SMN: POTENTIAL ROLE IN ALTERNATIVE mRNA SPLICING.

The Expression and Function of the SMN Protein

David Andrew Horn

A thesis submitted for the degree of Doctor of Philosophy

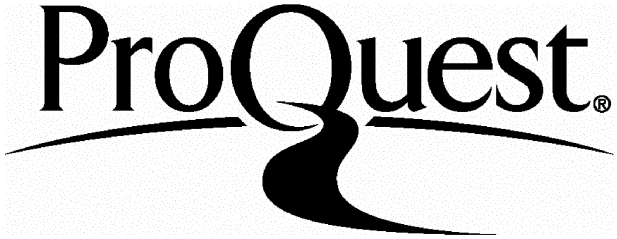
Medical Molecular Biology Unit

Department of Biochemistry

University College and Middlesex School of Medicine

University College London

January 1993



Acknowledgements

I would especially like to thank Prof. David Latchman for providing excellent supervision throughout the course of this research. I would also like to thank all of the members of the Medical Molecular Biology Unit who have provided advice and technical assistance. In particular I would like to thank John Estridge and Dr. Karen Lillycrop. I would also like to thank Dr. Keith Howard and Carolina Mailhos for providing comments on the text of this thesis and I would particularly like to thank Dr. Keith Howard for useful discussions concerning my Ph.D. work and also my future work.

This work was supported by a S.E.R.C. studentship.

Declaration

All of the work presented in this thesis is the work of David A. Horn. Contributions by other researchers to any of the work presented is listed below.

1). Dr. Georgio Terenghi of Hammersmith Hospital, London carried out *in-situ* hybridisation work seen in chapter 3.

2). Dr. Debbie Cumming of the National Heart and Lung Institute, London probed the Western blot seen in chapter 4. with the cTNT antiboby.

Abstract

The Sm proteins associate with small nuclear RNA (snRNA) molecules to form small nuclear ribonucleoprotein particles (snRNPs) which are essential for pre mRNA splicing. A recently identified Sm protein, SmN is expressed in a tissue specific manner and is closely related to the constitutively expressed SmB protein. SmN is abundant in neurons and is expressed at lower levels in cardiac muscle while being undetectable in other tissues; it is therefore the first example of a mammalian snRNP protein of this kind. Human, rat and mouse forms of the protein are 100% identical. Mutation at the SmN locus has also been implicated in Prader-Willi syndrome suggesting an important function for this protein.

Expression of SmN has been characterised in a number of neuronal and non-neuronal cell lines, in EC cells during differentiation and in tissues during embryonic development using the KSm5 antibody in conjunction with Western blotting. The KSm5 antibody has also been used to examine Sm protein distribution by immunoflourescence, and RNA probes have been used to examine SmN and SmB mRNA expression and SmN mRNA distribution in the mouse brain. The distribution of the closely related SmN, B and B' proteins has also been defined in human tissues.

SmN distribution, and its association with the splicing machinery has raised the possibility that it is involved in alternative splicing. In particular, cells and tissues which express SmN have been shown to have the ability to follow an alternative splicing pathway resulting in the production of the CGRP mRNA from the primary transcript of the CALC-I gene.

Characterisation of SmN expression in a number of cells and tissues allowed an investigation into the possible role of SmN in alternative splicing. A number of cell lines ectopically expressing SmN were also constructed in order to examine the role of SmN. Using a number of the cell lines and tissues characterised above, as well as the stable cell lines, the putatative role of SmN in alternatively splicing has been examined. Using quantitative PCR assays it appears that SmN is neither necessary nor sufficient to guide alternative splicing of a number of mRNAs which appeared to be candidates for regulation by SmN, from the literature, including CALC-I, NCAM (VASE exon) and c-src. A pair of CALC-I gene constructs have also been used to identify a *cis*-acting sequence not originally thought to be involved in CALC-I alternative splicing.

The data presented here and other recent results have been used to speculate on other possible roles for SmN. The SmN expressing stable cell lines provide a model in which to examine SmN functions further and the ND neuronal cell lines provide a good model to examine CALC-I alternative splicing.

Abbreviations

APS	Ammonium persulphate
bp	Base pairs
BSA	Bovine serum albumin
cAMP	N6,2'-O-Dibutyryladenosine 3':5'- cyclic monophosphate
CAT	Chloramphenicol acetyl transferase
cDNA	Complementary DNA
CGRP	Calcitonin gene-related peptide
CMV IE 94	Cytomegalovirus immediate early 94kd
CNS	Central nervous system
c.p.m.	Counts per minute
cTNT	Cardiac troponin-T
CTRP's	Calf thymus random primers
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethylsulphoxide
DNA	Deoxynucleic acid
DRG	Dorsal root ganglion
DTT	Dithiothreitol
EC	Embryonal carcinoma
ECL	Enhanced chemiluminescence
EDTA	Diaminoethanetetra-acetic acid, disodium salt
ES	Embryonal stem cell
FCS	Foetal calf serum
HBS	Hepes buffered saline
Hepes	N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid]
hnRNP	Heterogeneous nuclear RNP
kb	Kilobase

kd	Kilodalton
LGT	Low gelling temperature
LTR	Long terminal repeat
MMTV	Mouse mammary tumour virus
MoMLV	Moloney murine leukemia virus
MOPS	3-(N-morpholino) propanesulphonic acid
MPSV	Murine proliferative sarcoma virus
MWt	Molecular weight
N-CAM	Neuronal cell adhesion molecule
ND	Rat dorsal root ganglion derived cell lines
NGF	Nerve growth factor
pA	Polyadenylation signal
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol 6000
PNS	Peripheral nervous system
pre-mRNA	Precursor messenger RNA
PRP	pre-mRNA processing
PVP	Polyvinylpyrrolidone
RA	All-trans-retinoic acid
RNA	Ribonucleic acid
RNP	Ribonucleoprotein particle
r.p.m.	Revolutions per minute
rRNA	Ribosomal RNA
RSV	Rous sarcoma virus
RT	Reverse transcription
SDS	Sodium dodecyl sulphate
SLE	Systemic lupus erythematosis

Sm	Smith (patient sera)
snRNA	Small nuclear RNA
snRNP	Small nuclear RNP
SSC	Standard saline citrate
TAE	Tris-acetate EDTA buffer
TBE	Tris-borate EDTA buffer
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
Tm	Melting temperature
Tween 20	Polyoxyethylene-sorbitan monolaurate
UCMSM	University College & Middlesex School of Medecine
UV	Ultraviolet
X-gal	5-bromo-4-chloro-3-indolyl-*B*-galactopyranoside

Publications

1). Suburo, A. M; Wheatley, S. C; Horn, D. A; Gibson, S. J; Jahn, R; Fischer-Colbrie, R; Wood, J. N; Latchman, D. S. & Polak, J. M.
Intracellular redistribution of neuropeptides & secretory proteins during differentiation of neuronal cell lines.
(1992) Neuroscience 46 881-889.

2). Horn, D. A. & Latchman, D. S.
Neuronal expression of the tissue specific splicing protein SmN and its possible role in pre-mRNA splicing.
(1992) Biochem. Soc. Trans. 20 266.

3). Horn, D. A; Suburo, A; Terenghi, G; Hudson, L; Polak, J. M. & Latchman, D.S.
Expression of the tissue specific splicing protein in neuronal cell lines and in regions of the brain with different splicing capacities.
(1992) Mol. Brain Res. 16 13-19.

4). Wheatley, S. C; Suburo, A. M; Horn, D. A; Vucicevic, V; Terenghi, G; Polak, J. M. & Latchman, D. S.
Redistribution of secretory granule components precedes that of synaptic vesicle proteins during differentiation of a neuronal cell line in serum-free medium.
(1992) Neuroscience 51 575-582.

5). Grimaldi, K; Horn, D. A; Hudson, L. D; Terenghi, G; Polak, J. M. & Latchman, D.S.
Expression of the SmN splicing protein is regulated during development of the brain but not during heart development.
In press, Dev. Biol.

6). Gerrelli, D; Grimaldi, K; Horn, D. A; Mahadeva, U; Sharpe, N. & Latchman, D.S.

The cardiac form of the tissue specific splicing protein SmN is identical to the brain and embryonic forms of the protein.

In press, J. Mol. Cell. Cardio.

7). Horn, D. A. & Latchman, D. S.

The tissue specific SmN protein does not influence the alternative splicing of endogenous NCAM and c-src RNAs in transfected 3T3 cells.

In press, Mol. Brain Res.

8). Horn, D. A. & Latchman, D. S.

Alternative splicing of the mRNA encoding the α subunits of the G_o GTP-binding protein during brain development and in neuronal cell lines.

Submitted, Neurosci. Lett.

9). Horn, D. A. & Latchman, D. S.

Regions of the CALC-I gene necessary for regulated alternative splicing.

In preparation.

Contents

CHAPTER 1 : Introduction

CHAPTER 2 : Materials & Methods

2.2. METHODS

CHAPTER 4 : Results

Alternative splicing assays

CHAPTER 5 : Results

Production of stable cell lines and subsequent splicing assays

CHAPTER 6 : Discussion

List of figures

List of tables

Plasmid maps and details

(for references see materials page 56)

CHAPTER 1

INTRODUCTION

1.1. Mammalian pre-mRNA splicing

1.1.1. Splicing

Mammalian pre-mRNA processing (see fig.1.1.) involves the removal of intervening sequences or introns from the primary transcripts of RNA polymerase II genes to allow joining of the coding regions or exons (for reviews see Sharp, 1989; Mattaj, 1990; Lamond, 1991a). This process known as splicing is achieved in eukaryotes in the nucleus. The cleavage and ligation reactions that lead to the generation of mature mRNA proceed within the spliceosome. The spliceosome contains multiple RNA and protein components (see table 1.1.). Most coding genes in higher eukaryotes are interrupted by at least one and usually several introns all of which must be removed to allow production of the functional protein product. Some genes may be located within introns (Viskochil et al., 1991). Parallels can be drawn between this process and self splicing (Jaquier, 1990) as well as *trans*-splicing (see 1.1.4.) (Konarska et al., 1985).

Only a few genes in *S. cerevisiae* contain one or two small introns, whereas most are intron-free (for review see Ruby & Abelson, 1991) in contrast to vertebrate genes which contain multiple introns which can vary in length from approximately 50 to well over 1000 bases. The first intron of c-abl is more than 200 kilobases in length. Vertebrate exons are generally between 50 and 300 bases in length. Exons as short as three bases have been described however (Hamshere et al., 1991).

Mammalian pre-mRNA splicing as well as yeast group II self splicing reactions (Peebles et al., 1986) generate a branched RNA molecule (Ruskin et al., 1984). The first cleavage/ligation reaction in mRNA splicing leads to the formation of this circular lariat RNA. 5' splice site cleavage allows ligation of the 5' terminal guanosine of the intron usually to an adenosine residue at the branch point, through a 2'-5' phosphodiester bond. The multicomponent complex which forms prior to the generation of this intermediate has a sedimentation velocity coefficient of 40-60s. This complex is termed the spliceosome. A second cleavage/ligation reaction involves 3' splice site cleavage, and ligation of the 5' exon to the 3' exon via a 3'-5' phosphodiester bond. This reaction leads to intron removal in the form of a lariat and the production of a mature or partially processed transcript (see fig.1.1.).

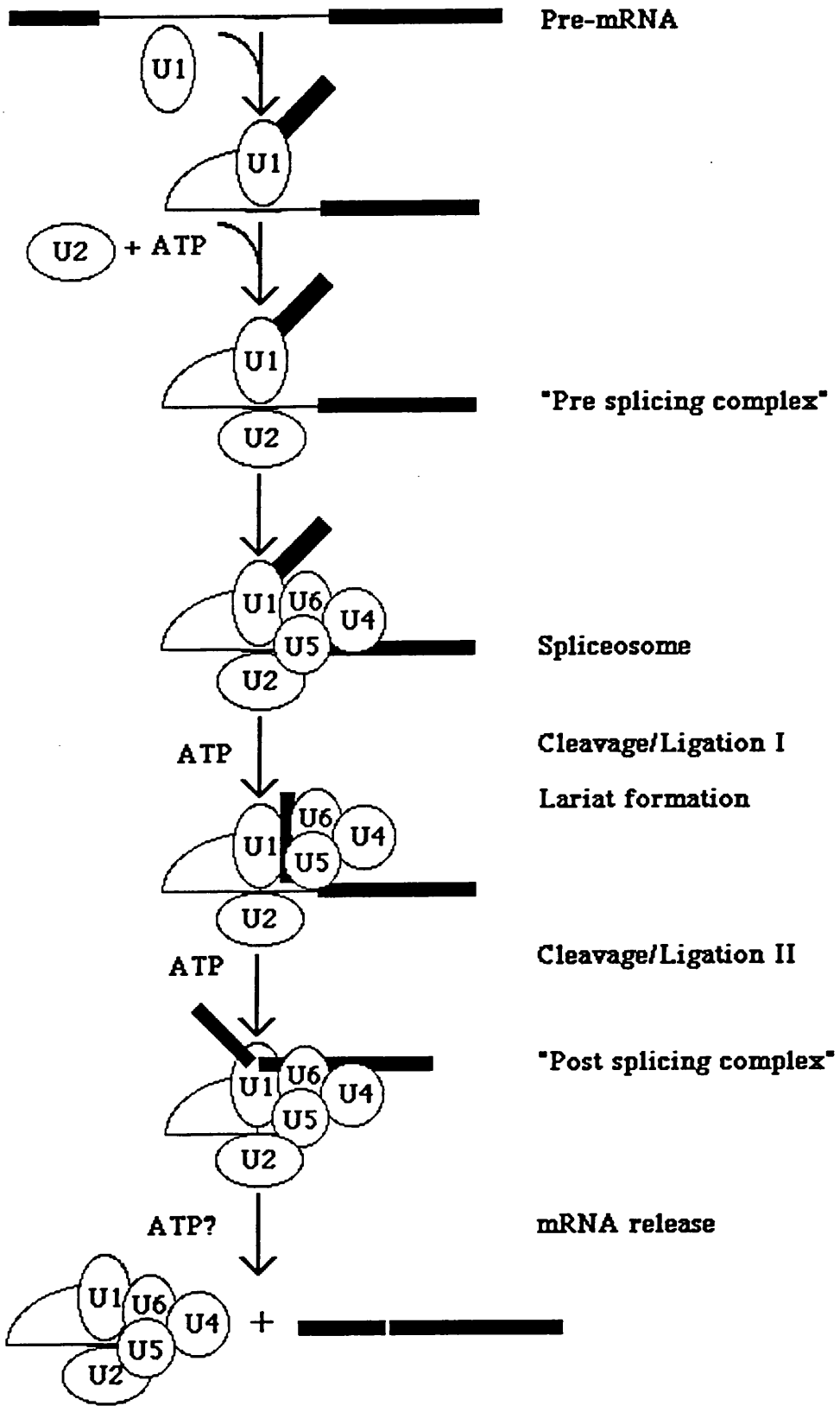


Figure 1.1. Model of the splicing pathway

The model of mammalian splicing shows removal of a single intron and ligation of two exons and the step-wise addition of snRNPs during spliceosome assembly. The branch point is indicated as a bold bar in the intron.

snRNA	U1	U2	U5	U4	U6
Size (nt)	164	187	116	145	106
Estimated abundance (copies/cell)	10^6	$6x10^5$	$2x10^5$	$2x10^5$	$4x10^5$
Cap structure	m3GpppA[a]	m3GpppA	m3GpppA	m3GpppA	mpppG[b]
Modified nucleosides[c]					
ribose methylation (2'-OMe)	3	10	5	2	8
base modification (m6A, m2G)	0	0	0	1	2
dual ribose + base modification	0	1	0	0	0
pseudouridine	2	13	3	3	3
Associated proteins					
SmG (9kd)	++[d]	++	++	++	
SmF (11kd)	+	+	+	+	
SmE (12kd)	+	+	+	+	
SmD3 (18kd)	+	+	+	+	
SmD2 (16.5kd)	+	+	+	+	
SmD1 (16kd)	+	+	+	+	
SmB (28kd)	+	+	+	+	
SmB' (29kd)	+	+	+	+	
C (22kd)	+				
A (34kd)	+				
54kd	+				
65kd	+				
70K (70kd)	+				
B" (28.5kd)		+			
A' (31kd)		+			
15kd			+		
40kd			+		
52kd			+		
11kd			+		
102kd			+		
116kd			+		
205kd			+	+	
220kd			+		
35kd				+	
43kd				+	
57kd				+	
71kd				+	
126kd				+	
20kd			+		
27kd			+		
60kd			++		
90kd			+		

Table 1.1. Characteristics of mammalian snRNP particles

a m3GpppA = 2,2,7 trimethyl Guanosine.

b mpppG = gamma-monomethyl phosphate ester.

c The nucleoside modifications listed may vary between different species and between different cell types.

d (++) indicates a double band on SDS-polyacrylamide gels.

Following the removal of all introns, the mature mRNA will subsequently be transported to the cytoplasm for translation (for review see Green, 1989).

Eukaryotic pre-mRNA molecules are modified at their 5' end with a monomethyl guanosine cap structure (m^7G) (Salditt-Georgieff et al., 1980) and at their 3' end with a 200-300 base poly(A) "tail". The poly(A) tail is added to the pre-mRNA by the action of a poly(A) polymerase, which has now been cloned from bovine and yeast cells (Raabe et al., 1991; Lingner et al., 1991) and an additional multisubunit complex (Takagaki et al., 1990). Poladenylation follows cleavage of the pre-mRNA downstream of the highly conserved polyadenylation signal (AAUAAA). The 5' and 3' end modifications appears to be to influence mRNA stability and the 5' cap is also necessary for translation. Generally the 5' cap structure is added to the pre-mRNA preceding polyadenylation (Gilmartin et al., 1988) and polyadenylation precedes splicing (Nevins & Darnell, 1978). The *in-vivo* substrate for splicing therefore is typically a capped and polyadenylated linear RNA. There are examples however where nascent transcripts are spliced so 3' cleavage and polyadenylation are not necessary for splicing to take place. The half life of introns within pre-mRNA can vary from a few seconds to 10-20 minutes. Following intron removal a lariat debranching enzyme leads to intron degradation.

1.1.2. *Cis*-acting sequences involved in splicing

There appear to be few conserved *cis*-acting sequences involved in mammalian pre-mRNA splicing. The 5' splice site consensus (CAG:GUAAGU) has a virtually invariant GU immediately downstream of the exon/intron boundary (Mount, 1982) and the 3' splice site consensus (PyPyPyPy-Py-CAG:G) has a virtually invariant AG preceded by a polypyrimidine tract, not found in yeast introns, immediately upstream of the intron/exon boundary. A weak consensus known as the branch point is the site of lariat formation (see above) (Ruskin et al., 1984) and is also found in mammalian introns located 18-37 bases upstream (Keller et al., 1981) of the 3' splice site. Complementarity between the branch point and the 5' splice site has in fact been suggested to aid lariat formation (Konarska et al., 1985). The 5' and 3' splice site and branch point sequences are all accurately and

efficiently processed, usually in a defined order in transcripts containing multiple introns (for review see Aebi and Weissmann, 1987).

'Cryptic' splice sites, which closely resemble normal splice sites, are efficiently used in the absense of the normal splice sites but are never used in their presence. Mutations of the normal branch point sequence, both *in-vitro* and *in-vivo* (Ruskin et al., 1985) also result in the activation of one or more cryptic branch point sequences. Naturally occuring mutations affecting mammalian splicing have been found only in the first six bases of the 5' splice site sequences and the last two bases of the 3' splice site sequences (Treisman et al., 1983; Rogers, 1984) while none have been found at the mammalian branch point. The branch point consensus in higher eukaryotic introns (Keller & Noon,1984) is less conserved than the invariant yeast UACUAAC (underlined A represents base at which branch is formed). The 5' and 3' splice sites, as well as the branch point sequence, are stringently conserved in yeast (Langford & Gallwitz, 1983; Pikielny et al., 1983). A number of mammalian branch point nucleotides are necessary for splicing to take place however (Reed & Maniatis, 1988). The above elements are all required for efficient splicing but are not sufficient to ensure usage of the correct pairs of 5' and 3' splice sites. This suggests that the sequence requirements for splicing in higher eukaryotes are not as stringent as in yeast with mutations in the natural splice site, or in distal sites resembling the consensus, allowing the activation of 'cryptic' sites. Mammalian introns therefore may be largely redundant unlike group II self splicing introns (Peebles et al., 1986) in which the intron acts as the splicing catalyst. This is reflected by the fact that higher eukaryotic introns are rarely accurately spliced in fungal or plant systems.

Mutational analysis of the rabbit β-globin gene showed that a minimum intron length of about 60 bases was required for efficient splicing with essentially random sequences except for the 5' most six bases and the 3' most 24 bases (Wieringa et al., 1983). Naturally occuring mutations in the splice site consensus sequences are in fact responsible for a number of disease phenotypes such as β-thalassemia (Treisman et al., 1982), haemophilia B (Rees et al., 1985) and phenylketonuria (DiLella et al., 1986).

1.1.3. Mechanism of splicing

It is unknown whether RNA or protein in the spliceosome, or a combination of both, catalyse the splicing reactions or whether the spliceosome simply aligns the substrate for "self splicing". The former is most likely in mammalian splicing due to the lack of conservation in mammalian introns. A number of catalytic RNAs (for review see Cech & Bass, 1986) or ribozymes do clearly exist (Lamond & Gibson, 1990), so it is possible that the proteins of the spliceosome are responsible for numerous structural and positional functions rather than being directly involved in intron removal. It has been suggested that nuclear pre-mRNA processing may be catalysed by RNA as in the case of self splicing but that the RNA catalyst acts in *trans* in the form of the snRNPs (Jaquier, 1990) (see 1.3.). U6, the most highly conserved of the snRNPs has been proposed to constitute part of a catalytic centre involved in splicing (Vankan et al., 1992).

Most of the details of mammalian pre-mRNA splicing have been elucidated using *in-vitro* whole cell or nuclear extracts of cultured HeLa cells using specific pre-mRNA transcripts produced by *in-vitro* transcription of DNA templates. These *in-vitro* splicing systems require ATP and magnesium ions (see fig.1.1.). The above lariat intermediates have however also been identified *in-vivo* in the splicing of the rabbit β-globin gene (Zeitlin & Efstratiadis, 1984). Despite the weak splicing consensus sequences found in mammalian genes the 5' and 3' splice sites and branch point sequences are precisely recognised to yield distinct products. Exon sequences appear to play an important role in the use of the adjacent 5' and 3' splice sites. 5' and 3' splice site selection is also affected by splice site proximity (Reed & Maniatis, 1986). The processing of pre-mRNA with multiple introns can follow highly preferred pathways (Lang & Spritz, 1987) and the rate of excision can be influenced by the presence or absence and proximity of other introns in the pre-mRNA (Reed & Maniatis, 1986).

Proximal sites are generally spliced together suggesting a one dimensional intron scanning model, after the binding of splicing factors at the splice sites, to the closest compatible splice site (Sharp, 1981; Reed & Maniatis, 1987). *Trans*-splicing (Konarska et al., 1985) between independent RNA molecules (see 1.1.4.) and mutually exclusive patterns of alternative splicing (see fig.1.2.) (Medford et al., 1984) provide evidence

against a one dimensional intron scanning model. Also intron scanning does not appear to play a dominant role in SV40 early (Manley et al., 1987) or dihydrofolate reductase (Mitchell et al., 1986) splice site selection. An alternative mechanism of splice site selection therefore may involve the assembly of splicing complexes on all potential splice sites allowing subsequent selection of distal sites (Kuhne et al., 1983).

It appears likely that splice site selection occurs by combinations of both of the above in individual cases. Relative stability of complexes, intron length, secondary structure (Munroe, 1984) and other RNA binding proteins or snRNAs or snRNPs can then exert an additional influence on the splicing pathway to be followed by altering the relative strength of potential splice sites. Individual elements within an intron are probably recognised seperately, before interaction between *trans*-acting components results in their assembly into a spliceosome. U2 snRNP (see 1.3.) can in fact bind to the branch point on substrate lacking a 5' splice site suggesting that intron elements may be assembled independently into partial complexes which are subsequently brought into proximity (Black et al., 1985). Some *cis*-acting elements may only therefore be recognised by reference to their position relative to other *cis*-acting elements. Exon scanning does appear to be involved in vertebrate splice site selection (Niwa et al., 1992).

1.1.4. Related RNA processing mechanisms

Self splicing

Self splicing involves the removal of introns via RNA catalysis, in which the intron is the catalyst (Bass & Cech, 1984). Group I introns are found in the ciliated protozoan *Tetrahymena*, yeast mitochondria and maize and bean chloroplasts (Cech, 1986). These introns can be removed *in-vitro* in the absence of proteins (Kruger et al., 1982). A different group of mitochondrial introns known as group II introns can also self-splice by a mechanism that involves formation of a lariat RNA (van der Veen et al., 1986). RNA splicing is known to occur rarely in prokaryotes and may also be via self-splicing (Schmidt, 1985).

RNA secondary structure (Waring & Davies, 1984) and base pairing interactions (Waring et al., 1986) are thought to be important in splice site selection during self splicing. Group I introns are in fact also capable of transposition via "reverse splicing" (Mohr & Lambowitz, 1991).

RNA editing

Some pre-mRNA transcripts derived from the mitochondrial genes in kinetoplastid protozoa and plants (Conklin et al., 1991) undergo RNA editing, which involves extensive insertion and, more rarely deletion of uridine residues (Benne, 1990). This process is thought to proceed via the action of a putatative "editosome" with uridine transfer from or to a "guide RNA" via a number of transesterification reactions.

1.2. Alternative pre-mRNA splicing

1.2.1. Alternative splicing

Expression of active proteins in eukaryotic cells is controlled at the level of transcription, pre-mRNA processing, translation, mRNA stability (for review see Cleveland & Yen, 1989) and post translational modification or processing (Darnell, 1982). Alternative splicing is a form of post transcriptional regulation, allowing the generation of sets of related proteins (see fig.1.2. for discussion see Smith et al., 1989). This thereby increases the coding capacity of the genome and is a means of regulating gene expression (for reviews see Breitbart et al., 1987; Smith et al., 1989; Latchman, 1990a). A functional role of introns is unclear but their existence is certainly necessary for the production of sets of related proteins from a single gene by alternative splicing.

Multiple mature mRNA transcripts can be derived from the same primary transcript via a number of possible pathways involving the differential processing of internal pre-mRNA sequences (see fig.1.2. 1-5). Multiple mRNA transcripts can also be derived from a single gene by utilization of different promoters or polyadenylation sites (Nabeshima et al., 1984; Early et al., 1980) (see fig.1.2. 6-7). This may alter 5' or 3' pre-mRNA ends leading to changes in secondary or tertiary structure. These altered 5' or 3' end sequences may allow the use of distinct splice sites, possibly via the action of *trans*-acting factors. Some alternative splicing events may not be regulated, and may reflect

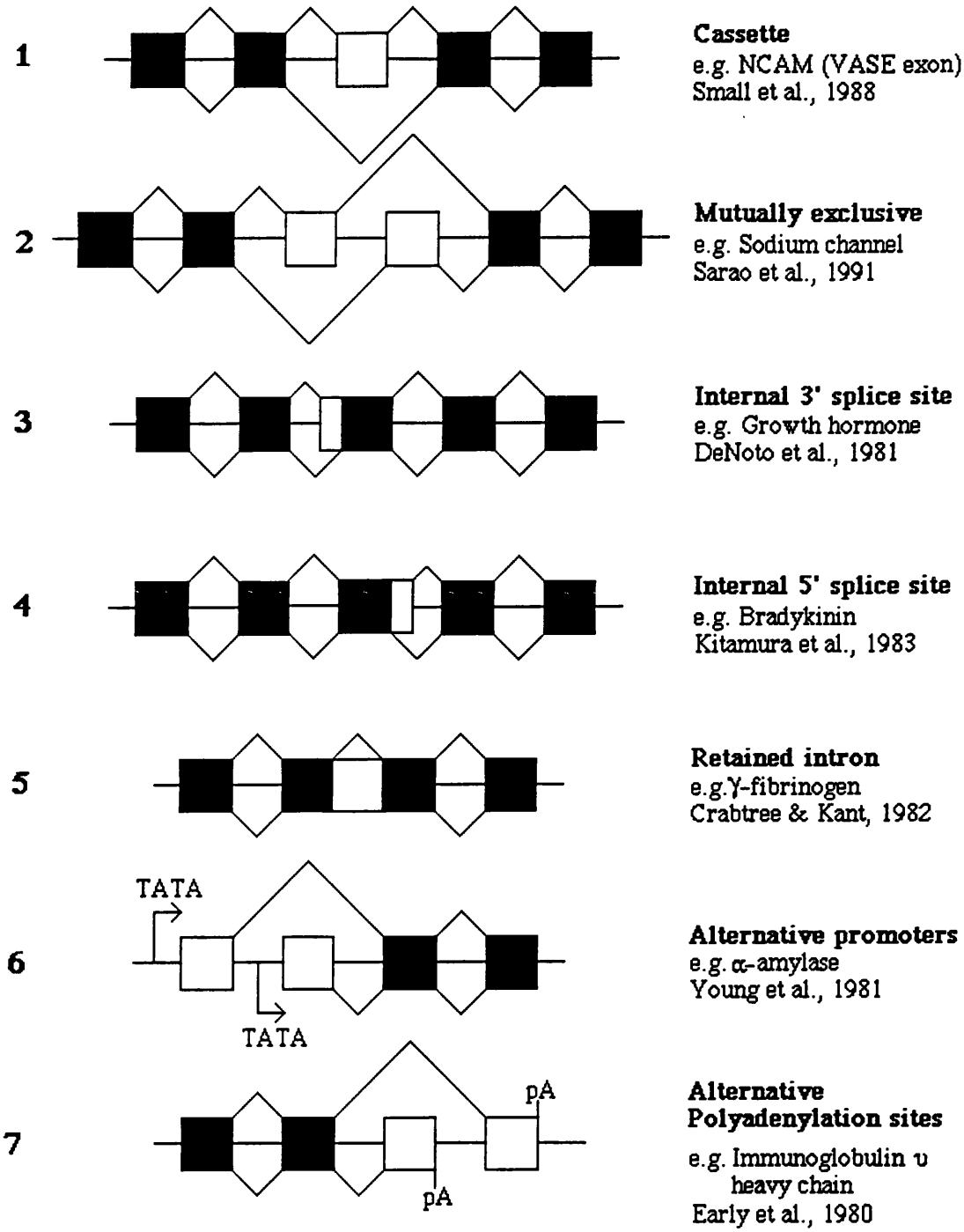


Figure 1.2. Patterns of mammalian alternative splicing

Schematic representation of the known patterns of alternative splicing. Constitutive exons (black), alternative sequences (white) and introns (solid lines). Alternative promoters (TATA) and polyadenylation (pA) signals are indicated. Examples of each pattern with references are indicated.

Adapted from Breitbart et al., 1987.

competition between processing sites for components of the constitutive splicing machinery based on complementarity to the consensus sequences. There are many examples however of regulated splicing of transcripts with the same 5' and 3' ends. Transcript ratios may vary during development or between tissues. Temporally and spatially regulated alternative splicing events clearly must involve the action of regulated *trans*-acting factors (e.g. Manley et al., 1987).

Some alternatively spliced genes contain an exceptionally large number of exons. Dystrophin has more than 65 exons spread over two million base pairs (Feener et al., 1989). Other genes can manifest exceptional diversity of alternative splicing. Troponin-T can generate as many as 64 different mRNAs from the same primary transcript (Breitbart et al., 1985) illustrating the requirement for extensive control of these splicing events.

Alternative splicing and in the case of immunoglobulin genes, DNA rearrangements, are capable of generating related protein isoforms. Alternative splicing is particularly suited to terminally differentiated cells that have lost DNA replicating capacity, and for cells that must respond rapidly to environmental stimuli (e.g. nerve and muscle cells). Unlike DNA rearrangement alternative splicing has the added advantage of being reversible.

1.2.2. Control of alternative splicing

Alternative exons must be distinguished from constitutive exons by *cis*-acting information (for review see Green, 1991), but all of the consensus sequences found to control alternative splicing are generally undistinguishable from those found to control constitutive splicing (Breitbart et al., 1987). *Cis* competition between 5' splice sites has shown sites more closely resembling the consensus sequence (see above) to be used preferentially (Eperon et al., 1986; Fu & Manley, 1987). The 5' splice site consensus of an α-crystallin alternative exon (King & Piatigorsky, 1983) and other examples (Aebi & Weissmann, 1987) have been found to diverge from GU to GC however. A number of pre-mRNA sequences have a role in alternative splice site selection including exon sequences (see below), intron sequences (Khoury et al., 1979; Deutscher et al., 1985; Helfman et al., 1990), the polypyrimidine stretch, which can influence 5' splice site selection (Fu et al., 1988) and splice junction sequences (Emeson et al., 1989) as well as

distal sequences such as the polyadenylation signal (Leff et al., 1987). Splice site proximity can also inluence selection (Reed & Maniatis, 1986; Fu & Manley, 1987).

Exon sequences, including coding sequences have also been shown to have a role in splice site selection. Competition assays have demonstrated a role for exon sequences in selection of β-globin exons one or two (Reed & Maniatis, 1985; Eperon et al., 1986). Deletions or insertions within exon sequences have been found to alter the efficiency of splice site usage (Somasekhar & Mertz, 1985; Ricketts et al., 1987). The inversion of an 81bp sequence in a fibronectin exon leads to exon skipping (Mardon et al., 1987). Insertions or deletions in adjacent exon sequences has also been shown to prevent use of a female-specific alternative 5' splice site within the *Drosophila double-sex* gene (Burtis & Baker, 1989). Mutation of as few as four bases within the cTNT alternative exon has been shown to disrupt selection of that exon (Cooper & Ordahl, 1989) and *cis*-elements in the alternative exon of the leukocyte common antigen pre-mRNA which are required for alternative splicing have been defined (Streuli & Saito, 1989).

Cis competition for constitutive factors could be responsible for constitutive alternative splicing patterns but temporally or spatially regulated alternative splicing requires the action of temporally or spatially regulated *trans*-acting factors. These factors may influence promoter usage, polyadenylation or splice site selection. A number of *trans*-acting factors involved in regulated alternative splicing have so far been identified (see 1.3.5. and table 1.3.).

It has been suggested that snRNA variants may influence alternative splicing choices (see 1.3.1.) and it is also possible that small antisense RNAs could potentially influence alternative splicing patterns. Such a molecule has been found derived from an alternatively spliced region of Troponin-T but this molecule was found to be expressed at high levels in all tissues examined (Medford et al., 1984).

Pre-mRNA secondary structures have been proposed to have a role in alternative splicing (Breitbart & Nadal-Ginard, 1986; Edlind et al., 1987; Estes et al., 1992). These structures could of course be stabilised or destabilised by *trans*-acting factors. Sequestration of exons into the loops of hairpin structures has been shown to result in exon skipping *in-vivo* (Solnick, 1985). These secondary structures must be exceptionally stable

however and do not appear to occur naturally (Solnick & Lee, 1987). The formation of certain secondary structures may be prevented *in-vivo* by the binding of hnRNP proteins to the pre-mRNA during transcription. This allows the distance between complementary regions in the pre-mRNA and the rate of transcription to have an additional effect on secondary structure formation (Eperon et al., 1988).

Steric hindrance of splicing factors has been proposed to be involved in the exclusion of certain small exons (e.g. c-src, Black, 1992) as well as in the mechanism of mutually exclusive exon selection in which the short distance between a 5' splice site and a downstream branchpoint prevent exons from splicing together (e.g. α-tropomyosin, Smith & Nadal-Ginard, 1989).

1.2.3. Genes which undergo alternative splicing

Regulated alternative splicing appears to be prevalent in the nervous system and in muscle (see below). A large number of genes with a wide variety of cellular functions undergo alternative splicing (Leff & Rosenfeld, 1986; Breitbart et al., 1987) including transcription factors (Hatzoupoulos et al., 1990), growth factors (Edwards et al., 1986), oncogenes (Stanton & Bishop, 1987), myosin (Rozek & Davidson, 1983) and peptide hormones (Kitamura et al., 1983). The DNA tumour viruses maximise the efficiency of their relatively small genomes by extensive regulated alternative splicing (Ziff, 1980).

A variety of genes have been found to undergo alternative splicing which is temporally and spatially regulated including CALC-I (Rosenfeld et al., 1983) (see fig 1.6), α-amylase (Young et al., 1981), Ca^{2+}-transporting ATPase (Greeb & Shull, 1989) fibronectin (Ffrench-Constant & Hynes, 1989) and synexin (Magendzo et al., 1991).

Alternative splicing can clearly result in the production of related proteins with different biological activities by inclusion or exclusion of various functional domains. These domains include different signal peptides, which may lead to differential localisation (see below). A number of specific examples have also been described in which related alternatively spliced products act as antagonists (e.g. Mumberg et al., 1991 and see below). In *Drosophila* a number of splicing choices influence eye colour, transposition (Laski et al., 1986) and sex determination (Baker, 1989).

Immunoglobulin genes are the best example of alternatively spliced genes in which the variable 3' end determines whether the protein will become membrane bound or secreted (Early et al., 1980). These immunoglobulin splice variants are also developmentally regulated (Rogers et al., 1980). Alternative splicing may also influence the localization of decay accelerating factor by allowing the introduction of a frameshift, which alters the hydrophobicity of the protein (Caras et al., 1987). Introduction of a stop codon by alternative splicing may also be responsible for regulating the production of active or inactive glutamic acid decarboxlase in adult and embryonic rat brain respectively (Bond et al., 1990).

The *Drosophila* transcription factor I-POU (Treacy et al., 1992) and the mammalian ERA-1 factor (LaRosa et al., 1988) both have alternatively spliced products with altered homeobox domains. The c-erb A gene also encodes alternative products which act as antagonists in the control of thyroid hormone action (Koenig et al., 1989). Mammalian CREB (cAMP-responsive element binding protein) products have different regulatory domains but the same DNA binding domain (Berkowitz & Gilman, 1990) and CREM (cAMP-responsive element modulator) products have alternative DNA binding domains. The CREM products are expressed in a cell specific manner (Foulkes et al., 1991) and are developmentally regulated in the mammalian testes, acting as an antagonist or activator (Foulkes et al., 1992). Developmentally regulated splice variants of synapsin also appear to affect synapse formation (Han et al., 1991). Alternative splicing of fibronectin, an extracellular matrix glycoprotein returns to the embryonic pattern in the region of wound healing in the adult rat (Ffrench-Constant et al., 1989). A specific splicing decision in sensory neurons may even be involved in reactivation of HSV (herpes simplex virus) from the latent state (Spivack et al., 1991).

Alternative splicing influences a number of mammalian, and *Drosophila* transcription factors as well as other gene products. The consequences of these splicing decisions indicates that important temporally and spatially regulated events are influenced by the outcome of this process. Cellular phenotype can clearly be influenced by alternative splicing (Rosenfeld et al., 1984).

1.3. *Trans*-acting factors involved in splicing

1.3.1. The spliceosome : snRNAs and snRNPs

The role of the *cis*-acting sequences mentioned above is mediated via the action of *trans*-acting factors. The spliceosome is composed of a number of such factors.

The spliceosome is a multicomponent ribonucleoprotein complex (Grabowski et al., 1985). This type of complex has also been proposed to have a role in site-specific recombination and DNA replication initiation. The snRNAs (for reviews of snRNAs and snRNPs see Maniatis & Reed, 1987; Guthrie & Patterson, 1988; Mattaj & Hamm, 1989; Lamond et al., 1990; Zeive, 1990) are highly conserved, abundant, uridine rich RNAs which are complexed with proteins to form ribonucleoprotein particles or snRNPs (see table 1.1.). The snRNA genes themselves have no introns and exist in multiple copies per genome with a number of related pseudogenes. The snRNAs are not polyadenylated, have a number of modified bases, particularly at their 5' end and exhibit a nuclear localization. A set of these snRNAs, U1, U2, U4/U6 complex (Hashimoto & Steitz, 1984), U5, U7, U9, U10, U11 and U12 (Reddy et al., 1988) are complexed with common (core Sm) proteins, which allow coimmunoprecipitation with antisera from SLE patients (see below). The major snRNAs have their 5' m^7G cap replaced with a trimethyl guanosine cap (m_3G) except U6 (see table 1.1.) and a number of snRNPs also have specific proteins associated (see 1.3.2.). SnRNP structure is sufficiently conserved between human and amphibian species to allow the formation of a functional hybrid snRNP beween human U1 and U2 snRNA and amphibian proteins (Pan & Prives, 1988). The yeast U1 (Kretzner et al., 1987), U5 (Patterson & Guthrie, 1987) and U4/U6 (Siliciano et al., 1987) homologs can also associate with common Sm proteins in *Xenopus* oocytes, probably via association with similar Sm binding sites. Human Sm antisera does not immunoprecipitate endogenous yeast snRNPs however suggesting that the putatative yeast Sm homologs have divergent epitopes. The yeast snRNPs also have specific proteins associated. The yeast U1 is much larger than its metazoan homolog and the yeast snRNAs are encoded in only single copies per genome and at only about 200 transcripts per cell (Wise et al., 1983). There appears to be a greater variety of snRNAs expressed in yeast cells however (Riedel et al., 1986).

U3 snRNP is involved in rRNA processing, has a nucleolar localization and was not immunoprecipitated by SLE patient antisera. U8 and U13 snRNPs are also involved in rRNA processing and do not assoociate with immunoreactive Sm proteins (Tyc & Steitz, 1989). The U1, U2, U4/U6 complex and U5 snRNPs are necessary for nuclear pre-mRNA splicing and together constitute the spliceosome (see fig.1.1.). As well as the U4/U6 complex a tri-snRNP U4/U6.U5 complex has also been observed (Behrens & Luhrmann, 1991).

RNase-H and deoxyoligonucleotides can be used to inactivate specific snRNPs and determine if they are essential for splicing. U1, (Kramer et al., 1984), U2 (Black et al., 1985), and U4/U6 (Black & Steitz, 1986) were all determined to be essential by this approach. U5 has also been shown to be essential by isolation of yeast mutants (Seraphin et al., 1991). None of these snRNPs were required for polyadenylation however (Berget & Robberson, 1986). 2'O-methyl and 2-aminoadenosine antisense RNA oligonucleotides have also been applied to the study of mammalian U1 (Barabino et al., 1990), U2 (Lamond et al., 1989; Barabino et al., 1989; Barabino et al., 1990), U5 (Lamm et al., 1991) and U4/U6 (Blencowe et al., 1989; Barabino et al., 1990; Wolff & Bindereif, 1992). Depletion and complementation experiments using these oligonucleotides reaffirm the essential nature of these snRNPs in the splicing reaction as well as defining specific functional snRNP regions, and yielding information on which stages of the splicing reaction each individual snRNP acts (see fig.1.1.).

A number of interactions have been identified within the spliceosome between snRNAs and between the pre-mRNA substrate and snRNAs. U1 snRNP protects a 15-17 nucleotide region spanning the 5' splice site (Mount et al., 1983) and promotes the splicing reaction (Zhuang & Weiner., 1986) and the U2 snRNPs were found to protect approximately 40 nucleotides spanning the branch point (Black et al., 1985). This U2 complex formation and hence splicing can be inhibited by mutation of the 3' splice site sequence and deletions in the polypyrimidine tract (Ruskin & Green, 1985). U5 appeared to bind and protect approximately 15 nucleotides at the 3' splice site (Chabot et al., 1985) and can also influence the specificity of 5' splice site cleavage in yeast (Newman & Norman, 1991). Protected regions correlate well with the lower size limit of an intron

(Weiringa et al., 1983). If introns are shortened below this length the splicing apparatus will find a proximal cryptic 5' splice site in the 5' exon. The proximity of the 5' splice site to the branch point is also limited by the ability to form the lariat.

Base pairing of snRNA with pre-mRNA has been directly demonstrated by producing compensatory base changes which supress splice site mutations. By this approach the 5' end of mammalian and yeast U1 have been shown to interact with the 5' splice site (Zhuang & Weiner, 1986; Seraphin et al., 1988; Siliciano & Guthrie, 1988; Seraphin & Rosbash, 1989) and U2 has been shown to interact with the branch point, allowing the acceptor adenosine to bulge out from the pre-mRNA (Parker et al., 1987; Wu & Manley, 1989; Zhuang & Weiner, 1989). U2 snRNA actually exhibits complementarity to the mammalian intron boundary and 3' splice site consensus (Keller & Noon, 1984). The formation of the U4/U6 complex involves base pairing and the U2 and U6 snRNAs have also been shown to exhibit base pairing interactions (Hausner et al., 1990) which are necessary for mammalian pre-mRNA splicing (Wu & Manley, 1991; Datta & Weiner, 1991). Yeast U1 can in fact influence both 5' and 3' splice site selection (Goguel et al., 1991).

Spliceosome formation follows an ordered pathway of snRNP interaction with the pre-mRNA (see fig.1.1.) which also requires a number of other protein factors. The model in figure.1.1. showing an ordered pathway of snRNP binding during spliceosome assembly (Bindereif & Green, 1987; Lamond, 1991a) was derived almost entirely from *in-vitro* experiments using splicing extracts. U1 binding to the 5' splice site is required before U2 can bind to the branch point (Barabino et al., 1990). U2 interacts with the branch point following U2 snRNP auxiliary factor (U2-AF) interaction with the polypyrimidine tract and 3' splice site (Zamore & Green, 1991). U4/U6.U5 can subsequently bind as a tripartite snRNP particle (Konarska & Sharp, 1987). U4 can however enter the complex alone while U6 can enter in the absence of U5 (Vankan et al., 1992). A number of *cis*-acting sequences are required for this stepwise assembly to take place (Frendewey & Keller, 1985) but assembly probably involves RNA/RNA, protein/RNA and protein/protein interactions. The U4/spliceosome association is weakened (Lamond et al., 1988) and the sensitivity of the 3' splice site to RNase-H digestion is altered as the splicing pathway proceeds (Sawa & Shimura, 1991) demonstrating the

dynamic nature of the complex involving a number of structural changes during the reaction. Mammalian U1 appears to associate weakly with the rest of the splicing machinery throughout. The U1 and U4 snRNPs do appear to remain associated with the spliceosome, however (Reed, 1990). Mammalian snRNP homologs appear to undertake splicing by following a similar pathway to that seen in yeast (Ruby & Abelson, 1988).

