail obtained reflected approximately the ratio of dCTP to ddCTP.

In Fig. 3.12 the kinetics of addition of a mixture of dCTP and ddCTP are presented. Three reactions were set up containing a constant amount of DNA termini (71.5nM), and a constant concentration of dCTP (10uM), plus (a) 0, (b) 0.4, or (c) 2μM ddCTP. Terminal transferase was added and aliquots were withdrawn at intervals. The number of residues added per end was calculated for each time point and the results plotted in Fig. 3.12. The initial rate of reaction was found to be the same irrespective of ddCTP concentration, but reactions performed in the presence of ddCTP plateaued when a number of ends had been added approximately equal to the ratio of dCTP to ddCTP in the reaction had been added. In the absence of ddCTP the reaction continued until about 50% of the available nucleotide had been incorporated, at which point the reaction also plateaued.

Using the dideoxynucleotide assay.

These observations have been used to develop a sensitive assay for the concentration of termini. In Fig. 3.13 are recorded the results of an experiment designed to calibrate the assay. Ten-fold dilutions of a standard reaction mix containing 100μM dTTP and 2.5μM ddTTP were incubated with varying amounts of DNA of known molar concentration (pCD5 cut by EcoRI = 2.9 Kb, 1μg = 1pmol). Terminal transferase was added and the samples incubated for 10min. Addition of dTTP to DNA is extremely rapid, this time of incubation being sufficient to incorporate 50% of the label in the absence of ddTTP. A linear relationship between concentration of termini and ^{3}H-labelled dTTP incorporated was observed. From the gradient of the line a value of 75 cpm per fmol of termini was calculated. The calibrated

FIGURE 3.12

Figure 3.12 Kinetics of Addition to DNA of a Mixture of dCTP and dTTP.

Three tailing reactions (25µl) were set up containing EcoRI-cut pCD5 (71.5nM of termini) supplemented with (a) 0, (b) 0.4 or (c) 2µM ddCTP. Terminal transferase was added to 1000 units ml^{-1} and the reactions incubated at 37°C. Samples (2.5µl) were withdrawn at the times shown and the number of residues added per DNA terminus calculated from the percentage incorporation into acid-insoluble material as described in Chapter 2.

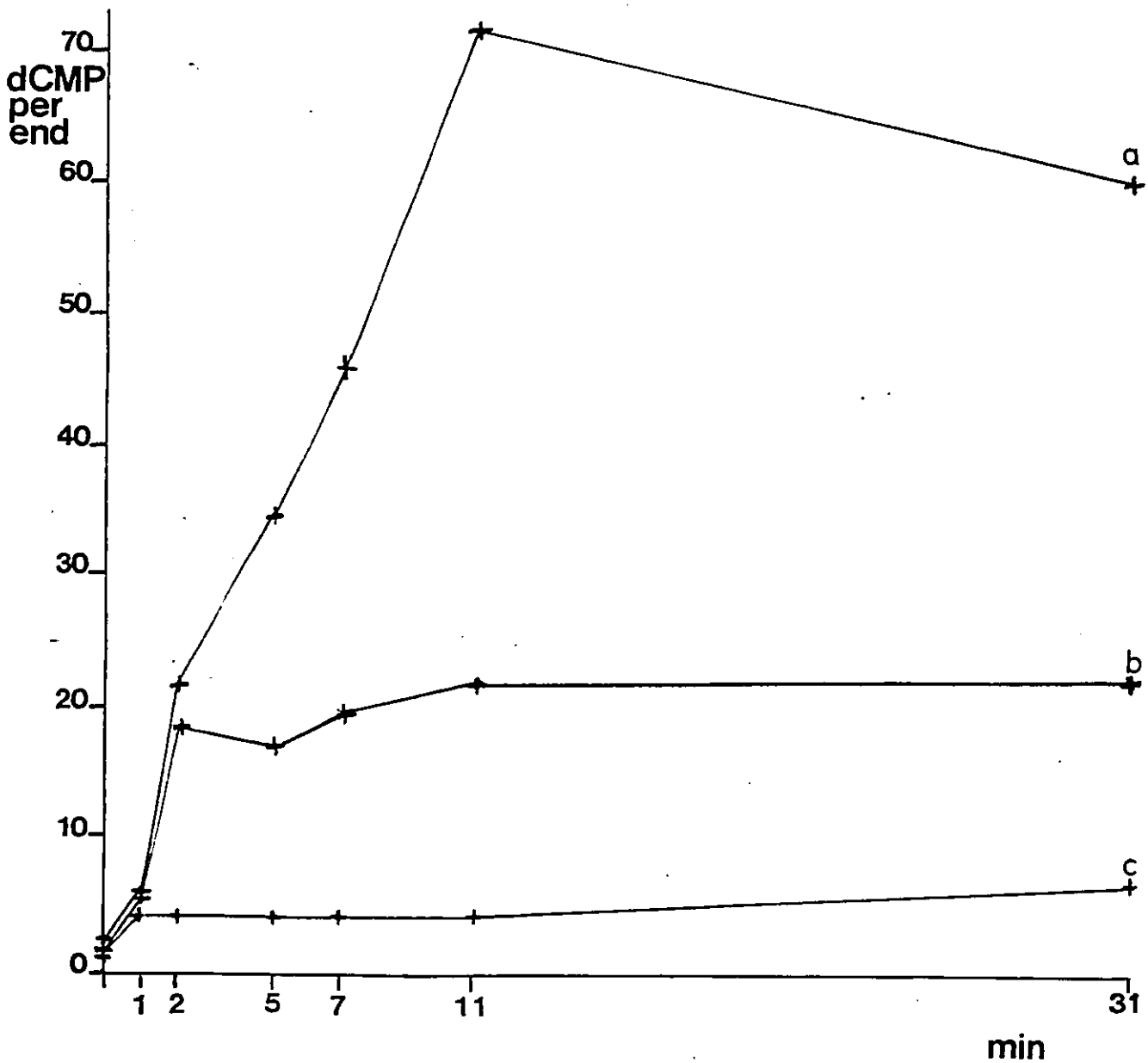
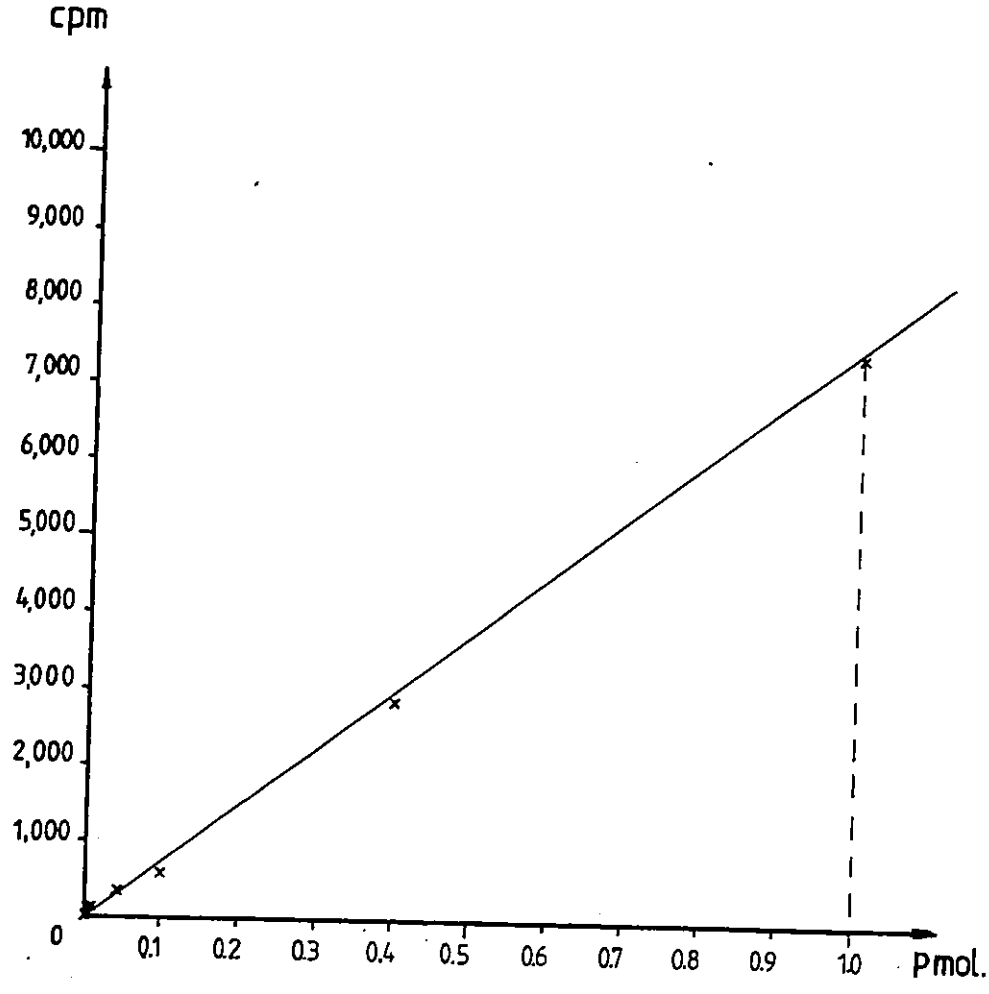

70
dCMP
per
end
60
50
40
30
20
10
0
a
b
c
1
2
5
7
11
31
min

FIGURE 3.13

Figure 3.13 Calibration of the Dideoxynucleotide Assay.

1µl aliquots of a standard mixture of ^{3}H-labelled dTTP (100µM sp. act. = 6 Cimmol^{-1}) and 2.5µM ddTTP were incubated in 10µl tailing reactions with the indicated quantities of EcoRI cut pCD5. Tailing was for 10 min. at 37^{o}C in the presence of 1000 units ml^{-1} of terminal transferase. The number of cpm incorporated into acid-insoluble material was determined for each sample and plotted as a function of pmol of DNA termini.


cpm
10,000
9,000
8,000
7,000
6,000
5,000
4,000
3,000
2,000
1,000
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Pmol.

assay mix was then used to calculate terminal concentrations during the construction of the SV3T3 C138 cDNA library, to be described in the following section.

Fig. 3.14 depicts the relationship between the vectors used in this study, pAT153 (Twigg and Sherratt, 1980) and pCD5 (M.R.D. Scott and P.W.J. Rigby, manuscript in preparation) and the plasmids from which they are respectively derived, pBR322 (Bolivar et al., 1977) and pXf3 (D. Hanahan, personal communication). pCD5 offers two advantages over pAT153. Firstly, since a linear fragment suitable for cloning may be produced by double-digestion with HindIII and BamI, background caused by partial vector cleavage of the type observed in Fig. 3.11 is negligible in cloning experiments. Secondly, inserts cloned between these sites may be excised using TaqI or, alternatively, a combination of SalI and either ClaI or EcoRI; in both cases the insert-specific fragments contain essentially no sequences from the plasmid vector. The latter facility greatly aids the characterization of inserts cloned by tailing into this vector.

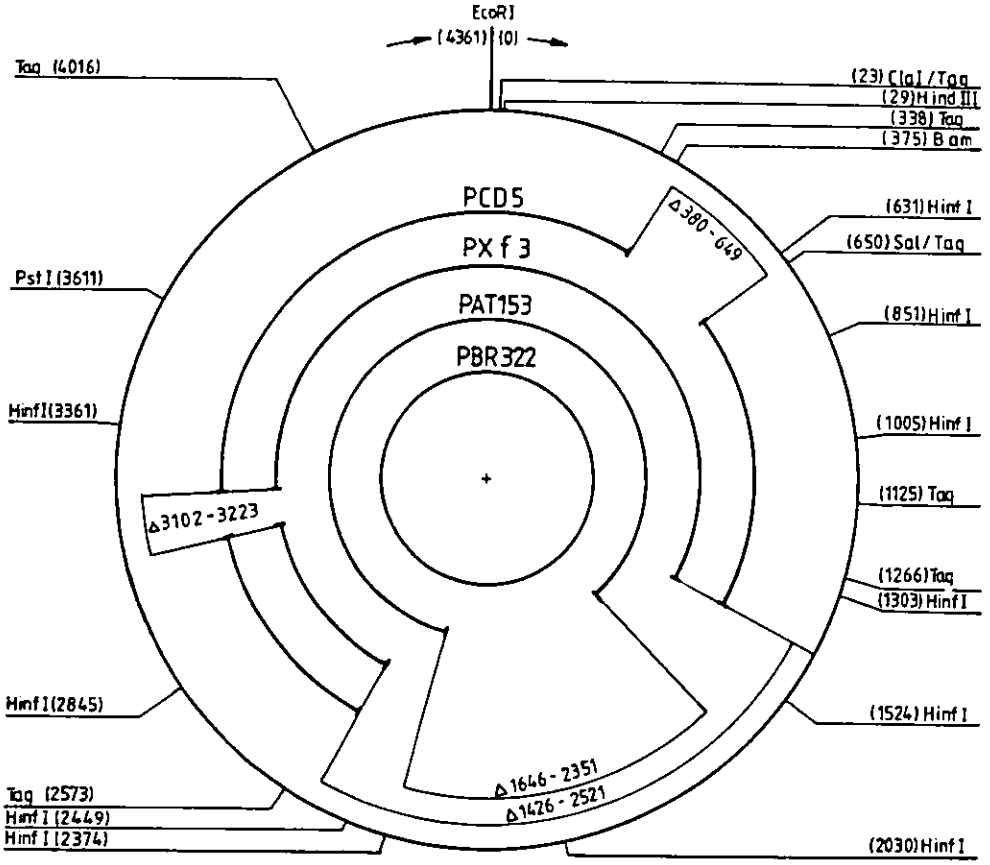
3.5 Construction of libraries of cDNA.

In this Section the construction of two libraries of cDNA will be described. Since the libraries were not produced simultaneously, some of the optimization procedures described in Section 3.4 were not applied during the construction of the Balb/c 3T3 library. For the same reason the SV3T3 C138 library was constructed in the improved cDNA cloning vector pCD5 (see Fig. 3.14 and above) rather than pAT153, and the E.coli host strain employed was changed from HB101 to DH1 because of the increased transfection efficiency offered by the latter (see Chapter 2).

FIGURE 3.14

Figure 3.14 Relationship Between the Plasmid Vectors Used in This Study and the Vectors From Which They Are Derived.

The outer ring represents the pBR322 genome with the positions of the cleavage sites for TaqI, HinfI, ClaI, SalI, EcoRI, HindIII and BamI identified by their nucleotide numbers (Sutcliffe, 1978). Note that in the studies described in this thesis the isoschizomer BstI was used instead of BamI. The inner rings represent the genomes of pCD5, pXf3, pAT153 and pBR322 respectively with the positions of the endpoints of any deletions (Δ) in pCD5, pXf3 and pAT153 identified by their nucleotide numbers.


EcoRI
(4361) (0)
Taq (4016)
(23) Cla I / Taq
(29) Hind III
(338) Taq
(375) Bam
PCD5
PXf3
PAT153
PBR322
Δ380 - 649
(631) Hinf I
(650) Sal / Taq
Pst I (3611)
(851) Hinf I
Hinf I (3361)
(1005) Hinf I
(1125) Taq
Δ3102 - 3223
(1266) Taq
(1303) Hinf I
Hinf I (2845)
(1524) Hinf I
Δ1646 - 2351
Δ1426 - 2521
Taq (2573)
Hinf I (2449)
Hinf I (2374)
(2030) Hinf I

The Balb/c 3T3 cDNA library.

Poly A^+, cytoplasmic, RNA was isolated from exponentially growing Balb/c 3T3 cells as described in Chapter 2. The first strand reaction contained 10 µg of poly A^+ RNA in a total volume of 1 ml. First strand synthesis was performed using the conditions described in Chapter 2, including rat-liver ribonuclease inhibitor but omitting NH_4^+ cation. Approximately 1µg of first strand was recovered following base treatment to remove RNA, chromatography over Sephadex G200M and precipitation with ethanol. Following second-strand synthesis, S1-nuclease treatment and chromatography over Biogel A150M, approximately 0.25 µg of double stranded cDNA were recovered following ethanol precipitation. The cDNA was tailed by measuring the rate of incorporation of dGMP residues with time in a pilot reaction, using an estimate of number average cDNA size of 0.25Kb, determined by gel electrophoresis of the pooled cDNA. Because of the necessity of this step for determining the terminal concentration, the cDNA was labelled during the first strand reaction to a specific activity of approximately 10^6 cpm.µg^{-1}. The optimal ratio of vector ends to cDNA ends was determined in the experiment shown in Fig. 3.11, comparison with the data of Fig. 3.10, suggesting that the estimate of ends concentration was correct. 0.25µg of dG-tailed cDNA was annealed to 3.6 µg of dC-tailed pAT153 and transfected into E.coli HB101 as described in Chapter 2; the transfection mix was divided into five aliquots which were plated separately onto nitrocellulose filters, the number of colonies per filter being determined by plating serial dilutions onto L-agar containing antibiotic. Approximately 50,000 colonies in all were obtained, which were stored by scraping off the cells from each filter and storing the five pools separately frozen at -70°C in the presence of 10% (v/v) glycerol. Dilutions of

these pools were plated as required. The overall efficiency in this instance was $2x10^5$cfu μg^{-1} of double-stranded cDNA and the "cloning efficiency" (defined as the number of colonies obtained per ug of input poly A^+ RNA) was therefore $5x10^3$ cfu μg^{-1} RNA.

The SV3T3 C138 library.

For the construction of a cDNA library derived from SV3T3 C138, poly A^+ RNA (5 μg) was used in a 0.5 ml first strand reaction as described in Chapter 2. NH_4^+ was included at 0.2M to improve the length of the product. About 1ug of single-stranded cDNA was recovered following alkali treatment and Sephadex G200M chromatography. Subsequent second-strand synthesis of the ethanol-precipitated cDNA and S1 treatment was followed by fractionation over Biogel A150M. The cDNA was divided into four pools of approximately equal mass to be cloned separately and the terminal concentration of each fraction was determined using the dTTP/ddTTP assay described in Section 3.4. The results of this experiment are shown in Table 3.4. The cDNA from each pool was tailed to saturation with dGTP as described in Chapter 2, using a 20-fold excess of nucleotide over termini and the tailed cDNA was then annealed to a 2-fold molar excess of HindIII plus BstI cleaved pCD5 DNA and transfected into the high efficiency of transfection host DH1, using the procedure documented in Chapter 2. The transfection frequencies obtained with each pool are shown in Table 3.4. Since the cloning efficiency was very high, around $2x10^5$ cfu μg^{-1} using an average value for the efficiency with each pool, the library was largely stored as tailed DNA at -70°C, aliquots being annealed and transfected as required. Subsequent analysis of the average insert size of clones derived from the four size pools showed very little difference

TABLE 3.4

Summary of data pertaining to the construction of two libraries of cDNA

Source of mRNA	Balb/c 3T3	SV3T3 C138
Input mRNA	10 µg	5 µg
Recovery of double-stranded cDNA	0.25 µg	0.17 + 0.15 + 0.19 + 0.22 µg = 0.73 µg
Molar yield from dideoxynucleotide assay	ND	0.52 + 0.48 + 0.82 + 1.0 pmol = 2.82 pmol
Calculated average size	ND	0.98, 0.94, 0.69, 0.64 kb
cfu μg^{-1} cDNA	2×10^5	1.8×10^6; 1.0×10^6, 1.0×10^6, 2.2×10^6 average = 1.5×10^6
cloning efficiency (cfu μg^{-1} mRNA)	5×10^3	average = 2×10^5

The molar yield of SV3T3 C138 double-stranded cDNA was calculated by subjecting 1/25 aliquot of each of the four size pools to dideoxynucleotide assay, using the established standard value of 75 cpm $fmol^{-1}$ of DNA (see Fig. 3.13)

in their relative insert sizes. The experiments described in the following Chapters were thus performed using equal numbers of clones derived from all four size pools.

3.6 Conclusion.

Improved methods of constructing libraries of cDNA have been described. Modifications to the first-strand reaction have produced a high proportion of full-length transcripts at acceptable yields. Methods for monitoring the behaviour of complex mixtures of reverse transcripts, coupled with sensitive gel assays on globin transcripts have allowed high fidelity conversion to double-stranded, blunt-ended cDNA which may be readily cloned using homopolymer tailing. A new strategy for the cloning of cDNA, involving accurate end titration and tailing to saturation, gave improved cloning efficiencies and greatly aided reproducibility. Finally, by employing the high efficiency transfection host *E.coli* DH1, a cloning efficiency of 2×10^5 cfu μg^{-1} of starting poly A^+ RNA was obtained.

With the cloning efficiencies described here it should prove possible to construct libraries of cDNA of representative of the entire poly A^+ RNA of any given cell type. Woods *et al.* (1980) have estimated that approximately 200,000 clones would be required for complete representation of lymphocyte poly A^+ RNA sequences. With the cloning efficiency achieved in this Chapter, this number of clones could be readily obtained using as little as 1 µg of poly A^+ RNA which is obtainable from only 10^6 to 10^7 cells.

CHAPTER IV

SCREENING FOR DIFFERENTIAL GENE EXPRESSION

4.1 Introduction

The first objective of the experiments described in this Chapter was the development of a model system for the identification and molecular cloning of sequences differentially expressed in two largely homologous populations of mRNA. In selecting a suitable system for this study, the following considerations were made.

(a) The mRNA profiles of the cell lines analysed should be very similar. The Balb/c 3T3 line and its SV40-transformed derivative SV3T3 C138 (P.W.J. Rigby, M. Lovett, P.H. Gallimore and C.N. Cole, manuscript in preparation) were selected for comparison since they are derived from the same parental line and, in light of the observations of Williams et al. (1977), it seemed unlikely that they would show extensive differences in the expression of mRNA.

(b) The system should contain a positive internal control. The SV40-transformed line is known to express early viral sequences absent from normal mouse cells (M. Lovett, personal communication). Furthermore, these mRNAs are found at very low levels in transformed cells (Tooze, 1980) offering a means of testing the limits of resolution of the approach. Thus, if no consistent changes in mRNA representation were observed, yet the technology could be developed to such a level that identification of viral sequences was accomplished, one could be confident of the efficiency of the screening procedure and that, in this hypothetical example, cellular gene expression was unaltered.

The second major objective of this work was to ask directly whether limited gene activation or other modulation of expression is induced by transforming agents such as SV40. If so, molecular clones of such affected sequences would be of great interest and their

analysis should further our understanding of SV40-mediated transformation and perhaps ultimately the basis of the neoplastic process.

4.2 Differential colony hybridization analysis

Variations of the colony hybridization technique of Grunstein and Hogness (1975) have been extensively used to screen for sequences differentially expressed in two RNA populations (Williams and Lloyd, 1979; Crampton et al., 1980; Derman et al., 1981). The experiments described in this Section were designed to test the applicability of the technique to this study. Four types of potential change in expression were envisaged: elevated levels in either the normal Balb/c 3T3 or SV3T3 line or de novo expression in one or other cell line. By screening libraries of cDNA clones derived from both cell lines all four possibilities should be detectable, within the confines of the sensitivity of the technique.

Screening of the Balb/c 3T3 library

This was accomplished using the technique of Hanahan and Meselson (1980), at a colony density of approximately 10,000 per 82mm diameter filter. At this density signal intensities obtained following hybridization with kinase-labelled RNA were comparable to those obtained using gridded colonies under identical conditions (Fig. 4.1). Long exposures resulted in extensive spreading of intense autoradiographic signals which could obscure responses from adjacent colonies. This was judged not to be a significant problem since parallel hybridization of a control filter bearing only vector-containing colonies showed considerable background with a number of false positive signals when exposed for a comparable period (Fig. 4.2). Thus any signals requiring such prolonged exposure that

FIGURE 4.1

Figure 4.1 Comparison of the Efficiency of Hybridization to Ordered and Unordered Colony Arrays.

A filter carrying approximately 10,000 clones (A) and a filter carrying 48 gridded clones (B) from the Balb/c 3T3 cDNA library were hybridized in parallel with 10^7 cpm ml^{-1} of $polyA^+$ Balb/c 3T3 RNA labelled *in vitro* with ^{32}P using T4 polynucleotide kinase. Exposure was for 4 days.

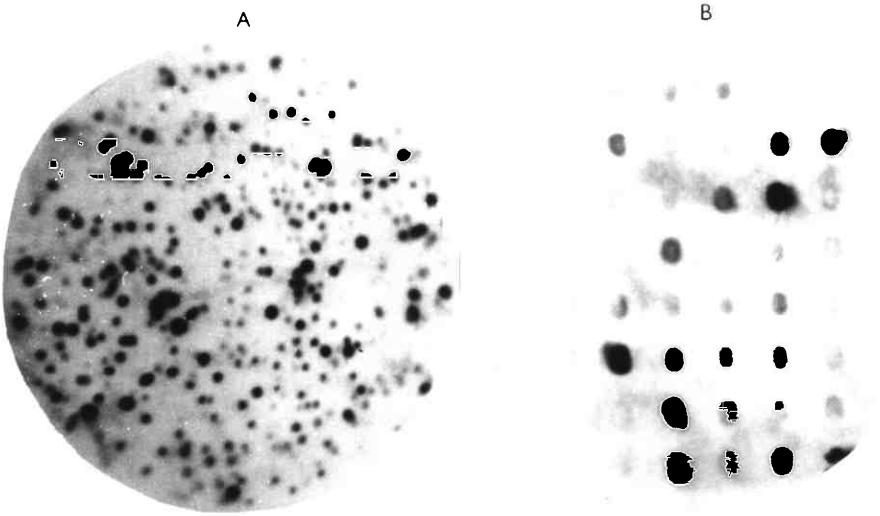
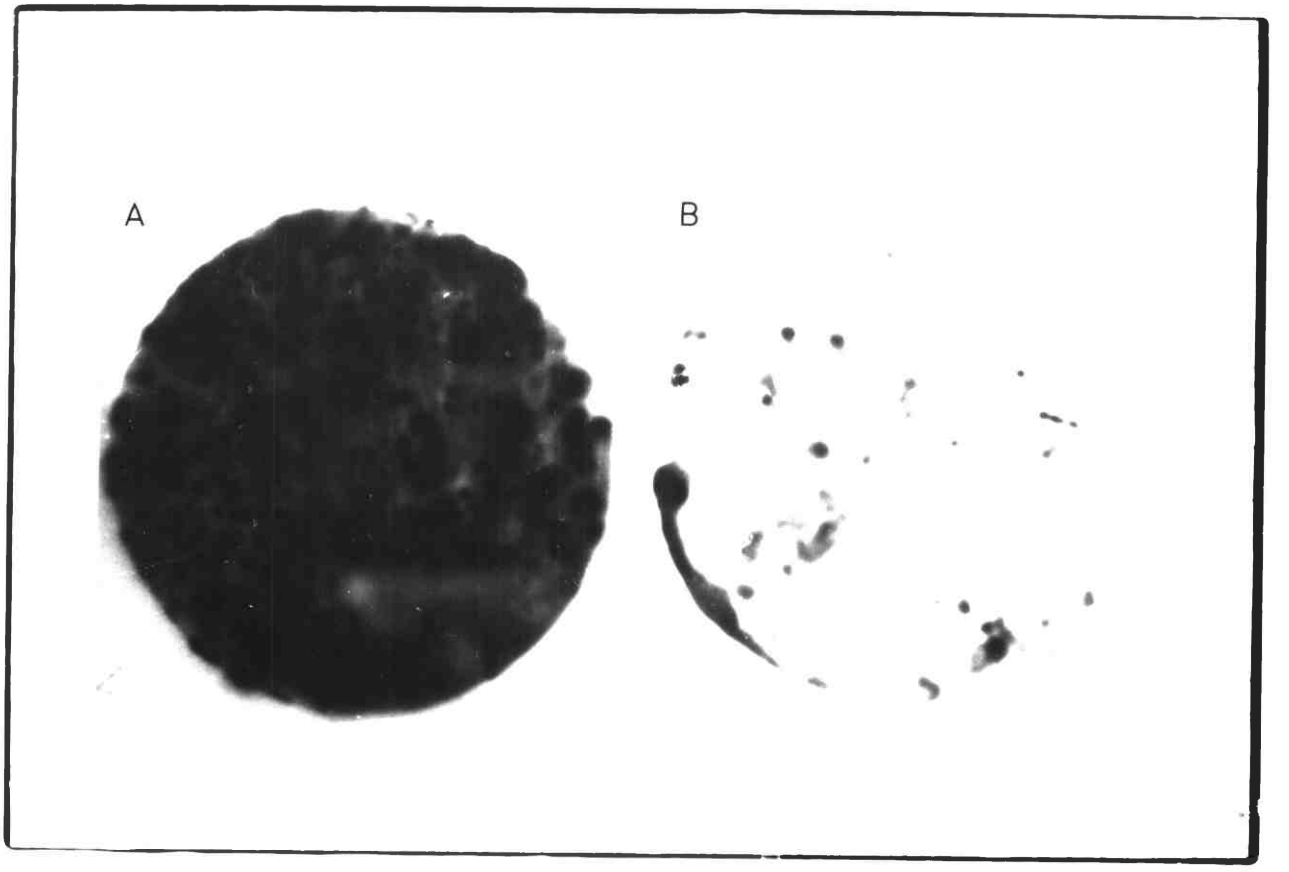

A
B

FIGURE 4.2

Figure 4.2 Background Noise in Colony Hybridization

A filter carrying approximately 10,000 clones from the Balb/c 3T3 cDNA library (A) and one carrying approximately 10,000 colonies of *E. coli* containing pAT153 were hybridized in parallel with 10^7 cpm ml^{-1} of $polyA^+$ Balb/c 3T3 RNA labelled *in vitro* with ^{32}P using T4 polynucleotide kinase. Exposure was for 21 days.


A
B

"swamping" becomes a significant problem would not be discriminated over background under the experimental conditions used. Screening at such high density allows very large numbers of colonies to be analysed to allow the detection of alterations in the expression of mRNAs which might be represented at very low frequency in the cDNA library.

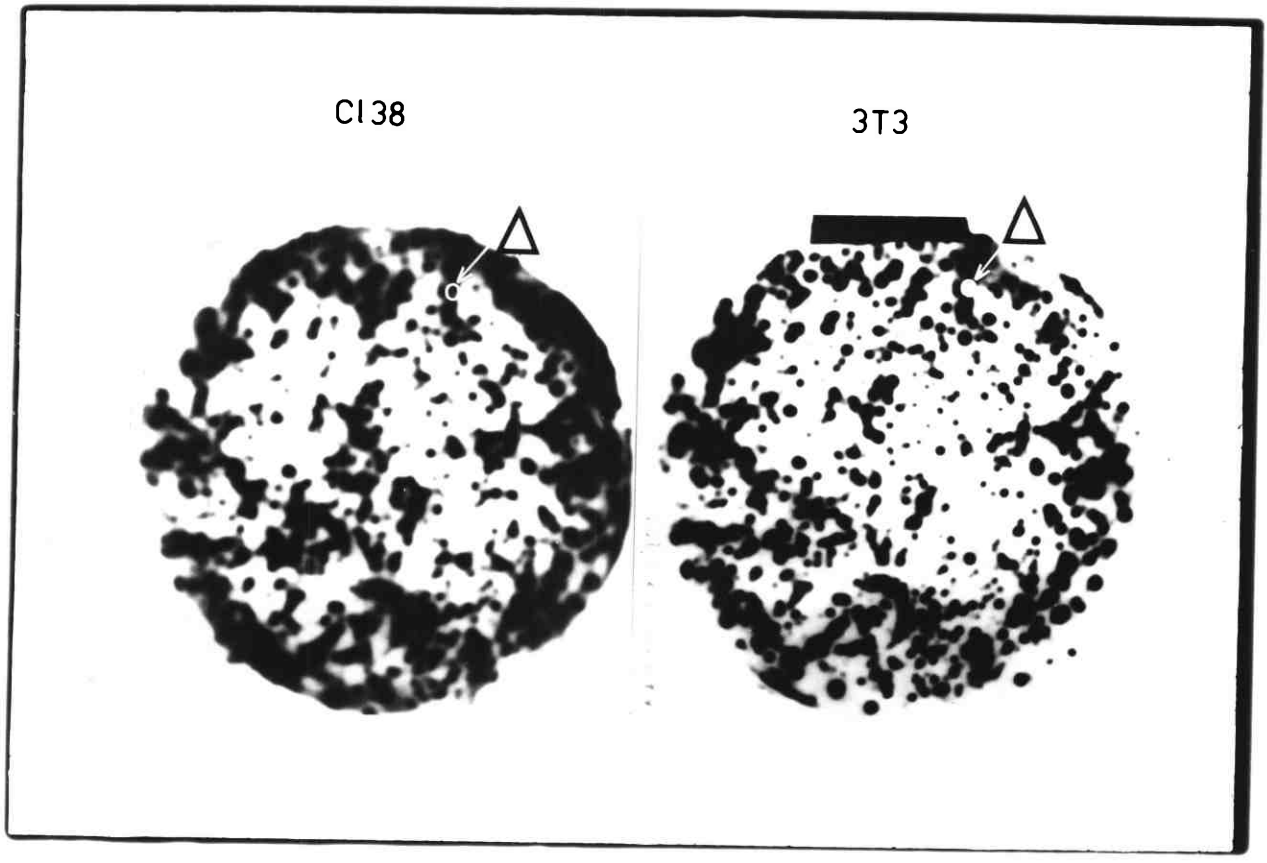
Preliminary experiments (my unpublished observations) indicated that the fidelity of replica plating using the high density screening technique of Hanahan and Meselson (1980), although high, yielded artifactual results where, at a frequency of about 1%, colonies failed to transfer to one or other replica filter. To reduce the occurrence of such artifacts, four replica filters were prepared, the first and last being hybridized with one probe and the second and third with the other (see Chapter 2). Approximately 50,000 colonies were screened to allow for the possibility that some mRNAs might be expressed at low levels in normal cells, perhaps below the detection threshold of the screening technique, yet might be expressed at sufficiently high levels in transformed cells to elicit a signal. In this experiment, five sets of replica filters were prepared, each comprising approximately 10,000 Balb/c 3T3 cDNA clones (see Chapter 3). Two replica filters were hybridized with kinase-labelled poly A^+ 3T3 RNA (see Chapter 2) and two with an SV3T3 C138 probe. For subsequent analysis, autoradiographs of the two replica filters obtained with each probe were superimposed before reproduction here.

The result obtained with one of the five filter sets is shown in Fig. 4.3. The vast majority of detectable responses were the same with each probe, indicating that no differences exist in the expression of normal cellular abundant poly A^+ mRNAs. In this, and in all subsequent analyses to be described, areas where several

FIGURE 4.3

Figure 4.3 Screening of the Balb/c 3T3 DNA Library by Differential Colony Hybridization.

Approximately 10,000 clones from the Balb/c 3T3 cDNA library were screened by differential colony hybridization with Balb/c 3T3 and SV3T3 C138 polyA$^+$ RNA labelled with ^{32}P *in vitro* using T4 polynucleotide kinase. Both probes were included at a concentration of 10^7 cpm ml^{-1} and exposure was for 6 days. Autoradiographs of the two replica filters hybridized with each probe were superimposed prior to reproduction. The only differential response is indicated (Δ).


Cl38
3T3

colonies showed different signals with each probe were ignored as being likely to represent artifacts of the screening procedure. Only single colonies (singlets) which exhibited a differential response and where neighbouring colonies gave an unaltered pattern of hybridization were considered for further analysis. Six potential candidates were identified from all five filter sets. One of these is shown in Fig. 4.3. In this case, an apparent reduction of expression level in C138, an area surrounding the colony was removed the master filter using a 3.0mm diameter cork borer and dispersed in medium. At the colony density used this corresponds to about 40 colonies. The cells from this area were then replaced at approximately 500 colonies per filter for rescreening. Using the formula $1-P = [(X-1)/X]^N$, (Clark and Carbon, 1976) where P = probability for having any sequence, X = total number of sequences (40) and N = the number of colonies plated (500), the probability of having the original clone represented is 99.99%. The result of the rescreen is shown in Fig. 4.4. No singlets were observed to respond in a manner consistent with the original signal. Rescreening of the remaining five putatives yielded the same negative result. Note that the error rate in the original screen (6/50,000) is approximately that expected assuming an error rate of 1% in the replica plating procedure, which would be expected to generate artifacts at a frequency of 0.01% using two replica filters with each probe.

