

MEDICAL LIBRARY,
ROYAL FREE HOSPITAL
HAMPSTEAD.

ProQuest Number: U049082

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

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




ProQuest U049082

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

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

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



2807400960

ROYAL FREE THESIS 1991

TRANSCRIPTIONAL REGULATION OF RETROVIRUS-LIKE VL30 GENETIC
ELEMENTS

LYNNE TRACEY EATON

SUBMITTED FOR THE DEGREE OF PhD.

SCHOOL
ROYAL FREE HOSPITAL OF MEDICINE

1990

This thesis is dedicated to Andy O'Brien

ABSTRACT

The nucleotide sequence of the long terminal repeats (LTRs) of the retrovirus-transmissible mouse VL30 cDNA clones, NVL-1 and NVL-2, were determined and compared with that of the prototype clone NVL-3. The three LTRs shared a typical U3 R U5 structure and showed the unusual features of redundancy in the tRNA^{sly} primer binding site and adjacent inverted repeat. The NVL-1 and NVL-2 LTRs were almost identical and differed from the NVL-3 LTR in the U3 domain harbouring transcriptional regulatory determinants.

The VL30 elements were serum responsive and showed elevated levels of expression in many transformed cell lines. S1 nuclease analysis showed that all the NVL VL30 elements responded to cellular transformation, but only NVL-1/2 elements were serum responsive. Both types of response were mediated via protein kinase C-independent pathways.

Transient expression assays showed that the NVL-1/2 and NVL-3 U3 domains possess promoter and enhancer activities. The U3 region of both types of LTR contained enhancer determinants important in controlling the response to N-ras as well as sequences important for controlling basal levels of expression. However the U3 domain of the NVL-1/2 LTRs did not confer serum responsiveness.

These results showed that the U3 domains of the NVL VL30 LTRs contain cis acting regulatory determinants important in controlling basal levels of expression and the response to cellular transformation. However the mechanisms controlling the serum response are at least in part dissociable from those responsible for cell transformation regulated expression.

TABLE OF CONTENTS

ABSTRACT	3
INTRODUCTION	15
<u>I.1 REGULATION OF GENE EXPRESSION</u>	16
I.1.1 TRANSCRIPTIONAL REGULATORY ELEMENTS	16
I.1.2 INDUCIBLE ENHANCERS	19
I.1.3 TRANSACTING FACTORS	20
I.1.4 POST-TRANSCRIPTIONAL MECHANISMS FOR CONTROLLING GENE EXPRESSION	23
<u>I.2. POLYPEPTIDE GROWTH FACTORS AND SIGNAL TRANSDUCTION</u>	24
I.2.1 GROWTH FACTOR INDUCED EVENTS AT THE CELL SURFACE AND IN THE CYTOPLASM	25
I.2.2 GROWTH FACTOR MEDIATED EVENTS IN THE NUCLEUS	29
<u>I.3 RETROELEMENTS</u>	31
I.3.1 NONVIRAL RETROPOSONS	31
I.3.1 RETROVIRUSES	32
I.3.2 RETROTRANSPOSONS	34
<u>I.4. MOUSE VL30 ELEMENTS</u>	36
I.4.1 CHARACTERISTIC FEATURES OF VL30 ELEMENTS	37
I.4.2 EXPRESSION OF VL30 ELEMENTS	38
I.4.3 PROMOTER STRENGTHS OF VL30 LTRs	39
I.4.4 THE NVL VL30 ELEMENTS	39
<u>I.5 AIMS</u>	40

MATERIALS	42
<u>MA.1 COMMERCIAL PRODUCTS</u>	43
<u>MA.2 SOLUTIONS AND BUFFERS</u>	48
METHODS	51
<u>M.1 CELL CULTURE</u>	52
M.1.1 MAINTENANCE OF CELL LINES	52
M.1.2 PASSAGING OF CELL LINES	52
M.1.3 STORAGE OF CELL LINES	53
M.1.4 CULTURING CELLS FROM FROZEN STOCKS	53
M.1.5 EXPERIMENTAL PROCEDURES WITH CULTURED CELLS	53
M.1.6 RETROVIRUS MEDIATED GENE TRANSFER	54
<u>M.2 PREPARATION OF PLASMID DNA</u>	54
M.2.1 SMALL SCALE PLASMID PREPARATIONS	55
M.2.2 LARGE SCALE PLASMID PREPARATIONS	55
M.2.2.1 Amplification of Low Copy Number Plasmids	55
M.2.2.2 Large Scale Cleared Lysate	56
M.2.2.3 G50 Column Purification	56
M.2.2.4 Purification of Plasmid DNA by PEG 8000 Precipitation	57
M.2.3 STORAGE OF BACTERIAL STOCKS	57
<u>M.3 MANIPULATION OF CLONED DNA</u>	57
M.3.1 RESTRICTION ENZYME DIGESTS	58
M.3.2 AGAROSE GEL ELECTROPHORESIS	58
M.3.3 ETHANOL PRECIPITATION OF NUCLEIC ACIDS	59
M.3.4 PHENOL/CHLOROFORM EXTRACTION	59
M.3.5 CTAB/BUTANOL EXTRACTION	60

M.3.6	PREPARATION OF VECTOR DNA	61
M.3.7	REMOVAL OF 5' PHOSPHATE GROUPS	61
M.3.8	PREPARATION OF INSERT DNA	61
M.3.9	LIGATION OF VECTOR AND INSERT	62
M.3.9.1	Ligations without LMA	62
M.3.9.2	Ligations in LMA	62
M.3.10	USE OF ARTIFICIAL LINKERS	62
M.3.10.1	Preparation of DNA for linker ligation	62
M.3.10.2	Phosphorylation of linkers	63
M.3.10.3	Ligation of linkers and insert DNA	63
M.3.10.4	Removal of excess linkers	64
M.3.11	BACTERIAL TRANSFORMATIONS	64
M.3.11.1	Preparation of competent cells	64
M.3.11.2	Transformation	64
<u>M.4</u>	<u>OLIGONUCLEOTIDE SYNTHESIS</u>	65
<u>M.5</u>	<u>PREPARATION OF ³²P LABELLED DNA PROBES</u>	66
M.5.1	RANDOMLY PRIMED PROBES	66
M.5.2	END-LABELLING OF DNA RESTRICTION FRAGMENTS	67
M.5.3	END-LABELLING OF OLIGONUCLEOTIDES	68
<u>M.6</u>	<u>COLONY SCREENING</u>	68
M.6.1	PREPARATION OF FILTERS	68
M.6.2	HYBRIDIZATION OF THE FILTERS	69
<u>M.7</u>	<u>DIDEOXY CHAIN TERMINATION DNA SEQUENCING</u>	69
M.7.1	PREPARATION OF M13 SUBCLONES	69
M.7.1.1	Cloning Manipulations	69
M.7.1.2	Transformation of <i>E.coli</i> with M13 Bacteriophage	69

M.7.2	PREPARATION OF SINGLE STRANDED DNA	
	TEMPLATES	70
M.7.3	SEQUENCING	70
	M.7.3.1 Annealing of Template and Primer	71
	M.7.3.2 Sequencing Reactions	71
	M.7.3.3 Gel Electrophoresis	72
M.7.4	ANALYSIS OF SEQUENCING GELS	72
<u>M.8</u>	<u>ANALYSIS OF CELLULAR RNA</u>	72
M.8.1	PREPARATION OF TOTAL CELLULAR RNA	72
M.8.2	NORTHERN BLOTS	73
M.8.3	RNA DOT BLOTS	74
M.8.4	HYBRIDIZATION OF NORTHERN AND RNA DOT BLOTS	74
M.8.5	STRIPPING PROBES FROM FILTERS	74
<u>M.9</u>	<u>NUCLEAR RUN ON TRANSCRIPTION</u>	75
M.9.2	PREPARATION OF ISOLATED NUCLEI	75
M.9.3	NUCLEAR RUN ON	75
M.9.4	PREPARATION OF DNA DOT BLOTS	76
M.9.4	HYBRIDIZATIONS	76
<u>M.10</u>	<u>TRANSIENT EXPRESSION ASSAYS</u>	77
M.10.1	TRANSIENT TRANSFECTIONS	77
M.10.2	CAT ASSAYS	78
<u>M.11</u>	<u>PREPARATION OF STABLE TRANSFECTANTS</u>	79
M.11.2	TRANSFECTIONS	79
M.11.3	SERUM RESPONSE EXPERIMENTS	79
<u>M.12</u>	<u>PRIMER EXTENSION</u>	80
M.12.1	PREPARATION OF PROBE	80
M.12.2	RNA PRIMER ANNEALING	80
M.12.3	PRIMER EXTENSION	81

RESULTS	82
<u>R.1 NUCLEOTIDE SEQUENCE DIVERSITY AMONGST</u>	
<u>DISTINCT VL30 ELEMENTS</u>	83
R.1.1 SEQUENCE HETEROGENEITY WITHIN THE NVL	
SUBGROUP OF VL30 ELEMENTS	83
R.1.2 COMPARISON BETWEEN NVL VL30 LTRs AND OTHER	
VL30 LTRs	89
R.1.3 HOMOLOGY WITH OTHER SEQUENCES	91
<u>R.2 VL30 ELEMENTS ARE RESPONSIVE TO STIMULATION</u>	
<u>BY SERUM GROWTH FACTORS</u>	91
R.2.1 SERUM STIMULATION OF VL30 ELEMENTS IN	
QUIESCENT NIH 3T3 CELLS	92
R.2.2 THE RESPONSE TO SERUM IS REGULATED AT THE	
TRANSCRIPTIONAL LEVEL	92
R.2.3 DIFFERENTIAL REGULATION OF NVL-3 AND NVL-1/2	
IN RESPONSE TO SERUM	95
R.2.4 THE REGULATION OF VL30 EXPRESSION IS MEDIATED	
VIA SEQUENCES WITHIN THE ELEMENT	97
R.2.4 THE SERUM RESPONSE IS NOT MEDIATED VIA THE	
PROTEIN KINASE C SIGNAL TRANSDUCTION	
PATHWAY	101
<u>R.3 VL30 ELEMENTS ARE RESPONSIVE TO CELLULAR</u>	
<u>TRANSFORMATION</u>	103
R.3.1 LEVELS OF VL30 TRANSCRIPTS IN CELL LINES	
TRANSFORMED VIA DIFFERENT MECHANISMS	103
R.3.2 AN AUTOCRINE/PARACRINE FACTOR DOES NOT	
MEDIATE UPREGULATION OF VL30 EXPRESSION IN	
TRANSFORMED CELLS	105

R.3.3	BOTH TYPES OF NVL VL30 ELEMENT ARE RESPONSIVE TO TRANSFORMATION	107
<u>R.4</u>	<u>VL30 LTRs POSSESS PROMOTER AND ENHANCER</u>	
	<u>ACTIVITY</u>	109
R.4.1	PREPARATION OF PROMOTER TESTING CONSTRUCTS	109
R.4.2	TRANSFECTION AND CAT ASSAY CONTROL EXPERIMENTS	112
R.4.3	DISTINCT VL30 LTRs HAVE DIFFERENT PROMOTER ACTIVITIES	115
R.4.4	PREPARATION OF ENHANCER TESTING PLASMIDS	117
R.4.5	BOTH VL30 LTRs EXHIBIT ENHANCER FUNCTION	117
<u>R.5</u>	<u>VL30 LTRs ARE TRANSACTIVATED BY ACTIVATED</u>	
	<u>N-RAS</u>	121
R.5.1	TITRATION OF THE TRANSACTIVATION BY ACTIVATED N-RAS	121
R.5.2	AN ENHANCER MEDIATES THE RESPONSE TO ACTIVATED N-RAS	124
R.5.3	TRANSACTIVATION BY ACTIVATED RAS GENES IS GENE DOSAGE DEPENDANT	124
R.5.4	THE PROTEIN KINASE C PATHWAY IS NOT A MAJOR PATHWAY MEDIATING THE RESPONSE TO N-RAS	124
<u>R.6</u>	<u>THE VL30 LTRs ARE NOT SUFFICIENT TO MEDIATE</u>	
	<u>THE SERUM RESPONSE</u>	127
R.6.1	A SERUM RESPONSE WAS NOT OBTAINED IN TRANSIENTLY TRANSFECTED NIH 3T3 CELLS	127
R.6.2	STABLY TRANSFECTED LTRs ARE NOT SERUM RESPONSIVE	128
<u>R.7</u>	<u>VL30-CAT TRANSCRIPTS ARE INITIATED FROM</u>	

<u>WITHIN THE LTR</u>	128
DISCUSSION	132
<u>D.1 STRUCTURAL ANALYSIS</u>	133
D.1.1 STRUCTURAL FEATURES OF THE NVL VL30 LTRs	133
D.1.2 NVL VL30 ELEMENTS ARE STRUCTURALLY DISTINCT FROM OTHER VL30 ELEMENTS	135
D.1.3 STRUCTURAL RELATIONSHIPS WITH OTHER RETROTRANSPOSABLE ELEMENTS	137
<u>D.2 EXPRESSION OF VL30 RNA</u>	138
D.2.1 SERUM RESPONSE	138
D.2.2 ELEVATED EXPRESSION IN TRANSFORMED CELL LINES	140
<u>D.3 PROMOTER AND ENHANCER PROPERTIES OF THE NVL VL30 U3 DOMAINS</u>	142
D.3.1 BASAL LEVELS OF PROMOTER AND ENHANCER ACTIVITY . . .	142
D.3.2 RESPONSE TO N-RAS IS MEDIATED VIA AN ENHANCER FUNCTION	143
D.3.3 LACK OF SERUM RESPONSE IN THE U3 DOMAIN OF NVL-1	144
<u>D.4 INDUCTION OF VL30 EXPRESSION BY SERUM AND RAS IS MEDIATED VIA PROTEIN KINASE C INDEPENDENT PATHWAYS</u>	146
<u>D.5 CONCLUSIONS</u>	148
AKNOWLEDGEMENTS	150
REFERENCES	151

FIGURES

Figure 1. Nucleotide sequencing strategy for NVL-1 and NVL-2.	84
Figure 2. Nucleotide sequence of NVL-1 and NVL-2 LTRs.	85
Figure 3. Comparison between the NVL-1/2 and NVL-3 LTRs.	88
Figure 4. Comparison between VL30 LTRs.	90
Figure 5. VL30 expression in quiescent and serum stimulated NIH 3T3 cells.	93
Figure 6. Nuclear run on transcription analysis of serum induced VL30 expression.	94
Figure 7. S1 nuclease analysis of RNA from NIH 3T3 cells.	96
Figure 8. Analysis of RNA from quiescent and serum stimulated VLNRK cells.	98
Figure 9. Analysis of RNA from Rat1 cell lines.	100
Figure 10. Effects of the Protein kinase C inhibitor H-7 on VL30 expression.	102
Figure 11. Dot blot analysis of RNA from normal and transformed NIH 3T3 cell lines.	104
Figure 12. The effect of conditioned medium from transformed cells on the expression of VL30 elements in normal NIH 3T3 cells.	106
Figure 13. S1 nuclease analysis of RNA from spontaneously transformed NIH 3T3 cells.	108
Figure 14. Preparation of the 3' LTR-CAT constructs.	110
Figure 15. Preparation of 5' LTR-CAT constructs.	111
Figure 16. Linear range of CAT assay.	113

Figure 17. Optimum input of DNA in transfections.	114
Figure 18. Promoter strengths of the NVL-1 and NVL-3 LTRs.	116
Figure 19. Preparation of enhancer testing constructs.	118
Figure 20. Enhancer strengths of the NVL-3 and NVL-1 LTRs.	119
Figure 21. VL30 promoter and enhancer function in N-ras transformed cells.	122
Figure 22. Transactivation of VL30 promoters in Ha-c-ras transformed cells.	125
Figure 23. The effect of the protein kinase C inhibitor H-7 on N-ras transactivation.	126
Figure 24. Expression of stably transfected LTR-CAT constructs in quiescent and stimulated NIH 3T3 cells.	129
Figure 25. Primer extension analysis of NIH 3T3 cells stably transfected with VL30 promoter and enhancer testing constructs.	131

TABLES

Table M.1	N° mixes	71
Table R.1	Promoter and enhancer strengths of NVL-1 and NVL-3	120
Table R.2	N-ras transactivation of NVL-1 and NVL-3 promoters and enhancers	123

ABBREVIATIONS

AcCoA	- Acetyl Coenzyme A
AP-1	- Activating Protein 1
bp	- base pairs
Budr	- bromodeoxyuridine
BSA	- bovine serum albumin
CTAB	- hexadecyltrimethylammonium bromide
CIAP	- calf intestinal alkaline phosphatase
CAT	- chloramphenicol acetyl transferase
DMEM	- Dulbecco's modified essential medium
DTT	- dithiothrietol
DEPC	- diethylpyrocarbonate
DMSO	- dimethylsulphoxide
FCS	- foetal calf serum
IPTG	- isopropyl- β -D-thio-galactopyranoside
IR	- inverted repeat
Kbp	- kilobase pairs
LMA	- low melting agarose
LTR	- long terminal repeat
MSV	- murine sarcoma virus
MLV	- murine leukaemia virus
MOPS	-3-[N-Morpholino]propanesulfonic acid
NP40	- Nonidet P40
NBS	- newborn calf serum
NRK	- normal rat kidney
nt	- nucleotides
PNK	- polynucleotide kinase
PBS	- phosphate buffered saline
PEG	- polyethylene glycol

RF - replicative form
TLC - thin layer chromatography
TEAA - triethylammonium acetate
TK - thymidine kinase
TPA - 12-O-tetradecanoyl-13-phorbol acetate
Tris - Tris(hydroxymethyl)aminomethane
UPE - upstream promoter element
VRC - vanadyl ribonucleoside complex
X-GAL - 5-bromo-4-chloro-3-indolyl- β -galactoside

INTRODUCTION

I.1 REGULATION OF GENE EXPRESSION

Although the cell contains a huge number of genes only some of them are expressed constitutively. These are the 'housekeeping' genes that are responsible for maintaining essential cellular functions. Other genes show regulated expression. For example some genes are only expressed at certain times during an organism's development or at certain stages during the cell cycle. Other genes respond to extracellular stimuli such as polypeptide growth factors and steroid hormones.

In an effort to understand the processes involved in, for example, an organism's development, neoplastic transformation or the cell cycle, a great deal of work has been done on patterns of gene expression in these situations. An important part of this work has been to try and understand the mechanisms that control gene expression.

By studying individual genes it is hoped that models can be described that will help in understanding the complex pathways involved in controlling various cellular responses. An understanding of when and where a gene is expressed is also important in understanding the function of that particular gene product.

Control of gene expression occurs at a number of levels. This project is particularly concerned with transcriptional regulatory mechanisms, in particular with the DNA sequences required for controlling gene expression. There are however important post-transcriptional that are essential for proper control of the expression of many gene products.

I.1.1 TRANSCRIPTIONAL REGULATORY ELEMENTS

From the analysis of a growing number of viral and cellular genes it has been possible to build up a picture of the mechanisms that control gene expression at the transcriptional level. Transcription is controlled

by the binding of specific transacting factors to specific DNA sequences, these factors interact to facilitate the initiation and efficient transcription of the gene by RNA polymerase II (1,2). The exact way that these DNA-protein interactions initiate and control transcription is unknown.

Two main regions have been described that are important for transcriptional regulation. In close proximity to the transcriptional start site is the promoter region (1,3). This region extends for up to 100bp upstream from the start site and consists of two types of element. The TATA box is a purine rich tract that occurs within 20-30bp of the start site (4). It appears to be responsible for accurate initiation of transcription as removal of the TATA box tends to result in initiation of transcription from multiple start sites (4). The other types of regulatory sequence within the promoter region are known as upstream promoter elements (UPEs) (1). Examples of UPEs include the CCAAT box and the G-C rich box. These elements are responsible for efficient initiation of transcription. A characteristic of the UPEs is that they are able to act in either orientation with respect to the gene (1,3).

The second major region is known as the enhancer. Enhancers tend to be found at a much greater distance from the transcriptional start site than the promoter elements. The enhancer is a cis-acting regulatory region that functions in either orientation and from upstream or downstream of the promoter. An enhancer can operate over a distance of many Kbp and can effect expression from an heterologous promoter (5,6). The enhancer's function is to increase the rate of transcription from the promoter (2). The enhancer regions of a number of genes have been studied in great detail and it has been shown that the enhancer is composed of a number of small (5-10bp) units that are the binding sites for specific

proteins (3,7,8). These small modules are called enhansons and they provide an extremely flexible system for controlling gene expression (3,8). Different combinations of the enhansons occur in different gene regulatory regions providing each gene with its own individual enhancer region but utilizing a relatively low number of different sequence motifs and hence a relatively low number of transcription factors (8).

The enhancer region of the simian virus 40 (SV40) early gene has been particularly well studied and has provided much of the understanding of the way in which eukaryotic gene expression is controlled. The SV40 enhancer was originally described as a tandemly repeated 72bp sequence (4). However closer study of the enhancer region has resulted in it being divided into a number of smaller domains (7,8). Each of the SV40 enhancers consists of at least three smaller regions of about 15-20bp which have been termed enhancer elements. These in turn are composed of the smaller subunits known as enhansons (8). The spacing of the enhansons within the enhancer elements is extremely important, but the positional requirements of the enhancer elements themselves is less stringent. Repetition at all levels of organisation within the SV40 enhancer is of great importance with full transcriptional activity of the SV40 enhancer only being achieved if the repetition of the enhansons, enhancer elements and the entire enhancer is seen. The enhancer elements can substitute for each other, in other words the deletion of one of the enhancer elements can be tolerated if another of the elements is duplicated.

Not all gene regulatory regions contain all of these types of sequence. In particular 'housekeeping' genes frequently have different regulatory requirements and may for example lack the TATA box region (9). However these genes may have extensive G-C rich regions containing multiple copies of the G-C box (10).

I.1.2 INDUCIBLE ENHANCERS

Many genes are responsive to external stimuli such as growth factors, steroid hormones, heat shock, and the presence of heavy metals. A number of inducible genes have now been studied and it has been found in many cases the response to the external stimulus is mediated via an enhancer function (1,5,6).

Of particular interest are genes that respond when quiescent cells are stimulated with serum or purified growth factors. The regulatory regions of a number of growth factor responsive genes have been studied. Many genes respond very rapidly to this stimulus (early response genes), whereas other genes take several hours to respond.

A particularly well studied example of a gene that is rapidly, but transiently stimulated in response to serum is the c-fos proto-oncogene (11). Basal levels of fos expression are very low, however within a few minutes of growth factor stimulation there is a dramatic increase in expression. Within two hours expression returns to the very low levels seen in unstimulated cells. Analysis of the 5' non-coding region of this gene shows that a short region of dyad symmetry is responsible for mediating the serum response (12). This region is known as the serum response element (SRE) and binds a transacting protein known as the serum response factor (SRF) (13,14). The analysis of another of these rapidly expressed genes, EGR2, shows that it possesses a sequence motif similar to parts of the c-fos regulatory sequences (15). This motif is the CArG box which forms part of the fos SRE. The SRF also binds to the EGR2 gene CArG box. The presence of this motif in more than one early response gene suggests that it may be a common mechanism for controlling these early response genes.

A number of genes showing a less rapid response to serum have also been studied. Many of these genes require protein synthesis in order to respond to serum stimulation suggesting that the gene products of the early response genes are responsible for the later responses shown by this other group of genes. In support of this is the observation that some of these genes possess binding sites for the trans-acting factor c-jun/AP-1 which is itself an early response gene (17,18). In some cases the presence of this protein binding site has been shown to be essential for serum responsiveness. For example the transin gene possesses an AP-1 site that is important in mediating growth factor responsive expression (19). Also the vimentin gene possesses two copies of an AP-1 binding site that are essential for mediating the response to growth factors, paradoxically this gene does not require protein synthesis for the response (20). A number of other genes possess AP-1-like sites although the role of this factor in mediating the response of these genes to serum or growth factors has not been shown. Many of these genes are however responsive to TPA which exerts its effects via protein kinase C (see section I.2) and the AP-1 binding sites (21). As protein kinase C is an important component of the signal transduction pathways that control the cell's response to growth factor stimulus (see I.2) it suggests that the AP-1 sites are important in mediating the growth factor responses of these genes. It is also interesting to note that in some cases, for example proliferin, the AP-1 site is required for a serum response, but other sites are also required (22).

I.1.3 TRANSACTING FACTORS

As mentioned above the various regulatory sequences described are the binding sites for specific transacting DNA binding proteins (2).

A number of DNA binding transacting factors have now been identified and isolated. These factors tend to fall into groups sharing similar structural characteristics of functional importance.

For example a number of transcription factors have been shown to possess a zinc-finger binding domain (23). Examples of these zinc finger proteins include the TATA box binding factor TFIIIA (23,24) and the G-C rich box binding factor SP1 (25). The zinc finger domain of these proteins has a repeating motif consisting of two closely spaced cysteines followed by two histidines. A zinc ion is tetrahedrally co-ordinated by the cysteines and histidines (23,24). This structure gives the protein its DNA sequence specificity.

Another commonly occurring motif is the leucine zipper (26). This structure is found in a number of transacting factors , such as AP-1/c-jun, c-fos, the cAMP response element binding (CREB) protein and the yeast transacting factor GCN4. The amino acid sequence forms an α -helix with the periodic repetition of leucine at every seventh position over a short portion of the molecule. The leucine side chains from one molecule are able to interdigitate with those from another molecule to allow dimerisation, both homo and heterodimers can be formed (26,27,28). These proteins also possess a region of basic amino acids immediately before the leucine zipper region which is thought to be important in the binding of the dimers to the DNA (26).

Often the DNA binding domain is separate from the activating function of the protein. For example the activating domain of Sp1 consists of glutamine rich domains, these are essential for the transcriptional activation function of this factor (29). The C-terminal end of c-fos is required for its activating function (30).

There is a number of theories on how the enhancer bound and promoter bound proteins (or protein complexes) interact to allow efficient transcription (31). The looping model is the most favoured of these models (31,32). This model suggests that the proteins are able to interact by the looping out of the intervening DNA. This model would require the proteins to lie on the same side of the DNA, thus making the position of the sequence motif of importance. The position of the enhancer with respect to the promoter is generally thought to be flexible in terms of the distance, however the insertion or deletion of a few nucleotides between the promoter and enhancer can drastically effect enhancer activity (1). This suggests that the exact spacing can be quite important. Further support for the looping model comes from the observation that an enhancer can act in trans if it is linked to the promoter via a protein bridge (33). This would not be possible if the mechanisms discussed below were used. Finally it has been shown that promoter usage is virtually unaffected by enhancer position. This supports the looping model, but is not compatible with the scanning/sliding model discussed below (34).

A number of other models have been proposed (31). The sliding model suggests that the protein recognises a binding site which it binds to and then moves along the DNA until it reaches another site where transcription is initiated. The oozing model suggests that the binding of a protein would facilitate the binding of another adjacent protein and so on until the transcription initiation site is reached, this process seems unlikely to be useful over long distances. The twisting model has been largely discounted. Changes in the conformation of the DNA have been proposed as a way in which regulation of expression could be mediated by transacting factors. In this model the binding of the factor would change

the DNA in some way resulting in activation. However as the DNA binding and activation domains of these molecules have been shown to be dissociable this model looks unlikely to be important.

Many of the rapid changes in gene expression seen in response to extracellular stimuli do not require protein synthesis implying that the transacting factors responsible are already present in the cell. The stimulus must activate these inactive proteins by some sort of post-translational modification. There is evidence that a number of modifications maybe important in controlling the activity of the factors (35). For example the c-fos SRF requires phosphorylation in order to bind to the SRE (36). Also CREB (cAMP response element binding protein) is phosphorylated by protein kinase A and protein kinase C (37). It has been suggested that phosphorylation is important because it increases the negative charge of the molecule, a negative charge is thought to be important in the activating regions (32,35). Another type of modification that is important is the glycosylation of Sp1 with O-linked N-acetylglucosamine monosaccharides (38). This modification may affect the transport of the Sp1 or its localisation within the nucleus, alternatively it may stabilize the factor within the cell. A number of other known trans-acting factors also undergo glycosylation, but no function has been shown for this type of modification in these factors.

I.1.4 POST-TRANSCRIPTIONAL MECHANISMS FOR CONTROLLING GENE

EXPRESSION

Although this project is primarily interested in transcriptional regulation of expression there are many important post-transcriptional mechanisms that can affect the levels of a particular mRNA or protein. For example the transient response of the c-fos gene to serum induction

is due in part to the instability of the mRNA (12). Also the deregulation of expression of the ornithine decarboxylase gene in c-Ha-ras transformed cells is as a result of changes in the stability and translation of the ODC mRNA (39). Also, as mentioned in the previous section, post-translational events such as phosphorylation can have important effects on the activity of a gene product.

I.2. POLYPEPTIDE GROWTH FACTORS AND SIGNAL TRANSDUCTION

There is a great deal of interest in the mechanisms that control cell growth and proliferation. It is hoped that an understanding of the mechanisms involved in the control of normal cellular growth will shed light on the aberrant processes that occur in cancer cells. Cancer is essentially a disease of uncontrolled cell growth.

One way in which the effects of growth factors have been studied has been to look at events at the cell surface and across the plasma membrane. An alternative approach has been to observe changes in gene expression that occur in the presence of growth factors.

Of relevance to these studies is the work on proto-oncogenes. Oncogenes were first isolated as parts of the genomes of acutely transforming retroviruses (40,41). However it was soon shown that these viral genes were copies of normal cellular genes introduced into the retrovirus by a process known as transduction (41). There are two main theories on the way that these normal cellular genes become tumorigenic. One is that they are over expressed because they are under viral control instead of their normal control regions (41). In some cases it has been shown that over expression of the normal cellular gene can cause transformation of cell lines. Also the viral oncogenes can differ from their cellular homologues in a number of ways, such as point mutations

or deletions. This appears to affect control of the biochemical activities of these genes thus resulting in their tumorigenicity (40). The analysis of naturally occurring tumours has indicated that changes in the cellular oncogenes are sometimes observed (42).

Over the last few years a large number of studies have shown that the normal cellular proto-oncogenes are components of the signal transduction pathways that carry messages from the cell surface to the nucleus (43). Oncogenes and their activated counterparts are therefore central to our understanding of normal control of cell growth and proliferation and of the processes that go awry in cancer cells.

Many of the studies that have been done on the effects of growth factors have utilized the ability of cultured cells, particularly 3T3 fibroblasts, to enter a quiescent state known as G0 when grown in serum free medium (44). If the quiescent cells are then induced with serum or purified growth factors they re-enter the G1 phase of the cell cycle where the cell prepares for DNA synthesis (44). By observing events at the cell surface, in the cytoplasm and in the nucleus as the cells go from G0 to G1 a great deal has been learnt about the signal transduction pathways involved in mediating the effects of growth factors. However there are still many stages of these pathways that are not well understood.

I.2.1 GROWTH FACTOR INDUCED EVENTS AT THE CELL SURFACE AND IN THE CYTOPLASM

Multiple pathways are responsible for mediating the effects of growth factor stimulation (45,46).

The growth factor or factors induce a response by binding to highly specific growth factor receptor molecules that traverse the plasma

membrane of the cell (45,47). EGF for example binds to a receptor which is a single peptide of 1186 amino acids. Approximately half of the molecule is outside the cell, 23 amino acids form the transmembrane region and the rest are in the cytoplasm (48). The binding of the growth factor to its receptor results in dimerisation of the receptor molecules and autophosphorylation by the molecule's intrinsic tyrosine kinase activity (47). This leads to a cascade of events in the cell membrane and the cytoplasm. As mentioned previously the protein products of a number of oncogenes have been shown to play an important role in signal transduction pathways. The v-erbB oncogene is a truncated form of the avian receptor for EGF (49). The viral gene lacks the EGF binding domain and it has been suggested that the receptor becomes constitutively activated by this mutation.

The molecules involved in the cascade of events mediating the signal through the plasma membrane and the cytoplasm are known as second messengers (48,50). Molecules that act as second messengers include Na^+ , Ca^{2+} , H^+ , diacylglycerol (DAG), and inositol^{tris}phosphate^(IP3). One event that appears to be particularly important in signal transduction is the increase in intracellular Ca^{2+} levels (45). This is achieved in a number of ways. For example EGF appears to stimulate an uptake of Ca^{2+} from extracellular sources (47). PDGF on the other hand stimulates the release of Ca^{2+} from intracellular stores (45). This occurs due to the activation of phospholipase C. This enzyme hydrolyses phosphatidylinositol 4,5-bisphosphate (PIP_2) which leads to the release of DAG and IP_3 . IP_3 stimulates the release of Ca^{2+} from the intracellular stores. The precise role of Ca^{2+} is not known but it does appear to be important in initiating DNA synthesis.

DAG causes the activation of protein kinase C, this is a Ca^{2+} -sensitive phospholipid dependent kinase which is responsible for a cascade of phosphorylation events that play an important role in signal transduction (45,51). Protein kinase C also effects the Na^+/H^+ exchanger in the plasma membrane (47). This results in a shift in pH which also appears to have a role as a second messenger. The activity of the Na^+/H^+ exchanger is also stimulated by growth factors that do not activate protein kinase C suggesting that more than one mechanism is responsible for the shift in pH (50).

The mechanisms activating phospholipase C are not fully understood, however it has been suggested that the ras proto-oncogenes may be important (52). This is interesting as it once again links proto-oncogenes with signal transduction pathways.

1.2.1.1 RAS PROTO-ONCOGENES

The ras proto-oncogenes encode a class of regulatory GTP-binding proteins involved in the control of cell proliferation. The ras genes are members of a large family of related genes found in all eukaryotic genomes (165).

The Ha-v-ras and Ki-v-ras oncogenes were discovered in the genomes of the Harvey and Kirsten rat sarcoma viruses, respectively. The cellular homologues of these genes are the Ki-c-ras and Ha-c-ras genes. These genes encode 21kDa proteins known as $\text{p21}^{\text{Ki-ras}}$ and $\text{p21}^{\text{Ha-ras}}$. A closely related gene, N-ras, was isolated from a neuroblastoma (53).

The p21 ras proteins are similar over much of their length, with major differences between the three mammalian proteins occurring in the carboxy terminal domain. Differences here presumably account for any possible differences in the function of the three proteins. There are a number of structural features of note in the ras proteins. The regions

important in the binding of GTP occur at residues 12, 13, 50, 61, 63, 116, 117, 119 and 146. Residues 32-40 form the effector loop, the conformation of this region is quite different in GTP- and GDP-bound forms. It is presumably the site where p21^{ras} interacts with its cellular target. The region between residues 61 and 75 also undergoes conformational change upon binding by GTP. This region is thought likely to interact with a protein important in the regulation of ras function. (165). The carboxy terminal region of the p21^{ras} proteins undergoes post-translational modification at several residues. In particular the CAAX box, found at the carboxy terminus of almost all ras-like proteins, is of interest. (C is cysteine, A is any aliphatic amino acid and X is any uncharged amino acid). Processing of this region is essential for localization of the p21^{ras} proteins to the plasma membrane, this is essential for the protein to function (165).

Proteins homologous to p21^{ras} have been found in a number of other eukaryotes. For example Saccharomyces cerevisiae, Schizosacchomyces pombe, Dictyostelium discoides and Drosophila melanogaster all possess at least one ras-like protein (53,165).

There are also a number of ras-related proteins (165,166). Some of these have 50-60% amino acid homology to the mammalian ras protein. This group includes Ral A and B, Rap 1, R-Ras and TC-21 in mammalian cells. Other ras-related proteins show lower levels of homology. Both the Rho family and the SEC4, YPT1 and Rab families show about 30% homology to ras.