Seven U1 snRNA variants have been described in *Xenopus laevis* oocytes which are differentially expressed and differ in sequence but not in length (Forbes et al., 1984). U1a and U1b variants have in fact been shown to be developmentally regulated in the mouse (Lund et al., 1985). The U1-A protein (see 1.3.3.) association differs between these U1 variants (Bach et al., 1990). Tissue specific and developmentally regulated variants of U4 (Korf et al., 1988) and U5 (Krol et al., 1981) have also been reported. These variants, which may associate with a variable complement of proteins to form functionally distinct snRNPs may concievably influence differential pre-mRNA processing, but no evidence of this has been reported (Mattaj & Hamm, 1989).

There are a large number of proteins associated with the snRNPs (see below). These proteins are thought to be associated with the need to maintain accuracy while undertaking constitutive and regulated splicing events, via interacting with sequences which contain an apparently low information content.

1.3.2. Mammalian snRNP and mRNA transport and localisation

Following processing in the nucleus mature mRNA transcripts must be transported to the cytoplasm for translation (Mattaj, 1990). SnRNA molecules are also transported to the cytoplasm prior to association with their core Sm protein components, 5' end processing, 3' end trimming prior to return to the nucleus where splicing takes place (Madore et al., 1984) (see fig.1.3.) (for reviews see Zieve & Sauterer, 1990; Andersen & Zieve, 1991; Nigg et al., 1991; Michaud & Goldfarb, 1992). The formation of splicing complexes on pre-mRNA appears to inhibit nuclear export until processing is complete (Chang & Sharp, 1989). Completion of splicing is not always necessary before export to the cytoplasm can occur however.

The U1, U2, U4 and U5 snRNAs like other RNA polymerase II transcripts have a 7 monomethyl guanosine (m^7G) cap added co-transcriptionally. In contrast U6 snRNA, which is an RNA polymerase III transcript, lacks an m^7G cap as well as an Sm protein binding site (see 1.3.) (for review of snRNP transcription and assembly see Parry et al 1989). The U1, U2, U4 and U5 snRNAs are exported to the cytoplasm and like other proteins the Sm proteins are synthesised in the cytoplasm. The Sm proteins and the snRNAs remain in the cytoplasm until forming a complex. These proteins as well as the U1-70K and U2-B" associate with snRNPs in the cytoplasm whereas U1-A, U1-C and U2-A' proteins move into the nucleus independently of the snRNP core proteins and associate with particles in the nucleus. By microinjection of U2 mutants into *Xenopus laevis* oocytes it was shown that association with Sm proteins at the Sm binding site, but not association with U2 specific proteins was necessary for nuclear targetting (Mattaj & De Robertis, 1985). Using a similar system it was also shown that trimethylation of the m^7G cap to form the trimethyl guanosine (m_3G) cap of U2 takes place in the cytoplasm of enucleated oocytes and is dependent upon the presence of an Sm protein binding site and association with Sm proteins, independent of the position of the site (Mattaj, 1986). U1 was also found to enter the *Xenopus* nucleus when complexed with core proteins and possessing an m_3G cap (Hamm et al., 1990). The cap structure is not such a strict requirement for U4 and U5 nuclear transport but U6 possessing an m_3G cap can enter the nucleus in the absence of Sm proteins and U1 lacking stably associated Sm proteins can still enter the nucleus at reduced efficiency. U6 appears to have a region structurally, and in terms of nuclear migration functionally equivalent to an Sm binding site (Hamm et al., 1990). Sm proteins other than those considered to be Sm antigens (B, B' and D) may be necessary or responsible for cap trimethylation however, as U3 snRNA is not immunoprecipitable with Sm antibodies (Lerner & Steitz, 1979) yet it has a trimethyl cap (Busch et al., 1982). Free m_3G profoundly inhibits snRNP nuclear uptake (Fischer & Luhrmann, 1990). The m_3G cap and Sm proteins, or an equivalent, appear therefore to constitute a bipartite karyophilic (nuclear targetting) signal (Hamm et al., 1990; Fischer & Luhrmann, 1990).

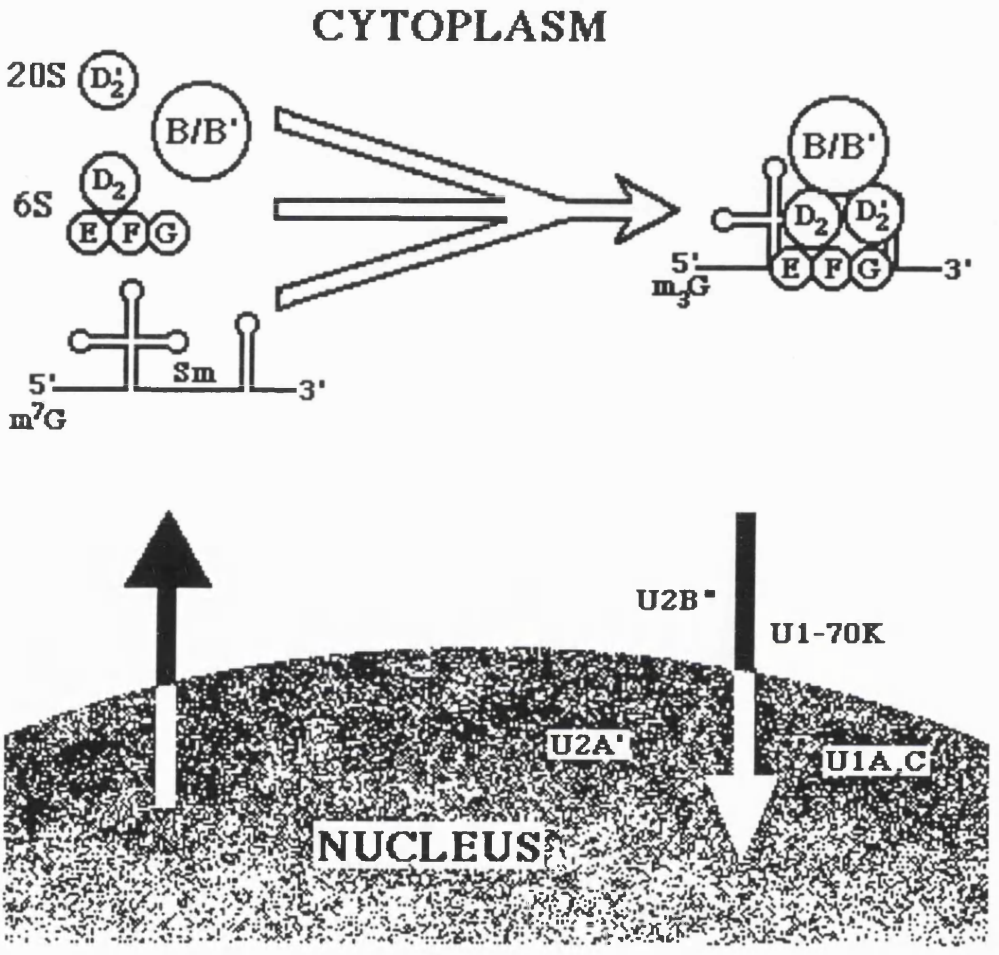


Figure 1.3. snRNP assembly and transport

Schematic representation of the cytoplasmic assembly and nuclear transport of pre-snRNP particles.

Adapted from Zieve & Sauterer, 1990.

SnRNPs are localized in large "sphere" organelles in amphibian oocyte nuclei (Gall & Callan, 1989). In mammalian cells snRNPs form a reticular network throughout the nucleoplasm. Sm staining was localised in 20-50 speckles in the nucleoplasm, complementary to, rather than coincident with sites of transcription and replication (Spector, 1990). Using 2'-*O*-Me oligonucleotides in conjunction with antibodies to probe the distribution of snRNPs U2, U4, U5 and U6 were detected in a small number of foci while U1 and U2-AF (see 1.3.4.) were widely distributed throughout the nucleoplasm, excluding nucleoli (Carmo-Fonseca et al., 1991a). The above foci were also present in unfixed HeLa and 3T3 cells and all spliceosomal snRNPs were co-localised (Carmo-Fonseca et al., 1991b). Neither SC-35 splicing factor (see 1.3.4.), hnRNP proteins (see 1.3.4.) or La antigen (see 1.4.1.) were localised to these foci. The above foci have now been defined as coiled bodies and snRNPs have also been shown to be present in "speckled" interchromatin granules which also contain SC-35 splicing factor (Fu & Maniatis, 1990; Carmo-Fonseca et al., 1992). Other snRNP-related antigens also exhibit overlapping patterns with Sm proteins during parts of the cell cycle (Nyman et al., 1991) during which the pattern of Sm antigen distribution also changes (Deng et al., 1981). The above nuclear organelles appear to be involved in some aspect of pre-mRNA processing; for example, pre-assembly of multi-snRNPs, sorting and degradation or export of splicing products, re-cycling of snRNPs or splicing itself. Proteins involved in 3' end processing exhibit a distinct localisation in the nucleus when compared to those involved in splicing (Takagaki et al., 1990).

1.3.3. snRNP specific proteins

Although specific regions of the RNA components of the snRNPs have been shown to be essential for certain aspects of splicing, little is known about possible functions of the protein components. The snRNPs involved in pre-mRNA processing all have common (see 1.4.2.) and specific proteins associated (see table 1.1.). Models of U1 and U2 snRNP structures have been postulated (Hamm et al., 1987; Mattaj et al., 1986) (see fig.1.4.). Three dimensional models of all the major snRNPs have also been proposed (Luhrmann et al., 1990; Kastner et al., 1992). The U1 specific proteins A, C and 70K and U2 specific

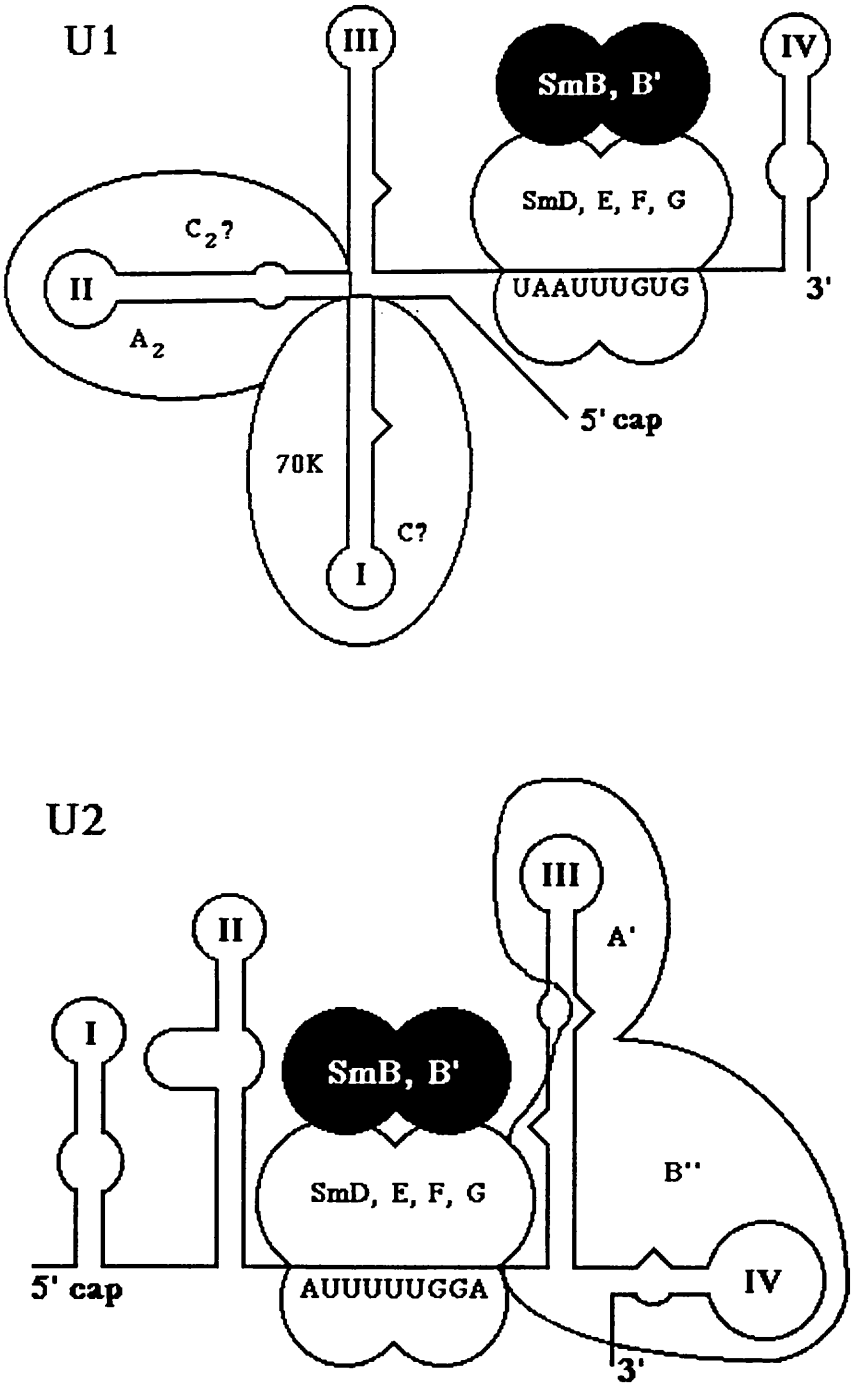


Figure 1.4. A model of the U1 and U2 snRNP structures

The proposed regions involved in RNA/protein and protein/protein interactions are indicated. The Sm proteins D, E, F and G are shown to bind directly to the Sm site while SmB and B' are proposed to bind indirectly.

Adapted from Hamm et al., 1987; Kastner et al., 1992; (U1) Mattaj, 1986; Hamm et al., 1989. (U2).

proteins A' and B" can associate with the snRNAs independently of the common Sm proteins (Hamm et al., 1987). The U1 specific protein binding sites are required in a specific configuration for production of a functional snRNP while the presence of A' and B" in the U2 snRNP do not appear to be essential for splicing (Pan & Prives, 1989; Hamm et al., 1990). U5 has a large number of specific associated proteins (Bach et al., 1989) including intron binding protein (IBP) which interacts with the 3' splice site (Tazi et al., 1986), is Sm antibody reactive and is the homolog of the yeast PRP8 (see 1.3.4.) protein which also binds pre-mRNA (Anderson et al., 1989). The U4/U6 complex also has a number of specific proteins associated and the U4/U6.U5 snRNP also has additional protein factors associated. These additional proteins confer a kinase activity to the tri-snRNP particle (Behrens & Luhrmann, 1991).

A number of snRNP proteins including U1-70K, U1A, U2B", IBP, U2AF, p(A)BP, hnRNP-C and *tra* have a 70-90 amino acid RNA recognition motif (RRM or RNP domain) (for reviews see Mattaj, 1989; Frankel et al., 1991). The crystal structure of the RRM has been defined (Nagai et al., 1990). An arginine rich domain is also known to mediate RNA binding (Lazinski et al., 1989). A charged domain found in the U1-70K protein is structurally similar to domains found in *Drosophila* proteins which are known to regulate alternative splicing (Query et al., 1989). U2B" actually requires U2A' for its interaction with the fourth hairpin of the U2 snRNP (Scherly et al., 1990).

1.3.4. Non-snRNP proteins involved in splicing

A number of hnRNP proteins (Pinol-Roma et al., 1988) appear to associate with nascent RNA transcripts in the nucleus preceding polyadenylation and splicing, and these proteins appear to be exchanged for a different set of associated proteins in the cytoplasm (for review see Dreyfuss et al., 1988). The hnRNP C protein which is a non-sequence specific RNA binding protein is in fact required for splicing and is present in splicing complexes (Choi et al., 1986). Also hnRNP proteins A1 and D bind to the 3' splice site (Swanson & Dreyfuss, 1988). Other protein components are also required for nuclear pre-mRNA splicing (Kramer et al., 1987). For example, anti-SC-35 antibodies (Fu & Maniatis, 1990) inhibit spliceosome formation and splicing *in-vitro*. The intact spliceosome

appears to contain more than 50 different proteins (Reed, 1990).

A large family of more than twenty yeast PRP genes (for review see Ruby & Abelson, 1991; Green, 1991), the products of which are necessary for splicing and some of which are snRNP associated, (e.g. PRP8, see above) and have mammalian homologs suggesting that these conserved genes play fundamental functional or structural roles in the spliceosome (e.g. Anderson et al., 1989). These proteins have been suggested to have a proof-reading role, improving the fidelity of splicing. None of the PRP genes have been found to be homologues of the core Sm proteins however. Some of these proteins contain RNA binding domains and motifs known to posess ATPase activity. Many of these proteins are in fact ATP dependent RNA helicases explaining the requirement for ATP during the splicing process.

Polypyrimidine tract binding protein enhances the U2 snRNP/pre-mRNA interaction (Garcia-Blanco et al., 1989) but is distinct from U2-AF which is a highly conserved factor necessary for U2 snRNP/pre-mRNA binding (Zamore & Green, 1991). SF1 and SF3 are also essential for the U2 snRNP interaction at the branch point (Kramer & Utans, 1991). A number of factors appear to interact with the polypyrimidine tract which may influence splice site selection.

1.3.5. *Trans* factors that influence regulated alternative slicing

Regulated alternative splicing is mediated by *cis*-acting sequences (see 1.2.4.) (e.g. Black, 1992) which interact with various diffusable *trans*-acting factors (e.g. Breitbart & Nadal-Ginard, 1987; Latchman, 1990a). Alternative splicing regulatory factors may bind directly to the RNA or may act indirectly by regulating the activity of other splicing components. Mammalian alternative exon *cis*-acting sequences are generally indistinguishable from those found in constitutive exons (Breitbart et al., 1987) but a number of splicing regulatory proteins which interact with the pre-mRNA directly have been defined. Some of these *trans*-acting factors can control a number of splicing choices (for review see Mattox et al., 1992) (see table 1.2.) suggesting that a number of regulated splicing decisions in mammalian cells could also be controlled by a single factor. It has been proposed that mammalian RNA binding proteins, specifically

Gene	Organism	Regulated pre-mRNA	Choice	Other functions
mer 1	*S. cerevisiae*	*mer 2*	Retained/spliced intron	
rpL 1	*X. laevis*	*rpL 1*		Ribosomal proteins
rpL 32	*S. cerevisiae*	*rpL 32*		
Sxl	*D. Melanogaster*	*Sxl*	Exon skipping/inclusion	Dosage compensation
		tra	Alternative 3' splice sites	
su(w^a)		*su(w^a)*	Retained/spliced intron	
		w^a	Polyadenylation/splicing	
tra		*dsx*	Alternative 3' exons	
tra-2		*dsx*		
		tra-2	Retained/spliced intron	
		exu	Polyadenylation/splicing	
rev	HIV	HIV	Retained/spliced introns	RNA transport
ASF/SF2	Human	SV40 T/t	Alternative 5' splice sites	General splicing factors
hnRNP A1		β-globin		
DSF		SV40 T/t		
		Adeno E1A		

Table 1.2. ***Trans*-acting factors which influence splice site selection**

Adapted from Mattox et al., 1992; Harper & Manley, 1991; (distal splicing fraction) and Mayeda & Krainer, 1992 (hnRNP A1). See text for discussion (1.3.5.).

hnRNP proteins could perform such a function (Bandziulis et al., 1989). HnRNP-A1 is in fact capable of regulating splice site selection (Mayeda & Krainer, 1992, see below). This protein is also capable of shuttling between the nucleus and the cytoplasm (Pinol-Roma & Dreyfuss, 1992). HnRNP diversity may allow these proteins to regulate other tissue specific splice choices (Burd et al., 1989) by associating with the pre-mRNA during transcription (Beyer & Osheim, 1988).

In the fruit fly *Drosophila melanogaster*, differential splicing of a cascade of regulatory genes determines sexual phenotype by producing functional proteins in the female (Baker, 1989). These genes encode RNA binding proteins that regulate splicing patterns of specific transcripts during somatic sexual differentiation (for review see Mattox et al., 1992). The *Sex-lethal* (*Sxl*) gene encodes a female-specific protein that binds to an upstream polypyrimidine tract in its own transcript to block splice site usage and promote usage of a downstream 3' splice site. This regulation acts as an autoregulatory positive feedback, allowing maintained production of functional protein. *Sxl* also promotes splicing of functional *transformer* (*tra*) transcripts (Inoue et al., 1990). *Tra*, in conjunction with *tra-2* promote use of the female specific *double-sex* (*dsx*) exon probably by binding directly to the exon (Hoshijima et al., 1991). *Drosophila* P-element transposition is also controlled at the level of splicing. The formation of a somatic cell specific complex which inhibits U1 snRNP binding at the 5' splice site prevents splicing and hence transposition (Siebel et al., 1992).

Although alternative splicing is common in mammals few mammalian proteins that can influence splice site selection have been identified. Alternative splicing factor (ASF) (Ge & Manley, 1990) was shown to be an essential splicing factor that can activate an SV40 early pre-mRNA proximal 5' splice site present in *cis* on a pre-mRNA with a single 3' splice site in a concentration dependent manner. Splicing factor-2 (SF-2) (Krainer et al., 1990) was shown to have a similar effect on β-globin pre-mRNA splicing and was shown to be a nuclear, non-snRNP associated factor. This effect had been seen previously in diluted HeLa splicing extracts which exhibited a relative increase in the use of distal splice sites in β-globin pre-mRNA (Reed & Maniatis, 1986). This factor was proposed to be important in preventing exon skipping, ensuring splicing accuracy and may be responsible

for alternative splicing of SV40 early t/T, β-globin (Ge & Manley, 1990; Krainer et al., 1990) and other pre-mRNAs *in-vivo*. Cloning of ASF and SF-2 have now shown them to be the same factor (Ge et al., 1991; Krainer et al., 1991). ASF/SF-2 (for review see Lamond, 1991b) is highly conserved between humans and mice and is 248 amino acids in length (approximately 28kd) and appears to exist itself in alternatively spliced forms. ASF/SF-2 shows homology to U1 snRNP 70K protein as well as known *Drosophila* splicing regulators (Mayeda et al., 1992) suggesting that it may function by regulating base pairing between U1 snRNA and 5' intron-exon junction sequences. An activity which antagonises the ASF/SF2 activity by promoting the use of distal 5' splice sites in adenovirus and SV40 pre-mRNA has now been found (Harper & Manley, 1991). The hnRNP-A1 protein may be responsible for this activity (Mayeda & Krainer, 1992). Another factor which influences splice site selection in mammalian cells is the HIV (human immunodeficiency virus) factor Rev, which inhibits splicing by blocking spliceosome formation on HIV pre-mRNA (Kjems et al., 1991).

All factors to date which are capable of influencing splice site selection appear to exert their effect by binding RNA directly. All of these factors appear to act by creating a steric block to spliceosome assembly which promotes the use of another splice site (Mattox et al., 1992). Human ASF/SF-2, hnRNP proteins A1, A2, and C1, U1 snRNP proteins 70K, and A, U2 snRNP protein B" and poly(A) binding protein as well as the *Drosophila sxl, tra-2* and *hdp* proteins all contain an RNA recognition motif (RRM) (Krainer et al., 1991). Another region of homology also exists within these proteins. An arginine and serine (RS) rich domain is found in human ASF/SF-2 and U1 snRNP 70K proteins as well as in *Drosophila tra, tra-2* and *su(w^a)* (Ge et al., 1991). The function of this domain is unknown but it is thought to constitute an activation domain. Generally it appears that a single factor, rather than a combination of *trans*-acting factors are required for regulated splicing events.

1.3.6. Other RNA processing mecanisms involving snRNPs

Trans-splicing

Trans-splicing involves the transfer of an RNA fragment from a "spliced leader" RNA molecule onto an independent pre-mRNA molecule in protozoa, nematodes and plants. All mRNAs are believed to arise via *trans*-splicing in trypanasomatids. *Trans*-splicing has also been found to occur in 10-15% of *C. elegans* transcripts, some of which undergo *cis* and *trans*-splicing (Bektesh et al., 1988) and also in plant mitochondria (Conklin et al., 1991). Spliced leader RNA found in trypanasomatids, *C. elegans*, *Crithidia fasciculata* and *leptomonas collosoma* all have an Sm binding site (see 1.4.2.) with a similar secondary structure to that found in mammalian snRNAs and can all associate with HeLa Sm proteins (Bruzik et al., 1988). Trypanasomatids contain homologs of U2, U4 and U6 snRNAs and these have all been shown to be assembled into snRNPs and to be essential for the *trans*-splicing reaction (Tscudi & Ulla, 1990). The spliced leader RNA does contain a region that resembles an Sm protein binding site and it also has an m_3G cap (Thomas et al., 1988). The spliced leader RNA has in fact been shown to associate with Sm proteins in the form of a snRNP (Bruzik et al., 1988; Van Doren & Hirsh, 1988) suggesting that the Sm proteins may be necessary for *cis* and *trans*-splicing. However, spliced leader RNA lacking an Sm binding site can substitute for U1 RNA in a mammalian *in-vitro* splicing system (Bruzik & Steitz, 1990). *Trans*-splicing has recently been shown to be possible in mammalian cells (Bruzik & Maniatis, 1992) showing functional conservation of *cis* and *trans*-splicing components.

Histone pre-mRNA 3' end formation

The 3' end stem loop structure found on histone mRNA affects the concentration of the histone mRNAs which is varied during the cell cycle (Marzluff & Pandey, 1988). The U7 snRNA is essential in histone pre-mRNA 3' end formation (Strub & Birnstiel, 1986). This snRNA posesses an m_3G cap and is also complexed with Sm proteins (Yang et al., 1981). U7 has three structural domains; an accessible 5' end, which forms base pairs with the pre-mRNA substrate (Schaufele et al., 1986); an Sm binding site, through which the Sm proteins bind; and a 3' hairpin. The first two domains are essential for U7 function (Gilmartin et al., 1988).

1.4. The common (core Sm) proteins

1.4.1. Systemic Lupus Erythematosus and Sm autoantibodies

Autoantibodies to a number of nuclear antigens are a feature of autoimmune diseases in man. These antigens are mostly nucleic acids, nucleic acid binding proteins or nucleoproteins (Tan, 1982) such as Sm and RNP. Sm antibodies are restricted to and are diagnostic for the human disease, SLE and the mouse strain, MRL. Double immunodiffusion analysis (Mattioli & Reichlin., 1973) and immunoaffinity purification (Buchanan et al., 1983) have shown that Sm antigens exist as a complex with RNP and also in an RNP free form. The Sm antigens were found to be closely associated with snRNA molecules but the anti-Sm antibodies found in patient sera and mouse monoclonal antibodies bind to the protein rather than the RNA component, specifically SmN, SmB, SmB' (see 1.4.3.) and SmD and sometimes SmE and U1-A (Guldner et al., 1983; Pettersson et al., 1984). RNP autoantibodies specifically recognise epitopes found on the U1-70K protein.

SLE is an autoimmune disease with many organs and tissues such as joints, skin, kidney, central nervous system, serous membranes, lungs, heart and skeletal muscle all susceptible to disease involvement. Particularly high levels of anti-Sm antibodies have been observed in patients with CNS dysfunction (Winfield et al., 1978). Other autoimmune diseases involving nuclear antigens include mixed connective tissue disease, Sjogren's syndrome, rheumatoid arthritis, scleroderma and dermatopolymyositis. The MRL-lpr mice exhibit an SLE-like autoimmune disease (for review see Cohen & Eisenberg, 1991) and this lymphoproliferation (lpr) disorder is due to defects in Fas antigen mediated apoptosis (programmed cell death) (Watanabe-Fukunaga, 1992; for review see Cohen & Eisenberg, 1992). It is unclear if similar defects are responsible for any cases of human SLE.

It is unclear if the aetiology of SLE is associated with a change in Sm epitopes or if circulating antibodies directed against Sm epitopes are a consequence of other unrelated factors. It is difficult therefore to speculate on the specific functional role of the Sm proteins in SLE based on knowledge of disease symptoms. HSV infection of primate cells has been shown to increase the levels of SmB, B' and D but not that of the nuclear antigen, La (Sharpe et al., 1989c) suggesting that lytic viral infections, allowing exposure of these

antigens to the immune system is unlikely to be responsible for SLE aetiology. HSV however may provide cross-reactive epitopes to the autogenous antigens (Fujinami et al., 1983). Other factors such as genetic background, immune regulation defects and polyclonal B-lymphocyte activation may also be responsible for SLE aetiology.

Since the discovery that sera from SLE patients contain antibodies directed against snRNP components (Lerner & Steitz, 1979), these antibodies have been extensively used to study various aspects of snRNP structure and function. These antibodies actually allowed the first demonstration that snRNPs are involved in splicing by inhibition of the splicing reaction *in-vitro*. Autoantibodies to the nuclear La or cytoplasmic Ro antigens did not inhibit splicing (Yang et al., 1981; Padgett et al., 1983).

1.4.2. Sm protein association with snRNA

The core Sm proteins are common to all the snRNPs particles known to be involved in mammalian pre-mRNA processing (U1, U2, U5 and U4/U6) (see table 1.1.) as well as with a number of minor snRNA molecules (Lerner & Steitz, 1979). Neither U6 or the nucleolar localised U3 snRNPs are directly associated with these proteins although U6 may transiently associate with a functionally related protein (Groning et al., 1991). Interactions between common and specific snRNP proteins and RNA within the U1 and U2 snRNP particles have been determined (Hamm et al., 1987; Mattaj, 1986) and models of U1 and U2 snRNP structure have been derived from this data (see fig.1.3.).

All of the snRNAs known to be involved in splicing have been sequenced and have a proposed secondary structure. The Sm protein binding site ($PuA[U]_nGPu$, where n is more than two in HeLa, rat, chicken and *Drosophila*) (Branlant et al., 1982) is a simple, short, single stranded RNA motif which serves as the site for assembly of the Sm antigens (Liautard et al., 1982) and is usually located between two conserved hairpins (Branlant et al., 1982) on each snRNA. The Sm proteins, B, B', D, D', E, F and G are common to all four snRNP particles. The Sm binding site is recognised and protected from nuclease digestion by a more tightly bound 6S core particle (Fisher et al., 1985) consisting of the smaller SmD, E, F and G proteins. SmF appears to interact directly with the Sm protein binding site (Woppmann et al., 1988). The SmB and B' proteins may bind to the snRNPs

via protein/protein interactions (Mattaj et al., 1986). U6 snRNA lacks an Sm site and associated Sm proteins, but can be Sm immunoprecipitated by virtue of its association with U4 snRNP (Hashimoto & Steitz., 1984) which does associate with the Sm proteins. The Sm protein binding site appears to be conserved between humans, *Drosophila*, plants and yeast (Mount & Steitz, 1981; Tollervey & Mattaj, 1987). Yeast snRNA can be immunoprecipitated with anti-Sm antibodies when associated with *Xenopus* Sm proteins (Riedel et al., 1987).

Whether the functions of the core proteins are restricted to roles in snRNP biosynthesis and metabolism remains unknown. The region of U4 containing the Sm binding site is dispensable for spliceosome assembly however (Wersig & Bindereif, 1990). As well as a proposed role in cytoplasm-nuclear transport (see 1.3.2.) Sm proteins have been suggested to have a stabilising function in U1 (Hamm et al; 1987) and U2 snRNPs (Mattaj & De Robertis, 1985).

1.4.3. The closely related proteins SmN, SmB & SmB'

SmB is a constitutive factor found in all tissues in both rodent and human. It has a predicted 231 amino acids and a predicted molecular weight of 23.7 kd. SmB' is an alternatively spliced product of the SmB gene (van Dam et al., 1989; Chu & Elkon, 1991) and is not found in rodent tissues and cells (Sharpe et al., 1989a; Schmauss & Lerner, 1990). The SmB' mRNA lacks 146 nt in the 3' region of the open reading frame. This 146 nucleotide stretch contains a termination codon in frame 9 nucleotides downstream from the 5' end, so SmB' mRNA actually yields a larger protein with a predicted 240 amino acids and a predicted molecular weight of 24.6 kd (van Dam et al., 1989; Elkon et al., 1990). The SmN mRNA also produces a protein of a predicted 240 amino acids in length with a predicted molecular weight of 24.6kd. This protein is expressed in a tissue specific fashion (see below). There are 17 different amino acids between SmB' and SmN. All charged amino acids are identical and six valine or isoleucine residues in SmN are replaced by methionine in SmB'.

Isolation and sequencing of SmN cDNA clones from rodent (McAllister et al., 1989 (rat); Li et al., 1989 (rat:Sm51); Gerrelli et al., 1991 (mouse)) and human cells (Rokeach et

al., 1989; Schmauss et al., 1989; Sharpe et al., 1989b) and of SmB and SmB' cDNA clones from rat (Li et al., 1989 (Sm11)) and human cells (Ohosone et al., 1989; van Dam et al., 1989) allowed a number of sequence comparisons to be carried out. SmN is closely related to SmB and SmB'. The coding regions of SmN in rat and human cDNA clones are 91.2% homologous at the nucleotide level and the predicted amino acid sequences are identical in rat, mouse and human (Schmauss et al., 1989. Gerrelli et al., 1991) (see fig.1.5.). On comparison of human SmN and SmB/B' sequences it was found that they are 81% homologous at the nucleotide level and 93% identical at the amino acid level. These proteins are therefore closely related but derived from distinct genes (van Dam et al., 1989; Schmauss et al., 1989). Predicted amino acid sequences of rat and human SmB clones also proved to be identical (van Dam et al., 1989). SmN and SmB' both contain 240 amino acids, so their co-migration on SDS polyacrylamide gels is predicted. SmB' is only expressed in human cells however and can be seperated from SmN by virtue of charge differences (SmB' is more basic than SmB or SmN) on two dimensional gels (Sharpe et al., 1989a).

40-45% of the amino acids at the carboxyl terminal third of these proteins are proline residues. Similar proline rich regions are present in the U1-A and U1-C proteins and some anti Sm antibodies recognise U1-A. There is no apparent sequence homology to SmD (Rokeach et al., 1988) however, which is recognised by some anti Sm sera (McAllister et al., 1989; van Dam et al., 1989). These Sm proteins do not contain an RRM (see 1.3.5.) which is present on U1-A (Query et al., 1989) but similar proteins have been shown to bind RNA directly and the 12% of positively charged residues alone could allow RNA binding (van Dam et al., 1989). The proline-rich domain could also be involved in RNA binding. SmN may interact with the same site as SmB and SmB' due to the similarity between the proteins (see fig.1.2.) and they need not necessarily interact directly with the RNA component of the snRNPs (see fig.1.2.). There is also homology with the Epstein Barr nuclear antigen and a poly-alanine homology is also seen with a yeast poly(A) binding protein which could be involved in protein-protein contacts (Rockeach et al., 1989). These Sm proteins may undergo post-translational modification *in-vivo* (Rockeach et al. 1990).

1). hSmN	MTVGKSSKML	QHIDYRMRCI	LQDGRIFIGT	FKAFDKHMNL	ILCDCDEFRK	IKPKNAKQPE
2). rSmN	----------	----------	----------	----------	----------	----------
3). mSmN	----------	----------	----------	----------	----------	----------
4). hSmB'	----------	----------	----------	----------	----------	-----S--A-
5). hSmB	----------	----------	----------	----------	----------	-----S--A-
	REEKRVLGLV	LLRGENLVSM	TVEGPPPKDT	GIARVPLAGA	AGGPGVGRAA	GRGVPAGVPI
	----------	----------	----------	----------	----------	----------
	----------	----------	----------	----------	----------	----------
	----------	----------	----------	----------	-----I----	---I-----M
	----------	----------	----------	----------	-----I----	---I-----M
	PQAPAGLAGP	VRGVGGPSQQ	VMTPQGRGTV	AAAAVAATAS	IAGAPTQYPP	GRGTPPPPVG
	----------	----------	----------	----------	----------	----------
	----------	----------	----------	----------	----------	----------
	----------	----------	----------	----A-----	----------	---G----M-
	----------	----------	----------	----A-----	----------	---G----M-
	RATPPPGIMA	PPPGMRPPMG	PPIGLPPARG	TPIGMPPPGM	RPPPPGIRGP	PPPGMRPPRP
	----------	----------	----------	----------	----------	----------
	----------	----------	----------	----------	----------	----------
	-GA----M-G	----------	--M-I--G--	--M-------	------M---	----------
	-GA----M-G	----------	--M-I--G--	--M-------	------M--L	L

1). Human SmN — Rokeach et al., 1989
2). Rat SmN — McAllister et al., 1989
3). Mouse SmN — Gerrelli et al., 1991
4). Human SmB' — van Dam et al., 1989
5). Human SmB — van Dam et al., 1989

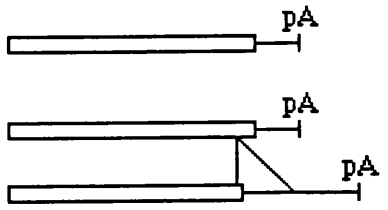


Figure 1.5. SmN, B and B' predicted amino acid sequence comparisons

Sequence comparisons of the closely related Sm proteins showing the constitutive factor SmB, the human factor SmB' and the tissue specific factor SmN (see text). References from which the sequences were derived are indicated and a schematic representation of the alternatively spliced SmB and SmB' as well as the SmN cDNA clones is shown.

The SmN and SmB proteins are therefore 100% conserved between species while SmB and SmN differ by approximately 10% within a species. This level of conservation suggests an important functional role for this closely related family of proteins.

1.4.4. SmN expression and distribution

SmN is expressed in a highly tissue specific manner in mouse rat and human. This protein is abundant in brain, expressed at low levels in heart and is usually undetectable in most other tissues (McAllister et al., 1988; McAllister et al., 1989) (see table 1.3.). The great majority of RNA and protein components of the spliceosome are ubiquitously expressed (Maniatis & Reed, 1987) making this the first example of a mammalian tissue specific snRNP protein (amphibian tissue specific and developmentally regulated proteins which were recognised by anti-Sm autoantibodies have been observed, Fritz et al., 1984). SmN expression has also been characterised in a number of cell lines and other tissues. Briefly, SmN expression decreases during EC and ES cell differentiation (Sharpe et al., 1990a) suggesting regulation during embryonic development. Also neuronal cells have been seen to express SmN protein (see table 1.3.) but none of the cell lines examined to date parallel the high expression levels seen in brain tissue. SmN expression is regulated at the level of transcription (Schmauss et al., 1992) and mRNA and protein levels have been found to correlate in the tissues and cells tested (McAllister et al., 1988; Li et al., 1989). Also SmN mRNA distribution in human tissues is similar to that found in the rat (McAllister et al., 1989). It is in fact possible to detect SmN mRNA transcripts in most tissues by PCR but their relatively low abundance in kidney and liver (Schmauss & Lerner, 1990; Delsert & Rosenfeld, 1992) is reflected by undetectable protein levels. SmN is snRNP associated but may not necessarily be present in all major snRNPs (McAllister et al., 1988). SmN may in fact have a lower affinity for the U1 snRNP relative to the U2 snRNP (personal communication, J. Huntriss, UCMSM).

SmN and SmB (SmB/SmB'human) are derived from distinct genes in humans and rodents (Schmauss et al., 1989; Schmauss & Lerner, 1990). Non-functional SmN pseudogenes have been isolated from rat (Schmauss & Lerner, 1990) and mouse (Grimaldi et al., 1992) but Southern blots of genomic DNA suggested that SmN pseudogenes are not

	McAllister et al., 1988	McAllister et al., 1989	Schmauss et al., 1989	Schmauss & Lerner, 1990	Schmauss et al., 1992	Li et al., 1989	Rokeach et al., 1989	Sharpe et al., 1989a	Sharpe et al., 1989b	Sharpe et al., 1990a	Sharpe et al., 1990b	Delsert & Rosenfeld, 1992	Consensus
Brain	+++	++	+	++		++	+++	f ++				++	+++
Pituitary						++							++
Cerebellum	++	+											++
Heart	+	+		+		+							+
Adrenal				+/-									+/-
Olfactory epithelium	-			+/-									+/-
Kidney				-		-						+a	-
Liver	-	-	-	-		-		f -	-			a + -	-
Lung	-			-		-							-
Prostate						-							-
Small intestine				-									-
Spleen	-					-							-
Testis				-									-
Thymus	-												-
Skeletal muscle						-							-
Stomach						-							-
F9	+			+		+				+ d		+	+
PCC3								+	+	+ d			+
PC12	+	+		+		+		+				+	+
Raji		-	+	+			+						+
TT		+	++	++	++								++
HeLa (M)											+		+
HeLa (I)											-		-
HeLa		-	-	-	+		-					-	-
3T3	-												-
A20	-	-				-		-					-

Table 1.3. SmN expression (relative to SmB/B') in tissues and cells

Abundance : +++ = high, ++ = intermediate, + = low, +/- = very low, - = undetectable.

SmN expression in rodent and human tissues and cells. Protein distribution was determined by Western blotting while mRNA distribution was determined by a combination of Northern blotting, RNase protection assays, PCR and *in-situ* hybridisation. See text for discussion.

a = Assay by PCR i.e. greatly increased sensitivity.

d = Differentiated cultured cells

f = Foetal tissue.

present in human tissue, (Schmauss et al., 1989) from which an intron containing SmN gene was isolated (Schmauss et al., 1992). Subsequently a human pseudogene has been found on chromosome six (Ozcelik et al., 1992) and the mouse pseudogene was localised to chromosome fourteen (Leff et al., 1992). The SmN and SmB genes appear to be derived from a common ancestor due to a duplication event. They have very similar structural organisation although they appear to be located on different chromosome. The SmN gene contains eight exons and exon one does not appear to be translated (Schmauss et al., 1992).

SmN has recently been shown to be subject to maternal imprinting with only the paternal copy being expressed in mouse brain. This gene exhibits linkage with the human Prader-Willi syndrome (PWS) region and presents a good candidate for involvement in the disease (Leff et al., 1992) i.e. PWS may be caused by an RNA processing defect in neurons. Other genes in this region are also likely to be involved in PWS. The SmN gene is located in conserved regions found on mouse chromosome seven and on human chromosome fifteen (Ozcelik et al., 1992) (not chromosome four as originally reported (Schmauss et al., 1992)). Mice with a maternal duplication of chromosome seven exhibit a loss of SmN expression due to maternal imprinting on both chromosomes. These mice present a model of PWS (Cattanach et al., 1992). Together with the conservation and the highly specific tissue distribution of SmN this suggests an important role for this protein.

1.4.5. A role for SmN in alternative splicing?

The similarity between SmN and SmB suggests that they may have similar functions within snRNP particles. The distribution of SmN and its association with the splicing machinery led a number of investigators to suggest a possible role for SmN in alternative splicing. Specifically a role in CALC-I processing (see 4.1.) was suggested (McAllister et al., 1989; Sharpe et al., 1989; Li et al., 1989) as well as in other splicing decisions (Schmauss et al., 1992; Delsert & Rosenfeld, 1992). A change in the complement of snRNP proteins in different tissues may therefore be involved in tissue specific splice choices. The U1-70K protein has been shown to interact with the Sm proteins which may in turn influence the 5' splice site interaction (Patton & Pederson, 1988). In addition U1

has been shown to play a role in 3' as well as 5' splice site selection (Kuo et al., 1991).

A number of genes have been reported to undergo alternative splicing in brain and heart relative to other tissues, exhibiting a distribution which correlates with SmN expression (e.g. dystrophin, Bies et al., 1992). Only a select number of these genes with well characterised expression patterns were chosen for assay however (see chapter 4 for a discussion of the selected splice choices found to correlate with SmN expression in the literature). Other genes found to exhibit splicing patterns which correlate to a degree with SmN expression include α and β-preprotachykinin transcripts which encode substance-P or both substance-P and substance-K respectively. The β transcript contains a cassette exon (see fig.1.2.) which is absent in the α transcript. The α transcript predominates in all brain regions examined while the β transcript predominates in the thyroid and intestine (Nawa et al., 1984). Clathrin light chains can contain brain-specific sequences due to alternative splicing (Jackson et al., 1987) and an amyloid A4 precursor mRNA isotype derived by alternative splicing is abundant in rat and human brain relative to a number of other tissues (Kang & Muller-Hill, 1990).

Aims

Molecular characterisation of Sm proteins can be helpful in understanding their central role in the splicing process and in SLE. Differences in primary structure between SmB, SmB' and SmN may give rise to differences in protein function. The distribution of SmN allows specific putatative roles of this protein to be more readily investigated.

The aim of this research therefore was to characterise SmN, SmB and SmB' distribution in detail. Cell and tissue models would subsequently be selected in which to examine the potential role of SmN in alternative splicing of the CALC-I gene. Other candidates for regulation by SmN were also to be selected from the literature, based on the expression data. It was also hoped that SmN expression data would present clues as to the role of this protein in alternative splicing (i.e. which genes are affected) or even clues as to another potential role in pre-mRNA processing or snRNP biosynthesis and metabolism.

The availability of the anti-Sm monoclonal antibodies, cDNA clones spanning the entire coding region of SmN and SmB' and CALC-I minigene constructs present the materials necessary to address the above problems.

CHAPTER 2

MATERIALS

&

METHODS

2.1. MATERIALS

2.1.1. Bacterial strains

Escherichia coli JM101 : *supE thi* Δ(*lac-proAB*) F' [*traD*36 *proAB*$^+$*lacI*^{q}*lacZ*ΔM15]

2.1.2. DNA

i). plasmids

The following were gifts from the investigators listed:

Bluescript-SK	(Short et al., 1988)	(Stratagene)
RSV-*cat*	(Gorman et al., 1982)	Dr. C. Gorman
SmN-201	(Gerrelli et al., 1991)	Dr. N. G. Sharpe
SmB'-16	(van Dam et al., 1989)	Dr. T. Cuypers
M5G-*neo*		Dr. M. Collins (Chester Beatty, London)
pHβApr1-*neo*	(Gunning et al., 1987)	Dr. L. Kedes
pJ4(*cat*) pJ5(*cat*) pJ7(*cat*)	(Morgenstern & Land, 1990)	Dr. H. Land
pRSV-Cal2 pRSV-Cal3		Dr. P. Broad (UCMSM)
SM-1	(Cooper et al., 1988)	Dr. T. Cooper

ii). recombinant phage

T3/2 in λgt11 (Williams et al., 1990) Dr. N. Sharpe

iii). oligonucleotides

pd(N)6 random hexanucleotides 0.9units/*ul* Pharmacia

The oligonucleotides listed below were either synthesised using an Applied Biosystems 381A DNA synthesiser or were gifts from the investigators listed. References indicate the source of the sequence.

a-, h-, m- or r- preceding the name of an oligonucleotide indicates the species from which each sequence was derived. i.e. a=avian, h=human, m=mouse and r=rat.

Antisense = A, Sense = S.