The potential sensitivity of the technique was examined by counting the number of signals obtained in an area picked at random, using exposures where background was insignificant (Fig. 4.5). Comparison with the total number of colonies present in the corresponding area of the master filter demonstrated that only 15% of

FIGURE 4.4

Figure 4.4 Rescreening of a Putative Differentially Hybridizing Clone.

Approximately 40 colonies from the area indicated in Fig. 4.3 were dispersed in L-broth and replated at a colony density of 500. Four replica filters were prepared as described in Section 2.8, the first and last being hybridized with the Balb/c 3T3 probe and the second and third were hybridized with the SV3T3 C138 probe. Autoradiographs of the two replica filters hybridized with each probe were superimposed prior to reproduction. The probes used are the same as those described in the legend to Fig. 4.3. Exposure was for 7 days.

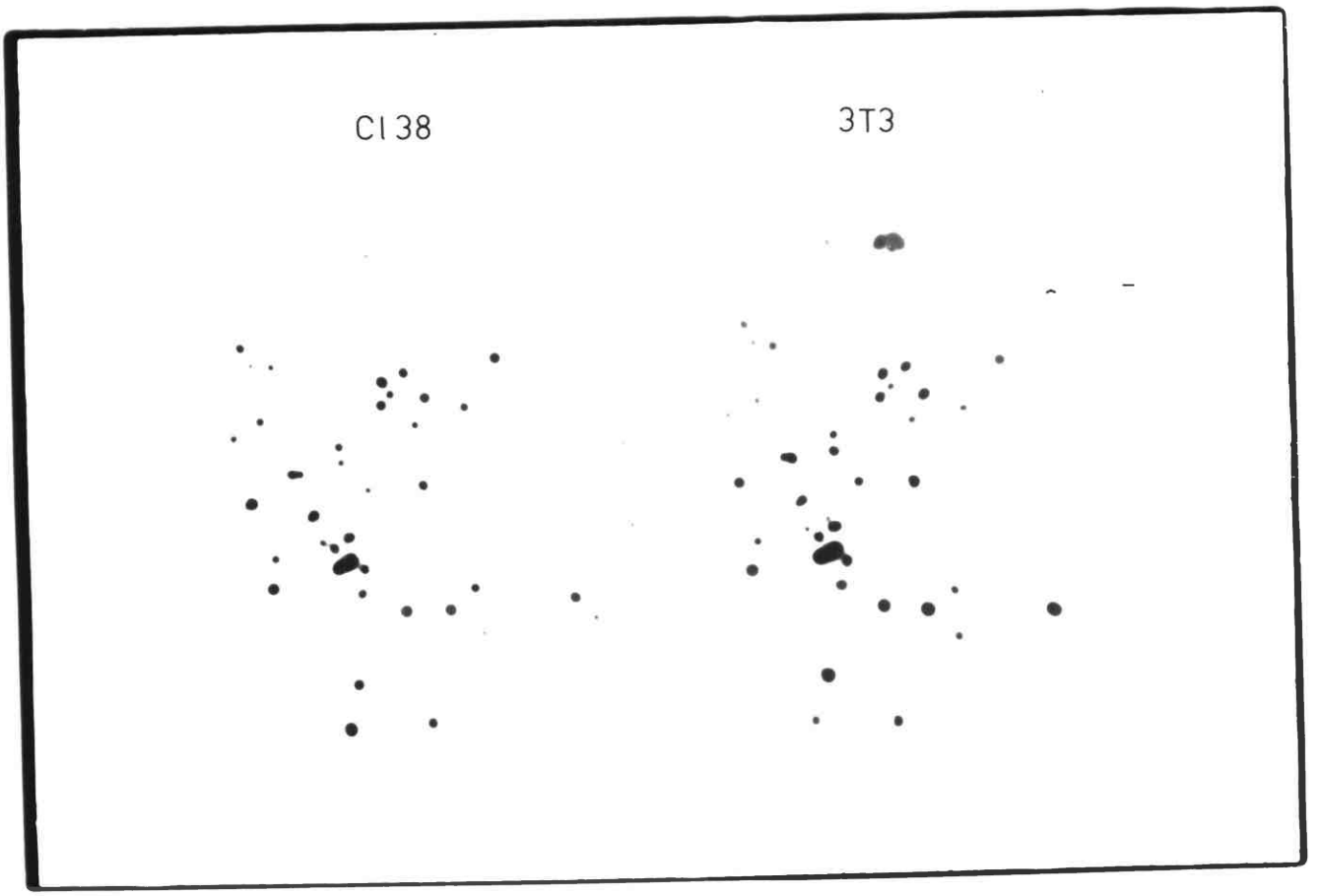
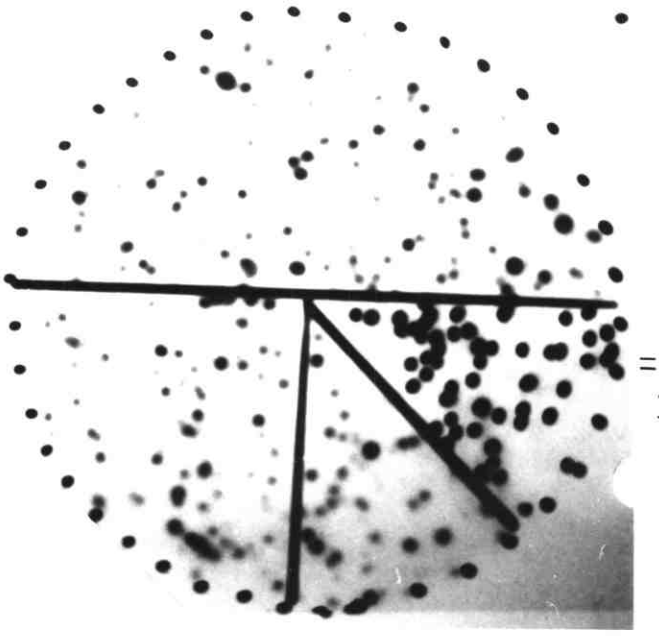

Cl 38
3T3

FIGURE 4.5

Figure 4.5 Sensitivity of Colony Hybridization.

A filter carrying approximately 3,000 clones from the Balb/c 3T3 cDNA library was screened with ^{32}P-labelled Balb/c 3T3 polyA$^+$ RNA, labelled in vitro with T4 polynucleotide kinase (Section 2.6). The sp. act. was 3×10^7 cpm μg^{-1} and the probe concentration was 350 ng ml^{-1}. Exposure was for 4 days. The filter was divided into eight equal sections and the number of detectable autoradiographic signals was counted in one section and divided by the number of colonies in the equivalent section of the master filter, giving the volume shown.



$= \frac{54}{356} = 15.1\%$

the clones elicited a signal with kinase-labelled RNA under the conditions of hybridization used, indicating that the vast majority of poly A^+ RNAs in the cell are inaccessible to analysis by this technique.

Screening of the SV3T3 C138 library

A reciprocal screen of clones from the transformed library was performed in the same way since, by definition, sequences found in transformed cells but absent from normal cells would not be represented in the 3T3 library.

Four replica filters bearing approximately 5,000 SV3T3 C138 cDNA clones were prepared and hybridized with 3T3 and C138 kinase-labelled RNA as described in Chapter 2. It was not considered necessary in this case to screen a larger number since the data obtained from the previous experiment confirmed the conclusion of other workers (Gergen et al., 1979; Crampton et al., 1980) that the technique can only monitor the behaviour of relatively abundant RNAs. The sensitivity achieved has usually been no better than 0.1% of total poly A^+ RNA. Thus, a screen of 5,000 colonies should ensure representation of any sequence abundant enough to lie within the detection threshold of the technique. Fig. 4.6 shows the result of this analysis. No differentially responding singlets were observed.

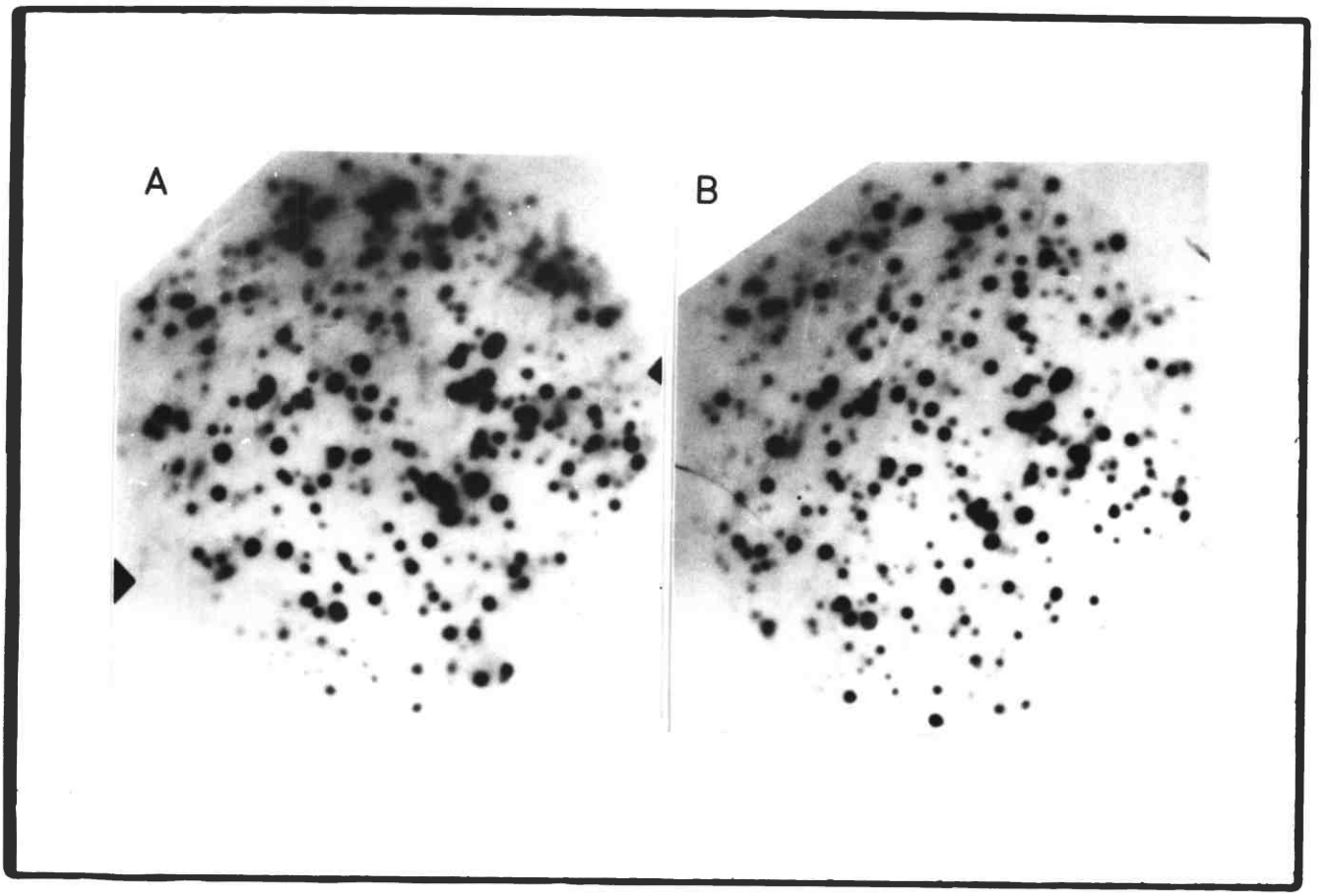
The results described in this Section yield two conclusions; that no alterations in the expression levels of abundant cellular poly A^+ RNAs have accumulated during the establishment and propagation of the C138 line and that the technique of differential colony hybridization is prone to artifactual results and is incapable of monitoring the behaviour of the majority of expressed poly A^+ RNAs.

Fig. 4.2 illustrates that the major barrier to high sensitivity

FIGURE 4.6

Figure 4.6 Screening of the SV3T3 C138 cDNA Library by Differential Colony Hybridization.

The legend to this Figure is essentially as for Fig. 4.3 except that the SV3T3 C138 cDNA library was used.


A
B

differential screening is the detection threshold imposed by signal-to-noise-ratio. In order to explore the expression of sequences represented at lower levels in mRNA populations this barrier must first be passed. Two ways of accomplishing this aim are obvious. The first is to improve the experimental conditions, perhaps by using a different host/vector system and modifying the conditions of hybridization and washing to ameliorate signal-to-noise ratio. This was not pursued as it was deemed unlikely ever to yield a significant improvement in sensitivity. The second approach is to purify a probe representative of the sequences of interest. Since "noise" in colony hybridization is a function of the great majority of probe which is not complementary to any given clone, prior removal of material unrelated to that to be scrutinized should dramatically reduce background noise. This rationale was applied to the 3T3/SV3T3 system.

4.3 Purification of probes using DNA-cellulose

For reasons of technical simplicity, attention will be concentrated upon the identification of one subset of the potential changes in gene expression documented at the beginning of Section 4.2, that is the *de novo* synthesis of specific RNAs in the SV3T3 line, or alternatively, expression at greatly elevated levels in cases where those sequences might be of very low abundance in the normal cell control. Considerations discussed in Chapter 1 indicate that gene activation could accompany transformation by SV40. Furthermore, the expression of viral-specific sequences in the transformed line could serve as an internal control in the analysis. If successful, the technique could readily be applied to search for normal cell-specific sequences and in Chapter 6 I shall discuss possible ramifications of

the principle of partial probe purification, with the increased sensitivity it confers, to the task of identifying small fluctuations in the abundance of cellular RNAs.

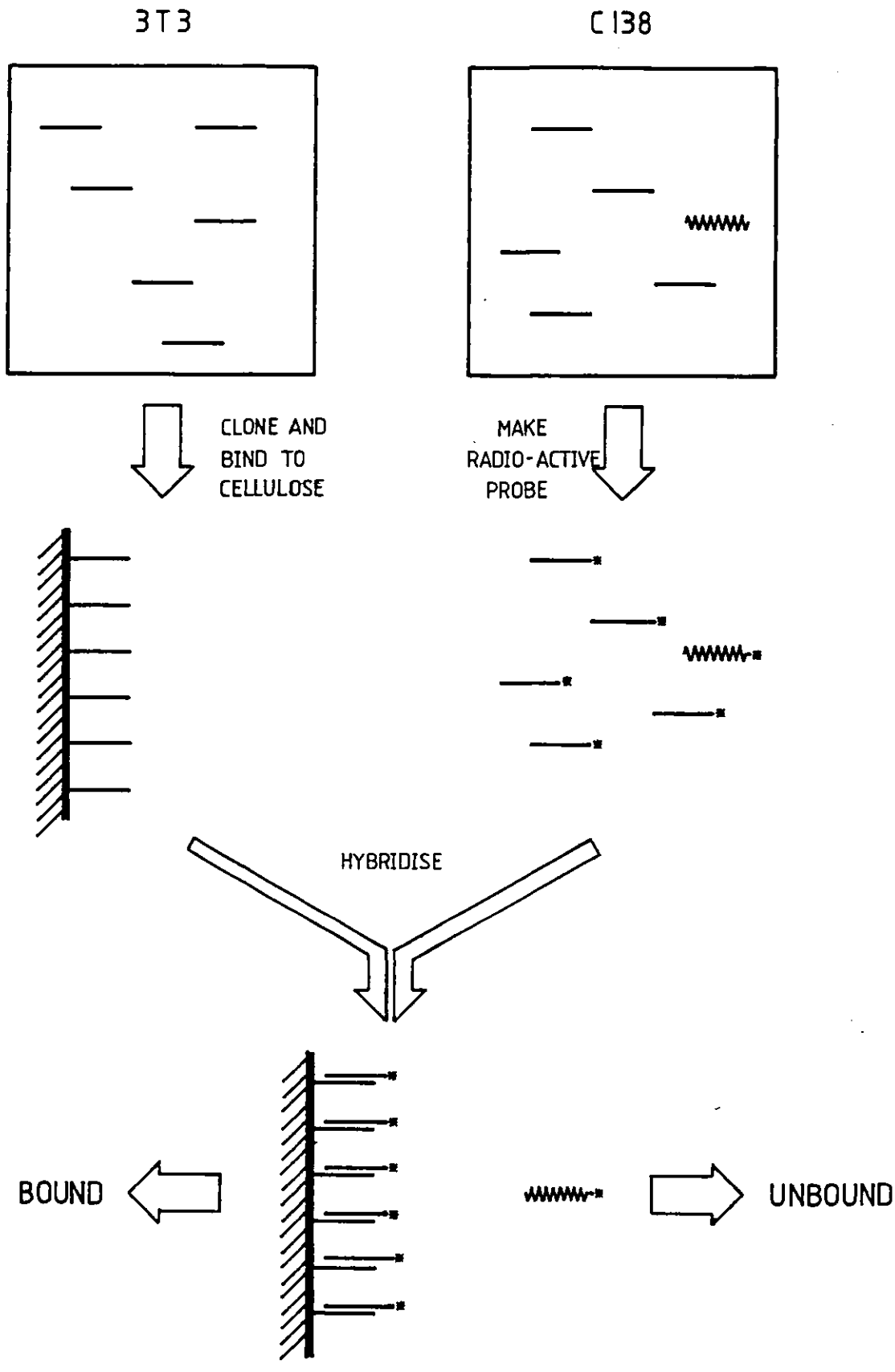
Partial purification of SV3T3 C138-specific probe

The strategy was to eliminate from a transformed cell probe sequences complementary to a normal Balb/c 3T3 cDNA library and is illustrated in schematic form in Fig. 4.7. The Balb/c 3T3 library of approximately 50,000 clones was pooled, amplified and plasmid DNA were fragmented, denatured and chemically coupled to cellulose (Noyes and Stark, 1975; see also Chapter 2). The efficiency of binding, monitored by inclusion of trace amounts of ^{32}P-labelled nick-translated bacteriophage DNA, indicated that 80.6% (3.2mg) of cDNA-containing plasmid DNA became covalently bound to the cellulose (72mg dry weight). Assuming an average insert size of approximately 0.5Kb, this corresponds to ~350ug of single-stranded 3T3 cDNA equivalents. All hybridizations were performed at a 3T3 cDNA concentration of ~1mg ml^{-1} and a ^{32}P-labelled C138 cDNA concentration of ~1μg ml^{-1}. There was thus, on average, a 1,000- fold excess of normal cell sequences; given that the frequency at which a particular sequence occurs in the library should reflect the abundance of that sequence in the mRNA population from which the library was derived, this 1,000-fold excess should hold for all sequences unaffected by transformation. In a preliminary experiment, one tenth of the cDNA-cellulose was subjected to several rounds of hybridization against 30ng of ^{32}P-labelled cDNA (sp. act. = 1.2 x 10^{8} cpm μg^{-1}). After each 24 hour round of hybridization the cellulose was exhaustively washed and the unhybridized cDNA was rehybridized to the 3T3 cDNA-cellulose, from which the previously hybridized material had

FIGURE 4.7

Figure 4.7 Scheme for Purification of a Probe Specific for SV3T3 C138 mRNA.

This scheme is described in the text of Section 4.3. Sequences represented in Balb/c 3T3 mRNA are indicated by ——— , and those that are specific for SV3T3 C138 RNA by ∿∿∿ . Radioactively labelled material is identified by *.


3T3
C138
CLONE AND
BIND TO
CELLULOSE
MAKE
RADIO-ACTIVE
PROBE
HYBRIDISE
BOUND
UNBOUND

been eluted. The results of this analytical experiment are shown in Table 4.1 and graphically in Fig. 4.8. Following five successive cycles of hybridization a limit of about 25% of input probe which failed to anneal to the cDNA-cellulose was attained.

Solution hybridization analysis of 3T3-depleted cDNA

To investigate the sequence complexity of this 3T3-depleted SV3T3 cDNA probe, the kinetics of hybridization to normal 3T3 and SV3T3 C138 RNA were examined using the established technique of mRNA/cDNA solution hybridization (Bishop et al., 1974b; Hastie and Bishop, 1976; Williams et al., 1977). An example of such an experiment, using unfractionated probe, is shown in the upper curve of Fig. 4.9. Here ^{32}P-labelled SV3T3 C138 cDNA (sp. act. 9 x 10^7 cpm μg^{-1}) was annealed to an excess of C138 poly A^+ RNA as described in Chapter 2. The percentage of probe hybridized, determined by endonuclease S1 assay, was plotted as a function of RNA concentration x time ($R_o t$ = moles nucleotide per litre x seconds). Although insufficient samples were assayed to allow the detailed analysis applied by other workers (Bishop et al., 1974b; Hastie and Bishop, 1976; Williams et al., 1977) the data provide an adequate representation of the overall trend of the hybridization reaction. Annealing was observed over 2 to 3 decades of $R_o t$, exposing a broad distribution in the abundance of cellular poly A^+ RNAs. More detailed analyses (Bishop et al., 1974b; Hastie and Bishop, 1976; Williams et al., 1977) have usually indicated the existence of discrete abundance classes of mRNA, although the generality of this phenomenon has been challenged (Quinlan et al., 1978; Woods et al., 1980). No obvious classes were perceived in this instance but the general trend was reasonably typical.

TABLE 4.1

Repeated Hybridization of C138 cDNA to 3T3 cDNA-cellulose

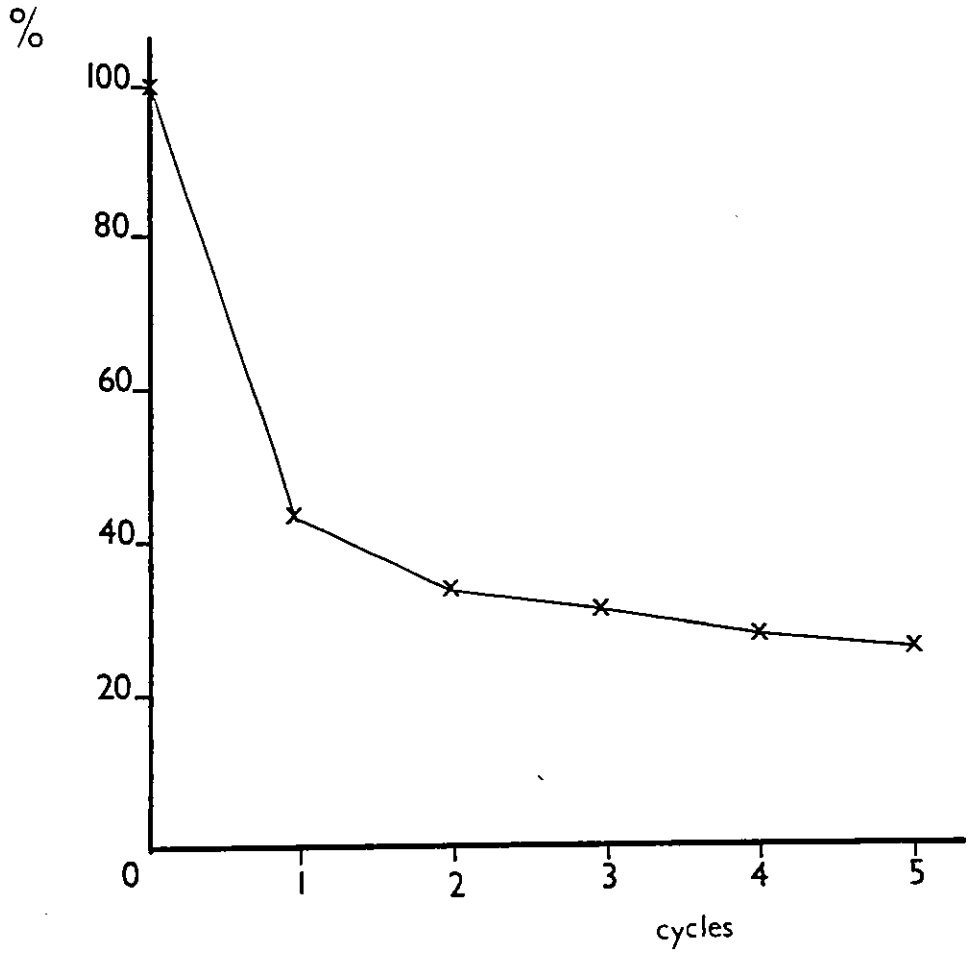
Hybridization Cycle	Input cDNA	Hybridized cDNA	Unhybridized cDNA
1	100%	57.7%	42.2%
2	42.2%	9.3%	32.9%
3	32.9%	2.9%	30.0%
4	30.0%	2.7%	27.3%
5	27.3%	0.8%	26.5%

^{32}P-labelled SV3T3 C138 cDNA was hybridized for 24 h periods to 3T3 cDNA-cellulose. The fraction of the probe hybridizing and failing to hybridize during each cycle is expressed as a percentage of original input cDNA.

FIGURE 4.8

Figure 4.8 Repeated Hybridization of C138 cDNA to cDNA-Cellulose.

The legend of this Figure is as for Table 4.1. The fraction of the probe that failed to hybridize during each cycle is shown.

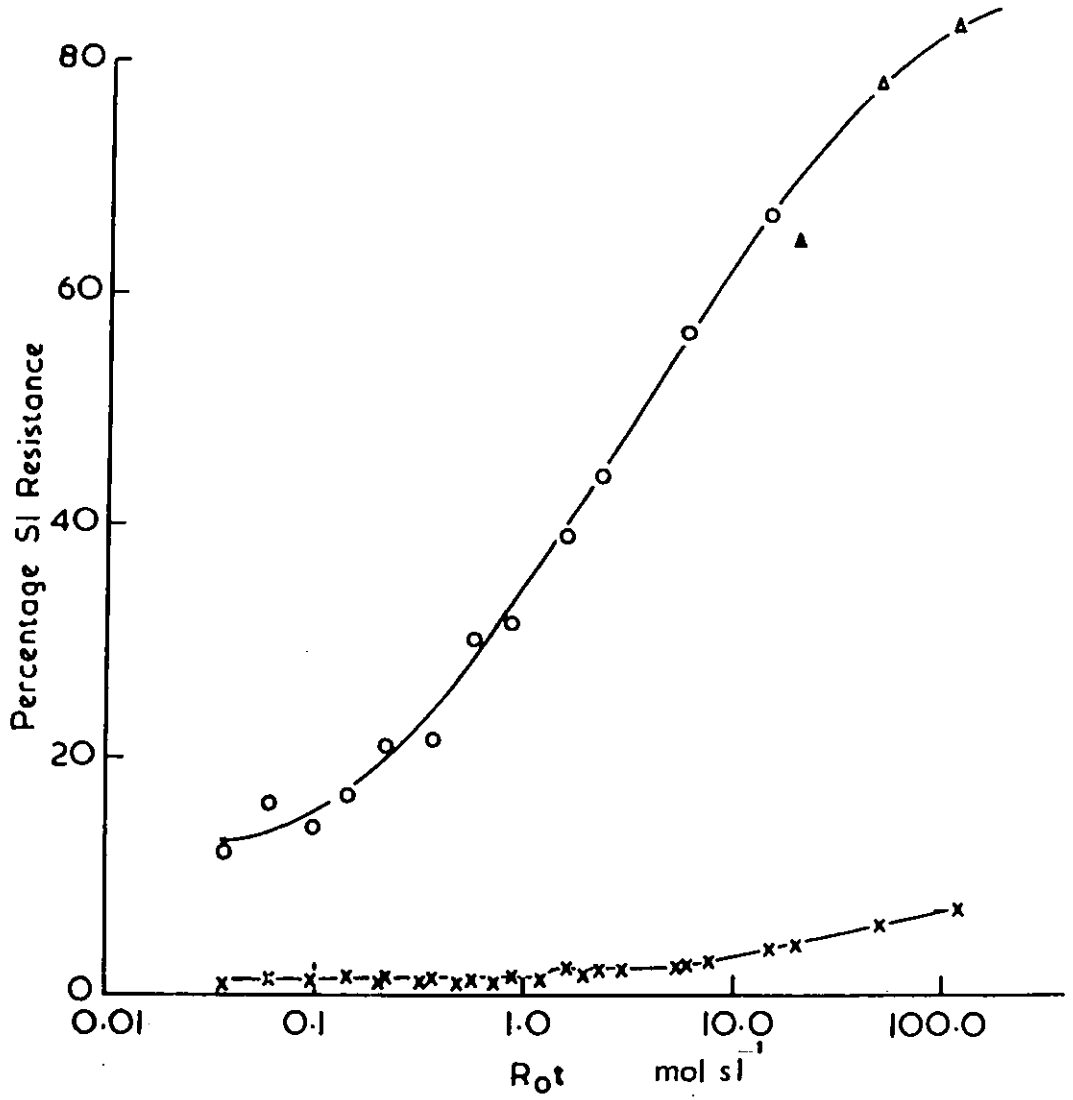

%
100
80
60
40
20
0
1
2
3
4
5
cycles

When the 3T3-depleted SV3T3 probe was annealed to C138 poly A^+ RNA, however, a strikingly different profile was obtained (Fig. 4.10). As expected no hybridization was observed at low R_0t values (<1mol $l^{-1}s$), indicating that the probe had been completed depleted of sequences complementary to abundant RNAs. At higher R_0t values (>1mol $l^{-1}s$), two distinct components were resolved. The first, with an estimated $R_0t_{\frac{1}{2}}$ of 1.3, amounts to about 4% of input probe and thus represents approximately 1% of C138 poly A^+ RNA. The second component, representing sequences of lower abundance, was evident at R_0t values in excess of 10mol $l^{-1}s$, beyond which a smooth and gradual transition was observed up to and presumably beyond the highest R_0t value assayed. It should be noted that mRNA/cDNA hybridization experiments usually plateau at about 80%, never showing complete hybridization. The most obvious explanation for this observation (for examples see Hastie and Bishop, 1976; Williams et al., 1977; Woods et al., 1980) is that the remaining fraction comprises sequences which are incapable of efficient hybridization. For example, very short probe fragments might fail to hybridize properly yet still precipitate with acid and register in the assay. Thus a figure of 25% for the fraction of probe that fails to anneal to cDNA-cellulose probably is an overestimate since a significant proportion of that fraction may consist of sequences that are too short to form stable hybrids. In Fig. 4.9 the result obtained from hybridization of 3T3-depleted C138 cDNA with its template RNA has been converted to percentage of original probe (prior to fractionation) to allow comparison with the data obtained using total C138 cDNA as probe. Comparison of the two curves demonstrates the power of the purification procedure.

FIGURE 4.9

Figure 4.9 Hybridization of ^{32}P-labelled SV3T3 C138 cDNA to Its Template mRNA.

cDNA (sp. act. = 1 x 10^8 cpm μg^{-1}, 1.2 x 10^7 cpm ml^{-1}) synthesized from SV3T3 C138 polyA$^+$ RNA was hybridized with 60μg ml^{-1} (O) and 200 μg ml^{-1} (Δ) of SV3T3 C138 polyA$^+$ RNA (upper curve). Aliquots were subjected to endonuclease S1 assay (see Section 2.12) following different times of incubation. Percentage resistance to S1 nuclease is shown as a function of $R_o t$ (mol l^{-1}s). The lower curve (X) is described in the text of Section 4.3.

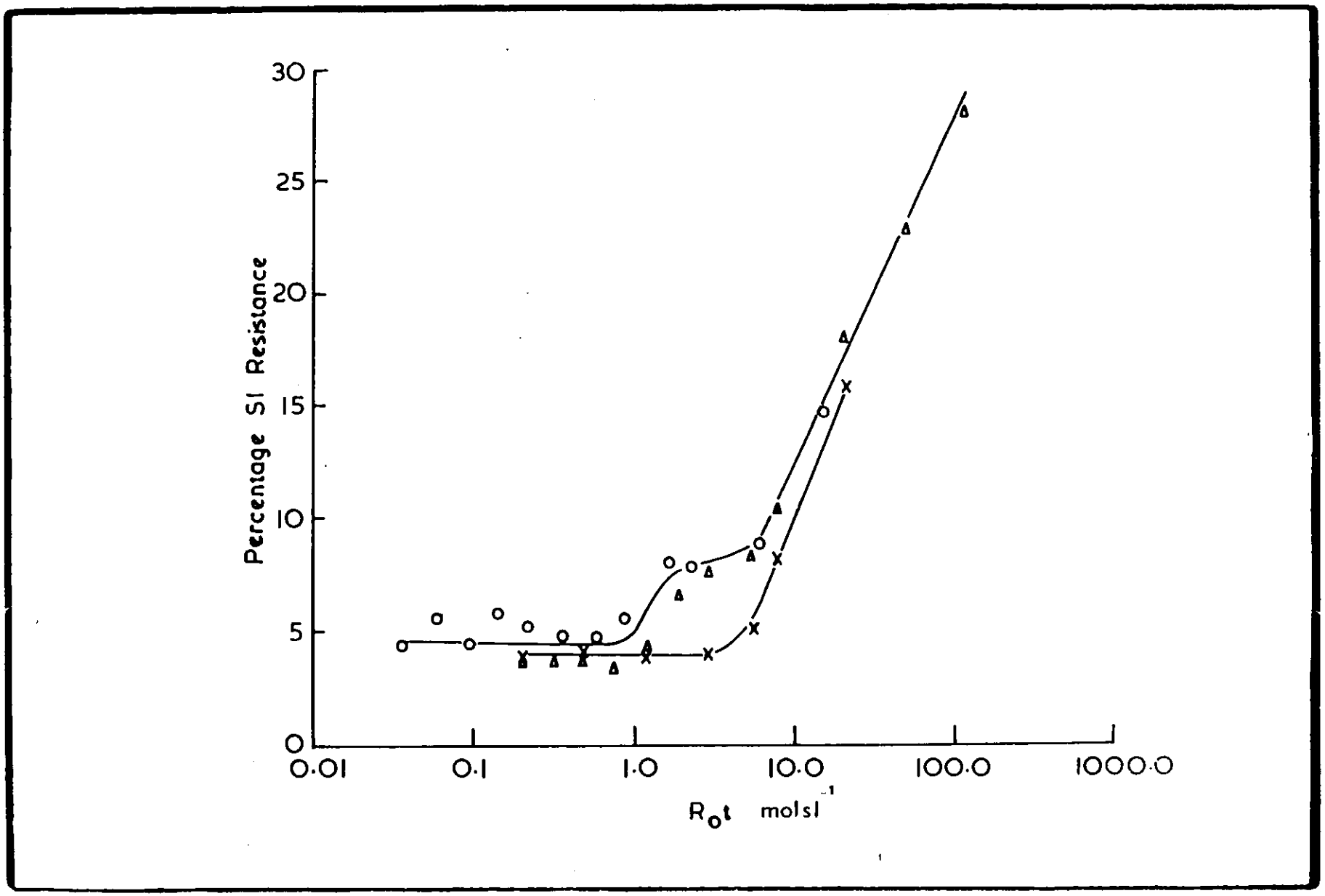

80
60
40
20
0
Percentage SI Resistance
0·01
0·1
1·0
10·0
100·0
R_0t mol sl^{-1}

In a parallel experiment, the 3T3-depleted probe was hybridized against Balb/c 3T3 poly A^+ RNA (the same preparation of RNA as was used to construct the 3T3 cDNA library). The more slowly annealing component of >10mol l^{-1}s was observed (Fig. 4.10) indicating that the probe still contained some sequences homolgous to low abundance normal 3T3 RNAs. This result is not unexpected since clones corresponding to RNAs of low abundance would be present infrequently, if at all, in the library. Hence the concentration of complementary DNA bound to cellulose would be low and hybridization would be inefficient. Strikingly, however, the more abundant ($R_0t_{1/2}$ = 1.3) transition was absent from the 3T3-driven reaction, indicating that these sequences represent a fraction either absent from or present at much lower levels in Balb/c 3T3 RNA than in SV3T3 C138. Assuming an $R_0t_{1/2}$ of 1.3, the transformation-specific component comprising 1% of total C138 mRNA would have an $R_0t_{1/2}$ when pure of 1.3×10^{-2}. Assuming that the average eukaryotic message has an $R_0t_{1/2}$ of 9×10^{-4} (Bishop et al., 1974b), the C138-specific transition would correspond to about 14 sequences each representing 0.07% of the total poly A^+ RNA. The results described in this Section thus offer evidence for the expression in an SV40-transformed cell line of a small number of poly A^+ RNAs of moderate abundance which are either absent from or expressed at much lower levels in normal cells. The identity and behaviour of these messages is of great interest from a standpoint of SV40-mediated transformation. Furthermore, the basic technology described above may be extended to allow molecular cloning of these differentially expressed sequences, satisfying both the primary objectives of Section 4.1.