Normal p21^{ras} is a GTP-binding protein with intrinsic GTPase activity which catalyzes the hydrolysis of GTP to GDP. The GDP bound form of p21^{ras} is 'inactive' and the GTP bound form is 'active' (53). Figure A shows the GTPase cycle. A Guanine Nucleotide Release Protein (GNRP) (discussed in greater detail below) interacts with the GDP-p21^{ras} to

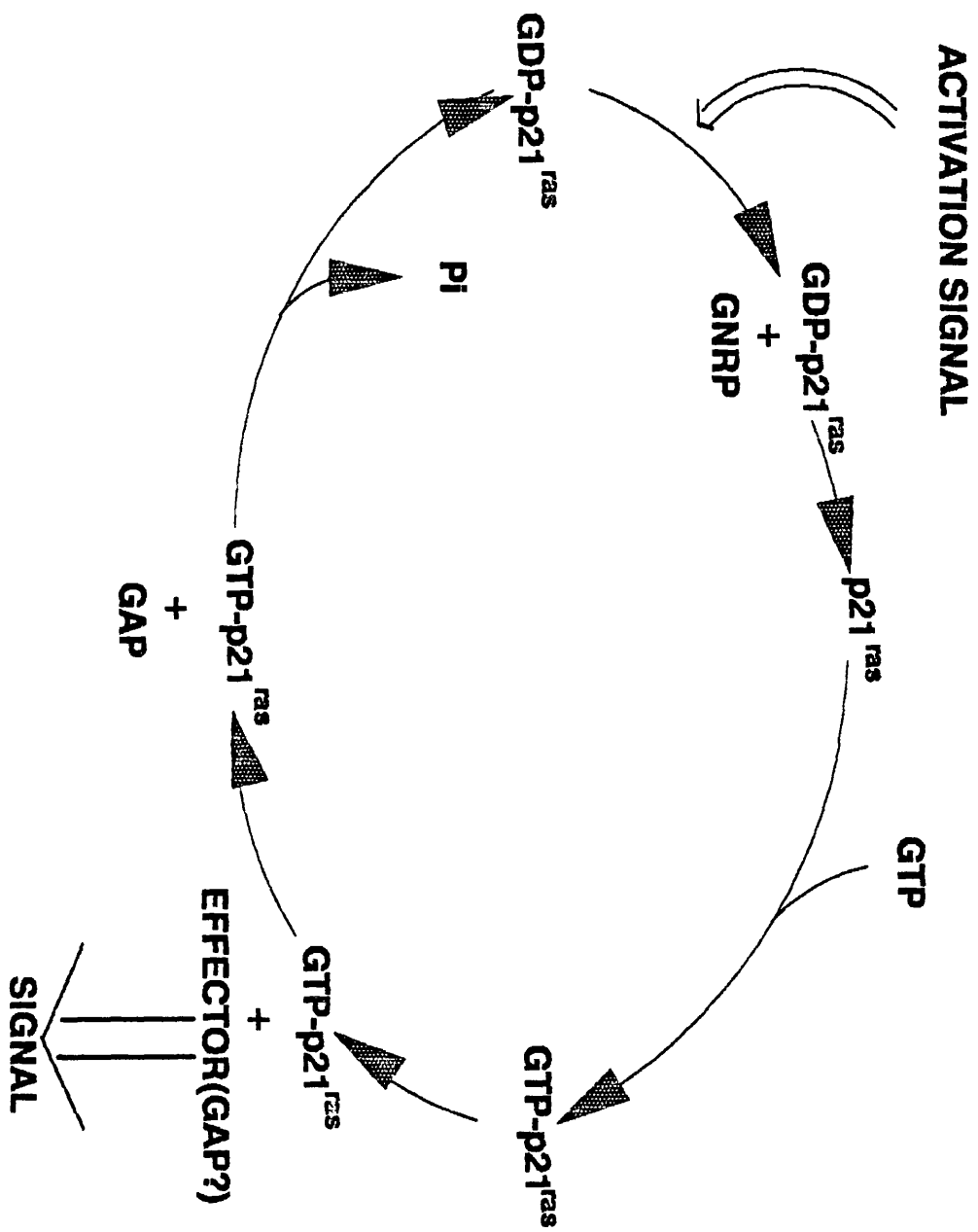


Figure A. The GTPase Cycle.

enhance the removal of GDP resulting in a transient 'empty' state (167). GTP is then able to bind to the p21^{ras} and it becomes active. Hydrolysis of the bound GTP results in the GTP-p21^{ras} reverting to the GDP-bound inactive state (167). This reaction is catalyzed by the GTPase activating protein (GAP)(167)(discussed in detail below).

The ras GAP appears to be the negative regulator of p21^{ras} function as it catalyzes the hydrolysis of GTP thus returning the p21^{ras} to its inactive GDP-bound form (54). It has been suggested that GAP may also be the effector molecule as it binds to the effector region of the p21^{ras}. However there is no direct evidence for this and it is possible that two overlapping proteins act at this region of the protein. Recently it has been noted that GAP is phosphorylated and it is thought that this may provide a link between p21^{ras} and growth factor receptors. However it is not known what effect that this phosphorylation has on GAP function.

The GNRPs appear to be the positive regulators of p21^{ras} function. Wolfman and Macara (168) have isolated a cytosolic protein from rat brains called Ras-GRF (Guanine nucleotide releasing factor) which enhances the removal of guanine nucleotides from p21^{ras}. Ras-GRF allows the rapid equilibration of p21 ras with the free guanine nucleotide pools in the cell thus controlling the fraction of p21^{ras} in the active state.

p21^{ras} is involved in the control of cellular proliferation. However its exact role is not yet understood. Ras function is essential during the first 8 hours of the G1 stage of the cell cycle (167). Also antibodies to ras injected into cells block the action of a wide variety of growth factors. Ras appears to have a general role in growth control rather than being linked to a particular receptor and signal transduction pathway (167). High protein kinase C activity in ras-transformed cells suggests that ras stimulates the pathway leading to protein kinase C activation, however there is no direct evidence for the stimulation of

phosphoinositide turnover (56). Therefore the mechanism for activating protein kinase C is unknown. Also protein kinase C is not required for morphological transformation by p21^{ras} (57). These results suggest that p21^{ras} activates multiple pathways and not just the pathway leading to protein kinase C.

The upstream and downstream regulation of the p21^{ras} is not yet well understood. For example any role that GAP may have as an effector is not yet clear. Also the role of the GNRPs is not yet fully characterised, however they are clearly important and may be stimulated by growth factor binding (167).

The cellular ras genes become oncogenic due to point mutations. mutation at residues 12, 13, 59, 61, or 63 reduce the GTPase activity of the ras proteins. Mutations at residues 116, 117, 119 or 146 decrease the affinity of p21^{ras} for nucleotide and thus increase its rate of bound GDP for GTP. This causes an accumulation of active GTP-p21^{ras} (165). Also GAP has no effect on oncogenic p21^{ras} with mutations at 12, 59 or 61 and therefore any residual intrinsic GTPase activity is not catalyzed by the GAP. These mutations all lead to the ras being in a constitutively active state which leads to a continuous flow of signal transduction leading to malignant transformation.

There are other signal transduction pathways that operate within the cell. For example some external stimuli can activate adenylate cyclase which converts ATP to cAMP (58). cAMP binds to protein kinase A which phosphorylates transacting proteins such as CREB which activate transcription (37). Interestingly in yeast ras has been shown to operate through adenylate cyclase, however this does not appear to be the situation in higher eukaryotes (54).

x H

I.2.2 GROWTH FACTOR MEDIATED EVENTS IN THE NUCLEUS

The signals generated by growth factor binding reach the nucleus very quickly. Within minutes a number of genes show increased levels of transcription. These genes have been termed immediate early or early response genes (59,60). These genes share a number of characteristics. For example they are all rapidly and transiently induced by growth factors. This requires no new protein synthesis and in fact the expression is superinduced by the protein synthesis inhibitor cycloheximide. A number of these genes have been studied in depth.

Several of them represent known or possible DNA binding proteins, for example c-fos (61,62), c-jun (63), and the zinc finger proteins EGR1 (64) and EGR2 (16).

Of great interest are those genes that are also known nuclear proto-oncogenes, such as c-fos and c-jun. The structure and function of the c-fos and c-jun oncogenes is well known and they are important in mediating the effects of growth factor stimulation, thereby providing another link between the signal transduction pathways and proto-oncogenes. c-fos is the cellular homologue of the oncogene carried by the FBJ and FBR murine osteosarcoma viruses (65). The gene encodes a nuclear phosphoprotein of 55kD (66). c-jun is the cellular homologue of the oncogene found in ASV17 (67). It appears to be identical to the transacting factor AP-1 (63). c-jun is a member of a family of factors which all bind to identical or similar DNA sequences (68,69). This oncogene has recently been shown to be part of the complex of proteins that binds with c-fos (70,71,72). In fact the c-jun region of the complex allows the association of fos with DNA, fos does not specifically bind to DNA but acts via binding with c-jun (70,71). Recently it has been shown that binding of c-jun to the AP-1 site is important in mediating the effects of serum, TPA and ras for a number of cellular and viral genes, for example polyoma (73). These observations provide a direct link between signal transduction and nuclear proto-oncogenes. S e r u m responsive genes have been isolated that respond much more slowly to serum stimulation. These genes frequently show a more sustained response with many hours needed for the levels of transcripts to return to normal. Like the immediate early genes these genes are controlled by transcriptional mechanisms, however a number of these genes have also been shown to have important post-transcriptional regulatory mechanisms,

such as proliferin (74) and TIMP (75). Unlike the immediate early genes these genes often require protein synthesis in order to respond to growth factor stimulation. This has led to the suggestion that these genes require transcription factors encoded by the immediate early genes in order to respond to the serum stimulation (see section 1).

I.3 RETROELEMENTS

Eukaryotic genomes contain a large amount of DNA encoding transposable genetic elements. These are elements that are able to move around the genome. Transposons can be grouped into those that transpose directly from DNA to DNA and those that transpose via reverse transcription of an RNA intermediate. The elements that utilize reverse transcriptase are termed retroelements or retroposons and can be further sub divided into two main groups, the non viral retroposons and retrotransposons which are characterised by their retrovirus provirus-like structures.

I.3.1 NONVIRAL RETROPOSONS

The non viral retroposons include the SINE and LINE gene families. SINES are short interspersed sequences of <500bp that are repeated 10^5 times in the genome (76). SINES are transcribed by RNA polymerase III and resemble other RNA polIII genes such as tRNA genes. Interestingly the rat ID class of SINES can repress or activate transcription from a nearby RNA polIII promoter in transient transfection experiments and have been shown to confer growth factor response and transformation response on a number of genes (77).

LINES are long interspersed sequences that are usually >5kb in length (78). These sequences possess open reading frames and encode a protein with similarities to reverse transcriptase (79).

I.3.1 RETROVIRUSES

Retroviral particles contain a dimeric RNA genome consisting of two apparently identical RNA molecules of 8-10Kb. Each genomic RNA molecule consists of a unique region flanked by a short directly repeated region (R). The R domain is important in transferring the nascent DNA chain and DNA polymerase from the 5' end to the 3' end of the genome during reverse transcription of the molecule. At the 5' end of the unique region is a domain known as U5, this domain separates the R region from the negative strand primer binding site (PBS). The minus strand PBS is complementary to the 3' terminal nucleotides of the tRNA^{po1} primer. The positive strand PBS occurs at the 3' end of the unique region and is separated from the R region by the U3 domain. The plus strand PBS has a purine rich tract that is highly conserved between viruses. The U3 domain contains sequence elements similar to those of eukaryotic genes, such as promoter elements and enhancer elements. During the replication cycle of the retrovirus the genomic RNA is reverse transcribed into a linear double stranded DNA molecule. This molecule is then able to integrate into the host cell DNA. The integrated virus is known as the provirus. When the linear DNA molecule is formed a long terminal repeat (LTR) is formed containing the U3 R U5 domains at each end of the molecule. The linear molecule is integrated into the host cell genome where it utilizes the host cell mechanisms to transcribe both full length genomic RNA and mRNAs for the retroviral genes. Oncoviruses possess three main retroviral genes. Gag encodes the internal structure proteins, pol encodes the reverse transcriptase and env encodes

the envelope proteins (80,81). Other retroviruses, such as HIV have much more complex genomes including genes that encode transacting factors.

The U3 domains of retroviral LTRs possess transcriptional regulatory determinants. The MoMSV LTRs have been particularly well studied and possess a TATA box, a CCAAT box and G-C rich box in the promoter region (82). They also possess imperfect 72 and 73bp direct repeats which behave in a similar way to the SV40 enhancers (83,84,85). Like the SV40 enhancer the retroviral enhancers are modular in structure. The enhancer consists of sequences responsible for mediating basal levels of expression as well as mediating the response to external stimuli such as glucocorticoids (86). Also the enhancer may play a role in negative regulation of expression in non permissive cell types such as undifferentiated embryonal carcinoma (EC) cells (87). The enhancers are also important in imparting a certain degree of host cell specificity. The MoMSV enhancers can functionally replace the SV40 enhancers (83), however the expression of these constructs is higher in mouse than in monkey cells (84).

The virus structure described above is known as replication competent as it possesses the genes necessary for the entire retroviral replication cycle. These viruses are also known as slowly oncogenic as infection with them can cause tumours of susceptible host cells, albeit after a quite considerable time. Strongly oncogenic retroviruses are often replication defective because portions of the unique region containing essential genes for replication have been replaced with modified versions of normal cellular genes known as oncogenes (see section I.2). These viral oncogenes are responsible for conferring tumorigenicity. In order to replicate these viruses require the presence of 'helper' replication competent viruses.

In the context of this project the slowly oncogenic viruses are of greater interest. These viruses can confer tumorigenicity by exerting transcriptional regulation on normal cellular genes by an insertion next to the gene (88). These insertion can work in two ways. In some cases, such as ALV-induced bursal lymphomas the promoter insertion model applies. The LTR sequences serve as a transcriptional promoter of the gene downstream of the provirus, for example c-myc in this case (89). However in many cases the provirus can effect expression of a cellular oncogene from either upstream or downstream, that is it utilizes the enhancers present in the LTR (88). This type of insertion is important because it relieves the provirus of any positional requirement. It can insert in either orientation with respect to the cellular gene and even at quite some distance.

I.3.2 RETROTRANSPOSONS

These elements possess a retrovirus provirus-like structure consisting of two LTRs surrounding a unique region. Well studied examples of retrotransposons include the *Drosophila* copia-like elements (90), the Ty elements of *Saccharomyces cerevisiae* (91,92) and mouse intracisternal A particles (IAPs) (93).

The LTRs of these elements have a retrovirus-like structure consisting of U3, R and U5 domains. Also, like retroviral LTRs, they possess inverted repeats at their ends and transcriptional regulatory sequences (94,95,96,97). In terms of actual DNA sequence homology there is very little similarity between the retroviral and retrotransposon LTRs.

The unique regions of the retrotransposons also show similarities to the retroviral unique regions. At the 5' end of the region there is

a tRNA PBS used to prime minus strand DNA synthesis and at the 3' end a purine tract which is the PBS for plus strand DNA synthesis. There are also open reading frames with homology to the retroviral gag and pol genes, frequently however there is no homology to the retroviral env gene (98,99,100).

These retrotransposons are able to form virus-like particles, but these particles do not appear to be infectious. The IAP particles for instance bud into intracellular membranes rather than through the plasma membrane (93). Also it has been shown that Ty elements transpose via an RNA intermediate (101).

Despite sharing many features with retroviruses the retrotransposons do not merely represent degenerate endogenous retroviruses, otherwise a higher level of sequence homology would be expected. Instead it has been suggested that these elements represent the basic units of retrotransposable elements with retroviruses themselves representing more highly evolved forms that are able to move from genome to genome (102). This fits in well with Temin's theory on retroviral evolution which suggests that retroviruses were formed by transposition of small mobile genetic elements around cellular genes, the retrovirus gradually picking up more information with each successive gene 'captured' (103). In this sort of model retrotransposons such as Ty and IAPs would represent an intermediate in retroviral evolution.

Retrotransposons are of interest for a number of reasons. In the last few years there have been a number of examples of the integration of complete or incomplete retrotransposons affecting the expression of cellular genes. For example yeast Ty elements have been shown to effect the transcription of cellular genes by inserting into the non-coding region of the genes (92). In activating transpositions the Ty element is

usually inserted in the opposite orientation to the gene in question. These insertions are known as ROAM mutations (regulated overproducing alleles under mating signals) as the Ty element expression and hence the cellular gene's expression is effected by the mating type of the yeast (104). Inactivating transpositions can also occur, these are caused by the separation of the target gene promoter elements (92).

There is also a number of examples of IAP insertions affecting the expression of cellular genes. For example an IAP has been shown to provide regulatory elements in cis for the c-mos oncogene in a myeloma cell line (105). In another example a solo IAP LTR affected the expression of CSF genes in transformed CSF dependent cells (106).

Clearly this sort of integration in a germ cell would result in the stable transmission of the 'new' regulatory region and gene. Obviously any such insertion could be undesirable resulting in the death of the organism. However there is at least one example of a retrovirus-like element stably affecting the expression of a cellular gene. The murine slp gene is androgen responsive and it has recently been shown that expression of this gene is controlled by the LTR of a degenerate retrovirus-like element (107). This insertion was tolerated because the 'parental' C4 gene had already been duplicated and therefore there was already a normally regulated copy of the gene present.

The widespread occurrence of retrotransposons suggests that they may be of evolutionary importance (108,109). It has been suggested that such elements could alter the genome in a number of ways. For, example, Ty elements have been found at the breakpoints of deletions, duplications, inversions and translocations. Presumably recombination between elements produced these changes (108).

I.4. MOUSE VL30 ELEMENTS

VL30 elements are a moderately repeated class of mouse retrotransposon. There are 150 copies of this element dispersed throughout the genome (110). Like the retrotransposable elements discussed in the previous section, VL30 elements possess a retrovirus provirus-like structure (111,112). Initially VL30 elements were isolated as a species of RNA subunit present in the virus particles of type C retroviruses propagated on some mouse cell lines, such as NIH 3T3 cells (113,114,115). However not all KiMLV particles possessed the 30S RNA (113). The 30S RNA therefore was clearly not an essential part of the retroviral genome as it was not always present. Therefore it was suggested that the 30S RNA was produced by a defective endogenous retrovirus that required the functions of helper independent MLV in order to be packaged (113,115,116).

I.4.1 CHARACTERISTIC FEATURES OF VL30 ELEMENTS

The VL30 elements all possess a provirus-like structure. This consists of a unique region surrounded by two LTRs. The entire element is about 5Kb with the LTRs being between 0.4 and 0.6Kb long (111,112). The LTRs are heterogenous and sequence analysis of a number of the LTRs suggests that there has been a considerable amount of recombination between members of the VL30 family (117,118,119). Conserved within the LTRs are a number of transcriptional regulatory sequence elements suggesting that the LTRs, like retroviral LTRs, are responsible for controlling expression of the VL30 elements. The sequence data also shows that the LTRs possess inverted repeat sequences at their ends, these are similar in sequence to the retroviral IRs (118,120). At the inner boundary of the 5' LTR the VL30 elements possess sequences homologous to

tRNA primer binding sites (PBS), however unlike retroviruses a number of different PBS sequences are found, this includes tRNA^{Pro}, tRNA^{Gly}, and tRNA^{Gln} (118,121). Also at the inner boundary of the 3' LTR there is a purine rich tract, this feature in retroviruses is important in plus strand DNA synthesis (120).

Hybridization studies indicated that in the unique region there is only very limited homology to retroviral sequences (122). This has been supported by the sequence data for the unique region of a VL30 clone (123). The sequence shows that there are some short regions of homology to retroviral gag and pol genes. However the VL30 elements have no open reading frames and no homology to retroviral env genes. These observations would tend to classify VL30 elements as retrotransposons rather than as defective retroviruses. However many retrotransposons, such as Ty and copia, have been shown to possess functional gag and pol genes.

I.4.2 EXPRESSION OF VL30 ELEMENTS

Although there are about 150 VL30 elements in the genome only a small subset are transcriptionally active. Methylation studies have shown that less than 5% of the elements are transcribed (124,110). VL30 expression in a number of cell lines has been studied. From these studies it is clear that the level of VL30 expression is correlated to the growth state of the cell (113,125). For example in cell lines with a more transformed phenotype higher levels of VL30 transcripts are seen. By contrast cells that show contact inhibition of growth tend to exhibit lower levels of VL30 RNA. VL30 expression is also induced by a variety of agents. For example cells treated with BuDR show elevated levels of VL30 expression (113) and chemically transformed AKR and C3H/10T1/2 cells

show elevated levels of expression (126). Quiescent cells stimulated with purified growth factors or TPA show a large induction of VL30 expression (117,127,128). Also in T cells VL30 elements have been shown to be responsive to glucocorticoids and cAMP (129).

I.4.3 PROMOTER STRENGTHS OF VL30 LTRs

As mentioned above the VL30 LTR sequences possess putative regulatory sequences such as a promoter and enhancer-like sequences. Rotman *et al* (130) have tested a number of LTRs obtained from genomic clones for promoter and enhancer activity using transient expression assays. The transcriptional activity of the clones is not known and this is reflected in the variable promoter strengths seen. However at least one of the clones, VL3, was shown to have a high promoter and enhancer activity. A number of other clones showed low promoter activities compared to the control SV40 early promoter.

I.4.4 THE NVL VL30 ELEMENTS

In order to clone the small subset of transcriptionally active retrovirus transmissible elements KiMLV particles propagated on NIH 3T3 cells were used in an endogenous reverse transcriptase reaction (112). Particles propagated on this cell line contained VL30 RNA which was reverse transcribed to produce full length copies of the retrovirus packaged elements. Twelve independently isolated clones were obtained and assigned to one of four classes on the basis of size and restriction enzyme analysis. These classes were designated NVL-1, NVL-2, NVL-3 and NVL-4 (NIH VL30). Eight of these clones were designated NVL-3, two were NVL-2 and one each NVL-1 and NVL-4. The restriction enzyme analysis indicated that as with all other VL30 elements, the NVL elements

consisted of a unique region surrounded by LTRs. The LTRs show the sequence only differences between the four types of NVL VL30 element. In fact NVL-1, NVL-2 and NVL-4 appear to be similar, with NVL-4 representing a truncated form of NVL-1. This was presumably caused as a result of incomplete reverse transcription of the NVL-4 element. The NVL-3 LTR appears to possess more substantial differences and is somewhat shorter than the other NVL elements. The unique regions of all four types of element appear to be identical according to the restriction enzyme analysis. Nucleotide sequence analysis of the NVL-3 LTR and determination of the functional domains shows that the NVL-3 LTRs are structurally similar to retroviral LTRs, however there is little sequence homology (118). Like a number of other VL30 LTRs the NVL-3 LTRs possess a number of sequence motifs similar to those seen in viral and cellular gene promoters and enhancers.

An analysis of mouse genomic DNA has shown that there are 2-4 copies of the NVL VL30 elements per haploid genome. This correlates well with the methylation data which shows 3-5 copies of NVL type elements are undermethylated indicating that the retrovirus transmissible NVL VL30 elements also represent the transcriptionally active elements (124).

Analysis of genomic DNA from heterologous NRK (normal rat kidney) cells infected with VL30 elements by retrovirus mediated gene transfer showed random integration of the VL30 elements (124).

I.5 AIMS

VL30 elements, particularly the transcriptionally active retrotransmissible NVL elements, are of interest for a number of reasons. As growth factor responsive and transformation responsive genes they are of interest because they may help to provide insight into mechanisms that

control responses to changes in the cells growth and in response to malignant transformation.

VL30 elements are also of interest because they are able to transpose if packaged as a pseudovirion complex by murine leukaemia viruses. There are growing number of examples of retrotransposons, such as Ty and IAPs, exerting transcriptional control over adjacent cellular genes via cis-acting enhancer sequences. It is therefore of interest to see whether VL30 elements possess cis regulatory functions that could potentially effect mouse cellular genes.

The purpose of this project therefore is to relate the structural features of the transcriptionally active, retrovirus transmissible NVL VL30 elements to their expression, particularly in response to serum and cellular transformation.

MATERIALS

MA.1 COMMERCIAL PRODUCTS

AMERSHAM

[α -³²P]dCTP-specific activity 370MBqml⁻¹

Adenosine 5'-[γ -³²P]triphosphate 370MBqml⁻¹

Klenow fragment of DNA polymerase I 4units μ l⁻¹

Sequencing kit

[α -³⁵S]dATP₃S (>1000 Cimmol⁻¹)

D-threo-[dichloroacetyl-1-C¹⁴]chloramphenicol 2GBqmmol⁻¹

ALDRICH

Dimethyl Sulphoxide (DMSO)

ATCC

NIH 3T3 cells

pSVOCAT

pSV2CAT

pSV2NEO

BDH

All chemicals were BDH Analar grade except where indicated.

Glass wool

Glass Beads 40 mesh

Polyethylene glycol (PEG)8000

REPELCOTE (2% dimethyldichlorosilane in 1,1,1-trichloroethane)

NNN'N'-tetramethylethylenediamine (TEMED)

Yeast RNA -stored as a 200mgml⁻¹ stock

Bromophenol blue

Xylene cyanol

Hexadecyltrimethylammonium bromide (CTAB)

Polyoxyethylene(20)sorbitan mono-oleate (TWEEN 80)

Nonidet P40 (NP40)

BECKMAN

MICROFUGE II

BOEHRINGER MANNHEIM

All restriction endonucleases were purchased from Boehringer Mannheim

Calf Intestinal Alkaline Phosphatase (CIAP) 24 units μl^{-1}

T4 DNA Ligase 1 unit μl^{-1}

Isopropyl- β -D-thio-galactopyranoside (IPTG)

Stored as a 100mM stock.

5-bromo-4-chloro-3-indolyl- β -galactoside (X-gal)

Stored as a 2% stock in dimethylformamide.

Bovine Serum Albumin (BSA) (DNAase free)

Polynucleotide Kinase (PNK) 9 units μl^{-1}

RNAase free DNAase I 23 units μl^{-1}

Vanadyl ribonucleoside complex (VRC) 20mM

Chloramphenicol acetyl transferase (CAT)

Stored as 1 units μl^{-1} in 10M Tris pH7.8

BIOSEARCH INC.

Cyclone DNA synthesizer

CANBERRA PACKARD

Emulsifier Scintillator

Filter-Count Scintillation Fluid

DIFCO

Bactotryptone

Yeast extract

Agar

FLOW

HeLa cells

GIBCO

Dulbecco's Modified Essential Medium (DMEM)

Foetal Calf Serum (FCS)

Newborn Calf Serum (NBS)

Gentamicin 20mgml⁻¹

Versene 1:5000 in isotonicly buffered saline

2.5% Trypsin

Geneticin (G418)

Stored as a 100mgml⁻¹ stock

KODAK

Film - X-OMAT AR or X-OMAT S

X-ray developer LX24

X-ray fixer FX40

KONTRON

Centrifuge tubes - 14ml polyallomer

Centrikon T-2070 ultracentrifuge

MERCK

Silica Gel Thin layer chromatography (TLC) plates

MSE

Microcentaur

Centaur 2

High Speed 18

NEN

Gene Screen Plus nylon filters

NUNC

Plastic tissue culture flasks and plates

PHARMACIA

Oligodeoxynucleotide hexamers

Sephadex G50

BindSilane (methacryloxypropyl-trimethoxy silane)

Placental RNAase inhibitor 27units μ l⁻¹

M-MLV reverse transcriptase 15units μ l⁻¹

SCHLEICHER AND SCHUELL

Dot blot apparatus

Nitrocellulose filters

SIGMA

12-0-tetradeconyl-13-phorbol acetate (TPA)

Stored as a 10⁻⁵M stock in 100% ethanol

1-(5-isoquinolinylsulphonyl)-2-methylpiperazine (H-7)

Agarose

Low Melting Point Agarose (LMA)

Ethidium Bromide

Stock solution 10mgml^{-1}

tRNA

Stock solution 10mgml^{-1}

RNAaseA

Stock solution 5mgml^{-1}

Lysozyme

Chloramphenicol

Stock solution- 34mgml^{-1} in ethanol.

dNTPs

Stored as 20mM stocks, working solutions were stored as 5mM stocks.

Dithiothrietol (DTT)

Diethyl pyrocarbonate (DEPC)

N-laurolylsarcosine

β -mercaptoethanol

Phosphoenol pyruvate

Pyruvate kinase

Fructose,1-6 diphosphate

Thymidine diphosphate

Heparin

Proteinase K

stored as a 10mgml^{-1} stock

Acetyl Coenzyme A (AcCoA)

Antifoam A

WHATMAN

3MM chromatography paper

DE81 paper

MA.2 SOLUTIONS AND BUFFERS

Freeze Solution I - 20% FCS in DMEM

Freeze Solution II - 40% FCS, 16% DMSO in DMEM

5 x Ligase buffer - 330mM Tris pH7.6, 33mM MgCl₂, 5mM ATP, 50mM DTT.

10 x Minigel buffer - 900mM Tris, 900mM Boric acid, 25mM EDTA.

10 x E buffer - 400mM Tris pH7.7, 200mM NaOAc, 20mM EDTA

L-broth - 10g bactotryptone, 5g yeast extract, 5g NaCl in 100ml water.

L-agar - 6% agar in L-broth.

TCM buffer - 10mM Tris pH7.5, 10mM MgCl₂, 1mM CaCl₂

S.O.C. medium - 2% bactotryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl₂, 10mM MgSO₄, 20mM glucose.

TELT - 50mM Tris pH7.5, 63mM EDTA, 0.4% (v/v) Triton-X100, 2.5M LiCl.

Lysis buffer - 50mM glucose, 25mM Tris pH8, 10mM EDTA.

10 x Hogness freezing medium - 36mM K₂HPO₄, 13mM KH₂PO₄, 20mM Na₃citrate, 10mM MgSO₄, 44% glycerol.

50 x Denhardt's solution - 5g Ficoll 400, 5g polyvinylpyrrolidone, 5g BSA (fraction V) in 500ml of water.

Phenol/Chloroform -equal volumes of phenol and chloroform make up 95% of the solution, water-5%, 8-hydroxyquinoline-0.1%w/v.

Formamide prehybridization solution - 50% formamide pH7, 5 x Denhardt's solution, 3 x SSC, 50mM HEPES pH7, 0.5% SDS, 200 μ gml⁻¹ yeast RNA, 200 μ gml⁻¹ salmon sperm DNA.

Formamide hybridization solution - 50% formamide pH7, 1 x Denhardt's solution, 3 x SSC, 50mM HEPES pH7, 0.5% SDS, 10% Dextran sulphate, 200 μ gml⁻¹ yeast RNA, 200 μ gml⁻¹ salmon sperm DNA.

TES - 20mM Tris pH7.5, 10mM NaCl, 0.1mM EDTA

40% acrylamide solution - 38g acrylamide, 2g bis acrylamide in 100mls of deionised water.

10 x Sequencing buffer - 1M Tris pH8.3, 1M Boric acid, 20mM EDTA.

Urea loading buffer - 8M urea, 10mM bromophenol blue, 10mM xylene cyanol.

1 x DNA extraction buffer (DEB) - 20mM Tris pH7.5, 200mM NaCl, 10mM EDTA.

GITC lysis buffer - 50g guanidine isothiocyanate, 50mM Tris pH7.5, 10mM EDTA, 2% sarcosyl, 1ml β -mercaptoethanol.

1 x MOPS - 200mM MOPS, 50mM NaOAc, 10mM EDTA pH7

Northern blot denaturing solution - 80% formaldehyde, 12% formamide pH7,

1 x MOPS

Dot blot denaturing solution - 12% formaldehyde, 69% formamide pH7, 100mM HEPES.

RNA annealing solution - 1M NaCl, 1%SDS, 10% dextran sulphate, 10mM HEPES pH7, 200 μ gml⁻¹ yeast RNA, 200 μ gml⁻¹ salmon sperm DNA.

Nuclei lysis buffer - 10mM HEPES pH7.9, 10mM NaCl, 50mM KCl, 3mM MgCl₂, 0.15mM spermine, 0.15mM spermidine.

Nuclei storage buffer - 50mM HEPES pH7.9, 5mM MgCl₂, 0.1mM EDTA.

2 x transcription mix - 50mM HEPES pH 7.9, 20mM KCl, 20mM DTT, 30mM β -mercaptoethanol, 50mM MgOAc, 2mM MnCl₂, 0.5mM EDTA, 8nM phosphoenol pyruvate, 6 μ gml⁻¹ pyruvate kinase, 2mM fructose,1-6 diphosphate, 1% TWEEN 80, 2mM thymidine diphosphate, 1mM CTP, 1mM ATP, 1mMGTP, 1000unitsml⁻¹ placental RNAase inhibitor, 0.4mgml⁻¹ heparin, 30Cimmol⁻¹ α ³²P UTP.

RES buffer - 0.5M LiCl, 1M urea, 0.25% SDS, 20mM Na₃citrate, 2.5mM EDTA.

LiCl/ethanol - 3:2 5M LiCl:95% Ethanol (v/v)

Transfection Buffer A - 0.5M CaCl₂, 0.1M HEPES pH7.05.

Transfection buffer B - 0.28M NaCl, 0.05M HEPES, 0.75mM NaH₂PO₄, 0.75mM Na₂HPO₄ pH 7.05.

Annealing solution - 80% formamide pH7, 400mM NaCl, 50mM HEPES pH7, 1mM EDTA.

5 x PE buffer - 50mM Tris pH8.3, 40mM KCl 8mM MgCl₂, 10mM DTT.

10 x NT - 500mM Tris pH7.5, 500mM NaCl, 100mM MgCl₂, 70mM β-mercaptoethanol, 50μg ml⁻¹ BSA

1 x restriction enzyme buffers:

1 x E buffer (BamHI, XbaI) - 6mM Tris pH7.9, 150mM NaCl 6mM MgCl₂.

1 x C buffer (EcoRI) - 100mM Tris pH7.4, 50mM NaCl, 10mM MgCl₂.

1 x A buffer (HindIII, PstI) - 10mM Tris pH7.4, 50mM NaCl, 10mM MgCl₂.

1 x G buffer (SacI, MspI) - 6mM Tris pH 7.4, 20mM NaCl, 6 mM MgCl₂.

1 x M (BglII) - 10mM Tris pH7.4, 60mM NaCl, 10mM MgCl₂.

1 x H (SmaI) - 6mM Tris pH8, 20mM KCl, 6mM MgCl₂.

All buffers were supplemented with 100μg ml⁻¹ BSA and 7mM β-mercaptoethanol.

METHODS

M.1 CELL CULTURE

M.1.1 MAINTENANCE OF CELL LINES

All the cell lines were grown as monolayers in Dulbecco's Modified Essential Medium (DMEM) in 25cm², 75cm² or 175cm² plastic tissue culture flasks. Normal and spontaneously transformed NIH 3T3 cells were grown in DMEM supplemented with 10% new born calf serum (NBS). The v-Ki-ras transformed NIH 3T3 cell lines CC1 and CA12 (131) were grown in DMEM plus 5% NBS. Phenotypically normal revertants, R5.5.1 and R2.1, derived from the CC1 cell line (131) were grown in DMEM plus 10%NBS. The c-Ha-ras transformed NIH 3T3 cell line EC34 (provided by Dr. J. Cunningham with the kind permission of Dr. E. Chang) was grown in DMEM plus 10% FCS, as were Rat1 embryo fibroblasts, Human HeLa cells and normal rat kidney cells in which VL30 had been introduced by retrovirus mediated gene transfer (VLNRK) (124). All cultures included the antibiotic gentamycin to a final concentration of 20µgml⁻¹. Cells were grown at 37°C in 95% air/5%CO₂.

The A1 cell line is an NIH 3T3 cell line that has become transformed spontaneously due to a high number of passages.

M.1.2 PASSAGING OF CELL LINES

Cells were passaged when they reached approximately 75% confluency. Cell monolayers were removed from the flask by digestion with trypsin. Stock trypsin (2.5%) was activated by dilution in versene, all NIH 3T3 cell lines and NIH 3T3 derived cell lines were trypsinized using a 1/25 dilution of the stock trypsin, all other cell lines were treated with a 1/10 dilution of the stock. The culture medium was removed from the flask, care was taken to remove as much medium as possible as the serum inhibits the action of trypsin. The activated trypsin was added to the flasks and gently washed over the cells, about 2ml of trypsin was required for the 75cm² flasks. The cells were incubated at 37°C for

approximately 5 minutes or until the cells could be gently dislodged from the surface of the flask. Fresh medium was added to inactivate the trypsin and the cells were seeded into fresh flasks. Usually the cells were split about 1 in 4 every 2-3 days.

M.1.3 STORAGE OF CELL LINES

Cell line stocks were stored in liquid nitrogen. To prepare cells for freezing, two confluent 75cm² flasks were trypsinized as described above and the cell suspension was added to 20 ml of fresh medium plus serum and pelleted at 2500rpm in an MSE Centaur 2 benchtop centrifuge. The cell pellet was resuspended in 2ml of freezing solution I and 2ml of freezing medium II was added drop by drop. 1ml aliquots were frozen by slowly reducing their temperature to -70°C before placing them in liquid nitrogen.

M.1.4 CULTURING CELLS FROM FROZEN STOCKS

The frozen cells were warmed rapidly to room temperature and 20ml of DMEM plus serum was added to the cells drop by drop. The cells were pelleted at 2500rpm and resuspended in 8ml of fresh culture medium. The cells were seeded into two 25cm² flasks and incubated at 37°C for 24 hours. The culture medium was then changed to remove any dead cells and the cells cultured as usual.