		5'					3'	
-40		GTT	TTC	CCA GTC	ACG	AC		Sequenase kit
act-1		CCG	GGC	TGT ATT	CCC	CTC	CC	Dr. K. Lillycrop
act-2		TCA	TTG	TAG AAG	GTG	TGG	TG	(UCMSM)
rib-5		ATC	GCT	CCT CAA	ACT	TGA	CC	Dr. K. Lillycrop
rib-3		AAC	TAC	AAC CAC	CTC	ATG	CC	
gt11-1		GGT	GGC	GAC GAC	TCC	TGG		Dr. K. Lillycrop
gt11-2		CAG	ACC	AAC TGG	TAA	TGG		
r-CGRP2	S	GAG	GCA	TCA TGG	GCT	TTC	TG	Amara et al., 1984
r-CGRP5	A	TCA	CGC	AGG TGG	CAG	TGT	T	"
RSV	S	CCA	TTC	ACC ACA	TTG	GTG	TGC	Yamamoto et al., 1980

h-Cal4	A	TAC	CAG	CCC	AAA	GAG	CCA	CC						Broad et al., 1989
h-CGRP5	A	TGG	AAC	CCA	CAT	TGG	TGG	GC						"
h-C4	A	ACC	GCT	TAG	ATC	TGG	GGC	TG						"
r-NCAM7	S	AGG	AGC	AAG	TCA	CTC	TGA	CT						Small et al., 1987
r-NCAM8	A	TTC	AAG	GTA	CAT	GGA	CTG	GG						"
m-src3	S	CTG	TCC	TTC	AAG	AAA	GGG	GAG	C					Martinez et al., 1987
m-src4	A	TGG	ATG	GAG	TCG	GAG	GGC	GC						"
m-Go6	S	TCT	TCT	GTG	TCG	CAC	TCA	GC						Strathmann et al., 1990
m-Go8	A	CTG	TAG	ACT	TCC	TTG	TGA	GC						"
m-Go9	A	CAG	CAA	AGA	GTC	CAT	GAA	GC						"
a-cTNT6	A	CCT	CGT	CTA	CCT	GAT	CCT	CC						Cooper & Ordahl, 1985

iv). molecular weight markers

Either 1kb DNA ladder (Gibco BRL), lambda DNA, Hind-III digest (Gibco BRL) or pGemini (Promega).

2.1.3. Monoclonal Antibodies

The KSm4 and KSm5 antibodies were gifts from Dr. D. Williams of the Kennedy Institute, London. Rabbit anti mouse peroxidase, flourescein and rhodamine conjugated antibodies were obtained from Dako.

2.1.4. Enzymes	units/*ul*	
Rnasin	40	Boehringer Mannheim
T7 RNA Pol	20	"
T4 DNA ligase	2.5	"
DNase 1	35	"
Alkaline phosphatase	1	"
Proteinase-K		"
DNA polymerase 1 (Klenow)	6	Gibco BRL
M-MLV Reverse Transcriptase	200	"
T3 RNA Pol	50	"
Restriction endonucleases		"
Taq DNA polymerase	4.2	Amersham
Lysozyme		Sigma
RNase-A		"

2.1.5. Cell lines

i). Neuronal

ND3, ND7, ND8, ND11 ND15, ND21 and N18Tg2 cells (Wood et al., 1990) were a gift from Dr. J. Wood, Sandoz Institute, London.

ii). Non-neuronal

PC12 cells (Greene & Tischler, 1976) were a gift from Prof. F. Walsh, Guy's hospital, London.

HeLa-I cells were a gift from the cell culture laboratories, I.C.R.F. laboratories, Lincoln's Inn Fields, London.

HeLa-M cells were a gift from Dr. D. Kioussis, N.I.M.R. London.

NIH-3T3 cells were a gift from Dr. N. G. Sharpe, UCMSM.

iii). EC cells

F9 EC cells (Liesi et al., 1983) were a gift fron Dr. N. LaThangue, N.I.M.R. London.

P19S18O1A1 (O1A1) EC cells (McBurney & Rogers, 1982) were a gift from Dr. J. den Hartog, Hubrecht laboratory, Utrecht.

2.1.6. Tissue samples and RNA

Rat and human tissues were a gift from Dr. P. Barton of the National Heart and Lung Institute, South Kensington, London. A developmental series of mouse brain RNA was a gift from Dr. Keith Grimaldi (UCMSM).

2.1.7. Chemicals, Equipment and Miscellaneous

All solid chemicals, organic solvents, alcohols and acids were obtained from BDH and were of AnalaR grade unless stated. Rainbow mid range protein molecular weight markers were obtained from Amersham. dATP, dCTP, dGTP and dTTP were obtained from Pharmacia. All radiochemicals ([α-^{32}P] dCTP (10*u*Ci/*u*l 3000Ci/mmol), [α-^{32}P] CTP (10*u*Ci/*u*l 800Ci/mmol) and chloramphenicol (1*u*Ci/*u*l) D-Thero-[dichloroacetyl-1,2-^{14}C]) were obtained from Dupont.

Nylon (Hybond N) and nitrocellulose (Hybond C) hybridization transfer membranes were obtained from Amersham. XAR 5 fast film was obtained from Kodak and Hyperfilm-MP was obtained from Amersham. Film fixer and developer were obtained from Photosol. All plasticware for tissue culture was obtained from Nunclon and all other disposable plasticware was obtained from Greiner. 3MM chromatography paper was obtained from Whatman and disposable filters were obtained from Sartorius.

2.1.8. Buffers and other stock solutions (All concentrations expressed at x1)

	Stock concentration	
Denhardts	x 100	0.02% w/v ficoll type 400 (Sigma), 0.02% w/v PVP (Sigma), 0.02% w/v BSA (Sigma).
Luria-Bertani medium (LB)	x 1	1% (w/v) bactotryptone, (Difco laboratories) 0.5% (w/v) bacto yeast extract, (Difco laboratories) 170mM sodium chloride pH 7.

HBS	x 10	140mM sodium chloride, 25mM Hepes, 1.1mM disodium orthophosphate.
MEA	x 10	200mM MOPS,1mM EDTA, 5mM sodium acetate pH 7.
PBS	x 10	104mM sodium chloride, 1.8mM potassium chloride, 5.4mM di-sodium orthophosphate dihydrate, 1.25mM potassium dihydrogen orthophosphate pH 7.
SSC	x 20	150mM sodium chloride, 15mM sodium citrate pH 8.
TBS	x 1	50mM Tris, 150mM sodium chloride pH 8.
TAE	x 50	40mM Tris, 20mM sodium acetate, 2mM EDTA pH 8.3.
TBE	x 10	90mM Tris, 90mM orthoboric acid, 2mM EDTA pH 8.3.

2.2. METHODS

The protocols listed below were generally adapted from those found in Sambrook et al., 1989. Protocols are individually referenced if they were obtained in whole or in the major part from elsewhere.

2.2.1. General

i). Chemicals

All solid chemicals were dissolved in dH_2O, adjusted to the correct pH with hydrochloric acid or sodium hydroxide, autoclaved or filter sterilized and stored at room temperature before use unless stated. Retinoic acid (Sigma) was dissolved in ethanol and dexamethasone (Sigma) was dissolved in DMSO. SDS and sodium hydroxide were not autoclaved, guanidinium isothiocyanate (BRL) was always filter sterilized and MOPS, chloroform, phenol, butanol, ethidium bromide (Sigma) and Tween-20 (Sigma) were

stored in the dark. Phenol, Hepes and dexamethasone were stored at 4°C and calcium chloride, deionised formamide, retinoic acid, cAMP (Sigma), ampicillin (Berk), X-gal (BRL), DTT (250mM aliquots), APS (10% aliquots) and all enzyme reaction buffers were stored at -20°C.

Chloroform always contained 4% v/v isoamyl alcohol (Sigma), phenol contained 0.1% w/v hydroxyquinoline (Sigma) and was buffered by shaking 3 times with an equal volume of 0.5M Tris pH 8 and with 0.1M Tris pH 8, 0.2% v/v β-mercaptoethanol twice, removing the aqueous layer each time. Formamide was deionised by stirring with 10% w/v duolite MB 6113 mixed resin for 30 minutes and subsequently filtering through 3MM paper. Solutions treated with DEPC (Sigma) were shaken for 3 hours containing 0.1% v/v DEPC and were then autoclaved for at least 1 hour.

ii). Cleaning nucleic acid preparations

Phenol/chloroform, phenol and chloroform extractions were carried out by addition to an equal volume, vortexing and spinning for 2 minutes at 12,000g. The top aqueous phase was retained. Ethanol precipitations of DNA and RNA were carried out by addition of at least 2.5 volumes of ethanol (Hayman Ltd) and 0.01 volumes of 4M sodium chloride to DNA and 0.1 volumes of 3M sodium acetate to RNA and storage at -20 or -70°C respectively overnight. Precipitated DNA was spun for 5 minutes while RNA was spun for 15 minutes at 12,000g. Pellets were washed in approximately 500*u*l of 70% v/v ethanol and spun as above. Pellets were then dryed in a freeze dryer and resuspended in dH_2O.

iii). Equalising DNA, RNA and protein concentrations

DNA and RNA concentrations were determined by optical density at 260nm using a Shimadzu UV-150-U2 double beam spectrophotometer (deuterium lamp). Protein concentrations were determined using the Pierce coomassie protein assay reagent (Pierce) (Sedmak & Grossberg, 1977). RNA concentrations were further equalised by visualization on agarose/ethidium bromide stained gels or by slot blot analysis using a Bio-Rad Bio-Dot SF slot blotting apparatus. RNA was cross linked to nylon membranes and hybridised to an rRNA cDNA (Kemp et al., 1986) oligo labelled (2.2.5.) probe using the same procedure as for Northern blots (2.2.13.).

iv). Miscellaneous

DNA and RNA were denatured by boiling for 3 minutes. Autoradiography was carried out at -70°C with an intensifying screen and all scintillation counting was carried out in 4ml of Ecoscint A scintillation solution (National Diagnostics) in an Intertechnique SL-30 liquid scintillation spectrometer. *E.coli* JM101were maintained on 1.6% w/v agar (Difco laboratories) made up in minimal medium (1M glucose, 350mM disodium hydrogen orthophosphate, 110mM potassium dihydrogen orthophosphate, 45mM sodium chloride and 100mM ammonium chloride). Herring testes DNA (Sigma) was prepared by sonication and boiling for 10 minutes. DNase-free RNase-A was prepared by boiling for 10 minutes at 10mg/ml in dH_2O and was stored at -20°C.

DNA

2.2.2. Competent cells and Transformation (all solution were sterile)

E. coli were maintained by growth on minimal medium. A single colony of *E. coli* JM101 cells was used to innoculate 2.5ml of LB. After shaking for 2 hours at 37°C these mid-log cells were diluted 1/20 in LB and incubated shaking at 37°C for a further 1.5 hours. Cells were pelleted at 3,000rpm for 10 minutes in an IEC Centra-4R benchtop centrifuge and resuspended in 25ml of 100mM calcium chloride at 4°C. The cells were again spun at 3,000rpm for 5 minutes at 4°C and were resuspended in 2.5ml of 100mM calcium chloride. Cells were either diluted with 0.5 volumes of glycerol and stored for up to a month at -70°C for transformation with homogeneous plasmid DNA or were left at 4°C overnight and transformed with newly ligated DNA the following day.

100*u*l of competent cells were transformed by addition of 10-20ng of homogeneous DNA or all 10*u*l of a ligation mix (see 2.2.7.) and left at 4°C for 30 minutes. After the addition of 200*u*l of LB, cells were incubated at 37°C for 30 minutes and plated on 1% agarose/LB plates containing 50*u*g/ml ampicillin and incubated at 37°C overnight. Plates also contained 0.02% w/v X-gal (Gibco BRL) for selection of Bluescript-SK clones. Single white colonies were picked for assay by miniprep (see 2.2.3.i) or by hybridisation assay (see 2.2.7.) prior to miniprep.

2.2.3. Plasmid Purification (all solutions were sterile)

i). Small scale preparation of plasmid DNA (miniprep)

Alkali lysis from Sambrook et al., 1989 except for the following modifications:-

Solution 1 = 20% w/v sucrose, 50mM EDTA, 50mM Tris pH 8

Cell pellets were resuspended in 150*u*l of solution 1 and were stored on ice for 10 minutes after addition of 350*u*l of solution 2.

Solution 3 = 3M sodium acetate. pH adjusted to 4.8 with glacial acetic acid.

After addition of 250*u*l of solution 3 lysates were left on ice for 20 minutes and plasmids were precipitated from the supernatent with 700*u*l of -20°C isopropanol at room temperature for 15 minutes. Precipitates were spun at 12,000g for 5 minutes and resuspended in 50*u*l of dH_2O. After phenol/chloroform extraction and ethanol precipitation pellets were resuspended in 20*u*l of dH_2O and typically 5*u*l was digested with the appropriate restriction enzymes in a 10*u*l volume containing 0.5*ug*/*u*l of RNase-A.

ii). Large scale preparation of plasmid DNA

Lysozyme/Triton X-100 lysis and polyethylene glycol 6000 precipitation

10mls of LB containing 50*ug*/ml ampicillin (Berk pharmaceuticals) was innoculated with 5*u*l of glycerol stock and shaken for 2-3 hours at 37°C. 0.5ml of these cultures was used to prepare glycerol stocks by addition of 0.5 volumes of glycerol and storage at -70°C. 400ml of LB/ampicillin was then innoculated with the remainder of this 10ml and left shaking at 37°C overnight. Cultures were spun at 7,000rpm for 10 minutes in a Sorvall RC-5B centrifuge using the GS3 rotor and pellets were resuspended in 4ml of 25% w/v sucrose, 50mM Tris pH 8 and 1mg/ml lysozyme, transferred to 50ml Sorvall tubes and left on ice for 15 minutes. EDTA was then added to 10mM and left on ice for another 10 minutes. 0.5 volumes of Triton solution (5.4% v/v Triton X-100, 270mM Tris pH 8, 340mM EDTA) was then added, mixed gently and the mixture replaced on ice for a further 30 minutes.

After spinning at 18,000rpm for 1 hour in a Sorvall RC-5B centrifuge using the SS-34 rotor the supernatent was retained and added to 1ml of 1M Tris pH 8 and 1ml of 4M

sodium chloride. This solution was then phenol/chloroform and chloroform extracted (spinning at 4,000rpm for 10 minutes in an IEC Centra-4R benchtop centrifuge. The phenol/chloroform extraction was shaken only gently).

Plasmid DNA was precipitated by the addition of 10% w/v polyethylene glycol 6000 which was dissolved at 37°C for 30 minutes and then left at 0-4°C for at least 1 hour. the precipitate was spun at 12,000g for 20 minutes and redissolved in 500*u*l of 0.1M Tris pH 8 and 0.2mg/ml RNase A and incubated at 37°C for 30 minutes. 1 volume of polyethylene glycol solution (20% w/v polyethylene glycol 6000, 1m sodium chloride, 10mM Tris pH 8, 1mM EDTA) was added and after leaving on ice for 1 hour the precipitate was spun at 12,000g for 15 minutes and resuspended in 400*u*l of 10mM Tris pH 8, 0.5M sodium chloride and left on ice for a further 30 minutes. After phenol and chloroform extraction the DNA was again treated with 0.125*u*g/ml RNase-A at 37°C for 1 hour. The DNA was then ethanol precipitated.

The purity and integrity of the plasmid DNA was always checked by restriction enzyme digestion and visualisation on agarose/ethidium bromide gels. Typical yields were approximately 750*u*g of DNA.

2.2.4. Restriction Enzyme Digests and Agarose Gels

Restrictions were carried out in 1 x the recommended Gibco BRL REact buffer with 4mM spermidine and typically 1 unit of enzyme per *u*g of DNA and incubated at 37°C for 1 hour unless stated. In order to digest PCR products, these were phenol/chloroform and chloroform extracted, ethanol precipitated and resuspended in water prior to digestion.

Agarose (Gibco BRL) or LGT agarose (IBI) gels between 1 and 2% w/v were made up using 1 x TBE or TAE and containing 0.5ug/ml ethidium bromide (Sigma). DNA or RNA samples were loaded in 2.5% w/v ficoll (Sigma) and 0.05% w/v orange-G (Sigma). Samples were typically electrophoresed at 50-100mA constant current in 1 x TBE or TAE and visualised using a UVP transilluminator.

2.2.5. Oligo labelling

Oligo labelling of DNA fragments was carried out according to the method of Fienberg and Vogelstein (Fienberg and Vogelstein, 1983). Appropriate restriction enzyme fragments were excised from LGT agarose gels and boiled in 3 volumes of dH_2O for 10 minutes.

Approximately 30ng of fragment was heat denatured at 100°C for 3 minutes and oligo labelling buffer was added before renaturation to a final concentration of 50mM Tris pH 8, 5mM $MgCl_2$, 0.1% v/v β-mercaptoethanol, 200*u*M dATP, dGTP and dTTP (Pharmacia), 200mM Hepes pH 6.6, 100*u*g/ml random hexanucleotides and 1*u*Ci/*u*l [α^{32}P] dCTP typically in a final volume of 40*u*l. Finally when cooled 6 units of Klenow were added and the mixture was incubated at room temperature overnight.

DNA probes were then spun through a 1ml Sephadex G50 (Pharmacia) column at 2,000rpm for 2 minutes in an IEC Centra 4R benchtop centrifuge and 1*u*l of the eluant was assayed by scintillation counting. Probes were typically >2 x 10^7 cpm.

2.2.6. Construction of Recombinant Plasmids

i). Vector Preparation

Typically 10*u*g of vector DNA was linearised with the appropriate restriction enzyme(s) and the efficiency of the enzyme(s) was tested on agarose/ethidium bromide gels. Linear vector was phenol/chloroform extracted and ethanol precipitated before resuspending in 90*u*l of 10mM Tris pH 8.3, 10*u*l of 10 x phosphatase buffer (Boehringer Mannheim) and 1 unit of calf intestinal alkaline phosphatase and incubating at 37°C for 30 minutes. The reaction was stopped by addition of 100*u*l of 2 x stop solution (1% SDS, 10mM EDTA and 200*u*g/ml proteinase-K) and incubation at 56°C for 30 minutes. Vector was then phenol/chloroform extracted and ethanol precipitated.

Bluescript-SK was not treated with alkaline phosphatase.

ii). DNA Fragment Purification

DNA fragments were purified using the Geneclean II kit (BIO 101 Inc). Purification was carried out according to the protocol obtained with the kit.

Briefly the appropriate fragments were cut from TAE gels in minimal volumes of agarose and dissolved in 3 volumes of sodium iodide at 50°C for 5 minutes. 5*u*l of glassmilk was added and the mixture left on ice for 5 minutes before pelleting the matrix at 12,000g for 10 seconds. The glassmilk matrix was resuspended in 500*u*l of -20°C New wash and respun and resuspended 3 times. After removing all the wash solution the pellet was resuspended in 10*u*l of dH_2O and incubated at 50°C for 3 minutes, then spun at 12,000g for 30 seconds. DNA in the supernatent was retained for ligation.

iii). Ligations (All solutions were sterile).

Approximately 200ng of vector DNA and 5 x the molar concentration of insert DNA were ligated in 10*u*l volumes containing 1 x ligation buffer (Boehringer Mannheim) and 1 unit of T4 DNA ligase at 14°C overnight. Molecular weight marker DNA (λ Hind-III digest) was also ligated and run on agarose/ethidium bromide gels to test the efficiency of the ligase.

iv). Recombinant Selection (all solutions were sterile).

Ligated DNA was used to transform competent cells (see 2.2.2.) and white colonies were picked and grown for 2 hours at 37°C in 200*u*l of LB containing 50*u*g/ml ampicillin (Berk) in 96 well plates. 0.5*u*l of each clone was then spotted onto Hybond N. To the remainder of each clone 2 drops of glycerol was added and these were frozen at -70°C. 1ng of Bluescript DNA was also used to transform competent cells in order to test their DNA uptake and colour selection efficiency.

Nylon filters were prepared for hybridisation with 5 minute washes, twice in 0.5M sodium hydroxide, twice in 1M Tris pH 7.5, twice in 1.5M sodium chloride and 0.5M Tris pH 7.5 and finally for 5 minutes in 2 x SSC. The DNA was then cross linked to the filters by baking for 2 hours at 80°C. 1*u*l of DNA insert was oligo labelled (see 2.2.5.) and used to probe filters, which were prehybridised in 6 x SSC, 1 x Denhardts and 100*u*g/ml denatured herring sperm DNA at 65°C for 30 minutes. Hybridisation was carried out in the

same solution plus denatured probe at 65°C overnight. Filters were washed twice in 1 x SSC and 0.5% SDS at 65°C for 30 minutes prior to autoradiography for 1 hour. Positive clones were grown up from glycerol stocks for miniprep and restriction enzyme analysis.

2.2.7. Transfections (All solution were filter sterilized).

8 x 10^5 cells (1 x 10^5 cells in conditioned medium for stable cell lines) on a 90mm plate were transfected by the calcium phosphate procedure of Gorman (Gorman, 1985). Briefly cells were transferred to DMEM supplemented with 10% v/v foetal calf serum 2 hours before transfection. Plasmids and herring sperm carrier DNA were sterilized by ethanol precipitation and 70% v/v ethanol wash and typically 10*u*g of each was added to each 90mm plate. 4 hours after transfection cells were washed twice in PBS and the standard growth medium was replaced. Cells were harvested by scraping from the plates 48 hours after transfection for the preparation of RNA or for CAT assays.

3T3 cells were transfected with the following modifications to increase the efficiency of DNA uptake. Calcium/phosphate/DNA precipitates were added to plates from which the media had been removed and left at 37°C for 15 minutes before replacing the medium. After 4 hours the medium was removed and replaced with 2.5ml of HBS and 15% v/v glycerol for 2 minutes at 37°C. The cells were then washed as above.

2.2.8. Southern Blotting

Agarose/ethidium bromide gels were prepared for Southern blotting by washing for 15 minutes in 0.5M sodium hydroxide and 1.5M sodium chloride, 15 minutes in 3M sodium chloride and 0.5M Tris pH 7 and 5 minutes in 2 x SSC. Blotting onto nylon filters was carried out overnight using 20 x SSC as blotting buffer. DNA was cross linked to blots by UV irradiation on a UVP transilluminator for 5 minutes. Filters were prehybridised in 6 x SSC, 10% w/v dextran sulphate (Pharmacia), 0.1% w/v SDS, 5 x Denhardts and 100*u*g/ml denatured herring sperm DNA (Sigma), for 2 hours at 65°C in a Hybaid MAXI hybridisation oven. Hybridisation to denatured, oligo labelled probe (2.2.5.) was carried out in the same solution overnight at 65°C.

30 minute washes were carried out typically at 65°C in 1 x SSC and 0.2% w/v SDS followed by decreasing concentrations of SSC to 0.1 x or until sufficient washing was achieved for autoradiography.

2.2.9. DNA Sequencing

DNA sequencing was performed using double stranded DNA templates with the 'Sequenase' (Tabor & Richardson, 1987) version 2.0 kit (United States Biochemical) and according to the instructions supplied with the kit. DNA was obtained by large scale plasmid preparation (2.2.3.ii.).

Sequencing products were seperated by electrophoresis through 6% polyacrylamide/urea gels prepared using sequagel solutions (National Diagnostics). These gels were typically run at 30W in 1 x TBE buffer.

RNA

2.2.10. Riboprobe Transcription and Hybridisation (see 3.2.2.)

Recombinant Bluescript-SK plasmids were linearised with a restriction enzyme cutting downstream of the insert with relation to the RNA polymerase binding site to be used. The efficiency of the enzyme was tested on agarose/ethidium bromide gels. Typically 40*u*g of plasmid was linearised, phenol/chloroform extracted and ethanol precipitated. To test the efficiency of transcription linear plasmid was incubated with RNA polymerase (T3 or T7) in the absense of radioactive label. 1*u*g of linear plasmid was incubated in 1 x RNA polymerase buffer (Boehringer Mannheim), 20mM DTT, 0.5mM ATP, CTP, GTP and TTP (Pharmacia), 20 units of Rnasin and 10 units of RNA polymerase at 37°C for 1 hour. Half of this reaction volume was removed and incubated at 37°C for a further 15 minutes in the presence of 18 units of DNase-I and a further 20 units of Rnasin. These samples along with linear plasmids were analysed on agarose/ethidium bromide gels and if efficient transcription was achieved the transcription reaction was repeated but with 30*u*M CTP and 2*u*Ci/*u*l [α^{32}P] CTP. After addition of 200*u*l of 8mM sodium chloride (DEPC treated) the probe was phenol/chloroform extracted, ethanol precipitated and resuspended in 400*u*l of dH_2O (DEPC treated) for scintillation counting. RNA probes were typically > 10^8 dpm.

RNA probes were hybridised to Northern blots (2.2.13.) prepared from non-denaturing agarose/ethidium bromide/TBE gels. RNA was cross linked to nylon blots by UV irradiation on a UVP transilluminator for 5 minutes then pre-hybridised in 5 x SSC, 60% v/v formamide, 1% w/v SDS, 20mM phosphate buffer (12mM disodium hydrogen orthophosphate, 8mM sodium dihydrogen orthophosphate) pH 6.8, 100*u*g/ml denatured herring sperm DNA (Sigma) and 100*u*g/ml total yeast RNA (Sigma) for 2 hours at 65°C in a Hybaid MAXI hybridisation oven. RNA probe hybridisation was carried out in the same solution plus 7% w/v dextran sulphate (Pharmacia) overnight at 65°C.

30 minute washes were carried out typically at 65°C in 1 x SSC and 0.2% w/v SDS followed by decreasing concentrations of SSC to 0.1 x or until sufficient washing was achieved for autoradiography.

2.2.11. RNA Purification

All solutions were prepared using DEPC (Sigma) treated dH_2O and all glassware was baked at 160°C overnight before use.

Total RNA was isolated from tissues or transfected or untransfected tissue culture cells by a rapid guanidinium isothiocyanate method (Wilkinson, 1988) with the following modifications. Whole PBS washed cell pellets or tissue samples were resuspended in at least 20 volumes of guanidinium lysis buffer. After grinding in liquid nitrogen tissues were homogenized using an Ultra Turrax T25 and dispersing tool. Sodium lauryl sarcosinate was then added to a final concentration of 0.5% w/v and the lysates were centrifuged at 5,000g for 10 minutes at room temperature to remove debris. Tissue culture cells were simply vortexed. Typically, 1ml of lysate was then layered onto a 220*u*l cushion of 5.7M caesium chloride and 0.1M EDTA, pH 5 and spun at 55,000 rpm for 3 hours at 17°C in a TL-100 Beckman ultracentrifuge. Pellets were resuspended in 150*u*l of H_2O and ethanol precipitated overnight.

2.2.12. Reverse Transcription-PCR (RNA-PCR) (RT-PCR)

Prior to reverse transcription mRNA was typically treated with 15 units of DNase I and 20 units of Rnasin in 50*u*l of dH_2O for 1 hour at 37°C to remove traces of genomic DNA. mRNA was then extracted with phenol/chloroform and chloroform and ethanol precipitated.

Reverse transcription was carried out typically in a 20*u*l reaction volume containing 1*u*g of total mRNA as determined by spectrophotometry, (and confirmed by agarose, ethidium bromide gel analysis) 0.25 units of random hexanucleotides, (allows priming of all transcripts so PCR amplifications can subsequently be carried out with various sets of specific oligonucleotides) 0.5mM dATP, dCTP, dGTP and dTTP (Pharmacia), 1 x Taq DNA polymerase buffer (Amersham), 12mM DTT, 100 units of M-MLV reverse transcriptase and 20 units of Rnasin. Samples were heated to 65°C for 5 minutes before oligonucleotide addition, cooled before enzyme addition and incubated at 37°C for 1 hour.

Oligonucleotide pairs were designed to have similar Tm values (Tm = (G+C x 4) + (A+T x 2)) and to have at least 2 G or C bases at their 3' end. Oligonucleotides were deprotected at 55°C overnight and stored at -20°C. After ethanol precipitation all oligonucleotides were resuspended in dH_2O and diluted to a stock concentration of 5*u*M for competetive PCR assays (Innis et al., 1990).

PCR analysis of cDNA was carried out by the method of Kawasaki (Kawasaki, 1990). Melting, annealing and elongation times and temperatures were designed to suit individual assays. Breifly, melting times were increased for longer fragments, oligonucleotides were annealed at approximately 5°C below their Tm and elongation times were 30+ seconds (45 seconds for fragments 400-500 base pairs long and 3 minutes for 1-3 kilobase pair fragments). Reactions were typically in a 50*u*l volume containing 30 picomoles of each primer, 0.1mM dATP, dCTP, dGTP and dTTP, 1 x Taq DNA polymerase buffer (Amersham) and 2 units of Taq DNA polymerase. These reaction were overlayed with 50*u*l of liquid parrafin (B.P.). Following PCR amplification using a Hybaid thermal reactor samples were extracted with 50*u*l of chloroform and loaded onto agarose, ethidium bromide gels.

In preliminary experiments PCR amplifications were carried out using different cDNA dilutions and numbers of cycles to identify conditions in which the PCR product signal was linearly related to the input cDNA.

2.2.13. Northern Blotting

All solutions were prepared using DEPC (Sigma) treated dH_2O and all glassware was baked at 160°C overnight before use.

RNA samples were resuspended in 33% v/v deionised formamide, 12% v/v formaldehyde, 1 x MEA, 8% v/v glycerol and 0.05% w/v bromophenol blue (Electran). Samples were heated to 65°C for 10 minutes before loading. Typically approximately 15*u*g of RNA per lane was electrophoresed on a 20 x 20cm, 1% w/v agarose gel containing 7% v/v formaldehyde and 1 x MEA (Lehrach et al., 1977). Gels were run at 300V constant voltage for 3 hours and subsequently transferred to nylon filters overnight using 20 x SSC

as blotting buffer.

RNA was cross linked to blots by UV irradiation on a UVP transilluminator for 5 minutes then pre-hybridised in 7% w/v SDS, 5 x SSC, 10 x Denhardts solution and 100*ug*/ml denatured herring sperm DNA (Sigma) for 2 hours at 65°C in a Hybaid MAXI hybridisation oven. Hybridisation to denatured, oligo labelled probe (2.2.5.) was carried out in the same buffer plus 10% w/v dextran sulphate (Pharmacia) overnight at 65°C.

30 minute washes were carried out typically at 65°C in 1 x SSC and 0.2% w/v SDS followed by decreasing concentrations of SSC to 0.1 x or until sufficient washing was achieved for autoradiography.

Protein

2.2.14. CAT assays

CAT assays were carried out according to the procedure of Gorman (Gorman, 1985).

2.2.15. Western Blotting

Tissues and cells were homogenised using a hand held homogeniser in 60mM Tris pH 6.8, 10% v/v glycerol, 2.3% w/v SDS 5% v/v β-mercaptoethanol and 0.25% w/v bromophenol blue (Electran) and were boiled prior to loading on a 12.5% w/v polyacrylamide gel. Tissue samples were also spun at 12,000g for 5 minutes to remove debris. Gels were made using 30% w/v acrylamide, 0.8% w/v bisacrylamide Protogel (National Diagnostics) in 0.4M Tris pH 8.8 and 0.1% w/v SDS and set with 0.04% w/v APS and 0.05% v/v TEMED (Sigma). The 12.5% gel was overlayed with dH_2O saturated butanol while setting. The 5% stacking gel also contained Protogel but with 0.125M Tris pH 6.8 and 0.1% SDS, 0.1% APS and 0.1% TEMED.

Protein samples were electrophoresed typically for 8 hours at 45mA constant current in 0.2M glycine (Sigma), 0.1% w/v SDS and 25mM Tris running buffer. Protein samples were equalised by visualising coomassie brilliant blue stained gels. Gels were soaked overnight in 50% v/v methanol, 6.5% v/v acetic acid and 2% w/v coomassie stain (Sigma). Destaining was carried out in the above solution minus the stain. Typically approximately 20*u*g of protein were loaded per lane.

Transfer to nitrocellulose filters was carried out overnight at 210mA constant current in 20% v/v methanol, 0.64M glycine and 80mM Tris transfer buffer (Towbin et al., 1979).

Filters were blocked for 1 hour in 5% w/v fat free dried milk (Marvel) and 0.2% v/v Tween-20 (Sigma) in TBS, washed 3 times in wash buffer (TBS with 2% w/v Marvel and 0.05% v/v Tween-20) and once in TBS for 5 minutes each time. KSm5 or KSm4 (ascites) primary antibody (Williams et al., 1986) was diluted 1/2,000 in wash buffer and filters were incubated in antibody for 90 minutes. The 4 washes were repeated as above and a secondary, horseradish peroxidase conjugated, rabbit anti-mouse immunoglobulins

(Dako) were diluted 1/1,000 in wash buffer and filters were again incubated for 90 minutes and washed as above. Autoradiography was carried out using the enhanced chemiluminescence (ECL) Western blotting detection system (Amersham) and Hyperfilm MP.

2.2.16. Immunoflourescence

Cells were seeded on glass coverslips and grown for 48 hours before fixing. Cells were air dryed and fixed in ice cold methanol for 3 minutes and were subsequently kept at room temperature and prevented from drying. Blocking was in PBS containing 2% w/v casien (Sigma) and 0.05% v/v Tween-20 (Sigma) for 5 minutes. KSm5 and KSm4 primary antibody was diluted 1/4,000 in blocking solution and 50*ul* was left on cells for 30 minutes. After 3 brief PBS washes cells were washed for 20 minutes in 50*ul* of PBS and again washed 3 times briefly in PBS. Rhodamine and Flourescien conjugated rabbit anti mouse secondary immunoglobulins (Dako) were diluted 1/2,000 in blocking buffer and again 50*ul* was left on the cells for 30 minutes. The above PBS washes were repeated and the cells were air dried and mounted in aquamount mountant (BDH). Cells were photographed using a Nikon Diaphot microscope and camera with the x40 objective under phase contrast and immunoflourescence optics.

Cell Lines

2.2.17. Maintenance and Differentiation of Cell Lines

Glassware was baked at 160°C overnight before use.

ND3, ND8, HeLa I, HeLa M, 3T3, LTA and PC12 cells were grown in DMEM (Gibco BRL with 0.11g/L sodium pyruvate) supplemented with 10% v/v foetal calf serum (Gibco BRL, mycoplasma and virus screened). N18Tg2 cells were grown in RPMI 1640 medium (Gibco BRL with L-glutamine) supplemented with 10% v/v foetal calf serum and 0.15% w/v sodium bicarbonate. ND7, ND11, ND15 and ND21 cells were grown in L15 (Liebovitz) medium (Gibco BRL with L-glutamine) supplemented with 10% v/v foetal calf serum, 0.35% w/v sodium bicarbonate and 0.35% w/v glucose.

To induce the ND7 cells to cease dividing and undergo morphological differentiation they were transferred to serum free medium consisting of a 1:1 mix of DMEM and Nutrient mix Ham's F12 (Gibco BRL with L-glutamine) supplemented with 5*u*g/ml human transferrin (Gibco BRL), 250ng/ml bovine insulin (Gibco BRL) and 30nm sodium selenite (Gibco BRL) for 3 days (Suburo et al., 1992). F9 teratocarcinoma cells were maintained using plasticware coated with 1% w/v gelatin (Sigma) and grown in DMEM and 10% v/v foetal calf serum. F9 cells were differentiated on untreated plasticware by the addition of 1mM cAMP (Sigma) and 0.1*u*M retinoic acid (Sigma) to the growth medium. This medium was changed every 48 hours. O1A1 embryonal carcinoma cells were grown in a 1:1 mix of DMEM and F12 supplemented with 7.5% v/v foetal calf serum. These cells were differentiated into a spectrum of cell types including cardiac muscle by aggregation in petri-grade culture dishes in the presence of 1% DMSO for 3 days and in tissue culture grade dishes for 5 days. Differentiation to neurons, glia and fibroblast-like cells was achieved by aggregation and culture as above but in the presence of 500nM retinoic acid (McBurney et al., 1982).

Cells were typically passaged every 3-4 days and not used after 30 passages. 2.5% w/v trypsin in saline (Gibco BRL) was used diluted 1:10 in versene (Gibco BRL). For long term cold storage cells were trypsinised for 5 minutes at 37°C, resuspended in 10ml of the standard medium and spun at 1,000rpm for 5 minutes in an IEC Centra 4R benchtop

centrifuge. Cell pellets were resuspended in 3.2ml of medium and 0.4ml of DMSO before freezing overnight in cryostat tubes at -70°C in dry ice. Tubes were then transferred to liquid nitrogen. Cells were thawed rapidly, resuspended and spun as above and again resuspended in 5ml of medium and returned to small tissue culture flasks.

2.2.18. Selection of Stable Cell Lines (see chapter 5)

Cells were transfected as in 2.2.9 if constructs contained the G418 (*neo*) resistance (aminoglycoside phosphotransferase) gene (Southern & Berg, 1982), or if not with 1.5*u*g of a plasmid containing the *neo* resistance gene (M5G*neo*), 15*u*g of the SmN cDNA construct and 10*u*g of carrier DNA.

Following transfection colonies were maintained in conditioned medium until only colonies resistant to G418 sulphate (Genetecin, Gibco BRL) remained. Conditioned medium was prepared by removal from sub-confluent cells, filter sterilization and supplementation with 50% of the standard concentration of foetal calf serum. This medium was replaced every two days. Resistant colonies were selected and expanded for assay by Western blotting of protein extracts. 3T3 and ND7 cells were selected in 200 and 100*u*g/ml Genetecin respectively and control, non-transfected cells grown in parallel and subjected to selection showed 100% cell death under these conditions. Isolated colonies were removed from 90mm diameter plates by removing the growth medium and incubating the colony in 0.5*u*l trypsin solution for 1 minute at 37°C. Cells were then removed in 5*u*l of growth medium and serially diluted in 96 well plates in 200*u*l of medium containing Genetecin.

CHAPTER 3

RESULTS

SmN, SmB and SmB' expression

3. Introduction

SmN may play a role in a number of splice choices. In order to test this hypothesis it was necessary to quantitate SmN/SmB ratios in a number of cells and tissues. Rodent samples were to be used in subsequent splicing assays as the absence of SmB' in these samples allows SmN levels to be more readily related to ratios of alternatively spliced products. The distribution of Sm peptides has also been examined in human tissues however. Characterisation of SmN expression may also provide further clues as to the function of SmN in alternative splicing or otherwise.

3.1. Protein

All of the data in this section is obtained using a pair of monoclonal anti-Sm antibodies derived from the mouse strain MRL/lpr (Williams et al., 1986). These antibodies, KSm4 and KSm5 have been shown to recognise the same epitopes as human antibodies specifically found on SmB, SmB' and SmD. KSm4 recognises an epitope found on SmB, SmB' and SmD as well as a 60 kilodalton protein found in rabbit thymus; this epitope is not found on SmN, while KSm5 recognises an epitope found on SmB, SmB' and SmN (Williams et al., 1985) (see 3.1.5.). The SmD proteins are not seen in any of the following blots as this protein has migrated off the end of the gels. The KSm5 antibody recognises a repeated proline rich epitope towards the carboxyl terminal end of SmB, SmB' and SmN. This region of these proteins is in fact recognised by most human and mouse anti-Sm antibodies (Elkon et al., 1990; Rockeach et al., 1990). KSm4 also recognises an epitope found on SmD which appears to be conformation dependent or due to a common post translational modification rather than primary sequence (Williams et al., 1990).

3.1.1. Quantitative nature of ECL detection

The ECL Western blotting detection system involves the use of a high sensitivity non-radioactive assay. As the KSm5 antiboby had not previously been used in conjunction with this assay it was desirable to assess the quantitative nature of the assay. As can be seen from this set of dilution series' using rodent cells the assay does appear to be quantitative, suggesting that SmN/SmB in F9 and ND7 cells is approximately 0.25 and 0.1

Fig 3.1.1.

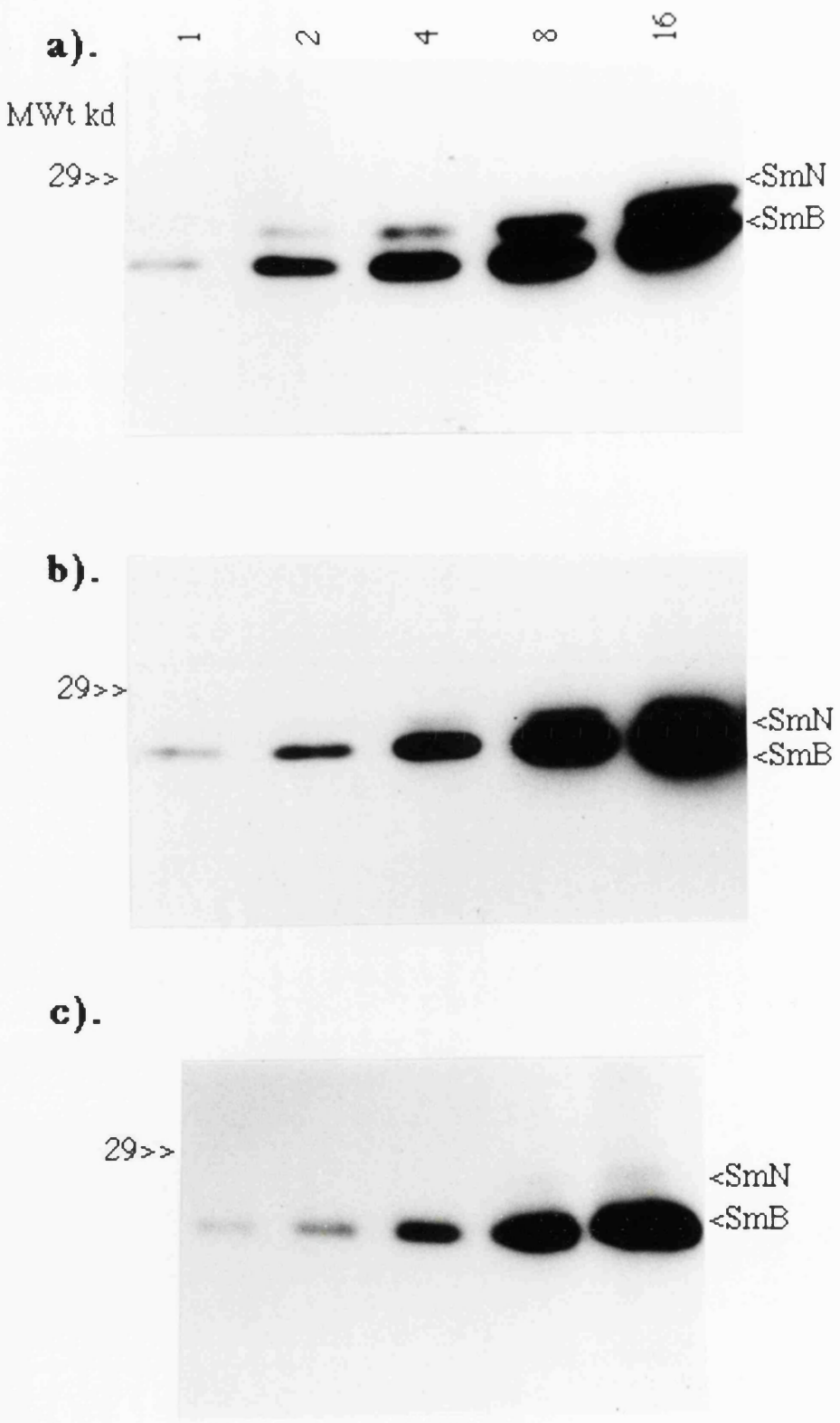
Quantitative Nature of ECL Detection

Western blots with the KSm5 antibody. Numbers represent volume of protein extract loaded in each track.

a). F9 EC cells.

b). ND7 neuronal cells.

c). 3T3 mouse fibroblasts.


a).
1
2
4
8
16
MWt kd
29>>
<SmN
<SmB
b).
29>>
<SmN
<SmB
c).
29>>
<SmN
<SmB

respectively. These estimated SmN levels are probably lower than this however as SmN has an additional (third) proline-rich motif at the 3' end which is also present on the human SmB' (Schmauss & Lerner, 1990) (see fig.1.4.) and is recognised by the KSm5 antibody (Williams et al., 1990).

The apparent molecular weights of SmN and SmB are approximately 29.5 and 28.5 kilodaltons respectively although the deduced molecular weights are 24.6 and 23.7. This observation has been reported previously and may be, in part due to post translational modification *in-vivo* (van Dam et al., 1989).

3.1.2. SmN expression in rodent tissues and cells

SmN has previously been reported to be expressed in a tissue specific manner (McAllister et al., 1989). SmN is abundant in brain tissue, expressed at low levels in the heart and at very low, but detectable levels in skeletal muscle as well as liver (see fig.3.1.2.a.). An attempt was made to equalise SmN (fig.3.1.2.a.i.) and SmB levels (fig.3.1.2.a.ii.) in each tissue to give a clearer indication of the SmN/SmB ratios. In fig.ii. approximately 50, 25 and 10*ug* of skeletal and cardiac muscle and brain protein were loaded respectively. Levels of SmN in the heart are not dramatically higher than those observed in skeletal muscle. This suggests that a subtype of cells in the heart may be expressing high levels of SmN, rather than all heart cells expressing low levels of SmN (see brain *in-situ* fig.3.2.2.). These cells may in fact be neurons. Heart tissue can be dissected almost free of neural and glial elements however. Cardiac nerve axons should not contribute any nuclear components such as mRNA and Sm proteins but some nuclear components may be derived from parasympathetic ganglia. Skeletal muscle is not easily dissected free of neural elements and as this tissue exhibits very low levels of SmN expression on Western blots. Therefore it appears unlikely that the heart SmN is derived from neuronal cells. Also a rat heart cell line, H9 (Kimes & Brandt, 1976) does reflect the *in-vivo* pattern of SmN and SmB expression in heart (personal communication, Dr. D. Gerrelli, UCMSM). The distribution of SmN in the heart is unclear however, so SmN function is best studied in neuronal cells where the protein is known to be expressed at naturally high levels *in-vivo* (Schmauss et al., 1992).

Fig 3.1.2.

SmN Expression in Rodent Tissues and Cells

Western blots with the KSm5 antibody.

a). SmN expression relative to SmB in adult rat tissues. It was attempted to equalise for each band in separate experiments.

i). SmN equalised

ii). SmB equalised

b). SmN expression relative to SmB in neuronal and non-neuronal cell lines.