FIGURE 4.10

Figure 4.10 Hybridization of ^{32}P-labelled 3T3-depleted C138 cDNA to Balb/c 3T3 and SV3T3 C138 mRNA.

The 3T3-depleted SV3T3 C138 cDNA probe obtained following five cycles of hybridization to 3T3 cDNA cellulose (see Table 4.1 and Fig. 4.8) was hybridized with 60µg ml^{-1} (O) and 200µg ml^{-1} (Δ) of SV3T3 C138 polyA$^+$ RNA and with 60µg ml^{-1} of Balb/c 3T3 polyA$^+$ RNA (X). Percentage resistance to S1 nuclease is shown as a function of R_0t (mol l^{-1}s).


30
25
20
15
10
5
0
Percentage SI Resistance
0·01
0·1
1·0
10·0
100·0
1000·0
R_0t molsl^{-1}

4.4 Colony screening with the purified probe

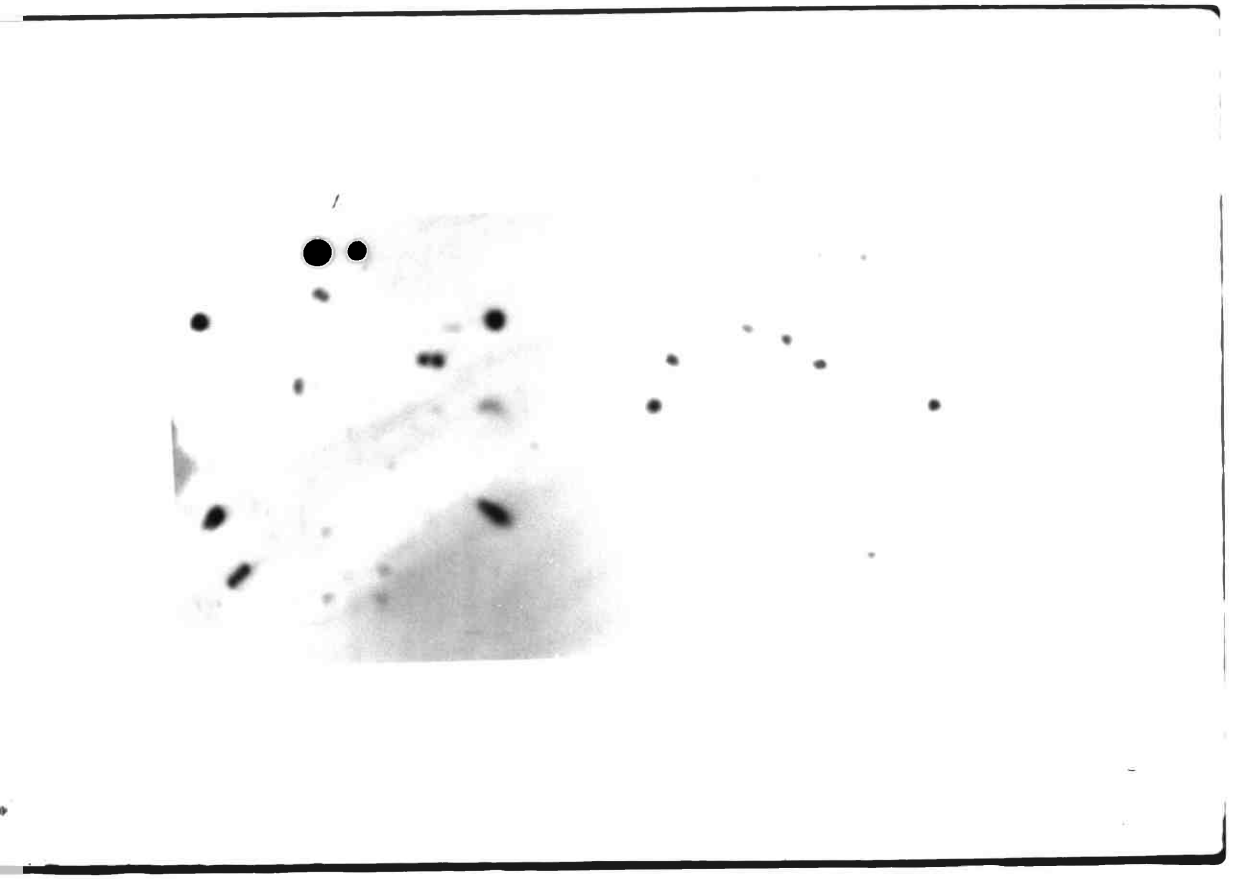
The probe depletion experiment was repeated on a preparative scale to yield sufficient probe for colony hybridization screening. 350ng of ^{32}P-labelled C138 cDNA (sp. act. = 1 x 10^8 cpm μg^{-1}, 3.65 x 10^7 cpm total) was hybridized against cellulose carrying 350μg of single-stranded 3T3 cDNA equivalents in a total volume of 0.4ml. Following five successive cycles of hybridization 9.5 x 10^6 cpm (26% of input probe) failed to anneal, in excellent agreement with the result of the analytical reaction. To further purify the $R_0t_{1/2}$= 1.3 component, the 3T3-depleted C138 cDNA was hybridized to its template mRNA to an R_0t value of 10 mol $l^{-1}s$. The hybridization reaction was quenched, unhybridized cDNA was digested with endonuclease S1, the undegraded cDNA was recovered by gel filtration on Sephadex G200 and RNA removed by alkaline hydrolysis. Further details are supplied in Chapter 2. 3.75% of input cDNA (1.37 x 10^6 cpm) was recovered, in excellent agreement with the value of 3.25% predicted from the R_0t analysis of Fig. 4.10. The transformation-specific sequences are thus 25% of the remaining probe and should constitute the most abundant sequences in the probe. This fraction was used to screen approximately 6400 SV3T3 C138 cDNA clones, at a colony density of 800 per 82mm diameter filter. The screen was performed at a probe concentration of $\sim$7 x 10^4 cpm ml^{-1}, corresponding to 10^3 - 10^4 cpm ml^{-1} of each individual species.

One hundred and fifty potential signals were picked following a four day exposure, arrayed on grids and rescreened with the same probe. Two typical filters from the first screen are shown in Fig. 4.11. Only the most intense signals are photographically

FIGURE 4.11

Figure 4.11 Colony Screening with the Purified Probe

Eight 82mm diameter filters, each carrying approximately 800 clones from the SV3T3 C138 library were screened as described in Sections 2.8 and 2.10 with the C138 specific cDNA probe described in the text of Section 4.4. Two representative filters are shown (5 day exposure).



reproducible with several very faint signals barely visible even by direct inspection of the autoradiograph. The result of the rescreen is shown in Fig. 4.12. Several control colonies carrying a plasmid containing several copies of the SV40 genome (W. Chia, M.R.D. Scott and P.W.J. Rigby, manuscript in preparation) were included to test whether the probe contained viral-specific sequences. Even after prolonged exposure (14 days) identification of many signals was extremely uncertain, although some of the clones gave clear responses. Forty-two possible positives in all were purified for further analysis to be described in Chapter 5. Note that no visible signal was detected from the SV40 clones. Note also that, as expected, a control filter bearing about 500 Balb/c 3T3 cDNA clones did not exhibit any detectable positive signals when hybridized in parallel.

4.5 Conclusion

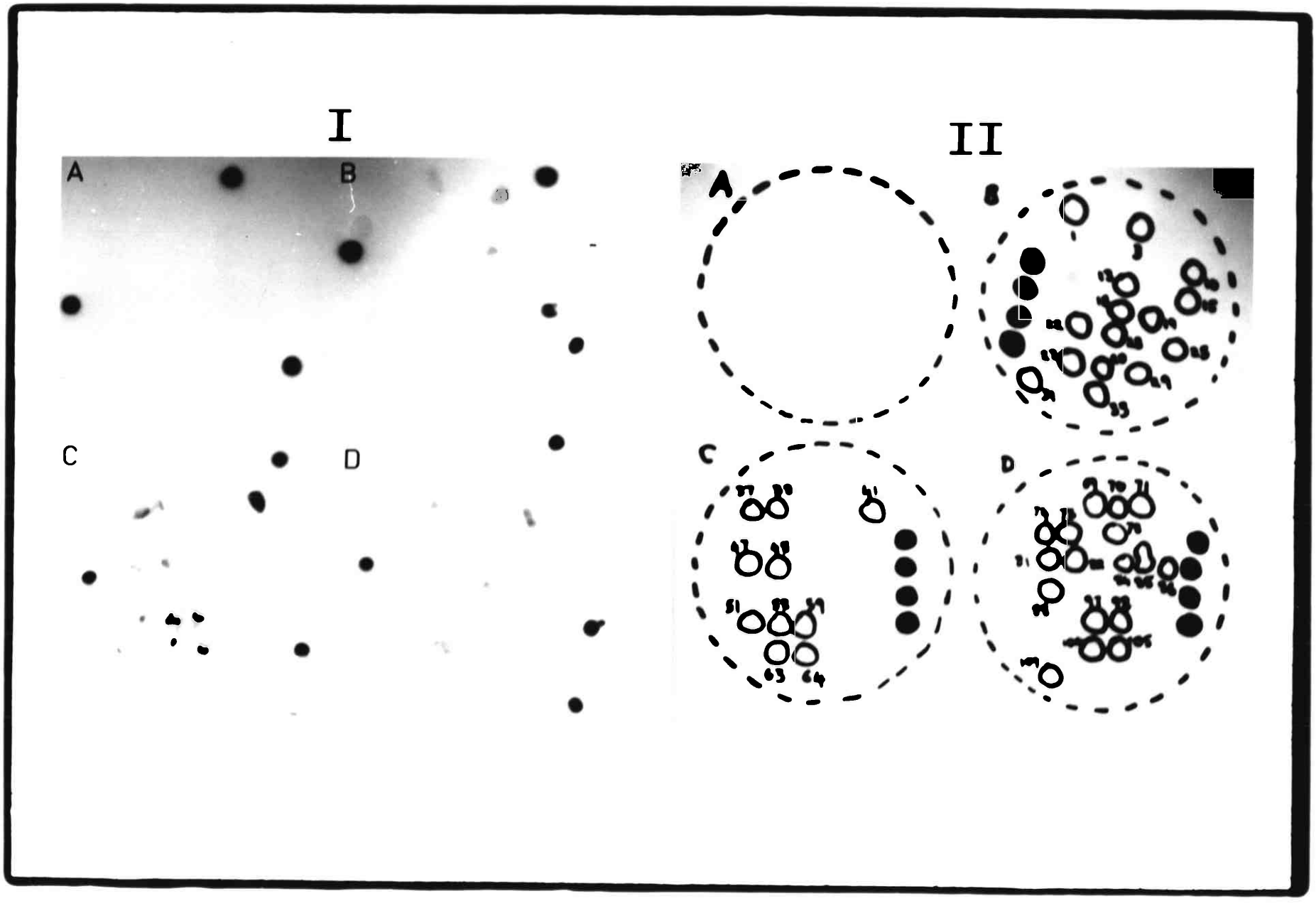
Approaches to the identification and molecular cloning of rare sequences differentially expressed in two grossly similar cell types have been explored. Conventional differential colony hybridization was found to be too insensitive to detect changes in the expression levels of the majority of cellular messenger RNAs. Purification of probe by prior depletion of common sequences allowed detection and cloning of a set of sequences expressed in an SV40-transformed cell line which are present at much lower abundance, if at all, in normal cells, and are represented below the detection threshold of differential colony hybridization. The disclosure of a small number of mRNAs of this nature yields support for the hypothesis that limited host gene activation may accompany transformation by SV40.

Detection of viral-specific sequences was not achieved in this experiment. This might have been accomplished by extending the final

FIGURE 4.12

Figure 4.12 Rescreening with the Purified Probe

One hundred and fifty putative positives from the screen of Fig. 4.11 were ordered in grids on filters (B, C, D) and rescreened with the C138-specific cDNA probe (see Section 4.4). A control filter (A) carrying approximately 500 clones from the Balb/c 3T3 cDNA library was included in the hybridization reaction. Twelve colonies containing tetramers of the SV40 genome cloned in a plasmid vector (W. Chia, M.R.D. Scott and P.W.J. Rigby, manuscript in preparation) were included with the putative positives of filters B, C and D. Part I shows the result of this experiment (the intense signals at the perimeter of each filter are radioactive orientation markers) and Part II shows a summary of the data. The positions of the clones selected for further analysis are identified by rings and the isolate number of each clone. The positions of the SV40 containing clones on filters B, C and D are identified by spots. Exposure was for 14 days.


I
II
A
B
C
D
A
B
C
D

probe purification step to higher R_0t values (>10mol $l^{-1}s$) but since some of the positive signals obtained were barely visible it would appear that a probe of considerably higher specific activity would have been required. The importance of having a low abundance internal control now becomes apparent. Failure to detect positive SV40 hybridization implies that de novo expression of other low abundance mRNAs would remain undetected. For example, a transformed cell specific component comprising 1% of total C138 sequences and hybridizing at an $R_0t_{1/2}$ of 10 mol $l^{-1}s$ could represent more than a hundred individual species. The possibility that much more extensive gene activation occurs in these cells than that observed can therefore not be excluded.

CHAPTER V

GENE ACTIVATION

5.1 Introduction

This Chapter will be concerned with preliminary characterization of the clones selected in Chapter 4 and the mRNA and genomic DNA sequences that they represent. Experiments to test the significance and generality of the phenomenon of gene activation will also be described in Sections 5.6 and 5.7.

5.2 Categorization of the pAG clones

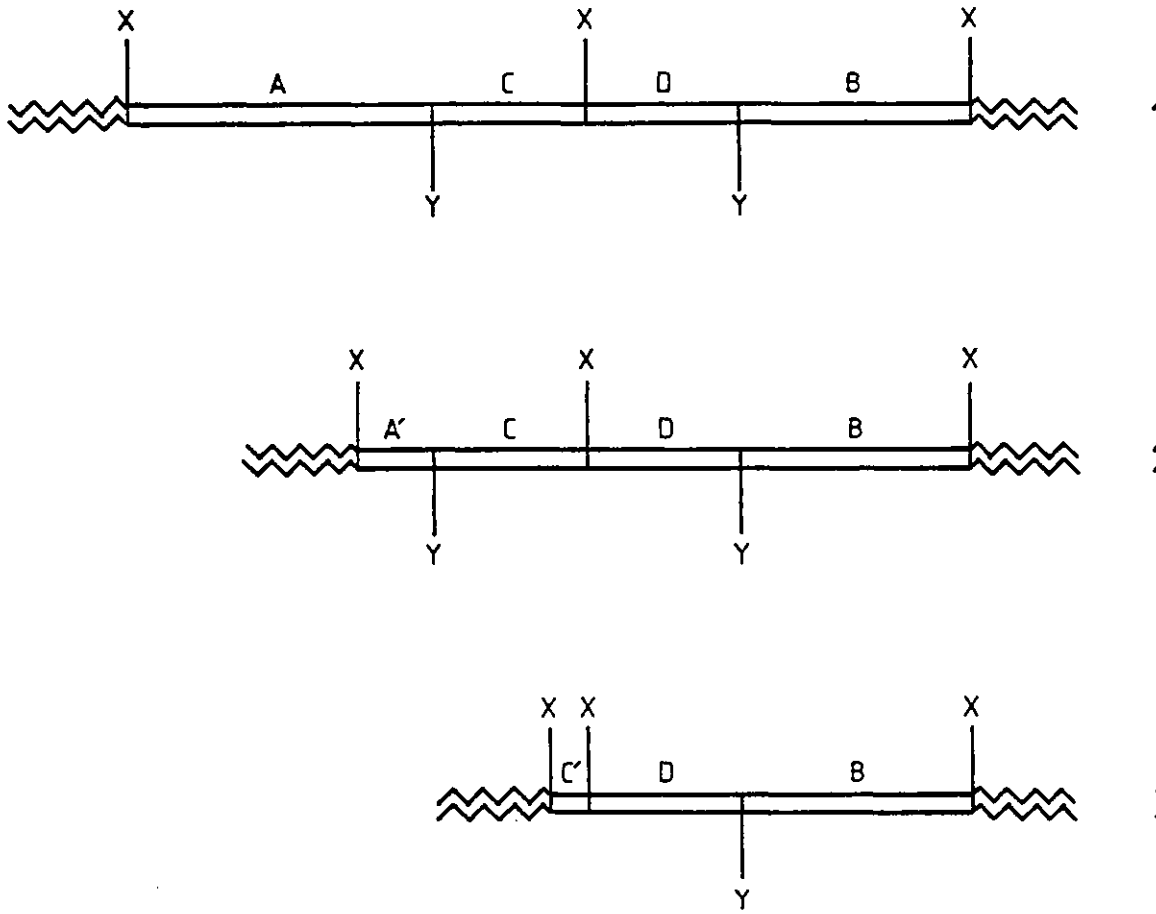
This section will address the problem of dividing the clone collection described in Chapter 4 into groups on the basis of sequence homology. A collection of forty-two clones corresponding to an estimated fourteen average-sized mRNAs should contain several independent isolates of any given sequence. Cross-hybridisation experiments could be used to identify members of the collection related to any particular clone but to perform the entire analysis by systematically screening all the clones in this way would be tedious. To accelerate the process, a preliminary "fingerprinting" step was devised. The rationale is as follows.

If the collection contains several independently isolated cDNA clones, each complementary to the same poly A^+ RNA, they should comprise an overlapping series, since it is improbable that they will all contain the entire mRNA sequence. Considering their mode of construction (see Chapter 3) it is likely that they will share sequences corresponding to the 3' end of the mRNA but may lack sequences from the 5' end due to premature termination of reverse transcription. Fig. 5.1 shows a schematic representation of the expected situation. To establish the relationship between these cloned segments, hydrolysis with restriction endonucleases that cleave several times within each insert would produce common internal fragments which could be identified by gel electrophoresis (see Fig.5.1).

FIGURE 5.1

Figure 5.1 Schematic Representation of a Set of Overlapping cDNA Clones.

Three cDNA clones with overlapping inserts (1, 2 and 3) are cleaved by restriction endonucleases (X and Y). Common fragments (B and D) may then be identified by gel electrophoresis.


X
X
X
A
C
D
B
Y
Y
1
X
X
X
A'
C
D
B
Y
Y
2
X X
X
C'
D
B
Y
3
endo X+Y and gel
1
2
3
A
B
C
D
A'
C'

One of the features of the cloning vector pCD5 (Fig. 5.2) is a facility for performing such a fingerprinting operation. Cleavage with the tetranucleotide recognition sequence endonuclease TaqI will completely separate vector sequences from cDNA inserted between the HindIII and BamI sites (see Fig. 5.2) since the insertion sites are immediately flanked by a ClaI site on one side and a SalI site on the other. These latter enzymes have hexanucleotide recognition sequences, the internal tetranucleotides of which are recognized by TaqI. To increase the probability of cleaving several times within the inserted cDNA, a double digest may be performed using TaqI and HinfI, giving a discrete set of known vector specific fragments in addition to any insert specific bands. The structure of pCD5 is very accurately known, being deduced from the published sequence of pBR322 (Sutcliffe, 1978) using knowledge of the way in which pXf3 was derived from pBR322 (D. Hanahan, personal communication) and the manner in which pCD5 was itself derived from pXf3 (my unpublished observations; see Fig. 5.2 for details of the relationships between these plasmids). Fine structure mapping experiments have confirmed the assignment of the deletion endpoints shown in Fig. 5.2 to be accurate to within 5 nucleotides (S. Campbell-Smith and M.R.D. Scott, unpublished observations). The sizes of fragments expected following clevage of pCD5 with TaqI, HinfI and TaqI + HinfI are shown in Table 5.1; also identified are the vector fragments eliminated during the cloning procedure. Both TaqI and HinfI produce 5' terminal extensions thus allowing end labelling to be performed using reverse transcriptase in the presence of ^{32}P-labelled deoxynucleoside triphosphates. This feature enables the entire procedure to be performed using the small amounts of plasmid DNA obtained from 1ml cultures (see Chapter 2).

FIGURE 5.2

Figure 5.2 Relationships Between the Plasmid Vectors Used in This Study and the Vectors from Which They Are Derived.

The legend to this Figure is as for Fig. 3.14 and the relevant features are described in the text of Section 5.2.

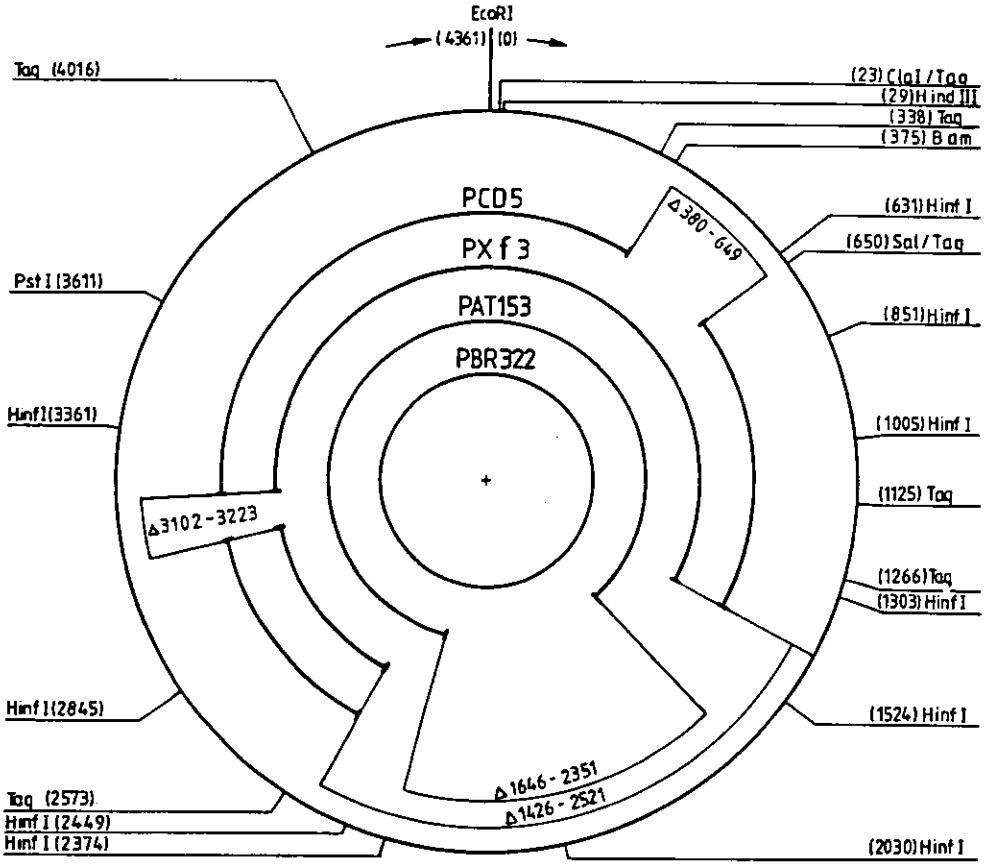

EcoRI
(4361) (0)
Taq (4016)
(23) ClaI / Taq
(29) Hind III
(338) Taq
(375) Bam
(631) Hinf I
(650) Sal / Taq
(851) Hinf I
(1005) Hinf I
(1125) Taq
(1266) Taq
(1303) Hinf I
(1524) Hinf I
(2030) Hinf I
Pst I (3611)
Hinf I (3361)
Hinf I (2845)
Taq (2573)
Hinf I (2449)
Hinf I (2374)
PCD5
PXf3
PAT153
PBR322
Δ380 - 649
Δ3102 - 3223
Δ1646 - 2351
Δ1426 - 2521

TABLE 5.1

Sizes of fragments expected following cleavage of pCD5 with *Taq* I, *Hinf* I or *Taq* I + *Hinf* I

Taq I	*Taq* I + *Hinf* I	*Hinf* I
1325	655	1580*
475	398	447
368	368	398
315*	315*	298
212	272	154
141	201	2877
41	175	
2877	154	
	141	
	120	
	41*	
	37	
	2877	

Sizes are given in base-pairs. Fragments eliminated or altered during cloning are indicated by *. Note that the *Taq* I cleavage site 1266 (see Fig. 5.2) does not cleave efficiently under the conditions used, thus an additional 616 bp *Taq* I fragment and a 261 bp *Taq* I + *Hinf* I fragment are usually present.

Fingerprinting studies on the pAG clones

Small scale (1ml) DNA preparations were performed (see Section 2.4) using forty of the clones isolated in Chapter 4 (two of the putatives, pAG33 and pAG70 were omitted at this stage but were included in the subsequent cross-hybridization analysis). One tenth of each of these DNAs was digested with a quantity of TaqI and HinfI sufficient to produce a limit digest of 0.2ug of pCD5 DNA; the resultant fragments were end labelled using reverse transcriptase and fractionated by electrophoresis on 8% polyacrylamide thin gels as described in Chapter 2. The results of this analysis are shown in Figs. 5.3 and 5.4.

To exemplify the interpretation, part of gel II has been enlarged (see Fig. 5.5). Several insert fragments are resolved in each track and the vector fragments provide convenient internal size markers. Five of the six clones shown share a 240 bp fragment (pAG88, 97, 98, 105 and 109). Two of these also have a 350 bp band (pAG88, 97) and the other three contain slightly smaller but similar sized fragments. The two clones containing both the 240 bp and the 350 bp fragments (pAG88, 97) possess additional fragments of slightly different sizes, 100 bp (pAG88) and 83 bp (pAG97). These data suggest strongly that pAG88, 97, 98, 105 and 109 represent an overlapping set of clones, of the type shown in Fig. 5.1, and allow deduction of the probable structures of these clones. These are shown in Fig. 5.6. Note that the sixth clone, pAG104, has no fragments in common with the other five and is, therefore, presumably unrelated.

Similar logic indicated homology between pAG85 and 104 (which have a 305 bp fragment in common), pAG57, 58, 59 and 69 (with 495 bp and 180 bp fragments in common) and pAG 75 and 82 (with common

FIGURE 5.3

Figure 5.3 Fingerprinting Studies on the pAG Clones - I.

This experiment is described in the text of Section 5.2 The samples electrophoresed are: (a) pAG1, (b) pAG3, (c) pAG10, (d) pAG13, (e) pAG15, (f) pAG18, (g) pAG19, (h) pAG22, (i) pAG23, (j) pAG25, (k) pAG27, (l) pAG28, (m) pAG29, (n) pAG31, (o) pAG37, (p) pAG38, (q) pAG41, (r) pAG47, (s) pAG48, (t) pAG57. 0.1μg aliquots of pCD5 digested with TaqI + HinfI were end labelled and electrophoresed in parallel with the clone digests. These are denoted by M.

pAG 1–57

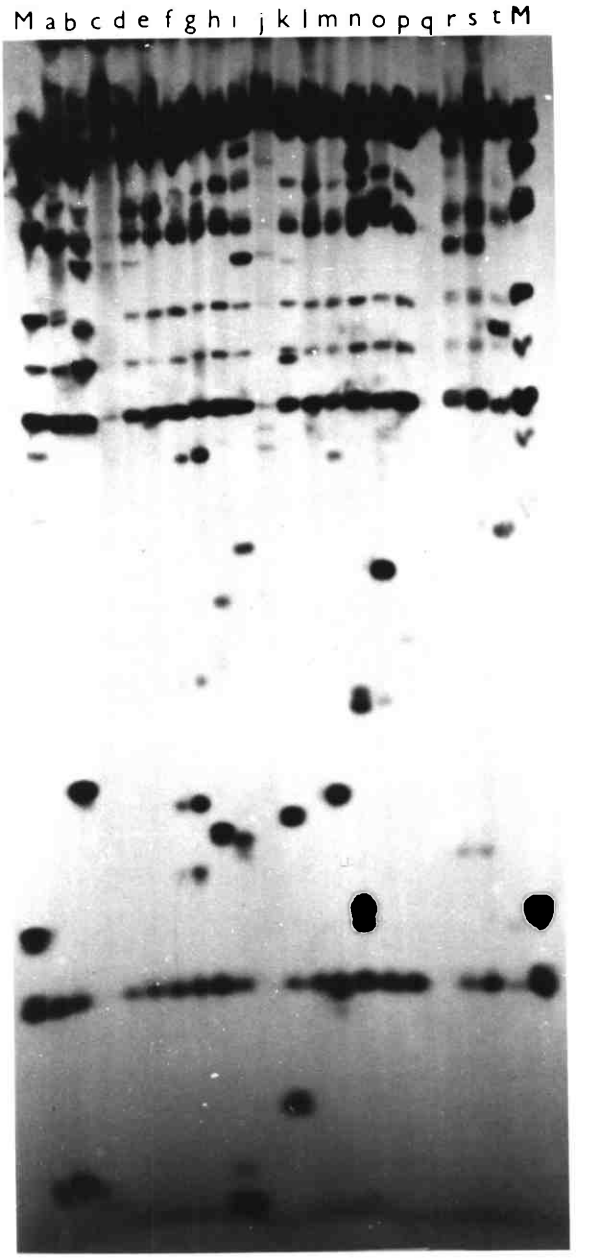


FIGURE 5.4

Figure 5.4 Fingerprinting Studies on the pAG Clones - II.

This experiment is described in the text of Section 5.2 The samples electrophoresed are: (a) pAG58, (b) pAG59, (c) pAG63, (d) pAG64, (e) pAG69, (f) pAG71, (g) pAG74, (h) pAG75, (i) pAG77, (j) pAG81, (k) pAG82, (l) pAG84, (m) pAG85, (n) pAG86, (o) pAG88, (p) pAG97, (q) pAG98, (r) pAG104, (s) pAG105, (t) pAG109. The marker pCD5 fragments are designated by M.

pAG58 – 109

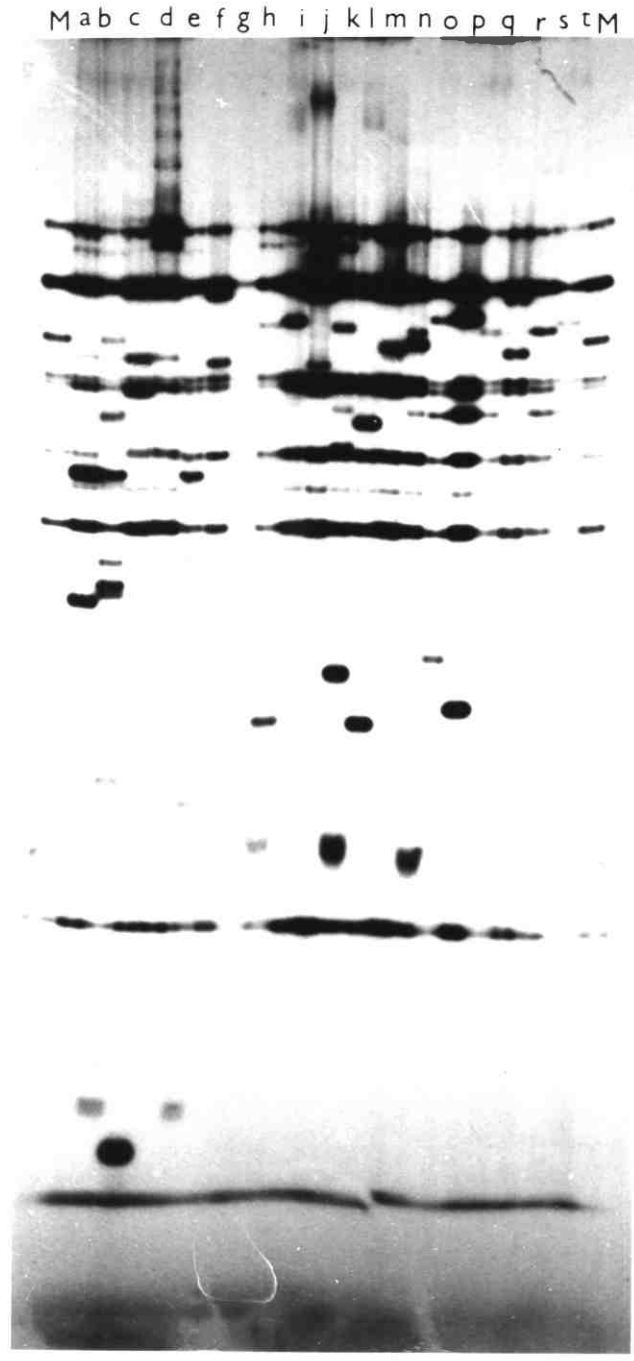


FIGURE 5.5

Figure 5.5 Example of a Fingerprint.

Part of Fig. 5.4 has been enlarged. The samples discussed in the text of Section 5.2 are (a) pAG88, (b) pAG97, (c) pAG98, (d) pAG104, (e) pAG105, (f) pAG109. The marker pCD5 fragments are designated M. The sizes of the fragments in base-pairs are also shown.