M.1.5 EXPERIMENTAL PROCEDURES WITH CULTURED CELLS

In order to obtain quiescent cultures the cells were seeded at a density that would give a confluent monolayer the next day. Confluent cultures were washed with serum free medium and then grown in serum free medium for 48 hours. In serum stimulation experiments quiescent cells

were stimulated with 15% FCS for 1-6 hours before harvesting. In experiments using the phorbol ester 12-0-tetradecanoyl-13-phorbol acetate (TPA) the drug was added to quiescent cells at a final concentration of 100nM for 1-6 hours before harvesting. The protein kinase C inhibitor, H-7 (132), was used in some serum and TPA stimulation experiments. The drug was added to cultures at a final concentration of 140 μ M 30 minutes prior to stimulation with either serum or TPA.

To obtain samples from steady state cells they were harvested when slightly subconfluent.

M.1.6 RETROVIRUS MEDIATED GENE TRANSFER

The fact that VL30 elements can be packaged by retrovirus particles (116,124) was utilized in order to transfer VL30 elements into heterologous cell types. The NIH 3T3 cell line A1 is infected with KiMLV and can therefore produce KiMLV(VL30). A 75cm² flask of A1 cells was grown until almost confluent and then fresh medium was placed on the cells for several hours. The RAT1 cells that were to be infected with the KiMLV(VL30) were treated with 25 μ gml⁻¹ DEAE-Dextran (in DMEM/10%FCS) for 30 minutes. The DEAE-Dextran was then removed and replaced with the filter sterilized conditioned medium from the A1 cells. The cells were incubated at 37°C overnight and then cultured as usual.

M.2 PREPARATION OF PLASMID DNA

Plasmid preparations were made using standard techniques or modifications of these techniques (133).

M.2.1 SMALL SCALE PLASMID PREPARATIONS

A single bacterial colony was picked and grown overnight in 3mls of L-broth plus $100\mu\text{gml}^{-1}$ ampicillin at 37°C with shaking. 1ml of the overnight culture was centrifuged at 10000rpm for 1 minute and the bacterial pellet resuspended in $100\mu\text{l}$ of TELT. $10\mu\text{l}$ of freshly prepared lysozyme solution (10mgml^{-1} in 20mM Tris pH7.5) was added, the sample was vortexed and incubated on ice for 1 minute. The sample was boiled for 1 minute and incubated on ice for a further 5 minutes. After spinning for 8 minutes at 10000rpm the supernatant was precipitated with 2.5 volumes of ethanol, pelleted, washed, and dried. The DNA pellet was redissolved in $20\mu\text{l}$ of 1/10 TE, this provided sufficient DNA for 2-3 small scale restriction enzyme digests.

M.2.2 LARGE SCALE PLASMID PREPARATIONS

M.2.2.1 Amplification of Low Copy Number Plasmids

In order to increase the yield of low copy number plasmids, such as pAT153, the cells were treated with chloramphenicol. This antibiotic prevents bacterial replication, but not plasmid replication. 10ml of L-broth plus ampicillin at $100\mu\text{gml}^{-1}$ were inoculated with the bacterial stock and incubated overnight at 37°C with shaking. 15ml of L-broth plus ampicillin was inoculated with $50\mu\text{l}$ of the overnight culture and grown until the OD_{600} was approximately 0.6. 250ml of L-broth plus ampicillin was then inoculated with 12ml of the late log culture and incubated at 37°C with shaking for 2.5 hours until the OD_{600} was 0.4, then chloramphenicol was added to final concentration of $170\mu\text{gml}^{-1}$. The culture was grown overnight at 37°C with shaking and plasmid DNA was prepared as described below.

M.2.2.2 Large Scale Cleared Lysate

250ml of L-broth plus $100\mu\text{gml}^{-1}$ ampicillin were inoculated with 0.5ml of an overnight culture and grown at 37°C with shaking overnight (unless amplified as described above).

250ml of the bacterial culture was centrifuged at 5000rpm in an MSE Highspeed 18 centrifuge. The pellet was resuspended in 10mls of lysis buffer containing 5mgml^{-1} lysozyme and incubated at room temperature for 5 minutes. 20 mls of 0.2M NaOH, 1%SDS were added and the tube inverted to mix the sample. After a 10 minute incubation on ice 15mls of ice cold 5M KOAc pH4.8 was added and the sample was incubated on ice for a further 10 minutes. The sample was centrifuged at 10000rpm for 10 minutes and the supernatant mixed with 0.6 volumes of ice cold isopropanol and incubated at room temperature for 15 minutes. The sample was pelleted by spinning at 12000rpm for 10 minutes at 4°C and the pellet washed with 70% ethanol and dried. The dry pellet was redissolved in 10mM Tris pH7.5, 10mM EDTA and treated with $20\mu\text{gml}^{-1}$ RNAaseA for 2 hours at 37°C followed by digestion with $100\mu\text{gml}^{-1}$ proteinase K for 2 hours at 37°C . The sample was extracted with phenol/chloroform, ethanol precipitated, washed twice with 70% ethanol and dried.

M.2.2.3 G50 Column Purification

The plasmid DNA was purified by chromatography through a Sephadex G50 column prepared in a plastic 10ml pipette. The tip of the pipette was blocked using glass wool and glass beads. The swollen G50 was carefully dripped onto the beads until the pipette was almost full. The column was washed exhaustively with TE before use in order to equilibrate it and to remove any impurities. The DNA sample was redissolved in 250-500 μl of TE and loaded onto the column. 1ml fractions were collected and monitored

on a 1% agarose mini-gel. Fractions containing DNA were pooled and ethanol precipitated, DNA pellets were washed twice with 70% ethanol and dried. The final pellet was redissolved in a small volume of 1/10TE and monitored on a 1% agarose minigel alongside known concentrations of DNA in order to determine the concentration.

M.2.2.4 Purification of Plasmid DNA by PEG 8000

Precipitation

Plasmid DNA to be used in transfections was further purified by precipitation with polyethylene glycol (PEG) 8000. The DNA pellet was redissolved in 0.5ml of TE and mixed well with 0.5ml of 20% PEG 8000, 2.5M NaCl. The DNA was precipitated at +4°C overnight and then spun at 10000rpm for 5 minutes. The supernatant was removed with great care and the DNA pellet redissolved in 0.5ml of TE. The DNA was precipitated with 50 μ l of 3M NaOAc pH5 and 550 μ l of isopropanol. The sample was snap frozen, pelleted and washed twice with 70% ethanol. The pellet was dried and redissolved in sterile water. The concentration was determined as described above.

M.2.3 STORAGE OF BACTERIAL STOCKS

450 μ l of the overnight culture was mixed with 50 μ l of 10x Hogness freezing solution and incubated at room temperature for 30 minutes. The samples were then frozen and stored at -70°C.

M.3 MANIPULATION OF CLONED DNA

All manipulations were carried out using standard techniques or slight modifications of these techniques (133).

M.3.1 RESTRICTION ENZYME DIGESTS

Small scale restriction digests using upto $1\mu\text{g}$ of DNA were usually carried out in a final volume of $20\mu\text{l}$. The DNA was added to $2\mu\text{l}$ of the appropriate 10 x restriction buffer, an approximately 5-fold excess of the enzyme and water to make the volume upto $20\mu\text{l}$. The digest was allowed to run for 1-2 hours before checking the extent of the digest by gel electrophoresis, if necessary further enzyme was added and the digest allowed to continue for a further hour before checking again. When larger amounts of DNA were to be digested ($10\mu\text{g}$) then the volume was increased to $100\text{-}200\mu\text{l}$, all the other reagents were increased accordingly. When two different enzymes were used the procedure depended on the assay conditions required. If both enzymes required similar assay conditions then the digests were done together. An alternative was to do the digests sequentially, sometimes it was possible to increase, for example, the salt concentration after the first digest and then add the second enzyme directly. If the conditions required for the second enzyme differ greatly from those for the first then the DNA has to be precipitated before doing the second digest.

M.3.2 AGAROSE GEL ELECTROPHORESIS

Agarose minigels were prepared using 1 x minigel buffer. The percent agarose used depended on the size of the fragments to be resolved, the larger the fragments the lower the percentage of gel used. For increased resolution larger gels were used, these were prepared with 1 x E buffer. All gels included ethidium bromide to a final concentration of $1\mu\text{gml}^{-1}$. Minigels were run for about 45 minutes at 50-60 volts and large gels were run overnight at 35 volts. In some cases it was desirable to recover the DNA from the gel, for example when preparing insert DNA,

in these cases low melting agarose was used, this agarose melts at 65°C. Gels prepared from LMA were made and used exactly as normal agarose gels, however care was taken not to run the gels at too high a voltage because this could lead to the gel melting. LMA gels tend to behave as though they were a much lower percentage than normal gels containing the percentage of agarose, this could cause problems if the fragments to be resolved were quite large as LMA gels much below 1% are very hard to handle.

M.3.3 ETHANOL PRECIPITATION OF NUCLEIC ACIDS

DNA and RNA samples were ethanol precipitated by increasing the concentration of NaCl to 200mM and adding 2.5 volumes of ethanol. Samples were then snap cooled at -70°C, spun at 10000rpm for 5 minutes and the pellet washed well with 70% ethanol. The pellet was dried in a dessicator and then stored either as a dry pellet or resuspended in an appropriate volume of 1/10 TE (DNA) or DEPC treated TE/0.5%SDS (RNA).

M.3.4 PHENOL/CHLOROFORM EXTRACTION

In order to remove proteins (such as after restriction digests) from DNA samples they were extracted with phenol/chloroform. The salt concentration was increased to 200mM and an equal volume of phenol/chloroform was added to the sample. The sample was mixed well and then spun at 10000 rpm for five minutes to separate the two phases. The aqueous phase was carefully removed and precipitated with 2.5 volumes of ethanol as described above.

M.3.5 CTAB/BUTANOL EXTRACTION

This process extracts DNA from LMA gels (134). The CTAB/butanol and the equilibrated water were prepared as follows. 150ml of water and 150ml of butanol were shaken well together and the phases allowed to separate. 1g of CTAB was added to 100ml of butanol. The butanol was then mixed with the water and 50 μ l of Antifoam A. The phases were allowed to separate out overnight and then bottled and stored separately. The plasmid harbouring the DNA fragment required was digested with the appropriate restriction enzyme or enzymes and electrophoresed on the correct percentage LMA gel. After electrophoresis the required fragment was cut carefully from the gel taking care to remove as much of the surrounding agarose as possible. The gel was melted at 65°C and 10 μ g of carrier tRNA added. The sample was immediately transferred to a 37°C waterbath and equal volumes of CTAB/butanol and equilibrated water were added to the sample. The sample was gently mixed for 3-5 minutes and then briefly spun to separate the phases. The butanol phase was removed and a further volume of butanol added to the sample, this process was repeated 3 times in total and the three butanol samples pooled. This sample was then extracted with an equal volume of the equilibrated water and the butanol phase saved. This butanol sample was then extracted twice with a 1/4 volume of 200mM NaCl. The NaCl samples were then pooled and mixed with an equal volume of ice cold chloroform. This sample was placed on ice for 30 minutes with frequent shaking to keep the two phases mixed. The sample was spun briefly at 2500rpm at 4°C to separate the phases and the aqueous phase was ethanol precipitated, washed with 70% ethanol and dried as described previously. The DNA pellet was redissolved in a small volume (10-20 μ l) of 1/10 TE.

M.3.6 PREPARATION OF VECTOR DNA

Usually the high copy number bluescript vector, pKS was used. $1\mu\text{g}$ of the vector was cut with the appropriate restriction enzyme or enzymes (section M.6.1). The digest was checked by mini gel electrophoresis, extracted with phenol/chloroform and ethanol precipitated. In certain situations, such as when there was a large insert to be removed from the vector, a larger amount of DNA was ^{digested} \wedge ($10\mu\text{g}$) and the appropriate band cut from an LMA gel after the 5' phosphate group had been removed (see below).

M.3.7 REMOVAL OF 5' PHOSPHATE GROUPS

The 5' phosphate groups were removed from vectors in order to prevent the recircularization of the vector without an insert. The DNA pellet was redissolved in $50\mu\text{l}$ of 20mM Tris pH8. The calf intestinal alkaline phosphatase (CIAP) ($24\text{units}\mu\text{l}^{-1}$) was diluted 1/50 in 20mM Tris pH8 and $0.5\mu\text{l}$ was added to the DNA and the reaction incubated at 37°C for 30 minutes. A further $0.5\mu\text{l}$ of the diluted enzyme was added and the reaction incubated for a further 30 minutes, this input is the minimum required to prevent recircularization. The reaction was phenolated twice and the DNA ethanol precipitated. The dry pellet was redissolved in $20\mu\text{l}$ of 1/10TE and stored at -20°C .

$1\mu\text{l}$ of the final DNA sample was monitored on a 1% agarose gel to check the concentration.

M.3.8 PREPARATION OF INSERT DNA

$1\mu\text{g}$ of DNA was cut with the appropriate restriction enzyme or enzymes and the digest monitored by gel electrophoresis. The ^{digested} \wedge DNA was electrophoresed on an LMA mini gel and the correct band cut carefully

from the gel. The DNA was either eluted from the gel by the CTAB/butanol method or used directly from the gel slice. The DNA was then used in ligations as described below.

M.3.9 LIGATION OF VECTOR AND INSERT

M.3.9.1 Ligations without LMA

Ideally ligations were carried out with approximately 50-100ng of DNA in a volume of 10 μ l. In these experiments the vector DNA was considered to account for almost all of the DNA input and the input of insert was negligible. 1-2 μ l of the vector DNA (50-100ng) were mixed with 1/2 of the insert (5 μ l), plus 2 μ l of 5 x ligase buffer, 1 unit of T4 DNA ligase and water to make up the volume to 10 μ l. The ligation was incubated overnight at room temperature. Those samples not used immediately could be stored at -20°C.

M.3.9.2 Ligations with LMA

If the insert was in LMA then the reaction volume was increased to 25 μ l, upto 1/2 of this volume could be LMA. When the vector was also in LMA 2 μ l of the DNA input was vector with the rest being insert. The reaction was carried out in 1 x ligase buffer and 1 unit of T4 DNA ligase, the volume was made up to 25 μ l with water. The reaction was incubated overnight at 37°C to prevent the LMA setting.

M.3.10 USE OF ARTIFICIAL LINKERS

M.3.10.1 Preparation of DNA FOR linker ligation

In some situations the insert DNA did not have the appropriate restrictions site or sites, in these cases artificial restriction sites could be created by ligating synthetic linkers containing the sequence

of the restriction site to blunt ended insert DNA. When linkers were to be used the initial DNA input was increased to 5-10 μ g. The DNA was cut and monitored as described above, but in a final volume of 100 μ l. In order to fill in the cut ends of the DNA 5 μ l of 15mM dNTPs and 1 μ l of the Klenow fragment of DNA polymerase I were added to the digest and the DNA incubated at room temperature for 15 minutes. The DNA was then extracted with phenol/chloroform and ethanol precipitated, washed with 70% ethanol and dried.

M.3.10.2 Phosphorylation of linkers

Synthetic linkers were prepared using an oligonucleotide synthesizer as described in section M.4. The linkers were treated with polynucleotide kinase to phosphorylate the 5' ends. 3 μ g of the linkers were mixed with 2 μ l of 5 x ligase buffer and 9 units of PNK. The volume was made up to 10 μ l with water and incubated at 37°C for 1 hour.

M.3.10.3 Ligation of linkers and insert DNA

The phosphorylated linkers were added to the dried insert DNA, 1 unit of T4 DNA ligase was added and the reaction was incubated at room temperature overnight. The reaction volume was increased to 100 μ l with 10 μ l of the correct 10 x restriction buffer and water and the DNA cut with a vast excess of the enzyme for that particular linker. If necessary this was followed by further digests required to give the correct ends. The ligated insert was then ethanol precipitated and the dry pellet redissolved in a small volume of 1/10TE.

M.3.10.4 Removal of excess linkers

Any linkers not ligated to the insert DNA must be removed before ligation of the insert and vector otherwise the linkers can be cloned in instead of the insert resulting in a large number of false positives. This was achieved by electrophoresing the DNA on two LMA gels. The correct band was cut from the first gel and melted at 65°C, the DNA was then loaded onto a second gel that had been prewarmed. The band was cut from the second gel and washed well to remove any remaining linkers present in the running buffer. The insert was then ligated in the usual way (see M.3.9.2).

M.3.11 BACTERIAL TRANSFORMATIONS

M.3.11.1 Preparation of competent cells

A 10ml culture of JM103 *E.coli* cells was grown up overnight at 37°C with shaking. 50mls of L-broth were inoculated with 0.5ml of the overnight culture and grown at 37°C with shaking until the OD₆₀₀ of the culture was about 0.5. 10 mls of the culture were pelleted at 2000rpm in a glass universal and the pellet resuspended in 5mls of ice cold 100mM MgCl₂. The cells were pelleted at 2rpm and resuspended in 1.5ml of 100mM CaCl₂. The cells were incubated on ice for 30 minutes with occasional gentle agitation. The cells were then competent for taking up DNA. The cells could be kept on ice or at 4°C for several hours, but tended to lose efficiency if left any longer.

M.3.11.2 Transformation

The DNA samples used in the transformation were made up to a final volume of 50µl with TCM buffer. A 1ng sample of uncut vector was included to monitor for transformation efficiency. The samples were kept on ice

unless they contained LMA, in which case they were kept at 37°C until just before use. 100 μ l of competent cells were added to each DNA sample and gently mixed in. The samples were incubated on ice for 40 minutes with occasional gentle agitation to keep the cells in suspension. The cells were then heat shocked at 42°C for 2 minutes and returned to ice for 20 minutes. 0.5ml of prewarmed S.O.C. medium was added to each sample and the samples were incubated at 37°C for 1 hour with shaking. 300 μ l of each transformation was spread onto L-broth agar plates containing 100 μ gml⁻¹ ampicillin. If colour selection was used then 50 μ l of 100mM IPTG and 20 μ l of 2% X-GAL were added before the transformation was plated out. The plates were allowed to dry and then were incubated at 37°C overnight. If the colour selection had been used then the plates were transferred to 4°C to allow maximum development of the colour. Blue colonies represented bacteria harbouring wild type plasmids, white colonies represented those containing recombinant plasmid molecules. Recombinants were checked by making small scale plasmid preparations and small scale restriction enzyme analysis. When colour selection was not available then the colonies were screened by hybridization (see section M.6).

M.4 OLIGONUCLEOTIDE SYNTHESIS

Synthetic oligonucleotides were prepared on a Biosearch Cyclone DNA synthesizer according to the manufacturers instructions. The column containing the oligonucleotide was stored at 4°C. The oligonucleotide was liberated from the solid support using ammonium hydroxide. The ammonium hydroxide was drawn through the column with a syringe and then pushed back and forth through the column using syringes attached to both ends of the column. The ammonium hydroxide was left in the column for 1.5 hours, after 45 minutes the ammonium hydroxide was pushed through the

column again. After this stage the ammonium hydroxide was withdrawn from the column taking the oligonucleotide with it. The solution was incubated at 55°C for 5 hours and then sodium hydroxide was added to a final concentration of 5mM and the sample was freeze dried overnight.

The oligonucleotide was further purified through a NENSorb Prep cartridge which removes the trityl group. The cartridge was activated with 10ml of methanol and pre-equilibrated with 5ml of 0.1M TEAA pH7. The sample was redissolved in 4ml of TEAA and pulled gently through the column. The column was washed with 10ml of 10% acetonitrile and 3 x 7ml of TEAA. The oligonucleotide was eluted from the column with 5ml of 35% methanol/65% water. Six 1ml fractions were collected and the OD₂₆₀ of each fraction was checked in order to locate the oligonucleotide. The samples were freeze dried and stored dry at -20°C.

M.5 PREPARATION OF ³²P LABELLED DNA PROBES

M.5.1 RANDOMLY PRIMED PROBES

Probes were prepared using a modified version of the oligo priming reaction (135). For each probe approximately 50-100ng of DNA was used. The DNA to be labelled was either in 1/10TE after CTAB/butanol extraction (M.3.5) or in an LMA gel slice. The DNA was denatured by boiling for 3 minutes (5 minutes if in LMA). The denatured DNA was then added to 1μl of dNTP mix, 1μl of oligodeoxynucleotide hexamers, 1μl of 10 x NT, 4μl of [α -³²P]dCTP and 1μl of Klenow, the final volume was made up to 10 μl and the reaction incubated at 20°C overnight. If the DNA was in LMA then the final volume was increased to 15μl and the input of the various reagents increased accordingly. The reaction was incubated at 25°C to prevent the LMA from setting.

Unincorporated labelled nucleotides were removed by spinning the sample through a 1ml G50 spun column. The column was prepared in a 1ml Gilson pipette tip, the end of the tip was blocked with glass wool and glass beads and the swollen G50 carefully dropped onto the plug. The tip was spun at 2000rpm very briefly to compact the G50, this procedure was repeated several times until the tip was almost full. The column was then spun for four minutes at 2000rpm, 100 μ l of TE was added to the column and it was spun at 3000rpm for 4 minutes. The volume of the probe was increased to 100 μ l with TE and 4 μ l of 5M NaCl and 1 μ l of carrier salmon sperm DNA were added. If the DNA sample was in LMA then the reaction was heated to 65°C for a few minutes before the next stage. The reaction was extracted with phenol/chloroform and spun through the column at 3000rpm for 5 minutes. 1 μ l of the final probe was counted in liquid scintillation fluid in the β counter. Before use the probe was denatured by boiling and rapidly cooled on ice.

M.5.2 END-LABELLING OF DNA RESTRICTION FRAGMENTS

1 μ g of DNA was cut with the appropriate restriction enzyme as described in section M.3.1. The DNA was then ethanol precipitated, washed with 70% ethanol and dried. The DNA was treated with CIAP (M.3.7) and extracted twice with phenol/chloroform. The sample was then reprecipitated, washed and dried. The pellet was redissolved in 3 μ l of water, 5 μ l of [γ -³²P]ATP, 1 μ l of 10 x PNK buffer and 9 units of PNK. The reaction was incubated at 37°C for 45 minutes and stopped by the addition of 50 μ l of TE and 1.5 μ l of 20% SDS.

This procedure was used to prepare ³²P labelled MspI fragments of pBR322 DNA for use as size markers on sequencing gels.

M.5.3 END-LABELLING OF OLIGONUCLEOTIDES

100ng of the synthetic oligonucleotide was labelled with [γ -³²P]ATP essentially as described above. However the reaction volume was decreased to 5- μ l and the PNK was added in two lots at intervals of 45 minutes. If the probe was to be used in filter hybridizations then it was spun through a G25 column in order to remove any unincorporated label, this column was prepared in the same way as the G50 column described above.

M.6 COLONY SCREENING

M.6.1 PREPARATION OF FILTERS

In some situations it was more appropriate to screen bacterial colonies for the correct insert using hybridization techniques, particularly when colour selection could not be used. In this case transformations were plated out directly onto nitrocellulose filters placed on L-broth agar plates supplemented with 100 μ gml⁻¹ ampicillin. The plates were incubated overnight at 37°C. Two replicas were made from each filter by placing fresh nitrocellulose filters over the original plate and gently transferring the colonies to the new filters. These new filters were then placed onto fresh plates and incubated overnight to allow the colonies to regrow. One copy of each filter was treated as follows to bind the DNA to the nitrocellulose. The filters were placed colony side up onto Whatman 3mm paper soaked in 0.5M NaOH for 5 minutes to denature the DNA followed by 5 minutes on 1.5M NaCl, 0.5M Tris pH7.4 to neutralize the filters. Finally the filters were placed onto 1.5M NaCl, 2xSSC for 5 minutes before air drying the filters for an hour and baking them for 2 hours at 80°C. The second plate was stored at 4°C.

M.6.2 HYBRIDIZATION OF THE FILTERS

The filters were prehybridized for several hours in formamide prehybridization solution at 42°C. Hybridizations were carried out in formamide hybridization solution plus $0.5-1 \times 10^6$ cpm/ml⁻¹ of the appropriate ³²P labelled probe at 42°C overnight. Filters were washed exhaustively in 2xSSC at room temperature, followed by a 20 minute wash in 0.1xSSC at 50°C. The filters were autoradiographed overnight at -70°C with intensifying screens. Positive colonies were picked from the second replica plate and checked further using small scale plasmid preparations and restriction enzyme analysis.

M.7 DIDEOXY CHAIN TERMINATION DNA SEQUENCING

M.7.1 PREPARATION OF M13 SUBCLONES

M.7.1.1 Cloning Manipulations

Fragments of the NVL-1 and NVL-2 5' LTRs were cloned into the replicative form (RF) of M13mp10 and M13mp11 using the standard techniques described in section M.3. Restriction sites were chosen that resulted in overlapping fragments of the LTRs being inserted and by cloning into both M13mp10 and M13mp11 it was possible to sequence both strands of the restriction fragments (136).

M.7.1.2 Transformation of *E.coli* with M13 Bacteriophage

The principles of transformation with M13 are similar to transformation with plasmid DNA. Competent JM103 *E.coli* cells were prepared and kept on ice. 300 μ l of cells were added to the ligated vector and insert DNA and incubated on ice for 40 minutes with occasional gentle agitation. The cells were heat shocked for 2 minutes at 42°C. Each transformation was mixed with 10 μ l of 100mM IPTG, 50 μ l of 2% X-GAL, 200 μ l

of exponentially growing JM103 cells and 3ml of top agar, the mixture was poured onto L-broth agar plates and allowed to set at room temperature before incubating overnight at 37°C. Blue plaques represented wild type M13 and white plaques represented recombinant phage.

M.7.2 PREPARATION OF SINGLE STRANDED DNA TEMPLATES

White plaques were picked into 1.5ml of L-broth and grown overnight at 37°C with shaking. 1ml of each culture was spun at 10000rpm for 10 minutes and the supernatant saved. 20 μ l of the supernatant was added to 1 μ l of 250mM EDTA and 1 μ l of 20% SDS and run on a 0.8% agarose gel overnight in parallel with wild type M13 in order to check for the presence of an insert. 300 μ l of 20% PEG 6000, 2.5M NaCl was added to the remaining supernatant and the sample was incubated at room temperature for 20 minutes. The tubes were spun for 7 minutes at 10000rpm and the PEG removed carefully. The pellet was resuspended in 200 μ l of 4% PEG 6000, 0.5M NaCl and pelleted at 10000rpm for 10 minutes. The PEG was removed carefully as it can act as a contaminant in the sequencing reactions. The pellet was redissolved in 100 μ l of TES and SDS added to a final concentration of 0.5% and incubated on ice for 5 minutes. The sample was extracted with 100 μ l of phenol/chloroform and 80 μ l of the aqueous layer was precipitated with 3 μ l of 3M NaOAc and 250 μ l of ethanol, washed twice with 70% ethanol and dried. The final pellet was resuspended in 20 μ l of 1/10 TE and 1 μ l was checked on a 0.8% agarose gel.

M.7.3 SEQUENCING

An Amersham DNA sequencing kit was used for these reactions.

M.7.3.1 Annealing of Template and Primer

The single stranded templates were annealed to the universal primer dGTAAAACGACGGCCAGT as follows. 5 μ l of template was mixed with 1 μ l of primer, 1.5 μ l of 100mM Tris pH8, 50mM MgCl₂ and 2.5 μ l of deionised water and heated to 95°C for 2-3 minutes. The samples were transferred to a 60°C waterbath for 1 hour before being allowed to cool slowly to room temperature.

M.7.3.2 Sequencing Reactions

N° mixes were prepared in advance from the stock dNTPs as shown in the table below.

	A°	C°	G°	T°
dCTP	20 μ l	1 μ l	20 μ l	20 μ l
dGTP	20 μ l	20 μ l	1 μ l	20 μ l
TTP	20 μ l	20 μ l	20 μ l	1 μ l
TE	20 μ l	20 μ l	20 μ l	20 μ l

Table M.1 N° mixes

The working solutions (N mixes) were prepared just prior to use by mixing equal volumes of the N° mix and the corresponding ddNTP. For each single stranded template the sequencing reaction was carried out as follows. 3 μ l of [α -³⁵S]dATP_γS and 1 μ l of the Klenow fragment of DNA polymerase I were added to the annealed template and primer. 2.5 μ l of this mixture was added to 2 μ l of each of the four N mixes and incubated at room temperature for 20 minutes. 2 μ l of the chase solution was added and the samples were incubated for a further 15 minutes at room temperature. The reactions were stopped by the addition of 4 μ l of urea loading buffer. If the reactions were not to be run immediately then they were stored at -20°C before adding the loading buffer.

M.7.3.3 Gel Electrophoresis

The sequencing gel was prepared as follows. The glass plates were washed with water and detergent and then rinsed with ethanol and distilled water. The notched plate was treated with the siliconising agent REPELCOTE. The other plate was treated with the silanising agent BindSilane which covalently binds the polyacrylamide to the glass. 8M urea, 5% polyacrylamide, and 1 x sequencing buffer gels were prepared. The gel was polymerised with 300 μ l of 25% ammonium persulphate and 50 μ l of TEMED. The gels were 40cm x 80cm and 0.3mm thick. Gels were pre-run for 60 minutes in 1 x sequencing gel buffer before the samples were loaded. The samples were denatured at 95°C for 3 minutes and 2 μ l were loaded. Gels were run at 35Watts until the bromophenol blue had reached the bottom, then another 2 μ l of each sample was loaded and the gel was run until the bromophenol blue of the second loadings had reached the bottom. The notched plate was removed and the gel fixed in 10% acetic acid, 10% methanol and dried at 80°C. Gels were autoradiographed at +4°C overnight or longer if necessary.

M.7.4 ANALYSIS OF SEQUENCING GELS

The DNA sequence for both strands of each fragment was read and the entire sequence assembled. A search of the NIH database was carried out using the DNASTAR sequence analysis programme.

M.8 ANALYSIS OF CELLULAR RNA

M.8.1 PREPARATION OF TOTAL CELLULAR RNA

Cell monolayers were washed briefly with DNA extraction buffer. Sufficient GITC lysis solution was added to cover the monolayer and the cells were incubated at room temperature for about 5 minutes, the lysed

cells were removed and the flask washed with a small volume of the GITC lysis solution. The samples were pooled and stored at -70°C if not to be used immediately. The RNA was isolated by centrifugation through a CsCl cushion (137), the cell lysate was carefully layered over 1.5ml of 5.7M CsCl in a 14ml polyallomer centrifuge tube. The remainder of the tube was filled with deionised water. The tubes were spun at 35000rpm for 20 hours at 20°C . The supernatant was removed carefully in order not to disturb the RNA which pellets at the bottom of the tube. The DNA occurs as a viscous band about 1cm above the base of the tube, great care was taken to remove all the DNA so that it did not contaminate the RNA. The pellet was washed carefully with 1ml of fresh CsCl and care was taken to remove any contaminating GITC from the lower sides of the tube. The RNA was dissolved in 400 μl of DEPC-treated TE, 0.5%SDS and the sample was spun at 10000rpm for 5 minutes to remove any insoluble contaminants. The RNA was then precipitated with ethanol, washed with 70% ethanol and dried. The final pellet was dissolved in 100 μl of DEPC-treated TE, 0.5%SDS and the concentration of the RNA determined by measuring the OD_{260} . The samples were monitored for degradation by minigel analysis.

M.8.2 NORTHERN BLOTS

1 μg samples of RNA were made up to a final volume of 7 μl and mixed with an equal volume of Northern blot denaturing solution. The samples were heated to 65°C for 4 minutes and then cooled rapidly in ice water. 2 μl of glycerol loading buffer was added to each sample and the samples were electrophoresed on 1% agarose, 0.8M formaldehyde gels in 1 x MOPS solution at 100volts for 60 minutes (138). The gels were stained for 30 minutes in the running buffer plus 1 μgml^{-1} ethidium bromide in order to

check the loading of the RNA. The RNA was blotted onto nylon filters using 10xSSC. Blots were baked at 80°C before use.

M.8.3 RNA DOT BLOTS

4 μ g of RNA was made up to a final volume of 10 μ l and an equal volume of dot blot denaturing solution was added. The samples were placed at 50°C for 15-30 minutes and then cooled rapidly in ice water. DEPC treated water was added to give a final volume of 400 μ l and 225 μ l, 75 μ l and 25 μ l volumes of each sample were blotted onto nylon filters by slow suction through a dot blot apparatus. The blots were baked at 80°C for 2 hours before use.

M.8.4 HYBRIDIZATION OF NORTHERN AND RNA DOT BLOTS

Filters were prehybridized for 1 hour in RNA annealing solution at 60°C in a shaking water bath. Hybridizations were carried out for 20 hours at 60°C in RNA annealing solution plus 0.5-2x10⁶ cpmml⁻¹ of ³²P labelled probe in sealed plastic bags. Filters were washed exhaustively in 2xSSC at room temperature. The filters were then washed in 2xSSC, 1%SDS at 60°C for 20 minutes, followed by a wash in 0.1xSSC for 20 minutes. The temperature of the final wash depended on the stringency required. Filters were wrapped in cling film and autoradiographed at -70°C with an intensifying screen.

M.8.5 STRIPPING PROBES FROM FILTERS

In some cases it was desirable to strip the probe from the filter and reprobe. This was done by incubating the filter in formamide hybridization solution for 30 minutes at 60°C followed by rinsing in 2xSSC. This process worked best if the filters were not allowed to dry

before trying to strip them. Stripped filters were then prehybridized and hybridized as described above.

M.9 NUCLEAR RUN-ON TRANSCRIPTION

M.9.1 PREPARATION OF ISOLATED NUCLEI

10^7 NIH 3T3 cells were serum starved for 48 hours and then stimulated with 15% FCS for 0, 2 or 5 hours. The cells were trypsinized as described in section M.2.2, added to 20 mls of ice cold PBS and pelleted at 2000rpm in an ice cold centrifuge bucket. All the following procedures were carried out on ice. The cell pellet was resuspended in 1ml of ice cold lysis buffer containing 1μ l placental RNAase inhibitor. The cells were gently disaggregated using a glass homogenizer and incubated on ice for 10 minutes. The cells were lysed by the addition of 10μ l of 20% NP40 and incubated on ice for a further 5 minutes. The cell lysate was layered over a 30% sucrose cushion prepared in lysis buffer and centrifuged at 10000rpm for 2.5 minutes. The pelleted nuclei were gently washed with storage buffer and resuspended in 18μ l of storage buffer and 12μ l of glycerol. The nuclei were stored at -70°C (139).

M.9.2 NUCLEAR RUN-ON

10μ l of nuclei were mixed with an equal volume of 2 x transcription mix (139) and incubated at room temperature for 50 minutes. Contaminating DNA was removed by treatment with 1μ l of RNAase free DNAase for 1.5 minutes at room temperature. 280μ l of RES, 6μ l of proteinase K and 10μ g of carrier *E.coli* RNA was added to the transcription mix and incubated at 50°C for 30 minutes. The RNA was precipitated with ethanol and the pellet redissolved in 90μ l of RES plus 6μ l proteinase K and incubated at 50°C for a further 30 minutes. The RNA was phenol/chloroform extracted

and the aqueous phase mixed with 0.75 μ l of 2M acetic acid and 100 μ l of LiCl/ethanol. The samples were precipitated overnight at -20°C, pelleted at 10000rpm and the RNA dissolved in 100 μ l of DEPC treated TE, 0.5% SDS. The RNA was reprecipitated with ethanol, washed with 70% ethanol and dried. The pellet was redissolved in 100 μ l of hybridization solution. 1 μ l of the sample was counted in liquid scintillation fluid in the β counter.

M.9.3 PREPARATION OF DNA DOT BLOTS

For each dot 5 μ g of DNA was used. Each sample was made up to 25 μ l with water and equal volume of 1M NaOH was added. The sample was incubated for 10 minutes at room temperature. 25 μ l of 1M Tris pH 7.5 was added to neutralize the sample and it was placed on ice immediately and 25 μ l of 1M HCl and 125 μ l of 5M NaCl were added. The DNA was immobilized on nitrocellulose filters in 10xSSC by gentle suction through a Schleicher and Schuell dot blot manifold. Filters were baked at 80°C for 2 hours before use.