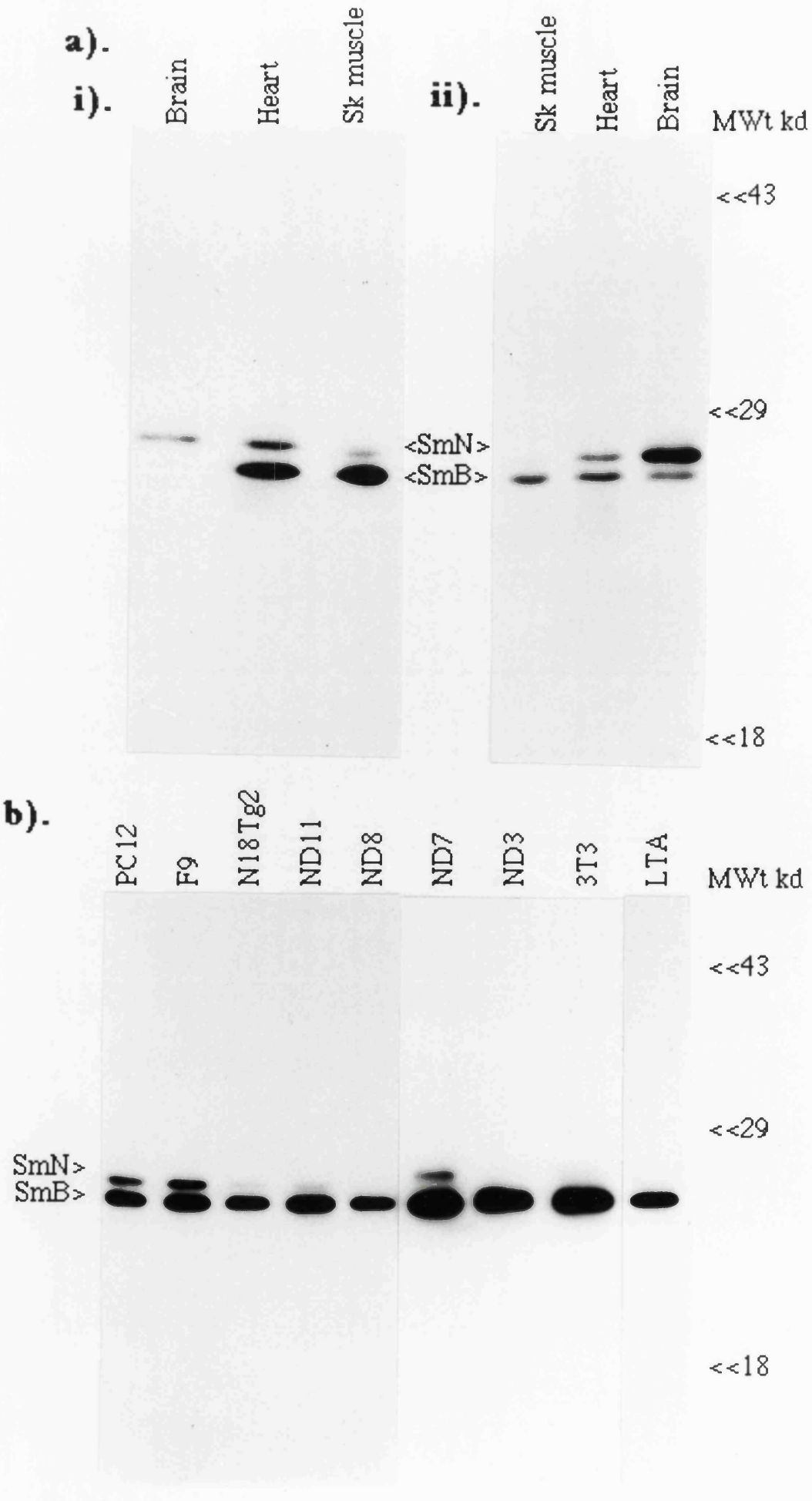
Neuronal = N18Tg2, ND11, ND8, ND7 and ND3.

Non-neuronal = PC12 rat adrenal pheochromocytoma

F9 embryonal carcinoma.

LTA mouse hepatocytes

3T3 mouse fibroblast.


a).
i).
Brain
Heart
Sk muscle
ii).
Sk muscle
Heart
Brain
MWt kd
<<43
<<29
<<18
<SmN>
<SmB>
b).
PC12
F9
N18Tg2
ND11
ND8
ND7
ND3
3T3
LTA
MWt kd
<<43
<<29
<<18
SmN>
SmB>

The ND series of cell lines are hybrid cell lines derived from neonatal rat DRG neurons fused with the mouse neuroblastoma N18Tg2. They provide a source of large quantities of clonal neuronal material, not available from other sources. They also display properties of sensory neurons and can therefore serve as a model for studying the regulation of expression and function of sensory-neuron specific neuropeptides (Wood et al., 1990) and proteins such as SmN. SmN is expressed in ND7 cells although at a low level (fig.3.1.2.b.). Low level expression can also be seen in ND11 and the parental N18Tg2 neuroblastoma cells while ND3 and ND8 cells as well as the non-neuronal 3T3 cells do not show detectable levels of SmN. PC12 cells are NGF responsive rat adrenal pheochromocytoma cells (Greene & Tischler, 1976). The PC12 cells and the F9 embryonal carcinoma cell line exhibit the highest levels of SmN expression. No cell lines were found to reflect the levels of SmN expression found in brain tissue (fig.3.1.2.b.).

3.1.3. SmN expression during rat tissue development

SmN can be seen to be developmentally regulated in the rat brain (fig.3.1.3.a.). A dramatic switch occurs during the days surrounding birth and SmN relative to SmB expression continues to rise until SmN greatly exceeds SmB expression in the adult brain. In the heart however SmN/SmB expression does not appear to be regulated after birth with the low level expression of SmN established by post natal day two (fig.3.1.3.b.). This also appears to be the case from embryonic day sixteen (Grimaldi et al., 1993). SmN therefore appears to be differentially regulated in two distinct tissues from the same gene (Gerrelli et al., 1992). This also suggests that SmN in the heart is not derived from neural cells as peripheral nerves are shown to express SmN in fig.3.2.2. *In-situ* hybridisation in heart tissue must be carried out to determine the distribution of heart SmN mRNA however.

Fig 3.1.3.

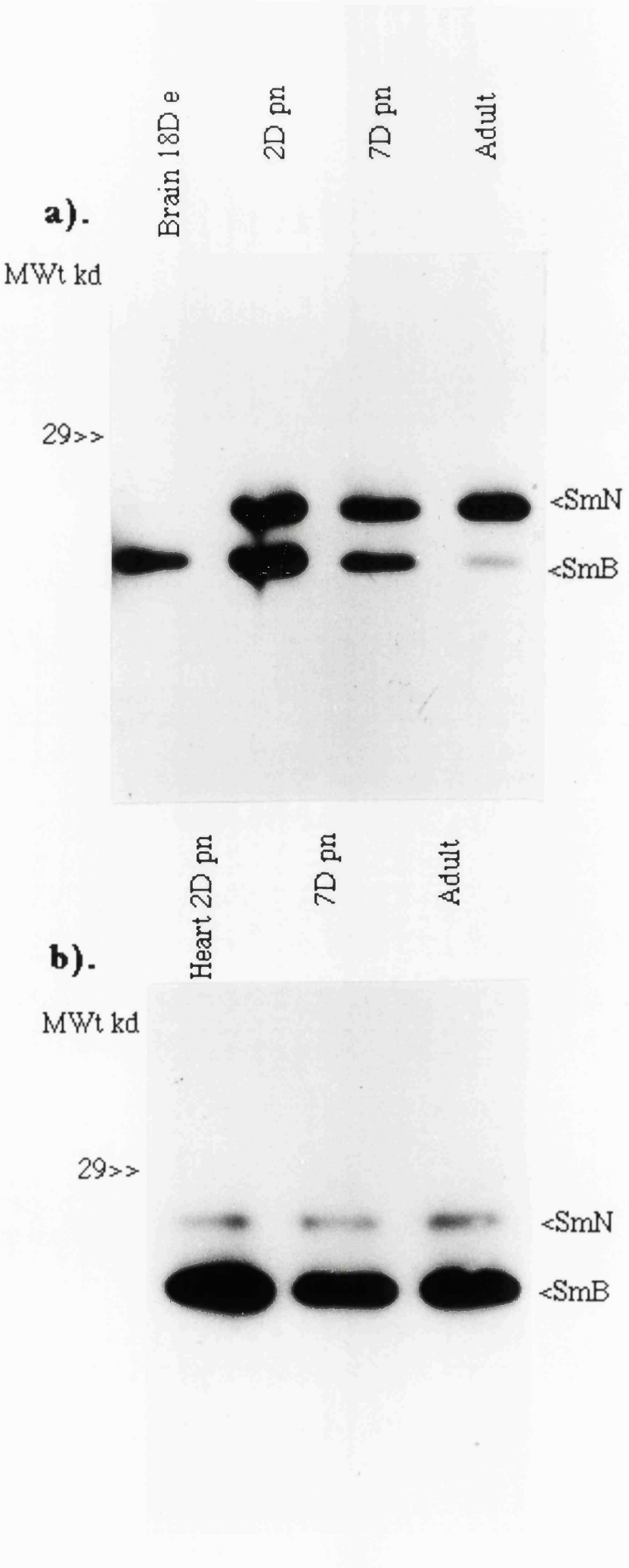
SmN Expression During Rat Tissue Development

Western blots with the KSm5 antibody.

D pn = days post natal D e = embryonic days.

a). Brain development.

b). Heart development.


a).
Brain 18D e
2D pn
7D pn
Adult
MWt kd
29>>
<SmN
<SmB
b).
Heart 2D pn
7D pn
Adult
MWt kd
29>>
<SmN
<SmB

3.1.4. SmN expression during rodent cell differentiation

ND7 neuronal cells can undergo morphological differentiation involving cessation of proliferation and neurite outgrowth (Wood et al., 1990; Suburo et al., 1992). The process of ND7 differentiation does not influence SmN or SmB expression (fig.3.1.4.c.).

Embryonal carcinoma (EC) cells are the pluripotent stem cells of murine teratocarcinomas. These EC cells can proliferate and differentiate in cell culture providing a system for the study of some early developmental events. Following treatment with NGF (nerve growth factor) and cAMP F9 EC cells will differentiate into endoderm-like cells (Liesi et al., 1983). Differentiation of F9 cells leads to downregulation of SmN relative to SmB (fig.3.1.4.a.) (Sharpe et al., 1990a). The level of SmB expression also appears reduced during differentiation possibly reflecting the reduced requirement for mRNA processing turnover in quiescent cells.

O1A1 EC cells were isolated from male C3H/He strain mouse embryos (McBurney & Rogers, 1982) and can be differentiated into a number of tissue types. DMSO treatment leads to morphological differentiation into approximately 10% cardiac cells, skeletal muscle, and other cell types, but no neurons or glia. Retinoic acid (RA) treatment however leads to morphological differentiation into neurons, astroglia and fibroblasts but not muscle. These differentiated cells have subsequently been shown to express a number of specific markers, confirming their differentiated phenotype (Edwards et al., 1983; Jones-Villeneuve et al., 1983). Based on the differentiated phenotype of the O1A1 cells it might be expected that SmN/SmB levels would increase and decrease in the RA and DMSO treated cells respectively but the opposite effect is seen (fig.3.1.4.b.). RA induces a dramatic reduction in the level of SmN expression while DMSO treatment results in an increased level of SmN expression.

The F9 and O1A1 cells do not appear to reflect *in-vivo* patterns of expression during development with respect to SmN (see 3.1.3.). These cells do show altered SmN/SmB ratios during differentiation however and could concievably be used as models in which to study SmN function. An extensive cascade of altered expression is likely to be responsible for differentiation of these cells. e.g. CALC-I expression is switched off during F9 differentiation (Evain-Brion et al., 1984) so cells stably transfected with SmN cDNA and expressing SmN present a favoured model for this purpose (see chapter 5.).

Fig 3.1.4.

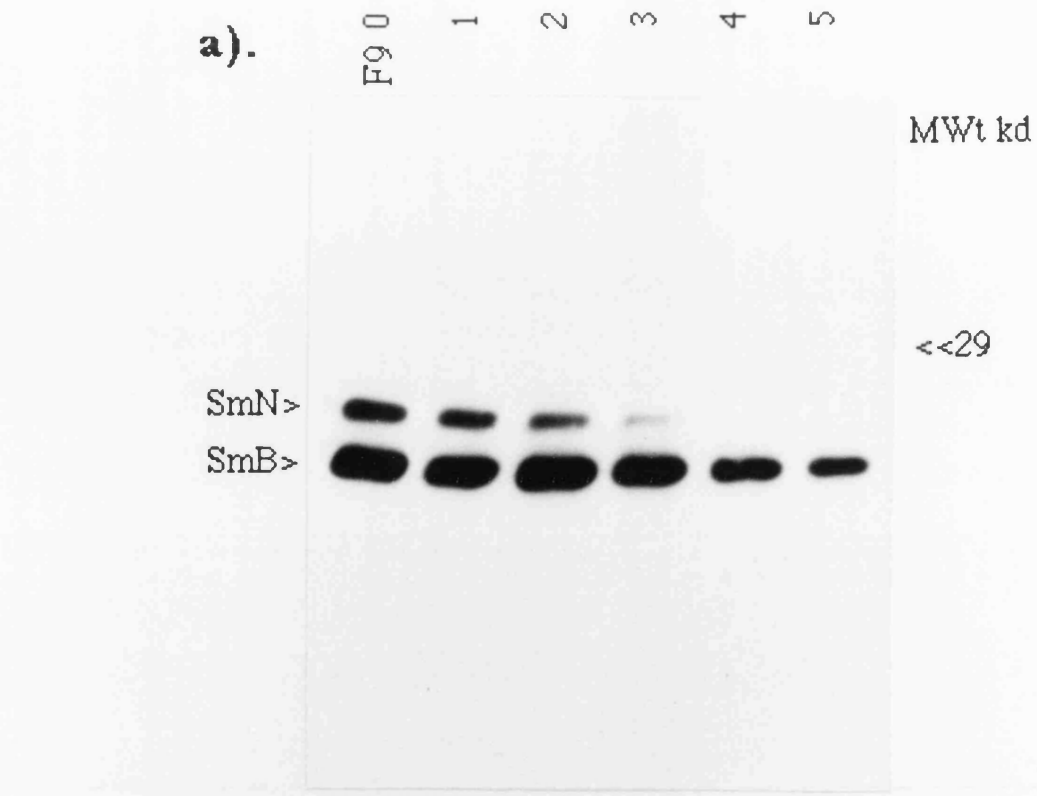
SmN Expression During Rodent Cell Differentiation

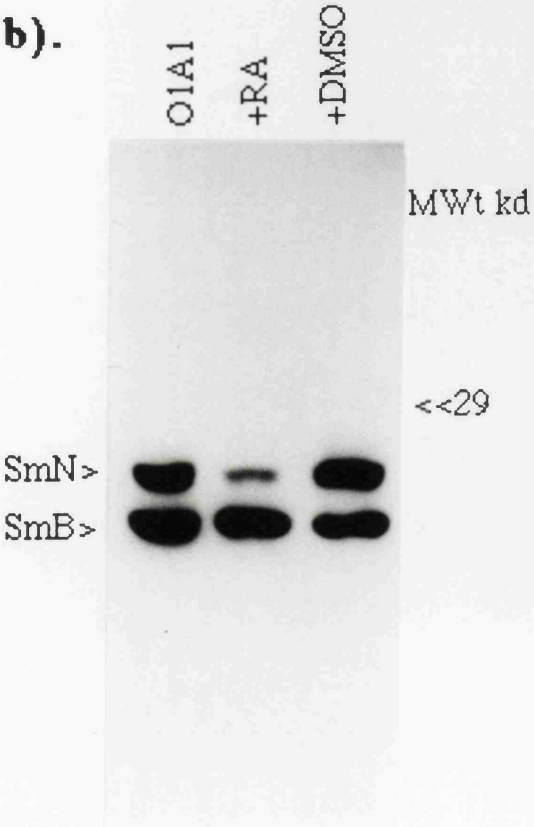
Western blots with the KSm5 antibody. Numbers represent days of growth in differentiation media.

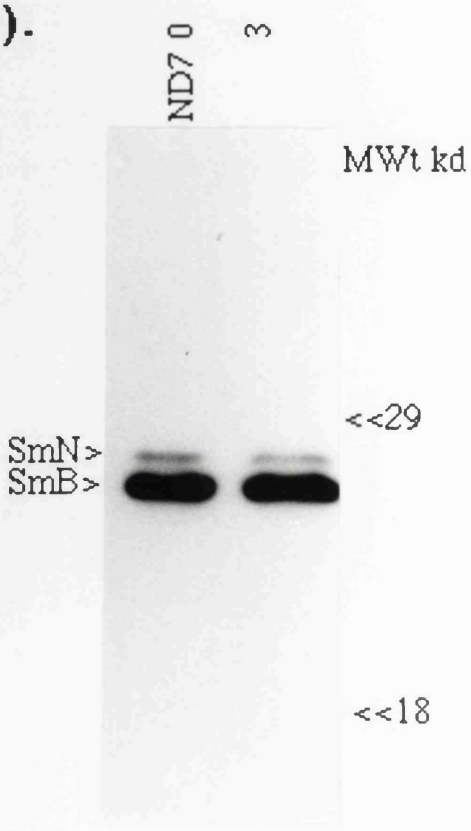
a). F9 EC cell differentiation.

b). O1A1 EC cell differentiation.

c). ND7 cell differentiation.


a).
F9 0
1
2
3
4
5
MWt kd
<<29
SmN>
SmB>


b).
O1A1
+RA
+DMSO
MWt kd
<<29
SmN>
SmB>


c).
ND7 0
3
MWt kd
<<29
SmN>
SmB>
<<18

3.1.5. SmN expression in human tissues and cells

SmB' which is closely related to SmN and SmB (see fig.1.4.), being an alternatively spliced form of SmB, is not expressed in rodent cells but is expressed in human cells (Schmauss & Lerner, 1990). Both SmN and SmB' proteins are 240 amino acids in length and have previously been reported to co-migrate on one-dimensional gels as well as sharing epitopes. These two proteins can be resolved by electrophoresis through two dimensional gels (Sharpe et al., 1989a) and the proteins can in fact be resolved by electrtophoresis through one-dimensional gels over an increased migration distance. The KSm5 antibody recognises a shared epitope on SmB, SmB' and SmN while the KSm4 antibody recognises a shared epitope on SmB, SmB' and SmD but not on SmN (Williams et al., 1986) (fig.3.1.5.a.). SmN can be seen to migrate at a slightly slower rate than SmB' (fig.3.1.5.b. and c.). This allows determination of the distribution of these Sm proteins in human cells and tissues. Duplicate Western blots were also incubated with KSm4 antibody to confirm the nature of the slowest migrating band as SmN. The epitopes recognised by the KSm antibodies do appear to be conserved between species as expected due to the 100% conservation between rodent and human SmN and SmB proteins (Schmauss et al., 1989; Schmauss & Lerner, 1990) (fig.3.1.5.a.).

SmN has a similar distribution in fetal human and in adult rodent tissues being expressed at an elevated level in brain relative to other tissues (fig.3.1.5.b.). All available human brain samples were fetal samples however and were derived from 11 to 17 week fetuses. These samples do not exhibit a trend of regulated Sm expression as in the rodent (fig.3.1.5.c.) (see 3.1.3.). SmB' does not appear to be expressed in a tissue specific manner in these fetal samples. SmB' mRNA has previously been reported to be tissue specific being relatively low in the brain (McAllister et al., 1989; Schmauss et al., 1992). This suggests that SmN may replace SmB' as well as SmB in this tissue. It is unclear whether levels of SmB or SmB' protein decline in human brain during later development from the samples available. The adult human heart samples do suggest developmental regulation however possibly with a differential distribution between ventricle and atria (fig.3.1.5.c.). It was also noted during these studies that the KSm5 antibody shows a stronger signal for SmB' than the KSm4 antibody relative to SmB signal probably due to

Fig 3.1.5.

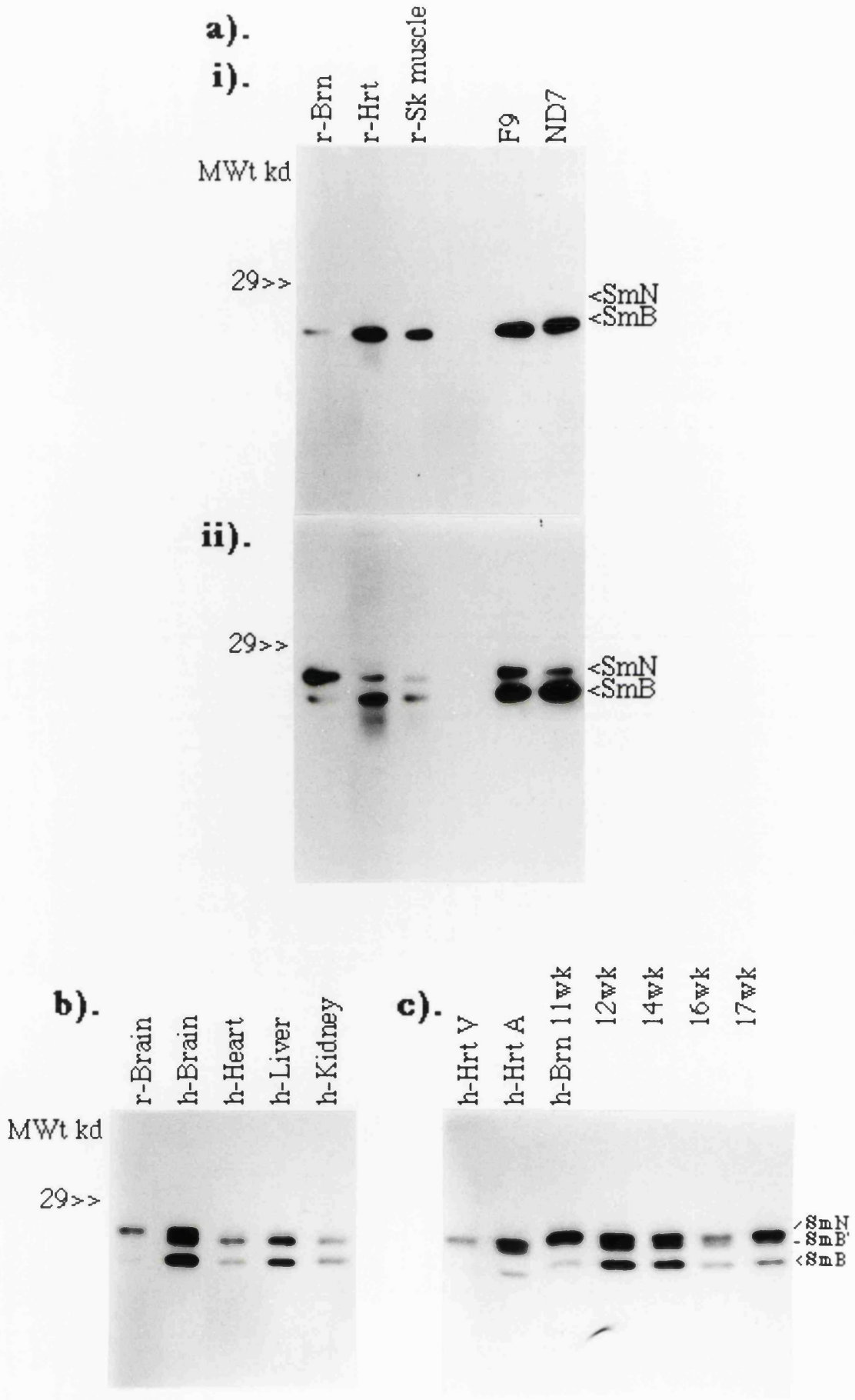
SmN Expression in Human Tissues

Western blots with the KSm antibodies.

a). Rodent tissue and cell proteins detected with the KSm4 (**i.**) and KSm5 antibodies (**ii.**).

b). Rat (r) and human (h) 12 week fetal tissue proteins detected with the KSm5 antibody.

c). Human adult heart, ventricle (V) and atria (A) and brain developmental series proteins detected with the KSm5 antibody.
Numbers indicate weeks of gestation.


a).
i).
r-Brn
r-Hrt
r-Sk muscle
F9
ND7
MWt kd
29>>
<SmN
<SmB
ii).
29>>
<SmN
<SmB
b).
r-Brain
h-Brain
h-Heart
h-Liver
h-Kidney
MWt kd
29>>
c).
h-Hrt V
h-Hrt A
h-Brn 11wk
12wk
14wk
16wk
17wk
<SmN
<SmB'
<SmB

the extra proline-rich motif present on SmB' (Schmauss & Lerner, 1990) (see fig.1.4.) which is recognised by this antibody (Williams et al., 1990). Also both HeLa (I) and (M), which were derived from different sources, were not found to express detectable levels of SmN although SmN mRNA levels appear to differ in these cells (Sharpe et al., 1990b).

3.1.6. Sm detection by immunoflourescence

Sm proteins are largely nuclear localised although there is a minor pool of Sm proteins in the cytoplasm (Zieve & Sauterer, 1990). These proteins have previously been shown to exhibit a punctate nuclear distribution (see 1.3.2.). This pattern includes a number of foci known as coiled bodies and interchromatin granules (Carmo-Fonseca et al., 1992). To determine if the SmN and SmB distribution is similar to that observed for other Sm proteins immunoflourescence staining of 3T3-MSN cells (mouse fibroblasts stably expressing SmN, see 5.3.) was carried out. These cells express endogenous SmB and express SmN from stably transfected constructs. Cells were incubated with the KSm4 antibody (fig.3.1.6.a.) which recognises an epitope found on SmB and SmD or the KSm5 antibody (fig.3.1.6.b.) which recognises an epitope found on SmB and SmN in these cells (Williams et al., 1986) (see 3.1.5.).

Photographs taken under phase contrast optics indicated that all nuclei are positively stained with the KSm antibodies. The nuclear distribution does reflect previously observed patterns but does not show any significant differences. Sm proteins do associate with snRNPs as dimers however so a mixed population of snRNPs would be expected. SmN expressed from a viral (CMV) promoter does localise to the nucleus (fig.3.1.6.b.) suggesting that the SmN protein is snRNP associated and that these cells are suitable for splicing assays aimed at determining the role of the SmN protein (see chapter 5).

To determine any effect of SmN on snRNP localisation it will be necessary to repeat these experiments with 3T3 parental cells and 3T3-MSN (SmN expressing) cells (and primary DRG cells) with the KSm5 antibody.

Fig 3.1.6.

SmN and SmB Detection by Immunoflourescence

The KSm antibodies were used in conjunction with 3T3-MSN cells (see fig.5.3.) to show Sm localization in nuclei.

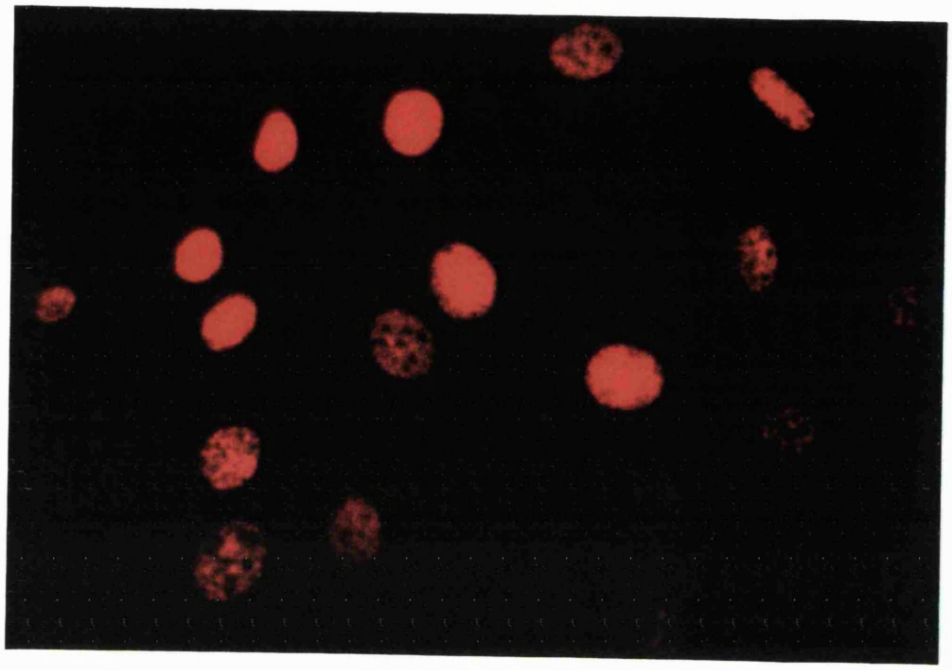
a). Primary antibody = KSm4.

Secondary antibody = rhodamine conjugated rabbit anti mouse.

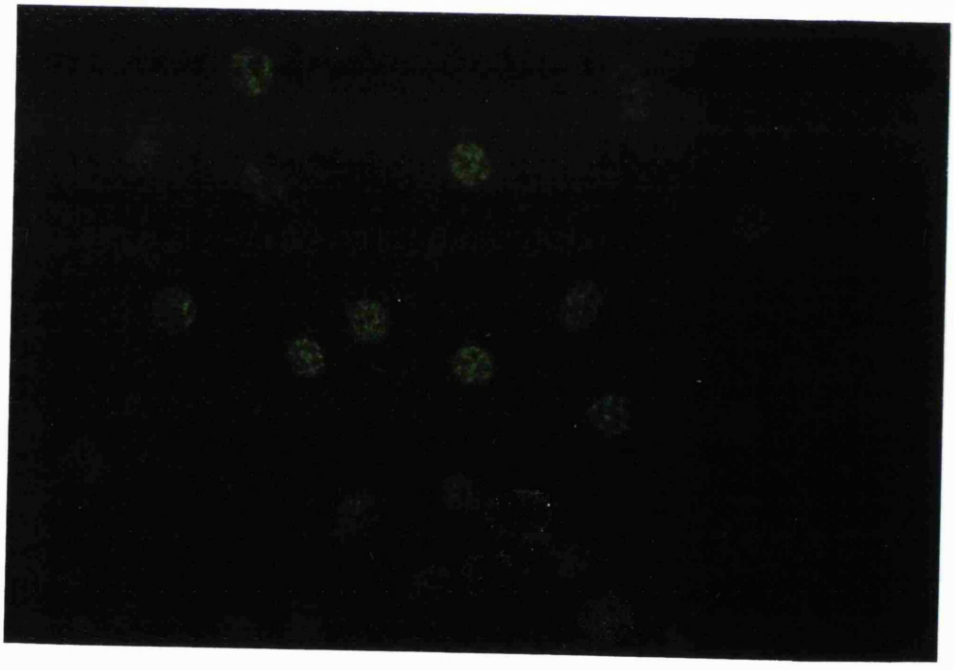
b). Primary antibody = KSm5.

Secondry antibody = flourescein conjugated rabbit anti mouse.

a).



b).



3.2.RNA

3.2.1. SmN mRNA expression on Northern blots

In the following experiments a mouse SmN (Gerrelli et al., 1991) and a human SmB' (van Dam et al., 1989) cDNA clone were used in order to detect rodent SmN and SmB transcripts. SmB' is not expressed in rodent cells and the SmB' cDNA is highly homologous to the SmB sequence (see fig.1.4.) so the SmB' cDNA should provide a suitable probe for SmB transcripts.

The primary control of eukaryotic gene expression lies at the level of transcription (Latchman, 1990b). Steady state mRNA levels may reflect transcriptional regulation. To determine if this is the case with the regulation of Sm genes the levels of SmN and SmB mRNA were determined in a number of cell lines as well as in rat brain tissue during development (see 3.4.). The SmB' clone was removed from the Sp65 vector and cloned into the EcoR1 site of Bluescript-SK for ease of manipulation. The orientation and nature of this clone was confirmed by DNA sequencing. Decreasing levels of SmN peptide during EC cell differentiation have already been shown to result from decreasing levels of mRNA (Sharpe et al., 1990a) and correlations of mRNA to protein levels in various tissues and cells have also been found elsewhere (McAllister et al., 1988; Li et al., 1989). The SmN, SmB and SmB' transcripts have an apparent molecular size of 1.6, 1.4 and 1.3kb respectively on denaturing gels (Delsert & Rosenfeld, 1992).

To determine the distribution of SmN in rodent cell lines the SmN cDNA was oligo-labelled and the probe was hybridised to Northern blots. The probe exhibits a certain degree of non-specific interaction with the 18S rRNA band providing a control for equal loading in each track (fig.3.2.1.a.). Cell lines were further characterised using oligo-labelled SmB' and SmN probes which were subsequently hybridised to the same filter (fig.3.2.1.b.). These blots indicate a direct correlation between the levels of SmN mRNA and SmN protein in each cell type (see 3.1.2.) suggesting that SmN expression is regulated at the level of transcription or RNA stability, and not translation in these cells. These blots also show that within the ND cells SmN/SmB mRNA levels differ as follows: ND3/ND21<ND8/ND11/ND15<ND7.

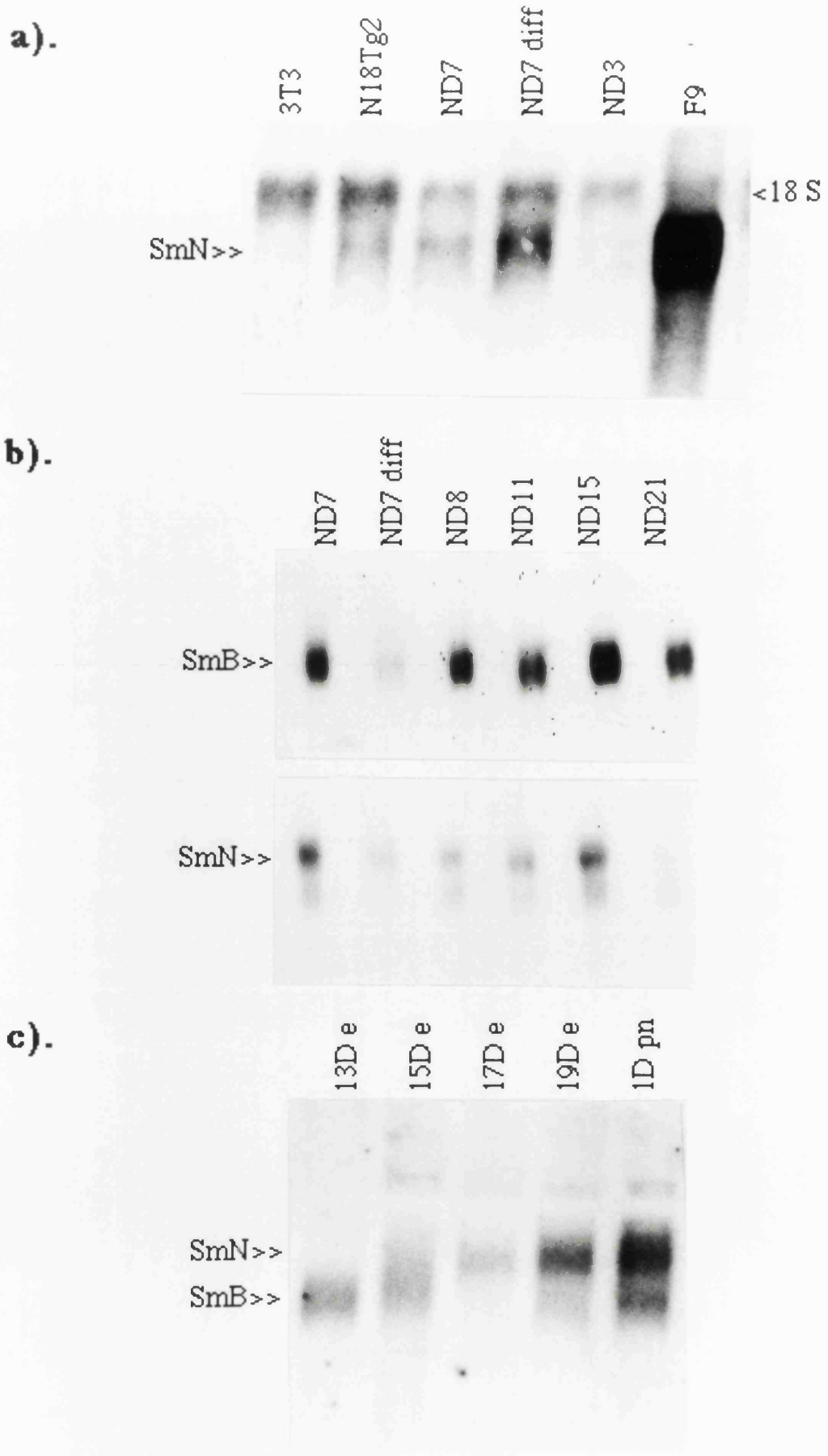
Fig 3.2.1.

SmN mRNA Expression on Northern Blots

a). Cell line total RNA hybridised with SmN cDNA. 18S rRNA indicates the loading in each track.

b). Cell line total RNA hybridised with SmB' cDNA and subsequently stripped and hybridised with SmN cDNA.
ND7 Diff = differentiated ND7 cells (see fig 3.1.4.c.).

c). Mouse tissue developmental series RNA hybridised with SmB' cDNA and washed at low stringency. The film was exposed for twelve days.
D pn = days post natal D e = embryonic days.


a).
3T3
N18Tg2
ND7
ND7 diff
ND3
F9
<18 S
SmN>>
b).
ND7
ND7 diff
ND8
ND11
ND15
ND21
SmB>>
SmN>>
c).
13D e
15D e
17D e
19D e
1D pn
SmN>>
SmB>>

To examine the developmental regulation of SmN and SmB mRNA in mouse brain the SmB' cDNA was oligo-labelled and the probe was hybridised to Northern blots. The filter was washed at low stringency (1 x SSC) to allow detection of both SmN and SmB transcripts. Levels of SmN and SmB mRNA in the mouse developmental series (fig.3.2.1.c.) show regulation as seen at the protein level in 3.1.3. i.e. SmN mRNA increases while SmB mRNA decreases during development. This suggests that primary regulation of SmN/SmB levels is at the level of transcription in tissues as well as in cells. SmN has recently been shown to be regulated at the level of transcription using 5' sequences derived from the SmN gene (Schmauss et al., 1992).

3.2.2. SmN expression *in-situ* in mouse brain

A short fragment from the 3' end of SmN was cloned into Bluescript-SK specifically for use in *in-situ* hybridisation. This clone shows the expression of SmN to be widely distributed (see fig.3.2.2.d.), being present in most areas of the brain as well as in DRG and spinal neurons (Horn et al., 1992). All hybridising cells were identified as neurons on the basis of cresyl violet staining with no hybridisation being observed in glial or other non-neuronal cells. Considerable variation was observed in the expression of the SmN mRNA in different brain regions. Particularly high levels of the mRNA were found in structures such as the pontine nuclei, amygdala and hypothalamus, whilst the mRNA was detectable only at very low levels in other regions such as the inferior colliculus and the inner granule cells of the olfactory bulb. Indeed even within a single structure such as the cerebellum considerable variation was observed in the levels of the SmN mRNA with the highest levels being observed in the Purkinje layer whilst the granular layer has a lower level and the SmN mRNA is virtually undetectable in the molecular layer (for more detail see table 3.1.). A similar SmN mRNA distribution has been described in rat brain while SmB mRNA was expressed at a diffuse, low level throughout (Li et al., 1989; Schmauss et al., 1992).

The large areas of brain tissue not exhibiting SmN mRNA expression suggest that SmN mRNA (and protein) is far more abundant in neurons than SmB mRNA (and protein) is in other cell types. SmN RNA was in fact suggested to be as abundant as actin RNA in

Fig 3.2.2.
SmN mRNA distribution in mouse brain

a). Characterisation of an SmN probe suitable for *in-situ* hybridisation:
T3/2 (Williams et al., 1990), a clone spanning 135 nucleotides at the 3' end of the SmN coding region was amplified by PCR from the λgt11 clone. 30 cycles were carried out at 94°C for 30 seconds, 50°C for 30 seconds and 72°C for 30 seconds using the gt11-1 and gt11-2 oligonucleotide primers. A band of approximately 200bp was digested with EcoRI and the resulting 135bp band was cut from a 2% LGT agarose/ethidium bromide stained gel and purified using Geneclean, then cloned into Bluescript-SK.

To determine the specificity of the probe for SmN mRNA and the orientation of the clone, transcriptions were carried out on mouse SmN (Gerrelli et al., 1991) and human SmB' (Van Dam et al., 1989) cDNA clones in sense and antisense directions. The resulting RNA was run on a 1% agarose gel and transferred to nylon filters. The blot was then incubated with RNA probes transcribed from the T3/2 clone under standard northern blotting conditions. T3/2 was transcribed from the T3 promoter and hybridised to filters which were subsequently washed to a stringency of 0.02 x SSC and 0.05% SDS at 65°C.

i). Agarose / ethidium bromide gel showing SmN and SmB RNA.
S = sense RNA A/S = antisense RNA

ii). 10 minute exposure of a blot showing the T3 transcript of T3/2 hybridising to SmN A/S RNA and weakly to SmB A/S RNA.

b). Schematic representation of the T3/2 fragment cloned into Bluescript-SK. The orientation of the T3/2 clone was derived from **a).**

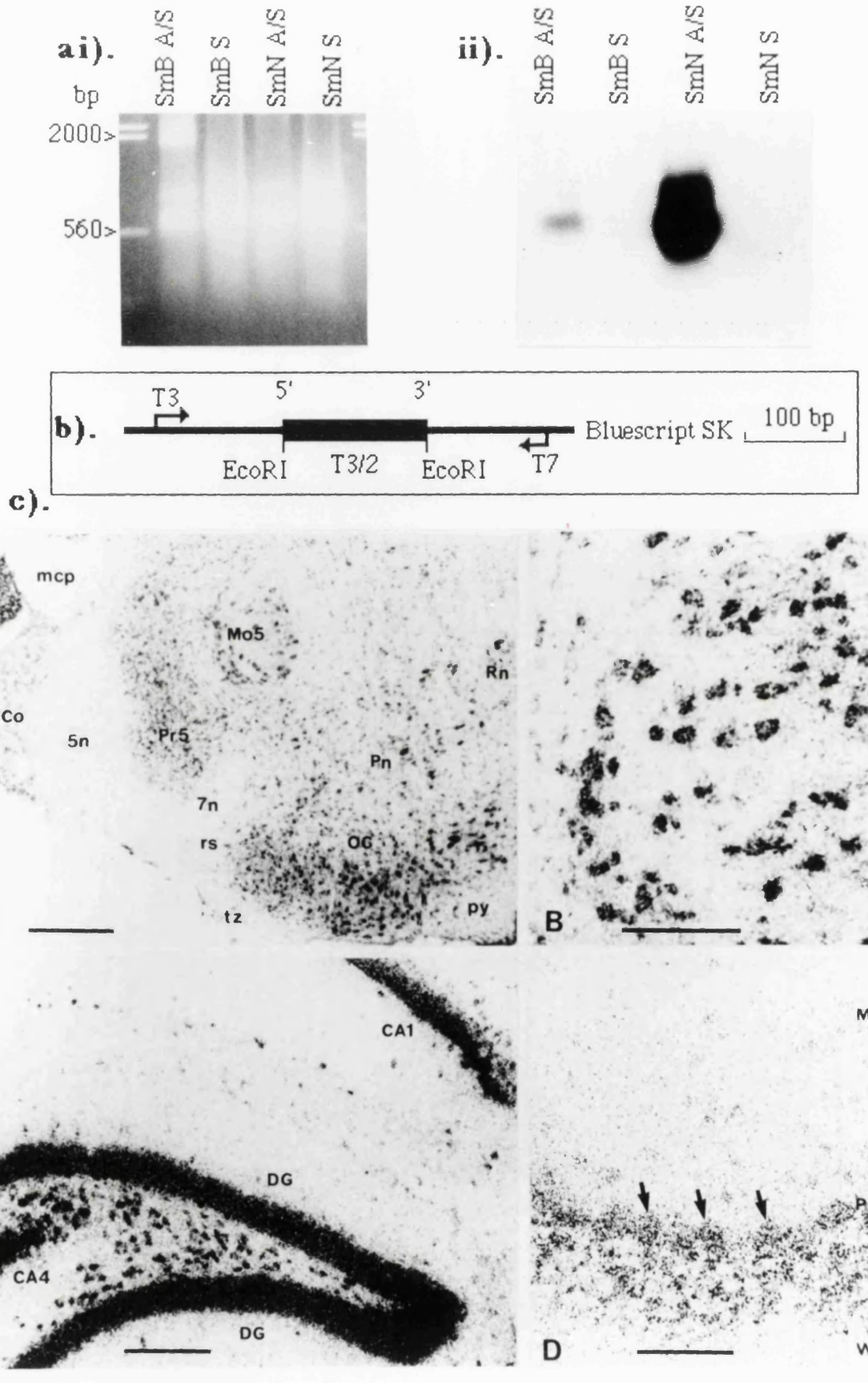
c). *In-situ* hybridisation on mouse adult brain sections using the antisense T3/2 probe transcribed from the T7 promoter.

A). Pons: Bar=1000*u*m.

B). Motor trigeminal nucleus (Mo5 in A): Bar=500*u*m.

C). Hippocampus: CA1 and CA4=Ammon's horn, DG=dentate gyrus. Bar=500*u*m.

D). Cerebellum: Arrows=neurons of the Purkinje layer (P) and granule layer (G), M=molecular layer and W=white matter. Bar=100*u*m.


ai).
SmB A/S
SmB S
SmN A/S
SmN S
bp
2000>
560>
ii).
SmB A/S
SmB S
SmN A/S
SmN S
b).
T3
5'
3'
Bluescript SK
100 bp
EcoRI
T3/2
EcoRI
T7
c).
mcp
Mo5
Rn
VCo
5n
Pr5
Pn
7n
rs
OG
tz
py
A
B
CA1
DG
CA4
DG
C
M
P
G
W
D

Region	Signal intensity
Olfactory bulb	
glomerular cells	—
tuft & mitral cells	+++
inner granular cells	—
Neocortex	++
Primary olfactory cortex	+++
Hippocampus	
Ammon's horn	+++
dentate gyrus	+++
Amygdala	++
Hypothalamic nuclei	++
Thalamic nuclei	++
Superior colliculus	++
Inferior colliculus	+
Substantia nigra	+++
Trigeminal nuclei	+++
Cochlear nuclei	++
Olivary complex nuclei	+++
Brainstem reticular nuclei	++
Brainstem sensory & motor nuclei	+++
Cerebellum	
molecular layer	—
Purkinje cells	++
granule cell layer	+
deep cerebellar nuclei	++
Spinal cord, ventral & dorsal horns	+++
Dorsal root ganglia	+++

Table 3.1 SmN mRNA in mouse neural tissue

Intensity of *in-situ* hybridisation signal for the T3/2, SmN probe in various parts of the brain and spinal cord. Results are the average from four different animals. The signal was assesed semiquantitatively by two separate observers according to an arbitrary scale of three levels of intensity: + = weak, ++ = moderate, +++ = strong.

neurons (Schmauss & Lerner, 1990). Experiments aimed at defining SmN mRNA distribution in the heart using this probe are currently underway.