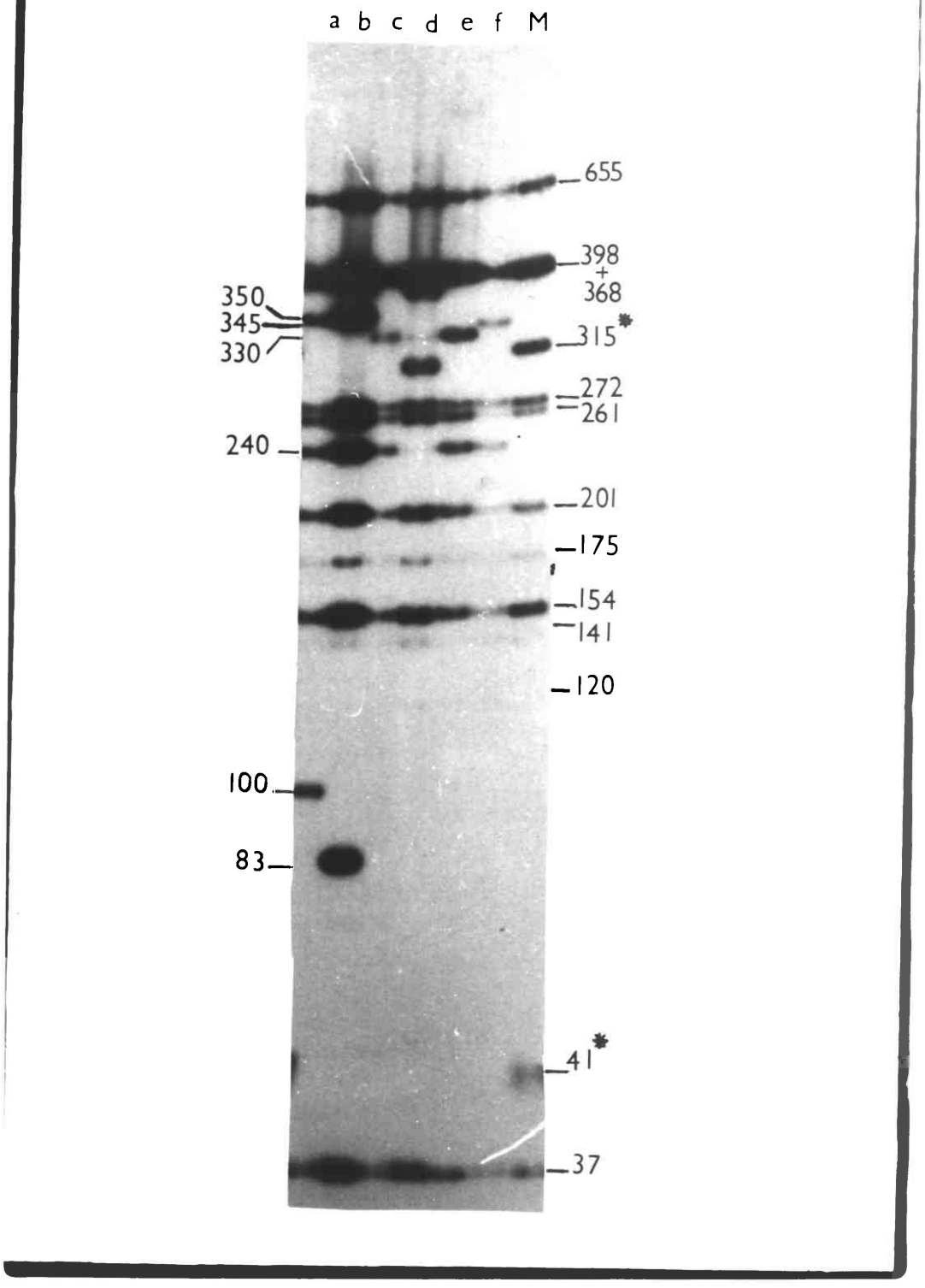
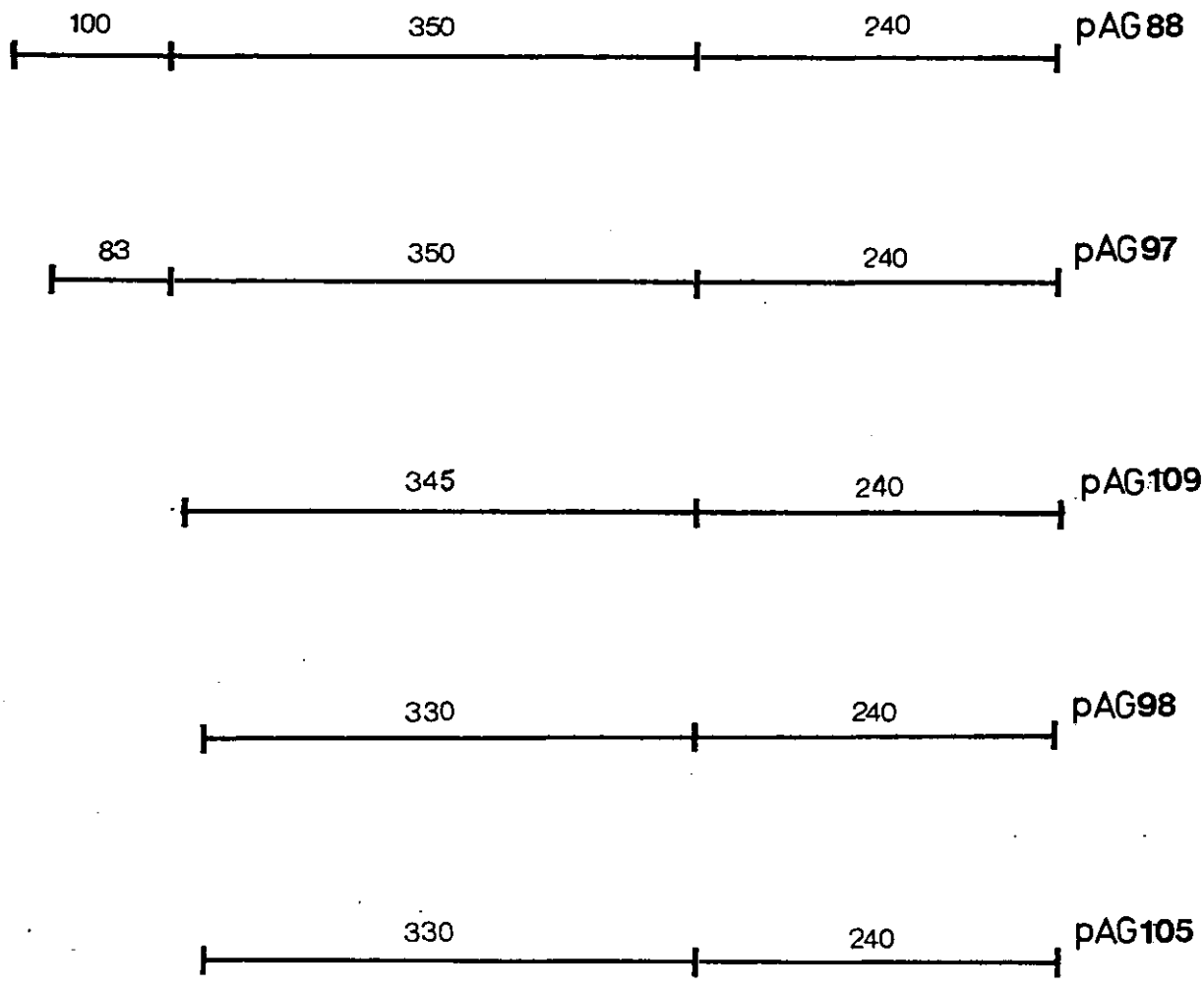

a b c d e f M
655
398
+
368
350
345
330
315*
272
261
240
201
175
154
141
120
100
83
41*
37

FIGURE 5.6

Figure 5.6 Probable Structures of pAG88, 97, 98, 105 and 109.

The symbol (|) denotes a cleavage site for TaqI or HinfI. The sizes of fragments (in base pairs) produced by double digestion with TaqI and HinfI are also shown. Only the insert fragments are depicted.


100
350
240
pAG88
83
350
240
pAG97
345
240
pAG109
330
240
pAG98
330
240
pAG105

fragments of 480 bp, 335 bp and 49 bp). No other groups of clones showing obvious homology were identified in this analysis.

Cross hybridization analysis of the pAG clones

One member of each of the four putative "sets" of clones was labelled with ^{32}P by nick-translation and cross-hybridized to digests of the entire collection (including pAG33 and 70 which were omitted from the fingerprinting study). One tenth of each small-scale plasmid DNA preparation was cleaved with a quantity of EcoRI and SalI sufficient to produce a limit digest of 0.2 ug of pCD5 DNA. An aliquot (1/10) of each digest was removed, end-labelled using reverse transcriptase and electrophoresed in a 2% agarose gel in Tris-acetate buffer which was dried down and autoradiographed, allowing the size of the insert-specific fragments to be determined (Fig. 5.7); further details of the procedure are supplied in Chapter 2). The remainder of each digest was divided into several aliquots, electrophoresed through a 2% agarose gel and transferred to a nitrocellulose filter. Nick translated DNA probes were prepared from one member of each of the putative sets identified by fingerprinting and hybridized to the immobilized fragments as described in Chapter 2. The results of this analysis are shown in Figs. 5.8 (pAG85; Set 1), 5.9 (pAG59; Set 2), 5.10 (pAG82; Set 3) and 5.11 (pAG88; Set 4). All four clones hybridized to insert-specific fragments from several other members of the collection. The conclusions of this study are summarised in Table 5.2. Over half, twenty-four in all, of the prospective clones fell within one or other of these four homology groupings indicating that they comprise the predominant "Sets". From the result of the original colony screen (see Fig. 4.12) it is apparent that, with the exception of pAG3, 10, 25 and 41 which will be discussed later, most of those clones not belonging to the four sets described above gave

FIGURE 5.7

Figure 5.7 Sizing of the cDNA Inserts in the pAG Clones.

This experiment is described in the text of Section 5.2. The samples electrophoresed are identified by the pAG isolate number. The flanking tracks are end labelled DNA markers used to size the insert fragments (S. Campbell-Smith and M.R.D. Scott, unpublished observations). The pAG25 digest was partial in this experiment and in those of Figs. 5.8, 5.9, 5.10, 5.11 and 5.13 but this clone has been studied in detail; these experiments are described in Sections 5.2 and 5.5. The 2.5Kb vector fragment is present in almost all the digests, and in most cases higher molecular weight species are observed produced by partial digestion. The insert specific bands are those represented in the lower part of each gel.

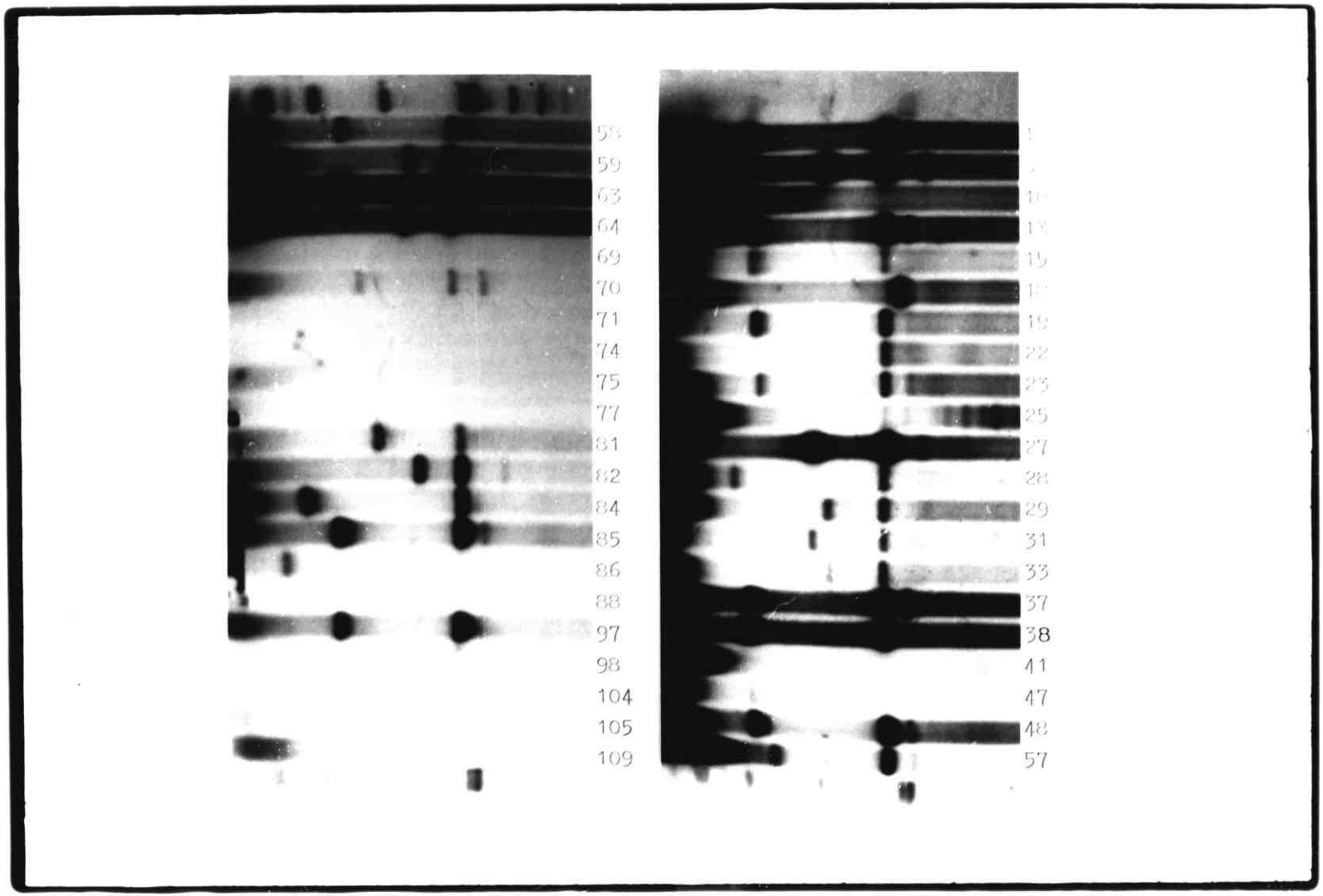

58
59
63
64
69
70
71
74
75
77
81
82
84
85
86
88
97
98
104
105
109
1
19
15
19
22
23
25
27
28
29
31
33
37
38
41
47
48
57

FIGURE 5.8

Figure 5.8 Cross Hybridization Analysis of the pAG Clones - I.

The legend to this Figure is basically as for Fig. 5.7 except that the DNA was not end labelled prior to electrophoresis but instead was transferred to nitrocellulose and hybridized with nick-translated ^{32}P-labelled pAG85 (sp. act. = 10^7 cpm μg^{-1}, the concentration of probe was = 10^6 cpm ml). Further details are supplied in Section 5.2 and in Chapter 2.

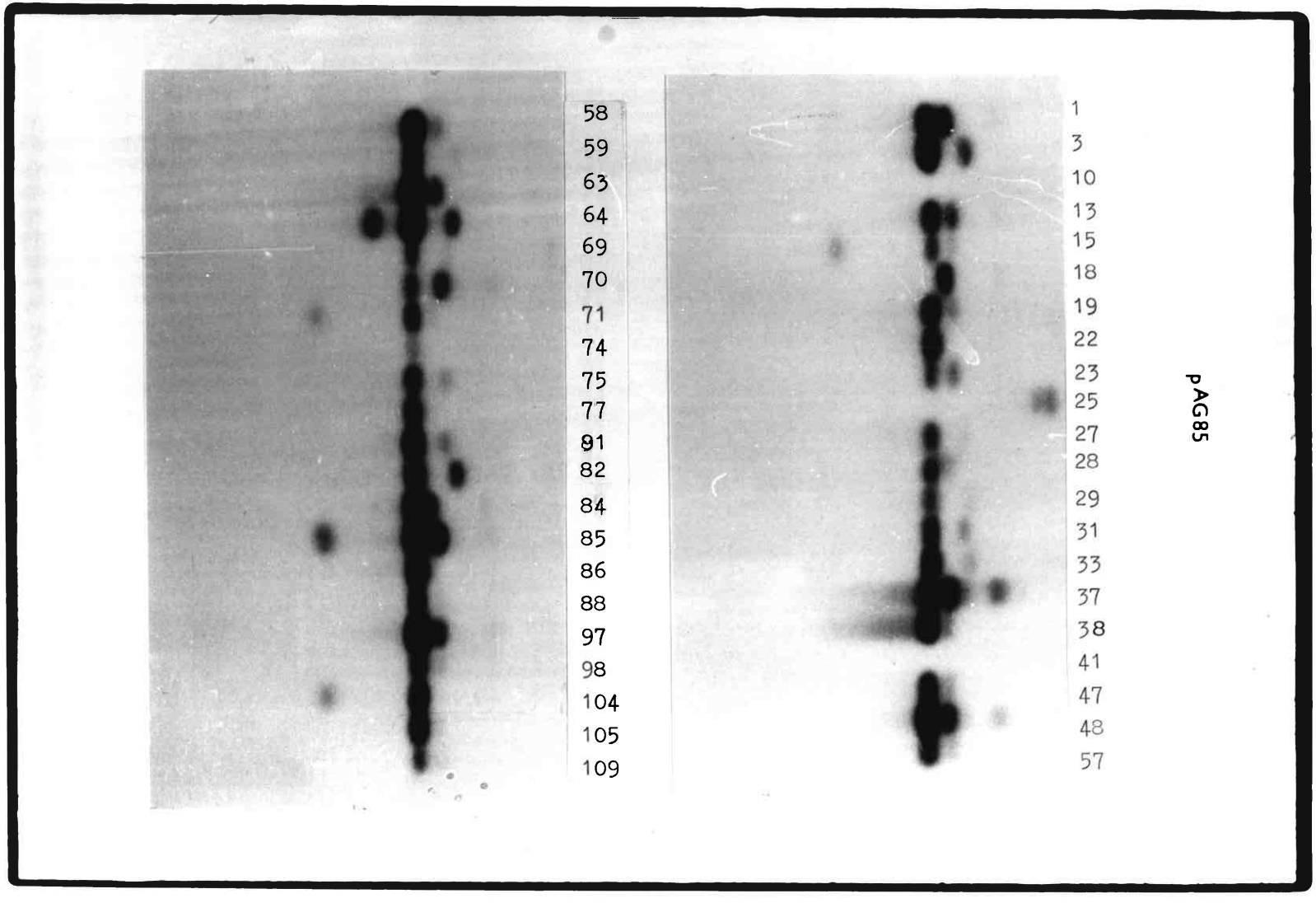

58
59
63
64
69
70
71
74
75
77
81
82
84
85
86
88
97
98
104
105
109
1
3
10
13
15
18
19
22
23
25
27
28
29
31
33
37
38
41
47
48
57
pAG85

FIGURE 5.9

Figure 5.9 Cross Hybridization Analysis of the pAG Clones - II.

The legend to this Figure is as for Fig. 5.8 except that the probe used was pAG59.

pAG 59

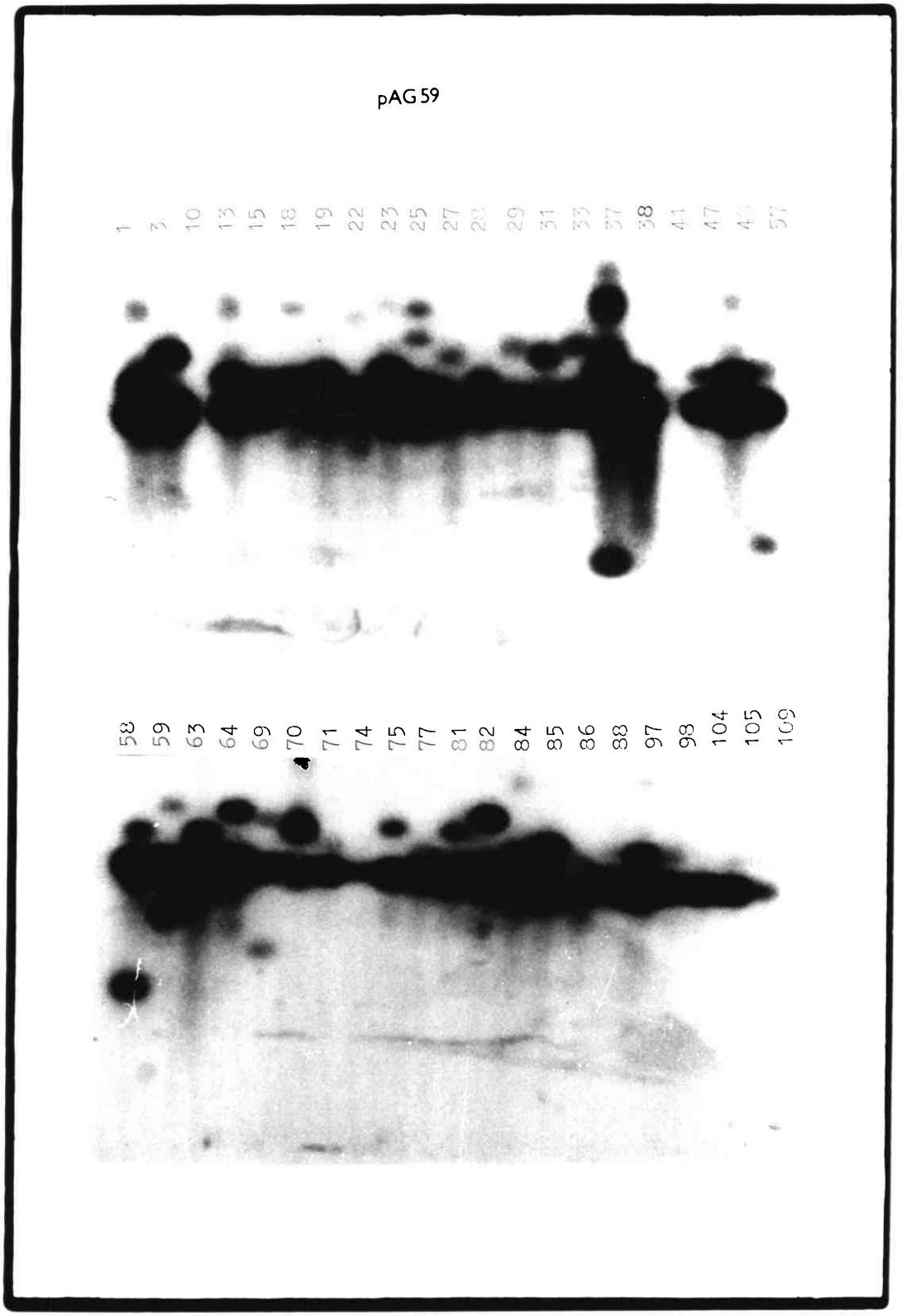

1
3
10
13
15
18
19
22
23
25
27
28
29
31
33
37
38
41
47
43
57
58
59
63
64
69
70
71
74
75
77
81
82
84
85
86
88
97
98
104
105
109

FIGURE 5.10

Figure 5.10 Cross Hybridization Analysis of the pAG Clones - III.

The legend to this Figure is as for Fig. 5.8 except that the probe used was pAG82.

58
59
63
64
69
70
71
74
75
77
81
82
84
85
86
88
97
98
104
105
109
1
3
10
13
15
18
19
22
23
25
27
28
29
31
33
37
38
41
47
48
57
pAG82

FIGURE 5.11

Figure 5.11 Cross Hybridization Analysis of the pAG Clones - IV

The legend to this Figure is as for Fig. 5.8 except that the probe used was pAG88.

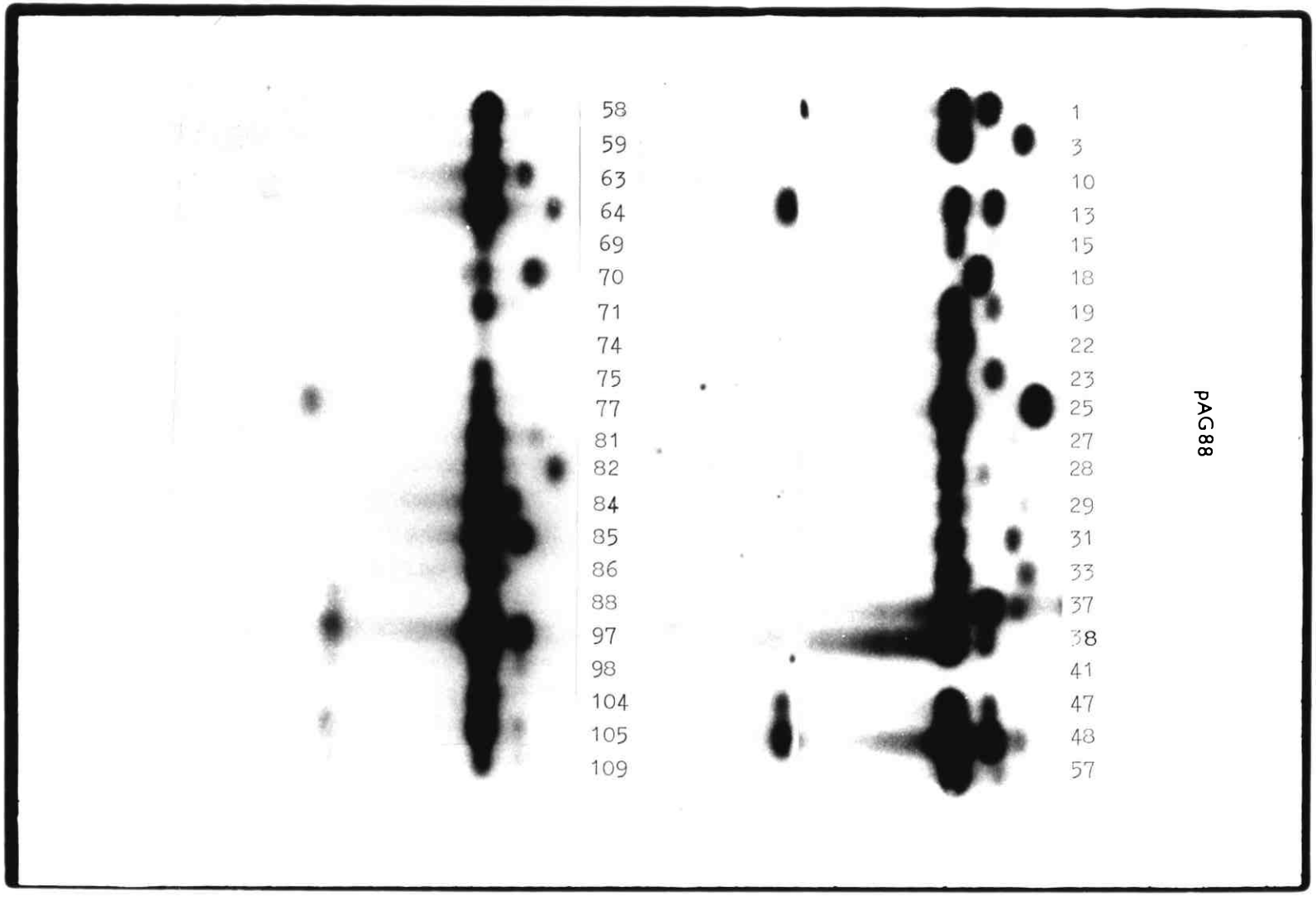

58
59
63
64
69
70
71
74
75
77
81
82
84
85
86
88
97
98
104
105
109
1
3
10
13
15
18
19
22
23
25
27
28
29
31
33
37
38
41
47
48
57
PAG88

TABLE 5.2

Summary of cross hybridization data.

Designation	Set 1	Set 2	Set 3	Set 4
Prototype	pAG64	pAG59	pAG82	pAG88
Length insert	1.55Kb	1.8Kb	1.8Kb	0.67Kb
pAG clones related by cross-hybridization	15, 64, 71, 85, 104.	37, 57, 58, 59, 69.	1, 31, 75, 82,	13, 47, 48, 77, 88, 97, 98, 105, 109.
Total	5	5	4	9

The "prototype" is the clone within each "Set" which contains the longest insert. Insert sizes were determined in the experiment of Figure 5.7.

dubious autoradiographic responses. Attention will therefore be concentrated entirely upon the four sets defined by the fingerprinting and cross hybridization experiments and those other clones that gave a clear signal in the colony screening experiment of Fig. 4.12; the remainder are thus poorly characterized.

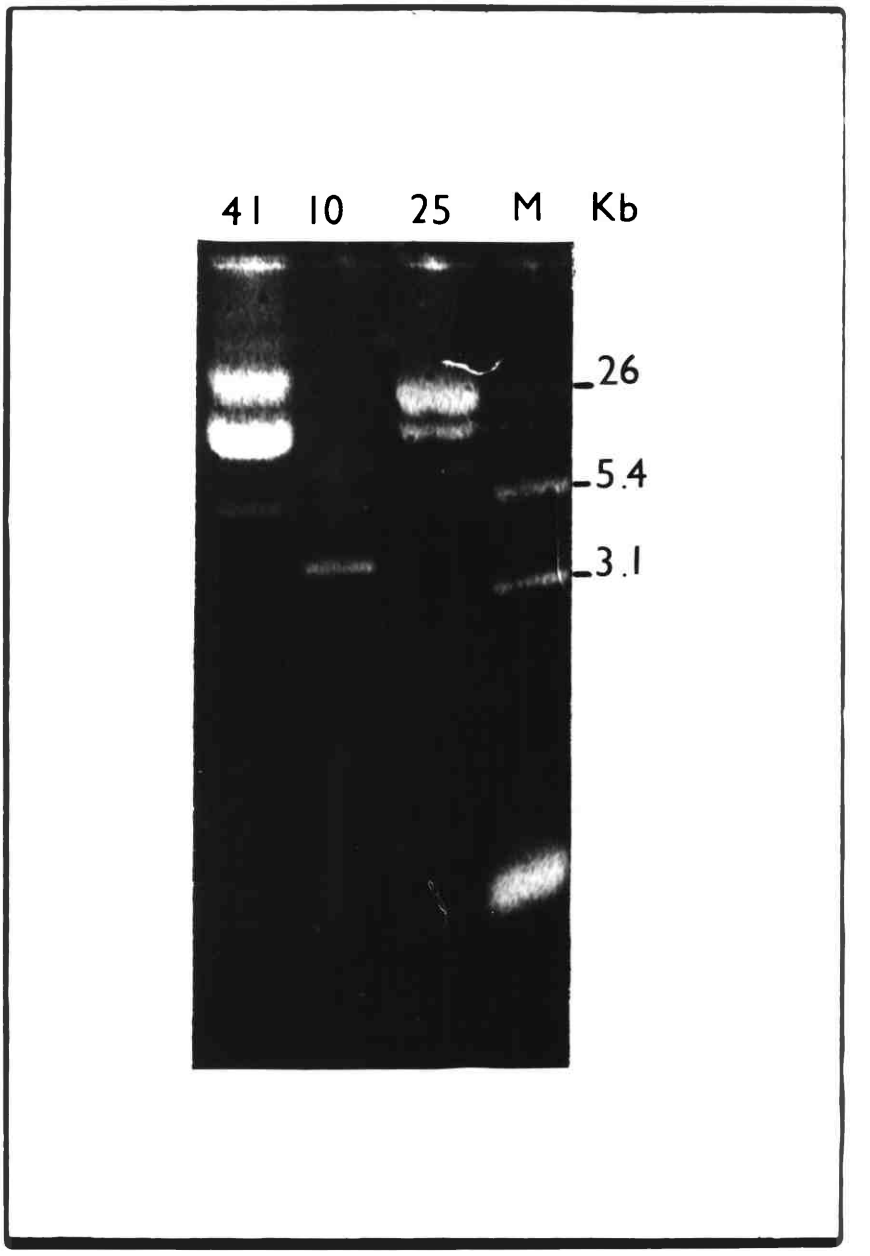
An unusual clone, pAG25

Two clones, pAG25 and pAG41 gave particularly strong responses in the original colony screen (Fig. 4.12). Unexpectedly, however, the fingerprinting study of Fig. 5.3 does not show a clear pattern of bands for either clone, although in the case of pAG25 an unusually large number of faint bands can be resolved. This is a result that could arise if the plasmid contains a very large insert. To test this hypothesis pAG25, pAG41 and another clone, pAG10 (which also gave a definite response in the colony screen but failed to show any fragments in the fingerprint) were analyzed by gel electrophoresis of the native plasmid DNA (see Fig. 5.12). Two of the clones, pAG25 and pAG41, contain a major supercoiled species of at least 10 Kb, indicating an insert size of $>$7.5Kb. Whereas pAG25 only contains the expected two bands of supercoiled and open circular DNA, pAG41 also contains a number of bands of higher molecular weight and one minor species of lower molecular weight. This kind of pattern implies an inherent instability of the plasmid during propagation so this clone was not subjected to further scrutiny. The third clone, pAG10, contains an average sized plasmid of 3 to 4Kb, but due to persistent difficulties in preparing appreciable quantities of the plasmid DNA further analysis was not performed. No problems of instability or yield were encountered with pAG25 and the extreme size of the insert (9.5Kb, see Section 5.5) in itself begs further investigation.

FIGURE 5.12

Figure 5.12 Gel Electrophoresis of Native Plasmid DNA From pAG10, 25 and 41.

Aliquots (0.1 - 0.5µg) of pAG10, 25 and 41 DNA were electrophoresed together with supercoiled plasmid DNA markers. The gel was 0.8% agarose, the clones are identified by their pAG isolate numbers and the sizes in Kb of the supercoiled DNA markers electrophoresed in track M are also shown.


41
10
25
M
Kb
26
5.4
3.1

Cross hybridization of pAG25 to Set 3 and Set 4

When cross hybridization to the clone collection as described in Chapter 2, homology was indicated to all members of both Set 3 and Set 4, plus an additional clone, pAG28 (see Fig. 5.13). The probable origin of pAG25 and its relationship to the members of Set 3 and Set 4 have been investigated in detail and will be discussed in Sections 5.5 and 6.2.

A control clone, pAG3

Comparison of the signal obtained with pAG3 (see Fig. 4.12) and the corresponding colony on the master plate showed that they were of a different size and shape, indicating that the original signal was an artifact; this has been confirmed by RNA gel transfer hybridization analysis (see below). However, this clone serves as a useful control in such transfer hybridization experiments.

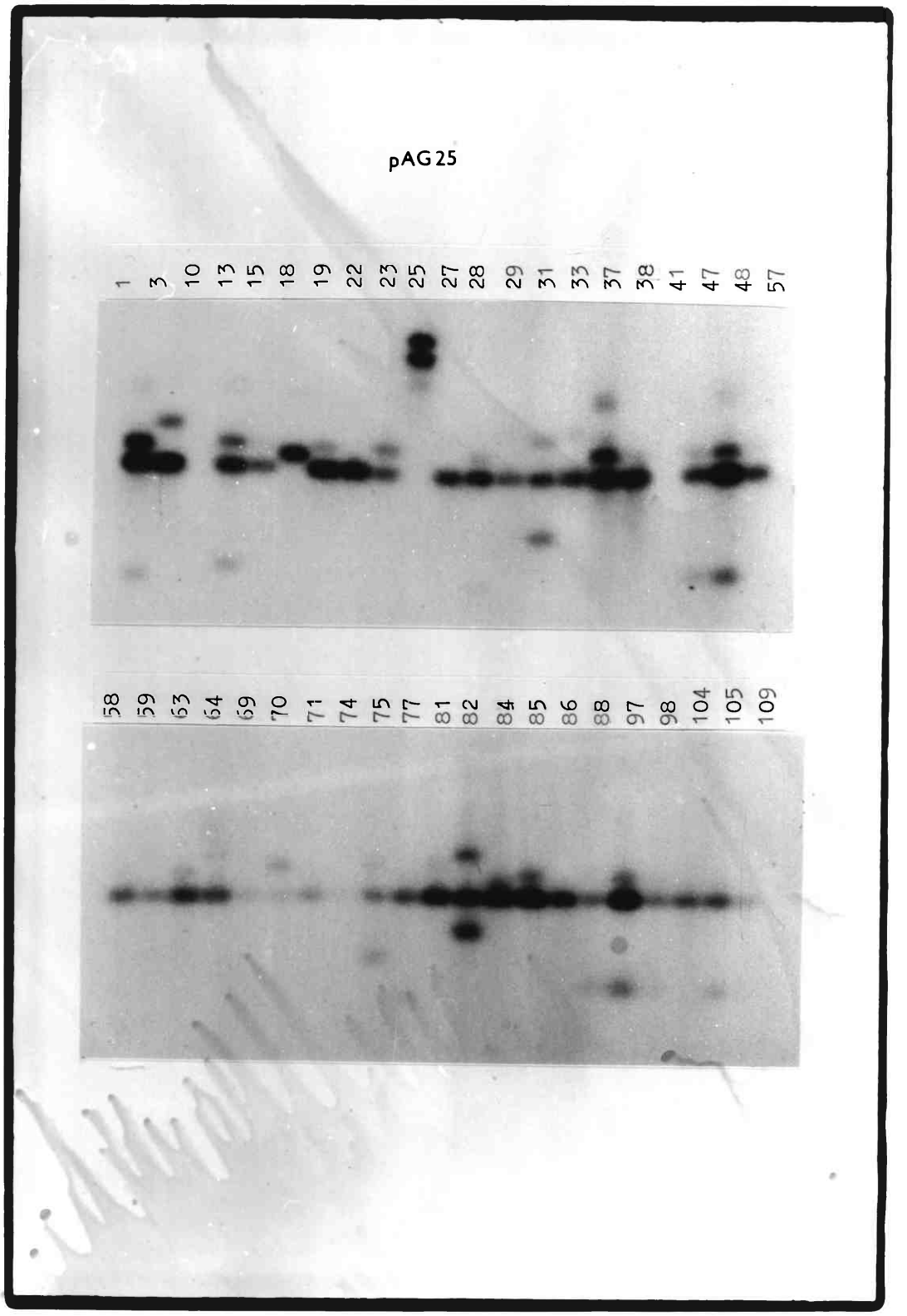
5.3 Transcripts homologous to the SV3T3 C138 specific clones

The cloning and screening protocol was designed to isolate cDNA clones corresponding to mRNAs which, while of relatively low abundance in SV3T3 C138, are present at a higher level in this transformed line than in the parental Balb/c 3T3 line. To confirm that the procedure was successful, and to effect a preliminary characterization of the relevant mRNAs, RNA gel transfer hybridization experiments (Alwine et al., 1977; Thomas, 1980) were performed. During these experiments, amounts of plasmid DNA size markers (homologous to the vector moiety of the clones) equivalent to 0.1% of total poly A^+ were run together with the RNA samples, allowing approximate quantitation of the levels of the relevant RNAs. However, no attempt was made to derive accurate quantitations from these experiments.