M.9.4 HYBRIDIZATIONS

DNA dot blots were prehybridized overnight at 42°C in the formamide prehybridization solution. Hybridisations were carried out for 48 hours in the formamide hybridization solution plus equal inputs of each labelled transcript. The final input of probe was approximately 2×10^6 cpmml⁻¹. After hybridization the filters were washed exhaustively in 2xSSC, 0.1% SDS at room temperature followed by two washes in 0.1xSSC, 0.1% SDS at 40°C. After several rinses in 2xSSC the filters were treated with 1 μ gml⁻¹ RNAaseA in 2xSSC at 20°C for 10 minutes. Filters were rinsed in 2xSSC, 0.5% SDS at room temperature and autoradiographed at -70°C with an intensifying screen.

M.10 TRANSIENT EXPRESSION ASSAYS

M.10.1 TRANSIENT TRANSFECTIONS

All transfections were carried out using the calcium phosphate precipitate procedure (140,141). 24 hours prior to transfection the cells were seeded into 60mm plates at a density that would give 75% confluent plates the following day. 2-3 hours before the transfection the cells were washed with fresh medium containing 10%FCS and gentamicin. Calcium phosphate precipitates were prepared as follows. The DNA sample was made

up to 120 μ l with sterile water and an equal volume of buffer A was added.

The sample was mixed well and incubated at room temperature for 10 minutes. The sample was added to 240 μ l of buffer B, mixed well and incubated at room temperature for 15 minutes. The precipitate was added to the culture medium and incubated overnight at 37°C in the usual culture conditions. The next day the cells were washed twice with fresh medium, at this stage NIH 3T3 cells were split into two plates, HeLa cells were not usually split. The cells were incubated for a further 48 hours and then harvested for analysis of CAT activity. In experiments using the protein kinase C inhibitor, H-7, the drug was added to a final concentration of 140 μ M after the cells had been split. Cells were then incubated for a further 48 hours as usual before being harvested and analyzed by CAT assay.

The monolayer was trypsinized as described (section M.2) and added to 10 mls of fresh medium and pelleted at 2500rpm. The cell pellet was resuspended in 1ml of PBS and spun down at 10000rpm, the cell pellet was resuspended in 50 μ l of 250mM Tris pH7.8 and the cells fractured by three rounds of freezing and thawing using dry ice/ethanol and a 37°C water bath. The cells were pelleted at 10000rpm and the supernatant saved. The

protein concentration was checked by spectrophotometry at OD_{280} so that equal inputs of protein could be used in each assay.

M.10.2 CAT ASSAYS

The CAT assays were done using a modified version of the method of Gorman *et al* (142). The cell extracts were heated to 65°C for 10 minutes before use in the CAT assay in order to inactivate any other enzymes that utilize AcCoA, the CAT enzyme is very stable and is not inactivated by this treatment. The cell extract was added to $5\mu\text{l}$ of 4mM AcCoA and $1\mu\text{l}$ of D-threo-[dichloroacetyl-1- ^{14}C]chloramphenicol, usually $25\mu\text{l}$ of cell extract was used and the volume was made up to $50\mu\text{l}$ with water. In some experiments a greater input of cell extract was used in order to see any response, also in some cases it was necessary to dilute the cell extracts in order to remain in the linear range determined by the stock enzyme, within any one experiment the same dilution of each cell extract was used. The linear range of the assay was determined using dilutions of the stock CAT enzyme in the standard assay (see Figure 16 in the results). The reactions were incubated at 37°C for 30 minutes to 24 hours and were stopped by the addition of $50\mu\text{l}$ of ethyl acetate. The samples were vortexed well and the phases separated by centrifugation at 10000rpm for 5 minutes. $35\mu\text{l}$ of the organic phase were loaded onto a silica gel thin layer chromatography plate. The samples were run in 19:1 chloroform:methanol in a chromatography tank that had been equilibrated for several hours, the samples were run until the solvent front was about 2/3 of the way up the plate. The plates were air dried for 30 minutes and then autoradiographed at room temperature. The results were quantitated by counting the spots corresponding to both the acetylated and unacetylated forms of the chloramphenicol in scintillation fluid on a β

counter. The results were expressed as percentage conversion and were normalized to the results seen for the controls pSV2CAT (142) or pBLCAT2 (143), constructs containing the CAT gene under the control of the SV40 enhancer and promoter or the TK promoter respectively.

M.11 PREPARATION OF STABLE TRANSFECTANTS

M.11.1 TRANSFECTIONS

Stable transfections were performed essentially in the same way as transient transfections. The plasmid DNA was prepared for transfection by digestion with the restriction enzyme PvuI, this enzyme cuts only within the plasmid sequences leaving the inserts intact. After digestion the DNA was extracted with phenol/chloroform, ethanol precipitated and the final dried pellet resuspended in a small volume of sterile water. For each transfection 4 μ g of linearized LTR-CAT plasmid was co-transfected with 0.5 μ g of linearized pSV2NEO (144), pSV2NEO expresses the bacterial neomycin resistance gene. The transfections were carried out as described (section M.10). After the cells had been washed they were split into 3 25cm² flasks containing 0.2mgml⁻¹, 0.6mgml⁻¹, or 1mgml⁻¹ G418 (geneticin), a mammalian homologue of the bacterial antibiotic neomycin. Three control flasks of untransfected cells were also cultured in these concentrations of G418. The cells were washed with fresh medium every 2-3 days until all the cells on the control flasks had died. The cells were then cultured in the medium plus antiobiotic until there were sufficient cells to use in serum stimulation experiments.

M.11.2 SERUM RESPONSE EXPERIMENTS

In order to investigate whether the transfected LTRs are responsive to serum, similar experiments were done to those described in section

M.1.6. For each time point two plates were prepared, the cells were made quiescent by growing in serum free medium for 48 hours and then stimulated for 0, 2 or 5 hours with 15% FCS. For each time point one plate was used to prepare protein extracts for analysis by CAT assay (section M.10.3) and the other was used to prepare RNA for Northern blotting (section M.8). In this way it was possible to check both the levels of VL30 driven CAT and the levels of endogenous VL30 RNA.

M.12 PRIMER EXTENSION

M.12.1 PREPARATION OF PROBE

A 24 base oligonucleotide corresponding to nt26-29 of the CAT gene mRNA was used in the primer extension analysis. The probe was labelled as described in section M.5.3 and run on an 8M urea 12% polyacrylamide sequencing gel. The gel was autoradiographed to detect the labelled oligonucleotide and the main labelled band cut from the gel. The acrylamide was mashed and incubated with 0.5ml of TE containing 20 μ g of carrier tRNA overnight at +4°C. The acrylamide was pelleted by spinning at 10000rpm for 5 minutes and the supernatant removed carefully and precipitated with ethanol as usual. The pellet was washed well with 70% ethanol and dried. The labelled probe was redissolved in 10 μ l of 1/10TE and 1 μ l of a 1/10 dilution counted in liquid scintillation fluid on the β counter. The probe could be stored at -20°C for a few days.

M.12.2 RNA PRIMER ANNEALING

15 μ g of RNA prepared from cells stably transfected with the NVL-CAT constructs was ethanol precipitated, washed and dried. 10 μ l of annealing solution and approximately 5x10⁵cpm of probe were added to each RNA sample. The samples were heated to 80°C for 3 minutes and then placed at

45°C for 4 hours. The volume of each sample was made up to 50 μ l with 1/10TE and ethanol precipitated. The samples were washed exhaustively with 70% ethanol and dried.

M.12.3 PRIMER EXTENSION

Each sample was incubated at 42°C for 40 minutes in the following 2 μ l 1/10TE, 27 units of placental RNAase inhibitor, 1 μ l of 5xPE buffer, 0.5 μ l of dNTPs and 7.5units of reverse transcriptase. The reactions were stopped by the addition of 5 μ l of urea loading buffer. The samples were denatured and run on an 8M urea, 8% polyacrylamide gel in 1 x sequencing buffer at 450volts until the bromophenol blue had run off the bottom. The samples were run alongside ³²P labelled MspI fragments of pBR322 DNA. The gel was fixed in 10% methanol, 10% acetic acid and dried onto DE81 paper. The gel was autoradiographed at -70°C with an intensifying screen.

RESULTS

R.1 NUCLEOTIDE SEQUENCE DIVERSITY AMONGST DISTINCT VL30

ELEMENTS

The NVL subset of VL30 elements represent the retrovirus transmissible, transcriptionally active members of this gene family (112). The LTRs of the clone NVL-3 have been sequenced previously (118) and it is of interest to see how the LTRs of the other members of the NVL group of elements, NVL-1 and NVL-2, compare to NVL-3. It is also of interest to see how the LTRs of this group compare with those of other members of the diverse VL30 gene family and other retrotransposons and retrovirus-like elements.

R.1.1 SEQUENCE HETEROGENEITY WITHIN THE NVL SUBGROUP OF VL30

ELEMENTS

Figure 1 shows the sequencing strategy used to determine the nucleotide sequences of NVL-1 and NVL-2. The same strategy was used for both clones. The BamHI EcoRI fragment was cloned into both M13mp10 and M13mp11, thus allowing the sequence to be determined for both strands of this fragment. The SacI BamHI fragment and the SacI PvuII fragments were both cloned into M13mp11. This strategy allowed overlapping sequences to be obtained for almost the entire length of the LTRs.

Figure 2 shows the nucleotide sequence for the 5' LTRs of the VL30 cDNA clones NVL-1 and NVL-2 (112). The two sequences are identical, with the exception that the NVL-2 LTR has 32bp at its 3' end that are absent in NVL-1, making the NVL-2 LTR 656bp long and the NVL-1 LTR 621bp long.

As mentioned previously the relationships between the various NVL LTRs are of interest. Figure 3 shows a comparison of the NVL-1/2 LTR with the previously sequenced NVL-3 LTR (118). The functional domains of the NVL-3 LTR were determined experimentally and the various regulatory

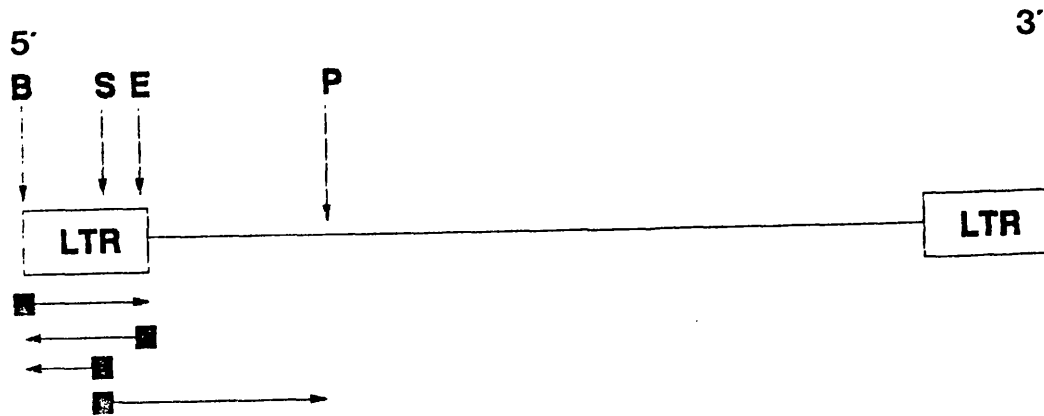


Figure 1. Nucleotide sequencing strategy for NVL-1 and NVL-2.
 B=BamHI, S=SacI, E=EcoRI, P=PvuII.

NVL-1

80
TGAAGAATGAAAAATTACTGGCCTCTTGTGAGAACATGAACTTTACCTCGGAGCCCACCCCTCCCATCTAGAAAAAT

160
TTTTGAGATAAAGGCCTCCTGGAACAACCTCAAAATGACATTGCCAAATGATAAGACATGAATCCTTAGTTACGTAGGTT

240
CCTTGATGGGACATGACTCCTTAGTTACGTAGGTTCCCTTGATAGGACATGACTCCTTAGTTACGTAGAATCCTTTGGCAG

320
AACCCCTTGTCCCTTGGCAGAACTCCCTAGTGATGAACTTGTACTTTCCCTGCCAGTTCTCCCCCTTTGAGTTTA

400
CTATATAAGCCTGTGAAAAATTTGGCTGGTCGTCGAGACTCCTCTACCCTGTGCAAAGGTGTATGAGTTTCGACCCAG

480
AGCTCTGTGTGCTTTCGTTTGTGCTTTATTTTCGACCCAGAGCTCTGGTCTGTGTGCTTTCATGTGCTGCTTTATTA

560
ATCTTACCTTCTACATTTTATGTATGGTCTCAGTGTCTTCTTGGGTACGGGCTGTCCCGGACTTGAGTGTCTGAGTGA

640
GGGTCTTCCCTCGAGGGTCTTTCATTTGCTACATGGCCCGGAATTCGAGAATCTTTCATTTGGTGCATTGGCCGGGAAT

693
TCGAAAAACAGCCGACCACCAGAGGTCTAGACCACTTAGAGGTAAGATTC

NVL-2

80
TGAAGAATGAAAAATTACTGGCCTCTTGTGAGAACATGAACTTTACCTCGGAGCCCACCCCTCCCATCTAGAAAAAT

160
TTTTGAGATAAAGGCCTCCTGGAACAACCTCAAAATGACATTGCCAAATGATAAGACATGAATCCTTAGTTACGTAGGTT

240
CCTTGATGGGACATGACTCCTTAGTTACGTAGGTTCCCTTGATAGGACATGACTCCTTAGTTACGTAGAATCCTTTGGCAG

320
AACCCCTTGTCCCTTGGCAGAACTCCCTAGTGATGAACTTGTACTTTCCCTGCCAGTTCTCCCCCTTTGAGTTTA

400
CTATATAAGCCTGTGAAAAATTTGGCTGGTCGTCGAGACTCCTCTACCCTGTGCAAAGGTGTATGAGTTTCGACCCAG

480
AGCTCTGTGTGCTTTCGTTTGTGCTTTATTTTCGACCCAGAGCTCTGGTCTGTGTGCTTTCATGTGCTGCTTTATTA

560
ATCTTACCTTCTACATTTTATGTATGGTCTCAGTGTCTTCTTGGGTACGGGCTGTCCCGGACTTGAGTGTCTGAGTGA

640
GGGTCTTCCCTCGAGGGTCTTTCATTTGCTACATGGCCCGGAATTCGAGAATCTTTCATTTGGTGCATTGGCCGGGAAT

720
TCGAAAAATCTTTCATTTGGTGCCTTGGCCGGGAAACAGCCGACCACCAGAGGTCTAGACCACTTAGAGGTAAGATTC

Figure 2. Nucleotide sequence of NVL-1 and NVL-2 LTRs. The features of these sequences are discussed in the text, the vertical lines show the 3' limits of the LTRs.

elements noted by their similarity to known sequence motifs (118). The two types of NVL LTR are similar with the same U3 R U5 pattern. These domains have not been determined experimentally for NVL-1 and NVL-2, but have been designated as such by their similarities to NVL-3. The U5 and R domains of the two types are identical and share the unusual characteristic feature of redundancy in the tRNA^{gly} primer binding site (PBS) and the adjacent inverted repeat (IR), the missing region in the NVL-1 LTR compared to NVL-2 corresponds to one of the extra copies of the IR and PBS. As previously noted the tRNA^{gly} PBS is unusual as this particular tRNA primer is not used by retroviruses (118). The two types of LTR also share an imperfect tandem repeat in the R region.

The U3 domains are of interest for two reasons. Firstly this domain harbours putative regulatory determinants, including promoter and enhancer like motifs. Secondly this domain is the only part of the two types of LTRs to show any sequence differences. This observation indicates that the NVL subset of VL30 elements can be further divided into NVL-3 type and NVL-1/2 type on the basis of the sequences in their U3 domains.

There are three major differences between the two types of U3 domain. First of all NVL-1/2 type LTRs have a 100bp insertion compared to NVL-3 occurring at nt144 of NVL-3 (see fig.3). This insert consists of three imperfect copies of a 35bp motif.

5' TGATA/GA/GGACATGACTCCTTAGTTACGTAGG/AT/ATCCT 3'

The sequence motif underlined in the consensus sequence shown above is similar to a motif known to bind the trans-acting factor AP-1 (21). This motif is conserved in all three copies of the 35bp repeat.

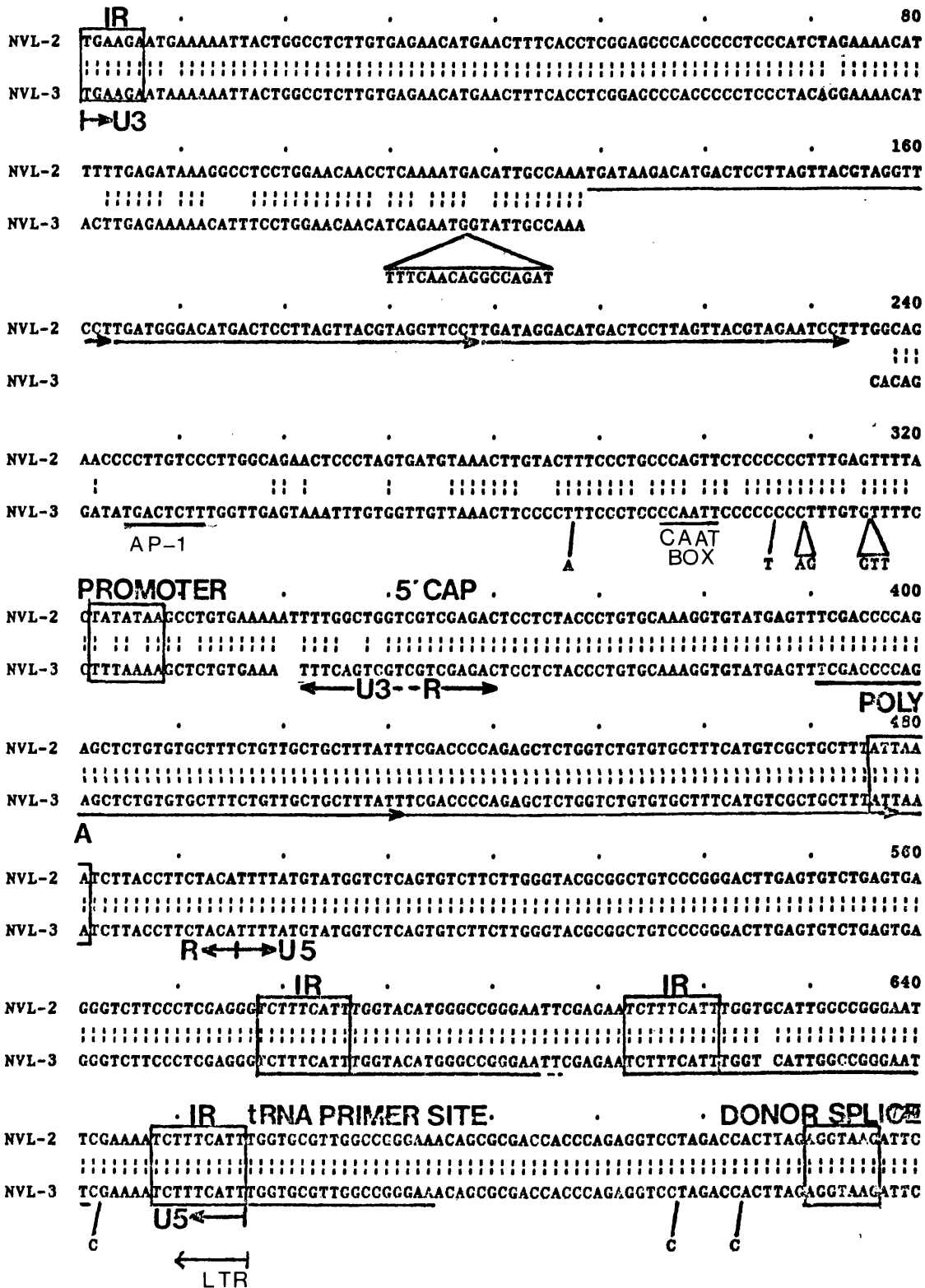


Figure 3. Comparison between the NVL-1/2 and NVL-3 LTRs. The various features on this figure were determined previously for NVL-3 (118) and are discussed in the text. IR = inverted repeat.

The NVL-3 LTR also possesses an insertion. This insert is 16bp long and occurs at nt117 relative to NVL-1/2 (see Figure 3).

The third difference between the two types of LTR occurs in the region between the insertions and the 5' cap. In this region there a number of small sequence differences between the two LTR types characterised by short regions of non-homology and short (1-3bp) insertions and deletions. This region in NVL-3 contains a putative CCAAT box and an AP-1-like motif not present in the NVL-1/2 type LTR. This region of the U3 domain also contains the putative TATA boxes, these differ slightly between the two types of LTR.

R.1.2 COMPARISON BETWEEN NVL VL30 LTRs AND OTHER VL30 LTRs

Sequence heterogeneity within the VL30 LTRs has been well documented (117,118,119). Sequence analysis of a number of cloned VL30 elements has shown the extent of this heterogeneity. Figure 4 shows a schematic comparison between the NVL-2 LTR and a number of other VL30 LTRs, mostly derived from genomic clones. The most striking similarities are between the U3 regions of NVL-2 and α 5 (125) and NVL-3 and VL3 (119). α 5 was isolated as a growth specific cDNA from SV40 transformed NIH 3T3 cells and VL3 is a genomic clone.

The other clones show a great deal of sequence heterogeneity in their U3 domains compared to the NVL U3 regions. However the U5 and R domains show a much greater degree of similarity between the different LTRs, with the major differences being due to deletions. For example some of the clones have a deletion in the region that has a repeat in the NVL type R domain. The exception to this is VL3 which has quite different U5 and R domains compared to the other clones studied. These comparisons clearly show however that the NVL VL30 elements form a distinct

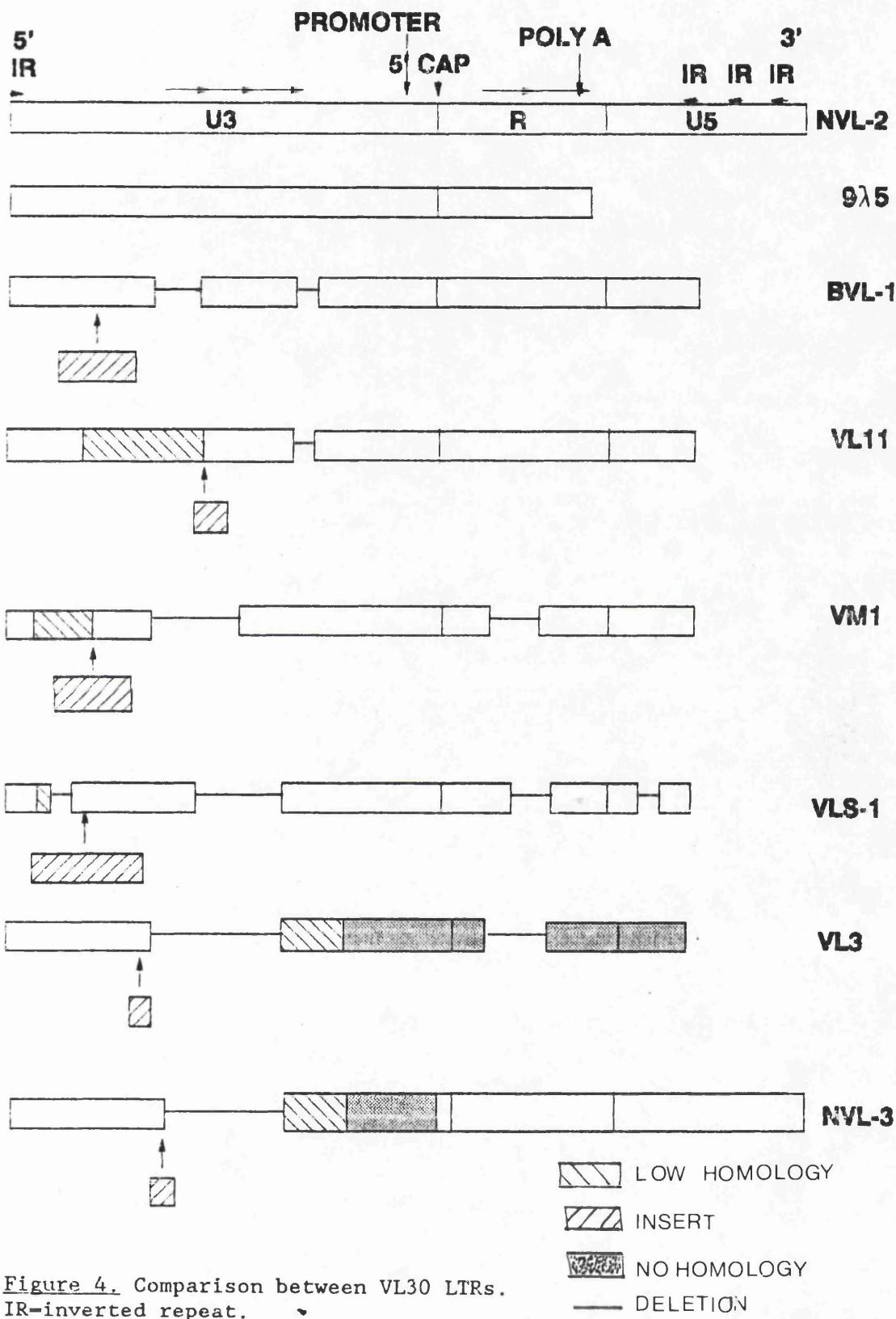


Figure 4. Comparison between VL30 LTRs. IR-inverted repeat.

structural subset as well as being functionally distinct from the other VL30 gene family members

R.1.3 HOMOLOGY WITH OTHER SEQUENCES

A search of the Genbank database indicates that the NVL LTRs share little or no homology with most other retrovirus-like elements. For the most part the only regions that showed homology with other sequences were the IRs, which were similar to retrovirus IRs.

R.2 VL30 ELEMENTS ARE RESPONSIVE TO STIMULATION BY SERUM

GROWTH FACTORS

Previous work has shown that the expression of VL30 elements is affected by whether the cells are actively growing or not (114). Also the inducibility of VL30 elements by polypeptide growth factors is well documented (127,128). In order to investigate the effect of changes in the growth state of the cell on VL30 expression further the effect of serum stimulation of quiescent NIH 3T3 cells has been studied. A number of aspects of the response are of interest. For example it is important to show whether any increase in expression is due an increase in the rate of transcription or whether post-transcriptional mechanisms are involved. Also given the diversity seen in the sequence of the U3 domain of the LTRs, it is of interest to see whether the two types of element (NVL-1/2 and NVL-3) show similar responses to serum stimulation. Another important question is whether any response is regulated by sequences within the VL30 element or whether sequences in flanking mouse cellular DNA are responsible. There is evidence that serum growth factors mediate their responses via the protein kinase C signal transduction pathway (45.47).

It would therefore be useful to know whether this pathway is involved in controlling any response to growth factor stimulation of VL30 elements.

R.2.1 SERUM STIMULATION OF VL30 ELEMENTS IN QUIESCENT NIH

3T3 CELLS

When NIH 3T3 cells are grown in serum free medium they become quiescent, that is they enter the G0 stage of the cell cycle. Figure 5 shows that VL30 expression in quiescent cells (lane b) is much lower than in normally growing cells (lane a). When fresh serum (15% FCS) is added to the cells VL30 expression is stimulated (lanes c-e). This response to serum peaks at about 5 hours post-stimulation. VL30 elements are quite different from the early response genes such as c-fos which are stimulated very rapidly but transiently when the cell is stimulated with serum (11). Instead VL30 expression is similar to that of genes such as proliferin (74) which are stimulated much later in the G1 phase of the cell cycle as the cells progress towards DNA synthesis.

R.2.2 THE RESPONSE TO SERUM IS REGULATED AT THE

TRANSCRIPTIONAL LEVEL

In order to find out whether the response to serum is regulated at the transcriptional level or whether post-transcriptional mechanisms are responsible, nuclear run on transcription analysis was done. Nuclei were isolated from NIH 3T3 cells that had been serum starved for 48 hours and from cells that had been starved and then stimulated with serum for 2 hours or 5 hours. No new transcription is initiated in isolated nuclei, however already initiated transcripts are completed *in vitro* giving a measure of the level of transcription at the time that the nuclei were

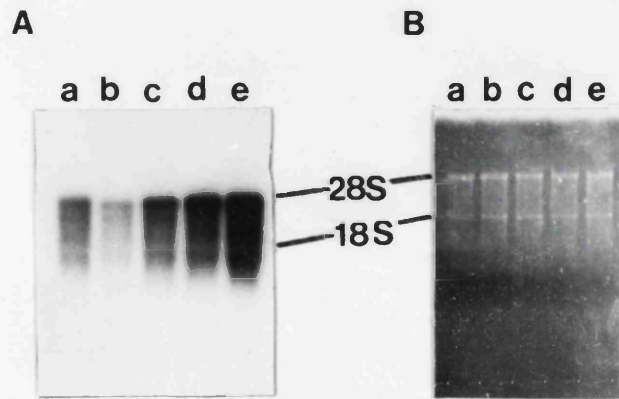


Figure 5. VL30 expression in quiescent and serum stimulated NIH 3T3 cells. A; Northern blot analysis of RNA from a) control NIH 3T3 cells, b) serum starved cells, c)-e) serum starved cells stimulated with 15% FCS for 1, 2 and 5 hours respectively. B; 1% agarose 0.8M formaldehyde gel used to prepare the blot (stained with ethidium bromide), a)-e) are as in A.

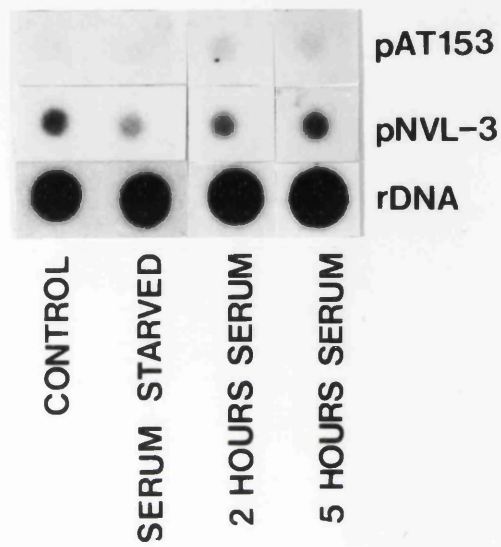


Figure 6. Nuclear run on transcription analysis of serum induced VL30 expression.

isolated. Clearly a difference in the levels of transcripts gives an indication of the extent of transcriptional control of expression of a given gene. Nuclear run on transcripts were prepared from these cells and the labelled transcripts were hybridized to filter bound DNA, Figure 6 shows the results of such an experiment. There is only a low level of hybridization to the VL30 DNA at time 0, but the signal increases to give an approximately 5 fold greater level of expression by 5 hours post-stimulation. This level of induction is somewhat lower than that seen in the Northern blot shown in Figure 5, however the level of induction does seem to vary from experiment to experiment.

The rDNA dots serve as a control for the input of labelled transcripts as this gene is not serum responsive. The pAT153 DNA dots serve as a control for levels of background hybridization. The result indicates that the response to serum is at least in part regulated at the level of transcription.

R.2.3 DIFFERENTIAL REGULATION OF NVL-3 AND NVL-1/2 IN RESPONSE TO SERUM

The data presented in this section was provided by Dr. J.D. Norton.

In order to analyze VL30 RNA transcribed from different NVL VL30 elements an S1 nuclease probe was derived from NVL-3 (see figure 7A and 145) that distinguishes transcripts from NVL-3 LTRs from transcripts initiated from NVL-1/2 LTRs by virtue of differences in their U3 domains. Figure 7B shows S1 nuclease analysis of RNA from NIH 3T3 cells digested with different inputs of enzyme. NVL-3 transcripts protect a fragment of about 500bp and NVL-1/2 transcripts protect two fragments of approximately 290nt and 320nt in length. These fragments map to positions of major mismatch between NVL-3 and NVL-1/2 at 118bp and 129bp in figure

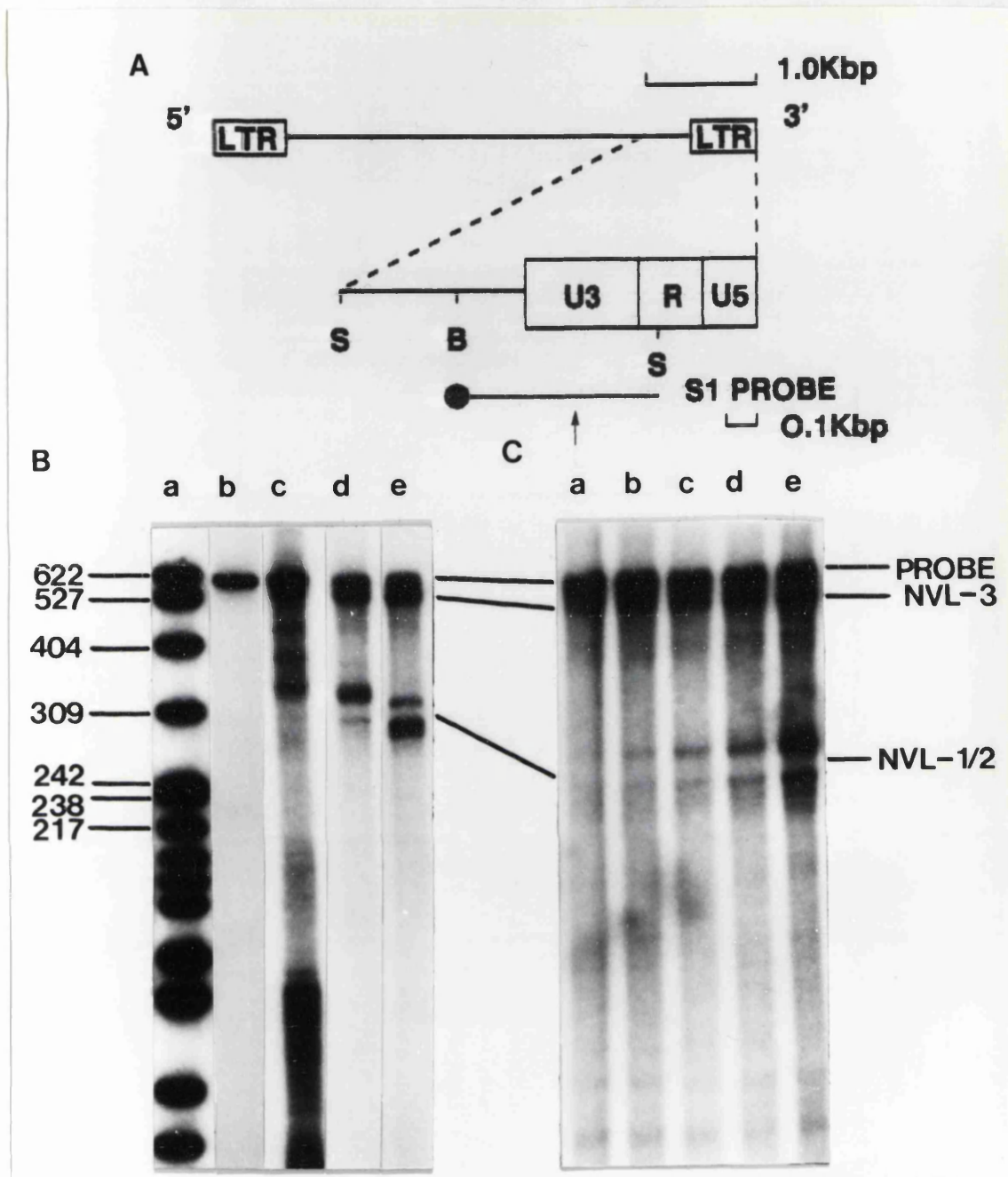


Figure 7. S1 nuclease analysis of RNA from NIH 3T3 cells. A; Structure of S1 nuclease mapping probe derived from NVL-3 (145). The solid circle denotes the labelled end of the probe and the arrow shows the approximate position of S1 cleavage distinguishing NVL-3 from NVL-1/2 transcripts. S=SacI, B=BglIII. B; 5 μ g of total cellular RNA was annealed with ³²P labelled probe and digested with c) 100 units, d) 250 units or e) 500 units of S1 nuclease. Protected fragments were run on a 5% polyacrylamide-8M urea sequencing gel. Lane a) MspI digest of pBR322 DNA and lane b) S1 probe alone (100cpm). C; 5 μ g of total cellular RNA was annealed with ³²P labelled probe and digested with 250 units of S1 nuclease, protected fragments were run on 5% polyacrylamide-8M urea sequencing gels. Lane a) serum starved cells, b)-e) cells stimulated with 15% FCS for 1, 2, 3 and 5 hours respectively.

3. The size heterogeneity can be attributed to variations in the extent of S1 cleavage at the points of mismatch, as an increase in input of enzyme resulted in more of the smaller fragment. There was also a small overall decrease in the size of the fragments which is probably due to terminal degradation by the enzyme.