SmB' fragments of 150 and 267bp in length derived from the 3' end of the SmB' coding region were excised from the human SmB' cDNA clone (Van Dam et al., 1989) by digestion with Bam HI followed by partial digestion with Pvu II. These were also cloned into Bluescript-SK with a view to defining SmB mRNA distribution in various rodent tissues and these experiments are also currently underway.

Levels of SmN mRNA can be directly correlated to SmN protein levels in all cells and tissues tested so far (McAllister et al., 1988; Sharpe et al., 1990a; see 3.1.2; 3.1.4. and 3.2.1.). Promoter elements derived from the SmN gene have now been shown to drive expression of SmN mRNA levels which correlate directly to protein levels in various cell types (Schmauss et al., 1992). It is probable therefore that the SmN mRNA distribution above can be directly correlated to SmN protein levels in these regions of the brain.

CHAPTER 4

RESULTS

Alternative splicing assays

4. Introduction

A number of alternative splicing pathways result from differential promoter usage or differential polyadenylation. However, many alternatively spliced products are also known to derive from the same primary transcript. SmN is thought to be present in the spliceosome and would therefore not be expected to affect promoter usage. Rather, SmN would be expected to exert a tissue specific effect on pre-mRNA transcripts during their association with the splicing machinery or other Sm associated snRNPs. Candidates for regulation by SmN, therefore would be expected to undergo alternative splicing from the same primary transcript. SmN could act as a negative factor inhibiting selection of a specific splice site, by steric interference for example, or could act as a positive factor promoting the use of a specific splice site. The interactions postulated above could be via recognition of specific sequences or recognition of secondary or tertiary pre-mRNA structures. SmN associated with various snRNAs could influence their conformation which could alter splice site specificity leading to altered splice site selection. SmN, B or B' have no RRM (RNA recognition motif) but direct interactions with pre-mRNA may not be necessary for SmN to fulfill this potential function.

Following an extensive literature search a number of alternatively spliced genes, which fit the above requirements were selected (see below). These genes all exhibited a distribution of alternatively spliced products which correlated with SmN expression in either brain or heart tissues or preferably both. These genes were therefore candidates for regulation of alternative splicing by SmN. In order to examine SmN function in the following assays tissue samples, including a rat brain and heart developmental series and cell lines expressing various levels of SmN were used. The ND neuronal cell lines provided a particularly useful model for examination of SmN function. Primary DRG cells can not be obtained in large quantities but the ND clones provide a source of large quantities of clonal neuronal material, not available from animal models. Specifically the ND3/ND7 cell pair provided a neuronal model in which to examine SmN function as they are derived from similar cell types (Wood et al., 1990) but express undetectable and detectable levels of SmN respectively (see 3.1.2.). Other neuronal and non-neuronal cells were also used in the following assays.

If SmN were to be an alternative splicing factor it was considered possible that it could be responsible for inclusion (e.g. NCAM and c-src) or exclusion (e.g. cTNT) of cassette exons or selection of proximal or distal (e.g. CALC-I and $G_o\alpha$ subunits) splice sites (see fig.1.2. and 4.1.). It was also considered likely that such a factor would be responsible for the alternative processing of a variety of genes in each cell and tissue in which it functioned.

Alternative splicing assays carried out on tissue samples or on cells in culture appear to be more valuable in comparison to those carried out using *in-vitro* splicing extracts as the efficiency of certain splice sites (Nelson & Green, 1990), including those of the CALC-I gene (Bovenberg et al., 1988; Emeson et al., 1989) have been shown to be altered *in-vitro*. This may be due to variations in salt concentration (Schmitt et al., 1987), loss of active regulatory factors or disruption of the structural organisation of certain components of the splicing machinery. Alternative splicing influenced by pre-mRNA secondary structure has also been shown to differ *in-vitro* and *in-vivo*. This effect could be due to the sequential interaction of hnRNP proteins with the nascent pre-mRNA during transcription *in-vivo* (Eperon et al., 1988; Beyer & Osheim, 1988). The following assays are therefore not carried out *in-vitro*. Splicing assays of transcripts which were found to be expressed endogenously were carried out using cultured cells as well as tissue samples (NCAM, c-src and G_o). Transcripts found to have a limited distribution were carried out in cultured cells by transfection of minigene constructs (CALC-I, see 4.1. and cTNT, see 4.5.).

PCR protocols were designed to optimise specificity and efficiency (Saiki et al., 1988). A number of modifications were also examined for their ability to improve specificity or remove contaminating DNA (D'Aquila et al., 1991; Hung et al., 1990; Sarkar et al., 1990; Jinno et al., 1990; Deragon et al., 1990). The nature of the DNA derived from PCR amplification of cDNA from alternatively spliced gene transcripts raised the possibility that heteroduplex DNA molecules would be formed which would have a complicated electrophoretic migration pattern on non-denaturing gels. This type of heteroduplex has previously been shown to have a relatively reduced mobility on non-denaturing gels (Small & Akeson, 1990). The assays below are all carried out on non-denaturing gels and in the cases of the NCAM (4.2.) and c-src (4.3.) assays this type of anomalously migrating

heteroduplex may be present (see below for discussion). Generally however heteroduplex formation did not interfere with quantification of alternatively spliced variants.

4.1. CALC-I alternative splicing

4.1. Introduction

A post transcriptional mechanism is responsible for the differential expression of CALC-I mRNAs from a single gene (Rosenfeld et al., 1984; Amara et al., 1984; Edbrooke et al., 1985). The rat (Amara et al., 1982) and human (Steenbergh et al., 1984) CALC-I (αCGRP) genes are conserved and undergo alternative splicing in a tissue specific manner from the same primary transcript, which terminates more than one kilobase downstream of the distal polyadenylation signal (Amara et al., 1984), and may lead to the generation of calcitonin or CGRP polypeptides (see fig.4.1.). Calcitonin is produced following splicing of the first three common exons to exon four in mammals in the C-cells of the thyroid while CGRP is produced by splicing of the first three exons to exons five and six in the mammalian central and peripheral nervous system (Rosenfeld et al., 1983).

Calcitonin is a small (32 amino acids) peptide hormone used as a biological marker of medullary thyroid carcinoma (Milhaud et al., 1974). It is a calcium regulatory hormone which is capable of raising cAMP levels (Laufer & Changeux, 1987) and appears to be involved in protection of the skeleton during times of calcium stress. CGRP (37 amino acids), based on expression pattens may be a neurotransmitter or neuromodulator (Marshall et al., 1986). However, CGRP can act as a potent vasodilator (Brain et al., 1985), inhibit glycogen turnover (Oetting Deems et al., 1991) and has a potent anti-inflammatory action (Raud et al., 1991). CGRP can also antagonise the effect of calcitonin on cell motility (Towhidul et al., 1991). Human and rat CGRP-I peptides differ at four amino acid positions (Morris et al., 1984) while the human CALC-I gene and the related CALC-II (βCGRP) gene produce CGRP peptides which differ at three amino acid positions (Steenbergh et al., 1985). An analagous rat CGRP-II has also been reported (Rosenfeld et al., 1984) and another related gene also exists (CALC-III). CGRP-II has been shown to be expressed in a sarcoma cell line (Hoppener et al., 1987). Consequently, immunochemical data on CGRP expression should be treated with caution.

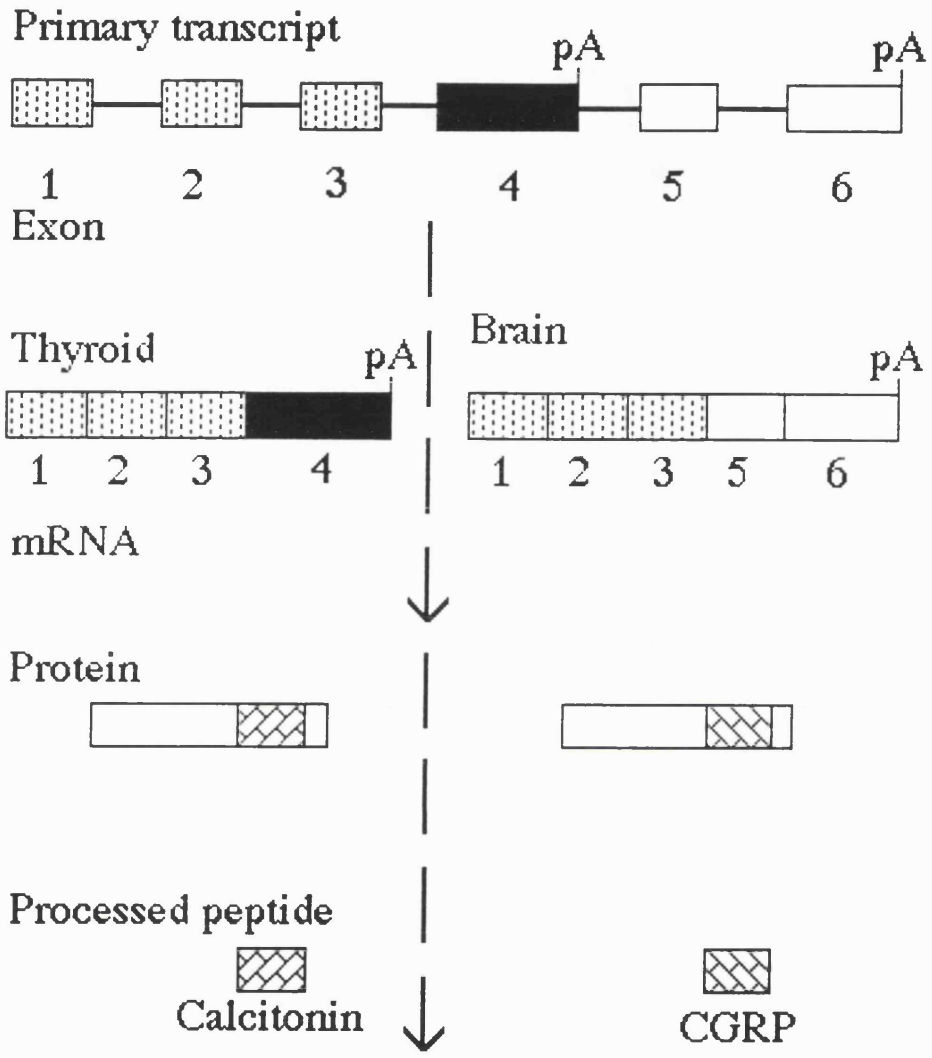


Figure 4.1. Schematic representation of CALC-I alternative splicing

The first three exons of the CALC-I gene are common to both the calcitonin and CGRP mRNA transcripts and consist of non-coding sequences. Exon 4 contains the calcitonin peptide coding sequence whilst exon 5 contains the CGRP peptide coding region. Hence alternative splicing followed by proteolytic cleavage of the protein produced in each tissue yields calcitonin in the thyroid and CGRP in the brain.

Immunoreactive CGRP found in the nervous system does appear to derive from the CALC-I gene however (Rosenfeld et al., 1983) and calcitonin mRNA does not appear to be expressed from the CALC-II or CALC-III genes (Alevizaki et al., 1986; Steenbergh et al., 1985).

The proteolytic cleavage of the human calcitonin precursor has been proposed to liberate a biologically active C terminal flanking peptide, (Amara et al., 1982; MacIntyre et al., 1982) but this peptide shows little phylogenetic conservation with the rat homolog (Craig et al., 1982). Another transcript is produced by splicing part of exon four to exons three and five (Minvielle et al., 1991) and another novel calcitonin encoding transcript has been reported (Adema & Baas, 1992).

The CALC-I pre-mRNA posesses alternative polyadenylation sites (see fig.4.1.). The absence of a classic splice donor consensus sequence at the 3' end of the calcitonin exon suggests that preferential use of a splice acceptor site gives rise to CGRP mRNA, whilst cleavage of a partially processed RNA transcript followed by polyadenylation gives rise to calcitonin mRNA (Emeson et al., 1989). CGRP expression from the CALC-I gene correlates well with the expression of SmN in a number of rodent and human cells and tissues (Crenshaw et al., 1987; Leff et al., 1987; McAllister et al., 1989). In transgenic mice expressing a metallothionein-CALC-I fusion gene CGRP mRNA was made predominantly in the brain whereas the only non-neural tissue to make significant amounts of CGRP mRNA was the heart; all other tissues produced calcitonin (Crenshaw et al., 1987). Extensive correlations between SmN expression and the ability to follow the CGRP splice choice have also been reported elsewhere (McAllister et al., 1989; Li et al., 1989; Sharpe et al., 1989).

In-vivo studies on expression of the rat CALC-I gene suggest that the use of a polyadenylation site is not a key regulatory event for the alternative splicing and that tissue-specific splicing is regulated by an alternative splice site selection (Emeson et al., 1989). This suggests that a specific regulatory machinery is necessary for the production of CGRP in neuronal cells (Crenshaw et al., 1987; Leff et al., 1987). This mechanism implicates the involvement of a *trans*-acting factor, and identifies a mechanism probably not unique to the calcitonin gene. A model has been proposed in which a *trans*-acting factor is

expressed specifically in cells producing CGRP. This factor has been proposed to alter the secondary structure of the pre-mRNA, inhibiting usage of the calcitonin polyadenylation site and leading to CGRP production (Leff et al., 1987; Emeson et al., 1989). *In-vitro* studies however suggest that a factor in HeLa cells may bind exon four to promote calcitonin splicing (Cote et al., 1992).

A number of factors can influence the transcription of the CALC-I gene, (Cote & Gagel,1986) but glucocorticoid (Cote & Gagel, 1986) and v-Ha-ras (Nakagawa et al., 1987), which induces differentiation of medullary thyroid carcinoma cells, have been found to affect the calcitonin/CGRP expression ratio. This suggests that dexamethasone affects a *trans*-acting machinery or factor involved in CALC-I alternative splicing in a quantitative or qualitative fashion. Calcium, antiglucocorticoid (Zeytin et al., 1987), steroid hormones (Gillardon et al., 1991) and cell division rate (Nelkin et al., 1989) may also affect CALC-I processing.

4.1.1. Splicing assay

CALC-I expression is limited to discreet tissues, namely thyroid C cells and nervous tissue (Rosenfeld et al., 1983). Very few cell lines express endogenous CALC-I transcripts, specifically the ND neuronal cell lines do not express this gene (Suburo et al., 1992). Undifferentiated F9 cells are one of the few cell lines known to express the CALC-I gene, producing predominantly CGRP transcripts. During F9 differentiation transcription of this gene is switched off (Evain-Brion et al., 1984).

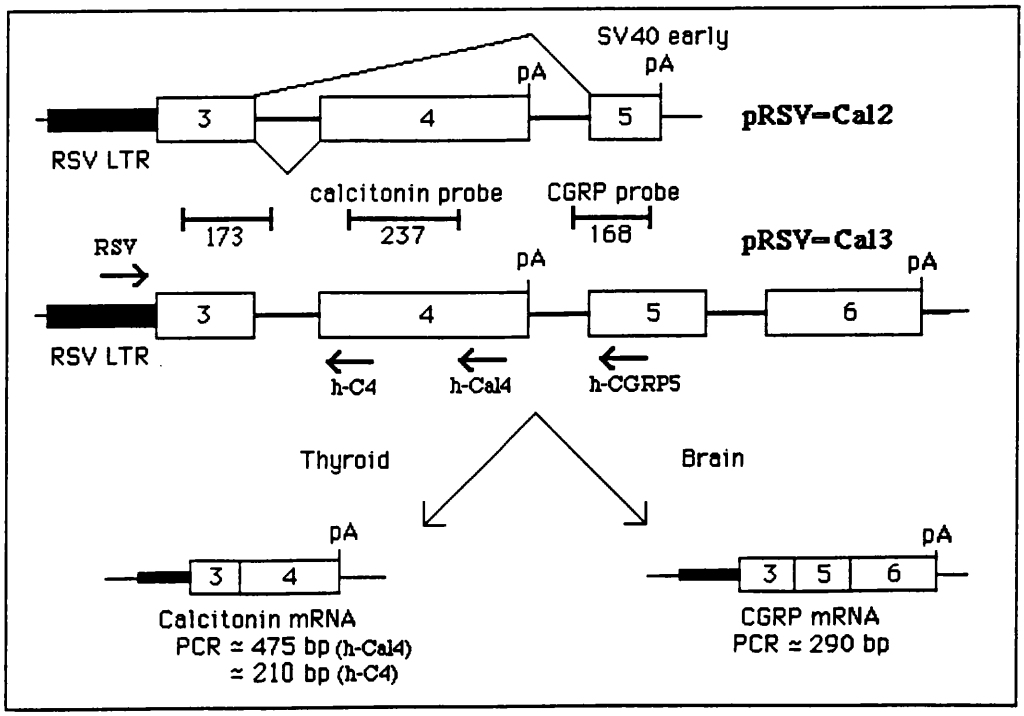
As endogenous expression from the CALC-I promoter is highly restricted, in order to express detectable levels of CALC-I pre-mRNA in diverse cell types and to examine the CALC-I splicing ratios in these cells it was necessary to obtain CALC-I minigene constructs. Constructs containing the alternative exons of the CALC-I gene downstream of the RSV-LTR were obtained for transfection and subsequent RT-PCR assay.

Human CALC-I minigene construct, pRSV-Cal3 was derived from a cosmid clone (Broad et al., 1989) and contained part of exon three and the whole of exons four through to six including both CALC-I polyadenylation signals. pRSV-Cal2 is similar to pRSV-Cal3 except that it was truncated within exon five and ligated to the SV40 early polyadenylation signal.

Fig 4.1.1.

CALC-1 alternative splicing assay

Schematic representation of the structure of the human CALC-I minigene constructs, pRSV-Cal2 and pRSV-Cal3, alternative processing of CALC-I pre-mRNA and the nature of the PCR assay. Small arrows indicate regions from which the human sequence and RSV oligonucleotide primers were designed for the splicing assay. Bars labelled 173, 237 and 168 represent the probes oligo-labelled for duplicate or independent detection respectively of the PCR products. The probe used for detection of both PCR products was derived from exon 3 of pRSV-Cal3 by Hpa I/Bam HI digestion and the resulting 173bp product was complementary with the PCR products through 67% of its length. The probe used for calcitonin PCR product detection was derived from exon 4 of pRSV-Cal3 by Bgl II, Nsi I digestion, and the resulting 237bp product (excised from 2% LGT agarose gel as 237/241bp doublet) was 100% complementary with the PCR product. The probe used for CGRP PCR product detection was derived from exon 5 of pRSV-Cal3 by Bgl II, Nar I digestion and the resulting 168bp product was complementary with the PCR product through 72% of its length.



RNA derived from transfected cells cultured on a 90mm plate was insufficient to detect calcitonin or CGRP transcripts on a Northern blot using the 168 and 237 probes. so a quantitative PCR assay was designed for this purpose (see fig.4.1.1.). The use of an oligonucleotide specific for the RSV-LTR ensures that PCR products are derived from the transfected minigene construct rather than any endogenous CALC-I transcripts which may be present in any of the cells to be examined.

4.1.2. PCR controls

Actin and ribosomal protein specific oligonucleotides (see materials) were used to equalise cDNA samples and these assays were shown to be quantitative by amplification of dilution series' (fig.4.1.2.a.). PCR amplification of CGRP cDNA derived from a rat brain and heart developmental series, using the r-CGRP2 and r-CGRP5 oligonucleotides showed endogenous CGRP expression to increase, relative to rRNA expression during brain development while the transcript was undetectable in all heart samples as expected (fig.4.1.2.b.).

The oligonucleotides were initially used in conjunction with the human CALC-I construct, pRSV-Cal3 to determine the optimal amplification conditions as well as to determine their specificity. This protocol yielded the expected products of approximately 1100 (RSV and h-C4), 1400 (RSV and h-Cal4) and 2400bp (RSV and h-CGRP5). The SM-I minigene construct to be used in the cTNT assay (see fig.4.5.1.) was also subjected to the same type of amplification with the RSV and a-cTNT6 oligonucleotides which also yielded a band, as expected, of approximately 3600bp.

The 475 and 290bp RT-PCR products were confirmed to be derived from accurately spliced products derived from the minigene constructs by restriction enzyme digestion with Hind III, Bgl II, Sph I and Hind III, Bgl II and Hinf I respectively. All of these digests yielded the expected size products.

The ratio of CALC-I RT-PCR products was also shown to be representative of input cDNA regardless of significant size differences by amplification of cDNA derived from ND7 cells transfected with pRSV-Cal3 (see fig.4.1.2.d.). Amplifications were carried out with RSV, h-CGRP5 and h-Cal4 (lane 1) or h-C4 (lane 2). Amplifications indicate a similar calcitonin/CGRP ratio with calcitonin bands of 210 or 475bp.

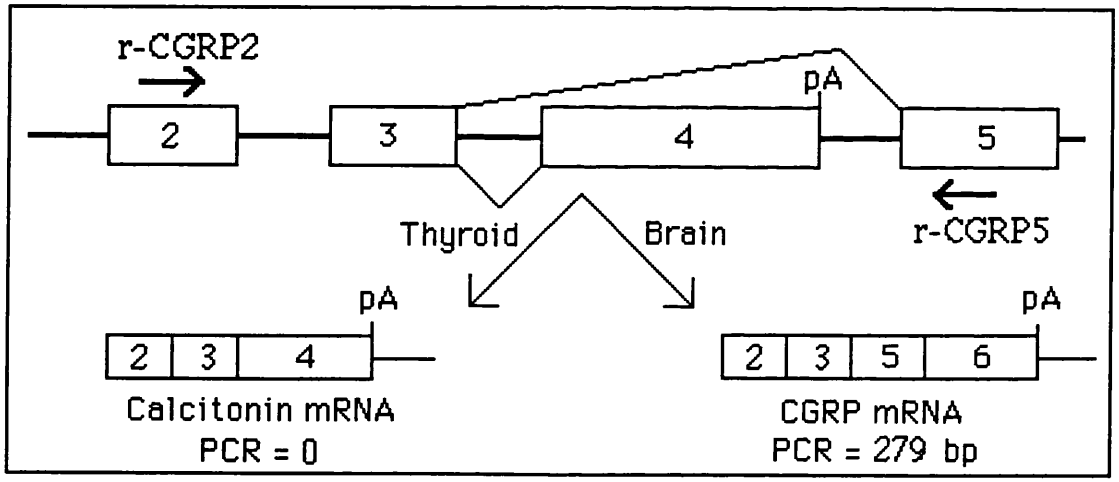
Fig 4.1.2.
PCR Controls

PCR amplifications of cDNA separated on agarose/ethidium bromide stained gels.

a). Amplification of a dilution series of ND7 cDNA. Numbers represent *ul* of input cDNA as do the numbers on all subsequent dilution series. Actin and rRNA cDNA was amplified using the act-1 and act-2 or rib-5 and rib-3 oligonucleotides respectively. 22 cycles were carried out at 94°C, 58°C and 72°C for 30 seconds each. The resulting products were separated on 2% gels.

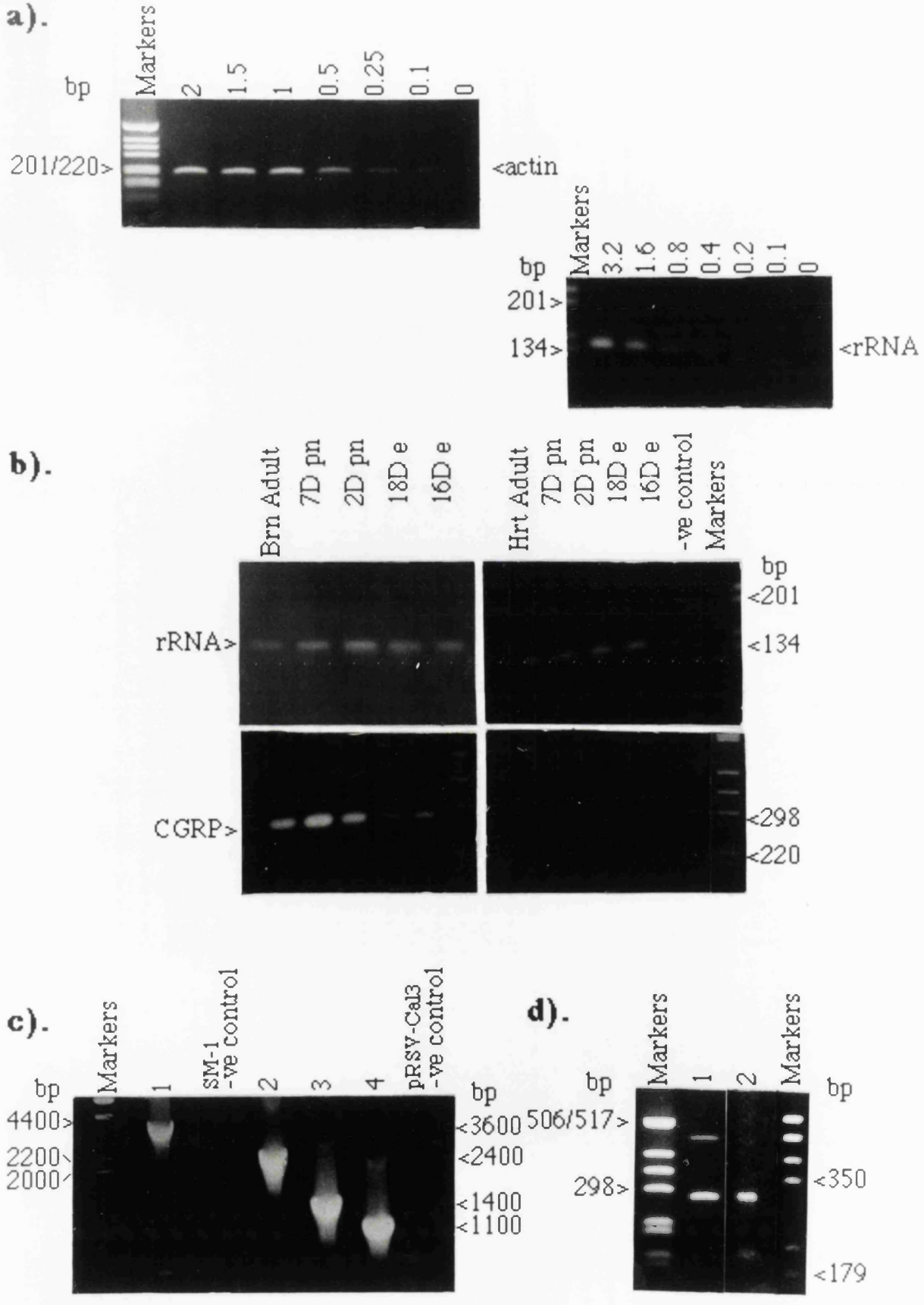
b). Amplifications of rat brain and heart tissue developmental cDNAs. rRNA cDNA was amplified as above and CGRP cDNA was amplified using the r-CGRP2 and r-CGRP5 oligonucleotides. 35 cycles were carried out at 94°C, 62°C and 72°C for 45 seconds each. The resulting products were separated on 1.6% gels.

Schematic representation of the rat CGRP splicing assay:



c). Amplifications of plasmid DNA with various plasmid pairs. SM-1 amplified with RSV and a-cTNT6 (1), 2-4 = pRSV-Cal3 amplified with RSV and h-CGRP5 (2), h-Cal4 (3), and h-C4 (4). The resulting products were separated on 1% gels. Amplifications were carried out for 30 cycles at 94°C for 1 minute, 64°C for 1 minute and 72°C for 3 minutes with approximately 1ng of plasmid DNA.

d). Amplifications of cDNA derived from cells transfected with pRSV-Cal3 using the oligonucleotides RSV, h-Cal4 (1) or h-C4 (2) and h-CGRP5 (see fig 4.1.1). 30 cycles were carried out at 94°C, 64°C and 72°C for 45 seconds each. The resulting products were seperated on 1.75% gels.


a).
bp
Markers
2
1.5
1
0.5
0.25
0.1
0
201/220>
<actin
bp
Markers
3.2
1.6
0.8
0.4
0.2
0.1
0
201>
134>
<rRNA
b).
Brn Adult
7D pn
2D pn
18D e
16D e
Hrt Adult
7D pn
2D pn
18D e
16D e
-ve control
Markers
bp
rRNA>
<201
<134
CGRP>
<298
<220
c).
bp
Markers
1
SM-1
-ve control
2
3
4
pRSV-Cat3
-ve control
bp
4400>
2200>
2000>
<3600
<2400
<1400
<1100
d).
bp
Markers
1
2
Markers
bp
506/517>
298>
<350
<179

4.1.3. Cell line assays

The PCR amplification is shown to be quantitative relative to input DNA by the dilution series (fig.4.1.3.a.).

On comparison of splicing ratios obtained using the human minigene constructs (Edbrooke et al., 1985), pRSV-Cal2 and pRSV-Cal3 it can be seen that the region deleted in pRSV-Cal2 prevents efficient production of CGRP transcripts, even from cells capable of producing predominantly CGRP transcripts from pRSV-Cal3. This may be due to the replacement of the exon six polyadenylation signal or simply due to the loss of the 3' end of exon five. Another human construct shows that exons one-three are not required for accurate regulation of CALC-I alternative splicing (Cote et al., 1991). It was also suggested that a human CALC-I construct truncated within exons three and five contained all of the *cis*-acting sequences necessary for accurate CALC-I processing (Bovenberg et al., 1986).

pRSV-Cal3 contains both natural polyadenylation signals and appears to contain all the *cis*-acting sequences responsible for the regulation of CALC-I splicing. This construct appears to reflect the splicing choices taken in neurons in the brain and in other tissues (Crenshaw et al., 1987). i.e. Neuronal cell lines, ND3 and ND7 produce predominantly CGRP transcripts when transfected with pRSV-Cal3 while 3T3 fibroblasts and HeLa-I epithelial cells produce predominantly calcitonin transcripts. This cell type specific splicing of transcripts derived from a transfected construct illustrates the presence of cell type specific factors which influence this splice choice. Accurate processing of a human construct in rodent cells also indicates a functional conservation of CALC-I processing *cis* and *trans* elements between rodents and human. Filters were also incubated with a 173bp probe (see fig.4.1.1.) which showed equal specificity for both PCR products. Amplification of cDNA from HeLa-I and ND7 cells also showed the same results with the RSV, h-C4 and h-CGRP5 oligonucleotides on agarose/ethidium bromide stained gel. The position of the larger band detected with the 168 probe in HeLa and ND7 cells corresponds to the size expected for the previously described CCP-II mRNA (Minvielle et al., 1991). Another novel calcitonin encoding transcript (Adema & Baas, 1992) was not detected in this assay.

Fig 4.1.3.

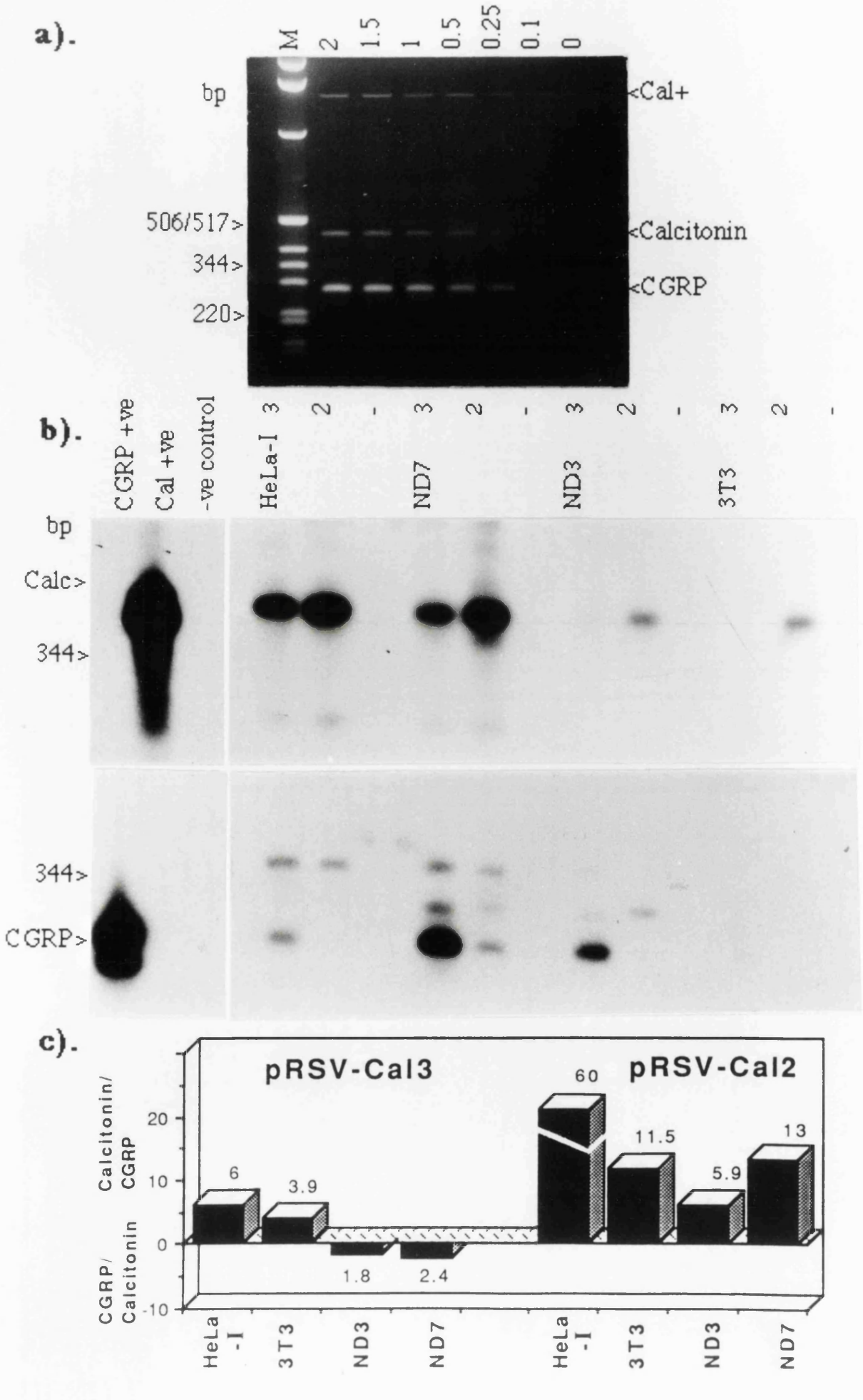
CALC-I Alternative Splicing in Cell Lines

PCR amplification of CALC-I cDNA from cells transfected with pRSV-Cal2 or pRSV-Cal3 were carried out as in 4.1.2.d. and separated on 1.75% agarose/ethidium bromide stained gels.

a). Agarose/ethidium bromide stained gel showing amplification of a dilution series of cDNA derived from ND7 cells transfected with pRSV-Cal3. Cal+ is a band amplified by the RSV and h-Cal4 oligonucleotide primers and is derived from plasmid DNA or unspliced precursor RNA.

b). Overnight exposures of duplicate Southern blots showing calcitonin and CGRP PCR products, of the correct size, hybridised to specific probes derived by restriction enzyme digestion of pRSV-Cal3 (see fig 4.1.1.). The calcitonin positive control is a 483bp Bgl II fragment while the CGRP positive control is a 231bp Nco I/Pvu II fragment derived from exons 4 and 5 of pRSV-Cal3 respectively. Filters were washed to a stringency of 0.25 x SSC and 0.1% SDS. 3 represents cells transfected with pRSV-Cal3, 2 represents cells transfected with pRSV-Cal2 and - represents untransfected cells.

c). Bar chart representing the ratios of CALC-I products obtained from the various cell lines in **b**). The figures were obtained by scintillation counting of fragments of filter containing the hybridising band. Cells appearing to produce more calcitonin are represented by bars above the origin while cells appearing to produce more CGRP are represented by bars below.


a).
M
2
1.5
1
0.5
0.25
0.1
0
bp
506/517>
344>
220>
<Cal+
<Calcitonin
<CGRP
b).
CGRP +ve
Cal +ve
-ve control
HeLa-I
ND7
ND3
3T3
bp
Calc>
344>
344>
CGRP>
c).
pRSV-Cal3
pRSV-Cal2
Calcitonin/ CGRP
CGRP/ Calcitonin
20
10
0
-10
6
3.9
1.8
2.4
60
11.5
5.9
13
HeLa -I
3T3
ND3
ND7
HeLa -I
3T3
ND3
ND7

The correlation between the expression of SmN and the ability to follow the CGRP splice choice appears to be lost in this example. ND3, 3T3 and HeLa-I (Sharpe et al., 1990b) cells do not produce detectable levels of SmN, ND3 cells however are able to produce predominantly CGRP transcripts from the minigene construct pRSV-Cal3, while the non neuronal cells produce predominantly calcitonin transcripts. SmN therefore does not appear to confer the ability to follow the CGRP splice choice.

Further evidence that SmN is not involved in the alternative splicing of CALC-I is that the mRNA encoding SmN is found in a few areas of the brain which can only carry out calcitonin-specific splicing in transgenic mice (Crenshaw et al.,1987), such as the Purkinje layer of the cerebellum and the inferior colliculus (see 3.2.2.). The undifferentiated O1A1 cell line also appears to present an example of a cell line expressing SmN but producing predominantly calcitonin transcripts, while undifferentiated F9 cells appear to have a similar SmN/SmB ratio (fig.3.1.4.) but produce predominantly CGRP transcripts (Adema & Baas, 1991). The best evidence against the involvement of SmN in CALC-I alternative splicing however is a cell line or tissue capable of the CGRP splice choice while expressing no SmN. The ND3 cells and 293 cells (Delsert & Rosenfeld, 1992) present examples of this.

SmN does not appear to be necessary for CGRP production but it remains possible that SmN controls other RNA processing events (see below).

4.2. NCAM (VASE exon) alternative splicing

4.2. Introduction

NCAMs are a group of cell surface sialoglycoproteins of the immunoglobulin gene superfamily which mediate calcium-independent interactions between cells (for review see Edelman, 1986). NCAM appears on early embryonic cells and is important in cell adhesion during morphogenesis. Later in development NCAM is found on various differentiated tissues and is a major adhesion molecule among neurons and between neurons and muscle. NCAM has been suggested to have a role in axon guidance, cell segregation, and formation and innervation of skeletal muscle. The chicken NCAM gene contains at least 20 exons spanning more than 75 kilobases (Cunningham et al., 1987). The nomenclature numbers the 5' exon as exon 0 and the 3' exon as exon 19.

The NCAM pre-mRNA undergoes tissue specific and developmentally regulated alternative splicing events in at least three distinct regions generating between 20 and 30 isoforms in heart development alone (Reyes et al., 1991). Exon 18 is a neural specific alternative exon (Tacke & Goridis, 1991) and other alternative splicing events between exons 15 to 19 also allows the production of variable cytoplasmic domains. An extracellular muscle-specific sequence domain (MSD-1) is present between exons 12 and 13 (Dickson et al., 1987; Thompson et al., 1989). This domain may introduce a proposed hinge region into the molecule and another extracellular 30 base pair alternative exon (cassette, see fig.1.2. N^{o} 1) known as VASE (immunoglobulin Variable domain-like Alternatively Spliced Exon) is present between exons seven and eight (Small et al., 1988) (see fig.4.2.1.c.). Inclusion of the VASE exon appears to crrelate well with SmN expression.

The large cytoplasmic domain is tissue specific and developmentally regulated (Murray et al., 1986) and alternative splicing and the use of two alternative polyadenylation signals (Barbas et al., 1988) at the 3' end appears to influence cellular localization (Powell et al., 1991). The muscle specific domain consists of four short exons, constituting 93 to 108 base pairs, which are included in NCAM mRNAs in the avian (Prediger et al., 1988), rat (Reyes et al., 1991), mouse (Santoni et al., 1989) and human (Thompson et al., 1989) striated muscle in an independently regulated manner. One of these exons, MSD1a has also

been found in NCAM mRNAs in mouse neural tissues (Santoni et al., 1989; Hamshere et al., 1991).

An NCAM cDNA clone derived from adult rat brain was found to be highly homologous to mouse (Barthels et al., 1987) and avian (Cunningham et al., 1987) cDNAs but contained a novel 30 base pair segment between exons seven and eight (Small et al., 1987). This segment was found to be present in four discreet sizes of NCAM mRNA in adult rat brain which differs significantly from the pattern in neonatal rat brain. Inclusion of this alternatively spliced VASE exon is developmentally regulated, with inclusion increasing during rat brain development (Small et al., 1988). The VASE exon was also found to be present in mouse cDNA but not in the chicken (Santoni et al., 1989). The mouse and rat exons are 100% conserved at the nucleotide level. The 10 amino acid insertion resulting from the inclusion of this exon would be in the fourth of five putatative immunoglobulin-like folds in the extracellular domain of the molecule. It appeared that regulation of this exon was independent of the more 3' alternative splicing events and therefore probably controlled by different *trans*-acting factors. The VASE exon, as well as being tissue specific, was shown to be independently developmentally regulated in rat brain and heart and was expressed only at low levels in a number of cell lines examined. Inclusion of the VASE exon is higher in the CNS relative to the PNS (Small & Akeson, 1990). During brain development from embryonic day 15 to adult the percentage of transcripts including the VASE exon relative to transcripts excluding the exon rose from 3 to 43% while levels rose from 10 to 50% during heart development (Reyes et al., 1991).

The protein containing the VASE insertion was shown to be expressed in cultured embryonic brain cells and the number of cells expressing the VASE insert appeared to increase during growth in culture (Small et al., 1988). Adult NCAM has significantly increased binding affinity than embryonic NCAM (Sadoul et al., 1983) so it is possible that the VASE exon influences these properties in some way, possibly via the action of polysialic acid association (Small et al., 1988). Inclusion of the VASE exon does not appear to inhibit adhesion but does downregulate the neurite growth promoting activity of NCAM (Doherty et al., 1992). This action was proposed to be via a multistep process

involving the activation of neuronal second messenger pathways leading to stable, long term cell-cell interactions in the CNS.

4.2.1. Quantitative splicing assay and cell line assays

NCAM transcripts were found to be expressed as endogenous transcripts in all the tissues and cells examined so it was not necessary to construct NCAM minigene constructs. Constitutive endogenous expression also allows splicing assays to be carried out using tissue mRNA samples. The PCR amplification is shown to be quantitative relative to input DNA by the dilution series (fig.4.2.1.a.).

On examination of NCAM alternative splicing in a number of cell lines (fig.4.2.1.b.), the correlation between SmN expression and inclusion of the VASE exon in NCAM transcripts appears to be lost. F9 and 3T3 cells can produce detectable quantities of transcripts including the VASE exon whereas ND7 cells do not produce detectable levels of this transcript. As 3T3 cells express lower levels of SmN than do ND7 cells this constitutes a loss of correlation. SmN therefore does not appear to confer the ability to follow the neuronal splice choice. PC12 cells have previously been shown to be unable to follow this splice choice despite expressing SmN (Small & Akeson, 1990) (fig.3.1.2.b.).

4.2.2. Tissue assays

The 30bp VASE exon is shown to be included in NCAM transcripts in a tissue specific manner in the mouse being more abundant in brain than in heart, and being undetectable in kidney and skeletal muscle (fig.4.2.2.a.). This reflects the distribution of SmN in these tissues. The developmental series of rat brain tissue cDNA (fig.4.2.2.b.) also shows a good correlation with developmental regulation of SmN (fig.3.1.3.). VASE exon inclusion and SmN levels increase during brain development. Hence inclusion of the VASE exon is spatially and temporally regulated (Small et al., 1988). The VASE exon has previously been shown to be more abundant in the central nervous system relative to the peripheral nervous system and most neural regions showed a correlation with SmN expression by *in-situ* hybridization (Horn et al., 1992) except for spinal cord and DRG. These regions express SmN but are unable to follow the VASE exon splicing pathway (Small & Akeson, 1990).

Fig 4.2.1.

NCAM, VASE Exon Alternative Splicing in Cell Lines

PCR amplifications of NCAM cDNA were carried out using the oligonucleotides r-NCAM7 and r-NCAM8. 30 cycles were carried out at 94°C for 30 seconds, 60°C for 30 seconds and 72°C for 30 seconds. The resulting products were then seperated on 2% agarose/ethidium bromide stained gels.

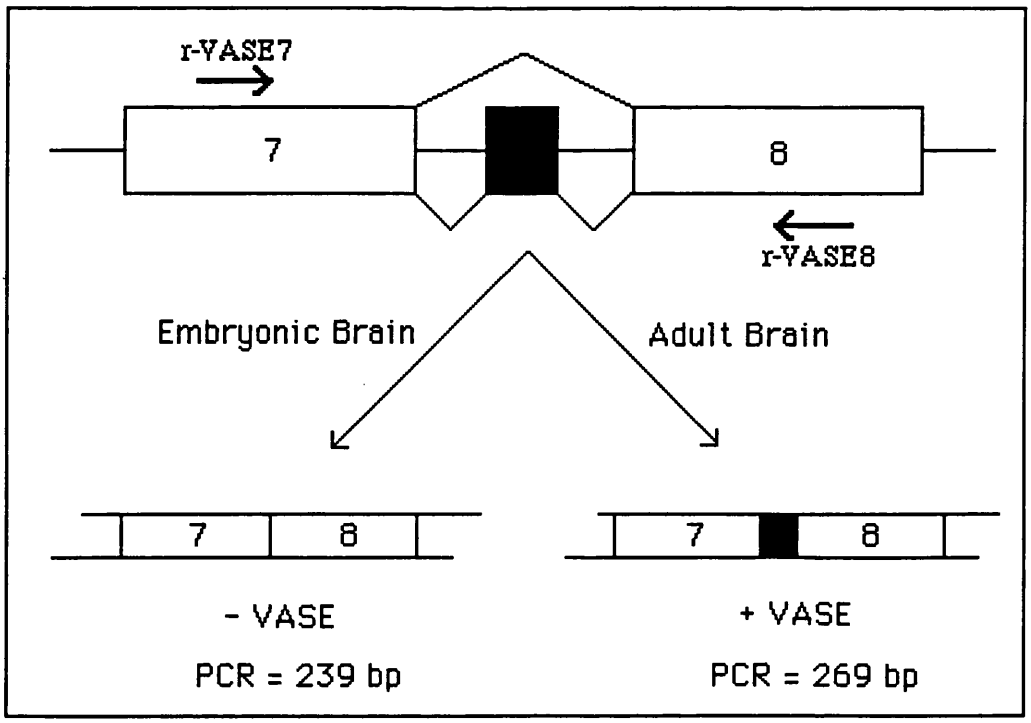
a). Amplification of a dilution series of 2 day post natal rat brain cDNA.

b). Amplification of rodent cell line cDNA.