FIGURE 5.13

Figure 5.13 Cross-Hybridization Analysis of the pAG Clones - V.

The legend to this Figure is as for Fig. 5.8, except that the probe used was pAG25.


pAG25
1 3 10 13 15 18 19 22 23 25 27 28 29 31 33 37 38 41 47 48 57
58 59 63 64 69 70 71 74 75 77 81 82 84 85 86 88 97 98 104 105 109

Transcripts complementary to Set 1 clones

Fig. 5.14 shows the result of an RNA gel transfer hybridization experiment performed using equivalent amounts (1μg) of poly $^{+}$ RNA from SV3T3 C138 and Balb/c 3T3 cells. The same preparations of RNA as were used in the experiments of Chapters 3 and 4 were used for all the experiments described in this Section, to obviate possible artifacts caused by variations between different RNA preparations made from the same cell line. One filter was hybridized with ^{32}P-labelled nick-translated pAG15 (Set 1), the other with the control clone, pAG3. pAG15 hybridized to a band of 1.85Kb, apparently present at very low levels in normal cells, but at greatly elevated levels (at least ten-fold increase) in the transformed line. By comparison with the plasmid DNA size markers the level of expression in the transformed line is approximately as expected ($\lesssim$ 0.1%). pAG3 hybridized to a 1.6Kb band of similar abundance but common to both cell lines, demonstrating that the amounts of RNA electrophoresed were indeed equivalent. For an obscure technical reason, two additional transformation-specific bands of 1.0Kb and 0.9Kb were not observed in this experiment using pAG15 but have been observed in subsequent experiments using the same RNA preparations; these experiments will be described in Sections 5.6.

Transcripts complementary to Set 2 Clones

The hybridization of ^{32}P-labelled pAG37 (Set 2) to SV3T3 C138 and Balb/c 3T3 poly A^{+} RNA is shown in Fig. 5.15. This clone gave a more complex pattern of hybridization than pAG15. One band, of 3.2Kb, is particularly prominent, being undetectable in Balb/c 3T3 mRNA but present at a level corresponding to 0.1% of total poly A^{+} RNA in the transformed line. Two less intense ands, of 6.5Kb and 7.6Kb, are also specific to the C138 RNA, while a fourth transcript of 1.8Kb is observed at the same level in both cell lines.

FIGURE 5.14

Figure 5.14 Transcripts Complementary to pAG64 and pAG3.

1μg aliquots of polyA$^+$ RNA extracted from SV3T3 C138 and Balb/c 3T3 cells were electrophoresed through 1% formaldehyde slab gels, transferred to nitrocellulose and hybridized with ^{32}P-labelled nick-translated plasmids. Marker tracks (M) contained amounts of plasmid fragments equivalent to 0.1% of polyA$^+$ RNA. The probe used in each experiment is designated; hybridization contained 10-25ng ml^{-1} of ^{32}P-labelled DNA at a sp. act. of 5 x 10^7 to 2 x 10^8 cpm μg^{-1}.

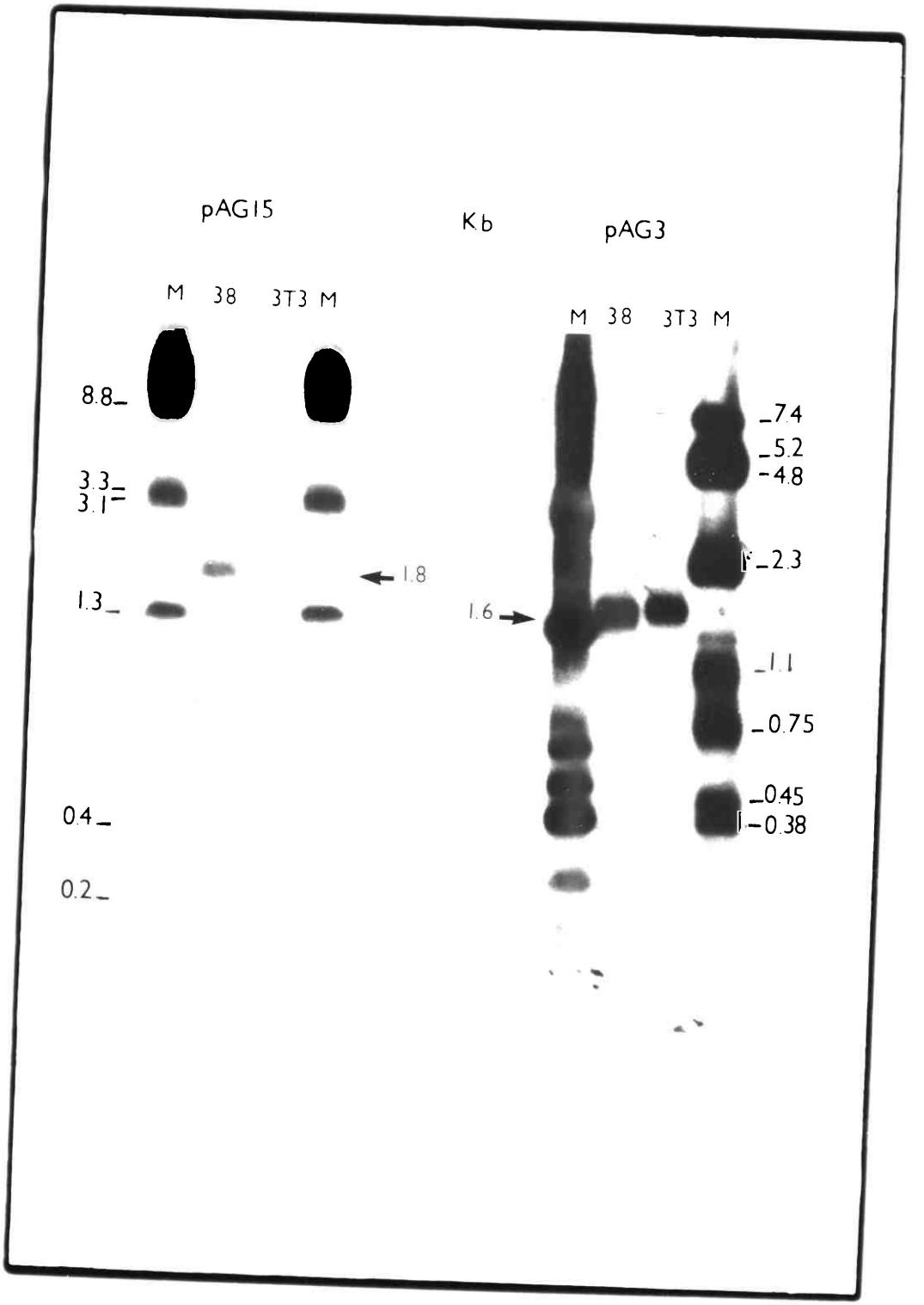
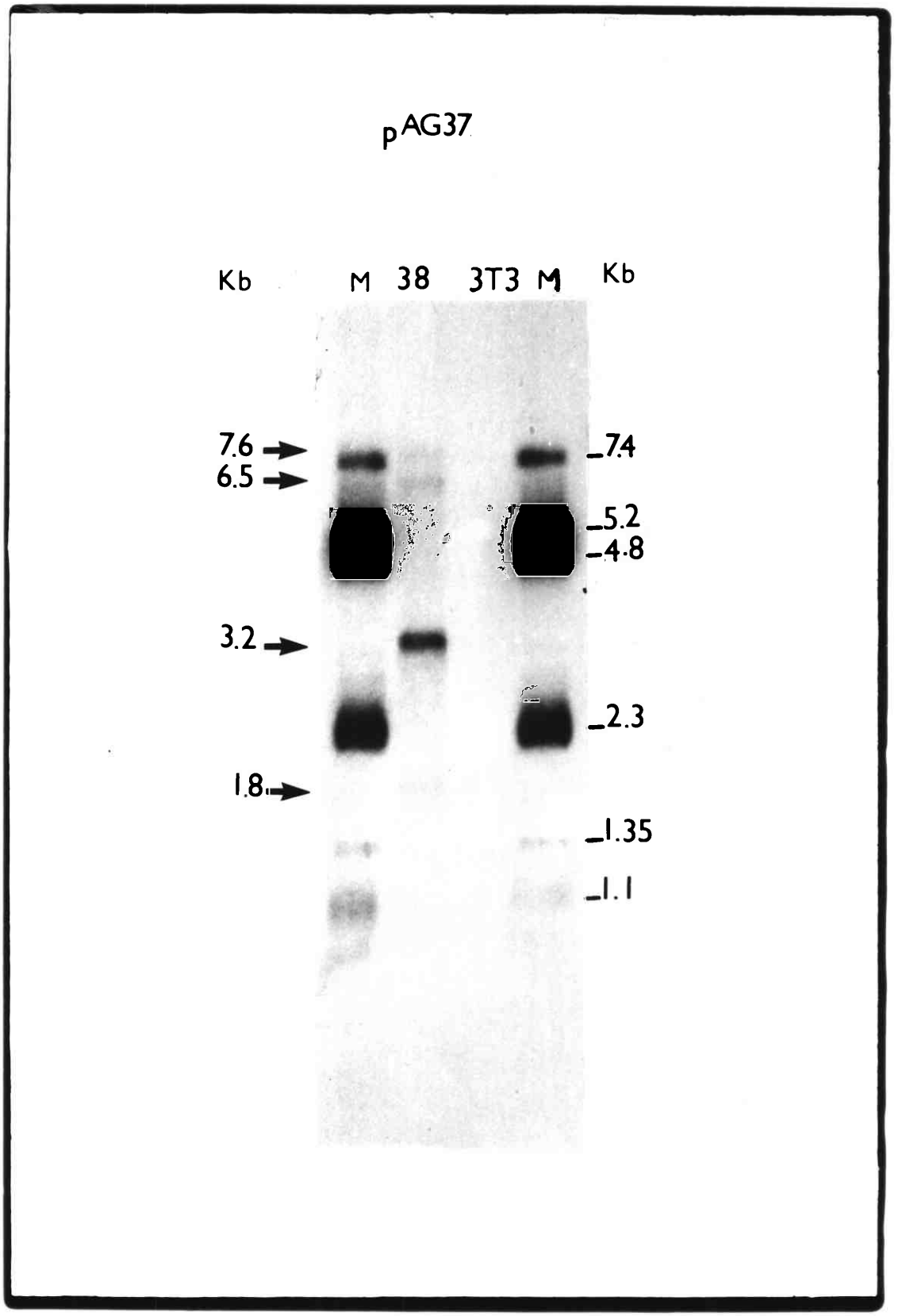

pAG15
Kb
pAG3
M
38
3T3
M
8.8
3.3
3.1
1.3
0.4
0.2
1.8
1.6
7.4
5.2
4.8
2.3
1.1
0.75
0.45
0.38

FIGURE 5.15

Figure 5.15 Transcripts Complementary to pAG59.

The legend to this Figure is as for Fig. 5.14 except that the probe used was pAG59.


pAG37
Kb
M
38
3T3
M
Kb
7.6
6.5
3.2
1.8
7.4
5.2
4.8
2.3
1.35
1.1

Transcripts complementary to Clones of Sets 3 and 4

In Fig. 5.16, pAG1 (Set 3) and pAG47 (Set 4) have been used to probe transformed and normal cell mRNA. pAG47 hybridized to one major species of 0.8Kb, common to both cell lines. Comparison with the semi-quantitative DNA markers indicates that this transcript is present at high levels, considerably in excess of 0.1% of total poly A^+ RNA. In SV3T3 C138, several additional transformation- specific bands are conspicuous, notably one of 3.0Kb and another of 5.6Kb. These transcripts are not evident in normal mRNA and each corresponds to 0.1% of total poly A^+ RNA, suggesting that they probably contain the transformation-specific sequences. pAG1 shows a similar pattern of hybridization, this time to an abundant common transcript of 1.8Kb with predominant C138 specific bands of 3.3Kb, 3.5Kb, 6.3Kb and 7.0Kb.

The pattern of hybridization seen with members of Sets 3 and 4 raises an important point. Since the clones hybridize to common abundant transcripts but were selected using a probe specific for C138 mRNA, they presumably contain sequences unique to the transformed cell line in addition to those shared with other more abundant mRNAs found also in normal cells. A hypothesis has been devised to explain these observations, and is diagrammed in Fig. 5.17. In this example, it is proposed that two distinct but related mRNAs exist. One species, designated B, has specific 3' terminal sequences but shares a sequence toward the 5' end with another transcript, species A. A cDNA probe prepared from this mixture should comprise an overlapping series of partial copies of each mRNA. Selection of this probe by hybridization to species A will remove all reverse transcripts containing A-complementary sequences, including those derived from species B that have extended into the A-complementary region. Only those partial

FIGURE 5.16

Figure 5.16 Transcripts Complementary to pAG1 and pAG47.

The legend to this Figure is as for Fig. 5.14 except that the probes used were pAG1 and pAG47.

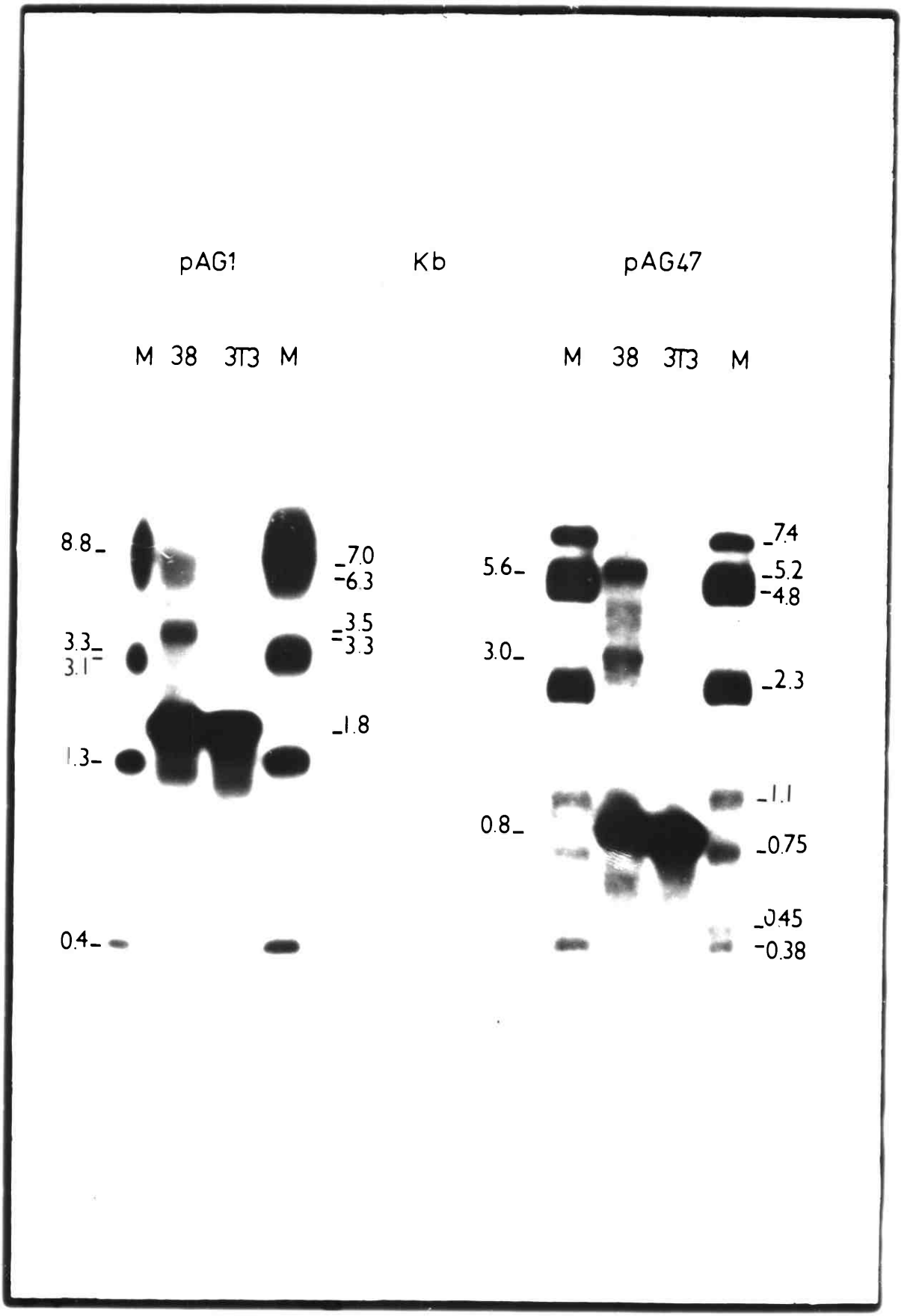
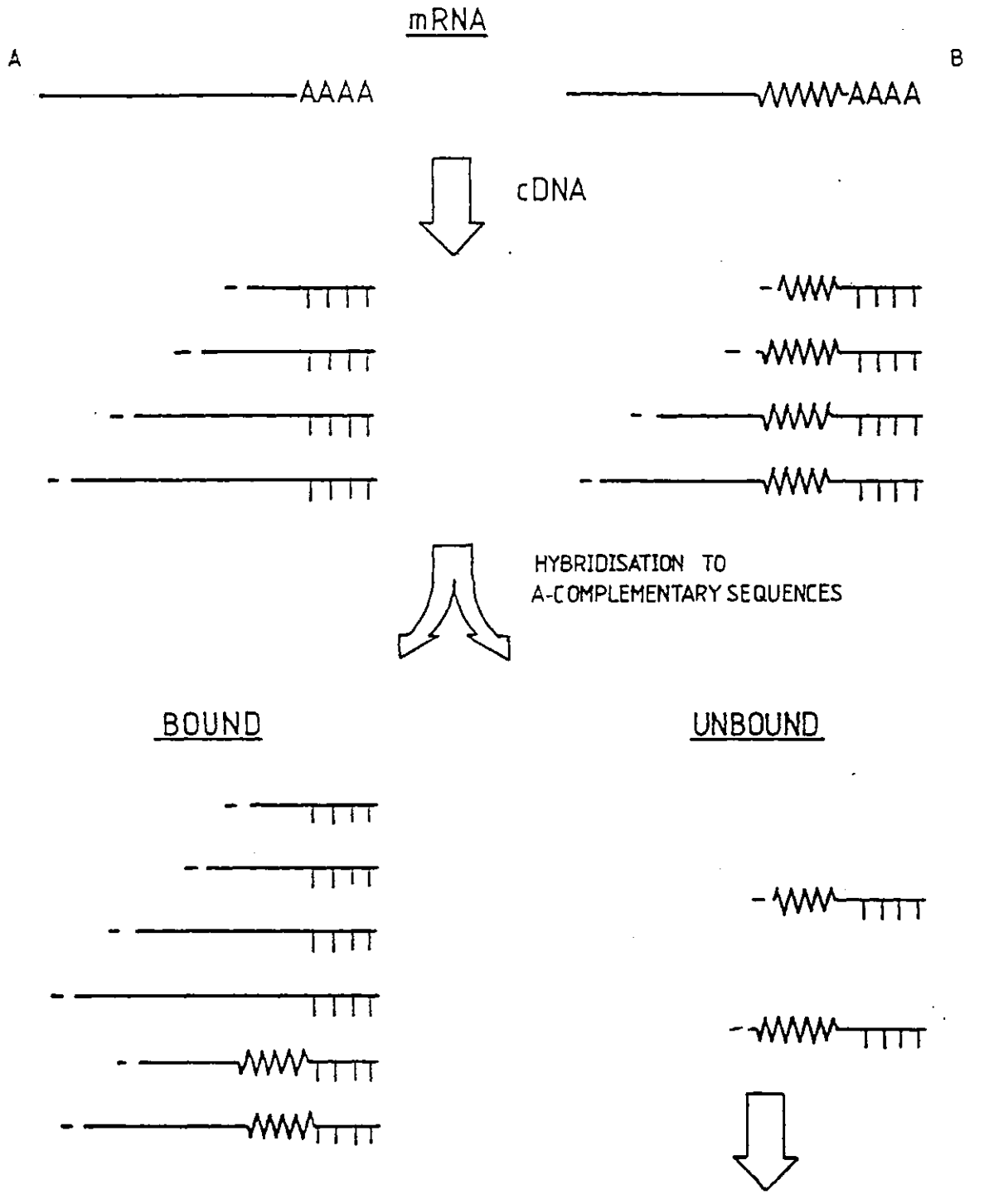

pAG1
Kb
pAG47
M 38 3T3 M
M 38 3T3 M
8.8
3.3
3.1
1.3
0.4
7.0
6.3
3.5
3.3
1.8
5.6
3.0
0.8
7.4
5.2
4.8
2.3
1.1
0.75
0.45
0.38

FIGURE 5.17

Figure 5.17 Hypothesis to Explain the Selection of Clones From Set 3 and Set 4.

This scheme is described in the text of Section 5.3.

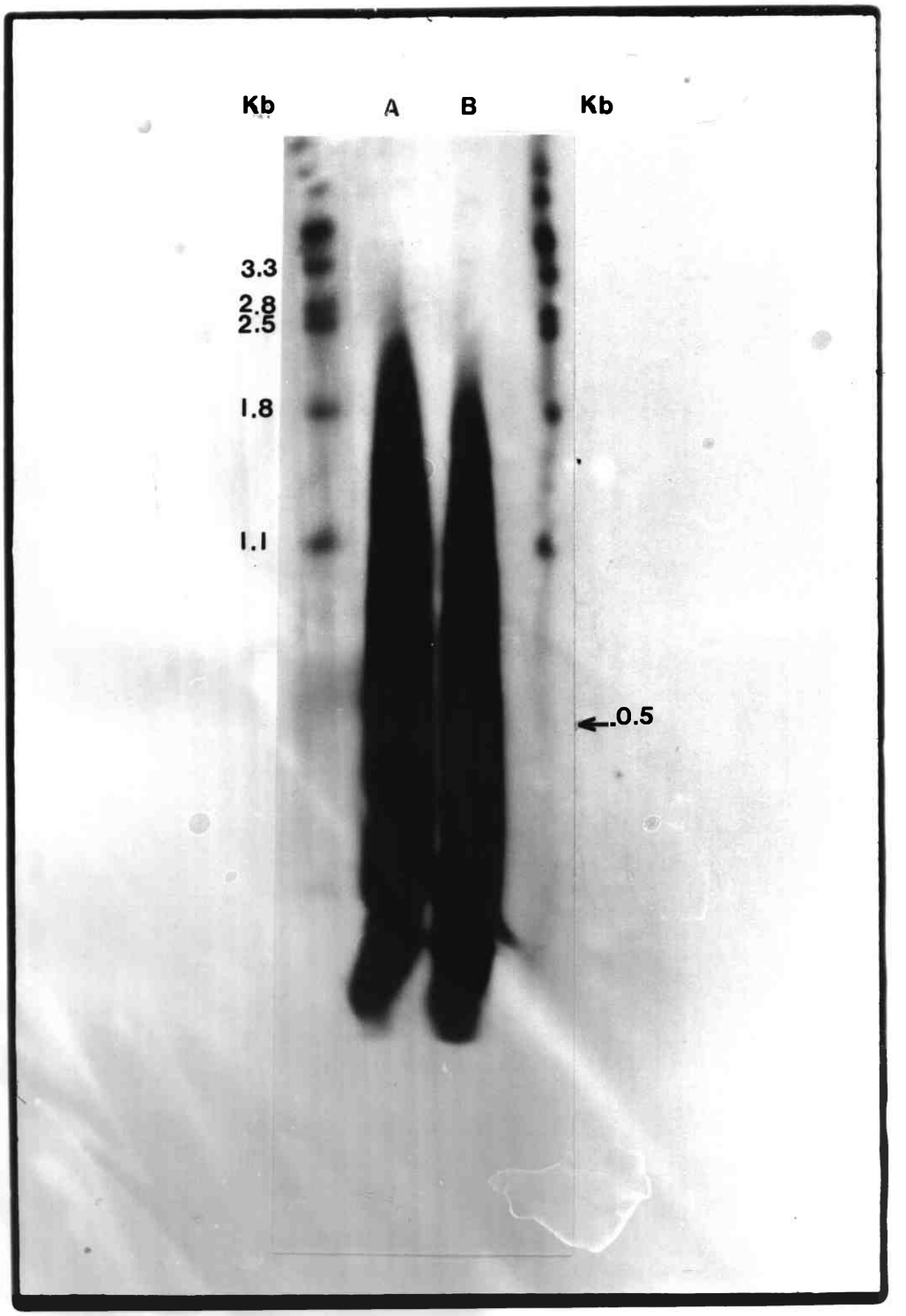

mRNA
A
AAAA
B
AAAA
cDNA
HYBRIDISATION TO
A-COMPLEMENTARY SEQUENCES
BOUND
UNBOUND
PROBE NOW LACKS A-COMPLEMENTARY
SEQUENCES AND WILL DETECT CLONES
OF B-TYPE TRANSCRIPTS BUT NOT
A-TYPE

reverse transcripts which have not extended into the A-complementary region will remain in the probe which will now be B-specific. Clones derived from species B will hybridize to this probe, even though they may contain sequences from the A-complementary region (see Fig. 5.17). Therefore the prediction is that one or more of the SV3T3 C138 specific transcripts observed in hybridization experiments using members of Sets 3 and 4 should contain a region specific to C138 RNA linked to sequences homologous to the common abundant mRNA. Logic dictates, considering the way in which the experiment was performed, that the region of homology must be remote from the 3' end of the transcript, to enable a specific probe to be prepared in the manner described in Fig. 5.17. Thus representative clones of the C138-specific transcripts would contain both the transformation-specific region and the region of homology to common abundant mRNA. Several lines of evidence are consistent with this hypothesis. The transformed cell-specific bands are observed in SV3T3 C138 mRNA at levels comparable to those of the bands homologous to clones of Sets 1 and 2. Accordingly the frequency with which Set 3 and 4 clones were identified was similar to those obtained for members of Sets 1 and 2 (see Table 5.1). Finally, Fig. 5.18 shows a denaturing agarose gel analysis of the highly-labelled SV3T3 C138 cDNA probe used for the selection and colony screening experiments of Chapter 4. Comparison with end-labelled λ-DNA markers clearly demonstrates that the majority of the cDNA is extremely short, with an average size of ~ 0.5Kb. This is not surprising since the conditions of labelling were designed to maximise the specific activity of the probe, hence the concentration of labelled nucleotide was suboptimal for production of full-length reverse transcripts.

FIGURE 5.18

Figure 5.18 Denaturing Agarose Gel Electrophoresis of Purified Probe

Aliquots (10^5 cpm) of the C138 cDNA probe used for the experiments described in the text of Section 4.4 were electrophoresed through a 2% alkaline agarose gel which was dried down and autoradiographed. The flanking tracks contain HindIII fragments labelled with ^{32}P using reverse transcriptase (see Section 2.6) which were electrophoresed in parallel to provide size markers. Tracks A and B contain C138 cDNA and 3T3-depleted C138 cDNA respectively.


Kb
A
B
Kb
3.3
2.8
2.5
1.8
1.1
0.5

Transcripts complementary to pAG25

Fig. 5.19 shows the result of an RNA gel transfer hybridization using ^{32}P-labelled nick-translated pAG25. The pattern of hybridization obtained is extremely complex. Several abundant species of 2.8Kb, 1.8Kb, 1.7Kb, 1.35Kb, 1.2Kb, 1.1Kb and 0.8Kb are common to both cell lines. Since pAG25 cross hybridizes to members of Sets 3 and 4 it seems likely that the 1.8Kb and 0.8Kb bands correspond to the common abundant transcripts described previously for Sets 3 and 4 respectively. Much transformed cell specific material is also evident, particularly a band at 3.6Kb and several larger than 5.0Kb.

5.4 DNA gel transfer hybridization analysis

Some preliminary genomic "blotting" studies were performed using members of the four clone Sets. To approximately quantitate the genomic fragments, plasmid size markers were included on the gels in amounts equivalent to approximately 100, 10 and 1 copies per cell.

DNA sequences complementary to clones of Set 1

When pAG64 (Set 1) was labelled with ^{32}P by nick translation and hybridized to restriction enzyme digests of SV3T3 C138 genomic DNA, the result presented in Fig. 5.20 was obtained. Hybridization is observed to a sequence represented many times in genomic DNA, evidenced by a heterogeneous "smear" of fragments following cleavage with EcoRI and BstI. This pattern of hybridization is diagnostic of the presence in this clone of a dispersed repeated sequence element (Jelinek et al., 1980; Crampton et al., 1981). These elements are families of short, repeated sequences that punctuate eukaryotic DNA. Following cleavage with EcoRI, two prominent bands of 5.8Kb and 3.6Kb are also observed with an estimated abundance of between one and ten copies per cell. No obvious bands are seen following BstI cleavage,

FIGURE 5.19

Figure 5.19 Transcripts Complementary to pAG25.

The legend to this Figure is as for Fig. 5.14 except that the probe used was pAG25.

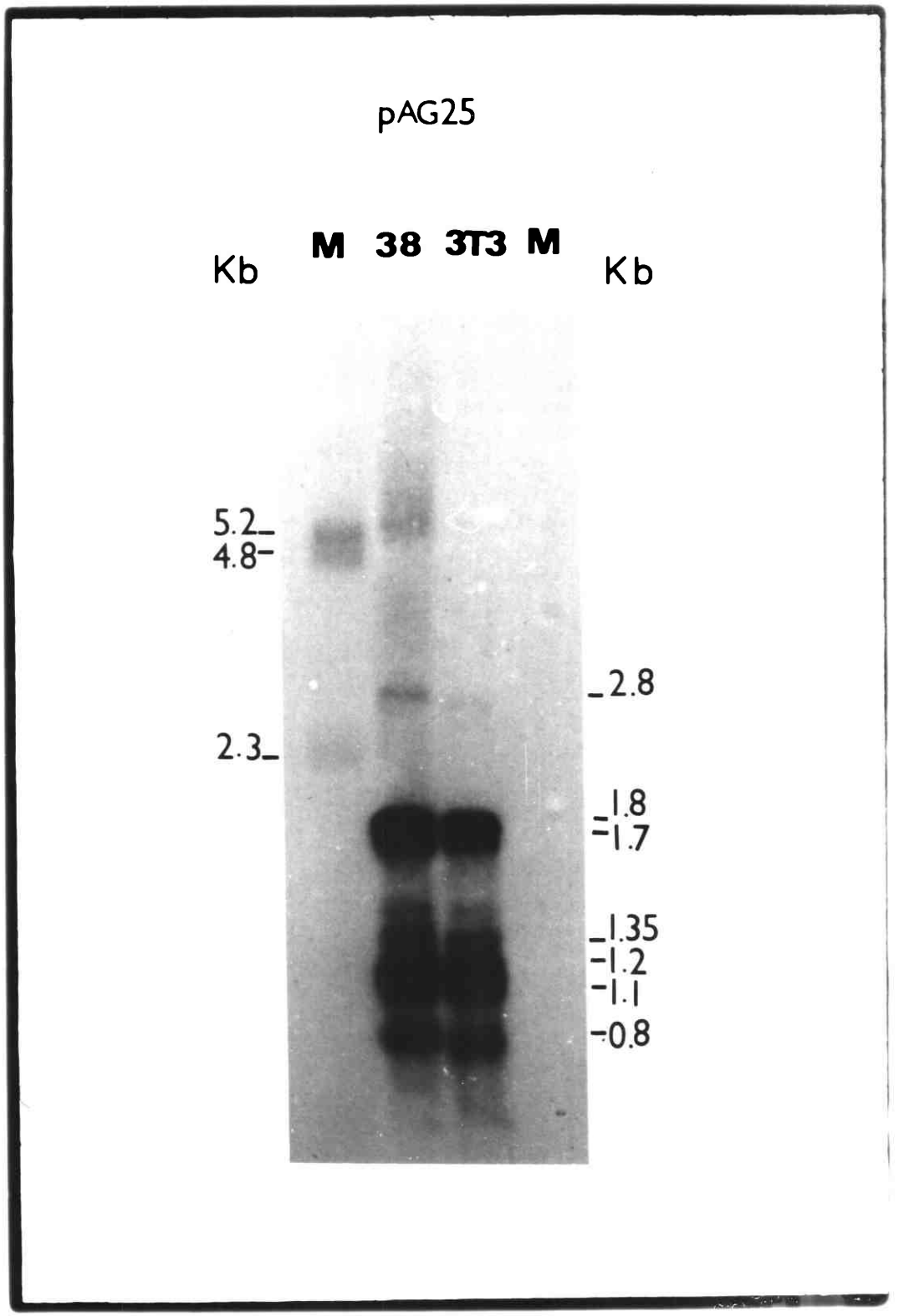
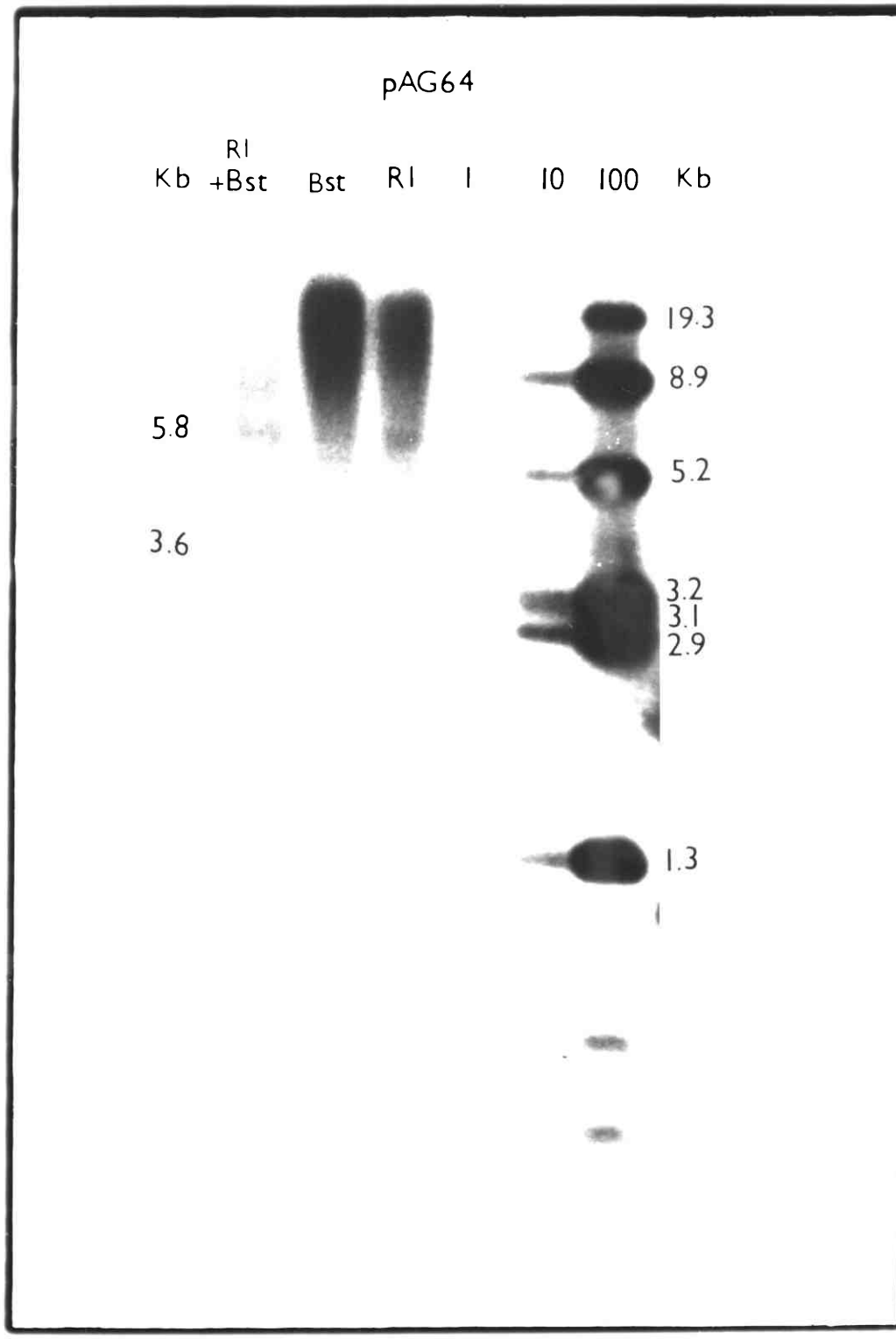

pAG25
M
38
3T3
M
Kb
Kb
5.2
4.8
2.3
2.8
1.8
1.7
1.35
1.2
1.1
0.8

FIGURE 5.20

Figure 5.20 DNA Sequences Homologous to pAG64.