As mentioned previously the sequence heterogeneity in the U3 domain of the NVL VL30 LTRs suggests that expression of these elements is regulated independently. In order to test this hypothesis RNA from serum starved and serum stimulated NIH 3T3 cells was analysed using the S1 nuclease mapping strategy described above. Figure 7C shows that the NVL-1/2 elements are responsible for almost all of the serum responsive expression. The NVL-3 elements are only slightly responsive and show quite high levels of expression even in quiescent cells.

R.2.4 THE REGULATION OF VL30 EXPRESSION IS MEDIATED VIA SEQUENCES WITHIN THE ELEMENT

The S1 nuclease mapping data in this section was provided by Dr. J.D. Norton.

The data presented so far does not show that the expression of NVL VL30 elements is controlled by sequences present within the elements, therefore there is a possibility that VL30 expression is regulated by sequences in flanking mouse cellular DNA. In order to investigate which of these two possibilities is true, the expression of VL30 elements stably integrated in heterologous cells was studied.

VLNRK cells are normal rat kidney cells into which VL30 elements have been introduced via retrovirus mediated gene transfer (124). Carter *et al* have previously shown that VLNRK cells have acquired VL30 proviruses at 1-2 copies per cell and that these elements are

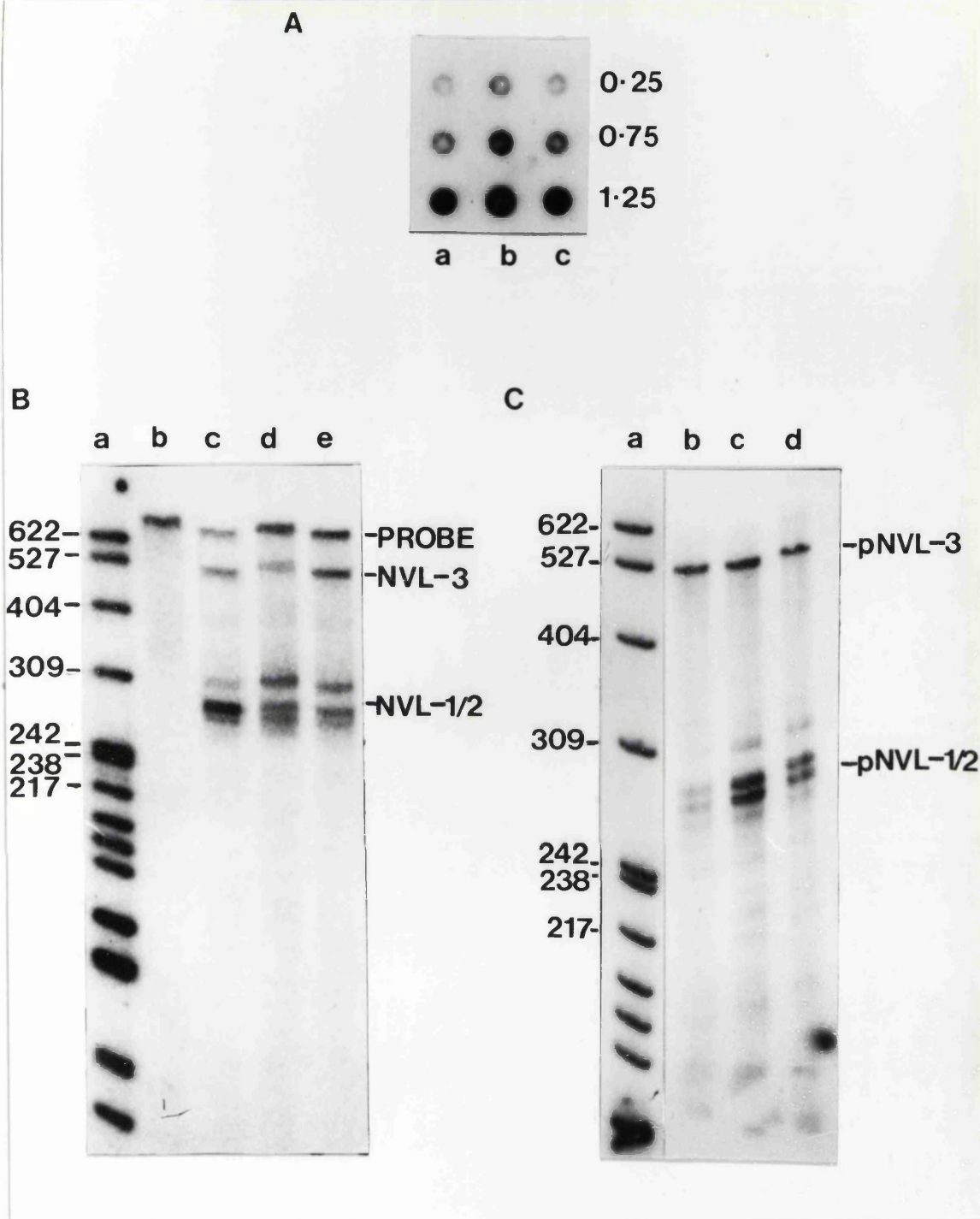


Figure 8. Analysis of RNA from quiescent and serum stimulated VLNRK cells. A; Dot blots were prepared with serial dilutions of RNA obtained from a) serum starved cells and b) and c) cells stimulated with 15% FCS for 2 hours and 5 hours respectively. B; S1 nuclease analysis was performed on c) 5 μ g of NIH 3T3 RNA, d) 2ng of MLV virion RNA and e) 4 μ g of VLNRK RNA. The RNA was annealed with 32 P labelled probe and digested with 250 units of S1 nuclease, protected fragments were run on 5% acrylamide-8M urea sequencing gels. Lane a) MspI fragments of pBR322 DNA, lane b) S1 probe alone (100 cpm). C; S1 nuclease analysis was performed as described on RNA from b) serum starved VLNRK cells and c) and d) cells stimulated with 15% FCS for 2 hours and 5 hours. Lane a) MspI fragments of pBR322 DNA.

transcriptionally active (124). Figure 8A shows a dot blot of RNA prepared from quiescent and serum stimulated VLNRK cells and probed with VL30. There appears to be a small response to the serum stimulation, however the kinetics are different from those seen in NIH 3T3 cells, with expression peaking at around 2 hours.

In order to investigate this response further, S1 nuclease mapping analysis was carried out on RNA obtained from VLNRK cells. Figure 8B shows the results of S1 nuclease analysis of RNA from the VLNRK cells (e) in parallel with RNA obtained from MLV virions propagated on NIH 3T3 cells (d) and RNA from NIH 3T3 cells (c). This figure shows that similar populations of NVL VL30 elements are found in all three RNA samples indicating that the NVL VL30 elements represent both the transcriptionally active and the retrovirus transmissible subsets of VL30 elements. Also this figure shows that the different NVL VL30 elements are packaged with equal efficiency. In order to look at the serum response in the VLNRK cells S1 nuclease analysis was carried out on RNA obtained from quiescent and serum stimulated cells. NVL-3 expression appears to be more or less constitutive, however the NVL-1/2 elements are quite responsive to serum stimulation. This data indicates that the serum response is controlled by sequences within the VL30 elements and not by flanking sequences in the mouse cellular DNA.

Similar experiments were conducted in a second rat cell line. The VLRat1 cell line was prepared by retrovirus mediated gene transfer of VL30 elements to Rat1 embryo fibroblasts. The VLRat1 cells were then serum starved and stimulated with 15% FCS for 5 hours. Figure 9 shows a northern blot analysis of RNA from the VLRat1 cells compared to the parental Rat1 RNA. Clearly no mouse VL30 is expressed in the parental cell line, whereas there is expression in the VLRat1 cells. However in this

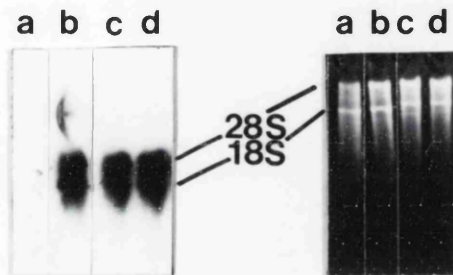


Figure 9. Analysis of RNA from Rat1 cell lines. Northern blot analysis of RNA from a) Rat1, b) VLRat1, c) quiescent VLRat1 and d) quiescent VLRat1 stimulated with 15% FCS for 5 hours. Alongside is the stained gel used to prepare the blot.

cell line VL30 elements appear not to be serum responsive. A further investigation of VL30 expression in this cell line using S1 nuclease analysis indicated that compared to both NIH 3T3 cells and VLNRK cells expression of VL30 in the VLRat1 cells was very low and no serum response was detected for either type of NVL element (data not shown).

R.2.4 THE SERUM RESPONSE IS NOT MEDIATED VIA THE PROTEIN

KINASE C SIGNAL TRANSDUCTION PATHWAY

The protein kinase C signal transduction pathway has been shown to be important in mediating some of the responses to growth factors (see introduction). It is therefore of interest to see what role, if any, protein kinase C plays in mediating the serum responsive expression of VL30 elements. The specific protein kinase C inhibitor, H-7 (132), was used to investigate the role of this pathway in the serum response.

As a control for this experiment serum starved cells were stimulated with the protein kinase C activator TPA in both the presence and absence of H-7. As figure 10A shows 140 μ M H-7 completely abolishes the response of VL30 elements to 100nM TPA. Lower inputs of H-7 gave only a partial inhibition of the response (data not shown). A further control was set up to check for cell viability in the presence of H-7. A number of identical plates of cells were incubated in a range of concentrations of H-7 from 0-175 μ M without a noticeable decrease in cell viability (data not shown).

Figure 10B shows the results of a serum response experiment in which three identical plates of cells were incubated in serum free medium for 48 hours, one of the plates was then stimulated with 15% FCS and one plate was treated with 15% FCS plus 140 μ M H-7. The serum response was not significantly inhibited by this concentration of H-7, implying that the

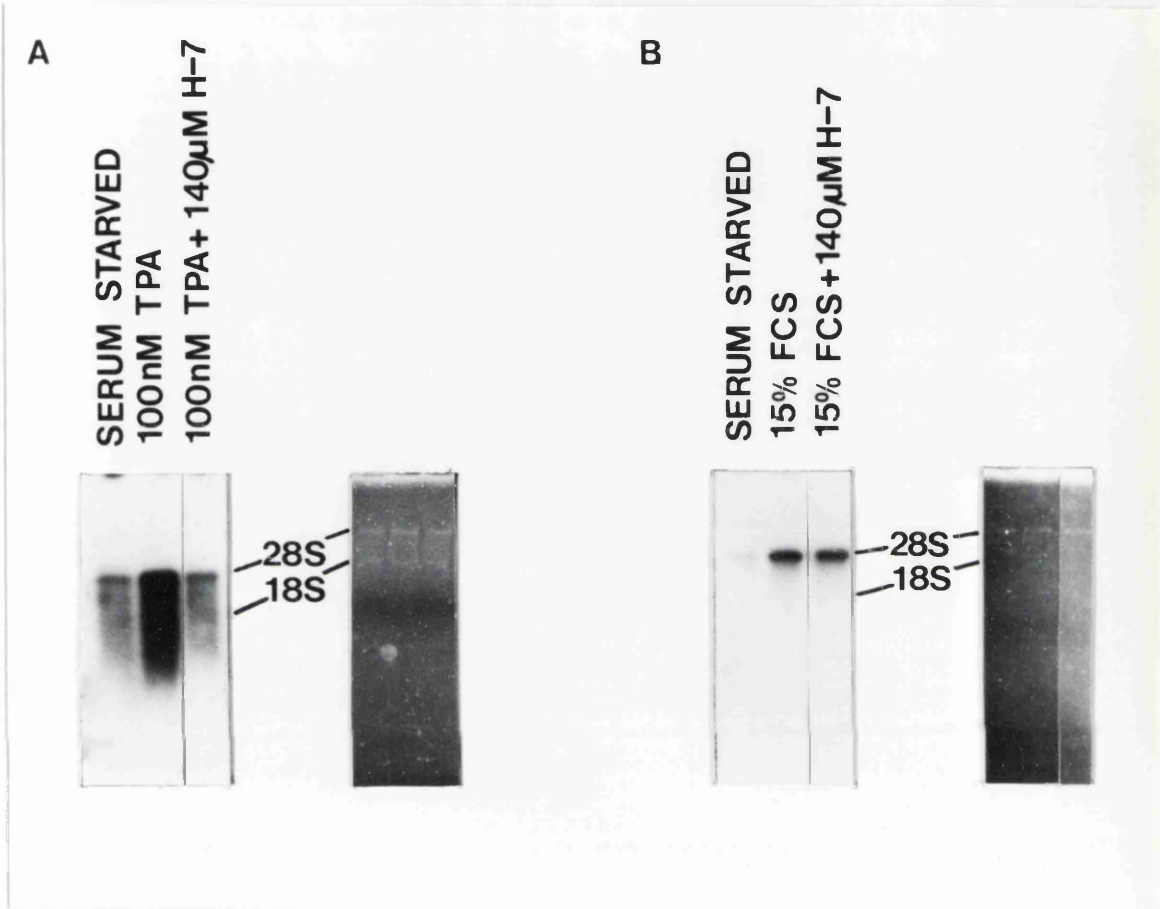


Figure 10. Effects of the Protein kinase C inhibitor H-7 on VL30 expression. Northern blot analysis of RNA from quiescent NIH 3T3 cells alongside RNA stimulated with 100nM TPA (A) or 15% FCS (B) for 6 hours either in the presence or absence of 140µM H-7. The ethidium bromide stained gels used to prepare the blots are shown alongside.

protein kinase C pathway is not an important pathway for mediating the response of VL30 elements to serum.

R.3 VL30 ELEMENTS ARE RESPONSIVE TO CELLULAR TRANSFORMATION

The effect of cellular transformation on the expression of VL30 elements by a variety of agents is quite well documented (114,126). In order to further investigate this response the levels of VL30 expression in several transformed cell lines was examined. Transformed cells are known to produce growth factors that act upon the cells in an autocrine/paracrine fashion (146), it was therefore of interest to investigate whether medium from transformed cell lines showing elevated levels of VL30 expression could effect VL30 expression in normal NIH 3T3 cells.

Only NVL-1/2 elements are responsive to serum stimulation (see figure 7C). It is therefore of interest to see whether both types of NVL VL30 element are responsive to cellular transformation.

R.3.1 LEVELS OF VL30 TRANSCRIPTS IN CELL LINES TRANSFORMED VIA

DIFFERENT MECHANISMS

The A1 cell line is an NIH 3T3 cell line that has become transformed spontaneously. Figure 11A shows a dot blot of RNA prepared from normal NIH 3T3 cells alongside RNA from A1 cells. The A1 cells have about a 20 fold higher level of expression of VL30 than the normal NIH 3T3 cells.

The cell lines CC1 and CA12 are KiMSV transformed non-producer cells (131). Both of these cell lines have a KiMSV provirus integrated at random in the genome. They have a transformed morphology and produce high levels of p21^{Ki-v-ras}(147). Two revertant cells lines, R.5.5.1 and

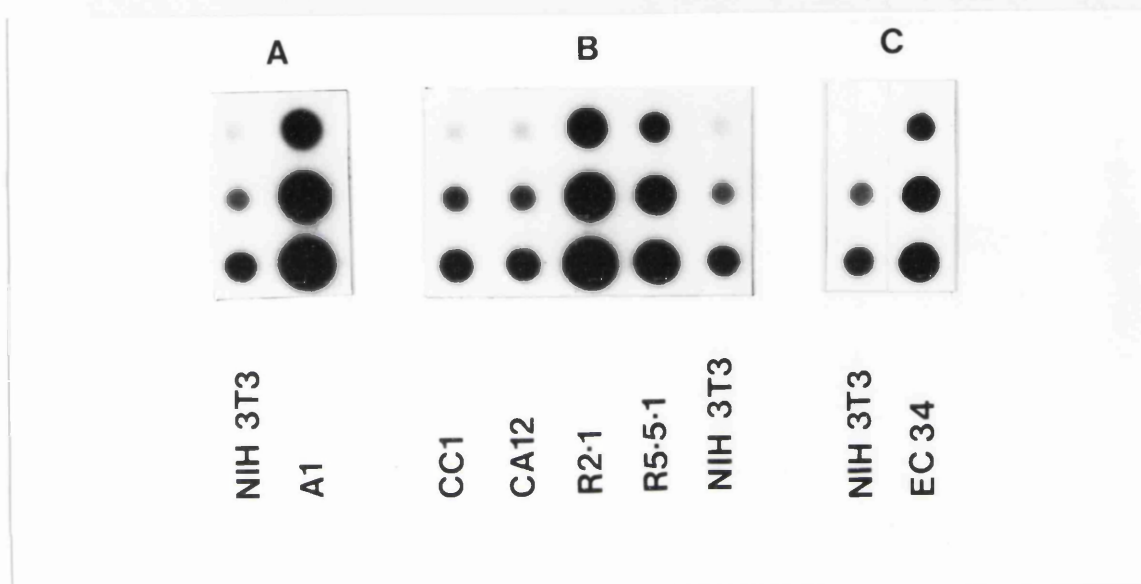


Figure 11. Dot blot analysis of RNA from normal and transformed NIH 3T3 cell lines. Blots were prepared with serial dilutions of RNA obtained from A; NIH 3T3 cells and spontaneously transformed cells (A1), B; NIH 3T3 cells and Ki-v-ras transformed non-producers (CC1 and CA12) and non-transformed revertants (R2.1 and R5.5.1) and C; NIH 3T3 cells and Ha-c-ras transformed cells (EC34).

R.2.1 were derived from the transformed cell line CCl by the BrdUrd killing method (131). These cell lines are morphologically similar to untransformed NIH 3T3 cells, however they retain fully functional integrated KiMSV proviruses (131) and express elevated levels of Ki-v-ras (147). Figure 11B shows dot blots of RNA obtained from these cell lines compared to RNA from normal NIH 3T3 cells. Both transformed cell lines express similar levels of VL30 to the normal NIH 3T3 cells. This implies that cellular transformation per se is not responsible for the elevated levels of VL30 seen in transformed cells. Interestingly the revertant cell lines both show levels of VL30 transcripts that are considerably higher than those in normal NIH 3T3 cells.

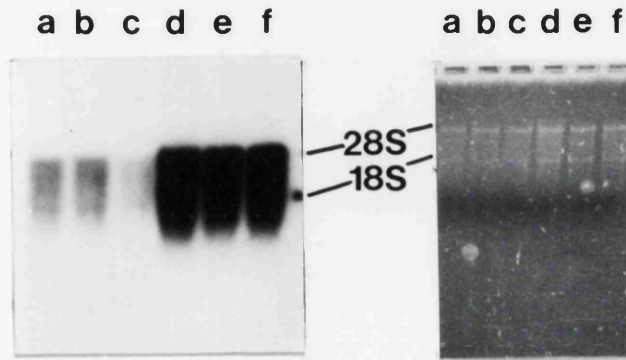
In order to investigate whether cells harbouring other activated ras oncogenes showed altered VL30 expression, expression in the cell line EC34 was examined. EC34 cells are mouse NIH 3T3 cells transformed by the activated Ha-c-ras oncogene found in human bladder carcinoma (EJ-ras) (42). As figure 11C shows, expression of VL30 elements in this cell line is higher than in normal NIH 3T3 cells.

R.3.2 AN AUTOCRINE/PARACRINE FACTOR DOES NOT MEDIATE

UPREGULATION OF VL30 EXPRESSION IN TRANSFORMED CELLS

Many transformed cell lines have been shown to produce growth factors that can act in an autocrine/paracrine fashion (146). The presence of such transforming growth factors abrogates the need for exogenous growth factors or serum. In order to see if an autocrine/paracrine mechanism is involved in stimulating VL30 expression in certain transformed cell lines, conditioned medium was taken from the transformed cells and placed onto normal NIH 3T3 cells. Figure 12A shows the results of an experiment where conditioned medium from the c-Ha-ras

A



B

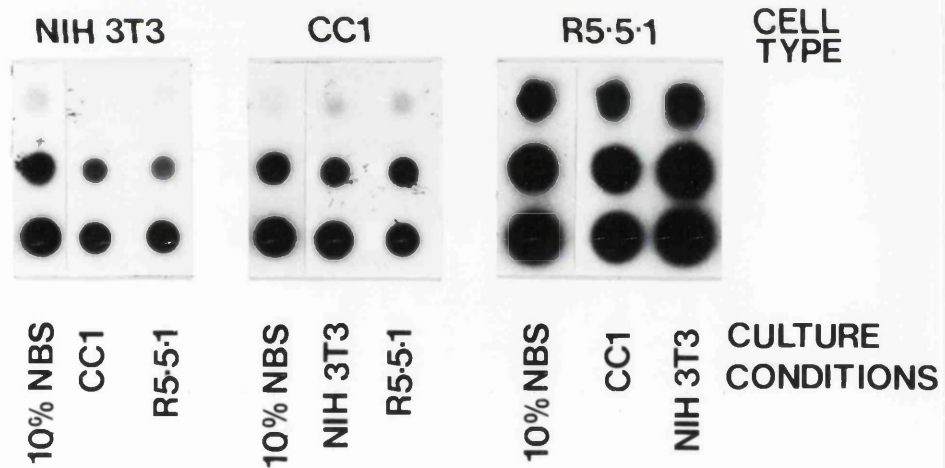


Figure 12. The effect of conditioned medium from transformed cells on the expression of VL30 elements in normal NIH 3T3 cells. A; Northern blot analysis of RNA from a) NIH 3T3 cells, b) and c) NIH 3T3 cells cultured in EC34 conditioned medium for 2 hours and 6 hours respectively, d) EC34 cells and e) and f) EC34 cells grown in NIH 3T3 conditioned medium for 2 hours and 6 hours respectively. B; Dot blot analysis of RNA from NIH 3T3 cells, CC1 cells and R5.5.1 cells cultured in 10% NBS and in conditioned medium for 5 hours.

transformed EC34 cell line, which shows elevated VL30 expression, was placed onto normal NIH 3T3 cells. Over a period of several hours the conditioned medium had no effect on VL30 expression in the normal cells.

Figure 12B shows the results of similar experiments using the Ki-v-ras transformed CC1 cells and the revertant R.5.5.1 cells. Conditioned medium from these lines did not effect VL30 expression in normal NIH 3T3 cells.

R.3.3 BOTH TYPES OF NVL VL30 ELEMENT ARE RESPONSIVE TO TRANSFORMATION

This data was provided by Dr. J.D. Norton.

S1 nuclease mapping analysis has shown that the NVL VL30 elements are regulated independently in response to serum stimulation of quiescent NIH 3T3 cells (see figure 7C). In order to see whether the response to cellular transformation is also regulated independently, S1 nuclease mapping analysis was carried out on RNA prepared from the spontaneously transformed cell line, A1.

Figure 13 shows that both types of element have elevated levels in A1 cells compared to those seen in normal NIH 3T3 cells. Lanes d) and e) show the levels of expression in serum starved and stimulated A1 cells compared to expression in steady state NIH 3T3 cells (lane c). When it is taken into account that 1/5 of the RNA input was used from the transformed cells it is clear that expression in this cell line is greatly elevated. This analysis also shows that neither type of element is serum responsive in the transformed cells.

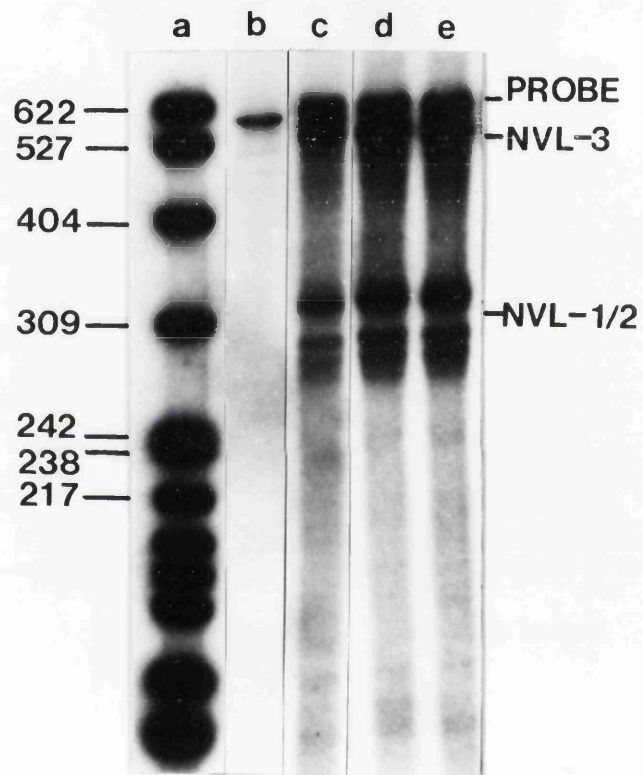


Figure 13. S1 nuclease analysis of RNA from spontaneously transformed NIH 3T3 cells. RNA samples were annealed to ^{32}P labelled probe and digested with 250 units of S1 nuclease, protected fragments were run on a 5% polyacrylamide-8M urea sequencing gel. Lane a) MspI digest of pBR322 DNA, b) undigested probe (100cpm), c) $5\mu\text{g}$ of RNA from NIH 3T3 cells, d) and e) $1\mu\text{g}$ of RNA from spontaneously transformed NIH 3T3 cells that have been serum starved (d) and stimulated with 15% FCS for 5 hours (e).

R.4 VL30 LTRs POSSESS PROMOTER AND ENHANCER ACTIVITY

The sequence data obtained for the NVL VL30 elements suggests that the NVL VL30 elements harbour transcriptional regulatory determinants within their U3 domains (see fig. 3 and 118). In order to see whether these regions contain promoter and enhancer activity constructs were prepared linking the U3 domains to the bacterial gene chloramphenicol acetyl transferase (CAT) (142) and the constructs were used in both transient and stable expression assays.

R.4.1 PREPARATION OF PROMOTER TESTING CONSTRUCTS

The NVL-1/2 and NVL-3 LTRs possess sequences in their U3 regions that are similar to transcriptional regulatory regions seen in other genes. In order to check the promoter activity of the NVL U3 regions they were linked to the bacterial gene chloramphenicol acetyl transferase (CAT). The promoter testing constructs were prepared as shown in figure 14. Both the NVL-1/2 and NVL-3 constructs were prepared in the same way. The 514bp BglIII/EcoRI fragment from pNVL-3LTR or the 600bp BglIII/EcoRI fragment from pNVL-1LTR (both provided by Dr.J.D. Norton) were cloned into the BamHI/EcoRI sites of the bluescript vector, pKS. These fragments contain a short piece of the unique region, the entire 3' U3 domain and part of the R domain upto the SacI site (see figure 3). Then the 1926bp StuI/ApaI fragment of pSV2CAT (144) containing the CAT gene coding sequence and the SV40 poly A and small t-antigen intron was inserted into the HindII/ApaI site downstream of the NVL U3 region.

A second set of constructs was prepared using U3 regions derived from the 5' LTRs of the NVL-1/2 and NVL-3 elements. These constructs were prepared by first removing the SacI fragment containing all the VL30 sequences from pNVL-3CAT (see figure 14). Then the 318bp BamHI/SacI

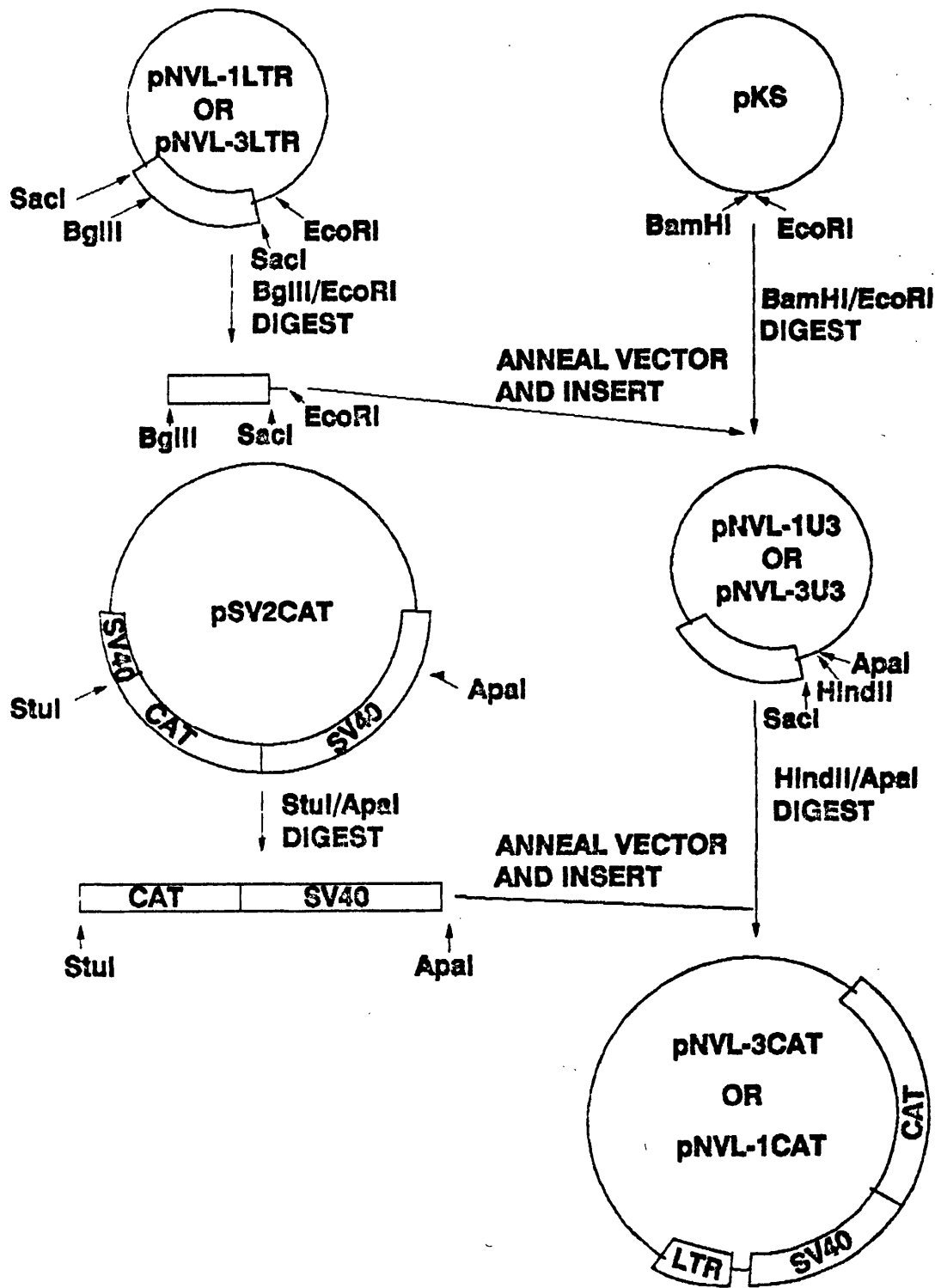


Figure 14. Preparation of the 3' LTR-CAT constructs.

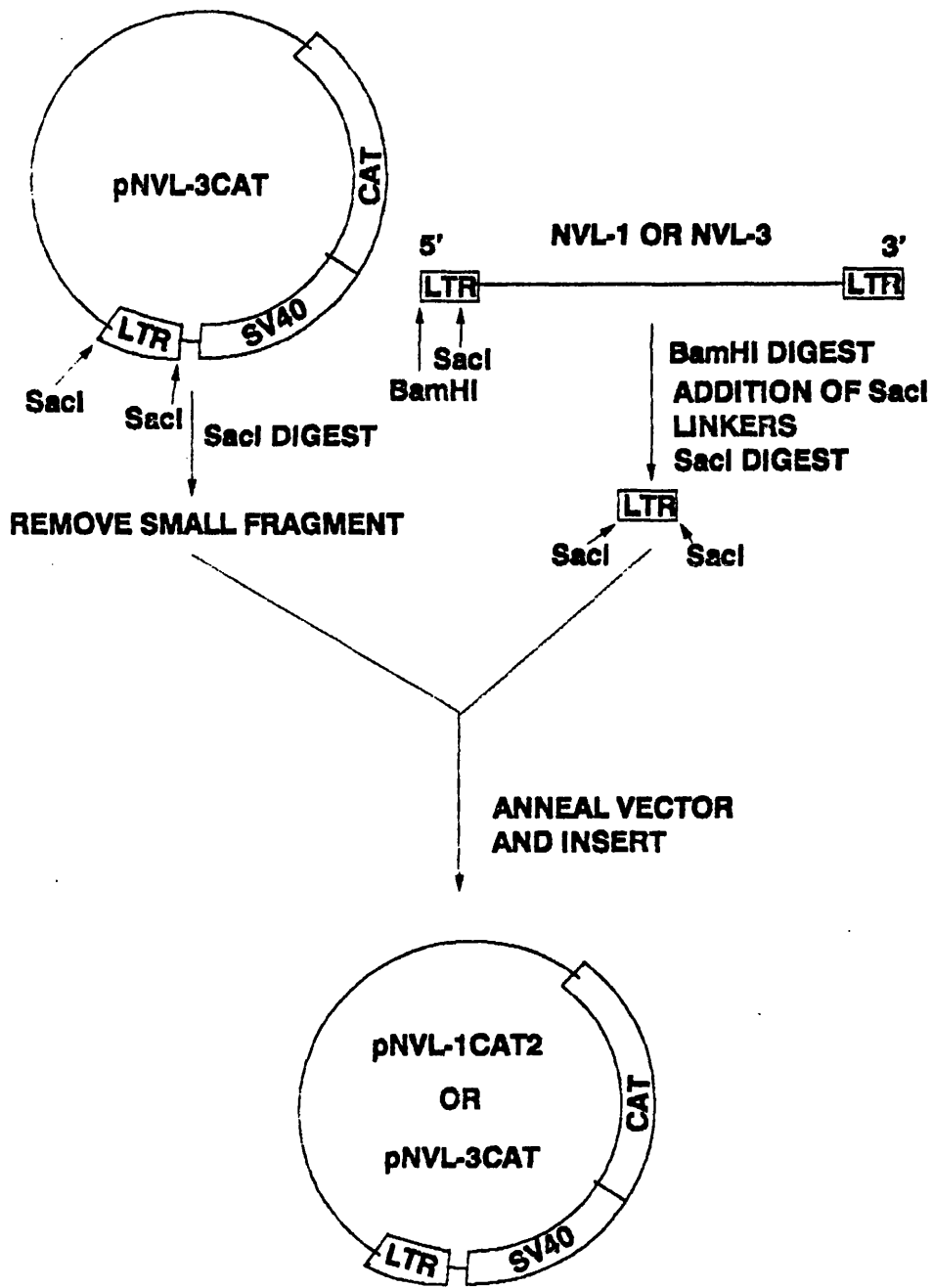


Figure 15. Preparation of 5' LTR-CAT constructs.

fragment of NVL-3 or the 404bp BamHI/SacI fragment of NVL-1/2 was inserted into this site after modifying the BamHI end to SacI with linkers. Figure 15 shows the preparation of these constructs schematically.

R.4.2 TRANSFECTION AND CAT ASSAY CONTROL EXPERIMENTS

Before embarking on transfections and CAT assays it was important to determine both the linear range of CAT activity and the optimum range of DNA input for transfections.

In order to estimate the amount of time to convert all the chloramphenicol to its acetylated form in a CAT assay four control reactions were set up in which 1 μ l of a 1/100 dilution of stock CAT enzyme was used in a standard reaction mix and incubated for 0, 15, 30, 45 or 60 minutes. This experiment indicated that the reaction was complete after about 30 minutes. This was therefore used as the standard assay time. However in experiments using constructs with very low promoter/enhancer activities longer assays were used.

The linear range of CAT activity was determined by using serial dilutions (0.5-0.001 units) of the stock CAT enzyme in a standard 30 minute CAT assay. Figure 16 shows the results of this assay. Similar control experiments were also done for longer CAT assays (data not shown).