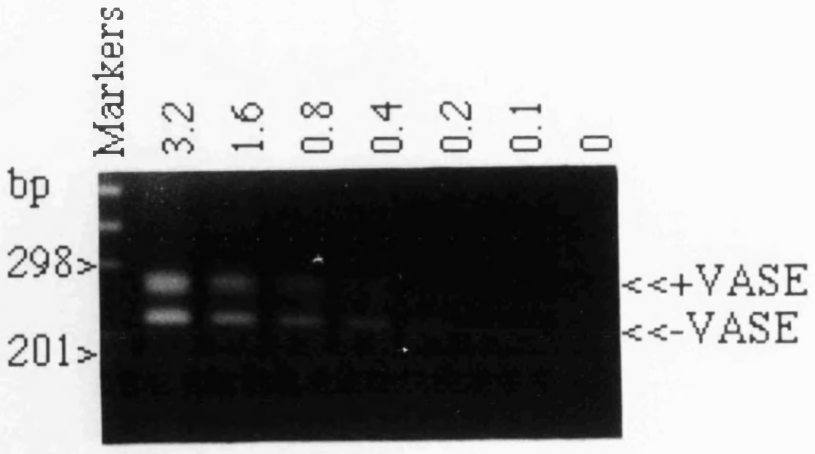
ND7 Diff = differentiated ND7 cells (see fig 3.1.4.a.).

c). Schematic representation of VASE exon alternative splicing in NCAM pre-mRNA. Small arrows indicate regions from which the oligonucleotide primers were designed for the splicing assay.

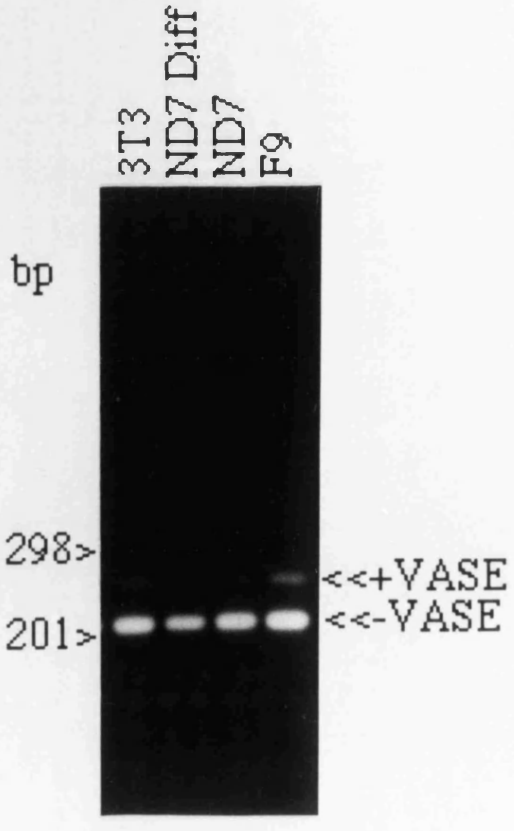
c).



a).


Markers
3.2
1.6
0.8
0.4
0.2
0.1
0
bp
298>
201>
<<+VASE
<<-VASE

b).


3T3
ND7 Diff
ND7
F9
bp
298>
201>
<<+VASE
<<-VASE

Resolution of the PCR products derived from tissue samples shows three bands in the heart and brain tracks. The spacing of the bands suggests that the smallest, 239bp band represents NCAM transcripts excluding the VASE exon while the central 269bp band represents NCAM transcripts containing the VASE exon. The largest band may represent another as yet unidentified, developmentally regulated NCAM isotype with inclusion of another alternative exon (approximately 12bp) between exons seven and eight. This band could derive from 10.5 kilobases of intron sequence between exons seven and eight (Small et al., 1988). The band also appears in some of the cell lines tested. After resolution under more denaturing conditions the three bands derived from mouse and rat adult brain resolve into two bands at the expected positions for fragments of 239 and 269bp in length. This suggests that the extra band is a heteroduplex formed between both NCAM isoforms as its appearance also appears to be quantitatively related to the appearance of the 269bp band. This type of heteroduplex has been reported previously after PCR amplification across the VASE exon region (Small & Akeson., 1990). It appears therefore that these experiments failed to detect any other insertion except for the VASE exon between NCAM exons seven and eight and inclusion of the VASE exon does not correlate with SmN expression.

Fig 4.2.2.

NCAM, VASE Exon Alternative Splicing in Tissues

PCR amplifications of NCAM cDNA and separation of products were carried out as in 4.2.1.

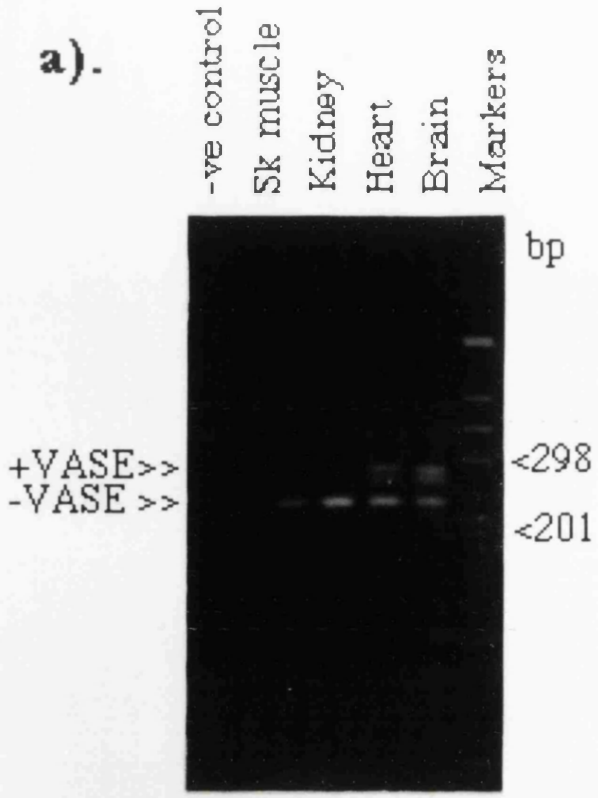
a). Amplification of mouse tissue cDNA.

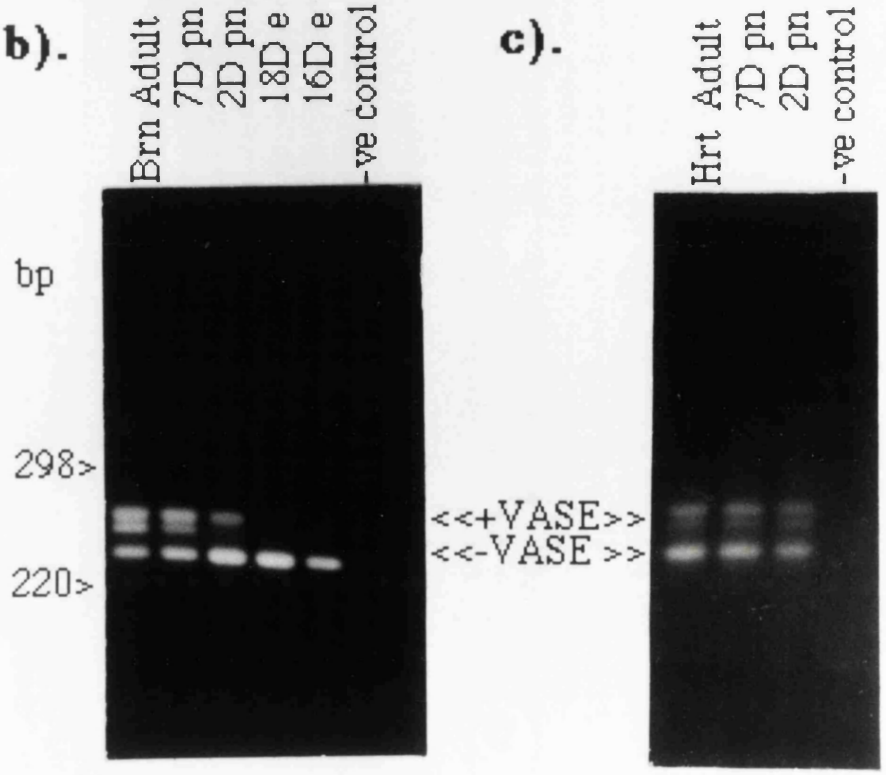
b). Amplification of rat brain developmental series cDNA.

D pn = days post natal D e = embryonic days

c). Amplification of rat heart developmental series cDNA.

D pn = days post natal


a).
-ve control
Sk muscle
Kidney
Heart
Brain
Markers
bp
+VASE>>
-VASE >>
<298
<201


b).
Brn Adult
7D pn
2D pn
18D e
16D e
-ve control
bp
298>
220>
c).
Hrt Adult
7D pn
2D pn
-ve control
<<+VASE>>
<<-VASE >>

4.3. c-src alternative splicing

4.3. Introduction

pp60$^{c\text{-}src}$ is produced in a distinct form which exhibits a tissue distribution corresponding to that of SmN. pp60$^{c\text{-}src}$ is the cellular homologue of the transforming 60 kilodalton phosphoprotein of RSV (for review see Hunter, 1987). pp60$^{c\text{-}src}$ is membrane associated and posesses a tyrosine-specific kinase activity, but its exact function and role in cellular regulation remains unclear.

A structurally distinct form of pp60$^{c\text{-}src}$, designated pp60$^{c\text{-}src+}$, from rat embryo brain cultures was found to be expressed at high levels and to have a high protein kinase activity (Brugge et al., 1985), and an extra serine phosphorylation site (Brugge et al., 1987), relative to constitutive pp60$^{c\text{-}src}$. The above kinase activity was shown to be regulated during fish, frog and avian development (Schartl et al., 1984) and c-src protein levels have been shown to be regulated during avian and human development being expressed at their highest levels in neuronal tissues and heart (Cotton & Brugge, 1983; Levy et al., 1984). This distinct neuronal specific protein was also expressed in mouse (Martinez et al., 1987) and avian (Levy et al., 1987) brain and was found to be encoded by a unique mRNA containing an 18 base pair alternative NI exon (cassette, see fig.1.2. N^{o} 1) encoding six amino acids (RKVDVR), between exons three and four (see fig.4.3.1.d.). Mouse c-src+ mRNA was undetectable in heart, kidney, lung, testis and spleen. The mouse and avian inserts encoded identical amino acids although there were four differences at the nucleotide level suggesting an important role for this alternative exon. pp60$^{c\text{-}src+}$ is specifically expressed in neuron-like cells derived from an EC cell line (Lynch et al., 1986) and several cell lines derived from human neuroblastomas, but not glioblastomas (Bolen et al., 1985).

The *cis* elements required for regulation of NI alternative splicing appear to be within the exon and its directly flanking sequences with the short size of the exon blocking its inclusion in non-neuronal cells (Black, 1991). Specifically a sequence downstream of the NI exon appears to promote neural specific inclusion of NI (Black, 1992).

It is possible that the biochemical properties of pp60$^{c\text{-}src+}$ are mediated by the inserted sequence (Levy & Brugge, 1989). Another exon designated NII was found to be inserted

into human brain c-src mRNA between exons three and four. This exon was 33 nucleotides in length, encoding eleven amino acids and a possible serine phosphorylation site (Pyper & Bolen, 1990). The NII exon was only found in mRNA also containing the NI exon in the form of c-src+NI/NII mRNA. The expression of a developmentally regulated neuronal c-src isoform suggests a role for this protein in differentiation and maintenance of the neuronal phenotype.

4.3.1. Quantitative splicing assay and cell line assays

c-src transcripts were found to be expressed as endogenous transcripts in all the tissues and cells examined except for F9 cells as shown previously (Black, 1991) so it was not necessary to obtain c-src minigene constructs for transfection. The PCR amplification is shown to be quantitative relative to input DNA by the dilution series (fig.4.3.1.a.). Fig.a. does not appear to show bands with the expected relative spacings for bands containing 18 and 51 additional base pairs respectively (see fig.4.3.1.d.). On separation of RT-PCR products under more denaturing conditions, i.e. in TAE buffer the assay does appear to produce bands with the expected spacings (fig.4.3.1.b.).

On examination of c-src alternative splicing in a number of neuronal and non-neuronal cell lines (fig.4.3.1.c.) it remains possible that SmN is responsible for regulating splicing of the NI exon, conferring the ability to follow the neuronal splice choice. Hence none of the cell lines examined showed inclusion of detectable levels of the NI exon as shown previously in the case of 3T3 and HeLa cells (Black,1991). The neuronal cell lines also exhibited a lack of NI exon inclusion however except the N18Tg2 cell line which appears to include both exons. This transcript was barely detectable in these cells however, possibly reflecting their low level of SmN expression. O1A1 EC cells have previously been shown to express $pp60^{c\text{-}src+}$ upon differentiation into neuron-like cells when grown in the presence of RA (Lynch et al., 1986). These cells however have reduced levels of SmN when differentiated in the same way (see fig.3.1.4.).

Fig 4.3.1.

c-src Alternative Splicing in Cell Lines

PCR amplifications of c-src cDNA were carried out using the oligonucleotide primers m-src3 and m-src4. 32 cycles were carried out at 94°C for 30 seconds, 65°C for 30 seconds and 72°C for 30 seconds. The resulting products were then separated on 2% agarose/ethidium bromide stained gels.

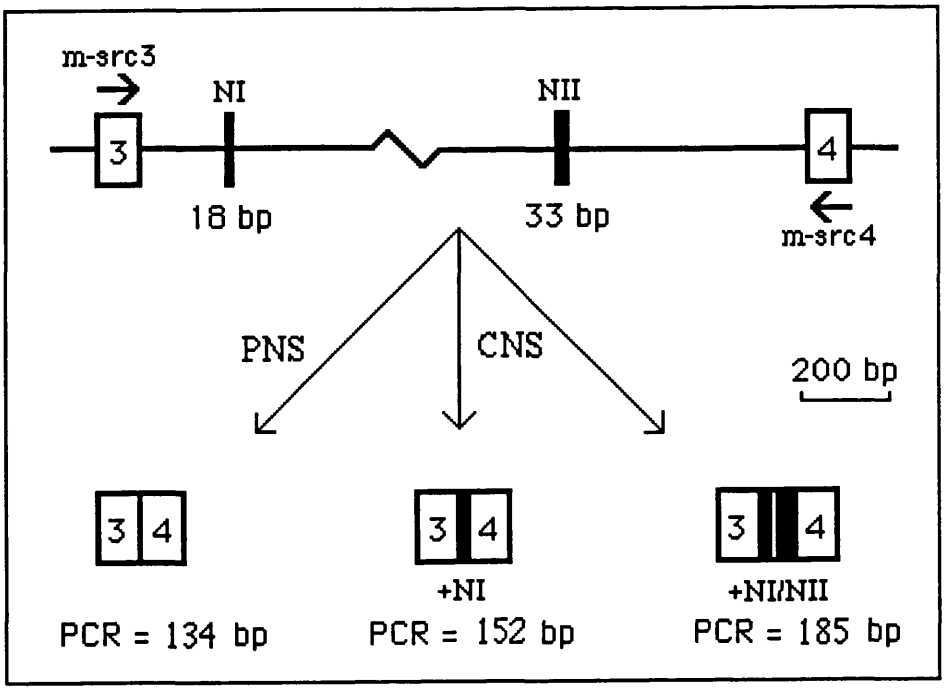
a). Amplification of a dilution series of 7 day post natal rat brain cDNA.

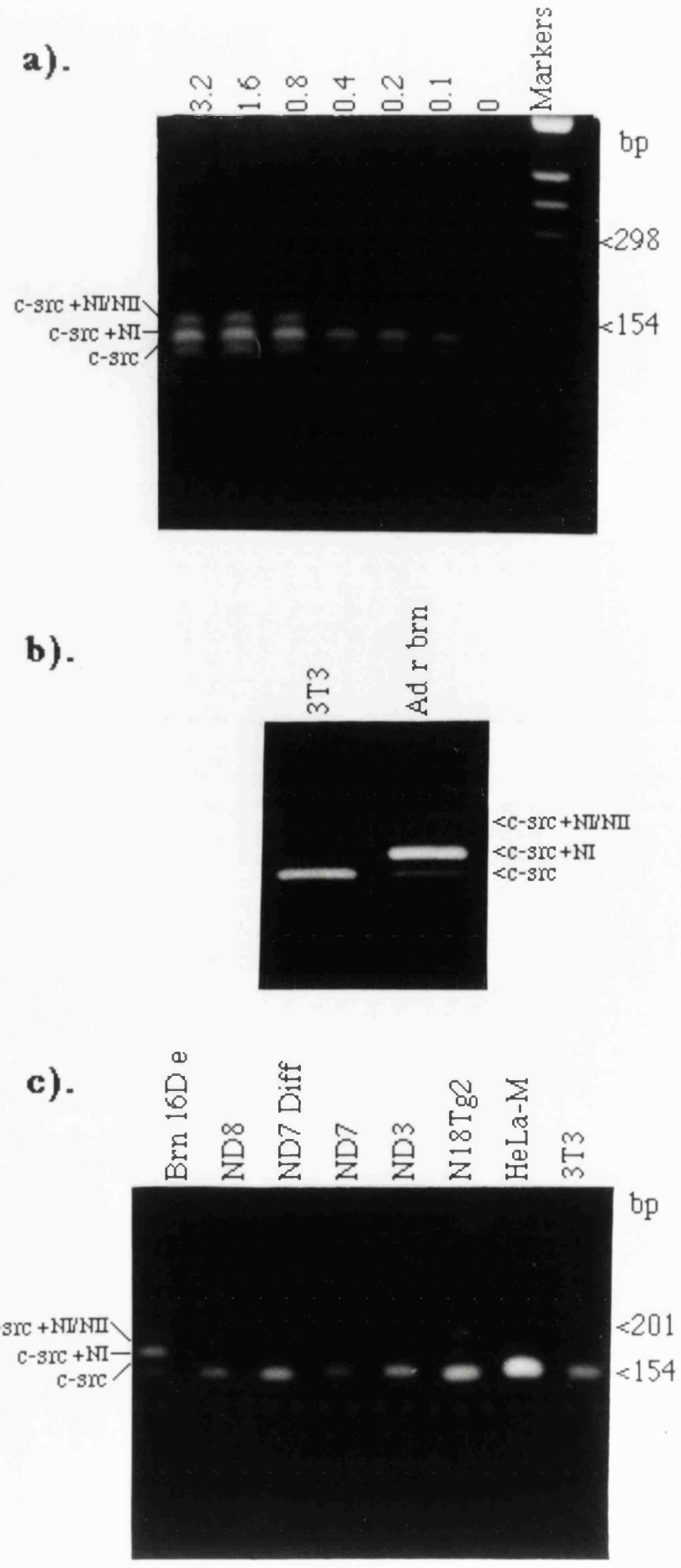
b). Resolution of the three c-src transcript derived products in TAE buffer.

c). Amplification of rodent cell line cDNA.
ND7 Diff = differentiated ND7 cells (see fig 3.1.4.c.).

d). Schematic representation of src alternative splicing. Small arrows indicate regions from which the oligonucleotide primers were designed for the splicing assay.

d).




a).
3.2
1.6
0.8
0.4
0.2
0.1
0
Markers
bp
298
154
c-src +NI/NII
c-src +NI
c-src
b).
3T3
Ad r brn
c-src +NI/NII
c-src +NI
c-src
c).
Brn 16D e
ND8
ND7 Diff
ND7
ND3
N18Tg2
HeLa-M
3T3
bp
201
154
c-src +NI/NII
c-src +NI
c-src

4.3.2. Tissue assays

The 18 base pair N1 and 33 base pair N2 exons are shown to be included in c-src transcripts in a tissue specific manner being more abundant in brain than in heart, and being undetectable in kidney and liver (fig.4.3.2.a.). Hence the distribution of this alternatively spliced transcript correlates well with SmN expression in rodent tissues. c-src+NI mRNA predominates in adult mouse and rat as well as human foetal brain. This transcript may only be abundant in the central nervous system as the c-src mRNA predominates in human spinal cord (fig.4.3.2.b.). The developmental series of rat brain and heart tissue cDNA (fig.4.3.3.c. and d.) shows developmental regulation of the N1 exon in brain development. Hence inclusion of this exon appears to increase during development, although not dramatically. The regulated splicing of the N1 exon in c-src transcripts does correlate with SmN expression during development as well as in the central and peripheral nervous system (fig.3.1.3. and 3.2.2.).

As in the case of NCAM above, an extra PCR product band is present in the brain tissue sample (see fig.4.3.2.a. *) possibly representing an as yet unidentified isotype of c-src which is expressed in a tissue specific and developmentally regulated fashion. The spacing of the bands suggests that it represents a transcript with a 60bp exon or combination of exons included between exons three and four. It may alternatively be a heteroduplex formed between the other bands for the same reasons mentioned for NCAM above.

Fig 4.3.2.

c-src Alternative Splicing in Tissues

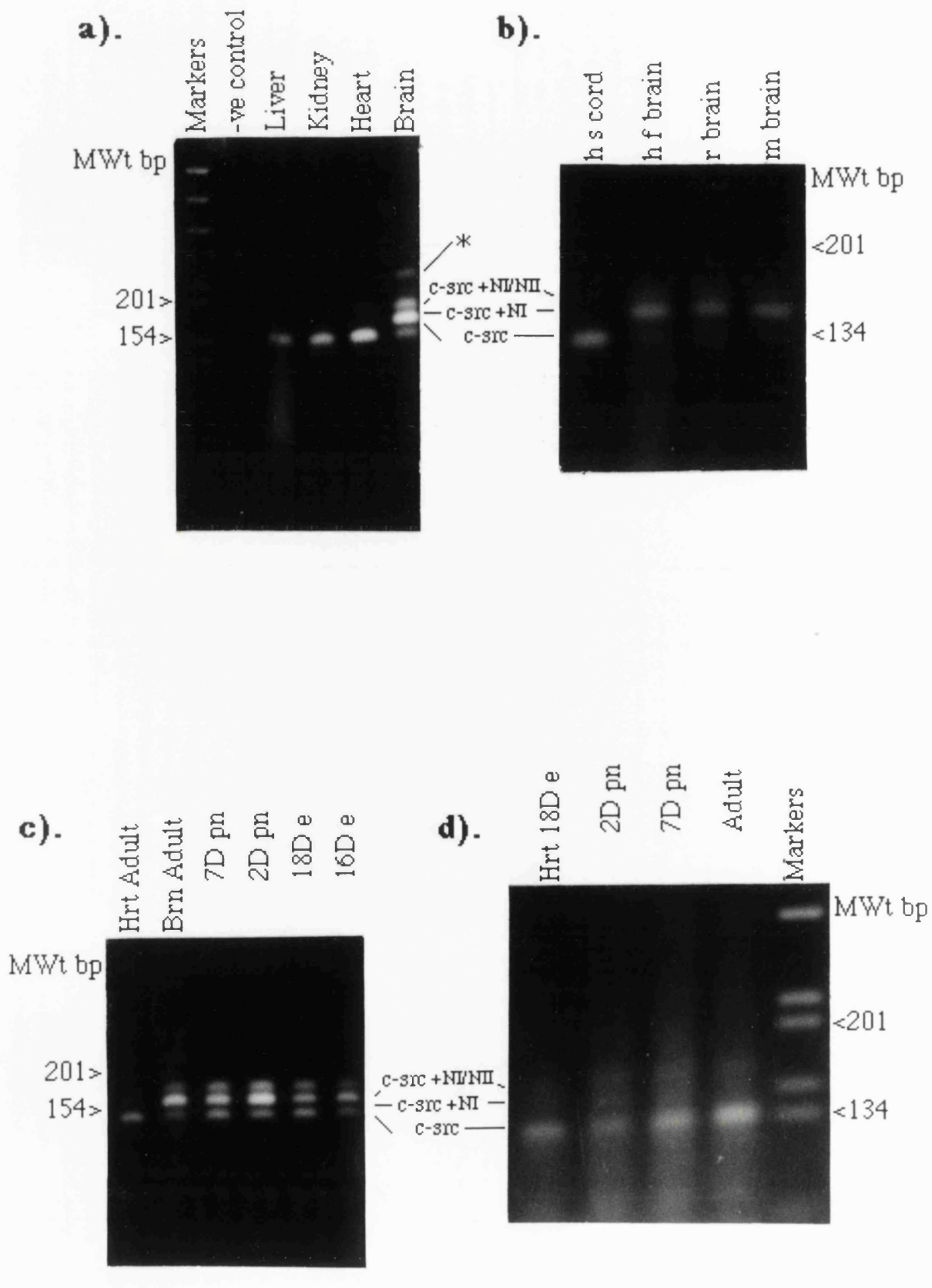
PCR amplifications of c-src cDNA and separation of products were carried out as in 4.3.1.

a). Amplification of mouse tissue cDNA.
(see text for discussion of band labelled *).

b). Amplification of mouse (m), rat (r) and human foetal (h f) brain and human spinal (h s) cord.

c). Amplification of rat brain developmental series cDNA.
D pn = days post natal D e = embryonic days

d). Amplification of rat heart developmental series cDNA.


a).
Markers
-ve control
Liver
Kidney
Heart
Brain
MWt bp
201>
154>
*
c-src +NI/NII
c-src +NI
c-src
b).
h s cord
h f brain
r brain
m brain
MWt bp
<201
<134
c).
Hrt Adult
Brn Adult
7D pn
2D pn
18D e
16D e
MWt bp
201>
154>
c-src +NI/NII
c-src +NI
c-src
d).
Hrt 18D e
2D pn
7D pn
Adult
Markers
MWt bp
<201
<134

4.4. G_0 GTP binding protein, α subunit alternative splicing

4.4. Introduction

The distribution of the $G_0\alpha$ subunit mRNAs shows some correlation with SmN expression in a number of tissues (Strathmann et al., 1990; Bertrand et al., 1990). G_0 is a heterotrimeric guanine nucleotide binding protein (G-protein) consisting of α, β and γ subunits. The α subunits are membrane bound and sensitive to inhibition via ADP-ribosylation by *Bordetella pertussis* exotoxin. G_0 may be a stimulator of phospholipase-C and potassium channels, an inhibitor of neuronal calcium channels and a mediator of membrane receptors (Bertrand et al., 1990). G_0 may also be activated by intracellular signals (Strittmatter et al., 1990). The human $G_0\alpha$ gene has at least 11 exons spread over more than 100 kilobase pairs.

Sixteen G-protein α subunits encoded by different genes are known, and this diversity of α subunits is increased by alternative splicing. In combination with a particular α subunit, at least three types of β and four types of γ subunits contribute to an even greater variety of heterotrimeric G-proteins. Alternative splicing of a single transcript encoding α_o subunits yields at least three mRNAs, designated α_{o1A}, α_{o1B} and α_{o2} (Bertrand et al., 1990). As α_{o1A} and α_{o1B} differ only in the 3' non-coding regions, they encode the same protein, namely the α_{o1} subunit. The α_{o2} mRNA differs from α_{o1} at its carboxyl end and encodes the α_{o2} subunit. Both proteins are the same length (see fig.4.4.1.c.). α_{o1} and α_{o2} mRNAs are generated by alternative splicing in hamster, rat (Hsu et al., 1990), mouse (Strathmann et al., 1990) and human (Tsukamoto et al., 1991) cells. Although the 5' and 3' regions of α_{o1A} mRNA are not translated the bovine, hamster and rat forms are almost identical.

The two α_o subunits differ only in their C-terminal regions, which is assumed to be the region responsible for determining interaction specificity for receptors and effectors. The G_{o1} and G_{o2} forms actually mediate inhibition of voltage dependent calcium channels through muscarinic and somatostatin receptors respectively (Kleuss et al., 1991).

4.4.1. Quantitative splicing assay and cell line assays

$G_o\alpha$ transcripts were found to be expressed as endogenous transcripts in most of the tissues and cells examined so it was not necessary to obtain $G_o\alpha$ minigene constructs. The PCR amplification is shown to be quantitative relative to input DNA by the dilution series and the oligonucleotides are shown to amplify bands of the correct sizes respectively (fig.4.4.1.a.). On examination of $G_o\alpha$ splicing patterns in various rodent cell lines SmN levels correlate well with expression of the O1A transcript. i.e. The O1A transcript is expressed at its highest level in PC12 cells which express more SmN than all other cell types tested (fig.4.4.1.b.).

4.4.2. Tissue assays

Expression of the $G_o\alpha$ transcripts is shown to be tissue specific being more abundant in brain and heart (fig.4.4.2.a.) as previously shown (Bertrand et al., 1990). Mouse brain appears to produce predominantly O2 mRNA whereas mouse heart as well as rat brain and heart generally produce similar quantities of O1A and O2 mRNA. The oligonucleotide sequences may be responsible for the absence of an α_{o2} band derived from human tissues. The antisense oligonucleotide sequences are derived from the mouse non-coding region and the sequence from which m-Go8 was derived (α_{o2} specific) does not appear to be conserved between mouse and human (Strathman et al., 1990; Tsukamoto et al., 1991). On examination of rat brain and heart developmental series' the relative abundance of the two transcripts does not appear to undergo any significant changes (fig.4.4.2.c.) indicating a loss of correlation between SmN expression and $G_o\alpha$ alternative splicing. i.e. SmN is expressed at increasing levels during brain development (3.1.3.).

It was also possible to detect the O1B (see 4.4.) transcript by increasing the elongation time during PCR amplification. This PCR product was approximately 1400 nucleotides in length indicated that this mRNA contains a retained, untranslated sequence of approximately 1000 nucleotides in length which is not present in the O1A mRNA.

Fig 4.4.1.

G_o, α Subunit Alternative Splicing in Cell Lines

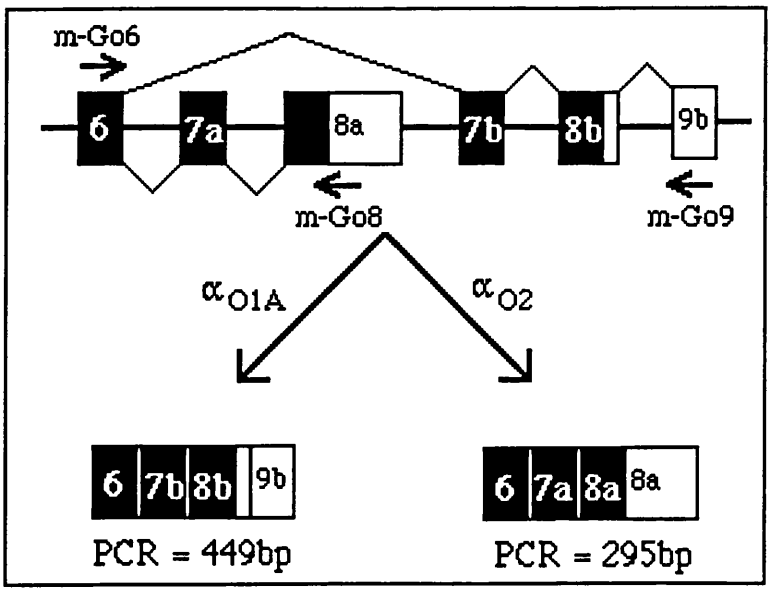
PCR amplifications of $G_o\alpha$ cDNA were carried out using the oligonucleotide primers m-Go6, m-Go8 and m-Go9. 32 cycles were carried out at 94°C for 45 seconds, 56°C for 45 seconds and 72°C for 45 seconds. The resulting products were then seperated on 1.6% agarose/ethidium bromide stained gels.

a). Amplification of a dilution series of 18 day embryo rat brain cDNA. mGo6+8 and m-Go6+9 represent amplifications with individual oligonucleotide pairs.

b). Amplification of rodent cell line cDNA.
ND7 Diff = differentiated ND7 cells (see fig 3.1.4.c.).

c). Schematic representation of $G_o\alpha$ alternative splicing. Small arrows indicate regions from which the oligonucleotide primers were designed for the splicing assay.

c).



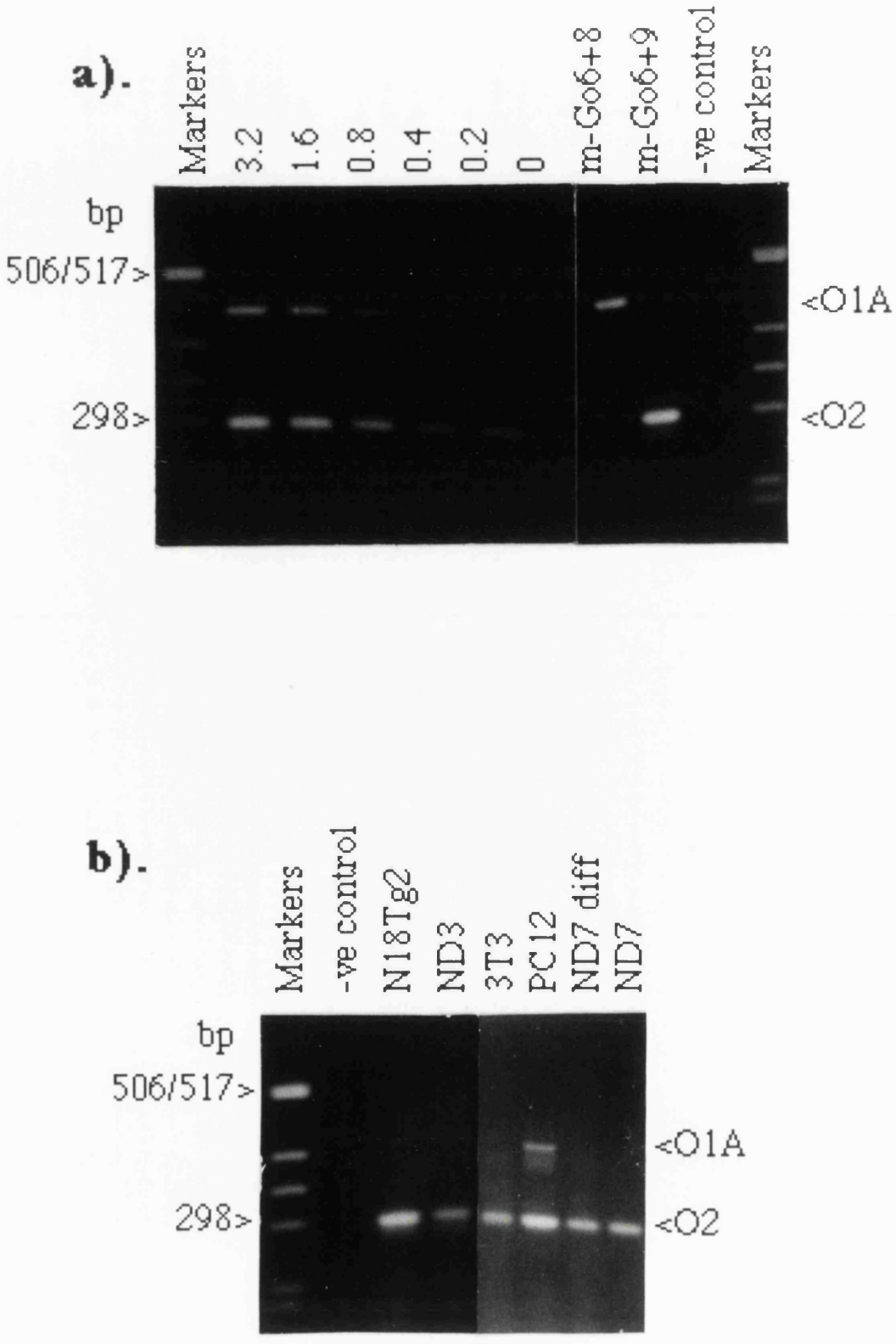

a).
Markers
3.2
1.6
0.8
0.4
0.2
0
m-Go6+8
m-Go6+9
-ve control
Markers
bp
506/517>
298>
<O1A
<O2
b).
Markers
-ve control
N18Tg2
ND3
3T3
PC12
ND7 diff
ND7
bp
506/517>
298>
<O1A
<O2

Fig 4.4.2.

G_o, α Subunit Alternative Splicing in Tissues

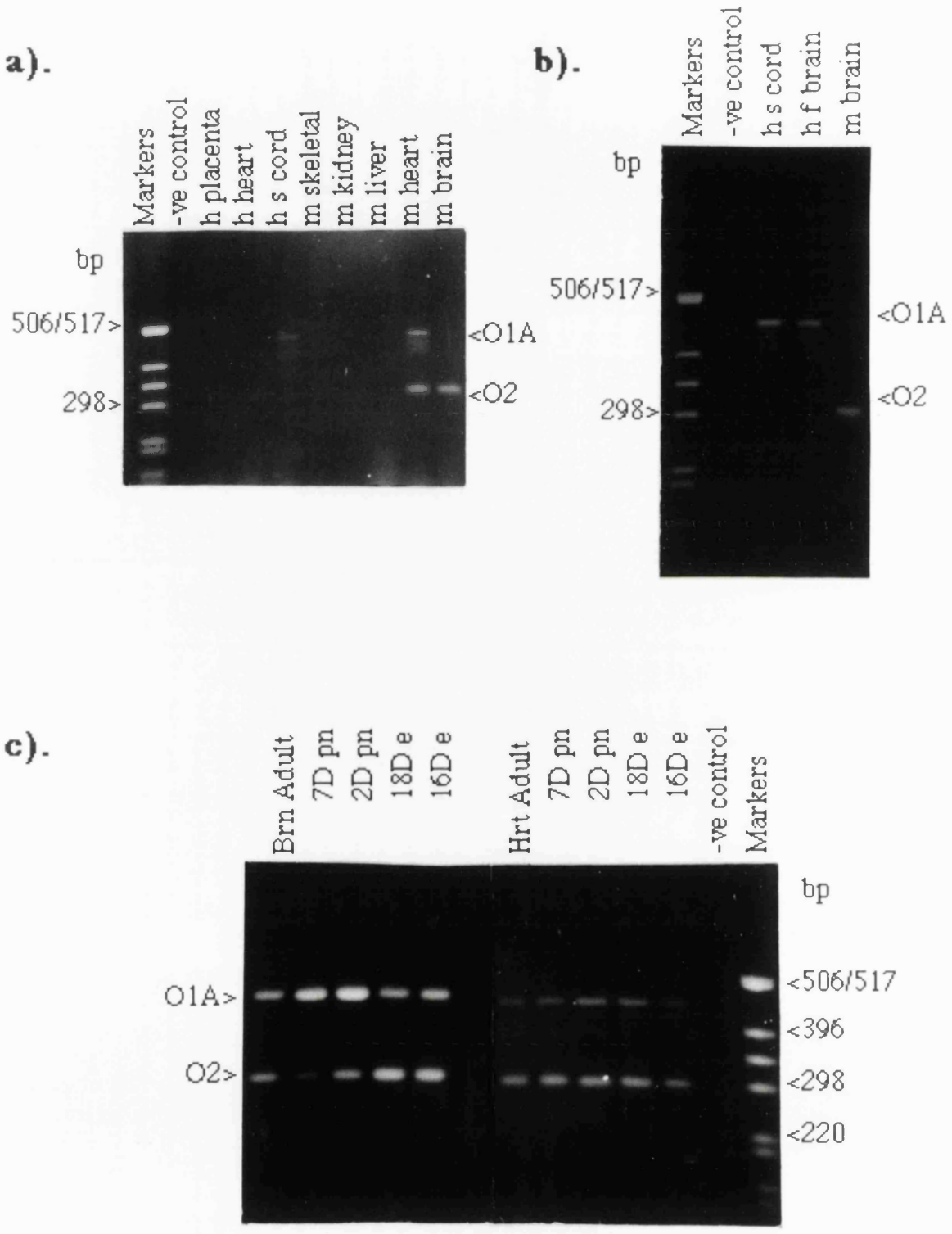
PCR amplifications of $G_o\alpha$ cDNA and separation of products were carried out as in 4.4.1.

a). Amplification of mouse and human tissue cDNA.

b). Amplification of mouse (m), and human foetal (h f) brain and human spinal (h s) cord.

c). Amplification of rat brain and heart developmental series cDNA.

D pn = days post natal D e = embryonic days


a).
Markers
-ve control
h placenta
h heart
h s cord
m skeletal
m kidney
m liver
m heart
m brain
bp
506/517>
298>
<O1A
<O2
b).
Markers
-ve control
h s cord
h f brain
m brain
bp
506/517>
298>
<O1A
<O2
c).
Brn Adult
7D pn
2D pn
18D e
16D e
Hrt Adult
7D pn
2D pn
18D e
16D e
-ve control
Markers
bp
O1A>
O2>
<506/517
<396
<298
<220

4.5. Cardiac troponin-T (cTNT) alternative splicing

4.5. Introduction

SmN is expressed at low levels in cardiac tissue suggesting that splice choices in this tissue may also be influenced by this protein. Troponin-T is the tropomyosin-binding subunit of the troponin complex (Flicker et al., 1982), playing an important role in regulating skeletal and cardiac muscle contraction (Adelstein & Eisenberg, 1980). The skeletal troponin-T gene undergoes alternative splicing (Breitbart & Nadal-Ginard, 1986) and the closely related avian and rat cTNT genes both generate two mRNA products by developmentally regulated alternative splicing (Cooper & Ordahl, 1985; Jin & Lin, 1989) although spatial regulation within the bovine heart is not seen (Jin & Lin, 1988).

The avian cTNT gene contains eighteen exons and is closely related to the skeletal muscle troponin-T. Exon five splicing is strictly developmentally regulated *in-vivo* being restricted to the embryonic stage of skeletal (Ordahl et al., 1980) and cardiac muscle (Cooper & Ordahl, 1985) and is responsible for the production of the two isotypes above. Exon 5 is a thirty base pair alternative exon (cassette, see fig 1.2. N^{o} 1 and fig 4.5.1.c.), but is only 50% conserved at the nucleotide level, between rat and chicken. Both encode highly acidic amino acid inserts however (rat=EDWSEEEEDE, avian=EEEEWLEEDD). It has been suggested that the role of embryonic cTNT is in the assembly of new sarcomeres (Cooper & Ordahl, 1985).

Exon five alone must be distinguished as an alternative exon and its frequency of inclusion must be modulated during the transition from embryo to adult. *Cis*-acting sequences involved in cTNT alternative splicing are limited to three small regions of the pre-mRNA containing exons four, five and six and constituting 1200 nucleotides (Cooper et al., 1988). Mutation of as few as four internal coding bases within the alternative exon can prevent exon recognition and a mutation within the 5' splice site of exon five, which increases its homology to the consensus sequence improves the efficiency of exon five inclusion (Cooper & Ordahl, 1989). It is not clear however if these are signals required for interaction with *trans*-acting factors involved in constitutive or alternative splicing or whether secondary or tertiary structure is disrupted by these mutations. Other secondary structures proposed to be involved in the splice choice had no influence when disrupted.

4.5.1. cTNT splicing in tissues and cells

SmN is expressed at low levels in cardiac muscle as well as being expressed at high levels in brain. SmN may influence the splicing of a number of genes in heart and brain therefore. cTNT undergoes alternative splicing in cardiac muscle and was also considered as a potential candidate for regulation by SmN. Expression from the cTNT promoter is restricted to striated muscle so to express abundant levels of cTNT pre-mRNA in diverse cell types a construct containing the alternative exon of the cTNT gene downstream of the RSV-LTR was obtained for transfection and subsequent assay by RT-PCR (see fig.4.5.1.c.). The avian cTNT minigene construct, SM-1 (Cooper & Ordahl, 1988) was used for this purpose and as in the case of the previous CALC-I assays, use of an oligonucleotide specific for the RSV-LTR in the PCR assay ensures that RT-PCR products are derived from the transfected minigene construct.

The SM-1 construct has previously been shown to produce processed transcripts including and excluding exon five in diverse cell types. Cultured fibroblasts and skeletal muscle cells have in fact been shown to produce significant quantities of mRNA lacking the alternative exon (Cooper et al., 1988). cTNT has already been shown to be developmentally regulated in the rat heart (Jin & Lin, 1989) To determine if this is also the case in exactly the rat heart samples which have been shown to express constant SmN levels during development (fig.3.1.3.b.) these samples were examined by Western blotting with a cTNT specific antibody (fig.4.5.1.a.). The *trans*-acting factor or factors involved in cTNT alternative splicing would be expected to be developmentally regulated in the heart. cTNT exon five splicing shows developmental regulation over a range during which SmN levels do not change suggesting that SmN does not regulate cTNT alternative splicing. SmN does not appear to be the factor involved in cTNT alternative processing.

If cTNT expression was to correlate with SmN expression, 3T3, HeLa and ND3 cells would express the embryonic transcript whereas ND7 cells may express the adult transcript (i.e. these cells most closely resemble the SmN/SmB ratio in adult heart). All cell types appear to express similar relative amounts of each transcript (fig.4.5.1.b.) again suggesting that SmN is not involved in the regulation of this splicing choice. This was seen previously in disparate cell types (Cooper et al., 1988). The proposed factor or machinery controlling

Fig 4.5.1.

cTNT Alternative Splicing

a). Western blot with a cTNT specific antibody with a rat heart developmental series.