10μg aliquots of SV3T3 C138 DNA were digested with EcoRI, BstI and EcoRI + BstI electrophoresed on a 0.8% neutral agarose gel, transferred to nitrocellulose and hybridized with ^{32}P-labelled nick-translated pAG64 DNA. Plasmid size markers were included in amounts equivalent to 100, 10 and 1 copy per cell. The hybridization contained 10ng ml^{-1} of labelled probe (sp. act. = $>10^8$ cpm μg^{-1} DNA) and exposure was for 4 days. All conditions were as described in Chapter 2.


pAG64
Kb
RI +Bst
Bst
RI
I
10
100
Kb
19.3
8.9
5.8
5.2
3.6
3.2
3.1
2.9
1.3

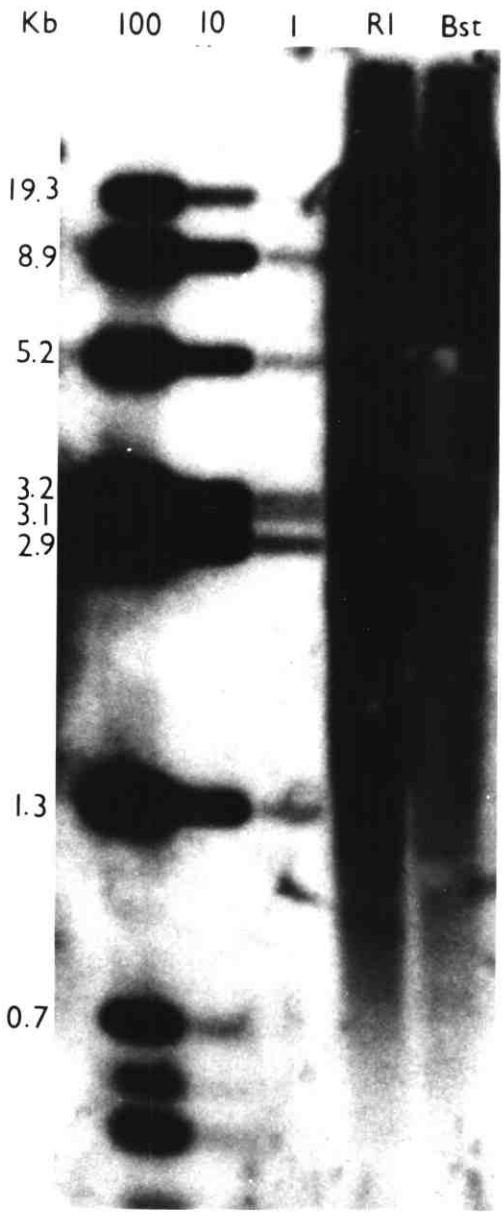
although much of the observed hybridization is to high molecular weight fragments (>10Kb). Double digestion with BstI and EcoRI removes the high molecular weight signals and exposes the 5.8Kb and 3.6Kb EcoRI fragments. Thus, although pAG64 shows homology to a dispersed, repeated sequence, two discrete species are produced by cleavage with EcoRI, apparently contained within one or more larger BstI fragments in the genomic DNA. The presence of dispersed repeated sequences in poly A^+ RNA has been observed before (Crampton et al., 1981; C.J. Watson and S. Legon, personal communication) although their function is as yet obscure. The presence of discrete fragments raises the question of whether they represent the pAG64 coding region or, alternatively, whether hybridization to pAG64 is purely via the repetitive sequence contained within this clone. Especially in the latter instance, localization of the repeated sequence within pAG64 would be a vital prerequisite for any analysis of gene organization. pAG 15 (also Set 1) showed a similar "smear" of signals following hybridization to SV3T3 C138 DNA (Fig. 5.21), indicating that the repeated sequence is also contained in this clone. pAG15 is a truncated, 0.67Kb member of the same set as pAG64. Thus, if pAG15 and pAG64 are derived from the same transcript, these results would imply that the repeated sequence is located towards the 3' end of the mRNA. The fact that both pAG64 and pAG15 hybridize to at least three transcripts in SV3T3 C138, one of 1.85Kb and two less abundant ones of 1.0Kb and 0.9Kb (D. Murphy, personal communication (see also Section 5.6)) also raises the question of whether the observed cross homology at the mRNA level is related to the dispersed repeated DNA sequence. If so, the possibility that pAG15 is in fact derived from either the 1.0Kb or 0.9Kb mRNA cannot be excluded at this stage although

FIGURE 5.21

Figure 5.21 DNA Sequences Homologous to pAG15.

The legend to this Figure is as for Figure 5.20 except that the probe used was pAG15.

pAG 15



preliminary restriction mapping data (my unpublished observations) indicate that these clones represent overlapping segments of the same transcript. A subclone of pAG64 lacking the repeated sequence would be desirable for further analysis of the structure of the genomic sequences that encode the 1.85Kb transcript.

DNA sequences homologous to clones of Set 2

The observation that members of Set 2 hybridize to several species of mRNA in SV3T3 C138 raises the question of whether this observed cross hybridization is due to a repeated sequence element of the type contained in pAG64 and pAG15 (Set 1). When pAG37 (Set 2) was used to probe EcoRI digested SV3T3 C138 DNA, hybridization to 3 fragments of 20Kb, 13Kb and 8.5Kb was observed, all judged to be represented once or a few times per genome equivalent (see Fig. 5.22). Thus, if the Set 2 transcripts contain a repeated sequence element, it is not represented in the portion encompassed by pAG37 and hence, by inference, is remote from the 3' end of the transcript from which this clone was derived. The experiment of Fig. 5.22 also rules out the possibility that any extensive gene amplification is responsible for the observed expression at greatly elevated levels of the corresponding mRNAs in the transformed cell line (see Section 6.2).

DNA sequences homologous to clones of Sets 3 and 4

DNA gel transfer hybridization to EcoRI and BstI digests of SV3T3 C138 DNA was performed using as probes members of Sets 3 and 4. Both pAG82 (Set 3) and pAG47 (Set 4) hybridized to the same 14.0Kb EcoRI and 7.0Kb BstI fragments, each estimated to be present at least 100 times per haploid genome (Fig. 5.23). This experiment therefore indicates that both clone Sets are related to or derived from the same segment of repeated DNA. This observation will be discussed further

FIGURE 5.22

Figure 5.22 DNA Sequences Homologous to pAG37.

The legend to this Figure is as for Figure 5.20 except that the probe used was pAG37.

pAG37

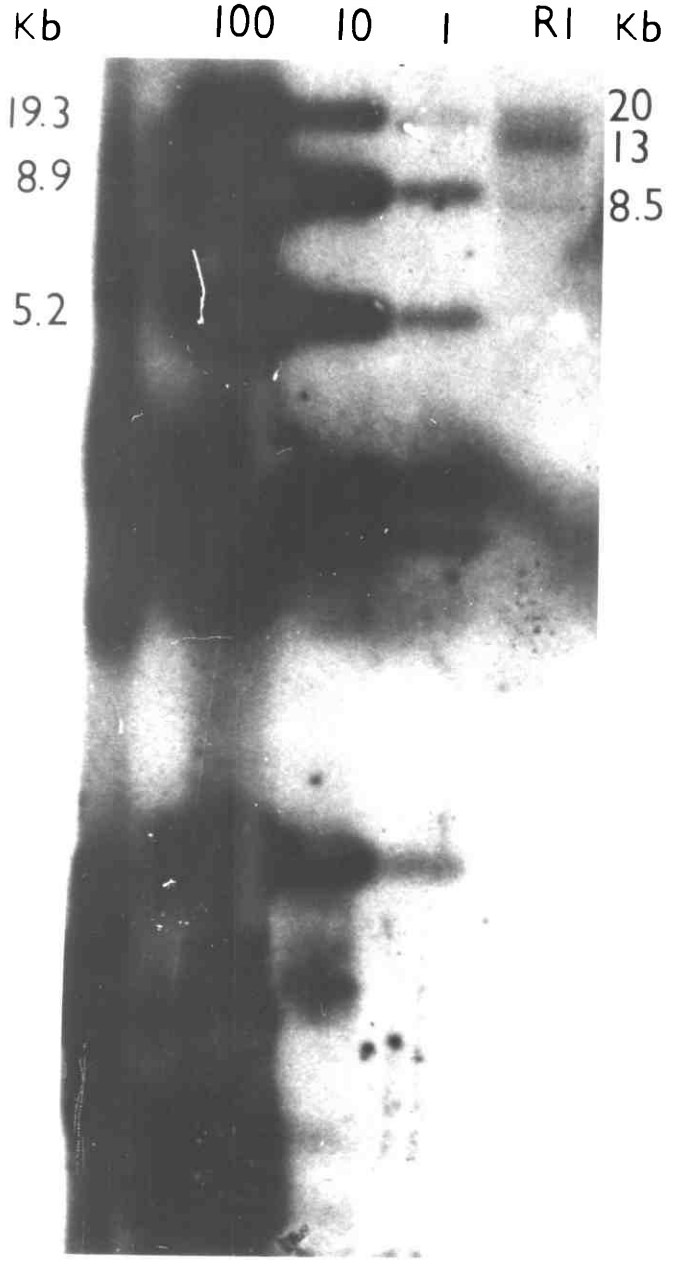
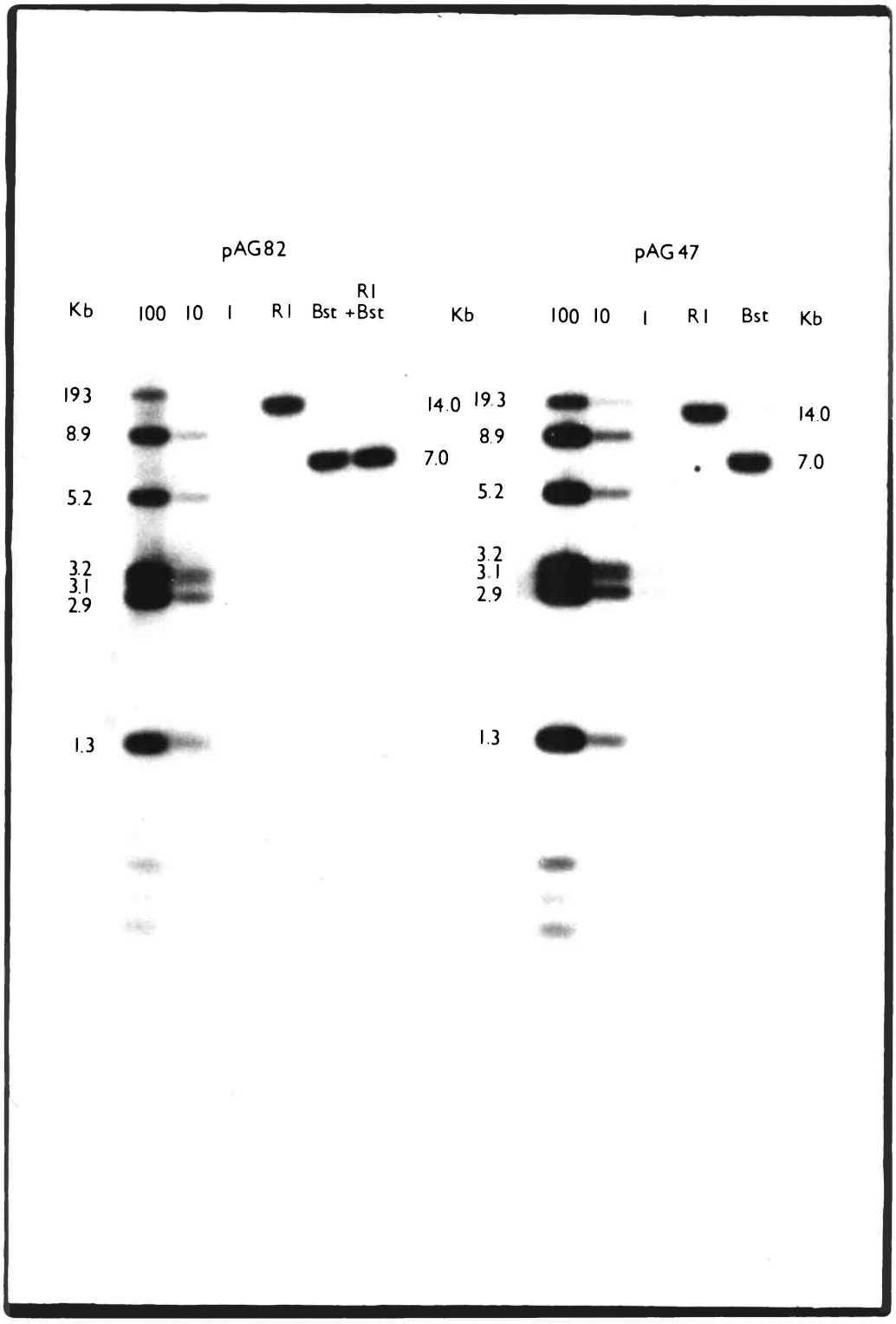


FIGURE 5.23

Figure 5.23 DNA Sequences Homologous to pAG82 and pAG47.

The legend to this Figure is as for Figure 5.20 except that the probes used were pAG82 and pAG47.


pAG82
Kb
100
10
1
RI
Bst
RI +Bst
19.3
8.9
5.2
3.2
3.1
2.9
1.3
14.0
7.0
pAG47
Kb
100
10
1
RI
Bst
Kb
19.3
8.9
5.2
3.2
3.1
2.9
1.3
14.0
7.0

in the next Section. Prolonged exposure of the pAG82 filter hybridization reveals two additional 4.1Kb and 3.4Kb BstI fragments (Fig. 5.24) which are of low abundance and may relate to the origin of those clones belonging to Sets 3 and 4, a notion that will be discussed in Section 6.2.

5.5 pAG25 and Sets 3 and 4 are of mitochondrial origin

A map of restriction endonuclease sites within the 9.5Kb insert of pAG25 is shown in Fig. 5.25. The data from which this map was constructed were largely derived by Stephanie Campbell-Smith (Imperial College, London) and thus will not be subject to consideration here. Cleavage of SV3T3 C138 DNA with EcoRI and BstI followed by transfer hybridization with pAG25 gave the result shown in Fig. 5.26. Hybridization to the internal 2.1Kb EcoRI and the internal 0.7 and 0.3Kb BstI fragments (see Fig. 5.25) was observed, indicating that the clone is at least partly colinear with genomic DNA. Comparison of the signal intensity of the 2.1Kb band (which is of comparable size to the vector moiety of the clone) with DNA markers suggests that the genomic DNA is highly reiterated, being present at least one hundred times per haploid genome. Two BstI fragments of 8.0 and 7.0Kb are apparent, presumably reflecting hybridization to the left and right arms of pAG25, as it is depicted in Fig. 5.25. Since Sets 3 and 4 share homology with pAG25, it seems likely that the 7.0Kb BstI band observed in this experiment is identical to that with the members of Sets 3 and 4 described in Section 5.4. Surprisingly, a single 14.0Kb band is additional to the internal 2.1 Kb EcoRI fragment implying that either the left and right arms of the clone hybridize to the same 14.0Kb band, or that they recognize different fragments of identical mobility. The former interpretation would be indicative of either a tandemly repeated or a circular structure for the DNA, a conclusion

FIGURE 5.24

Figure 5.24 Low Abundance DNA Sequences Homologous to pAG82.

The legend to this Figure is as for Fig. 5.20 except that the probe used was pAG82 and the exposure time was 21 days.

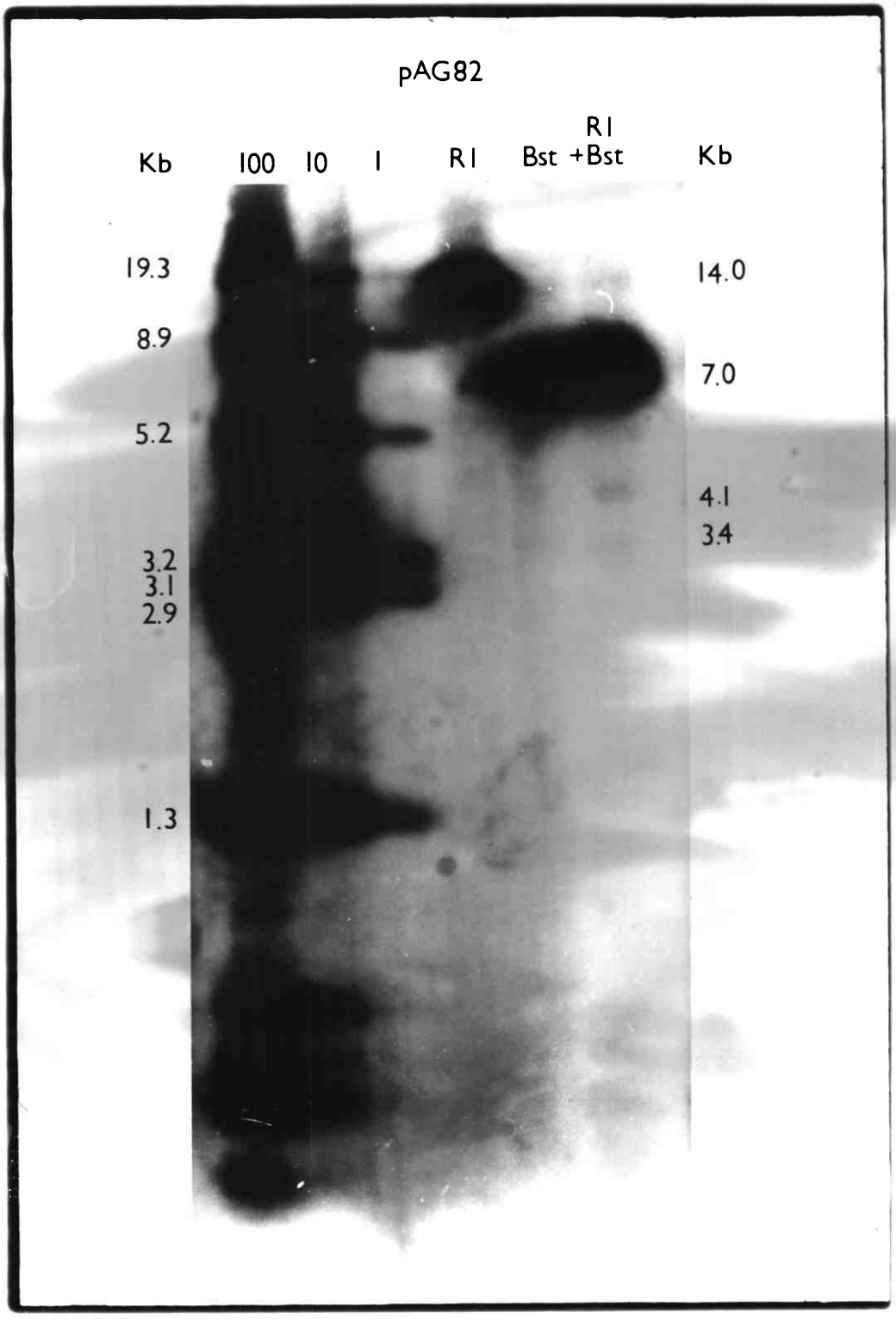
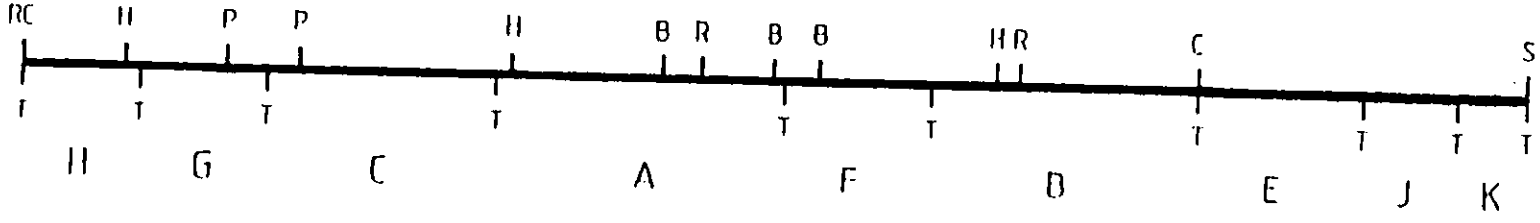
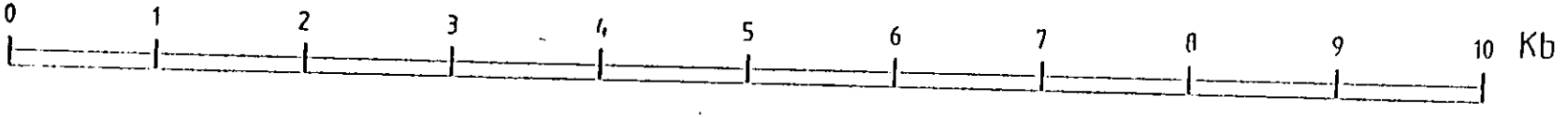

pAG82
Kb
100
10
1
R1
Bst
R1 +Bst
Kb
19.3
8.9
5.2
3.2
3.1
2.9
1.3
14.0
7.0
4.1
3.4

FIGURE 5.25

Figure 5.25 A Map of Restriction Endonuclease Sites Within the 9.5Kb Insert of pAG25.

Shown are the cleavage sites for EcoRI (R), BstI (B), ClaI (C), SalI (S), PvuII (P), HpaI (H) and TaqI (T). The 9 insert-specific fragments produced by TaqI cleavage are also indicated.


RC H P P H B R B B H R C S
T T T T T T T T T T
H G C A F D E J K


0 1 2 3 4 5 6 7 8 9 10 Kb

that is strongly supported by the observation that the sum of the sizes of fragments obtained following BstI cleavage (0.3 + 7.0 + 8.0 = 16.0Kb) is almost identical to the sum of the EcoRI fragment sizes (2.1 + 14.0 = 16.1).

The 2.1Kb ClaI-SalI fragment from the right arm of pAG25 (Fig. 5.25) was subcloned (see Chapter 2) and used to probe a genomic SV3T3 C138 DNA blot (Fig. 5.26). Hybridization was seen to the 14.0Kb EcoRI fragment, as expected, but, although more of the hybridization occurred to the 8.0Kb BstI fragment, surprisingly, the 7.0Kb fragment was also evident. Further cleavage with EcoRI reduced the 8.0Kb fragment to one of 7.0Kb. The 8.0Kb BstI fragment is therefore homologous to the right arm of pAG25 as it is depicted in Fig. 5.25. In addition some evidence exists for homology between the 7.0Kb and 8.0Kb BstI fragments, which are presumably represented in part by the left and right arms respectively of pAG25. Elucidation of the probable organization of pAG25-related genomic sequences in a repeated 16.0Kb unit which shows evidence of existing in either a tandemly arranged or a circular array raises the question whether this clone (and by inference also the members of Sets 3 and 4) could be of mitochondrial origin. The sequence of the 16.295Kb mouse L-cell mitochondrial genome has recently been determined (Bibb et al., 1981). Comparison of the restriction map of pAG25 with the map of mouse L-cell mitochondrial DNA deduced by inspection of the sequence (Bibb et al., 1981) demonstrates the mitochondrial origin of pAG25 (Fig. 5.27).

In an effort to more precisely localize the regions of homology between pAG25 and members of Sets 3 and 4, pAG25, pAG82 (Set 3) and pAG47 (Set 4) were hybridized to TaqI digested SV3T3 C138 DNA (Fig. 5.28) pAG82 hybridized to the 1.4Kb Taq C fragment and pAG47 to the

FIGURE 5.26

Figure 5.26 DNA Sequences Homologous to pAG25 and a Subcloned Fragment of pAG25.

The legend for this Figure is as for Fig. 5.20 except that the probes used were pAG25 and pAG25/CS (a subcloned 2.1Kb ClaI-SalI fragment of pAG25, described in the text).

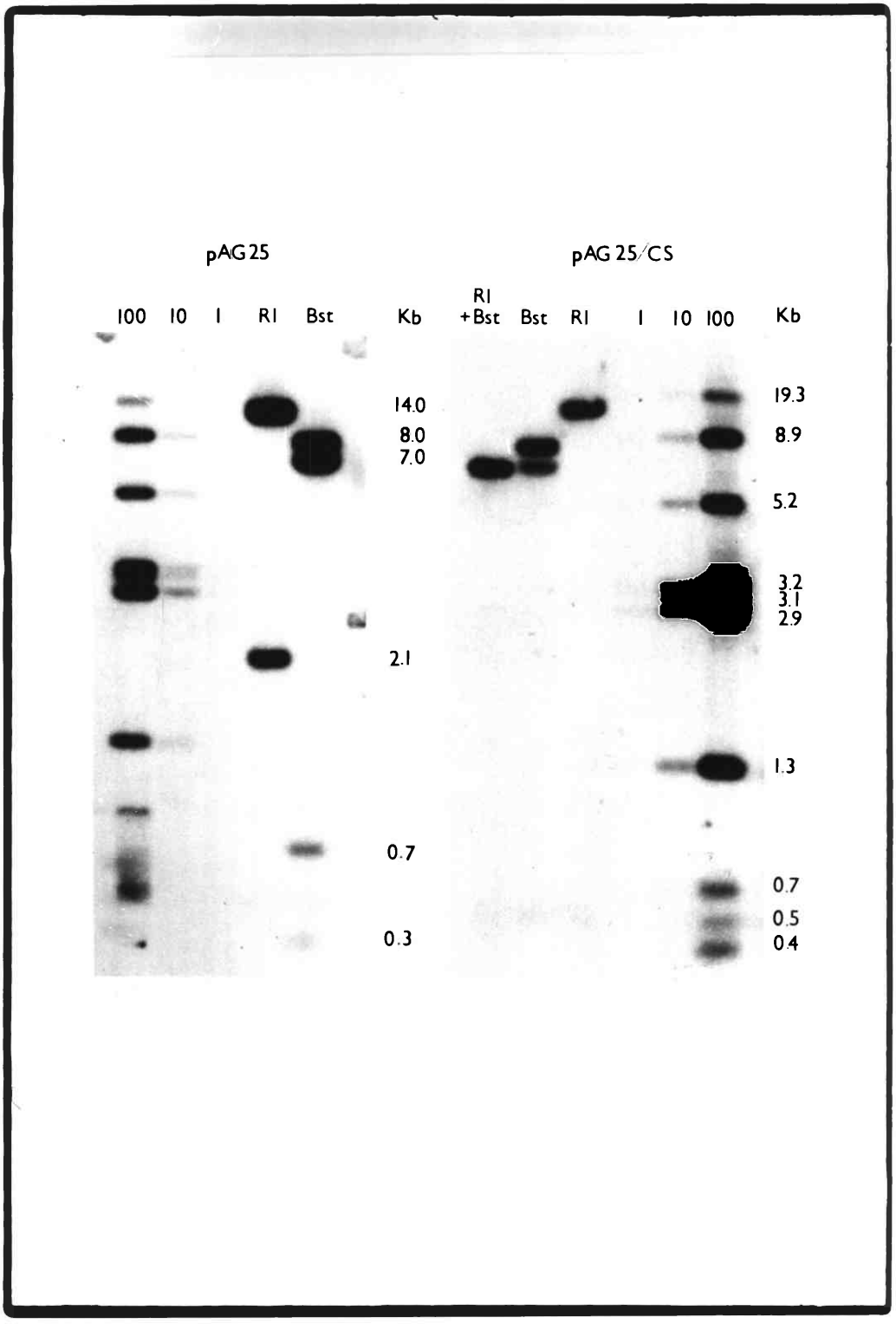

pAG25
100
10
1
RI
Bst
Kb
14.0
8.0
7.0
2.1
0.7
0.3
pAG25/CS
RI
+Bst
Bst
RI
1
10
100
Kb
19.3
8.9
5.2
3.2
3.1
2.9
1.3
0.7
0.5
0.4

FIGURE 5.27

Figure 5.27 Comparison of the Restriction Endonuclease Cleavage Map of pAG25 With That of Mouse Mitochondrial DNA.

Shown are the cleavage sites for EcoRI (R), BstI (B), ClaI (C), BglI, BglII (II), SalI (S), PvuII (P) and TaqI on pAG25 DNA and on the mouse L-cell mitochondrial genome. Also shown are the predominant regions of homology to pAG82 and pAG47 and the location of the genes for cytochrome oxidase subunit I (COI) and subunit II (COII). Note that the orientation of the pAG25 has been reversed relative to Fig. 5.25 so that the direction of transcription is from left to right.

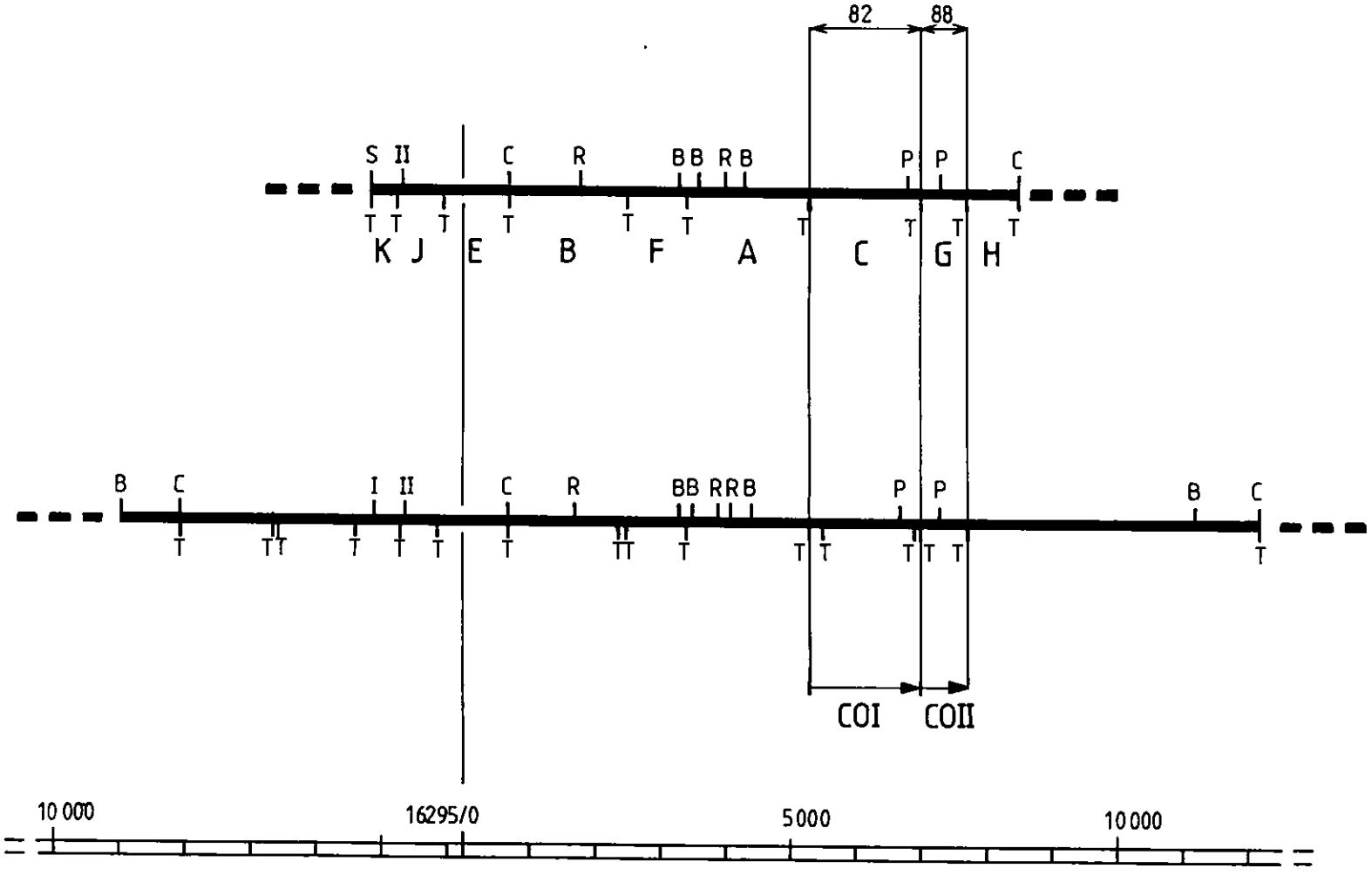
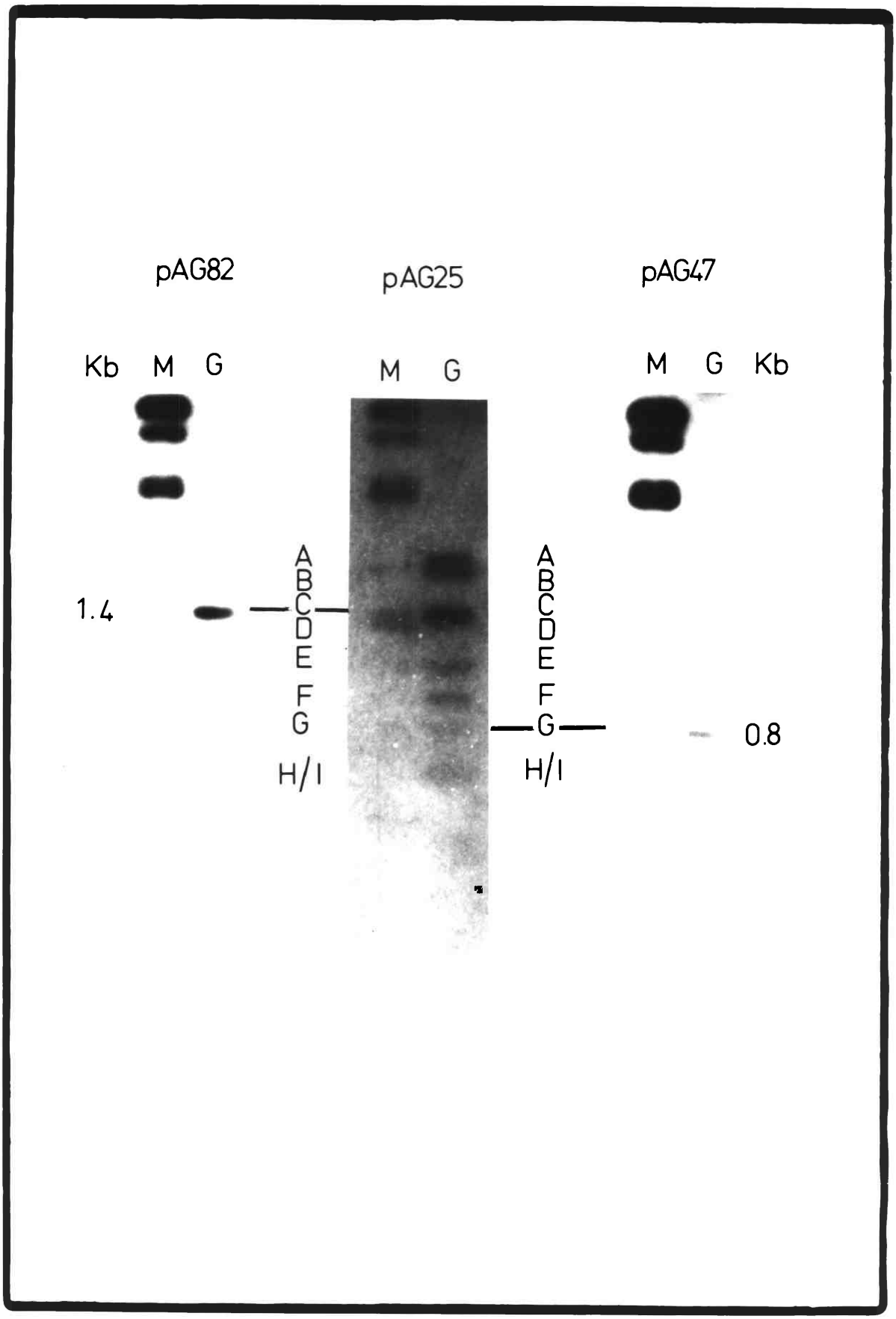

82
88
S II
C
R
B B R B
P P
C
K J E B F A C G H
B C
I II
C
R
B B R R B
P P
B C
COI
COII
10 000
16295/0
5000
10 000

FIGURE 5.28

Figure 5.28 Regions of Homology on pAG25 to pAG82 and pAG47

1μg aliquots of SV3T3 C138 DNA digested with TaqI were electrophoresed in a 1.5% neutral agarose gel, together with plasmid DNA size markers (M), transferred to nitrocellulose filters and hybridized as described in the legend to Fig. 5.20 with pAG25, pAG82 and pAG47. The genomic DNA fragments that correspond to TaqI fragments of pAG25 are indicated (see Fig. 5.25).


pAG82
pAG25
pAG47
Kb
M
G
M
G
M
G
Kb
1.4
A
B
C
D
E
F
G
H/I
A
B
C
D
E
F
G
H/I
0.8

0.8Kb Taq G fragment (see Fig. 5.25). Comparison with the data of Bibb et al. (1981) revealed that the 1.4Kb Taq C and 0.8Kb Taq G fragments comprise the majority of the regions of mitochondrial DNA that encode cytochrome oxidase subunit I (COI) and subunit II (COII) respectively.