In order to determine the optimum input of DNA to use in transfections six identical plates of NIH 3T3 cells were transfected with 0, 1, 3, 6, 10 or 15 μ g of pSV2CAT DNA. The transfections and CAT assays were done using the standard protocols described. Figure 17 shows that there is a decrease in transfection efficiency if greater than 6 μ g of DNA

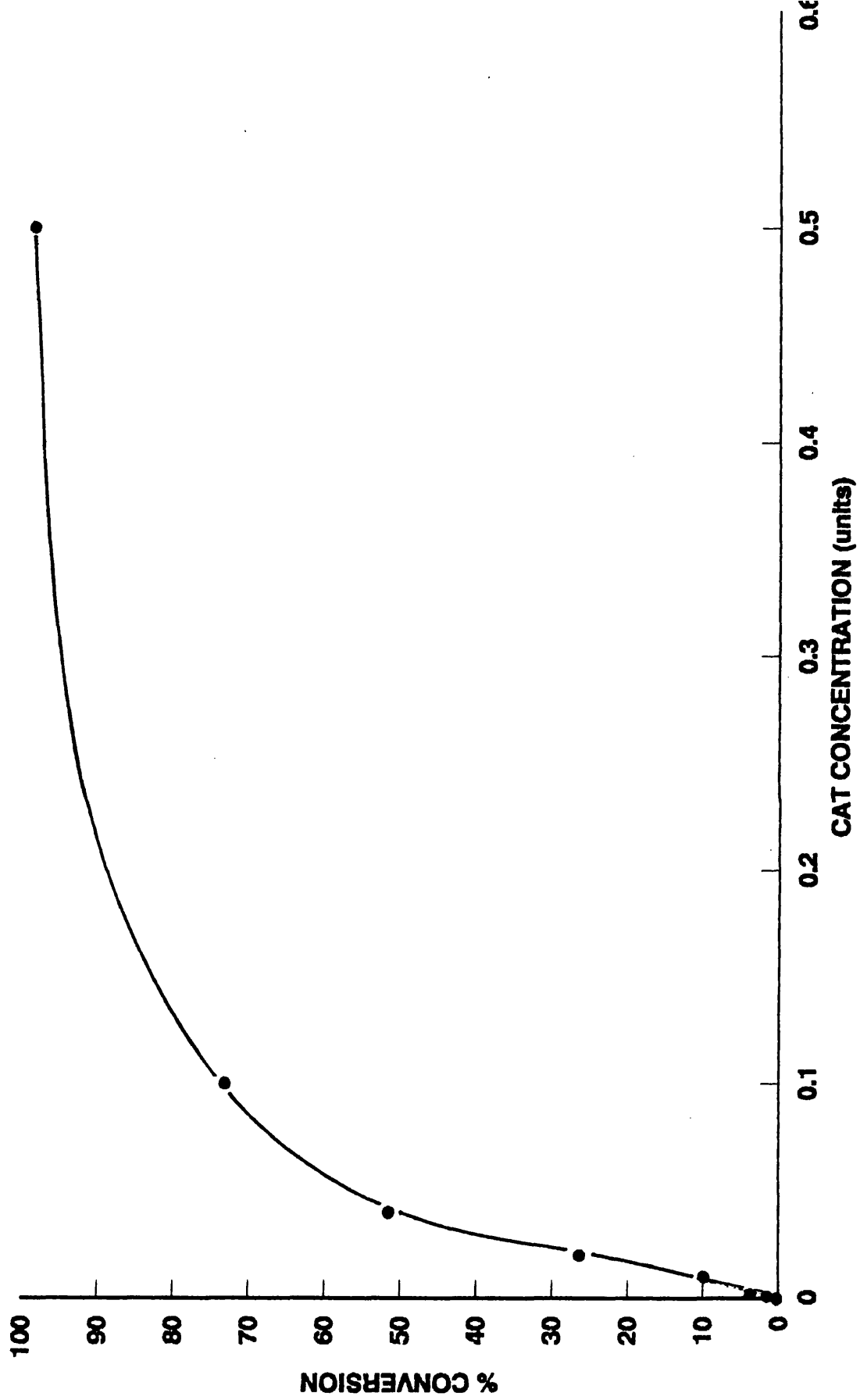


Figure 16. Linear range of CAT assay.

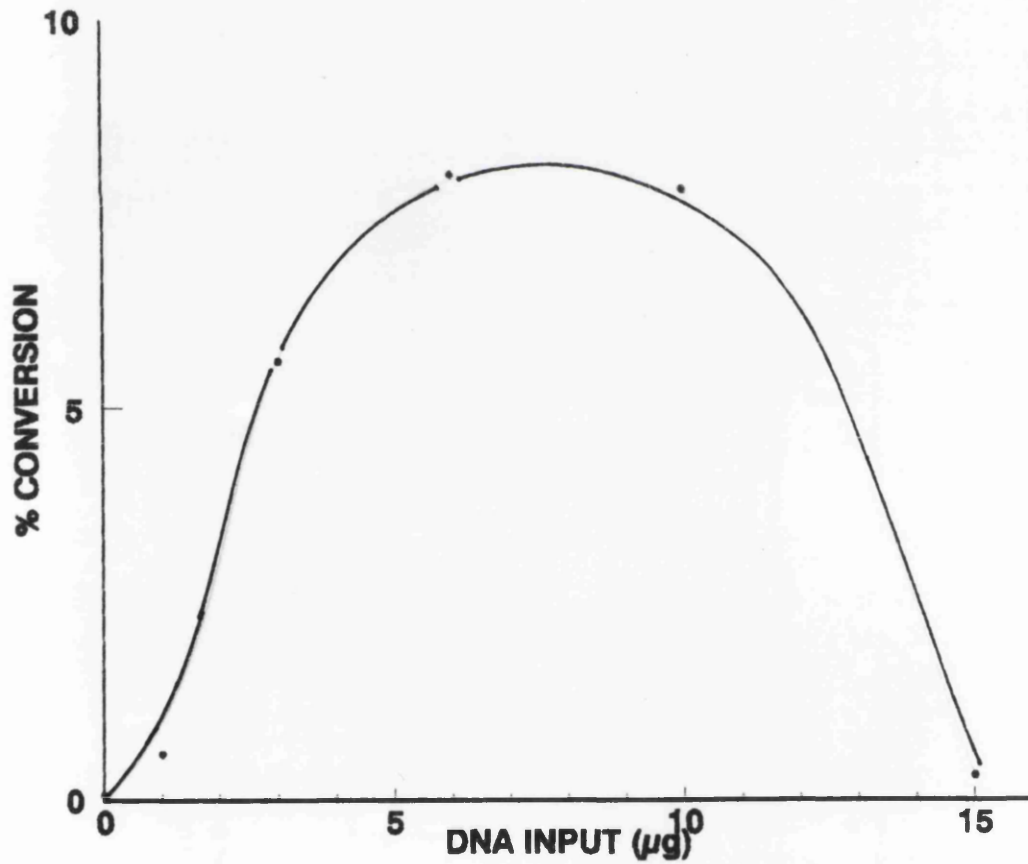
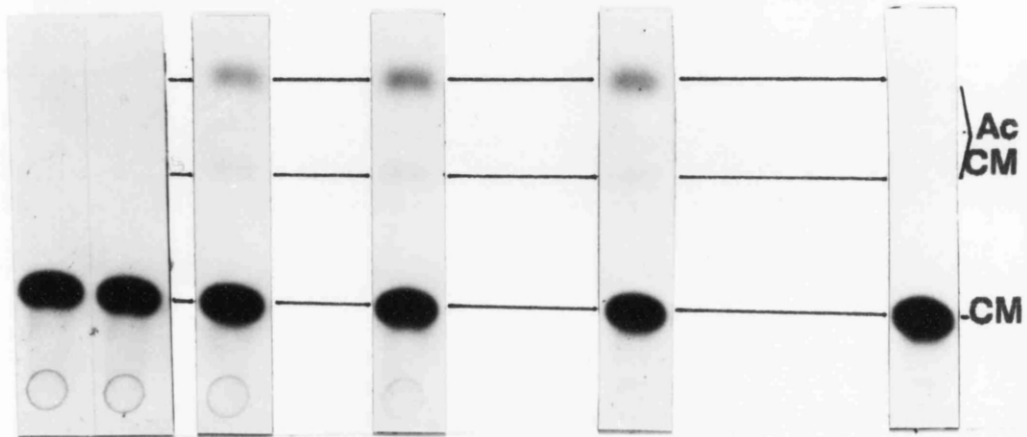


Figure 17. Optimum input of DNA in transfections.

was used. Therefore as a general rule all transfections were done using 3 µg of DNA.

R.4.3 DISTINCT VL30 LTRs HAVE DIFFERENT PROMOTER ACTIVITIES

In order to test the promoter activity of the NVL LTR constructs they were transfected into cultured cells and the cells analysed for transient expression of the CAT gene. Figure 18 shows the results of transfection into mouse NIH 3T3 cells and two heterologous cell lines, Human HeLa cells and rat1 cells. Table 1 shows quantitation of these results.

In NIH 3T3 cells the NVL-3 promoter is very strong, about 115 times stronger than the pSV2CAT control (see Table 1). However NVL-1 has low promoter activity compared to both NVL-3 and pSV2CAT. The NVL-3 promoter is about 230 times stronger than the NVL-1 promoter. The 5' and 3' LTRs show similar levels of promoter activity (data not shown), therefore the results for both LTRs have been pooled. The results for NVL-1 and NVL-2 driven CAT activity show similar levels (data not shown). Therefore the NVL-1 construct was used as a representative of the NVL-1/2 class of NVL VL30 elements.

In order to investigate whether the VL30 promoters are functional in heterologous cell types the NVL-CAT constructs were transfected into HeLa cells and into rat1 embryo fibroblasts and transient expression of CAT was assayed. In HeLa cells the VL30 promoters showed similar patterns of promoter activity to those seen in NIH 3T3 cells (see figure 18 and table 1). The NVL-1 promoter activity was much lower than that of NVL-3. However in this cell line the relative level of expression from the SV40 early promoter is somewhat higher. Interestingly, in the rat1 cells there was only a very low level of expression from all three promoters tested.

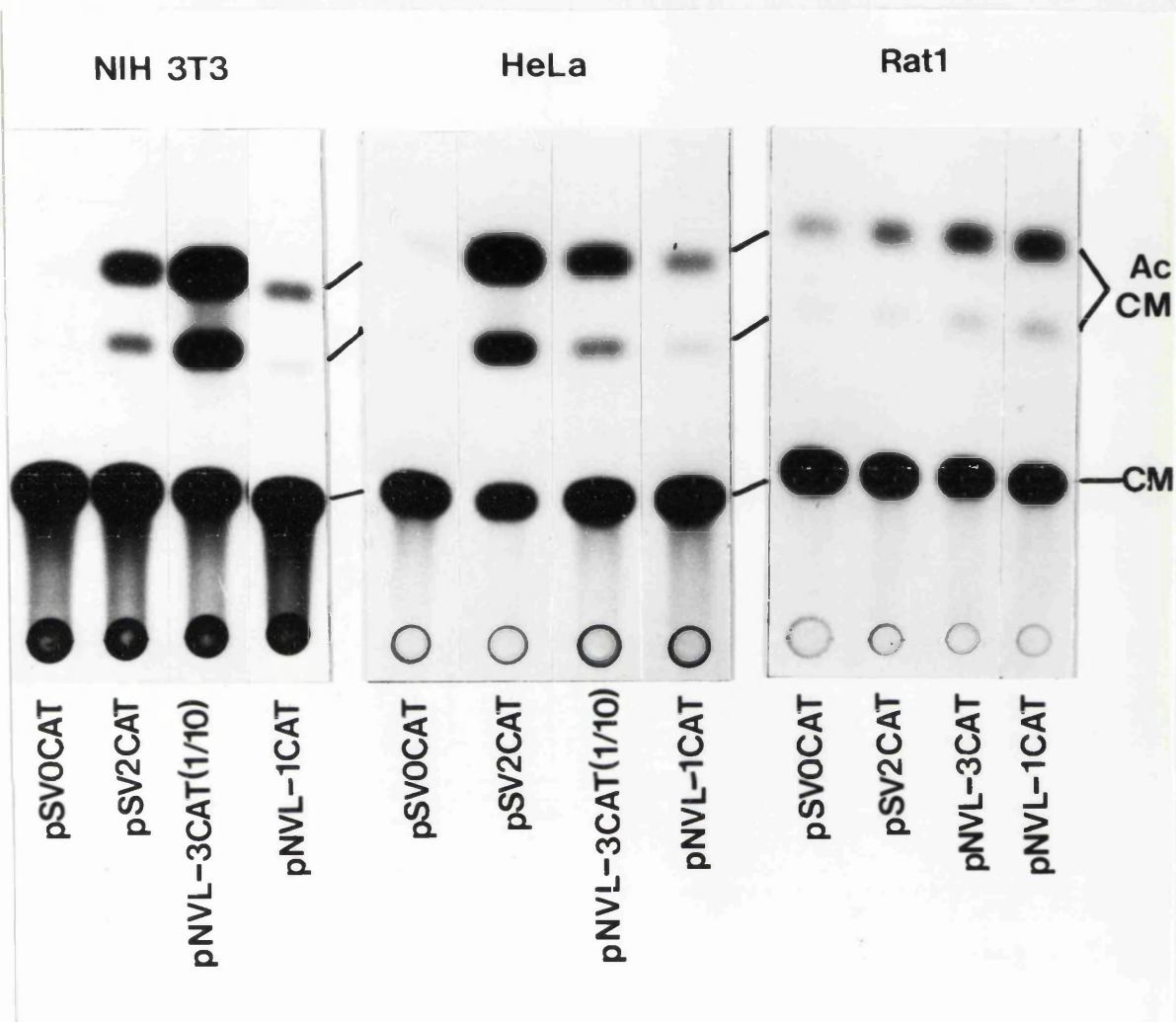


Figure 18. Promoter strengths of the NVL-1 and NVL-3 LTRs. $3\mu\text{g}$ samples of the LTR-CAT constructs were transfected in parallel with a promoterless control (pSV0CAT) and the CAT gene under the control of the SV40 early promoter and enhancer (pSV2CAT). Cells were harvested and assayed for CAT activity.

pNVL-3CAT(1/10) indicates that the cell extracts were diluted 1/10 before use.

As figure 18 shows none of the promoters initiated CAT transcripts at levels much above the background levels obtained for pSVOCAT, a promoterless control (142).

R.4.4 PREPARATION OF ENHANCER TESTING PLASMIDS

Enhancers are regulatory determinants that can function at a distance of several Kbp from the promoter, also they are able to function in an orientation-independent fashion and either upstream or downstream from the gene. In order to investigate whether the VL30 U3 regions possess enhancer function a further set of constructs was prepared. Figure 19 shows a schematic diagram of the preparation of these constructs. pBLCAT2 is a construct containing the CAT gene under the control of the TK promoter (143), the 318bp SacI fragment from pNVL-3CAT or the 595bp SacI fragment from pNVL-1CAT was inserted into the SacI site downstream of the CAT gene in pBLCAT2 and in either orientation with respect to the CAT gene. The orientation of the inserts was determined by restriction enzyme analysis.

R.4.5 BOTH VL30 LTRs EXHIBIT ENHANCER FUNCTION

In order to test the enhancer activity of the U3 domains the constructs described above were used in transient expression assays. As figure 20 shows both U3 domains possess enhancer activity. The enhancer activity of the NVL-3 LTR was much greater than that of the NVL-1 LTR. Table 1 shows quantitation of these results. The values obtained were normalized to those obtained for the control pBLCAT2 plasmid. It is interesting to note that the anti-sense orientation in both cases appears to give a higher level of enhancer activity. The NVL-3 enhancer testing

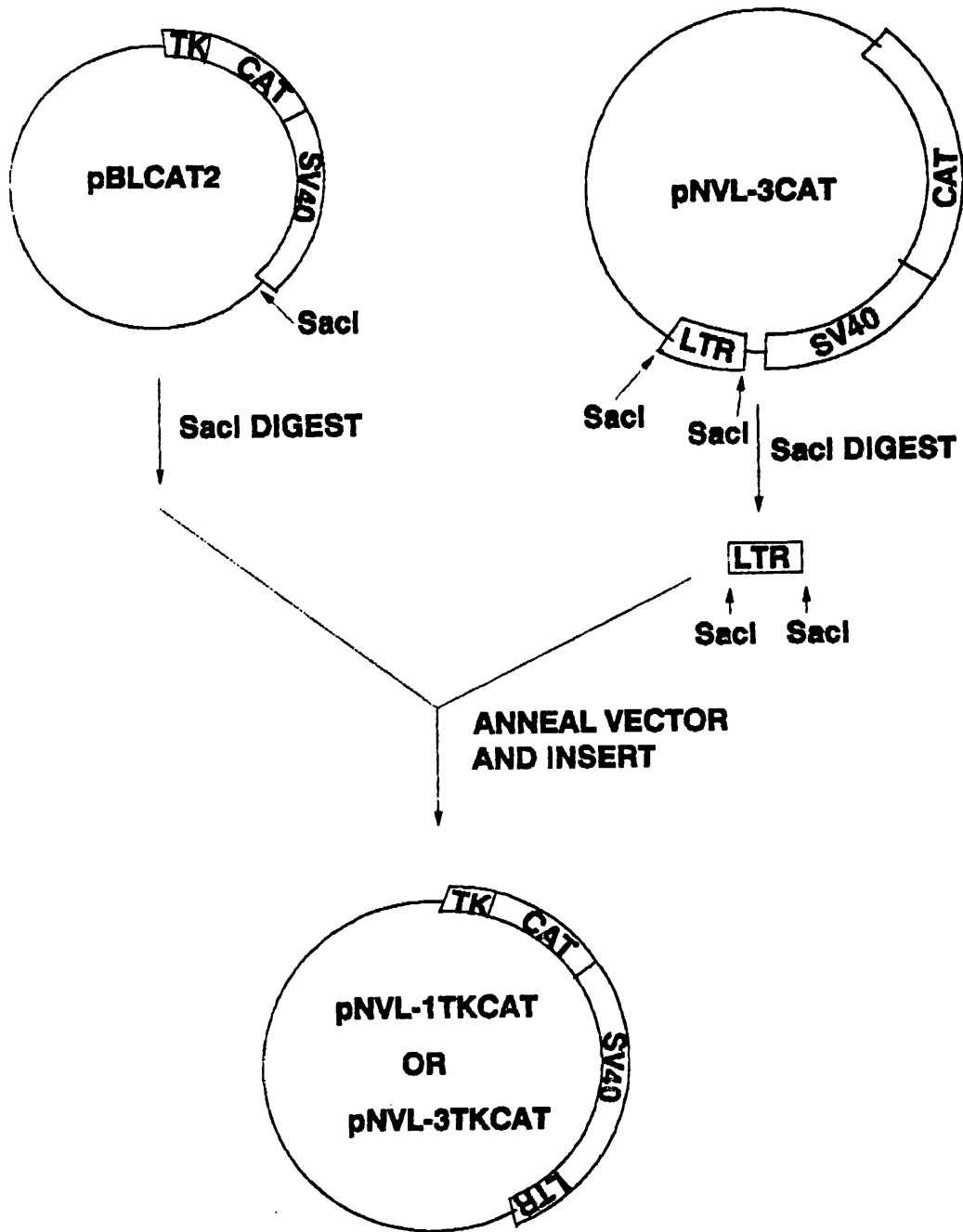


Figure 19. Preparation of enhancer testing constructs.

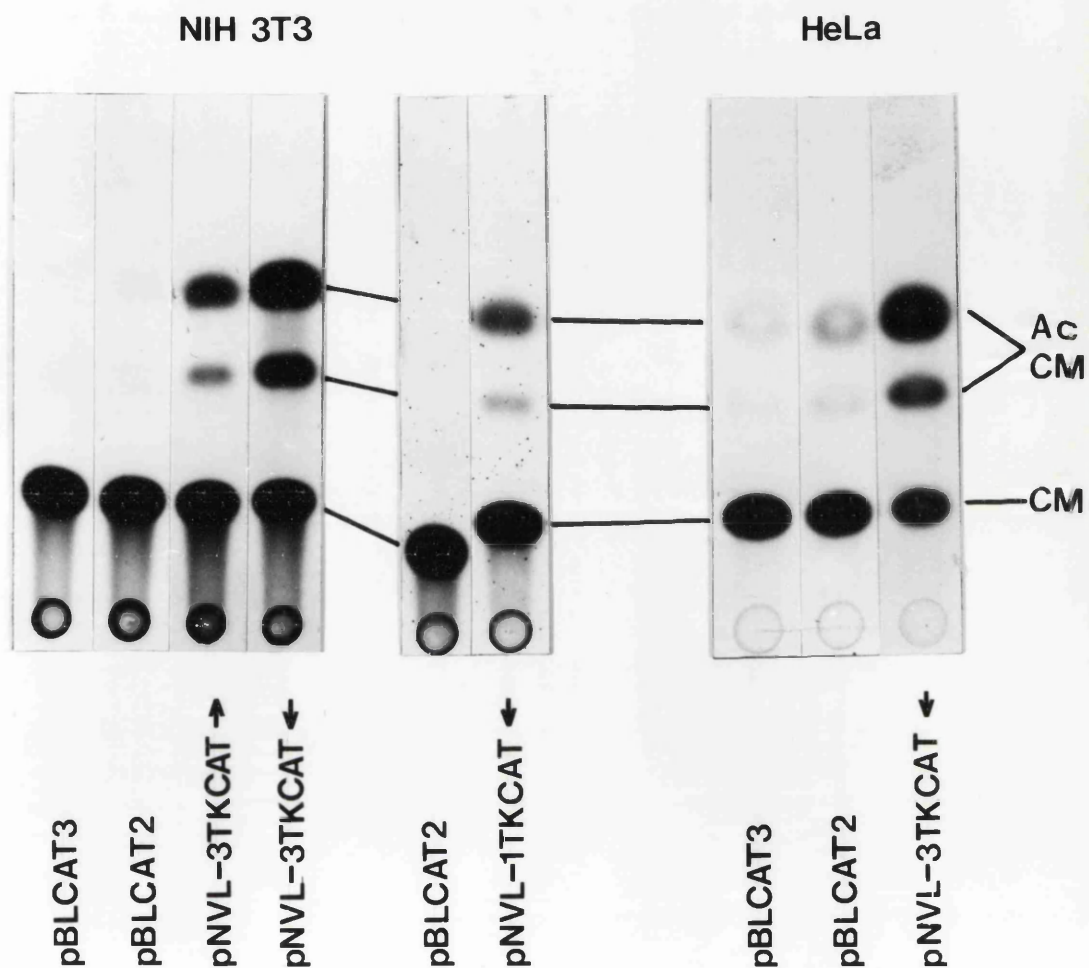


Figure 20. Enhancer strengths of the NVL-3 and NVL-1 LTRs. 3 μ g samples of the NVLTKCAT constructs were transfected in parallel with a promoterless control plasmid (pBLCAT3) and a control where the CAT gene is driven by the TK promoter (pBLCAT2). The arrows indicate that the LTRs are in the sense (→) and antisense (←) directions. Cells were harvested and assayed for CAT activity.

Cell Type	pSV2CAT	pNVL-3CAT	pNVL-1CAT
NIH 3T3	1	115	0.5
HeLa	1	0.9	0.05

Cell Type	pBLCAT2	pNVL-3TKCAT	pNVL-1TKCAT
NIH 3T3	1	518	13.5

Table R.1 Promoter and enhancer strengths of NVL-1 and NVL-3.

construct was also analyzed in HeLa cells, a similar pattern was seen in this cell line.

R.5 VL30 LTRs ARE TRANSACTIVATED BY ACTIVATED N-RAS

VL30 elements are responsive to cellular transformation including transformation mediated by the activated Ha-c-ras oncogene (see section R.3.1 and fig.11). In order to investigate the role of the VL30 U3 domains in regulating the response, co-transfections were carried out using either the promoter or enhancer testing constructs and an activated N-ras expression vector, pNRSac (148). This construct has the N-ras oncogene under the control of the SV40 promoter and enhancer.

R.5.1 TITRATION OF THE TRANSACTIVATION BY ACTIVATED N-RAS

The effect of the activated N-ras oncogene on promoter activity was investigated by co-transfecting either pNVL-3CAT or pNVL-1CAT with increasing inputs of pNRSac. Transient expression of the LTR driven CAT was then analysed by CAT assay. The results in figure 21A indicate that both NVL-3 and NVL-1 are responsive to N-ras. The greatest level of transactivation is achieved with 1 μ g of pNRSac to 2 μ g of LTR-CAT. This is an approximately molar ratio of the two plasmids. Table 2 shows quantitated results. The values obtained for the co-transfections are normalized to results obtained for the LTR-CAT constructs alone. The results indicate that the two types of LTR show similar levels of response to the activated N-ras oncogene.

The pNVL-3CAT and pNVL-1CAT constructs were also used in co-transfection experiments in HeLa cells. As figure 21B shows the LTRs are transactivated by N-ras in this cell line to a similar level to that seen in NIH 3T3 cells.

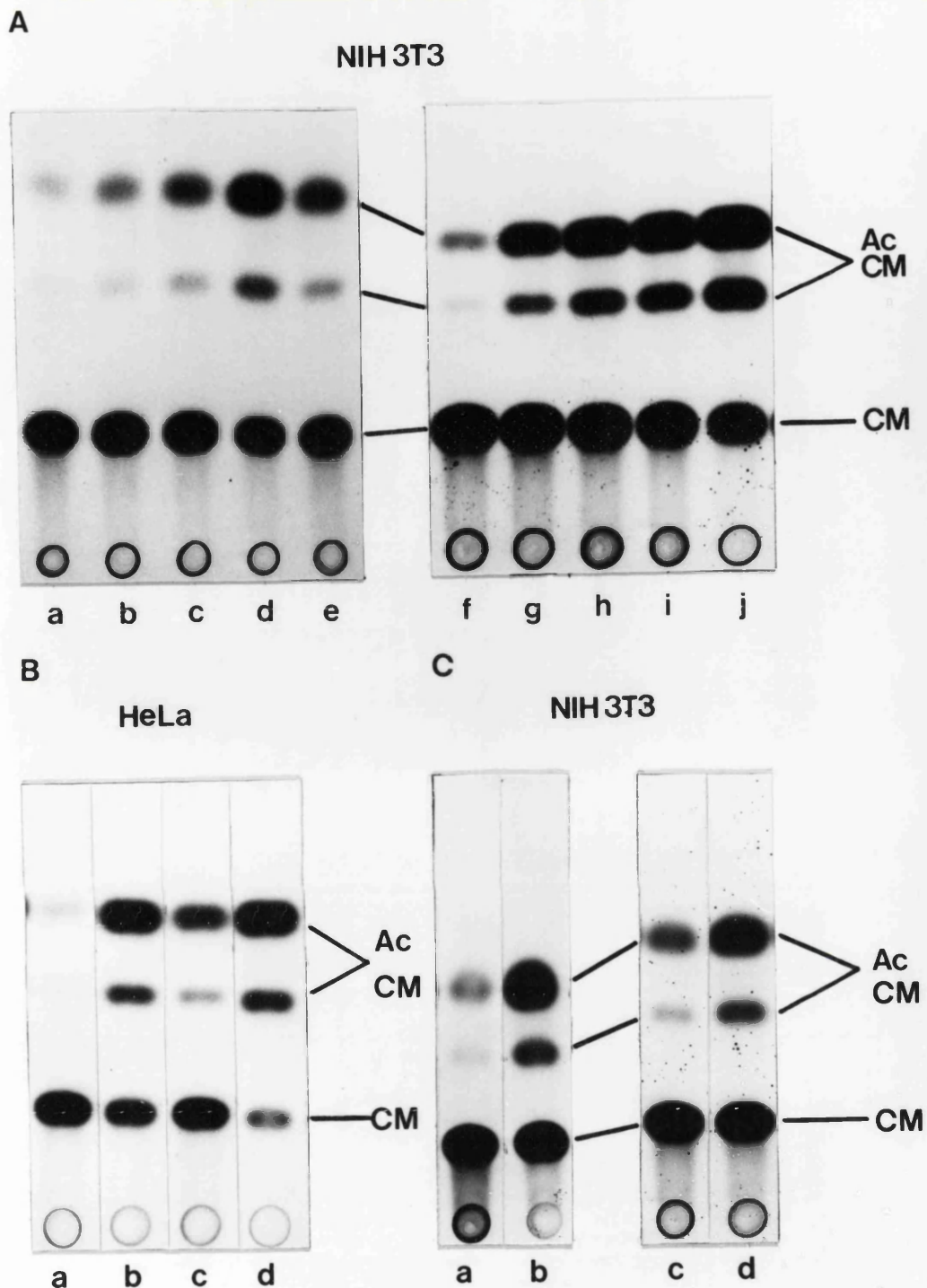


Figure 21. VL30 promoter and enhancer function in N-ras transformed cells. A; $2\mu\text{g}$ of the test plasmid (a-e pNVL-3CAT, f-j pNVL-1CAT) were co-transfected with 0 (a,f), $0.2\mu\text{g}$ (b,g), $0.5\mu\text{g}$ (c,h), $1\mu\text{g}$ (d,i) or $2\mu\text{g}$ (e,j) of pNRSac in NIH 3T3 cells. B; $2\mu\text{g}$ of test plasmid (a,b pNVL-3CAT, c,d pNVL-1CAT) were co-transfected with 0 (a,c) or $1\mu\text{g}$ (b,d) of pNRSac in HeLa cells. C; $2\mu\text{g}$ of the test plasmid (a,c pNVL-3TKCAT, b,d pNVL-1TKCAT) were co-transfected with 0 (a,c) or $1\mu\text{g}$ (b,d) of pNRSac in NIH3T3 cells. The total input of DNA was made up to $4\mu\text{g}$ (A) or $3\mu\text{g}$ (B,C) with carrier pBR322 DNA. After 48 hours the cells were harvested and assayed for CAT activity.

Cell Type	pNVL-3CAT	pNVL-3CAT +pNRSac	pNVL-1CAT	pNVL-1CAT +pNRSac
NIH 3T3	1	12	1	14.6

Cell Type	pNVL-3TKCAT	pNVL-3TKCAT +pNRSac	pNVL-1TKCAT	pNVL-1TKCAT +pNRSac
NIH 3T3	1	7	1	4

Table R.2 N-ras transactivation of NVL-1 and NVL-3 promoters and enhancers.

R.5.2 AN ENHANCER MEDIATES THE RESPONSE TO ACTIVATED N-RAS

To investigate whether the response to N-ras is mediated via an enhancer function, similar co-transfection experiments were carried out with the enhancer testing constructs pNVL-1TKCAT and pNVL-3TKCAT. In these experiments a roughly molar ratio of the LTR-CAT construct and pNRSac was used. Figure 21C shows that both enhancer testing constructs are transactivated by the activated N-ras oncogene. The results of these experiments are quantitated in table 2. These results show that the response to ras is mediated via an enhancer function.

R.5.3 TRANSACTIVATION BY ACTIVATED RAS GENES IS GENE DOSAGE

DEPENDANT

The EC34 cell line is a Ha-c-ras transformed NIH 3T3 cell line that shows elevated expression of VL30 (see figure 11C). In order to discover whether the expression of VL30 elements in this cell line could be further elevated co-transfection experiments were conducted using EC34 cells. Figure 22 shows the result of transfection of equal molar ratios of pNVL-3CAT or pNVL-1CAT with the N-ras expression vector pNRSac. Clearly the expression of NVL driven CAT in this cell line is further elevated when co-transfected with the pNRSac plasmid. This result indicates that the amount of increase in expression is related to the gene dosage of the activated ras.

R.5.4 THE PROTEIN KINASE C PATHWAY IS NOT A MAJOR PATHWAY

MEDIATING THE RESPONSE TO N-RAS

In order to investigate whether the protein kinase C signal transduction pathway is important in mediating the response to N-ras, co-

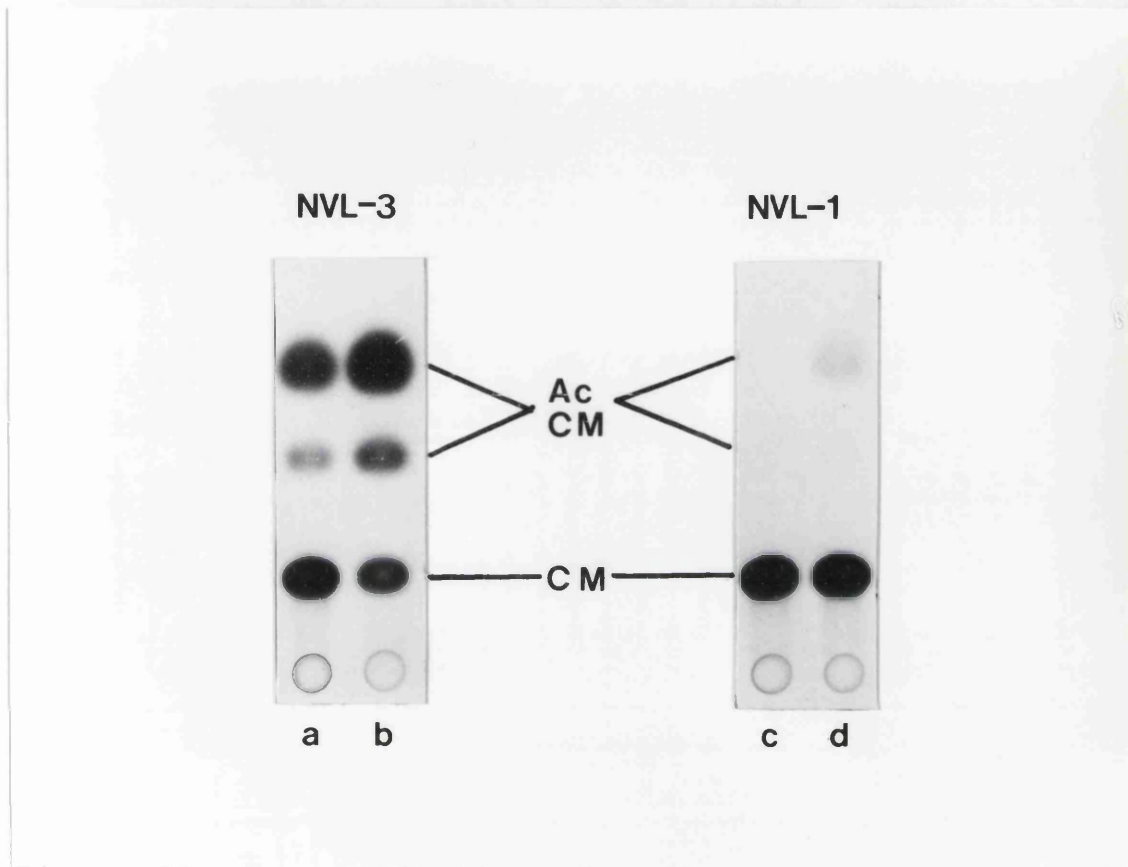


Figure 22. Transactivation of VL30 promoters in Ha-c-ras transformed cells. 2 μg of the LTR-CAT constructs were transfected into EC34 cells with either 0 (a,c) or 1 μg (b,d) of pNRSac. The final DNA input was adjusted to 3 μg with carrier pBR322 DNA. After 48 hours the cells were harvested and assayed for CAT activity.