D pn = days post natal

D e = embryonic days

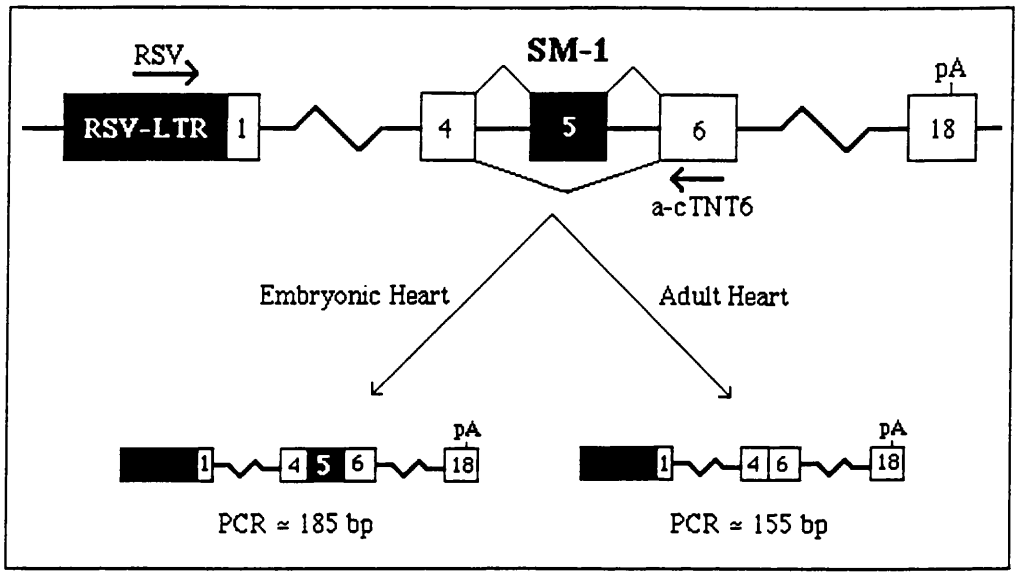
eTNT = embryonic cTNT (+ exon 5)

aTNT = adult cTNT (- exon 5)

b). PCR amplification of cTNT cDNA from rodent and human cells transfected with SM-1. PCR amplification was carried out using the oligonucleotides RSV and a-cTNT6. 40 cycles were carried out at 94°C for 30 seconds, 64°C for 30 seconds and 72°C for 30 seconds. The resulting products were seperated on a 2% agarose/ethidium bromide stained gel.

c). Schematic representation of the avian cTNT minigene construct, SM-1 and cTNT alternative splicing. Small arrows indicate regions from which the avian sequence and RSV oligonucleotide primers were designed for the splicing assay.

c).



a).
Hrt Adult
7D pn
2D pn
18D e
16D e
MWt kd
<<68
eTNT>
aTNT>
<<43
<<29
b).
bp
Markers
HeLa-M
HeLa-I
ND7
ND3
3T3
-ve control
201>
154>
<eTNT
<aTNT

cTNT alternative splicing appears to be present in muscle and non-muscle cells at equivalent active concentrations or the splicing decision in all these cells represents the constitutive choice while factors conferring bias to the splice choice are developmentally regulated and expressed only in the heart. Alternatively, the *cis*-acting regions controlling developmentally regulated splicing of exon five may have been lost during construction of the SM-1 plasmid (Cooper et al., 1988).

CHAPTER 5

RESULTS

Production of stable cell lines and subsequent splicing assays

5. Introduction

The nature of PCR assays involving the amplification of alternatively spliced products allows quantitative assays to be easily achieved (chapter 4). The same pair of oligonucleotides can generally be used for amplification of a number of different size transcripts which act as internal controls for one another in a competetive PCR assay (Gilliland et al., 1990). The dilution series' carried out above appear to indicate that the amplifications remain quantitative well into the visible range on ethidium bromide stained gels. The cell lines and assays above present models in which to test the relationship between SmN and alternative splicing further.

In order to directly test the ability of SmN to act as an alternative splicing factor it was necessary to alter the SmN/SmB ratio in cell lines and subsequently test the ratios of transcripts produced from alternatively spliced genes in these cells relative to the parental cells. Ideally a cell line could be produced in which SmN could be induced for subsequent assays of altered splicing patterns.

It was particularly necessary to test the src splice choice further as c-src+NI mRNA expression showed correlation with SmN expression in all the cells and tissues tested (see 4.3.).

5.1. Choice of expression vectors

In order to produce the above cell lines a series of expression vectors, pJ4-*cat* (MoMLV promoter), pJ5-*cat* (MMTV promoter) and pJ7-*cat* (CMV IE 94 promoter) (Morgenstern & Land, 1990) containing the CAT gene downstream of their respective promoters were assayed for CAT activity by transient transfection into ND7, 3T3 and HeLa-I cells. This was necessary to confirm the activity of the promoters in the cell lines as well as to assay the levels of induction possible from the MMTV promoter. These results may not relate directly to stable expression however as transient and stable expression from these promoters do show some differences. Transcription from most promoters is severely repressed in undifferentiated teratocarcinoma cells (Morgenstern & Land, 1990) so these cell types were not examined in these assays.

Fig 5.1.

pJ Vector Promoter Activity in Various Cell Lines

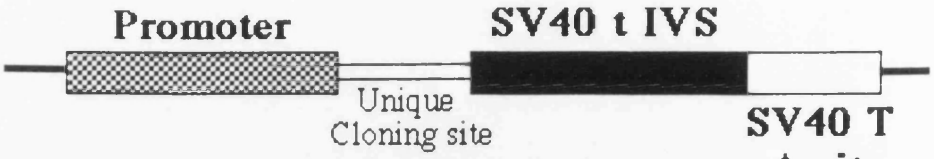
pJ4-*cat* (MoMLV promoter), pJ5-*cat* (MMTV promoter) and pJ7-*cat* (CMV IE 94 promoter) (Morgenstern & Land, 1990) were transfected into 3T3, ND7 and HeLa (I) cells in order to assay the activity of the various promoters in these cells. Dexamethasone was added 24 hours after transfection to a final concentration of 1*u*M to induce transcription from the MMTV promoter. 1, 10 and 100*u*M dexamethasone were previously found to show equivalent induction from the MMTV promoter. pJ5- = cells without dexamethasone. pJ5+ = cells incubated in dexamethasone. Cells were harvested 48 hours after transfection.

a). Schematic representation of the pJ series of vectors and their respective promoters.

b). 2 week exposure showing the CAT assay. Cell extracts were incubated for 3 hours at 37°C using 2, 40 and 75*u*g of ND7, HeLa and 3T3 protein respectively to correct for differences in transfection efficiency.

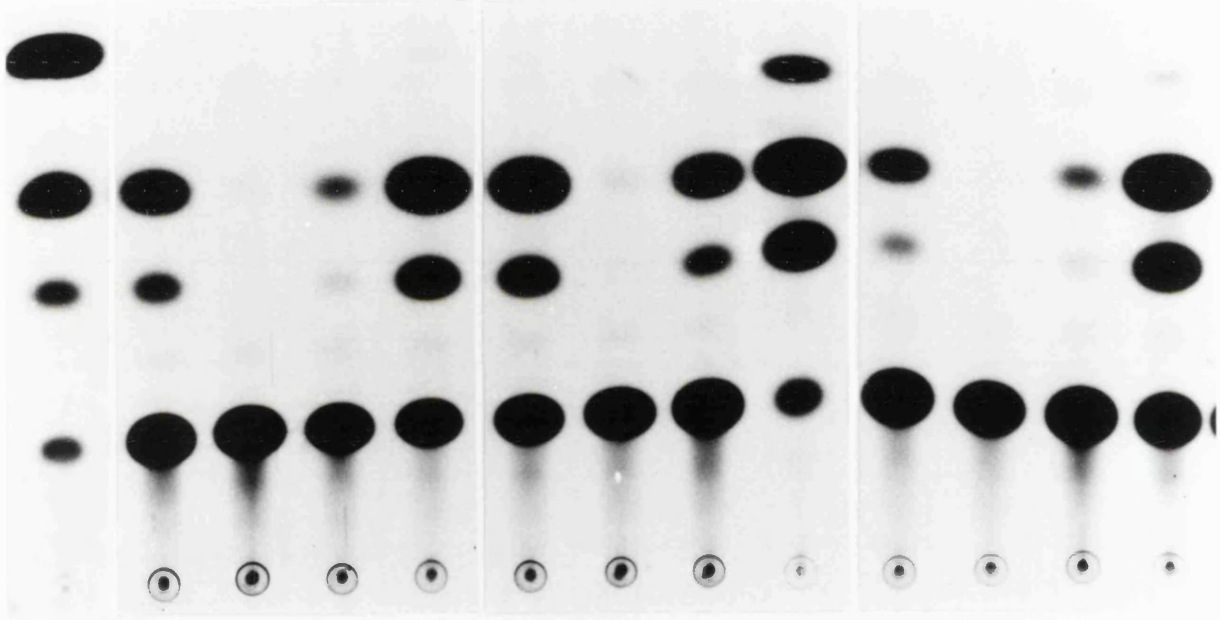
c). Bar chart showing % conversion values obtained after scintillation counting of excised bands. pJ7 is taken as 100% conversion in each case as comparisons between cell lines are not possible due to differences in transfection efficiency.

a).



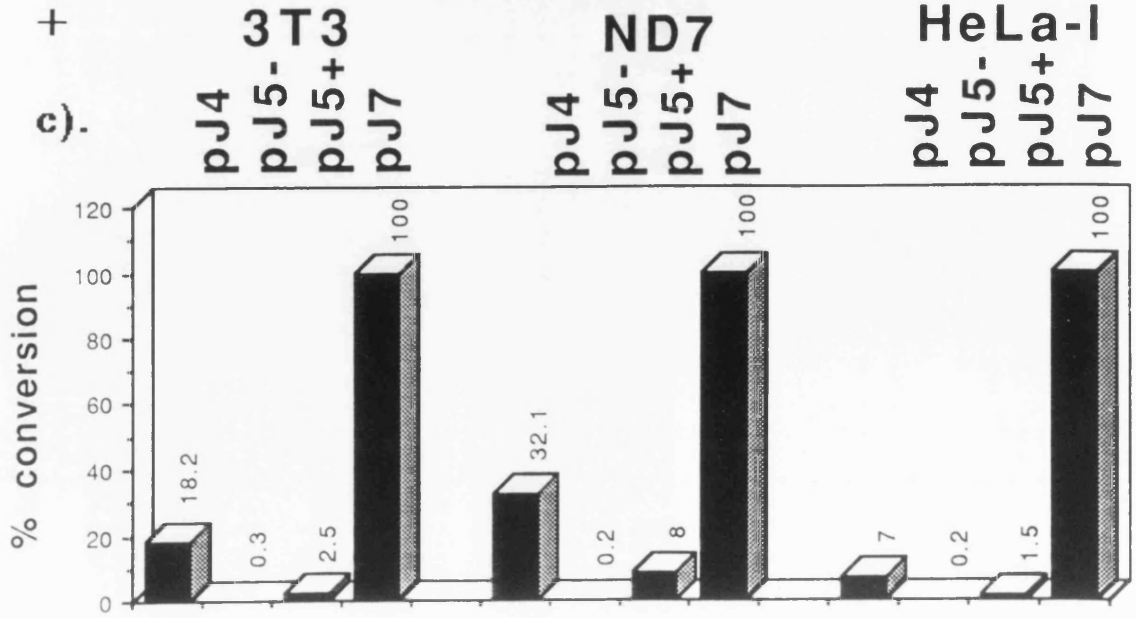
Vector	Promoter
pJ4	Mo-MULV LTR
pJ5	MMTV LTR
pJ7	Si CMV IE94

b).



\+

c).



The most efficient ratio of plasmid to carrier DNA was determined by transfecting various quantities of RSV-*cat* and carrier DNA into ND cells and the transfection efficiency of the various cell lines was then determined using RSV-*cat*.. The time of peak CAT activity after transfection was determined by extraction of ND8 cells at various times after transfection and the linearity of the CAT assay was also determined in a time course experiment. The CMV (pJ7) promoter was found to have the greatest activity in all three cell lines tested, while the MoMLV (pJ4) promoter showed intermediate activity and the MMTV (pJ5) promoter showed relatively weak activity after induction. The MMTV promoter was approximately 40 fold inducible in ND7 cells and approximately 8 fold inducible in 3T3 and HeLa-I cells. Dexamethasone has previously been shown to alter the calcitonin/CGRP splicing ratio (Cote & Gagel, 1986) but this did not appear to be the case in pRSV-Cal3 transfected ND7 cells. Also dexamethasone treatment did not affect the SmN/SmB ratio in these cells (see fig.5.2.).

5.2. Cloning and production of stable clones

The SmN cDNA (Gerrelli et al., 1991) and the SmB' cDNA (Van Dam et al., 1989) were cloned in the sense and antisense directions into a number of eukaryotic expression vectors. These vectors were pJ4, pJ5, pJ7, M5G-*neo* (MPSV promoter) and pHβApr1-*neo* (β actin promoter) (Gunning et al., 1987). Sense constructs were intended to increase the levels of expression while the antisense constructs were intended to inhibit expression (Izant & Weintraub, 1984; Kim & Wold, 1985; for review see Weintraub et al., 1985) or act as controls in 3T3 cells. The orientation of these clones was determined by restriction enzyme digestion. The cells chosen for alteration of SmN/SmB ratio were, i). 3T3 cells, as these cells are non-neuronal, express barely detectable levels of SmN and were previously shown to produce predominantly calcitonin when transfected with pRSV-Cal3 and, ii). ND7 cells, as these cells are neuronal, (the cells in which SmN is naturally expressed. i.e. these cells are more likely to express other factors which may be necessary for SmN function) express SmN and were previously shown to produce predominantly CGRP when transfected with pRSV-Cal3 (see 3.2. and 4.1.3.).

Fig 5.2.

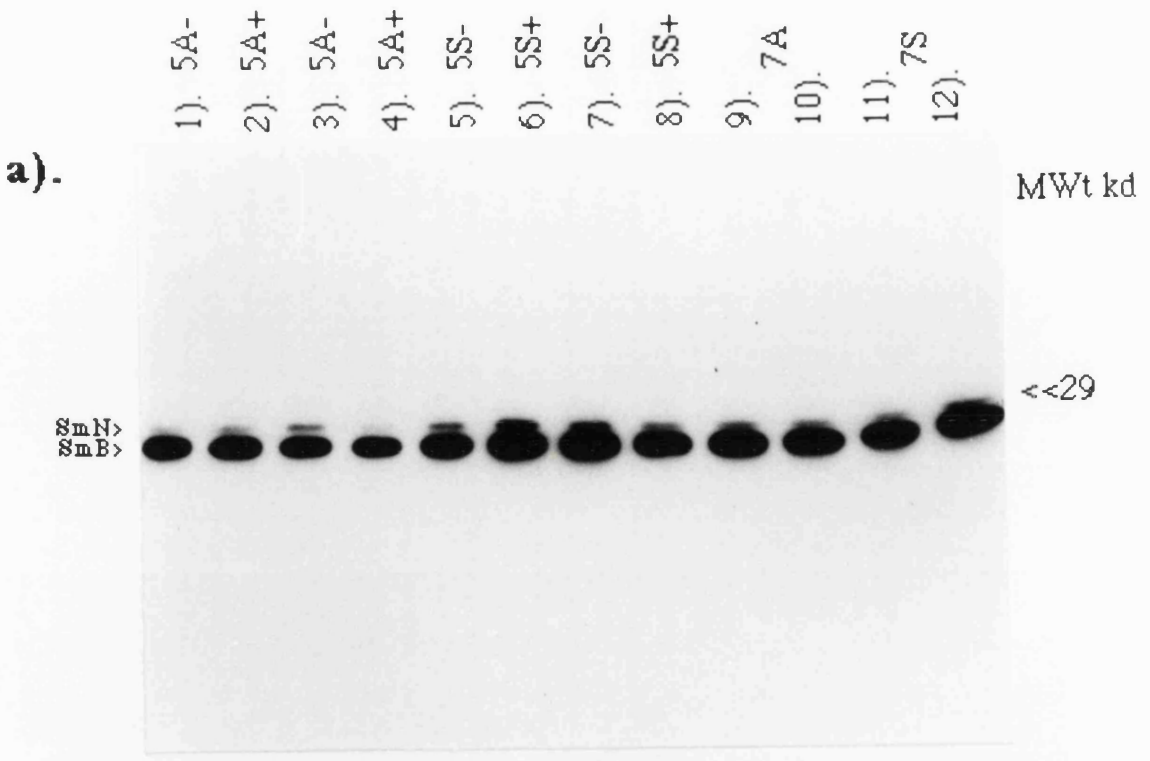
Cloning and Production of Stable Clones

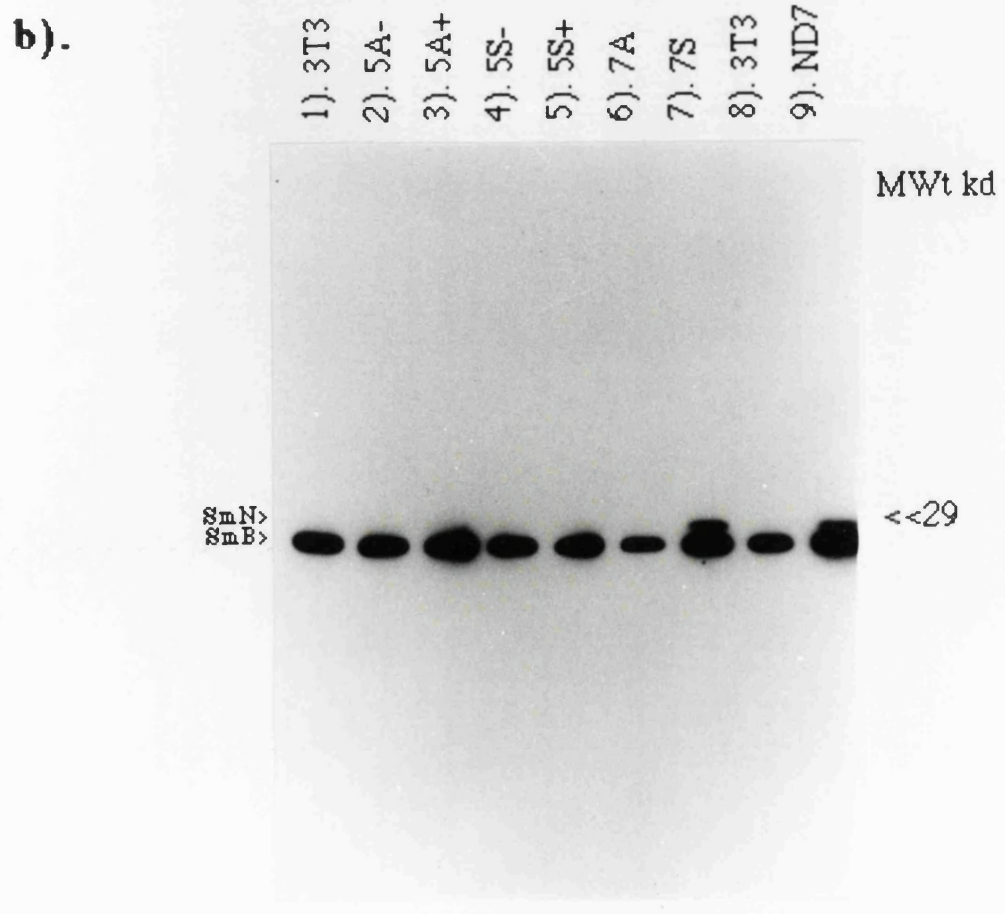
Clones of cells found to be resistant to Geneticin were assayed for alterations of SmN/SmB ratio by Western blotting with the kSm5 antibody. The SmN constructs used to produce each resistant clone is indicated.

e.g. 5A- indicates antisense (A, S = sense) SmN in pJ5 (5, 7 = pJ7). These cells were grown in the absence (-, + = presence of Geneticin) of Geneticin.

a). Geneticin resistant ND7 clones.

b). Geneticin resistant 3T3 clones.


1). 5A-
2). 5A+
3). 5A-
4). 5A+
5). 5S-
6). 5S+
7). 5S-
8). 5S+
9). 7A
10).
11). 7S
12).
a).
MWt kd
SmN>
SmB>
<<29


b).
1). 3T3
2). 5A-
3). 5A+
4). 5S-
5). 5S+
6). 7A
7). 7S
8). 3T3
9). ND7
MWt kd
SmN>
SmB>
<<29

Initially eight different constructs containing the SmN cDNA clone in the sense or antisense orientaion were cotransfected into ND7 cells with pRSV-Cal3. None of the SmN cDNA constructs was seen to alter the ratio of calcitonin/CGRP splicing. The same results were obtained with co-transfection experiments in HeLa cells (Li et al., 1989). Transient levels of Sm expression or antisense depletion were assayed after three days by Western blotting and none were found to have an altered SmN/SmB ratio. Transient expression or antisense depletion of SmN and SmB was therefore considered insufficient for quantitative assay of the effect of SmN/SmB ratio on alternative splicing. It therefore appeared that in order to alter the Sm ratio it was necessary to construct cell lines stably transfected, and stably expressing the SmN protein.

No ND7 clones found to be resistant to Geneticin selection showed sufficient alterations in SmN/SmB ratio. One ND7 clone transfected with antisense SmN under control of the MMTV promoter did appear to show a slight reduction in the relative expression of SmN after induction with dexamethasone (fig.5.2.a. track 4). The MMTV promoter may not be sufficiently active in stable clones, as in transiently transfected cultures, to show alterations in SmN/SmB ratio by Western blotting. Unfortunately no ND7 clones were found to exhibit significant antisense depletion of SmN. This may be due to the approximately 80% nucleotide homology between SmN and SmB clones within a species (coding region). Antisense SmN transcripts may therefore interfere with the production of both SmN and SmB peptides. To interfere with the production of a specific Sm peptide it may be necessary to produce antisense fragments derived from the non coding regions, which are more divergent (e.g. see Schmauss et al., 1992).

5.3. Stable expression of SmN

Some of the Genetecin resistant clones were found to have altered SmN/SmB ratios by Western blotting with the KSm5 antibody. The only constructs found to alter SmN/SmB ratios were pJ7 (3T3-7SN, see fig.5.3.b.) and M5G-*neo* (3T3-MSN, see fig.5.3.c.) containing SmN in the sense orientation. These constructs only appeared to be effective in 3T3 cells. High level expression driven from the CMV promoter has been shown previously (Boshart et al., 1985) but SmN expression driven from the MPSV

Fig 5.3.

Stable Expression of SmN Protein

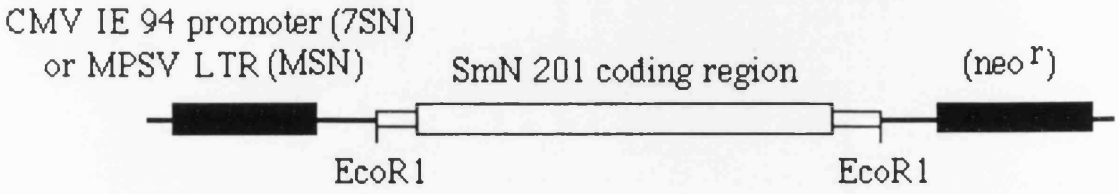
3T3 = parental cell line. Stable cell lines were selected by growth in 200*u*g/ml Geneticin.

a). Schematic representation of the SmN-201 cDNA clone (Gerrelli et al., 1991) inserted in the sense orientation into the M5G-*neo* vector or the pJ7 vector. The resulting constructs were MSN-*neo* and 7SN respectively.

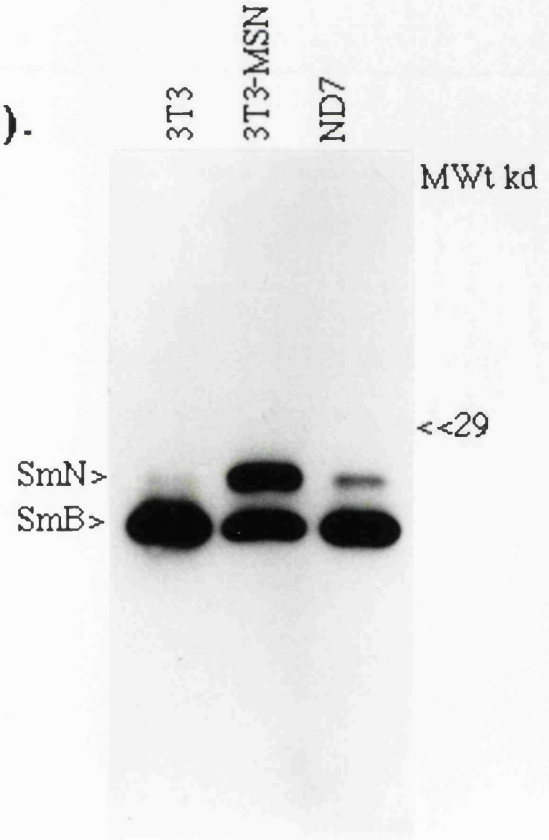
b). Western blot using the KSm5 antibody. 3T3-MSN = a clone of cells stably transfected with MSN-*neo*. ND7 cells were run in parallel to show the position of the SmN band and also for comparison of SmN/SmB ratios.

c). Western blot using the KSm5 antibody. 3T3-7SN = a clone of cells stably transfected with 7SN.

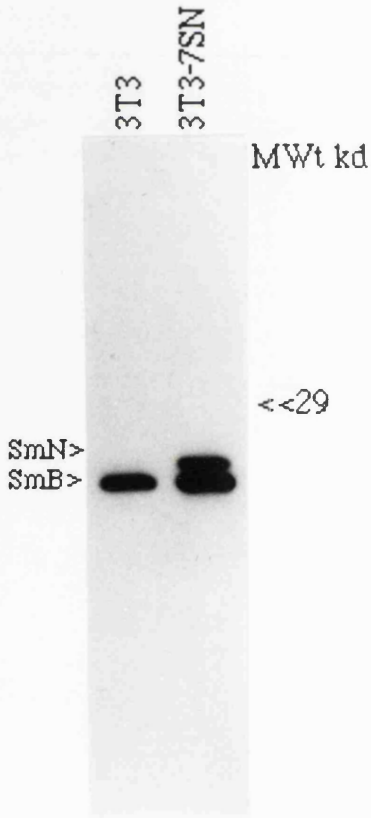
a).


CMV IE 94 promoter (7SN)
or MPSV LTR (MSN)
SmN 201 coding region
(neo r)
EcoR1
EcoR1

b).


3T3
3T3-MSN
ND7
MWt kd
<<29
SmN>
SmB>

c).


3T3
3T3-7SN
MWt kd
<<29
SmN>
SmB>

promoter (fig.5.3.c.) appeared to be greatest in this case. Antisense constructs (except as a control for sense constructs) and SmB' constructs were not used in conjunction with 3T3 cells as these cells naturally express only SmB. These 3T3 stable cell lines were suitable for assays of alternative splicing ratios.

Genetecin selection was removed after one month and the cells were once again assayed for SmN/SmB ratio by Western blotting after another four months. These cultures were found to be still expressing SmN and were therefore clonal and stable (3T3-MSN).

5.4. Splicing assays in stable cell lines

Prior to transfecting stable clones with pRSV-Cal3, DNA uptake efficiency was assayed by transfection with pJ7-*cat* (Morgenstern & Land, 1990). CAT assay showed all Geneticin resistant ND7 clones to have unaltered transfection efficiency. 3T3 clones however appeared to have reduced transfection efficiency relative to parental 3T3 cells, but as subsequent assays were to be carried out by PCR, DNA uptake levels were considered sufficient. Only the 3T3 stable cell lines were used in the subsequent assays.

To confirm that Calcitonin/CGRP (fig.5.4.a.), NCAM (fig.5.4.b.), c-src (fig.5.4.c.) and $G_o\alpha$ subunit (fig.5.4.d.) splicing are not influenced by SmN/SmB ratio these splicing decisions were examined in 3T3 parental cells in parallel with 3T3 cells stably expressing SmN (see fig.5.3.). All splicing assays were carried out as in chapter four. As suggested in chapter four SmN/SmB ratio does not affect the ratio of any of these splice products, including c-src.

Ectopic expression of SmN in HeLa cells has also shown that this protein is not sufficient to direct CGRP splicing (Delsert & Rosenfeld, 1992). Splicing assays in cells and tissues as well as in cells ectopically expressing SmN demonstrate that SmN is neither sufficient nor required for regulation of the above splice choices.

Fig.5.4.

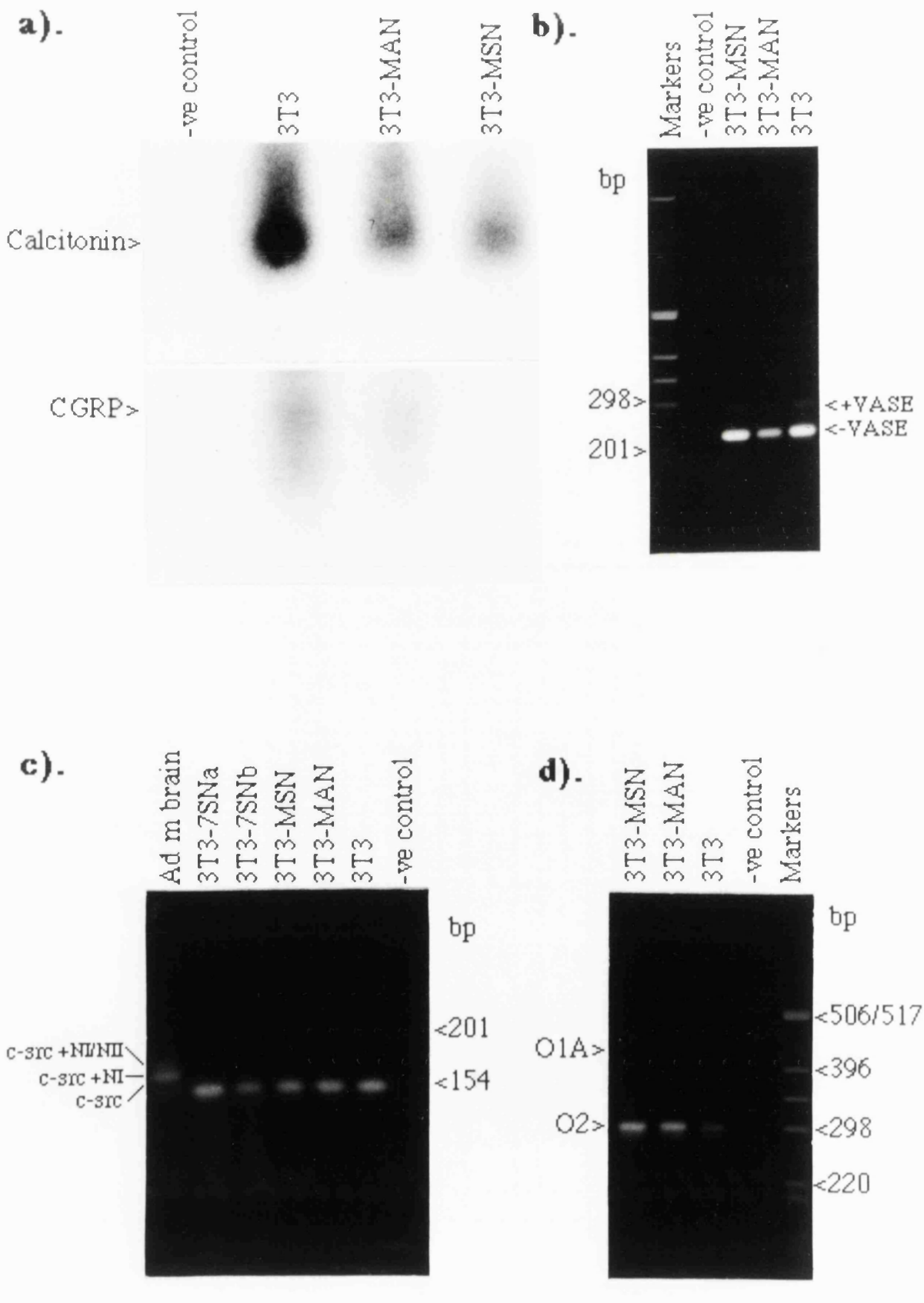
Splicing assays in stable cell lines

a). Calcitonin/CGRP assay carried out as in 4.1.

b). NCAM assay carried out as in 4.2.

c). c-src assay carried out as in 4.3.

d). $G_o\alpha$ assay carried out as in 4.4.


a).
-ve control
3T3
3T3-MAN
3T3-MSN
Calcitonin>
CGRP>
b).
Markers
-ve control
3T3-MSN
3T3-MAN
3T3
bp
298>
201>
<+VASE
<-VASE
c).
Ad m brain
3T3-7SNa
3T3-7SNb
3T3-MSN
3T3-MAN
3T3
-ve control
bp
<201
<154
c-src +NI/NII
c-src +NI
c-src
d).
3T3-MSN
3T3-MAN
3T3
-ve control
Markers
bp
<506/517
<396
<298
<220
O1A>
O2>

CHAPTER 6

DISCUSSION

6.1. SmN Expression

The expression pattern of the tissue specific splicing protein, SmN has been investigated in a number of cells and tissues, showing regulation during development and during EC cell differentiation. The abundance of the SmN protein and its distribution have been defined in the brain and the distribution of the closely related SmN, B and B' proteins has been defined in human and rodent tissues and cells. Clonal cell lines clearly are capable of expressing both SmN and SmB (see 3.1.2.) but it remains unclear whether individual rodent cells *in-vivo* express SmN or SmB or indeed if human cells express only one of the three closely related Sm peptides.

6.2. SmN and alternative splicing

SmN is 100% conserved from rodent to man (Schmauss et al., 1989) and is expressed in a highly tissue specific manner (McAllister et al., 1988) as well as being developmentally regulated (see 3.1.3.) being the first example of a mammalian snRNP protein of this kind. *Xenopus* also appears to express tissue specific and developmentally regulated Sm proteins (Fritz et al., 1984). The SmN gene structure has recently been resolved and SmN promoter fragments have been shown to drive differential expression of SmN at the level of transcription (Schmauss et al., 1992). Mutation at the SmN locus has also been implicated in Prader-Willi syndrome (Leff et al., 1992) suggesting an important function for this protein.

The SmN and SmB proteins are 90% related at the amino acid level suggesting that SmN in the brain replaces the constitutive functions of SmB in other tissues, presumably conferring a distinct neuronal role. SmB' presumably also has a similar but distinct function in human cells. The amino terminal 70% of the three peptides are in fact 96% identical (see fig.1.5.) suggesting that this region may constitute a domain responsible for the proposed constitutive role of these proteins while the carboxyl terminal 30% may confer specific functions. This may be particularly true in the case of the nine identical carboxyl terminal amino acids of SmN and SmB' which are absent in SmB.

SmN has been proposed to be a *trans*-acting component of the spliceosome which can confer the ability to follow an alternative splicing pathway in the processing of mRNAs which are alternatively spliced in a tissue specific manner. Characterisation of SmN expression in a number of cells and tissues allowed an investigation into the possible role of SmN in alternative splicing. A number of cell lines stably expressing SmN were also constructed in order to examine the role of SmN. These cell lines may also provide a model in which to test other potential functions of SmN. Assay of a variety of splicing decisions in cells and tissues and in cells ectopically expressing SmN demonstrate that SmN is neither sufficient nor required for regulation of a number of splice choices which previously correlated with SmN expression in the literature. These splicing decisions include CALC-I, NCAM (VASE exon), c-src, G_0 (α-subunit) and cTNT alternative splicing.

It is possible that SmN requires the action of another factor to influence alternative splicing decisions. This could explain why certain regions of the brain and certain cell lines including stable cell lines express SmN but are unable to produce the 'neuronal splice'. This does not however explain the ability of cells not expressing SmN to produce the 'neuronal splice' as in the case of ND3 (see 4.1.3.) and 293 cells (Delsert & Rosenfeld, 1992) which produce predominantly CGRP mRNA from the CALC-I gene.

Whether *trans*-acting factors which regulate these alternative splicing events are spliceosomal or are in fact independent of the spliceosome is unknown. Some such factors may bind directly to the pre-mRNA to block splicing via steric interference or to promote splice site usage by interaction with spliceosomal components. A number of *trans* factors which influence splice site selection have recently been identified (see table 1.3.). These factors appear to act by binding directly to the pre-mRNA substrate. If SmN does regulate other elusive alternative splicing events therefore (such as mutually exclusive splicing choices e.g. rat brain sodium channel (Sarao et al., 1991), see fig.1.2.), it probably does so by a mode of action distinct from other splicing regulatory factors. It seems likely however that an abundant factor such as SmN (Schmauss et al., 1992) would be present in the majority if not all active spliceosomes in the brain and would therefore exert its function on a variety or all genes expressed in neurons in the brain. It has been suggested that SmN

may have a lower affinity than SmB or SmB' for U1 (personal communication, J. Huntriss, UCMSM). Also SmB and SmB' in U1 may be antigenically distinct from those found in other snRNPs as they are recognised by anti-RNP antibodies (Petterson et al., 1984; Lelay-Taha et al., 1986).

It is probable that CALC-I alternative splicing as well as the splicing of other neuronal specific splice variants are produced via the action of an as yet unidentified tissue specific splicing factor that, like SmN, is present in most neuronal cells. This putatative factor or machinery probably acts by binding directly to the pre-mRNA in order to produce the CGRP splice choice in the case of the CALC-I gene (Leff et al., 1987). Regulation of CALC-I splicing may also require a factor or machinery which interacts directly with the pre-mRNA in order to produce the calcitonin splice choice (Cote et al., 1992). Both of the above putatative factors would appear to interact with *cis*-acting sequences either side of the calcitonin specific 3' splice site (Emeson et al., 1989; Cote et al., 1992). The autoimmune antigen HuD is only expresssed in brain and has RNA binding domains and homology to drosophila Elav and *Sxl* proteins (see table 1.3.). This protein is a candidate as a regulator of neuron specific pre-mRNA processing pathways (Szabo et al., 1991). HnRNP proteins also provide good candidates for regulation of alternative splicing events (Bandziulis et al., 1989; Mayeda & Krainer, 1992). The ND neuronal cell lines clearly provide a convenient source for the isolation of the regulatory factor or machinery responsible for control of CALC-I and other alternative splicing events as well as providing a model for a number of the splicing choices examined.

Characterisation of *cis*-acting elements involved in regulating alternative splicing and development of *in-vitro* systems in which these elements are still functional (e.g. c-src, Black, 1992) present the possibility of carrying out RNA/protein, U.V. cross-linking experiments. This type of experiment should soon lead to the isolation of a number of factors involved in the above and other regulated alternative splicing events.

6.3. Possible functions of Sm proteins

If SmN is not involved in alternative splicing it may be involved in another distinct step in snRNP action, biosynthesis or metabolism. An examination and discussion of the structures which contain Sm proteins and, for which functional information is available may provide clues as to the specific functions of the Sm proteins, specifically SmN, SmB and SmB'. A determination of the role of SmN is likely to lead to determination of the role of the related, SmB and SmB' proteins.

The spliceosome functions as a complex and it has therefore proved difficult to assign specific functions to spliceosomal components. The major immunoreactive Sm proteins are SmN, SmB, SmB' and SmD (i.e. the monoclonal antibodies, KSm4 recognises an epitope on SmB, SmB' and SmD; KSm5 recognises an epitope on SmB, SmB' and SmN and Y12 recognises an epitope on SmB, SmB', SmD and less efficiently on SmN (see Williams et al., 1986; Schmauss & Lerner, 1990; fig.3.1.5.)). SmD does not appear to be related to the other Sm proteins at the amino acid level but it shares epitopes with the other Sm proteins suggesting some similarity in higher order structure (Williams et al., 1986). A repeated proline rich domain appears to be responsible for SmN, SmB antigenicity and role in SLE (Williams et al., 1990). It is unclear whether this region is of particular importance in the proteins function.

Immunoreactive Sm proteins are necessary but not essential for cap trimethylation (Mattaj, 1986) (i.e. U3 is not complexed with Sm proteins but has an m_3G cap). Also a number of snRNPs which are complexed with Sm proteins do not appear to be involved in pre-mRNA processing, but are necessary for other processes (e.g. U7 (Yang et al., 1981), U11, U12 (Montzka & Steitz, 1988)). U7 is involved in histone pre-mRNA 3' end formation (Gilmartin et al., 1988). The cleavage reaction prior to polyadenylation is sensitive to inhibition by anti Sm monoclonal antibodies (Hashimoto & Steitz, 1986). U11 may be the snRNP involved in this step of polyadenylation. Also U12 is not involved in general splicing but is complexed with Sm proteins (Montzka & Steitz, 1988). The Sm binding site in U4 has been shown to be dispensable for spliceosome assembly (Wersig & Bindereif, 1990) and splicing (Bruzik & Steitz, 1990).

Due to the fact that Sm proteins are present in a number of snRNPs which do not

appear to be involved in splicing and the ability of spliceosomes to form with a deficit of Sm proteins, these proteins have been proposed to have a role in cytoplasm-nuclear transport and localisation (see 1.3.2.). Sm proteins have also been suggested to have a stabilising function in U1 (Hamm et al; 1987) U2 snRNPs (Mattaj & De Robertis, 1985), and yeast snRNPs (Jones & Guthrie, 1990).

Signals clearly exist to localise snRNPs to the nucleus as well as to specific positions within the nucleus (Carmo-Fonseca et al., 1992). The Sm proteins appear to provide part of a bipartite signal for nuclear targetting (Hamm et al., 1990; Fischer & Luhrmann, 1990) (see 1.3.2.). If Sm proteins provide part of a karyophilic signal it is possible that SmN confers altered properties to snRNPs, relating to interaction with nuclear pores, such as presenting a specific karyophilic domain. These properties would be specific to rodent neurons in particular. SmB' may also confer specific localisation signals to human cells expressing this protein.

Determining the role of the immunoreactive Sm proteins, as well as providing an insight into the role of the protein component of the spliceosome and its components, may lead to an understanding of the mechanisms responsible for the autoantibody response in SLE patients.

6.4. Future Work

This work clearly encourages further conjecture on the role of the SmN protein. The 3T3 cell lines ectopically expressing SmN, and the other cell lines examined provide a system in which to test further questions on the function of this protein. An SmN specific antibody would clearly be a useful tool in later experiments of this nature. The presence of high antibody titres in SLE patients with CNS dysfunction suggests that it may be possible to produce such a monoclonal antibody (see Winfield et al., 1978; Schmauss et al., 1992).

By constructing further stable cell lines expressing specific regions or putatative domains of SmN it may be possible to examine any effect on snRNP interaction and transport using m_3G antibodies.

Transgenic knockout mice not expressing SmN may provide a useful model in which to examine the function of SmN (Schmauss et al., 1992). These mice, in conjunction with mice which lack a functional SmN gene due to imprinting on both chromosomes (Cattanach et al., 1992) would also provide an insight into the role of SmN in Prader-Willi syndrome.

CHAPTER 7

REFERENCES

Adelstein, R. S. & Eisenberg, E. (1980) Ann. Rev. Biochem. 49 921-956.

Adema, G. J. & Baas, P. D. (1991) Biochem. Biophys. Res. Comm. 178 985-992.

Adema, G. J. & Baas, P. D. (1992) J. Biol. Chem. 267 7943-7948.

Aebi, M. & Weissmann, C. (1987) Trends in Gen. 3 102-107.

Alevizaki, M; Shiraishi, A; Rassool, F. V; Ferrier, G. J. M; Macintyre, I. & Legon, S. (1986) FEBS Lett. 206 47-52.

Amara, S. G; Evans, R. M. & Rosenfeld, M. G. (1984) Mol. Cell. Biol. 4 2151-2160.

Amara, S. G; Jonas, V; O'Neil, J. A; Vale, W; Riuier, J; Roos, B. A; Evans, R. M. & Rosenfeld, M. G. (1982) J. Biol. Chem. 257 2129-2132.

Amara, S. G; Jonas, V; Rosenfeld, M.G; Ong, E. S. & Evans, R. M. (1982) Nature 298 240-244.

Anderson, G. J; Bach, M; Luhrmann, R. & Beggs, J. D. (1989) Nature 342 819-821.

Andersen, J. & Zeive, G. W. (1991) Bioessays 13 57-64.

Bach, M; Krol, A. & Luhrmann, R. (1990) Nucl. Acids Res. 18 449-457.

Bach, M; Winkelmann, G. & Luhrmann, R. (1989) Proc. Natl. Acad. Sci. USA 86 6038-6042.

Baker, B. S. (1989) Nature 340 521-524.

Bandziulis, R. J; Swanson, M. S. & Dreyfuss, G. (1989) Genes & Dev. 3 431-437.

Barabino, S. M. L; Blencowe, B. J; Ryder, U; Sproat, B. S. & Lamond, A. I. (1990) Cell 63 293-302.

Barabino, S. M. L; Sproat, B. S; Ryder, U; Blencowe, B. J. & Lamond, A. I. (1989) EMBO J. 8 4171-4178.

Barbas, J. A; Chaix, J; Steinmetz, M. & Goridis, C. (1988) EMBO J. 7 625-632.

Barthels, D; Santoni, M; Wolfgang, W; Ruppert, C; Chaix, J; Hirsch, M; Fontecilla-Camps, J. C. & Goridis, C. (1987) EMBO J. 6 907-914.

Bass, B. L. & Cech, T. R. (1984) Nature 308 820-826.

Behrens, S. & Luhrmann, R. (1991) Genes & Dev. 5 1439-1452.

Bektesh, S; Van Doren, K. & Hirsh, D. (1988) Genes & Dev. 2 1277-1283.

Benne, R. (1990) T.I.G. 6 177-181.

Berget, S. M. & Robberson, B. L. (1986) Cell 46 691-696.

Berkovitz, L. A. & Gilman, M. Z. (1990) Proc. Natl. Acad. Sci. USA 87 5258-5262.

Bertrand, P; Sanford, J; Rudolph, U; Codina, J. & Birnbaumer, L. (1990) J. Biol. Chem. 265 18576-18580.

Beyer, A, L. & Osheim, Y. N. (1988) Genes & Dev. 2 754-765.

Bies, R. D; Phelps, S. F; Cortez, D. M; Roberts, R; Caskey, C. T. & Chamberlain, J. S. (1992) Nucl. Acids. Res. 20 1725-1731.

Bindereif, A. & Green, M. R. (1987) EMBO J. 6 2415-2424.

Black, D. L. (1991) Genes & Dev. 5 389-402.

Black, D. L. (1992) Cell 69 795-807.

Black, D. L; Chabot, B. & Steitz, J. A. (1985) Cell 42 737-750.

Black, D. L. & Steitz, J. A. (1986) Cell 46 697-704.

Blencowe, B. J; Sproat, B. S; Ryder, U; Barabino, S. & Lamond, A. I. (1989) Cell 59 531-539.

Bolen, J. B; Rosen, N. & Israel, M. A. (1985) Proc. Natl. Acad. Sci. USA 82 7275-7279.

Bond, R. W; Wyborski, R. J. & Gottlieb, D. I. (1990) Proc. Natl. Acad. Sci. USA 87 8771-8775.

Boshart, M; Weber, F; Jahn, G; Dorsch-Hasler, K; Fleckenstein, B. & Schaffner, W. (1985) Cell 41 521-530.

Bovenberg, K. A. L; Adema, G. J; Jansz, H. S. & Baas, P. D. (1988) Nucl. Acids Res. 16 7867-7883.