The observed homology between the ClaI-SalI fragment of pAG25 and the 8.0Kb BstI fragment (see Fig. 5.26) may be a consequence of cross-hybridization between tRNA genes which are distributed between protein-coding sequences on the mitochondrial DNA.

5.6 TheGenerality of the altered expression of transcripts homologous to the clone sets

It is now necessary to discuss the significance to SV40-mediated transformation of the SV3T3 C138 "activated" sequences. It it possible that although most cells in a normal line show nearly identical patterns of gene expression, individual variants exist which show slight differences in the expression of a small number of phenotypically silent genes. Selection of SV40- transformed derivatives inevitably involves subcloning, hence variants of this type could be the parents of the transformed cell lines. This possibility could be tested by subcloning the Balb/c 3T3 line under non-selective conditions and using the procedures described in this thesis to compare the normal cell line with its subclones. Even if this was so, it need not necessarily imply that the changes described here do not confer a selective advantage; indeed, it is possible that SV40 transformation is enhanced in cells with a particular backgound of host gene expression. What is of primary importance is whether the observed variations in cellular gene expression are unique to the cell line studied, SV3T3 C138. If the same changes were observed in other independently isolated SV40-transformed lines this

would provide firm evidence against the hypothesis that the changes described here might be phenotypically silent, irrespective of whether they occurred prior to or subsequent to transformation by SV40.

Generality of expression of transcripts homologous to members of Set 1

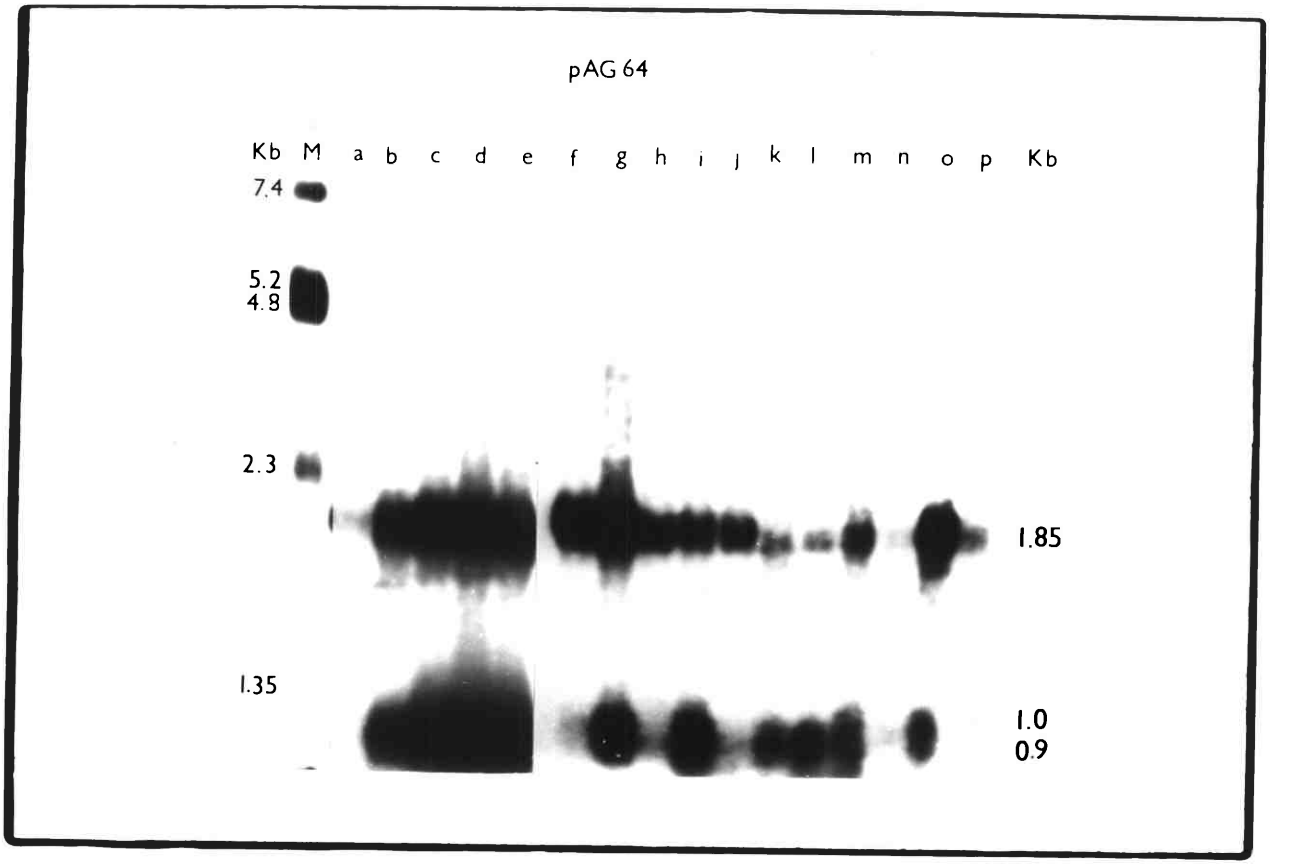
In Fig. 5.29, poly A^+ RNAs taken from Balb/c 3T3, SV3T3 C138, three methocell-derived subclones of C138 (C138A2, B4 and B5) and five independently isolated SV40-transformed Balb/c 3T3 lines, SV3T3 C120, C126, C149, C1H and C1M, have been hybridized in parallel to ^{32}P-labelled nick-translated pAG64 (Set 1). Also included are equivalent amounts of poly A^+ RNA taken from SV3T3 A209B4a, A255B1b and A7B4b. These three lines were isolated by Brockman (1978) following transformation of Balb/c 3T3 cells with temperature sensitive (tsA) mutants of SV40 (tsA209, tsA255 and tsA7, respectively). The effect of temperature-shift on the expression of transcripts related to members of Set 1 in the SV40(tsA)-transformed lines will be discussed in more detail in Section 5.7, but the same samples have been included in Fig. 5.29 to allow cross referencing between the results of this experiment and those of Section 5.7.

All the transformed cell lines display elevated expression of the 1.85Kb species when compared to the normal Balb/c 3T3 control (track a Fig. 5.29). The highest levels are observed in SV3T3 C138 and the three methocell-derived subclones of this line (tracks b, c, d and e) and in SV3T3 C120 (track f) and SV3T3 C126 (track g). SV3T3 C149 (track h) SV3T3 C1M (track i) and SV3T3 C1M (track j) contain slightly less of the 1.85Kb species although the level of expression is markedly higher than in the normal Balb/c 3T3 control. The three SV40(tsA)-transformed lines show a smaller yet significant increase in the amount of the 1.85Kb species (Fig. 5.29, the relevant tracks

FIGURE 5.29

Figure 5.29 Transcripts Homologous to pAG64 in Several Cell Lines

1μg aliquots (determined spectrophotometrically) of polyA$^+$ RNA isolated from a number of mouse cell lines were electrophoresed in a 1.5% formaldehyde agarose gel, transferred to nitrocellulose and hybridized with ^{32}P-labelled nick-translated pAG64. Plasmid DNA size markers were also included in amounts approximately equivalent to 0.1% of polyA$^+$ RNA (M). The samples are Balb/c 3T3 (track a), SV3T3 C138 (track b), SV3T3 C138/A2 (track c), SV3T3 C138/B4 (track d), SV3T3 C138/B5 (track e), SV3T3 C120 (track f), SV3T3 C126 (track g), SV3T3 C149 (track h), SV3T3 C1H (track i), SV3T3 C1M (track j), SV3T3 A255B1b at 32^{o}C (track k), SV3T3 A209B4a at 32^{o}C (track l), A7B4b at 32^{o}C (track m) and SV3T3 A255B1b, A209B4a and A7B4b at 39.5^{o}C (tracks n, o and p).


pAG 64
Kb M a b c d e f g h i j k l m n o p Kb
7.4
5.2
4.8
2.3
1.35
1.85
1.0
0.9

are k, l and m which contain poly A^+ RNA extracted from these cells at the permissive temperature of 32°C; at this temperature these cell lines are demonstrably transformed for several criteria (Brockman, 1978). The 0.9Kb and 1.0Kb species are poorly discriminated in this experiment but this is not a severe problem since the experiment of Section 5.7, which clearly resolves these two components reveals that they are probably coordinately expressed. The 0.9/1.0Kb species are undetectable in the normal Balb/c 3T3 control but are evident at levels comparable to the 1.85Kb species in SV3T3 C138 and its subclones (tracks b, c, d, e) and in SV3T3 C126 (track g), SV3T3 C1H (track i) and the SV40(tsA)-transformed lines (tracks, k, l and m). Three of the transformed cell lines, SV3T3 C120 (track f), SV3T3 C149 (track h) and SV3T3 C1M (track j), contain markedly lower but still detectable levels of the 0.9/1.0Kb species.

Thus the 1.85Kb 1.0Kb and 0.9Kb mRNAs show elevated expression in all the SV40-transformed cell lines studied and the 0.9/1.0Kb species are present at rather lower levels in three of the transformed cell lines (SV3T3 C120, SV3T3 C149 and SV3T3 C1M) than in the others. These observations will be discussed in Section 6.2 but three points should be noted here.

Firstly, elevated expression of poly A^+ RNAs related to members of Set 1 is such a common observation in SV40 transformed cells that it may be involved in the mechanism of maintenance of the transformed state.

Secondly, the fact that the SV40(tsA)-transformed lines show less pronounced over-expression of the 1.85Kb species than do those lines transformed by wild-type SV40 is consistent with the notion that activation of Set 1-related sequences is a consequence of a direct interaction between large T-antigen and the genes that encode these

these transcripts. Thus SV40 genomes carrying the tsA lesions, although able to effect transformation at 32°C, encode defective large T-antigens. If elevated expression of Set 1-related transcripts were a consequence of an interaction with large T- antigen, the inherent defect of tsA mutations might explain the less pronounced activation observed in cells transformed by SV40(tsA) mutants than in those transformed by wild type SV40.

Thirdly, the possibility that the observed effects are consequences of differences in the conditions of growth of the cell lines studied may be discounted, since all the RNA preparations used for the experiment of Fig. 5.29 were extracted from sub-confluent cultures of cells at equivalent cell densities and all the cell lines were maintained on medium with the same serum supplement (10% (v/v) FCS, see Section 2.1).

Expression in other cell lines of transcripts homologous to members of Set 2

Fig. 5.30 shows the result of an analogous experiment to the previous one using pAG59 (Set 2) as probe. The 3.2Kb transcript is observed in SV3T3 C138 and all its subclones (Fig. 5.30, tracks b, c, d and e) and, at even higher levels, in one of the other SV3T3 lines, C126 (track g). None of the other cell lines showed any evidence of increased expression of any Set 2-related species; this is true also for the SV40(tsA)-transformed lines which have not been included in Fig. 5.30. Surprisingly, the 6.5Kb transcript is present at higher levels in the three methocell-derived C138 subclones than in the parental line. Inspection of the data of Table 5.3 reveals that SV3T3 C138 forms colonies in methocell at a rather lower frequency than do most of the other SV3T3 lines studied; it is therefore possible that this transcript is expressed in a sub-population of

FIGURE 5.30

Figure 5.30 Transcripts Homologous to pAG59 in Several Cell Lines.

The legend for this Figure is as for Fig. 5.29 except that the probe used was pAG59 and the RNA samples isolated from the three SV40 (tsA) transformed lines (tracks k, l, m, n, o and p in Fig. 5.29) have been omitted.

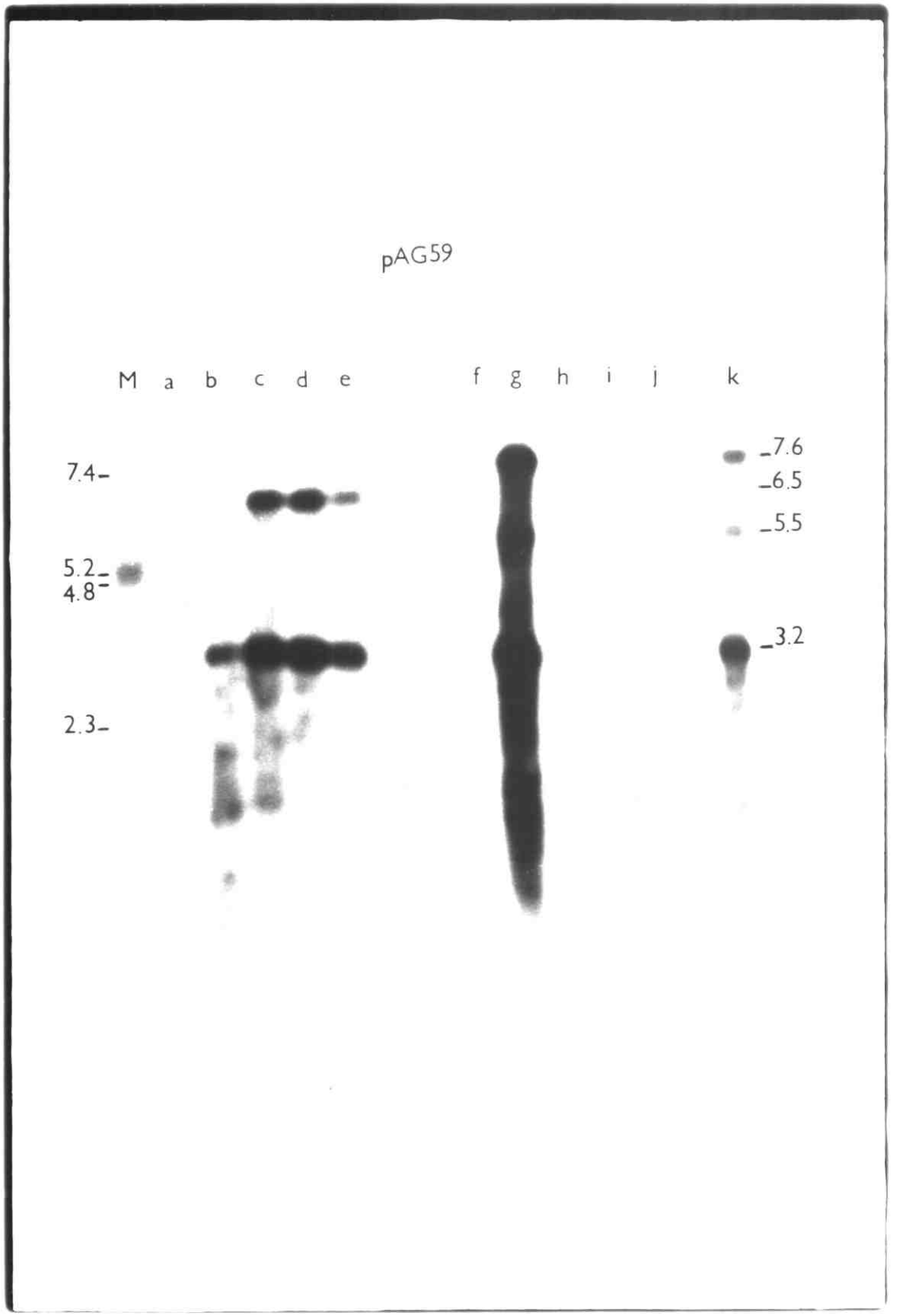

pAG59
M a b c d e f g h i j k
7.4
5.2
4.8
2.3
7.6
6.5
5.5
3.2

TABLE 5.3

Table 5.3 Biological Properties of a Number of SV40-transformed Mouse Cell Lines Contrasted with those of Normal Balb/c 3T3 Cells

The data for this Table are taken from Rigby et al. (1980).

Cell Line	Moi. of Transformation	% Efficiency of Plating on 3T3 Monolayers	% Efficiency of Plating in Methocell	Fibronectin % of 3T3	Tumorigenicity in Nude Mice Dose: 2 x 10^6 Cells Tumor positive / total	Latent Period
Balb/c 3T3	-	0.1	<0.0005	100	$^0/_4$	>150
SV3T3 C120	100	80	43	8.5	$^1/_6$	63
SV3T3 C126	10	6	11	15	$^3/_3$	20
SV3T3 C138	1	2	4.35	35	$^3/_3$	19
SV3T3 C149	1	8	7.2	18	$^0/_3$	>150
SV3T3 C1H	0.5	41	8.3	7	$^3/_3$	18
SV3T3 C1M	0.5	7	30	10	$^2/_3$	14

cells within the SV3T3 C138 line which are capable of efficient anchorage independent growth. The most easily testable prediction of this hypothesis is that the methocell-derived C138 subclones should plate with higher efficiency in methocell than the parental SV3T3 C138 line, but, unfortunately, this has not yet been tested.

In addition to the 3.2Kb RNA, SV3T3 C126 expresses a 7.6Kb species, which may correspond to the 7.6Kb transcript observed in SV3T3 C138, and an additional species of 5.5Kb which is not evident in SV3T3 C138. This makes it unlikely that the higher molecular weight Set 2-related RNAs are merely precursors to the 3.2Kb species and suggests that they are the products of four discrete but related genes. Consistent with the latter idea is the observation(Fig. 5.22) that pAG59 (Set 2) hybridizes to at least three large (>8Kb) EcoRI fragments of genomic DNA. Although activation of genes related to clones from Set 2 is certainly not a general phenomenon in SV40-transformed cell lines, the fact that it is observed in two independently derived SV3T3 lines, C138 and C126, argues that expression of Set 2-related transcripts may confer upon the cell a selectable change in growth behaviour; this hypothesis will be developed further in Section 6.2.

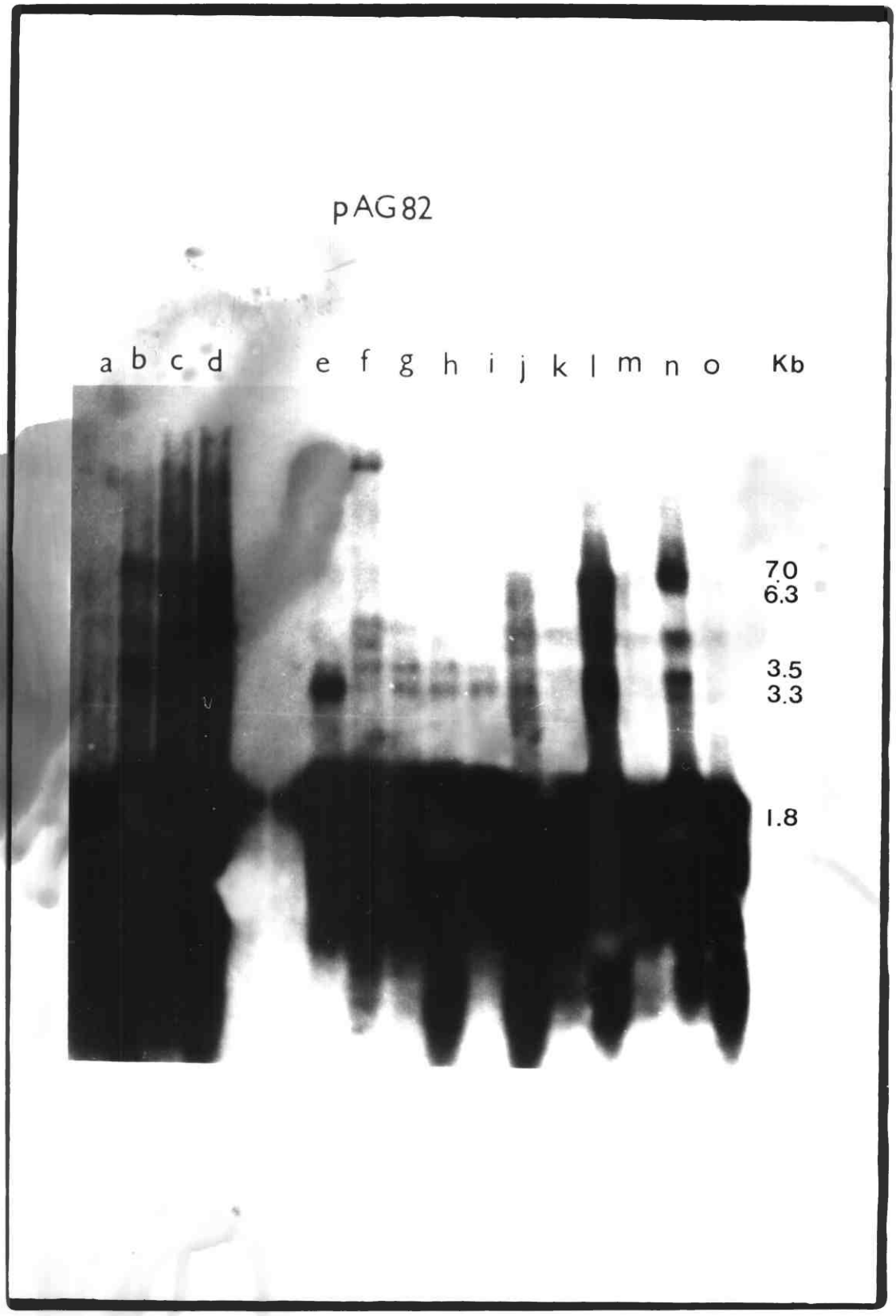
Expression of mitochondrial transcripts in other cell lines

Fig. 5.31 shows the result of a gel transfer hybridization experiment to monitor the expression of sequences related to pAG82 (Set 3) and hence to the mouse mitochondrial COI gene (Section 5.5). The SV3T3 C138 specific Set 3-related transcripts are apparent in poly A^+ RNA extracted from SV3T3 C138 and its subclones. Although similarly expressed in SV3T3 C120, SV3T3 C126, SV3T3 C149, SV3T3 C1H and SV3T3 C1M, these lines show some slight differences in the relative abundance of particular species, notably the 3.3Kb transcript

FIGURE 5.31

Figure 5.31 Transcripts Homologous to pAG82 in Several Cell Lines

The Legend to this figure is as for Figure 5.29 except that the probe used was pAG82 and the sample tracks are (a) SV3T3 C138, (b) SV3T3 C138/A2, (c) SV3T3 C138/B4, (d) SV3T3 C138/B5, (e) SV3T3 C120, (f) SV3T3 C126, (g) SV3T3 C149, (h) SV3T3 C1H, (i) SV3T3 C1M, (j, k and l) SV3T3 A255B1b, A209B4a and A7B4b at 32^{o}C and the same three cell lines at 39.5^{o}C (tracks m, n and o).


pAG82
a b c d e f g h i j k l m n o Kb
7.0
6.3
3.5
3.3
1.8

which is prominent in SV3T3 C120. The relevance of these observations to SV40 transformation is as yet obscure due to uncertainty about the relationship of the SV3T3 C138 specific transcripts to the 1.8Kb COI mRNA. Although from this experiment they appear to be characteristic of transformed cells, it is quite likely that the presence of the higher molecular weight species is a result of incomplete processing of mitochondrial RNA precursors, perhaps reflecting rather than contributing to the altered metabolism of transformed cells.

The expression of the Set 4-related transcripts was not investigated, but it seems reasonable to presume that an analogous behaviour would be observed.

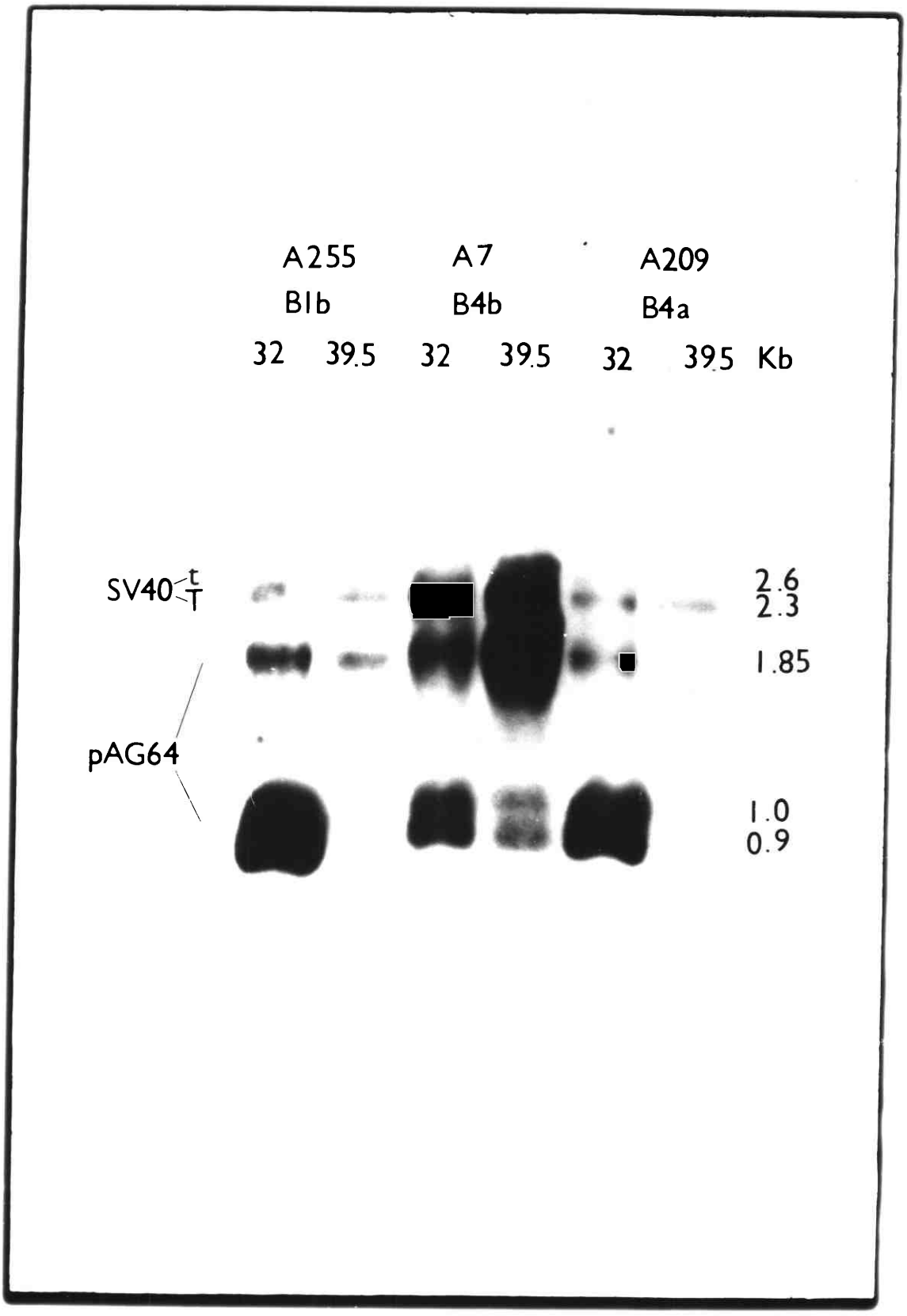
5.7 Reversion of SV40(tsA)-transformed cells is accompanied by altered expression of Set 1-related transcripts.

Fig. 5.32 shows the result of an experiment designed to monitor the effect of temperature-shift upon the expression of transcripts related to Set 1 clones in cells transformed by SV40(tsA) mutants. The three cell lines under study, SV3T3 A209B4a, SV3T3 A255B1b and SV3T3 A7B4b were described by Brockman (1978). Two of the cell lines, A255B1b and A209B4a, are completely temperature dependent, i.e., they revert to the growth behaviour and morphology characteristics of normal cells when maintained at the non-permissive temperature of 39.5°C. The other line, A7B4b, is partially temperature-independent in that it remains morphologically transformed at the non-permissive temperature and does not completely revert for other transformation parameters. In the experiment of Fig. 5.32, identical amounts of poly A^+ RNA (determined by ^3H-poly U assay, see Section 2.12) extracted from the three cell lines maintained at either the permissive (32°C) or non-permissive (39.5°C) temperature were electrophoresed on a

FIGURE 5.32

Figure 5.32 Transcripts Homologous to pAG64 in SV40 (tsA) Transformed Balb/c 3T3 Cells.

1µg aliquots (determined by ^{3}H-poly U assay, see Section 2.12) of polyA^{+} RNA isolated from three SV40 (tsA) transformed Balb/c 3T3 cell lines SV3T3 A209B4a, SV3T3 A255B1b and SV3T3 A7B4b maintained for 4 days at either 32^{o}C or 39.5^{o}C were electrophoresed in a 1% formaldehyde agarose gel, transferred to nitrocellulose and hybridized with ^{32}P-labelled nick-translated pAG64 and SV40 DNA.


A255
B1b
A7
B4b
A209
B4a
32
39.5
32
39.5
32
39.5
Kb
SV40 t
T
pAG64
2.6
2.3
1.85
1.0
0.9

formaldehyde agarose gel, transferred to nitrocellulose and hybridized sequentially with ^{32}P- labelled SV40 DNA and ^{32}P-labelled pAG64 DNA. SV40 transcription in SV40(tsA)-transformed cell lines is under investigation in this laboratory and will not be discussed in detail here, but the SV40 transcripts have been included in Fig. 5.32 to allow comparison with the expression of the Set 1 related species. The 0.9Kb and 1.0Kb Set 1-related transcripts which are clearly resolved in this experiment are not observed in the two temperature dependent SV40(tsA)-transformed lines, A209B4a and A255B1b, when they are maintained at 39.5°C. In contrast the 0.9Kb and 1.0Kb species are present, although at reduced levels, in the partially temperature-dependent line A7B4b following the temperature shift. This experiment therefore provides evidence for a connection between the viral transforming protein, large T-antigen, and the expression of two cellular transcripts and is thus consistent with the notion that expression of the 0.9Kb and 1.0Kb Set 1-related species is induced either directly by interaction with large T-antigen or indirectly via an intermediate mechanism; this could for example involve a complex between large T-antigen and a cellular protein.

The expression of the 1.85Kb Set 1-related transcript parallels that of the SV40-specific mRNAs. In both cases the temperature-dependent lines (A209B4a and A255B1b) show little change in the levels of expression of these RNAs. In contrast, however, the partially temperature-independent line (A7B4b) overproduces the 1.85Kb Set 1-related and SV40 specific species when shifted to the non-permissive temperature (39.5°C). Whether this effect could be related to the partial temperature dependence of this cell line is not clear from this experiment since the three cell lines were derived using different tsA mutants, but this possibility is being investigated (D.

Murphy, personal communication). Nevertheless, the parallel expression of the 1.85Kb Set 1-related species and the SV40 mRNAs provides further evidence for a link between products of the viral genome and expression of the Set 1-related transcripts.

5.8 Conclusion

The main conclusions to be drawn from the experiments described in this Chapter are:

(a) There is evidence for altered expression in SV40 transformed cells of sequences related to the genes for two cytochrome oxidase subunits, COI and COII, although the significance of this observation is as yet obscure.

(b) A group of cellular poly A^+ RNAs are induced in SV40-transformed cells. At least one of these transcripts is over-expressed in all SV40-transformed cells studied and there is evidence for a causal connection between expression of the viral transforming protein and the induction of these species.

(c) Another group of cellular poly A^+ RNAs are associated with SV40-transformation although their expression is clearly not mandatory for transformation by SV40.

One very important question remains, that is whether the observed cellular gene activation is a contributory factor in the expression of the transformed state. Thus, even in the case of the Set 1-related transcripts where there is evidence for a connection between expression of the viral genome and induction of these cellular mRNAs, it is not yet certain that this connection is part of the mechanism of transformation by SV40. This question will be discussed in Section 6.2.

CHAPTER VI

GENERAL DISCUSSION

6.1 Screening for differences between two mRNA populations

The customary method of differential colony hybridization has worked successfully on many occasions but as has been described in Chapter 4 of this thesis it suffers from severe technical limitations. These preclude analysis of the vast majority of expressed sequences which are insufficiently abundant to elicit a signal over background "noise" following hybridization. Precedents existed at the outset of this work which suggested it would be sensible to search for sequences showing altered expression following oncogenic transformation by SV40 (see Chapter 1). When this was tried the standard approach yielded a negative result. Common sense dictates that this is unlikely to be an exceptional circumstance; it therefore becomes imperative to be able to identify differential expression of mRNA sequences located in the abundance spectrum below the limits of sensitivity of conventional colony hybridization. One of several possible approaches to this problem has been applied, successfully, to identify and clone a number of sequences found at substantially increased levels in SV40-transformed cells. These will be discussed in the second part of this Chapter. In this section I will discuss some potential ways of alleviating the limitations of the technology as it has been applied in this study, together with the most salient applications to future research.