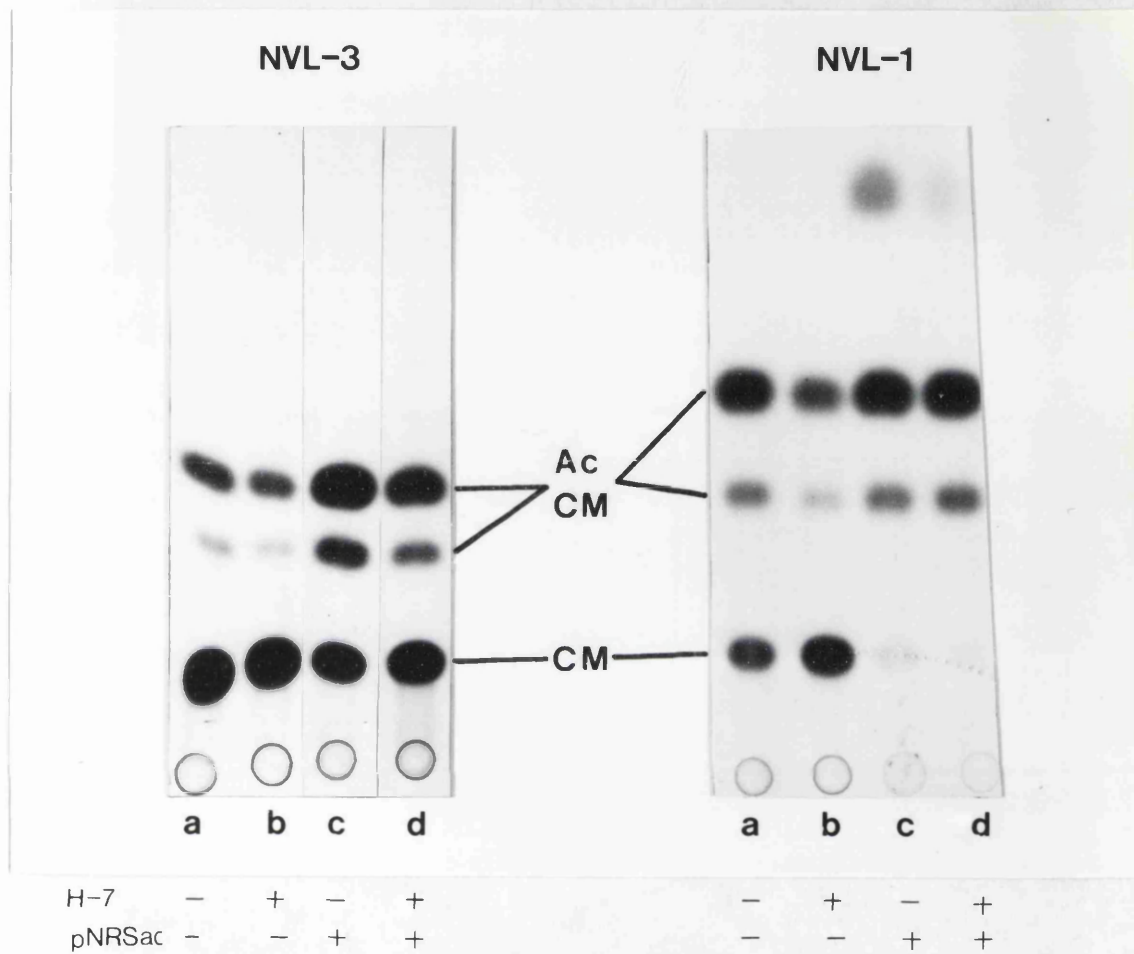


Figure 23. The effect of the protein kinase C inhibitor H-7 on N-ras transactivation. The LTR-CAT constructs were transfected with 0 (a,b) or 1 μ g (c,d) of pNRSac. Post-transfection the cells were split and incubated in fresh medium either in the presence (b,d) or absence (a,c) of H-7 for 48 hours before harvesting for CAT assay.

transfection experiments were done using molar ratios of pNVL-3CAT or pNVL-1CAT and pNRSac as described above. When the cells were split post-transfection, half were treated with the specific protein kinase C inhibitor H-7 at a final concentration of 140 μ M. Figure 23 shows the results of these experiments for both NVL-1 and NVL-3. Clearly the N-ras is still able to transactivate the LTRs, even in the presence of the inhibitor. However the overall expression appears to be somewhat lower when the inhibitor is present.

R.6 THE VL30 LTRs ARE NOT SUFFICIENT TO MEDIATE THE SERUM RESPONSE

The NVL-3 and NVL-1/2 VL30 elements respond differently to stimulation by serum, the response seen for the NVL-1/2 type of element is much greater than that seen for the NVL-3 elements (see fig 10C). The U3 domain is the only region of these elements that differs. Therefore it is of interest to see whether NVL CAT constructs containing only these regions show different responses to serum in transient or stable expression assays.

R.6.1 A SERUM RESPONSE WAS NOT OBTAINED IN TRANSIENTLY TRANSFECTED NIH 3T3 CELLS

Transient transfection assays were done in which the cells were transfected with either pNVL-1CAT or pNVL-3CAT and then cultured in serum free medium for 48 hours before stimulation with 15% FCS for 2 hours or 5 hours. Cells were harvested and both RNA and protein extracts were prepared. In several separate experiments it was impossible to obtain a serum response for either the endogenous elements or the VL30-CAT

constructs (data not shown). This data suggests that the transiently transfected cells were not able to become properly quiescent.

R.6.2 STABLY TRANSFECTED LTRs ARE NOT SERUM RESPONSIVE

As the transiently transfected cells apparently failed to achieve true quiescence in serum free medium, stably transfected cell lines were prepared containing the pNVL-1CAT or pNVL-3CAT constructs. A number of cell lines were prepared with either the NVL-3 or NVL-1 LTR/CAT constructs stably integrated. Stable transfectants were also prepared with the antisense enhancer testing constructs. The stably transfected cell lines containing the LTR/CAT promoter testing construct were serum starved and stimulated with 15% FCS for 2 hours or 5 hours as described previously. For each time point cells were harvested and both RNA and protein extracts were prepared. Figure 24 shows that the endogenous VL30 elements in these cell lines show the usual pattern of expression in quiescent and stimulated cells. However the CAT assay shows little or no response. These results indicate that the sequences required for mediating the serum response are not present in the LTR-CAT constructs. This is a surprising result as this region of the LTR contains the only sequence differences between the highly responsive NVL-1 and the less responsive NVL-3 elements.

R.7 VL30-CAT TRANSCRIPTS ARE INITIATED FROM WITHIN THE LTR

In order to confirm that transcription of the various gene constructs was initiated from the correct 5' cap, primer extension analysis was performed on RNA isolated from cells stably transfected with both the promoter and enhancer testing constructs. RNA from stably transformed cells was used because the low transfection efficiency

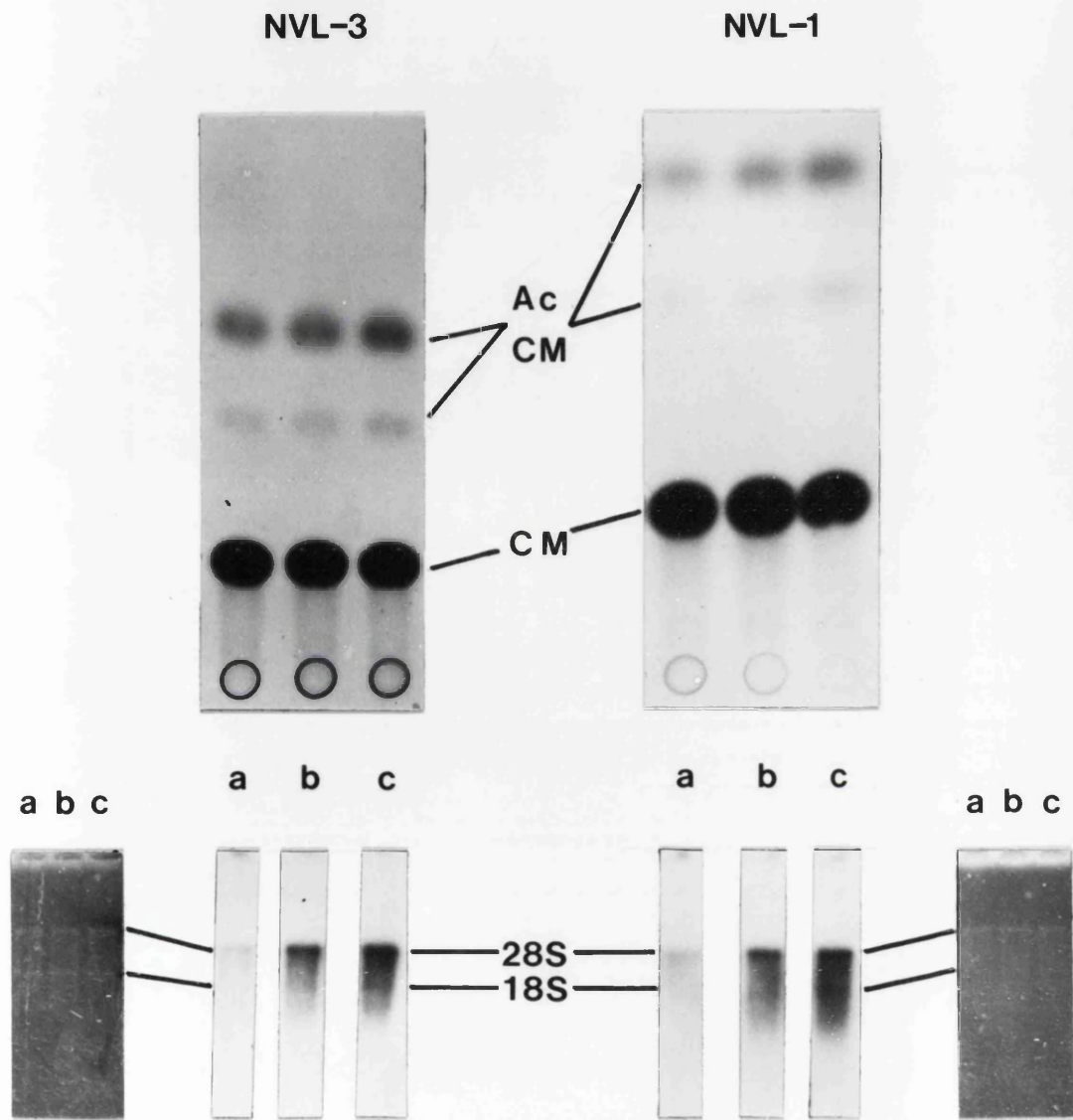


Figure 24. Expression of stably transfected LTR-CAT constructs in quiescent and stimulated NIH 3T3 cells. NIH 3T3 cells stably transfected with pNVL-3CAT or pNVL-1CAT were seeded into identical duplicate plates, serum starved for 48 hours (a) and stimulated with 15% FCS for 2 hours (b) and 5 hours (c) before harvesting for CAT assay and Northern analysis. The ethidium bromide stained gels alongside were used to prepare the blots.

obtained with the transient transfection system would not have given sufficient transcripts. A 26bp primer complementary to nt24-49 of the CAT gene mRNA was annealed to RNA prepared from the stably transformed cell lines. The results of the primer extensions are shown in figure 25. The largest primer extension products obtained from cells stably transfected with the pNVL-1CAT and pNVL-3CAT constructs are both about 145bp long. This maps the position of transcriptional initiation to within the LTRs, but it is about 40bp downstream of the putative 5' cap previously mapped in NIH 3T3 cells (see figure 3 and 118). From this experiment it is impossible to say whether this is the genuine position of initiation or whether it is a 'strong stop' band caused by secondary structure. However there are clearly no primer extension products larger than 145bp indicating that whatever the precise position of the 5' cap, CAT transcripts are initiated from within the LTRs.

The primer extension products obtained from cells stably transfected with pNVL-3TKCAT and pNVL-1TKCAT are about 145bp long. This maps the position of initiation of the transcripts to the 5' cap of the TK promoter region.

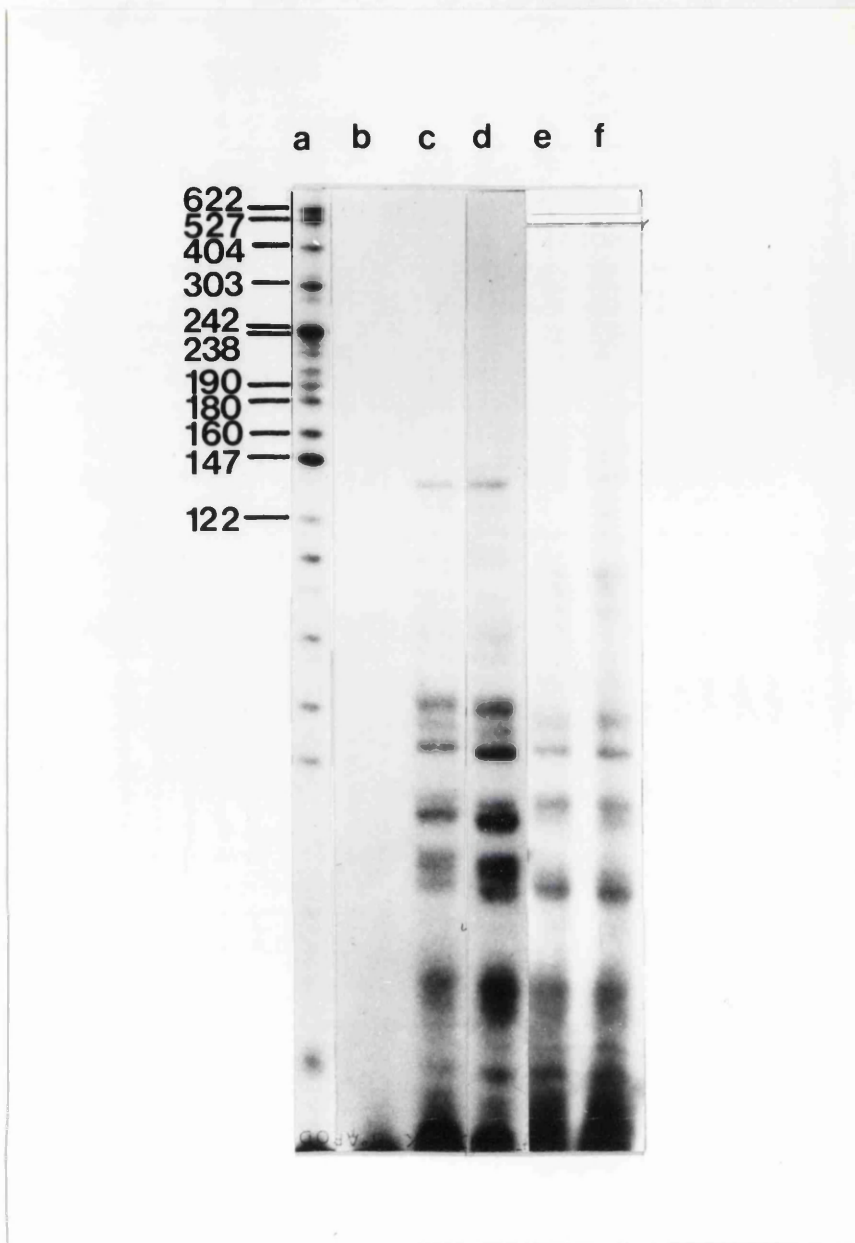


Figure 25. Primer extension analysis of NIH 3T3 cells stably transfected with VL30 promoter and enhancer testing constructs. 15µg of RNA from the stably transfected cells was annealed with 0.5×10^6 cpm of ^{32}P labelled probe, primer extension was performed and the products were run on an 8% polyacrylamide-8M urea sequencing gel. Lane a) MspI fragments of pBR322 DNA, b) no DNA, c) pNVL-3CAT, d) pNVL-1CAT e), pNVL-3TKCAT, f) pNVL-1TKCAT.

DISCUSSION

The purpose of this project has been to relate the structural features of the NVL VL30 LTRs to the expression of these elements, particularly with respect to stimulation by serum and to malignant cellular transformation.

D.1 STRUCTURAL ANALYSIS

D.1.1 STRUCTURAL FEATURES OF THE NVL VL30 LTRs

The NVL subset of the VL30 LTRs possess a number of interesting features. The NVL-3 LTRs and the regions flanking them were sequenced previously and various characteristic features were noted (118). Many of these features are conserved in the NVL-1 and NVL-2 LTRs and their flanking regions. For example these LTRs possess the overall retrovirus-like structure seen in NVL-3, that is U3 R and U5 domains flanked by short inverted repeat regions. The U5 and R domains are identical in all three elements. This includes the inverted repeat regions and the minus strand PBS at the 3' end of the 5' LTR. One of the most striking features observed in NVL-3 was the presence of two extra copies of the 3' IR and the PBS adjacent to it (118). This unusual feature is also seen in NVL-2, which has two extra copies, and NVL-1, which has one extra copy of this region. Norton *et al* have suggested that this duplication has arisen as a result of a successive series of duplications during reverse transcription (112). Conservation of this repetition implies that it is functionally important, perhaps in the process of retrotransposition. NVL-1 and NVL-2 also possess the imperfect tandem repeats seen in the R region of NVL-3. The sequence conservation of this repeat in all of the NVL LTRs suggests that it is also of functional significance.

Although the three LTRs are identical in the U5 and R domains, the U3 domains of the NVL-1/2 and NVL-3 elements differ in a number of ways.

The most obvious difference is the 105bp insertion seen in NVL-1/2 compared to NVL-3. The insert consists of three imperfect copies of a 35bp motif. Conserved within each copy of this motif is a short sequence similar to the binding site of the transacting factor AP-1 (21). The NVL-1/2 AP-1-like sequence is CATGACTCCT compared to a consensus sequence of G/ACTGAG/CTCAT. AP-1 is a member of a family of DNA binding proteins which all bind the same or similar sequences (68,69), therefore a small sequence divergence from the consensus may reflect the binding of a factor that is a member of this family rather than of AP-1 itself. The presence of this sequence motif is of interest because AP-1 sites in a number of genes have been shown to be essential for mediating the effects of growth factor stimulation. For example the vimentin gene is a serum responsive gene which has been shown to have a peak of expression about 8 hours after serum stimulation. A 24bp region containing two copies of an AP-1 binding site have been shown to be essential for mediating the response of this gene to serum (20). Also the AP-1 site has been shown to be important in mediating the effects of activated Ha-c-ras on a number of genes. For example, the polyoma virus enhancer contains an AP-1 site which is responsible for mediating the effects of Ha-c-ras and serum (73). It is interesting to note that the 35bp motif is repeated. As mentioned in the introduction the presence of more than one copy of a motif is a characteristic feature of many enhancers. It is also of interest that the NVL-3 LTR possesses a single copy of an AP-1-like site. This sequence is slightly different from the triple repeat AP-1-like site and occurs in the region between the 105bp insert and the promoter which is not present in NVL-1/2.

This region of NVL-3 also contains a CCAAT box motif not present in NVL-1/2. From this study it is not possible to say what effect, if

any, the absence of the CCAAT box has on expression of NVL-1/2. However this sequence motif is thought to be important in regulating efficient initiation of transcription (1) and therefore its absence in NVL-1/2 may be reflected in the lower basal levels of expression observed for NVL-1/2 in the transient expression assays.

Both NVL-1/2 and NVL-3 possess TATA box motifs, however the sequences of these elements differ in the two LTRs. Once again this project does not question the role of the different TATA boxes in regulating VL30 expression, however there is evidence that different TATA boxes respond to different stimuli, presumably due to interactions with different transacting factors (149,150). This type of regulation may be important in controlling the responses of the NVL VL30 elements to various stimuli such as the serum response where there are clear differences in the control of expression between the two types of element. Alternatively the factors binding to the TATA boxes could be present at different levels within the cell thus affecting expression. Another possibility is that the interaction between different TATA binding factors bound to the two different promoters and other regulatory proteins could be different, thus affecting expression.

D.1.2 NVL VL30 ELEMENTS ARE STRUCTURALLY DISTINCT FROM OTHER

VL30 ELEMENTS

A number of other VL30 elements have been cloned and their LTRs sequenced (117,119,125). For the most part these are genomic clones and therefore their functional relationship to the NVL VL30 elements is unknown. The sequence comparison in Figure 4 shows that there is a great deal of sequence homology between the NVL LTRs and the other VL30 LTRs. However the VL30 elements form a distinct subset, resembling each other

more than the other elements. This is particularly true of the U5 and R domains where the NVL elements are identical. Figure 4 does not show the small deletions and base changes that occur in the U5 and R regions in other VL30 elements relative to NVL elements.

Also it is interesting to note that none of the other elements have the unusual repetition of the IR and tRNA PBS seen in NVL elements. This may reflect the fact that these other elements are not transcribed or retrovirus transmissible and therefore would not undergo the reverse transcription that caused the duplication of these regions. Sequence conservation in the NVL elements suggests that this duplication may have functional significance in the retrotransposition of this subgroup of elements.

The distinct subset formed by the NVL LTRs with respect to sequence confirms the observations of Carter *et al* (151) who used hybridization studies to study the relationships between the various VL30 elements, they also showed that the NVL elements formed a distinct subset.

The U3 domains tend to be more heterogenous in terms of sequence. However even within this domain the various LTRs share a number of regions of homology. The two clones that show the closest sequence relationships to the NVL LTRs are λ 5 (125) and VL-3 (119).

λ 5 is a cDNA isolated from SV40 transformed NIH 3T3 cells and is described as growth specific (125). The regions of the LTR found in the clone are almost identical to those of NVL-1/2 and presumably λ 5 is derived from an NVL-1/2 transcript. However the λ 5 clone does not include any of the U5 domain, therefore it is possible that the element that this clone was derived from possesses a U5 domain that is not homologous to the NVL U5 domain.

The clone VL3 is also of interest as the U3 domain of this clone is almost identical to that of NVL-3. The U5 and R domains however are quite different from those of any other VL30 element sequenced (119). VL3 has been shown to possess strong promoter activity and therefore it is strange that this element was not cloned with the NVL elements. This observation suggests that although the VL3 element is transcribed it is not packaged by retroviral particles. The sequence divergence of this element from NVL-3 in the R and U5 domains indicates that these domains may possess essential signals required for the packaging of the NVL VL30 elements by retroviruses.

D.1.3 STRUCTURAL RELATIONSHIPS WITH OTHER RETRANSPOSABLE ELEMENTS

The VL30 elements are structurally similar to other retrotransposons. That is they possess a similar retrovirus-like structure consisting of U3 R and U5 domains within LTRs. The identical LTRs surround an internal unique region. However the sequence analysis shows that there is little sequence homology between the VL30 LTRs and those of other retrotransposons. This includes other murine elements, such as IAPs (93) and MuRRs (152) and elements from other organisms, such as *Drosophila* copia-like (94) and Yeast Ty (96) elements. This suggests that the VL30 elements are evolutionarily distinct from these other elements. The only homologies that can be seen are the IRs which are similar to those of retroviruses and retrotransposons. In fact sequence data from a large number of retroelements shows that these regions are very well conserved and it has been suggested that this is because all of these elements use a similar method of integration (81).

VL30 elements, like copia and Ty, appear to be some sort of less highly evolved form of retroelement, that is they appear to be an

intermediate entity between the cellular mobile elements and retroviruses. However unlike copia and Ty, VL30 does not possess ORFs encoding retroviral genes, although there are homologies to these genes in the unique region (123).

D.2 EXPRESSION OF VL30 RNA

D.2.1 SERUM RESPONSE

VL30 elements are responsive to serum stimulation of quiescent NIH 3T3 cells. The kinetics of the response indicate that the NVL VL30 elements are not early response genes like c-fos (11), c-jun (17,18) and a large number of other serum responsive genes (59,60). These immediate early response genes are induced rapidly but transiently, whereas VL30 elements are induced at a later time point and maintain elevated levels for several hours. These sort of expression characteristics have been seen for a few other genes. Examples of these genes include transin, a matrix degrading metalloproteinase (153), TIMP (tissue inhibitor of metalloproteinase) (75), proliferin, a member of the prolactin-growth hormone family (154) and vimentin, a subunit of a filament important in the cytoskeleton (20). Unlike the immediate early genes which all show very similar expression characteristics these genes form a more heterogeneous group. For example their kinetics of expression are all slightly different and although most of these genes require protein synthesis for expression, there are exceptions such as vimentin (20). VL30 elements themselves do not require protein synthesis for EGF stimulated expression (128).

Not all VL30 elements respond to serum. The S1 nuclease mapping data clearly shows that the NVL-1/2 elements are far more responsive than the NVL-3 elements. The independent regulation of different members of

the NVL VL30 elements is also seen during development. In most tissues there is a high degree of regulation of expression of NVL-1/2 elements, whereas NVL-3 elements are expressed relatively constitutively (145). The proliferin gene family also possesses members that are not responsive to serum stimulation as well as genes that are responsive (155). The regulatory regions of the responsive and unresponsive genes differ only in one small region. Also the 16C8 gene (TIMP) was shown to have differentially regulated transcripts (75). In this case there is only one gene and it is thought that there may be several transcriptional start sites that respond to different regulatory signals.

The serum response occurs at least in part at the transcriptional level. The magnitude of the response in the nuclear run on experiments was somewhat lower than expected, however the presence of a large number of transcripts produced by the relatively unresponsive NVL-3 elements could result in there being quite a high level of constitutive expression. This could obscure the effects of serum starvation and stimulation. Also the level of response does tend to vary somewhat from experiment to experiment. It is possible that there is an element of post-transcriptional control. This is supported by the fact that VL30 expression is slightly induced by cycloheximide (128) indicating that mRNA stability may play a role in VL30 expression.

The serum response is retained when VL30 elements are transferred to heterologous VLNRK cells indicating that the response is controlled by sequences within the VL30 elements. However it is interesting to note that the kinetics of the response in the rat cells is somewhat different from in NIH 3T3 cells. The peak in expression occurs somewhat earlier than in NIH 3T3 cells, but still occurs much later than the peak in expression seen for c-fos and other early response genes (11,59). Also

the magnitude of the response is somewhat lower than usually observed in NIH 3T3 cells. One possible explanation for both observations is that there is different availability of the necessary transcription factors in the NRK cells.

Interestingly only a very low level of expression was obtained in similar experiments with RAT1 cells. Also there was apparently no serum response in these cells. This result could be partly due to a very low copy number of VL30 elements resulting in very low levels of expression. However it is important to note that the parental Rat1 cells were also used for transfection assays to assess the promoter activity of the NVL LTRs and in this cell line there was only very low promoter activity for all the constructs tested. This data would suggest that the VL30 elements are not active in this cell line. This contradicts the results of Rodland *et al* who have transferred VL30 elements to RAT1 cells and obtained a response to EGF and to TPA (156). It is difficult to explain this apparent contradiction. However as these cells lines came from different sources it is possible that the cells used by Rodland *et al* and the ones used in this project may have quite different properties, reflected in the differences observed in VL30 expression.

D.2.2 ELEVATED EXPRESSION IN TRANSFORMED CELL LINES

It has long been acknowledged that the rate of growth of a cell line has an effect on the levels of VL30 expression (114,125). In particular many transformed cell lines have elevated levels of VL30 expression (114,126,157). The data presented here supports this to an extent as both spontaneously transformed cells and cells transformed by c-Ha-ras have elevated levels of VL30 transcripts. Also co-transfections with N-ras result in elevated levels of expression from VL30 enhancers

and promoters. However the data presented here agrees with earlier work which showed that VL30 expression is not elevated in Ki-v-ras transformed cells (114). These observations suggest a number of things. First of all clearly the elevated levels of VL30 in the spontaneously transformed cells and those cells transformed by N- or Ha-ras are not due merely to the uncontrolled growth seen in transformed cells, instead the VL30 elements are directly affected by the events involved in cell transformation. Also this result suggests that Ki-ras and Ha- and N-ras act, at least in part, via different mechanisms. Although these results suggest the interesting possibility that the expression of oncogenic ki-ras and Ha- and N-ras may have different effects on VL30 expression it is by no means conclusive. It would be most useful to study VL30 expression in a larger number of cells transformed by the three ras oncogenes in order to see whether these results can be reproduced. Also it would be interesting to repeat the co-transfection experiments using ki- and Ha-ras.

Interestingly the revertant cell lines derived from the Ki-v-ras transformed cells show elevated levels of expression. This may be attributed to the way that the cells were produced. The parental transformed cell line CCl was mutagenized with BuDR to give revertants (131). As mentioned in the introduction this drug has been shown to induce VL30 expression (114), however there is no evidence that there is any long term effect.

Unlike the serum response both types of VL30 element are responsive to cellular transformation. This suggests that different control mechanisms are responsible for growth factor induced and transformation induced changes in VL30 expression. This further supports the idea that the response to transformation is not merely a response to the elevated growth rate of these cells but that it is due to specific changes which directly effect VL30 expression.

A number of transformed cell lines have been shown to produce growth factor-like substances that can act in an autocrine/paracrine fashion to induce transformation in normal cells (157). However conditioned medium obtained from the Ha-c-ras transformed cells had no effect on the expression of VL30 elements in normal NIH 3T3 cells. The experiments could have been improved by a number of small changes such as using NIH 3T3 cells that had been serum starved before the addition of the conditioned medium. Also it may have been useful to have concentrated the conditioned medium before use. An alternative approach to addressing this question could have been to use purified transforming growth factors. The results suggest that changes in VL30 expression in transformed cells are due to a direct effect of the transforming agent, such as the oncogene or chemical mutagen, rather than to the presence of a transforming growth factor. This is supported by the work of Owen *et al* who showed that an increase in VL30 expression in ras transformed cells was directly due to ras gene expression (159).

D.3 PROMOTER AND ENHANCER PROPERTIES OF THE NVL VL30 U3

DOMAINS

In order to relate the structural aspects of the NVL VL30 LTRs to their expression it was necessary to test the LTRs for promoter and enhancer function.

D.3.1 BASAL LEVELS OF PROMOTER AND ENHANCER ACTIVITY

Clearly the differences in the structures of NVL-1/2 and NVL-3 are reflected in their promoter activities. The NVL-3 LTR promoter is very strong compared to both the NVL-1/2 promoter and the prototype promoter from SV40.

Both the elements possess quite strong enhancer activity when linked to a heterologous promoter. This is important because VL30 elements could possibly affect the expression of mouse cellular genes by an insertion mutagenesis type of mechanism (88). With the very high enhancer activity exhibited by these elements it is clear that an activating insertion by these elements could have a drastic effect. This type of effect has been shown for other retrotransposons such as Ty, which exerts its transcriptional regulatory effects on cellular genes via

enhancer-like sequences (104). Recently it was reported that in some tissues VL30 probes detected transcripts much larger than normal VL30 elements (129). It would be interesting to know whether these transcripts came from VL30 elements fused to cellular genes.

Higher CAT activities were obtained when the U3 domain was placed in the opposite orientation with respect to the promoter. A possible explanation is that the slight change in position with respect to the promoter that changing the orientation of the LTR would have caused, may facilitate better interactions with any promoter bound trans-acting factors.

The promoter and enhancer functions are also active in HeLa cells. Similar relative levels of promoter activity for NVL-1/2 and NVL-3 are seen, although the levels of SV40 driven CAT are somewhat higher than in NIH 3T3 cells. This indicates that VL30 expression is not species or cell type specific. However as noted previously there are only very low levels of expression from both LTRs in Rat1 cells. This suggests that there is a certain degree of tissue specificity.

D.3.2 RESPONSE TO N-RAS IS MEDIATED VIA AN ENHANCER FUNCTION

Both NVL-1/2 and NVL-3 LTRs show increased promoter activity when they are co-transfected with an N-ras expression vector. This is an enhancer mediated response and appears to be of a similar magnitude for both the LTRs. This suggests that the response is mediated by the same sequences in both NVL-3 and NVL-1/2. In fact this may not be the case. A recent report by Owen *et al* (159) shows that for NVL-3 a 53bp region is required for mediating high basal levels of expression and the response to Ha-c-ras in co-transfection experiments. The region that they show to be absolutely necessary for the response is in the region 5' to

the promoter and contains an AP-1-like site. There is only limited homology between NVL-1/2 and NVL-3 in this region suggesting that the ras response is mediated via different sequences in NVL-1/2 and NVL-3. Perplexingly Owen *et al* also show that Ki-v-ras mediated stimulation of NVL-3 is controlled by this region. This contrasts directly with the results presented here, where no Ki-v-ras induction of expression was seen for endogenous VL30 expression. Very preliminary results not presented here indicated that NVL-3 CAT constructs were not induced in co-transfections with a Ki-ras expression vector. It is difficult to explain these conflicting results especially in the light of previous reports of normal levels of VL30 expression in Ki-v-ras transformed cells (114). A possible explanation is the gene dosage, that is there may be a higher level of ras in the cells used by Owen *et al*. The data presented here (Figures 21 and 22) show that increasing the level of ras input increases the level of VL30 expression.

D.3.3 LACK OF SERUM RESPONSE IN THE U3 DOMAIN OF NVL-1

The most surprising result obtained was that the NVL-1 U3 domain was not able to confer a serum response in transfection assays. The U3 region contains the only known differences between the NVL-1/2 and the NVL-3 elements and it therefore seems reasonable to suggest that the sequence determinants responsible for controlling the response to serum would be found within this region. However the results clearly show that whilst endogenous VL30 elements are regulated as usual there is no serum mediated regulation of the LTR/CAT constructs.

The most simple explanation is that sequence requirements downstream from the SacI site are required for the serum response. There are a number of cases where downstream regulatory regions have been shown

to be important in controlling transcription. The two most relevant examples are the retrotransposons Ty and copia. Both class I and class II Ty elements have been shown to require downstream enhancer-like elements for full transcriptional activity (96,160). Copia elements have been shown to have control elements on either side of the transcriptional start site and a regulatory region within the unique region (95). The region within the LTR downstream of the transcriptional start site appears to be important in regulating tissue specific expression.

However it must be taken into consideration that the downstream sequences are the same in both NVL-1/2 and NVL-3 elements. The following model could explain how the serum response is controlled in VL30 elements. A downstream sequence, presumably common to both NVL elements, binds an essential factor for the serum response. Also sequences unique to NVL-1/2 bind another essential factor(s) and then protein protein interactions occur between the bound factors to facilitate serum induction. This model would allow for the low levels of induction seen for NVL-3 in the S1 mapping as it would allow for binding of a factor(s) to another site(s) that interacted less well with the downstream factor. This model is speculative and analysis of downstream sequences is required in order to find the serum regulatory regions. Control of serum responsive expression of VL30 elements seems likely to be quite complicated and it appears to involve at least two regulatory elements.

It is also possible that negative regulatory elements are involved. This is the case for c-fos expression where negative factors are involved in shutting off expression in quiescent cells and positive factors are required to stimulate expression upon addition of serum (161). Serum stimulation results in modification of the positive or negative factors in order to change their affinities for the serum response element.

There are some genes that are expressed only in quiescent cells (162). It would be interesting to know whether these genes are involved in controlling directly or indirectly the repression of serum responsive genes, such as NVL-1/2, in quiescent cells.

It was interesting to note that a true quiescence was not obtained in the transiently transfected cells and that endogenous VL30 expression was not down regulated or stimulated. A similar result was obtained by Rittling *et al* (20) for the serum inducible vimentin gene which shows similar kinetics to VL30. Their work showed that in transiently transfected cells they only obtained a 2 fold induction whereas in stably transformed cells they saw a normal response. They suggested that genes such as vimentin require a 'deep' quiescence not obtained in transiently transfected cells. Miska and Bosmann (163) have shown that there are different levels of quiescence. A deeper quiescence can be obtained which requires a longer lag period before DNA synthesis occurs after addition of serum or growth factors. It is possible that the transfections result in the cells being unable to reach this deep quiescence and that VL30 and other genes such as vimentin require this deep quiescence before they are completely downregulated.

D.4 INDUCTION OF VL30 EXPRESSION BY SERUM AND RAS IS MEDIATED VIA PROTEIN KINASE C INDEPENDENT PATHWAYS

The protein kinase C signal transduction pathway is important in mediating the effects of growth factor stimulus (45,47). Also it has been shown that protein kinase C is essential for mediating some of the effects of ras (56). Therefore it was of interest to find that this signal transduction pathway is unimportant in mediating either the effects of serum stimulation or ras mediated transformation for VL30

elements. Recently it has been shown that ras induces c-myc expression via a protein kinase C independent mechanism (57).

Unfortunately the pathways that are involved in mediating these responses are unknown. The presence of AP-1-like sites in the LTRs suggested that the protein kinase C pathway may be involved as AP-1 was first identified as the TPA responsive element (21). TPA is a potent stimulator of protein kinase C (51). The presence of AP-1-like sequences within the LTRs may indicate that another member of this large family of trans-acting factors is involved. As mentioned previously there is a family of proteins closely related to AP-1. The cAMP response element (CRE), for example, is very similar to the AP-1 binding site (164). The factor that binds to the CRE (CREB) is a member of the AP-1/c-jun family and possesses the characteristic leucine zipper motif (26). cAMP is a second messenger in another signal transduction pathway and mediates its effects via protein kinase A (164). Recently it has been shown that there is cross talk between the protein kinase A and protein kinase C signal transduction pathways. In certain cell types and for some genes AP-1 and CREB can bind to each others sites and activate transcription (164). A cDNA apparently derived from NVL-3 has been isolated from T cells treated with forskolin (129). Forskolin is an activator of cAMP and it would be interesting to see if VL30 elements in NIH 3T3 cells are stimulated by treatment with this drug.

In the transient transfection assays where H-7 was used there was a small overall decrease in CAT activity, although the magnitude of the ras response was not substantially altered. There are two explanations for this observation. The first is that the protein kinase C signal transduction pathway plays a small role in mediating basal levels of expression. The second is that the slight decrease in expression is due

to non-specific effects of the inhibitor or to a general slowing down in cell growth due to a lack of protein kinase C, rather than due to a specific lack of the enzyme having a direct effect on VL30 expression

D.5 CONCLUSIONS

A number of conclusions can be drawn from the results presented here.

Firstly the NVL class of transcriptionally active, retrovirus transmissible VL30 elements represents a subset of the VL30 family that is structurally as well as functionally distinct from the other VL30 elements.

The expression of the two types of NVL element, NVL-1/2 and NVL-3, is controlled by quite distinct mechanisms. This is reflected both in basal levels of expression and in their different responses to serum stimulation. Also there is indirect evidence that the response to ras is controlled by different sequences in the two types of element.