Bovenberg, R. A. L; van de Meerendonk, W. P. M; Baas, P. D; Steenbergh, P. H; Lips, C. J. M. & Jansz, H. S. (1986) Nucl. Acids Res. 14 8785-8803.

Brain, S. D; Williams, T. J; Tippins, J. R; Morris, H. R. & MacIntyre, I. (1985) Nature 313 54-56.

Branlant, C; Krol, A; Ebel, J; Lazar, E; Haendler, B. & Jacob, M. (1982) EMBO J. 1 1259-1265.

Breitbart, R. E; Andreadis, A. & Nadal-Ginard, B.(1987) Ann. Rev. Biochem. 56 467-495.

Breitbart, R. E. & Nadal-Ginard, B. (1986) J. Mol. Biol. 188 313-324.

Breitbart, R. E. & Nadal-Ginard, B. (1987) Cell 49 793-803.

Breitbart, R. E; Nguyen, H. T; Medford, R. M; Destree, A. ; Mahdavi, V. & Nadal-Ginard, B. (1985) Cell 41 67-82.

Broad, P. M; Symes, A. J; Thakker, R. V. & Craig, R. K. (1989) Nucl. Acids Res. 17 6999-7011.

Brugge, J. S; Cotton, P. C; Lustig, W; Yonemoto, L; Lipsich, P; Coussens, J. N; Barrett, D. N. & Keane, R. W. (1987) Genes & Dev. 1 287-296.

Brugge, J. S; Cotton, P. C; Queral, A. E; Barrett, J. N; Nonner, D. & Keane, R. W. (1985) Nature 316 554-557.

Bruzik, J. P. & Maniatis, T. (1992) Nature 360 692-694.

Bruzik, J. P. & Steitz, J. A. (1990) Cell 62 889-899.

Bruzik, J. P; Van Doren, K; Hirsh, D. & Steitz, J. A. (1988) Nature 335 559-562.

Buchanan, R. R. C; Venables, P. J. W; Morgan, A; Staines, N. A; Smith, P. R. & Maini, R. N. (1983) Clin. Exp. Immunol. 51 8.

Burd, C. G; Swanson, M. S; Gorlach, M. & Dreyfuss, G. (1989) Proc. Natl. Acad. Sci. U.S.A. 86 9788-9792.

Burtis, K. C. & Baker, B. S. (1989) Cell 56 997-1010.

Busch, H; Reddy, R; Rothblum, L. & Choi, Y. C. (1982) Ann. Rev. Biochem. 51 617-654.

Caras, I. W; Davitz, M. A; Rhea, L; Weddell, G; Martin, D. W. & Nuzsenzweig, V. (1987) Nature 325 545-549.

Carmo-Fonseca, M; Pepperbok, R; Carvalho, M. T. & Lamond, A. I. (1992) J. Cell Biol. 117 1-14.

Carmo-Fonseca, M; Pepperbok, R; Sproat, B. S; Ansorge, W; Swanson, M. S. & Lamond, A. I. (1991b) EMBO J. 10 1863-1873.

Carmo-Fonseca, M; Tollervey, D; Pepperbok, R; Barabino, S. M. L; Merdes, A; Brunner, C; Zamore, P. D; Green, M. R; Hurt, E. & Lamond, A. I. (1991a) EMBO J. 10 195-206.

Cattanach, B. M; Barr, J. A; Evans, E. P; Burtenshaw, M; Beechey, C. V; Leff, S. E; Brannan, C. I; Copeland, N. G; Jenkins, N. A. & Jones, J. (1992) Nature Genetics 2 270-274.

Cech, T. R. (1986) Cell 44 207-210.

Cech, T. R. & Bass, B. L. (1986) Ann. Rev. Biochem. 55 599-629.

Chabot, B; Black, D. L; LeMaster, D. M. & Steitz, J. A. (1985) Science 230 1344-1349.

Chang, D. D. & Sharp, P. A. (1989) Cell 59 789-795.

Choi, Y. D; Grabowski, P. J; Sharp, P. A. & Dreyfuss, G. (1986) Science 231 1534-1539.

Chu, J. & Elkon, K. B. (1990) Gene 97 311-312.

Cleveland, D. W. & Yen, T. J. (1989) New Biologist 1 121-126.

Cohen, P. L. & Eisenberg, R. A. (1991) Ann. Rev. Immunol. 9 243-269.

Cohen, P. L. & Eisenberg, R. A. (1992) Immunol. Today 13 427-428.

Conklin, P. L; Wilson, R. K. & Hanson, M. R. (1991) Genes & Dev. 5 1407-1415.

Cooper, T. A; Cardone, M. H. & Ordahl, C. P. (1988) Nucl. Acids Res. 16 8443-8465.

Cooper, T. A. & Ordahl, C. P. (1985) J. Biol. Chem. 260 11140-11148.

Cooper, T. A. & Ordahl, C. P. (1989) Nucl. Acids Res. 17 7905-7921.

Cote, G. J. & Gagel, R. F. (1986) J. Biol. Chem. 261 15524-15528.

Cote, G. J; Nguyen, I. N; Lips, C. J. M; Berget, S. M. & Gagel, R. F. (1991) Nucl. Acids Res. 19 3601-3606.

Cote, G. J; Stolow, D. T; Peleg, S; Berget, S. M. & Gagel, R. F. (1992) Nucl. Acids Res. 20 2361-2366.

Cotton, P. C. & Brugge, J. S. (1983) Mol. Cell. Biol. 3 1157-1162.

Crabtree, G. R. & Kant, J. A. (1982) Cell 31 159-166.

Craig, R. K; Hall, L; Edbrooke, M. R; Allison, J. & MacIntyre, I. (1982) Nature 295 345.

Crenshaw, E. B; Russo, A. F; Swanson, L. W. & Rosenfeld, M. G. (1987) Cell 49 389-398.

Cunningham, B. A; Hemperley, J. J; Murray, B. A; Prediger, E. A; Brackenbury, R. & Edelman, G. M. (1987) Science 236 799-806.

D'Aquila, R. T. et,al. (1991) Nucl. Acids Res. 19 3749.

Darnell, J. E. (1982) Nature 297 365-371.

Datta, B. & Weiner, A. M. (1991) Nature 352 821-824.

Delsert, C. D. & Rosenfeld, M. G. (1992) J. Biol. Chem. 267 14573-14579.

Deng, J. S; Takasaki, Y. & Tan, E. M. (1981) J. Cell Biol. 91 654-660.

DeNoto, F. M; Moore, D. D. & Goodman, H. M. (1981) Nucl. Acids Res. 9 3719-3730.

Deragon, J. M; Sinnett, D; Mitchell, G; Potier, M. & Labuda, D. (1990) Nucl. Acids Res. 18 6149.

Deutscher, S. L; Bhat, B. M; Pursley, M. H; Cladaras, C. & Wold, W. S. M. (1985) Nucl. Acids Res. 13 5771.

Dickson, G; Gower, H. J; Barton, C. H; Prentice, H. M; Elsom, V. L; Moore, S. E; Cox, R. D; Quinn, C; Putt, W. & Walsh, F. S. (1987) Cell 50 1119-1130.

DiLella, A. G; Marvit, J; Lidsky, A. S; Guttler, F. & Woo, S. L. C. (1986) Nature 322 799-803.

Doherty, P; Kitty Moolenaar, C. E. C; Ashton, S. V; Michalides, R. J. A. M. & Walsh, F. S. (1992) Nature 356 791-793.

Dreyfuss, G; Swanson, M. S. & Pinol-Roma, S. (1988) TIBS 13 86-91.

Early, P; Rogers, J; Davis, M; Calame, K; Bond, M; Wall, R. & Hood, L. (1980) Cell 20 313-319.

Edbrooke, M. R; Parker, D; McVey, J. H; Riley, J. H; Sorenson, G. D; Pettengill, O. S. & Craig, R. K. (1985) EMBO J. 4 715-724.

Edelman, G. M. (1986) Ann. Rev. Cell. Biol. 2 81-116.

Edlind, T. D; Cooley, T. E. & Ihler, G. M. (1987) Nucl. Acids Res. 15 8566.

Edwards, M. K. S; Harris, J. F. & McBurney, M. W. (1983) Mol. Cell. Biol. 3 2280-2286.

Edwards, R. H; Selby, M. J. & Rutter, W. J. (1986) Nature 319 784-787.

Elkon, K. B; Hines, J. J; Chu, J. & Parnassa, A. (1990) J. Immunol. 145 636-643.

Emeson, R. B; Hedjran, F; Yeakley, J. M; Guise, W. & Rosenfeld, M. G. (1989) Nature 341 76-80.

Eperon, L. P; Estibeiro, J. P. & Eperon, I. C. (1986) Nature 324 280-282.

Eperon, L. P; Graham, I. R; Griffiths, A. D. & Eperon, I. C. (1988) Cell 54 393-401.

Estes, P. A; Cooke, N. E. & Liebhaber, S. A. (1992) J. Biol. Chem. 267 14902-14908.

Evain-Brion, D; Binet, E; Donnadieu, M; Lawent, P. & Anderson, W. B. (1984) Dev. Biol. 104 406-412.

Feener, C. A; Koenig, M. & Kunkel, L. M. (1989) Nature 338 509-511.

Feinberg, A. P. & Vogelstein, B. (1983) Analytical Biochem. 132 6-13.

Ffrench-Constant, C. & Hynes, R. O.(1989) Development 106 375-388.

Ffrench-Constant, C; Van De Water, L; Dvorak, H. F. & Hynes, R. O. (1989) J. Cell Biol. 109 903-914.

Fischer, U. & Luhrmann, R. (1990) Science 240 787-790.

Fisher, D. E; Conner, G. E; Reeves, W. H; Wisniewolski, R. & Blobel, G. (1985) Cell 42 751-758.

Flicker, P. F; Phillips, Jr G. N. & Cohen, C. (1982) J. Mol. Biol. 162 495-501.

Forbes, D. J; Kirschner, M. W; Caput, D; Dahlberg, J. E. & Lund, E. (1984) Cell 38 681-689.

Foulkes, N. S; Borreli, E. & Sassone-Corsi, P. (1991) Cell 64 739-749.

Foulkes, N.S; Mellstrom, B; Benusiglio, E. & Sassone-Corsi, P. (1992) Nature 355 80-84.

Frankel, A. D; Mattaj, I. W. & Rio, D. C. (1991) Cell 67 1041-1046.

Frendeway, D. & Keller, W. (1985) Cell 42 355-367.

Fritz, A; Parisot, R; Newmeyer, D. & De Robertis, E. M. (1984) J. Mol. Biol. 178 273-285.

Fu, X; Colgan, J. D. & Manley, J. L. (1988) Mol. & Cell. Biol. 8 3582-3590.

Fu, X. & Manley, J. L. (1987) Mol. Cell. Biol. 7 238-248.

Fu, X. & Maniatis, T. (1990) Nature 343 437-441.

Fujinami, R. S; Oldstone, M. B. A; Wroblewska, Z; Frankel, M. E. & Koprowski, H. (1983) Proc. Natl. Acad. Sci. USA 80 2346-2350.

Gall, J. G. & Callan, H. G. (1989) Proc. Natl. Acad. Sci. USA 86 6635-6639.

Garcia-Blanco, M; Jamieson, S. F. & Sharp, P. A. (1989) Genes & Dev. 3 1874-1886.

Ge; H. & Manley, J. L. (1990) Cell 62 25-34.

Ge, H; Ping, Z. & Manly, J. C. (1991) Cell 66 373-382.

Gerrelli, D; Sharpe, G. & Latchman, D. S. (1991) Nucl. Acids Res. 19 (23) 6642.

Gillardon, F; Morano, I; Ganten, U. & Zimmerman, M. (1991) Neuroscience 125 77-80.

Gilliland, G; Perrin, S; Blanchard, K. & Bunn, H. F. (1990) Proc. Natl. Acad. Sci. USA 87 2725-2729.

Gilmartin, G. M; McDevitt, M. A. & Nevins, J. R. (1988) Genes & Dev. 2 578-587.

Goguel, V; Liao, X; Rymond, B. C. & Rosbash, M. (1991) Genes & Dev. 5 1430-1438.

Gorman, C. M. (1985) IRL Press. 2 143-190.

Gorman, C. M; Moffat, L. E. & Howard, B. H. (1982) Mol. Cell Biol. 2 1044-1051.

Grabowski, P. J; Seiler, S. R. & Sharp, P. A. (1985) Cell 42 345-353.

Greeb, J. & Shull, G. E. (1989) J. Biol. Chem. 264 18569-18576.

Green. M. R. (1989) Curr. Op. in Cell Biol. 1 519-525.

Green. M. R. (1991) Ann. Rev. Cell Biol. 7 559-599.

Greene, L. A. & Tischler, A. S. (1976) Proc. Natl. Acad. Sci. USA 73 2424-2428.

Grimaldi, K; Gerrelli, D; Sharpe, N. G; Lund, T. & Latchman, D. S. (1992) DNA Seq. 2 241-246.

Grimaldi, K; Horn, D. A; Hudson, L. D; Terenghi, G; Polak, J. M. & Latchman, D.S. (1993) Dev. Biol. In press.

Groning, K; Palfi, Z; Gupta, S; Cross, M; Wolff, T. & Bindereif, A. (1991) Mol. Cell. Biol. 11 2026-2034.

Guldner, H. H; Lakomek, H. J. & Bautz, F. A. (1983) J. Immunol. Meth. 64 45-59.

Gunning, P; Leavitt, J; Muscat, G; Ng, S. & Kedes, L. (1987) Proc. Natl. Acad. Sci. USA 84 4831-4835.

Guthrie, C. & Patterson, B. (1988) Ann. Rev. Gen. 22 387-419.

Hamm, J; Darzynkiewicz, E; Tahara, S. M. & Mattaj, I. W. (1990) Cell 62 569-577.

Hamm, J; Dathan, N. A. & Mattaj, I. W. (1989) Cell 59 159-169.

Hamm, J; Dathan, N. A; Scherly, D. & Mattaj, I. W. (1990) EMBO J. 9 1237-1244.

Hamm, J; Kasmaier, M. & Mattaj, I. W. (1987) EMBO J. 6 3479-3485.

Hamshere, M; Dickson, G. & Eperon, I. (1991) Nucl. Acids Res. 19 4709-4716.

Han; H; Nichols, R. A; Rubin, M. R; Bahler, M. & Greengard, P. (1991) Nature 349 697-700.

Harper, J. E. & Manley, J. L. (1991) Mol. Cell. Biol. 11 5945-5953.

Hashimoto, C. & Steitz, J. A. (1984) Nucl. Acids Res. 12 3283-3293.

Hashimoto, C. & Steitz, J. A. (1986) Cell 45 581-591.

Hatzopoulos, A. K; Stoykova, A. S; Erselius, J. R; Goulding, M; Neuman, T. & Gruss, P. (1990) Development 109 349-362.

Hausner, T; Giglio, L. M. & Weiner, A. M. (1990) Genes & Dev. 4 2146-2156.

Helfman, D. M; Roscigno, R. F; Mulligan, G. J; Finn, L. A. & Weber, K. S. (1990) Genes & Dev. 4 98-110.

Hoppener, J. W. M; Steenbergh, P. H; Slebos, R. J. C; Visser, A; Lips, C. J. M; Jansz, H. S; Bechet, J. M; Petermann, J. B. & Fischer, J. A. (1987) J. Clin. End. & Met. 64 809-817.

Horn, D. A; Suburo, A; Terenghi, G; Hudson, L; Polak, J. M. & Latchman, D. S. (1992) Mol. Brn. Res. 16 13-19.

Hoshijima, K; Inoue, K; Higuchi, I; Sakamoto, H. & Shimura, Y. (1991) Science 252 833-836.

Hsu, W. H; Rudolph, U; Sanford, J; Bertrand, P; Olate, J; Nelson, C; Moss, L. G; Boyd, A. E; Codina, J. & Birnbaumer, L. (1990) J. Biol. Chem. 265 11220-11226.

Hung, T; Mak, K. & Fong, K. (1990) Nucl. Acids Res. 18 4953.

Hunter, T. (1987) Cell 49 1-4.

Innis, M. A; Gelfand, D. H; Sninsky, J. J. & White, T. J. (1990) Academic press. PCR protocols.

Inoue, K; Hoshijima, K; Sakamoto, H. & Shimura, Y. (1990) Nature 344 461-463.

Izant, J. G. & Weintraub, H. (1984) Cell 36 1007-1015.

Jackson, A. P; Seow, H; Holmes, N; Drickamer, K. & Parham, P. (1987) Nature 326 154-159.

Jacquier, A. (1990) TIBS 15 351-354.

Jin, J. & Lin, J. J. (1988) J. Biol. Chem. 263 7309-7315.

Jin, J. & Lin, J. J. (1989) J. Biol. Chem. 264 14471-14477.

Jinno, Y; Yoshiura, K. & Niikawa, N. (1990) Nucl. Acids Res. 18 6739.

Jones, M. & Guthrie, C. (1990) EMBO J. 9 2555-2561.

Jones-Villeneuve, E. M; Rudnicki, M. A; Harris; J. F. & McBurney, M. W. (1983) Mol. Cell. Biol. 3 2271-2279.

Kang, J. & Muller-Hill, B. (1990) Biochem. Biophys. Res. Comm. 166 1192-1200.

Kastner, B; Kornstadt, U; Bach, M. & Luhrmann, R. (1992) J. Cell Biol. 116 839-849.

Kawasaki, E. S. (1990) Academic Press. San Diego. 21-27.

Keller, E. B. & Noon, W. A. (1981) Proc. Natl. Acad. Sci. USA 81 7417-7420.

Kemp, L. M; Brickell, P. M; LaThangue, N. B. & Latchman D. S. (1986) Biosci. Rep. 6 945-951.

Khoury, G; Gruss, P; Dhar, R. & Lai, C. (1979) Cell 18 85-92.

Kim, S. K. & Wold, B. J. (1985) Cell 42 129-138.

Kimes, B. W. & Brandt, B. L. (1976) Exp. Cell Res. 98 367-381.

King, C. R. & Piatigorsky, J. (1983) Cell 32 707-712.

Kitamura, N; Takagaki, Y, Furuto, S; Tanaka, T. & Nawa, H. (1983) Nature 305 545-549.

Kjems, J; Frankel, A. D. & Sharp, P. A. (1991) Cell 67 169-178.

Kleuss, C; Hescheler, J; Ewel, C; Rosenthal, W; Schultz, G. & Wittig, B. (1991) Nature 353 43-48.

Koenig, R. J; Lazar, M. A; Hodin, R. A; Brent, G. A; Larsen, P. R; Chin, W. W. & Moore, D. D. (1989) Nature 337 659-661.

Konarska, M. M; Grabowski, P. J; Padgett, R. A. & Sharp, P. A. (1985) Nature 313 552-557.

Konarska, M. M; Padgett, R. A. & Sharp, P. A. (1985) Cell 42 165-171.

Konarska, M. M. & Sharp, P. A. (1987) Cell 49 763-774.

Korf, G. M; Botros, I. W. & Stumph, W. E. (1988) Mol. Cell Biol 8 5566-5569.

Krainer, A. R; Conway, G. C. & Kozak, D. (1990) Cell 62 35-42.

Krainer, A. R; Mayeda, A; Kozak, D. & Binns, G. (1991) Cell 66 383-394.

Kramer, A; Frick, M. & Keller, W. (1987) J. Biol. Chem. 262 17630-17640.

Kramer, A; Keller, W; Appel, B. & Luhrmann, R. (1984) Cell 38 299-307.

Kramer, A. & Utans, U. (1991) EMBO J. 10 1503-1509.

Kretzner, L; Rymond, B. C. & Rosbash, M. (1987) Cell 50 593-602.

Krol, A; Gallinaro, H; Lazar, E; Jacob, M. & Branlant, C. (1981) Nucl. Acids Res. 9 769-

Kruger, K; Grabowski, P. J; Zaug, A. J; Sands, J; Gottscling, D. E. & Cech, T. R. (1982) Cell 31 147-157.

Kuhne, T; Wieringa, B; Reiser, J. & Weissmann, C. (1983) EMBO J. 2 727-733.

Kuo, H; Nasim, F. H. & Grabowski, P. J. (1991) Science 251 1045-1050.

Lamm, G. M; Blencowe, B. J; Sproat, B. S; Iribarren, A. M; Ryder, U. & Lamond, A. I. (1991) Nucl. Acids Res. 19 3193-3198.

Lamond, A. I. (1991a) Curr. Opin. Cell Biol. 3 493-501.

Lamond, A. I. (1991b) TIBS 16 452-453.

Lamond, A. I; Barabino, S. & Blencowe, B. J. (1990) Nucl. Acids & Mol. Biol. 4 245-257.

Lamond, A. I; Konarska; M. M; Grabowski, P. J. & Sharp, P. A. (1988) Proc. Natl. Acad. Sci. USA 85 411-415.

Lamond, A. I. & Gibson, T. J. (1990) TIG 6 145-149.

Lamond, A. I; Sproat, B; Ryder, U. & Hamm, J. (1989) Cell 58 383-390.

Lang, K. M. & Spritz, R. A. (1987) Mol. & Cell. Biol. 7 3428-3437.

Langford, C. J. & Gallwitz, D. (1983) Cell 33 519-527.

LaRosa, G. J. & Gudas, L. J. (1988) Mol. & Cell. Biol. 8 3906-3917.

Laski, F. A; Rio, D. C. & Rubin, G. M. (1986) Cell 44 7-19.

Latchman, D. S. (1990a) The New Biologist 2 297-303.

Latchman, D. S. (1990b) Biochem. J. 270 281-289.

Laufer, R. & Changeux, J. (1987) EMBO J. 4 901-906.

Lazinski, D; Grzadzielska, E. & Das, A. (1989) Cell 59 207-218.

Leff, S. E; Brannan, C. I; Reed, M. L; Ozcelik, T; Francke, U; Copeland, N. G. & Jenkins, N. A. (1992) Nature Genetics 2 259-264.

Leff, S. E. & Rosenfeld, M. G. (1986) Ann. Rev. Biochem. 55 1091-1117.

Leff, S. E; Evans, R. M. & Rosenfeld, M. G. (1987) Cell 48 517-524.

Lehrach, H; Diamond, D; Wozney, J. M. & Boedtker, J. (1977) Biochemistry 16 4743-4748.

Lelay-Taha, M; Reveillard, I; Sri-Widada, J; Brunel, C. & Jeanteur, P. (1986) J. Mol. Biol. 189 519-532.

Lerner, M. R. & Steitz, J. A. (1979) Proc. Natl. Acad. Sci. USA 76 5495-5499.

Levy, B. T; Sorge, L. K; Meymandi, A. & Maness, P. F. (1984) Dev. Biol. 104 9-17.

Levy, J. B. & Brugge, J. S. (1989) Mol. & Cell. Biol. 9 3332-3341.

Levy, J. B; Dorai, T; Wang, L. & Brugge, J. S. (1987) Mol. & Cell. Biol. 7 4142-4145.

Li, S; Crenshaw, E. B; Rawson, E. J; Simmons, D. M; Swanson, L. W. & Rosenfeld, M. G. (1990) Nature 347 528-533.

Li, S; Klein, E. S; Russo, A. F; Simmons, D. M. & Rosenfeld, M. G. (1989) Proc. Natl. Acad. Sci. USA 86 9778-9782.

Liautard, J; Sri-Widada, J; Brunel, C. & Jeanteur, P. (1982) J. Mol. Biol. 162 623-643.

Liesi, P; Rechardt, L. & Wartiovaara, J. (1983) Nature 306 265-267.

Lingner, J; Kellermann, J. & Keller, W. (1991) Nature 354 496-498.

Luhrmann R; Kastner, B. & Bach, M. (1990) Biocnim. Biophys. Acta 1087 265-292.

Lund, E; Kahan, B. & Dahlberg, J. E. (1985) Science 229 1271-1274.

Lynch, S. A; Brugge, J. S. & Levine, J. M. (1986) Science 234 873-876.

MacIntyre, I; Hillyard, C. J; Murphy, P. K; Reynolds, J. J; Gaines Das, R. E. & Craig, R. K. (1982) Nature 300 460-462.

Madore, S. J; Wieben, E. D. & Pederson, T. (1984) J. Cell Biol. 98 188-192.

Magandzo, K; Shirvan, A; Cultraro, C; Srivastava, M; Pollard, H. B. & Lee Burns, A. (1991) J. Biol. Chem. 266 3228-3232.

Maniatis, T. & Reed, R. (1987) Nature 325 673-678.

Manley, J. L. (1988) Biochim. Biophys. Acta 950 1-12.

Mardon, H. J; Sebastio, G. & Baralle, F. E. (1987) Nucl. Acids Res. 15 7725-7733.

Marshall, I; Al-kazwini, S; Holman, J. J; Broad, P. M; Edbrooke, M. R. & Craig, R. K. (1986) Plenum Neuroend. Mol. Biol. 1-16.

Martinez, R; Mathey-Prevot, B; Bernards, A; & Baltimore, D. (1987) Science 237 411-415.

Marzluff, W. F. & Pandey, N. B. (1988) T.I.B.S. 13 49-52.

Mattaj, I. W. (1986) Cell 46 905-911.

Mattaj, I. W. (1989) Cell 57 1-3.

Mattaj, I. W. (1990) Curr. Opin. Cell Biol. 2 528-538.

Mattaj, I. W. & Hamm, J. (1989) Development 105 183-189.

Mattaj, I. W. & DeRobertis, E. M. (1985) Cell 40 111-118.

Mattioli, M. & Reichlin, M. (1973) J. Immunol. 110 1318.

Mattox, W; Ryner, L. & Baker, B. S. (1992) J. Biol. Chem. 267 19023-19026.

Mayeda, A. & Krainer, A. R. (1992) Cell 68 365-375.

McAllister, G; Amara, S. G. & Lerner, M. R. (1988) Proc. Natl. Acad. Sci. USA 85 5296-5300.

McAllister, G; Roby-Shemkovitz, A; Amara, S. G. & Lerner, M. R. (1989) EMBO J. 8 1177-1181.

McBurney, M. W. & Rogers, B. J. (1982) Dev. Biol. 89 503-508.

Medford, R. M; Nguyen, H. T; Destree, A. T; Summers, E. & Nadal-Ginard, B. (1984) Cell 38 409-421.

Michaud, N. & Goldfarb, D. (1992) J. Cell Biol. 116 851-861.

Milhaud, G; Calmette, C; Taboulet, J; Julienne, A. & Moukhtar, M. S. (1974) Lancet 1 462-463.

Minvielle, S; Giscard-Dartevelle, S; Cohen, R; Taboulet, J; Labye, F; Jullienne, A; Rivaille, P; Milhaud, G; Moukhtar, M. S. & Lasmoles, F. (1991) J. Biol. Chem. 266 24627-24631.

Mitchell, P. J; Urlaub, G. & Chasin, L. (1986) Mol. Cell. Biol. 6 1926-1935.

Mohr, G. & Lambowitz, A. M. (1991) Nature 354 164-167.

Montzka, K. A. & Steitz, J. A. (1988) Proc. Natl. Acad. Sci. U.S.A. 85 8885-8889.

Morgenstern, J. P. & Land, H. (1990) Nucl. Acids Res. 18 1068.

Morris, H. R; Panico, M; Etienne, T; Tippins, J; Girgis, S. I. & MacIntyre, I. (1984) Nature 308 746-748.

Mount, S. M. (1982) Nucl. Acids Res. 10 459-472.

Mount, S. M; Petterson, I; Hinterberger, M; Karmas, A. & Steitz, J. A. (1983) 33 509-518.

Mount, S. M. & Steitz, J. A. (1981) Nucl. Acids Res. 9 6351-6368.

Mumberg, D; Lucibello, F. C; Schuermann, M. & Muller, R. (1991) Genes & Dev. 5 1212-1223.

Munroe, S. H. (1984) Nucl. Acids Res. 12 8437-8456.

Murray, B. A; Owens, G. C; Prediger, E. A; Crossin, K. L; Cunningham, B. A. & Edelman, G. M. (1986) J. Cell Biol. 103 1431-1439.

Nabeshima, Y; Fujii-Kuriyama, Y; Muramatsu, M. & Ogata, K. (1984) Nature 308 333-338.

Nagai, K; Oubridge, C; Jessen, T. H; Li, J. & Evans, P. R. (1990) Nature 348 515-520.

Nakagawa, T; Mabry, M; De Bustros, A; Ihle, J. N; Nelkin, B, D. & Baylin, S. B. (1987) Proc. Natl. Acad. Sci. USA 84 5923-5927.

Nawa, H; Kotani, H. & Nakanishi, S. (1984) Nature 312 729-734.

Nelkin, B. D; Chen, K. Y; de Bustros, A; Roos, B. A. & Baylin, S. B. (1989) Canc. Res. 49 6949-6952.

Nelson, K. K. & Green, M. R. (1990) Proc. Natl. Acad. Sci. USA 87 6253-6257.

Nevins, J. R. & Darnell, J. E. (1978) Cell 15 1477-1493.

Newman, A. & Norman, C. (1991) Cell 65 115-123.

Nigg, E. A; Baeuerle, P. A. & Luhrmann, R. (1991) Cell 66 15-22.

Niwa, M; MacDonald, C. C. & Berget, S. M. (1992) Nature 360 277-280.

Nyman, U; Jiang, W; Mellquist, E; Pettersson, I. & Ringertz, N. (1991) Exp. Cell Res. 197 307-313.

Oetting-Deems, R; Gardinaux, F; Deacon, R. W. & Young; D. A. (1991) Biochem. Biophys. Res. Comm. 181 116-120.

Ohosone, Y; Mimori, T; Griffith, A; Akizuki, M; Homma, M; Craft, J. & Hardin, J. A. (1989) Proc. Natl. Acad. Sci. USA 86 4249-4253.

Ordahl, C. P; Kioussis, D; Tilghman, S. M; Ovitt, C. & Fornwald, J. (1980) Proc. Natl. Acad. Sci. USA 77 4519-4523.

Ozcelik, T; Leff, S. E; Robinson, W; Donlon, T; Lalande, M; Sanjines, E; Schinzel, A. & Francke, U. (1992) Nature Genetics 2 265-269.

Padgett, R. A; Mount, S. M; Steitz, J. A. & Sharp, P. A. (1983) Cell 35 101-107.

Pan, Z. & Prives, C. (1988) Science 241 1328-1331.

Pan, Z. & Prives, C. (1989) Genes & Dev. 3 1887-1898.

Parker, R; Siliciano, P. G. & Guthrie, C. (1987) Cell 49 229-239.

Parry, H. D; Scherly, D. & Mattaj, I. W. (1989) TIBS 14 15-19.

Patterson, B. & Guthrie, C. (1987) Cell 49 613-624.

Patton, J. R. & Pederson, T. (1988) Proc. Natl. Acad. Sci. USA 85 747-751.

Peebles, C. L; Perlman, P. S; Mecklenburg, K. L; Petrillo, M. L; Tabor, J. H; Jarrell, K. A. & Cheng, H. L. (1986) Cell 44 213-223.

Pettersson, I; Hinter berger, M; Mimori, T; Gottlieb, E. & Steitz, J. A. (1984) J. Biol. Chem. 259 5907-5914.

Pikielny, C. W; Teem, J. L. & Rosbash, M. (1983) Cell 34 395-403.

Pinol-Roma, S; Do Choi, Y; Matunis, M. J. & Dreyfuss, G. (1988) Genes & Dev. 2 215-227.

Pinol-Roma, S. & Dreyfuss, G. (1992) Nature 355 730-732.

Powell, S. K; Cunningham, B. A; Edelman, G. M. & Rodriguez-Boulan, E. (1991) Nature 353 76-77.

Prediger, E. A; Hoffman, S; Edelman, G. M. & Cunningham, B. A. (1988) Proc. Natl. Acad. Sci. USA 85 9616-9620.

Pyper, J. M. & Bolen, J. B. (1990) Mol. Cell. Biol. 10 2035-2040.

Query, C. C; Bentley, R. C. & Keene, J. D. (1989) Cell 57 89-101.

Raabe, T; Bollum, F. J. & Manley, J. L. (1991) Nature 353 229-234.

Raud, J; Lundeberg, T; Brodda-Jansen, G; Theodorsson, E. & Hedquist, P. (1991) Biochem. Biophys. Res. Comm. 180 1429-1435.

Reddy, R. (1988) Nucl. Acids Res. 16 71-85.

Reed, R. (1990) Proc. Natl. Acad. Sci. USA 87 8031-8035.

Reed, R. & Maniatis, T. (1985) Cell 41 95-105.

Reed, R. & Maniatis, T. (1986) Cell 46 681-690.

Reed, R. & Maniatis, T. (1988) Genes & Dev. 2 1268-1276.

Rees, D. J. G; Rizza, C. R. & Brownlee, G. G. (1985) Nature 316 643-645.

Reyes, A. A; Small, S. J. & Akeson, R. (1991) Mol. Cell Biol. 11 1654-1661.

Ricketts, M. H; Simons, M. J; Parma, J; Mercken, L; Dong, Q. & Vassart, G. (1987) Proc. Natl. Acad. Sci. USA 84 3181-3184.

Riedel, N; Wise, J; Swerdlow, H; Mak, A. & Guthrie, C. (1986) Proc. Natl. Acad. Sci. USA 83 8097-8101.

Riedel, N; Wolin, S. & Guthrie, C. (1987) Science 235 328-331.

Rogers, J; Early, P; Carter, C; Calame, K; Bond, M; Hood, L. & Wall, R. (1980) Cell 20 303-312.

Rokeach, L. A; Haselby, J. A. & O. Hoch, S. (1988) Proc. Natl. Acad. Sci. U.S.A. 85 4832-4836.

Rokeach, L. A; Jannitapour, M; Haselby, J. A. & O. Hoch, S. (1989) J. Biol. Chem. 264 5024-5030.

Rokeach, L. A; Jannitapour, M. & O. Hoch, S. (1990) J. Immunol. 144 1015-1022.

Rosenfeld, M. G; Amara, S. G. & Evans, R. M. (1984) Science 225 1315-1320.

Rosenfeld, M. G; Mermod, J; Amara, S. G; Swanson, L. W; Sawchenko, P. E; Rivier, J; Vale, W. W. & Evans, R. M. (1983) Nature 304 129-135.

Rozek, C. E. & Davidson, N. (1983) Cell 32 23-34.

Ruby, S. W. & Abelson, J. (1988) Science 242 1028-1035.

Ruby, S. W. & Abelson, J. (1991) T.I.B.S. 7 79 86.

Ruskin, B. & Green, M. R. (1985) Nature 317 732-734.

Ruskin, B; Greene, J. M. & Green, M. R. (1985) Cell 41 833-844.

Ruskin, B; Krainer, A. R; Maniatis, T. & Green, M. R. (1984) Cell 38 317-331.

Sadoul, R; Hirn, M; Deagostini-Bazin, H; Rougon, G. & Goridis, C. (1983) Nature 304 347-349.

Saiki, R. K; Gelfand, D. H; Stoffel, S; Scharf, S. J; Higuchi, R; Horh, G. T; Mullis, K. B. & Erlich, H. A. (1988) Science 239 487-491.

Salditt-Georgieff, M; Harpold, M. & Chen-Kiang, S. (1980) Cell 19 69-78.

Sambrook, J;Fritsch, E. F. & Maniatis, T. (1989) 2nd ed. CSH press. Molecular Cloning.

Santoni, M. J; Barthels, D; Vopper, G; Boned, A; Goridis, C. & Wille, W. (1989) EMBO J. 8 385-392.

Sarao, R; Gupta, K; Auld, V. J. & Dunn, R. J. (1991) Nucl. Acids Res. 19 5673-5679.

Sarkar, G; Kapelner, S. & Sommer, S. S. (1990) Nucl. Acids Res. 18 7465.

Sawa, H. & Shimura, Y. (1991) Nucl. Acids Res. 19 6819-6821.

Schartl, M. & Barnekow, A. (1984) Dev. Biol. 105 415-422.

Schaufele, F; Gilmartin, G. M; Bannwarth, W. & Birnstiel, M. L. (1986) Nature 323 777-781.

Scherly, D; Boelens, w; Dathan, n. A; van Venrooij, W. J. & Mattaj, I. W. (1990) Nature 345 502-506.

Schmauss, C; Brines, M. L. & Lerner, M. R. (1992) J. Biol. Chem. 267 8521-8529.

Schmauss, C. & Lerner, M. R. (1990) J. Biol. Chem. 265 10733-10739.

Schmauss, C; McAllister, G; Ohosone, Y; Hardin, J. A. & Lerner, M. R. (1989) Nucl. Acids Res. 17 1733-1743.

Schmidt, F. J. (1985) Cell 41 339-340.

Schmitt, P; Gattoni, R; Keohavong, P. & Strevenin, J. (1987) Cell 50 31-39.

Sedmak, J. J. & Grossberg, S. E. (1977) Anal. Biochem. 79 544-552.

Seraphin, B; Kreutzner, L. & Rosbash, M. (1988) EMBO J. 7 2533-2538.

Seraphin, B; Abovich, N. & Rosbash, M. (1991) Nucl. Acids Res. 19 3857-3860.

Seraphin, B. & Rosbash, M. (1989) Gene 82 145-151.

Sharp, P. A. (1987) Science 235 766-771.

Sharpe, N. G; Isenberg, D; Williams, D. G. & Latchman, D. S. (1989c) J. Autoimm. 2 701-708.

Sharpe, N. G; Wall, W. & Latchman, D. S. (1990b) Nucl. Acids Res. 18 3425-3426.

Sharpe, N. G; Williams, D. G; Howarth, D. N; Coles, B. & Latchman, D. S. (1989b) FEBS Lett. 250 585-590.

Sharpe, N. G; Williams, D. G. & Latchman, D. S. (1990a) Mol. & Cell Biol. 10 6817-6820.

Sharpe, N. G; Williams, D. G; Norton, P. & Latchman, D. S. (1989a) FEBS Lett. 243 132-136.

Short, T. M; Fernandez, J. M; Sorge, J. A. & Huse, W. D. (1988) Nucl. Acids Res. 16 7583-7600.

Siebel, C. W; Fresco, C. D. & Rio, D. (1992) Genes & Dev. In press.

Siliciano, P. G; Brow, D. A; Roiha, H. & Guthrie, C. (1987) Cell 50 585-592.

Siliciano, P. G. & Guthrie, C. (1988) Genes & Dev. 2 1258-1267.

Small, S. J. & Akeson, R. (1990) J. Cell. Biol. 111 2089-2096.

Small, S. J; Haines, S. L. & Akeson, R. A. (1988) Neuron 1 1007-1017.

Small, S. J; Shull, G. E; Santoni, M. & Akeson, R. (1987) J. Cell Biol. 105 2335-2345.

Smith, C. W. J. & Nadal-Ginard, B. (1989) Cell 56 749-758.

Smith, C. W. J; Patton, J. G. & Nadal-Ginard, B. (1989) Ann. Rev. Gen. 23 527-577.

Solnick, D. (1985) Cell 43 667-676.

Solnick, D. & Lee, S. I. (1987) Mol. & Cell. Biol. 7 3194-3198.

Somasekhar, M. B. & Mertz, J. E. (1985) Nucl. Acids Res. 13 5591-5609.

Southern, P. J. & Berg, P. (1982) J. Mol. Appl. Gen. 1 327-341.

Spector, D. L. (1990) Proc. Natl. Acad. Sci. USA 87 147-151.

Spivack, J. G; Woods, G. M. & Fraser, N. W. (1991) J. Virol. 65 6800-6810.

Stanton, L. W. & Bishop, J. M. (1987) Mol. & Cell. Biol. 7 4266-4272.

Steenbergh, P. H; Hoppener, J. W. M; Zandberg, J; Lips, C. J. M. & Jansz, H. S. (1985) FEBS Lett. 183 403-407.

Steenbergh, P. H; Hoppener, J. W. M; Zandberg, J; Van de Veen, W. J. M; Jansz, H. S. & Lips, C. J. M. (1984) Endocrin. Metab. 59 358-360.

Strathmann, M; Wilkie, T. M. & Simon, M. I. (1990) Proc. Natl. Acad. Sci. USA 87 6477-6481.

Streuli, M. & Saito, H. (1989) EMBO J. 8 787-796.

Strittmatter, S. M; Valenzuela, D; Kennedy, T. E; Neer, E. J. & Fishman, M. C. (1990) Nature 344 836-841.

Strub, K. & Birnstiel, M. L. (1986) EMBO J. 5 1675-1682.

Suburo, A. M; Wheatley, S. C; Horn, D. A; Gibson, S. J; Jahn, R; Fischer-Colbrie, R; Wood, J. N; Latchman, D. S. & Polak, J. M. (1992) Neuroscience 46 881-889.

Swanson, M. S. & Dreyfuss, G. (1988) EMBO J. 7 3519-3529.

Szabo, A; Dalmau, J; Manley, G; Rosenfeld, M; Wong, E; Henson, J; Posner, J. B. & Furneaux, H. M. (1991) Cell 67 325-333.

Tabor, S. & Richardson, C. C. (1987) Proc. Natl. Acad. Sci. USA 84 4767-4771.

Tacke, R. & Goridis, C. (1991) Genes & Dev. 5 1416-1429.

Tan, E. M. (1982) Adv. Immunol. 33 167-237.

Takagaki, Y; Manley, J. L; MacDonald, C. C; Wilusz, J. & Shenk, T. (1990) Genes & Dev. 4 2112-2120.

Tazi, J; Alibert, C; Temsamani, J; Reveillaud, I; Cathala, G; Brunel, C. & Jeanteur, P. (1986) Cell 47 755-766.

Thomas, J. D; Conrad, R. C. & Blumenthal, T. (1988) Cell 54 533-539.

Thompson, M. A. & Ziff, E. B. (1989) Neuron 2 1043-1053.

Tollervey, D. & Mattaj, I. W. (1987) EMBO J. 6 469-476.

Towbin, H; Staehelin, T. & Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76 4350-4354.

Towhidul Alam, A. S. M; Moonga, B. S; Bevis, P. J. R; Huang, C. L. & Zaida, M. (1991) Biochem. Biophys. Res. Comm. 179 134-139.

Treacy, M. N; Neilson, L. I; Turner, E. E; He, X. & Rosenfeld, M. G. (1992) Cell 68 491-505.

Treisman, R; Orkin, S. H. & Maniatis, T. (1983) Nature 302 591-596.

Treisman, R; Proudfoot, N. J; Shander, M. & Maniatis, T. (1982) Cell 29 903-911.

Tsukamoto, T; Toyama, R; Itoh, H; Kozasa, T; Matsuoko, M. & Kaziro, Y. (1991) Proc. Natl. Acad. Sci. USA 88 2974-2978.

Tyc, K. & Steitz, J. A. (1989) EMBO J. 8 3113-3119.

Van Dam, A; Winkel, I; Zijlstra-Baalbergen, J; Smeenk, R. & Theo Cuypers, H. (1989) EMBO J. 8 3853-3860.

Van der Veen, R; Arnberg, A. C; van der Horst, G; Bonen, L; Tabak, H. F. & Grivell, L. A. (1986) Cell 44 225-234.

Van Doren, K. & Hirsh, D. (1988) Nature 335 556-559.

Vankan, P; McGuigan, C. & Mattaj, I. W. (1992) EMBO J. 11 335-343.

Viskochil, D; Cawthon, R; O'Connell, P; Xu, G; Stevens, J; Culver, M; Carey, J. & White, R. (1991) Mol. & Cell. Biol. 11 906-912.

Waring, R. B. & Davies, R. W. (1984) Gene 28 277-291.

Waring, R. B; Towner, P, Minter, S. J. & Davies, W. (1986) Nature 321 133-139.

Watanabe-Fukunaga, R; Brannan, C. I; Copeland, N. G; Jenkins, N. A. & Nagata, S. (1992) Nature 356 314-317.

Weintraub, H; Izant, J. G. & Harland, R. M. (1985) TIG 1 22-25.

Wersig, C. & Bindereif, A. (1990) Nucl. Acids Res. 18 6223-6229.

Wieringa, B; Hofer, E. & Weissmann, C. (1984) Cell 37 915-925.

Wilkinson, M. (1988) Nucl. Acids Res. 16 10934.

Williams, D. G; Sharpe, N. G; Wallace, G. & Latchman, D. S. (1990) J. Autoimm. 3 715-725.

Williams, D. G; Smith, P. R. & Maini, R. N. (1985) Biochem. Soc. Trans. 13 103-104.

Williams, D. G; Stocks, M. R; Smith, P. R. & Maini, R. N. (1986) Immunology 58 495-500.

Winfield, J. B; Brunner, C. M. & Koffler, D. (1978) Arth. & Rheum. 21 289-294.

Wise, J; Tollervey, D; Maloney, D; Swerdlow, H; Dunn, E. J. & Guthrie, C. (1983) Cell 35 743-751.

Wolff, T. & Bindereif, A. (1992) EMBO J. 11 345-359.

Wood, J. N; Bevan, S. J; Coote, P. R; Dunn, P. M; Harmar, A; Hogan, P; Latchman, D. S; Morrison, C; Rougon, G; Thevenian, M. & Wheatley, S. (1990) Proc. R. Soc. Lond. B. 241 187-194.

Woppmann, A; Rinke, J. & Luhrmann, R. (1988) Nucl. Acids Res. 16 10985-11004.

Wu, J. & Manley, J. L. (1989) Genes & Dev. 3 1553-1561.

Wu, J. & Manley, J. L. (1991) Nature 352 818-821.

Yamamoto, T; Jay, G. & Paston, I. (1980) Proc. Natl. Acad. Sci. USA 77 176-180.

Yang, V. W; Lerner, M. R; Steitz, J. A. & Flint, S. J. (1981) Proc. Natl. Acad. Sci. USA 78 1371-1375.

Young, R. A; Hagenbuchle, O. & Schibler, U. (1981) Cell 23 451-458.

Zamore, P. D. & Green, M. R. (1991) EMBO J. 10 207-214.

Zeitlin, S. & Efstratiadis, A. (1984) Cell 39 589-602.

Zeytin, F. N; Rusk, S. & Leff, S. E. (1987) Endocrinology 121 361-370.

Zieve, G. W. & Sauterer, R. A. (1990) Crit. Rev. Biochem. Mol. Biol. 25 1-46.

Ziff, E. B. (1980) Nature 287 491-499.

Zhuang, Y. & Weiner, A. M. (1986) Cell 46 827-835.

Zhuang, Y. & Weiner, A. M. (1989) Genes & Dev. 3 1545-1552.