The first limitation is apparent from consideration of the data of Chapter 5. The four clone Sets all show evidence of hybridizing to more than one transcript, and in three cases (Sets 2, 3 and 4) one of these RNAs is not differentially expressed. The 1.8Kb Set 2 transcript common to both normal and transformed cells is expressed at markedly lower levels than the 3.2Kb major transformation specific species and is of such low abundance that it would not deplete the

probe of sequences complementary to the 3.2Kb mRNA. If this was not the case, for example if the common mRNA was of comparable abundance to the transformation-specific mRNA, then its presence in the normal library could obscure detection of the related transformation-specific species. For example, if the region of homology was adjacent to the 3' terminus of the RNA, selection of a truncated probe in the manner described in Fig. 5.17 would clearly be impossible. To guard against this eventuality it might be appropriate to use a probe consisting of short, random fragments of cDNA. This could be synthesized from poly A^+ RNA using reverse transcriptase in the presence of random sequence oligonucleotide primers (Taylor et al., 1976). If necessary, dideoxynucleoside triphosphates could be included in the reaction to limit the length of the probe to an optimal size, probably of the order of 100-200 nucleotides. The selection procedures described in Chapter 4 would then be employed, yielding a purified probe specific for any activated sequences irrespective of whether or not they exist in conjunction with regions homologous to common abundant mRNAs.

A randomly primed cDNA probe might contribute toward solving another limitation in the technology, that of sensitivity. Although the detection threshold of signal-to-noise ratio has been passed, a further barrier to sensitivity is imposed by the maximum specific activity practicable. The cDNA probe used in these studies required 250uCi of α-^{32}P-labelled nucleotide, yet many autoradiographic signals were barely visible (see Chapter 4). For example, most of the Set 3 clones evinced particularly weak responses. Detection of sequences even ten-fold less abundant than these would become exponentially more expensive because, in addition to the obvious requirement for a probe of proportionately higher specific activity,

exists the need to screen proportionately more colonies. Thus cost-effectiveness would become an important consideration. A random sequence probe might provide a means of improving sensitivity without extravagant consumption of labelled nucleotide. Wahl et al. (1979) have reported dramatically increased sensivitity in filter hybridizations with nick-translated DNA probes when these were performed in the presence of dextran sulphate. They proposed that the major effect of the polymer was in accelerating the formation of probe "networks". This effect should be observed with any probe which, like nick-translated DNA, comprises a random assortment of partially complementary fragments. Supplementing the colony screening stage of an experiment using a cDNA probe prepared as above with an excess of randomly fragmented template mRNA (produced, for example, by partial base cleavage) and dextran sulphate would accelerate the rate of hybridization and encourage networking, presumably affording a concomitant improvement in sensitivity. It must be stressed, however, that although increases in sensitivity of up to 100-fold have been reported by Wahl et al. (1979), an improvement of this magnitude would be unlikely in this instance since the net amount of probe complementary to any given clone would be limiting, in contrast to the studies of Wahl et al. (1979) which were performed under conditions of massive probe excess. Despite this caveat, the concept remains worthy of consideration.

The third inherent flaw of the approach is that it presumes a particular type of change in expression, that is it can only detect sequences which are either absent from or present at very low levels in one cell type and hence ignores much smaller changes in the abundance of mRNAs. Such modulations in the abundance of cellular sequences might well be a common feature of those changes in gene

expression that control cell behaviour. Probe purification could be adapted to lower the detection threshold of differential colony hybridization. For example, a library could be divided into pools of colonies for screening as described in Chapter 4, except that an extra replica filter would be used to prepare plasmid DNA representative of each individual pool. This would then be coupled to cellulose (Noyes and Stark, 1977; see Chapters 2 and 4), nitrocellulose discs (Gillespie and Spiegelman, 1965) or derivatised paper discs (Stark and Williams, 1979). Probes specific to the two cell types to be examined would then be hybridized to the immobilized DNA and the fraction of each probe that hybridized to each pool of colonies would be used in a differential colony hybridization screen of their corresponding replica filters. Only the fraction of probe complementary to any given set of colonies would be used in their screening, thus reducing background caused by contaminating non-homologous species, the majority of which are of low abundance and would not be represented in any one pool. This procedure would presumably lower the threshold of detection in the colony screen.

There are therefore a number of ways in which the basic approaches to the analysis of differential gene expression described in Chapter 4 could be refined and adapted to overcome the inherent technical limitations. A truly comprehensive analysis of gene expression at the level of poly A^+ RNA now seems not only desirable but possible. Using the high efficiency cloning procedures described in Chapter 3, changes in the expression levels of sequences of relatively low abundance should be accessible to analysis even when only small amounts of starting material (of the order of a few micrograms of poly A^+ RNA) are available. The ability to clone such differentially-expressed sequences where no other assay exists

may help considerably in the study of complex problems like neoplasia and normal cell differentiation and development.

6.2 Gene activation in SV40-transformed cells

The experiments described in Chapters 4 and 5 have established that limited gene activation occurs in cells transformed by SV40. Furthermore, preliminary investigations of the generality of this phenomenon indicate that it is likely to be a common feature of transformation, at least when this is effected by SV40. In this final Section I will consider the following questions.

(a) Is gene activation part of the mechanism of transformation by SV40?

(b) How could gene activation be provoked by SV40?

Is gene activation part of the mechanism of transformation by SV40?

To answer this question it is necessary to seek correlations between the activation of cellular genes and the expression of growth properties characteristic of the transformed state. The first tentative evidence for this, presented in Section 5.6, is that subclones of SV3T3 C138 which had been selected by growth in semi-solid medium over-express the 6.5Kb Set 2-related species when compared to the parental line. It was therefore suggested that overproduction of the 6.5Kb mRNA might be associated with the ability of SV3T3 C138 to plate independently of anchorage. It is important to note that expression of this mRNA is not essential for growth in methocell since this transcript was not observed in other SV3T3 lines, including those that plate efficiently in semi-solid medium. However, it is surely unlikely that expression of any given transformation parameter is effected by a single gene product; it seems more reasonable that such characteristics will result from the participation of many different elements, including the amount and

distribution of different forms of viral gene products, the background of cellular gene expression and perhaps even factors supplied by the growth medium. An understanding of why some cells can, for example, grow in semi-solid medium will therefore require identification and analysis of each contributory element within the system under study; thus elevated expression of the 6.5Kb Set 2- related mRNA could enhance the growth properties of SV3T3 C138 cells in such a way that they grow more efficiently independently of anchorage whereas other lines will exhibit the same property for other reasons.

The Set 2-related transcripts show two other interesting properties. In Chapter 1 it was noted that a recent study (G. Winchester and D. Lane, personal communication) indicated a correlation between the overall amount of T-antigen in each of the six SV3T3 lines and their relative exponential rates of growth. Of particular relevance to this study is the fact that the two lines which contain the lowest levels of T-antigen and show the slowest growth rates are SV3T3 C138 and SV3T3 C126, i.e. the ones which express Set 2-related RNAs. It is also noteworthy that of the six lines tested, SV3T3 C138 and SV3T3 C126 are among the most highly tumorigenic when injected into animals (see Table 6.1). It is thus possible that cells which contain high levels of large T-antigen can proliferate sufficiently rapidly to form tumours whereas more slowly growing cells need to synthesize other cellular factors to achieve the same end. Thus expression of Set 2-related genes could enhance transformation of cells which underproduce large T-antigen. In light of the evidence discussed in Section 1.5 that SV40 could have a tumour promoting activity, it is tempting to speculate that expression of Set 2-related genes results from a promotional mechanism and that this expression allows the progression of transformation toward a more malignant state than would otherwise be observed.

TABLE 6.1

Correlation between tumorigenicity and gene activation in a number of SV40-transformed mouse cell lines

Cell line	Set 1		Set 2				Tumorigenicit
	0.9/1.0 kb	1.85 kb	3.2 kb	5.5 kb	6.5 kb	7.6 kb	+ve/total
Balb/c 3T3	-	-	-	-	-	-	0/4
SV3T3 C138	+	+	+	-	+	+	3/3
SV3T3 C138/A2	+	+	+	-	++	+	ND
SV3T3 C138/B4	+	+	+	-	++	+	ND
SV3T3 C138/B5	+	+	+	-	++	+	ND
SV3T3 C120	+/-	+	-	-	-	-	1/6
SV3T3 C126	+	+	++	+	-	++	3/3
SV3T3 C149	+/-	+	-	-	-	-	0/3
SV3T3 C1H	+	+	-	-	-	-	3/3
SV3T3 C1M	+/-	+	-	-	-	-	2/3

Where expression levels are comparable to SV3T3 C138 this is indicated by +. Where higher levels of expression are apparent this is indicated by ++. Where expression is undetectable this is indicated by - and where expression is observed but at markedly lower levels than in SV3T3 C138 this is indicated by +/-. The data on tumorigenicity have been extracted from Table 5.3.

Since the Set 1-related transcripts are observed at elevated levels in all SV40-transformed cell lines so far tested it is possible to correlate expression of these sequences both with the selection criterion used to isolate these cells (focus formation in low serum) and the transforming agent (SV40). This raises the problem that induction of these cellular RNAs might not be involved in transformation at all but instead could be related to the normal lytic cycle of the virus. There are two obvious ways of dissecting this problem. The first would be to investigate the expression of Set 1-related genes in cells transformed by agents other than SV40. If the same phenomenon was observed in other cell lines this would provide strong evidence for a mechanistic role in transformation. The second way is to try and connect expression of the various Set 1-related species with stringent parameters of transformation. When this is done one interesting correlation emerges. The data of Section 5.6 that pertain to the expression of all the Set 1- and Set 2-related transcripts in the six SV3T3 lines have been summarised in Table 6.1. It is apparent that all of the SV40 transformed lines which express high levels of the 0.9Kb and 1.0Kb Set 1-related RNAs are also the most tumorigenic when injected into nude mice. If this correlation holds true for other SV40-transformed cell lines it will offer further evidence of a key role in transformation for the Set 1-related genes.

Thus the available evidence is consistent with the notion that activation of the two families of cellular genes discussed in this Section is involved in the mechanism of transformation. It is further proposed that the Set 1 related genes play a central role in

SV40-transformation and that Set 2-related gene activation may enhance transformation in some cases, allowing the progression towards a more malignant state.

How could gene activation be provoked by SV40?

There are two obvious answers to this question. Firstly, some viral gene product, such as large T-antigen, might interact with cellular genes; this could be done either directly by binding to the DNA or indirectly *via* some cellular intermediate such as a complex with another protein. Certainly large T-antigen is known to possess many of the properties expected of a protein with such a role (see Section 1.3). The results of the experiments described in Sections 5.6 and 5.7 are also consistent with the idea that the Set 1-related genes are activated by a mechanism that involves large T-antigen. Thus large T polypeptide might interact directly with control sequences surrounding these genes to effect this elevated expression or alternatively it might serve to trigger a cellular mechanism which led to the induction of the Set 1-related gene(s). It would clearly be of interest to isolate clones of these genes and their flanking regions to allow T-antigen binding experiments (Tjian, 1978; Shalloway *et al.*, 1980) to be performed. It would also be of interest to test whether the presence of a dispersed repetitive DNA element within the poly A^+ RNA sequence of at least one of the Set 1-related transcripts is involved in the activation of Set 1 gene expression by SV40.

The second obvious mechanism for gene activation centres upon the idea that some SV40 function can increase the frequency with which genes become converted to an active state. Since transformation mediated by large T-antigen almost certainly involves interaction between the large T polypeptide and the cellular machinery for DNA

replication and cell division (see Chapter 1), it is possible to conceive of ways in which this interaction might lead to gene activation by an indirect route. For example, large T-antigen might cause origins of replication to fire out of synchrony leading to amplification of local regions of DNA. Such amplifications could become fixed by unequal crossover and this could lead to increased expression of genes carried within that local region. If this resulted in a selectable phenotypic change, cells containing the amplification could then become selected during passage of the cell line. In addition, rearrangements of the proviral DNA have been observed in SV40 transformed cells, (Hiscott et al., 1980, 1981; Sager et al., 1981). If this phenomenon also occurred in the cellular DNA it could lead to the activation of cellular genes, for example by transposing them to a site on the genome at which their expression was stimulated. It is quite likely that rearrangements of cellular DNA could also result from asynchronous DNA replication. A third possibility is that insertion of the SV40 genome adjacent to a cellular gene might serve to alter its expression (see Section 1.3). All these possibilities are testable, both by genomic DNA gel transfer hybridization experiments and by cloning of the relevant genes and their flanking regions, although it should be mentioned that the experiment of Fig. 5.22 argues that local DNA amplification is probably not directly involved in the activation of the Set 2-related genes.

The essential difference between the two routes to gene activation proposed here is that the first would be specific for particular genes whereas gene activation by the second route would be a general phenomenon but that cells which activated particular cellular genes would have a selective advantage over others, under the

conditions of growth and/or selection employed, and is therefore consistent with the view that SV40 exerts a tumour-promoting activity upon the host.

Expression of the mitochondrial genome in SV40-transformed cells

The nature of the altered mitochondrial gene expression observed in Chapter 5 is as yet obscure. However, it is perhaps appropriate to discuss some of the possible explanations for these observations. pAG25 has an unusually large, 9.5Kb,insert whose derivation is not yet clear but two main possibilities are apparent. Firstly, the insert of pAG25 could be derived by cloning of a reverse transcript of a large precursor mitochondrial RNA. Consistent with this notion is the fact that one end of the clone corresponds to approximately the region of the DNA in which transcription may be initiated (Bibb et al., 1981). Alternatively, pAG25 could be derived from a low level of contaminating DNA in the mRNA preparation. If the high molecular weight species hybridizing to the mitochondrial clones (Section 5.3) represented incompletely processed mitochondrial RNAs, it is possible that a probe of the type described in Fig.5.17 could be isolated which was specific for material such as tRNA sequences (Bibb et al., 1981) normally edited out of the mitochondrial RNA during processing and hence not present in RNA extracted from normal cells. This probe would then detect a clone like pAG25 which also contained these sequences. Similarly the clones of Sets 3 and 4 could be derived from incompletely processed RNA. Alternatively these clones might be derived from an altered subpopulation of mitochondria present in the transformed cell line, a concept that might explain the appearance of faint minor bands in DNA gel transfer hybridization experiments (see Fig. 5.24). An assessment of the significance of the altered mitochondrial gene expression observed in this study must therefore

await a more detailed analysis of the structure and relationship to the mitochondrial genome of the Set 3, Set 4 and pAG25 clones.

6.3 Conclusion

Presumably all mammalian viral genes, including those of SV40 itself, have been acquired from the host genome at some stage of evolution. Thus SV40 large T-antigen could have evolved from a cellular protein involved in the control of DNA replication and cell proliferation. Similarly small t-antigen might have evolved from a host serum growth factor; this could perhaps explain some of the observations discussed in Section 1.2 regarding the role of small t-antigen in transformation. In light of these hypotheses it is interesting to consider the observed correlation between the expression of the 1.85Kb Set 1-related transcript and that of the viral early mRNAs observed in cells transformed by SV40(tsA) mutants (Section 5.7). This consideration, together with the generality of the induction of the 1.85Kb Set 1-related transcript in SV40-transformed cells, invites speculation that the cellular gene that encodes the 1.85Kb RNA is evolved from the same ancestral gene as the viral transforming gene and hence is structurally and functionally related to the viral early region. Although no homology between the Set 1 clones and SV40 has been detected by DNA hybridization analysis (my unpublished observations), this possibility cannot be discounted until more sensitive tests, such as DNA sequence determination, have been applied. Indeed it must be mentioned that a 68000 molecular weight cellular protein has been identified which is immunologically related to large T-antigen and the expression of this protein is associated with cell proliferation (Lane and Hoeffler, 1980; D.P. Lane, personal communication).

It was proposed in Section 6.2 that, *via* a mechanism that is as yet obscure, SV40 transformation could result in the activation of cellular genes such as those that encode the Set 2-related transcripts; this would then confer a selective advantage upon cells expressing these gene products. The fact that two of the three SV3T3 cell lines studied in Section 5.6 which most efficiently form tumours when injected into animals also express the Set 2-related RNAs suggests that they may be the products of cellular transforming gene(s). In this regard it should be mentioned that Eva *et al.* (1982) observed transcripts homologous to several known retroviral "*onc*" genes (see Section 1.4) in a number of human tumour cell lines. It would clearly be of interest to determine whether the Set 2-related clones show any homology with a known retroviral transforming gene. Perhaps it is significant that in several cases Eva *et al.* (1982) found that the "*onc*"-specific probe detected several cellular transcripts giving a pattern of hybridization similar to that observed when Set 2-specific probes were hybridized to RNA extracted from SV40 transformed cells (see Section 5.6).

Clarification of the roles of the Set 1- and Set 2-related genes in transformation may be provided by three main lines of research.

(a) Further investigation of the generality of expression of transcripts homologous to clones of Sets 1 and 2 in other cell types, including lines transformed by other viruses, chemical carcinogens and cell lines derived from spontaneously arising tumours. Studies on the expression of these genes during development may also be of interest.

(b) Investigation of the transforming potential of the Set 1- and Set 2-related genes. Genomic DNA clones could be isolated and used in DNA transfection experiments (Wigler *et al.*, 1978; Shih *et*

al., 1979, 1981; Cooper et al., 1980, 1981; Murray et al., 1981). Alternatively, the mRNA coding regions derived either from genomic or from cDNA clones could be fused to a suitable retrovirus promoter (Oskarsson et al., 1980) before introduction into cells.

(c) More detailed structural and functional analysis of both genomic and cDNA clones, including DNA sequencing (Maxam and Gilbert, 1980) and T-antigen binding (Tjian, 1978; Shalloway et al., 1980; Myers et al., 1981a, b).

The true significance of the observations described here will not be clear until many experiments of the type described above have been performed, nevertheless the striking correlation between tumourigenicity and elevated expression of Set 1- and Set 2-related genes suggests that such further experimentation may prove to be very rewarding.

REFERENCES

References

Abrahams, P.J., Mulder, C., van de Voorde, A., Warnaar, S.O. and van der Eb, A.J. (1975) J. Virol. 16, 818.

Alt, F.W., Kellems, R.D., Bertino, J.R. and Schimke, R.T. (1978) J. Biol. Chem. 253, 1357.

Alwine, J.C., Kamp, D.J. and Stark, G.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5350.

Baird, W.M., Sedgwick, J.A. and Boutwell, R.K. (1971) Cancer Res. 31, 1434.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell 27, 299.

Baylin, S.B. and Mendelsohn, G. (1980) Endocrine Rev. 1, 1.

Berk, A.J. and Sharp, P.A.(1978) Proc. Natl.Acad. Sci. USA 75, 1274.

Bibb, M.J., van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167.

Bishop, J.O., Rosbash, M. and Evans, D.J. (1974a) J. Mol. Biol. 85, 75.

Bishop, J.O., Morton, J.G., Rosbash, M. and Richardson, M. (1974b) Nature 250, 199.

Bolivar, F., Rodriguez, R.L., Green, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.(1977) Gene 2, 95.

Botchan, M., Topp, W. and Sambrook, J. (1976) Cell 9, 269.

Bouck, N., Beales, N., Shenk, T., Berg, P. and di Mayorca, G. (1978) Proc. Natl. Acad. Sci. USA 75, 2473.

Brockman, W.W. (1978) J. Virol. 25, 860.

Brockman, W.W., Gutai, M.W. and Nathans, D. (1975) Virology 66, 36.

Brugge, J.S. and Butel, J.S. (1975) J. Virol. 15, 619.

Cairns, J. (1981) Nature 189, 353.

Casey, J. and Davidson, N. (1977) Nucleic Acids Res. 4, 1539.

Chou, J.Y., Avila, J. and Martin, R.G. (1974) J. Virol. 14, 116.

Chou, J.Y. and Martin, R.G. (1975) J. Virol. 15, 145.

Clark, L. and Carbon, J. (1976) Cell 9, 91.

Clayton, C.E. and Rigby, P.W.J. (1981) Cell 25, 547.

Colburn, N.H., Former, B.F., Nelson, K.A. and Yuspa, S.H. (1979) Nature 281, 589.

Conrad, S.E. and Botchan, M.R. (1981) J. Supramol. Struct. Cell Biochem. Supplement 5, 445.

Cooper, G.M., Okenquist, S. and Silverman, L. (1980) Nature 284, 418.

Cooper, G.M. and Neiman, P.E. (1981) Nature 287, 656.

Crampton, J.M., Humphries, S., Woods, D. and Williamson, R. (1980) Nucl. Acid Res. 8, 6007.

Crampton, J.M., Davies, K.E. and Knapp, T.F. (1981) Nucleic Acids Res. 9, 3821.

Crawford, L.M., Lane, D.P., Denhardt, D.T., Harlow, E.E., Nicklin, P.M., Osborn, K. and Pim, D.C. (1980) Cold Spring Harbor Symp. Quant. Biol. 44, 179.

Denhardt, D.T. (1966) Biochem. Biophys. Res. Commun. 23, 641.

Dicker, P. and Rozengurt, E. (1978) Nature 276, 723.

Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M. and Darnell, J.E. Jr. (1981) Cell 23, 731.

Diamond, L., O'Brien, T. and Baird, M. (1980) Adv. Cancer Res. 32,1.

Efstratiadis, A., Kafatos, F.C., Maxam, A.M. and Maniatis, T. (1976) Cell 7, 279.

Efstratiadis, A. and Villa-Komaroff, L. (1979) in Genetic Engineering 1, 15. (Setlow, J.K. and Hollaender, A. Eds.) Plenum Press, N.Y.

Fanning, E., Nowak, B. and Burger, C. (1981) J. Virol. 37, 92.

Favaloro, J., Triesman, R. and Kamen, R. (1980) in Methods in Enzymol. 65, 718. (Maldave, K. and Grossman, L., Eds.) Academic Press, N.Y.

Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., van de Voorde, A., van Heuverswyn, H., van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978) Nature 273, 113.

Fluck, M.M. and Benjamin, T.L. (1979) Virology 96, 205.

Frankel, A.E. and Fischinger, P.J. (1976) Proc. Natl. Acad. Sci. USA 73, 3705.

Gelb, L.D., Kohne, D.E. and Martin, M.A. (1971) J. Mol. Biol. 57, 129.

Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) Nucleic Acid Res. 7, 2115.

Gillespie, D. and Spiegelman, S. (1965) J. Mol. Biol. 12, 829.

Goff, S.P., Gilboa, E., Witte, O.N. and Baltimore, D. Cell 22, 777.

Gottesman, M.M. (1978) Proc. Natl. Acad. Sci. USA 75, 2767.

Gottesman, M.M. and Sobel, M.E. (1980) Cell 19, 449.

Graf, T. and Beug, H. (1978) Biochim. Biophys. Acta 516, 269.

Graham, F.L., Abrahams, P.J., Mulder, C., Heijneker, H.L., Warnaar, S.O., de Vries, F.A.J., Fiers, W. and van der Eb, A.J. (1974) Cold Spring Harbor Symp. Quant. Biol. 39, 637.

Groudine, M. and Weintraub, H. (1980 Proc. Natl. Acad. Sci. USA 77, 5351.

Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961.

Gutai, M.W. and Nathans, D. (1978) J. Mol. Biol. 126, 259.

Hanahan, D. and Meselson, M. (1980) Gene 10, 63.

Hassel, J.A., Lukanidin, E., Fey, G. and Sambrook, J. (1978) J. Mol. Biol. 120, 209.

Hastie, N.D. and Bishop, J.O. (1976) Cell 9, 761.

Hayman, M.J. (1981) J. Gen. Virol. 52, 1.

Hayward, W.S., Neel, B.G. and Astrin, S.M. (1981) Nature 290, 475.

Henry, P., Black, P.H., Oxman, M.N. and Weissman, S.M. (1966) Proc. Natl. Acad. Sci. USA 56, 1170.

Hiscott, J.B. and Defendi, V. (1979) J. Virol. 30, 590.

Hiscott, J., Murphy, D. and Defendi, V. (1980) Cell 22, 535.

Hiscott, J.B., Murphy, D. and Defendi, V. (1981) Proc. Natl. Acad. Sci. USA 78, 1736.

Jay, G., de Leo, A.B., Appella, E., Dubois, G.C., Law, L.W., Khoury, G. and Old, L.J. (1979) Cold Spring Harbor Symp. Quant. Biol. 44, 659.

Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Deininger, P.L. and Schmid, C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398.

Kacian, D.L., Spiegelman, S., Bank, A., Terada, M., Metafora, S., Dow, L. and Maks, P.A. (1972) Nature New Biol. 235, 167.

Kafatos, F.C., Jones, C.W. and Efstratiadis, A. (1979) Nucleic Acid. Res. 7, 1541.

Ketner, G. and Kelly, T.J. (1976) Proc. Natl. Acad. Sci. USA 73, 1102.

Kimura, G. and Itagaki, A. (1975) Proc. Natl.Acad. Sci. USA 72, 673.

Kit, S., de Torres, R.A., Dubbs, D.R. and Salvi, M.L. 1967) J. Virol. 1, 738.

Krontiris, T.G. and Cooper, G.M. (1981) Proc. Natl. Acad. Sci. USA 78, 1181.

Lane, D.P. and Crawford, L.V. (1979) Nature 278, 261.

Prives, C. and Beck, Y. (1977) INSERM Coloq. 69, 175.

Laskey, R.A. and Mills, R.D. (1976) FEBS Letters 82, 314.

Laskin, J.D., Mufson, A., Piccinini, L., Engelhardt, D.L. and Weinstein, B. (1981) Cell 25, 441.

Laszlo, A., Radke, K., Chin, S. and Bissel, M.J. (1981) Proc. Natl. Acad. Sci. USA 78, 6241.

Lehrach, H., Diamond, D., Wazney, J.M. and Boedtker, H. (1977) Biochemistry 16, 4743.

Leis, J.P. and Hurwitz, J. (1974) in Methods in Enzymology (Grossman, L. and Moldave, K. eds) 29E, 143. Academic Press, N.Y.

Lewis, A.M., Jr. and Martin, R.G. (1979) Proc. Natl. Acad. Sci. USA 76, 4299.

Linzer, D.I.M. and Levine, A.J. (1979) Cell 17, 43.

Maniatis, T. and Ptashne, M. (1973) Proc. Natl. Acad. Sci. USA 70, 1531.

Maniatis, T., Sim, G.K., Efstratiadis, A. and Kafatos, F.C. (1976) Cell 8, 163.

Martin, R.G., Setlow, V.P., Chepelinsky, A.B., Seif, R., Lewis, A.M., Jr., Scher, C.D., Stiles, C.D. and Avila, J. (1979a) Cold Spring Harbor Symp. Quant. Biol. 44, 311.

Martin, R.G., Setlow, V.P. and Edwards, C.A.F. (1979b) J. Virol. 31, 596.

Martin, R.G., Setlow, V.P., Edwards, C.A.F. and Vembu, D. (1979c) Cell 17, 635.

Martin, R.G. (1981) Adv. in Cancer Res. 34, 1.

Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymol. 65, 499. (Maldave, K. and Grossman, L. Eds.) Academic Press, N.Y.

McCormick, F. and Harlow, E. (1980) J. Virol. 34, 213.

McCutchan, T. and Singer, M. (1981) Proc. Natl.Acad. Sci. USA 78,95.

McDonnell, M.W., Simon, M.N. and Studier, F.W. (1977) J. Mol. Biol. 110, 119.

Mayer, B.J., Kleid, D.G. and Ptashne, M. (1975) Proc. Natl. Acad. Sci. USA 72, 4785.

Monahan, J.J., Harris, S.E., Woo, S.L.C., Robberson, D.L. and O'Malley, B.W. (1976) Biochemistry 15, 233.

Mougneau, E., Birg, F., Rassoulzadegan, M. and Cuzin, F. (1980) Cell 22, 917.

Murray, M.J., Shilo, B.-Z., Shih, C., Cowing, D,, Hsu, H.W. and Weinberg, R.A. (1981) Cell 25, 355.

Myers, R.M., Williams, R.C. and Tjian, R. (1981a) J. Mol. Biol. 148, 347.

Myers, R.M., Rio, D.C., Robbins, A.K. and Tjian, R. (1981b) Cell 25, 373.

Neel, B.G., Hayward, W.S., Robinson, H.L., Fang, J. and Astrin, S.M. (1981) Cell 23, 323.

Noori-Daloii, M.R., Swift, R.A., Kung, H.-J., Crittenden, L.B. and Witter, R.L. (1981) Nature 294, 574.

Noyes, B.E. and Stark, G.R. (1975) Cell 5, 301.

Opperman, H., Levinson, A.D., Varmus, H.E., Levintow, L. and Bishop, J.M. (1979) Proc. Natl. Acad. Sci. USA 76, 1804.

Oren, M., Maltzman, W. and Levine, A.J. (1981) Mol. Cell. Biol. 1, 101.

Oskarsson, M., McClements, W.L., Blair, D.G., Maizel, J.V. and van de Woude, G.F. (1980) Science 207, 1222.

Paucha, E., Harvey, R. and Smith, A.E. (1978a) J. Virol. 28, 154.

Paucha, E., Mellor, A., Harvey, R., Smith, A.E., Hewick, R.M. and Waterfield, M.D. (1978b) Proc. Natl. Acad. Sci. USA 75, 2165.

Payne, G.S., Courtneidge, S.A., Crittenden, L.B., Fadly, A.M., Bishop, J.M. and Varmus, H.E. (1981) Cell 23, 311.

Payne, G.S., Bishop, J.M. and Varmus, H.E. (1982) Nature 295, 209.

Pitot, M.C. and Sirica, A.G. (1980) Biochim. Biophys. Acta. 605, 191.

Ponder, B.A.J. (1980) Biochim. Biophys. Acta. 105, 369.

Prives, C., Beck, Y., Gidoni, D., Oren, M. and Shure, H. (1977) Cold Spring Harbor Symp. Quant. Biol. 44, 123

Quinlan, T.J., Beeler, G.W. Jr., Cox, R.F., Elder, P.K., Moses, H.L. and Getz, M.J. (1978) Nucl. Acid. Res. 5, 1611.

Reddy, V.B., Thimmappaya, B., Dhar, R., Subramanian, K.N., Zain, B.S., Pan, J., Ghosh, P.K., Celma, M.L. and Weissman, S.M. (1979) J. Virol. 30, 279.

Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237.

Rigby, P.W.J., Chia, W., Clayton, C.E. and Lovett, M. (1980) Proc. Roy. Soc. Lond. B. 210, 437.

Rio, D., Robbins, A., Myers, R. and Tjian, R. (1980) Proc. Natl. Acad. Sci. USA 77, 5706.

Risser, R. and Pollack, R. (1974) Virology 59, 477.

Ritzi, E. and Levine, A.J. (1969) J. Virol. 5, 686.

Roberts, R.J. (1978) Gene 4, 183.

Ross, J., Aviv, H., Scolnick, E. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 264.

Roussel, M., Saule, S., Legrou, C., Rommens, C., Beug, H., Graf, T and Stehelin, D. (1979) Nature 281, 452.

Roychoudry, R., Jay, E. and Wu, R. (1976) Nucl. Acid. Res. 3, 101.

Sanger, F. and Coulson, A.R. (1978) FEBS Letters 87, 107.

Sager, R., Anisowicz, A. and Howell, N. (1981) Cell 23, 41.

Sambrook, J., Westphal, H., Srinivasan, P.R. and Dulbecco, R. (1968) Proc. Natl. Acad. Sci. USA 60, 1288.

Sauer, G. and Defendi, V. (1966) Proc. Natl. Acad. Sci. USA 56, 452.

Scher, C.D. (1971) Virology 46, 956.

Seif, R. and Martin, R.G. (1979a) J. Virol. 31, 350.

Seif, R. and Martin, R.G. (1979b) J. Virol. 32, 979.

Setlow, V.P., Persico-DiLauro, M., Edwards, C.A.F. and Martin, R.G. (1980) Virology 101, 250.

Shalloway, D., Kleinberger, T. and Livingston, D.M. (1980) Cell 20, 411.

Shih, C., Shilo, B., Goldfarb, M.P., Dannenberg, A. and Weinberg, R.A. (1979) Proc. Natl. Acad. Sci. USA 76, 5714.

Shih, C., Padhy, L.C., Murray, M.J. and Weinberg, R.A. (1981) Nature 290, 261.

Shin, S.-I., Freedman, V.H., Risser, R. and Pollack, R. (1975) Proc. Natl. Acad. Sci. USA 72, 4435.

Smith, H.S., Scher, C.D. and Todaro, G.J. (1971) Virology 44, 359.

Smith, H.S., Gelb, L.D. and Martin, M.A. (1972) Proc. Natl. Acad. Sci. USA 69, 152.

Smith, D.F., Searle, P.F. and Williams, J.G. (1979) Nucleic Acids Res. 6, 487.

Soprano, K.J. and Baserga, R. (1980) Proc. Natl. Acad. Sci. USA 77, 1566.

Soprano, K.J., Dev, V.G., Croce, C.M. and Baserga, R. (1980a) Proc. Natl. Acad. Sci. USA 76, 3885.

Soprano, K.J., Rossini, M., Croce, C. and Baserga, R. (1980b) Virology 102, 317.

Southern, E.M. (1975) J. Mol. Biol. 98, 503.

Stark, G.R. and Williams, J.G. (1979) Nucleic Acids Res. 6, 195.

Stehelin, D., Varmus, H.E., Bishop, J.M. and Vogt, P.K. (1976) Nature 260, 170.

Steinberg, B.M. and Pollack, R. (1979) Virology 99, 302.

Subramanian, K.R. and Shenk, T.E. (1978) Nucleic Acid Res. 5, 3655.

Sutcliffe, G. (1978) Cold Spring Harbor Symp. Quant. Biol. 63, 81.

Taylor, J.M., Illmensee, R. and Summers, J. (1976) Biochim. Biophys. Acta. 442, 324.

Tegtmeyer, P. (1972) J. Virol. 10, 591.

Tegtmeyer, P. (1975) J. Virol. 15, 613.

Tegtmeyer, P., Schwartz, M., Collins, J.K. and Rundell, K. (1975) J. Virol. 16, 168.

Tegtmeyer, P., Rundell, K. and Collins, J.K. (1977) J. Virol. 21, 647.

Thomas, P.S. (1980) Proc. Natl. Acad. Sci USA 77, 5201.

Thummel, C., Tjian, R. and Grodzicker, T. (1981) Cell 23, 825.

Tjian, R. (1978) Cell 13, 165.

Todaro, G. and Green, H. (1963) J. Cell. Biol. 17, 299.

Tooze, J. (1980) DNA Tumor Viruses. Molecular Biology of Tumor Viruses, Second Edition, Part 2. Cold Spring Harbor Laboratory.

Twigg, A.J. and Sherratt, D. (1980) Nature 283, 216.

Varshavsky, A. (1981a) Proc. Natl. Acad. Sci. USA 78, 3673.

Varshavsky, A. (1981b) Cell 25, 561.

Verma, I., Temple, G.F., Fan, H. and Baltimore, D. (1972) Nature New Biol. 235, 163.

Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc. Natl. Acad. Sci. USA 76, 3683.

Wickens, M.P., Buell, G.N. and Schimke, R.T. (1978) J. Biol. Chem. 253, 2471.

Wigler, M. and Weinstein, I.B. (1976) Nature 259, 232.

Wigler, M., Pellicer, A., Silverstein, S. and Axel, R. (1978) Cell 14, 725.

Williams, J.G., Hoffman, R. and Penman, S. (1977) Cell 11, 901.

Woods, D., Crampton, J., Clarke, B. and Williamson, R. (1980) Nucl. Acid Res. 8, 5157.

Yuspa, S.H., Ben, T., Patterson, E., Michael, D., Elgjo, K. and Hennings, M. (1976a) Cancer Res. 36, 4062.

Yuspa, S.H., Lichti, U., Ben, T., Patterson, E., Hennings, H., Slaga, T.J., Colburn, N. and Kelsey, W. (1976b) Nature 262, 402.