The response to cellular transformation and the serum response are controlled independently in VL30 elements, unlike a number of other genes where both responses are mediated by the trans-acting factor AP-1. Both responses have in common that they are not regulated via the protein kinase C signal transduction pathway.

The serum response is not mediated by sequences within the LTR of the NVL-1/2 elements. This suggests that there are sequences downstream of the transcriptional start site responsible for mediating this response.

Although VL30 elements are induced by activated Ha-c and N- ras, Ki-v-ras has no effect on VL30 expression. This suggests that VL30 elements may be important in investigating the effects of the different

members of the ras gene family. Also VL30 responds directly to the ras oncogene rather than to a more indirect signal from the transformed cell such as autocrine/paracrine factor.

There are still a great many questions to be asked about the way in which VL30 expression is controlled. For example it would be interesting to find the downstream sequences that are required for the serum response. Also it would be useful to define essential regions of the LTRs required for expression at the basal level and in response to the various stimuli that induce VL30 expression.

Also it would be of interest to use DNAase I footprinting and gel retardation assays to show that particular sequences are protein binding sites. This would also give an indication of the proteins that are bound and the pathways involved.

Finally as mentioned above it would be interesting to investigate further the different responses VL30 elements show to transformation by Ki-ras and Ha- and N-ras.

Analysis of the transcriptional regulation of VL30 elements is of interest for several reasons. Clearly these elements provide a model for a serum responsive, transformation responsive gene. Also as mentioned above they are potentially useful in understanding the action of the ras gene family.

On a more biological level these elements are of interest because of their potential to move around the genome and affect the expression of cellular genes in either a positive or negative way. The evidence presented here for strong promoter and enhancer activity coupled with previous evidence of the retrovirus transmissibility of these elements suggests that this is possible.

AKNOWLEDGEMENTS

I would like to thank Dr. John Norton for his supervision of this project and his help and advice over the past three years and during the preparation of this thesis. Also I would like to thank Prof. A. V. Hoffbrand for his support.

I would like to thank all the members of the Molecular Biology group in the Haematology department for their help and encouragement, particularly Brian Leber and Miriam Deane. Special thanks go to John Yaxley who has been particularly helpful in the last few months with obtaining photographs and figures from illustrations.

This thesis would have been impossible to achieve without the support and encouragement of my family and friends. I would like to thank my Mum and Dad for always encouraging me in whatever I have done. Also thanks to Lisa and Tim for keeping me on my toes.

Thanks also to all my friends, especially Jeanette, Danny and Andy.

The biggest thanks must go to my husband, Andy, who has never failed to encourage and support me. He has put up with the antisocial hours and bad moods as well as providing the financial support needed to allow me to finish this thesis.

REFERENCES

1. Maniatis, T., S. Goodbourn and J. A. Fischer (1987).
Regulation of inducible and tissue-specific gene expression.
Science 236 p1237-1245.
2. Sassone-Corsi, P. and E. Borelli (1986).
Transcriptional regulation by trans-acting factors.
Trends in Genet. 2 p215-219.
3. Dynan, W. S. (1989).
Modularity in promoters and enhancers.
Cell 58 p1-4.
4. Benoist, C. and P. Chambon (1981).
In vivo sequence requirements of the SV40 early promoter region.
Nature 290 p304-310.
5. Marriot, S. J. and J. N. Brady (1989).
Enhancer function in viral and cellular gene regulation.
Biochem. and Biophys. Acta 989 p97-110.
6. Serfling, E., M. Jasin and W. Schaffner (1985).
Enhancers and eukaryotic gene transcription.
Trends in Genet. 1 p224-230.
7. Zenke, M., T. Grundström, H. Matthes, M. Winterzith, C. Schatz, A. Wildeman and P. Chambon (1986).
Multiple sequence motifs are involved in SV40 enhancer function.
EMBO 5 p387-398.
8. Ondek, B., L. Gloss and W. Herr (1988).
The SV40 enhancer contains two distinct levels of organisation.
Nature 333 p40-45.
9. Dynan, W. S. (1986)
Promoters for housekeeping genes.
Trends in Genet. 2 p196-197.

10. Gardiner-Garden, M. and M. Frommer (1987).
CpG islands in vertebrate genomes.
J. Mol. Biol. 196 p261-282.
11. Greenberg, M. E. and E. B. Ziff (1984).
Stimulation of 3T3 cells induces transcription of the c-fos
proto-oncogene.
Nature 311 p433-438.
12. Treisman, R. (1985).
Transient accumulation of c-fos RNA following serum stimulation
requires a conserved 5' element and c-fos 3' sequences.
Cell 42 p889-902.
13. Treisman, R. (1986).
Identification of a protein-binding site that mediates
transcriptional responses of the c-fos gene to serum factors.
Cell 46 p567-574.
14. Treisman, R. (1987).
Identification and purification of a polypeptide that binds to the
c-fos serum response element.
EMBO 6 p2711-2717.
15. Rangnekar, V. M., A. C. Aplin, and V. P. Sukhatme (1990).
The serum and TPA responsive promoter and intron-exon structure of
EGR2, a human early growth response gene encoding a zinc finger
protein.
Nucleic Acids Res. 18 p2749-2757.
16. Joseph, L. J., M. M. Le Beau, G. A. Jamieson, Jr, S. Acharya, T. B.
Shows, J. D. Rowley, and V. P. Sukhatme (1988).
Molecular cloning, sequencing, and mapping of EGR2, a human early
growth response gene encoding a protein with "zinc-binding finger"

structure.

PNAS 85 p7164-7168.

17. W. W. Lamph, P. Wamsley, P. Sassone-Corsi and I. M. Verma (1988).
Induction of proto-oncogene jun/AP-1 by serum and TPA.
Nature 334 p629-631.
18. Ryseck, R.-P., S. I. Hirai, M. Yaniv and R. Bravo (1988).
Transcriptional activation of c-jun during the G0/G1 transition in
mouse fibroblasts.
Nature 334 p535-537.
19. Kerr, L. D., J. T. Holt and L. M. Matrisian (1988).
Growth factors regulate transin gene expression by c-fos-dependent
and c-fos-independent pathways.
Science 242 p1424-1427.
20. Rittling, S. R., L. Coutinho, T. Amram and M. Kolbe (1989).
AP-1 jun binding sites mediate serum inducibility of the human
vimentin promoter.
Nucleic Acids Res. 17 p1619-1633.
21. Angel, P., M. Imagawa, R. Chui, B. Stein, R. J. Imbra, H. J.
Rahmsdorf, C. Jonat, P. Herrlich and M. Karin (1987).
Phorbol ester-inducible genes contain a common cis element
recognised by a TPA-modulated trans-acting factor.
Cell 49 p729-739 (1987).
22. Mordacq J. C. and D. I. H. Linzer (1989).
Co-localization of elements required for phorbol ester stimulation
and glucocorticoid repression of proliferin gene expression.
Genes Dev. 3 p760-769.
23. Vincent, A. (1986).
TFIIIA and homologous genes. The finger proteins.

- Nucleic Acids Res. 14 p4385-4391.
24. Klug, A. and D. Rhodes (1987).
'Zinc fingers': a novel protein for nucleic acid recognition.
Trends Biochem. Sci. 12 p464-469.
25. Kadonaga, J. T., K. R. Carner, F. R. Masiarz and R. Tjian (1987).
Isolation of cDNA encoding transcription factor Sp1 and functional
analysis of the DNA binding domain.
Cell 51 p1079-1090.
26. Landschultz, W. H., P. F. Johnson and S. L. McKnight (1988).
The leucine zipper: a hypothetical structure common to a new class
of DNA binding proteins.
Science 240 p1759-1764.
27. Nakabeppu, Y., K. Ryder and D. Nathans (1988)
DNA binding activities of three murine jun proteins: Stimulation by
fos.
Cell 55 p907-915.
28. Halazonetis, T. D., K. Georgopoulos, M. E. Greenberg and P. Leder
(1988).
c-jun dimerizes with itself and with c-fos, forming complexes of
different DNA binding affinities.
Cell 55 p917-924.
29. Courey, A. J. and R. Tjian (1988).
Analysis of Sp1 *in vivo* reveals multiple transcriptional domains,
including a novel glutamine-rich activation motif.
Cell 55 p887-898.
30. Sassone-Corsi, P., L. J. Ransone, W. W. Lamph and I. M. Verma
(1988).
Direct interaction between fos and jun nuclear oncoproteins: Role

- of the 'leucine zipper' domain.
Nature 336 p692-695.
31. Ptashne, M. (1986).
Gene regulation by proteins acting nearby and at a distance.
Nature 322 p697-701.
32. Ptashne, M. (1988).
How eukaryotic transcriptional activators work.
Nature 335 p683-689.
33. Müller, H.-P., J. M. Sogo and W. Schaffner (1989).
An enhancer stimulates transcription in trans when attached to the
promoter via a protein bridge.
Cell 58 p767-777.
34. Heuchel, R., P. Matthias and W. Schaffner (1989).
Two closely spaced promoter are equally activated by a remote
enhancer: evidence against a scanning model for enhancer action.
Nucleic Acids Res. 17 p8931-8947.
35. Berk, A. J. (1989).
Regulation of eukaryotic transcription factors by post-translational
modification.
Biochem. Biophys. Acta 1009 p103-109.
36. Prywes, R., A. Dutta, J. A. Cromlish and R. G. Roeder (1988).
Phosphorylation of serum response factor, a factor that binds to the
serum response element of the c-fos enhancer.
PNAS 85 p7206-7210.
37. Yamamoto, K. K., G. A. Gonzalez, W. H. Biggs III and M. R. Montminy
(1988).
Phosphorylation-induced binding and transcriptional efficacy of
nuclear factor CREB.

- Nature 334 p494-498.
38. Jackson, S. P. and R. Tjian (1988)
O-glycosylation of eukaryotic transcription factors: implications
for mechanisms of transcriptional regulation.
Cell 55 p125-133.
39. Hólttá, E., L. Sistonen and K. Alitalo (1988).
The mechanisms of ornithine decarboxylase deregulation in c-Ha-ras
oncogene-transformed NIH 3T3 cells.
J. Cell Biol. 263 p4500-4507.
40. Bishop, J. M. (1983).
Cellular oncogenes and retroviruses.
Ann. Rev. Biochem. 52 p301-354.
41. Bishop, J. M. (1987).
The molecular genetics of cancer.
Science 235 p305-311.
42. Tabin, C. J., S. M. Bradley, C. I. Bargmann, R. A. Weinberg, A. G.
Papageorge, E. M. Scolnick, R. Dhar, D. R. Lowy and E. H. Chang
(1982).
Mechanism of activation of a human oncogene.
Nature 300 p143-149.
43. Herrlich, P. and H. Ponta (1989).
'Nuclear' oncogenes convert extracellular stimuli into changes in
the genetic program.
Trends Genet. 5 p112-116.
44. Pardee, A. B., R. Dubrow, J. L. Hamlin and R. F. Kletzien (1978).
Animal cell cycle.
Ann. Rev. Biochem. 47 p715-756.
45. Rozengurt, E. (1986).

- Early signals in the mitogenic response.
Science 234 p161-166.
46. Rozengurt, E., J. Erusalimsky, H. Mehmet, C. Morris, E. Nånberg and J. Sinnett-Smith (1988).
Signal transduction in mitogenesis: Further evidence for multiple pathways.
Cold Spring Harbor Symp. Quant. Biol. 53 p935-944.
47. Schlessinger, J., A. Ullrich, A. M. Honegger and W. H. Moolenaar (1988).
Signal transduction by epidermal growth factor receptor.
Cold Spring Harbor Symp. Quant. Biol. 53 p515-519.
48. Pardee, A. B. (1989).
G1 events and regulation of cell proliferation.
Science 246 p603-608.
49. Downward, J., Y. Yarden, E. Mayes, G. Sciacca, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger and M. D. Waterfield (1984).
Close similarity of epidermal growth factor receptor and v-erbB oncogene protein sequences.
Nature 307 p521-527.
50. Berridge, M. J., K. D. Brown, R. F. Irvine and J. P. Heslop (1985).
Phosphoinositides and cell proliferation.
J. Cell Sci. Suppl. 3 p187-198.
51. Nishizuka Y. (1984).
The role of protein kinase C in cell surface signal transduction and tumour promotion.
Nature 308 p693-698.
52. Wakelam, M., S. A. Davies, M. D. Houslay, I. McKay, C. J. Marshall and A. Hall (1986).

- Normal p21^{N-ras} couples bombesin and other growth factor receptors to inositol phosphate production.
Nature 323 p173-176.
53. Barbacid, M. (1987).
Ras genes.
Annu. Rev. Biochem. 56 p779-827
54. McCormick, F. (1989).
ras GTPase activating protein: Signal transmitter and signal terminator.
Cell 56 p5-8.
55. McKay, I. A., C. J. Marshall, C. Calés and A. Hall (1986).
Transformation and stimulation of DNA synthesis in NIH 3T3 cells are a titrable function of normal p21^{N-ras} expression.
EMBO 5 p2617-2621.
56. Morriss, J. D. H., B. Price, A. C. Lloyd, A. J. Self, C. J. Marshall and A. Hall (1989).
Scrape loading of Swiss 3T3 cells with ras protein rapidly activates protein kinase C in the absence of phosphoinositide hydrolysis.
Oncogene 4 p27-31.
57. Lloyd, A. C., H. F. Paterson, J. D. H. Morriss, A. Hall and C. J. Marshall (1989).
p21^{H-ras} induced morphological transformation and increases in c-myc expression are independent of functional protein kinase C.
EMBO 8 p1099-1104.
58. Berridge, M. J. (1985)
The molecular basis of communication within the cell.
Sci. Am. 253 p124-134.
59. Lau, L. F. and D. Nathans (1987).

- Expression of a set of growth-related immediate early genes in Balb/c 3T3 cells: Co-ordinate regulation with c-fos or c-myc. PNAS 84 p1182-1186.
60. Almendral, J. E., D. Sommer, H. McDonald-Bravo, J. Burckhardt, J. Perera and R. Bravo (1988).
Complexity of the early genetic response to growth factors in mouse fibroblasts.
Mol. Cell. Biol. 8 p2140-2148.
61. Distel, R. J., H.-S. Ro, B. S. Rosen, D. L. Groves and B. M. Spiegelman (1987).
Nucleoprotein complexes that regulate gene expression in adipocyte differentiation: Direct participation of c-fos.
Cell 49 p835-844.
62. Franza, B. R., Jr, F. J. Rauscher III, S. F. Josephs and T. Curran (1988).
The fos complex and fos-related antigens recognize sequence elements that contain AP-1 binding sites.
Science 239 p1150-1153.
63. Angel, P., E. A. Allegretto, S. T. Okino, K. Hattori, W. J. Boyle, T. Hunter and M. Karin (1988).
Oncogene jun encodes a sequence-specific trans-activator similar to AP-1.
Nature 332 p166-171.
64. Sukhatme, V. P., X. Cao, L. C. Chang, C.-H. Tsai-Morris, D. Stamenkovich, P. C. P. Ferreira, D. R. Cohen, S. A. Edwards, T. B. Shows, T. Curran, M. M. Le Beau, and E. D. Adamson (1988).
A zinc finger-encoding gene co-regulated with c-fos during growth and differentiation, and after cellular depolarization.

- Cell 53 p37-43.
65. Curran, T., W. P. MacConnell, F. van Straaten and I. M. Verma (1983).
Structure of the FBJ murine osteosarcoma virus genome: molecular cloning of its associated helper virus and the cellular homolog of the v-fos gene from mouse and human cells.
Mol. Cell. Biol. 3 p914-921.
66. Curran, T., A. D. Miller, L. Zokas and I. M. Verma (1984).
Viral and cellular fos proteins: A comparative analysis.
Cell 36 p259-268.
67. Maki, Y., T. J. Bos, M. Starbuck and P. K. Vogt (1987).
Avian sarcoma virus 17 carries the jun oncogene.
PNAS 84 p2848-2852.
68. Ryder, K., L. F. Lau and D. Nathans (1988).
A gene activated by growth factors is related to the oncogene v-jun.
PNAS 85 p1487-1491.
69. Ryder, K., A. Lanahan, E. Perez-Albuerne and D. Nathans (1989).
Jun-D: A third member of the jun gene family.
PNAS 86 p1500-1503.
70. Rauscher, F. J. III, D. R. Cohen, T. Curran, T. J. Bos, P. K. Vogt, D. Bohmann, R. Tjian, and B. R. Franza, Jr. (1988)
Fos associated protein p39 is the product of the jun proto-oncogene.
Science 240 p1010-1016.
71. Chiu, R., W. J. Boyle, J. Meek, T. Smeal, T. Hunter and M. Karin (1988).
The c-fos protein interacts with c-jun/AP-1 to stimulate transcription of AP-1 responsive genes.
Cell 54 p541-552.

72. Sassone-Corsi, P., W. W. Lamph, M. Kamps and I. M. Verma (1988).
Fos-associated cellular p39 is related to nuclear transcription
factor AP-1.
Cell 54 p553-560. <
73. Imler, J. L., C. Schatz, C. Wasylyk, B. Chatton and B. Wasylyk
(1988).
A Harvey-ras responsive transcription is also responsive to a tumour
promoter and to serum.
Nature 332 p275-278.
74. Linzer, D. I. H. and E. L. Wilder (1987).
Control of proliferin gene expression in serum-stimulated mouse
cells.
Mol. Cell. Biol. 7 p2080-2086.
75. Edwards, D. R., P. Waterhouse, M. L. Holman and D. T. Denhardt
(1986).
A growth responsive gene (16C8) in normal mouse fibroblasts
homologous to a human collagenase inhibitor with erythroid-
potentiating activity: evidence for inducible and constitutive
transcripts.
Nucleic Acids Res. 14 p8863-8878 (1986).
76. Deininger, P. L. (1989)
SINEs: short interspersed repeated DNA elements in higher
eukaryotes.
In Mobile DNA (D. E. Berg and M. M. Howe Eds.).
77. Gaichenhaus, N. and F. Cuzin (1987).
A role for ID repetitive sequences in growth- and transformation-
dependent regulation of gene expression in rat fibroblasts.
Cell 50 p1081-1089.

78. Hutchinson, C. A. III, S. C. Hardies, D. D. Loeb, W. R. Shehee and M. E. Edgell (1989).
LINEs and related retroposons: long interspersed sequences in the eukaryotic genome.
In Mobile DNA (D. E. Berg and M. M. Howe Eds.).
79. Fanning, T. G. and M. Singer (1987).
The Line-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins.
Nucleic Acids Res. 15 p2251-2260.
80. Coffin, J. (1982).
Structure of retroviral genome.
In Molecular biology of tumour viruses: RNA tumour viruses. (R. Weiss, N. Teich, H. Varmus and J. Coffin Eds.)
Cold Spring Harbor Laboratory.
81. Varmus, H. and R. Swanstrom (1982).
Replication of retroviruses.
In Molecular biology of tumour viruses: RNA tumour viruses. (R. Weiss, N. Teich, H. Varmus and J. Coffin Eds.)
Cold Spring Harbor Laboratory.
82. Reddy, E. P., M. J. Smith and S. A. Aaronson (1981).
Complete nucleotide sequence and organisation of the Moloney murine sarcoma virus genome.
Science 214 p445-450.
83. Levinson, B., G. Khoury, G. Vande Woude and P. Gruss (1982).
Activation of SV40 genome by 72bp tandem repeats of Moloney sarcoma virus.
Nature 295 p568-572.
84. Laimins, L. A., G. Khoury, C. Gorman, B. Howard and P. Gruss (1982).

- Host specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus.
PNAS 79 p6453-6457.
85. Laimins, L. A., P. Gruss, R. Pozzatti and G. Khoury (1984).
Characterisation of enhancer elements in the long terminal repeat of Moloney murine sarcoma virus.
J. Virol. 49 p183-189.
86. Miksicek, R., A. Heber, W. Schmid, U. Danesch, G. Posseckert, M. Beato and Gunter Schütz (1986).
Glucocorticoid responsiveness of the transcriptional enhancer of the Moloney murine sarcoma virus.
Cell 46 p283-290.
87. Gorman, C. M., P. W. J. Rigby and D. P. Lane (1985).
Negative regulation of viral enhancers in undifferentiated embryonic stem cells.
Cell 42 p519-526.
88. Nusse, R. (1986).
The activation of cellular oncogenes by retroviral insertion.
Trends Genet. 2 p244-247.
89. Hayward, W. S., B. G. Neel, and S. M. Astrin (1981).
Activation of a cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis.
Nature 290 p475-480.
90. Finnegan, D. J. and D. H. Fawcett (1986).
Transposable elements in *Drosophila melanogaster*.
Oxford Surveys on Eukaryotic Genes 3 p1-62.
91. Kingsman, A. J. and S. M. Kingsman (1988).
Ty: A retroelement moving forward.

- Cell 53 p333-335.
92. Boeke, J. D. (1989).
Transposable elements in *Saccharomyces cerevisiae*.
In Mobile DNA (D. E. Berg and M. M. Howe Eds.)
93. Kuff, E. L. and K. K. Leuders (1988).
The intracisternal A-particle gene family: structure and functional aspects.
Adv. Cancer Res. 51 p183-276.
94. Flavell, A. J., R. Levis, M. A. Simon and G. M. Rubin (1981).
The 5' termini of RNAs encoded by the transposable element copia.
Nucleic Acids Res. 9 p6279-6291.
95. Sneddon, A. and A. J. Flavell (1989).
The transcriptional control regions of copia retrotransposon.
Nucleic Acids Res. 17 p4025-4035.
96. Fulton, A. M., P. D. Rathjen, S. M. Kingsman and A. J. Kingsman (1988).
Upstream and downstream transcriptional control signals in the yeast retrotransposon, Ty.
Nucleic Acids Res. 16 p5439-5458.
97. Falzon, M. and E. L. Kuff (1988).
Multiple protein-binding sites in an intracisternal A particle long terminal repeat.
J. Virol. 62 p4070-4077.
98. Mount, S. M. and G. M. Rubin (1985).
Complete nucleotide sequence of the *Drosophila* transposable element copia: homology between copia and retroviral proteins.
Mol. Cell. Biol. 5 p1630-1638.
99. Clare, J. and P. Farabaugh (1985).

- Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression.
PNAS 82 p2829-2833.
100. Adams, S. E., J. Mellor, K. Gull, R. B. Sim, M. F. Tuite, S. M. Kingsman and A. J. Kingsman (1987).
The functions and relationships of Ty-VLP proteins in yeast reflect those of mammalian retroviral proteins.
Cell 49 p111-119.
101. Boeke, J. D., D. J. Garfinkel, C. A. Styles and G. R. Fink (1985).
Ty elements transpose through an RNA intermediate.
Cell 40 p491-500.
102. Kingsman, A. J. and S. M. Kingsman (1988).
Ty: a retroelement moving forward.
Cell 53 p333-335.
103. Temin, H. M. (1980).
Origin of retroviruses from cellular mobile genetic elements.
Cell 21 p599-600.
104. Errede, B., M. Company and C. A. Hutchison III (1987).
Tyl sequence with enhancer and mating-type-dependent regulatory activities.
Mol. Cell. Biol. 7 p258-265.
105. Horowitz, M., S. Luria, G. Rechavi and D. Givol (1984).
Mechanism of activation of the mouse c-mos oncogene by the LTR of an intracisternal A-particle gene.
EMBO 3 p2937-2941.
106. Duhrsen, U., J. Stahl and N. M. Gough (1990).
In vivo transformation of a factor dependent haemopoietic cell line: role of IAP transposition for growth factor gene activation.

- EMBO 2 p1087-1096.
107. Stavenhagen, J. B. and D. M. Robins (1988).
An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene.
Cell 55 p247-254.
108. Finnegan, D. J. (1989).
Eukaryotic transposable elements and genome evolution.
Trends Genet. 5 p103-107.
109. Baltimore, D. (1985).
Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome.
Cell 40 p481-482.
110. Keshet, E. and A. Itin (1982).
Patterns of genomic distribution and sequence heterogeneity of a murine "retrovirus-like" multigene family.
J. Virol. 43 p50-58.
111. Keshet, E. and Y. Shaul (1981).
Terminal direct repeats in a retrovirus-like repeated gene family.
Nature 289 p83-85.
112. Carter, A. T., J. D. Norton and R. J. Avery (1983).
A novel approach to cloning transcriptionally active retrovirus-like elements from mouse cells.
Nucleic Acids Res. 11 p6243-6254.
113. Duesberg, P. H. and E. M. Scolnick (1977).
Murine leukaemia viruses containing a 30S subunit of unknown biological activity, in addition to the 38S subunit of the viral genome.
Virology 83 p211-216.

114. Howk, R. S., D. H. Troxler, D. Lowy, P. H. Duesberg, and E. M. Scolnick (1978).
Identification of a 30S RNA with properties of a defective type C virus in murine cells.
J. Virol. 25 p115-123.
115. Besmer, P., U. Olshevsky, D. Baltimore, D. Dolberg and H. Fan (1979).
Virus-like 30S RNA in mouse cells.
J. Virol. 29 p1168-1176.
116. Sherwin, S. A., U. R. Rapp, R. E. Benveniste, A. Sen and G. J. Todaro (1978).
Rescue of endogenous 30S retroviral sequences from mouse cells by baboon type C virus.
J. Virol. 26 p257-264.
117. Hodgson, C. P., P. K. Elder, T. Ono, D. N. Foster and M. J. Getz (1983).
Structure and expression of mouse VL30 genes.
Mol. Cell. Biol. 3 p2221-2231.
118. Norton, J. D., J. Connor and R. J. Avery (1984).
Unusual LTR sequence of a retrovirus transmissible mouse (VL30) genetic element: Identification of functional domains.
Nucleic Acids Res. 12 p3445-3460.
119. Itin, A. and E. Keshet (1986)
Diverse long terminal repeats are associated with murine retrovirus-like (VL30) elements.
Mol. Cell. Biol. 6 p1276-1282.
120. Itin, A. and E. Keshet (1983).
Nucleotide sequence analysis of the LTR of murine virus-like (VL30)

- and its adjacent sequences: Resemblance to retrovirus proviruses.
J. Virol. 47 p656-659.
121. Itin, A. and E. Keshet (1985).
Primer binding sites corresponding to several tRNA species are present in DNAs of different members of the same retrovirus-like gene family.
J. Virol. 54 p236-239.
122. Giri, C. P., C. P. Hodgson, P. K. Elder, M. G. Courtney and M. J. Getz (1983).
Discrete region of sequence homology between cloned rodent VL30 genetic elements and AKV-related MuLV provirus genomes.
Nucleic Acids Res. 11 p305-319.
123. Adams, S.E., P. D. Rathjen, C. A. Stanway, S. M. Fulton, M. H. Malim, W. Wilson, J. Ogden, L. King, S. M. Kingsman and A. J. Kingsman (1988).
Complete nucleotide sequence of a mouse VL30 retro-element.
Mol. Cell. Biol. 8 p2989-2998.
124. Carter, A. T., J. D. Norton, Y. Gibson and R. J. Avery (1986).
Expression and transmission of a rodent retrovirus-like (VL30) gene family.
J. Mol. Biol. 188 p105-108.
125. Singh, K., S. Saragosti and M. Botchan (1985).
Isolation of cellular genes differentially expressed in mouse NIH 3T3 cells and a Simian Virus 40-transformed derivative: Growth specific expression of VL30 genes.
Mol. Cell. Biol. 5 p2590-2598.
126. Courtney, M. G., L. J. Schmidt and M. J. Getz (1982).
Organisation and expression of endogenous virus-like (VL30) DNA

sequences in nontransformed and chemically transformed mouse embryo cells.

Cancer Res. 42 p569-576.

127. Foster, D. N., L. J. Schmidt, C. P. Hodgson, H. L. Moses and M. J. Getz (1982).

Polyadenylated RNA complementary to a mouse retrovirus-like multigene family is rapidly and specifically induced by epidermal growth factor stimulation of quiescent cells.

PNAS 79 p7317-7321.

128. Rodland, K. D., S. F. Jue and B. E. Magun (1986)

Regulation of gene expression by activators of protein kinase C.

J. Biol. Chem. 261 p5029-5033.

129. Harrigan, M. T., G. Baughman, N. F. Campbell and S. Bourgeois (1989).

Isolation and characterization of glucocorticoid- and cyclic AMP-induced genes in T lymphocytes.

Molec. Cell. Biol. 9 p3438-3446.

130. Rotman, G., A. Itin and E. Keshet (1986).

Promoter and enhancer activities of LTRs associated with cellular retrovirus-like (VL30) elements.

Nucleic Acids Res. 14 p645-658.

131. Morris, A. G., C. Clegg, J. Jones, B. Rodgers and R. Avery (1980).

The isolation and characterisation of a clonally related series of murine retrovirus-infected mouse cells.

J. Gen. Virol. 49 p105-113.

132. Hidaka, H., M. Inagaki, S. Kawamoto and Y. Sasaki (1984).

Isoquinolinesulfonamides, novel and potent inhibitors of cyclic nucleotide dependent protein kinase and protein kinase C.

- Biochemistry 23 p5036-5041.
133. Maniatis, T., E. F. Fritsch and J. Sambrook (1982).
Molecular Cloning: A Laboratory Manual.
Cold Spring Harbor Laboratory Press, Cold Spring Harbor,
New York.
134. Langridge J., P. Langridge and P. L. Berquist (1980).
Extraction of nucleic acids from agarose gels.
Analyt. Biochem. 103 p264-271.
135. Feinberg, A. P. and B. Vogelstein (1983).
Techniques for radiolabelling DNA restriction endonuclease fragments
to high specific activity.
Analyt. Biochem. 132 p6-11.
136. Messing, J. and J. Vieira (1982).
A new pair of M13 vectors for selecting either strand of double
digest restriction fragments.
Gene 19 p269-276.
137. Chirgwin, J. M., A. E. Przbyla, R. J. MacDonald and W. J. Rutter
(1979).
Isolation of biologically active ribonucleic acid from sources
enriched in ribonuclease.
Biochemistry 18 p5294-5299.
138. Lehrach, H., D. Diamond, J. M. Wozney and H. Boedtker (1977).
RNA molecular weight determination by gel electrophoresis under
denaturing conditions, a critical reexamination.
Biochemistry 16 p4743-4751.
139. Mason, I. J., D. Murphy, M. Munke, U. Franke, R. W. Elliot and B.
L. M. Hogan (1985).
Developmental and transformation-sensitive expression of the Sparc

- gene on mouse chromosome 11.
EMBO 5 1831-1837.
140. Graham, F. L. and A. J. Van Der Eb (1973).
A new technique for the assay of infectivity of human adenovirus 5
DNA.
Virology 52 p456-467.
141. Weber, F., J. de Villiers and W. Schaffner (1984).
An SV40 "enhancer trap" incorporates exogenous enhancers or
generates enhancers from its own sequences.
Cell 36 p983-992.
142. Gorman, C. M., L. F. Moffat and B. H. Howard (1982).
Recombinant genomes which express chloramphenicol acetyltransferase
in mammalian cells.
Mol. Cell. Biol. 2 p1044-1051.
143. Luckow, B. and G. Schütz (1987)
CAT constructions with multiple unique restriction sites for the
functional analysis of eukaryotic promoters and regulatory elements.
Nucleic Acids Res. 15 p5490.
144. Southern, P. and P. Berg (1982).
Transformation of mammalian cells to antibiotic resistance with a
bacterial gene under control of the SV40 early region promoter.
J. Mol. Appl. Genet. 1 p327-330.
145. Norton, J. D. and B. L. M. Hogan (1988).
Temporal and tissue-specific expression of distinct retrovirus-like
(VL30) elements during mouse development.
Dev. Biol. 125 p226-228.
146. De Larco, J. E. and G. J. Todaro (1978).
Growth factors from murine sarcoma virus-transformed cells.

Proc. Nat. Acad. Sci. USA 75 p4001-4005.

147. Norton, J. D., F. Cook, P. C. Roberts, J. P. Clewley and R. J. Avery (1984).

Expression of Kirsten murine sarcoma virus in transformed nonproducer and revertant NIH/3T3 cells: evidence for cell mediated resistance to a viral oncogene in phenotypic reversion.

J. Virol. 50 p439-444.

148. Murray, M. J., J. M. Cunningham, L. F. Parada, F. Dautry, P. Lebowitz and R. A. Weinberg (1983).

The HL-60 transforming sequence: a ras oncogene coexisting with altered myc genes in hematopoietic tumors.

Cell 33 p749-757.

149. Struhl, K. (1986).

Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms.

Molec. Cell. Biol. 6 p3847-3853 (1986).

150. Simon, M. C., T. M. Fisch, B. J. Benecke, J. R. Nevins and N. Heintz (1988).

Definition of multiple, functional distinct TATA elements, one of which is a target in the hsp 70 promoter for E1A regulation.

Cell 52 p723-729.

151. Carter, A. T., J. D. Norton, and R. J. Avery (1988).

The genomic DNA organisation and evolution of a retrovirus-transmissible family of mouse (VL30) genetic elements.

Biochem. Biophys. Acta 951 p130-138.

152. Schmidt, M., T. Wirth, B. Kröbger and I. Horak (1985).

Structure and genomic organization of murine retrovirus-related DNA sequences (MuRRS).

- Nucleic Acid Res. 13 p3461-3470.
153. Matrisian, L. M., N. Glaichenhaus, M-C. Gesnal and R. Breathnach (1985).
Epidermal growth factor and oncogenes induce transcription of the same cellular mRNA in rat fibroblasts.
EMBO 4 p1435-1440.
154. D. I. H. Linzer and E. L. Wilder (1987).
Control of proliferin expression in serum-stimulated mouse cells.
Molec. Cell. Biol. 7 p2080-2086.
155. D. I. H. Linzer and J. C. Mordacq (1987).
Transcriptional regulation of proliferin gene expression in response to serum in transfected mouse cells.
EMBO 6 p2281-2288.
156. Rodland, K. D., A. M. C. Brown and B. E. Magun (1987).
Individual mouse VL30 elements transferred to rat cells by viral pseudotypes retain their responsiveness to activators of protein kinase C.
Molec. Cell. Biol. 7 p2296-2298.
157. Owen, R. D. and M. C. Ostrowski (1987).
Rapid and selective alterations in the expression of cellular genes accompany conditional transcription of Ha-v-ras in NIH 3T3 cells.
Molec. Cell. Biol. 7 p2512-2520.
158. Todaro, G. J., C. Fryling and J. E. De Larco (1980).
Transforming growth factors produced by certain human tumour cells: polypeptides that interact with epidermal growth factor receptors.
Proc. Natl. Acad. Sci. USA 77 p5258-5262.
159. Owen, R. D., D. M. Bortner and M. C. Ostrowski (1990).
ras gene activation of a VL30 transcriptional element is linked to

transformation.

Molec. Cell. Biol. 10 p1-9.

160. Liao, X-B., J. J. Clare and P. J. Farabaugh (1987).

The upstream activation site of a Ty2 element of yeast is necessary but not sufficient to promote maximal transcription of the element.

Proc. Nat. Acad. Sci. 84 p8520-8524.

161. Sassone-Corsi, P. and I. M. Verma (1987).

Modulation of c-fos gene trnascrption by negative and positive cellular factors.

Nature 326 p507-510.

162. Schneider, C., R. M. King and L. Philipson. (1988).

Genes specifically expressed at growth arrest of mammalian cells.

Cell 54 p787-793.

163. Miska, D. and H. B. Bosmann (1980).

Existence of an upper limit of the prereplicative period in confluent cultures of C3H/10T1/2 cells.

Biochem. Biophys. Res. Comm. 93 p1140-1145.

164. Karin, M. (1989).

Complexities of gene regulation by cAMP.

Trends Genet. 5 p65-67.

165. Downward, J. (1990).

The ras superfamily of small GTP-binding proteins.

Trends Biochem. 15 p469-472.

166. Hall, A. (1990).

The cellular functions of small GTP-binding proteins.

Science 249 p635-640.

167. Bourne, H. R., D. A. Sanders and F. McCormick (1991).

The GTPase superfamily: conserved structure and molecular mechanism.

Nature 349 p117-127.

168. Wolfman, A. and I. G. Macara (1990).

A cytosolic protein catalyzes the release of GDP from p21^{ras}.

Science 248 p67-69.

PUBLICATIONS

Eaton, L. and Norton, J.D. (1988)

Growth Factor Regulation of Retrotransposon VL30 Gene Transcription
(abstract)

6th Tenovus Symposium 'Eukaryotic Gene Expression'

23-25 March 1988

Eaton, L. and Norton, J.D. (1990)

Independent Regulation of Mouse VL30 Retrotransposon Expression in
Response to Serum and Oncogenic Cell Transformation

Nucleic Acids Research 18 p2069-2077.